Modeloplossing examen Statistiek voor bedrijfskundigen-I,

versie groen/roze

1. [10 ptn] **Juist of fout?** Vul in met "juist" of "fout".

De score op deze vraag wordt toegekend aan de hand van de volgende tabel:

Aantal correct beantwoorde items	0	1	2	3	4	5	6	7	8	9	10
Score	0	1	1	2	3	3	4	5	6	8	10

1) Een stochastische variabele met een ordinale meetschaal is discreet. 1) Fout

2) Op een intervalschaal duidt de waarde 'nul' op afwezigheid van het bestudeerde kenmerk. 2) Fout

3) Bij een scheve gegevensverdeling met een lange staart aan de rechterzijde heeft de mediaan een negatieve z-score. 3) Juist

4) De standaardafwijking is ongevoelig aan extreme gegevenswaarden. 4) Fout

5) Voor twee gebeurtenissen A en B geldt $P(A|B)P(B) = P(B|A)P(A)$. 5) Juist

6) Disjuncte gebeurtenissen zijn niet onafhankelijk. 6) Juist

7) De kans dat een discrete stochast X een waarde a aanneemt is nul, en dit voor alle waarden a in het bereik van X. 7) Fout

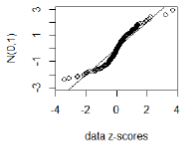
8) Het steekproefgemiddelde is een zuivere schatter van het populatiegemiddelde, omdat er steekproeven bestaan waarbij het steekproefgemiddelde gelijk is aan het populatiegemiddelde. 8) Fout

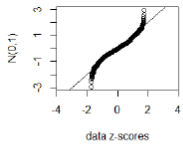
9) De absolute waarde van het verschil tussen $\bar{x}$ en populatiegemiddelde wordt de standaardfout van $\bar{x}$ genoemd. 9) Fout

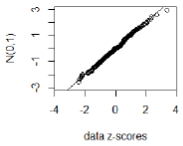
10) Voor kleine steekproeven ($n < 30$) is $\bar{x}$ normaal verdeeld indien x een normale verdeling heeft. 10) Juist

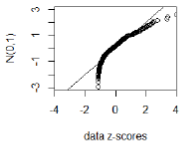
2. (a) [4 ptn] **Geef telkens aan welke normaliteitsplot bij de onderstaande uitspraken hoort.**
Bij elke uitspraak past slechts één grafiek. Een grafiek kan bij meerdere uitspraken passen.
De score op deze vraag wordt toegekend aan de hand van de volgende tabel:

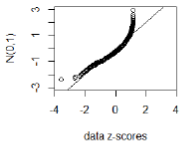
Aantal correct beantwoorde items	0	1	2	3	4
Score	0	0	2	3	4

A. 

B. 

C. 

D. 

E. 

1) De gegevensverzameling is vrijwel normaal verdeeld. 1) C
2) De gegevensverdeling is scheef met een lange staart naar links. 2) E
3) De gegevens variëren van min 4 tot plus 4 standaardafwijkingen van het gemiddelde. 3) A
4) Minder dan 1% van de gegevens is verder dan twee standaardafwijkingen van het gemiddelde verwijderd. 4) B

(b) [3 ptn] Tien studenten hebben meegedaan aan een IQ-test. Van één student is de score x verloren gegaan. Eén van de toetsanalisten herinnerde zich echter nog dat $x \geq 115$.
De resultaten in dalende volgorde zijn: x , 114 , 108 , 103 , 101 , 100 , 99 , 99 , 94 , 92.
Bepaal de kleinste gehele waarde van x zodat dit een *mogelijke uitschieter* zou zijn.

mediaan = 100,5

$Q_1 = 99$
$Q_3 = 108$
$\Rightarrow KA = 9$

$Q_3 + 1{,}5\ KA : 108 + 13{,}5 = 121{,}5$
$\Rightarrow$ kleinste gehele waarde zodat mogelijke uitschieter $\Rightarrow$ 122

3. In een bepaald academiejaar slaagde 30% van de eerstejaarsstudenten HW voor het vak *Wiskunde I-B* in eerste zittijd. Van deze studenten slaagde 80% eveneens voor het vak *Statistiek I* in eerste zittijd. Verder bleek dat 28% van de eerstejaarsstudenten HW beide vakken moest hernemen in tweede zittijd.

