UNIVERZA V LJUBLJANI
FAKULTETA ZA MATEMATIKO IN FIZIKO

Finančna matematika – 1. stopnja

Simon Škerjanec

Konstrukcija intervalov zaupanja z metodami ponovnega vzorčenja - kljukčeva metoda

Delo diplomskega seminarja

Mentor: doc. dr. Jaka Smrekar

Ljubljana, 2011

Kazalo

Povzetek

Delo diplomskega seminarja vključuje strnjen pregled kljukčevih metod. Poudarek ni toliko na teoretičnem vidiku, pač pa na osnovnih idejah in uporabah. Začne se z razlago ocenjevanja standardne napake po kljukčevi metodi, za primere, v katerih imamo en sam vzorec. Podanih je več zgledov, pri čemer nekateri od njih vsebujejo precej zapletene statistične postopke. Kljukčeva metoda je nato posplošena na druge mere statistične natančnosti, kot sta recimo pristranskost (*angl.* bias) in napaka napovedi (*angl.* prediction error). Podani so tudi zgledi teh idej. Zadnja tretjina diplomskega dela se ukvarja predvsem s tem, kako s kljukčevo metodo računati intervale zaupanja.

Abstract

This is a review of bootstrap methods, concentrating on basic ideas and applications rather than theoretical consideration. It begins with an exposition of the bootstrap estimate of standard error for one-sample situations. Several examples, some involving quite complicated statistical procedures, are given. The bootstrap is then extended to other measures of statistical accuracy such as bias and prediction error. Several more examples are presented illustrating these ideas. The last third of the paper deals mainly with bootstrap confidence intervals.

Math. Subj. Class. (2011): 65G40, 62G05, 62D99

Ključne besede: kljukčeva metoda, ocena standardne napake, intervali zaupanja, neparametrične metode.

Key words: Bootstrap method, estimated standard errors, approximate confidence intervals, nonparametric methods.

Del 1. **Uvod**

Tipičen problem v uporabni statistiki je oceniti neznani parameter θ. Bistveni vprašanji sta naslednji:

(1) Katero cenilko $\hat{\theta}$ moramo uporabiti?

(2) Ko smo cenilko $\hat{\theta}$ enkrat izbrali, kako natančna je kot cenilka za θ?

Kljukčeva metoda omogoča odgovor na drugo vprašanje v zelo splošnih modelih. Osnovana je na uporabi računalnika, ker zamenja ponavadi komplicirano teoretično analizo s precejšnjo količino računanja. Kot bomo videli, lahko kljukčeva metoda na rutinski način ponudi odgovore na vprašanja, ki so veliko prekomplicirana za tradicionalno statistično analizo. V dobi, ko cene računanja eksponentno padajo, so metode, ki v veliki meri vključujejo računalnike, vedno bolj privlačna izbira za obdelavo podatkov celo za relativno enostavne probleme.

Z zgledi in teorijo bom natančno opisal osnove kljukčevih metod. Dokazi in tehnični detajli so izpuščeni, le–te lahko bralec najde v podani literaturi, še posebej v [12].

1. Ocena vzorčnega povprečja

Preden se spustimo v glavnino obravnave, se poglobimo v delovanje kljukčeve metode na problemu, kjer je ne potrebujemo: pri ugotavljanju natančnosti vzorčnega povprečja. Denimo, da naši podatki sestojijo iz naključnega vzorca, katerega porazdelitve F ne poznamo. Naj bo F porazdelitev na $\mathbb{R}$.

$$X_1, X_2, \ldots, X_n \sim F \tag{1.1}$$

Naj bo $X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n$ realizacija vzorca. Iz nje lahko vzorčno povprečje izračunamo kot

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \quad ,$$

pri čemer nas zanima, kako natančno je vzorčno povprečje kot cenilka matematičnega upanja $\theta = E_F(X)$. Če je drugi centralni moment F enak

$$\mu_2(F) \equiv E_F(X^2) - E_F(X)^2 \quad ,$$

lahko standardno napako $\sigma(F; n, \bar{x})$, torej standardni odklon $\bar{x}$ za vzorec velikosti n, ki je porazdeljen po F, izračunamo kot

$$\sigma(F) = \sqrt{\frac{\mu_2(F)}{n}} \tag{1.2}$$

Skrajšan zapis $\sigma(F) = \sigma(F; n, \bar{x})$ lahko uporabljamo, ker je F edina neznanka, saj n in formulo za $\bar{x}$ poznamo. Standardna napaka je tradicionalna mera za natančnost $\bar{x}$, saj v normalnem modelu dobimo eksaktne rešitve oz. intervale. Na žalost enačbe (1.2) ne moremo uporabiti, ker ne poznamo $\mu_2(F)$, lahko pa uporabimo oceno za standardno napako

$$\bar{\sigma} = \sqrt{\frac{\bar{\mu}_2}{n}} \tag{1.3}$$

kjer je

$$\mu_2 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n-1} \quad ,$$

torej nepristranska cenilka $\mu_2(F)$.

Obstaja bolj očiten način za oceno $\theta(F)$. Naj bo $\hat{F}$ empirična porazdelitvena funkcija.

$$\hat{F} : \text{verjetnostna masa } 1/n \text{ na } x_1, x_2, \ldots, x_n \tag{1.4}$$

Tedaj lahko F v enačbi (1.2) preprosto zamenjamo z $\hat{F}$ in dobimo

$$\hat{\theta} \equiv \theta(\hat{F}) = \sqrt{\frac{\mu_2(\hat{F})}{n}} \tag{1.5}$$

za oceno standardne napake $\bar{x}$. To je kljukčeva ocena.

Ker je

$$\hat{\mu_2} \equiv \mu_2(\hat{F}) = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n} \quad , \tag{1.6}$$

$\hat{\sigma}$ ni čisto enako kot $\bar{\sigma}$, a je razlika v večini primerov premajhna, da bi bila pomembna.

Seveda v tem primeru v bistvu ne rabimo alternative za enačbo (1.3). Problemi nastanejo šele, ko želimo standardno napako za cenilke, ki so bolj komplicirane od $\bar{x}$, recimo mediano, kako korelacijo ali pa strmino regresije. V večini primerov ekvivalent formuli (1.2), ki standardno napako izrazi kot funkcijo porazdelitvene funkcije vzorca F, sploh ne obstaja. Posledično enačbe, kot je (1.3), za večino statistik ne obstajajo.

Na tem mestu vstopi računalnik. Izkaže se, da lahko vedno izračunamo kljukčevo oceno $\hat{\sigma} = \sigma(\hat{F})$, ne da bi poznali enostavno formulo za $\sigma(F)$. Izračun $\hat{\sigma}$ je enostavna uporaba metode Monte Carlo, ki je podrobneje opisana v nadaljevanju. V dobrem računskem okolju daje kljukčeva metoda statistiku enostavno formulo, kot je enačba (1.3), za katerokoli statistiko, ne glede na njeno kompleksnost.

2. Standardni inteval zaupanja

Standardne napake so grobe, a uporabne mere statistične natančnosti. Pogosto so v uporabi za ugotavljanje približnih intervalov zaupanja za neznani parameter θ

$$\theta \in \hat{\theta} \pm \hat{\sigma} z^{(\alpha)} \tag{2.1}$$

kjer je $z^{(\alpha)}$ $100 \cdot \alpha$-ti centil standardne normalne porazdelitve, npr $z^{0\,95} = 1{,}645$. Interval (2.1) je včasih dober, včasih pa spet ne. Razdelka 3 in 4 se ukvarjata z malo bolj sofisticirano uporabo kljukčeve metode, ki daje boljše približke za intervale zaupanja kot (2.1).

