Probabilità

Claudio Agostinelli
claudio.agostinelli@unitn.it
Dipartimento di Matematica
Università di Trento
http://www.science.unitn.it/~claudio

February 12, 2016

Indice

Prefazione

Questo materiale non è ancora nella sua forma definitiva.

La prima stesura risale all'anno accademico 2003-2004 per l'insegnamento di Probabilità e Statistica tenuto presso l'Università Ca' Foscari di Venezia agli studenti del corso di laurea in Informatica. È utilizzato come traccia per l'insegnamento di Probabilità e Statistica dell'Università di Trento, per il corso di laurea in Informatica dal 2015-2016.

Questo materiale non è un libro e neanche una dispensa! Viene attualmente distribuito agli studenti dell'insegnamento di Probabilità e Statistica (DISI, Trento) come materiale aggiuntivo al libro di testo per facilitarne l'apprendimento.

Trento, 12 Febbraio 2016

Claudio Agostinelli

Informazioni generali

- Nome dell'insegnamento: **Probabilità e Statistica**
- Crediti: 6cfu
- Corsi di Studi: Informatica
- Docente del corso: **Claudio Agostinelli**
  - email: `claudio.agostinelli@unitn.it`
  - pagina personale di Ateneo:
  - `http://www5.unitn.it/People/it/Web/Persona/PER0181898#INFO`
  - pagina personale:
  - `http://www.science.unitn.it/~claudio`
- Ricevimento studenti:
  - Sono disponibili in formato iCal all'indirizzo
  - `http://tinyurl.com/zhhz528`.
- Lezioni: Orario e luogo delle lezioni sono disponibili in formato iCal all'indirizzo `http://tinyurl.com/hwxtw8v`
- Materiale scaricabile:
- `http://www.science.unitn.it/~claudio/teaching/ps/index.html`
- Testo di riferimento: .

Modalità d'esame

- L'esame è composto da una prova scritta, che è
  - Obbligatoria.
  - Composta da quattro esercizi da svolgere e da 10 domande a risposta multipla.
  - Con durata di 2 ore.

Materiale utilizzabile durante l'esame

- **Si può usare**
  - Un formulario con alcune formule e tavole delle principali distribuzioni, scaricabili dal sito dell'insegnamento;
  - Calcolatrice non programmabile e/o Regolo calcolatore;
  - Penna nera o blu, matita e gomma.
- **Non si può usare** tutto quello non espressamente dichiarato nel punto precedente, in particolare
  - Telefonino;
  - Calcolatrici programmabili e/o Palmari;
  - Libri e/o appunti;
  - Fogli non forniti dal docente.

Capitolo 1

Probabilità

1.1 Probabilità

Tutte le volte in cui si ragiona per induzione, si fa, più o meno coscientemente, uso del calcolo delle probabilità

(H. Poincarè)

Programma del corso

- Concetto di Probabilità. Elementi di algebra degli insiemi e di calcolo combinatorio. Definizione assiomatica. Principali teoremi, probabilità condizionata, indipendenza stocastica, teorema di Bayes;
- Variabili aleatorie unidimensionali. Funzione di ripartizione e densità, momenti, funzione generatrice dei momenti;
- Particolari famiglie di distribuzioni unidimensionali: uniforme discreta, Bernoulliana, binomiale, ipergeometrica, Poisson, geometrica, uniforme continua, normale, beta, gamma, esponenziale;
- Variabili aleatorie multiple. Funzioni di ripartizione e densità congiunte, distribuzioni marginali e condizionate. Momenti. Funzione di regressione. Correlazione e indipendenza stocastica. Distribuzione normale multivariata.
- Trasformazione di variabili casuali e principali risultati;
- Successione di variabili aleatorie. Disuguaglianze notevoli: Jensen e Chebyshev. Convergenza: quasi certa, in probabilità, in media quadratica e in distribuzione. Legge dei grandi numeri. Teorema limite centrale;
- Statistica Descrittiva e
- Rappresentazioni grafiche
- Statistiche, stimatori, principio del campionamento ripetuto e proprietà degli stimatori;
- Stima puntuale: metodo dei minimi quadrati, metodo dei momenti e metodo di verosimiglianza;
- Elementi di verifica d'ipotesi e di stima intervallare;

Avvertenza

Queste note non sono da considerarsi appunti o dispense del corso e non possono sostituire l'uso del (dei) libro (libri) di testo. Sono in una forma non definitiva

Probabilità e Statistica

La Probabilità e la Statistica sono parti della Matematica.

Esse ci consentono di studiare fenomeni che presentano variabilità o incertezza nel loro esito:

- tra diversi casi o soggetti;
- tra diversi istanti temporali;
- tra diversi luoghi;
- tra diverse condizioni di misura;
- per l'impossibilità di separare l'osservazione del fenomeno di interesse da altri fenomeni di disturbo.

La Statistica, è spesso suddivisa in due aree principali:

- Descrittiva: usata per descrivere un fenomeno attraverso la sintesi di informazioni; Esplorazione statistica dei dati, statistica senza un modello probabilistico. Disponiamo di dati riferiti a tutta la popolazione di riferimento.
- Inferenziale: usata per costruire modelli (anche di natura probabilistica) interpretativi della variabilità di un fenomeno a partire da una informazione limitata e utilizzando il modello costruito formulare previsioni sull'andamento futuro (o non osservato) del fenomeno.

  I dati disponibili sono stati rilevati solamente su una parte delle unità statistiche (il campione da cui indagini campionarie). Vogliamo utilizzare le informazioni del campione per fare delle affermazioni sulle caratteristiche di tutta la popolazione.

La Probabilità fornisce:

- Teorie di base (in genere di natura astratta e quindi applicabile ad ampie classi di problemi);

- Strumenti matematici;
- Modelli probabilistici.

Entrambe presentano impostazioni diverse:

- Classica
- Frequentista (Neyman, Pearson)
- Bayesiana (Bayes)
  - soggettivista (Savage, De Finetti)
  - non soggettivista
- Logica (Carnap)
- Comparativa (Barnett)
- Fisheriana (Fisher, una impostazione che riguarda sopratutto la Statistica)

L'Approccio Assiomatico alla Probabilità (1933) dovuto a A.N. Kolmogorov (1903-1987) consente di sviluppare una teoria generale della probabilità che è staccata dalla sua interpretazione.

Concetto di Probabilità – O–ring

Supponiamo di avere aquistato una scatola di O–ring (guarnizioni) costituita da 90 pezzi di cui 30 pezzi con un diametro di $9mm$, 30 pezzi con diametro $10mm$ e le rimanenti con un diametro di $11mm$. Qual'è la probabilità che io estragga dalla scatola un O–ring di diametro $10mm$?

La riposta dipende molto da come avviene l'estrazione:

- scelgo senza guardare (e senza selezionare attraverso il tatto) la prima guarnizione con cui vengo in contatto;
- rovescio la scatola e considero la guarnizione che tocca per ultima il piano di appoggio;
- analizzo accuratamente le guarnizioni prima di sceglierle in modo da prendere quella con diametro $10mm$;
- estraggo senza guardare una guarnizione alla volta e decido di fermarmi appena ritengo di aver estratto una guarnizione di diametro $10mm$;

- Scelgo senza guardare (e senza selezionare attraverso il tatto) la prima guarnizione con cui vengo in contatto

In questa situazione è ragionevole pensare che ogni (singolo) O–ring abbia la stessa "probabilità", diciamo a, di venire estratto cioè:

$$\mathrm{P(un\ singolo\ O\text{–}ring\ di\ essere\ estratto)} = a = \frac{1}{90}$$

in conclusione

$$\begin{gathered}\mathrm{Pr(estrarre\ un\ O\text{–}ring\ di\ diametro\ pari\ a\ 10}mm) = \\ \frac{\text{numero di O–ring con diametro pari a } 10mm}{\text{numero di O–ring presenti nella scatola}} = \frac{30}{90} = \frac{1}{3}\end{gathered}$$

- Rovescio la scatola e considero la guarnizione che tocca per ultima il piano di appoggio

Sotto l'assunzione che le differenze fisiche tra gli O–ring di diversa misura non influiscano sull'esito dell'esperimento vale il ragionamento precedente.

- Analizzo accuratamente le guarnizioni prima di sceglierle in modo da prendere quella con diametro $10mm$

Se ad esempio confronto gli O–ring finché ne trovo 3 di misura diversa e poi scelgo quello con misura intermedia allora è ragionevole pensare che non ci sia aleatorietà (variabilità) nel fenomeno osservato, in futuro esprimeremo questo fatto scrivendo:

$$\mathrm{Pr(estrarre\ un\ O\text{–}ring\ di\ diametro\ pari\ a\ 10}mm) = 1$$

- Estraggo senza guardare una guarnizione alla volta e decido di fermarmi appena ritengo di aver estratto una guarnizione di diametro $10mm$;

Si noti che in questo ultimo caso, il risultato non dipende più dalla composizione della scatola (fatto salvo che vi sia almeno un O–ring di diametro $10mm$, che è il nostro caso) ma da chi esegue l'estrazione. In particolare dalla sua capacità di riconoscere un O–ring di diametro $10mm$ senza poterlo confrontare con gli altri. Quale sarà allora la probabilità?

O–ring – Conclusioni

Dall'esempio che abbiamo visto, possiamo concludere che la probabilità è subordinata alla definizione del problema che intediamo studiare in modo che non siano lasciati dubbi sulla sua interpretazione. Quando un problema così definito non ha un risultato certo parliamo di:

Definizione 1 (Esperimento aleatorio)**.** *Un Esperimento si dice* ***aleatorio****, per un certo individuo, in un certo instante, se l'individuo non è ancora in grado di indicarne con sicurezza il risultato (indipendentemente dal fatto che l'esperimento sia stato già eseguito o debba ancora essere eseguito)*

Se l'individuo che si trova in una tale situazione d'incertezza è interessato al risultato dell'esperimento, è naturale che egli si preoccupi innanzitutto di fissare un "ventaglio completo di eventualità, a due a due incompatibili", ossia un insieme Ω, i cui elementi rappresentino ipotetici risultati dell'esperimento, con la certezza che, comunque vadano le cose, il risultato effettivo dell'esperimento "cadrà in Ω" (nel senso che sarà rappresentato da uno e da un sol elemento di Ω)

Tale insieme Ω prende il nome di **Spazio Campionario** (in Statistica esso prende spesso il nome di *Popolazione*).

Palline Verdi, Bianche, Rosse

Costituisce un *Esperimento aleatorio* l'estrazione di una pallina da un'urna contenente palline dalle medesime caratteristiche fisiche ma contraddistinte da un diverso colore (Verde, Bianco, Rosso) da parte di un individuo bendato.

In questo caso lo Spazio Campionario è costituito dall'insieme $\Omega = \{Verde, Bianco, Rosso\}$

Ancora sull'esperimento aleatorio

Gli Esperimenti aleatori possono essere

- **unici**
- **ripetibili** nelle medesime condizioni.

Anche nel caso in cui l'esperimento sia unico, è spesso utile pensarlo come il risultato di un esperimento ripetibile, riprenderemo questo aspetto quando parleremo del **Principio del Campionamento Ripetuto**.

O–ring e Palline Verdi, Bianche e Rosse

Dal punto di vista teorico, non vi è nessuna differenza tra il problema di assegnare la probabilità di estrarre un O–ring di diametro $10mm$ o di estrarre una pallina Verde da una urna.

Questo fatto accade spesso e consente di sviluppare una teoria generale senza dover fare riferimento al problema specifico.

Eventi e risultati di un Esperimento aleatorio

Consideriamo uno spazio campionario Ω e un sottoinsieme A di Ω (scriveremo $A \subseteq \Omega$). Si può interpretare A come rappresentante un *evento legato al risultato dell'esperimento*: l'evento si realizzerà se e solo se codesto risultato "cadrà in A". (Si può anzi identificare questo evento con l'insieme A stesso).

Eventi e dadi

Consideriamo l'esperimento consista nel lanciare un dado, e per "risultato" s'intende il numero della faccia che uscirà. In questo caso $\Omega = \{1, 2, 3, 4, 5, 6\}$ ed esempi di eventi sono:

- uscita del numero uno: $A_1 = \{1\}$;
- uscita di un numero pari: $A_2 = \{2, 4, 6\}$;
- uscita di un numero più piccolo di 3: $A_3 = \{1, 2\}$;
- uscita di un numero primo: $A_4 : \{1, 2, 3, 5\}$.

Eventi e Insiemi

È facile notare che gli eventi sono (sotto) insiemi dell'insieme Ω e quindi, è conveniente per prima cosa prendere confidenza con il concetto di insieme e con le operazioni sugli insiemi.

- Diremo che x appartiene a un insieme A, e scriveremo $x \in A$ se x è un elemento di A;
- Due insiemi A, B sono uguali se hanno gli stessi elementi, cioè se $x \in A \Leftrightarrow x \in B$, indicheremo tale fatto scrivendo $A = B$ o anche $A \equiv B$.
- Gli insiemi possono essere definiti attraverso proposizioni: $\{x : \text{ proposizione riguardante } x\}$, ad esempio $A = \{x : x \in A\}$.

