

Uitwerkingen tentamen 6 juli 2012

1. We stellen T de gebeurtenis “test geeft positief resultaat”, F de gebeurtenis, “chauffeur heeft gefraudeerd”, V de gebeurtenis, “chauffeur heeft vergissing gemaakt” C de gebeurtenis, “chauffeur heeft correct ingevuld” Uit de gegevens volgt

$$P(T|F) = 0.9, P(T|V) = 0.09, P(T|C) = 0.01, P(F) = 0.03, P(V) = 0.07, P(C) = 0.9.$$

a)

$$P(T) = P(T|F)P(F) + P(T|V)P(V) + P(T|C)P(C) = 0.0423$$

b) Uit Bayes volgt, gebruik makend van onderdeel a)

$$P(V|T) = \frac{P(T|V)P(V)}{P(T)} = \frac{(0.09)(0.07)}{0.0423} = 0.149$$

c) Voor de tweede test geldt

$P(T_2|F) = 0.95$, $P(T_2|V) = 0.04$, $P(T_2|C) = 0.01$. Hieruit volgt, zoals in onderdeel a) dat $P(T_2) = 0.0403$, en dus, $P(T_2^c) = 1 - P(T_2) = 0.9597$. De gevraagde kans is $P(F|T_2^c \cap T_1)$: gegeven is immers dat het eerste testresultaat positief is en het tweede negatief. Gebruik makend van Bayes, en de onafhankelijkheid van de twee testen, geldt

$$\begin{aligned} P(F|T_2^c \cap T_1) &= \frac{P(T_2^c \cap T_1|F)P(F)}{P(T_2^c \cap T_1)} \\ &= \frac{P(T_2^c|F)P(T_1|F)P(F)}{P(T_2^c)P(T_1)} \\ &= \frac{(0.9)(0.05)(0.03)}{(0.04423)(0.9597)} = 0.00333 \end{aligned} \quad (1)$$

2. a) Continuïteit van F_X in $x = 0$, $x = 1$ geeft $a = 0, b = 1$. Normalisatie geeft $C = 6$.
- b) $f_X(x) = F'_X(x) = 6(x^2 - x^3)$ voor $x \in [0, 1]$ en $f_X(x) = 0$ elders.
- c) $\mathbb{E}(X) = \int_0^1 x6(x^2 - x^3) dx = 6(1/4 - 1/7) = 9/14$, $\mathbb{E}(X^2) = \int_0^1 x^2 6(x^2 - x^3) dx = 6(1/5 - 1/8) = 9/20$. $V(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = 9/20 - (9/14)^2 = 0.0367$.

- d) We berekenen eerst de cumulatieve verdeling. Gebruik makend van het feit dat X waarden aanneemt in $[0, 1]$, en dus niet-negatief kan zijn, volgt

$$\begin{aligned} P(-\ln(X^2) \geq x) &= P(X^2 \leq e^{-x}) = P(X \leq e^{-x/2}) \\ &= \begin{cases} 1, & x \leq 0 \\ 6((1/3)e^{-3x/2} - (1/6)e^{-3x}), & x \geq 0 \end{cases} \end{aligned}$$

Waarbij we in de tweede gelijkheid gebruikt hebben dat X niet negatief is.

De kansdichtheid is dan gelijk aan minus de afgeleide van $P(-\ln(X^2) \geq x)$, wat geeft

$$f_Y(y) = \begin{cases} 0 & \text{voor } y \leq 0 \\ (3e^{-3y/2} - 3e^{-3y}) & \text{voor } y \geq 0 \end{cases}$$

3. Alle kansdichtheden uit deze opgave zijn enkel verschillend van nul voor $x, y \in [0, 1]$, en zijn nul elders. We stellen verder dan ook dat $x, y \in [0, 1]$.

a) $f_X(x) = \int_0^1 f_{X,Y}(x, y) dy = 2x$, $f_Y(y) = \int_0^1 f_{X,Y}(x, y) dx = 3y^2$

- b) Deze kans is gelijk

$$\begin{aligned} P(X < Y) &= \int_0^1 \left(\int_0^y f_{X,Y}(x, y) dx \right) dy \\ &= \int_0^1 \left(\int_0^y dx (9xy^2 - 6x^3y^2 - 6xy^5 + 12x^3y^5) \right) dy \\ &= \int_0^1 ((9y^4/2) - (6y^6/4) - (6y^7/2) + (12y^9/4)) dy \\ &= (9/10 - 6/28 - 6/16 + 12/40) = 0.6107 \end{aligned}$$

