Soluzioni del Tutorato di Statistica 1 del 15/11/2010
Docente: Prof.ssa Enza Orlandi
Tutore: Dott.ssa Barbara De Cicco
Esercizio 1.

Sia $X_1, ..., X_n$ un campione casuale dalla distribuzione Poissoniana di parametro λ.

1. Determinare la funzione generatrice dei momenti e la distribuzione di $S = \sum_{i=1}^{n} X_i$.
   $m_X(t) = E[e^{tX}] = \sum_{i=1}^{n} e^{tx} \frac{e^{-\lambda}\lambda^x}{x!} = e^{\lambda(e^t-1)}$
   $S = \sum_{i=1}^{n} X_i$ quindi $S \sim Po(n\lambda)$ da cui $E[e^{tS}] = e^{n\lambda(e^t-1)}$.

2. Si calcoli lo stimatore per λ con il metodo dei momenti.
   $\mu_1' = \lambda$ e $M_1' = \bar{X}$ da cui otteniamo che $T_{\hat{\theta}} = \bar{X}$.

3. Si calcoli lo stimatore di massima verosimiglianza per λ.
   $L(\lambda) = \prod_{i=1}^{n} f(x_i, \lambda) = \prod_{i=1}^{n} \frac{e^{-\lambda}\lambda_i^x}{x_i!} = \frac{e^{-n\lambda}\lambda^{\sum_{i=1}^{n} X_i}}{\prod_{i=1}^{n} x_i!}$
   Passando al logaritmo si ottiene:
   $logL(\lambda) = -n\lambda + \sum_{i=1}^{n} x_i log(\lambda) - log \prod_{i=1}^{n} x_i!$
   derivando in λ si ha:
   $-n\lambda + \frac{\sum_{i=1}^{n} x_i}{\lambda} = 0$ da cui otteniamo che $\hat{\lambda} = \frac{\sum_{i=1}^{n} x_i}{\lambda} = \bar{X}$.

4. Trovare una statistica sufficiente.
   Poichè la distribuzione di Poisson appartiene alla famiglia esponenziale, infatti si può scrivere come:
   $L(\lambda) = \prod_{i=1}^{n} f(x_i, \lambda) = e^{-n\lambda} e^{\sum_{i=1}^{n} x_i log\lambda} \prod_{i=1}^{n} \frac{1}{x!}$
   Si ha quindi che $S = \sum_{i=1}^{n} X_i$ è una statistica sufficiente.

5. Determinare un UMVUE di λ.
   Dalla disuguaglianza di Cramer-Rao si ottiene:
   $Var[T] \geq \frac{(\tau' \lambda)^2}{nE[(\frac{\partial}{\partial\lambda} logf(x,\lambda))^2]} = \frac{1}{nE[(-1+\frac{X}{\lambda})]} = \frac{1}{nE[1+\frac{X^2}{\lambda^2}-\frac{2X}{\lambda}]} = \frac{\lambda}{n}$
   Ora poichè $\bar{X}$ è uno stimatore non distorto di λ e la sua varianza coincide con quella del limite inferiore di Cramer-Rao, allora possiamo concludere che $\bar{X}$ è un UMVUE per $\tau(\lambda) = \lambda$.

6. Dimostrare che le statistiche
   $T_1 = \frac{1}{n} \sum_{i=1}^{n} 1_{(0)}(X_i)$

$T_2 = (\frac{n-1}{n})^{\sum_{i=1}^{n}(X_i)}$

sono stimatori non distorti di $\tau(\lambda) = e^{-\lambda}$. Determinare un UMVUE di $\tau(\lambda)$.

$E[T_1] = E[\frac{1}{n}\sum_{i=1}^{n} 1_{(0)}(X_i)] = \frac{1}{n}\sum_{i=1}^{n} P(X_i = 0) = e^{-\lambda}$.
$E[T_2] = E[(\frac{n-1}{n})^{\sum_{i=1}^{n}(X_i)}] = E[(\frac{n-1}{n})^{S}] = \sum_{s=0}^{+\infty}(\frac{n-1}{n})^s e^{-n\lambda}\frac{(n\lambda)^s}{s!} = e^{-n\lambda}e^{\lambda(n-1)} = e^{-\lambda}$.

Quindi T_1 e T_2 sono due stimatori non distorti per $\tau(\lambda) = e^{\lambda}$.
Poichè T_2 è uno stimatore non distorto di $\tau(\lambda) = e^{\lambda}$ ed inoltre è funzione di S statistica sufficiente, allora T_2 è un UMVUE per $\tau(\lambda) = e^{\lambda}$.

