

Appunti di CPS

Omar Polo

15 febbraio 2020

Indice

1 Parti mancanti	1
2 Variabili casuali	1
2.1 Nel Discreto	1
2.1.1 Bivariate con legge discreta	1
2.2 Nel Continuo	2
2.2.1 Supporto di una v.c. con legge continua	2
2.2.2 Bivariate con legge continua	3
3 Indici di posizione	4
3.1 Moda	4
3.2 Mediana	4
3.3 Quantile- p	5
3.4 Valore atteso	5
3.4.1 Proprietà del valore atteso	5
4 Indici di variabilità	7
4.1 Varianza	7
4.1.1 Proprietà	7
4.2 Scarto quadratico medio	8
4.3 Range	8
4.4 Scarto interquantilico	8
4.5 Diseguaglianze di Markov e Čebyshev	8

4.6	Covarianza	9
5	Funzione di ripartizione e di sopravvivenza	9
5.1	Caso univariato	10
6	Funzione generatrice dei momenti	11
7	Leggi di tipo discreto	12
7.1	Leggi degeneri	13
7.2	Leggi binomiali	13
7.3	Leggi uniformi discrete	14
7.4	Leggi ipergeometriche	14
7.5	Leggi di Poisson	14
7.6	Leggi geometriche	16
8	Leggi di tipo continuo	18
8.1	Leggi uniformi continue	19
8.2	Leggi esponenziali	19
8.3	Funzione tasso di guasto	20
8.4	Leggi di Weibull	21
8.5	Leggi gamma	22
8.6	Leggi normali	23
9	Campionamento casuale semplice	26

1 Parti mancanti

- capitoli [1, 5]
- capitolo 13: “Leggi di v.c. trasformate”
- capitolo 15:
 - indici di posizione e variabilità per v.c. multivariate
 - varianza di una combinazione lineare
 - varianza di correlazione lineare

2 Variabili casuali

2.1 Nel Discreto

2.1.1 Bivariate con legge discreta

Una v.c. bivariata con legge discreta è definita da:

- il supporto congiunto $S_{X,Y} = \{(x_i, y_i) \in \mathbb{R}^2, i \in I \subseteq \mathbb{N}\}$ successione finita o numerabile di punti distinti in \mathbb{R}^2 senza punti di accumulazione all'infinito.
- la f.m.p congiunta $p_{X,Y} : S_{X,Y} \rightarrow [0, 1]$

La legge congiunta della v.c. (X, Y) è allora, per ogni $B \in \mathcal{B}_2$ data da:

$$P_{X,Y}(B) = P((X, Y) \in B) = \sum_{(x,y) \in B \cap S_{X,Y}} p_{X,Y}(x, y) \quad (1)$$

Leggi marginali Da una v.c. bivariata con legge discreta (X, Y) specificata da $S_{X,Y}$ e $p_{X,Y}(x, y)$ si possono “estrarre” varie leggi di v.c. univariate. Anzitutto conviene considerare le leggi delle componenti X e Y , dette **leggi marginali**.

La legge marginale di X ha:

- supporto marginale

$$S_X = \{x \in \mathbb{R} : (x, y) \in S_{X,Y} \text{ per qualche } y \in \mathbb{R}\}$$

- f.m.p. marginale

$$p_X(x) = P(X = x) = \sum_{y: (x,y) \in S_{X,Y}} p_{X,Y}(x, y), \text{ per } x \in S_X.$$

Similmente si può ottenere la legge marginale di Y in modo del tutto simmetrico.

Leggi condizionali Data una v.c. bivariata con legge discreta (X, Y) si possono “estrarre” anche due famiglie di leggi condizionali: la **legge condizionale di Y dato un valore osservabile di X** indicata con $Y|X = x, x \in S_X$:

- supporto condizionale

$$S_{Y|X=x} = \{y \in \mathbb{R} : (x, y) \in S_{X,Y}\}$$

- f.m.p. condizionale

$$\begin{aligned} P_{Y|X=x}(y) &= P(Y = y | X = x) = \frac{P(X = x, Y = y)}{P(X = x)} \\ &= \frac{p_{X,Y}(x, y)}{p_X(x)} \quad \text{per } y \in S_{Y|X=x}. \end{aligned}$$

e la legge condizionale di X dato un valore osservabile di Y che si deduce in modo del tutto analogo.

Componenti indipendenti e dipendenti

Definizione 1 La v.c. con legge discreta (X, Y) si dice con componenti indipendenti se

$$S_{X,Y} = S_X \times S_Y$$

e se per ogni $(x, y) \in S_{X,Y}$ si ha

$$p_{X,Y}(x, y) = p_X(x)p_Y(y).$$

Se invece le componenti di X e Y di (X, Y) non sono indipendenti si dice che la v.c. ha componenti dipendenti.

2.2 Nel Continuo

2.2.1 Supporto di una v.c. con legge continua

Il supporto S_X di una v.c. univariata X con legge continua è il più piccolo sottoinsieme chiuso di \mathbb{R} al quale P_X dà probabilità 1. Si ricorda che un insieme è chiuso se contiene tutti i suoi punti di accumulazione. Quindi $S_X \in \mathcal{B}_1$ è tale che

1. S_X è chiuso
2. $S_X \subseteq C$ per ogni $C \subseteq \mathbb{R}$ chiuso con $P_X(C) = P(X \in C) = 1$

In altre parole, S_X è la chiusura dell'insieme $x \in \mathbb{R} : p_X(x) > 0$. Si ricorda che la chiusura di un insieme numerico è l'unione fra l'insieme stesso e i punti non dell'insieme che sono però suoi punti di accumulazione.

2.2.2 Bivariate con legge continua

(X, Y) è una v.c. bivariata con legge continua se per ogni $B \in \mathcal{B}_2$, con $B = [a_1, b_1] \times [a_2, b_2]$, si ha:

$$\begin{aligned} P_{(X,Y)}(B) &= P((X, Y) \in B) \\ &= P(a_1 \leq X \leq b_1, a_2 \leq Y \leq b_2) \\ &= \iint_B p_{(X,Y)}(x, y) dx dy \end{aligned}$$

dove $p_{X,Y}(x, y)$ ha le proprietà di

- non negatività, per ogni $(x, y) \in \mathbb{R}^2$
- normalizzazione: $\iint_{\mathbb{R}^2} p_{X,Y}(x, y) dx dy = 1$.

Per una v.c. con legge continua vale che $P(X = x, Y = y)$ sia zero per ogni $(x, y) \in \mathbb{R}^2$ e che $S_{X,Y}$ sia la chiusura di $\{(x, y) \in \mathbb{R}^2 \mid p_{X,Y}(x, y) > 0\}$.

