Konfidensintervaller og vurdering af usikkerhed på estimerede størrelser

Rasmus Waagepetersen

October 26, 2018

Motivation

Eksempel: En landmåler får til opgave at måle længden λ fra A til B.

Entreprenøren ønsker resultatet med en nøjagtighed på $\pm$3mm.

Eksempel: Lægemiddelstyrelsen kræver, at et lægemiddelselskab estimerer sandsynlighed θ for bivirkninger af en medicin med en nøjagtighed på $\pm$ 0.01%

Problem: Hvordan fortolker og håndterer vi disse specifikationer i praksis, hvor alle målinger er behæftet med en (stokastisk) usikkerhed ?

Specifikationer af nøjagtighed under hensyntagen til tilfældighed

I praksis beregner landmåleren et estimat $\hat{\lambda}$ på basis af et antal målinger $l_1, \ldots, l_n$, f.eks.

$$\hat{\lambda} = \frac{1}{n} \sum_{i=1}^{n} l_i$$

I praksis er målingerne behæftede med målefejl e_i,

$$l_i = \lambda + e_i$$

Vi opfatter hver e_i som en realisation af en stokastisk variabel E_i og lader $L_i = \lambda + E_i$ angive de tilsvarende stokastiske udfald af målingerne.

Dermed (med lidt misbrug af notation) kan vi også anskue

$$\hat{\lambda} = \frac{1}{n} \sum_{i=1}^{n} L_i$$

som en stokastisk variabel.

Vi kan nu omformulere entreprenørens specifikation:

fejlen $|\hat{\lambda} - \lambda|$ skal være mindre end 3mm med en meget stor sandsynlighed - f.eks. 99.9%.

Dvs. forestiller man sig, at længdebestemmelsen blev gentaget et stort antal (hypotetiske) gange, så må fejlen kun overstige 3mm i 0.1% af gentagelserne.

Dette giver naturligvis ikke en absolut sikkerhed for længdebestemmelsen i et konkret tilfælde ! - men det vil sjældent gå galt...

Landmålerens opgave er nu at beregne sandsynligheden for, at $|\hat{\lambda} - \lambda| \geq 3$ og tjekke, at denne sandsynlighed er lille nok

Dvs. han har brug for at kende sandsynlighedsfordelingen af $\hat{\lambda}$ - anskuet som en stokastisk variabel.

Lidt beregninger med middelværdi og varians

Hvis X er en stokastisk variabel med middelværdi og varians

$$\mathbb{E}X = \mu \quad \mathbb{V}\mathrm{ar}X = \sigma^2$$

så gælder for et reelt tal a,

$$\mathbb{E}aX = a\mathbb{E}X \quad \mathbb{V}\mathrm{ar}(aX) = a^2\mathbb{V}\mathrm{ar}X$$

Hvis $X_1, \ldots, X_n$ er stokastiske variable med middelværdier μ_i og varianser σ_i^2 gælder

$$\mathbb{E}\sum_i X_i = \sum_i \mu_i$$

Hvis X_i'erne yderligere er *uafhængige*:

$$\mathbb{V}\mathrm{ar}\sum_i X_i = \sum_i \sigma_i^2$$

Normalfordelte stokastiske variable

Hvis X_i'erne er normalfordelte så er

$$aX_i \quad og \quad X. = \sum_i X_i$$

også normalfordelte

(middelværdi og varians for $X.$ bestemmes jf forrige slide)

Centrale grænseværdi-sætning

Hvis X_i'erne er uafhængige og identisk fordelte med fælles middelværdi μ og varians σ^2, så gælder for gennemsnittet

$$\bar{X}. = \frac{1}{n}X.$$

at

$$\bar{X}. \approx N(\mu, \frac{\sigma^2}{n})$$

eller ækvivalent,

$$\frac{1}{\sqrt{n}}(X. - n\mu) \approx N(0, \sigma^2) \quad \sqrt{n}(\bar{X}. - \mu) \approx N(0, \sigma^2)$$

Approksimationerne bliver bedre jo større n !

Tilbage til landmåleren

Antag fejlene E_i er uafhængige med middelværdi nul (dvs ingen systematisk målefejl) og kendt varians σ^2

Dermed bliver målingerne $L_1, \dots, L_n$ uafhængige med middelværdi λ og varians σ^2.

