

Kom ihåg: $\text{Cov}(X, Y) = E[(X - E[X])(Y - E[Y])]$

och

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \text{Var}(Y)}}$$

Sats: $\forall \text{ Cov}(aX, Y) = a \text{Cov}(X, Y)$

2/

$$\text{Var}(X - Y) = \text{Var}(X) + \text{Var}(Y) - 2 \text{Cov}(X, Y)$$

B: $\forall$ följer direkt av det.

2/

$$\begin{aligned} \text{Var}(X - Y) &= \text{Var}(X) + \text{Var}(-Y) + 2 \text{Cov}(X, -Y) \\ &= \text{Var}(X) + (-1)^2 \text{Var}(Y) - 2 \text{Cov}(X, Y) \quad \square \end{aligned}$$

Anm: Om X, Y är oberoende:

$$\text{Var}(X - Y) = \text{Var}(X) + \text{Var}(Y).$$

Sats: $-1 \leq \rho \leq 1$

B: Låt $\sigma_X = \sqrt{\text{Var}(X)}$ o $\sigma_Y = \sqrt{\text{Var}(Y)}$

vi har att

$$\begin{aligned} 0 &\leq \text{Var}\left(\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y}\right) = \text{Var}\left(\frac{X}{\sigma_X}\right) + \text{Var}\left(\frac{Y}{\sigma_Y}\right) - 2 \text{Cov}\left(\frac{X}{\sigma_X}, \frac{Y}{\sigma_Y}\right) \\ &= \frac{\text{Var}(X)}{\sigma_X^2} + \frac{\text{Var}(Y)}{\sigma_Y^2} - 2 \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} \\ &= 1 + 1 - 2\rho = 2(1 - \rho) \Rightarrow \rho \leq 1 \end{aligned}$$

p.s.s.

$$0 \leq \text{Var} \left(\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y} \right) = 2(1+\rho) \Rightarrow \rho \geq -1 \quad \square$$

Betingade tthner

Def:

Låt (X, Y) vara en två-dim. s.v. med gemensam tthn $f_{X,Y}$ och marg.-tthner $f_X(x) \leq f_Y(y)$.

Den betingade tthnen för X givet $Y=y$ (betecknad $f_{X|Y}$) ges av

$$f_{X|Y} = \begin{cases} \frac{f_{X,Y}(x,y)}{f_Y(y)} & \text{om } f_Y(y) > 0 \\ 0 & \text{annars.} \end{cases}$$

Anm: Om (X, Y) är diskreta ser vi att

$$f_{X|Y} = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{P(X=x, Y=y)}{P(Y=y)} = P(X=x | Y=y)$$

som i def. $P(A|B) = \frac{P(A \cap B)}{P(B)}$

Sats (Lagen om Total Sannolikhet):

Om (X, Y) är en två-dim. diskr. s.v. har vi att

$$P(X \in A | Y=y) = \sum_{x \in A} f_{X|Y}(x)$$

och att

$$P(X \in A) = \sum_y P(X \in A | Y=y) f_Y(y)$$

B: Vi har att

$$P(X \in A | Y=y) = \frac{P(X \in A, Y=y)}{P(Y=y)} = \frac{P\left(\bigcup_{x \in A} \{X=x, Y=y\}\right)}{P(Y=y)}$$

$$= \frac{\sum_{x \in A} P(X=x, Y=y)}{P(Y=y)} = \sum_{x \in A} P(X=x | Y=y) = \sum_{x \in A} f_{X|Y}(x, y)$$

Vi ser sedan att

$$\sum_y P(X \in A | Y=y) f_Y(y) \stackrel{y}{=} \sum_y \sum_{x \in A} f_{X|Y}(x, y) f_Y(y)$$

$$= \sum_{x \in A} \sum_y f_{X,Y}(x, y) = \sum_{x \in A} f_X(x) = P(X \in A) //$$

Sats: Om X, Y kont. s.v. $P(X \in A | Y=y) = \int_A f_{X|Y}(x, y) dx =$

$$P(X \in A) = \int_{-\infty}^{\infty} P(X \in A | Y=y) f_Y(y) dy \quad B: \text{analogt B}$$

