

Dispense essenziali di Probabilità e Statistica

Matteo Bitussi
Laurea in Informatica, Unitn

Anno accademico 2018-2019

Indice

1	Probabilità	3
1.1	Insieme delle parti di Ω : $P(\Omega)$	3
1.2	Tribù (o σ -algebra)	3
1.3	Spazio Probabilizzabile	3
1.4	Definizione di Probabilità	3
1.5	Spazio proabilizzato	3
1.6	Regole di calcolo delle probabilità	4
1.6.1	Regola 1	4
1.6.2	Regola 2	4
1.6.3	Regola 3	4
1.6.4	Regola 4 (Disuguaglianza di Bonferroni)	4
2	Calcolo combinatorio	5
2.1	Disposizioni con ripetizione	5
2.2	Disposizioni senza ripetizione	5
2.3	Permutazioni	5
2.4	Combinazioni	5
2.5	Cardinalità dell'insieme delle parti di un insieme finito	5
3	Probabilità sui reali	6
3.1	Tribù borelliana	6
3.2	Costruzione di una funzione di probabilità su $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$	6
3.3	Probabilità condizionale	7
3.4	Classe Completa di eventi	7
3.5	Teorema delle Probabilità Totali	7
3.6	Teorema di Bayes	7
3.7	Indipendenza stocastica	7
3.8	Tribù indipendenti	7
4	Variabili Aleatorie	8
4.1	Variabili aleatorie e Tribù	8
4.1.1	Teorema 10	8
4.1.2	Teorema 12	8
4.2	Variabili aleatorie e funzioni di probabilità	8
4.3	Variabili aleatorie discrete	9
4.3.1	Funzione di probabilità (o densità discreta)	9
4.3.2	Teorema	9
4.3.3	Distribuzione Binomiale	9
4.3.4	Funzione di ripartizione	9
4.3.5	Distribuzione Geometrica	10
4.3.6	Distribuzione Binomiale negativa (o di Pascal)	10
4.3.7	Distribuzione di Poisson	10
4.4	Variabili aleatorie continue	11
4.4.1	Densità	11
4.4.2	Variabili aleatorie assolutamente continue	11
4.4.3	Densità e funzione di ripartizione	11
4.4.4	Distribuzione Normale (o di Gauss)	11

4.4.5	Standardizzazione di una Normale	12
4.4.6	Distribuzione Esponenziale	12
4.4.7	Trasformazione di variabili aleatorie p.104 (manca)	12
4.5	Speranza matematica o valore atteso per v.a. discrete	12
4.6	Momenti	12
5	Variabili Aleatorie Doppie	13
5.1	Funzione di probabilità congiunta (discreta)	13
5.2	Variabili aleatorie doppie dotate di densità	13
5.2.1	Densità marginali	13
5.3	Distribuzioni condizionali per v.a. (Probabilità condizionale di $X Y = y$)	14
5.4	Distribuzioni condizionali e indipendenza per v.a. (p137 dispense B)	14
5.5	Funzioni di ripartizioni condizionali	14
5.6	Variabili aleatorie condizionali e speranza matematica	14
5.7	Speranza matematica della speranza matematica condizionale	14
5.8	Varianza e Varianza condizionale (Scomposizione della varianza)	15
5.9	Dipendenza in media	15
5.10	Rapporto di correlazione	15
5.11	Covarianza e correlazione	15
5.12	Varianza di una combinazione lineare di v.a.	15
6	Teoremi limite della probabilità	16
6.1	Convergenza in probabilità (o debole)	16
6.2	Convergenza in media quadratica	16
6.3	Disuguaglianza di Markov	17
6.4	Disuguaglianza di Chebychev	17
6.5	Somme di variabili casuali	17
6.6	Legge debole dei grandi numeri	17

Introduzione

Questa dispensa è pensata per raccogliere le informazioni essenziali necessarie per lo svolgimento degli esercizi durante l'anno e/o per l'esame finale. Per questo motivo non saranno approfondite e non potranno sostituire quelle fornite dal professore.

Capitolo 1

Probabilità

1.1 Insieme delle parti di Ω : $P(\Omega)$

Dato l'insieme Ω si dice **Insieme delle Parti** o **Insieme Potenza** di Ω l'insieme $P(\Omega)$ di tutti i possibili sottoinsiemi di Ω .

