Toets Kansrekening voor INF en BIT Module 4 (Data en Informatie - 201300180)

vrijdag 14 juni 2016, 8.45-10.45 uur (Docent Dick Meijer, mod. coord. Klaas Sikkel)

Deze toets bestaat uit 6 opgaven. Het formuleblad en de tabel van de standaardnormale verdeling zijn toegevoegd. Een gewone, niet programmeerbare rekenmachine is toegestaan (geen GR).

1. In een populatie zijn de volgende gebeurtenissen gedefinieerd: $A =$ "persoon bezit een *Apple* computer" en $H =$ "persoon is Hoger opgeleid."
Indien we een willekeurige persoon kiezen uit de populatie, geldt: $P(A) = 0.3$ en $P(H) = 0.4$.

- **a.** Bereken $P(A \cup H)$, als (voor dit onderdeel) aangenomen mag worden dat A en H onderling onafhankelijk zijn.
- **b.** Bereken de kans dat een *Apple* bezitter hoger opgeleid is, indien $P(A|H) = 0.5$.

2. Na uitgebreide selectie en screening is er een groep van 25 internationale studenten overgebleven, die allen evenzeer in aanmerking komen voor de 6 beschikbare beurzen voor een master CS in Twente. Van deze 25 studenten komen er 10 uit Azië, 5 uit Afrika en 10 uit Zuid-Amerika.
Men besluit de beurzen via loting toe te kennen.
X is het aantal studenten uit Azië die een beurs krijgen.

- **a.** Bereken $P(X = 2)$ en $E(X)$.
- **b.** Als $Y =$ "het aantal Afrikanen die een beurs krijgen", bereken dan de kans $P(X = 2 \text{ en } Y = 2)$.

3. Hiernaast is in een tabel de simultane kansfunctie $P(X = x \text{ en } Y = y)$ gegeven.

- **a.** Bereken $P(X > Y)$.
- **b.** Bepaal $E(X)$ en $var(X)$.
- **c.** Zijn X en Y (onderling) onafhankelijk? Motiveer je antwoord kort.

$x \backslash y$	0	2	4
0	0.09	0.12	0.09
2	0.12	0.16	0.12
4	0.09	0.12	0.09

4. De kansdichtheid van een variabele X wordt gegeven door: $f(x) = \begin{cases} \frac{1}{8}x, & \text{als } 0 < x < 4 \\ 0, & \text{elders} \end{cases}$

- **a.** Bereken $P(X \leq 2)$.
- **b.** Bereken $E(X)$ en $var(X)$.
- **c.** Bepaal de kansdichtheid van $Y = \sqrt{X}$ (Geef vooraf het waardenbereik van Y.)
- **d.** Stel we hebben de beschikking over 72 o.o. variabelen $X_1, \ldots, X_{72}$, die dezelfde verdeling hebben als X, bereken (benader) dan de kans $P(\sum_{i=1}^{72} X_i \leq 200)$.
(Als je b. niet hebt opgelost, neem dan $E(X) = 2.5$ en $var(X) = 1$)

5. Het aanvangssalaris X van startende IT-specialisten, in duizenden Euro's bruto per jaar, is in een land gemiddeld k€ 35, bij een standaardafwijking van k€ 3.
Uit onderzoek is gebleken dat de normale verdeling (bij benadering) van toepassing is op X.
Het land kent een belastingvrije voet van 5 duizend Euro en een "*flattax*" van 40%, dus bij een salaris van X duizend Euro is de belasting $Y = (X - 5) \cdot 0.4$, dus $\boldsymbol{Y = 0.4X - 2}$.
(Opmerking: de kans op een aanvangssalaris onder de 5 duizend Euro is verwaarloosbaar klein).

a. Bereken $P(X > 40)$.
b. Bereken $P(X > 40|X > 35)$.
c. Welke waarde heeft $\rho(X, Y)$? (**Beredeneer** je antwoord.)
d. Verifieer de in c. gegeven waarde door $\rho(X, 0.4X - 2)$ te bepalen m.b.v. de definitie van de correlatiecoëfficiënt en eigenschappen van covariantie en variantie.

6. Voor een loket ziet een klant twee andere klanten voor hem staan: deze moeten eerst bediend worden, alvorens hij zelf aan de beurt is. De bedieningsduren (in minuten) van de klanten voor hem beschouwen we als o.o., exponentieel verdeelde variabelen X en Y, beide met parameter $\lambda = \frac{1}{3}$. $W = X + Y$ is dan de wachttijd van de derde klant.

a. Geef $E(X)$ en bereken $P(X > E(X))$
b. Bepaal de mediaan, dus de waarde M zodat $P(X \geq M) = P(X \leq M) = \frac{1}{2}$
c. Leid de kansdichtheid van W af (met de convolutie-integraal) en geef $E(W)$.

