Statističke metode za usporedbu profila oslobađanja lijeka

Čolić, Ivan

Master's thesis / Diplomski rad

2018

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Science / Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:217:917301

Rights / Prava: In copyright / Zaštićeno autorskim pravom.

Download date / Datum preuzimanja: **2024-04-24**

Repository / Repozitorij:

Repository of the Faculty of Science - University of Zagreb

SVEUČILIŠTE U ZAGREBU

PRIRODOSLOVNO–MATEMATIČKI FAKULTET

MATEMATIČKI ODSJEK

Ivan Čolić

STATISTIČKE METODE ZA USPOREDBU PROFILA OSLOBAĐANJA LIJEKA

Diplomski rad

Voditelji rada:
doc.dr.sc. Pavle Goldstein i
dr.sc. Drago Špoljarić

Zagreb, veljača 2018.

Ovaj diplomski rad obranjen je dana _____ pred ispitnim povjerenstvom u sastavu:

1. _____, predsjednik
2. _____, član
3. _____, član

Povjerenstvo je rad ocijenilo ocjenom _____.

Potpisi članova povjerenstva:

1. _____
2. _____
3. _____

Zahvaljujem se mentorima doc.dr.sc Pavli Goldsteinu i dr.sc. Dragi Špoljariću na strpljenju, susretljivosti, razumijevanju te pomoći prilikom pisanja diplomskog rada.
Zahvaljujem se braći što su mi svojim primjerom pokazali da se trud isplati.
Zahvaljujem se roditeljima koji su mi bili neizmjerna podrška i bez kojih sve ovo što sam dosad postigao ne bi bilo moguće.
Posebno HVALA sestri na njenoj neiscrpnoj pozitivnoj energiji.
Ovaj diplomski posvećujem svim kolegama, prijateljicama i prijateljima koji su mi tijekom studiranja pomogli.

Sadržaj

Sadržaj **iv**

Uvod **1**

1 Metoda faktora **2**
1.1 Općenito 2
1.2 Faktori f_1 i f_2 2
1.3 Interval pouzdanosti procjenitelja od f_2 5
1.4 Primjena 7

2 Neparametarska metoda **11**
2.1 Mahalanobisova udaljenost 11
2.2 Regija pouzdanosti 14
2.3 Metoda Lagrangeovih multiplikatora 14
2.4 Interval pouzdanosti 14
2.5 Primjena 15

3 Parametarska metoda **19**
3.1 Koraci metode 19
3.2 Funkcije kandidati 20
3.3 Regija sličnosti određena s $Pooled_{SD}$ 21
3.4 Mahalanobisova udaljenost 21
3.5 Regija pouzdanosti 22
3.6 Primjena 23

Bibliografija **30**

Uvod

Proces istraživanja i razvoja novog lijeka je dugotrajan i skup, a vjerojatnost otkrivanja novih spojeva je mala. U pravilu novootkriveni lijekovi imaju patentnu zaštitu koja traje 20 godina. Nakon isteka patentne zaštite lijeka, aktivnu tvar osim tvrtke koja je nositelj patenta mogu proizvoditi i prodavati druge tvrtke. Iz tog razloga većina farmaceutskih firmi je usmjerila svoje poslovanje u razvoj i proizvodnju generičkih lijekova. Generički lijek identičan je originalnom u ljekovitoj tvari, dozi, učinkovitosti i nuspojavama. Ukratko, ima isto djelovanje kao original, ali njihov je razvoj brži i jeftiniji (vidi [4]).

Velika prednost generičkog lijeka je što on prolazi kroz skraćeno testiranje (ANDA-abbreviated new drug application) prije nego dobije dopuštenje od FDA (agencija za hranu i lijekove u SAD-u) za izlazak na tržište. Pri dokazivanju sigurnosti i učinkovitosti generičkog lijeka ne moramo kao za originalni lijek raditi testiranja na životinjama i ljudima. Dovoljno je pokazati da se ponaša isto kao originalni lijek. Generički lijek dobiti će FDA odobrenje za izlazak na tržište nakon što se pokaže da će se u istom vremenskom periodu u krvi pacijenta nalaziti jednaka količina aktivne supstance kao i za originalni lijek.

Tijekom razvoja i optimizacije testnih (generičkih) lijekova farmaceutski formulator (znanstvenik) koristi različite metode za usporedbu testnog i referentnog (originalnog) lijeka. U radu su opisane standardne statističke metode (metoda faktora, neparametarska metoda i parametarska metoda) za usporedbu profila oslobađanja testnog i referentnog lijeka (čvrstog oblika) u nekom relevantnom mediju (tekućini, otopini). U sljedećem poglavlju objašnjen je pojam profila oslobađanja.

Smatra se da su profili testnog i referentnog lijeka slični ukoliko je razlika između profila testnog i referentnog lijeka manja od razlike između profila dva referentna lijeka. Metode obrađene u ovom radu daju jasne rezultate preko kojih donosimo zaključke o sličnosti dvaju profila. U prvom poglavlju upoznajemo se sa faktorom sličnosti, njegovim teorijskim pretpostavkama, procjeniteljima i načinom tumačenja rješenja. U drugom poglavlju upoznajemo se sa neparametarskom metodom koja koristeći Mahalanobisovu udaljenost preko intervala pouzdanosti dolazi do zaključka o sličnosti dvaju profila. U trećem poglavlju obrađujemo parametarsku metodu koja odabranu funkciju prilagođava profilima oslobađanja i na temelju parametara odabrane funkcije konstruira regiju sličnosti i regiju pouzdanosti. Temeljem usporedbe regija na istom grafu donosimo zaključak o sličnosti.

Poglavlje 1

Metoda faktora

1.1 Općenito

Eksperiment (testiranje) uglavnom sadrži 6 - 12 tableta testnog i referentnog lijeka koje su ubačene u posude sa medijem. Aparatura koja se koristi kako bi se došlo do profila otapanja u većini slučajeva je USP Aparatura 2. Mijenjajući parametre aparature možemo simulirati različite uvjete. Koristeći lopatice koje se okreću određenom brzinom, oslobođeni sadržaj se raspršuje po cijelom mediju iz kojeg se uzimaju uzorci u definiranim vremenskim točkama. Pritom, profil oslobađanja tablete je vektor $x = (x_1, \ldots, x_P)$, gdje x_P predstavlja količinu oslobođenog sadržaja (u vremenskoj točki P) aktivne susptance (izraženo u postotku od propisanog sadržaja).

1.2 Faktori f_1 i f_2

Usporedbu profila oslobađanja preko faktora f_1 i f_2 predložili su 1996. godine Moore i Flanner ([3]). Predložene funkcije zahtijevaju podatke o oslobađanju prikupljene u istim uvjetima. Pod istim uvjetima podrazumjevamo:

- isti medij (tekućina, otopina)
- ista temperatura medija (37° C ± 0.5° C)
- ista pH vrijednost medija
- ista aparatura
- iste postavke aparature (npr. za USP II aparaturu jedan od parametara je brzina okretanja lopatica)

- ista vremena uzorkovanja.

Ljekovima koji spadaju u kategoriju brzo oslobađajućih, uzorkovanja se obično vrše u 15., 35., 45. i 60. minuti, dok ljekovima s kontroliranim oslobađanjem obično u prvom, trećem, petom i osmom satu (vidi [3]).

Neka je P broj vremenskih točaka u kojima se radi uzorkovanje i neka $(\mu_{r1}, \ldots, \mu_{rP})$ predstavlja količinu oslobođenog sadržaja (po točkama) aktivne susptance (izraženo u postotku od propisanog sadržaja (eng. *label claimed*)) referentnog lijeka i neka je $(\mu_{t1}, \ldots, \mu_{tP})$ isto to za testni lijek. Neka je $(|\mu_{r1} - \mu_{t1}|, \ldots, |\mu_{rP} - \mu_{tP}|)$ razlika u količini oslobođenog sadržaja između profila u P vremenskih točaka. Neka su D_1 i D_2 definirane kao:

$$D_1 := |\mu_{r1} - \mu_{t1}| + |\mu_{r2} - \mu_{t2}| + \cdots + |\mu_{rP} - \mu_{tP}| \tag{1.1}$$

$$D_2 := \sqrt{(\mu_{r1} - \mu_{t1})^2 + (\mu_{r2} - \mu_{t2})^2 + \cdots + (\mu_{rp} - \mu_{tp})^2} \tag{1.2}$$

Faktori f_1 i f_2 definirani su kao funkcije od D_1 i D_2 na sljedeći način ([3]):

$$f_1 = \left\{ \frac{\sum_{i=1}^{P} |\mu_{ti} - \mu_{ri}|}{\sum_{i=1}^{P} \mu_{ri}} \right\} \cdot 100 = \left\{ \frac{D_1}{\sum_{i=1}^{P} \mu_{ri}} \right\} \cdot 100 \tag{1.3}$$

$$f_2 = 50 \cdot \log \left\{ \left[1 + \frac{1}{P} \sum_{i=1}^{P} (\mu_{ti} - \mu_{ri})^2 \right]^{-\frac{1}{2}} \cdot 100 \right\} = 50 \cdot \log \left\{ \left[1 + \frac{1}{P} D_2^2 \right]^{-\frac{1}{2}} \cdot 100 \right\} \tag{1.4}$$

Formula (1.3) kao funkcija prosječne apsolutne razlike između dvaju profila oslobađanja se naziva **faktor razlike** (eng. *difference factor*). Poprima vrijednost 0 kada su testni i referentni profili identični, a raste proporcionalno različitosti profila ([9]). U nastavku ovog poglavlja bavimo se detaljnije f_2 faktorom koji se zbog njegove jednostavnosti često koristi pri dokazivanju sličnosti in vitro profila oslobađanja.

Faktor f_2 se zove **faktor sličnosti** (eng. *similarity factor*). Predstavlja mjeru sličnosti između dvaju profila oslobađanja. Pretpostavimo da su dva profila identična, tada je razlika 0 i f_2 poprima vrijednost $f_2 = 50 \cdot \log(100) = 100$. U slučaju kada je razlika u svakoj točki uzorkovanja jednaka 100 faktor sličnosti poprima vrijednost -0.001 koja se zaokruži na 0. Formula (1.4) je iskoristiva samo za usporedbu onih profila čija prosječna razlika manja od 100. Dakle, f_2 poprima vrijednosti između 0 i 100 gdje veća vrijednost od f2 označava veću sličnost profila.

Primjer

Pretpostavimo da je prosječna razlika ≥ 100, tada faktor sličnosti poprima vrijednost manju od nule. Prosječna razlika između testnog i referentnog lijeka može biti veća od 100 jer testni ili referentni lijek ne moraju imati u sebi jednaku propisanu količinu sastojaka. Stvar je u tome da se u procesu proizvodnje događa određeno odstupanje (uglavnom je to $\pm 10\%$ od propisane količine sadržaja - npr. ako na kutiji piše 100 mg ibuprofena, onda tablete sadrže $90 - 110\%$ od 100 mg). Recimo, referentna tableta ima 110 mg, a testna 98 mg i pritom za vrijeme uzorkovanja proces otapanja referentne je završio prije nego je proces otapanja testne uopće počeo. U praksi nema smisla uspoređivati takve profile.