Def: För två s.v. definieras detbetingade väntevärde av X givet $Y=y$
av

$$E[X | Y=y] = \begin{cases} \sum_x x f_{X|Y}(x, y) & \text{om } X, Y \text{ diskreta} \\ \int_{-\infty}^{\infty} x f_{X|Y}(x, y) dx & \text{om } X, Y \text{ kont.} \end{cases}$$

Sats: Vi har att

$$E[X] = \begin{cases} \sum_y E[X | Y=y] f_Y(y) & \text{om } (X, Y) \text{ diskr.} \\ \int_{-\infty}^{\infty} E[X | Y=y] f_Y(y) dy & \text{" " kont.} \end{cases}$$

B: övn (svär) D]

Tal: Låt (X, Y) ha gemensam tthn

$$f_{X,Y}(x,y) = y e^{1-xy} \quad \text{för } (x,y) \in [0,1]^2$$

Bestäm $f_{X|Y}(x)$ och använd denna för

att beräkna $E[X|Y=y]$.

L: Vi har att $f_{X|Y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$

och $f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dx = \int_0^1 y e^{1-xy} dx$

$$= e \left[-e^{-xy} \right]_0^1 = e(1 - e^{-y}) \quad \text{för } 0 \leq y \leq 1.$$

$$\Rightarrow f_{X|Y} = \frac{y e^{1-xy}}{e(1-e^{-y})} = \frac{y e^{-xy}}{1-e^{-y}} \quad \text{för } 0 \leq x, y \leq 1.$$

Vi har vidare att

$$E[X|Y=y] = \int_{-\infty}^{\infty} x f_{X|Y} dx = \int_0^1 \frac{xy e^{-xy}}{1-e^{-y}} dx$$

P.T.

$$= \frac{1}{1-e^{-y}} \left(\left[-x e^{-xy} \right]_0^1 - \int_0^1 -e^{-xy} dx \right)$$

$$= \frac{1}{1-e^{-y}} \left(0 - e^{-y} + \int_0^1 e^{-xy} dx \right) = \frac{-e^{-y}}{1-e^{-y}} + \frac{1}{1-e^{-y}} \left[\frac{-e^{-xy}}{y} \right]_0^1$$

$$= \frac{1}{1-e^{-y}} + \frac{1-e^{-y}}{y(1-e^{-y})} = \frac{1}{1-e^{-y}} + \frac{1}{y} //$$

Gränsvärdessatser (många obero. s.v.)

Vi betraktar en oändlig sekvens av
oberoende och likafördelade s.v.

$$X_1, X_2, \dots \quad (\text{i.i.d. sekr.})$$

Tag en i.i.d. sekvens $X_1, X_2, \dots$ med
 $\mu = E[X_k] < \infty$ och $\sigma^2 = \text{Var}(X_k) < \infty, \forall k$
 bildar medelvärdet (av de n första)

$$\bar{X}_n = \frac{X_1 + \dots + X_n}{n} = \frac{1}{n} \sum_{k=1}^n X_k.$$

Vad kan vi säga om $\bar{X}_n$?

Sats (Stora Talens Lag, STL, LLN):

För varje $\varepsilon > 0$ gäller att

$$\lim_{n \rightarrow \infty} P(|\bar{X}_n - \mu| > \varepsilon) = 0$$

Ann: Tag $\varepsilon = 10^{-100}$. Till slut kommer
 $\bar{X}_n$ vara närmare μ än 10^{-100} säger STL.

