OCENJEVANJE PARAMETROV V BAYESOVI STATISTIKI

ALEŠ TOMAN

Ekonomska fakulteta
Univerza v Ljubljani

Math. Subj. Class. (2010): 62F15, 62F10

Čeprav začetki Bayesove statistike segajo v drugo polovico 18. stoletja, je svoj razcvet doživela šele z razvojem računalnikov ob koncu 20. stoletja. V članku bomo na enostavnem zgledu prikazali ključne korake in lastnosti Bayesovega pristopa k ocenjevanju parametrov.

BAYESIAN PARAMETER ESTIMATION

Although the beginnings of Bayesian statistics date back to the second half of the 18$^{\text{th}}$ century, it started to flourish at the end of the 20$^{\text{th}}$ century when computers became widely available. In this paper we go through some basic steps and properties of Bayesian parameter estimation.

Uvod

Statistika je veda, ki razvija in proučuje metode zbiranja podatkov ter njihove analize in predstavitve. Statistiki pri svojem delu lahko le redkokdaj razpolagamo s podatki za celotno populacijo ali poznamo natančne lastnosti opazovanega pojava. Pogosteje imamo na voljo le informacije za nekaj na slepo izbranih enot, ki sestavljajo slučajni vzorec. Osrednja naloga sklepne statistike je opisovanje lastnosti populacije ali pojava na osnovi lastnosti, ki jih opazimo oziroma izmerimo na vzorcu [7].

Obstaja več pristopov k statističnemu sklepanju. Danes je bolj znan **frekventistični pristop**, ki so ga v prvi polovici 20. stoletja utemeljili R. A. Fisher (1890–1962), E. S. Pearson (1895–1980) in J. Neyman (1894–1981). Za ta pristop so značilne cenilke največjega verjetja, intervali zaupanja ter preizkušanje domnev [7].

Manj znan, a vse pomembnejši je **Bayesov pristop**. Za njegov začetek štejemo zapiske T. Bayesa (1702–1761), ki jih je leta 1763 objavil R. Price. V njih je nakazana znamenita Bayesova formula in razprava o tem, kako se ob opazovanju pojavov spreminjajo naša prepričanja [6]. Ob koncu 18. stoletja je P.-S. de Laplace (1749–1827) podrobneje (neodvisno od Bayesa) predstavil, kako Bayesovo formulo uporabimo v različnih statističnih problemih. Med drugim je iz podatkov o rojstvih v pariških porodnišnicah ocenil verjetnost za rojstvo deklice [4]. V nadaljevanju si bomo ogledali podoben primer.

Bayesov pristop je prevladoval v statistiki 19. stoletja. Čeprav je v teoriji omogočal enostavno analizo zapletenih modelov, je bil v praksi zaradi računskih zahtev manj uporaben. Razvoj učinkovitih simulacijskih algoritmov ter široka dostopnost računalnikov ob koncu 20. stoletja pa sta Bayesovemu pristopu vrnila mesto v statistični znanosti.

Bayesova formula

Bayesovo formulo pogosto povezujemo z **dvofaznimi poskusi**, kjer v prvi fazi nastopi natanko eden od dogodkov iz popolnega sistema dogodkov (hipotez) $H_1, \ldots, H_n$ in so od tega, kateri se je pripetil, odvisni pogoji poskusa v drugi fazi, v katerem opazujemo dogodek A [3].

Privzemimo, da poznamo verjetnosti $P(H_1), \ldots, P(H_n)$ vseh hipotez ter pogojne verjetnosti $P(A|H_1), \ldots, P(A|H_n)$ dogodka A glede na posamezne hipoteze. **Formula za popolno verjetnost** nam pove, kako izračunati brezpogojno verjetnost dogodka A v drugi fazi:

