

## Interpretacija diagramov

V članku je opisano, kaj lahko trdimo in česa ne moremo zagovarjati pri interpretaciji diagramov, če se držimo principov matematične statistike.

Opisana je tudi metoda ortogonalnih polinomov, s katero določamo, katera potenca polinoma je še statistično pomembna.

V Železarskem zborniku, pa tudi v drugih revijah se pogosto pojavljajo članki, v katerih avtorji delajo zaključke na podlagi rezultatov meritev, ki so prikazani z diagrami.

Diagramu navadno sledijo zaključki o povezanosti med odvisnimi in neodvisnimi spremenljivkami, ki pa so pogosto nekritični.

Velikokrat se namreč vidi, da določene trditve diagram sploh ne podpira. To velja tudi za nekatere referate na XV. posvetovanju metalurgov v Portorožu v oktobru 1972. Kritičnemu bralcu se vsiljuje vprašanje, kaj je prav: ali držijo zaključki, ki so morda avtorju znani iz literature, ali njegovi eksperimentalno dobljeni diagrami.

To zavaja tudi druge raziskovalce, če verjamejo razlagi, ki se morda pozneje izkaže kot nepravilna. To pa vnaša v naše delo nezaupanje.

Če bi vse zaključke statistično preverili, bi v svoje delo vnesli več kritičnosti in ne bi delali nepotrebnih napak.

V naslednjih primerih si bomo ogledali, kako lahko interpretiramo rezultate meritev, ki jih ponazarjamo z diagrami.

### Linearna regresija

Oglejmo si, kaj na primer lahko povemo o povezavi med spremenljivkama X in Y, če smo si X izbrali (neodvisna spremenljivka), Y pa merili (odvisna spremenljivka). Rezultate 11 meritev (tabela I) smo narisali kot točke na sliki 1.

Narišimo diagram!

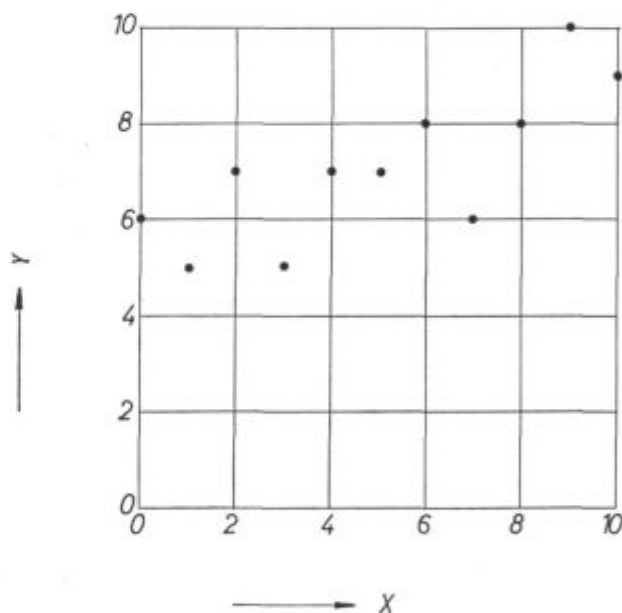
Predpostavimo, da že iz literature vemo, da je zveza linearna, t. j. da gre za premico s smernim koeficientom okrog 0,4.

Premico je treba narisati tako, da se njen potek kar najbolj ujema s potekom točk na sliki 1. Matematično to zahtevo izpolnimo tako, da določimo parametre  $b_0$  in  $b_1$  v regresijski enačbi  $\hat{Y} = b_0 +$

$+ b_1 X$  iz pogoja  $\sum_{i=1}^n (\hat{Y}_i - Y_i)^2 = \text{minimum}^1$ .

Tabela I

X	Y
0	6
1	5
2	7
3	5
4	7
5	7
6	8
7	6
8	8
9	10
10	9



Za naš primer dobimo sledeče vrednosti:

$b_0 = 5,227$ ,  $b_1 = 0,373$ , koeficient množične korelacije  $R = 0,781$ . To pomeni, da smo z linearno zvezo  $\hat{Y} = 5,227 + 0,373 X$  pojasnili približno 61 % celotnega razsipanja vrednosti Y.

Dobili smo torej enačbo premice, ki pa velja le za naš vzorec 11 meritev. Kaj lahko povemo o linearni zvezi med X in Y, ki naj velja za celo populacijo, t. j. za množico točk, ki bi jih lahko narisali, če bi jih izmerili.

O karakterističnih parametrih populacije lahko povemo le interval vrednosti, v katerem lahko dani parameter z neko določeno verjetnostjo pričakujemo.

Regressijsko enačbo za populacijo lahko zapišemo v obliki  $\hat{Y} = \beta_0 + \beta_1 X$ . Določili bomo interval, v katerem z 90 % gotovostjo pričakujemo vrednost parametra  $\beta_1$ .