Bemærk:

$$|\hat{\lambda} - \lambda| \leq 3 \Leftrightarrow -3 \leq \hat{\lambda} - \lambda \leq 3$$

Han/hun kan nu benytte CLT hvorved $\hat{\lambda} - \lambda \approx N(0, \sigma^2/n)$

Dvs. skal blot udregne sandsynligheden for, at en normalfordelt stokastisk variabel med middelværdi 0 og varians σ^2/n ligger mellem -3 og 3 (Excel, Maple, TI-89,...)

Ukendt σ^2

Hvis σ^2 er ukendt kan estimatet

$$s^2 = \frac{1}{n-1} \sum_i (L_i - \hat{\lambda})^2$$

anvendes i stedet

Tilfører ekstra approksimation til resultaterne (udover evt. brug af CLT).

Kvantificering af estimations-usikkerhed

Variansen af $\hat{\lambda}$ er σ^2/n.

Jo mindre varians des mindre er den forventede værdi af fejlen $\hat{\lambda} - \lambda$.

I praksis benyttes ofte standardafvigelsen

$$\text{sd} = \sqrt{\mathbb{V}\mathrm{ar}\hat{\lambda}} = \frac{\sigma}{\sqrt{n}}$$

som et mål for usikkerheden.

Intervaller og sandsynligheder

Hidtil: sandsynlighed for at estimationsfejl mindre end givet øvre grænse.

Nu: givet sandsynlighed, f.eks. 95%, hvad er så den øvre grænse ?

For en normalfordelt stokastisk variabel X med middelværdi μ og varians σ^2 gælder: sandsynlighed for at

$$|X - \mu| \leq k\sigma$$

er

$$95\%(k = 1.96) \quad 95.4\%(k = 2) \quad 99.7\%(k = 3) \quad 99.99\%(k = 4)$$

Bemærk, at vi direkte kan omsætte k til sandsynligheder, hvis vi bruger standardafgivelse σ som enhed.

Konfidensinterval

Med f.eks. $k = 1.96$ er der 95% sandsynlighed for

$$|X - \mu| \leq 1.96\sigma \Leftrightarrow -1.96\sigma \leq X - \mu \leq 1.96\sigma$$

Dette giver anledning til to forskellige intervaller:
a) (for X)

$$\mu - 1.96\sigma \leq X \leq \mu + 1.96$$

b) (for μ)

$$X - 1.96\sigma \leq \mu \leq X + 1.96\sigma$$

Bemærk: intervallet svarende til (b) er stokastisk - kaldes et 95% konfidensinterval for μ (baseret på data X)

Konfidensinterval fortsat

Antag $X_1, \ldots, X_n$ uafhængige og identisk fordelte med fælles middelværdi μ og varians σ^2.

Da er gennemsnit $\bar{X}.$ approksimativt $N(\mu, \sigma^2/n)$ og vi benytter $\bar{X}.$ som estimat for μ.

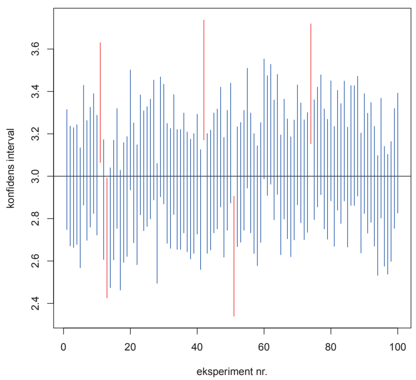
Jf forrige slide definerer

$$\bar{X}. - 1.96\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X}. + 1.96\frac{\sigma}{\sqrt{n}}$$

et 95% konfidensinterval for μ.

Simulationsstudie

100 simulerede 95% konfidensintervaller baseret på $\bar{X}.$ gennemsnit af 100 normalfordelte variable med middelværdi 3 og varians 2



Konfidensintervallerne indeholder sande middelværdi $\mu = 3$ i ca. 95% af tilfældende.

