* деф ($\boxed{\text{Генерална съвкупност}}$):

Генерална съвкупност (популация) се нарича съвкупността от всички изследвани обекти.

* деф ($\boxed{\text{случайна извадка}}$):

Случайният вектор $\vec{X} = (X_1, X_2, \dots, X_n)$ от независими и еднакво разпределени сл. вел. (i.i.d) с маргинална ф-я на разпределение $F_X(x)$ се нарича случайна извадка с размер n над X. $\vec{X}$ са копия на X

* деф ($\boxed{\text{точкова оценка}}$):

Функцията $\hat{\theta} = \hat{\theta}(\vec{X})$ се нарича статистика (още се нарича и точкова оценка на θ).

* деф ($\boxed{\text{ф-я на максималното правдоподобие}}$):

~~Съвместната плътност на $\vec{X} = (X_1, \dots, X_n)$~~

~~$f(x_1, \dots, x_n) = \prod_{i=1}^{n} f_{X_i}$~~

Функция на максималното правдоподобие наричаме съвместната плътност на $\vec{X}$, т.е.

$$L(\vec{X}, \theta) = \prod_{i=1}^{n} f_X(X_i, \theta).$$ → всеки път една и съща ф-я, но в различна сл. вел X_i, параметризирана по θ

→ Ако има конкретна реализация $\vec{x}$, то ф-ята на правдоподобие $L(\vec{x}, \theta) = \prod_{i=1}^{n} f_X(x_i, \theta)$ → съвместната плътност в конкретна точка в $\mathbb{R}^n$

Деф (оценка по метода на м.п.):

Статистиката $\hat{\theta}(\vec{x})$ се нарича оценка по метода на м.п., ако:
$$L(\vec{x}, \theta) = \sup_{\theta \in \Theta} L(\vec{x}, \theta), \text{ където } \Theta \text{ е}$$
множеството от всички допустими стойности за θ.

Ако $f_X(x, \theta)$ е диф. по θ, то $\hat{\theta}$ е решение на
$$\frac{\partial L}{\partial \theta} = 0 \iff \frac{\partial \ln L}{\partial \theta} = 0$$

Твърдение
(оценка по м.м.п. за $X \sim N(\mu, \sigma^2)$):

Нека $X \sim N(\mu, \sigma^2)$ и $\vec{x}$ е извадка над X. Тогава м.п.о. за μ е $\hat{\mu}(\vec{x}) = \overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} x_i$ (независимо дали σ^2 е известно или не).

МПО за σ^2 е:

i) $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2$, когато μ е известно

ii) $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{X_n})^2$, когато μ не е известно.

Деф (оценка по ММ):

Нека X е сл. вел. и има ф-я на разпределение $F_X(x, \theta)$, $\vec{x}$ е извадка над X и $\theta = (\theta_1, \dots, \theta_s)$.

$E(X^k) = \mu^{(k)}(\theta)$ → теоретичните моменти, $k = 1, 2, \dots$

$\overline{X_n}^{(k)} = \frac{1}{n}\sum_{i=1}^{n} x_i^k$ → емпирични моменти, $k = 1, 2, \dots$

Тогава решението на системата: $\mu^{(k)}(\theta) = \overline{X_n}^{(k)}$ за $k = 1, \dots, s$, се нарича оценка по м. м.

→ Свойства:

* Деф (неизместеност):

Казваме, че $\hat{\theta}$ е неизместена оценка на θ, ако $\mathbb{E}\hat{\theta}(\vec{X}) = \theta$. Когато $\theta = (\theta_1, \dots, \theta_s)$, равенството се разбира като $\mathbb{E}\hat{\theta}_j(\vec{X}) = \theta_j$, $1 \le j \le s$.

↳ Иначе, ако това не е изпълнено, то тогава $|\mathbb{E}\hat{\theta} - \theta|$ се нарича систематична грешка.

* $\mathbb{E}\hat{\theta} - \theta = 0$ → за да е неизместена

* Деф (състоятелност):

Ако $\hat{\theta}(\vec{X}) \xrightarrow[n\to\infty]{\mathbb{P}} \theta$, то $\hat{\theta}$ е състоятелна оценка.

↳ Какво означава това ⇒

$\theta \in \mathbb{R} \Rightarrow \mathbb{P}(|\hat{\theta}(\vec{X}) - \theta| > \varepsilon) \xrightarrow[n\to\infty]{} 0$

→ т.е. тази вероятност клони към 0, т.е. няма как да е $> \varepsilon$?

* Деф (силна състоятелност):

Ако $\hat{\theta}(\vec{X}) \xrightarrow[n\to\infty]{п.с.} \theta$, то $\hat{\theta}$ е силно състоятелна оценка.

* Деф: функция на МП наричаме съвместната плътност на $\vec{X}$, т.е.

$$L(\vec{X}, \theta) = \prod_{j=1}^{n} f_{X_j}(x_j, \theta)$$

→ понеже от сл. вел. са независими и еднакво разпределени

→ Неравенство на Бери–Есеен (равномерно)

$X_1 \sim Ber(p) \Rightarrow \mathbb{E}|X_1-p|^3 = p(1-p)(p^2+(1-p)^2) \le \frac{1}{16}$

Рецепта $n \ge 30 \Rightarrow$ ЦГТ е валидна ?

Следствие: $X_n \sim Bin(n,p)$

Тогава $\forall a<b$ $\lim\limits_{n\to\infty} P\left(a < \frac{X_n - np}{\sqrt{np(1-p)}} < b\right) =$

$$= \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{y^2}{2}}dy = \Phi(b)-\Phi(a)$$

$\sigma = \sqrt{p(1-p)}$

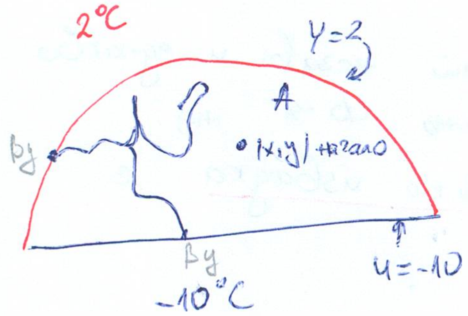
д-во:

$X_n = \sum_{i=1}^n Y_i \qquad Y_i \sim Ber(p)$

$$P\left(a < \frac{\sum_{i=1}^n Y_i - np}{\sqrt{np(1-p)}} < b\right) = P\left(\frac{\sum_{i=1}^n Y_i - np}{\sqrt{np(1-p)}} < b\right) - P\left(\frac{\sum_{i=1}^n Y_i - np}{\sqrt{np(1-p)}} \le a\right)$$

$$\xrightarrow[n\to\infty]{\text{ЦГТ}} \Phi(b) - \Phi(a)$$

→



Искаме да сметнем температурата

$$\begin{cases} \Delta f = 0, & A \\ f = u, & \partial A \end{cases}$$

$f(x,y)$ е темп. в (x,y)

$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

* Сл: $X_n \sim Bin(n,p)$. Тогава $\frac{X_n-np}{\sqrt{np(1-p)}} \xrightarrow[n\to\infty]{d} Z$.

Задача на Дирихле

Взимаме Брауново движение и гледаме къде излиза

$$\mathbb{E}_{(x,y)}[u(B)] = f(x,y) \approx \frac{1}{10}\sum_{j=1}^{10} u(B_j)$$

Изисква конкретна температура в границите (затова се нарича задача на Дирихле).

* фон Нойман казва за методи

Точкови оценки

23.05

задача: Изследване на характеристика (свойства) на съвкупност от обекти (индивиди).

деф: Генерална съвкупност (популация) се нарича съвкупността от всички изследвани обекти.

ГС → всички българи

Характеристики → ръст, тегло, преференция, размер

Нека предположим, че X (сл. вел) описва изследвана характеристика. Допълнително, X има ф-я на разп. $F_X(x)$

деф: Случайният вектор $\vec{X} = (x_1, \dots, x_n)$ от незав. и еднакво разпределени сл. вел. (н.е.р = iid) с маргинална ф-я на разпределение $F_X(x)$ се нарича случайна извадка с размер n над X.

$\vec{X}$ са копия на X

Пример: Digest 2.38 милиона читатели → Рузвелт 38%; 4000 сл. избрани гласоподаватели → Рузвелт 56%

Нека предположим, че $F_X(x) = F_X(x, \theta)$, т.е. разпределението на X зависи от неизвестен параметър θ.
↳ възможно е да е вектор от параметри

⊕ Пример: $X \sim N(\mu, \sigma^2)$

$\theta = (\mu, \sigma^2)$

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \qquad \mu \in \mathbb{R}, \ \sigma^2 \in [0, +\infty)$$

⊕ Пример: $X \sim Ber(p)$, $\theta = p$; $X \sim Bin(n, p)$, $\theta = (n, p)$

* деф: функцията $\theta = \theta(\vec{X})$ се нарича статистика (още се нарича и точкова оценка).

* Забележка: ф-ята θ се избира преди да измерим реализациите на $\vec{X}$

$\vec{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$

$\theta(\vec{X}) \to$ точков оценител

$\theta(\vec{x}) \to$ точкова оценка

A. Метод на максималното правдоподобие (М.О)

Предполагаме, че $\vec{X}$ е извадка на X с ф.р. $F_X(x, \theta)$ и плътност $f_X(x, \theta)$

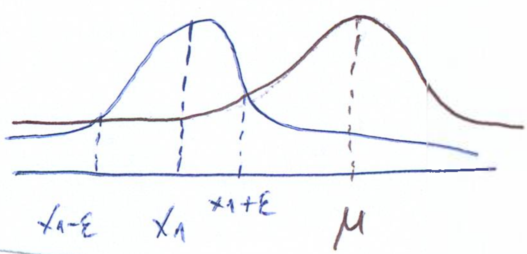
* деф: Съвместна плътност на $\vec{X} = (X_1, \dots, X_n)$

$$f_{X_1, \dots, X_n}(x_1, \dots, x_n) = \prod_{j=1}^{n} f_{X_j}(x_j) = \prod_{j=1}^{n} d\, x_1(x_j) := L = (\vec{x}, \theta)$$

се нарича ф-я на правоподобието.

⊕ Пример: $n = 1$ $X \sim N(\mu, 1)$ $\theta = \mu$

$$L(x, \theta) = L(x_1, \mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_1-\mu)^2}{2}}$$



$$\mathbb{P}(X \in (x_i-\varepsilon ; x_i+\varepsilon)) =$$

$$= \int_{x_i-\varepsilon}^{x_i+\varepsilon} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2}} dx \approx \underset{\substack{\downarrow \\ \text{дължината на интервала}}}{2\varepsilon}\, f_X(x_1, \mu)$$

$$\boxed{L(\vec{x}, \theta) = \prod_{j=1}^{n} f(x_j, \theta)}$$

$\mu \to \sup_{\mu} f_X(x_1, \mu)$

$\hat{\mu}(x_1) = x_1$

деф: Статистиката $\hat{\Theta}(\vec{x})$ се нар. оценка по метода на М.п., ако $L(\vec{x}, \hat{\theta}) = \sup_{\theta \in \Theta} L(\vec{x}, \theta)$, където Θ е мн-вото от всички допустими стойности за θ.

Ако $f_X(x, \theta)$ е диф. по θ, то $\hat{\theta}$ е реш. на

$$\frac{\partial L}{\partial \theta} = 0 \iff \frac{\partial \ln L}{\partial \theta} = 0$$ → лог. е мон. раст. ф-я и има с-во произведението до сбора Фума

4 Пример: $X \sim U(0, \theta)$, $\Theta = [0, +\infty)$

и $\vec{x}$ е извадка над X

$$f_X(x) = \begin{cases} \frac{1}{\theta}, & x \in [0, \theta] \\ 0, & x \notin [0, \theta] \text{ иначе} \end{cases}$$

$$L(\vec{x}, \theta) = \prod_{j=1}^{n} f_{x_j}(x_j) =$$

$$= \prod_{j=1}^{n} \frac{1}{\theta} 1_{\{x_j \in [0, \theta]\}}$$

$$= \frac{1}{\theta^n} 1_{\{\max(x_1, \dots, x_n) \in [0, \theta]\}}$$

$\frac{1}{\theta^n}$ → намаляваща по отн на θ

$\hat{\theta}(\vec{x}) = \max(x_1, \dots, x_n)$

$\hat{\theta}(\vec{x}) = \max(x_1, \dots, x_n) \to$ м.п.о.

