Generalizirana Gaussova distribucija

Kalkan, Domagoj

Sveučilište J.J. Strossmayera u Osijeku
Odjel za matematiku
Sveučilišni diplomski studij matematike; smjer: Financijska matematika i statistika

Domagoj Kalkan

Generalizirana Gaussova distribucija

Diplomski rad

Osijek, 2022.

Sveučilište J.J. Strossmayera u Osijeku
Odjel za matematiku
Sveučilišni diplomski studij matematike; smjer: Financijska matematika i statistika

Domagoj Kalkan

Generalizirana Gaussova distribucija

Diplomski rad

Mentor: izv. prof. dr. sc. Nenad Šuvak

Osijek, 2022.

Sadržaj

1 Uvod

Kako bismo mogli započeti analizu generalizirane Gaussove distribucije, podsjetimo se malo osnovnih svojstava Gaussove (normalne) distribucije.

Normalna distribucija najviše se koristi u statističkoj teoriji i primjeni. Teorija je pokazala da ona vrlo dobro opisuje pokuse čiji su ishodi posljedica sume mnogo međusobno nezavisnih i jednako distribuiranih utjecaja.

Definicija 1.1. *Za neprekidnu slučajnu varijablu X na vjerojatnosnom prostoru $(\Omega, \mathcal{F}, P)$ kažemo da ima Gaussovu ili normalnu distribuciju s paramatrima μ i σ^2 ako je njena funkcija gustoće dana izrazom*

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad x \in \mathbb{R},$$

gdje su μ i σ realni brojevi i $\sigma > 0$. Ako slučajna varijabla X ima normalnu distribuciju s parametrima μ i σ^2, koristimo oznaku $X \sim \mathcal{N}(\mu, \sigma^2)$.

Funkcija distribucije normalne slučajne varijable definirana je izrazom

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}}\, dt, \qquad x \in \mathbb{R}.$$

Normalna distribucija s parametrima $\mu = 0$ i $\sigma = 1$ zove se jedinična ili standardna normalna distribucija. Funkcija gustoće standardne normalne distribucije dana je izrazom

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \qquad x \in \mathbb{R}.$$

Neka je $X \sim \mathcal{N}(\mu, \sigma^2)$, tada je očekivanje slučajne varijable X

$$EX = \mu,$$

a varijanca

$$VarX = \sigma^2.$$

2 Osnovna svojstva generalizirane Gaussove distribucije

2.1 Funkcija gustoće

Za neprekidnu slučajnu varijablu X na vjerojatnosnom prostoru $(\Omega, \mathcal{F}, P)$ kažemo da ima generaliziranu Gaussovu distribuciju s parametrima μ, σ^2 i s ako je njena funkcija gustoće dana izrazom

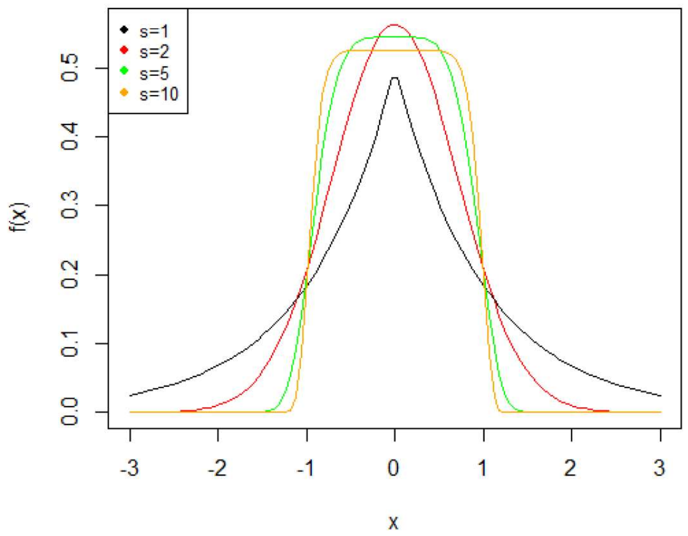
$$f(x) = Ke^{-\left|\frac{x-\mu}{\sigma}\right|^{s}}, \qquad x \in \mathbb{R}, \quad s > 0,$$

gdje je K normalizacijska konstanta sljededećg oblika:

$$K = \frac{s}{2\sigma\Gamma(1/s)},$$

a $\Gamma(\cdot)$ gama funkcija koja je za $z > 0$ definirana sljedećim izrazom:

$$\Gamma(z) = \int\limits_0^\infty t^{z-1}\, e^t\, dt.$$



Slika 1: Generalizirana Gaussova distribucija za $\mu = 0$ i $\sigma = 1$

U slučaju kada je s=2, dobijemo standardnu Gaussovu distribuciju, a u slučaju kada je s=1, dobijemo Laplaceovu distribuciju.

