

TMS136: Dataanalys och statistik – Tentamen 2011-10-21

Examinator och jour: Erik Jakobsson, tel. 031-772 53 79

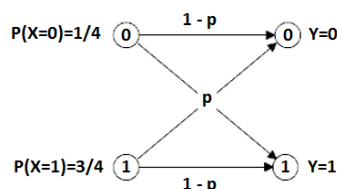
Hjälpmedel: Chalmersgodkänd räknare och formelsamling (formelsamling delas ut med tentan).

Betygsgränser: För betyg **3** krävs 14 poäng, för **4** krävs 20 poäng och för betyg **5** krävs 27 poäng. Totalt finns 35 poäng.

*Om du vill att jag skall mejla dig resultatet när jag rättat klart din tenta kan du skriva din mejladress på det **sista** pappret du lämnar in. Skriv i så fall gärna läsligt. Observera att tentan i detta fall kanske inte blir anonym, så strunta i det om det känns viktigt.*

Fullständiga och välmotiverade lösningar skall ges till varje uppgift.

- Definiera följande begrepp:
 - Slumpförsök. (1p)
 - Utfallsrum. (1p)
 - Stokastisk variabel. (1p)
 - Två stokastiska variabler X och Y är oberoende. (1p)
- Vad är sannolikheten att få kåk på given i en pokerhand? (3p)
- Vi tar emot bits som skickats till oss genom en kommunikationskanal. Varje bit kan dock förvanskas på vägen och ibland tar vi emot en 1:a när det egentligen var en 0:a som skickades, och tvärtom. Låt $X \sim \text{Ber}(3/4)$ vara den bit som skickas och Y den som tas emot. Som du ser i figuren så



Figur 1: Binär symmetrisk kanal

är sannolikheten att en bit förvanskas p .

- Bestäm Y s fördelning. (1p)
 - Om vi ser $Y = 0$, vad är sannolikheten att $X = 0$? (1p)
 - Vad är $\text{Cov}(X, Y)$? För vilket p är kovariansen 0? Vad tror du detta betyder för möjligheten att kommunicera genom kanalen? (2p)
 - För att försöka bestämma p skickar vi 10.000 nollor (oberoende) genom kanalen, och ser att 341 av dem kom fram som ettor. Ge en intervallskattning (ett konfidensintervall) för p . (2p)
- Låt T = tiden mellan två samtal till CSNs telefonväxel $\sim \text{Exp}(\lambda)$, med $\lambda = 1$ samtal per minut. Vad är sannolikheten att det kommer högst 55 samtal till CSN på en timma? *Motivera eventuella approximationer!* (3p)

5. Låt $X \sim Poi(\lambda)$. Och låt $X_1, \dots, X_n$ vara ett stickprov från X .
- Hitta momentskattaren (MOM) för λ . (1p)
 - Hitta maximum likelihood-skattaren (MLE) för λ . (2p)
 - För vilket k är $\hat{\lambda} = k \sum_{i=1}^{n-1} (X_{i+1} - X_i)^2$ en väntevärdesriktig skattare för λ ? (2p)

6. En viss produkt skall ha höjden 1195mm. På grund av olika fel i tillverkningsprocessen varierar höjden något. Anta att höjden för en slumpvis vald produkt är normalfördelad med väntevärde μ och varians σ^2 . Du mäter 5 produkter och får resultaten

1195.1 1197.7 1189.5 1195.7 1194.6

Testa om dessa data talar emot att $\mu = 1195$. (3p)

7. Antag att man har ett stickprov av storlek 20 från en normalfördelning med väntevärde μ och varians 2. Man vill testa $H_0: \mu = 10$ mot $H_1: \mu \neq 10$ på signifikansnivån $\alpha = 0.05$. Bestäm

$$\beta = P(\text{typ II fel}) = P(\text{inte förkasta } H_0 \mid \mu = 9)$$

om det verkliga värdet på μ är 9. (3p)

8. Du sommarjobbar på Kamratpostens frågespalt och har fått frågan "Är jag normal?" från en förtvivlad manlig läsare som lider av en generande åkomma. För att kunna svara sanningsenligt på frågan måste du hitta på ett sätt att uppskatta $p = \text{proportionen män som lider av åkomman}$. Åkomman är dock så generande att vi måste anta att de flesta män som lider av den inte gärna erkänner detta. Du funderar och kommer på att om du ger varje försöksperson ett mynt som de får singla i smyg och instruerar dem att

- Om det blir klave så måste de svara sanningsenligt
- Om det blir krona måste de svara "Ja, jag har åkomman!"

Du frågar n slumpmässigt valda män på detta sätt och får ett stickprov $X_1, \dots, X_n$, där $X_i = 1$ om person i svarar "Ja!" och $X_i = 0$ annars.

- Vad är $P(X_i = 1) = P(\text{En person svarar "Ja!"})$? (2p)
- Hitta på en bra skattare $\hat{p}$ för p . (2p)
- Är din $\hat{p}$ väntevärdesriktig och vad har den för varians? Och om du frågar många personer, vad har den för fördelning? (2p)
- Om du, för att försöka avgöra vad som är normalt (p_0) och inte, skulle göra ett test på $H_0: p = p_0$ mot $H_1: p < p_0$. Förklara hur du skulle gå till väga och vilken teststatistika du skulle använda? (2p)

(Förresten, detta kanske skulle bli ett bra projekt, med t.ex. otrohet bland chalmister (eller nåt annat snaskigt!))

Lycka till!