Б. Метод на моментите

От ЗГЧ X сл. вел. $\mathbb{E}X = \mu$

$\overline{X_n} \xrightarrow[n\to\infty]{\text{п.с.}} \mu$ (силен) $\quad \mathbb{P}(\lim_{n\to\infty} \overline{X_n} = \mu) = 1$

$\overline{X_n} \xrightarrow[n\to\infty]{\mathbb{P}} \mu$ (слаб) $\quad \lim_{n\to\infty} \mathbb{P}(|\overline{X_n} - \mu| > \varepsilon) = 0, \ \forall \varepsilon > 0$

* **Деф:** Нека X има ф.р. $F_X(x, \theta)$, $\vec{x}$ е извадка от X,

$\mathbb{E}(X^k) = \mu^{(k)}(\theta) \to$ теоретични моменти, $k = 1, 2, \dots$

$\overline{X_n}^{(k)} = \frac{1}{n} \sum_{j=1}^{n} x_j^k \to$ емпирични моменти, $k = 1, 2, \dots$

и $\theta = (\theta_1, \dots, \theta_s)$. Тогава, решение на системата

$\mu^{(k)}(\theta) = \overline{X_n}^{(k)}$ за $k = 1, \dots, s$

θ се нарича оценка по м.м.

Пример: $X \sim U(0, \theta)$, $\vec{x}$ е извадка на X

$\hat{\theta} = \max(x_1, \dots, x_n)$

$\mathbb{E}X = \frac{\theta}{2}$ $\quad \frac{\theta}{2} = \overline{x_n}$

$\overline{X_n} = \frac{1}{n} \sum_{j=1}^{n} x_j$ $\quad \tilde{\theta} = 2\overline{x_n} \neq \hat{\theta}$

Пример: $X \sim N(\mu, \sigma^2)$, $\vec{x}$ е извадка

$\mathbb{E}X = \mu$

$\mathbb{E}X^2 = \sigma^2 + \mu^2$

$\boxed{\mu = \overline{x_n}}$ $\quad \tilde{\mu} = \hat{\mu}$ $\quad$ заместваме

$\sigma^2 + \mu^2 = \overline{x_n}^{(2)} = \frac{1}{n} \sum_{j=1}^{n} x_j^2$

$\sigma^2 = \frac{1}{n} \sum_{j=1}^{n} x_j^2 - (\overline{x_n})^2 = \overline{x_n}^{(2)} - (\overline{x_n})^2 = \sigma^2$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{j=1}^{n}(X_j - \overline{X}_n)^2 = \frac{1}{n}\sum_{j=1}^{n}(X_j^2 - 2X_j\overline{X}_n + (\overline{X}_n)^2) =$$

$$= \frac{1}{n}\sum_{j=1}^{n} X_j^2 - \frac{2}{n}\overline{X}_n \sum_{j=1}^{n} X_j + \frac{1}{n}\cdot n\cdot(\overline{X}_n)^2 =$$

$$= \overline{X}_n^{(2)} - (\overline{X}_n)^2$$

* Свойства на оценките

* деф: Нека $\Theta(\vec{X})$ е оценител (оценка) за Θ, базиран на случайна извадка $\vec{X}$ над сл. вел. X.

$\Theta(\vec{X})$ се нарича неизместима оценка за Θ, ако

$$E(\Theta(\vec{X}) - \Theta) = 0$$

$\Theta(\vec{X}) - \Theta \to$ систематична грешка

* Пример: $X \sim N(\mu, \sigma^2)$ $\quad EX = \mu \quad DX = \sigma^2$

$$\hat{\mu} = \overline{X}_n$$

$$E(\hat{\mu} - \mu) = 0 \Leftrightarrow E\hat{\mu} = E\,\frac{1}{n}\sum_{j=1}^{n} X_j = \frac{1}{n} E\sum_{j=1}^{n} X_j =$$

$$= \frac{1}{n}\sum_{j=1}^{n} EX_j = \frac{1}{n}\cdot\sum_{j=1}^{n}\mu = \frac{1}{n}\cdot n\cdot\mu = \mu$$

извадката има същото разпределение като X

$\to \hat{\mu}$ е неизместима оценка

$$E(\hat{\sigma}^2) = E\left[\frac{1}{n}\sum_{j=1}^{n}(X_j - \overline{X}_n)^2\right] = E\left[\overline{X}_n^{(2)} - (\overline{X}_n)^2\right] =$$

$$= E\left[\left(\frac{1}{n}\sum_{j=1}^{n} X_j^2\right)\right] - E\left[\left(\frac{1}{n}\sum_{j=1}^{n} X_j\right)^2\right] =$$

$$= \frac{1}{n}\sum_{j=1}^{n} EX_j^2 - \frac{1}{n^2} E\left(\sum_{j=1}^{n} X_j^2 + \sum_{i\neq j} E(X_iX_j)\right) =$$

$$EX_i\,EX_j = \mu\cdot\mu = \mu^2$$

$$= \frac{1}{n}\cdot n(\sigma^2 + \mu^2) - \frac{1}{n^2}\sum_{j=1}^{n} EX_j^2 - \frac{1}{n^2}\sum_{i\neq j} E(X_iX_j) =$$

$$= \sigma^2 + \mu^2 - \frac{1}{n^2} \cdot n(\sigma^2+\mu^2) - \frac{1}{n^2} 2 \cdot \binom{n}{2} \cdot \mu^2$$

$n(n-1)$

$\uparrow$ за x_i $\uparrow$ за x_i зной ки x_i

$$= \sigma^2 + \mu^2 - \frac{1}{n}(\sigma^2+\mu^2) - \frac{1}{n^2} n(n-1)\mu^2 =$$

$$= (\sigma^2+\mu^2)\left(1-\frac{1}{n}\right) - \left(1-\frac{1}{n}\right)\mu^2 =$$

$$= \left(1-\frac{1}{n}\right)\cdot\sigma^2 = \frac{n-1}{n}\sigma^2 \neq \sigma^2 \Rightarrow \text{оц. е изместена}$$

$$\underbrace{\frac{n}{n-1}\hat{\sigma}^2 = S^2}_{\text{коригираме оц.}} \Rightarrow \mathbb{E}S^2 = \frac{n}{n-1}\mathbb{E}\hat{\sigma}^2 = \frac{n}{n-1}\cdot\frac{n-1}{n}\sigma^2 = \sigma^2$$

$\Rightarrow$ оценката е неизместена

коригираме оц.
за да е неизместена

Пример:

$X \sim U(0,\theta)$ $\quad \hat{\theta} = \max(X_1,\dots,X_n) = Y$ $\quad F_X(x) = \frac{x}{\theta}$

$$\mathbb{P}(Y \le y) = \mathbb{P}(\max(X_1,\dots,X_n) \le y)$$

$$F_Y(y) = \mathbb{P}(X_1 \le y, \dots, X_n \le y)$$

$$= \mathbb{P}(X_1 \le y) \dots \mathbb{P}(X_n \le y)$$

$$= \left(\frac{y}{\theta}\right)^n \quad \text{за } y\,[0,\theta]$$

$$\boxed{f_Y(y) = \frac{ny^{n-1}}{\theta^n}}$$

$$\mathbb{E}\hat{\theta} = \mathbb{E}Y = \int_0^\theta \frac{y \cdot ny^{n-1}}{\theta^n}\,dy =$$

$$= \frac{n}{\theta^n}\left.\frac{y^{n+1}}{n+1}\right|_0^\theta = \frac{n}{\theta^n}\cdot\frac{\theta^{n+1}}{n+1} = \frac{n}{n+1}\theta \neq \theta$$

$\hookrightarrow$ не е неизместена

за ДР: $\boxed{\mathbb{E}\hat{\theta} \stackrel{?}{=} \theta}$

* **деф**: Казваме, че $\theta(\vec{x})$ е състоятелна (силно същ.), ако

$$\theta(\vec{x}) \xrightarrow[n\to\infty]{P} \theta \quad \left(\theta(\vec{x}) \xrightarrow[n\to\infty]{n.c.} (\theta) \right)$$

θ е оц. по метод на моментите (м.м)

$\overline{X_n} \xrightarrow[n\to\infty]{P} \mathbb{E}X$ слаб ЗГЧ

$\overrightarrow{X_n}^{(k)} \xrightarrow[n\to\infty]{P} \mathbb{E}X^k$, ако $\mathbb{E}|X|^k < \infty$ същ. (краен е)

$\boxed{Y = X^k}$

Ако $\mathbb{E}X, \mathbb{E}X^2, \dots, \mathbb{E}X^k, \dots,$ са известни, то това еднозначно ли определя $F_X(x)$?

* **Твърдение**: Нека $X \sim N(\mu, \sigma^2)$ и $\vec{x}$ е извадка над X. Тогава м.п.о. за μ е $\hat{\mu}(\vec{x}) = \overline{x_n} = \frac{1}{n}\sum_{j=1}^{n} x_j$ (независимо дали σ^2 е известно или не), а за σ^2 е:

i) $\hat{\sigma}^2 = \frac{1}{n}\sum_{j=1}^{n}(x_j - \mu)^2$, когато μ е известно

ii) $\hat{\sigma}^2 = \frac{1}{n}\sum_{j=1}^{n}(x_j - \overline{x_n})^2$, когато μ е неизвестно

д-во: $\theta = (\mu, \sigma^2)$

$$L = (\vec{x}, \theta) = \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_j-\mu)^2}{2\sigma^2}} =$$

$$= \left(\frac{1}{2\pi}\right)^{n/2} \left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}} e^{-\sum_{j=1}^{n}\frac{(x_j-\mu)^2}{2\sigma^2}}$$

$$\ln L(\vec{x}, \theta) = \frac{n}{2}\ln\left(\frac{1}{2\pi}\right) - \frac{n}{2}\ln(\sigma^2) - \sum_{j=1}^{n}\frac{(x_j-\mu)^2}{2\sigma^2}$$

$$\frac{\partial \ln L}{\partial \mu} = \frac{\sum_{j=1}^{n}(x_j - \mu)}{\sigma^2} = 0 \Rightarrow \sum_{j=1}^{n} x_j - n\mu = 0$$

$$\Rightarrow \hat{\mu} = \frac{1}{n}\sum_{j=1}^{n} x_j = \overline{x_n}$$

$$\frac{\partial \ln L}{\partial \sigma^2} = -\frac{n}{2}\,\frac{1}{\sigma^2} + \frac{\sum_{j=1}^{n} (x_j-\mu)^2}{2\sigma^4} = 0 \quad \Big|\cdot 2\sigma^4 \; ?$$

$$-na^2 + \sum_{j=1}^{n} (x_j-\mu)^2$$

$$\sigma^2 = \frac{1}{n} \sum_{j=1}^{n} (x_j-\mu)^2$$

i) $\hat{\sigma}^2 = \frac{1}{n} \sum_{j=1}^{n} (x_j-\mu)^2$

ii) $\hat{\sigma}^2 = \frac{1}{n} \sum_{j=1}^{n} (x_j-\overline{x_n})^2$

M

Доверителни интервали

$\hat{\Theta} = \hat{\Theta}(\vec{X}), \ \vec{X} = (X_1, \dots, X_n) \quad (\vec{x} = (x_1, \dots, x_n) \Rightarrow \hat{\theta} = \hat{\theta}(\vec{x}))$

$\theta \in \mathbb{R}$ и ще търсим $(I_1(\vec{X}), I_2(\vec{X}))$, такъв, че

$P(\theta \in (I_1, I_2)) \geq \gamma$

ниво на доверие

$(\gamma = 0{,}9;\ 0{,}95;\ 0{,}99)$

$I_1 := I_1(\vec{X})$

$I_2 := I_2(\vec{X})$

(I_1, I_2) се нарича доверителни интервал

$\vec{x}, \ (I_1(\vec{x}), I_2(\vec{x}))$

$\vec{y}, \ (I_1(\vec{y}), I_2(\vec{y}))$

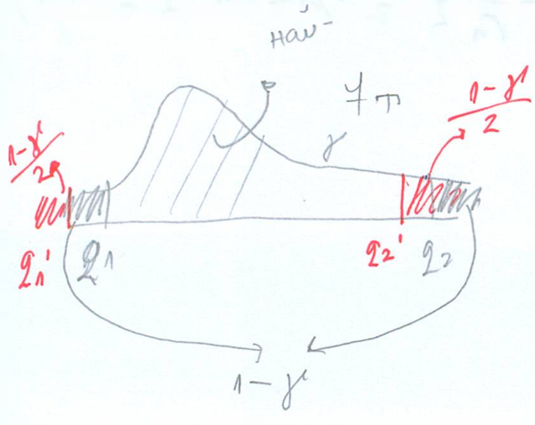
Централна статистика

Нека $T(\vec{X}, \theta)$ е статистика (функция). Деф: Казваме, че T е централна статистика, ако:

а) T е строго монотонна ($\uparrow \downarrow$) по θ

б) $P(T \leq x) = F_T(x)$ не зависи от θ

T е ф-я на θ, но разпределението ѝ



$P(q_1 < T < q_2) = F_T(q_2) - F_T(q_1) \geq \gamma$

$$= P(\underbrace{T^{-1}(q_1)}_{= I_1} < \theta < \underbrace{T^{-1}(q_2)}_{= I_2})$$

$\mathbb{P}(q_1 < T_1 < q_1) = \gamma$

Една задача е да се минимизира $q_2 - q_1$ или $I_2 - I_1$

① $X \sim \mathcal{N}(\mu, \sigma^2)$, където σ^2 е известно. $\theta = \mu$.

