# Sadržaj

# Poglavlje 10

# ZAKONI VELIKIH BROJEVA CENTRALNI GRANIČNI TEOREM *tko želi znati više*

Najvažniji teoremi koji će se koristiti u matematičkoj statistici označeni su s *važno*. Preporučujemo i motivirajuće primjere za Čebiševljevu nejednakost, za zakon velikih brojeva i centalni granični teorem.

Čebiševljeva nejednakost, zakoni velikih brojeva i cenralni granični teoremi su važan alati koji otkrivaju svojstva diskretnih ili kontinuiranih slučajnih varijabli koje imaju konačno očekivanje i varijancu ako nam i nije poznata njihova distribucija.

Čebiševljeva nejednakost daje ocjenu vjerojatnosti da se vrijednosti slučajne varijable razlikuje od očekivanja više od zadonog $\varepsilon$.

Zakoni velikih brojeva su skup teorema koji se odnosi na granične vrijednosti niza slučajnih varijabli.

Neka su $X_1, X_2, ..., X_n$ nezavisna mjerenja slučajne varijable $X$ u ponovljenim pokusima. Slučajne varijable $X_1, X_2, ..X_n$ su nezavisne i sve imaju istu distribuciju kao i slučajna varijabla $X$.
Može se uočiti da njihova aritmetička sredina $\overline{X} = \frac{1}{n}(X_1 + X_2 + .. + X_n)$ ima svojstvo stabilnosti distribucije i da je vjerojatnost da se vrijednosti od $\overline{X}$ razlikuje od očekivanja više od zadonog $\varepsilon$ jednaka nuli kad $n \to \infty$.

Centralni granični teoremi se odnose na granične zakone distribucije niza slučajnih varijabli. Suma velikog broja slučajnih varijabli ima standardnu normalnu distribuciju. Teoremi daju različite uvjete na pribrojnike u toj sumi.

## 10.1 ČEBIŠEVLJEVA NEJEDNAKOST

**MOTIV 10.1** *Neka slučajna varijabla ima varijancu (disperziju)* $Var(X) = 0.001$. *Kolika je vjerojatnost da slučajna varijabla odstupa od očekivanja manje od* $\varepsilon = 0.1$?

**TEOREM 10.1** *(MARKOVLJEVA NEJEDNAKOST)*

*Neka je X slučajna varijabla s nenegativnim vrijednostima i konačnim očekivanjem* $E(X)$.
*Tada* $\forall a > 0$ *vrijedi*

$$P(X \geq a) \leq \frac{E(X)}{a}$$

**Dokaz:**
Možemo pokazati za kontinuiranu slučajnu varijablu s funkcijom gustoće $f(x)$ (analogno za diskretnu).
Kako X poprima samo nenegativne vrijednosti uočimo one $x \in R(X)$,
$0 \leq x < a$ i $x \geq a$.
Prema definiciji očekivanja: $E(X) = \int_0^\infty xf(x)dx = \int_0^a xf(x)dx + \int_a^\infty xf(x)dx.$
Budući su integrali pozitivni vrijedi nejednakost

$$E(X) \geq \int_a^\infty xf(x)dx.$$

Ako u podintegralnoj funkciji zamijenimo x s konstantom koja je uvijek manja od x, $a \leq x$, zadržat će se znak nejednakosti

$$E(X) \geq a \int_a^\infty f(x)dx.$$

Prema definiciji funkcije gustoće vjerojatnosti f(x) dobivamo traženu nejednakost:

$$E(X) \geq a \cdot P(X \geq a).$$

**TEOREM 10.2** *(POOPĆENJE MARKOVLJEVE NEJEDNAKOSTI)*

*Neka je X slučajna varijabla i* $h : R \to R$ *nenegativna funkcija tako da postoji očekivanjem* $E(h(X))$.
*Tada* $\forall a > 0$ *vrijedi*

$$P(h(X) \geq a) \leq \frac{E(h(X))}{a}$$

**Dokaz:**
Kako h(X) poprima samo nenegativne vrijednosti uočimo one $x \in R(X)$,
$0 \leq h(x) < a$ , za x$\in$ $D_1 \subseteq R$, i $h(x) \geq a$, za $x \in D_2 \subseteq R$ :
Prema definiciji očekivanja funkcije slučajne varijable $X$:

$$E(h(X)) = \int_0^\infty h(x)f(x)dx = \int_{D_1} h(x)f(x)dx + \int_{D_2} h(x)f(x)dx.$$

Budući su integrali pozitivni vrijedi nejednakost

$$\begin{aligned} E(h(X)) &\geq \int_{D_2} h(x)f(x)dx \geq a\int_{D_2} f(x)dx \\ &\geq a \cdot P(h(X) \geq a). \end{aligned}$$

**TEOREM 10.3** *važno*
*(ČEBIŠEVLJEVA NEJEDNAKOST, engl. Chebyshev's inequality)*

*Neka je X slučajna varijabla s konačnom varijancom $Var(X)$.*
*Tada $\forall \varepsilon > 0$ vrijedi*

$$P(|X - E(X)| \geq \varepsilon) \leq \frac{Var(X)}{\varepsilon^2},$$

$$P(|X - E(X)| < \varepsilon) \geq 1 - \frac{Var(X)}{\varepsilon^2}.$$

*Ako označimo $Var(X) = \sigma^2, E(X) = \mu, \varepsilon = \lambda\sigma$, onda vrijedi*

$$P(|X - \mu| \geq \lambda\sigma) \leq \frac{1}{\lambda^2},$$

$$P(|X - \mu| < \lambda\sigma) \geq 1 - \frac{1}{\lambda^2}.$$

**Dokaz:**
(a) Dokaz pomoću generalizirane Markovljeve nejednakosti.
Pretpostavka teorema je da X ima varijancu tj. ima očekivanje
$E((X - E(X)^2) = Var(X)$ pa možemo primijeniti teorem za nenegativnu funkciju $h(x) = (x - E(X))^2$.

