Toets Kansrekening voor INF en BIT
Module 4 (Data en Informatie - 201300180) - maandag 29 juni 2015.

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. Bij een productielijn voor laptops wordt een geautomatiseerde routine-eindtest uitgevoerd, om te controleren of de laptop voldoet aan de kwaliteitseisen. Uit de statistieken van de fabriek blijkt dat 97% van de geproduceerde laptops worden goedgekeurd. Bij een evaluatie van de eindtest is gebleken dat van de goedgekeurde laptops 99% daadwerkelijk voldoen aan de kwaliteitseisen en dat van de afgekeurde laptops 15% toch aan de kwaliteitseisen voldoen.
Bereken de volgende kansen door relevante gebeurtenissen te definiëren, de gegeven en gevraagde kansen daarin uit te drukken en de rekenregels toe te passen:

a. de kans dat een laptop voldoet aan de eisen en

b. de kans dat een laptop die aan de eisen voldoet wordt goedgekeurd.

2. Uit een bak met 10 balletjes, genummerd 1 t/m 10, worden er lukraak en zonder terugleggen drie getrokken. Laat X het laagste en Y het hoogste nummer zijn op de drie getrokken balletjes.
Bereken:

a. $P(Y = 5)$

b. $P(X = 2|Y = 5)$

c. $E(X|Y = 5)$

3. Een (pseudo) random generator geeft ons een willekeurig getal X tussen 0 en 1.

a. Laat zien dat het k-de moment $E(X^k) = \frac{1}{k+1}$ (voor $k = 1,2,3, ...$) is.

b. Gebruik a. om de variantie van X af te leiden.

c. Toon aan dat $Y = -3\ ln(X)$ exponentieel verdeeld is en bepaal $E(Y)$.

4. De bedieningstijden van 25 klanten bij een loket zijn ieder exponentieel verdeeld met een verwachting van 2 minuten. Beschouw de bedieningstijden als onderling onafhankelijke stochastische variabelen $X_1, X_2, \dots, X_{25}$.

a. Bepaal de verwachting en de variantie van de totale bedieningstijd $\sum_{i=1}^{25} X_i$.

b. Bereken m.b.v. de Centrale Limiet Stelling een benadering voor de kans dat de totale bedieningstijd meer is dan een uur. Geef uw antwoord in tienden van procenten nauwkeurig.

c. Bereken de correlatiecoëfficiënt van de bedieningstijd van klant 1 (X_1) en de totale bedieningstijd $\sum_{i=1}^{25} X_i$.

5. Uit de populatie van 20-jarige mannen, van wie de gewichten $N(80, 100)$-verdeeld zijn, worden twee mannen willekeurig gekozen: hun gewichten noemen we X en Y.

a. Bereken $P(X > 90 \text{ en } Y > 90)$

b. Bereken $P(X + Y > 180)$

6. Een internet provider wil onderzoeken wat het effect is van een mogelijke wijziging in hun tariefsysteem. Daartoe wordt voor een groep van 54 personen, willekeurig gekozen uit het klantenbestand van de grootste concurrent, bepaald hoeveel van deze 54 personen bij introductie van het nieuwe tariefsysteem klant zouden worden. Zij X het aantal toekomstige klanten bij het nieuwe tariefsysteem onder de groep van 54 willekeurig gekozen personen.
Voor een positief besluit over invoering van het nieuwe tariefsysteem is nodig dat minstens de helft van de klanten van de concurrent willen overstappen. De vraag is of dit werkelijk het geval is, als minstens de helft van de mensen in de steekproef te kennen geeft te zullen overstappen.
Om een antwoord op die vraag te geven gaan we er bij onderstaande vragen vanuit dat **in werkelijkheid slechts 40%** van alle klanten van de concurrent zou overstappen.

a. Geef de kansverdeling van X (type +parameters) en bepaal $E(X)$ en $var(X)$.
b. Bereken (of benader) de kans $P(X \geq 27)$, de kans dat minstens de helft van de personen in de steekproef zou overstappen.