$\vec{X} = (X_1, \dots, X_n)$, $X_i \sim \mathcal{N}(\mu, \sigma^2)$, $1 \le i \le n$

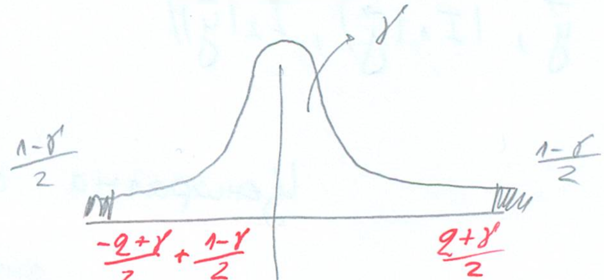
извадка, iid

$\overline{X_n} = \frac{1}{n} \sum_{j=1}^{n} X_j = \hat{\mu}$ ← независимо дали знаем σ^2 или не $\sim \hat{\mu} = \mathcal{N}(\mu, \frac{\sigma^2}{n})$

$\left(\sum_{i=1}^{n} X_i \sim \mathcal{N}(n\mu, n \cdot \sigma^2)\right)$

$T(\vec{X}, \mu) = \boxed{\dfrac{\overline{X_n} - \mu}{\sigma \sqrt{n}}} \sim \mathcal{N}(0, \frac{\sigma^2}{n}) \quad \sim \mathcal{N}(0,1)$

разделяме, защото искаме да е независимо от n



$\gamma = \mathbb{P}(q_1 < T < q_2) =$

$= \mathbb{P}\left(\frac{q_1 - \gamma}{2} < T < \frac{q_2 - \gamma}{2}\right) =$

$= \mathbb{P}\left(-\frac{q_1 - \gamma}{2} < T < \frac{q_1 + \gamma}{2}\right) = \mathbb{P}\left(-q_{\frac{1-\gamma}{2}} < \frac{\overline{X_n} - \mu}{\sigma \sqrt{n}} < q_{\frac{1-\gamma}{2}}\right) =$

$= \mathbb{P}\left(-q_{\frac{1+\gamma}{2}} \frac{\sigma}{\sqrt{n}} < \overline{X_n} - \mu < \frac{\sigma}{\sqrt{n}} q_{\frac{1+\gamma}{2}}\right) = \mathbb{P}\left(\overline{X_n} - \frac{\sigma}{\sqrt{n}} q_{\frac{1+\gamma}{2}} < \mu < \overline{X_n} + q_{\frac{1+\gamma}{2}} \frac{\sigma}{\sqrt{n}}\right)$

?

(1) $X \sim N(\mu, \sigma^2)$ $\quad \theta = \mu$ но нямаме информация за σ^2

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n) \text{ е неизместена оценка за } \sigma^2$$

хи-квадрат

$$s^2 \perp\!\!\!\perp \bar{X}_n \quad \text{и} \quad \frac{(n-1)s^2}{\sigma^2} = \sum_{i=1}^{n}\frac{(X_i - \bar{X}_n)^2}{\sigma^2} \sim \chi^2(n-1)$$

$$\bar{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

$$T(\vec{X}, \mu) = \frac{\bar{X}_n - \mu}{\frac{s}{\sqrt{n}}} = \frac{\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{s}{\sigma}}} = \frac{\overbrace{\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}}^{z}}{\sqrt{\frac{s^2}{\sigma^2}\cdot\frac{1}{n-1}}}$$

y

ако не целим за σ? справяли ли? И сите ще примери???

$Z \sim N(0,1)$

$Y \sim \chi^2(n-1)$

$$Z \perp\!\!\!\perp Y \qquad = \frac{Z}{\sqrt{\frac{Y}{n-1}}} \quad t \ldots$$

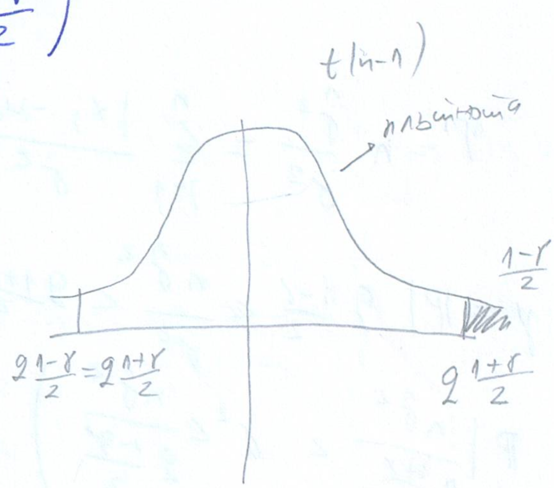
t разпределение

$$\gamma = \mathbb{P}\left(-q_{\frac{1+\gamma}{2}} < \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} < q_{\frac{1+\gamma}{2}}\right) =$$

$$= \mathbb{P}\left(\bar{X}_n - \frac{s}{\sqrt{n}}q_{\frac{1+\gamma}{2}} < \mu < \bar{X}_n + \frac{s}{\sqrt{n}}q_{\frac{1+\gamma}{2}}\right)$$

$$I_1 = \bar{X}_n - \frac{s}{\sqrt{n}}q_{\frac{1+\gamma}{2}}$$

$$I_2 = \bar{X}_n + \frac{s}{\sqrt{n}}q_{\frac{1+\gamma}{2}}$$



Приближаваме $t(n)$ с Z кога?

стандартно нормално

↳ когато $n-1 \geq 30$

↳ не работим с квантилите на t, а с тези на Z

- защо заместваме t със t

$$T = \frac{Z}{\sqrt{\frac{Y}{n-1}}}$$ — този знаменател приблизително изчезва

$$\frac{Y}{n-1}, \quad Y \sim \chi^2(n-1)$$

$$\frac{\sum_{j=1}^{n-1} Y_j}{n-1} \quad Y_j \sim \chi^2(1)$$

$$\downarrow n \to \infty$$

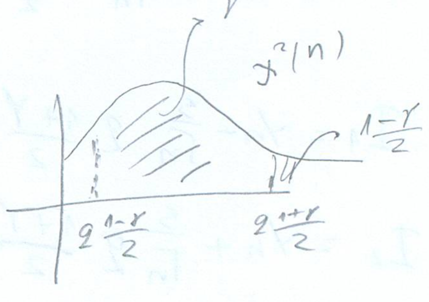
$$E Y_1 = 1$$ — и затова знаменателя не се вижда

Доверителен инт. за σ^2: $X \sim N(\mu, \sigma^2)$

а) μ → се знае, $\vec{X}$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{j=1}^{n}(x_j - \mu)^2 = \frac{1}{n}\underbrace{\sum_{j=1}^{n}\left(\frac{x_j-\mu}{\sigma}\right)^2}_{\chi^2(n)}$$

$\chi^2(n)$ → хи квадрат със степен n



$$T = n\frac{\hat{\sigma}^2}{\sigma^2} = \sum_{j=1}^{n}\frac{(x_j-\mu)^2}{\sigma^2} = \sigma^2(n)$$

$$\gamma = P\left(q_{\frac{1-\gamma}{2}} < \frac{n\hat{\sigma}^2}{\sigma^2} < q_{\frac{1+\gamma}{2}}\right) =$$

$$P\left(\frac{n\hat{\sigma}^2}{q_{\frac{1+\gamma}{2}}} < \sigma^2 < \frac{n\hat{\sigma}^2}{q_{\frac{1-\gamma}{2}}}\right)$$

$$I_1 = n \cdot \frac{1}{n}\sum_{j=1}^{n}(x_j-\mu)^2 = ?$$

$$I_2 = \frac{\sum_{j=1}^{n}(x_j-\mu)^2}{q_{\frac{1-\gamma}{2}}}$$

→ ако μ не се знае

$$I_1 = \frac{\sum_{i=1}^{n} (x_i - \overline{x}_n)^2}{q_{\frac{1+\gamma}{2}}}$$

$$I_2 = \frac{\sum_{i=1}^{n} (x_i - \mu)^2}{q_{\frac{1-\gamma}{2}}}$$

Проверка на хипотези

Пример: има странични ефекти, да ганем

Матем. постановка: X, $F_X(x, \theta)$; $\vec{X}$ – извадка. $\vec{X} = (X_1, \dots, X_n)$

$H_0 : \theta = \theta_0$ → нулева / базова хипотеза — проста срещу проста

$H_1 : \theta = \theta_1$ → алтернативна хипотеза

Цел: Търсим $W \subseteq \mathbb{R}^n$ (критична област) такава, че ако $\vec{X} \in W$ отхвърляме H_0, а ако $\vec{X} \in W^c$ приемаме H_0.

получаваме $\vec{X}$ и ако $\vec{x} \in W$ отхвърляме H_0

$\vec{x} \in W^c$ приемаме H_0

$\alpha = P(\vec{X} \in W / H_0) = \alpha_W$ → грешка от първи род

$\beta = P(\vec{X} \notin W^c / H_1) = \beta_W$ → грешка от втори род

Избираме $\alpha \in \{0{,}1;\ 0{,}05;\ 0{,}01\} \Rightarrow$ избираме $W: \alpha_W = \alpha$

$\beta_W = P(\vec{X} \in W^c \mid H_1)$ → допускаме че ваксината да е опасна, а всъщност не е

(не е толкова голяма драма, но обратното е!)

* Деф: Казваме, че $W^* \in \mathbb{R}^n$ е оптимална критична област (ОКО) за грешка от първи род α, ако

$$\beta_{W^*} = \inf_{\substack{\alpha_W = \alpha \\ \forall_W \subseteq \mathbb{R}^n}} \beta_W$$ → това β_{W^*}, което има оптимална (минимална) грешка

* Лема на Пирсън (Нейман):

$X \sim F_X(x, \theta)$ и плътност $f_X(x, \theta)$

$H_0: \quad \theta = \theta_0$

$H_1: \quad \theta = \theta_1$

$$L(\vec{X}, \theta) = \prod_{j=1}^{n} f(x_j, \theta)$$

$$L_1(\vec{X}) = L(\vec{X}, \theta_1)$$

$$L_0(\vec{X}) = L(\vec{X}, \theta_0)$$

$\vec{X} \quad L_1(x$

Лема:

Нека $X \sim f_X(x, \theta)$, $\vec{X} = (X_1, \dots, X_n)$ и разглеждаме

$H_0: \theta = \theta_0$

$H_1: \theta = \theta_1$. Тогава ако за зададена грешка от първи

род α е вярно, че $\alpha_W = \mathbb{P}(\vec{X} \in W \mid H_0) = \alpha$ и

$W^* = \{ \vec{x} \in \mathbb{R}^n : L_1(\vec{x}) \geq k L_0(\vec{x}) \}$

се влага α

$(W^*)^c = \{ \vec{x} \in \mathbb{R}^n : L_0(\vec{x}) \leq k L_0(\vec{x}) \}$, където $k > 0$ е константа, която може да зависи от ?

