Stockholms universitet
Statistiska institutionen
Per Gösta Andersson

Statistisk teori med tillämpningar, del 2

SKRIFTLIG TENTAMEN
Onsdagen den 26 oktober, 2022, kl 14.00-19.00.

Godkända hjälpmedel: Miniräknare, språklexikon, formelsamling med tabeller (bifogas).

Gräns för godkänt: 50 poäng av totalt 100.

För maximal poäng på varje uppgift krävs tydliga, utförliga och välmotiverade lösningar.

1. (20p) Från var och en av två stora populationer P_1 och P_2 har man slumpmässigt valt ut sex personer och tagit reda på deras månadslön (enhet: tusen kronor):

   P_1: 15.4 16.0 18.3 20.5 28.2 36.5

   P_2: 18.5 22.8 24.5 29.6 37.5 40.0

   Undersök med lämpliga test med tillhörande notation om man kan påstå att inkomsterna i P_2 är större än i P_1 på signifikansnivå 5% om

   (a) vi antar normalfördelade inkomster med samma varians i populationerna

   (b) vi inte kan anta normalfördelade inkomster i populationerna.

2. (25p) $X_1, \ldots, X_n$ är ett stickprov på $X \sim exp(\beta)$, där $\beta = E(X)$. Vi mäter livslängder med denna fördelning och vill intervallskatta den genomsnittliga populationslivslängden β. Vi utgår därför från punktskattningen $\bar{X} = (1/n)\sum_{i=1}^{n} X_i$.

   (a) Visa att $\bar{X}$ är en väntevärdesriktig och konsistent punktskattning av β.

   (b) Är $\sum_{i=1}^{n} X_i$ en tillräcklig ("sufficient") statistika för β? (Du kan utgå från att vi fått ett observerat stickprov $x_1, \ldots, x_n$.)

   (c) Man kan visa att $Y = 2n\bar{X}/\beta \sim \chi^2(2n)$.
   Vilka är de två kraven på en "pivot"? Uppfyller Y dessa krav?

   (d) Konstruera och beräkna ett exakt 95% symmetriskt två-sidigt konfidensintervall för β om $n = 10$ och $\bar{x} = 3.72$.

   (e) Konstruera och beräkna ett approximativt 95% symmetriskt två-sidigt konfidensintervall för β om $n = 100$ och $\bar{x} = 3.72$.

3. (15p) En fysikalisk parameter θ kan mätas indirekt genom observationer av diskreta oberoende stokastiska variabler $X_1, \ldots, X_{10}$, som alla har sannolikhetsfunktioner givna av

$$p(0) = P(X_i = 0) = e^{-\theta} \text{ och } p(1) = P(X_i = 1) = 1 - e^{-\theta}$$

   Bestäm ML-metodens skattning av θ för utfallet

$$(x_1, \ldots, x_{10}) = (0,\ 1,\ 0,\ 0,\ 1,\ 0,\ 0,\ 0,\ 0,\ 1).$$

   Bedöm också om vi fått en väntevärdesriktig skattning av θ. Inga beräkningar behövs.

   Tips: Hela uppgiften kan underlättas om man låter $X = \sum_{i=1}^{10} X_i$.

4. (20p) Låt $X_1, \ldots, X_{10}$ vara ett stickprov på $X \sim Poi(\lambda)$. Vi vill pröva

$$H_0 :\ \lambda = 0.1 \text{ mot } H_a :\ \lambda = 0.5$$

   med ett starkaste test. Visa att förkastelseområdet får formen
   $\sum_{i=1}^{10} x_i \geq c$ för ett observerat stickprov $x_1, \ldots, x_n$.

   Om vi låter $c = 3$, vad blir α och $1 - \beta(0.5)$, dvs vad blir testets signifikansnivå och styrka?

   Ledning: En summa av oberoende Poissonfördelade variabler är också Poissonfördelad.

5. (20p) Vid jämförelse mellan två fall under normafördelningsantagande med samma varians σ^2 används ju den sammanvägda stickprovsvariansen

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} \, ,$$

där n_1 och n_2 är stickprovsstorlekarna.

Visa att S_p^2 har minst varians av alla väntevärdesriktiga punktskattningar av variansen σ^2, som har formen $aS_1^2 + bS_2^2$, där a och b är konstanter.

Ledning: Bestäm först ett villkor för hur a och b ska förhålla sig till varandra för väntevärdesriktighet.