"Il Metodo del Maximum Likelihood applicato alle funzioni di luminosità".

Fanti R., Padrielli L., Rogora A..

IRA 47/81

1) INTRODUZIONE

Nella ricerca induttiva, fondamentale è il seguente problema: tra più ipotesi escludentisi a vicenda, e a priori possibili, accettarne una in base all'informazione che si ricava da un esperimento.

Una maniera di procedere nella ricerca della soluzione è il metodo del maximum likelihood. Questo è di generale applicazione e deriva dal teorema di Bayes che si può così brevemente illustrare.

Sia H_i la generica ipotesi possibile, e si supponga nota la probabilità $P(H_i)$ (detta a priori) che tale ipotesi si presenti.

Sia B l'esito di un esperimento e si supponga di conoscere la probabilità che B si manifesti nel caso dell'ipotesi H_i : sia $P(B/H_i)$ questa probabilità (detta condizionata).

Il teorema di Bayes, in base a questi elementi, determina la probabilità che sia proprio H_i l'ipotesi accettabile (posto che si sia verificato B).

E cioè:

$$P(H_i/B) = \frac{P(H_i) \cdot P(B / H_i)}{\sum_k P(H_k) \cdot P(B/H_k)}$$

Se consideriamo fissate le ipotesi a priori H_i e introduciamo una condizione p che renda verificati sia l'ipotesi H_i che il risultato sperimentale B, (cioè consideriamo p come se fosse un'ipotesi aggiuntiva), per il principio delle probabilità composte la formula di Bayes diventa:

$$P(H/p,B) = \frac{P(H,p/B)}{\sum P(H,p/B)} \qquad 1)$$

dove $P(H/p,B)$ è sempre la cosiddetta probabilità a priori, $P(H,p/B)$ è la probabilità condizionata e $P(H,p/B)$ è detta probabilità a posteriori.

Se chiamiamo $P(p/H,B)$ funzione di likelihood, allora il teorema di Bayes può essere riformulato così: la probabilità a posteriori varia come la probabilità a priori moltiplicata per il likelihood.

Possiamo riscrivere la 1) in questa forma:

$$P(H/p,B) \propto P\,(H/B)\mathcal{L}(p/H,B)$$

dove $\mathcal{L}(p/H, B)$ è appunto il likelihood.

Il principio di maximum likelihood stabilisce che, messi di fronte a una scelta di ipotesi H_i, noi scegliamo la p, se c'è, che massimizza la funzione $\mathcal{L}$. In altre parole noi scegliamo quella ipotesi che dà la massima probabilità di osservare l'evento.

2) IL MAXIMUM LIKELIHOOD APPLICATO ALLE FUNZIONI DI LUMINOSITA'

La funzione di luminosità è largamente usata negli studi statistici e cosmologici associati a campioni di radiosorgenti.

Essa viene definita come il numero di sorgenti per intervallo di logaritmo di potenza radio n (L).

Questo numero viene generalmente normalizzato o con il volume realmente osservato dal nostro campione radio (numero di sorgenti per unità di volume) o con il numero totale di oggetti della categoria che si sta studiando (percentuale di oggetti radioemittenti).

Per ogni intervallo di $\log L$ si osserva in genere un numero di oggetti piuttosto basso, da 0 a 20 circa, che costituisce il cosiddetto "esito dell'esperimento" (chiamato B nella 1).

L'ipotesi (H) nella 1) che si assume è che la funzione di luminosità n (L) sia una funzione esponenziale, genericamente rappresentata come:

$$n(L)\, d\log L = KL^{-\alpha}\, d\log L$$

K e α che rappresentano le costanti della funzione sono la condizione o ipotesi aggiuntiva (p nella 1) che vogliamo determinare con il metodo del maximum likelihood.

La funzione di likelihood è quindi la probabilità di avere l'insieme di osservati O_i per ogni intervallo ΔL_i , i=1,n, nelle condizioni sopra dette.

Per determinare tale funzione consideriamo l'iesimo intervallo. Dato un generico α e K, il numero aspettato di oggetti è:

$$E_i = KL_i^{-\alpha} \quad V_i \qquad 2)$$

dove V_i costituisce il fattore di normalizzazione.

