DIMOSTRAZIONE.

$$\begin{aligned}
E(|X^h|) &= \sum_{s\in S}|X(s)^h|P(s)\\
&= \sum_{s\in S:|X(s)^h|\leq 1}|X(s)^h|P(s)+\sum_{s\in S:|X(s)^h|>1}|X(s)^h|P(s)\\
&\leq P(|X^h|\leq 1)+\sum_{s\in S:|X(s)^h|>1}|X(s)^k|P(s)\\
&\leq P(|X^h|\leq 1)+\sum_{s\in S}|X(s)^k|P(s)<\infty
\end{aligned}$$

□

LEMMA 14. *Se X ed Y sono variabili aleatorie discrete tali che $E(X^2)$ ed $E(Y^2)$ esistono, allora esistono $E(X)$, $E(Y)$ ed $E(XY)$.*

DIMOSTRAZIONE. L'esistenza di $E(X)$ ed $E(Y)$ segue dal lemma precedente. Da $0 \leq (a-b)^2 = a^2+b^2-2ab$ si ha che per qualsiasi coppia di numeri reali a e b vale che $ab \leq (a^2+b^2)/2$, per cui

$$\begin{aligned}
E(|XY|) &= \sum_{s\in S}|X(s)||Y(s)|P(s)\\
&= \frac{1}{2}\sum_{s\in S}(|X(s)|^2+|Y(s)|^2)P(s)\\
&\leq \frac{1}{2}(\sum_{s\in S}|X(s)|^2P(s)+\sum_{s\in S}|Y(s)|^2P(s))<\infty
\end{aligned}$$

□

Ora abbiamo condizioni sufficienti per generalizzare i risultati sui valori attesi di variabili indipendenti.

ESERCIZIO 20. *Verificare che se X_i sono variabili aleatorie discrete tali che $E(X_i^2)$ esiste per ogni i allora vale il Teorema 8.*

Con il secondo momento è quindi possibile definire la varianza e la deviazione standard di X e vale anche in questo caso l'additività delle varianze per variabili indipendenti.

ESERCIZIO 21. *Verificare che se X è una variabile aleatoria discreta tale che $E(X^2)$ esiste allora la varianza e la deviazione standard di X esistono e vale il Teorema 9.*

5.3. Vettori aleatori e variabili aleatorie congiunte

In molti problemi, come quando abbiamo parlato di indipendenza di variabili aleatorie o linearità del valore atteso, si considerano più variabili aleatorie contemporaneamente. In questo caso conviene pensarle come un vettore aleatorio.

DEFINIZIONE 19. *Dato uno spazio di probabilità (discreta)* (S, P), *un* **vettore aleatorio** $\boldsymbol{X} = (X_1, \dots, X_n)$ *n-dimensionale è una funzione* $\boldsymbol{X} \colon S \to \mathbb{R}^n$ *tale che* $s \to \boldsymbol{X}(s) = ((X_1(s), \dots, X_n(s))$.

Si possono ripetere ora molte definizioni e proprietà delle variabili aleatorie, che si ottengono semplicemente sostituendo $\boldsymbol{X}$ ad X.

DEFINIZIONE 20. *La distribuzione di un vettore aleatorio* $\boldsymbol{X} = (X_1, \dots, X_n)$, *detta anche* **distribuzione congiunta** *delle* X_i, *è la coppia* $(S_{\boldsymbol{X}}, P_{\boldsymbol{X}})$ *con*

$$S_{\boldsymbol{X}} = \{\boldsymbol{x} \in \mathbb{R}^n | \text{ esiste } s \in S : \boldsymbol{X}(s) = \boldsymbol{x}\}$$

e

$$P_{\boldsymbol{X}}(\boldsymbol{x}) = P\{s \in S : \boldsymbol{X}(s) = \boldsymbol{x}\}$$

Si noti che per l'usuale proprietà delle funzioni, per ogni $B \subseteq S_{\boldsymbol{X}}$

$$P_{\boldsymbol{X}}(B) = \sum_{\boldsymbol{x} \in B} P_{\boldsymbol{X}}(\boldsymbol{x}).$$

ESERCIZIO 22. *Se* $\boldsymbol{X}$ *è un vettore aleatorio discreto, verificare che* $(S_{\boldsymbol{X}}, P_{\boldsymbol{X}})$ *è uno spazio di probabilità discreto.*

In un vettore aleatorio $\boldsymbol{X} = (X_1, \dots, X_n)$ ciascuna delle componenti X_i è una variabile aleatoria, con una sua distribuzione (S_{X_i}, P_{X_i}), ed è interessante ed utile in vari problemi studiare la relazione tra queste e la distribuzione congiunta $(S_{\boldsymbol{X}}, P_{\boldsymbol{X}})$. Le (S_{X_i}, P_{X_i}) sono dette **distribuzioni marginali**.

