Dr Dragoljub M. Brkić,
dipl. inž.
Tehnički opitni centar KoV,
Beograd

RASPODELA TAČKASTE OCENE DONJEG KVANTILA DVOPARAMETARSKE VEJBULOVE RASPODELE

UDC: 519.218.5:519.246

Rezime:

U radu je izveden izraz za funkciju gustine raspodele tačkaste ocene donjeg kvantila dvoparametarske Vejbulove raspodele. Predložena metoda njegovog intervalnog ocenjivanja, upoređena je sa odgovarajućom metodom datom u međunarodnom standardu IEC 61649:1997. Izvedeni analitički izrazi i predložena metoda ilustrovani su sa dva primera.

Ključne reči: dvoparametarska Vejbulova raspodela, donji kvantil, tačkasta ocena, metoda.

DISTRIBUTION OF THE LOWER QUANTILE POINT ESTIMATE IN THE TWO-PARAMETER WEIBULL DISTRIBUTION

Summary:

The expression for the distribution density function of the lower quantile point estimate in the two-parameter Weibull distribution has been derived. A method of lower quantile interval estimation has been proposed and then compared to the corresponding method given in the IEC 61649:1997 international standard. Two examples have been given to illustrate the derived analytical expression and the proposed method.

Key words: two-parameter Weibull distribution, lower quantile, point estimate, method.

Uvod

Tačkasta ocena t_p donjeg kvantila dvoparametarske Vejbulove raspodele, kao slučajna promenljiva, ima zakon raspodele koji se može izraziti funkcijom gustine raspodele $f(t_p)$. Vrednost tačkaste ocene t_p zavisi od parametara Vejbulove raspodele: parametra razmene — b i parametra oblika — k , kao i od usvojene vrednosti za red kvantila p koji predstavlja vrednost funkcije raspodele za $t=t_p$, tj. $p=F(t_p)$. Na osnovu vrednosti slučajno promenljive $t: t_1, t_2, \dots, t_N$, određuju se tačkaste

ocene parametara b i k : $\hat{b}$ i $\hat{k}$, respektivno. Kada se skup od N vrednosti za t uredi u rastućem poretku, tako da je $t_1 \leq t_2 \leq \dots \leq t_N$, tada se tačkasta ocena t_p može odrediti kao statistika, čiji je rang $[Np]$ ili kao srednja vrednost od dve statistike sa rangovima $[Np]$ i $[Np]+1$, gde $[Np]$ označava celobrojnu vrednost od Np . Iako se za p unapred usvaja vrednost, ona se može izračunati iz izraza za funkciju raspodele $F(t)$, unošenjem vrednosti za $\hat{b}$, $\hat{k}$ i t_p . Na taj način dobija se tačkasta ocena $\hat{p}$ za p .

Za tačkastu ocenu $\hat{p}$ pretpostavljena je beta raspodela u intervalu (0,1) sa parametrima A i B, čije se tačkaste ocene određuju na osnovu srednje vrednosti i standardne devijacije za $\hat{p}$. Pošto je u nizu provera na pseudoslučajnim brojevima potvrđena opravdanost navedene pretpostavke, na osnovu veze između $\hat{p}$ i $\hat{t}_p$ određena je funkcija gustine raspodele $f(\hat{t}_p)$, kojom se izražava zakon raspodele tačkaste ocene donjeg kvantila dvoparametarske Vejbulove raspodele. Pored određivanja tačkaste ocene $\hat{t}_p$ i njene funkcije gustine raspodele $f(\hat{t}_p)$, predložen je i postupak intervalne ocene za t_p , a dobijeni rezultati su upoređeni sa odgovarajućim rezultatima navedenim u međunarodnom standardu IEC 61649 [1]. Opravdanost navedene pretpostavke i primene predložene metode potvrđena je konkretnim primerima.