Praktisk brug af konfidensinterval

Hvis vi hver gang vi udfører et eksperiment hævder, at den ukendte middelværdi ligger i det beregnede 95% interval, så tager vi kun fejl i 5% af tilfældende

Hvis vi vil have større sikkerhed kan vi i beregningen af konfidensintervallet erstatte $1.96\sigma/\sqrt{n}$ med $3\sigma/\sqrt{n}$ eller $4\sigma/\sqrt{n}$ - giver 99.7% eller 99.99% intervaller.

Da tager vi kun fejl i 0.3% eller 0.01% af tilfældene.

Men intervallerne bliver da naturligvis bredere, så vores udsagn bliver mere udvandede

Den lidt tricky skelnen

Den 95% sandsynlighed vedrører et *fremtidigt ikke endnu gennemførte* eksperiment

Hvis vi f.eks. observerer en stikprøve

$$2.3\ 2.3\ 0.1\ 3.1\ 0.9\ -0.1\ 2.1\ 2.4\ 3.9\ 0.0$$

fra 10 stokastisk variable med $\sigma = 2$ er det konkrete udfald af konfidensintervallet

$$[1.7 - \frac{2}{\sqrt{10}}; 1.7 + \frac{2}{\sqrt{10}}] = [1.1; 2.3]$$

Dette interval er naturligvis ikke stokastisk, og det giver derfor ikke mening at tillægge en 95% sandsynlighed til dette interval (sandsynligheden er enten 0 eller 100% men vi ved ikke hvilken)

Vanskeligt, og måske heller ikke nødvendigt, at få studerende til at fange denne pointe

Stokastiske intervaller er erfaringsmæssigt lidt vanskelige for studerende.

Måske nemmere at udnytte at

$$\bar{X}. - 1.96\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X}. + 1.96\frac{\sigma}{\sqrt{n}} \Leftrightarrow |\bar{X}. - \mu| \leq 1.96\frac{\sigma}{\sqrt{n}}$$

Dvs. med 95% sandsynlighed ligger estimationsfejlen i det ikke-stokastiske interval $\pm 1.96\sigma/\sqrt{n}$

Ukendt varians

Vi har jf CLT

$$Z = \frac{\bar{X}. - \mu}{\sigma/\sqrt{n}} \approx N(0,1)$$

Dvs. Z er (approksimativt) en *pivot*-størrelse: fordelingen af Z afhænger hverken af μ eller σ.

Hvis σ erstattes af et estimat $\hat{\sigma}$ gælder i almindelighed stadig

$$\frac{\bar{X}. - \mu}{\hat{\sigma}/\sqrt{n}} \approx N(0,1)$$

hvorved

$$\left[\bar{X}. - 1.96\frac{\hat{\sigma}}{\sqrt{n}}; \bar{X}. - 1.96\frac{\hat{\sigma}}{\sqrt{n}}\right]$$

er et approksimativt 95% interval.

Dvs. hvis variansen er ukendt, kan vi i almindelighed blot erstatte den ukendte varians med et estimat, når vi udregner et 95% konfidensinterval.

Konfidensinterval for sandsynlighedsparameteren i en binomial fordeling

Antag Y er binomialfordelt med kendt antalsparameter n og ukendt sandsynlighedsparameter θ.

Da estimerer vi θ ved andelen $\hat{\theta} = Y/n$.

Mht konfidensinterval er der ikke noget nyt under solen da

$$Y = I_1 + \cdots + I_n$$

hvor I_i'erne er uafhængige binære variable med sandsynlighed $P(I_i = 1) = \theta$.

Dvs $\hat{\theta} = Y/n$ er gennemsnittet $\bar{I}. = \frac{1}{n}\sum_i I_i$

Hvis n ikke er for lille kan vi jf CLT (og slide 19) bruge at

$$\frac{\hat{\theta} - \theta}{\hat{\sigma}/\sqrt{n}} \approx N(0,1)$$

hvor

$$\hat{\sigma}^2 = \hat{\theta}(1 - \hat{\theta})$$

er et estimat af

$$\sigma^2 = \mathbb{V}\mathrm{ar} I_i = \theta(1 - \theta)$$

Derved bliver

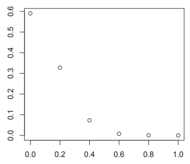
$$[\hat{\theta} - 1.96\frac{\hat{\sigma}}{\sqrt{n}}; \hat{\theta} + 1.96\frac{\hat{\sigma}}{\sqrt{n}}]$$

et (approksimativt) 95% konfidensinterval for θ

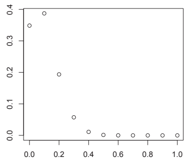
Hvornår er n stor nok ?

