Soluzioni Tutorato di Statistica 1 del 06/05/2010
Docente: Prof.ssa Enza Orlandi
Tutore: Dott.ssa Barbara De Cicco

Esercizio 1.
$X \sim Bernoulli(\theta)$; $\alpha = 0,0547$; $H_0 : \theta = \frac{1}{2}$; $H_1 : \theta = \frac{1}{4}$.

1. Per calcolare il test più potente usiamo il lemma di Neyman Pearson e quindi otteniamo:
   $\frac{L(1/2)}{L(1/4)} = \frac{\prod_{i=1}^{10} p_0^{x_i}(1-p_0)^{1-x_i}}{\prod_{i=1}^{10} p_1^{x_i}(1-p_1)^{1-x_i}} = \frac{(1/2)^{10}}{(1/4)^{10} 3^{10-\sum_{i=1}^{10} x_i}} \le k^*$
   $\iff \frac{2^{20}}{2^{10} 3^{10-\sum_{i=1}^{10} x_i}} \le k^* \iff -(10 - \sum_{i=1}^{10} x_i) log3 \le log(\frac{k^*}{2^{10}}) = k' \iff$
   $\sum_{i=1}^{10} x_i \le k$
   Allora la regione critica trovata è $C = \{(x_1, ..., x_{10}| \sum_{i=1}^{10} x_i \le k)$.
   Imponendo l'ampiezza del test otteniamo $\alpha = 0,0547$.
   $\alpha = P(RifiutareH_0|\theta_0) = P(\sum_{i=1}^{10} x_i \le k|\theta_0) = \sum_{i=0}^{k} \binom{k}{i} (\frac{1}{2})^i (\frac{1}{2})^{10-i}$
   L'uguaglianza risulta vera, come si può verificare dalle tavole della binomiale, per $k = 2$.
   Dunque: $C = \{(x_1, ..., x_{10}| \sum_{i=1}^{10} x_i \le 2)$.
   La potenza del test per $\theta = 1/4$ è:
   $\sum_{i=0}^{2} \binom{10}{i} (\frac{1}{4})^i (\frac{3}{4})^{10-i}$

2. $\pi_Y(\theta) = P(RifiutareH_0|\theta) = P(\sum_{i=1}^{10} x_i \ge 6) = \sum_{i=6}^{10} \binom{10}{i} \theta^i (1-\theta)^{10-i}$
   $sup_{\theta \le 1/2} \pi_Y(\theta) = \sum_{i=6}^{10} \binom{10}{i} (\frac{1}{2})^i (\frac{1}{2})^{10-i} = 0,377$

Esercizio 2.
X con densità $f_X(x) = \theta x^{\theta-1} 1_{(0,1)}(x)$

1. Usiamo il lemma di Neyman Pearson ed otteniamo:
   $\frac{L(\theta_0)}{L(\theta_1)} = \frac{\theta_0^2 (x_1 x_2)^{\theta_0 - 1} 1_{(0,1)}(x_1) 1_{(0,1)}(x_2)}{\theta_1^2 (x_1 x_2)^{\theta_1 - 1} 1_{(0,1)}(x_1) 1_{(0,1)}(x_2)} = \frac{1}{4 x_1 x_2} \le k^* \iff \frac{1}{x_1 x_2} \le 4k^{k*} = k$
   Allora le regione critica trovata è: $C\{(x_1, x_2)|x_1 x_2 \ge k\}$.
   Troviamo quindi k imponendo l'ampiezza del test:
   $\alpha = \frac{1}{2}(1-log2) = P(RifiutareH_0|\theta_0) = P(x_1 x_2 \ge k|\theta_0) = \int_k^1 dx_1 \int_{k/x_1}^1 dx_2 = 1 - k - klogk$
   L'uguaglianza risulva vera per $k = 1/2$, allora il test più potente è individuato dalla regione C:
   $C\{(x_1, x_2)|x_1 x_2 \ge 1/2\}$.

2. $\pi_Y(\theta) = P(RifiutareH_0|\theta) = P(\frac{3}{4} x_1 \le x_2) = \int_0^1 dx_1 \int_{\frac{3}{4} x_1}^1 \theta^2 x_1^{\theta-1} x_2^{\theta-1} dx_2 = 1 - \frac{1}{2}(\frac{3}{4})^\theta$
   $sup_{\theta \le 1} \pi_Y(\theta) = sup_{\theta \le 1}[1 - \frac{1}{2}(\frac{3}{4})^\theta] = 5/8$

Esercizio 3.
I risultati dei due test di matematica sono distribuiti rispettivamente come $N(\mu_X, \sigma^2)$, e $N(\mu_Y, \sigma^2)$ con σ^2 non nota. Il campione di ampiezza $n = 9$ di studenti del primo liceo ha dato come risultato $\bar{x} = 81,31$ e $S_x^2 = 60,76$. Dal sendondo liceo invece è stato estratto un campione di ampiezza $m = 15$ di studenti ed ha dato come risultato $\bar{y} = 78,61$ e $S_y^2 = 48,24$. L'intervallo di confidenza al 95% per $(\mu_x - \mu_y)$è

dato da:
$T = \frac{Z}{\sqrt{U/(m+n-2)}} \sim t(m+n-2)$
$\gamma = P(-t_{(1+\gamma)/2} \leq T \leq t_{(1+\gamma)/2}) = P(\bar{X} - \bar{Y} - t_{(1+\gamma)/2} S_p \sqrt{1/n + 1/m} \leq \mu_x - \mu_y \leq \bar{X} - \bar{Y} + t_{(1+\gamma)/2} S_p \sqrt{1/n + 1/m})$
Dove $S_p = \sqrt{\frac{(n-1)S_x^2 + (m-1)S_y^2}{m+n-2}}$.
Inoltre dalle tavole otteniamo che $t_0 = t_{0,025}(22) = 2,074$.
Quindi l'intervallo di confidenza per $\mu_x - \mu_y$ al 95% è : $(-3,65; 9,05)$.

Esercizio 4.
È dato un campione di v.a. di ampiezza 16 da una distribuzione $N(\mu, 25)$ e sia $\bar{X} = 73.8$
$Q = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$
$0,95 = P(-q_1 \leq Q \leq q_1) = P(-q_1 \frac{5}{4} \leq \bar{X} - \mu \leq q_1 \frac{5}{4}) = P(\bar{X} - q_1 \frac{5}{4} \leq \mu \leq \bar{X} + q_1 \frac{5}{4})$
Dalle tavole si ricava che $q_1 = 1,96$ e quindi l'intervallo cercato è: $(71,35; 76,25)$.