Volume 1

Richiami e Complementi di Calcolo delle Probabilità

Roberta Pappadà
Nicola Torelli

Università degli Studi di Trieste

Dipartimento di Scienze Economiche, Aziendali, Matematiche e Statistiche "Bruno de Finetti" (DEAMS)

Corso di laurea in Statistica e Informatica per l'Azienda, la Finanza e l'Assicurazione

a.a. 2020-2021

Indice

Capitolo 1

Probabilità e variabili aleatorie

1.1 Variabili aleatorie univariate e loro caratteristiche

Il termine *variabile aleatoria* (v.a.) indica una quantità il cui valore è determinato in relazione a un esperimento aleatorio. Le quantità d'interesse sono funzioni a valori reali definite sullo spazio campionario, e sono anche dette *variabili casuali*. Nell'ambito dell'inferenza statistica ha rilievo, ad esempio, l'esperimento aleatorio definito dall'estrazione di una (o più) unità da una popolazione. In genere poi l'interesse è su una o più caratteristiche di tale unità che è spesso possibile esprimere con un valore numerico. In base a tale idealizzazione sarà possibile riassumere tramite variabili aleatorie l'incertezza su quanto si osserva in un campione di unità, se questo viene selezionato dalla popolazione secondo schemi assimilabili a esperimenti casuali. Potremo quindi esprimere, attraverso una variabile aleatoria, il numero dei punti ottenuti con il lancio di due dadi, ma anche la durata di vita di una persona oppure il numero di sinistri che sperimenta un assicurato, o ancora il numero di coloro che rispondono ad un sondaggio di voler votare un determinato partito. Si adotterà la notazione comune di indicare con la lettera maiuscola $(X, Y, \dots)$ le variabili casuali e con la minuscola i valori che esse assumono $(x, y, \dots)$. Tali valori vengono chiamati anche "determinazioni" o "realizzazioni" della variabile. La trattazione riguarderà le variabili casuali cosiddette discrete e continue. Inoltre in questi brevi ed essenziali richiami non verranno trattati alcuni aspetti di carattere teorico nell'ambito della probabilità. Coloro che volessero approfondire le nozioni richiamate nei seguenti capitoli possono fare riferimento a manuali e libri di testo di calcolo delle probabilità; si veda, ad esempio [1].

1.1.1 Variabili aleatorie discrete e continue

Definizione 1.1 (Variabile aleatoria)**.** Una variabile aleatoria X è una funzione che associa a ciascun elemento ω di uno spazio campionario Ω un numero reale $X(\omega) = x$. L'insieme dei valori x che assume la variabile aleatoria è il *supporto* R_X.

Definizione 1.2 (Variabile aleatoria discreta)**.** Una variabile aleatoria X è detta *discreta* se il suo supporto R_X contiene un insieme finito o un'infinità numerabile di valori.

In questo caso è possibile associare, secondo una determinata regola, un numero ad ogni evento elementare dello spazio campionario generato dall'esperimento aleatorio (e.g., il lancio di una moneta, il lancio di due dadi). Ad ogni valore x del supporto di X, possiamo assegnare la *probabilità* P che X assuma il valore x, definendo così una *distribuzione di probabilità*. Il legame di dipendenza della probabilità dal valore di X è dato dalla *funzione di probabilità* o *funzione di densità discreta* di X.

Definizione 1.3 (Funzione di probabilità)**.** Data una variabile aleatoria discreta X con supporto R_X, chiamiamo *funzione di probabilità* di X la funzione

$$p(x) = P(X = x), \qquad \text{per } x \in R_X \tag{1.1}$$

Valgono le seguenti proprietà:

$$p(x) \geq 0, \qquad \sum_{x \in R_X} p(x) = 1$$

Introduciamo il concetto di funzione di ripartizione.

Definizione 1.4 (Funzione di ripartizione)**.** La *funzione di ripartizione (f.d.r.)* (o funzione di distribuzione) $F(x)$ di una variabile aleatoria X fornisce in corrispondenza di ciascun valore reale x la probabilità che X assuma valori minori o uguali a x

$$F(x) = P(X \leq x) \quad \forall x \in \mathbb{R} \tag{1.2}$$

Si ha una *variabile aleatoria continua* quando lo spazio campionario Ω su cui essa è definita è continuo; è costituito, cioè, da una infinità non numerabile di eventi elementari (si pensi, ad esempio, alla durata di una lampadina elettrica o di una conversazione telefonica). In altri termini, le variabili casuali continue possono assumere tutti i valori compresi in un intervallo reale (generalmente queste variabili producono risposte che derivano da un processo di misurazione). La definizione di variabile aleatoria continua può essere data in termini della funzione $F(\cdot)$ come segue.

Definizione 1.5 (Variabile aleatoria continua)**.** Se la variabile aleatoria X è dotata di funzione di ripartizione $F(x)$ continua in $x \in \mathbb{R}$, allora X è una *variabile aleatoria continua.*

Definizione 1.6 (Funzione di densità)**.** Sia X una variabile aleatoria continua con funzione di ripartizione $F(x)$. La *funzione di densità (di probabilità)* $f(x)$ di X è

$$f(x) = \frac{dF(x)}{dx}. \tag{1.3}$$

Per costruzione la funzione di densità deve essere non negativa, cioè $f(x) \geq 0$ per ogni x e deve avere un'area complessiva uguale a 1

$$\int_{-\infty}^{\infty} f(x)dx = 1.$$

L'insieme dei valori x per cui $f(x) > 0$ è il supporto della variabile aleatoria X. La funzione di densità della variabile aleatoria X è quindi tale che la probabilità che X sia compresa tra due valori a e b, $a < b$, è uguale all'area sotto la funzione $f(x)$ compresa tra a e b

$$P(a \leq X \leq b) = \int_a^b f(x)dx.$$

La funzione di ripartizione è un concetto di fondamentale importanza, in quanto ci consente di calcolare le probabilità di tutti gli eventi di interesse collegati alla variabile casuale. Essa gode delle seguenti proprietà:

1. $F(x)$ è non decrescente; quindi per $x_1 > x_0$ si avrà $F(x_1) \geq F(x_0)$;
2. $0 \leq F(x) \leq 1$ ed è pari a 0 se x tende a $-\infty$ mentre è pari a 1 se x tende a ∞;
3. $F(x)$ è continua a destra, cioè $\lim_{h \to 0^+} F(x+h) = F(x)$;
4. $F(x) = \int_{-\infty}^{x} f(t)dt$, nel caso continuo,

   $= \sum_{t \leq x} p(t)$, nel caso discreto,
5. Per ogni coppia di numeri a e b tali che $a < b$

   $F(b) - F(a) = P(a \leq X \leq b) = \int_a^b f(x)dx$, nel caso continuo,

   $= P(a < X \leq b) = \sum_A p(x)$, nel caso discreto,

   essendo $A = \{x \colon a < x \leq b\}$.

1.1.2 Valore atteso e altre misure sintetiche

Definizione 1.7 (Valore atteso o media di una variabile aleatoria discreta)**.** Sia X una variabile aleatoria discreta con funzione di probabilità $p(x)$. Il *valore atteso* o *media* della variabile X è definito come

$$E(X) = \sum_{x \in R_X} xp(x), \tag{1.4}$$

ammesso che tale serie sia assolutamente convergente. In caso contrario, il valore atteso non esiste.

Definizione 1.8 (Valore atteso o media di una variabile aleatoria continua)**.** Sia X una variabile aleatoria continua con funzione di densità $f(x)$. Il valore atteso o media $E(X)$ è

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx, \tag{1.5}$$

se l'integrale converge.

Spesso si usa il termine *speranza matematica* con riferimento a $E(X)$.

Definizione 1.9 (Valore atteso o media di una funzione di v.a.)**.** Sia $g(X)$ la funzione di una variabile aleatoria X. Il valore atteso di $g(X)$ è

$$\begin{aligned} E(g(X)) &= \sum_{x \in R_X} g(x)p(x) \qquad \text{se X è discreta} \\ &= \int_{-\infty}^{\infty} g(x)f(x)dx \quad \text{se X è continua} \end{aligned} \tag{1.6}$$

Teorema 1.1. Le proprietà fondamentali del valore atteso (o media) sono le seguenti

a. $E(c) = c$, ove c è un qualsiasi numero reale;

b. $E(cX) = cE(X)$ o più in generale $E(cg(X)) = cE(g(X))$;

c. $E(g(X) + h(X)) = E(g(X)) + E(h(X))$;

d. $E(g(X)) \leq E(h(X))$ se $g(x) \leq h(x)$ per ogni x.

L'indice fondamentale di variabilità per una v.a. X è la varianza che misura quanto distano in media i valori di X dalla loro media $E(X)$.

Definizione 1.10 (Varianza e deviazione standard)**.** Sia X una variabile aleatoria e sia $\mu_X = E(X)$. La varianza di X, denotata con $V(X)$ o talvolta σ^2, è definita come la media di $(X - \mu_X)^2$, cioè

$$V(X) = E[(X - \mu_X)^2] \tag{1.7}$$

La radice quadrata con segno positivo della varianza di X è detta *deviazione standard*. Esplicitando la definizione, si ha che

$$\begin{aligned} V(X) &= \sum_{x \in R_x} (x - \mu_X)^2 \, p(x) \qquad \text{se X è discreta} \\ &= \int_{-\infty}^{\infty} (x - \mu_X)^2 \, f(x) dx \quad \text{se X è continua} \end{aligned}$$

Teorema 1.2. Le proprietà fondamentali della varianza sono le seguenti

a. $V(c) = 0$, per c costante;

b. $V(cX) = c^2 V(X)$;

c. $V(X + c) = V(X)$;

d. $V(X) \geq 0$.

Inoltre la varianza ha l'espressione alternativa

$$V(X) = E(X^2) - [E(X)]^2. \tag{1.8}$$

Infatti, si può scrivere

$$\begin{aligned} V(X) &= E[(X - \mu_X)^2] = E(X^2 - 2X\mu_X + \mu_X^2) \\ &= E(X^2) - 2\mu_X E(X) + E(\mu_X^2) \\ &= E(X^2) - 2\mu_X^2 + \mu_X^2 = E(X^2) - \mu_X^2, \end{aligned}$$

per cui si ottiene la (1.8). Si noti che, per la proprietà d. della varianza, sia ha

$$E(X^2) \geq [E(X)]^2.$$

Sia X una variabile aleatoria e $Y = aX + b$, con a, b costanti. Per le proprietà della media e della varianza si ha

$$\begin{aligned} E(Y) &= aE(X) + b \\ V(Y) &= a^2 V(X). \end{aligned}$$

La prima uguaglianza segue dal Teorema 1.1 (proprietà a, b, c). Per la verifica della seconda uguaglianza, si può scrivere usando il Teorema 1.2

$$V(aX + b) = V(aX) = a^2 V(X)$$

In molti casi si richiede di individuare, data una certa probabilità p, qual è la soglia x_p tale che la probabilità che X sia non superiore a x_p risulti essere almeno uguale a p. I *quantili* di una distribuzione forniscono informazioni utili a descrivere e sintetizzare alcuni aspetti di una distribuzione di probabilità.

Definizione 1.11 (Quantili e percentili)**.** Fissato un valore p tale che $0 \leq p \leq 1$, il *quantile di ordine* p, x_p, o quantile p-esimo della variabile aleatoria X è il più piccolo numero che soddisfa

$$P(X \leq x_p) = F(x_p) \geq p$$

Se X è una variabile aleatoria continua con funzione di ripartizione F, allora il quantile p-esimo è tale che $F(x_p) = p$, ovvero

$$x_p = F^{-1}(p),$$

dove $F^{-1}(\cdot)$ corrisponde alla funzione inversa della funzione di ripartizione, la **funzione dei quantili**, tale che ad ogni livello di probabilità p, si associa un valore x tale che l'area a sinistra di x sotto la curva di densità è esattamente pari a p: $P(X \leq x) = p$. Si osservi che a $p = 0.5$ corrisponde la *mediana* della distribuzione.

Il valore x_p rappresenta inoltre il $100p$-esimo *percentile* $(0 \leq p \leq 1)$. Infine, particolarmente importanti sono il *I* e *III quartile* di una distribuzione, che corrispondono al 25-esimo e 75-esimo percentile, rispettivamente.

1.1.3 Due importanti disuguaglianze

Teorema 1.3 (Disuguaglianza di Markov)**.** Sia X una variabile aleatoria tale che $P(X \geq 0) = 1$. Allora

$$P(X \geq c) \leq \frac{E(X)}{c} \qquad \text{per ogni } c > 0. \tag{1.9}$$

Dimostrazione. Sia $u(\cdot)$ una funzione non negativa avente come dominio la retta reale; si procede dimostrando che vale la relazione

$$P(u(X) \geq c) \leq \frac{E(u(X))}{c}. \tag{1.10}$$

Indichiamo con A l'insieme dei valori di X per i quali si ha $u(X) \geq c$; nel caso discreto si può scrivere

$$E(u(X)) = \sum_A u(X)p(x) + \sum_{\overline{A}} u(X)p(x) \geq \sum_A u(X)p(x) \geq c\sum_A p(x)$$

e poiché $\sum\limits_A p(x) = P(u(X) \geq c)$, si perviene alla (1.10). La (1.9) si ricava banalmente per $u(X) = X$. Se X è continua, la dimostrazione è analoga. □

Teorema 1.4 (Disuguaglianza di Jensen)**.** Data una funzione ϕ convessa e una variabile aleatoria X, si ha (se esistono i valori medi considerati)

$$E(\phi(X)) \geq \phi(E(X)). \tag{1.11}$$

Dimostrazione. Poiché $\phi(x)$ è continua e convessa, esiste una retta $l(x) = a + bx$ che soddisfa

$$l(x) = a + bx \leq \phi(x)$$

e $l(E(X)) = \phi(E(X))$. Tale retta passa per il punto $(E(X), \phi(E(X)))$. Ora, si noti che $E(l(X)) = E(a + bX) = a + bE(X) = l(E(X))$; quindi

$$\phi(E(X)) = l(E(X)) = E(l(X))$$

e poiché deve essere $E(l(X)) \leq E(\phi(X))$ per la proprietà **d.** dei valori attesi, si ottiene la (1.11). □

In generale $\phi(E(X)) \neq E(\phi(x))$; per quanto appena visto, si ha ad esempio che $E(X^2) \geq [E(X)]^2$, essendo $\phi(x) = x^2$ convessa.

1.2 Momenti e funzioni generatrici

Definizione 1.12 (Momenti)**.** Il *momento r-mo* (o di ordine r) di una variabile aleatoria X è definito come il valore atteso di X^r, ovvero

$$m_r = E(X^r) \tag{1.12}$$

per $r = 1, 2, 3, \ldots$, posto che il valore atteso esista.

Si chiama *momento centrale di ordine* r di una variabile aleatoria X il valore atteso di $(X - E(X))^r$, ovvero

$$\overline{m}_r = E[(X - E(X))^r] \tag{1.13}$$

posto che tale valore atteso esista.

Pertanto il momento primo è nient'altro che la media di X

$$m_1 = E(X^1) = E(X) \tag{1.14}$$

Il momento secondo è $m_2 = E(X^2)$. Quindi la varianza di X è il secondo momento centrale, ovvero $V(X) = \overline{m}_2 = m_2 - m_1^2$. I primi due momenti danno quindi informazioni su valori centrali e varianza di X. I successivi momenti sono legati a caratteristiche specifiche della distribuzione. In particolare il momento centrale terzo è legato alla simmetria della distribuzione attraverso l'*indice di asimmetria o skewness*

$$\beta_1 = \frac{\overline{m}_3}{\overline{m}_2^{3/2}}.$$

Si verifica infatti che $\beta_1 = 0$ se la distribuzione è simmetrica, mentre $\beta_1 > 0$ o $\beta_1 < 0$ a seconda che la distribuzione presenti asimmetria positiva o negativa. La *curtosi* è la caratteristica della forma della distribuzione che misura lo spessore delle code di una distribuzione: il maggiore/minore spessore delle code della distribuzione determina un maggiore/minore appiattimento della forma della distribuzione. Il coefficiente di curtosi è dato dalla formula:

$$\gamma_2 = \beta_2 - 3, \quad \text{dove}\, \beta_2 = \frac{\overline{m}_4}{\overline{m}_2^2},$$

in termini dei momento centrali quarto e secondo.

Al fine di calcolare i momenti è possibile usare direttamente la definizione di valore atteso e le proprietà dei valori attesi di funzioni. Essi possono però essere più agevolmente calcolati in molti casi sfruttando una particolare funzione che caratterizza, quando esiste, in modo univoco una variabile aleatoria X, detta **funzione generatrice dei momenti**.

Definizione 1.13 (Funzione generatrice dei momenti)**.** La *funzione generatrice dei momenti* $m(t)$ per una variabile aleatoria X è definita $\forall t \in \mathbb{R}$ come

$$\begin{aligned} m(t) = E(e^{tX}) &= \sum_{x \in R_X} e^{tx} p(x) \qquad \text{se X è discreta,} \\ &= \int_{-\infty}^{\infty} e^{tx} f(x) dx \quad \text{se X è continua,} \end{aligned} \tag{1.15}$$

posto che il valore atteso esista per t in un intorno dello 0 (ovvero $E(e^{tX})$ esiste $\forall t$ in un intervallo $[-h, h]$ con $h > 0$).

Si ricorda che la funzione e^{tx} può essere sviluppata in serie come segue

$$e^{tx} = 1 + tx + \frac{t^2x^2}{2!} + \frac{t^3x^3}{3!} + \dots \tag{1.16}$$

pertanto in virtù delle proprietà del valore atteso si ha

$$\begin{aligned} m(t) = E(e^{tX}) &= E\left(1 + tX + \frac{t^2X^2}{2!} + \frac{t^3X^3}{3!} + \dots\right) \\ &= 1 + tE(X) + \frac{t^2E(X^2)}{2!} + \frac{t^3E(X^3)}{3!} + \dots \\ &= 1 + tm_1 + \frac{t^2m_2}{2!} + \frac{t^3m_3}{3!} + \dots \end{aligned} \tag{1.17}$$

dove m_1, m_2, m_3 sono i momenti di X. Quindi

$$\begin{aligned} \frac{dm(t)}{dt} &= m_1 + tm_2 + \frac{t^2 m_3}{2!} + \dots \\ \frac{d^2 m(t)}{dt^2} &= m_2 + tm_3 + \frac{t^2 m_4}{2!} + \dots \end{aligned} \tag{1.18}$$

e valutando tali derivate in $t = 0$ si ha

$$\begin{aligned} \left.\frac{dm(t)}{dt}\right|_{t=0} &= m_1 \\ \left.\frac{d^2 m(t)}{dt^2}\right|_{t=0} &= m_2 \end{aligned}$$

e così via. Se si considera la derivata r-esima di $m(t)$ e la si valuta poi nel punto $t = 0$ si ottiene il momento r-esimo di X:

$$m_r = E(X^r) = \left.\frac{d^r m(t)}{dt^r}\right|_{t=0} \tag{1.19}$$

Infatti se si può ammettere la derivabilità sotto il segno di integrale, come avviene nella generalità dei casi, si può scrivere

$$\frac{d^r m(t)}{dt^r} = E\left(\frac{d^r e^{tX}}{dt^r}\right) = E(X^r e^{tX}),$$

da cui, ponendo $t = 0$, si ottiene la (1.19).

Come accennato, la funzione generatrice dei momenti gode di alcune interessanti proprietà tra cui quella di **unicità**.

Proprietà di unicità della funzione generatrice dei momenti.
Se due variabili aleatorie X e Y hanno la stessa funzione generatrice dei momenti, allora esse hanno la stessa funzione di ripartizione. Viceversa, se X e Y hanno la stessa funzione di ripartizione, esse hanno la stessa funzione generatrice dei momenti, ammesso che esista.

La funzione generatrice dei momenti di una funzione lineare di X si può ottenere in modo semplice dalla funzione generatrice dei momenti di X. Infatti, si assuma X variabile aleatoria con funzione generatrice dei momenti

$m_X(t)$ e sia $Y = aX + b$, con a e b costanti. Allora

$$\begin{aligned} m_Y(t) &= E\left(e^{tY}\right) \\ &= E\left(e^{t(aX+b)}\right) \\ &= E\left(e^{atX}e^{bt}\right) \\ &= e^{bt}E\left(e^{atX}\right) \\ &= e^{bt}m_X(at). \end{aligned} \tag{1.20}$$

Da ciò deriva il seguente teorema.

Teorema 1.5. Sia X una variabile aleatoria con funzione generatrice dei momenti $m_X(t)$ e siano $a \neq 0$ e b due costanti. Se $Y = aX + b$, si ha

$$m_Y(t) = e^{bt}m_X(at).$$

Definizione 1.14 (Funzione generatrice di cumulanti)**.** Il logaritmo della funzione generatrice dei momenti di X (assumendo che esista) viene definito come la funzione generatrice di cumulanti di X:

$$c(t) = \ln m(t) \tag{1.21}$$

Per molte variabili aleatorie la *funzione generatrice di cumulanti* fornisce un modo semplice di ricavare media e varianza.

Il *cumulante r-esimo* di X è il coefficiente di $t^r/r!$ nello sviluppo in serie di Taylor della funzione generatrice dei cumulanti.

Le derivate prima e seconda di $c(t)$ valutate in $t = 0$ restituiscono la media e la varianza di X. Infatti

$$\begin{aligned} \frac{dc(t)}{dt} &= \frac{m'(t)}{m(t)} \\ \frac{d^2c(t)}{dt^2} &= \frac{m''(t)m(t) - (m'(t))^2}{m^2(t)} \end{aligned} \tag{1.22}$$

dove

$$\frac{dm(t)}{dt} = m'(t), \qquad \frac{d^2m(t)}{dt^2} = m''(t),$$

ed essendo $m(0) = 1$, si ha

$$\left.\frac{dc(t)}{dt}\right|_{t=0} = \frac{m'(0)}{m(0)} = m_1$$
$$\left.\frac{d^2c(t)}{dt^2}\right|_{t=0} = \frac{m''(0)m(0) - (m'(0))^2}{m^2(0)} = m_2 - m_1^2.$$

Una terza funzione generatrice è la *funzione generatrice per i momenti fattoriali.* Si ricorda che, per una variabile aleatoria X, il **momento fattoriale di ordine r** di X (r è un numero intero positivo) è definito come

$$E(X(X-1)\cdots(X-r+1)).$$

Definizione 1.15 (Funzione generatrice dei momenti fattoriali)**.** Sia X una variabile aleatoria. La funzione generatrice dei momenti fattoriali è definita come

$$\psi(t) = E(t^X), \quad t \in \mathbb{R} \tag{1.23}$$

posto che tale valore atteso esista.

La funzione generatrice dei momenti fattoriali viene utilizzata per originare momenti fattoriali in analogia ad $E(e^{tX})$ per i momenti semplici, e si rivela particolarmente utile nel caso di variabile aleatorie discrete a valori interi non negativi.

Sviluppando $\psi(t)$ in serie di Taylor nell'intorno di $t = 1$ si ottiene

$$\begin{aligned}\psi(t) &= E(1 + (t-1)X + \frac{(t-1)^2}{2!}X(X-1) + \frac{(t-1)^3}{3!}X(X-1)(X-2) + \dots \\ &= 1 + (t-1)E(X) + \frac{(t-1)^2}{2!}E(X(X-1)) \\ &\quad + \frac{(t-1)^3}{3!}E(X(X-1)(X-2)) + \dots\end{aligned}$$

da cui si ricava

$$
\begin{aligned}
\left.\frac{d\psi(t)}{dt}\right|_{t=1} &= E(X) \\
\left.\frac{d^2\psi(t)}{dt^2}\right|_{t=1} &= E(X(X-1)) = E(X^2) - E(X) \\
\left.\frac{d^3\psi(t)}{dt^3}\right|_{t=1} &= E(X(X-1)(X-2)) = E(X^3) - 3E(X^2) + 2E(X)
\end{aligned}
$$

e così via. Ora, si assuma X discreta con funzione di probabilità $p(x)$, per $x = 0, 1, \dots, n$. Allora

$$
\begin{aligned}
\psi(t) = E(t^X) &= \sum_{x=0}^{n} t^x p(x) \\
&= p(0) + tp(1) + t^2 p(2) + \cdots + t^n p(n)
\end{aligned}
$$

e quindi

$$
\left.\frac{d^k\psi(t)}{dt^k}\right|_{t=0} = k!p(k), \qquad k = 1, 2, \dots, n,
$$

ovvero le derivate di $\psi(t)$ valutate in $t = 0$ sono uguali ad una costante nota che moltiplica la funzione di probabilità di X.

La funzione generatrice dei momenti di una distribuzione può non esistere. Un'altra funzione che genera i momenti in modo simile e che invece esiste sempre è la *funzione caratteristica*.

Definizione 1.16 (Funzione caratteristica)**.** Data una variabile aleatoria X e la variabile reale t, si chiama funzione caratteristica di X il valore atteso di e^{itX}, dove i è l'unità immaginaria, $i = \sqrt{-1}$. In simboli,

$$
\begin{aligned}
\phi(t) = E(e^{itX}) &= \int_{-\infty}^{\infty} e^{itx} f(x)dx \quad \text{se X è continua,} \\
&= \sum_{x \in R_x} e^{itx} p(x) \qquad \text{se X è discreta.}
\end{aligned} \tag{1.24}
$$

Si osservi che si può scrivere

$$
e^{itX} = \cos(tx) + i \sin(tx)
$$

da cui

$$\phi(t) = E(e^{itX}) = E(\cos(tX)) + iE(\sin(tX)).$$

La funzione caratteristica esiste sempre in quanto $E(|e^{itX}|) = 1$, essendo $\left|e^{itX}\right|^2 = (\cos(tx))^2 + (\sin(tx))^2 = 1$.

Inoltre due distribuzioni di probabilità non condividono mai la stessa funzione caratteristica, a meno che non coincidano. Più precisamente, si ha che

a. due funzioni caratteristiche che sono uguali ovunque, eccetto che in un insieme numerabile di punti, sono funzioni caratteristiche della stessa distribuzione;

b. due distribuzioni che sono uguali ovunque, eccetto che in un insieme numerabile di punti, danno luogo alla stessa funzione caratteristica.

Sia $Y = aX + b$, con a e b costanti, e sia $\phi_X(t)$ la funzione caratteristica di X. Allora Y ha funzione caratteristica data da

$$\phi_Y(t) = E(e^{it(aX+b)}) = e^{itb}E(e^{i(taX)}) = e^{itb}\phi_X(at). \tag{1.25}$$

La funzione caratteristica può essere anche usata per trovare i momenti di una variabile casuale. A condizione che il momento r-esimo $E(X^r)$ esista, la funzione caratteristica può essere derivata r volte e si ha

$$E(X^r) = i^{-r}\frac{d^r\phi(t)}{dt^r}\bigg|_{t=0}. \tag{1.26}$$

Capitolo 2

Alcuni modelli statistici parametrici unidimensionali

Nelle pagine che seguono saranno introdotte alcune notevoli famiglie parametriche di distribuzioni di probabilità unidimensionali che hanno particolare rilevanza applicativa. Una *famiglia parametrica* di funzioni di densità è una collezione di funzioni di densità indicizzata da una quantità chiamata *parametro*. Ad esempio si consideri, per ogni $\lambda > 0$, la funzione di densità di probabilità $f(\cdot;\lambda)$. λ è il parametro e, al variare dei valori di λ, la collezione $\{f(\cdot;\lambda) : \lambda > 0\}$ è una famiglia parametrica di funzioni di densità. Nel capitolo verranno discussi alcuni modelli statistici nel discreto e nel continuo, riportando degli esempi di esperimenti casuali per i quali la famiglia parametrica introdotta può servire da modello realistico. Per la maggior parte dei modelli trattati saranno inoltre ricavate media, varianza e funzione generatrice dei momenti.

2.1 Modelli di distribuzione discreti

2.1.1 Variabili aleatorie Bernoulliane

Si consideri l'esperimento casuale più semplice possibile, ovvero quello che risulta in solo uno di due possibili esiti. Possiamo pensare a molti esempi di questo tipo di esperimento: il lancio di una moneta (testa o croce), l'esito di un esame (superato o fallito), il sesso di un neonato (maschio o femmina). Chiameremo un esperimento i cui esiti sono due eventi incompatibili A e il suo complemento $\overline{A}$ una *prova Bernoulliana*, associando il termine “successo” (s) all'evento A e “insuccesso” (f) all'evento $\overline{A}$. Inoltre sia p la probabilità di A e quindi $1-p$ la probabilità di $\overline{A}$.

Definizione 2.1 (Variabile aleatoria Bernoulliana)**.** Sia S lo spazio campionario di un esperimento, $A \subset S$ un evento per cui $p = P(A), 0 < p < 1$, e si definisca

$$\begin{aligned} X(w) &= 1 \quad \text{se} \quad w \in A \\ &= 0 \quad \text{se} \quad w \in \overline{A}. \end{aligned} \tag{2.1}$$

X è detta *variabile aleatoria Bernoulliana* di parametro p. Se l'esperimento è una prova Bernoulliana, allora $A = \{s\}$.

Una variabile aleatoria definita come in (2.1) è anche detta *variabile indicatrice* (dell'evento A) o variabile *binaria*.

La variabile casuale costruita in questo modo è una variabile aleatoria discreta, in quanto assume solo i valori 0 e 1 con la probabilità $P(X=1)=p$ e $P(X=0)=1-p$. Quindi la funzione di probabilità per una variabile aleatoria Bernoulliana X si scrive semplicemente come

$$p(k) = p^k(1-p)^{1-k}, \quad k = 0, 1. \tag{2.2}$$

Teorema 2.1. Se X ha una distribuzione di Bernoulli (2.2), allora

$$E(X) = p, \quad V(X) = p(1-p), \quad m(t) = 1 - p + pe^t. \tag{2.3}$$

Dimostrazione. Posto $q := 1-p$, il valore atteso di X è dato da

$$E(X) = 0 \cdot q + 1 \cdot p = p$$

Infatti, il momento di ordine r di X è

$$E(X^r) = 0^r \cdot q + 1^r \cdot p = p, \quad r = 1, 2, 3, \ldots$$

da cui segue

$$V(X) = E(X^2) - [E(X)]^2 = p - p^2 = pq.$$

La funzione generatrice dei momenti di X si ottiene come

$$m(t) = E(e^{tX}) = e^0 \cdot q + e^t \cdot p = q + pe^t.$$

□

Possiamo inoltre facilmente ricavare la funzione generatrice dei momenti fattoriali. Posto $q := 1 - p$, si ha

$$\psi(t) = E(t^X) = t^0 \cdot q + t^1 \cdot p = q + pt \tag{2.4}$$

La distribuzione di Bernoulli dipende unicamente dalla quantità p, e media, varianza, momenti e funzioni generatrici sono tutte espresse in funzione di tale parametro. Come si vedrà di seguito, le variabili aleatorie Bernoulliane sono alla base della costruzione di altri modelli discreti, tra cui la *distribuzione binomiale*. Il paragrafo che segue introduce tale modello, definendo la variabile aleatoria binomiale.

2.1.2 Variabili aleatorie binomiali

La variabile casuale binomiale si costruisce a partire dalla nozione di *esperimento casuale binomiale*, che consiste in un insieme di n prove ripetute con le seguenti caratteristiche:

i. ad ogni singola prova si hanno solo due esiti possibili, chiamati "successo" (s) ed "insuccesso" (f);

ii. la probabilità dell'evento che dà origine al "successo" è costante (si denoti con p);

iii. i risultati delle prove sono **indipendenti**.

Chiaramente la variabile aleatoria che descrive ogni singola prova ha distribuzione Bernoulliana di parametro p. L'indipendenza delle prove Bernoulliane deriva dal fatto che le prove sono "eventi indipendenti", ovvero l'esito di una prova non ha alcuna influenza sull'esito di ogni altra prova. Più precisamente, se A è un evento il cui verificarsi o non verificarsi dipende da un sottoinsieme delle prove e B è un evento il cui verificarsi o non verificarsi dipende da un sottoinsieme disgiunto delle prove, allora A e B sono indipendenti e vale $P(A \cap B) = P(A)P(B)$.

Lo spazio campionario naturale per un esperimento binomiale è costituito dal prodotto Cartesiano di n spazi campionari identici $S_j = \{s, f\}, j = 1, 2, \ldots, n$,

$$S = S_1 \times S_2 \times \ldots S_n.$$

Ciò significa che S è composto da tutte le n-uple $(x_1, x_2, \ldots, x_n)$ con $x_j = s$ o $x_j = f$, che rappresentano il successo o il fallimento nella j-ma prova, $j = 1, 2, \ldots, n$. Per ogni prova Bernoulliana si ha

$$P_j(\{s\}) = p, \quad P_j(\{f\}) = 1 - p = q$$

e poiché le prove sono indipendenti, possiamo calcolare la probabilità degli eventi rappresentati dalle singole n-uple moltiplicando le probabilità corrispondenti ai singoli eventi dello spazio campionario generato dall'esperimento. Quindi, ad esempio, all'evento elementare

$$A = \{(s, s, \ldots, s)\}$$

è associata la probabilità $p \cdot p \cdots p = p^n$, mentre all'evento costituito da una sequenza di n insuccessi

$$B = \{(f, f, \ldots, f)\}$$

è associata la probabilità $q \cdot q \cdots q = q^n$ e, analogamente, all'evento in cui il primo elemento rappresenta un successo e tutti gli altri un fallimento si assegna la probabilità $p \cdot q \cdots q = pq^{n-1}$, e così via. Si osservi che il medesimo livello di probabilità è associato a tutti gli eventi dello spazio campionario che presentano lo stesso numero di successi (e quindi di insuccessi), indipendentemente dall'ordine secondo cui questi si manifestano.

Definizione 2.2 (Variabile aleatoria binomiale)**.** Sia X il numero di successi ottenuti in un esperimento binomiale con n prove e parametro p. Allora X è detta *variabile aleatoria binomiale* di parametri n e p.

Poiché la variabile aleatoria binomiale conta il numero di successi in n prove di Bernoulli indipendenti e ripetute, essa è discreta ed assume i valori in $R_X = \{0, 1, 2, \ldots, n\}$. Vi è solo un elemento di S per cui $X = 0$ (l'"insuccesso" in ogni prova) e quindi $P(X = 0) = q^n$. Analogamente, vi è solo un elemento di S per cui $X = n$ (il "successo" in ogni prova) e si ha $P(X = n) = p^n$. Gli altri valori che X può assumere (gli interi da 1 a $n-1$) sono assegnati a più di un elemento in S, e pertanto la probabilità $P(X = x)$ sarà uguale alla somma delle probabilità degli eventi elementari composti da x "successi" ed $n - x$ "insuccessi". Tali ordinamenti hanno la stessa probabilità $p^x q^{n-x}$ ed il loro numero corrisponde al numero dei modi in cui possiamo scegliere x posizioni per i "successi" nelle n-uple (essendo le altre $n - x$ posizioni occupate dagli "insuccessi")

$$\frac{n!}{x!(n-x)!} = \binom{n}{x}.$$

Segue immediatamente che

$$P(X = x) = \binom{n}{x} p^x q^{n-x}, \quad x = 1, 2, \ldots, n-1.$$

Ricordando che $\binom{n}{0} = \binom{n}{n} = 1$, si ha il seguente teorema.

Teorema 2.2. Se X è una variabile aleatoria binomiale di parametri n e p la sua funzione di probabilità è espressa da

$$p(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \ldots, n, \tag{2.5}$$

Nel seguito, per indicare sinteticamente il modello, si userà la simbologia $X \sim Bin(n, p)$, che si spiega con "X è distribuita secondo una binomiale con parametri n e p".

La formula (2.5) è il generico termine dello sviluppo del binomio $[p + (1 - p)]^n$. Si può scrivere infatti

$$(p + (1-p))^n = \binom{n}{0}(1-p)^n + \binom{n}{1} p(1-p)^{n-1} + \cdots + \binom{n}{x} p^x (1-p)^{n-x} + \cdots + \binom{n}{n} p^n,$$

da cui si ha conferma della relazione $\sum_{x=0}^{n} p(x) = 1$, che deve valere per ogni funzione di probabilità. Si osservi inoltre che $\binom{n}{x} = \binom{n}{n-x}$ e quindi, nel

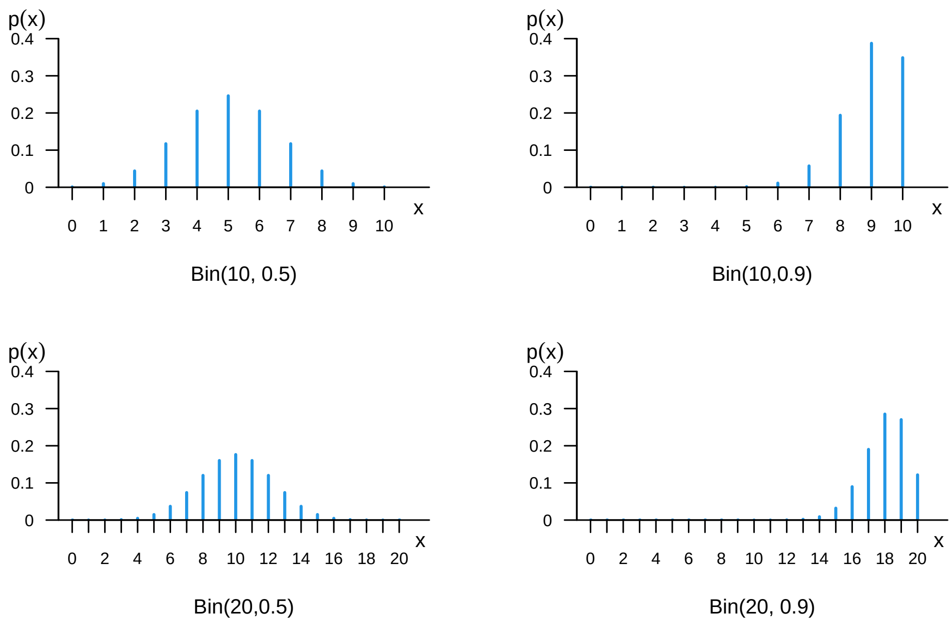


Figura 2.1: Funzione di probabilità binomiale per diversi valori di n e p.

caso particolare in cui $p = 1 - p = 1/2$, vale l'uguaglianza

$$p(x) = \binom{n}{x}\frac{1}{2^n} = \binom{n}{n-x}\frac{1}{2^n}, \quad x = 0, 1, \ldots, n.$$

Ne deriva che 0 ha la stessa probabilità di n, 1 la stessa probabilità di $n-1$, e così via. Nel caso in esame la distribuzione binomiale è *simmetrica.* in generale, i termini $p(x)$ prima aumentano monotonamente e poi diminuiscono monotonamente, come afferma il seguente teorema.

Teorema 2.3. Sia $p(x)$ la densità della distribuzione binomiale di X. Allora $p(x-1) < p(x)$ per $x < (n+1)p$; $p(x-1) > p(x)$ per $x > (n+1)p$, e $p(x-1) = p(x)$ se $x = (n+1)p$ e $(n+1)p$ è un intero, con x che varia da $1, \ldots, n$.

Dimostrazione.

$$\frac{p(x)}{p(x-1)} = \frac{n-x+1}{x}\frac{p}{q} = 1 + \frac{(n+1)p - x}{xq},$$

che è maggiore di 1 se $x < (n+1)p$, minore di 1 se $x > (n+1)p$ e uguale a 1 se l'intero $x = (n+1)p$. $\square$

Il seguente teorema fornisce valore atteso, varianza e funzione generatrice dei momenti della distribuzione binomiale.

Teorema 2.4. Se X ha una distribuzione binomiale di parametri n e p, allora

$$E(X) = np, \quad V(X) = np(1-p), \quad m(t) = (1 - p + pe^t)^n. \tag{2.6}$$

Dimostrazione. Posto $q = 1 - p$, il valore atteso di una variabile aleatoria binomiale è data da

$$\begin{aligned}
E(X) &= \sum_{x=0}^{n} x \binom{n}{x} p^x q^{n-x} = \sum_{x=1}^{n} x \frac{n!}{x!(n-x)!} p^x q^{n-x} \\
&= np \sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1} q^{n-x} \\
&= np \sum_{j=0}^{n-1} \binom{n-1}{j} p^j q^{n-1-j}, \quad j = x - 1 \\
&= np(p+q)^{n-1} = np.
\end{aligned}$$

dove è stato usato il risultato $(x+y)^k = \sum_{i=0}^{k} \binom{k}{i} x^i y^{k-i}$, per x e y numeri reali e k intero positivo.

La varianza si può ottenere con procedimento analogo:

$$\begin{aligned}
V(X) &= E(X^2) - [E(X)]^2 = \sum_{x=0}^{n} x^2 \binom{n}{x} p^x q^{n-x} - (np)^2 \\
&= \sum_{x=0}^{n} (x(x-1) + x) \frac{n!}{x!(n-x)!} p^x q^{n-x} - n^2 p^2 \\
&= \sum_{x=2}^{n} x(x-1) \frac{n!}{x!(n-x)!} p^x q^{n-x} + \sum_{x=1}^{n} x \frac{n!}{x!(n-x)!} p^x q^{n-x} - n^2 p^2 \\
&= \sum_{x=2}^{n} x(x-1) \frac{n(n-1)(n-2)!}{x(x-1)(x-2)!(n-x)!} p^x q^{n-x} + np - n^2 p^2 \\
&= n(n-1)p^2 \sum_{x=2}^{n} \frac{(n-2)!}{(x-2)!(n-x)!} p^{x-2} q^{n-x} + np - n^2 p^2 \\
&= n(n-1)p^2 \sum_{j=0}^{n-2} \binom{n-2}{j} p^j q^{n-2-j} + np - n^2 p^2, \quad j = x - 2 \\
&= n(n-1)p^2 (p+q)^{n-2} + np - n^2 p^2 = n^2 p^2 - np^2 + np - n^2 p^2 = np(1-p).
\end{aligned}$$

Infine è facile verificare che la funzione generatrice dei momenti binomiale è data da

$$
\begin{aligned}
m(t) = E(e^{tX}) &= \sum_{x=0}^{n} e^{tx}\binom{n}{x}p^x q^{n-x} \\
&= \sum_{x=0}^{n}\binom{n}{x}(pe^t)^x q^{n-x} \\
&= (pe^t + q)^n. \qquad \square
\end{aligned}
$$

Si osservi che la media e la varianza di X si possono ottenere facilmente considerando che

$$m'(t) = n(pe^t + q)^{n-1} \cdot pe^t$$

$$m''(t) = n(n-1)(pe^t)^2(pe^t + q)^{n-2} + npe^t(pe^t + q)^{n-1}$$

da cui

$$
\begin{aligned}
E(X) &= m'(0) = np \\
V(X) &= E(X^2) - [E(X)]^2 = m''(0) - (np)^2 \\
&= n(n-1)p^2 + np - (np)^2 = npq.
\end{aligned}
$$

Analogamente, se si considera la funzione generatrice dei momenti fattoriali

$$\psi(t) = E(t^X) = m(\ln t) = (pe^{\ln t} + q)^n = (pt + q)^n \tag{2.7}$$

da cui

$$\psi'(t) = np(pt + q)^{n-1}$$

$$\psi'(1) = E(X) = np$$

$$\psi''(t) = n(n-1)p^2(pt + q)^{n-2}$$

$$\psi''(1) = n(n-1)p^2 = E(X^2) - E(X)$$

e quindi

$$E(X^2) = n(n-1)p^2 + E(X) = n^2p^2 - np^2 + np$$

$$V(X) = E(X^2) - [E(X)]^2 = n^2p^2 - np^2 + np - n^2p^2 = npq.$$

È importante notare che una variabile aleatoria binomiale con parametri n e p è ottenuta come la somma di n variabili casuali Bernoulliane. Si supponga di voler eseguire n prove Bernoulliane indipendenti, ciascuna con probabilità p. Definiamo quindi

$$\begin{aligned} Y_i &= 1 \quad \text{se osserviamo un successo nella } i\text{-ma prova} \\ &= 0 \quad \text{altrimenti.} \end{aligned} \tag{2.8}$$

per $i = 1, 2, 3, \ldots, n$. Allora $\sum_{i=1}^{n} Y_i$ corrisponde al numero totale di successi nelle n prove; ne segue che $X = \sum_{i=1}^{n} Y_i$ è la variabile aleatoria binomiale con parametri n e p ed essendo $E(Y_i) = p, V(Y_i) = pq$, si ha $E(X) = \sum_{i=1}^{n} E(Y_i) = np$ e $V(X) = \sum_{i=1}^{n} V(Y_i) = npq$.

Esempio 2.1.

Consideriamo il campionamento con reimmissione da un'urna contenente M palline, K delle quali sono difettose. Sia X il numero di palline difettose in un campione di ampiezza n. Le singole estrazioni sono prove di Bernoulli dove il "successo" è identificato con "difettoso" e il campionamento con reimmissione consiste in n prove di Bernoulli indipendenti con probabilità di successo $p = K/M$. Ne segue che X ha distribuzione binomiale data da

$$\binom{n}{x} \left(\frac{K}{M}\right)^x \left(1 - \frac{K}{M}\right)^{n-x}, \quad x = 0, 1, \ldots, n.$$

2.1.3 Variabili aleatorie geometriche e binomiali negative

Le distribuzioni geometriche e binomiali negative rappresentano due famiglie di distribuzioni discrete molto importanti in statistica. Come per la distribuzione binomiale, anche queste posso essere definite a partire da una successione di prove Bernoulliane indipendenti. Vedremo inoltre che la distribuzione geometrica rappresenta un caso particolare della distribuzione binomiale negativa.

Si supponga di effettuare delle prove Bernoulliane indipendenti, in cui la probabilità dell'evento "successo" sia costante e pari a p, e sia X il numero di prove necessarie per ottenere il primo "successo". Ad esempio, si pensi di lanciare una moneta ripetutamente finché non si presenti l'evento "testa". Lo spazio campionario generato dall'esperimento è

$$\Omega = \{T, CT, CCT, CCCT, \ldots\}$$

Se si associa a ciascuno degli eventi elementari di Ω il numero di prove necessarie per ottenere per la prima volta l'evento "testa", si definisce una variabile aleatoria discreta che assume i valori $x = 1, 2, 3, \dots$.