Fordeling af $\hat{\theta}$ for forskellige værdier af n og θ:

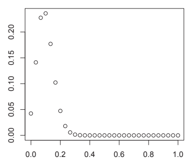
$\theta = 0.1$ $n = 5$



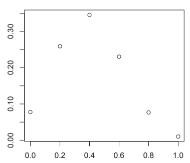
$\theta = 0.1$ $n = 10$



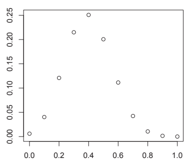
$\theta = 0.1$ $n = 30$



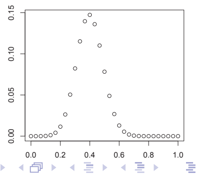
$\theta = 0.4$ $n = 5$



$\theta = 0.4$ $n = 10$



$\theta = 0.4$ $n = 30$



Generelt set-up

θ parameter i en parametrisk statistisk model.

Hvis $\hat{\theta}$ er (approksimativt) normalfordelt med middelværdi θ og spredning $\mathrm{sd}(\hat{\theta})$ så er

$$\hat{\theta} \pm 1.96\mathrm{sd}(\hat{\theta})$$

et (approksimativt) 95% konfidensinterval for θ.

(analogt med $\hat{\mu} = \bar{X}.$ og $\mathrm{sd}(\hat{\mu}) = \sigma/\sqrt{n}$)

Lineær regression

Antag $Y_1, \ldots, Y_n$ følger en regressionsmodel

$$Y_i = \alpha + \beta x_i + E_i$$

hvor $x_1, \ldots, x_n$ er kendte og E_i er uafhængige normalfordelte med middelværdi 0 og varians σ^2

Da er mindste kvadraters metode estimaterne

$$\hat{\beta} = \frac{\sum_i Y_i x_i - n\bar{Y}.\bar{x}.}{\sum_i (x_i - \bar{x}.)^2} \quad \hat{\alpha} = \bar{Y}. - \hat{\beta}\bar{x}.$$

Disse estimater er normalfordelte ! (lineære funktioner af data)

Konfidensintervaller for α og β

$\hat{\alpha}$ og $\hat{\beta}$ har varianser

$$\mathbb{V}\mathrm{ar}\hat{\alpha} = \sigma^2 \frac{\sum_i x_i^2}{n \sum_i (x_i - \bar{x}_\cdot)^2} \quad \mathbb{V}\mathrm{ar}\hat{\beta} = \frac{\sigma^2}{\sum_i (x_i - \bar{x}_\cdot)^2}$$

σ^2 estimeres ved

$$\frac{1}{n-2} \sum_i (Y_i - \hat{\mu}_i)^2$$

hvor $\hat{\mu}_i = \hat{\alpha} + \hat{\beta} x_i$ er estimatet af $\mu_i = \mathbb{E}[Y_i]$

Dermed lige ud af landevejen at konstruere 95% konfidensintervaller for α og β vha generel opskrift: estimat plus/minus 1.96 standardafvigelsen af estimatet

Konfidensinterval for $\mu_i = \mathbb{E} Y_i$

μ_i estimeres ved $\hat{\mu}_i = \hat{\alpha} + \hat{\beta} x_i$

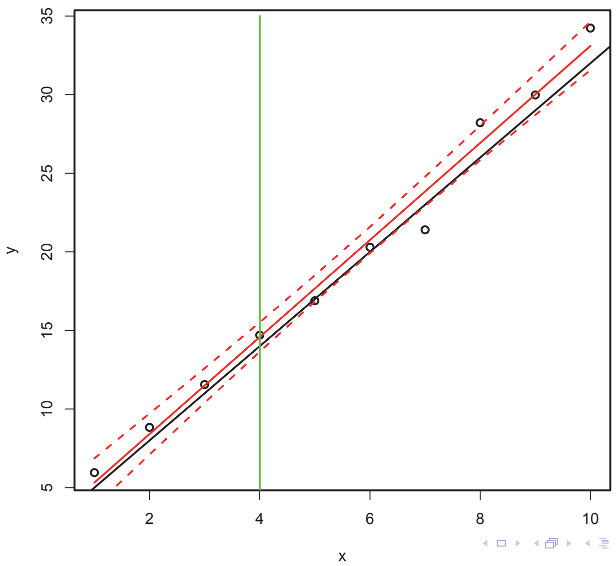
Variansen for $\hat{\mu}_i$:

$$\mathbb{V}\mathrm{ar}\hat{\mu}_i = \mathbb{V}\mathrm{ar}\hat{\alpha} + x_i^2 \mathbb{V}\mathrm{ar}\hat{\beta} + 2 x_i \mathbb{C}\mathrm{ov}(\hat{\alpha}, \hat{\beta}) \tag{1}$$

hvor

$$\mathbb{C}\mathrm{ov}(\hat{\alpha}, \hat{\beta}) = -\sigma^2 \frac{\bar{x}_\cdot}{\sum_i (x_i - \bar{x}_\cdot)^2}$$

Igen kan generel opskrift benyttes: estimat plus/minus 1.96 standardafvigelsen af estimatet.



Prædiktionsinterval for ny observation

Antag, at vi har estimeret en regressionslinje baseret på en stikprøve $(y_1, x_1), \ldots, (y_n, x_n)$.

Lad Y_{n+1} være en uobserveret variabel som vi gerne vil forudsige på baggrund af den observerede værdi x_{n+1} af den tilhørende forklarende variabel.

Vores bedste bud på Y_{n+1} er $\hat{\mu}_{n+1} = \hat{\alpha} + \hat{\beta} x_{n+1}$

Usikkerhed på prædiktion af Y_{n+1} vha $\hat{\mu}_{n+1}$ kvantificeres vha *prædiktionsinterval*

Der gælder

$$Y_{n+1} - \hat{\mu}_{n+1} = Y_{n+1} - \mu_{n+1} + \mu_{n+1} - \hat{\mu}_{n+1} = E_{n+1} + \mu_{n+1} - \hat{\mu}_{n+1}$$

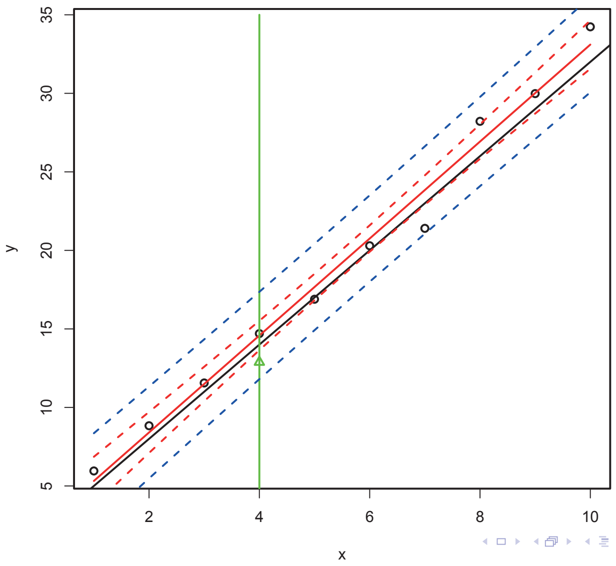
Her er E_{n+1} normalfordelt med middelværdi 0 og varians σ^2 og uafhængig af estimationsfejlen $\mu_{n+1} - \hat{\mu}_{n+1}$, som er normalfordelt med middelværdi nul og varians $\mathbb{V}\mathrm{ar}\hat{\mu}_{n+1}$ beregnet tidligere.

Dermed er $Y_{n+1} - \hat{\mu}_{n+1}$ normalfordelt med middelværdi nul og varians $\omega^2 = \sigma^2 + \mathbb{V}\mathrm{ar}\hat{\mu}_{n+1}$.

