STUDIJSKI MATERIALI

STATISTIČNO NAČRTOVANJE POSKUSOV

Dr. Marijan Blejec

IEOP 53/71

RCEF

RAZISKOVALNI CENTER EKONOMSKE FAKULTETE
UNIVERZE V LJUBLJANI

STATISTIČNO NAČRTOVANJE POSKUSOV

Dr. Marijan Blejec

IEOP 53/71

INSTITUT ZA EKONOMIKO IN ORGANIZACIJO PODJETJA

RCEF

Raziskovalni center Ekonomske fakultete

Univerze v Ljubljani

1971

K A Z A L O :

1. UVOD. PRVINE STATISTIČNEGA NAČRTOVANJA POSKUSOV.

1.1 Statistična metodologija načrtov poskusov se je razvila iz potreb poskusništva predvsem v agronomiji in biologiji. Iz tega izvira tudi vrsta ustaljenih izrazov, ki so sicer umestni v agronomskih poskusih, so pa v splošnem preživeli, ker se je uporaba metod statističnih poskusov razširila na najrazličnejša področja. Tako so metode razdeljenih površinic po imenu ne pa po načrtu brez smisla v psihologiji, industrijskih raziskavah ipd. Podobno je tudi s samim nazivom načrtovanje poskusov. Metode statističnega načrtovanja poskusov namreč s pridom uporabljamo tudi v problemih, ki nimajo značaj poskusa v ožjem smislu.

Poskus v ožjem smislu je umetno ustvarjanje določenih pogojev, ki delujejo na poskusne enote oziroma gradivo. V tem smislu s poskusom proučujemo vpliv določene vrste gnojila in agrotehnične mere tako, da parcelice, ki jih v poskusne namene pripravimo, gnojimo ali ne gnojimo z raziskovanim gnojilom in jih obdelujemo ali neobdelujemo na raziskovan način. Pri raziskovanju učinka različnih tehnoloških postopkov pri proizvodnji določenih izdelkov enako poskusne enote obdelujemo na predpisane načine pri različnih temperaturah, pri različnih recepturah mešanic ipd., jih obdelujemo na različne načine ipd. in po izvedenem poskusu merimo značilnosti izdelkov, kot so trdnost, življenjska doba ipd.

Nakazano agronomsko raziskavo pa moremo izvesti tudi na način, ki nima značilnosti pravega poskusa, pri katerem v poskusne namene tretiramo poskusne enote na določene načine. Ustrezno število parcel, obdelanih na en izmed štirih postopkov poskusa: GA_{oo} GA_{01} GA_{10} in GA_{11} zasejane z raziskovalno kulturo, moremo dobiti brez izvedbe poskusa. Iz obstoječe populacije gospodarstev oziroma parcel, ki so bile obdelovane po enem izmed štirih proučevanih postopkov moremo na slučajnosten način izbrati planu ustrezno število parcel, ki veljajo kot ponovitve poskusa. Ni potrebno posebej poudarjati, da je osnovno načelo, da iz obstoječe populacije oziroma štirih populacij izberemo potrebno število ponovitev slučajnostno. Meritve donosa na izbranih parcelah, so osnova za analizo podatkov enako, kot če bi izvedli poskus v pravem pomenu besede.

Kateri izmed obeh nakazanih načinov: pravi poskus na razmeroma homogenem poskusnem gradivu ali slučajno izbrane enote iz populacij posameznih postopkov iz obstoječih stvarnih populacij je boljši oziroma primernejši, zavisi od cilja raziskave. Pri pravem poskusu, ki je izveden na razmeroma homogenem poskusnem gradivu, odkrijemo značilnosti razlik že z razmeroma majhnim vzorcem oziroma poskusom. Ker pa ima tak poskus laboratorijski značaj, njegove rezultate razmeroma težko posplošimo na razmere, kakršne obstojajo v življenju. Po drugi strani pa zbrani podatki, ki so bili po načrtu raziskave izbrani iz stvarnih populacij, trpe na večji heterogenosti, ki izvira iz večjega kompleksa individualnih faktorjev. To je v škodo značilnosti izsledkov, kar moremo in moramo odpraviti z večjim številom ponovitev. Velika prednost drugega načina pa je v tem, da so izsledki praktično uporabnejši, ker so slika dejanskih odnosov v življenju, ne pa v laboratoriju. Če se dataknemo še težav pri izvedbi raziskave po enem ali drugem načinu, spoznamo, da je pravi poskus v toliko enostavnejši, ker podatke zberemo na enem mestu, težava pa je v tem, da je treba na umetno ustvarjenih poskusnih enotah poskuse izvesti, kar terja finančna sredstva za izvedbo in čas, kar zakasni rezultate.

Za iz obstoječih populacij izbrane vzorce pa je treba spoznati ustrezne populacije za posamezne postopke, izvesti izbor in zbrati podatke o enotah, ki so včasih geografsko na različnih mestih. Vendar pa je običajno čas, v katerem pridemo do rezultatov krajši kot pri izvedbi pravega poskusa, ker za izbrano enoto ob izboru preskočimo fazo izvedbe poskusa, ker je postopek na enoti že izveden. Možna je seveda tudi raziskava, ki ima elemente pravega in vzorčnega poskusa, če iz populacije širšega razmaka izberemo slučajnostno potrebno poskusno gradivo in na njem izvedemo poskus bodisi delno ali v celoti.

1.2 Metoda slučajnostnega izbora iz stvarnih populacij po posameznih postopkih je zelo razširila možnost uporabe metod statističnega načrtovanja poskusov. Imamo namreč primere, za katere ne moremo izbirati med alternativnima postopkoma, ker je npr. pravi poskus nemogoč. Ne moremo in nobenega smisla nima, da bi v poskusne namene umetno sestavljali gospodinjstva z določenimi značilnostmi in proučevali razlike v porabi v odvisnosti od proučevanih faktorjev. Po zgornjem načinu izberemo iz populacij posameznih postopkov določeno število gospodinjstev, ki imajo značilnosti postopka in izbrane podatke obdelamo, kot da bi bili dobljeni s poskusom v pravem pomenu besede. Vzemimo, da raziskujemo ali in kako je poraba mlečnih izdelkov odvisna od števila članov gospodinjstva, socialno ekonomske skupine in višine dohodkov. Za (4 velikosti gospodinjstev) (3 socialno ekonomske skupine) (5 skupin višine dohodkov) = 60 postopkov izberemo iz 60 populacij z značilnostmi postopkov slučajnostno npr. po $\bar{n} = 5$ gospodinjstev in zanje zberemo podatke o višini izdatkov za mlečne izdelke v proučevanem razdobju. Teh n = 300 podatkov predstavlja rezultate trofaktorskega poskusa Č.S.D. = 4.3.5. v petih ponovitvah. Element blokov bi mogli v tem primeru vključiti v našo raziskavo tako, da bi po 4.3.5 = 60 postopkov za eno ponovitev poiskali na homogenih področjih.

Podobno postopoma v raziskavi tržišča kjer načrtu raziskave oziroma poskusu ustrezne podatke dobimo bodisi z resničnim poskusom (npr. v izbrani trgovini v poskusne namene prodajamo določen izdelek na določen način) ali vzorčno, ko izberemo iz populacije vseh trgovin take, ki prodajajo na določen način in proučujemo njihovo prodajo. Tudi v teh primerih moremo kombinirati pravi poskus z vzorcem. Za postopke, ki jih je življenje samo ustvarilo, izberemo poskusne rezultate s slučajnostnim vzorcem, postopke, ki jih ni najti v stvarnosti, pa realiziramo s poskusom. Tako dobimo podatke o učinku standardne klasične prodaje z vzorcem obstoječih trgovin, medtem ko prodajo po novem načinu poskusno izvedeno v slučajnostno izbranih trgovinah. V praksi vzorčne poskuse, kot bi lahko imenovali drugo skupino poskusov, izvedemo tehnično tako, da iz celotne populacije vseh postopkov slučajnostno izbiramo enote in jih vgrajujemo glede na značilnosti v shemo poskusa. Za vsak postopek vnašamo rezultate za predpisano število ponovitev. Ko je za določen postopek doseženo ustrezno število ponovitev, v nadaljnjem vse enote za tak postopek izločimo oziroma jih ne upoštevamo. S to tehniko nadaljujemo dokler za vsak načrtovan postopek ne dosežemo ustrezno število ponovitev.

1.3 P o s k u s n o g r a d i v o . Poskusno gradivo sestoji iz množice fizičnih enot ali dogodkov, na katerih v toku poskusa glede na načrt poskusa izvedemo posamezne postopke z ali brez ponovitev. Tako v našem primeru predstavlja poskusno gradivo 30 delavcev, ki smo jih določili oziroma bolje rečeno izbrali, da bodo sodelovali v našem poskusu o vplivu načina dela na produktivnost dela. Poskusno gradivo mora predstavljati 64 kosov določene vrste lesa, na katerih bomo izvedli poskus 2^5 v dveh ponovitvah. Ti kosi lesa morejo biti vsi iz enega debla ali iz različnih dreves. V prvem primeru je poskusno gradivo bolj homogeno kot v drugem. Poskusno gradivo sestavimo tako, da iz obstoječih enot, ki so na razpolago, izberemo načrtu poskusa ustrezno število enot. Če je možen obseg enot, ki zadoščajo opredeljujočim pogojem, poskusa večji kot pa jih zahteva načrt poskusa, iz zaloge ustreznih enot izberemo potrebno število na slučajnosten način. Poskusno gradivo predstavlja skupina delavcev, ki opravljajo določeno delo po poskusnem načrtu. Poskusno gradivo predstavlja tudi določeno število parcelic na večji površini, ki jih obravnavamo z različnimi proizvodnimi faktorji, katerih učinek raziskujemo. Poskusno gradivo morejo v proizvodnih poskusih predstavljati vsakovrstne surovine. To gradivo je že vnaprej dano v diskretnih enotah, v primeru, da je zvezno, pa ga razdelimo v diskretne enote za potrebe poskusa. Če preskušamo npr. cement, moremo iz vreče cementa sestaviti poskusne enote po 10 dkg cementa, ki naprej predstavljajo poskusne enote.

1.4 Posamezne prvine statistično načrtovanega poskusa nakažimo s praktičnim primerom.

Vzemimo da proučujemo uspešnost treh načinov dela (N_1 N_2 N_3) na produktivnost dela (y), ki je izražena s številom kosov, izdelanimi v enoti časa. V ta namen v enostavni raziskavi vključimo 3o delavcev z enoletnim delovnim stažem in določimo, da po deset delavcev dela po vsakem izmed treh proučevanih načinih. Rezultati izvedenega poskusa y_{Ai} o individualnih produktivnostih dela so osnova za primerjalno analizo uspešnosti dela po treh proučevanih načinih. To analizo izvedemo s statistično tehniko analize variance (AVAR), s katero objektivno v mejah zanesljivosti, ki je zvezana s statističnimi metodami sklepanja, presodimo, ali in kolikšne so razlike v produktivnosti po različnih načinih dela. V poskusu eksplicitno nastopa samo dejavnik, katerega vpliv proučujemo - tj. način dela. Pri proučitvi problema pa takoj zasledimo, da individualne produktivnosti dela niso rezultat samo različnega načina dela, ampak niza drugih dejavnikov. Če poskus izvedemo na skupini delavcev, katerih staž je tri mesece, so rezultati o produktivnosti dela in razlike med posameznimi načini dela drugačne, kot pa pri skupini delavcev z enoletnim stažem. Če primerjamo med seboj podatke o produktivnosti dela za posamezne delavce, ki so delali po istem načinu dela, odkrijemo, da so ti podatki med seboj različni. Te razlike so rezultat vpliva najrazličnejših dejavnikov, ki vplivajo na produktivnost dela in niso enaki za vse delavce v poskusu, ampak se od delavca do delavca menjajo. Posamezni delavci so različno stari, različni so po prizadevnosti, imajo različne zmogljivosti, trenutno so različno razpoloženi ipd. Nekatere od teh motečih, a vedno prisotnih individualnih dejavnikov moremo vrednosti na posameznih enotah identificirati oziroma opredeliti. Razen opredeljivih dejavnikov pa nastopa še niz nedoločjivih, katerih vrednosti so težko, če že ne neizmerljive. Te vključimo v skupino slučajnostnih vplivov. Tako pojavljanje kot velikost slučajnostnih vplivov je vnaprej nedoločljiva.

V naš poskus je vključenih 30 delavcev in po deset delavcev dela po vsakem načinu dela. Pravimo, da smo poskus izvedli tako, da je vsaka vrednost faktorja uporabljena v desetih ponovitvah. Po zakonih o velikih številih se učinek individualnih faktorjev v povprečju zmanjša in sicer je varianca povprečja n̄-krat manjša od variance individualnih vrednosti. Seveda pa zakonitost zvezana za zakon velikih števil velja le za slučajnostne dogodke. Zato se vpliv starosti na produktivnost dela ne eliminira, če vzamemo, da so vsi delavci, ki delajo po načinu A_1 stari od 20 do 24 let, delavci, ki delajo po načinu A_2 od 25 do 29 let in delavci stari 30 do 34 let po načinu A_3. V tem primeru razlike v povprečjih kažejo združeno razliko zaradi dveh dejavnikov: starosti in načina dela. Jasno, da je taka analiza brez vrednosti. Da se izognemo sistematičnemu učinku dodatnih dejavnikov, ki s proučevanjem nimajo zveze, dodeljujemo posamezne načine dela posameznim delavcem na slučajnosten način. (Tako dobijo vsi dejavniki, ki imajo pod določenimi pogoji značaj sistematičnosti in zato moteč vpliv značaj slučajnostnih dejavnikov.)

Če po vrsti pogledamo nakazan načrt poskusa, moremo iz njega izdvojiti osnovne prvine statističnega poskusa.

Na kriterialnem znaku, ki je najobičajnejše numeričen, posredno ugotavljamo vpliv proučevanih faktorjev. Kriterialni znak, ki ga pogosto kvalificiramo kot odvisno spremenljivko, je vedno rezultativen znak. Njegova vrednost je pogojena z velikostjo faktorjev, ki na proučevan pojav delujejo. Kriterialni znak je v našem primeru produktivnost dela, izražena s številom kosov, izdelanih na enoto časa. Pri kmetijskih poskusih je kriterialni znak najpogosteje donos, ki ga dosežemo pod določenimi pogoji. V psiholoških raziskavah je kriterialni znak število točk, ki jih preizkušana oseba doseže pri določenem testu. V raziskavi kvalitete dela je kriterialni znak število napak, kijih tipkarica na določeno dolžino teksta napravi. V raziskavah proizvodnih sektorjev je kriterialni znak po pravilu znak, ki podaja kakovost izdelka, kot je trdnost, življenjska doba ipd. Kriterialni znak more biti tudi znak, ki ima ordinalni značaj. Tako more biti kriterialni znak rang, ki ga pod določenimi pogoji pripiše določenemu izdelku poskusni preskuševalec.

Faktorje, ki vplivajo na kriterialni znak, delimo v več vsebinsko in tehnično različnih skupin. Osnovna delitev faktorjev je v one, katerih vrednosti ne moremo na konkretnem primeru točno določiti. Tak faktor je npr. v našem primeru način dela, ki ga moremo točno predpisati in tudi izvajati. Določljiv faktor je v nakazanem primeru tudi starost. Pri kmetijskih poskusih je npr. določljiv faktor količina umetnega gnojila, agrotehnična mera i d. V psiholoških raziskavah je določljiv faktor, ki pogojujejo število točk pri določenem testu, starost testiranca, pogoji testiranja ipd. Pri raziskavi o kakovosti dela tipkarice je faktorialni znak hitrost diktiranja, staž daktilografinje, različen način rokopisa ipd.

V raziskavah o kakovosti izdelkov v proizvodnji so dosegljivi faktorji najrazličnejši proizvodni faktorji od kakovosti surovin do novega proizvodnega postopka v najširšem smislu.

V konkretnem primeru je faktor določen s svojo vrednostjo, ravnijo. Faktorji imajo po pravilu več vrednosti. Najmanjše število možnih vrednosti za določen faktor je dva. Število možnih vrednosti oziroma nivojev za določen faktor pa mora biti tudi neomejeno. Tako ima spol kot faktor dve vrednosti: moški ženski, faktor izmena tri vrednosti: dopoldanska, popoldanska in nočna izmena.

Faktor starost je zvezni znak in mora zavzeti neomejeno število vrednosti. Podobno je količina umetnega gnojila faktor z neomejenim številom vrednosti. Tudi hitrost tipkanja je faktor s teoretično neomejenim številom vrednosti. To velja po pravilu za vse faktorje, katerih vrednosti so zvezne, Kot faktor more v določenih raziskavah v gozdarstvu nastopiti tudi npr. drevo, ker je vrednost kriterijalnega znaka odvisna od drevesa. Prav tako more kot faktor nastopiti v določenih raziskavah delavec, če proučujemo kakšne so razlike v kakovosti dela v odvisnosti od delavcev. Po enaki logiki je enake vrste faktor tudi stroj, če proučujemo odvisnost produktivnosti dela od stroja. Faktor enakega tipa je v raziskavah trga kupec, prodajalec ali trgovina, če proučujemo odvisnost od višine prodanega blaga, od kupca, prodajalca ali trgovine.

Kot smo nakazali, faktorji v raziskavi nastopajo z različno učinkovitostjo, ki je podana z vrednostjo faktorja.

1.5 P o s k u s n i   n a č r t . Poskusni načrt sestoji iz predpisa za izvedbo poskusa. Cilj poskusnega načrta je, da za dan raziskovalni problem v danih okoliščinah da z najmanjšim obsegom raziskave oziroma z najmanjšimi stroški dobimo zadovoljive sklepe o vplivu proučevanih faktorjev na rezultativen ali kriterialen znak oziroma podatek. Za razliko od statistične raziskave, v kateri zberemo podatke o enotah proučevane populacije v celoti ali z vzorcem in dobljene podatke analiziramo, v statistično planiranem poskusu na enotah poskusnega gradiva zavestno po določenem vnaprej načrtovanem postopku izzovemo delovanje dejavnikov, katerih vpliv proučujemo.

Vzemimo, da proučujemo vpliv različnih načinov dela na produktivnost dela. Če vzamemo da proučujemo vpliv treh različnih načinov dela na produktivnost v najenostavnejšem zametku tak poskus izvedemo tako, da trije delavci delajo vsak po enem izmed proučevanih načinov dela. Tak načrt ima niz pomanjkljivosti in ugovorov. Prvo vprašanje je, katere delavce izberemo v poskus. Drugo vprašanje je, na kakšen način razmestiti proučevane načine posameznim delavcem. Takoj se pojavi tudi problem, ali so trije delavci oziroma ali je poskus, v katerem po vsakem postopku dela le po en delavec, zadosten, da objektivno odkrijemo eventuelne razlike v produktivnosti dela po posameznih načinih dela. Izkaže se, da razen načina dela vpliva na produktivnost dela še niz določljivih in nedoločljivih dejavnikov, ki vplivajo na produktivnost dela, ki je izražena npr. s številom izdelkov izdelanih v enoti časa.

1.6 O p r e d e l j u j o č i f a k t o r j i . Poskusno gradivo je določeno s posebnimi pogoji, ki so podani z opredeljujočimi pogoji poskusa. Opredeljujoči pogoji poskusa so določeni z vrsto omejitev in predpisov, ki jim morajo ustrezati posamezni poskusi. Te omejitve se nanašajo na poskusne enote, ki sestavljajo poskusno gradivo ali na samo izvedbo poskusa. Bistvo opredeljujočih pogojev je v tem, da so enaki za vse poskusne enote, torej so določeni z eno samo vrednostjo ali z omejenim variiranjem opredeljujočega pogoja. Tako pri raziskavah tržišča omejimo raziskavo oziroma poskus le na žensko prebivalstvo v določenem mestu v starosti od 25 do 30 let, ki ima visokošolsko izobrazbo. S temi opredeljujočimi faktorji, od katerih ima spol le eno vrednost, variiranje v vseh drugih opredeljujočih faktorjih pa je omejeno, dobimo razmeroma homogeno poskusno populacijo. Visoka homogenost poskusnega gradiva je v prid lažjega in preciznejšega ugotavljanja učinka proučevanih faktorjev. Zato je klasični poskus težil za čim bolj izenačenimi opredeljujočimi pogoji poskusa kar so dosegali z razmeroma težavnimi sredstvi in dragimi aparaturami. Čim večja homogenost poskusnega gradiva, ki je dana s čim popolnejšo opredelitvijo je osnovna komponenta pri statistično načrtovanem poskusu vendar ni alfa in omega dobrega poskusa iz dveh razlogov. Do zadovoljivih rezultatov o učinku proučevanih faktorjev pridemo tudi za ne preveč homogene populacije po drugi poti s ponavljanjem poskusa. S ponavljanjem poskusa se učinek heterogenosti po zakonu o velikih številih izravna. Učinek neenakosti v talnih pogojih pri poljskih poskusih omilimo s ponovitvijo poskusov. Poprečje, izračunano iz ponovitev, daje dosti zanesljivo sliko o vplivu poprečnih talnih pogojev. Drugi razlog, zaradi katerega ne težimo iz vsebinskih razlogov k pretirani homogenosti poskusnega gradiva, pa je v tem, da s čim natančnejšo opredelitvijo poskusnega gradiva sicer dosegamo večjo in večjo homogenost, obenem pa postaja poskusna populacija bolj in bolj specialna in selekcionirana. Raziskava takih populacij pa ima v večini primerov drugorazreden pomen, ker je zelo odmaknjena od življenja. Agronomski poskus v zelo izenačenih opredeljujočih pogojih v lončkih v toplih gredah da sicer razmeroma zanesljive razlike v učinku proučevanih faktorjev, vendar so ti rezultati omejenega, če že ne brez pomena za praktično uporabo v poljskih pogojih. Podobno je tudi z raziskavami tržišča, kjer zelo podrobno opredeljevanje kupcev, ki nastopajo v raziskavi, dovede do rezultatov, ki so praktično slabo uporabni, ker so reprezentativni le za zelo selekcionirano populacijo kupcev.

1.7 D o l o č l j i v i f a k t o r j i . V statističnih načrtih poskusov je večina faktorjev določljivih. Določljiv faktor je vsak faktor, katerega vrednosti moremo v konkretnem primeru realizirati ali ugotoviti. Tako je količina gnojila, s katerim pogojimo poskusno parcelo, določjiv faktor, ker moremo v vsakem primeru parcelo pognojiti s predpisano količino gnojila. Prav tako je starost kupca določljiv faktor, ker moremo njegovo starost v vsakem konkretnem primeru ugotoviti.

Določljiv faktor, ki ni opredeljujoč, ima v dani raziskavi najmanj dve vrednosti. Ali določen faktor učinkuje na kriterialni znak, moremo določiti le, če je faktor vsaj na dveh nivojih. Da ugotovimo ali določena količina gnojila učinkuje na donos, moramo uvesti faktor z vrednostima: gnojeno-negnojeno. Razlika donosu pri parcelicah, ki so bile gnojene v primerjavi s parcelicami, ki niso bile gnojene, poda sliko o učinkovitosti gnojila. Faktorji z samo dvema vrednostima so v poskusništvu zelo cenjeni iz dveh razlogov. Z alternativno vrednostjo, ki je običajno negacija določene vrednosti faktorja, uspemo iz faktorja z navidezno eno samo vrednostjo dobiti faktor z dvema vrednostima, kar omogoča proučitev učinka faktorja. Alternativna vrednost, pogosto imenovana kontrolna vrednost faktorja se v praksi pojavlja v najrazličnejših inačicah. Pri proučevanju učinka določenega zdravila imamo poleg vrednosti faktorja: "zdravilo uporabljeno" kontrolno vrednost: "zdravilo ni bilo uporabljeno". Pri "novem načinu" proizvodnje nastopa kot kontrolna vrednost "star način" proizvodnje. Poleg že omenjenega primera gnojeno-negnojeno imamo v kmetijstvu niz faktorjev, ki imajo dve vrednosti kot agrotehnična mera: "je bila" - "ni bila uporabljena" ipd.

Drug razlog za uporabo faktorjev s po dvema vrednostima pa je, da s poskusi, v katerih nastopajo faktorji s po dvema vrednostima analiziramo kompleksen učinek razmeroma velikega števila faktorjev. Medtem ko je ena ponovitev kompleksnega poskusa s šestimi faktorji s po dvema vrednostima $2^6 = 64$ postopkov, jih ima poskus istega obsega s faktorji po tri vrednosti ima $3^6 = 729$ postopkov.

1.8

Zelo pogosto v poskus ne vključimo vseh možnih vrednosti faktorjev. To je dostikrat nemogoče (število vrednosti neomejeno ali zelo veliko) pogosto pa to/za raziskavo niti ni potrebno. Zato v raziskavi faktor reduciramo na nekaj, najmanj na dve vrednosti. Reducirano število vrednosti predstavlja vrednosti, za katere nas zanimajo razlike v učinku. Tako v raziskavi trga za socioekonomski faktor vzamemo kot eno vrednost delavec kot drugo pa nameščenec. Poseben primer redukcije vrednosti faktorja so numerični faktorji. Da dobimo zadostno sliko o vplivu starosti ni treba vzeti osebe z vsemi starostmi, kar je tudi izvedljivo, temveč v proučevanje vključimo le nekaj vrednosti v razmaku starosti, ki je za raziskovanje zanimivo. Teh nekaj vrednosti, če ni drugih razlogov, vzamemo v enakih razmakih. Če je npr. za raziskavo zanimivo ponašanje kriterialnega znaka v odvisnosti od starosti v razmaku od 20 do 50 let dobimo zadostno sliko o vplivu starosti, če vzamemo starost v petletnih razmakih: 20 25 30 35 40 45 50. Namesto s sedmimi vrednostmi moremo proučiti vpliv starosti tudi s štirimi vrednostmi: 20 30 40 50 vendar je v tem primeru potek odvisnosti od starosti manj jasen, ker so razmaki širši. Problem je zelo podoben kot pri risanju funkcij, kjer z nekaj vrednostmi zadostno nakažemo potek zveznih funkcij. V teh primerih je starost nadomeščena z aritmetičnim zaporedjem starosti. Podoben primer je pri ugotavljanju kakovosti izdelkov v odvisnosti od temperature pri obdelavi izdelka. Poslužimo se iste tehnike in kompleten faktor nadomestimo z nekaj vrednostmi temperature v razmaku, za katerega proučujemo vpliv temperature. Enak problem je pri različnih dozah gnojila pri gnojilnih poskusih. Različne doze gnojila določimo tako, da tvorijo v proučevanem razmaku aritmetično zaporedje. Obnašanje kriterialnega znaka za nekaj vrednosti intervalnega znaka zadošča za opis vpliva v zveznem razmaku. Z interpolacijo moremo namreč dobiti ocene za vse interesantne vrednosti (npr. ekstrem).

Pri nominalnih faktorjih dostikrat izdvojimo samo določene vrednosti, ki nas v konkretni raziskavi zanimajo. Pri raziskavi o razlikah v ponašanju potrošnika navadno ne vzamemo vse socialno ekonomske skupine temveč le določene, ki so za planirano raziskavo zanimivi.

1.9

Vsi nakazani faktorji, ne glede na to, ali so v poskusu zastopane vse njihove vrednosti ali samo nekateri možni nivoji imajo, ko so enkrat odrejeni, značaj popolne opredelitve faktorja za dano raziskavo. Faktorje za katere so v okviru raziskave vključene vse vrednosti, imenujemo <u>fiksne faktorje</u>. V nadaljevanju bomo take faktorje in vse njihove elemente po pravilu zaznamovali z velikimi črkami. Fiksen faktor (P) ima P nivojev, ki jih moremo individualno zaznamovati s P_1 , P_2 ... P_A . Mogoče v začetku malo nenavadno zaznamovanje se izkaže v nadaljevanju kot izredno koristno. V poskusu imamo običajno kompleks faktorjev in je asociacija med elementi posameznega faktorja neposredna, ker je vezana na isto črko. Tako je gotovo manj prikladno pisanje da ima faktor P skupno f nivojev, da so poprečja za faktor P $\bar{y}_i$, pri čemer gre i = 1,2...r

Če vključimo v raziskavo faktorje, kot so npr. drevo v raziskavah v gozdarstvu, delavec v raziskavi o produktivnosti dela, stroj pri raziskavi o kakovosti dela, kupec ali trgovina pri raziskavi trga ipd. spoznamo, da je tip teh faktorjev različen od tipa faktorjev, ki smo jih obravnavali doslej. Dosedaj smo v raziskavo vključili ali vse vrednosti faktorja (npr. spol) ali pa določeno število fiksnih (odtod tudi ime faktorja) vrednosti, ki reprezentira vrednosti v določenem razmaku ali faktor v okviru poskusa opredeljuje. Za faktorje tipa kot so drevo, stroj, delavec, kupec, trgovina, pa je situacija drugačna. Ti faktorji morejo imeti dvojni značaj. V izolirani raziskavi more določeno število strojev v podjetju šteti kot fiksen faktor, če se omejimo na instalirane stroje. Drug značaj pa dobi stroj kot faktor v primeru, če skušamo rezultate poskusa posplošiti na populacijo določenih strojev na splošno. V tem primeru vseh strojev umišljene populacije ne moremo vključiti v raziskavo, ker je število strojev lahko tudi neomejeno, temveč vključimo samo nekaj strojev, nekaj dreves, nekaj delavcev, nekaj trgovin, ki naj reprezentirajo populacijo vseh vrednosti faktorjev iz umišljene populacije. Ker predpostavljamo, da je nekaj proučevanih vrednosti izbrano iz umišljene populacije na slučajnosten način, v takem primeru govorimo o <u>slučajnostnih faktorjih.</u> Slučajnostne faktorje in njihove ustrezne elemente po pravilu zaznamujemo z malimi črkami (a), število vrednosti v raziskavi je a, poprečja po vrednosti tega faktorja pa $\bar{y}_a$ ipd. Faktorji nakazanega tipa morejo imeti bodisi značaj fiksnega faktorja, če gre za izolirano proučitev vpliva tega faktorja ali značaj slučajnostnega faktorja, če gre za posplošitev učinka tega faktorja na umišljeno populacijo vrednosti. Tako je pet strojev fiksen faktor, če jih gledamo s stališča podjetja, ki z njimi razpolaga, ali slučajnosten faktor, če jih gledamo s stališča proizvajalca, ki proizvaja veliko teh strojev.

Druga možna varianta slučajnostnega faktorja je primer, da je število možnih vrednosti faktorjev končno npr. A za faktor (a), iz njih proučevano število vrednosti faktorja (a) v poskusu pa smatramo kot slučajnosten vzorec iz končnega števila vseh možnih vrednosti in o faktorjih modela III, če gre za slučajnosten faktor iz končne populacije vseh možnih vrednosti.

Terminološko govorimo o faktorjih modela I, če je faktor fiksen, o faktorjih modela II, če je faktor slučajnosten.

Nakazane lastnosti posameznih tipov faktorjev moremo strniti v naslednjo razpredelnico.

Tip faktorja	Model I	Model II	Model
Oznaka faktorja	(A)	(a)	(a)
Število vrednosti faktorja:			
vseh	A	∞	A
v poskusu	A	a	a

Analogno govorimo o poskusu modela I, če v poskusu nastopajo samo fiksni faktorji, o modelu II, če nastopajo samo slučajnostni faktorji tipa II, in o mešanih modelih, če v istem poskusu nastopajo faktorji različnih tipov.

1.10

Določljivih faktorjev, ki v poskusu vplivajo na pojav, je običajno veliko. Od njih je samo nekaj, katerih učinek proučujemo in zavestno ustvarjamo pogoje, da izzovemo njihovo delovanje in merimo njihov učinek. Vsi drugi določljivi, a za raziskavo nepomembni faktorji pa so v splošnem nezaželjeni, ker v glavnem zabrišejo jasnost delovanja proučevanih faktorjev.

Določljive, a za raziskavo nepomembne faktorje obravnavamo različno. Če tak faktor držimo v vsem poskusu na istem nivoju, preide v opredeljujoče pogoje.

Zelo uporabljana tehnika obravnavanja teh faktorjev je kontrola nad učinkom. Ta je izvedena preko načrta poskusa. Upoštevanje nepomembnih faktorjev v poskusu ima za cilj izločitev vpliva preko načrta poskusa. Ta tehnika je osnovna prvina načrtovanega poskusa. Tako z metodo blokov izločamo vpliv enega tipološkega faktorja, z latinskimi kvadrati uspemo kontrolirati vpliv dveh določljivih, a za raziskavo nepomembnih faktorjev ipd.

Tretji način izločitve sistematičnega vplivanja določljivih, a za raziskavo nepomembnih faktorjev, pa je slučajnostno dodeljevanje postopkov poskusnim enotam. S slučajnostnim dodeljevanjem postopkov poskusnemu gradivu spremenimo značaj preostalih določljivih, a za raziskavo nepomembnih faktorjev. Njih sicer ne kontroliramo in njihov učinek ne izločamo, pač pa s slučajnostnim dodeljevanjem uspemo spremeniti značaj teh faktorjev v slučajnostne. Za slučajnostne faktorje sicer ne moremo ugotavljati individualnega učinka, njihov sumarni učinek pa moremo meriti preko standardnega pogreška. Z vključevanjem nepomembnih faktorjev v slučajnostni pogrešek sicer povečujemo nezanesljivost rezultatov o učinku proučevanih faktorjev, vendar je ta pot še vedno korektnejša kot pa, da bi ti faktorji sistematično pačili rezultate poskusa. Na srečo pa imamo možnost zmanjševanja učinka slučajnostnih faktorjev z nadaljno prvino statistično načrtovanega poskusa s ponavljanjem poskusa. Izkoriščanje vseh navedenih prvin poskusa s ciljem, da v danih okoliščinah čim jasnejše pokažemo učinek vpliva proučevanih faktorjev, daje ustrezen načrt poskusa.

1.11 Postopek, kompletna množica postopkov.

V statistično načrtovanih poskusih po pravilu vključujemo v raziskavi več faktorjev hkrati. Pri izvajanju poskusa in pri statistični obdelavi podatkov poskusa upoštevamo v tem primeru vse kombinacije posameznih vrednosti nastopajočih faktorjev.

Iz faktorja: količina določenega gnojila s štirimi vrednostmi G_1 G_2 G_3 G_4 in faktorja: agrotehnična mera s tremi vrednostmi T_1 T_2 T_3 sestavimo $4 \times 3 = 12$ kombinacij vrednosti proučevanih faktorjev gnojilo - agrotehnična mera.

G_1T_1 G_1T_2 G_1T_3 G_2T_1 G_2T_2 G_2T_3 G_3T_1 G_3T_2 G_3T_3 G_4T_1 G_4T_2 G_4T_3

Vsako izmed možnih kombinacij vrednosti faktorjev imenujemo <u>postopek</u>, skupnost vseh možnih postopkov pa <u>kompletno množico postopkov</u>. V našem primeru sestoji kompletna množica postopkov iz $A = 12$ individualnih postopkov. Kombinacijo faktorjev v postopke moremo smatrati za samostojen kombiniran faktor gnojilo-agro mera GT z $A = 12$ vrednostmi-<u>postopki</u>.

Analogno ima poskus 2^6, v katerem proučujemo šest faktorjev s po dvema vrednostima, skupno 64 različnih postopkov. Vsak postopek pa sestoji iz kombinacije po šest faktorjev.

Ni nujno, da poskus vsebuje vse postopke kompletnega sistema postopkov. V poskus vključujemo včasih nekatere kombinacije različnih faktorjev, pač tiste, ki nas v konkretni raziskavi zanimajo. Tako bi mogli iz zgornjega kompletnega sistema postopkov, sestaviti okrnjen sistem postopkov: npr.

G_1A_1 G_2A_1 G_2A_2 G_3A_2 G_3A_3 G_4A_3

Seveda tak okrnjen sistem kombinacij faktorjev nima možnosti take analitične obdelave kot jo ima popoln sistem postopkov.

1.12

Odvisno od cilja poskusa dostikrat skupnost postopkov niti ne sestavljajo vse možne kombinacije vrednosti faktorjev. Tako morejo biti faktorji P, T, R kombinirani v sistem postopkov v obliki P + T. R. V tem primeru sistem postopkov sestoji iz faktorja P z vsemi njegovimi vrednostmi in iz kombinacij faktorjev T in R. Če je npr. vsak izmed proučevanih faktorjev na dveh nivojih, imamo v celoti šest postopkov:

$A_1 = P_1 \quad A_2 = P_2 \quad A_4 = TR_{11} \quad A_4 = TR_{12} \quad A_5 = TR_{21} \quad A_6 = TR_{22}$

Če gre za večfaktorski poskus, običajno vanj vključimo popoln sistem postopkov. V tem primeru dobimo s primerno analizo kompleksne in popolne informacije o delovanju vseh proučevanih faktorjev. Včasih pa v poskus ne vključimo vseh možnih postopkov. Razlogi za to so lahko različni. Ker število postopkov z kombiniranjem velikega števila faktorjev hitro raste, je uvedba otežkočena zaradi velikega obsega poskusa. Z velikim številom postopkov pa more postati gradivo tako heterogeno, da ogroža možnost izvrednotenja rezultatov poskusa. Tako govorimo o <u>frakcionalnih</u> poskusih, če izmed vseh možnih kombinacij postopkov vključimo v poskus samo nekatere, ali o poskusih v <u>delnih blokih,</u> če v določenih blokih za nekatere faktorje vključimo vse kombinacije faktorjev, a za nekatere faktorje samo posamezne vrednosti. V tehniki delnih blokov so npr. vrednosti faktorjev P(3) R(2) T(2) kombinirane v naslednji shemi:

Blok 1 $\quad PRT_{111} \quad PRT_{112} \quad PRT_{121} \quad PRT_{122}$

Blok 2 $\quad PRT_{211} \quad PRT_{212} \quad PRT_{221} \quad PRT_{222}$

Blok 3 $\quad PRT_{311} \quad PRT_{312} \quad PRT_{321} \quad PRT_{322}$

Faktorja P in T sta v vsakem izmed blokov kombinirana v vseh kombinacijah, medtem ko so vrednosti faktorja P povezane z blokom.

1.13

Če je v posameznih blokih povezanih le nekaj vrednosti enega faktorja govorimo o <u>nepopolnih blokih</u> v pravem smislu besede. Nepopolni bloki pridejo v poštev posebno v raziskavah, kjer ima en faktor (npr. sorta določene kulture ipd.) toliko različnih vrednosti, da je težko ustvariti homogeno poskusno gradivo za tako veliko število vrednosti enega znaka.

Enostaven primer nepopolnih blokov je npr. naslednja skupina:

Blok 1	A_1	A_2	A_3		A_1	A_2	A_3	
Blok 2		A_2	A_3	A_4	A_1	A_2		A_4
Blok 3	A_1	A_2		A_4	A_1		A_3	A_4
Blok 4	A_1		A_3	A_4		A_2	A_3	A_4

Pri delnih blokih so posamezne vrednosti določenega faktorja združene z bloki. V tem primeru je pri enem samem poskusu čisti učinek tega faktorja združen z blokom. Poseben razpored postopkov večfaktorskega poskusa pa omogoča, da z blokok ne združimo glavni učinek določenega faktorja, temveč vzajemni učinek oziroma interakcijo več faktorjev. Ker so interakcije višjih stopenj manj pomembne in dobivajo vse bolj značaj slučajnostnih vplivov, je ta rešitev dostikrat priporočljivejša kot delni bloki, če je le dana možnost izvedbe takega poskusa.

Formiranje postopkov poskusa 2^3 v dva nepopolna bloka, v katerih je z bloki združen učinek interakcije vseh treh faktorjev hkrati, je dana s shemo

Blok 1	PRT_{111}	PRT_{221}	PRT_{212}	PRT_{122}
Blok 2	PRT_{211}	PRT_{121}	PRT_{112}	PRT_{222}

1.14 P o n o v i t e v p o s t o p k a, p o n o v i t e v p o s k u s a.

Navedli smo že, da je ponovitev poskusa eden izmed osnovnih prvin statističnega načrtovanja poskusov. Ponovitev postopka je po pravilu ponovitev izvedbe poskusa oziroma postopka na različnih enotah poskusnega gradiva. Tako moremo npr. poskus o produktivnosti odvisnosti produktivnosti dela od treh načinov dela N_1 N_2 N_3 izvesti v dveh ponovitvah, če po vsakem načinu delata po dva/poskusnega gradiva (šest delavcev) na slučajnosten način izbrana delavca. Enako je raziskava o ocenjevanju kakovosti petih raznovrstnih pijač izvedena v treh ponovitvah, če izmed petnajstih poskuševalcev po trije na slučajnosten način izbrani preskuševalci ocenjujejo posamezno izmed petih pijač. V smislu gornje opredelitve moremo v poskusu posamezne postopke izvesti v enakem ali različnem številu ponovitev. Tako bi moglo po načinu N_1 delati $n_1 = 5$, po načinu N_2 $n_2 = 3$ in po načinu N_3 $n_3 = 8$ delavcev. Iz vsebinskih, a predvsem iz tehničnih razlogov vzamemo v posameznih primerih vsak postopek v poskusu z enakim številom ponovitev $\bar{n}$, če ni nekega posebnega razloga, da to pravilo kršimo. V tem primeru rečemo, da je bil poskus izveden v $\bar{n}$ ponovitvah.

Poseben poudarek pri opredelitvi ponovitve postopka je v tem, da mora biti postopek ponovljen na različnih enotah, če naj bo učinek ponovitve tak, kakor ga od ponovitve pričakujemo. Tako v primeru poskusa o produktivnosti dela ne moremo govoriti o ponovitvi poskusa, če bi po planu predvideli, da bi vsak izmed treh izbranih delavcev delal po enem in istem načinu dvakrat, da bi prvi delavec delal dvakrat po načinu N_1, drugi dvakrat po načinu N_2, tretji delavec pa dvakrat po načinu N_3. Ker gre v tem primeru za ponavljanje poskusa na istih poskusnih enotah (istih delavcih) je poprečna produktivnost po vsakem izmed načinov izraz učinka načina dela in individualnih sposobnosti delavca, ki dela po posameznem načinu. Če število merjenj po posameznem načinu še tako večamo, se bodo slučajnostni vplivi zaradi zakona o velikih številih zmanjševali, učinek individualne sposobnosti posameznega delavca pa se ne bo izravnaval, ker dela v vseh ponovitvah isti delavec. Šele pravo ponavljanje poskusa, pri katerem na slučajnosten način izberemo $\bar{n}$ delavcev, ki delajo po načinu N_1, enako število na slučajnosten način izbranih delavcev, ki delajo po načinu N_2 in isto za način N_3, vključimo individualne sposobnosti kot slučajnostno spremenljivko v poskus in se v poprečju izravnavajo tako, da pride v poprečju bolj in bolj do izraza učinek raziskovanega faktorja, tj. načina dela.

Še bolj pride do izraza neuspešnost ponovitve na istih poskusnih enotah v primeru ocenjevanja kakovosti pijač. Pri večkratnem poskušanju določene vrste pijače pride še v večji meri do izraza individualni okus preskuševalca. Razlike med ocenami merijo le trenutno menjanje okusa iste osebe, medtem ko so pri pravi ponovitvi razlike v ocenah predvsem izraz različnih okusov ljudi, kar je v konkretnih raziskavah zelo pomembno, če hočemo ugotavljati poprečno oceno kakovosti pijače.

V statistično načrtovanih poskusih sicer izvajamo kvazi-ponovitve na istih poskusnih enotah, vendar imajo te ponovitve v osnovi drugo nalogo in funkcijo kot pa prave ponovitve tj. ponovitve na različnih poskusnih enotah. Ponovitve postopka na istih enotah imajo namen, da zanesljiveje izvrednotimo rezultate poskusa v tem smislu, da iz individualnih meritev izločimo vpliv slučajnostnih faktorjev v ožjem smislu. Ker individualno meritev oziroma rezultat nadomestimo z več meritvami in te meritve predstavljajo vzorec iz umišljene populacije poskusa pod enakimi pogoji, s kvazi-ponovitvami na istih poskusnih enotah prispevamo sicer k zmanjšanju slučajnostnih vplivov predvsem pa na ta način ocenimo vpliv slučajnostnih faktorjev v poskusu.

1.15 Shema vključevanja prvin načrta v statistični načrt poskusa.

Osnovne prvine statističnega načrta poskusa: slučajnostnost dodeljevanja, ponavljanje postopkov, kontrola določljivih faktorjev v povezavi z različnimi faktorji, ki vplivajo na kriterialni znak, moremo strniti v naslednji shemi:

Faktorji / Primer proučevanja	Določljivi			Nedoločljivi
	opredeljujoči	vsebinsko		
		pomembni	nepomembni	slučajnostni
produktivnost dela	moški kvalificiran	način dela, staž	starost	kondicija ob poskusu
proizvodnja	surovina, stroj	način obdelave	delavec	
gnojilni poskus	nadmorska višina parcel poskusa, kultura	vrste gnojila	kraj poskusnih parcelic	mikrotalni pogoji na poskusnih parcelicah
psihološka raziskava	testni pogoji	spol, starost	predhodna izobrazba	razpoloženje ob testiranju
raziskava trga	mesto, izdelek	vrsta trgovine, način prodaje	kvalifikacija prodajalca, dan prodaje	razpoloženje kupca razpoloženje prodajalca
pomen in obravnavanje faktorjev	določajo populacijo, ki jo proučujemo prispevajo k homogenosti	raziskujemo vpliv z relativnimi primerjavami predstavljajo vsebinski del raziskave	vpliv nekaterih faktorjev izločamo z načrtom poskusa	njihov vpliv zmanjšamo s ponovitvijo poskusa Vpliv neizločenih s slučajnostno dodelitvijo postopkov prevedemo v slučajnostne faktorje

2. VZORČENJE. OCENJEVANJE. PRESKUŠANJE DOMNEV.

2.1 Statistika se bavi s proučevanjem množičnih pojavov. Preko zbiranja, obdelave in analize podatkov o populacijah. Značilnosti populacij podajamo z njihovimi parametri, ki so izraz opredeljujočih pogojev. Tako je npr. poprečna poraba določenega izdelka na gospodinjstvo rezultat kategorije gospodinjstva, višine dohodka, kraja in časa, za katerega veljajo podatki. S primerjavo parametrov za dve li več populacij, ki so različno opredeljene, analiziramo vpliv razlik med populacijami. Razlika v poprečjih porabe za dve skupini prebivalstva je izraz vpliva skupine na prodajo, razlika v poprečjih porabe za več zaporednih časovnih razdobij za isto skupino, podajajo dinamiko porabe ipd.

Eden najpogostnejših parametrov, preko katerega proučujemo vpliv spremenjenih pogojev, je vsekakor poprečje. Iz individualnih podatkov pridemo do poprečja s predpostavko, da se odkloni individualnih podatkov od poprečja v vsoti uničijo

$$\sum(y - M) = 0 \qquad (2.1)$$

iz česar sledi, da je

$$M = \frac{1}{N}\sum y \qquad (2.2)$$

Analitično najbolj nesporen sintetičen pokazovalec velikosti individualnih odklonov od poprečja je poprečen kvadratičen odklon ali varianca

$$\sigma_y^2 = \frac{1}{N}\sum(y - M_y)^2 \qquad (2.3)$$

Z varianco σ_y^2 in iz nje izpeljanim standardnim odklonom σ_y

$$\sigma_y = \sqrt{\sigma_y^2} \qquad (2.4)$$

merimo jakost individualnih, v posebnih primerih slučajnostnih vplivov.

2.2 Individualne vrednosti y_i se okrog aritmetične sredine ali popreč ja porazdeljujejo različno, odvisno od individualnih vplivov. To variiranje prikažemo s frekvenčno porazdelitvijo vrednosti. Frekvenčna porazdelitev pove pogostnost vrednosti v populaciji.

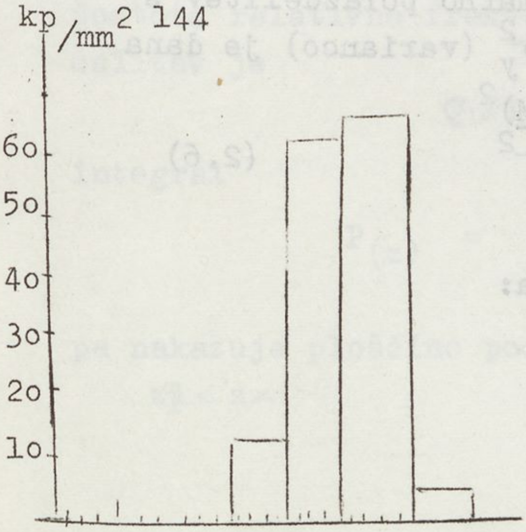
Frekvenčno porazdelitev podamo za končne populacije s frekvenčno porazdelitvijo, v kateri pokažemo, koliko enot - ali kolik del populacije ima vrednosti v posameznih razredih - grupi vrednosti. Za primer frekvenčne porazdelitve vzemimo trdnost patentirane jeseniške žice.

Tabela 2.1 Frekvenčna porazdelitev za trdnost patentirane jeseniške žice.

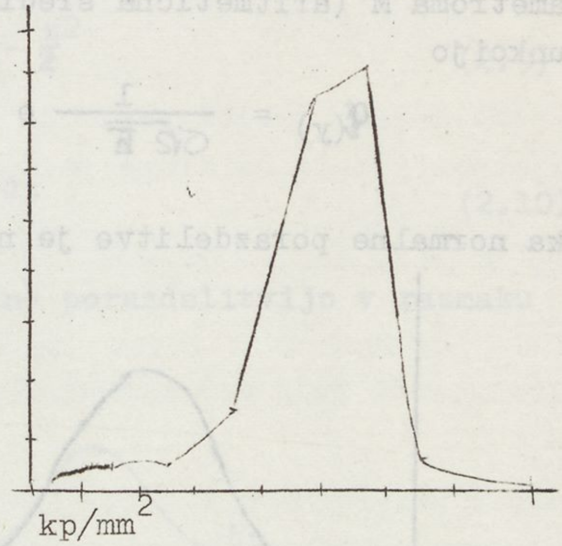
kp/mm^2	f_k	$f_k\%$
144 - 147	2	1.3
148 - 151	2	1.3
152 - 155	16	10.5
156 - 159	63	41,2
160 - 163	67	43.7
164 - 167	3	2.0
	N = 153 153	100.0

Frekvenčno porazdelitev prikažemo grafično s histrogramom ali s poligonom kot kažeta sliki 2.1 in 2.2

)



Sl.2.1 Histrogram



Sl.2.2. Frekvenčni poligon.

2.3 Normalna porazdelitev.

Pogosto se v praksi pojavljajo porazdelitve, za katere se vrednosti goste okrog poprečja simetrično, unimodolno (en sam vrh) njihovi grafikoni imajo zvonasto obliko. Take porazdelitve so dostikrat izraz homogenih populacij, na katere razen splošnih opredeljujočih faktorjev vplivajo samo slučajnostni vplivi. Tako zakonitost pojavljanja moremo ponazoriti z enostavnim linearnim modelom

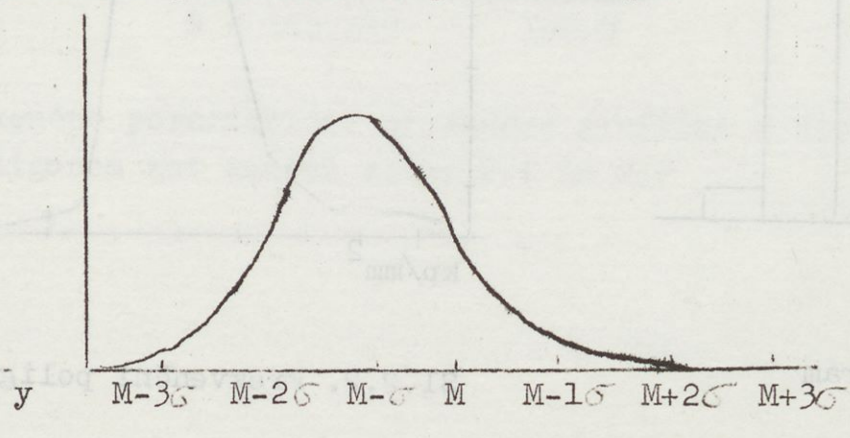
$$y_i = M + e_i \qquad e_i = :N(0, \sigma_e^2) \tag{2.5}$$

po katerem so posamezne vrednosti vsota M, ki je rezultat splošnih faktorjev in je za vse enote isti in slučajnostne komponente e_i, ki se spreminja po zakonitosti normalne porazdelitve z aritmetično sredino $M_e = 0$, in varianco σ_e^2.

Gostota relativne frekvence $\varphi(y)$ za normalno porazdelitev s parametroma M (aritmetična sredina) in σ_y^2 (varianco) je dana s funkcijo

$$\varphi(y) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(y-M)^2}{2\sigma^2}} \tag{2.6}$$

Slika normalne porazdelitve je naslednja:



Slika 2.3 Normalna porazdelitev y : N(M,σ)

Vse normalne porazdelitve so med seboj podobne, v odvisnosti od parametra M pa so pomaknjene v desno ali levo, v odvisnosti od σ pa je njihova variacija večja ali manjša.

Posebno vlogo igra standardizirana normalna porazdelitev

$$z = : N(0,1) \qquad (2.7)$$

za katero je $M_z = 0$ in $\sigma_z = 1$

Ta porazdelitev je tabelirana.

Preko standardiziranega z odklona

$$z = \frac{y - M}{\sigma} \qquad (2.8)$$

moremo vsako normalno porazdelitev transformirati v normalno porazdelitev in obratno.

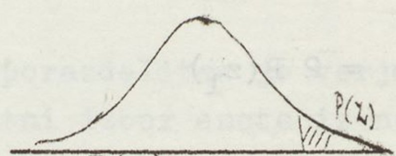
Gostota relativne frekvence za standardizirano normalno porazdelitev je

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \qquad (2.9)$$

integral

$$P_{(z)} = \int_{z_1}^{\infty} \varphi(z)\, dz \qquad (2.10)$$

pa nakazuje ploščino pod normalno porazdelitvijo v razmaku $z_1 < z < \infty$



Slika 2.4 Slika P(z)

kot funkcijo spodnje meje z.

2.4 Normalna porazdelitev kot verjetnostna porazdelitev.

Če iz normalno porazdeljene populacije slučajnostno (vsaka enota ima enako šanso, da je izbrana v izbor) izberemo enoto, je verjetnost, da je standardiziran odklon z-odklon večji od z enak P(z), kar pomeni, da je pri ponavljanju slučajnostnega izbora delež enot z vrednostjo večjo od z enak P(z).

Za nekaj karakterističnih vrednosti naznačimo odnose med z in P(z).

Tabela 2.2 Odnosi med z in P(z) za normalno porazdelitev

Z	$-\infty$	000	·674	1.000	1.645	1.960	2.000	2.376	2.576	3.000	3.070	3.2
P(z)	1.0000	·5000	·2500	·1587	·0500	·0250	·0227	·0100	·0050	·0013	·0010	·00

Iz P(z) moremo dobiti posredno tudi relativne frekvence ali verjetnosti tudi za druge razmake:

$$
\begin{aligned}
P_r\,(< /z/ < \pm z_1) &= 1-2P(z) \\
P_r\,(z < z_1) &= 1-P(z_1) \\
P_r\,(|z| > z_1) &= 2\,P(z_1) \qquad (2.11) \\
P_r\,(0 < z < z_1) &= P(z) - 0.5
\end{aligned}
$$

Tako moremo npr. z verjetnostjo

$$P_r\,(|z| < 1) = 1 - 2P(+1) = 1 - 2\ .\ ·1587 = ·6826$$

pričakovati, da za slučajnostno izbrano enoto in normalne porazdelitve vrednost standardiziranega z-odklona absolutno ni večja od 1.

Verjetnost, da za slučajnostno izbrano enoto iz normalne populacije absolutna vrednost za standardiziran odklon $z = \frac{y-M}{\sigma}$ ni večji od 1,96 je

$$P_r(|z|<1{,}96) = 1 - 2\,P(z=1{,}96) = 0.95$$

Ta verjetnost je razmeroma zelo velika šansa, da poskus (stvaren izbor) potrdi trditev: iz normalno porazdeljene populacije slučajnostno izbrana enota ima absolutno vrednost standardiziranega z-odklona manjšo kot 1.96 je 19:1, da se ne zgodi. Ker je ta šansa velika, računamo, da se bo napovedani dogodek zgodil, čeprav ni gotov, ampak le zelo verjeten. Take izjave delamo z določenim tveganjem. Tveganje je tem večje, čim večja je verjetnost, da se napovedani dogodek ne zgodi. Zato tveganje definiramo kot verjetnost, da se napovedan dogodek ne zgodi.

Tveganje za dogodek A, ki se zgodi z verjetnostjo Pr(A), je torej verjetnost, da se dogodek ne zgodi.

$$P_r(\bar{A}) = 1 - P_r(A) \qquad (2.12)$$

Za zgornji primer je tveganje izjave, da je $|z|<1{,}96$ enako $Pr(\bar{A}) = 1 - Pr(|z|<z_1) = 1 - 0.95 = 0.05$.

Statistični sklepi, napravljeni na osnovi slučajnostno izbranih delnih populacij, so vsi verjetnostni in veljajo z večjim ali manjšim tveganjem.

Standardizirana normalna porazdelitev je verjetnostna porazdelitev za dogodek: slušajnostni izbor enote iz normalno porazdeljene populacije.

2.5 Slučajnostni izbor je oznaka za izbor, v katerem ima vsaka enota enako šanso, da je izbrana. Slučajnostni izbor izvedemo različno. Običajno se poslužujemo tablic slučajnostnih števil, ki omogočajo slučajnostni izbor iz populacije, v kateri so enote označene z zaporednimi številkami. Iz tablic vzamemo kot slučajnostno število skupino iz toliko zaporednih števil, kolikor mestno je število N, ki označuje obseg populacije. V tabeli 2.3 je prikazana ena stran slučajnostnih števil. Če ima populacija npr. N = 800 enot, začnemo s prvo skupino tromestnih števil 535 in nadaljujemo z nadaljnjimi skupinami 491, 420. Številka 847 izpade, ker je večja kot N = 800.
Slučajnostne številke dobimo tudi s prekrivanjem tako, da se od tromestne do tromestne številke ne pomikamo za tri cifre temveč po eno. Tako dobimo
535 354 549 491 914 142 420 208 •084 •847
Od teh desetih slučajnostnih števil je v našem primeru uporabnih osem, ker sta 914 in 847 večji kot je N = 800.

Pseudoslučajnostne številke, ki dobro služijo v praktičnem primeru kot slučajnostne številke, dobimo jih na računalniku po enostavnem principu . Poljubno začetno slučajnostno število S_o pomnožimo z večmestnim številom K.
Zadnjih k mest produkta

$$S_o \cdot K$$

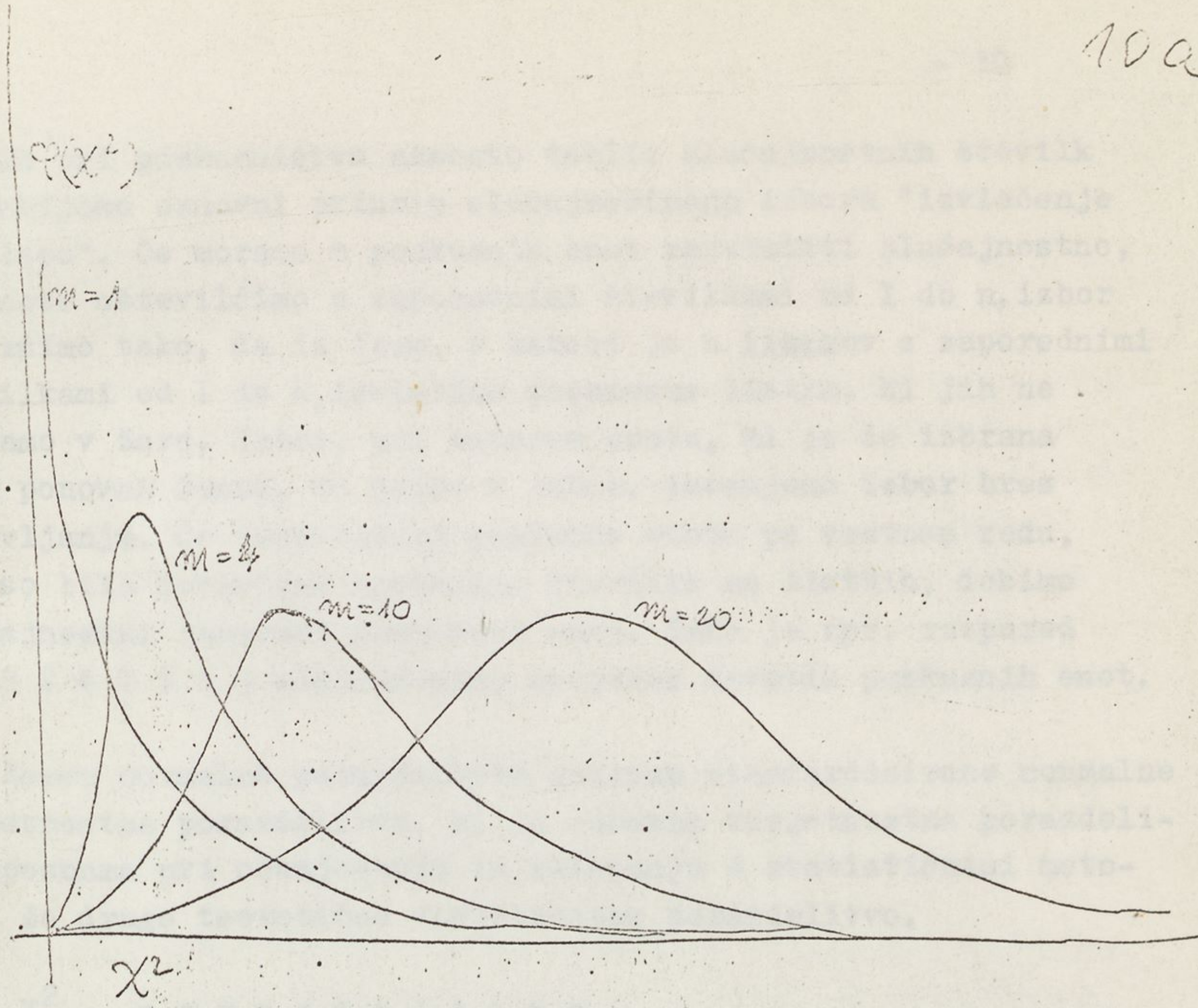
predstavlja naslednjo slučajnostno številko S_1. Če po istem načinu postopamo naprej, dobimo iz k -mestnega ostanka produkta

$$S_1 k$$

naslednje slučajnostno število S_2.

TABELA 2.3 Slučajnostna števila

5354	9142	0847	5393	5416	6505	7156	5634	9703	6221
0905	6986	9396	3975	9255	0537	2479	4589	0562	5345
1420	0470	8679	2328	3939	1292	0406	5428	3789	2882
3218	9080	6604	1813	8209	7039	2086	3386	4437	3798
9697	8431	4387	0622	6893	8788	2320	9358	5904	9539
0912	4964	0502	9683	4636	2861	2876	1273	7870	2030
4636	7072	4868	0601	3894	7182	8417	2367	7032	1003
2515	4734	9878	6761	5636	2949	3979	8650	3430	0635
5964	0412	5012	2369	6461	0678	3693	2928	3740	8047
7848	1523	7904	1521	1455	7089	8094	9872	0898	7174
5192	2571	3643	0707	3434	6818	5729	8614	2498	4129
8438	8325	9886	1805	0226	2310	3675	5058	2515	2388
8166	6349	0319	5436	6838	2460	6433	0644	7428	8556
9158	8263	6504	2562	1160	1526	1816	9690	1215	9590
6061	3525	4048	0382	4224	7148	2859	6526	5340	4064
5407	2818	0520	5941	1740	5140	9844	2847	1502	0763
0469	0435	2858	7116	3297	8454	5146	9803	1694	7949
7805	8428	5745	8141	8465	4795	1895	4487	2323	1068
7294	1214	0170	9643	7891	7304	8278	2315	7139	5594
5480	2843	8903	3828	1717	6312	0384	6552	1200	7264
1017	0106	1414	9736	3886	4753	3589	3864	0073	3626
0858	1727	3020	1831	2878	2838	1319	2199	6457	5798
8396	8903	2156	1031	6182	5094	1931	9188	1672	1510
7813	4209	5295	0605	9080	6940	9657	3423	2191	3636
2712	2516	0968	7526	2176	4057	9023	2327	4311	0281
7141	7871	2878	2990	3907	8375	6005	9452	7702	9468
2418	9661	0436	1223	9708	9354	0707	4238	0756	2190
5230	6208	2742	1087	9639	6813	1963	2620	8913	7777
3517	1376	7866	6584	6381	0218	1101	3192	5965	5250
0319	0951	3976	6372	3518	1859	9038	3474	5150	3621
7526	7460	5644	8640	0643	0916	3238	0177	2592	0264
5172	1998	6030	6677	8827	7821	9933	9523	4563	7391
2023	0950	1896	4729	1789	1111	1157	0266	0438	7535
3632	3816	4575	0738	4923	4131	0819	3361	3992	6702
0761	2838	6166	8534	5353	5737	1204	2325	2036	4714



Slika 2.5 χ^2 - porazdelitev

Včasih pri poskusništvu namesto tablic slučajnostnih števil uporabljamo osnovni princip slučajnostnega izbora "izvlačenje na slepo". Če moramo n poskusnih enot razvrstiti slučajnostno, te enote oštevilčimo z zaporednimi številkami od 1 do n, izbor pa vršimo tako, da iz žare, v kateri je n listkov z zaporednimi številkami od 1 do n, izvlačimo posamezne listke, ki jih ne vračamo v žaro. Izbor, pri katerem enota, ki je že izbrana nima ponovne šanse, da pride v izbor, imenujemo izbor brez ponavljanja. Če razporedimo poskusne enote po vrstnem redu, kot so bile izvlečene zaporedne številke na listkih, dobimo slučajnostni razpored poskusnih enot. Tako je npr. razpored 8 1 3 2 4 7 5 6 9 slučajnostni razpored devetih poskusnih enot.

2.6 Razen normalne porazdelitve oziroma standardizirane normalne verjetnostne porazdelitve, ki je osnovna verjetnostna porazdelitev, poznamo pri ocenjevanju in sklepanju s statističnimi metodami še druge teoretične verjetnostne porazdelitve.

2.7 X^2 - porazdelitev.

Iz verjetnostnega računa vemo, da je funkcija slučajnostnih spremenljivk tudi slučajnostna spremenljivka, ki ima svojo verjetnostno porazdelitev. Tako se vsota kvadratov m med seboj neodvisnih slučajnostnih spremenljivk z_i $i = 1 \dots m$, ki se porazdeljujejo standardizirano normalno

$$z_1^2 + z_2^2 + \dots z_m^2 = \sum_{i=1}^{m} z_i^2 = : X^2(m) \qquad (2.13)$$

porazdeljuje v porazdelitvi X^2 (hi kvadrat) z m stopinjami prostosti.

Število stopinj prostosti se v splošnem ravna po številu neodvisnih slučajnostnih spremenljivk v funkciji slučajnostnih spremenljivk. Ker so v izrazu

$$X^2 = \sum_{i=1}^{m} z_i^2$$

slučajnostne spremenljivke z_i med seboj neodvisne, je število stopinj prostosti za X^2 enako m.

Če pa je npr. n sicer neodvisnih spremenljivk vezanih na stalno aritmetično sredino, je število stopinj prostosti n - 1, ker je ena vrednost pogojena z n-1 vrednostmi. Podobno ima sistem n slučajnostnih spremenljivk, ki so razporejene v k grup m = n - k stopinj prostosti, če so grupni podatki vezani na stalne aritmetične sredine po grupah.

Gostota verjetnosti za X^2-porazdelitev je dana v funkciji

$$\varphi(X^2) = C_{X^2}\,(X^2)^{\frac{m-2}{2}}\, e^{-\frac{X^2}{2}} \qquad (2.14)$$

Z njo je pri danem m v celoti podana X^2- porazdelitev.

Če proučimo funkcijo za gostoto relativne frekvence, spoznamo, da je verjetnostna porazdelitev za X^2 v splošnem asimetrična, da pa se stopnja asimetrije pri večjem m manjša in X^2 bolj in bolj prehaja v simetrično unimodalno zvonasto - skratka normalno porazdelitev. Izkaže se, da se izraz

$$\sqrt{2X^2} - \sqrt{2m-1} = z = :N(0,1) \qquad (2.15)$$

asimptotično porazdeljuje v standardizirani normalni porazdelitvi. Uporaben približek dobimo za m >30.

Z uporabo zgornjega izraza dobimo, da je

$$X_P^2 = \frac{1}{2}(z_P + \sqrt{2m-1})^2 \qquad (2.16)$$

Verjetnostim P ustrezne vrednosti dobimo iz tabele o normalni porazdelitvi, ki je v skrčeni obliki dana za karakteristične vrednosti v naslednji tabeli.

Tabela 2.4 Tabela z_P .

P	.99	.95	.50	.30	.20	.10	.05	.02	.01	.001
Z_P	-2.3263	-1.6449	0.0000	0.5244	0.8416	1.2816	1.6449	2.0537	2.3263	3.0902

Ker se da poljubna normalna porazdelitev z enostavno linearno transformacijo $z = \frac{y-M}{\sigma}$ prevesti v standardizirano normalno, so s standardizirano normalno porazdelitvijo dane vse druge.