Normering:
Toetscijfer = $1 + 9 \times \frac{\text{aantal punten}}{36}$

1		**2**		**3**			**4**				**5**				**6**			**Tot**
a	b	a	b	a	b	c	a	b	c	c	a	b	c	d	a	b	c	
2	2	2	2	1	2	2	2	2	2	3	2	2	1	2	2	2	3	**36**

Formuleblad Kansrekening voor INF en BIT t.b.v. toetsen in module 4

Verdeling	$\boldsymbol{E(X)}$	$\boldsymbol{var(X)}$
Geometrisch	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Hypergeometrisch	$n \cdot \frac{R}{N}$	$n \cdot \frac{R}{N} \cdot \frac{N-R}{N} \cdot \frac{N-n}{N-1}$
Poisson $\quad P(X = x) = \frac{e^{-\mu}\mu^x}{x!}, x = 0, 1, 2, ...$	μ	μ
Uniform op (a, b)	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Exponentieel	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Erlang $\quad f_X(x) = \frac{\lambda(\lambda x)^{n-1} e^{-\lambda x}}{(n-1)!}, x \geq 0$	$\frac{n}{\lambda}$	$\frac{n}{\lambda^2}$

$$var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} var(X_i) + \sum_{i \neq j}\sum cov(X_i, X_j)$$

Uitwerkingen "Kansrekening voor INF/BIT"- module 4 (201300180) d.d. 14-06-2016

Opgave 1

a. $P(A \cup H) = P(A) + P(H) - P(AH) = P(A) + P(H) - P(A) \cdot P(H)$, wegens de onafhankelijkheid
Dus $P(A \cup H) = 0.3 + 0.4 - 0.3 \cdot 0.4 = 0.58$.

b. Gevraagd: $P(H|A) = \frac{P(A \cap H)}{P(A)}$, waarin $P(A) = 0.3$ en $P(A \cap H) = P(A|H)P(H) = 0.5 \cdot 0.4 = 0.2$
Dus $P(H|A) = \frac{0.2}{0.3} = \frac{2}{3}$

Opgave 2

a. Voor de verdeling van X is alleen van belang om onderscheid te maken tussen Aziaten en niet-Aziaten. X is dus hypergeometrisch verdeeld, omdat het trekkingen zonder terugleggen betreft (2 beurzen per persoon moet uitgesloten worden)

$P(X = 2) = \frac{\binom{10}{2}\binom{15}{4}}{\binom{25}{6}} \approx 34.7\%$.

$E(X) = n \cdot \frac{R}{N} = 6 \cdot \frac{10}{25} = 2.4$

b. Schematisch:

Azië	Afrika	Z.Amerika	Totaal
10	5	10	25
↓	↓	↓	↓
2	2	2	6

$P(X = 2 \text{ en } Y = 5) = \frac{\binom{10}{2}\binom{5}{2}\binom{10}{2}}{\binom{25}{6}} \approx 11.4\%$

Opgave 3

a. $P(X > Y) = P(X = 2 \text{ en } Y = 0) + P(X = 4 \text{ en } Y + 0) + P(X = 4 \text{ en } Y = 2)$
$= 0.12 + 0.09 + 0.12 = 33\%$

x	0	2	4	Totaal
$P(X = x)$	0.3	0.4	0.3	1
$x^2 P(X = x)$	0	1.6	4.8	6.4=$E(X^2)$

b. Rijgewijs optellen geeft de verdeling van X, zie tabel:
$E(X) = 2$ volgt meteen uit symmetrie van de verdeling
$var(X) = E(X^2) - (EX)^2 = 6.4 - 2^2 = 2.4$
(de waarde van $E(X^2)$ is in de laatste rij van de tabel bepaald)

c. X en Y zijn o.o. omdat steeds geldt $P(X = x \text{ en } Y = y) = P(X = x) \cdot P(Y = y)$,
Bijvoorbeeld als $x = 0$ en $y = 0$: $0.09 = P(X = 0 \text{ en } Y = 0) = P(X = 0) \cdot P(Y = 0) = 0.3 \cdot 0.3$

Opgave 4

a. $P(X \leq 2) = \int_0^2 \frac{1}{8} x dx = \left[\frac{1}{16} x^2\right]_{x=0}^{x=2} = \frac{1}{4}$
(of meetkundig uit de grafiek van f : opp. driehoek $= \frac{1}{2} \cdot 2 \cdot \frac{1}{4} = \frac{1}{4}$)

b. $E(X) = \int_{-\infty}^{\infty} x f(x) dx = \int_0^4 x \cdot \frac{1}{8} x dx = \left[\frac{1}{8} \cdot \frac{1}{3} x^3\right]_{x=0}^{x=4} = \frac{8}{3}$