Dato il basso numero di oggetti osservati, la statistica di Poisson descrive meglio di quella di Gauss gli eventi, per cui la probabilità di trovare un numero osservato O_i , quando il numero aspettato è E_i è:

$$P(O_i/E_i) = \frac{E_i^{O_i}\, e^{-E_i}}{O_i\,!} \qquad \text{(formula di Poisson)}$$

La funzione di likelihood è la probabilità composta delle n caselle, cioè:

$$\mathcal{L} = \Pi_i \frac{E_i^{O_i}\, e^{-E_i}}{O_i\,!}$$

Essendo le costanti moltiplicative degli esponenziali, sempre positive, poniamo $\mathcal{L} = e^{-S/2}$,

da cui:

$$S = -2 \ln \mathcal{L}$$

Il maximum likelihood applicato a questo caso consiste nel trovare il valore di K e α che rende massima $\mathcal{L}$, o, il che è lo stesso, minimo S.

3)LA FUNZIONE S E LE SUE DERIVATE

Nel capitolo precedente la funzione $S(\alpha,K)$ è stata definita come:

$$S = -2\ln \Pi_i \frac{E_i^{-O_i} e^{-E_i}}{O_i\,!}$$

Sostituendo E_i con la sua espressione 2) si ottiene:

$$S= -2\sum_i \left\{ O_i \ln (K L_i^{-\alpha} V_i) - K L_i^{-\alpha} V_i - \ln O_i! \right\} \qquad 3)$$

Per minimizzare S dobbiamo cercare i valori di α e K che annullano le derivate di S rispetto ad α e K.

$$\begin{cases} \dfrac{\delta S}{\delta K} = -\dfrac{1}{K}\sum_i O_i + \sum_i L_i^{-\alpha} V_i = 0 \\[2ex] \dfrac{\delta S}{\delta \alpha} = -2\sum_i \left\{- O_i \ln L_i + K L_i^{-\alpha} V_i \ln L_i\right\} = 0 \end{cases}$$

da cui:

$$\begin{cases} K = \dfrac{\sum_i O_i}{\sum_i L_i^{-\alpha} V_i} & 4) \\[2ex] \sum_i \ln L_i O_i - \dfrac{\sum_i O_i}{\sum_i L_i^{-\alpha} V_i} \sum_i L_i^{-\alpha} V_i \ln L_i = 0 & 5) \end{cases}$$

Purtroppo questo sistema non è analiticamente invertibile.

Per determinare α dall'equazione 4) è stato applicato un metodo iterativo che conduce alla soluzione con una precisione di 0,1%.

Il metodo applicato è il cosidetto metodo delle tangenti che si può così brevemente riassumere.

Poniamo l'equazione 4)

$$f(\alpha) = 0$$

Si consideri un punto iniziale qualsiasi α_o.

Sviluppando in serie la funzione $f(\alpha)$ intorno al punto α_o avremo:

$$f(\alpha) = f(\alpha_o) + (\alpha - \alpha_o)\left(\frac{\delta f}{\delta \alpha}\right)_{\alpha_o} = 0 \qquad \text{da cui}$$

$$(\alpha - \alpha_o) = -f(\alpha_o) \Big/ \left(\frac{\delta f}{\delta \alpha}\right)_{\alpha_o}$$

Si ottiene quindi un nuovo punto α^1

$$\alpha^1 = \alpha_o - f(\alpha_o) \Big/ \left(\frac{\delta f}{\delta \alpha}\right)_{\alpha_o}$$

che costituisce una migliore approssimazione alla soluzione ed è il punto di partenza per una seconda iterazione . Il procedimento continua fino a che le differenze tra α^n e α^{n-1} così calcolati non siano divenute minori di 0,1 % .

Una volta determinati α e K e l'S corrispondente, resta il problema di definire gli errori associati a questi valori .

Secondo la teoria, tutte le coppie di α e K a cui corrisponde un S (ricavabili dalla formula 3) $\leq S_{min} + 1$ sono da considerarsi entro 1 σ del valore trovato, mentre tutte le coppie con $S \leq S_{min} + 4$ definiscono l'insieme delle soluzioni entro 2 σ .

