EXAMEN STATISTIEK EN WISKUNDIGE DATA-ANALYSE (1^{ste} zit '19-'20, reeks B)

Opleiding industrieel ingenieur

T. Van Hecke

FACULTEIT INGENIEURSWETENSCHAPPEN EN ARCHITECTUUR

Nr.:

Naam : **Richting:** /20

Schrijf netjes. Vul in op de opengelaten plaatsen.
Eenvoudig rekenmachine toegelaten. Geen gsm, smartphone, ...
Geef uitleg bij de open vragen. Veel succes!

UNIVERSITEIT GENT

1. In een bepaalde stad is het aantal ongevallen per dag Poisson-verdeeld. De kans dat er op een dag 4 ongevallen gebeuren is even groot als de kans dat er op een dag 3 ongevallen gebeuren. In februari 2020 (schrikkeljaar) werd elke dag het aantal ongevallen geteld. Benader de kans dat er in deze maand minstens 13 dagen waren met meer dan 4 ongevallen. /4

x = # ongevallen per dag; x is Poisson verdeeld met gemiddelde λ;

$\lambda = ?$ Maar $P(x=4)=P(x=3) \Leftrightarrow e^{-\lambda}\dfrac{\lambda^4}{4!} = e^{-\lambda}\dfrac{\lambda^3}{3!} \Leftrightarrow \lambda = 4$

$\Rightarrow f(i) = P(x=i) = e^{-4}\dfrac{4^i}{i!}$

i = # dagen met meer dan 4 ongevallen; i : Bin$(29, p)$ met $p = P(x>4)$

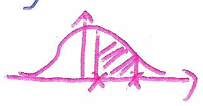
$P(x>4) = 1 - P(x \leq 4) = 1 - e^{-4}\sum_{i=0}^{4}\dfrac{4^i}{i!} = 1 - e^{-4}\left(1 + 4 + \dfrac{4^2}{2} + \dfrac{4^3}{6} + \dfrac{4^4}{24}\right) = 0.3712$

i is Bin$(29, 0.3712)$ met $np = \overbrace{29(0.3712)}^{10.76} > 5$ en $n(1-p) = \overbrace{29(1-0.3712)}^{18.24} > 5$

is is goed benaderd door y : $N(np, \sqrt{np(1-p)}) = N(10.76, 2.6)$

$P(13 \leq i \leq 29) \approx P(12.5 \leq y \leq 29.5) = P\left(\dfrac{12.5-10.76}{2.6} \leq \underbrace{\dfrac{y-10.76}{2.6}}_{z \text{ is } N(0,1)} \leq \dfrac{29.5-10.76}{2.6}\right)$

$= P(0.67 \leq z \leq 7.21) = 0.5 - 0.2486 = 0.2514$



OF

$P(i \geq 13) \approx P(y \geq 12.5) = P\left(\dfrac{y-10.76}{2.6} \geq \dfrac{12.5-10.76}{2.6}\right) = P(z \geq 0.67) = 0.5 - 0.2486$

$= 0.2514$

2. Bepaal de variantie van de continue variabele x met $\begin{cases} \frac{1}{8}x^3, & \text{als } x \in [0, 2] \\ 0, & \text{als } x < 0 \\ 1, & \text{als } x > 2 \end{cases}$ als cumulatieve distributiefunctie. /3

$$f(x) = \frac{dF}{dx} = \begin{cases} \frac{3}{8}x^2 & \text{als } x \in [0,1] \\ 0 & \text{elders} \end{cases}$$

$$\sigma^2 = E[x^2] - \mu^2 = \int_0^2 x^2 \cdot \frac{3}{8}x^2 dx - \left[\int_0^2 x \frac{3}{8}x^2 dx\right]^2 = \int_0^2 \frac{3}{8}x^4 dx - \left[\int_0^2 \frac{3}{8}x^3 dx\right]^2$$

$$= \frac{3}{8}\left[\frac{x^5}{5}\right]_0^2 - \left(\frac{3}{8}\left[\frac{x^4}{4}\right]_0^2\right)^2 = \frac{3}{8} \cdot \frac{32}{5} - \left(\frac{3}{8} \cdot \frac{6}{4}\right)^2 = \frac{12}{5} - \frac{9}{4} = \frac{3}{20}$$

