Sveučilište J.J. Strossmayera u Osijeku
Odjel za matematiku
Financijska matematika i statistika

**Ana Bulić**

# Procjena parametra Weibullove distribucije

Diplomski rad

Osijek, 2016.

Sveučilište J.J. Strossmayera u Osijeku
Odjel za matematiku
Financijska matematika i statistika

**Ana Bulić**

# Procjena parametara Weibullove distribucije

Diplomski rad

Mentor: prof. dr. sc. Mirta Benšić

Lektorirala: Danijela Pauković

Osijek, 2016.

# Sadržaj

# 1 Uvod

U ovom diplomskom radu bavit ćemo se Weibullovom distribucijom i procjenom parametara Weibullove distribucije. U prvom dijelu rada definirat ćemo kada slučajna varijabla ima Weibullovu distribuciju te navesti njezinu funkciju gustoće, funkciju distribucije, varijancu, očekivanje itd. Weibullova distribucija najčešće se koristi u teoriji pouzdanosti i teoriji životnog vijeka. Sukladno tome navest ćemo neke bitne funkcije vezane uz teoriju pouzdanosti kao što su funkcija rizika, funkcija otkaza i funkcija doživljenja.

Weibullova distribucija za neke vrijednosti svojih parametara može se prikazati kao neka druga distribucija, odnosno povezana je s nekim drugim distribucijama kao što je npr. eksponencijalna pa ćemo nešto više reći o povezanim distribucijama.

Glavni dio ovog diplomskog rada je procjena parametara Weibullove distribucije. Iako postoje mnoge metode za procjenu parametara mi ćemo se bazirati na najpoznatije metode i njihove modifikacije. Dakle, promatrat ćemo procjenu parametara metodom momenata, modificiranom metodom momenata zatim metodom maksimalne vjerodostojnosti, modificiranom metodom maksimalne vjerodostojnosti te na kraju procjenu parametara metodom kvantila ili tzv. minimalnom udaljenosti kvantila.

Na samom kraju rada kroz simulacije ćemo uspoređivati asimptotska svojstva procjenitelja.

# 2 Povijesni podatci

Weibullova distribucija nazvana je po švedskom fizičaru Waloddi Weibullu.
Waloddi Weibull rođen je 18. lipnja 1887. godine a umro je 12. listopada 1979. godine u Francuskoj.

Weibull distribuciju prvi put spominje u radu *A statistical theory of the strength of material* koji je objavljen 1939. godine. Weibullova distribucija je jedna od najpoznatijih distribucija koja se koristi u teoriji pouzdanosti i teoriji životnog vijeka.

Prije Weibulla sličnu distribuciju koristili su Rosin i Rammler (vidi [7]) za opisivanje distribucije veličine čestica, te se distribucija često naziva Rosin-Rammlerova distribucija. Najraniji poznati rad u kojem se javlja Weibullova distribucija jest rad Fishera i Tippeta (vidi [8]) iz 1928. u kojem je distribucija dobivena kao granična distribucija malih ekstrema u uzorku.

U području farmakologije Weibullova distribucija pojavljuje se pod nazivom Rosin-Rammler-Sperling-Bennet-Weibullova distribucija ili kratko RRSBW distribucija, iako se ponekad izostavlja Weibullovo ime.

U ruskoj literaturi ova distribucija se pojavljuje pod nazivom Weibull-Gnedenko distribucija, budući da je jedna od tri tipa graničnih distribucija po Gnedenku. Isto tako ponekad nosi i ime Frechet distribucija, po Frechetu koji je prvi put identificirao tu distribuciju kao graničnu.

Weibullova distribucija koristi se u tzv. teoriji životnog vijeka (lifetime studies) kao matematički model za opisivanje slučajnog vijeka trajanja nekog elementa (žarulje, otpornika, kondenzatora, itd.), u teoriji pouzdanosti (reliability theory) a isto tako ima široku upotrebu u biologiji, kemiji, medicini (modeliranje vremena do pojave tumora kod ljudi), farmakologiji, šumarstvu i injženjerskim istraživanjima.

U novije vrijeme Weibullova distribucija koristi se za modeliranje čvrstoće materijala, napuknuća u betonu, učestalosti poplava i potresa, distribuciju brzine vjetra, veličinu kapljica i slično. Kao što vidimo ova distribucija ima veliku primjenu.

# 3 Weibullova distribucija

## 3.1 Definicija i osnovni pojmovi

U ovom diplomskom radu najveća pažnja je posvećena procjeni parametara Weibullove distribucije. Da bismo mogli procijeniti parametre prvo moramo definirati kada neka slučajna varijabla ima Weibullovu distribuciju.

Slučajna varijabla $X$ ima troparametarsku Weibullovu distribuciju s parametrima $c > 0$, $\alpha > 0$,$\xi_0 > 0$ ako slučajna varijabla $Y$ :

$$Y = \left(\frac{X - \xi_0}{\alpha}\right)^c \tag{3.1}$$

ima standardnu eksponencijalnu distribuciju s funkcijom gustoće:

$$f_Y(y) = e^{-y}, y > 0 \tag{3.2}$$

i funkcijom distribucije:

$$F_Y(y) = 1 - e^{-y}, y > 0. \tag{3.3}$$

Funkcija distribucije troparametarske Weibullove slučajne varijable $X$ definirana je s:

$$F_X(x) = 1 - e^{-\left(\frac{x-\xi_0}{\alpha}\right)^c}, x > \xi_0. \tag{3.4}$$

Funkciju gustoće troparametarske Weibullove slučajne varijable $X$ dobijemo na sljedeći način:

$$\begin{aligned} f_X(x) &= \frac{dF_X(x)}{dx} \\ &= -e^{-\left(\frac{x-\xi_0}{\alpha}\right)^c}(-c)\left(\frac{x-\xi_0}{\alpha}\right)^{(c-1)}\frac{1}{\alpha} \\ &= \left(\frac{c}{\alpha}\right)\left(\frac{x-\xi_0}{\alpha}\right)^{(c-1)} e^{-\left(\frac{x-\xi_0}{\alpha}\right)^c}. \end{aligned}$$

Funkcija gustoće slučajne varijable $X$ koja ima troparametarsku Weibullovu distribuciju definirana je sa:

$$f_X(x) = \left(\frac{c}{\alpha}\right)\left(\frac{x-\xi_0}{\alpha}\right)^{c-1} e^{-\left(\frac{x-\xi_0}{\alpha}\right)^c}, x > \xi_0 \tag{3.5}$$

gdje se parametar $\xi_0$ naziva **parametar položaja** (engl. the location parameter), parametar $\alpha$ **parametar skaliranja** (engl. the scale parameter) i parametar $c$ **parametar oblika** (engl. the shape parameter).

Oznaka $X \sim W(c, \xi_0, \alpha)$.

U slučaju kada parametar $\xi_0$ iznosi 0, dobivamo dvoparametarsku Weibullovu distribuciju s funkcijom gustoće:

$$f_X(x) = \frac{c}{\alpha^c} x^{c-1} e^{-\left(\frac{x}{\alpha}\right)^c}, x > 0, \alpha > 0, c > 0. \tag{3.6}$$

Za Weibullovu distribuciju kažemo da je standardna Weibullova distribucija ili jednoparametarska Weibullova distribucija ako su vrijednosti parametara $\xi_0 = 0$ i $\alpha = 1$.