Allora $P(X = 1) = p$, poiché la probabilità di ottenere un successo è pari a p in ogni prova. Osserveremo $X = 2$ se e solo se si verifica l'evento "insuccesso" nella prima prova e un "successo" nella seconda prova, e quindi $P(X = 2) = p(1 - p)$. Analogamente, per ogni $k \geq 3$, osserveremo $X = k$ se e solo se si registrano $k - 1$ "insuccessi" seguiti da un "successo" nella k-ma prova, ovvero $P(X = k) = p(1 - p)^{k-1}$. Il numero aleatorio delle prove necessarie affinché si presenti una prima volta l'evento "successo" con probabilità costante e pari a p si chiama *variabile geometrica* di parametro p, poiché i valori della sua funzione di probabilità formano una progressione geometrica.

Definizione 2.3 (Variabile aleatoria geometrica)**.** Si consideri una successione di prove Bernoulliane indipendenti, ciascuna con probabilità di "successo" costante e pari a p, $p > 0$. Se X è il numero delle prove necessarie affinché si presenti una prima volta l'evento "successo", allora X è detta *variabile aleatoria geometrica* di parametro p.

Si può quindi enunciare il seguente risultato.

Teorema 2.5. Se X è una variabile aleatoria geometrica di parametro p la sua funzione di probabilità è espressa da

$$p(x) = p(1 - p)^{x-1}, \quad x = 1, 2, 3 \dots, \tag{2.9}$$

Vale la relazione

$$\begin{aligned} \sum_{x=1}^{\infty} p(x) &= p \sum_{x=1}^{\infty} (1 - p)^{x-1} = p[1 + (1 - p) + (1 - p)^2 + \dots] \\ &= p \left[\frac{1}{1 - (1 - p)} \right] = 1 \end{aligned}$$

dove si è applicata la proprietà di convergenza della serie geometrica $\sum_{i=0}^{\infty} r^i = \frac{1}{1-r}$, $|r| < 1$.
Nel seguito verrà adottata la notazione $X \sim Ge(p)$ per indicare una v.a. distribuita secondo una legge geometrica di parametro p.

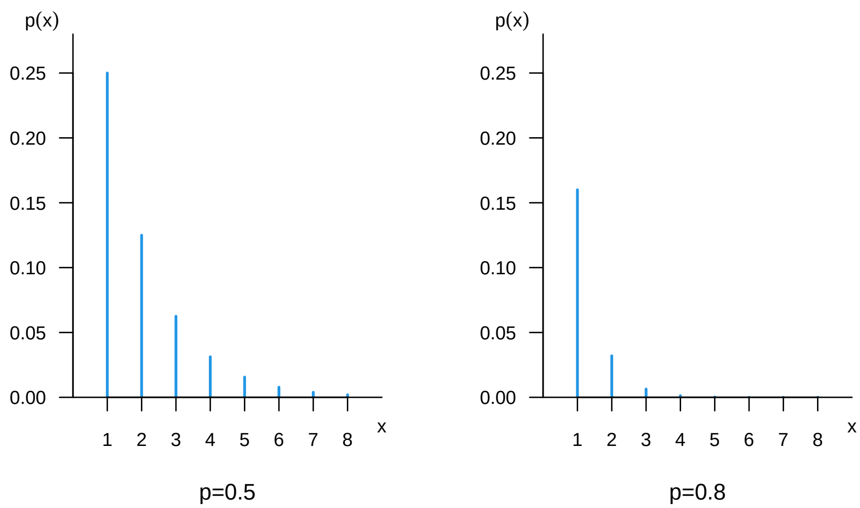


Figura 2.2: Funzione di probabilità geometrica di parametro $p = 0.5, 0.8$.

Teorema 2.6. Se X è una v.a. geometrica di parametro p, allora

$$E(X) = \frac{1}{p}, \quad V(X) = \frac{1-p}{p^2}, \quad m(t) = \frac{pe^t}{1-(1-p)e^t}. \tag{2.10}$$

Dimostrazione. Posto $q := 1 - p$, il valore atteso si può scrivere come

$$\begin{aligned} E(X) &= \sum_{x=1}^{\infty} xpq^{x-1} \\ &= p\sum_{k=0}^{\infty}(k+1)q^k, \quad k = x-1 \\ &= p\frac{1}{(1-q)^2} = \frac{1}{p} \end{aligned}$$

dove si è usato il risultato $\sum_{i=0}^{\infty}(i+1)r^i = \frac{1}{(1-r)^2}$.

Si ha inoltre

$$
\begin{aligned}
E(X^2) &= E(X(X-1)) + E(X) \\
&= \sum_{x=1}^{\infty} x(x-1)pq^{x-1} + \frac{1}{p} \\
&= 2pq \sum_{x=2}^{\infty} \frac{x(x-1)(x-2)!}{2!(x-2)!} q^{x-2} + \frac{1}{p} \\
&= 2pq \sum_{x=2}^{\infty} \binom{x}{2} q^{x-2} + \frac{1}{p} \\
&= \frac{2pq}{(1-q)^3} + \frac{1}{p} \\
&= \frac{2q}{p^2} + \frac{1}{p} \\
&= \frac{2}{p^2} - \frac{1}{p}
\end{aligned}
$$

ricordando che $\sum_{j=k}^{\infty} \binom{j}{k} r^{j-k} = \left(\frac{1}{1-r}\right)^{k+1}$, $|r| < 1$. Quindi si ha

$$
\begin{aligned}
V(X) &= E(X^2) - [E(X)]^2 \\
&= \frac{2}{p^2} - \frac{1}{p} - \frac{1}{p^2} \\
&= \frac{q}{p^2}.
\end{aligned}
$$

La funzione generatrice dei momenti si ricava facilmente come segue:

$$
\begin{aligned}
m(t) = E(e^{tX}) &= \sum_{x=1}^{\infty} e^{tx} pq^{x-1} \\
&= pe^t \sum_{x=1}^{\infty} (e^t)^{x-1} q^{x-1} \\
&= pe^t \sum_{x=1}^{\infty} (qe^t)^{x-1} \\
&= \frac{pe^t}{1-qe^t}, \quad (qe^t < 1).
\end{aligned}
$$

□

Con un procedimento simile si trova inoltre che

$$\begin{aligned}\psi(t) = E(t^X) &= \sum_{x=1}^{\infty} t^x pq^{x-1} \\ &= pt\sum_{x=1}^{\infty}(qt)^{x-1} \\ &= \frac{pt}{1-qt}, \quad (|qt| < 1).\end{aligned}$$

La funzione di ripartizione per la variabile aleatoria geometrica X si ottiene dalla funzione di probabilità come

$$\begin{aligned}P(X > t) &= \sum_{k=t+1}^{\infty} pq^{k-1} \\ &= p\sum_{i=0}^{\infty} q^{t+i}, \quad i = k-t-1 \\ &= pq^t\sum_{i=0}^{\infty} q^i = pq^t\frac{1}{1-q} = (1-p)^t, \quad t = 0,1,2,3,\ldots\end{aligned}$$

e quindi

$$F(t) = 1 - P(X > t) = 1 - (1-p)^t, \quad t = 0,1,2,\ldots$$

La distribuzione geometrica è caratterizzata da *assenza di memoria*, una proprietà che non si ritrova in nessun'altra distribuzione discreta. Per comprendere tale proprietà si consideri la probabilità condizionata che una variabile casuale geometrica sia maggiore di $a+b$, dato che è maggiore di a, ovvero $P(X > a+b|X > a)$, con a e b interi positivi. Se si denota con A l'evento $X > a$ e con B l'evento $X > a+b$, allora la probabilità cercata è espressa da

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

Poiché se $X > a+b$ allora deve essere $X > a$, si ha $B \subset A$ e quindi $A \cap B = B$, da cui

$$\begin{aligned} P(B|A) &= \frac{P(B)}{P(A)} \\ &= \frac{P(X > a+b)}{P(X > a)} \\ &= \frac{q^{a+b}}{q^a} = q^b = P(X > b). \end{aligned}$$

Questo risultato ci dice che la probabilità che sia necessario effettuare ancora più di b prove per osservare il primo successo sapendo che in a prove già effettuate l'evento successo non si sia verificato, è uguale alla probabilità non condizionata che siano necessari più di b tentativi prima del primo successo. Cioè, il fatto di avere già osservato a insuccessi consecutivi non cambia la distribuzione del numero di tentativi necessari per ottenere il primo successo.

Ora, si prenda nuovamente in esame una successione di prove di Bernoulli indipendenti con probabilità di successo in una singola prova p, e sia X il numero di tentativi necessari prima di osservare l'r-mo successo. X ha supporto $R_X = \{r, r+1, r+2, \dots\}$ poiché occorre effettuare almeno r prove per osservare r successi. Avremo $X = r$ se e solo se si è ottenuto un successo in ognuno dei primi r tentativi e quindi $P(X = r) = p^r$. Per ogni intero $k > r$, si avrà $X = k$ se e solo se la k-ma prova genera l'r-mo successo e quindi i primi $k-1$ tentativi devono contenere $r-1$ successi. Ne segue che

$$P(X = k) = \binom{k-1}{r-1} p^r q^{k-1-(r-1)}$$

Definizione 2.4 (Variabile aleatoria binomiale negativa)**.** Sia X il numero aleatorio di prove di Bernoulli indipendenti, con probabilità di successo della singola prova pari a p, che occorrono per osservare l'r-mo successo. Allora X è detta una *variabile aleatoria binomiale negativa* (detta anche di *Pascal*) di parametri r e p.

Teorema 2.7. Se X è una variabile aleatoria binomiale negativa di parametri r e p, allora X ha funzione di probabilità

$$p(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}, \quad x = r, r+1, r+2, \ldots. \tag{2.11}$$

$P(X = x) = p(x) \geq 0$ e inoltre si ha

$$\sum_{x=r}^{\infty} \binom{x-1}{r-1} p^r q^{x-r} = p^r \sum_{x=r}^{\infty} \binom{x-1}{r-1} q^{x-r} = p^r (1-q)^{-r} = 1,$$

dove si è posto $q := 1 - p$. Pertanto $p(x)$ è una funzione di probabilità. La (2.11) è detta *distribuzione binomiale negativa* $\mathcal{NB}(r, p)$ per la presenza dell'esponente negativo nello sviluppo binomiale. Si osservi che il caso $r = 1$ nella (2.11) corrisponde alla funzione di densità della v.a. geometrica. La funzione generatrice dei momenti fattoriali della variabile X è

$$\begin{aligned}
\psi(t) = E(t^X) &= \sum_{x=r}^{\infty} t^x \binom{x-1}{r-1} p^r q^{x-r} \\
&= (pt)^r \sum_{x=r}^{\infty} \binom{x-1}{r-1} (qt)^{x-r} \\
&= \left(\frac{pt}{1-qt} \right)^r, \quad (|qt| < 1).
\end{aligned}$$

da cui si ricava facilmente la media

$$E(X) = \frac{r}{p}$$

e la varianza

$$V(X) = \frac{rq}{p^2}.$$

In modo simile si ottiene la funzione generatrice dei momenti per X. In definitiva, si ha il seguente risultato.

Teorema 2.8. Se X è una v.a. binomiale negativa di parametri r e p, allora

$$E(X) = \frac{r}{p}, \quad V(X) = \frac{r(1-p)}{p^2}, \quad m(t) = \left(\frac{pe^t}{1-(1-p)e^t}\right)^r. \tag{2.12}$$

Una interpretazione alternativa deriva dal considerare la v.a. $X \sim \mathcal{NB}(r,p)$ come somma di r variabili aleatorie indipendenti

$$X = X_1 + X_2 + \cdots + X_r,$$

dove X_1 è il numero di prove necessarie per il primo successo, X_2 è il numero di prove occorrenti affinché essendosi già verificato l'evento successo per la prima volta, esso si verifichi per la seconda volta, e così via. Quindi X_i ha, per ogni i, funzione di probabilità

$$p(x) = p(1-p)^{x-1}, \quad x = 1, 2, \dots,$$

per cui la v.a. $X \sim \mathcal{NB}(r,p)$ può essere interpretata come la somma di r variabili aleatorie indipendenti e identicamente distribuite $X_i \sim Ge(p)$.

È interessante considerare la variabile aleatoria che descrive il numero totale di insuccessi che *precedono* l'r-mo successo in una successione di prove Bernoulliane indipendenti. Se Y è la v.a. *tempo d'attesa* (in termini di numero di insuccessi) per avere il successo r-mo, allora $Y = X - r$ e

$$P(Y = y) = \binom{y+r-1}{y} p^r q^y = \binom{-r}{y} p^r (-q)^y, \quad y = 0, 1, 2, \dots \tag{2.13}$$

La seconda formulazione segue da

$$\begin{aligned}
\binom{-r}{y} &= \frac{(-r)!}{y!(-r-y)!} \\
&= \frac{(-r)(-r-1)\cdots(-r-y+1)}{y!} \\
&= (-1)^y \frac{r(r+1)\cdots(r+y-1)}{y!} \\
&= (-1)^y \binom{y+r-1}{y}.
\end{aligned}$$

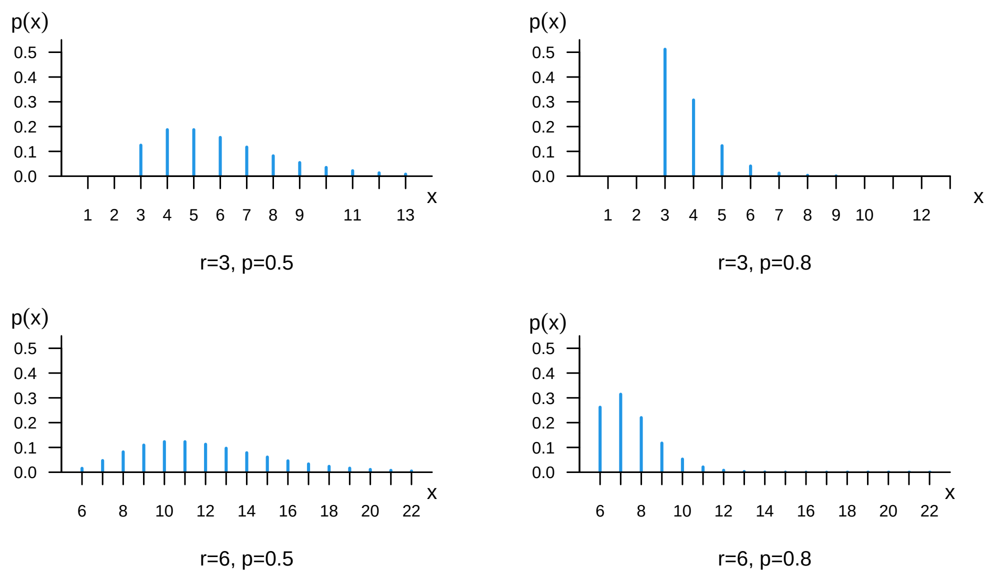


Figura 2.3: Funzione di probabilità binomiale negativa definita in (2.11) per diversi valori di r e p.

Y è chiaramente ancora una distribuzione binomiale negativa con la stessa varianza di X ma media diversa:

$$E(Y) = E(X - r) = \frac{rq}{p}$$

Teorema 2.9. Se Y ha distribuzione binomiale negativa espressa da (2.13), allora

$$E(Y) = \frac{rq}{p}, \quad V(Y) = \frac{rq}{p^2}, \quad m_Y(t) = \left(\frac{p}{1 - qe^t}\right)^r. \tag{2.14}$$

Dimostrazione.

$$\begin{aligned} m_Y(t) = E(e^{tY}) &= \sum_{y=0}^{\infty} e^{ty} \binom{-r}{y} p^r (-q)^y \\ &= p^r \sum_{y=0}^{\infty} \binom{-r}{y} (-qe^t)^y \\ &= p^r (1 - qe^t)^{-r} = \left(\frac{p}{1 - qe^t}\right)^r, \quad qe^t < 1. \end{aligned}$$

Allora

$$m'_Y(t) = p^r(-r)(1 - qe^t)^{-r-1}(-qe^t)$$

e

$$m''_Y(t) = rqp^r(q(r+1)e^{2t}(1 - qe^t)^{-r-2} + e^t(1 - qe^t)^{-r-1})$$

quindi

$$\begin{aligned} E(Y) &= m'_Y(0) = \frac{rq}{p} \\ V(Y) &= m''(0) - \left(\frac{rq}{p}\right)^2 \\ &= \frac{rq^2}{p^2} + \frac{rq}{p} = \frac{rq}{p^2} \end{aligned}$$

$\square$

Infine si osservi che dalla (2.13) per $r = 1$ si ottiene la distribuzione geometrica con densità nella forma

$$p(y) = p(1-p)^y, \quad y = 0, 1, 2, \ldots$$

dove la variabile aleatoria Y è intesa come numero aleatorio di insuccessi prima di ottenere un successo.

La distribuzione binomiale negativa ha i numeri interi non negativi come punti di massa e perciò si presta ad essere un modello per un esperimento casuale nel quale si debbano fare dei conteggi di un qualche tipo, ad esempio, nelle statistiche riguardanti gli incidenti e la salute.

Esempio 2.2.

Si supponga che un gruppo di ricerca medica voglia studiare l'effetto di un trattamento su una certa condizione, presente nel 10% della popolazione. In particolare, si vogliono individuare esattamente $r = 10$ individui che presentano la condizione oggetto di studio. Sia Y il numero di persone intervistate per ottenere un tale risultato. Allora assunto che Y abbia distribuzione binomiale negativa con $r = 10, p = 0.1$, il numero atteso di persone da intervistare è $10/0.1 = 100$, la varianza è $(10 \cdot 0.9)/(0.1)^2 = 900$, e la deviazione standard è $\sqrt{900} = 30$.

2.1.4 Variabile aleatoria di Poisson

Il modello probabilistico che viene introdotto in questa sezione è comune nelle applicazioni che riguardano fenomeni che evolvono nel tempo o nello spazio e che implicano conteggi delle realizzazioni di un evento aleatorio. Ad esempio, nel pianificare il numero di linee telefoniche o il tipo di attrezzature necessarie ad una certa attività occorre valutare il numero di telefonate entranti e uscenti in un certo periodo di tempo. Si pensi anche al numero di clienti che si presentano alla cassa di un negozio in un dato intervallo di tempo, il numero di difetti in un tratto di lunghezza fissa di un cavo ottico, o il numero di particelle radioattive emesse per unità di tempo. In ognuno di questi esempi siamo in grado di contare il numero di volte che un dato evento si verifica. Come possiamo modellare questi fenomeni? La distribuzione di Poisson fornisce un modello realistico per molti di questi fenomeni casuali, se sono soddisfatte particolari ipotesi che verranno discusse nel seguito.

Assumiamo di osservare le manifestazioni di un dato fenomeno in un certo intervallo di tempo; l'istante iniziale venga denotato con 0, che rappresenta quindi l'origine della nostra scala temporale. Assumiamo, per il momento, di osservare il fenomeno per un intervallo di tempo di lunghezza fissa t. Il numero di realizzazioni (arrivi) nell'intervallo $(0, t]$ è una variabile aleatoria X. Chiaramente X è discreta, essendo il numero di eventi che si verificano nell'intervallo. La funzione di probabilità di X dipende dal modo in cui gli eventi si verificano, sul quale si formuliamo le seguenti assunzioni:

(i) In un intervallo di tempo sufficientemente breve di lunghezza Δt, si possono verificare 0 eventi o 1 evento (non sono possibili due o più realizzazioni simultanee).

(ii) La probabilità che esattamente un evento si verifichi nell'intervallo Δt è proporzionale alla lunghezza dell'intervallo ed è pari a $\nu \Delta t$.

(iii) Eventi in intervalli di lunghezza Δt non sovrapposti sono indipendenti. Il verificarsi di eventi in intervalli disgiunti costituiscono quindi prove bernoulliane indipendenti.

Sotto queste condizioni, dette assunzioni per un **processo di Poisson** di parametro ν, possiamo suddividere il nostro intervallo di lunghezza t in $n = t/\Delta t$ piccoli intervalli della stessa ampiezza e considerare come successo il realizzarsi dell'evento aleatorio nel singolo intervallino. Sia p la probabilità, costante, dell'evento successo in ciascuno degli n intervalli. Si può ora concepire il numero di realizzazioni nell'intervallo di ampiezza t come il risultato di n prove Bernoulliane indipendenti, ciascuna con probabilità di successo $p = \nu \Delta t$. In altri termini, la variabile X numero di realizzazioni dell'evento aleatorio nell'intervallo t è equiparabile alla variabile casuale binomiale "numero di successi in n prove Bernoulliane indipendenti", avente media np, $p = \nu \Delta t = \nu t/n$. Quindi si ha

$$\begin{aligned} p(x) &= \binom{n}{x} (\nu \Delta t)^x (1 - \nu \Delta t)^{n-x} \\ &= \binom{n}{x} \left(\frac{\nu t}{n}\right)^x \left(1 - \frac{\nu t}{n}\right)^{n-x} \end{aligned}$$

Considerando ora il limite della funzione di probabilità per $\Delta t \to 0$, e quindi per n che tende all'infinito, si ottiene la funzione di probabilità del modello di Poisson, che fornisce la probabilità di esattamente x arrivi nell'intervallo continuo di lunghezza t, assumendo valide le assunzioni (i), (ii), (iii), sopra elencate. Si può scrivere

$$\begin{aligned}
p(x) &= \binom{n}{x}\left(\frac{\nu t}{n}\right)^x\left(1-\frac{\nu t}{n}\right)^{n-x} \\
&= \frac{n!}{x!(n-x)!}\frac{(\nu t)^x}{n^x}\left(1-\frac{\nu t}{n}\right)^n\left(1-\frac{\nu t}{n}\right)^{-x} \\
&= \frac{(\nu t)^x}{x!}\left(1-\frac{\nu t}{n}\right)^n\left(1-\frac{\nu t}{n}\right)^{-x}\frac{n(n-1)\cdots(n-x+1)(n-x)!}{n^x(n-x)!} \\
&= \frac{(\nu t)^x}{x!}\left(1-\frac{\nu t}{n}\right)^n\left(1-\frac{\nu t}{n}\right)^{-x}\frac{n(n-1)\cdots(n-x+1)}{n^x}
\end{aligned}$$

Al limite per $n \to \infty$,

$\left(1-\frac{\nu t}{n}\right)^n \to e^{-\nu t}$,

$\left(1-\frac{\nu t}{n}\right)^{-x} \to 1$,

$$\frac{n(n-1)\cdots(n-x+1)}{n^x} = 1 \cdot \left(\frac{n-1}{n}\right) \cdot \left(\frac{n-2}{n}\right) \cdots \cdot \left(\frac{n-x+1}{n}\right) \to 1$$

e quindi

$$p(x) \to \frac{(\nu t)^x}{x!}e^{-\nu t} \tag{2.15}$$

per ogni $x = 0, 1, 2, \ldots$.

Definizione 2.5 (Variabile aleatoria di Poisson)**.** Se gli assunti (i), (ii), (iii) sono soddisfatti, il numero X di eventi in un intervallo di tempo t è detta *variabile aleatoria di Poisson* di parametro $\lambda = \nu t$, dove ν è il numero medio di arrivi nell'unità di tempo.

Abbiamo dimostrato il seguente risultato.

Teorema 2.10. Per una variabile aleatoria di Poisson di parametro λ, la funzione di probabilità è

$$p(x) = \frac{\lambda^x}{x!}e^{-\lambda}, \quad x = 0, 1, 2, \ldots \tag{2.16}$$

Il parametro λ soddisfa $\lambda > 0$. Se X ha funzione di probabilità (2.16), X ha una *distribuzione di Poisson* indicata con la simbologia $X \sim Po(\lambda)$. La (2.16) è chiaramente non negativa e soddisfa la relazione

$$\sum_{x=0}^{\infty} p(x) = \sum_{x=0}^{\infty} \frac{\lambda^x}{x!}e^{-\lambda} = e^{-\lambda}e^{\lambda} = 1,$$

dove si è applicato lo sviluppo in serie di Taylor

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots = \sum_{j=0}^{\infty} \frac{x^j}{j!}.$$

Il risultato in (2.15) prova che il modello Poissoniano è il limite a cui tende la distribuzione binomiale quando, tenendo la media costante e pari al valore np, si fa tendere n ad infinito. Allo stesso tempo, i parametri individuali variano, poiché n aumenta al diminuire di $p = \nu t/n$ e il valore della funzione di probabilità binomiale converge al valore della funzione di probabilità di Poisson di parametro $\lambda = np = \nu t$. Ciò suggerisce che se n è "molto grande" e p è "molto piccolo", la distribuzione binomiale può essere approssimata con una Poisson con $\lambda = np$. Un criterio empirico che assicura una approssimazione molto buona si ha per valori di $n \geq 100$ e $np \leq 10$.

Teorema 2.11. Sia X una variabile aleatoria con distribuzione di Poisson di parametro $\lambda > 0$, allora

$$E(X) = \lambda, \quad V(X) = \lambda, \quad m(t) = e^{\lambda(e^t - 1)}. \tag{2.17}$$

Dimostrazione. Quanto al valore atteso, non sorprende che esso sia pari ad $np = \lambda$, dal momento che la distribuzione di Poisson è stata ottenuta come limite della binomiale con media np e mantenendo fissa la quantità np. Il calcolo si sviluppa come segue:

$$\begin{aligned}
E(X) &= \sum_{x=0}^{\infty} x \frac{\lambda^x}{x!} e^{-\lambda} \\
&= \sum_{x=1}^{\infty} x \frac{\lambda \lambda^{x-1}}{x(x-1)!} e^{-\lambda} \\
&= \lambda \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} e^{-\lambda} \\
&= \lambda e^{-\lambda} \sum_{y=0}^{\infty} \frac{\lambda^y}{y!} = \lambda,
\end{aligned}$$

dove si è posto $y = x - 1$. La varianza si ottiene come

$$\begin{aligned}
E(X^2) &= \sum_{x=0}^{\infty} x^2 \frac{\lambda^x}{x!} e^{-\lambda} \\
&= \sum_{x=0}^{\infty} (x(x-1) + x) \frac{\lambda^x}{x!} e^{-\lambda} \\
&= \sum_{x=0}^{\infty} x(x-1) \frac{\lambda^x}{x!} e^{-\lambda} + \sum_{x=0}^{\infty} x \frac{\lambda^x}{x!} e^{-\lambda} \\
&= \sum_{x=2}^{\infty} x(x-1) \frac{\lambda^2 \lambda^{x-2}}{x(x-1)(x-2)!} e^{-\lambda} + \lambda,
\end{aligned}$$

da cui si ottiene, ponendo $y = x - 2$,

$$E(X^2) = \lambda^2 e^{-\lambda} \sum_{y=0}^{\infty} \frac{\lambda^y}{y!} + \lambda = \lambda^2 + \lambda.$$

Pertanto si trova

$$V(X) = E(X^2) - E(X)^2 = \lambda,$$

cioè media e varianza coincidono. Si tratta di una proprietà caratteristica della distribuzione di Poisson. Infine, per la funzione generatrice dei momenti,

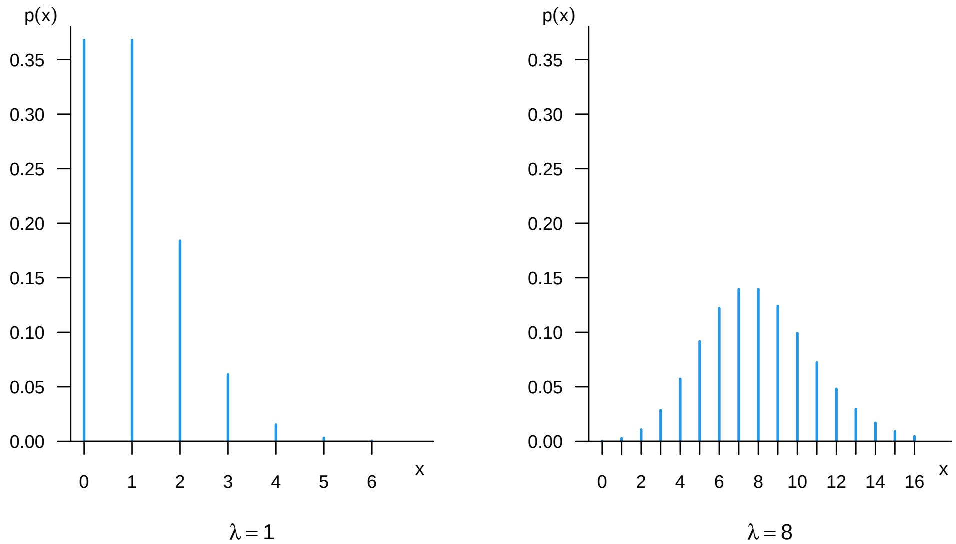


Figura 2.4: Funzione di probabilità di Poisson di parametri $\lambda = 1$ e $\lambda = 8$.

si osserva che

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

$\square$

Esempio 2.3.

Si assume spesso che fenomeni come incidenti che causano lesioni in un grande impianto industriale soddisfino le assunzioni di un processo di Poisson. Supponiamo che ciò sia vero per un impianto in particolare, per cui si ha un tasso medio di incidenti di $\nu = \frac{1}{2}$ a settimana. Se X rappresenta in numero di incidenti che si verificano in un periodo di $t = 6$ settimane, allora X ha una distribuzione di Poisson con parametro $\lambda = \frac{1}{2}6 = 3$ e la sua funzione di probabilità è data da

$$p(x) = \frac{3^x}{x!} e^{-3}, \quad x = 0, 1, 2 \ldots$$

Il valore atteso e la varianza di X sono entrambi pari a 3. La probabilità di esattamente 3 incidenti nello stesso periodo è

$$p(3) = \frac{3^3}{3!}e^{-3} = 0.224$$

e la probabilità di almeno 4 incidenti nello stesso periodo è

$$\sum_{x=4}^{\infty} \frac{3^x}{x!}e^{-3} = 1 - e^{-3}\left(1 + 3 + \frac{3^2}{2} + \frac{3^3}{6}\right) = 0.353$$

Infine la probabilità che non si verifichino incidenti si calcola per $x = 0$ ottenendo

$$p(0) = \frac{3^0}{0!}e^{-3} = 0.050.$$

Quest'ultima può essere calcolata anche in un modo alternativo. Se non vi sono incidenti nel periodo di 6 settimane, allora non possono verificarsi incidenti nei tre periodi successivi di 2 settimane. Il numero di incidenti nel periodo $t = 2$ settimane, Y, è una variabile aleatoria di Poisson con parametro $\lambda = \frac{1}{2}2 = 1$, e la probabilità di zero incidenti sarà allora $\frac{1^0}{0!}e^{-1} = e^{-1}$. Intervalli di tempo non sovrapposti in un processo di Poisson sono *prove indipendenti*, nello specifico, tre prove di Bernoulli indipendenti con probabilità di successo (nessun incidente) pari a e^{-1} in ogni prova. La probabilità di tre successi in fila è semplicemente $(e^{-1})^3 = e^{-3}$, che è lo stesso risultato ottenuto sopra.

Dalla (2.16) si deduce che la sequenza delle probabilità in corrispondenza degli interi $0, 1, 2, \ldots$ è $e^{-\lambda}, \lambda e^{-\lambda}, (\lambda^2/2!)\, e^{-\lambda}, \ldots$ e quindi la prima probabilità (di $X = 0$) è semplicemente $e^{-\lambda}$; il secondo termine si ottiene moltiplicando questo valore per λ, il terzo si ottiene moltiplicando $\lambda/2$ per il termine precedente, e così via. Infatti, vale

$$\frac{p(x+1)}{p(x)} = \frac{\dfrac{\lambda^{x+1}}{(x+1)!}e^{-\lambda}}{\dfrac{\lambda^x}{x!}e^{-\lambda}} = \frac{\lambda}{x+1},$$

ovvero ogni termine è espresso da λ diviso per un intero, moltiplicato per il termine precedente, rendendo più semplice valutare la funzione di ripartizione di X, $F(t) = \sum_{x=0}^{t} p(x)$. Si noti che se il parametro λ è intero, per $\lambda = x+1$ si ha

$$\frac{p(x+1)}{p(x)} = \frac{\lambda}{x+1} = 1$$

e le due probabilità $p(x+1)$ e $p(x)$ risultano uguali. Per valori di x più piccoli di $\lambda - 1$ le probabilità sono crescenti, mentre per x maggiore di λ si ottengono valori decrescenti.

2.1.5 Variabili aleatorie ipergeometriche

Supponiamo di dover estrarre un campione di n palline senza reinserimento da un'urna che contiene N palline, delle quali m sono bianche ed $N - m$ sono nere. Sia X il numero delle palline bianche presenti tra le n estratte. Se $m \geq n$ ed $N-m \geq n$, il supporto di X è dato dagli interi $0, 1, \ldots, n$. Il numero totale dei campioni che si possono estrarre dall'urna è $\binom{N}{n}$, che corrisponde al numero di tutti i possibili sottoinsiemi di cardinalità n che si possono costruire da un insieme di N elementi. Il campione è estratto casualmente dall'urna, e quindi ogni sottoinsieme ha la stessa probabilità $1/\binom{N}{n}$ di essere selezionato. Il numero dei sottoinsiemi contenenti esattamente x palline bianche (e quindi $n - x$ nere) è

$$\binom{m}{x}\binom{N-m}{n-x},$$

e pertanto si arriva alla funzione di probabilità di X, numero delle palline bianche nel campione estratto:

$$P(X = x) = \frac{\binom{m}{x}\binom{N-m}{n-x}}{\binom{N}{n}}, \quad x = 0, 1, 2, \ldots, n. \tag{2.18}$$

Definizione 2.6 (Variabile aleatoria ipergeometrica)**.** Dato un insieme di N oggetti, di cui m appartenenti alla categoria A ed $N - m$ alla categoria $\overline{A}$, si estraggono a sorte dall'insieme, senza reimmissione, n oggetti. La variabile aleatoria X, data dal numero di oggetti di tipo A che si presentano nelle n estrazioni è detta *variabile aleatoria ipergeometrica* di parametri N, m, n e si indica con la simbologia $X \sim Iper(N, m, n)$.

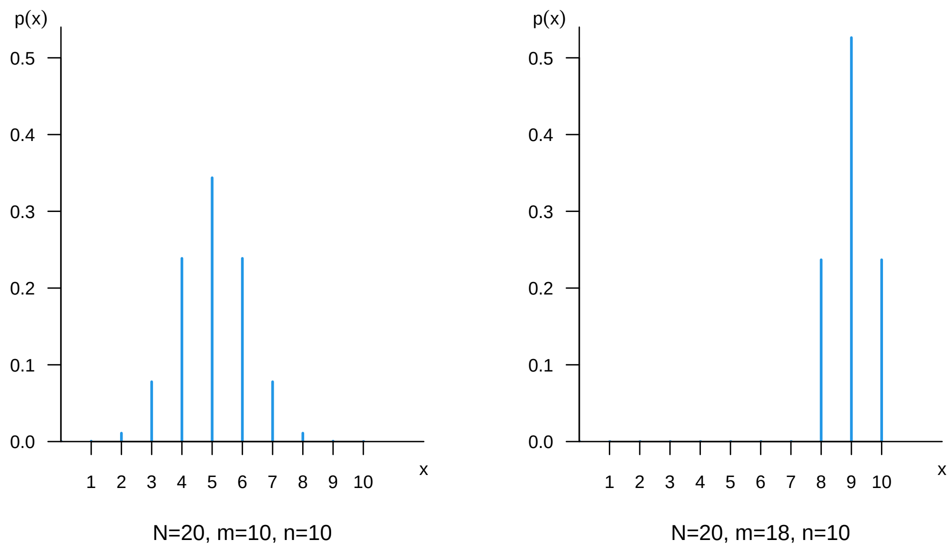


Figura 2.5: Funzione di probabilità ipergeometrica per $N = 20$, $n = 10$ e alcuni valori di m.

Per ricavare la (2.18) si è assunto che le palline bianche e nere nell'urna siano in numero almeno pari ad n, tuttavia alcuni dei valori nel supporto di X hanno probabilità pari a zero. Ad esempio, se consideriamo il caso di $m = 3$ palline bianche, $N - m = 2$ palline nere (quindi $N = 5$) ed $n = 4$, allora in ogni campione di 4 palline estratte dall'urna saranno presenti almeno due palline bianche e non più di tre palline bianche; la probabilità che il numero delle palline bianche nel campione estratto sia $0, 1, 4$ è dunque zero. Ciononostante, la (2.18) è sempre valida in base alla convenzione che $\binom{r}{k} = 0$ se $k < 0$ o $r < k$.

Teorema 2.12. Se X è una variabile aleatoria ipergeometrica di parametri N, m, n, allora X ha funzione di probabilità

$$p(x) = \frac{\binom{m}{x}\binom{N-m}{n-x}}{\binom{N}{n}}, \tag{2.19}$$

$\max\{0, n-(N-m)\} \leq x \leq \min\{n, m\}$.

La (2.19) è nota come *distribuzione ipergeometrica*. Si tratta sempre di un modello probabilistico associato a prove Bernoulliane ripetute, in cui però si è interessati a determinare la probabilità che in n prove si verifichino x "successi". L'elemento di differenziazione è che nel nuovo modello non si assume l'indipendenza delle prove.

Esempio 2.4.

Sia X il numero di esemplari difettosi in un campione di ampiezza n, in un campionamento fatto senza reimmissione da un'urna contenente N pezzi, m dei quali sono difettosi. Allora X ha una distribuzione ipergeometrica.

Si pensi, ad esempio, ad un rivenditore che acquista le componenti elettriche a lotti di 10. La sua politica è di controllare a caso 3 componenti di ogni lotto e di accettarlo solo se nessuno dei tre pezzi controllati risulta difettoso. Se il 30% dei lotti ha quattro pezzi difettosi e il 70% dei lotti ha un solo pezzo difettoso, che percentuale di lotti il rivenditore rifiuterà? Si denoti con A l'evento che il rivenditore accetti un lotto. Allora,

$$\begin{aligned} P(A) &= P(A|\text{il lotto ha 4 pezzi difettosi})\frac{3}{10} \\ &\quad + P(A|\text{il lotto ha un pezzo difettoso})\frac{7}{10} \\ &= \frac{\binom{4}{0}\binom{10-4}{3-0}}{\binom{10}{3}}\left(\frac{3}{10}\right) + \frac{\binom{1}{0}\binom{10-1}{3-0}}{\binom{10}{3}}\left(\frac{7}{10}\right) = \frac{54}{100} \end{aligned}$$

Da cui deduciamo che il 46% dei lotti verrà rifiutato.

Si noti che se la (2.18) è valida, allora deve essere

$$\sum_{x=0}^{n} \binom{m}{x}\binom{N-m}{n-x} = \binom{N}{n}. \tag{2.20}$$

Teorema 2.13. Se X ha una distribuzione ipergeometrica, allora

$$E(X) = \frac{nm}{N}, \quad V(X) = \frac{nm}{N}\frac{N-m}{N}\frac{N-n}{N-1}. \tag{2.21}$$

Dimostrazione.

$$E(X^k) = \sum_{x=0}^{n} x^k \binom{m}{x}\binom{N-m}{n-x} \Big/ \binom{N}{n}$$

Utilizzando le identità

$$x\binom{m}{x} = m\binom{m-1}{x-1},$$

$$n\binom{N}{n} = N\binom{N-1}{n-1},$$

si ha

$$\begin{aligned} E(X^k) &= \frac{nm}{N}\sum_{x=1}^{n} x^{k-1}\binom{m-1}{x-1}\binom{N-m}{n-x} \Big/ \binom{N-1}{n-1} \\ &= \frac{nm}{N}\sum_{y=0}^{n-1}(y+1)^{k-1}\binom{m-1}{y}\binom{N-m}{n-y-1} \Big/ \binom{N-1}{n-1}, \qquad y = x-1 \\ &= \frac{nm}{N}E[(Y+1)^{k-1}] \end{aligned}$$

dove Y è una variabile aleatoria ipergeometrica di parametri $n-1$, $N-1$, $m-1$. Ponendo $k=1$ si ottiene

$$E(X) = \frac{nm}{N}.$$

Ponendo $k = 2$ nella formula di $E(X^k)$

$$E(X^2) = \frac{nm}{N} E(Y+1) = \frac{nm}{N}\left[\frac{(n-1)(m-1)}{N-1} + 1\right]$$

Quindi

$$\begin{aligned} V(X) =& E(X^2) - [E(X)]^2 \\ =& \frac{nm}{N}\left[\frac{(n-1)(m-1)}{N-1} + 1\right] - \frac{n^2m^2}{N^2} \\ =& \frac{nm}{N}\left[\frac{(n-1)(m-1)}{N-1} + 1 - \frac{nm}{N}\right] \\ =& \frac{nm}{N}\left[\frac{(N-m)(N-n)}{N(N-1)}\right] \end{aligned}$$

$\square$

Si osservi che se denotiamo con $p = m/N$ la frazione di palline bianche estratte, allora si ottiene

$$V(X) = \frac{N-n}{N-1} np(1-p),$$

cioè la varianza della distribuzione ipergeometrica è $(N-n)/(N-1)$ volte la varianza della distribuzione binomiale.
La funzione generatrice dei momenti e la funzione generatrice dei momenti fattoriali non hanno grande utilità per la distribuzione ipergeometrica, in quanto, come si è visto, i momenti possono essere facilmente calcolati senza ricorrere all'uso delle funzioni generatrici.
Si supponga nuovamente di scegliere a caso n palline senza reinserimento da un insieme di N palline, delle quali la frazione $p = m/N$ è bianca; allora il numero di palline bianche selezionate è dato da una legge ipergeometrica. Appare ragionevole supporre che se m ed N sono "grandi" in rapporto ad n, allora il fatto che si effettui o meno il reinserimento a ogni estrazione possa essere trascurabile. Non tenendo conto delle palline precedentemente estratte, ogni successiva estrazione darà, quando m ed N sono grandi una pallina bianca con probabilità approssimativamente pari a p. In questo caso è quindi intuitivo supporre che la funzione di probabilità di X possa essere

ben approssimata da una variabile aleatoria binomiale di parametri n e p. Analiticamente, per $x \leq n$, si ha

$$
\begin{aligned}
p(x) &= \frac{\binom{m}{x}\binom{N-m}{n-x}}{\binom{N}{n}} \\
&= \frac{m!}{(m-x)!x!} \cdot \frac{(N-m)!}{(N-m-n+x)!(n-x)!} \cdot \frac{(N-n)!n!}{N!} \\
&= \frac{n!}{(n-x)!x!} \cdot \frac{m!}{(m-x)!N!} \cdot \frac{(N-m)!(N-n)!}{(N-m-(n-x))!} \\
&= \binom{n}{x} \frac{m!(N-x)!}{(m-x)!N!} \cdot \frac{(N-m)!(N-n)!}{(N-x)!(N-m-(n-x))!} \\
&= \binom{n}{x} \frac{m!/(m-x)!}{N!/(N-x)!} \cdot \frac{(N-m)!/(N-m-(n-x))!}{(N-x)!/(N-n)!} \\
&= \binom{n}{x} \frac{m(m-1)\cdots(m-x+1)}{N(N-1)\cdots(N-x+1)} \\
&\quad \cdot \frac{(N-m)(N-m-1)\cdots(N-m-(n-x)+1)}{(N-n+(n-x))\cdots(N-n+1)} \\
&= \binom{n}{x} \prod_{k=1}^{x} \frac{m-x+k}{N-x+k} \prod_{k=1}^{n-x} \frac{(N-m)-(n-x)+k}{N-n+k} \\
&\approx \binom{n}{x} p^x (1-p)^{n-x}, \\
&\quad \text{con}\, p = m/N \text{ ed } m, N \text{ grandi rispetto a } n \text{ e } x
\end{aligned}
$$

2.1.6 Variabili aleatorie uniformi discrete

Si pensi ad un'urna contenente N palline numerate da 1 a N. L'esperimento aleatorio consiste nell'estrarre a sorte una pallina. Le determinazioni della variabile aleatoria che ne risulta sono i primi N numeri interi, e la probabilità associata a ciascuno di questi valori è $1/N$. Si pensi all'esempio più semplice del lancio di un dado equilibrato, e si indichi con X il risultato. Allora la variabile aleatoria X può assumere i valori $1, 2, 3, 4, 5, 6$ e quindi

$$P(X = x) = \begin{cases} \frac{1}{6} & \text{per } x = 1, 2, 3, 4, 5, 6 \\ 0 & \text{altrimenti} \end{cases}$$

Definizione 2.7 (Variabile aleatoria uniforme discreta)**.** Una variabile aleatoria ha distribuzione *uniforme discreta* negli interi $1, \ldots, N$ se la sua funzione di probabilità è espressa da

$$p(x) = \frac{1}{N}, \quad x = 1, \ldots, N. \tag{2.22}$$

Quindi la distribuzione uniforme discreta è quella di un numero aleatorio discreto che può assumere (con probabilità positiva) solo i valori interi compresi in un certo intervallo e per il quale tali valori sono tutti *equiprobabili*.

Esempio 2.5.

Una scatola contiene 7 lampadine numerate da 1 a 7. Le lampadine vengono testate a sorte una dopo l'altra, senza reimmissione. Sia X il numero della lampadina che risulta da una estrazione. Dunque X assume come valori i numeri $1, 2, \ldots, 7$ e

$$p(x) = P(X = x) = \frac{1}{7}, \quad x = 1, \ldots, 7.$$

Teorema 2.14. La media, la varianza e la funzione generatrice dei momenti della distribuzione uniforme discreta sono date rispettivamente da

$$E(X) = \frac{N+1}{2}, \quad V(X) = \frac{N^2-1}{12},$$
$$m(t) = \frac{1}{N}(e^t + e^{2t} + \cdots + e^{Nt}) \tag{2.23}$$

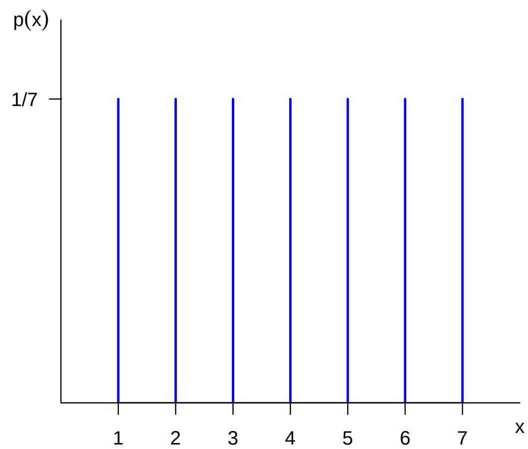


Figura 2.6: Funzione di probabilità uniforme discreta relativa all'Esempio 2.5.