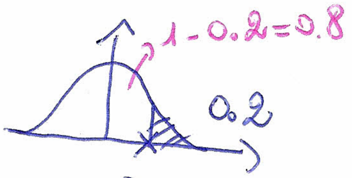
3. Gegeven de onafhankelijke veranderlijken $x : N(2, \sqrt{2})$, $y : N(4, \sqrt{7})$ en $z : \chi^2(9\,\text{d.f.})$. Bepaal α zodat $P(3\,x - y - 2 > \alpha\sqrt{z}) = 0.2$. /3

$$3x - y - 2 \text{ is } N\left(3 \cdot 2 - 4 - 2, \sqrt{(3)^2 \cdot 2 + (\sqrt{7})^2}\right) = N(0, 5)$$

$$\Rightarrow \frac{3x-y-2}{5} \text{ is } N(0,1) \text{ en } \frac{\frac{3x-y-2}{5}}{\sqrt{\frac{z}{9}}} = \frac{3}{5}\,\frac{3x-y-2}{\sqrt{z}} \text{ is } t(9df)$$

$$P(3x - y - 2 > \alpha\sqrt{z}) = 0.2 \Leftrightarrow P\left(\frac{3x-y-2}{\sqrt{z}}\right) = 0.2$$

$$\Leftrightarrow P\left(\underbrace{\frac{3}{5}\,\frac{3x-y-2}{\sqrt{z}}}_{t(9df)} > \alpha\frac{3}{5}\right) = 0.2$$



$$\frac{3}{5}\alpha = t_{0.8}(9df)$$

$$\Downarrow$$

$$\frac{3}{5}\alpha = 0.883 \Leftrightarrow \alpha = 1.472$$

4. Vul in onderstaande tabel met een hoofdletter de juiste oplossing aan. Er is telkens exact 1 juiste oplossing. Duid bij elke vraag een antwoord aan want standard setting wordt toegepast bij de evaluatie. /10

vraag	1	2	3	4	5	6	7	8
antwoord								

(1) Men beschikt over een kaartspel van 52 kaarten: 13 ♠, 13 ♡, 13 ♢ en 13 ♣. Men trek tegelijk 2 kaarten. Wat is de kans dat ze beiden van dezelfde soort zijn?

A. 0.25
B. 0.2353
C. 0.0625
D. 0.3529

$P(A) = \frac{C_4^1 \, C_{13}^2}{C_{52}^2} = \frac{4 \cdot \frac{13(12)}{2}}{\frac{52 \cdot 51}{2}} = \frac{12}{51} = 0.2353$

(2) Bij het vergelijken van gemiddeldes, waarbij data onderverdeeld is volgens 2 criteria (met 3 rijen en 4 kolommen), zullen de vrijheidsgraden van de interactie gelijk zijn aan

A. 9
B. 12
C. 8
D. 6

$(3-1) \times (4-1) = 6$

(3) Veronderstel dat de verdeling van de lonen in een bedrijf X een mediaan heeft van €35 000. Het eerste en derde kwartiel zijn respectievelijk €21 000 en €53 000. Zijn lonen van €100 000 en €1 dan uitschieters?