Z X^2- porazdelitvijo pa ni tako. Zato bi morali imeti tabelirano X^2-porazdelitev za vsako stopinjo prostosti m posebej.

Ker pa potrebujemo praktično le X_P^2 za nekatere karakteristične vrednosti P je možno podati te vrednosti za vsak m v eni sami vrstici. Te vrednosti za P = 0.99 , 0.95, 0.50 so podane v tabeli 2.5 do m = 30 za X^2-porazdelitev z m > 30 pa dobimo ustrezne vrednosti za X_P^2 preko zveze s standardiziranim z-odklonom.

Tabela 2.5 Kritične vrednosti za X^2 - porazdelitev.

2.8 Studentova t-porazdelitev.

Za med seboj neodvisni slučajnostni spremenljivki z = :N(0,1) in X^2 = :X^2(m), od katerih se z porazdeljuje standardizirano normalno, X^2 pa v X^2-porazdelitvi z m stopinjami prostosti, se izraz

$$\frac{z}{\sqrt{X^2/m}} = : t(m) \tag{2.17}$$

porazdeljuje v porazdelitvi, ki jo imenujemo Studentovo t-porazdelitev. Zanjo je število stopinj prostosti m, gostota verjetnosti pa je dana s funkcijo

$$\varphi(t) = C_t \left(1 + \frac{t^2}{m} \right)^{-\frac{m+1}{2}} \tag{2.18}$$

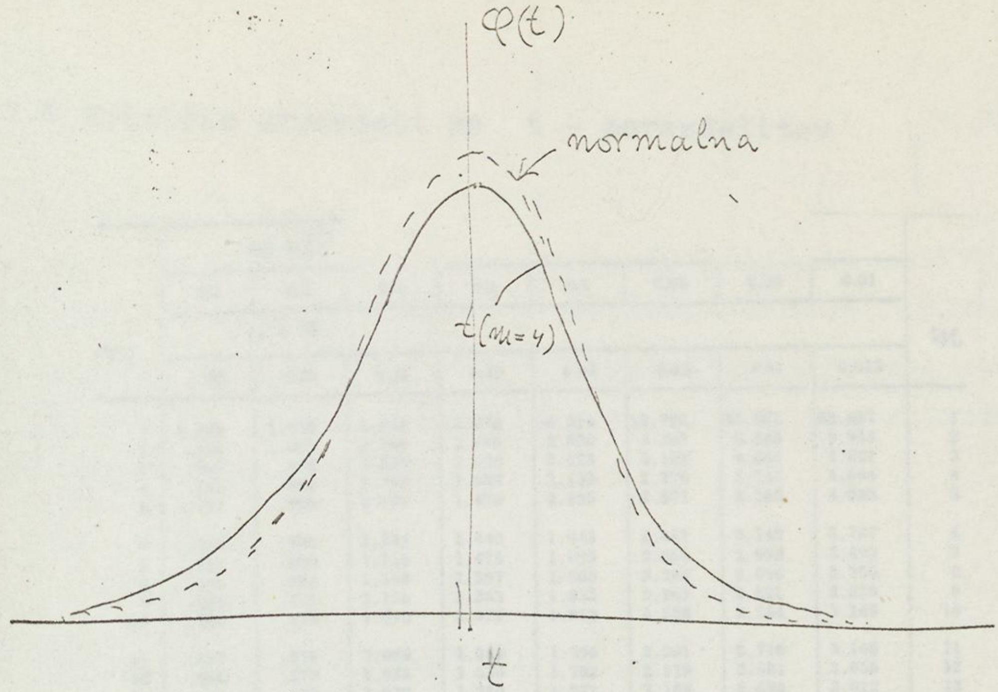
Kot kaže funkcija za gostoto verjetnosti $\varphi(t)$ in slika, je t-porazdelitev unimodalna, simetrična in zvonasta. Z večanjem m t-porazdelitev preide v standardizirano normalno porazdelitev.

Iz istega razloga kot za X^2 so tudi za t-porazdelitev tabelirane le t_P za določene karakteristične vrednosti P.

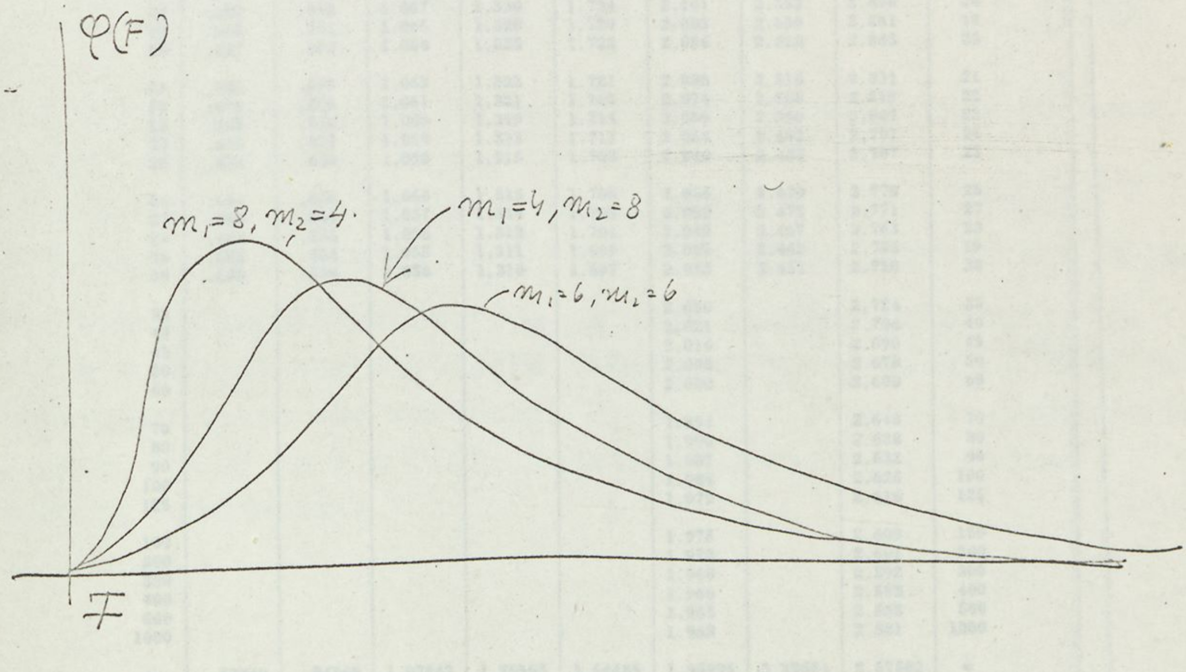
TABELA 2.5

Kritične vrednosti X^2-porazdelitev

m \ P	0.995	0.99	0.975	0.95	0.90	0.80	0.70	0.60	0.50	0.40	0.30	0.20	0.10	0.05	0.025	0.01	0.005	0.001
1	$0.0^4 39$	$0.0^3 16$	$0.0^3 98$	$0.0^2 39$	0.016	0.064	0.148	0.275	0.455	0.708	1.074	1.642	2.706	3.841	5.024	6.635	7.879	10.828
2	0.010	0.020	0.051	0.103	0.211	0.446	0.713	1.022	1.386	1.833	2.408	3.219	4.605	5.991	7.378	9.210	10.597	13.816
3	0.072	0.115	0.216	0.352	0.584	1.005	1.424	1.869	2.366	2.946	3.665	4.642	6.251	7.815	9.348	11.345	12.838	16.266
4	0.207	0.297	0.484	0.711	1.064	1.649	2.195	2.753	3.357	4.045	4.878	5.989	7.779	9.488	11.143	13.277	14.860	18.467
5	0.412	0.554	0.831	1.145	1.610	2.343	3.000	3.655	4.351	5.132	6.064	7.289	9.236	11.070	12.832	15.086	16.750	20.515
6	0.676	0.872	1.237	1.635	2.204	3.070	3.828	4.570	5.348	6.211	7.231	8.558	10.645	12.592	14.449	16.812	18.548	22.458
7	0.989	1.239	1.690	2.167	2.833	3.822	4.671	5.493	6.346	7.283	8.383	9.803	12.017	14.067	16.013	18.475	20.278	24.322
8	1.344	1.646	2.180	2.733	3.490	4.594	5.527	6.423	7.344	8.351	9.524	11.030	13.362	15.507	17.535	20.090	21.955	26.125
9	1.735	2.088	2.700	3.325	4.168	5.380	6.393	7.357	8.343	9.414	10.656	12.242	14.684	16.919	19.023	21.666	23.589	27.877
10	2.156	2.558	3.247	3.940	4.865	6.179	7.267	8.295	9.342	10.473	11.781	13.442	15.987	18.307	20.483	23.209	25.188	29.588
11	2.603	3.053	3.816	4.575	5.578	6.989	8.148	9.237	10.341	11.530	12.899	14.631	17.275	19.675	21.920	24.725	26.757	31.264
12	3.074	3.571	4.404	5.226	6.304	7.807	9.034	10.182	11.340	12.584	14.011	15.812	18.549	21.026	23.336	26.217	28.300	32.909
13	3.565	4.107	5.009	5.892	7.042	8.634	9.926	11.129	12.340	13.636	15.119	16.985	19.812	22.362	24.736	27.688	29.819	34.528
14	4.075	4.660	5.629	6.571	7.790	9.467	10.821	12.079	13.339	14.685	16.222	18.151	21.064	23.685	26.119	29.141	31.319	36.123
15	4.601	5.229	6.262	7.261	8.547	10.307	11.721	13.030	14.339	15.733	17.322	19.311	22.307	24.996	27.488	30.578	32.801	37.697
16	5.142	5.812	6.908	7.962	9.312	11.152	12.624	13.983	15.338	16.780	18.418	20.465	23.542	26.296	28.845	32.000	34.267	39.252
17	5.697	6.408	7.564	8.672	10.085	12.002	13.531	14.937	16.338	17.824	19.511	21.615	24.769	27.587	30.191	33.409	35.718	40.790
18	6.265	7.015	8.231	9.390	10.865	12.857	14.440	15.893	17.338	18.868	20.601	22.760	25.989	28.869	31.526	34.805	37.156	42.312
19	6.844	7.633	8.907	10.117	11.651	13.716	15.352	16.850	18.338	19.910	21.689	23.900	27.204	30.144	32.852	36.191	38.582	43.820
20	7.434	8.260	9.591	10.851	12.443	14.578	16.266	17.809	19.337	20.951	22.775	25.038	28.412	31.410	34.170	37.566	39.997	45.315
21	8.034	8.897	10.283	11.591	13.240	15.445	17.182	18.768	20.337	21.991	23.858	26.171	29.615	32.671	35.479	38.932	41.401	46.797
22	8.643	9.542	10.982	12.338	14.041	16.314	18.101	19.729	21.337	23.031	24.939	27.301	30.813	33.924	36.781	40.289	42.796	48.268
23	9.260	10.196	11.688	13.091	14.848	17.187	19.021	20.690	22.337	24.069	26.018	28.429	32.007	35.172	38.076	41.638	44.181	49.728
24	9.886	10.856	12.401	13.848	15.659	18.062	19.943	21.652	23.337	25.106	27.096	29.553	33.196	36.415	39.364	42.980	45.558	51.179
25	10.520	11.524	13.120	14.611	16.473	18.940	20.867	22.616	24.337	26.143	28.172	30.675	34.382	37.652	40.646	44.314	46.928	52.620
26	11.160	12.198	13.844	15.379	17.292	19.820	21.792	23.579	25.336	27.179	29.246	31.795	35.563	38.885	41.923	45.642	48.290	54.052
27	11.808	12.879	14.573	16.151	18.114	20.703	22.719	24.544	26.336	28.214	30.319	32.912	36.741	40.113	43.194	46.963	49.645	55.476
28	12.461	13.565	15.308	16.928	18.939	21.588	23.647	25.509	27.336	29.249	31.391	34.027	37.916	41.337	44.461	48.278	50.993	56.892
29	13.121	14.256	16.047	17.708	19.768	22.475	24.577	26.475	28.336	30.283	32.461	35.139	39.087	42.557	45.722	49.588	52.336	58.302
30	13.787	14.953	16.791	18.493	20.599	23.364	25.508	27.442	29.336	31.316	33.530	36.250	40.256	43.773	46.979	50.892	53.672	59.703
31	14.458	15.655	17.539	19.281	21.434	24.255	26.440	28.409	30.336	32.349	34.598	37.359	41.422	44.985	48.232	52.191	55.003	61.098
32	15.134	16.362	18.291	20.072	22.271	25.148	27.373	29.376	31.336	33.381	35.665	38.466	42.585	46.194	49.480	53.486	56.328	62.487
33	15.815	17.073	19.047	20.867	23.110	26.042	28.307	30.344	32.336	34.413	36.731	39.572	43.745	47.400	50.725	54.776	57.648	63.870
34	16.501	17.789	19.806	21.664	23.952	26.938	29.242	31.313	33.336	35.444	37.795	40.676	44.903	48.602	51.966	56.061	58.964	65.247
35	17.192	18.509	20.569	22.465	24.797	27.836	30.178	32.282	34.336	36.475	38.859	41.778	46.059	49.802	53.203	57.342	60.275	66.619
36	17.887	19.233	21.336	23.269	25.643	28.735	31.115	33.252	35.336	37.505	39.922	42.879	47.212	50.998	54.437	58.619	61.581	67.985
37	18.586	19.960	22.106	24.075	26.492	29.635	32.053	34.222	36.336	38.535	40.984	43.978	48.363	52.192	55.668	59.892	62.883	69.346
38	19.289	20.691	22.878	24.884	27.343	30.537	32.992	35.192	37.335	39.564	42.045	45.076	49.513	53.384	56.895	61.162	64.181	70.703
39	19.996	21.426	23.654	25.695	28.196	31.441	33.932	36.163	38.335	40.593	43.105	46.173	50.660	54.572	58.120	62.428	65.476	72.055
40	20.707	22.164	24.433	26.509	29.051	32.345	34.872	37.134	39.335	41.622	44.165	47.269	51.805	55.758	59.342	63.691	66.766	73.402



Sl. 2.6 Studentova t - porazdelitev



Slika 2.7 F-porazdelitev

Tabela 2.6 Kritične vrednosti za t - porazdelitev

m	$\alpha = 2P$								m
	0.5	0.4	0.3	0.2	0.1	0.05	0.02	0.01	
	$\alpha = P$								
	0.25	0.20	0.15	0.10	0.05	0.25	0.01	0.005	
1	1.000	1.376	1.963	3.078	6.314	12.706	31.821	63.657	1
2	.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	2
3	.765	.978	1.250	1.638	2.353	3.182	4.541	5.841	3
4	.741	.941	1.190	1.533	2.132	2.776	3.747	4.604	4
5	.727	.920	1.156	1.476	2.105	2.571	3.365	4.032	5
6	.718	.906	1.134	1.440	1.943	2.447	3.143	3.707	6
7	.711	.896	1.119	1.415	1.895	2.365	2.998	3.499	7
8	.706	.889	1.108	1.397	1.860	2.306	2.896	3.355	8
9	.703	.883	1.100	1.383	1.833	2.262	2.821	3.250	9
10	.700	.879	1.093	1.372	1.812	2.228	2.764	3.169	10
11	.697	.876	1.088	1.363	1.796	2.201	2.718	3.106	11
12	.695	.873	1.083	1.356	1.782	2.179	2.681	3.055	12
13	.694	.870	1.079	1.350	1.771	2.160	2.650	3.012	13
14	.692	.868	1.076	1.345	1.761	2.145	2.624	2.977	14
15	.691	.866	1.074	1.341	1.753	2.131	2.602	2.947	15
16	.690	.865	1.071	1.337	1.746	2.120	2.583	2.921	16
17	.689	.863	1.069	1.333	1.740	2.110	2.567	2.898	17
18	.688	.862	1.067	1.330	1.734	2.101	2.552	2.878	18
19	.688	.861	1.066	1.328	1.729	2.093	2.539	2.861	19
20	.687	.860	1.064	1.325	1.725	2.086	2.528	2.845	20
21	.686	.859	1.063	1.323	1.721	2.080	2.518	2.831	21
22	.686	.858	1.061	1.321	1.717	2.074	2.508	2.819	22
23	.685	.858	1.060	1.319	1.714	2.069	2.500	2.807	23
24	.685	.857	1.059	1.318	1.711	2.064	2.492	2.797	24
25	.684	.856	1.058	1.316	1.708	2.060	2.485	2.787	25
26	.684	.856	1.058	1.315	1.706	2.056	2.479	2.779	26
27	.684	.855	1.057	1.314	1.703	2.052	2.473	2.771	27
28	.683	.855	1.056	1.313	1.701	2.048	2.467	2.763	28
29	.683	.854	1.055	1.311	1.699	2.045	2.462	2.756	29
30	.683	.854	1.055	1.310	1.697	2.042	2.457	2.750	30
35						2.030		2,724	35
40						2.021		2.704	40
45						2.014		2.690	45
50						2.008		2.678	50
60						2.000		2.600	60
70						1.994		2.648	70
80						1.990		2.638	80
90						1.987		2.632	90
100						1.984		2.626	100
125						1.979		2.616	125
150						1.976		2.609	150
200						1.972		2.601	200
300						1.968		2.592	300
400						1.966		2.588	400
500						1.965		2.568	500
1000						1.962		2.581	1000
∞	.67449	.84162	1.03643	1.28155	1.64485	1.95996	2.32634	2.57582	∞

2.9 F- p o r a z d e l i t e v .

Tretja zveznih porazdelitev, ki je izvedena iz osnovnih normalno porazdeljenih slučajnostnih spremenljivk, je F-porazdelitev.

Pri pogoju, da sta $X_1^2(m_1)$ in $X_2^2\ (m_2)$ dve neodvisni slučajnostni spremenljivki, ki se porazdeljujeta v X^2-porazdelitvah z ustreznima stopinjama prostosti, se

$$\frac{X_1^2/m_1}{X_2^2/m_2} =: F(m_1,m_2) \qquad (2.19)$$

porazdeljuje v F-porazdelitvi.

Gostota verjetnosti zanjo je

$$\varphi(F) = C_F\, F^{\frac{m_1-2}{2}} \left(1 + \frac{m_1}{m_2} F\right)^{-\frac{m_1+m_2}{2}} \qquad (2.20)$$

Iz 2.19 in 2.20 vidimo, da je F-porazdelitev odvisna od dveh stopinj prostosti m_1 in m_2.
F-porazdelitev je v splošnem asimetrična v desno, stopnja asimetrije pa se manjša, če se število stopinj prostosti veča.

Ker je F-porazdelitev odvisna od dveh stopinj prostosti, je tabela $F_P(m_1 m_2)$ kombinacijska tabela $m_1\ m_2$ za vsak P. Zato jo tabeliramo le za one vrednosti P, ki so v praksi najpomembnejše tj za P = 0.05 in P = 0.01. Podrobnejše vrednosti dobimo v ustrezni literaturi.

TABELA 2.7

Kritične vrednosti za F - porazdelitev

α	m_2	m_1 = 1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞	m_2	α
0·1	1	39·9	49·5	53·6	55·8	57·2	58·2	58·9	59·4	59·9	60·2	60·7	61·2	61·7	62·0	62·3	62·5	62·8	63·1	63·3	1	0·1
0·05		161	199	216	225	230	234	237	239	241	242	244	246	248	249	250	251	252	253	254		0·05
0·01		4052	4999	5403	5625	5764	5859	5928	5982	6022	6056	6106	6157	6209	6235	6261	6287	6313	6339	6366		0·01
0·1	2	8·53	9·00	9·16	9·24	9·29	9·33	9·35	9·37	9·38	9·39	9·41	9·42	9·44	9·45	9·46	9·47	9·47	9·48	9·49	2	0·1
0·05		18·5	19·0	19·2	19·2	19·3	19·3	19·4	19·4	19·4	19·4	19·4	19·4	19·4	19·5	19·5	19·5	19·5	19·5	19·5		0·05
0·01		98·5	99·0	99·2	99·2	99·3	99·3	99·4	99·4	99·4	99·4	99·4	99·4	99·4	99·5	99·5	99·5	99·5	99·5	99·5		0·01
0·1	3	5·54	5·46	5·39	5·34	5·31	5·28	5·27	5·25	5·24	5·23	5·22	5·20	5·18	5·18	5·17	5·16	5·15	5·14	5·13	3	0·1
0·05		10·1	9·55	9·28	9·12	9·01	8·94	8·89	8·85	8·81	8·79	8·74	8·70	8·66	8·64	8·62	8·59	8·57	8·55	8·53		0·05
0·01		34·1	30·8	29·5	28·7	28·2	27·9	27·7	27·5	27·3	27·2	27·1	26·9	26·7	26·6	26·5	26·4	26·3	26·2	26·1		0·01
0·1	4	4·54	4·32	4·19	4·11	4·05	4·01	3·98	3·95	3·94	3·92	3·90	3·87	3·84	3·83	3·82	3·80	3·79	3·78	3·76	4	0·1
0·05		7·71	6·94	6·59	6·39	6·26	6·16	6·09	6·04	6·00	5·96	5·91	5·86	5·80	5·77	5·75	5·72	5·69	5·66	5·63		0·05
0·01		21·2	18·0	16·7	16·0	15·5	15·2	15·0	14·8	14·7	14·5	14·4	14·2	14·0	13·9	13·8	13·7	13·7	13·6	13·5		0·01
0·1	5	4·06	3·78	3·62	3·52	3·45	3·40	3·37	3·34	3·32	3·30	3·27	3·24	3·21	3·19	3·17	3·16	3·14	3·12	3·10	5	0·1
0·05		6·61	5·79	5·41	5·19	5·05	4·95	4·88	4·82	4·77	4·74	4·68	4·62	4·56	4·53	4·50	4·46	4·43	4·40	4·36		0·05
0·01		16·3	13·3	12·1	11·4	11·0	10·7	10·5	10·3	10·2	10·1	9·89	9·72	9·55	9·47	9·38	9·29	9·20	9·11	9·02		0·01
0·1	6	3·78	3·46	3·29	3·18	3·11	3·05	3·01	2·98	2·96	2·94	2·90	2·87	2·84	2·82	2·80	2·78	2·76	2·74	2·72	6	0·1
0·05		5·99	5·14	4·76	4·53	4·39	4·28	4·21	4·15	4·10	4·06	4·00	3·94	3·87	3·84	3·81	3·77	3·74	3·70	3·67		0·05
0·01		13·7	10·9	9·78	9·15	8·75	8·47	8·26	8·10	7·98	7·87	7·72	7·56	7·40	7·31	7·23	7·14	7·06	6·97	6·88		0·01
0·1	7	3·59	3·26	3·07	2·96	2·88	2·83	2·78	2·75	2·72	2·70	2·67	2·63	2·59	2·58	2·56	2·54	2·51	2·49	2·47	7	0·1
0·05		5·59	4·74	4·35	4·12	3·97	3·87	3·79	3·73	3·68	3·64	3·57	3·51	3·44	3·41	3·38	3·34	3·30	3·27	3·23		0·05
0·01		12·2	9·55	8·45	7·85	7·46	7·19	6·99	6·84	6·72	6·62	6·47	6·31	6·16	6·07	5·99	5·91	5·82	5·74	5·65		0·01
0·1	8	3·46	3·11	2·92	2·81	2·73	2·67	2·62	2·59	2·56	2·54	2·50	2·46	2·42	2·40	2·38	2·36	2·34	2·32	2·29	8	0·1
0·05		5·32	4·46	4·07	3·84	3·69	3·58	3·50	3·44	3·39	3·35	3·28	3·22	3·15	3·12	3·08	3·04	3·01	2·97	2·93		0·05
0·01		11·3	8·65	7·59	7·01	6·63	6·37	6·18	6·03	5·91	5·81	5·67	5·52	5·36	5·28	5·20	5·12	5·03	4·95	4·86		0·01
0·1	9	3·36	3·01	2·81	2·69	2·61	2·55	2·51	2·47	2·44	2·42	2·38	2·34	2·30	2·28	2·25	2·23	2·21	2·18	2·16	9	0·1
0·05		5·12	4·26	3·86	3·63	3·48	3·37	3·29	3·23	3·18	3·14	3·07	3·01	2·94	2·90	2·86	2·83	2·79	2·75	2·71		0·05
0·01		10·6	8·02	6·99	6·42	6·06	5·80	5·61	5·47	5·35	5·26	5·11	4·96	4·81	4·73	4·65	4·57	4·48	4·40	4·31		0·01

F - porazdelitev - nadaljevanje

α	m_2 \ m_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞	m_2	α
0·1	10	3·28	2·92	2·73	2·61	2·52	2·46	2·41	2·38	2·35	2·32	2·28	2·24	2·20	2·18	2·16	2·13	2·11	2·08	2·06	10	0·1
0·05		4·96	4·10	3·71	3·48	3·33	3·22	3·14	3·07	3·02	2·98	2·91	2·84	2·77	2·74	2·70	2·66	2·62	2·58	2·54		0·05
0·01		10·0	7·56	6·55	5·99	5·64	5·39	5·20	5·06	4·94	4·85	4·71	4·56	4·41	4·33	4·25	4·17	4·08	4·00	3·91		0·01
0·1	11	3·23	2·86	2·66	2·54	2·45	2·39	2·34	2·30	2·27	2·25	2·21	2·17	2·12	2·10	2·08	2·05	2·03	2·00	1·97	11	0·1
0·05		4·84	3·98	3·59	3·36	3·20	3·09	3·01	2·95	2·90	2·85	2·79	2·72	2·65	2·61	2·57	2·53	2·49	2·45	2·40		0·05
0·01		9·65	7·21	6·22	5·67	5·32	5·07	4·89	4·74	4·63	4·54	4·40	4·25	4·10	4·02	3·94	3·86	3·78	3·69	3·60		0·01
0·1	12	3·18	2·81	2·61	2·48	2·39	2·33	2·28	2·24	2·21	2·19	2·15	2·10	2·06	2·04	2·01	1·99	1·96	1·93	1·90	12	0·1
0·05		4·75	3·89	3·49	3·26	3·11	3·00	2·91	2·85	2·80	2·75	2·69	2·62	2·54	2·51	2·47	2·43	2·38	2·34	2·30		0·05
0·01		9·33	6·93	5·95	5·41	5·06	4·82	4·64	4·50	4·39	4·30	4·16	4·01	3·86	3·78	3·70	3·62	3·54	3·45	3·36		0·01
0·1	13	3·14	2·76	2·56	2·43	2·35	2·28	2·23	2·20	2·16	2·14	2·10	2·05	2·01	1·98	1·96	1·93	1·90	1·88	1·85	13	0·1
0·05		4·67	3·81	3·41	3·18	3·03	2·92	2·83	2·77	2·71	2·67	2·60	2·53	2·46	2·42	2·38	2·34	2·30	2·25	2·21		0·05
0·01		9·07	6·70	5·74	5·21	4·86	4·62	4·44	4·30	4·19	4·10	3·96	3·82	3·66	3·59	3·51	3·43	3·34	3·25	3·17		0·01
0·1	14	3·10	2·73	2·52	2·39	2·31	2·24	2·19	2·15	2·12	2·10	2·05	2·01	1·96	1·94	1·91	1·89	1·86	1·83	1·80	14	0·1
0·05		4·60	3·74	3·34	3·11	2·96	2·85	2·76	2·70	2·65	2·60	2·53	2·46	2·39	2·35	2·31	2·27	2·22	2·18	2·13		0·05
0·01		8·86	6·51	5·56	5·04	4·69	4·46	4·28	4·14	4·03	3·94	3·80	3·66	3·51	3·43	3·35	3·27	3·18	3·09	3·00		0·01
0·1	15	3·07	2·70	2·49	2·36	2·27	2·21	2·16	2·12	2·09	2·06	2·02	1·97	1·92	1·90	1·87	1·85	1·82	1·79	1·76	15	0·1
0·05		4·54	3·68	3·29	3·06	2·90	2·79	2·71	2·64	2·59	2·54	2·48	2·40	2·33	2·29	2·25	2·20	2·16	2·11	2·07		0·05
0·01		8·68	6·36	5·42	4·89	4·56	4·32	4·14	4·00	3·89	3·80	3·67	3·52	3·37	3·29	3·21	3·13	3·05	2·96	2·87		0·01
0·1	16	3·05	2·67	2·46	2·33	2·24	2·18	2·13	2·09	2·06	2·03	1·99	1·94	1·89	1·87	1·84	1·81	1·78	1·75	1·72	16	0·1
0·05		4·49	3·63	3·24	3·01	2·85	2·74	2·66	2·59	2·54	2·49	2·42	2·35	2·28	2·24	2·19	2·15	2·11	2·06	2·01		0·05
0·01		8·53	6·23	5·29	4·77	4·44	4·20	4·03	3·89	3·78	3·69	3·55	3·41	3·26	3·18	3·10	3·02	2·93	2·84	2·75		0·01
0·1	17	3·03	2·64	2·44	2·31	2·22	2·15	2·10	2·06	2·03	2·00	1·96	1·91	1·86	1·84	1·81	1·78	1·75	1·72	1·69	17	0·1
0·05		4·45	3·59	3·20	2·96	2·81	2·70	2·61	2·55	2·49	2·45	2·38	2·31	2·23	2·19	2·15	2·10	2·06	2·01	1·96		0·05
0·01		8·40	6·11	5·18	4·67	4·34	4·10	3·93	3·79	3·68	3·59	3·46	3·31	3·16	3·08	3·00	2·92	2·83	2·75	2·65		0·01
0·1	18	3·01	2·62	2·42	2·29	2·20	2·13	2·08	2·04	2·00	1·98	1·93	1·89	1·84	1·81	1·78	1·75	1·72	1·69	1·66	18	0·1
0·05		4·41	3·55	3·16	2·93	2·77	2·66	2·58	2·51	2·46	2·41	2·34	2·27	2·19	2·15	2·11	2·06	2·02	1·97	1·92		0·05
0·01		8·29	6·01	5·09	4·58	4·25	4·01	3·84	3·71	3·60	3·51	3·37	3·23	3·08	3·00	2·92	2·84	2·75	2·66	2·57		0·01

F porazdelitev (na-daljevanje)

α	n_2	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞	n_2	α
0·1	19	2·99	2·61	2·40	2·27	2·18	2·11	2·06	2·02	1·98	1·96	1·91	1·86	1·81	1·79	1·76	1·73	1·70	1·67	1·63	19	0·1
0·05		4·38	3·52	3·13	2·90	2·74	2·63	2·54	2·48	2·42	2·38	2·31	2·23	2·16	2·11	2·07	2·03	1·98	1·93	1·88		0·05
0·01		8·18	5·93	5·01	4·50	4·17	3·94	3·77	3·63	3·52	3·43	3·30	3·15	3·00	2·92	2·84	2·76	2·67	2·58	2·49		0·01
0·1	20	2·97	2·59	2·38	2·25	2·16	2·09	2·04	2·00	1·96	1·94	1·89	1·84	1·79	1·77	1·74	1·71	1·68	1·64	1·61	20	0·1
0·05		4·35	3·49	3·10	2·87	2·71	2·60	2·51	2·45	2·39	2·35	2·28	2·20	2·12	2·08	2·04	1·99	1·95	1·90	1·84		0·05
0·01		8·10	5·85	4·94	4·43	4·10	3·87	3·70	3·56	3·46	3·37	3·23	3·09	2·94	2·86	2·78	2·69	2·61	2·52	2·42		0·01
0·1	21	2·96	2·57	2·36	2·23	2·14	2·08	2·02	1·98	1·95	1·92	1·87	1·83	1·78	1·75	1·72	1·69	1·66	1·62	1·59	21	0·1
0·05		4·32	3·47	3·07	2·84	2·68	2·57	2·49	2·42	2·37	2·32	2·25	2·18	2·10	2·05	2·01	1·96	1·92	1·87	1·81		0·05
0·01		8·02	5·78	4·87	4·37	4·04	3·81	3·64	3·51	3·40	3·31	3·17	3·03	2·88	2·80	2·72	2·64	2·55	2·46	2·36		0·01
0·1	22	2·95	2·56	2·35	2·22	2·13	2·06	2·01	1·97	1·93	1·90	1·86	1·81	1·76	1·73	1·70	1·67	1·64	1·60	1·57	22	0·1
0·05		4·30	3·44	3·05	2·82	2·66	2·55	2·46	2·40	2·34	2·30	2·23	2·15	2·07	2·03	1·98	1·94	1·89	1·84	1·78		0·05
0·01		7·95	5·72	4·82	4·31	3·99	3·76	3·59	3·45	3·35	3·26	3·12	2·98	2·83	2·75	2·67	2·58	2·50	2·40	2·31		0·01
0·1	23	2·94	2·55	2·34	2·21	2·11	2·05	1·99	1·95	1·92	1·89	1·85	1·80	1·74	1·72	1·69	1·66	1·62	1·59	1·55	23	0·1
0·05		4·28	3·42	3·03	2·80	2·64	2·53	2·44	2·37	2·32	2·27	2·20	2·13	2·05	2·00	1·96	1·91	1·86	1·81	1·76		0·05
0·01		7·88	5·66	4·76	4·26	3·94	3·71	3·54	3·41	3·30	3·21	3·07	2·93	2·78	2·70	2·62	2·54	2·45	2·35	2·26		0·01
0·1	24	2·93	2·54	2·33	2·19	2·10	2·04	1·98	1·94	1·91	1·88	1·83	1·78	1·73	1·70	1·67	1·64	1·61	1·57	1·53	24	0·1
0·05		4·26	3·40	3·01	2·78	2·62	2·51	2·42	2·36	2·30	2·25	2·18	2·11	2·03	1·98	1·94	1·89	1·84	1·79	1·73		0·05
0·01		7·82	5·61	4·72	4·22	3·90	3·67	3·50	3·36	3·26	3·17	3·03	2·89	2·74	2·66	2·58	2·49	2·40	2·31	2·21		0·01
0·1	25	2·92	2·53	2·32	2·18	2·09	2·02	1·97	1·93	1·89	1·87	1·82	1·77	1·72	1·69	1·66	1·63	1·59	1·56	1·52	25	0·1
0·05		4·24	3·39	2·99	2·76	2·60	2·49	2·40	2·34	2·28	2·24	2·16	2·09	2·01	1·96	1·92	1·87	1·82	1·77	1·71		0·05
0·01		7·77	5·57	4·68	4·18	3·86	3·63	3·46	3·32	3·22	3·13	2·99	2·85	2·70	2·62	2·54	2·45	2·36	2·27	2·17		0·01
0·1	26	2·91	2·52	2·31	2·17	2·08	2·01	1·96	1·92	1·88	1·86	1·81	1·76	1·71	1·68	1·65	1·61	1·58	1·54	1·50	26	0·1
0·05		4·23	3·37	2·98	2·74	2·59	2·47	2·39	2·32	2·27	2·22	2·15	2·07	1·99	1·95	1·90	1·85	1·80	1·75	1·69		0·05
0·01		7·72	5·53	4·64	4·14	3·82	3·59	3·42	3·29	3·18	3·09	2·96	2·82	2·66	2·58	2·50	2·42	2·33	2·23	2·13		0·01

F- porazdelitev -(nadaljevanje)

α	m_2	m_1: 1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞	m_2	α
0·1	27	2·90	2·51	2·30	2·17	2·07	2·00	1·95	1·91	1·87	1·85	1·80	1·75	1·70	1·67	1·64	1·60	1·57	1·53	1·49	27	0·1
0·05		4·21	3·35	2·96	2·73	2·57	2·46	2·37	2·31	2·25	2·20	2·13	2·06	1·97	1·93	1·88	1·84	1·79	1·73	1·67		0·05
0·01		7·68	5·49	4·60	4·11	3·78	3·56	3·39	3·26	3·15	3·06	2·93	2·78	2·63	2·55	2·47	2·38	2·29	2·20	2·10		0·01
0·1	28	2·89	2·50	2·29	2·16	2·06	2·00	1·94	1·90	1·87	1·84	1·79	1·74	1·69	1·66	1·63	1·59	1·56	1·52	1·48	28	0·1
0·05		4·20	3·34	2·95	2·71	2·56	2·45	2·36	2·29	2·24	2·19	2·12	2·04	1·96	1·91	1·87	1·82	1·77	1·71	1·65		0·05
0·01		7·64	5·45	4·57	4·07	3·75	3·53	3·36	3·23	3·12	3·03	2·90	2·75	2·60	2·52	2·44	2·35	2·26	2·17	2·06		0·01
0·1	29	2·89	2·50	2·28	2·15	2·06	1·99	1·93	1·89	1·86	1·83	1·78	1·73	1·68	1·65	1·62	1·58	1·55	1·51	1·47	29	0·1
0·05		4·18	3·33	2·93	2·70	2·55	2·43	2·35	2·28	2·22	2·18	2·10	2·03	1·94	1·90	1·85	1·81	1·75	1·70	1·64		0·05
0·01		7·60	5·42	4·54	4·04	3·73	3·50	3·33	3·20	3·09	3·00	2·87	2·73	2·57	2·49	2·41	2·33	2·23	2·14	2·03		0·01
0·1	30	2·88	2·49	2·28	2·14	2·05	1·98	1·93	1·88	1·85	1·82	1·77	1·72	1·67	1·64	1·61	1·57	1·54	1·50	1·46	30	0·1
0·05		4·17	3·32	2·92	2·69	2·53	2·42	2·33	2·27	2·21	2·16	2·09	2·01	1·93	1·89	1·84	1·79	1·74	1·68	1·62		0·05
0·01		7·56	5·39	4·51	4·02	3·70	3·47	3·30	3·17	3·07	2·98	2·84	2·70	2·55	2·47	2·39	2·30	2·21	2·11	2·01		0·01
0·1	40	2·84	2·44	2·23	2·09	2·00	1·93	1·87	1·83	1·79	1·76	1·71	1·66	1·61	1·57	1·54	1·51	1·47	1·42	1·38	40	0·1
0·05		4·08	3·23	2·84	2·61	2·45	2·34	2·25	2·18	2·12	2·08	2·00	1·92	1·84	1·79	1·74	1·69	1·64	1·58	1·51		0·05
0·01		7·31	5·18	4·31	3·83	3·51	3·29	3·12	2·99	2·89	2·80	2·66	2·52	2·37	2·29	2·20	2·11	2·02	1·92	1·80		0·01
0·1	60	2·79	2·39	2·18	2·04	1·95	1·87	1·82	1·77	1·74	1·71	1·66	1·60	1·54	1·51	1·48	1·44	1·40	1·35	1·29	60	0·1
0·05		4·00	3·15	2·76	2·53	2·37	2·25	2·17	2·10	2·04	1·99	1·92	1·84	1·75	1·70	1·65	1·59	1·53	1·47	1·39		0·05
0·01		7·08	4·98	4·13	3·65	3·34	3·12	2·95	2·82	2·72	2·63	2·50	2·35	2·20	2·12	2·03	1·94	1·84	1·73	1·60		0·01
0·1	120	2·75	2·35	2·13	1·99	1·90	1·82	1·77	1·72	1·68	1·65	1·60	1·54	1·48	1·45	1·41	1·37	1·32	1·26	1·19	120	0·1
0·05		3·92	3·07	2·68	2·45	2·29	2·18	2·09	2·02	1·96	1·91	1·83	1·75	1·66	1·61	1·55	1·50	1·43	1·35	1·25		0·05
0·01		6·85	4·79	3·95	3·48	3·17	2·96	2·79	2·66	2·56	2·47	2·34	2·19	2·03	1·95	1·86	1·76	1·66	1·53	1·38		0·01
0·1	∞	2·71	2·30	2·08	1·94	1·85	1·77	1·72	1·67	1·63	1·60	1·55	1·49	1·42	1·38	1·34	1·30	1·24	1·17	1·00	∞	0·1
0·05		3·84	3·00	2·60	2·37	2·21	2·10	2·01	1·94	1·88	1·83	1·75	1·67	1·57	1·52	1·46	1·39	1·32	1·22	1·00		0·05
0·01		6·63	4·61	3·78	3·32	3·02	2·80	2·64	2·51	2·41	2·32	2·18	2·04	1·88	1·79	1·70	1·59	1·47	1·32	1·00		0·01

2.10 Z v e z e m e d z, X^2, t in F p o r a z d e l i t - v a m i .

Med obravnavanimi štirimi porazdelitvami so naslednje zveze:

$$z^2 =: X^2(1) \qquad (2.21)$$

$$X^2(m_1) + X^2(m_2) =: X^2(m_1 + m_2) \qquad (2.22)$$

$$t^2_{(m)} = F\ (1, m) \qquad (2.23)$$

$$X^2(m)/m =: F\ (m, \infty) \qquad (2.24)$$

2.11 V z o r č n i i z r a z i z a p o p r e č j a i n v a r i a n c e .

Če iz normalno porazdeljene populacije

$$y =: N(M_y, \sigma^2_y) \qquad (2.25)$$

s parametroma aritmetično sredino M_y in varianco σ^2_y izberemo slučajnosten vzorec z n enotami, dobimo naslednje zakonitosti za aritmetične sredine iz vzorca

$$\bar{y} = \frac{1}{n}\,Sy \quad \text{in varianco iz vzorca} \quad s^2 = \frac{S(y-\bar{y})^2}{n-1}$$

$$\frac{\bar{y} - M_y}{\sigma_y}\sqrt{n} =: z \qquad (2.26)$$

$$\frac{\bar{y} - M}{s}\sqrt{n} =: t\ (m = n-1) \qquad (2.27)$$

$$\frac{(n-1)s^2}{\sigma^2} =: X^2(m = n-1) \qquad (2.28)$$

Če iz dveh normalnih porazdelitev, v katerih se

$$y_1 =: N(M_1, \sigma^2) \text{ in } y_2 =: N(M_2, \sigma^2) \qquad (6.29)$$

izberemo neodvisna vzorca z n_1 in n_2 enotami, velja:

$$\frac{(\bar{y}_2 - \bar{y}_1) - (M_2 - M_1)}{\sigma} \sqrt{\frac{n_1 n_2}{n_1 + n_2}} = : z \qquad (2.30)$$

$$\frac{(\bar{y}_2 - \bar{y}_1) - (M_2 - M_1)}{s} \sqrt{\frac{n_1 n_2}{n_1 + n_2}} = : t(m = n_1 + n_2 - 2) \qquad (2.31)$$

pri čemer je

$$s^2 = \frac{(n_1-1)s_1^2 + (n_2-1)\, s_2^2}{n_1 + n_2 - 2} \qquad (2.32)$$

$$\frac{s_1^2}{s_2^2} = : F\,(m_1 = n_1 - 1\,,\ m_2 = n_2-1) \qquad (2.33)$$

Če iz dveh normalno porazdeljenih populacij $y_1 = ; N(M, \sigma_1^2)$ in $y_2 = : N(M, \sigma_2^2)$ izberemo slučajnostna vzorca z n_1 in n_2 enotami, velja

$$\frac{s_1^2 / \sigma_1^2}{s_2^2 / \sigma_2^2} = : F(m_1 = n_1-1,\ m_2 = n_2-1) \qquad (2.34)$$

Kot vidimo je izraz 2.33 samo poseben primer za izraz 2.34, če je $\sigma_1^2 = \sigma_2^2 = \sigma^2$

Če iz k populacij, v katerih se y_k v vseh populacijah enako normalno porazdeljuje z $y_k = : N(M, \sigma^2)$ z enakimi aritmetičnimi sredinami M in enakimi variancami, izberemo neodvisne vzorce z n_1 , n_2 ... n_k enotami, velja

$$\frac{S(n_j-1)s_j^2}{\sigma^2} = : X^2(m_1 = n-k) \qquad (2.35)$$

$$\frac{\dfrac{\overset{k}{\underset{j=1}{S}}\, n_i\, \bar{y}_j^{\,2} - \dfrac{(Sn_j\, \bar{y}_j)^2}{n}}{k-1}}{\dfrac{S(n_j-1)s_j^2}{n-k}} = : F(m_1 = k-1,\ m_2 = n-k) \qquad (2.36)$$

2.12 I n t e r v a l n e o c e n e .

Zgornje zakonitosti izkoriščamo za ocenjevanje parametrov iz slučajnostnih vzorcev in za preskušanje domnev o parametrih.

Iz izrazov, ki smo jih nakazali v 2.11 dobimo naslednje verjetnostne neenačbe za dvostransko ocenjevanje parametrov M in σ s tveganjem $\alpha = 2P$

$$\bar{y} - z_P \frac{\sigma}{\sqrt{n}} < M_y < \bar{y} + z_P \frac{\sigma}{\sqrt{n}} \tag{2.37}$$

$$\bar{y} - t_P(n-1) \frac{s}{\sqrt{n}} < M < \bar{y} + t_P(n-1) \frac{s}{\sqrt{n}} \tag{2.38}$$

$$\frac{(n-1)s^2}{X^2_P\ (n-1)} < \sigma^2 < \frac{(n-1)s^2}{X^2_{1-P}\ (n-1)} \tag{2.39}$$

$$(\bar{y}_2 - \bar{y}_1) - z_P \sigma \sqrt{\frac{n_1 + n_2}{n_1 \cdot n_2}} < (M_2 - M_1) < (\bar{y}_2 - \bar{y}_1) + z_P \sigma \sqrt{\frac{n_1 + n_2}{n_1 \cdot n_2}} \tag{2.40}$$

$$(\bar{y}_2 - \bar{y}_1) - t_P(n_1+n_2-2)\, s \sqrt{\frac{n_1 + n_2}{n_1 \cdot n_2}} < (M_2 - M_1) < (\bar{y}_2 - \bar{y}_1) + t_P(n_1+n_2-2)\, s \sqrt{\frac{n_1 + n_2}{n_1 \cdot n_2}} \tag{2.41}$$

$$\frac{s_1^2}{s_2^2} \cdot \frac{1}{F_P(n_1-1;n_2-1)} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{s_1^2}{s_2^2} F_P(n_2-1;n_1-1) \tag{2.42}$$

$$\frac{S(n_j-1)s_j^2}{\chi_P^2(n-k)} < \sigma^2 < \frac{S(n_j-1)s_j^2}{\chi_{1-P}^2(n-k)} \qquad (2.43)$$

Iz zgornjih dvostranskih ocen dobimo enostranske ocene, če v neenačbah levo ali desno stran izpustimo, ustrezno oceni. Tveganje za enostranske ocene je $\alpha = P$.

Nakazane ocene dobimo z slučajnostnimi vzorci. Slučajnostni vzorec je iz populacije slučajnostno izbrana delna populacija z obsegom n. Glede na način izbora je vzorec slučajnostni dogodek, izrazi kot $\bar{y}$, s^2 ipd. pa slučajnostne spremenljivke, ki imajo svoje verjetnostne porazdelitve.

2.13 Preskušanje domnev.

Statistično preskušanje domnev se ukvarja z preskušanjem domnev o statističnih parametrih in preko njih o značilnostih populacij. Ker domneve po pravilu preskušamo s slučajnostnimi vzorci, domneve sprejemamo oziroma zavračamo z določenim tveganjem.

2.14 Ničelna domneva.

Z vzorci moremo v principu domneve le zavračati. Zato osnovnim domnevam H_1, katere želimo po pravilu sprejemati, priredimo ustrezno protidomnevo H_o, ki jo imenujemo ničelno domnevo. Če zavrnemo H_o, smo s tem sprejeli H_1.

Domnevi H_1, da je trdnost izdelkov po dveh postopkih različna, ustreza ničelna domneva H_o, da trdnost ni različna.

$$H_1 : M_1 \neq M_2 \qquad H_o : M_1 = M_2$$

Domnevi H_1, da je odstotek izmeta večji kot 3% ustreza ničelna domneva H_o, da je odstotek izmeta enak ali manjši kot 3%.

$$H_1 : P > 3\% \qquad H_o : P \leq 3\%$$

2.15 Mehanizem preskušanja domnev razložimo na praktičnem primeru. Vzemimo v preskus enostavni domnevi : osnovni domnevi, da je za normalno porazdeljeno populacijo z $\sigma = 5$, aritmetična sredina $M_1 = 20$ $H_1 : M = M_1 = 20$ postavimo ustrezno ničelno domnevo, da je za normalno porazdeljeno populacijo z $\sigma = 6$ aritmetična sredina $M_o = 15$ $H_o : M = M_o = 15$. Domnevi skušamo preskusiti s slučajnostnim vzorcem z $n = 9$ enotami.

Iz zakonitosti o vzorcih iz normalnih populacij vemo, da velja za aritmetično sredino iz vzorcev

$$\frac{\bar{y} - M}{\sigma}\sqrt{n} = : z \qquad (2.44)$$

Če velja ničelna domneva, da prava aritmetična sredina ustreza ničelni domnevi $M = M_o = 15$ velja zakonitost

$$\frac{\bar{y} - M_o}{\sigma}\sqrt{n} = : z \qquad (2.45)$$

V tem primeru je velika verjetnost ($1 - \alpha = 0.95$), da je iz podatkov iz vzorca ($\bar{y}$) in domneve (M_o) in podatkov o ($n . \sigma$) izračunani izraz $z < z_o = + 1.645$ manjši od z_o in majhna verjetnost ($\alpha = 0.05$), da bo izračunani $z \geqslant z_o = 1,645$ enak ali večji od z_o.

Ker je velika verjetnost, da je v primeru, da velja ničelna domneva $z < z_o$, v tem primeru sprejmemo ničelno domnevo oziroma jo obratno zavrnemo, če je izračunani $z \geqslant z_o$.

Vrednost $z_o = + 1.645$, ki loči razmik vrednosti za z, v katerem ničelno domnevo sprejmemo od razmika vrednosti v katerem ničelno domnevo zavrnemo, imenujemo kritično mejo, razmik, v katerem ničelno domnevo zavrnemo pa kritični razmik.

V našem primeru je kritična meja $z_o = + 1.645$ kritični razmik pa razmik $z \geqslant z_o = + 1,645$.

V primeru, da ne velja ničelna domneva, temveč osnovna domneva, da je $M = M_1 = 20$, se v standardizirani normalni porazdelitvi porazdeljuje izraz

$$\frac{\bar{y} - M_1}{\sigma}\sqrt{n} = \frac{\bar{y} - 20}{6}\sqrt{9} = : z$$

in ne izraz 2.45

V novi situaciji je

$$\frac{\bar{y} - M_o}{\sigma}\sqrt{n} = \frac{\bar{y} - M_1}{\sigma}\sqrt{n} + \frac{M_1 - M_o}{\sigma}\sqrt{n} \qquad (2.46)$$

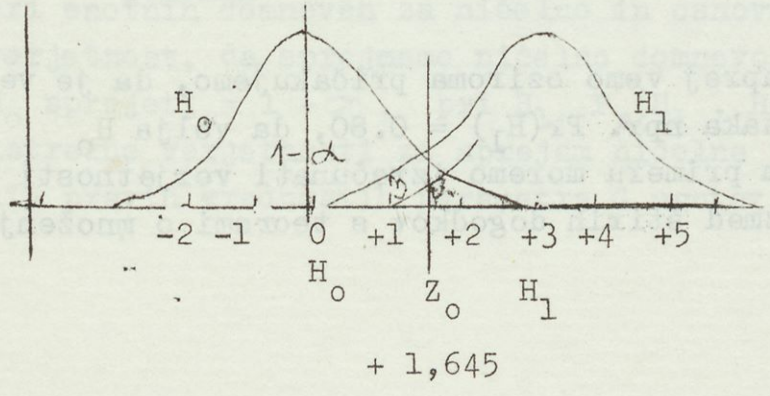
Če izračunamo novi porazdelitvi pod H_1 ustrezni z, za kritično vrednost $z_o = + 1.645$, dobimo

$$z_1 = \frac{\bar{y} - M_1}{\sigma}\sqrt{n} = \frac{\bar{y} - M_o}{\sigma}\sqrt{n} - \frac{M_1 - M_o}{\sigma}\sqrt{n} = z_o - \frac{M_1 - M_o}{\sigma}\, n =$$

$$= + 1.645 - \frac{20 - 14}{6}\sqrt{9} = - 1,455$$

Verjetnost, da bo v novi situaciji ($M = M_1 = 20$) vzorčni izraz $z = \frac{\bar{y} - M_o}{\sigma}\sqrt{n}$ pod kritično vrednostjo $z_o = + 1.645$, je torej $\beta = 1 - P(z_1 = -1.455) = 1 - 0.9272 = 0.0728$, da bo nad kritično vrednostjo $z_o = + 1,645$, pa $1 - \beta = P(z_1 = - 1.455) = 0.9272$.

2.16 Če narišemo za zgornje situacije verjetnostne porazdelitve dobimo nazornejšo sliko



$$z = \frac{\bar{y} - M_o}{\sigma}\sqrt{n}$$

Stanje \ Sklep	sprejmemo H_o zavrnemo H_1	Kritično vrednost	zavrnemo H_o sprejmemo H_1
H_o Velja	sklep pravilen verjetnost $1 - \alpha$		sklep nepravilen napaka α
H_1	sklep nepravilen napaka β		sklep pravilen verjetnost $1 - \beta$

Slika 2.8 Ponovitev odnosov pri preskušanju domnev

Iz slike in spodnje sheme vidimo, da ničelno domnevo sprejemamo, če ta velja z verjetnostjo $1 - \alpha$, a ničelno domnevo zavračamo kljub temu, da velja, z verjetnostjo α. To pomeni, da z verjetnostjo α napravimo tako imenovano napako prve vrste, ki je v tem, da ničelno domnevo zavrnemo, čeprav ta velja. Napako prve vrste predpišemo vnaprej z določitvijo kritične meje. V našem primeru smo jo določili na $\alpha = 0.05$.

Če velja osnovna domneva in je z pod + 1.645, sprejmemo H_o oziroma zavrnemo H_1. Sklep je nepravilen. Tovrstno napako, ki jo napravimo z verjetnostjo β imenujemo napako druge vrste. V našem primeru je verjetnost za napako druge vrste $\beta = 0.0728$. Če velja H_1 in je z nad + 1.645, H_1 sprejmemo. Sklep je torej pravilen. Verjetnost za tak sklep, če velja H_1, $1 - \beta$. Verjetnost, da sprejmemo H_1, če ta velja, imenujemo moč preskusa. Večja moč preskusa je v tem, da z večjo verjetnostjo sprejemamo osnovne domneve, če te veljajo.

2.17 Vzemimo, da vnaprej vemo oziroma pričakujemo, da je verjetnost, da velja H_1, enaka npr. $Pr(H_1) = 0.80$, da velja H_o $Pr(H_o) = 0.20$. V tem primeru moremo izračunati verjetnosti za nastop posameznih izmed štirih dogodkov s teoremi o množenju verjetnosti:

$Pr(H_oUH_o$ sprejet) $= P_r(H_o).P_r(H_o;H_o$ sprejet) $= 0.20 \; . \; 0.95 \; = 0.1900$

$P_r(H_oUH_o$ zavrnjen) $= P_r(H_o).P_r(H_o;H_o$zavrnjen) $= 0.20 \; . \; 0.05 \; = 0.0100$

$P_r(H_1UH_1$ sprejet) $= P_r(H_1).P_r(H_1;H_1$ sprejet) $= 0.80.0.9272 \; = 0.7418$

$P_r(H_1UH_1$ zavrnjen) $= P_r(H_1).P_r(H_1;H_1$zavrnjen) $= 0.80.0.0728 = 0.0582$

V tej situaciji pričakujemo, da bomo z verjetnostjo

$P_r(H_oUH_o$ sprejet) $+ P_r(H_1UH_1$sprejet) $= 0.1900 + 0.7418 = 0.9318$

sklepali pravilno in z verjetnostjo

$P_r(H_1UH_1$ zavrnjen) $+ P_r(H_oUH_o$ zavrnjen) $= 0.0582 + 0.0100 = 0.0682$

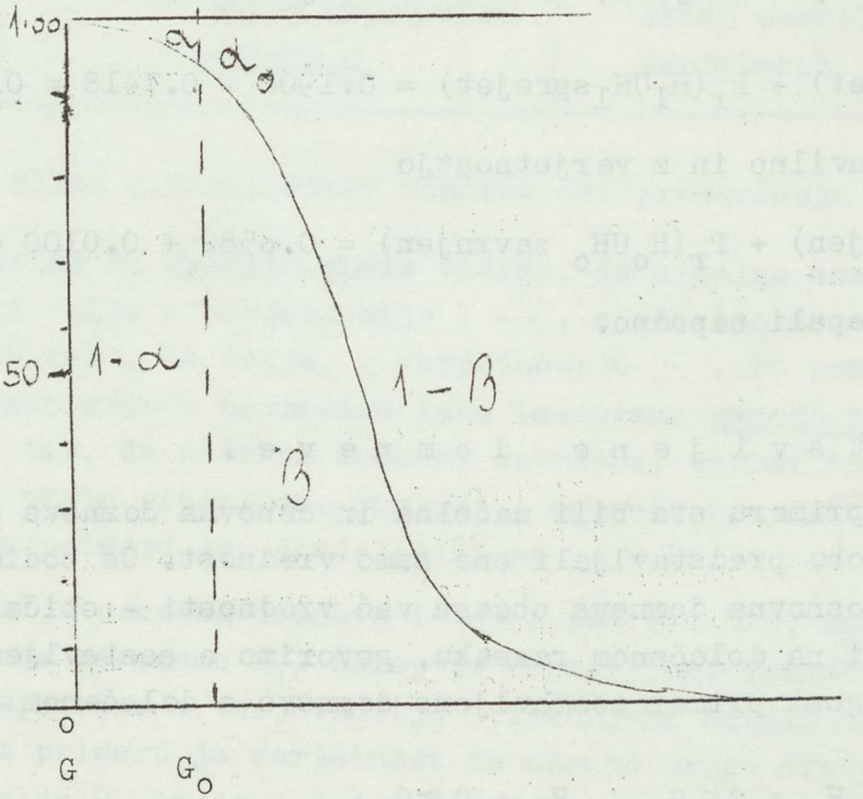
v celoti skepali napačno.

2.18 Sestavljene domneve.

V nakazanem primeru sta bili načelna in osnovna domneva enostavni, ker sta obe predstavljali eno samo vrednost. Če bodisi ničelna ali osnovna domneva obsega več vrednosti - običajno vse vrednosti na določenem razmaku, govorimo o sestavljenih domnevah. Pogost primer sestavljene domneve o določenem parametru G je

$$H_o : G \leq G_o \qquad H_1 : G > G_o \; .$$

Tako je npr. pri preverjanju kakovosti osnovna domneva, da je delež slabih izdelkov v pošiljki večji od P_o $(H_1 : P > P_o)$ ustrezna alternativna ničelna domneva pa je, da je delež defektnih izdelkov manjši ali enak P_o $(H_o : P \leq P_o)$. Kot moremo pri enotnih stavnih domnevah za ničelno in osnovno domnevo ugotoviti verjetnost, da sprejmemo ničelno domnevo (pri H_o $P_r(H_o$, H_o sprejet) $= 1 - \alpha$, pri H_1 $P_r(H_1$, H_o sprejet) $= \beta$, moremo ustrezne verjetnosti za sprejem ničelne vrednosti pri posameznih pravih vrednostih parametra G ugotoviti verjetnost za sprejem ničelne domneve. V grafikonu vse te vrednosti strnemo v operativno karakteristično krivuljo, ki kaže kako je verjetnost za sprejem ničelne domneve odvisna od vrednosti parametra. Za zgornji primer enostranske domneve je operativna karakteristična krivulja naslednja:



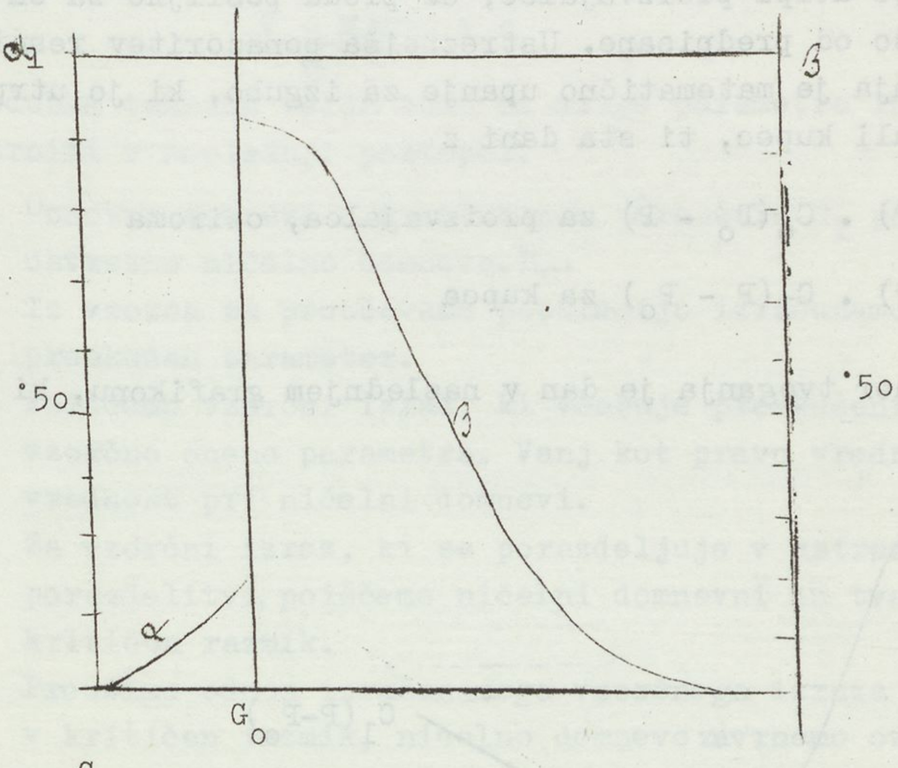
ničelna domneva	osnovna domneva

Slika 2.9 Operativna karakteristična krivulja za enostranske sestavljene preskuse

V razmaku $G < G_o$ ordinate za posamezne G kažejo verjetnost, da je ničelna domneva sprejeta $(1-\alpha)$. Ta se manjša, če se G bliža G_o in doseže za to vrednost najmanjšo vrednost $1-\alpha$.

Komplement do 1 predstavlja od zgoraj navzdol napako α, ki zavzame največjo vrednost (α_o) pri G_o. Verjetnost za sprejem ničelne domneve za prave vrednosti, ki so enake ali večje kot G_o je enaka napaki druge vrste β. Ta je največja $(1 - \alpha_o)$ pri $G = G_o$ in se manjša, čim večja je prava vrednost G, obratno pa se moč preskusa $1 - \beta$ z večanjem G veča.

Slika, ki prikaže odvisnost napake α in napake β v odvisnosti od prave vrednosti parametra, je podana v sliki 2.10.



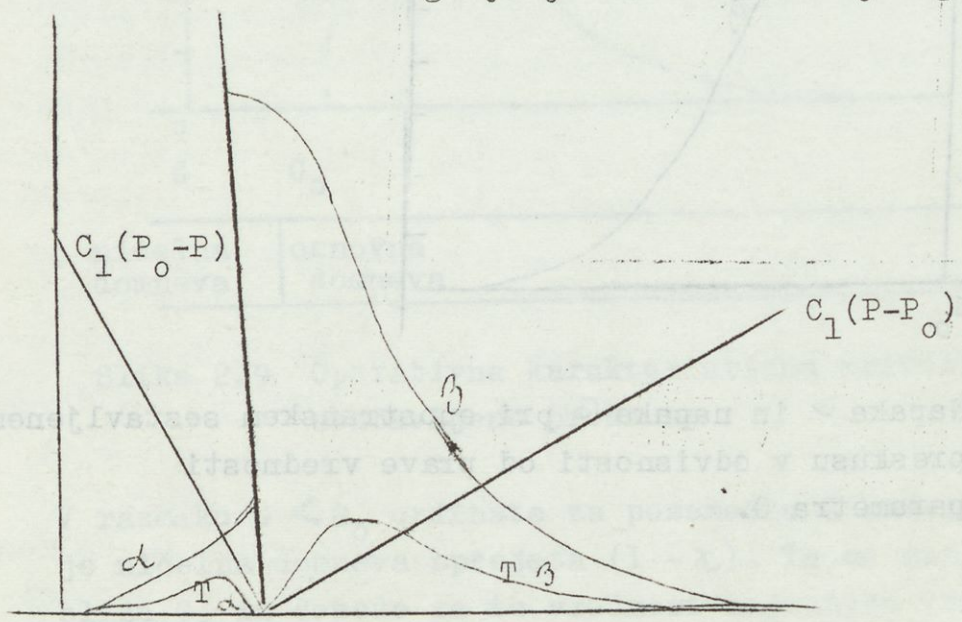
Slika 2.10 Napaka α in napake β pri enostranskem sestavljenem preskusu v odvisnosti od prave vrednosti parametra G.

2.19 Tveganje, podano z verjetnostjo nepravilnega sklepa, pa je enostransko prikazovanje stvarnega tveganja. Večji odkloni od mejne kakovosti, ki je še ustrezna, so vsebinsko pomembnejši kot pa manjši odkloni. Škoda, ki jo utrpi kupec, če kupi proizvodnjo neustrezne kakovosti, je proporcionalna razliki med stvarnim deležem P in mejnim deležem P_0, $C_1(P - P_0)$ pri čemer pomeni C_1 denarni ekvivalent izgube za en odstotek odklona kakovosti od predpisa. Obratno je izguba, ki jo utrpi proizvajalec, če boljšo kakovost P proda kot kakovost P_0 proporcionalna $C_0(P_0 - P)$ pri čemer pomeni C_0 denarni ekvivalent za škodo, ki jo utrpi proizvajalec, če proda pošiljko za en odstotek boljšo od predpisane. Ustreznejša ponazoritev resničnejšega tveganja je matematično upanje za izgubo, ki jo utrpi proizvajalec ali kupec, ti sta dani z

$$T_\alpha \quad \alpha(P) \cdot C_0(P_0 - P) \text{ za proizvajalca, oziroma}$$

$$T_\beta \quad \beta(P) \cdot C_1(P - P_0) \text{ za kupca}$$

Grafikon teh mer tveganja je dan v naslednjem grafikonu, ki



Slika 2.11 Pričakovane izgube

Iz slike vidimo, da je tako merjenje tveganja realnejše. Tveganje za P v neposredni okolici P_o je majhno, ker so majhni odkloni vsebinsko nepomembni, za velike razlike od P_o pa je tveganje za napačen sklep majhno, ker je verjetnost napačnega sklepa (bodisi α ali β) majhna.

Preskušanje domnev za parametre populacij.

Princip preskušanja domnev smo nakazali s preskusom za aritmetične sredine z izrazom

$$\frac{\bar{y} - M}{\sigma}\sqrt{n} = : z$$

Podobna tehnika velja tudi za druge parametre in jo moremo strniti v naslednji postopek.

1. Osnovni domnevi o preskušanem parametru H_1 postavimo ustrezno ničelno domnevo. H_o.
2. Iz vzorca za proučevano populacijo izračunamo oceno za preskušan parameter.
3. Poiščemo vzorčni izraz, ki vsebuje preskušani parameter in vzorčno oceno parametra. Vanj kot pravo vrednost postavimo vrednost pri ničelni domnevi.
4. Za vzorčni izraz, ki se porazdeljuje v ustrezni verjetnostni porazdelitvi, poiščemo ničelni domnevni in tveganju ustrezen kritičen razmik.
5. Proučimo odnos izračunanega vzorčnega izraza pod H_o. Če pade v kritičen razmik, ničelno domnevo zavrnemo oziroma sprejmemo osnovno domnevo
6. kar tolmačimo tako, da sklepamo, da je stanje od ničelne domneve značilno različno. V nasprotnem primeru je razlika neznačilna. Neznačilna razlika ne pomeni, da stvarno stanje ni različno od domnevnega, temveč le, da značilnost razlik z obstoječim preskusom, nismo odkrili.

V nadaljevanju podajamo vzorčne izraze 2.26 do 2.36 prirejene za preskušanje ustreznih domnev.

Za domneve o aritmetični sredini M_H

a) σ znan

$$\frac{\bar{y} - M_H}{\sigma}\sqrt{n} = z \quad \text{primerjamo z } z_P$$

b) σ neznan

$$\frac{\bar{y} - M_H}{s}\sqrt{n} = t \quad \text{primerjamo s } t_P(n-1)$$

c) razlike med aritmetičnima sredinama za dve populaciji z isto varianco $H_o : M_1 = M_2$

d) σ znan

$$\frac{\bar{y}_2 - \bar{y}_1}{\sigma}\sqrt{\frac{n_1\, n_2}{n_1 + n_2}} = z \quad \text{primerjamo z } z_P$$

e) σ neznan

$$\frac{\bar{y}_2 - \bar{y}_1}{s}\sqrt{\frac{n_1\, n_2}{n_1 + n_2}} = t \quad \text{primerjamo s } t_P(n_1+n_2-2)$$

f) razlike med aritmetičnimi sredinami za k populacij

g) domneve o varianci

$$\frac{Sn_j\bar{y}_j^2 - \frac{(Sn_j\bar{y}_j)^2}{n}}{k-1} \Bigg/ \frac{S(n_j-1)s_j^2}{n-k} = F \quad \text{primerjamo z}$$

$$F_P(m_1=k-1, m_2=n-k)$$

g) domneve o varianci

$$\frac{(n-1)s^2}{\sigma_H^2} = X^2 \quad \text{primerjamo z } X_P^2\ (n-1)$$

h) domneve o dveh variancah $H_o : \sigma_1^2 = \sigma_2^2$

$$\frac{s_1^2}{s_2^2} = F \quad \text{primerjamo z } F(m_1=n_1-1,\ m_2=n_2-1)$$

3. ANALIZA VARIANCE

3.1

Z analizo variance moremo preskušati domneve o razlikah med aritmetičnimi sredinami oziroma poprečji med več populacijami hkrati. Ker so aritmetične sredine izraz vpliva splošnih dejavnikov, oziroma opredeljujočih pogojev posameznih populacij, so razlike med aritmetičnimi sredinami izraz spremenjenih poskusnih pogojev ali tolmačeno vsebinsko izraz delovanja faktorjev, po katerih se posamezne populacije poskusov razlikujejo. Zaradi tega je analiza variance eden izmed osnovnih instrumentov pri obravnavanju poskusnih podatkov.

