

Jensenova nejednakost u teoriji vjerojatnosti

Danijel Krizmanić* Antonio Špac†

Sažetak

U ovom radu dokazana je Jensenova nejednakost za matematičko očekivanje, te je iz nje izvedena klasična Jensenova nejednakost. Na kraju je navedeno nekoliko primjena Jensenove nejednakosti.

Ključne riječi: *Jensenova nejednakost, matematičko očekivanje, konveksna funkcija, slučajna varijabla*

Jensen's inequality in probability theory

Abstract

In this paper Jensen's inequality for mathematical expectation is proved, and the classical Jensen's inequality is derived from it. At the end several applications of Jensen's inequality are given.

Keywords: *Jensen's inequality, mathematical expectation, convex function, random variable*

*Fakultet za matematiku, Sveučilište u Rijeci, email: dkrizmanic@math.uniri.hr

†Fakultet za matematiku, Sveučilište u Rijeci, email: antonio.spac@math.uniri.hr

1 Uvod

Kao i u drugim granama matematike, nejednakosti igraju važnu ulogu u teoriji vjerojatnosti i statistici. Pomoću njih procjenjujemo vjerojatnosti određenih događaja pomoću vjerojatnosti drugih događaja, momenata suma slučajnih varijabli pomoću sume momenata i tome slično. Tako se u teoriji vjerojatnosti često koriste Čebiševljeva, Jensenova i Cauchy-Schwarzova nejednakost, a u statistici Cramer-Raova nejednakost. Uporaba prikladne nejednakosti nerijetko je od ključne važnosti u rješavanju određenog problema, tako je npr. Berry-Esseenova nejednakost omogućila određivanje brzine konvergencije pri normalnoj aproksimaciji u centralnom graničnom teoremu. U ovom članku ćemo se pozabaviti s Jensenovom nejednakošću za matematičko očekivanje. Ova nejednakost je dobila ime po danskom matematičaru i inženjeru Johanu Ludwigu Jensenu (1859.–1925.), koji je pokazao da za konveksnu funkciju $f: I \rightarrow \mathbb{R}$ na otvorenom intervalu $I \subseteq \mathbb{R}$, realne brojeve x_1, x_2, \dots, x_n iz intervala I , te pozitivne realne brojeve a_1, a_2, \dots, a_n vrijedi

$$f\left(\frac{a_1x_1 + a_2x_2 + \dots + a_nx_n}{a_1 + a_2 + \dots + a_n}\right) \leq \frac{a_1f(x_1) + a_2f(x_2) + \dots + a_nf(x_n)}{a_1 + a_2 + \dots + a_n}. \quad (1)$$

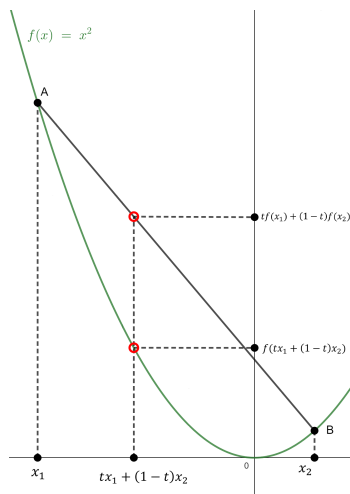
Prisjetimo se, realna funkcija $f: I \rightarrow \mathbb{R}$ na otvorenom intervalu I je konveksna ako za sve $x_1, x_2 \in I$ i $t \in [0, 1]$ vrijedi

$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2).$$

Grafički bi ovo značilo da je funkcija konveksna ako tetiva koja spaja dvije točke grafa te funkcije leži iznad grafa između tih točaka. Ako je funkcija f neprekidna na I , tada je dovoljno da gornja nejednakost vrijedi za $t = 1/2$ kako bi f bila konveksna. Ako u gornjoj nejednakosti zamijenimo " \leq " s " \geq " onda dobivamo konkavnu funkciju. Za funkcije koje imaju drugu derivaciju konveksnost se može ustanoviti pomoću diferencijalnog računa. Vrijedi sljedeći rezultat: Ako je $f: I \rightarrow \mathbb{R}$ dva puta diferencijabilna funkcija na otvorenom intervalu I , tada je ona konveksna na I ako i samo ako je $f''(x) \geq 0$ za svaki $x \in I$.