Dimostrazione. Dalla definizione si ha

$$E(X) = \sum_{x=1}^{N} x \frac{1}{N} = \frac{1}{N}(1 + 2 + \cdots + N)$$

da cui, ricordando che $\sum_{i=1}^{k} i = \dfrac{k(k+1)}{2}$, si ottiene

$$E(X) = \frac{1}{N} \frac{N(N+1)}{2} = \frac{N+1}{2}$$

La varianza è espressa da

$$V(X) = \sum_{x=1}^{N} x^2 \frac{1}{N} - \left(\frac{N+1}{2}\right)^2 = \frac{1}{N}(1 + 2^2 + \cdots + N^2) - \left(\frac{N+1}{2}\right)^2.$$

Ricordando che $\sum_{i=1}^{k} i^2 = \dfrac{k(k+1)(2k+1)}{6}$, si ha

$$\begin{aligned}
V(X) &= \frac{1}{N}\frac{N(N+1)(2N+1)}{6} - \frac{(N+1)^2}{4} \\
&= \frac{(2N+2)(2N+1) - 3(N^2+1+2N)}{12} \\
&= \frac{4N^2+2N+4N+2-3N^2-3-6N}{12} \\
&= \frac{N^2-1}{12}.
\end{aligned}$$

Infine si determina facilmente la funzione generatrice dei momenti:

$$m(t) = E(e^{tX}) = \sum_{x=1}^{N} e^{tx}\frac{1}{N} = \frac{1}{N}(e^t + e^{2t} + \cdots + e^{Nt}).$$

$\square$

2.1.7 Altri modelli discreti

Nei precedenti paragrafi sono state descritte alcune famiglie di modelli parametrici unidimensionali note e comunemente impiegate nelle applicazioni. Tuttavia è importante sottolineare che esistono molte altre famiglie di funzioni di probabilità discrete che possono essere formate per mezzo di procedimenti come il *troncamento* o la *traslazione*, a partire da quelle esistenti. Ne è un esempio la *variabile aleatoria di Poisson troncata in 0*, che si presenta nelle situazioni in cui il conteggio 0 non può essere osservato (si pensi ad un contatore automatico che avanza di uno scatto solo quando deve registrare un passaggio, una emissione, ecc.). Il fenomeno a partire dal conteggio 1 in poi è però ben modellizzato dalla distribuzione di Poisson. In tal caso si adotta la procedura di troncamento, applicabile anche ad altre distribuzioni, che consiste nel considerare tutti i punti massa della distribuzione di Poisson ad eccezione del punto $x = 0$ e, a ciascuno di essi, associare il valore della probabilitàà di Poisson moltiplicato per una costante k non negativa scelta in maniera tale che la somma dei valori di $p(x)$ per $x = 1, 2, \ldots$ sia uguale a

1, cioè che possa essere a sua volta una funzione di probabilità. Si ha:

$$p(x) = \begin{cases} k\dfrac{\lambda^x e^{-\lambda}}{x!} & \text{per } x = 1, 2, \ldots \\ \\ 0 & \text{altrove} \end{cases}$$

con k tale che $\sum_{x=1}^{\infty} k\frac{\lambda^x e^{-\lambda}}{x!} = 1$, cioè

$$1 = ke^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^x}{x!} = ke^{-\lambda}(e^{\lambda} - 1) \to k = \frac{1}{1 - e^{-\lambda}}$$

Quindi la distribuzione di Poisson troncata in 0 risulta essere

$$p(x) = \begin{cases} \dfrac{\lambda^x e^{-\lambda}}{x!(1 - e^{-\lambda})} & \text{per } x = 1, 2, \ldots \\ \\ 0 & \text{altrove} \end{cases}$$

Altre volte può accadere di avere un fenomeno ben modellizzato da una distribuzione di Poisson, ma il conteggio si ferma ad un valore k (in genere $<< \infty$) o perché non siamo in grado di osservarlo oltre tale valore o per carenze del contatore utilizzato o perché si decide a priori di troncare l'esperimento dopo un numero k di eventi manifestati (*prove censurate*). Si fa ricorso in questo caso ad una variabile casuale X detta *variabile casuale censurata* la cui distribuzione ha come punti massa i punti $x = 0, 1, 2, \ldots, k - 1$, mentre nel punto $x = k$ risulta essere concentrata tutta la massa rimanente:

$$p(x) = \begin{cases} \lambda^x e^{-\lambda}/x! & \text{per } x = 0, 1, 2, \ldots, k - 1 \\ 1 - \displaystyle\sum_{x=0}^{k-1} \frac{\lambda^x e^{-\lambda}}{x!} & \text{per } x = k \\ 0 & \text{altrove} \end{cases}$$

Naturalmente questo procedimento illustrato per la distribuzione di Poisson si può estendere in modo analogo ad altre distribuzioni discrete ogni qual volta si renda necessario.

2.2 Modelli di distribuzione continui

2.2.1 Variabili aleatorie rettangolari o uniformi

La distribuzione più semplice nel caso continuo è quella uniforme, che può essere vista come l'estensione al caso continuo della (2.22). Diremo che una variabile aleatoria continua è *uniformemente* distribuita nell'intervallo (a, b) se la probabilità che X assuma un valore in un dato intervallo di lunghezza Δt in (a,b), è proporzionale a Δt. La probabilità di osservare un valore di X più piccolo di a o più grande di b è zero. Ne segue che la densità di X è positiva solo per valori compresi tra a e b, ovvero $f(x) > 0$, per $a < x < b$. Inoltre, affinché la probabilità che X stia in un particolare intervallo sia proporzionale alla lunghezza dell'intervallo, $f(x)$ deve essere costante per $a < x < b$, ovvero $f(x) = c$, per $a < x < b$. Integrando $f(x)$ sul supporto di X si ha

$$\int_a^b c dx = c(b - a)$$

e ponendo uguale ad 1 la precedente espressione si ottiene $c = 1/(b - a)$.

Teorema 2.15. Se X è una variabile aleatoria continua, distribuita uniformemente nell'intervallo (a, b), la funzione di densità di probabilità di X è espressa da

$$f(x) = \frac{1}{b-a}, \quad a < x < b. \tag{2.24}$$

La distribuzione data dalla (2.24) è detta *distribuzione uniforme* di parametri a e b, $\infty < a < b < \infty$, e si utilizza la notazione $X \sim U(a, b)$.

Teorema 2.16. Se X è uniformemente distribuita nell'intervallo (a, b), allora

$$E(X) = \frac{a+b}{2}, \quad V(X) = \frac{(b-a)^2}{12}, \quad m(t) = \frac{e^{tb} - e^{ta}}{t(b-a)}$$

Dimostrazione.

$$E(X) = \int_a^b x\frac{1}{b-a}dx = \frac{b^2-a^2}{2(b-a)} = \frac{b+a}{2}.$$

$$\begin{aligned} V(X) = E(X^2) - [E(X)]^2 &= \int_a^b x^2\frac{1}{b-a}dx - \left(\frac{b+a}{2}\right)^2 \\ &= \frac{b^3-a^3}{3(b-a)} - \frac{(b+a)^2}{4} \\ &= \frac{(b-a)(b^2+ab+a^2)}{3(b-a)} - \frac{b^2+a^2+2ab}{4} \\ &= \frac{(b-a)^2}{12}. \end{aligned}$$

$$m(t) = E(e^{tX}) = \int_a^b \frac{e^{tx}}{b-a}dx = \frac{1}{b-a}\int_a^b e^{tx}dx = \frac{e^{tb}-e^{ta}}{t(b-a)}, \quad t \neq 0.$$

□

La deviazione standard per una variabile uniforme è quindi proporzionale alla lunghezza dell'intervallo su cui si distribuisce uniformemente, essendo

$$\sqrt{V(X)} = \frac{b-a}{\sqrt{12}}.$$

La distribuzione uniforme viene anche chiamata *rettangolare*, a causa della forma della funzione di densità. Essendo $F(t) = \int_{-\infty}^t f(x)dx$, la funzione di ripartizione di una variabile casuale uniforme è data da

$$F(t) = \begin{cases} 0, & t < a \\ \displaystyle\int_a^t \frac{1}{b-a}dx = \frac{t-a}{b-a}, & a \leq t \leq b \\ 1, & t > b. \end{cases} \tag{2.25}$$

La (2.25) fornisce un modello utile per alcuni fenomeni casuali. Se, ad esempio, sappiamo che una variabile casuale X può assumere valori solo in un intervallo finito (a, b) e se si assume che ogni sotto-intervallo di (a, b) di uguale lunghezza abbia la stessa probabilità di contenere X, allora X ha una distribuzione uniforme nell'intervallo (a, b). Quando si parla di un *numero casuale* nell'intervallo $(0, 1)$ si allude al valore di una variabile casuale uniformemente distribuita nell'intervallo $(0, 1)$.

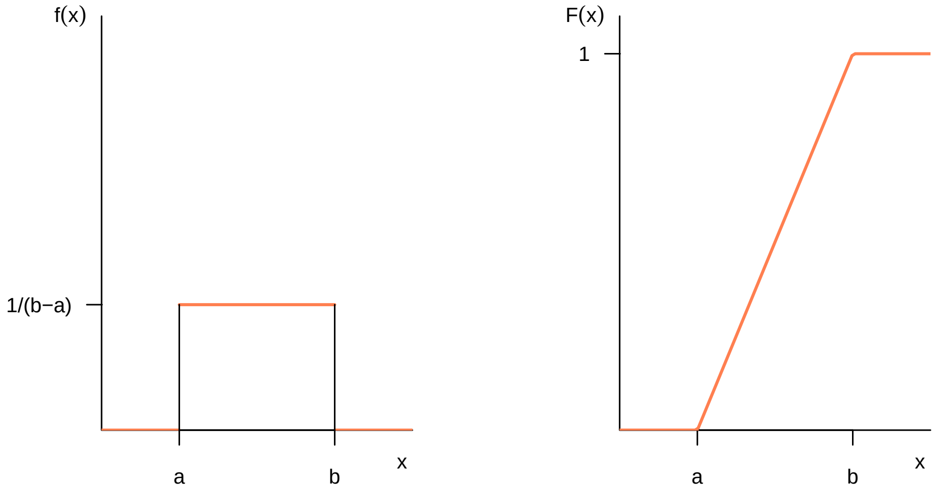


Figura 2.7: Funzione di densità e di ripartizione di una variabile aleatoria uniforme.

Esempio 2.6.

Si consideri l'esempio di una simulazione progettata per studiare gli effetti di vari tipi e posizioni di materiali isolanti per la progettazione di un palazzo. Assumiamo che si sia interessati ad osservare il valori di variabili casuali Bernoulliane di parametro p. Tali valori possono essere generati ricorrendo alla variabile aleatoria uniforme nell'intervallo $(0, 1)$. Sia X uniformemente distribuita nell'intervallo $(0, 1)$ e supponiamo di voler osservare il valore di Y, variabile casuale di Bernoulli con parametro p. Allora, se definiamo

$$\begin{aligned} Y &= 1 \quad \text{se} \;\; X \leq p \\ &= 0 \quad \text{se} \;\; X > p, \end{aligned}$$

si ha che

$$P(Y = 1) = P(X \leq p) = F_X(p) = p,$$

dove $F_X(\cdot)$ è la funzione di ripartizione di X. Quindi se, ad esempio, si desidera generare il valore osservato di una variabile aleatoria di Bernoulli con $p = 0.75$, si può generare un numero casuale nell'intervallo $(0, 1)$. Se il valore osservato risulta essere pari o inferiore a 0.75 allora si pone $Y = 1$, altrimenti $Y = 0$. Pertanto Y è una variabile aleatoria di Bernoulli con $p = 0.75$.

Abbiamo definito la distribuzione uniforme in un intervallo *aperto* (a, b), ma si potrebbe anche definirla nell'intervallo chiuso $[a, b]$ o in uno degli intervalli $[a, b)$, $(a, b]$. Si noti come, per ognuno di questi casi, si ha la stessa funzione di ripartizione. La funzione di probabilità $f(\cdot)$ di una variabile casuale X continua non è univocamente definita. Possono infatti esistere più di una funzione $f(\cdot)$ in grado di soddisfare la condizione $F(x) = \int_{-\infty}^{x} f(u)du$ per ogni x reale. Il concetto è che se il valore di una funzione cambia solo in "pochi" punti, il suo integrale rimane inalterato. Nella pratica, tuttavia, ci si limita a considerare *la* densità di probabilità come scelta univoca di $f(\cdot)$ in virtù di considerazioni di continuità.

2.2.2 Variabili aleatorie esponenziali

Possiamo introdurre la distribuzione esponenziale considerando il legame con un processo di Poisson in cui si osservano in media λ eventi in un intervallo di tempo unitario. Se l'esperimento consiste nel conteggio del numero di manifestazioni di un certo fenomeno nel tempo tale numero ha, sotto certe condizioni, una distribuzione Poisson con parametro, la media, proporzionale alla lunghezza dell'intervallo (si veda 2.1.4). Supponiamo ora che si sia appena verificata una di queste manifestazioni, ed indichiamo con T l'istante temporale in cui ciò si verifica. Allora T è una variabile aleatoria continua il cui supporto è dato da $R_T = \{t : t > 0\}$. Fissato t, si consideri l'evento "nessuna manifestazione si verifica nell'intervallo di lunghezza t, $(0, t]$", ovvero $\{T > t\}$. Ora, sappiamo che la probabilità che si verifichino zero eventi in ogni intervallo di lunghezza fissata t è $P(X = 0) = e^{-\lambda t}$, essendo X il numero degli eventi in $(0, t]$ e λ il numero medio di arrivi nell'unità di tempo. Quindi si ha

$$P(T > t) = 1 - F_T(t) = P(X = 0) = e^{-\lambda t},$$

da cui segue che la funzione di ripartizione di T è

$$F_T(t) = 1 - e^{-\lambda t}, \quad t > 0$$

e la funzione di densità è

$$f(t) = \frac{d}{dt} F_T(t) = \lambda e^{-\lambda t}, \quad t > 0.$$

Abbiamo così ottenuto la funzione di densità *esponenziale (negativa)*.

Teorema 2.17. Si assuma che il numero aleatorio di eventi osservati in un intervallo di tempo fissato di lunghezza t sia distribuito come una Poisson (con parametro λt). Allora il tempo T necessario per osservare un evento è distribuito come una variabile aleatoria esponenziale di parametro λ ($T \sim Esp(\lambda)$) avente densità

$$f(t) = \lambda e^{-\lambda t}, \quad t \geq 0, \tag{2.26}$$

dove $\lambda > 0$, e funzione di ripartizione

$$F(t) = \int_0^t \lambda e^{-\lambda x} dx = 1 - e^{-\lambda t}, \quad t \geq 0. \tag{2.27}$$

Quanto visto prova che la lunghezza dell'intervallo di tempo fra manifestazioni successive di un fenomeno ha una distribuzione esponenziale, se il numero delle manifestazioni in un dato intervallo di tempo ha una distribuzione di Poisson. Per questo la distribuzione esponenziale può essere impiegata come modello per la durata di vita di diversi fenomeni: il tempo (a partire da adesso) affinché avvenga un terremoto, il tempo di permanenza di un cliente ad uno sportello, il numero di chilometri che un'auto può percorrere prima che la batteria ceda.

Teorema 2.18. Sia X una variabile aleatoria distribuita secondo una distribuzione esponenziale di parametro $\lambda > 0$. Allora, si ha

$$E(X) = \frac{1}{\lambda}, \quad V(X) = \frac{1}{\lambda^2}, \quad m(t) = \frac{\lambda}{\lambda - t}. \tag{2.28}$$

Dimostrazione. Il valore atteso di X si ottiene integrando per parti come segue

$$\begin{aligned} E(X) &= \int_0^\infty x\lambda e^{-\lambda x} dx \\ &= (-xe^{-\lambda x})\Big|_0^\infty - \int_0^\infty (-e^{-\lambda x}) dx \\ &= 0 + \frac{1}{\lambda}\int_0^\infty \lambda e^{-\lambda x} dx \\ &= \frac{1}{\lambda}(-e^{-\lambda x})\Big|_0^\infty = \frac{1}{\lambda}. \end{aligned} \tag{2.29}$$

La funzione generatrice dei momenti è data da

$$\begin{aligned} m(t) = E(e^{tX}) &= \int_0^\infty e^{tx}\lambda e^{-\lambda x} dx \\ &= \int_0^\infty \lambda e^{-x(\lambda - t)} dx \\ &= -\frac{\lambda e^{-x(\lambda-t)}}{\lambda - t}\Big|_0^\infty \\ &= \frac{\lambda}{\lambda - t}, \quad t < \lambda \end{aligned} \tag{2.30}$$

da cui si deduce che la varianza è data da

$$\begin{aligned} V(X) &= E(X^2) - [E(X)]^2 = m''(0) - \frac{1}{\lambda^2} \\ &= \frac{d}{dt}\left(\frac{\lambda}{(\lambda-t)^2}\right)\Big|_{t=0} - \frac{1}{\lambda^2} = \frac{2\lambda}{\lambda^3} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}. \end{aligned}$$

Quindi la variabile aleatoria esponenziale X ha media e deviazione standard uguali e pari a $\frac{1}{\lambda}$. □

Esempio 2.7.

In un grande impianto industriale si verifica in media un incidente ogni due settimane. Considerando 5 giorni lavorativi a settimana, ciò equivale a un numero medio di 1/10 al giorno. Si assume che il numero di incidenti che si verificano in un giorno ($t = 1$) segua una distribuzione di Poisson di parametro λ. Se osserviamo il verificarsi di tali incidenti a partire da un lunedì per alcune settimane, e denotiamo con T il numero dei giorni fino al verificarsi del primo incidente, allora T è esponenziale di parametro $\lambda = 1/10$ (la densità e la funzione di ripartizione sono rappresentate di seguito). Possiamo, ad esempio, calcolare la probabilità che non si verifichino incidenti nella prima settimana

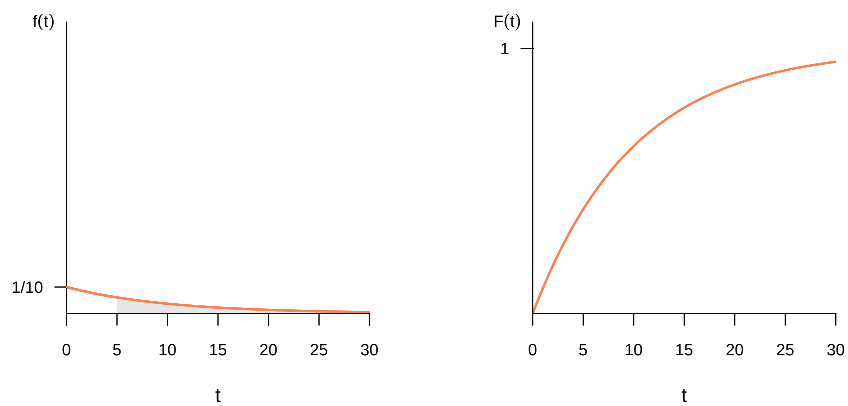
$$P(T > 5) = e^{-\frac{1}{10}\cdot 5} = 0.607$$

o la probabilità che il primo incidente si verifichi il mercoledì della seconda settimana

$$P(7 < T \leq 8) = F(8) - F(7) = 1 - e^{-\frac{1}{10}\cdot 8} - (1 - e^{-\frac{1}{10}\cdot 7}) = 0.047$$

Il numero medio di giorni per il primo incidente è dato da

$$E(T) = \frac{1}{\lambda} = \left(\frac{1}{10}\right)^{-1} = 10.$$



Come per la legge geometrica, anche per la distribuzione esponenziale parliamo di proprietà di assenza di memoria, nel senso che vale la relazione

$$P(T > a + b | T > a) = P(T > b)$$

con a e b costati positive. Quindi se nell'esempio precedente supponiamo di aver osservato $a = 4$ giorni senza incidenti allora la probabilità che trascorrano almeno $b = 2$ giorni prima di osservare il primo incidente è la stessa della probabilità dell'evento all'inizio dell'osservazione. La distribuzione esponenziale è la sola distribuzione continua a godere di tale proprietà. Il motivo per cui la legge esponenziale e quella geometrica condividono questa proprietà è da ricercare nel fatto che il processo di Poisson può essere derivato come limite di una successione di prove Bernoulliane indipendenti. Una variabile geometrica Y è il numero di tentativi che occorrono per il primo successo, e la variabile aleatoria esponenziale T è definita analogamente rispetto alla distribuzione di Poisson: essa rappresenta il tempo che occorre perché si verifichi il primo evento (successo). Infatti, se Y è geometrica di parametro p, allora

$$P(Y > n) = (1-p)^n.$$

D'altra parte, suddividendo un intervallo di lunghezza unitaria in n intervallini disgiunti, si ha che $p = \lambda/n$ e la probabilità che non si verifichino eventi in tale intervallo è $1 - \frac{\lambda}{n}$. Allora gli eventi $\{Y > n\}$ e $\{T > 1\}$ sono equivalenti e

$$\begin{aligned} P(T > 1) &= \lim_{n\to\infty} P(Y > n) \\ &= \lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^n \\ &= e^{-\lambda} \end{aligned}$$

ovvero la funzione di ripartizione esponenziale è il limite della funzione di ripartizione geometrica e si può dire che la variabile aleatoria esponenziale "eredita" la proprietà di assenza di memoria dalla legge geometrica.

2.2.3 Funzioni di rischio

Sia X una variabile aleatoria continua e positiva, che interpretiamo come il *tempo di vita* di qualche oggetto, con funzione di distribuzione F e densità f. La **funzione di rischio** (*hazard rate*) $\lambda(t)$ di F è definita da

$$\lambda(t) = \frac{f(t)}{\bar{F}(t)}, \quad \bar{F}(t) = P(X > t).$$

Per interpretare $\lambda(t)$, supponiamo che l'oggetto sia sopravvissuto per un tempo t e cerchiamo la probabilità che esso non sopravviva più di un tempo

dt. Consideriamo cioè

$$\begin{aligned} P(X \in (t, t+dt) \mid X > t) &= \frac{P(X \in (t, t+dt), X > t)}{P(X > t)} \\ &= \frac{P(X \in (t, t+dt))}{P(X > t)} \\ &= \frac{F(t+dt) - F(t)}{dt \cdot \bar{F}(t)}. \end{aligned}$$

Per $dt \to 0$ si ha quindi

$$P(X \in (t, t+dt) \mid X > t) \approx \frac{f(t)}{\bar{F}(t)} dt$$

e pertanto $\lambda(t)$ rappresenta la probabilità condizionata che un oggetto ceda all'età t. Supponiamo ora che la distribuzione del tempo di vita sia esponenziale. In virtù dell'assenza di memoria, si ha che la distribuzione del tempo di vita rimanente per un oggetto di età t è la stessa di quella di un oggetto nuovo, ovvero $\lambda(t)$ è costante:

$$\begin{aligned} \lambda(t) &= \frac{f(t)}{\bar{F}(t)} \\ &= \frac{\lambda e^{-\lambda t}}{e^{-\lambda t}} \\ &= \lambda. \end{aligned}$$

Di conseguenza la funzione di rischio di una variabile esponenziale è costante (λ viene spesso detto *tasso* della distribuzione).
La funzione di rischio determina univocamente la distribuzione F. Infatti esprimendo $\lambda(t)$ come

$$\lambda(t) = \frac{\frac{d}{dt} F(t)}{1 - F(t)}$$

e integrando ambo i membri si ha

$$-\log(1 - F(t)) = \int_0^t \lambda(u) du + k$$

da cui sostituendo con $t = 0$ si trova $k = 0$ e quindi

$$F(t) = 1 - \exp\left\{ -\int_0^t \lambda(u) du \right\}.$$

Se ad esempio una variabile aleatoria continua e positiva ha una funzione di rischio

$$\lambda(t) = a + bt$$

allora la sua funzione di ripartizione è data da

$$F(t) = 1 - e^{-at-bt^2/2}$$

e derivando si ottiene la densità $f(t) = (a + bt)\exp\{-(at + bt^2/2)\}$, $t \geq 0$.

2.2.4 Variabili aleatorie di Weibull

La *distribuzione di Weibull* è utilizzata ampiamente in ambito ingegneristico per la sua versatilità e nell'ambito dei fenomeni di vita, come la distribuzione del tempo di vita di qualche oggetto, in particolare quando si suppone che il suo funzionamento sia legato al funzionamento di una delle sue parti, quella "meno affidabile". In questi contesti, la distribuzione di Weibull fornisce una buona approssimazione della distribuzione del tempo di vita dell'oggetto.
Nel paragrafo precedente si è visto che una legge esponenziale è caratterizzata da una funzione di rischio costante. Vediamo ora un esempio di funzione di rischio non costante. Supponiamo che $\lambda(t)$ sia espressa da

$$\lambda(t) = \frac{\alpha}{\beta^\alpha} t^{\alpha-1}, \quad \alpha > 0, \beta > 0, t \geq 0.$$

Allora, ricordando che

$$F(t) = 1 - \exp\left\{-\int_0^t \lambda(t)dt\right\}.$$

si ottiene

$$\begin{aligned} F(t) &= 1 - \exp\left\{-\frac{\alpha}{\beta^\alpha}\int_0^t t^{\alpha-1}dt\right\} \\ &= 1 - \exp\left\{-\frac{\alpha}{\beta^\alpha}\frac{t^\alpha}{\alpha}\right\} \\ &= 1 - \exp\left\{-\left(\frac{t}{\beta}\right)^\alpha\right\} \end{aligned}$$

Pertanto la funzione di ripartizione di Weibull è del tipo

$$F(t) = 1 - e^{-\left(\frac{t}{\beta}\right)^\alpha}, \quad t \geq 0. \tag{2.31}$$

Una variabile aleatoria T si dice di Weibull di parametri α e β se la sua funzione di distribuzione è data dalla (2.31). Derivando, si trova

$$\begin{aligned} f(t) = \frac{d}{dt}F(t) &= -e^{-\left(\frac{t}{\beta}\right)^\alpha} \cdot (-\alpha)\left(\frac{t}{\beta}\right)^{\alpha-1} \cdot \frac{1}{\beta} \\ &= \frac{\alpha}{\beta^\alpha} t^{\alpha-1} e^{-\left(\frac{t}{\beta}\right)^\alpha}. \end{aligned}$$

La densità

$$f(t) = \frac{\alpha}{\beta^\alpha} t^{\alpha-1} e^{-\left(\frac{t}{\beta}\right)^\alpha}, \quad t \geq 0 \tag{2.32}$$

descrive la *famiglia di distribuzioni di Weibull*. Al variare di $\alpha > 0$ (parametro di forma) e $\beta > 0$ (parametro di scala) si ottengono curve molto diverse tra loro. La distribuzione di Weibull per $\alpha = 1$ corrisponde alla distribuzione esponenziale di parametro $1/\beta$. Media e varianza della distribuzione di Weibull posso essere calcolate per mezzo della funzione gamma $\Gamma(\cdot)$, definita nella sezione seguente dalla (2.34). Si trova che

$$E(T) = \beta\Gamma\left(\frac{1}{\alpha} + 1\right),$$

e

$$V(T) = \beta^2 \left[\Gamma\left(\frac{2}{\alpha} + 1\right) - \left\{\Gamma\left(\frac{1}{\alpha} + 1\right)\right\}^2\right].$$

2.2.5 Variabili aleatorie Gamma

Si è visto come il tempo di attesa per il primo evento in un processo di Poisson segue una distribuzione esponenziale. Si vuole ora considerare il tempo di attesa per il secondo evento. Più specificatamente, si consideri nuovamente un processo di Poisson di parametro λ nell'intervallo unitario e si denoti il tempo di attesa T_2 per il secondo evento. Allora in t unità di tempo ci si aspetterà λt eventi. Se $T_2 \geq t$ allora vi è al più un evento in t unità di tempo, ovvero

$$P(T_2 \geq t) = P(X \leq 1) = \sum_{x=0}^{1} \frac{e^{-\lambda t}(\lambda t)^x}{x!}$$

dove si è usato il fatto che $X \sim Po(\lambda t)$, da cui segue che

$$\begin{aligned} F(t) =& P(T_2 \leq t) = 1 - \sum_{x=0}^{1} \frac{(\lambda t)^x}{x!} e^{-\lambda t} \\ F(t) =& 1 - e^{-\lambda t} - \lambda t e^{-\lambda t}, \\ f(t) =& \frac{d}{dt} F(t) = \lambda^2 t e^{-\lambda t}, \quad t \geq 0. \end{aligned}$$

Generalizzando all'r-ma manifestazione in t unità di tempo si ottiene la funzione di ripartizione per la variabile casuale T_r, indicante il tempo di attesa per l'evento r-mo:

$$F(t) = P(T_r \leq t) = 1 - \sum_{x=0}^{r-1} \frac{(\lambda t)^x}{x!} e^{-\lambda t} \tag{2.33}$$

T_r è detta *variabile aleatoria Erlang* con parametri r e λ. La funzione di densità di T_r è

$$\begin{aligned} f(t) &= \frac{d}{dt} F(t) \\ &= \frac{d}{dt} \left(1 - e^{-\lambda t} - \lambda t e^{-\lambda t} - \frac{(\lambda t)^2}{2!} e^{-\lambda t} - \cdots - \frac{(\lambda t)^{r-1}}{(r-1)!} e^{-\lambda t} \right) \\ &= \lambda e^{-\lambda t} - \lambda e^{-\lambda t} + \lambda^2 t e^{-\lambda t} - \lambda^2 t e^{-\lambda t} + \frac{\lambda^3 t^2}{2!} e^{-\lambda t} - \cdots \\ &\quad - \frac{\lambda^{r-1} t^{r-2}}{(r-2)!} e^{-\lambda t} + \frac{\lambda^r t^{r-1}}{(r-1)!} e^{-\lambda t} \\ &= \frac{\lambda^r t^{r-1}}{(r-1)!} e^{-\lambda t}, \quad t > 0. \end{aligned}$$

La distribuzione Erlang è un caso particolare della più nota *distribuzione Gamma*. Si consideri innanzitutto la funzione $\Gamma : (0, \infty) \to \mathbb{R}$ detta "Gamma di Eulero", definita da:

$$\Gamma(\alpha) = \int_0^\infty u^{\alpha-1} e^{-u} du. \tag{2.34}$$

Integrando per parti $\Gamma(\alpha)$ si ottiene, per $\alpha > 1$,

$$\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$$

e quindi se $\alpha = n$ dove n è un intero positivo, si ottiene

$$\begin{aligned}\Gamma(n) &= (n-1)\Gamma(n-1)\\ &= (n-1)(n-2)\Gamma(n-2)\\ &= \cdots\\ &= (n-1)(n-2)\cdots 2\cdot\Gamma(1)\end{aligned}$$

essendo $\Gamma(1) = \int_0^\infty e^{-x}dx = 1$. Segue che

$$\Gamma(n) = (n-1)! \quad \forall n \in \mathbb{N}$$

La funzione Γ fornisce quindi un'estensione del fattoriale ai numeri reali positivi. Un altro valore notevole si ha per $\alpha = 1/2$; con il cambio di variabile $u = y^2$, si ottiene

$$\Gamma(1/2) = \int_0^{+\infty} u^{(-1/2)}e^{-u}du = \int_0^{+\infty} y^{-1}e^{-y^2}2ydy = 2\int_0^{+\infty} e^{-y^2}dy = \sqrt{\pi},$$

dove si è usato il valore dell'*integrale di Gauss*

$$\int_{-\infty}^{+\infty} e^{-y^2}dy = \sqrt{\pi}$$

Se nella definizione (2.34) operiamo il cambio di variabili $u = \lambda x$ con $\lambda > 0$, otteniamo

$$\Gamma(\alpha) = \int_0^\infty \lambda^\alpha x^{\alpha-1}e^{-\lambda x}dx$$

da cui si evince che

$$1 = \int_0^\infty \frac{\lambda^\alpha x^{\alpha-1}}{\Gamma(\alpha)}e^{-\lambda x}dx$$

Possiamo dare la seguente definizione.

Definizione 2.8 (Variabile aleatoria Gamma)**.** Se una variabile aleatoria X ha funzione di densità data da

$$f(x) = \frac{\lambda^\alpha x^{\alpha-1}}{\Gamma(\alpha)}e^{-\lambda x}, \quad x \geq 0 \tag{2.35}$$

dove $\alpha > 0$ e $\lambda > 0$, si dice che X ha *distribuzione Gamma* di parametri α e λ e si denota con $X \sim Ga(\alpha, \lambda)$.

La quantità λ è il reciproco del *parametro di scala* $\beta = 1/\lambda$, mentre α va sotto il nome di *parametro di forma*.

Vi sono alcuni casi particolari della distribuzione Gamma che riconducono ad altre distribuzioni e sono degni di nota:

- Per $\alpha = n$, intero positivo, la (2.35) coincide con la funzione di densità della distribuzione Erlang, ricavata precedentemente come la densità del tempo di attesa dell'n-esimo evento in un processo di Poisson di parametro λ (in un intervallo di tempo unitario).
- Per $\alpha = 1$ si ottiene la distribuzione esponenziale di parametro λ, $f(x) = \lambda e^{-\lambda x}$, $x \geq 0$.
- La distribuzione Gamma con $\lambda = \frac{1}{2}$ e parametro α uguale a $\frac{1}{2}, 1, \frac{3}{2}, 2, \frac{5}{2}, \ldots$ è detta *distribuzione chi-quadrato* con $\nu = 2\alpha$ gradi di libertà e si denota con χ^2_ν. Questa distribuzione è di grande rilevanza in statistica inferenziale e verrà approfondita più avanti.

Teorema 2.19. Se X ha una distribuzione Gamma con parametri α, λ, allora

$$E(X) = \frac{\alpha}{\lambda}, \quad V(X) = \frac{\alpha}{\lambda^2}, \quad m(t) = \left(\frac{\lambda}{\lambda - t}\right)^\alpha, \; t < \lambda.$$

Dimostrazione.

$$\begin{aligned}
m(t) = E(e^{tX}) &= \int_0^\infty e^{tx}\frac{\lambda^\alpha x^{\alpha-1}}{\Gamma(\alpha)} e^{-\lambda x} dx \\
&= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{\alpha-1} e^{-x(\lambda-t)} dx \\
&= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty \left(\frac{u}{\lambda - t}\right)^{\alpha-1} e^{-u} \frac{du}{\lambda - t}, \quad u = x(\lambda - t) \\
&= \left(\frac{\lambda}{\lambda - t}\right)^\alpha \frac{1}{\Gamma(\alpha)} \int_0^\infty u^{\alpha-1} e^{-u} du \\
&= \left(\frac{\lambda}{\lambda - t}\right)^\alpha .
\end{aligned}$$

Quindi $m'(t) = \lambda^\alpha \alpha (\lambda - t)^{-\alpha-1}$ e $m''(t) = \lambda^\alpha \alpha(\alpha + 1)(\lambda - t)^{-\alpha-2}$ da cui

$$E(X) = m'(0) = \frac{\alpha}{\lambda}$$

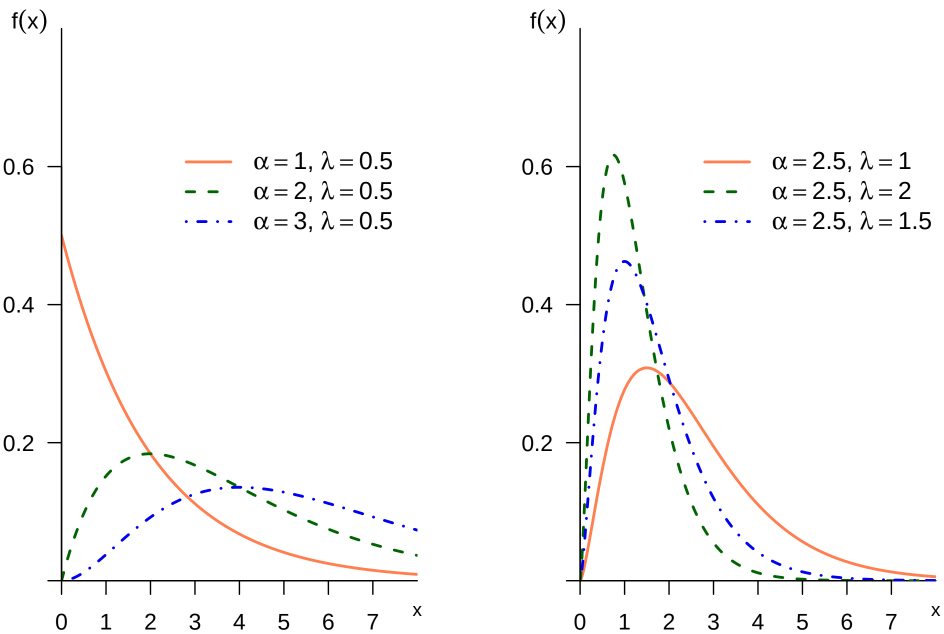


Figura 2.8: Funzione di densità della distribuzione Gamma per $\lambda = 1/2$ al variare di α, intero (a sinistra); densità della distribuzione Gamma per $\alpha = 2.5$ e diversi valori di λ (a destra).

e

$$V(X) = E(X^2) - [E(X)]^2 = m''(0) - \frac{\alpha^2}{\lambda^2} = \frac{\alpha^2+\alpha}{\lambda^2} - \frac{\alpha^2}{\lambda^2} = \frac{\alpha}{\lambda^2}.$$

□

Per $n \geq 1$, il grafico della densità della distribuzione Gamma di parametri (α, λ) ha un unico massimo in corrispondenza del valore $x = (\alpha - 1)/\lambda$. Per $0 < \alpha < 1$ la densità Gamma è illimitata per $x \to 0$.

Esempio 2.8.

Si assuma che le telefonate ad un numero verde gratuito seguano una distribuzione di Poisson con una media di 120 all'ora. Si consideri l'intervallo di tempo tra le 9:00 e le 12:00. Sia T_{10} l'istante nel quale viene ricevuta la decima telefonata, a partire dalle ore 9:00, utilizzando i minuti come unità temporale. Allora T_{10} ha distribuzione Erlang con $r = 10$ e $\lambda = 120/60 = 2$ e si ha

$$E(T_{10}) = \frac{10}{2} = 5,$$

quindi ci si aspetta che la decima chiamata arrivi alle ore 9:05. La probabilità che la decima chiamata arrivi prima delle 9:05 è data da

$$P(T_{10} < 5) = 1 - \sum_{x=0}^{9} \frac{10^x}{x!} e^{-10} = 0.542.$$

2.2.6 Variabile aleatoria normale (o gaussiana)

La distribuzione normale gioca un ruolo fondamentale in teoria della probabilità ed è alla base di molte delle tecniche usate in statistica applicata. La sua importanza deriva dal fatto che la distribuzione normale sembra essere un modello ragionevole del comportamento di molti fenomeni casuali e per questo è una delle distribuzioni che viene maggiormente impiegata nelle applicazioni. Inoltre, come si vedrà più avanti, il Teorema del Limite Centrale rafforza il ruolo della distribuzione normale, come distribuzione approssimata della somma di variabili aleatorie indipendenti, non necessariamente con la stessa distribuzione.

Per introdurre la distribuzione normale si consideri l'esempio di errori commessi misurando una quantità fisica. Si vedrà che, sotto determinate assunzioni, una distribuzione normale con opportuni parametri si presta come modello per l'errore di misurazione complessivo. Si assuma allora che l'errore finale della nostra misurazione sia una variabile aleatoria Y, somma di n errori incogniti da cui ciascuna misura è affetta, $X_1, \dots, X_n$, cioè $Y = \sum_{i=1}^n X_i$. Si assuma inoltre che ogni errore X_i possa assumere solo due valori, 1 o -1, con eguale probabilità; allora le X_i sono variabili aleatorie di Bernoulli *simmetriche* e

$$P(X_i = -1) = P(X_i = 1) = \frac{1}{2}, \quad i = 1, 2, \dots, n$$

e

$$E(X_i) = 0, \quad V(X_i) = 1.$$

Si noti che se V è una variabile aleatoria di Bernoulli di parametro $p = \frac{1}{2}$, allora $X = 2V - 1$ sarà pari a 1 se $V = 1$ e pari a -1 se $V = 0$. Quindi le variabili di Bernoulli simmetriche sono semplici funzioni lineari di v.a. Bernoulliane con $p = 1/2$. Possiamo esprimere Y nel seguente modo

$$Y = \sum_{i=1}^n X_i = \sum_{i=1}^n (2V_i - 1) = 2\sum_{i=1}^n V_i - n$$

in termini delle $V_1, \dots, V_n$. Sotto l'assunzione che queste ultime siano definite su prove di Bernoulli indipendenti, la loro somma $X = \sum_{i=1}^n V_i$ ha distribuzione binomiale di parametri n e $p = 1/2$, per cui possiamo considerare la variabile aleatoria $Y = 2X - n$, con X binomiale, e derivare la funzione di probabilità per Y come segue

$$\begin{aligned} P(Y = y) = P(2X - n = y) &= P\left(X = \frac{n+y}{2}\right) \\ &= \binom{n}{(n+y)/2}\left(\frac{1}{2}\right)^y \left(\frac{1}{2}\right)^{n-y} \\ &= \binom{n}{(n+y)/2}\frac{1}{2^n}, \quad y = -n, -n+2, -n+4, \dots, n-2, n. \end{aligned}$$

Possiamo inoltre scrivere

$$E(Y) = E(2X - n) = 2E(X) - n = 2\left(\frac{n}{2}\right) - n = 0$$

e

$$V(Y) = V(2X - n) = 4V(X) = 4n\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = n$$

All'aumentare di n, numero degli errori la cui somma dà Y, la distribuzione di Y rimane centrata in 0, dal momento che $E(Y) = 0$ per ogni n, ma aumenta la dispersione, poiché $V(X) = n$ aumenta al crescere di n. Allora, denotato con Z l'errore "standardizzato" $Z = Y/\sqrt{n}$, si avrà

$$E(Z) = 0 \quad \text{e} \quad V(Z) = 1 \quad \text{per ogni } n.$$

Cosa accade alla distribuzione di Z per n che cresce all'infinito? In particolare, cosa succede alla probabilità $P(z \leq Z \leq z + \Delta z)$, dove z è un qualsiasi valore fissato e $\Delta z > 0$ è un piccolo incremento? Si dimostra che al limite per $n \to \infty$ vale la relazione approssimata

$$P(z \leq Z \leq z + \Delta z) \doteq \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \Delta z,$$

La funzione

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad -\infty < z < +\infty. \tag{2.36}$$

è detta *funzione di densità normale standard*. È una funzione *simmetrica* rispetto a $z = 0$, assume la tipica forma a "campana", è centrata in 0, ed ha come asintoto l'asse delle x. Inoltre se Z ha distribuzione normale standard, allora

$$E(Z) = 0 \quad \text{e} \quad V(Z) = 1.$$

Definizione 2.9 (Variabile aleatoria normale standard)**.** Z è detta *variabile aleatoria normale standard* se la sua funzione di densità è espressa da

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad -\infty < z < +\infty, \tag{2.37}$$

e la sua funzione di ripartizione è data da

$$\Phi(z) = P(Z \leq z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \tag{2.38}$$

Alla funzione di ripartizione e alla densità di una variabile casuale normale standard si riservano i simboli $\Phi(\cdot)$ e $\phi(\cdot)$, rispettivamente. Nel seguito si

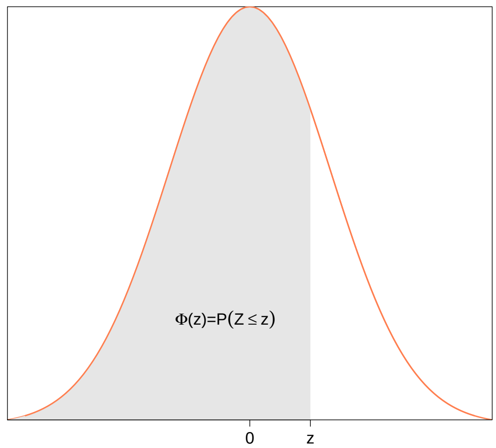


Figura 2.9: Funzione di densità della distribuzione normale standard. L'area in grigio è il valore della funzione di ripartizione $\Phi(z)$ nel punto z.

userà la notazione $\mathcal{N}(0,1)$ per denotare la distribuzione (2.37).
La (2.38) non è rappresentata da una forma esplicita, ma va valutata adottando tecniche di calcolo numerico. Per questo, sono state predisposte delle tavole (riportate in Appendice) che ne danno il livello per un insieme sufficientemente ampio di valori di z non negativi. Si noti che, in virtù della simmetria della densità normale vale la relazione

$$\Phi(z) = 1 - \Phi(-z)$$

che si dimostra utile in più situazioni.

Per affermare che la (2.37) è una funzione di densità deve essere

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = 1.$$

La dimostrazione può essere effettuata come segue.
Supponiamo di rappresentare con A l'area sotto la curva, quindi definiamo

$$A = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx.$$

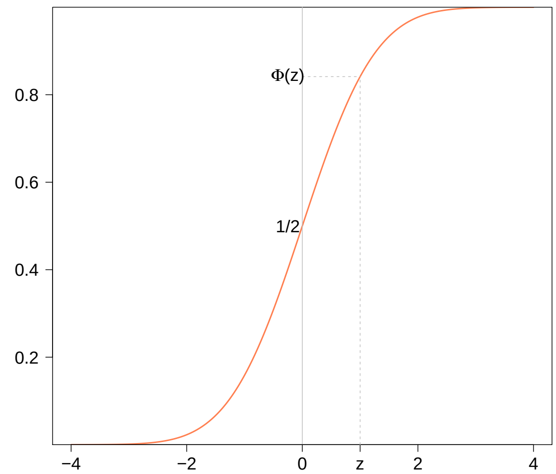


Figura 2.10: Funzione di ripartizione della distribuzione normale standard.