Dvs med 95% sandsynlighed gælder

$$-1.96\omega \leq Y_{n+1} - \hat{\mu}_{n+1} \leq 1.96\omega \Leftrightarrow \hat{\mu}_{n+1} - 1.96\omega \leq Y_{n+1} \leq \hat{\mu}_{n+1} + 1.96\omega$$

Dette interval er altid bredere end konfidensintervallet for μ_{n+1} da det både afspejler usikkerheden vedr. estimationen af μ_{n+1} samt variationen af Y_{n+1} omkring μ_{n+1}



Eksempel

```
> fit=lm(y~x)
> summary(fit)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   2.0933     0.7791   2.687   0.0276 *
x             2.9977     0.1256  23.873 1.01e-08 ***
---
```

95% konfidensinterval for β: $2.99 \pm 1.96 \cdot 0.1256 = [2.74; 3.24]$

Specielt kan vi bemærke, at konfidensintervallet ikke indeholder nul. Dvs. vi kan afvise at $\beta = 0$ og Y_i afhænger dermed signifikant af x_i.

Yderligere emner (til selvstudium)

- brug af eksakt fordeling af $\hat{\theta}$?
- konfidensinterval og transformation
- transformation for θ
- δ-metoden
- konfidensinterval for σ^2
- teoretisk middelværdi og varians
- **opgaver**

Opsummering

- For en ukendt parameter θ er et α-konfidensinterval et *stokastisk* interval som med sandsynlighed α indeholder θ
- Sandsynligheden refererer til et endnu ikke gennemført eksperiment
- Andelen af fremtidige eksperimenter hvor θ er indeholdt i konfidensintervallet knyttet til eksperimentet er ca. α.
- I almindelighed udregnes konfidensintervallet som $\hat{\theta} \pm k\text{sd}(\hat{\theta})$ hvor $k = 1.96, 3, 4$ svarer til $\alpha = 95\%, 99.7\%, 99.99\%$.
- For et konkret datasæt er konfidensintervallet *ikke stokastisk* - vi kan ikke tillægge sandsynlighed til et konkret udregnet konfidensinterval

Eksakt fordeling af $\hat{\theta}$

Forskellen $\hat{\theta} - \theta$ er fordelt som

$$\frac{Y}{n} - \theta$$

hvis fordeling let kan tabelleres.

Problem: fordelingen afhænger af ukendte θ - vi kan ikke finde nedre og øvre grænser, der ikke afhænger af θ.

Derimod er

$$\frac{\hat{\theta} - \theta}{\hat{\sigma}/\sqrt{n}} \approx N(0, 1)$$

som med 95% ssh sandsynlighed ligger indenfor grænserne -1.96 og 1.96, der ikke afhænger af θ.

Dermed bliver normalfordelingsapproksimationen for $\hat{\theta} - \theta$ en bekvem genvej til at etablere et konfidensinterval

Konfidensinterval og transformation

Antag $[L, U]$ er et α konfidensinterval for θ (f.eks. $\alpha = 95\%$)

Lad

$$\eta = g(\theta)$$

hvor g er en injektiv transformation

Da er $[g(L); g(U)]$ (eller $[g(U); g(L)]$) et α konfidensinterval for η

Anvendelse for binomial fordeling

For sandsynlighedsparameteren θ i en binomialfordeling benyttede vi estimatet

$$\hat{\theta} = \frac{Y}{n}$$

og konfidensintervallet

$$\hat{\theta} \pm 1.96\text{sd}(\hat{\theta})$$

hvor vi estimerede $\text{sd}(\hat{\theta})$ ved $\hat{\theta}(1 - \hat{\theta})/n$.

Problem: dette interval er ikke nødvendigvis indeholdt i $[0, 1]$!

Logit-transformation

Løsning: parameteren

$$\eta = g(\theta) = \log\left(\frac{\theta}{1-\theta}\right)$$

har variationsområde $]-\infty, \infty[$ når θ varierer i $]0,1[$ (og omvendt !)

