F10 MINSTA-KVADRATMETODEN, KONFIDENSINTERVALL, M.M.

Kap. 11.1–11.9, 12.1–12.3a | Linnéa Gustafsson
linneag2@kth.se

- Minsta-kvadratmetoden
- Väntevärdesriktighet, effektivitet, medelfel
- Konfidensintervall

KAP. 11

Repetition

En skattning av det rätta värdet θ kallas $\underbrace{\theta^*_{obs}}_{\text{Ett konkret värde som vi fått}}$ ("observerade skattningen av θ") och är ett utfall av den stokastiska variabeln θ^*.

$$\theta^*_{obs} = \underbrace{\theta^*_{obs}(x_1, x_2, \dots, x_n)}_{\text{Funktion av de mätdata vi fått}} \qquad \text{(punktskattningen)}$$

$$\theta^* = \theta^*(X_1, X_2, \dots, X_n)$$

Minsta-kvadratmetoden F.S. §9.2

Ex Anta att vi vill skatta arean av en kvadrat θ med MK-metoden och vi har mätt upp sidans längd och fått x_1 och x_2 och diagonalens längd och fått x_3. Då gäller att $E(X_1) = E(X_2) = \sqrt{\theta}$ och $E(X_3) = \sqrt{2}\sqrt{\theta}$. $V(X_i) = \sigma^2 \quad \forall i$.

Se §9.2 i F.S.

$$Q = \sum_{i=1}^{n} [\overbrace{\underbrace{x_i}_{\text{Det vi mätt upp}}}^{\text{Praktiska värdet}} - \overbrace{\underbrace{\mu_i(\theta_1, \theta_2, \dots, \theta_k)}_{\text{Det vi vill skatta}}}^{\text{Teoretiska värdet}}]^2$$

$\mu_i(\theta_1, \theta_2, \dots, \theta_k) = E(X_i)$ uttryckt i de parametrar vi vill skatta

Ju mindre skillnaden är mellan de uppmätta värdena x_i och våra teoretiska värden μ_i, desto bättre är teorin. De $\theta_1, \dots, \theta_k$ som minimerar Q väljs till MK-skattningarna av $\theta_1, \dots, \theta_k$.

"Lite grövre än maximum-likelihoodmetoden men funkar i princip alltid (behöver inte täthetsfunktionen / sannolikhetfunktionen)."

Tillbaka till vårt exempel:

Se § 9.2: $$Q = \sum_{i=1}^{n}[x_i - \underbrace{\mu_i(\theta_1, \dots, \theta_k)}_{E(X_i)}]^2 = \sum_{i=1}^{3}[x_i - \mu_i(\theta)]^2 =$$

$$= (x_1 - \sqrt{\theta})^2 + (x_2 - \sqrt{\theta})^2 + (x_3 - \sqrt{2}\sqrt{\theta})^2$$

$\frac{\partial Q}{\partial \theta} = 0$ ty minimera Q m.a.p. θ

$$\frac{\partial Q}{\partial \theta} = \cancel{2}(x_1 - \sqrt{\theta}) \cdot \left(\frac{-1}{\cancel{2\sqrt{\theta}}}\right) + \cancel{2}(x_2 - \sqrt{\theta}) \cdot \left(\frac{-1}{\cancel{2\sqrt{\theta}}}\right) + \cancel{2}(x_3 - \sqrt{2}\sqrt{\theta}) \cdot \left(\frac{-1}{\cancel{2\sqrt{\theta}}}\right)\sqrt{2} = 0$$

$$\Rightarrow x_1 + x_2 + \sqrt{2}x_3 = \sqrt{\theta} + \sqrt{\theta} + 2\sqrt{\theta} = 4\sqrt{\theta}$$

$$4\sqrt{\theta} = x_1 + x_2 + 2x_3 \Rightarrow \underbrace{\theta = \left(\frac{x_1 + x_2 + \sqrt{2}x_3}{4}\right)^2}_{\text{Ska eg. kolla 2:a-derivatan men struntar i det pga V-parabel}} \text{ minimerar } \theta$$

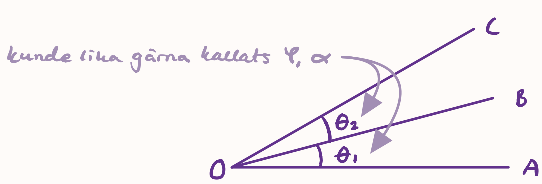
Lite smartare att mäta diagonalen än sidan (pga mindre relativt fel) $\Rightarrow$ viktar x_3 mer

$$\Rightarrow \theta^*_{obsMK} = \left(\frac{\underbrace{x_1 + x_2}_{} + \overbrace{\sqrt{2}\,x_3}}{4}\right)^2$$

Om bara mätt sidan $\Rightarrow$ medelvärde: $\frac{x_1 + x_2 + x_3}{3}$

Hur gör vi när vi ska skatta flera parametrar θ^*?