Тогава W^* е о.к.о

(4) $\alpha = 0{,}05$

$W_k = \{ \vec{x} \in \mathbb{R}^n : L_1(\vec{x}) \geq k L_0(\vec{x}) \}$

Ако намерим k^* : $\boxed{\alpha_{W_{k^*}} = \alpha = 0{,}05}$

$W^* = W_{k^*} = \{ \vec{x} \in \mathbb{R}^n : L_1(\vec{x}) \geq k^* L_0(\vec{x}) \}$

$\mathbb{P}(L_1(\vec{X}) \geq k^* L_0(\vec{X}) \mid H_0) = 0{,}05 \overset{\text{лема}}{\Longrightarrow} W^*$ е о.к.о g-60:

$$\alpha = P(\vec{x} \in w^* \mid H_0) = \int_{w^*} L_0(\vec{x})\, d\vec{x} = \alpha_w = P(\vec{x} \in w \mid H_0) =$$

$$\int_{w} L_0(\vec{x})\, d\vec{x}$$

$$\int_{w} L_0(\vec{x})\, d\vec{x} \qquad = \int_{w^* \cap w} L_0(\vec{x})\, d\vec{x} + \int_{w^* \cap w^c} L_0(\vec{x})\, d\vec{x} =$$

$$= \int_{w} L_0(\vec{x})\, d\vec{x} = \int_{w \cap w^*} L_0(\vec{x})\, d\vec{x} + \int_{w \cap w_*^c} L_0(\vec{x})\, d\vec{x} =$$

$$= \int_{w^* \cap w^c} L_0(\vec{x})\, d\vec{x} = \int_{w \cap w_x^c} L_0(\vec{x})\, d\vec{x} \qquad \beta_w \geq \beta_{w^*}$$

$$\beta_w = P(\vec{x} \in w^c \mid H_1) = \underbrace{\int_{w^c} L_1(\vec{x})\, d\vec{x} - \beta_{w^*}}_{\overset{?}{\geq 0}} + \beta_{w^*} \geq \beta_{w^*}$$

$$\int_{w^c} L_1(\vec{x})\, d\vec{x} = \int_{(w^*)^c} L_1(\vec{x})\, dx = \int_{w^* \cap w^c} L_1(\vec{x})\, d\vec{x} + \cancel{\int_{w^c \cap (w^*)^c} L_1(\vec{x})\, d\vec{x}} - \cancel{\int_{(w^*)^c \cap w^c} L_1(\vec{x})\, d\vec{x}}$$

$$- \int_{(w^*)^c \cap w} L_1(\vec{x})\, d\vec{x} = \int_{w^c \cap w^*} L_1(\vec{x})\, d\vec{x} - \int_{(w^*)^c \cap w} L_1(\vec{x})\, d\vec{x} \geq \int_{w^c \cap w^*} k L_0(\vec{x})\, d\vec{x} -$$

$$\int_{(w^*)^c \cap w} k L_0(\vec{x})\, d\vec{x} \overset{(A)}{=} 0$$

④ $X \sim N(\mu, \sigma^2)$, σ^2 е известна $\vec{X} = (X_1, \dots, X_n)$ и грешка от първи род α

$H_0: \mu = \mu_0$

$H_1: \mu = \mu_1$ $(\mu_0 < \mu_1)$

$$L_0(\vec{x}) = \frac{1}{(\sqrt{2\pi})^n} \cdot \prod_{j=1}^{n} e^{-\frac{(x_j-\mu_0)^2}{2\sigma^2}}$$

$$L_1(\vec{x}) = \frac{1}{(\sqrt{2\pi})^n} \cdot \prod_{j=1}^{n} e^{-\frac{(x_j-\mu_1)^2}{2\sigma^2}}$$

$$W_k = \{\vec{x} \in \mathbb{R} : L_1(\vec{x}) \geq k L_0(\vec{x})\} = \{\vec{x} \in \mathbb{R}^n : e^{-\sum_{j=1}^{n} \frac{(x_j-\mu_1)^2}{2\sigma^2}} \geq k e^{-\sum_{j=1}^{n} \frac{(x_j-\mu_0)^2}{2\sigma^2}} =$$

$$\alpha = \mathbb{P}(\vec{X} \in W_k \mid H_0) = \mathbb{P}\left(\frac{\sum_{j=1}^{n} X_j}{n} \geq k_2 \mid H_0\right) =$$

$$\cancel{\{\vec{x} \in \mathbb{R}^n : -\sum_{j=1}^{n} \frac{(x_j-\mu_1)^2}{2\sigma^2}}$$

$$= \{\vec{x} \in \mathbb{R}^n : -\sum_{j=1}^{n} (x_j-\mu_1)^2 \geq 2\sigma^2 \ln k - \Sigma$$

$X_0 \sim N(\mu_0, \sigma^2)$

$$= \mathbb{P}\left(\frac{\sum_{j=1}^{n} X_j - n\mu_0}{\sigma\sqrt{n}} \geq \sqrt{n}\, \frac{k_2 - n_0}{\sigma} \,\middle|\, H_0\right) = \mathbb{P}\left(Z \geq \frac{\sqrt{n}}{\sigma}(k_2 - n\mu_0)\right) =$$

$$= 1 - \Phi\left(\frac{\sqrt{n}}{\sigma}(k_2 - n\mu_0)\right) = \alpha \iff \Phi\left(\frac{\sqrt{n}}{\sigma}(k_2 - n\mu_0)\right) \Rightarrow$$

$$q_{1-\alpha} = \frac{\sqrt{n}}{\gamma}\left(K_1 - n\mu_0\right)$$

Лекция КН 06.06.2024

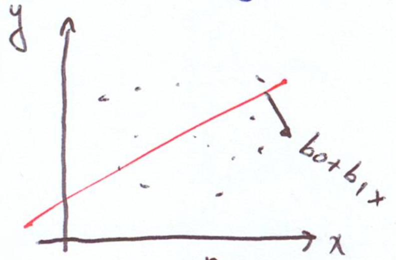
Линейна регресия

Галтон - искал да направи изследване дали има разлика между ръст на баща и син

$$y_{син} = x_{средно} + \beta(x_{баща} - x_{средно}), \quad \beta \approx 0{,}6$$

$$x_{баща} - x_{средно} = 10cm \Rightarrow y_{син} = x_{средно} + 6cm$$

→ регрес - нещо което ни връща обратно към средното



$(x_i, y_i)_{i=1}^{n}$

$$y_k = b_0 + b_1 x_k + \varepsilon_k$$

$$\hat{y} = \hat{b}_0 - \hat{b}_1 x$$

$$f(b_0, b_1) = \sum_{k=1}^{n} (y_k - \hat{y}_k)^2 = \sum_{k=1}^{n} (y_k - \hat{b}_0 - \hat{b}_1 x_k)^2$$

→ минимизираме по всички прави (b_0, b_1 ..)

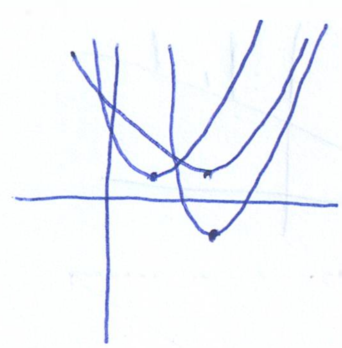
→ цел $\hat{b}_0, \hat{b}_1$: $\sum_{k=1}^{n} (y_k - \underbrace{\cancel{\hat{y}_k}}_{\hat{b}_0 - \hat{b}_1 x_k})^2 = \inf_{b_0, b_1} \sum_{k=1}^{n} (y_k - b_0 - b_1 x_k)^2$

→ защо се взима на квадрат?

$$0 = \frac{\partial f}{\partial b_0} = -2 \sum_{k=1}^{n} (y_k - b_0 - b_1 x_k)$$

$$0 = \frac{\partial f}{\partial b_1} = -2 \sum_{k=1}^{n} x_k (y_k - b_0 - b_1 x_k)$$

(=)



→ парабола, която има глобален минимум

$\textcircled{=} = n\bar{y} - nb_0 - b_1 n\bar{x}$

$\rightarrow \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

$\bar{y} = \frac{1}{n}\sum_{i=1}^{n}$

$= \sum_{k=1}^{n} x_k y_k \cancel{x} b_0 \bar{x} - b_1 \sum_{k=1}^{n} x_k^2$

$\hat{b}_1^* = \frac{\sum_{k=1}^{n}(x_k-\bar{x})y_k}{\sum_{k=1}^{n}(x_k-\bar{x})^2}$

$\bar{y} = b_0 + b_1\bar{x} \Rightarrow b_0 = \bar{y} - b_1\bar{x}$

$b_0 = \bar{y} - b_1\bar{x} = \sum_{k=1}^{n} y_k\left(\frac{1}{n} - \frac{\bar{x}(x_k-\bar{x})}{\sum_{k=1}^{n}(x_k-\bar{x})^2}\right)$

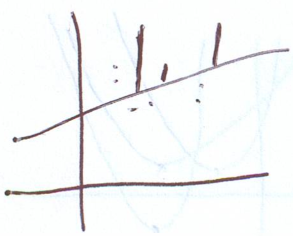
$\sum_{k=1}^{n} x_k y_k - b_1 \sum_{k=1}^{n} x_k^2 = n b_0 \bar{x} = n(\bar{y} - b_1\bar{x})\bar{x} \Rightarrow$

$\sum_{k=1}^{n} x_k y_k - n\bar{y}\bar{x} = b_1\left(\sum_{k=1}^{n} x_k^2 - n\bar{x}^2\right)$

$b_1 = \frac{\sum_{k=1}^{n} x_k y_k - n\bar{y}\bar{x}}{\sum_{k=1}^{n} x_k^2 - n\bar{x}^2} = \frac{\sum_{k=1}^{n}(x_k-\bar{x})y_k}{\sum_{k=1}^{n}(x_k-\bar{x})^2}$

$\sum_{k=1}^{n} x_k y_k - \bar{x}\sum_{k=1}^{n} y_k = \sum_{k=1}^{n}(x_k-\bar{x})y_k$

$b_0 = \bar{y} - b_1\bar{x} = \frac{1}{n}\sum_{k=1}^{n} y_k - \sum_{k=1}^{n}\frac{(x_k-\bar{x})\bar{x}}{\sum_{k=1}^{n}(x_k-\bar{x})^2} y_k =$



$$Y_k = \beta_0 + \beta_1 X_k + \mathcal{E}_k \to \text{грешка (стохастична)}$$ * стохастично, действителна

модел

предиктор / доза $\quad 1 \le k \le n$

отклик / резултат

↳ стохастични грешки са нез. в съвкупност

* **Допускания:**

1) $(\mathcal{E}_1, \dots, \mathcal{E}_n)$ са независими в съвкупност

2) $\mathbb{E}\mathcal{E}_k = 0, \ \forall 1 \le k \le n \quad (\Rightarrow \mathbb{E}Y_k = \beta_0 + \beta_1 x_k)$

среден отклик

↳ т.е. средно сме оценили добре

3) $D\mathcal{E}_k = \sigma^2, \ 1 \le k \le n,$ хомоскеластичност

4) $\mathcal{E}_k \sim N(0, \sigma^2), \ 1 \le k \le n$

↳ $DY_k = D\mathcal{E}_k = \sigma^2$

$$\hat{\beta}_1 = \sum_{k=1}^{n} \frac{(x_k - \bar{x}) Y_k}{\sum_{k=1}^{n} (x_k - \bar{x})^2}$$

$$\hat{Y}_k = \hat{\beta}_0 + \hat{\beta}_1 x_k \ ; \ Y_k = \beta_0 + \beta_1 x_k + \mathcal{E}_k$$

оценка

$$\hat{\beta}_0 = \sum_{k=1}^{n} Y_k \left(\frac{1}{n} - \frac{\bar{x}(x_k - \bar{x})}{\sum_{k=1}^{n} (x_k - \bar{x})^2} \right)$$

