STOCKHOLMS UNIVERSITET
MATEMATISKA INSTITUTIONEN
Avd. Matematisk statistik, MD

MT 3001
LÖSNINGAR
14 december 2022

Lösningar

14 december 2022

Uppgift 1

Rätt svar är

a) 1. $\mathbb{P}(A) > \mathbb{P}(B)$

b) 2. alltid falskt

c) 3. $\mathbb{P}(X = x) = F(x) - F(x-1)$

d) 1. $a = 1/2$ och $b = -1$

e) 5. $\mathrm{Cov}(X, Y) = 0$ men man kan inte avgöra om X och Y är oberoende

Uppgift 2

a) Villkoret $\int_{-\infty}^{\infty} f(x,y)dxdy = 1$ ger $c\int_0^2\int_0^2 xydxdy = 4c = 1$ och alltså $c = 1/4$.

b) Vi har att $f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy$ och får alltså för $x \in [0,1]$ att $f_X(x) = \int_0^2 \frac{1}{4}xydy = \frac{x}{2}$, och $f_X(x) = 0$ för övriga x. Av symmetriskäl fås också att $f_Y(y) = \frac{y}{2}$ för $y \in [0,2]$ och $f_Y(y) = 0$ för övriga y.

c) Vi har att $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$, där $\mathbb{E}[X] = \mathbb{E}[Y] = \int_{-\infty}^{\infty} yf_Y(y)dy = \int_0^2 \frac{y^2}{2}dy = \frac{4}{3}$, och alltså $\mathbb{E}[X+Y] = \frac{8}{3}$.

Uppgift 3

a) $\mathbb{P}(\text{alla visar samma färg}) = \mathbb{P}(\text{alla visar rött}) + \mathbb{P}(\text{alla visar gult}) + \mathbb{P}(\text{alla visar grönt}) = \left(\frac{3}{6}\right)^3 + \left(\frac{2}{6}\right)^3 + \left(\frac{1}{6}\right)^3 = \frac{1}{6}$.

b) $\mathbb{P}(\text{alla visar olika färg}) = \mathbb{P}(\text{en visar rött, en visar grönt, en visar gult}) = 3! \cdot \frac{3}{6} \cdot \frac{2}{6} \cdot \frac{1}{6} = \frac{1}{6}$, eftersom det finns 3! olika sätt att ordna de tre färgerna (dvs bestämma vilken färg som ska förekomma på vilken tärning) och varje sådan ordning förkommer med sannolikhet $\frac{3}{6} \cdot \frac{2}{6} \cdot \frac{1}{6}$.

Uppgift 4

a) Låt $Y = X^2$. Fördelningsfunktionen för Y är

$$F_Y(y) = \mathbb{P}(Y \le y) = \mathbb{P}(X^2 \le y) = \mathbb{P}(X \le \sqrt{y}) = F_X(\sqrt{y}).$$

Täthetsfunktionen för Y fås genom att derivera fördelningsfunktionen:

$$f_Y(y) = \frac{d}{dy} F_X(\sqrt{y}) = \frac{d\sqrt{y}}{dy} f_X(\sqrt{y}) = \frac{1}{2\sqrt{y}} \cdot 2\sqrt{y} e^{-y} = e^{-y}$$

för $y \ge 0$ och $f_Y(y) = 0$ för $y < 0$, dvs $Y \sim \text{Exp}(1)$.

b) Låt $Y_1 = X_1^2$ och $Y_2 = X_2^2$, och $Z = Y_1 + Y_2$. Faltningsformeln ger

$$f_Z(z) = \int_{-\infty}^{\infty} f_{Y_1}(z - y) f_{Y_2}(y) dy.$$

Integranden är positiv bara om $y > 0$ och $z - y > 0$, dvs $0 < y < z$, och vi får alltså

$$f_Z(z) = \int_0^z e^{-(z-y)} e^{-y} dy = \int_0^z e^{-x} dy = z e^{-z} \quad \text{för } z \ge 0$$

och $f_Z(z) = 0$ för $z < 0$.