Per dimostrare che $A = 1$, dimostriamo che $A^2 = 1$ dato che

$$\begin{aligned} A^2 &= \left\{ \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \right\} \left\{ \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy \right\} \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)/2}\, dx dy \end{aligned}$$

dove si è scritto il prodotto di due integrali come un integrale doppio. Cambiamo le variabili in coordinate polari ponendo

$$x = r\cos\theta, \quad y = r\sin\theta,$$

da cui $x^2 + y^2 = r^2$, $0 < r < \infty$, $0 < \theta < 2\pi$; l'integrale diventa

$$\begin{aligned} A^2 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)/2}\, dx dy \\ &= \frac{1}{2\pi} \int_{0}^{\infty} \left\{ \int_{0}^{2\pi} r e^{-r^2/2}\, d\theta \right\} dr \\ &= \int_{0}^{\infty} r e^{-r^2/2}\, dr \\ &= \Gamma(1) = 1, \end{aligned}$$

essendo $\int_0^{2\pi} re^{-r^2/2}\,d\theta = 2\pi re^{-r^2/2}$.

Sia Z una variabile aleatoria normale standard. Allora Z ha media 0 e varianza 1. Definiamo

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

dove $\sigma > 0$ e μ sono valori arbitrari, da cui

$$Z = \frac{X - \mu}{\sigma}.$$

Allora il valore atteso e la varianza di X sono date da

$$E(X) = \sigma E(Z) + \mu = \mu, \quad V(X) = \sigma^2 V(Z) = \sigma^2,$$

per cui la deviazione standard di X è la costante σ. La funzione di ripartizione della variabile aleatoria X si scrive come

$$F(x) = P(\sigma Z + \mu \leq x) = P\left(Z \leq \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

Derivando rispetto ad x si deduce che la densità di X è

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

Possiamo definire formalmente la distribuzione normale di media e deviazione standard arbitrarie come segue.

Definizione 2.10 (Variabile aleatoria normale (o gaussiana) di media μ e varianza σ^2)**.** Se X è una variabile aleatoria con densità data da

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-(x-\mu)^2/2\sigma^2}, \tag{2.39}$$

dove $\sigma > 0$, $\mu \in \mathbb{R}$, allora X ha *distribuzione normale (o gaussiana)* di media μ e varianza σ^2, e si denota con la simbologia $\mathcal{N}(\mu, \sigma^2)$. Il parametro μ è anche moda e mediana della distribuzione e σ è la deviazione standard (o scarto quadratico medio). La funzione di ripartizione di X è espressa da

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}}e^{-(y-\mu)^2/2\sigma^2}dy. \tag{2.40}$$

La (2.39) gode delle seguenti proprietà:

- è simmetrica, avendo come asse di simmetria la retta $x = \mu$;
- è crescente nell'intervallo $(-\infty, \mu)$ e decrescente nell'intervallo $(\mu, +\infty)$;
- ha due punti di flesso in $x = \mu - \sigma$ e $x = \mu + \sigma$;
- è concava nell'intervallo $(\mu - \sigma, \mu + \sigma)$ e convessa altrove;
- ha come asintoto l'asse delle x.

Si osservi che è possibile dimostrare che il rapporto $\beta_2 = \overline{m}_4/\sigma^4$ introdotto in Sezione 1.2 è uguale a 3 nella distribuzione normale, da cui si evince come l'indice di curtosi $\gamma_2 = \beta_2 - 3$ segnali lo scostamento dalla normalità, essendo pari a 0 per ogni distribuzione normale. Inoltre, data la simmetria della distribuzione normale, sono nulli i momenti centrati di ordine r, con r intero positivo dispari.

Le curve della densità normale sono centrate nel punto della media, ma presentano diversi livelli di variabilità: al crescere di σ la curva diviene sempre più schiacciata presentando una dispersione crescente rispetto alla media.

È utile enunciare e dimostrare formalmente il seguente risultato.

Teorema 2.20. Se X è una variabile casuale normale, allora

$$E(X) = \mu, \quad V(X) = \sigma^2, \quad m(t) = e^{t\mu + \sigma^2 t^2/2}.$$

Dimostrazione.

$$\begin{aligned} m(t) &= E(e^{tX}) = e^{t\mu} E(e^{t(X-\mu)}) \\ &= e^{t\mu} \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{t(x-\mu)} e^{-(x-\mu)^2/2\sigma^2} dx \\ &= e^{t\mu} \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(1/2\sigma^2)[(x-\mu)^2 - 2\sigma^2 t(x-\mu)]} dx. \end{aligned}$$

Se completiamo il quadrato entro parentesi, diventa

$$\begin{aligned} (x-\mu)^2 - 2\sigma^2 t(x-\mu) &= (x-\mu)^2 - 2\sigma^2 t(x-\mu) + \sigma^4 t^2 - \sigma^4 t^2 \\ &= (x - \mu - \sigma^2 t)^2 - \sigma^4 t^2 \end{aligned}$$

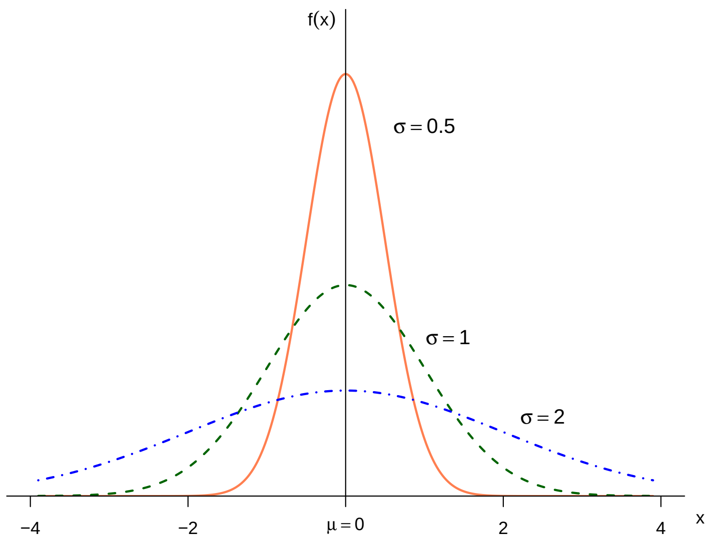


Figura 2.11: Distribuzione normale con media $\mu = 0$ e diversi valori della deviazione standard σ.

e si ha

$$\begin{aligned} m(t) &= e^{t\mu} \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(1/2\sigma^2)[(x-\mu-\sigma^2 t)^2 - \sigma^4 t^2]} dx \\ &= e^{t\mu} e^{\sigma^2 t^2/2} \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(x-\mu-\sigma^2 t)^2/2\sigma^2} dx \\ &= e^{t\mu} e^{\sigma^2 t^2/2}, \end{aligned}$$

dal momento che

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(x-\mu-\sigma^2 t)^2/2\sigma^2} dx = 1$$

perché è l'area sotto una distribuzione normale con media $\mu + \sigma^2 t$ e varianza σ^2. Quindi

$$m(t) = e^{t\mu + \sigma^2 t^2/2}. \tag{2.41}$$

Differenziando due volte $m(t)$ si ha

$$\begin{aligned} m'(t) &= e^{t\mu + \sigma^2 t^2/2} (\mu + \sigma^2 t) \\ m''(t) &= e^{t\mu + \sigma^2 t^2/2} (\mu + \sigma^2 t)^2 + \sigma^2 e^{t\mu + \sigma^2 t^2/2} \end{aligned}$$

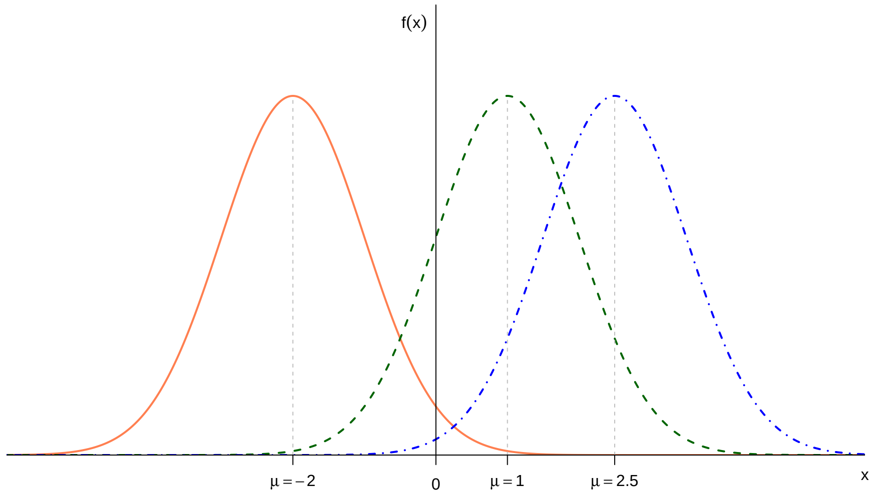


Figura 2.12: Distribuzione normale per $\sigma = 1$ e alcuni valori di μ.

e sostituendo $t = 0$ si trova

$$E(X) = m'(0) = \mu$$

e

$$V(X) = m''(0) - \mu^2 = \mu^2 + \sigma^2 - \mu^2 = \sigma^2.$$

□

Si consideri ora il problema della determinazione della probabilità della variabile aleatoria in un dato intervallo (a, b). Si ha

$$P(a \leq X \leq b) = \int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} dx, \tag{2.42}$$

da cui, posto $z = (x - \mu)/\sigma$ e quindi $dx = \sigma dz$, si ottiene

$$\begin{aligned} P(a \leq X \leq b) &= \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} \frac{1}{\sigma\sqrt{2\pi}} e^{-z^2/2} \sigma dz \\ &= \int_{-\infty}^{\frac{b-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz - \int_{-\infty}^{\frac{a-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \\ &= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right), \end{aligned}$$

dove $\Phi(\cdot)$ denota la funzione di ripartizione della normale standard.

Esempio 2.9.

Spesso i risultati di un test a risposta multipla vengono descritti utilizzando una distribuzione normale. Il docente utilizza i risultati del test per stimare i parametri normali μ e σ^2 e assegna la lettera A a quelli che hanno un punteggio superiore a $\mu+\sigma$, B a quelli che hanno un punteggio tra μ e $\mu+\sigma$, C a quelli che hanno un punteggio tra $\mu-\sigma$ e μ, D a quelli il cui punteggio è compreso tra $\mu-2\sigma$ e $\mu-\sigma$ e insufficiente a quelli che ottengono un punteggio inferiore a $\mu-2\sigma$.
Dato che

$$\begin{aligned} P(X>\mu+\sigma) &= P\left(\frac{X-\mu}{\sigma}>1\right) = 1-P\left(\frac{X-\mu}{\sigma}\leq 1\right) \\ &= 1-\Phi(1) \approx 0.1587 \end{aligned}$$

$$P(\mu<X<\mu+\sigma) = P\left(0<\frac{X-\mu}{\sigma}<1\right) = \Phi(1)-\Phi(0) \approx 0.3413$$

$$\begin{aligned} P(\mu-\sigma<X<\mu) &= P\left(-1<\frac{X-\mu}{\sigma}<0\right) = \Phi(0)-\Phi(-1) \\ &= \Phi(0)-1+\Phi(1) \approx 0.3413 \end{aligned}$$

$$\begin{aligned} P(\mu-2\sigma<X<\mu-\sigma) &= P\left(-2<\frac{X-\mu}{\sigma}<-1\right) = \Phi(-1)-\Phi(-2) \\ &= \Phi(2)-\Phi(1) \approx 0.1359 \end{aligned}$$

$$P(X<\mu-2\sigma) = P\left(\frac{X-\mu}{\sigma}<-2\right) = \Phi(-2) = 1-\Phi(2) \approx 0.0228$$

segue che il 16% circa della classe otterrà A, il 34% B, 34% C, e il 14% D; il 2% sarà insufficiente.

Esempio 2.10.

In alcune situazioni si conosce la funzione di ripartizione o si possiedono informazioni per risalire a questa e si richiede di determinare il valore corrispondente di Z o della variabile non standard X.
Si supponga, ad esempio, che in una data città la spesa mensile per i generi alimentari per le famiglie composte da quattro persone sia pari a 700 euro,

con una deviazione standard di 80 euro. Si assuma che tale spesa abbia distribuzione normale. Volendo determinare la spesa mensile massima sostenuta dal 3% delle famiglie più parsimoniose, occorre individuare il valore di X tale per cui la funzione di ripartizione nel punto x vale 0.03. Ora, essendo 0.03 minore di 0.5, il punto x si trova a sinistra della media e ciò significa che il corrispondente valore standardizzato di x è negativo. Disponendo delle tavole che riportano $\Phi(z)$ per valori positivi di z, si può ricorrere alla relazione $\Phi(z) = 1 - \Phi(-z)$. Nel nostro caso, posto $\Phi(-z) = 0.03$ si ottiene

$$\Phi(z) = 1 - \Phi(-z) = 0.97,$$

ovvero occorre ricercare all'interno delle tavole, la quantità 0.97 o la quantità ad essa più prossima e individuare il corrispondente valore di z. In definitiva, si trova che il valore di $\Phi(z)$ più prossimo a 0.97 è 0.9699 a cui corrisponde $z = 1.88$. Ciò significa che in valori standard la spesa del 3% delle famiglie più parsimoniose non supera $-z = -1.88$. La quantità cercata si trova risolvendo

$$-1.88 = \frac{x - 700}{80}$$

da cui $x = 549.6$ euro.

Si noti che, nella soluzione di problemi inversi del tipo di quello descritto, può verificarsi spesso che non sia disponibile nella tavola il livello esatto della funzione di ripartizione. Ad esempio, volendo determinare b tale che $\Phi(b) = 0.75$, ovvero il 75-esimo percentile della distribuzione, si ricorre alle tavole trovando

$$\Phi(0.67) = 0.7486, \quad \Phi(0.68) = 0.7517$$

da cui si deduce che b si trova tra 0.67 e 0.68. Si può scegliere, in questo caso, il valore più vicino a quello cercato, per cui si adotterà $b = 0.67$ come soluzione del quesito iniziale. Volendo essere ancora più precisi si può ricorrere all'interpolazione lineare. In generale, sia Φ il valore assegnato alla funzione di ripartizione e siano Φ' e Φ'' due valori contigui di $\Phi(z)$ tali che $\Phi' < \Phi < \Phi''$. Tali valori si trovano quindi nella stessa riga e corrispondono ad un certo z ed al successivo $z + 0.01$. Indicando con z' e z'' i valori di z che corrispondono a Φ' e Φ'', la quantità cercata risulta

$$z = z' + \frac{z'' - z'}{\Phi'' - \Phi'}(\Phi - \Phi').$$

L'importanza della normale nella teoria della probabilità è dovuta in gran parte ai cosiddetti *teoremi del limite centrale*. Tali teoremi stabiliscono le condizioni sotto le quali la somma di variabili casuali tende alla normale all'aumentare del numero delle variabili, come sarà approfondito più avanti.

2.2.7 Variabili aleatorie Beta

Una famiglia di densità di probabilità di variabili casuali continue a valori nell'intervallo $(0,1)$ è la famiglia di distribuzioni *Beta*.

Uno dei possibili modi per arrivare alla definizione della distribuzione Beta è considerare $X_1,\ldots,X_n$ variabili aleatorie uniformi nell'intervallo $(0,1)$. Fissato un punto t nell'intervallo $(0,1)$, si assuma che gli eventi $A_i=\{X_i\le t\}$, per $i=1,\ldots,n$, siano indipendenti. Tali variabili aleatorie definiscono quindi prove Bernoulliane indipendenti, dove il successo nella i-ma prova equivale al verificarsi dell'evento A_i, l'insuccesso al verificarsi dell'evento $\overline{A}_i$. Si ha quindi

$$P(A_i)=P(X_i\le t)=t,\quad i=1,\ldots,n.$$

Si denoti con Y_r l'r-mo più grande degli X_i e si consideri l'evento $\{Y_r>t\}$. Ciò si verifica se e solo se vi sono al più $r-1$ tra gli $X_1,\ldots,X_n$ più piccoli di t, ovvero se e solo se osserviamo $r-1$ o meno successi nelle n prove Bernoulliane. Poiché il numero di successi in n prove Bernoulliane indipendenti, ciascuna con probabilità di successo t, è binomiale di parametri n e t. Quindi

$$P(Y_r>t)=\sum_{k=0}^{r-1}\binom{n}{k}t^k(1-t)^{n-k},\quad 0<t<1$$

La funzione di ripartizione per Y_r è data da

$$\begin{aligned}F(t)&=1-\sum_{k=0}^{r-1}\binom{n}{k}t^k(1-t)^{n-k}\\&=\sum_{k=r}^{n}\binom{n}{k}t^k(1-t)^{n-k}\\&=\binom{n}{r}t^r(1-t)^{n-r}+\binom{n}{r+1}t^{r+1}(1-t)^{n-r-1}+\cdots+nt^{n-1}(1-t)+t^n\end{aligned}$$

e la funzione di densità è

$$\begin{aligned}
f(t) &= \frac{d}{dt}F(t) \\
&= \binom{n}{r} rt^{r-1}(1-t)^{n-r} - \binom{n}{r}(n-r)t^r(1-t)^{n-r-1} \\
&+ \binom{n}{r+1}(r+1)t^r(1-t)^{n-r-1} \\
&- \binom{n}{r+1}(n-r-1)t^r(1-t)^{n-r-2} \\
&+ \cdots + n(n-1)t^{n-2}(1-t) - nt^{n-1} + nt^{n-1} \\
&= \binom{n}{r} rt^{r-1}(1-t)^{n-r} = \frac{n!}{(r-1)!(n-r)!}t^{r-1}(1-t)^{n-r}, \quad 0 < t < 1.
\end{aligned}$$

Si osservi che le n v.a.

$$Y_1 \leq Y_2 \leq \cdots \leq Y_n$$

ottenute disponendo le X_i in ordine non decrescente, definiscono le cosiddette *statistiche d'ordine* corrispondenti alle variabili aleatorie $X_1, \ldots, X_n$ (denotate anche con $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$) dove $Y_1 \equiv \min(X_1, \ldots, X_n)$, $Y_n \equiv \max(X_1, \ldots, X_n)$ e, per ogni k tra 1 ed n, Y_k è la k-ma più grande determinazione fra $X_1, \ldots, X_n$.

Teorema 2.21. Sia Y_r l'r-ma determinazione più grande tra n variabili aleatorie indipendenti distribuite uniformemente in $(0, 1)$, $r = 1, \ldots, n$, allora la funzione di densità di Y_r è

$$f(t) = \frac{\Gamma(n+1)}{\Gamma(r)\Gamma(n-r+1)}t^{r-1}(1-t)^{n-r}, \quad 0 < t < 1, \tag{2.43}$$

ricordando che $\Gamma(a) = (a-1)!$, per a intero positivo.

Quello illustrato è un esempio particolare di *distribuzione Beta*.

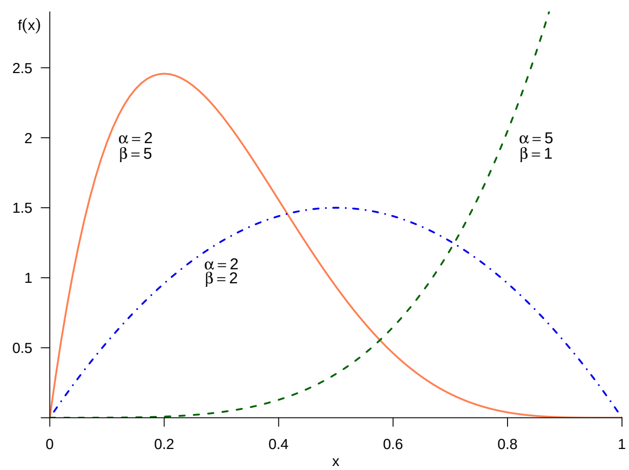


Figura 2.13: Funzione di densità della distribuzione Beta per diversi valori dei parametri α e β.

Definizione 2.11 (Variabile aleatoria Beta)**.** Se una variabile casuale X ha una densità data da

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \quad 0 < x < 1. \tag{2.44}$$

dove $\alpha, \beta > 0$, si dice che ha una *distribuzione Beta*, e si denoterà con $X \sim Beta(\alpha, \beta)$.

Per $\alpha = r, \beta = n - r + 1$, si ottiene la (2.43). Si dimostra inoltre che

$$\int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \tag{2.45}$$

per ogni $\alpha > 0, \beta > 0$, e pertanto $f(x)$ è una funzione di densità. Si osservi che l'integrale che compare al primo membro della (2.45) è detta *funzione Beta*, legata alla funzione Gamma dalla relazione

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

Si noti che se $\alpha = \beta = 1$, la distribuzione Beta si trasforma nella distribuzione uniforme su $(0, 1)$.

La funzione generatrice dei momenti della distribuzione Beta non ha una forma semplice; si possono però calcolare con facilità i momenti usando direttamente la definizione.

Teorema 2.22. Se X è una variabile casuale con funzione di densità Beta, allora

$$E(X) = \frac{\alpha}{\alpha + \beta}, \qquad V(X) = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}.$$

Dimostrazione.

$$\begin{aligned}
E(X^k) &= \int_0^1 x^k \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} dx \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 x^{k+\alpha-1}(1-x)^{\beta-1} dx \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} B(k+\alpha, \beta) \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(k+\alpha)\Gamma(\beta)}{\Gamma(k+\alpha+\beta)} \\
&= \frac{\Gamma(\alpha+\beta)\Gamma(k+\alpha)}{\Gamma(\alpha)\Gamma(k+\alpha+\beta)}
\end{aligned}$$

da cui si ottiene

$$E(X) = \frac{\Gamma(\alpha+\beta)\Gamma(1+\alpha)}{\Gamma(\alpha)\Gamma(1+\alpha+\beta)} = \frac{(\alpha+\beta-1)\Gamma(\alpha+\beta-1)\alpha\Gamma(\alpha)}{\Gamma(\alpha)(\alpha+\beta)(\alpha+\beta-1)\Gamma(\alpha+\beta-1)} = \frac{\alpha}{\alpha+\beta}$$

e

$$\begin{aligned}
V(X) = E(X^2) - [E(X)]^2 &= \frac{\Gamma(\alpha+\beta)\Gamma(2+\alpha)}{\Gamma(\alpha)\Gamma(2+\alpha+\beta)} - \left(\frac{\alpha}{\alpha+\beta}\right)^2 \\
&= \frac{(\alpha+1)\alpha}{(1+\alpha+\beta)(\alpha+\beta)} - \left(\frac{\alpha}{\alpha+\beta}\right)^2 \\
&= \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.
\end{aligned}$$

$\square$

2.2.8 Altri modelli continui

In questo paragrafo vengono introdotte altre famiglie parametriche di funzioni di densità di probabilità degne di nota; le introduzioni delle tre famiglie di distribuzioni conosciute col nome di distribuzione lognormale, distribuzione di Student, distribuzione chi-quadrato, e distribuzione F, vengono demandate ai capitoli successivi.

Distribuzione di Cauchy

La distribuzione di Cauchy ha funzione di densità

$$f(x) = \frac{1}{\pi\beta\{1 + [(x-\alpha)/\beta]^2\}}, \quad x \in \mathbb{R} \tag{2.46}$$

dove $-\infty < \alpha < +\infty$ e $\beta > 0$. Sebbene la densità di Cauchy sia simmetrica rispetto al parametro α, non ne esiste né la media né i momenti di ordine più elevato. La funzione di ripartizione è

$$\begin{aligned} F(x) &= \frac{1}{\pi}\int_{-\infty}^{x} \frac{1}{\pi\beta\{1 + [(u-\alpha)/\beta]^2\}} du \\ &= \frac{1}{2} + \frac{1}{\pi}\arctan\frac{x-\alpha}{\beta}. \end{aligned} \tag{2.47}$$

Distribuzione logistica

La distribuzione logistica è data in forma di funzione di ripartizione da

$$F(x) = \frac{1}{1 + e^{-(x-\alpha)/\beta}}, \tag{2.48}$$

dove $-\infty < \alpha < +\infty$ e $\beta > 0$. La media e la varianza della distribuzione sono

$$E(X) = \alpha, \quad V(X) = \frac{\beta^2\pi^2}{3}.$$

Si noti che la densità della logistica è simmetrica rispetto ad α. La distribuzione logistica è stata usata nei modelli dei livelli di tolleranza in problemi relativi a esperimenti biologici.

Capitolo 3

Cenni su modelli statistici multidimensionali

Nel capitolo precedente sono state trattate distribuzioni di probabilità di una sola variabile aleatoria, originata dall'associazione di un solo numero reale ad ogni evento elementare dello spazio campionario. Tale nozione può essere estesa al caso di due o più dimensioni.

3.1 Variabili aleatorie doppie discrete e continue

Si consideri l'esempio del lancio di due dadi, e si associ a ciascun evento elementare una coppia di numeri (x, y) il primo dei quali è la differenza in valore assoluto dei numeri che appaiono sulle due facce, ed il secondo è la somma dei medesimi numeri. Così all'evento elementare (1,6) è associata la coppia di numeri $(5, 7)$. Ad ogni coppia distinta di numeri è possibile associare un livello di probabilità, che è pari alla probabilità dell'evento o degli eventi elementari che producono tale coppia. Ovvero, indicano con X e Y le due variabili aleatorie, la probabilità $P(X = x, Y = y)$ risulta dalla somma delle probabilità degli eventi elementari per i quali si realizzano le uguaglianze $X = x$ e $Y = y$.

Definizione 3.1 (Variabile aleatoria doppia discreta)**.** Dato uno spazio campionario discreto, Ω, si chiama *variabile aleatoria doppia discreta* la funzione (X, Y) definita in $\mathbb{R}^2$ che associa ad ogni evento elementare ω dello spazio campionario la coppia di numeri reali (x, y), essendo $x = X(\omega)$ e $y = Y(\omega)$.

Ad ogni coppia (x, y) è possibile assegnare un livello di probabilità espresso dalla **funzione di probabilità congiunta**

$$p(x, y) = P(\{X = x\} \cap \{Y = y\}) = P(X = x, Y = y). \tag{3.1}$$

La (3.1) gode delle due evidenti proprietà

$$p(x, y) \geq 0, \quad \sum_x \sum_y p(x, y) = 1,$$

dove, nella seconda relazione le somme sono estese, rispettivamente, a tutti i possibili valori di X e a tutti i possibili valori di Y. La probabilità che la variabile (X, Y) assuma valori in un qualsiasi sottoinsieme A è data allora da

$$P((X, Y) \in A) = \sum_{(x,y) \in A} \sum p(x, y).$$

In particolare **la funzione di ripartizione congiunta** è espressa da

$$\begin{aligned} F(x, y) &= P(X \leq x, Y \leq y) \\ &= \sum_{u \leq x} \sum_{v \leq y} p(u, v). \end{aligned} \tag{3.2}$$

La (3.2), analogamente al caso di una sola variabile aleatoria, gode delle seguenti proprietà:

1. $F(x, y) \leq F(x', y)$ se $x \leq x'$; $F(x, y) \leq F(x, y')$ se $y \leq y'$;
2. $F(-\infty, y) = \lim_{x \to -\infty} F(x, y) = 0$ e $F(x, -\infty) = \lim_{y \to -\infty} F(x, y) = 0$;
3. $F(\infty, \infty) = \lim_{x,y \to \infty} F(x, y) = 1$;
4. $\lim_{h \to 0^+} F(x + h, y) = \lim_{h \to 0^+} F(x, y + h) = F(x, y)$, per ogni x e y.

La densità discreta di X può essere ottenuta da $p(x, y)$ tramite

$$p_X(x) = P(X = x) = \sum_y p(x, y) \tag{3.3}$$

e in maniera analoga

$$p_Y(y) = P(Y = y) = \sum_x p(x, y) \tag{3.4}$$

Le (3.3) e (3.4) definiscono le **funzioni di probabilità marginali** di X e Y rispettivamente. Infine, a partire dalle funzioni di probabilità marginali possiamo calcolare il valore atteso di X e Y, rispettivamente, come

$$E(X) = \sum_x x p_X(x) \tag{3.5}$$

e

$$E(Y) = \sum_y y p_Y(y) \tag{3.6}$$

Nel caso continuo si ha la seguente definizione.

Definizione 3.2 (Variabile aleatoria doppia continua)**.** Una variabile aleatoria doppia (X, Y) si dice *continua* se ad ogni sottoinsieme A dello spazio delle coppie di numeri reali è associata la probabilità

$$P((X, Y) \in A) = \iint_{(x,y)\in A} f(x, y)\, dx\, dy \tag{3.7}$$

essendo $f(x, y)$ la funzione di densità congiunta della variabile aleatoria.

La funzione di densità $f(x, y)$ gode delle due proprietà

$$f(x, y) \geq 0, \qquad \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, dx\, dy = 1$$

Come nel caso discreto, la funzione di ripartizione dà la probabilità dell'evento $\{X \leq x, Y \leq y\}$; sostituendo gli integrali alle somme si ha

$$F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u, v) du\, dv \tag{3.8}$$

Tale funzione ha le medesime proprietà viste per il caso discreto. Vale inoltre la relazione

$$f(x,y) = \frac{\partial^2}{\partial x \partial y} F(x,y).$$

Se X e Y sono *congiuntamente continue*, ovvero esiste una funzione $f(x,y)$ integrabile tale da verificare la (3.7), esse lo saranno anche individualmente. Per un qualunque sottoinsieme A della retta reale si ha

$$\begin{aligned} P(X \in A) &= P(X \in A, Y \in (-\infty, \infty)) \\ &= \int_A \int_{-\infty}^{\infty} f(x,y)\, dx\, dy \\ &= \int_A f_X(x)\, dx \end{aligned}$$

dove

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\, dy$$

è la **densità marginale** della variabile aleatoria X. In maniera analoga, la densità marginale di Y è data da

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\, dx.$$

In questo caso si ha

$$E(X) = \int_{-\infty}^{\infty} x f_X(x); \quad E(Y) = \int_{-\infty}^{\infty} y f_Y(y).$$

Data una variabile aleatoria doppia, è utile considerare la distribuzione di probabilità di Y per livelli assegnati di X, o viceversa. Quindi posto $X = x$ si vogliono determinare le probabilità $P(Y = y|X = x)$ per tutti i valori di Y. Siano X, Y discrete con funzione di probabilità $p(x,y)$. Ricorrendo alla definizione di probabilità condizionata si ottiene la **funzione di probabilità condizionata** di Y

$$P(Y = y|X = x) = p(y|x) = \frac{P(Y = y \cap X = x)}{P(X = x)} = \frac{p(x,y)}{p_X(x)} \tag{3.9}$$

per tutti i valori di x per cui $p_X(x) > 0$, e la funzione di probabilità condizionata di X dato $Y = y$

$$P(X = x|Y = y) = p(x|y) = \frac{P(X = x \cap Y = y)}{P(Y = y)} = \frac{p(x,y)}{p_Y(y)} \tag{3.10}$$

per tutti i valori di y per cui $p_Y(y) > 0$.

Le (3.9) e (3.10) implicano che i rispettivi denominatori siano non nulli. È inoltre immediato verificare che $\sum_y p(y|x) = \sum_x p(x|y) = 1$. La *funzione di ripartizione condizionata* di Y dato $X = x$ si definisce come

$$F(y|x) = P(Y \leq y|X = x) = \sum_{v \leq y} p(v|x),$$

e analogamente

$$F(x|y) = P(X \leq x|Y = y) = \sum_{u \leq x} p(u|y).$$

Nel caso continuo si procede nello stesso modo, sostituendo l'integrale alla sommatoria. Pertanto si ottengono le **funzioni di densità condizionate**

$$f(y|x) = \frac{f(x,y)}{f_X(x)}, \quad f(x|y) = \frac{f(x,y)}{f_Y(y)}, \tag{3.11}$$

dove $f(x,y)$ è la densità congiunta della variabile continua (X,Y) e $f_X(x)$, $f_Y(y)$ sono le densità marginali di X e Y rispettivamente.
Attraverso la densità condizionata possiamo definire la probabilità condizionata di eventi associati a una variabile quando è noto il valore di una seconda variabile; cioè per ogni sottoinsieme A della retta reale

$$P(X \in A|Y = y) = \int_A f(x|y)dx$$

Quindi, ponendo $A = (-\infty, a]$ possiamo definire la *funzione di ripartizione condizionata* di X dato $Y = y$ come

$$F(a|y) = P(X \leq a|Y = y) = \int_{-\infty}^{a} f(x|y)dx.$$

La nozione di distribuzione condizionata dà anche la possibilità di definire in modo naturale il concetto di *indipendenza* tra variabili casuali. Si ammetta che la funzione di probabilità (o densità) condizionata di Y dato $X = x$ non vari al variare del livello di X. Ciò implica che la conoscenza del valore assunto da X non aiuta in nessun modo a prevedere il comportamento di Y, ovvero la distribuzione di probabilità. Data dunque la variabile discreta (X, Y) descritta dalla funzione di probabilità congiunta $p(x, y)$, dalla (3.9) si desume che

$$p(x, y) = p(y|x)p_X(x);$$

tale relazione, se Y non dipende da X diviene

$$p(x, y) = p_Y(y)p_X(x).$$

Un analogo ragionamento può essere sviluppato sulla base della funzione di probabilità condizionata $p(x|y)$, e nel caso di variabili aleatorie congiuntamente continue.

Definizione 3.3 (Variabili aleatorie indipendenti)**.** Data una variabile aleatoria doppia (X, Y), le variabili aleatorie componenti X e Y, si dicono *indipendenti* se e solo se

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B) \tag{3.12}$$

per ogni coppia di sottoinsiemi della retta reale A e B.

In particolare, definendo $A = \{u : u \leq x\}$ e $B = \{v : v \leq y\}$, si ottiene

$$P(X \in A, Y \in B) = F(x, y) = F_X(x)F_Y(y), \quad \forall (x, y),$$

dove $F_X(x), F_Y(y)$ sono le *funzioni di ripartizione marginali* di X e Y rispettivamente.

Per (X, Y) variabile aleatoria doppia discreta, la (3.12) è equivalente a

$$p(x, y) = p_X(x)p_Y(y), \quad \forall (x, y), \tag{3.13}$$

e nel caso di variabili congiuntamente continue la condizione di indipendenza è equivalente a

$$f(x, y) = f_X(x)f_Y(y), \quad \forall (x, y). \tag{3.14}$$

Infatti se (3.12) è soddisfatta, allora otteniamo la (3.13) ponendo A e B uguali, rispettivamente, ai sottoinsiemi formati da un solo punto $A = \{x\}$ e $B = \{y\}$. Inoltre, se (3.13) è soddisfatta, allora per ogni coppia di sottoinsiemi A e B,

$$\begin{aligned} P(X \in A, Y \in B) &= \sum_{y \in B} \sum_{x \in A} p(x, y) \\ &= \sum_{y \in B} \sum_{x \in A} p_X(x) p_Y(y) \\ &= \sum_{y \in B} p_Y(y) \sum_{x \in A} p_X(x) \\ &= P(Y \in B) P(X \in A). \end{aligned}$$

Per definire i concetti di *covarianza* e *correlazione* occorre anzitutto introdurre la nozione di valore atteso di una funzione delle variabili aleatorie X e Y, componenti della variabile doppia (X, Y).

Definizione 3.4 (Valore atteso di una funzione di una v.a. doppia)**.** Sia $g(X, Y)$ una funzione che associa ad ogni coppia di numeri (x, y) un solo numero reale. Il valore atteso di $g(X, Y)$ è dato da

$$\begin{aligned} E(g(X,Y)) &= \sum_x \sum_y g(x, y) p(x, y), && \text{se (X,Y) è discreta} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y)\, dx\, dy && \text{se (X,Y) è continua} \end{aligned}$$

La (3.4) include le funzioni $g(X, Y) = X$, $g(X, Y) = X^2$, $g(X, Y) = Y$ e si osserva che tutti i momenti per X e Y possono essere ricavati dalla funzione di probabilità del vettore (X, Y). Inoltre, siamo anche in grado di valutare il momento $E(X^j Y^k)$, dove j, k sono interi non-negativi. Questi sono detti **momenti congiunti** della distribuzione di probabilità di (X, Y) e, se esistono, possono essere usati per descrivere diversi aspetti della distribuzione congiunta. Quindi se $j = 1$ e $k = 0$, il momento congiunto $(1, 0)$ è

$$E(X^1 Y^0) = E(X),$$

per $j = 0$, $k = 1$, si ha il momento congiunto $(0, 1)$

$$E(X^0 Y^1) = E(Y),$$

e i momenti congiunti $(2,0)$, $(0,2)$, $(1,1)$ corrispondono rispettivamente a

$$\begin{aligned} E(X^2Y^0) &= E(X^2) \\ E(X^0Y^2) &= E(Y^2) \\ E(X^1Y^1) &= E(XY). \end{aligned}$$

È utile definire la *funzione generatrice dei momenti congiunti* che viene di seguito introdotta nel caso bivariato.

Definizione 3.5 (Funzione generatrice dei momenti congiunti)**.** Siano X e Y due variabili aleatorie definite sullo stesso spazio campionario. La *funzione generatrice dei momenti congiunti* di X e Y, $m_{X,Y}$, è definita come segue

$$m_{X,Y}(s,t) = E(e^{sX+tY}), \tag{3.15}$$

per tutti i valori $s, t \in \mathbb{R}$ per i quali la variabile e^{sX+tY} ha valore atteso finito.

Si osservi che essendo $e^{0\cdot X+0\cdot Y} = 1$, la funzione generatrice dei momenti congiunti è definita in $s = t = 0$ per ogni coppia (X, Y). Inoltre $m_{X,Y}(0,0) = E(1) = 1$.

Ora, se esistono numeri positivi s_0, t_0 tali che $m_{X,Y}(s,t)$ è definita per tutti gli s e t con $-s_0 < s < s_0$ e $-t_0 < t < t_0$, allora X e Y hanno momenti di tutti gli ordini ed esistono le relative derivate di $m_{X,Y}$. Prendendo le opportune derivate parziali della funzione generatrice dei momenti congiunti, si ottengono i momenti singoli e congiunti di X, Y, ovvero

$$E(X^jY^k) = \frac{\partial^{j+k} m_{X,Y}(s,t)}{\partial s^j \partial t^k}\Big|_{(s,t)=(0,0)} \qquad j,k = 0,1,2\ldots \tag{3.16}$$

In altri termini, i momenti per (X, Y) possono essere generati dalle derivate parziali di $m_{X,Y}(s,t)$ con $s = t = 0$.

Il seguente risultato ha notevole rilevanza ed è valido per variabili aleatorie indipendenti. Si assuma X e Y variabili aleatorie indipendenti e continue con funzione di densità congiunta $f(x,y)$ e sia $g(X,Y) = g_1(X)g_2(Y)$. Allora possiamo scrivere

$$E(g(X,Y)) = E(g_1(X)g_2(Y)) = E(g_1(X))E(g_2(Y)) \tag{3.17}$$

Infatti, essendo X e Y variabili indipendenti, si ha

$$\begin{aligned} E(g_1(X)g_2(Y)) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f(x,y)\,dx\,dy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_1(x)g_2(y)f_X(x)f_Y(y)\,dx\,dy \\ &= \int_{-\infty}^{\infty} g_1(x)f_X(x)\,dx \int_{-\infty}^{\infty} g_2(y)f_Y(y)\,dy \\ &= E(g_1(X))E(g_2(Y)) \end{aligned}$$

e dunque il valore atteso di una qualsiasi funzione moltiplicativa di variabili aleatorie indipendenti è uguale al prodotto dei singoli valori attesi (lo stesso risultato è valido nel caso discreto).

Dalla definizione (3.4) posto $g(X,Y) = (X-\mu_X)(Y-\mu_Y)$, dove μ_X e μ_Y sono le medie delle distribuzioni marginali di X e Y, si perviene alla nozione di *covarianza*.

Definizione 3.6 (Covarianza tra due variabili aleatorie)**.** Data la variabile aleatoria doppia (X,Y) si chiama *covarianza* l'espressione

$$\begin{aligned} \mathrm{Cov}(X,Y) &= E[(X-\mu_X)(Y-\mu_Y)] \\ &= \sum_x\sum_y (x-\mu_X)(y-\mu_Y)p(x,y), \qquad \text{se (X,Y) è discreta} \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_X)(y-\mu_Y)f(x,y)dxdy, \ \text{se (X,Y) è continua} \end{aligned}$$

dove $\mu_X = E(X)$ e $\mu_Y = E(Y)$, se tali valori attesi esistono.

Sviluppando il membro a destra della precedente definizione vediamo che

$$\begin{aligned} \mathrm{Cov}(X,Y) &= E[(X-\mu_X)(Y-\mu_Y)] \\ &= E[XY - \mu_X Y - \mu_Y X + \mu_X\mu_Y] \\ &= E(XY) - \mu_X E(Y) - \mu_Y E(X) + \mu_X\mu_Y \\ &= E(XY) - \mu_X\mu_Y - \mu_Y\mu_X + \mu_X\mu_Y \\ &= E(XY) - \mu_X\mu_Y. \end{aligned}$$

e quindi $\text{Cov}(X,Y)$ è negativa, nulla o positiva a seconda che $E(XY)$ sia minore, uguale o maggiore di $E(X)E(Y)$. La covarianza gode inoltre delle seguenti proprietà:

a. $\text{Cov}(X,Y) = \text{Cov}(Y,X)$;

b. $\text{Cov}(X,X) = V(X)$;

c. $\text{Cov}(aX,bY) = ab\text{Cov}(X,Y)$, $a,b \in \mathbb{R}$;

d. $\text{Cov}(a+X,b+Y) = \text{Cov}(X,Y)$, $a,b \in \mathbb{R}$;

Se le variabili aleatorie X e Y presentano covarianza nulla si dicono **incorrelate**. È facile verificare che se X e Y sono indipendenti allora sono incorrelate, cioè $\text{Cov}(X,Y) = 0$. Infatti, per quanto visto prima, si deduce facilmente che

$$E(XY) = \sum_x \sum_y xy p_X(x) p_Y(y) = \sum_x x p_X(x) \sum_y y p_Y(y) = E(X)E(Y)$$

se (X,Y) è discreta, e

$$\begin{aligned} E(XY) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_X(x) f_Y(y)\, dx\, dy \\ &= \int_{-\infty}^{\infty} x f_X(x) dx \int_{-\infty}^{\infty} y f_Y(y) dy = E(X)E(Y) \end{aligned}$$

se (X,Y) è continua. Pertanto risulta $\text{Cov}(X,Y) = 0$. Il viceversa è invece falso, ovvero la incorrelazione non implica l'indipendenza.

Si è mostrato che l'indipendenza di variabili casuali si estrinseca in una qualche forma di fattorizzazione. Ad esempio, si è visto che due variabili casuali sono indipendenti se e solo se la loro funzione di probabilità (densità) congiunta è uguale al prodotto delle funzioni di probabilità (densità) marginali. Tale proprietà sussiste anche per la funzione generatrice dei momenti congiunti. Rimanendo nel caso bivariato, dalla (3.17) si ottiene, per X e Y indipendenti,

$$\begin{aligned} m_{X,Y}(s,t) = E(e^{sX+tY}) =& E(e^{sX} e^{tY}) \\ =& E(e^{sX}) E(e^{tY}) \\ =& m_X(s) m_Y(t) \end{aligned}$$

cioè la funzione generatrice dei momenti congiunti coincide con il prodotto delle funzioni generatrici dei momenti marginali.

Siano ora $Y_1 = g_1(X_1)$ e $Y_2 = g_2(X_2)$, due funzioni delle variabili aleatorie indipendenti X_1 e X_2. La funzione generatrice dei momenti congiunti di Y_1 e Y_2 è

$$\begin{aligned} m_{Y_1,Y_2}(t_1,t_2) = E(e^{t_1Y_1+t_2Y_2}) \\ =&E(e^{t_1g_1(X_1)+t_2g_2(X_2)}) \\ =&E(e^{t_1g_1(X_1)})E(e^{t_2g_2(X_2)}) \\ =&m_{Y_1}(t_1)m_{Y_2}(t_2), \end{aligned}$$

utilizzando il fatto che X_1 e X_2 sono indipendenti. Questo stabilisce che funzioni di variabili aleatorie indipendenti sono tra loro indipendenti.

Teorema 3.1. Siano X_1 e X_2 variabili aleatorie indipendenti e $Y_1 = g_1(X_1)$ e $Y_2 = g_2(X_2)$. Allora Y_1 e Y_2 sono indipendenti.

Il risultato appena visto è facilmente generalizzabile a n variabili aleatorie indipendenti $X_1, X_2, \ldots, X_n$: posto $Y_1 = g_1(X_1), Y_2 = g_2(X_2), \ldots, Y_n = g_n(X_n)$, allora $Y_1, Y_2, \ldots, Y_n$ sono indipendenti.

Un'altra costante caratteristica di interesse è il *coefficiente di correlazione di Bravais-Pearson*, indice che esprime un'eventuale relazione di linearità tra X e Y.

Definizione 3.7 (Coefficiente di correlazione di Bravais-Pearson)**.** Data la variabile aleatoria doppia (X, Y) il coefficiente di correlazione di Bravais-Pearson è definito, se $V(X), V(Y)$ non sono nulle, da

$$\rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{V(X)V(Y)}} \tag{3.18}$$

e assume valori nell'intervallo $[-1, 1]$.

I valori estremi sono assunti se e solo se una variabile aleatoria è funzione lineare dell'altra con probabilità 1. Si dimostra infatti che

$$\rho(X,Y) = \pm 1 \iff P(V = cU) = 1 \quad \text{per qualche costante c,}$$

essendo $U = \dfrac{X - \mu_X}{\sqrt{V(X)}}$ e $V = \dfrac{Y - \mu_Y}{\sqrt{V(Y)}}$.

Un valore di $\rho(X, Y)$ vicino a +1 o -1 indica un alto livello di linearità tra X e Y. In particolare, un valore positivo di $\rho(X, Y)$ indica che Y tende a crescere quando X cresce, mentre un valore negativo indica che Y decresce al crescere di X. Infine si osservi che $\rho(X, Y) = 0$ se e solo se X e Y sono incorrelate.

3.1.1 Disuguaglianza di Cauchy-Schwarz

Teorema 3.2 (Disuguaglianza di Cauchy-Schwarz)**.** Se X e Y hanno momenti secondi finiti, allora

$$[E(XY)]^2 = |E(XY)|^2 \leq E(X^2)E(Y^2), \tag{3.19}$$

con l'uguaglianza che vale se e solo se $P(Y = cX) = 1$, per una qualche costante c.