7. Calcolare il limite inferiore di Cramer-Rao per lo stimatore di $e^{-\lambda}$.
   $Var[T] \geq \frac{(-e^{-\lambda})^2}{nE[1+\frac{X^2}{\lambda}-\frac{2X}{\lambda}]} = \frac{\lambda e^{-2\lambda}}{n}$.

Esercizio 2.
Sia $X_1, ..., X_n$ un campione casuale da $f(x,\theta) = \frac{1}{\theta}e^{-x/\theta}1_{(0,\infty)}(x)$

1. Dimostrare che le statistiche $\hat{\theta}_1, ..., \hat{\theta}_4$ sono stimatori non distorti di θ e calcolarne gli errori quadratici medi relativi $MSE(\hat{\theta}_i)$.
   $\hat{\theta}_1 = X_1$; $\hat{\theta}_2 = \frac{X_1+X_2}{2}$; $\hat{\theta}_3 = \frac{X_1+2X_2}{3}$; $\hat{\theta}_4 = \bar{X}$
   $E[\hat{\theta}_1] = E[X_1] = \int_0^{+\infty} x\frac{1}{\theta}e^{-x/\theta}dx = \theta$
   $E[\hat{\theta}_2] = E[\frac{X_1+X_2}{2}] = \frac{1}{2}E[X_1] + \frac{1}{2}E[X_2] = \theta$
   $E[\hat{\theta}_3] = E[\frac{X_1+2X_2}{3}] = \frac{1}{3}E[X_1] + \frac{2}{3}E[X_2] = \theta$
   $E[\hat{\theta}_4] = E[\bar{X}] = \theta$
   Quindi $\hat{\theta}_1, ..., \hat{\theta}_4$ sono stimatori non distorti di θ.

2. Dimostrare che $\hat{\theta}_4$ è una statistica sufficiente e trovare l' UMVUE per θ.
   La distribuzione appartiene alla famiglia esponenziale infatti:
   $\prod_{i=1}^{n}(x_i,\theta) = (\frac{1}{\theta})^n e^{-\frac{1}{\theta}\sum_{i=1}^{n}x_i}\prod_{i=1}^{n}1_{0,\infty}(x_i)$
   Per cui si ha che $S = \sum_{i=1}^{n}X_i$ è una statistica sufficiente.
   Per la disuguaglianza di Cramer-Rao si ha:
   $Var[T] \geq \frac{1}{nE[(-\frac{1}{\theta}+\frac{X}{\theta^2})^2]} = \frac{1}{nE[\frac{1}{\theta^2}+\frac{X^2}{\theta^4}-\frac{2X}{\theta^3}]} = \frac{1}{\frac{n}{\theta^2}} = \frac{\theta^2}{n}$
   Poichè $\hat{\theta}_4$ è uno stimatore non distorto di θ e la sua varianza

coincide con il limite inferiore di Cramer-Rao, infatti $Var[\hat{\theta}_4] = \frac{\theta^2}{n}$, si ha che $\hat{\theta}_4$ è un UMVUE per $\tau(\theta) = \theta$.

3. Trovare un UMVUE per $Var(X_i)$.
   $Var[X_i] = E[X_i^2] - E[X_i]^2 = 2\theta^2 - \theta^2 = \theta^2$
   Scegliamo come possibile stimatore di θ^2, $\bar{X}^2$
   Poichè è distorto infatti:
   $E[\bar{X}^2] = Var[\bar{X}] - E[\bar{X}]^2 = \theta^2(\frac{1}{n} + 1)$
   lo correggiamo e poniamo: $T_{\hat{\theta^2}} = \frac{n}{n-1}\bar{X}^2$.
   Adesso $T_{\hat{\theta^2}}$ è uno stimaotre non distorto di θ^2 ed inoltre è funzione si $S = \sum_{i=1}^n X_i$, statistica sufficiente trovata al punto (2) dell'esercizio, quindi concludiamo che $T_{\hat{\theta^2}}$ è un UMVUE di $\tau(\theta) = \theta^2$.

Esercizio 3.

Sia $X_1, ..., X_n$ un campione casuale da:
$f(x, \theta) = \theta x^{\theta-1} 1_{(0,1)}(x)$, $\theta > 0$.