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

1		2			3			4			5		6		Totaal
a	b	a	b	c	a	b	c	a	b	c	a	b	a	b	
3	2	2	2	2	3	2	3	2	2	2	2	2	3	4	**36**

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

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
Uniform op (a, b)	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Exponentieel	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Erlang $f_X(x) = \frac{\lambda(\lambda x)^{n-1}e^{-\lambda x}}{(n-1)!}, x \geq 0$	$\frac{n}{\lambda}$	$\frac{n}{\lambda^2}$

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

Uitwerkingen Herkansing "Kansrekening voor INF/BIT"- module 4 (201300180) dd 29-06-2015

Opgave 1

a. Gebeurtenissen: V treedt op als de laptop "Voldoet aan eisen" en G als de laptop "Goedgekeurd"

$P(G) = 0.97$, $P(V|G) = 0.99$ en $P\left(V\middle|\overline{G}\right) = 0.15$

$P(V) = P(VG) + P(V\,\overline{G}) = P(V|G)P(G) + P\left(V\middle|\overline{G}\right)P(\overline{G}) = 0.99 \cdot 0.97 + 0.15 \cdot 0.03 = 0.9648$

b. $P(G|V) = \frac{P(GV)}{P(V)} = \frac{0.9603}{0.9648} \approx 0.9953$ (of eerst de regel van Bayes geven)

Opgave 2

a. $P(Y=5) = \frac{\binom{4}{2}}{\binom{10}{3}} = \frac{6}{120} = 0.05$ Schematisch:

nr. 1-4	nr. 5	nr. 6-10	Totaal
4	1	5	10
↓	↓	↓	↓
2	1	0	3

(of: $P(Y=5) = \frac{1}{10} \times \frac{4}{9} \times \frac{3}{8} \times 3$)

b. $P(X=2|Y=5) = \frac{P(X=2 \text{ en } Y=5)}{P(Y=5)} = \frac{\frac{2}{120}}{\frac{6}{120}} = \frac{1}{3}$

(als $Y = 5$ zijn er slechts $\binom{4}{2} = 6$ mogelijkheden: (1,2,5), (1,3,5), (1,4,5), (2,3,5), (2,4,5) en (3,4,5). Twee van die 6 mogelijkheden hebben 2 als laagste nummer: dan is $X = 2$ en $Y = 5$)

c. Analoog b vinden we: $P(X=1|Y=5) = \frac{3}{6}$ en $P(X=3|Y=5) = \frac{1}{6}$

Dus $E(X|Y=5) = \sum_x xP(X=x|Y=5) = 1 \times \frac{3}{6} + 2 \times \frac{2}{6} + 3\times \frac{1}{6} = \frac{5}{3}$

Opgave 3

a. $E(X^k) = \int_{-\infty}^{\infty} x^k f(x)dx = \int_0^1 x^k \cdot 1dx = \left[\frac{1}{k+1}x^{k+1}\right]_{x=0}^{x=1} = \frac{1}{k+1}$, voor $k = 1,2, \ldots$,

(dus $E(X) = \frac{1}{2}$ en $E(X^2) = \frac{1}{3}$.)

b. $var(X) = E(X^2) - (EX)^2 = \frac{1}{2+1} - \left(\frac{1}{2}\right)^2 = \frac{1}{12}$

c. $F_Y(y) = P(Y \le y) = P(-3\ln(X) \le y) = P\left(X \ge e^{-\frac{1}{3}y}\right) = 1 - P\left(X \le e^{-\frac{1}{3}y}\right) = 1 - F_X\left(e^{-\frac{1}{3}y}\right)$

$f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{d}{dy}\left(1 - F_X\left(e^{-\frac{1}{3}y}\right)\right) = +\frac{1}{3}e^{-\frac{1}{3}y}f_X\left(e^{-\frac{1}{3}y}\right)$

