UPPSALA UNIVERSITET
Matematiska Institutionen
Ingemar Kaj

PROV I MATEMATIK
Sannolikhet och statistik, 1MS321
2016-06-02
Suggested solutions

1. Lösningar till a), b) och c) fås genom att utföra beräkningarna

$$E(X) = \int_{-\infty}^{\infty} x f_X(x)\,dx = \int_0^4 x \cdot \frac{x}{8} = \frac{1}{8} \cdot \Big[\frac{x^3}{3}\Big]_{x=0}^4 = \frac{4^3}{24} = \frac{8}{3}$$
$$E(1/X) = \int_{-\infty}^{\infty} (1/x) f_X(x)\,dx = \int_0^4 \frac{1}{x} \cdot \frac{x}{8} = \frac{1}{8}\int_0^4 1\,dx = \frac{4}{8} = \frac{1}{2}$$
$$C(X, 1/X) = E(X \cdot 1/X) - E(X)\,E(1/X) = 1 - E(X)E(1/X) = 1 - \frac{8}{3}\cdot\frac{1}{2} = -\frac{1}{3}$$

2. Sätt
$$A = \{000 \text{ sänds}\}, \quad B = \{111 \text{ sänds}\}, \quad C = \{011 \text{ mottas}\}.$$
Enligt antagande har vi $P(A) = 0.4$ och $P(B) = 0.6$, och vidare
$$P(C|A) = 0.9 \cdot 0.1 \cdot 0.1 = 0.0090, \quad P(C|B) = 0.05 \cdot 0.95 \cdot 0.95 = 0.045125$$
a) Enligt lagen om total sannolikhet får vi
$$P(C) = P(C|A)P(A) + P(C|B)P(B) = 0.009 \cdot 0.4 + 0.045125 \cdot 0.6 = 0.0307$$
b) Bayes sats ger nu
$$P(A|C) = \frac{P(C|A)P(A)}{P(C)} = \frac{0.009 \cdot 0.4}{0.0307} = 0.1174$$
c) Komplement ger $P(B|C) = 1 - P(A|C) = 0.8826$.

3. a) Skattningens väntevärde är
$$E(\widehat{m}_1) = E[(X+Y)/2] = E(X)/2 + E(Y)/2 = (m_1 - m_2)/2 + (m_1 + m_2)/2 = m_1,$$
vilket ger väntevärdesriktigheten, per definition.
b) På grund av oberoendet är skattningens varians och standardavvikelse
$$V(\widehat{m}_1) = V(\frac{X+Y}{2}) = \frac{V(X)+V(Y)}{4} = \frac{9}{4}, \quad D(\widehat{m}_1) = \frac{3}{2}$$
c) Sätt $\widehat{m}_2 = (Y - X)/2$. Som ovan visas att $E(\widehat{m}_2) = m_2$ och $D(\widehat{m}_2) = 3/2$.

4. Enligt modellantagande ges antalet incidenter under tidslängden t av en Poissonprocess $N(t)$, $t \geq 0$, med intensitet $\lambda = 0.5$ stycken per vecka. Speciellt är $N(t)$ Poisson-fördelad med väntevärde $\lambda t = t/2$.
a) $P(N(4) = 0) = e^{-\lambda \cdot 4} = e^{-2} = 0.135$
b) Antalet incidenter per vecka är Poissonfördelat med väntevärde 1/2, och varje ny vecka är oberoende av föregående. Alltså
$$P(\text{en per vecka under fyra veckor}) = P(N(1) = 1)^4 = (0.5 \cdot e^{-0.5})^4 = e^{-2}/16 \approx 0.0085$$

5. Vi kan beskriva spelresultaten med variablerna
$$X_k = \text{vinst spelomgång } k = \begin{cases} -1, & \text{sannolikhet } 0.52 \\ +1, & \text{sannolikhet } 0.48 \end{cases} \qquad Y = \sum_{k=1}^{n} X_k = \text{vinst efter } n \text{ spel}$$

Vi har $E(X_k) = -1 \cdot 0.52 + 1 \cdot 0.48 = -0.04$, $E(X_k^2) = (-1)^2 \cdot 0.52 + 1 \cdot 0.48 = 1$ och $V(X_k) = 1 - (-0.04)^2 = 0.9984$. Med $n = 100$ fås $E(Y) = -4$ och $V(Y) = 99.84$. Enligt centrala gränsvärdessatsen är Y approximativt normalfördelad $N(-4\,,\,99.84)$, dvs $Z = (Y+4)/9.992$ är ungefär $N(0,1)$. Man finner

$$P(Y \geq -5) = P(Z \geq -1/9.992) = P(Z \geq -0.100) \approx \Phi(0.100) = 0.5398.$$

Alternativ: $X \sim \text{Bin}(100, 0.48)$ är antal vunna omgångar. Sökt: $P(X \geq 48) = 0.5393$.

