Föreläsning 7: 7.4, 8.1 – 8.2

- Konfidensintervall, Generell definition
- Konf. Int för
  - I) Väntevärde
  - II) Varians
- Relaterade fördelningar

Vi har metoder för att hitta punktskattningar, M.M., M.L., ~~mos~~ för parametrar men vi vet inte hur säkra dessa är.

Def (Konfidensinterval)
Låt X vara en s.v. och låt θ vara en parameter som beskriver en egenskap hos X. (väntevärde, varians, λ i Expl.)
Ett ~~konfidensintervall~~ $100(1-\alpha)\%$-igt konfidensintervall ges av ett slumpmässigt intervall $[L, U]$ s.a.

$$\mathbb{P}(L \leq \theta \leq U) = 1-\alpha$$

Ann:

I) L o U är slumpvariabler som i allmänhet kommer att bero på ett stickprov $(X_1, \dots, X_n)$ av X.

II) Det finns i huvudsak två metoder för att härlede konfidensintervall.

A) Pivotering

B) Likelihoodkvoter.

Härledningarna i boken bygger på A.
Generella idén för A är som följer:
Låt θ vara en parameter vi är intresserad av.
~~Låt $\hat{\theta}$ vara en~~.

Steg 1) Antag att vi har en väntevärdesriktig skattare av θ, d.v.s. $E[\hat{\theta}(X_1, X_2, \dots, X_n)] = \theta$, ~~dan~~
$X_1, X_2, \dots, X_n$ är ett stickprov av X.

Steg 2:

Hitta en 'pivotvariabel' som beror på $\hat{\theta}$:

D.v.s.

Hitta en s.v. $Y = Y(\hat{\theta})$ s.a. $P(Y \le y)$ inte beror på θ.

Ex Om $X \sim N(\mu, 1)$ så är $X - \mu$ en pivotvariabel för μ.

Sats: Låt $\alpha \in (0,1)$ och låt $z_{\alpha/2}$ vara definerad av

$$\mathbb{P}(Z_1 \le z_{\alpha/2}) = 1 - \frac{\alpha}{2} \quad \text{där } Z_1 \sim N(0,1).$$

Låt $X_1, X_2, \dots, X_n$ vara ett stickprov från en normalfördelning med väntevärde μ o varians σ^2.

Ett $100(1-\alpha)\%$-igt konfidensintervall för μ ges av

$$\left[\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}},\ \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right] \overset{\text{skrivs}}{=} \bar{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

Anm: Detta konf int antar att σ är känt.

Mera Ex:

En stålfabrik producerar skruvar med varierande vikt. Variationen/standardavvikelsen kan bestämmas till $\sigma = 0{,}025$ g.

Konstruera ett 95%-igt konfidensintervall för den genomsnittliga vikten, givet $\bar{x} = 0{,}9988$, $n = 10$

Lösn: 95%-igt $\Rightarrow \alpha = 0{,}025$, $z_{0{,}975} = 1{,}960$.

$z_{\alpha/2}$ fås ur tabell VI på s. 700, $\gamma = \infty$

$\Rightarrow$ konf int ges av

$$0{,}9988 \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 0{,}9988 \pm 1{,}960 \cdot \frac{0{,}025}{\sqrt{10}} = [0{,}9833, 1{,}0143]$$

I vissa situationer är man bara intresserad av en övre/undre gräns, ex. max last, minsta mängd näring vid odling.

<u>Def</u>: Låt X vara en s.v och låt θ vara en parameter.

Ett ~~konfidens~~ nedre/övre $100(1-\alpha)\%$-igt konfidens intervall definieras av

I) $[L, \infty)$, nedre $\quad P(L \leq \theta) = 1-\alpha$

II) $(-\infty, U]$, övre $\quad P(\theta \leq U) = 1-\alpha$

där L & U är två slumpvariabler.

<u>Sats</u>:

Låt $\alpha \in (0,1)$ och definiera z_α genom

$$P(Z \leq z_\alpha) = 1-\alpha, \quad Z \sim N(0,1)$$

Låt $X_1, X_2, \ldots, X_n$ vara ett stickprov av $N(\mu, \sigma^2)$.

Ett $100(1-\alpha)$-%igt konf. int. för μ ges av

övre: $\left(-\infty, \bar{X} + z_\alpha \frac{\sigma}{\sqrt{n}}\right]$

undre: $\left[\bar{X} - z_\alpha \frac{\sigma}{\sqrt{n}}, \infty\right)$

<u>Om σ är okänt görs följande:</u>

<u>Def</u> (T-fördelning).

X är T-fördelad med parameter $\gamma > 0$ om X har täthet

$$f(x) = c(\gamma)\left(1 + \frac{x^2}{\gamma}\right)^{-(\gamma+1)/2}, \quad x \in \mathbb{R}, \gamma > 0$$

där $c(\gamma)$ definieras av

$$\int f(x) = 1.$$

Parametern γ kallas ofta för 'frihetsgrader' 'degrees of freedom'

Sats: Låt $\alpha \in (0,1)$, och låt $t_{\alpha/2}(n-1)$ vara definierad av

$$P(Z_1 \leq t_{\alpha/2}(n-1)) = 1 - \frac{\alpha}{2}$$

där Z_1 är T-fördelad med parameter $n-1$. Låt $X_1, \dots, X_n$ vara ett stickprov från $N(\mu, \sigma^2)$. Ett $100(1-\alpha)\%$-igt konfidensintervall ges av

$$\overline{X} \pm t_{\alpha/2}(n-1) \frac{S}{\sqrt{n}}$$

där $S = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2} = \sqrt{S^2}$

är stickprovsvarians.

