

LÖSNINGAR till tentamen: Statistik och sannolikhetslära (LMA120)

Tid och plats: Tisdagen den 25 augusti, 8.30 - 12.30

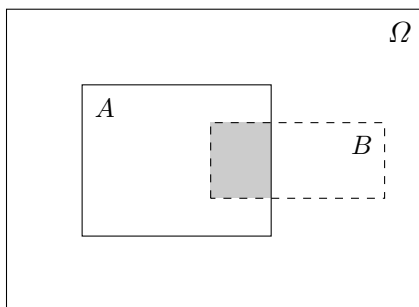
Hjälpmedel: Typgodkänd miniräknare, formelsamling, kursbok samt anteckningar

Betygsgränser: 3: 12 poäng, 4: 18 poäng, 5: 24 poäng. Maximalt antal poäng: 30

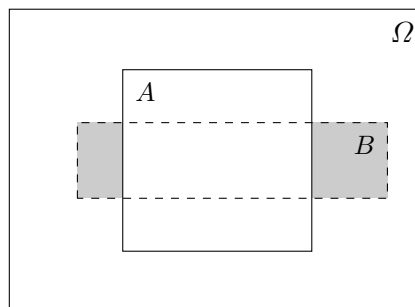
Examinator: Malin Palö

Telefonvakt: Petter Mostad (telefonnummer: 0768 30 36 35)

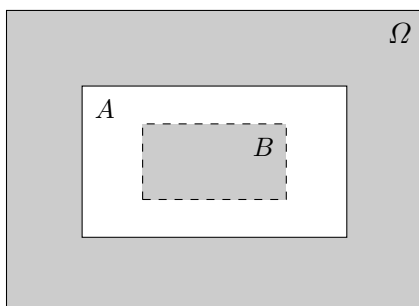
1. I respektive bild nedan, skugga det angivna området. (1+1+1+1 poäng)



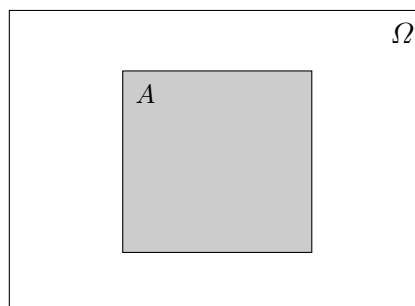
(a) $A \cap B$



(b) $A^c \cap B$



(a) $A^c \cup B$



(b) $\Omega^c \cup A$

2. I kortspelet poker består en hand av fem kort. Om fyra av korten har samma valör (exempelvis sju) så säger man att man har fyrtal. Vad är sannolikheten att få fyrtal om man slumpvis väljer fem kort ur en vanlig kortlek? (4 poäng)

Vi beräknar först antalet sätt vi kan få fyrtal på. Vi kan välja valören som de fyra korten med samma valör ska ha på totalt 13 olika sätt. När vi har valt denna valör kan vi välja det sista kortet på $52 - 4 = 48$ olika sätt. Enligt multiplikationsprincipen kan vi därför få fyrtal på totalt $13 \cdot 48$ olika sätt. Totalt finns det $\binom{52}{5}$ olika pokerhänder, och sannolikheten för att få fyrtal blir därför

$$\frac{13 \cdot 48}{\binom{52}{5}} = \frac{13 \cdot 48}{\frac{52 \cdot 51 \cdot 50 \cdot 49 \cdot 48}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5}} = \frac{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 13}{52 \cdot 51 \cdot 50 \cdot 49} = \frac{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 13}{(4 \cdot 13) \cdot (3 \cdot 17) \cdot (2 \cdot 5 \cdot 5) \cdot (7 \cdot 7)} = \frac{1}{5 \cdot 7^2 \cdot 17}$$

3. Ett hemmatest kan upptäcka laktosintolerans hos en person som är intolerant mot laktos med 94% sannolikhet. Med sannolikhet 0.01 ger testet ett positivt resultat även om den som gör testet inte är laktosintolerant. Anta att 5% av befolkningen är laktosintoleranta. Vad är sannolikheten att en slumpvist vald person som får ett positivt testresultat är laktosintolerant? (4 poäng)

