2018 05 21 Linjär regression

Modell: vi antar ett linjärt samband $y=\beta_0+\beta_1 x$ som modelleras av $\underline{Y}_i=\beta_0+\beta_1 x_i+\varepsilon_i$ där

1) $x_1, x_2, \dots, x_n$ är givna kända tal. De kallas för förklarande variabler.
2) $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ är ober. s.v. (störningar/fel) s.a. $E[\varepsilon_i]=0$ & $Var(\varepsilon_i)=\sigma^2$. Vi antar dessutom att $\varepsilon_i \sim N(0,\sigma^2)$.
3) Vi vill skatta β_0 = intercept & β_1 = slope (lutningskoeff.).
4) $\underline{Y}_1, \dots, \underline{Y}_n$ kallas för bero. variabler.
5) Linjen $y=\beta_0+\beta_1 x$ kallas för den teoretiska regressionslinjen.

Vi skattar β_0, β_1 genom att välja $\hat{\beta}_0, \hat{\beta}_1$ s.a.

$$Q(\beta_0,\beta_1)=\sum_{k=1}^{n}(y_k-(\beta_0+\beta_1 x_k))^2$$

minimeras. Här är $(x_1,y_1),\dots,(x_n,y_n)$ våra datapkter.

Vi minimerar $Q(\beta_0,\beta_1)$ genom att lösa $\frac{\partial Q}{\partial \beta_0}=\frac{\partial Q}{\partial \beta_1}=0$.

Om vi låter $S_{xy}:=\sum_{k=1}^{n}(x_k-\bar{x})(y_k-\bar{y})$, $S_{xx}=\sum_{k=1}^{n}(x_k-\bar{x})^2$, $S_{yy}=\sum_{k=1}^{n}(y_k-\bar{y})^2$

får vi $\hat{\beta}_1=\frac{S_{xy}}{S_{xx}}$ & $\hat{\beta}_0=\bar{y}-\hat{\beta}_1\bar{x}$.

Vår anpassade regressionslinje blir $\hat{y}=\hat{\beta}_0+\hat{\beta}_1 x$

Som ett mått på anpassningen beräknas förklaringsgraden

$$R^2=1-\frac{\sum_{k=1}^{n}(y_k-\hat{y}_k)^2}{\sum_{k=1}^{n}(y_k-\bar{y})^2}=\dots=\frac{S_{xy}^2}{S_{xx}S_{yy}}$$

Vi har att $R^2 \in (0,1)$ & att $R^2 \approx 1$ motsvarar att data stöder ett linjärt samband vdl.
Om $R^2 \approx 0$ så inte.

Under antagandet att $\varepsilon_i \sim N(0, \sigma^2)$ & iid kan man visa att
$\hat{\beta}_1 = \hat{\beta}_1(\varepsilon_1, \dots, \varepsilon_n) \sim N(\beta_1, \frac{\sigma^2}{S_{xx}})$ & $\hat{\beta}_0 = \hat{\beta}_0(\varepsilon_1, \dots, \varepsilon_n) \sim N(\beta_0, \sigma^2 \frac{\sum_1^n x_k^2}{n S_{xx}})$

OBS! $E[\hat{\beta}_i] = \beta_i$ så båda är VVR.
Med kännedom om fördelningarna kan vi skapa K.I.

$$1-2\alpha = P(-z_\alpha \le Z \le z_\alpha) = P\left(-z_\alpha \le \frac{\hat{\beta}_1 - \beta_0}{\sigma/\sqrt{S_{xx}}} \le z_\alpha\right)$$

$$\Rightarrow I_\beta = \hat{\beta}_1 \pm z_\alpha \frac{\sigma}{\sqrt{S_{xx}}}$$

Ex (F-halt) med insättning av data & om $\sigma = 1$ får vi

$$R^2 = \frac{S_{xy}^2}{S_{xx} S_{yy}} \approx 0.9436, \quad \hat{\beta}_1 \approx N(\beta_1, 0.1257) \text{ medan } \hat{\beta}_0 = N(\beta_0, 0.3627)$$

Oftast är σ^2 okänd. Isf. skattar vi σ^2 med $S_r^2 = \frac{1}{n-2}\sum_{k=1}^n (y_k - \hat{y}_k)^2 =$

$$= \frac{1}{n-2}\sum_{k=1}^n (y_k - \hat{\beta}_0 - \hat{\beta}_1 x_k)^2$$

Anm: 1) S_r^2 är VVR

2) $s^2 = \frac{1}{n-1}\sum_{k=1}^n (y_k - \bar{y})^2$ är ej bra här!