Se prej se pa vprašajmo, kakšne informacije lahko dobimo o naši populaciji parov (X, Y), če smo naredili omenjenih 11 meritev. Ali lahko na podlagi eksperimenta trdimo, da gre tudi v našem primeru za koeficient  $\beta_1 = \beta'_1 = 0,4$ , kot navaja literatura?

Označimo standardno napako ocene z  $S_E$ , ki je definirana z relacijo:

$$S_E = \sqrt{\frac{\sum_{i=1}^n (\hat{Y}_i - Y_i)^2}{n - 2}}$$

Standardno napako regresijskega koeficienta izračunamo po formuli:

$$S_{b_1} = \frac{S_E}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2}}$$

Izračunajmo parameter t:

$$t = \frac{b_1 - \beta'_1}{S_{b_1}}$$

za katerega velja Studentova distribucija (tabela II.). Hipotezo, da je naš vzorec iz populacije z vrednostjo  $\beta_1 = \beta'_1 = 0,4$  bomo zavrgli v primeru, če bo  $|t| \geq t_{(1 - \frac{\alpha}{2}), (n-2)}$ , pri čemer je  $\alpha$  verjetnost za napako prve vrste<sup>2</sup>, n pa število meritev v vzorcu.

V našem primeru je

- $S_E = 1,034$
- $S_{b_1} = 0,0986$
- $b_1 = 0,373$
- $\beta'_1 = 0,400$
- $t = 0,277$

Če se odločimo za  $\alpha = 0,05$ , je  $t_{(1 - \frac{\alpha}{2}), (n-2)} = t_{0,975; 9} = 2,262$ , kar je precej več kot 0,277.

Hipoteze torej ne moremo zavreči. To pa še ne pomeni, da smo že dokazali, da je  $\beta_1$  za našo populacijo enak  $\beta'_1 = 0,4$ . Upoštevati je treba še operacijske karakteristike<sup>2</sup> za tak primer. Vprašajmo se, ali nismo pri tem, ko smo sprejeli hipotezo, da je  $\beta_1 = \beta'_1 = 0,4$ , naredili preveliko napako druge vrste.

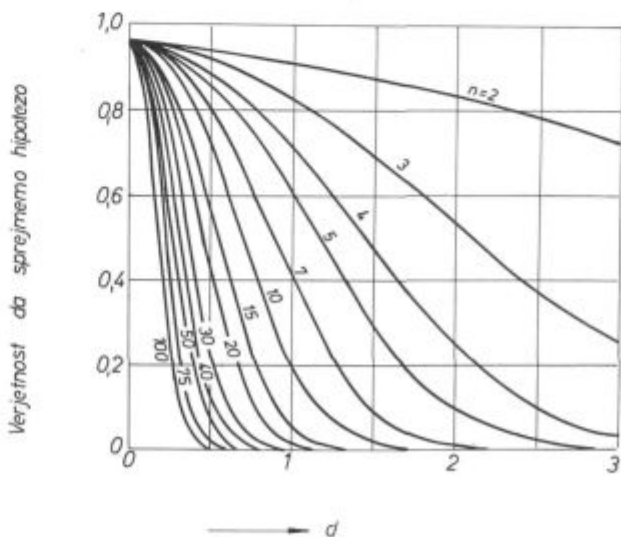
Tabela II

$t_{v, n}$	$v = 0,950$	$v = 0,975$
n		
1	6,314	12,706
2	2,920	4,303
3	2,353	3,182
4	2,132	2,776
5	2,015	2,571
6	1,943	2,447
7	1,895	2,365
8	1,860	2,306
9	1,833	2,262
10	1,812	2,228
11	1,796	2,201
12	1,782	2,179
13	1,771	2,160
14	1,761	2,145
15	1,753	2,131
16	1,746	2,120
17	1,740	2,110
18	1,734	2,101
19	1,729	2,093
20	1,725	2,086

**Operacijske karakteristike**

Pokazati se da,<sup>3</sup> da je za ta primer treba upoštevati drug tip operacijskih karakteristik, kot so opisane v članku<sup>2</sup>, saj imamo opravka s Studentovo distribucijo. Na sliki 2 je narisana familija operacijskih karakteristik za dvostransko preverjanje hipotez, če gre za Studentovo distribucijo, pri čemer smo si izbrali  $\alpha = 0,05$ .

$$d = \frac{|\beta_1 - \beta'_1| \cdot S_x}{\sigma_y}$$



Slika 2 Operacijske karakteristike za dvostranski t - test

Če privzamemo, da je standardna deviacija  $\sigma_y = S_y = 1,578$  in standardni odklon spremenljivke  $X$   $S_x = 3,317$ , lahko iz diagrama na sliki 2 odčitamo, da pripada 20 % verjetnosti za sprejeto hipotezo pri vzorcu (11 vrednosti) vrednost parametra  $d = 1$ . V diagramu je treba za ta primer upoštevati krivuljo s številko  $n-1$ , če je  $n$  število podatkov v vzorcu<sup>3</sup>.

