

FACIT: Tentamen L9MA30, LGMA30

Matematik 3 för lärare, åk 7-9, Sannolikhetslära och statistik,
Matematik 3 för gymnasielärare, Sannolikhetslära och statistik

2015-10-23 kl. 8.30-12.30

Examinator och Jour: Henrike Häbel, tel. 073 0671252 alt. 031 772 5380

Hjälpmedel: Valfri miniräknare, (dubbelsidigt) A4 ark med egna anteckningar samt med skrivningen utdelade tabellsidor och formelsamling.

Maxpoäng: 30. För godkänd krävs minst 15 p, för Väl Godkänd 24 p. Eventuell bonuspoäng från 2015 adderas. Lösningar skall redovisas till varje uppgift. Uppgifterna är ej ordnade efter svårighetsgrad. Rättningsprotokoll anslås senast tre veckor efter tentamen.

1. Visa att för två händelser A och B i något utfallsrum Ω , där $P(A \cup B) < 1$, gäller

$$P(A \cap B) > 1 - P(A^c) - P(B^c).$$

2p

Lösningsförslag:

$$\begin{aligned} P(A \cap B) &\stackrel{reg.}{=} P(A) + P(B) - P(A \cup B) \quad (0.5p) \\ &\stackrel{reg.}{=} 1 - P(A^c) + 1 - P(B^c) - P(A \cup B) \quad (0.5p) \\ &= 2 - \underbrace{P(A \cup B)}_{<1} - P(A^c) - P(B^c) \quad (0.5p) \\ &\quad \underbrace{\hspace{1.5cm}}_{>1} \\ &> 1 - P(A^c) - P(B^c) \quad (0.5p) \end{aligned}$$

2. I en urna finns 26 kulor med bokstäver A-Z (21 konsonanter och 5 vokaler). Vi drar 4 kulor en efter en utan återläggning. Beräkna sannolikheterna för händelserna

a) A = 'den första bokstaven är A',

b) B = 'vi drar exakt en konsonant och tre vokaler'

1p

1p

- c) C = 'den första bokstaven är A och bland de tre andra bokstäver finns C'.

2p

Lösningförslag:

a)

$$P(A) = \frac{1}{26} \approx 0.038 \quad (1p)$$

b)

$$P(B) = \frac{\binom{5}{3} \binom{21}{1}}{\binom{26}{4}} = \frac{10 \cdot 21}{14950} \approx 0.014 \quad (1p)$$

c)

$$P(C) = \underbrace{\frac{1}{26}}_A \cdot \underbrace{3 \cdot \frac{1}{25}}_{\text{bland de tre andra finns C}} \approx 0.0046 \quad (2p)$$

3. Ett företag tillverkar TV-apparater. En TV-apparat är felaktig med sannolikhet 0.06. Varje TV-apparat testas om den är felaktig innan den lämnar företaget. Testen anger ett fel för felaktiga TV-apparater med sannolikhet 0.85. Testen anger också ett fel för ej felaktiga TV-apparater med sannolikhet 0.1. En slumpmässigt utvald TV-apparat testas. Hur stor är sannolikheten att testen anger ett fel? Vad är sannolikheten att TV-apparaten är felaktig, om testen anger ett fel?

3p

Lösningförslag:

A_1 := 'TV felaktig', A_2 := 'TV inte felaktig', B := 'testen anger fel'.

$$P(A_1) = 0.06, P(A_2) = 1 - 0.06 = 0.94, P(B|A_1) = 0.85, P(B|A_2) = 0.1 \quad (0.5p)$$

$$\begin{aligned} P(A_1|B) &\stackrel{\text{betingad}}{=} \frac{P(A_1 \cap B)}{P(B)} \stackrel{\text{Bayes}}{=} \frac{P(A_1)P(B|A_1)}{P(A_1)P(B|A_1) + P(A_2)P(B|A_2)} \quad (2p) \\ &= \frac{0.06 \cdot 0.85}{0.06 \cdot 0.85 + 0.94 \cdot 0.1} = \frac{0.051}{0.145} \approx 0.352 \quad (0.5p) \end{aligned}$$

4. Antalet fel på en viss typ av lakansväv (som ligger i rullar) kan anses vara Poissonfördelad med i genomsnitt tre fel per tio meter väv. En kund köper 5 stycken vävbitar om vardera 2 meter. Vad är sannolikheten att det på dessa finns fler än 2 fel?

2p

Lösningförslag:

5 stycken vävbitar om vardera 2 meter blir en hel tygbit på 10 meter med i genomsnitt tre fel, d.v.s. för X = 'antal fel på kundens vävbitar' har vi $\mathbb{E}(X) = \lambda = 3$ (1p). Därmed får vi

$$P(X > 2) = 1 - P(X \leq 2) = 1 - (P(X = 0) + P(X = 1) + P(X = 2)) = 1 - 0.4232 = 0.5768 \quad (1p)$$

5. Låt X vara lånetid av böcker på biblioteken (i månader). Antag att X är en kontinuerlig stokastisk variabel med fördelningsfunktion

$$F_X(x) = \begin{cases} 0 & , x < 0 \\ x^2/4 & , 0 \leq x < 2, \\ 1 & , 2 \leq x \end{cases} .$$

- a) Vad är sannolikheten att lånetiden för en slumpmässigt utvald bok blir mellan en halv och en månad?