- c) De conditionele kansdichtheid is

$$f_{X|Y=y}(x) = \frac{f_{X,Y}(x, y)}{f_Y(y)} = \frac{9xy^2 - 6x^3y^2 - 6xy^5 + 12x^3y^5}{3y^2} = 3x - 2x^3 - 2xy^3 + 4x^3y^3$$

De conditionele verwachting is dan

$$\mathbb{E}(X|Y = y) = \int_0^1 x(3x - 2x^3 - 2xy^3 + 4x^3y^3) dx = \frac{3}{5} + \frac{2}{15}y^3$$

Bijgevolg, $\mathbb{E}(X|Y) = \frac{3}{5} + \frac{2}{15}Y^3$

4. a) Voor de Poisson verdeling met parameter λ geldt $\mathbb{E}(X^2) = V(X) + (\mathbb{E}(X))^2 = \lambda + \lambda^2$. Er geldt dus $\lambda^2 + \lambda = 2$, wat geeft (rekening houdend met de restrictie $\lambda \geq 0$) $\lambda = 1$. Bijgevolg $P(X = 0) = e^{-\lambda} = e^{-1}$. De uitspraak is dus juist.
- b) Deze uitspraak is juist. Niet gecorreleerd volgt uit

$$\begin{aligned} \text{cov}(X - Y, X + Y) &= \text{cov}(X, X) - \text{cov}(Y, Y) - \text{cov}(Y, X) + \text{cov}(X, Y) \\ &= V(X) - V(Y) = 0 \end{aligned}$$

Niet onafhankelijk volgt bijvoorbeeld via berekening van de moment genererende functie:

$$\begin{aligned} \mathbb{E}(e^{t(X+Y)} e^{s(X-Y)}) &= \mathbb{E}(e^{(t+s)X} e^{(t-s)Y}) \\ &= \left(\frac{e^{t+s} - 1}{t+s} \right) \left(\frac{e^{t-s} - 1}{t-s} \right) \\ &= \frac{e^{2t} - e^{t+s} - e^{t-s} + 1}{t^2 - s^2} \end{aligned}$$

Dit is duidelijk niet het product van een functie van s met een functie van t . Bijgevolg zijn $X + Y$ en $X - Y$ niet onafhankelijk.

- c) Rekenregels van de variantie geven

$$V(2X - Y + 3) = 4V(X) + (-1)^2 V(Y) = 4V(X) + V(Y) = 4 + 1 = 5$$

De uitspraak is dus juist.

5. a) $\mathbb{E}(X) = (M'_X(t))_{t=0} = ((2^3)3(2-t)^{-4})_{t=0} = (2^3)(3)/2^4 = 3/2$,
 $\mathbb{E}(X^3) = (M'''_X(t))_{t=0} = (2^3 \cdot 3 \cdot 4 \cdot 5 (2-t)^{-6})_{t=0} = (2^3)(60)/2^6 = 15/2$
- b)

$$\mathbb{E}(e^{t(3X+4X'+5)}) = \mathbb{E}e^{3tX} \mathbb{E}e^{4tX'} e^{5t} = \left(\frac{2}{2-3t} \right)^3 \left(\frac{2}{2-4t} \right)^3 e^{5t}$$

6. a) Er geldt $V(X_i) \leq 2$ voor i even of oneven. Bijgevolg, gebruik makend van onafhankelijkheid:

$$V\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \sum_{i=1}^n V(X_i) \leq \frac{1}{n^2} \sum_{i=1}^n 2 = \frac{2}{n}$$

Gebruik makend van de Chebychev ongelijkheid vinden we dan

$$P\left(\left|\frac{1}{n} \sum_{i=1}^n X_i - \mu\right| > \epsilon\right) \leq \frac{V\left(\frac{1}{n} \sum_{i=1}^n X_i\right)}{\epsilon^2} \leq \frac{2}{n\epsilon^2}$$

Bijgevolg, voor alle $\epsilon > 0$

$$\lim_{n \rightarrow \infty} P \left(\left| \frac{1}{n} \sum_{i=1}^n X_i - \mu \right| > \epsilon \right) = 0$$

- b) Indien X_i normaal verdeeld zijn, dan geldt dat $\sum_{i=1}^n (X_i - \mu)$ normaal verdeeld is met verwachting nul en variantie $\sum_{i=1}^n \sigma_i^2$. Bijgevolg is voor $a > 0$ $\frac{1}{\sqrt{an}} \sum_{i=1}^n (X_i - \mu)$ normaal verdeeld met verwachting nul en variantie $\frac{1}{an} \sum_{i=1}^n \sigma_i^2$. Er geldt

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \sigma_i^2 = 3/2$$

dus moeten we $a = 3/2$ kiezen: dan geldt immers voor $n \rightarrow \infty$: $\frac{1}{an} \sum_{i=1}^n \sigma_i^2 \rightarrow 1$.