Predpostavke, pod katerimi izvajamo analizo variance so naslednje. Iz A populacij, za katere se y_{Ai} porazdeljuje normalno z enakimi aritmetičnimi sredinami M in enakimi variancami σ^2

$$y_{Ai} = : N(M, \sigma^2) \tag{3.1}$$

izberemo vzorce z enakimi obsegi n_A $\bar{n}$. Iz podatkov vzorcev moremo dobiti dve neodvisni oceni za varianco σ^2. Če iz posameznega za posamezen vzorec ocenimo varianco

$$s^2_{yAi} = \frac{S(y_{Ai} - \bar{y}_A)^2}{\bar{n} - 1} \tag{3.2}$$

dobimo sestavljeno oceno variance za σ^2 kot tehtano varianco iz posameznih populacij

$$\hat{s}^2_y = \frac{\sum(\bar{n}-1)s^2y_{Ai}}{\sum(\bar{n}-1)} = \frac{\sum(y_{Ai} - \bar{y}_A)^2}{\sum(\bar{n}-1)} = \frac{\sum_{Ai} Sy^2_{Ai} - \sum_{iA} \frac{y^2_A}{}}{n - A} =$$

$$= \frac{Q_{Ai} - Q_A}{n - A} \tag{3.3}$$

Zaradi predpostavk, katerim zadoščajo posamezne populacije in zaradi adicijskega teorema o slučajnostnih porazdelitvah X^2, se

$$\frac{(n - A)\hat{s}_y^{\,2}}{\sigma^2} = : X^2 \quad (m = n - A) \tag{3.4}$$

porazdeljuje v X^2 porazdelitvi z $m = n$-A stopinjami prostosti. Iz znanih zakonitosti o aritmetičnih sredinah vzorcev in zaradi zgornjih predpostavk, velja, da je ocena variance aritmetičnih sredin vzorcev

$$s^2_{\bar{y}_A} = \frac{\sum(\bar{y}_A - \bar{y})^2}{A - 1} \tag{3.5}$$

ocena za $\sigma^2/\bar{n}$. Zato je

$$s_A^2 = \bar{n}\, s^2_{\bar{y}_A} = \frac{\bar{n} \sum(\bar{y}_A - \bar{y})^2}{A - 1} = \frac{\sum \frac{y_A^2}{\bar{n}} - \frac{\bar{y}^2}{A\bar{n}}}{A - 1} =$$

$$= \frac{Q_A - Q}{A - 1} \tag{3.6}$$

ocena variance σ^2, ki sledi zakonitostim, da se

$$\frac{(A - 1)\, s_A^2}{\sigma^2} = : X^2 (m = A-1) \tag{3.7}$$

porazdeljuje v X^2 porazdelitvi z $m = A - 1$ stopinjami prostosti. Po definiciji F-porazdelitve se izraz

$$\frac{X_1^2 / m_1}{X_2^2 / m_2} = : F(m_1\, m_2) \tag{3.8}$$

porazdeljuje v F-porazdelitvi, če sta slučajnostni porazdelitvi X_1^2 in X_2^2 med seboj neodvisni. Moremo dokazati, da sta s_A^2 in s_{Ai}^2 neodvisni oceni za σ^2, zaradi tega velja

$$s_A^2 / s_{Ai}^2 = : F(m_1 = A - 1,\ m_2 = n - A) \tag{3.9}$$

3.2 Če pa so aritmetične sredine proučevanih populacij med seboj različne $M_1 \neq M_2 \neq M_3 \ldots \neq M_A$

moremo dokazati, da s_A^2 ni nepristanska ocena za σ^2, temveč je matematično upanje za s_A^2

$$E(s_A^2) = \sigma^2 + \bar{n}\, S_A^2 \qquad (3.10)$$

pri čemer pomeni razen znanih količin

$$S_A^2 = \frac{1}{A-1} \sum (A)^2 \quad \text{če je} \quad (A) = (M_A - M) \qquad (3.11)$$

$\dfrac{(A-1)s_A^2}{\sigma^2}$ se v tem primeru ne porazdeljuje v X^2 porazdelitvi temveč v necentralni X^2 porazdelitvi.

S_A^2 je torej poprečen kvadratičen odklon grupnih aritmetičnih sredin, (A) pa odklon grupne aritmetične sredine od skupnega poprečja M sledi, da je $(A) = 0$.

Ker je tudi v tem primeru, da so aritmetične sredine populacij med seboj različne

$$E(s_e^2) = \sigma^2$$

sledi, da razmerje

$$s_A^2/s_e^2$$

v primeru različnih aritmetičnih sredin ni porazdeljeno v F-porazdelitvi, ampak v porazdelitvi, ki v posebnem primeru, da je $S_A^2 = 0$ oziroma, da so aritmetične sredine med seboj enake, preide v F-porazdelitev. To porazdelitev imenujemo necentralno F-porazdelitev. Necentralna F-porazdelitev zavisi razen od stopinj prostosti m_1 in m_2 še od parametra necentralnosti

$$\delta = \sqrt{(A-1)\, \bar{n} S_A^2/\sigma_e^2} \qquad (3.12)$$

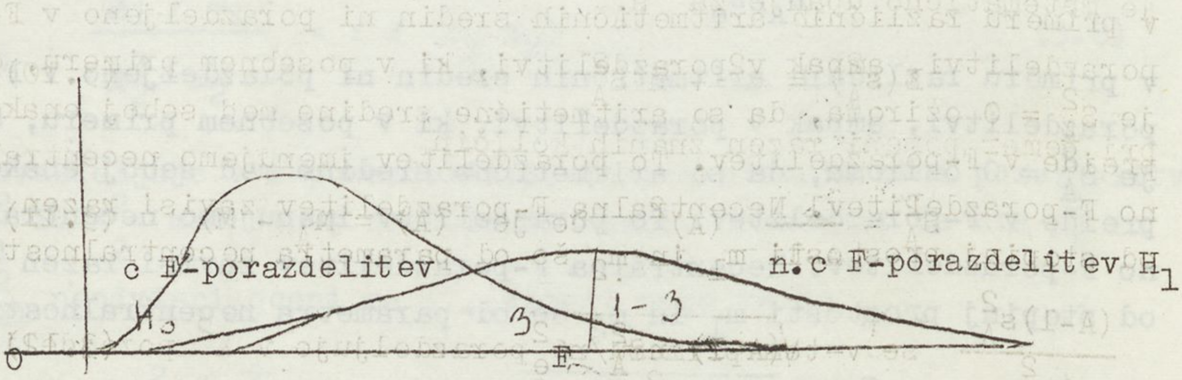
Razmerje matematičnih upanj

$$\frac{E(s_A^2)}{E(s_e^2)} = \frac{\sigma_e^2 + \bar{n}\, S_A^2}{\sigma_e^2} = 1 + \bar{n}S_A^2/\sigma_e^2 = 1 + \delta^2/(A-1) \quad (3.13)$$

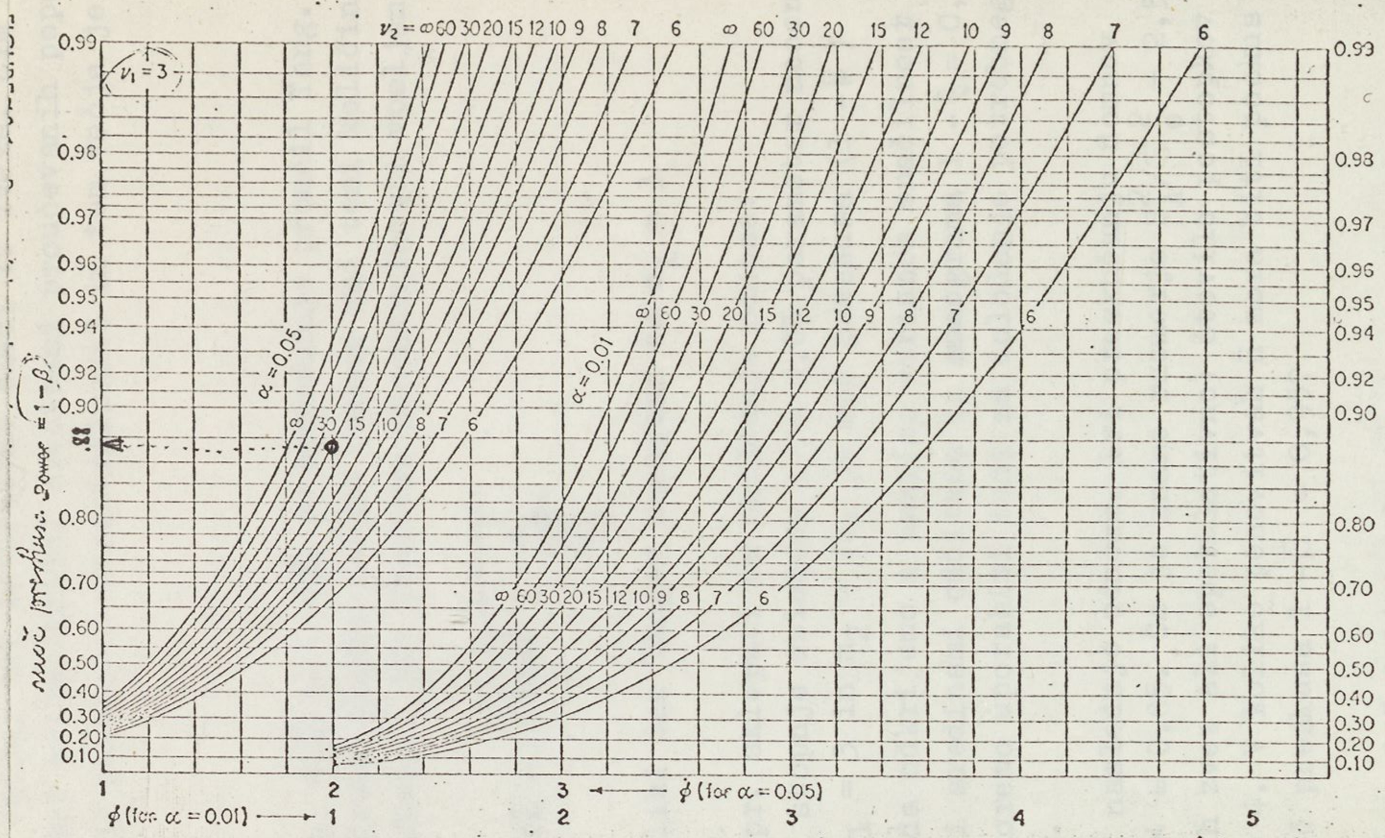
je odvisno od parametra necentralnosti in je 1, če je $\delta = 0$. Zgornje razmerje je tembolj različno od 1 čimvečja je necentralnost (čim večje so razlike med aritmetičnimi sredinami). Nakazane zakonitosti moremo koristiti za preskušanje domnev o aritmetičnih sredinah za več populacij.

V primeru, da postavimo ničelno domnevo, $H_o : M_1 = M_2 = \ldots M_A$, da so aritmetične sredine med seboj enake, kar je enako domnevi H_o : (A) = 0, da so komponente (A) = 0, nasproti osnovni domnevi, $H_1 : M_1 \neq M_2 \neq \ldots M_A$ ki je istovetna domnevi, da niso vse komponente (A) enake nič, moremo ti domnevi pisati tudi v obliki, $H_o : S_A^2 = 0$ $H_1 : S_A^2 \neq 0$ ali $H_o : \delta = 0$ $H_1 : \delta \neq 0$

Če velja domneva H_1 je razmerje s_A^2/s_e^2 sistematično večje, kot pa bi bilo v primeru, da velja H_o, pri kateri se razmerje $s_A^2/s_e^2 = : F(m_1,m_2)$. Situacijo v tem primeru ponazorimo grafično s sliko centralne (H_o) in necentralne (H_1) porazdelitve



Kritična vrednost F je določena iz centralne F-porazdelitve in ustreznih tablic : necentralna F-porazdelitev, ki se realizira v primeru H_1, je tembolj pomaknjena proti desni, čimvečji je parameter necentralnosti . V tem primeru je tudi moč



Slika 3.7 Tangov nomogram za $m_1 = 3$

preskusa $1-\beta$ tem večja, kar intuitivno sledi iz zveze. Čim večje so razlike med aritmetičnimi sredinami proučevanih populacij, tem večji je parameter necentralnosti in tem večja je moč preskusa.

3.3

Odnose med m_1, m_2, a, δ in $1-\beta$ je podrobneje proučil Tang. Z ustreznimi monogrami zlahka dobimo odnose med temi količinami. V grafikonu je namesto δ kot pa rametra necentralnosti vpeljan parameter

$$\phi = \delta/\sqrt{A} = \sqrt{\frac{(A-1)}{A} \text{ i } \frac{s_A^2}{\sigma_e^2}}$$

Kot primer je v sliki dan Tangov monogram za $m_1 = 3$

Iz njega moremo npr. sklepati na naslednje odnose:
Če je napaka prve stopnje določena z $\alpha = 0{,}05$ parameter necentralnosti je $\phi = 2$, $m_1 = 3$ in $m_2 = 20$, je moč preskusa $(1-\beta)$ to je verjetnost da odkrijemo z analizo variance značilnost razlik med štirimi sredinami odčitana iz monograma $1-\beta = 0{,}88$. Tangov monogram moremo uporabiti tudi za določanje potrebnega števila ponovitev.

Vzemimo za primer naslednje odnose. Pri preskušanju domnev postavimo, da je $\alpha = 0{,}05$. Če je znano razmerje $S_A^2/\sigma_e^2 = 2{,}5$ (določeno iz znanih zvez ali spekulativno) število postopkov v poskusu pa je $A = 4$. V koliko ponovitvah $\bar{n}$ mora biti poskus izveden, da je moč preskusa $1-\beta = 0{,}95$?

Iz razmerja $S_A^2/\sigma_e^2 = 2.5$ dobimo, da je

$$\phi = \sqrt{\bar{n} \cdot \frac{A-1}{A} \frac{S_A^2}{C_e^2}} = \sqrt{\bar{n} \frac{4-1}{4} 2.5} = \sqrt{1{,}875\, \bar{n}}$$

$$m_2 = A(\bar{n}-1) = 4(\bar{n}-1)$$

S poskusnim številom ponovitev ugotovimo iz Tangovega monograma

$\bar{n} = 3 \quad \phi = \sqrt{1,875.3} = \sqrt{5,625} = 2,37 \quad m_2 = 8; \; 1 - \beta = 0.89$

$\bar{n} = 4 \quad \phi = \sqrt{1,875.4} = \sqrt{7,5} = 2,74 \quad m_2 = 12; 1 - \beta = 0.98$

Iz rezultatov sklepamo, da se najbolj približamo postavljenim odnosom, če vzamemo število ponovitev $\bar{n} = 4$. V tem primeru je iz monograma odčitana vrednost za moč preskusa $1 - \beta = 0.98$ torej bliže zahtevane, medtem ko je za $\bar{n} = 3$ moč preskusa premajhna v primerjavi z zahtevano.

3.4 Homogenost varianc.

Osnovni pogoji, katerim morajo zadoščati poskusni podatki, da moremo na njih uporabiti analizo variance, je stavljena na poskusni pogrešek. Ta naj bi ustrezal pogoju $e_{Ai} = :N(0, \sigma_e^2)$

Poskusni pogrešek naj bi se porazdeljeval normalno, s konstantno varianco po posameznih grupah. Če rezultati poskusa tej zahtevi ne ustrezajo, analiza variance ni neposredno izvedljiva. V zvezi s tem se javljata dva metodološka problema. Prvi obstoji v tem, da preskusimo ali poskusni pogoji ustrezajo temu pogoju, drugi pa v tem, kako doseči, da bodo poskusni podatki ustrezali predpisanim pogojem.

Prvo nalogo rešujemo z ustreznimi preskusi o homogenosti varianc, od katerih bomo nakazali dva, ki se v praksi najpogosteje uporabljata: Bartlettov preskus in Cochranov preskus homogenosti varianc. Če odkrijemo da variance niso homogene in obstajajo med njimi značilne razlike, običajno skušamo z ustrezno transformacijo podatkov doseči homogenost varianc in normalnost poskusnega pogreška. Dostikrat pa je značaj razlik nakazan že s samim podatkom npr. odstotki, za katere je varianca

$\mathrm{Var}\ p = \frac{P(1-P)}{n}$ odvisna ne le od n, temveč tudi od velikosti parametra P.

3.5 B a r t l e t t o v p r e s k u s o h o m o g e n o s t i v a r i a n c .

Najpogosteje preskušamo homogenost varianc z Bartlettovim ali Cochranovim preskusom.

Bartlettov preskus o homogenosti varianc izvedemo na splošno na naslednji način. Z ničelno domnevo $H_o : \sigma_1^2 = \sigma_2^2 = \ldots \sigma_A^2$ preskušamo homogenost ocene varianc s_j^2 $j = 1 \ldots A$ s po m_j stopinjami prostosti z Bartlettovim preskusom z naslednjim izrazom

$$B = \frac{2,30259}{C}\left[m \log \overline{s^2} - \sum m_j \log s_j^2\right] = \frac{2,30259\ m}{C} \log \frac{Ms^2}{Gs^2} = \qquad (3.14)$$

Pri tem je

$$C = 1\ \frac{\sum\limits_j \frac{1}{m_j} - \frac{1}{m}}{3(A-1)} \qquad (3.15)$$

B se za primere, da je $m_j \geq 5$ približno porazdeljuje v $X^2(m = A-1)$

Vzemimo za primer ocene štirih grupnih varianc:

Grupa	s_j^2	m_j	$K = m_j s_j^2$	$\frac{1}{m_j}$	$\log s_j^2$	$m_j \log s_j^2$
1	84,2	8	673,6	·125	1,92531	15,40248
2	63,8	20	1276,0	·050	1,80482	36,09640
3	88,6	5	443,0	·200	1,94743	9,73715
4	72,1	10	721,0	·100	1,85794	18,57940
		43	3113,6	·475		79,81543
		m	$\sum m_j s_j^2$	$\sum \frac{1}{m_j}$		$\sum m_j \log s_j^2$

$$\overline{s^2} = \frac{\sum m_j s_j^2}{\sum m_j} = \frac{3113,6}{43} = 72,4093$$

$$C = 1 + \frac{\sum \frac{1}{m_j} - \frac{1}{m}}{3(A-1)} = 1 + \frac{\cdot 475 - \frac{1}{43}}{3(4-1)} = 1,05$$

$$\log \overline{s^2} = 1,85979$$

$$B = \frac{2,30259}{C} \left[m \log \overline{s^2} - \sum m_j \log s_j^2 \right] = \frac{2,30259}{1.05} \left[43.1,85979 - \right.$$

$$\left. - 79,81543 \right] = 3,410$$

Ker je $B = 3,410 < \chi^2_{0.10} (m = 4-1 = 3) = 6,25$

smatramo, da so razlike med ocenami varianc neznačilne.

$$C = \frac{\text{largest } \hat{\sigma}_j^2}{\Sigma \hat{\sigma}_j^2}$$

df for $\hat{\sigma}_j^2$	α	k = number of variances										
		2	3	4	5	6	7	8	9	10	15	20
1	.05	.9985	.9669	.9065	.8412	.7808	.7271	.6798	.6385	.6020	.4709	.3894
	.01	.9999	.9933	.9676	.9279	.8828	.8376	.7945	.7544	.7175	.5747	.4799
2	.05	.9750	.8709	.7679	.6838	.6161	.5612	.5157	.4775	.4450	.3346	.2705
	.01	.9950	.9423	.8643	.7885	.7218	.6644	.6152	.5727	.5358	.4069	.3297
3	.05	.9392	.7977	.6841	.5981	.5321	.4800	.4377	.4027	.3733	.2758	.2205
	.01	.9794	.8831	.7814	.6957	.6258	.5685	.5209	.4810	.4469	.3317	.2654
4	.05	.9057	.7457	.6287	.5441	.4803	.4307	.3910	.3584	.3311	.2419	.1921
	.01	.9586	.8335	.7212	.6329	.5635	.5080	.4627	.4251	.3934	.2882	.2288
5	.05	.8772	.7071	.5895	.5065	.4447	.3974	.3595	.3286	.3029	.2195	.1735
	.01	.9373	.7933	.6761	.5875	.5195	.4659	.4226	.3870	.3572	.2593	.2048
6	.05	.8534	.6771	.5598	.4783	.4184	.3726	.3362	.3067	.2823	.2034	.1602
	.01	.9172	.7606	.6410	.5531	.4866	.4347	.3932	.3592	.3308	.2386	.1877
7	.05	.8332	.6530	.5365	.4564	.3980	.3535	.3185	.2901	.2666	.1911	.1501
	.01	.8988	.7335	.6129	.5259	.4608	.4105	.3704	.3378	.3106	.2228	.1748
8	.05	.8159	.6333	.5175	.4387	.3817	.3384	.3043	.2768	.2541	.1815	.1422
	.01	.8823	.7107	.5897	.5037	.4401	.3911	.3522	.3207	.2945	.2104	.1646
9	.05	.8010	.6167	.5017	.4241	.3682	.3259	.2926	.2659	.2439	.1736	.1357
	.01	.8674	.6912	.5702	.4854	.4229	.3751	.3373	.3067	.2813	.2002	.1567
16	.05	.7341	.5466	.4366	.3645	.3135	.2756	.2462	.2226	.2032	.1429	.1108
	.01	.7949	.6059	.4884	.4094	.3529	.3105	.2779	.2514	.2297	.1612	.1248
36	.05	.6602	.4748	.3720	.3066	.2612	.2278	.2022	.1820	.1655	.1144	.0879
	.01	.7067	.5153	.4057	.3351	.2858	.2494	.2214	.1992	.1811	.1251	.0960
144	.05	.5813	.4031	.3093	.2513	.2119	.1833	.1616	.1446	.1308	.0889	.0675
	.01	.6062	.4230	.3251	.2644	.2229	.1929	.1700	.1521	.1376	.0934	.0709

Reprinted from chapter 15 of *Techniques of Statistical Analysis*, edited by C. Eisenhart, M.W. Hastay, and W. A. Wallis, McGraw-Hill Book Company, 1947.

TABELA 3.1. Kritične vrednosti za Cochranov preskus

3.6 C o c h r a n o v p r e s k u s o h o m o g e - n o s t i v a r i a n c .

Ta je zasnovan na preskusu razmerja med najvišjo oceno variance s^2_{max} in vsoto vseh ocenjenih varianc $\sum_j s_j^2$

$$C = \frac{s^2_{max}}{\sum_j s_j^2} \qquad (3.16)$$

Pri tem morajo biti vse variance izračunane z enakim številom stopenj prostosti $\bar{m}$. Razlike med variancami smatramo za značilne, če je izračunani C večji kot tabelirana vrednost $C_\alpha(A,\bar{m})$.

Vzemimo, da so variance s_j^2 v ponovitvah štirih postopkov z $\bar{m}$= 4 stopinjami prostosti

s_1^2 = 185,7; s_2^2 = 962,5; s_3^2 = 523,2; s_4^2 = 1218,2; $\sum s_j^2$ = 2889,7

Razmerje C je

$$C = \frac{s^2_{max}}{\sum s_j^2} = \frac{1218,2}{2889,7} = {\cdot}42$$

Iz tabele kritičnih vrednosti za C dobimo, da je $C_{o.o5}(4,4)$ = = ·6287 iz česar sklepamo, da razlike v variancah, čeprav so navidezno velike, niso značilne.

3.7 T r a n s f o r m a c i j e p o d a t k o v .

Nenormalnost in neenakost varianc v poskusnem gradivu pogosto odpravimo ali vsaj zmanjšamo z ustreznimi transformacijami.

3.8 Pogosto so poskusni podatki taki, da je po grupah konstanten koeficient variacije oziroma relativna varianca, ne pa standardni odklon ali absolutna varianca.

izraziti z

$$\sigma_e^2(A) = C^2 M_A^2 \qquad (3.17)$$

V takih primerih dosežemo z logaritemsko transformacijo podatkov

$$u = \log y \qquad (3.18)$$

da je varianca transformiranih podatkov σ_u^2 po grupah kontantna.

Da se izognemo težavam pri izračunu logaritmov, če je $y = 0$, namesto zgornje transformacije za take primere uporabljamo

$$u = \log (1 + y) \qquad (3.19)$$

3.9 Če je kriterialni podatki strukturni delež y%, vemo, da se p porazdeljuje v binomski porazdelitvi, varianca zanj pa zavisi razen od velikosti vzorca tudi od strukturnega deleža P.

$$E\ (y\%) = P\% \quad \mathrm{Var}\ y = \frac{P\ .\ Q}{n} = \frac{My\ (1 - My)}{n}$$

Za take primere s transformacijo

$$u = \arcsin \sqrt{y} \qquad (3.20)$$

dosežemo, da transformirani podatki zadoščajo normalnosti in enakosti varianc po grupah. Ustrezna tabela transformacije $u = \arcsin \sqrt{y}$ je dana v(dodatku) tabeli 3.3 .

ar = arcsin y %

%	0	1	2	3	4	5	6	7	8	9
0.0	0	0.57	0.81	0.99	1.15	1.28	1.40	1.52	1.62	1.72
0.1	1.81	1.90	1.99	2.07	2.14	2.22	2.29	2.36	2.43	2.50
0.2	2.56	2.63	2.69	2.75	2.81	2.87	2.92	2.98	3.03	3.09
0.3	3.14	3.19	3.24	3.29	3.34	3.39	3.44	3.49	3.53	3.58
0.4	3.63	3.67	3.72	3.76	3.80	3.85	3.89	3.93	3.97	4.01
0.5	4.05	4.09	4.13	4.17	4.21	4.25	4.29	4.33	4.37	4.40
0.6	4.44	4.48	4.52	4.55	4.59	4.62	4.66	4.69	4.73	4.76
0.7	4.80	4.83	4.87	4.90	4.93	4.97	5.00	5.03	5.07	5.10
0.8	5.13	5.16	5.20	5.23	5.26	5.29	5.32	5.35	5.38	5.41
0.9	5.44	5.47	5.50	5.53	5.56	5.59	5.62	5.65	5.68	5.71
1	5.74	6.02	6.29	6.55	6.80	7.04	7.27	7.49	7.71	7.92
2	8.13	8.33	8.53	8.72	8.91	9.10	9.28	9.46	9.63	9.81
3	9.98	10.14	10.31	10.47	10.63	10.78	10.94	11.09	11.24	11.39
4	11.54	11.68	11.83	11.97	12.11	12.25	12.39	12.52	12.66	12.79
5	12.92	13.05	13.18	13.31	13.44	13.56	13.69	13.81	13.94	14.06
6	14.18	14.30	14.42	14.54	14.65	14.77	14.89	15.00	15.12	15.23
7	15.34	15.45	15.56	15.68	15.79	15.89	16.00	16.11	16.22	16.32
8	16.43	16.54	16.64	16.74	16.85	16.95	17.05	17.16	17.26	17.36
9	17.46	17.56	17.66	17.76	17.85	17.95	18.05	18.15	18.24	18.34
10	18.44	18.53	18.63	18.72	18.81	18.91	19.00	19.09	19.19	19.28
11	19.37	19.46	19.55	19.64	19.73	19.82	19.91	20.00	20.09	20.18
12	20.27	20.36	20.44	20.53	20.62	20.70	20.79	20.88	20.96	21.05
13	21.13	21.22	21.30	21.39	21.47	21.56	21.64	21.72	21.81	21.89
14	21.97	22.06	22.14	22.22	22.30	22.38	22.46	22.55	22.63	22.71
15	22.79	22.87	22.95	23.03	23.11	23.19	23.26	23.34	23.42	23.50
16	23.58	23.66	23.73	23.81	23.89	23.97	24.04	24.12	24.20	24.27
17	24.35	24.43	24.50	24.58	24.65	24.73	24.80	24.88	24.95	25.03
18	25.10	25.18	25.25	25.33	25.40	25.48	25.55	25.62	25.70	25.77
19	25.84	25.92	25.99	26.06	26.13	26.21	26.28	26.35	26.42	26.49
20	26.56	26.64	26.71	26.78	26.85	26.92	26.99	27.06	27.13	27.20
21	27.28	27.35	27.42	27.49	27.56	27.63	27.69	27.76	27.83	27.90
22	27.97	28.04	28.11	28.18	28.25	28.32	28.38	28.45	28.52	28.59
23	28.66	28.73	28.79	28.86	28.93	29.00	29.06	29.13	29.20	29.27
24	29.33	29.40	29.47	29.53	29.60	29.67	29.73	29.80	29.87	29.93
25	30.00	30.07	30.13	30.20	30.26	30.33	30.40	30.46	30.53	30.59
26	30.66	30.72	30.79	30.85	30.92	30.98	31.05	31.11	31.18	31.24
27	31.31	31.37	31.44	31.50	31.56	31.63	31.69	31.76	31.82	31.88
28	31.95	32.01	32.08	32.14	32.20	32.27	32.33	32.39	32.46	32.52
29	32.58	32.65	32.71	32.77	32.83	32.90	32.96	33.02	33.09	33.15
30	33.21	33.27	33.34	33.40	33.46	33.52	33.58	33.65	33.71	33.77
31	33.83	33.89	33.96	34.02	34.08	34.14	34.20	34.27	34.33	34.39
32	34.45	34.51	34.57	34.63	34.70	34.76	34.82	34.88	34.94	35.00
33	35.06	35.12	35.18	35.24	35.30	35.37	35.43	35.49	35.55	35.61
34	35.67	35.73	35.79	35.85	35.91	35.97	36.03	36.09	36.15	36.21
35	36.27	36.33	36.39	36.45	36.51	36.57	36.63	36.69	36.75	36.81
36	36.87	36.93	36.99	37.05	37.11	37.17	37.23	37.29	37.35	37.41
37	37.47	37.52	37.58	37.64	37.70	37.76	37.82	37.88	37.94	38.00
38	38.06	38.12	38.17	38.23	38.29	38.35	38.41	38.47	38.53	38.59
39	38.65	38.70	38.76	38.82	38.88	38.94	39.00	39.06	39.11	39.17
40	39.23	39.29	39.35	39.41	39.47	39.52	39.58	39.64	39.70	39.76
41	39.82	39.87	39.93	39.99	40.05	40.11	40.16	40.22	40.28	40.34
42	40.40	40.46	40.51	40.57	40.63	40.69	40.74	40.80	40.86	40.92
43	40.98	41.03	41.09	41.15	41.21	41.27	41.32	41.38	41.44	41.50
44	41.55	41.61	41.67	41.73	41.78	41.84	41.90	41.96	42.02	42.07
45	42.13	42.19	42.25	42.30	42.36	42.42	42.48	42.53	42.59	42.65
46	42.71	42.76	42.82	42.88	42.94	42.99	43.05	43.11	43.17	43.22
47	43.28	43.34	43.39	43.45	43.51	43.57	43.62	43.68	43.74	43.80
48	43.85	43.91	43.97	44.03	44.08	44.14	44.20	44.25	44.31	44.37
49	44.43	44.48	44.54	44.60	44.66	44.71	44.77	44.83	44.89	44.94

Tabela 3,3 nadaljevanje

ar.= arcsin y%

%	0	1	2	3	4	5	6	7	8	9
50	45.00	45.06	45.11	45.17	45.23	45.29	45.34	45.40	45.46	45.52
51	45.57	45.63	45.69	45.75	45.80	45.86	45.92	45.97	46.03	46.09
52	46.15	46.20	46.26	46.32	46.38	46.43	46.49	46.55	46.61	46.66
53	46.72	46.78	46.83	46.89	46.95	47.01	47.06	47.12	47.18	47.24
54	47.29	47.35	47.41	47.47	47.52	47.58	47.64	47.70	47.75	47.81
55	47.87	47.93	47.98	48.04	48.10	48.16	48.22	48.27	48.33	48.39
56	48.45	48.50	48.56	48.62	48.68	48.73	48.79	48.85	48.91	48.97
57	49.02	49.08	49.14	49.20	49.26	49.31	49.37	49.43	49.49	49.54
58	49.60	49.66	49.72	49.78	49.84	49.89	49.95	50.01	50.07	50.13
59	50.18	50.24	50.30	50.36	50.42	50.48	50.53	50.59	50.65	50.71
60	50.77	50.83	50.89	50.94	51.00	51.06	51.12	51.18	51.24	51.30
61	51.35	51.41	51.47	51.53	51.59	51.65	51.71	51.77	51.83	51.88
62	51.94	52.00	52.06	52.12	52.18	52.24	52.30	52.36	52.42	52.48
63	52.53	52.59	52.65	52.71	52.77	52.83	52.89	52.95	53.01	53.07
64	53.13	53.19	53.25	53.31	53.37	53.43	53.49	53.55	53.61	53.67
65	53.73	53.79	53.85	53.91	53.97	54.03	54.09	54.15	54.21	54.27
66	54.33	54.39	54.45	54.51	54.57	54.63	54.70	54.76	54.82	54.88
67	54.94	55.00	55.06	55.12	55.18	55.24	55.30	55.37	55.43	55.49
68	55.55	55.61	55.67	55.73	55.80	55.86	55.92	55.98	56.04	56.11
69	56.17	56.23	56.29	56.35	56.42	56.48	56.54	56.60	56.66	56.73
70	56.79	56.85	56.91	56.98	57.04	57.10	57.17	57.23	57.29	57.35
71	57.42	57.48	57.54	57.61	57.67	57.73	57.80	57.86	57.92	57.99
72	58.05	58.12	58.18	58.24	58.31	58.37	58.44	58.50	58.56	58.63
73	58.69	58.76	58.82	58.89	58.95	59.02	59.08	59.15	59.21	59.28
74	59.34	59.41	59.47	59.54	59.60	59.67	59.74	59.80	59.87	59.93
75	60.00	60.07	60.13	60.20	60.27	60.33	60.40	60.47	60.53	60.60
76	60.67	60.73	60.80	60.87	60.94	61.00	61.07	61.14	61.21	61.27
77	61.34	61.41	61.48	61.55	61.62	61.68	61.75	61.82	61.89	61.96
78	62.03	62.10	62.17	62.24	62.31	62.37	62.44	62.51	62.58	62.65
79	62.72	62.80	62.87	62.94	63.01	63.08	63.15	63.22	63.29	63.36
80	63.44	63.51	63.58	63.65	63.72	63.79	63.87	63.94	64.01	64.08
81	64.16	64.23	64.30	64.38	64.45	64.52	64.60	64.67	64.75	64.82
82	64.90	64.97	65.05	65.12	65.20	65.27	65.35	65.42	65.50	65.57
83	65.65	65.73	65.80	65.88	65.96	66.03	66.11	66.19	66.27	66.34
84	66.42	66.50	66.58	66.66	66.74	66.81	66.89	66.97	67.05	67.13
85	67.21	67.29	67.37	67.45	67.54	67.62	67.70	67.78	67.86	67.94
86	68.03	68.11	68.19	68.28	68.36	68.44	68.53	68.61	68.70	68.78
87	68.87	68.95	69.04	69.12	69.21	69.30	69.38	69.47	69.56	69.64
88	69.73	69.82	69.91	70.00	70.09	70.18	70.27	70.36	70.45	70.54
89	70.63	70.72	70.81	70.91	71.00	71.09	71.19	71.28	71.37	71.47
90	71.56	71.66	71.76	71.85	71.95	72.05	72.15	72.24	72.34	72.44
91	72.54	72.64	72.74	72.84	72.95	73.05	73.15	73.26	73.36	73.46
92	73.57	73.68	73.78	73.89	74.00	74.11	74.21	74.32	74.44	74.55
93	74.66	74.77	74.88	75.00	75.11	75.23	75.35	75.46	75.58	75.70
94	75.82	75.94	76.06	76.19	76.31	76.44	76.56	76.69	76.82	76.95
95	77.08	77.21	77.34	77.48	77.61	77.75	77.89	78.03	78.17	78.32
96	78.46	78.61	78.76	78.91	79.06	79.22	79.37	79.53	79.69	79.86
97	80.02	80.19	80.37	80.54	80.72	80.90	81.09	81.28	81.47	81.67
98	81.87	82.08	82.29	82.51	82.73	82.96	83.20	83.45	83.71	83.98
99.0	84.26	84.29	84.32	84.35	84.38	84.41	84.44	84.47	84.50	84.53
99.1	84.56	84.59	84.62	84.65	84.68	84.71	84.74	84.77	84.80	84.84
99.2	84.87	84.90	84.93	84.97	85.00	85.03	85.07	85.10	85.13	85.17
99.3	85.20	85.24	85.27	85.31	85.34	85.38	85.41	85.45	85.48	85.52
99.4	85.56	85.60	85.63	85.67	85.71	85.75	85.79	85.83	85.87	85.91
99.5	85.95	85.99	86.03	86.07	86.11	86.15	86.20	86.24	86.28	86.33
99.6	86.37	86.42	86.47	86.51	86.56	86.61	86.66	86.71	86.76	86.81
99.7	86.86	86.91	86.97	87.02	87.08	87.13	87.19	87.25	87.31	87.37
99.8	87.44	87.50	87.57	87.64	87.71	87.78	87.86	87.93	88.01	88.10
99.9	88.19	88.28	88.38	88.48	88.60	88.72	88.85	89.01	89.19	89.43

Tabela 3.2 ŠTIRIMESTNI LOGARITMI

štirimestni logaritmi

N	0	1	2	3	4	5	6	7	8	9	PP 1	2	3	4	5	6	7	8	9
10	0000	0043	0086	0128	0170	0212	0253	0294	0334	0374	*4	8	12	17	21	25	29	33	37
11	0414	0453	0492	0531	0569	0607	0645	0682	0719	0755	4	8	11	15	19	23	26	30	34
12	0792	0828	0864	0899	0934	0969	1004	1038	1072	1106	3	7	10	14	17	21	24	28	31
13	1139	1173	1206	1239	1271	1303	1335	1367	1399	1430	3	6	10	13	16	19	23	26	29
14	1461	1492	1523	1553	1584	1614	1644	1673	1703	1732	3	6	9	12	15	18	21	24	27
15	1761	1790	1818	1847	1875	1903	1931	1959	1987	2014	*3	6	8	11	14	17	20	22	25
16	2041	2068	2095	2122	2148	2175	2201	2227	2253	2279	3	5	8	11	13	16	18	21	24
17	2304	2330	2355	2380	2405	2430	2455	2480	2504	2529	2	5	7	10	12	15	17	20	22
18	2553	2577	2601	2625	2648	2672	2695	2718	2742	2765	2	5	7	9	12	14	16	19	21
19	2788	2810	2833	2856	2878	2900	2923	2945	2967	2989	2	4	7	9	11	13	16	18	20
20	3010	3032	3054	3075	3096	3118	3139	3160	3181	3201	2	4	6	8	11	13	15	17	19
21	3222	3243	3263	3284	3304	3324	3345	3365	3385	3404	2	4	6	8	10	12	14	16	18
22	3424	3444	3464	3483	3502	3522	3541	3560	3579	3598	2	4	6	8	10	12	14	15	17
23	3617	3636	3655	3674	3692	3711	3729	3747	3766	3784	2	4	6	7	9	11	13	15	17
24	3802	3820	3838	3856	3874	3892	3909	3927	3945	3962	2	4	5	7	9	11	12	14	16
25	3979	3997	4014	4031	4048	4065	4082	4099	4116	4133	2	3	5	7	9	10	12	14	15
26	4150	4166	4183	4200	4216	4232	4249	4265	4281	4298	2	3	5	7	8	10	11	13	15
27	4314	4330	4346	4362	4378	4393	4409	4425	4440	4456	2	3	5	6	8	9	11	13	14
28	4472	4487	4502	4518	4533	4548	4564	4579	4594	4609	2	3	5	6	8	9	11	12	14
29	4624	4639	4654	4669	4683	4698	4713	4728	4742	4757	1	3	4	6	7	9	10	12	13
30	4771	4786	4800	4814	4829	4843	4857	4871	4886	4900	1	3	4	6	7	9	10	11	13
31	4914	4928	4942	4955	4969	4983	4997	5011	5024	5038	1	3	4	6	7	8	10	11	12
32	5051	5065	5079	5092	5105	5119	5132	5145	5159	5172	1	3	4	5	7	8	9	11	12
33	5185	5198	5211	5224	5237	5250	5263	5276	5289	5302	1	3	4	5	6	8	9	10	12
34	5315	5328	5340	5353	5366	5378	5391	5403	5416	5428	1	3	4	5	6	8	9	10	11
35	5441	5453	5465	5478	5490	5502	5514	5527	5539	5551	1	2	4	5	6	7	9	10	11
36	5563	5575	5587	5599	5611	5623	5635	5647	5658	5670	1	2	4	5	6	7	8	10	11
37	5682	5694	5705	5717	5729	5740	5752	5763	5775	5786	1	2	3	5	6	7	8	9	10
38	5798	5809	5821	5832	5843	5855	5866	5877	5888	5899	1	2	3	5	6	7	8	9	10
39	5911	5922	5933	5944	5955	5966	5977	5988	5999	6010	1	2	3	4	5	7	8	9	10
40	6021	6031	6042	6053	6064	6075	6085	6096	6107	6117	1	2	3	4	5	6	8	9	10
41	6128	6138	6149	6160	6170	6180	6191	6201	6212	6222	1	2	3	4	5	6	7	8	9
42	6232	6243	6253	6263	6274	6284	6294	6304	6314	6325	1	2	3	4	5	6	7	8	9
43	6335	6345	6355	6365	6375	6385	6395	6405	6415	6425	1	2	3	4	5	6	7	8	9
44	6435	6444	6454	6464	6474	6484	6493	6503	6513	6522	1	2	3	4	5	6	7	8	9
45	6532	6542	6551	6561	6571	6580	6590	6599	6609	6618	1	2	3	4	5	6	7	8	9
46	6628	6637	6646	6656	6665	6675	6684	6693	6702	6712	1	2	3	4	5	6	7	7	8
47	6721	6730	6739	6749	6758	6767	6776	6785	6794	6803	1	2	3	4	5	5	6	7	8
48	6812	6821	6830	6839	6848	6857	6866	6875	6884	6893	1	2	3	4	4	5	6	7	8
49	6902	6911	6920	6928	6937	6946	6955	6964	6972	6981	1	2	3	4	4	5	6	7	8
50	6990	6998	7007	7016	7024	7033	7042	7050	7059	7067	1	2	3	3	4	5	6	7	8
51	7076	7084	7093	7101	7110	7118	7126	7135	7143	7152	1	2	3	3	4	5	6	7	8
52	7160	7168	7177	7185	7193	7202	7210	7218	7226	7235	1	2	2	3	4	5	6	7	7
53	7243	7251	7259	7267	7275	7284	7292	7300	7308	7316	1	2	2	3	4	5	6	6	7
54	7324	7332	7340	7348	7356	7364	7372	7380	7388	7396	1	2	2	3	4	5	6	6	7
N	0	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9

interpolacija v tem odseku je nenatančna

N	0	1	2	3	4	5	6	7	8	9	PP 1	2	3	4	5	6	7	8	9
55	7404	7412	7419	7427	7435	7443	7451	7459	7466	7474	1	2	2	3	4	5	5	6	7
56	7482	7490	7497	7505	7513	7520	7528	7536	7543	7551	1	2	2	3	4	5	5	6	7
57	7559	7566	7574	7582	7589	7597	7604	7612	7619	7627	1	2	2	3	4	5	5	6	7
58	7634	7642	7649	7657	7664	7672	7679	7686	7694	7701	1	1	2	3	4	4	5	6	7
59	7709	7716	7723	7731	7738	7745	7752	7760	7767	7774	1	1	2	3	4	4	5	6	7
60	7782	7789	7796	7803	7810	7818	7825	7832	7839	7846	1	1	2	3	4	4	5	6	6
61	7853	7860	7868	7875	7882	7889	7896	7903	7910	7917	1	1	2	3	4	4	5	6	6
62	7924	7931	7938	7945	7952	7959	7966	7973	7980	7987	1	1	2	3	3	4	5	6	6
63	7993	8000	8007	8014	8021	8028	8035	8041	8048	8055	1	1	2	3	3	4	5	5	6
64	8062	8069	8075	8082	8089	8096	8102	8109	8116	8122	1	1	2	3	3	4	5	5	6
65	8129	8136	8142	8149	8156	8162	8169	8176	8182	8189	1	1	2	3	3	4	5	5	6
66	8195	8202	8209	8215	8222	8228	8235	8241	8248	8254	1	1	2	3	3	4	5	5	6
67	8261	8267	8274	8280	8287	8293	8299	8306	8312	8319	1	1	2	3	3	4	5	5	6
68	8325	8331	8338	8344	8351	8357	8363	8370	8376	8382	1	1	2	3	3	4	4	5	6
69	8388	8395	8401	8407	8414	8420	8426	8432	8439	8445	1	1	2	2	3	4	4	5	6
70	8451	8457	8463	8470	8476	8482	8488	8494	8500	8506	1	1	2	2	3	4	4	5	6
71	8513	8519	8525	8531	8537	8543	8549	8555	8561	8567	1	1	2	2	3	4	4	5	5
72	8573	8579	8585	8591	8597	8603	8609	8615	8621	8627	1	1	2	2	3	4	4	5	5
73	8633	8639	8645	8651	8657	8663	8669	8675	8681	8686	1	1	2	2	3	4	4	5	5
74	8692	8698	8704	8710	8716	8722	8727	8733	8739	8745	1	1	2	2	3	4	4	5	5
75	8751	8756	8762	8768	8774	8779	8785	8791	8797	8802	1	1	2	2	3	3	4	5	5
76	8808	8814	8820	8825	8831	8837	8842	8848	8854	8859	1	1	2	2	3	3	4	5	5
77	8865	8871	8876	8882	8887	8893	8899	8904	8910	8915	1	1	2	2	3	3	4	4	5
78	8921	8927	8932	8938	8943	8949	8954	8960	8965	8971	1	1	2	2	3	3	4	4	5
79	8976	8982	8987	8993	8998	9004	9009	9015	9020	9025	1	1	2	2	3	3	4	4	5
80	9031	9036	9042	9047	9053	9058	9063	9069	9074	9079	1	1	2	2	3	3	4	4	5
81	9085	9090	9096	9101	9106	9112	9117	9122	9128	9133	1	1	2	2	3	3	4	4	5
82	9138	9143	9149	9154	9159	9165	9170	9175	9180	9186	1	1	2	2	3	3	4	4	5
83	9191	9196	9201	9206	9212	9217	9222	9227	9232	9238	1	1	2	2	3	3	4	4	5
84	9243	9248	9253	9258	9263	9269	9274	9279	9284	9289	1	1	2	2	3	3	4	4	5
85	9294	9299	9304	9309	9315	9320	9325	9330	9335	9340	1	1	2	2	3	3	4	4	5
86	9345	9350	9355	9360	9365	9370	9375	9380	9385	9390	1	1	2	2	3	3	4	4	5
87	9395	9400	9405	9410	9415	9420	9425	9430	9435	9440	0	1	1	2	2	3	3	4	4
88	9445	9450	9455	9460	9465	9469	9474	9479	9484	9489	0	1	1	2	2	3	3	4	4
89	9494	9499	9504	9509	9513	9518	9523	9528	9533	9538	0	1	1	2	2	3	3	4	4
90	9542	9547	9552	9557	9562	9566	9571	9576	9581	9586	0	1	1	2	2	3	3	4	4
91	9590	9595	9600	9605	9609	9614	9619	9624	9628	9633	0	1	1	2	2	3	3	4	4
92	9638	9643	9647	9652	9657	9661	9666	9671	9675	9680	0	1	1	2	2	3	3	4	4
93	9685	9689	9694	9699	9703	9708	9713	9717	9722	9727	0	1	1	2	2	3	3	4	4
94	9731	9736	9741	9745	9750	9754	9759	9763	9768	9773	0	1	1	2	2	3	3	4	4
95	9777	9782	9786	9791	9795	9800	9805	9809	9814	9818	0	1	1	2	2	3	3	4	4
96	9823	9827	9832	9836	9841	9845	9850	9854	9859	9863	0	1	1	2	2	3	3	4	4
97	9868	9872	9877	9881	9886	9890	9894	9899	9903	9908	0	1	1	2	2	3	3	4	4
98	9912	9917	9921	9926	9930	9934	9939	9943	9948	9952	0	1	1	2	2	3	3	4	4
99	9956	9961	9965	9969	9974	9978	9983	9987	9991	9996	0	1	1	2	2	3	3	3	4
N	0	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9

Tabela 3.4 Rankiti Z_R(od A = 10 dalje so negativne vrednosti izpuščene, ker so simetrične)

n R \ A	2	3	4	5	6	7	8	9	10
1	0.564	0.864	1.029	1.163	1.267	1.352	1.424	1.485	1.539
2	−0.564	0.000	0.297	0.495	0.642	0.757	0.852	0.932	1.001
3		−0.864	−0.297	0.000	0.202	0.353	0.473	0.572	0.656
4			−1.029	−0.495	−0.202	0.000	0.153	0.275	0.376
5				−1.163	−0.642	−0.353	−0.153	0.000	0.123
6					−1.267	−0.757	−0.473	−0.275	−0.123
7						−1.352	−0.852	−0.572	−0.376
8							−1.424	−0.932	−0.656
9								−1.485	−1.001
10									−1.539

R \ A	11	12	13	14	15	16	17	18	19	20
1	1.586	1.629	1.668	1.703	1.736	1.766	1.794	1.820	1.844	1.867
2	1.062	1.116	1.164	1.208	1.248	1.285	1.319	1.350	1.380	1.408
3	0.729	0.793	0.850	0.901	0.948	0.990	1.029	1.066	1.099	1.131
4	0.462	0.537	0.603	0.662	0.715	0.763	0.807	0.848	0.886	0.921
5	0.225	0.312	0.388	0.456	0.516	0.570	0.619	0.665	0.707	0.745
6	0.000	0.103	0.191	0.267	0.335	0.396	0.451	0.502	0.548	0.590
7			0.000	0.088	0.165	0.234	0.295	0.351	0.402	0.448
8					0.000	0.077	0.146	0.208	0.264	0.315
9							0.000	0.069	0.131	0.187
10									0.000	0.062

R \ A	21	22	23	24	25	26	27	28	29	30
1	1.889	1.910	1.929	1.948	1.965	1.982	1.998	2.014	2.029	2.043
2	1.434	1.458	1.481	1.503	1.524	1.544	1.563	1.581	1.599	1.616
3	1.160	1.188	1.214	1.239	1.263	1.285	1.306	1.327	1.346	1.365
4	0.954	0.985	1.014	1.041	1.067	1.091	1.115	1.137	1.158	1.179
5	0.782	0.815	0.847	0.877	0.905	0.932	0.957	0.981	1.004	1.026
6	0.630	0.667	0.701	0.734	0.764	0.793	0.820	0.846	0.871	0.894
7	0.491	0.532	0.569	0.604	0.637	0.668	0.697	0.725	0.752	0.777
8	0.362	0.406	0.446	0.484	0.519	0.553	0.584	0.614	0.642	0.669
9	0.238	0.286	0.330	0.370	0.409	0.444	0.478	0.510	0.540	0.568
10	0.118	0.170	0.218	0.262	0.303	0.341	0.377	0.411	0.443	0.473
11	0.000	0.056	0.108	0.156	0.200	0.241	0.280	0.316	0.350	0.382
12			0.000	0.052	0.100	0.144	0.185	0.224	0.260	0.294
13					0.000	0.048	0.092	0.134	0.172	0.209
14							0.000	0.044	0.086	0.125
15									0.000	0.041

$\sum Z_R^2$					
2	•636192	11	8.978750	21	18.663020
3	1.492992	12	9.848536	22	19.652478
4	2.294100	13	10.820508	23	20.631938
5	3.195188	14	11.792974	24	21.624112
6	4.116514	15	12.773670	25	22.608800
7	5.051124	16	13.749302	26	23.595620
8	6.001636	17	14.726380	27	24.581300
9	6.953316	18	15.711630	28	25.578196
10	7.914726	19	16.692296	29	26.571152
		20	17.674844	30	27.562946

Tabela 3.4 nadaljevanje Rankiti Z_R

RANK ORDER	SIZE OF SAMPLE = N 31	32	33	34	35	36	37	38	39	40
1	2.056	2.070	2.082	2.095	2.107	2.118	2.129	2.140	2.151	2.161
2	1.632	1.647	1.662	1.676	1.690	1.704	1.717	1.729	1.741	1.753
3	1.383	1.400	1.416	1.432	1.448	1.462	1.477	1.491	1.504	1.517
4	1.198	1.217	1.235	1.252	1.269	1.285	1.300	1.315	1.330	1.344
5	1.047	1.067	1.087	1.105	1.123	1.140	1.157	1.173	1.188	1.203
6	0.917	0.938	0.959	0.979	0.998	1.016	1.034	1.051	1.067	1.083
7	0.801	0.824	0.846	0.867	0.887	0.906	0.925	0.943	0.960	0.977
8	0.694	0.719	0.742	0.764	0.786	0.806	0.826	0.845	0.863	0.881
9	0.595	0.621	0.646	0.670	0.692	0.714	0.735	0.755	0.774	0.793
10	0.502	0.529	0.556	0.580	0.604	0.627	0.649	0.670	0.690	0.710
11	0.413	0.442	0.469	0.496	0.521	0.545	0.568	0.590	0.611	0.632
12	0.327	0.358	0.387	0.414	0.441	0.466	0.490	0.514	0.536	0.557
13	0.243	0.276	0.307	0.336	0.364	0.390	0.416	0.440	0.463	0.486
14	0.161	0.196	0.228	0.259	0.289	0.317	0.343	0.369	0.393	0.417
15	0.080	0.117	0.151	0.184	0.215	0.245	0.273	0.300	0.325	0.350
16	0.000	0.039	0.076	0.110	0.143	0.174	0.203	0.232	0.258	0.284
17			0.000	0.037	0.071	0.104	0.135	0.165	0.193	0.220
18					0.000	0.035	0.067	0.099	0.128	0.156
19							0.000	0.033	0.064	0.094
20									0.000	0.031

	41	42	43	44	45	46	47	48	49	50
1	2.171	2.180	2.190	2.199	2.208	2.216	2.225	2.233	2.241	2.249
2	1.765	1.776	1.787	1.797	1.807	1.817	1.827	1.837	1.846	1.855
3	1.530	1.542	1.554	1.565	1.577	1.588	1.598	1.609	1.619	1.629
4	1.357	1.370	1.383	1.396	1.408	1.420	1.431	1.442	1.453	1.464
5	1.218	1.232	1.246	1.259	1.272	1.284	1.296	1.308	1.320	1.331
6	1.099	1.114	1.128	1.142	1.156	1.169	1.182	1.194	1.207	1.218
7	0.993	1.009	1.024	1.039	1.054	1.068	1.081	1.094	1.107	1.119
8	0.898	0.915	0.931	0.946	0.961	0.976	0.990	1.004	1.017	1.030
9	0.811	0.828	0.845	0.861	0.877	0.892	0.907	0.921	0.935	0.949
10	0.729	0.747	0.764	0.781	0.798	0.814	0.829	0.844	0.859	0.873
11	0.651	0.671	0.689	0.707	0.724	0.740	0.757	0.772	0.787	0.802
12	0.578	0.598	0.617	0.636	0.654	0.671	0.688	0.704	0.720	0.735
13	0.507	0.528	0.548	0.568	0.586	0.604	0.622	0.639	0.655	0.671
14	0.439	0.461	0.482	0.502	0.522	0.540	0.559	0.576	0.593	0.610
15	0.373	0.396	0.418	0.439	0.459	0.479	0.498	0.516	0.534	0.551
16	0.309	0.333	0.355	0.377	0.398	0.419	0.438	0.457	0.476	0.494
17	0.246	0.270	0.294	0.317	0.339	0.360	0.381	0.400	0.419	0.438
18	0.183	0.209	0.234	0.258	0.281	0.303	0.324	0.345	0.364	0.384
19	0.122	0.149	0.175	0.200	0.224	0.247	0.269	0.290	0.310	0.330
20	0.061	0.089	0.116	0.142	0.167	0.191	0.214	0.236	0.257	0.278
21	0.000	0.030	0.058	0.085	0.111	0.136	0.160	0.183	0.205	0.227
22			0.000	0.028	0.055	0.081	0.106	0.130	0.153	0.176
23					0.000	0.027	0.053	0.078	0.102	0.125
24							0.000	0.026	0.051	0.075
25									0.000	0.025

31	28.547370	41	38.477260
32	29.543360	42	39.466834
33	30.531542	43	40.464912
34	31.521988	44	41.453728
35	32.523292	45	42.526180
36	33.505148	46	43.441820
37	34.502814	47	44.437510
38	35.496144	48	45.427608
39	36.419058	49	46.426374
40	37.479478	50	47.420220

Tabela 3.5 Transformacija $u = \sqrt{y}$, $\sqrt{y+0.5}$

y	0	1	2	3	4	5	6	7	8	9
0	0.000 0.707	1.000 0.225	1.414 0.581	1.732 0.871	2.000 0.121	2.236 0.345	2.449 0.550	2.646 0.739	2.828 0.915	3.000 0.082
1	3.162 0.240	3.316 0.391	3.464 0.536	3.606 0.674	3.742 0.808	3.873 0.937	4.000 0.062	4.123 0.183	4.243 0.301	4.359 0.416
2	4.472 0.528	4.583 0.637	4.690 0.743	4.796 0.848	4.899 0.950	5.000 0.050	5.099 0.148	5.196 0.244	5.292 0.339	5.385 0.431
3	5.477 0.523	5.568 0.612	5.657 0.701	5.745 0.788	5.831 0.874	5.916 0.958	6.000 0.042	6.083 0.124	6.164 0.205	6.245 0.285
4	6.325 0.364	6.403 0.442	6.481 0.519	6.557 0.595	6.633 0.671	6.708 0.745	6.782 0.819	6.856 0.892	6.928 0.964	7.000 0.036
5	7.071 0.106	7.141 0.176	7.211 0.246	7.280 0.314	7.348 0.382	7.416 0.750	7.483 0.517	7.550 0.583	7.616 0.649	7.681 0.714
6	7.746 0.778	7.810 0.842	7.874 0.906	7.937 0.969	8.000 0.031	8.062 0.093	8.124 0.155	8.185 0.216	8.246 0.276	8.307 0.337
7	8.367 0.396	8.426 0.456	8.485 0.515	8.544 0.573	8.602 0.631	8.660 0.689	8.718 0.746	8.775 0.803	8.832 0.860	8.888 0.916
8	8.944 0.972	9.000 0.028	9.055 0.083	9.110 0.138	9.165 0.192	9.220 0.247	9.274 0.301	9.327 0.354	9.381 0.407	9.434 0.460
9	9.487 0.513	9.539 0.566	9.592 0.618	9.644 0.670	9.695 0.721	9.747 0.772	9.849 0.823	9.849 0.874	9.899 0.925	9.950 0.975

3.10

Če se y porazdeljuje v Poissonovi porazdelitvi, je zanjo

$E(y) = M$ in $\sigma_y^2 = M$. Varianca je proporcionalna poprečju. Za Poissonovo porazdelitev oziroma za vse primere, da je $\sigma_y^2 = C.M$ varianca proporcionalna M je ustrezna transformacija kvadratnega korena

$$u = \sqrt{y} \qquad (3.21)$$

Če nastopajo v poskusnih vrednostih tudi vrednosti $y \doteq 0$, uporabimo namesto zgornje transformacije transformacijo

$$u = \sqrt{y + 1/2} \qquad (3.22)$$

3.11

Če so poskusni podatki dani z rangi, porazdelitev rangov sicer ima enako varianco, če je število poskusnih enot po grupah enako, pač pa porazdelitev rangov ni normalna, ampak pravokotniška. S tabelo 3.4 transformiramo vrednosti rangov v rankite, za katere se vrednosti porazdeljujejo standardizirano normalno.

Da rankiti v zadovoljujoči meri ustrezajo standardizirani normalni porazdelitvi, preverimo na primeru rankitov za A = 50. Iz tabele rankitov dobimo pri A = 50 naslednjo frekvenčno porazdelitev rankitov:

n	R	f	f%	f'%
	1	1	2	2.3
-2				
	2-3	2	4	4.4
- 1.5				
	4-8	5	10	9.2
- 1				
	9-15	7	14	15.0

n	R	f	f%	f'%
- .5				
	16-25	10	20	19.1
0				
	26-35	10	20	19.1
.5				
	36-42	7	14	15.0
1.				
	43-47	5	10	9.2
1.5				
	48-49	2	4	4.4
2.				
	50	1	2	2.3
		50	100	100

Primerjava relativnih frekvenc rankitov f% z relativnimi frekvencami standardizirane normalne porazdelitve pokaže precejšnjo skladnost porazdelitve rankitov s standardizirano normalno porazdelitvijo.

4. ČISTI SLUČAJNOSTNI POSKUS

4.1 Kot že samo ime pove, pri čistem slučajnostnem poskusu od vseh prvin statističnega načrta vključimo le slučajnostno dodelitev postopkov poskusnim enotam in ponavljanje. Ker v tem primeru razen pravih slučajnostnih vplivov vključujemo v slučajnostne vplive in s tem v poskusni pogrešek vse določljive, a vsebinsko nepomembne faktorje, dosežemo zadovoljive rezultate o delovanju proučevanih postopkov na dva načina: a) s čim podrobnejšo opredelitvijo se kompleks določljivih, a vsebinsko nepomembnih faktorjev manjša in postane poskusno gradivo homogenejše. b) S ponavljanjem poskusa se vpliv slučajnostnih faktorjev (in z njimi tudi vpliv za raziskavo nepomembnih faktorjev) zmanjša. Na primeru čisto slučajnostnega poskusa, ki je najenostavnejši izmed načrtov statističnih poskusov, bomo raztolmačili tudi nekaj splošnih metod in postopkov analiziranja poskusnih podatkov.

4.2 Vzemimo za primer proučevanje vpliva postopka A na značilnost izdelka y kot kriterialni znak.

Da bi proučili vpliv proizvodnega postopka (A) z $A = 7$ različnimi načini proizvodnje, smo izdelali po $\bar{n} = 5$ izdelkov po posameznih načinih. Drugi pogoji proizvodnje so bili pri obravnavanju poskusnega gradiva z $n = A \cdot \bar{n} = 7 \cdot 5 = 35$ enotami enaki (opredeljujoči pogoji). $n = 35$ polizdelkov predstavlja poskusno gradivo, ki smo ga v poskusne namene iz velike množice izbrali na slučajnostni način. S tem smo zagotovili, da je poskusno gradivo kot slučajnosten vzorec reprezentant populacije. Drugič se srečamo z slučajnostjo pri dodeljevanju postopkov posameznim poskusnim enotam - polizdelkom. Če bi posamezne postopke v ponovitvah dodeljevali po vrsti sistematično po vrsti petim prvim izdelkom postopek A_1 naslednjim petim postopek A_2 in tako naprej do zadnjih petih postopek A_7, je nevarnost, da se v poskusu združi učinek proučevanega postopka s kakim določljivim, a za poskus nepomembnim faktorjem. V našem primeru bi mogel biti vrstni red izdelkov po času proizvodnje tak, da posamezni izdelki sistematično v daljšem zaporedju izvirajo iz različnih surovin, različnega splošnega tretiranja ipd. Ti vplivi se morejo združiti s proučevanim postopkom s čemer se pravi učinek proizvodnega faktorja popači. More biti npr. zadnjih 10 polizdelkov iz boljših surovin, kar nepravilno vpliva na učinek postopka A_6 in A_7. Razlike učinka A_6 in A_7 v primerjavi z učinkom drugih postopkov so nerealne. S slučajnostjo dodeljevanja postopkov poskusnim enotam se temu nezaželjenemu učinku izognemo.

4.3 Tehnično dosežemo slučajnost dodeljevanja postopkov poskusnim enotam na dva načina: ali posameznim postopkom dodelimo slučajnostno ustrezno število enot ali posameznim enotam slučajnostno dodelimo postopke. Iz oštevilčenega poskusnega gradiva (okvir vzorca) izberemo po vrsti na slučajnostnem načinu po pet polizdelkov in vsaki tako izbrani skupini po vrsti dodelimo postopek A_1, A_2 ... A_7. Npr. če so bili s tablicami slučajnostnih številk izbrane po vrsti slučajnostne številke 8 13 6 29 17, polizdelkom s temi zaporednimi številkami dodelimo postopek A_1. Podobno postopamo naprej seveda s tem, da je izbor brez ponavljanja, ker bi bil izbor s ponavljanjem brez smisla. Po drugem načinu pa izvedemo slučajnost razporeda tako, da po vrsti za vsak polizdelek slučajnostno iz skupine sedmih postopkov določimo, kateri postopek bomo uporabili na določenem polizdelku. Postopek ne pride več v konkurenco, ko je dodeljen že petkrat in je s tem število njegovih ponovitev izčrpano. Tako npr. polizdelku št. 1 ali prvemu po redu s slučajnostnim izborom postopka dodelimo, da ga bomo obravnavali po postopku A_3, drugega po postopku A_7, tretjega pa po postopku A_2 itd.

4.4 Po izvedenem poskusu, ki smo ga izvedli tako, da smo na slučajnostno dodeljenih polizdelkih dobili za proučevano kriterialno značilnost y naslednje rezultate:

y_{Ai}						y_A	$\bar{y}_A$
A_1	8	7	6	5	7	33	6,6
A_2	8	9	7	9	12	45	9,0
A_3	10	13	18	16	12	69	13,8
A_4	15	13	14	17	13	72	14,4
A_5	7	6	5	4	3	25	5,0
A_6	12	5	7	6	9	39	7,8
A_7	8	7	9	8	10	42	8,4
						$y = 325$	$9,3 = \bar{\bar{y}}$

Oznake v tabeli so dogovorno standardne. Medtem ko kriterialni znak vedno zaznamujemo z y, imajo indeksi tega znaka naslednji pomen: y_{Ai} pomeni posamezne meritve ali i-to meritev kriterialnega znaka y za postopek A. Medtem ko štejemo, da je število postopkov enako A, je število ponovitev za vsak postopek ali $\bar{n}$ (če je število ponovitev za vsak postopek enako) ali n_A, če je število ponovitev za postopke različno.

Nadalje pri vsotah podatkov izpuščamo indekse, po katerih smo seštevali, tako pomeni:

$$y_A = \sum_i y_{Ai} \qquad (4.1)$$

vsota za posamezne postopke, in

$$y = \sum_A y_A = \sum_A \sum_i y_{Ai} \qquad (4.2)$$

skupno vsoto podatkov v poskusu.

Analogno zaznamujemo s prečno črto ustrezna povprečja iz poskusnih rezultatov.

Tako pomeni

$$\bar{y}_A = y_A/n_A \qquad \bar{y} = y/n \qquad (4.3)$$

ali če je $n_A = \bar{n} = const$

$$\bar{y}_A = y_A/\bar{n} \qquad \bar{y} = y/A \cdot \bar{n} = y/n \qquad (4.4)$$

4.5 Linearen model za čisto slučajnostni načrt je sestavljen iz treh komponent: splošne aritmetične sredine M, ki je rezultat splošnih - stalnih pogojev, komponente (A), ki je rezultat vpliva faktorja oziroma postopkov (A) in slučajnostne komponente e_{Ai}. Če imajo postopki (A) značaj Modela I. je učinek postopkov dan z A vrednostmi (A), ki imajo lastnost, da je $\sum(A) = 0$. Za slučajnostno komponento e_{Ai} predpostavimo, da se porazdeljuje normalno $e_{Ai} = :N\ (0, \sigma^2)$ z matematičnim upanjem 0 in od A neodvisno varianco. V tem primeru je y_{Ai}, za kateregaje linearni model

$$y_{Ai} = M + (A) + e_{Ai} \qquad (4.5)$$

v okviru postopka normalno porazdeljena slučajnostna spremenljivka.