1.2 Tribù (o σ -algebra))

Una classe \mathcal{A} di parti di un insieme Ω si dice una **Tribù** se:

- $\Omega \in \mathcal{A}$
- Se $A \in \mathcal{A}$ allora $A^c \in \mathcal{A}$
- Se $A_1, \dots, A_i \in \mathcal{A}$ allora $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$

1.3 Spazio Probabilizzabile

Dato uno spazio campionario Ω e una tribù \mathcal{A} su Ω , la coppia (Ω, \mathcal{A}) è detto **Spazio Probabilizzabile**

1.4 Definizione di Probabilità

Dato uno spazio probabilizzabile (Ω, \mathcal{A}) , una **Probabilità** Pr è un'applicazione $Pr : \mathcal{A} \rightarrow \mathbb{R}^+$ tale che:

- (non negatività) se $A \in \mathcal{A}$ allora $Pr(A) \geq 0$
- (normalizzazione) $Pr(\Omega) = 1$
- (σ -additività) Se $\{A_i\}_{i=1}^{\infty}$ è una successione di eventi di \mathcal{A} a due a due incompatibili (cioè $A_i \cap A_j = \emptyset, i \neq j$), allora

$$Pr\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} Pr(A_i)$$

1.5 Spazio proabilizzato

La terna $(\Omega, \mathcal{A}, Pr)$ dove Ω è uno spaio campionario, \mathcal{A} è una Tribù su Ω e Pr è una funzione di probabilità $Pr : \mathcal{A} \rightarrow \mathbb{R}^+$, è detta **Spazio di Probabilità** o anche spazio di Kolmogorov.

1.6 Regole di calcolo delle probabilità

1.6.1 Regola 1

Se A è un evento di probabilità $Pr(A)$ allora la probabilità che A non si verifichi è

$$Pr(A^c) = 1 - Pr(A)$$

1.6.2 Regola 2

Se A e B sono due eventi, allora la probabilità che se ne verifichi almeno uno è data da

$$Pr(A \cup B) = Pr(A) + Pr(B) - Pr(A \cap B)$$

1.6.3 Regola 3

Se A è un evento che implica l'evento B , cioè se $A \subseteq B$, allora

$$Pr(B) = Pr(A) + Pr(B \cap A^c) \geq Pr(A)$$

1.6.4 Regola 4 (Disuguaglianza di Bonferroni)

Se A_1, A_2, \dots, A_n non sono eventi, allora

$$\sum_{i=1}^n Pr(A_i) - \sum_{1 \leq i < j \leq n} Pr(A_i \cap A_j) \leq Pr\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n Pr(A_i), n \geq 1$$

Capitolo 2

Calcolo combinatorio

2.1 Disposizioni con ripetizione

Dato un insieme $S = a_1, a_2, \dots, a_n$ di n oggetti distinti, il numero degli allineamenti che si possono formare con k oggetti scelti tra gli n - ritenendo diversi due allineamenti, o perchè contengono oggetti differenti o perchè gli stessi oggetti si susseguono in ordine diverso o, infine, perchè uno stesso oggetto si ripete un numero diverso di volte - è dato da

$$D_{n,k}^* = n^k$$

Ogni allineamento si dice disposizione con ripetizione di n oggetti di classe k .

2.2 Disposizioni senza ripetizione

Dato un insieme $S = a_1, a_2, \dots, a_n$ di n oggetti distinti, il numero degli allineamenti che si possono formare con $1 \leq k \leq n$ oggetti scelti tra gli n - ritenendo diversi due allineamenti o perchè contengono oggetti differenti o perchè gli stessi oggetti si susseguono in ordine diverso - è dato da

$$D_{n,k} = n(n-1)(n-2) \dots (n-k+1)$$

Ogni allineamento si dice disposizione semplice o senza ripetizione di n oggetti di classe k

2.3 Permutazioni

Dato un insieme $S = a_1, a_2, \dots, a_n$ di n oggetti distinti, il numero degli allineamenti che si possono formare con tutti essi - ritenendo diversi due allineamenti perchè gli oggetti si susseguono in ordine diverso - è dato da $n!$

2.4 Combinazioni

Dato un insieme $S = a_1, a_2, \dots, a_n$ di n oggetti distinti, il numero degli allineamenti che si possono formare con $1 \leq k \leq n$ oggetti scelti tra gli n - ritenendo diversi due allineamenti solo perchè contengono oggetti differenti - è dato da

$$C_{n,k} = \frac{D_{n,k}}{k!}$$

Ogni allineamento si dice combinazione senza ripetizione di n oggetti di classe k

2.5 Cardinalità dell'insieme delle parti di un insieme finito

Sia $S_n = a_1, a_2, \dots, a_n$ un insieme di n oggetti distinti, allora la cardinalità di $P(S)$ è 2^n

Capitolo 3

Probabilità sui reali

3.1 Tribù borelliana

Si chiama Tribù Borelliana di \mathbb{R} , e si denota con $\mathcal{B}(\mathbb{R})$, la tribù generata su \mathbb{R} dalla classe di tutti gli intervalli $(a, b]$ di \mathbb{R} . I suoi elementi si chiamano gli insiemi borelliani di \mathbb{B} . e lo spazio $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ è uno spazio probabilizzabile.

