UPPSALA UNIVERSITET
Matematiska institutionen
Örjan Stenflo

TENTAMEN I MATEMATIK
Sannolikhetsteori I, 1MS034
2021-10-25

Skrivtid: 8–13. Varje problem ger max 5 poäng. För betygen 3,4 resp. 5 krävs 18, 25 resp. 32 poäng (inklusive eventuella bonuspoäng). Tillåtna hjälpmedel: miniräknare samt formelsamling för kursen Sannolikhetsteori I, 1MS034. Lösningarna skall vara väl motiverade och försedda med förklarande text.

1. För två händelser A och B i ett utfallsrum Ω gäller $P(A) = 0.4$, $P(B) = 0.4$, och $P(A \cup B) = 0.5$.

   (a) Bestäm $P(A|B)$.

   (b) Låt $X(\omega) = \begin{cases} 1 & \text{om } \omega \in A \\ 0 & \text{annars} \end{cases}$ och $Y(\omega) = \begin{cases} 1 & \text{om } \omega \in B \\ 0 & \text{annars} \end{cases}$.
   Bestäm korrelationskoefficienten $\rho(X, Y)$.

   Lösning:

   (a) Då $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, ser vi att

   $$P(A \cap B) = P(A) + P(B) - P(A \cup B) = 0.4 + 0.4 - 0.5 = 0.3,$$

   så

   $$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.3}{0.4} = 0.75.$$

   (b) Notera att $X \sim Be(0.4)$ och $Y \sim Be(0.4)$ så $V(X) = P(A)(1 - P(A)) = 0.24$ och $V(Y) = P(B)(1 - P(B)) = 0.24$. Då kovariansen mellan X och Y är

   $$C(X, Y) = E(XY) - E(X)E(Y) = P(A \cap B) - P(A)P(B) = 0.3 - 0.4 \cdot 0.4 = 0.14,$$

   blir korrelationskoefficienten

   $$\rho(X, Y) = \frac{C(X, Y)}{D(X)D(Y)} = \frac{0.14}{0.24} = 7/12 \approx 0.583.$$

2. Låt X vara en stokastisk variabel med sannolikhetsfunktion:

   $$p_X(k) = P(X = k) = \begin{cases} \alpha & k = 1, 3 \\ \alpha\beta & k = 4 \\ 0 & \text{annars.} \end{cases},$$

   där $\alpha > 0$ och $\beta > 0$ är konstanter.

(a) Bestäm ett samband mellan α och β så att detta verkligen är en sannolikhetsfunktion.

(b) Antag att $E(X) = 3$. Bestäm exakta värden på α och β samt beräkna variansen av X.

Lösning:

(a) Då

$$P(X = 1) + P(X = 3) + P(X = 4) = \alpha(2 + \beta) = 1,$$

är

$$\alpha = \frac{1}{2 + \beta},$$

så

$$\beta = \alpha^{-1} - 2.$$

(b)

$$\begin{aligned} E(X) &= P(X = 1) + 3P(X = 3) + 4P(X = 4) \\ &= \alpha(1 + 3 + 4\beta) = \frac{4 + 4\beta}{2 + \beta} \\ &= 2 + \frac{2\beta}{2 + \beta} = 3, \end{aligned}$$

så $\beta = 2$ och $\alpha = 1/4$ och alltså är $P(X = 1) = P(X = 3) = 1/4$ och $P(X = 4) = 1/2$. Då

$$E(X^2) = P(X = 1) + 3^2 P(X = 3) + 4^2 P(X = 4) = 1/4 + 9/4 + 8 = 21/2,$$

blir

$$V(X) = E(X^2) - (E(X))^2 = 21/2 - 9 = 3/2.$$

3. Man har fyra kulor. Varje kula färgas vit med sannolikhet 1/2 och röd med sannolikhet 1/2 oberoende av varandra och placeras sedan i en urna som blandas om. Ur urnan drar man sedan fyra kulor med återläggning mellan varje dragning. Beräkna den betingade sannolikheten att alla kulor i urnan är vita givet att alla dragna kulor är vita.

Lösning: Låt $X \sim \text{Bin}(4, 1/2)$ beteckna antalet vita kulor i urnan, och låt Y beteckna antalet dragna vita kulor.

Då är enligt definitionen av betingad sannolikhet och lagen om total sannolikhet

$$\begin{aligned}
P(X=4|Y=4) &= \frac{P(X=4,Y=4)}{P(Y=4)} = \frac{P(Y=4|X=4)P(X=4)}{P(Y=4)} = \frac{P(X=4)}{P(Y=4)} \\
&= \frac{P(X=4)}{\sum_{k=0}^{4} P(Y=4|X=k)P(X=k)} = \frac{(1/2)^4}{\sum_{k=0}^{4}(k/4)^4 P(X=k)} \\
&= \frac{(1/2)^4}{\sum_{k=0}^{4}(k/4)^4\binom{4}{k}(1/2)^4} \\
&= \frac{1}{\sum_{k=1}^{4}(k/4)^4\binom{4}{k}} = \frac{4^4}{\sum_{k=1}^{4}k^4\binom{4}{k}} = \frac{4^4}{680} = \frac{32}{85} \approx 0.376.
\end{aligned}$$