4.6 Analizo variance za tak primer obračunamo kot enostavno analizo variance po naslednjem postopku:

Izračunamo iz osnovnih podatkov y_{Ai}

a) osnovne vsote

$$y_{Ai} \qquad y_A = \underset{i}{S}\, y_{Ai} \qquad y = {}_A\underset{i}{S}\, y_{Ai} = \sum y_A \qquad (4.6)$$

Iz y_{Ai}, y_A in y izračunamo pomožne količine

b)

$$Q_{Ai} = \underset{A}{\Sigma}\, S_i y^2_{Ai} \qquad Q_A = \frac{A}{n} \sum y^2_A \qquad Q = \frac{1}{n}\, y^2 \qquad (4.7)$$

c) S tem, da odštejemo od pomožnih količin Q, izračunamo pomožne **reducirane količine**.

$$q_{Ai} = Q_{Ai} - Q \qquad q_A = Q_A - Q \qquad (4.8)$$

d)
Analizo variance obračunamo po naslednji shemi: (4.9)

Vir variacije VV	Vsota kvadratov odklonov	Število stopenj prostosti m	Ocena variance s^2	$F = s^2/s^2_e$	$E(s^2)$
1	2	3			
Postopek	$K_A = q_A$	$m_A = A - 1$	$s^2_A = K_A/m_A$	$F_A = \frac{s^2_A}{s^2_e}$	$\sigma^2_e + \bar{n}\sigma^2_A$
poskusni pogrešek	$K_e = q_{Ai} - q_A$	$m_e = (\bar{n}-1)A$	$s^2_e = K_e/m_e$	1	σ^2_e
Total	$K = q_{Ai}$	$m = n - 1$	-	-	-

F_A preko kritičnih vrednosti za F-porazdelitev preskušamo značilnost σ^2_A oziroma razlik med (A). Iz izrazov za $E(s^2)$, dobimo tudi idejo, kako oceniti vrednost σ^2_A po obrazcu:

$$\hat{\sigma}^2_A \quad \frac{s^2_A - s^2_e}{\bar{n}} \tag{4.10}$$

4.7 Če izvedemo za naš primer po nakazanem postopku analizo variance, dobimo po vrsti:

a) ustrezne tabele y_{Ai} y_A y imamo podane že v tabeli ...

b) vrednosti pomožnih količin Q_{Ai}, Q_A , Q in Q_i Q_A so:

$$Q_{Ai} = \sum_A S_i\, y^2_{Ai} = 8^2 + 7^2 + \dots 8^2 + 10^2 = 3511$$

$$Q_A = \frac{A}{n} \sum y^2_A = \frac{7}{35} \left[33^2 + 45^2 + \dots + 42^2 \right] = \frac{1}{5}\, 16969 = 3393{,}8$$

$$Q = \frac{1}{n}\, y^2 = \frac{1}{35} 325^2 = 105625 = 3017{,}86$$

$$q_{Ai} = 3511 - 3017{,}86 = 495{,}14$$

$$q_A = 3393{,}8 - 3017{,}86 = 375{,}94$$

4.8 Intervalne ocene učinka postopka.

Ker je matematično upanje za aritmetično sredino postopka

$$E\bar{y}_A = E(M + (A) + \bar{e}_A) = M + (A) \qquad (4.11)$$

je $\bar{y}_A$ nepristranska ocena za skupen učinek, ki izvira iz opredeljujočih pogojev M in učinka postopka. Iz znanih zakonitosti o aritmetičnih sredinah vzorcev posnamemo, da je varianca

$$\mathrm{Var}\ \bar{y}_A = \frac{\sigma_e^2}{\bar{n}} \quad \text{ocena variance pa} \quad \mathrm{var}\bar{y}_A = \frac{s_e^2}{\bar{n}}$$

Iz zakonitosti t porazdelitve pa sledi, da je intervalna ocena za aritmetično sredino s tveganjem α

$$M_A = M + (A) = \bar{y}_A \pm t_{\alpha/2}\left[m_e = A(\bar{n} - 1)\right] \frac{s_e}{\sqrt{\bar{n}}} \qquad (4.12)$$

kritična vrednost za t se ravna po m_e, ker je s_e izračunan z m_e tolikimi stopinjami prostosti.

V našem primeru je standardni pogrešek povprečij

$$s_e(\bar{y}_A) = \sqrt{\frac{s_e^2}{\bar{n}}} = \sqrt{\frac{4,19}{5}} = 0.915$$

Maksimalni verjeten odklon pa

$$d\bar{y}_A = t_{\alpha/2}(me)\ se(\bar{y}_A) = t_{\alpha/2}(m_e = 28)\ se(\bar{y}_A) = 2,05 \cdot 0.915$$

4.9 Najmanjša značilna razlika NZR

Pri primerjavi ocene aritmetičnih sredin dveh populacij ocenimo varianco razlike dveh aritmetičnih sredin po obrazcu

$$\mathrm{var}\ \bar{y} = s_e^2\left[\frac{1}{n_1} + \frac{1}{n_2}\right] = s_e^2\ \frac{n_1 + n_2}{n_1\, n_2} \qquad (4.13)$$

pri čemer je

$$s_e^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)\, s_2^2}{n_1 + n_2 - 2} \text{ združena varianca.} \quad (4.14)$$

Pri primerjavi aritmetičnih sredin, dobljenih iz analize variance, izračunamo varianco razlike aritmetičnih sredin analogno le, da je s_e^2 varianca poskusnega pogreška. V posebnem primeru, da je število ponovitev vsakega postopka $n_A = \bar{n} = const$ je

$$\text{var } \Delta\bar{y} = \frac{2\, s_e^2}{\bar{n}} \quad (4.15)$$

maksimalna verjetna razlika (pri H_o) ali najmanjša značilna razlika (pri H_1) pa je analogno

$$d(\Delta\bar{y}_A) = t_{\alpha/2}(m_e) \sqrt{\frac{2\, s_e^2}{\bar{n}}} \quad (4.16)$$

za naš primer je $d(\Delta\bar{y}_A) = 2.05 \sqrt{\frac{2 \cdot 4.19}{5}} = 2,66$

Najmanjša značilna razlika je bila dolgo najpogosteje uporabljan pokazovalec, na osnovi katerega so odločali ali je razlika poprečij dveh postopkov značilna ali ne.

4.10 P o s t e r i o r n a a n a l i z a.

Vendar je pri uporabi najmanjše značilne razlike treba biti oprezen. Neoporečna je njena uporaba le, če sta v poskusu dva postopka. Če je postopkov $A > 2$ število vseh možnih razlik med popreči postopkov $\binom{A}{2} = \frac{A(A-1)}{2}$ hitro narašča. V tem primeru je velika verjetnost, da bo vsaj ena izmed razlik med popreči večja, kot je $d(\Delta\bar{y})$. Varia[illegible]cijski razmak $R\bar{y} = \bar{y}_{max} - \bar{y}_{min}$ v poskusu za primer da so $M_A = M$, namreč ne zavisi samo od m_e, s_e^2 in $\bar{n}$ temveč tudi od števila postopkov A, oziroma vzorcev. Z večanjem števila popreči se veča tudi $R(\bar{y})$. Zato moremo npr. z uporabo najmanjše značilne razlike $d(\Delta\bar{y})$ sklepati, da je npr. $\bar{y}_{max} - \bar{y}_{min}$ značilen, čeprav gre za popreča iz iste populacije ali za primer, da je $\sigma_A^2 = 0$.

Najmanjša značilna razlika je torej premajhna za vse razlike med ocenami popreča. Čim več je popreči med primerjanima aritmetičnima sredinama, tem večja je napaka α pri sklepanju na osnovi NZR. Poznamo več metod, s katerimi odklonimo pristranost pri sklepanju z NZR.

4.11 Ena izmed njih je Tukeyeva "resnično značilna razlika" (honestly significant difference).

Resnično značilna razlika (RZR) postavlja značilnost razlik na nivo napake α za vse pare razlik med sredinami. Ta preskus je vezan na vse omejitve, na katere je vezana analiza variance.

Po preskusu resnično značilne razlike popreči j dveh postopkov je razlika značilna če je absolutno večja kot

$$W_A = RZR = q_\alpha(m_e;A)\ se(\bar{y}) = q_\alpha(m_e;A)\sqrt{\frac{s_e^2}{\bar{n}}} \qquad (4.17)$$

$q_\alpha(m_e;A)$ je dobljen iz ocene studentiziranega razmaka

$$\frac{\bar{y}_{max} - \bar{y}_{min}}{s_{\bar{y}}}$$

Ustrezna tabela $q_\alpha(m_e;A)$ je dana v tabeli 4.1

Za naš primer je interpolirana vrednost

$$q_{\alpha=0.05}(m_e=28;A=7) = 4,49$$

in $RZR = q_{\alpha=0.05}(m_e=28;A=7)\sqrt{\frac{s_e^2}{\bar{n}}} = 4,49.0,915 = 4,1$

S tveganjem α = 0.05 resnično značilna razlika je 4,1 za razliko od najmanjše značilne razlike 2,66.

Če napravimo pregled vseh možnih razlik za naš primer dobimo iz rangiranih popreči j trikotniško matriko možnih razlik.

Tabela 4.1 KRITIČNE MEJE "STUDENTIZIRANEGA RAZMIKA"

$$q_{\alpha} = \frac{\bar{y}_{max} - \bar{y}_{min}}{s_{\bar{y}}}$$

me	α	p = število sredin postopkov								
		2	3	4	5	6	7	8	9	1o
5	.o5	3.64	4.6o	5.22	5.67	6.o3	6.33	6.58	6.8o	6.99
	.o1	5.7o	6.97	7.8o	8.42	8.91	9.32	9.67	9.97	1o.24
6	.o5	3.46	4.34	4.9o	5.31	5.63	5.89	6.12	6.32	6.49
	.o1	5.24	6.33	7.o3	7.56	7.97	8.32	8.61	8.87	9.1o
7	.o5	3.34	4.16	4.68	5.o6	5.36	5.61	5.82	6.oo	6.16
	.o1	4.95	5.92	6.54	7.o1	7.37	7.68	7.94	8.17	8.37
8	,o5	3.26	4.o4	4.53	4.89	5.17	5.4o	5.6o	5.77	5.92
	.o1	4.74	5.63	6.2o	6.63	6.96	7.24	7.47	7.68	7.87
9	.o5	3.2o	3.95	4.42	4.76	5.o2	5.24	5.43	5.6o	5.74
	.o1	4.6o	5.43	5.96	6.35	6.66	6.91	7.13	7.32	7.49
1o	.o5	3.15	3.88	4.33	4.65	4.91	5.12	5.3o	5.46	5.6o
	.o1	4.48	5.27	5.77	6.14	6.43	6.67	6.87	7.o5	7.21
11	.o5	3.11	3.82	4.26	4.57	4.82	5.o3	5.2o	5.35	5.49
	.o1	4.39	5.14	5.62	5.97	6.25	6.48	6.67	6.84	6.99
12	.o5	3.o8	3.77	4.2o	4.51	4.75	4.95	5.12	5.27	5.4o
	.o1	4.32	5.o4	5.5o	5.84	6.1o	6.32	6.51	6.67	6.81
13	.o5	3.o6	3.73	4.15	4.45	4.69	4.88	5.o5	5.19	5.32
	.o1	4.26	4.96	5.4o	5.73	5.98	6.19	6.37	6.53	6.67
14	.o5	3.o3	3.7o	4.11	4.41	4.64	4.83	4.99	5.13	5.25
	.o1	4.21	4.89	5.32	5.63	5.88	6.o8	6.26	6.41	6.54
15	.o5	3.o1	3.67	4.o8	4.37	4.6o	4.78	4.94	5.o8	5.2o
	.o1	4.17	4.83	5.25	5.56	5.8o	5.99	6.16	6.31	6.44
16	.o5	3.oo	3.65	4.o5	4.33	4.56	4.74	4.9o	5.o3	5.15
	.o1	4.13	4.78	5.19	5.49	5.72	5.92	6.o8	6.22	6.35
17	.o5	2.98	3.63	4.o2	4.3o	4.52	4.71	4.86	4.99	5.11
	.o1	4.1o	4.74	5.14	5.43	5.66	5.85	6.o1	6.15	6.27
18	.o5	2.97	3.61	4.oo	4.28	4.49	4.67	4.82	4.96	5.o7
	.o1	4.o7	4.7o	5.o9	5.38	5.6o	5.79	5.94	6.o8	6.2o
19	.o5	2.96	3.59	3.98	4.25	4.47	5.65	4.79	4.92	5.o4
	.o1	4.o5	4.67	5.o5	5.33	5.55	5.73	5.89	6.o2	6.14
2o	.o5	2.95	3.58	3.96	4.23	4.45	4.62	4.77	4.9o	5.o1
	.o1	4.o2	4.64	5.o2	5.29	5.51	5.69	5.84	5.97	6.o9
24	.o5	2.92	3.53	3.9o	4.17	4.37	4.54	4.68	4.81	4.92
	.o1	3.96	4.54	4.91	5.17	5.37	5.54	5.69	5.81	5.92
3o	.o5	2.89	3.49	3.84	4.1o	4.3o	4.46	4.6o	4.72	4.83
	.o1	3.89	4.45	4.8o	5.o5	5.24	5.4o	5.54	5.65	5.76
4o	.o5	2.86	3.44	3.79	4.o4	4.23	4.39	4.52	4.63	4.74
	.o1	3.82	4.37	4.7o	4.93	5.11	5.27	5.39	5.5o	5.6o

44

60	.05	2.83	3.40	3.74	3.98	4.16	4.31	4.44	4.55	4.65
	.01	3.76	4.28	4.60	4.82	4.99	5.13	5.25	5.36	5.45
120	.05	2.80	3.36	3.69	3.92	4.10	4.24	4.36	4.48	4.56
	.01	3.70	4.20	4.50	4.71	4.87	5.01	5.12	5.21	5.30
∞	.05	2.77	3.31	3.63	3.86	4.03	4.17	4.29	4.39	4.47
	.01	3.64	4.12	4.40	4.60	4.76	4.88	4.99	5.08	5.16

		A_4	A_3	A_2	A_7	A_6	A_1	A_5
		14.4	13.8	9.0	8.4	7.8	6.6	5.0
A_5	5.0	9.4	8.8	4.0	3.4	2.8	1.6	-
A_1	6.6	7.8	7.2	2.4	1.8	1.2	-	
A_6	7.8	6.6	6.0	1.2	0.6	-		
A_7	8.4	6.0	5.4	0.6	-			
A_2	9.0	5.4	4.8	-				
A_3	13.8	0.6	-					
A_4	14.4	-						

Po metodi najmanjših značilnih razlik značilne razlike so obkrožene z ... po Tukeyevi metodi resnično značilnih razlik pa z ... Razlike značilne po NRZ 4,0 3,4 in 2.8 so se po metodi RZR pokazale kot neznačilne.

4.12 Tukeyeva metoda resnično značilnih razlik bazira na poprečni značilni razliki s tveganjem $\alpha = 0.05$.

Kot zavisi značilnost celotnega variacijskega razmaka med sredinami od skupnega števila postopkov A, zavisi značilnost med dvema sredinama od tega, koliko poprečij se nahaja med primerjanima poprečjema. Metodo variacijskega razmaka moremo prenesti in specificirati na lokalne variacijske razmake med poprečji. Po tem principu se ravna značilnost razlik po tem, variacijski razmak kolikih poprečij ti sredini opredeljujeta. Če med njima ni nobene vrednosti opredeljujeta variacijski razmak 2 sredini, če je med njima ena sredina, opredeljujeta variacijski razmak treh sredin, če je med sredinama K - 2 sredin, opredeljujeta variacijski razmak K sredin.

Na tej metodi je osnovan Student - Newman - Keulov preskus razlik, pri čemer je

$$W_\alpha(m_e;K) = Q_\alpha(me;K).se(\bar{y}) \qquad (4.18)$$

4.13 Za naš primer so

$W_{\alpha}(m_e:K) = Q_{\alpha}(28;K)\ se(\bar{y})$ $\quad q_{0.01}(28;K)se(\bar{y})$

$W_{0.05}(28;2) = 2.90 \cdot .915 = 2,67$ $\quad 3.92 \cdot .915 = 3.59$

$W_{0.05}(28;3) = 3.50 \cdot .915 = 3.20$ $\quad 4.48 \cdot .915 = 4.10$

$W_{0.05}(28;4) = 3.86 \cdot .915 = 3.53$ $\quad 4.83 \cdot .915 = 4.42$

$W_{0.05}(28;5) = 4.13 \cdot .915 = 3.78$ $\quad 5.09 \cdot .915 = 4.66$

$W_{0.05}(28;6) = 4.32 \cdot .915 = 3.95$ $\quad 5.28 \cdot .915 = 4.83$

$W_{0.05}(28;7) = 4.48 \cdot .915 = 4.10$ $\quad 5.45 \cdot .915 = 4.99$

Pri tej metodi postopoma primerjamo razlike poprečij z ustreznimi W tako, da ugotavljamo ali so po vrsti diagonalne razlike (razlike z istimi K) večje od ustreznih kritičnih vrednosti W. Tako je razmak z K = 7 razmak W = 9.4 med A_5 in A_4 značilen, ker je večji od $W_{0.05}(28;7) = 4.10$. Enako sta značilna razmaka $A_1 - A_4$ in $A_5 - A_3$ z 7.8 in 8.8, ker sta večja kot $W_{0.05}(28,6) = 3.95$ in tako naprej. V tabeli so naznačene značilne razlike na stopnji $\alpha = 0.05$ (...) in na stopnji $\alpha = 0.01$ (—)

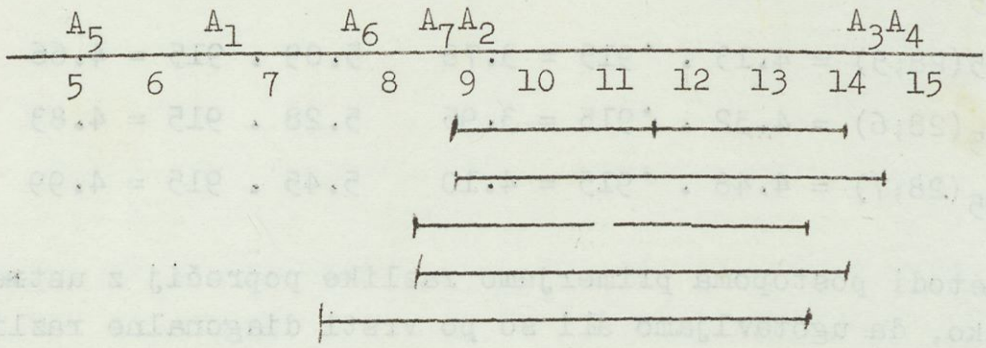
A		A_4	A_3	A_2	A_7	A_6	A_1	A_5
	$\bar{y}$	14.4	13.8	9.0	8.4	7.8	6.6	5.0
A_5	5.0	9.4	8.8	4.0	3.4	2.8	1.6	
A_1	6.6	7.8	7.2	2.4	1.8	1.2		
A_6	7.8	6.6	6.0	1.2	0.6			
A_7	8.4	6.0	5.4	0.6				
A_2	9.0	5.4	4.8					
A_3	13.8	0.6						
A_4	14.4							

Dobljene kritične diference W uporabljamo tudi za določanje intervalnih ocen.

Tako je npr. intervalna ocena

$$M_4 - M_5 = (\bar{y}_4 - \bar{y}_5) \pm W_{0.05}(28.7) = 9.4 \pm 4.1$$

Tabela značilnosti razlik more služiti tudi za združevanje postopkov v homogene skupine



Z analizo značilnosti razlik spoznamo, da tvorijo postopki A_5 A_1 A_6 A_7 A_2 homogeno celoto in ni med njimi značilnih razlik. Enako je s skupino A_3.A_4 poprečja prve skupine pa so od poprečij druge skupine značilno različna.

Za $K = 2$ je $W_\alpha(m_e;2) = t_\alpha(m_e) . \sqrt{2}$, kar je v skladu z definicijo najmanjše značilne razlike.

<u>A priori načrtovane multiple primerjave.</u>

4.14 Splošen F-preskus analize variance je pogosto šele prva stopnja v analizi podatkov, ki jih dobimo s poskusom. Splošen F-preskus pove samo ali so ali niso med postopki značilne razlike.

Z a posteriono analizo podatkov analiziramo značilnost razlik med dobljenimi poprečji po opravljenem poskusu. Pogosto pa moremo podobno kot napravimo splošno domnevo, da so med postopki razlike oziroma da proučevan faktor vpliva na pojav, vnaprej - apriori napraviti določene posebne domneve o zvezah med vrednostmi faktorjev na različnih nivojih.

Najenostavnejša taka a priori ničelna domneva je, da sta učinka dveh izmed več postopkov enaka, kar izrazimo z ničelno poddomnevo $H_o : M_1 = M_2$, ki jo moremo izraziti tudi kot
$H_o : M_1 - M_2 = 0$ Ustrezna protihipoteza je
$H_1 : M_1 - M_2 \neq 0$

Podobno moremo postaviti tudi ničelno poddomnevo da je učinek postopka A_3 enak poprečnemu učinku postopkov A_1 in A_2. To ničelno poddomnevo izrazimo z

$$H_o : \frac{M_1 + M_2}{2} = M_3 \quad \text{ali} \quad H_o : \frac{M_1 + M_2}{2} - M_3 = 0.$$

Ustrezna protidomneva je $H_1 : \dfrac{M_1 + M_2}{2} - M_3 \neq 0$

Ustrezne ocene teh količin dobimo iz ocen aritmetičnih sredin poskusa y_A, če prave vrednosti sredin zamenjamo z ocenami. Tako je ustrezna ocena za

$M_1 - M_2$ enaka $\bar{y}_2 - \bar{y}_1 = U_1$ ali za

$$\frac{M_1 + M_2}{2} - M_3; \quad \frac{\bar{y}_1 + \bar{y}_2}{2} - \bar{y}_3 = U_2$$

4.15 Ker v primeru $\bar{y}_2 - \bar{y}_1$ primerjamo $\bar{y}_2$ z $\bar{y}_1$, v primeru $\dfrac{\bar{y}_1 + \bar{y}_2}{2} - \bar{y}_3$ poprečje $\dfrac{\bar{y}_1 + \bar{y}_2}{2}$ primerjamo z $\bar{y}_3$ imenujemo te linearne izraze med poprečji <u>primerjave</u>.

Na splošno opredelimo kot primerjavo vsako linearno zvezo med ocenami aritmetičnih sredin

$$\sum_{k=1}^{A} c_k \bar{y}_k = U \qquad (4.19)$$

ki ima to lastnost, da je vsota koeficientov c_K enaka nič.

$$\sum_{K=1}^{A} c_k = 0 \qquad (4.20)$$

Za preskus ugotovimo, da je $u_1 = \bar{y}_1 - \bar{y}_2$ primerjava v smislu zgornje opredelitve.

$$u_1 = (1)\bar{y}_1 + (-1)\bar{y}_2 \quad \text{ali}$$

$$(1) + (-1) = 0$$

Iz istega razloga je primerjava tudi

$$u_2 = \frac{\bar{y}_1 + \bar{y}_2}{2} - \bar{y}_3 = (\tfrac{1}{2})\,\bar{y}_1 + (\tfrac{1}{2})\bar{y}_2 + (-1)\bar{y}_3;$$

ker je $\frac{1}{2} + \frac{1}{2} + (-1) = 0$

Tej primerjavi ekvivalentna primerjava je tudi

$$u_2 = \bar{y}_1 + \bar{y}_2 - 2\bar{y}_3\text{, ker je } 1 + 1 - 2 = 0$$

Vsaka primerjava nakazuje določeno vsebinsko utemeljeno zvezo med aritmetičnimi sredinami. Če upoštevamo opredelitev primerjave in zakonitosti linearnih zvez slučajnostnih spremenljivk dobimo, kako se porazdeljuje primerjava U kot linearna zveza med aritmetičnimi sredinami.

Če izhajamo iz modela

$$y_{Ai} = M + (A) + e_{Ai} \qquad (4.21)$$

so aritmetične sredine postopkov

$$\bar{y}_A = M + (A) + \bar{e}_A = M_A + \bar{e}_A \qquad (4.22)$$

primerjava U pa

$$U = C_A\bar{y}_A = C_A M_A + C_A\bar{e}_A \qquad (4.23)$$

Ker se $\bar{e}_A$ porazdeljuje, se porazdeljuje normalno tudi U, ker je linearna zveza neodvisnih slučajnostnih spremenljivk $\bar{e}_A$.

Pri tem je

$$E_n = E(\sum C_A M_A + \sum C_A\bar{e}_A) = \sum C_A M_A + \sum C_A E\bar{e}_A = \sum C_A M_A \qquad (4.24)$$

ker je $E\bar{e}_A = 0$ po definiciji.

$$Var_n = E\{U - E(U)\}^2 = E(\sum C_A\bar{e}_A)^2 = \sum C_A^2 \frac{\sigma_e^2}{i_A} \qquad (4.25)$$

V posebnem primeru, če je $n_A = \bar{n} =$ konst je

$$Var_n = \frac{\sigma_e^2}{\bar{n}} \sum C_A \qquad (4.26)$$

Če strnemo dobljene zveze, dobimo, da je U normalno porazdeljen z naslednjimi parametri

$$n = \sum C_A\bar{y}_A = : N(\sum C_A M_A ; \frac{\sigma_e^2}{\bar{n}} \sum C_A) \qquad (4.27)$$

Iz lastnosti normalnih porazdelitev povzamemo, da se

$$\frac{\bar{n}\,(U - \sum C_A M_A)^2}{\sigma_e^2 \sum C_A^2} = : x^2(1) \qquad (4.28)$$

porazdeljuje v x^2 - porazdelitvi z eno stopinjo prostosti. Ker je

$$m_e\, s_e^2/\sigma_e^2 = : x^2(m_e)$$

velja, da se izraz

$$\frac{\bar{n}(n - U)^2}{\sigma_e^2 \sum C_A^2} \Big/ \frac{s_e^2}{\sigma_e^2} = \frac{\bar{n}\,(n - U)^2}{s_e^2 \sum C_A^2} = \frac{x^2(1)}{x^2(m_e)} = : F(1:m_e) \quad (4.29)$$

porazdeljuje v F-porazdelitvi z $m_1 = 1$ in $m_2 \doteq m_e$ stopinjami prostosti.

Iz tega izraza dobimo dva pomembna rezultata:

1. Pri ničelni domnevi $H_o : U = \sum C_A M_A = 0$ uporabljamo izraz

$$\frac{\bar{n}\, n^2}{s_e^2 \sum C_A^2} = \frac{\bar{n}\,(\sum C_A \bar{y}_A)^2}{s_e^2 \sum C_A^2} = : F(1.m_e) \quad (4.30)$$

za preskušanje značilnosti primerjave n.

2. Ker pa velja, da je $\sqrt{F(1,m_e)} = t(m_e)$

moremo izraz

$$\frac{(n - U)\sqrt{\bar{n}}}{s_e \sqrt{\sum C_A^2}} = : t(m_e) \quad (4.31)$$

uporabiti za preskušanje domnev o primerjavi n, obenem pa tudi za določanje intervalne ocene za U.

Iz izraza 4.31 namreč sledi:

$$\Pr\left(n - t_{\alpha/2}(m_e)s_e\sqrt{\frac{C_A^2}{\bar{n}}} < U < n + t_{\alpha/2}(m_e)s_e\sqrt{\frac{\sum C_A^2}{\bar{n}}}\right) = 1 - \alpha \quad (4.32)$$

4.16 O r t o g o n a l n e   p r i m e r j a v e.

Primerjavi $n_1 = \sum C_{A1}\bar{y}_A$ in $n_2 = \sum C_{A2}\bar{y}_A$ sta ortogonalni, če

je na splošno
$$\sum_{A=1}^{A} \frac{C_{A1}\, C_{A2}}{n_A} = 0 \qquad (4.33)$$

ali v posebnem, če je število ponovitev $n_A = \bar{n} = const$, če je skalarni produkt koeficientov dveh primerjav
$$\sum_{A=1}^{AA} C_{A1}\, C_{A2} = 0 \qquad (4.34)$$
enak nič. Tako sta primerjavi $i_A = 0$ $n_1 = \bar{y}_1 - \bar{y}_2$ in $n_2 = \bar{y}_1 + \bar{y}_2 - 2\bar{y}_2$ ortogonalni primerjavi. Da sta primerjavi sledi iz tega da je $1 - 1 = 0$ in $1 + 1 - 2 = 0$, da sta ortogonalni pa iz tega, da je $(1)(1) + (-1)(1) + (0)(-2) = 0$.

Vsaka primerjava n prispeva k vsoti kvadratov postopkov K_A
$$K_{Cj} = \frac{(\sum_A C_{Aj}\bar{y}_A)^2}{\sum \frac{C_A^2}{n_A}} \qquad (4.35)$$
z eno stopnjo prostosti.

Ker je
$$EK_{Cj} = \sigma_e^2 + \frac{\left[\sum C_{Aj} M_A\right]^2}{\sum \frac{C_A^2}{\bar{n}_A}} \qquad (4.36)$$
je K_{Cj} v primeru, da velja $H_o : \sum C_{Aj} M_A = 0$ nepristrana ocena za σ_e^2 z eno stopinjo prostosti

Cochranov teorem: Z A - 1 med seboj ortogonalnimi primerjavami moremo razstaviti vsoto kvadratov K_A na A - 1 prispevkov ortogonalnih primerjav tako, da je

$$K_A = \sum_{j=1}^{A-1} K_{Cj} \qquad (4.37)$$

S tem postopkom moremo skupno (amorfno) vsoto kvadratov K_A2 z m_A stopinjami prostosti razstaviti na m_A delov K_{Cj} po eno stopinjo prostosti. Tako uspemo podrobneje analizirati odnose med učinki postopkov.

Vendar je treba opozoriti, da je metoda ortogonalnih primerjav korektna le, če primerjave načrtujemo vnaprej, ne pa da jih prilagodimo poskusnim rezultatom že izvršenega poskusa.

V praksi moremo razdeliti vseh m_A stopinj prostosti na m_A prispevkov z individualnimi stopinjami prostosti, ali pa izdvojimo le nekaj najpomembnejših.

4.17

Če vzamemo za primer poskus, v katerem preskušamo učinkovitost petih postopkov, od katerih se prva dva razlikujeta od naslednjih treh po določeni komponenti. Učinkovitost primerjamo s standardno tehniko, ki jo vključimo v poskus kot šesti postopek. Določiti je treba ortogonalni primerjavi, s katerima preskušamo: a) ali je skupina novih postopkov značilno različna od standardnega in b) ali kažeta grupi dveh in grupi treh postopkov značilne razlike. Če je število ponovitev enako za vse postopke, so ustrezne primerjave dane z naslednjimi koeficienti:

	A_1	A_2	A_3	A_4	A_5	A_6
	$\bar{y}_1$	$\bar{y}_2$	$\bar{y}_3$	$\bar{y}_4$	$\bar{y}_5$	$\bar{y}_6$
a	1	1	1	1	1	-5
b	3	3	-2	-2	-2	0
c	1	-1	0	0	0	0
d	0	0	1	-1	0	0
e	0	0	1	1	-2	0

Primerjava $n_a = \bar{y}_1 + \bar{y}_2 + \bar{y}_3 + \bar{y}_4 + \bar{y}_5 - 5\bar{y}_6$. Ta primerjava je identična s primerjavo

$$n_a = \frac{\bar{y}_1 + \bar{y}_2 + \bar{y}_3 + \bar{y}_4 + \bar{y}_5}{5} - \bar{y}_6$$

ki kaže na primerjavo popreči j novih postopkov s standardnim. Pri tem je treba opozoriti, da je bistvo primerjave, da je ničelna domneva

$$H_o : \left(\frac{\sum_{A=1}^{5} M_A}{5} = M_6 \right)$$

v tem, da je poprečje novih postopkov (prvih petih) enako poprečju standardnega postopka in ne, da so učinki posameznih postopkov enaki.

Analogno moremo primerjavo

$$n_2 = 3\bar{y}_1 + 3\bar{y}_2 - 2\bar{y}_3 - 2\bar{y}_4 - 2\bar{y}_5 \quad \text{prevesti v obliko}$$

$$n_2 = \frac{\bar{y}_1 + \bar{y}_2}{2} - \frac{\bar{y}_3 + \bar{y}_4 + \bar{y}_5}{3}$$

S to primerjavo preskušamo poprečen učinek prve vrste postopkov s poprečnim učinkom postopkov druge vrste.

Ortogonalne primerjave c d in e so primerjave med postopkoma prve vrste (1,2) d in e pa med postopki druge vrste (3,4,5).

Čisto slučajnostni poskus z neenakim številom ponovitev $n_A \neq$ const

4.18 V splošnem neenako število ponovitev pri posameznih poskusih povzroča težave pri obračunavanju poskusnih podatkov tako, da pogosto izločimo ustrezno število poskusnih rezultatov, da dosežemo izenačenost ponovitev ali pa določene podatke, če teh ni preveč, ocenimo, da dosežemo izenačenje.

V najenostavnejšem primeru čisto slučajnostnega poskusa pa pri obračunu ni posebnih težav. Zato se takih primerov, če zanje obstojajo pogoji, ne izogibamo. Model čisto slučajnostnega poskusa je slejkoprej isti

$$y_{Ai} = M + (A) + e_{Ai} \qquad (4.38)$$

s pogoji:

$$\sum_{A=1}^{A} n_A (A) = 0 \qquad S_A^2 = \frac{1}{A-1} \sum \frac{n_A^2}{n} (A)^2 \qquad (4.39)$$

$$e_{Ai} = : N(0, \sigma_e^2) \qquad \bar{e}_A = : N\left(0, \frac{\sigma_e^2}{n_A}\right) \qquad (4.40)$$

osnovne tabele so:

$$y_{Ai} \qquad Y_A \qquad Y \qquad (4.41)$$

Pomožne količine:

$$Q_{Ai} = \sum_{Ai} Sy_{Ai}^2 \qquad Q_A = \sum_{A} \frac{y_A^2}{n_A} \qquad Q = \frac{y^2}{n} \qquad (4.42)$$

$$q_{Ai} = Q_{Ai} - Q \qquad q_A = Q_A - Q$$

Analiza variance: (4.43)

VV	K	m	$s^2 = \frac{K}{m}$	F	Es^2
A	$K_A = q_A$	$m_A = A-1$	s_A^2	$F_A = s_A^2 / s_e^2$	$\sigma_e^2 + \bar{n}\,\sigma_A^2$
pp	$K_e = q_{Ai} - q_A$	$m_e = n-A$	s_e^2	1	σ_e^2
SK	$K = q_{Ai}$	$m = n-1$	-	-	-

4.19 Za primer vzemimo študijo o vplivu razvojnega materiala na prodajo krompirja.

Trije postopki so:

A_1 = predpakirano v prozornih plastičnih vrečkah

A_2 = pakirano v močnih papirnatih vrečah

A_3 = raztresen v razstavnih košarah

Kot poskusno gradivo smo izbrali 20 trgovin istega tipa in na slučajnosten način izbrali izmed njih

n_1 = 5 trgovin, ki naj bi prodajale po postopku A_1

n_2 = 10 trgovin, ki naj bi prodajale po postopku A_2

n_3 = 5 trgovin, ki naj bi prodajale po postopku A_3

Kot kriterialni znak je vzeta prodana količina na 10.000 din skupnega prometa. S to opredelitvijo smo iz kriterialnega znaka izločili velikost trgovine oziroma zaledja, ki bistveno vpliva na višino prodaje.

Rezultati izvršenega poskusa so:

Y_{Ai}											Y_A	n_A	$\bar{Y}_A$
A_1	49	36	47	23	40						195	5	39
A_2	29	26	30	39	45	13	32	18	38	40	310	10	31
A_3	12	16	23	28	16						95	5	19
											Y = 600	n=20	30

Pomožne količine so:

$$Q_{Ai} = \sum y_{Ai}^2 = 49^2+36^2+\ldots+28^2+16^2 = 20528 \qquad q_{Ai} = Q_{Ai}-Q = 2528$$

$$Q_A = \frac{Y_A^2}{n_A} = \frac{195^2}{5} + \frac{310^2}{10} + \frac{95^2}{5} = 19020 \qquad q_A = Q_A-Q = 1020$$

$$Q = Y^2/n = \frac{600^2}{20} = 18000$$

Analiza variance:

VV	K	m	s^2	F
A	$K_A = q_A = 1020$	A-1=2	$s_A^2 = 510$	$F_A = 510/88{,}70 = 5{,}75$
pp	q_{Ai} 2528 $-q_A$ 1020 K_e 1508	n-1=17	$s_e^2 = 88{,}70$	1
SK	$K = q_{Ai} = 2528$	m = 19	-	

Ker je kritična vrednost $F_{o.o5}^{(2,17)} = 3{,}59$ $F_{o.o1}^{(2,17)} = 6{,}11$, smatramo, da $F_A = 5{,}75$ nakazuje razlike v deležu prodaje v odvisnosti od načina embaliranja.

Vzemimo, da nas ne zanima le splošna ugotovitev, ali so razlike značilne, temveč smo vnaprej načrtovali, da primerjamo embalirano proti razsutemu blagu: Temu ~~problema~~ ustrezna primerjava je

$$u = \bar{y}_1 + \bar{y}_2 - 2\bar{y}_3 = 39 + 31 - 2.19 = 32$$

Tej primerjavi ustrezen del iz vsote kvadratov je:

$$\frac{u^2}{\sum \frac{C_K^2}{n_K}} = \frac{32^2}{\frac{1^2}{5} + \frac{1^2}{10} + \frac{2^2}{5}} = \frac{32^2}{\frac{11}{10}} = \frac{10240}{11} = 931$$

S to primerjavo je z eno stopnjo prostosti pojasnjene 931 od skupno 1020 vsote od K_A z dvema stopinjama prostosti.

Če ustrezno predelamo analizo variance, dobimo:

VV	K	m	s^2	F
A skupno	1020	2		
1+2:3	931	1	931	$10,5^{xxx}$
ostanek	89	1	89	1,0
pp	1508	17	88,70	10
SK	2528	19	-	-

Ker je $F_{o.ol}^{(1,17)} = 8,40$, je primerjava 1+2:3 značilna na stopnji $\alpha = o.ol$, ostanek pa je neznačilen. Predelana analiza variance je odkrila, da gre večina variabilnosti K_A na račun prednosti embaliranega blaga v primerjavi z razsutim blagom, Če izračunamo še meje zaupanja za

$$u = \frac{\bar{y}_1 + \bar{y}_2}{2} - \bar{y}_3$$ dobimo iz podatkov analize variance u = 16

$$u \pm t_{o.o5}^{(17)} \sqrt{s_e^2 \sum \frac{C_A^2}{n_A}} = 16 \pm 2,11 \cdot \sqrt{88,7 \left(\frac{1}{2^2.5} + \frac{1}{2^2.10} + \frac{1}{1^2.5} \right)} =$$

$= 16 \pm 11,33$

Če je število ponovitev za posamezne postopke različno

$n_A \neq \bar{n}$

je varianca za primerjavo $u = C_A\bar{y}_A$ enaka

$$\text{Var } u = \sigma_e^2 \sum \frac{C_A^2}{n_A}$$

primerjavi u_1 in u_2 pa sta v tem primeru ortogonalni, če je izpolnjen pogoj

$$\sum \frac{C_{A1}C_{A2}}{n_A} = 0$$

Za primer embalaže krompirja primerjava $u_2 = \bar{y}_1 - \bar{y}_2$ med različno emabalažo ni ortogonalna s primerjavo $u_1 = \bar{y}_1 + \bar{y}_2 - 2\bar{y}_3$, ker

$$\sum \frac{C_{A1}C_{A2}}{u_A} = \frac{(1).(1)}{5} + \frac{(-1).(1)}{10} + \frac{(0).(-2)}{5} = \frac{1}{10} \neq 0$$

$$K_{C2} = \frac{\sum(C_A\bar{y}_A)^2}{\sum \frac{C_{AA}^2}{u_A}} = \frac{(39 - 31)^2}{\frac{1}{5} + \frac{1}{10}} = 213$$

ker u_1 in u_2 nista ortogonalni primerjavi.

$K_{C2} + K_{C1} = 931 + 213 = 1144 \neq 1020$, kot bi dobili,če bi bili primerjavi ortogonalni.

$m_1 = \bar{y}_1 + \bar{y}_2 - 2\bar{y}_3$ ustrezna ortogonalna primerjava je $n_2 =$
$= 5\bar{y}_1 - 6\bar{y}_2 + \bar{y}_3$.

Če preskusimo, dobimo, da je

$$\sum \frac{C_{A1}C_{A2}}{n_A} = \frac{(1).(5)}{5} + \frac{(1).(-6)}{10} + \frac{(-2).(1)}{5} = 0$$

Prispevek te primerjave $\overset{k}{/K_A}$ pa je

$$K_{C1} = \frac{(5.39 - 6.31 + 1.19)^2}{\frac{5^2}{5} + \frac{6^2}{10} + \frac{1^2}{5}} = \frac{(-28)^2.10}{88} = 89$$

vsota $K_{C1} + K_{C2} = 931 + 89 = 1020$

4.20 O r t o g o n a l n i   p o l i n o m i .

Z ortogonalnimi primerjavami preskušamo vnaprej domnevane zveze in zakonitosti delovanja faktorjev. Če je proučevan faktor numeričen, morejo biti te domnevane zakonitosti ne samo primerjave med skupinami popreČij. Linearno povečavanje količine gnojila more imeti na kulturo tak vpliv, da donos narašča linearno ali krivuljčno. Čas učenja ima za rezultat linearno ali krivuljčno povečanje znanja.

Različna temperatura pri obdelavi določenega izdelka vpliva do določene temperature pozitivno in se kvaliteta izboljšuje, pri višjih temperaturah pa kakovost pada ipd.

Čeprav so krivulje, s katerimi opisujemo zvezo kriterialnega znaka y od numeričnega faktorja A najrazličnejše, najpogosteje opisujemo odvisnost y od A s polinomom določene stopnje

$$y = b_o + b_1A + b_2A^2 + b_3A^3 + \dots\dots\dots \qquad (4.44)$$

Če je zveza linearna, opišemo zvezo z linearno funkcijo

$$y = b_o + b_1A \qquad (4.45)$$

če je zveza kvadratična s parabolo druge stopnje

$$y = b_o' + b_1'A + b_2'A^2 \qquad (4.46)$$

Parabole višjih stopenj se poskusnim podatkom seveda bolje prilegajo kot parabole nižjih stopenj.

Za potrebe analize poskusnih podatkov pa so polinomi, dani v gornji obliki neprikladni. Primernejši so takoimenovani ortogonalni polinomi, ki so v ozki povezavi s pojmom primerjave. Polinom r-te stopnje moremo pisati v obliki

$$Y = B_oX_o + B_1X_1 + B_2X_2 + B_3X_3 \dots \qquad (4.47)$$

Pri tem so X_o X_1 X_2 X_3 X_K funkcije od A ustrezne stopnje

$$\begin{aligned} X_o &= a_{oo} \\ X_1 &= a_{1o} + a_{11}A \\ X_2 &= a_{2o} + a_{21}A + a_{22}A \end{aligned} \qquad (4.48)$$

z naslednjimi lastnostmi. Če imamo v poskusu numeričen faktor A vrednosti A_1 A_2 ... A_A, imajo polinomi X_K lastnost, da je vsota produktov

$$\begin{aligned} \sum_{A=1}^{A} X_K(A)K_j(A) &= 0 \quad \text{če je } K \neq j \qquad (4.49) \\ &= \sum_{A=1}^{A} X_K^2(A) \quad \text{če je } K = j \end{aligned}$$

Ti polinomi imajo torej lastnost ortogonalnih primerjav, zaradi česar jih imenujemo ortogonalni polinomi.

Za stvarno vrsto poprečij poskusnih podatkov $\bar{y}_A$ dobimo parametre polinoma izraženega z ortogonalnimi polinomi zlahka, če model

$$\bar{y}_A = \sum B_k X_k(A) + \bar{e}_A \tag{4.50}$$

pomnožimo z vrednostmi ortogonalnega polinoma $K_j(A)$ za vrednosti faktorja A dobimo

$$\sum_{A=1}^{A} \bar{y}_A X_A^j = \sum_{A=1}^{A} \sum_k B_k X_A^k X_A^j + \sum_A \bar{e}_A X_A^j \quad \text{in dalje zaradi} \tag{4.51}$$

4.49

$$\sum_A \bar{y}_A X_A^j = B_j \sum_A (X_k^j)^2 + \sum \bar{e}_A X_A^j \quad \text{in} \tag{4.52}$$

$$\frac{\sum \bar{y}_A X_A^j}{\sum (X_A^j)^2} = B_j + \frac{\sum \bar{e}_A X_A^j}{\sum (X_A^j)^2} \tag{4.53}$$

$\sum \bar{y}_A X_A^j$ oziroma $\hat{B}_j = \dfrac{\sum \bar{y}_A X_A^j}{\sum (X_A^j)^2}$ je ortogonalna primerjava,

katere vrednost pojasnjuje pomembnost komponente stopnje j v polinomu. Od vseh $\hat{B}_j$ edino $\hat{B}_o$ nima značaj primerjave, ker je $\sum X_A^o$ = konstanta in torej vsota

$$\sum_{A=1}^{A} X_A^o X^j \neq 0.$$ Ker pa je X_A^j primerjava, sledi,

$$\sum_{A=1}^{A} X_A^o X_A^j = X_A^o \sum_{A=1}^{A} X_A^j = 0 \qquad \text{če je } j \neq 0 \tag{4.54}$$

da imajo vrednosti vseh drugih ortogonalnih polinomov značaj primerjav. Prispevek komponente j k skupni vsoti K_A faktorja A je

$$K_{X^j} = \frac{\bar{n}(\sum \bar{y}_A X_A^j)^2}{\sum (X_A^j)^2} \qquad (4.55)$$

Nepristanska ocena parametra B_j je

$$\hat{B}_j = \frac{\sum \bar{y}_A X_A^j}{\sum (X_A^j)^2} \qquad (4.56)$$

Varianca za $\hat{B}_j$ je

$$\operatorname{Var} \hat{B}_j = \frac{e^2}{\bar{n}} \sum_{A=1}^{A} (X_A^j)^2 \qquad (4.57)$$

ocena variance pa

$$\operatorname{Var} \hat{B}_j = \frac{s_e^2}{\bar{n}} \sum_{A=1}^{A} (X_A^j)^2 \qquad (4.58)$$

Običajno so vrednosti faktorjev dane v enakih razmakih. Za take primere so ortogonalni polinomi do pete stopnje naslednji:

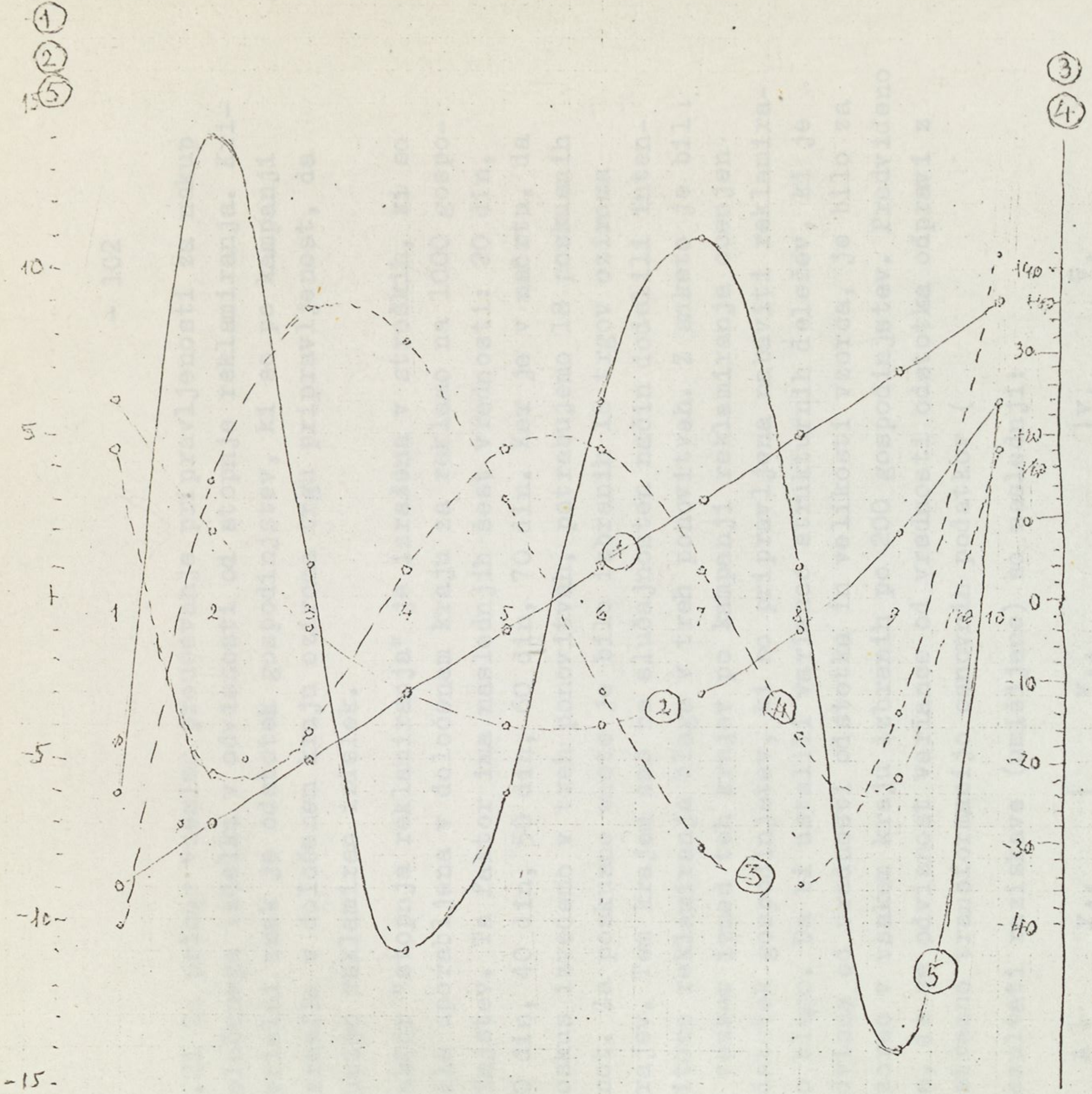
$$X^1 = \lambda_1 (x - \bar{x})$$
$$X^2 = \lambda_2 \left[(x - \bar{x})^2 - \frac{A^2 - 1}{12} \right]$$
$$X^3 = \lambda_3 \left[(x - \bar{x})^3 - \frac{1}{20} (3A^2 - 7)(x - \bar{x}) \right] \qquad (4.59)$$
$$X^4 = \lambda_4 \left[(x - \bar{x})^4 - \frac{1}{14} (3A^2 - 13)(x - \bar{x})^2 + \frac{3}{560} (A^2 - 1)(A^2 - 9) \right]$$

V tabeli so izračunane vrednosti polinomov X_A^j in vsote kvadratov $\sum (X_A^j)^2$

Tabela 4.1 Ortogonalni polinomi za faktorje z vrednostmi v enakih razmakih ($j=1...5$; $A=2$ ... 10)

število vrednosti faktorja	stopnja polinoma	vrednosti ortogonalnih polinomov za posamezne vrednosti faktorja X_A^j										$\sum(X_A^j)^2$
A	j	1	2	3	4	5	6	7	8	9	10	
2	1	1	-1									2
3	1	-1	0	1								2
	2	1	-2	1								6
4	1	-3	-1	1	3							20
	2	1	-1	-1	1							4
	3	-1	3	-3	1							20
5	1	-2	-1	0	1	2						10
	2	2	-1	-2	-1	2						14
	3	-1	2	0	-2	1						10
	4	1	-4	6	-4	1						70
6	1	-5	-3	-1	1	3	5					70
	2	5	-1	-4	-4	-1	5					84
	3	-5	7	4	-4	-7	5					180
	4	1	-3	2	2	-3	1					28
	5	-1	5	-10	10	-5	1					252
7	1	-3	-2	- 1	0	1	2	3				28
	2	5	0	- 3	-4	-3	0	5				84
	3	-1	1	1	0	-1	-1	0				6
	4	3	-7	1	6	1	-7	3				154
	5	-1	4	-5	0	+5	-4	1				84
8	1	-7	-5	-3	-1	1	3	5	7			168
	2	7	1	-3	-5	-5	-3	1	7			168
	3	-7	5	7	3	-3	-7	-5	7			264
	4	7	-13	-3	9	9	-3	-13	7			616
	5	-7	23	-17	-15	15	17	-23	7			2184

A	j	X_A^j 1	2	3	4	5	6	7	8	9	10	$\sum(X_A^j)^2$
9	1	-4	-3	-2	-1	0	1	2	3	4		60
	2	28	7	-8	-17	-20	-17	-8	7	28		2772
	3	-14	7	13	9	0	-9	-13	-7	14		990
	4	14	-21	-11	9	18	9	-11	-21	14		2002
	5	-4	11	-4	-9	0	9	4	-11	4		468
10	1	-9	-7	-5	-3	-1	1	3	5	7	9	330
	2	6	2	-1	-3	-4	-4	-3	-1	2	6	132
	3	-42	14	35	31	12	-12	-31	-35	-14	42	8580
	4	18	-22	-17	3	18	18	3	-17	-22	18	2860
	5	-6	14	1	-11	-6	6	11	1	-14	6	780



Slika 4.1. Ortogonalni polinomi stopenj 1-5 za A=1

4.21 Za primer vzemimo proučevanje pripravljenosti za nakup določenega izdelka v odvisnosti od stopnje reklamiranja. Kriterialni znak je odstotek gospodinjstev, ki so po kampanji izrazile v določenem kraju oziroma trgu pripravljenost, da kupijo reklamiran izdelek.

Faktor "stopnja reklamiranja" je izražena v stroških, ki so bila uporabljena v določenem kraju za reklamo na 1000 gospodinjstev. Ta faktor ima naslednjih šest vrednosti: 20 din, 30 din, 40 din, 50 din, 60 din, 70 din. Ker je v načrtu, da poskus izvedemo v treh ponovitvah, potrebujemo 18 poskusnih enot. Za poskusne enote je bilo izbranih 18 trgov oziroma krajev. Tem krajem smo na slučajnosten način dodelili intenzitete reklamiranja blaga v treh ponovitvah. Z anketo je bil v vsakem izmed teh krajev po kampanji reklamiranja ocenjen odstotek gospodinjstev, ki so pripravljena nabaviti reklamirano blago. Da bi ustalili varianco strukturnih deležev, ki je odvisna od vrednosti odstotka in velikosti vzorca, je bilo za vzorec v vsakem kraju izbranih po 200 gospodinjstev. Predvideno je, da se odvisnost variance od vrednosti odstotka odpravi z ustrezno transformacijo osnovnih podatkov (

Rezultati raziskave (umišljene) so naslednji:

A	y_{Ai}			v_{Ai}			v_A	$\bar{v}_A$
20	9	5	4	17,46	12,92	11,54	41.92	13.97
30	19	17	9	25,84	24,35	17,46	67,65	22,55
40	27	39	32	31,31	38,65	34,45	104,41	34,80
50	48	28	45	43,85	31,95	42,13	117,93	39,31
60	43	36	50	40,98	36,87	45,00	122,85	40,95
70	52	48	59	46,15	43,85	50,18	140,18	46,73
						v	594,94	

Če za transformirane podatke, za katere smatramo, da ustrezajo pogojem analize variance, obračunamo varianco, dobimo:

$$Q_{Ai} = \sum_{Ai} S\, v_{Ai}^2 = 17{,}46^2 + \ldots + 50{,}18^2 = 22184{,}3670$$

$$Q_A = \frac{1}{n} \sum v_A^2 = \frac{1}{3} (41{,}92^2 + \ldots + 140{,}18^2) = 21961{,}7656$$

$$Q = \frac{1}{n} v^2 = \frac{1}{63}\, 594{.}94^2 = 19664{,}0890$$

dalje je

$$q_{Ai} = Q_{Ai} - Q = 2520{,}2779$$

$$q_A = Q_A - Q = 2297{,}6765$$

a analiza variance

VV	K	m	s^2	F
strošek oglašanja	$K_A = q_A = 2297{,}6765$	5	459,54	$24{,}77^{xxx}$
pp	$K_e = q_{Ai} - q_A = 222{,}6014$	12	18,55	1
	$K = q_{Ai} = 2520{,}2779$			

Splošna analiza variance pokaže, da je intenziteta reklamiranja zelo značilen faktor ($F_{0.01}(5,12) = 8{,}89$). Vendar ta splošen rezultat nič ne pove o zakonitostih teh zvez. Ker je faktor "intenziteta oglašanja" numeričen faktor, za katerega so vrednosti v enakih razmakih, moremo skupen učinek analizirati z ortogonalnimi polinomi po komponentah.

Za A = 6 so koeficienti za linearno in kvadratično komponento z ortogonalnimi polinomi enaki

A	20	30	40	50	60	70	$(X_6)^2$	$\sum X_6\bar{y}_A$	$\frac{n(\sum X_6\bar{y}_A)^2}{(X_6)^2}$
X_6^1	-5	-3	-1	1	3	5	70	223,47	240,30
X_6^2	5	-1	-4	-4	-1 ...	5	84	-36,45	113,82
$\bar{y}_A$	13,97	22,55	34,80	39,31	40,95	46,73			

Če se zadovljimo s pojasnitvijo linearne in kvadratične komponente, dobimo naslednjo revidirano analizo variance

WV	K	m	s^2	F
strošek oglašanja	$K_A = 2297,6765$	5		
linearna komp	$K_{X1} = 2140,3000$	1	2140,30	$115,38^{xxx}$
kvadratična komp	$K_{X2} = 113,8207$	1	113,82	$6,14^{x}$
drugo (višje stop.)	$K_{AX1X2} = 43,5558$	3	14,52	-
pp e	$K_e = 222,6014$	12	18,55	1
Sk	$K = 2520,2779$	17		

Ker je $F_{0.05}(1,12) = 4.75$, $F_{0.01}(1,12) = 9,33$ in $F_{0.001}(1,12) = 18,6$ moremo sklepati, da je linearna komponenta za transformirane podatke visoko značilna ($\alpha = 0.001$), kvadratična komponenta pa značilna na nivoju ($\alpha = 0.05$). Komponente višjih stopenj (tretje, četrte in pete) so se v skupnem izkazale kot neznačilne.

Opomba: če pride s_A^2 ali za katerokoli njegovo komponento $s_A^2 < s_A^2$ včasih ta del kot čisto slučajnosten vključimo v poskusni pogrešek. Tako se združena ocena za pp zmanjša, obenem pa pridobimo na stopinjah prostosti. V našem primeru je revidirani

$$s_e^2 = \frac{43,5558 + 222,6014}{3 + 12} = \text{ z } m_e = 15 \text{ stopinjami prostosti}$$

in pridejo izračunani F_{X1} in F_{X2} nekoliko večji in število stopinj prostosti $m_e = 15$. Vendar razlika ni tolika, da bi bistveno vplivala na odnose značilnosti.

Če izračunamo ocene parametrov $\hat{B}_1$ in $\hat{B}_2$ dobimo, da je

$$\hat{B}_1 = \frac{\sum X_6^1 \bar{y}_A}{\sum (X_6^1)^2} = \frac{223,47}{70} = 3,192$$

$$\hat{B}_2 = \frac{\sum X_6^2 \bar{y}_A}{\sum (X_6^2)^2} = \frac{-36,45}{84} = -0.672$$

Iz teh parametrov spoznamo, da se transformirani strukturni delež $\bar{n}_A$ z večanjem intenzitete oglašanja na vsakih 10 din dvigne za 3,192 . 2 transformiranega deleža, da pa je naraščanje degresivno, ker je koeficient kvadratične komponente negativen.

Ker somkomponente tretje do pete stopnje neznačilne moremo smatrati

$$\hat{v}_A = \hat{B}_o X_o + \hat{B}_1 X_1 + \hat{B}_2 X_2$$

kot točkovno oceno poteka krivulje poprečnega transformiranega efekta oglaševanja.

Ocena $\hat{B}_o = \frac{\sum \bar{y}_A}{A}$ je v našem primeru

$$\hat{B}_o = \frac{223,47}{6} = 33,052$$

je

A	$\hat{B}_o X_o$	+ $\hat{B}_1 X_1$	+ $\hat{B}_2 X_2$	= $\bar{v}_A$	$\bar{y}_A\%$
20	33,052.(1)	+ 3,192(-5)	- ·672(5)	= 13,73	5.6
30	(1) +	.(-3)	.(-1)	= 24,15	16.7
40	(1) +	.(-1)	.(-4)	= 32.55	29.0
50	(1) +	.(+1)	.(-4)	= 38,93	39.5
60	(1) +	.(+3)	.(-1)	= 43,30	47,0
70	(1) +	.(+5)	.(5)	= 45.65	51.1

Iz stolpca za $\bar{y}_A\%$ vidimo, da ne samo $\bar{v}_A$ temveč tudi $\bar{y}_A\%$ degresivno narašča, več vložena sredstva imajo manjši učinek na delež pozitivnih, kar je v skladu s pričakovanjem.

Po obrazcu ( ) je ocena variance parametrov funkcije odvisnosti deleža od intenzitete oglašanja, naslednja

$$\mathrm{var}\ \hat{B}_1 = \frac{s_e^2}{\bar{n}} \sum (X_A^1)^2 = \frac{18,55}{3} . 70 = 432,833;$$

$$s_e\ \hat{B}_1 = \sqrt{\mathrm{var}\ \hat{B}} = 20,81$$

$$\mathrm{var}\ \hat{B}_2 = \frac{s_e^2}{\bar{n}} \sum (X_A^2)^2 = \frac{18.55}{3} = 519,40 ;$$

$$s_e \hat{B}_2 = \sqrt{\mathrm{var}\ \hat{B}_2} = 22,79$$

4.22 Čisto slučajnosten poskus za model II.

Če je faktor, ki ga proučujemo slučajnosten, pri čistem slučajnostnem poskusu ni neke posebne razlike s primerom fiksnega faktorja. Zaradi tega, da osvojimo simboliko, ki bo v kompleksnejših primerih prišla bolj do izraza, nakažimo analizo variance za primere $n_a = \bar{n}$ in $n_a \neq \bar{n}$.

Slučajnostne faktorje zaznamujemo z malimi črkami. Osnovna tabela podatkov in vse ustrezne količine izračunamo na enak način, le da A zamenjamo z a.

Osnovna tabela podatkov

$$\begin{matrix} y_{ai} & y_a & y \\ n & n_a & 1 \end{matrix} \qquad (4.60)$$

Model

$$y_{ai} = M + (a) + e_{ai} \qquad e_{ai} = : N(0, \sigma_e^2)$$

$$E_{(a)} = 0 \qquad Var_a = \sigma_a^2 \qquad (4.61)$$

Pomožne količine;

$$Q_{ai} = \underset{ai}{SS}\, y_{ai}^2 \qquad Q_a = \frac{1}{\bar{n}} \underset{a}{S}\, y_a^2/n_a \qquad Q = \frac{y^2}{n} \qquad (4.62)$$

$$q_{ai} = Q_{ai} - Q \qquad q_a = Q_a - Q$$

Analiza variance

VV	K	m	s^2	F	Es^2
a	$K_a = q_a$	a - 1	s_a^2	$F_a = \frac{s_a^2}{s_e^2}$	$\sigma_e^2 + \bar{n}\sigma_a^2$
pp	$K_e = q_{ai} - q_a$	n-a	s_e^2	1	σ_e^2
Sk	$K = q_{ai}$	n-1			

(4.63)

Vidimo, da σ_A^2 zamenja varianca od a σ_a^2

Če je število ponovitev postopkov enako je $Q_a = \frac{1}{\bar{n}} \Sigma y_a^2$ pri čemer je $\bar{n} = \frac{n}{a}$

Pri matematičnem upanju $E(s^2)$ v analizi variance je pri enakem številu ponovitev po postopkih $\bar{n} = \frac{n}{a}$ pri neenakem številu pa je $E\, s_a^2$ enak

$$\bar{n} = \frac{n^2 - Sn_a^2}{n(a-1)} \qquad (4.64)$$

Pri enakem številu ponovitev je

$$\bar{n} = \frac{n^2 - \underset{a}{S}\frac{n^2}{a^2}}{n(a-1)} = \frac{n}{a-1}\left[1 - \frac{a}{a^2}\right] = \frac{n}{a}$$

4.23 Zaradi primerjave uspešnosti različnih načrtov sistematičnega poskusa, čisto slučajnostnega načrta, metode popolnih blokov in latinskega kvadrata in analize kovariance, bomo za posamezne načrte na istem problemu aplicirali ob ustreznem poglavju na primeru, za katerega dobimo podatke z ustreznim simuliranjem poskusa.

S poskusom proučujemo učinek faktorja A, ki ima tri vrednosti; $A_1 = 40$; $A_2 = + 10$; $A_3 = - 50$

Poskus bo v vseh primerih izveden v $\bar{n} = 3$ ponovitvah. Poskusno gradivo sestoji torej iz $n = \bar{n} = 3.3 = 9$ enot. Poskusno gradivo je urejeno tako, da se od enote do enote pojav veča za 10 enot. Splošno poprečje je M = 100. Razen tega je predvidena za poskuse še slučajnostna komponenta, ki je take narave, da je $S_e = 0$. Če sestavimo iz teh elementov podatke za pojav, če faktor A ne apliciramo, dobimo naslednjo vrsto podatkov, ki bi jih dobili, če bi vršili slep poskus, tj. poskus brez delovanja faktorja A.

Poskusna enota	1	2	3	4	5	6	7	8	9
M	100	100	100	100	100	100	100	100	100
X	- 40	- 30	- 20	- 10	0	+ 10	+ 20	+ 30	+ 40
e	1	- 2	3	- 4	2	5	- 4	3	- 4
y_i=M+x+e	61	68	83	86	102	115	116	133	136

Problem poskusa na gradivu, ki je podobno zgornjemu v praksi pogosto srečamo. Obilo pojavov v časovnem proučevanju kaže konstanten trend(npr. v prodaji, številu turistov, spremembi temperature ipd.) ali pa je v zveznem enodimenzionalnem gradivu kot je npr. žica, parcele vzdolž določene linije, opaziti linearen trend v proučevani značilnosti.

Zato so nakazane metode v teh primerih zelo uspešno sredstvo proučevanja določenega faktorja ob izločanju vpliva trenda.

Vzemimo po vrsti posamezne metode poskusa.

a) Sistematično apliciranje faktorja A.

Pri sistematičnem apliciranju faktorja A po vrsti apliciramo prvim trem poskusnim enotam postopek A_1 z učinkom $A_1 = 40$, naslednjim trem postopek A_2 z učinkom $A_2 = 10$ in zadnjim trem postopek A_3 z učinkom $A_3 = -50$.

Rezultati poskusa z vsoto y_i + (A) so

i	1	2	3	4	5	6	7	8	9
	A_1	A_1	A_1	A_2	A_2	A_2	A_3	A_3	A_3
(A)	40	40	40	10	10	10	-50	-50	-50
M + x +e	61	68	83	86	102	115	116	133	136
y_{Ai}	101	108	123	96	112	125	66	83	86

$$\bar{y}_1 = \frac{101 + 108 + 123}{3} = 110,7 \qquad \hat{A}_1 = 10,7$$

$$\bar{y}_2 = \frac{96 + 112 + 125}{3} = 111 \qquad \hat{A}_2 = 11,0$$

$$\bar{y}_3 = \frac{66 + 83 + 86}{3} = 78,3 \qquad \hat{A}_3 = -21.7$$

$$\bar{y} = \quad = 100 \qquad \sum \hat{A} = 0$$

Ker je poskus sistematičen, ne moremo na poskusnih podatkih izvesti analize variance. Pač pa primerjava dobljenih rezultatov za oceno (A) kažejo, da se izračunane vrednosti bistveno odklanjajo od A(40, 10, -50).

b) Čisto slučajnostni poskus

Simulirajmo čisto slučajnostni poskus in dodelimo devetim poskusnim enotam postopke A_1 A_2 A_3 slučajnostne v treh ponovitvah.

i	1	2	3	4	5	6	7	8	9
	A_1	A_1	A_3	A_3	A_1	A_2	A_3	A_2	A_2
(A)	40	40	-50	-50	40	10	-50	10	10
M+x+e	61	68	83	86	102	115	116	133	136
y_{Ai}	101	108	33	36	142	125	66	143	146

$y_1 = 101 + 108 + 142 = 351$ $\quad$ $y_2 = 125 + 143 = 146 = 414$

$y_3 = 33 + 36 + 66 = 135$

$\bar{y}_1 = 351/3 = 117$ $\quad$ $\bar{y}_2 = 414/3 = 138$ $\quad$ $\bar{y}_3 = 135/3 = 45$

$\hat{A}_1 = + 17$ $\quad$ $\hat{A}_2 = + 38$ $\quad$ $\hat{A}_3 = - 55$

Če izvedemo na zgornjih podatkih analizo variance, dobimo;

$$Q_{Ai} = 101^2 + 108^2 + \dots + 146^2 = 106160$$

$$Q_A = \frac{1}{3}(351^2 + 414^2 + 135^2) = 104274$$

$$Q = \frac{1}{9}\, 900^2 = 90000$$

$$q_{Ai} = 106160 - 90000 = 16160 \qquad q_A = 104274 - 90000 = 14274$$

Analiza variance:

VV	K	m	s^2	F
A	$q_A = 14274$	2	7137	22.74
pp	$q_{Ai} - q_A = 1886$	6	314	1
Sk	$q_{Ai} = 106160$	8		

Analiza variance je z F-preskusom odkrila visoko značilne razlike med postopki, ocene postopkov $\hat{A}_A$ pa so bližji pravim vrednostim A_A kot pa pri sistematičnem načrtu.

5. <u>SLUČAJNOSTNI POSKUS V POPOLNIH BLOKIH</u>

5.1 Pri načrtovanju čisto slučajnostnega poskusa se srečamo s problemom heterogenosti poskusnega gradiva. Pogosto je obseg poskusa, ki je izveden v več ponovitvah, tolikšen, da je nemogoče zagotoviti homogenost poskusnega gradiva. Če je v poskusu z $A = 10$ postopki in $\bar{n} = 5$ ponovitvami, potrebujemo skupno $n = A \cdot \bar{n} = 10 \cdot 5$ poskusnih enot. Če gre za proizvodni poskus, je težko sestaviti 50 homogenih primerov surovine ali homogeno skupino 50 delavcev. Podobno je s poskusnimi parcelicami v agronomskih raziskavah, ali s poskusnimi živalicami pri bioloških raziskavah.