Vrijedi nejednakost $P((X - E(X))^2 \geq a) \leq \dfrac{E((X - E(X))^2)}{a}$, $\forall a > 0$.

Budući je $P((x - E(X))^2 \geq a) = P(|X - E(X)| \geq \sqrt{a})$, vrijedi nejednakost

$$P(|X - E(X)| \geq \sqrt{a}) \leq \frac{E((X - E(X))^2)}{a}, \ \forall a > 0.$$

Tada $\forall \varepsilon > 0$ vrijedi

$$P(|X - E(X)| \geq \varepsilon) \leq \frac{Var(X)}{\varepsilon^2}.$$

(b) Dokaz za diskretnu slučajnu varijablu (bez generalizirane Markovljeve nejednakosti)

Neka je X diskretna slučajna varijabla sa slikom $R(X) = \{x_1, x_2, ...\}$ i neka je $f(x)$ njena funkcija vjerojatnosti.

Pretpostavka teorema je da X ima konačnu varijancu (i očekivanje).

$Var(X) = \sum_{x_i}(x_i - (E(X))^2 f(x_i)$.

Uočimo one $x_i \in R(X)$ za koje je $|x_i - (E(X)| < \varepsilon$ i one za koje je $|x_i - (E(X)| \geq \varepsilon$.

$$\begin{aligned} Var(X) &\geq \sum_{x_i:\,|x_i-(E(X)|<\varepsilon} (x_i - (E(X))^2 f(x_i) + \sum_{x_i:\,|x_i-(E(X)|\geq\varepsilon} (x_i - (E(X))^2 f(x_i) \\ &\geq \sum_{x_i:\,|x_i-(E(X)|\geq\varepsilon} (x_i - (E(X))^2 f(x_i). \end{aligned}$$

Zamjenom svakog člana sume s manjom konstantom $\varepsilon$, nejednakost se zadržava pa vrijedi:

$$Var(X) \geq \sum_{x_i:\,|x_i-(E(X)|\geq\varepsilon} \varepsilon^2 f(x_i) \geq \varepsilon^2 \cdot \sum_{x_i:\,|x_i-(E(X)|\geq\varepsilon} f(x_i).$$

Suma $\displaystyle\sum_{x_i:\,|x_i-(E(X)|\geq\varepsilon} f(x_i)$ je prema definiciji funkcije distribuije jednaka $P(|X - (E(X)| \geq \varepsilon)$ i dobivamo konačnu nejednakost:

$$Var(X) \geq \varepsilon^2 \cdot P(|x_i - (E(X)| \geq \varepsilon).$$

Prisjetimo se motivacijskog primjera:

**PRIMJER 10.1** *motiv*

*Neka slučajna varijabla ima varijancu (disperziju)* $Var(X) = 0.001$. *Kolika je vjerojatnost da slučajna varijabla odstupa od očekivanja manje od* $\varepsilon = 0.1$*?*

**Rješenje:**

Trebamo izračunati $P(|X - E(X)| < \varepsilon)$.

Koristimo Čebiševljevu nejednakost u obliku $P(|X - E(X)| < \varepsilon) \geq 1 - \frac{Var(X)}{\varepsilon^2}$,

$$P(|X - E(X)| < 0.1) \geq 1 - \frac{0.001}{0.1^2} \geq 0.9.$$

**PRIMJER 10.2** *Slučajna varijabla ima očekivanje* $\mu = 3$ *i standardnu devijaciju* $\sigma = 0.1$. *Ocijenite* $P(2.5 < X < 3.5)$.

**Rješenje:**
Zadano je očekivanje $\mu = 3$ i uočimo da je
$P(2.5 < X < 3.5) = P(\mu - 0.5 < X < \mu + 0.5) = P(|X - \mu| < 0.5)$.
Koristimo Čebiševljevu nejednakost u obliku: $P(|X - \mu| < \varepsilon) \geq 1 - \frac{\sigma^2}{\varepsilon^2}$,
$P(|X - \mu| < 0.5) \geq 1 - \frac{0.1^2}{0.5^2} = 0.96$,

$$P(2.5 < X < 3.5) \geq 0.96.$$

## 10.2 ZAKON VELIKIH BROJEVA

**MOTIV 10.2** *Kontrolor uzima uzorak veličine $m = 1000$ iz skupa uređaja. Vjerojatnost da je uređaj neispravan je $p = 0.03$. U kojim granicama će biti broj neispravnih uređaja u uzorku s vjerojatnošću $\gamma = 0.99$.*

**Definicija 10.1** *(KONVERGENCIJA PO VJEROJATNOSTI)*

*Neka je $(X_n)$, $n \in N$ niz slučajnih varijabli. Ako postoji slučajna varijabla $X$ takva da*

$$\forall \varepsilon > 0 \ \ \lim_{n\to\infty} P(|X_n - X| < \varepsilon) = 1$$

*kažemo da niz $(X_n)$ slučajnih varijabli konvergira slučajnoj varijabli $X$ po vjerojatnosti i označavamo*

$$(P) \lim_{n\to\infty} X_n = X.$$

**Definicija 10.2** *Za niz slučajnih varijabli $(X_n), n \in N$ kažemo da zadovoljava zakon velikih brojeva ako postoji konstanta $C$ takva da vrijedi*