Če je  $\beta'_1 = 0,4$ , bi to odgovarjalo  $\beta_1 = 0,876$ .

Pri tako majhnem vzorcu lahko torej naredimo občutno napako druge vrste. Trdimo lahko, da je  $\beta_1 = 0,4$ , pri tem pa je še vedno 20 % verjetno, da smo se zmotili, ker je v resnici morda  $\beta_1 = 0,876$ .

Vprašamo se pa lahko tudi drugače. Kolikšna je verjetnost, da smo se zmotili, ko smo trdili, da je  $\beta_1 = 0,4$ , ker je le-ta morda 0,2 oziroma 0,6<sup>2</sup>.

Iz družine operacijskih karakteristik (slika 2) znaša za ta primer vrednost  $d = 0,420$  in verjetnost za takšno napako je potem približno 80 %.

Iz istega diagrama se vidi, da bi za preskus trditve, da je  $\beta_1 = \beta'_1 = 0,4$  potrebovali približno 40 podatkov, če bi hoteli zagotoviti, da bi pri  $\beta_1 = 0,2$  oziroma 0,6 znašala verjetnost za napako druge vrste le 20 %.

Pri tako majhnem vzorcu torej nikakor ni izključeno, da bi bil  $\beta_1$  za našo populacijo enak 0,4, vendar pa je tudi verjetnost za vrednosti  $\beta_1$  0,2 ali 0,6 zelo velika.

### Ničelna hipoteza

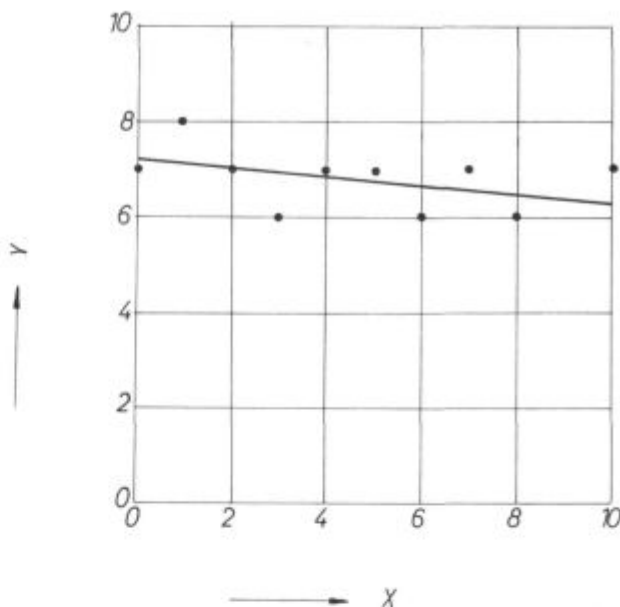
Oglejmo si še drug primer. Na sliki 3 so narisane točke ( $n = 11$ ), ki ponazarjajo odvisnost  $Y$  od  $X$  (tabela 3) pri nekem drugem poskusu.

Tabela III.

X	Y
0	7
1	8
2	7
3	6
4	7
5	7
6	6
7	7
8	6
9	6
10	7

Narisana je regresijska premica, ki jo karakterizirajo sledeči parametri:

$$R = 0,466$$



Slika 3

Primer za ničelno hipotezo

$$b_1 = -0,0909$$

$$S_{b_1} = 0,0575$$

$$S_x = 0,332$$

$$S_y = 0,647$$

Pojasnjenih je približno 22 % vseh odstopanj spremenljivke  $Y$ .

Ali smemo trditi na podlagi izračunane regresijske premice, da  $Y$  pada, ko  $X$  narašča od 0 proti 10? Ali gre sploh za kakšno odvisnost med  $X$  in  $Y$ ?

Ta primer testiramo tako, da postavimo ničelno hipotezo:

$$\beta'_1 = 0$$

Ali lahko trdimo, da je naš vzorec iz populacije, kjer je  $\beta_1 = 0$ , oziroma da ni odvisnosti med  $X$  in  $Y$ ? Izračunajmo vrednost parametra  $t$ !

$$t = \left| \frac{b_1 - \beta'_1}{S_{b_1}} \right|$$

V našem primeru je  $t = 1,581$ . Če si ponovno izberemo  $\alpha = 0,05$ , znaša  $t_{0,975; 9} = 2,262$ , kar je več kot 1,581. Hipoteze  $\beta_1 = \beta'_1 = 0$  ne moremo zavreči. Trdili bi torej, da med  $Y$  in  $X$  ni povezave.

Ali smo pa naredili napako druge vrste? Ali ni morda resnični  $\beta_1$  sicer majhen in negativen in je le število meritev premajhno? Spet si pomagajmo z operacijskimi karakteristikami na sliki 2.

Kolikšna je verjetnost, da je resnični  $\beta_1 = -0,1$ ?