Postavlja se pitanje koliko treba iznositi f_2 da bi se dva profila smatrala sličnima. U realnosti nije za očekivati da će vrijednost od f_2 biti blizu 100 čak i za tablete iz iste kutije. Stoga se smatra da je testni lijek sličan referentnom ukoliko razlika između profila testnog i referentnog lijeka nije veća od razlike između profila dva referentna lijeka. U praksi se pokazuje da je prosječna razlika ne veća od 10% prihvatljiva. Kada se prihvatljivih 10% razlike uvrsti u jednadžbu (1.4) dobije se sljedeće:

$$f_2 = 50 \cdot \log\left\{\left[1 + \frac{1}{P}\sum_{i=1}^{P}(10)^2\right]^{-\frac{1}{2}} \cdot 100\right\} = 50 \cdot \log\left\{\left[101\right]^{-\frac{1}{2}} \cdot 100\right\} = 50\log(9.95037) = 49.89.$$

Radi jednostavnosti rezultat se zaokružuje na 50.

Dakle, da bi se dva profila (testni i referentni) smatrala sličnima vrijednost faktora sličnosti mora biti najmanje 50.

Procjenitelj od f_2

U praksi za usporedbu profila koristimo dvije skupine po dvanaest tableta. Jedna skupina predstavlja tablete referentnog, a druga testnog lijeka. Svakoj od 12 tableta referentnog i testnog lijeka uzimamo uzorke kroz P vremenskih točaka kako bi se odredila količina oslobođenog sadržaja aktivne supstance. Usporedbu profila radimo na prosječnim profilima otapanja. Pri tome prosječni profil otapanja lijeka definiramo kao srednje vrijednosti količine oslobođenog sadržaja po vremenima. Stoga su $(\overline{x}_{r1}, \ldots, \overline{x}_{rP})$ i $(\overline{x}_{t1}, \ldots, \overline{x}_{tP})$ procjenitelji teorijskih vrijednosti $(\mu_{r1}, \ldots, \mu_{rP})$ odnosno $(\mu_{t1}, \ldots, \mu_{tP})$, gdje su $\overline{x}_{ti}$ i $\overline{x}_{ri}$ srednje vrijednosti u i-tom trenutku za 12 tableta testnog odnosno referentnog lijeka. Prema članku 2 procjenitelj od f_2 je

$$\hat{f}_2 = 50 \cdot \log\left\{\left[1 + \frac{1}{P}\sum_{i=1}^{P}(\overline{x}_{ti} - \overline{x}_{ri})^2\right]^{-\frac{1}{2}} \cdot 100\right\} \tag{1.5}$$

Nepristrani procjenitelj od f_2

Iz [3] znamo da pristranost $\hat{f}_2$ procjenitelja uzrokuje izraz $\sum_{i=1}^{P}(\sigma_{ti}^2 + \sigma_{ri}^2)/n$. Dakle, korekcija pristranog procjenitelja $\hat{f}_2$ bi bila oduzimanje nepristranog procjenitelja od $\sum_{i=1}^{P}(\sigma_{ti}^2 + \sigma_{ri}^2)/n$. Na taj način dolazimo do nepristranog procjenitelja za f_2:

$$\hat{f}_2^{\,*} = 50 \cdot \log\left(\left\{1 + \frac{1}{P}\left[\sum_{i=1}^{P}(\overline{x}_{ti} - \overline{x}_{ri})^2 - \sum_{i=1}^{P}\frac{s_{ti}^2 + s_{ri}^2}{n}\right]\right\}^{-\frac{1}{2}} \cdot 100\right) \tag{1.6}$$

gdje je n broj tableta (12), a $\overline{x}_{ti}$ i $\overline{x}_{ri}$ srednje vrijednosti u i-toj vremenskoj točki za 12 tableta testnog odnosno referentnog lijeka. Varijance uzoraka s_{ti}^2 i s_{ri}^2 su nepristrani procjenitelji varijanci populacije u i-toj vremenskoj točki za testni i referentni lijek. Uočimo da ako vrijedi

$$\sum_{i=1}^{P}(\overline{x}_{ti} - \overline{x}_{ri})^2 < \sum_{i=1}^{P}\frac{s_{ti}^2 + s_{ri}^2}{n}.$$

onda formula (1.6) nema smisla jer je varijabilnost po točkama velika i iz tog razloga treba posumnjati u same uzorke tableta koje imamo. Preciznije, testiranje nema smisla jer nemamo kvalitetne podatke.

1.3 Interval pouzdanosti procjenitelja od f_2

Bootstraping - osnovna ideja

Neka je $(x_1, \ldots, x_n)$ realizacija slučajnog uzorka $(X_1, \ldots, X_n)$ iz nepoznate distribucije F. Osnovna ideja je kako realizacije sadrže sve dostupne informacije o distribuciji F i stoga očekujemo da će ponovno uzorkovanje iz tog uzorka dati uzorak koji odgovara uzorkovanju iz distribucije F. Recimo da promatramo parametar θ iz distribucije F i njegov procjenitelj $\hat{\theta} = S(X_1, \ldots, X_n)$. Često nas zanima distribucija tog procjenitelja zbog pouzdanih intervala. Kada bi F bila poznata mogli bismo odrediti uzorkovnu distribuciju od $\hat{\theta}$ opetovanim uzorkovanjem iz F. Budući da je F nepoznata, uzorkujemo iz empirijske distribucije $\hat{F}$ dobivene na temelju realizacije $(x_1, \ldots, x_n)$ ([10]).

Bootstrap uzorkovanje u našem slučaju se radi posebno na 12 profila referentnog i posebno na 12 profila testnog lijeka. Ideja je napraviti M bootstrap uzoraka (uzorkovanje s vraćanjem) za testni i referentni lijek. Uspoređivali bi m-ti bootstrap uzorak referentnog lijeka s m-tim bootstrap uzorkom testnog lijeka. Važno je napomenuti da neki od profila u uzorku mogu biti isti ili nemogući, ali to ovisi o načinima generiranja uzoraka koji su detaljnije objašnjeni kasnije.

Jackknife procjena parametara - osnovna ideja

Jackknife procjena parametara se bazira na sistematičnom izostavljanju jednog podatka iz skupa podataka, računanju procjenitelja bez tog podatka i računanju prosjeka za te procjenitelje.([1]) U našem primjeru napravit ćemo n uzoraka za testni i referentni lijek gdje će svaki imati n-1 profil otapanja (n je broj tableta). Prvi uzorci ne sadrže podatke otapanja (profil) prve tablete, isto tako m-ti uzorak ne sadrži podatke otapanja m-te tablete itd.. Tada se računa procjenitelj od f_2 kako bi usporedili m-ti uzork testnog i m-ti uzorak referentnog lijeka, a nakon toga računamo prosjek za te procjenitelje.

Interval pouzdanosti

Procjenitelj od f_2 za bootstrap uzorak označavamo s $\hat{f}_2{}'$, a $\hat{f}_2{}'(m)$ je rezultat usporedbe m-tog referentnog bootstrap uzorka s m-tim testnim. Devedeset postotni interval pouzdanosti od $\hat{f}_2{}'$ je definiran sa $[L\hat{f}_2{}', U\hat{f}_2{}']$, gdje $L\hat{f}_2{}'$ i $U\hat{f}_2{}'$ predstavljaju 5-ti i 95-ti percentil od $E\hat{f}_2{}'$. Budući da je distribucija od $\hat{f}_2{}'$ pristrana potrebno je napraviti određene prilagodbe za određivanje intervala pouzdanosti. Prilagođeni interval pouzdanosti $[\hat{f}_{2(M\alpha 1)}{}', \hat{f}_{2(M\alpha 2)}{}']$ je definiran sa:

$$\alpha_1 = \phi\left(Z_0 + \frac{Z_0 + Z^{\alpha}}{1 - a(Z_0 + Z^{\alpha})}\right) \tag{1.7}$$

$$\alpha_2 = \phi\left(Z_0 + \frac{Z_0 + Z^{1-\alpha}}{1 - a(Z_0 + Z^{1-\alpha})}\right) \tag{1.8}$$

$$Z_0 = \phi^{-1}\left(\frac{\#(\hat{f}_2{}'(m) < \hat{f}_2)}{M}\right) \tag{1.9}$$

$$a = \frac{\sum(\hat{f}_{2(\cdot)} - \hat{f}_{2(i)})^3}{6(\sum(\hat{f}_{2(\cdot)} - \hat{f}_{2(i)})^2)^{\frac{3}{2}}} \tag{1.10}$$

gdje je α vrijednost greške tipa 1, $\hat{f}_{2(i)}$ i-ti jackknife procjenitelj, $\hat{f}_{2(\cdot)}$ srednja vrijednost svih jackknife procjenitelja, $\hat{f}_2{}'(m)$ bootstrap procjenitelj m-tog bootstrap uzorka, $\hat{f}_2$ je vrijednost procjenitelja originalnog uzorka i Z^{α} je α percentil standardne normalne distribucije. U intervalu $[\hat{f}_{2(M\alpha_1)}{}', \hat{f}_{2(M\alpha_2)}{}']$, $\hat{f}_{2(M\alpha_1)}{}'$ predstavlja vrijednost $\hat{f}_2{}'$ bootstrap procjenitelja koji se u sortiranom nizu svih bootstrap procjenitelja nalazi na mjestu s indeksom $\lfloor M\alpha_1 \rfloor$, ekvivalentno vrijedi za $\hat{f}_{2(M\alpha_2)}{}'$.

1.4 Primjena

U sljedećem primjeru uspoređujemo profil oslobađanja testne i referentne tablete. Bootstrap metodom odrediti ćemo interval pouzdanosti na temelju kojeg možemo zaključiti o sličnost profila oslobađanja referentne i testne tablete. Također, usporedit ćemo rezultate nepristranog i pristranog procjenitelja za f_2.

Primjer

U tablicama Referentni Lijek i Testni Lijek (iz [3]) nalaze se 12 profila oslobađanja referentnih i 12 profila testnih tableta. Vrijednosti u tablicama predstavljaju količinu oslobođenog sadržaja aktivne susptance (izraženo u postotku) u 30., 60., 90. i 180. minuti.

Referentni Lijek				
Tableta	30 minuta	60 minuta	90 minuta	180 minuta
1	36.1	58.6	80	93.3
2	33	59.5	80.8	95.7
3	35.7	62.3	83	97.1
4	32.1	62.3	81.3	92.8
5	36.1	53.6	72.6	88.8
6	34.1	63.2	83	97.4
7	32.4	61.3	80	96.8
8	39.6	61.8	80.4	98.6
9	34.5	58	76.9	93.3
10	38	59.2	79.3	94
11	32.2	56.2	77.2	96.3
12	35.2	58	76.7	96.8
prosjek	34.92	59.5	79.27	95.08

Testni Lijek				
Tableta	30 minuta	60 minuta	90 minuta	180 minuta
1	38.75	61.79	85.14	100.2
2	36.16	61.21	84.25	97.3
3	38.49	63.89	84.94	96.39
4	37.27	62.52	85.65	95.47
5	48.12	77.18	95.32	99.3
6	48.45	80.62	95.05	98.94
7	41.08	67.62	84.94	99.03
8	39.64	63.68	80.73	95.63
9	36.06	61.59	82.22	96.12
10	36.69	63.6	84.5	98.42
11	39.95	67.98	87.4	98.1
12	43.41	74.07	93.95	97.8
prosjek	40.34	67.15	87.01	97.73

Bitno je uočiti da se srednje vrijednosti testnog i referentnog lijeka ne razlikuju više od 8%. Na temelju ranije navedene činjenice da je razlika do 10% prihvatljiva, možemo naslutiti da će zaključak našega testiranja ići u korist sličnosti. Važno je napomenuti da zbog varijabilnosti uzoraka, sličnost profila ne zaključuje se direktnim uspoređivanjem $\hat{f}_2$ sa razinom sličnosti (eng. *similarity limit*, kasnije u tekstu SL). Za $\hat{f}_2$ SL je 50 (spominjali ranije). Ovom metodom sličnost profila zaključuje se preko intervala pouzdanosti od $E\hat{f}_2{}'$, preciznije donju granicu tog intervala uspoređujemo sa SL.