4) ISTRUZIONI PER L'USO DEI PROGRAMMI

Sono disponibili sul calcolatore VAX11/780 due programmi che applicano le teorie esposte: SIGMA e SSS.

-SIGMA-

Questo programma, dato un set di osservati O_i per intervalli di log L_i, determina la coppia α e k che meglio approssima i dati O_i con l'ipotesi di $O_i \propto KL_i^{-\alpha}$, con il metodo del maximum likelihood.

INPUT- L'Input free format, è sull'unità logica 5. Consiste di un primo record con due numeri Pρ, z che il programma non utilizza altro che per stamparli in testa all'output. L'utente può utilizzare questi due numeri come commenti. Segue una serie di record free format, tanti quante sono le caselle di Δ log L_i con i seguenti dati:

C_{1i}, C_{2i}, log L_i, V_i, O_i.

C_{1i} e C_{2i} non vengono utilizzati altro che nella stampa e costituiscono ulteriori informazioni che l'utente voglia aggiungere. Log L_i, V_i, O_i, sono rispettamente il logaritmo in base 10 della potenza radio, il coefficiente di normalizzazione (moltiplicativo) e il numero di oggetti osservati (può essere anche 0) per ogni casella.

OUTPUT- L'Output è sull'unità logica 6 . Dopo l'intestazione e i dati input il programma scrive una serie di valori che servono per la soluzione del sistema con il metodo delle tangenti.
Il controllo di questi numeri può essere utile per assicurarsi che il programma non lavori al limite dell'overflow.
Al termine viene stampato:

i) il numero di iterazioni che sono servite per determinare α

ii) α denominato "pendenza"

iii) $\log_{10} K$ chiamato "normalizzazione"

iv) il valore S_{min} corrispondente denominato SS

Nella Appendice 1 viene riportato il testo in FORTRAN del programma SIGMA.

Una copia si può ottenere con il comando copy; il nome completo del file è LUCIA SIGMA.FOR.

-SSS-

Il programma SSS serve per determinare quanto la differenza fra una coppia α',K' qualsiasi, e α_o, K_o determinati con il programma precedente sia da considerarsi significativa dato un set di osservati.

Può servire per il confronto di diverse funzioni di luminosità.

INPUT

unità 5: è previsto un input nell'unità 5 identico a quello del programma SIGMA.

INPUT

unità 9: Il programma prevede una seconda unità di lettura UNITA' 9 da cui vuol leggere il valore di α e $\log_{10}$ K da confrontare con i dati.

Sulla stessa scheda c'è un terzo numero intero I, tale che, se I=0 viene calcolato S relativo ad α e K letti, e se $I \neq 0$ il programma non considera il K letto, ma calcola il fattore di normalizzazione con la formula 3) e determina l'S relativo all'α letto e al K calcolato.

Quindi stampa α, $\log_{10}$ K, S con le diciture α =AL, $\log_{10}K$ = AK, S=SS.

Dal confronto di questo S con l'S_{min} calcolato dal programma SIGMA, l'utente sarà in grado di valutare se i valori di varie funzioni siano o no significativamente diverse tra loro.

Questo programma è presentato in Appendice 2 e si trova sul VAX con il nome [LUCIA] SSS.FOR.