Pripadna funkcija gustoće standardne Weibullove distribucije definirana je s:

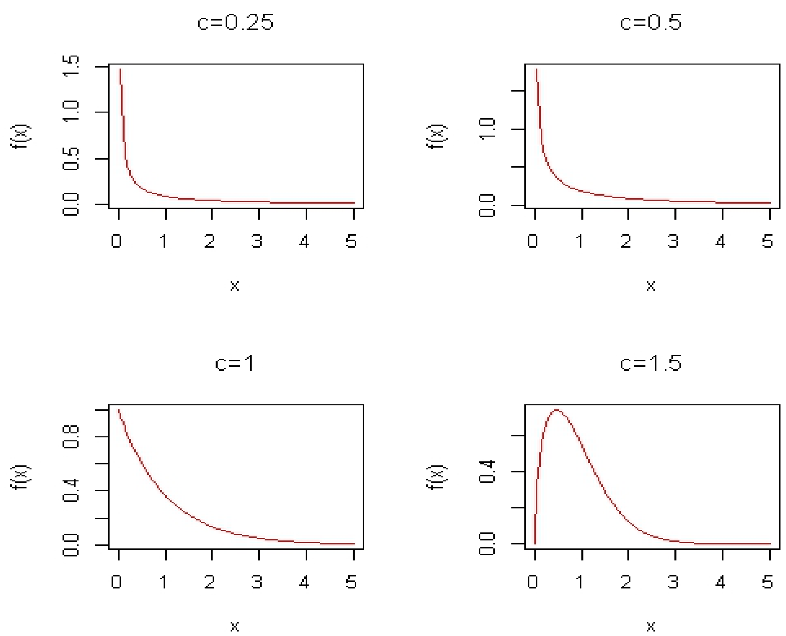
$$f_X(x) = cx^{c-1} e^{-x^c}, x > 0, c > 0. \tag{3.7}$$

Pripadna funkcija distribucije iznosi:

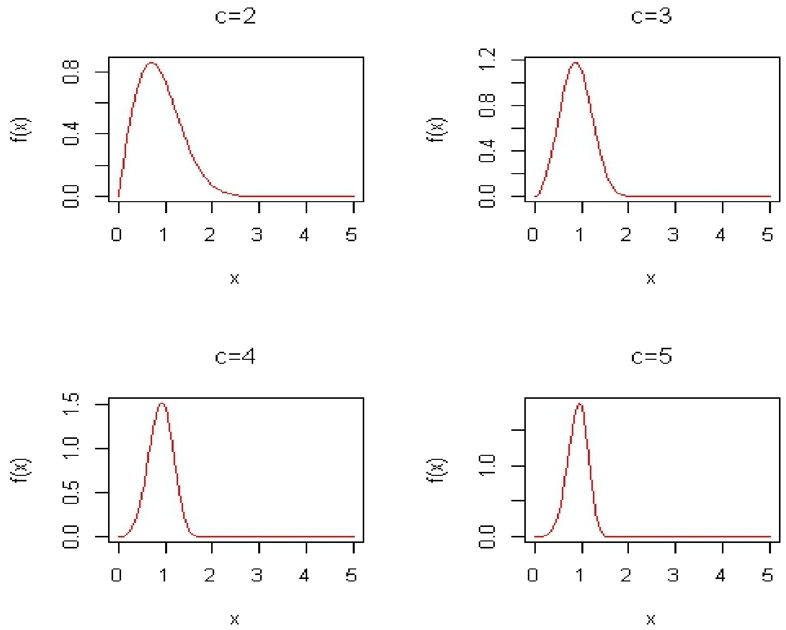
$$F_X(x) = 1 - e^{-x^c}, x > 0, c > 0. \tag{3.8}$$

U tom slučaju distribucija slučajne varijable $X$ ovisi samo o parametru $c$.

Na slici 1. i slici 2. prikazana je funkcija gustoće standardne Weibullove slučajne varijable $X$ s obzirom na različite vrijednosti parametra $c$.




Slika 1: Funkcija gustoće standardne Weibullove distribucije




Slika 2: Funkcija gustoće standardne Weibullove distribucije

Momenti odgovarajuće troparametarske funkcije gustoće u (3.5) lako se izvedu iz odgovarajuće standardne funkcije gustoće (3.7) koristeći transformaciju $X^{'} = \xi_0 + \alpha X$.

Iz (3.1) vrijedi:

$$Y = \left(\frac{X^{'} - \xi_0}{\alpha}\right)^c = \left(\frac{\xi_0 + \alpha X - \xi_0}{\alpha}\right)^c = X^c, \tag{3.9}$$

tj. $X^c = Y$, odnosno $X^c$ ima standarnu eksponencijalnu distribuciju jer slučajna varijabla $Y$ ima standardnu eksponencijalnu distribuciju. Kako je $X^c = Y$ slijedi:

$$(X^c)^{\frac{r}{c}} = Y^{\frac{r}{c}} = X^r. \tag{3.10}$$

Iz toga slijedi da je $\frac{r}{c}$-ti moment od $Y$ ujedno i r- ti moment od $X$ i on iznosi:

$$E(X^r) = \Gamma\left(\frac{r}{c} + 1\right) \tag{3.11}$$

iz čega dobijemo da je očekivanje slučajne varijable $X$ definirano s:

$$E(X) = \Gamma\left(\frac{1}{c} + 1\right) \tag{3.12}$$

gdje je $\Gamma$ zadana s:

$$\Gamma(c) = \int_0^\infty x^{c-1} e^{-x} dx. \tag{3.13}$$

Varijanca slučajne varijable $X$ definirana je s:

$$Var(X) = \Gamma\left(\frac{2}{c} + 1\right) - \left\{\Gamma\left(\frac{1}{c} + 1\right)\right\}^2. \tag{3.14}$$

Generalno, ako je $X$ troparametarska Weibullova distribucija očekivanje je dano sa:

$$E(X) = \xi_0 + \alpha \left[ \Gamma \left( \frac{1}{c} + 1 \right) \right], \tag{3.15}$$

a varijanca iznosi:

$$Var(X) = \alpha^2 \left\{ \Gamma \left( \frac{2}{c} + 1 \right) - \Gamma \left[ ( \frac{1}{c} + 1 \right) \right]^2 \right\}. \tag{3.16}$$

## 3.2 Teorija pouzdanosti

Da bismo što bolje mogli razumjeti mogućnost primjene Weibullove distribucije potrebno je uvesti neke pojmove iz teorije pouzdanosti. Pouzdanost nekog sustava je vjerojatnost da će taj sustav uspješno, bez "otkaza", obaviti zadaću koja mu je namijenjena.

Neka je $X$ slučajna varijabla koja modelira vrijeme pojave otkaza, čija je funkcija gustoće $f_X(x)$ i funkcija distribucije $F_X(x)$.

U teoriji pouzdanosti funkcija $F$ zove se **funkcija distribucije otkaza** (engl. failure distribution ili life distribution). Dakle, u teoriji pouzdanosti

$$F_X(x) = P(X \leq x) \tag{3.17}$$

predstavlja vjerojatnost da će sustav otkazati do trenutka x.

**Funkcija pouzdanosti ili doživljenja** (engl. reliability function ili survivol function) $S_X(x)$ definira se kao vjerojatnost da sustav neće otkazati do vremenskog trenutka x. Funkcija pouzdanosti ili doživljenja definirana je s:

$$S_X(x) = 1 - F_X(x) = P(X > x). \tag{3.18}$$

Ako je $X$ troparametarska Weibullova distribucija onda je:

$$S_X(x) = e^{-\left(\frac{X-\xi_0}{\alpha}\right)^c}, x > \xi_0. \tag{3.19}$$

Osim funkcije doživljenja, uz Weibullovu distribuciju veže se i funkcija hazarda ili rizika.

Funkcija hazarda ili funkcija rizika (engl. hazard function ili instantaneous failure rate) $h_X(x)$ definirana je s:

$$h_X(x) = \lim_{\Delta x \to 0} \frac{P(x < X \leq x + \Delta x |\ X > x)}{\Delta x} \tag{3.20}$$

gdje $P(x < X \leq x + \Delta x |\ X > x)$ označava vjerojatnost da slučajna varijabla X poprimi vrijednost iz intervala $(x, x + \Delta x]$ ako je njezina vrijednost veća od x.

Kako je

$$P(x < X \leq x + \Delta x |\ X > x) = \frac{P(x < X \leq x + \Delta x)}{P(X > x)} = \frac{F_X(x + \Delta x) - F_X(x)}{S_X(x)} \tag{3.21}$$

slijedi

$$h_X(x) = \frac{f_X(x)}{S_X(x)} = \frac{f_X(x)}{1 - F_X(x)}. \tag{3.22}$$

Za troparametarsku Weibullovu distribuciju funkcija hazarda ili rizika definirana je na sljedeći način:

$$h_X(x) = \left(\frac{c}{\alpha}\right)\left(\frac{x - \xi_0}{\alpha}\right)^{c-1}, x > \xi_0. \tag{3.23}$$

Možemo uočiti da je funkcija hazarda ili rizika troparametarske Weibullove distribucije opadajuća ako je parametar $c < 1$, konstantna ako je parametar $c = 1$ i rastuća ako je vrijednost parametra $c > 1$.