Velik del heterogenosti poskusnega gradiva pa izločimo z metodo blokov. Čisto slučajnostni razpored posameznih poskusov se umakne do določene mere načrtovane razmestitve. Enote poskusnega gradiva namreč združimo v homogene skupine s toliko enotami, kolikšno je število postopkov v poskusu. Tako sestavimo iz heterogenega poskusnega gradiva toliko homogenejših skupin, kolikor je ponovitev poskusa. Na enotah ene take homogene skupine - <u>bloka</u> izvedemo en popoln slučajnostni poskus v eni ponovitvi. Ker v okviru enega bloka izvedemo poskus z vsemi postopki, imenujemo tak blok <u>popoln blok.</u> V okviru bloka pa je poskus še vedno slučajnosten, ker na že znan način posameznim enotam bloka dodeljujemo postopke slučajnostno. Popolni bloki so ena izmed najpogostejših tehnik, ki jih uporabljamo pri statističnih načrtih poskusa. Blok more sestavljati pri poskusu v agronomiji skupina poskusnih parcelic na isti njivi, ker so pogoji na isti njivi bolj izenačeni kot pa na večih njivah, ki predstavljajo celotno poskusno gradivo. Pri poskusih z živalmi predstavljajo blok živali iz istega gnezda. Pri analizi trga more biti vodilo za združevanje trgovin v bloke kraj. Take skupine trgovin imajo skupnih niz značilnosti trgovanja, ki so pogojene s krajem.

Pri proizvodnih poskusih npr. z žico, more blok predstavljati 5 x 50 cm žice iz istega kolobarja 2,50 m dolg kos žice iz istega kolobarja. Ta kos razrežemo v pet kosov po 50 cm, ki predstavljajo poskusne enote bloka.

Ko primerjamo rezultate poskusa, razlike med postopki niso pod vplivom razlik med bloki. Tako pri agronomskih poskusih poskusi niso več odvisni od razlik med njivami, pri poskusu z živalmi nič od genetskih razlik živali različnih gnezdov. Pri analizi trga z zgornjim načinom izločimo iz poskusa faktorje, ki so pogojeni s krajem, pri proizvodnem postopku z žico pa vse razlike, ki izvirajo iz razlik v kakovosti žice med kolobarji ali mesta v kolobarju (začetek, sredina, konec).

5.2 Predpostavke in model slučajnostnih popolnih blokov.

Pri načrtih, v katerih sestavljamobloke, predpostavljamo, da so proučevani faktorji in faktorji, ki so vodilo za sestavljanje blokov med seboj neodvisni, oziroma njihovi učinki v kolikor pa so odvisni, vzamemo, da je ta odvisnost slučajna.

Načrt slučajnostnega poskusa s tremi postopki A_1 A_2 A_3 v 4 popolnih blokih (SPPB) moremo prikazati shematično:

B_1			B_2			B_3			B_4		
A_1	A_2	A_3	A_3	A_2	A_1	A_3	A_2	A_1	A_2	A_1	A_3

Če z y_{BA} zaznamujemo podatek za kriterialni znak za postopek A v bloku B, moremo pri pogoju o aditivnosti oziroma neodvisnosti pisati model SPB.

$$y_{BA} = M + (B) + (A) + e_{BA} . \qquad (5.1)$$

Ničelna domneva H_o : (A) = 0 je, da so vse vrednosti A enake nič, medtem ko je osnovna domneva H_1 (A) $\neq$ 0 da niso vse vrednosti komponente enako 0. Dalje veljajo predpostavke

$$\sum(B) = \sum(A) = 0 \qquad e_{BA} = : N(0, \sigma_e^2) \qquad (5.2)$$

da je vsota vseh vrednosti komponente A in vsota vseh vrednosti učinka bloka enaka 0, e_{BA} slučajnostna komponenta pa se porazdeljuje normalno s stalno varianco σ_e^2.

Kot je razvidno iz oznak (velike črke B in A) in iz predpostavk, nakazani model velja za primer, da imajo bloki in postopki značaj stalnega faktorja.

Potrebne tabele vsot, izračunane iz tabele osnovnih podatkov y_{BA} so naslednje:

$$\begin{matrix} y_{BA} & y_B \\ y_A & y \end{matrix} \qquad (5.3)$$

Pomožne količine, izračunane iz teh tabel, so:

$$Q_{BA} = \sum_{BA} y_{BA}^2 \qquad Q_A = \frac{A}{n} \sum y_A^2 \qquad Q_B = \frac{B}{n} \sum y_B^2 \qquad Q = \frac{1}{n} y^2$$

$$q_{BA} = Q_{BA} - Q \qquad q_A = Q_A - Q \qquad q_B = Q_B - Q \qquad (5.4)$$

Analizo variance pa obračunamo po naslednjem postopku:

Vir variacije	vsota kvadratov odklonov K	m	s^2	F	Es^2
Bloki (B)	$K_B = q_B$	$m_B = B-1$	s_B^2	-	$\sigma_e^2 + A\sigma_B^2$
Postopki (A)	$K_A = q_A$	$m_A = A-1$	s_A^2	$F_A = \frac{s_A^2}{s_e^2}$	$\sigma_e^2 + B\sigma_A^2$
pp=e	$K_e = q_{BA} - q_B - q_A$	$m_e (B-1)(A-1)$	s_e^2	1	σ_e^2
Sk	$K = q_{BA}$	$m = BA - 1$			

Iz analize variance je razviden obračun in F-preskus. Ker bloki niso formirani slučajnostno, nismo upravičeni z F-preskusom preskušati značilnosti razlik med bloki. Ne glede na to pa tak preskus tudi vsebinsko ni smiseln, ker bloki niso vsebinske ampak tehnične tvorbe, ki jih sestavljamo zato, da izločimo iz analize vpliv faktorjev, ki so z njimi povezani.

5.3 Vzemimo za primer slučajnostnih blokov naslednji problem. Proizvajalec mlečnih izdelkov se zanima ali in kako je število bakterij odvisno od štirih načinov vskladiščenja A. Pokazovalec je indeks števila bakterij po 72 urah vskladiščenja. Ker je število bakterij v veliki meri odvisno od pošiljke mleka, smo izbrali 5 pošiljk mleka in od vsake pošiljke en del vskladiščili po vsakem izmed štirih postopkov vskladiščenja. V tem primeru predstavlja pošiljka mleka blok, ki ima tipičen značaj faktorja modela II, ker je pet pošiljk izbranih slučajnostno iz umišljene populacije pošiljk.

Model za ta primer je

$$y_{pV} = M + (p) + (V) + e_{pV} \qquad (5.7)$$

Če zaradi asociacije zaznamujemo komponente z ustreznimi črkami (p = pošiljka, V = vskladiščenje

Osnovni rezultati raziskave so naslednji:

Pošiljka / y_{pV}	Vskladiščenje V_1	V_2	V_3	V_4	y_p
p_1	2	2	5	3	12
p_2	7	9	12	8	36
p_3	5	11	11	9	36
p_4	4	4	6	6	20
p_5	2	4	6	4	16
y_V	20	30	40	30	120 = y

$$Q_{pV} = S \sum_{p\,V} y_{pV}^2 = 2^2 + 2^2 + \dots + 6^2 + 4^2 = 904$$

$$Q_p = \frac{p}{n} Sy_p^2 = \frac{5}{20}\, 12^2 + 36^2 + 36^2 + 20^2 + 16^2 = 848$$

$$Q_V = \frac{V}{n} S y_V^2 = \frac{4}{20}\, 20^2 + 30^2 + 40^2 + 30^2 = 760$$

$$Q = \frac{1}{n} y^2 = \frac{120^2}{20} = 720$$

$q_{pV} = 904 - 720 = 184$

$q_p = 848 - 720 = 128$

$q_V = 760 - 720 = 40$

Ustrezna analiza variance je:

VV	K	m	s^2	F	$E(s^2)$
(p)	$K_p = q_p = 128$	$m_p = 5-1 = 4$	$s_p^2 = 32{,}00$	24	$\sigma_e^2 + v\sigma_{p}^2$
(v)	$K_v = q_v = 40$	$m_v = 4-1=3$	$s_v^2 = 13{,}33$	10	$\sigma_e^2 + p\sigma_v^2$
pp	q_{pv} 184 $- q_p$ −128 $- q_v$ − 40 $= Ke$ 16	$m_e = (5-1)(4-1) = 12$	$s_e^2 = 1{,}33$	1	σ_e^2
Sk	$K = q_{pv} = 184$	$m = 5.4-1 = 19$			

Če vzamemo vnaprej izdelane primerjave in pri tem upoštevamo, da se V_1 in V_2 razlikujeta od V_3 in V_4 v nekih posebnih elementih, moremo to izraziti s primerjavami

$\bar{y}_V$	$\bar{y}_1$	$\bar{y}_2$	$\bar{y}_3$	$\bar{y}_4$	$C_V \bar{y}_V$	C_V^2	$\frac{p(\sum C_V \bar{y}_V)^2}{C_V^2}$	F
	4	6	8	6				
1+2 : 3+4	1	1	- 1	-1	- 4	4	20	15 =20/1,33
1 : 2	1	-1	0	0	- 2	2	10	7.5=10/1.33
3 : 4	0	0	1	-1	2	2	10	7.5=10/1.33
							40	

Ker je $F_{0.05}(1,12) = 4.75$ $F_{0.01}(1,12) = 9.33$ sklepamo, da je poprečni učinek vskladiščenja po načinu 1 in 2 značilno različen od poprečnega učinka vskladiščenja po načinu 3 in 4 na stopnji $\alpha = 0.01$, 1 od 2 in 2 od 3 pa na stopnji $\alpha = 0.05$. Iz primera je razvidno, da je vsota prispevkov individualnih stopenj prostosti enaka skupni vsoti $K_V = 20 + 10 + 10 = 40$ kar je posledica tega, da so vse tri primerjave med seboj ortogonalne.

Če upoštevamo $E(s^2)$, moremo oceniti posamezne komponente variabilnosti.

Tako je ocena za σ_p^2 enaka

$$\hat{\sigma}_p^2 = \frac{s_p^2 - s_e^2}{V} = \frac{32 - 1,33}{4} = 7,67$$

$$\text{ocena za } \hat{\sigma}_V^2 = \frac{1}{V-1} \sum (V)^2 = \frac{s_V^2 \cdot V - s^2}{5} = \frac{13,33 - 1,33}{5} = 2,4$$

5.4 Obračun in analiza variance se v primeru, da so bloki ali bloki in faktor, ki ga proučujemo, slučajnostni, bistveno ne spremeni. Oznake za posamezne komponente pišemo ustrezno z velikimi črkami za fiksne in z malimi črkami za slučajnostne. Bistveno drugačno pa je seveda tolmačenje dobljenih rezultatov in pomen posameznih oznak.

Za primer, da so bloki fiksni, so vrednosti bloka vezane na pogoj

$$\sum B = 0 \qquad \sigma^2_B \text{ pa je } \sigma^2_B = \frac{1}{B-1}\sum (B)^2 \tag{5.7}$$

Če pa so bloki slučajnosten faktor in predpostavljamo, da je b blokov vzorec iz populacije neomejenega števila blokov, za katere predpostavljamo, da so vrednosti bloka slučajnostna spremenljivka $b = : N(0, \sigma^2_b)$

enako velja za faktor A

A fiksen $\quad \sum A = 0 \qquad \sigma^2_A = \frac{1}{A-1}\sum A^2$ (5.8)

a slučajnosten $\quad a = N:(0, \sigma^2_a)$ (5.9)

Ustrezna matematična upanja za posamezne kombinacije pa so:

VV	Model I		Mešan model				Model II	
	B	A	b	A	B	a	b	a
blok	$\sigma^2_e + A\sigma^2_B$		$\sigma^2_e + A\sigma^2_b$		$\sigma^2_e + a\sigma^2_B$		$\sigma^2_e + a\sigma^2_b$	
postopek	$\sigma^2_e + B\sigma^2_A$		$\sigma^2_e + b\sigma^2_A$		$\sigma^2_e + B\sigma^2_a$		$\sigma^2_e + b\sigma^2_a$	
pogrešek	σ^2_e		σ^2_e		σ^2_e		σ^2_e	

(5.10)

5.5 Če ne predpostavljamo, da so učinki blokov in proučevanega faktorja vezani aditivno oziroma, da so neodvisni $^2_{ab} \neq 0$, je model slučajnostnega poskusa v popolnih blokih z i meritvami na poskusni enoti

$$Y_{bai} = M + (b) + (a) + (ba) + e_{bai} \qquad (5.11)$$

analiza variance pa

VV	K	m	s^2	$E\,s^2$ (5.12)
b	q_b	$b-1$	s_b^2	$\sigma_e^2 + a\bar{n}\sigma_{ba}^2 + a\bar{n}\sigma_b^2$
a	q_a	$a-1$	s_a^2	$\sigma_e^2 + b\bar{n}\sigma_{ba}^2 + b\bar{n}\sigma_a^2$
pp	$q_{ba}-q_b-q_a$	$(b-1)(a-1)$	s_{ba}^2	$\sigma_e^2 + \bar{n}\sigma_{ba}^2$
vp	$q_{bai} - q_{ba}$	$(i-1)ba$	s_e^2	σ_e^2

Za posamezne kombinacije vrst faktorjev moremo dobiti model I.
(B,A blok in faktor fiksen faktor)
mešana modela (b, A in B, a) od katerih v praksi pogosto nas
prvi
model II. (b, a blok in faktor slučajnostna)
$E(s^2)$ za posamezne modele so naslednji: (5.13)

Model I. BA			Mešan model b, A		Model II. b, a	
VV		$E(s^2)$	VV	$E(s^2)$	VV	$E(s^2)$
B	s_B^2	$\sigma_e^2 + A\bar{n}\sigma_B^2$	b	$\sigma_e^2 + \quad + A\bar{n}\sigma_b^2$	b	$\sigma_e^2+\bar{n}\sigma_{ba}^2+a\bar{n}\sigma_b^2$
A	s_A^2	$\sigma_e^2 + B\bar{n}\sigma_A^2$	A	$\sigma_e^2 + \bar{n}\sigma_{bA}^2 + b\bar{n}\sigma_A^2$	a	$\sigma_e^2 + \bar{n}\sigma_{ba}^2 + b\bar{n}\sigma_a^2$
pp	s_p^2	$\sigma_e^2 + \bar{n}\sigma_{AB}^2$	pp	$\sigma_e^2 + \bar{n}\sigma_{bA}^2$	pp	$\sigma_e^2 + \bar{n}\sigma_{ba}^2$
vp	s_v^2	σ_e^2	vp	σ_e^2	vp	σ_e^2

V obrazcih za $E(s^2)$ nakazane variance imajo naslednji pomen

$$ {}^{2}_{A} = \frac{1}{A-1} \; {}_{A}(A)^2 \qquad {}^{2}_{AB} = \frac{1}{(A-1)(B-1)} (AB)^2 $$

$$ {}^{2}_{bA} = \frac{1}{A-1} \; {}_{A=1}^{A} \; {}^{2}_{bA} \quad \text{ipd.} $$

5.6 V primeru, da na posamezni poskusni enoti izvedemo eno samo meritev ($\bar{n} = 1$) je shema analize variance naslednja:

VV	K	m	s^2	Es^2 (5.14)
b	q_b	$b-1$	s_b^2	$\sigma_e^2 + \dot{a}\sigma_{ba}^2 + a\sigma_b^2$
a	q_a	$a-1$	s_a^2	$\sigma_e^2 + \dot{b}\sigma_{ba}^2 + b\sigma_a^2$
pp	$q_{ba}-q_b-q_a$	$(b-1)(a-1)$	s_{ba}^2	$\sigma_e^2 + \sigma_{ba}^2$
Sk	q_{ba}	$ab-1$		

Ker je $\bar{n} = 1$, odpade zadnja vrstica za vzorčni pogrešek. Shemo moremo s kombiniranjem, da male črke zamenjamo z velikimi, preurediti za Model I ali za mešan model.

Pri tem je $\dot{a} = 1 - \frac{a}{A}$; $\dot{b} = 1 - \frac{b}{B}$; če je faktor ali blok fiksen spoznamo, da sta količini

$$ \dot{a} = 1 - \frac{A}{A} = 0 \quad \text{ali} \quad \dot{b} = 1 - \frac{B}{B} = 0 $$

Če pa je faktor slučajnosten ($A = \infty$) ali blok slučajnosten ($B = \infty$) je

$$ \dot{a} = 1 - \frac{a}{\infty} = 1 \qquad \dot{b} = 1 - \frac{b}{\infty} = 1 $$

Izrazi kot so $\dot{a}$ in $\dot{b}$ z naznačenim pomenom se javljajo v različnih načrtih. Če je faktor take narave, da iz končnega števil možnih vrednosti A izberemo v poskus a slučajnostno izbranih vrednosti faktorja je seveda na splošno

$$\dot{a} = 1 - \frac{a}{A} \;.$$

Če upoštevamo pomen nakazanih simbolov, dobimo naslednje variante za proučevani primer

(5.15)

Model I.		Mešani modeli		Model II.
BA	Es^2	B,a Es^2	b,A $E\,s^2$	ba Es^2
B	$\sigma_e^2 + A\sigma_B^2$	B $\sigma_e^2 + \sigma_{Ba}^2 + a\sigma_B^2$	b σ_e^2 +	$A_b^2\sigma^b\ \sigma_e^2 + {}^2_{ba} + a\sigma_b^2$
A	$\sigma_e^2 + B\sigma_A^2$	a σ_e^2 + + $B\sigma_a^2$	A σ_e^2 + $_{bA} + b\sigma_A^2$	a $\sigma_e^2 + \sigma_{ba}^2 + b\sigma_a^2$
pp	$\sigma_e^2 + \sigma_{BA}^2$	pp $\sigma_e^2 + \sigma_{Ba}^2$	pp $\sigma_e^2 + \sigma_{bA}^2$	pp $\sigma_e^2 + \sigma_{ba}^2$

Iz $E(s^2)$ za različne kombinacije spoznamo, da edino mešan model (b, A) in Model II (b, a) omogočata korekten F-preskus za varianco faktorja, medtem, ko za model I. (B,A) in mešan model (B,a) to ni mogoče in dobimo v splošnem F premajhen, ker je σ_{BA}^2 oziroma σ_{Ba}^2 navzoč le v imenovalcu, ne pa tudi v števcu. Ker v praktičnih primerih pogosto predpostavljamo, da so bloki b slučajnostni, moremo za te primere obračunavati analizo variance neglede na to ali je $\sigma_{ba}^2 = 0$ ali ne.

Podobna analiza variance in sklepi slede tudi v primeru, da je $\bar{n} \cdot 1$ le da moremo v tem primeru še z $\dfrac{s_e^2}{s_v^2}$ preskušati značilnost komponente σ_{ba}^2 .

Problematika analiz odnosov posameznih postopkov je pri pop- nih blokih enaka kot pri čisto slučajnostnem poskusu. Tudi v tem primeru veljajo apriorne in aposteriorna analiza, ortogonalne primerjave, ortogonalni polinomi za numerične faktorje in...

5.7 Relativna uspešnost slučajnostnih popolnih blokov.

Uspešnost določenega poskusnega načrta se pokaže na zmanjšanju poskusnega pogreška pri stalnem skupnem obsegu poskusa n. Tako kot merilo uspešnosti služi razmerje

$$\frac{s_e^2 , ČSP}{s_e^2 . SPB} \qquad (5.16)$$

med poskusnim pogreškom pri čisto slučajnostnem poskusu in poskusnim pogreškom pri slučajnostnih popolnih blokih. Kolik bi pri istem poskusu bila s_e^2.ČSP, dobimo z združenjem s_e^2 in seveda tehtano z ustreznim številom stopinj prostosti

$$s_e^2 . ČSP = \frac{m_b s_b^2 + (m_A + m_e) s_e^2}{m} \qquad (5.17)$$

Fisherjeva korektura razmerja iz 5.16 je

$$\frac{(m_b+1)}{(m_b+3)} \cdot \frac{(m_b+m_e+3)}{(m_b+m_e+1)} \cdot \frac{s_e^2 . ČSP}{s_e^2} \qquad (5.18)$$

V našem primeru je uspešnost naslednja:

$$s^2_{e.ČSP} = \frac{4,32 + (3+12)1,33}{19} = \frac{148}{19} = 7,79$$

$$E_{SB:ČSP} = \frac{(4+1)(16+3)}{(4+3)\cdot(16+1)} \cdot \frac{7.79}{1.33} = 4,66$$

Izračun pokaže, da je uspešnost slučajnostnih popolnih blokov E = 4,66 kar pomeni, da bi morali vzeti 4,66 krat večji čisto slučajnostni poskus, da bi dosegli tako natančnost, kot smo dosegli s slučajnostnimi popolnimi bloki.

5.8 Preskusimo metodo slučajnostnih blokov še na simuliranem poskusu, Ker predvidevamo, da je v gradivu sistematičen trend, združimo v bloke po tri zaporedne enote poskusnega gradiva
Blok 1 (1,2,3) Blok 2 (4,5,6) Blok 3 (7,8,9)
Znotraj blokov pa je razmestitev postopkov slučajnostna.

	Blok 1			Blok 2			Blok 3		
i	1	2	3	4	5	6	7	8	9
	A_3	A_2	A_1	A_1	A_3	A_2	A_1	A_2	A_3
(A)	- 50	10	40	40	- 50	10	40	10	- 50
M+x+e	61	68	83	86	102	115	116	133	136
y_{BA}	11	78	123	126	52	125	156	143	86
y_B	212			303			385		900 = y

	A_1	A_2	A_3
y_A	405	346	149
$\bar{y}_A$	135	115.3	49.7
$\hat{A}_A$	35	15.3	- 50.3

Analiza variance zahteva naslednje posamezne količine

$$Q_{BA} = 11^2 + 78^2 + \ldots + 86^2 = 107720$$

$$Q_B = \frac{1}{3}\, 212^2 + 303^2 + 385^2 = 94992.7$$

$$Q_A = \frac{1}{3}\left[405^2 + 346^2 + 149^2\right] = 101980{,}7$$

$$Q = \frac{900^2}{9} = 90000$$

$q_{BA} = 17720$ $\qquad q_B = 4992.7$ $\qquad q_A = 11980.7$

a analiza variance je

VV	K	m	s^2	F
B	$q_B = 4992.67$	2	2496,33	
A	$q_A = 11980.67$	2	5990.33	31.6
pp	$q_{BA} = q_B - q_A = 746.67$	4	186.67	
Sk	$q_{BA} = 17720$	8		

F-preskus je odkril visoko značilne razlike. Ocene efektov

$\hat{A}_A(35,\ 15.3,\ -50.4)$ so se znatno približale pravim vrednostim

$A_A\ (40,\ 10,\ -50)$

Da preskusimo relativno uspešnost slučajnostnih popolnih blokov v našem simuliranem primeru, izračunajmo koeficient uspešnosti:

Ocena variance pri čisto slučajnostnem poskusu je:

$$s^2_{e.ČSP} = \frac{m_B\ s^2_B + (m_A + m_e)s^2_e}{m} = \frac{2 \cdot 2496,3 + (2+4)186,6}{8} = 764,0$$

$$E_{ČSB:ČSP} = \frac{(2+1).(2+4+3)\ 764,0}{(2+3).(2+4+1)\ 186,6} = 3,16$$

5.9 Manjkajoče vrednosti pri slučajnostnih popolnih blokih

Že na več mestih smo opomnili, da je enako število ponovitev tehnično in vsebinsko utèmeljeno, ker olajša izračune in interpretacijo rezultatov. Posebej velja to za slučajnostne popolne bloke pri katerih iz definicije blokov sledi, da mora celoten poskus vsebovati n = b.A enot. Včasih pa se le zgodi, da iz kateregakoli vzroka (posamezen poskus neuspel, ga ni bilo možno izvesti, donos uničen ipd.) iz popolnega poskusa izpade eden ali več rezultatov. Za take primere moremo za slučajnostne popolne bloke po posebni tehniki, ki je podlaga tudi pri manjkajočih podatkih pri drugih statističnih načrtih, en ali več podatkov, oceniti. Seveda je tako ocenjevanje uspešno le, če manjkajočih podatkov ni preveč.

Pri ocenjevanju manjkajočega podatka vzamemo kot osnovno načelo, da je ocena za manjkajoči podatek najboljša tista vrednost, pri kateri je s_e^2 najmanjši. Ker s_e^2 zavisi od vseh vrednosti poskusa, jasno zavisi tudi od manjkajočega. To načelo je logično, ker vemo, da bi ocena, ki bi bila zelo nerealna in izven stvarnih rezultatov poskusa, s_e^2 samo povečala.

Če z y'_{BA} zaznamujemo ocenjeno vrednost za postopek A v bloku B, je po zgornjem načelu

$$y'_{BA} = \frac{By'_B + Ay'_A - y'}{(B-1)(A-1)} \qquad (5.19)$$

pri tem je y'_B vsota znanih vrednosti v bloku v katerem podatek manjka, y'_A analogna vsota znanih vrednosti v vsoti za postopek A in y' vsota vseh znanih vrednot v celotnem poskusu seveda brez ocenjenega y'_{BA}.

Vzemimo za primer, da v našem poskusu število bakterij v mleku nimamo podatka za V_3, v drugi pošiljki (izmerjena vrednost $y_{3,2} = 12$). Potrebne vsote (vsote brez ocenjevane) so

$$y'_{B3} = 24 \qquad y'_{A2} = 28 \qquad y' = 109$$

ker je B = 5 in A = 4 dobimo

$$y'_{B3A2} = \frac{5.24+4.28-109}{(5-1)(4-1)} = \frac{123}{12} = 10,25$$

Ocena torej ne odstopa dosti od prave vrednosti (le za 1,75). Če obračunamo analizo variance z ocenjenim podatkom y'_{BA} je s'^2_A pristransko precenjen in ga je treba zmanjšati za

$$\frac{(y'_B - (b-1)y'_{BA})^2}{b(b-1)} \tag{5.20}$$

analogno moramo znižati število stopinj prostosti za standardni pogrešek m_e za 1 $\quad m'_e = m_e - 1 = (b-1)(A-1)-1$

varianca razlike med dvema povprečjema za postopke

$$\mathrm{Var}\ A\bar{y}_A = \sigma_e^2 \left[\frac{2}{b} + \frac{b}{A(A-1)(b-1)} \right]$$

namesto

$$\mathrm{Var}\ A\bar{y}_A = \frac{2\sigma_e^2}{b} \tag{5.21}$$

v primeru, da ni manjkajočih podatkov.

5.10 Metoda parov.

Poseben primer metode blokov je metoda parov. Če sta v poskusu samo dva postopka A_1 in A_2, ki ju proučujemo v blokih, se analiza variance in proučitev rezultatov takega poskusa zelo poenostavi.

Priredimo model za poskus s slučajnostnimi bloki za A = 2.

$$y_{b1} = M + A_1 + (B) + e_{b1} \qquad e_{bA} = : N(0, \sigma_e^2) \qquad (5.22)$$

$$y_{b2} = M + A_2 + (B) + e_{b2}$$

$$d_b = y_{b2} - y_{b1} = A_2 - A_1 + (e_{b2} - e_{b1})$$

Iz razlike vrednosti poskusov iz istega bloka, $d_b = y_{b2} - y_{b1}$ je izginila splošno poprečje M in kar je še važnejše, učinek bloka.

Iz zgornjega modela sklepamo, da se

$$d_b = : N(A_2 - A_1 : 2\sigma_e^2 = \sigma_d^2) \qquad (5.23)$$

in poprečje

$$\bar{d} = \frac{Sd_b}{b} = : N(A_2 - A_1 : \frac{\sigma_d^2}{b}) \qquad (5.24)$$

Nepristranska ocena za s_d^2 je

$$s_d^2 = \frac{S(d_b - \bar{d})}{b-1} \qquad (5.25)$$

Celoten problem se je reduciral na problem ocene poprečja $\bar{d}$, ki služi za osnovo ničelne domneve $H_o: A_2 = A_1$ oziroma $H_o: D \doteq 0$, ali za intervalno oceno razlike $A_2 - A_1$.

Preko analize variance pridemo do enakega rezultata po naslednjih stopnjah

$$y_{bA} \qquad y_A \qquad y_b = y_{bA1} + y_{bA2} \tag{5.26}$$

$$y_b \qquad y$$

$$Q_{bA} = \underset{bA}{S}\underset{}{S} y_{bA}^2 = \underset{b}{S}(y_{b1}^2 + y_{b2}^2)$$

$$Q_A = \frac{1}{b} \underset{A}{S}\, y_A^2 = \frac{1}{b}(y_1^2 + y_2^2)$$

$$Q_b = \frac{1}{2} \underset{b}{S} y_b^2 = \frac{1}{2} \underset{b}{S}(y_{b1} + y_{b2})^2 \tag{5.27}$$

$$Q = \frac{1}{2b}\, y^2 = \frac{1}{2b}(y_1 + y_2)^2$$

(5.27)

VV	K	m	s^2	F
B	$Q_b - Q = \frac{1}{2} \underset{b}{S} (y_{b1} + y_{b2})^2 - \frac{1}{b}(y_1+y_2)^2$			
A	$Q_A - Q = \frac{1}{b} \; y_1^2 + y_2^2 - \frac{1}{2}(y_1+y_2)^2 = \frac{1}{2b}(y_2 - y_1)^2$	1	$\frac{1}{2b}(y_2-y_1)^2 = \frac{1}{2b} D^2 = \frac{b}{2} \bar{d}^2$	$F_d = \frac{b\bar{d}^2}{s_d^2}$
pp	$Q_{bA} - Q_b = \underset{b}{S} (y_{b1}^2 + y_{b2}^2 - \frac{1}{2}(y_{b1}+y_{b2})^2 = \frac{1}{2b} S(y_{b2}-y_{b1})^2$	b-1	$\frac{1}{2} \frac{S(y_{b2}-y_{b1})^2}{(b-1)} = \frac{1}{2} s_d^2$	
Sk	$Q_{bA} - Q$	2b-1		

Ker je $F(1.b-1) = t^2(b-1)$ je

$$\sqrt{\frac{b\,\bar{d}^2}{s_d^2}} = : t(b-1) \qquad (5.29)$$

Dobljeni rezultat je identičen s prvotnim rezultatom. Metoda parov je zaradi svoje enostavnosti zelo uporabljana metoda. S pridom jo uporabljamo povsod tam, kjer proučujemo faktor z dvema vrednostima in je podana možnost za sestavljanje blokov oziroma parov vrednosti. Tako pri proučevanju značilnosti razlik v produktivnosti po dveh načinih dela (stari - novi) vzamemo kot blok poskusno enoto delavca, ki dela po starem in novem načinu. Pri proučevanju uspešnosti reklamiranja določenega blaga na prodajo vzamemo kot blok posamezno trgovino, v kateri prodajamo po starem in novem načinu in proučujemo dalje razliko v indeksih prodaje. Pri proučevanju uspešnosti prepariranja lesa vzamemo kot blok košček lesa, ki ga razpolovimo. En delček prepariramo; drugega pa ne in proučujemo razlike v karakteristikah obeh delcev, ali poskusne parcele razpolovimo in eno polovico posejemo negnojeno, drugo pa gnojeno. V posameznih primerih iz diferenc, ki jih izračunamo iz parov - bloka izločimo iz poskusa heterogenosti med bloki npr. med delavci, med trgovinami, med različnimi njivami. Pri poskusih, v katerih so poskusne enote ljudje, z metodo parov iz poskusa izločimo osebnostno komponento.

5.11

Za primer vzemimo proučitev spremembe cen določenega kmetijskega izdelka A v dveh različnih razdobjih T_1 in T_2. V ta namen smo na slučajnosten način izbrali b = 15 trgov in zanje zbrali podatke o cenah ob času T_1 in T_2. Rezultati registracije so naslednji:

	Trg	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Cena	T_1	8	7	9	13	8	6	9	7	5	13	11	10	9	7	8
	T_2	12	10	14	15	12	10	12	8	10	14	15	14	14	13	13
	d	4	3	5	2	4	4	3	1	5	1	4	4	5	6	5

Proučiti je treba, ali je podražitev v času T_2 značilna in če je, ugotoviti, za koliko se je cena s tveganjem $\alpha = 0.05$ najmanj povečala.

$$s_{d}^{2} = \frac{S_{d}^{2} - \frac{D^2}{n}}{n - 1} = \frac{240 - \frac{56^2}{15}}{15 - 1} = 2.210$$

$$se_{\bar{d}} = \sqrt{\frac{s_d^2}{n}} = \sqrt{\frac{2.210}{15}} = 0.384$$

$$\bar{d} = \frac{D}{n} = \frac{56}{15} = 3.733$$

$$\bar{d}_s = \bar{d} - t_{\alpha = P = 0.05}^{(14)} \; se_{\bar{d}} = 3{,}733 - 1{,}761 . 0.384 = 3{,}057$$

5.12 R a n k i t i v s l u č a j n o s t n i h b l o k i h .

Pri preskušanju kakovosti izdelkov glede na okus potrošnika moremo postopati tako, da vsak preskuševalec oceni eno samo vrsto izmed preskušanih izdelkov in da o njem določeno oceno v točkah (od 1-5, 1-10 ali 1-100 točk). Izkaže pa se, da so rezultati uspešnejši, če ena in ista oseba primerjalno ocenjuje več oziroma vse proučevane vrste izdelka in ne samo eno; ta način ima vse elemente metode blokov, pri čemer ima ocenjevalec vlogo bloka. Tako v veliki meri iz raziskave izločimo splošno osebnostno komponento ocenjevalcev, ki ocenjujejo bolj ali manj kritično in so njihove ocene v splošnem manjše ali večje.

Pogosto od ocenjevalca niti ne zahtevamo, da za vsako vrsto ocenjevanega izdelka da oceno s točkami, temveč se zadovoljimo s tem, da preskušana oseba določi vrstni red ocenjevanih izdelkov, tj. da jih rangira po kakovosti oziroma okusu. Z rangom izdelka dobimo pogosto zadostno primerjalno informacijo o kakovosti posameznih vrst blaga. Običajno nas ne zanima točkovna ocena za določeno vrsto blaga, temveč vrstni red izdelkov glede na oceno potrošnikov. Tak rezultat daje indikacijo o tem, katera vrsta blaga bolj ali manj ustreza okusu in željam potrošnika.

Za take vrste raziskav je v smislu statistične metodologije posamezen ocenjevalec blok, poskusna enota posamezno ocenjevanje, postopek - ocenjevanje določene vrste blaga, kriterialni znak pa rang, ki ga posamezen ocenjevalec določi v primerjalnem ocenjevanju različnim vrstam blaga. Rang kot kriterialni znak pa ne ustreza pogojem analize variance. Iz pravokotniške porazdelitve rangov pa dobimo s transformacijo rangov standardizirano normalno porazdelitev rankitov (glej tabelo 3.4). Ker rankiti ne zadoščajo pogojem normalnosti, jih moremo obravnavati z analizo variance.

5.13 Če postopek analiziranja podatkov poskusa slučajnostnih popolnih blokov prenesemo na naš primer je postopek naslednji:

Da bi proučili kakovost A različnih vrst blaga, izberemo b ocenjevalcev od katerih vsak oceni z rangi R_{bA} od 1 - A ocenjevane vrste blaga - postopke. Kako izberemo b ocenjevalcev je odvisno od cilja raziskave. V specializiranih raziskavah so to ocenjevalci ali preskuševalci - strokovnjaki. Če pa želimo dobiti rezultate, veljavne za populacijo potrošnikov, izberemo b potrošnikov - ocenjevalcev, slučajnostno. b kot faktor je v drugem primeru slučajnosten, medtem ko ima v prvem primeru običajno značaj fiksnega faktorja.

Range R_{bA} s transformacijo prevedemo v tabelo rankitov Z_{bA}. Iz definicije rankitov (standardizirani) povzamemo, da je

$$\sum_{A=1}^{A} Z_{bA} = Z_b = 0 \quad \text{in} \quad \sum_b Z_b = 0 \tag{5.30}$$

$$Z_{bA} \quad Z_b = 0$$

$$Z_A \quad Z = 0 \tag{5.31}$$

$$Q_{bA} = \sum_b \sum_A Z^2_{bA} = b \sum_{R=1}^{A} Z^2_R \qquad Q_A = \frac{1}{b} \sum Z^2_A \tag{5.31}$$

$$Q_b = 0 \qquad Q = 0$$

$$q_{bA} = Q_{bA} = b \sum^{A} Z^2_R \qquad q_A = Q_A$$

Številu postopkov A ustrezne vrednosti $\sum Z^2_R$ so tabelirane v tabeli rankitov.

Analiza variance rankitov je naslednja:

VV	K	m	s^2	F
A	Q_A	A-1	s^2_{ZA}	$F = s^2_{ZA}/s^2_{ze}$
pp	$Q_{bA} - Q_A$	(b-1)(A-1)	s^2_{Ze}	1
$\sum$	Q_{bA}	b(A-1)		

(5.32)

5.14 Vzemimo za primer analizo okusnosti štirih vrst sladoleda: A_1 = mešanica naravne in umetne vanilije; A_2 = naravna vanilija; A_3 = umetna vanilija; A_4 = brez dodatka vanilije. Vnaprej so planirane naslednje tri ortogonalne primerjave. a) ali dodana vanilija daje značilno boljši okus (neglede na to v kakšni kombinaciji), b) ali je mešanica značilno različna od čistega dodatka vanilije (neglede kakšnega), c) ali da naravna vanilija v primerjavi z umetno drug rezultat.

V preskus je vključeno b = 29 na slučajnosten način izbranih otrok, od katerih je vsak poskusil vsako izmed štirih vrst sladoleda in njihov okus rangiral z 1 - 4. Da preskuševalci ne bi vedeli za posamezne vrste, katere so, so bile vse vrste sladoleda na oko med seboj enake. Posamezne vrste so bile dane poskusnim osebam na slučajnosten način v tem smislu, da je bil vrstni red poskušanja slučajnosten. S tem je bil dodatni določljiv faktor, ki bi mogel vplivati na oceno - vrstni red vključen v slučajnostne vplive.

Rezultati poskusa so naslednji:

R	1	2	3	4			
Z_R	1.029	·297	-·297	- 1.029			
b							
1	A_1	A_2	A_4	A_3			
2	A_2	A_1	A_3	A_4			
·	·	·	·	·			
·	·	·	·	·			
·	·	·	·	·			
29	A_1	A_3	A_2	A_4	f_A	$f_{AR}Z_R = Z_A$	$Z_b=0$
f_{AR} A_1	15	7	2	5	29	11,80	
A_2	6	9	10	4	29	1,76	
A_3	3	6	4	16	29	- 12,79	
A_4	5	7	13	4	29	- ·77	
f_R	29	29	29	29	116	$0 = Z$	

$$Q_{bA} = \underset{b}{S}\underset{A}{\sum} Z^2_{bA} = b\sum Z^2_R = 29\left[1{,}03^2 + \cdot 30^2 + (-\cdot 30)^2 + (-1\cdot 03)^2\right] = 66{,}7522$$

$$Q_A = \frac{1}{b}\sum Z^2_A = \frac{1}{29}\left[11{,}80^2 + 1{,}76^2 + (-12{,}79)^2 + (-0.77)^2\right] = 10{,}5695$$

Analiza variance:

VV	K	m	s^2	F
A	$Q_A = 10{,}5695$	3	3.5119	$5{,}26^{xx}$
pp	$Q_{bA} - Q_A = 56{,}1827$	84	·6688	1,
Sk	$Q_{bA} = 66{,}7522$	87		

	Z_A	a	b	c
A_1	11,80	1	2	0
A_2	1,76	1	- 1	1
A_3	- 12,79	1	- 1	- 1
A_4	- •77	- 3	0	0
	0	+ 3,08	+ 34,63	14,55

			F
Primerjave	a	$3,08^2/29 \cdot \lbrack 1^2+1^2+1^2+(-3)^2 \rbrack = 0,0273$	-
	b	$34,63^2/29 \cdot \lbrack 2^2+(-1)^2+(-1)^2 \rbrack = 6,8922$	10,30
	c	$14,55^2/29 \cdot \lbrack 1^2+(-1)^2 \rbrack = \underline{3,6500}$	5,46
		10,5695	

Raziskava je pokazala, da je mešanica značilno boljša od poprečja čiste vanilije (umetne in naravne) in da je izmed čistih dodatkov naravna značilno boljša od umetne.

5.15 Friedmanova analiza variance.

Razen z rankiti proučujemo ordinalne podatke tudi na druge načine. Friedmanova analiza variance je npr. zasnovana na analizi vsot rangov po postopkih. Če imamo npr. b poskusnih oseb, ki rangirajo A postopkov po določenem kriteriju (npr. ocene po okusu, po izgledu itd.), dobimo za vsak postopek A b ocen, ki so dane z rangi. Če med postopki ni razlik, se vsote rangov za posamezen postopek $R_A = \sum_b R_{AB}$ odklanjajo od povprečja rangov $\bar{R} = \frac{b(A+1)}{2}$ samo zaradi slučajnostnih vplivov. V nasprotnem primeru, če ne velja ničelna domneva $H_o: A_1 = A_2 = A_3 = \dots = A_A$ pa variabilnost med vsotami rangov za postopke naraste. Ničelno domnevo preskušamo z Friedmanovo analizo variance po naslednjem postopku:

b oseb, ki rangira A postopkov, predstavlja b blokov. Če tabelo rangov zaznamujemo z R_{bA}, dobimo iz nje naslednje količine:

$$R_{bA} \qquad \sum_b R_{bA} = R_A \qquad R = \sum_A R_A = b\,\frac{A(A+1)}{2} \qquad \bar{R} = \frac{b(A+1)}{2} \qquad (5.33)$$

Dalje velja:

$$\frac{b\sum(R_A-\bar{R})^2}{R} = \frac{12}{bA(A+1)}\sum R_A^2 - 3b(A+1) = : X^2(m = A-1) \qquad (5.34)$$

da se zgornji izraz porazdeljuje približno v X^2- porazdelitvi z $m = A-1$ stopinjami prostosti, če velja ničelna domneva, da je $A_1 = A_2 = A_3 \ldots = A_A$.

Vzemimo za primer, da pet ocenjevalcev ocenjuje z razvrstitvijo po kvaliteti 4 različne vrste izdelkov. Ocena za vsako vrsto, da z rangi od 1 do 4.

Rezultati raziskave so naslednji:

ocenjevalec (b)

R_{bA}	1	2	3	4	5	R_A
A_1	1	2	1	1	1	6
A_2	4	3	4	3	4	18
A_3	3	4	2	4	3	16
A_4	2	1	3	2	2	10

$$R = 50 = \frac{5.(4)(4+1)}{2} = 50$$

Preskusni izraz je:

$$\frac{12}{5(4+1).4}\left[6^2+18^2+16^2+10^2\right] - 3.5(4+1) = 10,92$$

$\chi^2_{0.05}(m=4-1=3) = 9{,}21 \qquad \chi^2_{0.001}(m=3) = 13{,}08$

Preskus je pokazal, da so razlike značilne na stopnji $\alpha = 0.01$ tveganja.

Seveda Friedmanova analiza variance odkrije samo značilnost razlik med rangi.

5.16 Značilnost apriornega vrstnega reda $H_1 : A_1 > A_2 > A_3 .. > A_4$ pa preskušamo z L - preskusom. V ta namen iz vsot rangov R_{Ai}, ki so urejene po vrstnem redu, kot predvideva H_1, izračunamo izraz

$$L = \sum_{i=1}^{A} iR_{Ai} \tag{5.35}$$

Kritične vrednosti za L pri danem b in A so za b = 2 do 25 in A = 3 do 10 dane v tabeli 5.1.

Če za naš primer vnaprej domnevamo, da je vrstni red kakovosti $H_i : A_1 < A_4 < A_2 < A_3$ dobimo, da je

$$L = 1.R_{A1} + 2R_{A4} + 3R_{A2} + 4R_{A3} =$$

$$= 1.6 + 2.10 + 3.18 + 4.16 = 144$$

Ker je $\quad L_{0.01}(b=5, A=4) = 141 \quad$ in $\quad L_{0.001}(b=5, A=4) = 145$

vzamemo, da se je apriorna razvrstitev izkazala kot značilna na stopnji $\alpha = 0.01$

Za $A > 10$ in velike b moremo vzeti kot približek, da se

$$\frac{12L - 3bA(A+1)^2}{bA^2(A^2-1)(A+1)} = : \chi^2(m=1) \tag{5.36}$$

s kritičnimi vrednostmi

$\chi_{0.10}(1) = 2{,}71 \quad \chi^2_{0.05}(1) = 3{,}84 \quad \chi^2_{0.01}(1) = 6{,}63 \quad \chi^2_{0.001}(1) = 10{,}8$

Tabela 5.1.

Kritične vrednosti za L

k	A: 3	4	5	6	7	8	9	10
	—	—	109	178	269	388	544	726
	—	60	106	173	261	376	520	696
	28	58	103	160	252	362	500	670
	—	89	160	260	394	567	790	1056
3	42	87	155	252	382	549	761	1019
	41	84	150	244	370	532	736	987
	56	117	210	341	516	743	1032	1382
4	55	114	204	331	501	722	999	1339
	54	111	197	321	487	701	971	1301
	70	145	259	420	637	917	1273	1704
5	68	141	251	409	620	893	1236	1656
	66	137	244	397	603	869	1204	1614
	83	172	307	499	757	1090	1512	2025
6	81	167	299	486	737	1063	1472	1972
	79	163	291	474	719	1037	1436	1927
	96	198	355	577	876	1262	1750	2344
7	93	193	346	563	855	1232	1706	2288
	91	189	338	550	835	1204	1668	2238
	109	225	403	655	994	1433	1987	2662
8	106	220	393	640	972	1401	1940	2602
	104	214	384	625	950	1371	1900	2549
	121	252	451	733	1113	1603	2223	2980
9	119	246	441	717	1088	1569	2174	2915
	116	240	431	701	1065	1537	2131	2859
	134	278	499	811	1230	1773	2459	3296
10	131	272	487	793	1205	1736	2407	3228
	128	266	477	777	1180	1703	2361	3169
	147	305	546	888	1348	1943	2694	3612
11	144	298	534	869	1321	1905	2639	3541
	141	292	523	852	1295	1868	2592	3478
	160	331	593	965	1465	2112	2929	3927
12	156	324	581	946	1437	2072	2872	3852
	153	317	570	928	1410	2035	2822	3788

prva vrsta $\alpha = 0{\cdot}001$. druga vrsta $\alpha = 0{\cdot}01$.
treća vrsta $\alpha = 0{\cdot}05$

Kritične vrednosti za d_w

b	A: 3	4	5	6	7	8	9	10
	172	358	642	1044	1585	2285	3163	4241
13	169	350	628	1022	1553	2240	3104	4164
	165	343	615	1003	1525	2201	3052	4097
	185	384	689	1121	1702	2453	3397	4556
14	181	376	674	1098	1668	2407	3335	4475
	178	368	661	1078	1639	2367	3281	4405
	197	410	736	1197	1818	2622	3631	4869
15	194	402	721	1174	1784	2574	3567	4786
	190	394	707	1153	1754	2532	3511	4714
	210	436	783	1274	1935	2790	3864	5183
16	206	427	767	1249	1899	2740	2740	3798
	202	420	754	1228	1868	2697	3741	5022
	223	463	830	1350	2051	2958	4098	5496
17	218	453	814	1325	2014	2907	4029	5407
	215	445	800	1303	1982	2862	3970	5330
	235	489	876	1427	2167	3126	4330	5808
18	231	479	860	1401	2130	3073	4260	5717
	227	471	846	1378	2097	3028	4199	5638
	248	515	923	1503	2283	3294	4563	6121
19	243	505	906	1476	2245	3240	4491	6027
	239	496	891	1453	2217	3193	4428	5946
	260	541	970	1579	2399	3461	4796	6433
20	256	531	953	1552	2360	3406	4722	6337
	251	522	937	1528	2325	3358	4657	6253
	273	567	1017	1656	2515	3629	5028	6745
21	268	556	999	1628	2475	3572	4952	6647
	263	547	983	1603	2439	3523	4886	6561
	285	593	1063	1732	2631	3796	5260	7057
22	280	582	1045	1703	2589	3738	5182	6956
	275	573	1029	1678	2553	3687	5115	6868
	298	619	1110	1808	2747	3963	5492	7368
23	292	608	1091	1778	2704	3904	5413	7265
	288	598	1075	1753	2667	3852	5343	7176

6. LATINSKI KVADRATI

6.1 Poskuse oziroma poskusne enote združujemo v bloke po različnih določljivih, a za raziskavo nepomembnih faktorjih. Pri določenih problemih more več faktorjev služiti kot vodilo za sestavljanje blokov. Če proučujemo odvisnost višine prodaje od različnih načinov prodaje, more biti blok trgovina, v kateri prodajamo zaporedoma po vseh načinih prodaje, ki predstavljajo postopke. Ker pa višina prodaje zavisi tudi od dneva, more kot vodilo za sestavljanje blokov biti tudi dan. Če je dan vzet kot popoln blok, moramo na isti dan prodajati po vseh štirih načinih. Ker popolne bloke sestavljamo po enem, čeprav tipološkem znaku, izgleda, da se je potrebno odločiti za tistega, s katerim uspemo pojasniti največ variabilnosti in s tem zmanjšati poskusni pogrešek. S posebnim dodajanjem postopkov poskusnim enotam pa uspemo, da so postopki v popolnih blokih po dveh faktorjih hkrati. Če npr. preskušamo razlike v štirih načinih prodaje, moremo s štirimi trgovinami kot bloki pri primerno načrtovanem poskusu vključiti v bloke tudi čas oziroma dan. To dosežemo, če trgovinam vsak dan določamo prodajne načine tako, da se vsak dan v štirih trgovinah zvrste vsi štirje postopki. Načrt takega poskusa je dan v naslednji shemi:

	Trg 1	Trg 2	Trg 3	Trg 4
Torek	A_1	A_2	A_3	A_4
Sreda	A_2	A_3	A_4	A_1
četrtek	A_3	A_4	A_1	A_2
petek	A_4	A_1	A_2	A_3

V tej shemi je prikazan načrt, ki ga zaradi kvadrataste oblike sheme imenujemo latinski kvadrat. V njem so postopki (A) uporabljeni tako, da so postopki združeni v bloke po dveh faktorjih (dan in trgovina).

6.2 Za latinski kvadrat velja naslednji linearni model

$$Y_{VSA} = M + (V) + (S) + (A) + e_{VSA} \; ; \; e_{VSA} = : N(0, \sigma_e^2) \quad (6.1)$$

Pri tem en določljiv, a nepomemben faktor zaznamujemo z (V) (vrstica v kvadratu) drug določljiv, a nepomemben faktor z (S) (stolpec v kvadratu), z A pa postopek. Zaradi konstrukcije latinskega kvadrata je število nivojev za vse tri faktorje (vrstico, stolpec in faktor A) enako številu postopkov A. Če vzamemo, da so vsi trije faktorji faktorji modela I. zanje velja $\sum V = 0$, $\sum S = 0$ $\sum A = 0$, za poskusni pogrešek pa velja predpostavka $e_{VSA} = : N(0, \sigma_e^2)$.

Iz konstrukcije latinskega kvadrata sledi, da moremo iz osnovnih podatkov Y_{VSA} sestaviti vsote y_V, y_S , y_A . Ker so v vsotah y_V zastopane vse vrednosti stolpcev in vsi postopki y_V variira zaradi razlik v učinku faktorja V in poskusnega pogreška. Podobno velja za (S) in (A).

Analizo variance za latinski kvadrat obračunamo po naslednjem postopku.

Če izhjamo iz modela

$$Y_{VSA} = M + (V) + (S) + (A) + e_{VSA} \qquad \sum(V) = 0 \qquad \sum(S) = 0$$

$$\sum(A) = 0 \qquad e_{VSA} = : N(0, \sigma_e^2) \qquad (6.2)$$

v katerem smatramo, da so učinki vseh treh faktorjev vezani aditivno, so torej med seboj neodvisni, izračunamo najprej iz Y_{VSA} ustrezne vsote.

$$
\begin{array}{ll}
y_{VSA} & y_V \\
y_S & y \\
y_A &
\end{array}
\qquad (6.3)
$$

Iz teh vrst izračunamo pomožne količine:

$$Q_{VSA} = \sum_V \sum_S \sum_A y^2_{VSA} \qquad q_{VSA} = Q_{VSA} - Q \qquad (6.4)$$

$$Q_V = \frac{1}{A} \sum_V y^2_V \qquad q_V = Q_V - Q$$

$$Q_S = \frac{1}{A} \sum_S y^2_S \qquad q_S = Q_S - Q$$

$$Q_A = \frac{1}{A} \sum_A y^2_A \qquad q_A = Q_A - Q$$

$$Q = \frac{1}{A^2} y^2$$

Iz njih pa obračunamo analizo variance po naslednji shemi:

(6.5)

VV	K	m	s^2	F	E_{s^2}
V	q_V	A - 1	s^2_V		$\sigma^2_e + A\sigma^2_V$
S	q_S	A - 1	s^2_S		$\sigma^2_e + A\sigma^2_S$
A	q_A	A - 1	s^2_A	$F_A = s^2_A/s^2_e$	$\sigma^2_e + A\sigma^2_A$
pp	$q_{VSA} - q_V - q_S - q_A$	(A-1)(A-2)	s^2_e	1	σ^2_e
Sk	q_{VSA}	$A^2 - 1$			

V shemi nakazan primer predpostavlja, da je za raziskavo pomemben samo faktor (A). More pa biti za raziskavo pomemben tudi faktor (V) ali (S) ali oba. Vendar moramo paziti, da moremo v latinskem kvadratu kombinirati raziskavo dveh ali treh faktorjev le, če moremo predpostavljati, da so ti faktorji med seboj neodvisni. V tem primeru faktor (V) in (S) primerjamo z

$$F_V = s_V^2/s_e^2 \qquad \text{in } F_S = s_S^2/s_e^2$$

6.3 S t a n d a r d n i   l a t i n s k i   k v a d r a t i .

Latinski kvadrat imenujemo standarden, če sta prvi stolpec in prva vrsta urejena po zaporedju postopkov (ali po abecedi ali numerično npr. standarden latinski kvadrat 3 x 3 predstavlja poskus z naslednjo ureditvijo postopkov:

$$\begin{matrix} A_1 & A_2 & A_3 \\ A_2 & A_3 & A_1 \\ A_3 & A_1 & A_2 \end{matrix} \tag{6.6}$$

Dva latinska kvadrata sta konjugirana, če so vrstice enega kvadrata identične stolpcem drugega. Konjugiran v sebi pa je kvadrat, če je konjugiran kvadrat identičen osnovnemu kvadratu. Tako je npr. kojugiran v sebi naslednji latinski kvadrat 4 x 4.

$$\begin{matrix} A_1 & A_2 & A_3 & A_4 \\ A_2 & A_1 & A_4 & A_3 \\ A_3 & A_4 & A_2 & A_1 \\ A_4 & A_3 & A_1 & A_2 \end{matrix} \tag{6.7}$$

Iz vsakega standardnega latinskega kvadrata moremo s preureditvijo vrst ali stolpcev dobiti A ! (A-1)! različnih latinskih kvadratov.

Za latinske kvadrate od 2 x 2 do 6 x 6 je število različnih kvadratov naslednje:

Latinski kvadrat	Število standardnih latinskih kvadratov	Število možnih razvrstitev
2 x 2	1	2
3 x 3	1	12
4 x 4	4	576
5 x 5	56	161.280
6 x 6	9.408	812.851.200

Število standardnih LK, še bolj pa število možnih razvrstitev z večanjem LK močno narašča. Slučajnostne LK dobimo tako, da najprej izberemo na slučajnosten način standarden LK nato pa slučajnostno razporedimo vrste in stolpce. Tako npr. iz standardnega LK 4 x 4

2	A_1	A_2	A_3	A_4
4	A_2	A_1	A_4	A_3
3	A_3	A_4	A_2	A_1
1	A_4	A_3	A_1	A_2

najprej z nakazano slučajnostno permutacijo vrstic dobimo

	4	3	1	2
1	A_4	A_3	A_1	A_2
2	A_1	A_2	A_3	A_4
3	A_3	A_4	A_2	A_1
4	A_2	A_1	A_4	A_3

nato pa s slučajnostnimi permutacijami stolpcev

	1	2	3	4
1	A_1	A_2	A_3	A_4
2	A_3	A_4	A_2	A_1
3	A_2	A_1	A_4	A_3
4	A_4	A_3	A_1	A_2

3 × 3

A	B	C
B	C	A
C	A	B

4 × 4

1

A	B	C	D
B	A	D	C
C	D	B	A
D	C	A	B

2

A	B	C	D
B	C	D	A
C	D	A	B
D	A	B	C

3

A	B	C	D
B	D	A	C
C	A	D	B
D	C	B	A

4

A	B	C	D
B	A	D	C
C	D	A	B
D	C	B	A

5 × 5

A	B	C	D	E
B	A	E	C	D
C	D	A	E	B
D	E	B	A	C
E	C	D	B	A

6 × 6

A	B	C	D	E	F
B	F	D	C	A	E
C	D	E	F	B	A
D	A	F	E	C	B
E	C	A	B	F	D
F	E	B	A	D	C

7 × 7

A	B	C	D	E	F	G
B	C	D	E	F	G	A
C	D	E	F	G	A	B
D	E	F	G	A	B	C
E	F	G	A	B	C	D
F	G	A	B	C	D	E
G	A	B	C	D	E	F

8 × 8

A	B	C	D	E	F	G	H
B	C	D	E	F	G	H	A
C	D	E	F	G	H	A	B
D	E	F	G	H	A	B	C
E	F	G	H	A	B	C	D
F	G	H	A	B	C	D	E
G	H	A	B	C	D	E	F
H	A	B	C	D	E	F	G

9 × 9

A	B	C	D	E	F	G	H	I
B	C	D	E	F	G	H	I	A
C	D	E	F	G	H	I	A	B
D	E	F	G	H	I	A	B	C
E	F	G	H	I	A	B	C	D
F	G	H	I	A	B	C	D	E
G	H	I	A	B	C	D	E	F
H	I	A	B	C	D	E	F	G
I	A	B	C	D	E	F	G	H

10 × 10

A	B	C	D	E	F	G	H	I	J
B	C	D	E	F	G	H	I	J	A
C	D	E	F	G	H	I	J	A	B
D	E	F	G	H	I	J	A	B	C
E	F	G	H	I	J	A	B	C	D
F	G	H	I	J	A	B	C	D	E
G	H	I	J	A	B	C	D	E	F
H	I	J	A	B	C	D	E	F	G
I	J	A	B	C	D	E	F	G	H
J	A	B	C	D	E	F	G	H	I

11 × 11

A	B	C	D	E	F	G	H	I	J	K
B	C	D	E	F	G	H	I	J	K	A
C	D	E	F	G	H	I	J	K	A	B
D	E	F	G	H	I	J	K	A	B	C
E	F	G	H	I	J	K	A	B	C	D
F	G	H	I	J	K	A	B	C	D	E
G	H	I	J	K	A	B	C	D	E	F
H	I	J	K	A	B	C	D	E	F	G
I	J	K	A	B	C	D	E	F	G	H
J	K	A	B	C	D	E	F	G	H	I
K	A	B	C	D	E	F	G	H	I	J

Tabela 6.1 Izbrani latinski kvadrati.

V tabeli 6.1 je danih nekaj izbranih LK iz katerih na opisani način dobimo s permutiranjem vrstic in stolpcev slučajnostne LK.

6.4 Za primer latinskega kvadrata 4 x 4 vzemimo raziskavo avtomobilskih gum s cestnim testom. Faktor, ki ga proučujemo je vrsta gume, uporabljena za izdelavo avtomobilskih plaščev (4 vrste: A_1 A_2 A_3 A_4). Kriterialni znak je debelina sloja gume na kolesih po določenem številu kilometrov. 4 plašči vsake vrste (n = 16) so bili montirani na 4 kolesa štirih avtomobilov tako, da je en faktor predstavljala pozicija kolesa SD, SL, ZD, ZL. Poskus za faktor (A) je bil načrtovan v latinskem kvadratu s faktorjema vozilo in pozicija (V) in (S) in vrsto plašča kot faktorja, ki ga raziskujemo.

Rezultati, pisani poleg ustreznih oznak postopkov, so naslednji:

Pozicija

avtomobile

y_{VSA}	S_1	S_2	S_3	S_4	y_V
V_1	A_1 4	A_2 6	A_3 12	A_4 17	39
V_2	A_2 8	A_3 14	A_4 18	A_1 8	48
V_3	A_3 12	A_4 18	A_1 5	A_2 8	43
V_4	A_4 19	A_1 5	A_2 6	A_3 12	42
y_S	43	43	41	45	172

	A_1	A_2	A_3	A_4
y_A	22	28	50	72

$$Q_{VSA} = \sum_V \sum_S \sum_A y_{VSA}^2 = 2256 \qquad q_{VSA} = 407$$

$$Q_V = \frac{1}{4} \sum_V y_V^2 = \frac{1}{4} \left[39^2 + 48^2 + 43^2 + 42^2 \right] = 1859,5 \qquad q_V = 10,5$$

$$Q_S = \frac{1}{4}\sum_S y_S^2 = \frac{1}{4}\left[43^2+43^2+41^2+45^2\right] = 1851 \qquad q_S = 2,0$$

$$Q_A = \frac{1}{4}\sum_A y_A^2 = \frac{1}{4}\left[22^2+28^2+50^2+72^2\right] = 2238 \qquad q_A = 389,0$$

$$Q = \frac{1}{4^2}\, y^2 = \frac{1}{4^2}\, 172^2 = 1849$$

Analiza variance:

Vir variacije	K	m	s^2	
vozilo V	$q_V = 10,5$	3	$s_V^2 = 3,5$	
pozicija S	$q_S = 2,0$	3	$s_S^2 = 0,67$	
guma A	$q_A = 389,0$	3	$s_A^2 = 129,67$	$F_A = \frac{129,67}{2,58} = 50$
	$q_{VSA} = 407$ $- q_V = 10,5$ $- q_S = 2,0$ $- q_A = 389,0$			
pp	$K_e = 15,5$	6	$s_e^2 = 2,58$	1

Ker je $F_{\alpha=0.001}(3,6) = 9,18$ zmatno manjši kot izračunani F, sklepamo, da so razlike v kvaliteti gum visoko značilne.

Analiza poprečij za posamezne vrste gum nakaže naslednje sklepe. Ker nismo vnaprej načrtovali nobenih primerjav, izvedemo posteriorno analizo

$$\bar{y}_A = \frac{y_A}{4}$$

$$\bar{y}_{A1} = \frac{22}{4} = 5,5 \qquad \bar{y}_{A2} = \frac{28}{4} = 7 \qquad \bar{y}_{A3} = \frac{50}{4} = 12,5 \qquad \bar{y}_{A4} = \frac{72}{4} = 18$$

$$s_{\bar{y}} = \sqrt{\frac{s_e^2}{A}} = \sqrt{\frac{2,58}{4}} = \sqrt{6,45} = 2,54$$

	α = o.o5	α = o.ol	$W^{(K)}_{o.o5}$	$W^{(K)}_{o.ol}$
q (K=2,m=6) =	3,46	5,24	8.79	13,31
q (K=3,m=6) =	4,34	6,33	11,02	16,08
q (K=4,m=6) =	4,90	7,03	12,45	17,86

	A_1 5.5	A_2 7	A_3 12.5	A_4 18
A_4 = 18	12.5	11	5.5	
A_3 = 12.5	7 $^{K=3}$	5.5		
A_2 = 7	1.5 $^{K=2}$			
A_1 = 5.5				

Kot značilno različni α = o.o5 so se pokazali plašči kakovosti A_4 od kakovosti A_1 in A_2.

6.5 Uspešnost poskusov v Latinskih kvadratih.

Uspešnost poskusov v latinskih kvadratih v primerjavi s čisto slučajnostnim poskusom ali poskusom v popolnih blokih enake velikine dobimo s primerjavo ocen varianc poskusnega pogreška z vključitvijo doprinosov izključenih faktorjev. Tako je ocena variance poskusnega pogreška pri čisto slučajnostnem poskusu $s^2_{e.VS}$ iz podatkov analize variance

$$s^2_{e.VS} = \frac{s_V^2 + s_S^2 + (A-1)s_e^2}{(A+1)} \qquad (6.8)$$

Ocena variance poskusnega pogreška pri metodi blokov, če so bloki vrstice

$$s^2_{e.S} = \frac{s^2_S + (A-1)s^2_e}{A} \qquad (6.9)$$

Če pa so bloki stolpci

$$s^2_{e.V} = \frac{s^2_S + (A-1)s^2_e}{A} \qquad (6.10)$$

glede na to so koeficienti uspešnosti LK : ČSP

$$E_{LK:ČSP} = \frac{[(A-1)(A-2)+1][A(A-1)+3]}{[(A-1)(A-2)+3][A(A-1)+1]} \frac{s^2_{e.VS}}{s^2_e} = D_{ČS}(A) \frac{s^2_{e.VS}}{s^2_e} \qquad (6.11)$$

$$E_{LK:SB.V} = \frac{[(A-1)(A-2)+1][(A-1)^2+3]}{[(A-1)(A-2)+3][(A-1)^2+1]} \frac{s^2_{e.S}}{s^2_e} = D_{SB}(A) \frac{s^2_{e.s}}{s^2_e} \qquad (6.12)$$

$$E_{LK:SB.S} = \frac{[(A-1)(A-2)+1][(A-1)^2+3]}{[(A-1)(A-2)+3][(A-1)^2+1]} \frac{s^2_{e.V}}{s^2_e} = D_{SB}(A) \frac{s^2_{e.V}}{s^2_e} \qquad (6.13)$$

Korekturni členi $D_{ČS}$ in D_{SB}, ki so funkcije velikosti latinskih kvadratov so blizu 1 in imajo za A=3 do A=8 naslednje vrednosti:

A	$D_{ČS}$	D_{SB}
3	·7714	·8400
4	·8974	·9333
5	·9492	·9686
6	·9719	·9833
7	·9831	·9902
8	·9891	·9938

6.6 V našem primeru raziskave z avtomobilskimi gumami se izkaže naslednje:

$$E_{LK.ČSP} = D_{ČS}(A)\,\frac{s_V^2 + s_S^2 + 3\,s_e^2}{5\,s_e^2} =$$

$$= \cdot 8974\,\frac{3.5 + 0{,}67 + 3.2{,}58}{5\,.\,2{,}58} = \cdot 7179$$

$$E_{LK:SB.S} = D_{SB}(4)\,\frac{s_S^2 + 3\,s_e^2}{4\,s_e^2} =$$

$$= \cdot 9333\,\frac{0{,}67 + 3\,.\,2{,}58}{4\,.\,2{,}58} = \cdot 7605$$

$$E_{LK:SB.V} = D_{SB}(4)\,\frac{s_V^2 + 3\,s_e^2}{4.\,s_e^2} =$$

$$= \cdot 9333\,\frac{3{,}5 + 3\,.\,2{,}58}{4.\,2{,}58} = 1{,}0165$$

Koeficienti uspešnosti pokažejo, da je latinski kvadrat dal le v primerjavi z s slučajnostnimi bloki vrstic za 1,65% boljše rezultate.

6.7 Za simulirani primer s trendom v nosnovnem gradivu ima smisel uporabiti latinski kvadrat. V njem z razvrstitvijo osnovnega gradiva, ki je urejen po velikosti

	S_1	S_2	S_3
V_1	1	2	3
V_2	4	5	6
V_3	7	8	9

z vrsticami izločimo vpliv dolgoročnejšega vpliva linearnega trenda (prva vrstica 1 - 3 so enote na začetku, druga vrstica 4 - 6 so enote v sredini in tretja vrstica so enote na koncu gradiva) s stolpci pa kratkoročni vpliv trenda (v prvem stolpcu 1,: 4,7 so prve enote v blokih - vrsticah, v drugem stolpcu 2, 5, 8 so srednje enote v blokih in v tretjem stolpcu so zadnje enote v blokih 3, 6, 9. Če tako razporejenim poskusnim enotam pripišemo postopke, v shemi latinskega kvadrata, dobimo načrt:

$1/A_1$	$2/A_2$	$3/A_3$
$4/A_3$	$5/_{A1}$	$6/A_2$
$7/A_2$	$8/_{A_3}$	$9/A_1$

S sestavljanjem komponent dobimo:

	1 A_1	2 A_2	3 A_3	4 A_3	5 A_1	6 A_2	7 A_2	8 A_3	9 A_1
(A)	40	10	-50	-50	40	10	10	-50	40
M+x+e	61	68	83	86	102	115	116	133	136
y_{VSA}	101	78	33	36	142	125	126	83	176

Če prepišemo zgornje podatke v standardno obliko latinskega kvadrata, dobimo:

y_{SVA}	S_1	S_2	S_3	y_V
	A_1 101	A_2 78	A_3 33	212
	A_3 36	A_1 142	A_2 125	303
	A_2 126	A_3 83	A_1 178	385
y_S	263	303	334	900 y

	A_1	A_2	A_3
y_A	419	329	152
$\bar{y}_A$	139.7	109.7	50,7
$\hat{A}$	39.7	9.7	-49,3

Pomožne količine, potrebne za obračun analize variance, so:

$$Q_{VSA} = 101^2 + 178^2 = 108200 \qquad q_{VSA} = Q_{VSA} - Q = 18200$$

$$Q_V = \frac{1}{3}\left[212^2+303^2+385^2\right] = 94992,67 \qquad q_V = Q_V - Q = 4992,67$$

$$Q_S = \frac{1}{3}\left[263^2+303^2+334^2\right] = 90844,67 \qquad q_S = Q_S - Q = 844,67$$

$$Q_A = \frac{1}{3}\left[419^2+329^2+152^2\right] = 102302 \qquad q = Q_A - Q = 12302$$

$$Q = \frac{1}{9}\, 900^2 = 90000$$

Analiza variance pa je

VV	K	m	s^2	F
(V)	$q_V = 4992.67$	2	2496,33	
(S)	$q_S = 844.67$	2	422,33	
(A)	$q_A = 12302$	2	6151	202,80
pp e	$q_{VSA} - q_V - q_S - q_A = 60,66$	2	30,33	1.00
Sk	$q_{VSA} = 18200$	8		

Iz analize variance se jasno vidi, da so bloki po vrsticah (stratumi) pojasnili znatno več od skupnega K (4992.67) kot pa stolpci (samo 844.67).

