Probleme recapitulative

① Aplicații la teorema lui Slutsky

Teorema lui Slutsky

Fie $(X_n)_n, (Y_n)_n$, $n \in \mathbb{N}$ două șiruri de v.a. și $c \in \mathbb{R}$. Atunci:

a) $\left.\begin{array}{l} X_n \xrightarrow{\mathcal{D}} X \\ Y_n \xrightarrow{\mathbb{P}} c \end{array}\right\} \Rightarrow X_n + Y_n \xrightarrow{\mathcal{D}} X + c$

b) $\left.\begin{array}{l} X_n \xrightarrow{\mathcal{D}} X \\ Y_n \xrightarrow{\mathbb{P}} c \end{array}\right\} \Rightarrow \begin{array}{l} X_n Y_n \xrightarrow{\mathcal{D}} c \cdot X, \text{ pt. } c \neq 0 \\ X_n Y_n \xrightarrow{\mathbb{P}} 0, \text{ pt. } c = 0. \end{array}$

c) $\left.\begin{array}{l} X_n \xrightarrow{\mathcal{D}} X \\ Y_n \xrightarrow{\mathbb{P}} c \end{array}\right\} \Rightarrow \dfrac{X_n}{Y_n} \xrightarrow{\mathcal{D}} \dfrac{X}{c}$ pentru $c \neq 0$.

$\boxed{1}$ Fie $X_1, X_2 \dots X_n$ i.i.d. $\sim N(0,1)$. Determinați repartiția limită pentru v.a. W_n definită astfel:

$$\boxed{W_n = \sqrt{n} \cdot \frac{\sum_{i=1}^{n} X_i}{\sum_{i=1}^{n} X_i^2}}$$

Cum gândim?

→ vedem sume de v.a. și ne gândim ce rezultate cunoaștem despre repartiția unei sume

ex. sumă de Bernoulli(p) ⟶ Binomială(n,p) etc.

→ vedem un raport de v.a., deci putem folosi c) de la teorema lui Slutsky

→ vedem o sumă de pătrate de v.a. normale standard și ne gândim la repartiția χ^2

Pași:

1) Construim două v.a. a.î. raportul lor să fie egal cu W_n

Fie $V_n = \dfrac{\sum_{i=1}^{n} X_i^2}{n}$, $U_n = \dfrac{\sum_{i=1}^{n} X_i}{\sqrt{n}}$

Vedem că $W_n = \dfrac{U_n}{V_n}$.

Pentru a găsi repartițiile limită pentru U_n și V_n este util uneori (și în cazul de față!) să ne folosim de funcția generatoare de momente

↳ deoarece, pentru repartițiile cunoscute se știu aceste funcții și în felul acesta identificăm repartiția v.a. de interes

2) Calculăm m.g.f. (moment generating function) pt. U_n și V_n:

- $$M_{U_n}(t) = E\left(e^{t \cdot U_n}\right) = E\left(e^{t \cdot \frac{\sum_{i=1}^{n} X_i}{\sqrt{n}}}\right) = E\Big(\underbrace{e^{\frac{t}{\sqrt{n}} \cdot X_1} \cdot \ldots \cdot e^{\frac{t}{\sqrt{n}} X_n}}_{\text{i.i.d. pt. că } X_1 \ldots X_n \text{ sunt i.i.d.}}\Big)$$

$$= \prod_{i=1}^{n} E\left(e^{\frac{t}{\sqrt{n}} X_i}\right)$$

Cum $X_i \sim N(0,1)$ $\boxed{M_{X_i}(t) = E\left(e^{tX_i}\right) = e^{\frac{1}{2}t^2}}$

vezi tabel cu m.g.f.

$$E\left(e^{\frac{t}{\sqrt{n}} \cdot X_i}\right) = \int_{-\infty}^{\infty} e^{\frac{t}{\sqrt{n}} x} \cdot \underbrace{f(x)}_{\text{densitatea normalei}}\, dx = \int_{-\infty}^{\infty} e^{\frac{t}{\sqrt{n}} x} \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}x^2} dx$$

Reamintim: $\boxed{\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}}$ integrala Euler-Poisson

$$= \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(x^2 - 2 \cdot \frac{t}{\sqrt{n}} \cdot x + \frac{t^2}{n} - \frac{t^2}{n}\right)} dx$$

vrem să facem un pătrat perfect pentru a putea să ne folosim de integrala Euler-Poisson

$$I = \frac{1}{\sqrt{2\pi}} \cdot e^{\frac{1}{2} \cdot \frac{t^2}{n}} \cdot \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(x - \frac{t}{\sqrt{n}}\right)^2} dx = \frac{1}{\sqrt{2\pi}} \cdot e^{\frac{1}{2}\frac{t^2}{n}} \cdot \int_{-\infty}^{\infty} e^{-\left[\frac{1}{\sqrt{2}}\left(x - \frac{t}{\sqrt{n}}\right)\right]^2} dx$$

