

6. Granične teoreme i njihove primene

Profesor Milan Merkle
emerkle@etf.rs milanmerkle.etf.rs

Matematička statistika- jesen 2019

Konvergenција slučajnih promenljivih

U primenama matematike konvergenција se koristi za aproksimaciju. Na primer, iz

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

sledi da je

$$(1 + 3/100)^{100} \approx e^3 \quad \text{ili} \quad e^3 \approx (1 + 3/100)^{100}$$

Kod slučajnih promenljivih imamo komplikovaniju situaciju zbog toga što bliskost slučajnih promenljivih može da se posmatra u smislu bliskosti njihovih vrednosti ili u smislu bliskosti njihovih raspodela.

Konvergencija slučajnih promenljivih

U primenama matematike konvergencija se koristi za aproksimaciju. Na primer, iz

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

sledi da je

$$(1 + 3/100)^{100} \approx e^3 \quad \text{ili} \quad e^3 \approx (1 + 3/100)^{100}$$

Kod slučajnih promenljivih imamo komplikovaniju situaciju zbog toga što bliskost slučajnih promenljivih može da se posmatra u smislu bliskosti njihovih vrednosti ili u smislu bliskosti njihovih raspodela.

Konvergenција u strogom smislu

- *Konvergenција skoro svuda:*

$$P\left(\lim_{n \rightarrow +\infty} X_n = X\right) = 1,$$

gde je limes definisan kao u matematici.

U primenama se češće koriste alternativni koncepti bliskosti slučajnih promenljivih.

Konvergenција u strogom smislu

- *Konvergenција skoro svuda:*

$$P\left(\lim_{n \rightarrow +\infty} X_n = X\right) = 1,$$

gde je limes definisan kao u matematici.

U primenama se češće koriste alternativni koncepti bliskosti slučajnih promenljivih.

Konvergenција u raspodeli

$X_n \sim F_n, n = 1, 2, \dots; X \sim F.$

Definicija 6.1. Kažemo da niz slučajnih promenljivih X_n konvergira u raspodeli ka X ako važi:

$$\lim_{n \rightarrow +\infty} F_n(x) = F(x) \quad (F_n = P(X_n \leq x), F(x) = P(X \leq x))$$

u svakoj tački $x \in \mathbb{R}$ u kojoj je funkcija F neprekidna.

Teorema 6.2 (teorema o neprekidnosti, Lévy continuity theorem). Niz slučajnih promenljivih X_n konvergira u raspodeli ka X ako i samo ako je

$$\lim_{n \rightarrow +\infty} \varphi_n(t) = \varphi(t) \quad \text{za svako } t \in \mathbb{R}. \quad \checkmark$$

Konvergenција u raspodeli

$X_n \sim F_n, n = 1, 2, \dots; X \sim F.$

Definicija 6.1. Kažemo da niz slučajnih promenljivih X_n konvergira u raspodeli ka X ako važi:

$$\lim_{n \rightarrow +\infty} F_n(x) = F(x) \quad (F_n = P(X_n \leq x), F(x) = P(X \leq x))$$

u svakoj tački $x \in \mathbb{R}$ u kojoj je funkcija F neprekidna.

Teorema 6.2 (teorema o neprekidnosti, Lévy continuity theorem). Niz slučajnih promenljivih X_n konvergira u raspodeli ka X ako i samo ako je

$$\lim_{n \rightarrow +\infty} \varphi_n(t) = \varphi(t) \quad \text{za svako } t \in \mathbb{R}. \quad \checkmark$$

Apksimacija binomne raspodele Puasonovom ✓

Primer 121. Za $X_n \sim \text{Bin}(n, p_n)$,

$$\varphi_n(t) = \sum_{k=0}^n \binom{n}{k} e^{itk} p_n^k (1-p_n)^{n-k} = (1 + p_n (e^{it} - 1))^n$$

Ako $n \rightarrow \infty$ i $np_n \rightarrow \lambda > 0$:

$$\lim_{n \rightarrow +\infty} \varphi_n(t) = \exp(\lambda (e^{it} - 1)) = \varphi_X, \quad X \sim \text{Pois}(\lambda)$$

Teorema o neprekidnosti: $\lim_{n \rightarrow +\infty} P(X_n = k) = e^{-\lambda} \frac{\lambda^k}{k!}$

Empirijsko pravilo: Za veliko n i $np \leq 5$, $X \sim \text{Bin}(n, p)$:

$$P(X = k) \approx e^{-np} \frac{(np)^k}{k!} \quad \checkmark$$

Apksimacija binomne raspodele Puasonovom ✓

Primer 121. Za $X_n \sim \text{Bin}(n, p_n)$,

$$\varphi_n(t) = \sum_{k=0}^n \binom{n}{k} e^{itk} p_n^k (1-p_n)^{n-k} = (1 + p_n (e^{it} - 1))^n$$

Ako $n \rightarrow \infty$ i $np_n \rightarrow \lambda > 0$:

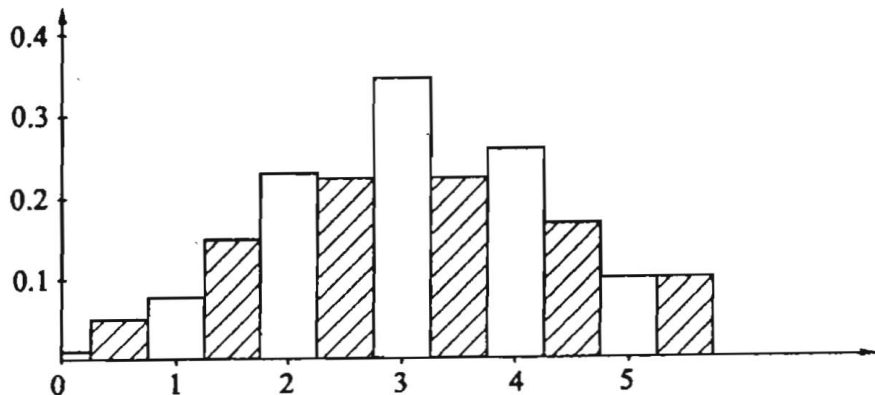
$$\lim_{n \rightarrow +\infty} \varphi_n(t) = \exp(\lambda (e^{it} - 1)) = \varphi_X, \quad X \sim \text{Pois}(\lambda)$$