Uspešnost latinskih kvadratov v primerjavi z enotnim slučajnostnim poskusom je naslednja;

$$s^2_{e.VS} = \frac{2496,33+422,33+(3-1).30,33}{3+1} = 744,83$$

$$E_{LK:ČSP} = ,7714 \cdot \frac{744,83}{30,33} = 18,94$$

Uspešnost latinskih kvadratov je v tem primeru zelo velika. To ni slučajno. Zato je uporaba latinskih kvadratov, v vseh primerih poskusnega gradiva, v katerem je trend če že ne linearen, vsaj monotono naraščajoč ali padajoč, priporočljiva.

6.8 Manjkajoči podatek pvo latinskem kvadratu.

Iz istih razlogov kot pri slučajnostnih popolnih blokih more tudi pri latinskem kvadratu manjkati en ali več osnovnih podatkov. Če manjka le eden, ga moremo po načelu, da je najboljša ocena tista, ki da najmanjši poskusni pogrešek, oceniti po obrazcu

$$y'_{VSA} = \frac{A(y'_V + y'_S \mp y'_A) - 2y}{(A-1)(A-2)} \qquad (6.14)$$

Seveda je treba število stopinj prostosti za poskusni pogrešek m_e zmanjšati za 1 $m'_e = (A-1)(A-2)-1)$, vsoto kvadratov ~~xx~~ K_A za postopek pa popraviti po obrazcu

$$K'_A = K_A - \frac{\left[y' - y'_V - y_S - (A-1)\, y'_A\right]^2}{\left[(A-1)(A-2)\right]^2} \qquad (6.15)$$

y'_V, y'_S, y'_A in y' so ustrezne vsote za vrsto, stolpec in postopek, v katerem manjka podatek.

Če vzamemo za primer, da v simuliranem primeru manjka podatek y_{111} to je podatek iz prve vrstice, prvega stolpca za postopek A_1 je ocena y'_{111} naslednja:

$$y'_{V1} = 78+33 = 111 \quad y'_{S1} = 36+126 = 162 \quad y'_{A1} = 142+178 = 320$$

$$y' = 78+33+ \ldots 178 = 799 \;, \quad A = 3$$

če vstavimo te podatke v zgornji obrazec dobimo

$$y'_{111} = \frac{3(111+162+162+320) - 2.799}{(3-1)(3-2)} = 90.5$$

prava vrednost pa je 101.

6.9 G r š k o - l a t i n s k i k v a d r a t i .

Z latinskim kvadratom izločimo dva določljiva, a za raziskavo nepomembna faktorja, ki ju v dani shemi tehnično vključimo enega v vrstice, drugega pa v stolpce kvadrata. Z grško-latinskim kvadratom izločimo tri določljive, a za raziskavo nepomembne faktorje. Grško-latinski kvadrat sestoji iz dveh ortogonalnih kvadratov. V njem so prirejene vrednosti tretjega nepomembnega faktorja T tako, da so ortogonalne faktorjem vrstic, stolpcev in proučevanega faktorjaA.

Kot primer grško-latinskega kvadrata v originalni obliki, v kateri so vrednosti enega faktorja v kvadratu zaznamovane z latinskimi, drugega pa z grškimi črkami, vzemimo grško latinski kvadrat 4 x 4

stolpci

vrstice	1	2	3	4
1	$A\alpha$	$B\beta$	$C\gamma$	$D\delta$
2	$B\delta$	$A\gamma$	$D\beta$	$C\alpha$
3	$C\beta$	$D\alpha$	$A\delta$	$B\gamma$
4	$D\gamma$	$C\delta$	$B\alpha$	$A\beta$

(6.16)

Če vzamemo npr. postopek A, dobimo da so z A vezani $A\alpha$, $A\gamma$, $A\delta$, $A\beta$, torej vse vrednosti faktorja zaznamovanega z grškimi črkami. Enako so npr. z vrednostjo γ vezani $C\gamma$ $A\gamma$ $B\gamma$ $D\gamma$ vse vrednosti faktorja zaznamovanega z latinskimi črkami. V grško-latinskem kvadratu vezane vrednosti štirih faktorjev moremo za naš primer 4 x 4 pisati tudi v obliki:

Poskusna enota	V	S	α	A
1	1	1	1	1
2	1	2	2	2
3	1	3	3	3
4	1	4	4	4
5	2	1	4	2
6	2	2	3	1
7	2	3	2	4
8	2	4	1	3
9	3	1	2	3
10	3	2	1	4
11	3	3	4	1
12	3	4	3	2
13	4	1	3	4
14	4	2	4	3
15	4	3	1	2
16	4	4	2	1

z α_4 so vezane enote

Enota	V	S	α	A
4	(1	4	4	4)
5	(2	1	4	2)
11	(3	3	4	1)
14	(4	2	4	3)

Iz primera vidimo, da so z α_4 vezane vse štiri vrednosti V, S in A.

6.10 Model poskusa s ponovitvami v grško-latinskem kvadratu je

$$Y_{VS\alpha A} = M+(V)+(S)+(\alpha)+(A)+e_{VS\alpha A}$$

$$e_{VS\alpha Ai} = : N(0,\sigma_e^2) \quad (6.17)$$

A analiza variance za grško-latinski kvadrat A x A fiksnih faktorjev (Model I) je naslednja:

VV	K	m	s^2	$E(s^2)$ (6.18)
V	$K_V = q_V$	A-1	s_V^2	$\sigma_e^2 + A\sigma_V^2$
S	$K_S = q_S$	(A-1)	s_S^2	$\sigma_e^2 + A\sigma_S^2$
α	$K_\alpha = q_\alpha$	(A-1)	s^2	$\sigma_e^2 + A\sigma_\alpha^2$
A	$K_A = q_A$	(A-1)	s_A^2	$\sigma_e^2 + A\sigma_A^2$
pp	$K_e = q_{VS\alpha A} - q_V - q_S - q_\alpha - q_A$	(A-1)(A-3)	s_e^2	σ_e^2
Sk	$K = q_{VS\alpha A}$	$A^2 - 1$		

Če je za raziskavo pomemben le faktor A, z F-preskusom, primerjamo le $F_A = s_A^2/s_e^2$. Podobno kot v latinskem kvadratu pa morajo biti vsi ali nekaj faktorjev, ne pa samo eden, za raziskavo pomembni faktorji; če le moremo predpostaviti, da so med seboj neodvisni. V tem primeru je F-preskus uporaben za vse za raziskavo pomembne faktorje.

6.11 Prednosti in pomanjkljivosti latinskih in grško-latinskih kvadratov.

Glavne prednosti latinskih in grško-latinskih kvadratov so v tem, da moremo v njih z razmeroma majhnim obsegom poskusa izločiti tudi do tri določljive, a za raziskavo nepomembnih faktorjev, kar vpliva na zanesljivost poskusnih rezultatov. Obratno pa moremo z obsegom poskusa A^2 enot proučiti do štiri faktorje z po A vrednostmi. Razen razmeroma majhnega obsega je tudi analiza enostavna.

Glavne pomanjkljivosti pa so v tem, da mora biti število vrednosti faktorjev za vse faktorje enako. Glede na to so latinski in grško-latinski kvadrati večji kot 8 x 8 redki. Pri latinskih in grško-latinskih kvadratih, ki so manjši kot 5 x 5 odpade malo število stopinj prostosti na poskusni pogrešek.

Število stopinj prostosti na enostavne latinske in grško-latinske kvadrate:

	2x2	3x3	4x4	5x5	6x6	7x7	8x8
LK	0	2	6	12	20	30	42
GLK	-	0	3	8	15	24	35

Pogoj, da so faktorji neodvisni oziroma, da so učinki aditivni, je precejšnja omejitev za uporabo LK in GLK.

Slučajnostni razpored v LK posebno pa v GLK je razmeroma kompleksen.

Tabela 6.1 Primerjala formiranja komponent

$\sum Y_A^2 2$						Q_A^2 $n = A^2$	$n = A^2$
$\sum (Y_A^2 2 - \bar{Y})^2$					Q $m_a = 1$	$K = q_A^2$ $m = (A^2-1)(A-1)$	K $m = A^2 - 1$
čisto slučajnostni poskus	$K_A = q_A$ $m_A = A - 1$					$K_e = q_A 2 - q_A$ $m_e = (A-1)\ A$	α K_A $(A-1)$ K_e $m_e = (A-1)A$
slučajnostni popolni bloki	$K_A = q_A$ $m_A = A - 1$	$K_B = q_B$ $m_B = A - 1$				$K_e = q_{AB} - q_A - q_B$ $m_e = (A-1)(A-1)$	0_1 K_A $(A-1)$ $K_B (A-1)$ m_e K_e $(A - 1)$
latinski kvadrat	$K_A = q_A$ $m_A = A - 1$	$K_V = q_V$ $m_V = A - 1$				$K_e = q_A 2 - q_A -$ $q_V - q_s$ $m_e = (A-1)(A-2)$	1 $K_V (A - 1)$ K_A $K_s (A - 1)$ $(A-1)$ K_e $m_e =$ $(A-1)(A-2)$
grško-latinski kvadrat	$K_A = q_A$ $m_A = A - 1$	$K_V = q_V$ $m_V = A - 1$	$K_s = q_s$ $m_s = A - 1$	$K_\alpha = q_\alpha$ $m_\alpha = A - 1$		$K_e = q_A 2 - q_A$ $- q_V - q_s - q_\alpha$ $m_e = (A-1)(A-3)$	1 $Kv\ (A-1)$ K_A $K_s\ (A-1)$ $(A-1)$ $K_\alpha\ (A-1)$ $Ke\ me =$ $(A-1)(A-3)$

Iz sheme je razvidno, kako se skupna vsota kvadratov od načrta do načrta vse bolj razstavlja in kako se na ta način manjšata poskusni pogrešek (K_e in m_e)

6.12 Kot primerjalni povzetek formiranja komponent v

a) čisto slučajnostnem poskusu A postopki v A ponovitvah
b) slučajnostnih blokih A postopkov v A blokih
c) latinskem kvadratu A X A
d) grško-latinskem kvadratu A x A

vzemimo poskus z A^2 poskusnimi enotami.

Tabela 6.1 Primerjava formiranja komponent

Iz sheme je razvidno, kako se skupna vsota kvadratov od načrta do načrta vse bolj razstavlja in kako se na ta način manjšata poskusni pogrešek (K_e in m_e).

6.13 Latinski kvadrati v blokih.

Poskus v latinskem kvadratu ima za standardni pogrešek $m_e = (A-1)(A-2)$ stopinj prostosti. Tako je pri $A = 2$ $m_e = 0$, pri $A=3$ $m_e = 2$, pri $A = 4$ je $m_e = 6$, pri $A = 5$ $m_e = 12$. Od skupnega števila stopinj prostosti $m = A^2 - 1$ odpade razmeroma malo na standardni pogrešek. Zato pogosto poskuse v latinskih kvadratih ponavljamo, da pridobimo na številu stopinj prostosti za standardni pogrešek A. V tem primeru načrtujemo ponovitve tako, da predstavlja vsak LK v celoti samostojen blok. Tako dosežemo, da iz poskusa izločimo heterogenost poskusnega gradiva, ki mora pri b ponovitvah oziroma blokih imeti $n = bA^2$ enot. Tako moremo latinski kvadrat, s katerim proučujemo več vrst gnojila, ponoviti na več njivah - blokih. V tem primeru predstavlja vsaka njiva blok - razlike med njivami - bloki pa s takim poskusom izločamo.

Shema latinskega kvadrata v blokih je naslednja:

blok 1

	s_1	s_2	s_3
v_1	A_1	A_2	A_3 .....
v_2	A_2	A_3	A_1 .....
v_3	A_3	A_1	A_2 .....

blok 2

	s_4	s_5	s_6
v_4	A_3	A_1	A_2
v_5	A_1	A_2	A_3
v_6	A_2	A_3	A_1

(6.19)

model latinskega kvadrata v blokih - oziroma ponovitvah pa je:

$$y_{bvsA} = M+(b)+(bv)+(bs)+(A)+e_{bvsA} \qquad e_{bvsA} = :N(0, \sigma^2_e) \qquad (6.20)$$

Iz sheme in modela je razvidno, da tako vrstice kot stolpci nimajo med bloki nobene zveze oziroma skupne značilnosti. Če npr. faktor v predstavljajo 4 trgovine, so v posameznih blokih po štiri različne trgovine. Stolpci oziroma vrstice predstavljajo v okviru glavnega bloka ppodbloke, zato je zanje komponenta (bv) in (bs), ne pa (v) in (s).

Iz osnovnih podatkov poskusa v latinskih kvadratih v blokih y_{bvsA} izračunamo naslednje vrste oziroma tabele.

$$y_{bvsA} \quad y_{bv} \quad y_A$$

$$y_{bs} \quad y_b \quad y \tag{6.21}$$

in pomožne količine

$$Q_{bvsA} = \sum_b \sum_v \sum_s \sum_A y^2_{bvsA} \qquad Q_{bv} = \frac{1}{A} \sum_b \sum_v y^2_{bv} \qquad Q_A = \frac{1}{bA} \sum_A y^2_A \tag{6.22}$$

$$Q_{bs} = \frac{1}{A} \sum_b \sum_s y^2_{bs} \qquad Q_b = \frac{1}{A^2} \sum_b y^2_b \qquad Q = y^2/bA^2$$

ustrezne q_v izračunamo, da od Q_v odštejemo Q.

Analiza variance LK v blokih pa je:

(6.23)

VV	K	m	s^2	$E(s^2)$
(b)	q_b	b-1	s^2_b	$\sigma^2_e+A\sigma^2_{bv}+A\sigma^2_{bs}+A^2\sigma^2_b$
(bv)	$q_{bv}-q_b$	b(A-1)	s^2_{bv}	$\sigma^2_e + A\sigma^2_{bv}$
(bs)	$q_{bs}-q_b$	b(A-1)	s^2_{bs}	$\sigma^2_e + A\sigma^2_{bA}$
(A)	q_A	(A-1)	s^2_A	$\sigma^2_e + bA\sigma^2_A$
pp	$q_{bvsA}-q_{bv}-q_{bs}-q_A+q_b$	(A-1)[b(A-1)-1]	s^2_e	σ^2_e
Sk	q_{bvsA}	bA^2-1		

Iz sheme za analizo variance povzamemo, da moremo značilnost za (bv), (bs) in (A) z F-preskusom ugotoviti tako, da ustrezen s^2 delimo z s_e^2. To pa ni slučaj za medbločno komponento (b), ker ni direktno dana ocena za $\sigma_e^2 + A\sigma_{bv}^2 + A\sigma_{bs}^2$. Ker pa analiziranje značilnosti za b v tem primeru ni vprašanje, ta problem, ki je v drugih primerih pereč, ne bomo proučili na tem mestu. Iz analize variance spoznamo, da vsak nadaljnji blok doprinese skupno A^2 stopinj prostosti, od teh pa gre na povečanje m_e $(A-1)^2$.

Ocena manjkajočega podatka v latinskem kvadratu v blokih.

Oceno y'_{bvsA} za podatek v bloku b, vrstici bv in stolpcu bs, v katerem je po planu apliciran postopek A, je po metodi najmanjših kvadratov za K_e naslednje:

$$y'_{bvsA} = \frac{bAy'_{bv} + bAy'_{bs} + Ay'_{A} - by'_{b} - y'}{(A-1)[b(A-1) - 1]} \qquad (6.24)$$

Podobno kot pri slučajnostnih blokih oziroma latinskem kvadratu brez ponovitev, se m_e zmanjša za 1 oziroma za toliko, kolikor je ocenjenih podatkov. Ocena za več manjkajočih podatkov se znatno zaplete. Razen tega pa taki poskusi izgubljajo na pomenu.

6.14 P o n a v l j a n j e m e r i t e v v l a t i n s k e m k v a d r a t u .

V primeru, da gre za ponavljanje meritev na poskusnih enotah, se poskusni pogrešek razdeli v dve komponenti: pravi poskusni pogrešek e_{bvsA} in vzorčni pogrešek e_{bvsAi} .

Model latinskih kvadratov v blokih z $\bar{n}$ ponovitvami na posameznih poskusnih enotah je naslednji

$$y_{bvsAi} = M + (b) + (bs) + (bv) + (A) + e_{bvsA} + e_{bvsAi} \tag{6.25}$$

analiza variance pa naslednja:

(6.26)

VV	K	m	s^2	$E(s^2)$
(b)	q_b	b-1	s_b^2	$\sigma_i^2 + \bar{n}\sigma_e^2 + \bar{n}A\sigma_{bv}^2 + \bar{n}A\sigma_{bs}^2 + \bar{n}A^2\sigma_b^2$
(bv)	$q_{bv}-q_b$	b(A-1)	s_{bv}^2	$\sigma_i^2 + \bar{n}\sigma_e^2 + A\sigma_{bv}^2$
(bs)	$q_{bs}-q_b$	b(A-1)	s_{bs}^2	$\sigma_i^2 + \bar{n}\sigma_e^2 + \bar{n}A\sigma_{bs}^2$
(A)	q_A	(A-1)	s_A^2	$\sigma_i^2 + \bar{n}\sigma_e^2 + \bar{n}bA\sigma_A^2$
pp	$q_{bsvA}-q_{bv}-q_{bs}-q_A+q_b$	(A-1)[b(A-1)-1]	s_e^2	$\sigma_i^2 + \bar{n}\sigma_e^2$
vp	$q_{bvsAi} - q_{bvsA}$	(i-1)(b.A)	s_i^2	σ_i^2
Sk	q_{bvsAi}	bA^2i-1		

Pomožne količine q so izračunane po znani splošni shemi

$q_u = \frac{u}{n}\sum y_u^2 - \frac{1}{n}y^2$, pri čemer je v našem primeru skupno število vseh meritev $n = bA^2\,\bar{n}$.

Tako je npr.

$$q_{bv} = \frac{bA}{bA^2\bar{n}}\sum y_{bv}^2 - \frac{y^2}{bA^2\bar{n}}$$

S pomočjo $E(s^2)$ moremo oceniti tudi posamezne komponente variacije.

Kot primer poskusa moremo vzeti gnojilni poskus na več njivah (bloki) na katerih izvedemo poskus v latinskih kvadratih (da izločimo spreminjanje kvalitete zemlje v dveh pravokotnih smereh v,s) na vsaki ploskvi latinskega kvadrata pa merimo donos na več različnih rastlinah.

6.15 Kot je razvidno iz $E(s^2)$, v shemi za analizo variance, komponente (bv) in (A) z F-preskusom preskušamo tako, da s^2_{bv} s^2_{bs} in s^2_A primerjamo s s^2_e, medtem ko značilnost variance čistega poskusnega pogreška preskušamo z razmerjem

$$F_e = \frac{s_e^2}{s_i^2}$$

Model in poskuse z latinskim kvadratom v blokih z več meritvami na posameznih poskusnih enotah in ustrezno shemo za analizo variance moremo uporabiti za štiri različne načrte z latinskimi kvadrati

	b	A	$\bar{n}$
Navaden LK	1	A	1
LK v blokih	b	A	1
LK z $\bar{n}$ meritvami na poskusnih enotah	1	A	$\bar{n}$
LK v blokih z $\bar{n}$ meritvami na poskusnih enotah	b	A	$\bar{n}$

Če ni blokov ali več meritev vzamemo za b oziroma $\bar{n}$ kot faktor 1, b oziroma $\bar{n}$ kot indeks pa odpade.

6.16 S p o j e n i   l a t i n s k i   k v a d r a t i.

Pri latinskih kvadratih v blokih se z vsakim blokom m_e poveča za $(A-1)^2$. Če je $A = 2$ se m poveča za 1, za $A = 3$ za 4, za $A = 4$ za 9. To povečanje se sicer z večanjem A veča s kvadratom, a je za poskuse z majhnim številom postopkov (A=2, A=3) razmeroma majhno; m_e za LK v blokih razmeroma hitreje povečamo s spojenimi latinskimi kvadrati. Faktor v je včasih take narave, da so njegove vrednosti enake v vseh blokih. Če je faktor v dan v tednu, je jasno, da faktor (bv) preide v faktor (v), ker je neodvisen od bloka, ker so v vseh blokih isti dnevi. Tako je tudi v primeru poskusa z avtomobilskimi gumami, ki je nakazan v primeru. Ta poskus bi mogli izvesti v blokih tako, da bi blok predstavljali po štirje avtomobili iste znamke, na teh avtomobilih pa bi po načrtu, ki je dan z latinskim kvadratom, montirali štiri različne vrste gum. Ker ima vsak avto, neglede na znamko, štiri točno podane pozicije koles (SL, SD, ZL, ZD), so te pozicije skozi vse bloke iste. Zato v analizi variance ta faktor (v) sodeluje le z $m_v = A-1$ stopinjami prostosti, a ne z $m_{bv} = b(A-1)$ stopinjami prostosti in preide $m_{bv} - m_v = (b-1)(A-1)$ stopinjo prostosti v m_e. M_e se torej poveča na

$$m_e = (A-1)\,|\,b(A-1)-1\,| + (b-1)(A-1) = (A-1)(bA-2) \qquad (6.27($$

S povečanjem števila blokov za 1 se m_e poveča za A(A-1).

Shema, model in AVAR za poskus s spojenimi latinskimi kvadrati je:

	s_{11}	s_{12}	s_{13}	s_{21}	s_{22}	s_{23}
V_1	A_1	A_2	A_3	A_3	A_1	A_2
V_2	A_2	A_3	A_1	A_2	A_3	A_1
V_3	A_3	A_1	A_2	A_1	A_2	A_3

(6.28)

$$y_{bVsA} = M + (b) + (V) + (bs) + (A) + e_{bVsA}$$

$$e_{bVsA} = : N(0, \sigma_e^2) \qquad (6.29)$$

a analiza variance

VV	K	m	s^2	$E(s^2)$ (6.30)
(b)	q_b	b-1	s_b^2	$\sigma_e^2 + A\sigma_s^2 + A^2\sigma_s^2$
(V)	q_V	A-1	s_v^2	$\sigma_e^2 + bA\sigma_v^2$
(bs)	$q_{bs}-q_b$	b(A-1)	s_A^2	$\sigma_e^2 + A\sigma_s^2$
A	q_A	(A-1)	s_A^2	$\sigma_e^2 + bA\sigma_A^2$
pp	$q_{bVsA}-q_{bs}-q_v-q_A$	(A-1)(bA-2)	s_e^2	σ_e^2
Sk	q_{bVsA}	bA^2-1		

Po značaju je V običajno fiksen faktor, zato je kot tak vzet tudi v zgornji shemi.

6.17 V našem primeru je kot načelo za sestavljanje blokov vzeta znamka avtomobila. Isti poskus pa bi mogli izvesti tudi na bA avtomobilih iste znamke. V tem primeru bi kriterij za sestavljanje blokov odpadel in bi dobili glede na avtomobile čisto slučajnostno ponavljanje poskusa. Takih primerov je zelo veliko in so taki načrti poskusov zelo pogosti. V tem primeru meje blokov odpadejo in je važno le, da je vsak A v vsaki vrsti b kratzastopan in da so v stolpcih vsi postopki.

Shema, model in AVAR za tak poskus (Change-over poskus) je:

	s_1	s_2	s_3	s_4	s_5	s_6
V_1	A_1	A_1	A_3	A_2	A_3	A_2
V_2	A_2	A_3	A_1	A_3	A_2	A_1
V_3	A_3	A_2	A_2	A_1	A_1	A_3

(6.31)

$$y_{VsA} = M + (V) + (s) + (A) + e_{VsA} \qquad e_{VsA} = : N(0, \sigma_e^2) \quad (6.32)$$

Analiza variance je: $\qquad s = bA$

VV	K	m	s^2	(6.33)
V	q_V	A-1	s_V^2	
s	q_s	s-1	s_s^2	
A	q_A	A-1	s_A^2	$F_A = s_A^2/s_e^2$
pp	$q_{VsA}-q_V-q_s-q_A$	(A-1)(s-2)	s_e^2	1
	q_{VsA}	$s_A - 1$		

Po svoji obliki je celoten načrt sličen poskusu z enostavnim latinskim kvadratom (v njega tudi preide, če je b=1 in je s = bA = A). V drugih primerih pa je to posebna oblika spojenega latinskega kvadrata.

Za poskuse, v katerih je A=2 je tak način edino možen, ker je $m_e = 0$ za enostaven latinski kvadrat in tudi za latinske kvadrate v blokih. V primeru tako spojenega latinskega kvadrata je za A=2 na splošno $m_e = (A-1)(s-2) = s-2 = 2(b-1)$.

6.18 Za primer poskusa z opuščenimi spojenimi bloki v latinskem kvadratu vzemimo naslednji poskus. Da bi proučili, ali način prodaje (2 načina prodaje) značilno vplivata na prodajo, so v osmih s = b . A = 4 . 2 = 8 slučajnostno izbranih trgovinah prodajali poskusno blago po obeh načinih. Trgovina v tem primeru predstavlja blok. Da bi izločili vpliv mesta, na katerem prodajajo blago, posamezen način prodaje ni bil določen posameznemu mestu (pri pultu, pri vratih, pri pultu v notranjosti trgovine) čisto slučajnostno, temveč v skladu z načelom opuščenih spojenih latinskih kvadratov tako, da je vsak način prodaje bil apliciran na enakem mestu b krat (v našem primeru štirikrat).

y_{VSA} Mesto	trgovina s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	y_V
V_1	A_1 71	A_2 66	A_1 49	A_2 59	A_1 50	A_1 50	A_2 48	A_2 59	452
V_2	A_2 50	A_2 42	A_1 49	A_1 50	A_2 49	A_2 43	A_1 31	A_1 48	362
y_S	121	108	98	109	99	93	79	107	814=y

$y_{A1} = 391 \qquad y_{A2} = 423$

$q_{VSA} = \sum_{VSA} y^2_{VSA} = 71^2 + \ldots + 48^2 = 42744 \qquad q_{VSA} = 1321,75$

$Q_S = \frac{1}{2}\left[121 + \ldots + 107\right] = 41965 \qquad q_S = 542,75$

$Q_V = \frac{1}{8}\left[452^2 + 362^2\right] = 41918,5 \qquad q_V = 496,25$

$Q_A = \frac{1}{8}\left[391^2 + 423^2\right] = 41476,25 \qquad q_A = 54,00$

$Q = \frac{1}{16}\, 814^2 = 41422,25$

Analiza variance je:

VV	K	m	s^2	
Mesto (V)	q_V 496,25	1	496,25	-
Trgovina (S)	q_S = 542,75	7	77,54	-
Način prodaje (A)	q_A = 54,00	1	54,00	1,416
pp	$q_{VSA} - q_V - q_S - q_A$			
Sk	228,75	6	38,125	1,000

Ker je $F_A = 1,416 < F_{0.10}(1,6) = 3,78$, sklepamo, da način prodaje ne vpliva značilno na višino prodaje.

7. FAKTORSKI POSKUSI

7.1 Po ustaljeni terminologiji govorimo o večfaktorskih poskusih takrat, kadar s poskusom kompleksno proučujemo dva ali več za raziskavo pomembnih faktorjev. Ta terminologija ne ustreza povsem, ker je tudi poskus, v katerem proučujemo en faktor, faktorski poskus. Tako je faktorski poskus sinonim za večfaktorski poskus.

Eden izmed bistvenih prvin in značilnosti statističnega proučevanja poskusnih podatkov je v tem, da se je proučevanje enega samega faktorja s klasičnim poskusom razširilo na večfaktorski poskus, v katerem z istim poskusom obseženo analizo, vpliva več faktorjev hkrati. Kot bomo videli, trofaktorski poskus ni istoveten s tremi enofaktorskimi poskusi in dá trofaktorski poskus kompleksnejše rezultate o vplivu faktorjev. Tako npr. s trofaktorskim gnojilnim poskusom analiziramo in ugotovimo, katera kombinacija količine treh gnojil daje najboljši pridelek. Podobno z večfaktorskim poskusom dodatkov določeni brezalkoholni pijači ugotovimo kompleksen učinek mešanice na okus pijače. Ta rezultat ni v vseh primerih vsota za ocene dodatkov. Enako je s kompleksnim učinkom različnih proizvodnih dejavnikov, ki vplivajo na kakovost izdelkov in podobno. Zato so včasih večfaktorski poskus imenovali tudi kompleksni poskus.

7.2 Glavni učinek. Interakcija.

Vzemimo za primer dva faktorja P s tremi vrednostmi in B z dvema vrednostima. S kombiniranjem vrednosti teh dveh faktorjev dobimo A=6 postopkov.

Učinek vsakega postopka (obenem z opredeljujočimi pogoji je dan z aritmetično sredino za kriterialni znak M_{AB}. Tako imamo toliko sredin, kolikor je postopkov oziroma kombinacij vrednosti faktorjev.

$$\begin{array}{c|ccc|c}
M_{PR} & & & & M_R \\
\hline
 & M_{11} & M_{12} & M_{13} & M_{1\cdot} \\
 & M_{21} & M_{22} & M_{23} & M_{2\cdot} \\
\hline
M_P & M_{\cdot 1} & M_{\cdot 2} & M_{\cdot 3} & M
\end{array} \qquad (7.1)$$

M_{PR} so poprečja postopkov in izražajo vpliv opredeljujočih faktorjev in vpliv faktorjev P in R.

Aritmetične sredine $M_P = \frac{1}{R} \sum_R M_{PR}$ so sredine, ki variirajo samo zaradi vpliva p in $M_R = \frac{1}{P} \sum_P M_{PR}$ sredine, ki variirajo zaradi vpliva faktorja R.

$M = \frac{1}{P.R} \sum\sum_{PR} M_{PR}$ pa je skupna sredina, ki je izraz splošnih faktorjev.

Tako je torej

$$\begin{aligned}
M &= M \\
M_P &= M + (P) \\
M_R &= M + (R) \qquad\qquad (7.2) \\
M_{PR} &= M + (P) + (R) + (PR)
\end{aligned}$$

M = M rezultat opredeljujočih pogojev

M_P = vsota rezultatov opredeljujočih pogojev M in "glavnega učinka" faktorja (P)

M_R = vsota opredeljujočih pogojev M in glavnega učinka faktorja (R).

Poprečja M_{PR} so rezultat kompleksnega učinka opredeljujočih faktorjev M in faktorjev P in R.

Vendar so poprečja $M_{PR} = M + (P) + (R)$ vsota rezultatov opredeljujočih pogojev M in glavnih učinkov (P) in (R) le v primeru, če je vpliv obeh faktorjev neodvisen in so učinki aditivni. V splošnem pa je kompleksen učinek faktorjev med seboj odvisen. Vpliv enega faktorja na kriterialni znak je drugačen pri različnih nivojih drugega faktorja. Vpliv različnih doz gnojila R je bistveno odvisen od doze gnojila P. To se pokaže v tem, da v M_{PR} razen od M in glavnih učinkov (P) in (R) nastopa še dodatna komponenta (PR), ki je izraz vzajemnega delovanja faktorjev P in R in ga imenujemo <u>interakcijo</u> med faktorjema P in R.

Iz gornjih zvez moremo izvrednotiti iz M, M_P, M_R, M_{PR} glavna učinka obeh faktorjev (P) in (R) in interakcijo (PR). Tako zlahka dobimo, da je

$$
\begin{aligned}
(P) &= M_P - M \\
(R) &= M_R - M \\
(PR) &= M_{PR} - M_P - M_R + M
\end{aligned}
\qquad (7.3)
$$

Iz znanih zvez med M, M_P , M_R in M_{PR} dobimo, da je

$$
\begin{aligned}
\sum_P (P) &= \sum_P (M_P - M) = 0 \\
\sum_R (R) &= \sum_R (M_R - M) = 0
\end{aligned}
\qquad (7.4)
$$

$$
\begin{aligned}
\sum_P (PR) &= \sum_P \left[(M_{PR} - M_R) - (M_P - M)\right] = 0 \\
\sum_R (PR) &= \sum_R \left[(M_{PR} - M_P) - (M_R - M)\right] = 0
\end{aligned}
\qquad (7.5)
$$

Vsota ali poprečje interakcije po katerikoli vrednosti faktorja P ali R je enaka 0 po definiciji.

7.3 Vzemimo shematične vrednosti za M_{PR} za primer 2 x 3 poskus brezalkoholne pijače.

Faktor S je količina sladkorja v treh dozah:

S_1 = nizka koncentracija; S_2 = srednja koncentracija;
S_3 = visoka koncentracija

faktor D pa dodatek z dvema vrednostima: D_o = brez dodatka
D_1 = z dodatkom.

S kombinacijo zgornjih dveh faktorjev dobimo A = 6 različnih pijač.

Shematične vrednosti poprečne ocene za okus šestih pijač M_{DS} za zgornji primer so:

M_{DS}	S_1	S_2	S_3	M_D
D_o	32	58	72	54
D_1	40	92	60	64
M_S	36	75	66	59 M

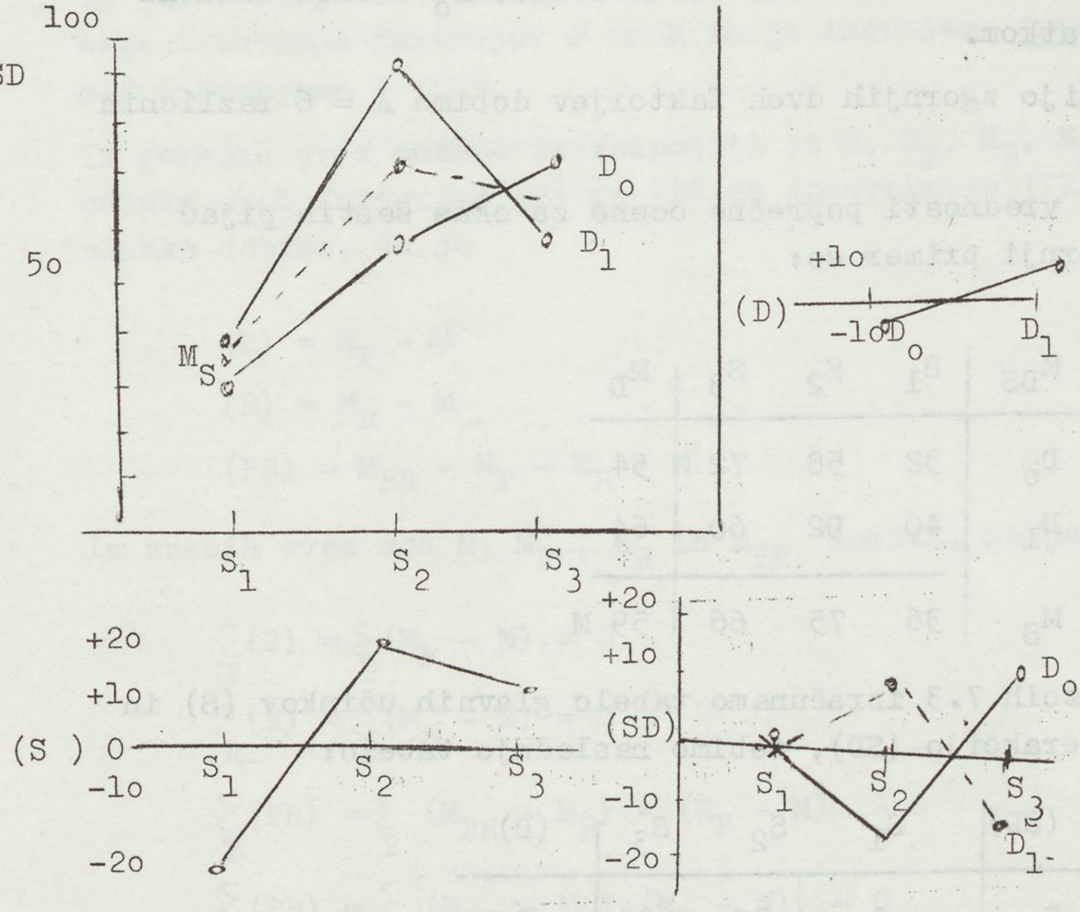
Če po obrazcih 7.3 izračunamo tabelo glavnih učinkov (S) in (D) in interakcijo (SD), dobimo naslednjo tabelo:

(SD)	S_1	S_2	S_3	(D)
D_o	+ 1	- 12	+ 11	- 5
D_1	- 1	+ 12	- 11	+ 5
(S)	- 23	+ 16	+ 7	59 = M

Če za primer M_{21} = 92 kontroliramo dane zveze, dobimo

$M_{21} = M + S_2 + D_1 + SD_{21} = 59 + (+16) + (+5) + (+12) = 92$

Za zgornji primer vrednosti ocen za okus šestih pijač, ki jih dobimo s kombiniranjem faktorjev sladkor in dodatek, dobimo naslednjo sliko



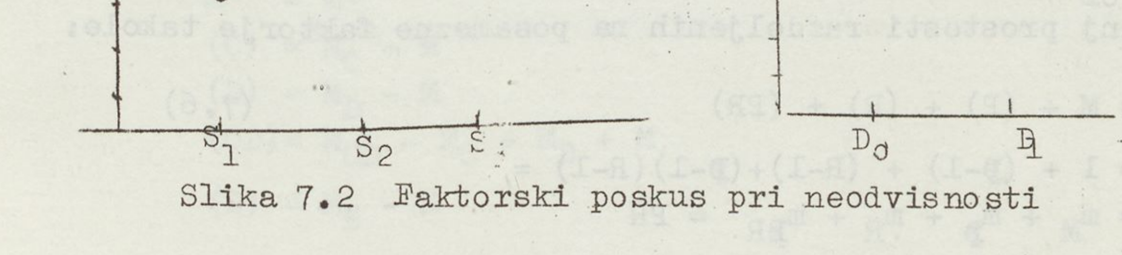
Slika 7.1 Komponente faktorskega poskusa 2 x 3

Analiza grafikona poprečnih ocen za šest kombiniranih pijač pokaže, da je najvišjo oceno dosegla kombinacija: srednja koncentracija sladkorja z dodanim dodatkom ($M_{21} = 92$).

Iz grafikona za M_{SD} vidimo neposredno, da v učinku faktorjev S D obstaja interakcija. To sklepamo iz tega, ker je potek krivulje za oceno pri različnih koncentracijah sladkorja drugačen, če je dodan ali ni dodan dodatek D. Če pri istih glavnih učinkih (S) in (D) in sredin M sestavimo vrednosti M_{SD} brez interakcije, dobimo naslednjo tabelo za M_{SD}

M_{SD}	S_1	S_2	S_3	M_D
D_o	31	70	61	54
D_1	41	80	71	61
M_S	36	75	66	59

Iz tabele vidimo, da so M_D in M_S ostali zaradi nakazanih lastnosti interakcije enaki kot v osnovnem primeru, v katerem je bila tudi interakcija. Če narišemo grafikon za M_{SD} za primer, da interakcije ni, dobimo naslednjo značilno sliko:



Slika 7.2 Faktorski poskus pri neodvisnosti

Če interakcije ni, je zveza med učinkoma obeh faktorjev aditivna $M_{SD} = M + (S) + (D)$, črte za M_{SD} enega faktorja pri različnih vrednostih drugega faktorja pa vzporedne.

Če kritično premotrimo nakazani primer, je na splošno zelo pomemben naslednji sklep. Če interakcije ni, poprečna črta ... vedno podaja zakonitost poteka vpliva faktorja pri različnih vrednostih drugega faktorja, samo da je na poprečnem nivoju, razlike učinkov pa ostanejo iste. V tem primeru je poprečno ponašanje faktorja (S) MS slika za ponašanje faktorja pri posameznih nivojih drugega faktorja M_{SDo} ali M_{SD1} in

$M_{SD} = M_S + M_D - M$

Če pa je interakcija obstaja, poprečje M_S ne kaže poteka niti pri D_o niti pri D_1. Tako je ugotovitev, da $(SD) \neq 0$ znak, da je treba proučiti celo tabelo poprečij M_{SD}, če hočemo ugotoviti zakonitosti vpliva dveh faktorjev, ker samo z M_S in M_D ne moremo pojasniti M_{SD} . Zato so v tem primeru M_S in M_D drugotnega pomena.

7.4 Ker je glavni učinek (P) vezan na pogoj (P) = 0, je število stopinj prostosti za glavni učinek $m_p = P - 1$.

V tabeli interakcije so interakcije vezane na pogoj, da so vrste po vrsticah in stolpcih enake 0. Zato je tabela interakcij določena že s tabelo vrednosti interakcij, ki ima en stolpec in eno vrstico manj torej z (P-1)(R-1) vrednostmi, vrednosti zadnje vrstice in zadnjega stolpca pa so določene z robnimi pogoji. Zato je število stopinj prostosti za interakcijo $m_{CD} = (C-1)(R-1)$. V celoti je v dvofaktorskem poskusu PR stopinj prostosti razdeljenih na posamezne faktorje takole:

$$M_{PR} = M + (P) + (R) + (PR) \qquad (7.6)$$

$$PR = 1 + (P-1) + (R-1)+(P-1)(R-1) =$$

$$m = m_M + m_P + m_R + m_{PR} = PR$$

7.5 Glavni učinki in interakcije pri trofaktorskem poskusu.

Če razširimo sklepanje iz dvofaktorskega poskusa na trofaktorski poskus, dobimo iz tabele popreči j M_{SDE} za $n = CDE$ postopkov in tabele sumarnih popreči j

$$M_{CDE} \quad M_{CD} \quad M_{CE} \quad M_{C} \qquad (7.7)$$

$$M_{DE} \quad M_{D} \quad M_{E} \quad M$$

naslednje zveze

$$
\begin{aligned}
M &= M \\
M_C &= M + (C) \\
M_D &= M + (D) \\
M_{CD} &= M + (C) + (D) + (CD) \qquad (7.8) \\
M_E &= M + (E) \\
M_{CE} &= M + (C) + (E) + (CE) \\
M_{DE} &= M + (D) + (E) + (DE) \\
M_{CDE} &= M + (C) + (D) + (CD) + (E) + (CE) + (DE) + (CDE)
\end{aligned}
$$

M_{CDE} je torej rezultat M, glavnih učinkov vseh faktorjev (C), (D), (E) vseh dvojnih interakcij (CD), (CE) in (DE) in trojne interakcije (CDE). To je izraz kompleksnega delovanja vseh treh faktorjev hkrati.

Iz zgornjih enačb dobimo:

$$
\begin{aligned}
M &= M \\
(C) &= M_C - M \\
(D) &= M_D - M \qquad (7.9) \\
(CD) &= M_{CD} - M_C - M_D + M \\
(E) &= M_E - M
\end{aligned}
$$

$$(CE) = M_{CE} - M_C - M_E + M$$
$$(DE) = M_{DE} - M_D - M_E + M$$
$$(CDE) = M_{CDE} - M_{CD} - M_{CE} - M_{DE} + M_C + M_D + M_E - M$$

V praktičnih primerih more iz modela M_{CDE} katera komponenta izpasti, če so učinki faktorjev neodvisni in in kompleksnih učinkov. Pri popolni neodvisnosti dobimo

$$M_{CDE} = M + (C) + (D) + (E) \qquad (7.10)$$

Če za kontrolo po obrazcih za ta model izračunamo iz sumarnih poprečij interakcije, se izkaže, da so enake nič.

Število stopinj prostosti za glavne učinke in za dvojne interakcije poznamo. Z enakim sklepanjem kot zanje dobimo, da je število stopinj prostosti za trojno interakcijo $m_{(CDE)} = (C-1)(D-1)(E-1)$

ali skupno:

$$CDE = m_M + m_C + m_D + m_{CD} + m_E + m_{CE} + m_{DE} + m_{CDE} \qquad)7.11)$$

$$1 + (C-1)+(D-1)+(C-1)(D-1)+(E-1)+(C-1)(E-1)+(D-1)(E-1)+ + (C-1)(D-1)(E-1) = CDE$$

S podobnim sklepanjem razširimo sistematiko na faktorske poskuse z več kot tremi faktorji.

Interakcija druge stopnje je razmeroma lahko predstavljiva. Čim višja pa je stopnja interakcije, tem težje je njeno praktično tolmačenje. Ta nepredstavljivost včasih doseže tako stopnjo, da interakcije bodisi vključimo v poskusni pogrešek kot slučajnostno komponento ali celo interakcije višjih stopenj uporabljamo samostojno kot poskusni pogrešek.

7.6 D v o f a k t o r s k i   p o s k u s i .

Dvofaktorski je vsak poskus z dvema pomembnima faktorjema. V zvezi z dvofaktorskim poskusom kot tudi v zvezi z večfaktorskimi poskusi uporabljamo vse dosedaj nakazane metode čisto slučajnostnega poskusa, bloke, latinske kvadrate ipd. kot tudi še niz drugih postopkov, ki izboljšajo sklepe, ki jih moremo narediti iz poskusa.

Prva naloga večfaktorskega poskusa je, da iz poskusnih podatkov z analizo variance ugotovimo, katere od komponent (glavni učinki, interakcije) so značilni, da na osnovi teh rezultatov modificiramo model poskusa in nadalje s primerjavami ali z analizo poprečij analiziramo poskusne podatke.

Na splošno je vsota kvadratov postopkov $K_A \equiv cd = q_{cd}$. Vsota kvadratov odklonov za glavne učinke c in d $K_c = q_c$ in $K_d = q_d$. Ker je $K_{A = cd}$ rezultat glavnih učinkov in interakcije, gre na račun interakcije

$$K_{cd} = K_A - K_c - K_d = q_{cd} - q_c - q_d \qquad)7.12)$$

Zveza

$$K_A = K_c + K_d + K_{cd}$$

$$q_{cd} = q_c + q_d + (q_{cd} - q_c - q_d) \qquad (7.13)$$

$$m_A = m_c + m_d + m_{cd}$$

$$cd-1 = (c-1) + (d-1) + (c-1)(d-1)$$

so splošne zveze med vsotami kvadratov in med številom stopinj prostosti posameznih komponent učinkov v dvofaktorskem poskusu.

Modeli in analiza variance dvofaktorskega poskusa.

7.7 Č i s t o s l u č a j n o s t n i d v o f a k t o r - s k i p o s k u s .

Če vzamemo, da vršimo dvofaktorski poskus kot čisto slučajnostni poskus, zgoraj nakazano razdelitev vgradimo v model in analizo variance za čisto slučajnostni poskus. Tako dobimo:

$$y_{CDi} = M + (C) + (D) + (CD) + e_{CDi}$$

$$K = K_C + K_D + K_{CD} + K_e$$

$$Q_{CDi} = Q + q_C + q_D + (q_{CD} - q_C - q_D) + (q_{CDi} - q_{CD}) \quad (7.14)$$

$$\bar{n} = CD\bar{n} = 1 + (C-1) + (D-1) + (C-1)(D-1) + CD(\bar{n}-1)$$

Vse štiri enačbe so identitete zveze med posameznimi komponentami pa dobimo, če jih čitamo v navpični smeri.

Zgornje zveze strnemo v standardno obliko analize variance:

VV	K	m	s^2	E	$E(s^2)$
(C)	$K_C = q_C$	$m_C = (C-1)$	s_C^2	$F_C = s_C^2/s_e^2$	$\sigma_e^2 + \bar{n}d\,\sigma_C^2$
(D)	$K_D = q_D$	$m_D = (D-1)$	s_D^2	$F_D = s_D^2/s_e^2$	$\sigma_e^2 + \bar{n}C\,\sigma_D^2$
(CD)	$K_{CD} = q_{CD} - q_C - q_D$	$m_{CD} = (C-1)(D-1)$	s_{CD}^2	$F_{CD} = s_{CD}^2/s_e^2$	$\sigma_e^2 + \bar{n}\,\sigma_{CD}^2$
pp	$K_e = q_{CDi} - q_{CD}$	$m_e = CD(\bar{n}-1)$	s_e^2	1	σ_e^2
Sk	$K = q_{CDi}$	$m = CD\bar{n} - 1$			

7.8 Za primer dvofaktorskega poskusa v čisto slučajnostnem poskusu vzemimo raziskavo gostote briketov iz določene surovine v odvisnosti od pritiska in temperature pri proizvodnji. Temperatura pri proizvodnji je bila na treh nivojih ($T_1 = 1900^o$, $T_2 = 2100^o$, $T_3 = 2300^o$) pritisk kot drug faktor na treh nivojih ($P_1 = 5.0$ at, $P_2 = 12.5$ at, $P_3 = 20.0$ at).

Vsak izmed $A = T.P = 3.3 = 9$ postopkov, ki obstojajo iz vseh možnih kombinacij faktorjev T in P, je bil izveden v $\bar{n} = 2$ ponovitvah. Poskusnemu gradivu (surovini) so bile v $n = 18$ poskusih dodeljeni po dve ponovitvi vsakega postopka. Osnovni rezultati poskusa, urejeni v tabeli, so naslednji:

y_{PTi}	Temperatura		
Pritisk	$T_1 = 1900^o$	$T_2 = 2100^o$	$T_3 = 2300^o$
$P_1 = 5,0$	2 280 ali 244 260 } 484	338	266
	224	354	234
$P_2 = 12,5$	322	350	234
	342	370	258
$P_3 = 20,0$	330	290	330
	298	336	304

y_{PT}	T_1	T_2	T_3	y_P
P_1	504	692	500	1696
P_2	664	720	492	1876
P_3	628	626	634	1888
y_T	1796	2038	1626	5460 = y

Pomožne količine so:

$$Q_{PTi} = \sum\sum_{PTi} y^2_{PTi} = 260^2+224^2+\ldots+330^2+304^2 = 1689592$$

$$Q_{PT} = \frac{P.T}{n}\sum\sum_{PT} y^2_{PT} = \frac{1}{2}\left[504^2+692^2+\ldots 634^2\right] = 1686228$$

$$Q_P = \frac{P}{n}\sum_{P} y^2_P = \frac{1}{6}\left[1696^2+1876^2+1888^2\right] = 1660056$$

$$Q_T = \frac{T}{n}\sum_{T} y^2_T = \frac{1}{6}\left[1796^2+2038^2+1626^2\right] = 1670489{,}33$$

$$Q = \frac{1}{n} = \frac{1}{18}(5460)^2 = 1656200$$

$$q_{PTi} = 1689592$$
$$-\ \underline{1656200}\quad 33392$$
$$q_{PT} = 30028$$
$$q_P = 3856$$
$$q_T = 14289{,}33$$

AVAR je naslednja:

AVAR je naslednja:

VV	K	m	s^2	F
(P)	$K_P = q_P = 3856{,}00$	$P-1 = 2$	$s_P^2 = 1928$	$F_P = \frac{1928}{373{,}78} = 5{,}16^{x}$
(T)	$K_T = q_T = 14289{,}33$	$T-1 = 2$	$s_T^2 = 7144{,}67$	$F_T = \frac{7144{,}67}{373{,}78} = 19{,}11^{xxx}$
(PT)	$q_{PT} = 30028{,}00$ $- \ q_P = -3856{,}00$ $- \ q_T = -14289{,}33$ $K_{PT} = 11882{,}67$	$(P-1)(T-1) = 4$	$s_{PT}^2 = 2680{,}33$	$F_{PT} = \frac{2680{,}33}{373{,}78} = 7{,}17^{xx}$
pp	$q_{PTi} = 33392$ $- \ q_{PT} = -30028$ $K_e = 3364$	$(\bar{n}-1)PT = 9$	$s_e^2 = 373{,}78$	1
Sk	$q_{PTi} = 33392$	$n - 1 = 17$	-	

F \ α	0.05 [*]	0.01 [**]	0.001 [***]
$F_{\alpha}(2,9)$	4126	8,02	16,4
$F_{\alpha}(4,9)$	3,63	6,42	12,6

Kot sklepamo iz dobljenih razmerij F in kritičnih vrednosti iz tabel F porazdelitve, so razlike glavnega učinka pritiska značilno različne na stopnji $\alpha = 0.05$, razlike v temperaturi na stopnji $\alpha = 0.001$ in interakcije na stopnji $\alpha = 0.01$, kar je nakazano z ustreznim številom * pri vrednostih za F. Ker je značilna tudi interakcija (PT) moramo v analizi upoštevati kombinacijsko tabelo poprečij $\bar{y}_{PT}$, ne pa sumarna poprečja.

$\bar{y}_{PT}$	T_1	T_2	T_3
P_1	252	346	250
P_2	332	360	246
P_3	314	313	317

$$se(\bar{y}_{PT}) = \sqrt{\frac{s_e^2}{\bar{n}}} = \sqrt{\frac{373,78}{2}} = 41,071$$

$$d_{0.05}(\bar{y}_{PT}) = t(9)_{0.05}\, se\bar{y}_{PT} = 1,833 \cdot 41,071 = 75,28$$

7.9 R a z l i č n i m o d e l i č i s t o s l u č a j - n o s t n e g a d v o f a k t o r s k e g a p o s k u s a .

Nakazana analiza faktorskega poskusa in dan primer sta modela I. Oba faktorja C, D ali P,T imata namreč značaj fiksnih faktorjev.

Če nakažemo problematiko splošnejše vpeljimo v naš model za faktorja splošni oznaki c,d. Ta simbola sta v posebnem mali črki, če je faktor slučajnosten (model II.) in veliki črki, C in D, če je faktor fiksen (Model I.) po znanih zvezah je $\dot{c} = 1$, $\dot{d} = 1$ in $\dot{C} = 0$, $\dot{D} = 0$.

Nakažimo postopek analize variance za ta splošen primer.

Model: $y_{cdi} = M + (c) + (d) + (cd) + e_{cdi}$ (7.16)

Osnovna tabela podatkov

$$y_{cdi}$$

$$\begin{matrix} y_{cd} & y_c \\ y_d & y \end{matrix} \qquad (7.17)$$

Pomožne količine:

$$Q_{cdi} = SSS\, y^2_{cdi} \qquad Q_{cd} = \frac{cd}{n} SS y^2_{cd} \qquad Q_c = \frac{c}{n} S\, y^2_c$$

$$Q_d = \frac{d}{n} S\, y^2_d \qquad Q = \frac{1}{n}\, y^2 \qquad (7.18)$$

$$q_{cdi} = Q_{cdi} - Q\,;\; q_{cd} = Q_{cd} - Q\,;\; q_c = Q_c - Q\,;\; q_d = Q_d - Q$$

Analiza variance je: (7.19)

VV	K	m	s^2	$E(s^2)$
(c)	$K_c = q_c$	$m_c = c-1$	s^2_c	$\sigma^2_e + \bar{n}\dot{d}\sigma^2_{cd} + \bar{n}\, d\sigma^2_c$
(d)	$K_d = q_d$	$m_d = d-1$	s^2_d	$\sigma^2_e + \bar{n}\dot{c}\sigma^2_{cd} + \bar{n} c\sigma^2_d$
(cd)	$K_{cd} = q_{cd} - q_c - q_d$	$m_{cd} = (c-1)(d-1)$	s^2_{cd}	$\sigma^2_e + \bar{n}\cdot\sigma^2_{cd}$
pp	$K_e = q_{cdi} - q_{cd}$	$m_e = cd(\bar{n}-1)$	s^2_e	σ^2_e
Sk	$K = q_{cdi}$	$m = cd\bar{n} - 1$		

V koloni $E(s^2)$ so količine $\dot{c}$ in $\dot{d}$, kar nakazuje različne F primerjave za različne plane.

Če jih navedemo eksplicitno, dobimo: $E(s^2)$ za različne modele.

$E(s^2)$ za različne modele:

Model I. C,D fiksna | Model II. c,d slučajnostna (7.20)

$y_{CDi} = M + (C) + (D) + (CD) + e_{CDi}$ | $y_{cdi} = M + (c) + (d) + (cd) + e_{cdi}$

VV	$E(s^2)$	F	VV	$E(s^2)$	F
(C)	$\sigma_e^2 + \bar{n}D\,\sigma_C^2$	$F_C = s_C^2/s_e^2$	(c)	$\sigma_e^2 + \bar{n}\,\sigma_{cd}^2 + nd\,\sigma_c^2$	$F_c = s_c^2/s_{cd}^2$
(D)	$\sigma_e^2 + \bar{n}C\,\sigma_D^2$	$F_D = s_D^2/s_e^2$	(d)	$\sigma_e^2 + \bar{n}\,\sigma_{cd}^2 + nc\,\sigma_d^2$	$F_d = s_d^2/s_{cd}^2$
(CD)	$\sigma_e^2 + \bar{n}\,\sigma_{CD}^2$	$F_{CD} = s_{CD}^2/s_e^2$	(cd)	$\sigma_e^2 + \bar{n}\,\sigma_{cd}^2$	$F_{cd} = s_{cd}^2/s_e^2$
pp	σ_e^2	1	pp	σ_e^2	

Mešana modela (7.21)

C fiksen d slučajnosten | c slučajnosten, D fiksen

$y_{Cdi} = M + (C) + (d) + (Cd) + e_{Cdi}$ | $y_{cDi} = M + (c) + (D) + (cD) + e_{cDi}$

VV	$E(s^2)$	F	VV	$E(s^2)$	F
(C)	$\sigma_e^2 + \bar{n}\,\sigma_{cd}^2 + \bar{n}d\,\sigma_C^2$	$F_C = s_C^2/s_{Cd}^2$	(c)	$\sigma_e^2 + \bar{n}d\,\sigma_c^2$	$F_c = s_c^2/s_e^2$
(d)	$\sigma_e^2 + \bar{n}\,c\,\sigma_d^2$	$F_d = s_d^2/s_e^2$	(D)	$\sigma_e^2 + \bar{n}\,\sigma_{cD}^2 + \bar{n}C\,\sigma_D^2$	$F_D = s_D^2/s_{cD}^2$
(Cd)	$\sigma_e^2 + \bar{n}\,\sigma_{Cd}^2$	$F_{cd} = s_{Cd}^2/s_e^2$	(cD)	$\sigma_e^2 + \bar{n}\,\sigma_{cD}^2$	$F_{cD} = s_{cD}^2/s_e^2$
pp	σ_e^2		pp	σ_e^2	

V vseh štirih možnih variantah se jasno vidi, na katero komponento (s_{cd}^2 ali s_e^2) moramo primerjati s^2 za glavne učinke, da odkrivamo značilnost razlik med glavnimi učinki faktorjev C in D. Razlike so tudi v pomenu komponent varianc.

Tabela varianc v dvofaktorskem poskusu (7.22)

	d	D	
c	σ^2_{cd}	$\sigma^2_{cD} = \frac{1}{D-1} \sum_{D} \sigma^2_{cD}$	σ^2_{c}
C	$\sigma^2_{Cd} = \frac{1}{C-1} \sum_{C} \sigma^2_{Cd}$	$\sigma^2_{CD} = \frac{1}{(C-1)(D-1)} \sum_{C}\sum_{D} (CD)^2$	$\sigma^2_{C} = \frac{1}{C-1} \sum_{C} (C)^2$
	σ^2_{d}	$\sigma^2_{D} = \frac{1}{D-1} \sum_{D} (D)^2$	σ^2_{e}

Variance tipa I. σ^2_C, σ^2_D σ^2_{CD} so poprečni kvadrati učinka faktorjev

Variance tipa II. σ^2_c σ^2_d σ^2_{cd} so prave variance slučajnostnih faktorjev

Mešane variance σ^2_{cD} σ^2_{Cd} pa so poprečja grupnih varianc.

7.10 A p r i o r n e p r i m e r j a v e v d v o f a k t o r s k e m p o s k u s u .

Problematika primerjanja poskusnih podatkov dvofaktorskih poskusov je ista kot primerjave pri analizi učinka postopkov na splošno. Enako z apriornimi primerjavami preskušamo vnaprej predvidene zakonitosti, z aposteriornimi primerjavami pa razlike med poprečji ali skupinami poprečij.

7.11 Z n a č i l n o s t i r a z l i k m e d u č i n k i e n e g a f a k t o r j a p r i r a z l i č n i h n i v o j i h d r u g e g a f a k t o r j a.

Pod vplivom interakcije se učinek faktorja C različno ponaša pri različnih nivojih faktorja D in obratno. S preureditvijo analize variance moremo preskusiti značilnost razlik po enem faktorju na različnih nivojih drugega faktorja.

q_{CD}, ki je izraz delovanja glavnih učinkov (C), (D) in interakcije (CD), moremo pisati v obliki:

$$q_{CD} = q_C + (q_{CD} - q_C) =$$

$$K_{CD} = K_C + K_D + K_{CD} . \qquad (7.23)$$

$q_{CD} - q_C$ ima C(D-1) stopinj prostosti in ga moremo pisati

$$q_{CD} - q_C = \sum_C (q_D - q)/.C = K_D + K_{CD} \qquad (7.24)$$

$(q_D - q)/.C$ = doprinos $q_{D/C}$ faktorja D na nivoju C k skupni vsoti K

vsota doprinosov faktorja D na vseh nivojih je enaka $K_D + K_{CD}$ torej vsoti kvadratov odklonov glavnega učinka faktorja (D) in interakcije (CD). Vsak izmed teh delov ima $m_{D/C} = (D-1)$ stopinj prostosti, skupno torej C(D-1) stopinj prostosti.

Operativno izračunamo doprinos faktorja D na nivoju C drugega faktorja po obrazcu

$$q_{D/C} = \frac{1}{n} \sum_D y_{CD}^2 - \frac{1}{nD} y_C^2 \qquad (7.25)$$

in obratno

$$q_{C/D} = \frac{1}{n} \sum_C y_{CD}^2 - \frac{1}{n\,C} y_D^2 \qquad (7.26)$$

7.12 Nakazano analizo izvedemo za naš primer analize učinka temperature in pritiska takole:

$$q_{T/P_j} = \frac{1}{n} \sum_j y^2_{P_jT} - \frac{1}{n\,T} y^2_{P_j}$$

$$q_{P/Ti} = \frac{1}{n} \sum_P y^2_{PT_i} - \frac{1}{n\,P} y^2_{T_i}$$

$$q_{T/P1} = \frac{1}{2} \left[504^2+692^2+500^2 \right] - \frac{1}{2.3} 1696^2 = 12037.33$$

$$q_{T/P2} = \frac{1}{2} \left[664^2+720^2+492^2 \right] - \frac{1}{2.3} 1876^2 = 14117.33$$

$$q_{T/P3} = \frac{1}{2} \left[628^2+626^2+634^2 \right] - \frac{1}{2.3} 1888^2 = \underline{17.33}$$

$$2671.99$$

$$q_{P/T1} = \frac{1}{2} \left[504^2+664^2+628^2 \right] - \frac{1}{2.3} 1796^2 = 7045.33$$

$$q_{P/T2} = \frac{1}{2} \left[692^2+720^2+626^2 \right] - \frac{1}{2.3} 2038^2 = 2329.33$$

$$q_{P/T3} = \frac{1}{2} \left[500^2+492^2+634^2 \right] - \frac{1}{2.3} 1626^2 = \underline{6364.00}$$

$$15738.66$$

Analiza variance nivojskih razlik je

VV	K	m	s^2	F
(PT)	K_{PT} = 11882,67	4		
(T)	K_T = 14289,33	2		
(PT)+(T)	$K_{PT}+K_T$ = 26172	6		
$(P)_{P1}$	$q_{T/P1}$ = 12037,33	2	6018,67	16.102^{xx}
	$q_{T/P2}$ = 14117,33	2	7058,67	18.885^{xxx}
	$q_{T/P3}$ = 17,33	2	8,67	0.023
(P)	K_P = 3856,00	2		
(PT)+(P)	15738,67	6		
$(P)/T_1$	$q_{(P)/T1}$ = 7045,33	2	3522,67	9.424^{xx}
$(P)/T_2$	$q_{(P)/T2}$ = 2329,33	2	1164,67	3.116^{o}
$(P)/T_3$	$q_{(P)/T3}$ = 6364,00	2	3182,oo	8.513^{xx}
pp	K_e = 3364	9	373,78	1.000

$F_{0.10}(2.9) = 3{,}01$

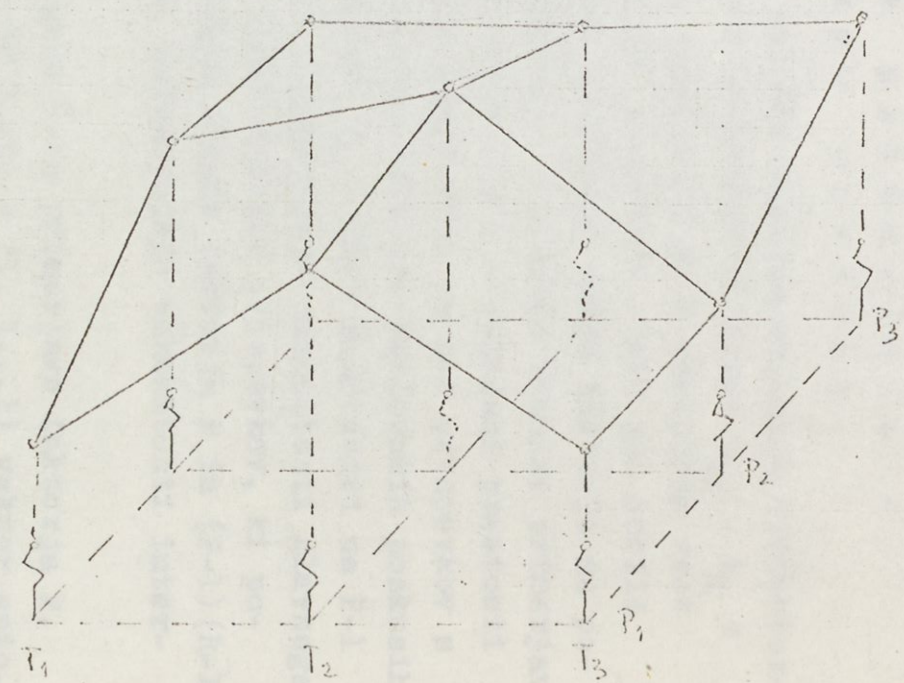
$F_{0.05}(2.9) = 4{,}26^{o}$ o

$F_{0.01}(2.9) = 8{,}02^{x}$ x

$F_{0.001}(2.9) = 16{,}4$ xx

xxx

Analiza je pokazala, da so razlike v gostoti značilno odvisne od temperature na nivoju pritiska P_1 in P_2, medtem ko so razlike neznačilne pri pritisku P_3. Gostota pa se je obratno pokazala značilno odvisna od pritiska pri temperaturi T_1 in T_3 medtem ko pri T_2 obstaja le sum na značilnost.



Sl. 7.3 Faktorski poskus z interakcijo (ploskev odgovorov).

Iz analize variance sta tudi razvidni zvezi:

$$K_{PT} + K_T = q_{T/P1} + q_{T/P2} + q_{T/P3}$$

$$11882,67+14289,33 = 12037,33+14117,33+17,34 = 26172$$

in

$$K_{PT} + K_P = q_{P/T1} + q_{P/T2} + q_{P/T3}$$

$$11882,67+3856,00 = 7045,33+2329,33+6364,00 = 15738,67$$

7.13 Ortogonalne primerjave v dvofaktorskem poskusu.

V prejšnjem primeru smo iz splošne analize variance dvofaktorskega poskusa vsoto kvadratov odklonov $K_{PT} + K_P$ z $P(T-1)$ stopinjami prostosti razstavili na P nivojskih vsot kvadratov s po $(T-1)$ stopinjami prostosti. Tako smo dobili osnovo za podrobnejšo analizo ponašanja enega faktorja na posameznih nivojih drugega faktorja. S podobno tehniko primerjav kot pri poskusih z enim faktorjem z $A-1$ stopinjami prostosti uspemo skupno vsoto kvadratov razstaviti na $A-1$ prispevkov s po eno stopinjo prostosti, moremo tudi pri faktorskih poskusih skupno število stopinj prostosti $m_A = PR-1$ razstaviti na $P-1$ individualnih prispevkov, ki pojasnjujejo zakonitosti glavnega učinka faktorja (P), $(R-1)$ individualnih prispevkov, ki pojasnjujejo zakonitosti glavnega učinka faktorja R in $(P-1)(R-1)$ individualnih prispevkov, ki pojasnjujejo zakonitosti interakcije (PR).

Z $C_P(\acute{P})$ zaznamujemo koeficiente $\acute{P}$-te primerjave faktorja P. V sistem primerjav vključimo še $C_P(0) = (1, 1 \ldots 1)$ vektor enic. Ta posreduje izraz $U_B(0) = \sum C_P(0) y_P = \sum 1.y_P = \sum_P y_P = y$.

y ni primerjava v pravem smislu, pač pa je ortogonalen z vsemi primerjavami, ker je $\sum_{P=1}^{P} C_P(0)\ C_P(P) = \sum C_P(P) = 0.$

Analogno z $C_R(\acute{R})$ zaznamujemo ustrezne koeficiente za ortogonalen sistem primerjav z $C_R(0) = (1,1,1)$ z enakim pomenom kot $C_P(0)$.