Ex 11.19 sid. 261



AOB : x_1, x_2, x_3

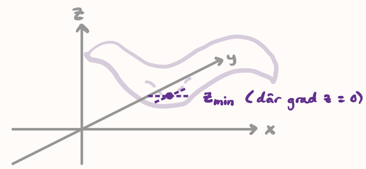
BOC : x_4, x_5

AOC : x_6, x_7

$$Q = \sum_{i=1}^{n}[x_i - \mu_i(\theta_1, \theta_2, \dots, \theta_k)]^2 = \sum_{i=1}^{7}[x_i - \mu_i(\theta_1, \theta_2)]^2 =$$

$$= (x_1 - \theta_1)^2 + (x_2 - \theta_1)^2 + (x_3 - \theta_1)^2 + (x_4 - \theta_2)^2 + (x_5 - \theta_2)^2 +$$
$$+ [x_6 - (\theta_1 + \theta_2)]^2 + [x_7 - (\theta_1 + \theta_2)]^2$$

min Q fås där grad $Q = \left(\frac{\partial \theta}{\partial \theta_1}, \frac{\partial \theta}{\partial \theta_2}\right) = (0, 0)$



Hoppar över att kolla om det verkligen är min och inte max / sadelpunkt

$$\left.\begin{array}{l} \frac{\partial Q}{\partial \theta_1} = 0 \\ \frac{\partial Q}{\partial \theta_2} = 0 \end{array}\right\} \Rightarrow \text{MK-skattningarna av } \theta_1 \text{ och } \theta_2$$

$$\Leftrightarrow [\text{se boken}] \Leftrightarrow \left.\begin{array}{l} x_1 + x_2 + x_3 + x_6 + x_7 = 5\theta_1 + 2\theta_2 \\ x_4 + x_5 + x_6 + x_7 = 2\theta_1 + 4\theta_2 \end{array}\right\} \Rightarrow$$

$$\Rightarrow \begin{cases} \theta^*_{1\,obsMK} = \dfrac{2x_1 + 2x_2 + 2x_3 - x_4 - x_5 + x_6 + x_7}{8} \\ \theta^*_{2\,obsMK} = \dfrac{-2x_1 - 2x_2 - 2x_3 + 5x_4 + 5x_5 + 3x_6 + 3x_7}{16} \end{cases}$$

Tre viktiga begrepp: väntevärdesriktighet, effektivitet och medelfel

Väntevärdesriktighet

Def. En skattning θ^*_{obs} av θ är väntevärdesriktig (vvr) $\Leftrightarrow$ $E(\theta^*) = \theta$

Ex $\mu^*_{obs} = \dfrac{x_1 + x_2 + x_3}{3}$ om $E(X_i) = \mu$

$$E(\mu^*) = E\left[\frac{X_1 + X_2 + X_3}{3}\right] = \frac{1}{3}\Big(E(X_1) + E(X_2) + E(X_3)\Big) = \frac{\mu + \mu + \mu}{3} = \mu$$

d.v.s. skattningen är vvr

Ex $\mu^*_{obs} = \dfrac{5x_3 + 3x_1 + 2x_2}{10}$ om $E(X_i) = \mu$

$$E(\mu^*) = E\left[\frac{5X_3 + 3X_1 + 2X_2}{10}\right] = \frac{1}{10}\Big(5\mu + 3\mu + 2\mu\Big) = \frac{10\mu}{10} = \mu$$

d.v.s. skattningen är vvr

Effektivitet

Def Om vi har flera väntevärdesriktiga skattningar av θ så är den skattningen som har lägst varians den effektivaste skattningen av θ.

Ex

$$\mu^*_{obs} = \frac{x_1 + x_2 + x_3}{3}$$

$$\hat{\mu}_{obs} = \frac{5x_3 + 3x_1 + 2x_2}{10}$$

Vi antar att variansen för $X_i = \sigma^2 \ \forall \ i$ och X_i:na oberoende.

$$V[\mu^*] = V\left[\frac{X_1 + X_2 + X_3}{3}\right] = \frac{1}{9} \, 3 \, V(X_i) = \frac{V(X_i)}{3} = \frac{\sigma^2}{3}$$

$$V[\hat{\mu}] = V\left[\frac{5X_3 + 3X_1 + 2X_2}{10}\right] = \frac{1}{100}\left[5^2 V(X_3) + 3^2 V(X_1) + 2^2 V(X_2)\right] =$$

$$= \frac{38}{100}\sigma^2$$

$V(\mu^*) < V(\hat{\mu})$ d.v.s. μ^*_{obs} är den effektivaste skattningen av μ.