↳ **Неизместеност:**

$$\underset{\overset{\parallel ?}{\beta_1}}{\mathbb{E}\hat{\beta}_1} = \mathbb{E}\left[\sum_{k=1}^{n} \frac{(x_k - \bar{x})}{\sum_{k=1}^{n} (x_k - \bar{x})^2} Y_k\right] = \sum_{k=1}^{n} \frac{(x_k - \bar{x})}{\sum_{k=1}^{n} (x_k - \bar{x})^2} \cdot \underbrace{(\beta_0 + \beta_1 x_k)}_{\mathbb{E}Y_k} =$$

$$= \beta_0 \frac{\sum_{k=1}^{n} (x_k - \bar{x})}{\sum_{k=1}^{n} (x_k - \bar{x})^2} + \beta_1 \sum_{k=1}^{n} \frac{(x_k - \bar{x}) x_k - \bar{x}}{\sum_{k=1}^{n} (x_k - \bar{x})^2} = \beta_2$$

$\sum_{k=1}^{n}(x_k - \bar{x}) \overset{\parallel}{=} 0$ ← центъра на тежестта

$$* \ \sum \bar{x}(\bar{x} - x_k) = 0$$

$\rightarrow \hat{\sigma}^2 = \frac{1}{n-2} \sum_{k=1}^{n} (Y_k - \hat{\beta}_0 - \hat{\beta}_1 X_k)^2 = \frac{\sigma^2 \sum_{k=1}^{n} \frac{(Y_k - \hat{\beta}_0 - \hat{\beta}_1 X_k)^2}{\sigma^2}}{n-2}$

$$\sum_{k=1}^{n} \frac{(Y_k - \hat{\beta}_0 - \hat{\beta}_1 X_k)^2}{\sigma^2} = \sum_{k=1}^{n} V_k, \quad V_k \sim \chi^2(1)$$

$$\stackrel{d}{=} \frac{\sigma^2 \sum_{k=1}^{n} V_k}{n-2} \xrightarrow[n \to \infty]{3\Gamma 4} \sigma^2$$

Теорема: Нека $Z_1, \ldots, Z_n$ са стандартни нормални разпределения и независими в съвкупност.

Тогава $Y = \sum_{j=1}^{n} |Z_j|^2 \sim \chi^2(n)$

Мотивация:

$$H_j = \text{детерминистично} + Z_j$$

$$(H_j - \text{детерминистично})^2 = Z_j^2$$

Що от тая година го правим само за 2, 1 - ??

→ Точкови оценки

последен запис

Постановка: X описва някакво свойство на някаква генерална съвкупност, X, $F_X(x) = P(X<x)$, $f_X(x)$

Идея: $\vec{X} = (X_1, \dots, X_n)$, $(X_j)_{j=1}^{n}$ са н.е.р. сл. вел.

$X_i \overset{d}{=} X$, На базата на $\vec{X}$ искаме да определим/приближим F_X, f_X

→ $\hat{\theta}$ е функция на наблюденията
↳ е точкова оценка за θ

→ трябва да изберане такава, така че да максимизира нещо

→ функция на максималното правдоподобие, наричаме изв. съвместната плътност на $\vec{X}$, т.е.

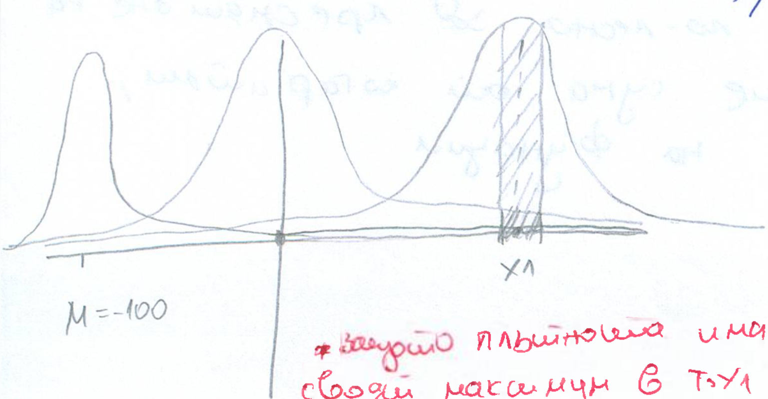
$$L(\vec{X}, \theta) = \prod_{j=1}^{n} f_X(x_j, \theta)$$

→ МПО

$X \sim N(\mu, 1)$; $\vec{X} = X_1$; $X_1 = x_1$

$$L(\vec{x}; \mu) = f_X(x_1, \mu)$$

$$L(x_1; \mu) = f_X(x_1; \mu) \xrightarrow{\text{избираме}} \hat{\mu} = x_1 \;\; \to \text{оценката}$$



*виждаме плътността има своя максимум в т. x_1

→ ако изберем μ да е някаква μ, вероятността да се падне x_1 е много малка

→ ако вземем да е в 0, малко по-добре е, но пак е малка

→ ако вземем $\mu = x_1$, вероятността да се падне точно около x_1 е много по-голяма

* Взимам този вектор на θ, който ни максимизира функцията на максимално правдоподобие

* **Деф:** Нека X е сл. вел. с плътност $f_X(x,\theta)$,
Нека $\vec{X}$ са n независими наблюдения над X.
Тогава под МПО за θ разбираме $\hat{\theta}$:

$$L(\vec{X},\hat{\theta}) = \sup_{\theta\in\Theta} L(\vec{X},\theta)$$

→ т.е. избираме този параметър (набор от параметри), който максимизира по всички възможни стойности за θ, нашата функция на правдоподобие (или ни дава максимална вероятност да видим нашата реализация)

→ Ако f_X е диференцируема по θ, то решаваме $\frac{\partial L}{\partial \theta} = 0$, за да намерим екстремума $\hat{\theta}$

→ понеже L е произведение

$L(\vec{X},\theta) = \prod_{j=1}^{n} f_X(x_j,\theta)$, се взима логаритъм,

защото произведението става сума и понеже логаритъмът е растяща функция, еквивалентно на решаването на $\frac{\partial L}{\partial \theta} = 0$ е

$$\frac{\partial \ln L}{\partial \theta} = 0$$

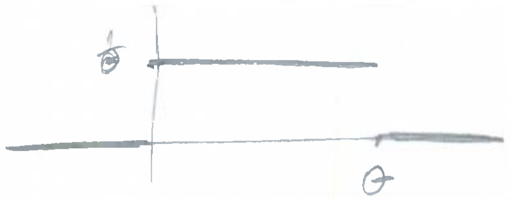
* Когато вземем логаритъм от произведението, ставаме в сума, което е доста по-лесно за пресмятане на производни, когато имаме сума от логаритми, отколкото произведение на функции

(4) $X \sim U(0,\theta)$, $\theta > 0$ $\qquad f_X(x) = \frac{1}{\theta} \cdot 1_{[0,\theta]}(x)$

$\vec{X}$

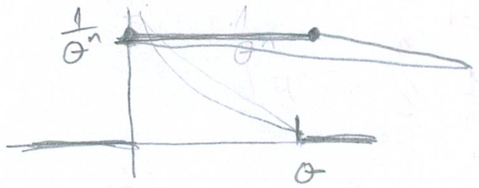
$$L(\vec{x},\theta) = \prod_{j=1}^{n} \frac{1}{\theta} \cdot 1_{[0,\theta]}(x_j) =$$

$$= \frac{1}{\theta^n} \cdot 1_{[0,\theta]}(x^*) = \begin{cases} \frac{1}{\theta^n}, & x^* \le \theta \\ 0, & x^* > \theta \end{cases}$$



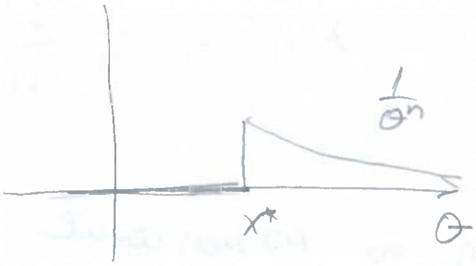
тази ф-я не е диференцируема, тя е индикаторна, приема стойност в диапазона $0, \theta$ и 0 извън диапазона

⇒ графиката е



не е диференцируема в тези точки

$\frac{1}{\theta^n}$ е намаляваща по θ



$$L(\vec{x},\theta) = \sup_{\theta > 0} L(\vec{x},\theta)$$

$$\hat{\theta} = \max_{j \le n}(x_j), \quad \hat{\theta} = \hat{\theta}(\vec{x})$$

⇒ $\hat{\theta}$ е максимума на всички наблюдения

+ Твърдение за МПО за μ и σ^2

g-во: $L(\vec{x},\theta) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \sigma^n e^{-\sum_{j=1}^{n} \frac{(x_j-\mu)^2}{2\sigma^2}}$ $\theta = (\mu, \sigma^2)$

$$\ln L(\vec{x},\theta) = n \ln \frac{1}{\sqrt{2\pi}} - \frac{n}{2}\ln\sigma^2 - \sum_{j=1}^{n} \frac{(x_j-\mu)^2}{2\sigma^2}$$

за да може да сметнем $\frac{\partial \ln L}{\partial \sigma^2}$ =>

$$\left| \begin{array}{l} \frac{\partial \ln L}{\partial \mu} = \cancel{\frac{1}{\sigma^2}} \sum_{j=1}^{n} (x_j-\mu) = 0 \\ \frac{\partial \ln L}{\partial \sigma^2} = \frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{j=1}^{n} (x_j-\mu)^2 = 0 \end{array} \right.$$

$$=> \left| \begin{array}{l} \hat{\mu} = \frac{1}{n} \sum_{j=1}^{n} x_j = \overline{x_n}^{(1)} \\ \hat{\sigma}^2 = \frac{1}{n} \sum_{j=1}^{n} (x_j-\mu)^2 \end{array} \right.$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{j=1}^{n} (x_j-\overline{x}_n)^2$$

→ Метод на моментите

$$\mathbb{E}X = \mu^{(1)}(\theta) \quad ; \quad \theta = (\mu^{(1)})^{-1}(\mathbb{E}X)$$

$$\overline{X_n}^{(1)} = \frac{\sum_{j=1}^{n} x_j}{n} \xrightarrow[n\to\infty]{п.с.} \mathbb{E}X \quad (УЗГЧ)$$

$$\mu^{(1)}(\hat{\theta}) = \overline{X_n}^{(1)}, \qquad \hat{\theta} = (\mu^{(1)})^{-1}(\overline{X_n}^{(1)})$$

$$\downarrow \qquad \downarrow$$

$$\theta = (\mu^{(1)})^{-1}(\mathbb{E}X)$$

$X \sim U(0, \theta)$, $\vec{X}$ м.м.п. $\hat{\theta} = \max\limits_{j \le n} X_j$

↳ максимумо от наблюденията

$$\mathbb{E}X = \mu^{(1)}(\theta) = \frac{\theta}{2} \Rightarrow \theta = 2\mathbb{E}X$$

$$\overline{X}_n^{(1)} = \frac{\theta}{2} \Rightarrow \hat{\theta} = 2\overline{X}_n^{(1)} \rightarrow \text{ММО}$$

средното аритметично

$X \sim N(\mu, \sigma^2)$ $\quad \mathbb{E}X = \mu \quad \mathbb{E}X^2 = DX + (\mathbb{E}X)^2 = \sigma^2 + \mu^2$

$\overline{X}_n^{(1)} = \hat{\mu}$, ММО = ММП $\quad * \overline{X}_n^{(2)} = \frac{1}{n}\sum\limits_{j=1}^{n} X_j^2$

$$\overline{X}_n^{(2)} = \mu^2 + \sigma^2 \Rightarrow \hat{\sigma}^2 = \overline{X}_n^{(2)} - (\overline{X}_n^{(1)})^2 = \frac{1}{n}\sum_{j=1}^{n}(X_j - \overline{X}_n^{(1)})^2$$

ММП = ММО

↳ В случай на нормално разпределение и двете оценки съвпадат

* Свойства

— неизместеност

↳ че очакваната стойност на оценката трябва да съвпада с истинската

↳ разликата се нарича систематична грешка $|\mathbb{E}\hat{\theta} - \theta|$

$X \sim N(\mu, \sigma^2)$, $\hat{\mu} = \frac{1}{n}\sum\limits_{j=1}^{n} X_j$ $\quad \mathbb{E}\hat{\mu} = \frac{1}{n}\sum\limits_{j=1}^{n}\mathbb{E}X_j = \frac{n\mu}{n} = \mu$