Eventi e Insiemi

- $\emptyset$ indica l'**insieme vuoto**, cioè l'insieme che non ha nessun elemento: $x \notin \emptyset$ per qualsiasi x. Esso può essere definito da una proposizione che è sempre falsa, ad esempio: $\emptyset = \{x : x \neq x\}$.

- Denotiamo con $A = \{x_1, x_2, \cdots, x_n\} = \{x_i\}_{i=1}^n$ con $n \in \mathbb{N}$ un insieme **finito**. L'insieme $\emptyset$ è considerato un insieme finito.

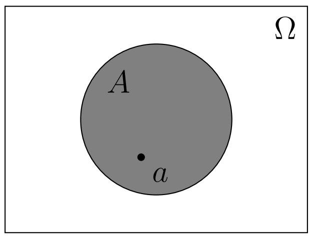
- Indicheremo con $\#A$ il numero di elementi dell'insieme finito A. Per esempio $\#\emptyset = 0$ e $\#\{1, 2, 3\} = 3$.

- Denotiamo con $A = \{x_1, x_2, \cdots\} = \{x_i\}_{i=1}^\infty$ un insieme numerabile (*countable set*) consistente nella succesione di elementi $x_1, x_2, \cdots$.

Insiemi e Diagramma di Venn

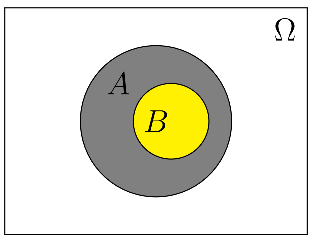
In molte situazioni è comodo rappresentare gli insiemi e le operazioni su di esse attraverso i Diagrammi di Venn.

Ad esempio, dato uno spazio, diciamo Ω, e un insieme A su tale spazio, e un punto $a \in A$ possiamo rappresentare il tutto come segue



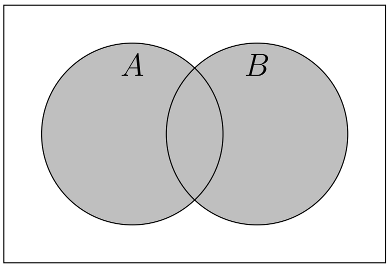
Sottoinsieme

Definizione 2 (Sottoinsieme)**.** *Diremo che B è un sottoinsieme di A e lo indicheremo scrivendo $B \subseteq A$ se $\forall x \in B$ abbiamo $x \in B \Rightarrow x \in A$:*



Unione

Definizione 3 (Unione)**.** *Diremo che C è l'insieme unione di A e B e lo indicheremo $C = A \cup B$ se $\forall x \in \Omega$ abbiamo $x \in A \Rightarrow x \in C$ e $x \in B \Rightarrow x \in C$ cioè $x \in C$ se $x \in A$ o $x \in B$:*



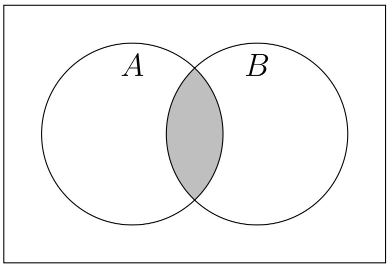
$$C = A \cup B = \{x : x \in A \text{ o } x \in B\}$$

Unione

- Per una successione di insiemi $A_1, A_2, \cdots$ diremo che l'insieme unione
$$\cup_{i=1}^{\infty} A_i = A_1 \cup A_2 \cup \cdots$$
consiste di tutti gli elementi x tali che $x \in A_i$ per qualche $i = 1, 2, \cdots$.
- Più generalmente, per una arbitraria famiglia di insiemi $\mathcal{A}$, l'unione $\cup_{A \in \mathcal{A}} A$ consiste di tutti gli elementi x tali che $x \in A$ per **qualche** $A \in \mathcal{A}$.

Intersezione

Definizione 4 (Intersezione)**.** *Diremo che C è l'insieme intersezione di A e B e lo indicheremo $C = A \cap B$ se $\forall x \in \Omega$ abbiamo $x \in C \Rightarrow \{x \in A$ e $x \in B\}$ cioè $C = \{x : x \in A$ e $x \in B\}$:*



$$C = A \cap B = \{x : x \in A \text{ e } x \in B\}$$

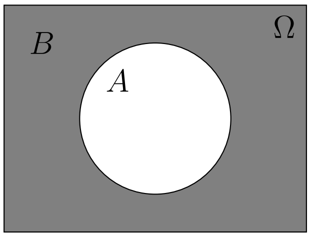
Intersezione

- Per una successione di insiemi $A_1, A_2, \cdots$ diremo che l'insieme intersezione
$$\cap_{i=1}^{\infty} A_i = A_1 \cap A_2 \cap \cdots$$
consiste di tutti gli elementi x tali che $x \in A_i$ per **tutti** $i = 1, 2, \cdots$.

- Più generalmente, per una arbitraria famiglia di insiemi $\mathcal{A}$, l'intersezione $\cap_{A \in \mathcal{A}} A$ consiste di tutti gli elementi x tali che $x \in A$ per tutti $A \in \mathcal{A}$.

Complementare

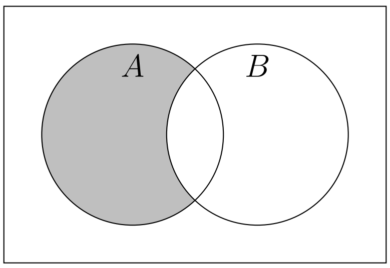
Definizione 5 (Complementare)**.** *Diremo che B è l'insieme complementare di A e lo indicheremo scrivendo $B = A^c$ se $\forall x \in \Omega$ abbiamo $x \in B \Rightarrow x \notin A$ cioè $B = \{x : x \notin A\}$:*



Valgono le seguenti proprietà: $A \cap A^c = \emptyset$ e $A \cup A^c = \Omega$.

Differenza

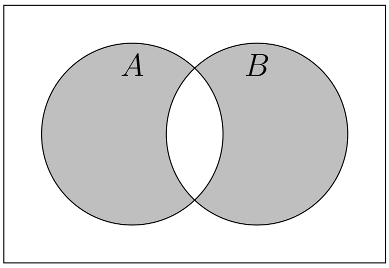
Definizione 6 (Differenza)**.** *Diremo che C è l'insieme differenza di A e B o in altre parole il complementare di B in A e lo indicheremo $C = A/B = \{x : x \in A \text{ e } x \notin B\}$ oppure con $A - B$:*



$$A/B = A \cap B^c$$

Differenza simmetrica

Definizione 7 (Differenza simmetrica)**.** *Diremo che* C *è l'insieme differenza simmetrica di* A *e* B *se* $C = A \triangle B = (A/B) \cup (B/A)$*:*



Prodotto Cartesiano

Definizione 8 (Prodotto Cartesiano)**.** $C = A \times B$ *è detto Prodotto Cartesiano (o prodotto semplice) degli insiemi* A *e* B *e consiste di tutte le coppie ordinate* (x, y) *tale che* $x \in A$ *e* $y \in B$*.*

Il prodotto $A_1 \times A_2 \times \cdots \times A_n$ di n insiemi $A_1, A_2, \cdots, A_n$ consiste delle n-uple ordinate $(x_1, x_2, \cdots, x_n)$ tali che $x_i \in A_i$ per ogni $i = 1, \cdots, n$.

In particolare, $A^n = A \times \cdots \times A$ è il prodotto di n copie dell'insieme A.

Insiemi disgiunti

Definizione 9 (Insiemi disgiunti)**.** *Due insiemi* A *e* B *si dicono disgiunti se* $A \cap B = \emptyset$*.*

Gli insiemi nella famiglia $\mathcal{A}$ è detta disgiunta a coppie se $A \cap B = \emptyset$ per ogni $A, B \in \mathcal{A}$ e $A \neq B$.

Proprietà delle operazioni

- Idempotenza: $A \cap A = A$ e $A \cup A = A$;
- Commutativa: $A \cup B = B \cup A$ e $A \cap B = B \cap A$;
- Associativa: $A \cup B \cup C = A \cup (B \cup C)$ e $A \cap B \cap C = A \cap (B \cap C)$;
- Distributiva (dell'unione rispetto all'intersezione): $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$;
- Distributiva (dell'intersezione rispetto all'unione): $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$;

- Evento impossibile: $A \cap \emptyset = \emptyset$ e $A \cup \emptyset = \emptyset$.
- Spazio campionario: $A \cap \Omega = A$ e $A \cup \Omega = \Omega$.
- Regola di De Morgan:

$$\begin{aligned} (A \cup B)^c &= A^c \cap B^c \ , \\ (A \cap B)^c &= A^c \cup B^c \ , \\ (A^c)^c &= A \ . \end{aligned}$$

  Più in generale, dati gli insiemi $A_1, A_2, \ldots, A_n$, si ha che

$$\begin{aligned} \left(\bigcup_{i=1}^{n} A_i \right)^c &= \bigcap_{i=1}^{n} A_i^c \ , \\ \left(\bigcap_{i=1}^{n} A_i \right)^c &= \bigcup_{i=1}^{n} A_i^c \ , \end{aligned}$$

Esercizio 1. *Rappresentare attraverso i Diagrammi di Venn le proprietà esposte. E dimostrare le proprietà.*

Ma quanti eventi possiamo avere e dadi

Definizione 10 (Insieme delle Parti)**.** *Dato l'insieme Ω si dice **Insieme delle Parti** o **Insieme Potenza** di Ω l'insieme $\mathcal{P}(\Omega)$ di tutti i possibili sottoinsiemi di Ω.*

Quindi, siamo portati a concludere che il numero degli eventi è pari al numero di elementi in $\mathcal{P}(\Omega)$.

Ma quanti eventi possiamo avere e dadi

Nel caso del lancio di un dado gli eventi sono $2^6 = 64$ (vedremo più avanti come mai usando il calcolo combinatorio) alcuni eventi sono:

- uscita del numero due: $A_5 = \{2\}$;
- uscita di un qualsiasi numero: $A_6 = \{1, 2, 3, 4, 5, 6\}$;
- uscita di un numero non meno grande di 4: $A_7 = \{4, 5, 6\}$;
- non esce nessun numero: $A_8 = \emptyset$.

Esercizio 2. *Scrivere tutti i 64 eventi e assegnare ad ogni evento la sua probabilità.*

Ma quanti eventi possiamo avere e dadi

Tuttavia può darsi che certe parti (sottoinsiemi) di Ω corrispondano a eventi non interessanti (ai fini di un determinato problema) oppure troppo complicati per poter essere studiati.

Nel problema del lancio di un dado dove si registra la faccia uscita gli eventi interessanti sono:

- $B_1 = \{1\}$;
- $B_2 = \{2\}$;
- $B_3 = \{3\}$;
- $B_4 = \{4\}$;
- $B_5 = \{5\}$;
- $B_6 = \{6\}$;

ma è sufficiente considerare solo questi eventi?

Teorema 1. *Siano A e B due sottoinsiemi di Ω. Allora ognuna delle seguenti condizioni che segue è vera se e solo se sono vere le rimanenti:*

- $A \subseteq B$*;*
- $A \cap B = A$*;*
- $A \cup B = B$*;*
- $B^c = \Omega / B \subseteq \Omega / A = A^c$*;*
- $A \cap B^c = \emptyset$*;*
- $A^c \cup B = \Omega$*.*

Dimostriamo che la prima implica la seconda e viceversa, il resto è lasciato come esercizio.

Dimostrazione 1. • $A \subseteq B \Rightarrow A \cap B = A$*;*

$A \subseteq B$ *significa che tutti i punti di* A *sono anche in* B *cioè* $x \in A \Rightarrow x \in B$*. Quindi tutti i punti di* A *appartengono all'intersezione.*

Rimane da mostrare che solo questi punti appartengono all'intersezione. A tale scopo consideriamo i punti $x \in B$ *ma* $x \notin A$*: essi non possono appatenere all'intersezione. Rimangono i punti* $x \notin B$*. Ma ovviamente questi non posso appatenere all'intersezione.*

Proposizione 1. • *$A \cap B = A \Rightarrow A \subseteq B$;*

Dimostrazione 2. *$A \cap B = A$ significa che $x \in A \cap B \Leftrightarrow x \in A$ e quindi non esiste $x \in A : x \notin B$ cioè $\{x \in A : x \notin B\} = \emptyset$. Quindi tutti i punti di A sono in B cioè $A \subseteq B$*

Ma quanti eventi possiamo avere e dadi

In generale, non è sufficiente considerare solo questi insiemi come eventi, in ciascun caso e per ciascun problema da studiare, converrà scegliere una determinata classe $\mathcal{A}$ (non vuota), di parti di Ω, e di riservare il nome di **eventi** agli elementi di questa classe.