Elementi della tribù Borelliana

La tribù di Borel su \mathbb{R} contiene anche i seguenti Elementi

- $(a, b]$
- $[a, b]$
- $[a, b)$
- $(-\infty, b]$
- (a, ∞)
- i singoletti di \mathbb{R}
- gli insiemi finiti di \mathbb{R}
- gli insiemi numerabili di \mathbb{R}

3.2 Costruzione di una funzione di probabilità su $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$

Per procedere all'assegnazione di una funzione di Probabilità agli eventi di $\mathcal{B}(\mathbb{R})$, si fissa la probabilità da attribuire agli intervalli $(a, b]$ mediante una funzione $F(x)$ che è

- non decrescente
- continua da destra per ogni $x \in \mathbb{R} : \lim_{x \rightarrow x_0^+}(x) = F(x_0)$ per ogni $x_0 \in \mathbb{R}$
- $\lim_{x \rightarrow +\infty} F(x) = 1$
- $\lim_{x \rightarrow -\infty} F(x) = 0$

ponendo

$$Pr((a, b]) = F(b) - F(a)$$

Ad ogni insieme di $\mathcal{B}(\mathbb{R})$ è quindi possibile attribuire una probabilità. Il calcolo effettivo di $Pr(A)$ può essere fatto in modo semplice quando A è

- un intervallo
- un'unione numerabile di intervalli disgiunti

$$Pr\left(\bigcup_{i=1}^{\infty} (a_i, b_i]\right) = \sum_{i=1}^{\infty} Pr((a_i, b_i]) = \sum_{i=1}^{\infty} (F(b_i) - F(a_i))$$

3.3 Probabilità condizionale

Sia $(\Omega, \mathcal{A}, Pr)$ uno spazio probabilizzato. Fissato un elemento h di \mathcal{A} con $Pr(H) \neq 0$, si chiama funzione di probabilità dedotta da Pr sotto la condizione H la funzione di probabilità Pr_H sullo spazio (Ω, \mathcal{A}) Probabilizzabile

$$Pr_H(A) = \frac{Pr(A \cap H)}{Pr(H)}$$

Per ogni evento $A \in \mathcal{A}$.

La probabilità $Pr_H(A)$ si chiama **Probabilità Condizionale** di A , secondo Pr , sotto la condizione H e si denota

$$Pr(A|H)$$

3.4 Classe Completa di eventi

Dato uno spazio probabilizzabile (Ω, \mathcal{A}) la famiglia di eventi $\{A_i\}_{i=1}^{\infty}$ è detta Classe Completa se

- $\bigcup_{n=1}^{\infty} A_n = \Omega$
- $A_i \cap A_j = \emptyset, i \neq j$

3.5 Teorema delle Probabilità Totali

Sia $\{A_i\}_{i=1}^{\infty}$ una famiglia di eventi che costituisce una Classe Completa di Ω tale che

$$Pr(A_i) > 0, i = 1, 2, \dots$$

Sia B un qualunque evento. allora

$$Pr(B) = \sum_{i=1}^{\infty} Pr(A_i \cap B) = \sum_{i=1}^{\infty} Pr(A_i) Pr(B|A_i)$$

3.6 Teorema di Bayes

Sia $\{A_i\}_{i=1}^{\infty}$ una Classe Completa di eventi tale che:

$$Pr(A_i) > 0, i = 1, 2, \dots$$

e B un qualunque evento con $Pr(B) > 0$. allora

$$Pr(A_i|B) = \frac{Pr(A_i) Pr(B|A_i)}{\sum_{j=1}^{\infty} Pr(A_j) Pr(B|A_j)} \quad j = 1, 2, \dots$$

3.7 Indipendenza stocastica

In uno spazio probabilizzato (Ω, \mathcal{A}, P) due eventi A, B si dicono tra loro stocasticamente indipendenti se e solo se

$$Pr(A \cap B) = Pr(A) \cdot Pr(B)$$

In particolare si noti che dati due eventi stocasticamente indipendenti A, B allora:

$$Pr(A|B) = \frac{Pr(A \cap B)}{Pr(B)} = Pr(A)$$

e lo stesso vale per $Pr(B|A) = Pr(B)$

La nozione di indipendenza può essere estesa a più di due eventi. Vedi NOTE-B P.61

3.8 Tribù indipendenti

Dato uno spazio probabilizzato $(\Omega, \mathcal{A}, Pr)$. Due Tribù contenute in \mathcal{A} si dicono tra loro indipendenti se ogni elemento dell'uno è indipendente da ogni elemento dell'altra.

Capitolo 4

Variabili Aleatorie

Sia dato lo spazio probabilizzabile (Ω, \mathcal{A}) . Si dice **Variabile aleatoria** (v.a.) ogni funzione a valori reali definita in Ω , $y = X(\omega)$, tale che

$$\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{A}$$

per ogni valore reale x .

