Tutorato di Statistica 1 del 13/05/2010
Docente: Prof.ssa Enza Orlandi
Tutore: Dott.ssa Barbara De Cicco

Esercizio 1.

$X_1, ..., X_n$ è un campione casuale da $N(\mu, 25)$.
Il test più potente è dato dal lemma di Neyman Pearson:
$\frac{L(\mu_0)}{L(\mu_1)} = \frac{\prod_{i=1}^n \frac{e^{-(x-\mu_0)^2/50}}{\sqrt{50\pi}}}{\prod_{i=1}^n \frac{e^{-(x-\mu_1)^2/50}}{\sqrt{50\pi}}} =$
$e^{-1/50\{\sum_{i=1}^n (x_i-\mu_0)^2 - \sum_{i=1}^n (x_i-\mu_1)^2\}}$
Passando al logaritmo si ottiene:
$-\{\sum_{i=1}^n (x_i - \mu_0)^2 - \sum_{i=1}^n (x - \mu_1)^2\} \leq 50k \Longleftrightarrow \frac{\sum_i x_1}{n} \leq k^*$
L'ampiezza del test è:
$\alpha = 0,025 = P(Rifiutare H_0|\mu_0) = P(\bar{X} \leq k^*|\mu_0) = P(\frac{\bar{X}-\mu_0}{5/\sqrt{n}} \leq \frac{k^*-\mu_0}{5/\sqrt{n}}) = \Phi(\frac{k^*-\mu_0}{5/\sqrt{n}})$
Dalle tavole si ricava che $\frac{k^*-\mu_0}{5/\sqrt{n}} = -1,96$ da cui $k^* = 10 - \frac{9,8}{\sqrt{n}}$.
$\beta = P(Accettare H_0|\mu_1) = P(\bar{X} > k^*|\mu_1) = P(\frac{\bar{X}-\mu_1}{5/\sqrt{n}} > \frac{k^*-\mu_1}{5/\sqrt{n}}|\mu_1) = 1 - \Phi(\frac{k^*-\mu_1}{5/\sqrt{n}})$
Dalle tavole si ricava che $\frac{k^*-\mu_1}{5/\sqrt{n}} = 1,96$ quindi $k^* = 5 + \frac{9,8}{\sqrt{n}}$.
Ora $\alpha = \beta$ quindi risolvendo otteniamo:
$10 - \frac{9,8}{\sqrt{n}} = 5 + \frac{9,8}{\sqrt{n}}$
Da cui $n = 16$.

Esercizio 2.

x è una singola osservazione proveniente da una variabile casuale $N(\mu, \sigma^2)$.

1. Si assuma: $\theta_0 = \{\mu_0 = 0, \sigma_0^2 = 1\}$, $\theta_1 = \{\mu_1 = 1, \sigma_1^2 = 2\}$ $P(Rifiutare H_0|\theta_0) = P(|X| > 2|\theta_0) = P(X < -2) + P(X > 2) = 2\Phi(-2) = 0,0456$
   La potenza del test sotto H_1 è data da:
   $P(Rifiutare H_0|\theta_1) = P(|X| > 2|\theta_1) = P(\frac{X-1}{\sqrt{2}} < \frac{-3}{\sqrt{2}}) + P(\frac{X-1}{\sqrt{2}} > \frac{1}{\sqrt{2}}) = \Phi(\frac{-3}{\sqrt{2}}) + \Phi(\frac{-1}{\sqrt{2}}) = 0,2594.$

2. Dal lemma di Neyman Pearson si ottiene:
   $\frac{L(\theta_0)}{L(\theta_1)} = \sqrt{2}e^{-\frac{x^2+2x-1}{4}} \leq k^* \Longleftrightarrow -(x^2 + 2x - 1) \leq 4log(k^*/2) = k_1 \Longleftrightarrow x^2 + 2x - 1 \geq k_1$
   Allora la regione critica è:$C = \{x|x^2 + 2x - 1 \geq k_1\}$.

Esercizio 3.

X è l'ammontare di succo d'arancia in grammi per giorno consumato dagli americani. $\sigma = 96$; $n = 576$ e $\bar{X} = 133$.
$Q = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ Allora:
$0,90 = P(-q_1 < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < q_1) = P(-1,645\sigma/\sqrt{n} < \bar{X} - \mu < 1,645\sigma/\sqrt{n}) = P(\bar{X} - 1,645\sigma/\sqrt{n} < \mu < \bar{X} + 1,645\sigma/\sqrt{n}) = P(133 - 1,645\frac{96}{\sqrt{576}} < \mu < 133 + 1,645\frac{96}{\sqrt{576}})$
Allora l'intervallo per μ al 90% è: $(126,42; 139,58)$.

Esercizio 4.
Siano $X_1, ..., X_n$ e $Y_1, .., Y_m$ i rispettivi campioni casuali di ampiezza $n = 15$ ed $m = 8$ per le distribuzioni $N(\mu_x, \sigma_x^2)$ ed $N(\mu_y, \sigma_y^2)$. I campioni sono indipendenti e tali che $\bar{x} = 70.1$, $\bar{y} = 75.3$, $\sigma_x = 60$, $\sigma_y = 40$. Troviamo l'intervallo di confidenza al 90% per $\mu_x - \mu_y$.
$0,90 = P(-z_{0,05} < \frac{(\bar{X}-\bar{Y})-(\mu_x-\mu_y)}{\sqrt{\frac{\sigma_x^2}{n}+\frac{\sigma_y^2}{m}}} < z_{0,05}) =$
$P((\bar{X}-\bar{Y})-z_{0,05}\sqrt{\frac{\sigma_x^2}{n}+\frac{\sigma_y^2}{m}} < \mu_x-\mu_y < (\bar{X}-\bar{Y})+z_{0,05}\sqrt{\frac{\sigma_x^2}{n}+\frac{\sigma_y^2}{m}})$ Allora l'intervallo di confidenza è: $(70,1-75,3-1,645\sqrt{\frac{60^2}{15}+\frac{40^2}{8}}; 70,1-75,3+1,645\sqrt{\frac{60^2}{15}+\frac{40^2}{8}}) =$
$(-10,135; -0,265)$

Esercizio 5.
$X_1, ..., X_n$ è un campione casuale di ampiezza n estratto da $f(x;\theta) = \theta^2 x e^{-\theta x} 1_{(0,+\infty)}(x)$
Osserviamo che le variabili sono delle $\Gamma(2, \theta)$. Inoltre la densità appartine alla famiglia esponenziale, dove: $a(\theta) = \theta^2$,$b(x) = x1_{(0,+\infty)}(x)$, $c(\theta) = -\theta$ funzione decrescente in θ e $d(x) = x$.
Allora $T = \sum_{i=1}^n X_i$ è la statistica.
$T = \sum_{i=1}^n X_i \sim \Gamma(2n, \theta)$, allora:
$\alpha = P(T \leq k^*) = P(\sum_{i=1}^n X_i \leq k^*)$.
Allora la regione critica del test uniformemente più potente è:
$C = \{(x_1, ..., x_n) : T \leq k^*\}$ con k^* che verifica la relazione precedente, ovvero:
$sup_{\theta \leq \theta_0} int_0^{k^*} \frac{\theta^{2n}}{\Gamma(2n)} x^{2n-1} e^{\theta x} dx =$
$\int_0^{k^*\theta} \frac{y^{2n-1}}{\Gamma(2n)e^{-y}} dy =$
Avendo eseguito il cambio di variabili: $y = \theta x$. Allora:
$= int_0^{k^*} \Gamma(2n, \theta_0)(x) dx$