7. a) $a \neq 0$ dan $(X_1, X_2) \perp X_3$. Indien $a = 0$ dan $X_1 \perp X_2 \perp X_3$.
b)

$$\begin{aligned} & cov(X_1 - X_2, X_1 + X_2) \\ &= cov(X_1, X_1) - cov(X_2, X_1) + cov(X_1, X_2) - cov(X_2, X_2) \\ &= 1 - a + a - 1 = 0 \end{aligned}$$

Omdat het koppel $(X_1 - X_2, X_1 + X_2)$ multivariaat normaal verdeeld is impliceert dit ook onafhankelijkheid.

- c) Er geldt $V(X_1 - X_2) = V(X_1) + V(X_2) - 2cov(X_1, X_2) = 1 + 1 - 2(1/2) = 1$, $V(X_1 + X_2) = V(X_1) + V(X_2) + 2cov(X_1, X_2) = 1 + 1 + 2(1/2) = 3$ en $V(X_3) = 2$. Bijgevolg: $\alpha = 1, \beta = 1/3, \gamma = 1/2$.
8. a) $P^2 > 0$ (i.e., alle elementen van P^2 zijn strikt positief), dus irreducibel. Vanwege symmetrie is de periode van alle toestanden identiek. We kunnen van 1 naar 1 in 2, 3, 4, .. etc. stappen, dus periode is 1.
- b) Gebruik makend van de Markov eigenschap is deze kans

$$P(X_5 = 4 | X_3 = 1)P(X_7 = 4 | X_5 = 4) = (P^2)_{14}(P^2)_{44} = 2/27$$

- c) Indien we $f(i)$ de kans noemen dat vertrekkend van toestand i toestand 4 eerder getroffen wordt dan toestand 3, dan geldt

$$Pf = f$$

(met f de kolom met elementen $f(i), i = 1, \dots, 4$) en $f(4) = 1, f(3) = 0$. Stellen we $f(1) = x, f(2) = y$ dan geeft dit de vergelijkingen

$$\begin{cases} \frac{x}{3} + \frac{1}{3} = x \\ \frac{x}{2} + \frac{1}{4} = y \end{cases}$$

wat geeft $x = y = 1/2$. (Dit kan ook uit symmetrieoverwegingen worden afgeleid)

9. a) Deze uitspraak is fout. De bijbehorende covariantiematrix zou dan gelijk zijn aan

$$\begin{pmatrix} 2 & -4 \\ -4 & 2 \end{pmatrix}$$

Deze matrix heeft determinant -12 en is dus niet positief definit.

- b) Fout, neem

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

dan geldt

$$P^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Bijgevolg geldt voor *iedere kansverdeling* $\mu = (\mu_1, \mu_2)$ en voor iedere n dat $\mu P^{2n} = \mu$, en dus ook $\lim_{n \rightarrow \infty} \mu P^{2n} = \mu$. Deze Markov keten heeft echter als unieke stationaire verdeling $\mu = (1/2, 1/2)$.

- c) Deze uitspraak is juist. Voor het bewijs nemen we aan dat een $M > 0$ bestaat zodanig dat $|X_n|, |Y_n|, |X|, |Y| \leq M$. Er geldt dan

$$\begin{aligned} & P(|X_n Y_n - XY| > \epsilon) \\ & \leq P(|X_n(Y_n - Y)| > \epsilon/2) + P(|(X_n - X)Y| \geq \epsilon/2) \\ & \leq P(M|Y_n - Y| > \epsilon/2) + P(|X_n - X|M \geq \epsilon/2) \\ & \leq P\left(|Y_n - Y| > \frac{\epsilon}{2M}\right) + P\left(|X_n - X| \geq \frac{\epsilon}{2M}\right) \end{aligned}$$

De bewering volgt dan door het nemen van de limiet $n \rightarrow \infty$.