2.2 Funkcija distribucije

U slučaju kada je $x \leq \mu$, tada je funkcija distribucije od X dana izrazom:

$$F(x) = \frac{s \int_{-\infty}^{x} \exp\{-(\mu - y)^s / \sigma^s\} dy}{2\sigma\Gamma(1/s)},$$

što uz supstituciju $z = ((\mu - y)/\sigma)^s$ daje sljedeći izraz:

$$F(y) = \frac{\int_{((\mu-y)/\sigma)^s}^{\infty} z^{1/s-1} \exp(-z) dz}{2\Gamma(1/s)}.$$

Korištenjem nepotpune Gama funkcije za koju vrijedi:

$$\Gamma(a, x) = \int_{x}^{\infty} t^{a-1} \exp(-t) dt,$$

dobijemo sljedeći izraz:

$$F(x) = \frac{\Gamma(1/s, ((\mu - x)/\sigma)^s)}{2\Gamma(1/s)}.$$

Slično, za $x > \mu$ funkcija distribucije dana je izrazom:

$$F(x) = 1 - \frac{\Gamma(1/s, ((x - \mu)/\sigma)^s)}{2\Gamma(1/s)}.$$

2.3 Očekivanje i varijanca

Ukoliko slučajnu varijablu Z zapišemo kao $Z = (X - \mu)/\sigma$, tada je njezina funkcija gustoće

$$f(z) = \frac{s \exp\{-|z|^s\}}{2\Gamma(1/s)}.$$

Pokazuje se da je k-ti moment slučajne varijable Z

$$\mathrm{E}(Z^k) = \frac{1 + (-1)^k}{2\Gamma(1/s)} \Gamma\left(\frac{k+1}{s}\right)$$

Nadalje, n-ti moment od X tada izgleda:

$$\begin{aligned} \mathrm{E}(X^n) &= \mathrm{E}[(\mu + \sigma Z)^n] \\ &= \sum_{k=0}^{n} \binom{n}{k} \mu^{n-k} \sigma^k E(Z^k) \\ &= \frac{\mu^n \sum_{k=0}^{n} \binom{n}{k} (\sigma/\mu)^k \{1 + (-1)^k\} \Gamma((k+1)/s)}{2\Gamma(1/s)}. \end{aligned}$$

Tada, jednostavno za prvi i drugi moment imamo:

$$\mathrm{E}(X) = \mu$$

$$\mathrm{E}(X^2) = \mu^2 + \frac{\sigma^2\Gamma(3/s)}{\Gamma(1/s)}.$$

Nakon toga, dolazimo do n-tog centralnog momenta

$$\mathrm{E}((X-\mu)^n) = \sigma^n\,(1+(-1)^n)\int_{\mu}^{\infty}\left(\frac{x-\mu}{\sigma}\right)^n \frac{s\exp\left(-((x-\mu)/\sigma)^s\right)}{2\sigma\Gamma(1/s)}dx$$

$$= \frac{s\,\sigma^n\,(1+(-1)^n)}{2\sigma\Gamma(1/s)}\int_0^{\infty} z^n\exp(-z^s)dz$$

$$= \frac{\sigma^n\,(1+(-1)^n)}{2\sigma\Gamma(1/s)}\int_0^{\infty} y^{(n+1)/(s-1)}\exp(-y)dy$$

$$= \frac{\sigma^n\,(1+(-1)^n)\,\Gamma((n+1)/s)}{2\Gamma(1/s)}.$$

Posebno, drugi centralni moment, tj. varijanca generalizirane Gaussove distribucije je :

$$\mathrm{E}((X-\mu)^2) = \frac{\sigma^2\Gamma(3/s)}{\Gamma(1/s)} = VarX.$$

2.4 Karakteristična funkcija

Fokus ovog odjeljka je na karakterističnoj funkciji generalizirane Gaussove distribucije. Karakteristična funkcija za $\mu = 0$ i $\sigma = 1$ se može napisati u sljedećem obliku:

- za s>0:

$$\phi_s(t) = 4K\int_0^{\infty}\cos(tx)e^{-\frac{x^s}{2}}dx, \quad t\in\mathbb{R}$$

- za $s \in (0,2]$:

$$\phi_s(t) = \mathrm{E}\left(e^{-\frac{t^2V_{G,s}^2}{2}}\right), \quad t\in\mathbb{R},$$

gdje je funkcija gustoće slučajne varijable $V_{G,s}$ dana u sljedećoj propoziciji.

Propozicija 1. *Označimo s $X_q \sim \mathcal{N}_q(0,1)$ varijablu koja ima generaliziranu Gaussovu distribuciju s parametrima 0 i 1, tj. funkcija gustoće slučajne varijable je sljedećeg oblika:*

$$f_{X_q}(x) = Ke^{-|x|^q}, \quad x \in \mathbb{R}, \quad q > 0.$$

Za bilo koje $0 < q \leq s \leq 2$ vrijedi

$$X_q \stackrel{d}{=} V_{s,q} \cdot X_{\frac{2q}{s}},$$

gdje je $V_{s,q}$ pozitivna slučajna varijabla neovisna o $X_{\frac{2q}{s}} \sim \mathcal{N}_{\frac{2q}{s}}(0,1)$ i gdje $\stackrel{d}{=}$ označava jednakost distribucija. Štoviše, $V_{s,q}$ ima sljedeća svojstva:

