

Statistik och sannolikhetslära, TNIU66

Tentamen, 23 augusti 2025, Svar och lösningsskisser

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

Del B: Sannolikhetslära

1. En vanlig kortlek med 52 kort blandas.

- a) Vad är sannolikheten att dra ett ess i första draget?

Lösning: Det finns 4 ess i en kortlek med 52 kort. Sannolikheten att dra ett ess i första draget är:

$$\Pr(\text{ess}) = \frac{4}{52} = \frac{1}{13}$$

- b) Vad är sannolikheten att dra två ess i rad utan återläggning?

Lösning: Sannolikheten att dra ett ess i första draget är $\frac{4}{52}$. Om det första kortet är ett ess, finns det 3 ess kvar i en kortlek med 51 kort. Sannolikheten att dra ett ess i andra draget är $\frac{3}{51}$. Den totala sannolikheten är:

$$\Pr(2 \text{ ess}) = \frac{4}{52} \times \frac{3}{51} = \frac{1}{13} \times \frac{1}{17} = \frac{1}{221}$$

- c) Vad är sannolikheten att bland fem drag med återläggning och ordentlig blandning mellan varje drag få exakt två ess?

Lösning: Detta är ett binomialförsök med $n = 5$ och $p = \frac{1}{13}$. Sannolikheten att få exakt 2 ess är:

$$\Pr(X = 2) = \binom{5}{2} \left(\frac{1}{13}\right)^2 \left(1 - \frac{1}{13}\right)^{5-2} = 10 \times \left(\frac{1}{13}\right)^2 \times \left(\frac{12}{13}\right)^3 = 17280/371293 \approx 0,04654$$

2. Täthetsfunktion

En slumpvariabel X har täthetsfunktionen

$$f(x) = \begin{cases} \frac{1}{10}, & \text{om } -5 \leq x \leq 5 \\ 0, & \text{annars} \end{cases}$$

- a) Visa att f verkligen är en täthetsfunktion.

Lösning: För att f ska vara en täthetsfunktion måste den uppfylla två villkor:

- $f(x) \geq 0$ för alla x
- $\int_{-\infty}^{\infty} f(x) dx = 1$

Det första villkoret är uppfyllt eftersom $f(x) = \frac{1}{10}$ är positivt för $-5 \leq x \leq 5$ och 0 annars. För det andra villkoret:

$$\int_{-5}^5 \frac{1}{10} dx = \frac{1}{10} \times (5 - (-5)) = \frac{1}{10} \times 10 = 1$$

Så f är en giltig täthetsfunktion.

- b) Bestäm $\Pr(-1 \leq X \leq 1)$

Lösning: Sannolikheter är areor är integraler.

$$\Pr(-1 \leq X \leq 1) = \int_{-1}^1 f(x) dx = \int_{-1}^1 \frac{1}{10} dx = \frac{2}{10} = \frac{1}{5}$$

- c) Antag att vi tar ett medelvärde av 40 oberoende observationer från X . Bestäm sannolikheten att detta medelvärde ligger mellan -1 och 1.

Lösning: Enligt centrala gränsvärdessatsen är medelvärdet av 40 oberoende observationer från X approximativt normalfördelat med väntevärde $E(X)$ och standardavvikelse $\sigma(X)/\sqrt{40}$. Här är X likformigt fördelat så $E(X) = (5 + (-5))/2 = 0$ och $\sigma(X) = (5 - (-5))/\sqrt{12} = 5/\sqrt{3}$. Därmed

$$\sigma(\bar{X}) = \frac{\sigma(X)}{\sqrt{40}} = \frac{5}{\sqrt{3} \times 40} = \frac{5}{2\sqrt{30}}$$

Vi vill bestämma $\Pr(-1 \leq \bar{X} \leq 1)$:

$$\Pr(-1 \leq \bar{X} \leq 1) = \Pr\left(\frac{-1-0}{\frac{5}{2\sqrt{30}}} \leq Z \leq \frac{1-0}{\frac{5}{2\sqrt{30}}}\right) = \Pr(-2,19 \leq Z \leq 2,19) = 0,97148$$

Använd z-tabell för att hitta sannolikheten.

Del C: Statistik

3. En forskare undersöker hur lång tid (i minuter) det tar för studenter att lösa en viss typ av uppgift.

Ett stickprov på elva studenter ger följande tider:

Tider: 18, 22, 19, 25, 21, 20, 23, 24, 19, 20

Hon matar in serien i sin miniräknare (Casio fx-82) och får följande:

$$\sum x^2 = 4501, \sum x = 211, n = 10, \bar{x} = 21.1, \sigma = 2.211, s = 2.331$$

- a) Beräkna ett 95% dubbelsidigt konfidensintervall för den genomsnittliga lösningstiden i populationen.

