Linköpings universitet
Institutionen för teknik och naturvetenskap
Michael Hörnquist

Statistik och sannolikhetslära, TNIU66

Tentamen, 25 mars 2020, Svar och lösningsskisser

Nedanstående är svar, med viss diskussion och antydningar till hur uppgifterna *kan* lösas.

1. (a) Medelvärdet ökar, eftersom totala lönesumman ökar och antalet anställda är oförändrat. (b) Medianen är oförändrad, eftersom den delar stapelarean i två lika stora delar, och ingen area flyttar sig förbi mitten. (c) Typvärdet minskar, eftersom det är det vanligaste värdet, och här blir 25 kkr vanligast istället för 32 kkr när de tre personerna med lön 23 kkr får sin höjning. (d) Standardavvikelsen minskar, eftersom spridningen av löner minskar.

2. Låt A vara händelsen att få en 1:a eller 2:a på tärningen och B händelsen att svara "JA" på den efterföljande frågan.
   (a) Sökt sannolikhet är $p = \Pr(B|A^c)$ (eftersom vi bara är intresserade av de som svarat "ja" på frågan om narkotikabruk). Satsen om total sannolikhet ger oss $\Pr(B) = \Pr(A)\Pr(B|A) + \Pr(A^c)\Pr(B|A^c)$ och eftersom $\Pr(A) = 2/6$ (tärningskast, likformig sannolikhet), $\Pr(B|A) = 1/2$ (slantsingling, likformig sannolikkhet) och $\Pr(B) = 240/600$ (givet som förutsättning) får vi ekvationen $240/600 = (2/6)(1/2) + (4/6)p \Leftrightarrow p = 35/100$.
   (b) Betingad sannolkhet (Bayes sats med förenklad nämnare) $\Pr(A^c|B) = \Pr(B|A^c)\Pr(A^c)/\Pr(B) = p\ (4/6)/(240/600) = 35/60 \approx 58\%$.

3. (a) Låt X vara en slumpvariabel som beskriver längden hos *en* bräda, $X \sim N(\mu = 3,5; \sigma =?)$. Det är σ som är sökt. Givet är att $\Pr(3,45 \leq X \leq 3,55) = 0,80$. Eftersom intervallet är symmetriskt kring väntevärdet, med 10% i respektive "svans" i fördelningen till vänster om 3,45 och till höger om 3,55, kan vi uttrycka sambandet som $\Pr(X \leq 3,55) = 0,90$. Sedvanlig transformering till $N(0,1)$ ger $\Pr(Z \leq 0,05/\sigma) = 0,90$ och ur normalfördelningstabell (Appendix B Wahlin) fås $0,05/\sigma = 1,282$ (lättast ur nedersta raden i t-tabellen). Vi får därmed $\sigma = 3,9$ cm. (b) Låt X_1 och X_2 vara längd hos varsin bräda av samma slag som tidigare. Då är $X_1 \sim N(\mu = 3,5; \sigma = 0,039)$ och $X_2 \sim N(\mu = 3,5; \sigma = 0,039)$. Sätt $Y = X_1 + X_2$ så vet vi att $Y \sim N(\mu = 3,5+3,5; \sigma = 0,039 \cdot \sqrt{2}) = N(\mu = 7,0; \sigma = 0,055)$. Sökt $\Pr(Y < 6,9) = \Pr(Z < (6,9 - 7,0)/0,055) = \Pr(Z < -1,81) = 3,5\%$.

4. Stickprovets storlek är $n = 110+55+35 = 200$, populationens storlek är ovidkommande (sålänge den är substantiellt större än stickprovet). Punktskattning av andelen "ja"-väljare är $p = 110/200 = 0,55$. Eftersom $np(1-p) = 49,5 > 5$ går den vanliga normalapproximationen bra att använda. Konfidensnivån 90% svarar mot $\alpha = 0,10$.
   (a) Intervallets nedre gräns ges av $p - z_{1-\alpha}\sqrt{p(1-p)/n}$, vilket med tabellvärdet $z_{0,90} = 1,282$ ger intervallet $\pi > 50,5\%$.
   (b) Intervallgränserna ges av $p \pm z_{1-\alpha/2}\sqrt{p(1-p)/n}$, vilket med tabellvärdet $z_{0,95} = 1,645$ ger intervallet $49,2\% < \pi < 60,8\%$.

(c) Eftersom man redan innan stickprovet gjordes hade formulerat att syftet var att se om det fanns en majoritet för stängning, så är det rimligt att se på det enkelsidiga intervallet. Och då är svaret ”ja”.

5. (a) Falskt.
   (b) Falskt.
   (c) Falskt.
   (d) Falskt.
   (e) Sant.
   (f) Falskt.

6. (a) Uppåt begränsat intervall för riktningskoefficienten ges av $\beta_1 < b_1 + t_{n-2;1-\alpha}\mathrm{SE}(\beta_1) = 0,493 + 1,860 \cdot 0,089 = 0,66$. Alltså klart mindre än ett (1) på nivån 95%.
   (b) Prognosintervallet ges som $\hat{y}_{x^*} \pm t_{n-2;1-\alpha/2} s\sqrt{1 + 1/n + (x^* - \bar{x})^2 / \sum_i (x_i - \bar{x})^2}$. Punktskattningen $\hat{y}_{181} = 92,507 + 0,493 \cdot 181 = 181,7$ cm. Eftersom $\sum_i (x_i - \bar{x})^2 = s^2_{\text{fäder}}(n-1) = 568,8$ cm^2 (568,4 om någon räknar ut det direkt från x-värdena), medelfelet $s = 2,1235$ cm och tabellvärde $t_{8;0,975} = 2,306$ blir intervallet $[176,6;\ 186,9]$ cm.