Solo in casi particolari avremo $\mathcal{A} \equiv \mathcal{P}(\Omega)$.

In ogni caso sarà opportuno scegliere la classe $\mathcal{A}$ in modo tale che essa possieda buone doti di *stabilità*.

Tribù

Un possibile modo di costruire la classe $\mathcal{A}$ è di sceglierla in modo che sia una **Tribù**

Definizione 11 (Tribù (o σ–algebra))**.** *Una classe $\mathcal{A}$ di parti di un insieme Ω si dice una **Tribù** se*

- *(1) $\Omega \in \mathcal{A}$;*
- *(2) Se $A \in \mathcal{A}$ allora $A^c \in \mathcal{A}$;*
- *(3) Se $A_1, \cdots, A_i, \cdots \in \mathcal{A}$ allora $\cup_{i=1}^{\infty} A_i \in \mathcal{A}$.*

Tribù e Eventi

Definizione 12 (Evento)**.** *Ogni elemento A di una Tribù $\mathcal{A}$ su Ω è detto **Evento**.*

Un Evento A si è verificato (si dice anche che l'evento è Vero) se il risultato dell'esperimento aleatorio è un sottoinsieme dell'evento.

Corrispondenza tra logica degli eventi e algebra degli insiemi

Eventi	**Insiemi**
Certo	Universale
Impossibile	Vuoto
Implicazione	Inclusione
Negazione	Complementare
Somma logica	Unione
Prodotto logico	Intersezione
Incompatibilità	Disgiunzione
Classe completa	Partizione

Tribù e algebra

Definizione 13 (Algebra). *Una classe $\mathcal{A}$ di parti di un insieme Ω si dice un'**algebra** se*

- *(1) $\Omega \in \mathcal{A}$;*
- *(2) Se $A \in \mathcal{A}$ allora $A^c \in \mathcal{A}$;*
- *(3) Se $A_1, \cdots, A_i, \cdots, A_n \in \mathcal{A}$ allora $\cup_{i=1}^n A_i \in \mathcal{A}$.*

Esercizio 3. *Data una Tribù (o un'Algebra) $\mathcal{A}$ su Ω mostrare che $\emptyset \in \mathcal{A}$.*

Soluzione 1. *Usando la (1) abbiamo $\Omega \in \mathcal{A}$ e usando la (2) per $A = \Omega$ abbiamo che $\emptyset = \Omega/\Omega \in \mathcal{A}$*

Esercizio 4. *Mostrare che dato un insieme finito Ω, l'insieme $\mathcal{P}(\Omega)$ è un'algebra di Ω.*

Esempi di Tribù

Considerando l'esperimento aleatorio dell'estrazione di un dado e lo spazio campionario $\Omega = \{1, 2, 3, 4, 5, 6\}$ esempi di tribù sono:

- $\mathcal{A}_1 = \{\emptyset, \Omega\}$;
- $\mathcal{A}_2 = \{\emptyset, \{1\}, \{2, 3, 4, 5, 6\}, \Omega\}$;
- $\mathcal{A}_3 = \{\emptyset, \{1\}, \{2\}, \{1, 3, 4, 5, 6, \}, \{2, 3, 4, 5, 6\}, \{1, 2\}, \{3, 4, 5, 6\}, \Omega\}$;
- $\mathcal{A}_4 = \mathcal{P}(\Omega)$.

Esercizio 5. *Data una Tribù (o un'Algebra) $\mathcal{A}$ su Ω mostrare che dati due insiemi $A, B \in \mathcal{A}$ abbiamo $A \cap B \in \mathcal{A}$.*

Soluzione 2. *Dalla Regola di De Morgan abbiamo*

$$A \cap B = [A^c \cup B^c]^c$$

Ora, $A^c, B^c \in \mathcal{A}$ per la (2), $A^c \cup B^c \in \mathcal{A}$ per la (3) e $[A^c \cup B^c]^c \in \mathcal{A}$ per la (2)

Esercizio 6. *Mostrare che dato un insieme finito Ω un'algebra $\mathcal{A}$ su di esso e n insiemi $A_1, A_2, \cdots, A_n \in \mathcal{A}$ l'insieme $\cap_{i=1}^n A_i \in \mathcal{A}$.*

Esercizio 7. *Data una Tribù (o un'Algebra) $\mathcal{A}$ su Ω mostrare che dati due insiemi $A, B \in \mathcal{A}$ abbiamo $A/B \in \mathcal{A}$.*

Soluzione 3. *Possiamo scrivere $A/B = A \cap B^c$, ma $B^c \in \mathcal{A}$ per (2) e $A \cap B^c \in \mathcal{A}$ per l'esercizio precedente*

Esercizio 8. *Data una Tribù (o un'Algebra) $\mathcal{A}$ su Ω mostrare che dati due insiemi $A, B \in \mathcal{A}$ abbiamo $A \triangle B \in \mathcal{A}$.*

Teorema 2. *Date due Algebre $\mathcal{A}_1, \mathcal{A}_2$ di sottoinsiemi di Ω allora $\mathcal{A} = \mathcal{A}_1 \cap \mathcal{A}_2$ è un'algebra su Ω.*

Dimostrazione 3.

- *Siccome $\Omega \in \mathcal{A}_1$ e $\Omega \in \mathcal{A}_2$ allora $\Omega \in \mathcal{A}$*
- *Dato un insieme $A \in \mathcal{A}$ allora $A \in \mathcal{A}_1$ e $A \in \mathcal{A}_2$ ma questo per (2) implica che $A^c \in \mathcal{A}_1$ e $A^c \in \mathcal{A}_2$ cioè $A^c \in \mathcal{A}$*
- *Dati due insiemi $A, B \in \mathcal{A}$ allora $A, B \in \mathcal{A}_1$ e $A, B \in \mathcal{A}_2$ inoltre per la (3) abbiamo $A \cup B \in \mathcal{A}_1$ e $A \cup B \in \mathcal{A}_2$ e quindi $A \cup B \in \mathcal{A}$*

Esercizio 9. *Mostrare che lo stesso risultato vale per una Tribù.*

Esercizio 10. *Trovare due Algebre la cui unione non è un'Algebra.*

Soluzione 4. *Consideriamo $\Omega = \{1,2,3\}$, $\mathcal{A}_1 = \{\emptyset, \Omega, \{1\}, \{2,3\}\}$ e $\mathcal{A}_2 = \{\emptyset, \Omega, \{2\}, \{1,3\}\}$.*

Allora $\mathcal{A} = \mathcal{A}_1 \cup \mathcal{A}_2 = \{\emptyset, \Omega, \{1\}, \{2\}, \{1,3\}, \{2,3\}\}$ la quale non è un'algebra perché

$$\{1\} \cup \{2\} = \{1,2\} \notin \mathcal{A}$$

Lo stesso vale per una Tribù.

Esercizio 11. $\mathcal{A} = \{A \subseteq \Omega : A$ *è un insieme finito*$\}$ *è sempre un'Algebra?*

Soluzione 5. *No, ad esempio quando Ω è un'insieme non finito esso non appartiene a $\mathcal{A}$ e la (1) non è soddisfatta*

Esercizio 12. *Sia Ω un insieme. Mostrare che la famiglia $\mathcal{A}$ di tutti i sottoinsiemi finiti di Ω e di tutti gli insiemi che hanno un complementare finito è un'algebra.*

Soluzione 6. *La (1) è soddisfatta perchè il complementare di Ω è finito ($\emptyset$) e quindi $\Omega \in \mathcal{A}$.*

Per ogni $A \in \mathcal{A}$ abbiamo

- *se A è finito allora $A^c \in \mathcal{A}$;*
- *se A non è finito allora il suo complementare A^c deve essere finito e quindi $A^c \in \mathcal{A}$;*

e la (2) è soddisfatta.

Infine sia $A, B \in \mathcal{A}$.

- *se A, B sono entrambi finiti allora $A \cup B$ è finito;*
- *se almeno uno dei due, diciamo A non è finito allora A^c è finito ed è finito pure $(A \cup B)^c$ perché $(A \cup B)^c \subseteq A^c$ (siccome $A \subseteq (A \cup B)$).*

Tribù e Algebra

Definizione 14. *Un'Algebra si dice* ***finita*** *se è costituita da un numero finito di elementi.*

Teorema 3. *Sia $\mathcal{A}$ un'Algebra finita su Ω allora è anche una Tribù.*

Dimostrazione 4. *Le prime due condizioni sono ovviamente verificate, rimane da verificare la (3). A questo scopo consideriamo una successione $A_1, \cdots, \in \mathcal{A}$. Siccome $\mathcal{A}$ è finita allora questa successione contiene un numero finito di insiemi distinti (k), $A_{i_1}, \cdots, A_{i_k}$ e quindi*

$$\cup_{m=1}^{\infty} A_m = \cup_{l=1}^{k} A_{i_l} \in \mathcal{A}$$

perchè l'unione non cambia se noi aggiungiamo un numero arbitrario di insiemi già inclusi

Tribù generata

Definizione 15 (Tribù generata a partire da una famiglia di insiemi)**.** *Data una famiglia $\mathcal{F}$ di sottoinsiemi di Ω poniamo*

$$\mathcal{A}_{\mathcal{F}} = \cap\{\mathcal{A} : \mathcal{A} \text{ è una Tribù e } \mathcal{F} \subseteq \mathcal{A}\}$$

chiamiamo $\mathcal{A}_{\mathcal{F}}$ la Tribù generata da $\mathcal{F}$.

Esempio 1. *Sia*

$$\mathcal{F} = \left\{ \{0,1\}, [\frac{1}{2^{n+1}}, \frac{1}{2^n}) : n = 0,1,2,\cdots \right\}$$

Dire quale dei seguenti insiemi appartiene a $\mathcal{A}_{\mathcal{F}}$*: (1)* $\{0\}$*, (2)* $\{1\}$*, (3)* $\{1/2\}$*, (4)* $\{1/3\}$*, (5)* $\{[0,1]\}$*, (6)* $\{(1/4,1]\}$*, (7)* $\{[0,1/2]\}$*, (8)* $\{[1/4,1)\}$*, (9)* $\{(0,1/2)\}$*.*

Soluzione 7. *Notiamo che tutti gli insiemi che compongono la famiglia sono digiunti e che*

- *per* $n = 0$ *abbiamo l'intervallo* $[1/2, 1)$
- *mentre per* $n = 1$ *abbiamo* $[1/4, 1/2)$
- *dalla loro unione abbiamo che (8) è in* $\mathcal{A}_{\mathcal{F}}$
- *Il (5) è in* $\mathcal{A}_{\mathcal{F}}$ *perché* $\{[0,1]\} = \{0,1\} \cup \cup_{n=0}^{\infty} \left\{ [\frac{1}{2^{n+1}}, \frac{1}{2^n}) \right\}$
- *infine (9) è in* $\mathcal{A}_{\mathcal{F}}$ *perché* $\{(0,1/2)\} = \{[0,1]\}/([1/2,1) \cup \{0,1\})$

I rimanenti insiemi non fanno parte di $\mathcal{A}_{\mathcal{F}}$*.*

Limite di successioni

Sia $(A_n)_{n=1}^{\infty}$ una successione di insiemi.

Definizione 16 (Limite superiore)**.** *Si definisce* ***limite superiore*** *di* A_n *l'insieme dato da*

$$\text{limsup} A_n = \cap_{n=1}^{\infty} \cup_{m=n}^{\infty} A_m = \cap_{n=1}^{\infty} \cup_{m=n}^{\infty} \{\omega \in \Omega; \omega \in A_m\}$$

Tale insieme è costituito da tutti i punti $\omega \in \Omega$ che appartengono ad A_n per **infiniti indici** n. A rappresentare detto insieme si usa anche la scrittura alternativa $\{\omega : \omega \in A_n \text{i.o}\}$ (i.o. = infinitly often) cioè $\forall n \exists m \geq n : \omega \in A_m$.

Limite di successioni

Sia $(A_n)_{n=1}^{\infty}$ una successione di insiemi.

Definizione 17 (Limite inferiore)**.** *Si definisce* ***limite inferiore*** *di* A_n *l'insieme dato da*

$$\text{liminf} A_n = \cup_{n=1}^{\infty} \cap_{m=n}^{\infty} A_m = \cup_{n=1}^{\infty} \cap_{m=n}^{\infty} \{\omega \in \Omega; \omega \in A_m\}$$

Tale insieme è costituito da tutti i punti $\omega \in \Omega$ che appartengono a tutti gli insiemi A_n **tranne che a un numero finito di esso**. A rappresentare detto insieme si usa anche la scrittura alternativa $\{\omega : \omega \in A_n \text{a.e.}\}$ (a.e. = almost everywhere) cioè $\exists n \forall m \geq n : \omega \in A_m$.

Limite di successioni

Sia $(A_n)_{n=1}^{\infty}$ una successione di insiemi.

