

LÖSNINGAR TILL TENTAMEN: Statistik och sannolikhetslära (LMA120)**Tid:** Onsdagen den 15 april 2015, 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:** Petter Mostad**Telefonvakt:** Malin Palö Forsström (telefonnummer: 0768 30 36 35)

1. Arbetsdagen är totalt åtta timmar lång, och Håkan kan påbörja besiktningen vid varje hel timma förutom när det bara är en timma kvar av arbetsdagen, dvs. han kan schemalägga dessa två timmar på totalt 7 olika sätt. Håkan har sedan 6 en-timmas-pass kvar av arbetsdagen, varför han kan placera ut tiden han lägger på beräkningar på totalt $\binom{6}{3}$ olika sätt. Resten av dagen måste han sedan lägga på möten. Enligt multiplikationsprincipen finns det därför totalt

$$7 \cdot \binom{6}{3} = 7 \cdot \frac{6 \cdot 5 \cdot 4}{1 \cdot 2 \cdot 3} = 7 \cdot 20 = 140$$

olika sätt Håkan kan planera sin dag på.

2. Låt X_1 vara livslängden för det första batteriet, X_2 vara livslängden för det andra batteriet, osv. Då är

$$X_1 \sim N(2, 1)$$

$$X_1 + X_2 \sim N(2 + 2, \sqrt{1^2 + 1^2}) = N(2 \cdot 2, \sqrt{2})$$

$$X_1 + X_2 + X_3 \sim N(2 + 2 + 2, \sqrt{1^2 + 1^2 + 1^2}) = N(3 \cdot 2, \sqrt{3})$$

och generellt gäller att

$$X_1 + X_2 + \dots + X_k \sim N(2k, \sqrt{k})$$

osv. Vi får därför att

$$\begin{aligned} P(X_1 \leq 10) &< P(X_1 + X_2 \leq 10) \\ &< P(X_1 + X_2 + X_3 \leq 10) \\ &< P(X_1 + X_2 + X_3 + X_4 \leq 10) \\ &< P(X_1 + X_2 + X_3 + X_4 + X_5 \leq 10) = \Phi\left(\frac{10 - 10}{\sqrt{5}}\right) = \Phi(0) = 0.5 \end{aligned}$$

$$P(X_1 + X_2 + X_3 + X_4 + X_5 + X_6 \leq 10) = \Phi\left(\frac{10 - 12}{\sqrt{6}}\right) = \Phi\left(\frac{-2}{\sqrt{6}}\right) = 0.207$$

$$P(X_1 + X_2 + X_3 + X_4 + X_5 + X_6 + X_7 \leq 10) = \Phi\left(\frac{10 - 14}{\sqrt{7}}\right) = \Phi\left(\frac{-4}{\sqrt{7}}\right) = 0.0653$$

$$P(X_1 + X_2 + X_3 + X_4 + X_5 + X_6 + X_7 + X_8 \leq 10) = \Phi\left(\frac{10 - 16}{\sqrt{8}}\right) = \Phi\left(\frac{-6}{\sqrt{8}}\right) = 0.0169$$

Inköparen måste alltså beställa minst åtta batterier.

3. (a) Eftersom att $n = 24$ är ganska stort, så kan vi använda en normalfördelning istället för en χ^2 -fördelning när vi räknar ut konfidensintervallet, trots att vi inte känner till standardavvikelsen. Ett 95%-igt konfidensintervall ges då av

$$[\bar{x} - 1.96s/\sqrt{n}, \bar{x} + 1.96s/\sqrt{n}],$$

där

$$\bar{x} = \frac{\sum x_i}{24} = 28.38$$

och

$$s = \sqrt{\frac{\sum x_i^2 - (\sum x_i)^2/n}{n-1}} = 1.334.$$

Konfidensintervallet blir alltså

$$[27.85, 28.92].$$

- (b) Nej, eftersom att konfidensintervallet inte innehåller talet 29.
(c) Ja! Ju högre konfidensgrad vi väljer desto bredare blir konfidensintervallet. Om vi räknar ut ett dubbelsidigt konfidensintervall med konfidensgrad p så ges konfidensintervallets övre gräns av

$$\bar{x} + \Phi^{-1}(1 - (1 - p)/2) \cdot s/\sqrt{n}.$$

Vi behöver därför lösa ekvationen

$$\bar{x} + \Phi^{-1}(1 - (1 - p)/2) \cdot s/\sqrt{n} = 29, .$$

Om vi löser ut p så ser vi att gränsen för vilken minsta konfidensgrad vi behöver ta för att vårt konfidensintervall ska innehålla talet 29 är

$$p = 1 - 2 \left(1 - \Phi \left((29 - \bar{x}) \cdot \sqrt{n}/s \right) \right) = 2\Phi \left((29 - \bar{x}) \cdot \sqrt{n}/s \right) - 1 = 0.976.$$

Ett 98%-igt konfidensintervall hade alltså innehållit talet 29, och i det fallet hade företaget behövt sätta upp bullerplank.

