Riposso di statistica e definizione di Ergodicità

$$\langle\ \rangle_t = \langle\ \rangle_S = \langle\ \rangle_e$$

Sia data una variabile casuale x e sia $f(x)$ la sua pdf cioè $p(x_0 \leq x \leq x_0+dx) = f(x_0)dx$ $\int_{D_x} f(x)\,dx = 1$ (certezza)

Sono rilevanti le seguenti grandezze derivate dalla conoscenza di f

$$E[g] = \int_{D_x} g(x)\cancel{f_0}(x)\,dx$$ si dice valore di aspettazione di g

Nel caso la pdf sia quella di più variabili casuali es x, y, z

$$E[g] \iiint_{D_{xyz}} g(x,y,z)\, f(x,y,z)\, dx\,dy\,dz$$

Osservazione

Generalmente $f(x)$ non è nota quindi si eseguono dei campionamenti $x_1, x_2, \dots x_N$ e si prova a determinare le proprietà, la forma o tutte l'informazione di f.

Si costruiscono gli stimatori delle grandezze ottenibili conoscendo la f

Valore di aspettazione di x

$$E[x] := \int_{D_x} x\, f(x)\, dx$$

la media pesata della variabile casuale dove i pesi sono la probabilità che x si manifesti

Lo stimatore di $E[x]$ è $\bar{x} := \frac{1}{N}\sum_{i=1}^{N} x_i$

Il valore di aspettazione di $\bar{x}$ è proprio $E[x]$

Dimostriamolo utilizzando la linearità dell'operatore integrale valore di aspettazione

$$E[\bar{X}] = E[\frac{1}{N}\sum_{i=1}^{N} x_i] = \frac{1}{N}\sum_{i=1}^{N} E[x_i] = \frac{1}{N} \cdot N E[X] = E[X]$$

Consideriamo i momenti statistici che sono definiti come segue

$$\mu^n := E[X^n]$$

Ci sono anche i momenti centrali che sono i seguenti

$$\hat{\mu}^n = E[(X-\mu^1)^n] = E[(X-E[X])^n]$$

Consideriamo il momento centrale del secondo ordine

$$\hat{\mu}^2 = E[(X-\mu^1)^2] = E[(X-E[X])^2] = Var[X] \quad \text{varianza di } x$$

Forma notevole della varianza di x

$$Var[X] = E[(X-E[X])^2] = E[(X^2 - 2XE[X] + E[X]^2)] =$$

$$= E[X^2] - 2E[X]E[X] + E[X]^2 = E[X^2] - E[X]^2$$

Lo stimatore statistico della varianza è $S^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2$
perché dividiamo per $N-1$ e non per N?

$$E[\sum_{i=1}^{N}(x_i-\bar{x})^2] = E[(\sum_{i=1}^{N}(x_i^2 - 2x_i\bar{x} + \bar{x}^2))] =$$

$$= E[\sum_{i=1}^{N} x_i^2 - 2\frac{1}{N}\sum_{j=1}^{N} x_j \sum_{i=1}^{N} x_i + \frac{N}{N^2}\sum_{j=1}^{N} x_j \cdot \sum_{k=1}^{N} x_k] =$$

$$= \sum_{i=1}^{N} E[X^2] - \frac{2}{N}\sum_{j,i=1}^{N} E[x_j x_i] + \frac{\cancel{N}}{\cancel{N^2}}\sum_{j,k=1}^{N} E[x_j x_k]] =$$

$$= N E[X^2] - \frac{2}{N}\sum_{i=1}^{N} E[x_i^2] - \frac{2}{N}\sum_{\substack{i,j=1 \\ i\neq j}}^{N} [x_j x_i] +$$

$$+ \frac{1}{N}\sum_{j=1}^{N} E[x_j^2] + \frac{1}{N}\sum_{\substack{j,k=1 \\ j\neq k}}^{N} E[x_j x_k] =$$

$$= N E[X^2] - \frac{N}{N} E[X^2] - \frac{1}{N}\sum_{\substack{j,k \\ j\neq k}}^{N} [x_j x_k]$$

