

## Tentamen L9MA30, LGMA30

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

2017-08-21 kl. 14-18

**Examinator och Jour:** Claes Andersson, tel. 0734 031540 alt. 031 772 8294

**Hjälpmedel:** Valfri miniräknare (tomt minne och ingen möjlighet att koppla upp sig till nätet; kontrolleras av examinatoren), (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 2016 adderas. Lösningar skall redovisas till varje uppgift. Uppgifterna är ej ordnade efter svårighetsgrad. Rättningsprotokoll anslås senast tre veckor efter tentamen.

1. Antag att  $A$  och  $B$  är två händelser i något utfallsrum  $\Omega$  sådana att  $P(A|B) > P(A)$ . Visa att  $P(A|B^c) < P(A)$ .

2p

**Lösningsförslag:**

$$P(A|B^c) = \frac{P(A \cap B^c)}{P(B^c)} = \frac{P(A) - P(A \cap B)}{1 - P(B)} < \frac{P(A)(1 - P(B))}{1 - P(B)} = P(A),$$

där olikheten följer av att  $P(A|B) > P(A)$ .

2p

2. Otto spelar på Lotto. Han väljer en rad som består av sex olika nummer mellan 1 och 30. Sedan dras sex nummer mellan 1 och 30 utan återläggning, och varje nummer i dragningen som finns med i Ottos rad ger ett rätt. Vad är sannolikheten att Otto får sitt fjärde rätt på det femte dragna numret?

3p

**Lösningsförslag:**

Vi söker  $P(\text{Femte dragna numret ger Otto hans fjärde rätt})$ . Vi kan skriva detta som

$$P(\text{Fjärde rätt på femte dragna}) = P(\text{Femte rätt} | \text{Tre rätt på fyra första})P(\text{Tre rätt på fyra första})$$

På femte dragningen finns 26 nummer att dra ur och, givet att tre ur Ottos rad redan dragits, är 3 av dessa 26 med i Ottos rad. Alltså är  $P(\text{Femte rätt} | \text{Tre rätt på fyra första}) = 3/26$ . De fyra första numren kan dras på totalt  $\binom{30}{4}$  sätt, och antalet sätt att välja de fyra första så att tre av dem ingår i Ottos rad är  $\binom{6}{3} \binom{24}{1}$ . Alltså är

$$P(\text{Tre rätt på fyra första}) = \frac{\binom{6}{3} \binom{24}{1}}{\binom{30}{4}} = \frac{24 \cdot \binom{6}{3}}{\binom{30}{4}}$$

och svaret ges därför av

$$P(\text{Femte dragna numret ger Otto sitt fjärde rätt}) = \frac{3 \cdot 24 \cdot \binom{6}{3}}{26 \cdot \binom{30}{4}} \approx 9.33 \cdot 10^{-5}$$

3. Ett företag vill komma till rätta med stölder bland personalen. Av denna anledning överväger man att använda en lögn-detektor och helt enkelt screena personalen genom att fråga samtliga om de någon gång stulit från arbetsplatsen. Man tror sig veta att 1% av personalen begår stölder. Lögn-detektorn man använder ger utslag med 99% sannolikhet om den utfrågade ljugar, men även om den utfrågade talar sanning ger den utslag med 5% sannolikhet. Antag att en person som stulit skulle ljuga vid utfrågningen och en person som inte stulit skulle tala sanning. Om en slumpvis vald person frågas ut och lögn-detektorn ger utslag, vad är sannolikheten att den utfrågade ljugar?

3p

#### Lösningsförslag:

Låt  $S$  vara händelsen att den person som frågas ut har stulit och  $U$  händelsen att lögn-detektorn ger utslag. Vi vet då utifrån informationen vi fått att

$$\begin{aligned} P(S) &= 0.01, \\ P(U|S) &= 0.99 \text{ och} \\ P(U|S^c) &= 0.05 \end{aligned}$$

Det vi söker är

$$P(S|U) \stackrel{\text{Bayes}}{=} \frac{P(U|S)P(S)}{P(U)} = \frac{P(U|S)P(S)}{P(U|S)P(S) + P(U|S^c)P(S^c)} = \frac{0.99 \cdot 0.01}{0.99 \cdot 0.01 + 0.05 \cdot 0.99} = \frac{1}{6}$$

4. Låt  $X$  vara en slumpvariabel med täthetsfunktion

$$f_X(x) = \begin{cases} c(x-1)^2, & 0 \leq x \leq 2, \\ 0, & \text{annars} \end{cases}$$

där  $c \in \mathbb{R}$  är en konstant.

- a) Bestäm värdet på  $c$ .

1p

- b) Beräkna väntevärdet,  $\mathbb{E}[X]$ , och variansen,  $V(X)$ .

