

STATISTIK OCH DATAANALYS III

Tentamen 1, 8.5 hp, ST2101

Formalia

Datum: 4 mars 2025

Skrivtid: 14.00 - 19.00

Resultat meddelas: senast 25 mars 2025

Tillåtna hjälpmedel

- R/R studio
- Miniräknare
- Språklexikon

Tentamen består av fem uppgifter på maximalt 20 poäng per uppgift. För maximal poäng på en uppgift krävs tydliga, utförliga och väl motiverade lösningar. Lösningarna skrivs och lämnas in på papper. Ingen digital inlämning ska göras. R får användas som hjälpmedel under tentamen men alla steg i lösningarna ska tydligt redovisas på papper.

Tentan ska göras anonymt, kontrollera därför innan du lämnar in din tentamen att du bara angett din anonymkod och INTE skrivit ditt namn någonstans i tentan.

Uppgift 1 (20 poäng)

Antag att Y har följande täthetsfunktion.

$$f(y) = \begin{cases} cy, & 1 \leq y \leq 3, \\ 0, & \text{annars.} \end{cases}$$

- a) Bestäm konstanten c .
- b) Vad är sannolikheten att
 - i) $Y \leq 2$?
 - ii) $Y = 2$?
- c) Beräkna väntevärdet för Y .
- d) Beräkna medianen för Y .

Uppgift 2 (20 poäng)

Ett café säljer egen färskpressad apelsinjuice. Mängden apelsiner som går åt varje dag (Y) antas följa en exponentialfördelning med väntevärde 7 kilo.

- a) Bestäm fördelningsfunktionen för den dagliga åtgången av apelsiner.
- b) Beräkna sannolikheten att mängden apelsiner som går åt i caféet en dag överstiger 10 kilo.
- c) Kostnaden för juicen per dag beräknas enligt funktionen

$$U = 6Y + 100.$$

Bestäm fördelningsfunktionen för kostnaden.

Uppgift 3 (20 poäng)

Ett särskilt råmaterial klassificeras efter mått på fuktighet X (i procent) och orenhet Y (i procent). Antag att X och Y följer nedanstående simultana fördelning.

Y	X			
	1	2	3	4
1	0.05	0.05	0.15	0.10
2	0.10	0.20	0.30	0.05

- Beräkna väntevärdet och variansen för X respektive Y .
- Är X och Y stokastiskt oberoende?
- Beräkna korrelationen mellan X och Y .
- För att kunna använda råmaterialet krävs upphettning för att motverka fuktighet, och filterering för att bli av med orenhet. Det medför en extra kostnad per sats på $C = 2X + 10Y^3$ hundra kronor. Beräkna förväntad extra kostnad.

Uppgift 4 (20 poäng)

En produktionschef som ska besluta om att köpa in en tillverkningsmaskin av design A eller av design B, undersöker genom ett experiment tiden det tar att slutföra en viss uppgift på varje maskin. Åtta tekniker deltog i experimentet, där varje tekniker använde både maskin A och maskin B i en slumpmässig ordning. Tiderna (i sekunder) som krävs för att slutföra uppgiften anges i följande tabell.

Tekniker	Maskin A	Maskin B
1	32	30
2	40	39
3	42	42
4	26	23
5	35	36
6	29	27
7	45	41
8	22	21

- Testa med hjälp av ett lämpligt paramteriskt test, om det går att påvisa någon skillnad i slutförande-tid mellan maskin A och maskin B. Vilka antaganden krävs?
- Testa med hjälp av ett teckentest, om det går att påvisa någon skillnad i slutförande-tid mellan maskin A och maskin B. Beräkna p -värdet för teckentestet. Vilka antaganden krävs?

- c) De 8 teknikerna använder båda maskiner i experimentet. Ett alternativt upplägg vore att slumpmässigt tilldela 16 tekniker att använda antingen maskin A eller maskin B. Vilket av uppläggen är att föredra? Förklara varför.

Uppgift 5 (20 poäng)

Studenter kan vända sig till *Ellicas statistikmottagning* för att få hjälp med sina statistikstudier. Under mottagningens öppettider anländer studenter enligt en slumpmässig process, där antal studenter som kommer till mottagningen under en timme (Y) antas följa en Poisson-fördelning.

- a) Antag att ett slumpmässigt urval av timmar görs där antal studenter observeras. Härled maximumlikelihood-estimatoren för väntevärdet i Poisson-fördelningen.
- b) Vad är MSE för maximumlikelihood-estimatoren?
- c) Vad är maximumlikelihood-estimatoren för sannolikheten att det inte kommer några studenter under en timme?
- d) Nämn en konjugerande prior för Poisson-modellen och förklara kort hur en bayesiansk estimator för väntevärdet kan tas fram.