Standardni interval (2.1) je osnovan na ideji, da dobesedno vzamemo normalno porazdelitev na velikem vzorcu

$$\frac{\hat{\theta} - \theta}{\hat{\sigma}} \sim N(0{,}1) \quad .$$

Uporabni statistiki uporabljajo različne trike, da to oceno izboljšajo. Na primer, če je θ korelacijski koeficient in $\hat{\theta}$ vzorčna korelacija, transformacija

$$\phi = \operatorname{arc}\tanh(\phi)$$
$$\hat{\phi} = \operatorname{arc}\tanh(\hat{\theta})$$

zelo izboljša normalno aproksimacijo, vsaj v primerih, ko je resnična porazdelitev vzorca dvorazežna normalna. V takem primeru je pravilen pristop, da najprej transformiramo, nato izračunamo (2.1) za ϕ, nato pa ta interval spremenimo nazaj v merilo θ.

Videli bomo, da lahko intervali, izračunani po ključkevi metodi, avtomatsko inkorporirajo take trike, torej od analitika ne zahtevajo uporabe posebnih metod, kot je prejšnja transformacija z $\operatorname{arc\,tanh}$, za vsako novo situacijo.

Pomembna tema nadaljevanja je zamenjava teoretične analize s surovo računsko močjo. To seveda ni argument proti vsej teoriji, le proti nepotrebni teoriji. Najpogostejše statistične metode so bile izpopolnjene v dvajseth in tridesetih letih 20. stoletja, ko je bilo računanje počasno in drago. Zdaj, ko je računanje hitro in poceni, lahko pričakujemo in upamo na spremembe v statistični metodologiji. Opisal bom le eno od potencialnih sprememb, [8] pa se ukvarja z nekaterimi drugimi.

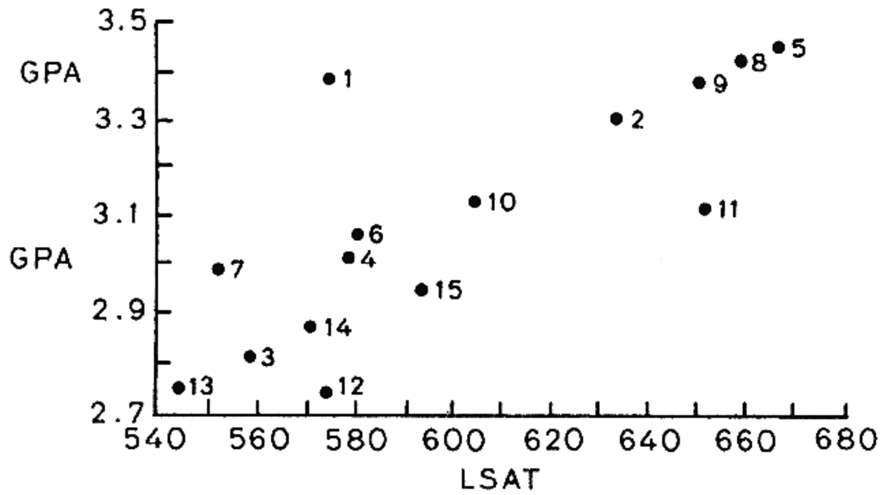
Del 2. Kljukčeva ocena standardne napake

Ta razdelek je podrobnejši opis kljukčeve ocene standardne napake. Za zdaj predpostavimo, da je realizacija vzorca $y = (x_1, x_2, \ldots, x_n)$ sestavljena iz neodvisnih in enako porazdeljenih (nep) opažanj slučajnega vektorja $X_1, X_2, \ldots, X_n$, kot v (1.1). Naj bo F neznana porazdelitev na $\mathscr{H}$, prostoru skupnih vzorcev opažanj. Poznamo želeno statistiko $\theta(y)$, ki ji hočemo prirediti oceno standardne napake.

Na sliki 1 je prikazan primer. $\mathscr{H}$ je v tem primeru $\mathbb{R}^2_+$, pozitivni kvadrant ravnine. Imamo $n = 15$ opažanj v dveh spremenljivkah, katerih vsaka predstavlja eno od ameriških pravnih fakultet. Vsaka točka x_i je sestavljena iz dveh končnih statistik za letnik, ki je vpisal leta 1973. (LSAT so standardizirani testi jezika, GPA pa je povprečje ocen).

$$x_i = (LSAT_i, GPA_i) \tag{2.2}$$

Izmerjeni koeficient Pearsonove korelacije za teh 15 točk je $\hat{\theta} = 0{,}776$. Vedeti želimo standardno napako te ocene.



Slika 1. Podatki pravnih fakultet [8]. Točke (po vrstnem redu fakultet) so: (576, 3,39), (635, 2,30), (558, 3,81), (578, 3,03), (666, 3,44), (580, 3,07), (555, 3,00), (661, 3,43), (651, 3,36), (605, 3,13), (653, 3,12), (575, 2,74), (545, 2,76), (572, 2,88), (594, 2,96),

Naj bo $\sigma(F)$ standardna napaka $\hat{\theta}$ kot funkcija neznane porazdelitve vzorca F:

$$\sigma(F) = \sqrt{\mathrm{Var}_F(\hat{\theta}(y)} \ . \tag{2.3}$$

Seveda je $\sigma(F)$ tudi funkcija velikosti vzorca n in oblike statistike $\hat{\theta}(y)$, a ker sta obe znani, tega ni treba poudarjati v notaciji. Kljukčeva ocena standardne napake je

$$\hat{\sigma} = \sigma(\hat{F}) \tag{2.4}$$

kjer je $\hat{F}$ empirična porazdelitvena funkcija (1.3), z verjetnostjo $^1\!/\!_n$ za vsako od opažanj x_i. V primeru pravnih fakultet je $\hat{F}$ porazdelitev, ki vsaki od točk da težo 1/15 kot v sliki 1, $\hat{\sigma}$ pa je standardna deviacija od korelacijskega koeficienta za 15 neodvisnih in enako porazdeljenih meritev iz $\hat{F}$.

3. Algoritem Monte Carlo

V večini primerov, tudi v prejšnjem, ni enostavnega izraza za funkcijo $\theta(F)$ v (2.3). Kljub temu je $\hat{\theta} = \sigma(F)$ enostavno izračunati numerično s pomočjo Monte Carlo algoritma, za katerega bomo potrebovali naslednje oznake: Naj bo $y^* = (x_1^*, x_2^*, \dots x_n^*)$ n neodvisnih opažanj iz porazdelitve $\hat{F}$, ki jim rečemo kljukčev vzorec. Ker je $\hat{F}$ empirična porazdelitvena funkcija, je kljukčev vzorec ista stvar kot naključen vzorec s ponavljanjem velikosti n, izbranih po realizaciji vzorca $\{x_1, x_2, \dots, x_n\}$.

Monte Carlo algoritem je sestavljen iz treh korakov:

(i) z uporabo generatorja naključnih števil neodvisno izberi veliko število B kljukčevih vzorcev $y^*(1), y^*(2), \dots, y^*(B)$

(ii) za vsak vzorec $y^*(b)$ izračunaj relevantno statistiko $\hat{\theta}(y^*(b))$

(iii) izračunaj standardni odklon vzorca za vrednosti $\hat{\theta}(b)$

$$\hat{\sigma}_B = \sqrt{\frac{\sum_{b=1}^{B} \left(\hat{\theta}(b) - \hat{\theta}^*(\cdot)\right)}{B-1}} \tag{3.1a}$$

$$\hat{\theta}^*(\cdot) = \frac{\sum_{b=1}^{B} \hat{\theta}^*(b)}{B} \quad . \tag{3.1b}$$

Enostavno je videti, da se $\hat{\sigma}_B$ približuje $\hat{\sigma} = \sigma(\hat{F})$, ko gre $B \to \infty$, torej kljukčevi oceni standardne napake. Postopek je v bistvu izračun standardnega odklona z Monte Carlo vzorčenjem. Obravnava potrebne velikosti B je v razdelku 9, za večino primerov pa je

B med 50 in 200 dovolj dober. V nadaljevanju bomo razliko med $\hat{\sigma}_B$ in $\hat{\sigma}$ ignorirali ter obema rekli kar $\hat{\sigma}$.