Teorema o neprekidnosti: $\lim_{n \rightarrow +\infty} P(X_n = k) = e^{-\lambda} \frac{\lambda^k}{k!}$

Empirijsko pravilo: Za veliko n i $np \leq 5$, $X \sim \text{Bin}(n, p)$:

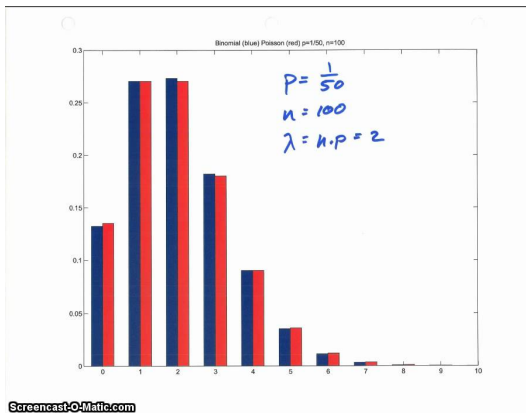
$$P(X = k) \approx e^{-np} \frac{(np)^k}{k!} \quad \checkmark$$

Primer



Slika 28. Binomna raspodela ($n = 5, p = 3/5$) i njena Poissonova aproksimacija (osenčeno).

Primer



Slika 117. Histogrami binomne raspodele (plavo) i Puasonove aproksimacije za $n = 100$ i $np = 2$

Limes u verovatnoći i u srednjem kvadratnom smislu

Niz slučajnih promenljivih X_n

- **Konvergencija u verovatnoći**:-bliskost vrednosti X_n i X za veliko n .

$$\lim_{n \rightarrow +\infty} P(|X_n - X| \geq \varepsilon) = 0 \quad \text{za svako } \varepsilon > 0.$$

- **Srednja kvadratna konvergencija**: -bliskost vrednosti X_n i X za veliko n .

$$\lim_{n \rightarrow +\infty} E(X_n - X)^2 = 0$$

Primer: $E(\hat{\mu}_n - \mu)^2 = \text{Var}(\hat{\mu}_n) \rightarrow 0$ (model merenja).

Srednja kvadratna \implies u verovatnoći \implies u raspodeli

*Konvergencija se koristi za aproksimacije raspodela i njihovih momenata. Teoreme koje se odnose na konvergenciju (**granične teoreme**) utvrđuju uslove pod kojima se aproksimacija može primeniti.*

Limes u verovatnoći i u srednjem kvadratnom smislu

Niz slučajnih promenljivih X_n

- **Konvergencija u verovatnoći**:-bliskost vrednosti X_n i X za veliko n .

$$\lim_{n \rightarrow +\infty} P(|X_n - X| \geq \varepsilon) = 0 \quad \text{za svako } \varepsilon > 0.$$

- **Srednja kvadratna konvergencija**: -bliskost vrednosti X_n i X za veliko n .

$$\lim_{n \rightarrow +\infty} E(X_n - X)^2 = 0$$

Primer: $E(\hat{\mu}_n - \mu)^2 = \text{Var}(\hat{\mu}_n) \rightarrow 0$ (model merenja).

Srednja kvadratna \implies u verovatnoći \implies u raspodeli

*Konvergencija se koristi za aproksimacije raspodela i njihovih momenata. Teoreme koje se odnose na konvergenciju (**granične teoreme**) utvrđuju uslove pod kojima se aproksimacija može primeniti.*

Limes u verovatnoći i u srednjem kvadratnom smislu

Niz slučajnih promenljivih X_n

- **Konvergencija u verovatnoći**:-bliskost vrednosti X_n i X za veliko n .

$$\lim_{n \rightarrow +\infty} P(|X_n - X| \geq \varepsilon) = 0 \quad \text{za svako } \varepsilon > 0.$$

- **Srednja kvadratna konvergencija**: -bliskost vrednosti X_n i X za veliko n .

$$\lim_{n \rightarrow +\infty} E(X_n - X)^2 = 0$$

Primer: $E(\hat{\mu}_n - \mu)^2 = \text{Var}(\hat{\mu}_n) \rightarrow 0$ (model merenja).

Srednja kvadratna \implies u verovatnoći \implies u raspodeli

*Konvergencija se koristi za aproksimacije raspodela i njihovih momenata. Teoreme koje se odnose na konvergenciju (**granične teoreme**) utvrđuju uslove pod kojima se aproksimacija može primeniti.*

Limes u verovatnoći i u srednjem kvadratnom smislu

Niz slučajnih promenljivih X_n

- **Konvergencija u verovatnoći**:-bliskost vrednosti X_n i X za veliko n .

$$\lim_{n \rightarrow +\infty} P(|X_n - X| \geq \varepsilon) = 0 \quad \text{za svako } \varepsilon > 0.$$

- **Srednja kvadratna konvergencija**: -bliskost vrednosti X_n i X za veliko n .

$$\lim_{n \rightarrow +\infty} E(X_n - X)^2 = 0$$

Primer: $E(\hat{\mu}_n - \mu)^2 = \text{Var}(\hat{\mu}_n) \rightarrow 0$ (model merenja).