Če za vsak nivo faktorja R izračunamo primerjavo

$$\sum_{P=1}^{P} C_P(\acute{P}) y_{PR} = U_R(\acute{P}) \tag{7.27}$$

dobimo za $\acute{P} = 0, 1, 2 \ldots (P-1)$ vrsto primerjav. $U_R(\acute{P})$ so odvisni od R, če se primerjave $\acute{P}$ različno obnašajo po nivojih. Podobno kot iz poprečij ali vsot moremo tudi iz vrste primerjav po faktorju P izračunati dalje primerjave po faktorju R

$$\sum_{R} C_R(\acute{R}) U_R(\acute{P}) = \sum_{R=1}^{R} C_R(\acute{R}) \sum_{P=1}^{P} C_P(\acute{P}) y_{PR} = \sum_{R=1}^{R} \sum_{P=1}^{P} C_R(\acute{R}) C_P(\acute{P}) y_{PR} =$$

$$= U(\acute{P},\acute{R}) = U(\acute{R},\acute{P}) \tag{7.28}$$

$U(\acute{P},\acute{R})$ je primerjava, ki pojasnjuje interakcijo $\acute{P}$-te primerjave faktorja P z $\acute{R}$-to primerjavo faktorja R.

Prispevek komponente $(\acute{P},\acute{R})$ k vsoti kvadratov K_{PR} je

$$K_{(\acute{P},\acute{R})} = \frac{[U(\acute{P},\acute{R})]^2}{\bar{n} \sum_{P} C_P^2(\acute{P}) \sum_{R} C_R^2(\acute{R})} \tag{7.29}$$

Če najprej proučimo izraz K(0,0), dobimo, da je

$$K_{(0,0)} = \frac{\left(\sum_{PR} C_P(0) C_R(0) y_{PR}\right)^2}{\bar{n} \sum_{P} C_P^2(0) . \sum_{R} C_R^2(0)} = \frac{\left(\sum_{PR} y_{PR}\right)^2}{\bar{n} . P.R} = Q \tag{7.30}$$

korekturni faktor v analizi variance.

Dalje so

$$U(P',0) = \sum_{PR} c_P(P)c_R(0)y_{PR} = \sum_{P} c_P(P)y_P = U(P') \qquad (7.31)$$

$$U(0,R') = \sum_{P}\sum_{R} c_P(0)c_R(R)y_{PR} = \sum_{R} c_R(R)y_R = U(R') \qquad (7.32)$$

primerjave, s katerimi so pojasnjeni glavni učinki faktorjev P in R, njihovi prispevki k Q_{PR} pa so:

$$K(P',0) = \frac{\left(\sum_{PR} c_P(P')c_R(0)y_{PR}\right)^2}{\bar{n}\sum c_P^2(P') \,.\, \sum c_R^2(0)} = \frac{\left(\sum_{P} c_P(P)y_P\right)^2}{\bar{n}R \,.\, \sum c_P^2(P)} = K(P') \qquad (7.33)$$

$$P' = (1,2...P-1)$$

oziroma

$$K(0,R') = \frac{\left(\sum_{PR} c_P(0)c_R(R')y_{PR}\right)^2}{\bar{n} \,.\, \sum c_P^2(0) \sum_{R} c_R^2(R')} = \frac{\left(\sum_{R=1}^{R} c_R(R')y_R\right)^2}{\bar{n} \,.P.\, \sum_{R=1} c_R^2(R')} = K(R') \qquad (7.34)$$

$$R' = (1,2...\ R-1)$$

H Q z eno stopinjo prostosti, k (P-1) členom K(P') in (R-1) členi K(R') s po eno stopinjo prostosti za primerjave med glavnimi učinki, se pridruži še (P-1)(R-1) izrazov

$$K(P',R') = \frac{U(P',R')^2}{\bar{n}\sum c_P^2(P')\sum c_R^2(R')} \qquad P' = 1,2...(P-1), \quad R' = 1.2...(R-1) \qquad (7.35)$$

ki pojasnjujejo (P-1)(R-1) stopinj prostosti interakcije med (P) in (R).

7.14 Za primer vzemimo poskus, v katerem preskušamo uspešnost pri delu v odvisnosti od načina učenja (stari N_o novi N_1) in časa uvajanja ($T_1 = 5$ dni, $T_2 = 10$ dni, $T_3 = 15$ dni, $T_4 = 20$ dni). Poskus je bil izveden v dveh ponovitvah. V poskusu je sodelovalo $n = \bar{n}N . T = 2.2.4 = 16$ na slučajnosten način izbranih delavcev, katerim so bili postopki dodeljeni slučajnostno.

Osnovni podatki in ustrezne tabele so:

y_{NT2}	Čas učenja T_1	T_2	T_3	T_4
N_o	32	43	48	55
	29	47	52	59
N_1	23	35	50	72
	18	32	55	77

y_{NT}	T_1	T_2	T_3	T_4	y_N
N_o	61	90	100	114	365
N_1	41	67	105	149	362
y_T	102	157	205	263	727= = y

Standardna analiza variance je:

$$Q_{NTi} = \sum_{NTi} y^2_{NTi} = 32^2 + 29^2 + \ldots + 77^2 = 37177$$

$$Q_{NT} = \frac{NT}{n} \sum_{NT} y^2_{NT} = \frac{8}{16} \left[61^2 + 90^2 + \ldots + 149^2\right] = 37106{,}5$$

$$Q_N = \frac{N}{n} \sum y^2_N = \frac{1}{16} \left[365^2 + 372^2\right] = 33033{,}625$$

$$Q_T = \frac{T}{n} \sum y^2_T = \frac{4}{16} \left[102^2 + \ldots + 263^2\right] = 36561{,}75$$

$$Q = \frac{1}{n} y^2 = \frac{1}{16} \; 727^2 = 33033{,}0625$$

q_{NTi} = 4143,9375

q_{NT} = 4073,4375

q_{N} = 0,5625

q_{T} = 3528,6875

VV	K	m	s^2	F
(N)	q_N = 0,5625	1	0,5625	-
(T)	q_T = 3528,6875	3	1176,2292	133,47 xxx
(NT)	q_{NT} = 4073,4375			
	- q_N - 0,5625			
	- q_T - 3528,6875			
	K_{NT} = 544,1875	3	181,3958	20,58 xxx
pp	q_{NTi} 4143,9375			
	- q_{NT} - 4073,4375			
	K_e = 70,5000	8	8,8125	1,00
Sk	q_{NTi} = 4316,94	15	-	-

Standardna analiza variance je odkrila splošno visoko značilno odvisnost uspeha od časa učenja, ne pokaže pa značilnih razlik v glavnem učinku načina učenja. Pač pa je visoko značilna interakcija med načinom učenja in časom učenja, kar kaže na to, da je časovno uspeh bistveno različen glede na metodo učenja.

7. 15 Z analizo individualnih primerjav skušamo pojasniti zakonitosti, ki se nakazujejo v splošni analizi variance. Za faktor način učenja N z dvema vrednostima sta dve ortogonalni zvezi $\bar{y}_1 + \bar{y}_2 = {}_0U_N$ in $\bar{y}_2 - \bar{y}_1 = {}_1U_N$ katerih koeficiente strnemo v matriko

$${}'_{\acute{N}}C_N = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$ pri čemer $\acute{N}$ pomeni stopnjo primerjave

$\acute{N} = 0,1,2\ldots N-1$, N pa indeks koeficienta primerjave z $N = 1,2\ldots N$.

Podobno velja za drugi faktor T-čas, ki je numeričen v enakih časovnih razmakih. Zanj vzemimo kot sistem ortogonalnih primerjav ortogonalne polinome:

$${}'_{\acute{T}}C_T = \begin{pmatrix} 1 & 1 & 1 & 1 \\ -3 & -1 & 1 & 3 \\ 1 & -1 & -1 & 1 \\ -3 & 3 & -3 & 1 \end{pmatrix}$$

$\acute{T} = 0,1,2\ldots T-1$ stopnja primerjave (vrstica)

$T = 1,2\ldots N$ indeks koeficienta primerjave (stolpec)

Dalje je kvadrat matrike ${}'_{\acute{N}}C_N$: ${}'_{\acute{N}}C_N C'_{\acute{N}} = \begin{matrix} 2 & 0 \\ 0 & 2 \end{matrix} = {}'_{\acute{N}}C'_{\acute{N}}$

matrika kvadratov koeficientov in ${}'_{\acute{N}}C^{-1}_{\acute{N}} = \begin{matrix} \frac{1}{2} & 0 \\ 0 & 1/2 \end{matrix}$ inverzna matrika.

Za faktor T je kvadrat matrike ${}_T{}'C_T$ enak

$${}'_{\acute{T}}C_T C'_{\acute{T}} = \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 20 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 20 \end{pmatrix} = {}'_{\acute{T}}C'_{\acute{T}}$$ in inverzna matrika

$${}'_{\acute{T}}C^{-1}_{\acute{T}} = \begin{pmatrix} 1/4 & 0 & 0 & 0 & 0 \\ 0 & 1/20 & 0 & 0 \\ 0 & 0 & 1/4 & 0 \\ 0 & 0 & 0 & 1/20 \end{pmatrix}$$

Komponente z individualnimi stopinjami prostosti obračunamo iz tabele vsot $_Ny_T$ po naslednji shemi (v matriki pomeni indeks levo vrstico, indeks desno pa stolpec).

	T	$\acute N$	$\acute T$
N	$_Ny_T$	$_NC'_{\acute N}$	$_Ny_TC'_{\acute T} = {}_NU'_{\acute T}$ (7.36)
$'\dot T$	$_{\dot T}'C_T$	$_{\dot N}'C_Ny_TC'_{\acute T} = {}_{\dot N}'U'_{\acute T}$	$_{\dot T}'C_{\acute T}^{-1}$
$'\dot N$	$_{\dot N}'C_Ny_T = {}_{\dot N}'U_T$	$_{\dot N}'C_{\acute N}^{-1}$	$\dfrac{_{\dot N}'C_{\acute N}^{-1}\,U_{\acute T}^{2}\,C_{\acute T}^{-1}}{\bar n} = {}_{\dot N}'K'_{\acute T}$

V našem primeru dobimo

		T 1	2	3	4	$\acute N$ 0	1	$\acute T$ 0	1	2	3
N	1	61	90	100	114	1	- 1	365	169	- 15	23
	2	41	67	105	149	1	1	362	362	18	- 6
$'\dot T$	0	1	1	1	1	727	- 3	1/4	0	0	0
	1	-3	-1	1	3	539	193	0	1/20	0	0
	2	1	-1	-1	1	3	33	0	0	1/4	0
	3	-1	3	-3	1	17	-29	0	0	0	1/20
$'\dot N$	0	102	157	205	263	1/2	0	33033,1	3524,5^{xxx}	·6	3.6^{o}
	1	-20	-23	5	35	0	1/2	4,6	465,6^{xxx}	68,1^{xxx}	10,5^{x}

Iz zadnje matrike posnamemo: glavni učinek načina z eno stopinjo prostosti je enak K_N = ·5625 v standardni analizi variance. K_T = 3528,6875, ki odpade v standardni analizi variance z m_T = 3 stopinjami prostosti na faktor T_1 je razdeljen na tri individualne stopinje prostosti K_T = 3528,6775 = 3524.5125+·5625+3.6125, kar kaže na to, da je splošna linearna odvisnost od časa visoko značilna, kvadratična komponenta neznačilna (0.5625) medtem ko je kubična komponenta značilna na stopnji α = o.lo (obstaja sum/za značilnost). Obratno pa se izkažejo vse tri komponente interakcije od skupno 544,1875 = = 465,6125 + 68,0625 + 10,5125 kot značilne, kar kaže na to, da sta krivulji po obeh načinih v odvisnosti od časa bistveno različni v vseh komponentah. Največje razlike so v linearni komponenti, visoko značilne so tudi razlike v kvadratični komponenti, ki je za prvi način učenja N_o negativno (- 15) pri drugem načinu pa pozitivna (18). Na stopnji tveganja α = o.o5 so značilne tudi razlike v kubični komponenti.

7.16 Če po isti tehniki obdelamo še primer faktorskega poskusa 3 x 3 o gostoti briketov v odvisnosti od pritiska in temperature ob stiskanju, dobimo:

		T			P'			T'		
$P^y{}_T$		1	2	3	0	1	2	0	1	2
P	1	504	692	500	1	-1	1	1696	- 4	- 380
	2	664	720	492	1	0	-2	1876	-172	- 284
	3	628	626	634	1	1	1	1888	6	10
T'	0	1	1	1	5460	-170	-654	1/3	0	0
	1	- 1	0	1	192	10	390	0	1/2	0
	2	1	-2	1	-168	346	198	0	0	1/6
P'	0	1796	2038	1626	1/3	0	0			
	1	124	- 66	134	0	1/2	0			
	2	- 196	- 122	150	0	0	1/6			

·165000,000 11881,00 xxx

2708,3333 xx

3072,00 12,5000 6337,5000

784,00 4988,1667 544,50000

Matrika F		0	1	2
	0		$6{,}44^{xx}$	$31{,}79^{xxx}$
P	1	$8{,}22^{x}$	$0{,}03$	$16{,}96^{xx}$
	2	$2{,}10$	$13{,}35^{xx}$	$1{,}46$

Če člene v matriki K(P̈T̈) delimo z s_e^2 = 373,78, dobimo matriko F za preskus značilnosti za posamezne komponente. Te vrednosti primerjamo s kritičnimi vrednostmi F(1,9).

$F_{0.05}(1.9) = 5{,}12$ $F_{0.01}(1.9) = 10{,}6$ in $F_{0.001} = 22{,}9$.

Z x so zaznamovane stopnje značilnosti posameznih komponent. Za glavni učinek pritiska je linearna komponenta značilna na stopnji = o.o5, medtem, ko je kvadratična komponenta neznačilna. Iz predznaka ustreznega H = 192 vidimo, da se z naraščanjem pritiska na splošno gostota briketov linearno veča. Linearna komponenta glavnega učinka temperature je negativna in značilna na stopnji α = o.o5, prav tako je negativna in visoko značilna kvadratična komponenta (α = o.ool). Od interakcij sta se pokazali značilni na nivoju α = o.ol in pozitivni komponenti U (1,2) in U (2,1), kar kaže na to, da se linearna komponenta pritiska z večanjem temperature spreminja pozitivno kvadratično, to je, da izkaže od najnižje do srednje temperature padec, od srednje do najvišje pa porast ($P_{(1)}$ $T_1 = 124$, $T_2 = -62$, $T_3 = 134$) obratno pa kvadratična komponenta za temperaturo $T_{(2)}$ kaže značilen linearen porast $T_{(2)}$($P_1 = -380$, $P_2 = -284$, $P_3 = 10$).

Podobna značilnost U(P̍=2, T̍=1) kaže na to, da kvadratična komponenta pritiska z večanjem temperature linearno narašča U(P̍=2, T̍=1) = 3,46, ($P_{(2)}T_1$ = - 196, - 122, + 150) oziroma, da se linearna komponenta temperature z večanjem pritiska spreminja pozitivno kvadratično ($T_{(1)}$ - $P_1 = -4$, $P_2 = -172$, $P_3 = 6$).

7.17 T r o f a k t o r s k i s l u č a j n o s t n i p o s k u s s p o n o v i t v a m i .

Faktorski poskusi niso omejeni na dva faktorja, temveč moremo v poskus vključiti več faktorjev hkrati. Seveda se problem tehnično zamota, ker potrebno število postopkov pri večjih faktorjih hitro narašča, razen če je število vrednosti faktorjev razmeroma majhno.

Da nakažemo problematiko razširitve faktorskega poskusa, vzemimo splošen problem trofaktorskega poskusa. V poskusu s faktorji p,r, t z $\bar{n}$ ponovitvami obsega poskus $n = \bar{n}p.\ r\ .\ t$ poskusnih enot, v katerem $\bar{n}$ krat slučajnostno dodelimo p.r.t postopkov. Splošen linearen model čisto slučajnostnega trofaktorskega poskusa s ponovitvami je

$$y_{prti} = M+(p)+(r)+(pr)+(t)+(pt)+(rt)+(prt)+e_{prti} \qquad (7.37)$$

$$e_{prti} =: N(0,\sigma_e^2)$$

Vse komponente: glavni učinki (p), (r), (t), dvojne interakcije (pr),(pt),(rt) in trojna interakcija (prt) so splošne in morejo biti v konkretnem posamezni faktorji Modela I. in Modela II. Na splošno predpostavljamo, da je $e_{prti} =: N(0,\sigma_e^2)$

Iz tabele osnovnih podatkov y_{prti} izračunamo naslednje tabele vsot:

$$y_{prti}$$

$$y_{prt} \quad y_{pr} \quad y_{pt} \quad y_{p} \qquad (7.38)$$

$$y_{rt} \quad y_{r} \quad y_{t} \quad y$$

Iz tabel vsot izračunamo po znanem splošnem obrazcu pomožne količine $Q_z = \frac{z}{n}\sum_z y_z^2$ z = kompleks simbolov

$$Q_{prti} = \sum_{prti} y_{prti}^2 \qquad (7.39)$$

$$Q_{prt} = \frac{prt}{n} \sum_{prt} y^2_{prt} \; ; \quad Q_{pr} = \frac{pr}{n} \sum_{pr} y^2_{pr} \; ; \quad Q_{pt} = \frac{pt}{n} \sum_{pt} y^2_{pt} \; ;$$

$$Q_p = \frac{p}{n} \sum_p y^2_p \; ; \quad Q_r = \frac{r}{n} \sum_r y^2_r \; ; \quad Q_t = \frac{t}{n} \sum_t y^2_t \; ; \quad Q = \frac{1}{n} y^2$$

Iz njih izračunamo: $q_z = Q_z - Q$

Iz dobljenih količin pa obračunamo analizo variance po shemi:

VV	K	m	s^2	(7.40)
(p)	$K_p = q_p$	$p - 1$	s^2_p	
(r)	$K_r = q_r$	$r - 1$	s^2_r	
(pr)	$K_{pr} = q_{pr} - q_p - q_r$	$(p-1)(r-1)$	s^2_{pr}	
(t)	$K_t = q_t$	$(t-1)$	s^2_t	
(pt)	$K_{pt} = q_{pt} - q_p - q_t$	$(p-1)(t-1)$	s^2_{pt}	
(rt)	$K_{rt} = q_{rt} - q_r - q_t$	$(r-1)(t-1)$	s^2_{rt}	
(prt)	$K_{prt} = q_{prt} - q_{pr} - q_{pt} - q_{rt} + q_p + q_r + q_t$	$(p-1)(r-1)(t-1)$	s^2_{prt}	
(pp)	$K_e = q_{prti} - q_{prt}$	$prt(\bar{n}-1)$	s^2_e	
Sk	$K = q_{prti}$	$prt\bar{n}-1$		

s^2	$E(s^2)$	(7.41)
s^2_p	$\sigma^2_e + \bar{n}\dot{r}t\sigma^2_{prt} + \bar{n}\dot{r}t\sigma^2_{pr} + \bar{n}r\dot{t}\sigma^2_{pt} + \bar{n}rt\sigma^2_p$	
s^2_r	$\sigma^2_e + \bar{n}\dot{p}\dot{t}\sigma^2_{prt} + \bar{n}\dot{p}t\sigma^2_{pr} + \bar{n}p\dot{t}\sigma^2_{rt} + \bar{n}pt\sigma^2_r$	
s^2_{pr}	$\sigma^2_e + \bar{n}\dot{t}\sigma^2_{prt} + \bar{n}t\sigma^2_{pr}$	
s^2_t	$\sigma^2_e + \bar{n}\dot{p}\dot{r}\sigma^2_{prt} + \bar{n}\dot{p}r\sigma^2_{pt} + \bar{n}p\dot{r}\sigma^2_{rt} + \bar{n}pr\sigma^2_t$	

s^2	$E(s^2)$
s^2_{pt}	$\sigma^2_e + \bar{n}\,\dot{r}\,\sigma^2_{prt} + \bar{n}\,r\,\sigma^2_{pt}$
s^2_{rt}	$\sigma^2_e + \bar{n}\,\dot{p}\,\sigma^2_{prt} + \bar{n}\,p\,\sigma^2_{rt}$
s^2_{prt}	$\sigma^2_e + \bar{n}\,\sigma^2_{prt}$
s^2_e	σ^2_e

V shemi za analizo variance so za $E(s^2)$ podani splošni izrazi. V njih so faktorji označeni z malimi črkami. Če gre za fiksne faktorje modela I, male črke zamenjamo z velikimi črkami P,R,T in simboli $\acute{P} = \acute{R} = \acute{T} = 0$, če pa so faktorji slučajnostni (Model II) pa male črke p,r,t ohranimo in je $\dot{p} = \dot{r} = \dot{t} = 1$.

Če po teh navodilih sestavimo $E(s^2)$ za posamezne modele, dobimo za različne modele naslednje slike:

Matematično upanje za s^2, $E(s^2)$ pri različnih modelih za trofaktorski poskus (7.42)

Model I P, R, T			Mešan model P, R, t		
s^2	$E(s^2)$	F	s^2	$E(s^2)$	F
s_P^2	$\sigma_e^2 + \bar{n}\,RT\,\sigma_P^2$	s_P^2/s_e^2	s_P^2	$\sigma_e^2 + \bar{n}R\,\sigma_{Pt}^2 + \bar{n}Rt\,\sigma_P^2$	$F_P = s_P^2/s_{Pt}^2$
s_R^2	$\sigma_e^2 + \bar{n}PT\,\sigma_R^2$	s_R^2/s_e^2	s_R^2	$\sigma_e^2 \quad \bar{n}P\,\sigma_{Rt}^2 + nPt\,\sigma_R^2$	$F_R = s_R^2/s_{Rt}^2$
s_{PR}^2	$\sigma_e^2 + \bar{n}T\,\sigma_{PR}^2$	s_{PR}^2/s_e^2	s_{PR}^2	$\sigma_e^2 + \bar{n}\,\sigma_{PRt}^2 + \; + \bar{n}t\,\sigma_{PR}^2$	$F_{PR} = s_{PR}^2/s_{PRt}^2$
s_T^2	$\sigma_e^2 + \bar{n}PR\,\sigma_T^2$	s_T^2/s_e^2	s_t^2	$\sigma_e^2 + nPR\,\sigma_t^2$	$F_t = s_t^2/s_e^2$
s_{PT}^2	$\sigma_e^2 + \bar{n}R\,\sigma_{PT}^2$	s_{PT}^2/s_e^2	s_{Pt}^2	$\sigma_e^2 + \bar{n}R\,\sigma_{Pt}^2$	$F_{Pt} = s_{Pt}^2/s_e^2$
s_{RT}^2	$\sigma_e^2 + \bar{n}P\,\sigma_{RT}^2$	s_{RT}^2/s_e^2	s_{Rt}^2	$\sigma_e^2 + \bar{n}P\,\sigma_{Rt}^2$	$F_{Rt} = s_{Rt}^2/s_e^2$
s_{PRT}^2	$\sigma_e^2 + \bar{n}\,\sigma_{PRT}^2$	s_{PRT}^2/s_e^2	s_{PRt}^2	$\sigma_e^2 + \bar{n}\,\sigma_{PRt}^2$	$F_{PRt} = s_{PRt}^2/s_e^2$
s_e^2	σ_e^2	1	s_e^2	σ_e^2	

Matematično upanje za s^2, $E(s^2)$ pri različnih modelih za trofaktorski poskus (nadaljevanje)

Mešan model (7.43)

P, r, t

s^2	$E(s^2)$	F
s_P^2	$\sigma_e^2 + \sigma_{Prt}^2 + \bar{n}\,t\sigma_{Pr}^2 + \bar{n}r\sigma_{Pt}^2 + \bar{n}rt\sigma_P^2$	$F'_P = s_P^2/(s_{Pr}^2 + s_{Pt}^2 - s_{Prt}^2)$ $m'_1 = p-1;\ m'_2 = m'_{Prt} = \dfrac{(s_{Pr}^2 + s_{Pt}^2 - s_{Prt}^2)}{\dfrac{s_{Pr}^4}{m_{Pr}} + \dfrac{s_{Pt}^4}{m_{Pt}} + \dfrac{s_{Prt}^4}{m_{Prt}}}$
s_r^2	$\sigma_e^2 + \bar{n}P\sigma_{rt}^2 + \bar{n}Pt\sigma_r^2$	$F_r = s_r^2 / s_{rt}^2$
s_{Pr}^2	$\sigma_e^2 + \bar{n}\sigma_{Prt}^2 + \bar{n}t\sigma_{Pr}^2$	$F_{Pr} = s_{Pr}^2 / s_{Prt}^2$
s_t^2	$\sigma_e^2 + \bar{n}P\sigma_{rt}^2 + \bar{n}Pr\sigma_t^2$	$F_t = s_t^2 / s_{rt}^2$
s_{Pt}^2	$\sigma_e^2 + \bar{n}\sigma_{Prt}^2 + \bar{n}r\sigma_{Pt}^2$	$F_{Pt} = s_{Pt}^2/s_{Prt}^2$
s_{rt}^2	$\sigma_e^2 + \bar{n}\,r\sigma_{rt}^2$	$F_{rt} = s_{rt}^2/s_e^2$
s_{Prt}^2	$\sigma_e^2 + \bar{n}\sigma_{Prt}^2$	
s_e^2	σ_e^2	

Matematično upanje za s^2 $E(s^2)$ pri različnih modelih za trofaktorski poskus (nadaljevanje)

(7.44)

Model II.

p, r, t

s^2	$E(s^2)$	F
s_p^2	$\sigma_e^2 + \bar{n}\sigma_{prt}^2 + \bar{n}t\sigma_{pr}^2 + \bar{n}r\sigma_{pt}^2 + \bar{n}r\sigma_p^2$	$F'_p = s_p^2 / (s_{pr}^2 + s_{pt}^2 - s_{prt}^2)$ $m'_1 = (p-1)$ $m'_2 = m_{p.rt}$
s_r^2	$\sigma_e^2 + \bar{n}\sigma_{prt}^2 + \bar{n}t\sigma_{pr}^2 + np\sigma_{rt}^2 + \bar{n}pt\sigma_r^2$	$F'_r = s_r^2 / (s_{pr}^2 + s_{rt}^2 - s_{prt}^2)$ $m'_1 = (r-1)$ $m'_2 = m_{r.pt}$
s_{pr}^2	$\sigma_e^2 + \bar{n}\sigma_{prt}^2 + \bar{n}t\sigma_{pr}^2$	$F_{pr} = s_{pr}^2 / s_{prt}^2$
s_t^2	$\sigma_e^2 + \bar{n}\sigma_{prt}^2 + \bar{n}r\sigma_{pt}^2 + \bar{n}p\sigma_{rt}^2 + \bar{n}pr\sigma_t^2$	$F'_t = s_t^2 / (s_{pt}^2 + s_{rt}^2 - s_{prt}^2)$ $m'_1 = (t-1)$ $m'_2 = m_{t.pr}$
s_{pt}^2	$\sigma_e^2 + \bar{n}\sigma_{prt}^2 + \bar{n}r\sigma_{pt}^2$	$F_{pt} = s_{pt}^2 / s_{prt}^2$
s_{rt}^2	$\sigma_e^2 + \bar{n}\sigma_{prt}^2 + \bar{n}p\sigma_{rt}^2$	$F_{rt} = s_{rt}^2 / s_{prt}^2$
s_{prt}^2	$\sigma_e^2 + \bar{n}\sigma_{prt}$	$F_{prt} = s_{prt}^2 / s_e^2$
s_e^2	σ_e^2	

$$m'_{p.rt} = \frac{(s_{pr}^2 + s_{pt}^2 - s_{prt}^2)^2}{\frac{s_{pr}^4}{m_{pr}} + \frac{s_{pt}^4}{m_{pt}} + \frac{s_{prt}^4}{m_{prt}}}$$

analogno $m'_{r.pt}$ in $m'_{t.pr}$

7.18 Posebnost pri mešanem poskusu P,r,t, v katerem je en faktor (P) fiksen, dva (r in t) pa slučajnostna in pri Modelu II. p,r,t, v katerem so vsi trije faktorji slučajnostni, je v tem, da F-preskus za σ_p^2 pri mešanem poskusu in F-preskusi za σ_p^2a, σ_p^2in σ_t^2 pri Modelu II. ne moremo izvesti direktno. Niti v enem od teh primerov namreč ni dana zveza, na katero bi mogli te količine direktno primerjati. Če preskušamo v mešanem modelu (P, r, t) značilnost za σ_p^2, moramo s_p^2 primerjati z oceno, za katero je matematično upanje $\sigma_e^2 + \bar{n}\,\sigma_{Prt}^2 + \bar{n}t\,\sigma_{Pr}^2 + \bar{n}\,r\,\sigma_{Pt}^2$. Take količine pa v analizi variance ni.

Gornji izraz pa je matemtično upanje kombinacije $s_{Pr}^2 + s_{Pt}^2 - s_{Prt}^2$. Velja namreč:

$$E(s_{Pr}^2 + s_{Pt}^2 - s^2_{Prt}) = E(s_{Pr}^2) + E(s_{Pt}^2) - E(s_{Prt}^2) =$$

$$= (\sigma_e^2 + \bar{n}\sigma_{Prt}^2 + \bar{n}\,t\,\sigma_{Pr}^2) + (\sigma_e^2 + \bar{n}\,\sigma_{Prt}^2 + \bar{n}\,r\,\sigma_{Pt}^2) -$$

$$- (\sigma_e^2 + \bar{n}\,\sigma_{Prt}^2) = \sigma_e^2 + \bar{n}\,\sigma_{Prt}^2 + \bar{n}\,t\,\sigma_{Pr}^2 + \bar{n}\,r\,\sigma_{Pt}^2 \qquad (7.45)$$

Kot je razvidno iz sheme za F-preskus, si v tem primeru pomagamo s "quasi F-preskusom" tako, da v imenovalec v F vstavimo ustrezno linearno kombinacijo iz s^2. Število stopinj prostosti m_2 je izračunano iz števila stopinj prostosti posameznih s^2 po obrazcu

$$m'_{P.rt} = \frac{(s_{pr}^2 + s_{Pt}^2 - s_{Prt}^2)^2}{\frac{s_{pr}^4}{m_{pr}} + \frac{s_{pt}^4}{m_{pt}} + \frac{s_{Prt}^4}{m_{prt}}} \qquad (7.46)$$

Analogno izračunamo tudi $m'_{p.rt}$, $m'_{r.pt}$ in $m_{t.pr}$ pri modelu II. (prt).

$$m'_{p.rt} = \frac{(s^2_{pr} + s^2_{pt} - s^2_{prt})^2}{\frac{s^4_{pr}}{m_{pr}} \cdot \frac{s^4_{pt}}{m_{pt}} \cdot \frac{s^4_{prt}}{m_{prt}}}$$

$$m'_{r.pt} = \frac{(s^2_{pr} + s^2_{rt} - s^2_{prt})^2}{\frac{s^4}{m_{pr}} + \frac{s^4_{pt}}{m_{pt}} + \frac{s^4_{prt}}{m_{prt}}} \qquad (7.47)$$

$$m'_{t.pr} = \frac{(s^2_{pt} + s^2_{rt} - s^2_{prt})^2}{\frac{s^4_{pt}}{m_{pt}} + \frac{s^4_{rt}}{m_{rt}} + \frac{s^4_{prt}}{m_{prt}}}$$

7.19 Ti izrazi so posebni izrazi splošnega pravila, po katere linearni kombinaciji ocen varianc $s^2_e = \sum e_z s^2_z$, ki nastopa v F-preskusu, ustreza

$$m'_e = \frac{(\sum_z e_z s^2_z)^2}{\sum_z \frac{e^2_z s^4_z}{m_z}} \qquad (7.48)$$

To pravilo velja na splošno in ne samo za faktorske poskuse.

7.20 Ustavimo se/pri simboliki za variance in izračun ocen varianc.

V matematičnih upanjih $E(s^2)$ pri faktorskih poskusih naletimo na variance samo fiksnih faktorjev npr. σ^2_P, σ^2_{PR}, σ^2_{PRT}, samo slučajnostnih npr. σ^2_p, σ^2_{pr}, σ^2_{prt} ali mešani σ^2_{Pr}, $\sigma^2_{P.rt}$.

Za primere, da gre za varianco samo fiksnih faktorjev, izrazi npr. σ^2_P, σ^2_{PR}, σ^2_{PRT} pomenijo:

$$\sigma^2_P = \frac{1}{P-1}\sum_P (P)^2 \; ; \; \sigma^2_{PR} = \frac{1}{(P-1)(R-1)}\sum_{PR}(PR)^2 \; ;$$

$$\sigma^2_{PRT} = \frac{1}{(P-1)(R-1)(T-1)}\sum_{PRT}(PRT)^2 \qquad (7.49)$$

Za ocene variabilnosti samo slučajnostnih faktorjev so matematična upanja kvadratov učinkov npr. σ^2_p, σ^2_{pr}, σ^2_{prt}, medtem ko so variance z mešanimi faktorji npr. σ^2_{Pr}, $\sigma^2_{P.Rt}$, $\sigma^2_{P.rt}$

enaki

$$\sigma^2_{Pr} = \frac{1}{(P-1)}\sum_P \sigma^2_{Pr} \; ; \; \sigma^2_{PRt} = \frac{1}{(P-1)(R-1)}\sum_{PR}\sigma^2_{PRt}$$

$$\sigma^2_{Prt} = \frac{1}{(P-1)}\sum_P \sigma^2_{Prt} \qquad (7.50)$$

Ocene variance, ki nastopajo v faktorskih poskusih, često izračunavamo kot samostojne ocene odnosov med faktorji, ali kot elemente za načrtovanje bodočih poskusov. Če upoštevamo matematična upanja različnih komponent variance s^2 moremo s primerno linearno kombinacijo iz s^2-tov dobiti nepristranske ocene določenih komponent varianc. Te ocene morejo računsko ispasti tudi negativne. V tem primeru ocenimo varianco z 0. Če hočemo npr. iz mešanega poskusa P, r, t oceniti σ^2_t, dobimo to oceno po obrazcu

$$\hat{\sigma}^2_t = \frac{s^2_t - s^2_{rt}}{\bar{n}\, p\, r} \qquad (7.51)$$

Če upoštevamo odnose iz $E(s^2)$ v analizi variance, spoznamo da je

$$E\,\hat{\sigma}^2_t = \sigma^2_t \qquad (7.52)$$

7.21 F a k t o r s k i   p o s k u s i 2^p.

Z vsebinskega in metodološkega vidika predstavljajo faktorski poskusi 2^p, to so poskusi s faktorji, ki imajo po dve vrednosti, posebno kategorijo. Že pri razpravi o faktorjih na splošno smo nakazali poseben pomen faktorjev, ki imajo po dve vrednosti, v katerih nastopata alternativni vrednosti, kot so: a_o negnojeno - a_1 gnojeno, b_o star način - b_1 nov način, določena komponenta tehnološkega postopka: c_o ni bila uporabljena, c_1 je bila uporabljena; določen dodatek pijači: d_o = ni bil dodan, d_1 je bil dodan.

7.22 M o d e l   d v o f a k t o r s k e g a 2^2 p o s k u s a

Splošen model čisto slučajnostnega poskusa

$$y_{ABi} = M + (A) + (B) + (AB) + e_{ABi} \quad \text{in} \tag{7.53}$$

$$\bar{y}_{AB} = M + (A) + (B) + (AB) + \bar{e}_{AB}$$

dobi v poskusu 2^2 za posamezne postopke naslednjo obliko

$$\bar{y}_{a_1b_1} = M + \frac{1}{2}A + \frac{1}{2}B + \frac{1}{2}AB + \bar{e}_{11} \tag{7.54}$$

$$\bar{y}_{a_1b_o} = M + \frac{1}{2}A - \frac{1}{2}B - \frac{1}{2}AB + \bar{e}_{10}$$

$$\bar{y}_{a_ob_1} = M - \frac{1}{2}A - \frac{1}{2}B - \frac{1}{2}AB + \bar{e}_{01}$$

$$\bar{y}_{a_ob_o} = M - \frac{1}{2}A - \frac{1}{2}B + \frac{1}{2}AB + \bar{e}_{oo}$$

Vrednosti komponent glavnih učinkov (A), (B) in (AB) so pogojene s tem, da na vsako odpade ena sama stopinja prostosti in da je vsota $\sum(A) = \sum(B) = \sum_A(AB) = \sum_B(AB) = 0$, razlike v učinku pa A, B in AB.

Iz zgornjih enačb dobimo dalje:

$$\bar{y}_{a_1} = \frac{1}{2}\left[\bar{y}_{a_1b_1} + \bar{y}_{a_1b_o}\right] = M + \frac{1}{2}A + \bar{e}_{1.} \qquad (7.55)$$

$$\bar{y}_{a_o} = \frac{1}{2}\left[\bar{y}_{a_ob_1} + \bar{y}_{a_ob_o}\right] = M - \frac{1}{2}A + \bar{e}_{o}$$

$$\bar{y}_{\underline{a}} = \bar{y}_{a_1} - \bar{y}_{a_o} = A + (\bar{e}_{1.} - \bar{e}_{o.})$$

$$\bar{y}_{b_1} = \frac{1}{2}\left[(\bar{y}_{a_1b_1} + \bar{y}_{a_ob_1})\right] = M + \frac{1}{2}B + \bar{e}_{.1}$$

$$\bar{y}_{b_o} = \frac{1}{2}(\bar{y}_{a_1b_o} + \bar{y}_{a_ob_o}) = M - \frac{1}{2}B + \bar{e}_{.o}$$

$$\bar{y}_{\underline{b}} = \bar{y}_{b_1} - \bar{y}_{b_o} = B + \bar{e}_{.1} - \bar{e}_{.o} \qquad (7.57)$$

$$\frac{1}{2}(\bar{y}_{a_1b_1} - \bar{y}_{a_1b_o} - \bar{y}_{a_ob_1} + \bar{y}_{a_ob_o}) = AB + \frac{1}{2}(\bar{e}_{11} - \bar{e}_{1o} - \bar{e}_{o1} + \bar{e}_{oo})$$

Dalje dobimo:

$$\frac{1}{2}\;(\bar{y}_{a_1b_1} - \bar{y}_{a_1b_o}) - (\bar{y}_{a_ob_1} - \bar{y}_{a_ob_o}) = \frac{1}{2}\,\bar{y}_{a_1\underline{b}} - \bar{y}_{a_1\underline{b}} =$$

$$= \frac{1}{2}\,\bar{y}_{\underline{a}b_1} - \bar{y}_{\underline{a}b_o} = AB + e \ldots \qquad (7.58)$$

Interakcija AB je polovica razlike med učinkom faktorja b na nivoju a_1 in nivoju a_o in obratno polovica razlike med učinkom faktorja a na nivoju b_1 in nivoju b_o.

7.23 Zaradi tega, ker je število vrednosti posameznih faktorjev v poskusih 2^p enako dva, se računski postopek analize variance in interpretacije podatkov znatno poenostavi, če vpeljemo posebno simboliko in definicije. Neglede na to, da smo črke aA uporabljali sistematično za zaznamovanje postopkov, in b B za zaznamovanje blokov, uporabljamo črke po vrsti po abecedi a,b,c,d,e za faktorje v dvofaktorskih poskusih. Druga splošna poenostavitev v simboliki je v tem, da vsote poskusnih podatkov po postopkih zaznamujemo kar z ustreznimi indeksi tako, da v simbolu uporabimo ustrezno npr. črko a če je faktor vrednost a_1 in indeks izpustimo, če ima faktor vrednost a_o. Tako npr. v trofaktorskem poskusu 2^3 treh faktorjev a,b,c npr. vsoto $y_{a_1 b_1 c_o}$ zaznamujemo kar z "ab". S tem vemo, da je to vsota podatkov, na katerih so bili od treh faktorjev aplicirane vrednosti a_1 b_1 in c_o. Vrednost $y_{a_o b_o c_o} \equiv (1)$ postavimo v tej simboliki sistematično enako (1). Tako so v trofaktorskem poskusu po vrsti vsote po postopkih zaznamovane z:

$y_{a_1 b_1 c_1}$	$y_{a_1 b_1 c_o}$	$y_{a_1 b_o c_1}$	$y_{a_1 b_o c_o}$	$y_{a_o b_1 c_1}$	$y_{a_o b_1 c_o}$	$y_{a_o b_o c_1}$
abc	ab	ac	a	bc	b	c

$$y_{a_o b_o c_o} \qquad (7.59)$$

(1)

V tej simboliki je a-1 efekt razlike faktorja a, česta b in c na nivoju b_o in c_o .

Enako je ab - b učinek faktorja a, pri b_1 in c_o.

Podobno je ac - c učinek faktorja a, pri b_o in c_1 , in abc - bc učinek faktorja a pri b_1 in c_1.

Glavni učinek faktorja a za ves poskus je poprečje teh delnih učinkov

$$\frac{1}{4}\left[(abc - bc) + (ac - c) + (ab - b) + (a - 1)\right] \qquad (7.60)$$

Analogno velja za glavni učinek b

$$\frac{1}{4}\left[(abc - ac) + (bc - c) + (ab - a) + (b - 1)\right] \qquad (7.61)$$

in za glavni učinek faktorja c

$$\frac{1}{4}\left[(abc - ab) + (bc - b) + (ac - a) + (c - 1)\right] \qquad (7.62)$$

Interakcija (ab) na nivoju c_o je definirana s polovično razliko učinka faktorja a in nivoju b_1 in b_o na nivoju c_o

$$\frac{1}{2}\left[(ab - b) - (a-1)\right] \qquad (7.63)$$

interakcija (ab) na nivoju c_1 pa s polovično razliko učinka faktorja a na nivoju b_1 in b_o pri c_1

$$\frac{1}{2}\left[(abc - bc) - (ac - c)\right] \qquad (7.64)$$

Interakcija ab v celem poskusu pa je definirana s poprečjem med obema prispevkoma, torej

$$AB = \frac{1}{2}\left[(ab - b) - (a - 1)\right] + \frac{1}{2}\left[(abc - bc) - (ac - c)\right] \qquad (7.65)$$

Analogno sta definirani tudi ostali dve dvojni interakciji AC in BC, trojna interakcija ABC pa je dalje polovična razlika med interakcijo AB na nivoju c_1 in na nivoju c_o, torej

$$ABC = \frac{1}{2}\left\{\frac{1}{2}\left[(abc - bc) - (ac - c)\right] - \frac{1}{2}\left[(ab - b) - (a - 1)\right]\right\} \qquad (7.66)$$

Zgornja simbolika omogoča izredno enostaven matematični izraz za formiranje linearnih zvez kombinacij postopkov. Za trofaktorski poskus formalni produkt

$$(a \pm 1)(b \pm 1)(c \pm 1) \qquad (7.67)$$

omogoča sestavljanje ustreznih linearnih zvez tako, da upoštevamo "+", če faktor ne nastopa v učinku komponente in "-" če nastopa v učinku komponente.

Tako je npr. za interakcijo

$$BC = \frac{1}{2^2}(a + 1)(b - 1)(c - 1) = \frac{1}{4}\left[abc - bc - ac + c - ab + b + a - 1\right] \text{ ipd.}$$

7.24 Sistematično moremo prikazati formiranje posameznih komponent tudi v matriki koeficientov vrednosti +1 in -1, ki jih zaradi nazornosti zaznamujemo z ustreznimi predznaki + in - . Za 2^3 je matrika naslednja:

Postopek		Faktor a	b	c	Komponenta M	(a)	(b)	(ab)	(c)	(ac)	(bc)	(abc)
$a_0b_0c_0$	1	-	-	-	+	-	-	+	-	+	+	-
$a_1b_0c_0$	a	+	-	-	+	+	-	-	-	-	+	+
$a_0b_1c_0$	b	-	+	-	+	-	+	-	-	+	-	+
$a_1b_1c_0$	ab	+	+	-	+	+	+	+	-	-	-	-
$a_0b_0c_1$	C	-	-	+	+	-	-	+	+	-	-	+
$a_1b_0c_1$	ac	+	-	+	+	+	-	-	+	+	-	-
$a_0b_1c_1$	bc	-	+	+	+	-	+	-	+	-	+	-
$a_1b_1c_1$	abc	+	+	+	+	+	+	+	+	+	+	+

Z
7.25/Yatesovo metodo izračunavamo učinke posameznih komponent sistematično izredno enostavno. Potek izračuna in struktura izračuna je nakazana za 2^3 poskus.

Tabela ima 2^3 vrstic (kolona 1)

1. V koloni 3 po vrsti pišemo vsote podatkov sistematično po posameznih kombinacijah od 1 dalje.
2. Seštevamo po dva in dva podatka iz kolone 3 in jih po vrsti vpisujemo v kolono 4. Za dvojnimi vsotami vpisujemo dalje razlike po dva zaporedna podatka (glej shemo v koloni 2).
3. Izračun iz točke 2 ponovimo na podatkih iz kolone 4 in rezultate vpišemo v kolono 5.
4. Če ponovimo izračun iz točke 2 na podatkih iz kolone 5 dobimo v koloni 6 linearne zveze, ki ustrezajo posameznim komponentam.
5. Če podatke iz kolone 6 delimo z $\bar{n}2^2$ dobimo v koloni 7 posamezne komponente; $\bar{n}$ pomeni število ponovitev ali blokov.
6. Če podatke iz kolone (6) kvadriramo in delimo z $\bar{n}\ 2^3$, dobimo v koloni 8 prispevke posameznih komponent k skupni vsoti kvadratov abc, ki izvira iz delovanja vseh treh faktorjev.

Shematičen preskus Yatesove metode za izračun komponent za poskus 2^3 je naslednji:

Shematični preskus Yatesove metode za izračun komponent za poskus 2^3 je naslednji:

(7.70)

Yates						Efekt $(6)/\bar{n}\, 2^2$	K $(6)/\bar{n}\, 2^3$	m
1	2	3	4	5	6	7	8	9
1	2+1	1	a+1	(ab+b)+(a+1)	(abc+bc+ac+c)+(ab+b+a+1)	M	K_M	1
2	4+3	a	ab+b	(abc+bc)+(ac+c)	(abc-bc+ac-c)+(ab-b+a-1)	A	K_A	1
3	6+5	b	ac+c	(ab-b)+(a-1)	(abc+bc-ac-c)+(ab+b-a-1)	B	K_B	1
4	8+7	ab	abc+bc	(abc-bc)+(ac-c)	(abc-bc-ac+c)+(ab-b-a+1)	AB	K_{AB}	1
5	2-1	c	a-1	(ab+b)-(a+1)	(abc+bc+ac+c)-(ab+b+a+1)	C	K_C	1
6	4-3	ac	ab-b	(abc+bc)-(ac+c)	(abc-bc+ac-c)-(ab-b+a-1)	AC	K_{AC}	1
7	6-5	bc	ac-c	(ab-b)-(a-1)	(abc+bc-ac-c)-(ab+b-a-1)	BC	K_{BC}	1
8	8-7	abc	abc-bc	(abc-bc)-(ac-c)	(abc-bc-ac+c)-(ab-b-a+1)	ABC	K_{ABC}	1
							Q_{ABC}	2^3

Če premotrimo koeficiente, s katerimi množimo vsote postopkov, da dobimo posamezne komponente, vidimo, da so te linearne zveze ortogonalne primerjave, ki pojasnjujejo z $2^3-1 = 7$ individualnimi stopinjami prostosti sedem komponent trafaktorskega poskusa, z osmo (vsota podatkov) pa še sredino.

7.26 Razširitev metodologije na 2^p poskus.

Vse kar smo nakazali za trofaktorski poskus 2^3, moremo zlahka razširiti na 2^p poskus.

Formalni produkt za komponente moremo z istim tolmačenjem kot za poskus 2^3 razširiti na poskus 2^p

$$(a \pm 1)(b \pm 1)(c \pm 1)(d \pm 1)(e \pm 1) \ldots \qquad (7.71)$$

Tudi tabela komponent se da zlahka razširiti, če sistematično napišemo vse postopke a_i b_i c_i d_i e_i ... pripišemo predznake za individualne faktorje a b c d e in postavimo predznak "+" , če postopek vsebuje vrednost na nivoju 1 in "-", če vsebuje vrednost na nivoju 0. Pri predznakih za komponente (glavne učinke in interakcije do stopnje p)pa algebrično množimo vrednosti 1 z ustreznimi predznaki za vse tiste faktorje, ki nastopajo v komponenti.

Tabela komponent do poskusa 2^5 je prikazana v tabeli 7.2.

Prav tako moremo uporabiti Yatesovo metodo, da sistematično po vrsti napišemo vse postopke, potrebno seštevanje in odštevanje pa izvedemo p-krat. Po p-kratnem ponavljanju procesa vsot in razlik, dobimo komponentam ustrezne linearne zveze postopkov. Če dobljene u_k delimo z $\bar{n} \cdot 2^{p-1}$ dobimo izvrednotene posamezne komponente u_k $u_k^2/\bar{n} \cdot 2^p = K$ pa so individualni prispevki posameznih komponent k skupni vsoti kvadratov zaradi faktorjev z 2^{p-1} stopinjami prostosti.

Tabela 7.2 ortogonalne primerjave za glavne učinke in interakcije faktorskih poskusov 2^2 do 2^5

učinki

2^2

postepki	I	A	B	AB
(1)	+	−	−	+
a	+	+	−	−
b	+	−	+	−
ab	+	+	+	+
c	+	−	−	+
ac	+	+	−	−
bc	+	−	+	−
abc	+	+	+	+
d	+	−	−	+
ad	+	+	−	−
bd	+	−	+	−
abd	+	+	+	+
cd	+	−	−	+
acd	+	+	−	−
bcd	+	−	+	−
abcd	+	+	+	+
e	+	−	−	+
ae	+	+	−	−
be	+	−	+	−
abe	+	+	+	+
ce	+	−	−	+
ace	+	+	−	−
bce	+	−	+	−
abce	+	+	+	+
de	+	−	−	+
ade	+	+	−	−
bde	+	−	+	−
abde	+	+	+	+
cde	+	−	−	+
acde	+	+	−	−
bcde	+	−	+	−
abcde	+	+	+	+

2^3

postepki	C	AC	BC	ABC
(1)	−	+	+	−
a	−	−	+	+
b	−	+	−	+
ab	−	−	−	−
c	+	−	−	+
ac	+	+	−	−
bc	+	−	+	−
abc	+	+	+	+
d	−	+	+	−
ad	−	−	+	+
bd	−	+	−	+
abd	−	−	−	−
cd	+	−	−	+
acd	+	+	−	−
bcd	+	−	+	−
abcd	+	+	+	+
e	−	+	+	−
ae	−	−	+	+
be	−	+	−	+
abe	−	−	−	−
ce	+	−	−	+
ace	+	+	−	−
bce	+	−	+	−
abce	+	+	+	+
de	−	+	+	−
ade	−	−	+	+
bde	−	+	−	+
abde	−	−	−	−
cde	+	−	−	+
acde	+	+	−	−
bcde	+	−	+	−
abcde	+	+	+	+

2^4

postepki	D	AD	BD	ABD	CD	ACD	BCD	ABCD
(1)	−	+	+	−	+	−	−	+
a	−	−	+	+	+	+	−	−
b	−	+	−	+	+	−	+	−
ab	−	−	−	−	+	+	+	+
c	−	+	+	−	−	+	+	−
ac	−	−	+	+	−	−	+	+
bc	−	+	−	+	−	+	−	+
abc	−	−	−	−	−	−	−	−
d	+	−	−	+	−	+	+	−
ad	+	+	−	−	−	−	+	+
bd	+	−	+	−	−	+	−	+
abd	+	+	+	+	−	−	−	−
cd	+	−	−	+	+	−	−	+
acd	+	+	−	−	+	+	−	−
bcd	+	−	+	−	+	−	+	−
abcd	+	+	+	+	+	+	+	+
e	−	+	+	−	+	−	−	+
ae	−	−	+	+	+	+	−	−
be	−	+	−	+	+	−	+	−
abe	−	−	−	−	+	+	+	+
ce	−	+	+	−	−	+	+	−
ace	−	−	+	+	−	−	+	+
bce	−	+	−	+	−	+	−	+
abce	−	−	−	−	−	−	−	−
de	+	−	−	+	−	+	+	−
ade	+	+	−	−	−	−	+	+
bde	+	−	+	−	−	+	−	+
abde	+	+	+	+	−	−	−	−
cde	+	−	−	+	+	−	−	+
acde	+	+	−	−	+	+	−	−
bcde	+	−	+	−	+	−	+	−
abcde	+	+	+	+	+	+	+	+

2^5

postepki	E	AE	BE	ABE	CE	ACE	BCE	ABCE	DE	ADE	BDE	ABDE	CDE	ACDE	BCDE	ABCDE
(1)	−	+	+	−	+	−	−	+	+	−	−	+	−	+	+	−
a	−	−	+	+	+	+	−	−	+	+	−	−	−	−	+	+
b	−	+	−	+	+	−	+	−	+	−	+	−	−	+	−	+
ab	−	−	−	−	+	+	+	+	+	+	+	+	−	−	−	−
c	−	+	+	−	−	+	+	−	+	−	−	+	+	−	−	+
ac	−	−	+	+	−	−	+	+	+	+	−	−	+	+	−	−
bc	−	+	−	+	−	+	−	+	+	−	+	−	+	−	+	−
abc	−	−	−	−	−	−	−	−	+	+	+	+	+	+	+	+
d	−	+	+	−	+	−	−	+	−	+	+	−	+	−	−	+
ad	−	−	+	+	+	+	−	−	−	−	+	+	+	+	−	−
bd	−	+	−	+	+	−	+	−	−	+	−	+	+	−	+	−
abd	−	−	−	−	+	+	+	+	−	−	−	−	+	+	+	+
cd	−	+	+	−	−	+	+	−	−	+	+	−	−	+	+	−
acd	−	−	+	+	−	−	+	+	−	−	+	+	−	−	+	+
bcd	−	+	−	+	−	+	−	+	−	+	−	+	−	+	−	+
abcd	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
e	+	−	−	+	−	+	+	−	−	+	+	−	+	−	−	+
ae	+	+	−	−	−	−	+	+	−	−	+	+	+	+	−	−
be	+	−	+	−	−	+	−	+	−	+	−	+	+	−	+	−
abe	+	+	+	+	−	−	−	−	−	−	−	−	+	+	+	+
ce	+	−	−	+	+	−	−	+	−	+	+	−	−	+	+	−
ace	+	+	−	−	+	+	−	−	−	−	+	+	−	−	+	+
bce	+	−	+	−	+	−	+	−	−	+	−	+	−	+	−	+
abce	+	+	+	+	+	+	+	+	−	−	−	−	−	−	−	−
de	+	−	−	+	−	+	+	−	+	−	−	+	−	+	+	−
ade	+	+	−	−	−	−	+	+	+	+	−	−	−	−	+	+
bde	+	−	+	−	−	+	−	+	+	−	+	−	−	+	−	+
abde	+	+	+	+	−	−	−	−	+	+	+	+	−	−	−	−
cde	+	−	−	+	+	−	−	+	+	−	−	+	+	−	−	+
acde	+	+	−	−	+	+	−	−	+	+	−	−	+	+	−	−
bcde	+	−	+	−	+	−	+	−	+	−	+	−	+	−	+	−
abcde	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

* A larger table will be found in [15.4].

7.27 Za primer vzemimo raziskavo o vplivu načina prodaje (a_o - stari, a_1 - novi) razlika v ceni (b_o = višja b_1 = nižja) in kakovosti (c_1 = boljša c_o = slabša) na višino prodaje.

Poskus je bil izveden tako, da so bili trije kraji izbrani kot bloki, v vsakem kraju pa po 2^3=8 po značaju in splošnem prometu podobnih trgovin, katerim je bila slučajnostno določena po ena izmed A = 2^3 kombinacij prodaje.

Rezultati raziskave so naslednji:

y_{AB} A	B_1	B_2	B_3	y_A
1	8	13	11	32
a	9	18	9	36
b	15	24	22	61
ab	14	22	28	64
c	12	28	13	63
ac	9	17	22	48
bc	18	26	27	71
abc	17	29	18	64
y_B	102	177	150	429

Standardna analiza variance je naslednja:

Iz zgornjih podatkov dobimo

$$Q_{AB} = \sum y_{AB}^2 = 8^2 + 13^2 + \ldots + 18^2 = 8723$$

$$Q_A = \frac{1}{3}\sum y_A^2 = \frac{1}{3}\left[32^2 + 36^2 + \ldots + 64^2\right] = 8129$$

$$Q_B = \frac{1}{8}\sum y_B^2 = \frac{1}{8}\left[102^2 + 177^2 + 150^2\right] = 8029{,}12$$

$$Q = \frac{1}{3.8}\, y^2 = \frac{429^2}{3.8} = \quad = 7668{,}37$$

$q_{AB} = 1054,63$

$q_A = 460,63$

$q_B = 360,75$

Standardna analiza variance:

VV	K	m	s^2	F
B	$q_B = 360,75$	2	180,37	
A	$q_A = 460,63$	7	65,80	$3,95^x$
pp	$q_{AB}-q_A-q_B = 233,25$	14	16,66	
Sk	$q_{AB} = 1054,63$	23		

Razlike med postopki, ki so kombinacije treh faktorjev, so se pokazale značilne na stopnji $\alpha = 0.05$, ker je

$$F_{0.05}(7,14) = 2.76 < F_A = 3.95 < F_{0.01}(7,14) = 4,28$$

$K_A = 460,63$ bomo z Yatesovo metodo razstavili na $m_A = 7$ individualnih stopinj prostosti, od katerih vsaka pokaže po eno komponento trofaktorskega poskusa.

A	y_A	y_1	y_2	y_3	U_k	$\frac{u_n^2}{3.2^3} = K$	A
1	32	32+36= 68	68+125=193	193+236=429	35,720	-	
a	36	61+64=125	101+135=236	7+(-12)=-5	- .417	1,04	(a)
b	61	53+48=101	4+ 3= 7	57+ 34= 91	7.583	345,04	(b)
ab	64	71+64=135	-5 - 7=-12	-1+(-2)=- 3	- .250	,38	(ab)
c	53	36-32= 4	125 -68= 57	236-193= 43	3.583	77,04	(c)
ac	48	64-61= 3	135-101= 34	-12-(-7)=-19	-1.583	15,04	(ac)
bc	71	48-53= -5	3- 4 = -1	34-57 = -23	-1.917	22,04	(bc)
abc	64	64-71= -7	-7-(-5)=-2	-2-(-1)= -1	.083	,04	(abc)
				$\Sigma = 512 = 8 \cdot 64$		$460,62 = K_A$	

Kontrola: $\Sigma y_3 = 2^3 \cdot y_{abc}$

Iz algebrične identitete sledi, da je vsota končnih y = u , ki jih dobimo z Yatesovo metodo enak na splošno

$$\Sigma Y_{(p)} = 2^p y_{abcd} \dots$$

V našem primeru je resnično: $\Sigma Y_{(3)} = 512 = 2^3 \cdot 64$

Enako se sklada vsota prispevka vseh sedmih komponent s skupno vsoto kvadratov postopkov. Če te vrednosti vnesemo v tabelo za analizo variance, dobimo:

VV	K	m	s^2	F
Bloki	360,75	2	180,37	
Postopki skupaj	460,62	7		
(a)	1,04	1	1.04	
(b)	345,04	1	345,04	20,71 xxx
(ab)	.38	1	.38	
(c)	77.04	1	77.04	4,62 x
(ac)	15.04	1	15.04	
(bc)	22.04	1	22.04	1,32
(abc)	.04	1	.04	
pp	233.25	14	16.66	
Sk		23		

Analiza variance faktorjev je pokazala značilnost glavnega učinka faktorja (b) cene in faktorja (c) kakovosti. Interakcija (bc) ni značilna, ker je $F_{0.05}(1.14) = 4,60$. Iz tega sklepamo, da je prodaja značilno večja pri nižjih cenah in boljši kakovosti.

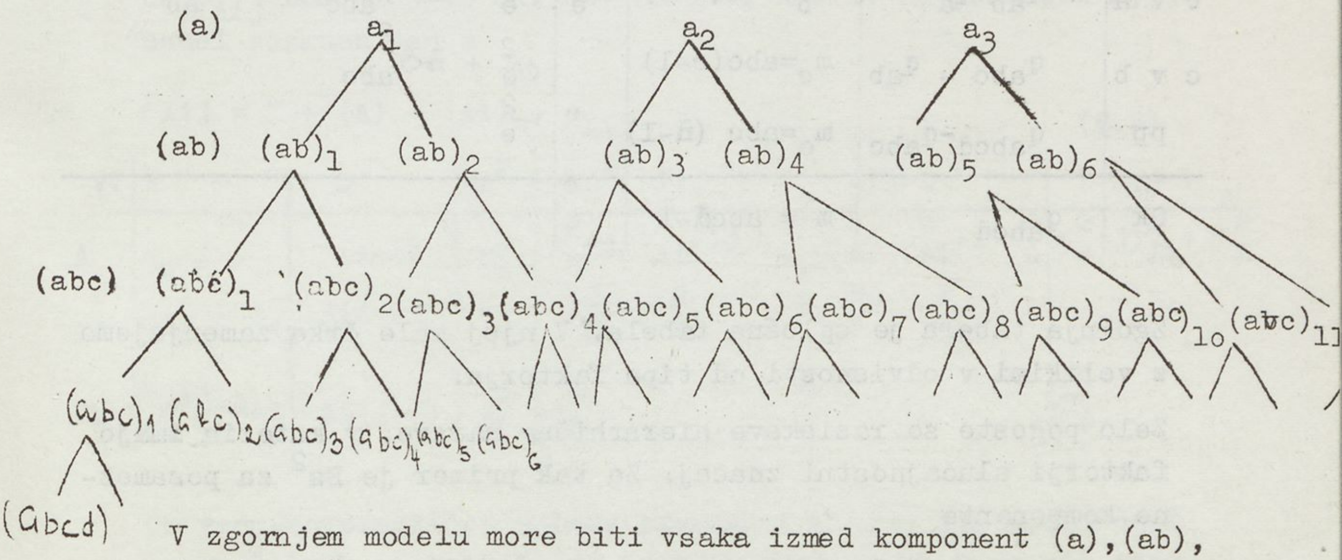
8. ČISTI HIERARHIČNI POSKUS

8.1 Vzemo, da proučujemo za določen rudnik delež pepela v premogu. Če vzamemo posamezne kope, iz njih izberemo slučajnostno vagončke, iz vagončkov pa slučajnostne kose premoga, je raziskava v veliki meri slična vzorčenju v več stopnjah. Variabilnost pepela v premogu v celoti je v tem primeru odvisna od razlik med kopi, od razlik med vagončki v kopih in razlik med kepami premoga na posameznih vagončkih.

Če z y_{abc} zaznamujemo c-to kepo b-tega vagončka v a-tem kopu, moremo sestaviti model

$$y_{abcd} = M + (a)+(ab)+(abc)+e_{abcd} \qquad e_{abcd} = :N(0, \sigma_e^2) \quad (8.1)$$

Ker je vrstni red virov variacije urejen - hierarhičen: kopa, vagončki v kopih, kepe v vagončkih, imenujemo tak poskus hierarhičen ali gnezdast. V njem se enote nižje stopnje kopičijo v gnezdih - enotah višje stopnje. Shema takega poskusa je:



V zgornjem modelu more biti vsaka izmed komponent (a),(ab), (abc) bilo tipa I. ali II. razen v zadnji stopnji, za katero predpostavljamo, da je faktor slučajnostni z $e_{abcd} = :N(0, \sigma_e^2)$

Pomožne količine, potrebne za obračun analize variance so

$y_{abcd} \quad y_{abc} \quad y_{ab} \quad y_a \quad y$

$$Q_{abcd} = \underset{abcd}{S}\, y^2_{abcd} \qquad Q_{abc} = \frac{abc}{n} \underset{abc}{S}\, y^2_{abc} \qquad Q_{ab} = \frac{ab}{n} \underset{ab}{S} y^2_{ab}$$

$$Q_a = \frac{a}{n} \underset{a}{S}\, y^2_a \qquad Q = \frac{1}{n} y^2 \qquad (8.3)$$

$$q_{abcd} = Q_{abcd} - Q; \quad Q_{abc} = Q_{abc} - Q\ ; \quad q_{ab} = Q_{ab} - Q\ ; \quad q_a = Q_a - Q \qquad (8.4)$$

Analiza variance za hierarhičen poskus je:

(8.5)

VV	K	m	s^2	Es^2
a	q_a	$m_a = a-1$	s_e^2	$\sigma_e^2 + \bar{n}\cdot c\,\sigma^2_{abc} + \bar{n}cb\,\sigma^2_{ab} + \bar{n}cb\,\sigma^2_a$
b v a	$q_{ab} - q_a$	$m_b = a(b-1)$	s_e^2	$\sigma_e^2 + \bar{n}c\,\sigma^2_{abc} + \bar{n}c\,\sigma^2_{ab}$
c v b	$q_{abc} - q_{ab}$	$m_c = abc(c-1)$		$\sigma_e^2 + \bar{n}\sigma^2_{abc}$
pp	$q_{abcd} - q_{abc}$	$m_e = abc\,(\bar{n}-1)$		σ_e^2
Sk	q_{abcd}	$m = abc\bar{n}-1$		

Zgornja tabela je splošna tabela. V njej male črke zamenjujemo z velikimi v odvisnosti od tipa faktorja.

Zelo pogoste so raziskave hierarhične narave, v katerih imajo faktorji slučajnostni značaj. Za tak primer je Es^2 za posamezne komponente

s^2	$E(s^2)$	(8.6)
s_a^2	$\sigma_e^2 + \bar{n}\sigma_{abc}^2 + \bar{n}\,c\,\sigma_{ab}^2 + \bar{n}\,cb\,\sigma_a^2$	
s_b^2	$\sigma_e^2 + \bar{n}\sigma_{abc}^2 + \bar{n}c\sigma_{ab}^2$	
s_c^2	$\sigma_e^2 + \bar{n}\sigma_{abc}^2$	
s_e^2	σ_e^2	

Iz teh zvez zlahka dobimo, da so ocene za variance komponent

$$\hat{\sigma}_a^2 = \frac{s_a^2 - s_b^2}{\bar{n}\ cb} \qquad \hat{\sigma}_{ab}^2 = \frac{s_b^2 - s_c^2}{\bar{n}\ c}\ ; \qquad \hat{\sigma}_{abc}^2 = \frac{s_c^2 - s_e^2}{\bar{n}} \tag{8.7}$$

8.2 Da ponovitev meritve na isti poskusni enoti ni ponovitev v pravem smislu in da nima istega učinka kot prava ponovitev, moremo raztolmačiti na modelu za hierarhičen poskus.

Če y_{Aij} pomeni j-to meritev na i-ti enoti, za postopek A, je model poskusa dan s

$$y_{Aij} = M + (A) + (Ai) + e_{Aij} \tag{8.8}$$

VV	K	m	s^2	Es^2 i;j	i;j=1	i=1;j
A	q_A	(A-1)	s_A^2	$\sigma_e^2 + j\sigma_{Ai}^2 + ij\sigma_A^2$	$\sigma_e^2 + \sigma_{Ai}^2 + i\sigma_A^2$	$\sigma_e^2 + j\sigma_{Ai}^2 + \sigma_A^2$
pp	$q_{Ai} - q_A$	A(i-1)	s_e^2	$\sigma_e^2 + j\sigma_{Ai}^2$	$\sigma_e^2 + \sigma_{Ai}^2$	-
vp	$q_{Aij} - q_{Ai}$	Ai(j-1)	s_v^2	σ_e^2	-	σ_e^2

Če vzamemo praktičen primer preskušanja produktivnosti dela po različnih postopkih, je poskus organiziran v pravih ponovitvah če i na slučajnostni način izbranih delavcev dela po posameznih postopkih. Ker je v vsakem primeru po eno samo merjenje, (j=1) velja v analizi varianceE(s) za j=1. Če primerjamo

s_A^2 z s_e^2 vidimo, da s tem razmerjem resnično preskušamo značilnost σ_A^2. Če pa je poskus organiziran tako, da izberemo A delavcev in vsakemu na slučajnosten način dodelimo po eden izmed postopkov, vsak delavec pa po tem postopku dela j-krat, v tem primeru to niso prave ponovitve, temveč le j meritev na isti enoti. Iz zadnjega stolpa spoznamo, da je za primer s_e^2 nedoločen, razmerje

$$E\, s_A^2/s_V^2 = \frac{\sigma_e^2 + j\,\sigma_{Ai}^2 + \sigma_A^2}{\sigma_e^2}$$

pa odkrije združen učinek variabilnosti zaradi razlik med delavci σ_{Ai}^2 in razlik med postopki σ_A^2. F-preskus v tem primeru ni upravičen.

9. DELNI BLOKI (Split-plot)

9.1 Vzemimo kot primer poskus v kmetijstvu, v katerem nastopata dva faktorja: (p) agrotehnična mera (npr. 2 različni globini oranja) in (r) gnojenje (3 različne doze gnojenja). Ker z določeno globino orjemo le večje površine, z določeno dozo gnojenja pa lahko gnojimo tudi na manjših površinah, faktorski poskus p x r izvedimo na naslednji način: poskusno gradivo predstavlja $\bar{n}p$ večjih poskusnih parcel, na katerih na slučajnosten način apliciramo p agrotehničnih mer v $\bar{n}$ ponovitvah. Vsako od $\bar{n}p$ poskusnih parcel pa razdelimo na r enakih podparcelic, na katerih uporabimo slučajnostno r doz gnojenja. Poskus je glede na faktor p čisto slučajnosten poskus. S poskusno enoto večjo parcelo, glede na način gnojenja pa predstavlja vsaka večja poskusna parcela blok, v katerem so na r poskusnih podparcelicah aplicirane vse vrednosti faktorja r. Poskus je faktorski poskus, kombiniran tako, da je faktor (p) načrtovan v čisto slučajnostnem poskusu, faktor (r) pa v slučajnostnih blokih. Ker so osnovne površine razdeljene v manjše, se ta proces poskusa imenuje split-plot poskus, ker pa niso vsi faktorji v slučajnostnih blokih, imenujemo poskus, poskus v delnih blokih.