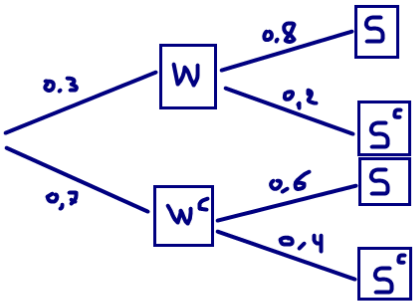
Los de volgende vragen op met correct gebruik van kansnotaties en gebruik volgende afkortingen voor de gebeurtenissen:
W = "Geslaagd voor *Wiskunde I-B* in eerste zittijd" en S = "Geslaagd voor *Statistiek I* in eerste zittijd".

(a) [3 ptn] Stel dat 500 studenten HW examens van de vakken *Wiskunde I-B* en *Statistiek I* aflegden in eerste zittijd. Hoeveel studenten slaagden voor beide vakken?

(b) [3 ptn] Hoe groot was de kans dat iemand die niét slaagde voor *Statistiek I*, wél geslaagd was voor *Wiskunde I-B*?

(c) [3 ptn] Zijn de gebeurtenissen 'Geslaagd voor *Wiskunde I-B* in eerste zittijd' en 'Geslaagd voor *Statistiek I* in eerste zittijd' van elkaar (on-)afhankelijk? Motiveer aan de hand van kansregels.

gegeven:



$P(W) = 0,3$

$P(S \mid W) = 0,8$

$P(W^c \cap S^c) = 0,28$

a) $P(W \cap S) = P(W) \cdot P(S \mid W) = \frac{3}{10} \cdot \frac{8}{10} = \frac{24}{100} = 24\%$

$24\% \cdot 500 =$ **120 studenten**

b) $P(W \mid S^c) = \frac{P(W \cap S^c)}{P(S^c)} = \frac{\frac{3}{10} \cdot \frac{2}{10}}{\frac{3}{10} \cdot \frac{2}{10} + \frac{28}{100}} = \frac{\frac{6}{100}}{\frac{34}{100}} = \frac{3}{17}$

c) de gebeurtenissen zijn **niet onafhankelijk**,

bv. $P(W) \neq P(W \mid S^c)$ ofwel: $P(S^c \mid W) \neq P(S^c)$

$\frac{3}{10} \neq \frac{3}{17}$ ofwel: $\frac{2}{10} \neq \frac{34}{100}$

ofwel: $P(W \cap S) \neq P(W) \cdot P(S)$ ofwel: $P(S) \neq P(S \mid W)$

$\frac{24}{100} \neq \frac{3}{10} \cdot \frac{66}{100}$ ofwel: $\frac{66}{100} \neq \frac{8}{10}$

4. [4 ptn] **Zijn de volgende stochastische variabelen binomiaal?** Duid telkens aan met "ja" of "nee".

De score op deze vraag wordt toegekend aan de hand van de volgende tabel:

Aantal correct beantwoorde items	0	1	2	3	4
Score	0	0	0	2	4

1) Het aantal mislukkingen in een binomiaal experiment. 1) ja

2) Het aantal doelpunten in een serie van 5 strafschoppen, nadat een match van de Rode Duivels eindigt op een gelijkspel, ook na verlengingen. 2) nee

3) Het aantal keer dat je met Euromillions moet spelen alvorens je wint. 3) nee

4) Het aantal personen zonder abonnement op Netflix bij 100 op straat aangesproken volwassenen. 4) ja

Toelichting:

2) niet elke strafschop heeft dezelfde kans op een doelpunt, wegens andere speler

3) het aantal pogingen is vooraf niet gekend

5. Het invoeren van flexi-jobs en de witte kassa hebben het zwartwerk in de horeca gevoelig teruggedrongen, zo blijkt uit een onderzoeksrapport (eind 2017). Toch blijft deze sociale en fiscale fraude een groot probleem, zowel voor de overheid (die belastingsinkomsten misloopt) als voor de betrokken werknemer (die geen sociale rechten geniet). Een restaurantuitbater denkt dat zowat 30% van de horecazaken nog steeds fraude pleegt. We beschouwen een aselecte steekproef van n horecazaken in Vlaanderen.