Bootstrap generiranje uzoraka napravili smo na dva načina:

- Prvi način je nasumično biranje profila s vraćanjem. Tim načinom dobivamo bootstrap uzorke koji sadrže 12 profila nasumično izabranih iz tablice Referentni Lijek i 12 profila nasumično izabranih iz tablice Testni Lijek.

- Drugi način je nasumično biranje (s vraćanjem) vrijednosti po točkama, gdje je za svaku vremensku točku P nasumično izabrano odgovarajućih 12 vrijednosti iz tablice Referentni Lijek, isto tako izabrano je 12 vrijednosti iz tablice Testni Lijek za svaku vremensku točku P.

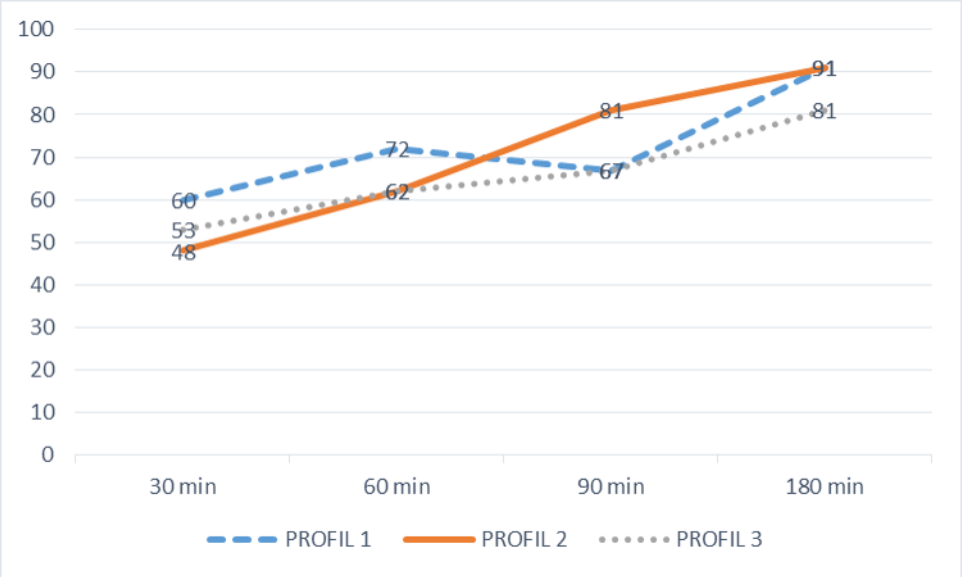
Želimo izračunati $\hat{f}_2{}'$ za testne i referentne profile svakog bootstrap uzorka. Računajući jackknife procjenitelje dobije se srednja vrijednost jackknife procjenitelja $\hat{f}_{2(\cdot)} = 60.07967$.

Jackknife procjenitelji	
Bez Tablete	$\hat{f}_{2(i)}$
1	59.30443
2	58.89351
3	58.64955
4	58.99472
5	64.00611
6	61.99628
7	59.83042
8	58.28485
9	59.17057
10	59.08660
11	60.79834
12	61.94066
$\hat{f}_{2(\cdot)}$	60.07967

Koristeći jednadžbu (1.10) dobijemo da je $a = -0.05226787$. Vrijednost pristranog procjenitelja $\hat{f}_2$ je 60.03 dok nepristrani procjenitelj $\hat{f}_2^*$ poprima vrijednost 60.73. Koristeći jednadžbe (1.7),(1.8)i (1.9) i sortirajući $\hat{f}_2'(m)$ vrijednosti dolazimo do intervala pouzdanosti. U sljedećoj tablici prikazani su 90% intervali pozdanosti dobiveni za različit broj bootstrap uzoraka, podijeljeni po načinima generiranja.

Intervali pouzdanosti				
Način generiranja	200 boot.uz.	400 boot.uz.	600 boot.uz.	1000 boot.uz.
1. način	(51.93, 67.60)	(52.59, 67.14)	(52.15, 67.36)	(52.33, 67.37)
2. način	(55.61, 64.59)	(55.05, 64.58)	(55.65, 65.72)	(55.70, 64.90)

Prema [8] 400 bootstrap uzoraka dovoljno je za precizan interval pouzdanosti, ali kako bi bili što precizniji i sigurniji u naš interval nakon 400 bootstrap uzoraka napravili smo još 600 i 1000 uzoraka. Gledajući donje granice intervala primjećujemo da su sve veće od SL. Temeljem toga zaključujemo da su profili oslobađanja referentnog i testnog lijeka slični. Primijetimo kako se drugim načinom dobila veća donja granica intervala pouzdanosti, odnosno drugim načinom dobijemo uži interval pouzdanosti. Naravno, treba imati na umu kako se generiranjem na drugi način stvaraju nemogući profili, odnosno krivulje profila nisu stalno neopadajuće.



Slika 1.1: Prikaz nemogućeg profila

Važno je spomenuti da za usporedbu profila oslobađanja preko faktora sličnosti treba obratiti pozornost na vremena uzorkovanja. Nakon što jedna od tableta postigne 85% oslobađanja određuje se još samo jedna vremenska točka uzorkovanja. Limitiranje broja točaka nakon 85% oslobađanja važno je radi donošenja kvalitetnijeg zaključka o sličnosti. Problem nastaje kada se nakon 85% stavi više točaka uzorkovanja jer su tada krivulje profila vrlo slične, a vrijednost faktora sličnosti raste te nas može navesti na krivi zaključak.

Faktor sličnosti f_2 predstavlja jednostavnu mjeru sličnosti dvaju profila oslobađanja koju u praksi procjenjujemo sa $\hat{f}_2$. Zbog jednostavnosti računanja i tumačenja rezultata iznimno je popularna metoda koju koriste proizvođači lijekova.

Poglavlje 2

Neparametarska metoda

Neparametarska metoda opisana u ovom poglavlju koristi se za testiranje globalne sličnosti dvaju profila oslobađanja. Ona kao i svaka druga metoda ima određene uvjete koji trebaju biti zadovoljeni prije njenog korištenja.

Jedan od uvjeta je dobro definirani SL (razina sličnosti) kojeg određujemo na temelju karakteristika samog lijeka ili na temelju iskustva. SL može biti definiran kao globalni SL ili lokalni SL. Globalni SL D_g je definiran kao prihvatljiva razlika između testnog i referentnog lijeka po svim vremenskim točkama (uglavnom 10%). Lokalni SL D_t je definiran kao prihvatljiva razlika između između testnog i referentnog lijeka za svaku vremensku točku posebno. Recimo u trenutku t prihvatljiva razlika je 12% dok u trenutku $t + 1$ može biti 7%.

Podaci oslobađanja se trebaju prikupljati u istim uvjetima (isto kao u prvoj metodi) na istom broju tableta (uobičajeno je to 12 tableta za testni i 12 za referentni lijek). Prema [7] tablete referentnog lijeka korištene u testiranju moraju pripadat grupi najnovije proizvedenih tableta.

Uzorkovanja se trebaju napraviti u vremenskim točkama koje adekvatno reprezentatiraju profil oslobađanja lijeka, a zaključak o sličnosti donosi se nakon što smo sigurni da raspolažemo s kvalitetnim predstavnicima tableta koje testiramo.

2.1 Mahalanobisova udaljenost

Statistička udaljenost često korištena za izračun razlike između dvaju multivarijantnih srednjih vrijednosti je Mahalanobisova udaljenost ([7]):

$$D_M = \sqrt{(x_t - x_r)' S^{-1}_{pooled} (x_t - x_r)}. \tag{2.1}$$

gdje je

- $x_r = (x_{r1}, \ldots, x_{rP})$ vektor srednjih vrijednosti referentnog lijeka, gdje x_{ri} predstavlja srednju vrijednost u i-toj vremenskoj točki za uobičajeno 12 tableta, a analogno vrijedi za $x_t = (x_{t1}, \ldots, x_{tP})$ vektor srednjih vrijednosti testnog lijeka

- $S_{pooled} = \frac{S_{ref}+S_{test}}{2}$ je kovarijacijska matrica uzoraka, združena za testni i referentni lijek, gdje su S_{ref} i S_{test} kovarijacijske matrice vremenskih točaka testnog i referentnog lijeka

Pokažimo da je Mahalanobisova udaljenost stvarno udaljenost (metrika). Definirajmo prvo metriku [2] .

Definicija 2.1.1. *Neka je $X \neq \emptyset$ neprazan skup i $d : X \times X \to \mathbb{R}$ preslikavanje sa Kartezijevog produkta $X \times X$ u skup realnih brojeva $\mathbb{R}$ za koje vrijedi:*

1. $d(a, b) \geq 0, \forall a, b \in X$ *(pozitivnost),*
2. $d(a, b) = 0 \Leftrightarrow a = b$ *(strogost),*
3. $d(a, b) = d(b, a), \forall a, b \in X$ *(simetričnost),*
4. $d(a, b) \leq d(a, c) + d(c, b) \forall a, b, c \in X$ *(nejednakost trokuta).*

Kažemo da je d funkcija udaljenosti ili razdaljinska funkcija, odnosno ***metrika*** *na skupu X.*

Da bi dokazali da je Mahalanobisova udaljenost metrika prvo obratimo pažnju da sljedeće stvari.

Korolar 2.1.2. *Kovarijacijska matrica je pozitivno semidefinitna.*

Znamo da su svojstvene vrijednosti (pozitivno semidefinitne matrice) ≥ 0. Također, kako je kovarijacijska matrica S_{pooled} simetrična postoji (po Spektralnom Teoremu) ortogonalna matrica U ($U' = U^{-1}$) i dijagonalna matrica $D = diag(d_1, \ldots, d_P)$ t.d. $S_{pooled} = U'DU$. Dijagonalna matrica D na dijagonali ima svojstvene vrijednosti matrice S_{pooled}. Primijetimo da se Mahalanobisova udaljenost računa sa S^{-1}_{pooled}. Ukoliko postoji inverz od S_{pooled} onda je on $S^{-1}_{pooled} = (U'DU)^{-1} = U^{-1}D^{-1}U'^{-1} = U'D^{-1}U$. Treća jednakost vrijedi zbog toga što je matrica U ortogonalna. Također, obratimo pažnju na to da je inverz dijagonalne matrice $D^{-1} = diag(1/d_1, \ldots, 1/d_P)$. Primijetimo da inverz postoji ako su svi $d_i \neq 0$, iz toga slijedi da su svi $1/d_i > 0$ jer je S_{pooled} pozitivno semidefinitna matrica. Dakle, S^{-1}_{pooled} je pozitivno definitna matrica. Važno je spomenuti da zbog pozitivne definitnosti matrice vrijedi $X'S^{-1}_{pooled}X > 0, \forall X \neq 0$.

Vratimo se sada na svojstva metrike. Lako je za uočiti da svojstvo 1) vrijedi. Svojstvo 2) dokazujemo na sljedeći način:

$$D_M = \sqrt{(x_t - x_r)' S^{-1}_{pooled}(x_t - x_r)} = 0 \Leftrightarrow (x_t - x_r)' S^{-1}_{pooled}(x_t - x_r) = 0,$$

a kako znamo da je S^{-1}_{pooled} pozitivno definitna matrica slijedi

$$(x_t - x_r) = 0 \Rightarrow x_t = x_r.$$

Izraz ispod korijena nije ništa drugo nego skalarni produkt $< S^{-1}_{pooled}(x_t - x_r), (x_t - x_r) >$, preko simetričnosti matrice dokazujemo svojstvo 3).