Zakaj je vsak kljukčev vzorec enake velikosti n kot že osnovni vzorec? Spomnimo se, da je $\sigma(F)$ pravzaprav $\sigma(F; n, \hat{\theta})$, torej standardna napaka za statistiko $\hat{\sigma}()$, temelječo na naključnem vzorcu velikosti n iz neznane porazdelitve F. Kljukčeva ocena $\hat{\sigma}$ je v bistvu $\sigma(F; n, \hat{\sigma})$, izračunana na $F = \hat{F}$. Monte Carlo algoritem ne bo konvergiral k $\hat{\sigma}$, če se velikost kljukčevih vzorcev razlikuje od pravega n. V [3] je pokazano, kako lahko algoritem popravimo, da dobimo pravi $\hat{\sigma}$ v primeru, da je velikost kljukčevih vzorcev različna od n, a se do sedaj potreba po tem še ni pokazala.

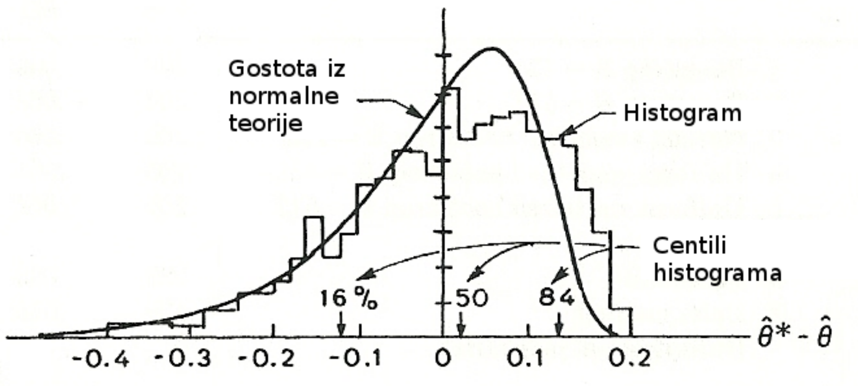
Slika 2 kaže histogram za $B = 1000$ za korelacijski koeficient podatkov pravnih fakultet. Zaradi jasnosti je abscisa narisana v enotah $\hat{\theta}^* - \hat{\theta} = \hat{\theta}^* - 0{,}776$. Formula (3.1) da za kljukčevo oceno napake $\hat{\sigma} = 0{,}127$. To lahko primerjamo z običajno oceno napake za $\hat{\theta}$, ki jo da normalna teorija:

$$\hat{\theta}_{\mathrm{NORM}} = \frac{1 - \hat{\theta}^2}{\sqrt{n-3}} = 0{,}115 \tag{3.2}$$

Za računanje porabimo za faktor B več časa kot za izračun statistike za originalni vzorec.

Opomba: Implementacija Monte Carlo algoritma, ki vodi do $\hat{\sigma}_B$ (3.1) je enostavna. V Stanfordovi verziji programskega jezika S je prof.Arthur Owen dodal en sam ukaz, ki za vsako statistiko v S-ovem katalogu izračuna kljukčevo oceno. Kot primer so podatki iz slike 2 dobljeni preprosto z ukazom

```
tboot(lawsdata, correlation, B=1000) .
```



Slika 2. Histogram za statistiko $\hat{\theta}^*$ pri $B = 1000$ za podatke pravnih fakultet. Porazdelitvena gostota, ki jo dobimo iz normalne teorije, ima podobno obliko, a na gornjem repu pada hitreje.

4. Neparametrična ocena največjega verjetja

Obstaja še en način za opis klјukčeve standardne napake: $\hat{F}$ je neparametrična ocena največjega verjetja (ONV) neznane porazdelitve F [18]. To pomeni da je klјukčeva ocena $\hat{\sigma} = \sigma(\hat{F})$ neparametrična CNV za $\sigma(F)$, resnične standardne napake.

V resnici ni nobene zahteve, da bi morali klјukčevo metodo izvajati neparametrično. Recimo da verjamemo, da je pravilna porazdelitev vzorca F normalna v dveh spremenljivkah. V tem primeru lahko F ocenimo iz parametrične CNV za $\hat{F}_{\text{NORM}}$, normalne porazdelitve v dveh spremenljivkah z enakim vektorjem povprečja in kovariančno matriko kot podatki. V tem primeru lahko klјukčeve vzorce v koraku (i) jemljemo kar iz F_{NORM} namesto iz empirične porazdelitvene funkcije, koraka (ii) in (iii) pa izpeljemo nespremenjena.

Gladka krivulja na sliki 2 kaže rezultat tega "Normalnega klјukčevega" postopka na podatkih pravnih fakultet. V resnici je klјukčevo vzorčenje tu nepotrebno, saj lahko zaradi Fisherjeve formule za vzorčno gostoto korelacijskega koeficienta v primeru dvorazsežne normalne porazdelitve (gl. 32. poglavje Johnson in Kotz, 1970).

O tej gostoti lahko razmišljamo kot o klјukčevi porazdelitvi za $B = \infty$. Izraz (3.2) je dober približek $\hat{\sigma}_{\text{NORM}} = \sigma(\hat{F}_{\text{NORM}})$, parametrične klјukčeve ocene standardne napake.

5. Prednosti, slabosti in natančnost klјukčeve metode

Ko primerjamo prednosti in slabosti klјukčeve metode, si je vredno zapomniti, da so vse običajne formule za ocenjevanje standardnih napak, kot npr. $\hat{\mathscr{I}}^{-1/2}$, kjer je $\mathscr{I}$ opažena Fisherjeva informacija, so v bistvu klјukčeve ocene, izpeljane v parametričnem okviru. To je natančno razloženo v [13], razdelek 5. Očitni neparametrični algoritem (i)-(iii) ima to odliko, da se izogne vsem parametričnim predpostavkam, vsem približkom (kot recimo ti v povezavi z izrazom za standardno napako CNV s pomočjo Fisherjeve informacije) in v resnici vsem analitičnim problemom vseh oblik. Analitik ima tako možnost, da dobi standardne napake za izredno komplicirane cenilke, pri čemer mu je edina ovira čas za računanje.

TABELA 1. Eksperiment v vzorčenju, pri katerem primerjamo oceni standardne napake po kljukčevi in *jackknife* metodi, za 25% odrezano povprečje. Velikost vzorca je $n = 15$.

	Standardna normalna F			Negativna eksponentna F		
	Povpr.	SD	CV	Povpr.	SD	CV
kljukčeva $\hat{\sigma}$ ($B = 200$)	0,287	0,071	0,25	0,242	0,078	0,32
Jackknife $\hat{\sigma}_J$	0,280	0,084	0,30	0,224	0,085	0,38
Resnična (minimalni CV)	0,286		(0,19)	0,232		(0,27)

Kako dobro deluje kljukčeva metoda? Tabela 1 nam prikazuje odgovor v enem primeru. Tu je $\mathscr{H}$ realna premica, $n = 15$, statistika $\hat{\theta}$, ki nas zanima, pa je 25% odrezano povprečje. Če je resnična porazdelitev vzorca $F = N(0{,}1)$, je resnična standardna napaka $\sigma(F) = 0{,}286$. Kljukčeva ocena $\hat{\sigma}$ je skoraj nepristranska, in povprečno enaka 0,287 v eksperimentu z velikim vzorcem. Standardni odklon kljukčeve ocene $\hat{\sigma}$ je v tem primeru sam po sebi le 0,071, s koeficientom variacije $0{,}071/0{,}287 = 0{,}25$. (Opazimo, da sta v tabeli 1 dva nivoja Monte Carlo: najprej izbira dejanskih vzorcev $y = (x_1, x_2, \ldots, x_n)$ iz F, nato pa še izbira kljukčevih vzorcev $(x_1^*, x_2^*, \ldots x_n^*)$ pri fiksnem y. S kljukčevimi vzorci izračunamo $\hat{\sigma}$ za fiksno vrednost y. Standardni odklon z vrednostjo 0,071 je mišljen kot raznolikost $\hat{\sigma}$ zaradi naključne izbire y.)