(a) [3 ptn] Hoe groot is de kans dat er in minstens één derde van $n = 20$ horecazaken zwartwerk wordt vastgesteld?

(b) [3 ptn] Je inspecteert lukraak $n = 20$ horecazaken en stelt zwartwerk vast in maximaal 3 ervan. Zijn er statistische redenen om te twijfelen aan de 30%-hypothese van de restaurantuitbater? Leg uit.

(c) [4 ptn] Onderstel $n = 100$. De kans dat in minstens de helft van deze 100 horecazaken fraude geconstateerd wordt, is het best te benaderen door $P(z > a)$ waarbij a gelijk is aan:
A. $\frac{19,5}{\sqrt{0,21}}$ B. $\frac{19,5}{\sqrt{21}}$ C. $\frac{20,5}{\sqrt{21}}$ D. $\frac{20,5}{\sqrt{0,21}}$ E. Kans is niet benaderbaar via de normale verdeling.
Verklaar.

a) $P(x \geq \frac{20}{3}) = P(x \geq 7) = 1 - P(x \leq 6) = 1 - 0,608 = \boxed{0,392}$

b) $P(x \leq 3) = 0,107 = \boxed{10,7\ \%}$

$\Rightarrow$ geen reden tot twijfel, aangezien $10,7\% > 5\%$

c) $\mu = 100 \cdot 0,3 = \boxed{30}$ $\sigma = \sqrt{100 \cdot 0,7 \cdot 0,3} = \boxed{\sqrt{21}} \approx 4,5$

$\mu \pm 3\sigma < \begin{matrix} 43,5 \\ 16,5 \end{matrix} \overset{!}{\in} [0,100]$

continuïteitscorrectie

$P(x \geq \frac{100}{2}) \approx P\left(z > \frac{50 - 0,5 - 30}{\sqrt{21}}\right) = P\left(z > \frac{19,5}{\sqrt{21}}\right) \Rightarrow$ antwoord B

6. Een fabrikant van wasmachines garandeert zijn klanten gratis vervanging in geval van defect binnen de twee jaar en vervanging tegen halve prijs bij defect binnen de 4 jaar (maar niet binnen de eerste 2 jaar). Een wasmachine wordt desgevallend slechts één keer vervangen. De levensduur van een wasmachine is normaal verdeeld met een gemiddelde van 12 jaar en een standaardafwijking van 4 jaar.

(a) [2 ptn] Welk percentage van de wasmachines zal kosteloos moeten vervangen worden?

(b) [3 ptn] Hoeveel wasmachines zullen aan halve prijs moeten vervangen worden op een totale levering van 10.000 stuks?

(c) [3 ptn] Beschrijf de vorm en de parameters van de kansverdeling van de gemiddelde levensduur van een wasmachine in een steekproef met $n = 16$.

(d) [2 ptn] Welk percentage van de wasmachines heeft een levensduur van exact 12 jaar?

stel x = de levensduur van een wasmachine in jaren

$x \sim N(12,4)$

a) $P(x<2) \approx P\left(z < \frac{2-12}{4}\right) = P(z<-2,5) = 0,5 - P(0<z<2,5)$

$= 0,5 - 0,4938 = 0.0062 = \boxed{0,62\%}$

b) $P(2<x<4) \approx P\left(\frac{2-12}{4} < z < \frac{4-12}{4}\right) = P(-2,5<z<-2)$

$= P(0<z<2,5) - P(0<z<2) = 0,4938 - 0,4772 = 0,0166$

$\Rightarrow 0,0166 \cdot 10000 = \boxed{166}$

c) $\bar{x}$ is normaal verdeeld met $\begin{cases} \mu_{\bar{x}} = 12 \text{ jaar} \\ \sigma_{\bar{x}} = \frac{4}{\sqrt{16}} = 1 \text{ jaar} \end{cases}$

de vorm is klokvormig

d) $P(x=12) = \boxed{0}$

Modeloplossing examen Statistiek voor bedrijfskundigen-I

1. [10 ptn] **Juist of fout?** Vul in met "juist" of "fout".