Come nel caso discreto, si possono ottenere le leggi univariate indotte, che saranno di tipo continuo:

- marginale di X con f.d.p.

$$p_X(x) = \int_{-\infty}^{+\infty} p_{X,Y}(x, y) dy$$

- marginale di Y con f.d.p.

$$p_Y(y) = \int_{-\infty}^{+\infty} p_{X,Y}(x, y) dx$$

- condizionale $Y|X = x, x \in S_X$ con f.d.p.

$$p_{Y|X=x}(y) = \frac{p_{X,Y}(x, y)}{p_X(x)}$$

- condizionale $X|Y = y, y \in S_Y$ con f.d.p.

$$p_{X|Y=y}(x) = \frac{p_{X,Y}(x, y)}{p_Y(y)}$$

Infine, i supporti si ottengono con le identiche formule del caso discreto:

$$\begin{aligned} S_X &= \{x \in \mathbb{R} \mid (x, y) \in S_{X,Y} \text{ per qualche } y\} \\ S_Y &= \{y \in \mathbb{R} \mid (x, y) \in S_{X,Y} \text{ per qualche } x\} \\ S_{Y|X=x} &= \{y \in \mathbb{R} \mid (x, y) \in S_{X,Y}\} \\ S_{X|Y=y} &= \{x \in \mathbb{R} \mid (x, y) \in S_{X,Y}\} \end{aligned}$$

Componenti indipendenti In modo del tutto analogo al caso continuo, si dice che una v.c. bivariata (X, Y) con legge continua ha componenti indipendenti se per ogni $(x, y) \in \mathbb{R}^2$ vale che

$$p_{X,Y}(x, y) = p_X(x)p_Y(y).$$

3 Indici di posizione

3.1 Moda

Definizione 2 Sia X una v.c. con legge discreta o continua e f.m.p./f.d.p. $p_X(x)$. Si dice **moda** di X , indicata con x_{mo} , un valore del supporto di X per cui

$$p_X(x_{mo}) \geq p_X(x) \quad \forall x \in S_X$$

Nel caso del continuo si richiede inoltre che la densità di X sia continua almeno da destra o da sinistra in x_{mo} .

3.2 Mediana

Definizione 3 Sia X una v.c. univariata con f.r. $F_X(x)$. Si dice **mediana** di X , indicata con $x_{0.5}$, un valore reale tale che valgano simultaneamente

$$P(X \leq x_{0.5}) \geq 0.5 \quad e \quad P(X \geq x_{0.5}) \geq 0.5$$

La mediana non è necessariamente unica. Tutte le soluzioni dell'equazione $F_X(x) = 0.5$ sono mediane di X . Se invece l'equazione non ha soluzioni, la mediana di X è unica e risulta essere il più piccolo valore di x per cui $F_X(x) \geq 0.5$.

3.3 Quantile- p

Definizione 4 Sia X una v.c. univariata con f.r. $F_X(x)$. Per $p \in (0, 1)$ si dice **quantile- p** di X , indicato con x_p , un valore reale tale che valgano simultaneamente:

$$P(X \leq x_p) \geq p \quad e \quad P(X \geq x_p) \geq 1 - p.$$

Si tratta di una generalizzazione del concetto di mediana, infatti la mediana è anche detta quantile-0.5.

Come la mediana, anche il quantile- p non è necessariamente unico.

3.4 Valore atteso

Il valore atteso $E(X)$ è la media aritmetica ponderata dei valori assumibili dalla v.c. con pesi dati dalla f.m.p. Se la variabile ha supporto continuo, la ponderazione è data dalla f.d.p. e la somma viene intesa in senso continuo, ovvero come un integrale.

$E(X)$ è quindi definita come

$$E(X) = \begin{cases} \sum_{x \in S_X} xp_X(x) & \text{se } X \text{ ha legge discreta} \\ \int_{-\infty}^{+\infty} xp_X(x) dx & \text{se } X \text{ ha legge continua} \end{cases}$$

Si richiede che la somma (o l'integrale) convergano assolutamente, quindi:

$$\sum_{x \in S_X} |x|p_X(x) < +\infty \quad \text{oppure} \quad \int_{-\infty}^{+\infty} |x|p_X(x) dx < +\infty$$

3.4.1 Proprietà del valore atteso

Valore atteso di trasformate Siano X e Y v.c. univariate con $Y = g(X)$, allora vale che

$$E(Y) = E(g(X)) = \begin{cases} \sum_{x \in S_X} g(x)p_X(x) & \text{se } X \text{ ha legge discreta} \\ \int_{-\infty}^{+\infty} g(x)p_X(x) dx & \text{se } X \text{ ha legge continua} \end{cases}$$

Valore atteso del prodotto di v.c. indipendenti Sia $X = (X_1, \dots, X_d)$ una v.c. con componenti indipendenti, allora

$$E(X_1, \dots, X_d) = \prod_{i=1}^d E(X_i)$$

Proprietà di Cauchy Quando esiste finito, il valore atteso può non essere un punto del supporto di X , ma è sempre intermedio fra i punti del supporto. Supponendo, senza perdita di generalità che $S_X = \{x_1, \dots, x_k\}$

$$x_1 < x_2 < \dots < x_k$$

si ha

$$x_1 \leq x_i \leq x_k \quad \forall i \in \{1, 2, \dots, k\}$$

e quindi

$$x_1 p_X(x_i) \leq x_i p_X(x_i) \leq x_k p_X(x_i) \quad \forall i \in \{1, 2, \dots, k\}.$$

Sommando le disuguaglianze si ottiene

$$x_1 \sum_{i=1}^k p_X(x_i) \leq \sum_{i=1}^k x_i p_X(x_i) \leq x_k \sum_{i=1}^k p_X(x_i)$$

da cui, per la normalizzazione, si ottiene

$$x_1 \leq E(X) \leq x_k.$$

Proprietà di linearità Siano X e Y v.c. univariate con $Y = aX + b$, allora

$$E(Y) = E(aX + b) = a E(X) + b.$$

Proprietà di linearità generalizzata Se (X, Y) è una v.c. bivariata e $T = aX + bY$ una combinazione lineare delle componenti di (X, Y) con $a, b \in \mathbb{R}$, allora

$$E(t) = E(aX + bY) = a E(X) + b E(Y).$$

Proprietà del baricentro Si tratta di un caso particolare della linearità

$$E(X - E(X)) = 0.$$

Proprietà dei minimi quadrati Se X è una v.c. univariata e i valori attesi indicati esistono, allora per ogni $c \in \mathbb{R}$

$$E((X - c)^2) \geq E((X - E(X))^2)$$

dove

- c è una predizione puntuale della realizzazione futura di X
- $X - c$ è l'errore di predizione
- $(X - c)^2$ è la perdita quadratica dovuta all'errore di predizione
- $E((X - c)^2)$ è la perdita quadratica media, detta rischio quadratico

4 Indici di variabilità

Sia X una v.c. univariata con legge discreta e supporto finito. Si possono cogliere aspetti importanti della distribuzione di X attraverso indici sintetici.

Definizione 5 (Indice di variabilità) *Un indice di variabilità di una v.c. univariata X è una predizione di quanto disterà il valore futuro di X da una sua particolare predizione puntuale, quindi in sostanza un indice di prevedibilità di X .*

4.1 Varianza

L'indice di variabilità principale, la varianza di X , è definito come la media aritmetica ponderata del quadrato degli scarti di X dal proprio valore atteso.

$$\text{Var}(X) = E((X - E(X))^2) = \begin{cases} \sum_{x \in \mathcal{S}_X} (x - E(X))^2 p_X(x) & \text{legge continua} \\ \int_{-\infty}^{+\infty} (x - E(X))^2 p_X(x) dx & \text{legge discreta} \end{cases}$$

4.1.1 Proprietà

Non negatività Per ogni X con varianza finita $\text{Var}(X) \geq 0$. $\text{Var}(X) = 0$ solo per $X \sim \mathcal{D}(x_0)$.

