STOCKHOLMS UNIVERSITET
MATEMATISKA INSTITUTIONEN
Avd. Matematisk statistik, TB

MT4001
LÖSNINGAR
26 februari 2026

Lösningar

Tentamen i Statistisk analys, 26 februari 2026

Uppgift 1

a) Sant

b) Falskt

c) Falskt

d) Sant

e) Falskt

Uppgift 2

a) Vi beräknar $t = (\bar{x} - 10)/(s/\sqrt{n}) = 1.857$. Vår referensvariabel har under $H_0 : \mu = 10$ en t-fördelning med $n - 1 = 9$ frihetsgrader. Vi förkastar H_0 om $|t|$ är stor. Från Tabell ser vi att vi ska förkasta H_0 på 95%-nivån om $|t| > 2.26$. Slutsatsen blir således att vi *inte* kan förkasta H_0. Visserligen är medelvärdet $\bar{x} = 10.32$ en bit ifrån hypotesvärdet 10 men med tanke på datamaterialets spridning kan denna avvikelse gott och väl ske av ren slump.

b) Ett 99% konfidensintervall för μ ges av: $\bar{x} \pm t_{\alpha/2}(n-1)s/\sqrt{n} = 10.32 \pm 3.25 * 0.545/\sqrt{10} = [9.76,\ 10.88]$.

Uppgift 3

a) Vi antar modellen enkel linjär regression. Denna modell antar att de olika växterna växer till sig oberoende av varandra (fullt rimligt) och att tillväxten men inte temperaturen är slumpmässigt (fullt rimligt). Dessutom antar modellen att den genomsnittliga tillväxten ökar linjärt med tiden, och att observationernas varians inte beror på temperaturen x vilket ju inte går att "bevisa", men om man ritar ut data punkterna finns i alla fall inget som talar emot detta antagande. Dvs $Y = \alpha + \beta x + \epsilon$, där $E(\epsilon) = 0$ och $V(\epsilon) = \sigma^2$.

Modellen har tre parametrar, α, β och σ. Från formelsamlingen ges skattningarna av $\hat{\beta} = S_{xy}/S_{xx} = 15.8/40 = 0.395$, $\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} = 66.88$. Vi skattar σ med $s^2 = SSE/(n-2) = \left(S_{yy} - 15.8^2/40\right)/3 = (6.368 - 15.8^2/40)/3 = 0.0423$, så $s = 0.206$.

b) Effekten av att öka temperaturen är β. Denna skattas såklart med $\hat{\beta} = 0.395$ och ett 95% konfidensintervall ges av

$$\hat{\beta} \pm t_{0.025}(n-2)\frac{s}{\sqrt{S_{xx}}} = 0.395 \pm 3.18\frac{0.206}{\sqrt{40}} = 0.395 \pm 0.104 = [0.291,\ 0.499]$$

c) En skattning av medellängden vid 19.5 grader ges av $\hat{\alpha} + \hat{\beta}19.5 = 74.58$.

Uppgift 4

a) Om man tittar på fördelningarna så är de lite skeva med några extra långa personer. Därmed kan dessa få för stort genomslag om man använder medelvärden och standardavvikelser.

b) Vi änvänder Wilcoxons 2-stickprovstest. Vi rangordnar spelarnas längder gemensamt. Det svenska laget får då rangerna 1, 2, 3, 5 och 8, medan Serbiens spelare får rangerna 4, 6, 7, 9, 10. Summan av de svenska rangerna blir $R = 19$. Eftersom vi använder två sidigt test så förkastar vi H_0 om summan är för stor eller för liten, där dessa gränser $k_{low} = 17$ och $k_{high} = 38$ fås från Tabell 9 med $n_1 = n_2 = 5$ och 2.5% i varje ända. Eftersom varken $R \leq k_{low}$ eller $R \geq k_{high}$ så förkastar vi *inte* H_0. Så trots att Serbiens spelare tycks vara längre så är stickprovet så pass litet så att detta skulle kunna vara en ren slump.

Uppgift 5

a) Eftersom vi i förväg inte har någon kunskap om näbblängden så är det rimligt att vilket värde som helst på p är lika rimligt, dvs att apriorifördelningen för p är rektangelfördelad på 0 till 1, $p \sim U[0,1]$ med täthet $f(p) = 1,\ 0 \leq p \leq 1$.

b) Om vi låter X vara antal blåmesar vars näbbar är längre än 1.2 cm så är denna fördelning, givet p, $X \sim Bin(n = 10, p)$. Vi har observerat $x = 2$. Likelihooden blir således $L(p) = \binom{10}{2}p^2(1-p)^{10-2}$.

c) Aposteriorifördelningen $f(p|x=2)$ är proportionell mot produkten av likelihood och apriorifördelning, så $f(p|x=2) \propto \binom{10}{2}p^2(1-p)^8 \cdot 1 \propto p^2(1-p)^8$ för $p \in [0,1]$ (och 0 för andra p). Notera att detta nu är en fördelning för p.

d) Om man ritar upp denna kurva så är den störst för värden nära 0.2 och avtar snabbt för värden både mindre och större än 0.2. Det verkar således rimligt att anta att aposteriorifördelningens väntevärde är nära 0.2. Detta stämmer, man kan visa att väntevärdet är 0.25.

Uppgift 6

a) Baserat på ett stickprov $x_1, \ldots, x_n$ från $\Gamma(k, \beta)$ så blir likelihooden

$$L(\beta) = \prod_i \beta^k x_i^{k-1} e^{-\beta x_i}/(k-1)! = \frac{\beta^{kn}}{((k-1)!)^n} e^{-\beta \sum_i x_i} \prod_i x_i^{k-1}.$$

Loglikelihooden blir således

$$\ell(\beta) = kn \log(\beta) - n \log\left((k-1)!\right) - \beta \sum_i x_i + \sum_i (k-1) \log(x_i).$$

b) Om vi deriverar loglikelhooden får vi $\ell'(\beta) = \frac{kn}{\beta} - \sum_i x_i$. Om vi sätter derivatan till 0 får vi lösningen $\hat{\beta} = \frac{k}{\bar{x}}$.

c) Eftersom $E(X) = k/\beta$ så bör $\bar{x} \approx k/\beta$, vilket i sin tur leder till att $\hat{\beta} \approx \beta$ vilket ju är vad vi önskar.