Per semplificare la notazione assumeremo che $S_{\boldsymbol{X}} = \mathbb{R}^n$ ponendo $P_{\boldsymbol{X}}(\boldsymbol{x}) = 0$ se $\boldsymbol{x}$ non era originalmente appartenente al codominio di $\boldsymbol{X}$.

LEMMA 15. *Per ogni un vettore aleatorio* $\boldsymbol{X} = (X_1, \dots, X_n)$ *vale*

$$P_{X_1}(x_1) = \sum_{(x_2, \dots, x_n) \in \mathbb{R}^{n-1}} P_{\boldsymbol{X}}(x_1, \dots, x_n).$$

DIMOSTRAZIONE. Per ogni $x_1 \in \mathbb{R}$, gli eventi

$$A_{(x_2, \dots, x_n)} = \{s \in S : \boldsymbol{x}(s) = (x_1, \dots, x_n)\} \subseteq S$$

sono disgiunti, per cui

$$\begin{aligned} \sum_{(x_2, \dots, x_n) \in \mathbb{R}^{n-1}} P_{\boldsymbol{X}}(x_1, \dots, x_n) &= \sum_{(x_2, \dots, x_n) \in \mathbb{R}^{n-1}} P(A_{(x_2, \dots, x_n)}) \\ &= P(\cup_{(x_2, \dots, x_n) \in \mathbb{R}^{n-1}} A_{(x_2, \dots, x_n)}) \\ &= P(s : X_1(s) = x_1) = P_{X_1}(x_1) \end{aligned}$$

□

Naturalmente il Lemma precedente si applica ad ogni proiezione o marginale X_i, per cui dalle congiunte si determinano le marginali.

ESERCIZIO 23. *Dimostrare tramite un esempio che l'opposto non è vero e vi sono distribuzioni congiunte diverse che danno luogo alle stesse marginali.*

L'indipendenza delle variabili aleatorie si può esprimere in termini del rapporto tra le distribuzioni congiunta e marginali, nel senso che le componenti di un vettore aleatorio $\boldsymbol{X}= (X_1, \dots, X_n)$ sono indipendenti se e solo se

$$P_{\mathbf{X}}(\boldsymbol{x}) = \prod_{i=1}^{n} P_{X_i}(x_i)$$

per tutti gli $\boldsymbol{x}= (x_1, \dots, x_n) \in \mathbb{R}^n$. Quindi nel caso indipendente le marginali permettono la ricostruzione della distribuzione congiunta.

Se di una o più variabili conosciamo il valore assunto abbiamo delle distribuzioni condizionali.

DEFINIZIONE 21. *Dato un vettore aleatorio* $\boldsymbol{X}= (X_1, \dots, X_n)$*, per* $k = 1, \dots, n$*, se* $(x_1, \dots, x_k)$ *è tale che* $P(X_1 = x_1, \dots, X_k = x_k) \neq 0$*, si dice* **distribuzione condizionata o condizionale** *di* $X_{k+1}, \dots, X_n$ *dato che* $(X_1, \dots, X_k) = (x_1, \dots, x_k)$ *lo spazio di probabilità costituito da*

$$\begin{aligned} &S_{X_{k+1},\dots,X_n|(X_1,\dots,X_k)=(x_1,\dots,x_k)} \\ &\qquad = \{(x_{k+1}, \dots, x_n) : (x_1, \dots, x_k, x_{k+1}, \dots, x_n) \in S_{\mathbf{X}}\} \end{aligned}$$

e

$$\begin{aligned} &P_{X_{k+1},\dots,X_n|X_1,\dots,X_k}((x_{k+1}, \dots, x_n)|(x_1, \dots, x_k)) \\ &\qquad = \frac{P_{\mathbf{X}}(\boldsymbol{x})}{P(X_1 = x_1, \dots, X_k = x_k)} \end{aligned}$$

Naturalmente si può prendere $S_{X_{k+1},\dots,X_n|(X_1,\dots,X_k)=(x_1,\dots,x_k)} = \mathbb{R}^{n-k}$.

