FINORA ABBIAMO INTRODOTTO I CONCETTI DI

- SPAZIO DI PROBABILITÀ $(\Omega, \mathcal{A}, P)$
- VARIABILE ALEATORIA $X: \Omega \to \mathbb{R}$
- DISTRIBUZIONE $P_X: \mathcal{B}(\mathbb{R}) \to [0,1]$ DI X
- INDIPENDENZA TRA VARIABILI ALEATORIE

LE DEFINIZIONI DATE ERANO GENERALI, ANCHE SE FINORA CI SIAMO CONCENTRATI SULLE V.A. DISCRETE, CIOÈ CHE ASSUMONO SOLO UN NUMERO FINITO O NUMERABILE DI VALORI. ANCHE PER GLI SPAZI DI PROBABILITÀ, L'INSIEME Ω È STATO QUASI SEMPRE FINITO.

VEDIAMO ORA DI INTRODURRE UN'ALTRA CLASSE DI VARIABILI ALEATORIE. PER QUESTA CLASSE DAREMO LA CONTROPARTE DI CONCETTI VISTI PER LE VARIABILI ALEATORIE DISCRETE.

- DENSITÀ
- VALORE MEDIO
- VARIANZA
- COVARIANZA

NONCHÉ UN ELENCO DELLE DENSITÀ PIÙ UTILIZZATE (COME LO SONO LE DENSITÀ BINOMIALE, GEOMETRICA, IPERGEOMETRICA E POISSON TRA LE DENSITÀ DI V.A. DISCRETE).

Definizione Sia $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità. Una v.a. $X: \Omega \to \mathbb{R}$ si dice **assolutamente continua** se esiste una funzione $f: \mathbb{R} \to [0, +\infty)$ integrabile su $\mathbb{R}$ tale che

$$P\{X \in B\} = \int_B f \, dx \qquad \forall B \in \mathcal{B}(\mathbb{R}) \qquad (*)$$

dove ricordo che $\mathcal{B}(\mathbb{R})$ è la σ-algebra di Borel su $\mathbb{R}$ (la famiglia degli insiemi "decenti"). Equivalentemente

$$P_X(B) = \int_B f \, dx \qquad \forall B \in \mathcal{B}(\mathbb{R})$$

dove P_X è la distribuzione di X.

In tal caso la funzione f è detta **densità** di X (oppure si dice che X ammette f come densità).