Osim hazard funkcije ili funkcije rizika spomenimo još i kumulativnu funkciju hazarda ili rizika.
Kumulativna funkcija rizika (engl. cumulative hazard rate) $H(t)$ definirana je izrazom:

$$H_X(x) = \int_0^x h(t)dt. \tag{3.24}$$

Kumulativna funkcija rizika za troparametarsku Weibullovu distribuciju definirana je s:

$$H_X(x) = \left(\frac{x - \xi_0}{\alpha}\right)^c \tag{3.25}$$

## 3.3 Uređajna statistika

Neka su $X_1$,$X_2$,...,$X_n$ nezavisne jednako distribuirane troparametarske Weibullove slučajne varijable s funkcijom gustoće (3.5) i funkcijom distribucije (3.4). Njihovim preuređenjem u rastućem poretku $X_{(1)} \leq X_{(2)} \leq X_{(3)} \leq ... \leq X_{(n)}$ dobivamo tzv. uređajne statistike.

Funkcija gustoće uređajne statistike $X_{(r)}$, gdje je $1 \leq r \leq n$, definirana je s:

$$f_r(x) = \frac{n!}{(r-1)!(n-1)!} \left(F_X(x)\right)^{r-1} \left(1 - F_X(x)\right)^{n-r} f_X(x). \tag{3.26}$$

Koristeći formulu (3.26) za vrijednost $r = 1$ funkcija gustoće najmanje uređene statistike $X_{(1)}$ definirana je s:

$$f_{X_{(1)}}(x) = n\left[1 - F_X(x)\right]^{n-1} p_X(x) = \frac{nc}{\alpha}\left(\frac{x-\xi_0}{\alpha}\right)^{c-1} e^{-n\left(\frac{x-\xi_0}{\alpha}\right)^c}, x > \xi_0. \tag{3.27}$$

Iz formule (3.27) jasno je da slučajna varijabla $X_{(1)}$ isto ima troparametarsku Weibullovu distribuciju, ako $\alpha$ u (3.5) zamijenimo sa $\alpha n^{\frac{-1}{c}}$.

Neka je $X$ troparametarska Weibullova slučajna varijabla. Koristeći linearnu transformaciju $X = \xi_0 - \alpha X_{(r)}$ funkcija gustoće slučajne varijable $X_{(r)}$, gdje je $1 \leq r \leq n$, iznosi:

$$\begin{aligned}
f_{X_{(r)}}(x) &= \frac{n!}{(r-1)!(n-1)!} \left(F_X(x)\right)^{r-1} \left(1 - F_X(x)\right)^{n-r} f_X(x) \\
&= \frac{n!}{(r-1)!(n-1)!} \left(1 - e^{-\left(\frac{x-\xi_0}{\alpha}\right)^c}\right)^{r-1} \left(1 - 1 - e^{-\left(\frac{x-\xi_0}{\alpha}\right)^c}\right)^{n-r} f_X(x) \\
&= \frac{n!}{(r-1)!(n-1)!} \left(1 - e^{-\left(\frac{\xi_0 - \alpha X_{(r)} - \xi_0}{\alpha}\right)^c}\right)^{r-1} \left(1 - 1 - e^{-\left(\frac{\xi_0 - \alpha X_{(r)} - \xi_0}{\alpha}\right)^c}\right)^{n-r} f_X(x) \\
&= \frac{n!}{(r-1)!(n-1)!} \left(1 - e^{-(x)^c}\right)^{r-1} \left(1 - 1 - e^{-(x)^c}\right)^{n-r} f_X(x) \\
&= \frac{n!}{(r-1)!(n-1)!} \left(1 - e^{-(x)^c}\right)^{r-1} \left(1 - 1 - e^{-(x)^c}\right)^{n-r} cx^{c-1} e^{-x^c} \\
&= \frac{n!}{(r-1)!(n-r)!} \left(1 - e^{-x^c}\right)^{r-1} e^{-x^c(n-r+1)} cx^{c-1}.
\end{aligned} \tag{3.28}$$

Iz funkcije gustoće slučajne varijable $X_{(r)}$ i linearne transformacije $\xi_0 + \alpha X_{(r)}$ dobijemo k-ti moment od $X_{(r)}$ koji iznosi:

$$\begin{aligned} E\left[\left(X_{(r)}\right)^k\right] &= \frac{n!}{(r-1)!(n-r)!} \int_0^\infty x^k \left(1 - e^{-x^c}\right)^{r-1} e^{-x^c(n-r+1)} c x^{c-1} dx \\ &= \frac{n!}{(r-1)!(n-r)!} \sum_{i=0}^{r-1} (-1)^i \binom{r-1}{i} \int_0^\infty e^{-x^c(n-r+i+1)} x^k c x^{c-1} dx \\ &= \frac{n!}{(r-1)!(n-r)!} \Gamma\left(1 + \frac{k}{c}\right) \sum_{i=0}^{r-1} \frac{(-1)^i \binom{r-1}{i}}{(n-r+i+1)^{1+(k/c)}}. \end{aligned} \tag{3.29}$$

# 4 Povezane distribucije

Kao što smo već ranije spomenuli troparametarska Weibullova distribucija za različite vrijednosti parametara poprima neke druge oblike i dovodimo je u vezu s nekim drugim distribucija. U ovom poglavlju ćemo nešto više reći o povezanosti Weibullove distribucije s ostalim distribucijama, te ćemo navesti neke distribucije koje su proizašle iz Weibullove distribucije.

Podsjetimo se još jednom kako izgledaju funkcija gustoće i fukcija distribucije slučajne varijable $X$ koja ima troparametarsku Weibullovu distribuciju:

- $f_X(x) = \left(\frac{c}{\alpha}\right)\left(\frac{x-\xi_0}{\alpha}\right)^{c-1} e^{-\left(\frac{x-\xi_0}{\alpha}\right)^c}, x > \xi_0$
- $F_X(x) = 1 - e^{-\left(\frac{x-\xi_0}{\alpha}\right)^c}, x > \xi_0$

Weibullova distribucija postaje **dvoparametarska eksponencijalna distribucija** kada je parametar $c = 1$. Pripadna fukcija distribucije izgleda ovako:

$$F_X(x) = 1 - e^{-\frac{x-\xi_0}{\alpha}}, \quad x > \xi_0. \tag{4.1}$$

U slučaju kada je parametar $c = 2$, Weibullova distribucija postaje **dvoparametarska Rayleighova distribucija** s pripadnom funkcijom distribucije:

$$F_X(x) = 1 - e^{-\left[\frac{x-\xi_0}{\alpha}\right]^2}, \quad x > \xi_0. \tag{4.2}$$

Iz troparametarske Weibullove distribucije možemo izvesti tzv. **reflektirajuću Weibullovu distribuciju** s funkcijom gustoće:

$$f_{X_r}(x) = \frac{c}{\alpha}\left(\frac{\xi_0 - x}{\alpha}\right)^{c-1} e^{-\left[\frac{(\xi_0-x)}{\alpha}\right]^c}, x < \xi_0, \alpha > 0, c > 0, \tag{4.3}$$

i funkcijom distribucije:

$$F_{X_r}(x) = e^{-\left[\frac{(\xi_0-x)}{\alpha}\right]^c}, x < \xi_0, \alpha > 0, c > 0. \tag{4.4}$$

Reflektirajući funkciju gustoće Weibullove slučajne varijable na desnu stranu Balakrishnan i Kocherlakota su 1985. godine (vidi [3]) definirali **dvostruku Weibullovu distribuciju** s funkcijom gustoće:

$$f_X(x) = \frac{c}{2\alpha}\left|\frac{x-\xi_0}{\alpha}\right|^{c-1} e^{-\left|\frac{(x-\xi_0)}{\alpha}\right|^c}, \quad -\infty < x < \infty. \tag{4.5}$$

1984. godine Zacks (vidi [3]) je uveo **troparametarsku eksponencijalnu Weibullovu distribuciju** čija je funkcija distribucije definirana sa:

$$F_X(x;\lambda,c,\tau) = 1 - e^{-\lambda x - \left[\lambda (x-\tau)^+\right]^c}, \quad x \geq 0, \tag{4.6}$$

gdje je $(x-\tau)^+ = \max(0, x-\tau)$, $\lambda > 0$ parametar skaliranja, $c \geq 1$ parametar oblika i $\tau \geq 0$ parametar položaja. Svojstvo ove distribucije je da ima neopadajuću funkciju hazarda ili rizika (više o ovom možete pronaći u [4]).