Upotrebljene oznake

A, B	— parametri beta raspodele,
$\hat{A}, \hat{B}$	— tačkaste ocene za A i B, respektivno,
N	— ukupan broj vrednosti slučajne promenljive (veličina populacije),
$n = \lfloor N/2 \rfloor$	— broj na slučajan način izdvojenih vrednosti iz skupa od N vrednosti slučajne promenljive (veličina potpopulacije),
N_t	— broj provera (broj ponavljanja izdvajanja potpopulacije),
t_p	— kvantil reda p dvoparametarske Vejbulove raspodele,

$\hat{t}_p$	— tačkasta ocena za t_p ,
b, k	— parametar razmere, oblika Vejbulove raspodele, respektivno,
$\hat{b}, \hat{k}$	— tačkaste ocene za b i k, respektivno,
$f(\hat{t}_p)$	— funkcija gustine raspodele tačkaste ocene $\hat{t}_p$,
$p = F(t_p)$	— funkcija raspodele za $t = t_p$,
$\hat{p} = F(\hat{t}_p)$	— tačkasta ocena za p,
$f(\hat{p})$	— funkcija gustine raspodele za $\hat{p}$,
$[p_1, p_2]$	— interval poverenja za p,
$[t_{p1}, t_{p2}]$	— interval poverenja za t_p ,
$\bar{p}, \bar{s}_p$	— srednja vrednost i varijansa za $\hat{p}$, respektivno,
B(A, B)	— beta funkcija sa parametrima A i B,
$x_{v1;v2;\gamma}$	— gornji kvantil beta raspodele; v_1, v_2 — parametri raspodele, γ — gornji kvant (rizik),
α_1, α_2	— donji, gornji rizik,
$\alpha = \alpha_1 + \alpha_2$	— ukupan rizik,
$P = 1 - \alpha$	— poverenje,
$\lfloor Q \rfloor$	— celobrojna vrednost od Q.

Teorijske postavke

Donji kvantil dvoparametarske Vejbulove raspodele definisan je izrazom:

$$p = F(t_p) = 1 - \exp \left[- \left(\frac{t_p}{b} \right)^k \right]; \quad (1)$$

$b; k > 0, t_p \geq 0$

gde je:

b — parametar razmere,

k — parametar oblika,
 t_p — donji kvantil reda p .

Neka je u toku ispitivanja slučajna promenljiva t , koja ima dvoparametarsku Vejbulovu raspodelu sa parametrima b i k , uzela sledeće vrednosti: $t_1, t_2, \dots, t_N$. Na osnovu njih, primenom neke od poznatih metoda za statističko ocenjivanje parametara raspodela, na primer metode maksimalne verodostojnosti, određuju se tačkaste ocene parametara b i k : $\hat{b}$ i $\hat{k}$.

Zatim se iz skupa vrednosti $\{t_1, t_2, \dots, t_N\}$ izdvoji na slučajan način podskup od n vrednosti: $\{t_1, t_2, \dots, t_n\}$. Pogodno je da se usvoji da je $n = \lfloor N/2 \rfloor$, tj. celobrojna vrednost od $N/2$. Na osnovu vrednosti za t iz tog podskupa određuje se tačkasta ocena kvantila $t_p = \hat{t}_p$. Određivanje tačkaste ocene $\hat{t}_p$ obavlja se na sledeći način:

vrednosti za t u podskupu uredi se u rastućem poretку: $t_1 \leq t_2 \leq \dots \leq t_n$, a zatim se odrede susedne vrednosti čiji su redni brojevi $[np]$ i $[np] + 1$, tj. $t_{[np]}$ i $t_{[np] + 1}$, gde je p usvojena vrednost reda donjeg kvantila. Na primer, ako je usvojena vrednost reda kvantila $p = 0,10$ ili (10%), a broj vrednosti u podskupu $n = 50$, tada su $[50 \times 0,10] = 5$ i $[50 \times 0,10] + 1 = 6$, tj. tražene dve susedne vrednosti su t_5 i t_6 . Aritmetička sredina te dve susedne vrednosti za t može se definisati kao tačkasta ocena kvantila t_p , tj.:

$$\hat{t}_p = \frac{1}{2} (t_{[np]} + t_{[np] + 1}) \quad (2)$$

Ako se u izrazu (1), b , k i t_p zameni tačkastim ocenama $\hat{b}$, $\hat{k}$ i $\hat{t}_p$, respektivno, dobiće se tačkasta ocena za p u obliku:

$$\hat{p} = 1 - \exp \left[- \left(\frac{\hat{t}_p}{\hat{b}} \right)^{\hat{k}} \right] \quad (3)$$