Srednja kvadratna \implies u verovatnoći \implies u raspodeli

*Konvergencija se koristi za aproksimacije raspodela i njihovih momenata. Teoreme koje se odnose na konvergenciju (**granične teoreme**) utvrđuju uslove pod kojima se aproksimacija može primeniti.*

Nejednakosti Markova i Čebiševa

Teorema 6.3 (Markov) *Ako je $X \geq 0$ i postoji $E X$:*

$$P(X \geq \varepsilon) \leq \frac{E(X)}{\varepsilon} \quad \text{za svako } \varepsilon > 0.$$

Teorema 6.4 (Čebišev) *Za svaku slučajnu promenljivu X koja ima varijansu,*

$$P(|X - E X| \geq \varepsilon) \leq \frac{\text{Var } X}{\varepsilon^2} \quad \text{za svako } \varepsilon > 0.$$

Univerzalno pravilo tri sigme

Za $\varepsilon = 3\sqrt{\text{Var}X} = 3\sigma$:

$$P(|X - E X| \leq 3\sigma) \geq \frac{8}{9} = 0.89$$

Važi za svaku raspodelu koja ima varijansu.

Nejednakosti Markova i Čebiševa

Teorema 6.3 (Markov) *Ako je $X \geq 0$ i postoji $E X$:*

$$P(X \geq \varepsilon) \leq \frac{E(X)}{\varepsilon} \quad \text{za svako } \varepsilon > 0.$$

Teorema 6.4 (Čebišev) *Za svaku slučajnu promenljivu X koja ima varijansu,*

$$P(|X - E X| \geq \varepsilon) \leq \frac{\text{Var } X}{\varepsilon^2} \quad \text{za svako } \varepsilon > 0.$$

Univerzalno pravilo tri sigme

Za $\varepsilon = 3\sqrt{\text{Var}X} = 3\sigma$:

$$P(|X - E X| \leq 3\sigma) \geq \frac{8}{9} = 0.89$$

Važi za svaku raspodelu koja ima varijansu.

Nejednakosti Markova i Čebiševa

Teorema 6.3 (Markov) *Ako je $X \geq 0$ i postoji $E X$:*

$$P(X \geq \varepsilon) \leq \frac{E(X)}{\varepsilon} \quad \text{za svako } \varepsilon > 0.$$

Teorema 6.4 (Čebišev) *Za svaku slučajnu promenljivu X koja ima varijansu,*

$$P(|X - E X| \geq \varepsilon) \leq \frac{\text{Var } X}{\varepsilon^2} \quad \text{za svako } \varepsilon > 0.$$

Univerzalno pravilo tri sigme

Za $\varepsilon = 3\sqrt{\text{Var}X} = 3\sigma$:

$$P(|X - E X| \leq 3\sigma) \geq \frac{8}{9} = 0.89$$

Važi za svaku raspodelu koja ima varijansu.

Zakon velikih brojeva

Teorema 6.5 (Slabi zakon velikih brojeva) *Neka su X_1, X_2, \dots nezavisne slučajne promenljive sa $E X_k = \mu$ i neka postoji konstanta V takva da je $\text{Var } X_k \leq V$ za svako k . Tada niz aritmetičkih sredina $\hat{\mu}_n = (X_1 + \dots + X_n)/n$ konvergira u verovatnoći ka μ , tj.*

$$\lim_{n \rightarrow +\infty} P(|\hat{\mu}_n - \mu| \geq \varepsilon) = 0 \quad \text{za svako } \varepsilon > 0. \quad \checkmark$$

Dokaz preko nejednakosti Čebiševa.

- $\hat{\mu}_n$ je ocena matematičkog očekivanja μ
- Obratite pažnju da X_i mogu imati različite raspodele.

Zakon velikih brojeva

Teorema 6.5 (Slabi zakon velikih brojeva) *Neka su X_1, X_2, \dots nezavisne slučajne promenljive sa $E X_k = \mu$ i neka postoji konstanta V takva da je $\text{Var } X_k \leq V$ za svako k . Tada niz aritmetičkih sredina $\hat{\mu}_n = (X_1 + \dots + X_n)/n$ konvergira u verovatnoći ka μ , tj.*

$$\lim_{n \rightarrow +\infty} P(|\hat{\mu}_n - \mu| \geq \varepsilon) = 0 \quad \text{za svako } \varepsilon > 0. \quad \checkmark$$

Dokaz preko nejednakosti Čebiševa.

- $\hat{\mu}_n$ je *ocena matematičkog očekivanja* μ
- *Obratite pažnju da X_i mogu imati različite raspodele.*

Zakon velikih brojeva

Teorema 6.5 (Slabi zakon velikih brojeva) *Neka su X_1, X_2, \dots nezavisne slučajne promenljive sa $E X_k = \mu$ i neka postoji konstanta V takva da je $\text{Var } X_k \leq V$ za svako k . Tada niz aritmetičkih sredina $\hat{\mu}_n = (X_1 + \dots + X_n)/n$ konvergira u verovatnoći ka μ , tj.*

$$\lim_{n \rightarrow +\infty} P(|\hat{\mu}_n - \mu| \geq \varepsilon) = 0 \quad \text{za svako } \varepsilon > 0. \quad \checkmark$$

Dokaz preko nejednakosti Čebiševa.

- $\hat{\mu}_n$ je *ocena matematičkog očekivanja μ*
- *Obratite pažnju da X_i mogu imati različite raspodele.*

Zakon velikih brojeva

Teorema 6.5 (Slabi zakon velikih brojeva) *Neka su X_1, X_2, \dots nezavisne slučajne promenljive sa $E X_k = \mu$ i neka postoji konstanta V takva da je $\text{Var } X_k \leq V$ za svako k . Tada niz aritmetičkih sredina $\hat{\mu}_n = (X_1 + \dots + X_n)/n$ konvergira u verovatnoći ka μ , tj.*

$$\lim_{n \rightarrow +\infty} P(|\hat{\mu}_n - \mu| \geq \varepsilon) = 0 \quad \text{za svako } \varepsilon > 0. \quad \checkmark$$

Dokaz preko nejednakosti Čebiševa.