Definizione 18 (Limite). *Se*

$$\text{limsup} A_n = \text{liminf} A_n$$

*allora si dice che la successione ammette **limite** e lo si indica con* $\lim A_n$

Limite di successioni e proprietà

Teorema 4. *Sia* $(A_n)_{n=1}^{\infty}$ *una successione di insiemi. Allora*

$$\text{liminf} A_n \subseteq \text{limsup} A_n$$

Se A_n accade per tutti gli n tranne che per un numero finito di essi allora accade per un numero infinito di n

Nota: quindi per dimostrare che il limite di una successione esiste è sufficiente mostrare che $\text{limsup} A_n \subseteq \text{liminf} A_n$.

Limite di successioni e proprietà

Esercizio 13. *Sia* $(A_n)_{n=1}^{\infty}$ *una successione di insiemi da una Tribù* $\mathcal{A}$. *Allora*

- $\text{liminf} A_n \in \mathcal{A}$
- $\text{limsup} A_n \in \mathcal{A}$

Soluzione 8. *Dalla definizione di limite inferiore e superiore è sufficiente mostrare che*

$$\cup_{m=n}^{\infty} A_m \in \mathcal{A} \ e \ \cap_{m=n}^{\infty} A_m \in \mathcal{A} \qquad \forall n = 1, 2, \cdots$$

Limite di successioni e proprietà

Esercizio 14. *Sia* $(A_n)_{n=1}^{\infty}$ *una successione di insiemi da una Tribù* $\mathcal{A}$. *Allora*

- $(\text{liminf} A_n)^c = \text{limsup} A_n^c$
- $(\text{limsup} A_n)^c = \text{liminf} A_n^c$

Soluzione 9. *Infatti, per la Regola di De Morgan, abbiamo*

$$\begin{aligned} (\text{liminf} A_n)^c = & \quad (\cup_{n=1}^{\infty} \cap_{m=n}^{\infty} A_m)^c \\ = & \quad \cap_{n=1}^{\infty} (\cap_{m=n}^{\infty} A_m)^c \\ = & \quad \cap_{n=1}^{\infty} \cup_{m=n}^{\infty} A_m^c = \text{limsup} A_n^c \end{aligned}$$

e in maniera analoga per l'altro caso.

Limite di successioni e proprietà

Definizione 19 (Successione non decrescente)**.** *Se una successione di insiemi* $(A_n)_{n=1}^{\infty}$ *è tale che* $A_1 \subseteq A_2 \subseteq \cdots$ *allora la succesione è detta* ***non decrescente****; si scrive allora* $A_n \uparrow$*.*

Definizione 20 (Successione non crescente)**.** *Se una successione di insiemi* $(A_n)_{n=1}^{\infty}$ *è tale che* $A_1 \supseteq A_2 \supseteq \cdots$ *allora la succesione è detta* ***non crescente****; si scrive allora* $A_n \downarrow$*.*

Teorema 5. *Una successione* $(A_n)_{n=1}^{\infty}$ *non descrescente (o non crescente) ammette limite e risulta*

- $\lim A_n = \cup_{n=1}^{\infty} A_n$ *se* $A_n \uparrow$
- $\lim A_n = \cap_{n=1}^{\infty} A_n$ *se* $A_n \downarrow$

Limite di successioni

Esercizio 15. *Trovare il limite inferiore e superiore della successione*

$$A_n = \begin{cases} \left(\frac{1}{3} - \frac{1}{n+1}, 1 + \frac{1}{n+1}\right) & se\ n = 1, 3, 5, \cdots \\ \left(\frac{1}{n+1}, \frac{2}{3} - \frac{1}{n+1}\right) & se\ n = 2, 4, 6, \cdots \end{cases}$$

Soluzione 10. *La prima successione di intervalli (n dispari) converge all'intervallo* $[\frac{1}{3}, 1]$ *mentre la seconda successione di intervalli (n pari) converge all'intervallo* $(0, \frac{2}{3})$*.*

Quindi

- $\liminf A_n = [\frac{1}{3}, \frac{2}{3})$
- $\limsup A_n = (0, 1]$

Spazio Probabilizzabile

Definizione 21 (Spazio Probabilizzabile)**.** *Dato un spazio campionario* Ω *e una tribù* $\mathcal{A}$ *su* Ω*, la coppia* $(\Omega, \mathcal{A})$ *è detto* ***Spazio Probabilizzabile****.*

Nella scelta di $(\Omega, \mathcal{A})$ c'è sempre una dosa di *arbitrarietà* e ciò per due ragioni principali:

- gli elementi di Ω rappresentano ipotetici risultati dell'esperimento, secondo un opportuno *codice*: è chiaro che la scelta di tale codice è spesso arbitraria.

Se l'esperimento consiste nel lancio di una moneta, e se ci s'interessa solo alla faccia che apparirà (testa o croce), si potrà prendere $\Omega = \{0, 1\}$, con la convenzione che 0 significhi croce e 1 testa. Ma equalmente legittima è la convenzione inversa (0 =testa, 1 =croce)

- Una volta fissato l'insieme Ω, la scelta della Tribù è arbitraria perché dipende dal criterio in base al quale si decide di considerare certi eventi come interessanti e altri no.

  Ritornando all'esperimento del lancio di un dado, fissato $\Omega = \{1, 2, 3, 4, 5, 6\}$ se il nostro interesse è nel solo esito è uscita la faccia con il numero 1, allora è sufficiente usare la Tribù $\{\emptyset, \Omega, \{1\}, \{1\}^c\}$. In questo caso però possiamo anche usare il seguente spazio $\Omega' = \{0, 1\}$ dove 0 indica che non è uscita la faccia 1 e 1 invece che è uscita. In questo caso la Tribù diventa $\{\emptyset, \Omega', \{0\}, \{1\}\}$.

Probabilità

Definizione 22 (Probabilità)**.** *Dato un spazio probabilizzabile* $(\Omega, \mathcal{A})$*, una* ***Probabilità*** $\Pr$ *è un'applicazione* $\Pr : \mathcal{A} \to \mathbb{R}^+$ *tale che*

- *(1) (non negatività) se* $A \in \mathcal{A}$ *allora* $\Pr(A) \geq 0$*;*

- *(2) (normalizzazione)* $\Pr(\Omega) = 1$*;*

- *(3) (σ–additività) Se* $\{A_i\}_{i=1}^{\infty}$ *è una successione di eventi di* $\mathcal{A}$ *a due a due incompatibili (cioè* $A_i \cap A_j = \emptyset, i \neq j$*), allora*

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

Questa è la definizione assiomatica di Kolmogorov.

Spazio di Probabilità

Definizione 23 (Spazio di Probabilità)**.** *La terna* $(\Omega, \mathcal{A}, \Pr)$ *dove* Ω *è uno spazio campionario,* $\mathcal{A}$ *è una Tribù su* Ω *e* $\Pr$ *è una funzione di Probabilità* $\Pr : \mathcal{A} \to \mathbb{R}^+$*, è detta* ***Spazio di Probabilità*** *o anche Spazio di Kolmogorov.*

Sulla probabilità di uscita di una faccia di un dado

Consideriamo l'esperimento aleatorio del lancio di un dado regolare. È ragionevole considerare lo spazio probabilizzabile dato da $\Omega = \{1, 2, 3, 4, 5, 6\}$ e $\mathcal{A} = \mathcal{P}(\Omega)$. Resta da determinare la probabilità Pr. Per la natura del problema si può supporre che i possibili risultati (le faccie) si presentino con uguale probabilità, cioè

$$\Pr(\{1\}) = \Pr(\{2\}) = \Pr(\{3\}) = \Pr(\{4\}) = \Pr(\{5\}) = \Pr(\{6\}) = p$$

Il numero p è determinato dal fatto che

$$\begin{aligned} 1 = &\quad \Pr(\Omega) \\ = &\quad \Pr(\{1\} \cup \{2\} \cup \{3\} \cup \{4\} \cup \{5\} \cup \{6\}\} \\ = &\quad \Pr(\{1\}) + \Pr(\{2\}) + \Pr(\{3\}) + \Pr(\{4\}) + \Pr(\{5\}) \\ = &\quad 6\ p \end{aligned}$$

da cui

$$p = \frac{1}{6}$$

Siamo ora in grado di calcolare la probabilità di tutti gli eventi. Ad esempio se $A = \{2, 4, 6\}$ (l'evento è uscita una faccia con un numero pari) allora

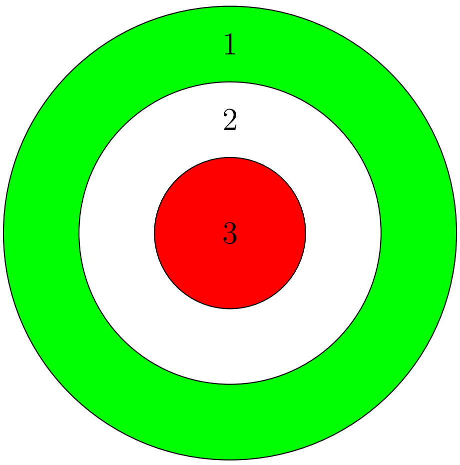
$$\Pr(A) = \Pr(\{2\}) + \Pr(\{4\}) + \Pr(\{6\}) = \frac{1}{2}$$

e più in generale per ogni evento A sarà

$$\Pr(A) = \frac{\#A}{\#\Omega} = \frac{\#A}{6}$$

Sulla probabilità di colpire una corona circolare

Qual'è la probabilità di colpire una corona circolare?



Lo spazio campionario può essere scelto uguale a $\Omega = \{1, 2, 3\}$, cioè i punteggi che identificano le corone (si noti che si assume che sicuramente venga colpito il tabellone!). La Tribù può essere quella generata a partire dalla famiglia $\mathcal{F} = \{\{1\}, \{2\}, \{3\}\}$. Infine per determiante la funzione di probabilità è ragionevole considerare il fatto che la probabilità sia proporzionale all'area della corona. Per cui (dato r il raggio del tabellone)

$$\Pr(\{1\}) = \frac{\pi r^2 - \pi \left(\frac{2}{3}\ r\right)^2}{\pi\ r^2} = 1 - \frac{2}{3}^2 = \frac{5}{9}$$

$$\Pr(\{2\}) = \frac{\pi \left(\frac{2}{3}\ r\right)^2 - \pi \left(\frac{1}{3}\ r\right)^2}{\pi\ r^2} = \frac{2}{3}^2 - \frac{1}{3}^2 = \frac{3}{9}$$

$$\Pr(\{3\}) = \frac{\pi \left(\frac{1}{3}\ r\right)^2}{\pi\ r^2} = \frac{1}{3}^2 = \frac{1}{9}$$

e in generale

$$\Pr(\{n\}) = \frac{(4-n)^2 - (3-n)^2}{3^2} \qquad n = 1, 2, 3$$

Si noti che il risultato non dipende dalla dimensione del tabellone, cioè da r. Inoltre è facile notare che ogni evento in $\mathcal{A}$ è ottenibile come unione degli eventi $\{1\}, \{2\}, \{3\}$, per cui il calcolo della sua probabilità è semplice.

Regole di Calcolo

Regola 1. *Se A è un evento di probabilità $\Pr(A)$ allora la probabilità che A* ***non*** *si verifichi è*

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

Dimostrazione 5. *Poiché $\Omega = A \cup A^c$ e $A \cap A^c = \emptyset$*

- *dalla (3) segue che* $\Pr(\Omega) = \Pr(A) + \Pr(A^c)$
- *e quindi per (2)* $1 = \Pr(A) + \Pr(A^c)$

In particolare questo fatto implica che $\Pr(\emptyset) = 0$.

Regola 2. *Se A e B sono due eventi, allora la probabilità che se ne verifichi almeno uno è data da*

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

Dimostrazione 6. *Decomponiamo i due eventi $A \cup B$ e B secondo*

- $A \cup B = A \cup (B \cap A^c)$ *con* A *e* $B \cap A^c$ *incompatibili;*
- $B = (A \cap B) \cup (B \cap A^c)$ *con* $A \cap B$ *e* $B \cap A^c$ *incompatibili;*

Dall'assioma (3) abbiamo

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

e sottraendo membro a membro

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

Una conseguenza è

$$\Pr(A \cup B) \leq \Pr(A) + \Pr(B)$$

perché $\Pr(A \cap B) \geq 0$ per (1).

