Toets Kansrekening voor INF en BIT
Module 4 (Data en Informatie - 201300180) - Maandag 30 juni 2014, 9.45-11.45 uur

Deze toets bestaat uit 5 opgaven. Het formuleblad en de tabel van de binomiale en standaardnormale verdeling zijn toegevoegd. Een gewone, niet programmeerbare rekenmachine is toegestaan (geen GR).

1. Bereken $P(\overline{A}|B)$ als bekend is dat $P(A \cup B) = 0.8$, $P(A \cap B) = 0.3$ en $P(A) = 0.35$.

2. (Deze opgave is gebaseerd op een rechtszaak in de V.S.)
De politie is bij een inval 496 pakketjes op het spoor gekomen die mogelijkerwijs cocaïne bevatten. Het kan echter ook zijn dat er onschuldige pakketjes met wit poeder bij zitten. Er worden aselect 4 pakketjes uit de partij gehaald en deze worden in het laboratorium onderzocht. Veronderstel dat van de 496 pakketjes er 331 cocaïne bevatten en dus 165 een onschuldig wit poeder.

a. Bereken, uitgaande van bovengenoemde veronderstelling, in drie decimalen nauwkeurig de kans dat alle 4 pakketjes die in het laboratorium onderzocht worden, cocaïne bevatten.

De 4 pakketjes blijken alle 4 cocaïne te bevatten. Van de 492 overige pakketjes besluit de politie er twee op de markt te brengen via undercover agenten. De kopers worden gearresteerd, maar hebben kans gezien het "bewijsmateriaal" te doen verdwijnen (door de wc gespoeld o.i.d.). De politie beweert dat het wel zeer waarschijnlijk is dat ook deze twee pakketjes cocaïne hebben bevat.

b. Bereken in drie decimalen nauwkeurig de kans dat deze twee volgende pakketjes géén cocaïne bevatten.

Veronderstel nu dat van de 496 pakketjes er c pakketjes cocaïne bevatten (met $c \geq 4$) en dus $496 - c$ een onschuldig wit poeder.

c. Geef een uitdrukking voor de kans (als functie van c) dat de eerste 4 pakketjes cocaïne bevatten en de twee volgende niet. (Deze kans blijkt maximaal te zijn voor $c = 331$; dat is dus de gunstigste situatie voor de gearresteerde kopers.)

3. Sommige economen pleiten voor een drastische vereenvoudiging van het belastingstelsel, om de economie te stimuleren en de administratieve rompslomp sterk terug te dringen: een "*flat tax*" van 30%, één belastingvrije voet van 10000 Euro en geen aftrekposten meer.
Zij X het belastbaar inkomen van een willekeurige werknemer in duizenden Euro's ($X >$ 10 duizend) en $f_X(x)$ de kansdichtheid, waarmee deze gemodelleerd wordt.
Dan is $\boldsymbol{Y} = 0.3(X - 10) = \boldsymbol{0.3X - 3}$ de te betalen belasting van die werknemer in dit voorstel.

a. Als $E(X) = 30$ en $\sigma_X = 10$, wat is dan de verwachte belastingaanslag $E(Y)$ en σ_Y?
b. Leid af hoe je de kansdichtheid $f_Y(y)$ van Y kan uitdrukken in de kansdichtheid van X.
c. Geef de definitie van de correlatiecoëfficiënt $\rho(X, Y)$ en toon aan dat in dit voorbeeld $\rho(X, Y) = 1$

4. De programmeerafdeling van een software leverancier schat, op grond van ervaringen met een bepaald groot softwarepakket, dat er gemiddeld ééns in de twee weken een bug gemeld zal worden: de verwachte tussenaankomsttijd tussen twee opeenvolgende bugs bedraagt dus 2 weken.
Verondersteld wordt dat de tussenaankomsttijden $X_1, X_2, \ldots\ldots$ (in weken) van de bugs vanaf het uitbrengen van de software onderling onafhankelijk en exponentieel verdeeld zijn. X_1 is dus de tijd (in weken) die verstrijkt tussen het uitbrengen van de software en de melding van de eerste bug.
De n-de bug komt binnen na $S_n = \sum_{i=1}^{n} X_i$ weken.

a. Bereken $P(X_1 > 2)$
b. Bepaal $E(S_n)$ en $var(S_n)$

c. Bereken (of benader) $P(S_{30} > 52)$, de kans dat er binnen 52 weken minder dan 30 bugs binnenkomen.

d. Leid m.b.v. de convolutie-integraal de kansdichtheid van $X_1 + X_2$ af.

5. Bij de fabricage van een product treedt met een kans 0.1 een defect op. De defecte producten worden opnieuw bekeken en met een kans van 0.4 lukt het om een defect product weer in orde te krijgen. We bekijken 100 producten die op deze wijze behandeld worden.
Zij X het aantal producten (onder deze 100), die meteen goed zijn. Het aantal defecte producten die vervolgens opnieuw bekeken worden, is dus $100 - X$. Zij Y het aantal producten van deze $100 - X$ die in tweede instantie in orde komen. Is dus bijvoorbeeld $X = 92$, dan zijn er in eerste instantie 8 defecte producten. Deze worden opnieuw bekeken en Y is het aantal onder deze 8 waarvan het lukt om ze weer in orde te krijgen.

a. Geef de kansverdeling (type + parameters) van $X, E(X)$ en $var(X)$.

b. Bereken $P(Y = 3|X = 92)$.

c. Bepaal $P(X + Y = 95)$.
(Volstaan kan worden met een sommatie waarvan de termen bestaan uit zo ver mogelijk uitgewerkte numerieke formules).