ESERCIZIO 24. *Verificare che la coppia definita nella definizione precedente è uno spazio di probabilità .*

Naturalmente la medesima definizione si poteva dare permutando gli indici. Dal teorema delle probabilità totali si vede che

$$\begin{aligned} &P(X_1 = x_1, \dots, X_n = x_n) \\ &\quad = \sum_{(x_1,\dots,x_k)\in S_{X_1,dots,X_k}} P_{X_{k+1},\dots,X_n|X_1,\dots,X_k}((x_{k+1}, \dots, x_n)|(x_1, \dots, x_k)) \\ &\qquad\qquad \cdot P(X_1 = x_1, \dots, X_k = x_k). \end{aligned}$$

Per $k = 1$ questa osservazione indica come ricostruire la distribuzione congiunta dalla conoscenza delle distribuzioni condizionate e della relativa marginale.

DEFINIZIONE 22. *Il valore atteso calcolato rispetto alla probabilità condizionale di una variabile aleatoria date le altre si chiama* **valore atteso condizionale** *e si denota*

$$E(X_n|(X_1,\ldots,X_{n-1})=(x_1,\ldots,x_{n-1}).$$

Vogliamo introdurre ora una misura della dipendenza di variabili aleatorie, limitandoci al caso di due variabili per semplicità . Abbiamo visto che nel caso di variabili aleatorie indipendenti il valore atteso del prodotto si fattorizza, per cui viene naturale di studiare la quantità seguente:

DEFINIZIONE 23. *Dato un vettore aleatorio* (X,Y)*, ossia due variabili aleatorie e la loro distribuzione congiunta, tali che* $E(X^2),E(Y^2)<\infty$*, si dice* **covarianza** *di* X *ed* Y *il valore*

$$Cov(X,Y)=E(XY)-E(X)E(Y).$$

Si noti che per il Lemma 14 la definizione è ben posta e che vale $Cov(X,X)=Var(X)$.

LEMMA 16.

$$Cov(X,Y)=E((X-E(X)(Y-E(Y)).$$

DIMOSTRAZIONE. Se $Cov(X,Y)$ esiste allora $E(X^2),E(Y^2)<\infty$ ed anche $E(|XY|),E(|X|),E(|Y|)<\infty$, il che implica che

$$\begin{aligned} E(|X-E(X)||Y-E(Y)|) &\leq E(|XY|+|X||E(Y)|+|Y||E(X)| \\ &\qquad +|E(X)||E(Y)|) \\ &\leq E(|XY|)+3E(|X|)E(|Y|)<\infty \end{aligned}$$

per cui $E((X-E(X))(Y-E(Y))$ esiste. Ora

$$\begin{aligned} E((X-E(X))(Y-E(Y)) &\leq E(XY)-XE(Y)-YE(X) \\ &\qquad +E(X)E(Y) \\ &\leq E(XY)+E(X)E(Y)<\infty \end{aligned}$$

□

Dimensionalmente $Cov(X,Y)$ è il prodotto delle dimensioni di X ed Y. Per ottenere un numero puro dividiamo per il prodotto delle deviazioni standard, o equivalentemente, consideriamo la correlazione tra le variabili standardizzate.

DEFINIZIONE 24. *Dato un vettore aleatorio* (X,Y) *tale che* $E(X^2),E(Y^2)<\infty$*, si dice* **correlazione** *di* X *ed* Y *il valore*

$$r=r(X,Y)=\frac{Cov(X,Y)}{SD(X)SD(Y)}.$$

Si ricorda che la variabili standardizzata è definita, per una variabile aleatoria discreta X con secondo momento finito, da $\bar{X}=\frac{X-E(X)}{SD(X)}$ e che valgono $E(\bar{X})=0$ e $Var(\bar{X})=E(\bar{X})=1=SD(\bar{X})$

LEMMA 17. $r = E(\bar{X}\bar{Y}) = Cov(\bar{X}\bar{Y})$

DIMOSTRAZIONE. Dalle proprietà di $\bar{X}$ e $\bar{Y}$ si ha

$$\begin{aligned} Cov(\bar{X}\bar{Y}) &= E(\bar{X}\bar{Y}) \\ &= E((\frac{X - E(X)}{SD(X)})(\frac{Y - E(Y)}{SD(Y)})) \\ &= \frac{Cov(X, Y)}{SD(X)SD(Y)} \end{aligned}$$

□

Oltre ad essere un numero puro la correlazione è limitata:

TEOREMA 13. *Per ogni vettore aleatorio* (X, Y) *con secondo momento finito si ha*

$$-1 \leq r(X, Y) \leq 1;$$

inoltre $r = 1$ *se e solo se* $Y = cX + d$ *con* $c > 0$ *ed* $r = -1$ *se e solo se* $Y = -cX + d$ *con* $c > 0$.