Tabela 2 se vrne k primeru korelacijskega koeficienta $\hat{\theta}$. Namesto realnih podatkov imamo eksperiment v vzorčenju, v katerem je resnični F dvorazsežna normalna porazdelitev, resnična korelacija $\theta = 0{,}50$ in velikost vzorca $n = 14$. Tabela 2 je skrajšana iz večje tabele v [10], v kateri so nekatere metode za ocenjevanje standardne napake zahtevale, da je n sod.

Leva stran tabele 2 se nanaša na $\hat{\theta}$, medtem ko se desna nanaša na

$$\hat{\phi} = \operatorname{arc}\tanh(\hat{\theta}) = \frac{\log(1+\hat{\theta})}{2(1-\hat{\theta})} \quad .$$

Za vsako cenilko standardne napake je srednja kvadratna napaka (*angl.* mean squared error) $\sqrt{E(\hat{\sigma} - \sigma)^2}$ podana v stolpcu z naslovom $\sqrt{\text{MSE}}$.

Kljukčeva metoda je bila uporabljena prvič z $B = 128$ in drugič z $B = 512$, pri čemer je večja vrednost ocene le malo izboljšala, v skladu z rezultati v razdelku 9. Dodatno povečevanje B bi bilo brezpredmetno. Da se pokazati, da je $\sqrt{\text{MSE}}$ za $\hat{\theta}$ pri $B = \infty$ enak 0,63, le 0,001 manj kot tisti za $B = 512$. Cenilka iz normalne teorije (3.2), za katero vemo, da je za ta eksperiment idealna, ima $\sqrt{\text{MSE}} = 0{,}056$.

Lahko tudi naredimo kompromis med popolnoma neparametrično $\hat{\sigma}$ ter popolnoma parametrično kljukčevo oceno $\hat{\sigma}_{\text{NORM}}$. To je narejeno v 3., 4., in 5,vrstici tabele 2. Naj bo

$$\hat{\Sigma} = \sum_{i=1}^{n} \frac{(x_i - \bar{x})(x_i - \bar{x})'}{n}$$

matrika vzorčne kovariance opažanja. *Normalna glajena kljukčeva metoda* izbere vzorec iz $\hat{F} \oplus N_2(0{,}0\ 25\hat{\Sigma})$, kjer $\oplus$ pomeni konvolucijo. Ta postopek oceni F z mešanico n enakovrednih porazdelitev $N_2(x_i, 0{,}25\Sigma)$, kar je enako, kot če bi F ocenili z normalnim oknom. Vsaka točka x_i^* v zglajenem vzorcu je vsota naključno izbranih originalnih točk x_j, ki jim prištejemo dvorazsežno normalno točko $z_j \sim N_2(0, 0{,}25\Sigma)$. Glajenje nima velikega učinka na levi strani tabele, v primeru $\hat{\phi}$ pa je izjemno efektivno. Ta rezultat je sumljiv, saj je resnična porazdelitev dvorazsežna normalna, funkcija $\hat{\phi} = \operatorname{arc}\tanh(\hat{\theta})$ pa je izbrana specifično zato, da bi v družini dvorazsežnih normalnih porazdelitev imela praktično konstantno standardno napako. *Enakomerno glajena kljukčeva metoda* vzorči iz $\hat{F} \oplus \mathfrak{U}(0, 0{,}25\hat{\Sigma})$, kjer je $\mathfrak{U}$ enakomerna porazdelitvena funkcija na rombu, izbranem tako, da ima $\mathfrak{U}$ povprečni vektor 0 in kovariančno matriko $0{,}25\hat{\Sigma}$. Ta metoda nekoliko zmanjša $\sqrt{\text{MSE}}$ za obe strani tabele.

Tabela 2. Ocene standardne napake korelacijskega koeficienta $\hat{\theta}$ in $\hat{\phi} = \operatorname{arc}\tanh\hat{\theta}$, pri velikosti vzorca $n = 14$, vzetega iz dvorazsežne normalne porazdelitve F z resnično korelacijo $\rho = 0{,}5$ (iz večje tabele v [10])

		Skupna statistika za 200 poskusov							
		Ocene standardne napake $\hat{\theta}$				Ocene standardne napake $\hat{\phi}$			
	B	Povpr.	SD	CV	$\sqrt{\text{MSE}}$	Povpr.	SD	CV	$\sqrt{\text{MSE}}$
1. kljukčeva	128	0,206	0,066	0,32	0,067	0,301	0,065	0,22	0,065
2. kljukčeva	512	0,206	0,063	0,31	0,064	0,301	0,062	0,21	0,062
3. NGK	128	0,200	0,060	0,30	0,063	0,296	0,041	0,14	0,041
4. EGK	128	0,205	0,061	0,030	0,062	0,298	0,058	0,19	0,058
5. EGK	512	0,205	0,059	0,29	0,060	0,296	0,052	0,18	0,052
8. NT	/	0,217	0,056	0,26	0,056	0,302	0	0	0,003

Legenda:
NGK - Normalna glajena kljukčeva
EGK - Enakomerno glajena kljukčeva
NT - Normalna teorija

Del 3. Ključčevi intervali zaupanja

6. Standardna metoda

V tem razdelku bom predstavil tri tesno povezane metode uporabe ključčeve metode za določanje intervalov zaupanja. Obravnava je osnovana na enostavnih parametričnih modelih, kjer je logične osnove ključčevih metod najlažje razumeti.

Obravnavali smo že način za izračun $\hat{\sigma}$, torej oceno standardne napake cenilke $\hat{\theta}$. V praksi sta $\hat{\theta}$ in $\hat{\sigma}$ ponavadi uporabljeni skupaj, da z njuno pomočjo lahko dobimo približne intervale zaupanja $\theta \in \hat{\theta} \pm \hat{\sigma} z^{(\alpha)}$, (2.1), kjer je $z^{(\alpha)}$ $100 \cdot \alpha$-ti centil standardne normalne porazdelitve. Interval (2.1) ima približno verjetnost pokritja $1 - 2\alpha$. Za primer pravnih fakultet iz razdelka 2 nam za resnični korelacijski koeficient vrednosti

$$\hat{\theta} = 0{,}776$$
$$\hat{\sigma} = 0{,}115$$
$$z^{(0{,}05)} = -1{,}645$$

kot 90% centralni interval zaupanja dajo $\theta \in [0{,}587, 0{,}965]$.

Intervalu iz (2.1) bomo rekli *standardni interval za* θ. Kadar delamo znotraj parametričnih družin, kot je dvorazsežna normalna, dobimo $\hat{\sigma}$ iz (2.1) ponavadi tako, da odvajamo logaritem funkcije verjetja (gl. [Rao (1973), razdelek 5a]), čeprav v kontekstu tega članka raje uporabimo parametrično ključčevo oceno σ, npr.$\hat{\sigma}_{\mathrm{NORM}}$ iz razdelka 2.