2 Matematičko očekivanje

U ovom članku dokazat ćemo "vjerojatnosnu" verziju Jensenove nejednakosti, iz koje će (1) lagano slijediti. Prije iskaza ovog rezultata prisjetimo se osnovnih pojmova o matematičkom očekivanju iz teorije vjerojatnosti. Matematičko očekivanje diskretne slučajne varijable X koja poprima prebrojivo mnogo različitih vrijednosti x_1, x_2, x_3, \dots (tu ćemo uključivati i slučaj



Slika 1. Primjer konveksne funkcije

kada X poprima konačno mnogo vrijednosti), s vjerojatnostima

$$p_i = P(X = x_i),^1 \quad i \in \mathbb{N},$$

dano je formulom

$$EX = \sum_{i=1}^{\infty} x_i p_i, \quad (2)$$

ukoliko red na desnoj strani gornje jednakosti apsolutno konvergira. Intuitivno možemo reći da je očekivanje "srednja vrijednost" slučajne varijable. Za apsolutno neprekidnu slučajnu varijablu X s funkcijom gustoće f , što znači da vrijedi

$$P(X \leq x) = \int_{-\infty}^x f(t) dt, \quad x \in \mathbb{R},$$

matematičko očekivanje je jednako

$$EX = \int_{-\infty}^{\infty} t f(t) dt,$$

¹Pri čemu je $p_i \geq 0$ za svaki $i \in \mathbb{N}$, te $\sum_{i=1}^{\infty} p_i = 1$.

ukoliko integral na desnoj strani zadnje jednakosti postoji. Varijanca slučajne varijable X definira se s

$$\text{Var } X = E[(X - EX)^2], \quad (3)$$

ukoliko očekivanja u (3) postoje. Iz definicije i svojstava matematičkog očekivanja lako se pokazuje da je $\text{Var} X = EX^2 - (EX)^2$, te da je varijanca nenegativna, što povlači $EX^2 \geq (EX)^2$. Ako stavimo $g(x) = x^2$, $x \in \mathbb{R}$, tada zadnju nejednakost možemo pisati u obliku $Eg(X) \geq g(EX)$. Uočimo da je funkcija g konveksna. Jensenova nejednakost kaže da zadnja nejednakost vrijedi za svaku konveksnu funkciju g .

3 Jensenova nejednakost

U ovom poglavlju iskazat ćemo i dokazati Jensenovu nejednakost za matematičko očekivanje. Dokaz je preuzet iz [7], a temelji se na primjeni određenih rezultata iz matematičke analize za nizove realnih brojeva (detalji se mogu pogledati u npr. [4]) i teorije vjerojatnosti za matematičko očekivanje [6]. Onda ćemo pomoću ovog rezultata izvesti diskretnu Jensenovu nejednakost (1).

Teorem 3.1 (Jensenova nejednakost). *Neka je X slučajna varijabla, $f: I \rightarrow \mathbb{R}$ konveksna funkcija te pretpostavimo da postoje očekivanja EX i $Ef(X)$. Tada vrijedi*

$$f(EX) \leq Ef(X).$$

Dokaz. Za početak pokažimo da je funkcija

$$g(x) = \frac{f(x) - f(x_0)}{x - x_0}, \quad x \neq x_0,$$

monotono rastuća na $S = \mathbb{R} \setminus \{x_0\}$, za proizvoljni $x_0 \in \mathbb{R}$. Neka su $x, y \in S$ tako da je $x < y$. Imamo 3 slučaja:

(i) Prvi slučaj: $x < y < x_0$. Zapišimo y na sljedeći način

$$y = \frac{x_0 - y}{x_0 - x} \cdot x + \frac{y - x}{x_0 - x} \cdot x_0.$$

Tada, uz oznaku $\alpha = \frac{x_0 - y}{x_0 - x} \in \langle 0, 1 \rangle$, imamo

$$\begin{aligned} f(y) &= f(\alpha x + (1 - \alpha)x_0) \\ &\leq \alpha f(x) + (1 - \alpha)f(x_0) \\ &= \frac{x_0 - y}{x_0 - x} f(x) + \frac{y - x}{x_0 - x} f(x_0), \end{aligned}$$

pri čemu nejednakost u gornjoj relaciji slijedi iz definicije konveksnosti. Odavde slijedi

$$\begin{aligned} f(y) - f(x_0) &\leq \frac{x_0 - y}{x_0 - x} f(x) + \frac{y - x}{x_0 - x} f(x_0) - f(x_0) \\ &= \frac{x_0 - y}{x_0 - x} f(x) + (1 - \alpha - 1) f(x_0) \\ &= \frac{x_0 - y}{x_0 - x} f(x) - \frac{x_0 - y}{x_0 - x} f(x_0) \\ &= \frac{x_0 - y}{x_0 - x} (f(x) - f(x_0)). \end{aligned}$$