Formula per il calcolo

$$\text{Var}(X) = E(X^2) - (E(X))^2$$

Invarianza rispetto a traslazioni

$$\text{Var}(X + b) = \text{Var}(X)$$

Omogeneità di secondo grado

$$\text{Var}(aX) = a^2 \text{Var}(X)$$

4.2 Scarto quadratico medio

Lo scarto quadratico medio, o deviazione standard, è definito come la radice quadrata aritmetica della varianza. In simboli

$$\sigma_X = \sqrt{\text{Var}(x)}$$

4.3 Range

Il *range* di una v.c. univariata X , indicato con R_X è definito per variabili con supporto limitato come

$$R_X = \max(S_X) - \min(S_X).$$

4.4 Scarto interquantilico

Lo scarto interquantilico di una v.c. univariata X , indicato con IQR_X è definito come

$$IQR_X = x_{0.75} - x_{0.25}$$

4.5 Diseguaglianze di Markov e Čebyshev

Teorema 1 (Diseguaglianza di Markov) *Sia X una v.c. non negativa con valore atteso $\mu = E(X) > 0$ finito. Allora per ogni $c > 0$ vale*

$$P(X \geq c\mu) \leq \frac{1}{c}.$$

Teorema 2 (Diseguaglianza di Čebyshev) *Sia X una v.c. univariata con valore atteso $\mu = E(X)$ finito e varianza $\sigma^2 = \text{Var}(X) > 0$ anch'essa finita. Allora, per ogni $k > 0$ vale*

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}.$$

Entrambe le diseguaglianze sono poco informative per $c \in (0, 1]$ o $k \in (0, 1]$, ma diventano molto informative negli altri casi.

4.6 Covarianza

La covarianza è un indice sintetico della dipendenza delle componenti di una v.c. bivariata. Se indichiamo il supporto come successione dipendente di due indici:

$$S_{X,Y} = (x_i, y_i) \quad i \in I, j \in J \text{ con } I \text{ e } J \text{ finiti o numerabili}$$

ed esprimere la f.m.p. in forma abbreviata come applicazione:

$$(x_i, y_i) \rightarrow p_{ij} = P(X = x_i, Y = y_i).$$

Allora la covarianza, indicata con il simbolo $\text{Cov}(X, Y)$, è definita come media ponderata del prodotto di scarti:

$$\begin{aligned} \text{Cov}(X, Y) &= \sum_{i \in I} \sum_{j \in J} (x_i - E(X))(y_j - E(Y))p_{ij} \\ &= E((X - E(X))(Y - E(Y))) \end{aligned}$$

Per il calcolo della covarianza è nota una formula per il calcolo analoga a quella che si usa per la varianza

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y).$$

5 Funzione di ripartizione e di sopravvivenza

Definizione 6 Si dice *funzione di ripartizione* di $X = (X_1, \dots, X_d)$ la funzione

$$F_X : \mathbb{R}^d \rightarrow [0, 1]$$

che a ciascun punto $x = (x_1, \dots, x_d)$ di \mathbb{R}^d fa corrispondere il valore d'immagine

$$\begin{aligned} F_X(x) &= P(X_1 \leq x_1, \dots, X_d \leq x_d) \\ &= P_X((-\infty, x_1] \times \dots \times (-\infty, x_d]) \\ &= P(\cap_{i=1}^d \{s \in S \mid X_i(s) \leq x_i\}). \end{aligned}$$

Se X ha componenti indipendenti, per ogni $x \in \mathbb{R}^d$ vale la relazione

$$F_X(x) = F_{X_1, \dots, X_d}(x_1, \dots, x_d) = \prod_{i=1}^d F_{X_i}(x_i)$$

Teorema 3 (Proprietà strutturali) Sia X una v.c. univariata con legge qualsiasi. La funzione di ripartizione $F_X(x)$ è una applicazione $F_X : \mathbb{R} \rightarrow [0, 1]$ che gode delle seguenti proprietà:

- è monotona non decrescente

$$x_1 < x_2 \Rightarrow F_X(x_1) \leq F_X(x_2)$$

- è continua da destra in ogni punto $x \in \mathbb{R}$

$$\forall x \in \mathbb{R} \quad \lim_{\varepsilon \rightarrow 0^+} F_X(x + \varepsilon) = F_X(x)$$

- i limiti agli estremi del dominio sono zero ed uno:

$$\begin{aligned} \lim_{x \rightarrow -\infty} F_X(x) &= 0 \\ \lim_{x \rightarrow +\infty} F_X(x) &= 1 \end{aligned}$$

Teorema 4 (Caratterizzazione) Se $F : \mathbb{R} \rightarrow [0, 1]$ ha le proprietà 1, 2 e 3 del Teorema 3 allora esiste una v.c. univariata X con legge di probabilità P_X di cui F ne è la funzione di ripartizione.

5.1 Caso univariato

Se X è una v.c. univariata la sua $F_X(x)$ permette di calcolare agevolmente le probabilità degli intervalli.

$$P(a < X \leq b) = F_X(b) - F_X(a) \quad \text{con } a < b$$

Funzione di sopravvivenza La funzione $P(X > x)$ viene detta funzione di sopravvivenza di X

$$P(X > x) = 1 - P(X \leq x) = 1 - F_X(x).$$

Dalla $p_X(x)$ alla $F_X(x)$ Sia X una v.c. univariata. Dalla definizione si ha subito che

$$F_X(x) = P(X \leq x) = \begin{cases} \sum_{t \in S_X | t \leq x} p_X(t) & \text{se } X \text{ ha legge discreta} \\ \int_{-\infty}^x p_X(t) dt & \text{se } X \text{ ha legge continua} \end{cases}$$

Si possono quindi fare le seguenti osservazioni:

- se X ha legge discreta univariata, la $F_X(x)$ è costante a tratti, con punti di salto posizionati ai punti del supporto, e con valore del salto pari alla massa di probabilità posta sul punto;
- se X ha legge continua, la sua funzione di ripartizione è una funzione continua di $x \in \mathbb{R}$, in quanto primitiva di una funzione integrabile.

Le varie $F_X(x)$ calcolate verranno mostrate nelle sezioni delle leggi.