Lösning: Konfidensintervallet för medelvärdet blir t-fördelat eftersom stickprovet kommer från en normalfördelning med okänd standardavvikelse.

$$\bar{x} \pm t_{1-\alpha/2, n-1} \times \frac{s}{\sqrt{n}}$$

För $n = 10$ och $\alpha = 0.05$ blir $t_{0,975,9} = 2,262$ från t-tabell. Konfidensintervallet blir:

$$21,1 \pm 2,262 \times \frac{2.331}{\sqrt{10}} = 21,1 \pm 1,67 = [19,4, 22,8].$$

- b) Förklara vad detta konfidensintervall innebär i praktiken.

Lösning: Ett 95% konfidensintervall innebär att intervallet med 95% sannolikhet rymmer det sanna medelvärdet.

- c) Hur påverkas konfidensintervallet om man istället hade haft ett stickprov på 100 studenter med samma medelvärde och standardavvikelse? Hur många gånger kortare eller längre blir det?

Lösning: Konfidensintervallet blir smalare eftersom standardfelet minskar med större stickprov. Standardfelet för $n = 100$ är:

$$\frac{s}{\sqrt{100}} = \frac{2.331}{10} = 0.2331$$

Jämfört med standardfelet för $n = 10$:

$$\frac{s}{\sqrt{10}} = \frac{2.331}{\sqrt{10}} = 0.737$$

Konfidensintervallet blir ungefär $\frac{0.737}{0.2331} \approx 3.16$ gånger smalare. Tar man hänsyn till att även t -värdet minskar när n ökar tillkommer en faktor $2,262/1,984 = 1,140$.

4. Ett företag hävdar att deras produkter har en felfrekvens på högst 5%.

- a) Formulera noll- och alternativhypotes för ett test där man kan statistiskt bevisa att företaget har högre felfrekvens.

Lösning: Nollhypotes (H_0): $\pi \leq \pi_0 = 0,05$

Alternativhypotes (H_a): $\pi > \pi_0 = 0,05$

- b) I ett stickprov på 200 produkter hittas 16 defekta. Genomför ett hypotestest på 5% signifikansnivå.

Lösning: Börjar med att kontrollera $np(1-p) > 5$ så att vår vanliga formalism fungerar (effektivt att testvariabeln blir normalfördelad). Med $n = 200$ och $p = 16/200$ är kravet uppfyllt (14,72). Testvariabeln är:

$$z = \frac{p - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}$$

Där $p = \frac{16}{200} = 0,08$, $\pi_0 = 0,05$, $n = 200$:

$$z = \frac{0,08 - 0,05}{\sqrt{\frac{0,05 \times 0,95}{200}}} = \frac{0,03}{\sqrt{0,0002375}} = \frac{0,03}{0,0154} \approx 1,95$$

Jämför med kritiskt värde från z -tabell för $\alpha = 0,05$, $z_{1-\alpha} = 1,645$, ser att $z > z_{0,95}$ och därmed hamnar vi i det kritiska området och *nollhypotesen kan förkastas*. Undersökningen bevisar statistiskt att företaget talar osanning.

- c) Vad är risken för typ I-fel i detta sammanhang?

Lösning: Risken för typ I-fel är signifikansnivån, vilket är 5%. Det innebär att risken att förkasta en sann nollhypotes är 5%, dvs något som i genomsnitt sker vid var 20:e stickprovsundersökning.

5. Regression

En mäklare undersöker sambandet mellan antal kvadratmeter x och månadshyra y för 6 lägenheter. En regressionsanalys har gjorts i Excel.

- a) Förklaringsgraden ges av R^2 , som i detta fall är:

$$R^2 = 0,9659$$

- b) Punktskattningen för hyran ges av $\hat{y} = b_0 + b_1 \cdot x = 1577,28 + 77,88 \cdot 55 = 5860,68$. Felmarginalen för en lägenhet 55 m^2 ges uttrycket för prognosintervall

$$t_{1-\alpha/2; n-2} s \sqrt{1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}},$$

vilken med observationen $\bar{x} = 55 = x$ blir

$$t_{1-\alpha/2; n-2} s \sqrt{1 + \frac{1}{n}}.$$

Ur utdatasammanfattningen fås $s = 306,45$ och ur tabell $t_{0,975;4} = 2,776$, sammantaget blir felmarginalen 918,87. Sökt prognosintervall blir [4942; 6780] kr.

- c) Punktskattningen för hyran ges av $\hat{y} = b_0 + b_1 \cdot x = 1577,28 + 77,88 \cdot 55 = 5860,68$. Felmarginalen för *medelvärdet* av lägenheter 55 m^2 ges ur uttrycket för konfidensintervall

$$t_{1-\alpha/2;n-2}s\sqrt{\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}},$$

vilken med observationen $\bar{x} = 55 = x$ blir

$$t_{1-\alpha/2;n-2}s\sqrt{\frac{1}{n}}.$$

Ur utdatasammanfattningen fås $s = 306,45$ och ur tabell $t_{0,975;4} = 2,776$, sammantaget blir felmarginalen $347,40$. Sökt prognosintervall blir $[5513; 6208]$ kr.