$$P(A) = \sum_{i=1}^{n} P(H_i) \cdot P(A|H_i).$$

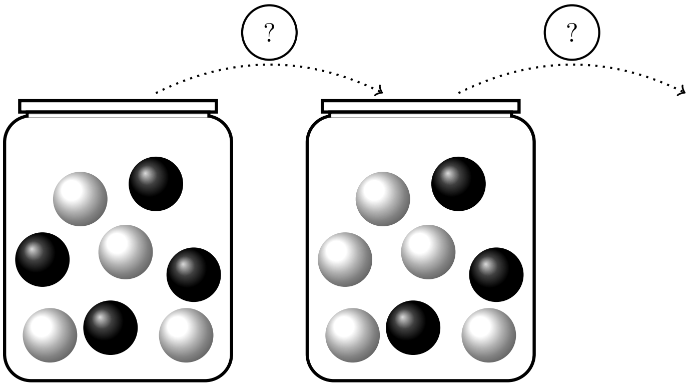
Postavimo si še obratno vprašanje: Če vemo, da se je dogodek A v drugi fazi zgodil, kolikšna je pogojna verjetnost, da se je v prvi fazi zgodila hipoteza H_i? Odgovor nam da **Bayesova formula**

$$P(H_i|A) = \frac{P(H_i) \cdot P(A|H_i)}{P(A)}.$$

Verjetnosti $P(H_i)$ pravimo apriorna, verjetnosti $P(H_i|A)$ pa aposteriorna verjetnost hipoteze H_i. Opazimo, da imenovalec v Bayesovi formuli ni odvisen od i in je pri vseh hipotezah enak. Potrebujemo ga za to, da se aposteriorne verjetnosti vseh n hipotez seštejejo v 1. Zato v Bayesovi statistiki pogosto zapišemo samo sorazmerje

$$P(H_i|A) \propto P(H_i) \cdot P(A|H_i).$$

Zgled 1. V prvem kozarcu so dobro premešane 4 črne in 4 bele kroglice, v drugem pa 3 črne in 5 belih. Iz prvega kozarca na slepo izvlečemo eno kroglico in jo preložimo v drugi kozarec. Tega pretresemo in nato iz njega na slepo izvlečemo eno kroglico. Kolikšne so verjetnosti dogodkov A, da iz prvega kozarca izvlečemo belo kroglico, B, da iz drugega kozarca izvlečemo belo kroglico, in C, da smo iz prvega kozarca izvlekli belo kroglico, če vemo, da smo iz drugega kozarca izvlekli belo kroglico?



Slika 1. Shema dvofaznega poskusa s kroglicami.

Naloga opisuje dvofazni poskus, njegova shema je prikazana na sliki 1. V prvi fazi iz prvega kozarca izvlečemo kroglico in jo preložimo v drugi kozarec. Pri tem sta možni dve hipotezi: H_1, da izvlečemo črno kroglico, in H_2, da izvlečemo belo kroglico. Ker so v prvem kozarcu 4 črne in 4 bele kroglice, sta apriorni verjetnosti hipotez

$$P(H_1) = P(H_2) = \frac{4}{8} = \frac{1}{2}.$$

V drugi fazi poskusa izvlečemo kroglico iz drugega kozarca. Možna sta dva dogodka (je črna ali bela); v nalogi nas zanima dogodek B, da izvlečemo belo kroglico. Po formuli za popolno verjetnost izračunamo

$$P(B) = P(H_1) \cdot P(B|H_1) + P(H_2) \cdot P(B|H_2) = \frac{1}{2} \cdot \frac{5}{9} + \frac{1}{2} \cdot \frac{6}{9} = \frac{11}{18}.$$

Pri računanju $P(B|H_1)$ smo si pomagali z vsebinsko interpretacijo dogodkov. Če v prvi fazi izvlečemo črno kroglico (pogoj H_1), imamo nato v drugem kozarcu 4 črne in 5 belih kroglic. Verjetnost, da iz njega izvlečemo belo kroglico (dogodek B), je zato $P(B|H_1) = \frac{5}{9}$. Podobno izračunamo še $P(B|H_2) = \frac{6}{9}$.

