Universiteit Utrecht — Boedapestlaan 6
Mathematisch Instituut — 3584 CD Utrecht

Inleiding Kansrekening en Statistiek Uitwerkingen Tentamen, 2015-16

* **Elke opgave dient op een apart blad ingeleverd te worden**.
* Zet op elk blaadje dat je inlevert je naam en collegekaartnummer. Zet op het eerste blad ook de naam van je werkcollegebegeleider.
* Je mag een eenvoudige rekenmachine gebruiken, het informatie A4tje, de standaard normale tabel en de t-verdeling tabel.
* Als je een onderdeel niet kan oplossen, ga dan verder met het volgende. Je mag gerust gebruik maken van wat er in de tekst van een onopgelost onderdeel staat. Geef niet alleen antwoorden, maar laat de hele redenering zien die tot het antwoord leidt.

(1) Stel dat de simultane (gezamenlijke) kansdichtheid van X en Y gegeven wordt door

$$f(x,y) = \begin{cases} x+y & \text{als } 0<x<1,\ 0<y<1 \\ \\ 0 & \text{elders.} \end{cases}$$

(a) Laat zien dat de marginale kansdichtheden f_X en f_Y van X en Y gegeven zijn door

$$f_X(x) = \begin{cases} x+1/2 & 0<x<1 \\ 0 & \text{elders,} \end{cases}$$

en

$$f_Y(y) = \begin{cases} y+1/2 & 0<y<1 \\ 0 & \text{elders,} \end{cases}$$

(0.5 punten)

(b) Bepaal $\text{Cov}(X,Y)$. (0.5 punten)

(c) Bepaal $P(X+Y>1/3)$. (0.5 punten)

(d) Bepaal $P(XY>1/2)$. (0.5 punten)

Uitwerking (a): Eerst $f_X(x)=0$ voor $x\notin(0,1)$. Als $x\in(0,1)$, dan

$$f_X(x) = \int_0^1 (x+y)dy = xy + y^2/2|_0^1 = x+1/2.$$

Evenzo, $f_Y(y)=0$ voor $y\notin(0,1)$, en voor $0<y<1$, $f_Y(y)=y+1/2$.

Uitwerking (b): $\text{Cov}(X,Y)=E(XY)-E(X)E(Y)$. Nu

$$E(XY) = \int_)^1 \int_0^1 xy(x+y)\,dy\,dx = \int_0^1 (x^2/2+x/3)\,dx = 1/3,$$

en

$$E(X) = \int_0^1 x(x+1/2)\,dx = 7/12,$$

en door symmetrie $E(Y)=7/12$. Dus

$$\text{Cov}(X,Y) = 1/3 - 49/144 = -1/144.$$

Uitwerking (c):

$$P(X+Y>1/3) = 1-P(X+Y\le 1/3) = 1-\int_0^{1/3}\int_0^{1/3-x}(x+y)\,dy\,dx = 80/81.$$

Uitwerking (d):

$$P(XY>1/2) = \int_{1/2}^1\int_{1/2x}^1 (x+y)\,dy\,dx = \int_{1/2}^1 x - \frac{1}{8x^2}\,dx = 1/4.$$

(2) (a) Zij X_1, X_2, X_3 onafhankelijke geometrisch verdeelde stochasten met parameter p, en $Z = X_1 + X_2 + X_3$. Laat zien dat voor $k \geq 3$,

$$P(X_1 + X_2 + X_3 = k) = \frac{1}{2}(k-2)(k-1)p^3(1-p)^{k-3}.$$

(1 punt)

(b) Zij $X_1, X_2, \ldots, X_n$ onafhankelijke stochasten met $E(X_i) = \mu \in (-\infty, \infty)$ en $\mathrm{Var}(X_i) = \sigma^2 \in (0, \infty)$ voor $i = 1, \cdots, n$. Schrijf $\overline{X}_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$. Laat zien dat voor alle $a > 0$,

$$P\Big(n^{1/4}\Big|\frac{\overline{X}_n - \mu}{\sigma}\Big| \geq a\Big) \leq \frac{1}{a^2\sqrt{n}}.$$

(1 punt)

Uitwerking(a): We kunnen Z interpreteren als de benodigde tijd om tot drie successen te komen. Dus in de gebeurtenis $\{Z = k\}$, de laatste herhaling is een succes en de eerste twee successen in de eerste $k-1$ herhalingen plaatsvinden. Dus

$$P(Z = k) = p\, C_{k-1,2}p^2(1-p)^{k-3} = \frac{1}{2}(k-2)(k-1)p^3(1-p)^{k-3}.$$