Facem S.V.:

$$\frac{1}{\sqrt{2}}\left(x-\frac{t}{\sqrt{n}}\right)=y \Rightarrow x=2y+\frac{t}{\sqrt{n}} \Rightarrow dx=\sqrt{2}\,dy$$

$x\to-\infty \Rightarrow y\to-\infty$

$x\to\infty \Rightarrow y\to\infty$

$$I=\frac{1}{\sqrt{2\pi}}\cdot e^{\frac{1}{2}\frac{t^2}{n}}\cdot\int_{-\infty}^{\infty} e^{-y^2}\cdot\sqrt{2}\,dy = e^{\frac{1}{2}\frac{t^2}{n}}\cdot\frac{1}{\cancel{\sqrt{2}}}\cdot\cancel{\sqrt{2}}\cdot\frac{1}{\cancel{\sqrt{\pi}}}\cdot\cancel{\sqrt{\pi}}$$

Deci $M_{U_n}(t)=\prod_{i=1}^{n} e^{\frac{1}{2}\frac{t^2}{n}} = \underbrace{e^{\frac{1}{2}t^2}}$

↳ observăm că este m.g.f.-ul unei normale standard

Așadar $U_n \xrightarrow{\mathcal{D}} Z\sim N(0,1)$

• $M_{V_n}(t)=E\left(e^{tV_n}\right)=E\left(e^{\frac{t}{\sqrt{n}}\cdot\sum_{i=1}^{n}X_i^2}\right)\overset{X_1,\dots X_n}{\underset{i.i.d.}{=}}\prod_{i=1}^{n}E\left(e^{\frac{t}{\sqrt{n}}X_i^2}\right)$

Cum $X_i\sim N(0,1) \Rightarrow X_i^2\sim\chi_1^2$ (chi-pătrat cu un grad de libertate)

$$\boxed{M_{X_i^2}(t)=E\left(e^{tX_i^2}\right)=\frac{1}{(1-2t)^{1/2}} \text{ pentru } t<\frac{1}{2}.}$$

(vezi tabelul cu m.g.f. pt. χ^2)

Analog, obținem:

$$M_{V_n}(t)=\underbrace{\left(1-\frac{2t}{n}\right)^{-\frac{n}{2}}} \text{ pentru } t<\frac{1}{2}$$

↳ observăm că este m.g.f.-ul unei v.a. gama cu parametrii $\alpha=\frac{n}{2}$ și $\beta=\frac{2}{n}$

Așadar densitatea lui V_n este dată de:

$$f_{V_n}(x)=\begin{cases}\frac{1}{\Gamma(\frac{n}{2})}\cdot\frac{1}{(\frac{2}{n})^{\frac{n}{2}}}\cdot x^{\frac{n}{2}-1}\cdot e^{-\frac{nx}{2}}, & x>0\\ 0, & \text{în rest}\end{cases}$$

Cum $E(V_n) = \alpha \cdot \beta = 1$ putem aplica inegalitatea lui Cebâșev și obținem:

$$\mathbb{P}(|V_n - 1| > \varepsilon) \leq \frac{Var(V_n)}{\varepsilon^2} = \frac{\alpha \beta^2}{\varepsilon^2} = \frac{\frac{n}{2} \cdot \left(\frac{2}{n}\right)^2}{\varepsilon^2}$$

$$= \frac{2}{n\varepsilon^2} \xrightarrow{n \to \infty} 0$$

Așadar, $V_n \xrightarrow{P} 1$.

3) Invocăm teorema lui Slutsky:

$$\left.\begin{array}{l} U_n \xrightarrow{\mathcal{D}} Z \sim N(0,1) \\ V_n \xrightarrow{P} 1 \end{array}\right\} \xRightarrow{T.} W_n = \frac{U_n}{V_n} \xrightarrow{\mathcal{D}} 1 \cdot Z = Z \sim N(0,1)$$

Așadar repartiția limită a v.a. W_n este normala standard.