Più in generale abbiamo

$$\Pr(\cup_{i=1}^{n} A_i) \leq \sum_{i=1}^{n} \Pr(A_i)$$

Regola 3. *Se* A *è un evento che implica l'evento* B*, cioè se* $A \subseteq B$*, allora*

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

Dimostrazione 7. *Poiché possiamo scrivere* $B = A \cup (B \cap A^c)$ *e* A *e* $B \cap A^c$ *sono incompatibili allora per (3) abbiamo* $\Pr(B) = \Pr(A) + \Pr(B \cap A^c)$*. Inoltre per (1) e il fatto che* $B \cap A^c$ *è un evento abbiamo* $\Pr(B \cap A^c) \geq 0$ *da cui il risultato.*

Questa regola ha due importanti conseguenze

- $\Pr(A) \leq 1$, (basta porre $B = \Omega$)
- Se A e B sono due eventi equivalenti, cioè $A \subseteq B$ e $B \subseteq A$ allora $\Pr(A) = \Pr(B)$

Regola 4 (Disuguaglianza di Bonferroni)**.** *Se* $A_1, A_2, \cdots, A_n$ *sono eventi allora*

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

Dimostrazione 8. *Basta mostrare la disuguaglianza di sinistra. (Dimostrazione per induzione, noi vediamo soltanto l'inizio). Per $n = 1$ è banalmente verificata. Essa è vera per $n = 2$ in quanto*

$$\Pr(A_1) + \Pr(A_2) - \Pr(A_1 \cap A_2) = \Pr(A_1 \cup A_2)$$

Se $n = 3$ si ha

$$\begin{aligned}
&= \Pr(A_1 \cup A_2 \cup A_3) \\
&= \Pr(A_1 \cup A_2) + \Pr(A_3) - \Pr(A_3 \cap (A_1 \cup A_2)) \\
&= \Pr(A_1) + \Pr(A_2) - \Pr(A_1 \cap A_2) + \Pr(A_3) - \Pr((A_3 \cap A_1) \cup (A_3 \cap A_2)) \\
&= \sum_{i=1}^{3} \Pr(A_i) - \Pr(A_1 \cap A_2) \\
&- [\Pr(A_3 \cup A_1) + \Pr(A_2 \cup A_3) - \Pr(A_1 \cap A_2 \cap A_3)] \\
&= \sum_{i=1}^{3} \Pr(A_i) - \sum_{1 \leq i \leq j \leq 3} \Pr(A_i \cap A_j) + \Pr(A_1 \cap A_2 \cap A_3)
\end{aligned}$$

Limiti e Probabilità

Teorema 6.

- *Se $A_1 \subseteq A_2 \subseteq \cdots \in \mathcal{A}$ è una successione di eventi non decrescente e $A = \lim A_n = \cup_{n=1}^{\infty} A_n$ allora*

$$\Pr(A) = \Pr(\lim A_n) = \lim \Pr(A_n)$$

- *Se $A_1 \supseteq A_2 \supseteq \cdots \in \mathcal{A}$ è una successione di eventi non crescente e $A = \lim A_n = \cap_{n=1}^{\infty} A_n$ allora*

$$\Pr(A) = \Pr(\lim A_n) = \lim \Pr(A_n)$$

Dimostrazione 9. *Poiché $A = \cup_{n=1}^{\infty} A_n = \cup_{n=1}^{\infty} A_n / A_{n-1} = \cup_{n=1}^{\infty} (A_n \cap A_{n-1}^c)$ (avendo definito $A_0 = \emptyset$) ed è $A_{n-1} \subseteq A_n$, $n = 1, 2, \cdots$, gli eventi $A_n \cap A_{n-1}^c$ sono incompatibili e dunque, per la σ–additività e per la prima parte della Regola 3*

$$\begin{aligned}
\Pr(A) &= \Pr\left(\cup_{n=1}^{\infty} (A_n / A_{n-1})\right) \\
&= \sum_{n=1}^{\infty} \Pr(A_n / A_{n-1}) \\
&= \sum_{n=1}^{\infty} (\Pr(A_n) - \Pr(A_{n-1}))
\end{aligned}$$

ma

$$
\begin{aligned}
&= \sum_{n=1}^{\infty} (\Pr(A_n) - \Pr(A_{n-1})) \\
&= \lim_{N\to\infty} \sum_{n=1}^{N} (\Pr(A_n) - \Pr(A_{n-1})) \\
&= \lim_{N\to\infty} \Pr(A_N)
\end{aligned}
$$

Per la seconda parte, si considerano i complementari in modo che $A_1^c \subseteq A_2^c \subseteq \cdots$, *abbiamo*

$$\lim_{n\to\infty} A_n^c = \cup_{n=1}^{\infty} A_n^c = (\cap_{n=1}^{\infty} A_n)^c = A^c$$

allora

$$\lim_{n\to\infty} \Pr(A_n^c) = \Pr(\lim_{n\to\infty} A_n^c) = \Pr(A^c)$$

e quindi

$$
\begin{aligned}
&= \lim_{n\to\infty} \Pr(A_n) = \lim_{n\to\infty} (1 - \Pr(A_n^c)) = \\
&= 1 - \lim_{n\to\infty} \Pr(A_n^c) = 1 - \Pr(\lim_{n\to\infty} A_n^c) \\
&= 1 - \Pr(A^c) = \Pr(A)
\end{aligned}
$$

Costruzione della funzione di Probabilità

Adesso ci occupiamo del modo in cui definire la funzione Pr nel caso che Ω sia al più numerabile e si consideri come Tribù $\mathcal{A} = \mathcal{P}(\Omega)$. Supponiamo che $\Omega = \{\omega_1, \omega_2, \cdots\}$. Ad ogni evento $\omega_i \in \Omega$ si assegna un peso $p(\omega_i)$, $i = 1, 2, \cdots$ in modo che

$$p(\omega_i) \geq 0 i = 1, 2, \cdots ; \qquad \sum_{i=1}^{\infty} p(\omega_i) = 1$$

Per ogni evento $A \in \mathcal{A}$ si definisce la sua probabilità come somma di tutti i pesi relativi agli eventi ω_i che appartengono ad A

$$\Pr(A) = \sum_{\omega_i \in A} p(\omega_i) = \sum_{\omega_i \in A} \Pr(\{\omega_i\}) \ .$$

È facile vedere (provate come esercizio) che il procedimento proposto porta alla definizione di una funzione di Probabilità coerente con gli assiomi enunciati.

Quando Ω non è finito né numerabile, il problema della costruzione della funzione di Probabilità diviene più complesso. Poiché uno dei più importanti insiemi non numerabili è quello dei numeri reali il problema è assai rilevante. Lo affronteremo un po' più avanti.

Sulla probabilità di avere testa

Esercizio 16. *Una moneta è lanciata ripetutamente fino a che non compare una testa. Il numero di croci prima che compaia una testa può essere* $0, 1, 2, \cdots$. *Questo fatto può essere descritto con uno spazio di probabilità* $(\Omega, \mathcal{A}, \Pr)$ *così costituito*

- $\Omega = \{0, 1, 2, \cdots\}$;
- $\mathcal{A} = \mathcal{P}(\Omega)$;
- $\Pr(\{n\}) = \frac{1}{2^{n+1}}$ *che è la probabilità di avere n croci.*

Vogliamo calcolare la probabilità dell'evento $\{n \textit{ sia pari}\}$.

Soluzione 11. *Siccome abbiamo scelto* $\mathcal{A} = \mathcal{P}(\Omega)$ *allora* $\{n \textit{ sia pari}\}$ *è un evento. Inoltre*

$$\{n \textit{ sia pari}\} = \{0\} \cup \{2\} \cup \{4\} \cup \{6\} \cup \cdots$$

allora

$$\Pr(\{n \textit{ sia pari}\}) = \Pr(\{0\} \cup \{2\} \cup \{4\} \cup \{6\} \cup \cdots)$$

A questo scopo è più semplice considerare la seguente probabilità

$$\Pr(\{n \textit{ sia dispari}\}) = \Pr(\{1\} \cup \{3\} \cup \{5\} \cup \cdots)$$

infatti,

$$\begin{aligned}\Pr(\{n \textit{ sia dispari}\}) &= \frac{1}{2^2} + \frac{1}{2^4} + \frac{1}{2^6} + \frac{1}{2^8} + \cdots \\ &= \frac{1}{1-q} - 1 = \frac{4}{3} - 1 = \frac{1}{3}\end{aligned}$$

con $q = 1/4$, *perché è una progressione geometrica di ragione* $q = 1/4$ *dove manca il primo termine (1). Siccome* $\{n \textit{ sia pari}\} = \Omega/\{n \textit{ sia dispari}\}$ *abbiamo*

$$\Pr(\{n \textit{ sia pari}\}) = 1 - \frac{1}{3} = \frac{2}{3}$$

Prodotto di una famiglia di Tribù

Definizione 24. *Data una qualsiasi famiglia* $\{(\Omega_i, \mathcal{A}_i)\}_{i \in I}$ *di spazi probabilizzabili si considerino tutte le famiglie* $\{A_i\}_{i\in I}$ *con* $A_i \in \mathcal{A}_i$ *per ogni indice i, e* $A_i \neq \Omega_i$ *al più per un numero finito di indici. Per ogni siffatta famiglia, si consideri il "rettangolo"* $\prod_{i\in I}$. *La Tribù generata, sull'insieme* $\prod_{i\in I} \Omega_i$, *dalla classe di tutti questi rettangoli si chiama la Tribù prodotto della famiglia di Tribù* $\{\mathcal{A}_i\}_{i\in I}$, *e si denota con* $\otimes_{i\in I}\mathcal{A}_i$. *Quando sia* $I = \{1, 2, \cdots, n\}$, *si scrive più semplicemente* $\mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_n$.

Prodotto di una famiglia di Spazi probabilizzati

Definizione 25. *Sia data una famiglia* $\{(\Omega_i, \mathcal{A}_i, \mathrm{Pr}_i)\}_{i\in I}$ *di spazi probabilizzati. Si dimostra che, sulla Tribù prodotto* $\otimes_{i\in I}\mathcal{A}_i$*, esiste un'unica funzione di probabilità* Pr *che verifichi la relazione*

$$\mathrm{Pr}\left(\prod_{i\in I} A_i\right) = \prod_{i\in I} \underset{i}{\mathrm{Pr}}(A_i)$$

per ogni famiglia $\{A_i\}_{i\in I}$ *(come definita al punto precedente). Questa probabilità si chiama probabilità prodotto della famiglia* $\{\mathrm{Pr}_i\}_{i\in I}$ *e si denota con* $\otimes_{i\in I}\mathrm{Pr}_i$*. Quando sia* $I = \{1, 2, \cdots, n\}$*, si scrive più semplicemente* $\mathrm{Pr}_1 \otimes \cdots \otimes \mathcal{P}_n$*.*

Esercizio 17 (Quale prodotto?)**.** *In un esperimento aleatorio si è chiesto a tre signore, scelte casualmente, se utilizzano un certo prodotto.*

1. *Elencare gli elementi dello spazio campionario, usando le lettere* S *per "sì" e* N *per "no";*
2. *Usando l'algebra* $\mathcal{P}(\Omega)$ *elencare i risultati che implicano l'evento* $E_1 =$*"almeno due donne usano il prodotto";*
3. *Definire l'evento i cui elementi sono:* $E_2 = \{(S, S, S), (N, S, S), (S, S, N), (N, S, N)\}$.

Esercizio 18 (Probabilità di eventi)**.** *Dati due sottoinsiemi* A *e* B *di uno spazio campionario* Ω*, con* $\mathrm{Pr}(A^c) = 0.3$*,* $\mathrm{Pr}(B) = 0.4$ *e* $\mathrm{Pr}(A \cap B^c) = 0.5$.

1. *Definire in maniera appropriata, per studiare il problema, lo spazio campionario* Ω *e un'algebra* $\mathcal{A}$ *su di esso;*

Si determinino le probabilità

2. $\mathrm{Pr}(A)$*;*
3. $\mathrm{Pr}(A \cap B)$*;*
4. $\mathrm{Pr}(A \cup B)$.

Esercizio 19 (Mazzo di carte)**.** *Da un mazzo di 52 carte se ne sceglie una in modo casuale.*

1. *Qual è la probabilità di estrarre una figura qualsiasi o una carta di fiori?*
2. *E di estrarre una figura di fiori?*

Esercizio 20 (Esami da superare). *Uno studente deve sostenere due esami E_1 e E_2. Sia A l'evento che egli superi l'esame E_1 e B l'evento che superi l'esame E_2. Si ammetta che* $\Pr(A) = 0.8$, $\Pr(B) = 0.9$ *e* $\Pr(A \cap B) = 0.75$.