# 5 Metode procjene parametara

Cilj ovog diplomskog rada je procjena parametara Weibullove distribucije, tj. procjena parametara $c$, $\alpha$ i $\xi_0$. Postoje dvije vrste procjene: grafička i analitička procjena. Zbog svoje jednostavnosti i brzine grafičke metode procjene su popularnije, ali analitičke metode procjene su preciznije. Mi ćemo se bazirati na analitičku procjenu podataka. Nepoznate parametre $c$, $\alpha$ i $\xi_0$ troparametarske Weibullove distribucije treba procijeniti na temelju uzorka $x_1$, $x_2$,..., $x_n$ koji se sastoji od $n$ nezavisnih opažanja slučajne varijable $X \sim W(c, \xi_0, \alpha)$. U tu svrhu razvijene su razne statističke metode. U ovom radu nešto više ćemo reći o:

- Metodi momenata
- Modificiranoj metodi momenata
- Metodi maksimalne vjerodostojnosti
- Modificiranoj metodi maksimalne vjerodostojnosti
- Metodi "Minimum quantile distance" ili metodi kvantila
- Ostalim metodama

## 5.1 Metoda momenata

Metoda momenata je jedna od najzastupljenih statističkih metoda za procjenu parametara neke distribucije. Smatra se da je to i jedna od najstarijih metoda. Ukratko, metoda momenata temelji se na pretpostavci da su vrijednosti statističkih momenata izračunatih na danom uzorku $x_1$, $x_2$,..., $x_n$ bliske vrijednostima teorijskih momenata vjerojatnosne distribucije, odnosno statistički momenti izračunati na danom uzroku (uzorački momenti) izjednačavaju se s teorijskim momentima. Nedostatak metode momenata je taj što uzorak mora biti dovoljno velik.

Koristeći metodu momenata, parametre Weibullove distribucije procjenjujemo koristeći:

- Očekivanje: $\mu = E(X) = \xi_0 + \alpha \left[\Gamma\left(\frac{1}{c}+1\right)\right]$
- Varijancu: $Var(X) = \alpha^2 \left\{\Gamma\left(\frac{2}{c}+1\right) - \Gamma\left[\left(\frac{1}{c}+1\right)\right]^2\right\}$
- Treći centralni moment:
  $\mu_3 = E(X-\mu)^3) = \alpha^3 \left\{\Gamma(1+\frac{3}{c}) - 3\Gamma(1+\frac{1}{c})\Gamma(1+\frac{2}{c}) + 2\left[\Gamma(1+\frac{1}{c})\right]^3\right\}$

Budući da Weibullova distribucija ima tri parametra, procjenitelji parametra Weibullove distribucije mogu biti određeni koristeći očekivanje uzroka $\overline{x}$, varijancu uzorka $s^2$ i treći centralni moment uzorka $\widehat{\mu}^3$.
Nepoznate parametre $c$, $\alpha$ i $\xi_0$ Weibullove distribucije dobivamo rješavanjem sljedećih sustava jednadžbi koji se dobiva izjednačavanjem uzoračkih momenata s odgovarajućim teorijskim momentima:

$$\overline{x} = \xi_0 + \alpha \left[\Gamma\left(\tfrac{1}{c}+1\right)\right],$$

$$s^2 = \alpha^2 \left\{\Gamma\left(\tfrac{2}{c}+1\right) - \Gamma\left[\left(\tfrac{1}{c}+1\right)\right]^2\right\},$$

$$\widehat{\mu}^3 = E(X^3) = \alpha^3 \left\{\Gamma(1+\tfrac{3}{c}) - 3\Gamma(1+\tfrac{1}{c})\Gamma(1+\tfrac{2}{c}) + 2\left[\Gamma(1+\tfrac{1}{c})\right]^3\right\},$$

gdje su:

$$\overline{x} = \textstyle\sum_{i=1}^{n} \frac{x_i}{n},$$

$$s^2 = \textstyle\sum_{i=1}^{n} \frac{(x_i-\overline{x})^2}{n-1},$$

$$\widehat{\mu}^3 = \textstyle\sum_{i=1}^{n} \frac{(x_i-\overline{x})^3}{n}.$$

Kako bi se dobila nepristranost procjenitelja varijance uzorka, u nazivnik od $s^2$ stavlja se $n-1$, a ne $n$.
Koristeći formule za momente od $X$ i uz pretpostavku da je $\xi_0$ poznat godine 1963. Menon (vidi [3]) je predložio jednostavnu formulu za procjenu parametara $c$ dvoparametarske Weibullove distribucije :

$$\widetilde{c}^{-1} = \frac{\sqrt{6}}{\pi} \times \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(\ln x_i - \overline{\ln x}\right)^2} = \frac{\sqrt{6}}{\pi} \times \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(\ln x_i - \sum_{j=1}^{n}\ln\left(\frac{x_j}{n}\right)\right)^2} \tag{5.1}$$

gdje je $\widetilde{c}^{-1}$ asimptotski normalan i nepristran procjenitelj za $c^{-1}$.

Također, statistika

$$\frac{1}{\gamma}\left\{\ln\alpha - \frac{1}{n}\sum_{j=1}^{n}\ln x_j\right\} \tag{5.2}$$

navodi se kao procjenitelj od $c^{-1}$, s varijancom

$$\left(c^2 n\right)^{-1}\frac{\pi^2}{6\gamma^2} = 4.39\frac{c^{-2}}{n} \tag{5.3}$$

gdje je $\gamma$ Eulerova konstanta (vidi [3]).

1963. godine Menon je istaknuo da procjenitelj u (5.2), osim svoje nepristranosti, nema neka značajnija svojstva.

S druge strane, procjenitelj od $\alpha$ izvodi se iz (5.2) i iznosi:

$$\widetilde{\alpha} = exp\left\{\frac{1}{n}\sum_{i=1}^{n}\ln x_i + \gamma\left(\widetilde{c^{-1}}\right)\right\} \tag{5.4}$$

$\widetilde{\alpha}$ je asimptotski nepristran procjenitelj za $\alpha$.

Osim Menona, metodu momenata za troparametarsku Weibullovu distribuciju proučavali su i drugi kao što su Newby (1980,1984) i Cran (1998) (vidi [3]).
Cran je 1988. godine (vidi [6]) predložio metodu koja uključuje prva četiri centralna momenta i dao formule za procijenjene parametre:

$$\widehat{c} = \frac{ln2}{ln\left(\widehat{\mu_1} - \widehat{\mu_2}\right) - ln\left(\widehat{\mu_2} - \widehat{\mu_4}\right)}, \tag{5.5}$$

$$\widehat{\xi_0} = \frac{\widehat{\mu_1}\widehat{\mu_4} + \widehat{\mu_2}^2}{\widehat{\mu_1} + \widehat{\mu_4} - 2\widehat{\mu_2}^2}, \tag{5.6}$$

$$\widehat{\alpha} = \frac{\widehat{\mu_1} - \widehat{\xi_0}}{\Gamma\left(1 + \frac{1}{\widehat{c}}\right)}. \tag{5.7}$$

## 5.2 Modificirana metoda momenata

Za troparametarsku Weibullovu distribuciju Cohen, Whitten i Ding su 1984. godine (vidi [3]) predložili modificiranu metodu momenata, koja je bazirana na običnoj metodi momenata. Razlika je u tome što se treći centralni moment u ovoj metodi zamjenjuje s najmanjom uređajnom statistikom. U modificiranoj metodi nepoznate parametre $c$, $\alpha$ i $\xi_0$ procjenjujemo koristeći:

- Očekivanje: $E(X) = \xi_0 + \alpha\left[\Gamma\left(\frac{1}{c} + 1\right)\right]$
- Varijancu: $Var(X) = \alpha^2\left\{\Gamma\left(\frac{2}{c} + 1\right) - \Gamma\left[\left(\frac{1}{c} + 1\right)\right]^2\right\}$
- Očekivanje najmanje uređajne statistike: $E(X_1') = \xi_0 + \frac{\alpha}{n^{1/c}}\Gamma\left(\frac{1}{c} + 1\right)$

Nepoznate parametre $c$, $\alpha$ i $\xi_0$ Weibullove distribucije dobivamo rješavanjem sljedećih sustava jednadžbi koji se dobiva izjednačavanjem uzoračkih momenata s odgovarajućim teorijskim momentima:

$$\begin{aligned} \overline{x} &= \xi_0 + \alpha\left[\Gamma\left(\frac{1}{c} + 1\right)\right], \\ s^2 &= \alpha^2\left\{\Gamma\left(\frac{2}{c} + 1\right) - \Gamma\left[\left(\frac{1}{c} + 1\right)\right]^2\right\}, \\ x_{(1)} &= \xi_0 + \frac{\alpha}{n^{1/c}}\Gamma\left(\frac{1}{c} + 1\right) \end{aligned} \tag{5.8}$$