Za ta primer znaša vrednost  $d$ :

$$d = \frac{|\beta_1 - \beta'_1| S_x}{\sigma_y} = 0,512$$

To pomeni, da je približno 70 % verjetno, da smo zgrešili premico z regresijskim koeficientom  $-0,1$ . 20 % verjetnost za napako druge vrste bi bila šele pri  $d = 1$ , oziroma  $\beta_1 = -0,195$ .

Potrebovali bi vsaj 32 podatkov, da bi lahko rekli, da je  $\beta_1 = \beta'_1 = 0$  in bi bilo pri tem 20 % verjetno, da bi zgrešili pri takem testiranju resbila šele pri  $d = 1$ , oziroma  $\beta_1 = -0,195$ .

**Populacija**

Kaj torej lahko povemo o parametru  $\beta_1$  za primer na sliki 1. Točne vrednosti ne moremo povedati. Lahko pa določimo interval, v katerem z 90 % gotovostjo pričakujemo vrednost parametra  $\beta_1$  po formuli<sup>3, 4, 5, 6</sup>:

$$\beta_1 = b_1 \pm t_{\frac{1+\gamma}{2}, n-2} \cdot S_{b_1}$$

Pri tem je  $\gamma$  verjetnost (0,90),  $n$  pa število podatkov v vzorcu (11) in  $t_{0,95; 9} = 1,833$  po tabeli II:  
 $0,193 < \beta_1 < 0,553$

Zaradi velikega procenta nepojasnjenih vplivov (slučajnostnih) tudi za določeno vrednost  $X'$  ne moremo povedati, kolikšna bi bila pripadajoča vrednost  $Y'$ . Določeni vrednosti  $X'$  pripada veliko število možnih vrednosti  $Y'$ . Vse te vrednosti se porazdeljujejo okrog neke srednje vrednosti  $\bar{Y}'$ . Predpostavljamo, da je distribucija vrednosti  $Y'$ , ki pripada nekemu določenemu  $X'$ , normalna in da je standardna deviacija teh  $Y'$  enaka pri vseh  $X'$ .

Interval vrednosti, v katerem lahko z neko določeno verjetnostjo ( $\gamma$ ) pričakujemo naš  $\bar{Y}'$ , se lahko izračuna po formuli:

$$\bar{Y}' = \hat{Y}(X') \pm t_{\frac{1+\gamma}{2}, n-2} \cdot S_{\bar{Y}'}$$

pri čemer je

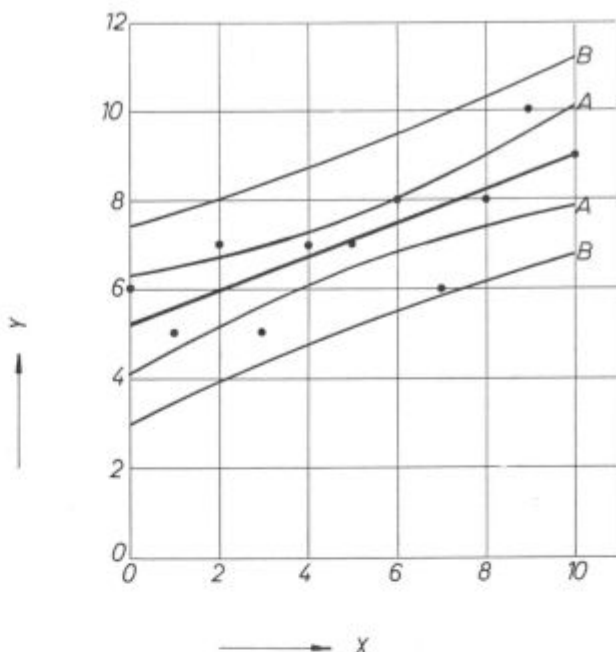
$$S_{\bar{Y}'} = S_E \cdot \sqrt{\frac{1}{n} + \frac{(X' - \bar{X})^2}{\sum_i (X_i - \bar{X})^2}}$$

Na sliki 4 sta narisani obe meji (A) za 90 % verjetnost za podatke iz tabele 1. Če bi torej naredili zelo veliko meritev, bi lahko na primer pri  $X' = 7$  dobili množico pripadajočih vrednosti  $Y'$ . Srednja vrednost teh  $Y'$  bi bila nekje med 7,1 in 8,5. To pričakujemo z 90 % gotovostjo.

Vprašajmo se še, kaj lahko pričakujemo za posamezno vrednost  $Y'$ , ki pripada nekemu  $X'$ . Z verjetnostjo  $\gamma$  lahko pričakujemo, da bo  $Y'$  ležal v intervalu

$$Y' = \hat{Y}(X') + t_{\frac{1+\gamma}{2}, n-2} \cdot S_{Y'}$$

pri čemer je  $S_{Y'} = S_E \cdot \sqrt{1 + \frac{1}{n} + \frac{(X' - \bar{X})^2}{\sum_i (X_i - \bar{X})^2}}$



Slika 4  
 Regresijska premica in meje 90 % gotovosti

Na sliki 4 sta narisani obe mejni vrednosti (B) tudi za 90 % gotovost za vrednost  $Y'$ . Pri  $X' = 7$  lahko torej z 90 % gotovostjo pričakujemo, da bomo pripadajočo vrednost  $Y'$  našli nekje med 5,8 in 9,8.