Som innan när σ skattas får vi $\frac{\hat{\beta}_1 - \beta_1}{S_r/\sqrt{S_{xx}}} \sim t(n-2)$ &

$\frac{\hat{\beta}_0 - \beta_0}{S_r \sqrt{\frac{\sum_{k=1}^n x_k^2}{n S_{xx}}}} \sim t(n-2)$ dvs. t-fördelningar.

Ex (F-halt) om σ^2 är okänd skattar vi & får $s_r^2 \approx 0.239$
Ett 95% K.I. för β_1 blir då (med $t_{0.025}(6) \approx 2.45$)

$$I_{\beta_1} = \hat{\beta}_1 \pm 0.45 \frac{s_r}{\sqrt{S_{xx}}} \approx [1.31, 2.16] \leftarrow 95\% \text{ numeriskt K.I.}$$

Vi kan använda K.I. för att göra hypotestest.

Ex (F-halt) vi vill testa $H_0: \beta_1 = 0$, $H_1: \beta_1 \neq 0$, på signif. 0.05. Enl. förra gången kan vi välja RR s.a. $\hat{\beta}_1 \in RR \iff 0 \notin I_{\beta_1}$, där I_{β_1} är ett K.I. för β_1 med samma värde på α.
Vi har enl. ovan att $0 \notin I_{\beta_1} \approx [1.31, 2.16]$, så vi förkastar H_0 på signif. 0.05.

Rimlighetsanalys av modellen

Betrakta residualerna $e_k = y_k - \hat{\beta}_0 - \hat{\beta}_1 x_k = y_k - \hat{y}_k$.
Om vi ser mönster i residulalerna bör vi vara skeptiska till modellantagandet.

Q018 05 22 Gamla tentauppgifter

1) Låt X vara en s.v. med täthetsfunk. $f(x)=(x+C_1)e^{-C_0x}$ $x\geq 0$

a) Bestäm ett samband mellan C_1 & C_0 s.a. $f(x)$ verkligen är en tfkn. Bestäm eventuella restriktioner på C_1, C_0.

Använd att $\int_{-\infty}^{\infty} f(x)dx=1$ & $f(x)\geq 0$ $\forall x\in\mathbb{R}$.

Vi har att:

$$1=\int_{-\infty}^{\infty} f(x)\,dx=\int_0^{\infty}(x+C_1)e^{-C_2x}\,dx \overset{P.I.}{=}\left[(x+C_1)\frac{e^{-C_0x}}{C_0}\right]_0^{\infty}+\frac{1}{C_0}\int_0^{\infty}e^{-C_0x}\,dx=$$

$$=\frac{C_1}{C_0}+\frac{1}{C_0}\left[\frac{e^{-C_0x}}{C_0}\right]_0^{\infty}=\frac{C_1}{C_0}+\frac{1}{C_0^2}.$$

Vi får att $C_1=C_0-1/C_0$. Vi har att $C_1\geq 0 \Rightarrow C_0^2\geq 1 \Rightarrow C_0\geq 1$ //

b) Beräkna $\mathbb{E}[X^n]$ för alla värden på $n\in\mathbb{N}$.

