Lösningar till tentamen FMSF75 – matstat för W – 2019-01-07

Modell:

① Z = skillnad i sträcka (utan-med luftkond.)

$Z \in N(\Delta, \sigma)$

$z_i = 16 \quad 45 \quad 30 \quad 34 \quad 31$

a) $\Delta^*_{obs} = \bar{z} = 31.2$

$\sigma^*_{obs} = s = 10.3779$

$I_\Delta = \left(\bar{z} \pm \underbrace{t_{0.025}(4)}_{2.776} \frac{s}{\sqrt{5}}\right) = (18.31, 44.09)$ 95% intervall

b) Nej, eftersom intervallet täcker över 20.

② a) X_i = energiförbr. dygn i; $X_i \in N(2500, 100)$

En veckas totala energiförbr. $= \sum_{i=1}^{7} X_i$

$\sum_{i=1}^{7} X_i \in N(7 \cdot 2500, \sqrt{7} \cdot 100)$

$P\left(\sum_{i=1}^{7} X_i > 18000\right) = 1 - \Phi\left(\frac{18000 - 7 \cdot 2500}{\sqrt{7} \cdot 100}\right)$

$= 1 - \Phi(1.8898) = \underline{\underline{0.0294}}$

b) X = energiförbr. $\in N(2500, 100)$

Y = energitillf. $\in N(2400, 80)$

$X - Y \in N(2500 - 2400, \sqrt{100^2 + 80^2})$

$P(X - Y > 0) = 1 - \Phi\left(\frac{0 - 100}{\sqrt{100^2 + 80^2}}\right) = \underline{\underline{0.7826}}$

3) X = antalet av de 53 mätningarna som tillhör klass stark påverkan

a) $X \in Bin(53, p)$ där $p = P$(en mätning tillhör stark...)

$H_0: p \geq 0.48$ (som tidigare)

$H_1: p < 0.48$ (minskad försurningspåverkan)

Vi observerade 20 mätningar i klassen stark påverkan.

Direktmetoden:

P-värde $= P(X \leq 20$ då $X \in Bin(53, 0.48))$

$= 0.0867$

Eftersom P-värde > 0.05 (standardgräns) kan H_0 ej förkastas på nivå 5%.

Nej, dessa data stärker inte misstanken om minskad försurningspåverkan

b) $p^*_{obs} = \frac{20}{53}$

$P(\text{precis en tillhör klassen}) = \binom{2}{1} p^1 (1-p)^1 = 2p(1-p)$

$\approx 2 p^*_{obs}(1 - p^*_{obs}) = 2 \cdot \frac{20}{53} \cdot \frac{33}{53} = \underline{\underline{0.4699}}$

4 Modell: y = antal familjegrupper

a) x = år

$y_i = \alpha + \beta \cdot x_i + \varepsilon_i, \quad \varepsilon_i \in N(0, \sigma)$

Enligt regrni-kommandot blir

$I_\beta = (-7.338, 1.275)$ vilket innehåller 0.

Hypotesen $H_0: \beta = 0$ kan alltså ej förkastas och en signifikant trend kan ej påvisas.

Residualerna uppvisar inget mönster och tycks vara normalfördelade.

Eftersom β är ej signifikant skilt från 0 är modellen $y_i = \alpha + \varepsilon_i, \quad \varepsilon_i \in N(0, \sigma)$ förmodligen rimligare.

b) Sökt är ett konfidensintervall för $x = 2017$

Om linjära regressionsmodell används fås den från regrni till $(160.5, 232.2)$.

Om den enklare modellen $y_i = \alpha + \varepsilon_i$ används gäller det att göra ett konfidensintervall för α:

$$\left(\bar{y} \pm t_{0.025}^{(16-1)} \frac{s}{\sqrt{16}}\right) = \left(221.03 \pm 2.13 \cdot \frac{38.57}{\sqrt{16}}\right) =$$

där s fås genom std(y)

$= (200.48, 241.59)$

4 c) Sökt är ett prediktionsintervall för $6 \cdot y$ då $x = 2018$.

