МАТЕМАТИЧКИ ФАКУЛТЕТ
УНИВЕРЗИТЕТ У БЕОГРАДУ

КАРАКТЕРИЗАЦИЈА РАСПОДЕЛА ПОМОЋУ СТАТИСТИКА ПОРЕТКА

-Мастер рад-

Студент **Никола Јевтић**

Ментор **Др Слободанка Јанковић**

Октобар, 2015.

Садржај

§Увод

1 Увод

Карактеризација функција расподела *(енг. Characterization of Probability Distributions)* спада у групу математичких теорија, чији развој и даље траје, тј. и данас је предмет изучавања многих научника. Појам карактеризације расподеле је најједноставније објаснити примером. Наиме, уколико нам је дата популација X, и њена одговарајућа функција расподеле $F(x)$, тада на основу типа расподеле можемо издвојити нека њена својства у класу $\mathfrak{S}$. Карактеризација представља обрнут поступак, уколико су нам познате неке од особина популације X из класе $\mathfrak{S}$, да ли можемо закључити која је расподела $F(x)$ у питању. Како различите расподеле имају многа својства, тако и теорема о карактеризацији постоји много. У овом раду биће описана карактеризација заснована на особинама статистика поретка. Тако да су у првом поглављу дефинисани појам статистика поретка, њихове функције и густине расподела, као и њихове важне особине које се користе у теоремама карактеризације. Такође, највећу пажњу посвећујемо карактеризацији експоненцијалне расподеле (поглавље 3.1), као расподели чије су теореме карактеризације најбројније, и из које се одређеним монотоним трансформацијама могу добити и друге расподеле, за које ће важити сличне карактеризационе особине (поглавље 3.5). Поглавље 3.2 је посвећено најважнијим теоремама карактеризације експоненцијалне расподеле, помоћу једнако расподељених статистика поретка. Да би се касније у поглављима 3.3 и 3.4, у којима је представљена карактеризација на основу независности статистика поретка и очекивања статистика поретка, направило међусобно поређење, тј. њихова повезаност са теоремама из поглавља 3.2. Последња два поглавља посвећена су карактеризацији неких других расподела.

§Статистике поретка

2 Статистике поретка

2.1 Увод

Како је већ споменуто у претходном поглављу, теорема о карактеризацији има јако пуно, засноване су на различитим својствима функција расподела и особинама њихових случајних величина *(одсуство меморије, независност...)*, као и на особинама њима одговарајућих статистика поретка, функција ризика, математичког очекивања, ентропија, итд. У овом раду су описане теореме карактеризације расподеле добијене на основу особина статистика поретка. За овај уводни део користимо литературу [1],[2] и [3]. Типови карактеризација помоћу статистика поретка поређане су у групама на основу својства:

- *Једнакости расподела* одговарајућих функција статистика поретка

- *Независности* функција расподела статистика поретка

- *Особина момената* статистика поретка.

Све теореме наведене су по хронолошком реду и осликавају поступни развој теорије карактеризације, кроз њихове полазне идеје, како се међусобно унапређују, и колика је заједничка веза између горе наведених група карактеризација помоћу особина статистика поретка. Зато је у уводном делу пажња посвећена статистикама поретка (појам, дефиниција, функција расподеле - једне статистике поретка или заједничка расподела две статистике поретка, математичко очекивање). За почетак, из нотацијских разлога, кратак осврт на основне концепте теорије вероватноће.
Под ознаком $(\Omega, \mathcal{A}, P)$ подразумевамо простор вероватноћа, где је Ω простор догађаја, $\mathcal{A} \subseteq \Omega$ скуп догађаја ($\sigma-$алгебра) и P је мера вероватноће дефинисана на $\mathcal{A}$. Великим словом X означићемо случајну промењиву, која представља пресликавање на простору Ω и која је мерљива у односу на скуп $\mathcal{A}$, тј. за сваки реални број x, скуп величина из Ω за које важи $X < x$ припада скупу $\mathcal{A}$.
За дату случајну величину X, функцију

$$F(x) = P\{X \leq x\}$$

називамо *функцијом расподеле* случајне величине X. Уколико је потребно нагласити на коју се случајну величину односи расподела, користићемо и ознаку $F_X(x)$, нпр. функцију расподеле минималне статистике поретка означићемо са $F_{X_{(1,n)}}(x)$ (види 2.2). За овако дефинисану функцију расподеле важи да је неопадајућа, непрекидна здесна и $F(-\infty) = 0$, $F(+\infty) = 1$. Како посебну пажњу посвећујемо карактеризацији експоненцијалне расподеле, рећићемо још да је функција $F(x)$ апсолутно непрекидна уколико је непрекидна и диференцијабилна. У том случају можемо дефинисати функцију $f(x) = F'(x)$ за свако x, коју називамо функцијом густине.

Појам и дефиниција статистика поретка:

Претпоставимо да су $X_1, X_2, ..., X_n$ n независних и једнако расподељених случајних величина са функцијом расподеле $F(x)$. Одговарајуће *статистике поретка* су случајне величине X_i поређане у неопадајућем поретку. Такав низ називамо *варијациони низ.* Најмањи међу X_i-јевима означићемо са $X_{(1,n)}$, други најмањи по реду означићемо са $X_{(2,n)}$,. . . ,и најзад, највећи означавамо са $X_{(n,n)}$. Отуда

$$X_{(1,n)} \leq X_{(2,n)} \leq ... \leq X_{(n,n)}.$$

Величина која се налази на k-тој позицији, називамо k-та статистика поретка. Статистике поретка $X_{(1,n)}$ и $X_{(n,n)}$ још називамо и минималном, односно максималном статистиком поретка:

$$X_{(1,n)} = \min(X_1, X_2, ..., X_n)$$

$$X_{(n,n)} = \max(X_1, X_2, ..., X_n)$$

Уколико су вредности X_i и X_j једнаке за неко $i \neq j$, онда не правимо разлику која од величина стоји испред друге у варијационом низу. Уколико је $F(x)$ непрекидна тада је $P(X_i = X_j) = 0$ за неко $i \neq j$, тј. у овом случају се статистике поретка разликују са вероватноћом 1. Како у овом раду посебну пажњу посвећујемо карактеризацији експоненцијалне расподеле, у већини случајева претпостављамо да је функција расподеле $F(x)$ непрекидна, осим уколико није другачије напоменуто (дискретни случај). Особине раподеле функција ових статистика поретка показаће се од посебне важности и послужиће као идеја за карактеризацију експоненцијалне расподеле. Зато у даљем тексту говоримо о функцији расподеле статистика поретка. Симболом $F_{X_{(k,n)}}(x)$ (скраћено $F_{(k,n)}(x)$) означићемо функцију расподеле $k-$те статистике поретка $X_{(k,n)}$, док $f_{X_{(k,n)}}(x)$ (скраћено $f_{(k,n)}(x)$) представља њену одговарајућу густину расподеле. У наредна два поглавља су описане функције расподела једне статистике поретка $X_{(k,n)}$, и заједничка расподела две статистике поретка $X_{(i,n)}$ и $X_{(j,n)}$, њихова извођења и докази, одговарајуће густине расподела и њихове особине.

2.2 Расподела статистике поретка

У овом одељку изводимо формулу за функцију расподеле $k-$те статистике поретка, под претпоставкама наведеним у претходном делу: $X_1, X_2, \ldots, X_n$ низ независних и једнако расподељених случајних величина из популације X са непрекидном функцијом расподеле $F(x)$. Тада важи

Теорема 2.1. *Нека су $X_{(1,n)} \leq X_{(2,n)} \leq \ldots \leq X_{(n,n)}$ статистике поретка случајног узорка обима n из расподеле непрекидног типа, чија је функција расподеле $F(x)$ и њој одговарајућа густина расподеле $f(x)$. Тада је функција расподеле $k-$те статистике поретка дата са*

$$F_{X_{(k,n)}}(x) = \sum_{j=k}^{n} \binom{n}{j} F^j(x)(1-F(x))^{n-j},$$

где је $F(x) = P(X \leq x)$.

Доказ. Нека је I случајна величина која представља број случајних величина $X_1, X_2, \ldots, X_n$ које су мање или једнаке x. Дакле, за сваки од $X_1, X_2, \ldots, X_n$ рећићемо да је $X_k \leq x$ успешан догађај, и $X_k \geq x$ неуспешан догађај за $0 \leq k \leq n$. Тада I представља број успешних догађаја од n покушаја. Вероватноћа успеха је $p = P(X \leq x) = F(x)$ јер су случајне величине једнако расподељене. Због њихове независности вероватноћа да $k-$та статистика поретка буде мања или једнака x једнака је догађају да најмање k од n случајних величина $X_1, X_2, \ldots, X_n$ буде мање од x. Одатле, I има биномну расподелу са параметрима n и p, $I \sim \mathcal{B}(n,p)$, где је $p = F(x)$. Дакле, догађај да $X_{(k,n)} \leq x$ је еквивалентан догађају $I \geq k$, тј. барем k случајних величина је мање или једнако од x. Сада можемо одредити функцију расподеле $X_{(k,n)}$:

$$\begin{aligned} F_{X_{(k,n)}}(x) &= P(X_{(k,n)} \leq x) \\ &= P(I \geq k) \\ &= \sum_{j=k}^{n} \binom{n}{j} p^j (1-p)^{n-j} \\ &= \sum_{j=k}^{n} \binom{n}{j} F^j(x)(1-F(x))^{n-j} \quad -\infty < x < +\infty \qquad (2.2.1) \end{aligned}$$ ∎

Последица 2.1. *Користећи формулу 2.2.1 можемо одредити функцију расподеле минималне и максималне статистике поретка*

$$F_{X_{(1,n)}}(x) = 1-(1-F(x))^n, \ -\infty < x < \infty, \qquad (2.2.2)$$

$$F_{X_{(n,n)}}(x) = F^n(x), \ -\infty < x < \infty. \qquad (2.2.3)$$

Доказ. На основу Теореме(2.1) имамо

$$\begin{aligned}
F_{X_{(1,n)}}(x) &= \sum_{j=1}^{n}\binom{n}{j}F^j(x)(1-F(x))^{n-j}\\
&= \sum_{j=0}^{n}\binom{n}{j}F^j(x)(1-F(x))^{n-j} - (1-F(x))^n\\
&= 1-(1-F(x))^n, \ \ -\infty < x < \infty,\\
F_{X_{(n,n)}}(x) &= \sum_{j=n}^{n}\binom{n}{j}F^j(x)(1-F(x))^{n-j}\\
&= \binom{n}{n}F^n(x)(1-F(x))^0\\
&= F^n(x), \ \ -\infty < x < \infty \ \ \blacksquare
\end{aligned}$$

За случајну величину апсолутно непрекидног типа, можемо одредити густину расподеле $k-$те статистике поретка.

Теорема 2.2. *Нека су $X_{(1,n)} \le X_{(2,n)} \le \ldots \le X_{(n,n)}$ статистике поретка случајног узорка обима n из расподеле непрекидног типа, чија је фунцкија расподеле $F(x)$ и њој одговарајућа густина расподеле $f(x)$. Тада је густина расподеле $k-$те статистике поретка дата са*

$$f_{X_{(k,n)}}(x) = k\binom{n}{k}f(x)F^{k-1}(x)(1-F(x))^{n-k}, \tag{2.2.4}$$

где је $F(x) = P(X \le x)$.