Navedeni postupak određivanja tačkaste ocene $\hat{p}$ ponavlja se N_t puta ($N_t \geq 50$), tako da se dobije skup vrednosti za tačkastu ocenu p : $\{\hat{p}_1, \hat{p}_2, \dots, \hat{p}_{N_t}\}$.

Tačkasta ocena $\hat{p}$, kao slučajna promenljiva, ima odgovarajući zakon raspodele. Pretpostavimo da je to beta raspodela sa parametrima A i B , čija je funkcija gustine raspodele data sledećim izrazom:

$$f(\hat{p}) = \frac{1}{B(A, B)} \hat{p}^{A-1} (1 - \hat{p})^{B-1}; \quad \hat{p} \in [0, 1] \quad (4)$$

Da bi se odredili parametri A i B potrebno je iz skupa $\{\hat{p}_1, \hat{p}_2, \dots, \hat{p}_{N_t}\}$ odrediti srednju vrednost i standardnu devijaciju za $\hat{p}$: $\bar{\hat{p}}$, $s_{\hat{p}}^2$. Na osnovu veze između parametara A , B i $\bar{\hat{p}}$, $s_{\hat{p}}^2$ određuju se tačkaste ocene za A i B , pomoću sledećih izraza:

$$\hat{A} = \left(\frac{\bar{\hat{p}} - \bar{\hat{p}}^2}{s_{\hat{p}}^2} - 1 \right) \bar{\hat{p}} \quad (5)$$

$$\hat{B} = \left(\frac{\bar{\hat{p}} - \bar{\hat{p}}^2}{s_{\hat{p}}^2} - 1 \right) (1 - \bar{\hat{p}}) \quad (6)$$

Zamenom parametara A i B u izrazu (4) njihovim tačkastim ocenama $\hat{A}$ i $\hat{B}$, respektivno, funkcija $f(\hat{p})$, data izrazom (4), postaje potpuno određena.

Opravdanost pretpostavke da $\hat{p}$ ima beta raspodelu sa parametrima $\hat{A}$

$\hat{B}$, može se dokazati proverama na pseudoslučajnim brojevima koji imaju dvoparametarsku Vejbulovu raspodelu sa poznatim vrednostima parametara b i k . Opravdanost navedene pretpostavke potvrđena je u primerima.

Određivanje funkcije gustine raspodele tačkaste ocene

Na osnovu veze između t_p i p , date izrazom (1), može se odrediti funkcija gustine raspodele tačkaste ocene $f(\hat{t}_p)$ donjeg kvantila $\hat{t}_p$, polazeći od sledeće jednakosti:

$$f(t_p)dt_p = f(p)dp \quad (7)$$

Radi jednostavnijeg označavanja, izostavljena je oznaka, $\hat{\cdot}$, za tačkaste ocene. Na osnovu izraza (7) dobija se:

$$f(t_p) = f(p) \frac{dp}{dt_p} \quad (8)$$

Diferenciranjem izraza (1) po t_p , dobija se:

$$\frac{dp}{dt_p} = -\frac{k}{b} \left(\frac{t_p}{b} \right)^{k-1} e^{-\left(\frac{t_p}{b} \right)^k} \quad (9)$$

Zamenom $f(t_p)$ i dp/dt_p iz izraza (4) i (9), respektivno, u izraz (8) dobija se:

$$f(t_p) = \frac{k}{bB(A, B)} \left[1 - e^{-\left(\frac{t_p}{b} \right)^k} \right]^{A-1} \left[e^{-\left(\frac{t_p}{b} \right)^k} \right]^{B-1} \left(\frac{t_p}{b} \right)^{k-1} e^{-\left(\frac{t_p}{b} \right)^k} \quad (10)$$

gde je:

$\hat{p} = p$ u izrazu (4) zamenjeno izrazom (1).