Låt L vara händelsen att en slumpvist vald person är laktosintolerant och låt T vara händelsen att samma person testas positivt för laktosintolerans med hemmatestet. Enligt uppgiften är

då $P(T | L) = 0.94$, $P(T | L^c) = 0.01$ och $P(L) = 0.05$. Vi behöver räkna ut vad $P(L | T)$ är. Enligt Bayes sats är

$$P(L | T) = \frac{P(T | L) \cdot P(L)}{P(T)}$$

Det enda vi saknar för att kunna använda denna är $P(T)$. Vi använder därför omskrivningen

$$\begin{aligned} P(T) &= P(T \cap L) + P(T \cap L^c) \\ &= P(T | L) \cdot P(L) + P(T | L^c) \cdot P(L^c) \\ &= P(T | L) \cdot P(L) + P(T | L^c) \cdot (1 - P(L)) \\ &= 0.94 \cdot 0.05 + 0.01 \cdot (1 - 0.05) = 0.0565 \end{aligned}$$

Vi får därför

$$P(L | T) = \frac{P(T | L) \cdot P(L)}{P(T)} = \frac{0.94 \cdot 0.05}{0.0565} = 0.8319$$

Dvs. ungefär 83% av de som testar positivt för laktosintolerant är intoleranta mot laktos.

4. Antag att tiden det tar för en kund att betala för sina matvaror i en affär är exponentialfördelad med medeltid två minuter. Antag att det står tre personer före dig i kön. Vad är sannolikheten att du får vänta i mer än fem minuter innan det blir din tur att betala? (4 poäng)

Eftersom att det är tre personer före dig i kön blir medeltiden innan det är din tur att betala $3 \cdot 2 = 6$ minuter. Eftersom att tiden för varje enskild kund att bli klar är exponentialfördelad, så är även den totala tiden T för de tre personerna som står före dig i kön exponentialfördelad. Eftersom att $\mathbb{E}[T] = 6$ så får vi alltså att $T \sim \exp(1/6)$. Eftersom att fördelningsfunktionen av en exponentialfördelad variabel med parameter λ ges av

$$F(x) = \begin{cases} 0 & \text{då } x \leq 0 \\ 1 - e^{-\lambda x} & \text{då } x > 0 \end{cases}$$

får vi att

$$P(T > 5) = 1 - P(T \leq 5) = 1 - F(5) = 1 - (1 - e^{-1/6 \cdot 5}) = e^{-5/6} = 0.435$$

dvs. sannolikheten att vi får vänta i minst fem minuter är ungefär 44%.

5. Låt X vara en kontinuerlig stokastisk variabel med täthetsfunktion

$$f(x) = \begin{cases} 0 & \text{om } x < 1 \\ x - 1 & \text{om } 1 \leq x < 2 \\ 1/2 & \text{om } 2 \leq x < 3 \\ 0 & \text{om } x \geq 3 \end{cases}$$

(a) Vad är $P(X < 0)$?

(b) Vad är $\mathbb{E}[X]$?

(2+2 poäng)

(a) Eftersom att $f(x) = 0$ då $x < 0$ så är $P(X < 0) = 0$.

(b) Enligt definitionen av förväntat värde är

$$\begin{aligned} \mathbb{E}[X] &= \int_{-\infty}^{\infty} x f(x) dx = \int_1^3 x f(x) dx = \int_1^2 x(x-1) dx + \int_2^3 x \cdot \frac{1}{2} dx \\ &= \left[\frac{x^3}{3} - \frac{x^2}{2} \right]_1^2 + \left[\frac{x^2}{2} \cdot \frac{1}{2} \right]_2^3 = \left(\frac{8}{3} - \frac{4}{2} \right) - \left(\frac{1}{3} - \frac{1}{2} \right) + \frac{9}{4} - \frac{4}{4} \\ &= \frac{25}{12} \end{aligned}$$

6. Antag att man vid tillverkningen av kullager får diametrar med väntevärdet 47 mm och standardavvikelsen 0.3 mm. Uppskatta sannolikheten att den genomsnittliga diametern vid 20 slumpmässigt utvalda mätningar överstiger 47.1 mm.