Z Bayesovo formulo izračunamo še aposteriorni verjetnosti hipotez

$$P(H_1|B) = \frac{P(H_1) \cdot P(B|H_1)}{P(B)} = \frac{(1/2) \cdot (5/9)}{11/18} = \frac{5}{11} < \frac{1}{2},$$

$$P(H_2|B) = \frac{P(H_2) \cdot P(B|H_2)}{P(B)} = \frac{(1/2) \cdot (6/9)}{11/18} = \frac{6}{11} > \frac{1}{2}.$$

Ob tem opazimo:

- Aposteriorni verjetnosti hipotez se razlikujeta od apriornih. Razlog za razliko je informacija o dogodku, ki se je zgodil v drugi fazi poskusa.
- Razmerje aposteriornih verjetnosti hipotez $(5 : 6)$ je enako razmerju pogojnih verjetnosti dogodka B v drugi fazi. To je posledica enakih apriornih verjetnosti hipotez.

Za končno rešitev zapišimo še verjetnosti dogodkov A in C. Ker je $A = H_2$, je $P(A) = \frac{1}{2}$, in ker je $C = H_2|B$, velja $P(C) = \frac{6}{11}$.

Evrski kovanec in neznani parameter p

Opišimo enostaven statistični problem, na katerem bomo prikazali vse korake Bayesovega pristopa k ocenjevanju neznanih parametrov ter ga na kratko primerjali s frekventističnim pristopom. Zamislimo si stavo z ne nujno poštenim evrskim kovancem. Kovanec bomo vrgli enkrat, dobitek pa prejmemo, če smo pred tem napovedali pravi izid. Na kaj bomo stavili? Splača se staviti na izid, ki je bolj verjeten.

Pri metu kovanca sta možna dva dogodka: dogodek C, da pade cifra, ter dogodek G, da pade grb. Ker drugih možnosti ni, je

$$P(C) + P(G) = 1.$$

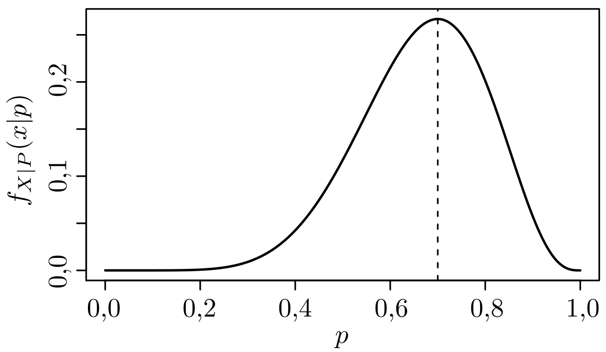
Ker ne vemo, ali je kovanec pošten, označimo $P(C) = p$ ter $P(G) = 1-p$. Število p je neznani populacijski parameter (neznana verjetnost) oziroma lastnost kovanca. Od njega bo odvisna naša stava. Naloga statistike je zanj podati čim boljšo oceno.

Privzemimo, da smemo pred stavo z nekaj meti preizkusiti kovanec. Kovanec smo vrgli $n = 10$-krat in pri tem je cifra padla $x = 7$-krat. Zaporedje 10 črk, med katerimi se 7-krat pojavi C in 3-krat G, je slučajni vzorec, s katerim bomo ocenili neznani parameter p. Pri tem bomo uporabili frekventistični ter Bayesov pristop.

V obeh pristopih je pomembna **funkcija verjetja** $f_{X|P}(x|p)$. Ta podaja verjetnost na danem vzorcu izmerjenih vrednosti (x cifer v n metih) v odvisnosti od neznanega parametra p. V našem primeru je verjetnost, da v 10 metih kovanca opazimo 7 cifer, enaka

$$f_{X|P}(x|p) = \binom{10}{7} p^7 (1-p)^3.$$

To je binomsko verjetje. Graf funkcije verjetja je na sliki 2. V frekventistični statistiki lahko oceno $\hat{p}$ parametra p določimo z metodo največjega verjetja. Ta za $\hat{p}$ izbere tisto vrednost p, pri kateri funkcija verjetja $f_{X|P}(x|p)$ doseže najvišjo vrednost. Zlahka preverimo, da je to pri $\hat{p} = \frac{x}{n} = 0{,}7$. Vrednost prikazuje črtkana navpičnica na grafu na sliki 2.