Nakon par algebarskih operacija, prethodne jednadžbe postaju:

$$\frac{s^2}{\left(\overline{x} - x_{(1)}\right)^2} = \frac{\widehat{\Gamma}_2 - \widehat{\Gamma}_1^2}{\left[\left(1 - n^{-1/\widehat{c}}\right)\widehat{\Gamma}_1\right]^2}, \tag{5.9}$$

$$\widehat{\xi_0} = \frac{n^{\frac{1}{\widehat{c}}}x_{(1)} - \overline{x}}{n^{\frac{1}{\widehat{c}}} - 1}, \tag{5.10}$$

$$\widehat{\alpha} = \frac{n^{\frac{1}{\widehat{c}}}\left(\overline{x} - x_{(1)}\right)}{\left(n^{\frac{1}{\widehat{c}}} - 1\right)\widehat{\Gamma}_1}, \tag{5.11}$$

gdje je $\widehat{\Gamma}_j = \Gamma_j(\widehat{c}) = \Gamma\left(1 + \frac{j}{\widehat{c}}\right)$.

Radi jednostavnosti koriste se sljedeće formule za procjenu parametara:

$$\widehat{\alpha} = S \cdot C(\widehat{c}), \tag{5.12}$$

$$\widehat{\xi}_0 = \overline{x} - S \cdot D(\widehat{c}), \tag{5.13}$$

gdje je $C(c) = \frac{1}{\sqrt{\Gamma_2 - \Gamma_1{}^2}}$, $D(c) = \frac{\Gamma_1}{\sqrt{\Gamma_2 - \Gamma_1{}^2}}$.

Kada su nam poznati $\overline{x}$, $s^2$ i $x_{(1)}$ onda iz jednadžbe (5.9) lako računamo $\widehat{c}$, a isto tako $\widehat{\xi}_0$ iz (5.10), te $\widehat{\alpha}$ iz (5.11).
Iako je ova metoda primjenjiva za sve vrijednosti parametra $c$, Cohnen, Whitten i Ding (vidi [3]) istaknuli su da bi moglo doći do nekih računskih problema u slučaju kada je $c < 0.5$, no taj slučaj nije od neke velike važnosti za praksu.

## 5.3 Metoda maksimalne vjerodostojnosti

Metoda maksimalne vjerodostojnosti ili kraće MLE metoda (engl. Maximum Likelihood Estimation) jedna je od najpoznatijih statističkih metoda za procjenu parametara neke distribucije. Razlog tomu je taj što se može primijeniti na većinu teorijskih distribucija i što pod određenim uvjetima ima jako dobra statistička svojstva kao što su invarijantnost, konzistentnost i asimptotska nepristranost.

Neka su $x_1, x_2, ..., x_n$ nezavisne realizacije slučajne varijable $X > 0$ s funkcijom gustoće $f(x; \boldsymbol{\theta})$. Odgovarajuća funkcija vjerodostojnosti se definirana kao:

$$L(\boldsymbol{\theta}) = \sum_{i=1}^{n} f(x_i; \boldsymbol{\theta}). \tag{5.14}$$

gdje $L : \boldsymbol{\theta} \to [0, \infty\rangle$. Bez smanjenja općenitosti možemo pretpostaviti da je:

$$x_1 < x_2 < ... < x_n$$

Procjena metodom maksimalne vjerodostojnosti je vrijednost parametra $\widehat{\boldsymbol{\theta}}$ koji maksimizira $L(\boldsymbol{\theta}) = \sum_{i=1}^{n} f(x_i; \boldsymbol{\theta})$ na skupu svih mogućih vrijednosti vektora parametara $\boldsymbol{\theta}$. Budući da je logaritamska funkcija strogo rastuća, problem maksimizacije funkcije $L(\boldsymbol{\theta})$ ekvivalentan je problemu maksimizacije funkcije $\ln L(\boldsymbol{\theta})$.

Nas zanima procjena paramatera Weibullove distribucije. Krenimo prvo od dvoparametarske zato što je uobičajeno da je $\xi_0$ poznat, npr. $\xi_0 = 0$ bez smanjenja općenitosti. Za dvoparametarsku Weibullovu distribuciju, odnosno kada je $\xi_0 = 0$, funkcija vjerodostojnosti zadana je sa:

$$L(c,\alpha) = \sum_{i=1}^{n} \left(\frac{c{x_i}^{c-1}}{\alpha^c}\right) e^{-\left(\frac{x_i}{\alpha}\right)^c}, \tag{5.15}$$

a njezin prirodni logaritam glasi:

$$\ln L(c,\alpha) = n - nc\ln\alpha + (c-1)\sum_{i=1}^{n} i - \frac{1}{\alpha^c}\sum_{i=1}^{n} {x_i}^c. \tag{5.16}$$

Parcijalnim deriviranjem prethodne jednadžbe po $\alpha$ i $c$ i izjednačavanjem s 0 dobivamo:

$$\frac{\partial \ln L}{\partial c} = \frac{n}{c} - n\ln n + \sum_{i=1}^{n} \ln x_i + \frac{\ln\alpha}{\alpha^c}\sum_{i=1}^{n} {x_i}^c - \frac{1}{\alpha^c}\sum_{i=1}^{n} {x_i}^c \ln x_i = 0, \tag{5.17}$$

$$\frac{\partial \ln\alpha}{\partial\alpha} = \frac{-nc}{\alpha} = \frac{c}{\alpha^{c+1}}\sum_{i=1}^{n} {x_i}^c = 0. \tag{5.18}$$

Supstitucijom (5.18) u (5.17) i sređivanjem izraza, jednadžba koju treba zadovoljavati ML-procjenitelj $\widehat{c}$ glasi:

$$\widehat{c} = \left[\left(\sum_{i=1}^{n} {x_i}^{\widehat{c}} \ln x_i\right)\left(\sum_{i=1}^{n} {x_i}^{\widehat{c}}\right)^{-1} - \frac{1}{n}\sum_{i=1}^{n} {x_i}^{\widehat{c}} \ln x_i\right]^{-1}. \tag{5.19}$$

Uvrštavanjem tog rješenja u (5.18) dobiva se ML-procjenitelj $\widehat{\alpha}$:

$$\widehat{\alpha} = \left(\frac{1}{n}\sum_{i=1}^{n} {x_i}^{\widehat{c}}\right)^{\frac{1}{\widehat{c}}}. \tag{5.20}$$

ML-procjenitelj za dvoparametarsku Weibullovu distribuciju postoji i jedinstven je, dok za troparametarsku Weibullovu distribuciju to nije slučaj. Pogledajmo zašto:

Za bilo koju realizaciju

$$0 < x_1 < x_2 < ... < x_n$$

troparametarske Weibullove slučajne varijable $X$ standarni ML- procjenitelj ne postoji.

Naime, ako fiksiramo $c \in (0,1)$ i $\alpha \in (0,\infty)$ tada

$$f(x_1;\xi_0,c,\alpha) = \left(\frac{c}{\alpha}\right)\left(\frac{x_1-\xi_0}{\alpha}\right)^{c-1} e^{-\left(\frac{x_1-\xi_0}{\alpha}\right)^c} \to \infty$$

kada $\xi_0 \to x_1$ slijeva.

Osim toga, $\forall i = 2, ...n$ vrijedi:

$$f(x_i;\xi_0,c,\alpha) = \left(\frac{c}{\alpha}\right)\left(\frac{x_i-\xi_0}{\alpha}\right)^{c-1} e^{-\left(\frac{x_i-\xi_0}{\alpha}\right)^c} \to \left(\frac{c}{\alpha}\right)\left(\frac{x_i-x_1}{\alpha}\right)^{c-1} e^{-\left(\frac{x_i-x_1}{\alpha}\right)^c} > 0$$

Zbog toga je:

$$\lim_{\xi_0 \to x_1} L(\xi_0, c, \alpha) \to \infty.$$

Da bi ML-procjenitelj postojao limes fukcije vjerodostojnosti mora biti konačan, što kod nas nije slučaj.