Dalla $F_X(x)$ alla funzione di massa di probabilità Data la $F_X(x)$ di una v.c. discreta X si può recuperare il supporto come insieme dei punti di salto

$$S_X = \{x \in \mathbb{R} \mid F_X(x) - \lim_{\varepsilon \rightarrow 0^+} F_X(x - \varepsilon) > 0\}$$

da cui poi dedurre

$$p_X(x) = P(X = x) = \lim_{\varepsilon \rightarrow 0^+} P(x - \varepsilon < X \leq x) = F_X(x) - \lim_{\varepsilon \rightarrow 0^+} F_X(x - \varepsilon)$$

Dalla $F_X(x)$ alla funzione di densità di probabilità Nei punti x in cui $p_X(x)$ è continua, la $F_X(x)$ è derivabile e vale

$$p_X(x) = \frac{d}{dx} F_X(x).$$

6 Funzione generatrice dei momenti

Definizione 7 (Momenti) Si dicono *momenti* di una v.c. univariata X i valori

$$\mu_r = E(X^r), r = 1, 2, \dots$$

In particolare, μ_r è detto **momento r -esimo** di X .

Si ha subito che:

$$\begin{aligned} \mu_1 &= \mu = E(X) \\ \mu_2 - \mu_1^2 &= E(X^2) - (E(X))^2 = \text{Var}(X) \end{aligned}$$

Definizione 8 (Funzione generatrice dei momenti) Sia X una v.c. univariata con f.m.p. o f.d.p. $p_X(x)$. La funzione generatrice dei momenti di X , indicata con $M_X(t)$, è una funzione di variabile reale definita da

$$M_X(t) = E(e^{tX}) = \begin{cases} \sum_{x \in S_X} e^{tx} p_X(x) & \text{se } X \text{ ha legge discreta} \\ \int_{-\infty}^{+\infty} e^{tx} p_X(x) dx & \text{se } X \text{ ha legge continua} \end{cases}$$

Tale funzione gode delle seguenti proprietà:

- nell'origine vale sempre $M_X(0) = 1$ per ogni X ;
- il dominio di finitezza di $M_X(t)$, ovvero $D_X = \{t \in \mathbb{R} \mid M_X(t) < +\infty\}$, è **convesso** (ovvero è un intervallo, una semiretta o l'intera retta reale);

- $M_X(t) > 0 \quad \forall t \in D_X$

Definizione 9 (Funzione generatrice dei momenti propria) Si dice che la v.c. univariata X ha funzione generatrice dei momenti propria se il dominio di finitezza di $M_X(t)$ include l'origine come punto interno.

Una proprietà interessa delle funzioni generatrici dei momenti proprie è che nell'origine hanno derivate di ogni ordine, i cui valori “generano” i momenti di X

$$\mu_r = E(X^r) = \left. \frac{d^r}{dt^r} M_X(t) \right|_{t=0} = M_X^{(r)}(0).$$

che rende in alcuni casi più semplice calcolare valore atteso e varianza.

Definizione 10 (Generatrice della somma di v.c. indipendenti) Sia $X = (X_1, \dots, X_d)$ una v.c. multivariata con componenti X_i indipendenti che hanno funzione generatrice dei momenti propria $M_{X_i}(t)$. Allora $S = \sum_{i=1}^d X_i$ ha funzione generatrice dei momenti propria

$$M_S(t) = \prod_{i=1}^d M_{X_i}(t).$$

7 Leggi di tipo discreto

Le leggi discrete sono in generale individuate da due ingredienti:

1. un insieme senza punti di accumulazione al finito

$$S_X = \cup_{x \in I} x_i, \quad x_i \in \mathbb{R}^d, \quad i \in I \subseteq \mathbb{N}^+$$

detto **supporto** della variabile casuale.

2. una applicazione

$$p_X : S_X \rightarrow [0, 1]$$

detta **funzione di massa di probabilità** che soddisfi le condizioni:

- $p_X(x) > 0$ per ogni $x \in S_x$
- $\sum_{x \in S_x} p_X(x) = 1$

Dato il supporto e la funzione massa di probabilità, la legge discreta corrispondente è definita da

$$P_X(B) = P(X \in B) = \sum_{x \in S_X \cap B} p_X(x).$$

7.1 Leggi degeneri

Si dice che una v.c. d -variata X ha legge degenerare in $x_0 \in \mathbb{R}^d$, valore prefissato, e si scrive $X \sim \mathcal{D}(x_0)$ se per $B \in \mathcal{B}_d$ la legge di probabilità di X è

$$P_X(B) = P(X \in B) = \begin{cases} 1 & \text{se } x_0 \in B \\ 0 & \text{se } x_0 \notin B. \end{cases}$$

Funzione di ripartizione

$$F_X(x) = \begin{cases} 0 & \text{se } x < x_0 \\ 1 & \text{se } x \geq x_0 \end{cases}$$

Generatrice dei momenti, valore atteso e varianza

$$M_X(t) = E(e^{tX}) = \sum_{x \in S_X} e^{tx} p_X(x) = e^{tx_0}(1) = e^{tx_0}$$

$$M'_X(t) = x_0 e^{tx_0} \Rightarrow E(X) = M'_X(0) = x_0$$

$$M''_X(t) = x_0^2 e^{tx_0} \Rightarrow E(X^2) = M''_X(0) = x_0^2$$

da cui $\text{Var}(X) = E(X^2) - (E(X))^2 = x_0^2 - (x_0)^2 = 0$.

7.2 Leggi binomiali

Si dice che la v.c. univariata X ha legge binomiale con indice $n \in \mathbb{N}^+$ e parametro $p \in (0, 1)$, e si scrive $X \sim Bi(n, p)$, se per ogni $B \in \mathcal{B}_1$ vale

$$P_X(B) = P(X \in B) = \sum_{x \in \{0, 1, \dots, n\} \cap B} \binom{n}{x} p^x (1-p)^{n-x}.$$

Quando $n = 1$ le leggi binomiali sono dette di **di Bernoulli** o **binomiali elementari**.

Alcuni risultati noti:

$$F_X(x) = \begin{cases} 0 & \text{se } x < 0 \\ 1-p & \text{se } 0 \leq x < 1 \\ 1 & \text{se } x \geq 1 \end{cases} \quad \text{supponendo } X \sim Bi(1, p)$$

Proprietà additiva Sia $X = (X_1, \dots, X_d)$ una v.c. multivariata con componenti X_i indipendenti e legge marginale $X_i \sim Bi(n_i, p)$, dove $n_i \in \mathbb{N}^+$ e $p \in (0, 1)$ per cui

$$M_{X_i}(t) = (1 - p + pe^t)^{n_i}$$

allora $S = \sum_{i=1}^d X_i$ ha f.g.m. propria

$$\begin{aligned} M_S(t) &= \prod_{i=1}^d M_{X_i}(t) \\ &= \prod_{i=1}^d (1 - p + pe^t)^{n_i} \\ &= (1 - p + pe^t)^{\sum_{i=1}^d n_i} \end{aligned}$$

per cui

$$S \sim Bi\left(\sum_{i=1}^d n_i, p\right).$$

7.3 Leggi uniformi discrete

La v.c. d -variata X ha legge uniforme discreta in $D = \cup_{i=1}^k x_i, x_i \in \mathbb{R}^d$ dove gli x_i sono k punti distinti, e si scrive in breve $X \sim Ud(X_1, \dots, x_k)$ se

- $S_X = D$
- $p_X(x) = 1/k$ per ogni $x \in S_X$.

