

TENTAMEN: Matematisk statistik för K (TMA073 och TMA072)

Fredagen den 25 maj 2012, kl 14:00-18:00

Lärare och jour: Aila Särkkä, telefon 772 3542

Hjälpmedel: Formelsamling, tabeller (även BETA, Physics Handbook, skoltabeller, t.ex. TEFYMA), valfri miniräknare med tömda minnen.

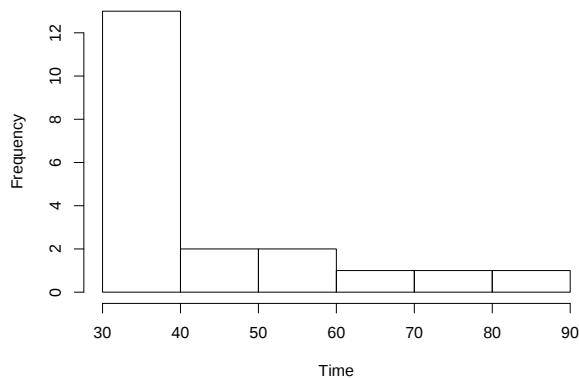
- 1) Ett automatiskt sprinklersystem har två aktiveringsapparater. En av apparaterna fungerar (dvs. att den aktiverar sprinklern) med sannolikhet 0.9 och den andra fungerar med sannolikhet 0.8 oberoende av den första. Om minst en av apparaterna fungerar, aktiveras sprinklern vid brand. Antag att det börjar brinna nära sprinklern.
 - a) Vad är sannolikheten att sprinklern kommer att aktiveras?
 - b) Vad är sannolikheten att sprinklern inte kommer att aktiveras?
 - d) Vad är sannolikheten att bara den apparaten som har sannolikhet 0.9 att fungera, fungerar? (3p)
- 2) Låt X vara en kontinuerlig stokastisk variabel med fördelningsfunktionen

$$F(x) = \begin{cases} 0 & \text{för } x < 0 \\ \frac{1}{4}x^3 & \text{för } 0 \leq x \leq 1 \\ \frac{3}{4}x - \frac{1}{2} & \text{för } 1 < x \leq 2 \\ 1 & \text{för } x > 2 \end{cases}$$

- a) Vad är sannolikheten att värdet av X är mellan 0.5 och 1.5?
 - b) Antag att vi har 10 oberoende stokastiska variabler från samma fördelning som X . Vad är sannolikheten att exakt 2 av dem antar ett värde som är minst 1.5? (4p)
- 3)
 - a) Härleda ett 90% konfidensintervall för väntevärdet baserad på ett stickprov av storlek n från $N(\mu, \sigma)$ -fördelningen då variansen inte är känd.
 - b) Man är intresserad av att veta den genomsnittliga nivån av en viss hormon. Man har mätt hormonnivån från 15 personer och fått medelvärdet 37.3 (pg/mL) och stickprovsstandardavvikelsen 13.9 (pg/mL). Hitta ett 90% konfidensintervall för den genomsnittliga hormonnivån. Ge antaganden som du har gjort.
 - c) I b) gjordes mätningarna under dagen. Enligt en tidigare studie är den genomsnittliga hormonnivån under natten 25.5 (pg/mL). Använd intervallet i b) för att säga om hormonnivån under dagen verkar skillja sig från hormonnivån under natten. Förklara. (5p)

- 4) Man kan använda minsta kvadratsmetoden för att skatta regressionsparametrarna i enkel linjär regression.
- Vad är iden bakom metoden?
 - Vilka antaganden måste man göra för att kunna använda metoden?
 - Varför plottar man residualplottar? (4p)
- 5) Man undersöker hur länge det tar att få någons parkeringsplats (dvs. hur länge det tar att den som har platsen man väntar på kör i väg). Man tror att det tar högst 50 sekunder. För att testa detta har man mätt tiden (i sekunder) det tar för 20 bilar och fått

49.48 43.30 85.97 46.92 49.18 79.30 47.35 46.52 59.68 42.89 49.29 68.69
41.61 46.81 43.75 46.55 42.33 71.48 78.95 42.06



Medelvärde av väntetiderna är 54.11 och stickprovsvariansen är 207.19.

- Använda T eller Z testet (det som är lämpligast) för att testa påståendet att det tar högst 50s att få platsen man väntar på. Använd signifikansnivån 5%.
- Testa hypotesparet i a) genom att använda teckentestet för medianen. Använd signifikansnivån 5%.
- Jämför resultaten i a) och b).
- Ge antaganden du har gjort i a) och b).
- Vilket test skulle du rekommendera? Varför? (5p)

- 6) Taxiföretagets ledning vill bestämma om de vill byta däcktyp (från gammal till ny) för att minska bränslekonsumption. Var och en av de 10 taxibilarna utrustades av en av däcktyperna och föraren gjorde en testkörning. Efter körningen byttes däcken till den andra typen och testkörningen upprepades. Man mätte antalet "miles per gallon" och resultaten av testkörningen var följande:

Bil	Ny däcktyp (x_1)	Gammal däcktyp (x_2)
1	32	27
2	36	31
3	32	30
4	29	26
5	34	29
6	31	28
7	33	30
8	34	31
9	35	33
10	32	29

($\bar{x}_1 = 32.8$, $s_1^2 = 4.18$, $\bar{x}_2 = 29.4$, $s_2^2 = 4.27$, $\bar{d} = 3.4$ och $s_D^2 = 1.38$, där $D = X_1 - X_2$)

Taxiföretaget vill inte byta däcktyp om man inte kan hitta starkt bevis om att man kan köra fler "miles per gallon" med de nya däcken än med de gamla.

- Vilket test (parvist eller baserat på oberoende stickprov) bör utföras? Varför?
- Ge hypoteserna och utför ett hypotestest för att testa om den nya däcktypen ökar mängden "miles per gallon". Använd signifikansnivån 0.01.
- Skulle man ha färkastat nollhypotesen med signifikansnivån 0.05? Varför/varför inte? (4p)

- 7) Man analyserar dricksvattnet i två olika delar av staden för att undersöka om blynivåerna är olika. 100 vattenprov har tagits i vardera stadsdel. Medelvärde av blynivån är 34.1 (parts per million) och stickprovsstandardavvikelsen 5.9 (parts per million) i den första stadsdelen. Motsvarande siffror i den andra stadsdelen är 36.0 (parts per million) och 6.0 (parts per million).
- a) Utför ett signifikanstest för att testa om den genomsnittliga blynivån skilljer sig mellan de två stadsdelarna. Är skillnaden statistiskt signifikant på signifikansnivån 5%?
 - b) Hitta ett 95% konfidensintervall för differensen av de två blynivåerna. Verkar blynivåerna skillja sig från varandra enligt konfidensintervallet?
 - c) Vilka antaganden har du gjort i a) och i b)?
 - d) Antag att stadingenjörerna blir oroliga om de upptäcker en skillnad som är större än 5 (parts per million). Baserad på konfidensintervallet i b), är den statistiska signifikansen i a) av praktisk signifikans för ingenjörerna? Förklara. (5p)

Lycka till!