Zato moramo biti zelo previdni pri opisovanju poteka odvisnosti  $Y(X)$ . Neutemeljena bi bila trditev, da na primer  $Y$  močno narašča, ko gre  $X$  od 7 proti 9, nato pa hitro pada.

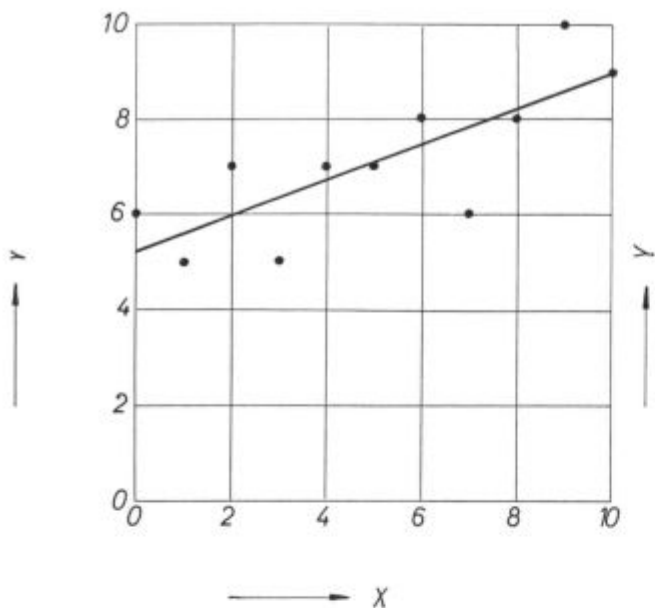
**Parabolična odvisnost**

Zastavimo si še eno vprašanje. Predpostavimo, da so točke na sliki 1 rezultat nekih meritev vrednosti  $Y$  pri posameznih vrednostih  $X$ . Očitno je, da so vrednosti  $Y$  do neke mere odvisne od vrednosti  $X$ . Toda, ali je odvisnost zares linearna? Če nam tega ni nihče prej povedal, bi bilo morda pravilneje narisati kakšno parabolo, ki bi lepše ponazarjala potek teh točk na sliki 1.

Po metodi najmanjših kvadratov<sup>1</sup> smo poiskali regresijske koeficiente v naslednjih enačbah:

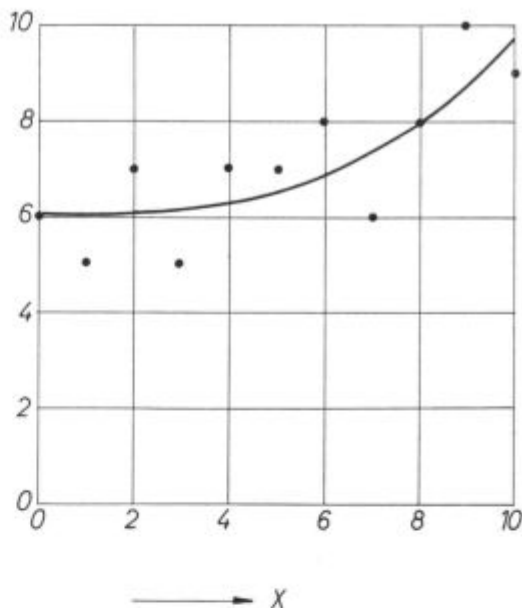
- Linearna  $\hat{Y}_1 = b_{01} + b_{11}X$  (61 %)
- Kvadratična  $\hat{Y}_2 = b_{02} + b_{12}X^2$  (65 %)
- Kubična  $\hat{Y}_3 = b_{03} + b_{13}X^3$  (63 %)
- Parabola četrte stopnje  $\hat{Y}_4 = b_{04} + b_{14}X^4$  (59 %)

V oklepaju je naveden procent raztrosa vrednosti  $Y$ , ki ga pojasnjuje posamezna enačba ( $R^2$ ). Parabola druge stopnje  $\hat{Y}_2$  pojasnjuje največji del raztrosa, vendar gre le za 4 % izboljšanja glede na linearno regresijsko funkcijo  $\hat{Y}_1$ .



