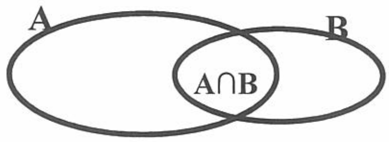
Uitwerking Tentamen Kansrekening en Statistiek (191530082) voor INF, donderdag 10 april 2014



Opgave 1

a. Als A en B disjunct zijn, dan geldt: $P(A \cap B) = 0$,
dus $P(A \cup B) = P(A) + P(B) = 0.2 + 0.5 = \mathbf{0.7}$

b. Als A en B onafhankelijk zijn, dan is $P(A \cap B) \overset{\text{o.o.}}{=} P(A) \times P(B) = 0.2 \times 0.5 = 0.10$
Dus: $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.2 + 0.5 - 0.1 = \mathbf{0.60}$

c. $P(A|B) = \frac{P(A \cap B)}{P(B)} = 0.3$,
ofwel: $P(A \cap B) = 0.3 \times P(B) = 0.3 \times 0.5 = 0.15$.
Dus: $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.2 + 0.5 - 0.15 = \mathbf{0.55}$

Opgave 2

a. X is binomiaal verdeeld met $n = 10$ en $p = 0.2$: $P(X > 0) = 1 - P(X = 0) = 1 - 0.8^{10} = 89.3\%$

b. voor $n = 250$ en $p = 0.01$ kunnen we bij benadering de Poisson verdeling gebruiken met $\lambda = np = 2.5$
Dus $P(X \geq 5) = 1 - P(X \leq 4) = 1 - 0.890 = 11.0\%$

c. voor $n = 400$ en $p = 0.1$ is X bij benadering N(np, $np(1\text{-}p)$)- dus N(40, 36)-verdeeld
$P(X < 30) = P(X \leq 29.5)$ (*continuïteitscorrectie*)
$= P\left(\frac{X-40}{\sqrt{36}} \leq \frac{29.5-40}{6}\right) \approx P(Z \leq -1.75) = \frac{1}{2} - P(0 \leq Z \leq 1.75) = \frac{1}{2} - 0.4599 = 4.01\%$

Opgave 3

a. $P(X = 0) = P(X = 2) = 0.3$ en $P(X = 1) = 0.4$
Dus $E(X) = 1$ wegens symmetrie
En $var(X) = E(X^2) - EX^2 = [0^2 \cdot 0.3 + 1^2 \cdot 0.4 + 2^2 \cdot 0.3] - 1^2 = 0.6$

b. $\rho(X, Y) = \frac{E(XY) - EX \cdot EY}{\sigma_X \sigma_Y}$
Hierin zijn $EX = EY = 1$ en $\sigma_X = \sigma_Y = \sqrt{0.6}$, omdat X en Y dezelfde verdeling hebben (zie a).
$E(XY) = \sum\sum x \cdot y \cdot P(X = x \text{ en } Y = y) = [1 \cdot 1 + 1 \cdot 2 + 2 \cdot 1 + 2 \cdot 2] \cdot 0.15 = 1.35$
Dus $\rho(X, Y) = \frac{E(XY) - EX \cdot EY}{\sigma_X \sigma_Y} = \frac{1.35 - 1 \cdot 1}{0.6} = \frac{7}{12}$ (≈ 0.58)
Er is een matige positieve correlatie tussen het aantal zieken in de ochtend en de avondploeg.

c. $E(X + Y) = E(X) + E(Y) = 2$ en
$var(X + Y) = var(X) + varY) + 2cov(X, Y) = 0.6 + 0.6 + 2 \cdot 0.35 = 1.9$

d. $P(Y = 0|X = 0) = \frac{P(Y=0 \text{ en } X=0)}{P(Y=0)} = \frac{0.2}{0.3} = \frac{2}{3}$ en evenzo: $P(Y = 1|X = 0) = \frac{1}{3}$
$E(Y|X = 0) = \sum y P(Y = y|X = 0) = 0 \cdot \frac{2}{3} + 1 \cdot \frac{1}{3} = \frac{1}{3}$.

Opgave 4

a. $E(X) = \frac{1}{\lambda} = 2$, dus $\lambda = 0.5$
$P(X_1 > 3) = \int_3^\infty 0.5e^{-0.5x} dx = -e^{-0.5x}\big|_{x=3}^{\infty} = e^{-1.5} \approx 22.3\%$ en
$P(X_1 > 5 \mid X_1 > 2) = P(X_1 > 3) = 22.3\%$ wegens geheugenloosheid.

b. $P(X_1 > 3 \text{ en } X_2 > 3) \overset{\text{o.o.}}{=} P(X_1 > 3) \cdot P(X_2 > 3) = 0.2231^2 \approx 5.0\%$

c. $P(X_1 + X_2 > 3) = 1 - P(X_1 + X_2 \leq 3) = 1 - \int_0^3 \int_0^{3-y} 0.5e^{-0.5x}\, 0.5e^{-0.5y} dx dy$
$= 1 - \int_0^3 0.5e^{-0.5y} \cdot \left[-e^{-0.5x}\big|_{x=0}^{3-y}\right] dy$
$= 1 - \int_0^3 0.5e^{-0.5y} \cdot [e^{-1.5+0.5y} - 1] dy$