Slika 2. Graf funkcije verjetja.

Dobljena **točkovna ocena** parametra p je korektna, a je navedba zgolj točkovne ocene lahko zavajajoča. Če bi pri samo enem izmed metov kovanca opazili drugačen izid, bi se naša ocena parametra p spremenila za 0,1. Več informacij (natančnost naše ocene) o neznanem parametru p predstavimo z **intervalom zaupanja**. Konstrukcij intervalov zaupanja za neznano verjetnost je več. Tu bomo uporabili le dve, ki ju zaradi njunih lastnosti najpogosteje priporočajo za uporabo v praksi [1, 2]. Wilsonov[1] 95 % interval zaupanja za p je [0,397; 0,892], Clopper-Pearsonov[2] 95 % interval zaupanja pa [0,348; 0,933].

Bayesov pristop

Bayesova statistika neznane parametre obravnava kot (zvezne) slučajne spremenljivke, naše védenje o njihovih vrednostih pa opiše s porazdelitvami, najlažje z gostoto verjetnosti te slučajne spremenljivke. Bayesov pristop k ocenjevanju parametrov idejno sledi Bayesovi formuli. Najprej privzamemo neko **apriorno gostoto verjetnosti** $f_P(p)$ za neznani parameter p in izračunamo verjetje $f_{X|P}(x|p)$ na danem vzorcu izmerjenih vrednosti v odvisnosti od parametra p. Pri tem lahko apriorno gostoto verjetnosti izberemo na osnovi preteklih analiz ali splošnih dognanj. Nato določimo **aposteriorno**

[1] Wilsonov $100(1-\alpha)\,\%$ interval zaupanja omejujeta vrednosti $\frac{x+z^2/2}{n+z^2} \pm \frac{z\sqrt{n}}{n+z^2}\sqrt{\hat{p}(1-\hat{p})+\frac{z^2}{4n}}$, kjer z označuje $(1-\alpha/2)$-kvantil standardne normalne porazdelitve.

[2] Spodnja meja Clopper-Pearsonovega $100(1-\alpha)\,\%$ intervala zaupanja je $\alpha/2$-kvantil porazdelitve $\mathrm{Beta}(x, n-x+1)$, zgornja meja pa $(1-\alpha/2)$-kvantil porazdelitve $\mathrm{Beta}(x+1, n-x)$.

gostoto verjetnosti $f_{P|X}(p|x)$ parametra p po Bayesovi formuli

$$f_{P|X}(p|x) = \frac{f_P(p) \cdot f_{X|P}(x|p)}{f_X(x)}.$$

Imenovalec $f_X(x)$ predstavlja robno (brezpogojno) porazdelitev vzorčnih vrednosti in ni odvisen od p, zagotovi pa, da je integral aposteriorne gostote verjetnosti enak 1. Poenostavljeno lahko zapišemo samo sorazmerje

$$f_{P|X}(p|x) \propto f_P(p) \cdot f_{X|P}(x|p)$$

in nato določimo ustrezno normirno konstanto. Aposteriorna gostota verjetnosti predstavlja naše védenje o neznanem parametru p, potem ko smo združili naše apriorno znanje in informacije iz podatkov.

Vrnimo se k našemu kovancu. Funkcijo verjetja smo že spoznali. Ker bomo normirno konstanto v aposteriorni gostoti verjetnosti določili na koncu, zapišemo samo funkcijski del funkcije verjetja

$$f_{X|P}(x|p) \propto p^7(1-p)^3.$$

Izbrati moramo še apriorno porazdelitev parametra p. Jasno je, da p leži med 0 in 1, zato je smiselna izbira gostote verjetnosti iz družine porazdelitev beta [4]. Ta ima dva parametra, označimo ju z α in β. Oglejmo si njene lastnosti. Naj bo $p \sim \mathrm{Beta}(\alpha, \beta)$. Gostota verjetnosti spremenljivke p je tedaj

$$f_P(p) = \frac{1}{B(\alpha,\beta)} p^{\alpha-1}(1-p)^{\beta-1} \text{ za } p \in [0,1],$$