Za troparametarsku Weibullovu distribuciju funkcija vjerodostojnosti zadana je s:

$$L_3(\xi_0, c, \alpha) = \frac{c^n}{\alpha^c n} \left[ \prod_{i=1}^{n} (x_i - \xi_0)^{c-1} \right] e^{-\frac{1}{\alpha^c} \sum_{i=1}^{n} (x_i - \xi_0)^c}. \tag{5.21}$$

Unatoč nepostojanju ML- procjenitelja za troparametarsku Weibullovu slučajnu varijablu $X$ ipak se promatra problem procjene parametara, na način da se promatra sustav jednadžbi koji se dobiva izjednačavanjem parcijalnih derivacija (5.22) s nulom. Tj. dobije se sljedeći sustav jednadžbi:

$$\alpha^c - \frac{1}{n} \sum_{i=1}^{n} (x_i - \xi_0)^c = 0, \tag{5.22}$$

$$\frac{\sum_{i=1}^{n} (x_i - \xi_0)^c \ln (x_i - \xi_0)}{\sum_{i=1}^{n} (x_i - \xi_0)} - \frac{1}{c} \sum_{i=1}^{n} \ln (x_i - \xi_0) = 0, \tag{5.23}$$

$$(c-1) \sum_{i=1}^{n} (x_i - \xi_0)^{-1} - c\alpha^{-c} \sum_{i=1}^{n} (x_i - \xi_0)^{c-1} = 0. \tag{5.24}$$

Ove jednadžbe nazivaju se jednadžbe vjerodostojnosti.

Rješavanjem sustava jednadžbi (5.23)- (5.25) procjenitelji dobiveni metodom maksimalne vjerodostojnosti zadovoljavaju sljedeće jednažbe:

$$\widehat{\alpha} = \left[ \frac{1}{n} \sum_{i=1}^{n} \left( x_i - \widehat{\xi}_0 \right)^{\widehat{c}} \right]^{\frac{1}{\widehat{c}}}, \tag{5.25}$$

$$\widehat{c} = \left\{ \left[ \sum_{i=1}^{n} \left( x_i - \widehat{\xi}_0 \right)^{\widehat{c}} \ln \left( x_i - \widehat{\xi}_0 \right) \right] \left[ \sum_{i=1}^{n} \left( x_i - \widehat{\xi}_0 \right)^{\widehat{c}} \right]^{-1} - \frac{1}{n} \sum_{i=1}^{n} \ln \left( x_i - \widehat{\xi}_0 \right) \right\}^{-1}, \tag{5.26}$$

$$(\widehat{c} - 1) \sum_{i=1}^{n} \left( x_i - \widehat{\xi}_0 \right)^{-1} = \widehat{c} \widehat{\alpha}^{-\widehat{c}} \sum_{i=1}^{n} \left( x_i - \widehat{\xi}_0 \right)^{\widehat{c}-1}. \tag{5.27}$$

Ako vrijednost od $\widehat{\xi}_0$ zadovoljava jednažbe (5.26), (5.27) i (5.28) i ako je vrijednost od $\widehat{\xi}_0$ veća od najmanje uređene statistike $X_{(1)}$, tada je $\widehat{\xi}_0$ ML- procjenitelj od $\xi_0$. Inače je $\widehat{\xi}_0 = x_{(1)}$.

Postojanje rješenja za sustav jednadžbi (5.23) - (5.25) privuklo je mnogo pažnje, zbog toga što taj sustav može imati više od jednog rješenja ili pak ne mora imati niti jedno rješenje (više o egzistenciji rješenja možete pogledati u [6]), ili dobiveno rješenje predstavlja točku lokalnog maksimuma ili lokalnog minimuma s obzirom na $\xi_0$ (vidi [4]).

Što se tiče asimptotskih svojstava ML-procjenitelja, ML-procjenitelji Weibullove distribucije su asimptotski normalni samo kada je vrijednost parametra $c > 2$. Kada je $c \in (0,1)$, tada je $X_{(1)}$ "superefikasan" procjenitelj za $\widehat{\xi}_0$.

Ako su ML- procjenitelji asimptotski normalni, i ako je $\xi_0$ poznat parametar, Dubey je 1965. godine (vidi [3]) za velike n dao formule:

$$nVar(\widehat{\alpha}) = \left\{1 + \frac{[\psi(2)]^2}{\psi'(1)}\right\}\left(\frac{\alpha}{c}\right)^2 = 1.109\left(\frac{\alpha}{c}\right)^2, \tag{5.28}$$

$$nVar(\widehat{c}) = c^2\left[\psi'(1)\right]^{-1} = \frac{6c^2}{\pi^2} = 0.608c^2, \tag{5.29}$$

$$Corr(\widehat{\alpha}, \widehat{c}) = \frac{\psi(2)}{\left\{\psi'(1) + [\psi(2)]^2\right\}^{\frac{1}{2}}} = 0.313, \tag{5.30}$$

gdje su $\psi(\cdot)$ i $\psi'(\cdot)$ digama i trigama funkcije definirane s:

$$\psi(x) = \frac{d}{dx}\ln\Gamma(x),$$

$$\psi'(x) = \frac{d^2}{dx^2}\ln\Gamma(x).$$

Kako jednadžbe (5.26) - (5.28) ne daju eksplicitna rješenja za procjenitelje, Cohen je 1965. godine (vidi [3]) predložio uklanjanje $\alpha^c$ iz jednadžbi (5.26) i (5.27). Iz čega dobivamo:

$$\left[\frac{\sum_{i=1}^n\left(x_i - \widehat{\xi}_0\right)^{\widehat{c}} ln\left(x_i - \widehat{\xi}_0\right)}{\sum_{i=1}^n\left(x_i - \widehat{\xi}_0\right)^{\widehat{c}}} - \frac{1}{c}\right] - \frac{1}{n}\sum_{i=1}^n ln\left(x_i - \widehat{\xi}_0\right) = 0. \tag{5.31}$$

- Ako je $\xi_0$ poznat, onda iz prethodne jednadžbe izračunamo procjenitelj $\widehat{c}$.
- Ako $\xi_0$ nije poznat Cohen predlaže sljedeće:
  - odabrati prvu aproksimaciju za $\widehat{\xi}_0 < X_{(1)}$
  - riješiti jednadžbe (5.32) i (5.26) za $\widehat{\alpha}$ i $\widehat{c}$
  - zatim provjeriti zadovoljava li jednadžba za procjenu $\widehat{\xi}_0$ ove dvije vrijednosti za procjenu $\widehat{\alpha}$ i $\widehat{c}$
  - nakon toga isprobavati sve moguće kombinacije dok procjenitelj od $\xi_0$ ne bude dovoljno blizu predloženom.

## 5.4 Modificirana metoda maksimalne vjerodostojnosti

Cohen i Whitten su 1982. godine (vidi [3]) predložili modificiranu metodu maksimalne vjerodostojnosti za procjenu parametara $c$, $\alpha$ i $\xi_0$ kao alternativu za metodu maksimalne vjerodostojnosti kada je vrijednost parametra $c < 2$ ili $c < 2.2$. Kako bismo postigli asimptotsku normalnost procjenitelja nastalih metodom maksimalne vjerodostojnosti koristimo modificiranu metodu maksimalne vjerodostojnosti.

Metoda se zasniva na tome da se jednadžba vjerodostojnosti

$$(\widehat{c}-1)\sum_{i=1}^{n}\left(x_i-\widehat{\xi}_0\right)^{-1}=\widehat{c}\widehat{\alpha}^{-\widehat{c}}\sum_{i=1}^{n}\left(x_i-\widehat{\xi}_0\right)^{\widehat{c}-1} \tag{5.32}$$

za $\xi_0$ zamijeni s alternativnim jednadžbama u sljedećim oblicima:

$$-ln\left(\frac{n}{n+1}\right)=\frac{\left(x_{(1)}-\xi_o\right)^c}{\alpha^c}, \tag{5.33}$$

$$\xi_0+\left(\frac{\alpha}{n^{\frac{1}{c}}}\right)\Gamma\left(1+\frac{1}{c}\right)=x_{(1)}, \tag{5.34}$$

$$\xi_0+\alpha\Gamma\left(1+\frac{1}{c}\right)=\overline{x}, \tag{5.35}$$

$$\alpha 2\left[\Gamma\left(1+\frac{2}{c}\right)-\Gamma^2\left(1+\frac{1}{c}\right)\right]=s^2, \tag{5.36}$$

$$\xi_0+\alpha(ln2)^{\frac{1}{c}}=x_{med}, \tag{5.37}$$

gdje se $x_{med}$ odnosi na uzorački medijan.