Da bi dokazali svojstvo 4) definirajmo $D_M(x) = \sqrt{x' S^{-1}_{pooled} x}$ i $D_M(y) = \sqrt{y' S^{-1}_{pooled} y}$ te $D_M(x + y) = \sqrt{(x + y)' S^{-1}_{pooled}(x + y)}$. Od ranije znamo da je $S^{-1}_{pooled} = U' D^{-1} U$. Označimo svojstvene vrijednosti od S^{-1}_{pooled} sa λ_i i neka je $Q = diag(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_P})U$. Primjetimo da

$$S^{-1}_{pooled} = Q'Q.$$

Neka je $z = Qx$ i $e = Qy$ te neka je $\|v\|_E = \sqrt{v_1^2 + \ldots + v_n^2}$ euklidova norma. Tada

$$D_M(x) = \sqrt{x' S^{-1}_{pooled} x} = \sqrt{x' Q' Q x} = \sqrt{z'z} = \|z\|_E$$

analogno vrijedi za $D_M(y)$.

$$D_M(x + y) = \sqrt{(x + y)' S^{-1}_{pooled}(x + y)} = \sqrt{(x + y)' Q' Q (x + y)} = \sqrt{(z + e)'(z + e)} = \|z + e\|_E$$

Kako znamo da za euklidovu normu vrijedi

$$\|z + e\|_E \leq \|z\|_E + \|e\|_E$$

automatski slijedi

$$D_M(x + y) \leq D_M(x) + D_M(y)$$

odnosno

$$D_M(x - z) = D_M((x - y) + (y - z)) \leq D_M(x - y) + D_M(y - z).$$

Time smo dokazali da je Mahalanobisova udaljenost zaista metrika.

2.2 Regija pouzdanosti

Devedeset postotna (90%) regija pouzdanosti (CR) za razliku između dvaju multivarijantnih srednjih vrijednosti definirana je na sljedeći način ([7]):

$$CR = K[(y - (x_t - x_r))' S^{-1}_{pooled}(y - (x_t - x_r))] \leq F_{P,2n-P-1,.90} \tag{2.2}$$

gdje $y = (\mu_{t1} - \mu_{r1}, \ldots, \mu_{tP} - \mu_{rP})$ predstavlja razliku koja se nalazi u CR, a

$$K = \frac{n}{2}\frac{2n - P - 1}{(2n - 2)P} \tag{2.3}$$

i $F_{P,2n-P-1,.90}$ je 90-ti percentil F-distribucije sa stupnjevima slobode P i $2n - P - 1$ gdje je n broj tableta.

Ukoliko je $P \leq 3$ regiju pouzdanosti možemo grafički usporediti sa regijom sličnosti i na temelju grafa zaključiti o sličnosti dvaju profila oslobađanja. U sljedećem poglavlju detaljnije opisujem usporedbu dvaju profila oslobađanja preko grafa.

2.3 Metoda Lagrangeovih multiplikatora

Pretpostavimo da su zadane dvije funkcije više varijabli $f, \varphi : D \to \mathbb{R}$ definirane na skupu $D \subseteq \mathbb{R}^k$. Funkciji φ pridružimo implicitnu jednadžbu $\varphi(y_1, \ldots, y_k) = 0$ i pripadajući skup $S \subseteq D$ definiran tom jednadžbom $S = \{(y_1, \ldots, y_k) \in D \mid \varphi(y_1, \ldots, y_k) = 0\}$. Problem tipa

$$\begin{cases} z = f(x_1, \ldots, x_k) \to max, min \\ \varphi(x_1, \ldots, x_k) = 0 \end{cases}$$

često rješavamo uvođenjem Lagrangeove funkcije $L(x_1, \ldots, x_k, \lambda)$:

$$L(x_1, \ldots, x_k, \lambda) = f(x_1, \ldots, x_k) + \lambda\varphi(x_1, \ldots, x_k), \quad (x_1, \ldots, x_k) \in D, \quad \lambda \in \mathbb{R}.$$

Parametar λ se zove **Lagrangeov multiplikator** (iz [6]). U našem slučaju k=P, a

$$\varphi(x_1, \ldots, x_p) = 0 \Leftrightarrow (rub\ od\ CR) - F_{P,2n-P-1,.90} = 0.$$

2.4 Interval pouzdanosti

Preko devedeset postotne (90%) regije pouzdanosti koja je definirana oko razlike vektora $x_r = (x_{r1}, \ldots, x_{rP})$ i $x_t = (x_{t1}, \ldots, x_{tP})$ dolazimo do 90% pouzdanog intervala Mahalanobisove udaljenosti (D_M) za tu razliku. Koristimo metodu Lagrangeovog multiplikatora

kako bi pronašli točke $(\mu_{t1} - \mu_{r1}, \ldots, \mu_{tP} - \mu_{rP})$ na granici regije pouzdanosti za koje Mahalanobisova udaljenost od ishodišne točke $(0, 0, ...)$ poprima minimum i maksimum. Točke na granici regije pouzdanosti za koje D_M poprima minimum i maksimum su ([8]):

$$y_1^* = (x_t - x_r)\left(1 + \sqrt{\frac{F_{P,2n-P-1,.90}}{K(x_t - x_r)' S^{-1}_{pooled}(x_t - x_r)}}\right) \tag{2.4}$$

$$y_2^* = (x_t - x_r)\left(1 - \sqrt{\frac{F_{P,2n-P-1,.90}}{K(x_t - x_r)' S^{-1}_{pooled}(x_t - x_r)}}\right) \tag{2.5}$$

Interval (D_M^l, D_M^u) je 90% pouzdani interval Mahalanobisove udaljenosti gdje su:

$$D_M^l = Min\left(\sqrt{(y_1^*)' S^{-1}_{pooled}(y_1^*)},\ \sqrt{(y_2^*)' S^{-1}_{pooled}(y_2^*)}\right) \tag{2.6}$$

$$D_M^u = Max\left(\sqrt{(y_1^*)' S^{-1}_{pooled}(y_1^*)},\ \sqrt{(y_2^*)' S^{-1}_{pooled}(y_2^*)}\right) \tag{2.7}$$

Primjetimo da D_M^l i D_M^u nisu ništa drugo nego Mahalanobisove udaljenosti ishodišta od točaka y_1^* i y_2^*. Mahalanobisova udaljenost ostalih točaka u regiji pouzdanosti poprima vrijednost unutar intervala (D_M^l, D_M^u). Ukoliko nul vektor (ishodište) zadovoljava uvjete od CR (2.2) onda $D_M^l = 0$.

Kažemo da su dva profila oslobađanja globalno slična ako se interval (D_M^l, D_M^u) nalazi unutar intervala

$$\left(-\sqrt{[D_g]' S^{-1}_{pooled}[D_g]},\quad \sqrt{[D_g]' S^{-1}_{pooled}[D_g]}\right) \tag{2.8}$$

gdje je $[D_g]'$ vektor $1 \times P$ sa svim vrijednostima jednakim globalnom SL D_g (uobičajeno $D_g = 10\%$).

2.5 Primjena

U sljedećem primjeru uspoređujemo profil otapanja testne i referentne tablete. Izračunat ćemo 90% interval pouzdanosti Mahalanobisove udaljenosti i na temelju njega donijeti zaključak o globalnoj sličnosti profila dvaju lijekova.

Primjer

U tablicama Referentni Lijek i Testni Lijek nalaze se 12 profila oslobađanja referentnih i 12 profila testnih tableta. Vrijednosti u tablicama predstavljaju količinu oslobođenog sadržaja aktivne susptance (izraženo u postotku) u 10., 15., 30. i 45. minuti.

Referentni Lijek				
Tableta	10 minuta	15 minuta	30 minuta	45 minuta
1	34	47	68	79
2	34	49	72	81
3	38	54	76	83
4	29	41	64	75
5	29	42	66	77
6	24	38	63	74
7	24	37	62	74
8	10	34	58	71
9	22	33	59	72
10	22	34	59	72
11	24	36	61	73
12	23	37	62	75

Testni Lijek				
Tableta	10 minuta	15 minuta	30 minuta	45 minuta
1	69	80	87	89
2	71	82	90	91
3	71	80	88	90
4	67	80	87	89
5	70	80	87	88
6	67	79	86	88
7	71	82	90	91
8	72	83	91	93
9	71	82	91	93
10	70	81	90	92
11	70	82	91	93
12	72	83	91	92

Primjetimo kako je broj vremenskih točaka $P = 4 \succ 3$ pa regiju pouzdanosti ne možemo prikazati grafički. Prvo idemo izračunati Mahalanobisovu udaljenost D_M preko formule (2.1). Imamo

$$x_r = (26.08333, 40.16667, 64.16667, 75.50000)$$

i

$$x_t = (70.08333, 81.16667, 89.08333, 90.75000)$$

iz čega slijedi

$$(x_t - x_r) = (44.00000, 41.00000, 24.91667, 15.25000).$$

Izračunajmo još S_{pooled}

$$S_{pooled} = \begin{bmatrix} 28.62879 & 22.71212 & 19.44318 & 13.715909 \\ 22.71212 & 23.15152 & 19.34091 & 13.522727 \\ 19.44318 & 19.34091 & 16.93561 & 12.056818 \\ 13.71591 & 13.52273 & 12.05682 & 8.965909 \end{bmatrix}$$

iz čega slijedi

$$S_{pooled}^{-1} = \begin{bmatrix} 0.164896977 & -0.07767432 & -0.1036772 & 0.004313442 \\ -0.077674324 & 1.14178018 & -1.7204574 & 0.710316544 \\ -0.103677202 & -1.72045736 & 4.2820686 & -3.004806198 \\ 0.004313442 & 0.71031654 & -3.0048062 & 3.074291971 \end{bmatrix}.$$

Jednadžbom (2.1) dolazimo do $D_M = 14.13404$.

Znamo da je $n = 12$ i $P = 4$, koristeći formulu (2.3) dolazimo do $K = 1.295455$. Sljedeći korak je izračunati 90% interval pouzdanosti Mahalanobisove udaljenosti. Kako bi to mogli odrediti prvo trebamo izračunati točke na granici regije pouzdanosti za koje D_M poprima minimum i maksimum. Preciznije, računamo y_1^* i y_2^*. Koristeći jednadžbe (2.4) i (2.5) dolazimo do

$$y_1^* = (48.11751, 44.83677, 27.24836, 16.67709)$$

i

$$y_2^* = (39.88249, 37.16323, 22.58497, 13.82291).$$

Konačno, 90% interval pouzdanosti Mahalanobisove udaljenosti je $(D_M^l, D_M^u) = (12.81138, 15.4567)$. Ostaje nam još provjeriti da li se interval pouzdanosti nalazi unutar intervala ($D_g = (10, 10, 10, 10)$)

$$\left(- \sqrt{[D_g]' S_{pooled}^{-1} [D_g]}, \quad \sqrt{[D_g]' S_{pooled}^{-1} [D_g]} \right) = (-5.282684, 5.282684).$$

Kako je D_M^u veći od 5.282684 zaključujemo da profili otapanja dvaju lijekova nisu globalno slični.

Poglavlje 3

Parametarska metoda

Parametarska metoda opisana u ovom poglavlju koristi se za testiranje sličnosti profila oslobađanja. Metoda se temelji na prilagodbi matematičke krivulje na profile oslobađanja i daljnjem promatranju tih profila kroz parametre funkcije.