1. *Si definisca lo spazio campionario Ω;*
2. *Si definisca un'algebra adatta a studiare i risultati interessanti dell'esperimento;*
3. *Si rappresentino gli eventi con i diagrammi di Venn;*
4. *Si descrivano verbalmente gli eventi: $A \cup B$, $A^c \cup B^c$, $A \cap B^c$, $A^c \cap B^c$, $(A \cup B)^c$;*
5. *Si calcolino le probabilità degli eventi di cui al punto precedente.*

Esercizio 21 (Tempo libero). *Da un'indagine svolta presso una certa scuola è emerso che nel tempo libero il 10% degli studenti studia musica, il 20% pratica sport, il 5% studia una lingua straniera. Inoltre il 5% studia musica e pratica anche uno sport, il 3% studia musica e una lingua straniera, il 2% studia una lingua e fa sport e l'1% fa tutte tre le cose. Scegliendo in modo casuale uno studente,*

1. *qual è la probabilità che pratichi solo sport?*
2. *Che studi musica e una lingua ma non pratichi nessuno sport?*

Esercizio 22 (Due urne). *Un'urna contiene due palle nere e una rossa. Una seconda urna ne contiene una bianca e due rosse. Si estrae a caso una palla da ciascuna urna.*

1. *Descrivete uno spazio campionario per quest'esperimento;*
2. *Definite un'algebra adatta a studiare gli eventi descritti nei punti successivi;*
3. *Descrivete l'evento "prima pallina nera";*
4. *Qual è la probabilità che entrambe le palline siano dello stesso colore?*
5. *E che siano di colore diverso?*

Esercizio 23 (Moneta equilibrata). *Una moneta equilibrata viene lanciata 3 volte consecutivamente.*

1. *Qual è la probabilità che si presenti esattamente una testa?*

2. *Qual è la probabilità che si presenti almeno una testa?*

Esercizio 24 (Due dadi, ancora!). *Si lanciano due dadi a sei facce bilanciati.*

1. *Si definisca per ogni domanda che segue uno spazio probabilizzato adatto a studiare i risultati di interesse;*
2. *Qual è la probabilità che la somma dei risultati sia un numero pari?*
3. *E che sia uguale a 5?*
4. *Qual è la probabilità che la differenza in modulo fra i due risultati sia uguale a 3?*

1.2 Calcolo Combinatorio

Calcolo Combinatorio

Il Calcolo combitanorio, all'inizio, si interessava principalmente all'enumerazione di permutazioni, combinazioni e partizioni di un insieme finito sotto varie condizioni. Oggi il Calcolo combinatorio è una parte della matematica discreta che studia l'esistenza, la costruzione, l'enumerazione e esamina le proprietà di configurazioni che soddisfano delle particolari condizioni.

Siano $S_1, S_2, \cdots, S_r$, r insiemi di cardinalità $n_1, n_2, \cdots, n_r$ formati da oggetti distinti e si consideri il prodotto cartesiano

$$\Omega_r = S_1 \times S_2 \times \cdots \times S_r = \{(s_1, s_2 \cdots, s_r) : s_1 \in S_1, s_2 \in S_2, \cdots s_r \in S_r\} \ .$$

Gli elementi di Ω_r sono evidentemente gli *allineamenti* $(s_1, s_2, \cdots, s_r)$ con $s_i \in S_i$, $i = 1, 2, \cdots, r$.

Quante r–uple $(s_1, s_2, \cdots, s_r)$ contiene Ω_r? Ovvero quanti sono gli allineamenti che si possono formare prendendo un oggetto da S_1, un oggetto da S_2, ..., un oggetto da S_r?

Disposizioni con ripetizione

Poiché si può scegliere $s_1 \in S_1$ in n_1 modi e per ognuna di tali scelte si può scegliere $s_2 \in S_2$ in n_2 modi e così di seguito il numero complessivo degli allineamenti possibili sarà dato da

$$n_1 \cdot n_2 \cdots n_r \ .$$

Se in particolare, $S_1 = S_2 = \cdots = S_r$ ed $n_i = n > 1$, $i = 1, 2, \cdots, r$ allora il numero delle r–uple $(s_1, s_2, \cdots, s_r)$ di Ω_r è dato da n^r ed ogni allineamento si dice *disposizione con ripetizione.*

Regola 5 (Disposizioni con ripetizione)**.** *Dato un insieme $S = \{a_1, a_2, \cdots, a_n\}$ di n oggetti distinti, il numero degli allineamenti che si possono formare con r oggetti scelti tra gli n – ritenendo diversi due allineamenti o perché contengono oggetti differenti o perché gli stessi oggetti si susseguono in ordine diverso o, infine, perché uno stesso oggetto si ripete un numero diverso di volte – è dato da*

$$D^*_{n,r} = n^r$$

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

Esempio 2. *Ad esempio, $S = \{a_1, a_2, a_3\}$ ed $r = 2$, allora tali allineamenti sono*

$$(a_1, a_1), (a_1, a_2), (a_2, a_1), (a_1, a_3), (a_3, a_1), (a_2, a_2), (a_2, a_3), (a_3, a_2), (a_3, a_3)$$

in numero di $3^2 = 9$.

Esercizio 25. *Un'urna contiene* 20 *palline numerate da* 1 *a* 20*. Si eseguono* 5 *estrazioni rimettendo, dopo ogni estrazione, la pallina nell'urna.*

Quanti sono gli allineamenti che si possono ottenere come risultato delle 5 *estrazioni? E quanti sono gli allineamenti in cui non compare il numero* 20*?*

Soluzione 12. *Poiché* $S_1 = \{1, 2, \cdots, 20\}$ *la risposta alla prima domanda sarà* $D^*_{20,5} = 20^5 = 3200000$ *essendo* $n = 20$ *e* $r = 5$*. Per rispondere alla seconda domanda basta osservare che gli oggetti a disposizione sono quelli contenuti in* $S_2 = \{1, 2, \cdots, 19\}$ *e pertanto* $D^*_{19,5} = 19^5 = 2476099$

Disposizioni senza ripetizione

Se invece si vogliono contare gli allineamenti che si possono formare con gli oggetti contenuti in $S = \{a_1, a_2, \cdots, a_n\}$ – presi a gruppi di $1 \leq r \leq n$ – ma in modo che uno stesso oggetto non appaia più di una volta, si hanno le *disposizioni senza ripetizione*. Il numero di tali allineamenti equivale a contare i modi in cui si possono occupare r caselle con gli oggetti dati (e senza ripetizione). La prima casella potrà essere occupata in n modi. Per ognuno di questi vi sono $n-1$ modi per occupare la seconda casella e così di seguito fino all'ultima che potrà essere occupata in $(n-r+1)$ modi.

Regola 6 (Disposizioni senza ripetizione)**.** *Dato un insieme* $S = \{a_1, a_2, \cdots, a_n\}$ *di* n *oggetti distinti, il numero degli allineamenti che si possono formare con* $1 \leq r \leq n$ *oggetti scelti tra gli* n *– ritenendo diversi due allineamenti o perché contengono oggetti differenti o perché gli stessi oggetti si susseguono in ordine diverso – è dato da*

$$D_{n,r} = n(n-1)(n-2)\cdots(n-r+1)$$

Ogni allineamento si dice disposizione semplice o senza ripetizione di n *oggetti di classe* r.

Esempio 3. *Ad esempio, se* $S = \{a_1, a_2, a_3\}$ *ed è* $r = 2$*, allora tali allineamenti saranno*

$$(a_1, a_2), (a_2, a_1), (a_1, a_3), (a_3, a_1), (a_2, a_3), (a_3, a_2)$$

in numero di $3 \cdot 2 = 6$*. Si noti che non sono più presenti le coppie nella forma* (a_i, a_i)*,* $i = 1, 2, 3$.

Esercizio 26. *Un'urna contiene* 20 *palline numerate da* 1 *a* 20*. Si eseguono* 5 *estrazioni senza rimettere, dopo ogni estrazione, la pallina nell'urna. Quanti sono gli allineamenti che si possono ottenere come risultato delle* 5 *estrazioni? E quanti quelli in cui non compare il numero* 20*?*

Soluzione 13. *Abbiamo* $n = 20$ *e* $r = 5$*, poiché non si rimette la pallina nell'urna, abbiamo*

$$D_{20,5} = 20 \cdot 19 \cdot 18 \cdot 17 \cdot 16 = 1860480$$

mentre per rispondere al secondo quesito

$$D_{19,5} = 19 \cdot 18 \cdot 17 \cdot 16 \cdot 15 = 1395360 \ .$$

Permutazioni

Se, in particolare, si considerano le disposizioni senza ripetizione con $r = n$ allora $D_{n,n} = n(n-1)(n-2)\cdots 1 = n!$ e tali allineamenti, che possono differire solo per l'ordine con cui si susseguono gli oggetti, sono detti *permutazioni* di n oggetti.

Regola 7. *Dato un insieme* $S = \{a_1, a_2, \cdots, a_n\}$ *di* n *oggetti distinti, il numero degli allineamenti che si possono formare con tutti essi – ritenendo diversi due allineamenti perchè gli oggetti si susseguono in ordine diverso – è dato da* $n!$ *(si pone* $0! = 1$*).*

Esempio 4. *Ad esempio, se* $S = \{a_1, a_2, a_3\}$*, allora le permutazioni sono*

$$(a_1, a_2, a_3), (a_1, a_3, a_2), (a_3, a_2, a_1), (a_3, a_1, a_2), (a_2, a_3, a_1), (a_2, a_1, a_3)$$

in numero di $3! = 6$.

Esercizio 27. *Si abbiano* $n \geq 1$ *palline numerate da* 1 *a* n *ed altrettante scatole egualmente numerate. In quanti modi possiamo disporre le* n *palline nelle* n *scatole? E in quanti modi si può farlo quando si richiede che la pallina numerata con* $1 \leq m \leq n$ *occupi la scatola col numero* m*?*

Soluzione 14. *I modi sono ovviamente* $n!$*. I modi in cui si possono distribuire le* n *palline nelle* n *scatole con la pallina col numero* m *nella scatola* m *sono* $(n-1)!$

Combinazioni

Le disposizioni senza ripetizione si distinguono sia per l'ordine sia perchè contengono oggetti differenti. Se ora le disposizioni non vengono più distinte secondo l'ordine, gli allineamenti si chiamano *combinazioni* ed il loro numero viene indicato con

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

Regola 8. *Dato un insieme $S = \{a_1, a_2, \cdots, a_n\}$ di n oggetti distinti, il numero degli allineamenti che si possono formare con $1 \leq r \leq n$ oggetti scelti tra gli n – ritenendo diversi due allineamenti solo perché contengono oggetti differenti – è dato da*

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

Ogni allineamento si dice combinazione senza ripetizione di n oggetti di classe r.

Il numero $C_{n,r}$ è spesso indicato con $\binom{n}{r}$ e si chiama coefficiente binomiale (o anche binomio di Newton).

Esempio 5. *Se, ad esempio $S = \{a_1, a_2, a_3\}$ ed è $r = 2$, le combinazioni saranno date da*

$$(a_1, a_2), (a_1, a_3), (a_2, a_3)$$

in numero di $\frac{3 \cdot 2}{2} = 3$

Esercizio 28. *Si vuole costituire un comitato di 5 membri scelti tra 10 persone. Quanti differenti comitati si possono formare?*

Soluzione 15. *Poiché ognuna delle persone disponibili potrà apparire al più una volta nel comitato e poiché due gruppi costituiscono uno stesso comitato se contengono le stesse persone (anche se esse appaiono in un ordine diverso) allora il numero richiesto è*

$$C_{10,5} = \frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5} = 252$$

Cardinalità dell'insieme delle parti di un insieme finito

Teorema 7. *Sia $S_n = \{a_1, a_2, \cdots, a_n\}$ un insieme di n oggetti distinti allora la cardinalità di $\mathcal{P}(S)$ è 2^n.*

Dimostrazione 10. *Dimostriamo questo asserto per induzione. Cominciamo con il verificare che l'asserto vale quando $n = 0, 1, 2$ infatti*

- $n = 0$ *e* $\mathcal{P}(S_0) = \mathcal{P}(\emptyset) = \{\emptyset\}$ *ha cardinalità* $2^0 = 1$*;*

- $n = 1$ *e* $\mathcal{P}(S_1) = \{\emptyset, \{a_1\}\}$ *ha cardinalità* $2^1 = 2$*;*

- $n = 2$ *e* $\mathcal{P}(S_2) = \{\emptyset, \{a_1\}, \{a_2\}, S_2\}$ *ha cardinalità* $2^2 = 4$*.*

Ora assumiamo che $\#\mathcal{P}(S_k) = 2^k$ *e mostriamo che vale* $\#\mathcal{P}(S_{k+1}) = 2^{(k+1)}$*. Infatti i sottoinsiemi di* S_{k+1} *possono essere suddivisi in due categorie*

- *quelli che non contengono* a_{k+1}*, diciamo* A*;*

- *quelli che contengono* a_{k+1}*, diciamo* B*.*

Gli insiemi A *sono* 2^k *per l'ipotesi di induzione. Rimane da mostrare che gli insiemi* B *sono* 2^k*.*