- *$V_{s,q}$ je neograničena slučajna varijabla za $s < 0$ i $V_{s,q} = 1$ za $s = 2$*
- *za $s > 2$, $V_{s,q}$ je kontinuirana slučajna varijabla sa sljedećom funkcijom gustoće:*

$$f_{V_{s,q}}(v) = \frac{1}{2\pi} \frac{\Gamma(\frac{s}{2q})}{\Gamma(\frac{1}{q})} \int_{\mathbb{R}} v^{-it-1} \frac{2^{\frac{it}{q}}\Gamma(\frac{it+1}{q})}{2^{\frac{its}{2q}}\Gamma(\frac{p(it+1)}{2q})} dt, \quad v > 0.$$

Dokaz se nalazi u [2], poglavlje 4, str. 11.

Kao posljedica pozitivne definitnosti, za $\phi_s(t)$ za $s \in (0,2]$ imamo već definirani oblik. Međutim, za $s > 2$ čini se da se $\phi_s(t)$ ne može napisati u podlognijem obliku i najbolje pojednostavljenje koje se može izvesti je trivijalna simetrizacija koja pretvara Fourierovu transformaciju u kosinusnu transformaciju. Pored toga, kosinusna reprezentacija nam dopušta da pojednostavimo implementaciju za numeričko izračunavanje $\phi_s(t)$.

2.4.1 Analitička svojstva karakteristične funkcije

Važno pitanje, posebno za numeričke metode, je kada se karakteristična funkcija slučajne varijable može predstaviti kao red potencija oblika

$$\sum_{k=0}^{\infty} \frac{(it)^k}{k!} \mathrm{E}(X^k)?$$

Gornji izraz je posebno koristan jer momenti generalizirane Gaussove distribucije postoje za svaki k.

Propozicija 2. *$\phi_s(t)$ je realna analitička funkcija za*

- *$t \in \mathbb{R}$ za $s > 1$*
- *$|t| < \frac{1}{2}$ za $s = 1$.*

Za $s < 1$ funkcija $\phi_s(t)$ nije analitička realna funkcija.

Dokaz se nalazi u [2], poglavlje 5.2, str. 18.

Rezultati prethodne propozicije također dovode do zaključka da za $s > 1$ funkcija izvodnica momenata od $X_s \sim \mathcal{N}_s(0,1)$, $M_s(t) = \mathrm{E}(e^{tX_s})$ postoji za sve $t \in \mathbb{R}$.

2.4.2 Asimptotsko ponašanje karakteristične funkcije

U ovom poglavlju proučit ćemo asimptotsko ponašanje od $\phi_s(t)$ kad $t \to \infty$. Zapravo, sljedeći rezultat daje asimptotsko ponašanje ne samo od $\phi_s(t) = \mathrm{E}\left(e^{-\frac{t^2 V_{G,s}^2}{2}}\right)$, nego i općenitije funkcije

$$t \mapsto \mathrm{E}\left(V_{G,s}^m e^{-\frac{t^2 V_{G,s}^2}{2}}\right), \quad m > 0.$$

Analiza ove funkcije također omogućuje pronalaženje asimptotskih ponašanja na derivacijama višeg reda funkcije $\phi_s(t)$. Na primjer, derivacija prvog reda može se povezati s prethodnom funkcijom na sljedeći način:

$$\phi_s'(t) = -t\mathrm{E}\left(V_{G,s}^2 e^{-\frac{t^2 V_{G,s}^2}{2}}\right).$$

Propozicija 3. *Neka je $m \in (0, \infty)$. Tada*

$$\lim_{t\to\infty} t^{m+s+1}\mathrm{E}\left(V_{G,s}^m e^{-\frac{t^2 V_{G,s}^2}{2}}\right) = A_m = \frac{s}{2}\Gamma\left(\frac{m+s+1}{2}\right)2^{\frac{m+2s}{2}-\frac{s+1}{s}}.$$

Dokaz se nalazi u [2], poglavlje 5.4, str. 18.

Koristeći Propoziciju 3, možemo dati točno ponašanje repova za $\phi_s(t)$.

Propozicija 4. *Za $s \in (0, 2)$*

$$\lim_{t\to\infty} \phi_s(t)t^{s+1} = A_0,$$

gdje je A_0 dano u prethodnoj propoziciji. Štoviše, za $0 < q, s < 2$ i neki $\sigma > 0$

$$\lim_{t\to\infty} \frac{\phi_q(\sigma t)}{\phi_s(t)} = \begin{cases} 0, & q > s \\ \frac{1}{\sigma^{-q-1}}, & q = s \\ \infty, & q < s \end{cases}.$$

Dokaz se nalazi u [2], poglavlje 5.4, str. 19.

3 Procjena parametara

U ovom radu je dan jednostavan slučajni uzorak $X_1, ..., X_n$ iz generalizirane Gaussove distribucije. Želimo proučiti svojstva procjenitelja nepoznatih parametara i njihovu ovisnost o parametrima i veličini uzorka. U tu svrhu, koristimo metodu momenata i metodu maksimalne vjerodostojnosti za procjenu parametara funkcije gustoće generalizirane Gaussove distribucije.