Primjena ove metode se preporuča u situacijama kada imamo puno podataka, odnosno profila i kada se uzorkovanje radi u 4 ili više vremenskih točaka. U primjeru na kraju poglavlja radimo sa jednom testnom serijom i tri referentne serije gdje svaka serija ima po 12 tableta.

3.1 Koraci metode

1. korak - Odredimo funkciju koja se najbolje prilagođava profilima iz referentnih serija.

2. korak - Odabranu funkciju prilagođavamo svakom profilu oslobađanja iz referentnih serija. Profile oslobađanja promatramo kroz parametre odabrane funkcije.

3. korak - Izračunamo varijancu parametara za svaku referentnu seriju i varijancu između referentnih serija. Izračunamo $Pooled_{SD}$ i na temelju toga definiramo regiju sličnosti.

4. korak - Koristeći funkciju iz prvog koraka, opisujemo svaki profil iz testne serije i referentne serije. Pritom, testna serije ima n profila (uobičajeno 12) nekog lijeka kojeg uspoređujemo s referentnim lijekom.

5. korak - Izračunati statističku udaljenost (Mahalanobisovu udaljenost) između dvaju multivarijantnih srednjih vrijednosti parametara.

6. korak - Postaviti devedeset postotnu (90%) regiju pouzdanosti (CR) oko razlike dvaju multivarijantnih srednjih vrijednosti.

7. korak - Usporediti regiju pouzdanosti sa regijom sličnosti (definiranom u 3. koraku) kako bi odredili da li su slični.

3.2 Funkcije kandidati

Ideja je pronaći što jednostavniju funkciju (s malo parametara) koja najbolje odgovara našim profilima. Neke od funkcija koje se koriste za fitanje profila su:

1. Probit
$$\frac{X}{100} = \phi(\alpha + \beta \cdot (\log(t)))$$
gdje je ϕ=standardna normalna distirbucija

2. Logistic
$$\frac{X}{100} = \frac{e^{(\alpha+\beta\cdot\log(t))}}{1 + e^{(\alpha+\beta\cdot\log(t))}}$$

3. Weibull
$$\frac{X}{100} = 1 - e^{-\alpha\cdot t^{\beta}}$$

4. Quadratic
$$\frac{X}{100} = \alpha + \beta 1 \cdot (t - \bar{t}) + \beta 2 \cdot (t - \bar{t})^2$$
gdje je $\bar{t}$ = prosjek od svih uzorkovanih vremena

5. Exponential
$$\frac{X}{100} = 1 - e^{-\alpha\cdot t}$$

Metodu koja će se koristiti radi odabira funkcije (koja najbolje odgovara našim podacima) odabire sam ispitivač. Funkcija kojom se bavimo u nastavku poglavlja i koja se često koristi u praksi je Weibull-ova funkcija. Weibull-ovom funkcijom možemo opisati profil tablete čije oslobađanje kreće s odgodom, a preko linearizirane forme funkcije lako dolazimo do parametara koristeći metodu najmanjih kvadrata. Linearizirana forma Weibull-ove funkcije je

$$\ln\left(-\ln\left(1 - \frac{X_{(t)}}{X_{(inf.)}}\right)\right) = \ln(\alpha) + \beta \ln(t) \tag{3.1}$$

gdje $X_{(t)}$ predstavlja količinu oslobođenog sadržaja (u vremenu t) aktivne susptance (izraženo u postotku), $X_{(inf.)}$ predstavlja količinu oslobođenog sadržaja (u nekoj beskonačno udaljenoj vremenskoj točki) aktivne susptance (izraženo u postotku, uobičajeno 100%), α je faktor skaliranja i β je faktor koji utječe na oblik krivulje.

3.3 Regija sličnosti određena s $Pooled_{SD}$

Weibull-ovom funkcijom opisujemo svaki profil (uobičajeno 12 profila) iz referentnih serija. Na taj način dolazimo do parametara na temelju kojih promatramo naše profile. Kako naša funkcija ima dva parametra α i β, svaki profil je opisan svojim parom (α, β).

Regiju sličnosti definiramo preko varijanci logaritmiranih parametara unutar serija i varijanci između serija. Preciznije nakon logaritamske transformacije parametara, varijance unutar serija i varijance između serija združujem na sljedeći način ([5]):

$$Pooled_{SD} = \sqrt{\left(\frac{Var_1 + Var_2 + \cdots + Var_n}{n}\right) + Var(\overline{x})} \tag{3.2}$$

gdje se $Pooled_{SD}$ računa posebno za parametar α i posebno za parametar β. Primjerice, prilikom izračuna $Pooled_{SD}$ za α, Var_i označava varijancu logaritmiranog parametra α za i-tu seriju (imamo n serija), a $Var(\overline{x})$ je varijanca prosjeka logaritmiranih parametara α za svaku od n serija referentnog lijeka. Na temelju $Pooled_{SD}$ (za oba parametra) konstruiramo regiju sličnosti.

Uz pretpostavku da su parametri nezavisni i normalno distribuirani možemo konstruirati regiju sličnosti kao pravokutnik kojem su stranice na x-osi i y-osi veličine $2Pooled_{SD}$, $4Pooled_{SD}$ i $6Pooled_{SD}$. Najmanji pravokutnik obuhvaća 68% slučajeva, srednji 95% ,a najveći 99% slučajeva. Na slici 3.1 $Pooled_{SD}$ za α (x-os) je 0.3, a za β (y-os) je 0.1.

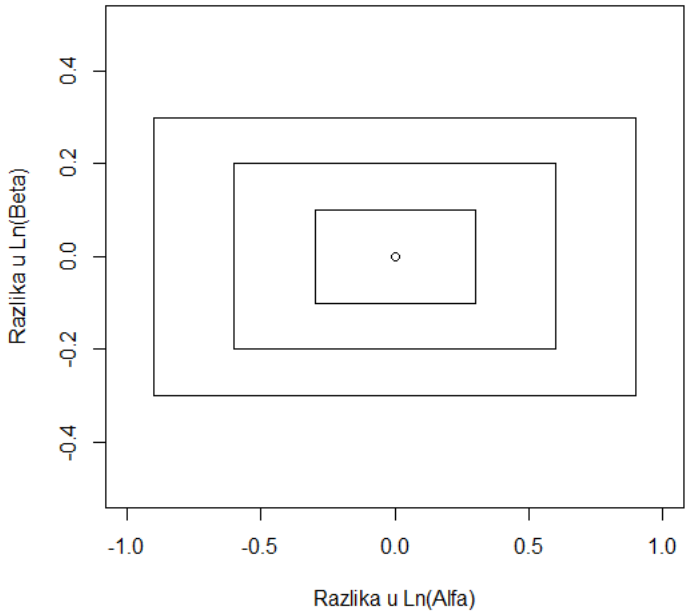
3.4 Mahalanobisova udaljenost

Mahalanobisovu udaljenost definirali smo ranije na stranici 11 ovog diplomskog rada s formulom (2.1). U ovoj metodi oznake iz formule (2.1) predstavljaju druge vrijednosti nego su predstavljale u neparametarskoj metodi. Uz raniju pretpostavku da naša funkcija ima dva parametra (α, β) vrijedi sljedeće:

- $x_r = (x_{r1}, x_{r2})$ vektor srednjih vrijednosti referentnog lijeka, gdje x_{ri} predstavlja srednju vrijednost logaritmiranog i-tog parametra za uobičajeno 12 tableta, a analogno vrijedi za $x_t = (x_{t1}, x_{t2})$ vektor srednjih vrijednosti testnog lijeka

- $S_{pooled} = \frac{S_{ref}+S_{test}}{2}$ je kovarijacijska matrica uzoraka, združena za testni i referentni lijek, gdje su S_{ref} i S_{test} kovarijacijske matrice logaritmiranih parametara testnog i referentnog lijeka

Označimo s D^2 Kvadriranu Mahalanobisovu udaljenost.



Slika 3.1: Regija sličnosti određena s $Pooled_{SD}$

Prema [5] Mahalanobisova udaljenost nakon skaliranja odgovara Hotelling-ovoj T^2 statistici koja pod predpostavkom da je $x_t = x_r$ ima F distribuciju (potrebno za definiciju regije pouzdanosti). Definirajmo **Faktor skaliranja** K' na sljedeći način:

$$K' = \frac{N_1 + N_2 - P - 1}{(N_1 + N_2 - 2)P} \cdot \frac{N_1 N_2}{N_1 + N_2 - P} \tag{3.3}$$

gdje N_1 i N_2 predstavljaju broj tableta u testnoj i referentnoj seriji, a P predstavlja broj parametara. Skalirana (normalizirana) Mahalanobisova udaljnost je definirana na sljedeći način ([5]):

$$T^2 = K' * D^2. \tag{3.4}$$

3.5 Regija pouzdanosti

Devedeset postotna (90%) regija pouzdanosti (CR) za razliku između dvaju multivarijantnih srednjih vrijednosti u ovoj metodi definirana je na sljedeći način ([5]):

$$CR = K'[(y - (x_t - x_r))' S^{-1}_{pooled}(y - (x_t - x_r))] \leq F_{P,N_1+N_2-P-1,.90} \tag{3.5}$$

gdje y predstavlja multivarijantnu razliku parametara koja se nalazi u CR, a $F_{P,N_1+N_2-P-1,.90}$ je 90-ti percentil F-distribucije sa stupnjevima slobode P i $N_1 + N_2 - P - 1$. Ideja je nacrtati regiju pouzdanosti i pogledati da li se nalazi unutar regije sličnosti. Ukoliko se nalazi unutar regije sličnosti, zaključujemo da su profili otapanja testnog i referentnog lijeka slični.

3.6 Primjena

U sljedećem primjeru uspoređujemo profil otapanja testne i referentne tablete. Kreirati ćemo 90% regiju pouzdanosti i na temelju usporedbe sa regijom sličnosti donijeti zaključak o sličnosti profila dvaju lijekova.

Primjer

U tablicama Referentni Lijek (1, 2, 3) i Testni Lijek nalaze se 12 profila oslobađanja referentnih i 12 profila testnih tableta. Vrijednosti u tablicama predstavljaju količinu oslobođenog sadržaja aktivne susptance (izraženo u postotku) u 10., 15., 30. i 45. minuti. Nakon što odredimo regiju sličnosti uspoređivat ćemo profile iz tablica Referentni Lijek 3 i Testni lijek.

