Tentamen Inleiding Kansrekening en Statistiek

2 juli 2015

- Elke opgave dient **op een apart blad** ingeleverd te worden.
- Zet **op elk blad** je naam en studentnummer.
- Het gebruik van rekenmachines, het boek, aantekeningen of andere hulpmiddelen is absoluut niet toegestaan. Het is i.h.b. ook absoluut niet toegestaan een telefoon o.i.d. op tafel te hebben.
- Geef telkens een beknopt bewijs van je antwoord of een berekening/toelichting, tenzij duidelijk in de opgave staat aangegeven dat dit niet hoeft.
- **NOTA BENE:** het gebruiken van aparte bladen voor elke opgave met daarop telkens je naam en studentnummer is 10 van de 100 punten waard.

Opgave 1. (a: 6pt, b: 6pt, c: 6pt, d: 6pt, e: 6pt).
De stochasten X, Y hebben een gemeenschappelijke kansdichtheid gegeven door de volgende functie:

$$f_{X,Y}(x,y) = \begin{cases} e^{-y+c} & \text{als } 0 < x < y, \\ 0 & \text{anders.} \end{cases}$$

Hier is $c \in \mathbb{R}$ een nader te bepalen constante.

(a) Bewijs of weerleg: X en Y zijn onafhankelijk.

(b) Geef de waarde van c.

(c) Geef de (marginale) cumulatieve verdelingsfuncties F_X en F_Y van X en Y.

(d) Definieer $D := Y - X$. Laat zien dat D een exponentiële verdeling heeft met parameter 1.

(e) Laat zien dat X en D onafhankelijk zijn.

(*Hint:* Je mag zonder bewijs gebruiken dat twee stochasten A, B onafhankelijk zijn dan en slechts dan als $\mathbb{P}(A \leq a, B \leq b) = \mathbb{P}(A \leq a)\mathbb{P}(B \leq b)$ voor alle $a, b \in \mathbb{R}$.)

Z.O.Z.

Opgave 2. (a: 6pt, b: 6pt, c: 6pt, d: 6pt, e: 6pt).
In een zonnig land wordt een referendum gehouden waarbij de kiezers moeten kiezen tussen twee alternatieven A en B. We zijn geïnteresseerd in de fractie p van de stemmers die op alternatief A zullen stemmen. We vragen hiertoe aan een volledig willekeurige steekproef van n kiezers waar ze op zullen stemmen. Hun antwoorden kunnen we modelleren met i.i.d. Bernouilli(p) stochasten $X_1, \ldots, X_n$. Hierbij is $X_i = 1$ als de i-de kiezer op A stemt.

(a) Beschouw de schatter $T := \overline{X}$ van p. Bereken MSE(T),

(b) Laat zien dat het interval

$$\left(\frac{2\overline{X} + \frac{z^2_{\alpha/2}}{n} - \sqrt{4\overline{X}(1-\overline{X})\frac{z^2_{\alpha/2}}{n} + \frac{z^4_{\alpha/2}}{n^2}}}{2(1+\frac{z^2_{\alpha/2}}{n})}, \frac{2\overline{X} + \frac{z^2_{\alpha/2}}{n} + \sqrt{4\overline{X}(1-\overline{X})\frac{z^2_{\alpha/2}}{n} + \frac{z^4_{\alpha/2}}{n^2}}}{2(1+\frac{z^2_{\alpha/2}}{n})} \right)$$

een benaderend betrouwbaarheidsinterval voor p is met betrouwbaarheid $1-\alpha$, waarbij $z_\alpha = \Phi^{-1}(1-\alpha)$ het $(1-\alpha)$-kwantiel van de standaard normale verdeling voorstelt.

(c) Stel nu we willen de nulhypothese $H_0 : p \leq \frac{1}{2}$ toetsen tegen de alternatieve hypothese $p > \frac{1}{2}$. Laat met het oogmerk hierop zien dat

$$\sup_{0<p\leq\frac{1}{2}} \left\{ 1 - \Phi\left(\frac{\frac{1}{2} + \frac{z_\alpha}{2\sqrt{n}} - p}{\sqrt{\frac{p(1-p)}{n}}} \right) \right\} = \alpha,$$

en leid hieruit af dat de toets die H_0 verwerpt als $\overline{X} > \frac{1}{2} + \frac{z_\alpha}{2\sqrt{n}}$ (bij benadering) significatieniveau α heeft. Hier is Φ natuurlijk als altijd de cdf van de standaard-normale verdeling. (*Hint:* Merk eerst op dat Φ stijgend is en beredeneer dat $(\frac{1}{2} + \frac{z_\alpha}{2\sqrt{n}} - p)/\sqrt{\frac{p(1-p)}{n}}$ dalend is voor $0 < p \leq 1/2$. Je hoeft hier niet per se voor te differentiëren.)

Het blijkt dat het aantal inwoners van het land dusdanig klein is dat het u lukt om alle kiezers te spreken voordat zij hun stem uitbrengen en u weet nu dus precies het aantal a van mensen dat op alternatief A stemt en het aantal b van mensen dat op alternatief B stemt. (We mogen ervan uitgaan dat niemand zijn of haar keuze wijzigt nadat ze u hebben gesproken en dat niemand tegen u liegt.) Het blijkt dat $a > b$. Tijdens het tellen van de stemmen worden alle stembriefjes op een grote hoop gegooid, en worden ze vervolgens in een volledig willekeurige volgorde geteld. Er is een groot bord waarop tijdens het tellen steeds de huidige stand wordt bijgehouden (er staat hoeveel van de stemmen tot nu toe voor A waren en hoeveel voor B). We zijn geïnteresseerd in de kans dat, tijdens de hele telling, het alternatief A altijd (strikt) meer van de tot dan toe getelde stemmen heeft dan alternatief B.