Procjenitelje modificiranom metodom maksimalne vjerodostojnosti dobivamo tako da riješimo sustav jednadžbi koji se sastoji od jednadžbi:

$$\widehat{\alpha}=\left[\frac{1}{n}\sum_{i=1}^{n}\left(x_i-\widehat{\xi}_0\right)^{\widehat{c}}\right]^{\frac{1}{\widehat{c}}}, \tag{5.38}$$

$$\widehat{c}=\left\{\left[\sum_{i=1}^{n}\left(x_i-\widehat{\xi}_0\right)^{\widehat{c}}\ln\left(x_i-\widehat{\xi}_0\right)\right]\left[\sum_{i=1}^{n}\left(x_i-\widehat{\xi}_0\right)^{\widehat{c}}\right]^{-1}-\frac{1}{n}\sum_{i=1}^{n}\ln\left(x_i-\widehat{\xi}_0\right)\right\}^{-1}, \tag{5.39}$$

i jedne od jednadžbi (5.33)- (5.37), u ovisnosti koju odaberemo.

Više o ovoj metodi se može pronaći u [4].

## 5.5 Metoda kvantila

Da bismo nešto više mogli reći o ovoj metodi, prvo ćemo reći što je to funkcija kvantila i kako ona izgleda za troparametarsku Weibullovu distribuciju.

Neka je $X$ slučajna varijabla s funkcijom gustoće $F_X(x)$ i neka je $p \in (0,1)$. Broj $x_p$ se naziva p- ti kvantil ako je

$$F(x_p) = p.$$

Funkciju kvantila $Q(p)$ definiramo kao inverz kumulativne funkcije gustoće, odnosno:

$$Q(p) = F^{-1}(p).$$

Metodu kvantila (engl. Minimum quantile distance estimation) predložili su Carmody, Eubank i LaRiccia 1984. godine (vidi [3]). Za troparametarsku Weibullovu distribuciju pomoću minimalne udaljenosti kvantila procijenit ćemo parametar $\boldsymbol{\theta} = (\xi_0, \alpha, c)$.

Neka je $Q(p) = -\ln(1-p), 0 < p < 1$ funkcija kvantila od eksponencijalne distribucije. Otprije znamo da su nam eksponencijalna i Weibullova distribucija u vezi pa iz funkcije kvantila eksponencijalne distribucije dobijemo funkciju kvantila troparametarske Weibullove distribucije :

$$\begin{aligned}
&F_X(x) = p \\
&1 - e^{-y} = p \\
&1 - e^{-Q} = p \\
&1 - e^{-p(\frac{Q-\xi_0}{\alpha}^{c}} = p \\
&1 - p = e^{-\left(\frac{Q-\xi_0}{\alpha}\right)^c} \quad / \ln \\
&\ln(1-p) = -\left(\frac{x-\xi_0}{\alpha}\right)^c \quad /^{\frac{1}{c}} \\
&(\ln(1-p))^{\frac{1}{c}} = -\left(\frac{Q-\xi_0}{\alpha}\right) \qquad /(-\alpha) \\
&-\alpha\,(\ln(1-p))^{\frac{1}{c}} = Q - \xi_0 \\
&Q = \xi_0 + \alpha\left[-\ln(1-p)\right]^{\frac{1}{c}},
\end{aligned} \tag{5.40}$$

tj. funkcija kvantila troparametarske Weibullove distribucije iznosi:

$$Q(p;\xi_0,\alpha,c) = \xi_0 + \alpha\left[-\ln(1-p)\right]^{\frac{1}{c}}. \tag{5.41}$$

Neka je uzoračka funkcija kvantila definirana s:

$$\widetilde{Q}(p) = x_{(i)}, \qquad \frac{i-1}{n} < p \leq \frac{i}{n}, i = 1, 2, ..., n. \tag{5.42}$$

Za dane $k < n$ neka je $\boldsymbol{P} = (p_1, p_2, ..., p_k)$ tako da je $0 < p_1 < ... < p_k < 1$.

Neka je

$$\widetilde{\boldsymbol{Q}}_{\boldsymbol{p}} = \left(\widetilde{Q}(p_1), \widetilde{Q}(p_2), ..., \widetilde{Q}(p_k)\right)^T. \tag{5.43}$$

vektor funkcija $\widetilde{Q}(p_1), \widetilde{Q}(p_2), ..., \widetilde{Q}(p_k)$ gdje je $\widetilde{Q}(p_i)$ definiran kao u (5.42) $\forall i = 1, ..., k$ i neka je

$$\boldsymbol{Q}_{\boldsymbol{p}}(\boldsymbol{\theta}) = (Q(p_1; \boldsymbol{\theta}), Q(p_2; \boldsymbol{\theta}), ..., Q(p_k; \boldsymbol{\theta}))^T \tag{5.44}$$

vektor funkcija $Q(p_1; \boldsymbol{\theta}), Q(p_2; \boldsymbol{\theta}), ..., Q(p_k; \boldsymbol{\theta})$ gdje je $Q(p_i; \boldsymbol{\theta})$ definiran kao u (5.41) $\forall i = 1, ..., k$.

Procjenitelj od $\boldsymbol{\theta}$ dobiven metodom kvantila ili metodom najmanje udaljenosti kvantila je onaj vektor koji minimizira kvadratnu formu:

$$\left(\widetilde{\mathbf{Q}}_{\boldsymbol{p}} - \boldsymbol{Q}_{\boldsymbol{p}}(\boldsymbol{\theta})\right)^T W(\boldsymbol{\theta}) \left(\widetilde{\mathbf{Q}}_{\boldsymbol{p}} - \boldsymbol{Q}_{\boldsymbol{p}}(\boldsymbol{\theta})\right) \tag{5.45}$$

kao funkciju od $\boldsymbol{\theta}$. $W(\theta)$ je $k \times k$ težinska matrica koju odabiremo u ovisnosti o $\boldsymbol{\theta}$.

Dobar izbor za težinsku matricu $W$ predložili su 1984. godine Carmody, Eubank i LaRiccia (vidi [3]) i ona izgleda ovako:

$$W^*(c) = H_p(c) {R_p}^{-1} H_p(c) \tag{5.46}$$

gdje je :

- ${R_p}^{-1} = \min(p_i, p_j) - p_i p_j$
- $H_p(c)$ je $k \times k$ dijagonalna matrica koja na i-toj dijagonali ima elemente $c(1-p_i) \left[Q(p_i)\right]^{\frac{(c-1)}{c}}$

Elementi matrice $W^*(c)$ mogu se laku izračunati budući da je matrica ${R_p}^{-1}$ tridijagonalna inverzna matrica.

Kako procjenitelj $\widetilde{\boldsymbol{\theta}}_{\boldsymbol{p}}(W)$ nema zatvorenu formu, procjenitelj $\widetilde{\boldsymbol{\theta}}$ se može lako izračunati koristeći neke standardne procedure minimizacije. Procjenitelj dobiven metodom minimalne udaljednosti kvantila ima i neka aspimptotska svojstva kao što su jedinstvenost, konzistentnost i asimptotska normalnost (vidi [3]).

## 5.6 Ostale metode procjene parametara

Osim prethodno navedenih metoda procjene parametara postoji još mnogo metoda za procjenu parametara Weibullove distribucije kao što je metoda najmanjih običnih kvadrata, metoda najmanjih potpunih kvadrata, hibridne metode, grafičke metode, Bayesova metoda, Shrinkageova metoda itd. Nešto više o ovim metodama možemo pronaći u [4].

# 6 Uspoređivanje asimptotskih svojstava procjenitelja kroz simulacije

U ovom poglavlju usporedit ćemo asimptotska svojstva procjenitelja kroz simulacije, te ćemo kroz primjer vidjeti koja nam je metoda najbolja. Svi izračuni su provedeni u programskom paketu R. Simulacije provodimo za dvoparametarsku Weibullovu distribuciju s funkcijom gustoće:

$$f_X(x) = \frac{c}{\alpha^c} x^{c-1} e^{-\left(\frac{x}{\alpha}\right)^c}, x > 0, \alpha > 0, c > 0 \tag{6.1}$$

Parametre ćemo procijeniti metodom momenata, metodom maksimalne vjerodostojnosti i metodom kvantila. Kod metode momenata koristili smo očekivanje uzorka $\overline{x}$ te i varijancu uzorka $s^2$ i naredbu `uniroot.all` iz paketa `rootSolve` kako bismo riješili jednadžbe i time procijenili parametre $\alpha$ i $c$. Kod metode maksimalne vjerodostojnosti za procjenu parametara koristili smo naredbu `fitdistr` iz paketa `MASS` dok smo kod metode kvantila koristili naredbu `qmedist` iz paketa `fitdistrplus`. Početni parametri nam iznose $c = 4$, $\alpha = 10$ i broj ponavljanja simulacija je 500.