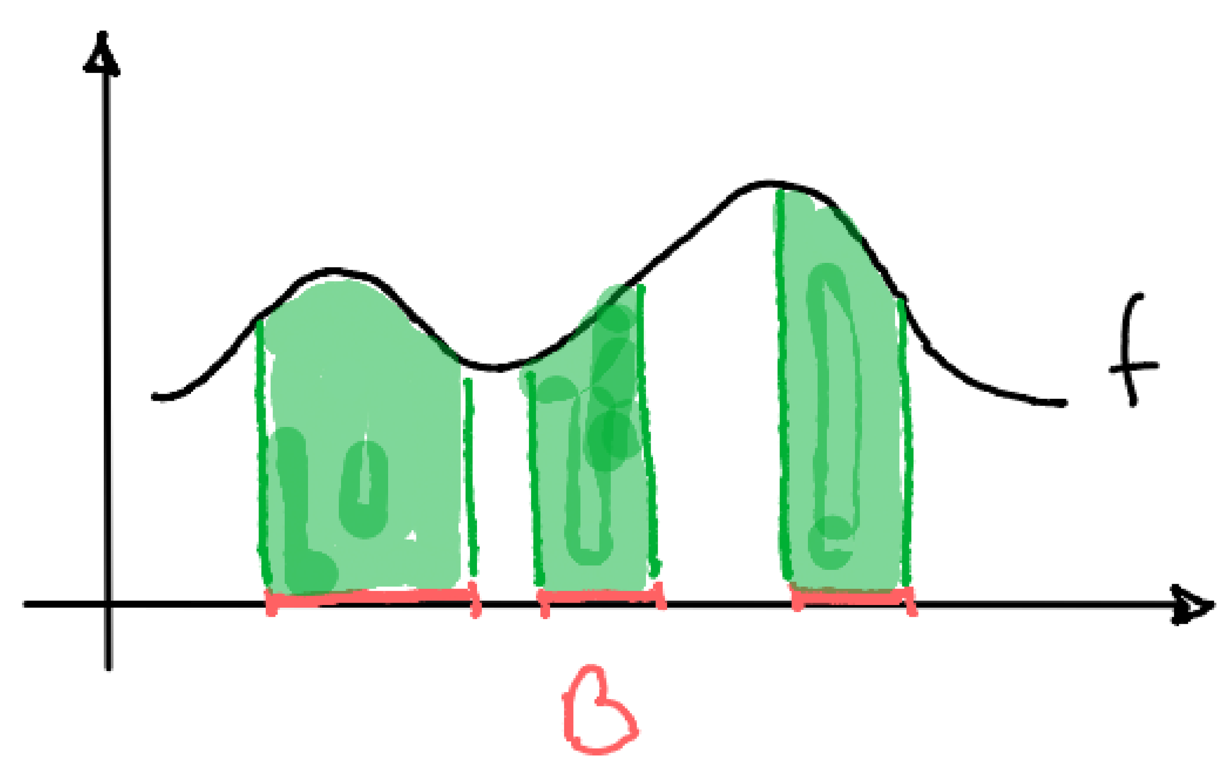
Nota Affinché una funzione X sia v.a. basta verificare che $\{X \in B\}$ è un evento quando $B = (a,b)$ $\forall a, b \in \mathbb{R}$, oppure $B = (-\infty, a)$ $\forall a \in \mathbb{R}$. Similmente, affinché f sia una densità per X basta verificare che $(*)$ vale quando $B = (a,b)$ oppure $B = (-\infty, a)$.

Nota di lavoro: Assumo che conosciate i concetti di integrale generalizzato di una funzione su un intervallo non limitato.

NOTA TECNICA Mentre è chiaro il concetto di misura di un intervallo (a,b) (ovviamente $b-a$), così come vi è noto il concetto di $\int_a^b f\,dx$, estendere il concetto di misura agli insiemi decenti $B \in \mathcal{B}$ così come dare un significato a $\int_B f\,dx$ richiede uno sforzo tecnico.

Non approfondiamo: per i nostri scopi B sarà al massimo una unione di intervalli.

Dal punto di vista geometrico visualizzate $\int_B f\,dx$ come l'area del sottografico di f con base B



NOTA TECNICA Cambiando il valore di f su un insieme finito di punti, l'integrale in $(*)$ non cambia. Quindi in generale la densità di una v.a. assolutamente continua non è unica. Volendo essere più precisi, due densità f_1 ed f_2 associate alla stessa v.a. X possono differire al più su un insieme di misura nulla (per esempio su un numero finito o numerabile di punti). Quindi la densità è unica a meno di cambiamenti irrilevanti.

<u>OSSERVAZIONE</u> MENTRE PER UNA V.A. DISCRETA ABBIAMO CHE LA DENSITÀ q SODDISFA $\sum_{x\in\mathbb{R}} q(x)=1$, E QUINDI $q(x)\in[0,1]$ $\forall x\in\mathbb{R}$, PER UNA DENSITÀ f DI UNA V.A. ASSOLUTAMENTE CONTINUA ABBIAMO

$$\int_{\mathbb{R}} f\,dx = P\{X\in\mathbb{R}\} = P(\Omega)=1$$

QUESTO PERÒ NON SIGNIFICA CHE $f(x)\in[0,1]$ $\forall x\in\mathbb{R}$.