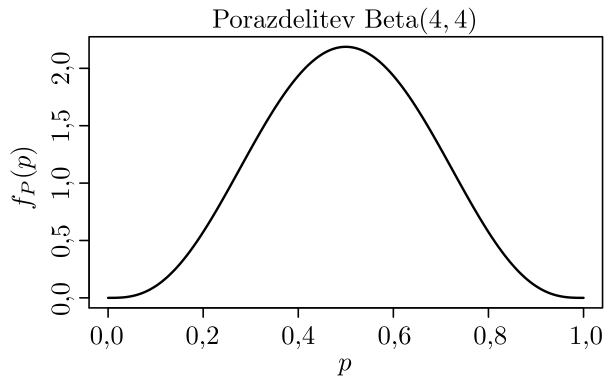
upanje $E(p) = \frac{\alpha}{\alpha+\beta}$ ter varianca $\mathrm{var}(p) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$. Tu smo z $B(\alpha, \beta)$ označili vrednost funkcije beta. Tudi pri apriorni porazdelitvi lahko ohranimo samo njen funkcijski del, zato zapišemo

$$f_P(p) \propto p^{\alpha-1}(1-p)^{\beta-1}.$$

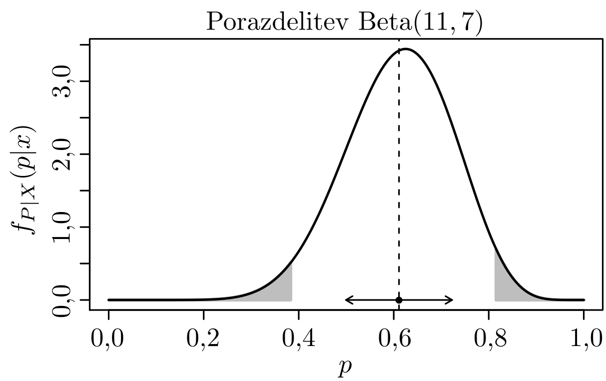
Določiti moramo še parametra α in β; parametrom apriorne porazdelitve pravimo **hiperparametri**. Privzemimo, da iz izkušenj vemo, da naj bi bil kovanec pošten, zato α in β določimo tako, da bo $E(p) = \frac{1}{2}$. Torej mora biti $\alpha = \beta$. Varianca je tedaj enaka $\mathrm{var}(p) = \frac{1}{4(2\alpha+1)}$, standardni odklon pa $\sigma(p) = \frac{1}{2\sqrt{2\alpha+1}}$. Z njim povemo, kako prepričani smo v poštenost kovanca. Zaradi enostavnosti izberimo $\alpha = 4$, kjer dobimo $\sigma(p) = \frac{1}{6}$. Graf apriorne gostote verjetnosti parametra p je na sliki 3.

Z uporabo Bayesove formule ugotovimo, da je aposteriorna gostota verjetnosti parametra p sorazmerna

$$f_{P|X}(p|x) \propto \underbrace{p^{4-1}(1-p)^{4-1}}_{\text{apriorna}} \cdot \underbrace{p^7(1-p)^3}_{\text{verjetje}} = p^{10}(1-p)^6.$$



Slika 3. Apriorna gostota verjetnosti parametra p.



Slika 4. Aposteriorna gostota verjetnosti parametra p.

V zadnjem izrazu prepoznamo obliko gostote porazdelite beta, zato velja

$$p|x \sim \mathrm{Beta}(11, 7) = \mathrm{Beta}(\alpha', \beta').$$

Graf njene gostote verjetnosti je na sliki 4. Za točkovno oceno parametra najpogosteje izberemo matematično upanje aposteriorne gostote verjetnosti; to znaša $\frac{\alpha'}{\alpha'+\beta'} = 0{,}611$ in ga prikazuje črtkana navpičnica na grafu na sliki 4. Natančnost točkovne ocene lahko podamo s standardnim odklonom aposteriorne gostote; ta je $\sqrt{\frac{\alpha'\beta'}{(\alpha'+\beta')^2(\alpha'+\beta'+1)}} = 0{,}012$ in ga prikazujeta vodoravni puščici na grafu; še pogosteje pa s **centralnim intervalom** aposteriorne gostote verjetnosti. 95 % centralni interval dobimo tako, da na levem in desnem repu porazdelitve odrežemo po 2,5 % verjetnosti [4]. Ustrezna repa sta na grafu označena s sivo, med njima (med kvantiloma) pa nam ostane intervalna ocena [0,383; 0,816].