B: Kom ihåg Chebyshev's olikhet:

$$P(|X - \mu_X| \geq a) \leq \frac{\text{Var}(X)}{a^2} \quad \forall a > 0.$$

$$\begin{aligned} \text{Här är } E[\bar{X}_n] &= E\left[\frac{1}{n} \sum_{k=1}^n X_k\right] = \frac{1}{n} \sum_{k=1}^n E[X_k] \\ &= \frac{1}{n} \sum_{k=1}^n \mu = \frac{1}{n} \cdot n \mu = \mu \end{aligned}$$

$$\begin{aligned}
 \Rightarrow P(|\bar{X}_n - \mu| > \varepsilon) &= P(|\bar{X}_n - E[\bar{X}_n]| > \varepsilon) \\
 &\leq \frac{\text{Var}(\bar{X}_n)}{\varepsilon^2} = \frac{1}{\varepsilon^2} \text{Var}\left(\frac{1}{n} \sum_{k=1}^n X_k\right) = \frac{1}{n^2 \varepsilon^2} \text{Var}\left(\sum_{k=1}^n X_k\right) \\
 &\stackrel{\text{ober}}{=} \frac{1}{n^2 \varepsilon^2} \sum_{k=1}^n \text{Var}(X_k) = \frac{1}{n^2 \varepsilon^2} \cdot n \sigma^2 = \frac{\sigma^2}{n \varepsilon^2} \rightarrow 0 \quad \text{da } n \rightarrow \infty.
 \end{aligned}$$

Ex: Kasta en vanlig tärning många ggr. Medelvärdet kommer närma sig väntevärdet som är 3.5.

Sats (Centrala gränsvärdesatsen, CLT):

Låt $X_1, X_2, \dots$ vara en iid sekvens med

$\mu = E[X_1] < \infty$ o $\sigma^2 = \text{Var}(X_1) < \infty$. Låt

sedan $S_n = \sum_{k=1}^n X_k$. Vi har att

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \rightarrow Z \sim N(0, 1)$$

B: Svårt. Använder mgt + Taylor $\square$

Har använder vi CLT i praktiken?

För stora n är

$$\frac{1}{\sigma\sqrt{n}} \sum_{k=1}^n (X_k - \mu) \hat{\sim} N(0, 1).$$

$$2/ \quad \sum_{k=1}^n \bar{X}_k \approx N\left(\mathbb{E}\left[\sum_{k=1}^n \bar{X}_k\right], \text{Var}\left(\sum_{k=1}^n \bar{X}_k\right)\right)$$

$$3/ \quad \bar{X}_n = \frac{1}{n} \sum_{k=1}^n \bar{X}_k \approx N\left(\mathbb{E}[\bar{X}_n], \text{Var}(\bar{X}_n)\right)$$

Ex: Låt $Z \sim \text{Bin}(n, p)$. Vi kan skriva

$$Z = X_1 + X_2 + \dots + X_n \quad \text{där } X_1, X_2, \dots, X_n \text{ är oberoende,}$$

$$X_k \sim \text{Be}(p) \quad (X_k \in \{0, 1\} \text{ o. } \mathbb{P}(X_k = 1) = p).$$

CLT ger att om $np, n(1-p)$ är stora

(typ > 5 alt > 10) så är summan

$$Z \approx N(\mathbb{E}[Z], \text{Var}(Z)) = N(np, np(1-p)).$$

$$\uparrow$$

$$Z \sim \text{Bin}(n, p)$$

Parametrisk punktskattningVår statistiska modell:

1/ Vi drar ett slumpmässigt stickprov (sample) $\underline{X}_1, \underline{X}_2, \dots, \underline{X}_n$ (av storlek n) ur en population. S.v. $\underline{X}_1, \dots, \underline{X}_n$ antas vara oberoende.

2/ Vi antar att fördelningen av $\underline{X}_i$ är känd såväl som på värdet av en eller flera parametrar.

Ex: $\underline{X}_i \sim \text{Exp}(\beta)$ där β är okänd.

Ex: $\underline{X}_i \sim N(\mu, \sigma^2)$ med μ och/eller σ^2 okända.

Vi vill skatta (gissa) värdet på parametrarna, och låter θ beteckna en generisk parameter.

Def: En (punkt-)skattare för θ är en funktion $\hat{\theta} = \hat{\theta}(\underline{X}_1, \dots, \underline{X}_n)$ av stickprovet $\underline{X} = (\underline{X}_1, \dots, \underline{X}_n)$.

Ett experiment genomförs och man samlar in datapunkter $x_1, x_2, \dots, x_n$. Dessa är uttallen av $\underline{X}_1, \dots, \underline{X}_n$.