- $\hat{\mu}_n$ je *ocena matematičkog očekivanja μ*
- *Obratite pažnju da X_i mogu imati različite raspodele.*

Konvergenција skoro svuda?

Teorema 6.8-Strogi zakon velikih brojeva Kolmogorova.

Neka je X_1, X_2, \dots, X_n nezavisni uzorak iz raspodele slučajne promenljive X i $\hat{\mu}_n = (X_1 + \dots + X_n)/n, n = 1, 2, \dots$.

1 a) $E X = \mu$ ako i samo ako

$$\lim_{n \rightarrow +\infty} \hat{\mu}_n = \mu \text{ skoro svuda.}$$

2 b) $E X$ ne postoji ako i samo ako $\hat{\mu}_n$ divergira skoro svuda.

U primenama:

- 1 Posmatramo niz $\hat{\mu}_n$ za rastuće vrednosti n
- 2 Ako iz niza $\hat{\mu}_n$ može zaključiti da postoji $\lim \hat{\mu}_n$, onda $\mu \approx \hat{\mu}_n$ za dovoljno veliko n
- 3 Ako postoje indikacije da niz ne konvergira, onda zaključujemo da raspodela nema matematičko očekivanje.

Konvergenција skoro svuda?

Teorema 6.8-Strogi zakon velikih brojeva Kolmogorova.

Neka je X_1, X_2, \dots, X_n nezavisni uzorak iz raspodele slučajne promenljive X i $\hat{\mu}_n = (X_1 + \dots + X_n)/n$, $n = 1, 2, \dots$.

1 a) $E X = \mu$ ako i samo ako

$$\lim_{n \rightarrow +\infty} \hat{\mu}_n = \mu \quad \text{skoro svuda.}$$

2 b) $E X$ ne postoji ako i samo ako $\hat{\mu}_n$ divergira skoro svuda.

U primenama:

- 1 Posmatramo niz $\hat{\mu}_n$ za rastuće vrednosti n
- 2 Ako iz niza $\hat{\mu}_n$ može zaključiti da postoji $\lim \hat{\mu}_n$, onda $\mu \approx \hat{\mu}_n$ za dovoljno veliko n
- 3 Ako postoje indikacije da niz ne konvergira, onda zaključujemo da raspodela nema matematičko očekivanje.

Konvergenција skoro svuda?

Teorema 6.8-Strogi zakon velikih brojeva Kolmogorova.

Neka je X_1, X_2, \dots, X_n nezavisni uzorak iz raspodele slučajne promenljive X i $\hat{\mu}_n = (X_1 + \dots + X_n)/n$, $n = 1, 2, \dots$.

1 a) $E X = \mu$ ako i samo ako

$$\lim_{n \rightarrow +\infty} \hat{\mu}_n = \mu \quad \text{skoro svuda.}$$

2 b) $E X$ ne postoji ako i samo ako $\hat{\mu}_n$ divergira skoro svuda.

U primenama:

- 1 Posmatramo niz $\hat{\mu}_n$ za rastuće vrednosti n
- 2 Ako iz niza $\hat{\mu}_n$ može zaključiti da postoji $\lim \hat{\mu}_n$, onda $\mu \approx \hat{\mu}_n$ za dovoljno veliko n
- 3 Ako postoje indikacije da niz ne konvergira, onda zaključujemo da raspodela nema matematičko očekivanje.

Konvergenција skoro svuda?

Teorema 6.8-Strogi zakon velikih brojeva Kolmogorova.

Neka je X_1, X_2, \dots, X_n nezavisni uzorak iz raspodele slučajne promenljive X i $\hat{\mu}_n = (X_1 + \dots + X_n)/n$, $n = 1, 2, \dots$.

1 a) $E X = \mu$ ako i samo ako

$$\lim_{n \rightarrow +\infty} \hat{\mu}_n = \mu \quad \text{skoro svuda.}$$

2 b) $E X$ ne postoji ako i samo ako $\hat{\mu}_n$ divergira skoro svuda.

U primenama:

- 1 Posmatramo niz $\hat{\mu}_n$ za rastuće vrednosti n
- 2 Ako iz niza $\hat{\mu}_n$ može zaključiti da postoji $\lim \hat{\mu}_n$, onda $\mu \approx \hat{\mu}_n$ za dovoljno veliko n
- 3 Ako postoje indikacije da niz ne konvergira, onda zaključujemo da raspodela nema matematičko očekivanje.

Konvergenција skoro svuda?

Teorema 6.8-Strogi zakon velikih brojeva Kolmogorova.

Neka je X_1, X_2, \dots, X_n nezavisni uzorak iz raspodele slučajne promenljive X i $\hat{\mu}_n = (X_1 + \dots + X_n)/n$, $n = 1, 2, \dots$.

① a) $E X = \mu$ ako i samo ako

$$\lim_{n \rightarrow +\infty} \hat{\mu}_n = \mu \quad \text{skoro svuda.}$$

② b) $E X$ ne postoji ako i samo ako $\hat{\mu}_n$ divergira skoro svuda.

U primenama:

① Posmatramo niz $\hat{\mu}_n$ za rastuće vrednosti n

② Ako iz niza $\hat{\mu}_n$ može zaključiti da postoji $\lim \hat{\mu}_n$, onda $\mu \approx \hat{\mu}_n$ za dovoljno veliko n

③ Ako postoje indikacije da niz ne konvergira, onda zaključujemo da raspodela nema matematičko očekivanje.

Konvergenција skoro svuda?