Ако знаем μ, $\mathbb{E}\hat{\sigma}^2 = \frac{1}{n}\mathbb{E}\sum\limits_{j=1}^{n}(X_j - \mu)^2 = \frac{1}{n}\sum\limits_{j=1}^{n}\mathbb{E}(X_j - \mu)^2 = \frac{n\sigma^2}{n} = \sigma^2$

това е точно дисперсията

ако не знаем μ, то $\hat{\sigma}^2 = \frac{1}{n}\sum_{j=1}^{n}(X_j - \overline{X}_n^{(n)})^2$

$$\mathbb{E}\hat{\sigma}^2 = \frac{1}{n}\sum_{j=1}^{n}\mathbb{E}(X_j - \overline{X}_n^{(n)})^2 = \frac{1}{n}\sum_{j=1}^{n}\mathbb{E}(X_j^2 - 2X_j\overline{X}_n^{(n)} + (\overline{X}_n^{(n)})^2) =$$

$$= \frac{1}{n}\sum_{j=1}^{n}\mathbb{E}X_j^2 - \frac{2}{n} \quad \dots$$ няма смисъл :)

$$\Rightarrow \mathbb{E}\hat{\sigma}^2 = \left(\frac{n-1}{n}\right)\sigma^2 \neq \sigma^2$$ → т.е. оценката е изместена

и за да е неизместена взимаме $\frac{1}{n-1}$, при голямо n няма значение а ни помага за оценката

– асимптотично неизместване

д) състоятелност на точкова оценка

$\forall \theta \in \mathbb{R} \quad \mathbb{P}(|\hat{\theta}(\vec{x}) - \theta| > \varepsilon) \xrightarrow[n\to\infty]{} 0$

силна състоятелност

$X \sim N(\mu, \sigma^2)$ $\quad \hat{\mu} = \overline{X}_n^{(n)} \xrightarrow[n\to\infty]{n.c} \mu$ (УЗГЧ)

$\hat{\mu}$ освен неизместена е и силно състоятелна

D! с увеличаване на извадката, напето средно ще се схожда към μ

$X \sim U(0, \theta)$ $\quad \hat{\theta}_1 = \max_{j\le n}(X_j) = X^*$; $\hat{\theta}_2 = 2\overline{X}_n^{(n)}$

$$\mathbb{E}\hat{\theta}_2 = 2\mathbb{E}\overline{X}_n^{(n)} = \frac{2}{n}\sum_{j=1}^{n}\mathbb{E}X_j = \frac{2}{n}\cdot n\underbrace{\frac{\theta}{2}}_{\mathbb{E}X_j} = \theta$$

$\hat{\theta}_2 = 2\overline{X}_n^{(n)} \xrightarrow[n\to\infty]{n.c} 2\cdot\frac{\theta}{2} = \theta$ ⇒ силно състоятелна независимост

$$\to F_{X^*}(x) = \mathbb{P}(X^* < x) = \mathbb{P}\left(\bigcap_{j=1}^{n} \{X_j < x\}\right) = \prod_{j=1}^{n} \mathbb{P}(X_j < x) \;\circled{=}$$

за максимума да е по-малък, трябва те всички да са по-малки от x и понеже са независими, това е произведение

еднакви

$$\circled{=} \; \mathbb{P}(X_1 < x)^n = \left(\frac{x}{\theta}\right)^n, \quad x \in (0, \theta)$$

$$f_X(x) = F_X(x)' = n\frac{x^{n-1}}{\theta^n}, \quad x \in (0, \theta)$$

$$\mathbb{E}X^* = \frac{n}{\theta^n}\int_0^\theta x \cdot x^{n-1}\,dx = \frac{n}{\theta^n(n+1)}\theta^{n+1} = \frac{n}{n+1}\theta$$

≠ изместена оценка

$\to \frac{n+1}{n}\hat{\theta}_n$ е неизместена оценка

4) $X \sim N(\mu, 1)$: $\vec{X} = x_1$

↳ Имаме сл. величина X, която следва нормално разпределение с неизвестна средна стойност μ и известна дисперсия $\sigma^2 = 1$. Имаме и едно наблюдение $X_1 = x_1$ и искаме да намерим оценката на μ по М.М.П.

1. Определяме функцията на правдоподобие

$$f_X(x, \mu) = \frac{1}{\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2}}$$

↳ за едно наблюдение $X_1 = x_1$

$$L(x_1, \mu) = f_X(x_1, \mu) = \frac{1}{\sqrt{2\pi}} e^{\frac{-(x_1-\mu)^2}{2}}$$

2. Намиране на логаритмичната правдоподобност

$$\ln L(x_1, \mu) = \ln\left(\frac{1}{\sqrt{2\pi}}\right) - \frac{(x_1-\mu)^2}{2}$$

↳ $\ln\left(\frac{1}{\sqrt{2\pi}}\right)$ не зависи от μ => може да се игнорира

$$\ln L(x_1, \mu) = -\frac{(x_1-\mu)^2}{2}$$

3. Максимизиране на лог-правдоподобността:

$$\frac{d \ln L}{d\mu} = \frac{d}{d\mu}\left(-\frac{(x_1-\mu)^2}{2}\right) = x_1 - \mu$$

4. Приравняване производната на нула

$$x_1 - \mu = 0 \Rightarrow \mu = x_1 \Rightarrow \hat{\mu} = x_1$$

↳ Това е логично, защото при едно наблюдение най-правдоподобната стойност за средната стойност на нормално разпределение е самото наблюдение.

$$\frac{d}{d\mu}\left(-\frac{(x_1-\mu)^2}{2}\right)$$ ~~-р/ Цего от тегия $\mu = x_1 - M$~~

~~$-1 \frac{d}{d\mu}\left(-\frac{\mu^2}{2}\right) =$~~

$$\Rightarrow \frac{d}{d\mu}\left(-\frac{(x_1^2 - 2x_1\mu + \mu^2)}{2}\right) = -\frac{1}{2}\cdot(0 + 2\cancel{x_1} + 2\mu)$$

това по M е 0 това е $2x$, щы като положиме $M = 1$

$$\Rightarrow \frac{d}{d\mu}\left(-\frac{(x_1-\mu)^2}{2}\right) = -\frac{1}{2}(-2\cancel{x_1} + 2\mu) = -\frac{\cancel{2}}{\cancel{2}}(-x_1+\mu) = \underline{(x_1 - \mu)}$$

→ Кво решение

$$L(\vec{x}, \mu) = f_X(x_1, \mu)$$

$$L(x_1, \mu) = f_X(x_1, \mu) \Longrightarrow \hat{\mu} = x_1$$ + картинката

1 $X \sim U(0, \theta)$, $\theta > 0$, $\Theta = [0, +\infty)$

$\vec{X} = (X_1, X_2, \dots, X_n)$

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & a \le x \le \\ 0, & \text{иначе} \end{cases}$$

↳ цел: да намерим м.п.о за θ

→ функцията на правдоподобие

$$f_X(x, \theta) = \begin{cases} \frac{1}{\theta}, & \text{ако } 0 \le x \le \theta \\ 0, & \text{иначе} \end{cases}$$

$$\Rightarrow L(\vec{x}, \theta) = \prod_{j=1}^{n} f_X(x_j, \theta) = \prod_{j=1}^{n} \frac{1}{\theta} \cdot \mathbb{1}\{x_j \in [0, \theta)\} =$$

$$= \frac{1}{\theta^n} \cdot \mathbb{1}\{\max(x_1, \dots, x_n) \in [0, \theta)\}$$

↳ или

$$L(\vec{x}, \theta) = \prod_{j=1}^{n} f_X(x_j, \theta) = \begin{cases} \frac{1}{\theta}, & 0 \le x \le \theta \\ 0, & \text{иначе} \end{cases}$$

⇒ функцията на правдоподобието е ненулево само ако $\theta \ge \max(x_1, \dots, x_n)$

$\Rightarrow L(\vec{x}, \theta) = \frac{1}{\theta^n}$, за $\theta \ge X_{(n)}$ и 0 иначе

→ трябва да максимизираме по θ, но понеже $\frac{1}{\theta^n}$ е намаляваща по θ, тя ще бъде най-голяма когато θ е възможно най-малка, но удовлетворява $\theta \ge X_{(n)}$

→ $\hat{\theta} = X_{(n)}$ → ~~м.п.о~~ м.п.о за θ е макс. наблюдавана стойност от наблюденията

$$\hat{\theta} = \max(X_1, X_2, \dots, X_n) = X_{(n)}$$

М.М.О

1) $X \sim U(0, \theta)$, $\vec{x} = (x_1, x_2, \dots, x_n)$

→ намираме теоретичните моменти (в случая е един):

$$X \sim U(0, \theta) \Rightarrow EX = \frac{a+b}{2} = \frac{0+\theta}{2} = \frac{\theta}{2}$$

→ намираме на емпиричните моменти (в случая един)

* средната стойност на наблюденията в извадката

$$\overline{X} = \frac{1}{n} \sum_{j=1}^{n} x_j$$

→ приравняваме теоретичните и емпиричните моменти

$$EX = \overline{X}$$

$$\frac{\theta}{2} = \overline{X}$$

$$\Rightarrow \theta = 2\overline{X} \qquad \Rightarrow \hat{\theta}_{MM} = 2\overline{X}$$

2) $X \sim N(\mu, \sigma^2)$, $\vec{x} = (x_1, x_2, \dots, x_n)$ извадка над X

→ намираме теоретичните моменти

$$X \sim N(\mu, \sigma^2) \Rightarrow EX = \mu$$

$$DX = EX^2 - (EX)^2 \Rightarrow \sigma^2 = EX^2 - \mu^2$$

$$EX^2 = \sigma^2 + \mu^2$$

→ намираме емпиричните моменти

$$\overline{X}^{(1)} = \frac{1}{n} \sum_{j=1}^{n} x_j$$

$$\overline{X}^{(2)} = \frac{1}{n} \sum_{j=1}^{n} x_j^2$$

→ приравняваме

- $\mathbb{E}X = \mu = \overline{X} \Rightarrow \hat{\mu} = \overline{X}^{(n)}$

- $\mathbb{E}X^2 = \sigma^2 + \mu^2$

$$\frac{1}{n}\sum_{j=1}^{n} x_j^2 = \sigma^2 + \overline{X}^2 \Rightarrow \sigma^2 = \frac{1}{n}\sum_{j=1}^{n} x_j^2 - \overline{X}^2$$

*Твърдение за м.п.о. при $X \sim N(\mu, \sigma^2)$

$$\hat{\mu} = \overline{X} = \frac{1}{n}\sum_{j=1}^{n} x_j$$

$\hat{\sigma}^2$ има две опции:

i) $\hat{\sigma}^2 = \frac{1}{n}\sum_{j=1}^{n}(x_j - \mu)^2$ – когато μ е известно

ii) $\hat{\sigma}^2 = \frac{1}{n}\sum_{j=1}^{n}(x_j - \overline{X}_n)^2$ – когато μ не е известно

Доказателство:

$$L(\vec{X}, \theta) = \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_j-\mu)^2}{2\sigma^2}}, \quad \theta = (\mu, \sigma^2)$$

$$\tilde{L}(\vec{X}, \theta) = \ln L(\vec{X}, \theta) = n \ln \frac{1}{\sqrt{2\pi}} - \frac{n}{2} \ln \sigma^2 - \sum_{j=1}^{n} \frac{(x_j-\mu)^2}{2\sigma^2}$$

$$\frac{\partial \ln L}{\partial \mu} = \cancel{\frac{1}{\sigma^2}} \sum_{j=1}^{n}(x_j - \mu) = 0$$

$$\frac{\partial \ln L}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{j=1}^{n}(x_j-\mu)^2 = 0$$

$$\Rightarrow \hat{\mu} = \frac{1}{n}\sum_{j=1}^{n} x_j = \overline{X}_n^{(n)}$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{j=1}^{n}(x_j - \mu)^2$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{j=1}^{n}(x_j - \overline{X}_n^{(n)})^2$$

* Свойства:

* деф (неизместеност):

Казваме, че $\hat{\theta}$ е неизместена оценка на θ, ако $E\hat{\theta}(\vec{X}) = \theta$. Когато $\theta = (\theta_1, \dots, \theta_s)$, равенството се разбира като $E\hat{\theta}_j(\vec{X}) = \theta_j$, $1 \le j \le s$.