Izraz (10) predstavlja funkciju gustine raspodele tačkaste ocene donjeg kvantila $t_p = \hat{t}_p$ dvoparametarske Vejbulove raspodele. U izrazu (10), parametri b , k , A i B se zamenjuju njihovim tačkastim ocenama $\hat{b}$, $\hat{k}$, $\hat{A}$ i $\hat{B}$, respektivno.

Određivanje intervalne ocene za tačkastu ocenu

Rešavanjem izraza (1) po t_p , dobija se:

$$t_p = b[-\ln(1-p)]^{1/k} \quad (11)$$

Ako se u izrazu (11) b i k zamene tačkastim ocenama $\hat{b}$ i $\hat{k}$, respektivno, a p sa granicama poverenja p_1 i p_2 , dobiće se granice poverenja za t_p :

$$t_{p1} = \hat{b}[-\ln(1-p_1)]^{1/\hat{k}} \quad (12)$$

$$t_{p2} = \hat{b}[-\ln(1-p_2)]^{1/\hat{k}} \quad (13)$$

gde su granice poverenja p_1 i p_2 date sledećim izrazima:

$$p_1 = 1 - X_{n-r; r+1; \alpha_1} \quad (14)$$

$$p_2 = X_{r+1; n-r; \alpha_2} \quad (15)$$

U izrazima (14) i (15), $x_{[.]}$ je gornji kvantil beta raspodele, n je ukupan broj vrednosti za $\hat{p}$, $r = [n\hat{p}]$, α_1 i α_2 su donji i gornji rizik, respektivno.

Primeri

Primer 1

Primenom elektronskog računara generisano je $N=100$ pseudoslučajnih brojeva koji imaju dvoparametar-

sku Vejbulovu raspodelu sa parametrima $b=1000$ i $k=2,5$. Na osnovu tih brojeva primenom metode maksimalne verodostojnosti, određene su tačkaste ocene parametara b i k : $\hat{b}=1008,78$ i $\hat{k}=2,367$.

Iz skupa od $N=100$ vrednosti pseudoslučajnih brojeva izdvajano je, na slučajan način, $n=50$ vrednosti ovih brojeva i pomoću izraza (2) određivana je tačkasta ocena donjeg kvantila, $\hat{t}_p$, (za $p=0,10$). Postupak određivanja $\hat{t}_p$ ponavljan je $N_t=50$ puta i pomoću izraza (3) određeno je 50 vrednosti za $\hat{p}$. Pod pretpostavkom da $\hat{p}$ ima beta raspodelu, na osnovu tih 50 vrednosti za $\hat{p}$ i pomoću izraza (5) i (6) određene su tačkaste ocene za parametre A i B beta raspodele: $\hat{A}=5,67$ i $\hat{B}=47,87$. Na slici 1 prikazan je histogram i funkcija gustine raspodele $f(\hat{p})$. Grafički rezultati proračuna dobijeni su posredstvom elektronskog računara. Kao što se vidi na slici 1, podudarnost histograma i funkcije $f(\hat{p})$ je dobra, što ukazuje na opravdanost pretpostavke da $\hat{p}$ ima beta raspodelu sa parametrom A i B . Na slici 2 grafički je prikazana funkcija gustine raspodele $f(\hat{t}_p)$ i odgovarajući histogram za $\hat{t}_p$. I u ovom slučaju podudarnost funkcije $f(\hat{t}_p)$ i odgovarajućeg histograma za $\hat{t}_p$ je zadovoljavajuća, što ukazuje na to da je korektno izveden izraz (10) za funkciju gustine raspodele $f(\hat{t}_p)$. Na slici 1 prikazane su granice varijacije za $\hat{p}$: $p_1=0,0564$ i $p_2=0,1619$ sa rizicima $\alpha_1=\alpha_2=0,10$, a na slici 2 gra-

nice varijacije za $\hat{t}_p$: $t_{p1}=303,5$ i $t_{p2}=485,5$ sa rizicima $\alpha_1=\alpha_2=0,10$.