7.4 Leggi ipergeometriche

Si dice che la v.c. univariata X ha legge ipergeometrica con indice $n \in \mathbb{N}^+$ e parametri N e D , dove $n \leq N \in \mathbb{N}^+$ e $D \in \mathbb{N}^+$ con $D \leq N$, e si scrive in breve $X \sim IG(n; D, N)$ se vale

$$P_X(\{x\}) = P(X = x) = \frac{\binom{D}{x} \binom{N-D}{n-x}}{\binom{N}{n}}$$

per tutti i valori x per cui hanno senso i coefficienti binomiali e 0 altrimenti. Usando il simbolo S_x per i valori di x per cui $P(X = x) > 0$ le leggi P_X ipergeometriche sono tali che per ogni $B \in \mathcal{B}_1$

$$P_X(B) = P(X \in B) = \sum_{x \in S_x \cap B} \frac{\binom{D}{x} \binom{N-D}{n-x}}{\binom{N}{n}}$$

7.5 Leggi di Poisson

Definizione 11 Si dice che X ha legge di Poisson con parametro $\lambda > 0$ e si scrive $X \sim P(\lambda)$ se è una v.c. univariata con legge discreta con supporto $S_X = \mathbb{N}$ e f.m.p. per $x \in S_X$ pari a

$$p_X(x) = e^{-\lambda} \frac{\lambda^x}{x!}.$$

Per verificare che si tratti di una buona definizione occorre controllare le solite due condizioni:

- la positività di $p_X(x)$ sul supporto $S_X = \mathbb{N}$ è banale perché $p_X(x)$ è il prodotto di tre fattori positivi.
- la normalizzazione segue da

$$\begin{aligned} \sum_{x \in S_X} p_X(x) &= \sum_{x=0}^{+\infty} e^{-\lambda} \frac{\lambda^x}{x!} \\ &= e^{-\lambda} \sum_{x=0}^{+\infty} \frac{\lambda^x}{x!} \quad \text{noto: } \sum_{x=0}^{+\infty} \frac{\lambda^x}{x!} = e^\lambda \\ &= e^{-\lambda} e^\lambda = e^{-\lambda+\lambda} = e^0 = 1. \end{aligned}$$

Si usa tipicamente $X \sim P(\lambda), \lambda > 0$ per modellare la distribuzione di una variabile casuale che esprime un **conteggio** con supporto illimitato superiormente (o molto più grande dei valori tipicamente assunti dal conteggio). È questo il caso di $Bi(n, p)$ per $p = \lambda/n$ ed n sufficientemente grande.

Valore atteso

$$E(x) = \sum_{x \in S_X} x p_X(x) = \sum_{x=0}^{+\infty} x e^{-\lambda} \frac{\lambda^x}{x!} = \lambda \sum_{x=1}^{+\infty} e^{-\lambda} \frac{\lambda^{x-1}}{(x-1)!} = \lambda \sum_{x=0}^{+\infty} e^{-\lambda} \frac{\lambda^x}{x!} = \lambda$$

Funzione generatrice dei momenti e varianza

$$\begin{aligned} M_X(t) &= E(e^{tX}) = \sum_{x \in S_X} e^{tx} p_X(x) \\ &= \sum_{x=0}^{+\infty} (e^t)^x e^{-\lambda} \frac{\lambda^x}{x!} \\ &= e^{-\lambda} \sum_{x=0}^{+\infty} \frac{(\lambda e^t)^x}{x!} \\ &= e^{-\lambda} e^{\lambda e^t} \\ &= e^{\lambda(e^t-1)} \end{aligned}$$

da cui

$$\begin{aligned} M'_X(t) &= e^{\lambda(e^t-1)} \lambda e^t \implies E(X) = M'_X(0) = \lambda \\ M''_X(t) &= e^{\lambda(e^t-1)} \lambda^2 e^{2t} + e^{\lambda(e^t-1)} \lambda e^t \implies E(X^2) = M''_X(0) = \lambda^2 + \lambda \end{aligned}$$

e pertanto

$$\text{Var}(X) = \text{E}(X^2) - (\text{E}(X))^2 = \lambda^2 + \lambda - (\lambda)^2 = \lambda.$$

Proprietà additiva Sia $X = (X_1, \dots, X_d)$ una v.c. multivariata con componenti X_i indipendenti e legge marginale $X_i \sim P(\lambda_i)$ dove $\lambda_i > 0$ per $i = 1, \dots, d$

$$M_{X_i}(t) = e^{\lambda_i(e^t - 1)}$$

allora $S = \sum_{i=1}^d X_i$ ha f.g.m. propria

$$\begin{aligned} M_S(t) &= \prod_{i=1}^d M_{X_i}(t) \\ &= \prod_{i=1}^d e^{\lambda_i(e^t - 1)} \\ &= \exp \left\{ \left(\sum_{i=1}^d \lambda_i \right) (e^t - 1) \right\} \end{aligned}$$

da cui si evince che

$$S \sim P \left(\sum_{i=1}^d \lambda_i \right).$$

7.6 Leggi geometriche

Definizione 12 Si dice che X ha legge geometrica con parametro $p \in (0, 1)$ e si scrive $X \sim \text{Ge}(p)$ se è una v.c. univariata con legge discreta che ha supporto $S_X = \mathbb{N}^+$ e f.m.p. per $x \in S_X$ pari a

$$p_X(x) = p \times (1 - p)^{x-1}.$$

Verifichiamo che sia una buona definizione:

- positività di $p_X(x)$ sul supporto $S_X = \mathbb{N}^+$ banale perché è il prodotto di fattori positivi

- la normalizzazione segue da

$$\begin{aligned}
 \sum_{x \in S_X} p_X(x) &= \sum_{x=1}^{+\infty} p(1-p)^{x-1} \\
 &= p \sum_{x=1}^{+\infty} (1-p)^{x-1} \\
 &= p \sum_{i=0}^{+\infty} (1-p)^i \\
 &= p \frac{1}{1 - (1-p)} = 1.
 \end{aligned}$$

dove si è usata il risultato della serie geometrica con ragione x , $|x| < 1$:

$$\begin{aligned}
 \sum_{i=0}^{\infty} x^i &= \lim_{n \rightarrow \infty} \sum_{i=0}^n x^i \\
 &= \lim_{n \rightarrow \infty} \frac{1 - x^{n+1}}{1 - x} \\
 &= \frac{1}{1 - x}.
 \end{aligned}$$

Per una legge geometrica $X \sim Ge(p)$ il parametro p rappresenta $P(X = 1)$.