3.1 Metoda momenata

U ovom poglavlju, funkciju gustoće slučajne varijable X koja ima generaliziranu Gaussovu distribuciju ćemo definirati na sljedeći način:

$$f(x,\theta) = [2\Gamma(1+1/s)A(s)]^{-1}\exp\{-[|x-\mu|/A(s)]^s\}, \qquad x \in \mathbb{R},$$

gdje je $A(s) = [\sigma^2\Gamma(1/s)\Gamma(3/s)]^{1/2}$.

Procjenu parametra $\theta = (\mu, \sigma, s)$ metodom momenata označimo s $\tilde{\theta}_n = (\tilde{\mu}_n, \tilde{\sigma}_n^2, \tilde{s}_n)$, a odgovarajući procjenitelj s $\tilde{\Theta}_n = (\tilde{M}_n, \tilde{V}_n, \tilde{S}_n)$. Stoga, $\tilde{\mu}_n$ i $\tilde{\sigma}_n^2$ su jednostavno aritmetička sredina i varijanca uzorka. Procjena $\tilde{s}_n$ se dobiva iz r-tog centralnog momenta (r paran i ≥ 4):

$$\frac{1}{n}\sum_{i=1}^{n}(X_i - \tilde{\mu}_n)^r = \frac{[A(\tilde{s}_n)]^r\Gamma[(r+1)/\tilde{s}_n]}{\Gamma(1/\tilde{s}_n)}.$$

Točan izraz za s nije moguće odrediti, i zbog toga se koristi iterativna tehnika za pronalaženje korijena, s ograničenjem da je $s > 0$. $\tilde{\Theta}_n$ je asimptotski normalan, odnosno

$$\sqrt{n}(\tilde{\Theta}_n - \theta) \xrightarrow{d} \mathcal{N}(0, \Sigma(\theta)),$$

što se može pronaći u [4].

Asimptotska svojstva procjenitelja metodom momenata su, dakle, u potpunosti određena sa $\Sigma(\theta)$, koja u slučaju $r = 4$ ima sljedeći oblik:

$$\Sigma(\theta) = \begin{bmatrix} \epsilon_\mu & 0 & 0 \\ 0 & \epsilon_{\sigma^2} & \epsilon_{\sigma^2 s} \\ 0 & \epsilon_{\sigma^2 s} & \epsilon_s \end{bmatrix},$$

gdje su elementi matrice dani sljedećim izrazima:

$$\epsilon_\mu = \sigma^2$$

$$\epsilon_{\sigma^2} = \sigma^4[g_4(s) - 1]$$

$$\epsilon_s = \{[g_4'(s)]^2\}^{-1}\{4g_4(s)[g_4'(s) - g_6(s)] + g_8(s) - g_4^2(s)\}$$

$$\epsilon_{\sigma^2 s} = [\sigma^2/g_4'(s)][g_4(s) - 2g_4^2(s) + g_6(s)].$$

Funkcija $g_r(s)$ ovisi samo o r i s, i definirana je izrazom

$$g_r(s) := \sigma^{-r}[A(s)]^r\Gamma(r+1/s)\Gamma(1/s).$$

Primijetimo da je procjenitelj metodom momenata parametra μ asimptotski neovisan o procjeniteljima za σ^2 i s, dok su prethodni procjenitelji međusobno ovisni. Asimptotska varijanca procjenitelja od σ^2 se povećava kako se s smanjuje za $s<1$. Na kraju, $\Sigma(\theta)$ ne ovisi o lokacijskom parametru μ, dok asimptotska varijanca procjenitelja od s ne ovisi ni o σ^2.

3.2 Metoda maksimalne vjerodostojnosti

Logaritam funkcije vjerodostojnosti (log-likelihood funkcija) za slučajni uzorak $X_1, ..., X_n$ iz generalizirane Gaussove distribucije dana je izrazom

$$\log L(\mu,\sigma,s) = n\log\left\{\frac{s}{2\sigma\Gamma(1/s)}\right\} - \sum_{i=1}^{n}\left|\frac{X_i-\mu}{\sigma}\right|^2$$

Derivacije log-likelihood funkcije po parametrima μ, σ, s dane su sljedećim izrazima:

$$\frac{\partial \log L}{\partial \mu} = \frac{s}{\sigma^s}\left\{\sum_{X_i\geq\mu}(X_i-\mu)^{s-1} - \sum_{X_i<\mu}(\mu-X_i)^{s-1}\right\}$$

$$\frac{\partial \log L}{\partial \sigma} = -\frac{n}{\sigma} + \frac{s}{\sigma^{s+1}}\sum_{i=1}^{n}|X_i-\mu|^s$$

$$\frac{\partial \log L}{\partial s} = \frac{n}{s}\left\{\frac{1}{s}\Psi\left(\frac{1}{s}\right)+1\right\} - \sum_{i=1}^{n}\left|\frac{X_i-\mu}{\sigma}\right|\log\left|\frac{X_i-\mu}{\sigma}\right|,$$

gdje je $\Psi(x) = \frac{d}{dx}\log\Gamma(x)$.