Podijelimo li lijevu i desnu stranu s $x_0 - y$ slijedi

$$\frac{f(y) - f(x_0)}{x_0 - y} \leq \frac{f(x) - f(x_0)}{x_0 - x},$$

te pomnožimo li s -1 dobivamo

$$\frac{f(y) - f(x_0)}{y - x_0} \geq \frac{f(x) - f(x_0)}{x - x_0},$$

tj.

$$g(y) \geq g(x).$$

(ii) Drugi slučaj: $x < x_0 < y$. Kako je

$$x_0 = \frac{y - x_0}{y - x} \cdot x + \frac{x_0 - x}{y - x} \cdot y,$$

uz oznaku $\beta = \frac{y - x_0}{y - x} \in \langle 0, 1 \rangle$, slijedi

$$f(x_0) = f(\beta x + (1 - \beta)y) \leq \beta f(x) + (1 - \beta)f(y),$$

tj.

$$(\beta + (1 - \beta))f(x_0) \leq \beta f(x) + (1 - \beta)f(y)$$

$$\beta f(x_0) - \beta f(x) \leq (1 - \beta)f(y) - (1 - \beta)f(x_0)$$

$$\frac{y - x_0}{y - x} (f(x_0) - f(x)) \leq \frac{x_0 - x}{y - x} (f(y) - f(x_0)).$$

Pomnožimo zadnju nejednakost s $\frac{y-x}{(y-x_0)(x_0-x)}$, pa dobivamo

$$\frac{f(x_0) - f(x)}{x_0 - x} \leq \frac{f(y) - f(x_0)}{y - x_0},$$

tj.

$$\frac{f(x) - f(x_0)}{x - x_0} \leq \frac{f(y) - f(x_0)}{y - x_0}.$$

Dakle, dobili smo

$$g(x) \leq g(y).$$

(iii) Treći slučaj: $x_0 < x < y$. Uočimo da je

$$x = \frac{y-x}{y-x_0} \cdot x_0 + \frac{x-x_0}{y-x_0} \cdot y,$$

pa uz oznaku $\gamma = \frac{y-x}{y-x_0} \in \langle 0, 1 \rangle$, imamo

$$f(x) = f(\gamma x_0 + (1-\gamma)y) \leq \gamma f(x_0) + (1-\gamma)f(y).$$

Oдавde slijedi

$$\begin{aligned} f(x) - f(x_0) &\leq \frac{y-x}{y-x_0} f(x_0) - f(x_0) + \frac{x-x_0}{y-x_0} f(y) \\ &= \left(\frac{y-x}{y-x_0} - 1 \right) f(x_0) + \frac{x-x_0}{y-x_0} f(y) \\ &= \frac{x_0-x}{y-x_0} f(x_0) + \frac{x-x_0}{y-x_0} f(y) \\ &= -\frac{x-x_0}{y-x_0} f(x_0) + \frac{x-x_0}{y-x_0} f(y). \end{aligned}$$

Podijelimo li lijevu i desnu stranu s $x - x_0$ dobivamo

$$\frac{f(x) - f(x_0)}{x - x_0} \leq \frac{-f(x_0)}{y - x_0} + \frac{f(y)}{y - x_0} = \frac{f(y) - f(x_0)}{y - x_0},$$

odnosno

$$g(x) \leq g(y).$$

Kako smo dokazali da za proizvoljne $x, y \in S$ takve da je $x < y$ vrijedi $g(x) \leq g(y)$, zaključujemo da je g monotono rastuća funkcija na skupu S .