Dimostrazione. Dall'esistenza dei valori attesi $E(X^2)$ ed $E(Y^2)$, segue l'esistenza di $E(X), E(Y)$ e $E(XY)$. Definiamo $0 \leq h(t) = E((tX - Y)^2) = E(t^2X^2 - 2tXY + Y^2) = E(X^2)t^2 - 2E(XY)t + E(Y^2)$. Se $h(t) > 0$, le radici di $h(t)$ non sono reali, ovvero $4[E(XY)]^2 - 4E(X^2)E(Y^2) < 0$, cioè $[E(XY)]^2 < E(X^2)E(Y^2)$. Se $h(t) = 0$ per un qualche t, supponiamo $t = t_0$, allora $E((t_0X - Y)^2) = 0$, che implica $P(t_0X = Y) = 1$. □

Utilizzando la disuguaglianza di Cauchy-Schwarz si dimostra che $|\rho(X, Y)| \leq 1$, con l'uguaglianza che vale se e solo se una variabile è funzione lineare dell'altra con probabilità 1. Si può riscrivere la (3.19) come $|E(UV)| \leq \sqrt{E(U^2)E(V^2)}$. Ponendo

$$U = X - \mu_X, \quad U = Y - \mu_Y$$

si ha infatti $|\mathrm{Cov}(X, Y)| = |E(UV)| \leq \sqrt{E(U^2)E(V^2)} = \sqrt{V(X)V(Y)}$.

3.2 Variabili aleatorie multiple

Possiamo definire in maniera analoga al caso di due variabili, la distribuzione di probabilità congiunta di n variabili aleatorie. In generale, la variabile aleatoria multipla $(X_1, X_2, \ldots, X_n)$ è generata dall'associazione ad ogni evento elementare ω dello spazio campionario Ω di una n-upla di numeri reali $(x_1, x_2, \ldots, x_n)$, dove $x_1 = X_1(\omega)$, $x_2 = X_2(\omega), \ldots,$ $x_n = X_n(\omega)$. La *funzione di ripartizione della variabile aleatoria multipla* $(X_1, X_2, \ldots, X_n)$ si definisce come

$$F(x_1, x_2, \ldots, x_n) = P(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n) \tag{3.20}$$

Occorre, anche in questo caso distinguere il caso discreto dal caso continuo. Se $(X_1, X_2, \ldots, X_n)$ è una variabile aleatoria multipla discreta, allora la funzione di probabilità congiunta è definita come

$$p(x_1, x_2, \ldots, x_n) = P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)$$

e deve risultare

$$p(x_1, x_2, \ldots, x_n) \geq 0; \quad \sum_{x_1} \cdots \sum_{x_n} p(x_1, x_2, \ldots, x_n) = 1,$$

dove la singola sommatoria è estesa a tutti i valori della variabile X_i $(i = 1, \ldots, n)$ a cui si riferisce. Se $(X_1, X_2, \ldots, X_n)$ è una variabile aleatoria multipla continua, si ha la funzione di densità congiunta $f(x_1, x_2, \ldots, x_n)$ che deve soddisfare

$$f(x_1, x_2, \ldots, x_n) \geq 0; \quad \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, x_2, \ldots, x_n)\, dx_1 dx_2 \ldots dx_n = 1.$$

Se A è un sottoinsieme dello spazio campionario, la probabilità che la variabile aleatoria $(X_1, \ldots, X_n)$ appartenga ad A è espressa da

$$P((X_1, \ldots, X_n) \in A) = \sum \cdots \sum_{(x_1, \ldots, x_n) \in A} p(x_1, \ldots, x_n) \tag{3.21}$$

se $(X_1, \ldots, X_n)$ è discreta e

$$P((X_1, \ldots, X_n) \in A) = \int \cdots \int_{(x_1, \ldots, x_n) \in A} f(x_1, \ldots, x_n)\, dx_1 \ldots dx_n \tag{3.22}$$

se $(X_1, \ldots, X_n)$ è continua. Data la funzione di probabilità o densità congiunta è possibile scrivere la funzione di probabilità o densità *marginale* della singola variabile X_j, sommando (o integrando) rispetto a tutte le variabili eccetto che la j-ma:

$$\begin{aligned} f_{X_j}(x_j) &= \sum_{\forall i \neq j} \cdots \sum p(x_1, \ldots, x_n), \\ &= \int_{\forall i \neq j} \cdots \int f(x_1, \ldots, x_n)\, dx_1 \ldots dx_n, \end{aligned} \tag{3.23}$$

rispettivamente nel discreto e nel continuo. Infine, il concetto di indipendenza può essere così esteso al caso $n > 2$.

Definizione 3.8 (Indipendenza)**.** Le variabili aleatorie X_1, $X_2, \ldots,$ X_n si dicono *indipendenti* se e solo se vale la relazione

$$f(x_1, x_2, \ldots, x_n) = f_{X_1}(x_1) f_{X_2}(x_2) \cdots f_{X_n}(x_n), \quad \forall (x_1, x_2, \ldots, x_n),$$

dove $f_{X_j}(x_j)$ è la funzione di probabilità (densità) marginale di X_j.

3.2.1 Distribuzione multinomiale

La distribuzione multinomiale si ottiene quando si ripete n volte un esperimento aleatorio, i cui esiti sono più di due. Può essere perciò considerata come un'estensione della binomiale.
Supponiamo che ogni esperimento possa avere come risultato uno qualsiasi tra k possibili esiti, con probabilità $p_1, p_2, \ldots, p_k$, $\sum_{i=1}^{k} p_i = 1$. Se denotiamo con X_i il numero degli esperimenti tra gli n che hanno dato come risultato l'esito numero i, allora

$$P(X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k) = \frac{n!}{x_1! x_2! \cdots x_k!} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k}, \tag{3.24}$$

dove $x_i = 0, \ldots, n$ e $n = x_1 + x_2 + \ldots x_k$. La (3.24) definisce la *distribuzione multinomiale*, ed è verificata notando che ogni successione degli esiti degli n esperimenti che portano al fatto che l'esito i si verifichi esattamente x_i volte, per $i = 1, \ldots, k$, avrà probabilità pari a $p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k}$ grazie all'ipotesi di indipendenza. Essendoci $n!/(x_1! x_2! \cdots x_k!)$ di tali sequenze, si perviene alla (3.24). Si noti che la distribuzione marginale della singola variabile X_i è la binomiale di parametri n e p_i.

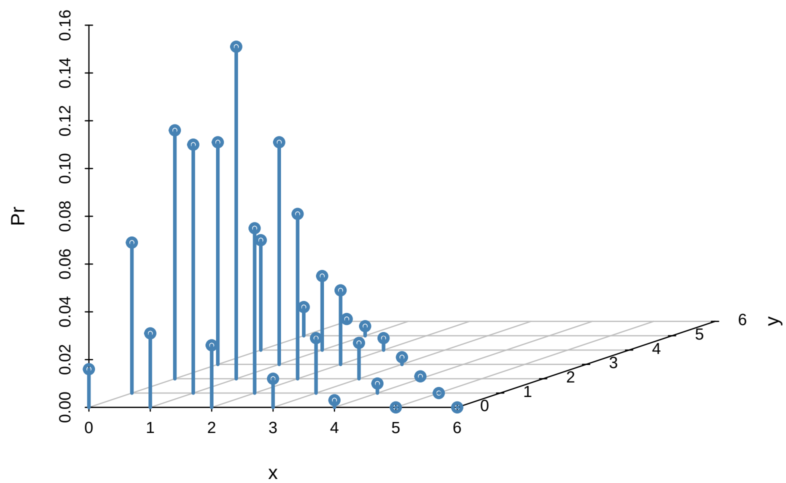


Figura 3.1: Funzione di probabilità della distribuzione multinomiale con $k = 3$, di parametri $n = 6$ e $p_1 = 1/6$, $p_2 = 1/3$.

Esempio 3.1.

Supponiamo di lanciare 9 volte un dado equilibrato. La probabilità che 1 appaia tre volte, 2 e 3 due volte ciascuno, 4 e 5 una sola volta ciascuno e il 6 mai è data da

$$p(3,2,2,1,1,0) = \frac{9!}{3!2!2!1!1!0!}\left(\frac{1}{6}\right)^3\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^1\left(\frac{1}{6}\right)^1\left(\frac{1}{6}\right)^0 = 0.0015$$

3.2.2 Distribuzione normale multivariata

La distribuzione normale si presta alla generalizzazione al caso multidimensionale, di cui si considera dapprima il caso particolare di due dimensioni.

Definizione 3.9 (Distribuzione normale bivariata)**.** Le variabili aleatorie X e Y sono dette avere una *distribuzione normale bivariata* (o doppia) se la loro densità congiunta è data da

$$f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}\, e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2+\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2-2\rho\frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}\right]} \tag{3.25}$$

con parametri $\mu_X = E(X)$, $\mu_Y = E(Y)$, $\sigma_X, \sigma_Y > 0$ e $\rho \in [-1,1]$ coefficiente di correlazione lineare.

In termini matriciali, occorre definire il vettore dei parametri media $\boldsymbol{\mu} = (\mu_X, \mu_Y)'$ e la matrice 2×2, $\boldsymbol{\Sigma}$, di varianza e covarianza simmetrica

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix}$$

che contiene sulla diagonale principale la varianza di X e Y, rispettivamente, mentre al di fuori della diagonale principale vi è la covarianza tra X e Y, $\mathrm{Cov}(X,Y) = \rho\sigma_X\sigma_Y$. Allora la variabile aleatoria doppia (X,Y) si distribuisce secondo una normale multivariata in 2 dimensioni di parametri $\boldsymbol{\mu}$ e $\boldsymbol{\Sigma}$

$$(X,Y)' \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

La (3.25) naturalmente è positiva. Si può inoltre dimostrare che è in effetti una funzione di densità:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = 1.$$

Infatti sostituendo

$$u = \frac{x-\mu_X}{\sigma_X}, \quad v = \frac{y-\mu_Y}{\sigma_Y},$$

l'integrale doppio diventa

$$\begin{aligned}
&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{2\pi\sqrt{1-\rho^2}}\, e^{-\frac{1}{2(1-\rho^2)}[u^2+v^2-2\rho uv]}dudv \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{2\pi\sqrt{1-\rho^2}}\, e^{-\frac{1}{2(1-\rho^2)}[(u-\rho v)^2+(1-\rho^2)v^2]}dudv
\end{aligned}$$

e se sostituiamo

$$w = \frac{u - \rho v}{\sqrt{1-\rho^2}}, \quad dw = \frac{du}{\sqrt{1-\rho^2}},$$

l'integrale può essere scritto come il prodotto di due integrali semplici

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-w^2/2} dw \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-v^2/2} dv$$

entrambi uguali ad 1, come si è visto studiando la distribuzione normale unidimensionale.
Integrando la (3.25) rispetto a y (x) si verifica che se (X, Y) ha distribuzione normale bidimensionale, allora X e Y sono entrambe variabili aleatorie normali di parametri, rispettivamente, μ_X, σ_X^2, μ_Y, σ_Y^2 e le distribuzioni marginali di X e Y sono date da

$$f_X(x) = \frac{1}{\sigma_X\sqrt{2\pi}} e^{-(x-\mu_X)^2/2\sigma_X^2}, \quad f_Y(y) = \frac{1}{\sigma_Y\sqrt{2\pi}} e^{-(y-\mu_Y)^2/2\sigma_Y^2}.$$

Per variabili casuali normali, la covarianza nulla è una condizione necessaria e sufficiente per l'indipendenza (questa condizione non è valida in generale):

$$\text{Cov}(X, Y) = 0 \iff X, Y \text{ indipendenti}$$

Infatti essendo $\text{Cov}(X, Y) = \rho\sigma_X\sigma_Y$, se $\rho = 0$ possiamo scrivere la (3.25) come

$$f(x, y) = f_X(x) f_Y(y);$$

d'altra parte, se X e Y sono indipendenti, allora sono incorrelate e quindi $\text{Cov}(X, Y) = 0$.

Un piano parallelo al piano (x, y) interseca orizzontalmente la superficie campanulare di una normale bivariata, formando un ellissoide tale per cui, in ogni punto (x, y) appartenente all'ellisse, la funzione di densità è costante $f(x, y) = k$ (si veda Fig. 3.2).

Un piano perpendicolare al piano (x, y) interseca verticalmente la superficie e, per un fissato valore di x o y definisce una funzione di densità univariata condizionata $f(y|X = x)$ o $f(x|Y = y)$. Ad esempio, la densità condizionata di Y dato $X = x$ è data da

$$\begin{aligned} f(y|x) &= \frac{f(x, y)}{f_X(x)} \\ &= \frac{1}{\sigma_Y\sqrt{2\pi(1-\rho^2)}} e^{\left[-\left(y-\mu_Y-\rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X)\right)^2 \Big/ 2\sigma_Y^2(1-\rho^2)\right]} \end{aligned}$$

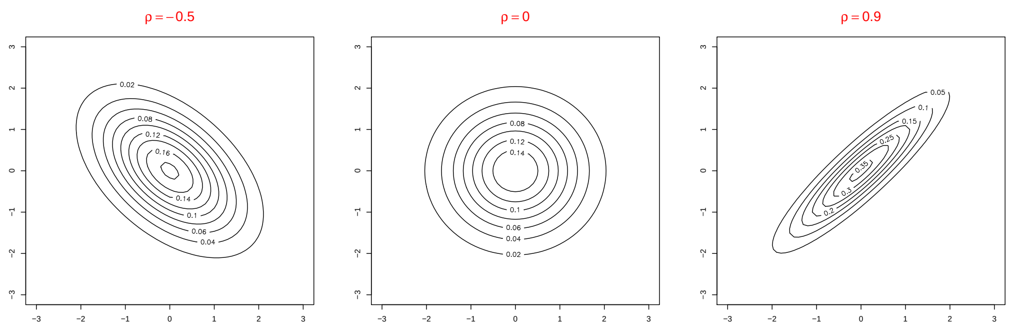


Figura 3.2: Curve di livello della normale bivariata di parametri $\mu_X = \mu_Y = 0$, $\sigma_X = \sigma_Y = 1$ e alcuni valori di ρ.

che è la densità di una legge normale di media $\mu_Y + \rho \frac{\sigma_Y}{\sigma_X}(x - \mu_X)$ e varianza $\sigma_Y^2(1 - \rho^2)$.

Si consideri ora la variabile aleatoria multipla $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)'$. Il vettore $\boldsymbol{X}$ si distribuisce come una normale multivariata a n dimensioni se la sua funzione di densità di probabilità è

$$f(\boldsymbol{x}) = \frac{1}{(2\pi)^{n/2}(\det\boldsymbol{\Sigma})^{1/2}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}, \tag{3.26}$$

per $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, dove $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_n)'$ è il vettore dei parametri media e $\boldsymbol{\Sigma}$ è la matrice simmetrica, definita positiva di dimensione $(n \times n)$:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \ldots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \ldots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \ldots & \sigma_n^2 \end{bmatrix}$$

con $\sigma_{ij} = \sigma_{ji}$, $i, j = 1, \ldots, p$, $\sigma_{ij} = \mathrm{Cov}(X_i, X_j)$, $\sigma_i^2 = V(X_i)$.

Data la funzione di densità (3.26), per ogni punto $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ possiamo calcolare attraverso un integrale multiplo la probabilità congiunta che ogni variabile X_i assuma un valore inferiore a x_i, contemporaneamente per ogni $i = 1 \ldots, p$,

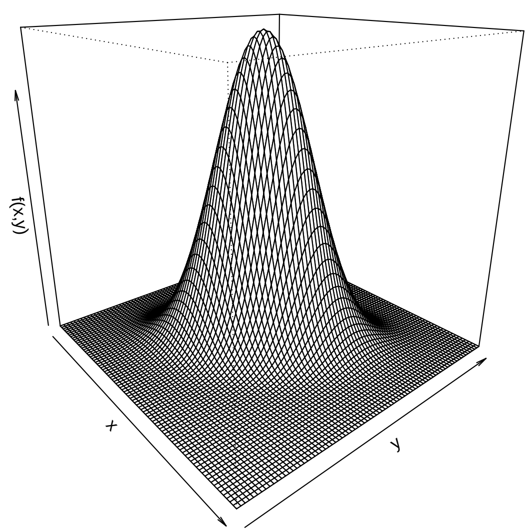


Figura 3.3: Densità normale bivariata di parametri $\mu_X = \mu_Y = 0$, $\sigma_X = \sigma_Y = 1$ e $\rho = 1/2$.

$$P(\boldsymbol{X} < \boldsymbol{x}) = P(X_1 < x_1, \dots X_n < x_n) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_n} f(\boldsymbol{x}) d\boldsymbol{x}$$

Se il vettore $\boldsymbol{X}$ si distribuisce come una normale multivariata a n dimensioni, $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, ogni singola variabile X_i si distribuisce come una Normale univariata, ovvero $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$. Se la matrice $\boldsymbol{\Sigma}$ è diagonale allora tutte le covarianze sono nulle $\sigma_{ij} = 0$ che, nel caso normale, è una condizione necessaria e sufficiente per l'indipendenza; vi è, cioè, indipendenza fra tutte le possibili coppie di variabili X_i e X_j. Questo implica, come visto nel caso bidimensionale, che la funzione di densità per il vettore casuale $\boldsymbol{X}$ è il prodotto delle funzioni di densità di ogni singola variabile X_i, $i = 1, \dots, n$. Infine si osservi che se $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, allora

$$\boldsymbol{Y} = \boldsymbol{\Sigma}^{-1/2}(\boldsymbol{X} - \boldsymbol{\mu}) \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$$

è una normale multipla a componenti indipendenti e standardizzate.

Capitolo 4

Trasformazione di variabili aleatorie e convergenza di successioni

Si presenta di frequente il problema di determinare le proprietà statistiche di una variabile casuale $Y(\omega)$ che è il risultato della trasformazione applicata ad un'altra variabile casuale $X(\omega)$ con distribuzione nota. Ad esempio, $Y(\omega)$ può essere espressa matematicamente dalla funzione di variabile reale $g(x)$, definita in R_X, che fa corrispondere all'evento ω il numero reale $y \in R_Y$, generica realizzazione della funzione di variabile aleatoria:

$$Y(\omega) = g(X(\omega)). \tag{4.1}$$

Generalizzando la formulazione di queste trasformazioni al caso multidimensionale, si può definire una variabile aleatoria m-dimensionale $\boldsymbol{Y}(\omega) = \boldsymbol{g}(\boldsymbol{Z}(\omega))$, che con legge definita dalla funzione vettoriale $\boldsymbol{g}$ si può esprimere come funzione di un vettore aleatorio n-dimensionale $\boldsymbol{Z}(\omega)$ avente densità di probabilità nota.

Nel paragrafo che segue si studierà la trasformazione (4.1), e si mostrerà come si determina la legge probabilistica della funzione di una variabile casuale $Y = g(X)$, utilizzando il *metodo della funzione di ripartizione*, il *metodo della trasformazione* della densità di probabilità $f_X(x)$ e le proprietà della *funzione generatrice dei momenti.*

4.1 Distribuzione di funzioni di una variabile aleatoria

Un metodo generale per ottenere la distribuzione della funzione della variabile aleatoria X, $Y = g(X)$, consiste nel determinare la sua funzione di ripartizione $F_Y(y)$ mediante mediante una opportuna integrazione della densità di X, che per ipotesi è nota e definita sul supporto R_X. Infatti la funzione di ripartizione di Y è data da

$$F_Y(y) = P(Y \leq y), \quad \forall y \in \mathbb{R},$$

ed essendo $Y = g(X)$, l'evento $\{Y \leq y\}$ è equivalente a $\{g(X) \leq y\}$, da cui

$$F_Y(y) = P(g(X) \leq y).$$

In molti casi quest'ultima probabilità può essere facilmente calcolata essendo espressa in termini della funzione data g e della variabile casuale data X.

Esempio 4.1.

Sia X uniformemente distribuita in $[3, 5]$, $f_X(x) = \frac{1}{2}$, $3 \leq x \leq 5$, e si consideri la trasformazione $Y = \dfrac{X-2}{3}$. Allora si ha

$$F_Y(y) = P(Y \leq y) = P\left(\frac{X-2}{3} \leq y\right) = P(X \leq 3y + 2) = \frac{3y-1}{2}.$$

La densità di probabilità $f_Y(y)$ si ottiene quindi, come ci è noto, per derivazione

$$f_Y(y) = dF(y)/dy = 3/2$$

Inoltre, da $3 \leq x \leq 5$ si ricava facilmente $\frac{1}{3} \leq y \leq 1$, e pertanto Y è ancora una variabile uniformemente distribuita con funzione di densità

$$f_Y(y) = \frac{3}{2}, \quad \frac{1}{3} \leq y \leq 1.$$

La trasformazione lineare nell'esempio è un caso particolare di una funzione monotona (crescente). Se la funzione $y = g(x)$ è continua e monotona crescente (decrescente) allora esiste la funzione inversa $g^{-1}()$. Ovvero, se $y = t$

il valore di x che dá questo valore di y è $g^{-1}(t)$. Si osservi che se $g(x)$ è crescente, l'insieme dei valori y tali che $y \leq t$ corrisponde all'insieme dei valori di x tali che $x \leq g^{-1}(t)$, mentre se $g(x)$ è decrescente, l'insieme dei valori $y \leq t$ corrisponde a $x \geq g^{-1}(t)$. Inoltre se $y = g(x)$ è monotona crescente (decrescente) allora tale è $x = g^{-1}(y)$ e pertanto la sua derivata $dg^{-1}(y)/dy$ è positiva o negativa a seconda che $g(x)$ sia crescente o decrescente.

Teorema 4.1. Sia X una variabile aleatoria continua con funzione di densità f_X e sia $y = g(x)$ una funzione derivabile per ogni $x \in R_X$ e tale che $dg(x)/dx = g'(x) > 0$ (oppure $g'(x) < 0$) per ogni $x \in R_X$. Allora la funzione di ripartizione della variabile aleatoria $Y = g(X)$ è

$$\begin{aligned} F_Y(y) &= F_X(g^{-1}(y)), \; \alpha < y < \beta \qquad \text{se } g'(x) > 0 \;\; \forall x \in R_X \\ &= 1 - F_X(g^{-1}(y)), \; \alpha < y < \beta \quad \text{se } g'(x) < 0 \;\; \forall x \in R_X \end{aligned} \tag{4.2}$$

e la densità di Y è

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{dg^{-1}(y)}{dy} \right|, \; \alpha < y < \beta \tag{4.3}$$

dove $\alpha = \min\{\lim_{x \to -\infty} g(x), \lim_{x \to \infty} g(x)\}$ e $\beta = \max\{\lim_{x \to -\infty} g(x), \lim_{x \to \infty} g(x)\}$.

Dimostrazione. Se $g'(x) > 0$ per ogni $x \in R_X$, allora la funzione $y = g(x)$ è continua e strettamente crescente; la funzione inversa $x = g^{-1}(y)$ esiste ed è derivabile. Quindi possiamo scrivere

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = P(X \leq g^{-1}(y)) = F_X(g^{-1}(y)),$$

per $\alpha < y < \beta$, ed è nulla altrove. Derivando si ottiene

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} F_X(g^{-1}(y)) = f_X(g^{-1}(y)) \frac{dg^{-1}(y)}{dy}, \; \alpha < y < \beta.$$

Se $g'(x) < 0$ per ogni $x \in R_X$, allora la funzione $y = g(x)$ è continua e strettamente decrescente; ancora la funzione inversa $x = g^{-1}(y)$ esiste ed è derivabile. Quindi possiamo scrivere

$$\begin{aligned} F_Y(y) = P(Y \leq y) = P(g(X) \leq y) &= P(X \geq g^{-1}(y)) \\ &= 1 - P(X < g^{-1}(y)) \\ &= 1 - F_X(g^{-1}(y)), \end{aligned}$$

per $\alpha < y < \beta$, ed è nulla altrove, da cui

$$\begin{aligned} f_Y(y) = \frac{d}{dy} F_Y(y) &= \frac{d}{dy}(1 - F_X(g^{-1}(y))) \\ &= -f_X(g^{-1}(y)) \frac{dg^{-1}(y)}{dy} \\ &= f_X(g^{-1}(y)) \left| \frac{dg^{-1}(y)}{dy} \right|, \; \alpha < y < \beta, \end{aligned}$$

essendo $dg^{-1}(y)/dy < 0$ in questo caso. Pertanto se $Y = g(x)$ e $g'(x) > 0$ (oppure $g'(x) < 0$) per ogni $x \in R_X$ la funzione di densità di Y risulta espressa dalla (4.3). □

Esempio 4.2.

La funzione di densità di una trasformazione lineare della variabile aleatoria X, $Y = aX + b$ con $a > 0$ è data da

$$f_Y(y) = f_X\left(\frac{y-b}{a}\right) \left| \frac{d}{dy}\left(\frac{y-b}{a}\right) \right| = \frac{1}{a} f_X\left(\frac{y-b}{a}\right).$$

Esempio 4.3.

Sia X una variabile aleatoria normale $\mathcal{N}(\mu, \sigma^2)$ e sia $Y = e^X$ la variabile aleatoria definita dalla funzione $g(x) = e^x$, derivabile e tale che $g'(x) > 0$ per ogni x.

Si ha inoltre che $\min\{\lim_{x\to-\infty} e^x, \lim_{x\to\infty} e^x\} = 0$ e $\max\{\lim_{x\to-\infty} e^x, \lim_{x\to\infty} e^x\} = \infty$. La funzione inversa è

$$x = g^{-1}(y) = \ln y, \quad y > 0$$

e quindi $dg^{-1}(y)/dy = 1/y$. Applicando la (4.3) si ha

$$f_Y(y) = \begin{cases} \dfrac{1}{\sqrt{2\pi}\sigma y} e^{-(\ln y - \mu)^2/2\sigma^2}, & 0 < y < \infty \\ 0, & \text{altrove.} \end{cases} \tag{4.4}$$

La distribuzione (4.4) va sotto il nome di *distribuzione lognormale.*

Esempio 4.4.

Sia X una variabile aleatoria con distribuzione uniforme in $(0, 1)$:

$$f_X(x) = \begin{cases} 1, & 0 < x < 1 \\ 0, & \text{altrove.} \end{cases}$$

Sia $Y = -\ln X$. Si noti che essendo il supporto di X dato da $R_X = \{x : 0 < x < 1\}$, Y assumerà valori in $R_Y = \{y : y > 0\}$. La trasformazione inversa è $X = e^{-Y}$, cioè $y = g(x) = -\ln x$, $g^{-1}(y) = e^{-y}$ e $dg^{-1}(y)/dy = -e^{-y}$. Allora la funzione di densità Y è

$$f_Y(y) = f_X(e^{-y})\left|\frac{d}{dy}e^{-y}\right| = e^{-y}, \quad y > 0.$$

D'altra parte, se Y è una variabile aleatoria con densità

$$f_Y(y) = e^{-y}, \quad y > 0$$

e definita $X = e^{-Y}$, si trova che X ha densità

$$f_X(x) = f_Y(-\ln x)\left|\frac{d}{dx}(-\ln x)\right| = e^{\ln x}\frac{1}{x} = 1, \quad 0 < x < 1,$$

ovvero $X \sim U(0,1)$.

Nell'Esempio 4.4 si è trattato un caso particolare della *Trasformazione Integrale di Probabilità* che si rivela utile nel generare numeri casuali Y con distribuzione assegnata F_Y. Infatti, se $X \sim U(0,1)$ e F_Y è continua e invertibile, possiamo usare la funzione inversa di F_Y e definire

$$Y = F_Y^{-1}(X)$$

la cui inversa è $X = F_Y(Y)$. Ora, $g(x) = F_Y^{-1}(x)$ ed essendo $F_X(x) = x$, per ogni $0 < x < 1$, preso $t \in \mathbb{R}$, applicando il Teorema 4.1, si ha

$$P(Y \leq t) = F_X(g^{-1}(t)) = F_X(F_Y(t)) = F_Y(t), \quad 0 < F_Y(t) < 1.$$

e dunque la v.a. Y così costruita ha la legge desiderata.

È interessante osservare che se $U \sim U(0,1)$ allora $V = 1 - U \sim U(0,1)$: infatti, se $v \in (0,1)$, si ha

$$F_V(v) = P(V \leq v) = P(1 - U \leq v) = 1 - P(U \leq 1 - v) = v.$$

Viceversa, se Y è una variabile aleatoria avente funzione di ripartizione continua F_Y, e $X = F_Y(Y)$, allora X ha distribuzione uniforme in $(0,1)$:

$$F_X(x) = F_Y(F_Y^{-1}(x)) = x, \quad 0 < x < 1.$$

Si osservi ora che la condizione del Teorema 4.1 per cui $g(x)$ debba avere derivata prima sempre positiva o sempre negativa può essere rimossa se si considerano le radici $x_1, x_2, \ldots, x_n$ dell'equazione $y = g(x)$ e si scrive la densità di $Y = g(x)$ come

$$f_Y(y) = \sum_{i=1}^{n} f_X(g_i^{-1}(y)) \left| \frac{dg_i^{-1}(y)}{dy} \right|, \tag{4.5}$$

dove $g_i^{-1}(y) = x_i$.

Esempio 4.5.

Sia $X \sim \mathcal{N}(0,1)$ e si definisca la variabile aleatoria $Y = X^2$. Poiché la funzione $y = g(x) = x^2$ è non monotona ed ha radici $\pm\sqrt{y}$, possiamo definire le funzioni inverse $g_1^{-1}(y) = -\sqrt{y}$ e $g_2^{-1}(y) = \sqrt{y}$ e ottenere la densità di Y come segue:

$$f_Y(y) = f_X(g_1^{-1}(y)) \left| \frac{dg_1^{-1}(y)}{dy} \right| + f_X(g_2^{-1}(y)) \left| \frac{dg_2^{-1}(y)}{dy} \right|$$

ottenendo così

$$f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-y/2} \left| -\frac{1}{2\sqrt{y}} \right| + \frac{1}{\sqrt{2\pi}} e^{-y/2} \left| \frac{1}{2\sqrt{y}} \right| = \frac{1}{\sqrt{2\pi}} y^{-\frac{1}{2}} e^{-\frac{y}{2}}, \quad y \geq 0 \tag{4.6}$$

Nella (4.6) si riconosce la funzione della densità di distribuzione Gamma di parametri $n = 1/2$, $\lambda = 1/2$. Si perviene quindi al risultato che se Y è il quadrato di una variabile aleatoria normale standard, allora

$$Y \sim \chi_1^2,$$

ovvero Y ha distribuzione chi-quadrato con 1 grado di libertà.

Il risultato ottenuto nell'Esempio 4.5 può essere verificato anche utilizzando il cosiddetto *metodo della funzione generatrice dei momenti*, sfruttandone la proprietà di unicità discussa precedentemente. Se $Y = X^2$, $X \sim \mathcal{N}(0,1)$, la sua funzione generatrice dei momenti sarà espressa da

$$
\begin{aligned}
m_Y(t) = E(e^{tY}) &= \int_{-\infty}^{\infty} e^{tx^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}(1-2t)} dx \\
&= \frac{1}{\sqrt{2\pi}} \frac{(1-2t)^{-\frac{1}{2}}}{(1-2t)^{-\frac{1}{2}}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}(1-2t)} dx \\
&= (1-2t)^{-\frac{1}{2}} = \left(\frac{\frac{1}{2}}{\frac{1}{2} - t} \right)^{1/2} \qquad \text{per} \quad t < 1/2
\end{aligned}
$$

essendo

$$
\frac{1}{\sqrt{2\pi}(1-2t)^{-\frac{1}{2}}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}(1-2t)} dx = 1.
$$

Si ottiene quindi $m_Y(t) = (1-2t)^{-\frac{1}{2}}$ in cui riconosciamo la funzione generatrice dei momenti di una variabile aleatoria $Ga(1/2, 1/2)$ che, come visto, corrisponde ad una chi-quadrato con 1 grado di libertà, e pertanto deve essere $Y = X^2 \sim \chi_1^2$.

Con un procedimento analogo si perviene al seguente risultato. Sia $X \sim Ga(r, \lambda)$, $a > 0$ e $Y = aX$. Allora

$$
Y \sim Ga\left(r, \frac{\lambda}{a}\right).
$$

Basta infatti osservare che la funzione generatrice dei momenti di Y

$$
m_Y(t) = m_X(at) = \left(\frac{\lambda}{\lambda - at} \right)^r = \left(\frac{\lambda/a}{\lambda/a - t} \right)^r, \quad t < \lambda/a
$$

corrisponde alla funzione generatrice dei momenti della distribuzione Gamma di parametri r e λ/a.

4.2 Valori attesi di funzioni di variabili aleatorie

Nel Capitolo 1 è stata data la definizione di valore atteso di una funzione di una variabile aleatoria (si veda (1.6)). In generale, date le variabili casuali $X_1, X_2, \ldots, X_n$, si ponga $Y = g(X_1, X_2, \ldots, X_n)$; allora

$$
E(Y) = E(g(X_1, X_2, \ldots, X_n)),
$$

dove (per variabili casuali continue congiunte)

$$E(g(X_1,\dots,X_n)) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} g(x_1,\dots,x_n)f(x_1,\dots,x_n)\,dx_1\dots\,dx_n$$

dove $f(x_1,\dots,x_n)$ è la funzione di densità congiunta del vettore aleatorio $(X_1,X_2,\dots,X_n)$. La definizione è analoga nel caso discreto, considerando la funzione di probabilità congiunta e le somme estese a tutti i valori di $X_1,\dots,X_n$. Le propritä del valore atteso viste in precedenza possono essere facilmente estese al caso di funzioni di più variabili aleatorie. In particolare:

1. se c è una costante e $H(\cdot)$ è una funzione, allora

$$E(c\,H(X_1,\dots,X_n)) = cE(H(X_1,\dots,X_n))$$

2. sia $g(X_1,\dots,X_n) = h_1(X_1,\dots,X_n) + h_2(X_1,\dots,X_n)$, allora

$$E(g(X_1,\dots,X_n)) = E(h_1(X_1,\dots,X_n)) + E(h_2(X_1,\dots,X_n))$$

3. se $X_1,X_2,\dots,X_n$ sono variabili aleatorie continue indipendenti aventi funzione di densità congiunta

$$f(x_1,\dots,x_n) = f_{X_1}(x_1)f_{X_2}(x_2)\cdots f_{X_n}(x_n)$$

e $g(\cdot)$ è una funzione di $X_1,\dots,X_n$ tale che

$$g(X_1,X_2,\dots,X_n) = h_1(X_1)h_2(X_2)\cdots h_n(X_n),$$

allora

$$\begin{aligned} E(g(X_1,X_2,\dots,X_n)) &= E(h_1(X_1)h_2(X_2)\cdots h_n(X_n)) \\ &= E(h_1(X_1))\,E(h_2(X_2))\cdots E(h_n(X_n)). \end{aligned} \tag{4.7}$$

L'ultimo risultato esprime una importante proprietà del valore atteso che consente di scrivere il valore atteso del prodotto di funzioni di variabili aleatorie indipendenti come il prodotto dei valori attesi dei singoli fattori, in analogia a quanto visto nel caso bivariato nel Capitolo 3. Vedremo a breve l'utilità di questa proprietà anche nello studio della somma di variabili aleatorie.

4.3 Combinazioni lineari di variabili aleatorie

Si consideri la variabile aleatoria multipla $(X_1, X_2, \ldots, X_n)$ e siano $a_1, a_2, \ldots, a_n \in \mathbb{R}$. La combinazione lineare

$$Y = a_1X_1 + a_2X_2 + \cdots + a_nX_n$$

è una variabile aleatoria di cui è importante determinare la media, la varianza e, sotto certe ipotesi, la distribuzione di probabilità.

Teorema 4.2. Sia $(X_1, X_2, \ldots, X_n)$ una variabile aleatoria multipla e siano $a_i \in \mathbb{R}$, $i = 1, \ldots, n$. Allora la variabile aleatoria $Y = \sum_{i=1}^n a_iX_i$ ha media

$$E(Y) = E\left(\sum_{i=1}^n a_iX_i\right) = \sum_{i=1}^n a_iE(X_i) \tag{4.8}$$

e varianza data da

$$V(Y) = V\left(\sum_{i=1}^n a_iX_i\right) = \sum_{i=1}^n a_i^2V(X_i) + 2\sum\sum_{i<j} a_ia_j\mathrm{Cov}(X_i, X_j) \tag{4.9}$$

Dimostrazione. Forniamo la dimostrazione per il caso $n = 2$; si procederà poi per induzione. Si consideri allora la variabile aleatoria $Y = a_1X_1 + a_2X_2$ e consideriamo il caso di (X_1, X_2) discreta con funzione di probabilità congiunta $f(x_1, x_2)$ (per il caso continuo il procedimento è equivalente). Il valore atteso si ricava come

$$\begin{aligned}
E(Y) &= \sum_{x_1}\sum_{x_2}(a_1x_1 + a_2x_2)f(x_1, x_2) \\
&= a_1\sum_{x_1}\sum_{x_2} x_1f(x_1, x_2) + a_2\sum_{x_1}\sum_{x_2} x_2f(x_1, x_2) \\
&= a_1\sum_{x_1} x_1f_{X_1}(x_1) + a_2\sum_{x_2} x_2f_{X_2}(x_2)
\end{aligned}$$

per cui si ottiene

$$E(Y) = a_1E(X_1) + a_2E(X_2). \tag{4.10}$$

Per la varianza si ha

$$\begin{aligned}
V(Y) &= E[(a_1X_1 + a_2X_2 - a_1E(X_1) - a_2E(X_2))^2] \\
&= E[(a_1(X_1 - E(X_1)) + a_2(X_2 - E(X_2)))^2] \\
&= E[a_1^2(X_1 - E(X_1))^2 + a_2^2(X_2 - E(X_2))^2 \\
&\quad + 2a_1a_2(X_1 - E(X_1))(X_2 - E(X_2))] \\
&= a_1^2 \sum_{x_1} \sum_{x_2} (x_1 - E(X_1))^2 f(x_1, x_2) + a_2^2 \sum_{x_1} \sum_{x_2} (x_2 - E(X_2))^2 f(x_1, x_2) \\
&\quad + 2a_1a_2 \sum_{x_1} \sum_{x_2} (x_1 - E(X_1))(x_2 - E(X_2)) f(x_1, x_2) \\
&= a_1^2 \sum_{x_1} (x_1 - E(X_1))^2 f_{X_1}(x_1) + a_2^2 \sum_{x_2} (x_2 - E(X_2))^2 f_{X_1}(x_2) \\
&\quad + 2a_1a_2 E[(x_1 - E(X_1))(x_2 - E(X_2))] f(x_1, x_2)
\end{aligned}$$

Quest'ultima espressione si può chiaramente scrivere come

$$V(Y) = a_1^2 V(X_1) + a_2^2 V(X_2) + 2a_1a_2 \mathrm{Cov}(X_1, X_2). \tag{4.11}$$

Procedendo per induzione, la (4.8) vale per $n = 1$ e, come dimostrato, per $n = 2$. Ammesso ora che valga per $n = k - 1$, proviamo che essa è valida anche per $n = k$. Poniamo

$$U = a_1X_1 + a_2X_2 + \cdots + a_{k-1}X_{k-1}.$$

Quindi in base alla (4.10)

$$E(U + a_kX_k) = E(U) + a_kE(X_k),$$

e poiché si è ammesso che la (4.8) valga per $n = k - 1$, allora la (4.8) è dimostrata.

Procedendo in maniera analoga, la (4.9) è valida per $n = 1$ e per $n = 2$. Assumiamo che valga per $n = k - 1$. Allora si può mostrare che vale anche per $n = k$. Si ha

$$V(U + a_kX_k) = V(U) + a_k^2 V(X_k) + 2a_k \mathrm{Cov}(U, X_k),$$

da cui sostituendo a $V(U)$ l'espressione data dalla (4.11) per $n = k - 1$ ed essendo

$$\mathrm{Cov}(U, X_k) = \sum_{i=1}^{k-1} a_i \mathrm{Cov}(X_i, X_k),$$

è facile vedere che $V(U + a_k X_k)$ ha la stessa struttura della (4.9), come volevasi dimostrare. □

Se le variabili $X_1, X_2, \ldots, X_n$ sono incorrelate, la (4.9) diviene

$$V(Y) = \sum_{i=1}^{n} a_i^2 V(X_i) \tag{4.12}$$

In definitiva si ha immediatamente il seguente corollario del Teorema 4.2.

Corollario 4.1. Siano $X_1, X_2, \ldots, X_n$ variabili aleatorie incorrelate, $a_1, \ldots, a_n$, n costanti reali e $Y = \sum a_i X_i$. Allora

$$E(Y) = \sum_{i=1}^{n} a_i \mu_i, \qquad V(Y) = \sum_{i=1}^{n} a_i^2 \sigma_i^2$$

dove $\mu_i = E(X_i)$ e $\sigma_i^2 = V(X_i)$.

Dal corollario sopra enunciato segue che se $X_1, X_2, \ldots, X_n$ sono variabili aleatorie incorrelate, ciascuna con valore atteso μ e varianza σ^2, allora la variabile aleatoria $Y = \sum a_i X_i$ ha media μ_y e varianza σ_Y^2 date da

$$\mu_Y = \mu \sum_{i=1}^{n} a_i, \qquad \sigma_Y^2 = \sigma^2 \sum_{i=1}^{n} a_i^2.$$

Inoltre vale il seguente risultato.

Teorema 4.3. Siano $X_1, X_2, \ldots, X_n$ variabili aleatorie incorrelate con medie $\mu_1, \mu_2, \ldots, \mu_n$ e varianze $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$. Siano $a_1, \ldots, a_n$ e $b_1, \ldots, b_n$ due insiemi di costanti. Allora, definite le combinazioni lineari

$$U = \sum_{i=1}^{n} a_i X_i \quad \text{e} \quad W = \sum_{i=1}^{n} b_i X_i$$

si ha

$$\mathrm{Cov}(U, W) = \sum_{i=1}^{n} a_i b_i \sigma_i^2 \tag{4.13}$$

Dimostrazione.

$$\begin{aligned}
\mathrm{Cov}(U,W) &= E[(U-E(U))(W-E(W))] \\
&= E\left[\left(\sum_{i=1}^{n} a_i X_i - \sum_{i=1}^{n} a_i \mu_i\right)\left(\sum_{i=1}^{n} b_i X_i - \sum_{i=1}^{n} b_i \mu_i\right)\right] \\
&= E\left[\left(\sum_{i=1}^{n} a_i (X_i - \mu_i)\right)\left(\sum_{i=1}^{n} b_i (X_i - \mu_i)\right)\right] \\
&= \sum_{i=1}^{n} a_i b_i E[(X_i-\mu_i)^2] + 2\sum\sum_{i<j} a_i b_i E[(X_i-\mu_i)(X_j-\mu_j)] \\
&= \sum_{i=1}^{n} a_i b_i \sigma_i^2
\end{aligned}$$

essendo la covarianza nulla per ogni coppia di X_i, $i = 1, \ldots, n$. □

Esempio 4.6.

La scelta appropriata delle costanti $a_1, \ldots, a_n$ e $b_1, \ldots, b_n$ può rendere la covarianza (e la correlazione) uguale a un qualunque valore desiderato, incluso lo zero quando a_i e b_i sono tali che $\sum_{i=1}^n a_i b_i \sigma_i^2 = 0$. Siano X_1 e X_2 variabili aleatorie incorrelate, aventi stessa media μ e varianza σ^2. Se poniamo

$$\begin{aligned}
U &= X_1 + X_2 \\
V &= X_1 - X_2,
\end{aligned}$$

allora $\mu_U = 2\mu$, $\mu_V = 0$, $\sigma_U^2 = \sigma_V^2 = 2\sigma^2$, e $\mathrm{Cov}(U,V) = 0$, ovvero U e V sono a loro volta incorrelate. Ciò spiega il fatto che la somma di due variabili aleatorie non dice nulla circa la loro differenza, essendo queste incorrelate.

Da quanto visto discende il seguente corollario.

Corollario 4.2. Siano $X_1, X_2, \ldots, X_n$ variabili aleatorie incorrelate con medie $\mu_1, \mu_2, \ldots, \mu_n$ e varianza σ^2. Siano $U = \sum_{i=1}^n a_i X_i$ e $V = \sum_{i=1}^n b_i X_i$, con $a_1, \ldots, a_n, b_1, \ldots, b_n \in \mathbb{R}$. Allora, condizione necessaria e sufficiente affinché U e V siano incorrelate è che $\sum_{i=1}^n a_i b_i = 0$.

Prima di affrontare l'ultima parte di questa sezione che riguarda la funzione generatrice dei momenti di una combinazione lineare di variabili aleatorie indipendenti, si osservi che se due variabili sono correlate (ovvero il coefficiente definito dalla (3.18) non è nullo), allora tra di esse deve esistere un certo grado di dipendenza. Tuttavia, l'implicazione inversa non vale. A titolo di controesempio, consideriamo le variabili $X = U + V$ e $Y = U - V$, dove U e V sono variabili uniformi in $[0, 1]$; X e Y sono dipendenti, perché non vale la (3.14), ma la covarianza tra X e Y è nulla, e quindi è nullo il coefficiente di correlazione lineare. D'altra parte si ricorda che se due variabili sono statisticamente indipendenti allora sono certamente incorrelate.

Si consideri nuovamente una combinazione lineare $Y = \sum_{i=1}^{n} a_i X_i$ di variabili aleatorie indipendenti $X_1, X_2, \ldots, X_n$. Il seguente risultato è molto utile nel determinare la legge di probabilità di funzioni lineari di variabili aleatorie indipendenti. Se $a_i \in \mathbb{R}$ per $i = 1, \ldots, n$, la funzione generatrice dei momenti di Y è espressa in termini delle funzioni generatrici dei momenti di $X_1, X_2, \ldots, X_n$:

$$\begin{aligned} m_Y(t) &= E(e^{tY}) = E(e^{t(a_1X_1 + a_2X_2 + \cdots + a_nX_n)}) \\ &= E(e^{ta_1X_1} e^{ta_2X_2} \cdots e^{ta_nX_n}) \\ &= E(e^{ta_1X_1}) E(e^{ta_2X_2}) \cdots E(e^{ta_nX_n}) \\ &= m_{X_1}(ta_1) m_{X_2}(ta_2) \cdots m_{X_n}(ta_n), \end{aligned}$$

dove, usando la (4.7), il valore atteso del prodotto $e^{ta_1X_1} e^{ta_2X_2} \cdots e^{ta_nX_n}$ è stato espresso come prodotto dei valori attesi dei singoli fattori.