3p

**Lösningförslag:**

- a) Eftersom  $f(x)$  är en täthetsfunktion måste den uppfylla  $\int_{-\infty}^{\infty} f(x)dx = 1$ . Detta ger

$$1 = c \int_0^2 (x-1)^2 dx = c \left[ \frac{(x-1)^3}{3} \right]_0^2 = \frac{2c}{3} \Rightarrow c = \frac{3}{2}$$

- b) För väntevärdet kan vi notera att tätheten är symmetrisk kring 1, och direkt få att  $E[X] = 1$ . Alternativt beräknar man

$$E[X] = \frac{3}{2} \int_0^2 x(x-1)^2 dx = \frac{3}{2} \int_0^2 (x^3 - 2x^2 + x) dx = \frac{3}{2} \left[ \frac{x^4}{4} - \frac{2x^3}{3} + \frac{x^2}{2} \right]_0^2 = 1$$

För variansen är det för en gångs skull enklast att använda  $V(X) = E[(X - E[X])^2]$ , vilket ger

$$V(X) = E[(X-1)^2] = \frac{3}{2} \int_0^2 (x-1)^2 (x-1)^2 dx = \frac{3}{2} \int_0^2 (x-1)^4 dx = \frac{3}{2} \left[ \frac{(x-1)^5}{5} \right]_0^2 = \frac{3}{5}$$

5. Antag att  $X$  är en kontinuerlig stokastisk variabel som är likformigt fördelad på intervallet  $(-1, 1)$ . Antag vidare att  $Y = X^2$ .

- a) Beräkna  $C(X, Y)$ .

2p

- b) Är  $X$  och  $Y$  oberoende?

1p

**Lösningförslag:**

- a) Vi använder  $C(X, Y) = E[XY] - E[X]E[Y]$ . Eftersom  $X$  är likformigt fördelad på  $(-1, 1)$  är  $E[X] = 0$ , så andra termen är 0. Återstår att beräkna  $E[XY]$ , men eftersom  $Y = X^2$  får vi

$$E[XY] = E[X^3] = \frac{1}{2} \int_{-1}^1 x^3 dx = \frac{1}{2} \left[ \frac{x^4}{4} \right]_{-1}^1 = 0.$$

Således får vi  $C(X, Y) = 0$ .

- b) Nej, eftersom värdet på  $Y$  bestäms helt av värdet på  $X$  är de beroende. Om man vill försäkra sig ytterligare kan man texta villkoret i definitionen av oberoende, alltså att  $X, Y$  är oberoende om  $F_{XY}(x, y) = F_X(x)F_Y(y)$  för alla värden på  $x$  och  $y$ . Om vi väljer  $x = -2/3$  och  $y = 1/4$  får vi att  $F_X(-2/3) = 1/6$  och  $F_Y(1/4) = P(Y \leq 1/4) = P(-1/2 \leq X \leq 1/2) = 1/2$ , så  $F_X(-2/3)F_Y(1/2) = 1/12 > 0$ . Men  $F_{XY}(-2/3, 1/4) = P(X \leq -2/3 \cap Y \leq 1/4) = P(X \leq -2/3 \cap -1/2 \leq X \leq 1/2) = 0$ , eftersom  $X$  inte kan vara både mindre än  $-2/3$  och större än  $-1/2$ .

6. Visa att

$$f_{XY}(x, y) = \begin{cases} x^3 + \frac{xy}{2} & , 0 \leq x \leq 1, 0 \leq y \leq 2 \\ 0 & , \text{annars} \end{cases}$$

är en simultan täthetsfunktion för en två-dimensionell kontinuerlig stokastisk variabel  $(X, Y)$  och beräkna väntevärdet av  $X + Y$ .

4p

### Lösningsförslag:

För att  $f_{XY}$  ska vara en täthetsfunktion måste vi visa att den integrerar till 1.

$$\int_0^1 \int_0^2 \left( x^3 + \frac{xy}{2} \right) dy dx = \int_0^1 \left( 2x^3 + \left[ \frac{xy^2}{2} \right]_0^2 \right) dx = \int_0^1 2x^3 + x dx = \left[ x^4/2 + x^2/2 \right]_0^1 = 1$$

Väntevärdet  $E[X + Y]$  beräknas som

$$\begin{aligned} E[X + Y] &= \int_0^1 \int_0^2 (x + y) \left( x^3 + \frac{xy}{2} \right) dy dx = \int_0^1 \int_0^2 \left( x^4 + \frac{x^2 y}{2} + x^3 y + \frac{xy^2}{2} \right) dy dx \\ &= \int_0^1 \left( 2x^4 + \frac{x^2}{2} \left[ \frac{y^2}{2} \right]_0^2 + x^3 \left[ \frac{y^3}{3} \right]_0^2 + \frac{x}{2} \left[ \frac{y^3}{3} \right]_0^2 \right) dx \\ &= \int_0^1 \left( 2x^4 + x^2 + 2x^3 + \frac{4}{3}x \right) dx \\ &= \left[ \frac{2}{5}x^5 + \frac{x^3}{3} + \frac{x^4}{2} + \frac{2}{3}x^2 \right]_0^1 = \frac{19}{10} \end{aligned}$$

7. Elaine köper frozen yoghurt (froyo) som påstås innehålla 10 g fett per 100 g froyo. Av olika anledningar misstänker Elaine att detta inte stämmer. Mängden fett,  $X$ , mäts i 21 prover av 100 g froyo. Värdena kan antas vara normalfördelade,  $\mathcal{N}(\mu, \sigma)$ , och observationerna kan sammanfattas med följande summor:

$$\sum x_i = 238.1694, \quad \sum x_i^2 = 2939.242$$

- a) Bilda ett 95% konfidensintervall för värdet på  $\mu$ .