1p

- b) Beräkna täthetsfunktionen $f_X(x) = F'_X(x)$.

1p

- c) Beräkna den förväntade lånetiden $\mathbb{E}(X)$ och variansen $V(X)$.

3p

Lösningsförslag:

- a)

$$P(0.5 \leq X \leq 1) = F_X(1) - F_X(0.5) = \frac{1}{4} - \frac{1}{16} = \frac{3}{16} \quad (1p)$$

- b)

$$f_X(x) = F'_X(x) = \frac{1}{2}x \quad \forall 0 \leq x \leq 2 \quad \text{och } 0 \text{ annars} \quad (1p)$$

- c)

$$\mathbb{E}(X) = \int_0^2 \frac{x^2}{2} dx = \left[\frac{x^3}{6} \right]_0^2 = \frac{4}{3} \quad (1p)$$

och

$$\mathbb{E}(X^2) = \int_0^2 \frac{x^3}{2} dx = \left[\frac{x^4}{8} \right]_0^2 = 2 \quad (1p)$$

och därmed

$$V(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = 2 - \frac{16}{9} = \frac{2}{9} \quad (1p)$$

6. Betrakta följande simultana sannolikhetsfunktion för X och Y med värden i $\{1, 2, 4\}$ respektive $\{-1, 0, 2\}$.

(X, Y)	-1	0	2
1	$\frac{1}{8}$	$\frac{1}{6}$	$\frac{1}{12}$
2	$\frac{1}{12}$	$\frac{1}{24}$	a
4	$\frac{1}{8}$	b	0

- a) Bestäm konstanterna $a, b \in \mathbb{R}$. Du kan använda att $P(X = 2) = \frac{1}{4}$.

1p

- b) Beräkna kovariansen $C(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$.

3p

- c) Är händelserna $X = 1$ och $Y = -1$ oberoende? OBS, du behöver inte testa om X och Y är oberoende!

1p

Lösningsförslag:

- a)

$$a = \frac{1}{4} - \frac{1}{12} - \frac{1}{24} = \frac{1}{8} \quad (0.5p)$$

och

$$b = 1 - \frac{3}{8} - \frac{5}{24} - \frac{1}{6} = \frac{1}{4} \quad (0.5p)$$

- b)

$$\mathbb{E}(X) = \left(\frac{1}{8} + \frac{1}{6} + \frac{1}{12}\right) + 2\frac{1}{4} + 4\left(\frac{1}{8} + \frac{1}{4}\right) = \frac{19}{8} \quad (0.5p)$$

och

$$\mathbb{E}(Y) = -\left(\frac{1}{8} + \frac{1}{12} + \frac{1}{8}\right) + 2\left(\frac{1}{12} + \frac{1}{8}\right) = \frac{1}{12} \quad (0.5p)$$

och

$$\mathbb{E}(XY) = -\frac{1}{8} - \frac{2}{12} - \frac{4}{8} + \frac{2}{12} + \frac{4}{8} = -\frac{1}{8} \quad (1p)$$

och därmed

$$C(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = -\frac{1}{8} - \frac{19}{8} \frac{1}{12} = -\frac{31}{96} \quad (1p)$$

- c)

$$P(X = 1, Y = -1) = \frac{1}{8} = \frac{3}{8} \cdot \frac{1}{3} = P(X = 1)P(Y = -1)$$

$\Rightarrow$ händelserna $X = 1$ och $Y = -1$ är oberoende.

1p

7. TV-företaget ville få en uppfattning om hur stor andelen felaktiga TV-apparater, som förekommer i ett stort varulager, är. Därför drogs ett slumpmässigt urval omfattande 400 apparater. Av dessa var 24 felaktiga. Bestäm ett 95% konfidensintervall för andel felaktiga TV-apparater. Ledning: Anta en oändlig population. Ger detta intervallet grund för påståendet att mindre än 10% av TV-apparaterna i det stora varulagret är felaktiga?

Lösningförslag:

Antal felaktiga TV-apparater antas $\text{Bin}(400, p)$ -fördelade, där p är den sanna proportionen studenter som fuskar i hela (oändlig) populationen. Skatta p med $\hat{p} = \frac{24}{400} = 0.06$. Eftersom $n\hat{p}(1-\hat{p}) = 22.56 > 10$ använder vi normalapproximationen (0.5p) och får approximativt $100(1-\alpha)\%$ konfidenstervall med $\alpha = 0.05$ och $\lambda_{0.025} = 1.96$ för p :

$$CI = \left[\hat{p} \pm \lambda_{0.025} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right] = \left[0.06 \pm 1.96 \sqrt{\frac{0.06(1-0.06)}{400}} \right] = [0.037, 0.083] \text{ (1.5p)}$$