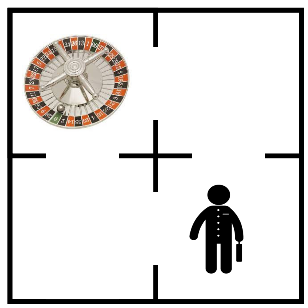
(d) Laat zien dat $\mathbb{P}(\text{de laatst getelde stem is voor } A) = \frac{a}{a+b}$.

(e) Bewijs de volgende stelling van Bertrand uit 1887:

$$\mathbb{P}(\text{alternatief } A \text{ ligt voor tijdens de hele telling}) = \frac{a-b}{a+b},$$

als $a \geq b \geq 0$ en $(a,b) \neq (0,0)$. (Als $a < b$ dan is de genoemde kans natuurlijk gelijk aan nul.)
(*Hint:* Gebruik inductie naar a en b met als inductie hypothese "de bewering is waar voor alle paren (a', b') met $a' \leq a, b' \leq b, a' \geq b' \geq 0$ en $(a', b') \notin \{(0,0),(a,b)\}$". Splits uit naar de waarde van de laatst getelde stem, en gebruik het antwoord uit het vorige onderdeel.)

Opgave 3. (a: 6pt, b: 6pt, c: 6pt, d: 6pt, e: 6pt).
Om te vieren dat u het tentamen inleiding kansrekening en statistiek met vlag en wimpel heeft gehaald, gaat u zodirect natuurlijk eerst stevig drinken en daarna naar het casino. In het casino aangekomen strompelt u in uw dronkenschap van kamer naar kamer, telkens volkomen willekeurig met gelijke kans één van de aangrenzende kamers kiezend. De plattegrond van het casino is als hieronder aangegeven. In één van de kamers bevindt zich een roulettetafel, en de kamer waarin u begint is aangegeven met een icoontje van een persoon met een boekentas (dit icoontje stelt uw persoon voor).



(a) Stel een stelsel vergelijkingen op voor het verwachte aantal deuren (deuropeningen) waar u doorheen zult gaan alvorens u voor de eerste keer in de kamer met de roulettetafel komt. (**Nota bene:** een toelichting is hier niet nodig.)

(b) Los dit stelsel op.

Zoals algemeen bekend is bij roulette de kans op rood (of zwart) gelijk aan $p := 18/37$; en als u 1 euro op rood (of zwart) inzet krijgt u bij winst uw inzet plus een extra euro terug en bent u bij verlies uw euro kwijt. U heeft een aanzienlijk vermogen van N euro op zak. Eenmaal bij de roulettetafel aangekomen blijft u telkens 1 euro op rood inzetten totdat u bankroet bent. (Uit de handout “gamblers ruin” en bijbehorende opgaven weten we dat u met kans 1 uiteindelijk bankroek zult gaan.) We zijn geinteresseerd in het maximale vermogen M dat u ooit op zak zult hebben. Dus als u bijvoorbeeld het eerste spel wint en daarna alleen maar verliest dan is $M = N + 1$.

(c) Laat zien dat

$$\mathbb{P}(M \geq m) = \begin{cases} 1 & \text{als } m \leq N, \\ \frac{\left(\frac{1-p}{p}\right)^N - 1}{\left(\frac{1-p}{p}\right)^m - 1} & \text{als } m > N. \end{cases}$$

(**Nota bene:** u mag zonder bewijs resultaten uit de handout gamblers ruin gebruiken mits u zeer duidelijk vermeldt wat u precies gebruikt.)

Z.O.Z.

(d) Stel X is een stochast die waarden in $\mathbb{N} \cup \{0\}$ aanneemt. Laat zien dat

$$\mathbb{E}X = \sum_{n=1}^{\infty} \mathbb{P}(X \geq n).$$

(*Hint:* Merk op dat $\mathbb{E}X = \sum_{k=1}^{\infty} k\mathbb{P}(X = k) = \sum_{k=1}^{\infty} \sum_{\ell=1}^{k} \mathbb{P}(X = k)$ en verwissel de volgorde van sommatie.)

(e) Uit de vorige twee onderdelen volgt dat

$$\mathbb{E}M = N + \sum_{m>N} \frac{\left(\frac{1-p}{p}\right)^N - 1}{\left(\frac{1-p}{p}\right)^m - 1}.$$

Laat zien dat $\mathbb{E}M \leq N + 18$.
(*Hint:* merk op dat $\frac{a-1}{b-1} \leq \frac{a}{b}$ als $b > a > 1$.)

Moraal: een mogelijke gokstrategie zou kunnen zijn om een zeer groot bedrag te lenen en dan zodra als men "in de plus" staat te stoppen, het geleende bedrag terug te betalen en het verschil te houden. Uit dit laatste antwoord zien we echter dat u bij een dergelijke strategie niet meer dan 18 euro kunnen verwachten te verdienen, zelfs als u in de toekomst kunt kijken en precies kunt weten vanaf welk moment uw vermogen nooit meer hoger zal worden dan de huidige waarde.