Доказ. Густина расподеле $f_{X_{(k,n)}}(x)$ случајне величине $X_{(k,n)}$ је

$$f_{X_{(k,n)}}(x) = \frac{d}{dx}F_{X_{(k,n)}}(x).$$

Тада, из Теореме (2.1) имамо

$$\begin{aligned}
f_{X_{(k,n)}}(x) =& \frac{d}{dx}\sum_{j=k}^{n}\binom{n}{j}F^j(x)(1-F(x))^{n-j}\\
=& \sum_{j=k}^{n}\binom{n}{j}\Big(jF^{j-1}(x)(1-F(x))^{n-j}f(x)\\
&-(n-j)F^j(x)(1-F(x))^{n-j-1}f(x)\Big)\\
=& \sum_{j=k}^{n}j\binom{n}{j}F^{j-1}(x)(1-F(x))^{n-j}f(x)\\
&-\sum_{j=k}^{n}(n-j)\binom{n}{j}F^j(x)(1-F(x))^{n-j-1}f(x),
\end{aligned}$$

Из прве суме извлачимо први сабирак, и како је за $k = n$ сабирак у последњој суми једнак нули, остаје

$$\begin{aligned}
f_{X_{(k,n)}}(x) =& k\binom{n}{k}f(x)F^{k-1}(x)(1-F(x))^{n-k} \\
&+ \sum_{j=k+1}^{n} j\binom{n}{j}F^{j-1}(x)(1-F(x))^{n-j}f(x) \\
&- \sum_{j=k}^{n-1}(n-j)\binom{n}{j}F^{j}(x)(1-F(x))^{n-j-1}f(x) \\
=& k\binom{n}{k}f(x)F^{k-1}(x)(1-F(x))^{n-k} \\
&+ \sum_{j=k}^{n-1}(j+1)\binom{n}{j+1}F^{j}(x)(1-F(x))^{n-j-1}f(x) \\
&- \sum_{j=k}^{n-1}(n-j)\binom{n}{j}F^{j}(x)(1-F(x))^{n-j-1}f(x)
\end{aligned}$$

Како је

$$(j+1)\binom{n}{j+1} = \frac{n!}{j!(n-j-1)!} = (n-j)\frac{n!}{j!(n-j)!} = (n-j)\binom{n}{j}$$

суме се потиру, чиме је доказ завршен ■

Према Теореми 2.1, функција расподеле $X_{(k,n)}$ представљена је у облику репа Биномне расподеле (сума полази од k), са вероватноћом $F(x)$. Осим ове репрезентације, функцију расподеле, на основу следеће леме, можемо приказати у облику Пирсонове непотпуне Бета функције.

Лема 2.1. *Нека је $F(x)$ функција расподеле за $x \geq 0$, тада важи*

$$\sum_{j=k}^{n}\binom{n}{j}F^{j}(x)(1-F(x))^{n-j} = I_{F(x)}(k, n-k+1) \tag{2.2.5}$$

где је

$$I_x(a,b) = \frac{1}{\beta(a,b)}\int_0^x t^{a-1}(1-t)^{b-1}dt$$

непотпуна Бета функција, и важи

$$\frac{1}{\beta(k, n-k+1)} = k\binom{n}{k}.$$

Доказ. Једнакост формуле (2.2.5) изводимо здесна на лево користећи парцијалну интеграцију

$$I = \int_0^{F(x)} t^{k-1}(1-t)^{n-k}dt$$

нека је

$$u = (1-t)^{n-k} \qquad du = -(n-k)(1-t)^{n-k-1}$$

$$dv = t^{k-1}dt \qquad v = \frac{1}{k}t^k$$

тада

$$I = \frac{1}{k}F^k(x)(1-F(x))^{n-k} + \frac{n-k}{k}\int_0^{F(x)} t^k(1-t)^{n-k-1}dt.$$

У следећем кораку, за

$$u = (1-t)^{n-k-1} \qquad du = -(n-k-1)(1-t)^{n-k-2}$$

$$dv = t^k dt \qquad v = \frac{1}{k+1}t^{k+1}$$

добијамо

$$\begin{aligned} I = \frac{1}{k}F^k(x)(1-F(x))^{n-k} + \frac{n-k}{k}\Bigg(&\frac{1}{k+1}F^{k+1}(x)(1-F(x))^{n-k-1} + \\ &+ \frac{n-k-1}{k+1}\int_0^{F(x)} t^{k+1}(1-t)^{n-k-2}dt\Bigg) = \\ = \frac{1}{k}F^k(x)(1-F(x))^{n-k} + &\frac{n-k}{k(k+1)}F^{k+1}(x)(1-F(x))^{n-k-1} + \\ &+ \frac{(n-1)(n-k-1)}{k(k+1)}\int_0^{F(x)} t^{k+1}(1-t)^{n-k-2}dt. \end{aligned}$$

Понављајући поступак $n-k$ пута, у последњем кораку добијамо

$$\begin{aligned} I = \frac{1}{k}F^k(x)(1-F(x))^{n-k} &+ \frac{n-k}{k(k+1)}F^{k+1}(x)(1-F(x))^{n-k-1} + \\ &+ \frac{(n-k)(n-k-1)}{k(k+1)(k+2)}F^{k+2}(x)(1-F(x))^{n-k-2} + \\ &+ \ldots + \frac{(n-k)(n-k-1)\cdots 1}{k(k+1)(k+2)\cdots(n-1)}\int_0^{F(x)} t^{n-1}dt. \end{aligned}$$

Решавањем последњег интеграла добијамо суму

$$(2.2.6)\quad I = \frac{1}{k}F^k(x)(1-F(x))^{n-k} + \frac{n-k}{k(k+1)}F^{k+1}(x)(1-F(x))^{n-k-1} +$$

$$+ \ldots + \frac{(n-k)!}{k(k+1)\cdots n}F^n(x).$$

Добијени резултат (2.2.6) можемо записати у скраћеном облику

$$(2.2.7)\qquad I = \frac{1}{k}\sum_{j=k}^{n}\frac{k!}{j!}\frac{(n-k)!}{(n-j)!}F^j(x)(1-F(x))^{n-j}.$$

Дакле, интеграл из једнакости (2.2.5), користећи (2.2.7), можемо записати као

$$\begin{aligned}
I_{F(x)}(k, n-k+1) &= k\binom{n}{k}\int_0^{F(x)} t^{k-1}(1-t)^{n-k}dt \\
&= k\binom{n}{k}\cdot\frac{1}{k}\sum_{j=k}^{n}\frac{k!}{j!}\frac{(n-k)!}{(n-j)!}F^j(x)(1-F(x))^{n-j} \\
&= \frac{n!}{(n-k)!k!}\cdot\sum_{j=k}^{n}\frac{k!}{j!}\frac{(n-k)!}{(n-j)!}F^j(x)(1-F(x))^{n-j} \\
&= \sum_{j=k}^{n}\frac{n!}{(n-j)!j!}F^j(x)(1-F(x))^{n-j} \\
&= \sum_{j=k}^{n}\binom{n}{j}F^j(x)(1-F(x))^{n-j}.
\end{aligned}$$

Чиме је доказ завршен ■

На основу Леме (2.1), функцију расподеле можемо записати на следећи начин

$$\begin{aligned}
F_{X_{(k,n)}}(x) &= \int_0^{F(x)}\frac{n!}{(k-1)!(n-k)!}t^{k-1}(1-t)^{n-k}dt \\
(2.2.8)\qquad &= I_{F(x)}(k, n-k+1),\ \ -\infty < x < \infty
\end{aligned}$$

Потребно је још нагласити да формула за $F_{X_{(k;n)}}(x)$ у једначини (2.2.8) важи и за дискретну, и непрекидну расподелу популације. Међутим, уколико претпоставимо да је расподела апсолутно непрекидног типа, диференцирањем једнакости (2.2.8) добијамо израз за густину расподеле величине $X_{(k;n)}$, који је идентичан оном у (2.2.4).

Пример 2.1. *Посматрамо стандардну униформну расподелу са функцијом густине* $f(u) = 1,\ 0 \le u \le 1$, *и функцијом расподеле* $F(u) = u,\ 0 \le u \le 1$. *Тада према Теореми(2.1), функција расподеле* $U_{(k,n)}$ $(1 \le k \le n)$ *једнака је*

(2.2.9)

$$F_{U_{(k,n)}}(u) = \sum_{j=k}^{n} \binom{n}{j} u^j (1-u)^{n-1} = \int_0^u \frac{n!}{(k-1)!(n-k)!} t^{k-1}(1-t)^{n-k} dt, \quad 0 \le u \le 1.$$

Диференцирањем једнакости (2.2.9) добијамо функцију густине

$$f_{U_{(k,n)}}(u) = \frac{n!}{(k-1)!(n-k)!} t^{k-1}(1-t)^{n-k}, \ \ 0 \le u \le 1 \qquad (2.2.10)$$

Из горње фунцкије густине можемо израчунати $m-$*ти моменат* $U_{(k,n)}$:

$$\begin{aligned} E(U_{(k,n)}^m) &= \int_0^1 u^m f_{U_{(k,n)}}(u) du \\ &= \beta(k+m, n-k+1) \,/\, \beta(k, n-k+1) \end{aligned}$$

где је $\beta(p,q)$ *потпуна бета функција дифинисана са*

$$\beta(p,q) = \int_0^u t^{p-1}(1-t)^{q-1} dt, \ \ p, q > 0,$$

одакле, за $m = 1$, *добијамо математичко очекивање од* $U_{(k,n)}$

$$E(U_{(k,n)}) = \frac{k}{n+1}.$$

2.3 Заједничка расподела две статистике поретка

При извођењу заједничке функције расподеле две статистике поретка $X_{(i,n)}$ и $X_{(j,n)}$ $(1 \le i < j \le n)$, прво графички представљамо догађај $(x_i < X_{(i,n)} \le x_i + \delta x_i, x_j < X_{(j,n)} \le x_j + \delta x_j)$:

где укупно $i-1$ X_k-ова $(1 \le k \le n)$ је мање или једнако од x_i, тачно један припада интервалу $(x_i, x_i + \delta x_i]$, укупно $j-i-1$ X_k-ова је веће од $x_i + \delta x_i$ а мање или једнако од x_j, интервалу $(x_j, x_j + \delta x_j]$ припада тачно један, и преосталих $n-j$ су већи од $x_j+\delta x_j$. За δx_i и δx_j довољно мало, вероватноћа овог догађаја једнака је

$$
\begin{aligned}
(2.3.1)\quad P\Big(x_i < X_{(i,n)} \le x_i + \delta x_i, x_j < X_{(j,n)} \le x_j + \delta x_j\Big) = \\
=&\frac{n!}{(i-1)!1!(j-i-1)!1!(n-j)!}\Big(F(x_i)\Big)^{i-1} \\
&\times \Big(F(x_i+\delta x_i)-F(x_i)\Big)\Big(F(x_j)-F(x_i+\delta x_i)\Big)^{j-i-1} \\
&\times \Big(F(x_j+\delta x_j)-F(x_j)\Big)\Big(1-F(x_j+\delta x_j)\Big)^{n-j} \\
&+ O\Big((\delta x_i)^2\delta x_j\Big) + O\Big(\delta x_i(\delta x_j)^2\Big)
\end{aligned}
$$