De score op deze vraag wordt toegekend aan de hand van de volgende tabel:

Aantal correct beantwoorde items	0	1	2	3	4	5	6	7	8	9	10
Score	0	1	1	2	3	3	4	5	6	8	10

1) Gegevens op een ratioschaal bezitten de eigenschap dat gelijke verschillen in waarden wijzen op gelijke verschillen in het bestudeerde kenmerk. 1) Juist

2) Een continue stochastische variabele kan ook een eindig aantal mogelijke uitkomsten hebben. 2) Fout

3) De modus is een maatstaf die niet bruikbaar is bij kwantitatieve gegevens. 3) Fout

4) In elke gegevensverzameling geldt: $\sum_{i=1}^{n}(x_i - \bar{x}) = 0$. 4) Juist

5) De gemiddelde score op een examen is 73 (op 100) , met een standaardafwijking gelijk aan 2. Dan zijn ten minste 75% van de scores begrepen tussen 68 en 79. 5) Juist

6) Gegeven: twee onafhankelijke gebeurtenissen A en B met $P(A) = 0{,}3$ en $P(B) = 0{,}1$. Dan is $P(A \cup B) = 0{,}4$. 6) Fout

7) Bij een continue kansverdeling geldt: $P(x \geq a) = P(x > a)$ voor alle waarden a. 7) Juist

8) Het kanshistogram van een binomiaal verdeelde stochastische variabele is met $p = 1/2$ is symmetrisch. 8) Juist

9) Steekproeven met grootte n worden genomen uit een populatie met populatieverwachting μ. De verwachting van het steekproefgemiddelde is dan gelijk aan $\mu/\sqrt{n}$. 9) Fout

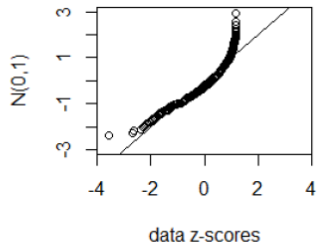
10) De centrale limietstelling zegt dat de steekproefverdeling van $\bar{x}$ exact normaal verdeeld is als $n \geq 30$. 10) Fout

2. (a) [4 ptn] **Juist of fout?** Vul in met "juist" of "fout", op je antwoordbundel.

De score op deze vraag wordt toegekend aan de hand van de volgende tabel:

Aantal correct beantwoorde items	0	1	2	3	4
Score	0	1	1	2	4

Een gegevensverzameling wordt gekenmerkt door de volgende normaliteitsplot:



1) De mediaan en het rekenkundig gemiddelde liggen dicht bij elkaar. 1) Juist
2) Minder dan 1% van de gegevens is verder dan twee standaardafwijkingen van het gemiddelde verwijderd. 2) Fout
3) De gegevens variëren van min 3 tot plus 3 standaardafwijkingen van het rekenkundig gemiddelde. 3) Fout
4) De gegevensverdeling is scheef met een lange staart naar links. 4) Juist

(b) [3 ptn] **Kies het juiste alternatief.** Slechts 1 alternatief is correct. Je dient geen uitwerking te tonen.

Tien studenten hebben meegedaan aan een IQ-test. De resultaten zijn:

100 , 102 , 92 , 99 , 103 , 101 , 114 , 99 , 94 , 116.

A. Enkel 116 is een uitschieter.
B. 92 is een mogelijke uitschieter, 114 en 116 zijn vrijwel zekere uitschieters.
C. Enkel 114 en 116 zijn uitschieters.
D. 92 en 114 zijn mogelijke uitschieters, 116 is een vrijwel zekere uitschieter.
E. Er zijn geen uitschieters.

Toelichting:

gerangschikt: 92, 94, 99, 99, 100, 101, 102, 103, 114, 116

$Q_1 = 99$
$Q_3 = 103$ $\Rightarrow KA = 4$

$Q_1 - 1{,}5\,KA$: $99 - 6 = 93$
$Q_1 - 3\,KA$: $99 - 12 = 87$
$Q_3 + 1{,}5\,KA$: $103 + 6 = 109$
$Q_3 + 3\,KA$: $103 + 12 = 115$

3. [4 ptn] **Zijn de volgende stochastische variabelen binomiaal?** Duid telkens aan met "ja" of "nee".