Standardni intervali so izredno uporabno statistično orodje. Imajo to odliko, da so avtomatski – napišemo lahko namreč program, ki (2.1) dobi neposredno iz podatkov y in oblike porazdelitvene gostote za y, zato od statistika ne zahteva nobenega dodatnega dela. Po drugi strani so standardni intervali lahko precej nenatančni, kot je lepo vidno iz tabele 3. Standardni interval (2.1) pri uporabi $\hat{\sigma}_{\mathrm{NORM}}$ (3.2), je presenetljivo različen od eksaktnih intervalov, ki nam jih da normalna teorija pri predpostavki dvorazsežne normalne porazdelitve vzorca F.

V tem primeru je dobro znano, da je bolje narediti transformacijo $\hat{\phi} = \operatorname{arc\,tanh}(\hat{\theta})$, $\phi = \operatorname{arc\,tanh}(\theta)$, uporabiti (2.1) v merilu ϕ in nato dobljeni interval zaupanja transformirati nazaj v merilo θ. Rezultat tega postopka, prikazan v tretji vrstici tabele, je precej bliže dejanskemu, a na tej transformaciji ni ničesar avtomatskega. Za kakšno drugo statistiko, torej kakšno, ki ni ravno korelacijski koeficient, ali pa kako drugo družino porazdelitev, bi bržko rabili kakšen drugačen trik, da bi se (2.1) obnašala zadovoljivo.

Ključčevo metodo lahko uporabimo za avtomatsko določanje približnih intervalov zaupanja. Sledeča obravnava je obnovljena iz [14][15] in [11], 10. poglavje. Četrta

Tabela 3. Natančni in približni 90% centralni intervali zaupanja za θ, pravi korelacijski koeficient, za podatke pravnih fakultet iz Slike 1. D/L je razmerje med dolžino desne in leve stranji intervala okoli $\hat{\theta}$.

1. Eksaktni (normalna teorija)	$[0{,}496, 0{,}898]$	D/L = 0,44
2. Standardni (2.1)	$[0{,}587, 0{,}966]$	D/L = 1,00
3. Transformiran standardni	$[0{,}508, 0{,}907]$	D/L = 0,49
4. Parametrični ključčevi (BC)	$[0{,}488, 0{,}900]$	D/L = 0,43
5. Neparametrični ključčevi (BC_a)	$[0{,}43, 0{,}92]$	D/L = 0,42

vrstica tabele 3 kaže, da je parametričen kljukčev interval za korelacijski koeficient θ praktično identičen pravemu. V tem smislu "parametričen" pomeni, da kljukčeva metoda začne iz CNV dvorazsežne normalne porazdelitve $\hat{F}_{\text{NORM}}$, kot pri krivulji na sliki 2. Ta odličen rezultat ni le naključje. kljukčeva metoda v uporabi v 4.vrstici $\hat{\theta}$ v bistvu transformira v najboljše (najbolj normalno) merilo, poišče pripadajoči interval ter rezultat transformira nazaj v merilo θ. Vse to naredi avtomatsko s kljukčevim algoritmom, pri čemer od statistika ne zahteva dodatne pozornosti. Cena tega je morda velika količina računanja, recimo $B = 1000$, kot opisano v razdelku 9.

Naj bo $\hat{G}(s)$ parametrična kljukčeva kumulativna porazdelitvena funkcija $\hat{\theta}^*$,

$$\hat{G}(s) = P_* \left\{ \hat{\theta}^* < s \right\} \tag{6.1}$$

kjer je P_* verjetnost, izračunana na podlagi kljukčeve porazdelitve $\hat{\theta}^*$. Na sliki 2 dobimo $\hat{G}(s)$ kot integral krivulje iz normalne teorije. V nadaljevanju bomo predstavili tri različne kljukčeve intervale zaupanja, od najmanj pa do najbolj splošnega. Razlikujejo se predvsem v tem, kateri centili so uporabljeni.

7. Centilna metoda

Pri najpreprostejši metodi kot oceno centralnega $(1 - 2\alpha)$ intervala zaupanja za θ vzamemo kar

$$\theta \in \left[\hat{G}^{-1}(\alpha),\ \hat{G}^{-1}(1-\alpha) \right] \ .$$

Tej metodi pravimo *centilna metoda* [11], razdelek 10.4. Interval centilne metode je le interval med $100 \cdot \alpha$ in $100 \cdot (1 - \alpha)$ centiloma kljukčeve porazdelitve $\hat{\theta}^*$.

Vpeljimo oznako $\theta[\alpha]$ za končno točko približnega intervala zaupanja za θ, tako da je

$$\theta \in [\theta[\alpha], \theta[1 - \alpha]]$$

centralni $1 - 2\alpha$ interval. Različne metode bomo ločili po subskriptih; tako bo imel centilni interval končni točki

$$\theta_P[\alpha] \equiv \hat{G}^{-1}(\alpha) \ , \tag{7.1}$$

standardni interval zaupanja pa

$$\theta_S[\alpha] = \hat{\theta} + \hat{\sigma} z^{(\alpha)} \ . \tag{7.2}$$

Vrstici 1 in 2 v tabeli 4 povzemata ti dve definiciji.

Recimo, da je kljukčeva kumulativna porazdelitvena funkcija $\hat{G}$ popolnoma normalna, recimo

$$\hat{G}(s) = \Phi \left(\frac{s - \hat{\theta}}{\hat{\sigma}} \right) \ , \tag{7.3}$$

kjer je

$$\Phi(s) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{s} e^{\frac{-t^2}{2}} \mathrm{d}t \ ,$$

torej standardna normalna porazdelitev. Z drugimi besedami, naj bo $\hat{\theta}^*$ porazdeljena po kljukčevi porazdelitvi $N(\hat{\theta}, \hat{\sigma}^2)$. V tem primeru se standardna metoda in centilna metoda ujemata: $\theta_S[\alpha] = \theta_P[\alpha]$. V primerih, kot je npr.tisti na sliki 2, kjer je $\hat{G}$ precej nenormalna, pa je standardni interval precej različen od (7.1). Kaj je torej boljše?

Tabela 4. Štiri metode za postavljanje približnih intervalov zaupanja za realen parameter θ. Vsaka metoda je pravilna pod splošnejšimi predpostavkami kod prejšnja. Metode 2, 3 in 4 so definirane s pomočjo centilov $\hat{G}$, kljukčeve porazdelitve.

Metoda	Oznaka	α-končna točka	Pravilna če	Konst.
1. S	$\theta_S[\alpha]$	$\hat{\theta} + \hat{\sigma} z^{(\alpha)}$	$\hat{\theta} \sim N(\theta, \sigma^2)$	σ
			Obstaja monotona transformacija, da $\hat{\phi} = g(\hat{\theta})$, $\phi = g(\theta)$, ki zadošča:	
2. C	$\theta_P[\alpha]$	$\hat{G}^{-1}(\alpha)$	$\hat{\phi} \sim N(\phi, \tau^2)$	τ
3. BC	θ_{BC}	$\hat{G}^{-1}(\Phi\{2z_0 + z^{(\alpha)}\})$	$\hat{\phi} \sim N(\phi - z_0\tau, \tau^2)$	z_0, τ
4. BC_a	θ_{BC_a}	$\hat{G}^{-1}\left(\Phi\left\{z_0 + \frac{z_0 + z^{(\alpha)}}{1 - a(z_0 + z^{(\alpha)})}\right\}\right)$	$\hat{\phi} \sim N(\phi - z_0\tau_\phi, \tau_\phi^2)$ kjer $\tau_\phi = 1 + a\phi$	z_0, a

Legenda:
S - Standardna
C - Centilna
BC - Centilna metoda, popravljena za pristranskost

Da bi lako odgovorili na to vprašanje, si poglejmo najenostavnejšo možno situacijo, kjer je za vse θ

$$\hat{\theta} \sim N(\theta, \sigma^2) \tag{7.4}$$

Imamo torej en sam neznan parameter θ, brez ostalih nadležnih parametrov in eno samo statistiko $\hat{\theta}$, ki je okoli θ porazdeljena normalno s konstantno standardno napako σ. V tem primeru je parametrična kljukčeva kumulativna porazdelitvena fukncija podana z (7.3), tarej je $\sigma_S[\alpha] = \sigma_P[\alpha]$. (Kljukčeva ocena $\hat{\sigma}$ je enaka σ).

