Soluzioni del Tutorato di Statistica 1 del 06/12/2010
Docente: Prof.ssa Enza Orlandi
Tutore: Dott.ssa Barbara De Cicco

Esercizio 1.

Sia $X_1, ..., X_n$ un c.c. estratto da Poisson di parametro λ. Trovare il test più potente di ampiezza α per $H_0 : \lambda = \lambda_0$ contro $H_1 : \lambda = \lambda_1$ con $\lambda_0 < \lambda_1$. Come cambia il test nel nel caso $\lambda_0 > \lambda_1$?

$$f(x, \theta) = \frac{\lambda^x e^{-\lambda}}{x!} 1_{(0,1,2..)}(x)$$

Il test più potente di ampiezza α é : $\lambda(x_1, x_2, ..) = \frac{L(\theta_0)}{L(\theta_1)} \leq k^*$
$\lambda(x_1, x_2, ..) = \frac{\prod_{i=1}^n \lambda_0^{x_i} e^{-\lambda_0} \frac{1}{x_i!} 1_{(0,1,..)}(x_i)}{\prod_{i=1}^n \lambda_1^{x_i} e^{-\lambda_1} \frac{1}{x_i!} 1_{(0,1,..)}(x_i)} = \left(\frac{\lambda_0}{\lambda_1}\right)^{\sum_{i=1}^n x_i} e^{-n(\lambda_0 - \lambda_1)} \leq k^*$
passando al logaritmo otteniamo:

$$log\left(\frac{\lambda_0}{\lambda_1}\right) \sum_{i=1}^n x_i \leq logk^* + n(\lambda_0 - \lambda_1)$$

$$\sum_{i=1}^n x_i \geq \frac{logk^* + nlog(\lambda_0 - \lambda_1)}{log\left(\frac{\lambda_0}{\lambda_1}\right)}$$

da cui $\sum_{i=1}^n x_i \geq k'$.

Allora la regione critica per $\lambda_0 \leq \lambda_1$ è: $C^* = \{x_1, x_2, .. : \sum_{i=1}^n x_i \geq k'\}$
Il test più potente di ampiezza α è: si rifiuti $H_0 \iff \sum_{i=1}^n \geq k'$

$$\alpha = P(rifiuto H_0) = P(\sum_{i=1}^n x_i \geq k') = 1 - P(\sum_{i=1}^n x_i \leq k')$$

poichè $\forall i, X_i \sim Po(\lambda_0)$ allora $\sum_{i=1}^n x_i \sim Po(n\lambda_0)$ quindi k' è il quantile della $Po(n\lambda_0)$ di livello $1 - \alpha$.
Nel caso in cui $\lambda_0 > \lambda_1$ si procede in maniera analoga sempre applicando il lemma di Neyman-Pearson, il test diventerà: si rifiuti $H_0 \iff \sum_{i=1}^n < k'$ quindi
$\alpha = P(\sum_{i=1}^n < k')$ dove k' è il quantile della $Po(n\lambda_0)$ di livello α.

Esercizio 2.

Sia X una singola osservazione dalla densità: $f(x,\theta) = \theta x^{\theta-1} 1_{(0,1)}(x)$, $\theta > 0$

1. Nel verificare $H_0 : \theta \leq 1$ in alternativa a $H_1 : \theta > 1$ determinate la funzione di potenza e l'ampiezza di un test del tipo: si rifiuti H_0 se e solo se $X \geq 1/2$.
   La funzione di potenza è:

   $\Pi_Y(\theta) = P(X \geq \frac{1}{2}) = \int_{\frac{1}{2}}^{1} \theta x^{\theta-1} 1_{(0,1)}(x) dx = 1 - \left(\frac{1}{2}\right)^{\theta}$
   $sup_{\theta \geq 1} \pi(\theta) = 1/2.$

2. Determinate un test più potente di ampiezza α per $H_0 : \theta = 2$ in alternativa a $H_1 : \theta = 1$.
   Applichiamo il lemma di Neyman-pearson:

   $\lambda(X_1) = \frac{\theta_0 x^{\theta_0 - 1}}{\theta_1 x^{\theta_1 - 1}} = 2x \leq k$

   Da cui $x \leq k^*$, dove $k^* = k/2$.
   La regione critica è quindi $C^* = \{x : x \leq k^*\}$.
   L' ampiezza del test è:
   $\alpha = P(rifiutareH_0|H_0) = P(X \leq k^*) = \int_0^{k^*} \theta x^{\theta-1} = (k^*)^{\theta}$ quindi $k^* = \sqrt{\alpha}$.

