Soluzioni - foglio 6

2 maggio 2023

E6.1 Siano $X_1, ..., X_n$, n misure dell'altezza μ di una persona (in centimetri). Assumiamo che X_i siano indipendenti e identicamente distribuite con media μ e deviazione standard $\sigma = 1$ cm. La media delle misure $n^{-1}\sum_i X_i$ costituisce una stima dell'altezza μ. Utilizzando la disuguaglianza di Chebyshev, calcolare il numero di misure n necessarie per determinare μ con una precisione di 0.5 cm e con una confidenza pari al 90%.
Soluzione:
Data la disuguaglianza di Chebyshev:

$$\forall \epsilon > 0,\ P\{|X - \mu| \geq \epsilon\} \leq \frac{\sigma^2}{\epsilon^2},$$

nel nostro caso abbiamo che $E(\sum_i X_i/n) = \mu$ e $Var(\sum_i X_i/n) = \sigma^2/n$, quindi:

$$P\{|\sum_i \frac{X_i}{n} - \mu| \geq 0.5\} \leq \frac{\sigma^2}{n0.25} = \frac{4\sigma^2}{n}.$$

Quindi per imporre una precisione di 0.5 cm con una confidenza pari al 90% bisogna porre $4\sigma^2/n = 10\%$. Il numero di misure necessarie è $n = 40\sigma^2$.

E6.2 Il bosone di Higgs H è una particella fondamentale che può decadere in diversi stati finali con le seguenti preobabilità: due quark bottom $b\bar{b}$ (P=0.57), due bosoni W^+W^- (P=0.21), due gluoni gg (P=0.09), due leptoni tau $\tau\bar{\tau}$(P=0.06), due quark charm $c\bar{c}$(P=0.03), due bosoni ZZ (P=0.03) o altro (chiamiamo questo stato γ) con P=0.01. Calcolare l'informazione di Shannon per ogni stato, l'entropia di Shannon e l'entropia grezza.
Soluzione:
- $S(b\bar{b}) = -\log_2(0.57) = 0.811$, $S(W^+W^-) = 2.25$, $S(gg) = 4.06$, $S(\tau\bar{\tau}) = 4.06$, $S(c\bar{c}) = 5.06$, $S(ZZ) = 5.06$, $S(\gamma) = 6.64$.
- $H = \sum_{i=1}^{7} P_i \log P_i^{-1} = 1.79$.
- $H_0 = \log N = 2.81$, con $N = 7$ possibili stati finali.

E6.3 Se $H(X) = 4$, $H(Y) = 3$ e $H(X,Y) = 5$, calcola le entropie condizionate.
Soluzione:
Poichè

$$H(X,Y) = H(Y) + H(X|Y),$$

per trovare $H(X|Y)$ e $H(Y|X)$ posso scrivere che

$$H(X|Y) = H(X,Y) - H(Y) = 2 \text{ e } H(Y|X) = H(X,Y) - H(X) = 1.$$

E6.4 Siano X e Y variabili casuali discrete indipendenti. Usando solo la definizione di entropia, dimostrare che $H(X,Y) = H(X) + H(Y)$.
Soluzione:
Dalla definizione di entropia possiamo scrivere che

$$H(X,Y) = \sum_i \sum_j p(x_i, y_j) \log_2 p(x_i, y_j)^{-1},$$

se X e Y sono indipendenti si avrà che $p(x_i, y_j) = p(x_i)p(y_j)$, per cui sostituendo nella definizione si ottiene:

$$H(X,Y) = \sum_i \sum_j p(x_i, y_j) \log_2(p(x_i)p(y_j))^{-1}.$$

Possiamo quindi scrivere

$$H(X,Y) = \sum_i \sum_j p(x_i)p(y_j) \left[\log_2 p(x_i)^{-1} + \log_2 p(y_j)^{-1}\right]$$

$$= \sum_i p(x_i) \left\{ \sum_j p(y_j) \log_2 p(x_i)^{-1} + \sum_j p(y_j) \log_2 p(y_j)^{-1} \right\}$$

$$= \sum_i p(x_i) \left\{ \sum_j p(y_j) \log_2 p(x_i)^{-1} + H(Y) \right\}$$

$$= \sum_i p(x_i) \log_2 p(x_i)^{-1} + \sum_i p(x_i) H(Y) = H(X) + H(Y).$$

E6.5 Sia dato $X = \{x_1, x_2, x_3, x_4\}$e l'insieme di interi $\{L_1, L_2, L_3, L_4\}$ con $L_1 = 1, L_2 = 2, L_3 = 2$ e $L_4 = 2$. Per quale motivo non può esistere una codifica istantanea C che abbia gli interi L_i come lunghezze delle rappresentazioni $C(x_i)$?
Soluzione:
$2^{-1} + 2^{-2} + 2^{-2} + 2^{-2} = 1.25 > 1$.

E6.6 Supponiamo di avere 27 palline uguali per forma e colore, una delle quali è di peso inferiore delle altre 26. Avendo a disposizione una bilancia a due piatti, determina una strategia capace di individuare la pallina più leggera con tre pesate.
Soluzione:
Sono sufficienti 3 pesate. La pallina più leggera è una qualunque delle 27. Per massimizzare l'informazione di Shannon che otteniamo con ogni pesata dobbiamo dividere le 27 palline in tre gruppi da 9. Confrontando i primi due gruppi riduciamo il problema al caso di 9 palline (la pallina è nel gruppo più leggero o nel terzo se i piatti sono in equilibrio). Dividiamo le 9 palline in tre gruppi da 3. Confrontando i primi due gruppi riduciamo il problema al caso di 3 palline.