Vi har att:

$$\mathbb{E}[X^n]=\int_{-\infty}^{\infty}x^n f(x)\,dx=\int_0^{\infty}(x^{n+1}+C_1x^n)e^{-C_0x}\,dx$$

$$\int_0^{\infty}x^{n+1}e^{-C_0x}\,dx=\left[x^{n+1}\frac{e^{-C_0x}}{-C_0}\right]_0^{\infty}+\int_0^{\infty}\frac{n+1}{C_0}x^n e^{-C_0x}\,dx=\int_0^{\infty}\frac{n+1}{C_0}x^n e^{-C_0x}\,dx$$

$$\Rightarrow \mathbb{E}[X^n]=\left(\frac{n+1}{C_0}+C_1\right)\int_0^{\infty}x^n e^{-C_0x}\,dx=\left(\frac{n+1}{C_0}+C_1\right)\left(\left[x^n\frac{e^{-C_0x}}{-C_0}\right]_0^{\infty}+\int_0^{\infty}nx^{n-1}\frac{e^{-C_0x}}{C_0}\,dx\right)$$

$$=\left(\frac{n+1}{C_0}+C_1\right)\left(0+\frac{n}{C_0}\int_0^{\infty}x^{n-1}e^{-C_0x}\,dx\right)=\ldots=\left(\frac{n+1}{C_0}+C_1\right)\frac{n!}{C_0^n}\int_0^{\infty}e^{-C_0x}\,dx$$

$$=\left(\frac{n+1}{C_0}+C_1\right)\frac{n!}{C_0^n}\ //$$

Mgf: $M_X(t)=M_X(0)+M_X'(0)t+M_X''(0)\frac{t^2}{2!}+\ldots+M_X^{(k)}(0)\frac{t^k}{k!}+\ldots$

$M_X^{(k)}(0)=\mathbb{E}[X^k]$

1) beräkna mgf, 2) Taylorutv. 3) läs av $M_X^{(k)}(0)$.

Alt. lösning till b)

2) Betrakta händelserna A, B, C.

a) Antag att $\mathbb{P}(A \cap B | C) = 0.4$, $\mathbb{P}(B \cap C) = 0.5$ & $\mathbb{P}(C) = 0.9$. Beräkna $\mathbb{P}(A | B \cap C)$

L Vi har att:

$$\mathbb{P}(A | B \cap C) = \frac{\mathbb{P}(A \cap B \cap C)}{\mathbb{P}(B \cap C)} = \frac{\mathbb{P}(A \cap B | C)\,\mathbb{P}(C)}{\mathbb{P}(B \cap C)} = \frac{0.4 \cdot 0.9}{0.5} = \frac{36}{60} \;//$$

b) Finns det tre händelser A, B, C s.a. $\mathbb{P}(A \cap B \cap C) = 0.24$, $\mathbb{P}(C) = 0.9$ & $\mathbb{P}(B | C) = 0.2$?

L Svar: Nej! Vi får att:

$$\mathbb{P}(A \cap B | C) = \frac{\mathbb{P}(A \cap B \cap C)}{\mathbb{P}(C)} = \frac{0.24}{0.9}, \; 0.2 = \mathbb{P}(B | C),$$

vilket motsäger att $A \cap B \subset B$ //

---

3) Antag att s.v. U har sannolikhetsfunktion

$$\mathbb{P}(U = k) = \frac{2k}{N(N+1)}, \quad k = 1, 2, \ldots, N$$

där N är en parameter. Antag vidare att den betingade fördelningen för X givet $U = k$ är $Exp(k)$.

a) Antag att $X_1, \ldots, X_n$ är iid alla med samma fördelning som X. Hitta MME'n (momentskattaren) för N uttryckt i $X_1, \ldots, X_n$.

b) Låt $U_1, \ldots, U_n$ vara iid med samma fördelning som U. Hitta MLE'n (maximum likelihood-skattaren) för N uttryckt i $U_1, \ldots, U_n$.

L a) Vi ska (enl. momentmetoden) beräkna $\mathbb{E}[X]$. Vi kan hitta marginalfördelningen & sedan använda def. av väntevärde.

