Quesito 1 – Laboratorio

Vedi notebook

Quesito 2 - Laboratorio

Vedi notebook

Quesito 1 - Teoria

a) Sia data una successione di v.a. indipendenti $(X_n)_{n\in\mathbb{N}}$ definite tutte nello stesso spazio di probabilità. Assumiamo che ciascuna delle X_n abbia media μ e varianza σ^2. Sia X una v.a. definita nello stesso spazio di probabilità. Diremo allora che la successione $(X_n)_{n\in\mathbb{N}}$ converge in probabilità a X se

$$\lim_{n\to\infty} P(|X_n - X| \geq \eta) = 0 \quad \forall\, \eta > 0.$$

La convergenza in probabilità si indica con

$$X_n \xrightarrow{\mathcal{P}} X.$$

b) Data una successione di v.a. come sopra avremo che la media campionaria $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ convergerà in probabilità alla media teorica μ. In formule avremo

$$\lim_{n\to\infty} P(|\overline{X}_n - \mu| \geq \eta) = 0 \quad \forall\, \eta > 0.$$

c) La disuguaglianza di Chebishev ci assicura che se X è una v.a. con media e varianza finiti

$$P(|X - E[X]| \geq \eta) \leq \frac{\mathrm{Var}(X)}{\eta^2} \quad \forall\, \eta > 0.$$

Riscriviamo la disuguaglianza di Markov per la media campionaria

$$P(|\overline{X}_n - E[\overline{X}_n]| \geq \eta) \leq \frac{\mathrm{Var}(\overline{X}_n)}{\eta^2} \quad \forall\, \eta > 0.$$

Ricordandoci che $E[\overline{X}_n] = \mu$ e $\mathrm{Var}(\overline{X}_n) = \frac{\sigma^2}{n}$ (**Commento**: non è importante dimostrarlo se lo sapete. Se vi rimane tempo e lo volete dimostrare certamente non mi offendo) avremo

$$P(|\overline{X}_n - \mu| \geq \eta) \leq \frac{\sigma^2}{n\eta^2} \quad \forall\, \eta > 0,$$

che chiaramente tende a zero per $n \to \infty$.

d) Immaginiamo di avere un campione di dati provenienti da una distribuzione che non conosciamo. La legge dei grandi numeri ci rassicura che se il nostro campione è abbastanza grande (in generale $n \geq 30$) potremo stimare la media teorica con la media campionaria. In più, maggiore è l'ampiezza del campione, maggiore sarà l'accuratezza della stima.

Quesito 2 - Teoria

a) Sia X una v.a., discreta o continua. Definiamo funzione di ripartizione la funzione $F_X(t) : \mathbb{R} \to [0, 1]$ tale che

$$F_X(t) = P(X \leq t).$$

Nel caso in cui X sia una v.a. discreta di densità $p(x)$, avremo $F_X(t) = P(X \leq t) = \sum_{x\leq t} p(x)$.

Nel caso in cui X sia una v.a. continua di densità $f(x)$, avremo $F_X(t) = P(X \leq t) = \int_{-\infty}^t dx\, f(x)$.

b) La funzione è una funzione che ad ogni valore dell'insieme dei reali asocia una probabilità. Nel caso in cui la v.a. sia continua, la funzione di ripartizione misura l'area sottesa alla densità di probabilità nell'intervallo $]-\infty, t]$.

c) Si definisce quantile di ordine α un valore q_α che divide il campione in due parti proporzionali ad α e $1-\alpha$. Data una v.a. X continua, il quantile q_a indica il punto dell'asse x che divide l'area sottesa alla densità in due parti proporzionali a α (alla sinistra di q_α) e a $1-\alpha$ (a destra di q_α). Quantile e funzione di ripartizione sono collegati dalla relazione $F_X(q_\alpha) = \alpha$.

d) $P(X \in [a, b]) = P(a \leq X \leq b) = F_X(b) - F_X(a)$.