2. Aplicații la metoda delta

Metoda delta

Fie $(X_n)_n$ un șir de v.a., $(v_n)_n$ un șir numeric a.î. $v_n \xrightarrow{n\to\infty} \infty$. Presupunem că există v.a. X și $a \in \mathbb{R}$ a.î.:

$$v_n(X_n - a) \xrightarrow{\mathcal{D}} X$$

Atunci, dacă g este o funcție derivabilă în a:

$$\boxed{v_n(g(X_n) - g(a)) \xrightarrow{\mathcal{D}} g'(a)X}$$

În particular, dacă $v_n = \sqrt{n}$ și $X \sim N(0, \sigma^2)$ iar

$$\sqrt{n}(X_n - a) \xrightarrow{\mathcal{D}} X \sim N(0, \sigma^2) \text{ atunci}$$

$$\sqrt{n}(g(X_n) - g(a)) \xrightarrow{\mathcal{D}} Y \sim N(0, (g'(a)\sigma)^2)$$

---

1 Calculați $\mathbb{P}(\overline{X}(1-\overline{X}) \leq y)$ când $n \to \infty$.

Cum gândim?

→ ni se cere să calculăm o probabilitate când $n \to \infty$, deci vrem să vedem la cine tinde funcția de repartiție a v.a. de interes

→ convergența unui șir de funcții de repartiție ne duce cu gândul la convergența în distribuție

→ convergența în distribuție ne duce cu gândul la mai multe ~~probl~~ teoreme, printre care și metoda delta.

Pași:

1) Știm din T.L.C. că $\overline{X} \xrightarrow{\mathcal{D}} N(\mu, \frac{\sigma^2}{n})$

Fie $g(\overline{X}) = \overline{X}(1-\overline{X})$ unde $\overline{X}$ este media de selecție dintr-o populație cu medie μ și dispersie σ^2

Alegem $v_n = \sqrt{n}$ și $a = \mu$

Știm $\sqrt{n}(\overline{X}-\mu) \xrightarrow{\mathcal{D}} N(0, \sigma^2)$, deci ne aflăm în cazul particular al metodei delta, deci:

$$\sqrt{n}(g(\overline{X}) - g(\mu)) \xrightarrow{\mathcal{D}} N(0, (g'(\mu)\cdot\sigma)^2)$$

g e derivabilă și $g'(\mu) = 1-2\mu$

Deci $\sqrt{n}(\overline{X}(1-\overline{X}) - \mu(1-\mu)) \xrightarrow{\mathcal{D}} N(0, [(1-2\mu)\sigma]^2)$

2) Exprimăm probabilitatea de interes în raport cu relația de mai sus:

$$\mathbb{P}(\overline{X}(1-\overline{X}) \leq y) = \mathbb{P}(\overline{X}(1-\overline{X}) - \mu(1-\mu) \leq y - \mu(1-\mu))$$

$$= \mathbb{P}\left(\frac{\overline{X}(1-\overline{X}) - \mu(1-\mu)}{|1-2\mu|\sigma/\sqrt{n}} \leq \frac{y-\mu(1-\mu)}{|1-2\mu|\sigma/\sqrt{n}}\right) \quad \text{pt. } \mu \neq \frac{1}{2}$$

am împărțit prin abaterea medie pătratică pentru a ajunge la o normală standard pt. $n \to \infty$

$$\overset{n\to\infty}{\simeq} \Phi\left(\frac{y-\mu(1-\mu)}{\frac{|1-2\mu|\cdot\sigma}{\sqrt{n}}}\right)$$

funcția de repartiție a normalei standard

3 Aplicații la intervale de încredere

Obs. O metodă foarte accesibilă pentru a construi intervale de încredere constă în folosirea inegalității lui Cebâșev (deși de obicei lungimea acestor intervale nu este optimă!):

Inegalitatea lui Cebâșev

$$\mathbb{P}\left(|X-E(X)| < \varepsilon\sqrt{Var(X)}\right) > 1-\frac{1}{\varepsilon^2} \quad \forall \varepsilon > 0$$

$\boxed{1}$ Fie $\hat{\theta}$ un estimator pentru un parametru θ, cu varianță (nu neapărat nedeplasat) finită. Atunci:

$$\mathbb{P}\left(|\hat{\theta}-\theta| < \varepsilon \cdot \sqrt{E(\hat{\theta}-\theta)^2}\right) > 1-\frac{1}{\varepsilon^2}$$

$$\Leftrightarrow \mathbb{P}\left(\hat{\theta}-\varepsilon\sqrt{E(\hat{\theta}-\theta)^2} < \theta < \hat{\theta}+\varepsilon\sqrt{E(\hat{\theta}-\theta)^2}\right) > 1-\frac{1}{\varepsilon^2}$$

Deci $\left(\hat{\theta}-\varepsilon\sqrt{E(\hat{\theta}-\theta)^2}, \hat{\theta}+\varepsilon\sqrt{E(\hat{\theta}-\theta)^2}\right)$ este un interval de încredere de nivel $1-\frac{1}{\varepsilon^2}$ de încredere pentru θ.