Величине $O\Big((\delta x_i)^2\delta x_j\Big)$ и $O\Big((\delta x_i)^2\delta x_j\Big)$ представљају вероватноћу догађаја да се у интервалима $(x_i, x_i+\delta x_i]$, и $(x_j, x_j+\delta x_j]$, појави више од једне случајне величине X_k $(1 \le k \le n)$. Диференцирањем израза (2.3.1) изводимо формулу заједничке фунцкије густине за $X_{(i,n)}$ и $X_{(j,n)}$ $(1 \le i < j \le n)$

$$
\begin{aligned}
f_{(i,j;n)}(x_i,x_j) =& \lim_{\substack{\delta x_i \to 0 \\ \delta x_j \to 0}} \left(\frac{P(x_i < X_{i,n} \le x_i+\delta x_i, x_j < X_{j,n} \le x_j+\delta x_j)}{\delta x_i \delta x_j}\right) \\
=&\frac{n!}{(i-1)!(j-i-1)!(n-j)!}\Big(F(x_i)\Big)^{i-1} \\
(2.3.2)\qquad &\times \left(\lim_{\delta x_i \to 0}\frac{F(x_i+\delta x_i)-F(x_i)}{\delta x_i}\right)\Big(F(x_j)-F(x_i)\Big)^{j-i-1} \\
&\times \left(\lim_{\delta x_j \to 0}\frac{F(x_j+\delta x_j)-F(x_j)}{\delta x_j}\right)\Big(1-F(x_j)\Big)^{n-j} \\
=&\frac{n!}{(i-1)!(j-i-1)!(n-j)!}\Big(F(x_i)\Big)^{i-1} \\
&\times \Big(F(x_j)-F(x_i)\Big)^{j-i-1}\Big(1-F(x_j)\Big)^{n-j} f(x_i)f(x_j) \\
&\qquad -\infty < x_i < x_j < \infty.
\end{aligned}
$$

Уколико изаберемо $i=1$ и $j=n$, из (2.3.2) можемо одредити заједничку густину расподеле минималне и максималне статистике поретка

$$
\begin{aligned}
(2.3.3)\quad f_{(1,n;n)}(x_i,x_j) = n(n+1)F(x_i)(F(x_i)-F(x_j))^{n-2}f(x_i)f(x_j) \\
-\infty < x_i < x_j < \infty;
\end{aligned}
$$

Слично, за $j = i+1$, из (2.3.2) можемо одредити заједничку густину расподеле суседних статистика поретка, $X_{(i,n)}$ и $X_{(i+1,n)}$ $(1 \le i \le n-1)$

(2.3.4)

$$f_{(i,i+1;n)}(x_i, x_{i+1}) = \frac{n!}{(i-1)!(n-i-1)!}\big(F(x_i)\big)^{i-1}\big(1-F(x_{i+1})\big)^{n-i-1} f(x_i) f(x_{i+1})$$

$$-\infty < x_i < x_{i+1} < \infty.$$

На овај начин, можемо одредити функцију густине у генерализованом случају:
Заједничка функција густине $X_{(i_1;n)}, X_{(i_2;n)}, \dots, X_{(i_k;n)}$ $(1 \le i_1 < i_2 < \dots < i_k \le n, 1 \le k \le n)$ за $x_1 \le x_2 \le \dots \le x_k$ је

$$\begin{aligned} f_{(i_1,i_2,\dots,i_k;n)}(x_1, x_2, \dots, x_k) = & \frac{n!}{(i_1-1)!(i_2-i_1-1)!\cdots(n-i_k)!} \\ & \times \big(F(x_1)\big)^{i_1-1}\big(F(x_2)-F(x_1)\big)^{i_2-i_1-1}\cdots\big(1-F(x_k)\big)^{n-i_k} \\ & \times f(x_1) f(x_2)\cdots f(x_k) \end{aligned} \tag{2.3.5}$$

Уколико дефинишемо $x_0 = -\infty, x_{k+1} = +\infty, i_0 = 0, n_{k+1} = n+1,$ funkcija gustine (2.3.5) може се записати у следећем облику:

$$n!\left[\prod_{j=1}^{k} f(x_j)\right]\prod_{j=0}^{k}\left\{\frac{\big(F(x_{j+1})-F(x_j)\big)^{i_{j+1}-i_j-1}}{(i_{j+1}-i_j-1)!}\right\}.$$

Из претходног следи да је функција густине n статистика поретка

$$f_{X_{(1;n)},X_{(2;n)},\dots,X_{(n;n)}}(x_1, x_2, \dots, x_n) = n!\prod_{j=1}^{n} f(x_j) \quad x_1 \le x_2 \le \dots \le x_n. \tag{2.3.6}$$

Резултат (2.3.6) изгледа очигледно, собзиром да постоји $n!$ могућности рангирања x_i−ова, зато у неким радовима полазимо од фунцкије густине n статистика поретка да би се извела формула за заједничку расподелу две или више статистика поретка [1].

§Карактеризација

3 Карактеризација

Међу теоремама карактеризације помоћу статистика поретка, најзаступљеније су оне које се односе на експоненцијалну расподелу. Тако да се у овом раду претежно бавимо карактеризацијом експоненицјалне расподеле. Монотоним трансформацијама експоненцијалне расподеле добијају се многе друге (униформна, Паретова, Вејбулова расподела, итд.), за које ће, у већини случајева, важити сличне карактеризационе особине као код експоненцијалне расподеле. У уводу овог одељка, описујемо особине експоненцијалне расподеле, битне за њену карактеризацију.

3.1 Експоненцијална расподела

Нека је X случајна величина са функцијом расподеле

(3.1.1) $$F(x) = 1 - e^{-\lambda x}, \ \ x \geq 0, \ \ \lambda > 0.$$

Тада кажемо да случајна величина има екпоненцијалну расподелу са параметром $\lambda > 0$. Из једнакости (3.1.1), следе следећа својства фунцкије $F(x)$:

(S1) $$(1 - F(x))' = -\lambda(1 - F(x)), \quad x > 0, \quad F(0+) = 0;$$

(S2) $$\int_x^{+\infty} (1 - F(z))dz = \frac{1}{\lambda}(1 - F(x)), \quad x > 0, \quad F(0+) = 0;$$

(S3) $$1 - F(x + z) = (1 - F(x))(1 - F(z)) \text{ за свако } x, z \geq 0.$$

Услов $F(0+) = 0$ нас ограђује од случаја дегенерисане експоненцијалне расподеле. Наиме, уколико изоставимо услов $F(0+) = 0$, тада су својства (S1) и (S2) задовољена и у случају функције расподеле облика:

$$F_c(x) = \begin{cases} 1 - ce^{-\lambda x} & \text{за } x > 0 \\ 0 & \text{у супротном,} \end{cases}$$

где је константа $0 \leq c \leq 1$. Наш циљ је, да на основу наведених својстава експоненцијалне расподеле, добијемо теореме њене карактеризације. Како за $c = 0$, својства (S1) и (S2) не важе, услов $F(0+) = 0$ нас ограничава само на недегенерисани случај експоненцијалне расподеле.
Надаље, искључујемо могућност дегенерисане расподеле (у појединим

теоремама та претпоставка је експлицитно наведена, или услов $F(0+) = 0$). Приметимо да својство (S1) можемо записати у следећем облику

$$(\text{S3*}) \qquad P(X \geq x+z | X \geq z) = P(X \geq x) \text{ за свако } x, z \geq 0,$$

које је познато као *својство одсуства меморије (lack of memory)*. Како је, из својстава (S1) и (S2), очигледно да (3.1.1) важи (што сваку особину појединачно чини карактеризационом теоремом), еквиваленција (S3) и (3.1.1) је мање очигледна. У овом раду нећемо директно изводити тај доказ (видети [3, стр.8]), али показаћемо да је (S3) еквивалентан услову Десуове теореме, одакле следи еквивалентност и са (3.1.1). Такође, важно је споменути и домен на коме важи својство (S3). Уколико претпоставимо да (S3), или његова еквивалентна форма (S3*), важи само за позитивне целе бројеве x и z, тада је геометријска расподела једина расподела за коју ово својство важи (ово је пример карактеризације дискретне случајне величине).

Доказаћемо још међусобну еквивалентност својстава $(S_j), j = 1,2,3$. На почетку докажимо еквивалентност особина (S1) и (S2). Интеграцијом (S1), у границама од x до бесконачности, добијамо (S2). Са друге стране, како је лева страна једнакости (S2) непрекидна, а тиме и десна страна, следи да су обе стране диференцијабилне, па диференцирањем (S2) добијамо (S1). Сада је довољно доказати еквивалентност (S2) и (S3), одакле следи међусобна еквивалентност сва три услова. Прво, покажимо да је за функцију $F(x)$, која задовољава услов (S3), интеграбилна функција $1 - F(x)$, тј. да је интеграл

$$\int_0^{+\infty} (1 - F(x))dx = \frac{1}{\lambda}$$

коначан. Заиста, уколико (S3) применимо n пута за $x = y$, добијамо да за свако реално $z > 0$ важи

$$1 - F(nz) = (1 - F(z))^n,$$

али, и за сваки реалан број $x = nz > 0$ важи

$$1 - F(x) = \left(1 - F\left(\frac{x}{n}\right)\right)^n.$$

Отуда, примењујући горњу једнакост за $x = n$, имамо

$$\int_0^{+\infty} (1 - F(x))dx \leq \sum_{n=0}^{+\infty} (1 - F(n)) = \sum_{n=0}^{+\infty} (1 - F(1))^n = \frac{1}{F(1)},$$

где је $F(1) > 0$, јер би у супротном важило да је $F(x) = 0$ за свако x. Сада интеграцијом (S3) по z од 0 до бесконачности добијамо

$$\begin{aligned}\int_0^{+\infty}(1-F(x+z))dz &= \int_0^{+\infty}(1-F(x))(1-F(z))dz\\ &= \frac{1}{\lambda}(1-F(x)).\end{aligned}$$

Увођењем смене $x+z=t$ добијамо да својство (S2) важи. Сада, докажимо да важи и супротно. Претпоставимо да важи својство (S2), из претходног знамо да тада важи и (S1). За произвољну функцију $g(x) = e^{\lambda x}(1-F(x))$ важиће $g'(x) = 0$. Решавањем диференцијалне једначине добијамо да је $g(x) = C$, тј. $\frac{g(x)}{C} = 1$ за неку константу $C \neq 0$, и за свако $x > 0$. Тада важи

$$\frac{g(x+z)}{C} = \frac{g(x)}{C}\frac{g(z)}{C} \text{ за свако } x, z > 0,$$

тј. за свако $x, z > 0$ важи

$$C(1-F(x+z)) = (1-F(x))(1-F(z)).$$

Када $z \to 0$, добијамо да је $C = 1$, одакле видимо да важи (S3).

