Prova scritta di "Analisi statistica dei dati sperimentali"
10/1/2003

Problema 1

La radiazione cosmica al suolo e' composta per il 90% di muoni (μ) e per il 10% di elettroni (e). Un rivelatore di particelle e' efficiente al 95% per gli e ed al 15% per i μ.
a) Calcolare la probabilita' di misurare un evento qualunque (Trigger)
b) Calcolare la percentuale di eventi di e e di μ rivelati
c) Calcolare la percentuale di eventi di e e di μ rivelati nel caso di un doppio rivelatore in coincidenza con le stesse caratteristiche del primo.

Soluzione

La formula delle probabilita' totali

$$P(T) = \sum_i P(T|H_i)P(H_i)$$

con $\sum_i P(H_i) = H_1 + H_2 = 1,\ H_1 = \mu,\ H_2 = e$. Dai dati del problema:

$$\begin{array}{rcl} P(\mu) & = & 0.90 \quad ; \quad P(e) \ = \ 0.10 \\ P(T|\mu) & = & 0.15 \quad ; \quad P(T|e) \ = \ 0.95 \end{array}$$

a)

$$P(T) = P(T|\mu)P(\mu) + P(T|e)P(e) = 0.15 \cdot 0.90 + 0.95 \cdot 0.10 = 0.23$$

b)

$$P(e|T) = (P(T|e)P(e))/P(T) = 0.095/0.23 = 0.41$$

$$P(\mu|T) = 1 - P(e|T) = 0.59$$

c)

$$P(TT|\mu) = P(T|\mu)P(T|\mu) = 0.15 \cdot 0.15 = 0.0225$$

(in quanto i rivelatori hanno le stesse caratteristiche e sono comunque indipendenti). Analogamente:

$$P(TT|e) = P(T|e)P(T|e) = 0.95 \cdot 0.95 = 0.9025$$

$$P(TT) = P(TT|\mu)P(\mu) + P(TT|e)P(e)$$

$$P(e|TT) = \frac{P(TT|e)P(e)}{P(TT)} = \frac{0.9025 \cdot 0.10}{0.0225 \cdot 0.9 + 0.9025 \cdot 0.10} = 0.82$$

$$P(\mu|TT) = 1 - P(e|TT) = 0.18$$

Problema 2

La grandezza x_i e' funzione delle grandezze u e v misurate, con deviazione standard s_u e s_v. a) Determinare l'incertezza relativa nei seguenti casi:
$x_1 = (u+v)/2$
$x_2 = (u-v)/2$
$x_3 = u^2 + v^2$
$x_4 = uv^2$
$x_5 = u \ln v$
b) Se u e v sono misure a statistica poissoniana, determinare l'espressione analitica dell'incertezza relativa.
c) se $u = 16$ e $v = 9$ conteggi, quale delle x_i ha la maggiore incertezza relativa?

Soluzione

$$s_x^2 = \left(\frac{\partial x}{\partial u}\right)^2 s_u^2 + \left(\frac{\partial x}{\partial v}\right)^2 s_v^2$$

a) x_3 e' sempre positivo, indipendentemente dai valori misurati di u e v

$s_{x_1}^2 = \frac{1}{4}s_u^2 + \frac{1}{4}s_v^2 \Longrightarrow s_{x_1}/|x_1| = \sqrt{s_u^2 + s_v^2}/|u+v|$
$s_{x_2}^2 = \frac{1}{4}s_u^2 + \frac{1}{4}s_v^2 \Longrightarrow s_{x_2}/|x_2| = \sqrt{s_u^2 + s_v^2}/|u-v|$
$s_{x_3}^2 = (2u)^2 s_u^2 + (2v)^2 s_v^2 \Longrightarrow s_{x_3}/x_3 = 2\sqrt{u^2 s_u^2 + v^2 s_v^2}/(u^2+v^2)$
$s_{x_4}^2 = (v^2)^2 s_u^2 + (2uv)^2 s_v^2 \Longrightarrow s_{x_4}/|x_4| = \sqrt{s_u^2/u^2 + 4s_v^2/v^2}$
$s_{x_5}^2 = (\ln v)^2 s_u^2 + (-u/v)^2 s_v^2 \Longrightarrow s_{x_5}/|x_5| = \sqrt{s_u^2/u^2 + s_v^2/(v \ln v)^2)}$

b) se u e v sono conteggi, escludendo $v = 0, 1$, i valori sono sempre positivi, a parte x_2 che puo' essere anche negativo