Teorema 4.4. Siano $X_1, X_2, \ldots, X_n$ variabili aleatorie indipendenti e sia $Y = \sum_{i=1}^{n} a_i X_i$, con $a_i \in \mathbb{R}$ per $i = 1, \ldots, n$. La funzione generatrice dei momenti di Y è

$$m_Y(t) = \prod_{i=1}^{n} m_{X_i}(ta_i). \tag{4.14}$$

Tale fattorizzazione ha la seguente implicazione. Come si è visto, se due variabili casuali su uno stesso spazio di probabilità X e Y hanno stessa funzione generatrice dei momenti allora X e Y hanno la stessa legge di probabilità: quindi se usiamo la (4.14) per valutare la funzione generatrice dei momenti di $\sum_{i=1}^{n} a_i X_i$ e in questa riconosciamo la funzione generatrice dei momenti di una certa distribuzione, allora $\sum_{i=1}^{n} a_i X_i$ avrà necessariamente la stessa distribuzione.

Esempio 4.7.

Siano X_1, X_2 e X_3 tre variabili aleatorie indipendenti con distribuzione geometrica di parametro p. Allora $m_{X_i}(t) = pe^t/(1-qe^t)$. Quindi, in virtù del Teorema (4.4) è facile verificare che se $Y = X_1 + X_2 + X_3$ allora

$$m_Y(t) = m_{X_1}(t)m_{X_2}(t)m_{X_3}(t) = \left(\frac{pe^t}{1-qe^t}\right)^3, \quad t < -\ln(q)$$

che corrisponde alla funzione generatrice dei momenti di una distribuzione binomiale negativa $\mathcal{NB}(3, p)$.

Esempio 4.8.

Siano X_1 e X_2 due variabili aleatorie esponenziali indipendenti di parametro λ. È noto che la funzione generatrice dei momenti di ciascuna X_i è $\lambda/(\lambda - t)$, per $t < \lambda$. Sia $U = X_1 + X_2$. La funzione generatrice dei momenti di U è data da

$$m_U(t) = m_{X_1}(t)m_{X_2}(t) = \frac{\lambda}{\lambda - t}\frac{\lambda}{\lambda - t} = \left(\frac{\lambda}{\lambda - t}\right)^2$$

ed è quella di una distribuzione Gamma (Erlang, in realtà) di parametri $(2, \lambda)$. È facile verificare che se X_1, X_2, ..., X_n sono indipendenti e $X_i \sim Esp(\lambda)$, allora $\sum_{i=1}^n X_i \sim Ga(n, \lambda)$.

Ricordando che la distribuzione esponenziale di parametro λ corrisponde alla distribuzione $Ga(1, \lambda)$, l'Esempio 4.8 può essere considerato un caso particolare di un risultato più generale, enunciato e dimostrato di seguito.

Teorema 4.5. Siano $X_1, X_2, \ldots, X_n$ variabili aleatorie indipendenti tali che $X_i \sim Ga(\alpha_i, \lambda)$, per $i = 1, \ldots, n$. Allora la variabile aleatoria

$$Y = X_1 + X_2 + \cdots + X_n$$

ha distribuzione Gamma con parametri $\sum_i \alpha_i$ e λ:

$$Y \sim Ga\left(\sum_i \alpha_i, \lambda\right). \tag{4.15}$$

Dimostrazione. Per la (4.14) la funzione generatrice dei momenti di $Y = X_1 + X_2 + \cdots + X_n$ è data da

$$m_Y(t) = \prod_{i=1}^{n} m_{X_i}(t)$$

ed essendo $X_i \sim Ga(\alpha_i, \lambda)$, si ha $m_{X_i} = \left(\frac{\lambda}{\lambda - t}\right)^{\alpha_i}$, da cui

$$m_Y(t) = \prod_{i=1}^{n} \left(\frac{\lambda}{\lambda - t}\right)^{\alpha_i} = \left(\frac{\lambda}{\lambda - t}\right)^{\sum_i \alpha_i},$$

in cui riconosciamo la funzione generatrice dei momenti della legge Gamma di parametri $\sum_i \alpha_i$, λ. □

Con un procedimento analogo, il Teorema 4.4 permette di verificare (come accennato nel Capitolo 2) che la somma di n variabili aleatorie indipendenti di Bernoulli, ciascuna di parametro p, è binomiale di parametri n e p (si lascia la dimostrazione per esercizio).

La proprietà per cui la somma di variabili aleatorie indipendenti e distribuite secondo una stessa legge di probabilità segue ancora la stessa distribuzione è solitamente indicata con il nome di *proprietà di riproducibilità*. Oltre al risultato relativo alla distribuzione Gamma, si verifica facilmente che:

- se $X_1, X_2, \ldots, X_n$ sono variabili aleatorie binomiali indipendenti $X_i \sim Bin(n_i, p)$, allora
$$\sum_{i=1}^{n} X_i \sim Bin\left(\sum_{i=1}^{n} n_i, p\right); \tag{4.16}$$
- se $X_1, X_2, \ldots, X_n$ sono binomiali negative indipendenti $X_i \sim \mathcal{NB}(r_i, p)$, allora
$$\sum_{i=1}^{n} X_i \sim \mathcal{NB}\left(\sum_{i=1}^{n} r_i, p\right); \tag{4.17}$$
- se $X_1, X_2, \ldots, X_n$ sono variabili aleatorie di Poisson indipendenti $X_i \sim Po(\lambda_i)$, allora
$$\sum_{i=1}^{n} X_i \sim Po\left(\sum_{i=1}^{n} \lambda_i\right), \tag{4.18}$$

così la distribuzione di una somma di variabili aleatorie indipendenti aventi distribuzione di Poisson è ancora una variabile di Poisson con parametro uguale alla somma dei singoli parametri.

Concludiamo questa sezione osservando che il Teorema 4.5 si applica al caso particolare di variabili aleatorie chi-quadrato indipendenti. Vale cioè il seguente corollario.

Corollario 4.3. Siano $X_1, X_2, \dots, X_n$ variabili aleatorie indipendenti tali che $X_i \sim \chi^2_{r_i}$, per $i = 1, \dots, n$. Allora la variabile aleatoria

$$Y = X_1 + X_2 + \cdots + X_n$$

ha distribuzione chi-quadrato con $r_1 + r_2 + \cdots + r_n$ gradi di libertà.

Dimostrazione. La dimostrazione è immediata dalla (4.15) considerando la distribuzione chi-quadrato come caso particolare della distribuzione Gamma per $r_i = n/2$, per n intero positivo, e $\lambda = 1/2$. □

Corollario 4.4. Siano $X_1, X_2, \dots, X_n$ variabili aleatorie indipendenti tali che $X_i \sim \mathcal{N}(0, 1)$, per $i = 1, \dots, n$. Allora la variabile aleatoria

$$Y = X_1^2 + X_2^2 + \cdots + X_n^2$$

ha distribuzione chi-quadrato con n gradi di libertà:

$$Y \sim \chi^2_n \tag{4.19}$$

Dimostrazione. Si è visto che il quadrato di una variabile aleatoria normale standard ha distribuzione chi-quadrato con un grado di libertà, quindi $X_i^2 \sim \chi^2_1$. Per il corollario 4.3 la somma di n variabili chi-quadrato con un grado di libertà ha distribuzione chi-quadrato con n gradi di libertà, da cui segue la (4.19). □

4.3.1 Combinazioni lineari di variabili normali

Una qualunque funzione lineare di variabili aleatorie normali indipendenti è ancora una variabile aleatoria normale. In particolare, se $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ e $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ e X_1 e X_2 sono indipendenti, allora si ha

$$X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2),$$
$$X_1 - X_2 \sim \mathcal{N}(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2).$$

In generale, per variabili normali e indipendenti vale il seguente risultato.

Teorema 4.6. Siano $X_1, X_2, \dots, X_n$ variabili aleatorie indipendenti tali che $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$, $i = 1, \dots, n$. Allora la variabile aleatoria

$$Y = a_1X_1 + a_2X_2 + \cdots + a_nX_n$$

ha distribuzione normale. In particolare

$$Y \sim \mathcal{N}\left(\sum_i a_i\mu_i, \sum_i a_i^2\sigma_i^2\right). \tag{4.20}$$

Dimostrazione. sia $Y = \sum_{i=1}^n a_iX_i$, $a_i \in \mathbb{R}$. Allora per la (4.14)

$$\begin{aligned} m_Y(t) &= \prod_{i=1}^n m_{X_i}(ta_i) \\ &= \prod_{i=1}^n e^{ta_i\mu_i + (\sigma_i^2 t^2 a_i^2)/2} \\ &= e^{t\sum_i a_i\mu_i + \frac{t^2}{2}\sum_i a_i^2\sigma_i^2} \end{aligned} \tag{4.21}$$

L'ultima espressione rappresenta la funzione generatrice dei momenti con media e varianza pari a quelle della v.a. Y. Ne consegue, per la proprietà di unicità della funzione generatrice dei momenti che Y è distribuita normalmente. $\square$

Si noti che nel caso di variabili aleatorie normali non indipendenti (e quindi non incorrelate), una combinazione lineare di tali variabili ha ancora distribuzione normale. Ad esempio, se X_1 e X_2 hanno covarianza non nulla e

$$X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2) \qquad X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$$

allora

$$X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2 + 2\mathrm{Cov}(X_1, X_2)),$$

e

$$X_1 - X_2 \sim \mathcal{N}(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2 - 2\mathrm{Cov}(X_1, X_2)).$$

In generale, quindi, la famiglia delle v.a. normali è chiusa rispetto ad ogni combinazione lineare. Se $X_1, X_2, \dots, X_n$ sono v.a. normali di media

μ_i e varianza σ_i^2, $i = 1, \ldots, n$, allora la variabile aleatoria $\sum_{i=1}^n a_i X_i$, per $a_1, a_2, \ldots, a_i$ costanti, ha distribuzione normale con media e varianza espressi da

$$E\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i \mu_i$$
$$V\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i^2 \sigma_i^2 + 2\sum_{i<j}\sum a_i a_j \mathrm{Cov}(X_i, X_j).$$

4.4 Variabili casuali ordinate

Siano $X_1, X_2, \ldots, X_n$, n variabili aleatorie date, indipendenti e con stessa funzione di densità $f(x)$ e funzione di ripartizione $F(x)$. Si ricordi che, dalla definizione di variabile aleatoria, per ogni $\omega \in \Omega$, $X_i(\omega)$ è un numero reale. Allora possiamo considerare le variabili aleatorie

$$X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)} \tag{4.22}$$

che si ottengono ordinando le X_i dal valore più piccolo al più grande, in modo tale che

$$X_{(1)} = \min\{X_1, X_2, \ldots, X_n\},$$
$$X_{(n)} = \max\{X_1, X_2, \ldots, X_n\}.$$

Come accennato in precedenza, la famiglia di v.a. ordinate definita come in (4.22) è detta delle *statistiche d'ordine*, e si denota con $X_{(k)}$ la v.a. posizionata al k-mo ordine della graduatoria non decrescente.

È interessante determinare le distribuzioni di $X_{(1)}$ e $X_{(n)}$. Si consideri, innanzitutto, che l'evento $\{X_{(n)} \leq x\}$ si verifica se, e soltanto se, risulta $\{X_1 \leq x, X_2 \leq x, \ldots, X_n \leq x\}$. La funzione di ripartizione di $X_{(n)}$ può essere espressa in termini delle distribuzioni marginali di $X_1, X_2, \ldots, X_n$, assumendo la loro indipendenza:

$$\begin{aligned} F_{X_{(n)}}(x) &= P(X_{(n)} \leq x) \\ &= P(X_1 \leq x, X_2 \leq x, \ldots, X_n \leq x) \\ &= P(X_1 \leq x) P(X_2 \leq x) \cdots P(X_n \leq x); \end{aligned}$$

considerando che le X_i hanno stessa distribuzione, si ha

$$F_{X_{(n)}}(x) = [F(x)]^n. \tag{4.23}$$

La funzione di densità di $X_{(n)}$ si ricava facilmente come segue:

$$\begin{aligned} f_{X_{(n)}}(x) &= \frac{d}{dx} F_{X_{(n)}}(x) \\ &= \frac{d}{dx}[F(x)]^n \\ &= n[F(x)]^{n-1} f(x), \end{aligned}$$

essendo $dF(x)/dx = f(x)$. Si osservi che sia la funzione di ripartizione che la densità del massimo di $X_1, X_2, \ldots, X_n$, $X_{(n)}$, dipendono esclusivamente dalla densità e dalla funzione di ripartizione che accomuna le X_i, $i = 1, \ldots, n$.

Con un ragionamento analogo si ricava la distribuzione del minimo $X_{(1)}$ di $X_1, X_2, \ldots, X_n$. Si consideri l'equivalenza degli eventi $\{X_{(1)} > x\}$ e $\{X_1 > x, X_2 > x, \ldots, X_n > x\}$. Allora risulta

$$\begin{aligned} P(X_{(1)} > x) &= P(X_1 > x, X_2 > x, \ldots, X_n > x) \\ &= P(X_1 > x) P(X_2 > x) \cdots P(X_n > x) \\ &= [1 - F(x)]^n \end{aligned}$$

dalle stesse considerazioni fatte in precedenza sulla distribuzione delle variabili X_i. Ne segue che

$$F_{X_{(1)}}(x) = P(X_{(1)} \leq x) = 1 - [1 - F(x)]^n. \tag{4.24}$$

La funzione di densità di $X_{(1)}$ è data da

$$\begin{aligned} f_{X_{(1)}}(x) &= \frac{d}{dx} F_{X_{(1)}}(x) \\ &= \frac{d}{dx}(1 - [1 - F(x)]^n) \\ &= n[1 - F(x)]^{n-1} f(x). \end{aligned}$$

Esempio 4.9.

Siano $X_1, X_2, \ldots, X_n$ n variabili aleatorie indipendenti con distribuzione esponenziale,

$$\begin{aligned} f(x) &= \lambda e^{-\lambda x}, \quad x > 0 \\ F(x) &= 1 - e^{-\lambda x}, \quad x > 0. \end{aligned}$$

Allora, se $X_{(1)}$ è il minimo delle n variabili esponenziali, si ha

$$F_{X_{(1)}}(x) = 1 - [1 - F(x)]^n = 1 - e^{-n\lambda x}, \quad x > 0$$

e inoltre

$$f_{X_{(1)}}(x) = n\lambda e^{-n\lambda x}, \quad x > 0.$$

Ne segue che $X_{(1)}$ è una variabile aleatoria esponenziale di parametro $n\lambda$. Il massimo, $X_{(n)}$, ha funzione di ripartizione

$$F_{X_{(n)}}(x) = [F(x)]^n = [1 - e^{-\lambda x}]^n, \quad x > 0$$

e funzione di densità

$$f_{X_{(n)}}(x) = n\lambda e^{-\lambda x}[1 - e^{-\lambda x}]^{n-1}, \quad x > 0,$$

da cui si deduce che il massimo non ha distribuzione esponenziale.

Esempio 4.10.

Siano $X_1, X_2, \ldots, X_n$ n variabili aleatorie indipendenti $X_i \sim U(0, \theta)$, $i = 1, \ldots, n$. Allora la funzione di densità della v.a. minimo $X_{(1)}$ è:

$$f_{X_{(1)}}(x) = \frac{n}{\theta}\left(1 - \frac{x}{\theta}\right)^{n-1} \quad \text{per} \quad x \in (0, \theta),$$

e 0 altrove. Possiamo ricavare i momenti della v.a. $X_{(1)}$ come segue:

$$\begin{aligned} E(X_{(1)}^r) =& \int_0^\theta x^r \frac{n}{\theta}\left(1 - \frac{x}{\theta}\right)^{n-1} dx = \int_0^1 (w\theta)^r \frac{n}{\theta}(1 - w)^{n-1}\theta dw \\ =& n\theta^r \int_0^1 w^r (1 - w)^{n-1} dw = n\theta^r B(r + 1, n) \\ =& \theta^r \frac{n\Gamma(n)\Gamma(r + 1)}{\Gamma(r + 1 + n)} = \frac{\theta^r}{\binom{r+n}{r}} \quad r = 0, 1, 2, \ldots \end{aligned}$$

In particolare, risulta $E(X_{(1)}) = \theta/(n + 1)$.

In modo analogo, la funzione di densità della v.a. massimo $X_{(n)}$ è:

$$f_{X_{(n)}}(x) = \frac{n}{\theta}\left(\frac{x}{\theta}\right)^{n-1} \quad \text{per} \quad x \in (0, \theta),$$

e 0 altrove. I momenti della v.a. $X_{(n)}$ sono:

$$E(X_{(n)}^r) = \int_0^\theta x^r \frac{n}{\theta}\left(\frac{x}{\theta}\right)^{n-1} dx = \frac{n}{\theta^n} \int_0^\theta x^{r+n-1} dx = \frac{n\theta^r}{r + n}.$$

In particolare, si ha $E(X_{(n)}) = n\theta/(n+1)$.

Se $X_i \sim U(0,1)$ allora la v.a. ordinata $X_{(k)}$ si distribuisce come una v.a. Beta $Beta(k, n-k+1)$ (si veda il paragrafo 2.2.7) e la funzione di densità è:

$$f_{X_{(k)}}(x) = n\binom{n-1}{k-1}x^{k-1}(1-x)^{n-k} = \frac{1}{B(k, n-k+1)}x^{k-1}(1-x)^{n-k}, x \in (0,1).$$

Il valore atteso della variabile $X_{(k)}$ è $E(X_{(k)}) = k/(n+1)$, $k = 1, 2, \dots, n$.

4.5 La disuguaglianza di Chebyshev e la legge dei grandi numeri

In questa sezione vengono introdotte alcune disuguaglianze probabilistiche notevoli, per prima quella attribuita al matematico russo Chebyshev. L'utilità di tale risultato sarà evidente nel seguito nella discussione della nota *legge dei grandi numeri*, che fornisce una giustificazione all'approccio frequentista alla probabilità che è stato adottato in queste dispense.

Teorema 4.7 (Disuguaglianza di Chebyshev)**.** Sia X una variabile aleatoria qualsiasi tale che $E(X^2)$ esista. Allora

$$P(|X| \geq \varepsilon) \leq \frac{1}{\varepsilon^2}E(X^2). \tag{4.25}$$

Dimostrazione. Sia X una variabile aleatoria e ε una qualsiasi costante positiva. Si definisca la variabile aleatoria Y_ε come

$$\begin{aligned} Y_\varepsilon = & 1 \quad \text{se} \quad |X| \geq \varepsilon \\ & 0 \quad \text{se} \quad |X| < \varepsilon \end{aligned}$$

Quindi

$$E(Y_\varepsilon) = 1 \cdot P(|X| \geq \varepsilon) + 0 \cdot P(|X| < \varepsilon) = P(|X| \geq \varepsilon)$$

e sicuramente vale che

$$X^2 \geq X^2 Y_\varepsilon \geq \varepsilon^2 Y_\varepsilon$$

Da ciò segue che

$$E(X^2) \geq E(\varepsilon^2 Y_\varepsilon) = \varepsilon^2 E(Y_\varepsilon) = \varepsilon^2 P(|X| \geq \varepsilon),$$

che prova la disuguaglianza (4.25). □

Si noti che questa semplice disuguaglianza ha validità per ogni variabile aleatoria, discreta o continua. Se X è uniforme in $(-1, 1)$, allora

$$\begin{aligned} f(x) &= \frac{1}{2}, \quad -1 < x < 1 \\ E(X^2) &= \int_{-1}^{1} \frac{x^2}{2} dx = \frac{1}{3} \\ P(|X| \geq \varepsilon) &= 0, \qquad \varepsilon \geq 1 \\ &= 1 - \varepsilon, \qquad \varepsilon < 1, \end{aligned}$$

e si ha

$$1 - \varepsilon \leq \frac{1}{3\epsilon^2}.$$

Se X è normale standard,

$$\begin{aligned} E(X^2) &= 1 \\ P(|X| \geq \varepsilon) &= 1 - P(|X| \leq \varepsilon) \\ &= 1 - P(-\varepsilon < X < \varepsilon) \\ &= 1 - (2\Phi(\varepsilon) - 1) \\ &= 2 - 2\Phi(\varepsilon) \leq 1/\varepsilon^2 \end{aligned}$$

Ora sia Y una variabile aleatoria tale che

$$E(Y) = \mu \quad \text{e} \quad V(Y) = \sigma^2$$

e definiamo $X = (Y - \mu)/\sigma$. Allora

$$E(X^2) = \frac{1}{\sigma^2} E((Y - \mu)^2) = 1,$$

e la disuguaglianza di Chebyshev diventa

$$P\left(\left|\frac{Y - \mu}{\sigma}\right| \geq \varepsilon\right) \leq \frac{1}{\varepsilon^2},$$

cioè

$$P(|Y - \mu| \geq \varepsilon\sigma) \leq \frac{1}{\varepsilon^2}, \tag{4.26}$$

ammesso che la varianza di Y, σ^2, esista. Possiamo anche scrivere che, per ogni numero reale $z > 0$,

$$P(|Y - \mu| < z) \geq 1 - \frac{\sigma^2}{z^2}. \tag{4.27}$$

La disuguaglianza di Chebyshev nella forma (4.27) permette quindi di valutare la probabilità che una variabile casuale ha di assumere valori che differiscano dalla media per un quantità non più grande di un numero dato, fissato a priori. Si noti anche che la disuguaglianza di Chebyshev è, in realtà, una diretta conseguenza della disuguaglianza di Markov. Infatti dalla (1.9), considerata la variabile aleatoria

$$W = (Y - E(Y))^2$$

ed essendo soddisfatta l'ipotesi alla base della disuguaglianza di Markov $P(W \geq 0) = 1$, si ha

$$P(W \geq z^2) \leq \frac{E(W)}{z^2}.$$

Essendo $E(W) = E[(Y - E(Y))^2] = V(Y)$ possiamo anche scrivere che, per ogni numero reale $z > 0$

$$P(|Y - E(Y)| \geq z) \leq \frac{V(Y)}{z^2},$$

che è equivalente alla (4.27).

Come si vedrà la disuguaglianza di Chebyshev è utile per stabilire le cosiddette *leggi deboli dei grandi numeri.* Prima di introdurre queste ultime è opportuno dare alcuni cenni sulla convergenza di una successione di variabili aleatorie, propedeutici rispetto ad alcuni argomenti dei prossimi capitoli.

Esempio 4.11.

Si consideri la variabile aleatoria $Beta(0.5, 0.5)$ definita dalla funzione di densità

$$f(x) = \frac{1}{\pi} x^{-1/2}(1-x)^{-1/2}, \quad 0 < x < 1.$$

che ha varianza pari a $1/8$. Supponiamo di voler determinare il limite inferiore della probabilità $P(\mu - 1.1\sigma < X < \mu + 1.1\sigma)$ con la disuguaglianza di Chebyshev. Dalla (4.26) si ottiene immediatamente

$$P(|X-\mu| < \varepsilon\sigma) \geq 1 - \frac{1}{\varepsilon^2},$$

per cui si ottiene

$$P(\mu - 1.1\sigma < X < \mu + 1.1\sigma) = P(|X-\mu| < 1.1\sqrt{1/8}) \geq 1 - \frac{1}{1.1^2} = 0.17.$$

Per arrivare a definire il concetto di convergenza di una successione di variabili aleatorie si può partire da un esempio. Si considerino n variabili aleatorie indipendenti $(X_1, X_2, \ldots, X_n)$, in cui X_i ha distribuzione uniforme in (0,1). Sia $Y_n = \max\{X_1, X_2, \ldots, X_n\}$. La funzione di ripartizione di Y_n è

$$F_n(y) = P(Y_n \leq y) = P(X_1 \leq y, X_2 \leq y, \ldots, X_n \leq y)$$

ed essendo $P(X_i \leq y) = y$, $i = 1, \ldots, n$, dall'indipendenza segue

$$F_n(y) = P(X_1 \leq y)P(X_1 \leq y)\cdots P(X_n \leq y) = y^n, \quad 0 < y < 1.$$

Al variare di n, Y_n costituisce una successione di variabili aleatorie $\{Y_n\}$ di cui ha senso studiare il comportamento al divergere di n. Si osservi che

$$\lim_{n\to\infty} F_n(y) = F(y),$$

dove

$$F(y) = \begin{cases} 0, & \text{se} \quad y < 1 \\ 1, & \text{se} \quad y \geq 1, \end{cases}$$

si può dire che la distribuzione limite di Y_n è $F(y)$.

In generale. esistono vari modi per definire la convergenza di una successione X_1, X_2, ... di variabili aleatorie definite su uno stesso spazio campionario Ω, verso la variabile aleatoria X.

Definizione 4.1 (Convergenza *quasi certa*)**.** Si dice che la successione $\{X_n\}$ converge quasi certamente a X (e si scrive $X_n \overset{qc}{\to} X$) se

$$P(\lim_{n\to\infty} X_n = X) = 1.$$

Tale convergenza viene detta anche *convergenza quasi ovunque*, *convergenza forte*, o "con probabilità 1". La convergenza quasi certa comporta che, per quasi tutti gli elementi di Ω vale $\underset{n\to\infty}{lim} X_n = X$ e costituisce un criterio molto forte di convergenza tra variabili aleatorie.

Definizione 4.2 (Convergenza in probabilità)**.** Si dice che la successione $\{X_n\}$ converge a X in probabilità (e si scrive $X_n \overset{p}{\to} X$) se

$$\lim_{n\to\infty} P(|X_n - X| < \varepsilon) = 1.$$

In questo caso si parla anche di *convergenza in misura* o *convergenza debole*. Questa convergenza implica una vicinanza probabilistica tra X_n e X nel senso che, al crescere di n, l'evento $(|X_n - X| < \varepsilon)$ diventa certo, cioè la v.a X_n tende ad assumere valori che si discosteranno sempre meno, in valore assoluto, dalla v.a. X verso cui converge.

Può essere più semplice, in alcuni casi, studiare la convergenza dei momenti, e in particolare la convergenza definita mediante i momenti del secondo ordine.

Definizione 4.3 (Convergenza in media quadratica)**.** Si dice che la successione $\{X_n\}$ converge a X in media quadratica (e si scrive $X_n \overset{m}{\to} X$) se

$$\lim_{n\to\infty} E(X_n - X)^2 = 0.$$

La definizione afferma, in sostanza, che la differenza quadratica media tende ad annullarsi, per n che tende all'infinito.

La più semplice forma di convergenza è quella in distribuzione.

Definizione 4.4 (Convergenza in distribuzione)**.** Sia $F_n(x)$ la funzione di ripartizione della generica variabile aleatoria X_n della successione $\{X_n\}$ e sia $F(x)$ la funzione di ripartizione di X. Allora si dice che la successione $\{X_n\}$ converge in distribuzione a X se in ogni punto di continuità x di $F(x)$ si ha

$$\lim_{n\to\infty} F_n(x) = F(x)$$

e scriveremo $X_n \xrightarrow{d} X$.

La nozione di convergenza in distribuzione può essere formulata in altri modi equivalenti. La successione $\{X_n\}$ converge in distribuzione a X se, per ogni coppia di punti a e b, con $a < b$, ove $F(x)$ è continua, al tendere di n all'infinito

$$P(a < X_n \leq b) = F_n(b) - F_n(a) \to F(b) - F(a) = P(a < X \leq b) \qquad (4.28)$$

Un'altra definizione di convergenza in distribuzione che sarà utile in seguito è quella che si evince dal teorema seguente.

Teorema 4.8. Sia $\{X_n\}$ una successione di variabili aleatorie; siano $F_n(x)$ ed $m_{X_n}(t)$, rispettivamente, la funzione di ripartizione e la funzione generatrice dei momenti della generica variabile aleatoria della successione. Se $F(x)$ è la funzione di ripartizione di X avente come funzione generatrice dei momenti $m_X(t)$, allora la successione $\{X_n\}$ converge in distribuzione X se e solo se

$$\lim_{n\to\infty} m_{X_n}(t) = m_X(t),$$

per ogni t compreso in un intervallo aperto contenente lo zero.

Questo risultato, la cui dimostrazione è omessa, afferma che se la successione di funzioni generatrici di momenti corrispondente ad una successione di variabili aleatorie converge alla funzione generatrice di momenti di una data variabile aleatoria, allora a questa converge la successione della variabili aleatorie.

Tra le quattro definizioni di convergenza sussiste una precisa connessione: la convergenza in media quadratica, in probabilità e la convergenza quasi certa implicano tutte la convergenza in distribuzione, definita tramite le funzioni di ripartizione. Viceversa, la convergenza in distribuzione non implica gli altri tipi di convergenza. Inoltre la convergenza in media quadratica e la convergenza quasi certa implicano quella in probabilità. Tali implicazioni possono essere così schematizzate

$$\begin{array}{ccccc} \boxed{X_n \stackrel{qc}{\to} X} & \Rightarrow & \boxed{X_n \stackrel{p}{\to} X} & \Leftarrow & \boxed{X_n \stackrel{m}{\to} X} \\ & & \Downarrow & & \\ & & \boxed{X_n \stackrel{d}{\to} X} & & \end{array}$$

Ad esempio, è facile verificare che la convergenza in media quadratica implica la convergenza in probabilità considerando la disuguaglianza di Chebyshev nella forma

$$P(|X_n - X| \geq \varepsilon) \leq \frac{1}{\varepsilon^2} E(X_n - X)^2,$$

da cui si comprende che se il secondo membro della disuguaglianza tende a 0 anche il primo tende a 0, per $n \to \infty$, che implica $X_n \stackrel{p}{\to} X$.

4.5.1 Leggi dei grandi numeri

Nella teoria della probabilità sono stati dimostrati alcuni teoremi chiamati *leggi dei grandi numeri*, che riguardano il comportamento della successione delle medie campionarie. Sia $\{X_n\}$ una successione di variabili aleatorie e sia $\{\overline{X}_n\}$ la corrispondente *successione delle medie campionarie*

$$\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i.$$

Sotto quali condizioni si può affermare che la successione $\{\bar{X}_n\}$ converge ad una costante μ? O più in generale, sotto quali condizioni la successione $\{\overline{X}_n - \bar{\mu}_n\}$ converge a 0, essendo $\bar{\mu}_n$ la media aritmetica relativa alla successione $\{\mu_n\}$? Le leggi dei grandi numeri rispondono a questi interrogativi.

Il risultato più importante in questo ambito è il seguente.

Teorema 4.9 (Legge debole dei grandi numeri)**.** Sia $\{X_n\}$ una successione di variabili aleatorie indipendenti e identicamente distribuite con media $E(X_i) = \mu$ e varianza $V(X_i) = \sigma^2 < \infty$. Allora, definita la variabile aleatoria

$$\overline{X}_n = \sum_{i=1}^{n} \frac{X_i}{n}$$

per ogni $\delta > 0$, si ha

$$\lim_{n\to\infty} P(|\overline{X}_n - \mu| < \delta) = 1.$$

cioÃ¨ la successione $\{\overline{X}_n\}$ converge in probabilità a μ.

Dimostrazione. Si noti anzitutto che il valore atteso e la varianza della media delle n variabili aleatorie sono espressi da

$$E(\overline{X}_n) = \sum_{i=1}^{n} \frac{\mu}{n} = \mu, \quad V(\overline{X}_n) = \frac{1}{n^2}\sum_{i=1}^{n} \sigma^2 = \frac{\sigma^2}{n}.$$

Pertanto, dalla disuguaglianza di Chebyshev, si ricava

$$P(|\overline{X}_n - \mu| \geq \delta) \leq \frac{E(\overline{X}_n - \mu)^2}{\delta^2} = \frac{\sigma^2}{n\delta^2}$$

Quindi si ottiene

$$P(|\overline{X}_n - \mu| < \delta) = 1 - P(|\overline{X}_n - \mu| \geq \delta) \geq 1 - \sigma^2/(n\delta^2)$$

Considerando il limite per $n \to \infty$ si ha

$$\lim_{n\to\infty} P(|\overline{X}_n - \mu| < \delta) \geq \lim_{n\to\infty} (1 - \sigma^2/(n\delta^2)) = 1$$

per ogni $\delta > 0$.

□

Una formulazione equivalente della legge debole dei grandi numeri è la seguente

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

Una formulazione molto nota della legge dei grandi numeri si deve a Bernoulli, che storicamente ha costituito il primo nucleo attorno al quale altri studiosi (Poisson, Chebyshev, Khintchin) hanno poi sviluppato le formulazioni più generali della legge dei grandi numeri.

Si considerino prove ripetute di un esperimento, indipendenti e caratterizzate dalla stessa distribuzione di probabilità (n misure della stessa grandezza, n lanci della stessa moneta, ecc). Si assuma inoltre che la probabilità che un certo evento A si verifichi, $P(A)$, sia costante in ogni prova. Considerando una serie di prove ripetute un gran numero di volte nelle stesse condizioni, ci si attende, dunque, che la frequenza relativa osservata con cui l'evento A si manifesta nelle n prove (ovvero il rapporto tra il numero di volte che l'evento A si verifica ed n) possa essere all'incirca uguale alla probabilità dell'evento $P(A)$. In altre parole, l'esperienza suggerisce che all'aumentare del numero delle osservazioni la media aritmetica tende a stabilizzarsi intorno ad una costante, che può allora assumersi come l'entità incognita della grandezza. La tendenza verso la probabilità della frequenza relativa con cui un evento si manifesta in una lunga serie di prove è nota come *legge empirica del caso* o *legge empirica dei grandi numeri.*

Sia X una variabile aleatoria Bernoulliana, che può assumere quindi i valori 1 e 0 con probabilità p e $1-p=q$, rispettivamente, e si consideri la variabile aleatoria

$$S_n = X_1 + X_2 + \cdots + X_n,$$

somma di n variabili aleatorie indipendenti ciascuna di tipo Bernoulliano. Poiché $E(X_i) = p$ e $V(X_i) = pq$, $i = 1, \ldots, n$, si avrà $E(S_n) = np$ e $V(S_n) = npq$. Si noti come S_n è semplicemente il numero di "successi" in n repliche di un esperimento con due possibili risultati dove la probabilità di "successo" è p e quella di "insuccesso" è q. Si consideri la frequenza relativa dei successi in n prove $R_n = S_n/n$; si ha $E(R_n) = p$ e $V(R_n) = pq/n$. Bernoulli prova il teorema seguente che stabilisce le condizioni sufficienti perché sia legittimo aspettarsi che valga la legge empirica del caso.

Teorema 4.10. Sia R_n la frequenza relativa dei "successi" in n prove indipendenti di un esperimento a due alternative con probabilità p per l'evento "successo" e $1 - p = q$ per l'evento "insuccesso". Allora

$$\lim_{n\to\infty} P(|R_n - p| < \delta) = 1.$$

Dimostrazione. La dimostrazione del teorema è immediata dall'applicazione della disuguaglianza di Chebyshev alla variabile aleatoria $R_n = S_n/n$ di media p e varianza pq/n:

$$P(|R_n - p| \geq \varepsilon\sqrt{pq/n}) \leq \frac{1}{\varepsilon^2},$$

da cui

$$P(|R_n - p| < \varepsilon\sqrt{pq/n}) > 1 - \frac{1}{\varepsilon^2}$$

per ogni $\varepsilon > 0$. Posto $\varepsilon = \frac{\delta\sqrt{n}}{\sqrt{pq}}$

$$P(|R_n - p| < \delta) > 1 - \frac{pq}{n\delta^2}$$

da cui la tesi. □

Esempio 4.12.

Si assuma che sia disponibile un nuovo processo per la produzione di chip informatici in silicio, e che sia p la probabilità (ignota) che ciascun chip prodotto in tal modo sia difettoso. Inoltre si sa che un chip può essere difettoso o non difettoso indipendentemente dagli altri. Si richiede di determinare il numero n di chip da produrre e testare in modo tale che la proporzione di pezzi difettosi osservati non differisca da p per più di 0.01 con probabilità almeno pari a 0.9. Ovvero si richiede n tale che

$$P(|\overline{X}_n - p| < 0.01) > 0.9.$$

dove $\overline{X}_n = \sum_{i=1}^n X_i/n$, per $X_1, \ldots, X_n$ variabili aleatorie bernoulliane indipendenti di parametro p. Per la disuguaglianza di Chebyshev, possiamo scrivere

$$P(|\overline{X}_n - p| \geq \varepsilon\sqrt{p(1-p)/n}) \leq \frac{1}{\varepsilon^2},$$

cioè

$$P(|\overline{X}_n - p| < \varepsilon\sqrt{p(1-p)/n}) > 1 - \frac{1}{\varepsilon^2},$$

per ogni $\varepsilon > 0$. Il valore p è tale che $0 \leq p \leq 1$, e chiaramente $p(1-p) \leq 1/4$. Quindi si ha

$$\sqrt{\frac{p(1-p)}{n}} \leq \sqrt{\frac{1}{4n}} = \frac{1}{2\sqrt{n}}.$$

Pertanto deve essere verificata la condizione

$$P\left(|\overline{X}_n - p| < \frac{\varepsilon}{2\sqrt{n}}\right) > 1 - \frac{1}{\varepsilon^2}$$

da cui, ponendo $\delta = \varepsilon/2\sqrt{n}$, si ottiene

$$P(|\overline{X}_n - p| < \delta) > 1 - \frac{1}{4n\delta^2}.$$

Affinché tale probabilità sia almeno pari a γ poniamo $\gamma = 1 - 1/4n\delta^2$ e risolvendo per n si trova

$$n = \frac{1}{4\delta^2(1-\gamma)}.$$

Sostituendo $\delta = 0.01$ e $\gamma = 0.9$ si ottiene

$$n = \frac{1}{4(0.01)^2(0.1)} = 25000.$$

Si supponga allora di aver effettivamente prodotto e successivamente testato $n = 25000$ chip, trovando che 500 di essi sono difettosi. Pertanto la proporzione osservata di pezzi difettosi risulta essere $500/25000 = 0.02$. La probabilità che tale numero non differisca da p per più di 0.01 è quindi almeno pari a 0.9, cioè è almeno 0.9 la probabilità che il valore incognito p si trovi tra 0.01 e 0.03 (che sono i valori che che distano 0.01 dal valore trovato 0.02). La probabilità esatta può essere, in realtà, molto maggiore di 0.9, per le seguenti due ragioni: (i) il valore di n è stato ricavato mediante la disuguaglianza di Chebyshev, che fornisce un limite per la probabilità considerata; (ii) derivando n si è sostituito $p(1-p)$ con $1/4$, che rende la varianza più grande possibile. Nella prossima sezione, si utilizzerà il teorema del limite centrale per determinare, in modo più accurato, l'ampiezza campionaria n necessaria affinché sia soddisfatta una certa condizione sulla probabilità.

È importante osservare che le formulazioni della legge dei grandi numeri fin qui considerate descrivono il valore limite per una successione di probabilità, che è diverso dal considerare il valore limite di $\overline{X}_n$, ammesso che esso esista. Vale a dire, calcolando $|\overline{X}_n - \mu|$, per valori sempre più grandi di n, la *probabilità* che tale differenza ecceda un δ fissato diventa via via più piccola e converge a 0. Si parla quindi di *leggi deboli dei grandi numeri*, che vanno distinte dalle cosiddette *leggi forti* che implicano, invece, la convergenza quasi certa delle stesse successioni. Di queste ultime si enuncia di seguito (senza darne dimostrazione) il risultato principale.

Teorema 4.11. Sia $\{\overline{X}_n\}$ una successione di variabili aleatorie indipendenti tali che $E(X_i) = \mu$ e $V(X_i) = \sigma^2 < \infty$. Definita $\overline{X}_n = (1/n)\sum_{i=1}^n X_i$, per ogni $\varepsilon > 0$ risulta

$$P(\lim_{n\to\infty} |\overline{X}_n - \mu| < \varepsilon) = 1,$$

cioè la successione $\{\overline{X}_n\}$ converge quasi certamente a μ.

4.6 Teorema del limite centrale e approssimazioni di distribuzioni di probabilità

Il teorema del limite centrale, di fondamentale importanza nel calcolo delle probabilità e nell'inferenza statistica, riguarda la legge di probabilità di somme di variabili aleatorie quando n, il numero dei termini sommati, cresce all'infinito. Nel paragrafo 2.2.6 si è considerata la somma di variabili aleatorie Bernoulliane, e visto come la legge di tale somma fosse descritta, al limite, da una funzione di densità, la densità normale. Questo risultato verrà generalizzato, considerando somme di variabili aleatorie la cui distribuzione non si limita a quella Bernoulliana.

Teorema 4.12 (Teorema del limite centrale)**.** Siano $X_1, X_2, \dots, X_n$ n variabili aleatorie indipendenti e identicamente distribuite con media $E(X_i) = \mu$ e $V(X_i) = \sigma^2$. Sia $S_n = X_1 + X_2 + \cdots + X_n$, con $E(S_n) = n\mu$ e $V(S_n) = n\sigma^2$. Allora la successione di variabili aleatorie $\{Z_n\}$, essendo Z_n la variabile standardizzata corrispondente a S_n data da

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}, \quad n = 1, 2, 3, \dots$$

converge in distribuzione ad una normale standardizzata $\mathcal{N}(0, 1)$, ovvero

$$\lim_{n\to\infty} P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \le z\right) = \Phi(z), \tag{4.29}$$

dove $\Phi(\cdot)$ è la funzione di ripartizione della normale standardizzata.

Dimostrazione. Si tratta di dimostrare che

$$\lim_{n\to\infty} P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \le z\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-y^2/2} dy.$$

Posto $Z_n = (S_n - n\mu)/\sigma\sqrt{n}$, la dimostrazione è condotta utilizzando la funzione generatrice dei momenti di Z_n, per dimostrare che la successione $\{Z_n\}$ converge in distribuzione alla $\mathcal{N}(0, 1)$. Si consideri quindi l'identità

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} = V_1 + V_2 + \cdots + V_n,$$

con $V_i = (X_i - \mu)/\sigma\sqrt{n}$. Assumendo che X_i ammetta funzione generatrice dei momenti, data l'indipendenza delle V_i, $i = 1, \dots, n$, scriviamo la funzione generatrice dei momenti di Z_n come

$$m_{Z_n}(t) = m_{V_1}(t) m_{V_2}(t) \cdots m_{V_n}(t). \tag{4.30}$$

Inoltre, essendo $E(X_i) = \mu$ e $V(X_i) = \sigma^2$ per ogni i, si ha che ciascuna delle V_i ha media 0, varianza $1/n$ e funzione generatrice dei momenti data da (in virtù della (1.20))

$$m_{V_i}(t) = E(e^{tV_i}) = E(e^{t(X_i - \mu)/\sigma\sqrt{n}}) = m_{X_i - \mu}\left(\frac{t}{\sigma\sqrt{n}}\right).$$

I fattori che appaiono nel secondo membro della (4.30) sono tutti uguali, facendo riferimento alla prima possiamo scrivere

$$m_{Z_n}(t) = \left[m_{X_1-\mu}\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n. \tag{4.31}$$

Sappiamo inoltre che

$$m_{X_1-\mu}\left(\frac{t}{\sigma\sqrt{n}}\right) = 1+E(X_1-\mu)\frac{t}{\sigma\sqrt{n}}+E(X_1-\mu)^2\frac{t^2}{2\sigma^2 n}+E(X_1-\mu)^3\frac{t^3}{3!\sigma^3 n^{3/2}}+\dots$$

ed essendo $E(X_1-\mu) = 0$ e $E(X_1-\mu)^2 = \sigma^2$, l'espressione precedente diviene

$$m_{X_1-\mu}\left(\frac{t}{\sigma\sqrt{n}}\right) = 1 + \frac{t^2}{2n} + \frac{\overline{m}_3 t^3}{3!\sigma^3 n^{3/2}} + \dots$$

Sostituendo nella (4.31) si ha

$$m_{Z_n}(t) = \left[1 + \frac{t^2}{2n} + \frac{\overline{m}_3 t^3}{3!\sigma^3 n^{3/2}} + \dots\right]^n.$$

Un risultato noto dell'analisi afferma che se $\lim_{n\to\infty} w_n = w$, allora

$$\lim_{n\to\infty}\left(1 + \frac{w_n}{n}\right)^n = e^w.$$

Quindi se

$$w_n = \frac{t^2}{2} + \frac{\overline{m}_3 t^3}{3!\sigma^3 n^{1/2}} + \dots$$

si ha

$$\lim_{n\to\infty} w_n = \frac{t^2}{2}$$

e pertanto

$$\lim_{n\to\infty} m_{Z_n}(t) = e^{t^2/2}.$$

Il risultato dell'ultimo limite è la funzione generatrice dei momenti di una normale $\mathcal{N}(0,1)$ e si può quindi concludere che la distribuzione limite di Z_n è una normale standardizzata. □

Il rilievo che il teorema del limite centrale ha sul piano applicativo è dovuto al fatto che la (4.29) è valida in modo approssimato anche per n finito; ovvero, per n sufficientemente elevato ci si attende che

$$P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq z\right)$$

sia ben approssimata da $\Phi(z)$. Ciò è equivalente ad affermare che la distribuzione della somma $S_n = \sum_{i=1}^n X_i$ si può approssimare, purché n sia sufficientemente elevata, con la distribuzione normale di media $n\mu$ e varianza $n\sigma^2$:

$$\begin{aligned} F_{S_n}(s) &= P(S_n \leq s) \\ &= P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq \frac{s - n\mu}{\sigma\sqrt{n}}\right) \\ &\doteq \Phi\left(\frac{s - n\mu}{\sigma\sqrt{n}}\right). \end{aligned}$$

Esempio 4.13.