Teorema 6.8-Strogi zakon velikih brojeva Kolmogorova.

Neka je X_1, X_2, \dots, X_n nezavisni uzorak iz raspodele slučajne promenljive X i $\hat{\mu}_n = (X_1 + \dots + X_n)/n$, $n = 1, 2, \dots$.

① a) $E X = \mu$ ako i samo ako

$$\lim_{n \rightarrow +\infty} \hat{\mu}_n = \mu \quad \text{skoro svuda.}$$

② b) $E X$ ne postoji ako i samo ako $\hat{\mu}_n$ divergira skoro svuda.

U primenama:

① Posmatramo niz $\hat{\mu}_n$ za rastuće vrednosti n

② Ako iz niza $\hat{\mu}_n$ može zaključiti da postoji $\lim \hat{\mu}_n$, onda $\mu \approx \hat{\mu}_n$ za dovoljno veliko n

③ Ako postoje indikacije da niz ne konvergira, onda zaključujemo da raspodela nema matematičko očekivanje.

Konvergenција skoro svuda?

Teorema 6.8-Strogi zakon velikih brojeva Kolmogorova.

Neka je X_1, X_2, \dots, X_n nezavisni uzorak iz raspodele slučajne promenljive X i $\hat{\mu}_n = (X_1 + \dots + X_n)/n$, $n = 1, 2, \dots$.

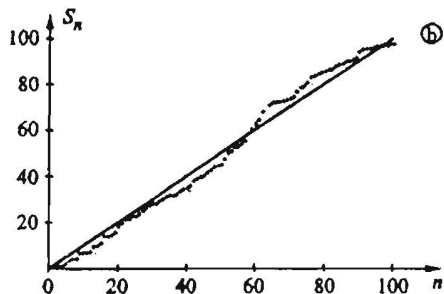
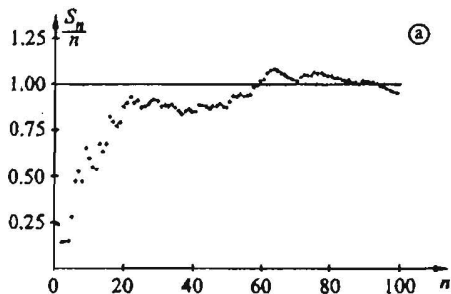
① a) $E X = \mu$ ako i samo ako

$$\lim_{n \rightarrow +\infty} \hat{\mu}_n = \mu \quad \text{skoro svuda.}$$

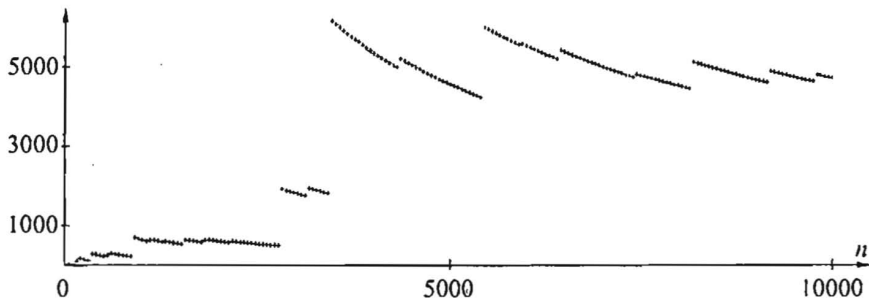
② b) $E X$ ne postoji ako i samo ako $\hat{\mu}_n$ divergira skoro svuda.

U primenama:

- ① Posmatramo niz $\hat{\mu}_n$ za rastuće vrednosti n
- ② Ako iz niza $\hat{\mu}_n$ može zaključiti da postoji $\lim \hat{\mu}_n$, onda $\mu \approx \hat{\mu}_n$ za *dovoljno veliko n*
- ③ Ako postoje indikacije da niz ne konvergira, onda zaključujemo da raspodela nema matematičko očekivanje.



Slika 29. Ponašanje $\hat{\mu}_n$ i S_n za eksponencijalnu raspodelu sa parametrom 1 (računarska simulacija).



Slika 30. Ponašanje ocene drugog momenta CAUCHYjeve raspodele (računarska simulacija).

Statistička definicija verovatnoće preko ZVB

- Radi ocene verovatnoće događaja A sa nepoznatom verovatnoćom $p = P(A)$, eksperiment se ponavlja n puta.
- Neka je $X_i = I_A$ indikator događaja A u i -tom ponavljanju.
- Definišemo $m_n = X_1 + \dots + X_n$ (broj realizacija događaja A u n ponavljanja).
- $\hat{p}_n = \frac{m_n}{n} = \frac{X_1 + \dots + X_n}{n}$ je *ocena nepoznate verovatnoće p*
- ZVB $\implies \hat{p}_n \rightarrow p$ u verovatnoći

Statistička definicija verovatnoće preko ZVB

- Radi ocene verovatnoće događaja A sa nepoznatom verovatnoćom $p = P(A)$, eksperiment se ponavlja n puta.
- Neka je $X_i = I_A$ indikator događaja A u i -tom ponavljanju.
- Definišemo $m_n = X_1 + \dots + X_n$ (broj realizacija događaja A u n ponavljanja).
- $\hat{p}_n = \frac{m_n}{n} = \frac{X_1 + \dots + X_n}{n}$ je *ocena nepoznate verovatnoće p*
- ZVB $\implies \hat{p}_n \rightarrow p$ u verovatnoći

Statistička definicija verovatnoće preko ZVB

- Radi ocene verovatnoće događaja A sa nepoznatom verovatnoćom $p = P(A)$, eksperiment se ponavlja n puta.
- Neka je $X_i = I_A$ indikator događaja A u i -tom ponavljanju.
- Definišemo $m_n = X_1 + \dots + X_n$ (broj realizacija događaja A u n ponavljanja).
- $\hat{p}_n = \frac{m_n}{n} = \frac{X_1 + \dots + X_n}{n}$ je *ocena nepoznate verovatnoće p*
- ZVB $\implies \hat{p}_n \rightarrow p$ u verovatnoći

