Linköpings universitet
Institutionen för teknik och naturvetenskap
Michael Hörnquist

Statistik och sannolikhetslära, TNIU66

Tentamen, 21 mars 2024, Svar och lösningsskisser

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

1. (a) Direkt räkning utifrån sambandet för medelvärde $\sum$ antal $\cdot$ värde/(totalt antal) ger $(3\cdot(-2)+0\cdot(-1)+6\cdot 0+4\cdot 1+1\cdot 2+4\cdot 3+1\cdot 4+2\cdot 5+1\cdot 6)/(3+0+6+4+1+4+1+2+1) = 32/22 \approx 1,45$. Med 22 element i mängden fås medianen som medelvärdet av talen på position 11 och 12 om de ordnas i storleksordning. Eftersom båda elementen har värdet 1, blir det också medianens värde. Typvärdet är det vanligast värdet, dvs högsta stapeln, dvs 0.

   (b) Om allt multipliceras med två blir allt dubbelt så stort (inklusive standardavvikelsen, men det frågades inte efter). Medelvärde blir $32/11 \approx 2,91$, median blir 2, och typvärde blir 0.

2. (a) $P(X \geq 5) = /\text{oförenliga händelser}/ = P(X=5)+P(X=6) = 0,54+0,14 = 0,68$

   (b) $E(X) = 0\cdot 0,00+1\cdot 0,10+2\cdot 0,02+3\cdot 0,07+4\cdot 0,12+5\cdot 0,54+6\cdot 0,14 = 4,37$
   $SD(X) = \sqrt{(0-4,37)^2\cdot 0,00+(1-4,37)^2\cdot 0,10+\cdots+(6-4,37)^2\cdot 0,14} = 1,41$

   (c) Inte alls, staplarnas höjd ges inte av sambandet för punktsannolikheter för binomialsannolikheter.

3. Multiplikationssatsen för beroende händelser.

   (a) Pr(båda svarta)=Pr(första svart)Pr(andra svart|första svart)= $(2/5)(1/4) = 1/10$

   (b) Pr(båda vita)=Pr(första vit)Pr(andra vit|första vit)= $(3/5)(2/4) = 3/10$

   (c) Pr(en vit och en svart)=Pr(första vit och andra svart)+Pr(första svart och andra vit)= Pr(första vit)Pr(andra svart|första vit) + Pr(första svart)Pr(andra vit|första svart) = $(3/5)(2/4)$ $+(2/5)(3/4) = 6/10$.

   Notera (c) kan också lösas med komplementsatsen och resultaten från (a) och (b). Alla tre deluppgifterna kan även lösas med resonemang kring ”dragning utan återläggning och utan ordning”.

4. (a) Vikter tillhör sådant som brukar vara normalfördelat, vilket vi antar att även dessa är. Eftersom standardavvikelsen är beräknad från ett stickprov var den inte känd i förväg, och vi har därför ett konfidensintervall för medelvärde som ges av $[\bar{x}+t_{n-1;\alpha/2}s/\sqrt{n},\ \bar{x}+t_{n-1;1-\alpha/2}s/\sqrt{n}]$. Med $\bar{x} = 57,5$g, $s = 5,0$g, $1-\alpha = 0,95$, $n = 10$ och $t_{10-1;0,05/2} = 2,262$ fås $[57,5-2,262\cdot 5,0/\sqrt{10},\ 57,5+2,262\cdot 5,0/\sqrt{10}]$ (gram), dvs intervallet mellan 53,9 gram och 61,1 gram innehåller medelvikten med 95% säkerhet.

   (b) Medelfelet ges av $s/\sqrt{n} = 1,58$ g.

   (c) Om Fritz väger stickprovet i klump tapppar han information om variationen, och allt han får fram är en punktskattning av medelvikten som 575 g/10 = 57,5 gram. Han har inte en aning om hur pass säker han kan vara på det värdet.

5. Hypotest, enkelsidigt, för andel.

   Låt π beteckna den sanna men okända andelen pizzor med omkrets över en meter. Eftersom Jonas vill bevisa att pizzorna är större än en meter är alternativhypotesen, den som ska bevisas, $H_a : \pi > \pi_0 = 80\%$. Nollhypotesen är $H_0 : \pi \leq \pi_0 = 80\%$, där vi räknar i ”värsta-fall scenario” för pizzerian, vilket är gränsen π_0. Hypotestestning är konservativ till nollhypotesens fördel.

   Stickprovets storlek är $n = 60$, vilket för 54 pizzor av tillräcklig storlek ger punktskattningen $p = 54/60 = 0,9$. Därmed följer $np(1-p) = 5,4 > 5$ och normalapproximationsformlerna kan användas.

   Vi bildar $z = (p - \pi_0)/\sqrt{\pi_0(1-\pi_0)/n} = 1,936$ och jämför med normalfördelningstabellens värde $z_{1-0,05} = 1,645$. Eftersom $z > z_{0,95}$ följer att nollhypotesen *kan* förkastas.

   Slutsats: På signifikansnivån 5% är det statistiskt bevisat att minst 80% av pizzorna har en omkrets som överstiger en meter.

6. (a) Falskt. Orsakssamband får vi överhuvud taget aldrig ut med statistiska metoder.

   (b) Falskt. Går alltid att räkna ut en korrelationskoefficient.

   (c) Falskt. $-0,4$ ligger mellan -1 och 1 och är ett högst legitimt värde.

   (d) Falskt. b_1 kan anta vilken reellt värde som helst.

   (e) Falskt. b_0 är skärningen mellan regressionslinjen och y-axeln.

   (f) Sant.