APPENDICE n°1 SIGMA.FOR

```
      DIMENSION RO(20),ERO(20),PLO(20),V(20),NOS(20),P(20)
      REAL*8 EE,EEE,DD,PP,AL,VV,CC,DF,OO,AK
      K=1
      READ(5,*)PO,Z
      WRITE(6,*) PO,Z
 2    READ(5,*,END=100)RO(K),ERO(K),PLO(K),V(K),NOS(K)
      WRITE(6,7) RO(K),ERO(K),PLO(K),V(K),NOS(K)
 7    FORMAT(10X,2E10.2,F7.3,E10.2,I5)
      K=K+1
      GO TO 2
 100  NO=K-1
      DO 3 K=1,NO
      P(K)=10**PLO(K)
 3    PLO(K)=PLO(K)*2.30259
      CC=0
      DO 5 K=1,NO
 5    CC=CC+NOS(K)*PLO(K)
      OO=0.
      IT=1.
      DO 6 K=1,NO
 6    OO=OO+NOS(K)
      AL=1.
 200  IF (IT.GT.50) GO TO 300
      EE=0
      DD=0
      EEE=0
      DO 71 K=1,NO
      PP=P(K)**(-AL)*V(K)
      EE=EE+PP*PLO(K)
      EEE=EEE+PP*PLO(K)**2
 71   DD=DD+PP
      WRITE(6,22) EE,EEE,PP,DD
 22   FORMAT(2X,'EE',E12.4,2X,'EEE',E12.4,2X,'PP',E12.4,2X,'DD',E12.4)
      F=CC-(OO/DD)*EE
      DF=-(EE**2)*OO/(DD**2)+EEE*OO/DD
      WRITE(6,21) F,DF,AL
 21   FORMAT(2X,'F',E12.4,2X,'DF',E12.4,F6.2)
      IF(ABS(F).LT.0.00001) GO TO 10
      IF(DF.GT.0.0001.AND.DF.LT.10000) GO TO 20
      AL=AL+.1
      IT=IT+1
      GO TO 200
 20   AL=AL-F/DF
      IT=IT+1
      GO TO 200
 10   AK=LOG10(OO/DD)
      WRITE(6,11) IT,AL,AK
 11   FORMAT(2X,'NUMERO DI ITERAZIONE',I4,2X,'PENDENZA',F6.2,2X,
     1'NORMALIZ',E12.4)
      GO TO 400
 300  WRITE(6,12)
 12   FORMAT('NON CONVERGE')
 400  VV=0
      DO 50 K=1,NO
 50   VV=VV+NOS(K)*LOG(V(K))
      DF=0
      DO 31 K=1,NO
      DFF=1
      IFF=NOS(K)
```

```
      IF(JFF.LT.1)GO TO 33
      DO 32 J=1,JFF
32    OFF=OFF*J
33    OFF=LOG(OFF)
31    OF=OF+OFF
      SS=-2*OO*AK*2.30259+2*AL*CC-2*VV+2*(10**AK)*DD+2*OF
      WRITE(6,26)SS
26    FORMAT(10X,'SS',E12.4)
      STOP
      END
```

APPENDICE n° 2 SSS.FOR

```
      DIMENSION RO(20),ERO(20),PLO(20),V(20),NOS(20),P(20)
      REAL*8 DD,PP,VV,OF,OFF,CC,AL,OO,AK
      K=1
      READ(5,*)PO,Z
      WRITE(6,*)PO,Z
2     READ(5,*,END=100)RO(K),ERO(K),PLO(K),V(K),NOS(K)
      WRITE(6,7) RO(K),ERO(K),PLO(K),V(K),NOS(K)
7     FORMAT(10X,2E10.2,F7.3,E10.2,I5)
      K=K+1
      GO TO 2
100   NO=K-1
      DO 3 K=1,NO
      P(K)=10**PLO(K)
3     PLO(K)=PLO(K)*2.30259
      CC=0
      DO 5 K=1,NO
5     CC=CC+NOS(K)*PLO(K)
      OO=0.
      DO 6 K=1,NO
6     OO=OO+NOS(K)
      READ(9,*) AL,AK,I
      DD=0
      DO 71 K=1,NO
      PP=P(K)**(-AL)*V(K)
71    DD=DD+PP
      IF(I.EQ.0)GO TO 72
      AK=LOG10(OO/DD)
72    VV=0
      DO 50 K=1,NO
50    VV=VV+NOS(K)*LOG(V(K))
      OF=0
      DO 31 K=1,NO
      OFF=1
      JFF=NOS(K)
      IF(JFF.LT.1)GO TO 33
      DO 32 J=1,JFF
32    OFF=OFF*J
33    OFF=LOG(OFF)
31    OF=OF+OFF
      SS=-2*OO*AK*2.30259+2*AL*CC-2*VV+2*(10**AK)*DD+2*OF
      WRITE(6,30)AL,AK,SS
30    FORMAT(5X,'AL',F6.2,5X,'AK',F6.2,5X,'SS',F6.2)
      STOP
      END
```