Def: En (punkt-)skattning av θ får man genom att använda data $x_1, \dots, x_n$ i skattaren, dvs $\hat{\theta}(x_1, \dots, x_n)$.

Ann: $\hat{\theta}(\bar{x})$ är en s.v. men $\hat{\theta}(x)$ är ett tal.

Ex: 1) Om $\theta = \mu = E[X_1]$ så använder

vi skattaren $\hat{\theta}(\bar{x}) = \bar{x} = \frac{x_1 + \dots + x_n}{n}$

(sample mean).

2) Om $\theta = \sigma^2 = \text{Var}(X_1)$ så används ofta

$$\hat{\theta} = \widehat{\sigma^2} = s^2 = \frac{1}{n-1} \sum_{k=1}^n (x_k - \bar{x})^2$$

(sample variance)

Def: Skattaren $\hat{\theta}$ är väntevärdesriktig (VVR) om $E[\hat{\theta}(\bar{x})] = \theta$ (unbiased).

Effektiviteten (av VVR-skattare)

bedöms genom variansen $\text{Var}(\hat{\theta}(\bar{x}))$.

(Liten = bra).

Fördelningen för $\hat{\theta}(\mathbf{X})$ är av central betydelse.

Betrakta $\hat{\theta}(\mathbf{X}) = \bar{X} = \frac{1}{n} \sum_{k=1}^n X_k$.

Ex 1:

Fördelningen av $\bar{X}$ är i allmänhet svåråtkomlig men:

i)
$$E[\bar{X}] = E\left[\frac{1}{n}(X_1 + \dots + X_n)\right] = \frac{1}{n} \sum_{k=1}^n E[X_k]$$

$$= \frac{1}{n} \sum_{k=1}^n \mu = \mu \quad \text{så } \bar{X} \text{ är en VVIZ skattare för } \mu.$$

ii) Om $\sigma^2 = \text{Var}(X_k) \Rightarrow$

$$\text{Var}(\bar{X}) = \text{Var}\left(\frac{1}{n} \sum_{k=1}^n X_k\right) \underset{\text{ober.}}{=} \frac{1}{n^2} \sum_{k=1}^n \text{Var}(X_k) = \frac{1}{n^2} n \sigma^2 = \frac{\sigma^2}{n}$$

Ex 2: Betrakta
$$S^2 = \frac{1}{n-1} \sum_{k=1}^n (X_k - \bar{X})^2$$

Lång räkning (appendix C) ger att $E[S^2] = \sigma^2$, så att S^2 är en VVIZ skattare för σ^2 .

Ex 3: Vi vill skatta $p = P(A)$ där

$A \subset \Omega$ är någon händelse (utvald person har en viss gen etc). Vi vill skatta p

och låter
$$X_k = \begin{cases} 1 & \text{om } A \text{ inträffar på försök } k \\ 0 & \text{annars.} \end{cases}$$

Vi låter $\hat{p} = \bar{X}$, Vi observerar att

$X_1 + X_2 + \dots + X_n \sim \text{Bin}(n, p)$ så att

$$E[\hat{p}] = E\left[\frac{1}{n} \sum_{k=1}^n X_k\right] = \frac{1}{n} E\left[\underbrace{\sum_{k=1}^n X_k}_{\text{bin-förd}}\right] = \frac{1}{n} \cdot np = p \quad (\text{vvr})$$

och

$$\begin{aligned} \text{Var}(\hat{p}) &= \text{Var}\left(\frac{1}{n} \sum_{k=1}^n X_k\right) = \frac{1}{n^2} \text{Var}\left(\sum_{k=1}^n X_k\right) \stackrel{\downarrow}{=} \frac{1}{n} np(1-p) \\ &= \frac{p(1-p)}{n}. \end{aligned}$$

Vi är ofta intresserade av $\text{Var}(\hat{\theta}(X))$,

Def: Standardfelet är $\sqrt{\text{Var}(\hat{\theta}(X))}$

och betecknas ibland $D(\hat{\theta})$.

Anm: Ofta beror $\text{Var}(\hat{\theta}(X))$ på den

okända parametern θ . Därför måste

$D(\hat{\theta})$ skattas. Vi skattar en egenskap

hos en skattare!