3.2 Карактеризација експоненцијалне расподеле

Споменули смо да карактеризационе теореме се односе на нека позната својства функција расподела, на основу којих доносимо закључак о самој расподели. Међу најзначајнијим резултатима карактеризације експоненцијалне расподеле је Десуова теорема. Главни мотив за решавање овог проблема је својство (2.2.2). Наиме, уколико је расподела F, случајне величине X, експоненцијално расподељена са параметром $\lambda > 0$, и ако је $(X_1, X_2, ..., X_n)$ случајни узорак из расподеле F, тада из својства (2.2.2) следи

$$\begin{aligned}F_{(1;n)}(x) &= 1-(1-F(x))^n\\ &= 1-e^{-\lambda n x}, \text{ за } x \geq 0.\end{aligned} \tag{3.2.1}$$

Тада,

$$P(nX_{(1;n)} \leq x) = P\left(X_{(1;n)} \leq \frac{x}{n}\right)$$

на основу (3.2.1) важи

$$\begin{aligned}&= 1-e^{-\lambda n\frac{x}{n}}\\ &= 1-e^{-\lambda x}, \text{ за } x \geq 0.\end{aligned} \tag{3.2.2}$$

Дакле, на основу израза (3.2.2) можемо закључити да уколико је случајна величина експоненцијално расподељена са параметром $\lambda > 0$, онда је њена одговарајућа минимална статистика поретка такође експоненцијално расподељена са параметром $\lambda > 0$, тј. једнако је расподељена као и X. Поставља се питање, да ли важи обратно?

Теорема 3.1 (Десу[1]). *Уколико је $F(x)$ недегенерисана функција расподеле, тада су за сваки природан број $n \geq 1$ случајне величине $nX_{(1;n)}$ и X једнако расподељене ако и само ако је $F(x)$ облика (3.1.1).*

Доказ. У претходном разматрању је доказано да су $F_{nX_{(1;n)}}(x)$ и $F_X(x)$ идентичне када је $X \sim \varepsilon(\lambda)$. Сада ћемо доказати импликацију

$$(3.2.3) \qquad F_{nX_{(1;n)}} = F_X(x) \Rightarrow X \sim \varepsilon(\lambda).$$

Једнакост у исказу (3.2.3) можемо записати у следећем облику

$$\left.\begin{aligned} P(nX_{(1;n)} \leq x) = 1 - \left(1 - F\left(\frac{x}{n}\right)\right)^n, \\ F_X(x) = P\left(X \leq x\right); \end{aligned}\right\} \quad \Rightarrow \quad \left(1 - F\left(\frac{x}{n}\right)\right)^n = 1 - F(x)$$

Уколико уведемо ознаку $G(x) = 1 - F(x)$, важи

$$(3.2.4) \qquad G(x) = G^n\left(\frac{x}{n}\right), \text{ или } G^{1/n}(x) = G\left(\frac{x}{n}\right).$$

Услов $nX_{(1;n)} \stackrel{d}{=} X$ је еквивалентан $X_{(1;n)} \stackrel{d}{=} \frac{1}{n}X$, па важи

$$\left.\begin{aligned} P(X_{(1;n)} \leq x) = 1 - (1 - F(x))^n, \\ P\left(\frac{X}{n} \leq x\right) = P(X \leq nx); \end{aligned}\right\} \quad \Rightarrow \quad 1 - F(nx) = (1 - F(x))^n$$

одакле следи

$$(3.2.5) \qquad G(nx) = G^n(x), \text{ за свако } x \geq 0, n \geq 1.$$

На основу својстава (3.2.5) и (3.2.4) можемо закључити да за свака два природна броја m и $n \neq 0$, када је $x = 1$, важи

$$(3.2.6) \qquad G\left(\frac{m}{n}\right) = G^m\left(\frac{1}{n}\right) = G^{\frac{m}{n}}(1).$$

Из (3.2.6) следи да за сваки рационалан број облика $x = \frac{m}{n} \in \mathbb{Q}$, важи

$$G(x) = G^x(1) = e^{x \ln G(1)}.$$

[1] M.M.Desu (1971) - амерички статистичар

Приметимо да је $0 < G(1) < 1$. Заиста, уколико би $G(1) = 0$ или $G(1) = 1$, тада постоји број $k \neq 0$, на основу претходног, за који је $G(k) = 0$ или $G(k) = 1$, што је у контрадикцији са $G(k) = 1 - F(k)$ и чињеницом да је $F(k)$ недегенерисана расподела, тј. $F(0) = 0$. Увођењем смене $\lambda = -\ln G(1)$, где је λ позитиван и коначан број, следи

$$G(x) = e^{-\lambda x} \text{ за сваки рационалан број } x.$$

Докажимо још да горња једнакост важи и за ирационалне бројеве, чиме ће важити за све позитивне реалне бројеве. Како је скуп рационалних бројева свуда густ, постоје низови $\{x_1^{(n)}\}$ и $\{x_2^{(n)}\}$ из скупа $\mathbb{Q}$, за које важи

$$x_1^{(n)} < x < x_2^{(n)} \text{ и } \lim_{n\to+\infty} x_1^{(n)} = \lim_{n\to+\infty} x_2^{(n)} = x,$$

где је x произвољан ирационалан број. Како је $G(x)$ нерастућа функција, следи

$$e^{-\lambda x_1^{(n)}} \leq G(x) \leq e^{-\lambda x_2^{(n)}}.$$

Када $n \to +\infty$, добијамо да је $G(x) = e^{-\lambda x}$ за свако $x \geq 0$, што је и требало доказати ■

Приказани доказ се најчешће појављује у радовима(видети [4],[3]). Поступак доказа, који је описан и у овом раду, су првобитно користили Коши[2] и Дарбу[3], при решавању Кошијевих функционалних једначина (видети [5, стр.237]).

Као што смо већ споменули у претходном поглављу, услов теореме (3.1) и својство (S3) су еквивалентни. Доказали смо да при услову $nX_{(1;n)} \stackrel{d}{=} X$, за свако $n \geq 1$ и $x \geq 0$ важи

$$1 - F(nx) = (1 - F(x))^n, \tag{3.2.7a}$$

$$1 - F(x) = \left(1 - F\left(\frac{x}{n}\right)\right)^n. \tag{3.2.7b}$$

Да из својства (S3) следи (3.2.7a) доказујемо помоћу индукције за $z = x$. Сада је довољно показати да $nX_{(1;n)} \stackrel{d}{=} X$ имплицира (S3), тј. показаћемо да за фунцкију која испуњава услове (3.2.7a,b), важи својство (S3). Прво, покажимо да је таква функција непрекидна за свако $x \geq 0$, и да задовољава (S3) за све рационалне бројеве $x, z > 0$. На основу ових чињеница, следиће да таква функција задовољава (S3) за све реалне бројеве. Под

[2]A.C. Cauchy (1821) - француски математичар

[3]G. Darbaux (1875) - француски математичар

претпоставком да је наша функција неопадајућа и дефинисана на позитивном домену, важиће да је $F(0) = 0$, уз претпоставку да је недегенерисана следи непрекидност у нули, $F(0+) = 0$. Покажимо да својство (S3) важи за све рационалне бројеве $x, z > 0$. Нека је $G(x) = 1 - F(x)$, тада на основу (3.2.7a,b), за све природне бројеве m и n важи

$$G(m+n) = G^{m+n}(1) = G^{m}(1)G^{n}(1) = G(m)G(n),$$

одакле, и за све рационалне бројеве x и z облика $\frac{p}{q}, \frac{s}{t} \in \mathbb{Q}$, важи

$$\begin{aligned} G(x+z) =& G\left(\frac{pt+sq}{qt}\right) = G^{pt}\left(\frac{1}{qt}\right)G^{sq}\left(\frac{1}{qt}\right) \\ =& G\left(\frac{pt}{qt}\right)G\left(\frac{sq}{qt}\right) = G(x)G(z). \end{aligned}$$

Остаје да докажемо непрекидност функције $G(x)$ за произвољно x_0. Нека су x и z рационални бројеви такви да је $x < x_0 < x + z$, и нека $x \to x_0$ и $z \to 0$. На основу претходно доказаног, и $G(0) = 1$, следи да је $G(x_0 + 0) = G(x_0 - 0) = G(x_0)$, тј. функција $G(x)$ је непрекидна у тачки x_0. Овим смо доказали еквивалентност својства *одсуства меморије* и $nX_{(1;n)} \overset{d}{=} X$.

Даља уопштења Десуове теореме, су спровођена заменом услова *за свако* $n \geq 1$, слабијим условима. У доказу наредне теореме користимо следећу лему.

Лема 3.1. *Нека су n_1 и n_2 ($n_1 < n_2$) природни бројеви такви да је $\ln n_1/\ln n_2$ ирационалан број. Тада за произвољне целе бројеве m и n (било позитивне, или негативне), скуп бројева $z_{m,n} = m\ln n_1 + n\ln n_2$ је свуда густ на реалној правој.*

Доказ. Претпоставимо супротно. Нека је z реалан број за који постоји најмањи $z^*_{m,n}$ такав да је $z < z^*_{m,n}$ и $z^*_{m,n} \neq z$. Другим речима, интервал $(z, z^*_{m,n})$ не садржи бројеве облика $z_{m,n}$, тј. за сваки $z_{u,v} < z^*_{m,n}$ важи $z_{u,v} \leq z$ (нпр. такав број би био $z_{u,v} = z^*_{m,n} - \ln n_1$). Означимо са $z^{\backprime}_{u,v}$ највећи од бројева из $\{z_{m,n}\}$, који не прелази z. Тада, никоја два броја из $\{z_{m,n}\}$ нису на мањем растојању од $\alpha = z^*_{m,n} - z^{\backprime}_{u,v} > 0$, јер разлика два члана скупа $\{z_{m,n}\}$ припада том скупу. У супротном би постојали бројеви $z_{m_1,n_1} > z_{m_2,n_2}$, за које је $z_{u,v} = z_{m_1,n_1} - z_{m_2,n_2} < \alpha$, чиме би бројеви $z_{u,v} + z^{\backprime}_{u,v}$ и $z^*_{m,n} - z_{u,v}$ били у контрадикцији са дефиницијом бројева $z^*_{m,n}$ и $z^{\backprime}_{u,v}$ (јер интервал $(z, z^*_{m,n})$ не садржи бројеве облика $z_{m,n}$). На исти начин можемо закључити да никоја два узастопна члана $\{z_{m,n}\}$ не могу бити на растојању већем од α. Последица до сада свега изведеног је да се разлика свака два члана $\{z_{m,n}\}$ може представити као умножак од α. Тада постоје природни бројеви k и p за које важи

$$\begin{aligned}(\ln n_1 + \ln n_2) - \ln n_1 &= k\alpha\\(\ln n_1 + \ln n_2) - \ln n_2 &= p\alpha\end{aligned}$$

Количник ове две једнакости је у контрадикцији са претпоставком да је $\ln n_1/\ln n_2$ ирационалан број. Дакле, избор броја z, тако да је $\{z_{m,n}\}$ свуда густ на реалној правој, је могућ ∎

Наредна теорема је резултат америчког статистичара, индијског порекла, Џајарама Сетурамана[4].