DIMOSTRAZIONE. La dimostrazione che segue può essere letta in due modi, prima considerando solo i simboli in alto poi quelli in basso nelle coppie $\pm$ e $\mp$. Si ha che

$$\begin{aligned} 0 \leq E(\bar{X} \pm \bar{Y})^2 &= E(\bar{X}^2) + E(\bar{Y}^2) \pm 2E(\bar{X}\bar{Y}) \\ &= 2 \pm 2r(X, Y). \end{aligned} \tag{5.22}$$

Questo implica che

$$-1 \leq r(X, Y) \leq 1;$$

inoltre, se $r(X, Y) = \mp 1$ vale il segno di uguaglianza in (5.22) per cui $E(\bar{X} \pm \bar{Y})^2 = 0$. Poiché $(\bar{X} \pm \bar{Y})^2 \geq 0$ ne discende che $\bar{X} = \mp \bar{Y}$ e di conseguenza

$$Y = \mp \frac{SD(Y)}{SD(X)} X - \frac{SD(Y)}{SD(X)} E(X) + E(Y) = \mp cX + d$$

con $c = \frac{SD(Y)}{SD(X)} \geq 0$. □

CAPITOLO 6

PROBABILITÀ NEL CONTINUO

6.1. Variabili aleatorie continue

In molti problemi l'insieme dei valori che possono essere assunti da una variabile aleatoria può avere la cardinalità del continuo. In questo caso la formalizzazione di spazi di probabilità adeguati non è elementare e qui ci limitiamo a trattare un caso particolare in cui l'insieme è $\mathbb{R}$ e la probabilità è determinata da funzioni continue a tratti; questi spazi di probabilità costituiscono la distribuzione di parecchie variabili aleatorie continue che permettono di studiare numerosi problemi. In un certo senso seguiamo la linea storica di studiare prima le variabili aleatorie continue definite dalla loro distribuzione e rimandare la chiarificazione dei fondamenti.

Sarebbe possibile utilizzare funzioni con al più una infinità numerabile di punti di discontinuità , ma per semplicità consideriamo il caso in cui questi sono al più un numero finito.

DEFINIZIONE 25. *Uno* **spazio di probabilità sul continuo** *è una coppia* $(\mathbb{R}, f)$ *con* $\mathbb{R}$ *insieme dei numeri reali ed* $f : \mathbb{R} \to \mathbb{R}$ *una funzione continua tranne al più un numero finito di punti di discontinuità tale che*

(i) $f(x) \geq 0$ *per ogni* $x \in \mathbb{R}$
(ii) $\int_{-\infty}^{\infty} f(x)dx = 1$

Si noti che l'integrale nella definizione è ben definito, come integrale improprio. Infatti l'integrale esiste in ogni intervallo $[a, b]$ in cui f è continua; inoltre, detto $D(f) = \{x_1, \ldots, x_n\} < \infty$, $x_1 < \cdots < x_n$, l'insieme finito dei punti di discontinuità di f e considerate successioni $\bar{x}_j(0) \searrow_j -\infty$, $x_j(i) \nearrow_j x_i$, $\bar{x}_j(i) \searrow_j x_i$ e $x_j(n) \nearrow_j \infty$ esistono gli integrali in $[\bar{x}_j(i), x_j(i+1)]$, $i = 0, \ldots, n-1$. Per definizione di integrale improprio, (ii) dice che il limite seguente esiste, vale quanto indicato:

$$\lim_{j \to \infty} \sum_{i=0}^{n-1} \int_{\bar{x}_j(i)}^{x_j(i+1)} f(x)dx = 1$$

e non dipende dalle successioni scelte. Allo stesso modo sono definiti gli integrali per ogni intervallo $[a, b]$. Conviene in realtà considerare gli intervalli semiaperti della forma $]a, b]$, $a < b$, $a, b \in \overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ perchè si puo' esprimere un ulteriore intervallo come unione di due disgiunti.