Ex: Om $\bar{X}$ skattar μ har vi $\text{Var}(\bar{X}) = \frac{\sigma^2}{n}$.

Vi skattar $\sqrt{\text{Var}(\bar{X})}$ med $s_{\bar{X}} = \frac{s}{\sqrt{n}}$ där

$$s^2 = \frac{1}{n-1} \sum_{k=1}^n (X_k - \bar{X})^2,$$

Ex: Om $\hat{p} = \bar{X}$ skattar p låter vi

$$S_{\hat{p}} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \quad \text{skatta} \quad \text{Var}(\hat{p}) = \sqrt{\frac{p(1-p)}{n}}.$$

Hur skall vi hitta en bra skattare?

Momentmetoden

Om $X_1, X_2, \dots$ är iid (ober. o. likafördelade) med parameter θ är det ofta så att $E[X_i] = g(\theta)$ för ngn fun g .

Idé:

1) Observera att STL säger att för n stort så är $\bar{X} \approx E[X_i]$.

2) Lös ut $\hat{\theta}$ ur ekvationen $\bar{X} = g(\hat{\theta})$.

Vi ser att $g(\theta) = E[X_i] \approx \bar{X} = g(\hat{\theta})$ och därmed bör också $\theta \approx \hat{\theta}$ s.a. $\hat{\theta}$ från 2) ovan är en bra skattare.

Detta kallas för en MME- (Method of Moment Estimator) skattare.

Tal: Väntetiden på en buss anses

vara $U[0, \theta]$.

a) Hitta MME'n för θ

b) Om vi fick data (mätvärdena)

$x_1=2, x_2=7, x_3=4, x_4=15, x_5=11$ minuter,
vilket värde får vår skattning?

L: a) Momentmetoden:

i) Hitta g s.a. $E[X_1] = g(\theta)$

ii) Lös ut $\hat{\theta}$ ur sambandet $\bar{X} = g(\hat{\theta})$

Vi har att $X_1 \sim U[0, \theta]$ s.a. $E[X_1] = \frac{\theta}{2}$,

$\Rightarrow \bar{X} = \frac{\hat{\theta}}{2} \Rightarrow \hat{\theta} = 2\bar{X}$ är vår MME,

b) Skattningen blir $\hat{\theta}(x_1, \dots, x_5) = 2\bar{X}$

$$= 2 \cdot \frac{x_1 + \dots + x_5}{5} = 15.6 \text{ minuter.}$$

Om vi istället har två okända parametrar θ_1, θ_2 beräknar vi

$$E[X_1] = g_1(\theta_1, \theta_2)$$

$$E[X_1^2] = g_2(\theta_1, \theta_2).$$

Vi löser sedan ut $\hat{\theta}_1$ och $\hat{\theta}_2$ ur sambanden

$$\bar{X} = g_1(\hat{\theta}_1, \hat{\theta}_2)$$

$$\bar{X}^2 = g_2(\hat{\theta}_1, \hat{\theta}_2)$$

där $\bar{X}^k = \frac{1}{n} \sum_{i=1}^n X_i^k$

Kom ihåg: $\mathbf{X} = (X_1, \dots, X_n)$ är ett stöckprov av oberoende glikatördelade (i.i.d) s.v, och vi vill skatta värdet på parametern θ .

Notation: Om θ är okänd skriver vi $f(x_1, \dots, x_n | \theta)$ för den gemensamma tthn för s.v. $X_1, \dots, X_n$.

Anm: 1/ Då $X_1, \dots, X_n$ är oberoende ser vi att

$$f(x_1, \dots, x_n | \theta) = f(x_1 | \theta) \cdots f(x_n | \theta) = \prod_{k=1}^n f(x_k | \theta).$$

2/ Vi skriver $f(\cdot | \theta)$ för att poängtera att $\theta = (\theta_1, \dots, \theta_m)$ är okänd.