$$\forall \varepsilon > 0 \ \ \lim_{n\to\infty} P(|X_n - C| < \varepsilon) = 1$$

*i označavamo*

$$(P) \lim_{n\to\infty} X_n = C.$$

**TEOREM 10.4** *važno*
*(ZAKON VELIKIH BROJEVA*
*(specijalan slučaj za aritmetičku sredinu))*

*Neka je $\{X_n\}, n \in N$ niz slučajnih varijabli takvih da za svaki $n$ slučajne varijable $X_1, X_2, ..X_n$ su nezavisne, imaju ograničenu varijancu*
*$Var(X_i) = \sigma^2 \le M > 0$ i $E(X_i) = \mu,\ i = 1, ..., n$.*
*Tada za aritmetičku sredinu*

$$\overline{X} = \frac{1}{n}(X_1 + X_2 + .. + X_n)$$

*vrijedi*

$$\forall \varepsilon > 0 \ \ \lim_{n\to\infty} P(|\overline{X} - \mu| < \varepsilon) = 1,$$

$$(P) \lim_{n\to\infty} \overline{X} = \mu.$$

**Dokaz:**
$\overline{X} = \dfrac{1}{n}(X_1 + X_2 + .. + X_n),\ E(\overline{X}) = \mu,\ Var(\overline{X}) = \dfrac{\sigma^2}{n}.$
Koristimo Čebiševljevu nejednakost za slučajnu varijablu $\overline{X}$ u obliku:

$P(|\overline{X} - E(\overline{X})| < \varepsilon) \geq 1 - \dfrac{Var(\overline{X})}{\varepsilon^2},$

$$P(|\overline{X} - \mu| < \varepsilon) \geq 1 - \frac{\sigma}{n \cdot \varepsilon^2},$$

$$\lim_{n\to\infty} P(|\frac{1}{n}(X_1 + X_2 + .. + X_n) - \mu| < \varepsilon) = 1.$$

**TEOREM 10.5** *(BERNOULLIJEV SLABI ZAKON VELIKIH BROJEVA - za rel.frekv binomne sl. varijable)*

*Neka je u Bernoullijevoj shemi slučajna varijabla X=broj uspjeha događaja A u m nezavisnih ponavljanja,* $P(A) = p$. $X \sim B(m, p)$. *Slučajna varijabla* $Y = \frac{X}{m}$ *zove se relativna frekvencija uspjeha događaja A u Bernoullijevoj shemi.*
*Tada vrijedi*

$$\lim_{m\to\infty} \sum_{|\frac{k}{m} - p| < \varepsilon} \binom{m}{k} p^k (1-p)^{m-k} = \lim_{m\to\infty} P(|\frac{X}{m} - p| < \varepsilon) = 1,$$

$$(P) \lim_{m\to\infty} \frac{X}{m} = p.$$

**Dokaz:**
Za binomnu slučajnu varijablu $X \sim B(m, p)$, $E(X) = mp$,
$Var(X) = mp(1-p)$.
Za relativnu frekvenciju uspjeha $Y = \frac{X}{m}$ odredimo očekivanje i varijancu:
$E(\frac{X}{m}) = p,\ Var(\frac{X}{m}) = \frac{p(1-p)}{m}.$
Primijenimo Čebiševljevu nejednakost za slučajnu varijablu $Y = \frac{X}{m}$ nejednakost u obliku: $P(|Y - E(Y)| < \varepsilon) \geq 1 - \frac{Var(Y)}{\varepsilon^2}$,

$$P(|\frac{X}{m} - p| < \varepsilon) \geq 1 - \frac{p(1-p)}{m \cdot \varepsilon^2},$$

$$\lim_{m\to\infty} P(|\frac{X}{m} - p| < \varepsilon) = 1.$$

Kako je funkcija vjerojatnosti binomne slučajne varijable
$f(x) = \binom{m}{x} p^x (1-p)^{m-x}$

$$\lim_{m\to\infty} \sum_{|\frac{k}{m}-p|<\varepsilon} \binom{m}{k} p^k (1-p)^{m-k} = \lim_{m\to\infty} P(|\frac{X}{m} - p| < \varepsilon) = 1.$$

**NAPOMENA 10.1** *važno*

*Oblik Bernoullijevog slabog zakona velikih brojeva*

$$P(|\frac{X}{m} - p| < \varepsilon) \geq 1 - \frac{p(1-p)}{m \cdot \varepsilon^2}$$

*često se koristi u zadacima za određivanje*
*(a) minimalnog broja pokusa m*
*(b) odstupanja* $\varepsilon$
*da bi za zadani* $\gamma$, $P(|\frac{X}{m} - p| < \varepsilon) \geq \gamma$.

*Rješenje:*

*(a)* $m \geq \dfrac{1}{1-\gamma} \cdot \dfrac{1}{\varepsilon^2} \cdot p(1-p)$.

*(b)* $\varepsilon^2 \geq \dfrac{1}{1-\gamma} \cdot \dfrac{1}{m} \cdot p(1-p)$.

*Ako je p nepoznato onda se procjenjuje da je* $p(1-p) \leq \frac{1}{4}$, *i*

*(a)* $m \geq \dfrac{1}{1-\gamma} \cdot \dfrac{1}{4 \cdot \varepsilon^2}$,

*(b)* $\varepsilon^2 \geq \dfrac{1}{1-\gamma} \cdot \dfrac{1}{4 \cdot m}$.

**PRIMJER 10.3** *U Bernoullijevoj shemi vjerojatnost događaja A je* $p = P(A) = 1/3$. *Odredite minimalan broj ponavljanja tako da s vjerojatnošću ne manjom od* $\gamma = 0.99$ *apsolutno ostupanje relativne frekvencije od p bude najviše* $\varepsilon = 0.01$.