- Giova osservare che nella definizione la probabilità non gioca alcun ruolo e che quando \mathcal{A} è la classe di tutti i sottoinsiemi di Ω la condizione nella definizione è sempre soddisfatta.
- Per rendersi conto della necessità di imporre alla funzione $X(\omega)$ la condizione riportata sopra, basterà dire che, intendendo assegnare una probabilità agli insiemi $\{\omega \in \Omega : X(\omega) \leq x\}$ per ogni reale x ed avendo probabilizzato la classe \mathcal{A} , occorre che tali insiemi appartengano ad \mathcal{A} .

4.1 Variabili aleatorie e Tribù

Siano $\tilde{\Omega}$ e Ω due insiemi arbitrari e sia $X : \tilde{\Omega} \rightarrow \Omega$ una funzione. Se \mathcal{A} è una Tribù su Ω allora:

$$\tilde{\mathcal{A}} = \{X^{-1}(A) : A \in \mathcal{A}\}$$

è una Tribù su $\tilde{\Omega}$.

4.1.1 Teorema 10

Siano $\tilde{\Omega}$ e Ω due insiemi arbitrari e sia $X : \tilde{\Omega} \rightarrow \Omega$ una funzione. Se \mathcal{A} è una Tribù su Ω allora:

$$\tilde{\mathcal{A}} = \{A \subseteq \tilde{\Omega} : X^{-1}(A) \in \mathcal{A}\}$$

4.1.2 Teorema 12

Ogni funzione continua oppure monotona crescente o decrescente $f : (\mathbb{R}, \mathcal{B}(\mathbb{R})) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ è una variabile aleatoria.

4.2 Variabili aleatorie e funzioni di probabilità

Il valore che assume la funzione $y = X(\omega) : \Omega \rightarrow \mathbb{R}$ in corrispondenza di un esperimento è aleatorio in quanto dipende dal particolare risultato conseguito nell'esperimento $\omega \in \Omega$; Ci si potrà chiedere con quale probabilità la funzione $X(\omega)$ assuma valore nell'intervallo $(a, b]$ cioè, dare un significato alla scrittura

$$\text{Probabilità di } (a < X \leq b) = Pr(X \in (a, b]) = Pr(\{\omega \in \Omega : a < X \leq b\})$$

Si osservi a tale scopo che l'intervallo $(a, b]$ e l'insieme A

$$A = \{\omega \in \Omega : a < X(\omega) \leq b\} \in \mathcal{A}$$

sono in un certo senso equivalenti giacchè quando si verifica A , cioè $\omega \in \mathcal{A}$, allora $X \in (a, b]$ e viceversa. Dato che all'evento A è assegnata $Pr(A)$, si potrà porre, per ogni $a < b$,

$$Pr_X((a, b]) = Pr(X \in (a, b]) = Pr(\{\omega \in \Omega : a < X \leq b\})$$

La funzione di probabilità P_X , definita sulla classe di Borel di \mathbb{R} , è nota col nome di distribuzione della v.a. X e mediante essa sarà possibile determinare
 $Pr_X((a, b]) = Pr(X \in (a, b]) = Pr(\{\omega \in \Omega : a < X \leq b\})$

4.3 Variabili aleatorie discrete

Una v.a. X definita su (Ω, \mathcal{A}) è detta discreta se i valori distinti dell'insieme $\bigcup_{\omega \in \Omega} \{X(\omega)\}$ costituiscono un insieme R_X finito o numerabile.

4.3.1 Funzione di probabilità (o densità discreta)

Se X è una v.a. discreta con $R_X = \{x_1, x_2, \dots\}$, allora la funzione, definita in \mathbb{R} , data da

$$p(x) = \begin{cases} Pr(X = x_i) > 0 & x = x_i \in R_X \\ 0 & x \notin R_X \end{cases}$$

è detta funzione di probabilità (o densità discreta) della v.a. X , R_X viene detto supporto della v.a. X .

4.3.2 Teorema

Se X è una v.a. discreta con $R_X = \{x_1, x_2, \dots\}$ allora

$$p(x) \geq 0$$

per ogni x reale e

$$\sum_{x \in R_X} p(x) = 1$$

4.3.3 Distribuzione Binomiale

Si dice che una v.a. X si distribuisce secondo la distribuzione di probabilità (o legge) binomiale di parametri $N \geq 1$ (intero) e $0 \leq p \leq 1$, se

$$Pr(X = x) = \begin{cases} \binom{N}{x} p^x (1-p)^{N-x} & x = 0, 1, \dots, N \\ 0 & \text{altrimenti} \end{cases}$$

E scriveremo $X \sim Bi(N, p)$, dove n è il numero di prove effettuate, e p è la probabilità di successo della singola prova.

