STOCKHOLMS UNIVERSITET
MATEMATISKA INSTITUTIONEN
Avd. Matematisk statistik, MD

MT 3001
LÖSNINGAR
28 oktober 2020

Lösningar

28 oktober 2020

Uppgift 1

Rätt svar är

a) $A \cap B$

b) $a = 6$

c) $\text{NegBin}(r_1 + r_2, p)$ om $p_1 = p_2 = p$

d) $\mathrm{P}(X \geq 1) = 1/2$

e) Som det är formulerat kan påståendet vara både sant och falskt. Påståendet skulle ha varit: ”Om $A, B \subset \mathbb{R}^2$ och $|A| \leq |B|$ (där $|\cdot|$ betecknar area), så gäller *alltid* att $\mathrm{P}((X,Y) \in A) \leq \mathrm{P}((X,Y) \in B)$”, vilket hade varit falskt.

Uppgift 2

Låt S, N och I beteckna händelserna att personen är storrökare, normalrökare respektive icke-rökare, och D händelsen att personen avlider under året. Givet i uppgiften är att $\mathrm{P}(S) = 0.1$, $\mathrm{P}(N) = 0.25$, $\mathrm{P}(I) = 0.65$ och, om vi betecknar $\mathrm{P}(D|I) = p$, att $\mathrm{P}(D|S) = 6p$ och $\mathrm{P}(D|N) = 3p$. Enligt definitionen av betingad sannolikhet gäller att

$$\mathrm{P}(I|D) = \frac{\mathrm{P}(D|I)\mathrm{P}(I)}{\mathrm{P}(D)},$$

och lagen om total sannolikhet ger att

$$\begin{aligned}\mathrm{P}(D) &= \mathrm{P}(D|S)\mathrm{P}(S) + \mathrm{P}(D|N)\mathrm{P}(N) + \mathrm{P}(D|I)\mathrm{P}(I) \\ &= 6p \cdot 0.1 + 3p \cdot 0.25 + p \cdot 0.65 \\ &= 2p.\end{aligned}$$

Vi får alltså

$$\mathrm{P}(I|D) = \frac{p \cdot 0.65}{2p} = 0.325.$$

Uppgift 3

Låt X_i beteckna vikten hos passagerare i och Y_i vikten hos hens bagage. Definiera $V_i = X_i + Y_i$.

a) $\mathbb{E}[V_i] = \mathbb{E}[X_i] + \mathbb{E}[Y_i] = 75 + 15 = 90$
$\mathrm{Var}(V_i) \overset{ober.}{=} \mathrm{Var}(X_i) + \mathrm{Var}(Y_i) = 10^2 + 5^2 = 125$
Standardavvikelsen blir $\sqrt{125} = 11.2$

b) Eftersom den sammanlagda vikten V_{tot} är en summa av 100 st oberoende likafördelade stokastiska variabler kan vi använda oss av centrala gränsvärdessatsen (C.G.S.) för att beräkna sannolikheten att $V_{tot} > 9300$ enligt

$$P(V_{tot} > 9300) = P\left(\frac{V_{tot} - 100 \cdot 90}{11.2\sqrt{100}} > \frac{9300 - 100 \cdot 90}{11.2\sqrt{100}}\right) \overset{C.G.S.}{\approx} P\left(Z > 2.68\right)$$

där $Z \sim \mathrm{N}(0,\, 1)$, vilket tillsammans med symmetriargument ger att

$$P(V_{tot} > 9300) \approx 1 - \Phi(2.68) = 1 - 0.9963 = 0.0037$$

Uppgift 4

a) Låt T_i beteckna livslängden för komponent i $(i = 1, \ldots, 4)$. Vi har att $P(T_i \leq 2) = \int_2^\infty \frac{1}{x^2} dx = 0.5$ och den sökta sannolikheten blir alltså:

$$P(\cap_{i=1}^4 \{T_i \leq 2\}) \overset{ober}{=} \prod_{i=1}^{4} P(T_i \leq 2) = 0.5^4 = 0.0625.$$

b) Antalet komponenter X som fortfarande fungerar efter två år är $\mathrm{Bin}(4, 0.5)$-fördelat, och vi får alltså:

$$P(X \geq 3) = P(X = 3) + P(X = 4) = \binom{4}{3} 0.5^4 + \binom{4}{4} 0.5^4 = 0.3125.$$

Uppgift 5

a) Faltningsformeln ger, för $z \geq 0$, att

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)dx = \int_0^z e^{-x} e^{-(z-x)}dx = \int_0^z e^{-z}dx = ze^{-z}.$$

För $z < 0$ gäller att $f_Z(z) = 0$ (eftersom Z bara antar positiva värden).

b) Korrelationen mellan X och Z definieras som

$$\rho(X,Z) = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)\text{Var}(Z)}} = \frac{\mathbb{E}[XZ] - \mathbb{E}[X]\mathbb{E}[Z]}{\sqrt{\text{Var}(X)\text{Var}(Z)}}.$$

Eftersom X och Y är exponentialfördelade med parameter 1 så har vi att $\mathbb{E}[X] = \text{Var}(X) = 1$ och $\mathbb{E}[Z] = \mathbb{E}[X+Y] = 2$. Oberoende ger också att $\text{Var}(Z) = \text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y) = 2$. Väntevärdet av produkten ges av

$$\mathbb{E}[XZ] = \mathbb{E}[X(X+Y)] = \mathbb{E}[X^2] + \mathbb{E}[XY] = \mathbb{E}[X^2] + \mathbb{E}[X]^2 = \text{Var}(X) + 2\mathbb{E}[X]^2 = 3,$$

där den tredje olikheten motiveras av att X och Y är oberoende och lika fördelade. Sammantaget får vi att

$$\rho(X,Z) = \frac{3 - 1\cdot 2}{\sqrt{1\cdot 2}} = \frac{1}{\sqrt{2}}.$$

Uppgift 6

Låt X_1 beteckna antalet oanvända pennor vi får vid första dragningen, och X_2 antalet oanvända pennor vi får vid andra dragningen. Vi har att $X_1 \sim \text{HypGeo}(n, N, m)$ där $n = 2$ st dragna pennor bland totalt $N = 10$ pennor, av vilka $m = 6$ är oanvända vid första dragningen. Vad gäller den andra dragningen så beror den på utfallet av den första dragningen, men i övrigt så drar vi pennor enligt samma princip. Vi får att $X_2 \mid X_1 = x \sim \text{HypGeo}(2, 10, 6-x)$ där $x = 0, 1, 2$. Den sökta sannolikheten är $P(X_2 = 2)$ och denna får vi fram genom att använda lagen om total sannolikhet tillsammans med sannolikhetsfunktionen för en hypergeometrisk stokastisk variabel:

$$P(X_2 = 2) = \sum_{x=0}^{2} P(X_2 = 2 \mid X_1 = x)P(X_1 = x) = \sum_{x=0}^{2} \frac{\binom{6-x}{2}\binom{4+x}{0}}{\binom{10}{2}} \frac{\binom{6}{x}\binom{4}{2-x}}{\binom{10}{2}}$$

$$= 2\frac{\binom{6}{2}\binom{4}{2}}{\binom{10}{2}^2} + \frac{\binom{5}{2}\binom{6}{1}\binom{4}{1}}{\binom{10}{2}^2} \approx 0.2074.$$