Lastnosti Bayesovega pristopa

Konjugirane apriorne porazdelitve. V primeru s kovancem smo opazili, kako prikladna je bila izbira apriorne gostote verjetnosti iz družine porazdelitev beta. Tudi aposteriorna porazdelitev je pripadala isti družini porazdelitev. Temu rečemo, da je družina porazdelitev beta konjugirana k binomski funkciji verjetja. Normirne konstante v aposteriorni gostoti verjetnosti sploh nismo eksplicitno zapisali, upanje in standardni odklon te porazdelitve pa smo določili brez uporabe računalnika.

Pri izbiri drugačne apriorne porazdelitve analiza ne bi bila več tako enostavna. Ker moramo integrirati funkcijo $f_P(p) \cdot f_{X|P}(x|p)$, lahko račun analitično sploh ni mogoč. To nas privede do uporabe računalniških simulacij in pojasni, zakaj je Bayesova statistika svoj razcvet doživela šele dve stoletji po svojem nastanku.

Aposteriorna porazdelitev je kompromis med apriorno porazdelitvijo in informacijami iz podatkov. Pri apriorni gostoti verjetnosti $\mathrm{Beta}(\alpha, \beta)$ in na vzorcu opaženih 7 cifrah in 3 grbih, je aposteriorna gostota verjetnosti oblike

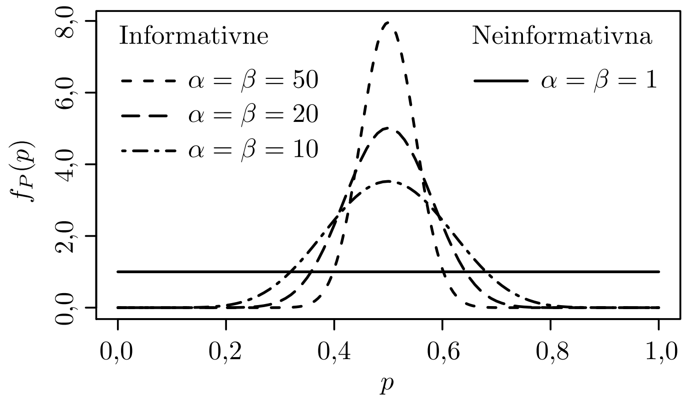
$$f(p|x) \propto p^{\alpha-1}(1-p)^{\beta-1} \cdot p^7(1-p)^3.$$

Pri majhnih α in β imata v aposteriorni porazdelitvi vodilno vlogo vrednosti 7 in 3 iz vzorca, pri velikih α in β pa ima vodilno vlogo apriorna porazdelitev neznanega parametra. V Bayesovi statistiki so aposteriorne porazdelitve vselej kompromis med apriornimi porazdelitvami in podatki. Pri tem z rastočo velikostjo vzorca čedalje večji pomen pridobivajo podatki.