Alt. $\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|U]]$ - ★

$\boxed{\mathbb{E}[X|U=k] = \frac{1}{k} \quad (\text{ty } X \text{ givet } U=k \text{ är Exp}(k))}$

$\boxed{\Rightarrow \mathbb{E}[X|U] = \frac{1}{U}}$

$$\bigstar = \mathbb{E}[1/U] = \sum_{k=1}^{N} \frac{1}{k} \mathbb{P}(U=k) = \sum_{k=1}^{N} \frac{1}{k} \frac{2k}{N(N+1)} = \frac{2}{N+1}$$

momentmetoden: $\bar{X} = \frac{2}{\hat{N}+1} \Rightarrow \hat{N} = \frac{2}{\bar{X}} - 1$ //

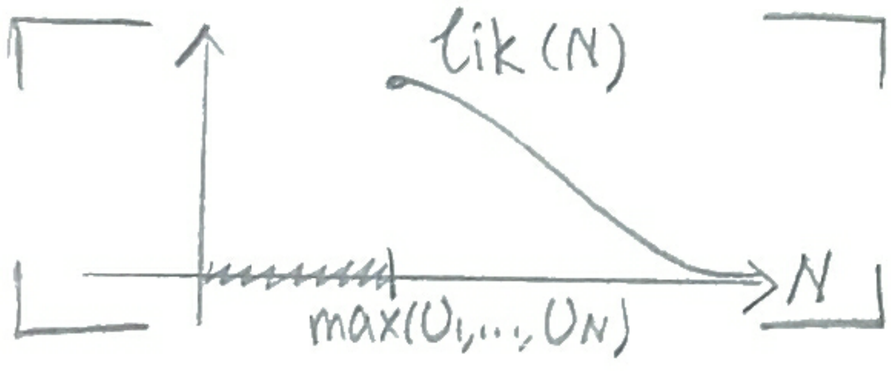
b) Vi ska hitta likelihooden & maximera den. Likelihooden ges av den gemensamma sf för $U_1, \dots, U_n$.

$$lik(N) = f(U_1, \dots, U_n | N) \overset{\text{oberoende}}{=} \prod_{k=1}^{n} f(U_k|N) = \prod_{k=1}^{n} \frac{2U_k}{N(N+1)} I(U_k \in \{1, \dots, N\}) =$$

$$\boxed{\mathbb{P}(U=k) = \frac{2k}{N(N+1)} \underbrace{I(1 \le k \le N)}_{\text{indikatorfunk.}} = \frac{2k}{N(N+1)} I(k \in \{1, \dots, N\})}$$

$$= \frac{2^n}{N^n (N+1)^n} U_1 \cdot U_2 \dots U_n I(U_1, U_2, \dots, U_n \in \{1, \dots, N\}).$$

Vilket N maximerar $lik(N)$? Jo, $\hat{N} = \max(U_1, \dots, U_n)$, vilket där vår MLE.



---

4)

x (dag)	1	2	3	4	5	6	7	8	9	10
y (deltagare)	120	118	112	112	108	106	97	90	80	78

11 12

78 64

a) Vi ska skatta β_0 & β_1 och använder $\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}$ & $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$.
Med insatta data får vi $\hat{\beta}_1 \approx -5,213$ & $\hat{\beta}_0 \approx 130,3$.

c) Vi delar upp data i 2 delar: vecka 1 resp. 2. Vi får då:

$y = -2.857x + 122.67$ v. 1

$y = -6.314x + 140.15$ v. 2

2018 05 23 Räkneövning 8
Linjär regression & delar av gammal tenta

14.13 En linje är passad till n observationer med minsta-kvadrat-metoden. Anta att statistiska standardmodellen håller. Vi vill skatta värdet vid en ny pkt x_0. Kalla värdet μ_0. Skattaren är $\hat{\mu}_0 = \hat{\beta}_0 + \hat{\beta}_1 x_0$.

a) Ange ett uttryck för $\mathrm{Var}(\hat{\mu}_0)$ som funk. av $x_0 - \bar{x}$.