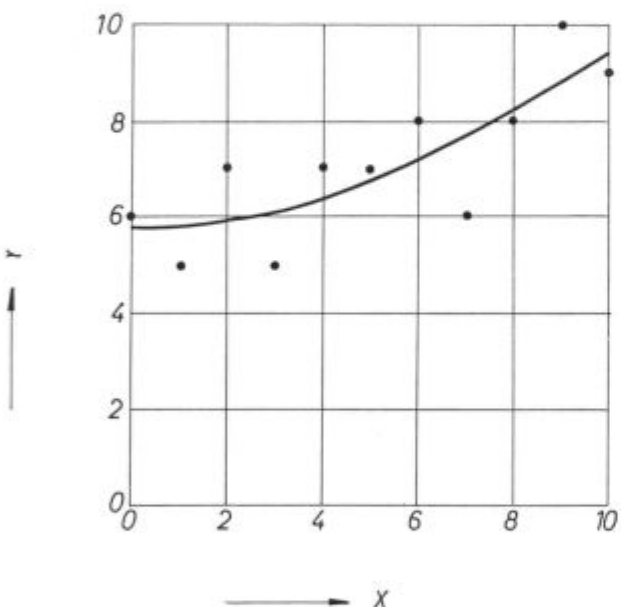
Slika 5

Regresijska enačba:  $\hat{Y}_1 = b_{01} + b_{11} X$



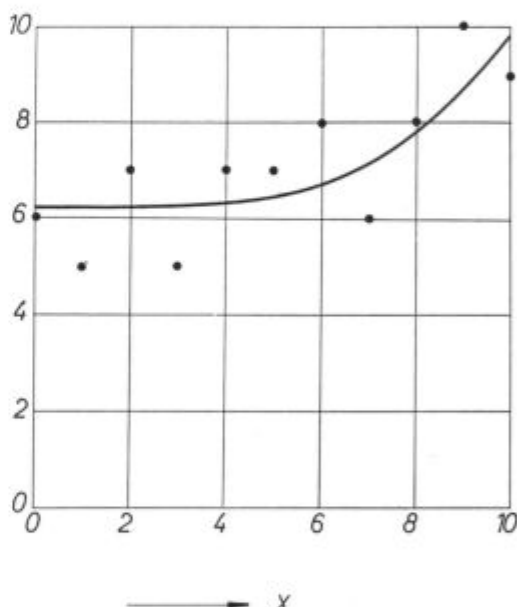
Slika 7

Regresijska enačba:  $\hat{Y}_3 = b_{03} + b_{13} X^2$



Slika 6

Regresijska enačba:  $\hat{Y}_2 = b_{02} + b_{12} X^2$



Slika 8

Regresijska enačba:  $\hat{Y}_4 = b_{04} + b_{14} X^4$

Naredimo enako analizo glede vrednosti  $\bar{Y}'$  in  $Y'$ , ki jih lahko pričakujemo z 90 % gotovostjo v primeru, ko gre za kvadratično parabolo. Narišemo lahko podoben diagram (slika 9), kot ga prikazuje slika 4. V tem primeru lahko z 90 % gotovostjo pri  $X' = 7$  pričakujemo  $Y'$  med 7,0 in 8,2.

Posamezno vrednost  $Y'$  pa lahko z 90 % verjetnostjo najdemo v intervalu med 5,7 in 9,5.

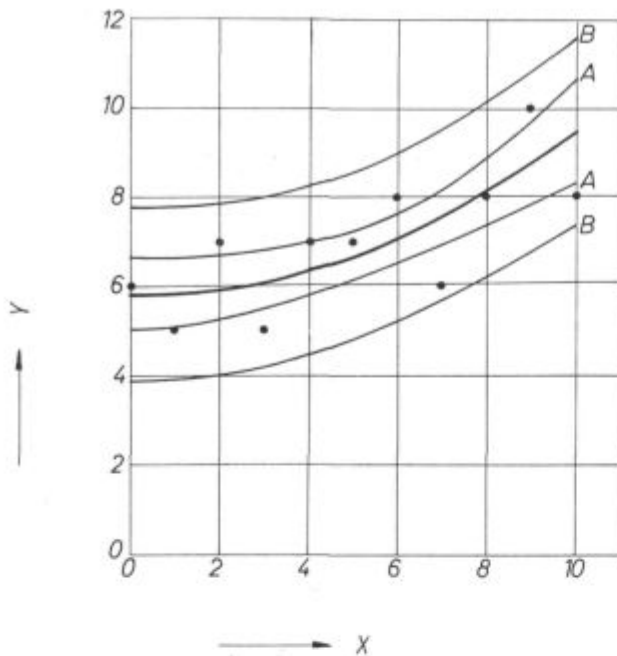
S tega stališča je kaj malo važno, kaj v resnici narišemo skozi točke na sliki 1 — premico ali parabolo. Obe sta dokaj neostro definirani. Če že potegnemo parabolo, moramo pa biti zelo previdni pri zaključkih.

Nujno je treba upoštevati statistično gotovost, oziroma negotovost pri vrednostih funkcije  $Y$ , ki jih želimo napovedati. Zaradi velikega raztrosa točk okrog parabole težko trdimo za populacijo, da se na primer krivulja odvisnosti med  $Y$  in  $X$  pri večjih  $X$  bolj strmo dviga kot pri majhnih vrednostih  $X$ .