Kaj pa če bi namesto (7.4) imeli za vse θ

$$\hat{\phi} \sim N(\phi, \tau^2) \quad , \tag{7.5}$$

za neko monotono transformacijo $\hat{\phi} = g(\hat{\theta})$, $\phi = g(\theta)$, kjer je τ konstanta? V primeru korelacijskega koeficienta je bila funkcija g kar $\operatorname{arc\,tanh}$. Standardne končne točke (7.1) so sedaj lahko izjemno netočne. Vseeno je enostavno preveriti, da so centilne točke še vedno pravilne. "Pravilne" v tem primeru pomeni, da je (7.1) slika očitnega intervala za ϕ, torej $\hat{\phi} \pm \tau z^{(\alpha)}$ nazaj v merilo θ:

$$\theta_P[\alpha] = g^{-1}\left(\hat{\phi} + \tau z^{(\alpha)}\right) \quad .$$

Pravilne so tudi v smislu, da imajo točno zahtevano verjetnost konvergence $1 - 2\alpha$.

Rekli bi lahko tudi, da so centilni intervali invariantni za transformacije oblike

$$\phi_P[\alpha] = g(\theta_P[\alpha]) \tag{7.6}$$

za katerokoli monotono transformacijo g. To pomeni, da pravilnost v kakšnem transformiranem merilu $\phi = g(\theta)$ implicira tudi točnost v originalnem merilu θ. Statistiku zato ni potrebno poznati normalizacijske transformacije g, le da ta obstaja. Definicija (7.1) avtomatsko poskrbi za vse pisanje, potrebno za uporabo narmalizacijskih transformacij za intervale zaupanja.

Fisherjeva teorija ocenjevanja največjega verjetja pravi, da smo vsaj do prvega reda asimptotične aproskimacije vedno v situaciji (7.4). Poleg tega smo za isti red aproksimacije za vsak g tudi v situaciji (7.5). Da bi lahko razlikovali med standardnimi in ključkevimi intevali zaupanja, potrebujemo asimptotično teorijo višjega reda, kot v [14] in [15]. Točno ti členi višjega reda pogosto naredijo točne intervale zelo asimetrične okoli CNV $\hat{\theta}$, kot v tabeli 3. Ključkevi intervali dobro ujamejo to asimetrijo.

Centilna metoda avtomatsko vključi normalizacijske transformacije, kot pri prehodu iz (7.4) v (7.5). Izkaže se, da obstajata dva načina, na katera je lahko predpostavka (7.4) zavajujoča, prvi od katerih se nanaša na možno pristranskost $\hat{\theta}$. Kot primer si poglejmo $f_\theta(\hat{\theta})$, družino porazdelitvenih gostot za opaženi korelacijski koeficient $\hat{\theta}$ pri vzorčenju $n = 15$ vrednosti iz dvorazsežne normalne porazdelitve z resničnim korelacijskim koeficientom θ. V resnici je enostavno videti, da monotona transformacija $\hat{\phi} = g(\hat{\theta})$, $\phi = g(\theta)$, ki bi to družino slikala v $\hat{\phi} \sim N(\phi, \tau^2)$ kot v (7.5), ne obstaja. Če bi obstajala, bi bila verjetnost

$$P_\theta\left\{\hat{\theta} < \theta\right\} = P_\phi\left\{\hat{\phi} < \phi\right\} = 0{,}50 \quad ,$$

obenem pa za $\theta = 0{,}776$ pri integriranju porazdelitvene gostote $f_{0,776}(\hat{\theta})$ dobimo $P_{\theta=0,776}\{\hat{\theta} < \theta\} = 0{,}431$.

8. Centilna metoda, popravljena za pristranskost

Centilna metoda, popravljena za pristranskost (metoda BC, *bias-corrected*), v 3. vrstici tabele 4, naredi popravek za ta tip pristranskosti. Naj bo

$$z_0 \equiv \Phi^{-1}\left\{\hat{G}(\hat{\theta})\right\} \tag{8.1}$$

kjer je Φ^{-1} inverzna funkcija standardne normalne kumulativne porazdelitvene funkcije. Metoda BC ima α-končno točko

$$\theta_{BC}[\alpha] \equiv \hat{G}^{-1}\left(\Phi\left\{2z_0 = z^{(\alpha)}\right\}\right) \quad . \tag{8.2}$$

Opomba: če je $\hat{G}(\hat{\theta}) = 0{,}50$, tj.če je polovica ključkeve porazdelitve θ^* manj kot opažena vrednost $\hat{\theta}$, potem je $z_0 = 0$ in $\theta_{BC}[\alpha] = \theta_P[\alpha]$. V nasprotnem primeru definicija (8.2) popravi pristranskost.

V [11], razdelek 10.7 je pokazano, da je interval BC za θ točno pravilen, če je

$$\hat{\phi} \sim N(\phi - z_0\tau, \tau^2) \tag{8.3}$$

za neko monotono transformacijo $\hat{\phi} = g(\hat{\theta})$, $\phi = g(\theta)$ in neko konstanto z_0. Na prvi pogled morda ne izgleda, kot da je (8.3) splošnješa od (7.4), a v resnici je ta popravek pristranskosti dostikrat pomemben.

V primeru tabele 3 nam da centilna metoda (7.1) 90% interval $[0{,}536, 0{,}911]$, v primerjavi z intervalom BC, ki je $[0{,}496, 0{,}898]$. Po definiciji končne točke eksaktnega intervala zadoščajo

$$\begin{aligned} P_{\theta=496}\{\hat{\theta} > 0{,}776\} &= 0{,}05 \\ &= P_{\theta=0,898}\{\hat{\theta} < 0{,}776\} \quad . \end{aligned} \tag{8.4}$$

1. Eksaktni	$[0{,}631 \cdot \hat{\theta}, 1{,}88 \cdot \hat{\theta}]$	D/L = 2,38
2. Standardni (2.1)	$[0{,}466 \cdot \hat{\theta}, 1{,}53 \cdot \hat{\theta}]$	D/L = 1,00
3. BC (8.2)	$[0{,}580 \cdot \hat{\theta}, 1{,}69 \cdot \hat{\theta}]$	D/L = 1,64
4. BC_a (8.8)	$[0{,}630 \cdot \hat{\theta}, 1{,}88 \cdot \hat{\theta}]$	D/L = 2,37
5. Neparametrični BC_a	$[0{,}640 \cdot \hat{\theta}, 1{,}68 \cdot \hat{\theta}]$	D/L = 1,88

TABELA 5. 90% Centralni intervali zaupanja za θ, po ugotovitvi, da $\hat{\theta} \sim \theta(\chi^2_{19}/19)$. Eksaktni interval je zelo nagnjen proti desni $\hat{\theta}$. Metoda BC predstavlja le delno izboljšanje glede na standardni interval. BC_a za $a = 0{,}108$ pa se skoraj popolnoma ujema z eksaktnim.