1. Trovate lo stimatore di massima verosimiglianza di $\mu = \frac{\theta}{1+\theta}$.
   $L(\theta) = \prod_{i=1}^n f(x_i, \theta) = \prod_{i=1}^n \theta x_i^{\theta-1} 1_{(0,1)}(x_i) = \theta^n \prod_{i=1}^n x_i^{\theta-1} 1_{(0,1)}(x_i)$
   Passando al logaritmo si ottiene:
   $logL(\theta) = nlog\theta + (\theta - 1)\sum_{i=1}^n logx_i + \sum_{i=1}^n log1_{(0,1)}(x_i)$
   Derivando in θ otteniamo:
   $\frac{\partial}{\partial} logL(\theta) = \frac{n}{\theta} + \sum_{i=1}^n logx_i = 0$ da cui: $\hat{\theta} = -\frac{n}{\sum_{i=1}^n} logX_i$
   Per cui invertendo la funzione, si ha che lo stimatore di μ è: $\hat{M} = \frac{-n}{\sum_{i=1}^n logX_i - n}$.

2. Trovate una statistica sufficiente.
   La distibuzione appartiene alla famiglia esponenziale infatti:
   $\prod_{i=1}^n f(x_i, \theta) = \theta^n e^{(\theta-1)\sum_{i=1}^n logx_i} \prod_{i=1}^n 1_{(0,1)}(x_i)$
   Per cui $S = \sum_{i=1}^n logX_i$ è una statistica sufficiente.

3. C' è una funzione di θ per la quale esiste uno stimatore non distorto la cui varianza coincide con il limite inferiore di Cramer-Rao?
   Dalla disuguaglianza di Cramer-Rao si ottiene:
   $Var[T] \geq \frac{1}{nE[(\frac{1}{\theta} + logX)^2]} = \frac{1}{nE[\frac{1}{\theta^2} + \log^2 X + \frac{2}{\theta} logX]}$
   Calcoliamo quindi la distribuzione di $Y = -logX$
   $P(Y \leq y) = P(-logX \leq y) = P(X \geq e^{-y}) = \int_{e^{-y}}^1 x\theta x^{\theta-1} = \frac{\theta}{\theta+1}(1 - e^{-y(\theta+1)})$
   La densità è: $f_Y(y) = \frac{d}{dy} F_Y(y) = \theta e^{-\theta y} 1_{(0,+\infty)}(y)$.
   Per cui $Y \sim Exp(\theta)$.
   Allora $Var[T] \geq \frac{1}{nE[\frac{1}{\theta^2} + \log^2 X + \frac{2}{\theta} logX]} = \frac{1}{nE[\frac{1}{\theta^2} + Y^2 - \frac{2}{\theta} Y]} = \frac{\theta^2}{n}$. Ora sia

$Z = \frac{Y}{n}$, $\sum_{i=1}^{n} Z_i \sim \Gamma(n, n\theta)$ infatti:
$P(Z \leq z) = P(\frac{Y}{n}) \leq z) = P(Y \leq nz) = \int_0^{nz} \theta e^{\theta x} dx = \int_0^{z} \theta e^{-\theta n y} dy$
Avendo effettuato il cambio di variabili $x = ny$.
Quindi $Z \sim Exp(n\theta)$ da cui $Z^* = \sum_{i=1}^{n} Z_i \sim \Gamma(n, n\theta)$.
$E[Z^*] = \frac{1}{\theta}$
$Var[Z^*] = \frac{1}{n\theta^2}$
Osserviamo quindi che Z^* è uno stimatore non distorto di $\frac{1}{\theta}$, dalla disuguaglianza di Cramer-Rao si ha:
$Var[T] \geq \frac{(\frac{1}{\theta^2})^2}{\frac{n}{\theta^2}} = \frac{1}{n\theta} = Var[Z^*]$.
Quindi una funzione del parametro θ per la quale esiste uno stimatore non distorto che raggiunge il limite inferiore di Cramer-Rao è $\frac{1}{\theta}$.

4. Trovate l'UMVUE di θ e di $\frac{1}{\theta}$.
Troviamo ora l'UMVUE per θ, cerchiamo quindi uno stimatore non distorto di θ:
$E[\hat{\theta}] = \int_0^{+\infty} \frac{1}{x} \frac{(n\theta)^n}{\gamma(n)} x^{n-1} e^{-n\theta x} dx = n\theta \frac{\Gamma(n-1)}{\Gamma(n)} = \frac{n}{n-1}\theta$.
Lo correggiamo:
$\hat{\theta}^* = \frac{n-1}{n}\hat{\theta}$ è quindi lo stimatore non distorto di θ ed essendo funzione della statistica sufficiente S trovata al punto (2), concludiamo che $\hat{\theta}^*$ è l'UMVUE di θ.