Uppgift 5

Låt n vara antalet biljetter som säljs och låt $Y_i = 1$ om personen som köper den $i$te biljetten dyker upp och $Y_i = 0$ annars. Enligt flygbolagets uppskattningar och givna antaganden gäller att $Y_i \sim \text{Be}(0.95)$ för alla $i = 1, \dots, n$ och att $\{Y_i\}$ är oberoende. Låt $S_n = \sum_{i=1}^n Y_i$ beteckna antalet passagerare som dyker upp. Då gäller att $S_n \sim \text{Bin}(n, 0.95)$. Vi ska bestämma n

så att $\mathbb{P}(S_n \leq 300) > 0.99$. Enligt Centrala gränsvärdessatsen är S_n approximativt normalfördelad, eftersom det är en summa av många oberoende lika fördelade stokastiska variabler. Låt Φ beteckna fördelningsfunktionen för N(0,1)-fördelningen. Vi får då

$$\mathbb{P}\left(S_n \leq 300\right) = \mathbb{P}\left(\frac{S_n - \mathbb{E}[S_n]}{\mathrm{Var}(S_n)} \leq \frac{300 - \mathbb{E}[S_n]}{\sqrt{\mathrm{Var}(S_n)}}\right) \approx \Phi\left(\frac{300 - n \cdot 0.95}{\sqrt{n \cdot 0.05 \cdot 0.95}}\right).$$

Ska denna sannolikhet vara 0.99 måste $(300-n\cdot 0.95)/\sqrt{n \cdot 0.05 \cdot 0.95}$ ges av 0.01-kvantilen i N(0,1)-fördelningen, dvs $(300 - n \cdot 0.95)/\sqrt{n \cdot 0.05 \cdot 0.95} = 2.3263$. Detta leder till en andragradsekvation i $x = \sqrt{n}$ som har lösning $x = 17.5$, dvs $x^2 = 306.25$. Bolaget kan alltså sälja max 306 biljetter.

Uppgift 6

a) Låt $f(a)$ vara den förväntade förlusten när vi gissar a, dvs

$$f(a) = \mathbb{E}[(X-a)^2] = \mathbb{E}[X^2] + a^2 - 2a\mathbb{E}[X].$$

För att minimera $f(a)$ med avseende på a deriverar vi och söker nollställe till derivatan. Vi får $f'(a) = -2\mathbb{E}[X] + 2a$ och alltså $f'(a) = 0$ omm $a = \mathbb{E}[X]$. Detta är ett minimum eftersom $f''(a) = 2 > 0$.

b) Låt $h(a)$ vara förväntad förlust då man gissar a, dvs $h(a) = \mathbb{E}[|X-a|]$, vilket kan skrivas som

$$\begin{aligned} h(a) &= \int_{-\infty}^{a} (a-x)dx + \int_{a}^{\infty} (x-a)dx \\ &= \int_{-\infty}^{a} af(x)dx - \int_{-\infty}^{a} xf(x)dx + \int_{a}^{\infty} xf(x)dx - \int_{a}^{\infty} af(x)dx. \end{aligned}$$

De termer där integranden är $af(x)$ ger

$$\int_{-\infty}^{a} af(x)dx - \int_{a}^{\infty} af(x)dx = a[F(a) - (1-F(a))] = 2aF(a) - a$$

och de termer där integranden är $xf(x)$ ger

$$-\int_{-\infty}^{a} xf(x)dx + \int_{a}^{\infty} xf(x)dx = \mathbb{E}[X] - 2\int_{-\infty}^{a} xf(x)dx.$$

Alltså kan vi skriva

$$h(a) = 2aF(a) - a + \mathbb{E}[X] - 2\int_{-\infty}^{a} xf(x)dx.$$

Derivatan blir $h'(a) = 2af(a) + 2F(a) - 1 - 2af(a) = 2F(a) - 1$ och alltså $h'(a) = 0$ omm $F(a) = 1/2$, dvs om a är medianen i fördelningen. Detta är ett minimum eftersom $h''(a) = 2f(a) > 0$.