

TENTAMEN: Statistik och sannolikhetslära (LMA120)

Tid och plats: Onsdagen den 22 januari, MVH12, 10.00 - 12.00

Hjälpmedel: Typgodkänd miniräknare, formelsamling, kursbok samt anteckningar

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

1. För vilken av följande fördelningar är $P(X = 3)$ störst?

- En normalfördelning med medelvärde 3 och standardavvikelse 1
- En normalfördelning med medelvärde 3 och standardavvikelse 0.1
- En exponentialfördelning, med parameter $1/2$
- En binomialfördelning, med $n = 4$ och $p = 0.5$
- En Poissonfördelning, med $\lambda = 4$ (8 poäng)

Eftersom att X i de tre första alternativen är en kontinuerlig stokastisk variabel ser vi direkt att $P(X = 3) = 0$ i dessa fall.

Om $X \sim \text{bin}(4, 0.5)$ så är

$$P(X = 3) = \binom{4}{3} \cdot 0.5^3 \cdot (1 - 0.5)^{4-3} = 4 \cdot 0.5^4 = 0.25$$

Om $X \sim \text{Poisson}(4)$ så är

$$P(X = 3) = \frac{e^{-4} \cdot 4^3}{3!} = \frac{e^{-4} \cdot 4 \cdot 4 \cdot 4}{3 \cdot 2} = \frac{4 \cdot 4 \cdot 2}{3 \cdot e^4} \approx 0.19$$

Alltså är $P(X = 3)$ störst om X är binomialfördelad med $n = 4$ och $p = 0.5$.

2. Malin har som sommarjobb att räkna antalet bilar som passerar en bro under två timmar varje dag. Tiden i minuter mellan två bilar som passerar antas vara exponentialfördelad med parameter $\lambda = 1/20$.

- (a) Hur länge måste Malin i snitt vänta mellan att två bilar dyker upp?
- (b) Vad är sannolikheten att det dyker upp ingen, en eller två bilar under ett arbetspass?
- (c) Vad är sannolikheten att inga bilar dyker upp under två dagar i följd?
- (d) När det är 5 minuter kvar av Malins arbetspass, och det inte har dykt upp några bilar den senaste halvtimmen bestämmer sig Malin för att gå hem. Vad är sannolikheten för att det dyker upp en bil under de fem minuter av Malins arbetspass som hon inte är där? (10 poäng)

(a) Om X_1 är tiden mellan två bilar så är $X_1 \sim \exp \lambda_1$, där $\lambda_1 = 1/20$. Eftersom att X_1 är exponentialfördelad så är

$$\mathbb{E}[X_1] = 1/\lambda_1 = 20$$

dvs. i snitt får Malin vänta i 20 minuter mellan två bilar.

- (b) Om X_2 är antalet bilar som passerar under ett arbetspass, så är X_2 Poisson-fördelad med parameter $\lambda_2 = 120 \cdot \lambda_1 = 120/20 = 6$, eftersom att ett arbetspass varar i 120 minuter och det i snitt passerar en bil var 20:e minut enligt (a). Vi får därför att

$$\begin{aligned} P(X_2 \leq 2) &= P(X_2 = 0) + P(X_2 = 1) + P(X_2 = 2) \\ &= \frac{e^{-\lambda_2} \lambda_2^0}{0!} + \frac{e^{-\lambda_2} \lambda_2^1}{1!} + \frac{e^{-\lambda_2} \lambda_2^2}{2!} \\ &= \frac{e^{-6} \cdot 6^0}{1} + \frac{e^{-6} \cdot 6^1}{1} + \frac{e^{-6} \cdot 6^2}{2} \\ &= \frac{e^{-6}}{1} + \frac{e^{-6} \cdot 6}{1} + \frac{e^{-6} \cdot 36}{2} \\ &= 25 \cdot e^{-6} \approx 0.062 \end{aligned}$$

dvs. sannolikheten för att det passerar max 2 bilar under ett arbetspass är ungefär 6%.

- (c) Sannolikheten att ingen bil dyker upp under en dag är $P(X_2 = 0) = \frac{e^{-6} \cdot 6^0}{0!} = e^{-6}$. Eftersom att antalet bilar som passerar under dag 2 kan antas vara oberoende av antalet bilar som passerade under dag 1, så blir sannolikheten att inga bilar passerar under varken dag 1 eller dag 2

$$P(X_2 = 0) \cdot P(X_2 = 0) = e^{-6} \cdot e^{-6} = e^{-12} \approx 6 \cdot 10^{-6} \approx 0$$