(4 poäng)

Låt X_i vara diametern på det i :te kullagret man mäter och $X = (X_1 + \dots + X_{20})/20$ så att X är den genomsnittliga diametern. Enligt uppgiften är $\mathbb{E}[X_i] = 47$ och $\sigma(X_i) = 0.3$ för $i = 1, \dots, 20$, så

$$\mathbb{E}[X] = \mathbb{E}\left[\frac{X_1 + \dots + X_{20}}{20}\right] = \frac{\mathbb{E}[X_1 + \dots + X_{20}]}{20} = \frac{\mathbb{E}[X_1] + \dots + \mathbb{E}[X_{20}]}{20} = \frac{20 \cdot 47}{20} = 47$$

och

$$\begin{aligned} \text{Var}(X) &= \text{Var}\left(\frac{X_1 + \dots + X_{20}}{20}\right) = \frac{\text{Var}(X_1 + \dots + X_{20})}{20^2} = \frac{\text{Var}(X_1) + \dots + \text{Var}(X_{20})}{20^2} \\ &= \frac{20 \cdot 0.3^2}{20^2} = 0.0045 \end{aligned}$$

Vi använder nu den centrala gränsvärdessatsen, och får

$$\begin{aligned} P(\bar{X} > 47.1) &= P\left(\frac{\bar{X} - 47}{\sqrt{0.0045}} > \frac{47.1 - 47}{\sqrt{0.0045}}\right) \approx 1 - \Phi\left(\frac{47.1 - 47}{\sqrt{0.0045}}\right) \\ &\approx 1 - \Phi(1.4907) \\ &= 1 - 0.695 \\ &= 0.305 \end{aligned}$$

dvs. sannolikheten att den genomsnittliga diametern vid 20 slumpmässigt utvalda mätningar överstiger 47.1 mm är ungefär 30.5%.

7. För att ta reda på en produkts fryspunkt gjordes följande mätningar: $x_1 = 1.1^\circ \text{ C}$, $x_2 = 1.6^\circ \text{ C}$, $x_3 = 2.2^\circ \text{ C}$, $x_4 = 0.6^\circ \text{ C}$, $x_5 = 1.5^\circ \text{ C}$ och $x_6 = 1.7^\circ \text{ C}$.

- Beräkna ett 98%-igt dubbelsidigt konfidensintervall för fryspunktens medelvärde, om standardavvikelsen $\sigma = 0.2$ man får vid mätningar är känd sedan innan.
- Beräkna ett 98%-igt dubbelsidigt konfidensintervall för fryspunktens medelvärde, om standardavvikelsen är okänd och måste skattas från datan. (3+3 poäng)

(a) Vi räknar först ut att stickprovsmedelvärdet är

$$\bar{x} = \frac{1.1 + 1.6 + 2.2 + 0.6 + 1.5 + 1.7}{6} = 1.45$$

Konfidensintervallet blir därför

$$\left[1.45 - 2.33 \cdot \frac{0.2}{\sqrt{6}}, 1.45 + 2.33 \cdot \frac{0.2}{\sqrt{6}}\right] = [1.26, 1.64]$$

(b) Vi börjar med att räkna ut stickprovsstandardavvikelsen s ;

$$\begin{aligned} s &= \sqrt{\frac{\sum_{i=1}^6 (x_i - \bar{x})^2}{n - 1}} \\ &= \sqrt{\frac{(1.1 - 1.45)^2 + (1.6 - 1.45)^2 + (2.2 - 1.45)^2 + (0.6 - 1.45)^2 + (1.5 - 1.45)^2 + (1.7 - 1.45)^2}{6 - 1}} \\ &= 0.55 \end{aligned}$$

Eftersom att vi har $6 - 1 = 5$ frihetsgrader (dvs. $df = 5$) och $\alpha = 1 - 0.98 = 0.02$ får vi konfidensintervallet

$$\left[1.45 - 3.36 \cdot \frac{0.55}{\sqrt{6}}, 1.45 + 3.36 \cdot \frac{0.55}{\sqrt{6}} \right] = [1.18, 1.72]$$