Def: Givet observerade data $x_1, \dots, x_n$ (dvs $X_1 = x_1, X_2 = x_2, \dots, X_n = x_n$) definierar vi likelihooden av θ som

$$L(\theta) = f(x_1, \dots, x_n | \theta)$$

Anm: 1/ Observera att $L(\theta)$ ses som en funktion av θ , ej $x_1, \dots, x_n$.

2/ Ibland skriver man $lik(\theta)$ istället för $L(\theta)$.

3/ Om s.v. $X_1, \dots, X_n$ är diskreta så blir

$$L(\theta) = f(x_1, \dots, x_n | \theta) = \prod_{k=1}^n f(x_k | \theta) = \prod_{k=1}^n P(X_k = x_k | \theta).$$

Alltså är $L(\theta)$ sann. att få de givna värdena $x_1, \dots, x_n$ som funktion av θ .

Def: ML-skattaren (Maximum Likelihood Estimator, MLE) av θ är det värde $\hat{\theta}$ som maximerar $L(\theta)$.

Anm: I det diskreta fallet är $\hat{\theta}$ det värde på θ som gör de observerade data mest troligt (most likely).

2/ Det kont. fallet är mer svårtolkat.

Vi vill alltså maximera $L(\theta)$. Ofta är det enklast att maximera

$$\begin{aligned} l(\theta) &= \log(L(\theta)) = \log\left(\prod_{k=1}^n f(x_k | \theta)\right) \\ &= \sum_{k=1}^n \log(f(x_k | \theta)). \end{aligned}$$

OBS: Vi söker värdet på θ som maximerar $L(\theta)$, inte själva maximum av $L(\theta)$.

Metod (för MLE):

1/ Hitta ett stickprov $x_1, \dots, x_n$ från en fördelning med tthn $f_{\mathbf{X}}(x)$ med parameter θ .

2/ Bestäm
$$L(\theta) = \prod_{k=1}^n f_{\mathbf{X}}(x_k | \theta)$$

3/ Hitta det θ som maximerar $L(\theta)$ (alt $l(\theta)$). Kalla detta värde $\hat{\theta}(x_1, \dots, x_n)$ (en skattning).

4/ Ersätt $x_1, \dots, x_n$ med $\bar{x}_1, \dots, \bar{x}_n$ för att hitta motsvarande skattare $\hat{\theta}(\bar{x}_1, \dots, \bar{x}_n)$.

Tal: Låt $\bar{x}_1, \dots, \bar{x}_n$ vara iid $\text{Poi}(\lambda)$.

a) Hitta MME för λ

b) Hitta MLE för λ

L: a) Om $\bar{x}_k \sim \text{Poi}(\lambda)$ så är $E[\bar{x}_k] = \lambda$.

MME: Lös ekvationen $\bar{x} = \hat{\lambda}$ (trivialt)

s.a. vår MME blir $\hat{\lambda} = \bar{x}$.

b) Givet stickprovet $x_1, \dots, x_n$ blir

$$L(\lambda) = \prod_{k=1}^n f(x_k | \lambda) = \prod_{k=1}^n \frac{\lambda^{x_k}}{x_k!} e^{-\lambda}$$

$$\begin{aligned}
 \Rightarrow \ell(\lambda) &= \log(L(\lambda)) = \sum_{k=1}^n \log\left(\frac{\lambda^{x_k}}{x_k!} e^{-\lambda}\right) \\
 &= \sum_{k=1}^n (x_k \log(\lambda) - \log(x_k!) - \lambda) \\
 &= \log(\lambda) \sum_{k=1}^n x_k - n\lambda - \sum_{k=1}^n \log(x_k!)
 \end{aligned}$$

Detta skall maximeras:

$$\ell'(\lambda) = \frac{1}{\lambda} \sum_{k=1}^n x_k - n$$

$$\ell''(\lambda) = -\frac{1}{\lambda^2} \sum_{k=1}^n x_k < 0 \quad \text{om} \quad \sum_{k=1}^n x_k \neq 0,$$

$$\text{Så } \ell'(\lambda) = 0 \Rightarrow \lambda = \frac{1}{n} \sum_{k=1}^n x_k = \bar{x} \quad \text{om} \quad \sum_{k=1}^n x_k \neq 0$$

och $\ell''(\lambda) < 0$ ger att $\hat{\lambda} = \bar{x}$ maximeras $\ell(\lambda)$. Om nu $\sum_{k=1}^n x_k = 0 \Rightarrow L(\lambda) = e^{-n\lambda}$

som maximeras då $\lambda = 0$ s. a. återigen blir $\hat{\lambda} = \bar{x}$.