Metoda ortogonalnih polinomov daje točnejši odgovor na vprašanje, ali sploh gre za parabolo ali ne.

**Metoda ortogonalnih polinomov**

V primeru, ko so posamezne vrednosti  $X$  enakomerno (v enakih intervalih) razporejene in, ko



Slika 9  
Regresijska parabola Y: in meje 90 % gotovosti

vsaki vrednosti X pripada le en Y, je mogoče narediti polinomno regresijo s pomočjo ortogonalnih polinomov.<sup>6, 7, 8</sup>

Metoda je zelo podobna postopni regresiji<sup>1</sup>, oziroma polinomni regresiji<sup>9</sup>.

Največja prednost te metode pred običajno polinomno regresijo je v tem, da lahko postopoma dodajamo člene višjega reda, ne da bi pri tem morali spreminjati tudi koeficiente pri ostalih stopnjah polinoma.

Ce smo na primer poskušali najprej opisati zvezo Y (X) z linearno enačbo  $\hat{Y} = b_0 + b_1 X$  po programu za polinomno regresijo<sup>9</sup>, nato pa smo jo izboljšali z nastankom  $\hat{Y} = b'_0 + b'_1 X + b'_2 X^2$ , smo morali vso obdelavo narediti še enkrat. Vrednosti  $b_0$  in  $b_1$  nam ne koristijo več, saj je  $b_0 \neq b'_0$  in  $b_1 \neq b'_1$ .

Slabost programa<sup>9</sup> je tudi v tem, da ne moremo vedeti, kolikšen je prispevek dodanega kvadratnega člana.

Metoda ortogonalnih polinomov nam pa da drugačno sliko. Namesto običajnega izraza

$$\hat{Y} = b_0 + b_1 X + \dots + b_k X^k$$

lahko uporabimo izraz:

$$\hat{Y} = A_0 + A_1 \xi_1 + \dots + A_k \xi_k,$$

kjer so  $\xi_i (i = 1 \rightarrow k)$  ortogonalni polinomi i-te stopnje spremenljivke X,  $A_i (i = 0 \rightarrow k)$  pa konstante.

Za ortogonalne polinome  $\xi_i$  velja:

$$\sum_n \xi_i \xi_k = 0 \text{ če je } i \neq k$$

$$\sum_n \xi_i = 0$$

Pri tem seštevamo po vseh vrednostih spremenljivke X.

Zaradi ortogonalnih lastnosti lahko izračunamo koeficiente  $A_i$ :

$$A_0 = \sum_n \frac{\hat{Y}}{n} = \bar{Y}$$

$$A_i = \frac{\sum_n Y \cdot \xi_i}{\sum_n \xi_i^2} \quad i = 1 \dots k$$

V primeru, ko vrednosti X naraščajo od 1 do  $n^{11}$  v enakih intervalih (po 1), lahko izračunamo prve tri ortogonalne polinome po formulah:

$$\xi_1 = \lambda_1 (X - \bar{X})$$

$$\xi_2 = \lambda_2 \left[ (X - \bar{X})^2 - \frac{n^2 - 1}{12} \right]$$

$$\xi_3 = \lambda_3 \left[ (X - \bar{X})^3 - (X - \bar{X}) \cdot \frac{3n^2 - 7}{20} \right]$$

$\lambda$  si izberemo tako, da je  $\xi$  čim manjše celo število.

Analiza variance daje sliko o prispevku posameznih ortogonalnih polinomov, ki nastopajo v regresijski enačbi. Vsota kvadratov, ki pripada i-temu polinomu je enaka:

$$A_i \sum_n (Y \xi_i)$$

### Regresijske enačbe s pomočjo ortogonalnih polinomov

Po tej metodi smo obdelali primer iz tabele I in se prepričali, da gre pri tem le za linearno odvisnost (Tabela IV.)

Opraviti imamo z 11 vrednostmi X, ki naraščajo po 1 od 0 do 10. Omejili smo se le na prve tri ortogonalne polinome.

Iz tabel<sup>10</sup> se da videti, da znaša vrednost parametra F za 95 % verjetnost za ta primer  $F_{1; 7; 0,05} = 5,5914$ , za 90% verjetnost pa 12,246.

Očitno je torej, da je edina smiselna regresijska enačba — enačba premice. Kvadratični prispevek ni statistično pomemben.

Regresijsko enačbo zapišemo z ortogonalnimi polinomi takole:

$$\hat{Y} = 7,0909 + 0,3727 \cdot (X - 5),$$

oziroma:

$$\hat{Y} = 5,2273 + 0,3727 X$$

### ZAKLJUČEK

Opisali smo torej nekaj osnovnih primerjav, ki jih mora upoštevati vsak, ki želi interpretirati svoje diagrame. Nekritično tolmačenje takšnih diagramov lahko zavaja raziskovalce pri njihovem delu, da pogosto ugotavljajo dejstva, za katera nimajo nobene osnove.



Tabela IV.

