Universiteit Utrecht
Mathematisch Instituut

Universiteit Utrecht

Boedapestlaan 6
3584 CD Utrecht

Oefententamen kansrekening 2010/11

1. Stel dat de simultane dichtheid van X en Y gegeven wordt door

$$f(x,y) = \begin{cases} \frac{4y}{x} & \text{als } 0 < y < x < 1 \\ 0 & \text{elders.} \end{cases}$$

   (a) Bepaal de marginale kansdichtheden f_X en f_Y van X en Y.
   (b) Bepaal $E(XY)$.
   (c) Bepaal $P(X+Y \leq 1)$.
   (d) Bepaal de kansdichtheid van $Z = X^2$.

uitwerking (a): We bepalen eerst de marginale dichtheid van X. Voor $0 < x < 1$,

$$f_X(x) = \int_0^x \frac{4y}{x}\,dy = \frac{2y^2}{x}|_0^x = 2x.$$

Dus

$$f_X(x) = \begin{cases} 2x & 0 < x < 1; \\ 0 & \text{elders.} \end{cases}$$

Voor $0 < y < 1$,

$$f_Y(y) = \int_y^1 \frac{4y}{x}\,dy = 4y \ln y|_y^1 = -4y \ln y.$$

Dus

$$f_Y(y) = \begin{cases} -4y \ln y & 0 < y < 1; \\ 0 & \text{elders.} \end{cases}$$

uitwerking (b):

$$E(XY) = \int_0^1 \int_0^x xy\frac{4y}{x}\,dy\,dx = \int_0^1 \frac{4y^3}{3}|_0^x dx = \int_0^1 \frac{4x^3}{3} = 1/3$$

.

uitwerking (c):

$$P(X+Y \leq 1) = \int_0^{1/2} \int_0^x \frac{4y}{x}\,dy\,dx \; + \int_{1/2}^1 \int_0^{1-x} \frac{4y}{x}\,dy\,dx \; = \; \ln 4 - 1 \approx 0.386.$$

uitwerking (d): Voor $0 < z < 1$,

$$P(Z \leq z) = P(X^2 \leq z) = P(X \leq \sqrt{z}) = \int_0^{\sqrt{z}} 2x\,dx = x^2|_0^{\sqrt{z}} = z.$$

Dus

$$f_Z(z) = \begin{cases} z & 0 < z < 1; \\ 0 & \text{elders.} \end{cases}$$

In andere worden Z is uniform verdeeld op het interval $(0, 1)$.

2. Zij $X_1, X_2, \ldots$ een rij van onafhankelijke gelijk verdeelde stochasten met $P(X_i = 1) = P(X_i = 0) = 1/2$. Definieer de stochasten Y_i, $i = 1, 2, \ldots$, door

$$Y_i = \begin{cases} X_i, & \text{als } X_i = 1, \\ X_i - 1, & \text{als } X_i = 0. \end{cases}$$

(a) Bepaal $P(\sum_{i=1}^{10} Y_i = 0)$.

(b) Bepaal $\lim_{n\to\infty} P(\sum_{i=1}^{n} Y_i < \sqrt{n})$.

(c) Bepaal $\lim_{n\to\infty} P(\sum_{i=1}^{n} Y_i > 0)$.

uitwerking (a): De stochasten $Y_1, \cdots, Y_n$ zijn onafhenakelijk gelijk verdeeld met $P(Y_i = 1) = P(Y_i = -1) = 1/2$. De gebeurtenis $\{\sum_{i=1}^{10} Y_i = 0\}$ plaats vindt dan en slechts dan als precies 5 van de stochasten $Y_1, \cdots, Y_n$ waarde 1 heeft. Vanwege onafhenkelijkheid ieder uitkomst met precies 5 enen kans $1/2^{10}$ heeft. Dus

$$P(\sum_{i=1}^{10} Y_i = 0) = \binom{10}{5} 2^{-10}.$$

uitwerking (b): Merk op dat $E(Y_i) = 0$, $\mu = E(Y_i^2) = 1$, en $\sigma^2 = \text{Var}(Y_i) = 1$. Dus $E(\sum_{i=1}^{n} Y_i) = 0$ en $\text{Var}(\sum_{i=1}^{n} Y_i) = \sqrt{n}$. Door de Centrale Limiet Stelling,

$$\lim_{n\to\infty} P(\sum_{i=1}^{n} Y_i < \sqrt{n}) = \lim_{n\to\infty} P(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i < 1) = \int_{-\infty}^{1} \frac{1}{2\pi} e^{-\frac{x^2}{2}}\,dx = 0.1587.$$

uitwerking (c):

$$\lim_{n\to\infty} P(\sum_{i=1}^{n} Y_i > 0) = \lim_{n\to\infty} P(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i > 0) = 1 - \int_{-\infty}^{0} \frac{1}{2\pi} e^{-\frac{x^2}{2}}\,dx = 1/2.$$