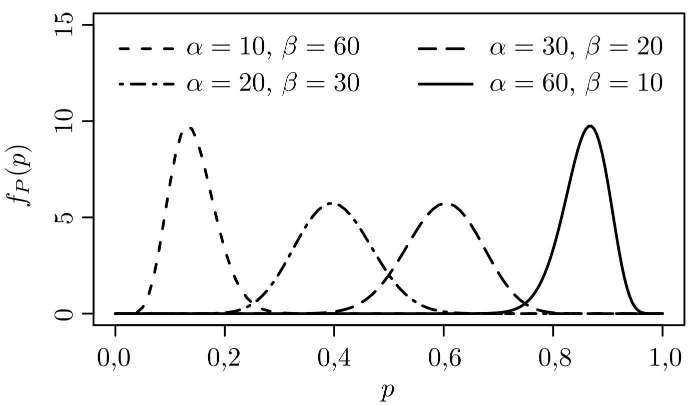
Bayesova statistika je subjektivna. Z izbiro hiperparametrov v Bayesovo analizo vnesemo naše apriorno znanje o proučevanem pojavu, kar je vsekakor lahko subjektivno, ni pa nujno. Če smo v preteklosti že analizirali drug evrski kovanec in ugotovili, da je bil pošten, lahko podobno pričakujemo tudi od kovanca, ki ga proučujemo sedaj, in to izrazimo z izbiro hiperparametrov apriorne porazdelitve. Kadar vsebinsko podobnih raziskav ne poznamo, lahko uporabimo neinformativne apriorne porazdelitve. Statistično analizo lahko ponovimo z različnimi apriornimi porazdelitvami in tako natančno analiziramo njihov vpliv na končne rezultate [4].

Informativne in neinformativne apriorne porazdelitve

Da bi zmanjšali očitke subjektivnosti, so v Bayesovi statistiki vpeljali **neinformativne** apriorne porazdelitve. Vrnimo se k družini porazdelitev beta. Z različnimi izbirami hiperparametrov α in β lahko opišemo zelo različna apriorna znanja. Simetrične porazdelitve dobimo pri izbiri $\alpha = \beta$. Grafe gostot verjetnosti pri različnih parametrih prikazuje slika 5. Opazimo, da pri



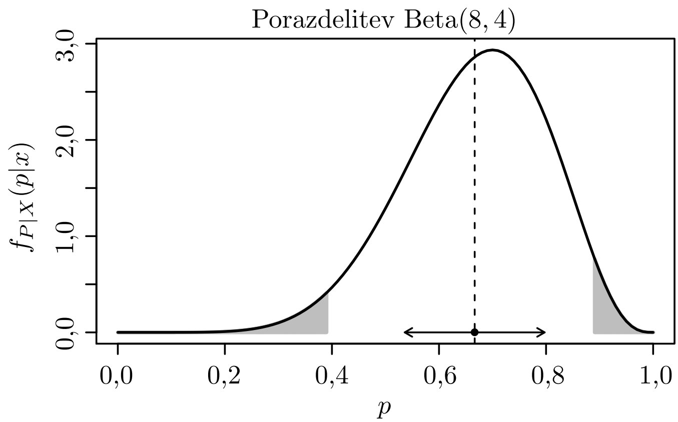
Slika 5. Družina simetričnih porazdelitev beta.



Slika 6. Družina nesimetričnih porazdelitev beta.

izbiri $\alpha = \beta = 1$ dobimo zvezno enakomerno porazdelitev na intervalu $[0, 1]$. Uporaba te porazdelitve pomeni, da nimamo nikakršnih apriornih znanj o možnih vrednostih parametra p. Taka porazdelitev je neinformativna.

Če sta parametra α in β različna, porazdelitve beta postanejo nesimetrične. S takšno izbiro hiperparametrov lahko izrazimo apriorno znanje, da kovanec ni pošten. Grafe gostot verjetnosti pri različnih parametrih prikazuje slika 6.



Slika 7. Aposteriorna gostota verjetnosti parametra p pri neinformativni apriorni porazdelitvi.

Analiza kovanca z neinformativno apriorno porazdelitvijo

Če je apriorna porazdelitev zvezno enakomerna na intervalu $[0, 1]$ oziroma Beta$(1, 1)$, je aposteriorna gostota verjetnosti

$$f_{P|X}(p|x) \propto \underbrace{f_P(p)}_{=1} \cdot f_{X|P}(x|p) = f_{X|P}(x|p)$$

sorazmerna funkciji verjetja [7]. V našem primeru dobimo $p|x \sim$ Beta$(8, 4)$. Njeno upanje je 0,667, standardni odklon 0,131 in 95 % centralni interval aposteriorne verjetnosti [0,390; 0,891]. Njen graf je prikazan na sliki 7.