Ex: Gör om stålfabriksexemplet med σ okänd.

$\overline{x} = 0{,}9988$, $\alpha = 0{,}025$, $t_{\alpha/2}(10-1) = 2.262$ Tabell VI s. 699

$r = n-1 = 9$

Gå till rad 9

och kolumn 0,975

Simulering: KI för μ, σ känt vs σ okänt.

$N(0,1)$ vs T-fördelning ($\gamma > 0$).

Anm:

Generellt så är inte data normalfördelad, men centrala gränsvärdessatsen låter oss approximera konfidensintervall.

Konfidensintervall för σ^2:

Def

X är χ^2-fördelad med parameter $k = 1, 2, \ldots$

om X har tätheten

$$f(x) = c(k)\, x^{k/2-1} e^{-x/2}, \quad x > 0.$$

Obs f är ej symmetrisk.

Sats: Låt $\alpha \in (0,1)$ och definera $\chi^2_{\alpha/2}$, $\chi^2_{1-\alpha/2}$ genom

$$P(Y \leq \chi^2_{\alpha/2}) = 1 - \frac{\alpha}{2}, \quad P(Y \leq \chi^2_{1-\frac{\alpha}{2}}) = \frac{\alpha}{2}$$

där Y är χ^2-fördelad med parameter $n-1$.

Låt $X_1, X_2, \ldots, X_n$ vara ett stickprov från $N(\mu, \sigma^2)$.

Ett $100(1-\alpha)\%$-igt konf. int. för σ^2 ges av

$$\frac{(n-1)\,\delta^2}{\chi^2_{\alpha/2}} \leq \sigma^2 \leq \frac{(n-1)\,\delta^2}{\chi^2_{1-\alpha/2}}$$

där δ^2 är stickprovsvariansen.

Simulering av KI $\simeq$ χ^2-tätheten.

Räkne Ex: ~~[illegible]~~

Givet följande observerade stickprov

$x = (5.4, -1.6, 2.4, -0.2, 3.2, -0.4, 4.0, 5.0, 8.8, 1.2)$, $n = 10$

$\delta^2 = 10{,}0573$

Beräkna ett 95%-igt konf intervall för σ^2.

Lösn: $\alpha = 0{,}025$, Tabel IV: $\nu = n-1 = 9$, $F = 0{,}025 = 0{,}975$

$$\underset{\substack{\uparrow \\ F = 0{,}975}}{\chi^2_{0{,}025}} = 19{,}0, \quad \underset{\substack{\uparrow \\ F = 0{,}025}}{\chi^2_{0{,}975}} = 2{,}7$$

Konf. int för σ^2 ges därmed av

$$\left[\frac{9 \cdot 10{,}0573}{19{,}0}, \frac{9 \cdot 10{,}0573}{2{,}7}\right] \approx [4.8, 33.5]$$

Härledning av konfidensinterval av μ då σ är känt:

Låt $X_1, \dots, X_n$ vara ett stickprov av $N(\mu, \sigma^2)$.

Då vet vi att $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ är en väntevärdesriktig skattare av μ

dvs

$$\mathbb{E}[\overline{X}] = \mu.$$

Vidare så vet vi att eftesom $X_i \sim N(\mu, \sigma^2)$ så gäller

$$\overline{X} \sim N(\mu, \sigma^2/n).$$

Standardisering av $\overline{X}$ ger att

$$\frac{\overline{X} - \mu}{\sqrt{\sigma^2/n}} = \underbrace{\frac{\overline{X} - \mu}{\sigma/\sqrt{n}}}_{\sim N(0,1)} \text{ är vår pivot varuchel.}$$

Låt nu $z_{\alpha/2}$ vara s.a.

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

V.L.:

$$P\left(\left| \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \right| \leq z_{\alpha/2} \right) \overset{\text{def av absolutbelopp}}{=} P\left(-z_{\alpha/2} \leq \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \leq z_{\alpha/2} \right)$$

$$= \text{// Multiplicera } \frac{\sigma}{\sqrt{n}} \text{ plån. //} = P\left(-\frac{\sigma}{\sqrt{n}} z_{\alpha/2} \leq \overline{X} - \mu \leq \frac{\sigma}{\sqrt{n}} z_{\alpha/2} \right)$$

$$= \text{// subtrahera } \overline{X} \text{ och multiplicera med } -1 \text{ //} = P\left(-\overline{X} - \frac{\sigma}{\sqrt{n}} z_{\alpha/2} \leq -\mu \leq -\overline{X} + \frac{\sigma}{\sqrt{n}} z_{\alpha/2} \right)$$

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