A

Prova intracorso 13/12/2007
Metodi per l'Analisi dei Dati Sperimentali AA 2007/08

Nome…..…………………….. Cognome……………………………
Matricola…………………….

---

5 punti 1) Si consideri una VA caratterizzata dalla funzione densità di probabilità

$$f(y) = \frac{1}{\beta^2} y e^{-\frac{y}{\beta}}.$$

Si determini lo stimatore basato sulla massima verosimiglianza per il parametro β.

La funzione verosimiglianza per un campione di dimensione *n* è:

$$L(\theta) = \prod_{i=1}^{n} \frac{x_i e^{-x_i/\theta}}{\theta^2}$$

Lo stimatore basato sulla massima verosimiglianza si ottiene massimizzando rispetto al parametro la funzione verosimiglianza calcolata in corrispondenza dei dati sperimentali. Come visto a lezione conviene farlo dopo aver applicato il logaritmo.

$$\ln L(\theta) = \sum \ln(x_i) - \sum \frac{x_i}{\theta} - 2n \ln \theta$$

$$\frac{\partial \ln L(\theta)}{\partial \theta} = \frac{1}{\theta^2} \sum x_i - \frac{2n}{\theta}$$

Ponendo pari a zero la derivata e risolvendo per θ si ottiene lo stimatore:

$$\hat{\theta} = \frac{\sum_{i=1}^{n} x_i}{2n}$$

5 punti 2) Una macchina produce barre metalliche. Un campione di 15 barre è selezionato a caso nell'ambito della produzione giornaliera. Il diametro in millimetri delle 15 barre è riportato in tabella.

8.24	8.25	8.20	8.23	8.24	8.21	8.26	8.26	8.20	8.25	8.23	8.23	8.19	8.28	8.24

Nell'ipotesi che il diametro sia una VA distribuita normalmente si determini, l'intervallo fiduciario al 95% per la media del diametro delle barre. Illustrate e motivate tutti i passaggi.

Si deve determinare l'intervallo fiduciario della media di un campione di cui non si conosce la varianza. Questa andrà stimata dai dati sperimentali. La variabile aleatoria da studiare sarà quindi una t di Student a 14 gradi di libertà.

Fissiamo γ=95%. Si calcola c tale che F(c)=0.5*(1+γ), con F la CDF di una t di Student a 14 gdl. Con Matlab o con le tabelle si ricava c=2.1448
L'intervallo da determinare è:

$$\left[\bar{y}-\frac{c\,s}{\sqrt{n}},\bar{y}+\frac{c\,s}{\sqrt{n}}\right]$$

Dove è la media del campione ed s è la deviazione standard non distorta stimata dai dati: $\bar{y}=\frac{\sum_{i=1}^{n} y_i}{n}$; $s^2=\frac{\sum_{i=1}^{n}\left(y_i-\bar{y}\right)^2}{n-1}$

8 punti In tabella sono riportati i dati sperimentali a 230°C di viscosità a zero shear (in Poise) contro peso molecolare per un certo polimero. La viscosità può dipendere dal peso molecolare secondo una legge lineare (y=*a*×M) o con una potenza 3.4 (y=*a*×$M^{3.4}$). Eseguendo una regressione lineare si determini quale delle due dipendenze è compatibile con i dati, e si stimi il valore del parametro *a* e la varianza sperimentale.

M	1.00E+03	2.00E+03	3.00E+03	4.00E+03	5.00E+03	6.00E+03	7.00E+03	8.00E+03	9.00E+03	1.00E+04
Viscosità	1.58E+05	1.67E+06	6.64E+06	1.77E+07	3.77E+07	7.01E+07	1.18E+08	1.86E+08	2.78E+08	3.98E+08

Si devono effettuare due stime di parametri di modelli **lineari** nei parametri e quindi confrontare i risultati. Utilizzando Matlab, la funzione da utilizzare è regress.
Con i dati del compito A si ottiene nel caso del primo modello $\hat{a}=24674$, e con il modello a legge di potenza $\hat{a}=9.99\times10^{-6}$. Diagrammando le due regressioni ed i dati sperimentali si può concludere che il modello a legge di potenza è compatibile con i dati sperimentali. (Nel caso del compito B la situazione era opposta).
La varianza si determina con lo stimatore non distorto:

$$s^2=\frac{\sum_{i=1}^{n}\left(y_i-\hat{a}x_i\right)^2}{n-1} \quad \textit{oppure} \quad s^2=\frac{\sum_{i=1}^{n}\left(y_i-\hat{a}x_i^{3.4}\right)^2}{n-1}$$

Prova intracorso 13/12/2007
Metodi per l'Analisi dei Dati Sperimentali AA 2007/08

Nome…..…………………….. Cognome……………………………
Matricola…………………….