Effektivitet: ett jämförande (/komparativt) mått

Medelfel

Def En skattning av $D(\theta^*)$ kallas medelfelet för θ^* och kallas $d(\theta^*)$ Se § 9.3 i F.S.

Medelfelet för skattningen av θ är alltså $D^*(\theta^*)_{obs}$

Ex Anta att X_i:na $\in N(\mu, \sigma)$. Vi har $\mu^*_{obs} = \bar{x}$. Vad är medelfelet för denna skattning?

Vi ska ha $D^*(\mu^*)_{obs}$

$$\mu^* = \bar{x} = \frac{1}{n}\sum X_i$$

$$D(\mu^*) = D(\bar{x}) = \sqrt{V(\bar{x})} = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}}$$

Medelfelet

$$D^*(\mu^*)_{obs} = \left(\frac{\sigma}{\sqrt{n}}\right)^*_{obs} = \frac{\sigma^*_{obs}}{\sqrt{n}} = \left[\sigma^*_{obs} = \sqrt{\text{stickprovsvariansen}} = s\right] = \frac{s}{\sqrt{n}} = \frac{1}{\sqrt{n}} \cdot \sqrt{\frac{1}{n-1}\sum(x_i - \bar{x})^2}$$

Ex $X \in Bin(n, p)$, $p^*_{obs} = \frac{x}{n}$. Vad är $d(p^*)$?

$$d(p^*) = D^*(p^*)_{obs}$$

$$D(p^*) = D\left(\frac{X}{n}\right)$$

$$V\left(\frac{X}{n}\right) = \frac{1}{n^2} V(X) = [\text{ se § 3 }] = \frac{1}{n^2} np(1-p) = \frac{p(1-p)}{n}$$

$$\Rightarrow D(p^*) = \sqrt{\frac{p(1-p)}{n}}$$

$$\text{Medelfelet} = D^*(p^*)_{obs} = \left(\sqrt{\frac{p(1-n)}{n}}\right)^*_{obs} = \sqrt{\frac{p^*_{obs}(1-p^*_{obs})}{n}} = \sqrt{\frac{\frac{x}{n}(1-\frac{x}{n})}{n}}$$

Ex $X \in Po(\mu)$

Se F.S. § 3 $\Rightarrow E(X) = \mu$, $\mu^*_{obs} = \bar{x}$

$$\text{Medelfelet} = D^*(\mu^*)_{obs}$$

$$D(\mu^*) = \sqrt{V(\mu^*)}$$

$$V(\mu^*) = V(\bar{X}) = V\left[\frac{1}{n} \Sigma X_i\right] = [\text{ober.}] = \frac{1}{n^2} n V(X_i) = \frac{V(X_i)}{n} = [X_i \in Po(\mu), \text{ se § 3}] =$$

$$= \frac{\mu}{n}$$

$$\Rightarrow D(\mu^*) = \sqrt{\frac{\mu}{n}}$$

$$D^*(\mu^*)_{obs} = \left(\sqrt{\frac{\mu}{n}}\right)^*_{obs} = \sqrt{\frac{\mu^*_{obs}}{n}} = \sqrt{\frac{\bar{x}}{n}}$$

KAP. 12

Konfidensintervall

Def Ett intervall I_θ som med sannolikhet $1-\alpha$ täcker över θ kallas för ett konfidensintervall för θ med konfidensgrad $1-\alpha$.

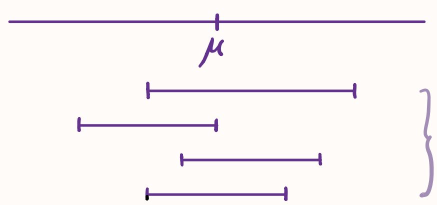
D.v.s. $P(\theta \in I_\theta) = 1-\alpha$

Ex Anta att vi vill skatta μ med $\bar{x}$ och att vi vill bilda ett konf-int så att vi med 95%-ig sannolikhet täcker över μ.

$I_\mu = [\,\bar{x}-a,\ \bar{x}+a\,]$

$P(\bar{X}-a < \mu < \bar{X}+a) = 0.95$

Observera att det är $\bar{X}$ som är slumpmässigt!