Statistička definicija verovatnoće preko ZVB

- Radi ocene verovatnoće događaja A sa nepoznatom verovatnoćom $p = P(A)$, eksperiment se ponavlja n puta.
- Neka je $X_i = I_A$ indikator događaja A u i -tom ponavljanju.
- Definišemo $m_n = X_1 + \dots + X_n$ (broj realizacija događaja A u n ponavljanja).
- $\hat{p}_n = \frac{m_n}{n} = \frac{X_1 + \dots + X_n}{n}$ je *ocena nepoznate verovatnoće p*
- ZVB $\implies \hat{p}_n \rightarrow p$ u verovatnoći

Statistička definicija verovatnoće preko ZVB

- Radi ocene verovatnoće događaja A sa nepoznatom verovatnoćom $p = P(A)$, eksperiment se ponavlja n puta.
- Neka je $X_i = I_A$ indikator događaja A u i -tom ponavljanju.
- Definišemo $m_n = X_1 + \dots + X_n$ (broj realizacija događaja A u n ponavljanja).
- $\hat{p}_n = \frac{m_n}{n} = \frac{X_1 + \dots + X_n}{n}$ je **ocena nepoznate verovatnoće p**
- ZVB $\implies \hat{p}_n \rightarrow p$ u verovatnoći

Statistička definicija verovatnoće preko ZVB

- Radi ocene verovatnoće događaja A sa nepoznatom verovatnoćom $p = P(A)$, eksperiment se ponavlja n puta.
- Neka je $X_i = I_A$ indikator događaja A u i -tom ponavljanju.
- Definišemo $m_n = X_1 + \dots + X_n$ (broj realizacija događaja A u n ponavljanja).
- $\hat{p}_n = \frac{m_n}{n} = \frac{X_1 + \dots + X_n}{n}$ je **ocena nepoznate verovatnoće p**
- ZVB $\implies \hat{p}_n \rightarrow p$ u verovatnoći

Koliko veliko n treba da bude?

Iz ZVB sledi da je, za dato $\varepsilon > 0$ i za svako $\alpha \in (0, 1)$ postoji n takvo da je $P(|\hat{\mu}_n - \mu| \geq \varepsilon) \leq \alpha$.

- Čebišev: $P(|\hat{\mu}_n - \mu| \geq \varepsilon) \leq \frac{V}{n\varepsilon^2}$

- $n \geq \frac{V}{\varepsilon^2 \alpha}$

- U slučaju ocene verovatnoće, $V = p(1 - p)$ zavisi od p . Tada uzimamo najveću vrednost $V = 1/4$, ili aproksimaciju $V \approx \hat{p}_n(1 - \hat{p}_n)$.
- Slično se postupa i kod ocene očekivanja ako varijansa zavisi od μ .

Koliko veliko n treba da bude?

Iz ZVB sledi da je, za dato $\varepsilon > 0$ i za svako $\alpha \in (0, 1)$ postoji n takvo da je $P(|\hat{\mu}_n - \mu| \geq \varepsilon) \leq \alpha$.

- Čebišev: $P(|\hat{\mu}_n - \mu| \geq \varepsilon) \leq \frac{V}{n\varepsilon^2}$
- $n \geq \frac{V}{\varepsilon^2 \cdot \alpha}$

- U slučaju ocene verovatnoće, $V = p(1 - p)$ zavisi od p . Tada uzimamo najveću vrednost $V = 1/4$, ili aproksimaciju $V \approx \hat{p}_n(1 - \hat{p}_n)$.
- Slično se postupa i kod ocene očekivanja ako varijansa zavisi od μ .

Koliko veliko n treba da bude?

Iz ZVB sledi da je, za dato $\varepsilon > 0$ i za svako $\alpha \in (0, 1)$ postoji n takvo da je $P(|\hat{\mu}_n - \mu| \geq \varepsilon) \leq \alpha$.

- Čebišev: $P(|\hat{\mu}_n - \mu| \geq \varepsilon) \leq \frac{V}{n\varepsilon^2}$
- $n \geq \frac{V}{\varepsilon^2 \cdot \alpha}$

- U slučaju ocene verovatnoće, $V = p(1 - p)$ zavisi od p . Tada uzimamo najveću vrednost $V = 1/4$, ili aproksimaciju $V \approx \hat{p}_n(1 - \hat{p}_n)$.
- Slično se postupa i kod ocene očekivanja ako varijansa zavisi od μ .

Koliko veliko n treba da bude?

Iz ZVB sledi da je, za dato $\varepsilon > 0$ i za svako $\alpha \in (0, 1)$ postoji n takvo da je $P(|\hat{\mu}_n - \mu| \geq \varepsilon) \leq \alpha$.

- Čebišev: $P(|\hat{\mu}_n - \mu| \geq \varepsilon) \leq \frac{V}{n\varepsilon^2}$
- $n \geq \frac{V}{\varepsilon^2 \cdot \alpha}$

- U slučaju ocene verovatnoće, $V = p(1 - p)$ zavisi od p . Tada uzimamo najveću vrednost $V = 1/4$, ili aproksimaciju $V \approx \hat{p}_n(1 - \hat{p}_n)$.
- Slično se postupa i kod ocene očekivanja ako varijansa zavisi od μ .

Koliko veliko n treba da bude?

Iz ZVB sledi da je, za dato $\varepsilon > 0$ i za svako $\alpha \in (0, 1)$ postoji n takvo da je $P(|\hat{\mu}_n - \mu| \geq \varepsilon) \leq \alpha$.