Referentni Lijek 1				
Tableta	10 minuta	15 minuta	30 minuta	45 minuta
1	60	72	85	91
2	49	63	81	89
3	53	62	75	81
4	41	54	71	80
5	63	72	87	94
6	64	75	90	96
7	37	45	61	70
8	63	70	83	89
9	60	69	85	92
10	38	47	65	76
11	46	56	75	84
12	62	70	86	93

Referentni Lijek 2				
Tableta	10 minuta	15 minuta	30minuta	45 minuta
1	43	51	69	79
2	43	51	67	76
3	40	47	63	72
4	52	62	79	88
5	39	46	61	70
6	59	69	85	92
7	61	69	84	90
8	65	73	88	93
9	36	43	57	64
10	59	67	85	92
11	46	53	71	79
12	41	46	59	66

Referentni Lijek 3				
Tableta	10 minuta	15 minuta	30 minuta	45 minuta
1.	41	51	68	79
2.	44	55	73	82
3.	14	17	23	39
4.	52	63	79	85
5.	31	39	56	68
6.	52	60	73	81
7.	52	62	78	85
8.	38	47	62	72
9.	39	48	65	75
10.	49	56	72	86
11.	57	61	83	85
12.	54	62	78	85

Testni Lijek				
Tableta	10 minuta	15 minuta	30 minuta	45 minuta
1.	57	65	78	84
2.	56	64	76	83
3.	55	63	75	81
4.	53	61	74	81
5.	50	58	72	80
6.	53	60	73	80
7.	55	63	75	83
8.	53	61	74	81
9.	56	64	77	84
10.	52	59	72	79
11.	59	66	79	86
12.	55	63	75	82

Prateći korake metode rješavamo dani primjer. Funkciju koju smo primjenili u ovom primjeru je Weibulova funkcija. Koristeći lineariziranu formu Weibulove funkcije (3.1) i primjenom metode najmanjih kvadrata dolazimo do parametara funkcije za svaki profil. Koristeći logaritamsku transformaciju parametara dolazimo do sljedećih setova podataka:

Referentni Lijek 1			Referentni Lijek 2		
Tableta	$\ln(\alpha)$	$\ln(\beta)$	Tableta	$\ln(\alpha)$	$\ln(\beta)$
1.	-1.508256	-0.4609435	1.	-2.170972	-0.3779172
2.	-2.160416	-0.2475182	2.	-2.014865	-0.4742177
3.	-1.468060	-0.6478766	3.	-2.098417	-0.4876208
4.	-2.278910	-0.3154796	4.	-1.932476	-0.3530154
5.	-1.610798	-0.3716011	5.	-2.084473	-0.5185720
6.	-1.726323	-0.2776344	6.	-1.714072	-0.3668421
7.	-2.245302	-0.4462509	7.	-1.459122	-0.5049776
8.	-1.248367	-0.6254457	8.	-1.417093	-0.4596420
9.	-1.661480	-0.3886996	9.	-2.083865	-0.5853805
10.	2.417425	-0.3189708	10.	-1.770137	-0.3470310
11.	-2.167488	-0.3145597	11.	-1.965061	-0.4548032
12.	-1.620468	-0.3880633	12.	-1.773907	-0.7223893
Varijanca	0.1494633	0.0162210	Varijanca	0.0623383	0.0117434
Prosjek	-1.842775	-0.4002536	Prosjek	-1.873705	-0.4710341

Referentni Lijek 3		
Tableta	ln(α)	ln(β)
1.	-2.278513	-0.3381868
2.	-2.188481	-0.3290305
3.	-3.654255	-0.3058735
4.	-1.746210	-0.4545839
5.	-2.711710	-0.2964238
6.	-1.548356	-0.6199056
7.	-1.758200	-0.4558082
8.	-2.211943	-0.4406836
9.	-2.281459	-0.3780925
10.	-2.052430	-0.3588490
11.	-1.599245	-0.5042319
12.	-1.649687	-0.5057977
Varijanca	0.3507198	0.0096232
Prosjek	-2.140041	-0.4156222

Kako se $Pooled_{SD}$ računa posebno za parametar α i posebno za parametar β računamo $Var(\overline{x})$ posebno za logaritmirani parametar ln(α) i logaritmirani parametar ln(β). $Var(\overline{x})$ za ln(α) je 0.02670978, a za ln(β) je 0.001386089. $Pooled_{SD}$ za parametar α računamo na sljedeći način:

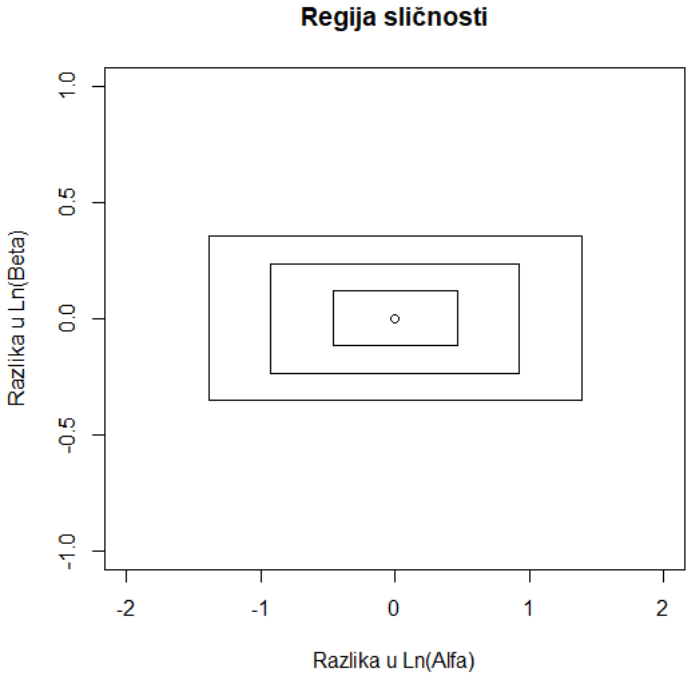
$$Pooled_{SD}(\alpha) = \sqrt{\left(\frac{0.1494633 + 0.06233834 + 0.3507198}{3}\right) + 0.02670978} = 0.4628358.$$

Također $Pooled_{SD}$ za parametar β računamo na sljedeći način:

$$Pooled_{SD}(\beta) = \sqrt{\left(\frac{0.01622102 + 0.01174336 + 0.00962319}{3}\right) + 0.001386089} = 0.117963.$$

Sada kada smo izračunali $Pooled_{SD}$ za parametre konstruiramo regiju sličnosti 3.2.

Kako smo ranije našli parametre za profile iz tablica Referentni lijek (1, 2, 3), analogno tražimo parametre za profile iz tablice Testni Lijek.



Slika 3.2: Regija sličnosti određena s $Pooled_{SD}$

Testni Lijek		
Tableta	ln(α)	ln(β)
1.	-1.356663	-0.6582509
2.	-1.358794	-0.6804876
3.	-1.332472	-0.7224197
4.	-1.480890	-0.6485953
5.	-1.655027	-0.5818176
6.	-1.449202	-0.6826394
7.	-1.424859	-0.6517629
8.	-1.480890	-0.6485953
9.	-1.422076	-0.6306511
10.	-1.472990	-0.6857867
11.	-1.339170	-0.6405842
12.	-1.378130	-0.6868768
Prosjek	-1.4292636	-0.6598723

Sljedeći korak je izračunati Mahalanobisovu udaljenost između

$$x_r = (-2.1400408, -0.4156222)$$

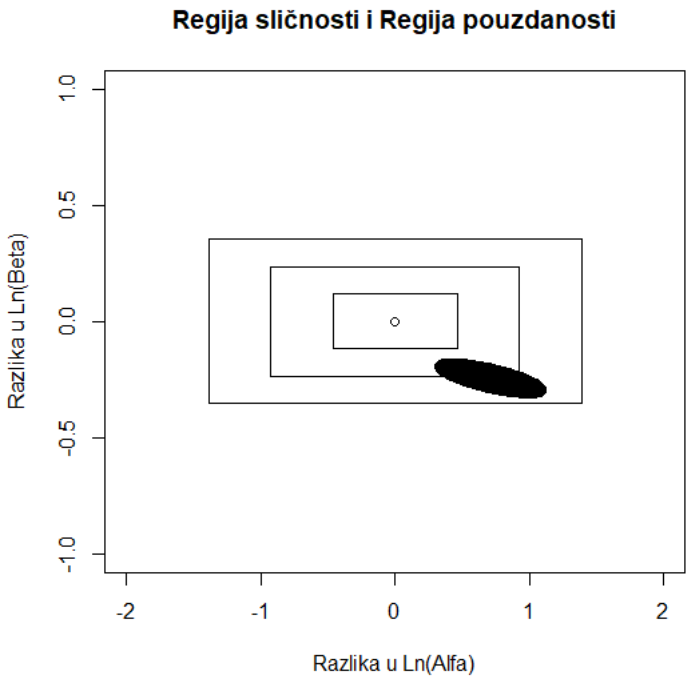
i

$$x_t = (-1.4292636, -0.6598723).$$

Mahalanobisova udaljenost izračunata na ekvivalentan način kao u neparametarskoj metodi, poprima vrijednost 3.52652. Ostaje još samo regiju pouzdanosti usporediti sa regijom sličnosti. Faktor skaliranja K' (3.3) poprima vrijednost 3.124. Neka $Y = (x_\alpha, y_\beta)$ predstavlja multivarijantnu razliku parametara koja se nalazi u regiji pouzdanosti, tada vrijedi (3.5):

$$CR = 39.768(x_\alpha^2)+(1310.705(y_\beta^2))+(27.094x_\alpha)+(397.11y_\beta)+38.884+342.327x_\alpha y_\beta \leq 2.57457.$$

Skup svih Y koji zadovoljavaju gornji uvjet čini regiju pouzdanosti koju na slici 3.3 uspoređujemo s regijom sličnosti.



Slika 3.3: Regija sličnosti i regija pouzdanosti

Vidimo da se naša regija pouzdanosti nalazi unutar trećeg pravokutnika, na temelju toga možemo zaključiti da su profili testnog i referentnog lijeka slični.

Parametarska metoda je dosta fleksibilna jer ne zahtijeva ista vremena uzorkovanja za testne i referentne tablete. Važno je spomenuti da nije jasno kako protumačiti slučaj u kojem je elipsa malo izašla izvan granica regije sličnosti. Također, nije jasno niti koliko $Pooled_{SD}$ treba uzeti za kreiranje regije sličnosti, a i tu je problem zavisnosti između parametara.

Bibliografija

[1] *Jackknife resampling*, dostupno je na https://en.wikipedia.org/wiki/Jackknife_resampling, (siječanj 2018.).

[2] Boris Guljaš, *METRIČKI PROSTORI*, dostupno je na https://web.math.pmf.unizg.hr/~guljas/skripte/metprost.pdf, (siječanj 2018.).

[3] V. P. Shah, Y. Tsong, P. Sathe i J.P. Liu, *In Vitro Dissolution Profile Comparison - Statistics and Analysis of the Similarity Factor,* f_2, Pharmaceutical Research **15** (1998), br. 6, 889–896.

[4] Melita Petrović, *ORIGINALNI ILI GENERIČKI LIJEK?*, dostupno je na http://www.plivazdravlje.hr/aktualno/clanak/8920/Originalni-ili-generic ki-lijek.html, (siječanj 2018.).

[5] P. Sathe, Y. Tsong i V. P. Shah, *In Vitro Dissolution Profile Comparison: Statistics and Analysis, Model Dependent Approach*, Pharmaceutical Research **13** (1996), br. 12, 1799–1803.

[6] M. Tepić, *KOMPLEKSNOST SKRIVENIH MARKOVLJEVIH MODELA*, Diplomski rad, Sveučilište u Zagrebu, Prirodoslovno-Matematički fakultet, Matematički odsjek, 2015.

[7] Y. Tsong, T. Hammerstrom, P. Sathe i V. P. Shah, *STATISTICAL ASSESSMENT OF MEAN DIFFERENCES BETWEEN TWO DISSOLUTION DATA SETS*, Drug Information Journal **30** (1996), 1105–1112.

[8] Y. Tsong, P. Sathe i V. P. Shah, *COMPARING TWO DISSOLUTION DATA SETS FOR SIMILARITY*, **30**, 129–134.

[9] J. W. Moore i H. H. Flanner, *Mathematical Comparison of Dissolution Profiles*, (1996), 64–74.

[10] Vanja Wagner, *BOOTSTRAP*, dostupno je na https://web.math.pmf.unizg.hr/~wagner/userfiles/nastava/sp2_vjezbe10.pdf, (siječanj 2018.).

Sažetak

Jedna od tehnika usporedbe testnog i referentnog lijeka je usporedba brzine oslobađanja lijeka u nekom relevantnom mediju.