La funzione di sopravvivenza è, per $x \in \mathbb{N}^+$

$$P(X > x) = P(C_1 \cap \dots \cap C_x) = P(C_1) \times \dots \times P(C_x) = (1-p)^x$$

per cui la funzione di ripartizione è

$$F_X(x) = P(X \leq x) = 1 - P(X > x) = 1 - (1-p)^x.$$

Per $x \in \mathbb{R}$ invece si ha

$$F_X(x) = \begin{cases} 0 & \text{se } x < 1 \\ 1 - (1-p)^{\lfloor x \rfloor} & \text{se } x \geq 1. \end{cases}$$

Se $0 < p < 1$ la v.c. $X \sim Ge(p)$ può essere vista come il tempo d'attesa nel tempo discreto che gode della proprietà di assenza della memoria, ovvero per ogni $s, t \in \mathbb{N}^+$

$$P(X > s+t \mid X > s) \quad \text{non dipende da } s.$$

Valore atteso

$$E(X) = 1/p.$$

Generatrice dei momenti e varianza

$$\begin{aligned}
M_X(t) &= E(e^{tX}) = \sum_{x \in S_X} e^{tx} p_X(x) \\
&= \sum_{x=1}^{+\infty} e^{t(x-1+1)} p(1-p)^{x-1} \\
&= pe^t \sum_{x=1}^{+\infty} e^{t(x-1)} (1-p)^{x-1} \\
&= pe^t \sum_{i=0}^{+\infty} (e^t(1-p))^i \quad \text{posto } i = x-1 \\
&= \frac{pe^t}{1 - (1-p)e^t}
\end{aligned}$$

Saltando alcuni passaggi si trova che

$$M'_X(t) \Big|_{t=0} = \frac{pe^t}{(1 - (1-p)e^t)^2} \Big|_{t=0} = \frac{1}{p}.$$

Con calcoli lasciati al lettore¹

$$\text{Var}(X) = \frac{1-p}{p^2}.$$

8 Leggi di tipo continuo

Si dice che X è una **v.c. univariata con legge di tipo continuo** se per ogni $B \in \mathcal{B}_1$ si può esprimere $P_X(B)$ in forma integrale come

$$P_X(B) = P(X \in B) = \int_B p_X(x) dx = \int_a^b p_X(x) dx$$

se $B = [a, b]$ con $a < b$, dove la funzione $p_X(x)$ detta funzione di densità di probabilità soddisfa le condizioni:

1. $p_X(x) \geq 0$ per ogni $x \in \mathbb{R}$
2. $\int_{\mathbb{R}} p_X(x) dx = \int_{-\infty}^{+\infty} p_X(x) dx = 1$

Una P_x definita in questo modo soddisfa gli assiomi di Kolmogorov.

¹dal professore, non da me. Prendetevela con lui.

8.1 Leggi uniformi continue

Si dice che X ha legge uniforme continua in (a, b) , dove $a < b$, e si scrive $X \sim U(a, b)$ se la f.d.p. di X è

$$p_X(x) = \begin{cases} \frac{1}{b-a} & \text{se } x \in (a, b) \\ 0 & \text{se } x \notin (a, b). \end{cases}$$

da cui valori atteso e varianza si ricavano facilmente come:

$$E(X) = \int_{-\infty}^{+\infty} x p_X(x) dx = \int_a^b x \frac{1}{b-a} dx = \frac{1}{b-a} \left[\frac{1}{2} x^2 \right]_a^b = \frac{a+b}{2}$$

$$E(X^2) = \int_a^b x^2 \frac{1}{b-a} dx = \frac{1}{b-a} \left[\frac{1}{3} x^3 \right]_a^b = \frac{a^2 + ab + b^2}{3}$$

$$\text{Var}(X) = E(X^2) - (E(X))^2 = \dots = \frac{(b-a)^2}{12}.$$

8.2 Leggi esponenziali

Si dice che X ha legge esponenziale con parametro $\lambda > 0$, e si scrive $X \sim \text{Esp}(\lambda)$, se la f.d.p. di X è

$$p_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{se } x \geq 0 \\ 0 & \text{se } x < 0 \end{cases}$$

La verifica della non-negatività è banale, e anche la condizione di normalizzazione si verifica facilmente:

$$\begin{aligned} \int_{-\infty}^{+\infty} p_X(x) dx &= \int_0^{+\infty} p_X(x) dx = \int_0^{+\infty} \lambda e^{-\lambda x} dx = \left[-e^{-\lambda x} \right]_0^{+\infty} \\ &= -\lim_{x \rightarrow +\infty} e^{-\lambda x} - (-e^0) = 1. \end{aligned}$$

Calcolo del valore atteso:

$$E(X) = \int_{-\infty}^{+\infty} x p_X(x) dx = \int_0^{+\infty} x \lambda e^{-\lambda x} dx = \frac{1}{\lambda} \int_0^{+\infty} \lambda x e^{-\lambda x} dx = 1$$

dove l'ultimo integrale vale 1 (si calcola per parti dopo aver applicato la sostituzione $\lambda x = t$).

Le leggi esponenziali sono la modellazione di *default* per tempi d'attesa. Il risultato sul valore atteso dà un significato al parametro λ . Per un tempo d'attesa esponenziale, $1/\lambda$ è il valore medio dell'attesa. Quindi:

$$\lambda = \frac{1}{\text{media dell'attesa}} = \frac{1}{E(X)}$$

Funzione di ripartizione

$$F_X(x) = \begin{cases} 0 & \text{se } x < 0 \\ 1 - e^{-\lambda x} & \text{se } x \geq 0 \end{cases}$$

Funzione generatrice dei momenti, valore atteso e varianza

$$\begin{aligned} M_X(t) = E(e^{tX}) &= \int_0^{+\infty} e^{tx} \lambda e^{-\lambda x} dx \\ &= \lambda \int_0^{+\infty} e^{-(\lambda-t)x} dx \\ &= \lambda \times \frac{\lambda-t}{\lambda-t} \int_0^{+\infty} e^{-(\lambda-t)x} dx \\ &= \frac{\lambda}{\lambda-t} \underbrace{\int_0^{+\infty} (\lambda-t) e^{-(\lambda-t)x} dx}_{\substack{\text{ovvero la f.d.p della legge} \\ \text{Esp}(\lambda-t) \text{ che per la} \\ \text{normalizzazione vale 1}}} \\ &= \frac{\lambda}{\lambda-t} \end{aligned}$$

da cui

$$\begin{aligned} M'_X(t) &= - \left(1 - \frac{t}{\lambda}\right)^{-2} \left(-\frac{1}{\lambda}\right) = \frac{1}{\lambda} \left(1 - \frac{t}{\lambda}\right)^{-2} \\ M''_X(t) &= \frac{-2}{\lambda} \left(1 - \frac{t}{\lambda}\right)^{-3} \left(-\frac{1}{\lambda}\right) = \frac{2}{\lambda^2} \left(1 - \frac{t}{\lambda}\right)^{-3} \end{aligned}$$

da cui infine si ottiene

$$\text{Var}(X) = E(X^2) - (E(X))^2 = \frac{1}{\lambda^2}.$$

8.3 Funzione tasso di guasto

Sia T un tempo d'attesa, quindi una v.c. univariata con $P(T \geq 0) = 1$ e legge continua. Siano date

funzione di ripartizione $F_T(t) = P(T \leq t)$

funzione di sopravvivenza $\bar{F}_T(t) = 1 - F_T(t) = P(T > t)$

f.d.p. $p_T(t) = \frac{d}{dt} F_T(t)$ supposta continua ovunque, salvo che in un numero finito di punti.