Si assume frequentemente che la durata di vita di una certa componente elettronica sia una variabile aleatoria esponenziale di parametro λ (per cui $\mu = \sigma = 1/\lambda$). In particolare si consideri un macchinario il cui funzionamento dipende esclusivamente da una certa componente elettronica; si assuma inoltre che la durata di vita della componente i-ma, X_i, sia esponenziale di parametro $\lambda = 0.002$ (per ora), $i = 1, \ldots, 20$. Le X_i sono indipendenti, e quando una data componente smette di funzionare viene istantaneamente sostituita da una nuova. La durata di vita del macchinario è allora la durata complessiva delle 20 componenti, $T = \sum_{i=1}^{20} X_i$. Per il teorema del limite centrale T ha distribuzione approssimata da una normale con

$$\mu_T = \frac{20}{0.002} = 10000 \text{ ore}, \quad \sigma_T = \sqrt{20/(0.002)^2} = 2236 \text{ ore}.$$

Volendo calcolare la probabilità che il macchinario funzioni per almeno un anno (8760 ore) si utilizzerà quindi l'approssimazione normale come segue:

$$\begin{aligned} P(T > 8760) &= 1 - P(T \leq 8760) \\ &\doteq 1 - \Phi\left(\frac{8760 - 10000}{2236}\right) \\ &= 1 - \Phi(-0.55) = 0.71. \end{aligned}$$

Spesso si è più interessati alla media $(1/n)\sum_i^n X_i$ che alla somma. Si noti però che se è nota la distribuzione della somma, si può facilmente derivare la distribuzione della media. Infatti, si ha

$$F_{(1/n)\sum X_i}(z) = P\left(\frac{1}{n}\sum_{i=1}^n X_i \le z\right) = P\left(\sum_{i=1}^n X_i \le nz\right) = F_{\sum X_i}(nz).$$

Il teorema del limite centrale fornisce quindi una *distribuzione approssimata della media*: se $X_1, X_2, \ldots, X_n$ sono variabili aleatorie indipendenti, ognuna con la stessa distribuzione dotata di media μ e varianza σ^2, allora $\overline{X}_n = (1/n)\sum_{i=1}^n X_i$ "standardizzata" ha una distribuzione che tende a una normale standardizzata. Ovvero, essendo $E(\overline{X}_n) = \mu$ e $V(\overline{X}_n) = \sigma^2/n$, si ha

$$\frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \,\dot{\sim}\, \mathcal{N}(0,1), \tag{4.32}$$

ovvero $\overline{X}_n$ ha approssimativamente una distribuzione $\mathcal{N}(\mu, \sigma^2/n)$.

È opportuno qui analizzare e distinguere due termini connessi al teorema del limite centrale. Una funzione di ripartizione viene chiamata *funzione di ripartizione limite* se è la funzione limite di una successione di funzioni di ripartizione. La (4.29) ne è un esempio: $\Phi(z)$ è la funzione di ripartizione limite di una successione di funzioni di ripartizione $F_{Z_1}(\cdot), F_{Z_2}(\cdot), \ldots, F_{Z_n}(\cdot)$. Una *funzione di ripartizione asintotica*, o distribuzione asintotica, di una variabile casuale Y_n, in una successione di variabili casuali $Y_1, Y_2, \ldots, Y_n, \ldots$ è una distribuzione approssimativamente uguale alla vera distribuzione di Y_n, per n abbastanza grande. Quindi, ad esempio, diciamo che $\overline{X}_n$ ha una distribuzione asintotica normale con media μ e varianza σ^2/n. Si noti che una distribuzione asintotica può dipendere da n, mentre una distribuzione limite no. I due concetti, pur essendo distinti, sono strettamente connessi: se la distribuzione di Z_n converge a $\Phi(z)$, allora per n grande la distribuzione di Z_n deve essere approssimativamente $\mathcal{N}(0,1)$.

Esempio 4.14.

Si supponga di effettuare n prove Bernoulliane indipendenti, ciascuna con probabilità di successo p. La proporzione di successi osservata è data da $\overline{X}_n = n^{-1}\sum_{i=1}^n X_i$ ed è approssimativamente normale con $\mu = p$ e $\sigma =$

$\sqrt{p(1-p)/n}$; come osservato nell'esempio 4.12, per ogni p tale che $0 \leq p \leq 1$, si ha che $p(1-p) \leq 1/4$, da cui $\sigma \leq 1/2\sqrt{n}$. Ponendo $\sigma = 1/2\sqrt{n}$ si adotta un approccio "conservativo", in quanto si considera la più ampia varianza possibile σ^2, che a sua volta restituisce il più piccolo valore possibile (approssimato) per la probabilità che la variabile aleatoria assuma un valore in un intervallo fissato centrato in $\mu = p$. Si supponga, anche in questo caso, di voler determinare il più piccolo valore di n tale che sia verificata

$$P(|\overline{X}_n - p| < \delta) > \gamma.$$

Poiché si utilizza l'approssimazione normale (e quindi continua) fornita dal teorema del limite centrale (assumendo che il valore che verrà determinato si sufficientemente grande), possiamo affrontare il problema determinando n tale che

$$P(|\overline{X}_n - p| \leq \delta) = \gamma.$$

Si ha quindi

$$\begin{aligned} P(|\overline{X}_n - p| \leq \delta) &= P(-\delta \leq \overline{X}_n - p \leq \delta) \\ &= P\left(\frac{-\delta}{1/2\sqrt{n}} \leq \frac{\overline{X}_n - p}{1/2\sqrt{n}} \leq \frac{\delta}{1/2\sqrt{n}}\right) \\ &\doteq \Phi(2\delta\sqrt{n}) - \Phi(-2\delta\sqrt{n}) \\ &= 2\Phi(2\delta\sqrt{n}) - 1, \end{aligned}$$

e imponendo che tale probabilità sia uguale a γ risulta

$$\Phi(2\delta\sqrt{n}) = \frac{1}{2}(1+\gamma)$$

che è soddisfatta da $2\delta\sqrt{n} = z_{(1+\gamma)/2}$. Il valore richiesto è dato da

$$n = \frac{z^2_{(1+\gamma)/2}}{4\delta^2}.$$

Per $\delta = 0.01$, $\gamma = 0.9$, $z_{0.95} = 1.64$ infine si trova

$$n = \frac{1.64^2}{4(0.01)^2} = 6724,$$

un valore molto inferiore al valore $n = 25000$ ottenuto nell'esempio 4.12 applicando la disuguaglianza di Chebyshev.

4.6.1 Approssimazione della distribuzione chi-quadrato

Nella sezione 4.3 si è discusso il risultato per cui una variabile aleatoria χ_n^2 può essere vista come la somma dei quadrati di n variabili aleatorie normali standard indipendenti, che hanno quindi distribuzione χ_1^2 con media 1 e varianza 2. È facile verificare che se $X \sim \chi_n^2$, allora $\mu_X = n$ e $\sigma_X^2 = 2n$. Un'utile applicazione del teorema del limite centrale è quella che consente il calcolo approssimato, per n sufficientemente grande ($n > 30$), delle probabilità concernenti la variabile aleatoria $X \sim \chi_n^2$:

$$P(a \leq X \leq b) \doteq \frac{1}{\sqrt{2\pi}} \int_{(a-n)/\sqrt{2n}}^{(b-n)/\sqrt{2n}} e^{-y^2/2} dy = \Phi\left(\frac{b-n}{\sqrt{2n}}\right) - \Phi\left(\frac{a-n}{\sqrt{2n}}\right).$$

Ne segue che per determinare il quantile di ordine γ della distribuzione chi-quadrato con n (elevato) gradi di libertà, ovvero quel valore $\chi_{n,\gamma}^2$ tale che

$$P(X \leq \chi_{n,\gamma}^2) = \gamma,$$

si può considerare l'approssimazione

$$P(X \leq \chi_{n,\gamma}^2) \doteq P\left(Z \leq \frac{\chi_{n,\gamma}^2 - n}{\sqrt{2n}}\right),$$

per cui si ha

$$\frac{\chi_{n,\gamma}^2 - n}{\sqrt{2n}} \approx z_\gamma,$$

dove z_γ è il quantile di ordine γ della normale standardizzata; ne segue che

$$\chi_{n,\gamma}^2 \approx n + z_\gamma \sqrt{2n}.$$

4.6.2 Approssimazione della distribuzione binomiale

Abbiamo già visto che la distribuzione normale fu trovata come limite cui tende la binomiale al tendere del numero n delle prove all'infinito. A tale risultato si può pervenire applicando il teorema del limite centrale.

Si è detto che se $X_1, X_2, \ldots, X_n$ sono variabili aleatorie Bernoulliane indipendenti, ciascuna di parametro p, allora $Y = \sum_{i=1}^n X_i$ è binomiale di parametri n e p; inoltre, essendo $E(X_i) = p$ e $V(X_i) = p(1-p)$, $i = 1, \ldots, n$, si ha che $E(Y) = np$ e $V(Y) = np(1-p)$. In virtù del teorema del limite centrale, vale l'approssimazione (per n sufficientemente grande)

$$P(Y \leq y) \doteq \Phi\left(\frac{Y - np}{\sqrt{np(1-p)}}\right)$$

La velocità di convergenza, all'aumentare di n, della binomiale alla normale dipende dal valore di p: l'approssimazione è, infatti, tanto migliore quanto più è simmetrica la distribuzione binomiale (in pratica, l'approssimazione è da considerarsi adeguata se np ed $n(1-p)$ sono entrambi maggiori di 5).

Si osservi che la binomiale è una variabile aleatoria discreta, mentre l'approssimazione considerata è continua; per migliorare l'approssimazione si può considerare, in corrispondenza di ciascun valore $p(x)$ della binomiale, un rettangolo avente per base il segmento di estremi $x - 1/2$, $x + 1/2$, e altezza $p(x)$ (si veda la figura 4.1). Facendo corrispondere la probabilità (approssimata) che la variabile aleatoria assuma un valore minore o uguale a 4 all'area sottesa alla normale nell'intervallo $(-\infty, 4.5)$ (essendo 4.5 l'estremo destro dell'intervallo che rappresenta il numero 4) si adotta la cosiddetta *correzione per continuità*. Essa fornisce approssimazioni più accurate per n moderatamente elevato, mentre è trascurabile per n molto grande.

In generale, per Y binomiale di parametri n e p, l'approssimazione normale con la correzione per continuità è

$$P(Y \leq b) \doteq \Phi\left(\frac{b + 0.5 - np}{\sqrt{np(1-p)}}\right),$$

$$P(Y \geq a) = 1 - P(Y \leq a-1) \doteq 1 - \Phi\left(\frac{a - 0.5 - np}{\sqrt{np(1-p)}}\right),$$

$$\begin{aligned} P(a \leq Y \leq b) &= P(Y \leq b) - P(Y \leq a-1) \\ &\doteq \Phi\left(\frac{b + 0.5 - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{a - 0.5 - np}{\sqrt{np(1-p)}}\right). \end{aligned}$$

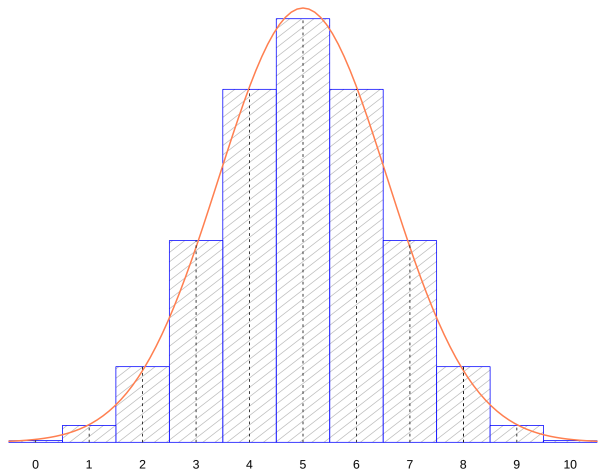


Figura 4.1: Correzione per continuità della distribuzione binomiale $Bin(10, 0.5)$

Esempio 4.15.

Sia $Y \sim Bin(n = 15, p = 0.6)$. La funzione di probabilità è quindi

$$p(y) = \binom{15}{y} 0.6^y 0.4^{15-y}, \quad y = 0, 1, \ldots, 15 \tag{4.33}$$

I valori delle probabilità sono rappresentati in figura 4.2, insieme con la densità normale di media $np = 9$ e varianza $np(1-p) = 3.6$. Il calcolo della probabilità esatta $P(6 \leq Y \leq 10)$ corrisponde a

$$P(6 \leq Y \leq 10) = P(Y = 6) + P(Y = 7) + P(Y = 8) + P(Y = 9) + P(Y = 10)$$

e dalla (4.33) si ottiene il valore 0.76; d'altra parte l'approssimazione fornisce la probabilità

$$\begin{aligned} P(6 \leq Y \leq 10) &= F_Y(10) - F_Y(5) \\ &\doteq \Phi\left(\frac{10-9}{\sqrt{3.6}}\right) - \Phi\left(\frac{5-9}{\sqrt{3.6}}\right) \\ &= \Phi(0.53) - \Phi(-2.11) \\ &= 0.7019 - 0.0174 = 0.6845. \end{aligned}$$

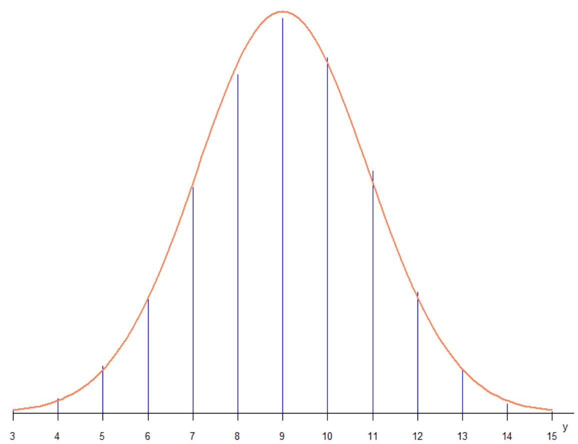


Figura 4.2: Funzione di probabilità binomiale $Bin(n = 15, p = 0.6)$ e curva della densità normale di parametri $\mu = np = 9$, $\sigma^2 = np(1-p) = 3.6$.

Applicando la correzione per continuità si ottiene

$$\begin{aligned} P(6 \leq Y \leq 10) &= F_Y(10) - F_Y(5) \\ &\doteq \Phi\left(\frac{10.5-9}{\sqrt{3.6}}\right) - \Phi\left(\frac{5.5-9}{\sqrt{3.6}}\right) \\ &= \Phi(0.79) - \Phi(-1.84) = 0.7523. \end{aligned}$$

Si conclude il capitolo trattando due distribuzioni di particolare rilievo nell'ambito dell'inferenza statistica, la distribuzione *t di Student* ed *F di Fisher*.

4.7 Le distribuzioni t di Student ed F di Fisher

Introduciamo la distribuzione t di Student mediante il seguente importante risultato.

Teorema 4.13. Sia Z una variabile aleatoria normale standard e sia V una variabile aleatoria indipendente con distribuzione χ^2 con d gradi di libertà. Allora la variabile aleatoria

$$T = \frac{Z}{\sqrt{V/d}}$$

ha funzione di densità

$$f(t) = \frac{\Gamma((d+1)/2)}{\sqrt{\pi d}\,\Gamma(d/2)} \frac{1}{(1+t^2/d)^{(d+1)/2}}, \quad -\infty < t < \infty \tag{4.34}$$

nota come distribuzione t di Student con d gradi di libertà, $T \sim t_d$.

La distribuzione t di Student è simmetrica, ha media $E(T) = 0$ per $d > 1$ e varianza $V(T) = d/(d-2)$ per $d > 2$. In figura 4.3 è rappresentata la distribuzione t di Student per $d = 2$, $d = 5$ e $d = 60$ gradi di libertà. Per $d \to \infty$, la (4.34) converge al valore della densità normale standard per ogni numero reale. Inoltre, come per la normale, alcuni quantili della distribuzione per diversi gradi di libertà ($d \leq 200$) sono riportati nella tavola in Appendice. La riga in corrispondenza di ∞ restituisce, in realtà, i valori dei quantili di una $\mathcal{N}(0,1)$. Per $d = 15$ si trovano, ad esempio, i quantili di ordine 0.90 e 0.99 che corrispondono ai valori $t_{15,0.90} = 1.3406$ e $t_{15,0.99} = 2.6025$. Se $T \sim t_{15}$ allora si avrà

$$P(T \leq 1.3406) = 0.9, \quad P(T \leq 2.6025) = 0.99.$$

Inoltre, per la simmetria rispetto a $t = 0$, si ha

$$P(T \leq -1.3406) = 0.1, \quad P(T \geq -2.6025) = 0.99.$$

Esempio 4.16.

Siano Z_1 e Z_2 variabili aleatorie indipendenti tali che $Z_1 \sim \mathcal{N}(0,1)$ e $Z_2 \sim \mathcal{N}(0,1)$. Definiamo $Y_1 = Z_1 - Z_2$ e $Y_2 = Z_1 + Z_2$. Allora Y_1 e Y_2 sono v.a. Gaussiane e inoltre

$$E(Y_1) = E(Y_2) = 0, \quad V(Y_1) = V(Y_2) = 2,$$

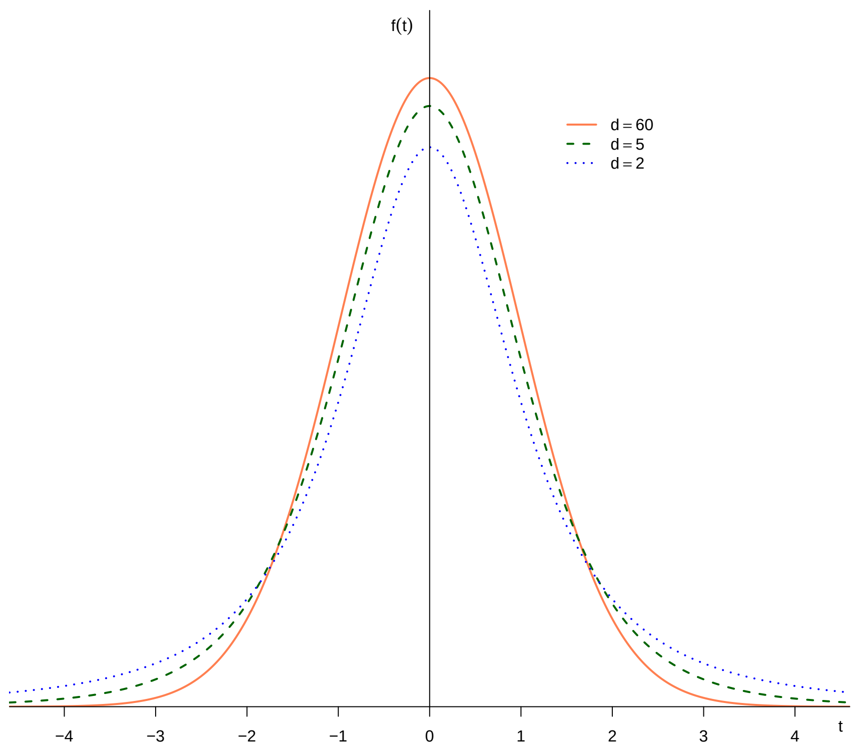


Figura 4.3: Densità t di Student per alcuni valori di d.

e, usando il Teorema 4.3

$$\mathrm{Cov}(Y_1, Y_2) = \mathrm{Cov}(Z_1 - Z_2, Z_1 + Z_2) = 0.$$

Poiché Y_1 e Y_2 sono normali con $\rho = 0$, se ne deduce che sono indipendenti. Inoltre le variabili aleatorie $Y_1/\sqrt{2}$ e $Y_2/\sqrt{2}$ sono normali standard e indipendenti. Pertanto

$$\left(\frac{Y_2}{\sqrt{2}}\right)^2 = \frac{Y_2^2}{2} \sim \chi_1^2$$

(e $Y_2^2/2$ indipendente da $Y_1/\sqrt{2}$). Allora, il rapporto

$$T = \frac{Y_1/\sqrt{2}}{\sqrt{Y_2^2/2}} = \frac{Y_1}{|Y_2|} = \frac{Z_1 - Z_2}{|Z_1 + Z_2|}$$

ha distribuzione t di Student con 1 grado di libertà.

Un'altra distribuzione che può essere derivata da variabili aleatorie normali è la distribuzione F.

Teorema 4.14. Siano U e V due variabili aleatorie indipendenti tali che $U \sim \chi^2_{d_1}$ e $V \sim \chi^2_{d_2}$. Allora la variabile aleatoria

$$Y = \frac{U/d_1}{V/d_2}$$

ha funzione di densità, definita per $y > 0$

$$f(y) = \frac{\Gamma((d_1 + d_2)/2)}{\Gamma(d_1/2)\Gamma(d_2/2)} \left(\frac{d_1}{d_2}\right)^{d_1/2} \left(1 + \frac{d_1 y}{d_2}\right)^{-(d_1+d_2)/2} y^{(d_1/2)-1} \quad (4.35)$$

nota come F di Fisher con d_1 e d_2 gradi di libertà, e si utilizza la notazione $Y \sim F_{d_1,d_2}$.

La distribuzione F di Fisher ha due parametri, d_1 e d_2, che corrispondono ai gradi di libertà delle variabili chi-quadrato presenti nel rapporto; si noti che è importante che i gradi di libertà vengano ordinati in modo tale che il primo sia

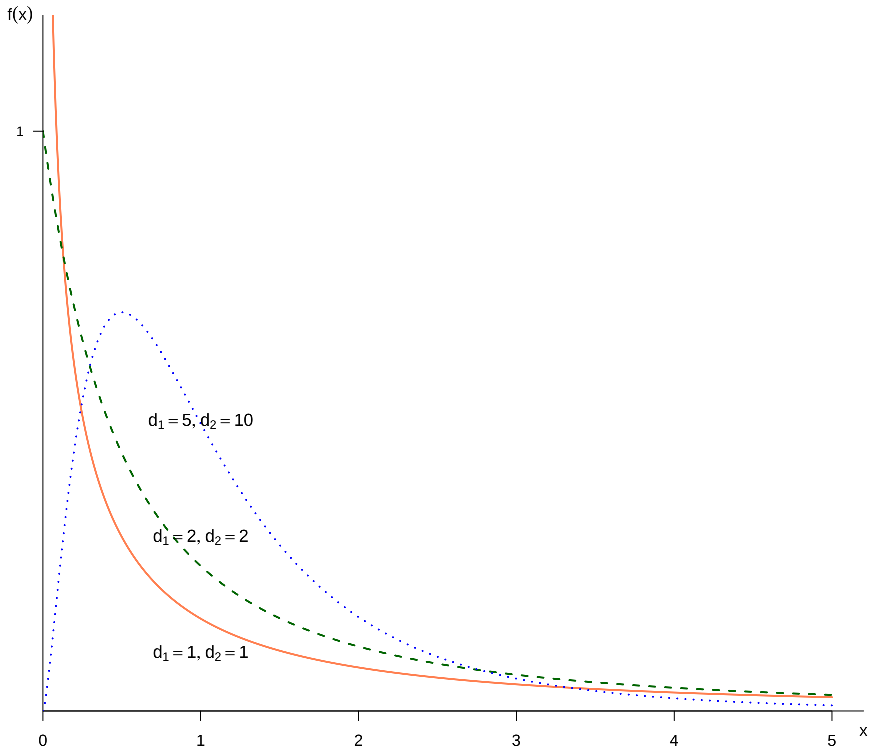


Figura 4.4: Densità F di Fisher per alcuni valori di d_1 e d_2.

quello che compare al numeratore. Il reciproco di F, $1/F = (V/d_2)/(U/d_1)$, è, infatti, ancora una variabile aleatoria con distribuzione F con d_2 e d_1 gradi di libertà.

L'andamento della distribuzione F di Fisher per alcuni valori dei parametri è illustrato in figura 4.4. La variabile aleatoria $Y \sim F_{d_1,d_2}$ ha media e varianza date da

$$E(Y) = \frac{d_2}{d_2 - 2}, \quad d_2 > 2$$

e

$$V(Y) = \frac{2d_2^2(d_1 + d_2 - 2)}{d_1(d_2 - 2)^2(d_2 - 4)}, \quad d_2 > 4.$$

Come per la distribuzione t, di questa distribuzione è particolarmente utile ricavare i quantili f_γ per un prefissato livello di γ, per tutte le coppie (d_1, d_2) (si vedano le tavole in Appendice, in cui sono riportati i valori per $\gamma = 0.9, 0.95, 0.99$). Per determinare il quantile di ordine $(1 - \gamma)$ si utilizza il fatto che

$$P\left(F_{d_2,d_1} < \frac{1}{f_{d_1,d_2;\gamma}}\right) = 1 - \gamma,$$

ovvero il quantile $(1 - \gamma)$ di F_{d_2,d_1} è $1/f_{d_1,d_2;\gamma}$.

In virtù del Teorema 4.13 sappiamo che il rapporto di una variabile aleatoria normale standard con la radice quadrata di una χ^2_d, indipendente dalla prima, divisa per i suoi gradi di libertà, ha distribuzione t di Student con d gradi di libertà. È facile osservare che tale variabile aleatoria al quadrato ha, dunque, distribuzione $F_{1,d}$ essendo il rapporto tra una normale standard al quadrato, cioè una χ^2_1 e una χ^2_d divisa per i suoi gradi di libertà. Pertanto, se $F = T^2$, dove $T \sim t_d$, allora $F \sim F_{1,d}$.

Esempio 4.17.

Siano Z_1, Z_2, Z_3, Z_4 e Z_5 variabili aleatorie normali standard indipendenti. Definiamo $Y_1 = Z_1^2 + Z_2^2$ e $Y_2 = Z_3^2 + Z_4^2 + Z_5^2$. Allora Y_1 e Y_2 sono indipendenti e distribuite, rispettivamente, secondo una χ^2 con 2 e 3 gradi di libertà. Il rapporto

$$F = \frac{Y_1/2}{Y_2/3}$$

ha distribuzione F con $d_1 = 2$ e $d_2 = 3$ gradi di libertà. Dalle tavole dei quantili della distribuzione F si trova

$$f_{2,3;0.95} = 9.552,$$

$$f_{2,3;0.05} = \frac{1}{f_{3,2;0.95}} = \frac{1}{19.16} = 0.052,$$

e quindi si deduce che

$$P(0.052 \leq F \leq 9.552) = 0.90;$$

ne segue che è pari a 0.9 la probabilità che il valore osservato di

$$F = \frac{(Z_1^2 + Z_2^2)/2}{(Z_3^2 + Z_4^2 + Z_5^2)/3}$$

si trovi nell'intervallo $(0.052, 9.552)$.

Bibliografia

[1] Francesco Caravenna, Paolo Dai Pra. Probabilità: Un'introduzione attraverso modelli e applicazioni, Springer Milan (2013)

Appendice A

Tavole

Nelle pagine successive compaiono sei tavole statistiche: descriviamo brevemente come si usano.

Tavola della funzione di ripartizione della distribuzione N(0,1). La tavola fornisce, per $0 \leq x \leq 4.49$ il valore della funzione

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt.$$

Ad esempio, $\Phi(0.63) = 0.73565$. Per valori di x superiori a 4.49, si pone $\Phi(x) = 1$; invece per valori di x negativi si ricorre alla relazione

$$\Phi(-x) = 1 - \Phi(x).$$

Per ottenere il quantile q_α della variabile $\mathcal{N}(0,1)$ (cioè il numero q_α tale che si abbia $P(X \leq q_\alpha) = \Phi(q_\alpha) = \alpha$, si usa la tavola della funzione di ripartizione al rovescio, cioè cercando per quale valore di x si ha $\Phi(x) = \alpha$: ad esempio $q_0.95 = 1.65$. In questo modo si possono ottenere solo i valori di q_α per $\alpha \geq 0.5$; per valori inferiori a $1/2$ si usa l'eguaglianza

$$q_\alpha = -q_{1-\alpha}.$$

Tavola dei quantili della variabile di Student t_n. Questa tavola permette di trovare direttamente, in funzione dei gradi di libertà, n, e del numero α con $0 < \alpha < 1$, il valore del quantile $t_{n,\alpha}$. Ad esempio, il quantile 0.975 della distribuzione t di Student con 5 gradi di libertà è l'elemento che si trova in corrispondenza di $n = 5$, nella colonna con intestazione 0.975: $t_{5,.975} = 2.5706$. Poiché anche la densità di Student è una funzione pari, vale l'eguaglianza

$$t_{n,\alpha} = -t_{n,1-\alpha}.$$

Tavola dei quantili della distribuzione χ^2_n. Questa tavola riporta i quantili di ordine $0 < \alpha < 1$ della distribuzione chi-quadrato al variare dei gradi di libertà n. Ad esempio, il quantile 0.1 della distribuzione χ^2 con 7 gradi di libertà è l'elemento che si trova in corrispondenza di $n = 7$, nella colonna con intestazione 0.1: $\chi^2_{7,0.1} = 2.8331$. La densità χ^2 non è una funzione pari, di conseguenza non si possono ricavare i quantili con $\alpha \leq 1/2$ da quelli con $\alpha \geq 1/2$ per questo motivo tale tavola riporta i quantili $\chi^2_{n,\alpha}$ per α vicino a 1 e vicino a 0.

Tavole dei quantili della distribuzione $F_{m,n}$. Le tavole degli quantili della legge di Fisher riportano, in funzione dei gradi di libertà m ed n e limitatamente ad $\alpha = 0.90, 0.95, 0.99$, i valori dei quantili della legge F di Fisher: per ogni coppia (m, n) la tavola dà la quantità f_α che soddisfa

$$P(F \leq f_\alpha) = \alpha$$

Ad esempio, il quantile di ordine 0.95 della distribuzione F di Fisher con $m = 7$ e $n = 3$ gradi di libertà è $f_{0.95} = 8.887$. Inoltre vale la relazione

$$f_{m,n;1-\alpha} = \frac{1}{f_{n,m;\alpha}}$$

cioè, ad esempio, il quantile di ordine 0.05 della distribuzione F di Fisher con $m = 3$ e $n = 7$ gradi di libertà si ottiene come $f_{3,7;0.05} = 1/f_{7,3;0.95} = 1/8.887 = 0.1125$.

Tavola della funzione di ripartizione della distribuzione N(0,1)										
z	**0.00**	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**
0.0	0.50000	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.52790	0.53188	0.53586
0.1	0.53983	0.54380	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.62930	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.65910	0.66276	0.66640	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.70540	0.70884	0.71226	0.71566	0.71904	0.72240
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.75490
0.7	0.75804	0.76115	0.76424	0.76730	0.77035	0.77337	0.77637	0.77935	0.78230	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.0	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.86650	0.86864	0.87076	0.87286	0.87493	0.87698	0.87900	0.88100	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.3	0.90320	0.90490	0.90658	0.90824	0.90988	0.91149	0.91308	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.92220	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.94520	0.94630	0.94738	0.94845	0.94950	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.96080	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.97320	0.97381	0.97441	0.97500	0.97558	0.97615	0.97670
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.98030	0.98077	0.98124	0.98169
2.1	0.98214	0.98257	0.98300	0.98341	0.98382	0.98422	0.98461	0.98500	0.98537	0.98574
2.2	0.98610	0.98645	0.98679	0.98713	0.98745	0.98778	0.98809	0.98840	0.98870	0.98899
2.3	0.98928	0.98956	0.98983	0.99010	0.99036	0.99061	0.99086	0.99111	0.99134	0.99158
2.4	0.99180	0.99202	0.99224	0.99245	0.99266	0.99286	0.99305	0.99324	0.99343	0.99361
2.5	0.99379	0.99396	0.99413	0.99430	0.99446	0.99461	0.99477	0.99492	0.99506	0.99520
2.6	0.99534	0.99547	0.99560	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632	0.99643
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.99720	0.99728	0.99736
2.8	0.99744	0.99752	0.99760	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801	0.99807
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861
3.0	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99896	0.99900
3.1	0.99903	0.99906	0.99910	0.99913	0.99916	0.99918	0.99921	0.99924	0.99926	0.99929
3.2	0.99931	0.99934	0.99936	0.99938	0.99940	0.99942	0.99944	0.99946	0.99948	0.99950
3.3	0.99952	0.99953	0.99955	0.99957	0.99958	0.99960	0.99961	0.99962	0.99964	0.99965
3.4	0.99966	0.99968	0.99969	0.99970	0.99971	0.99972	0.99973	0.99974	0.99975	0.99976
3.5	0.99977	0.99978	0.99978	0.99979	0.99980	0.99981	0.99981	0.99982	0.99983	0.99983
3.6	0.99984	0.99985	0.99985	0.99986	0.99986	0.99987	0.99987	0.99988	0.99988	0.99989
3.7	0.99989	0.99990	0.99990	0.99990	0.99991	0.99991	0.99992	0.99992	0.99992	0.99992
3.8	0.99993	0.99993	0.99993	0.99994	0.99994	0.99994	0.99994	0.99995	0.99995	0.99995
3.9	0.99995	0.99995	0.99996	0.99996	0.99996	0.99996	0.99996	0.99996	0.99997	0.99997
4.0	0.99997	0.99997	0.99997	0.99997	0.99997	0.99997	0.99998	0.99998	0.99998	0.99998
4.1	0.99998	0.99998	0.99998	0.99998	0.99998	0.99998	0.99998	0.99998	0.99999	0.99999
4.2	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999
4.3	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999
4.4	0.99999	0.99999	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000

Tavola dei quantili della distribuzione T(n)

	Valore della funzione di ripartizione									
n	**0.75**	**0.8**	**0.85**	**0.9**	**0.95**	**0.975**	**0.99**	**0.995**	**0.999**	**0.9995**
1	1.0000	1.3764	1.9626	3.0777	6.3137	12.7062	31.8210	63.6559	318.2888	636.5776
2	0.8165	1.0607	1.3862	1.8856	2.9200	4.3027	6.9645	9.9250	22.3285	31.5998
3	0.7649	0.9785	1.2498	1.6377	2.3534	3.1824	4.5407	5.8408	10.2143	12.9244
4	0.7407	0.9410	1.1896	1.5332	2.1318	2.7765	3.7469	4.6041	7.1729	8.6101
5	0.7267	0.9195	1.1558	1.4759	2.0150	2.5706	3.3649	4.0321	5.8935	6.8685
6	0.7176	0.9057	1.1342	1.4398	1.9432	2.4469	3.1427	3.7074	5.2075	5.9587
7	0.7111	0.8960	1.1192	1.4149	1.8946	2.3646	2.9979	3.4995	4.7853	5.4081
8	0.7064	0.8889	1.1081	1.3968	1.8595	2.3060	2.8965	3.3554	4.5008	5.0414
9	0.7027	0.8834	1.0997	1.3830	1.8331	2.2622	2.8214	3.2498	4.2969	4.7809
10	0.6998	0.8791	1.0931	1.3722	1.8125	2.2281	2.7638	3.1693	4.1437	4.5868
11	0.6974	0.8755	1.0877	1.3634	1.7959	2.2010	2.7181	3.1058	4.0248	4.4369
12	0.6955	0.8726	1.0832	1.3562	1.7823	2.1788	2.6810	3.0545	3.9296	4.3178
13	0.6938	0.8702	1.0795	1.3502	1.7709	2.1604	2.6503	3.0123	3.8520	4.2209
14	0.6924	0.8681	1.0763	1.3450	1.7613	2.1448	2.6245	2.9768	3.7874	4.1403
15	0.6912	0.8662	1.0735	1.3406	1.7531	2.1315	2.6025	2.9467	3.7329	4.0728
16	0.6901	0.8647	1.0711	1.3368	1.7459	2.1199	2.5835	2.9208	3.6861	4.0149
17	0.6892	0.8633	1.0690	1.3334	1.7396	2.1098	2.5669	2.8982	3.6458	3.9651
18	0.6884	0.8620	1.0672	1.3304	1.7341	2.1009	2.5524	2.8784	3.6105	3.9217
19	0.6876	0.8610	1.0655	1.3277	1.7291	2.0930	2.5395	2.8609	3.5793	3.8833
20	0.6870	0.8600	1.0640	1.3253	1.7247	2.0860	2.5280	2.8453	3.5518	3.8496
21	0.6864	0.8591	1.0627	1.3232	1.7207	2.0796	2.5176	2.8314	3.5271	3.8193
22	0.6858	0.8583	1.0614	1.3212	1.7171	2.0739	2.5083	2.8188	3.5050	3.7922
23	0.6853	0.8575	1.0603	1.3195	1.7139	2.0687	2.4999	2.8073	3.4850	3.7676
24	0.6848	0.8569	1.0593	1.3178	1.7109	2.0639	2.4922	2.7970	3.4668	3.7454
25	0.6844	0.8562	1.0584	1.3163	1.7081	2.0595	2.4851	2.7874	3.4502	3.7251
26	0.6840	0.8557	1.0575	1.3150	1.7056	2.0555	2.4786	2.7787	3.4350	3.7067
27	0.6837	0.8551	1.0567	1.3137	1.7033	2.0518	2.4727	2.7707	3.4210	3.6895
28	0.6834	0.8546	1.0560	1.3125	1.7011	2.0484	2.4671	2.7633	3.4082	3.6739
29	0.6830	0.8542	1.0553	1.3114	1.6991	2.0452	2.4620	2.7564	3.3963	3.6595
30	0.6828	0.8538	1.0547	1.3104	1.6973	2.0423	2.4573	2.7500	3.3852	3.6460
40	0.6807	0.8507	1.0500	1.3031	1.6839	2.0211	2.4233	2.7045	3.3069	3.5510
50	0.6794	0.8489	1.0473	1.2987	1.6759	2.0086	2.4033	2.6778	3.2614	3.4960
60	0.6786	0.8477	1.0455	1.2958	1.6706	2.0003	2.3901	2.6603	3.2317	3.4602
70	0.6780	0.8468	1.0442	1.2938	1.6669	1.9944	2.3808	2.6479	3.2108	3.4350
80	0.6776	0.8461	1.0432	1.2922	1.6641	1.9901	2.3739	2.6387	3.1952	3.4164
90	0.6772	0.8456	1.0424	1.2910	1.6620	1.9867	2.3685	2.6316	3.1832	3.4019
100	0.6770	0.8452	1.0418	1.2901	1.6602	1.9840	2.3642	2.6259	3.1738	3.3905
120	0.6765	0.8446	1.0409	1.2886	1.6576	1.9799	2.3578	2.6174	3.1595	3.3734
140	0.6762	0.8442	1.0403	1.2876	1.6558	1.9771	2.3533	2.6114	3.1495	3.3613
200	0.6757	0.8434	1.0391	1.2858	1.6525	1.9719	2.3451	2.6006	3.1315	3.3398
∞	0.6745	0.8416	1.0364	1.2816	1.6449	1.9600	2.3263	2.5758	3.0902	3.2905

Tavola dei quantili della distribuzione $\chi^2(n)$

	Valore della funzione di ripartizione																				
n	**0.0005**	**0.001**	**0.005**	**0.01**	**0.025**	**0.05**	**0.1**	**0.15**	**0.2**	**0.25**	**0.5**	**0.75**	**0.8**	**0.85**	**0.9**	**0.95**	**0.975**	**0.99**	**0.995**	**0.999**	**0.9995**
1	3.929E-07	1.570E-06	3.927E-05	1.571E-04	9.821E-04	3.932E-03	0.0158	0.0358	0.0642	0.1015	0.4549	1.3233	1.6424	2.0722	2.7055	3.8415	5.0239	6.6349	7.8794	10.8274	12.1153
2	9.997E-04	2.001E-03	0.0100	0.0201	0.0506	0.1026	0.2107	0.3250	0.4463	0.5754	1.3863	2.7726	3.2189	3.7942	4.6052	5.9915	7.3778	9.2104	10.5965	13.8150	15.2014
3	0.0153	0.0243	0.0717	0.1148	0.2158	0.3518	0.5844	0.7978	1.0052	1.2125	2.3660	4.1083	4.6416	5.3170	6.2514	7.8147	9.3484	11.3449	12.8381	16.2660	17.7311
4	0.0639	0.0908	0.2070	0.2971	0.4844	0.7107	1.0636	1.3665	1.6488	1.9226	3.3567	5.3853	5.9886	6.7449	7.7794	9.4877	11.1433	13.2767	14.8602	18.4662	19.9977
5	0.1581	0.2102	0.4118	0.5543	0.8312	1.1455	1.6103	1.9938	2.3425	2.6746	4.3515	6.6257	7.2893	8.1152	9.2363	11.0705	12.8325	15.0863	16.7496	20.5147	22.1057
6	0.2994	0.3810	0.6757	0.8721	1.2373	1.6354	2.2041	2.6613	3.0701	3.4546	5.3481	7.8408	8.5581	9.4461	10.6446	12.5916	14.4494	16.8119	18.5475	22.4575	24.1016
7	0.4849	0.5985	0.9893	1.2390	1.6899	2.1673	2.8331	3.3583	3.8223	4.2549	6.3458	9.0371	9.8032	10.7479	12.0170	14.0671	16.0128	18.4753	20.2777	24.3213	26.0179
8	0.7104	0.8571	1.3444	1.6465	2.1797	2.7326	3.4895	4.0782	4.5936	5.0706	7.3441	10.2189	11.0301	12.0271	13.3616	15.5073	17.5345	20.0902	21.9549	26.1239	27.8674
9	0.9718	1.1519	1.7349	2.0879	2.7004	3.3251	4.1682	4.8165	5.3801	5.8988	8.3428	11.3887	12.2421	13.2880	14.6837	16.9190	19.0228	21.6660	23.5893	27.8767	29.6669
10	1.2651	1.4787	2.1558	2.5582	3.2470	3.9403	4.8652	5.5701	6.1791	6.7372	9.3418	12.5489	13.4420	14.5339	15.9872	18.3070	20.4832	23.2093	25.1881	29.5879	31.4195
11	1.5870	1.8338	2.6032	3.0535	3.8157	4.5748	5.5778	6.3364	6.9887	7.5841	10.3410	13.7007	14.6314	15.7671	17.2750	19.6752	21.9200	24.7250	26.7569	31.2635	33.1382
12	1.9345	2.2141	3.0738	3.5706	4.4038	5.2260	6.3038	7.1138	7.8073	8.4384	11.3403	14.8454	15.8120	16.9893	18.5493	21.0261	23.3367	26.2170	28.2997	32.9092	34.8211
13	2.3049	2.6172	3.5650	4.1069	5.0087	5.8919	7.0415	7.9008	8.6339	9.2991	12.3398	15.9839	16.9848	18.2020	19.8119	22.3620	24.7356	27.6882	29.8193	34.5274	36.4768
14	2.6966	3.0407	4.0747	4.6604	5.6287	6.5706	7.7895	8.6963	9.4673	10.1653	13.3393	17.1169	18.1508	19.4062	21.0641	23.6848	26.1189	29.1412	31.3194	36.1239	38.1085
15	3.1073	3.4825	4.6009	5.2294	6.2621	7.2609	8.5468	9.4993	10.3070	11.0365	14.3389	18.2451	19.3107	20.6030	22.3071	24.9958	27.4884	30.5780	32.8015	37.6978	39.7173
16	3.5357	3.9417	5.1422	5.8122	6.9077	7.9616	9.3122	10.3090	11.1521	11.9122	15.3385	19.3689	20.4651	21.7931	23.5418	26.2962	28.8453	31.9999	34.2671	39.2518	41.3077
17	3.9800	4.4162	5.6973	6.4077	7.5642	8.6718	10.0852	11.1249	12.0023	12.7919	16.3382	20.4887	21.6146	22.9770	24.7690	27.5871	30.1910	33.4087	35.7184	40.7911	42.8808
18	4.4391	4.9048	6.2648	7.0149	8.2307	9.3904	10.8649	11.9462	12.8570	13.6753	17.3379	21.6049	22.7595	24.1555	25.9894	28.8693	31.5264	34.8052	37.1564	42.3119	44.4337
19	4.9125	5.4067	6.8439	7.6327	8.9065	10.1170	11.6509	12.7727	13.7158	14.5620	18.3376	22.7178	23.9004	25.3289	27.2036	30.1435	32.8523	36.1908	38.5821	43.8194	45.9738
20	5.3978	5.9210	7.4338	8.2604	9.5908	10.8508	12.4426	13.6039	14.5784	15.4518	19.3374	23.8277	25.0375	26.4976	28.4120	31.4104	34.1696	37.5663	39.9969	45.3142	47.4977
21	5.8954	6.4467	8.0336	8.8972	10.2829	11.5913	13.2396	14.4393	15.4446	16.3444	20.3372	24.9348	26.1711	27.6620	29.6151	32.6706	35.4789	38.9322	41.4009	46.7963	49.0096
22	6.4041	6.9829	8.6427	9.5425	10.9823	12.3380	14.0415	15.2787	16.3140	17.2396	21.3370	26.0393	27.3015	28.8224	30.8133	33.9245	36.7807	40.2894	42.7957	48.2676	50.5105
23	6.9240	7.5291	9.2604	10.1957	11.6885	13.0905	14.8480	16.1219	17.1865	18.1373	22.3369	27.1413	28.4288	29.9792	32.0069	35.1725	38.0756	41.6383	44.1814	49.7276	51.9995
24	7.4528	8.0847	9.8862	10.8563	12.4011	13.8484	15.6587	16.9686	18.0618	19.0373	23.3367	28.2412	29.5533	31.1325	33.1962	36.4150	39.3641	42.9798	45.5584	51.1790	53.4776
25	7.9905	8.6494	10.5196	11.5240	13.1197	14.6114	16.4734	17.8184	18.9397	19.9393	24.3366	29.3388	30.6752	32.2825	34.3816	37.6525	40.6465	44.3140	46.9280	52.6187	54.9475
26	8.5374	9.2222	11.1602	12.1982	13.8439	15.3792	17.2919	18.6714	19.8202	20.8434	25.3365	30.4346	31.7946	33.4295	35.5632	38.8851	41.9231	45.6416	48.2898	54.0511	56.4068
27	9.0929	9.8029	11.8077	12.8785	14.5734	16.1514	18.1139	19.5272	20.7030	21.7494	26.3363	31.5284	32.9117	34.5736	36.7412	40.1133	43.1945	46.9628	49.6450	55.4751	57.8556
28	9.6558	10.3907	12.4613	13.5647	15.3079	16.9279	18.9392	20.3857	21.5880	22.6572	27.3362	32.6205	34.0266	35.7150	37.9159	41.3372	44.4608	48.2782	50.9936	56.8918	59.2990
29	10.2266	10.9861	13.1211	14.2564	16.0471	17.7084	19.7677	21.2468	22.4751	23.5666	28.3361	33.7109	35.1394	36.8538	39.0875	42.5569	45.7223	49.5878	52.3355	58.3006	60.7342
30	10.8040	11.5876	13.7867	14.9535	16.7908	18.4927	20.5992	22.1103	23.3641	24.4776	29.3360	34.7997	36.2502	37.9902	40.2560	43.7730	46.9792	50.8922	53.6719	59.7022	62.1600
40	16.9058	17.9166	20.7066	22.1642	24.4331	26.5093	29.0505	30.8563	32.3449	33.6603	39.3353	45.6160	47.2685	49.2438	51.8050	55.7585	59.3417	63.6908	66.7660	73.4029	76.0963
50	23.4611	24.6736	27.9908	29.7067	32.3574	34.7642	37.6886	39.7539	41.4492	42.9421	49.3349	56.3336	58.1638	60.3460	63.1671	67.5048	71.4202	76.1538	79.4898	86.6603	89.5597
60	30.3393	31.7381	35.5344	37.4848	40.4817	43.1880	46.4589	48.7587	50.6406	52.2938	59.3347	66.9815	68.9721	71.3411	74.3970	79.0820	83.2977	88.3794	91.9518	99.6078	102.6971
70	37.4671	39.0358	43.2753	45.4417	48.7575	51.7393	55.3289	57.8443	59.8978	61.6983	69.3345	77.5766	79.7147	82.2553	85.5270	90.5313	95.0231	100.4251	104.2148	112.3167	115.5766
80	44.7917	46.5197	51.1719	53.5400	57.1532	60.3915	64.2778	66.9938	69.2070	71.1445	79.3343	88.1303	90.4053	93.1058	96.5782	101.8795	106.6285	112.3288	116.3209	124.8389	128.2636
90	52.2768	54.1559	59.1963	61.7540	65.6466	69.1260	73.2911	76.1954	78.5584	80.6247	89.3342	98.6499	101.0537	103.9040	107.5650	113.1452	118.1359	124.1162	128.2987	137.2082	140.7804
100	59.8946	61.9182	67.3275	70.0650	74.2219	77.9294	82.3581	85.4406	87.9453	90.1332	99.3341	109.1412	111.6667	114.6588	118.4980	124.3421	129.5613	135.8069	140.1697	149.4488	153.1638
120	75.4654	77.7555	83.8517	86.9233	91.5726	95.7046	100.6236	104.0374	106.8056	109.2197	119.3340	130.0546	132.8063	136.0620	140.2326	146.5673	152.2113	158.9500	163.6485	173.6184	177.6006
140	91.3894	93.9253	100.6547	104.0343	109.1368	113.6594	119.0293	122.7476	125.7580	128.3800	139.3339	150.8941	153.8537	157.3517	161.8270	168.6130	174.6478	181.8405	186.8465	197.4498	201.6804
200	140.6591	143.8420	152.2408	156.4321	162.7280	168.2785	174.8353	179.3550	183.0028	186.1717	199.3337	213.1022	216.6088	220.7441	226.0210	233.9942	241.0578	249.4452	255.2638	267.5388	272.4220