De score op deze vraag wordt toegekend aan de hand van de volgende tabel:

Aantal correct beantwoorde items	0	1	2	3	4
Score	0	1	1	2	4

1) Het aantal politici met een Twitter-account uit 100 willekeurig gekozen politici. 1) ja
2) Het aantal eigenaars van zonnepanelen in een steekproef van 1000 Vlamingen. 2) ja
3) Het aantal pogingen nodig om te slagen voor het theoretisch rij-examen. 3) nee
4) Het aantal vakken uit BA1 handelswetenschappen waarvoor een modeltrajectstudent tweede zit heeft. 4) nee

Toelichting:

3) per vak is de slaagkans niet even groot
4) het aantal pogingen is vooraf niet gekend

4. [10 ptn] Een bepaald bedrijf heeft in het voorbije jaar 50 sollicitatiebrieven ontvangen, waarvan 10 geschreven door werkzoekenden met een masterdiploma. Van deze masters werden er 3 uitgenodigd tot een sollicitatiegesprek. Verder werden 70% van de ontvangen sollicitatiebrieven door het bedrijf verticaal (in de prullenmand) geklasseerd.

Los de volgende vragen op aan de hand van kansnotaties en kansregels. Gebruik hierbij de volgende afkortingen voor de gebeurtenissen:
M = "Sollicitant met masterdiploma" en U = "Sollicitant kreeg uitnodiging voor een gesprek".

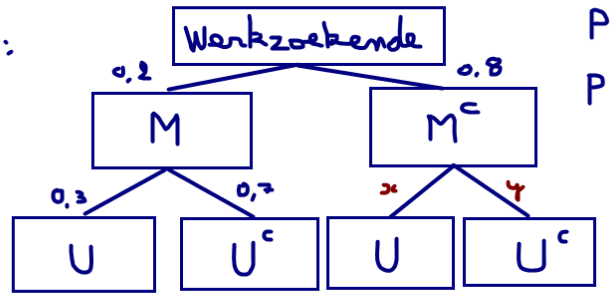
(a) [1 ptn] Bepaal de kans dat een aselect gekozen sollicitatiebrief werd geschreven door een werkzoekende zonder masterdiploma.

(b) [3 ptn] Bepaal de kans dat een lukraak gekozen sollicitatiebrief afkomstig is van een werkzoekende zonder masterdiploma en dat deze sollicitant niet werd uitgenodigd voor een gesprek.

(c) [3 ptn] Bepaal de kans dat een willekeurig in de prullenmand gegooide sollicitatiebrief door een master werd geschreven.

(d) [3 ptn] *Onderzoek* of het uitnodigen tot een sollicitatiegesprek afhankelijk is van het diploma van de schrijver van de sollicitatiebrief.

kansboom:



$P(U^c) = 0,7$

$P(U^c) = P(U^c \cap M) + P(U^c \cap M^c)$

$0,7 = 0,2 \cdot 0,7 + 0,8 \cdot y$

$\Rightarrow y = 0,7 \Rightarrow x = 0,3$

OF kruistabel:

50	M	M^c	Σ
U	3	12	15
U^c	7	28	35
Σ	10	40	50

* = gegeven
* = berekend

a) $P(M^c) = 1 - P(M) = 1 - 0,2 = \boxed{0,8}$

b) $P(M^c \cap U^c) = P(U^c \mid M^c) \cdot P(M^c) = 0,7 \cdot 0,8 = \boxed{0,56}$

c) $P(M \mid U^c) = \dfrac{P(U^c \mid M) \cdot P(M)}{P(U^c)} = \dfrac{0,7 \cdot 0,2}{0,7} = \boxed{0,2}$

d) $P(U \mid M) \overset{?}{=} P(U)$

$P(U \mid M) = 0,3$

$P(U) = P(U \cap M) + P(U \cap M^c) = \dfrac{3}{50} + \dfrac{12}{50} = \dfrac{15}{50} \overset{!}{=} 0,3$

$\Rightarrow$ U en M zijn onafhankelijk

5. [9 ptn] Voor het Vlaams Parlement stemden 30% van de kiezers in Oost-Vlaanderen in 2014 voor de N-VA. We beschouwen een aselecte steekproef van $n = 20$ kiesgerechtigden uit Oost-Vlaanderen.