Om den linjära regressionsmodellen anv. fås intervallet från reggui: $6 \cdot (105.6,\ 284.9) = (633.6,\ 1709.4)$

Om den enklare modellen används kan prediktionsintervallet approximeras med

$$6 \cdot \left(\bar{y} \pm t_{0.025}^{(16-1)} \cdot s\right) = 6 \cdot (221.03 \pm 2.13 \cdot 38.57)$$

$$= 6 \cdot (138.81,\ 303.24) = (832.9,\ 1819.5)$$

d) En skattad autokorrelationsfunktion på gruppantal tyder inte på att mätningarna av gruppantal är beroende.
Inte heller residualerna från den linjära regressionsmodellen tycks vara beroende.

5

(5) Modell: $x_1, \ldots x_{30}$ obs av X = antal dygn med mindst 40 mm

$X \in Po(\lambda)$

OBS! Av misstag finns det 40 observationer i filen punkt-nederbörd.mat. Modellen baseras alltså på 40 år: $x_1, \ldots x_{40}$ obs. av X = antal dygn med minst 40 mm.

a) $\lambda^*_{obs} = \bar{x} = 13.45$

b) $P(X \geq 20) = 1 - P(X \leq 19) = 0.0562$ $\quad X \in Po(13.45)$

c) Man vill testa hypoteserna

$H_0: \lambda \leq 13.45$

$H_1: \lambda > 13.45$

Antag att k är det lägsta antal då H_0 förkastas på nivå 1%.

Det innebär att P-värde =

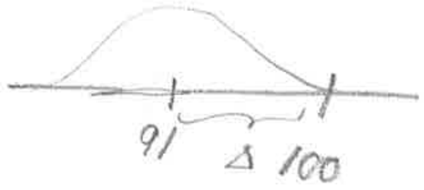
$P(X \geq k$ då $X \in Po(13.45)) < 0.01$

Genom att pröva olika värden på k fås

k	P-värde
22	0.0197
23	0.0110
24	0.0059 ← H_0 förk. på nivå 1%

Man kan alltså acceptera 23 fall. Överstiger antalet fall 23 kommer H_0 att förkastas på nivå 1%.

6 a)



Det systematiska felet är $91 - 100 = -9$.

b) $I_\mu = \left(\bar{x} \pm \lambda_{\alpha/2} \cdot \frac{2}{\sqrt{10}}\right) = (3{,}37,\ 6.63)$

Detta ger $\lambda_{\alpha/2} \cdot \frac{2}{\sqrt{10}} = \frac{6.63 - 3{,}37}{2} = 1{,}63$

och $\lambda_{\alpha/2} = \frac{1{,}63 \cdot \sqrt{10}}{2} = 2.5773$ vilket innebär att $\alpha = 0.01$ och konfidensgraden 99%.

c) X_i = mätfel nr i $\quad E(X_i) = 0 \quad D(X_i) = \frac{1}{\sqrt{12}}$

Enligt centrala gränsvärdessatsen gäller

$$\sum_1^{50} X_i \overset{\approx}{\in} N\left(50 \cdot 0,\ \sqrt{50} \cdot \frac{1}{\sqrt{12}}\right) = N\left(0, \sqrt{\frac{50}{12}}\right)$$

$$P\left(\sum_1^{50} X_i > 0{,}1\right) \approx 1 - \Phi\left(\frac{0{,}1}{\sqrt{\frac{50}{12}}}\right) = 0{,}4805$$

d) $P(\text{förkasta } H_0 \text{ då } \mu = 3{,}5) = 0{,}2 = S(3{,}5)$ Genom att pröva i kommandot styrkefkn för $n = 10$

e) X = tid mellan störningar, $X \in Exp(1/0{,}25)$

$$P(X > 0{,}5) = \int_{0{,}5}^{\infty} 0{,}25\, e^{-0{,}25x}\, dx = 0{,}1353$$