$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x) dx = \int_0^4 x^2 \cdot \frac{1}{8} x dx = \left[\frac{1}{8} \cdot \frac{1}{4} x^4\right]_{x=0}^{x=4} = 8$

$var(X) = E(X^2) - (EX)^2 = 8 - \left(\frac{8}{3}\right)^2 = \frac{8}{9}$

c. Als $Y = \sqrt{X}$ en $0 \leq X \leq 4$, dan geldt $0 \leq Y \leq \sqrt{4}$. Dus voor $0 \leq y \leq 2$ geldt:
$F_Y(y) = P(Y \leq y) = P(\sqrt{X} \leq y) = P(X \leq y^2) = F_X(y^2)$.
Dus $f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy}[F_X(y^2)] = 2y \cdot f_X(y^2)$.

Omdat $f_X(x) = \frac{1}{8}x$ voor $0 \le x \le 4$, is $f_X(y^2) = \frac{1}{8}y^2$ voor $0 \le y \le 2$,
dus $f_Y(y) = 2y \cdot \frac{1}{8}y^2 = \frac{1}{4}y^3$, voor $0 \le y \le 2$ (*ga na dat* $\int_{-\infty}^{\infty} f_y(y)\, dy = \int_0^2 \frac{1}{4}y^3 dy = 1$)

d. Volgens de CLS geldt bij benadering: $\sum_{i=1}^{72} X_i \sim N(n\mu, n\sigma^2) = N\left(72 \cdot \frac{8}{3}, 72 \cdot \frac{8}{9}\right) = N(190{,}64)$
$P(\sum_{i=1}^{72} X_i \le 200) \approx P\left(Z \le \frac{200-192}{8}\right) = \Phi(1.00) = 84.13\%.$

Opgave 5

a. $\sigma = 3$ dus $\sigma^2 = 9$: X is $N(35{,}9)$, dus $P(X > 40) = P\left(Z > \frac{40-35}{3}\right) \approx 1 - \Phi(1.67) = 4.75\%.$

b. $P(X > 40 | X > 35) = \frac{P(X>40 \text{ en } X>35)}{P(X>35)} = \frac{P(X>40)}{P(X>35)} = \frac{0.0475}{0.5} = 9.5\%.$
opmerking: de normale verdeling is ***niet geheugenloos****, dus* $P(X > 40|X > 35) \ne P(X > 5)$

c. $\rho(X,Y) = +1$, omdat er een positief lineair verband tussen X en Y bestaat: $Y = 0.4X - 2$

d.

$$\rho(X,Y) = \rho(X, 0.4X - 2) = \frac{cov(X,\ 0.4X - 2)}{\sigma_X \cdot \sigma_{0.4X-2}} = \frac{0.4cov(X,X)}{\sigma_X \cdot 0.4\sigma_X} = \frac{0.4var(X)}{0.4\sigma_X^2} = 1$$

Opgave 6

a. $E(X) = \frac{1}{\lambda} = 3$ dus $P\big(X > E(X)\big) = \int_3^{\infty} \lambda e^{-\lambda x} dx = \left[-e^{-\lambda x}\right]_{x=3}^{x\to\infty} = 0 - \left(-e^{-\frac{1}{3}\cdot 3}\right) = e^{-1} \approx 36.8\%$

b. $P(X \ge M) = e^{-\frac{1}{3}\cdot M} = \frac{1}{2}$, als $-\frac{1}{3} \cdot M = ln\left(\frac{1}{2}\right)$, dus als $M = -3 \cdot \ln\left(\frac{1}{2}\right) \approx 2.08$

c. Convolutie-integraal toepassen:

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)dx = \int_0^z \frac{1}{3}e^{-\frac{1}{3}x} \cdot \frac{1}{3}e^{-\frac{1}{3}(z-x)}dx = \int_0^z \frac{1}{9}e^{-\frac{1}{3}z}dx = \left[\frac{1}{9}e^{-\frac{1}{3}z} \cdot x\right]_{x=0}^{x=z}$$

$= \frac{1}{9}e^{-\frac{1}{3}z} \cdot z$, als $z \ge 0$ (en $f_{X+Y}(z) = 0$, als $z < 0$)

$E(W) = E(X+Y) = E(X) + E(Y) = 3 + 3 = 6.$

($E(W)$ *kun je ook met de kansdichtheid* $f_{X+Y}(z)$ *uitrekenen (2 maal partieel integreren) maar deze oplossing is natuurlijk veel eenvoudiger*)