2 "Standardmodellen": $\mu_i = \beta_0 + \beta_1 x_i$, $i = 1, \dots, n$

$e_i \sim N(0, \sigma^2)$, ober. x_i deterministiska variabler (ej stokastiska)

$$\mathrm{Var}(\hat{\mu}_0) = \mathrm{Var}(\hat{\beta}_0 + \hat{\beta}_1 x_0) = \mathrm{Var}(\hat{\beta}_0) + x_0^2 \mathrm{Var}(\hat{\beta}_1) + 2x_0 \mathrm{cov}(\hat{\beta}_0, \hat{\beta}_1).$$

Sats B: $\mathrm{Var}(\hat{\beta}_0) = \dfrac{\sigma^2 \sum x_i^2}{n\sum x_i^2 - (\sum x_i)^2}$, $\mathrm{Var}(\hat{\beta}_1) = \dfrac{n\sigma^2}{n\sum x_i^2 - (\sum x_i)^2}$,

$$\mathrm{cov}(\hat{\beta}_0, \hat{\beta}_1) = \frac{-\sigma^2 \sum x_i}{n\sum x_i^2 - (\sum x_i)^2}.$$

Kort räkning (analogt med $\mathbb{E}(X - \mathbb{E}(X)) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$)

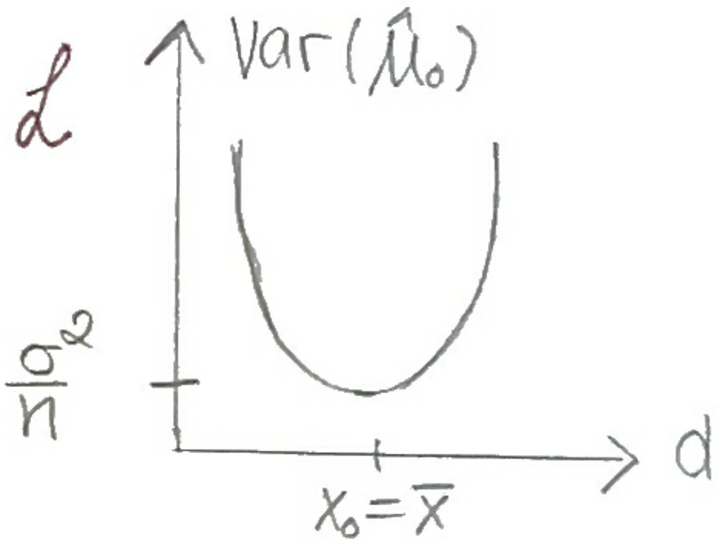
$$\sum x_i^2 - \frac{1}{n}\left(\sum x_i\right)^2 = \sum (x_i - \bar{x})^2$$

$$= \frac{\sigma^2\left(\sum x_i^2 - \frac{1}{n}\left(\sum x_i\right)^2\right) + \frac{\sigma^2}{n}\left(\sum x_i\right)^2 + x_0^2 n\sigma^2 - 2x_0\sigma^2 \sum x_i}{n\left(\sum x_i^2 - \frac{1}{n}\left(\sum x_i\right)^2\right)} =$$

$$= \frac{\sigma^2}{n} + \frac{\sigma\left(\left(\frac{1}{n}\sum x_i\right)^2 + x_0^2 - 2x_0 \frac{1}{n}\sum x_i\right)}{\sum (x_i - \bar{x})^2} = \frac{\sigma^2}{n} + \frac{\sigma^2(x_0^2 - 2x_0\bar{x} + \bar{x}^2)}{\sum (x_i - \bar{x})^2} =$$

$$= \frac{\sigma^2}{n} + \frac{\sigma^2 (x_0 - \bar{x})^2}{\sum (x_i - \bar{x})^2} = \sigma^2\left(\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum (x_i - \bar{x})^2}\right)$$

b) Skissa $\mathrm{Var}(\hat{\mu}_0)$ som funk. av $d = x_0 - \bar{x}$



c) Hitta ett 95% K.I. för μ_0

$\hat{\mu}_0$, β_1 normalfördelade då de är linjärkomb. av e_i, som ober. normalförd. $\Rightarrow \hat{\mu}_0$ normalförd.