Теорема 3.2 (Сетураман). *Нека је $F(x)$ недегенерисана функција расподеле случајне величине X. Тада $X \sim \varepsilon(\lambda)$ ако и само ако $n_i X_{(1;n_i)} \sim \varepsilon(\lambda)$ за произвољне $1 < n_1 < n_2$ за које је $\ln n_1/\ln n_2$ ирационалан број.*

Доказ. Уведимо већ познату смену, $G(x) = 1 - F(x)$. Тада, као и у претходној теореми, из услова теореме (3.2), важи

$$G^n(x) = G(nx) \text{ за свако } x \geq 0 \text{ и за различите } n = n_1 \text{ или } n_2.$$

Такође знамо да важи и $G\left(\frac{x}{n}\right) = G^{\frac{1}{n}}(x)$, тада индукцијом добијамо да

$$G^N(x) = G(Nx) \tag{3.2.8}$$

важи за сваки цео број $N = n_1^s n_2^t$, где су s и t произвољни цели бројеви (било позитивни, или негативни). Тада, на основу Леме 3.1, знамо да је скуп бројева облика

$$u = \ln N = s \ln n_1 + t \ln n_2$$

свуда густ на реалној правој, за произвољне целе бројеве s и t. Стога, постоје низови $s = s(k)$ и $t = t(k)$, за $k \geq 1$, за које $u(k) \to 0$. Нека је $u = u(k)$. Тада, уколико је $g(x) = \ln(-\ln G(e^x))$, $z = \ln x$ и $u = \ln N$, из једнакости (3.2.8) имамо

$$\begin{aligned}g(u+z) - g(z) &= \ln(-\ln G(e^{u+z})) - \ln(-\ln G(e^z))\\&= \ln(-\ln G(xN)) - \ln(-\ln G(x))\\&= \ln(-N \ln G(x)) - \ln(-\ln G(x))\\&= \ln\left(\frac{-N \ln G(x)}{-\ln G(x)}\right)\\&= \ln N = u\end{aligned}$$

[4] Jayaram Sethuraman (1965) - Department of Statistics, Florida State University

тј.

$$g(u+z)-g(z) \over u =1. \tag{3.2.9}$$

Отуда, ако $u \to 0$, следи да је $g'(z) = 1$ уколико извод постоји (то не следи из последње једнакости јер u није произвољно). Како је $G(z)$ монотона, онда је и $g(z)$ монотона. Већ смо споменули да је низ бројева u свуда густ на реалној правој. Тада, нека је v произвољан низ бројева такав да $v \to 0$, па можемо конструисати низове бројева $u_1 \leq v \leq u_2$ такве да $u_i \to 0$ и $\frac{u_i}{v} \to 1$. Због монотоности функције $g(z)$ важи

$$\frac{g(u_1+z)-g(z)}{u_1}\frac{u_1}{v} \leq \frac{g(v+z)-g(z)}{v} \leq \frac{g(u_2+z)-g(z)}{u_2}\frac{u_2}{v},$$

где из (3.2.9) следи

$$\frac{u_1}{v} \leq \frac{g(v+z)-g(z)}{v} \leq \frac{u_2}{v}.$$

Тада, према Теореми о два жандарма, $g'(z)$ постоји и једнак је јединици, када $v \to 0$. Из решења ове диференцијалне једначине следи да је $\ln(-\ln G(e^z)) = C$, $C = const$. Тада је

$$G(e^z) = \exp\left\{-e^{z+C}\right\}, \text{ тј. } G(x) = e^{-\lambda x},\ \lambda = e^C > 0,$$

што је и требало доказати ■

У књизи [3] је услов теореме, $n_i X_{(1;n_i)} \sim \varepsilon(\lambda)$, замењен својством (S3), међутим доказ је исти у оба случаја, јер смо показали еквивалентност ова две особине. Каснија уопштења Теореме 3.2, доказао је Б. Арнолд[5], претпостаљајући $n_2 = n_1 + 1$. Ово нису једина уопштења Десуове теореме, следећа теорема је такође резултат Б. Арнолда.

Теорема 3.3 (Арнолд)**.** *Претпоставимо да за неко* $n \geq 2$, $nX_{(1;n)}$ *има исту расподелу као и функција расподеле популације* $F(x)$*. Уколико за функција* $F(x)$ *важи да је* $\lim\limits_{x\to 0+} \frac{F(x)}{x} = \lambda > 0$ *коначан, тада је* $F(x)$ *облика (3.1.1).*

Доказ. Из услова теореме $nX_{(1;n)} \stackrel{d}{=} X$, следи до сада већ утврђена особина

$$(1-F(x))^n = 1 - F(nx). \tag{3.2.7a}$$

[5] Barry C. Arnold (1971) - амерички статистичар

Претпоставимо да ово важи за неко $n \geq 2$, такво да је $n(k) = n^k, k \geq 1$. Од низа случајних величина $X_1, X_2, ..., X_{n(k)}$ формирамо блокове

$$\begin{aligned}
&\left(X_1, X_2, ..., X_{n(k-1)}\right)\\
&\left(X_{n(k-1)+1}, X_{n(k-1)+2}, ..., X_{2n(k-1)}\right)\\
&\vdots\\
&\left(X_{(n-1)n(k-1)+1}, X_{(n-1)n(k-1)+2}, ..., X_{n(k)}\right)
\end{aligned}$$

и за сваки од блокова одредимо минимум и означимо са

$$X^{(1)}_{(1;n(k-1))}, X^{(2)}_{(1;n(k-1))}, ..., X^{(n)}_{(1;n(k-1))},$$

тада је и

$$X_{(1;n(k))} = \min\left(X^{(1)}_{(1;n(k-1))}, X^{(2)}_{(1;n(k-1))}, ..., X^{(n)}_{(1;n(k-1))}\right). \tag{3.2.10}$$

Како су случајне величине X_j независне и једнако расподељене, онда су такве и $X^{(j)}_{(1;n(k-1))}$. Тада за $k = 2$, сваки од $X^{(j)}_{(1;n(k-1))}$ има расподелу као и $X_{(1;n)}$, чија је функција расподеле $F(nx)$. Функција расподеле за $X_{(1;n(2))}$ из (3.2.10) је

$$P(X_{(1;n(2))} \geq x) = (1 - F(nx))^n = 1 - F(n \cdot nx) = 1 - F(n^2 x)$$

Индукцијом по k добијамо

$$P(X_{(1;n(k))} \geq x) = (1 - F(x))^{n(k)} = 1 - F\left(n(k)x\right)$$

На основу овога, следи

$$\begin{aligned}
P\left(X_{(1;n(k))} \leq \frac{x}{n(k)}\right) &= 1 - \left(1 - F\left(\frac{x}{n(k)}\right)\right)^{n(k)}\\
&= 1 - \left(1 - F\left(n(k)\frac{x}{n(k)}\right)\right) = F(x)
\end{aligned}$$

где је $n \geq 2$, и $n(k) = n^k$, за $k \geq 1$. Из претходног видимо да важи

$$F(x) = 1 - \left(1 - F\left(\frac{x}{n(k)}\right)\right)^{n(k)} \quad \text{за свако } k \geq 1. \tag{3.2.11}$$

По претпоставци теореме, када $k \to +\infty$, $F(xn^{-k}) = \lambda x n^{-k} + o(n^{-k})$. Тада, познату релацију

$$\lim_{n\to+\infty}\left(1 + \frac{x}{n} + o(\frac{1}{n})\right)^n = e^x,$$

можемо применити на функцију расподеле $F(xn^{-k})$. Па за свако $x > 0$, важи

$$F(x) = 1 - \lim_{k\to+\infty}\left(1 - \frac{\lambda x}{n^k} + o\left(\frac{1}{n^k}\right)\right)^{n^k} = 1 - e^{-\lambda x}, \quad x > 0 \blacksquare$$

Метода коју смо користили у доказу је *метода граничних закона*, и користи се у многим другим доказима. Теорема (3.3) важи и за друге расподеле уколико изузмемо претпоставку $\lim\limits_{x\to 0+}\frac{F(x)}{x} = \lambda > 0$, нпр. можемо посматрати расподелу облика $F(x) = 1 - e^{g(x)}$, где је $g(x) < 0$ нерастућа функција. Да би за њу важили особина (3.2.11), потребно је да важи

$$g(x) = n^k g\left(\frac{x}{n^k}\right), \text{ за свако } x \text{ и свако } k \geq 1, \tag{3.2.12}$$

где је $n \geq 2$ фиксиран број. Хуанг[6] је показао да се таква функција може конструисати. Један пример такве функције је

$$g(x) = \begin{cases} 0, & \text{за свако } x \leq 0 \\ -2^j\left(\frac{\ln x - j\ln 2}{\ln 2} + 1\right), & x > 0, \text{ и } j \in \mathbb{Z} \text{ такво да } 2^j < x \leq 2^{j+1}. \end{cases}$$

Овде (3.2.12) важи за $n = 2$, међутим, $F(x)$ у овом случају није експоненцијална.
Да се све карактеризационе теореме не своде само на особине минималне статистике поретка, показују наредне теореме. Пре него што наведемо њихове исказе, кратко упознавање са појмом *нормализованих размака.*

Дефиниција 3.1. *За случајни узорак $X_1, X_2, ..., X_n$ обима n, из популације X са недегенерисаном функцијом расподеле $F(x)$, дефинишемо одговарајући варијациони низ*

$$X_{(1;n)} \leq X_{(2;n)} \leq ... \leq X_{(n;n)}.$$

Под појмом нормализованих размака подразумевамо случајне величине облика

$$D_{(1;n)} = X_{(1;n)} \text{ и } D_{(k;n)} = (n-k+1)\left(X_{(k;n)} - X_{(k-1;n)}\right), \ 2 \leq k \leq n. \tag{3.2.13}$$

За теорију карактеризације, најважније су оне особине нормализованих размака при екпоненцијалној расподели популације. Зато је следећи резултат Сукхатмеа[7], од посебне важности за карактеризацију екпоненцијалне расподеле.

[6] J. S. Huang (1974) - кинески статистичар
[7] Pandurang Vasudeo Sukhatme (1937) - индијски статистичар

Теорема 3.4. *Нека је* $F(x) = 1 - e^{-\lambda x}, x \geq 0, \lambda \geq 0$. *Тада су разлике*

$$X_{(k;n)} - X_{(k-1;n)},\ \ 0 \leq k \leq n,\ X_{(0;n)} = 0, \tag{3.2.14}$$

независне и једнако расподељене променљиве са експоненцијалном расподелом

$$P\left(X_{(k;n)} - X_{(k-1;n)} \leq x\right) = 1 - e^{-\lambda(n-k)x},\ x \geq 0,$$

тј.
Уколико је $F \sim \varepsilon(\lambda)$, *онда су* $D_{(1;n)}, D_{(2;n)}, ..., D_{(n;n)}$ *једнако расподељене и независне случајне величине са експоненцијалном расподелом* $\varepsilon(\lambda)$.

Доказ. На основу једнакости (2.3.6), густина расподеле вектора $(X_{(1;n)}, X_{(2;n)}, ..., X_{(n;n)})$ је

$$f_{X_{(1;n)},X_{(2;n)},...,X_{(n;n)}}(x_1, x_2, ..., x_n) = n!\lambda^n e^{-\lambda \sum\limits_{i=1}^{n} x_i}.$$

Јакобијан трансформације $T_k = X_{(k;n)} - X_{(k-1;n)}$, је детерминанта горње троугаоне јединичне матрице, и једнак је један. Одатле је заједничка густина расподеле вектора $T_1, T_2, ..., T_n$ једнака

$$f_{T_1,T_2,...,T_n}(t_1, t_2, ..., t_n) = n!\lambda^n \exp\left\{-\lambda \sum_{i=1}^{n}\sum_{j=1}^{i} t_j\right\}.$$

Како је

$$\sum_{i=1}^{n}\sum_{j=1}^{i} t_j = \sum_{i=1}^{n}(n-i+1)t_i,$$

$$\begin{aligned} f(t_1, t_2, ..., t_n) &= \prod_{j=1}^{n}(n-j+1)\lambda^n \exp\left\{-\lambda\sum_{j=1}^{n}(n-j+1)t_j\right\} \\ &= \prod_{j=1}^{n}\left((n-j+1)\lambda e^{-\lambda(n-j+1)t_j}\right),\ t_j \geq 0,\ j = 1, ..., n\ \ . \end{aligned}$$

Како је заједничка густина расподеле производ маргиналних густина сваке променљиве појединачно, може се доказати да су случајне величине независне (последица Фубинијеве теореме), и једнако расподељене са функцијом густине $f(t_k) = \lambda(n-k+1)e^{-\lambda(n-k+1)t_k}$, $1 \leq k \leq n$, $t_k \geq 0$. Тада, из $T_k \sim \varepsilon((n-k+1)\lambda)$ следи да су случајне величине $D_{(k;n)} = (n-k+1)T_k$, $1 \leq k \leq n$, независне и једнако расподељене са $\varepsilon(\lambda)$ расподелом ∎

Слично тврђење важи и за униформну расподелу.

