Hertentamen Inleiding Kansrekening en Statistiek (WISB161) 2017-2018
19 april 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 10 extra punten met de bonusvraag 7c. De puntenverdeling per vraag is: 1 - 15, 2 - 15, 3 - 10, 4 - 10, 5 - 15 , 6 - 15, 7 - 10(+10).*

Veel succes!

Opgave 1 Stel dat er verkiezingen zijn geweest waarbij N individuen gestemd hebben. Er zijn 2 kandidaten, kandidaat A en kandidaat B. Kandidaat A heeft 53% van de stemmen behaald en kandidaat B 47%. Voor de verkiezingen is er een representatieve steekproef gehouden van grootte $n = 1000$ en veronderstel dat alle deelnemers aan de steekproef gestemd hebben dat ze correct aangeven op wie ze zullen stemmen. Definieer de stochastische variabelen p_A en p_B als de fractie van de deelnemers in de steekproef die, respectievelijk, op kandidaat A en kandidaat B zullen stemmen. Ga ervan uit dat $N \gg n$.

a 8pt) Wat is de variantie van p_A?

b 7pt) Bepaal $Cov(p_A, p_B)$.

Opgave 2 X en Y zijn discrete stochastische variabelen op de uitkomstenruimte $(\mathbb{Z}^+)^2$, met $\mathbb{Z}^+$ de verzameling van alle positieve gehele getallen, i.e., $\mathbb{Z}^+ = \{1, 2, 3, \ldots\}$. De gezamenlijke kansfunctie van X en Y is gegeven door:

$$P(X = i, Y = j) = \left(\frac{1}{2}\right)^i i \left(\frac{1}{i+1}\right)^j$$

U mag in deze opgave gebruiken dat $\sum_{i=1}^{\infty} \frac{i}{i+1} (\frac{1}{2})^i = 2 - \log 4$.

a 8pt) Bepaal de marginale kansfunctie voor X.

b 7pt) Zijn X en Y onafhankelijk?

Opgave 3 Een experimentator wil graag de concentratie C van stof A in een vloeistof bepalen. Daartoe neemt hij 4 monsters van de vloeistof en voegt aan elk monster een reagens toe die de vloeistof donkerder maken, hoe hoger de concentratie van stof A, des de donkerder de vloeistof wordt. Bij alle 4 de monsters bepaalt hij welk percentage van laserlicht door de verdunning komt. De experimentator weet dat de theoretische relatie tussen de concentratie C in het monster en het percentage doorgelaten licht I gegeven is door:

$$I = 100e^{-\frac{C}{2}}$$

De experimentator veronderstelt dat er bij elke meting van het percentage een standaard-normaal-verdeelde meetfout is t.o.v. van de theoretische waarde, en dat alle metingen onafhankelijk van elkaar zijn. De gemeten percentages doorgelaten licht zijn: $34\%, 35\%, 36\%, 31\%$.

a 10pt) Bepaal de meest waarschijnlijke schatter van de concentratie C.

Opgave 4 Zij X en Y twee onafhankelijke uniform verdeelde stochasten op $(0, 1)$. Definieer de stochast Z door $Z = -\frac{\log Y}{X}$.

a 10pt) Bepaal $P(Z > 1)$.

Opgave 5 Stel dat het jaarinkomen van wiskundigen normaal verdeeld is volgens $N(\mu, \sigma^2)$. Een willekeurige steekproef onder 25 wiskundigen geeft als steekproefgemiddelde van het jaarinkomen €50.000,- en S_{25}=€5.000,-.

a 9pt) Bepaal de waarde van het jaarinkomen x waarvoor geldt dat we 95% zekerheid weten dat $\mu > x$.

b 6pt) Bepaal een tweezijdig 90%-betrouwbaarheidsinterval voor μ.

Opgave 6 Zij Y het aantal auto's in een willekeurig gekozen huishouden. De kansfunctie van Y is gegeven in onderstaande tabel:

Y	0	1	2	3	4
P(Y)	0.25	0.44	0.2	0.08	0.03

Jan wil wedden dat als hij een willekeurige steekproef van grootte 225 uitvoert, dat het aantal auto's groter is dan 225. Hij is zo zeker van zijn zaak dat hij €1,- wint als hij gelijk heeft en €100,- verliest als er 225 of minder auto's worden geteld.

a 9pt Wat is de kans dat Jan deze weddenschap wint? Je mag ervan uitgaan dat de steekproefgrootte groot genoeg is om asymptotische benaderingen te mogen gebruiken.

b 6pt Wat is de verwachtingswaarde van de weddenschap? Als je a) niet hebt beantwoord, neem dan aan dat de kans op het winnen van de weddenschap 0.85 is.

Opgave 7 Zij A een verzameling van n verschillende reële getallen, i.e., $A = \{x_1, x_2, \ldots x_n\}$. Zij $x_{(i)}$ het element van A zodanig dat er precies $(i-1)$ elementen van A kleiner zijn dan $x_{(i)}$, $x_{(1)}$ is dus het kleinste element van A en $x_{(n)}$ het grootste element van A. Stel dat de elementen van A oorspronkelijk in een willekeurige volgorde geordend zijn waarbij elke volgorde even waarschijnlijk is.

a 4pt) Hoeveel mogelijke ordeningen van de verzameling A bestaan er?

b 6pt) Bubblesort is een algoritme om n elementen op volgorde te leggen, zodanig dat het het kleinste element op positie 1 is, het een-na-kleinste element op positie 2, ... , en het grootste element op positie n. Bubble-sort kijkt naar naast-elkaar gelegen elementen en verwisselt ze van volgorde als het element op positie i groter is dan het element op positie $i+1$. Het algoritme vergelijkt (en verwisselt indien nodig) eerst de elementen op positie 1 en 2, dan op positie 2 en 3, tot op positie $n-1$ en n. Dan weet je zeker dat het grootste element op positie n is. Vervolgens start je opnieuw en vergelijk je (en verwissel je indien nodig) eerst de elementen op positie 1 en 2, dan op positie 2 en 3, tot op positie $n-2$ en $n-1$. Nu ligt ook het element $x_{(n-1)}$ op positie $n-1$. Deze procedure wordt herhaald tot alle elementen op volgorde liggen.

Voor deze vraag zijn we geïnteresseerd in het aantal keren k dat elementen verwisseld worden. Het aantal verwisselingen is gelijk aan het aantal paren (i, j), met $i < j$, in de oorspronkelijke configuratie waarvoor $x_i > x_j$.

Definieer $P_n(k)$ als de kans dat je k keer 2 elementen van positie moet wisselen voordat de n elementen op volgorde liggen.

Geef de verwachtingswaarde van het aantal verwisselingen voordat alle elementen op volgorde liggen.

c 10pt) Bonusvraag: Bepaal de variantie van het aantal verwisselingen. (Advies, ga pas aan de bonusvraag werken als je alle andere opgaven af hebt).

Einde.