- (d) Antalet bilar X_3 som dyker upp under de sista fem minuterna är helt oberoende av antalet bilar som passerat tidigare under dagen, eftersom att exponentialfördelade och Poissonfördelade stokastiska variabler saknar minne. Därför är $X_3 \sim \text{Poisson}(\lambda_3)$, där $\lambda_3 = 5/20 = 0.25$, och sannolikheten att inga iblar passerar under de sista fem minuterna blir

$$P(X_3 = 0) = \frac{e^{-\lambda_3} \cdot \lambda_3^0}{0!} = e^{-0.25} \approx 0.7788$$

3. Johan tycker mycket om pannkakor, och när han gör pannkakor brukar han köpa medelstora ägg, som enligt förpackningen ska väga 53–63 gram. Den senaste tiden tycker dock Johan att alla ägg han köpt varit förvånansvärt små, och han bestämmer sig därför för att väga alla ägg han köper. De senaste 20 äggen Johan har köpt har haft följande vikter:

46	47	47	48	49	49	49	50	52	52
52	52	53	53	55	55	56	56	57	59

Johan räknar ut att $\sum v_i = 1037$ och $\sum v_i^2 = 54\,027$, där v_i är vikten på ägg nummer i .

- (a) Givet Johans insamlade data, beräkna ett ensidigt uppåt begränsat 95%-igt konfidensintervall för medelvikten på ett ägg. Baserat på detta, finns det skäl för Johan att kontakta äggproducenten för att klaga?

- (b) Johan läser på äggproducentens hemsida att äggen som klassas som normalstora har en vikt som är normalfördelad med medelvärde 58 gram och standardavvikelse 2,5 gram. Givet att de här siffrorna stämmer, hur sannolikt är det att få ett ägg som väger mindre än eller lika med 53 gram?
- (c) Hur sannolikt är det, givet att äggproducentens data, att stickprovsmedelvärdet är så litet som, eller mindre än, i Johans fall?
- (d) Totalt väger 19 av äggen Johan vägt minde än 58 gram. Givet att producentens uppgifter stämmer, hur stor är sannolikheten att minst 19 av 20 ägg väger mindre än 58 gram? (12 poäng)

- (a) Vi har stickprovsstorlek $n = 20$, stickprovsmedelvärde

$$\bar{x} = \sum v_i/n = 1037/20 = 51.85$$

och stickprovsvarians

$$s^2 = \frac{\sum v_i^2 - n(\bar{x})^2}{n-1} = \frac{54\,027 - 20 \cdot 51.85^2}{20-1} = 13.61$$

Eftersom att antalet frihetsgrader är $n-1 = 19$ och $\alpha = 0.05$ så får vi $z = 1.7291$. Detta ger konfidensintervallet

$$\left(-\infty, \bar{x} + \frac{z s}{\sqrt{n}}\right) = (-\infty, 53.2764)$$

- (b) Om X är vikten hos ett slumpvist valt ägg, så är

$$P(X \leq 53) = \Phi\left(\frac{53-58}{2.5}\right) = \Phi(-2) = 0.02$$

dvs. sannolikheten att få ett ägg som väger mindre än 53 gram är enligt äggproducenten mindre än 2%.

- (c) I ett stickprov om 20 ägg, låt X_i vara vikten på ägg nummer i . Då är enligt äggproducenten $X_i \sim N(58, 2.5)$, så

$$\mathbb{E}[\bar{X}] = \mathbb{E}\left[\frac{X_1 + \dots + X_{20}}{20}\right] = \frac{\mathbb{E}[X_1] + \dots + \mathbb{E}[X_{20}]}{20} = \frac{20 \cdot 58}{20} = 58$$

och

$$\text{Var}(\bar{X}) = \text{Var}\left(\frac{X_1 + \dots + X_{20}}{20}\right) = \frac{\text{Var}(X_1) + \dots + \text{Var}(X_{20})}{20^2} = \frac{2.5^2}{20}$$

så $\bar{X} \sim N\left(58, \frac{2.5}{\sqrt{20}}\right)$. Vi får därför att

$$P(\bar{X} \leq 51.85) = \Phi\left(\frac{51.85 - 58}{2.5/\sqrt{20}}\right) \approx \Phi(-11) \approx 10^{-28}$$

(d) Låt X vara antalet ägg som väger mer än 58 gram. Då är $X \sim \text{bin}(20, 0.5)$ eftersom att normalfördelningen är symmetrisk. Vi får därför att

$$\begin{aligned} P(\text{minst 19 ägg väger mindre än 58 gram}) &= P(X \leq 1) \\ &= \binom{20}{0} \cdot 0.5^0 \cdot 0.5^{20-0} + \binom{20}{1} \cdot 0.5^1 \cdot 0.5^{20-1} \\ &= 21 \cdot 0.5^{20} \approx 2 \cdot 10^{-5} \end{aligned}$$