$= 1 - [-e^{-0.5y} - 0.5ye^{-1.5}]_{y=0}^{3} = 1 - [-2.5e^{-1.5} + 1] = 2.5e^{-1.5} \approx \mathbf{55.8\%}$

(of gebruik dat $X_1 + X_2$ Erlang verdeeld is met $n = 2$ en $\lambda = 0.5$ (zie formuleblad)

$P(X_1 + X_2 > 3) = \int_3^\infty 0.25xe^{-0.5x}dx =$...partiële integratie...= 55.8%)

d. $X_1 + X_2 + \cdots + X_{36}$ is volgens de CLS bij benadering $N\left(36 \cdot 2,\ 36 \cdot \frac{1}{0.5^2}\right)$ – verdeeld

$P(X_1 + X_2 + \cdots + X_{36} > 90) \approx P\left(Z > \frac{90-72}{\sqrt{144}}\right) = 1 - \Phi(1.5) =$ 1- 0.9332 ≈ 6.7%

e. $F_X(x) = P(-2\ln(U) \leq x) = P(U \geq e^{-0.5x}) = 1 - F_U(e^{-0.5x})$

$f_X(x) = \frac{d}{dx}F_X(x) = 0.5e^{-0.5x}f_U(e^{-0.5x})$

$0 < e^{-0.5x} < 1$ als $\boldsymbol{x} > \mathbf{0}$: dan is$f_U(e^{-0.5x}) = 1$ en dus $f_X(x) = 0.5e^{-0.5x}$ (anders is $f_X(x) = 0$)

X is dus exponentieel verdeeld met parameter $\lambda = 0.5$.

Opgave 5

a. $E(c\overline{X}) = cE(\overline{X}) = c\mu$ en

$var(c\overline{X}) = c^2 \cdot var(\overline{X}) = c^2 \cdot \frac{\mu^2}{10}$

b. $T = c\overline{X}$ is een zuivere schatter van μ als $E(T) = \mu$, dus als $E(c\overline{X}) = cE(\overline{X}) = c \cdot \mu = \mu \to c = 1$

c. De verwachte kwadratische fout is $E(T - \mu)^2 = (ET - \mu)^2 + var(T) = (c - 1)^2\mu^2 + c^2 \cdot \frac{\mu^2}{10}$

Daar μ > 0 is deze minimaal als $g(c) = (c - 1)^2 + \frac{c^2}{10}$ minimaal is:

$g`(c) = 2(c - 1) + \frac{2c}{10} = 0$ als $c = \frac{10}{11}$.

Dit is een minimum, omdat g onbegrensd en positief is (of omdat $g``\left(\frac{10}{11}\right) = 2.3 > 0$),

dus $T = \frac{10}{11}\overline{X}$ is de beste schatter.

(Merk op dat het blijkbaar beter is een (kleine) onderschatting van μ toe te staan om zo ook de variantie van de schatter te laten afnemen)

Opgave 6

a. We gaan na of het berekende gemiddelde lengte van de cyclus significant afwijk van 29.5

1. De lengtes $X_1, X_2, ..., X_{15}$ zijn o.o. en $N(\mu, \sigma^2)$-verdeeld, onbekende μ en σ^2
2. We toetsen $H_0 : \mu = 29.5$ tegen $H_1 : \mu \neq 29.5$ met α = 5%
3. Toetsingsgrootheid: $T = \dfrac{\overline{X} - 29.5}{S/\sqrt{15}}$ is onder H_0 t_{14} – verdeeld
4. $\overline{x} = 27.667$ en $s = 2.743$, dus de uitkomst van T is $\dfrac{27.667 - 29.5}{2.743/\sqrt{15}} = \dfrac{-1.833}{0.7082} = -2.59$
5. We verwerpen H_0 als $T \geq c$ of $T \leq -c$
   Onbetrouwbaarheidsdrempel 5%, $c =$ 2.14 uit de t_{14}-tabel, zodat $P(T \leq c) = 0.975$
6. Uitkomst –2.59 ligt in het kritieke gebied $\Rightarrow H_0$ verwerpen
7. We achten bewezen, bij onbetrouwbaarheidsdrempel 5%, dat het (populatie)gemiddelde van de lengte van de menstruatiecyclus afwijkt van de lengte van een maanmaand.

b. 95%-BI(σ^2) = $\left(\frac{(n-1)s^2}{c_2}, \frac{(n-1)s^2}{c_1}\right)$,

met $s^2 = 2.743^2$ en $c_1 = 5.63$ en $c_2 = 26.1$ uit de χ^2 – tabel met $df = n - 1 = 14$ bij de staartkansen van 2.5%

Dus 95%-BI(σ^2) ≈ (4.0, 18.7)