A. zowel €100 000 als €1
B. enkel €100 000
C. enkel €1
D. noch €100 000 noch €1

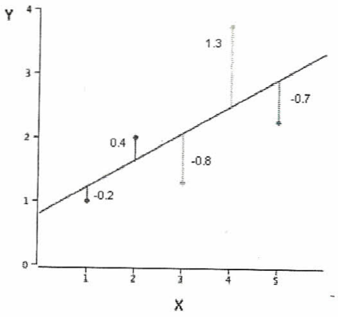
IQR = 3de kw - 1ste kw = 53000 - 21000 = 32 000
Uitschieters: (als < -27000 of > 101 000)
$21000 - \frac{3}{2}(32\,000) = -27000$
$53000 + \frac{3}{2}(32\,000) = 101\,000$

(4) De standaarddeviatie van een groep metingen is 10. Als 5 wordt afgetrokken van elke meetwaarde, wat is dan de variantie van de nieuwe metingen

A. 5
B. 10
C. 100
D. $\sqrt{10}$

$V(x) \underset{\text{eig}}{=} V(x-5) = 10^2 = 100$

(5) Onderstaande tekening toont de regressielijn die y voorspelt in functie van x. De waarden op de tekening stellen de residues voor. Wat is de waarde van SSE?

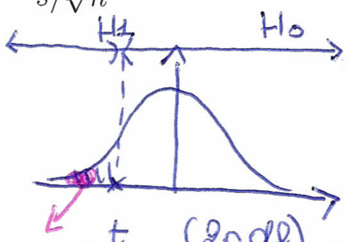


A. 1.01
B. 3.02
C. 0
D. 0.75

$$SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$
$$= (-0.2)^2 + (0.4)^2 + (-0.8)^2 + (1.3)^2 + (-0.7)^2$$
$$= 3.02$$

(6) Bij het testen van $H_0 : \mu = 10$ (met $H_1 : \mu < 10$) op basis van een steekproef met grootte $n = 21$ en $t_{ber} = \dfrac{\overline{x} - \mu}{s/\sqrt{n}}$, verwerpen we de nulhypothese met 90% betrouwbaarheid als

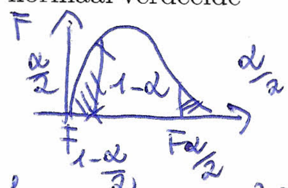
eën-zijdige test (links)



A. $t_{ber} > 1.32$
B. $t_{ber} > 1.72$
C. $t_{ber} < -1.72$
D. $t_{ber} < -1.32$

t_{ber} $\quad -t_{0.90}(20\,df) = -1.32$

(7) We beschikken over 2 onafhankelijke steekproeven komende uit normaal verdeelde populaties:



steekproef 1:	$n_1 = 8$	$\overline{x_1} = 10$	$s_1 = 4$
steekproef 2:	$n_2 = 9$	$\overline{x_2} = 7$	$s_2 = 5$

Het 95% betrouwbaarheidsinterval voor $\dfrac{\sigma_2^2}{\sigma_1^2}$ is

$$\frac{\sigma_2^2}{\sigma_1^2} \in \left[\frac{1}{F_{0.025}(8,7\,df)} \cdot \frac{s_2^2}{s_1^2}, \frac{s_2^2}{s_1^2} F_{0.025}(7,8\,df)\right]$$

A. [0.3449, 7.6563]
B. [0.3584, 6.4063]
C. [0.3189, 7.0781]
D. [0.3811, 6.8125]

$$= \left[\frac{1}{4.9} \cdot \frac{25}{16}, \frac{25}{16}(4.53)\right]$$
$$= [0.3189, 7.0781]$$

(8) Beschouw een lengte $x : N(1, 0.2)$ en een lengte $y : N(3, 0.1)$. Welke verdeling heeft dan $x - 2\,y$?

A. $N(-5, \sqrt{0.02})$
B. $N(-5, \sqrt{0.08})$
C. $N(-5, 0)$
D. $N(-5, \sqrt{0.06})$

$$x - 2y : N\left(1 - 2(3), \sqrt{(0.2)^2 + 2^2(0.1)^2}\right)$$
$$= N(1-6, \sqrt{0.04 + 4(0.01)})$$
$$= N(-5, \sqrt{0.08})$$