Теорема 3.5. *Нека је* $F(x) = x$ *за* $0 \leq x \leq 1$. *Тада су размаци једнако расподељени са функцијом расподеле* $F_1(x) = 1 - (1-x)^n$.

Приметимо да у претходној теореми размаци нису међусобно независни. Штавише, показаћемо да је независност нормализованих размака $D_{(k;n)}$ карактеризациона особина експоненцијалне расподеле. Посебну важност имају следеће теореме, које се тичу функције густине размака две статистике поретка и условне функције расподеле.

Теорема 3.6. *Нека је функција расподеле $F(x)$ дефинисана на сегменту $[0,+\infty)$. Тада за свако $i,j=1,2,...,n$, такво да је $i<j$, и $u>0$, важи*

$$\tilde{f}(u) := f_{X_{(j;n)}-X_{(i;n)}}(u) = \int_0^{+\infty} f_{X_{(i;n)},X_{(j;n)}}(x,x+u)dx. \tag{3.2.15}$$

Доказ. Нека је S област на којој је $X_{(j;n)} > X_{(i;n)}$ и $X_{(j;n)} \leq X_{(i;n)}+u$. Тада је

$$\begin{aligned} F_{X_{(j;n)}-X_{(i;n)}}(u) &= \int\int_S f_{X_{(i;n)},X_{(j;n)}}(x,y)dxdy \\ &= \int_0^{+\infty}\int_x^{x+u} f_{X_{(i;n)},X_{(j;n)}}(x,y)dxdy. \end{aligned}$$

Користећи Лајбницово правило диференцирања интеграла, добијамо

$$\begin{aligned} \tilde{f}_{i,j}(u) &= \frac{\partial}{\partial u}F_{X_{(j;n)}-X_{(i;n)}}(u) = \frac{\partial}{\partial u}\int_0^{+\infty}\int_x^{x+u} f_{X_{(i;n)},X_{(j;n)}}(x,y)dxdy \\ &= \int_0^{+\infty}\left(\frac{\partial}{\partial u}\int_x^{x+u} f_{X_{(i;n)},X_{(j;n)}}(x,y)dy\right)dx \\ &= \int_0^{+\infty} f_{X_{(i;n)},X_{(j;n)}}(x,x+u)dx \quad \blacksquare \end{aligned}$$

Теорема 3.7. *Условна функција расподеле $X_{(j;n)}$, када је дато $X_{(j;n)}=x_i$ за $i<j$, је једнака функцији расподеле $(j-i)-$те статистике поретка из узорка обима $n-i$, из популације чија је функција расподеле облика*

$$F^*(x_j) = \begin{cases} \frac{F(x_j)-F(x_i)}{1-F(x_i)}, & \text{ако је } x_j \geq x_i \\ 0, & \text{иначе.} \end{cases} \tag{3.2.16}$$

Доказ. Из маргиналне функције расподеле $X_{(i;n)}$ и заједничке функције густине $X_{(i;n)}$ и $X_{(j;n)}$, добијамо да је условна функција густине $X_{(j;n)}$, када је $X_{(i;n)} = x_i$, дата са

$$\begin{aligned} f_{X_{(j;n)}}(x_j|X_{(i;n)} = x_i) =& f_{X_{(i;n)},X_{(j;n)}}(x_i, x_j)/f_{X_{(i;n)}}(x_i) \\ =& \frac{(n-i)!}{(j-i-1)!(n-j)!} \left(\frac{F(x_j) - F(x_i)}{1 - F(x_i)} \right)^{j-i-1} \\ & \times \left(\frac{1 - F(x_j)}{1 - F(x_i)} \right)^{n-j} \frac{f(x_j)}{1 - F(x_i)}, \end{aligned}$$

где је $i < j \le n$ и $x_i \le x_j < \infty$. Закључак следи из чињенице да $\frac{F(x_j)-F(x_i)}{1-F(x_i)}$ представља нашу функцију расподеле $F^*(x_j)$, и $\frac{f(x_j)}{1-F(x_i)}$ њену одговарајућу функцију густине ■

Следећу теорему доказали су Пури и Рубин (1970).

Теорема 3.8 (Пури и Рубин). *Нека су X_1 и X_2 независне, ненегативне и једнако расподељене случајне величине са апсолутно непрекидном и диференцијабилном функцијом расподеле $F(x)$, обележја X, и ако је*

$$X \stackrel{d}{=} X_{(2;2)} - X_{(1;2)}, \tag{3.2.17}$$

$F(x)$ је експоненцијално расподељена.

Исказ Теореме 3.8 је идентичан оном у [6], док се у оригиналном раду (видети [7]) претпоставља да ако за две независне и једнако расподељене, ненегативне и апсолутно непрекидне случајне величине, X и Y, важи $X \stackrel{d}{=} |X - Y|$, тада X има експоненцијалну расподелу. Међутим, како за две случајне величине важи $|X_1 - X_2| = X_{(2;2)} - X_{(1;2)}$, свеједно је који ћемо услов писати у теореми.

Доказ Теореме 3.8 може се наћи у [7]. Иначе без претпоставке о апсолутној непрекидности расподеле, услов (3.2.17) је испуњен и у случају дискретне и сингуларне функције расподеле, али не и мешавина те две расподеле.

Напомена. Слично тврђење у коме претпостављамо да $X_{(2;2)} - X_{(1;2)}$ има експоненцијалну расподелу, није карактеризација. Розберг[8] је навео пример функције расподеле облика

$$F(x) = 1 - e^{-x} \left(1 + \frac{4}{\alpha^2} \left(1 - cos(\alpha x)\right) \right),\ x \ge 0,\ \alpha \ge 2\sqrt{2},$$

код које $D_{(2;2)}$ такође има експоненцијалну расподелу, али расподела популације није експоненцијална. У општем случају, из претпоставке

[8] H. J. Rossberg (1972) - немачки математичар

о експоненцијалној расподељености размака $D_{(k,n)}, 1 \le k \le n$, не следи карактеризација експоненцијалне расподеле.
Розберг је такође препознао разлог због чега расподеле другачије од експоненцијалне могу имати својство да су размаци експоненцијално расподељени. Наиме, он је уочио да за све расподеле $F(x)$, такве да је $F(0+) = 0$, за које је $D_{(k,n)}$ из експоненцијалне расподеле, Лапласова трансформација

$$L_k(s) = \int_0^{+\infty} e^{-sx} dF^k(x)$$

има нуле за неко $s \in \mathbb{C}$ (скуп свих комплексних бројева), такво да $\Re s \ge 0$. Отуда, поставља питање када ће, под претпоставком да је $L_k(s) \ne 0$ за $\Re s \ge 0$, експоненцијалност размака $D_{(k,n)}$ имплицирати експоненцијалну расподељеност популације.

Теорема 3.9 (Розберг). *Нека је функција расподеле популације $F(x)$ таква да је $F(0+) = 0$ и $F(x) > 0$ за све $x > 0$. Претпоставимо да Лапласова трансформација $L_k(s)$ нема нула за $\Re s \ge 0$. Ако, за неко $k \ge 1$, $D_{(k;n)} \sim \varepsilon(1)$ онда је $F(x) = 1 - e^{-x}$, $x \ge 0$.*

Нешто другачији исказ Теореме 3.9 може се наћи у књизи [8]. Доказ ове теореме (може се наћи у [3, стр. 42]) је прилично компликован, и може се поједноставити уз додатне претпоставке. Зато у следећој дефиницији уводимо појам стопе ризика (енг. *hazard rate*).

Дефиниција 3.2 (Hazard rate). *Нека је $F(x)$ апсолутно непрекидна функција расподеле. Тада је стопа ризика (hazard rate) дефинисана са*

$$r(x) = \frac{F'(x)}{1 - F(x)} \tag{3.2.18}$$

У случају експоненцијалне расподеле стопа ризика је константна. Кажемо још да функција расподеле $F(x)$ има растућу стопу ризика (IHR-*increasing hazard rate*) ако је $r(x) \le r(x+y), x, y \ge 0$ и $F(x)$ има опадајућу стопу ризика (DHR-*decreasing hazard rate*) ако је $r(x) \ge r(x+y), x, y \ge 0$. Кажемо и да функција $F(x)$ припада класи $\mathfrak{D}$ уколико је IHR, односно DHR. Следеће теорема укључује претпоставку о стопи ризика, и резултат је Ахсанулаха[9].

Теорема 3.10 (Ахсанулах). *Нека је $F(x)$ апсолутно непрекидна са функцијом густине $f(x) = F'(x)$. Претпоставимо да је $F(0) = 0$, $F(x)$*

[9] Mohammad Ahsanullah (1976) - амерички статистичар, индијског порекла

строго растућа за свако $x > 0$ и да је стопа ризика $r(x)$ монотона за свако $x \geq 0$. Уколико су за неко $n \geq 2$, $D_{(1;n)}$ и $D_{(2;n)}$ једнако расподељене, онда је $F(x) = 1 - e^{-\lambda x}$, $x \geq 0$, за неко $\lambda > 0$.