U radu su opisane standardne statističke metode (metoda faktora, neparametarska metoda i parametarska metoda) za usporedbu profila oslobađanja testnog i referentnog lijeka. Primjenom metoda (koristeći programski jezik R) na različitim primjerima detaljnije je objašnjena njihova upotreba.

Summary

One of the techniques for comparison of a test and a reference drug is to compare a dissolution rate of drugs in some dissolution medium.

In this thesis, we describe standard statistical methods (factor method, model independent method and model dependent method) for comparison between dissolution profiles of a test and a reference drug. By applying the methods (using programming language R) to different examples we are describing their application in detail.

Dodatak

U nastavku su dodani kodovi (napisani u programskom jeziku R) korišteni za rješavanje ranije navedenih primjera.

```
Primjer iz prvog poglavlja.
Prvi način:

f2<-function(R,T,P)
{   razlika<-T-R
    kvadrat<-razlika^2
    suma1<-sum(kvadrat)
    suma2<-1+((1/P)*suma1)
    korjen<-suma2^(-(1/2))
    logarit<-log10(korjen*100)
    f2<-50*logarit
    return(f2)
}

podaci1<-read.table("Aref.csv",sep=";")
refmatrica<-matrix(0,nrow=12,ncol=4,byrow=TRUE)
podaci2<-read.table("Btest.csv",sep=";")
testmatrica<-matrix(0,nrow=12,ncol=4,byrow=TRUE)

for ( i in 1:12)
{
for (j in 1:4)
{
refmatrica[i,j]<-podaci1[i,j]
testmatrica[i,j]<-podaci2[i,j]
}
}

### matrice za referentni lijek i za testni lijek
r1<-refmatrica[,1]
r2<-refmatrica[,2]
r3<-refmatrica[,3]
r4<-refmatrica[,4]

t1<-testmatrica[,1]
t2<-testmatrica[,2]
t3<-testmatrica[,3]
t4<-testmatrica[,4]

#r1,r2,r3,r4<-originalni podatci referentnog lijeka po vremenima
#t1,t2,t3,t4<-originalni podatci testnog lijeka po vremenima
#1-30,2-60,3-90,4-180

###
percentilL<-0.05
percentilU<-0.95
P<-4 #vremenske točke
velicinaboot<-1000#brojbootstrapa

#radim vektor s indeksima
vektorindex<-c(1:12)
```

```
vektorindex

# radim bootstrap uzorak tog indeks vektora
# gdje svaki broj(indeks) predstavlja jednu referentnu i testnu krivulju

refbs<-replicate(velicinaboot,sample(vektorindex,12,replace=TRUE))
testbs<-replicate(velicinaboot,sample(vektorindex,12,replace=TRUE))

f2kapica<-numeric(velicinaboot)
f2kapica
for ( i in 1:velicinaboot){
  amatrica<-matrix(0,nrow=12,ncol=4) #referentna
  bmatrica<-matrix(0,nrow=12,ncol=4) #testna
  for (r in 1:12){
     if(refbs[r,i]==1){
        amatrica[r,]<-refmatrica[1,]
} else if (refbs[r,i]==2) {
     amatrica[r,]<-refmatrica[2,]
}else if (refbs[r,i]==3) {
     amatrica[r,]<-refmatrica[3,]
}else if (refbs[r,i]==4) {
     amatrica[r,]<-refmatrica[4,]
}else if (refbs[r,i]==5) {
     amatrica[r,]<-refmatrica[5,]
}else if (refbs[r,i]==6) {
     amatrica[r,]<-refmatrica[6,]
}else if (refbs[r,i]==7) {
     amatrica[r,]<-refmatrica[7,]
}else if (refbs[r,i]==8) {
     amatrica[r,]<-refmatrica[8,]
}else if (refbs[r,i]==9) {
     amatrica[r,]<-refmatrica[9,]
}else if (refbs[r,i]==10) {
     amatrica[r,]<-refmatrica[10,]
}else if (refbs[r,i]==11) {
     amatrica[r,]<-refmatrica[11,]
}else if (refbs[r,i]==12) {
     amatrica[r,]<-refmatrica[12,]
}

if(testbs[r,i]==1){
        bmatrica[r,]<-testmatrica[1,]
} else if (testbs[r,i]==2) {
     bmatrica[r,]<-testmatrica[2,]
}else if (testbs[r,i]==3) {
     bmatrica[r,]<-testmatrica[3,]
}else if (testbs[r,i]==4) {
     bmatrica[r,]<-testmatrica[4,]
}else if (testbs[r,i]==5) {
     bmatrica[r,]<-testmatrica[5,]
}else if (testbs[r,i]==6) {
     bmatrica[r,]<-testmatrica[6,]
}else if (testbs[r,i]==7) {
     bmatrica[r,]<-testmatrica[7,]
}else if (testbs[r,i]==8) {
     bmatrica[r,]<-testmatrica[8,]
}else if (testbs[r,i]==9) {
```

```
     bmatrica[r,]<-testmatrica[9,]
}else if (testbs[r,i]==10) {
     bmatrica[r,]<-testmatrica[10,]
}else if (testbs[r,i]==11) {
     bmatrica[r,]<-testmatrica[11,]
}else if (testbs[r,i]==12) {
     bmatrica[r,]<-testmatrica[12,]
}
}

    R<-c(mean(amatrica[,1]),mean(amatrica[,2]),mean(amatrica[,3]),mean(amatrica[,4]))
    T<-c(mean(bmatrica[,1]),mean(bmatrica[,2]),mean(bmatrica[,3]),mean(bmatrica[,4]))
    f2kapica[i]<-f2(R,T,P)
}
f2kapica

# f2kapica(m) je bootstrap procjenitelj m-tog bootstrap uzorka

# sada racunamo f2kapicaorig gdje je to procjenitelj
# originalnih vrijednosti testnih i referentnih lijekova
orig1<-c(mean(r1),mean(r2),mean(r3),mean(r4))
orig2<-c(mean(t1),mean(t2),mean(t3),mean(t4))

f2kapicaorig<-f2(orig1,orig2,P)

###
## f2kapjack[i] je i-ta jackknife statistic
###
# i. jackknife statistika (bez i-te tablete)
# radim na originalnim podatcima
f2kapjack<-numeric(12)

for (i in 1:12){
rj<-matrix(0,nrow=11,ncol=4) #referentna
tj<-matrix(0,nrow=11,ncol=4) #testna
rj[,1]<-r1[-i]
rj[,2]<-r2[-i]
rj[,3]<-r3[-i]
rj[,4]<-r4[-i]

tj[,1]<-t1[-i]
tj[,2]<-t2[-i]
tj[,3]<-t3[-i]
tj[,4]<-t4[-i]

mrj<-c(mean(rj[,1]),mean(rj[,2]),mean(rj[,3]),mean(rj[,4]))
mtj<-c(mean(tj[,1]),mean(tj[,2]),mean(tj[,3]),mean(tj[,4]))

f2kapjack[i]<-f2(mrj,mtj,P)
}

f2kapjack
## f2kapica(.) je srednja vrijednost jackknife statistike
f2srjack<-mean(f2kapjack)

a=sum((f2srjack-f2kapjack)^3)/(6*(sum((f2srjack-f2kapjack)^2)^(3/2)))
```

```
br<-0
for (i in 1:velicinaboot){
      if (f2kapica[i]<f2kapicaorig)
         {br=br+1}
}

M<-velicinaboot
M
z0<-qnorm(br/M)
z0
zalfa1<-qnorm(percentilL) # 5-ti percentil qnorm(0.05)
zalfa2<-qnorm(percentilU)     # 95-ti percentil  qnorm(0.95)

alfa1<-pnorm(z0+((z0+zalfa1)/(1-(a*(z0+zalfa1)))))
alfa2<-pnorm(z0+((z0+zalfa2)/(1-(a*(z0+zalfa2)))))

index1<-floor(M*alfa1)
index2<-floor(M*alfa2)

sortf2kapica<-sort(f2kapica)
sortf2kapica

interval1<-sortf2kapica[index1]
interval2<-sortf2kapica[index2]
interval1
interval2
#####################################################

#####################################################
Drugi način (naveden je samo kod za drugi način bootstrap generiranja, ostalo slijedi analogno):

podaci1<-read.table("Aref.csv",sep=";")
podaci1
refmatrica<-matrix(0,nrow=12,ncol=4,byrow=TRUE)
refmatrica
podaci2<-read.table("Btest.csv",sep=";")
testmatrica<-matrix(0,nrow=12,ncol=4,byrow=TRUE)
testmatrica

for ( i in 1:12)
{
for (j in 1:4)
{
refmatrica[i,j]<-podaci1[i,j]
testmatrica[i,j]<-podaci2[i,j]
}
}
r1<-refmatrica[,1]
r2<-refmatrica[,2]
r3<-refmatrica[,3]
r4<-refmatrica[,4]
t1<-testmatrica[,1]
t2<-testmatrica[,2]
t3<-testmatrica[,3]
t4<-testmatrica[,4]
#r1,r2,r3,r4<-originalni podatci referentnog lijeka po vremenima
#t1,t2,t3,t4<-originalni podatci testnog lijeka po vremenima
```

```
#1-30,2-60,3-90,4-180

###
percentilL<-0.05
percentilU<-0.95
P<-4 #vremenske točke
brbootstrapa<-1000# broj bootstrap uzoraka

#neka su r30 ,r60 ,r90 ,r180 matrice
#bootstrap uzoraka za 30,60,90,180 minutu
r30=replicate(brbootstrapa,sample(r1,12,replace=TRUE))
r60=replicate(brbootstrapa,sample(r2,12,replace=TRUE))
r90=replicate(brbootstrapa,sample(r3,12,replace=TRUE))
r180=replicate(brbootstrapa,sample(r4,12,replace=TRUE))

# objašnjenje: prvi stupac iz r30 je dio prvog bootstrap uzorka referentnog
# lijeka, predstavlja vrijednosti otapanja od 12 krivulja u 30 minuti
# tako isto kada gledamo prvi stupac u r60,r90,r180.
# Gdje god da te život nosi
  uvik moraš znati tko si#
# Kada se spoje svi ti stupci imamo 12 krivulja kroz 4 vremenske točke
# u prvom redu je prva krivulja u drugom druga i tako do 12.
# neka su  svr30 ,svr60 ,svr90 ,svr180 matrice(vektor)srednjih vrijednosti
# bootstrap uzorka u vremenu 30,60,90,180

svr30<-numeric(brbootstrapa)
for (i in 1:brbootstrapa){
 svr30[i]<-mean(r30[,i])
}
svr60<-numeric(brbootstrapa)
for (i in 1:brbootstrapa){
 svr60[i]<-mean(r60[,i])
}
svr90<-numeric(brbootstrapa)
for (i in 1:brbootstrapa){
 svr90[i]<-mean(r90[,i])
}
svr180<-numeric(brbootstrapa)
for (i in 1:brbootstrapa){
 svr180[i]<-mean(r180[,i])
}
# neka su t30 ,t60 ,t90 ,t180 matrice
# bootstrap uzoraka za 30,60,90,180 minutu

t30=replicate(brbootstrapa,sample(t1,12,replace=TRUE))