Za intervalnu procjenu od (μ,σ,s) i testiranje uz njih vezanih hipoteza, potrebna je Fisherova informacijska matrica. Neka je n=1, tada su derivacije drugog reda zadane s:

$$\frac{\partial^2 \log L}{\partial \mu^2} = -\frac{s(s-1)}{\sigma^2}\left|\frac{X-\mu}{\sigma}\right|^{s-2}$$

$$\frac{\partial^2 \log L}{\partial \mu\partial\sigma} = \left(\frac{s}{\sigma}\right)^2 \mathrm{sign}(\mu-X)\left|\frac{X-\mu}{\sigma}\right|^{s-1}$$

$$\frac{\partial^2 \log L}{\partial\mu\partial s} = \frac{1}{X-\mu}\left|\frac{X-\mu}{\sigma}\right|^s \left\{1 + s\log\left|\frac{X-\mu}{\sigma}\right|\right\}$$

$$\frac{\partial^2 \log L}{\partial\sigma^2} = \frac{1}{\sigma^2}\left\{1 - s(s+1)\left|\frac{X-\mu}{\sigma}\right|^s\right\}$$

$$\frac{\partial^2 \log L}{\partial\sigma\partial s} = \frac{1}{\sigma}\left|\frac{X-\mu}{\sigma}\right|^s \left\{1 + s\log\left|\frac{X-\mu}{\sigma}\right|\right\}$$

$$\frac{\partial^2 \log L}{\partial s^2} = -\frac{1}{s^2}\left\{1 + \frac{2}{s}\Psi\left(\frac{1}{s}\right) + \frac{1}{s^2}\Psi\left(\frac{1}{s}\right)\right\} - \left|\frac{X-\mu}{\sigma}\right|^s \left(\log\left|\frac{X-\mu}{\sigma}\right|\right)^2.$$

Sada, koristeći činjenice

$$\begin{aligned}
\mathrm{E}\left(\left|\frac{X-\mu}{\sigma}\right|^m\right) &= \frac{s}{\Gamma(1/s)}\int_0^\infty z^m \exp(-z^s)\,\mathrm{dz} \\
&= \frac{1}{\Gamma(1/s)}\int_0^\infty y^{(m+1)/(s-1)} \exp(-y)\,\mathrm{dy} \\
&= \frac{\Gamma(m+(1/s))}{\Gamma(1/s)}
\end{aligned}$$

$$\begin{aligned}
\mathrm{E}\left(\left|\frac{X-\mu}{\sigma}\right|^m \log\left|\frac{X-\mu}{\sigma}\right|\right) &= \frac{s}{\Gamma(1/s)}\int_0^\infty z^m \log z \exp(-z^s)\,\mathrm{dz} \\
&= \frac{1}{s\Gamma(1/s)}\int_0^\infty y^{(m+1)/(s-1)} \log y \exp(-y)\,\mathrm{dy} \\
&= \frac{\Gamma((m+1)/s)\Psi((m+1)/s)}{s\Gamma(1/s)}
\end{aligned}$$

$$\begin{aligned}
\mathrm{E}\left(\left|\frac{X-\mu}{\sigma}\right|^m \left\{\log\left|\frac{X-\mu}{\sigma}\right|\right\}^2\right) &= \frac{s}{\Gamma(1/s)}\int_0^\infty z^m (\log z)^2 \exp(-z^s)\,\mathrm{dz} \\
&= \frac{1}{s^2\Gamma(1/s)}\int_0^\infty y^{(m+1)/(s-1)} (\log y)^2 \exp(-y)\,\mathrm{dy}
\end{aligned}$$

$$= \frac{\Gamma((m+1)/s)\{\Psi^2((m+1)/s) - \Psi'((m+1)/s)}{s^2\Gamma(1/s)},$$

može se pokazati da su elementi Fisherove informacijske matrice dani sljedećim izrazima:

$$\mathrm{E}\left(-\frac{\partial^2 \log L}{\partial\mu^2}\right) = \frac{s(s-1)\Gamma(1-(1/s))}{\sigma^2\Gamma(1/s)}$$

$$\mathrm{E}\left(-\frac{\partial^2 \log L}{\partial\mu\partial\sigma}\right) = 0$$

$$\mathrm{E}\left(-\frac{\partial^2 \log L}{\partial\mu\partial s}\right) = 0$$

$$\mathrm{E}\left(-\frac{\partial^2 \log L}{\partial\sigma^2}\right) = \frac{s}{\sigma^2}$$

$$\mathrm{E}\left(-\frac{\partial^2 \log L}{\partial\sigma\partial s}\right) = -\frac{1}{s\sigma}\left\{1+\Psi\left(1+\frac{1}{s}\right)\right\}$$

$$\mathrm{E}\left(-\frac{\partial^2 \log L}{\partial s^2}\right) = \frac{1}{s^2}\left[1+\frac{2}{s}\Psi\left(\frac{1}{s}\right)+\frac{1}{s^2}\Psi'\left(\frac{1}{s}\right)\right.$$
$$\left.+\frac{1}{s}\Gamma\left(1+\frac{1}{s}\right)\left\{\Psi^2\left(1+\frac{1}{s}\right)-\Psi'\left(1+\frac{1}{s}\right)\right\}\right].$$