Det vill säga att konfidenstervall inte innehåller värden ≥ 0.1 , så det ger grund för påståendet att mindre än 10% av TV-apparaterna i det stora varulagret är felaktiga. ($H_0 : p \geq 0.1$ vs. $H_1 : p < 0.1$, H_0 förkastas.) (1p)

8. Ett oberoende forskningsinstitut undersöker om elever får mer poäng på matematiktentor om de arbetar med ett matematiskt datorspel som heter Matika utanför de vanliga skolaktiviteterna. För detta ändamål har man samlat det genomsnittliga antalet poäng på en särskild matematiktentamen (med 100 poäng totalt) från 10 olika, slumpmässigt utvalda skolor där 50% av eleverna studerade med och 50% utan Matika. Tabell 1 visar genomsnittliga antalet poäng med Matika och utan Matika för varje skola. Antag att dessa värden är observationer av parvis beroende slumpvariabler

Skola	1	2	3	4	5	6	7	8	9	10
X	76	59	56	60	61	60	55	58	54	63
Y	68	57	51	57	56	61	53	55	54	59
$\Delta = X - Y$	8	2	5	3	5	-1	2	3	0	4

Tabell 1: Genomsnittliga antalet poäng med Matika (X) och utan Matika (Y) för 10 skolor

$(X_1, Y_1), \dots, (X_{10}, Y_{10})$ sådana att differenserna $Z_i = X_i - Y_i$ är oberoende och $N(\Delta, \sigma)$ -fördelade. Bestäm ett 99% konfidenstervall för skillnaden Δ i genomsnittlig tentamenpoäng. Kan man säga att användning av datorspelet Matika ger högre antal poäng?

Lösningförslag:

Med den nya tabellen och med $\Delta = X - Y$ får man

$$\bar{\Delta} = 3.1 \text{ och } s_{\Delta} = \sqrt{6.767} \approx 2.601 \quad (1.5p)$$

och därmed

$$I_{\Delta} = \bar{\Delta} \mp t(n-1)_{\alpha/2} \frac{s_{\Delta}}{\sqrt{n}} = 3.1 \mp t(9)_{0.005} \frac{2.601}{\sqrt{10}} \quad (1p)$$

och med $t(9)_{0.005} = 3.250$ blir det

$$I_{\Delta} = [0.427, 5.773]. \quad (0.5p)$$

Det vill säga att konfidenstervall inte innehåller noll, så det finns grund för påståendet att användning av datorspelet Matika ger högre antal poäng. ($H_0 : \Delta = 0$ vs. $H_1 : \Delta \neq 0$, H_0 förkastas.) (1p)

9. Antag att man har ett stickprov $(x_1, \dots, x_n)$ av storlek n , d.v.s. x_i är observationer av oberoende och likafördelade stokastiska variabler X_i med väntevärde μ och varians σ^2 . När man skattar variansen σ^2 finns det flera möjliga alternativ, till exempel

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad \text{och} \quad s_\star^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2.$$

Visa minst en skillnad mellan s^2 och $s_\star^2$ (utom att man delar med $n-1$ i ena fallet och med n i det andra). Ledning: Tänk på hur man kan jämföra olika skattare.

2p

Lösningförslag:

- a) s^2 är väntevärdesriktig, men $s_\star^2$ är det inte.

$$\begin{aligned} \mathbb{E}(s^2) &\stackrel{reg.}{=} \mathbb{E} \left(\frac{1}{n-1} \left(\sum_{i=1}^n (X_i^2) - n\bar{X}^2 \right) \right) \\ &\stackrel{reg.}{=} \frac{1}{n-1} \left(\sum_{i=1}^n \underbrace{\mathbb{E}(X_i^2)}_{=V(X_i) + (\mathbb{E}(X_i))^2} - n \underbrace{\mathbb{E}(\bar{X}^2)}_{=V(\bar{X}) + (\mathbb{E}(\bar{X}))^2} \right) \\ &\stackrel{def.}{=} \frac{1}{n-1} \left(\sum_{i=1}^n (V(X_i) + (\mathbb{E}(X_i))^2) - n(V(\bar{X}) + (\mathbb{E}(\bar{X}))^2) \right) \\ &= \frac{1}{n-1} \left(\sum_{i=1}^n (\sigma^2 + \mu^2) - n \left(\frac{\sigma^2}{n} + \mu^2 \right) \right) \\ &= \frac{1}{n-1} (n\sigma^2 + n\mu^2 - \sigma^2 - n\mu^2) \\ &= \frac{1}{n-1} (n\sigma^2 - \sigma^2) \\ &= \frac{n-1}{n-1} \sigma^2 = \sigma^2 \end{aligned}$$

men

$$\mathbb{E}(s_\star^2) = \mathbb{E} \left(\frac{n-1}{n} s^2 \right) = \frac{n-1}{n} \mathbb{E}(s^2) = \frac{n-1}{n} \sigma^2 \neq \sigma^2 \quad (2p)$$

- b) $s_\star^2$ är effektivare än s^2 :

$$V(s_\star^2) = V \left(\frac{n-1}{n} s^2 \right) = \underbrace{\left(\frac{n-1}{n} \right)^2}_{<1} V(s^2) < V(s^2) \quad (2p)$$

Lycka till!