$s_{x_1}/x_1 = \sqrt{u+v}/(u+v)$
$s_{x_2}/|x_2| = \sqrt{u+v}/|u-v|$
$s_{x_3}/x_3 = 2\sqrt{u^3+v^3}/(u^2+v^2)$
$s_{x_4}/x_4 = \sqrt{1/u + 4/v}$
$s_{x_5}/x_5 = \sqrt{1/u + 1/(v(\ln v)^2)}$

c) $s_{x_1}/x_1 = 20\%$
$s_{x_2}/x_2 = 71.4\%$ (incertezza maggiore)
$s_{x_3}/x_3 = 41\%$
$s_{x_4}/x_4 = 71.2\%$
$s_{x_5}/x_5 = 29\%$

Problema 3

Un gruppo di 10 studenti ha conseguito i seguenti risultati nelle prove scritta ed orale:

prova scritta	x	27	20	14	30	5	7	16	9	21	27
prova orale	y	27	21	20	30	14	20	23	26	30	24

a) Se il voto finale e' basato per il 70% sul voto dell'orale, valutare il valore medio e la varianza della distribuzione dei voti finali (ricordarsi che il voto finale di ogni singolo studente e' un numero intero...)
b) Verificare l'ipotesi che un buon scritto sia il preludio di un buon orale (i.e. se vi sia correlazione...).
c) valutare i coefficienti della relazione lineare $y = a + bx$ (si assuma $s_{y_i} = s$ per l'incertezza sui valori y, mentre si assuma nulla l'incertezza su x).

Soluzione

a) Il voto finale e' dato da $f = x * 0.3 + y * 0.7$ arrotondando il risultato all'intero piu' vicino. Di conseguenza, i valori di f diventano:

prova scritta	x	27	20	14	30	5	7	16	9	21	27
prova orale	y	27	21	20	30	14	20	23	26	30	24
prova finale	f	27	21	18	30	11	16	21	21	27	25

$$\overline{f} = \frac{1}{N}\sum_i f_i$$

$$s^2 = \frac{1}{N-1}\sum_i (f_i - \overline{f})^2$$

$$\overline{f} = 22;\ s = 6$$

b) per verificare se vi e' correlazione tra lo scritto e l'orale, si utilizza il calcolo del coefficiente di correlazione r

$$N = 10;\ \sum x_i = 176;\ \sum x_i^2 = 3806;\ \sum y_i = 235;\ \sum y_i^2 = 5747;\ \sum x_i y_i = 4419$$

$$r = \frac{N\sum x_i y_i - \sum x_i \sum y_i}{[N\sum x_i^2 - (\sum x_i)^2]^{1/2}[N\sum y_i^2 - (\sum y_i)^2]^{1/2}} = 0.70976$$

Considero ora l'ipotesi "i dati non sono correlati", e faccio il test sul valore del coefficiente di correlazione lineare:

$$P_N(|r| > 0.7096) \simeq 0.02 < 0.05$$

Il risultato implica che l'ipotesi puo' essere rigettata al livello di confidenza del 5% (ovvero ammettere che i dati sono correlati).

c) Nell'ipotesi che le incertezze su y_i siano uguali e quelle su x_i trascurabili, le espressioni per il calcolo di a e b si semplificano notevolmente e non hanno bisogno della conoscenza di $s_y = s$; possiamo utilizzare le comuni calcolatrici tascabili:

$$\Delta = N\sum_i x_i^2 - (\sum_i x_i)^2 = 7084$$

$$a = \frac{\sum_i x_i^2 \cdot \sum_i y_i - \sum_i x_i \cdot \sum_i x_i y_i}{\Delta} = 16.46$$

$$b = \frac{N\sum_i x_i y_i - \sum_i x_i \cdot \sum_i y_i}{\Delta} = 0.399$$

Poiche' non si conosce l'incertezza su y, questa puo' essere ricavata mediante l'espressione

$$s_y^2 = s^2 = \frac{1}{N-2}\sum_i (y_i - a - b \cdot x_i)^2 = 13.9$$

e successivamente valutare le incertezze su a e b:

$$s_a = \sqrt{\frac{s^2 \sum_i x_i^2}{\Delta}} = 2.74; \; s_b = \sqrt{\frac{Ns^2}{\Delta}} = 0.14$$

$$a = 16 \pm 3; \; b = 0.40 \pm 0.14$$