Linköpings universitet
Institutionen för teknik och naturvetenskap
Michael Hörnquist

Statistik och sannolikhetslära, TNIU66

Tentamen, 26 mars 2021 em, Svar och lösningsskisser

1. (a) Direkt avläsning i respektive lådagram ger för A: max=6,93, min=−2,96, median=2,32 och medel=2,17 samt för B: max=6,91, min=−1,61, median=2,57 och medel=2,62. (b) Histogram 1 visar att värdena är mer koncentrerade kring medelvärdet, mindre spridning, vilket passar ihop med lådagram för datamängd B.

2. (a) Låt G vara händelsen att ett frö gror. Då blir enligt satsen om total sannolikhet $\Pr(G) = \Pr(G|A)\Pr(A) + \Pr(G|B)\Pr(B)$. Låt sökt sannolikhet $\Pr(G|B) = x$ och stoppa in givna värden: $0,80 = 0,70 \cdot 0,35 + x \cdot (1 - 0,35)$. Ekvationslösning ger $x = \Pr(G|B) = 0,8538 \approx 85\%$. (b) Betingad sannolikhet ger $\Pr(B|G) = \Pr(G|B)\Pr(B)/\Pr(G) = 0,8538 \cdot 0,65/0,80 = 0,6937 \approx 69\%$.

3. (a) $\Pr(X > 8) = \int_8^\infty \lambda e^{-\lambda x}\ dx = [-e^{\lambda x}]_8^\infty = e^{-8\lambda} = e^{-16} \approx 1,1 \cdot 10^{-7}$. (b) $E(X+Y) = E(X) + E(Y) = 1/2 + 1/2,5 = 0,9$, oberoende ger $SD(X+Y) = \sqrt{SD(X)^2 + SD(Y)^2} = \sqrt{(1/2)^2 + (1/2,5)^2} = \sqrt{0,41} \approx 0,64$. (c) Summan av exponentialfördelade variabler är varken normalfördelad eller exponentialfördelad, så det återstår att det är ”något annat” (man kan visa att täthetsfunktionen blir $f(x) = 10(e^{-2,0x} - e^{-2,5x})$).

4. Antag att värdena kommer från en normalfördelning (som vi brukar när det gäller vikter) och är oberoende (rimligt då de är från olika restauranger). Konfidensintervall för medelvärde kan då fås från detta stickprov med t-fördelning enligt sidan 154 i Wahlin och värdena $\bar{x} = 89,43$g, $s = 1,512$g, $n = 7$, $t_{n-1;1-0,10/2} = 1,943$ till $[88,3;\ 90,5]$g. Innan du lägger ut ett angrepp på Instagram riktat mot den internationella hamburgerkedjan, att de fuskar med vikten på sina quarter pounder, bör du dock beakta att kött tappar i vikt ca 15 – 20% vid tillagning.

5. (a) Falskt. Aldrig sannolikheten för en hypotes.
   (b) Falskt. Se (a).
   (c) Falskt. Vi säger aldrig att en nollhypotes godtas (och även om vi så skulle göra är det inte den sannolikheten)
   (d) Falskt. Detta är fel av typ-II, β.
   (e) Sant.
   (f) Falskt. Aldrig att nollhypotesen anses bevisad.

6. (a) Punktskattningen för vinsten är $\hat{y}_{800} = b_0 + b_1 \cdot 800 = 354,64$ tkr, där b_0 och b_1 tas ur utdatasammanfattningen som ”konstant” respektive ”X-variabel 1”. Felmarginalen för prognosintervall ges av $t_{n-2;1-\alpha/2} s \sqrt{1 + (1/n) + (x^* - \bar{x})^2 / \sum_{i=1}^n (x_i - \bar{x})^2}$. Med $n = 29$ och $1 - \alpha = 0,95$ fås tabellvärdet $t_{n-2;1-\alpha/2} = 2,052$ och ur utdatasammanfattningen $s = 55,5456$. Ur texten hittar vi $\bar{x} = 630$ timmar och $s_x = 245$ timmar, detta ger bidragen $(x^* - \bar{x})^2 = (800 - 630)^2 = 28900$ timmar^2 och $\sum_{i=1}^n (x_i - \bar{x})^2 = (n-1)s_x^2 = 1680700$ timmar^2. Därmed blir prognosintervallet 239 tkr $\leq$ vinsten $\leq$ 470 tkr.
   (b) Angivet p-värde för ”X-variabel 1” står för sannolikheten att få ett värde $|b_1| \geq 0,509946$ (det värde vi faktiskt fick, oavsett tecken, eller något ännu mer extremt) om ”nollhypotesen är sann”, dvs om det sanna värdet på riktningskoefficienten är noll (underförstådd nollhypotes vid regressionsanalys).