Definizione 26. *In uno spazio di probabilità continuo* $(\mathbb{R}, f)$ *si definiscono* **eventi** *le unioni finite di intervallini semiaperti disgiunti e si indica con* $\mathcal{C}$ *la famiglia di tali insiemi. Per ogni* $A = \cup_{i=1}^{n}]a_i, b_i]$ *si pone*

$$P_f(A) = \sum_{i=1}^{n} \int_{a_i}^{b_i} f(x)dx.$$

Si noti che, benchè strettamente parlando la probabilità di un punto $x \in \mathbb{R}$ sia minore di qulla di qualsiasi intervallo che contenga questo punto, ossia 0.

Si noti inoltre che (i) ed (ii) della definizione di spazio di probabilità continuo sono l'analogo di (1) e (2) nella definizione 13. L'additività numerabile invece è qui una conseguenza delle proprietà di additività dell'integrazione rispetto al dominio. Verificheremo per semplicità solo l'additività finita.

Lemma 18. *In uno spazio di probabilità* $(\mathbb{R}, f)$ *se* $A, B \in \mathcal{C}$ *sono tali che* $\cup_{I \in A} I = \cup_{J \in B} J$ *allora* $P_f(A) = P_f(B)$.

Dimostrazione. Si noti che l'intersezione di due intervallini semiaperti non disgiunti è un intervallino semiaperto. Se $A = \{I_1, \ldots, I_n\}$ e $B = \{J_1, \ldots, J_k\}$ allora $\{L_{i,j}, i = 1, \ldots, n, j = 1, \ldots, k\} \in \mathcal{C}$ e per l'additività dell'integrale rispetto al dominio

$$P_f(A) = \sum_{i=1}^{n} \int_{I_i} f(x)dx = \sum_{i=1}^{n} \sum_{j=1}^{k} \int_{L_{i,j}} f(x)dx = \sum_{j=1}^{k} \int_{J_j} f(x)dx = P_f(B)$$

in cui la seconda e la terza uguaglianza dipendono dal fatto che se un intervallo $]a, b]$ è unione finita di intervallini $]a_i, b_i]$ disgiunti allora questi si possono ordinare in modo che $b_i = a_{i+1}$ e

$$\sum_{i=1}^{n} \int_{a_i}^{b_i} f(x)dx = \int_{a}^{b} f(x)dx.$$

□

Uno spazio di probabilità $(\mathbb{R}, f)$ può essere utilizzato per definire la distribuzione di una variabile aleatoria continua.

Definizione 27. *Si dice che una variabile aleatoria continua* X *ha distribuzione* $(\mathbb{R}, f)$ *se per ogni* $A \in \mathcal{C}$ *vale*

$$P(X \in A) = \int_{A} f(x)dx.$$

In tal caso f *è detta la densità di* X.

Si noti che l'espressione di sinistra è soltanto formale in quanto non è stato definito lo spazio di probabilità su cui è definito X e su cui agisce P.

ESEMPIO 53. *Una variabile aleatoria con densità* $f(x) = \mathbb{I}_{[0,1]}$ *è detta uniforme in* $[0, 1]$*. In generale, la variabile aleatoria con densità* $f(x) = \frac{1}{b-a}\mathbb{I}_{[a,b]}$ *è detta* **uniforme** *in* $[a, b]$ *e si indica* $X \sim U([a, b])$*.*

ESEMPIO 54. *Una variabile aleatoria con densità* $f(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ *è detta Gaussiana standard. In generale, la variabile aleatoria con densità* $f(x) = \frac{1}{\sqrt{2\pi}\sigma}e^{\frac{-(x-\mu)^2}{\sigma}}$ *è detta* **gaussiana o normale** *di media* m *e varianza* σ^2 *e si indica* $X \sim N(\mu, \sigma^2)$*.*

ESERCIZIO 25. *Verificare che se* $X \sim N(\mu, \sigma^2)$ *allora* $E(X) = \mu$ *e* $Var(X) = \sigma^2$*.*

Definiamo ora il valore atteso di una variabile aleatoria continua.

DEFINIZIONE 28. *Data una variabile aleatoria continua* X *con densità* f *si dice* **valore atteso** *di* f*:*

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx$$

se esiste

$$E(|X|) = \int_{-\infty}^{\infty} |x|f(x)dx < \infty.$$

Con questa definizione valgono tutte le proprietà viste per il valore atteso, la varianza e la deviazione standard di variabili aleatorie.

6.2. Funzioni di distribuzione

6.3. Vettori aleatori continui