Tavola dei quantili 0.9 della distribuzione F(m,n)

																		n																	
m	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	22	24	26	28	30	40	50	60	70	80	90	100	120	140	200	∞
1	39.86	8.526	5.538	4.545	4.060	3.776	3.589	3.458	3.360	3.285	3.225	3.177	3.136	3.102	3.073	3.048	3.026	3.007	2.990	2.975	2.949	2.927	2.909	2.894	2.881	2.835	2.809	2.791	2.779	2.769	2.762	2.756	2.748	2.742	2.731	2.706
2	49.50	9.000	5.462	4.325	3.780	3.463	3.257	3.113	3.006	2.924	2.860	2.807	2.763	2.726	2.695	2.668	2.645	2.624	2.606	2.589	2.561	2.538	2.519	2.503	2.489	2.440	2.412	2.393	2.380	2.370	2.363	2.356	2.347	2.341	2.329	2.303
3	53.59	9.162	5.391	4.191	3.619	3.289	3.074	2.924	2.813	2.728	2.660	2.606	2.560	2.522	2.490	2.462	2.437	2.416	2.397	2.380	2.351	2.327	2.307	2.291	2.276	2.226	2.197	2.177	2.164	2.154	2.146	2.139	2.130	2.123	2.111	2.084
4	55.83	9.243	5.343	4.107	3.520	3.181	2.961	2.806	2.693	2.605	2.536	2.480	2.434	2.395	2.361	2.333	2.308	2.286	2.266	2.249	2.219	2.195	2.174	2.157	2.142	2.091	2.061	2.041	2.027	2.016	2.008	2.002	1.992	1.985	1.973	1.945
5	57.24	9.293	5.309	4.051	3.453	3.108	2.883	2.726	2.611	2.522	2.451	2.394	2.347	2.307	2.273	2.244	2.218	2.196	2.176	2.158	2.128	2.103	2.082	2.064	2.049	1.997	1.966	1.946	1.931	1.921	1.912	1.906	1.896	1.889	1.876	1.847
6	58.20	9.326	5.285	4.010	3.405	3.055	2.827	2.668	2.551	2.461	2.389	2.331	2.283	2.243	2.208	2.178	2.152	2.130	2.109	2.091	2.060	2.035	2.014	1.996	1.980	1.927	1.895	1.875	1.860	1.849	1.841	1.834	1.824	1.817	1.804	1.774
7	58.91	9.349	5.266	3.979	3.368	3.014	2.785	2.624	2.505	2.414	2.342	2.283	2.234	2.193	2.158	2.128	2.102	2.079	2.058	2.040	2.008	1.983	1.961	1.943	1.927	1.873	1.840	1.819	1.804	1.793	1.785	1.778	1.767	1.760	1.747	1.717
8	59.44	9.367	5.252	3.955	3.339	2.983	2.752	2.589	2.469	2.377	2.304	2.245	2.195	2.154	2.119	2.088	2.061	2.038	2.017	1.999	1.967	1.941	1.919	1.900	1.884	1.829	1.796	1.775	1.760	1.748	1.739	1.732	1.722	1.714	1.701	1.670
9	59.86	9.381	5.240	3.936	3.316	2.958	2.725	2.561	2.440	2.347	2.274	2.214	2.164	2.122	2.086	2.055	2.028	2.005	1.984	1.965	1.933	1.906	1.884	1.865	1.849	1.793	1.760	1.738	1.723	1.711	1.702	1.695	1.684	1.677	1.663	1.632
10	60.19	9.392	5.230	3.920	3.297	2.937	2.703	2.538	2.416	2.323	2.248	2.188	2.138	2.095	2.059	2.028	2.001	1.977	1.956	1.937	1.904	1.877	1.855	1.836	1.819	1.763	1.729	1.707	1.691	1.680	1.670	1.663	1.652	1.645	1.631	1.599
11	60.47	9.401	5.222	3.907	3.282	2.920	2.684	2.519	2.396	2.302	2.227	2.166	2.116	2.073	2.037	2.005	1.978	1.954	1.932	1.913	1.880	1.853	1.830	1.811	1.794	1.737	1.703	1.680	1.665	1.653	1.643	1.636	1.625	1.617	1.603	1.570
12	60.71	9.408	5.216	3.896	3.268	2.905	2.668	2.502	2.379	2.284	2.209	2.147	2.097	2.054	2.017	1.985	1.958	1.933	1.912	1.892	1.859	1.832	1.809	1.790	1.773	1.715	1.680	1.657	1.641	1.629	1.620	1.612	1.601	1.593	1.579	1.546
13	60.90	9.415	5.210	3.886	3.257	2.892	2.654	2.488	2.364	2.269	2.193	2.131	2.080	2.037	2.000	1.968	1.940	1.916	1.894	1.875	1.841	1.814	1.790	1.771	1.754	1.695	1.660	1.637	1.621	1.609	1.599	1.592	1.580	1.572	1.558	1.524
14	61.07	9.420	5.205	3.878	3.247	2.881	2.643	2.475	2.351	2.255	2.179	2.117	2.066	2.022	1.985	1.953	1.925	1.900	1.878	1.859	1.825	1.797	1.774	1.754	1.737	1.678	1.643	1.619	1.603	1.590	1.581	1.573	1.562	1.553	1.539	1.505
15	61.22	9.425	5.200	3.870	3.238	2.871	2.632	2.464	2.340	2.244	2.167	2.105	2.053	2.010	1.972	1.940	1.912	1.887	1.865	1.845	1.811	1.783	1.760	1.740	1.722	1.662	1.627	1.603	1.587	1.574	1.564	1.557	1.545	1.537	1.522	1.487
16	61.35	9.429	5.196	3.864	3.230	2.863	2.623	2.454	2.330	2.233	2.156	2.094	2.042	1.998	1.961	1.928	1.900	1.875	1.852	1.833	1.798	1.770	1.747	1.726	1.709	1.649	1.613	1.589	1.572	1.559	1.550	1.542	1.530	1.522	1.507	1.471
17	61.46	9.433	5.193	3.858	3.223	2.855	2.615	2.446	2.320	2.224	2.147	2.084	2.032	1.988	1.950	1.917	1.889	1.864	1.841	1.821	1.787	1.759	1.735	1.715	1.697	1.636	1.600	1.576	1.559	1.546	1.536	1.528	1.516	1.508	1.493	1.457
18	61.57	9.436	5.190	3.853	3.217	2.848	2.607	2.438	2.312	2.215	2.138	2.075	2.023	1.978	1.941	1.908	1.879	1.854	1.831	1.811	1.777	1.748	1.724	1.704	1.686	1.625	1.588	1.564	1.547	1.534	1.524	1.516	1.504	1.495	1.480	1.444
19	61.66	9.439	5.187	3.848	3.212	2.842	2.601	2.431	2.305	2.208	2.130	2.067	2.014	1.970	1.932	1.899	1.870	1.845	1.822	1.802	1.768	1.739	1.715	1.694	1.676	1.615	1.578	1.553	1.536	1.523	1.513	1.505	1.493	1.484	1.468	1.432
20	61.74	9.441	5.184	3.844	3.207	2.836	2.595	2.425	2.298	2.201	2.123	2.060	2.007	1.962	1.924	1.891	1.862	1.837	1.814	1.794	1.759	1.730	1.706	1.685	1.667	1.605	1.568	1.543	1.526	1.513	1.503	1.494	1.482	1.473	1.458	1.421
21	61.81	9.444	5.182	3.841	3.202	2.831	2.589	2.419	2.292	2.194	2.117	2.053	2.000	1.955	1.917	1.884	1.855	1.829	1.807	1.786	1.751	1.722	1.698	1.677	1.659	1.596	1.559	1.534	1.517	1.503	1.493	1.485	1.472	1.464	1.448	1.410
22	61.88	9.446	5.180	3.837	3.198	2.827	2.584	2.414	2.287	2.189	2.111	2.047	1.994	1.949	1.911	1.877	1.848	1.823	1.800	1.779	1.744	1.715	1.690	1.669	1.651	1.588	1.551	1.526	1.508	1.495	1.484	1.476	1.463	1.454	1.438	1.401
23	61.94	9.448	5.178	3.834	3.194	2.822	2.580	2.409	2.282	2.183	2.105	2.041	1.988	1.943	1.905	1.871	1.842	1.816	1.793	1.773	1.737	1.708	1.683	1.662	1.644	1.581	1.543	1.518	1.500	1.487	1.476	1.468	1.455	1.446	1.430	1.392
24	62.00	9.450	5.176	3.831	3.191	2.818	2.575	2.404	2.277	2.178	2.100	2.036	1.983	1.938	1.899	1.866	1.836	1.810	1.787	1.767	1.731	1.702	1.677	1.656	1.638	1.574	1.536	1.511	1.493	1.479	1.468	1.460	1.447	1.438	1.422	1.383
25	62.05	9.451	5.175	3.828	3.187	2.815	2.571	2.400	2.272	2.174	2.095	2.031	1.978	1.933	1.894	1.860	1.831	1.805	1.782	1.761	1.726	1.696	1.671	1.650	1.632	1.568	1.529	1.504	1.486	1.472	1.461	1.453	1.440	1.431	1.414	1.375
26	62.10	9.453	5.173	3.826	3.184	2.811	2.568	2.396	2.268	2.170	2.091	2.027	1.973	1.928	1.889	1.855	1.826	1.800	1.777	1.756	1.720	1.691	1.666	1.644	1.626	1.562	1.523	1.498	1.479	1.465	1.455	1.446	1.433	1.424	1.407	1.368
27	62.15	9.454	5.172	3.823	3.181	2.808	2.564	2.392	2.265	2.166	2.087	2.022	1.969	1.923	1.885	1.851	1.821	1.795	1.772	1.751	1.715	1.686	1.660	1.639	1.621	1.556	1.517	1.492	1.473	1.459	1.448	1.440	1.427	1.417	1.400	1.361
28	62.19	9.456	5.170	3.821	3.179	2.805	2.561	2.389	2.261	2.162	2.083	2.019	1.965	1.919	1.880	1.847	1.817	1.791	1.767	1.746	1.711	1.681	1.656	1.634	1.616	1.551	1.512	1.486	1.467	1.453	1.442	1.434	1.421	1.411	1.394	1.354
29	62.23	9.457	5.169	3.819	3.176	2.803	2.558	2.386	2.258	2.159	2.080	2.015	1.961	1.916	1.876	1.843	1.813	1.787	1.763	1.742	1.706	1.676	1.651	1.630	1.611	1.546	1.507	1.481	1.462	1.448	1.437	1.428	1.415	1.405	1.388	1.348
30	62.26	9.458	5.168	3.817	3.174	2.800	2.555	2.383	2.255	2.155	2.076	2.011	1.958	1.912	1.873	1.839	1.809	1.783	1.759	1.738	1.702	1.672	1.647	1.625	1.606	1.541	1.502	1.476	1.457	1.443	1.432	1.423	1.409	1.400	1.383	1.342
40	62.53	9.466	5.160	3.804	3.157	2.781	2.535	2.361	2.232	2.132	2.052	1.986	1.931	1.885	1.845	1.811	1.781	1.754	1.730	1.708	1.671	1.641	1.615	1.592	1.573	1.506	1.465	1.437	1.418	1.403	1.391	1.382	1.368	1.357	1.339	1.295
50	62.69	9.471	5.155	3.795	3.147	2.770	2.523	2.348	2.218	2.117	2.036	1.970	1.915	1.869	1.828	1.793	1.763	1.736	1.711	1.690	1.652	1.621	1.594	1.572	1.552	1.483	1.441	1.413	1.392	1.377	1.365	1.355	1.340	1.329	1.310	1.263
60	62.79	9.475	5.151	3.790	3.140	2.762	2.514	2.339	2.208	2.107	2.026	1.960	1.904	1.857	1.817	1.782	1.751	1.723	1.699	1.677	1.639	1.607	1.581	1.558	1.538	1.467	1.424	1.395	1.374	1.358	1.346	1.336	1.320	1.309	1.289	1.240
70	62.87	9.477	5.149	3.786	3.135	2.756	2.508	2.333	2.202	2.100	2.019	1.952	1.896	1.849	1.808	1.773	1.742	1.714	1.690	1.667	1.629	1.597	1.570	1.547	1.527	1.455	1.412	1.382	1.361	1.344	1.332	1.321	1.305	1.294	1.273	1.222
80	62.93	9.479	5.147	3.782	3.132	2.752	2.504	2.328	2.196	2.095	2.013	1.946	1.890	1.843	1.802	1.766	1.735	1.707	1.683	1.660	1.622	1.590	1.562	1.539	1.519	1.447	1.402	1.372	1.350	1.334	1.321	1.310	1.294	1.282	1.261	1.207
90	62.97	9.480	5.145	3.780	3.129	2.749	2.500	2.324	2.192	2.090	2.009	1.942	1.886	1.838	1.797	1.761	1.730	1.702	1.677	1.655	1.616	1.584	1.556	1.533	1.512	1.439	1.395	1.364	1.342	1.325	1.312	1.301	1.284	1.272	1.250	1.195
100	63.01	9.481	5.144	3.778	3.126	2.746	2.497	2.321	2.189	2.087	2.005	1.938	1.882	1.834	1.793	1.757	1.726	1.698	1.673	1.650	1.611	1.579	1.551	1.528	1.507	1.434	1.388	1.358	1.335	1.318	1.304	1.293	1.277	1.264	1.242	1.185
120	63.06	9.483	5.143	3.775	3.123	2.742	2.493	2.316	2.184	2.082	2.000	1.932	1.876	1.828	1.787	1.751	1.719	1.691	1.666	1.643	1.604	1.571	1.544	1.520	1.499	1.425	1.379	1.348	1.325	1.307	1.293	1.282	1.265	1.252	1.228	1.169
140	63.10	9.484	5.141	3.773	3.120	2.739	2.490	2.313	2.181	2.078	1.996	1.928	1.872	1.824	1.782	1.746	1.714	1.686	1.661	1.638	1.599	1.566	1.538	1.514	1.493	1.418	1.372	1.340	1.317	1.299	1.285	1.273	1.256	1.243	1.218	1.156
200	63.17	9.486	5.139	3.769	3.116	2.734	2.484	2.307	2.174	2.071	1.989	1.921	1.864	1.816	1.774	1.738	1.706	1.678	1.652	1.629	1.590	1.556	1.528	1.504	1.482	1.406	1.359	1.326	1.302	1.284	1.269	1.257	1.239	1.225	1.199	1.130
∞	63.33	9.491	5.134	3.761	3.105	2.722	2.471	2.293	2.159	2.055	1.972	1.904	1.846	1.797	1.755	1.718	1.686	1.657	1.631	1.607	1.567	1.533	1.504	1.478	1.456	1.377	1.327	1.291	1.265	1.245	1.228	1.214	1.193	1.176	1.144	1.000

Tavola dei quantili 0.95 della distribuzione F(m,n)

																	n																		
m	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**22**	**24**	**26**	**28**	**30**	**40**	**50**	**60**	**70**	**80**	**90**	**100**	**120**	**140**	**200**	**∞**
1	161.45	18.513	10.128	7.709	6.608	5.987	5.591	5.318	5.117	4.965	4.844	4.747	4.667	4.600	4.543	4.494	4.451	4.414	4.381	4.351	4.301	4.260	4.225	4.196	4.171	4.085	4.034	4.001	3.978	3.960	3.947	3.936	3.920	3.909	3.888	3.841
2	199.50	19.000	9.552	6.944	5.786	5.143	4.737	4.459	4.256	4.103	3.982	3.885	3.806	3.739	3.682	3.634	3.592	3.555	3.522	3.493	3.443	3.403	3.369	3.340	3.316	3.232	3.183	3.150	3.128	3.111	3.098	3.087	3.072	3.061	3.041	2.996
3	215.71	19.164	9.277	6.591	5.409	4.757	4.347	4.066	3.863	3.708	3.587	3.490	3.411	3.344	3.287	3.239	3.197	3.160	3.127	3.098	3.049	3.009	2.975	2.947	2.922	2.839	2.790	2.758	2.736	2.719	2.706	2.696	2.680	2.669	2.650	2.605
4	224.58	19.247	9.117	6.388	5.192	4.534	4.120	3.838	3.633	3.478	3.357	3.259	3.179	3.112	3.056	3.007	2.965	2.928	2.895	2.866	2.817	2.776	2.743	2.714	2.690	2.606	2.557	2.525	2.503	2.486	2.473	2.463	2.447	2.436	2.417	2.372
5	230.16	19.296	9.013	6.256	5.050	4.387	3.972	3.688	3.482	3.326	3.204	3.106	3.025	2.958	2.901	2.852	2.810	2.773	2.740	2.711	2.661	2.621	2.587	2.558	2.534	2.449	2.400	2.368	2.346	2.329	2.316	2.305	2.290	2.279	2.259	2.214
6	233.99	19.329	8.941	6.163	4.950	4.284	3.866	3.581	3.374	3.217	3.095	2.996	2.915	2.848	2.790	2.741	2.699	2.661	2.628	2.599	2.549	2.508	2.474	2.445	2.421	2.336	2.286	2.254	2.231	2.214	2.201	2.191	2.175	2.164	2.144	2.099
7	236.77	19.353	8.887	6.094	4.876	4.207	3.787	3.500	3.293	3.135	3.012	2.913	2.832	2.764	2.707	2.657	2.614	2.577	2.544	2.514	2.464	2.423	2.388	2.359	2.334	2.249	2.199	2.167	2.143	2.126	2.113	2.103	2.087	2.076	2.056	2.010
8	238.88	19.371	8.845	6.041	4.818	4.147	3.726	3.438	3.230	3.072	2.948	2.849	2.767	2.699	2.641	2.591	2.548	2.510	2.477	2.447	2.397	2.355	2.321	2.291	2.266	2.180	2.130	2.097	2.074	2.056	2.043	2.032	2.016	2.005	1.985	1.938
9	240.54	19.385	8.812	5.999	4.772	4.099	3.677	3.388	3.179	3.020	2.896	2.796	2.714	2.646	2.588	2.538	2.494	2.456	2.423	2.393	2.342	2.300	2.265	2.236	2.211	2.124	2.073	2.040	2.017	1.999	1.986	1.975	1.959	1.947	1.927	1.880
10	241.88	19.396	8.785	5.964	4.735	4.060	3.637	3.347	3.137	2.978	2.854	2.753	2.671	2.602	2.544	2.494	2.450	2.412	2.378	2.348	2.297	2.255	2.220	2.190	2.165	2.077	2.026	1.993	1.969	1.951	1.938	1.927	1.910	1.899	1.878	1.831
11	242.98	19.405	8.763	5.936	4.704	4.027	3.603	3.313	3.102	2.943	2.818	2.717	2.635	2.565	2.507	2.456	2.413	2.374	2.340	2.310	2.259	2.216	2.181	2.151	2.126	2.038	1.986	1.952	1.928	1.910	1.897	1.886	1.869	1.858	1.837	1.789
12	243.90	19.412	8.745	5.912	4.678	4.000	3.575	3.284	3.073	2.913	2.788	2.687	2.604	2.534	2.475	2.425	2.381	2.342	2.308	2.278	2.226	2.183	2.148	2.118	2.092	2.003	1.952	1.917	1.893	1.875	1.861	1.850	1.834	1.822	1.801	1.752
13	244.69	19.419	8.729	5.891	4.655	3.976	3.550	3.259	3.048	2.887	2.761	2.660	2.577	2.507	2.448	2.397	2.353	2.314	2.280	2.250	2.198	2.155	2.119	2.089	2.063	1.974	1.921	1.887	1.863	1.845	1.830	1.819	1.803	1.791	1.769	1.720
14	245.36	19.424	8.715	5.873	4.636	3.956	3.529	3.237	3.025	2.865	2.739	2.637	2.554	2.484	2.424	2.373	2.329	2.290	2.256	2.225	2.173	2.130	2.094	2.064	2.037	1.948	1.895	1.860	1.836	1.817	1.803	1.792	1.775	1.763	1.742	1.692
15	245.95	19.429	8.703	5.858	4.619	3.938	3.511	3.218	3.006	2.845	2.719	2.617	2.533	2.463	2.403	2.352	2.308	2.269	2.234	2.203	2.151	2.108	2.072	2.041	2.015	1.924	1.871	1.836	1.812	1.793	1.779	1.768	1.750	1.738	1.717	1.666
16	246.47	19.433	8.692	5.844	4.604	3.922	3.494	3.202	2.989	2.828	2.701	2.599	2.515	2.445	2.385	2.333	2.289	2.250	2.215	2.184	2.131	2.088	2.052	2.021	1.995	1.904	1.850	1.815	1.790	1.772	1.757	1.746	1.728	1.716	1.694	1.644
17	246.92	19.437	8.683	5.832	4.590	3.908	3.480	3.187	2.974	2.812	2.685	2.583	2.499	2.428	2.368	2.317	2.272	2.233	2.198	2.167	2.114	2.070	2.034	2.003	1.976	1.885	1.831	1.796	1.771	1.752	1.737	1.726	1.709	1.696	1.674	1.623
18	247.32	19.440	8.675	5.821	4.579	3.896	3.467	3.173	2.960	2.798	2.671	2.568	2.484	2.413	2.353	2.302	2.257	2.217	2.182	2.151	2.098	2.054	2.018	1.987	1.960	1.868	1.814	1.778	1.753	1.734	1.720	1.708	1.690	1.678	1.656	1.604
19	247.69	19.443	8.667	5.811	4.568	3.884	3.455	3.161	2.948	2.785	2.658	2.555	2.471	2.400	2.340	2.288	2.243	2.203	2.168	2.137	2.084	2.040	2.003	1.972	1.945	1.853	1.798	1.763	1.737	1.718	1.703	1.691	1.674	1.661	1.639	1.586
20	248.02	19.446	8.660	5.803	4.558	3.874	3.445	3.150	2.936	2.774	2.646	2.544	2.459	2.388	2.328	2.276	2.230	2.191	2.155	2.124	2.071	2.027	1.990	1.959	1.932	1.839	1.784	1.748	1.722	1.703	1.688	1.676	1.659	1.646	1.623	1.571
21	248.31	19.448	8.654	5.795	4.549	3.865	3.435	3.140	2.926	2.764	2.636	2.533	2.448	2.377	2.316	2.264	2.219	2.179	2.144	2.112	2.059	2.015	1.978	1.946	1.919	1.826	1.771	1.735	1.709	1.689	1.675	1.663	1.645	1.632	1.609	1.556
22	248.58	19.450	8.648	5.787	4.541	3.856	3.426	3.131	2.917	2.754	2.626	2.523	2.438	2.367	2.306	2.254	2.208	2.168	2.133	2.102	2.048	2.003	1.966	1.935	1.908	1.814	1.759	1.722	1.696	1.677	1.662	1.650	1.632	1.619	1.596	1.542
23	248.82	19.452	8.643	5.781	4.534	3.849	3.418	3.123	2.908	2.745	2.617	2.514	2.429	2.357	2.297	2.244	2.199	2.159	2.123	2.092	2.038	1.993	1.956	1.924	1.897	1.803	1.748	1.711	1.685	1.665	1.650	1.638	1.620	1.607	1.583	1.529
24	249.05	19.454	8.638	5.774	4.527	3.841	3.410	3.115	2.900	2.737	2.609	2.505	2.420	2.349	2.288	2.235	2.190	2.150	2.114	2.082	2.028	1.984	1.946	1.915	1.887	1.793	1.737	1.700	1.674	1.654	1.639	1.627	1.608	1.595	1.572	1.517
25	249.26	19.456	8.634	5.769	4.521	3.835	3.404	3.108	2.893	2.730	2.601	2.498	2.412	2.341	2.280	2.227	2.181	2.141	2.106	2.074	2.020	1.975	1.938	1.906	1.878	1.783	1.727	1.690	1.664	1.644	1.629	1.616	1.598	1.585	1.561	1.506
26	249.45	19.457	8.630	5.763	4.515	3.829	3.397	3.102	2.886	2.723	2.594	2.491	2.405	2.333	2.272	2.220	2.174	2.134	2.098	2.066	2.012	1.967	1.929	1.897	1.870	1.775	1.718	1.681	1.654	1.634	1.619	1.607	1.588	1.575	1.551	1.496
27	249.63	19.459	8.626	5.759	4.510	3.823	3.391	3.095	2.880	2.716	2.588	2.484	2.398	2.326	2.265	2.212	2.167	2.126	2.090	2.059	2.004	1.959	1.921	1.889	1.862	1.766	1.710	1.672	1.646	1.626	1.610	1.598	1.579	1.566	1.542	1.486
28	249.80	19.460	8.623	5.754	4.505	3.818	3.386	3.090	2.874	2.710	2.582	2.478	2.392	2.320	2.259	2.206	2.160	2.119	2.084	2.052	1.997	1.952	1.914	1.882	1.854	1.759	1.702	1.664	1.637	1.617	1.601	1.589	1.570	1.557	1.533	1.476
29	249.95	19.461	8.620	5.750	4.500	3.813	3.381	3.084	2.869	2.705	2.576	2.472	2.386	2.314	2.253	2.200	2.154	2.113	2.077	2.045	1.990	1.945	1.907	1.875	1.847	1.751	1.694	1.656	1.629	1.609	1.593	1.581	1.562	1.549	1.524	1.467
30	250.10	19.463	8.617	5.746	4.496	3.808	3.376	3.079	2.864	2.700	2.570	2.466	2.380	2.308	2.247	2.194	2.148	2.107	2.071	2.039	1.984	1.939	1.901	1.869	1.841	1.744	1.687	1.649	1.622	1.602	1.586	1.573	1.554	1.541	1.516	1.459
40	251.14	19.471	8.594	5.717	4.464	3.774	3.340	3.043	2.826	2.661	2.531	2.426	2.339	2.266	2.204	2.151	2.104	2.063	2.026	1.994	1.938	1.892	1.853	1.820	1.792	1.693	1.634	1.594	1.566	1.545	1.528	1.515	1.495	1.481	1.455	1.394
50	251.77	19.476	8.581	5.699	4.444	3.754	3.319	3.020	2.803	2.637	2.507	2.401	2.314	2.241	2.178	2.124	2.077	2.035	1.999	1.966	1.909	1.863	1.823	1.790	1.761	1.660	1.599	1.559	1.530	1.508	1.491	1.477	1.457	1.442	1.415	1.350
60	252.20	19.479	8.572	5.688	4.431	3.740	3.304	3.005	2.787	2.621	2.490	2.384	2.297	2.223	2.160	2.106	2.058	2.017	1.980	1.946	1.889	1.842	1.803	1.769	1.740	1.637	1.576	1.534	1.505	1.482	1.465	1.450	1.429	1.414	1.386	1.318
70	252.50	19.481	8.566	5.679	4.422	3.730	3.294	2.994	2.776	2.609	2.478	2.372	2.284	2.210	2.147	2.093	2.045	2.003	1.966	1.932	1.875	1.828	1.788	1.754	1.724	1.621	1.558	1.516	1.486	1.463	1.445	1.430	1.408	1.393	1.364	1.293
80	252.72	19.483	8.561	5.673	4.415	3.722	3.286	2.986	2.768	2.601	2.469	2.363	2.275	2.201	2.137	2.083	2.035	1.993	1.955	1.922	1.864	1.816	1.776	1.742	1.712	1.608	1.544	1.502	1.471	1.448	1.429	1.415	1.392	1.376	1.346	1.273
90	252.90	19.485	8.557	5.668	4.409	3.716	3.280	2.980	2.761	2.594	2.462	2.356	2.267	2.193	2.130	2.075	2.027	1.985	1.947	1.913	1.856	1.808	1.767	1.733	1.703	1.597	1.534	1.491	1.459	1.436	1.417	1.402	1.379	1.363	1.332	1.257
100	253.04	19.486	8.554	5.664	4.405	3.712	3.275	2.975	2.756	2.588	2.457	2.350	2.261	2.187	2.123	2.068	2.020	1.978	1.940	1.907	1.849	1.800	1.760	1.725	1.695	1.589	1.525	1.481	1.450	1.426	1.407	1.392	1.369	1.352	1.321	1.243
120	253.25	19.487	8.549	5.658	4.398	3.705	3.267	2.967	2.748	2.580	2.448	2.341	2.252	2.178	2.114	2.059	2.011	1.968	1.930	1.896	1.838	1.790	1.749	1.714	1.683	1.577	1.511	1.467	1.435	1.411	1.391	1.376	1.352	1.335	1.302	1.221
140	253.41	19.489	8.546	5.654	4.394	3.700	3.262	2.961	2.742	2.574	2.442	2.335	2.246	2.171	2.107	2.052	2.004	1.961	1.923	1.889	1.830	1.782	1.741	1.706	1.675	1.567	1.502	1.457	1.424	1.399	1.380	1.364	1.340	1.322	1.289	1.204
200	253.68	19.491	8.540	5.646	4.385	3.690	3.252	2.951	2.731	2.563	2.431	2.323	2.234	2.159	2.095	2.039	1.991	1.948	1.910	1.875	1.817	1.768	1.726	1.691	1.660	1.551	1.484	1.438	1.404	1.379	1.358	1.342	1.316	1.298	1.263	1.170
∞	254.31	19.496	8.526	5.628	4.365	3.669	3.230	2.928	2.707	2.538	2.404	2.296	2.206	2.131	2.066	2.010	1.960	1.917	1.878	1.843	1.783	1.733	1.691	1.654	1.622	1.509	1.438	1.389	1.353	1.325	1.302	1.283	1.254	1.232	1.189	1.000

Tavola dei quantili 0.99 della distribuzione F(m,n)

	n																																		
m	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**22**	**24**	**26**	**28**	**30**	**40**	**50**	**60**	**70**	**80**	**90**	**100**	**120**	**140**	**200**	**∞**
1	4052	98.50	34.116	21.198	16.258	13.745	12.246	11.259	10.562	10.044	9.646	9.330	9.074	8.862	8.683	8.531	8.400	8.285	8.185	8.096	7.945	7.823	7.721	7.636	7.562	7.314	7.171	7.077	7.011	6.963	6.925	6.895	6.851	6.819	6.763	6.635
2	4999	99.00	30.816	18.000	13.274	10.925	9.547	8.649	8.022	7.559	7.206	6.927	6.701	6.515	6.359	6.226	6.112	6.013	5.926	5.849	5.719	5.614	5.526	5.453	5.390	5.178	5.057	4.977	4.922	4.881	4.849	4.824	4.787	4.760	4.713	4.605
3	5404	99.16	29.457	16.694	12.060	9.780	8.451	7.591	6.992	6.552	6.217	5.953	5.739	5.564	5.417	5.292	5.185	5.092	5.010	4.938	4.817	4.718	4.637	4.568	4.510	4.313	4.199	4.126	4.074	4.036	4.007	3.984	3.949	3.925	3.881	3.782
4	5624	99.25	28.710	15.977	11.392	9.148	7.847	7.006	6.422	5.994	5.668	5.412	5.205	5.035	4.893	4.773	4.669	4.579	4.500	4.431	4.313	4.218	4.140	4.074	4.018	3.828	3.720	3.649	3.600	3.563	3.535	3.513	3.480	3.456	3.414	3.319
5	5764	99.30	28.237	15.522	10.967	8.746	7.460	6.632	6.057	5.636	5.316	5.064	4.862	4.695	4.556	4.437	4.336	4.248	4.171	4.103	3.988	3.895	3.818	3.754	3.699	3.514	3.408	3.339	3.291	3.255	3.228	3.206	3.174	3.151	3.110	3.017
6	5859	99.33	27.911	15.207	10.672	8.466	7.191	6.371	5.802	5.386	5.069	4.821	4.620	4.456	4.318	4.202	4.101	4.015	3.939	3.871	3.758	3.667	3.591	3.528	3.473	3.291	3.186	3.119	3.071	3.036	3.009	2.988	2.956	2.933	2.893	2.802
7	5928	99.36	27.671	14.976	10.456	8.260	6.993	6.178	5.613	5.200	4.886	4.640	4.441	4.278	4.142	4.026	3.927	3.841	3.765	3.699	3.587	3.496	3.421	3.358	3.305	3.124	3.020	2.953	2.906	2.871	2.845	2.823	2.792	2.769	2.730	2.639
8	5981	99.38	27.489	14.799	10.289	8.102	6.840	6.029	5.467	5.057	4.744	4.499	4.302	4.140	4.004	3.890	3.791	3.705	3.631	3.564	3.453	3.363	3.288	3.226	3.173	2.993	2.890	2.823	2.777	2.742	2.715	2.694	2.663	2.641	2.601	2.511
9	6022	99.39	27.345	14.659	10.158	7.976	6.719	5.911	5.351	4.942	4.632	4.388	4.191	4.030	3.895	3.780	3.682	3.597	3.523	3.457	3.346	3.256	3.182	3.120	3.067	2.888	2.785	2.718	2.672	2.637	2.611	2.590	2.559	2.536	2.497	2.407
10	6056	99.40	27.228	14.546	10.051	7.874	6.620	5.814	5.257	4.849	4.539	4.296	4.100	3.939	3.805	3.691	3.593	3.508	3.434	3.368	3.258	3.168	3.094	3.032	2.979	2.801	2.698	2.632	2.585	2.551	2.524	2.503	2.472	2.450	2.411	2.321
11	6083	99.41	27.132	14.452	9.963	7.790	6.538	5.734	5.178	4.772	4.462	4.220	4.025	3.864	3.730	3.616	3.518	3.434	3.360	3.294	3.184	3.094	3.021	2.959	2.906	2.727	2.625	2.559	2.512	2.478	2.451	2.430	2.399	2.377	2.338	2.248
12	6107	99.42	27.052	14.374	9.888	7.718	6.469	5.667	5.111	4.706	4.397	4.155	3.960	3.800	3.666	3.553	3.455	3.371	3.297	3.231	3.121	3.032	2.958	2.896	2.843	2.665	2.563	2.496	2.450	2.415	2.389	2.368	2.336	2.314	2.275	2.185
13	6126	99.42	26.983	14.306	9.825	7.657	6.410	5.609	5.055	4.650	4.342	4.100	3.905	3.745	3.612	3.498	3.401	3.316	3.242	3.177	3.067	2.977	2.904	2.842	2.789	2.611	2.508	2.442	2.395	2.361	2.334	2.313	2.282	2.260	2.220	2.130
14	6143	99.43	26.924	14.249	9.770	7.605	6.359	5.559	5.005	4.601	4.293	4.052	3.857	3.698	3.564	3.451	3.353	3.269	3.195	3.130	3.019	2.930	2.857	2.795	2.742	2.563	2.461	2.394	2.348	2.313	2.286	2.265	2.234	2.212	2.172	2.082
15	6157	99.43	26.872	14.198	9.722	7.559	6.314	5.515	4.962	4.558	4.251	4.010	3.815	3.656	3.522	3.409	3.312	3.227	3.153	3.088	2.978	2.889	2.815	2.753	2.700	2.522	2.419	2.352	2.306	2.271	2.244	2.223	2.191	2.169	2.129	2.039
16	6170	99.44	26.826	14.154	9.680	7.519	6.275	5.477	4.924	4.520	4.213	3.972	3.778	3.619	3.485	3.372	3.275	3.190	3.116	3.051	2.941	2.852	2.778	2.716	2.663	2.484	2.382	2.315	2.268	2.233	2.206	2.185	2.154	2.131	2.091	2.000
17	6181	99.44	26.786	14.114	9.643	7.483	6.240	5.442	4.890	4.487	4.180	3.939	3.745	3.586	3.452	3.339	3.242	3.158	3.084	3.018	2.908	2.819	2.745	2.683	2.630	2.451	2.348	2.281	2.234	2.199	2.172	2.151	2.119	2.097	2.057	1.965
18	6191	99.44	26.751	14.079	9.609	7.451	6.209	5.412	4.860	4.457	4.150	3.910	3.716	3.556	3.423	3.310	3.212	3.128	3.054	2.989	2.879	2.789	2.715	2.653	2.600	2.421	2.318	2.251	2.204	2.169	2.142	2.120	2.089	2.066	2.026	1.934
19	6201	99.45	26.719	14.048	9.580	7.422	6.181	5.384	4.833	4.430	4.123	3.883	3.689	3.529	3.396	3.283	3.186	3.101	3.027	2.962	2.852	2.762	2.688	2.626	2.573	2.394	2.290	2.223	2.176	2.141	2.114	2.092	2.060	2.038	1.997	1.905
20	6209	99.45	26.690	14.019	9.553	7.396	6.155	5.359	4.808	4.405	4.099	3.858	3.665	3.505	3.372	3.259	3.162	3.077	3.003	2.938	2.827	2.738	2.664	2.602	2.549	2.369	2.265	2.198	2.150	2.115	2.088	2.067	2.035	2.012	1.971	1.878
21	6216	99.45	26.664	13.994	9.528	7.372	6.132	5.336	4.786	4.383	4.077	3.836	3.643	3.483	3.350	3.237	3.139	3.055	2.981	2.916	2.805	2.716	2.642	2.579	2.526	2.346	2.242	2.175	2.127	2.092	2.065	2.043	2.011	1.988	1.947	1.854
22	6223	99.46	26.639	13.970	9.506	7.351	6.111	5.316	4.765	4.363	4.057	3.816	3.622	3.463	3.330	3.216	3.119	3.035	2.961	2.895	2.785	2.695	2.621	2.559	2.506	2.325	2.221	2.153	2.106	2.070	2.043	2.021	1.989	1.966	1.925	1.831
23	6229	99.46	26.617	13.949	9.485	7.331	6.092	5.297	4.746	4.344	4.038	3.798	3.604	3.444	3.311	3.198	3.101	3.016	2.942	2.877	2.766	2.676	2.602	2.540	2.487	2.306	2.202	2.134	2.086	2.050	2.023	2.001	1.969	1.946	1.905	1.810
24	6234	99.46	26.597	13.929	9.466	7.313	6.074	5.279	4.729	4.327	4.021	3.780	3.587	3.427	3.294	3.181	3.083	2.999	2.925	2.859	2.749	2.659	2.585	2.522	2.469	2.288	2.183	2.115	2.067	2.032	2.004	1.983	1.950	1.927	1.886	1.791
25	6240	99.46	26.579	13.911	9.449	7.296	6.058	5.263	4.713	4.311	4.005	3.765	3.571	3.412	3.278	3.165	3.068	2.983	2.909	2.843	2.733	2.643	2.569	2.506	2.453	2.271	2.167	2.098	2.050	2.015	1.987	1.965	1.932	1.909	1.868	1.773
26	6245	99.46	26.562	13.894	9.433	7.281	6.043	5.248	4.698	4.296	3.990	3.750	3.556	3.397	3.264	3.150	3.053	2.968	2.894	2.829	2.718	2.628	2.554	2.491	2.437	2.256	2.151	2.083	2.034	1.999	1.971	1.949	1.916	1.893	1.851	1.755
27	6249	99.46	26.546	13.878	9.418	7.266	6.029	5.234	4.684	4.283	3.977	3.736	3.543	3.383	3.250	3.137	3.039	2.955	2.880	2.815	2.704	2.614	2.540	2.477	2.423	2.241	2.136	2.068	2.019	1.983	1.956	1.934	1.901	1.877	1.836	1.739
28	6253	99.46	26.531	13.864	9.404	7.253	6.016	5.221	4.672	4.270	3.964	3.724	3.530	3.371	3.237	3.124	3.026	2.942	2.868	2.802	2.691	2.601	2.526	2.464	2.410	2.228	2.123	2.054	2.005	1.969	1.942	1.919	1.886	1.863	1.821	1.724
29	6257	99.46	26.517	13.850	9.391	7.240	6.003	5.209	4.660	4.258	3.952	3.712	3.518	3.359	3.225	3.112	3.014	2.930	2.855	2.790	2.679	2.589	2.514	2.451	2.398	2.215	2.110	2.041	1.992	1.956	1.928	1.906	1.873	1.849	1.807	1.710
30	6260	99.47	26.504	13.838	9.379	7.229	5.992	5.198	4.649	4.247	3.941	3.701	3.507	3.348	3.214	3.101	3.003	2.919	2.844	2.778	2.667	2.577	2.503	2.440	2.386	2.203	2.098	2.028	1.980	1.944	1.916	1.893	1.860	1.836	1.794	1.696
40	6286	99.48	26.411	13.745	9.291	7.143	5.908	5.116	4.567	4.165	3.860	3.619	3.425	3.266	3.132	3.018	2.920	2.835	2.761	2.695	2.583	2.492	2.417	2.354	2.299	2.114	2.007	1.936	1.886	1.849	1.820	1.797	1.763	1.738	1.694	1.592
50	6302	99.48	26.354	13.690	9.238	7.091	5.858	5.065	4.517	4.115	3.810	3.569	3.375	3.215	3.081	2.967	2.869	2.784	2.709	2.643	2.531	2.440	2.364	2.300	2.245	2.058	1.949	1.877	1.826	1.788	1.759	1.735	1.700	1.675	1.629	1.523
60	6313	99.48	26.316	13.652	9.202	7.057	5.824	5.032	4.483	4.082	3.776	3.535	3.341	3.181	3.047	2.933	2.835	2.749	2.674	2.608	2.495	2.403	2.327	2.263	2.208	2.019	1.909	1.836	1.785	1.746	1.716	1.692	1.656	1.630	1.583	1.473
70	6321	99.48	26.289	13.626	9.176	7.032	5.799	5.007	4.459	4.058	3.752	3.511	3.317	3.157	3.022	2.908	2.810	2.724	2.649	2.582	2.469	2.377	2.301	2.236	2.181	1.991	1.880	1.806	1.754	1.714	1.684	1.659	1.623	1.596	1.548	1.435
80	6326	99.48	26.269	13.605	9.157	7.013	5.781	4.989	4.441	4.039	3.734	3.493	3.298	3.138	3.004	2.889	2.791	2.705	2.630	2.563	2.450	2.357	2.281	2.216	2.160	1.969	1.857	1.783	1.730	1.690	1.659	1.634	1.597	1.570	1.521	1.404
90	6331	99.49	26.253	13.590	9.142	6.998	5.766	4.975	4.426	4.025	3.719	3.478	3.284	3.124	2.989	2.875	2.776	2.690	2.614	2.548	2.434	2.342	2.265	2.200	2.144	1.952	1.839	1.764	1.711	1.671	1.639	1.614	1.576	1.549	1.499	1.379
100	6334	99.49	26.241	13.577	9.130	6.987	5.755	4.963	4.415	4.014	3.708	3.467	3.272	3.112	2.977	2.863	2.764	2.678	2.602	2.535	2.422	2.329	2.252	2.187	2.131	1.938	1.825	1.749	1.695	1.655	1.623	1.598	1.559	1.532	1.481	1.358
120	6340	99.49	26.221	13.558	9.112	6.969	5.737	4.946	4.398	3.996	3.690	3.449	3.255	3.094	2.959	2.845	2.746	2.660	2.584	2.517	2.403	2.310	2.233	2.167	2.111	1.917	1.803	1.726	1.672	1.630	1.598	1.572	1.533	1.505	1.453	1.325
140	6343	99.49	26.208	13.545	9.099	6.956	5.725	4.934	4.385	3.984	3.678	3.437	3.242	3.082	2.947	2.832	2.733	2.647	2.571	2.503	2.389	2.296	2.219	2.153	2.096	1.902	1.786	1.710	1.654	1.613	1.580	1.554	1.514	1.485	1.432	1.299
200	6350	99.49	26.183	13.520	9.075	6.934	5.702	4.911	4.363	3.962	3.656	3.414	3.219	3.059	2.923	2.808	2.709	2.623	2.547	2.479	2.365	2.271	2.193	2.127	2.070	1.874	1.757	1.678	1.622	1.579	1.546	1.518	1.477	1.447	1.391	1.247
∞	6366	99.50	26.125	13.463	9.020	6.880	5.650	4.859	4.311	3.909	3.602	3.361	3.165	3.004	2.868	2.753	2.653	2.566	2.489	2.421	2.305	2.211	2.131	2.064	2.006	1.805	1.683	1.601	1.540	1.494	1.457	1.427	1.381	1.346	1.279	1.000