In altre parole

La distribuzione binomiale descrive la probabilità di avere esattamente x successi, provando N volte, con p probabilità di vittoria di un singolo evento.

Proprietà

- Media: $\mathbb{E}(X) = Np$
- Varianza: $\text{Var}(X) = Np(1-p)$

4.3.4 Funzione di ripartizione

Sia X una v.a.. Si dice funzione di ripartizione della v.a. X la funzione $y = F(x)$, definita per ogni x reale, data da

$$F(x) = Pr(X \leq x) \quad x \in \mathbb{R}$$

Funzione di ripartizione e funzione di probabilità

Per una v.a. discreta, si osservi, a conferma delle proprietà generali della funzione di ripartizione, come i punti di discontinuità di $F(x)$ coincidano con i punti di R_X della v.a. e che l'ampiezza del salto in detti punti corrisponde alla funzione di probabilità, cioè

$$p(X = x) = F(x) - F(X^-)$$

4.3.5 Distribuzione Geometrica

La distribuzione Geometrica nasce con riferimento allo stesso schema che ha condotto alla distribuzione Binomiale ma ora, anziché contare il numero di successi in N prove indipendenti, interessa il numero delle prove necessarie per ottenere il primo successo.

Si dice che una v.a. X si distribuisce secondo una distribuzione geometrica di parametro $0 \leq p \leq 1$ se la sua funzione di probabilità è

$$Pr(X = x) = \begin{cases} p(1-p)^{x-1} & x = 1, 2, 3, \dots \\ 0 & \text{altrove} \end{cases}$$

e scriveremo $X \sim Ge(p)$.

Proprietà

- Funzione di ripartizione: $F(x) = 1 - (1-p)^x$
- Momento secondo: $\mathbb{E}(X^2) = \frac{2-p}{p^2}$
- Varianza: $\mathbb{V}ar(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \frac{1-p}{p^2}$

4.3.6 Distribuzione Binomiale negativa (o di Pascal)

Si dice che una v.a. X si distribuisce secondo la distribuzione binomiale negativa di parametri $0 < p \leq 1$ e $r \geq 1$ (intero) se la sua funzione di probabilità è data da

$$Pr(X = x) = \begin{cases} \binom{x-1}{r-1} p^r (1-p)^{x-r} & x = r, r+1, r+2, \dots \\ 0 & \text{altrove} \end{cases}$$

e indichiamo con $X \sim BiNe(r, p)$.

In altre parole

La distribuzione di Pascal dà la probabilità che siano necessari esattamente x fallimenti per avere r successi. p è la probabilità di un singolo successo.

Relazione tra Binomiale e Binomiale negativa (Teorema)

Sia $X \sim BiNe(r, p)$ e $Z \sim Bi(N, p)$ allora

$$Pr(Z \geq r) = Pr(X \leq N)$$

4.3.7 Distribuzione di Poisson

La distribuzione di Poisson (o poissoniana) è una distribuzione di probabilità discreta, che esprime le probabilità per il numero di eventi che si verificano successivamente e indipendentemente in un dato intervallo di tempo, sapendo che mediamente se ne verifica un numero λ .

Si dice che una v.a. X si distribuisce secondo la distribuzione di Poisson di parametri $\lambda \geq 0$ se la sua funzione di probabilità è data da

$$P(X = n) = \frac{\lambda^n}{n!} e^{-\lambda}$$

Proprietà

- $\mathbb{E}(X) = \lambda$
- $(V)ar(X) = \lambda$

Se Y_1 e Y_2 sono due variabili aleatorie indipendenti con distribuzioni di Poisson di parametri λ_1 e λ_2 rispettivamente, allora:

- la loro somma $Y = Y_1 + Y_2$ segue ancora una distribuzione di Poisson, di parametro $\lambda = \lambda_1 + \lambda_2$;
- la distribuzione di Y_1 condizionata da $Y = n$ è la distribuzione binomiale di parametri $\frac{\lambda_1}{\lambda}$ e n .

4.4 Variabili aleatorie continue

Una v.a. X definita su (Ω, \mathcal{A}) è detta continua se la sua funzione di ripartizione è continua.

4.4.1 Densità

Si dice che la v.a. X è dotata di densità se la probabilità con cui X assume valori nell'intervallo $(a, b]$ è data mediante la formula

$$Pr(X \in (a, b]) = Pr(a < X \leq b) = \int_a^b f(x)dx$$

in cui $f(x)$ prende il nome di funzione di densità di probabilità della v.a. X e deve avere le seguenti caratteristiche

- $f(x) > 0$ per ogni $x \in \mathbb{R}$
- $\int_{-\infty}^{+\infty} f(x)dx = 1$

4.4.2 Variabili aleatorie assolutamente continue

Una v.a. X definita su (Ω, \mathcal{A}) è detta assolutamente continua se la sua funzione di ripartizione è continua e la sua v.a. X ammette densità.