3p

- b) Kan nollhypotesen  $\mu = 10$  förkastas? Vilken signifikansnivå har testet?

1p

**Lösningsförslag:**

- a) Konfidensintervallet ges av

$$\bar{x} \pm t_{n-1, 0.975} \frac{s}{\sqrt{n}}$$

där  $n = 20$ ,  $\bar{x} = \frac{1}{n} \sum x_i = 11.34$ ,  $s = \sqrt{\frac{1}{n-1} (\sum x_i^2 - n\bar{x}^2)} = 3.45$  och  $t_{20, 0.975} = 2.086$  vilket ger intervallet  $(9.77, 12.91)$ .

- b) Nej, eftersom konfidensintervallet innehåller 10 kan nollhypotesen inte förkastas. Signifikansnivån är 5%.

8. NBA-slutspelet 2017 vanns av Golden State Warriors. Deras kanske största stjärna är Stephen Curry, främst känd för sitt trepoängsskytte. I slutspelet lyckades han med 72 av 172 trepoängsförsök.

- a) Beräkna ett 90% konfidensintervall för sannolikheten att Stephen Curry lyckas med ett trepoängsförsök baserat på datan från slutspelet.

3p

- b) Hur många trepoängsförsök hade man behövt observera för att få ett 90% konfidensintervall av längd högst 0.06 (som alltså sträcker sig högst 0.03 åt varje håll från punktskattningen)?

2p

### Lösningförslag:

- a) Konfidensintervallet ges av

$$\hat{p} \pm \lambda_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}},$$

där  $\hat{p} = 72/172 = 0.419$ ,  $\alpha = 0.1$ ,  $\lambda_{0.05} = 1.645$ ,  $n = 172$ . Detta ger intervallet

$$0.419 \pm 0.0619 = (0.357, 0.480)$$

- b) För att konfidensintervallet ska vara av längd högst 0.06 måste vi ha att

$$\lambda_{0.05} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq 0.03 \Rightarrow n \geq \lambda_{0.05}^2 \frac{\hat{p}(1-\hat{p})}{0.03^2} = 731.7,$$

alltså behövs minst 732 observationer. Detta är under antagandet att det observerade  $\hat{p}$  blir detsamma som vi fick från datan, om man vill vara helt på den säkra sidan kan man ersätta  $\hat{p}$  med det värde som maximerar uttrycket som  $n$  måste överskrida. Det fås genom  $\hat{p} = 0.5$ , och skulle ge att man behöver 752 observationer.

9. Antag att vi har två stickprov av storlek  $n_X$  respektive  $n_Y$ , med stickprovsvarians  $s_X^2$  respektive  $s_Y^2$ . Om man antar att de två stickproven kommer från fördelningar med samma varians  $\sigma^2$ , så skattar man  $\sigma^2$  med så kallad pooled variance

$$s_p^2 = \frac{(n_X - 1)s_X^2 + (n_Y - 1)s_Y^2}{n_X + n_Y - 2}.$$

Visa att  $s_p^2$  är väntevärdesriktig.

2p

### Lösningförslag:

Vi börjar med att konstatera att  $E[s_X^2] = E[s_Y^2] = \sigma^2$  eftersom

$$\begin{aligned} E[s_X^2] &= E \left[ \frac{1}{n_X - 1} \left( \sum X_i^2 - n_X \bar{X}^2 \right) \right] = \frac{1}{n_X - 1} E \left[ E[X_i^2] - n_X E[\bar{X}^2] \right] = \\ &= \frac{1}{n_X - 1} \left( \sum_{i=1}^{n_X} (\sigma^2 + \mu_X^2) - n_X \left( \frac{\sigma^2}{n_X} + \mu_X^2 \right) \right) \\ &= \frac{1}{n_X - 1} (n_X \sigma^2 + n_X \mu_X^2 - \sigma^2 - n_X \mu_X^2) = \sigma^2 \end{aligned}$$

Med detta etablerat kan vi betrakta väntevärdet av  $s_p^2$ :

$$\begin{aligned} E[s_p^2] &= E\left[\frac{(n_X - 1)s_X^2 + (n_Y - 1)s_Y^2}{n_X + n_Y - 2}\right] = \frac{(n_X - 1)E[s_X^2] + (n_Y - 1)E[s_Y^2]}{n_X + n_Y - 2} = \\ &= \frac{(n_X - 1)\sigma^2 + (n_Y - 1)s_Y^2}{n_X + n_Y - 2} = \sigma^2, \end{aligned}$$

och alltså är  $s_p^2$  väntevärdesriktig.

Lycka till!