Definirajmo sada niz $(x_n)_n$ s

$$x_n = x_0 - \frac{1}{n}, \quad n \in \mathbb{N}.$$

Kako je funkcija g monotono rastuća, te vrijedi $x_n \leq x_{n+1}$ za svaki $n \in \mathbb{N}$, slijedi $g(x_n) \leq g(x_{n+1})$, pa zaključujemo da je niz $(g(x_n))_n$ monotono rastući. Kako za svaki prirodan broj n vrijedi

$$x_n = x_0 - \frac{1}{n} < x_0 < x_0 + 1,$$

primijenivši još jednom svojstvo rasta funkcije g dobivamo

$$g(x_n) \leq g(x_0 + 1) \quad \text{za svaki } n \in \mathbb{N},$$

što znači da je niz $(g(x_n))_n$ omeđen odozgo (pri čemu mu je $g(x_0 + 1)$ jedna gornja međa). Iz realne analize znamo da je svaki monotono rastući i odozgo omeđen niz realnih brojeva konvergentan, pa zaključujemo da je niz $(g(x_n))_n$ konvergentan. Označimo njegov limes s

$$a = \lim_{n \rightarrow \infty} g(x_n).$$

Stavimo sada

$$y_n = x_0 + \frac{1}{n}, \quad n \in \mathbb{N}.$$

Slično kao i prije zaključuje se da je niz $(g(y_n))_n$ monotono padajući i omeđen odozdo te iz toga slijedi da je i konvergentan. Označimo njegov limes s

$$b = \lim_{n \rightarrow \infty} g(y_n).$$

Kako za svaki prirodan broj n vrijedi

$$x_n < x_0 < y_n,$$

zbog monotonog rasta funkcije g imamo $g(x_n) \leq g(y_n)$ za svaki $n \in \mathbb{N}$. Odavde, puštajući n u beskonačno dobivamo

$$\lim_{n \rightarrow \infty} g(x_n) \leq \lim_{n \rightarrow \infty} g(y_n),$$

tj. $a \leq b$. Neka je sada $c \in [a, b]$ proizvoljan (u slučaju da je $a = b$ uzmemo $c = a = b$). Definirajmo funkciju $h : \mathbb{R} \rightarrow \mathbb{R}$ s

$$h(x) = c(x - x_0) + f(x_0).$$

Primijetimo da vrijedi $h(x_0) = f(x_0)$. Dokažimo da za svaki $x \in \mathbb{R}$ vrijedi $h(x) \leq f(x)$. Imamo tri slučaja:

(i) Prvi slučaj: $x > x_0$. Kako za dovoljno veliki prirodan broj n vrijedi

$$x_0 < x_0 + \frac{1}{n} = y_n < x,$$

zbog svojstva monotonog rasta funkcije g slijedi

$$c \leq b \leq g(y_n) \leq g(x) = \frac{f(x) - f(x_0)}{x - x_0}.$$

Pomnoživši lijevu i desnu stranu s $x - x_0$ dobivamo

$$c(x - x_0) \leq f(x) - f(x_0)$$

$$c(x - x_0) + f(x_0) \leq f(x),$$

tj.

$$h(x) \leq f(x).$$

(ii) Drugi slučaj: $x < x_0$. Za dovoljno veliki prirodan broj n vrijedi

$$x < x_0 - \frac{1}{n} = x_n < x_0,$$

pa iz toga dobivamo

$$c \geq a \geq g(x_n) \geq g(x) = \frac{f(x) - f(x_0)}{x - x_0}.$$

Pomnožimo lijevu i desnu stranu s $x - x_0$, pa imamo

$$c(x - x_0) \leq f(x) - f(x_0)$$

$$h(x) \leq f(x).$$

(iii) Treći slučaj: $x = x_0$. U ovom slučaju je $h(x) = f(x)$ pa ujedno i $h(x) \leq f(x)$.

Dakle, dokazali smo da je $h(x) \leq f(x)$ za svaki $x \in \mathbb{R}$.

Neka je Ω skup (prostor elementarnih događaja) na kojem je definirana slučajna varijabla X , tj. $X: \Omega \rightarrow \mathbb{R}$. Uzmimo $\omega \in \Omega$ proizvoljan. Stavimo $x_0 = EX$ i $x = X(\omega)$. Po prethodno dobivenom vrijedi $h(x) \leq f(x)$, tj.