4.4.3 Densità e funzione di ripartizione

Per una v.a. X assolutamente continua con densità $f(x)$ e con funzione di ripartizione $F(x)$ abbiamo:

$$Pr(X \in (a, b]) = \int_a^b f(x)dx = F(b) - F(a)$$

4.4.4 Distribuzione Normale (o di Gauss)

Si dice che una v.a. X si distribuisce con legge di probabilità Normale (o Gaussiana) di parametri $-\infty < \mu < +\infty$ e $0 < \sigma < +\infty$ se possiede la seguente densità.

$$f(x, \mu, \sigma) = \frac{1}{\sqrt{(2\pi\sigma^2)}} e^{(-\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2})}$$

e la indichiamo con $X \sim N(\mu, \sigma^2)$. La v.a. $X \sim N(0, 1)$ è chiamata Normale Standard.

Proprietà

- Valore atteso: $\mathbb{E}(X) = \mu$
- Varianza: $\mathbb{V}ar(X) = \sigma^2$

4.4.5 Standardizzazione di una Normale

Data una $X \sim N(\mu, \sigma^2)$, Allora

$$Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$$

Questa operazione viene chiamata Standardizzazione

4.4.6 Distribuzione Esponenziale

Si dice che una v.a. X ha legge Esponenziale con parametro $\lambda > 0$ se la sua funzione di densità

$$f(x; \lambda) = \begin{cases} \lambda e^{(-\lambda x)} & x > 0 \\ 0 & \text{altrove} \end{cases}$$

e la indichiamo nel seguente modo $X \sim \text{Exp}(\lambda)$. La distribuzione Esponenziale è senza memoria.

Proprietà

- Media: $\mathbb{E}(X) = \frac{1}{\lambda}$
- Varianza: $\text{Var}(X) = \frac{1}{\lambda^2}$
- Funzione di ripartizione: $F(x) = 1 - e^{-\lambda x}$
- Il minimo $Y = \min\{X_1, \dots, X_n\}$ tra n variabili aleatorie indipendenti con distribuzioni esponenziali di parametri $\lambda_1, \dots, \lambda_n$ è ancora una variabile aleatoria con distribuzione esponenziale, di parametro $\lambda = \lambda_1 + \dots + \lambda_n$.

4.4.7 Trasformazione di variabili aleatorie p.104 (manca)

4.5 Speranza matematica o valore atteso per v.a. discrete

Sia X una v.a. discreta con funzione di probabilità $p_X(x)$. Allora, si chiama speranza matematica di X la quantità (finita)

$$\mathbb{E}(X) = \sum_{x \in R_X} x p_X(x)$$

Sia X una v.a. dotata di densità $f_X(x)$ e funzione di ripartizione $F_X(x)$. Si chiama speranza matematica di X la quantità (finita).

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} x f_X(x) dx$$

4.6 Momenti

Data la v.a. X si dice momento non centrato di ordine r (intero positivo) il valore

$$\mu_r = \mathbb{E}(X^r)$$

e si dice momento centrato dalla media di ordine r

$$\bar{\mu}_r = \mathbb{E}((x - \mu_1)^r)$$

Valori di sintesi basati sui momenti

- Media: $\mu = \mu_1 = \mathbb{E}(X)$
- Varianza: $\text{Var}(X) = \sigma^2 = \bar{\mu}_2 = \mathbb{E}((x - \mu_1)^2) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$
- Deviazione standard: $\sigma = \sqrt{\sigma^2}$

Capitolo 5

Variabili Aleatorie Doppie

Sia $(\Omega, \mathcal{A}, Pr)$ uno spazio probabilizzato. Siano $X(\omega)$ e $Y(\omega)$ due v.a. definite su Ω in modo che:

$$Z(\omega) = (X(\omega), Y(\omega)) : \Omega \rightarrow \mathbb{R}^2$$

$Z(\omega)$ è detta v.a. doppia e $R_Z = R_{X,Y} = \{(x, y) : x \in R_X, y \in R_Y\}$

Resta da definire la funzione di probabilità di $Z(\omega)$. Le funzioni di ripartizione $F_X(x)$ e $F_Y(y)$ di X e Y rispettivamente, in genere non sono sufficienti per determinare tale proprietà.