**Rješenje:**

Trebamo odrediti $m$ tako da $P(|\frac{X}{m} - p| < \varepsilon) \geq 0.99$.
Koristimo Bernoullijev slabi zakon velikih brojeva: $P(|\frac{X}{m} - p| < \varepsilon) \geq 1 - \frac{p(1-p)}{m \cdot \varepsilon^2}$

$$P(|\frac{X}{m} - \frac{1}{3}| < 0.01) \geq 1 - \frac{\frac{2}{9}}{m \cdot 0.01^2}$$

Broj ponavljanja ćemo odrediti iz zadane vjerojatnosti $\gamma$ i uvjeta
$1 - \frac{\frac{2}{9}}{m\cdot(0.01)^2} \geq 0.99$

$$m \geq \frac{1}{1-\gamma} \cdot \frac{1}{\varepsilon^2} \cdot p(1-p) = \frac{2}{9} \cdot \frac{1}{1-0.99} \cdot \frac{1}{(0.01)^2} \quad \Rightarrow \quad m \geq 222222.$$

Napomena (zadatak ćemo riješiti i koristeći Moivre-Laplaceov teorem-poslije).

Prisjetimo se motivacijskog primjera:

**PRIMJER 10.4** *motiv*

*Kontrolor uzima uzorak veličine* $m = 1000$ *iz skupa uređaja. Vjerojatnost da je uređaj neispravan je* $p = 0.03$. *U kojim granicama će biti broj neispravnih uređaja u uzorku s vjerojatnošću* $\gamma = 0.99$.

**Rješenje:**
Treba oderditi a i b takve da za $X$ = broj neispravnih uređaja u uzorku veličine $m$, $X \sim B(m,p)$, $P(a < X < b) \geq 0.99$.
Koristimo Bernoullijev slabi zakon velikih brojeva:
$P(|\frac{X}{m} - p| < \varepsilon) \geq 1 - \frac{p(1-p)}{m\cdot\varepsilon^2}$,

$$P(|\frac{X}{1000} - 0.03| < \varepsilon) \geq 1 - \frac{0.03(1-0.03)}{1000\cdot\varepsilon^2}.$$

Odstupanje $\varepsilon$ ćemo odrediti iz zadane vjerojatnosti $\gamma$ i uvjeta
$1 - \frac{0.03(1-0.03)}{1000\cdot\varepsilon^2} \geq 0.99$.

$$\varepsilon^2 \geq \frac{1}{1-\gamma} \cdot \frac{1}{m} \cdot p(1-p) = \frac{1}{1-0.99} \cdot \frac{1}{1000} \cdot 0.03 \cdot (1-0.03) = 0.003.$$

$$\varepsilon \geq \sqrt{0.003} = 0.054 \quad \Rightarrow \quad |\frac{X}{1000} - 0.03| < 0.054 \quad \Rightarrow \quad 0 < X < 80.$$

Napomena (zadatak ćemo riješiti i koristeći Moivre-Laplaceov teorem-poslije).

## 10.3 CENTRALNI GRANIČNI TEOREM

Motivirajući primjer je isti kao i poglavlju Zakon velikih brojeva, ali će se sad riješiti primjenom centralnog graničnog teorema (Moivre - Laplaceov teorem).

**MOTIV 10.3** *Kontrolor uzima uzorak veličine $m = 1000$ iz skupa uređaja. Vjerojatnost da je uređaj neispravan je $p = 0.03$. U kojim granicama će biti broj neispravnih uređaja u uzorku s vjerojatnošću $\gamma = 0.99$*

**TEOREM 10.6** *(CENTRALNI GRANIČNI TEOREM-CGT; specijalni slučaj)*

*Neka je $\{X_n\}, n \in N$ niz slučajnih varijabli takvih da za svaki $n$ slučajne varijable $X_1, X_2, ..., X_n$ su nezavisne, imaju ograničenu varijancu*
*$Var(X_i) = \sigma^2 \leq M > 0$ i $E(X_i) = \mu,\ i = 1, ..., n$.*
*Tada*

$$\lim_{n\to\infty} P(a < \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\cdot\sigma} < b) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{1}{2}x^2}dx,$$

$$\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\cdot\sigma} \sim N(0,1), \quad n \to \infty.$$

*Za velike $n$ vrijedi*

$$P(a < \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\cdot\sigma} < b) \approx F^*(b) - F^*(a).$$

*($F^*(x)$ funkcija distribucije standardne normale distribucije).*

**Dokaz:** (literatura)

**TEOREM 10.7** *važno*
*(CGT za aritmetičku sredinu)*

*Neka je $\{X_n\}, n \in N$ niz slučajnih varijabli takvih da za svaki $n$ slučajne varijable $X_1, X_2, ..X_n$ su nezavisne, imaju ograničenu varijancu*
*$Var(X_i) = \sigma^2 \leq M > 0$ i $E(X_i) = \mu,\ i = 1, ..., n$.*
*Tada za aritmetičku sredinu $\overline{X} = \frac{1}{n}(X_1 + X_2 + .. + X_n)$ vrijedi*

$$\lim_{n\to\infty} P(a < \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} < b) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{1}{2}x^2}dx,$$

$$\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1), \qquad n \to \infty,$$

*odnosno*

$$\overline{X} \sim N(\mu, \frac{\sigma^2}{n}), \qquad n \to \infty.$$

*Za velike n vrijedi*

$$P(a < \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} < b) \approx F^*(b) - F^*(a).$$

**Dokaz:**
Primijenimo CGT (specijalni slučaj) za $\overline{X} = \frac{1}{n}(X_1 + X_2 + .. + X_n)$.

**PRIMJER 10.5** *Neka su $X_1, X_2, ..., X_n$ nezavisne slučajne varijable, imaju ograničenu varijancu $Var(X_i) = \sigma^2 = 2$ i $E(X_i) = \mu = 3,\ i = 1, ..., n = 3200$. Za aritmetičku sredinu $\overline{X} = \frac{1}{n}(X_1 + X_2 + .. + X_n)$ odredite*
$P(2.95 < \overline{X} < 3.075)$.