Ako se u izrazima (12) i (13) unesu vrednosti za p_1 i p_2 , dobija se $t_{p1}=303,06$ i $t_{p2}=484,93$, što ukazuje na veliku podudarnost sa vrednostima t_{p1} i t_{p2} , prikazanim na slici 2. Za $p=0,10$, $\hat{b}=1008,78$ i $\hat{k}=2,367$, pomoću izraza (11), dobija se tačkasta ocena donjeg kvantila $\hat{t}_p : \hat{t}_p=389,85$. Ta vrednost približno je jednaka srednjoj vrednosti $\hat{t}_p=393,31$ prikazanoj na slici 2.

Granice poverenja za donji kvantil t_p mogu se odrediti pomoću izraza (12) i (13), koristeći granice poverenja p_1 i p_2 date izrazima (14) i (15). Za $p=0,10$, $n=50$, $r=np=5$ i $\alpha_1=\alpha_2=0,10$ određene su vrednosti za p_1 i p_2 : $p_1=0,064260$ i $p_2=0,177618$. Unoseći vrednosti za p_1 i p_2 u izraze (12) i (13) dobijaju se granice poverenja: $t_{p1}^*=320,8$ i $t_{p2}^*=506,25$. Dakle, sa poverenjem $P=1-(\alpha_1+\alpha_2)=0,80$, može se tvrditi da interval poverenja $[320,8; 506,25]$ prekriva stvarnu vrednost donjeg kvantila t_p ($p=0,10$).

Primer 2

Radi ispitivanja pouzdanosti, iz jedne proizvodne partije izdvojeno je metodom slučajnog uzorka, $N=40$ uređaja. Ispitivanje pouzdanosti obavljeno je bez zamene i popravke, a trajalo je do dvadesetog otkaza. Vremena do pojave stanja »u otkazu« sistematizovana u rastućem poretku, jesu:

$t_1=5$, $t_2=10$, $t_3=17$, $t_4=32$, $t_5=32$, $t_6=33$, $t_7=34$, $t_8=36$, $t_9=54$, $t_{10}=55$, $t_{11}=55$, $t_{12}=58$, $t_{13}=58$, $t_{14}=61$, $t_{15}=64$, $t_{16}=65$, $t_{17}=65$, $t_{18}=66$, $t_{19}=67$, $t_{20}=68$.

Sl. 1 — Grafički prikaz histograma i funkcije gustine raspodele $f(p)$

Sl. 2 — Grafički prikaz histograma i funkcije gustine raspodele $f(p)$

Za uređaje koji nisu otkazali, vremena do prekida ispitivanja su:

$$t_{21} = t_{22} = \dots = t_{39} = t_{40} = 68.$$

Pod pretpostavkom da vremena do pojave stanja »u otkazu« imaju dvo-

parametarsku Weibulovu raspodelu sa parametrima b i k , na osnovu datih podataka određene su tačkaste ocene tih parametara primenom metode maksimalne verodostojnosti: $\hat{b} = 84$ i $\hat{k} = 2,091$ [1].

Istim postupkom, kao i u *Primeru 1*, iz skupa od $N=40$ vrednosti vremena do pojave stanja »u otkazu«, odnosno do prekida ispitivanja, izdvajano je na slučajan način $n=Np=20$ vrednosti i izračunavane su tačkaste ocene donjeg kvantila $\hat{t}_p$, ($p=0,10$), odnosno vrednosti za $\hat{p}$. Postupak je ponavljan $N_t=50$ puta. Pod pretpostavkom da tačkasta ocena $\hat{p}$ ima beta raspodelu, određene su tačkaste ocene ove raspodele: $\hat{A}=1,94$ i $\hat{B}=12,05$.