$\boxed{2}$ Fie $X_1, X_2 \dots X_n \sim Unif(0,\theta)$ i.i.d. Să se construiască un interval de încredere pentru θ.

Metoda I (folosim inegalitatea lui Cebâșev)

Folosind metoda verosimilității maxime obținem că $\hat{\theta} = X_{(n)}$.

Vrem să aplicăm inegalitatea lui Cebâșev, motiv pentru care calculăm:

$$E_\theta(X_{(n)}) = \frac{n}{n+1}\cdot\theta \quad \text{și} \quad E_\theta(X_{(n)}-\theta)^2 = \frac{2\theta^2}{(n+1)(n+2)}$$

Deci, $\mathbb{P}\left(|X_{(n)} - \theta| < \varepsilon \cdot \frac{\sqrt{2}\,\theta}{\sqrt{(n+1)(n+2)}}\right) > 1 - \frac{1}{\varepsilon^2}$

$\Leftrightarrow \mathbb{P}\left(\frac{|X_{(n)} - \theta|}{\theta} \cdot \sqrt{\frac{(n+1)(n+2)}{2}} < \varepsilon\right) > 1 - \frac{1}{\varepsilon^2}$

Cum $X_{(n)} \xrightarrow{P} \theta$ putem scrie, pentru un n suficient de mare:

$\mathbb{P}\left(\frac{|X_{(n)} - \theta|}{X_{(n)}} \cdot \sqrt{\frac{(n+1)(n+2)}{2}} < \varepsilon\right) > 1 - \frac{1}{\varepsilon^2}$

Deci $\left(X_{(n)} - \varepsilon \cdot X_{(n)} \cdot \frac{\sqrt{2}}{\sqrt{(n+1)(n+2)}},\ X_{(n)} + \varepsilon X_{(n)} \cdot \frac{\sqrt{2}}{\sqrt{(n+1)(n+2)}}\right)$

este un interval de încredere cu un nivel $1 - \frac{1}{\varepsilon^2}$ de încredere.

Metoda a II-a

Densitatea v.a. $X_{(n)}$ este $f_n(y) = n \cdot \frac{y^{n-1}}{\theta^n}, 0<y<\theta$

(vezi rezultate curs / seminar)

Atunci v.a. $T = \frac{X_{(n)}}{\theta}$ are densitatea $h(t) = nt^{n-1}$ $\forall 0<t<1$

---

def. Fie $X \sim f_\theta$. O v.a. $T(X, \theta)$ se numește $\boxed{\text{pivot}}$ dacă repartiția lui T nu depinde de θ.

Teoremă: Fie $T(X, \theta)$ un pivot cu proprietatea că pentru fiecare θ $T(X, \theta)$ este o statistică care, privită ca funcție de θ e fie strict crescătoare, fie strict descr. în orice x.

Atunci se poate construi interval de încredere pt. θ de orice nivel.

Vom construi un interval de încredere $\left(\frac{X_{(n)}}{b}, \frac{X_{(n)}}{a}\right)$ cu lungimea $L = X_{(n)}\left(\frac{1}{a} - \frac{1}{b}\right)$

Vrem să minimizăm lungimea intervalului de încredere păstrând aceeași probabilitatea asociată, adică impunând condiția ca $\int_a^b h(t)\,dt = 1-\alpha \iff \int_a^b n t^{n-1}\,dt = 1-\alpha$
$\iff b^n - a^n = 1-\alpha$

Cum $(1-\alpha)^{\frac{1}{n}} < b \le 1$ și $\frac{dL}{db} = X_{(n)} \cdot \left(-\frac{1}{a^2} \cdot \frac{da}{db} + \frac{1}{b^2}\right)$
$= X_{(n)} \cdot \left(\frac{a^{n+1} - b^{n+1}}{b^2 a^{n+1}}\right) < 0$ rezultă că minimul se atinge pentru $b = 1$.

Deci intervalul de încredere de lungime minimă cu nivelul de încredere $1-\alpha$ este: $\left(X_{(n)}, \frac{X_{(n)}}{\alpha^{1/n}}\right)$.