

Tentamenskrivning: TMS145 - Grundkurs i matematisk statistik och bioinformatik, 7,5p.

Tid: Tisdagen den 10 april, 2012 kl 14.00 - 18.00.

Examinator: Erik Kristiansson

Jour: Erik Kristiansson, tel 070-5259751.

Hjälpmedel: kalkylator, egen handskrivna formelsamling (fyra A4 sidor) samt med skrivningen utdelade tabellsidor.

Max är 32 poäng. För godkänt krävs minst 15 poäng, för betyget 4 krävs 21 poäng och för 5 krävs 26 poäng. Uppgifterna kommer inte i svårighetsordning.

1. Låt X vara en stokastisk variabel med täthetsfunktion

$$f_X(x) = 1 - \frac{x}{2}, \quad 0 < x < 2.$$

- (a) Beräkna väntevärdet för X .
- (b) Beräkna standardavvikelsen för X .
- (c) Låt X_1, \dots, X_{100} vara ett oberoende stickprov från en fördelning med täthetsfunktion $f_X(x)$. Låt

$$Y = \frac{X_1 + \dots + X_{100}}{100}.$$

Beräkna approximativt $P(|Y - E[Y]| > 0.1)$.

(5 p)

2. Till en telefonsupportavdelning inkommer två olika typer av ärenden: handhavandefel och reklamationer. Sannolikheten att ett inkommande ärende är en handhavandefel är 0,75 medan sannolikheten för att ett inkommande ärende är en reklamation är 0,25. Tiden det tar för att avsluta ett ärende antas vara exponentialfördelad där väntevärdet beror på typen av ärende. För handhavandefel är väntevärdet 4 och för reklamationer är väntevärdet 8.

- (a) Beräkna sannolikheten att ett ärende tar mer än 6 minuter att avsluta.
- (b) Beräkna sannolikheten för att ett ärende är av typen handhavandefel om det tar mer än 6 minuter att avsluta.

(4p)

3. Mängden mRNA av genen PROM1 misstänks vara sammankopplad med aggressiviteten hos en viss typ av tumörer. För att undersöka om så är fallet uppmättes mRNA-nivån hos PROM1 hos tio patienter med den aggressiva formen (x_1, \dots, x_{10}) och hos 8 patienter med den snällare formen (y_1, \dots, y_8) av tumören. Medelvärdet och stickprovstandardavvikelsen för de två grupperna beräknades till $\bar{x} = 7,76$ och $s_X = 2,02$ samt $\bar{y} = 5,82$ och $s_Y = 0,90$. Mätningarna från de två grupperna kan antas vara normalfördelade men med olika varians.
- (a) Undersök om det finns en skillnad i mRNA-nivå för PROM1 mellan de två grupperna genom att genomföra ett lämpligt statistiskt test. Glöm inte att formulera fördelningsantaganden samt noll- och alternativ-hypotes. Använd signifikansnivån $\alpha = 0.05$.
- (b) Beräkna testets p-värde.

(5 p)

4. Låt (X, Y) vara en tvådimensionell stokastisk variabel med täthetsfunktion

$$f_{X,Y}(x, y) = x + y, \quad 0 < x < 1, 0 < y < 1.$$

- (a) Är X och Y oberoende?
- (b) Beräkna korrelationen mellan X och Y . Tolka resultatet.
- (c) Beräkna $P(X > 2Y)$.

(5 p)

5. Låt X_1, \dots, X_n vara ett stickprov från en normalfördelning med väntevärde μ och känd varians σ^2 .

- (a) Härled ett konfidensintervall för μ . Konfidensgraden ska vara $1 - \alpha$.
- (b) Bestäm minsta antalet observationer som krävs för att konfidensintervallets längd inte överstiger $\sigma/2$.

(4 p)

6. Livslängden (x) hos en komponent misstänks vara exponentialfördelad med väntevärde μ . För att undersöka saken genomfördes ett test av 100 komponenter där livslängden observerades.

Livslängd	Antal
$0 \leq x \leq 10$	47
$10 < x \leq 20$	31
$x > 20$	22

Observationerna kan antas vara oberoende. Medellivslängden av de 100 komponenterna beräknades till $\bar{x} = 13,5$.

- (a) Använd ett χ^2 -test för att avgöra om observationerna kommer från en exponentialfördelning med $\mu = 10$.
- (b) Använd ett χ^2 -test för att avgöra om observationerna kommer från en exponentialfördelning.

(5 p)

7. Låt X_1, \dots, X_n vara ett stickprov från en normalfördelning med väntevärde μ och varians σ^2 . Täthetsfunktionen är då

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{\sigma^2}}, \quad -\infty < x < \infty.$$

- (a) Beräkna maximumlikelihoodskattningarna $\hat{\mu}_{ML}$ och $\hat{\sigma}_{ML}^2$ av μ och σ^2 .
- (b) Hur förhåller sig $\hat{\sigma}_{ML}^2$ till stickprovsstandardavvikelsen

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 ?$$

Är maximumlikelihoodskattaren av σ^2 väntevärdesriktig?

(4 p)