Na slici 3 prikazan je histogram i funkcija gustine raspodele $f(\hat{p})$, a na slici 4 funkcija gustine raspodele $f(\hat{t}_p)$ i odgovarajući histogram za $\hat{t}_p$. Kao što se može videti na slici 3, dobijene granice varijacije za $\hat{p}$ su: $p_1=0,0391$ i $p_2=0,2619$ sa rizicima $\alpha_1=\alpha_2=0,10$. Na slici 4 prikazane su granice varijacije za $\hat{t}_p$: $t_{p1}=17,99$ i $t_{p2}=47,53$ sa rizicima $\alpha_1=\alpha_2=0,10$.

Kada se u izrazima (12) i (13) uvrste vrednosti za p_1 i p_2 , dobija se $t_{p1}=17,99$ i $t_{p2}=47,5$, što ukazuje na veliku podudarnost sa vrednostima t_{p1} i t_{p2} prikazanim na slici 4. Za $p=0,10$, $\hat{b}=84$ i $\hat{k}=2,091$, pomoću izraza (11), dobija se tačkasta ocena donjeg kvantila t_p : $\hat{t}_p=28,63$.

Granice poverenja za donji kvantil t_p mogu se odrediti pomoću izraza (12) i (13), koristeći granice poverenja p_1 i p_2 , date izrazima (14) i (15). Za $p=0,10$, $n=20$, $r=np=2$ i $\alpha_1=\alpha_2=0,10$ određene su vrednosti za p_1 i p_2 : $p_1=0,056418$ i $p_2=0,244765$. Unoseći ove vrednosti za p_1 i p_2 u izraze (12) i (13) dobijaju se granice poverenja: $t_{p1}^*=21,53$ i $t_{p2}^*=45,75$.

Ako se donja granica poverenja $t_{p1}^*=21,53$, dobijena ovim postupkom, uporedi sa donjom granicom poverenja $B_{10}=20,43$, dobijenom navedenom metodom u međunarodnom standardu IEC 61649:97 (Primer primene metode),

Sl. 3 — Grafički prikaz histograma i funkcije gustine raspodele $f(\hat{p})$

Sl. 4 — Grafički prikaz histograma i funkcije gustine raspodele $f(p)^{\wedge}$

može se zaključiti da je slaganje rezultata dobro. Međutim, prednost metode predložene u ovom radu jeste što se za p može uzeti bilo koja vrednost iz intervala $(0,1)$, i što se može odrediti gornja granica za t_p , dok se postojećom metodom u navedenom međunarodnom standardu određuje samo donja granica poverenja za t_p i to samo za vrednost reda kvantila $p=0,10$.

Zaključak

Valjanost izvedenog izraza za funkciju gustine raspodele, $f(t_p)^{\wedge}$, tačkaste ocene donjeg kvantila dvoparametar-ske Vejbulove raspodele proveravana

je na pseudoslučajnim brojevima, koji imaju dvoparametarsku Vejbulovu raspodelu, a rezultati provere bili su zadovoljavajući, što je potvrđeno navedenim primerima.

Predložena metoda intervalnog ocenjivanja donjeg kvantila t_p upoređena je sa odgovarajućom metodom datom u međunarodnom standardu IEC 61649:1997. Dobijeni su približni rezultati, što ukazuje na opravdanost primenjivosti predložene metode. Ona omogućava određivanje donje i gornje granice poverenja za donji kvantil t_p za bilo koju vrednost za p ($0 < p < 1$), za razliku od metode po međunarodnom standardu IEC 61649 koja obezbeđuje određivanje samo donje granice poverenja, B_{10} , za t_p i to jedino za $p=0,10$ (10%).

Literatura:

- [1] International standard IEC 61649 — Goodness-of-fit tests, confidence intervals and lower confidence limits for Weibull distributed data, First edition, 1997-05.
- [2] Chapouille, P., De Pazzis, R.: Fiabilité des Systèmes, Masson, Paris, 1968.

- [3] Van Der Waerden, B. L.: Mathematische Statistik, Springer-Verlag, Berlin, 1965.
- [4] Brkić, D. M.: A Method for Evaluation of Number Class Intervals of Histogram, Microelectronics and Reliability, Vol. 32, No 2/3, pp 245—248, 1991.