Kom ihåg: Vi hade $X_1, X_2, \ldots$ iid med okänd parameter θ. Vi (punkt-)skattar θ mha $\hat{\theta} = \hat{\theta}(X_1, \ldots, X_n)$. Här är $\hat{\theta}$ en s.v.

Efter mätning då data $x_1, x_2, \ldots, x_n$ har blivit observerade blir den (numeriska) realiserade gissningen (av θ) $\hat{\theta}(x_1, \ldots, x_n)$.

OBS: Ofta kallas både $\hat{\theta}(X_1, \ldots, X_n)$ och $\hat{\theta}(x_1, \ldots, x_n)$ för skattning (estimate). Vissa skiljer på de två (skattning/skattare att estimate/estimator)

Def: För $Z \sim N(0,1)$ och $0 \le \alpha \le 1$ definierar vi z_α via $\mathbb{P}(Z \ge z_\alpha) = \alpha$

Anm: Symmetri ger $\mathbb{P}(Z \le -z_\alpha) = \alpha$

$\Rightarrow \mathbb{P}(-z_\alpha \le Z \le z_\alpha) = 1 - 2\alpha \leftarrow$ vanlig!

Sats: Låt $X_1 \sim N(\mu_1, \sigma_1^2)$ och $X_2 \sim N(\mu_2, \sigma_2^2)$ vara oberoende $\Rightarrow X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

B: $M_{X_i}(t) = e^{\mu_i t + \frac{1}{2}\sigma_i^2 t^2}$ $\quad i = 1, 2$

$$\Rightarrow \quad M_{X_1+X_2}(t) \underset{\text{Ober.}}{=} M_{X_1}(t)\, M_{X_2}(t) =$$

$$= e^{\mu_1 t + \frac{1}{2}\sigma_1^2 t^2} e^{\mu_2 t + \frac{1}{2}\sigma_2^2 t^2} = e^{(\mu_1+\mu_2)t + \frac{1}{2}(\sigma_1^2+\sigma_2^2)t^2} \qquad \square$$

Vi har lärt oss hur vi gör en punktskattning, men hur bra är den?

Konfidensintervall (K.I. fall 1):

Antag $X_1, X_2, \ldots$ iid $N(\mu, \sigma^2)$ med σ känd. Vi skattar μ med $\hat{\mu}(X_1,\ldots,X_n) = \overline{X}$.

Vi vet att $X_1 + \cdots + X_n \sim N(n\mu, n\sigma^2)$ s.a.

$$\overline{X} = \frac{X_1 + \cdots + X_n}{n} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \Rightarrow \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1).$$

[Kom ihåg $X \sim N(\mu, \sigma^2) \Rightarrow \frac{X-\mu}{\sigma} \sim N(0,1)$]

Vi får

$$1 - 2\alpha = \mathbb{P}(-z_\alpha \le Z \le z_\alpha) = \mathbb{P}\left(-z_\alpha \le \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \le z_\alpha\right)$$

$$= \mathbb{P}\left(-z_\alpha \frac{\sigma}{\sqrt{n}} \le \overline{X} - \mu \le z_\alpha \frac{\sigma}{\sqrt{n}}\right) = \mathbb{P}\left(\overline{X} - z_\alpha \frac{\sigma}{\sqrt{n}} \le \mu \le \overline{X} + z_\alpha \frac{\sigma}{\sqrt{n}}\right)$$

Dvs. med sann. $1-2\alpha$ tillhör μ det slumpmässiga intervallet

$$I_\mu = \left[\bar{X} - z_\alpha \frac{\sigma}{\sqrt{n}}, \bar{X} + z_\alpha \frac{\sigma}{\sqrt{n}}\right].$$

Detta är ett K.I. för μ med <u>konfidensgrad</u> $1-2\alpha$.

Efter att data $x_1, x_2, \ldots, x_n$ har observerats bildas det <u>numeriska</u> K.I.

$$I_\mu = \left[\bar{x} - z_\alpha \frac{\sigma}{\sqrt{n}}, \bar{x} + z_\alpha \frac{\sigma}{\sqrt{n}}\right].$$

<u>Def:</u> Låt $\underline{\theta} < \bar{\theta}$ vara skattningar av θ. Om $\mathbb{P}\left(\underline{\theta}(\bar{X}) \leq \theta \leq \bar{\theta}(\bar{X})\right) = 1-\alpha$ sägs intervallet $I_\theta = \left[\underline{\theta}(\bar{X}), \bar{\theta}(\bar{X})\right]$ vara ett K.I. för θ med konfidensgrad $1-\alpha$.