E' necessario considerare la seguente funzione di ripartizione (detta congiunta)

$$F_Z(z) = F_{X,Y}(x, y) = Pr(\{X \leq x\} \cap \{Y \leq y\}) \quad (x, y) \in R_{X,Y}$$

5.1 Funzione di probabilità congiunta (discreta)

Per due v.a. discrete X e Y , la v.a. doppia $Z = (X, Y)$ (che è discreta) ha funzione di probabilità (congiunta)

$$P_Z(z) = \begin{cases} p_{X,Y}(x, y) = Pr(\omega : \{X(\omega) = x\} \cap \{Y(\omega) = y\}) & (x, y) \in R_{X,Y} \\ 0 & \text{altrove} \end{cases}$$

5.2 Variabili aleatorie doppie dotate di densità

La v.a. doppia $Z = (X, Y)$ si dirà dotata di densità se esiste una funzione $f_{X,Y}(x, y)$ tale che

- $f_{X,Y}(x, y) \geq 0, \quad \forall (x, y) \in \mathbb{R}^2$

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_{X,Y}(x, y) dx dy$$

-

$$Pr(a < x \leq b, c < y \leq d) = \int_{-\infty}^x \int_{-\infty}^y f_{X,Y}(u, v) du dv$$

tale funzione è chiamata densità congiunta $f_Z(z) = f_{X,Y}(x, y)$.

5.2.1 Densità marginali

Dalle formule di prima abbiamo che

$$F_{X,Y}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{X,Y}(u, v) du dv$$

E quindi

- Densità marginale della v.a. X

$$f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x,v) dv$$

- Densità marginale della v.a. Y

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(u,y) du$$

5.3 Distribuzioni condizionali per v.a. (Probabilità condizionale di $X|Y = y$)

Sia (X, Y) una v.a. doppia discreta con funzione di probabilità

$$p_{X,Y}(x,y) = Pr(X = x, Y = y)$$

allora in accordo con la definizione di probabilità condizionale

$$p_{X|Y}(X = x|Y = y) = Pr(\{X = x\}|\{Y = y\}) = \frac{p_{X,Y}(x,y)}{p_Y(y)} \quad y \in R_Y (p_Y(y) > 0)$$

Per ogni valore fissato di $y \in R_Y$ la funzione $p_{X|Y}(X = x|Y = y)$ prende il nome di probabilità condizionale di $X|Y = y$

5.4 Distribuzioni condizionali e indipendenza per v.a. (p137 dispense B)

5.5 Funzioni di ripartizioni condizionali

Dalla funzione di probabilità condizionale (nel caso discreto) e dalla densità condizionale (nel caso assolutamente continuo), possiamo costruire le funzioni di ripartizione condizionale

$$F_{X|Y}(x|y) = \sum_{u \leq x : u \in R_X} p_{X|Y}(u|y)$$

e

$$F_{X|Y}(x|y) = \int_{-\infty}^x f_{X|Y}(u|y) du$$

5.6 Variabili aleatorie condizionali e speranza matematica

Data la v.a. doppia (X, Y) allora la funzione $X|Y = y$ ($y \in R_Y$) è una v.a. con funzione di probabilità $P_{X|Y}(x|y)$. Quindi alla definizione di speranza matematica e di varianza abbiamo

$$\mathbb{E}(X|Y = y) = \sum_{x \in R_X} x p_{X|Y}(x|y)$$

$$var(X|Y = y) = \sum_{x \in R_X} (x - \mathbb{E}(X|Y = y))^2 p_{X|Y}(x|y)$$

e in maniera del tutto analoga nel caso di v.a. dotate di densità.

5.7 Speranza matematica della speranza matematica condizionale

Ad esempio per v.a. doppie discrete (il risultato vale nel caso generale??)

$$\mathbb{E}(\mathbb{E}(X|Y)) = \mathbb{E}(X)$$

5.8 Varianza e Varianza condizionale (Scomposizione della varianza)

Sia (X, Y) una v.a. doppia, allora

$$\text{Var}(X) = \mathbb{E}(\text{Var}(X|Y)) + \text{Var}(\mathbb{E}(X|Y))$$

5.9 Dipendenza in media

La v.a. X si dice indipendente in media da Y se

$$\mathbb{E}(X|Y = y) = \mathbb{E}(X) \quad \forall y \in R_Y$$

Si noti che se X è indipendente stocasticamente da Y allora è anche indipendente in media. Viceversa non è vero, in generale.

5.10 Rapporto di correlazione

Sia (X, Y) una v.a. doppia discreta, si chiama rapporto di correlazione di X dato Y

$$\eta_{X|Y}^2 = \frac{\text{Var}(\mathbb{E}(X|Y))}{\text{Var}(X)}$$

E in modo analogo si definisce $\eta_{Y|X}^2$. Dalla formula della scomposizione della varianza è facile vedere che:

$$0 \leq \eta_{X|Y}^2 \leq 1$$

inoltre

- se $\eta_{X|Y}^2 = 0$ allora X è indipendente in media da Y
- se $\eta_{X|Y}^2 > 0$ allora X è dipendente in media da Y
- $\eta_{X|Y}^2 = 1$ se e solo se $\Pr(X = \mathbb{E}(X|Y)) = 1$

5.11 Covarianza e correlazione

La covarianza e la correlazione sono altri due indici di dipendenza (lineare) tra due v.a.

