Tentamen Inleiding Kansrekening en Statistiek (WISB161) 2017-2018
1 februari 2018

I *Schrijf je naam en studentnummer op elk vel dat je inlevert.*
II *Elke opgave van het tentamen wordt door 1 docent nagekeken. Maak daarom opgaven 1 en 2, opgaven 3 en 4, opgave 5, en opgave 6 en 7 op een apart blad.*
III *Elektronische apparatuur is niet toegestaan. Het boek van Dekking et al. mag wel gebruikt worden.*
IV *U mag in ieder onderdeel de conclusies van voorgaande onderdelen gebruiken, ook als u die (nog) niet bewezen hebt. Als u een antwoord op een vorige deelvraag niet heeft kunnen vinden, mag u een antwoord naar keuze veronderstellen en daarmee verder rekenen. Geef duidelijk aan als u dit doet. Als de vraag door de aanname eenvoudiger wordt kan dit tot puntenaftrek leiden.*
V *Motiveer steeds uw antwoord door duidelijk aan te geven welke argumenten en welke resultaten u gebruikt om een bepaalde conclusie te trekken.*
VI *Uitdrukkingen hoeven niet numeriek bepaald te worden tenzij hier expliciet om wordt gevraagd. Vereenvoudig uw antwoord wel indien mogelijk.*
VII *U heeft 3 uur de tijd voor het tentamen.*
VII *Achter elke deelvraag staat het aantal punten dat met de deelvraag te behalen is. In totaal zijn er 90 punten te behalen en 5 extra punten met de bonusvraag 5c. De puntenverdeling per vraag is: 1 - 10, 2 - 10, 3 - 21, 4 - 15, 5 - 10(+5), 6 - 14, 7 - 10.*

Veel succes!

Opgave 1 Stel dat twee stochastische variabelen X en Y een gezamenlijke kansfunctie hebben gegeven door Tabel 1.

Tabel 1: Gezamenlijke kansen $P(X = x, Y = y)$

		$x=0$	$x=1$	$x=2$
y	1	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{16}$
	2	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{8}$
	3	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{16}$

a 4pt) Wat is de gezamenlijke uitkomstenruimte Ω van (X, Y)?

b 6pt) Zijn X en Y onafhankelijk? Licht je antwoord toe.

Opgave 2 Stel dat het IQ van studenten die het vak WISB161 volgen normaal verdeeld is, maar dat het gemiddelde en de variantie van de normale verdeling niet bekend is. Stel dat je wilt testen, met een significantieniveau gelijk aan $\alpha = 0.05$, of het IQ van deze populatie hoger is dan het landelijke gemiddelde van 100. Daartoe meet je het IQ van 5 willekeurig gekozen studenten, en de gevonden IQ-waarden zijn: $\{121, 97, 122, 119, 101\}$.

a 10pt) Bereken een geschikt betrouwbaarheidsinterval voor het gemiddelde IQ van WISB161-studenten en concludeer op basis van dit betrouwbaarheidsinterval of het gemiddelde IQ van deze populatie hoger is dan 100.

Opgave 3 Zij $-2 \leq \theta \leq 2$ een parameter en zij X_θ een continue stochastische variabele op de reële getallen met kansdichtheidsfunctie

$$f_\theta(x) = \begin{cases} \alpha(\theta) + \theta x & \text{voor } x \in (0,1) \\ 0 & \text{als } x \notin (0,1). \end{cases}$$

Stel dat we een willekeurig steekproef $X_1, X_2, \ldots, X_n$ hebben van grootte n hebben waarbij alle X_i onafhankelijk en identiek verdeeld zijn als X_θ.

a 4pt) Gebruik dat f_θ een kansdichtheid is om te bepalen hoe $\alpha(\theta)$ van θ afhangt.

b 2pt) Bepaal $P(X_\theta = \frac{1}{2})$.

c 3pt) Bepaal $E(X_\theta)$.

d 5pt) Bepaal een zuivere schatter van θ op basis van het steekproefgemiddelde.

e 5pt) Bepaal een formule waaraan de meest-waarschijnlijke schatter (maximum likelihood schatter) van θ moet voldoen als we aan mogen nemen dat de absolute waarde van de meest waarschijnlijke schatter kleiner is dan 2.

f 2pt) Stel $n = 2$, $x_1 = \frac{1}{3}$ en $x_2 = \frac{3}{4}$. Bepaal de meest waarschijnlijke schatting van θ.