Иначе, ако това не е изпълнено, тогава

$|E\hat{\theta} - \theta|$ — систематична (систематичната) грешка.

$E\hat{\theta} - \theta = 0 \to$ неизместена

* деф (състоятелност)

Ако $\hat{\theta}(\vec{X}) \xrightarrow[n\to\infty]{P} \theta$, то $\hat{\theta}$ е състоятелна оценка

↳ какво означава това ⇒

$\theta \in \mathbb{R} \Rightarrow P(|\hat{\theta}(\vec{X}) - \theta| > \varepsilon) \xrightarrow[n\to\infty]{} 0$

↳ т.е. тази вероятност клони към 0, т.е. няма как да е > ε ?

примери за $U(0, \theta)$???

* деф (силна състоятелност):

Ако $\hat{\theta}(\vec{X}) \xrightarrow[n\to\infty]{п.с.} \theta$, то [illegible] на

* деф: функция на максималното правдоподобие наричаме съвместната плътност на $\vec{X}$, т.е.

$L(\vec{X}, \theta) = \prod_{j=1}^{n} f(x_j, \theta)$ ↳ понеже тези n сл. вел. са нез. и еднакво разпределени

$X \sim U(a,b)$ MM

→ теоретични

$$\mathbb{E}X = \frac{a+b}{2}$$

$$\mathbb{E}X^2 = \frac{a^2+ab+b^2}{3} \quad \text{или} \quad = \frac{(b-a)^2}{12}$$

→ емпирични

$$\overline{X}^{(1)} = \frac{1}{n}\sum_{j=1}^{n} x_j$$

$$\overline{X}^{(2)} = \frac{1}{n}\sum_{j=1}^{n} x_j^2$$

$$\mathbb{E}X = \overline{X}^{(1)}$$

$$\frac{a+b}{2} = \overline{X}^{(1)} \Rightarrow a+b = 2\overline{X}^{(1)}$$

$$\mathbb{E}X^2 = \overline{X}^{(2)} \Rightarrow \frac{a^2+ab+b^2}{3} = \frac{1}{n}\sum_{j=1}^{n} x_j^2 \quad (*)$$

$$\Rightarrow a+b = 2\overline{X}^{(1)} \Rightarrow b = 2\overline{X}^{(1)} - a$$

заместване в $(*)$ $\dfrac{a^2 + a(2\overline{X}^{(1)} - a) + (2\overline{X}^{(1)} - a)^2}{3} = \dfrac{1}{n}\sum_{j=1}^{n} x_j^2$

$$\cancel{\frac{a^2 + 2a\overline{X}^{(1)} - a^2 + 4\overline{X}^{(1)2} - 4a\overline{X}^{(1)} + a^2}{3}}$$

...

$X \sim N(\mu, 1)$ $\quad \vec{x} = (x_1, x_2, \dots, x_n)$

$$f_X(\vec{x}, \theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2}}$$

$$L(\vec{x}, \theta) = \prod_{j=1}^{n} f_X(x_j, \theta) = \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(x_j-\mu)^2}{2}}$$

$$\Rightarrow \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\sum_{j=1}^{n}(x_j-\mu)^2}$$

$$\ln L = -\frac{n}{2} \ln (2\pi) - \frac{1}{2}\sum_{j=1}^{n}(x_j-\mu)^2$$

не зависи от μ, може да се игнорира

$$\frac{\partial \ln L}{\partial \mu} = \frac{\partial}{\partial \mu} - \frac{1}{2}\sum_{j=1}^{n}(x_j-\mu)^2 = -\frac{1}{2} \cdot 2 \sum_{j=1}^{n}(x_j-\mu) = -\sum_{j=1}^{n}(x_j-\mu)$$

$\rightarrow$ приравняване на 0

$$-\sum_{j=1}^{n}(x_j-\mu) = 0$$

$$\sum_{j=1}^{n} x_j = n\mu \Rightarrow \hat{\mu} = \frac{1}{n}\sum_{j=1}^{n} x_j = \overline{X_n}$$

$$\underline{\hat{\mu} = \overline{X_n}}$$

+ Доверителни интервали

$P(I_1 \le \theta \le I_2) \ge \gamma$ → ниво на доверие

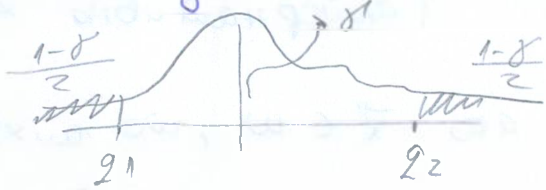
$\gamma = 0,99,\ 0,9$ → $0,95$ by default

– (I_1, I_2) се нарича доверителен интервал, с ниво на доверие γ

* Задача: Имаме някакъв модел, искаме да оценим θ, който описва модела, който имаме. Конструираме лява и дясна граница на случаен интервал (зависи от извадката), така че вероятността да намерим истинското, т.е. интервала (I_1, I_2) да хваща θ, да е поне γ $(0,95)$

→ θ е фиксиран параметър, а стохастичността е в I_1, I_2

→ Най-често се избира $1-\gamma$ да се слага в двата края, по половина т.е.



* Деф (Централна статистика): Казваме, че $T = T(\vec{X}, \theta)$ е централна статистика, ако:

1) T е строго монотонна по θ $(\uparrow\downarrow)$
2) $P(T \le x) = F_T(x)$ не зависи от θ

(T е функция на θ, но разпределението и не зависи от θ)

▸ Проверка на хипотези

~~Постановка~~ – имаме някаква хипотеза (твърдение), че свежда до някакво разпределение $F_X(x, \theta)$

– имаме някаква извадка $\vec{x}$

– и искаме да конструираме някакво подмножество на $\mathbb{R}^n \Rightarrow W \subseteq \mathbb{R}^n$, така че

– ако нашата конкретна реализация $\vec{x}$ (числата които получим в извадката) $\in W \Rightarrow$ отхвърляме хипотезата

$\vec{x} \notin W \Rightarrow$ приемаме хипотезата

→ математическа постановка

X, $F_X(x, \theta)$

права срещу права хипотеза (защото допускаме, че има някакво равенство)

$H_0: \theta = \theta_0$ (нулева хипотеза

$H_1: \theta = \theta_1$ (алтернативна хипотеза)

⇒ $\vec{x}$, W ако $\vec{x} \in W$, то отхвърляме H_0 и приемаме H_1

→ сложно е когато казваме, че θ е в някакъв интервал, множество, но не е конкретна стойност

▸ Грешка от първи род: $\mathbb{P}(\vec{x} \in W \mid H_0) = \alpha = \alpha_W$

Всички наблюдения да попаднат в W, т.е. да отхвърлим нулевата хипотеза, при положение, че тя е вярна

→ Грешка от втори род: $\mathbb{P}(\vec{X} \notin W \mid H_1) = \beta_\Gamma = \beta_W$

приели сме нулевата хипотеза, но е вярна алтернативата

* едната грешка плаща за цената на другата

$W_1 \subseteq W_2 \Rightarrow \begin{matrix} \alpha_{W_1} \leq \alpha_{W_2} \\ \beta_{W_1} \geq \beta_{W_1} \end{matrix}$ ⇒ имаме нужда от баланс

⇒ α е контролируема грешка (тя е обратно и е опасно)!

β не е толкова опасна грешка (казваме че ваксината влияе на 15%, а всъщност тя влияе на 1%.)

→ W - критична област за грешка от $I^{ви}$ род α

* деф (О.К.О): ~~[illegible]~~

Казваме, че $W^* \in \mathbb{R}^n$ е оптимална критична област (О.К.О) за фиксирана грешка от $I^{ви}$ род α, ако

$$\beta_{W^*} = \min_{\alpha_W = \alpha_{W^*} = \alpha} \beta_W$$

→ т.е. фиксирали сме грешка от първи род, и търсим такава критична област, така че грешката от втори род да е минимална

! Нулевата хипотеза е тази която ни е важно, да не я зъбркаме много

→ разделяме пространството $\mathbb{R}^n$ на 2 и ако нашата извадка $\vec{x}$:

$\vec{x} \in W^* \Rightarrow$ приемаме алтернативната

$\vec{x} \in (W^*)^c = \mathbb{R}^n / W^*$ → приемаме нулевата хипотеза

допълнението

* Лемата на Нейман-Пирсън ще ни позволи да видим, как се конструира оптимална критична област, така че имаме фиксирана грешка от първи род и минимална грешка от втори род, т.е. имаме възможно най-добрия случай

* Лема (Нейман-Пирсън): → изисква плътност, т.е. X да е непрекъсната сл. вел.

Нека X, $f_X(x,\theta)$, $L(\vec{x},\theta)$ и извадка $\vec{x}$.

Ако $W^* \subseteq \mathbb{R}^n$ е к.о. с грешка от първи род α и

$$\left.\begin{aligned} W^* &\subseteq \{L_1(\vec{x}) \geq k L_0(\vec{x})\} \\ (W^*)^c &\subseteq \{L_1(\vec{x}) \leq k L_0(\vec{x})\} \end{aligned}\right\} \text{ за някое } k>0 .$$

Тогава W^* е О.К.О за $\begin{aligned} H_0&: \theta=\theta_0 \\ H_1&: \theta=\theta_1 \end{aligned}$

и казва, че ако искаме да конструираме О.К.О, то трябва да търсим региони от този тип

→ Линейна регресия

- регрес = връщане към средното

$$Y = \beta_0 + \beta_1 X + \varepsilon \rightarrow \text{грешка}$$

Y ↓ отклик; $\beta_1 X$ ↓ предиктор

↳ метода на най-малките квадрати

↳ $(\varepsilon_1, \dots, \varepsilon_n)$ - са взаимно независими

↳ $E\varepsilon_j = 0$, $\forall j \le n$ - т.е. липсва систематична грешка в модела

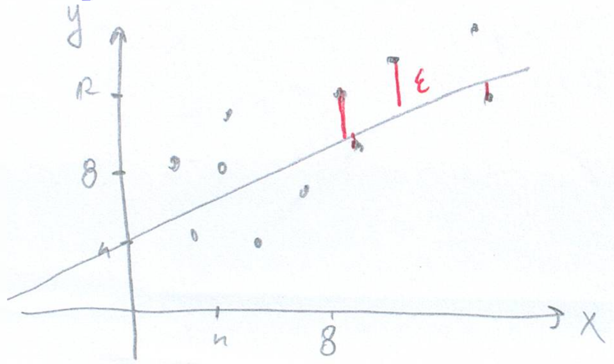
↳ $D\varepsilon_j = \sigma^2$ - хомо скедастичност

↳ $\varepsilon_j \sim N(0, \sigma^2)$ - нормалност на грешката

→ Two main objectives

- establish if there is a relationship between two variables
- forecast new observations
  - ↳ Can we use what we know about the relationship to forecast unobserved values?

* We call it "linear" because it equation represents a straight line in 2-dimensional plot



→ the linear regression is going to draw a line that minimizes this errors

$$y = \beta_0 + \beta_1 X + \varepsilon$$

y - dependent variable

x - independent variable

- β_0 is the constant or intercept
- β_1 is x's slope or coefficient
- ε is the error term

+ Confidence Intervals

↳ overall we're going to choose models with greater value for R-squared

Точкови оценки — запис видео 12

* ф-я на максимално правдоподобие

$X \sim N(\mu, \sigma^2)$, $\vec{X}$; $\underset{(\mu, \sigma^2)}{\theta} \in \Theta = \mathbb{R} \times \mathbb{R}^+$ → реални → реални неотрицателни числа

$$L(\vec{x}, \theta) = (2\pi)^{-\frac{n}{2}} (\sigma^2)^{-\frac{n}{2}} e^{-\sum_{j=1}^{n} \frac{(x_j - \mu)^2}{2\sigma^2}}$$

* логаритъм е растяща ф-я

↳ решава се по $\frac{\partial \ln L}{\partial \theta} = 0$ → тази ф-я има само глобален максимум

* Твърдение: $\vec{X}$ е извадка (n - н.е. р.сл. вел.) на X както $X \sim N(\mu, \sigma^2)$.