NOTARE ANCHE CHE PER UNA V.A. DISCRETA $P\{X=x\}=q(x)$ MENTRE PER UNA V.A. ASSOLUTAMENTE CONTINUA $P\{X=x\}=0$

INTUITIVAMENTE, UNA V.A. DISCRETA HA DENSITÀ CHE SI CON= CENTRA SU UN NUMERO AL PIÙ NUMERABILE DI PUNTI (I VALORI DI X), MENTRE IN UNA V.A. ASSOLUTAMENTE CONTINUA LA DENSI= È DIFFUSA. CONVIENE INTRODURRE IL SEGUENTE CONCETTO

<u>DEFINIZIONE</u> DATA UNA V.A. X, CHIAMEREMO <u>FUNZIONE DI RIPARTIZIONE</u> DI X LA FUNZIONE $F:\mathbb{R}\to[0,1]$ DEFINITA DA

$$F(t) := P\{X\le t\} \qquad \forall t\in\mathbb{R}$$

PER LA PROPRIETÀ DI MONOTONIA DELLA PROBABILITÀ ABBIAMO CHE F È CRESCENTE. INOLTRE PER LE PROPRIETÀ DI CONTINUITÀ DELLA PROBABILITÀ

$$\lim_{t\to-\infty} F(t)=0 \qquad \lim_{t\to+\infty} F(t)=1$$

NEL CASO DI UNA V.A. ASSOLUTAMENTE CONTINUA X CON DENSITÀ f LA FUNZIONE DI RIPARTIZIONE RISULTA ESSERE

$$F(t) = \int_{-\infty}^{t} f\,dx$$

PER IL TEOREMA FONDAMENTALE DEL CALCOLO INTEGRALE, SE f È CONTINUA ALLORA F È DERIVABILE E $F' = f$.

LA COSA SI GENERALIZZA UN POCO: SE f È CONTINUA CON ECCEZIONE DI UN NUMERO FINITO DI PUNTI $x_1 < \dots < x_n$ ALLORA F È DERIVABILE IN (x_k, x_{k+1}), $k = 1, \dots, n-1$, E LÌ $F' = f$. PER <u>ESEMPIO</u>

$$f(x) = \begin{cases} 1/2 & \text{SE } x \in (0,2) \\ 0 & \text{ALTRIMENTI} \end{cases} \qquad F(x) = \begin{cases} 0 & \text{SE } x \in (-\infty, 0] \\ x/2 & \text{SE } x \in (0,2) \\ 1 & \text{SE } x \in [2, +\infty) \end{cases}$$

VICEVERSA, CONDIZIONE SUFFICIENTE AFFINCHÉ UNA V.A. X CON FUNZIONE DI RIPARTIZIONE F SIA ASSOLUTAMENTE CONTINUA È CHE F SIA C^1 A TRATTI, CIOÈ

- F SIA CONTINUA
- F SIA DERIVABILE OVUNQUE, TRANNE AL PIÙ UN NUMERO FINITO DI PUNTI $x_1, \dots, x_n$

IN QUESTO CASO LA DENSITÀ DI X È $f = F'$ (DEFINITA IN MODO ARBITRARIO SUI PUNTI $\{x_i\}$, IL COME È IRRILEVANTE)

DATI UNO SPAZIO DI PROBABILITÀ $(\Omega, \mathcal{A}, P)$, UNA V.A. $X: \Omega \to \mathbb{R}$ E UNA FUNZIONE $\psi: \mathbb{R} \to \mathbb{R}$, POSSIAMO CONSIDERARE LA COMPOSIZIONE $Y = \psi \circ X$.

SE X È DISCRETA, ALLORA Y SARÀ A SUA VOLTA UNA V.A. DISCRETA. IN GENERALE PERÒ VA VERIFICATO CHE $\{Y \in B\}$ SIA UN EVENTO PER OGNI INSIEME DECENTE $B \in \mathcal{B}(\mathbb{R})$.

UNA CONDIZIONE SUFFICIENTE È CHE ψ SIA CONTINUA.