- Čebišev: $P(|\hat{\mu}_n - \mu| \geq \varepsilon) \leq \frac{V}{n\varepsilon^2}$
- $n \geq \frac{V}{\varepsilon^2 \cdot \alpha}$

- U slučaju ocene verovatnoće, $V = p(1 - p)$ zavisi od p . Tada uzimamo najveću vrednost $V = 1/4$, ili aproksimaciju $V \approx \hat{p}_n(1 - \hat{p}_n)$.
- Slično se postupa i kod ocene očekivanja ako varijansa zavisi od μ .

Koliko veliko n treba da bude?

Iz ZVB sledi da je, za dato $\varepsilon > 0$ i za svako $\alpha \in (0, 1)$ postoji n takvo da je $P(|\hat{\mu}_n - \mu| \geq \varepsilon) \leq \alpha$.

- Čebišev: $P(|\hat{\mu}_n - \mu| \geq \varepsilon) \leq \frac{V}{n\varepsilon^2}$

- $n \geq \frac{V}{\varepsilon^2 \cdot \alpha}$

- U slučaju ocene verovatnoće, $V = p(1 - p)$ zavisi od p . Tada uzimamo najveću vrednost $V = 1/4$, ili aproksimaciju $V \approx \hat{p}_n(1 - \hat{p}_n)$.
- Slično se postupa i kod ocene očekivanja ako varijansa zavisi od μ .

Koliko veliko n treba da bude?

Iz ZVB sledi da je, za dato $\varepsilon > 0$ i za svako $\alpha \in (0, 1)$ postoji n takvo da je $P(|\hat{\mu}_n - \mu| \geq \varepsilon) \leq \alpha$.

- Čebišev: $P(|\hat{\mu}_n - \mu| \geq \varepsilon) \leq \frac{V}{n\varepsilon^2}$

- $n \geq \frac{V}{\varepsilon^2 \cdot \alpha}$

- U slučaju ocene verovatnoće, $V = p(1 - p)$ zavisi od p . Tada uzimamo najveću vrednost $V = 1/4$, ili aproksimaciju $V \approx \hat{p}_n(1 - \hat{p}_n)$.
- Slično se postupa i kod ocene očekivanja ako varijansa zavisi od μ .

Centralna granična teorema

Neka je X_i , $i = 1, 2, \dots$ niz nezavisnih slučajnih promenljivih sa istom raspodelom, $E X = \mu$, $\text{Var } X = \sigma^2$. Normirani zbir

$$Z_n = \frac{X_1 + \dots + X_n - n\mu}{\sigma\sqrt{n}}$$

ima $E Z_n = 0$, $\text{Var } Z_n = 1$.

Teorema 6.9 Niz Z_n konvergira u raspodeli ka $Z \sim \mathcal{N}(0, 1)$, tj.

$$\lim_{n \rightarrow +\infty} P(Z_n \leq x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$$

za svako $x \in \mathbb{R}$.

Centralna granična teorema

Neka je X_i , $i = 1, 2, \dots$ niz nezavisnih slučajnih promenljivih sa istom raspođelom, $E X = \mu$, $\text{Var } X = \sigma^2$. Normirani zbir

$$Z_n = \frac{X_1 + \dots + X_n - n\mu}{\sigma\sqrt{n}}$$

ima $E Z_n = 0$, $\text{Var } Z_n = 1$.

Teorema 6.9 Niz Z_n konvergira u raspođeli ka $Z \sim \mathcal{N}(0, 1)$, tj.

$$\lim_{n \rightarrow +\infty} P(Z_n \leq x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$$

za svako $x \in \mathbb{R}$.

CGT - dokaz

$$Z_n = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} = \frac{(X_1 - \mu) + \cdots + (X_n - \mu)}{\sigma\sqrt{n}}$$

- Možemo pretpostaviti da je $\mu = 0$. Karakteristična funkcija za Z_n je $\varphi_n(t) = \left(\varphi\left(\frac{t}{\sigma\sqrt{n}}\right)\right)^n$.
- Treba dokazati da je $\lim_{n \rightarrow +\infty} \varphi_n(t) = e^{-t^2/2}$, za svako $t \in \mathbb{R}$.
- Smena: $\tau = \frac{t}{\sigma\sqrt{n}}$; $\Psi(\tau) = \log(\varphi_n(t))$
- $\lim_{n \rightarrow +\infty} \varphi_n(t) = e^{-t^2/2} \iff \lim_{\tau \rightarrow 0} \Psi(\tau) = -t^2/2$
- $\Psi(\tau) = \Psi(0) + \Psi'(0)\tau + \Psi''(0)\frac{\tau^2}{2} + o(\tau^2)$ ($\tau \rightarrow 0$)
- $\Psi(0) = 0$, $\Psi'(0) = 0$, $\Psi''(0) = -n\sigma^2$.
- $\Psi(\tau) = -n\sigma^2 \cdot \tau^2 + o(\tau^2) = -\frac{t^2}{2} + o(t^2/\sigma^2 n)$ ($n \rightarrow +\infty$)
- Dokazano!

Procena greške

Za fiksno n , neka su X_1, \dots, X_n nezavisne slučajne promenljive sa istom raspodelom, i neka je F_n funkcija raspodele slučajne promenljive Z_n (normiranog zbira). CGT tvrdi da $F_n(t) \rightarrow \Phi(t)$ za svako t . To znači da za veliko n možemo da kažemo da je $F_n(t) \approx \Phi(t)$. Koliko n treba da bude da bi se greška aproksimacija bila ispod zadatog nivoa?

