Tenta MT1011 8 januari 2024 lösningar

1 (a) Av 50 individer, varav 20 rökare, dras 25 individer slumpmässigt (utan återläggning) för att ingå i gruppen som får läkemedel.

$$X = \#\text{ rökare i denna grupp.}$$

Då är $X \sim \text{Hyp}(N=50, n=25, p=\frac{20}{50})$, varmed

$$E[X] = np = 25 \cdot \frac{20}{50} = 10$$

$$\text{Var}(X) = np(1-p)\frac{N-n}{N-1} = 25 \cdot \frac{20}{50} \cdot \frac{30}{50} \cdot \frac{25}{49} = \frac{150}{49}$$

5p

(b) Notera att samtliga rökare hamnar i samma grupp om antingen $X=0$ eller $X=20$. Vi får

$$P(X=0) = \frac{\binom{20}{0}\binom{30}{25}}{\binom{50}{25}} = \frac{30!}{\cancel{25!}\,5!} \cdot \frac{25!\cancel{25!}}{50!} \approx 1.127 \cdot 10^{-9}.$$

På samma sätt fås

$$P(X=20) = \frac{\binom{20}{20}\binom{30}{5}}{\binom{50}{25}} = \frac{30! \cdot 25!}{5! \cdot 50!} = P(X=0).$$

5p

Därmed blir svaret $\approx 2.25 \cdot 10^{-9}$.

2 Drag kulorna en och en och placera dem på rad. Var kula är lika sannolik att hamna på en given position. Då det finns 10 kulor så gäller därmed för varje $k = 1, 2, \ldots, 10$ att

$$P(\text{röd på position } k) = \frac{1}{10}.$$

4p

Det följer att:

(a) $P(\text{röda sist}) = 1/10.$ 2p

(b) $P(\text{röda först eller sist}) = P(\text{röda först}) + P(\text{röda sist})$

$$= \frac{1}{10} + \frac{1}{10} = 1/5.$$

2p

(c) $P(\text{första två olika}) = P(\text{röda först}) + P(\text{röda två})$

$$= \frac{1}{10} + \frac{1}{10} = 1/5.$$

2p

$\boxed{3}$ (a) En täthet är positiv och integrerar till 1. Därmed gäller

$$1 = \int_{-1}^{1} c(1-x^2)\,dx = c\left[x - \frac{x^3}{3}\right]_{-1}^{1} = c\,2\left(1-\frac{1}{3}\right) = c\,\frac{4}{3}.$$

Vi får $c = \frac{3}{4}$. 3p

(b) Då X är kontinuerlig gäller

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x\, f_X(x)\,dx = \int_{-1}^{1} \frac{3}{4}(x-x^3)\,dx = \frac{3}{4}\left[\frac{x^2}{2} - \frac{x^4}{4}\right]_{-1}^{1} = 0.$$

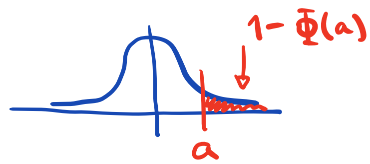
$$\mathbb{E}[X^2] = \int_{-1}^{1} x^2 \cdot \frac{3}{4}(1-x^2)\,dx = \int_{-1}^{1} \frac{3}{4}(x^2-x^4)\,dx$$

$$= \frac{3}{4}\left[\frac{x^3}{3} - \frac{x^5}{5}\right]_{-1}^{1} = \frac{3}{4}\cdot 2\left(\frac{1}{3}-\frac{1}{5}\right) = \frac{3}{2}\cdot\frac{2}{15} = \frac{1}{5}$$ 3p

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = 1/5.$$

(c) För oberoende och likafördelade variabler med väntevärde μ och varians σ^2 gäller enligt centrala gränsvärdessatsen för stora n approximativt

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1).$$



Vi får därmed

$$\mathbb{P}(\bar{X} > 1/5) = \mathbb{P}\left(\frac{\bar{X}}{1/\sqrt{5\cdot 50}} > \frac{1}{5/\sqrt{5\cdot 50}}\right)$$ 4p

$$\approx 1 - \Phi(\sqrt{10}) \approx 1 - \Phi(3.16)$$

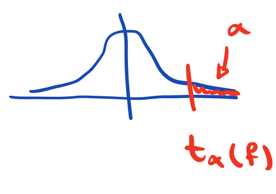
ur tabell $\approx 1 - 0.9992 = 0.0008$

$\boxed{4}$ (a) Då beroende kan misstänkas mellan mätningar i samma sjö så betraktar vi differenserna, och analysear data som ett stickprov i par. Differenserna efter − före blir 2p