Vår MLE blir $\hat{\lambda}(\underline{x}) = \bar{\underline{x}}$.

Tal: Låt $X_1, \dots, X_n$ vara i.i.d. $N(\mu, \sigma^2)$.

a) Hitta MME för μ, σ^2

b) Hitta MLE för μ, σ^2 .

L: a) Vi beräknar två moment:

$$E[X] = \mu \quad \text{och} \quad E[X^2] = \text{Var}(X) + E[X]^2 = \sigma^2 + \mu^2.$$

Enligt momentmetoden skall vi sätta

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

$$\overline{X^2} = \hat{\sigma}^2 + \hat{\mu}^2$$

$$\begin{aligned} \Rightarrow \hat{\sigma}^2 &= \overline{X^2} - \hat{\mu}^2 = \frac{1}{n} \left(\sum_{k=1}^n X_k^2 \right) - \overline{X}^2 \\ &= \frac{1}{n} \left(\sum_{k=1}^n X_k^2 - n \overline{X}^2 \right) = \frac{1}{n} \sum_{k=1}^n (X_k^2 + \overline{X}^2 - 2X_k \overline{X}) \\ &= \frac{1}{n} \sum_{k=1}^n (X_k - \overline{X})^2 // \end{aligned}$$

b) Givet stickprovet $x_1, \dots, x_n$ blir

$$L(\mu, \sigma^2) = \prod_{k=1}^n f(x_k | \mu, \sigma^2) = \prod_{k=1}^n \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{1}{2} \frac{(x_k - \mu)^2}{\sigma^2}\right)$$

$$\Rightarrow \ell(\mu, \sigma^2) = \sum_{k=1}^n \log\left(\frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{1}{2} \frac{(x_k - \mu)^2}{\sigma^2}\right)\right)$$

$$= -n \log \sigma - n \log \sqrt{2\pi} - \frac{1}{2\sigma^2} \sum_{k=1}^n (x_k - \mu)^2.$$

Vi hittar maximum:

$$\frac{\partial \ell}{\partial \mu} = \frac{1}{\sigma^2} \sum_{k=1}^n (x_k - \mu)$$

$$\frac{\partial \ell}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3} \sum_{k=1}^n (x_k - \mu)^2$$

$$\text{Om } \frac{\partial \ell}{\partial \mu} = 0 \Rightarrow 0 = \sum_{k=1}^n (x_k - \mu) = -n\mu + \sum_{k=1}^n x_k$$

$$\Rightarrow \mu = \bar{x}$$

Vidare är $0 = \frac{\partial l}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3} \sum_{k=1}^n (x_k - \mu)^2$

$$= \left\{ \mu = \bar{x} \right\} = -\frac{n}{\sigma} + \frac{1}{\sigma^3} \sum_{k=1}^n (x_k - \bar{x})^2$$

$$\Rightarrow \sigma^2 = \frac{1}{n} \sum_{k=1}^n (x_k - \bar{x})^2$$

Dvs $\hat{\mu}(\mathbf{x}) = \bar{x}$ och $\hat{\sigma}^2(\mathbf{x}) = \frac{1}{n} \sum_{k=1}^n (x_k - \bar{x})^2$

blir våra MLE,

Ann: Observera att $E[\hat{\sigma}^2] = E\left[\frac{n-1}{n} S^2\right] = \frac{n-1}{n} \sigma^2$

så att vår MLE ovan ej är VVR,

Observera även att (med $S^2 = \frac{1}{n-1} \sum_{k=1}^n (x_k - \bar{x})^2$)

$$E[S] = E[\sqrt{S^2}] \neq \sqrt{E[S^2]} = \sqrt{\sigma^2} = \sigma$$

så S är en VVR skattare för σ .