$$\text{cov}(X, Y) = \mathbb{E}(X * Y) - \mathbb{E}(X) * \mathbb{E}(Y)$$

mentre

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\text{Var}(X) * \text{Var}(Y)}}$$

5.12 Varianza di una combinazione lineare di v.a

Sia (X, Y) una v.a. doppia e a e b due costanti. Allora

$$\text{Var}(aX + bY) = a^2 \text{Var}(X) + b^2 \text{Var}(Y) + 2ab \text{cov}(X, Y)$$

Capitolo 6

Teoremi limite della probabilità

6.1 Convergenza in probabilità (o debole)

Ci sono diversi modi per esprimere il fatto che S_n/n si avvicina a p . Potremmo ad esempio scrivere che, per n grande e per ϵ piccolo a piacere

$$Pr\{|S_n/n - p| \geq \epsilon\} \approx 0$$

o equivalentemente

$$\lim_{n \rightarrow +\infty} Pr\{|S_n/n - p| \geq \epsilon\} = 0$$

in simboli questo tipo di convergenza si denota con

$$p$$

$$S_n/n \longrightarrow \mu$$

e si legge **converge in probabilità (o in senso debole)** ad una v.c. Y se, per ogni $\epsilon > 0$,

$$\lim_{n \rightarrow \infty} Pr(|Y_n - Y| \geq \epsilon) = 0,$$

ovvero

$$\lim_{n \rightarrow \infty} Pr(|Y_n - Y| \leq \epsilon) = 1,$$

6.2 Convergenza in media quadratica

Un'altra formalizzazione del concetto di "vicinanza" potrebbe richiedere che in media gli scostamenti (al quadrato) di S_n/n da p siano piccoli, quando n è grande:

$$\mathbb{E}[(S_n/n - p)^2] \approx 0,$$

o equivalentemente

$$\lim_{n \rightarrow \infty} \mathbb{E}[(S_n/n - p)^2] = 0.$$

In simboli questo tipo di convergenza si denota con

$$m.q.$$

$$S_n/n \longrightarrow p$$

e si legge **"converge in media quadratica"**.

Più in generale diremo che una successione Y_1, Y_2, \dots **converge in media quadratica** ad una v.c. Y se

$$\lim_{n \rightarrow \infty} \mathbb{E}[(Y_n - Y)^2] = 0.$$

Proposizione

La convergenza in media quadratica implica la convergenza in Probabilità:

$$\begin{aligned} m.q. \quad & P \\ Y_n \rightarrow Y & \Rightarrow Y_n \rightarrow Y \end{aligned}$$

6.3 Disuguaglianza di Markov

Sia Y una v.c. che assume valori non negativi allora per ogni numero reale $a > 0$

$$Pr(Y \geq a) \leq \frac{\mathbb{E}(Y)}{a}$$

6.4 Disuguaglianza di Chebychev

Sia Y una v.c. con valore atteso $\mathbb{E}(Y) = \mu$ e varianza $\mathbb{V}ar(Y) = \sigma^2$. Allora

$$Pr(|Y - \mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2}$$

6.5 Somme di variabili casuali

Proposizione

Siano Y_1, \dots, Y_n v.c. con valore atteso rispettivamente μ_1, \dots, μ_n . allora

$$\mathbb{E}(Y_1 + \dots + Y_n) = \mu_1 + \dots + \mu_n$$

Proposizione

Siano Y_1, \dots, Y_n v.c. indipendenti con varianza $\sigma_1^2, \dots, \sigma_n^2$ rispettivamente. Allora

$$\mathbb{V}ar(Y_1 + \dots + Y_n) = \sigma_1^2 + \dots + \sigma_n^2$$

Proposizione

Siano Y_1, \dots, Y_n v.c. indipendenti, tutte con valore atteso μ e varianza σ^2 e sia $\bar{Y}_n = \sum_{i=1}^n Y_i/n$. Allora

$$\begin{aligned} \mathbb{E}(\bar{Y}_n) &= \sum_{i=1}^n \frac{\mathbb{E}(Y_i)}{n} = n \frac{\mu}{n} = \mu, \\ \mathbb{V}ar(\bar{Y}_n) &= \sum_{i=1}^n \frac{\mathbb{V}ar(Y_i)}{n^2} = n \frac{\sigma^2}{n^2} = \frac{\sigma^2}{n}. \end{aligned}$$

6.6 Legge debole dei grandi numeri

Sia Y_1, Y_2, \dots una successione di v.c. indipendenti, ciascuna con valore atteso μ e varianza σ^2 . Allora, per ogni $\epsilon > 0$,

$$\lim_{n \rightarrow \infty} Pr\{|\bar{Y}_n - \mu| \geq \epsilon\} = 0$$

ovvero $\bar{Y}_n \rightarrow \mu$
p.151