6. a) Sannolikheterna för alla vägar ut från en nod måste summera till ett. Alltså fås $\star = 1/2$, $\star\star = 3/4$.

   b) Motsvarande övergångsmatris blir

$$\mathbf{P} = \begin{pmatrix} 1/2 & 0 & 1/2 \\ 3/4 & 1/4 & 0 \\ 1/4 & 3/4 & 0 \end{pmatrix}$$

   c) Den stationära fördelningen $\pi = (\pi_0\ \pi_1\ \pi_2)$, fås från balansekvationerna $\pi = \pi\mathbf{P}$, vilka utskrivna i detta fall blir

$$\begin{aligned} \frac{1}{2}\pi_0 + \frac{3}{4}\pi_1 + \frac{1}{4}\pi_2 &= \pi_0 \\ \frac{1}{4}\pi_1 + \frac{3}{4}\pi_2 &= \pi_1 \\ \frac{1}{2}\pi_0 \qquad\qquad\qquad &= \pi_2 \end{aligned}$$

   dvs $\pi_2 = \pi_0/2$, $\pi_1 = \pi_2$. Normering enligt $\pi_0 + \pi_1 + \pi_2 = 1$ ger nu lösningen $\pi = (1/2\ \ 1/4\ \ 1/4)$.

7. Vi har en observation $x = 178$ av variabeln

$$X = \text{antal detekterade fel} \sim \text{Bin}(n, p).$$

   där $n = 200$. Punktskattningen av detekteringssannolikheten p är $\widehat{p} = X/n$ och motsvarande numeriska skattning $\widehat{p} = x/n = 178/200 = 0.890$. Skattningens varians och medelfel ges av

$$V(\widehat{p}) = \frac{p(1-p)}{200}, \quad d(\widehat{p}) = \sqrt{\frac{\widehat{p}(1-\widehat{p})}{200}} = 0.0221.$$

   Med normalapproximation av binomialfördelningen (som fungerar bra eftersom $n\widehat{p}(1-\widehat{p}) \approx 19.5 > 10$) fås nu ett approximativt 95% konfidensintervall genom att sätta

$$\begin{aligned} I_p: \quad & [\widehat{p} - \lambda_{0.025}\, d(\widehat{p})\ ,\ \widehat{p} + \lambda_{0.025}\, d(\widehat{p})] \\ & = [0.890 - 1.96 \cdot 0.0221\ ,\ 0.890 + 1.96 \cdot 0.0221] \approx [0.847\ ,\ 0.933] \end{aligned}$$

8. Med den angivna paketankomstfördelningen X finner man $E(X) = 3 \cdot p + 0 \cdot (1-p) = 3p$, $E(X^2) = 3^2 \cdot p = 9p$. Under jämviktsförhållanden gäller då

$$P(Q = 0) = \frac{1-3p}{1-p}, \quad E(Q) = \frac{9p - 3p}{2(1-3p)} = \frac{3p}{1-3p}.$$

   Genom att betrakta funktionerna ser man att vi måste ha $0 \leq p < 1/3$. Om $p < 1/3$ så kan systemet förr eller senare ta hand om de paket som får vänta, dvs alla inkommande paket kommer också att levereras, om än med viss fördröjning. Inga förluster, så "output" är lika med "input". Systemets throughput är alltså samma som förväntat antal inkommande paket per slot, dvs $E(X) = 3p$. Om $p \geq 1/3$ blir både sannolikheten $P(Q = 0)$ och köstorleken $E(Q)$ negativa, vilket är orimligt. Tolkningen blir att om p är för stor så kommer det in mer arbete till systemet än vad dess kapacitet klarar av att utföra per tidsenhet, dvs kön växer okontrollerat och det finns ingen chans för länken att beta av de jobb som väntar.