Uitwerking(b): Definieer $Y = \frac{n^{1/4}}{\sigma}\overline{X}_n$. Dan $E(Y) = \frac{n^{1/4}}{\sigma}\mu$, en door onafhankelijkheid

$$\mathrm{Var}(Y) = \mathrm{Var}\Big(\frac{n^{1/4}}{\sigma}\overline{X}_n\Big) = \frac{n^{1/2}}{\sigma^2}\frac{1}{n}\sigma^2 = \frac{1}{n^{1/2}}.$$

Met behulp van Chebyshev's ongelijkheid krijgen we

$$P\Big(n^{1/4}\Big|\frac{\overline{X}_n - \mu}{\sigma}\Big| \geq a\Big) = P\Big(\Big|Y - E(Y)\Big| \geq a\Big) \leq \frac{Var(Y)}{a^2} = \frac{1}{a^2 n^{1/2}}.$$

(3) Gegeven zijn twee onafhankelijke stochasten X en Y.
(a) Stel dat X en Y standaard normaal verdeeld zijn. Bepaal $P(2X \leq 1 - 3Y)$. (1 punt)
(b) Stel nu dat X en Y uniform verdeeld op $(0,1)$ zijn. Bepaal $P(Y \geq 2X)$, en $P(Y \geq 2X \,|\, X + Y \leq 1)$. (1 punt)

Uitwerking (a): Merk op dat $2X$ normaal $N(0,4)$ verdeeld is, en $3Y$ normaal $N(0,9)$ verdeeld is. Dus $2X + 3Y$ is normaal $N(0,13)$ verdeeld, en de stochast $\frac{2X+3Y}{\sqrt{13}}$ is standaard normaal verdeeld. Hieruit volgt dat

$$P(2X \leq 1 - 3Y) = P(2X + 3Y \leq 1) = P(\frac{2X+3Y}{\sqrt{13}} \leq \frac{1}{\sqrt{13}}) = .6103,$$

waarbij de numerieke waarde uit de normale tabel is afgelezen.

Uitwerking (b): Merk op dat de simultane kansdichtheid van X en Y wordt gegeven door

$$f(x,y) = \begin{cases} 1 & 0 \leq x, y \leq 1 \\ 0, & \text{anders.} \end{cases}$$

Dus,

$$P(Y \geq 2X) = \int_0^{1/2}\int_{2x}^{1} 1\, dy\, dx = \frac{1}{4},$$

$$P(X + Y \leq 1) = \int_0^{1}\int_{0}^{1-x} 1\, dy\, dx = \frac{1}{2},$$

$$P(Y \geq 2X, X + Y \leq 1) = \int_0^{1/3}\int_{2x}^{1-x} 1\, dy\, dx = \frac{1}{6},$$

en

$$P(Y \geq 2X \,|\, X + Y \leq 1) = \frac{P(Y \geq 2X, X + Y \leq 1)}{P(X + Y \leq 1)} = \frac{1}{3}.$$

(4) Gegeven is een rij $X_1, X_2, \ldots, X_n$ van onafhankelijk gelijk verdeelde stochasten uit een uniforme verdeling op het interval $[-\theta, \theta]$, waarbij θ onbekend is.

(a) Toon aan dat $T = \dfrac{3}{n}(X_1^2 + X_2^2 + \cdots + X_n^2)$ een zuivere schatter is voor θ^2. (0.5 punten)

(b) Bepaal de *mean squared error* MSE(T) van T. (1 punt)

Uitwerking (a): Merk op dat $E(X_i^2) = \theta^2/3$, dus

$$E(T) = \frac{3}{n}\left(\frac{\theta^2}{3} + \cdots + \frac{\theta^2}{3}\right) = \frac{3}{n} \cdot n \cdot \frac{\theta^2}{3} = \theta^2.$$

Omdat $E(T) = \theta^2$, is T een zuivere schatter voor θ^2.