Pripadajoče vrednosti za končne točke BC so

$$P_{\theta=0{,}488}\{\hat{\theta} > 0{,}776\} = 0{,}0465 \qquad (8.5)$$
$$P_{\theta=0{,}900}\{\hat{\theta} < 0{,}776\} = 0{,}0475 \quad ,$$

v primerjavi s končnimi točkami za centilno metodo

$$P_{\theta=0{,}536}\{\hat{\theta} > 0{,}776\} = 0{,}0725 \qquad (8.6)$$
$$P_{\theta=0{,}911}\{\hat{\theta} < 0{,}776\} = 0{,}0293 \quad .$$

Popravek pristranskosti je precej pomemben pri enačenju verjetnosti napake pri obeh končnih točkah. Če lahko dobimo natančen približek za z_0 (kot omenjeno v razdelku 9), bomo torej raje uporabili intervale BC.

V tabeli 5 najdemo enostaven zgled, pri katerem je metoda BC manj uspešna. Podatki sestojijo iz ene same realizacije vzorca $\hat{\theta} \sim \theta(\chi^2_{19}/19)$, torej naključne spremenljivke, pozrazdeljene po $\chi^2_{19}/19$, raztegnjene z neznanim raztegom θ. (Ta definicija naredi $\hat{\theta}$ nepristransko za θ.) Želimo interval zaupanja za razteg θ. V tem primeru je interval BC, osnovan na $\hat{\theta}$ definitivno velik napredek od standardnega intervala (2.1), a gre le približno pol toliko kot bi moral, da bi pravilno odražal asimetrijo eksaktnega intervala.

Izkaže se, da je parametrično družino $\hat{\theta} \sim \theta(\chi^2_{19}/19)$ nemogoče transformirati v (8.3), še več, tega se ne da narediti niti približno, z aproksimacijo. Rezultat v [12] kaže, da obstaja monotona transformacija g, tako da $\hat{\phi} = g(\hat{\theta})$, $\phi = g(\theta)$ do visoke stopnje aproksimacije zadoščata

$$\hat{\phi} \sim N(\phi - z_0\tau_\phi, \tau_\phi^2) \ \ (\tau_\phi = 1 + a\phi) \quad . \qquad (8.7)$$

Konstanti v (8.7) sta $z_0 = 0{,}1082$, $a = 0{,}1077$.

Metoda BC_a [14], 4.vrstica tabele 4, je metoda za izračun kljukčevih intervalov zaupanja, ki so točno pravilni za probleme, ki jih lahko spravimo v obliko (8.7). Ta metoda ima α-končno točko

$$\theta_{\mathrm{BC}_a}[\alpha] \equiv \hat{G}^{-1}\left(\Phi\left\{z_0 + \frac{z_0 + z^{(\alpha)}}{1 - a(z_0 + z^{(\alpha)})}\right\}\right) \quad . \qquad (8.8)$$

Če je $a = 0$, potem je $\theta_{\mathrm{BC}_a}[\alpha] = \theta_{\mathrm{BC}}[\alpha]$, v ostalih primerih pa BC_a intervali predstavljajo znatno izboljšavo metode BC, kot je razvidno iz tabele 5.

Konstanta z_0 v (8.8) je dana z $z_0 = \Phi^{-1}\{\hat{G}(\hat{\theta})\}$, (8.1), zato jo lahko izračunamo direktno iz kljukčeve porazdelitve. Kako pa izvemo a? Izkaže se, da za enoparametrične družine $f_\theta(\hat{\theta})$ dobro aproksimacijo dobimo z

$$a \doteq \frac{\mathrm{SKEW}_{\theta=\hat{\theta}}(i_\theta(t))}{6} \tag{8.9}$$

kjer je $\mathrm{SKEW}_{\theta=\hat{\theta}}(i_\theta(t))$ asimetričnost statistike rezultata

$$i_\theta(t) = \frac{\partial}{\partial\theta} \log f_\theta(t)$$

pri parametru $\theta = \hat{\theta}$. Za $\hat{\theta} \sim \theta(\theta(\chi^2_{19}/19))$ to da $a \doteq 0{,}1081$, v primerjavi z dejansko vrednostjo $a = 0{,}1077$, ki je izpeljana v [14]. Za korelacijo normalne družine iz tabele 3 je $a \doteq 0$, kar razloži, zakaj se metoda BC, za katero je $a = 0$, tam tako dobro obnaša.

Prednost formule (8.9) je v tem, da nam ni treba poznati transformacije g, ki nas pripelje do (8.7), da bi lahko aproksimirali a. V resnici je $\theta_{\mathrm{BC}_a}[\alpha]$ tudi transformacijsko invariantna (kot v (7.6)), kot že $\theta_{\mathrm{BC}}[\alpha]$ in $\theta_P[\alpha]$. Kot pri kljukčevih metodah tudi intervale BC_a lahko izračunamo direktno iz oblike porazdelitvene gostote $f_\sigma(\cdot)$, za θ blizu $\hat{\theta}$.

Formula (8.9) se nanaša na primer, kjer je θ edini parameter. V razdelku 9 na kratko obravnavamo težji problem postavljanja intervalov zaupanja za θ v večparametrični družini, in celo v neparametričnih situacijah, kjer je število nadležnih parametrov praktično neskončno.

[11], 10.poglavje, obravnava še nekaj ostalih načinov uporabe kljukčeve metode za konstrukcijo približnih intervalov zaupanja, ki tu niso predstavljeni. Ena od teh metod je recimo "kljukčeva t", ki pa ni tema mojega seminarja.

9. Velikosti kljukčevih vzorcev

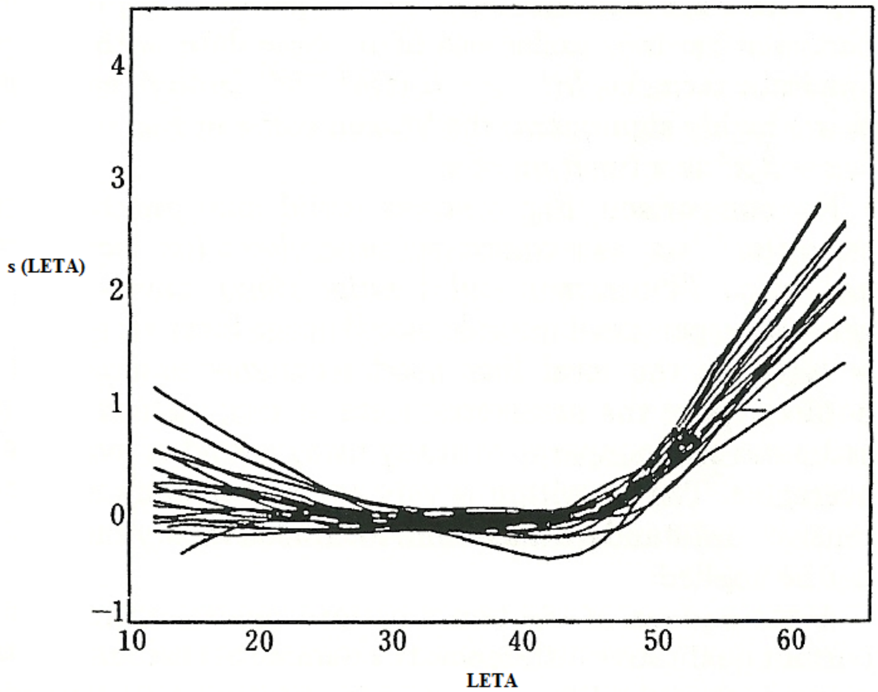
Koliko Monte Carlo replikacij moramo izvesti pri kljukčevi metodi (torej, kolikšen naj bo B)? Oglejmo si cenilnko standardne napake $\hat{\sigma}_B$, osnovano na B kljukčevih vzorcih (3.1). Ko gre $B \to \infty$, se $\hat{\sigma}_B$ približuje $\hat{\sigma}$, kljukčevi cenilki standardne napake, kot je bila na začetku definirana v (2.4). Ker $\hat{F}$ ni popolna cenilka za F, bo $\hat{\sigma} = \sigma(\hat{F})$ imela neničeln koeficient variacije (CV) za oceno resnične standardne napake $\sigma = \sigma(F)$; $\hat{\sigma}_B$ bo imela še večji koeficient variacije zaradi naključnosti, ki jo prinese kljukčevo Monte Carlo vzorčenje.