Bayes in Laplace sta v svojih delih uporabila neinformativne apriorne porazdelitve. Laplace jo je utemeljil s **principom nezadostnega razloga**. Ta pravi, da moramo, kadar odločamo v popolni negotovosti, vse možne vrednosti neznanega parametra obravnavati kot enako verjetne [4].

Posodabljanje ocen v Bayesovem pristopu

Denimo, da smo isti evrski kovanec vrgli še 5-krat ter pri tem dobili 2 cifri in 3 grbe. Skupaj smo torej v 15 metih dobili 9 cifer in 6 grbov. Oglejmo si, kako v statistično analizo vključimo dodatne podatke. Naša informativna apriorna porazdelitev je bila Beta$(4, 4)$. Z upoštevanjem 7 cifer in 3 grbov v desetih metih smo v prejšnjih razdelkih prišli do aposteriorne porazdelitve Beta$(4 + 7, 4 + 3) =$ Beta$(11, 7)$. Podobno lahko sklepamo z razširjenimi podatki (skupaj 15 metov) in dobimo aposteriorno porazdelitev Beta$(4 + 9, 4 + 6) =$ Beta$(13, 10)$.

Obstaja še druga pot. Namesto da analizo začnemo od začetka in združimo stare in nove podatke, lahko nove podatke le dodamo zaključkom stare analize. Pri tem aposteriorno porazdelitev Beta$(11, 7)$ na osnovi začetnega vzorca uporabimo kot apriorno porazdelitev pri analizi dodatnih podatkov. Ko upoštevamo 2 cifri in 3 grbe, pridemo do iste končne aposteriorne porazdelitve Beta$(11+2, 7+3) =$ Beta$(13, 10)$. Zaradi enostavnega dodajanja novih podatkov v že obstoječo analizo je Bayesova statistika uporabna v aplikacijah, za katere je značilen avtomatski zajem podatkov.

Sklepne misli

V članku smo predstavili glavne korake Bayesovega pristopa k ocenjevanju neznanih parametrov. Obravnavali smo neinformativne in informativne apriorne porazdelitve, s katerimi lahko v statistično analizo vključimo naše predhodno (ne)védenje o neznanem parametru. Na vzorcu zbrane informacije predstavlja funkcija verjetja, ki ima pomembno vlogo tudi v frekventistični statistiki. Z Bayesovo formulo nato predhodno znanje in vzorčne informacije združimo v aposteriorno porazdelitev, ki opisuje naše védenje o neznanem parametru po končani raziskavi. Tega najpogosteje povzamemo z matematičnim upanjem in centralnim intervalom aposteriorne verjetnosti. Postopek smo si ogledali na primeru ocenjevanja neznane verjetnosti. Izbira primernih apriornih porazdelitev je v praksi težka naloga in zahteva sodelovanje med statistiki in drugimi znanstveniki. Pravi potencial Bayesovega pristopa zato lahko izkoristimo le z interdisciplinarnim sodelovanjem.

LITERATURA

[1] A. Agresti in B. A. Coull, *Approximate is better than »exact« for interval estimation of binomial proportions*, The American Statistician **52** (1998), 119–126.

[2] L. D. Brown, T. T. Cai in A. DasGupta, *Interval estimation for a binomial proportion*, Statistical Science **16** (2001), 101–133.

[3] J. A. Čibej, *Matematika. Kombinatorika, verjetnostni račun, statistika*, DZS, Ljubljana, 1994.

[4] A. B. Gelman, J. B. Carlin, H. S. Stern in D. B. Rubin, *Bayesian data analysis*, 2. izdaja, Chapman & Hall/CRC, 2004.

[5] H. Hoijtink, *Bayesian data analysis*, v: *The SAGE handbook of quantitative methods in psychology* (R. E. Millsap in A. Maydeu-Olivares), The SAGE, 2009.

[6] S. B. McGrayne, *The theory that would not die*, Yale University Press, New Haven & London, 2011.

[7] J. A. Rice, *Mathematical statistics and data analysis*, 3. izdaja, Thomson Brooks/Cole, 2007.