Uitwerking (b): Omdat T een zuivere schatter is, dan is MSE$(T) = \text{Var}(T)$. Merk op dat

$$E(X_i^4) = \int_{-\theta}^{\theta} x^4 \frac{1}{2\theta}\, dx = \theta^4/5,$$

zodat

$$\text{Var}(X_i^2) = E(X_i^4) - (E(X_i^2))^2 = 4\theta^4/45.$$

Uit onafhankelijkheid van de stochasten X_i volgt nu dat

$$\text{MSE}(T) = \text{Var}(T) = \left(\tfrac{3}{n}\right)^2 \sum_{i=1}^{n} \text{Var}(X_i^2) = \frac{9}{n} \cdot \frac{4\theta^4}{45} = \frac{4\theta^4}{5n}.$$

(5) Zij $x_1, x_2, \cdots, x_n$ een realisatie (dataset) van n onafhankelijke gelijk verdeelde stochasten $X_1, X_2, \cdots, X_n$. Neem aan dat ieder X_i Poisson verdeeld is met onbekende parameter μ.

(a) Bepaal de maximum likelihood schatter van μ gebasseerd op de data. (1 punt)

(b) Veronderstel dat n voldoende groot is zodat de stochast $\dfrac{\overline{X}_n - E(\overline{X}_n)}{\sqrt{\text{Var}(\overline{X}_n)}}$ gezien kan worden als standaard normaal verdeeld. Bepaal een 95% betrouwbaarheids interval voor μ gebasseerd op de gegeven data. (1.5 punten)

Uitwerking(a) De likelihood functie wordt gegeven door

$$L(\mu) = P(X_1 = x_1) \cdots P(X_n = x_n) = e^{-n\mu} \frac{\mu^{x_1 + \cdots + x_n}}{x_1! \cdots x_n!},$$

en de loglikelihood wordt gegeven door

$$\ell(\mu) = -n\mu + \sum_{i=1}^{n} x_i \ln \mu - \ln(x_1! \cdots x_n!).$$

We nemen de afgeleide,

$$\ell'(\mu) = -n + \frac{\sum_{i=1}^n x_i}{\mu}.$$

Nu leidt $\ell'(\mu) = 0$ tot $\mu = \dfrac{\sum_{i=1}^n x_i}{n}$. Omdat $\ell''(\frac{\sum_{i=1}^n x_i}{n}) = \frac{-1}{\sum_{i=1}^n x_i} < 0$, dan is $\dfrac{\sum_{i=1}^n x_i}{n}$ de maximum likelihood schatting van μ.

Uitwerking(b) Merk op dat $E(X_i) = \mu = \text{Var}X_i$. Dus $E(\overline{X}_n) = \mu$, $\text{Var}(\overline{X}_n) = \dfrac{\mu}{n}$ en $\dfrac{\overline{X}_n - E(\overline{X}_n)}{\sqrt{\text{Var}(\overline{X}_n)}} = \dfrac{\overline{X}_n - \mu}{\sqrt{\mu}/\sqrt{n}}$. Voor n voldoende groot

$$P\Big(-z_{0.025} < \frac{\overline{X}_n - \mu}{\sqrt{\mu}/\sqrt{n}} < z_{0.025}\Big) \approx 0.95.$$

Nu

$$\begin{aligned} -z_{0.025} < \frac{\overline{X}_n - \mu}{\sqrt{\mu}/\sqrt{n}} < z_{0.025} \quad &\Leftrightarrow \quad \left(\frac{\overline{X}_n - \mu}{\sqrt{\mu}/\sqrt{n}}\right)^2 < z_{0.025}^2 \\ &\Leftrightarrow \quad \mu^2 - \mu(2\overline{X}_n + z_{0.025}^2/n) + \overline{X}_n^2 < 0. \end{aligned}$$

De laatste ongelijkheid is equivalent met

$$\frac{2\overline{X}_n + \frac{z^2_{0.025}}{n} - \sqrt{(2\overline{X}_n + \frac{z^2_{0.025}}{n})^2 - 4\overline{X}_n^2}}{2} < \mu < \frac{2\overline{X}_n + \frac{z^2_{0.025}}{n} + \sqrt{(2\overline{X}_n + \frac{z^2_{0.025}}{n})^2 - 4\overline{X}_n^2}}{2}.$$

Dus een 95% betrouwbaarheids interval voor μ gebasseerd op de gegeven data wordt gegeven door

$$\Big(\frac{2\overline{x}_n + \frac{3.8416}{n} - \sqrt{(2\overline{x}_n + \frac{3.8416}{n})^2 - 4\overline{x}_n^2}}{2}, \frac{2\overline{x}_n + \frac{3.8416}{n} + \sqrt{(2\overline{x}_n + \frac{3.8416}{n})^2 - 4\overline{x}_n^2}}{2}\Big).$$

waarbij $\overline{x}_n = \dfrac{x_1 + \cdots + x_n}{n}$, en $z^2_{0.025} = (1.96)^2 = 3.8416$.