3.3 Asimptotska svojstva prosjeka, minimuma i maksimuma

Ako je $X_1, ..., X_n$ slučajni uzorak iz generalizirane Gaussove distribucije i ako $\overline{X}_n = (X_1 + ... + X_n)/n$ označava aritmetičku sredinu uzorka, tada je prema centralnom graničnom teoremu za nizove nezavisnih jednako distribuiranih slučajnih varijabli slijedi da je asimptotska distribucija standardizirane aritmetičke sredine slučajnog uzorka standardna normalna distribucija. Označimo ekstremne vrijednosti slučajnog uzorka s $M_n = max(X_1, ..., X_n)$ i $m_n = min(X_1, ..., X_n)$.

Primjetimo da je

$$1 - F(t) \sim \frac{1}{2\Gamma(1/s)} \left(\frac{t-\mu}{\sigma} \right)^{1-s} \exp\left\{ - \left(\frac{t-\mu}{\sigma} \right)^{s} \right\}$$

i

$$F(-t) \sim \frac{1}{2\Gamma(1/s)} \left(\frac{t+\mu}{\sigma} \right)^{1-s} \exp\left\{ - \left(\frac{t+\mu}{\sigma} \right)^{s} \right\}$$

za velike vrijednosti nezavisne varijable t. Tada slijedi:

$$\lim_{t\to\infty} \frac{1 - F(t + xg(t))}{1 - F(t)} = \lim_{t\to\infty} \left(\frac{t + xg(t) - \mu}{t-\mu} \right)^{1-s} \exp\left\{ \left(\frac{t-\mu}{\sigma} \right)^{s} - \left(\frac{t + xg(t) - \mu}{\sigma} \right)^{s} \right\},$$

$$\lim_{t\to\infty} \frac{F(-t - xg(t))}{F(-t)} = \lim_{t\to\infty} \left(\frac{t + xg(t) + \mu}{t+\mu} \right)^{1-s} \exp\left\{ \left(\frac{t+\mu}{\sigma} \right)^{s} - \left(\frac{t + xg(t) + \mu}{\sigma} \right)^{s} \right\}.$$

Ako se uzme $g(t) \sim (\sigma^s/s)t^{(1-s)}$ za $t \to \infty$, tada bi se oba prethodna limesa svela na e^{-x} za $t \to \infty$. Dakle, iz Teorema 1.6.2 iz Leadbetter *et al.* (1983) slijedi da moraju postojati normirajuće konstante $a_n, b_n, c_n, d_n > 0$ takve da je

$$\mathrm{P}\{a_n(M_n - b_n) \leq x\} \to \exp\left(-e^{-x}\right),$$

$$\mathrm{P}\{c_n(m_n - d_n) \leq x\} \to 1 - \exp\left(-e^{x}\right).$$

kad $n \to \infty$. Također, može se odrediti oblik normirajućih konstanti. Na primjer, koristeći Korolar 1.6.3 iz Leadbetter *et al.* (1983), može se odrediti da je $a_n = (s/\sigma)(\log n)^{1-(1/s)}$ i $b_n = \sigma(\log n)^{1/s}$.

3.4 Procjena parametara temeljena na simuliranim uzorcima

Unutar ovog odjeljka bavimo se raznim simulacijama uzoraka iz generalizirane Gaussove distribucije. Uvidjet ćemo kako promjene dimenzija uzorka iz generalizirane Gaussove distribucije (n) te promjene broja simuliranih uzoraka (m) utječu na procjenu parametara, točnije, na njihove prosjeke i srednje kvadratne greške procjena. Prosjeke procjena ranije definiranih parametara iz generalizirane Gaussove distribucije ćemo računati pomoću metode maksimalne vjerodostojnosti i metode momenata te im dati oznake: $\overline{\mu}, \overline{\sigma}, \overline{s}$. Osim toga, za svaku prosječnu procjenu parametra, računamo pripadajući korijen srednje kvadratne greške (RMSE) koji daje informaciju o tome koliko u prosjeku procjene odstupaju od stvarnih veličina. Napomenimo kako unutar metode momenata ne računamo prosječnu procjenu parametra s, nego uzimamo procjenu dobivenu metodom maksimalne vjerodostojnosti.
Prethodno opisani postupak ćemo primjeniti na tri specifične generalizirane Gaussove distribucije.

3.4.1 Standardna normalna distribucija

Standardna normalna distribucija je specijalni slučaj generalizirane Gaussove distribucije s parametrima $\mu = 0, \sigma = 1, s = 2$.