Omdat $f_X(x) = 1$ als $0 \le x \le 1$, is $f_X\left(e^{-\frac{1}{3}y}\right) = 1$ als $y \ge 0$, dus:

$f_Y(y) = \frac{1}{3}e^{-\frac{1}{3}y}$, voor $y \ge 0$, dus Y is exponentieel verdeeld met $\lambda = \frac{1}{3}$

Dus $E(Y) = \frac{1}{\lambda} = 3$

Opgave 4

a. Omdat $E(X_i) = 2 = \frac{1}{\lambda}$ en $var(X_i) = \frac{1}{\lambda^2} = 4$, geldt: $E\left(\sum_{i=1}^{25} X_i\right) = \sum_{i=1}^{25} E(X_i) = 25 \times 2 = 50$

en $var\left(\sum_{i=1}^{25} X_i\right) = \sum_{i=1}^{25} var(X_i) = 25 \times 4 = 100$.

b. Volgens de Centrale Limiet Stelling geldt dat $\sum_{i=1}^{25} X_i$ bij benadering N(25×2, 25×4)-verdeeld is:

$P\left(\sum_{i=1}^{25} X_i > 60\right) = P\left(\frac{\sum_{i=1}^{25} X_i - 50}{\sqrt{100}} > \frac{60-50}{\sqrt{100}}\right) \approx 1 - \Phi(1) = 1 - 0.8413 = 15.87\%$

c. $\varrho\left(X_1, \sum_{i=1}^{25} X_i\right) = \frac{cov\left(X_1,\ \sum_{i=1}^{25} X_i\right)}{\sqrt{varX_1}\sqrt{\sum_{i=1}^{25} X_i}}$

Er geldt: $cov\left(X_1, \sum_{i=1}^{25} X_i\right) = cov(X_1, X_1) + cov(X_1, X_2) + \cdots + cov(X_1, X_{25})$

$= var(X_1) + 0 + \cdots + 0 = 4$

en (zie a.) $var(X_1) = 4$ resp. $var\left(\sum_{i=1}^{25} X_i\right) = 100$.
Dus $\varrho\left(X_1, \sum_{i=1}^{25} X_i\right) = \frac{4}{\sqrt{4}\sqrt{100}} = \frac{1}{5}$

Opgave 5

a. $P(X > 90 \text{ en } Y > 90) \overset{\text{o.o.}}{=} P(X > 90) \cdot P(Y > 90) = \left(1 - \Phi\left(\frac{90-80}{10}\right)\right)^2 \approx 2.52\%$

b. $X + Y$ is N(80+80, 100+100)-verdeeld, dus
$P(X + Y > 180) = P\left(Z > \frac{180-160}{\sqrt{200}}\right) \approx 1 - \Phi(1.41) = 7.93\%$

Opmerking: bij opgave 5 en 4b géén continuïteitscorrectie toepassen: de variabelen zijn al continu!

Opgave 6

a. Zij X het aantal toekomstige klanten bij het nieuwe tariefsysteem onder de groep van n = 54 willekeurig gekozen klanten van de concurrent: $X \sim B(54, 0.40)$.
$E(X) = np = 54 \cdot 0.4 = 21.6$ en $var(X) = np(1-p) = 54 \cdot 0.4 \cdot 0.6 = 12.96$

b. Omdat $n > 25$, en ook $np = 21.6 > 5$ en $n(1-p) = 32.4 > 5$, kunnen we X normaal benaderen met $\mu = 21.6$ en $\sigma = \sqrt{var(X)} = \sqrt{12.96} = 3.6$
$P(X \geq 27) \overset{\text{cont. corr.}}{=} P(X \geq 26.5) \approx P\left(Z \geq \frac{26.5-21.6}{3.6}\right) \approx 1 - \Phi(1.36) = 1 - 0.9131 = 8.69\%$