A questo scopo si noti che ogni insieme B *è esprimibile come* $A \cup \{a_{k+1}\}$ *cioè esiste una funzione bijettiva (iniettiva e suriettiva) del tipo*

$$A \mapsto A \cup \{a_{k+1}\}$$

e questo implica che gli insiemi B *siano* 2^k*.*

In conclusione abbiamo

$$2^k + 2^k = 2^{k+1}$$

Cardinalità di $\mathcal{P}$ e Combinazioni

È importante notare che vale la seguente identità

$$2^n = \sum_{r=0}^{n} \binom{n}{r}, \qquad n \geq 0$$

ed inoltre vale la formula di Van der Monde

$$\binom{m+n}{r} = \sum_{i=0}^{r} \binom{m}{i} \cdot \binom{n}{r-i}$$

infine è facile vedere che

$$\binom{n}{r} = \frac{n!}{r! \cdot (n-r)!} = \binom{n}{n-r}$$

Esercizio 29 (Scimmie e tastiere)**.** *Una scimmia digita* 3 *tasti in maniera casuale da una tastiera per computer (supponiamo che la tastiera abbia* 100 *tasti). Qual'è la probabilità che la scimmia digiti una parola di tre lettere che inizia con una consonante (*20*) e finisce con due vocali (consideriamo anche la "j" come vocale e quindi abbiamo* 6 *vocali). La parola non deve avere senso compiuto.*

Soluzione 16. *Consideriamo di codificare i tasti con numeri da* 1 *a* 100 *e abbiamo* $\Omega_i = \{1, 2, \cdots, 100\}$ *per ogni prova* $i = 1, 2, 3$. *La cardinalità dello spazio prodotto* $\Omega = \Omega_1 \times \Omega_2 \times \Omega_3$ *dove starà il nostro risultato è* $D^*_{100,3} = 100^3$. *Usiamo come Tribù su* Ω_i *l'insieme delle parti e come Tribù su* Ω *la Tribù prodotto. Dato che tutti gli eventi elementari presenti in* Ω *sono equiprobabili e quindi*

$$\Pr_{\Omega}(\textit{evento elementare}) = \frac{1}{100^3}$$

allora la probabilità dell'evento desiderato è pari a $1/100^3$ *per il numero di eventi elementari che implicano tale evento. La prima consonante può essere scelta in* 20 *modi diversi, la vocale in* 6 *modi diversi e l'ultima vocale in* 6 *modi diversi quindi*

$$\begin{aligned} &= \Pr_{\Omega}(\textit{una parola di tre lettere che inizia con una consonante e finisce con due vocali}) \\ &= \frac{20 \cdot 6 \cdot 6}{100^3} = 0.00072\ . \end{aligned}$$

Esercizio 30 (Abracadabra)**.** *Qual'è la probabilità che compaia la parola ABRACADABRA se le lettere A,A,A,A,A,B,B,C,D,R,R sono sorteggiate in maniera casuale?*

Soluzione 17. *I casi possibili sono tutte le permutazioni dell'insieme delle lettere cioè* 11!. *Se tutte le lettere fossero diverse allora avremmo un caso favorevole. Ma nel nostro caso ci sono lettere che compaiono più volte così certe permutazioni portano tutte alla stessa parola.*

Etichettiamo le lettere con un numero da 1 *a* 11. *Lo spazio campionario è dato dall'insieme delle permutazioni di* 11 *numeri così* $\#\Omega = 11!$. *I casi favorevoli (quelli che implicano l'evento desiderato) sono calcolati nel modo seguente*

- *La lettera A può essere posizionata in* 5! *modi diversi*

- *Le lettere B e R in* 2! *modi*

così

$$\Pr(ABRACADABRA) = \frac{5! \cdot 2! \cdot 2!}{11!} = \frac{1}{83160}$$

Esercizio 31 (Estrazioni da un'urna)**.** *Un'urna contiene* 100 *palline di cui* 30 *bianche e* 70 *rosse. Si vuole conoscere la probabilità di estrarre* 5 *palline bianche in una successione di* 10 *estrazioni senza reimmissione.*

Soluzione 18. *I possibili risultati delle* 10 *estrazioni sono ovviamente gli allineamenti* $(a_1, a_2, \cdots, a_{10})$ *in cui* $a_i = B$ *(*B *=bianca) oppure* $a_i = R$ *(*R *=rossa), per* $i = 1, 2, \cdots, 10$.

Quindi Ω *sarà costituito da tutti questi allineamenti.*

Si può assumere che ognuno di tali allineamenti abbia la stessa probabilità di venire estratto e il problema si riduce nel calcolare il numero di eventi elementati che implicano l'evento desiderato.

A tale scopo osserviamo che la cardinalità di Ω *è*

$$\#\Omega = C_{100,10} = \binom{100}{10}$$

in quanto non è importante l'ordine con cui compaiono. Vi sono poi 30 *palline bianche da cui possiamo estrarne* 5 *sicché esse possono essere selezionate in*

$$C_{30,5} = \binom{30}{5}$$

modi diversi. Per ogni scelta delle palline bianche, le 5 *rosse possono essere scelte tra le* 70 *disponibili, in* $C_{70,5} = \binom{70}{5}$ *modi. Pertanto il numero degli allineamenti che contengono* 5 *palline bianche è*

$$\binom{30}{5} \cdot \binom{70}{5}$$

e la probabilità dell'evento desiderato è

$$\Pr(\{\textit{5 palline bianche e 5 rosse}\}) = \frac{\binom{30}{5} \cdot \binom{70}{5}}{\binom{100}{10}} \simeq 0.0996\ .$$

Un altro modo per determinare la probabilità cercata è quello di fare riferimento al numero degli allineamenti che tengono conto anche dell'ordine in cui si presentano le palline. Il numero degli allineamenti di 100 palline prese a gruppi di 10 è $D_{100,10}$.

Le cinque posizioni occupate dalle palline bianche nella successione delle 10 *estrazioni possono essere scelte in* $C_{10,5}$ *modi.*

Quando una tale scelta è fatta, la prima pallina bianca può essere scelta in 30 *modi, la seconda in* 29*, la terza in* 28*, la quarta in* 27 *e la quinta in* 26 *cioè* $D_{30,5}$ *e nello stesso modo per le palline rosse:* $D_{70,5}$.

In definitiva il numero degli allineamenti contenenti 5 *palline bianche e* 5 *rosse sarà*

$$\binom{10}{5} \cdot D_{30,5} \cdot D_{70,5}$$

e la probabilità

$$\Pr(\{\textit{5 palline bianche e 5 rosse}\}) = \frac{\binom{10}{5} \cdot D_{30,5} \cdot D_{70,5}}{D_{100,10}}$$

Qual'è lo spazio probabilizzabile che stiamo considerando in questo modo?

Esercizio 32 (Mazzo di 52 carte)**.** *Una mano di poker è formata da cinque carte estratte a caso senza reimmissione da un mazzo di* 52 *carte. Determinare le probabilità dei seguenti eventi*

- $E_1 = \{$*la mano contiene* 5 *carte dello stesso colore (che possono essere messe) in scala,* 10*,* J*,* Q*,* K*,* A *(scala reale)*$\}$;
- $E_2 = \{$*la mano contiene* 5 *carte, di uno stesso colore, con valori in successione (ad esempio,* A*,* 2*,* 3*,* 4*,* 5*, ecc.) che non sia una scala reale*$\}$;
- $E_3 = \{$*la mano contiene quattro carte di eguale valore*$\}$;
- $E_4 = \{$*la mano contiene due carte di eguale valore e tre carte di eguale valore (Full)*$\}$.

Soluzione 19. *Consideriamo come spazio campionario* $\Omega = \{$*combinazioni di* 5 *carte da un mazzo di* 52$\}$.

- *Per calcolare la probabilità di* E_1 *basta pensare che (in questa formulazione) vi sono solo* 4 *combinazioni che implicano* E_1 *quindi*

$$\Pr(E_1) = \frac{4}{\binom{52}{5}}$$

- *L'evento* E_2 *è implicato da* $40 - 4$ *combinazioni perché per ogni colore abbiamo* 10 *possibili successioni di* 5 *carte ma l'ultima è la scala reale e quindi*

$$\Pr(E_2) = \frac{40}{\binom{52}{5}} - \frac{4}{\binom{52}{5}}$$

- *Per* E_3 *possiamo ragionare nel modo seguente. Fissiamo per ora il valore, diciamo* 1*, abbiamo* $\binom{4}{4}$ *modi per scegliere le quattro carte* 1 *e* $\binom{48}{1}$ *modi per scegliere l'ultima carta (si badi che utilizzando le combinazioni non teniamo conto dell'ordine con cui le carte compaiono). Ora ci sono* 13 *valori nel mazzo e quindi gli eventi favorevoli sono*

$$13 \cdot \binom{4}{4} \cdot \binom{48}{1} = 13 \frac{4!}{4!\ 0!} \frac{48!}{47!\ 1!} = 13 \cdot 48 = 624$$

e la probabilità è

$$\Pr(E_3) = \frac{624}{\binom{52}{5}} \simeq 0.240 \cdot 10^{-3}$$

- *Per* E_4 *ragioniamo allo stesso modo. Fissiamo i due valori* 1 *e* 2 *ad esempio, e quindi il numero di allineamenti (senza tenere conto dell'ordine in cui compaiono) è*

$$\binom{4}{3} \cdot \binom{4}{2}$$

rimane da calcolare in quanti modi possiamo scegliere i due valori: $D_{13,2} = 13 \cdot 12$ *e quindi il numero di eventi favorevoli è*

$$13 \cdot 12 \cdot \binom{4}{3} \cdot \binom{4}{2} = 13\ 12 \frac{4!}{3!\ 1!} \frac{4!}{2!\ 2!} = 13 \cdot 12 \cdot 2^3 \cdot 3 = 3744$$

e la probabilità è

$$\Pr(E_4) = \frac{3744}{\binom{52}{5}} \simeq 1.441 \cdot 10^{-3}$$

- *Si calcoli la probabilità dell'evento* $E_5 = \{$*almeno due assi*$\}$*; Dato che la probabilità di avere* $k = 0, 1, \cdots, 4$ *assi è*

$$\Pr(\{\textit{esattamente } k \textit{ assi}\}) = p_k = \frac{\binom{4}{k}\binom{48}{5-k}}{\binom{52}{5}}$$

allora $\Pr(\{\textit{almeno due assi}\}) = p_2 + p_3 + p_4$.

- *Si calcoli la probabilità di fare Colore servito (escono 5 carte dello stesso seme).*

$$\Pr(\{\textit{Colore}\}) = 4 \cdot \frac{\binom{13}{5}\binom{39}{0}}{\binom{52}{5}}$$

- *Si calcoli la probabilità di fare Poker servito*

$$\Pr(\{\textit{Poker servito}\}) = 13 \cdot \frac{\binom{4}{4}\binom{48}{5-4}}{\binom{52}{5}}$$

Esercizio 33 (Impiegati e uffici)**.** *In quanti modi si possono assegnare venti impiegati a quattro uffici se ad ogni ufficio devono essere assegnati cinque impiegati?*

Soluzione 20. *Possiamo ragionare nel modo seguente, pensando inizialmente che il problema sia di assegnare ogni impiegato ad una delle 20 sedie disponibili. Il numero di queste assegnazioni è pari al numero di permutazioni cioè* $20!$. *È chiaro che all'interno di queste assegnazioni (alle sedie) ci sono permutazioni che portano alla stessa assgnazioni in termini di ufficio. Per ogni ufficio le permutazioni equivalenti sono pari a* $5!$ *e in definitiva il numero di modi è*

$$\frac{20!}{(5!)^4}$$

Esercizio 34 (Ragazzi e ragazze)**.** *Da una lista di* 10 *ragazzi e* 7 *ragazze si deve formare un comitato comprendente* 5 *ragazzi e* 3 *ragazze. Quanti possibili comitati si possono formare?*

Soluzione 21. *Possiamo scegliere i ragazzi in* $\binom{10}{5}$ *modi diversi e per ognuno di questi modi possiamo scegliere le ragazza in* $\binom{7}{3}$ *modi e quindi*

$$\binom{10}{5} \cdot \binom{7}{3} = 8820$$

Esercizio 35 (Codici)**.** *Quanti codici si possono formare utilizzando tre cifre e due lettere dell'alfabeto inglese (26 lettere), se le lettere occupano le prime due posizioni e cifre e lettere possono ripetersi.*

Quanti codici presentano qualche ripetizione?

Soluzione 22. *Nel primo quesito usiamo le Disposizioni con ripetizione e quindi abbiamo*

$$D^*_{26,2} \cdot D^*_{10,3} = 26^2 \cdot 10^3$$

Per la seconda parte possiamo calcolare facilmente il numero di codici che non presentano ripetizioni che sono $D_{26,2} \cdot D_{10,3} = 26 \cdot 25 \cdot 10 \cdot 9 \cdot 8$ *e quindi il numero di quelli che presentano una qualche ripetizione è*

$$D^*_{26,2} \cdot D^*_{10,3} - D_{26,2} \cdot D_{10,3}$$

Esercizio 36 (Compleanni)**.** *Determinare la probabilità che n persone scelte a caso abbiano tutte compleanno diverso (si ignorino gli anni bisestili).*

Soluzione 23. *Facciamo innanzitutto un'assunzione di equiprobabilità del giorno di nascita (questa assunzione non è verificata nella realtà perché ci sono periodi dell'anno in cui ci sono più nascite.) Sotto questa ipotesi la probabilità è*

$$\frac{D_{365,n}}{D^*_{365,n}} = \frac{\frac{365!}{(365-n)!}}{365^n}$$

se $n \leq 365$ *altrimenti la probabilità è* 0

1.3 Probabilità sui Reali

Costruzione di uno spazio probabilizzato quando Ω è più che numerabile

Per semplicità, limiteremo la nostra attenzione al caso $\Omega = \mathbb{R}$ l'insieme dei numeri reali. I risultati potranno essere estesi (facendo uso degli spazi prodotto) al caso $\Omega = \mathbb{R}^n$ per qualche n intero positivo fissato.