$\Rightarrow \frac{\hat{\mu}_0 - \mu_0}{S_{\hat{\mu}_0}} \sim t_{n-2}$ (kap 6. lemma B),

där $S_{\hat{\mu}_0} = \sqrt{s^2\left(\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum (x_i - \bar{x})^2}\right)}$ $\qquad s^2 = \frac{1}{n-2}\sum (y_i - \hat{y}_i)^2$

$\Rightarrow K.I. = \hat{\mu}_0 \pm S_{\hat{\mu}_0} t_{n-2}(0.025)$

Tenta 20100816

5. Dioximhalt i sjöar

sjö nr.	1	2	3	4 ...	11
halt (ppm)	13,74	9,04	...		

Anta att förekomsten av dioxider är ober, likaförd. $N(\mu, \sigma^2)$

a) Skatta stickprovsmedelvärde & varians.

$\hat{\mu} = \bar{x} = \frac{1}{n}\sum x_i = 9.856$, $\sigma^2 = s^2 = \frac{1}{n-1}\sum (x_i - \bar{x})^2 \approx 20.89$

b) Sätt upp lämpligt hypotestest för att testa om $\mu > 8$.

$H_0: \mu = 8$, $H_1: \mu > 8$.

c) Kan vi förkasta nollhypotesen på 5%-nivån?

Under H_0 gäller $T = \frac{\bar{x} - 8}{\sqrt{s^2/11}} \sim t_{10}$

$\bar{x}$ & s ovan ger $t = \frac{9.856 - 8}{\sqrt{20.89/11}} \approx 1,346$

Tabell ger $0.05 = \mathbb{P}(T \geq 1.812)$

$\Rightarrow \mathbb{P}(T \geq t) > 0.05 \Rightarrow$ kan ej förk. H_0!

a) Vad är p-värdet?

$\mathbb{P}(T \geq t) = \mathbb{P}(T \geq 1.372) \approx 0.1 = p$

---

6. Låt $X_1, \dots, X_n$ vara ober. likaförd. s.v. som tar värden $\{-1, 0, 2\}$ med sanno. $1/2$, $1/4$ resp. $1/4$.

a) Hitta mgf.

$M_X(t) = \mathbb{E}(e^{tX}) = \frac{1}{2}e^{-t} + \frac{1}{4}e^{0t} + \frac{1}{4}e^{2t} = \frac{1 + 2e^{-t} + e^{2t}}{4}$

b) Låt $S_n = \frac{\sum X_i}{\sqrt{n}}$ & hitta mgf för S_n

$M_{S_n}(t) = \mathbb{E}(e^{tS_n}) = \mathbb{E}(e^{\frac{t}{\sqrt{n}}\sum_{i=1}^{n} X_i}) = M_X(\frac{t}{\sqrt{n}})^n = \left(\frac{1 + 2e^{-t/\sqrt{n}} + e^{2t/\sqrt{n}}}{4}\right)^n$

c) Bestäm $\lim_{n \to \infty} M_{S_n}(t)$. Mot vilken fördelning konv. fördelningen för S_n?

$f(t) = e^{t/\sqrt{n}}$, $f'(t) = \frac{e^{t/\sqrt{n}}}{\sqrt{n}}$, $f'' = \frac{e^{t/\sqrt{n}}}{n}$ & $f''' = \frac{e^{t/\sqrt{n}}}{n^{3/2}}$

$f(t) = f(0) + f'(0)t + f''(0)\frac{t^2}{2} + O(n^{3/2}) = 1 + \frac{t}{\sqrt{n}} + \frac{t^2}{2n} + \dots$

$M_{S_n}(t) = \left(\frac{1 + 2(1 - \frac{t}{\sqrt{n}} + \frac{t^2}{2n} + O(n^{3/2})) + (1 + \frac{2t}{\sqrt{n}} + \frac{4t^2}{2n} + O(n^{3/2}))}{4}\right)^n =$

$= \left(1 + \frac{3t^2}{4n} + O(n^{3/2})\right)^n \to e^{3t^2/4}$.

mgf för $N(\mu, \sigma^2)$ är $e^{\mu t + \frac{1}{2}\sigma^2 t^2} \Rightarrow$ förd. konv. till $N(0, 3/2)$

---

7. Vi observerar följande

Utfall:	1	2	...	11	12	0
Frekvens:	10	22	...	1	1	2 (tot. 201 obser.)