Enostavno je izpeljati naslednji približek:

$$\mathrm{CV}(\hat{\sigma}_B) \doteq \left\{ \mathrm{CV}(\hat{\sigma})^2 + \frac{E[\hat{\delta} + 2}{4B} \right\}^{1/2} , \tag{9.1}$$

kjer je $\hat{\delta}$ kurtosis kljukčeve porazdelitve $\hat{\theta}^*$ pri danem podatkovnem vektorju y in $E[\hat{\theta}]$ njeno matematično upanje, povprečeno čez y. Za tipične situacije $\mathrm{CV}(\hat{\sigma})$ leži med 0,10 in 0,30. Recimo, če je $\hat{\theta} = \bar{x}$, $n = 20$, $x_i \sim_{\mathrm{nep}} N(0{,}1)$, tedaj je $\mathrm{CV}(\hat{\sigma}) \doteq 0{,}16$.

Tabela 6 kaže $\mathrm{CV}(\hat{\theta}_B)$ za različne vrednosti B in $\mathrm{CV}(\hat{\theta})$, če privzamemo, da je v (9.1) $E[\hat{\delta}] = 0$. Za vrednosti $\mathrm{CV}(\theta) > 0{,}10$ *je za B nad 100 izboljšanje le majhno.* Precej dobre rezultate dobimo že za $B = 25$. Tudi še manjše vrednosti B so lahko že precej informativne, kar je lepo vidno na sliki 3, v Stanfordovih podatkih presaditev srca. Situacija je precej različna za postavljanje kljukčevih intervalov zaupanja. Računi v [14], razdelek 8, kažejo, da je $B = 1000$ nekakšen grob minimum za število



Slika 3. 20 kljukčevih replikacij lokalne ocene za Stanfordove podatke presaditev srca

Monte Carlo replikacij, potrebnih za izračun intervala BC ali BC_a. Nekoliko manjše vrednosti, recimo okoli $B = 250$, so dovoljšnje za uporabne centilne intervale, saj nam ni treba računati konstante z_0. Intervali zaupanja so v celoti bolj zahtevna mera statistične natančnosti kot standardne napake, zato potreba po večji računski moči ni presenetljiva.

		$B \rightarrow$				
		25	50	100	200	∞
CV($\hat{\sigma}$)	0,25	0,29	0,27	0,26	0,25	0,25
$\downarrow$	0,20	0,24	0,22	0,21	0,21	0,20
	0,15	0,21	0,18	0,17	0,16	0,15
	0,10	0,17	0,14	0,12	0,11	0,10
	0,05	0,15	0,11	0,09	0,07	0,05
	0	0,14	0,10	0,07	0,05	0

Tabela 6. Koeficienti variacije $\hat{\sigma}_S$, kljukčeve ocene standardne napake na B Monte Carlo replikacijah, kot funkcija B in CV($\hat{\sigma}$), tudi v limitah obojih. Osnovano na (9.1), ob predpostavki $E[\hat{\delta}] = 0$.

Del 4. Zaključek

Kljukčeva metoda je splošna metodologija za odgovarjanje na vprašanje kako natančna je cenilka $\hat{\theta}$ za oceno neznanega parametra θ. Poudariti velja, da je temeljna uporaba kljukčeve metode in posledično ocena neznanih parametrov odvisna od računalniške zmogljivosti, natančnost le te pa se s tehnološkim napredkom povečuje. Sklepamo lahko, da bo v statistični metodologiji s časom prišlo do korenitih sprememb, saj bo teoretično analizo statistika kmalu izpodrinila in zamenjala surova računalniška moč in programi numeričnih algoritmov.

Primer omenjenega algoritma, je metoda Monte Carlo, s katero se srečamo v razdelku 3, ki opisuje kljukčevo oceno standardne napake. Metoda je pri izračunu napake ključnega pomena. Z Metodo Monte Carlo imenujemo vse metode, ki pri reševanju problemov uporabljajo slučajnost. Vredno je poudariti, da izračunane vrednosti niso točne, so le približki količin, ki nas zanimajo. V našem primeru torej približek ocene standardne napake.

Kar zadeva kljukčeve intervale zaupanja, je napredovanje od osnovnih standardnih intervalov do metode BC_a osnovano predvsem na zaporedju vedno manj restriktivnih predpostavk, kot je prikazano v tabeli 4. Vsaka naslednja metoda v tabeli 4 zahteva več računanja. Najprej kljukčeva porazdelitev $\hat{G}$, nato še konstanta za popravek pristranskosti z_0, nato pa še a. K sreči so vsi ti izračuni po naravi algoritmični in kot taki dovzetni za avtomatsko obravnavo.

Literatura

[1] BAHADUR, R. and SAVAGE, L. (1956). *The nonexistence of certain statistical procedures in nonparametric problems.* Ann. Math. Statist. **27** 1115 - 1122.
[2] BERAN, R. (1984). *Bootstrap methods in statistics.* Jahrb. Math. Ver. **86** 14-30.
[3] BICKEL, P. J., FREEDMAN, D.A. (1981). *Some asymptotic theory for the bootstrap.* Ann. Statist. **9** 1196-1217
[4] BOX, G. E. P., COX, D. R. (1964). *An analysis of transformations.* J. R. Statist. Soc. Ser. B **26** 211 - 252.
[5] BREIMAN, L., FRIEDMAN, J. H. (1985). *Estimating optimal transformations for multiple regression and correlation.* J. Amer. Statist. Assoc. **80** 580 - 619.
[6] COX, D. R. (1972). *Regression models and life tables.* J. R. Statist. Soc. Ser. B **34** 197 - 202.
[7] CRAMER, H. (1946). *Mathematical Methods of Statistics.* Princeton University Press, Princeton, New Jersey.
[8] EFRON, B. (1979b). *Computers and the theory of statistics: thinking the unthinkable.* Soc. Ind. Appl. Math. **21** 460 - 480.
[9] EFRON, B. (1981a). *Censored data and the bootstrap.* J. Amer. Statist. Assoc. **76** 312 - 319.
[10] EFRON, B. (1981b). *Nonparametric estimates of standard error: the jackknife, bootstrap, and other resampling methods.* Biometrika **68** 589 - 599.
[11] EFRON, B. (1982a). *The Jackknife, the bootstrap, and other resampling plans.* Soc. Ind. Appl. Math. CBMS-Natl. Sci. Found. Monogr. **38**.
[12] EFRON, B. (1982b). *Transformation theory: how normal is a one parameter family of distributions?* Ann. Statist. **10** 323 - 339.
[13] EFRON, B. (1982c). *Maximum likelihood and decision theory.* Ann. Statist. **10** 340 - 356.
[14] EFRON, B. (1984). *Better bootstrap confidence intervals.* Tech. Rep. Stanford Univ. Dept. Statist.
[15] EFRON, B. (1985). *Bootstrap confidence intervals for a class of parametric problems.* Biometrika **72** 45 - 58.
[16] EFRON, B., GONG, G. (1983). *A leisurely look at the bootstrap, the jackknife, and cross-validation.* Amer. Statistican **37** 36 - 48.
[17] FIELLER, E. C. (1954). *Some problems in interval estimation.* J. R. Statist. Soc. Ser. B **16** 175 - 183
[18] KEIFER, J., WOLFOWITZ, J. (1956). *Consistency of the Maximum Likelihood Estimator in the Presence of Infinitely Many Incidental Parameters.* Ann. Math. Statist. Volume 27, Number 4 (1956), 887-906.