4. (a) Låt X_A och X_B vara längden i mm på en slumpvist vald skruv från företag A respektive B. Då är $X_A \sim N(10, 1)$ och $X_B \sim N(9, 2)$. Sannolikheten p_A för att en skruv från företag A kan användas är

$$\begin{aligned} P(9.5 \leq X_A \leq 10.5) &= P(X_A \leq 10.5) - P(X_A \leq 9.5) \\ &= \Phi \left(\frac{10.5 - 10}{1} \right) - \Phi \left(\frac{9.5 - 10}{1} \right) \\ &= 0.3829 \end{aligned}$$

och sannolikheten p_B att en skruv från företag B kan användas är

$$\begin{aligned} P(9.5 \leq X_B \leq 10.5) &= P(X_B \leq 10.5) - P(X_B \leq 9.5) \\ &= \Phi \left(\frac{10.5 - 9}{2} \right) - \Phi \left(\frac{9.5 - 9}{2} \right) \\ &= 0.1747 \end{aligned}$$

- (b) Låt Y_A och Y_B vara antalet skruvar från företag A respektive B man måste testa innan man hittar en som har rätt längd. Då är Y_A och Y_B geometriskt fördelade med parameter p_A respektive p_B , så i snitt måste man testa $1/p_A$ respektive $1/p_B$ skruvar. Utslaget över de skruvar som går att använda blir i längden kostnaden per skruv därför

$$50 \cdot \mathbb{E}[X_A] = 50/p_A = 130.6 \text{ öre}$$

om man köper skruvar från företag A och

$$50 \cdot \mathbb{E}[X_B] = 20/p_B = 114.5 \text{ öre}$$

om man köper skruvar från företag B . I längden blir det alltså billigast att köpa skruvar från företag B .

- (c) För varje använd skruv har vi i snitt behövt kassera

$$\mathbb{E}[Y_A] - 1 = 1/p_A - 1 = 1.6116$$

skruvar om vi använder skruvar från företag A , och

$$\mathbb{E}[Y_B] - 1 = 1/p_B - 1 = 4.7241$$

skruvar om vi använder skruvar från företag B . De nya snittkostnaderna per användbar skruv blir därför

$$50 \cdot \mathbb{E}[Y_A] + 6 \cdot (\mathbb{E}[Y_A] - 1) = 140.3 \text{ öre}$$

om vi använder skruvar från företag A , och

$$20 \cdot \mathbb{E}[Y_B] + 6 \cdot (\mathbb{E}[Y_B] - 1) = 142.8 \text{ öre}$$

om vi använder skruvar från företag B , dvs. nu blir företag A billigast att använda i längden.

5. (a) Om tiden i minuter mellan två bilar som passerar är exponentialfördelad med parameter λ , så antalet bilar som passerar under en minut Poissonfördelad med samma parameter λ . Eftersom att medelvärdet av en Poissonfördelning är dess parameter λ , så passerar det i snitt 5 bilar per minut.
- (b) Vi ska välja μ så att $\mathbb{E}[X] = \lambda^{-1} = 1/5 = 0.2$.

$$\mathbb{E}[X] = \mathbb{E}[0.1 + Y] = 0.1 + \mathbb{E}[Y] = 0.1 + \mu^{-1}$$

så vi får ekvationen

$$0.2 = 0.1 + \mu^{-1} \Leftrightarrow \mu^{-1} = 0.1 \Leftrightarrow \mu = 10.$$

6. (a) Stickprovsmedelvärdet är

$$\bar{x} = \frac{7 + 4 + 4 + 6 + 3 + 8 + 10 + 10 + 8 + 5 + 4 + 9}{12} = 6.5,$$

stickprovsstandardavvikelsen är

$$s = \sqrt{\frac{\sum_i (x_i - \bar{x})^2}{n - 1}} = 2.5045$$

och stickprovsvariansen är

$$s^2 = \frac{\sum_i (x_i - \bar{x})^2}{n - 1} = 6.2727$$

- (b) Om X är Poissonfördelad så är $\mathbb{E}[X] = \lambda$, så $\bar{x}$ skulle kunna vara en bra skattare till λ , dvs. vi borde ha att $\lambda \approx \bar{x} = 6.5$. För att visa att $\bar{x}$ är en bra skattare kontrollerar vi att den är väntevärdesriktig;

$$\begin{aligned}\mathbb{E}\left[\frac{X_1 + X_2 + \dots + X_n}{n}\right] &= \frac{\mathbb{E}[X_1 + X_2 + \dots + X_n]}{n} \\ &= \frac{\mathbb{E}[X_1] + \mathbb{E}[X_2] + \dots + \mathbb{E}[X_n]}{n} \\ &= \frac{\lambda + \lambda + \dots + \lambda}{n} \\ &= \frac{n\lambda}{n} \\ &= \lambda\end{aligned}$$

dvs. $\bar{x}$ är en väntevärdesriktig skattare för λ .

- (c) Det finns ett antal saker som indikerar att man kanske inte ska lite alltför mycket på skattningen av parametern λ , exempelvis:
- Stickprovet är ganska litet
 - Variansen för en Poissonfördelad variabel är λ , och $6.5 \neq 6.27$ som var stickprovsvariansen.
 - Om vi ritar ett stapeldiagram för stickprovet, så det inte ut som grafen för $P(X = x)$ då $X \sim \text{Poisson}(6.5)$.