Anta att utfallen är ober. & följer en Poissonförd. med param. λ

a) Skatta λ

$\mathcal{L}$ $X \sim Poi(\lambda) \Rightarrow \mathbb{E}(X) = \lambda$. $\mathbb{E}(x)$ kan skattas med $\bar{x}$. Välj skattaren $\hat{\lambda} = \bar{x} = \frac{2 \cdot 0 + 1 \cdot 10 + 2 \cdot 22 + \ldots + 12 \cdot 1}{201} = \frac{987}{201}$

b) Skatta standardfelet för skattaren från a)

$\mathcal{L}$ $Var(\hat{\lambda}) = Var(\bar{x}) = \frac{Var(X)}{n} = \frac{\lambda}{n}$.

Skatta $Var(\hat{\lambda})$ med $\hat{\sigma}^2 = \frac{\hat{\lambda}}{n} = \frac{987}{201^2}$.

Standardfelet skattas till $\hat{\sigma}_{\bar{x}} = \frac{\sqrt{987}}{201}$

c) Ge ett approx. K.I. för λ

$\mathcal{L}$ $\bar{x}$ innehåller summa av 201 ober. likafördel. variabler

$\Rightarrow \bar{x}$ approx. normalfördel.

$\bar{x} \overset{appr.}{\sim} N(\lambda, s_{\bar{x}}^2) \Rightarrow \frac{\bar{x} - \lambda}{s_{\bar{x}}} \sim N(0,1)$

$0,95 \approx \mathbb{P}(-1,96 \leq \frac{\bar{x} - \lambda}{s_{\bar{x}}} \leq 1.96) = \mathbb{P}(\bar{x} - 1,96 s_{\bar{x}} \leq \lambda \leq 1.96 s_{\bar{x}})$

$\Rightarrow$ K.I. $\approx [4.315, 4.909]$

---

8. 140 deltagare, 35 numrerade bord, alla bord har lika många platser.

a) På hur många sätt kan deltagare placeras?
Ta hänsyn till vilket bordsnr. man sitter vid.

$\mathcal{L}$ 4/bord, ett bord i taget.

Tilldelningarna för bord 1 kan göras på $\binom{140}{4}$ olika sätt.

Bord 2: $\binom{136}{4}$, bord i: $\binom{140 - 4(i-1)}{4}$ sätt.

Totalt $\binom{140}{4}\binom{136}{4}\binom{132}{4} \cdots \binom{4}{4} = \frac{140!\,136!\,132! \ldots 4!}{(4!)^{35}(140-4)!(136-4)! \ldots (4-4)!} =$

$= \frac{140!}{(4!)^{35}}$

b) Som a) men utan hänsyn till bordsnr.

∂ Borden kan permuteras på 35! olika sätt.

Antalet placeringar blir $\frac{140!}{(4!)^{35}35!}$

c) 20 deltagare är fransmän. Hur många placeringar finns då bord 1-5 ockuperas av enbart fransmän?

Ta hänsyn till bordsnr.

∂ Fransmännen kan placeras på $\frac{20!}{(4!)^5}$ sätt.

Övriga deltagare $\frac{120!}{(4!)^{30}}$

Totalt $\frac{20!\,120!}{(4!)^{35}}$

d) Samma som c) men vi kräver inte att fransmännen sitter på just bord 1-5, bara att de sitter på 5 bord.

∂ De 5 borden kan väljas på $\binom{35}{5}$ sätt.

Svaret blir $\binom{35}{5}\frac{20!\,120!}{(4!)^{35}}$