$$c(X(\omega) - EX) + f(EX) \leq f(X(\omega)).$$

Iz ove nejednakosti koristeći svojstva monotonosti² i linearnosti³ matematičkog očekivanja, te činjenicu da je očekivanje konstante ponovno ta ista konstanta⁴ dobivamo

$$c(EX - EX) + f(EX) \leq E[f(X)],$$

tj. $f(EX) \leq Ef(X)$. □

Pokažimo sada nejednakost (1). Neka je $f: I \rightarrow \mathbb{R}$ konveksna funkcija na otvorenom intervalu $I \subseteq \mathbb{R}$, x_1, x_2, \dots, x_n realni brojevi iz I , te a_1, a_2, \dots, a_n pozitivni realni brojevi. Neka je X slučajna varijabla koja poprima vrijednosti x_1, x_2, \dots, x_n s vjerojatnostima

$$p_i = P(X = x_i) = \frac{a_i}{a_1 + a_2 + \dots + a_n}, \quad i = 1, 2, \dots, n.$$

Iz definicije matematičkog očekivanja (2) slijedi

$$EX = \sum_{i=1}^n x_i p_i = \sum_{i=1}^n x_i \frac{a_i}{a_1 + a_2 + \dots + a_n} = \frac{a_1 x_1 + a_2 x_2 + \dots + a_n x_n}{a_1 + a_2 + \dots + a_n}.$$

Standardni rezultat diskretne teorije vjerojatnosti kaže da je $f(X)$ slučajna varijabla s očekivanjem

$$Ef(X) = \sum_{i=1}^n f(x_i) p_i.$$

Stoga imamo

$$Ef(X) = \frac{a_1 f(x_1) + a_2 f(x_2) + \dots + a_n f(x_n)}{a_1 + a_2 + \dots + a_n}.$$

Sada Teorem 3.1 povlači

$$f\left(\frac{a_1 x_1 + a_2 x_2 + \dots + a_n x_n}{a_1 + a_2 + \dots + a_n}\right) \leq \frac{a_1 f(x_1) + a_2 f(x_2) + \dots + a_n f(x_n)}{a_1 + a_2 + \dots + a_n},$$

dakle zaista vrijedi (1).

²Svojstvo monotonosti očekivanja kaže da ako za slučajne varijable X i Y definirane na skupu Ω za koje postoje očekivanja, vrijedi $X \leq Y$, tj. $X(\omega) \leq Y(\omega)$ za svaki $\omega \in \Omega$, tada je $EX \leq EY$.

³Svojstvo linearnosti očekivanja kaže da za slučajne varijable X i Y definirane na skupu Ω za koje postoje očekivanja vrijedi $E(\alpha X + \beta Y) = \alpha EX + \beta EY$ za sve $\alpha, \beta \in \mathbb{R}$.

⁴Ako slučajna varijabla X poprima samo jednu realnu vrijednost x , tada je i njezino očekivanje jednako x . To lagano slijedi iz (2), jer je $x_1 = x$ i $p_1 = 1$, pa slijedi $EX = x_1 p_1 = x$. Konkretno, u našem slučaju su c, EX i $f(EX)$ konstante, pa su njihova očekivanja te iste vrijednosti, tj. $E(c) = c, E(EX) = EX$ i $E[f(EX)] = f(EX)$.

4 Neke primjene Jensenove nejednakosti

Funkcije $x \mapsto e^x$ i $x \mapsto |x|^p$ (za $p \geq 1$) su konveksne funkcije na skupu \mathbb{R} , pa iz Teorema 3.1 odmah dobivamo da za slučajnu varijablu X vrijedi

$$e^{EX} \leq Ee^X \quad \text{i} \quad (E|X|)^p \leq E|X|^p,$$

ukoliko navedena očekivanja postoje.

Vratimo se sada na nejednakost (1). U slučaju kada su realni brojevi a_1, a_2, \dots, a_n svi međusobno jednaki, nejednakost (1) postaje

$$f\left(\frac{x_1 + x_2 + \dots + x_n}{n}\right) \leq \frac{f(x_1) + f(x_2) + \dots + f(x_n)}{n}. \quad (4)$$

Može se pokazati da jednakost u relaciji (4) vrijedi ako i samo ako je $x_1 = x_2 = \dots = x_n$ ili je f linearna funkcija.