Provest ćemo simulacije za različite veličine uzorka ($n \in \{200, 300, 500\}$).

Rezultati simulacija dani su tablicom:

| | | MOM | ML | MQD |
|---|---|---|---|---|
| **n=200** | | | | |
| prosjek | c | 4.005082 | 4.015945 | 4.066991 |
| pristranost | c | 0.00508246 | 0.01594485 | 0.06699131 |
| MSE | c | 0.05191178 | **0.04792537** | 0.1401134 |
| prosjek | $\alpha$ | 10.00259 | 10.00102 | 10.00316 |
| pristranost | $\alpha$ | 0.002590835 | 0.001021713 | 0.003156106 |
| MSE | $\alpha$ | 0.03429688 | **0.03410405** | 0.04182995 |
| **n=300** | | | | |
| prosjek | c | 4.013545 | 4.021018 | 4.063832 |
| pristranost | c | 0.01354517 | 0.02101772 | 0.06383243 |
| MSE | c | 0.03463069 | **0.03286436** | 0.08704107 |
| prosjek | $\alpha$ | 10.00678 | 10.00583 | 9.996596 |
| pristranost | $\alpha$ | 0.006777979 | 0.005828408 | -0.003404481 |
| MSE | $\alpha$ | 0.02327236 | **0.02315176** | 0.03015525 |
| **n=500** | | | | |
| prosjek | c | 4.002251 | 4.010216 | 4.026388 |
| pristranost | c | 0.002251405 | 0.01021582 | 0.02638809 |
| MSE | c | 0.02266875 | **0.0213783** | 0.05206197 |
| prosjek | $\alpha$ | 9.999662 | 9.998993 | 9.996482 |
| pristranost | $\alpha$ | -0.0003377381 | -0.001006584 | -0.003518408 |
| MSE | $\alpha$ | 0.01487662 | **0.01486563** | 0.01855873 |

Tablica 1: Rezultati simulacija za dvoparametarsku Weibullovu distribuciju, $c = 4$, $\alpha = 10$

Ako pogledamo rezultate simulacija u Tablici 1. i vrijednosti za srednje kvadratnu grešku (MSE), možemo uočiti kako nam se među ispitivanim metodama metoda maksimalne vjerodostojnosti pokazala kao najbolja metoda za procjenu parametara jer u svim slučajevima imamo najmanju srednje kvadratnu grešku.
Treba također napomenuti da je kod metode kvantila korištena matrica identiteta kao težinska matrica. Teoretski rezultati opisani u poglavlju 5.5 sugeriraju da bi ta metoda dala bolje rezultate uz primjenu matrice definirane u (5.46).
Uspoređivanjem navedena tri procjenitelja kroz simulacije na velikom skupu podataka pokazuje se da metoda maksimalne vjerodostojnosti daje najbolje rezultate iako valja naglasiti da metoda momenata ne odstupa u velikoj mjeri od metode maksimalne vjerodostojnosti u srednjekvadratnom smislu. Za očekivati je, također, da bi i metoda kvantila dala bolje rezultate prikladnijim odabirom težinske matrice.

# 7 Literatura


[1] Abernethy R. B.: *The New Weibull Handbook* , Abernethy R. B., Florida(2010).

[2] Eubank R. L., LaRiccia V. N.: *Weighted $L^2$ Quantitle Distance Estimators for Randomly Censored Data* , Department of Statistic, Dalas, Texas (1981.)

[3] Johnson N. L., Kotz S., Balakrishnan N.: *Continuous Univariate Distributions, Volume 1, Second Edition* , John Wiley  Sons, New York (1994.)

[4] Marković D.: *Problem procjene parametara u Weibullovom modelu* , Sveučilište u Zagrebu, Zagreb (2009.)

[5] Meeker W. Q., Escobar L. A.: *Statistical Methods for Reliability Data* , Wiley-Interscience, New York (1998.)

[6] Murthy D. N. Prabhakar , Xie M., Jiang R.: *Weibull Models* , John Wiley  Sons, New Jersey (2004.)

[7] Rosin, P., Rammler, E.: *The laws governing the neness of powdered coal* , Inst. Fuel (1933.)

[8] Tippett L.H.C.,Fisher R.A.: *Limiting forms of the frequency distribution of the largest and smallest members of a sample* , Proc. Camb. Philos. Soc. (1928.)

# 8 Sažetak

Weibullova distribucija se najčešće koristi u teoriji pouzdanosti te u teoriji životnog vijeka kao matematički model za opisivanje slučajnog vijeka trajanja nekog elementa a isto tako i u farmakologiji, biologiji, elektrotehnici, itd.
U ovom diplomskom radu glavni cilj nam je procjena parametara Weibullove distribucije. Kako bi što bolje pristupili tom problemu prvo smo nešto više rekli o samoj Weibullovoj distribuciji i njezinoj primjeni. Weibullova distribucija za različite vrijednosti parametara $c$,$\alpha$ i $\xi_0$ poprima različite oblike i možemo je dovesti u vezu s nekim drugim distribucijama (npr. dvoparametarska eksponencijalna kada je $c = 1$, dvoparametarska Rayleigh-ova kada je $c = 2$, itd.).
Problem procjene parametara promatrali smo kroz 5 metoda: metode momenata, modificirane metode momenata, metode maksimalne vjerodostojnosti, modificirane metode maksimalne vjerodostojnosti te metode kvantila odnosno najmanje udaljenosti kvantila.
Na samom kraju rada smo kroz simulacije uspoređivali asimptotska svojstva procjenitelja. Metodom momenata, metodom maksimalne vjerodostojnosti i metodom minimalne udaljenosti kvantila smo u programskom paketu R procijenili parametre dvoparametarske Weibullove distribucije.

**Ključne riječi**: Weibullova distribucija, standardna Weibullova distribucija, funkcija gustoće, funkcija distribucije, procjena parametara, metode za procjenu parametara.

# 9 Summary

Weibbull's distribution is frequently used in theory of reliability and in theory of life expectancy as a mathematical model as well as in pharmacology, biology, electrical engineering etc. The main aim of this degree essay is the estimate parameters of Weibull distribution. In order to analyse this problem it is important to point out general facts about this distribution and it's use. Weibull's distribution for different values of the parameters c takes different forms and can be related to some other distributions (for example two- parameter exponential when $c = 1$, two-parameter Rayleigh's when $c = 2$, etc.). The problem of the parameters and their estimation was analysed in five methods: moment estimation, modified moment estimation, maximum likelihood estimation, modified maximum likelihood estimation and method of quantile respectively minimum quantile distance estimation. Eventually we have compared asymptotic characteristics of estimators through simulation. With these three methods- method of moments, maximum likelihood estimation and method of minimum quantile distance we have estimated the parameters of two - parameter Weibull's distribution in program package R.

**Key words**: Weibull distribution, standard Weibull distribution, density function, distribution function, parameter estimation, methods of parameter estimation.

# 10 Životopis

Rođena sam 26. studenog 1991. u Sarajevu, Bosna i Hercegovina. Osnovnu školu "Ivan Goran Kovačić" pohađala sam u Livnu, Bosna i Hercegovina, koju sam završila 2006. kada upisujem Opću Gimnaziju u Livnu. Nakon položene državne mature 2010. godine upisujem Preddiplomski studij matematike na Odjelu za matematiku Sveučilišta u Rijeci. Godine 2014. završila sam Preddiplomski studij izradom završnog rada Klizne krivulje iz područja Diferencijalne geometrije te sam stekla akademski stupanj prvostupnica matematike. Iste godine upisujem Diplomski studij matematike na Odjelu za matematiku Sveučilišta u Osijeku, smjer Financijska matematika i statistika. Tijekom završne godine diplomskog studija obavila sam stručnu studentsku praksu u Koncernu Agram, točnije u osiguravajućem društvu Euroherc i osiguravajućem društvu Bosna Sunce Osiguranje kao aktuar. Trenutno sam zaposlena u osiguravajućem društvu Agram Life kao aktuar pripravnik.