0.4 0.5 0.6 0.6 0.9 1.0 0.3 −0.1

Vi kan därmed betrakta detta som ett stickprov $X_1, X_2, \ldots, X_8$ av en $N(\mu, \sigma^2)$-fördelning. Ett 95 % konfidensintervall för μ ges av

$$\bar{X} \pm t_{0.025}(7)\cdot\frac{s}{\sqrt{8}}.$$ 2p



Beräkning ger

$$\bar{x} = 0.525 \qquad s = 0.3454$$
$$t_{0.025}(7) \approx 2.365 \qquad \text{(ur tabell)}$$

så vi får att 95% KI för μ ges av

$$(0.525 \pm 0.289) = (0.236, 0.814).$$ 2p

(b) Vi vill testa hypotesen att kalkning påverkar pH-värdet i sjöarna. Att kalkning saknar effekt motsvarar $\mu = 0$. Vi ställer upp

$$H_0: \mu = 0$$
$$H_1: \mu \neq 0$$ 2p

Vid test av ovanstående hypotes kan H_0 förkastas på signifikansnivå 5% om och endast om motsvarande 95%-iga KI för μ innehåller värdet för μ under H_0, dvs värdet 0. Det gör det ej, så H_0 förkastas. 2p

$\boxed{5}$ Drag en individ slumpmässigt ur populationen. Låt

$S = \{\text{individ sjuk}\}$
$+ = \{\text{test positivt}\}$
$- = \{\text{test negativt}\}$.

Givet ur uppgiften är

$$\mathbb{P}(S) = 0.001$$
$$\mathbb{P}(+ \mid S) = 0.995$$ 4p
$$\mathbb{P}(- \mid S^c) = 0.970.$$

(a) Lagen om total sannolikhet ger

$$\mathbb{P}(+) = \mathbb{P}(+ \mid S)\mathbb{P}(S) + \mathbb{P}(+ \mid S^c) \cdot \mathbb{P}(S^c)$$ 3p
$$= 0.995 \cdot 0.001 + 0.030 \cdot 0.999 \approx 0.0310$$

(b) Bayes sats ger

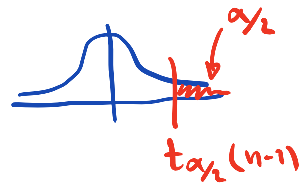
$$\mathbb{P}(S \mid +) = \frac{\mathbb{P}(S \cap +)}{\mathbb{P}(+)} = \frac{\mathbb{P}(+ \mid S) \cdot \mathbb{P}(S)}{\mathbb{P}(+)}$$ 3p
$$\approx \frac{0.995 \cdot 0.001}{0.0310} \approx 0.0321$$

6

(a) Falskt. $\bar{X}$ är normalfördelat med väntevärde μ och varians σ^2/n. **2p**

(b) Sant. Vi har

$$\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)\text{-fördelat}$$



och därmed att

$$\bar{X} \pm t_{\alpha/2}(n-1)\cdot\frac{S}{\sqrt{n}}$$

ger ett $(1-\alpha)\cdot 100\,\%$-igt konfidensintervall för μ. **2p**

Då n ökar så minskar felmarginalen, och intervallet blir mindre. Dessutom minskar $t_{\alpha/2}(n-1)$.

(c) Falskt. Ökad konfidens motsvarar ett lägre α, som i sin tur ger ett större värde på $t_{\alpha/2}(n-1)$. **2p**

(d) Falskt. Nollhypotesen H_0 är antingen sann eller ej. Det p-värdet anger är sannolikheten att observera ett så extremt resultat som vi gjort, givet att H_0 är sann. Dvs. **2p**

$$p\text{-värde} = \mathbb{P}(\text{data} \mid H_0 \text{ sann}).$$

(e) Sant. Under H_0 gäller att

$$T = \frac{\bar{X}}{S/\sqrt{n}} \sim t(n-1)\text{-fördelat.}$$

H_0 förkastas om vi observerar ett värde på $|T|$ som överstiger $t_{0.025}(n-1)$. Det är detsamma som att **2p**

$$\bar{X} > t_{0.025}(n-1)\cdot\frac{S}{\sqrt{n}} \quad \text{eller} \quad \bar{X} < -\,t_{0.025}(n-1)\cdot\frac{S}{\sqrt{n}}.$$

Ovan gäller om och endast om

$$\left(\bar{X} - t_{0.025}(n-1)\frac{S}{\sqrt{n}},\ \bar{X} + t_{0.025}(n-1)\frac{S}{\sqrt{n}}\right)$$

innehåller värdet 0.