Procjena parametara metodom maksimalne vjerodostojnosti

n	**m**	$\overline{\mu}$	**RMSE**	$\overline{\sigma}$	**RMSE**	$\overline{s}$	**RMSE**
1000	100	0.00293	0.02145	1.00334	0.02576	2.00894	0.15879
1000	500	-0.00032	0.02261	1.00353	0.02324	2.01770	0.14192
1000	1000	0.00019	0.02269	1.00267	0.02230	2.01521	0.13923
10000	100	0.00016	0.00718	0.99891	0.00710	1.99386	0.04643
10000	500	$9.70579 \cdot 10^{-5}$	0.00695	1.00009	0.00705	2.00195	0.04210
10000	1000	0.00016	0.00676	0.99988	0.00684	1.99996	0.04473

Tablica 1: Procjena parametara standardne normalne distribucije metodom maksimalne vjerodostojnosti

Iz tablice vidimo da povećanjem dimenzije uzorka prosječne vrijednosti procjene parametra μ teže 0, prosječne vrijednosti procjene σ teže prema 1 i prosječne vrijednosti procjene s teže prema 2. Također, povećanjem dimenzije uzorka, kao i povećanjem broja simuliranih uzoraka, RMSE vrijednosti za $\overline{\mu}, \overline{\sigma}, \overline{s}$ redom teže ka 0.

Procjena parametara metodom momenata
U ovom slučaju, za prosječnu procjenu parametra s uzimamo vrijednost $s = 2$.

n	**m**	$\overline{\mu}$	**RMSE**	$\overline{\sigma}$	**RMSE**
1000	100	-0.00324	0.09713	0.99822	0.03342
1000	500	0.00163	0.09082	0.99954	0.03405
1000	1000	0.00166	0.08954	0.99885	0.03199
10000	100	0.00078	0.02939	0.99892	0.00989
10000	500	0.00029	0.02697	0.99935	0.01012
10000	1000	0.00052	0.02721	1.00030	0.00994

Tablica 2: Procjena parametara standardne normalne distribucije metodom momenata

Iz tablice vidimo da povećanjem dimenzije uzorka prosječna vrijednost procjene parametara μ teži vrijednosti 0, dok $\overline{\sigma}$ ide prema 1. Također, povećanjem dimenzije uzorka, RMSE vrijednosti za $\overline{\mu}$ i $\overline{\sigma}$ idu prema 0.

3.4.2 Laplaceova distribucija

Laplaceova distribucija je specijalni slučaj generalizirane Gaussove distribucije s parametrima $\mu = 0, \sigma = 1, s = 1$.

Procjena parametara metodom maksimalne vjerodostojnosti

n	**m**	$\overline{\mu}$	**RMSE**	$\overline{\sigma}$	**RMSE**	$\overline{s}$	**RMSE**
1000	100	-0.00124	0.03111	1.01708	0.04111	1.00924	0.06110
1000	500	-0.00113	0.03132	1.00491	0.03560	1.00339	0.05830
1000	1000	-0.00029	0.03175	1.00481	0.03510	1.0036	0.05998
10000	100	-0.00011	0.00534	1.00450	0.01275	0.99411	0.01823
10000	500	-0.00064	0.00512	0.99925	0.01189	1.00065	0.01924
10000	1000	$-9.40499 \cdot 10^{-5}$	0.00503	0.99972	0.01032	1.00171	0.01836

Tablica 3: Procjena parametara Laplaceove distribucije metodom maksimalne vjerodostojnosti

Iz tablice vidimo da povećanjem dimenzije uzorka prosječne vrijednosti procjene parametra μ teže 0, prosječne vrijednosti procjene σ teže prema 1 i prosječne vrijednosti procjene s teže prema 1. Također, povećanjem dimenzije uzorka, kao i povećanjem broja simuliranih uzoraka, RMSE vrijednosti za $\overline{\mu}, \overline{\sigma}, \overline{s}$ redom teže ka 0.

Procjena parametara metodom momenata

U ovom slučaju, za prosječnu procjenu parametra s uzimamo vrijednost $s = 1$.

n	m	$\overline{\mu}$	RMSE	$\overline{\sigma}$	RMSE
1000	100	-0.00701	0.04431	1.00539	0.03498
1000	500	0.00269	0.04575	0.99967	0.03810
1000	1000	-0.00072	0.04545	1.00058	0.03617
10000	100	-0.00165	0.01432	1.00227	0.01134
10000	500	-0.00025	0.01430	0.99989	0.01208
10000	1000	0.00023	0.01402	0.99919	0.01192

Tablica 4: Procjena parametara Laplaceove distribucije metodom momenata

Iz tablice vidimo da povećanjem dimenzije uzorka prosječna vrijednost procjene parametara μ teži vrijednosti 0, dok $\overline{\sigma}$ ide prema 1. Također, povećanjem dimenzije uzorka, RMSE vrijednosti za $\overline{\mu}$ i $\overline{\sigma}$ idu prema 0.

3.4.3 Generalizirana Gaussova distribucija s repovima težima od repova normalne distribucuje

Distribucija s težim repovima od normalne ima primjerice, sljedeće vrijednosti parametara $\mu = 0, \sigma = 1, s = 4$.