Definizione 13 Si dice *funzione tasso di guasto di T* (o *hazard rate, failure rate*) la funzione $r_T(\cdot)$ definita per i valori di t per cui $F_T(t) < 1$ da

$$r_T(t) = \frac{p_T(t)}{\bar{F}_T(t)} = -\frac{d}{dt} \log \bar{F}_T(t).$$

Nei punti in cui $r_T(t)$ è continua, è anche proporzionale alla probabilità che l'attesa, ancora viva al tempo t , termini entro il tempo $t + \varepsilon$, ossia nel tempuscolo immediatamente successivo a t .

Dalla funzione tasso di guasto si può determinare la funzione di ripartizione e la funzione di densità di probabilità di T , infatti:

$$\int_0^t r_T(u) du = -\log \bar{F}_T(t)$$

da cui si ottiene che per $t > 0$

$$F_T(t) = 1 - \exp \left\{ - \int_0^t r_T(u) du \right\}$$

e

$$p_T(t) = r_T(t) \times \exp \left\{ - \int_0^t r_T(u) du \right\}.$$

8.4 Leggi di Weibull

Un tempo d'attesa T_0 ha legge di Weibull monoparametrica con parametro di forma $c > 0$ se per $t > 0$ il suo tasso di guasto è

$$r_{T_0}(t) = c \times t^{c-1}$$

Interessante è notare come $Esp(1)$ faccia parte delle leggi di Weibull, ma $Esp(\lambda), \lambda \neq 1$ no. Per questo motivo, si introduce un secondo parametro alle leggi di Weibull:

$$r_T(t) = \lambda c (\lambda t)^{c-1}.$$

Le esponenziali diventano un caso particolare $Esp(\lambda) \sim W(1, \lambda)$.

Si noti come $r_{T_0}(t)$ sia decrescente per $0 < c < 1$, costante per $c = 1$ e crescente se $c > 1$.

Si può calcolare la funzione di ripartizione tenendo a mente che $\int_0^t r_T(u) du = (\lambda t)^c$:

$$F_T(t) = 1 - \exp\{-(\lambda t)^c\}$$

e la corrispondente f.d.p.

$$p_T(t) = \lambda c (\lambda t)^{c-1} \exp\{-(\lambda t)^c\}.$$

8.5 Leggi gamma

Si tratta di un secondo modo per modellare tempi d'attesa nel continuo con tasso di guasto monotono.

Definizione 14 *Un tempo d'attesa T_0 ha legge gamma monoparametrica con parametro di forma $\alpha > 0$ se per $t > 0$ la sua funzione di densità di probabilità è*

$$p_{T_0}(t) = \frac{1}{\Gamma(\alpha)} t^{\alpha-1} e^{-t}.$$

La funzione $\Gamma(\alpha)$ è la funzione gamma di Eulero ed è definita dall'integrale convergente per $\alpha > 0$

$$\Gamma(\alpha) = \int_0^{+\infty} t^{\alpha-1} e^{-t} dt.$$

che è l'estensione della nozione di fattoriale per i numeri reali positivi.

Anche in questo caso introdurre un parametro *scala* $\lambda > 0$:

$$p_T(t) = \frac{\lambda^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\lambda t}.$$

in modo che le leggi esponenziali diventi un caso particolare di leggi gamma $Exp(\lambda) \sim Ga(1, \lambda)$.

Si dimostra (non qui) che $r_{T_0}(t)$ è decrescente se $0 < \alpha < 1$, costante se $\alpha = 1$ e crescente se $\alpha > 1$

Funzione generatrice dei momenti, valore atteso e varianza Alcuni passaggi sono stati omessi per brevità

$$\begin{aligned} M_X(t) &= E(e^{tX}) = \int_0^{+\infty} e^{tx} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} dx \\ &= \lambda^\alpha \times \int_0^{+\infty} \frac{x^{\alpha-1}}{\Gamma(\alpha)} e^{-(\lambda-t)x} dx \\ &= \frac{\lambda^\alpha}{(\lambda-t)^\alpha} \times \int_0^{+\infty} \frac{(\lambda-t)^\alpha}{\Gamma(\alpha)} e^{-(\lambda-t)x} dx \\ &= \frac{\lambda^\alpha}{(\lambda-t)^\alpha} \end{aligned}$$

Si ottiene

$$\begin{aligned} E(X) &= M'_X(0) = \frac{\alpha}{\lambda} \\ E(X^2) &= M''_X(0) = \frac{\alpha(\alpha+1)}{\lambda^2} \end{aligned}$$

da cui

$$\text{Var}(X) = E(X^2) - (E(X))^2 = \frac{\alpha}{\lambda^2}.$$

Proprietà additiva Sia $X = (X_1, \dots, X_d)$ una v.c. multivariata con componenti X_i indipendenti e legge marginale $X_i \sim Ga(\alpha_i, \lambda)$ dove $\alpha_i > 0$ per ogni $i = 1, \dots, n$ e $\lambda > 0$ per cui

$$M_{X_i}(t) = \left(1 - \frac{t}{\lambda}\right)^{-\alpha_i}$$

allora $S = \sum_{i=1}^d X_i$ ha f.g.m. propria

$$\begin{aligned} M_S(t) &= \prod_{i=1}^d M_{X_i}(t) \\ &= \prod_{i=1}^d \left(1 - \frac{t}{\lambda}\right)^{-\alpha_i} \\ &= \left(1 - \frac{t}{\lambda}\right)^{-\sum_{i=1}^d \alpha_i} \end{aligned}$$

da cui si evince che

$$S \sim Ga\left(\sum_{i=1}^d \alpha_i, \lambda\right).$$

8.6 Leggi normali

Definizione 15 (Legge normale standard) Una v.c. univariata Z con supporto $S_Z = \mathbb{R}$ e f.d.p

$$p_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

è detta con legge normale standard, in breve $Z \sim N(0, 1)$.

Definizione 16 (Legge normale con parametri) Una v.c. univariata $X = \mu + \sigma Z$ con $Z \sim N(0, 1)$ è detta normale con parametro di posizione μ e parametro di scala σ , in breve $X \sim N(\mu, \sigma^2)$ e a f.d.p.

$$p_X(x) = \frac{1}{\sqrt{2\pi} \sigma} \exp \left\{ -\frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2 \right\}$$

Un'applicazione delle leggi normali è lo studio degli errori di misurazione. Si supponga di effettuare misurazioni ripetute con lo stesso strumento di una certa quantità μ . Le misure x_i , affette da errore, possono essere modellate

come realizzazione della variabile casuale $X \rightarrow x_i$. Gli stessi errori di misurazione (ignoti) possono essere modellati come variabile casuale $Z \rightarrow z_i$. In questo ultimo caso però conviene usare una scala *standard*², e perciò

$$x_i = \mu + \sigma z_i \quad i = 1, \dots, n \quad \sigma \in \mathbb{R}$$

da cui per confronto si può dedurre che X non è altro che una trasformata

$$X = \mu + \sigma Z.$$

Chiusura sotto trasformazioni affini Se $X \sim N(\mu, \sigma^2)$ e $T = a + bX$, con $b \neq 0$ allora

$$T \sim N(a + b\mu, b^2\sigma^2).$$

La dimostrazione segue dall'applicazione della definizione di X , dal notare che Z è simmetrica $Z \sim -Z$ e dal confronto:

$T = a + bX$	definizione di T
$\sim a + b(\mu + \sigma Z)$	definizione di X
$\sim a + b\mu + b\sigma Z$	
$\sim a + b\mu + b \sigma Z$	per simmetria
$\sim N(a + b\mu, b^2\sigma^2)$	per confronto.