Esercizio 3.

Sia $X_1, ..., X_n$ un campione casuale di ampiezza n estratto da:

$$f(x;\theta) = \theta^2 x e^{-\theta x} 1_{(0,+\infty)}(x)$$

Vedere se esiste un test uniformemente più potente di ampiezza α per verificare: $H_0 : \theta \leq 1$ contro $H_1 : \theta > 1$

Osserviamo che le variabili sono delle $\Gamma(2,\theta)$. Inoltre la densità appartiene 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:

$= \int_0^{k^*} \Gamma(2n, \theta_0)(x) dx.$

Esercizio 4.

Siano $X_1, ..., X_n \sim f(x, \theta) = (\theta + 2)x^{\theta+1}$ con: $0 < x < 1$ e $\theta > 2$.
Si vuole testare il test di ipotesi: $H_0 : \theta \leq \theta_0$ contro $H_1 : \theta > \theta_0$. Trovare il test uniformemente più potente di ampiezza α.
Riscriviamo la funzione di densità:

$f(x, \theta) = (\theta + 2)x^{\theta+1} = (\theta + 2)e^{(\theta+1)logx}$, quindi appartiene alla famiglia esponenziale e possiamo applicare il teorema: $c(\theta) = \theta + 1$ è una funzione monotona crescente e la statistica da usare è $T = \sum_{i=1}^n logX_i$.

Calcoliamo la distribuzione di $Y = -logX$.

$F_Y(y) = P(Y \leq y) = P(-logX \leq y) = P(X \geq e^{-y}) = \int_{e^{-y}}^1 (\theta + 2)x^{\theta+1} dx = 1 - e^{-y(\theta+2)}$

$f_Y(y) = \frac{d}{dy} F_Y(y) = (\theta + 2)e^{-y(\theta+2)}$ da cui $Y \sim Exp(\theta + 2)$.

Allora $\sum_{i=1}^n Y_i = -\sum_{i=1}^n logX_1 \sim \Gamma(n, \theta + 2)$.

$\alpha = P(\sum_{i=1}^n logX_i > k|\theta_0) = P(-\sum_{i=1}^n logX_i < k_1|\theta_0) = P(\sum_{i=1}^n Y_i < k_1|\theta_0) = \int_0^{k_1} \frac{(\theta_0+2)^n}{\Gamma(n)} x^{n-1} e^{-(\theta_0+2)x} dx$
Con $k_1 = -k$.

Il test uniformemente più potente di ampiezza α eè dato dalla regione critica:
$C^* = \{(x_1, ..., x_n) t.c. \sum_{i=1}^n logX_i > k\}$ dove k_1 è il quantile di livello α di una $\Gamma(n, \theta_0 + 2)$.

Esercizio 5.
Sia $X_1, ..., X_n$ un campione casuale dalla distribuzione uniforme sull' intervallo $(0, \theta)$. Sia $Y_{(1)} := min\{X1, ..., Xn\}$ e $Y_{(n)} := max\{X1, ..., Xn\}$ e sia $a > 1$.

1. Dimostra che $(aY_{(1)}, aY_{(n)})$ è un intervallo di confidenza per θ.
   (Y_1, Y_n) è un intervallo di confidenza se $Y_1 \leq Y_n$ e $P(Y_1 < \theta) = \gamma$, con γ indipendente da θ. Abbiamo che:
   $1 - \gamma = P(\theta \neq (aY_1, aY_n)) = P(\theta < aY_1) + P(\theta > aY_n) = (1 - \frac{1}{a})^n + (\frac{1}{a})^n$, che non dipende da θ.

2. Determinare a tale che il livello di confidenza di questo intervallo sia massimo.
   Dobbiamo trovare il minimo della funzione:
   $f(a) = (1 - \frac{1}{a})^n + (\frac{1}{a})^n$
   Abbiamo che $f(1) = 1$ e $lim_{a \to +\infty} f(a) = 1$ è un unico punto critico, che è un minimo, che otteniamo calcolando la sua derivata:
   $f^{'}(a) = \frac{n}{a^{n+1}}[(a-1)^{n-1} - 1]$
   che è nulla in $a = 2$ dove la funzione vale:
   $f(2) = 2^{-(n-1)}$
   Dunque il valore ottimale per a è 2 e corrisponde ad un livello di confidenza $\gamma = 1 - 2^{-(n-1)}$.