(a) [3 ptn] Bepaal de kans dat er in deze steekproef minder dan één derde voor de N-VA heeft gestemd.

(b) [3 ptn] Bepaal de kans dat minstens twee en hoogstens tien personen uit deze steekproef voor de N-VA hebben gestemd.

(c) [3 ptn] Is je antwoord in vraag (b) consistent met de regel van Tsjebysjev? Leg uit.

Stel x = aantal NVA-kiezers $\quad x \sim B(20;0,3)$

a) $P(x \leq 6) = \boxed{0,608}$

b) $P(2 \leq x \leq 10) = P(x \leq 10) - P(x \leq 1) = 0,983 - 0,008 = \boxed{0,975}$

c) $\mu = 20 \cdot 0,3 = 6 \qquad \sigma = \sqrt{20 \cdot 0,7 \cdot 0,3} = \sqrt{4,2} \approx 2,1$

Tsjebysjev: $P(\mu - 2\sigma < x < \mu + 2\sigma) \geq 75\%$

$P(1,8 < x < 10,2) = P(x \leq 10) - P(x \leq 1) = 0,983 - 0,008 = \boxed{0,975}$

$97,5\% > 75\% \Rightarrow$ consistentie

6. [10 ptn] De maandelijkse huurprijs van een appartement in een bepaalde deelgemeente van Gent anno 2017 is normaal verdeeld met een gemiddelde van €750 en een standaardafwijking van €200.

(a) [3 ptn] Je hebt net je masterdiploma behaald en werk gevonden rond het Gentse en wenst daar een appartement te huren. Wat is de kans dat je een appartement zal vinden van max. €600 per maand?

(b) [4 ptn] Bepaal de minimumprijs van de 10% duurste appartementen in deze deelgemeente van Gent.

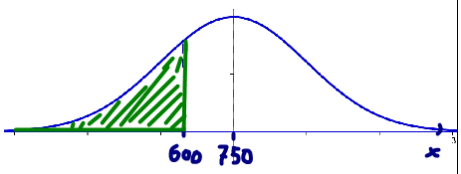
(c) [3 ptn] In een steekproef van $n = 25$ willekeurig geselecteerde appartementen vinden we $\bar{x} = €850$ Is dit een zeldzame gebeurtenis? Leg uit.

a) stel x = maandelijkse huurprijs appartement

$x \sim N(750, 200)$

$P(x \leq 600) \approx P\left(z < \frac{600-750}{200}\right) = P(z < -0{,}75)$

$= 0{,}5 - P(0 < z < 0{,}25) = 0{,}5 - 0{,}2734 = \boxed{0{,}2266}$



b) stel x_0 = gevraagde minimumprijs

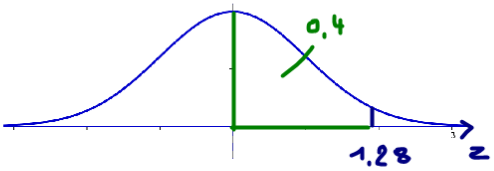
$P(x > x_0) \leq 0{,}1$

$\Leftrightarrow P\left(z > \frac{x_0-750}{200}\right) \leq 0{,}1$

$\Leftrightarrow \frac{x_0-750}{200} \geq 1{,}28$

$\Leftrightarrow x_0 \geq 750 + 1{,}28 \cdot 200$

$\Leftrightarrow \boxed{x_0 \geq 1006\ €}$



c) $\bar{x} \sim N(750, 40)$ want $\bar{x}$ is normaal verdeeld met $\begin{cases} \mu_{\bar{x}} = 750 \\ \sigma_{\bar{x}} = \frac{200}{\sqrt{25}} = 40 \end{cases}$

$P(\bar{x} \geq 850) \approx P\left(z \geq \frac{850-750}{40}\right) = P(z \geq 2{,}5) = 0{,}5 - 0{,}4938 = \boxed{0{,}62\%}$

$\Rightarrow$ zeer kleine kans $\Rightarrow$ **ja**