Funkcija $f(x) = -\ln x$ je konveksna funkcija na skupu pozitivnih realnih brojeva, pa iz (4) slijedi

$$-\ln\left(\frac{x_1 + x_2 + \dots + x_n}{n}\right) \leq -\frac{\ln x_1 + \ln x_2 + \dots + \ln x_n}{n},$$

odakle koristeći elementarna svojstva logaritamske funkcije dobivamo

$$\frac{x_1 + x_2 + \dots + x_n}{n} \geq \sqrt[n]{x_1 x_2 \dots x_n},$$

što je dobro poznata aritmetičko-geometrijska nejednakost. Osim ove nejednakosti između aritmetičke i geometrijske sredine, iz diskretne Jensenove nejednakosti (1) se mogu izvesti i neke druge poznate nejednakosti među sredinama, te Youngova nejednakost, Cauchyjeva nejednakost, Hölderova nejednakost i nejednakost Minkowskog. Više o tome može se naći u [2] i [5]. Različite primjene Jensenove nejednakosti ilustrirane su u [3].

Pokažimo i jednu primjenu Jensenove nejednakosti u geometriji. Neka su a_1, a_2, \dots, a_n duljine stranica (jednostavnog) mnogokuta, a o njegov opseg. Pokažimo da vrijedi nejednakost

$$\frac{a_1^2}{o - a_1} + \frac{a_2^2}{o - a_2} + \dots + \frac{a_n^2}{o - a_n} \geq \frac{o}{n - 1}.$$

Promotrimo funkciju

$$f(x) = \frac{x^2}{o - x}, \quad x \neq o.$$

Kako je $f'(x) = \frac{2ox-x^2}{(o-x)^2}$ i $f''(x) = \frac{2o^2}{(o-x)^3}$, to je $f''(x) > 0$ za $0 < x < o$. Iz toga slijedi da je f konveksna na intervalu $(0, o)$. Prema nejednakosti (4) vrijedi

$$\frac{\left(\frac{a_1+a_2+\dots+a_n}{n}\right)^2}{o - \frac{a_1+a_2+\dots+a_n}{n}} \leq \frac{\frac{a_1^2}{o-a_1} + \frac{a_2^2}{o-a_2} + \dots + \frac{a_n^2}{o-a_n}}{n},$$

te sređivanjem ovog izraza dobivamo

$$\frac{a_1^2}{o-a_1} + \frac{a_2^2}{o-a_2} + \dots + \frac{a_n^2}{o-a_n} \geq \frac{o}{n-1}.$$

Specijalno, za trokut sa stranicama duljina a, b, c te poluopsegom s dobije se

$$\frac{a^2}{b+c} + \frac{b^2}{a+c} + \frac{c^2}{a+b} \geq s.$$

Spomenimo na kraju kako se u teoriji informacija pomoću Jensenove nejednakosti pokazuje nenegativost relativne entropije ili Kullback-Leiblerove divergencije jedne vjerojatnosne distribucije u odnosu na drugu (više o ovim pojmovima može se pročitati u [1]). Ova veličina mjeri udaljenost između dviju vjerojatnosnih distribucija, te je u slučaju diskretnih slučajnih varijabi X i Y s istim skupom vrijednosti $\{x_1, x_2, \dots\}$ i vjerojatnostima $p_i = P(X = x_i)$ i $q_i = P(Y = x_i)$, definirana s

$$D_{KL}(X||Y) = - \sum_{i=1}^{\infty} p_i \ln \left(\frac{q_i}{p_i} \right).$$

Relativna entropija nije simetrična, tj. općenito $D_{KL}(X||Y) \neq D_{KL}(Y||X)$.

Literatura

- [1] Z. Bertić, *Neke metode za analizu kvalitete procijenjene particije*, diplomski rad, Odjel za matematiku Sveučilišta J. J. Strossmayera u Osijeku, Osijek, 2015.
- [2] K. Bošnjak, *Jensenova nejednakost i nejednakosti izvedene iz nje*, diplomski rad, Odjel za matematiku Sveučilišta J. J. Strossmayera u Osijeku, Osijek, 2012.
- [3] I. Ilišević, *Jensenova nejednakost*, Osječki matematički list, 5 (2005), 9-19.
- [4] S. Kurepa, *Matematička analiza 2*, Tehnička knjiga, Zagreb, 1987.

- [5] M. Ribičić Penava, K. Bošnjak, *Jensenova nejednakost i nejednakosti izvedene iz nje*, Osječki matematički list, **16** (2016), 15-25.
- [6] N. Sarapa, *Teorija vjerojatnosti*, Školska knjiga, Zagreb, 2002.
- [7] A. Špac, *Vjerojatnosne nejednakosti*, diplomski rad, Fakultet za matematiku Sveučilišta u Rijeci, Rijeka, 2023.