

Tenta Sannolikhets teori I 251204 lösningar

- [1] (a) Då A, B, C är parvis disjuncta följer det från sannolikhetsaxiomen att

$$P(A \cup B \cup C) = P(A) + P(B) + P(C).$$

Vi får

$$P(C) = 0.9 - 0.2 - 0.35 = \boxed{0.35}.$$

2p

- (b) Då A, B, C är parvis disjuncta och $D \subseteq A \cup B \cup C$ gäller

$$\begin{aligned} D &= [D \cap A] \cup [D \cap B] \cup [D \cap C] \\ &= [D \cap A] \cup [D \cap B], \end{aligned}$$

samt

$$P(D) = P(D \cap A) + P(D \cap B) = 0.1 + 0.15 = \boxed{0.25}.$$

3p

- [2] (a) Vi bestämmer

$$\begin{aligned} F_M(x) &= P(M \leq x) = 1 - P(M > x) \\ &= 1 - P(X > x, Y > x) \\ &= 1 - P(X > x)P(Y > x) = 1 - (1-x)^2 = \boxed{2x - x^2}. \end{aligned}$$

$$P(X > x) = \int_x^1 dz = 1 - x$$

$$\begin{aligned} F_M(x) &= P(M \leq x) = P(X \leq x, Y \leq x) \\ &= P(X \leq x)P(Y \leq x) = \boxed{x^2}. \end{aligned}$$

3p

Ovanstående gäller för $x \in [0, 1]$.

- (b) Sätt $A = \{M \leq 0.5\}$ och $B = \{M \leq 0.5\}$. Vi vill då beräkna $P(A \cap B^c)$. Vi har

$$P(A) = P(A \cap B) + P(A \cap B^c).$$

Då $m \leq M$ gäller $B \subseteq A$, så $A \cap B = B$ och

$$\begin{aligned} P(A \cap B^c) &= P(A) - P(B) \\ &= P(M \leq 0.5) - P(M \leq 0.5) \\ &= F_M(0.5) - F_M(0.5) = (1 - \frac{1}{4}) - \frac{1}{4} = \boxed{\frac{1}{2}}. \end{aligned}$$

2p

- [3] (a) Centrala gränsvärdessatsen säger att om $X_1, X_2, \dots$ är oberoende och likafördelade slumpvariabler med väntevärde μ och varians σ^2 , vilken är ändlig, så gäller för $x \in \mathbb{R}$

$$P\left(\frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \leq x\right) \rightarrow \Phi(x) \quad \text{då } n \rightarrow \infty. \quad 2p$$

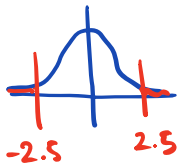
- (b) Då $n=100$ fr anses stort nog för att centrala gränsvärdessatsen skall ge god approximation så får vi

$$P(\bar{X}_{100} \leq 9.5) = P\left(10 \cdot \frac{\bar{X}_{100} - 10}{2} \leq 10 \cdot \frac{9.5 - 10}{2}\right)$$

$$\approx \Phi(-2.5) = 1 - \Phi(2.5) \quad 3p$$

$$\approx 1 - 0.9938 = 0.0062.$$

↑ or tabell



- [4] Variabeln sample genereras av funktionen get_sample, som återopats med input 3.

Då get_sample återopats med input 3 svarar den med summan av X_2 och X_3 . 2p

X_2 och X_3 är två vektorer av längd 10000 som innehåller oberoende observationer från en Poisson-fördelning med parameter 2 resp. 4.

Summan av två oberoende Poisson-variabler är igen Poisson, med parametern som är summan av parametrarna.

Ovanstående innebär att sample är en vektor av längd 10000 med oberoende observationer från en Poisson-fördelning med $\lambda = 2 + 4 = 6$. Då parametern motsvarar fördelningens väntevärde, och väntevärdet motsvarar histogrammets jämviktspunkt, så kan vi utesluta histogram 1, 2, 3 och 6. I histogram 4 ser 3p

Histogram 5 bör vara korrekt svar.

5 (a) Drag en nyexaminerad student på måfå. Låt

$H = \{ \text{student högpresterande} \}$

$A = \{ \text{student program A} \}$

$B = \{ \text{student program B} \}$

Från uppgift gäller

$$P(A) = 0.4$$

$$P(H|A) = 0.8$$

$$P(B) = 0.6$$

$$P(H|B) = 0.8$$

5p

Vi får från lagen om total sannolikhet

$$P(H) = P(H|A)P(A) + P(H|B)P(B) = \boxed{0.8}$$

(b) Bayes sats ger

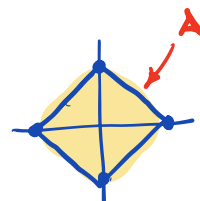
$$P(A|H) = \frac{P(H|A)P(A)}{P(H)} = \frac{0.8}{0.8} P(A) = \boxed{0.4}$$

2p

(c) Undersökningen som nyligen utförts visar att rekryteraren har fel - andelen högpresterande är lika inom de två utbildningsprogrammen. Dock har program B en större marknadsandel, vilket gör att bland högpresterande studenter är det en högre andel som kommer från program B, vilket svaret i (b) visar. Rekryteraren har sannolikt träffat fler högpresterande från program B och felaktigt fått intrycket att andelen högpresterande där är högre. 3p

6 (a) Den likformiga fördelningen på A ges av tätheten

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{|A|} & \text{om } (x,y) \in A \\ 0 & \text{annars} \end{cases}$$



Arean för A är $|A| = 2$.

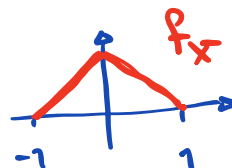
3p

- (b) För att bestämma väntevärde och varians bestämmer vi först marginaltäteten.