X	Y	$\xi_1$	$\xi_2$	$\xi_3$	$Y \cdot \xi_1$	$Y \cdot \xi_2$	$Y \cdot \xi_3$	
0	6	-5	+15	-30	-30	+90	-180	$A_0 = 7,0909$
1	5	-4	+6	+6	-20	+30	+30	
2	7	-3	-1	+22	-21	-7	+154	
3	5	-2	-6	+23	-10	-30	+115	$A_1 = 0,3727$
4	7	-1	-9	+14	-7	-63	+98	
5	7	0	-10	0	0	-70	0	$A_2 = 0,0338$
6	8	+1	-9	-14	8	-72	-112	
7	6	+2	-6	-23	12	-36	-138	$A_3 = 0,0002$
8	8	+3	-1	-22	24	-8	-176	
9	10	+4	+6	-6	40	+60	-60	
10	9	+5	+15	+30	45	+135	+270	
$\Sigma$	78	0	0	0	+41	+29	+1	
$\lambda$		1	1	$\frac{5}{6}$				
$\Sigma \xi^2$		110	858	4290				

## Analiza variance za primer iz tabele I.

Prispevek polinomov	Vsota kvadratov	Stopnje prostosti	Povprečni kvadrat	F-vrednost
Linearni	15,2818	1	15,2818	12,37
Kvadratični	0,9802	1	0,9802	0,79
Kubični	0,0002	1	0,0002	0,00
Ostane	8,6469	7	1,2353	
Vsota	24,9091	10		

Metoda ortogonalnih polinomov pa kaže, kako učinkovito je mogoče tudi brez modernega računalnika določiti osnovne karakteristike povezave med X in Y.

## Literatura

1. B. Rode: Statistična analiza regresije z uporabo elektronskih računalnikov. Zelezarski zbornik št. 3, leto 1969, stran 221.

2. B. Brudar: Preverjanje statističnih hipotez s pomočjo operacijskih karakteristik. Zelezarski zbornik št. 3, 1972, stran 175.

3. A. H. Bowker, G. J. Lieberman: Engineering Statistics, Prentice Hall, Inc. 1959.

4. M. Fisz: Wahrscheinlichkeitsrechnung und mathematische Statistik VEB Deutsche Verlag der Wissenschaften, Berlin 1970.

5. N. W. Smirnov, I. W. Dunin — Barkovski: Mathematische Statistik in der Technik, VEB Deutsche Verlag, der W., Berlin 1970.

6. B. Ostle: Statistics in Research, The Iowa State University Press 1969.

7. A. Linder: Planen und Auswerten von Versuchen, Birkhäuser Verlag 1969.

8. O. L. Davies: Design and Analysis of Industrial Experiments, Hafner Publishing Company, New York 1971.

9. System/360 Scientific Subroutine Package (360A-CM-03X) Version III, Programmer's Manual, IBM Application Program

10. B. Rode, J. Rodič: Statistično planiranje in vrednotenje metalurških raziskav, Zel. zbornik št. 2, 1968, stran 99.

## ZUSAMMENFASSUNG

Die Messergebnisse werden üblich in einem Diagramm als Punkte aufgezeichnet. Diese Punkte werden dann mit einer entsprechen Kurve verbunden.

Was für eine Kurve durch diese Punkte aufgezeichnet werden kann und welche Schlussfolgerungen über die Zusammenhänge zwischen der abhängigen und unabhängigen Variable gezogen werden können wenn wir die Prinzipien der mathematischen Statistik befolgen, ist in diesem Artikel

enthalten. Es sind zwei Beispiele beschrieben, bei denen mit Hilfe der Operationskarakteristiken die Zuverlässigkeitsgrenzen der Regresionsgleichungsparametern bestimmt werden.

Mit der Methode der orthogonalen Polynomen ist die Wichtigkeit der einzelnen Glieder im Polynom, welche die Abhängigkeit Y von X darstellt bestimmt.

## SUMMARY

Experimental results are usually plotted in a diagram. Singular points are connected by a most suitable curve.

The paper describes which curve may be drawn and which conclusions on the relationship between the dependent and the independent variable can be made if the principles of mathematical statistics are obeyed. Two

examples are described where confidence intervals for the parameters of the regression equation are determined by means of operating characteristic curves.

Method of orthogonal polynomials determines the importance of single terms in polynomials which represents the relationship between Y and X.

## ЗАКЛЮЧЕНИЕ

Результаты измерений как следует рисуем в форме точек в диаграмму. Точки соединяем в наиболее подходящую кривую.

Дано описание какую кривую можно нарисовать, и какие заключения можно вывести о соотношении между зависимыми и независимыми изменяемыми величинами, придерживаясь принципов математической статистики.

Описаны два примера при которых, при помощи оперативных характеристик, определены границы надёжности параметров уравнения регрессии. Методом ортогональных многочленов определено значение отдельных членов, что представляет собой зависимость Y от X.