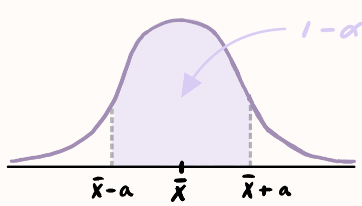


Då vi slumpar $\bar{X}$ får vi en massa konfidensintervall

Om vi slumpar $\bar{X}$ 100 ggr så kommer konfidensintervallen i genomsnitt täcka över μ 95 ggr.

Varför ser våra konfidensintervall i § 12.1, § 12.2, § 12.3 ut som de gör?

Ex Anta att X_i:na ober. och $N(\mu, \sigma)$. Vi gör n st mätningar och vill bilda ett konf-int av graden $1-\alpha$ för μ. Vi antar att σ är känt. Vi skattar μ med $\bar{x}$, $\mu^*_{obs} = \bar{x}$.



Vi ska hitta a.

$P(\mu \text{ finns i intervallet}) = 1-\alpha \quad \Rightarrow \quad P(|\bar{X}-\mu| < a) = 1-\alpha$

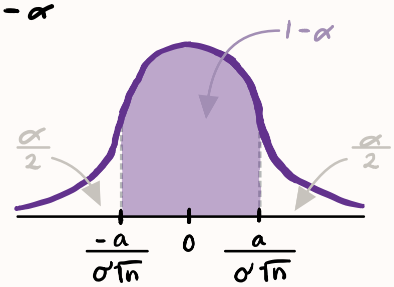
$P(-a < \bar{X}-\mu < a) = 1-\alpha$

$P(\mu-a < \bar{X} < \mu+a) = 1-\alpha$

Gör om till $N(0,1)$:

$$P\left(\frac{\mu-a-\mu}{\sigma/\sqrt{n}} < \underbrace{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}}_{Y \in N(0,1)} < \frac{\mu+a-\mu}{\sigma/\sqrt{n}}\right) = 1-\alpha$$

$$P\left(-\frac{a}{\sigma/\sqrt{n}} < Y < \frac{a}{\sigma/\sqrt{n}}\right) = 1-\alpha$$



Vi har en sannolikhet och söker gränsen $\Rightarrow$ anpassa till tab. 2:

"$P(Y > \lambda_x) = \alpha$"

Fig. $\Rightarrow$ $P\left(Y > \frac{a}{\sigma/\sqrt{n}}\right) = \frac{\alpha}{2}$ $\Rightarrow$ $\frac{a}{\sigma/\sqrt{n}} = \lambda_{\alpha/2}$ $\Rightarrow$ $a = \frac{\sigma}{\sqrt{n}}\lambda_{\alpha/2}$

$$\Rightarrow I_\mu = \bar{x} \pm \frac{\sigma}{\sqrt{n}}\lambda_{\alpha/2}$$

Användning av § 12

Vi ska bilda konf-int för väntevärde när σ är känd $\Rightarrow$ § 12.1

Enl. § 12.1 gäller att $I_\theta = \theta^*_{obs} \pm D \cdot \lambda_{\alpha/2}$

I vårt fall:

$\theta = \mu$

$\theta^*_{obs} = \mu^*_{obs} = \bar{x}$

Enl. § 12 gäller $\theta^* \in N(\theta, D)$

I vårt fall:

$D = D(\theta^*) = D(\mu^*) = D(\bar{X}) = [\text{ se § } 11.1\,a] = \frac{\sigma}{\sqrt{n}}$

$$\Rightarrow I_\mu = \bar{x} \pm \frac{\sigma}{\sqrt{n}}\lambda_{\alpha/2}$$

Konfidensintervallets bredd beror på tre saker

1) En stor spridning hos mätdata (stort σ) ger bredare intervall

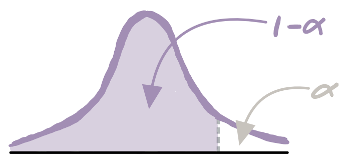
2) Ju fler observationer (stort n) ju smalare konf-int

3) Ju större säkerhet vi vill ha ju bredare intervall ($\lambda_{\alpha/2}$)

Ensidigt konfidensintervall

I exemplet ovan fås det ensidigt uppåt resp. ensidigt uppåt begränsade konf-intervallet till:

$$\left(-\infty,\ \bar{x} + \frac{\sigma}{\sqrt{n}}\lambda_\alpha\right)$$



resp.

$$\left(\bar{x} - \frac{\sigma}{\sqrt{n}}\lambda_\alpha,\ \infty\right)$$