**Rješenje:**
Prema CGT za aritmetičku sredinu n slučajnih varijabli
$\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1), \qquad n \to \infty,$

$$\lim_{n\to\infty} P(a < \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} < b) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{1}{2}x^2} dx.$$

$$\begin{aligned}
P(2.95 < \overline{X} < 3.075) &= P(\frac{2.95-\mu}{\frac{\sigma}{\sqrt{n}}} < \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} < \frac{3.075-\mu}{\frac{\sigma}{\sqrt{n}}}) \\
&= P(\frac{2.95-3}{\frac{\sqrt{2}}{\sqrt{3200}}} < \frac{\overline{X}-3}{\frac{\sqrt{2}}{\sqrt{3200}}} < \frac{3.075-3}{\frac{\sqrt{2}}{\sqrt{3200}}}) \\
&= P(-2 < \frac{\overline{X}-3}{\frac{\sqrt{2}}{\sqrt{3200}}} < 3) = F^*(3) - F^*(-2) = 0.975.
\end{aligned}$$

**TEOREM 10.8** *(integralni MOIVRE - LAPLACEOV TEOREM, CGT za binomnu sl. varijablu)*

*Neka je u Bernoullijevoj shemi slučajna varijabla X=broj uspjeha događaja A u m nezavisnih ponavljanja $P(A) = p$. $X \sim B(m, p)$.*
*Tada vrijedi*

$$\lim_{m\to\infty} P(a < \frac{X - m\cdot p}{\sqrt{mp(1-p)}} < b) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{1}{2}x^2}dx,$$

$$\frac{X - m\cdot p}{\sqrt{mp(1-p)}} \sim N(0,1), \qquad m \to \infty,$$

*odnosno*

$$X \sim N(mp, mp(1-p)), \qquad m \to \infty.$$

*Za velike n vrijedi*

$$P(a < \frac{X - m\cdot p}{\sqrt{mp(1-p)}} < b) \approx F^*(b) - F^*(a),$$

*odnosno*

$$P(a < X < b) \approx F^*(\frac{b - mp}{\sqrt{mp(1-p)}}) - F^*(\frac{a - mp}{\sqrt{mp(1-p)}}).$$

**Dokaz:**

Promatrajmo slučajne varijable $X_i \sim B(1,p),\ i = 1,...,m,$ kad $m \to \infty$.
$Var(X_i) = p(1-p)$ i $E(X_i) = p,\ i = 1,...,m.$
Integralni Moivre-Laplaceov teorem je specijalan slučaj CGT za niz slučajnih varijabli $X_i = 1,...,m,$ kad $m \to \infty$

$$\lim_{m\to\infty} P(a < \frac{\sum_{i=1}^{m} X_i - m\mu}{\sqrt{m}\cdot\sigma} < b) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{1}{2}x^2}dx,$$

$$\lim_{m\to\infty} P(a < \frac{\sum_{i=1}^{m} X_i - mp}{\sqrt{m\cdot p(1-p)}} < b) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{1}{2}x^2}dx.$$

Slučajna varijabla $X$=broj uspjeha događaja A u $m$ nezavisnih ponavljanja $P(A) = p$. $X \sim B(m,p)$. Tada je $X = \sum_{i=1}^{m} X_i$ pa vrijedi

$$\lim_{m\to\infty} P(a < \frac{X - mp}{\sqrt{m\cdot p(1-p)}} < b) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{1}{2}x^2}dx.$$

**PRIMJER 10.6** *Neka je X binomna slučajna varijabla* $X \sim B(m,p)$, $m = 3200,\ p = \frac{1}{2}$. *Izračunajte vjerojatnost da slučajna varijabla poprimi vrijednosti u intervalu* $(1550, 1650)$.

**Rješenje:**
Trebamo izračunati $P(1550 < X < 1650)$.
Prema integralnom Moivre-Laplaceovom teoremu $\frac{X-m\cdot p}{\sqrt{mp(1-p)}} \sim N(0,1)$, $m \to \infty$,
i vrijedi aproksimacija: $P(a < X < b) \approx F^*(\frac{b-mp}{\sqrt{mp(1-p)}}) - F^*(\frac{a-mp}{\sqrt{mp(1-p)}})$.

$$\begin{aligned} P(1550 < X < 1650) &\approx F^*(\frac{1650-3200\cdot 0.5}{\sqrt{3200\cdot 0.25}}) - F^*(\frac{1550-3200\cdot 0.5}{\sqrt{3200\cdot 0.25}}) \\ &= F^*(1.767) - F^*(-1.767) = 2F^*(1.767) - 1 \\ &= 2\cdot 0.961 - 1 = 0.922. \end{aligned}$$

**PRIMJER 10.7** *Vjerojatnost da novorođenče bude muško ili žensko je 1/2. Kolika je vjerojatnost da među 1000 novorođenčadi bude barem 490 muških?*

**Rješenje:** $X \sim B(m,p)$, $m = 1000$, $p = \frac{1}{2}$
Trebamo izračunati $P(X \geq 490) = 1 - P(X \leq 490)$.
Prema integralnom Moivre-Laplaceovom teoremu $\frac{X-m\cdot p}{\sqrt{mp(1-p)}} \sim N(0,1)$, $m \to \infty$,
i vrijedi aproksimacija: $P(a < X < b) \approx F^*(\frac{b-mp}{\sqrt{mp(1-p)}}) - F^*(\frac{a-mp}{\sqrt{mp(1-p)}})$.