t60=replicate(brbootstrapa,sample(t2,12,replace=TRUE))
t90=replicate(brbootstrapa,sample(t3,12,replace=TRUE))
t180=replicate(brbootstrapa,sample(t4,12,replace=TRUE))
svt30<-numeric(brbootstrapa)
for (i in 1:brbootstrapa){
 svt30[i]<-mean(t30[,i])
}
svt60<-numeric(brbootstrapa)
for (i in 1:brbootstrapa){
 svt60[i]<-mean(t60[,i])
}
svt90<-numeric(brbootstrapa)
```

```
for (i in 1:brbootstrapa){
 svt90[i]<-mean(t90[,i])
}
svt180<-numeric(brbootstrapa)
for (i in 1:brbootstrapa){
 svt180[i]<-mean(t180[,i])
}

f2kapica<-numeric(brbootstrapa)
for (i in 1:brbootstrapa){
    R<-c(svr30[i],svr60[i],svr90[i],svr180[i])
    T<-c(svt30[i],svt60[i],svt90[i],svt180[i])
    f2kapica[i]<-f2(R,T,P)
    }
#f2kapica(m) je bootstrap procjenitelj m-tog bootstrap uzorka.
#####################################################################

#####################################################################
Primjer iz drugog poglavlja.

podaci1<-read.table("REF34.csv",sep=";")
podaci2<-read.table("TEST4.csv",sep=";")

# RACUNAM MAHALANOBISOVU UDALJENOST (M-distance)
# D=sqrt((xt-xr)'*(spooledˆ-1)*(xt-xr))
# xt i xr su vektori srednjih vrijednosti po vremenima
xr<-c(mean(podaci1[,1]),mean(podaci1[,2]),mean(podaci1[,3]),mean(podaci1[,4]))
xt<-c(mean(podaci2[,1]),mean(podaci2[,2]),mean(podaci2[,3]),mean(podaci2[,4]))

spooled1<-matrix(0,nrow=4,ncol=4)
spooled1
spooled2<-matrix(0,nrow=4,ncol=4)
spooled2
for ( i in 1:4)
{
  for( j in 1:4)
  {
    spooled1[i,j]=cov(podaci1[,i],podaci1[,j])
    spooled2[i,j]=cov(podaci2[,i],podaci2[,j])

  }
}
spooled<-(spooled1+spooled2)/2
# inverz od spooled
invspooled<-solve(spooled)

DM=sqrt(t(xt-xr)%*%invspooled%*%(xt-xr))

DM # mahalanobisova udaljenost

# idemo odrediti (DML,DMU) lower 90% limit i upper 90% limit
P<-4  # broj mjerenja
n<-12 # broj tableta

K<-((n)/(2))*(2*n-P-1)/((2*n-2)*P)
F<-qf(0.90,P,2*n-P-1)
podkor1<-F/(K*t(xt-xr)%*%invspooled%*%(xt-xr))
```

```
y1<-(xt-xr)*(1+sqrt(podkor1))
y2<-(xt-xr)*(1-sqrt(podkor1))
DML<-min(sqrt(t(y1)%*%invspooled%*%y1),sqrt(t(y2)%*%invspooled%*%y2))
DMU<-max(sqrt(t(y1)%*%invspooled%*%y1),sqrt(t(y2)%*%invspooled%*%y2))

DG<-rep(10,times=P)
DG
RD<-sqrt(t(DG)%*%invspooled%*%DG)
######################################################################

######################################################################
Primjer iz trećeg poglavlja.

podaci1<-read.table("REF1.csv",sep=";")
A<-matrix(0,nrow=12,ncol=4,byrow=TRUE)

podaci2<-read.table("REF2.csv",sep=";")
B<-matrix(0,nrow=12,ncol=4,byrow=TRUE)

podaci3<-read.table("REF3.csv",sep=";")
C<-matrix(0,nrow=12,ncol=4,byrow=TRUE)

podaci4<-read.table("TEST1.csv",sep=";")
D<-matrix(0,nrow=12,ncol=4,byrow=TRUE)

for ( i in 1:12)
{
for (j in 1:4)
{
A[i,j] <-podaci1[i,j]
B[i,j] <-podaci2[i,j]
C[i,j]<-podaci3[i,j]
D[i,j]<-podaci4[i,j]
}
}
# A,B,C <- su referentne
# D je testna

# A,B i C se koriste kako bi se odredio similarity region

# for petlja koja ce izbaciti sve alfe i sve bete  prve reference A

t<-c(10,15,30,45)  # vremena
Aalfa<-numeric(12)
Abeta<-numeric(12)

for (i in 1:12)
{
x<-A[i,]
LS<-log(-log(1-x/100)) #lijeva strana
rez1=lm(LS~log(t))
#rez1$coef # koeficijenti, gdje je intercept
          # ln(alfa) ,a koeficijent ispred log(t)
          # je beta

# da bi došao od ln(alfa) do alfa trebam sve
# dignoti u eksponent od e ( prirodni logaritam)
```

```
Aalfa[i]<-exp(1)^rez1$coef[1]
Abeta[i]<-rez1$coef[2]
}
# isto tako idemo odrediti za drugu referencu B
Balfa<-numeric(12)
Bbeta<-numeric(12)

for (i in 1:12)
{
x<-B[i,]
LS<-log(-log(1-x/100)) #lijeva strana
rez1=lm(LS~log(t))
#rez1$coef
Balfa[i]<-exp(1)^rez1$coef[1]
Bbeta[i]<-rez1$coef[2]
}
# isto tako idemo odrediti za treću referencu C
Calfa<-numeric(12)
Cbeta<-numeric(12)

for (i in 1:12)
{
x<-C[i,]
LS<-log(-log(1-x/100)) #lijeva strana
rez1=lm(LS~log(t))
#rez1$coef
Calfa[i]<-exp(1)^rez1$coef[1]
Cbeta[i]<-rez1$coef[2]
}
#########################
## logaritamska transformacija parametara
#########################
LAalfa<-log(Aalfa)
LAbeta<-log(Abeta)
LBalfa<-log(Balfa)
LBbeta<-log(Bbeta)
LCalfa<-log(Calfa)
LCbeta<-log(Cbeta)
#### cilj je izračunati pooled SD=SQRT[{(var1+var2+var3)/3}+var(xpotez)]
####  za alfu i betu
Avaralfa<-var(LAalfa)
Avarbeta<-var(LAbeta)
Bvaralfa<-var(LBalfa)
Bvarbeta<-var(LBbeta)
Cvaralfa<-var(LCalfa)
Cvarbeta<-var(LCbeta)

#### sada smo izračunali var1, var2, var3 za alfu i betu
Ameanalfa<-mean(LAalfa)
Ameanbeta<-mean(LAbeta)
Bmeanalfa<-mean(LBalfa)
Bmeanbeta<-mean(LBbeta)
Cmeanalfa<-mean(LCalfa)
Cmeanbeta<-mean(LCbeta)

## varijance od xpotez za alfu i betu
alfaxpotez<-var(c(Ameanalfa,Bmeanalfa,Cmeanalfa))
```

```
betaxpotez<-var(c(Ameanbeta,Bmeanbeta,Cmeanbeta))

pooledalfa<-sqrt({(Avaralfa+Bvaralfa+Cvaralfa)/3}+alfaxpotez)
pooledbeta<-sqrt({(Avarbeta+Bvarbeta+Cvarbeta)/3}+betaxpotez)

## radimo sliku

plot(0,0,xlim=c(-2, 2),
      ylim=c(-1, 1),xlab='Razlika u Ln(Alfa)', ylab='Razlika u Ln(Beta)',
main='Regija sličnosti i Regija pouzdanosti')
lines(c(-pooledalfa,-pooledalfa),c(-pooledbeta,pooledbeta))
lines(c(-pooledalfa,pooledalfa),c(-pooledbeta,-pooledbeta))
lines(c(pooledalfa,pooledalfa),c(-pooledbeta,pooledbeta))
lines(c(-pooledalfa,pooledalfa),c(pooledbeta,pooledbeta))

## to je slika od 1STD SIMILARITY REGION
lines(c(-2*pooledalfa,-2*pooledalfa),c(-2*pooledbeta,2*pooledbeta))
lines(c(-2*pooledalfa,2*pooledalfa),c(-2*pooledbeta,-2*pooledbeta))
lines(c(2*pooledalfa,2*pooledalfa),c(-2*pooledbeta,2*pooledbeta))
lines(c(-2*pooledalfa,2*pooledalfa),c(2*pooledbeta,2*pooledbeta))

## to je slika od 2STD SIMILARITY REGION
lines(c(-3*pooledalfa,-3*pooledalfa),c(-3*pooledbeta,3*pooledbeta))
lines(c(-3*pooledalfa,3*pooledalfa),c(-3*pooledbeta,-3*pooledbeta))
lines(c(3*pooledalfa,3*pooledalfa),c(-3*pooledbeta,3*pooledbeta))
lines(c(-3*pooledalfa,3*pooledalfa),c(3*pooledbeta,3*pooledbeta))
## Poštenim putem ići bit će teško, znaj:)
## to je slika od 3STD SIMILARITY REGION

Dalfa<-numeric(12)
Dbeta<-numeric(12)
for (i in 1:12)
{
x<-D[i,]
LS<-log(-log(1-x/100)) #lijeva strana
rez1=lm(LS~log(t))
#rez1$coef
Dalfa[i]<-exp(1)^rez1$coef[1]
Dbeta[i]<-rez1$coef[2]
}

LDalfa<-log(Dalfa)
LDbeta<-log(Dbeta)
xr<-(c(mean(LCalfa),mean(LCbeta)))
xr
xt<-(c(mean(LDalfa),mean(LDbeta)))
xt

sref<-matrix(0,nrow=2,ncol=2)
sref
sref[1,1]<-cov(LCalfa,LCalfa)
sref[1,2]<-cov(LCalfa,LCbeta)
sref[2,1]<-cov(LCbeta,LCalfa)
sref[2,2]<-cov(LCbeta,LCbeta)
sref

stest<-matrix(0,nrow=2,ncol=2)
```

```
stest
stest[1,1]<-cov(LDalfa,LDalfa)
stest[1,2]<-cov(LDalfa,LDbeta)
stest[2,1]<-cov(LDbeta,LDalfa)
stest[2,2]<-cov(LDbeta,LDbeta)
stest

##### sada određujem spooled=(stest+sref)/2

spooled<-(stest+sref)/2
spooled

# solve(spooled) je inverz matrice spooled
solve(spooled)%*%spooled

t(xt-xr)%*%solve(spooled)%*%(xt-xr)

# prava mahalanobisova udaljenost
mahud<-sqrt(t(xt-xr)%*%solve(spooled)%*%(xt-xr))
mahud
x<-seq(0.3,1.2,length.out = 200)
y<-seq(-0.15,-0.40,length.out = 200)

elipsa<-(39.768*(x^2))+(1310.705*(y^2))+(27.094*x)+(397.11*y)+38.884+342.327*x*y <=2.57457
elipsa

## radimo sliku

for (i in 1:200)
{
   for (j in 1:200)
 {
elipsa<-(39.768*(x[i]^2))+(1310.705*(y[j]^2))+(27.094*x[i])+(397.11*y[j])+38.884+342.327*x[i]*y[j]
     if(elipsa<=2.57457)
      {
       points(x[i],y[j])
      }

 }

}
```

Životopis

Rođen sam 08.03.1993. u Zagrebu. Završio sam Osnovnu školu Zapruđe 2007. godine. Iste godine upisujem Gornjogradsku gimnaziju koju završavam 2011. godine. Tijekom osnovnoškolskog i srednjoškolskog obrazovanja uspješno sam sudjelovao na državnom natjecanju iz debate i državnom prvenstvu srednjih škola u malom nogometu. Nakon završetka gimnazije upisujem preddiplomski studij Matematike (smjer: nastavnički) na Prirodoslovno-matematičkom fakultetu u Zagrebu. Godine 2015. upisujem diplomski sveučilišni studij Matematička statistika. Tijekom studija držao sam demonstrature iz Linearne Algebre i predstavljao fakultet na sveučilišnim natjecanjima u futsalu.