

TMS136: Dataanalys och statistik – Tentamen 2010-08-21

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 12 poäng, för **4** krävs 17 poäng och för betyg **5** krävs 24 poäng.

Fullständiga och välmotiverade lösningar skall ges till varje uppgift.

1. Förklara följande begrepp inom punktskattning:
 - (a) Stickprov. (1p)
 - (b) Skattare och skattning. (1p)
 - (c) Väntevärdesriktig (unbiased) skattare. (1p)
 - (d) Vad är ett lämpligt kriterium för val mellan två väntevärdesriktiga skattare? (1p)
2. Antag att händelserna A och B är oberoende. Visa att A och B^c är oberoende. Är A^c och B^c oberoende? (3p)
3. Betrakta den kontinuerliga stokastiska variabeln X med täthetsfunktion $f(x) = c \sin(x)$, för $0 \leq x \leq \pi$ där c är en konstant.
 - (a) Bestäm värdet på c . (1p)
 - (b) Vad har X för väntevärde? (1p)
 - (c) Vad har X för varians? (1p)
4. I de flesta städer brukar större taxibolag numrera sina bilar med $1, 2, \dots, N$. Under en promenad i en ny stad observerar du 7 taxibilar från ett och samma bolag med nummer
$$97, \quad 234, \quad 166, \quad 7, \quad 65, \quad 17, \quad 4.$$
 - (a) Vilka antaganden på dina observationer behöver du göra för att skatta bolagets totala antal taxibilar N ? (1p)
 - (b) Tag fram Maximum Likelihood-skattaren för antalet taxibilar och skatta N med den. (3p)
5. På en loppcircus tvingas loppan Leo gå på lina inför en engagerad publik. Leo är höjdrädd och uppträder därför både tveksamt och irrationellt. Varje sekund tar Leo ett steg (1 cm) framåt, med sannolikhet $1/2$, eller så står han still och samlar mod, även det med sannolikhet $1/2$. Leo kan inte välja att gå bakåt. Varje sekund i Leos liv kan lustigt nog anses oberoende.
 - (a) Vad är sannolikheten att Leo kommit över den 40 cm långa linan inom en minut? Motivera eventuella approximationer. (3)
 - (b) Ungefär hur många sekunder måste cirkusdirektören vänta för att kunna vara åtminstone 95% säker på att Leo har kommit över? (3)

6. I USA uppskattas fyra av tusen individer vara HIV-smittade. ELISA-testet används som standard för att testa en person för HIV. En stor studie vid 752 olika laboratorier i USA rapporterade att ELISA-testet har *sensitivitet* (andel positiva resultat bland smittade individer) 0.997 och *noggrannhet* (andel negativa resultat bland friska individer) 0.985. (Källa: Wikipedia)
- (a) Antag att en blodcentral låter samtliga donatorer som testas negativt för HIV få lämna blod. Utav en miljon (godtyckliga) amerikaner som kommer in till kliniken för donation, hur många förväntas få lämna blod trots att de är bärare av HIV? (1p)
 - (b) Om du anmäler dig för donation och får beskedet att ditt test var positivt, hur orolig behöver du vara (i procent)? (2p)
 - (c) ELISA-testet utvecklades för att på hög konfidensnivå kunna utesluta att donerat blod är HIV-smittat. Borde man därför blir förvånad om ovanstående sannolikhet är låg? Motivera ditt svar. (1p)
 - (d) I termer av hypotesprövning. Vad sätter ELISA-testet som nollhypotes respektive alternativ hypotes? Kan du uttrycka sannolikheten för *Typ 1* och *Typ 2* fel för testet i termer av ELISA-testets *sensitivitet* och *noggrannhet*. (2p)
7. Din personlige apelsinleverantör påstår att hans lådor med apelsiner innehåller ett Binomialfördelat antal ruttna apelsiner. Varje låda innehåller 24 apelsiner. Du tror ingalunda på honom och bestämmer dig för att ställa honom mot väggen. Du testar 75 lådor apelsiner med följande resultat:

39 lådor innehöll 0 ruttna apelsiner
23 lådor innehöll 1 ruttan apelsin
12 lådor innehöll 2 ruttna apelsiner
1 låda innehöll 4 ruttna apelsiner

Ställ upp en lämplig hypotes och testa. Rapportera alla data du anser viktiga. Diskutera resultatet, vad bör du säga till leverantören. (4)

Lycka till!