<u>Anm:</u> i) $\underline{\theta}, \bar{\theta}$ är "under"- resp "över"- skattningar av θ.

ii) Om $\underline{\theta} = -\infty$ (alt $\bar{\theta} = \infty$) kallas intervallen $(-\infty, \bar{\theta}]$ (alt $[\underline{\theta}, \infty)$) <u>ensidiga</u> K.I. (annars tvåsidigt).

iii) Fall 1 fungerade för att

<u>referensvariabeln</u>

$$R := \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$$ innehöll μ och hade känd

<u>fördelning.</u>

<u>Fall 2:</u> Återigen är $X_1, X_2, \ldots$ iid $N(\mu, \sigma^2)$ men

här är σ^2 okänd. Vi skattar σ^2 med

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$$ och bildar

$$R = \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim \, ?$$

"För att komma vidare måste vi veta

fördelningen för R."

<u>Def:</u> En s.v. X med täthet

$$f(s) = c\left(1+\frac{s^2}{n}\right)^{-(n+1)/2}$$

är t-fördelad med n frihetsgrader

($X \sim t(n)$).

TMA 321

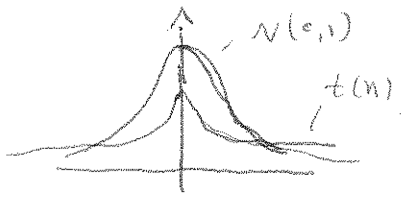
$\boxed{F16}$ 5

<u>Sats:</u> För $X_1, \dots, X_n$ iid $N(\mu, \sigma^2)$

gäller att $$R = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t(n-1).$$

<u>B:</u> Orgie i mgf, se kap 6 $\square$

<u>Anm:</u> 1) $t(n)$ är symmetrisk



2) Om $X_n \sim t(n) \Rightarrow X_n \xrightarrow[n \to \infty]{d} Z$ där $Z \sim N(0,1)$.

(Dvs $t(n) \to N(0,1)$)

3) Vi använder tabeller för $t_\alpha(n)$ där

$\alpha = \mathbb{P}(T \geq t_\alpha(n))$ om $T \sim t(n)$.

Åter till fall 2:

Med $R = \dfrac{\bar{X} - \mu}{S/\sqrt{n}} \sim t(n-1)$ får vi att

$$1 - \alpha = \mathbb{P}(R \leq t_\alpha(n-1)) = \mathbb{P}\left(\frac{\bar{X} - \mu}{S/\sqrt{n}} \leq t_\alpha(n-1)\right)$$

$$= \mathbb{P}\left(\mu \geq \bar{X} - \frac{S}{\sqrt{n}} t_\alpha(n-1)\right)$$ så att

$$I_\mu = \left[\bar{X} - \frac{S}{\sqrt{n}} t_\alpha(n-1), \infty\right)$$ är ett

ensidigt K.I. för μ med konf. grad $1-\alpha$.

<u>Anm:</u> Om $\mathbb{P}(\underline{\mu}(\mathbf{X}) < \mu < \overline{\mu}(\mathbf{X})) = 1-\alpha$

$$\Rightarrow \mathbb{P}\left(\frac{1}{\overline{\mu}(\mathbf{X})} < \frac{1}{\mu} < \frac{1}{\underline{\mu}(\mathbf{X})}\right) = 1-\alpha \quad \text{s.a.}$$

$\left[\frac{1}{\overline{\mu}(\mathbf{X})}, \frac{1}{\underline{\mu}(\mathbf{X})}\right]$ är ett K.I. för $\frac{1}{\mu}$

med konf. grad $1-\alpha$.

<u>Tal:</u> Utfallen av en serie experiment anses vara $N(\mu, \sigma^2)$. Med $n=13$ datapunkter erhölls

2,81 5.14 2.39 2.63 7.47 7.23 7.25

5.01 −0.62 5.15 7.89 4.47 6.1

a) Skatta μ, σ^2

b) Hitta 99 % (tvåsidigt) K.I. för μ.

<u>L:</u> a) Då $X_1, \ldots, X_{13} \sim N(\mu, \sigma^2)$ (oberoende!?)

använder vi $\hat{\mu} = \overline{X}$ och $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$

som är VVR. Insättning av data ger

$$\hat{\mu}(x) = \overline{x} \approx 4.84 \quad , \quad s^2(x) \approx 6.2082.$$

b) Vi har att $\frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t(n-1) = t(12)$

Tabell ger

$$0.99 = \mathbb{P}\left(-3.05 \leq \frac{\overline{X} - \mu}{S/\sqrt{n}} \leq 3.05\right)$$

$$= \mathbb{P}\left(\overline{X} - 3.05\frac{S}{\sqrt{n}} \leq \mu \leq \overline{X} + 3.05\frac{S}{\sqrt{n}}\right)$$

s.a. ett 99 % numeriskt K.I. för μ

blir $I_\mu = 4.84 \pm 3.05 \frac{\sqrt{6.21}}{\sqrt{13}} = [2.73, 6.95]$.