SE PERÒ X È UNA V.A. ASSOLUTAMENTE CONTINUA, AFFINCHÉ LO SIA ANCHE Y VANNO FATTE RICHIESTE PIÙ STRINGENTI SULLA ψ.

<u>PROPOSIZIONE</u> SIA f UNA DENSITÀ PER X. SUPPONIAMO CHE ψ SIA STRETTAMENTE MONOTONA E REGOLARE.

ALLORA Y È ASSOLUTAMENTE CONTINUA CON DENSITÀ $g: \mathbb{R} \to \mathbb{R}$ DATA DA

$$g(x) = f(\psi^{-1}(x)) \, |\psi'(\psi^{-1}(x))|^{-1}$$

PROOF OMESSA. BASTA COMUNQUE TENER CONTO CHE ψ È INVERTIBILE ED UTILIZZARLA PER UN CAMBIO DI VARIABILE.

<u>ESEMPIO</u> SIA $\psi(x) = ax + b$, CON $a, b \in \mathbb{R}$ E $a \neq 0$. SE X È UNA V.A. ASS. CONTINUA, ALLORA ANCHE $Y = aX + b$ LO È, E SUA DENSITÀ È

$$g(x) = f\left(\frac{x-b}{a}\right) \frac{1}{|a|}$$

DEFINIZIONE Una v.a. assolutamente continua X con densità f si dice ammettere valore medio finito se

$$\int_{\mathbb{R}} |x| f(x)\, dx < \infty$$

In tal caso si chiama VALORE MEDIO la quantità

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

Questa definizione assomiglia molto a quella che abbiamo visto per una v.a. discreta X con densità q:

$$E[X] := \sum_{x \in \mathbb{R}} x\, q(x)$$

Possiamo dare una definizione di valore medio che abbia senso per ogni v.a. e che includa le due precedenti?

DEFINIZIONE Sia X una v.a. non negativa. Si defi= nisce VALORE MEDIO di X la quantità

$$E[X] := \sup \left\{ E[Y] : \begin{array}{l} Y: \Omega \to \mathbb{R} \text{ è una v.a. discreta} \\ \text{tale che } 0 \leq Y \leq X \end{array} \right\}$$

ammesso che sia finita.

PER UNA V.A. X QUALUNQUE, SE ENTRAMBE

$$X^+ := \max\{X, 0\} \qquad X^- := -\min\{X, 0\}$$

HANNO VALORE MEDIO FINITO DIREMO CHE X AMMETTE VALORE MEDIO FINITO. IN TAL CASO SI DEFINISCE

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

QUESTA DEFINIZIONE RESTITUISCE SIA QUELLA PER LE V.A. DISCRETE (OVVIAMENTE) SIA QUELLA PER LE V.A. ASSOLUTAMENTE CONTINUE (ANDANDOLE A CONCENTRARE).

CONVINCIAMOCI DI QUEST'ULTIMA AFFERMAZIONE.

SIA X UNA V.A. NON NEGATIVA E ASSOLUTAMENTE CONTINUA, E SIA f LA SUA DENSITÀ. SUPPONIAMO PER SEMPLICITÀ CHE $f=0$ FUORI DA UN INTERVALLO COMPATTO $[a,b]$. RISCALANDO E TRASLANDO POSSIAMO ANCHE SUPPORRE CHE $[a,b]=[0,1]$ (CI DÀ SOLO IL VANTAGGIO DI POTER DESCRIVERE PIÙ FACILMENTE LA SUDDIVISIONE IN INTERVALLINI).