$$\begin{aligned} P(X \leq 490) &\approx F^*(\frac{490-1000\cdot 0.5}{\sqrt{1000\cdot 0.25}}) = F^*(-\frac{10}{\sqrt{250}}) = 1 - F^*(\frac{10}{\sqrt{250}}) \\ &= 1 - F^*(0.63) = 1 - 0.7357 = 0.2643. \end{aligned}$$

$$P(X \geq 490) = 1 - P(X \leq 490) \approx 1 - 0.2643 = 0.7357.$$

**TEOREM 10.9** *(integralni MOIVRE - LAPLACEOV TEOREM za rel. frekv. binomne sl. varijable) CGT za binomnu =Bernoullijev slabi ZVB za rel. frekvencije binomne*

*Neka je u Bernoullijevoj shemi slučajna varijabla X=broj uspjeha događaja A u m nezavisnih ponavljanja, $P(A) = p$. $X \sim B(m,p)$. Slučajna varijabla $Y = \frac{X}{m}$ zove se relativna frekvencija uspjeha događaja A u Bernoullijevoj shemi.*
*Tada vrijedi*

$$\lim_{m\to\infty} P(|\frac{X}{m} - p| < \varepsilon) = \lim_{m\to\infty} 2F^*(\varepsilon\sqrt{\frac{m}{p(1-p)}}) - 1 = 1.$$

*Za velike m vrijedi*

$$P(|\frac{X}{m} - p| < \varepsilon) \approx 2F^*(\varepsilon\sqrt{\frac{m}{p(1-p)}}) - 1.$$

**Dokaz:** $P(|\frac{X}{m}-p|<\varepsilon)=P(-\varepsilon\sqrt{\frac{m}{p(1-p)}}<\frac{X-mp}{\sqrt{mp(1-p)}}<\varepsilon\sqrt{\frac{m}{p(1-p)}})$

Koristimo Integralni Moivre-Laplaceov teorem; za $X\sim B(m,p)$ je
$P(a<\frac{X-m\cdot p}{\sqrt{mp(1-p)}}<b)\approx F^*(b)-F^*(a)$,

$$P(-\varepsilon\sqrt{\frac{m}{p(1-p)}}<\frac{X-mp}{\sqrt{mp(1-p)}}<\varepsilon\sqrt{\frac{m}{p(1-p)}})\approx F^*(b)-F^*(a),$$

gdje je $b=-a=\varepsilon\sqrt{\frac{m}{p(1-p)}}$.
Koristimo svojstvo $F^*(x)=1-F^*(-x)$

$$P(-\varepsilon\sqrt{\frac{m}{p(1-p)}}<\frac{X-mp}{\sqrt{mp(1-p)}}<\varepsilon\sqrt{\frac{m}{p(1-p)}})\approx 2F^*(b)-1,$$

i dobivamo željenu tvrdnju

$$P(|\frac{X}{m}-p|<\varepsilon)\approx 2F^*(\varepsilon\sqrt{\frac{m}{p(1-p)}})-1.$$

**PRIMJER 10.8** *Kolika je vjerojatnost da se prilikom bacanja simetričnog novčića $m=3600$ puta relativna frekvencija pojavljivanja pisma po apsolutnoj vrijednosti razlikuje od $p=1/2$ za $\varepsilon=0.01$?*

**Rješenje:**
X=broj pojavljivanja pisma u Bernoullijevoj shemi bacanja novčića
$X\sim B(m,p)$, $m=3600$, $p=\frac{1}{2}$.
Trebamo izračunati $P(|\frac{X}{3600}-\frac{1}{2}|<0.01)$.
Primijenit ćemo integralni Moivre-Laplaceov teorem za rel. frekv. binomne sl. varijable u obliku $P(|\frac{X}{m}-p|<\varepsilon)\approx 2F^*(\varepsilon\sqrt{\frac{m}{p(1-p)}})-1$.

$$\begin{aligned}P(|\frac{X}{3600}-\frac{1}{2}|<0.01) &\approx 2F^*(\varepsilon\sqrt{\frac{m}{p(1-p)}})-1=2F^*(0.01\sqrt{\frac{3600}{\frac{1}{2}\cdot\frac{1}{2}}})-1\\ &= 2F^*(1.2)-1=2\cdot 0.8849-1=0.7698.\end{aligned}$$

**NAPOMENA 10.2** *važno*

*Oblik integralnog Moivre-Laplaceovog teorema*
$P(|\frac{X}{m}-p|<\varepsilon)\approx 2F^*(\varepsilon\sqrt{\frac{m}{p(1-p)}})-1$ *često se koristi u zadacima za određivanje*
*(a) minimalnog broja pokusa m i*

*(b) odstupanja $\varepsilon$*
*da bi za zadani $\gamma$, $P(|\frac{X}{m} - p| < \varepsilon) \geq \gamma$.*

*Rješenje:*
*(a) $m \geq \frac{(z_{\frac{1+\gamma}{2}})^2}{\varepsilon^2} \cdot p(1-p)$, $F^*(z_{\frac{1+\gamma}{2}}) = \frac{1+\gamma}{2}$,*
*(b) $\varepsilon^2 \geq \frac{(z_{\frac{1+\gamma}{2}})^2}{m} \cdot p(1-p)$.*
*Ako je p nepoznato onda se procjenjuje da je $p(1-p) \leq \frac{1}{4}$:*
*(a) $m \geq \frac{(z_{\frac{1+\gamma}{2}})^2}{4\varepsilon^2}$, $\quad F^*(z_{\frac{1+\gamma}{2}}) = \frac{1+\gamma}{2}$,*
*(b) $\varepsilon^2 \geq \frac{(z_{\frac{1+\gamma}{2}})^2}{4m}$.*

**PRIMJER 10.9** *U Bernoullijevoj shemi vjerojatnost događaja A je $p = P(A) = 1/3$. Odredite minimalan broj ponavljanja tako da s vjerojatnošću ne manjom od $\gamma = 0.99$ apsolutno ostupanje relativne frekvencije od p bude najviše $\varepsilon = 0.01$.*