Procjena parametara metodom maksimalne vjerodostojnosti

n	m	$\overline{\mu}$	RMSE	$\overline{\sigma}$	RMSE	$\overline{s}$	RMSE
1000	100	-0.00255	0.01516	0.99714	0.01856	4.04258	0.35018
1000	500	0.00152	0.01609	1.00082	0.01722	4.05368	0.37604
1000	1000	0.00033	0.01568	1.00188	0.01743	4.08320	0.39637
10000	100	$-7.91401 \cdot 10^{-5}$	0.00613	0.99851	0.00541	3.99678	0.11112
10000	500	-0.00018	0.00591	0.99976	0.00545	4.00195	0.11023
10000	1000	$1.52060 \cdot 10^{-5}$	0.00613	1.00004	0.00558	4.00933	0.11230

Tablica 5: Procjena parametara distribucije s težim repovima od normalne metodom maksimalne vjerodostojnosti

Iz tablice vidimo da povećanjem dimenzije uzorka prosječne vrijednosti procjene parametra μ teže 0, prosječne vrijednosti procjene σ teže prema 1 i prosječne vrijednosti procjene s teže prema 4. Također, povećanjem dimenzije uzorka, kao i povećanjem broja simuliranih uzoraka, RMSE vrijednosti za $\overline{\mu}, \overline{\sigma}, \overline{s}$ redom teže ka 0.

Procjena parametara metodom momenata
U ovom slučaju, za prosječnu procjenu parametra s uzimamo vrijednost $s = 4$.

n	**m**	$\overline{\mu}$	**RMSE**	$\overline{\sigma}$	**RMSE**
1000	100	-0.00261	0.01779	0.99964	0.03306
1000	500	0.00186	0.01882	0.99635	0.03407
1000	1000	0.00058	0.01862	0.99833	0.03283
10000	100	0.00056	0.00625	0.99864	0.01039
10000	500	0.00027	0.00596	0.99964	0.01009
10000	1000	-0.00012	0.00564	1.00019	0.00989

Tablica 6: Procjena parametara distribucije s težim repovima od normalne metodom momenata

Iz tablice vidimo da povećanjem dimenzije uzorka prosječna vrijednost procjene parametara μ teži vrijednosti 0, dok $\overline{\sigma}$ ide prema 1. Također, povećanjem dimenzije uzorka, RMSE vrijednosti za $\overline{\mu}$ i $\overline{\sigma}$ idu prema 0.

Popis slika

Popis tablica

Literatura

[1] M. Benšić, N. Šuvak, *Uvod u vjerojatnost i statistiku*, Sveučilište J.J. Strossmayera, Odjel za matematiku, 2014.

[2] A. Dytso, R. Bustin, H. Vincent Poor, S. Shamai, *Analytical properties of generalized Gaussian distribution*, Department of Electrical Engineering, Princeton University, Princeton, USA, 2018.

[3] S. Nadarajah, *A generalized normal distribution*, University of Nebraska, Lincoln, USA, 2005.

[4] M. K. Varanasi, B. Aazhang, *Parametric generalized Gaussian density estimation*, Department of Electrical and Computer Engineering, Rice University, Houston, Texas, 1989.

Sažetak

U ovom smo radu definirali generaliziranu Gaussovu distribucije te analizirali njezina svojstva. Nakon toga, uz dani slučajni uzorak, smo napravili procjenu parametara koristeći metodu momenata i metodu maksimalne vjerodostojnosti. Osim toga, dotakli smo se asimptotike i karakteristične funkcije. Unutar same karakteristične funkcije, definirali smo njezina analitička svojstva. Naposljetku, napravili smo razne simulacije slučajnih uzoraka iz generalizirane Gaussove distribucije, na temelju njih procijenili parametre te distribucije i procjene usporedili s obzirom na dimenziju uzorka.

Ključne riječi: generalizirana Gaussova distribucija, procjena parametara, asimptotika, karakteristična funkcija, slučajni uzorak, simulacija

Generalized Gaussian distribution

Summary

In this paper, we defined the generalized Gaussian distribution and analyzed on its properties. After that, with the given random sample, we estimated the parameters using the method of moments and the maximum likelihood method. In addition, we touched on asymptotics and the characteristic function. Within the characteristic function itself, we have defined its analytical properties. Finally, we made various simulations of random samples from the generalized Gaussian distribution, based on them we estimated the parameters of that distribution and compared the estimates with respect to the sample size.

Keywords: generalized Gaussian distribution, parameter estimation, asymptotics, characteristic function, random sample, simulation

Životopis

Rođen sam 31. listopada 1996. godine u Osijeku. Pohađao sam osnovnu školu "Bratoljuba Klaića" u Bizovcu. 2011. godine upisujem Opću gimnaziju u Osijeku koju završavam 2015. godine. Iste godine upisujem se na Odjelu za matematiku Sveučilišta Josipa Jurja Strossmayera u Osijeku na preddiplomskom studiju matematike kojeg završavam 2019. godine i stječem naziv prvostupnika matematike s temom završnog rada *Matematika u sportskom klađenju* pod mentorstvom prof. dr. sc. Mirte Benšić. Nakon toga, 2019. godine upisujem diplomski studij matematike, smjer Financijska matematika i statistika.