FISSIAMO n E CONSIDERIAMO LE DUE V.A. DISCRETE

$$\underline{Y}(\omega) = \begin{cases} k/n & \text{SE} \quad \omega \in \left\{ \frac{k}{n} \leq X < \frac{k+1}{n} \right\} \quad k=0,\dots,n-1 \\ 0 & \text{ALTRIMENTI} \end{cases}$$

$$\overline{Y}(\omega) = \begin{cases} \frac{k+1}{M} & \text{SE} \quad \omega \in \left\{ \frac{k}{M} < X \leq \frac{k+1}{M} \right\} \quad k = 0, \dots, M-1 \\ 0 & \text{ALTRIMENTI} \end{cases}$$

ABBIAMO $\underline{Y} \leq X \leq \overline{Y}$, INOLTRE $\underline{Y}$ HA DENSITÀ

$$\underline{q}(x) = \begin{cases} P\left\{ k/M \leq X \leq k+1/M \right\} & \text{SE} \quad x = \frac{k}{M} \quad k = 0, \dots, M-1 \\ 0 & \text{ALTRIMENTI} \end{cases}$$

MENTRE $\overline{Y}$ HA DENSITÀ

$$\overline{q}(x) = \begin{cases} P\left\{ k/M \leq X \leq k+1/M \right\} & \text{SE} \quad x = \frac{k+1}{M} \quad k = 0, \dots, M-1 \\ 0 & \text{ALTRIMENTI} \end{cases}$$

PER COME ABBIAMO COSTRUITO $\underline{q}$ E $\overline{q}$

$$E[\underline{Y}] = \sum_{x \in \mathbb{R}} x \, \underline{q}(x) = \sum_{k=0}^{M-1} \frac{k}{M} P\left\{ \frac{k}{M} \leq X \leq \frac{k+1}{M} \right\}$$

$$= \sum_{k=0}^{M-1} \frac{k}{M} \int_{k/M}^{k+1/M} f(x)\,dx \leq \sum_{k=0}^{M-1} \int_{k/M}^{k+1/M} x f(x)\,dx = \int_0^1 x f(x)\,dx$$

$$= E[X] = \sum_{k=0}^{M-1} \int_{k/M}^{k+1/M} x f(x)\,dx \leq \sum_{k=0}^{M-1} \frac{k+1}{M} \int_{k/M}^{k+1/M} f(x)\,dx$$

$$= \sum_{k=0}^{M-1} \frac{k+1}{M} P\left\{ \frac{k}{M} \leq X \leq \frac{k+1}{M} \right\} = \sum_{x \in \mathbb{R}} x \, \overline{q}(x) = E[\overline{Y}]$$

INOLTRE

$$E[X]-E[\underline{Y}] \leq E[\bar{Y}]-E[\underline{Y}] = \sum_{x\in\mathbb{R}} x\left(\bar{q}(x)-\underline{q}(x)\right)$$

$$= \sum_{k=0}^{m-1} \frac{1}{m}\, P\left\{\frac{k}{m} \leq X \leq \frac{k+1}{m}\right\} = \frac{1}{m}$$

QUESTO PROVA CHE POSSIAMO APPROSSIMARE DAL BASSO BENE QUANTO VOGLIAMO $E[X]$ CON IL VALORE MEDIO DI UNA V.A. DISCRETA Y TALE CHE $0 \leq Y \leq X$ $\blacksquare$

MOLTI DEI RISULTATI VISTI PER IL VALORE MEDIO DI V.A. DISCRETE SONO VALIDI IN QUESTO AMBITO PIÙ GENERALE.

- <u>LINEARITÀ</u>: SE X_1 ED X_2 HANNO VALORE MEDIO FINITO E $c_1, c_2 \in \mathbb{R}$, ALLORA $X = c_1 X_1 + c_2 X_2$ HA VALORE MEDIO FINITO E

$$E[X] = c_1 E[X_1] + c_2 E[X_2]$$

- <u>MONOTONIA</u>: SE X_1 ED X_2 HANNO VALORE MEDIO FINITO E $X_1 \leq X_2$, ALLORA

$$E[X_1] \leq E[X_2]$$

- SE X_1 ED X_2 HANNO VALORE MEDIO FINITO, E SONO INDIPENDENTI, ALLORA $X = X_1 X_2$ HA VALORE MEDIO FINITO E

$$E[X] = E[X_1]\,E[X_2]$$

IL TEOREMA VISTO PER LE V.A. DISCRETE, RELATIVO ALLA COMPOSIZIONE CON APPLICAZIONI, VALE ANCHE PER V.A. ASSOLUTAMENTE CONTINUE. LO VEDIAMO SOLO IN VERSIONE SCALARE

