TENTAMEN: Matematisk statistik för K (28 maj, 2010)

Kortfattade lösningar:

1) a) $P(A \cup B) = \frac{113}{500}$

b) $P(A \setminus B) = P(A) - P(A \cap B) = \frac{40}{500}$

c) $P(A^C \cup B^C) = 1 - P(A \cap B) = \frac{458}{500}$

2) a) Sannolikheten att nollhypotesen förkastas även om den ä sann.

b) Sannolikheten att nollhypotesen förkastas då mothypotesen är sann.

c) Man vill att signifikansnivån är nära 0 och styrkan nära 1. Man behöver ett stort stickprov fï att kunna garantera detta.

3) a) $\bar{x} \pm t_{0.01}^{(n-1)} S/\sqrt{n} = 3410.14 \pm 2.998 \cdot 1.018/\sqrt{8} = 3410.14 \pm 1.08$

b) Med 98% sannolikhet har man hittat ett intervall som täcker den sanna smältpunkten.

c) Mätningarna är ett stickprov från $N(\mu, \sigma)$.

d) Intervallet skulle bli kortare därför att variansen av medelvärdet skulle minska.

4) Låt vikten av burk vara en stokastisk variabel X.

a) $X \sim N(\mu, 10)$. Bestäm μ genom att beräkna $P(X \geq 340) = 0.90$. Man får att $\Phi((340 - \mu)/10) = 0.10$ och $(340 - \mu)/10 = -1.28$, vilket ger $\mu = 352.8$g. Medelvikten är $\mu = 352.8$g.

b) Låt Y v ara antalet burkar som väger minst 340g. Då är $Y \sim Bin(10, 0.9)$ och $P(Y > 3) = 1 - P(Y \leq 3) = 1 - \sum_{y=0}^{3} \binom{10}{y} 0.9^y (1 - 0.9)^{10-y} \approx 1$.

5) a) Man vill testa $H_0 : \mu_A = \mu_B$ mot $H_1 : \mu_A < \mu_B$ genom att använda ett 2 stickprovs T-test. Teststatistikan är $t = (\bar{x}_A - \bar{x}_B)/(S_p\sqrt{1/n_A + 1/n_b})$, där $S_P^2 = ((n_A - 1)S_A^2 + (N_B - 1)S_B^2)//n_A + n_B - 2) = 158.31$. Om H_0 är sann är teststatistikan T_{173}-fördelad, dvs. N(0,1)-fördelad. Teststatistikans värde blir -1.04, vilket inte är på kritiska området (< -1.645 eller > 1.645). Man kan inte förkasta H_0 på signifikansnivån 0.05 och kan inte säga att läkemedel B sänker kolesterolnivån mer än läkemedel A.

b) Varianserna lika, dvs. $\sigma_A^2 = \sigma_B^2$; två oberoende stickprov.

6) a) Ett parvist T-test. Samma bilar har används i de två temperaturen. $H_0 : \mu_{40} = \mu_{80}$ mot $H_0 : \mu_{40} \neq \mu_{80}$.

b) Rad 1: Ett dubbelsidigt parvist T-test med signifikansnivån 5%
Rad 2-3: Mothypotesen accepteras (och H_0 förkastas)
Rad 4-5: p-värdet blir 0.0239
Rad 6-7: Ett 95% konfidensintervall f ör skillnaden av de två väntevärdena är $(0.0060, 0.0666)$
Rad 8-1: Man har valt att skriva ut teststatistikans värde 2.7130, antal frihetsgrader 9 och standardavvikelsen för differensen av medelvärdena 0.0423

c) $2P(T > 2.7130|H_0 \text{ sann})$ där $T = \bar{d}/(s_D/\sqrt{10})$, $\bar{d}$ är medelvärdet av deifferenserna och s_D standardavvikelsen av differenserna.

d) På signifikansnivån 5% kan man förkasta H_0, och det verkar som den genomsnittliga emissionen skiljer sig mellan de två temeraturerna. Det verkar som den var större i 40°F än i 80°F.

7) a) För att mäta linjärt samband mellan de två poängtalen.

b) Det verkar som det finns positiv linjär beroende mellan de två poängtalen, dvs. att när en ökar, ökar också den andra.

c) Därför att intervallet täcker 0 verkar det som att det inte finns ett linjärt samband mellan de två poängtalen.

d) Inga (utom att man har två stokastiska variabler man tittar på).