Тогава μ е $\overline{X_n} = \frac{\sum_{j=1}^{n} X_j}{n}$

↓ средното аритметично

М.П.О. за σ^2 е

а) $\hat{\sigma}^2 = \frac{1}{n} \sum_{j=1}^{n} (X_j - \mu)^2$ ако μ се знае

б) $\hat{\sigma}^2 = \frac{1}{n} \sum_{j=1}^{n} (X_j - \overline{X_n})^2$ ако μ не се знае

↓ винаги вярно, независимо дали имаме информация за дисперсията или не

д-во: $\ln L(\vec{X}, \theta) = -\frac{n}{2} \ln 2\pi - \frac{1}{2} \ln \sigma^2 - \sum_{j=1}^{n} \frac{(X_j - \mu)^2}{2\sigma^2}$ $\theta = (\mu, \sigma^2)$

$$\frac{\partial \ln L}{\partial \mu} = \frac{1}{\sigma^2} \sum_{j=1}^{n} (X_j - \mu) = 0 \Rightarrow \mu = \overline{X_n}$$

↳ по μ

↳ мак. прав. оценка на μ винаги е средното аритметично

$$\frac{\partial \ln L}{\sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{j=1}^{n} (x_j - \mu)^2 = 0$$

$\mu = \overline{x_n}$

↳ диференцира по σ^2

→ Ако знаем $\mu \Rightarrow \hat{\sigma}^2 = \frac{1}{n} \sum_{j=1}^{n} (x_j - \mu)^2$

но ако не знаем $\mu \Rightarrow \hat{\sigma}^2 = \frac{1}{n} \sum_{j=1}^{n} (x_j - \overline{x_n})^2$

↳ Ако е само един неизвестен параметър, решаваме само $\frac{\partial \ln L}{\sigma^2}$ и получаваме $\hat{\sigma}^2$ → точкова оценка за дисперсията

↳ Ако не знаем и двата параметри, решаваме двете уравнения (производна по първото, производна по второто)

→ М.П. работи защото там има най-много концентрирана вероятност

→ $\vec{X}$, имат м-во от хипотези H за плътността на X, сл. вел., чиято извадка е $\vec{X}$. За всяка хипотеза имаме плътност $f_X(h)$, $h \in H$

Знаем, че $h_t \in H$ такава, че $f_X(x) = f(x, h_t)$
↳ истинска, true хипотеза

$H = \mathbb{R} \times \mathbb{R}_+$ $h = (\mu, \sigma^2)$, $X \sim N(\mu, \sigma^2)$

$h = \theta$
↓
защото хипотезите не е нужно да се изчерпват с числа или вектори от числа

$h_t = (0, 1)$
↓
може да е 0 или 1, ако зад всичко стои стандартно нормална сл. вел.

* Ентропия на Y

$$Ent(Y) = \int \log f_Y(y)\, f_Y(y)\, dy$$

↳ трябва да вземаме хипотези, които са близки до истинската, че ентропията да се вижда добре

$$Ent(X) = \int \log f(y, h_t)\, f(y, h_t)\, dy$$

$$\| \qquad > \mathbb{E} \log f(x, h_t), \quad \forall h \neq h_t$$

$$\mathbb{E} \log f(x, h_t)$$

↳ $h \in \mathcal{H} : \dfrac{\ln L(\vec{X}, h)}{n} \xrightarrow[n \to \infty]{} \mathbb{E} \log f(x, h)$

$h_t \in \mathcal{H} : \dfrac{\ln L(\vec{X}, h_t)}{n} \xrightarrow[n \to \infty]{} Ent(X)$

!→ максимума ни гарантира, че ще се схождаме към най-истинската

↳ гарантира, че при много хипотези ще се схождаме към истинската хипотеза

Метод на моментите

→ При М.М.П избираме нещо, което максимизира съвместната плътност

↳ тук при М.М. не е необходимо да изисваме плътност

$X, \; F(x, \Theta), \; \vec{X}_n$

$$\mathbb{E}X \overset{п.с.}{=} \lim_{n \to \infty} \frac{\sum_{i=1}^{n} X_i}{n}$$

$\mathbb{E}X = \mu(\Theta_0)$

, да допуснем че $\mathbb{E}Y = \mu(\Theta)$, ако Y има $F(x, \Theta)$

когато увеличаваме данните ще се схождаме към истинското очакване т.е. $\mu(\Theta_0)$

$$\overline{X_n}^{(1)} := \frac{\sum_{j=1}^{n} x_j}{n} = \mu(\theta) \Rightarrow \hat{\theta}\, \mu^{-1}\left(\overline{X_n}^{(1)}\right)$$

+ Деф: Нека X, $F(X,\theta)$, $\theta = (\theta_1, \dots, \theta_s)$ и $\vec{X}$,

поставяме $\quad X_n^{(j)} = \dfrac{\sum_{j=1}^{n} X_k^{(j)}}{n} \xrightarrow[n\to\infty]{} \mathbb{E} X^j$

Решението на уравненията $\underbrace{\mathbb{E}X^j}_{\mu^j(\theta)} = X_n^j$, $1 \le j \le s$

$$\hat{\theta} = \mu^{-1}\left(X_n^{(\cdot)}\right), \quad X_n^{(\cdot)} = \left(X_n^{(1)}, \dots, X_n^{(s)}\right)$$

се нарича оценка по М.М.

* Всеки един от тези моменти, да кажем, j на брой е някаква ф-я на неизвестните параметри

Пр. при нормалното, първият момент е просто μ

⊕ $X \in U(0,\theta)$, $\theta > 0$, $\vec{X}$

$$X_n^{(1)} = \frac{\sum_{j=1}^{n} x_j}{n} = \frac{\theta}{2} = \mathbb{E}X \Rightarrow \hat{\theta} = 2 X_n^{(1)}$$

$$X_n^{(1)} = \frac{\hat{\theta}}{2} \xrightarrow[n\to\infty]{\text{п.с.}} \mathbb{E}X = \frac{\theta_0}{2} \Rightarrow \hat{\theta} \to \theta_0$$

от ЗГЧ

М.П.О $\hat{\theta} = \max(X_i) \neq 2X_n^{(1)}$

⊕ $X \sim N(\mu, \sigma^2)$

първият момент е средното аритметично от първите степени

$$\mathbb{E}X = \mu = X_n^{(1)} = \frac{\sum_{j=1}^{n} X_j}{n} \Rightarrow \hat{\mu} = X_n^{(1)} \equiv$$ оценка по М.М.П.

$$\mathbb{E}X^2 = \sigma^2 + \mu^2 = \underbrace{\frac{\sum_{j=1}^{n} X_j^2}{n}}_{X_n^{(2)}} \Rightarrow \hat{\sigma}^2 = \frac{\sum_{j=1}^{n} X_j^2}{n} - \left(X_n^{(1)}\right)^2 = X_n^{(2)} - \left(X_n^{(1)}\right)^2$$

вторият момент е средното аритметично от вторите степени

$\equiv$ оценка по М.М.П

* Свойства на точковите оценки

а) Неизместеност

$\theta_0 = \mathbb{E}\hat{\theta}(\vec{X})$ → т.е. не искаме точкова оценка, която с увеличаването на данните, или която има систематична грешка в средното, такава отклонение от истинската оценка

истинския параметър

⊕ $X \sim N(\mu, \sigma^2)$ $\hat{\mu} = X_n^{(1)}$ $\mathbb{E}\hat{\mu} = \mathbb{E}\,\frac{1}{n}\sum_{j=1}^{n} X_j = \mu$

реално няма как да му стигнем истинското очакване

→ Ако μ е известно, а не знаем σ^2, то

$$\mathbb{E}(\hat{\sigma}^2) = \mathbb{E}\left(\frac{1}{n}\sum_{j=1}^{n}(X_j - \mu)^2\right) = \frac{1}{n}\sum_{j=1}^{n}\mathbb{E}(X_j - \mu)^2 = \frac{1}{n}\sum_{j=1}^{n} DX_j = \sigma^2$$

→ когато знаем μ, $\hat{\sigma}$, даден с този израз е неизместена оценка на истинския параметър σ^2

↳ Ако μ и σ са неизвестни, то

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{j=1}^{n}\left(X_j - X_n^{(1)}\right)^2 = \frac{1}{n}\sum_{j=1}^{n} X_j^2 - \left(\frac{1}{n}\sum_{j=1}^{n} X_j\right)^2 = X_n^{(2)} - \left(X_n^{(1)}\right)^2$$

очакването на този израз е п пъти по втория момент на X_1, защото те са еднакво разпределени

$$\mathbb{E}\hat{\sigma}^2 = \mathbb{E}X_n^{(2)} - \mathbb{E}\left(X_n^{(1)}\right)^2 \ominus$$

ще ни даде ли истинската стойност σ^2, която е зад модела

$$\mathbb{E}\hat{\sigma}^2 = \left(\frac{n-1}{n}\right)\sigma^2 \approx 1 \cdot \sigma^2$$

при големи извадки 2

$$\ominus\ \mathbb{E}X_1^2 - \frac{1}{n^2}\mathbb{E}\left(\sum_{j=1}^{n} X_j^2 + \sum_{i\neq j} X_i X_j\right) =$$

$$= \mathbb{E}X_1^2 - \frac{n}{n^2}\mathbb{E}X_1^2 - \frac{1}{n^2}\sum_{i\neq j}\mathbb{E}X_iX_j =$$

толкова пъти го имаме μ

$$= \left(\frac{n-1}{n}\right)\mathbb{E}X_1^2 - \frac{1}{n^2}\sum_{i\neq j}\underbrace{\mathbb{E}X_i}_{\mu}\,\underbrace{\mathbb{E}X_j}_{\mu} = \left(\frac{n-1}{n}\right)\mathbb{E}X_1^2 - \frac{1\cdot\mu^2}{n^2}\cdot n(n-1) =$$

$$= \left(\frac{n-1}{n}\right)\underbrace{\left(\mathbb{E}X_1^2 - (\mathbb{E}X_1)^2\right)}_{\sigma^2} = \frac{n-1}{n}\sigma^2$$

↳ $\hat{\sigma}^2$ е неизместената оценка на σ^2, защото нейното очакване е $\left(\frac{n-1}{n}\right)\sigma^2$ истинското σ

↳ това се нарича асимптотично неизвестна оценка, ~~неизместена~~ е защото не е точно равно на σ^2, но за голямо n, ще е близу до истинското σ

б) Състоятелност на точкова оценка

деф: $X \sim F(x, \theta)$, $\theta = (\theta_1, \dots, \theta_s)$, $\vec{X}$, то казваме, че

$\theta = \theta(\vec{X})$ е състоятелна оценка на θ, $\hat{\theta} \xrightarrow[n \to \infty]{P} \theta$ или

еквивалентно $\hat{\theta}_j \xrightarrow[n \to \infty]{} \theta_j$, $\forall 1 \le j \le s$

* всяка една трябва да се схожда, по вероятност, към истинската θ

Казваме, че $\hat{\theta}$ е силно състоятелна, ако $\hat{\theta} \xrightarrow[n \to \infty]{п.с.} \theta$ или

еквивалентно $\hat{\theta}_j \xrightarrow[n \to \infty]{п.с.} \theta_j$, $\forall 1 \le j \le s$

Пр $X \sim N(\mu, \sigma^2)$ и $\hat{\mu} = X_n^{(1)} \xrightarrow[n \to \infty]{УЗГЧ} \mu$ $\qquad X_j = Y_j$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{j=1}^{n} (X_j - \bar{X}_n)^2 = \frac{1}{n} Y_n^{(2)} - \left(\frac{Y_n^{(1)}}{2}\right)^2 =$$

$$= \frac{\sum_{j=1}^{n} X_j^2}{n} - \left(\frac{\sum_{j=1}^{n} X_j}{n}\right)^2$$

$$\xrightarrow[n \to \infty]{УЗГЧ} \mathbb{E} X_1^2 - \mu^2 = \sigma^2 + \cancel{\mu^2} - \cancel{\mu^2} = \sigma^2$$