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

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

(b) Laat zien dat $\text{Var}X_i^2 = 4\theta^4/45$.

(c) Is $\sqrt{T}$ een zuivere schatter voor θ? Zo ja, toon dit aan. Zo nee, geef aan of de onzuiverheid $E(\sqrt{T}) < \theta$ positief of negatief is.

uitwerking (a): Merkop 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): Merkop dat $E(X_i^4) = \theta^4/5$, zodat $\mathrm{Var}(X_i^2) = E(X_i^4) - (E(X_i^2))^2 = 4\theta^4/45$.

uitwerking (c): Merk op dat de functie $g(x) = -\sqrt{x}$ is strict convex, dus door Jensen's ongelijkheid $E(-\sqrt{T}) > -\sqrt{E(T)}$ ofwel $E(\sqrt{T}) < \sqrt{E(T)} = \sqrt{\theta^2} = \theta$. Dus heeft $\sqrt{T}$ een negatieve onzuiverheid.

4. Een exploratie maatschappij doet boringen in een gebied waarbij er kans θ is dat er olie wordt aangeboord. Stel dat er 5 boringen verricht worden totdat er eindelijk olie wordt gevonden. Op een iets andere plaats in het gebied wordt opnieuw (en onafhankelijk van de eerdere boringen) een serie boringen uitgevoerd totdat bij de 10e boring voor het eerst olie gevonden wordt. Steeds op iets andere plekken in het gebied worden er nog vier series van dit soort boringen uitgevoerd: de benodigde aantallen pogingen in deze vier series zijn 8, 7, 6, en 8.

   Dus alles tezamen zijn

$$5 \quad 10 \quad 8 \quad 7 \quad 6 \quad 8$$

   de aantallen pogingen in elke serie tot en met succes.

   (a) Laat zien dat de likelihood functie die bij deze steekproef van grootte 6 hoort gelijk is aan:

$$L(\theta) = \theta^6(1-\theta)^{38}, \qquad 0 \le \theta \le 1.$$

   (b) Wat is de maximum likelihood schatting voor θ?

uitwerking (a): Als in een serie van metingen X het aantal pogingen is totdat voor het eerst succes optreedt, dan is X geometrisch verdeeld met parameter θ, en dus geldt er voor de likelihood dat

$$\begin{aligned} L(\theta) &= P(X_1 = 5, X_2 = 10, X_3 = 8, X_4 = 7, X_5 = 6, X_6 = 8) \\ &= P(X_1 = 5)P(X_2 = 10)P(X_3 = 8)P(X_4 = 7)P(X_5 = 6)P(X_6 = 8) \\ &= (1-\theta)^4\theta(1-\theta)^9\theta(1-\theta)^7\theta(1-\theta)^6\theta(1-\theta)^5\theta(1-\theta)^7\theta \\ &= (1-\theta)^{38}\theta^6, \end{aligned}$$

met $0 \le \theta \le 1$.

uitwerking (b): Omdat we net zagen dat $L(\theta) = \theta^6(1-\theta)^{38}$, is het voldoende om te bepalen voor welke waarde van $\theta \in [0,1]$ $L(\theta)$ een maximum heeft. Dit kunnen

we doen door L' te bepalen, maar eenvoudiger is het om eerst de loglikelihood $\ell(\theta)$ te bepalen, en deze dan te differentieren naar θ. Er geldt dat:

$$\ell(\theta) = \ln L(\theta) = 6 \ln \theta + 38 \ln(1 - \theta).$$

Maar dan is

$$\ell'(\theta) = \frac{6}{\theta} - \frac{38}{1 - \theta},$$

zodat $\ell'(\theta) = 0$ dan en slechts dan als $\theta = \frac{6}{44}$. Ga zelf even na dat dit inderdaad een maximum is.