TEOREMA SIA X UNA V.A. ASSOLUTAMENTE CONTINUA CON DENSITÀ f, E SIA $\varphi : \mathbb{R} \to \mathbb{R}$ UNA APPLICAZIONE CONTINUA. CONSIDERIAMO LA V.A. $Y = \varphi \circ X$. LA Y HA VALORE MEDIO FINITO SE E SOLO SE

$$\int_{\mathbb{R}} |\varphi(x)| f(x) dx < \infty$$

IN TAL CASO

$$E[Y] = \int_{\mathbb{R}} \varphi(x) f(x) dx$$

PROOF OMESSA

OSSERVAZIONI

- IL FATTO CHE φ SIA CONTINUA ASSICURA CHE Y SIA EFFETTIVAMENTE UNA V.A. QUESTA IPOTESI PUÒ ESSERE INDEBOLITA (PER ESEMPIO BASTA CHE φ SIA CONTINUA TRANNE AL PIÙ IN UN NUMERO FINITO DI PUNTI).
- IN GENERE NON POSSIAMO AFFERMARE CHE Y SIA ASSOLUTAMENTE CONTINUA. PER ESEMPIO SE $\varphi(x) = c$ ALLORA Y È UNA V.A. COSTANTE, DUNQUE DISCRETA.

Possiamo definire il concetto di varianza anche per le v.a. assolutamente continue.

<u>Definizione</u> Sia X una v.a. assolutamente continua con densità f. Diremo che X ha momento secondo finito se X^2 ha valore medio finito, cioè (per il teorema precedente) se

$$\int_{\mathbb{R}} x^2 f(x)\, dx < \infty$$

In tal caso chiamiamo <u>varianza</u> di X la quantità

$$\mathrm{Var}\, X := E[X^2] - E[X]^2$$

Questa definizione è ben posta perché (similmente a quanto visto nel caso discreto)

momento secondo finito $\Rightarrow$ valore medio finito

Inoltre $\mathrm{Var}\, X = E[(X - E[X])^2]$: come nel caso discreto questo segue dalla linearità del valor medio. La varianza misura quindi, anche in questo caso, come in media X si distribuisce intorno al suo valore medio.

Molte delle proprietà della varianza viste per le v.a. discrete rimangono confermate per le v.a. ass. continue

- $\mathrm{Var}(cX) = c^2\, \mathrm{Var}\, X$ $\qquad \forall c \in \mathbb{R}$
- $\mathrm{Var}(X+c) = \mathrm{Var}\, X$

DEFINIZIONE Date due v.a. assolutamente continue X e Y, entrambe aventi momento secondo finito, chiameremo covarianza delle due la quantità

$$\mathrm{Cov}(X,Y) := E[XY] - E[X]E[Y]$$

Diremo che X e Y sono scorrelate se $\mathrm{Cov}(X,Y)=0$.

Se X,Y sono indipendenti allora sono scorrelate.

Anche per la covarianza rimangono confermate le proprietà viste nel caso di v.a. discrete.

- La covarianza gode delle proprietà di simmetria e bilinearità: se X,Y,Z sono tre v.a. ass. continue e $a,b \in \mathbb{R}$ allora

$$\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$$

$$\mathrm{Cov}(aX+bY,Z) = a\,\mathrm{Cov}(X,Z) + b\,\mathrm{Cov}(Y,Z)$$

- La covarianza e la varianza sono legate dalla relazione

$$\mathrm{Var}(X+Y) = \mathrm{Var}\,X + \mathrm{Var}\,Y + 2\,\mathrm{Cov}(X,Y)$$

Come già osservato, la covarianza fornisce un indice dell'indipendenza tra due v.a. in termini di valori medi.