4. Låt N vara antalet tärningskast man behöver göra tills man fått 6:a 1000 gånger. Beräkna approximativt

$$P(|N-6000|<300).$$

Lösning: Då $N = \sum_{i=1}^{1000} X_i$, där $X_i \sim$ ffg$(1/6)$ är oberoende är enligt centrala gränsvärdessatsen N approximativt normalfördelad med väntevärde

$$E(N) = 1000E(X_1) = 1000\cdot(1/(1/6) = 6000,$$

och varians

$$V(N) = 1000V(X_1) = 1000(1-1/6)/(1/6)^2 = 30000.$$

Från centrala gränsvärdessatsen (med kontinuitetskorrektion (halvkorrektion)) följer därför att

$$\begin{aligned}
P(|N-6000|<300) &= P(5700<N<6300) = P(5700.5<N\le 6299.5) \\
&= P(\frac{5700.5-6000}{\sqrt{30000}} < \frac{N-6000}{\sqrt{30000}} \le \frac{6299.5-6000}{\sqrt{30000}})) \\
&\approx \Phi(299.5/\sqrt{30000}) - \Phi(-299.5/\sqrt{30000}) \\
&= 2\Phi(\underbrace{299.5/\sqrt{30000}}_{\approx 1.73}) - 1 \underbrace{\approx}_{\text{tabell}} 2\cdot 0.958 - 1 \approx 0.92,
\end{aligned}$$

där Φ betecknar fördelningsfunktionen av en standardnormal fördelad stokastisk variabel.

5. Från en fruktskål med 2 äpplen 8 päron och 8 apelsiner tar man slumpmässigt 3 frukter.

   (a) Beräkna sannolikheten att minst en av de 3 tagna frukterna är ett päron.

   (b) Låt X beteckna antalet olika fruktsorter som finns bland de 3 tagna frukterna. Beräkna $E(X)$.

Lösning:

(a) Låt $Y \sim \text{Hyp}(18, 3, 8)$ vara antalet dragna päron.

$$P(Y \geq 1) = 1 - P(Y = 0) = 1 - \frac{\binom{10}{3}}{\binom{18}{3}} = 1 - \frac{10 \cdot 9 \cdot 8}{18 \cdot 17 \cdot 16} = 29/34 \approx 0.853.$$

(b) Låt $X_1 = \begin{cases} 1 & \text{minst ett äpple väljs} \\ 0 & \text{annars} \end{cases}$, $X_2 = \begin{cases} 1 & \text{minst ett päron väljs} \\ 0 & \text{annars} \end{cases}$ och $X_3 = \begin{cases} 1 & \text{minst en apelsin väljs} \\ 0 & \text{annars} \end{cases}$. Då

$$E(X_1) = P(\text{minst ett äpple väljs}) = 1 - \frac{16 \cdot 15 \cdot 14}{18 \cdot 17 \cdot 16} = 16/51,$$

och enligt uppgift (a) (och symmetri) $E(X_2) = E(X_3) = 29/34$, blir

$$E(X) = E(X_1 + X_2 + X_3) = E(X_1) + E(X_2) + E(X_3) = 16/51 + 29/34 + 29/34 = 103/51.$$

Alternativ lösning:

$$P(X = 1) = P(\text{"endast päron väljs"}) + P(\text{"endast apelsiner väljs"}) = 2\frac{\binom{8}{3}}{\binom{18}{3}} = 7/51.$$

$$\begin{aligned} P(X = 2) &= P(\text{"exakt 2 äpplen"}) + P(\text{"exakt 2 päron"}) + P(\text{"exakt 2 apelsiner"}) \\ &= \frac{\binom{2}{2}\binom{16}{1} + 2 \cdot \binom{8}{2}\binom{10}{1}}{\binom{18}{3}} = 36/51 = 12/17. \end{aligned}$$

$$P(X = 3) = P(\text{"1 äpple, 1 päron, 1 apelsin väljs"}) = \frac{\binom{2}{1}\binom{8}{1}\binom{8}{1}}{\binom{18}{3}} = 8/51$$

(Kontroll: $\sum_{k=1}^{3} P(X = k) = (7 + 36 + 8)/51 = 1.$)
Per definition är därför

$$E(X) = \sum_{k=1}^{3} kP(X = k) = 1 \cdot (7/51) + 2 \cdot (36/51) + 3 \cdot (8/51) = 103/51.$$

6. Låt X och Y vara oberoende Exp(1)-fördelade stokastiska variabler och definiera $Z = X + Y$.

   (a) Bestäm den momentgenererande funktionen för Z.

   (b) Bestäm $E(Z^3)$.