Доказ. При рачунању расподеле $D_{(2;n)}$, користимо формулу потпуне вероватноће при услову $X_{(1;n)} = y$

$$P(D_{(2;n)} < x) = \int_0^{+\infty} P(D_{(2;n)} < x | X_{(1;n)} = y) f_{(1;n)}(y) dy.$$

Сада замењујемо израз за густину $f_{(1;n)}(y)$ из (2.2.4), и условну функцију расподеле из Теореме 3.7 и 3.6. Добијамо

$$\text{(D2)} \quad \begin{aligned} P(D_{(2;n)} < x) &= \int_0^{+\infty} \left(1 - \left(\frac{1 - F\left(y + \frac{x}{n-1}\right)}{1 - F(y)}\right)^{n-1}\right) n(1 - F(y))^{n-1} f(y) dy \\ &= 1 - n \int_0^{+\infty} \left(\frac{1 - F\left(y + \frac{x}{n-1}\right)}{1 - F(y)}\right)^{n-1} (1 - F(y))^{n-1} f(y) dy, \end{aligned}$$

јер је $\int_0^{+\infty} n(1 - F(y))^{n-1} f(y) dy = \int_0^{+\infty} f_{(1;n)}(y) dy = 1$, одакле такође следи да и расподелу $D_{(1;n)}$ можемо записати у облику

$$\begin{aligned} F_{(1;n)}(x) &= 1 - \left(1 - F\left(\frac{x}{n}\right)\right)^n \\ \text{(D1)} \qquad &= 1 - n \int_0^{+\infty} \left(1 - F\left(\frac{x}{n}\right)\right)^n (1 - F(y))^{n-1} f(y) dy. \end{aligned}$$

На основу (D1) и (D2), и Теореме 3.4 о једнакој расподељености $D_{(1;n)}$ и $D_{(2;n)}$, следи

$$\text{(3.2.19)} \qquad \int_0^{+\infty} K(x, y, n) (1 - F(y))^{n-1} f(y) dy = 0, \text{ за свако } x \geq 0,$$

где је

$$K(x, y, n) = \left(1 - F\left(\frac{x}{n}\right)\right)^n - \left(\frac{1 - F\left(y + \frac{x}{n-1}\right)}{1 - F(y)}\right)^{n-1}.$$

Како је стопа ризика $r(x) = -\frac{d}{dx} \ln(1 - F(x))$ монотона, следи да је $\ln(1 - F(x))$ било конкавна или конвексна, у зависности од тога да ли је $r(x)$ растућа, односно опадајућа. Претпоставимо да је стопа ризика $r(x)$ растућа. Тада из конкавности функције $\ln(1-F(x))$ следи

$$\begin{aligned}\ln\left(1-F\left(y+\frac{x}{n}\right)\right)=&\ln\bar{F}\left(\frac{1}{n}y+\frac{n-1}{n}\left(y+\frac{x}{n-1}\right)\right)\\ \geq&\frac{1}{n}\ln(1-F(y))+\frac{n-1}{n}\ln\left(1-F\left(y+\frac{x}{n-1}\right)\right),\end{aligned}$$

што је еквивалентно

$$\left(1-F\left(y+\frac{x}{n}\right)\right)^{n}\geq(1-F(y))\left(1-F\left(y+\frac{x}{n-1}\right)\right)^{n-1}.$$

Када неједнакост помножимо са $(1-F(y))^{-n}\geq 0$ следи

$$K(x,y,n)\geq\left(1-F\left(\frac{x}{n}\right)\right)^{n}-\left(\frac{1-F\left(y+\frac{x}{n}\right)}{1-F(y)}\right)^{n}\geq 0,$$

где последња неједнакост такође следи из својства конкавности функције $\ln(1-F(x))$. Примењујући добијену неједнакост, следи да (3.2.19) важи само када је

$$K(x,y,n)=0 \text{ за свако } x,y\geq 0.$$

Тада је

$$1-F\left(y+\frac{x}{n-1}\right)=(1-F(y))\left(1-F\left(\frac{x}{n}\right)\right)^{\frac{n}{n-1}} \text{ за свако } x,y\geq 0.$$

Одакле, при смени $z=\frac{x}{n-1}$, добијамо особину *одсуства меморије*, тј.

$$\frac{1-F(y+z)}{1-F(y)} \text{ не зависи од } y.$$

У претходном одељку је доказана еквивалентност својства *одсуства меморије* и Десуове теореме, што је чини карактеризационом особином експоненцијалне расподеле, чиме је овај део доказа, при претпоставци да је $r(x)$ растућа, завршен. У случају када је $r(x)$ опадајућа, доказ је аналоган ∎

У овој теореми смо претпоставили једнаку расподељеност две статистике поретка, и то $D_{(1;n)}$ и $D_{(2;n)}$. Међутим, резултат остаје исти и при претпоставци о једнакој расподељености $D_{(i;n)}$ и $D_{(i+1;n)}$ за $2\leq i\leq n$, јер доказ концептуално остаје исти. Ахсанулах је доказао да важе и знатна уопштења Теореме 3.10. У наставку су представљени искази ових теорема, док се докази могу наћи у радовима [9] и [10].

Теорема 3.11. *Нека је X ненегативна случајна величина са апсолутно непрекидном строго растућом функцијом расподеле $F(x)$ за свако $x > 0$, и $F(x) < 1$ за свако x. Тада су следећа својства еквивалентна*

(а) *X има експоненцијалну расподелу са параметром $\lambda > 0$;*

(б) *$F(x)$ је из класе $\mathfrak{D}$, и за неко i, такво да је $2 \le i \le n$, статистике $D_{(i;n)}$ и $D_{(1;n)}$ су једнако расподељене.*

Теорема 3.12. *Нека је X ненегативна случајна величина са апсолутно непрекидном функцијом расподеле $F(x)$ која је строго растућа на $[0, +\infty)$. Тада су следећа својства еквивалентна*

(а) *X има експоненцијалну расподелу са параметром $\lambda > 0$;*

(б) *$F(x)$ је из класе $\mathfrak{D}$, и за неко i и j, $1 \le i < j < n$, статистике $D_{(i;n)}$ и $D_{(j;n)}$ су једнако расподељене.*

Наредна теорема о карактеризацији, заснованој на једнакости расподела, не садржи претпоставку о стопи ризика, и представља такође значајан резултат међу теоремама карактеризације помоћу статистика поретка вишег реда.

Теорема 3.13 (Гатер[10]). *Нека је $F(x)$ непрекидна и строго растућа функција расподеле за свако $x > 0$. Онда је F експоненцијална функција расподеле акко је*

$$X_{(j-i;n-i)} \stackrel{d}{=} X_{(j;n)} - X_{(i;n)} \tag{3.2.20}$$

за две различите вредности j_1 и j_2 од j и неко i и $n \ge 3$, такве да је $1 \le i < j_1 < j_2 \le n$.

Доказ теореме дат је у [11]. Као и у Теореми 3.10 доказано је да на основу услова (3.2.20), за функцију расподеле $F(x)$, важи својство *одсуства меморије*. Обрнуто, уколико је F експоненцијална функција расподеле, особина (3.2.20) следи на основу (2.2.4) и Теореме 3.6.

За крај овог поглавља који се односи на карактеризацију експоненцијалне расподеле помоћу једнако расподељених статистика поретка наводимо Лему о репрезентацији статистка поретка, која се још назива *Сукхатмеова репрезентација*, и Теорему која се односи на репрезентацију.

[10] Ursula Gather (1989) - немачка статистичарка

Лема 3.2. *За статистике поретка узорка из експоненцијалне расподеле може се показати да важи следећа особина*

$$X_{(k;n)} = \sum_{j=n-k+1}^{n} \frac{X_j}{j}, \text{ за } 1 \le k \le n. \tag{3.2.21}$$

Доказ. Важи да је

$$\begin{aligned} X_{(k;n)} =& X_{(k;n)} - X_{(k-1;n)} + X_{(k-1;n)} - \dots + X_{(1;n)} - X_{(0;n)} \\ =& \frac{1}{n-k+1}\left((n-k+1)(X_{(k;n)} - X_{(k-1;n)})\right) + \dots + \frac{1}{n}\left(n(X_{(1;n)} - X_{(0;n)})\right) \\ =& \frac{1}{n-k+1} D_{(k;n)} + \frac{1}{n-k+2} D_{(k-1;n)} + \dots + \frac{1}{n} D_{(1;n)}, \end{aligned}$$

јер је $X_{(0;n)} = 0$, и према Теореми 3.4 важи да су $D_{(1;n)}, D_{(2;n)}, \dots, D_{(n;n)}$ независне и једнако расподељене случајне величине са $\varepsilon(\lambda)$, одакле следи формула

$$X_{(k;n)} = \sum_{j=n-k+1}^{n} \frac{X_j}{j}, \text{ за } 1 \le k \le n,$$

где су $X_j \sim \varepsilon(\lambda)$ ■

Следеће тврђење доказао је Ахсанулах (1972).

Теорема 3.14. *Потребан и довољан услов да недегенерисана, ненегативна случајна величина X, са апсолутно непрекидном функцијом расподеле F, има експоненцијалну расподелу је да се њена статистика поретка $k-$тог реда може представити као*

$$X_{(k;n)} = \sum_{j=1}^{k} \frac{X_j}{n-j+1},$$

за свако $k \in \mathbb{N}$ такво да је $1 \le k \le n$, где су $X_j, j = 1, 2, \dots, k$ независне и једнако расподељене случајне величине са заједничком функцијом расподеле F.

3.3 Карактеризација експоненцијалне расподеле, заснована на независности статистика поретка

У овом поглављу представићемо неке од теорема карактеризације које се односе на независност статистика поретка. У претходном поглављу 3.2 смо већ споменули, да је независност размака статистика поретка, карактеризациона особина експоненцијалне расподеле (на основу Теорема 3.4 и 3.5). Отуда, ово поглавље почињемо теоремом коју је доказао М. Фис[11].

[11] Marek Fisz (1958) - пољски математичар

Теорема 3.15 (Фис)**.** *Нека су X_1 и X_2 независне случајне величине са непрекидном функцијом расподеле $F(x)$. Претпоставимо да је $F(0)=0$ и да је $F(x)$ строго растућа за свако $x>0$. Тада су $X_{(2;2)}-X_{(1;2)}$ и $X_{(1;2)}$ независне ако и само ако је $F(x)=1-e^{-\lambda x}$ за неко $\lambda>0$.*

Доказ. У специјалном случају $n=2$ из Теореме 3.4 следи да су $X_{(2;2)}-X_{(1;2)}$ и $X_{(1;2)}$ заиста независне ако је $F(x)$ експоненцијална функција расподеле. Отуда, остаје да докажемо обрнуто тврђење.
Уколико су $X_{(2;2)}-X_{(1;2)}$ и $X_{(1;2)}$ независне, тада је

$$P(X_{(2;2)}-X_{(1;2)}<x|X_{(1;2)}=z)=P(X_{(2;2)}-X_{(1;2)}<x), \tag{3.3.1}$$

за свако $z>0$. С друге стране, према Теореми 3.7 важи

$$P(X_{(2;2)}-X_{(1;2)}<x|X_{(1;2)}=z)=P(X_{(2;2)}<x+z|X_{(1;2)}=z)=P(X^*_{(1;1)}<x+z),$$

где је $X^*_{(1;1)}$ статистика поретка из популације чија је функција расподеле облика

$$F^*(x)=\begin{cases}\frac{F(x)-F(z)}{1-F(z)}, & \text{ако је } x\geq z\\ 0, & \text{иначе.}\end{cases}$$

Уколико десну страну једнакости (3.3.1) означимо са $H(x)$, тада је

$$\begin{aligned}H(x)=&\frac{F(x+z)-F(z)}{1-F(z)}=\frac{F(x+z)-1+1-F(z)}{1-F(z)}\\ =&\frac{1-F(x+z)}{1-F(z)} \text{ за свако } x\geq 0 \text{ и } z>0.\end{aligned} \tag{3.3.2}$$

Када $z\to 0$ следи да је $F(x)=H(x)$. Отуда, из (3.3.2) следи да је

$$F(x)=1-\frac{1-F(x+z)}{1-F(z)} \text{ тј. } 1-F(x)=\frac{1-F(x+z)}{1-F(z)},$$

за свако $x\geq 0$ и $z>0$. Одакле следи да за функцију $F(x)$ важи својство одсуства меморије, што значи да $F(x)$ мора бити екпоненцијална ■

Теорема 3.15 се односи само на случај $n=2$. Следе уопштења Фисове теореме.