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dy$$

$$= \begin{cases} \int_{-x-1}^{x+1} \frac{1}{2} dy = \frac{1}{2}(2x+2) = x+1, & -1 \leq x \leq 0 \\ \int_{x-1}^{-x+1} \frac{1}{2} dy = \frac{1}{2}(-2x+2) = -x+1, & 0 < x \leq 1 \end{cases}$$

Av symmetriskäl gäller $\mathbb{E}[X] = 0$.
Därför för vi



$$\begin{aligned} \text{Var}(X) &= \mathbb{E}[X^2] = \int_{-1}^1 x^2(1-|x|) dx \\ &= \int_{-1}^0 (x^2 + x^3) dx + \int_0^1 (x^2 - x^3) dx \\ &= \left[\frac{x^3}{3} + \frac{x^4}{4} \right]_{-1}^0 + \left[\frac{x^3}{3} - \frac{x^4}{4} \right]_0^1 \\ &= \frac{1}{3} - \frac{1}{4} + \frac{1}{3} - \frac{1}{4} = \frac{2}{3} - \frac{1}{2} = \frac{1}{6}. \end{aligned}$$

4p

- (c) Av symmetriskäl gäller $\mathbb{E}[Y] = \mathbb{E}[X] = 0$. Därmed för vi

$$\text{Cov}(X,Y) = \mathbb{E}[XY] = \iint_A xy \cdot \frac{1}{2} dx dy$$

För $y \geq 0$ gäller

$$\int_{-(1-y)}^{1-y} \frac{1}{2} xy dy = 0,$$

da vi integrerar en udda funktion. Samma för $y < 0$

$$\int_{-(1+y)}^{1+y} \frac{1}{2} xy dy = 0.$$

Alltså för vi

$$\text{Cov}(X,Y) = \int_{-1}^0 \int_{-(1+y)}^{1+y} \frac{1}{2} xy dx dy + \int_0^1 \int_{-(1-y)}^{1-y} \frac{1}{2} xy dx dy = 0.$$

3p

- 7 (a) Då kortleken är välblandad kan vi anta att varje val av fem kort är lika sannolikt. Den klassiska sannolikhetsdefinitionen ger då

$$P(\text{precis ett par}) = \frac{\# \text{ händer med ett par precis}}{\# \text{ händer totalt}}$$

Det totala antalet händer om fem kort är $\binom{52}{5}$. Antalet som ger precis ett par är

$$\underbrace{\binom{13}{1} \binom{4}{2}}_{\text{val av valör och färg för par}} \underbrace{\binom{12}{3} \binom{4}{1} \binom{4}{1} \binom{4}{1}}_{\text{val av valörer och färg för övriga}} = 13 \cdot 6 \cdot 220 \cdot 4 \cdot 4 \cdot 4$$

3 p

Vi får

$$P(\text{precis ett par}) = \frac{13 \cdot 6 \cdot 220 \cdot 4 \cdot 4 \cdot 4}{52 \cdot 51 \cdot 50 \cdot 49 \cdot 48} \cdot 5! \approx \boxed{0.42}$$

- (b) För att få "dubbelpar eller bättre" behöver vi antingen två par och ett ensamt kort, eller tre kort av en valör och två av en annan. Antalet händer av vardera av dessa:

$$\underline{2+2+1}: \underbrace{\binom{13}{2} \binom{4}{2} \binom{4}{2}}_{\text{val av de två paren}} \underbrace{\binom{11}{1} \binom{4}{1}}_{\text{val av ensamt}} = 13 \cdot 6 \cdot 6 \cdot 6 \cdot 11 \cdot 4$$

$$\underline{3+2}: \underbrace{\binom{13}{1} \binom{4}{3}}_{\text{val av tripp}} \underbrace{\binom{12}{1} \binom{4}{2}}_{\text{val av par}} = 13 \cdot 4 \cdot 12 \cdot 6$$

4 p

Vi får

$$P(\text{dubbelpar el. bättre}) = \frac{13 \cdot 6 \cdot 6 \cdot 6 \cdot 11 \cdot 4 + 13 \cdot 4 \cdot 12 \cdot 6}{\binom{52}{5}} \approx \boxed{0.049}$$

(c)

$$P(\text{tre av samma} \mid \text{dubbelpar el. bättre}) = \frac{P(\text{tre} + \text{två})}{P(\text{dubbelpar el. bättre})}$$

$$= \frac{13 \cdot 4 \cdot 12 \cdot 6}{13 \cdot 4 \cdot 12 \cdot 6 + 13 \cdot 6 \cdot 6 \cdot 6 \cdot 11 \cdot 4} \approx \boxed{0.029}$$

3 p

8 (a) Enligt centrala gränsvärdesatsen kommer Z_n vara approximativt $N(0,1)$ -fördelad då n är stort, oavsett fördelning på X_n , så länge variansen är ändlig. Påstående (i) är korrekt.

Om X_n antar ändligt många värden, och därmed är diskret, så kommer Z_n oavsett n anta endast ändligt många värden, och därmed vara diskret. Då en normalfördelning är kontinuerlig så behöver inte likhet i fördelning uppstå för ändligt n . Påstående (ii) är felaktigt. 5p

(b) Då m är stort kommer histogrammen väl återge respektive fördelning. Dock krävs det att n är stort för att de två fördelningarna skall vara approximativt lika.

Histogrammen kan därmed komma att skilja sig både då n och m är små, men av olika anledningar. Då n är litet ty fördelningarna skiljer sig. Då m är litet ty histogrammen ej återger väl fördelningarna. 5p