TENTAMEN: Matematisk statistik för K (20 augusti, 2004)

Kortfattade lösningar:

1) a) $P(A \cap B) = P(A)P(B)$ och då är $P(B) = P(A \cap B)/P(A) = \frac{3}{4}$
$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.85$

b) $P(A|B) = P(A \cap B)/P(B) = 0.6$ och man får att
$P(B) = P(A \cap B)/P(A|B) = \frac{1}{2}$ och då är $P(A \cup B) = 0.6$

2) a) Man kan säga någonting om noggranheten av skattningen.

b) Man kan inte säga att det finns en skillnad mellan pojkar och flickor därför att konfidensintervallet innehåller värdet 0.

3) Låt X_1 vara korrosionshastigheten av kokillhärdat gjutjärn med 0.1% krominnehåll och X_2 korrosionshastigheten med 0.2% krominnehåll. Man antar att X_1 och X_2 är normalfördelade med okända varianser $\sigma_1^2 = \sigma_2^2$.

a) Testa hypotesen $H_0 : \mu_1 = \mu_2$ mot $H_1 : \mu_1 \neq \mu_2$. Man får att $S_p^2 = 0.0935$, $T = (\bar{X}_1 - \bar{X}_2)/S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \sim T_{23}$ och att T får ett värde 1.356. Nu är $t = 1.356 < 1.714 = t_{.05}$ så att H_0 kan inte förkastas på signifikansnivån 0.1 (eller mindre). Dvs. att krominnehållet verkar inte påverka korrositionshastigheten. (Du kan också räkna p-värdet om du vill.)

b) $\bar{x}_1 - \bar{x}_2 \pm t_{0.005} s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = 0.165 \pm 0.355$. Därför att $0 \in [-0.17, 0.5]$ kan man säga att det finns ingen skillnad mellan de olika krominnehållen enligt 99% konfidensintervallet.

c) Man drar samma slutsats i a) och b): ingen skillnad mellan korrositionshastigheterna beroende på krominnehållet.

4) a) $H_1 : \mu \neq 1$.

b) $T = \frac{\bar{X}-1}{S/\sqrt{n}} \sim T_{15}$, där $S = \sqrt{\frac{1}{n-1}\sum(X_i - \bar{X})^2}$ (X är vikten av frös) och dess värde är 3.999.

c) p=0.0012 vilket är väldigt litet. Man skulle förkasta H_0 på signifikansnivå 0.12% och högre

5) $\sigma_A = \sigma_B = 1$

a) Man testar $H_0 : \mu_A = \mu_B$ mot $H_1 : \mu_A \neq \mu_B$ genom att använda teststatistikan $Z = (\bar{X}_A - \bar{X}_B)/\sigma\sqrt{\frac{1}{10} + \frac{1}{12}}$. Då är teststatistikan $Z \sim N(0,1)$ och den får värdet 2.009. Då är $P(Z > 2.009|H_0 \text{ sann}) = 1 - P(Z \leq 2.009|\mu = 0) = 1 - \Phi(2.009) = 1 - 0.9778 = 0.0222$, där Φ är fördelningsfunktionen av den standardiserade normalfördelningen. Det exakta dubbelsidiga p-värdet är $2 \cdot 0.0222 = 0.044$.

b) $X_A \sim N(\mu_A, 1)$ och $X_B \sim N(\mu_B, 1)$.

6) Låt X vara densiteten. Då är $X \sim N(0.0046, 9.6 \cdot 10^{-8})$.

a) $P(0.004 < X < 0.005) = P(\frac{0.004-0.0046}{0.0003} < \frac{X-0.00046}{0.0003} < \frac{0.005-0.0046}{0.0003}) = P(-2 < Z < 1.33) = \Phi(1.33) - \Phi(-2) = 0.8854$, där $Z \sim N(0,1)$ och Φ är fördelningsfunktionen av den standardiserade normalfördelningen.

b) $P(X > x_0) = 0.05$ och man borde hitta x_0.
$P(X \leq x_0) = P(\frac{X-0.0046}{0.0003} \leq \frac{x_0-0.0046}{0.0003}) = \Phi(\frac{x_0-0.0046}{0.0003}) = 0.95 = \Phi(1.645)$ och $x_0 = 0.0051$.

7) a) Plotta m i x-axeln och f i y-axeln.

b) $f = -3.05 + 0.108m$ dvs. att desto fler fågelholkar besatta av mesar desto fler döda flugsnappare.

c) Testa $H_0 : \beta_1 = 0$ mot $H_1 : \beta_1 \neq 0$. Teststatistikan får ett värde $b_1/(s/\sqrt{\sum(m_i - \bar{m})^2}) = 1.159$, där $s^2 = \sum(f_i - b_0 - b_1 m_i)^2/(n-2) = 14.315$. Teststatistikan är T_{12}-fördelad och jämförs med värde 2.179. Nu är $t_{0.025} = 2.179 > 1.159$ och man kan inte förkasta H_0 på signifikansnivå 0.05 eller mindre. Man kan inte säga att β_1 skiljer sig från 0, dvs. att det inte finns signifikant regression.