Normering: toetscijfer = $1 + \frac{aantal\ punten}{30} \times 9$

1	2			3			4				5			Tot
	a	b	c	a	b	c	a	b	c	d	a	b	c	
3	2	2	2	2	2	2	2	2	3	2	2	2	2	30

Formuleblad Kansrekening voor INF en BIT t.b.v. toetsen

Verdeling	$E(X)$	$var(X)$
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 $P(X = x) = \frac{e^{-\mu}\mu^x}{x!}, x = 0, 1, 2, \ldots$	μ	μ
Exponentieel	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Uniform op (a, b)	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$

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

Uitwerkingen:

Opgave 1

$P(A\cup B)=P(A)+P(B)-P(A\cap B)$, waarin $P(A\cup B)=0.8, P(A\cap B)=0.3$ en $P(A)=0.35$, dus $P(B)=0.8-0.35+0.3=0.75$

$P(\overline{A}|B)=\frac{P(\overline{A}\cap B)}{P(B)}=\frac{P(B)-P(AB)}{P(B)}=\frac{0.75-0.3}{0.75}=0.6$

(Of gebruik de complement regel: $P(\overline{A}|B)=1-P(A|B)=1-\frac{P(AB)}{P(B)}=1-\frac{0.3}{0.75}=0.6$)

Opgave 2

a. Gebruikmakned van de hypergeometrische formule is de gevraagde kans: $\frac{\binom{331}{4}\binom{165}{0}}{\binom{496}{4}}\approx 0.197$

b. Deze kans is gelijk aan : $\frac{\binom{327}{0}\binom{165}{2}}{\binom{492}{2}}\approx 0.112$

c. De kans is $\frac{\binom{c}{4}\binom{496-c}{0}}{\binom{496}{4}}\times\frac{\binom{c-4}{0}\binom{496-c}{2}}{\binom{492}{2}}\quad\left(=\frac{c(c-1)(c-2)(c-3)(496-c)(495-c)}{496\times495\times494\times493\times492\times491}\right)$

Opgave 3

a. $Y=0.3X-3$, dus $E(Y)=E(0.3X-3)=0.3E(X)-3=0.3\cdot30-3=6$ (1000 Euro)
$var(Y)=var(0.3X-3)=0.3^2var(X)=0.3^2\cdot10^2=9$ dus $\sigma_Y=3$

b. $F_Y(y)=P(Y\le y)=P(0.3X-3\le y)=P\left(X\le\frac{y+3}{0.3}\right)=F_X\left(\frac{y+3}{0.3}\right)$,
dus $f_Y(y)=\frac{d}{dy}F_Y(y)=\frac{1}{0.3}f_X\left(\frac{y+3}{0.3}\right)$

c. $\rho(X,Y)=\frac{cov(X,Y)}{\sigma_X\sigma_Y}$, dus
$\rho(X,Y)=\rho(X,0.3X-3)=\frac{cov(X,0.3X-3)}{\sigma_X\sigma_{0.3X-3}}=\frac{0.3\cdot cov(X,X)}{\sigma_X\cdot0.3\sigma_X}=\frac{0.3\sigma_X^2}{0.3\sigma_X^2}=1$

Opgave 4

a. $E(X_1)=2=\frac{1}{\lambda}$, dus $\lambda=\frac{1}{2}$. $P(X>2)=\int_2^\infty\frac{1}{2}e^{-0.5x}dx=-e^{-0.5x}|_2^{x\to\infty}=0-(-e^{-1})\approx36.8\%$

b. $E(S_n)=nE(X_1)=n\cdot2=2n$ en $var(S_n)=n\cdot var(X_1)=n\cdot\frac{1}{0.5^2}=4n$

c. $S_{30}=\sum_{i=1}^{30}X_i$ is volgens de CLS bij benadering $N(30\cdot2,\ 30\cdot4)=N(60,120)$-verdeeld
$P(S_{30}>52)\overset{\text{CLS}}{\approx}P\left(Z\ge\frac{52-60}{\sqrt{120}}\right)\approx P(Z\ge-0.73)=\Phi(0.73)=0.7673$

d. $f_{X_1+X_2}(z)=\int_{-\infty}^\infty f_{X_1}(x)f_{X_2}(z-x)dx=\int_0^z0.5e^{-0.5x}\cdot0.5e^{-0.5(z-x)}dx=\int_0^z0.5^2e^{-0.5z}\,dx$
$=0.25e^{-0.5z}\cdot x|_{x=0}^{x=z}=0.25xe^{-0.5z}$, voor $z\ge0$

Opgave 5

a. $X\sim B(100,0.9)$, dus: $E(X)=np=90$ en $var(X)=np(1-p)=9$

b. Gegeven $X=92$ is $Y\sim B(8,0.4)$ −verdeeld.
Dus (m.b.v. binomiale tabel): $P(Y=3|X=92)=0.5941-0.3154\approx27.9\%$
of ook (rechtstreeks): $P(Y=3|X=92)=\binom{8}{3}0.4^30.6^5\approx27.9\%$

c. $P(X+Y=95)=\sum_{x=0}^{95}P(X=x\text{ en }Y=95-x)=\sum_{x=0}^{95}P(X=x)P(Y=95-x\,|\,X=x)$

$$=\sum_{x=0}^{95}\binom{100}{x}0.9^x0.1^{100-x}\binom{100-x}{95-x}0.4^{95-x}0.6^5.$$

(ook met de trinomiale kansfunctie te geven)