Funzione generatrice dei momenti

$M_X(t) = E(e^{tX})$	def. di f.g.m.
$= E(e^{t(\mu + \sigma Z)})$	def. di X
$= E(e^{t\mu} e^{t\sigma Z})$	
$= e^{t\mu} E(e^{t\sigma Z})$	linearità di E
$= e^{t\mu} M_Z(t\sigma)$	def. di f.g.m.

²che quindi va *scalata* di caso in caso con l'ausilio di un fattore σ

esplicitiamo M_Z

$$\begin{aligned}
 M_Z(t) &= E(e^{tZ}) = \int_{-\infty}^{+\infty} e^{tz} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz \\
 &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z^2 - 2tz + t^2 - t^2)} dz \\
 &= e^{\frac{1}{2}t^2} \underbrace{\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z-t)^2} dz}_{\substack{\text{si tratta della f.d.p. di} \\ N(t, 1) \text{ e quindi per la} \\ \text{normalizzazione vale 1}}} \\
 &= e^{\frac{1}{2}t^2}
 \end{aligned}$$

e quindi

$$M_X(t) = e^{t\mu} e^{\frac{1}{2}(t\sigma)^2} = e^{t\mu + \frac{1}{2}t^2\sigma^2}$$

Essendo $M_Z(t)$ finita per ogni $t \in \mathbb{R}$, Z ha f.g.m. propria.

Ora è possibile calcolare valore atteso e varianza:

$$\begin{aligned}
 E(X) &= M'_X(t) \Big|_{t=0} = \mu \\
 E(X^2) &= M''_X(t) \Big|_{t=0} = \sigma^2 + \mu^2 \\
 \text{Var}(X) &= E(X^2) - (E(X))^2 = \sigma^2 + \mu^2 - (\mu)^2 = \sigma^2.
 \end{aligned}$$

Proprietà additiva Se $X \sim N(\mu_X, \sigma_X^2)$ e $Y \sim N(\mu_Y, \sigma_Y^2)$ sono indipendenti allora

$$S = X + Y \sim N(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2).$$

Funzione di ripartizione La funzione di ripartizione di $Z \sim N(0, 1)$ verrà indicata con $\Phi(z)$ definita come:

$$\Phi(z) = P(Z \leq z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du.$$

È possibile ricondurre la f.r. di una v.c. normale $X \sim N(\mu, \sigma^2)$ a quella della normale standard Z :

$$\begin{aligned}
 F_X(x) &= P(X \leq x) && \text{def. di f.r.} \\
 &= P(\mu + \sigma Z \leq x) && \text{def. di } X \\
 &= P\left(Z \leq \frac{x - \mu}{\sigma}\right) \\
 &= \Phi\left(\frac{x - \mu}{\sigma}\right).
 \end{aligned}$$

9 Campionamento casuale semplice

Un esperimento di campionamento casuale semplice è l'osservazione ripetuta in circostanze indipendenti di un dato fenomeno che manifesta, di ripetizione in ripetizione, una apprezzabile variabilità. Il numero di repliche n dell'osservazione si dice **numerosità**.

Effettuato un esperimento di c.c.s. si ottiene $y = (y_1, \dots, y_n)$, che può essere vista come realizzazione di una v.c. n -variata $Y = (Y_1, \dots, Y_n)$ con componenti Y_i indipendenti e identicamente distribuite.

Il supporto di Y è detto spazio campionario, indicato con \mathcal{Y} ed è definito come

$$\mathcal{Y} = S_{Y_1} \times \dots \times S_{Y_n} = S_{Y_1}^n.$$

Possiamo vedere $p_{Y_1}(y_1)$ come una funzione dipendente da un parametro θ ignoto. Y ha quindi legge

$$p_Y(y; \theta) = p_{Y_1, \dots, Y_n}(y_1, \dots, y_n; \theta) = \prod_{i=1}^n p_{Y_1}(y_i; \theta)$$

dove $y \in \mathcal{Y} = S_{Y_1}^n$.

Statistiche riassuntive Dato che nel mondo reale il valore del parametro θ è solitamente ignoto, il campionamento casuale semplice viene usato come mezzo per determinarlo. Il campione osservato permetterà di illuminare aspetti altrimenti ignoti della legge di popolazione.

Una volta ottenuto il campione y , allo scopo di determinare θ , si calcola una opportuna statistica riassuntiva $t = t(y)$, tipicamente della stessa dimensione di θ . Tale statistica riassuntiva può dipendere dalla forma di $p_Y(y; \theta)$.

Definizione 17 *La distribuzione campionaria di una statistica $t = t(y)$ è la legge di probabilità della v.c. trasformata $T = t(Y)$ quando la legge di Y è descritta da $P_Y(y; \theta)$.*

Quest'ultima definizione sottolinea come anche la nostra $t = t(y)$, dato che dipende dal campione y , è realizzata dal caso, e quindi non sarà in grado di indicare con certezza il valore di θ .

C.c.s. da popolazione normale con σ noto

$$\begin{aligned}
 p_Y(y; \mu) &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{1}{2} \frac{(y_i - \mu)^2}{\sigma^2} \right\} \\
 &= \left(\frac{1}{\sqrt{2\pi}\sigma} \right)^n \times \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \mu)^2 \right\} \\
 &= \left(\frac{1}{\sqrt{2\pi}\sigma} \right)^n \times \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i^2 + \mu^2 - 2\mu y_i) \right\} \\
 &= \left(\frac{1}{\sqrt{2\pi}\sigma} \right)^n \times \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n y_i^2 - \frac{n\mu^2}{2\sigma^2} + \frac{\mu}{\sigma^2} \sum_{i=1}^n y_i \right\} \\
 &= \underbrace{\left(\frac{1}{\sqrt{2\pi}\sigma} \right)^n \times \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n y_i^2 \right\}}_{\text{che chiameremo } h(y) \text{ per semplicità}} \times \underbrace{\exp \left\{ -\frac{n\mu^2}{2\sigma^2} + \frac{\mu}{\sigma^2} \sum_{i=1}^n y_i \right\}}_{\text{dato che non dipende da } \mu} \\
 &\quad \text{che chiameremo } g(\cdot, \mu) \text{ dato che dipende da } \mu \\
 &= h(y) \times g \left(\sum_{i=1}^n y_i; \mu \right)
 \end{aligned}$$

L'informazione portata su μ dalla statistica somma è equivalentemente portata dalla statistica media campionaria $\bar{y}_n = s_n/n$ analogo empirico di $E(Y)$. La distribuzione campionaria della statistica media campionaria è

$$\bar{Y}_n \sim N \left(\mu, \frac{\sigma^2}{n} \right).$$