Teorema 6.12 (Berry-Eseén) Neka je $m_3 = \mathbb{E}|X_1 - \mu|^3$. Tada je

$$\sup_{x \in \mathbb{R}} |F_n(x) - \Phi(x)| \leq C \cdot \frac{m_3}{\sigma^3 \sqrt{n}},$$

gde je $C < \frac{1}{2}$ konstanta koja ne zavisi od raspodele za X_k .

Procena greške

Za fiksno n , neka su X_1, \dots, X_n nezavisne slučajne promenljive sa istom raspodelom, i neka je F_n funkcija raspodele slučajne promenljive Z_n (normiranog zbira). CGT tvrdi da $F_n(t) \rightarrow \Phi(t)$ za svako t . To znači da za veliko n možemo da kažemo da je $F_n(t) \approx \Phi(t)$. Koliko n treba da bude da bi se greška aproksimacija bila ispod zadatog nivoa?

Teorema 6.12 (Berry-Eseén) Neka je $m_3 = \mathbb{E} |X_1 - \mu|^3$. Tada je

$$\sup_{x \in \mathbb{R}} |F_n(x) - \Phi(x)| \leq C \cdot \frac{m_3}{\sigma^3 \sqrt{n}},$$

gde je $C < \frac{1}{2}$ konstanta koja ne zavisi od raspodele za X_k .

Apksimacije preko CGT

Za bilo koju početnu raspodelu i za $n \geq 20$ (empirijsko pravilo):

Normirani zbir Z_n

$$Z_n = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \sim \mathcal{N}(0, 1)$$

Zbir

$$X_1 + \cdots + X_n \sim \mathcal{N}(n\mu, n\sigma^2).$$

Aritmetička sredina

$$\hat{\mu}_n = \frac{X_1 + \cdots + X_n}{n} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right), \quad \frac{\hat{\mu}_n - \mu}{\sigma} \sqrt{n} \sim \mathcal{N}(0, 1)$$

Apromimacije preko CGT

Za bilo koju početnu raspodelu i za $n \geq 20$ (empirijsko pravilo):

Normirani zbir Z_n

$$Z_n = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \sim \mathcal{N}(0, 1)$$

Zbir

$$X_1 + \cdots + X_n \sim \mathcal{N}(n\mu, n\sigma^2).$$

Aritmetička sredina

$$\hat{\mu}_n = \frac{X_1 + \cdots + X_n}{n} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right), \quad \frac{\hat{\mu}_n - \mu}{\sigma} \sqrt{n} \sim \mathcal{N}(0, 1)$$

Aproksimacije preko CGT

Za bilo koju početnu raspodelu i za $n \geq 20$ (empirijsko pravilo):

Normirani zbir Z_n

$$Z_n = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \sim \mathcal{N}(0, 1)$$

Zbir

$$X_1 + \cdots + X_n \sim \mathcal{N}(n\mu, n\sigma^2).$$

Aritmetička sredina

$$\hat{\mu}_n = \frac{X_1 + \cdots + X_n}{n} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right), \quad \frac{\hat{\mu}_n - \mu}{\sigma} \sqrt{n} \sim \mathcal{N}(0, 1)$$

Koliko n treba da bude?

Primer 126. Uzimamo uzorak X_1, \dots, X_n iz raspodele sa (nepoznatim) matematičkim očekivanjem μ i (poznatom) varijansom σ^2 i računamo ocenu $\hat{\mu}$. Koliko n treba da bude da bi bio zadovoljen uslov

$$P(|\hat{\mu}_n - \mu| \geq \varepsilon) \leq \alpha$$

za date parametre $\varepsilon > 0$ i $\alpha \in [0, 1]$? Naći brojnu vrednost za n sa $\sigma^2 = 4$, $\varepsilon = 0.1$ i $\alpha = 0.1$

R: Preko CGT: $n \geq \frac{\sigma^2 K_{1-\alpha/2}^2}{\varepsilon^2} = 1100$ Preko Čebiševa: $n \geq \frac{\sigma^2}{\alpha \varepsilon^2} = 4000$

Aproksimacije raspodela preko CGT

Binomna raspodela

Ako je početna raspodela Bernulijeva sa verovatnoćom uspeha p , onda je raspodela zbira $\text{Bin}(n, p)$.

$$\text{Bin}(n, p) \sim \mathcal{N}(np, np(1 - p))$$

za $n \geq 20$. Ova aproksimacija primenjuje se za $np > 5$ (za $np < 5$ - Puasonova aproksimacija).

Primer 128 Ako je $X \sim \text{Bin}(100, 1/2)$, naći približnu vrednost verovatnoća: a) $P(X > 60)$, b) $P(40 < X < 60)$, c) $P(X = 50)$.

Aproksimacije raspodela preko CGT

Binomna raspodela

Ako je početna raspodela Bernulijeva sa verovatnoćom uspeha p , onda je raspodela zbira $\text{Bin}(n, p)$.

$$\text{Bin}(n, p) \sim \mathcal{N}(np, np(1 - p))$$

za $n \geq 20$. Ova aproksimacija primenjuje se za $np > 5$ (za $np < 5$ - Puasonova aproksimacija).

Primer 128 *Ako je $X \sim \text{Bin}(100, 1/2)$, naći približnu vrednost verovatnoća: **a)** $P(X > 60)$, **b)** $P(40 < X < 60)$, **c)** $P(X = 50)$.*

Aproksimacije raspodela preko CGT-nastavak

Puasonova raspodela

Ako je početna raspodela $\text{Poiss}(\frac{\lambda}{n})$, onda je raspodela zbira $\text{Poiss}(\lambda)$.

$$\text{Poiss}(\lambda) \sim \mathcal{N}(\lambda, \lambda)$$

Ova aproksimacija koristi se za $\lambda \geq 10$ (za male vrednosti λ nema ni potrebe za aproksimacijom).

Za vežbu: Primeri 126,128,129. Zadaci: 125-139, 307-311.