Shema nakazanega poskusa izvedenega v $\bar{n}$ = 3 ponovitvah je naslednja:

p_1	p_2	p_2	p_1	p_1	p_2	
r_1	r_3	r_1	r_2	r_1	r_2	
r_3	r_2	r_2	r_1	r_3	r_1	(9.1)
r_2	r_1	r_3	r_3	r_2	r_3	

9.2 Z delnimi bloki v čisto slučajnostnem poskusu rešimo tudi naslednji primer. Da bi kompleksno proučili vpliv tipa trgovine in načina prodaje na velikost prodaje, smo na slučajnosten način izbrali po $\bar{n}$ trgovini določenega tipa, v vsaki izmed teh pa slučajnostno predpisali vrstni red r načinov prodaje. Glede na tip trgovine je poskus čisto slučajnosten poskus trgovin, glede na način prodaje pa je vsaka trgovina blok, v katerem prodajamo po slučajnostnem vrstnem redu na vse proučevane načine.

9.3 Če proučujemo odvisnost produktivnosti dela v odvisnosti od dveh stopenj staža in dveh načinov dela na slučajnosten način, za vsako stopnjo staža izberemo $\bar{n}$ delavcev, od katerih vsak delavec dela po slučajnostnem vrstnem redu po obeh načinih dela. V tem primeru je faktor staž v čisto slučajnostnem poskusu z delavci kot poskusnimi enotami, faktor "način dela" pa faktor v blokih - delavcih.

9.4 Zaradi dvojnega značaja faktorjev nastopajo v poskusih v delnih blokih dve vrsti poskusnih pogreškov: $e_{pi} = : N(0,\sigma_1^2)$ je pogrešek med poskusnimi enotami čisto slučajnostnega dela poskusa s faktorjem p,

$e_{pir} =: N(0,\sigma_2^2)$ pa pogrešek, ki izvira iz poskusov z vključitvijo drugega faktorja r.

Linearni model dvofaktorskega poskusa p x r v delnih blokih je torej:

$$y_{pir} = M + (p) + e_{pi} + (r) + (pr) + e_{pir} \qquad \begin{array}{l} e_{pi} = : N(0,\sigma_1^2) \\ e_{pir} = : N(0,\sigma_2^2) \end{array} \tag{9.2}$$

Osnovne podatke y_{pir} za analizo variance seštejemo v naslednje tabele:

$$\begin{array}{ccc} y_{pir} & y_{pr} & y_r \\ y_{pi} & y_p & y \end{array} \qquad (9.3)$$

Tabele v prvi vrsti sheme služijo za obračun komponent v blokih, v drugi pa za obračun čisto slučajnostnega poskusa faktorja p.

Standardne pomožne količine $Q_{\check{z}}$ so:

$$Q_{pir} = \sum_{pir} y_{pir}^2 \qquad Q_{pr} = \frac{1}{\bar{n}} \sum_{pr} y_{pr}^2 \qquad Q_r = \frac{1}{np} \sum_r y_r^2 \qquad (9.4)$$

$$Q_{pi} = \frac{1}{r} \sum_{pi} y_{pi}^2 \qquad Q_p = \frac{1}{\bar{n}\, r} \sum_p y_p^2 \qquad Q = \frac{1}{\bar{n}\, pr}\, y^2$$

iz njih je $q_{\check{z}} = Q_{\check{z}} - Q$

Analiza variance za delne bloke v čisto slučajnostnem poskusu, je naslednja:

(9.5)

VV	K	m	s^2	$E(s^2)$
(p)	q_p	p-1	s_p^2	$\sigma_2^2 + \bar{n}\sigma_1^2 + \bar{n}r\,\sigma_{pr}^2 + \bar{n}r\,\sigma_p^2$
$pp_1 = e_1$	$q_{pi} - q_p$	$p(\bar{n}-1)$	s_1^2	$\sigma_2^2 + \bar{n}\sigma_1^2$
(r)	q_r	(r-1)	s_r^2	$\sigma_2^2 + \quad + \bar{n}p\,\sigma_{pr}^2 + \bar{n}p\,\sigma_r^2$
(pr)	$q_{pr} - q_p - q_r$	(p-1)(r-1)	s_{pr}^2	$\sigma_2^2 \quad + \bar{n}\,\sigma_{pr}^2$
$pp_2 = e_2$	$q_{pri} - q_{pr} - q_{pi} + q_p$	$p(\bar{n}-1)(r-1)$	s_e^2	σ_2^2
Sk	q_{pri}	pir - 1		

Splošna shema analize variance nakazuje različne F-preskuse glede na značaj faktorjev p in r (fiksen ali slučajnosten). Če navedemo štiri možne variante, dobimo naslednje odnose:

Model I.: P,R fiksna faktorja

(9.6)

VV	s^2	$E(s^2)$	F
(P)	s_P^2	$\sigma_2^2 + \bar{n}\sigma_1^2 + \bar{n}r\sigma_P^2$	$F_P = s_P^2/s_1^2$
pp_1	s_1^2	$\sigma_2^2 + \bar{n}\sigma_1^2$	$F_1 = s_1^2/s_2^2$
(R)	s_R^2	$\sigma_2^2 + \bar{n}p\sigma_R^2$	$F_R = s_R^2/s_2^2$
(PR)	s_{PR}^2	$\sigma_2^2 + \bar{n}\sigma_{PR}^2$	$F_{PR} = s_{PR}^2/s_2^2$
pp_2	s_2^2	σ_2^2	

Mešan model: P fiksen, r slučajnosten (9.7)

VV	s^2	$E(s^2)$	F
(P)	s_P^2	$\sigma_2^2 + \bar{n}\sigma_1^2 + \bar{n}\sigma_{Pr}^2 + \bar{n}r\sigma_P^2$	$F_P = s_P^2/(s_1^2+s_{Pr}^2-s_2^2)$
pp_1	s_1^2	$\sigma_2^2 + \bar{n}\sigma_1^2$	$F_1 = s_1^2/s_2^2$
(r)	s_r^2	$\sigma_2^2 + \bar{n}p\sigma_r^2$	$F_r = s_r^2/s_2^2$
(Pr)	s_{Pr}^2	$\sigma_2^2 + \bar{n}\sigma_{pr}^2$	$F_{Pr} = s_{Pr}^2/s_2^2$
pp_2	s_2^2	σ_2^2	

Mešan model p slučajnosten , R fiksen

(9.8)

VV	s^2	$E(s^2)$	F
(p)	s_p^2	$\sigma_2^2 + \bar{n}\sigma_1^2 + \bar{n}\, r\sigma_p^2$	$F_p = s_p^2/s_1^2$
pp_1	s_1^2	$\sigma_2^2 + \bar{n}\sigma_1^2$	$F_1 = s_1^2/s_2^2$
(R)	s_R^2	$\sigma_2^2 + \bar{n}\sigma_{pR}^2 + \bar{n}p\sigma_R^2$	$F_R = s_R^2/s_{pR}^2$
(pR)	s_{pR}^2	$\sigma_2^2 + \bar{n}\sigma_{pR}^2$	$F_{pR} = s_{pR}^2/s_2^2$

Model II. p slučajnosten, r slučajnosten (9,9)

VV	s^2	$E(s^2)$	F
(p)	s_p^2	$\sigma_2^2 + \bar{n}\sigma_1^2 + \bar{n}\sigma_{pr}^2 + \bar{n}\, r\sigma_p^2$	$F_p' = s_p^2/(s_1^2+s_{pr}^2-s_2^2)$
pp_1	s_1^2	$\sigma_2^2 + \bar{n}\sigma_1^2$	$F_1 = s_1^2/s_2^2$
(r)	s_r^2	$\sigma_2^2 + \quad + \bar{n}\sigma_{pr}^2 + \bar{n}\, p\sigma_r^2$	$F_r = s_r^2/s_{pr}^2$
(pr)	s_{pr}^2	$\sigma_2^2 \quad + \bar{n}\sigma_{pr}^2$	$F_{pr} = s_{pr}^2/s_2^2$
pp_2	s_2^2	σ_2^2	

9.5 V zgornjem pregledu so iz splošnega modela in $E(s^2)$ za posamezne komponente izpeljane ocene s^2 in $E(s^2)$ za vse štiri možne kombinacije vrst faktorjev P,R; P,r; p,R in p,r. Iz $E(s^2)$ so izpeljana ustrezna razmerja F. Iz pregleda vidimo, da je v modelih, v katerih je faktor r slučajnosten, možen za preskus faktorja p ali P samo "quasi F-preskus".

Analogno kot pri quasi-F-preskusu pri običajnem faktorskem poskusu, tudi v tem primeru izračunamo linearni $(s_1^2 + s_{pr}^2 - s_2^2)$ ustrezno približno število stopinj prostosti m_2'

$$m_2' = \frac{(s_1^2 + s_{pr}^2 - s_2^2)^2}{\dfrac{s_1^4}{m_1} + \dfrac{s_{pr}^4}{m_{pr}} + \dfrac{s_2^4}{m_2}}$$

9.6 Glede na dvojnost poskusnega pogreška so variance za primerjave sredin odvisne od tega, katere sredine primerjamo. Če shematično nakažemo sredine v tabeli delnih blokov

$$\begin{array}{ccc|c}
\bar{y}_{p1r1} & -3- & \bar{y}_{p1r2} & \bar{y}_{p1} \\
4 & 4 & 4 & 1 \\
\bar{y}_{p2r1} & -3- & \bar{y}_{p2r2} & \bar{y}_{p2} \\
\hline
\bar{y}_{r1} & -2- & \bar{y}_{r2} &
\end{array} \qquad (9.10)$$

1 Ocena variance razlik sredin prvega faktorja p

$$\text{var}\, \Delta\bar{y}_p = \frac{2\, s_1^2}{\bar{n}\, r} \qquad (9.11)$$

2 Ocena variance razlik sredin faktorja v blokih r

$$\text{var}\, \Delta\bar{y}_r = \frac{2 s_2^2}{\bar{n}\, p} \qquad (9.12)$$

3 Ocena variance razlik sredin v faktorju r na istem nivoju faktorja p

$$\text{var}\, \Delta\bar{y}_{plr} = \frac{2\, s_2^2}{\bar{n}} \qquad (9.13)$$

4 Ocena variance razlik sredin v faktorju p na istem ali različnem nivoju faktorja r

$$\text{var}\,\Delta y_{pr} = \frac{2\left[(r-1)s_2^2 + s_1^2\right]}{\bar{n}\,r} \qquad (9.14)$$

9.7 Metodo delnih blokov uporabljamo takrat, kadar je en faktor mogoče aplicirati le na večjih poskusnih enotah, drug faktor pa na pred enotah. Zanesljivost ocen učinka faktorja (p) je manjša kot za ocene učinka faktorja (r) in interakcije (pr).

9.8 Vzemimo za primer proučevanje odvisnosti produktivnosti dela od ropota in pozicije delovnega mesta. Faktor ropot ima dve vrednosti (tišina - ropot), pozicija delovnega mesta pa tri vrednosti (svetlo, srednje svetlo, temačno). Po načrtu je predvideno, da poskus izvedemo v $\bar{n} = 3$ ponovitvah.

Ker moremo pogoje ropot - tišina ustvariti le za več delavcev v istem prostoru, ne pa za posameznega delavca, je načrtovan poskus v delnih blokih tako, da je faktor ropot v čisto slučajnostnem poskusu v treh ponovitvah, faktor pozicija delovnega mesta pa v bloku. Poskus je izveden na $\bar{n}p = 3.2 = 6$ osnovnih enotah prve stopnje - prostorih in na $\bar{n}p.r = 3.2.3 = 18$ osnovnih enotah druge stopnje - delavcih.

Od šestih prostorov smo slučajnostno pripisali trem ropot, trem tišino in slučajnostno dodelili 18 delavcev pozicijam po prostorih.

Urejeni rezultati poskusa so naslednji:

Pogoj dela	y_{PIR} R_1	R_2	R_3	y_{PI}
P_{11}	33	32	28	93
P_{12}	38	35	32	105
P_{13}	36	30	31	97
P_{21}	30	27	20	77
P_{22}	30	24	18	72
P_{33}	28	25	17	70
y_{PR} P_1	107	97	91	295 y_P
P_2	88	76	55	219
y_R	195	173	146	514 Y

$$Q_{PIR} = \sum_{PIR} y^2_{PIR} = 33^2 + \ldots + 17^2 = 15274 \qquad q_{PIR} = 596,45$$

$$Q_{PI} = \frac{1}{R} \sum_{PI} y^2_{PI} = \frac{1}{3}\left[93^2 + \ldots + 70^2\right] = 15032 \qquad q_{PI} = 354,45$$

$$Q_{PR} = \frac{1}{\bar{n}} \sum_{PR} y^2_{PR} = \frac{1}{3}\left[107^2 + \ldots + 55^2\right] = 15228 \qquad q_{PR} = 450,45$$

$$Q_P = \frac{1}{\bar{n}\ R} \sum_{P} y^2_P = \frac{1}{3.3}\left[295^2 + 219^2\right] = 14998,44 \qquad q_P = 320,89$$

$$Q_R = \frac{1}{\bar{n}P} \sum_R y_R^2 = \frac{1}{3.2}\left[195^2+173^2+146^2\right] = 14878{,}33 \qquad q_R = 200{,}78$$

$$Q = \frac{1}{\bar{n}\;PR}\, y^2 = \frac{1}{3.2.3}\, 514^2 = 14677{,}55$$

Analiza variance pa je po shemi 9.4

VV	K	m	s^2	F
Pogoj dela (P)	$q_P = 320{,}89$	1	320,89	38,247 xxx
pp_1	$q_{PI} - q_P = 33{,}56$	4	8,39	1
Razporeditev (R)	$q_R = 200{,}78$	2	100,39	10,089 xxx
(PR)	$q_{PR}-q_P-q_R = 28{,}78$	2	14,39	1,446
pp_2	$q_{PRI}-q_{PR}-q_R+q_{RI} = 79{,}56$	8	9,95	1
Sk		17		

Analiza variance pokaže, da sta glavna učinka obeh faktorjev (pogoja dela in pozicije) visoko značilna. Glede na različne vrste pogreškov pogoj dela preskušamo

$F_P = s_P^2/s_1^2 = 320{,}89/8{,}39$ pozicijo pa z

$F_R = s_R^2/s_2^2 = 100{,}39/9{,}95 = 10{,}089.$

Ker je $F_{0.001}(1,4) = 25{,}4$ in $F_{0.01}(2,8) = 8{,}65$ $F_{0.001}^{(2,8)} = 18{,}5$, sta glavna učinka obeh faktorjev visoko značilna, interakcija pa neznačilna. V okviru poskusa smatramo, da sta učinka faktorjev pogoj dela in pozicija aditivna.

V tabeli poprečij veljajo odnosi

$\bar{y}_{PR}$	35,7	32,3	30,3	32,8	$\bar{y}_P$
	29,3	25,3	18,3	24,3	
$\bar{y}_R$	32,5	28,8	24,3	28,6	$\bar{y}$

$$\text{var } \Delta\bar{y}_P = \frac{2\, s_1^2}{\bar{n}\, .\, r} = \frac{2.8,39}{3.3} = 1,864 \qquad \text{se}\Delta\bar{y}_P = 1,36$$

$$\text{var } \Delta\bar{y}_r = \frac{2\, s_2^2}{\bar{n}\, .\, p} = \frac{2.9,95}{3.2} = 3,317 \qquad \text{se}\Delta\bar{y}_r = 1,82$$

$$\text{var } \Delta\bar{y}_{plr} = \frac{2\, s_2^2}{\bar{n}} = \frac{2.9,95}{3} = 6,633 \qquad \text{se}\Delta\bar{y}_{Plr} = 2,57$$

$$\text{var}\Delta\bar{y}_{pr} = \frac{2\,[(r-1)s_2^2 + s_1^2]}{\bar{n}\ r} = \frac{2\,[(3-1)9,95+8,39]}{3.3} = 6,287$$

$$\text{se}\Delta\bar{y}_{pr} = 2,51$$

9.9 Delni bloki v drugih načrtih

Delne bloke ne uporabljamo samo pri čisto slučajnostnih poskusih, temveč jih kombiniramo tudi z drugimi elementi poskusov. S pridom kombiniramo delne bloke s splošno tehniko slučajnostnih blokov, z latinskimi kvadrati in podobno.

Shema poskusa delnih blokov v slučajnostnih blokih je

b_1		b_2		b_3	
p_1	p_2	p_1	p_1	p_1	p_2
r_1	r_3	r_3	r_1	r_1	r_2
r_3	r_1	r_2	r_3	r_3	r_1
r_2	r_2	r_1	r_2	r_2	r_3

(9.15)

Ustrezen linearni model pa:

$$Y_{bpr} = M + (b) + (p) + e_{bp} + (p) + (pr) + e_{bpr}$$

$$e_{bp} = : N(0, \sigma_1^2) \qquad e_{bpr} = : N(0, \sigma_2^2) \qquad (9.16)$$

Shema poskusa delnih blokov v latinskem kvadratu je:

	s_1	s_2	s_3
v_1	r_1 p_1 r_1	r_2 p_2 r_1	r_2 p_3 r_1
v_2	r_1 p_2 r_2	r_2 p_3 r_1	r_1 p_1 r_2
v_3	r_1 p_3 r_2	r_2 p_1 r_1	r_1 p_2 r_2

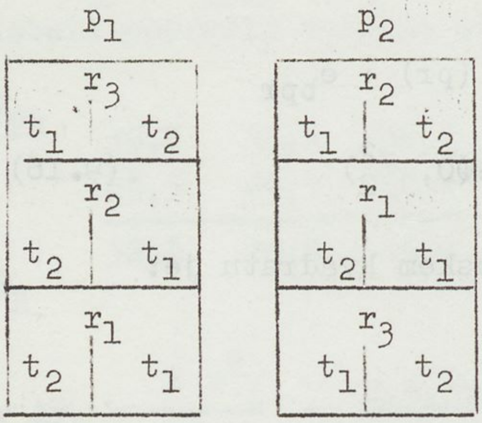
(9.17)

ustrezen linearen model pa:

$$Y_{vspr} = M + (v) + (s) + (p) + e_{vsp} + (r) + (pr) + e_{vspr}$$

$$e_{vsp} = : N(0, \sigma_1^2) \qquad e_{vspr} = : N(0, \sigma_2^2) \qquad (9.18)$$

Podobno kot imamo vzorčenje v več stopnjah tudi poskuse v delnih blokih razširimo na več faktorjev tako, da je poskusna enota višjega faktorja blok za naslednji faktor. Shema ene ponovitve takega načrta (split-split-plota) s tremi faktorji: p, r, t je:



(9.20)

linearni model split-split-plota pa je

$$Y_{pirt} = M + (p) + e_{pi} + (r) + (pr) + e_{pir} + (t) + (pt) + (rt) + (prt) + e_{pirt}$$

$$e_{pi} = : N(0, \sigma_1^2) \qquad e_{pir}^2 = : N(0, \sigma_2^2) \quad e_{pirt} = N(0, \sigma_3^2)$$

V tem načrtu imamo torej tri poskusne pogreške: e_{pi}, e_{pir} in e_{pirt}. Temu ustrezna je tudi analiza variance in F-preskusi.

9.10 D e l n i b l o k i v s l u č a j n o s t n i h b l o k i h .

Ker pogosto delne bloke kombiniramo s slučajnostnimi bloki, bomo prikazali problematiko in metodologijo le-teh. Pri proučevanju vpliva agrotehnične mere in gnojenja, more biti blok posamezno kmetijsko gospodarstvo, na katerem izvedemo vse postopke poskusa v tehniki delnega bloka.

Pri proučevanju vpliva tipa trgovine in načina prodaje na velikost prodaje, morejo biti bloki posamezni kraji ali regije. Pri proučevanju odvisnosti produktivnosti dela od staža in načina dela morejo predstavljati bloke delavci enake starosti ipd.

Če ponovimo shemo in model delnih blokov v slučajnostnih blokih, dobimo dalje:

b_1		b_2		b_3	
p_1	p_2	p_2	p_1	p_1	p_2
r_1	r_1	r_3	r_2	r_3	r_1
r_3	r_2	r_2	r_1	r_1	r_3
r_2	r_3	r_1	r_3	r_2	r_2

(9.22)

Model delnih blokov v slučajnostnih blokih je:

$$Y_{bpr} = M + (b)+(p)+e_{bp} +(r)+(pr)+ e_{bpr} \qquad (9.23)$$

$$e_{bp} = :N(0, \sigma_1^2)$$

$$e_{bpr} = :N(0, \sigma_2^2)$$

Osnovne tabele so:

$$y_{bpr} \quad y_{bp} \quad y_{br} \quad y_b \qquad (9.24)$$

$$y_{pr} \quad y_p \quad y_r \quad Y$$

Ustrezne pomožne količine so:

$$Q_{bpr} = \sum_{bpr} y^2_{bpr} \qquad Q_{bp} = \frac{1}{r}\sum_{bp} Y^2_{bp} \qquad Q_{br} = \frac{1}{p}\sum_{br} Y^2_{br}$$

$$Q_b = \frac{1}{pr}\sum_b y_b^2 \qquad Q_{pr} = \frac{1}{b}\sum_{pr} Y^2_{pr} \qquad Q_p = \frac{1}{br}\sum_p y_p^2 \qquad (9.25)$$

$$Q_r = \frac{1}{bp}\sum_r y_r^2 \qquad Q = \frac{1}{bpr} Y^2$$

Analiza variance: (9.26)

VV	K	m	s^2
(b)	q_b	b - 1	s_b^2
(p)	q_p	p - 1	s_p^2
pp_1	$q_{bp}-q_b-q_p$	(b-1)(p-1)	s_1^2
(r)	q_r	(r - 1)	s_r^2
(pr)	$q_{pr}-q_p-q_r$	(p-1)(r-1)	s_{pr}^2
pp_2	$q'_{bpr}-q_{bp}-q'_{br}+q_p$	(b-1)(r-1)	s_2^2
Sk	q_{bpr}	bpr-1	

9.11 Bistvo popolnih blokov je v tem, da vsebuje vsak blok vse možne postopke oziroma kombinacije faktorjev, ki jih proučujemo. Delni bloki temu pogoju ne ustrezajo. Od A = p.r postopkov, kolikor jih je v primeru enostavnih delnih blokov, je v samem bloku le r postopkov. V vsakem bloku so namreč postopki, ki vsebujejo vse vrednosti faktorja r na posameznih nivojih faktorja p. Za razliko od popolnih blokov, v katerih nastopajo vsi možni postopki spadajo delni bloki v skupino poskusov z nepopolnimi bloki, ker v posameznih blokih zasledimo le nekatere postopke.

10 DELNI BLOKI Z ŽRTVOVANIMI INFORMACIJAMI.

10.1 Delne bloke uporabljamo iz dveh razlogov. Najpogosteje jih uporabljamo v primerih, kjer zaradi narave enega ali več faktorjev ne moremo izvesti postopke na posameznih enotah neodvisno. Take situacije so bile nakazane v različnih primerih npr. v gnojilnih poskusih določeno agrotehnično mero ne moremo izvesti na osnovnih poskusnih parcelicah, ampak le na večji površini ipd.

Drug razlog za uporabo delnih blokov pa je obsežnost popolnih blokov. Če je število postopkov v faktorskem poskusu veliko, to je bodisi zaradi velikega števila vrednosti posameznih faktorjev, še večkrat pa zaradi velikega števila proučevanih faktorjev, so bloki heterogeni. S tem pa je porušen smisel blokov, ki naj bi bili čimbolj homogeni. Heterogenost blokov vpliva na poskusni pogrešek, zaradi česar so ocene in sklepi iz takega poskusa nezanesljivejši. Delni bloki so bolj homogeni kot popolni bloki, ker so po obsegu manjši. Kot je razvidno iz analize variance v našem primeru, sta komponenti (r) za glavni učinek in interakcija (pr) samo pod vplivom poskusnega pogreška, ki izvira iz variabilnosti v delnih blokih (σ_2^2), medtem ko je glavni učinek faktorja (p) pod vplivom poskusnega pogreška, ki izvira iz variabilnosti v delnih blokih (σ_2^2) in med delnimi bloki (σ_1^2) v skupnem poprečnem iznosu $\sigma_2^2 + \bar{n}\sigma_1^2$.

Z delnimi bloki smo v našem primeru dosegli, da celotna variabilnost poskusnega gradiva vpliva le na glavni učinek faktorja (p), poskusni pogrešek za (r) in (pr) pa se je zmanjšal. Pri delnih blokih z več faktorji je situacija podobna in se zaradi homogenosti delnih blokov manjša poskusni pogrešek vseh komponent, razen za glavni učinek (p) tistega faktorja, ki je v načrtu v čisto slučajnostnem poskusu. Ker je v takih poskusih zanesljivost za glavni učinek nekega faktorja manjša kot za vse druge komponente (za glavne učinke drugih faktorjev in interakcije vseh stopenj) je vprašanje, ali je tak poskus najracionalnejši. Vsebinsko so gotovo najpomembnejše informacije o glavnih učinkih in interakcije med prevelikim številom faktorjev (npr. druge stopnje), medtem ko interakcije izgubljajo na pomenu in predstavljivosti, čim višje stopnje so. Iz tega se je rodila ideja, da prevzame vlogo glavnega učinka faktorja (p) npr. ena izmed interakcij višjega reda, npr. interakcija najvišje stopnje, katere informacijo brez škode žrtvujemo v tem smislu, da je zanjo poskusni pogrešek večji kot za druge komponente.

Problematiko poskusa z žrtvovanimi informacijami proučimo na poskusu 2^3.

10.2 D e l n i   b l o k i   v   p o s k u s u 2^3.

Shema poskusa 2^3 = axbxc v katerem sta faktorja a in b v delnem bloku, faktor c pa čisto slučajnosten, je:

Blok	1		2		3		
	(1)	ac	ab	bc	ab	bc	
	ab	c	a	abc	1	ac	(10.1)
	b	abc	(1)	ac	a	abc	
	a	bc	b	c	b	c	
delni blok	1	2	3	4	5	6	

v delnih blokih 1,3,5 so vse kombinacije faktorjev a in b na nivoju c_o (1, a, b, ab) v delnih blokih 2 4 in 6 pa na nivoju c_1 (c, ac, bc, abc).

Model zgornjega poskusa je:

$$Y_{Bcab} = M+(B)+(c)+e_{Bc}+(a)+(b)+(ab)+(ac)+(bc)+(abc)+e_{Bcba} \quad (10.2)$$

$e_{Bc} = :N(0,\sigma_1^2)$.

$e_{Bcba} = :N(0,\sigma_2^2)$

Pri tem pomeni: (B) blok, a,b in c pa faktorje.

Analiza variance za ta poskus je:

$$
\begin{array}{l|c|l}
VV & m & E(s^2) \\
\hline
(B) & 2 & \sigma_1^2 + 4\sigma_2^2 + 8\sigma_B^2 \\
(c) & 1 & \sigma_1^2 + 4\sigma_2^2 + 12\sigma_c^2 \\
pp_1 & 2 & \sigma_1^2 + 4\sigma_2^2 \\
(a) & 1 & \sigma_1^2 + 12\sigma_A^2 \\
(b) & 1 & \sigma_1^2 + 12\sigma_B^2 \\
(ab) & 1 & \sigma_1^2 + 6\sigma_{AB}^2 \\
(ac) & 1 & \sigma_1^2 + 6\sigma_{Ac}^2 . \\
(bc) & 1 & \sigma_1^2 + 6\sigma_{Bc}^2 \\
(abc) & 1 & \sigma_1^2 + 3\sigma_{ABc}^2 \\
pp_2 & 12 & \sigma_1^2 \\
\hline
Sk & 23 &
\end{array}
\tag{10.3}
$$

Iz $E(S^2)$ je razvidna diskriminirana vloga glavnega učinka faktorja c.

Če premotrimo razporeditev posameznih postopkov po delnih blokih, so v blokih 1.3.5 od glavnega učinka vrednosti c_o (sami -)

in v blokih 2,4,6 c_1 (sami +). Za izračun vseh drugih komponent je v ortogonalnih primerjavah v vsakem delnem bloku enako število podatkov pozitivnih in negativnih. Zato se v okviru delnega bloka razen drugih učinkov, s katerimi je primerjava ortogonalna, uniči tudi vpliv delnega bloka, e_{Bp} , ker je za vse enote bloka isti.

10.3 P o s k u s 2^3 z žrtvovano interakcijo (ABC).

Analogno kot s tem, da vse postopke s c_1 ali + v ortogonalni primerjavi pripišemo enemu delnemu bloku in vse postopke s c_o ali - v ortogonalni primerjavi pripišemo drugemu delnemu bloku in s tem združimo glavni učinek faktorja c z razlikami med delnimi bloki, lahko združimo iz poskusa 2^3 učinek poljubne druge komponente z razlikami med delnimi bloki, če v ene delne bloke vključimo vse postopke s +, v druge delne bloke pa vse postopke z - iz ustrezne ortogonalne primerjave, katere učinek združimo.

Če se odločimo, da z variabilnostjo med delnimi bloki združimo učinek trojne interakcije (ABC), (za katero je interes najmanjši), dobimo naslednji načrt.

Iz sheme zvez med postopki in komponentami dobimo:

	Delni blok 1				Delni blok 2				
Komponenta	abc	a	b	c	ab	ac	bc	(1)	
(ABC)	+	+	+	+	-	-	-	-	
(A)	+	+	-	-	+	+	-	-	
(B)	+	-	+	-	+	-	+	-	(10.4)
(C)	+	-	-	+	-	+	+	-	
(AB)	+	-	-	+	-	+	-	+	
(AC)	+	-	+	-	-	+	-	+	
(BC)	+	+	-	-	-	-	+	+	

Če v delni blok 1 vključimo: abc a b c, v drugega pa: ab ac bc (l), dobimo željeno situacijo, za interakcijo (ABC) so v prvem bloku sami +, v drugem - za vse druge kombinacije pa je število pozitivnih in negativnih vrednosti postopkov enako tako v prvem kot v drugem delnem bloku.

Shema načrta 2^3 v delnih blokih v štirih ponovitvah, v katerem je interakcija (ABC) žrtvovana, je:

(10.5)

Ponovitev	I		II		III		IV	
	a	bc	ac	c	(l)	a	abc	bc
	abc	ac	ab	b	ac	c	c	ab
	b	(l)	bc	a	ab	abc	b	(l)
	c	ab	(l)	abc	bc	b	a	ac
delni blok	1	2	3	4	5	6	7	8

Model poskusa:

(10.6)

$$Y_{Pabc} = M+(P)+(ABC)+e_{P(ABC)}+(A)+(B)\text{Š}(AB)+(C)+(AC)+(BC)+e_{Pabc}$$

Analiza variance pa je:

VV	m	$E(s^2)$
(P) pon.	3	$\sigma_1^2 + 4\sigma_2^2 + 8\sigma_P^2$
(ABC)	1	$\sigma_1^2 + 4\sigma_2^2 + 4\sigma_{ABC}^2$
med delnimi bloki pp_1	3	$\sigma_1^2 + 4\sigma_2^2$
(A)	1	$\sigma_1^2 + 16\sigma_A^2$
(B)	1	$\sigma_1^2 + 16\sigma_B^2$
(AB)	1	$\sigma_1^2 + 8\sigma_{AB}^2$
(C)	1	$\sigma_1^2 + 16\sigma_C^2$
(AC)	1	$\sigma_1^2 + 8\sigma_{AC}^2$
(BC)	1	$\sigma_1^2 + 8\sigma_{BC}^2$
pp_2	18	σ_1^2
Sk	31	

(10.7)

10.4 P o s k u s 2^3 z d e l n o ž r t v o v a n i m i i n t e r a k c i j a m i (AB),(AC),(BC), (ABC)

V prejšnjem načrtu je bila žrtvovana interakcija (ABC) povezana z variabilnostjo med delnimi bloki.

Če je poskus večkrat ponovljen, pa moremo postopke kombinirati tako, da je v vsaki ponovitvi žrtvovana druga interakcija. Tako moremo pri poskusu 2^3 v štirih ponovitvah po vrsti žrtvovati (ABC),(AB),(AC) in (BC). Vsaka izmed štirih interakcij je v tem primeru žrtvovana le v eni ponovitvi oziroma ohranjena v treh ponovitvah. Ker je vsaka interakcija le delno žrtvovana, poskus imenujemo poskus z delno žrtvovanimi interakcijami. Prednost takega načrta je v tem, da se nezanesljivost rezultatov, ki je pri poskusu z eno žrtvovano interakcijo skoncentrirana na eno samo interakcijo, enakomerno porazdeli na več interakcij.

Shema poskusa 2^3 z delno žrtvovanimi interakcijami (ABC),(AB),(AC) (BC) je

Ponovitev	1		2		3		4		
	a	(1)	c	bc	abc	bc	abc	b	
	c	bc	abc	ac	(1)	c	bc	c	(10.8)
	abc	ab	(1)	a	ac	ab	(1)	ab	
	b	ac	ab	b	b	a	a	ac	
Žrtvovana interakcija	(ABC)		(AB)		(AC)		(BC)		
delni blok	1	2	3	4	5	6	7	8	

Analiza variance za zgornji poskus je

VV	m	$E(s^2)$
Med ponovitvami	3	$\sigma_1^2 + 4\sigma_2^2 + 8\sigma_P^2$
Med bloki v ponovitvah	4	$\sigma_1^2 + 4\sigma_2^2 + 1/4\sigma_{ABC}^2 + 1/2\sigma_{AB}^2 + 1/2\sigma_{AC}^2 + 1/2\sigma_{BC}^2$
(A)	1	$\sigma_1^2 + 16\sigma_A^2$
(B)	1	$\sigma_1^2 + 16\sigma_B^2$
(AB)	1	$\sigma_1^2 + 6\sigma_{AB}^2$
(C)	1	$\sigma_1^2 + 16\sigma_C^2$
(AC)	1	$\sigma_1^2 + 6\sigma_{AC}^2$
(BC)	1	$\sigma_1^2 + 6\sigma_{BC}^2$
(ABC)	1	$\sigma_1^2 + 6\sigma_{ABC}^2$
pp_2	17	σ_1^2
Sk	31	

(10.9)

V zgornji shemi poskusa sta po dva delna bloka ene ponovitve sestavljena iz + pozitivnih in - negativnih postopkov v ortogonalni primerjavi komponente, ki je žrtvovana. To moremo kontrolirati s pomočjo matrike predznakov za 2^3 poskus.

Za razliko s poskusom, v katerem je črtovana ena sama interakcija (ABC) in je zanjo analiza manj zanesljiva, ker je njen poskusni pogrešek $\sigma_1^2 + 4\sigma_2^2$, je v primeru delnega žrtvovanja varianca med bloki v ponovitvah rezultat šestih komponent (torej nejasna), možna pa je analiza vseh sedmih komponent s pp_2, ki vsebuje samo variabilnost znotraj delnih blokov.

11. FRAKCIONALNI POSKUSI

11.1 V delnih blokih vzamemo v en blok samo del postopkov celotnega poskusa, da zagotovimo homogenost v blokih, ki je ogrožena, če je število postopkov veliko, ker kombiniramo večje število faktorjev. Poskusi v pravih delnih blokih in v blokih z žrtvovanimi informacijami pa le obsegajo vseh A postopkov poskusa.

Pri večjem številu faktorjev pa je število postopkov lahko tolikšno, da tako obsežen poskus niti v eni ponovitvi ne moremo izvesti z vsemi postopki. Tako se število postopkov v poskusu 2^p z dodatkom novega faktorja podvoji

p	1	2	3	4	5	6	7	8
A	2,	4,	8,	16	32	64	120	256

pri poskusih 3^p pa potroji

p	1	2	3	4	5	6	7	8
A	3	9	27	81	243	729	5103	15309 .

V teh primerih si pomagamo s frakcionalnimi-delnimi poskusi, v katerih v poskus vključimo samo del vseh možnih kombiniranih postopkov.

11.2 Nakažimo splošne principe frakcionalnih poskusov na enostavnem primeru poskusa 2^3, čeprav je tak poskus po obsegu premajhen, da bi bil frakcionalen poskus vsebinsko upravičen.

V poskusu 2^3 imamo skupno 8 stopinj prostosti za oceno splošnega poprečja in sedmih učinkov. Če v poskus ne vključimo vseh osem postopkov, temveč samo nekaj, npr. polovico, se ne moremo izogniti določeni izgubi informacij in natančnosti. Ker je interes za najvišje interakcije zaradi tega, ker jih težko interpretiramo, najmanjši, v našem primeru žrtvujemo informacijo o interakciji najvišjega reda (ABC).

Osem členov ortogonalne primerjave, s katero ocenjujemo interakcijo (ABC) moremo grupirati v dve skupini, prvo s pozitivnimi in drugo z negativnimi znaki:

$$(ABC) \quad + a \quad + b \quad + c \quad + abc \qquad (11.1)$$
$$- 1 \quad - ab \quad - ac \quad - bc$$

Vzemimo v frakcionalen poskus postopke s pozitivnimi predznaki in poglejmo, kaj je ostalo v tem primeru od drugih komponent. Če iz tabele 7.2 sestavimo tabelo predznakov za ta primer, dobimo:

(11.2)

Postopek	(1) M	(A)	(B)	(AB)	(C)	(AC)	(BC)	(ABC)
				Matrika učinkov				
a	+	+	-	-	-	-	+	+
b	+	-	+	-	-	+	-	+
c	+	-	-	+	+	-	-	+
abc	+	+	+	+	+	+	+	+

Interakcija (ABC) je v tem delu identična z M, ki ima tudi same +. Če primerjamo predznake za druge učinke, vidimo, da so kontrasti po dveh komponent identični in sicer:

$$\begin{aligned} (ABC) &\equiv (1) \\ (A) &\equiv (BC) \\ (B) &\equiv (AC) \\ (AB) &\equiv (C) \end{aligned} \qquad (11.3)$$

Drugače pa so druge komponente med seboj ortogonalne npr. (A) in (B)

$$(+1)(-1) + (-1)(+1) + (-1)(-1) + (+1)(+1) = 0$$

Enako velja za druge kombinacije.

Vsaki komponenti ustreza druga komponenta, ki ima v okviru frakcionalnega poskusa enake predznake koeficiente primerjave. Vsaka komponenta ima v tem smislu svojega "družabnika", ki ima to lastnost, da sta si v okviru frakcionalnega poskusa primerjavi, ki sta si družabnika, identični in zato njihova učinka spojena. Določeni komponenti dobimo družabnika, če jo formalno pomnožimo z ABC in vzamemo, da je $A^2 = 1$, $B^2 = 1$, $C^2 = 1$

npr. komponenti (A) ustreza "družabnik"

$$(A).(ABC) \equiv A^2\ BC \equiv (BC) \qquad (11.4)$$

Če izračunamo za določeno komponento ortogonalno primerjavo, zaradi značaja družabnikov, primerjava vsebuje skupen učinek ustreznih družabnikov, učinek drugih komponent pa se zaradi ortogonalnosti uniči.

S frakcionalnim poskusom ne moremo dobiti torej čistih učinkov posameznih komponent, ampak le za skupen učinek družabnikov.

11.3 Splošna praksa, ki omogoča uporabo delnih poskusov v praksi je, da zanemarimo vpliv interakcije višjih redov.

Če vzamemo za primer frakcionalni poskus 2^6, pri čemer vključimo v poskus $2^6 = 32$ postopkov določujočega kontrasta - najvišje interakcije ABCDEF. Navedimo komponente poskusa in njihove družabnike.

	Družabnika	Pojasnjuje	Družabnika		Pojasnjuje
(1)	ABCDEF	M			
(A)	BCDEF	(A)	ABC	DEF	e_1
B	ACDEF	(B)	ABD	CEF	e_2
C	ABDEF	(C)	ABE	CDF	e_3
D	ABCEF	D	ABF	CDE	e_4
E	ABCDF	E	ACD	BEF	e_5
F	ABCDE	F	ACE	BDF	e_6
AB	CDEF	AB	ACF	BDE	e_7
AC	BDEF	AC	ADE	BCF	e_8
AD	BCEF	AD	ADF	BCE	e_9
AE	BCDF	AE	AEF	BCD	e_{10}
AF	BCDE	AF			
BC	ADEF	BC			
BD	ACEF	BD			
BE	ACDF	BE			
BF	ACDE	BF			
CD	ABEF	CD			
CE	ABDF	CE			
CF	ABDE	CF			
DE	ABCF	DE			
BF	ABCE	BF			
EF	ABCD	EF			

(11.5)

V prikazanem primeru moremo s predpostavko, da interakcij tretje in višjih stopenj nič, iz opredeljujočega kontrasta interakcije ABCDEF:

1 ab ac ad ae af bc bd be bf cd ce cf de df

ef abcd abce abcf abde abdf abcf acde acdf acef adef

bcde bedf bcef bdef edef abcdef (11.6)

analizirati vseh$\binom{6}{1}$ = 6 glavnih učinkov, vseh $\binom{6}{2}$ = 15 dvojnih interakcij, a za oceno poskusnega pogreška ostane še $\frac{1}{2}\binom{6}{3}$ = 10 parov družabnikov interakcij tretje stopnje. Iz zgornjega primera je jasno razvidno tudi formiranje komponente - družabnika: Dobimo jo bodisi iz zveze npr. za (BC)

$$ABCDEF \cdot BC = A.B^2.C^2.DEF = ADEF \qquad (11.7)$$

ali pa tako, da je določeni interakciji ustrezen družabnik interakcija, katere faktorji dopolnjujejo do kompleta faktorjev. Od kompleta ABCDEF je torej BC ustrezen družabnik ADEF, ker je $A\,{}^{BC}\,{}_{DEF}$.

11.4 Podobno kot s 1/2-ponovitvijo poskusa, moremo izvesti tudi 1/4-ponovitev poskusa, v katerem je 2^{p-2} postopkov. Frakcionalne-delne poskuse moremo razširiti tudi na faktorske poskuse 3^p, v katerih imajo faktorji po 3 vrednosti, razširjeni pa so frakcionalni poskusi za primere s faktorji z več kot tremi vrednostmi.

12. ANALIZA KOVARIANCE.

12.1 Faktorje smo v vseh dosedanjih primerih poskusnih načrtov dajali klasifikatorno s tem, da so faktorji dani le z nekaj, čeprav numeričnimi vrednostmi. Dosti numeričnih faktorjev pa je podanih za posamezno poskusno enoto z vrednostmi, ki so na določenem razmaku zvezne in morejo zavzeti vse vrednosti v tem razmaku. Tako moremo npr. dohodek pri proučevanju porabe določenega izdelka podati z zneskom X dohodka za vsako posamezno raziskovano osebo, prav tako starost pri proučevanju navad prebivalstva. Take narave so tudi podatki o proizvodnih faktorjih pri proučevanju proizvodnje, proizvodnosti dela ipd.

Probleme, v katerih iščemo odvisnost kriterialnega znaka od faktorialnih znakov rešujemo z regresijsko analizo. Faktorji, katerih vrednosti moremo izraziti numerično pa so v poskusih lahko tudi faktorji, ki za raziskavo niso pomembni in hočemo njih vpliv izločiti. To pa izvedemo z analizo kovariance tako, da iz variacije izločimo tisti del, ki gre na račun kovariance med kriterialnom znakom y in kovariantnim znakom x.

Če raziskujemo, ali so razlike med porabo določenega izdelka v odvisnosti od tega, ali oseba živi v mestu ali na deželi, je gotovo višina dohodka faktor, ki vpliva na porabo, a je za raziskavo v konkretnem primeru nepomemben. Podoben primer je z velikostjo parcele pri poskusnih v agronomskih raziskavah. Da izločimo vpliv velikosti parcele na donos, vključimo velikost v opredeljujoče pogoje tako, da vzamemo poskusne parcelice vse enako velike. Če pa gre za vzorčni poskus, v katerem smo izbrali parcele iz obstoječih parcel, pa velikost parcele variira. Vpliv velikosti parcele na donos in variacijo, ki izvira iz tega, izločimo z analizo kovariance. Podobno moremo izločiti vpliv števila dreves, rastlin ipd. Pri bioloških raziskavah je tak spreminjajoč faktor, ki ga obravnavamo z analizo kovariance, npr. teža poskusnih živali.

Če je splošen kompleks vpliva faktorjev, ki so obravnavani kot faktorji s končnim številom vrednosti nivojev M_G, je linearni model poskusa

$$y_{Gi} = M_G + e_{Gi} \qquad (12.1)$$

Če poskusni pogrešek razdelimo v dva dela: kovariantni, ki je izražen z linearnim delom $b(x_{Gi} - \bar{x})$ in $e_{Gi.x}$, ki je slučajnostni po izločitvi vpliva faktorja x, dobimo model za analizo kovariance

$$y_{Gi} = M_G + b(x_{Gi} - \bar{x}) + e_{Gi.x} \qquad (12.2)$$

Če izločamo vpliv dveh ali večih kovariant spremenljivk, govorimo o multipli analizi kovariance, ki ima v primeru, da izločamo, dva spremljajoča faktorja x in u za naslednji model:

$$y_{Gi} = M_G + b_1(x_{Gi} - \bar{x}) + b_2(U_{Gi} - \bar{U}) + e_{Gi.xU} \qquad (12.3)$$

12.2 Splošen model čisto slučajnostnega poskusa z kovariantnim znakom x, je:

$$y_{Ai} = M + (A) + b(x_{Ai} - \bar{x}) + e_{Ai.x} \qquad e_{Ai.x} = :N(0, \sigma_e^2) \qquad (12.4)$$

Iz modela vidimo, da je poskusni pogrešek iz navadnega čisto slučajnostnega poskusa razstavljen v dve komponenti kovariatni $b(x_{Ai} - \bar{x})$ in slučajnostni del e_{Ai}. Analizo kovariance pri čisto slučajnostnem poskusu pa obračunamo po naslednjem postopku:

iz podatkov x_{Ai} in y_{Ai} izračunamo ustrezne tabele kot pri čisto slučajnostnem poskusu za y_{Ai}

$$\begin{array}{cc} x_{Ai} & y_{Ai} \\ x_A & y_A \\ x & y \end{array} \qquad (12.5)$$

Iz teh osnovnih podatkov izračunamo pomožne količine, ki so sorodne Q_z.

$$Q^{xx}_{Ai} = \sum_{Ai} x^2_{Ai} \qquad Q^{xy}_{Ai} = \sum_{Ai} x_{Ai} y_{Ai} \qquad Q^{yy}_{Ai} = \sum_{Ai} y^2_{Ai}$$

$$Q^{xx}_{A} = \frac{1}{\bar{n}} \sum_{A} x^2_{A} \qquad Q^{xy}_{A} = \frac{1}{\bar{n}} \sum_{A} x_{A} y_{A} \qquad Q^{yy}_{A} = \frac{1}{\bar{n}} \sum_{A} y^2_{A} \quad (12.6)$$

$$Q^{xx} = \frac{1}{\bar{n}A} x^2 \qquad Q^{xy} = \frac{1}{\bar{n}A} x\, y \qquad Q^{yy} = \frac{1}{\bar{n}A} y^2$$

Iz teh količin izračunamo po znanem pravilu q^{xx}_u, q^{xy}_u in q^{yy}_u tako, da od Q_u odštejemo ustrezen Q npr. $q^{xy}_{Ai} = Q^{xy}_{Ai} - Q^{xy}$

Sestavimo trojno tabelo xx, xy in yy po isti shemi kot pri analizi variance za čisto slučajnostni poskus:

V stolpcu 1, 2 in 3 so nakazani osnovni elementi analize variance pri čisto slučajnostnem poskusu. Stolpec 3 je simboličen in nakazuje, kako izračunamo količine v stolpcih 4, 5, 6. V stolpcu 3 nakazane operacije izvedemo za stolpce 4, 5 in 6 na xx xy in yy. V stolpcu 7 so vsote kvadratov kriterialnega znaka y če smo eliminirali vpliv x. $K^{yy}_{e.x}$ in $K^{yy}_{A+e.x}$ izračunamo s korekturo $\frac{(K^{xy})^2}{K^{xx}}$ korigirano vsoto kvadratov za kovariatni znak.

$K^{yy}_{A.x} = K^{yy}_{A+e.x} - K^{yy}_{e.x}$ pa dobimo kot razliko $K^{yy}_{A+e.x} \cdot K^{yy}_{e.x}$ ima m_e-1 stopinj prostosti. m_e je zmanjšan za eno stopinjo prostosti zaradi korekture $B(x_{Ai} - \bar{x})$, ki ima eno stopinjo prostosti. Ocene varianc $s^{yy}_{A.x}$ in $s^{yy}_{e.x}$ dobimo, če podatke iz stolpca 7 delimo z ustreznimi stopinjami prostosti iz stolpca 8. F_A preskus izvršimo na standardni način in ga primerjamo s kritičnimi vrednostmi $F(m_1 = A-1;\ m_e = n-A-1)$.

(12.7)

1	2	3	4	5	6	7	8	9	10
VV	m	K	K^{xx}	K^{xy}	K^{yy}	$K^{yy}_{.x}$	m	$s^{yy}_{.x}$	F
A	$m_A = A-1$	$K_A = q_A$	K^{xx}_A	K^{xy}_A	K^{yy}_A	$K^{yy}_{A.x} = K^{yy}_{A+e.x} - K^{yy}_{e.x}$	$m_A = A-1$	$s^{yy}_{A.x}$	$F'_A = s^{yy}_{A.x}/s^{yy}_{e.x}$
pp	$m_e = n-A$	$K_e = q_{Ai} - q_A$	K^{xx}_e	K^{xy}_e	K^{yy}_e	$K^{yy}_{e.x} = K^{yy}_e - \frac{(K^{xy}_e)^2}{K^{xx}_e}$	$m_e - 1$	$s^{yy}_{e.x}$	1
	$m = n-1$	$K_{A+e} = K_A + K_e$	K^{xx}_{A+e}	K^{xy}_{A+e}	K^{yy}_{A+e}	$K^{yy}_{A+e.x} = K^{yy}_{A+e} - \frac{(K^{xy}_{A+e})^2}{K^{xx}_{A+e}}$	$m_A + m_e - 1$		

Če iz modela izračunamo popreča, dobimo:

$$\bar{y}_A = M + (A) + b(\bar{x}_A - \bar{x}) + \bar{e}_{A.x} \qquad b = K_e^{xy}/K_e^{xx} \tag{12.8}$$

$$\text{Ocena } (\widehat{M+(A)}) = \bar{y}_A - b(\bar{x}_A - \bar{x}) = \bar{y}_{A.x}$$

$$\text{var}\left[(M+(A) = \bar{y}_{A.x})\right] = s_{e.x}^{yy}\left[\frac{1}{n_A} + \frac{(\bar{x}_A-\bar{x})^2}{K_e^{xx}}\right] \tag{12.9}$$

$$\text{var}\,\Delta\,\bar{y}_{A.x} = s_{e.x}^{yy}\left[\frac{1}{\bar{n}_1} + \frac{1}{\bar{n}_2} + \frac{(\bar{x}_{A1} - \bar{x}_{A2})^2}{K_e^{xx}}\right] \tag{12.10}$$

ali če je $n_A = \bar{n} = \text{const}$

$$\text{var}\,\Delta\,\bar{y}_{A.x} = s_{e.x}^{yy}\left[\frac{2}{\bar{n}} + \frac{(\bar{x}_{A1} - \bar{x}_{A2})^2}{K_e^{xx}}\right] \tag{12.11}$$

12.3 Za primer analize kovariance pri enostavnem slučajnostnem poskusu vzemimo simulirani poskus treh postopkov $A_1 = 40$ $A_2 = 10$ $A_3 = -50$. Kot kovariatno spremenljivko vzemimo linearen trend, ki se pojavlja v poskusnem gradivu in se od enote do enote veča za 10.

- Iz primera simuliranega čisto slučajnostnega poskusa iz odstavka 4.23 so osnovni podatki x_{Ai} in y_{Ai} naslednji:

		A_1			A_2			A_3		
x_{Ai} y_{Ai}	i	x_{1i}	y_{1i}	i	x_{2i}	y_{2i}	i	x_{3i}	y_{3i}	
	1	- 40	101	5	10	125	3	- 20	33	
	2	- 30	108	8	30	143	4	- 10	36	
	5	0	142	9	40	146	6	20	66	
x_A y_A		- 70	351		80	414		- 10	135	$x = 0$
$\bar{x}_A$ $\bar{y}_A$		- 23,333	117		26,607	138		- 3.333	45	$y = 900$

Pomožne količine Q za analizo kovariance so:

$$Q^{xx}_{Ai} = \sum_{Ai} x^2_{Ai} = (-40)^2+(-30)^2+ \dots (+20)^2 = 6000$$

$$Q^{xy}_{Ai} = \sum_{Ai} x^2_{Ai} y_{Ai} = (-40)(101) + (-30).108+\dots +(20).66 = 4400$$

$$Q^{yy}_{Ai} = \sum_{Ai} y^2_{Ai} = 101^2 + 108^2 + \dots + 66^2 = 106160$$

$$Q^{xx}_{A} = \frac{1}{\bar{n}}\sum_{A} x^2_{A} = \frac{1}{3}\left[(-70)^2+(80)^2+(-10)^2\right] = 3800$$

$$Q^{xy}_{A} = \frac{1}{\bar{n}}\sum_{A} x_A y_A = \frac{1}{3}\left[(-70).351+80.414+(-10).135\right] = 2400$$

$$Q^{yy}_{A} = \frac{1}{\bar{n}}\sum_{A} y^2_{A} = \frac{1}{3}\left[351^2+414^2+135^2\right] = 104274$$

$$Q^{xx} = \frac{1}{\bar{n}A}.x^2 = \frac{1}{3.3}.0^2 = 0$$

$$Q^{xy} = \frac{1}{\bar{n}A}\, x.y = \frac{1}{3.3}\, .\, 0.900 = 0$$

$$Q^{yy} = \frac{1}{\bar{n}\,A}\, y^2 = \frac{1}{3.3}\, 900^2 = 90000$$

Korigirane vrednosti q_z so $Q_z - Q$

$$q^{xx}_{Ai} = Q^{xx}_{Ai} - Q^{xx} = 6000 - 0 = 6000$$

$$q^{xy}_{Ai} = Q^{xy}_{Ai} - Q^{xy} = 4400 - 0 = 4400$$

$$q^{yy}_{Ai} = Q^{yy}_{Ai} - Q^{yy} = 106160 - 90000 = 16160$$

$$q^{xx}_{A} = Q^{xx}_{A} - Q^{xx} = 3800 - 0 = 3800$$

$$q^{xy}_{A} = Q^{xy}_{A} - Q^{xy} = 2400 - 0 = 2400$$

$$q^{yy}_{A} = Q^{yy}_{A} - Q^{yy} = 104274 - 90000 = 14274$$

Analiza kovariance:

	K^{xx}	K^{xy}	K^{yy}	$K^{yy}_{.x}$	m	$s^2_{.x}$	F
A	3800	2400	14274	$K^{yy}_{A.x}=12933-68=12865$	2	$s^2_{A.x}=6432,5$	F=473,0
pp	2200	2000	1886	$K^{yy}_{e.x}=1886-\frac{2000^2}{2200}=68$	5	$s^2_{e.x}=13,6$	1
Sk	6000	4400	16160	$K^{yy}_{e+A.x}=16160-\frac{4400^2}{6000}=12933$			

$$\bar{y}_{A.cov} = \bar{y}_A - b(\bar{x}_A - \bar{x}) = \bar{y}_A - 0.909\, x_A \qquad b = \frac{K_{xy}}{K_{xx}} = \frac{2000}{2200} = 0.909$$

$$\bar{y}_{1.cov} = 117-0.909 = 138,2 \qquad A_1 = 38,2$$

$$\bar{y}_{2.cov} = 138 - 0.909(26.677) = 113,8 \qquad A_2 = 13,8$$

$$\bar{y}_{3.cov} = 45 - 0.909(-3.333) = 48,0 \qquad A_3 = - 52.0$$

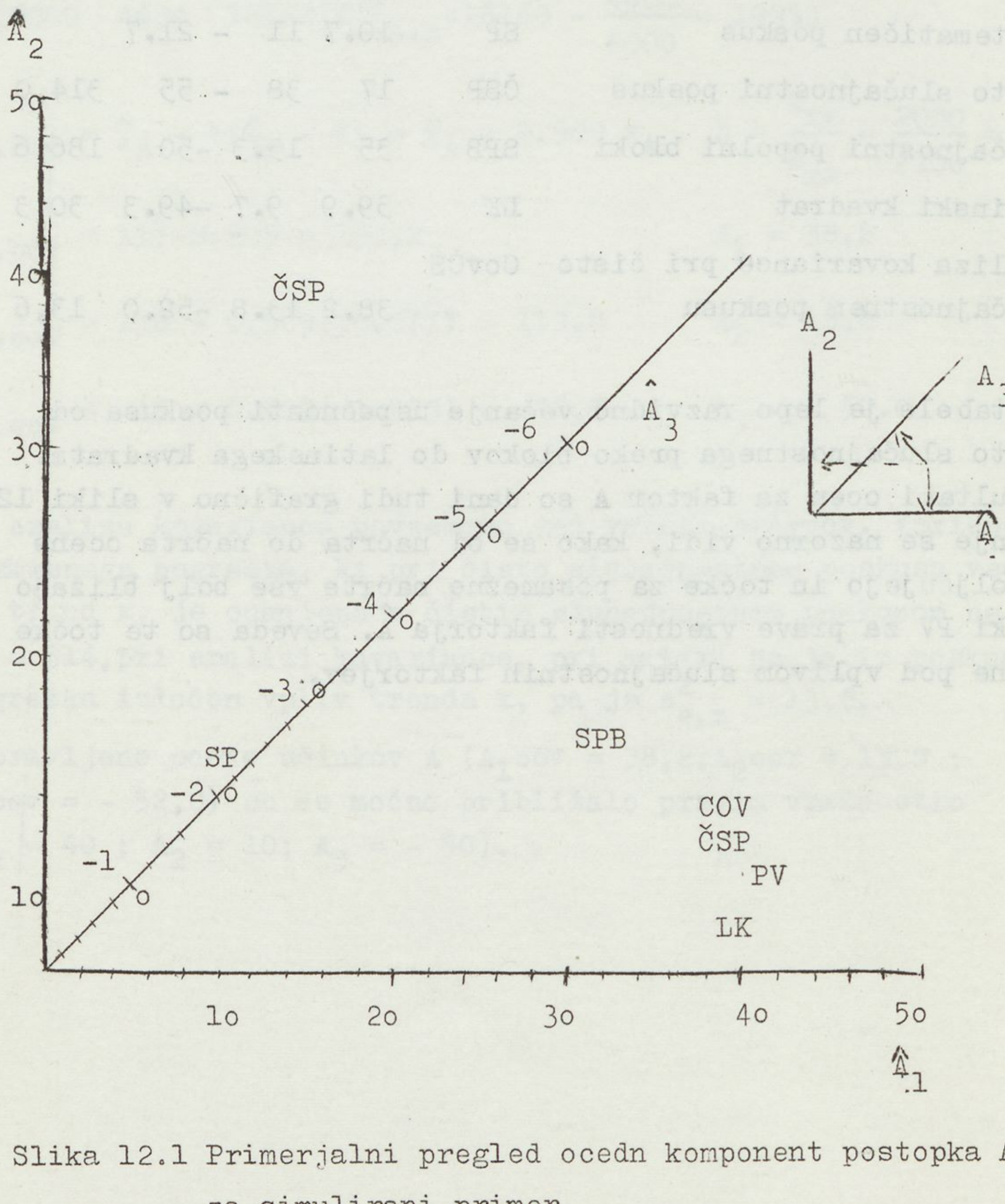
Iz analize kovariance povzamemo več važnih sklepov. Variance poskusnega pogreška, ki pri čisto slučajnostnem poskusu vsebuje še trend x, je ocenjena s čistim slučajnostnim poskusom na $s_e^2 = 314$, pri analizi kovariance, pri kateri pa je iz poskusnega pogreška izločen vpliv trenda x, pa je $s^2_{e.x} = 13,6$.

Popravljene ocene učinkov A (A_1cov = 38,2; A_2cor = 13,8 ; A_3cov = - 52,0) so se močno približale pravim vrednostim (A_1 = 40 ; A_2 = 10; A_3 = - 50).

12.4 Če strnemo rezultate za simulirani primer iz vseh obravnavanih metod, dobimo naslednjo tabelo ocen efektov, uspešnosti E in

		$\hat{A}_1$	$\hat{A}_2$	$\hat{A}_3$	s_e^2	E
Prave vrednosti	PV	40	10	- 50	-	-
Sistematičen poskus	SP	10.7	11	- 21.7		
Čisto slučajnostni poskus	ČSP	17	38	- 55	314,0	1
Slučajnostni popolni bloki	SPB	35	15.3	-50	186,6	3,16
Latinski kvadrat	LK	39.9	9.7	-49.3	30,3	18,94
Analiza kovariance pri čisto slučajnostnem poskusu	CovČS	38.2	13.8	-52.0	13,6	

Iz tabele je lepo razvidno večanje uspešnosti poskusa od čisto slučajnostnega preko blokov do latinskega kvadrata. Rezultati ocen za faktor A so dani tudi grafično v sliki 12.1. Iz nje se nazorno vidi, kako se od načrta do načrta ocene izboljšujejo in točke za posamezne načrte vse bolj bližajo toški PV za prave vrednosti faktorja $\hat{A}$. Seveda so te točke kot ocene pod vplivom slučajnostnih faktorjev.



Slika 12.1 Primerjalni pregled ocedn komponent postopka A za simulirani primer

12.5 Analiza kovariance v slučajnostnih popolnih blokih.

Analogno čisto slučajnostnem poskusu vpeljemo kovariatno kontrolno spremenljivko v model slučajnostnih popolnih blokov, če v model slučajnostnih popolnih blokov vnesemo linearni člen $b(x_{BA} - \bar{x})$

$$y_{Ba} = M + (B) + (A) + b(x_{BA} - \bar{x}) + e_{BA.x} \qquad (12.12)$$

Poskusni pogrešek se torej tudi tu razdeli v dva dela: pojasnjeni - kovariatni $b(x_{BA} - \bar{x})$ in preostali - slučajnostni e_{BA}. Analizo kovariance obračunamo podobno kot v primeru čisto slučajnostnega poskusa, tako, da vzamemo za osnovo shemo analize variance za poskus v slučajnostnih popolnih blokih.

Analiza kovariance za slučajnostne popolne bloke je zelo slična analizi kovariance pri čistem slučajnostnem poskusu in je njena svota enaka kot pri ČSP.

Analiza kovariance pri slučajnostnih blokih :

VV	K	K^{xx}	K^{xy}	K^{yy}	m	$K^{yy}_{.x}$	m	s^2	F
(B)	$K_B = q_B$	K_B^{xx}	K_B^{xy}	K_B^{yy}	$m_B = B-1$				
(A)	$K_A = q_A$	K_A^{xx}	K_A^{xy}	K_A^{yy}	$m_A = A-1$	$K^{yy}_{A.x} = K^{yy}_{A+e.x} - K^{yy}_{e.x}$	$A-1$	$s^{yy}_{A.x}$	$F'_A = \dfrac{s^{yy}_{A.x}}{s^{xx}_{e.x}}$
pp	$K_e = q_{BA}-q_B-q_A$	K_e^{xx}	K_e^{xy}	K_e^{yy}	$m_e = (B-1)(A-1)$	$K^{yy}_{e.x} = K^{yy}_e - \dfrac{(K^{xy}_e)^2}{K^{xx}_e}$	$m'_e = m_e - 1$	$s^{2yy}_{e.x}$	
	$K_{A+e} = K_A + K_e$	K^{xx}_{A+e}	K^{xy}_{A+e}	K^{yy}_{A+e}	$m=(A-1)B$	$K^{yy}_{A+e.x} = K^{yy}_{e+A} - \dfrac{(K^{xy}_{A+e})^2}{K^{xx}_{A+e}}$	m_A+m_e-1		

12.6 A n a l i z a   k o v a r i a n c e   p r i   l a t i n s k e m   k v a d r a t u .

Iz prejšnjih primerov moremo razširiti tehniko obračuna analize kovariance pri poljubnem načrtu, med drugim tudi za latinski kvadrat. Po standardni shemi za analizo variance pri ustreznem načrtu obračunamo K^{xx} K^{xy} in K^{yy}. Iz K_A in K_e dobimo $K_A+K_e = K_{A+e}$. Za K_A in K_{A+e} obračunamo reducirani vsoti kvadratov

$K^{yy}_{e.x} = K^{yy}_{e} - \frac{(K^{xy}_{e})^2}{K^{xx}_{e}}$ in $K^{yy}_{A+e.x} - \frac{(K^{xy}_{A+e})^2}{K^{xx}_{A+e}}$ iz njih pa reduciramo vsoto kvadratov $K^{yy}_{A.x} = K^{yy}_{A+e.x} - K^{yy}_{e.x}$. Iz teh količin izračunamo

$$s^{yy}_{A.x} = \frac{K^{yy}_{A.x}}{A-1} \quad \text{in} \quad s^{yy}_{e.x} = \frac{K^{yy}_{e}}{m_e-1}$$

pa poskusni izraz

$$F_A = \frac{s^{yy}_{A.x}}{s^{yy}_{e.x}} \text{, ki ga preskusimo z } F_{\alpha} \; (m_1=m_A; \; m_2= m_e-1)$$

Če po tem postopku reproduciramo analizo kovariance v latinskem kvadratu, dobimo:

V shemi je razen za faktor (A) nakazana analiza kovariance tudi za (V) in (S), čeprav običajno analiziramo le faktor A. Razširjenje analize kovariance tudi na (V) in (S) je izvedeno zato, da vidimo, kako moremo razširiti analizo kovariance na več komponent. Razširitev za S in V je v našem primeru smiselna, če sta glede na raziskavo tudi faktor (V) in (S) za raziskavo vsebinsko pomembna.

Za korigirano poprečje

$\bar{y}_{A.cov} = \bar{y}_A - b(\bar{x}_A - \bar{x})$ je ocena variance

$$\mathrm{var}\ \bar{y}_{A.cov.} = s^{yy}_{e.x}\left(\frac{1}{A} + \frac{(\bar{x}_A - \bar{x})^2}{K^{xx}_e}\right)$$

za razliko dveh poprečij

$\bar{y}_{A.cov} = \bar{y}_{A_2.cov} - \bar{y}_{A_1.cov}$ pa $\mathrm{var}\,\Delta\bar{y}_{A.cov} = s^{yy}_{e.x}\left[\frac{2}{A} + \frac{(\bar{x}_{A2} - \bar{x}_{A1})^2}{K^{xx}_e}\right]$

Analiza kovariance pri latinskem kvadratu

(12.13)

VV	K	K^{xx}	K^{xy}	K^{yy}	m	K^{yy}	m_r	s^{yy}_{x}	F
V	$K_V = q_v$	K^{xx}_V	K^{xy}_V	K^{yy}_V	$m_V = A-1$	$K^{yy}_{U.x} = K^{yy}_{U+e.x} - K^{yy}_{e.x}$	$A-1$	$s^{yy}_{V.x}$	F_V
S	$K_S = q_S$	K^{xx}_S	K^{xy}_S	K^{yy}_S	$m_S = A-1$	$K^{yy}_{S.x} = K^{yy}_{S+e.x} - K^{yy}_{e.x}$	$A-1$	$s^{yy}_{S.x}$	F_S
A	$K_A = q_A$	K^{xx}_A	K^{xy}_A	K^{yy}_A	$m_A=(A-1)$	$K^{yy}_{A.x} = K^{yy}_{A+e.x} - K^{yy}_{e.x}$	$(A-1)$	$s^{yy}_{A.x}$	F_A
pp	$K_e=q_{VSA}-q_V$	K^{xx}_e	K^{xy}_e	K^{yy}_e	$m_e=(A-1)(A-2)$	$K^{yy}_{e.x} = K^{yy}_e - \frac{(K^{xy}_e)^2}{K^{xx}_e}$	A_φ^2-3A+1	$s^{yy}_{e.x}$	1
	$K_{A+e} = K_A + K_e$	K^{xx}_{A+e}	K^{xy}_{A+e}	K^{yy}_{A+e}	$m_{A+e}=(A-1)^2$	$K^{yy}_{A+e.x} = K^{yy}_{A+e} + \frac{(K^{xy}_{A+e})^2}{K^{xx}_{A+e}}$	$A(A-2)$		
	$K_{V+e} = K_V+K_e$	K^{xx}_{V+e}	K^{xy}_{V+e}	K^{yy}_{V+e}	$m_{V+e}=(A-1)^2$	$K^{yy}_{V+e.x} = K^{yy}_{V+e} - \frac{(K^{xx}_{V+e})^2}{K^{xx}_{V+e}}$	$A(A-2)$		
	$K_{S+e}=K_S + K_e$	K^{xx}_{S+e}	K^{xy}_{S+e}	K^{yy}_{S+e}	$m_{S+e}=(A-1)^2$	$K^{yy}_{S+e.x} \doteq K^{yy}_{S+e} - \frac{(K^{xy}_{S+e})^2}{K^{xx}_{S+e}}$	$A(A-2)$		