**Rješenje:**
Trebamo odrediti $m$ tako da $P(|\frac{X}{m} - p| < \varepsilon) \geq 0.99$.
Koristimo Integralni Moivre-Laplaceov teorem za relativnu frekvenciju $\frac{X}{m}$ u Bernoullijevoj shemi.
$P(|\frac{X}{m} - p| < \varepsilon) \approx 2F^*(\varepsilon\sqrt{\frac{m}{p(1-p)}}) - 1$,

$$P(|\frac{X}{m} - \frac{1}{3}| < 0.01) \approx 2F^*(0.01\sqrt{\frac{m}{\frac{2}{9}}}) - 1.$$

Broj ponavljanja ćemo odrediti iz zadane vjerojatnosti $\gamma$ i uvjeta
$2F^*(0.01\sqrt{\frac{m}{\frac{2}{9}}}) - 1 \geq 0.99 \quad \Rightarrow \quad F^*(0.01\sqrt{\frac{m}{\frac{2}{9}}}) \geq 0.995$,
$F^*(z) = 0.995 \quad \Rightarrow \quad z = 2.6 \quad \Rightarrow \quad 0.01\sqrt{\frac{m}{\frac{2}{9}}} \geq 2.6$,
$m \geq \frac{(z_{\frac{1+\gamma}{2}})^2}{\varepsilon^2} \cdot p(1-p) = \frac{2.6^2}{0.01^2} \cdot \frac{2}{9} \quad \Rightarrow \quad m \geq 15022$.
(Bernoullijev SZVB za rel. frekv. dao je ocjenu $m \geq 222222$.)

**PRIMJER 10.10** *Koliko puta treba baciti simetričnu kocku da bi relativna frekvencija pojavljivanja broja 6 bila između $\frac{19}{120}$ i $\frac{21}{120}$ s vjerojatnošću $\gamma = 0.95$.*

**Rješenje:**
X=broj pojavljivanja 6 u Bernoullijevoj shemi bacanja kocke
$X \sim B(m,p), m, \ p = \frac{1}{6}$.
Trebamo odrediti $m$ tako da $P(\frac{19}{120} < \frac{X}{m} < \frac{21}{120}) \geq 0.95$.

$$\begin{aligned} P(\frac{19}{120} < \frac{X}{m} < \frac{21}{120}) &= P(\frac{19}{120} - \frac{1}{6} < \frac{X}{m} - \frac{1}{6} < \frac{21}{120} - \frac{1}{6}) \\ &= P(-\frac{1}{120} < \frac{X}{m} - \frac{1}{6} < \frac{1}{120}) = P(|\frac{X}{m} - \frac{1}{6}| < \frac{1}{120}). \end{aligned}$$

Koristimo Integralni Moivre-Laplaceov teorem za relativnu frekvenciju $\frac{X}{m}$ u Bernoullijevoj shemi.
$P(|\frac{X}{m} - p| < \varepsilon) \approx 2F^*(\varepsilon \sqrt{\frac{m}{p(1-p)}}) - 1$,

$$P(|\frac{X}{m} - \frac{1}{6}| < \frac{1}{120}) \approx 2F^*(\frac{1}{120} \sqrt{\frac{m}{\frac{5}{36}}}) - 1$$

Broj ponavljanja ćemo odrediti iz zadane vjerojatnosti $\gamma$ i uvjeta
$2F^*(\frac{1}{120} \sqrt{\frac{m}{\frac{5}{36}}}) - 1 \geq 0.95 \quad \Rightarrow \quad F^*(\frac{1}{120} \sqrt{\frac{m}{\frac{5}{36}}}) \geq 0.975$
$F^*(z) = 0.975 \quad \Rightarrow \quad z = 1.96 \quad \Rightarrow \quad \frac{1}{120} \sqrt{\frac{m}{\frac{5}{36}}} \geq 1.96,$
$m \geq \frac{(z_{\frac{1+\gamma}{2}})^2}{\varepsilon^2} \cdot p(1-p) = \frac{1.96^2}{(\frac{1}{120})^2} \cdot \frac{5}{36} = 7684 \quad \Rightarrow \quad m \geq 7684.$

**PRIMJER 10.11** *Simetričnu kocku bacamo $m = 4500$ puta. U kojim granicama s vjerojatnošću $\gamma = 0.9$ treba očekivati relativne frekvencije pojavljivanja boja 6?*

**Rješenje:**
$X$ =broj pojavljivanja 6 u Bernoullijevoj shemi bacanja kocke,
$X \sim B(m,p), m, p = \frac{1}{6}$.
Treba oderditi a i b takve da za $X$ =broj 6 u $m$ bacanja $P(a < \frac{X}{m} < b) \geq 0.9$.
Koristimo Integralni Moivre-Laplaceov teorem za relativnu frekvenciju $\frac{X}{m}$
u Bernoullijevoj shemi $P(|\frac{X}{m} - p| < \varepsilon) \approx 2F^*(\varepsilon \sqrt{\frac{m}{p(1-p)}}) - 1$.

$$P(|\frac{X}{4500} - \frac{1}{6}| < \varepsilon) \approx 2F^*(\varepsilon \sqrt{\frac{4500}{\frac{1}{6}(1 - \frac{1}{6})}}) - 1$$

Odstupanje $\varepsilon$ ćemo odrediti iz zadane vjerojatnosti $\gamma$ i uvjeta
$2F^*(\varepsilon \sqrt{\frac{4500}{\frac{5}{36}}}) - 1 \geq 0.9 \quad \Rightarrow \quad F^*(\varepsilon \sqrt{36 \cdot 900}) \geq 0.95.$
$F^*(z) = 0.9 \Rightarrow z = 1.65 \quad \Rightarrow \quad \varepsilon \sqrt{36 \cdot 900} \geq 1.65 \Rightarrow \varepsilon \geq 0.0091$