Opgave 4 Stel dat 100 onderzoeksteams elk een dataset hebben verzameld en daarmee elk een 90%-betrouwbaarheidsinterval voor een parameter θ geconstrueerd hebben. Veronderstel dat de datasets onafhankelijk zijn.

a 5pt) Bepaal de kansfunctie van het aantal keren dat de daadwerkelijke parameterwaarde θ_0 niet in het 90% betrouwbaarheidsinterval valt.

b 10pt) Stel dat je twijfelt of elke onderzoeksgroep in staat is om een correct 90%-betrouwbaarheidsinterval te construeren. Je kiest als nulhypothese dat elke onderzoeksgroep correct een 90%-betrouwbaarheidsinterval kan construeren en de alternatieve hypothese is dat niet alle groepen dit kunnen. Om deze test uit te voeren heb je zelf θ_0 met grote nauwkeurigheid bepaalt en op basis hiervan geconcludeerd dat θ_0 bevat was in 80 van de 100 90% betrouwbaarheidsintervallen.

Voer een geschikte test uit met een significantieniveau van $\alpha = 0.0456$. Je mag ervan uitgaan dat de steekproefgrootte groot genoeg is om asymptotsiche benaderingen te mogen gebruiken.

Opgave 5 Zij $X_1, X_2, \ldots, X_n$ onafhankelijke, identiek verdeelde Bernoulli stochasten met parameter p. Definieer een run als een maximale hoeveelheid opeenvolgende enen. In het onderstaande realisatie met $n = 12$ zijn er 3 runs.

$$0 \quad 0 \quad 0 \quad \underbrace{1 \quad 1}_{\text{run 1}} \quad 0 \quad \underbrace{1}_{\text{run 2}} \quad 0 \quad 0 \quad \underbrace{1 \quad 1 \quad 1}_{\text{run 3}}$$

Zij $I_i = \begin{cases} 1 & \text{als een run begint op positie } i \\ 0 & \text{anders} \end{cases}$ voor $1 \leq i \leq n$ en zij $R = \sum_{i=1}^{n} I_i$ het aantal runs.

a 5pt) Bepaal de marginale kansverdeling van $I_1, I_2, \ldots, I_n$.

b 5pt) Bepaal de verwachtingswaarde van R.

c 5pt) **Bonusopgave** Bepaal de variantie van R.

Opgave 6 Zij X_1, X_2, …, X_n onafhankelijke identiek verdeelde stochasten met kansdichtheidsfunctie

$$f_\theta(x) = \begin{cases} \frac{1}{1+\theta} e^{-\frac{x}{1+\theta}} & \text{als } x > 0 \\ 0 & \text{als } x \leq 0 \end{cases}$$

met $1 \leq \theta \leq 2$. Zij T_1 een schatter van θ gedefinieerd door $T_1(X_1, X_2, \ldots, X_n) = \bar{X}_n - 1$ met $\bar{X}_n := \frac{1}{n} \sum_{i=1}^{n} X_i$.

a 6pt) Bereken de gemiddelde kwadratische fout van T_1.

b 6pt) De schatter T_1 kan schattingen produceren die buiten het interval $[1, 2]$ liggen. We definiëren daarom een tweede schatter T_2 door

$$T_2(X_1, X_2, \ldots, X_n) = \min\{2, \max\{1, T_1(X_1, X_2, \ldots, X_n)\}\}$$

met andere woorden, als de schatting $t_1(x_1, x_2, \ldots, x_n)$ groter is dan 2, dan nemen we 2 als schatting, als de schatting kleiner is dan 1, dan nemen we 1 als schatting. Laat zien of de gemiddelde kwadratische fout van T_2 groter, kleiner, of gelijk is aan de gemiddelde kwadratische fout van T_1. Merk op, ook als je vraag **a** niet hebt kunnen beantwoorden kun je vraag **b** beantwoorden.

c 2pt) Beargumenteer welke schatter de voorkeur heeft?

Opgave 7 Zij X uniform verdeeld op $(0, 1)$ en Y een Pareto verdeling met parameter $\alpha = 1$, i.e., de kansdichtheidsfunctie van Y wordt gegeven door:

$$f_Y(y) = \begin{cases} \frac{1}{y^2} & \text{als } y \geq 1 \\ 0 & \text{anders.} \end{cases}$$

X en Y zijn onafhankelijke stochasten. Definieer de stochast Z door $Z = X^Y$.

a 10pt) Bepaal de cumulatieve verdelingsfunctie $F_Z(z)$.

Einde.