---

14 punti 4) Sono state effettuate misure della temperatura all'interno di un reattore al variare della pressione nell'intervallo [2, 3] atm. I dati sono di seguito riportati:

p° [atm]	2.0	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	3.0
T [°K]	773.3	811.4	861.1	916.1	947.2	959.2	1001.6	1043.1	1091.0	1079.6	1115.6

Si suppone che un buon modello per descrivere i dati sia:

$$T = \frac{A \cdot p^2}{B \cdot p^2 + C \cdot p + 1}$$

dove A, B e C sono parametri del modello.

a) il modello è linearizzabile?
b) In caso affermativo, si effettui una stima puntuale e per intervallo dei parametri del modello linearizzato.
c) Si effettui una stima dei parametri del modello non lineare
d) Si effettui una stima della varianza del modello non lineare
Si diagrammino i dati sperimentali e la curva del modello non lineare

---

a) Il modello è linearizzabile con una trasformazione iperbolica $\frac{1}{T} = k_1 + \frac{k_2}{p} + \frac{k_3}{p^2}$.

b) La stima puntuale dei parametri del modello linearizzato con la massima verosimiglianza si ottiene utilizzando il metodo dei minimi quadrati pesati. I pesi sono necessari perché la trasformazione non lineare altera le accuratezze. I pesi si determinano attraverso la derivata della trasformazione: $\gamma_i^2 = \frac{1}{\left(\left.\frac{\partial g}{\partial T}\right|_i\right)^2} = T^4$. La matrice dei pesi è $\underline{\underline{W}} = diag\left(1/\gamma_i^2\right)$.

La stima dei parametri k_i è un problema di regressione multilineare:

$$\hat{\boldsymbol{\theta}} = \left(\boldsymbol{X}^T \cdot \boldsymbol{W} \cdot \boldsymbol{X}\right)^{-1} \cdot \left(\boldsymbol{X}^T \cdot \boldsymbol{W} \cdot \boldsymbol{y}\right) \qquad con \qquad \boldsymbol{X} = \begin{bmatrix} 1 & 1/p_1 & 1/p_1^2 \\ 1 & 1/p_2 & 1/p_2^2 \\ \ldots & \ldots & \ldots \end{bmatrix}$$

Con Matlab, procedendo in modo diretto, per il compito A si ottiene $\hat{k}_1$ =0.0009, $\hat{k}_2$ =-0.0018, $\hat{k}_3$ =0.005 da cui si ricava $\hat{A}$=200, $\hat{B}$=0.18, $\hat{C}$=-0.36.
Per la stima per intervallo dei parametri ki si fissa γ=0.95. La stima per intervallo coinvolge una VA di tipo t di Student: $P\left(\frac{\hat{k}_i - k_i}{s_{\hat{k}_i}} \in \left[-t_{N-3,\alpha/2}, t_{N-3,\alpha/2}\right]\right) = 1-\alpha$. Per determinare la

deviazione standard dello stimatore del parametro $s_{\hat{k}_i}$ si deve fare riferimento alla matrice di covarianza dello stimatore dei parametri dei modelli multilineari: $\underline{\underline{V}}_{\underline{\hat{k}}} = \sigma^2 \left(\underline{\underline{X}}^T \cdot \underline{\underline{W}} \cdot \underline{\underline{X}} \right)^{-1}$. Gli elementi sulla diagonale principale di tale matrice sono le varianze degli stimatori dei parametri.

c) La stima dei parametri del modello non lineare si effettua con Matlab (nlinfit o fminsearch) partendo dal primo tentativo ottenuto attraverso la linearizzazione. Per il compito A si perviene a $\hat{A}$=202.05, $\hat{B}$=0.186, $\hat{C}$=-0.348.

d) La stima della varianza imparziale si ottiene dalla relazione:

$$s^2 = \frac{\sum_{i=1}^{n} \left(y_i - \frac{\hat{A} \cdot p_i^2}{\hat{B} \cdot p_i^2 + \hat{C} \cdot p_i + 1} \right)^2}{n-3}$$