Assumiamo che $E[x_j x_k] = E[x_j]E[x_k] = E[x]^2$ ovvero le variabili casuali sono scorrelate

$$E\left[\sum_{i=1}^{N}(x_i-\bar{x})^2\right] = (N-1)E[x^2] - \frac{1}{\cancel{N}}\cancel{N}(N-1)E[x]^2 =$$

$$= (N-1)E[x^2] - (N-1)E[x]^2 =$$

$$= (N-1)\left(E[x^2] - E[x]^2\right)$$

$$\stackrel{!}{=} (N-1)\,E[(x-E[x])^2]$$

$$\stackrel{!}{=} (N-1)\,\mathrm{Var}[x]$$

Quindi $$E\left[\frac{1}{N-1}\sum_{i=1}^{N}(x_i-\bar{x})^2\right] = \mathrm{Var}[x]$$

se N è molto grande $N \gg 10$ allora non c'è molta differenza nell'usare lo stimatore biassato

$$s'^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})^2$$

<u>Correlazione e covarianza</u>

Siano x, y due variabili casuali e siano $f(x)$ e $g(y)$ tali che $h(x,y) = f(x)\cdot g(y)$ la Jpdf è il prodotto delle due ovvero sono <u>variabili scorrelate</u> $\Rightarrow f(x)\,g(y)\,dx\,dy$

$$P(x_0 \le x < x_0+dx \wedge y_0 \le y \le y_0+dy) =$$

$$= f(x_0)dx\, g(y_0)\,dy$$

Allora si ha quanto segue

$$E[t(x)\,z(y)] = \int_{D_x}dx\int_{D_y}dy\, f(x)\,g(y)\,t(x)\,z(y) = \int_{D_x}dx\, f(x)\,t(x)\int_{D_y}dy\, g(y)\,z(y)$$

$E[t]$ $E[z]$

$$E[(x-E[x])(y-E[y])] = E[(xy - yE[x] - xE[y] + E[x]E[y])] =$$

$$= E[xy] - E[x]E[y] - E[x]E[y] + E[x]E[y] = E[xy] - E[x]E[y]$$

Forma notevole per la covarianza ~~se x y scorrelate~~

$$\mathrm{cov}[x,y] = E[(x-E[x])(y-E[y])] = E[xy] - E[x]E[y]$$

Se le grandezze sono scorrelate $E[xy] = E[x]E[y]$

da cui $\mathrm{cov}[x,y] = 0$

In generale si può definire il coefficiente di correlazione

$$r := \frac{\mathrm{cov}[x\,y]}{\mathrm{Var}[x]^{1/2}\,\mathrm{Var}[y]^{1/2}}$$

$$1 \leq r \leq 1$$

Se x e y sono scorrelate $\mathrm{cov}[xy] = 0 \rightarrow r = 0$

Se $x = \alpha y + \beta$ $\quad \mathrm{cov}[xy] = \alpha \mathrm{Var}[x]$

correlate linearmente $\quad \mathrm{Var}[y] = \alpha^2 \mathrm{Var}[x]$

da cui $r = \dfrac{\alpha \mathrm{Var}[x]}{\mathrm{Var}[x]^{1/2}\,\alpha\,\mathrm{Var}[x]^{1/2}} = 1$

Se $x = -\alpha y + \beta$ $\quad \mathrm{cov}[x,y] = -\alpha \mathrm{Var}[x]$

anticorrelate linearmente $\quad \mathrm{Var}[y] = \alpha^2 \mathrm{Var}[x]$

da cui $r = \dfrac{-\alpha \mathrm{Var}[x]}{\mathrm{Var}[x]^{1/2}\,\alpha\,\mathrm{Var}[x]^{1/2}} = -1$