Soluzioni del Tutorato di Statistica 1 del 22/11/2010
Docente: Prof.ssa Enza Orlandi
Tutore: Dott.ssa Barbara De Cicco

Esercizio 1.
Sia $X_1, ..., X_n$ un campione casuale da $f(x, \theta) = \frac{1}{\theta} 1_{[0,\theta]}(x)$, con $\theta > 0$.

1. Stimate θ con il metodo dei momenti, chiamate lo stimatore T_1 e trovate la sua media e il suo errore quadratico medio.

   $E[X] = \int_0^\theta x \frac{1}{\theta} dx = \frac{\theta}{2}$
   Si ha quindi che $T_1 = T_{\hat{\theta}} = 2\bar{X}$.
   Dalla definizione di errore quadratico medio si ha:
   $MSE_\theta = Var[T_1] + \{\tau(\theta) - E[T_1]\}^2$
   poichè si verifica che T_1 è uno stimatore non distorto di θ, allora $MSE_\theta = Var[T_1] = Var[2\bar{X}] = \frac{4}{3} \frac{\theta^2}{n}$.

2. Trovate lo stimatore di massima verosomiglianza di θ, chiamate lo stimatore T_2 e trovate la sua media e il suo errore quadratico medio.

   $L(\theta) = \prod_{i=1}^n \frac{1}{\theta} 1_{[0,\theta]}(x_i) = \frac{1}{\theta^n} \prod_{i=1}^n 1_{[0,\theta]}(x_i)$

   Si ha qiundi che lo stimatore di massima verosimiglianza $T_2 = max\{X_1, ..., X_n\}$
   Chiamiamo $Y = max\{X_1, ..., X_n\}$ e ne calcoliamo la distribuzione:

   $P(Y \leq y) = P(max\{X_1, ..., X_n\} \leq y) = P(X \leq y)^n = (\int_0^y \frac{1}{\theta} dx)^n = \frac{y^n}{\theta^n}$
   $F_Y(y) = n \frac{y^{n-1}}{\theta^n} 1_{(0,\theta)}(y)$
   $E[Y] = E[T_2] = \int_0^\theta ny \frac{y^{n-1}}{\theta^n} dy = \frac{n\theta}{n+1}$
   $E[Y^2] = E[T_2^2] = \int_0^\theta ny^2 \frac{y^{n-1}}{\theta^n} dy = \frac{n\theta^2}{n+2}$
   $Var[Y] = Var[T_2] = \frac{n\theta^2}{(n+1)^2(n+2)}$
   $MSE_\theta = Var[T_2] - \{\theta - E[T_2]\}^2 = \frac{2\theta^2}{(n+1)(n+2)}$.

Esercizio 2.
Sono stati eseguiti 400 lanci di una moneta e sono stata ottenute 175 Teste e 225 croci. Trovate un intervallo di confidenza di livello $\gamma = 0.90$ per la probabilità di testa, ed un altro di livello $\gamma = 0,99$.
La variabile X è quindi una $Bernoulli(p)$ quindi
$P(X = 1) = p$ e $P(X = 0) = 1 - p$
Sia $\hat{p} = \frac{175}{400} = \frac{\sum_{i=1}^{400} X_i}{400}$
$0.90 = P(-q_1 \leq \frac{\bar{X}-p}{\sqrt{\sigma^2/n}} \leq q_1) = P(-q_1 \leq \frac{\bar{X}-p}{\sqrt{p(1-p)/n}} \leq q_1) \sim$

$P(-1,645\sqrt{\hat{p}(1-\hat{p})/n} \leq p \leq 1,645\sqrt{\hat{p}(1-\hat{p})/n})$
Avendo usato l'approssimazione per grandi campioni poichè n è grande.

Esercizio 3.
Si consideri un campione casuale di ampiezza n da una distribuzione $N(\mu, \sigma^2)$.

1. Determinare un intervallo di confidenza per μ quando σ^2 è nota. Come varia l'ampiezza dell'intervallo se σ^2 si raddoppia?
   L'intervallo si trova considerando come quantità pivotale $Q = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ da cui ricaviamo che $\gamma = P(\bar{X} - q_1\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + q_1\frac{\sigma}{\sqrt{n}})$, dove q_2 è il quantile di una Normale standard di livello $\frac{1+\gamma}{2}$. L'intervallo di confidenza per μ di livello γ è quindi:
   $(\bar{X} - q_1\frac{\sigma}{\sqrt{n}}; \bar{X} + q_1\frac{\sigma}{\sqrt{n}})$.
   L'intervallo ha ampiezza $2q_1\frac{\sigma}{\sqrt{n}}$. Raddoppiando σ, si raddoppia l'ampiezza.

2. Determinare un intervallo di confidenza per σ^2 quando μ è noto. Come varia l'ampiezza di questo intervallo al variare di μ? Per quale valore di μ la sua ampiezza è minima?
   La quantità pivotale da considerare è: $Q = \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2}$ quindi si ha :
   $\gamma = P(q_1 \leq Q \leq q_2) = P(\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{q_2} \leq \sigma^2 \leq \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{q_1})$
   dove q_1 e q_2 sono i quantili di una distribuzione χ^2_n.
   L'intervallo di confidenza è quindi : $(\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{q_2}; \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{q_1})$.
   L'ampiezza dell'intervallo è: $(\frac{1}{q_1} - \frac{1}{q_2})\sum_{i=1}^{n}(X_i-\mu)^2$.
   Questo è minimo per $\mu = \bar{X}$.

Esercizio 4.
Sia $X_1, ..., X_n$ un campione casuale estratto da una $N(\mu, 25)$. Applica il metodo pivotale per trovare un intervallo di confidenza per μ al 90%. Quanto deve essere grande il campione per avere l'ampiezza dell'intervallo < 1?
La quantità pivotale è: $Q = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$
$0.90 = P(-q_1 \leq Q \leq q_1) = P(\bar{X} - q_1\frac{5}{\sqrt{n}} \leq \mu \leq \bar{X} + q_1\frac{5}{\sqrt{n}})$
Dalle tavole si ricava che $q_1 = 1,645$ e quindi l'intervallo per μ al 90% è:
$(\bar{X} - 1,645\frac{5}{\sqrt{n}}; \bar{X} + 1,645\frac{5}{\sqrt{n}})$.
L'ampiezza dell'intervallo è: $\bar{X} + 1,645\frac{5}{\sqrt{n}} - \bar{X} + 1,645\frac{5}{\sqrt{n}}$.
Si ha quindi che:
$\bar{X} + 1,645\frac{5}{\sqrt{n}} - \bar{X} + 1,645\frac{5}{\sqrt{n}} < 1$ per $n > 270$

Esercizio 5.
Trovate un intervallo di confidenza al 90% per la media di una distribuzione normale con $\sigma = 3$ dato il campione $(3.3, 0.3, 0.6, 0.9)$. Quale sarebbe l'intervallo di confidenza se σ fosse noto?

$\bar{X} = \frac{1}{4}(3.3 + 0.3 + 0.6 + 0.9) = 1,277$
$0,90 = P(-q_1 \leq \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \leq q_1) = P(\bar{X} - q_1 \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + q_1 \frac{\sigma}{\sqrt{n}})$
Dalle tavole si ricava che $q_1 = 1,645$
Quindi l'intervallo di confidenza per μ è:
$(1,195; 3,74)$.