Теорема 3.16. *Уколико је F апсолутно непрекидна функција расподеле, и ако су $X_{(1;n)}$ и $\sum\limits_{i=2}^{n}(X_{(i;n)}-X_{(1;n)})$ независне, онда је F експоненцијална.*

Теорема 3.17. *Уколико је F апсолутно непрекидна функција расподеле, и ако су за неко $m=2,3,\ldots,n$, $X_{(m;n)}$ и $X_{(m+1;n)}-X_{(m;n)}$ независне, онда је F експоненцијална.*

Овај резултат је проширио Говиндараџулу[12].

Теорема 3.18. *Уколико је F апсолутно непрекидна функција расподеле и ако су $X_{(m;n)}$ и $X_{(t;n)} - X_{(s;n)}$ независне за неке $m, s, t, 1 \leq m \leq s < t \leq n$, тада је F експоненцијална.*

3.4 Карактеризација експоненцијалне расподеле, заснована на моментима статистика поретка

У теорији карактеризације се тежи ка слабљењу и смањењу броја претпоставки у теореми. С тим у вези, вратимо се на Десуову Теорему 3.1, и покушајмо да услов *једнаких расподела* заменимо условом *једнаких очекивања.* За почетак, наведимо неке основне особине момената статистика поретка.
Из поглавља (2.2) знамо да се функција расподеле $F_{(k;n)}$ може представити у облику непотпуне Бета расподеле, и као таква она задовољава рекурентну формулу

$$\left(1 - \frac{k}{n}\right) F_{(k;n)} + \frac{k}{n} F_{(k+1;n)} = F_{(k;n-1)}.$$

Одатле, када очекивање постоји, важи слична рекурентна формула

$$\left(1 - \frac{k}{n}\right) E[X_{(k;n)}] + \frac{k}{n} E[X_{(k+1;n)}] = E[X_{(k;n-1)}],$$

где је

$$E[X_{(k;n)}] = \int_{-\infty}^{+\infty} x dF_{(k;n)}(x) = \int_{0}^{1} F^{-1}(t) k \binom{n}{k} t^{k-1} (1-t)^{n-k} dt,$$

и F^{-1} представља инверзну функцију од F, дефинисану са

$$F^{-1}(t) = \inf\{x | F(x) \geq t\}.$$

Теорема 3.19 (Хуанг)**.** *Уколико је F недегенерисана функција расподеле, и ако је*

$$E[nX_{(1;n)}] = E[X_1] < \infty, \text{ за свако } n = 2, 3, \ldots,$$

тада је F експоненцијална функција расподеле.

Хуанг је показао да упркос његовој претпоставци $E[X_1] < \infty$, ова претпоставка је мање рестриктивна од претпоставке коју је имао Десу (Теорема 3.1). Зато презентујемо јачу верзију Теореме 3.19.

[12] Z. Govindarajulu (1978) - индијски статистичар

Теорема 3.20. *Уколико је F недегенерисана функција расподеле и ако су*

(а) *$nX_{(1;n)}$ и X_1 једнако расподељене за неко $n \geq 2$*

(б) *$E[nX_{(1;n)}] = E[X_1]$ за остале n–ове*

тада је F експоненцијална функција расподеле.

3.5 Монотоне трансформације експоненцијалне расподеле

Већ смо споменули да се одређеним монотоним трансформацијама експоненцијалне расподеле могу добити неке друге расподеле, за које ће важити сличне карактеризационе теореме, као и код експоненцијалне расподеле. Од посебне важности је следећа теорема.

Теорема 3.21. *Нека је X случајна величина са непрекидном функцијом расподеле $F(x)$, тада*

$$Y = -\ln F(x), \tag{3.5.1}$$

и

$$W = -\ln(1 - F(x)), \tag{3.5.2}$$

имају експоненцијалну расподелу, са параметром 1.

Доказ.

$$\begin{aligned}
P(Y \leq x) =& P(-\ln F(x) \leq x) = P(\ln F(x) \geq -x) = P(e^{\ln F(x)} \geq e^{-x}) \\
=& P(F(x) \geq e^{-x}) = 1 - P(F(x) < e^{-x}) = 1 - P(X < F^{-1}(e^{-x})) \\
=& 1 - F(F^{-1}(e^{-x})) = 1 - e^{-x},\ x \geq 0 \\
P(W \leq x) =& P(-\ln(1 - F(x)) \leq x) = P(\ln(1 - F(x)) \geq -x) = P(1 - F(x) \geq e^{-x}) \\
=& P(F(x) \leq 1 - e^{-x}) = P(X \leq F^{-1}(1 - e^{-x})) = F(F^{-1}(1 - e^{-x})) \\
=& 1 - e^{-x},\ x \geq 0.
\end{aligned}$$

На основу ове теореме, трансформацијом других расподела добијамо експоненцијалну расподелу, чиме се њене карактеризационе особине преносе и на друге расподеле. Навешћемо примере одређених расподела које се најчешће користе, и за сваку од њих по једну карактеризацију.

1. *Униформна расподела.* Уколико је $F(x) = x$ за $0 \le x \le 1$, тада на основу (3.5.1) следи да $Y = -\ln X$ има експоненцијалну расподелу. Својство одсуства меморије постаје

$$P(-\ln X \ge x + z | -\ln X \ge z) = P(-\ln X \ge x), \text{ за свако } x, z > 0.$$

Увођењем смене $e^{-x} = u$ и $e^{-z} = v$, добијамо да за униформну расподелу важи карактеризација у облику мултипликативне особине одсуства меморије

$$P(X \le uv | X \le v) = P(X \le u), \; 0 \le u, v \le 1.$$

2. *Вејбулова расподела.* Нека је $a > 0$. Тада Вејбулову расподелу дефинишемо у облику

$$F(x) = 1 - e^{-x^a}, \; x \ge 0.$$

Из (3.5.2), следи да X^a има експоненцијалну расподелу. Како је $a > 0$, следи да за Вејбулову расподелу важи теорема аналогна Десуовој теореми. Уколико је $X_{(1;n)}$ минимална статистика поретка од n независних величина из популације X, тада је $n^{1/a} X_{(1;n)}$ једнако расподељена као и X.

3. *Паретова расподела.* За случајну величину X са функцијом расподеле

$$F(x) = 1 - x^{-a}, \; a > 0, x \ge 1,$$

кажемо да има Паретову расподелу. Из трансформације (3.5.2) добијамо да је $W = a \ln X$ експоненцијално расподељена, и из својства одсуства меморије, при смени $e^{x/a} = u$ и $e^{z/a} = v$, следи

$$P(X \ge uv | X \ge v) = P(X \ge u), \; u, v \ge 1.$$

Добили смо да ова особина представља карактеризацију Паретове расподеле.

4. *Гумбелова расподела.* Уколико X има функцију расподеле $F(x) = e^{-e^{-x}}$, тада из (3.5.1) следи да $Y = e^{-X}$ има екпоненцијалну расподелу. Карактеризација Гумбелове расподеле добија се на основу Десуове теореме, $X_{(n;n)} - \ln n$ има исту расподелу као X, где је $X_{(n;n)}$ максимална статистика поретка.

3.6 Карактеризација неких других расподела

У претходном одељку смо описали неке од трансформација којима из експоненцијалне расподеле добијамо друге расподеле. На тај начин одређене особине и теореме које су важиле за експоненцијалну расподелу, важе и за изведене расподеле. Паретову расподелу смо добили трансформацијом $Y = e^X$, тиме се њене карактеризације добијају модификацијама Десуове, Пури-Рубинове, Ахсанулахове теореме, итд. Искази наредних теорема су цитирани из рада [12].

Теорема 3.22. *Нека су X и Y једнако расподељене, ненегативне и апсолутно непрекидне случајне величине. Тада, X и $\max\left(\frac{X}{Y}, \frac{Y}{X}\right)$ имају исту расподелу ако и само ако случајна величина X има Паретову расподелу.*

Доказ. Може се наћи у раду [12].

Теорема 3.23. *Нека је $X_1, X_2, ..., X_n$ узорак из ненегативне расподеле. Ако за неко j статистике $X_{(j+1;n)}/X_{(j;n)}$ и $X_{(1;n-j)}$ имају исту расподелу, тада X_1 има Паретову расподелу.*

Теорема 3.24. *Нека је $X_1, X_2, ..., X_n$ узорак из ненегативне, апсолутно непрекидне расподеле из класе $\mathfrak{D}$. Тада, $\left(X_{(j+1;n)}/X_{(j;n)}\right)^{n-j}$ и $\left(X_{(k+1;n)}/X_{(k;n)}\right)^{n-k}$, $1 \leq j < k < n$, имају исту расподелу ако и само ако X_1 има Паретову расподелу.*

Доказ. Може се наћи у раду [12].

Закључак

Овај рад, представља увод у област карактеризације расподела. Он је базиран на већ установљеним и доказаним теоремама, које чине неопходну основу за даље познавање и разумевање ове области, која је предмет изучавања многих научника. Карактеризације се мало изучавају током основних студија, али оне су повезане са многим предметима који јесу део основних студија. Сама по себи ова област се може сврстати и у вероватноћу, и у статистику, и она их међусобно повезује. У овом раду претежно изучавамо карактеризацију експоненцијалне расподеле помоћу статистика поретка.

Карактеризација расподела се користи за формирање непараметарских тестова за испитивање расподела и тестова сагласности, чиме ова теорија постаје знатно компликованија. То није предмет изучавања у овом раду, али многе описане теореме служе као основа за формирање тест статистика и хипотеза за тестове о расподели обележја.

Литература

[1] Barry C. Arnold, N. Balakrishnan, H. N. Nagarala, *A first course in order statistics*, Society for Industrial and Applied Mathematics, Philadelphia, 1954.

[2] Herbert A. David, H. N. Nagaraja, *Order Statistics*, Wiley-Intersience, 1925.

[3] Janos Galambos, Samuel Kotz, *Characterizations of Probability Distributions*, Springer-Verlag, Berlin, Heidelberg, New York, 1978.

[4] M. M. Desu *A Characterization of the Exponential Distribution by Order Statistics*, State University of New York at Buffalo, The Annals of Mathematical Statistics, pp. 837-838, 1971.

[5] J. Galambos *Advanced Probability Theory*, CRC Press, 1995.

[6] K. Balakrishnan, *Exponential Distribution: Theory, Methods and Applications*, CRC Press, 1996.

[7] Prem S. Puri and Herman Rubin, *A Characterization Based On the Absolute Difference of Two iid. Random Variables*, Purdue University, The Annals of Mathematical Statistics, 1970, Vol. 41, No.6, 2113-2122.

[8] Ahsanullah, M. and Hamedani, *G. G. Exponential Distribution: Theory and Methods*, NOVA Science, New York (2010).

[9] M. Ahsanullah, *On a Characterization of the Exponential Distribution by Order Statistics*, Journal of Applied Probability, Vol. 13, No. 4 (Dec., 1976), pp. 818-822

[10] M. Ahsanullah, *On a Characterization of the Exponential Distribution by Spacings*, Ann. Inst. Statist. Math. 30 (1978), Part A, 163-166

[11] Ursula Gather, *On a Characterization of the Exponential Distribution by Properties of Order Statistics*, Statistics and Probability Letters 7(1989), 93-96, North-Holland.

[12] Марко Обрадовић, *Карактеризације Неких Расподела и Бахадурова Асимптотска Ефикасност Тестова Сагласности*, Докторска дисертација, Београд, 2015.