Vi kan danne konfidensinterval for η og dernæst transformere tilbage vha den inverse transformation

$$\theta = g^{-1}(\eta) = \frac{\exp(\eta)}{1+\exp(\eta)}$$

Vores estimat for η er

$$\hat{\eta} = g(\hat{\theta})$$

og variansen approksimeres ved (δ-metoden)

$$\mathbb{V}\mathrm{ar}\hat{\eta} \approx (g'(\theta))^2 \mathbb{V}\mathrm{ar}\hat{\theta} = \frac{1}{n\theta(1-\theta)}$$

Konfidensintervallet for η bliver

$$\hat{\eta} \pm 1.96\sqrt{\frac{1}{n\theta(1-\theta)}}$$

og konfidensintervallet for θ bliver

$$\left[g^{-1}\left(\hat{\eta} - 1.96\sqrt{\frac{1}{n\theta(1-\theta)}}\right); g^{-1}\left(\hat{\eta} + 1.96\sqrt{\frac{1}{n\theta(1-\theta)}}\right)\right]$$

som med garanti er indeholdt i $[0,1]$!

δ-metoden

Lad $X = g(Y)$ hvor Y er en stokastisk variabel med middelværdi μ og varians σ^2 og g er differentiabel.

Da gælder (lineær approksimation)

$$X \approx g(\mu) + g'(\mu)(Y-\mu) \Rightarrow \mathbb{V}\mathrm{ar}X \approx (g'(\mu))^2 \mathbb{V}\mathrm{ar}Y$$

(igen vha. regneregler for middelværdi og varians)

Konfidensinterval for varians

For en lineær regressionsmodel med normalfordelte observationer, kan man vise, at

$$(n-2)\frac{\hat{\sigma}^2}{\sigma^2}$$

er $\chi^2(n-2)$ fordelt (og dermed en pivot-størrelse).

Dermed gælder med 95% sandsynlighed, at

$$\chi^2_{0.025}(n-2) \leq (n-2)\frac{\hat{\sigma}^2}{\sigma^2} \leq \chi^2_{0.975}(n-2)$$

hvor $\chi^2_p(n-2)$ angiver p-fraktilen for en $\chi^2(n-2)$-fordeling $(0 \leq p \leq 1)$.

Dette giver konfidensintervallet

$$\left[\hat{\sigma}^2\frac{n-2}{\chi^2_{0.975}(n-2)}; \hat{\sigma}^2\frac{n-2}{\chi^2_{0.0.025}(n-2)}\right]$$

for σ^2.

Opgaver

1. Check resultaterne på slide 8 når $\approx$ er erstattet med $=$ og X_i'erne antages at være normalfordelte.
2. (vedr. slide 9) Hvad gør landmåleren hvis sandsynligheden ikke er under 99.9% ?
3. Check (jf slide 21) at
   3.1
   $$\mathbb{V}\mathrm{ar} I_i = \theta(1-\theta)$$
   3.2
   $$\hat{\theta}(1-\hat{\theta}) = \frac{1}{n}\sum_i (I_i - \bar{I}.)^2$$
   Dvs variansestimatet er $(n-1)/n$ gange det sædvanlige variansestimat
   $$s^2 = \frac{1}{n-1}\sum_i (I_i - \bar{I}.)^2$$
4. Antag $\theta = 0.1$. Hvor stor skal n være, for at bredden på konfidensintervallet bliver mindre end 0.02 ?

4. Mindste kvadraters estimaterne af α og β kan udregnes som
   $$\begin{bmatrix}\hat{\alpha}\\ \hat{\beta}\end{bmatrix} = (X^{\mathsf{T}}X)^{-1}X^{\mathsf{T}}Y$$
   hvor X er $n \times 2$ matricen med søjler $(1, 1, \ldots, 1)^{\mathsf{T}}$ og $(x_1, \ldots, x_n)^{\mathsf{T}}$ og $Y = (Y_1, \ldots, Y_n)^{\mathsf{T}}$. Brug dette til at udregne kovariansmatricen for $(\hat{\alpha}, \hat{\beta})^{\mathsf{T}}$

   Bemærk: kovarians-matricen Σ for en stokastisk vektor Z har indgange $\Sigma_{ij} = \mathbb{C}\mathrm{ov}(Z_i, Z_j)$. Hvis A er en matrix af passende dimension, så har AZ kovariansmatrix $A\Sigma A^{\mathsf{T}}$.
5. Vis (1) på slide 26 ved at bruge resultatet og vinket fra forrige opgave.
6. (slide 35) Vis at $[g(L); g(U)]$ er et α konfidensinterval for η når g er voksende
7. Check resultaterne på slide 37-39