Prema formuli $\varepsilon^2 \geq \frac{(z_{\frac{1+\gamma}{2}})^2}{m} \cdot p(1-p) = \frac{1.65^2}{4500} \cdot \frac{5}{36} = 7.566 \times 10^{-5}$

$$\varepsilon \geq \frac{1.65}{30} \cdot \frac{1}{6} = 9.1667 \times 10^{-3} \quad \Rightarrow \quad \varepsilon \geq 9.1667 \times 10^{-3}.$$

Odredili smo $\varepsilon$ tako da $P(|\frac{X}{4500} - \frac{1}{6}| < 0.00916) \geq 0.9$.

$$|\frac{X}{4500} - \frac{1}{6}| < 0.00916 \quad \Rightarrow \quad \frac{1}{6} - 0.00916 < \frac{X}{4500} < \frac{1}{6} + 0.00916$$

$$\Rightarrow \quad 0.15751 < \frac{X}{4500} < 0.17583.$$

Prisjetimo se motivacijskog primjera:

**PRIMJER 10.12** *motiv*

*Kontrolor uzima uzorak veličine* $m = 1000$ *iz skupa uređaja. Vjerojatnost da je uređaj neispravan je* $p = 0.03$. *U kojim granicama će biti broj neispravnih uređaja u uzorku s vjerojatnošću* $\gamma = 0.99$

**Rješenje:**

Treba oderditi a i b takve da za $X$ =broj neispravnih uređaja u uzorku veličine $m$, $X \sim B(m, p)$, $P(a < X < b) \geq 0.99$.

Koristimo Integralni Moivre-Laplaceov teorem za relativnu frekvenciju $\frac{X}{m}$ u Bernoullijevoj shemi $P(|\frac{X}{m} - p| < \varepsilon) \approx 2F^*(\varepsilon\sqrt{\frac{m}{p(1-p)}}) - 1$,

$$P(|\frac{X}{1000} - 0.03| < \varepsilon) \approx 2F^*(\varepsilon\sqrt{\frac{1000}{0.03(1-0.03)}}) - 1.$$

Odstupanje $\varepsilon$ ćemo odrediti iz zadane vjerojatnosti $\gamma$ i uvjeta

$2F^*(\varepsilon\sqrt{\frac{1000}{0.03(1-0.003)}}) - 1 \geq 0.99 \quad \Rightarrow \quad F^*(\varepsilon\sqrt{\frac{1000}{0.03(1-0.003)}}) \geq 0.995$.

$F^*(z) = 0.995 \;\Rightarrow\; z = 2.6 \quad \Rightarrow \quad \varepsilon\sqrt{\frac{1000}{0.03(1-0.003)}} \geq 2.6$.

Prema formuli $\varepsilon^2 \geq \frac{(z_{\frac{1+\gamma}{2}})^2}{m} \cdot p(1-p) = \frac{2.6^2}{1000} \cdot 0.03 \cdot (1-0.03) = 1.9672 \times 10^{-4}$

$$\varepsilon \geq 0.014$$

Odredili smo $\varepsilon$ tako da $P(|\frac{X}{1000} - 0.03| < 0.014) \geq 0.99$.

$$|\frac{X}{1000} - 0.03| < 0.014 \quad \Rightarrow \quad 16 < X < 44.$$

(Bernoullijev SZVB za rel. frekv. dao je ocjenu $0 < X < 80$.)

## 10.4 Ponovimo

ČEBIŠEVLJEVA NEJEDNAKOST

| | |
|---|---|
| za svaki $\varepsilon > 0$ | $P(\lvert X - E(X)\rvert \geq \varepsilon) \leq \frac{Var(X)}{\varepsilon^2}$, |
| za svaki $\lambda$ | $P(\lvert X - \mu\rvert \geq \lambda\sigma) \leq \frac{1}{\lambda^2}$, |

ZAKONI VELIKIH BROJEVA (ZVB)

| | |
|---|---|
| ZVB za $\overline{X}$ | $\lim_{n\to\infty} P(\lvert\overline{X} - \mu\rvert < \varepsilon) = 1$, |
| ZVB za $X \sim B(m,p)$ | $\lim_{n\to\infty} P(\lvert\frac{X}{m} - p\rvert < \varepsilon) = 1$, |
| Bernoull. slabi ZVB | $P(\lvert\frac{X}{m} - p\rvert < \varepsilon) \geq 1 - \frac{p(1-p)}{m\cdot\varepsilon^2}$ |

CENTRALNI GRANIČNI TEOREMI (CGT)

| | |
|---|---|
| CGT za niz nez. sl, var. $X_i$ | $n \to \infty$, |
| | $\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\cdot\sigma} \sim N(0,1)$ |
| CTG za $\overline{X}$ | $n \to \infty$, |
| | $\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$ |
| CTG za $X \sim B(m,p)$, $m \to \infty$ | = Moivre-Laplace |
| $\frac{X-m\cdot p}{\sqrt{mp(1-p)}} \sim N(0,1)$, | $P(a < X < b) \approx F^*(\frac{b-mp}{\sqrt{mp(1-p)}}) - F^*(\frac{a-mp}{\sqrt{mp(1-p)}})$ |
| CGT za rel. frekv. od $X \sim B(m,p)$ | = Moivre-Laplace za rel. frekv. |
| $\frac{\frac{X}{m}-p}{\sqrt{\frac{p(1-p)}{m}}} \sim N(0,1)$, | $P(\lvert\frac{X}{m} - p\rvert < \varepsilon) \approx 2F^*(\varepsilon\sqrt{\frac{m}{p(1-p)}}) - 1$ |