

Statistik och sannolikhetslära, TNIU66

Tentamen, 21 mars 2019, Svar och lösningsskisser

Nedanstående är svar, med viss diskussion och antydningar till hur uppgifterna *kan* lösas. Det är *inte* nödvändigtvis fullständiga lösningar.

1. Medelvärde för respektive mängd blir 10, 1000 samt 0. De har dessutom samma stickprovsstandardavvikelse $3/\sqrt{2} \approx 2,12$. Eftersökta kvoter blir för de första två mängderna $3/(10\sqrt{2}) \approx 0,212$ och $3/(1000\sqrt{2}) \approx 2,12 \cdot 10^{-3}$, medan det för den tredje mängden blir division med noll och därmed inte går att beräkna.

Fördelen med spridningsmättet ”standardavvikelse delat med medelvärde” är att det tar hänsyn till hur stora talen i mängden är, det är i viss mening uppenbart att variationen i den första mängden är större än variationen i den andra mängden, trots att standardavvikelserna är lika. En nackdel är dock att måttet inte fungerar om det är mängder vars medelvärde ligger nära noll (med extremfallet som i uppgiften att det inte ens är definierat när medelvärdet är exakt noll).

2. De första tre bokstäverna kan kombineras på $(29 - 6)^3 - 91 = 12076$ sätt. Följande tre tecken (två siffror och en siffra eller bokstav) på $10^2 \cdot (10 + 29 - 6) - 1 = 3299$ sätt. Totalt fås alltså $12076 \cdot 3299 = 39\,838\,724$ kombinationer. Detta kan jämföras med situationen fram till och med 2018, när det fanns 12 063 924 kombinationer att tillgå.

3. Vi har $X \sim N(\mu = 3, 0; \sigma = 0, 75)$.

(a) $\Pr(X > 5) = 1 - \Pr(Z \leq (5 - 3)/0,75) \approx 0,38\%$.

(b) $\Pr(3,0 \leq 3,5) = \Pr(Z \leq (3,5 - 3)/0,75) - \Pr(Z \leq (3 - 3)/0,75) \approx 25\%$.

4. Stickprovets storlek är $n = 2342$ medan populationens storlek är ovidkommande (sålänge den är substantiellt större än stickprovet, vilket Sveriges befolkning med rösträtt är). Punktskattningen av andel är den som opinionsundersökningen ger, dvs $p = 0,034$.

(a) Uppåt begränsat enkelsidigt konfidensintervall för andel ges av $\pi \leq p + z_{1-\alpha} \sqrt{p(1-p)/n}$. För $\alpha = 5\%$ fås $z_{1-\alpha} = 1,645$. Detta är giltigt eftersom $np(1-p) > 5$. Instoppade värde ger $\pi \leq 4,016\%$. På konfidensnivån 95% finns alltså ingen täckning för rubriken i tidningen.

(b) Testvariabeln $z = (p - \pi_0)/\sqrt{\pi_0(1 - \pi_0)/n} \approx -1,48$. Kritiska området ligger till vänster om z_α , vilket ur tabell ger oss att $\alpha = 0,06944$. Signifikansnivån måste alltså vara $\alpha = 6,9\%$ eller högre för att rubriken ska vara statistiskt korrekt.

5. (a) Falskt. Krävs fyra gånger så stort ($\sqrt{n}$ i nämnaren).
 - (b) Falskt. Mest konstigt, konfidensnivå är ofta 95% eller så.
 - (c) Falskt. Ett enkelsidigt intervall är ju i princip oändligt stort.
 - (d) Falskt. Det är *aldrig* sannolikheten att en hypotes är sann eller falsk som beräknas (även om vi önskar att kunna ta fram den).
 - (e) Falskt. Se (d).
 - (f) Sant.
6. (a) Står i texten att förklaringskraften $r^2 = 71\%$, och då kurvan lutar nedåt blir korrelationskoefficienten $r \approx -0,84$.
 - (b) Ur utdatasammanfattningen framgår direkt $r = \text{Multipel-R} = 0,97$.
 - (c) Styckpriset 250 kr motsvarar $x = 1/250 = 0,004$, vilket insatt i regressionslinjen blir $\hat{y}_{0,004} = 0,8015 + 2128,9497 \cdot 0,004 \approx 9,3$.
 - (d) Konfidenstintervallet ges av uttrycket

$$b_0 + b_1x \pm t_{n-2;1-\alpha/2} s \sqrt{\frac{1}{n} + \left(\frac{x - \bar{x}}{s/\text{SE}(b_1)} \right)^2}.$$

Medelvärde $\bar{x}$ är medelvärdet av alla observationer längs x -axeln i figur 2, dvs medelvärdet av alla inverterade försäljningspris. Räkande ger $\bar{x} = \frac{1}{13}(\frac{1}{60} + \frac{1}{70} + \dots + \frac{1}{300}) \approx 0,008909454$. Med tabellvärdet $t_{11;0,95} = 1,796$ och $x = 0,004$ fås med värden från utdatasammanfattningen $9,317 \pm 1,589 = [7,7; 10,9]$.

Bonusfrågan: Den totala summan pengar för varje prisnivå blir som störst för just priset 250 kr, i försöket som Jonas genomförde.