   Lösning:

(a) Då Z kan skrivas som en summa av två oberoende Exp(1)-fördelade stokastiska variabler gäller $Z \sim \Gamma(2,1)$, så Z har (enligt formelsamlingen) momentgenererande funktion $\psi(t) = (\frac{1}{1-t})^2 = (1-t)^{-2}$. (Alternativt kan vi se detta som produkten av de momentgenerande funktionerna för två Exp(1)-fördelade stokastiska variabler.)

(b) Eftersom $E(Z^n) = \psi^{(n)}(0)$, och

$$\psi^{(3)}(t) \quad = \quad 2 \cdot 3 \cdot 4(1-t)^{-5},$$

är

$$E(Z^3) \quad = \quad \psi^{(3)}(0) = 24.$$

7. I land 1 är kroppslängden hos invånarna (i cm) normalfördelad med väntevärde 165 och varians 49 och i land 2 är kroppslängden hos invånarna (i cm) normalfördelad med väntevärde 175 och varians 64. Antag att man slumpmässigt väljer ut 3 invånare från land 1 och 2 invånare från land 2. Låt Z beteckna kroppslängden av den längsta bland de 5 utvalda invånarna.

(a) Bestäm fördelningsfunktionen för Z.

(b) Bestäm approximativt 0.05-kvantilen för Z.

Lösning:

(a) Låt $X_i \sim N(165, 49)$, $i = 1, \ldots, 3$ och $Y_i \sim N(175, 64)$, $i = 1, 2$ där X_1, X_2, X_3, Y_1, Y_2 är oberoende och låt $Z = \max(X_1, X_2, X_3, Y_1, Y_2)$. Då $P(X_1 \leq t) = \Phi((t-165)/7)$ och $P(Y_1 \leq t) = \Phi((t-175)/8)$ blir fördelningsfunktionen

$$\begin{aligned} F_Z(t) = P(Z \leq t) &= P(X_1 \leq t, X_2 \leq t, X_3 \leq t, Y_1 \leq t, Y_2 \leq t) \\ &= (P(X_1 \leq t))^3 (P(Y_1 \leq t))^2 \\ &= (\Phi(\frac{t-165}{7}))^3 (\Phi(\frac{t-175}{8}))^2. \end{aligned}$$

(b) Vi söker t så att

$$P(Z \leq t) = (\Phi(\frac{t-165}{7}))^3 (\Phi(\frac{t-175}{8}))^2 = 0.95.$$

Då $(\Phi((t-175)/8))^2 = 0.95$ om $\Phi((t-175)/8) = \sqrt{0.95} \approx 0.9747$ ser vi från tabell att $(\Phi((t-175)/8))^2 = 0.95$ om $(t-175)/8 \approx 1.95$ alltså $t \approx 190.6$. Då $(\Phi((190.6-165)/7))^3 \approx 1$ är alltså 0.05-kvantilen för Z ungefär 190.6.

8. Låt $X_1, X_2, X_3, \ldots$ vara oberoende och lika fördelade, med väntevärde 0 och varians 1. Antag vidare att det finns ett $h > 0$ så att den momentgenererande funktionen $\psi_{X_1}(t)$ är ändlig för $|t| < h$.

Vad säger centrala gränsvärdessatsen i denna situation

(a) i. formulerad i termer av fördelningsfunktioner?
    ii. formulerad i termer av momentgenererande funktioner?

(b) Ge ett bevis av det senare. (Du kan anta att Taylorutvecklingen $\psi_{X_1}(t) = 1 + E(X_1)t + E(X_1^2)t^2/2 + O(t^3)$ gäller då $t \to 0$.)

Lösning:

Låt

$$Z_n = \frac{\sum_{i=1}^n X_i}{\sqrt{n}}.$$

(a) Om F_{Z_n} betecknar fördelningsfunktionen för Z_n så gäller $F_{Z_n}(t) \to \Phi(t)$, då $n \to \infty$, där Φ betecknar fördelningsfunktionen för en standard normalfördelad stokastisk variabel.

(b) Om ψ_{Z_n} betecknar den momentgenererande funktionen för Z_n så gäller $\psi_{Z_n}(t) \to \psi(t)$, då $n \to \infty$, där $\psi(t) = e^{t^2/2}$ betecknar den momentgenererande funktionen för en standard normalfördelad stokastisk variabel.

(c) Då $\psi_{X+Y}(t) = \psi_X(t)\psi_Y(t)$, och $\psi_{aX}(t) = \psi_X(at)$ om X och Y är oberoende stokastiska variabler och a är en konstant, och $E(X_1) = 0$ och $V(X_1) = E(X_1^2) - (E(X_1))^2 = 1$, gäller

$$\psi_{Z_n}(t) = \psi_{\sum_{i=1}^n X_i}(t/\sqrt{n}) = (\psi_{X_1}(t/\sqrt{n}))^n = (1 + \frac{t^2}{2n} + O(\frac{t^3}{n^{3/2}}))^n \to e^{t^2/2},$$

då $n \to \infty$, och då $\psi(t) = e^{t^2/2}$ är den momentgenererande funktionen för en $N(0,1)$-fördelad stokastisk variabel har vi därmed visat CGS formulerad i termer av momentgenererande funktioner.