Il primo passo è quello di cercare un spazio probabilizzabile adeguato. Una possibilità risulta il considerare $\mathcal{A} = \mathcal{P}(\mathbb{R})$ che è sicuramente una Tribù. Tuttavia questa Tribù è troppo vasta e comprende molti eventi che non sono di interesse.

È il caso di considerare una Tribù meno fine (che abbia meno elementi di $\mathcal{P}(\mathbb{R})$). Ad esempio possiamo richiedere che la Tribù contenga tutti gli intervalli nella forma

$$(a, b] \qquad a \leq b \ .$$

A questo scopo possiamo considerare la Tribù generata a partire da questa classe di insiemi cioè, posto $\mathcal{F} = \{(a, b] : \ a \leq b; a, b \in \mathbb{R}\}$,

$$\mathcal{A}_{\mathcal{F}} = \cap\{\mathcal{A} : \mathcal{A} \text{ è una Tribù e } \mathcal{F} \subseteq \mathcal{A}\}$$

che è la più piccola Tribù che contiene tutti i suddetti intervalli.

Definizione 26 (Tribù Boreliana)**.** *Si chiama Tribù Boreliana di $\mathbb{R}$, e si denota con $\mathcal{B}(\mathbb{R})$, la Tribù generata su $\mathbb{R}$ dalla classe di tutti gli intervalli $(a, b]$ di $\mathbb{R}$. I suoi elementi si chiamano gli insiemi boreliani di $\mathbb{B}$.*

e lo spazio $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ è uno spazio probabilizzabile.

Tribù Boreliana

La Tribù di Borel su $\mathbb{R}$ contiene anche i seguenti elementi

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

- gli insiemi numerabili di $\mathbb{R}$.

Ma $\mathcal{B}(\mathbb{R})$ contiene anche sottoinsiemi di $\mathbb{R}$ che non si lasciano descrivere con la stessa semplicità. In buona sostanza, $\mathcal{B}(\mathbb{R})$ contiene tutti i sottoinsiemi di $\mathbb{R}$ per i quali abbia senso parlare della loro *lunghezza* (o nel caso di $\mathbb{R}^2$, $\mathbb{R}^3$, ecc., di *area*, *volume*, ecc.).

Naturalmente, $\mathcal{B}(\mathbb{R})$ *non contiene tutti* i sottoinsiemi di $\mathbb{R}$; alcuni sottoinsiemi di $\mathbb{R}$ non appartengono a $\mathcal{B}(\mathbb{R})$ ma sono difficili a costruirsi e non hanno alcuna concreta rilevanza applicativa (Ad esempio si veda Billingsley (1995), pag. 45).

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

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

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

ponendo

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

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

- un intervallo;
- un'unione numerabile di intervalli disgiunti

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

Se però l'evento A non può ricondursi ai casi precedenti, la determinazione di $\Pr(A)$, mediante $F(x)$, richiede l'uso della nozione di integrale di Lebesgue–Stieltjes.

Sull'assunzione di *equiprobabilità*

La valutazione di equiprobabilità degli eventi elementari *non è compatibile* con un insieme Ω numerabile. In questo caso infatti si dovrebbe avere $\sum_{i=1}^{\infty} \Pr(\omega_i) = \Pr(\Omega) = 1$ con $\Pr(\omega_i) = c$, $i = 1, 2, \cdots$ ma

- se $c > 0$ è
$$\sum_{i=1}^{\infty} \Pr(\omega_i) = \infty$$
- se $c = 0$ è
$$\sum_{i=1}^{\infty} \Pr(\omega_i) = 0$$

Per essere più chiari si dovrebbe dire che l'ipotesi di equiprobabilità non è compatibile con un insieme Ω numerabile se si *chiede che* Pr *sia* σ*–additiva*!.

Esempio 6. *Si consideri l'esperimento consistente nello scegliere un punto dell'intervallo* $[0, 1]$*. Come Tribù utilizziamo la Tribù di Borel dei sottoinsiemi di* $[0, 1]$*, cioè*

$$\mathcal{B}\left([0,1]\right) = \{[0,1] \cap A : A \in \mathcal{B}(\mathbb{R})\} \ .$$

Assegniamo la probabilità che il punto scelto appartenga all'intervallo $(a, b]$ *(di* $[0, 1]$*) mediante la lunghezza dell'intervallo stesso*

$$\Pr((a,b]) = F(b) - F(a) = b - a, \qquad 0 \le a \le b \le 1$$

cioè prendendo $F(x) = x$*,* $0 \le x \le 1$*.*

Per prima cosa notiamo che gli intervalli $(a, b]$*,* (a, b)*,* $[a, b)$*,* $[a, b]$ *hanno tutti la stessa probabilità. Infatti, la funzione* $F(x) = x$ *è continua ed inoltre abbiamo, ad esempio*

$$\Pr([a,b]) = \Pr(\lim_{n\to\infty} (a - \frac{1}{n}, b])$$

ma gli intervalli $(a - \frac{1}{n}, b]$ *per* $n \to \infty$ *costituisce una successione non crescente e quindi*

$$\begin{aligned}
\Pr(\lim_{n\to\infty} (a - \frac{1}{n}, b]) &= \lim_{n\to\infty} \Pr((a - \frac{1}{n}, b]) \\
&= \lim_{n\to\infty} \left(F(b) - F(a - \frac{1}{n}) \right) = F(b) - \lim_{n\to\infty} F(a - \frac{1}{n}) \\
&= F(b) - F(a) + \lim_{n\to\infty} \frac{1}{n} = F(b) - F(a)
\end{aligned}$$

Esempio 7. *Quali sono le probabilità dei seguenti eventi?*

- $A_1 = \{(0, \frac{1}{4})\}$;
- $A_2 = \{a\}$;
- $A_3 = \{a_1, a_2, \cdots\}$;
- $A_4 = \{[\frac{1}{2}, \frac{3}{4}] \cup [\frac{4}{5}, 1]\}$;
- $A_5 = [0, 1]/A_3$;
- $A_6 = \cup_{n=1}^{\infty}[\frac{k}{k+1}, \frac{k}{k+1} + \frac{1}{10^k}]$

dove $a \in [0, 1]$ *e* $a_i \in [0, 1]$, $i = 1, 2, \cdots$.

Soluzione 24.

- A_1 *è un intervallo e pertanto* $\Pr(A_1) = F(\frac{1}{4}) - F(0) = \frac{1}{4}$;
- $A_2 = \{a\} = \cap_{n=1}^{\infty}(a - \frac{1}{n}, a] = \lim_{n\to+\infty}(a - \frac{1}{n}, a]$ *e pertanto (ricordando che la successione degli intervalli è non crescente) per la proprietà di continuità,*

$$\begin{aligned}
\Pr(A_2) = \Pr(\{a\}) &= \Pr\left(\lim_{n\to+\infty}(a - \frac{1}{n}, a]\right) = \lim_{n\to+\infty} \Pr\left((a - \frac{1}{n}, a]\right) \\
&= \lim_{n\to+\infty}\left(F(a) - F(a - \frac{1}{n})\right) \\
&= \lim_{n\to+\infty} \frac{1}{n} = 0
\end{aligned}$$

- $A_3 = \{a_1, a_2, \cdots\} = \cup_{i=1}^{\infty}\{a_i\}$ *e quindi*

$$\Pr(A_3) = \Pr(\{a_1, a_2, \cdots\}) = \Pr(\cup_{i=1}^{\infty}\{a_i\}) = \sum_{i=1}^{\infty} \Pr(\{a_i\}) = 0$$

- *Per* A_4 *si consideri che i due intervalli sono disgiunti allora*

$$\begin{aligned}
\Pr(A_4) = \Pr\left([\frac{1}{2}, \frac{3}{4}] \cup [\frac{4}{5}, 1]\right) &= \Pr\left([\frac{1}{2}, \frac{3}{4}]\right) + \Pr\left([\frac{4}{5}, 1]\right) \\
&= F\left(\frac{3}{4}\right) - F\left(\frac{1}{2}\right) + F(1) - F\left(\frac{4}{5}\right) = \frac{1}{4} + \frac{1}{5} = \frac{9}{20}
\end{aligned}$$

-

$$\Pr(A_5) = \Pr([0, 1]/A_3) = \Pr([0, 1]) - \Pr(A_3) = \Pr([0, 1]) = 1$$

- *A_6 è un'unione numerabile di intervalli disgiunti e pertanto*

$$\begin{aligned}\Pr(A_6) &= \sum_{k=1}^{\infty} \Pr\left(\left[\frac{k}{k+1}, \frac{k}{k+1} + \frac{1}{10^k}\right]\right) \\ &= \sum_{k=1}^{\infty} \left(F\left(\frac{k}{k+1} + \frac{1}{10^k}\right) - F\left(\frac{k}{k+1}\right)\right) \\ &= \sum_{k=1}^{\infty} \frac{1}{10^k} = \frac{1}{1-\frac{1}{10}} - 1 = \frac{\frac{1}{10}}{1-\frac{1}{10}} = \frac{1}{9}\end{aligned}$$

1.4 Probabilità Condizionale

Probabilità Condizionale

Definizione 27. *Sia* $(\Omega, \mathcal{A}), \mathrm{Pr})$ *uno spazio probabilizzato. Fissato un elemento* H *di* $\mathcal{A}$*, con* $\Pr(H) \neq 0$*, si chiama funzione di probabilità dedotta da* $\Pr$ *sotto la condizione* H *la funzione di probabilità* $\Pr_H$ *sullo spazio* $(\Omega, \mathcal{A})$ *probabilizzabile*

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

per ogni evento $A \in \mathcal{A}$*.*

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

$$\Pr(A|H) \ .$$

Si tratta evidentemente di una funzione di probabilità concentrata su H. Essa coincide con Pr nel caso banale in cui l'evento H sia quasi certo secondo P.

È facile mostrare che la funzione di probabilità $\Pr_H(A) = \Pr(A|H)$ rispetta tutti gli assiomi della probabilità.

Definizione 28 (Spazio di Probabilità Condizionale)**.** *Lo spazio di probabilità* $(\Omega, \mathcal{A}, \Pr_H)$ *è detto Spazio di Probabilità Condizionale.*

Spazio di Probabilità Condizionale

Dopo che è accaduto l'evento H, ogni evento $A \in \mathcal{A}$ tale che $A \cap H = \emptyset$ si ha $\Pr_H(A) = \Pr(A|H) = 0$.

Quindi diventano due modi equivalenti di probabilizzare lo stesso problema gli spazi

$$(\Omega, \mathcal{A}, \Pr_H) \qquad (\Omega_H, \mathcal{A}_H, \Pr_H)$$

dove, chiaramente

$$\Omega_H = \{A \subseteq H : A \in \Omega\} = H$$

perché gli eventi (elementari) che possono essere accaduti sono solo quelli che implicano H e

$$\mathcal{A}_H = \{A \cap H : A \in \mathcal{A}\}$$

Esercizio 37. *Siano* $\mathcal{A}$ *una Tribù di sottoinsiemi di* Ω *e* $H \subset \Omega$ *un qualunque insieme. Mostrare che* $\mathcal{A}_H = \{A \cap H : A \in \mathcal{A}\}$ *è una Tribù di sottoinsiemi di* H *ma non una Tribù di sottoinsiemi di* Ω.

Esercizio 38 (Lancio di due monete bilanciate)**.** *Consideriamo l'esperimento del lancio di due monete bilanciate. Sappiamo che nel primo lancio è uscita una testa. Qual è la probabilità di avere due teste?*

Soluzione 25. *Lo spazio probabilizzato è* $(\Omega = \{\{(C,C)\}, \{(C,T)\}, \{(T,C)\}, \{(T,T)\}\}, \mathcal{A} = \mathcal{P}(\Omega), \Pr(\{\textit{evento elementare}\}) = \frac{1}{4})$.

L'evento accaduto H *è* $\{\{(T,C)\}, \{(T,T)\}\}$ *quindi*

$$\mathcal{A}_H = \{\emptyset, \{(T,C)\}, \{(T,T)\}, \Omega_H\}$$

e la funzione di probabilità condizionale $\Pr_H$ *è calcolata nel modo seguente per i due eventi elementari presenti in* H

$$\Pr_H(\{(T,C)\}) = \frac{\Pr(\{(T,C)\}\cap H)}{\Pr(H)} = \frac{\Pr(\{(T,C)\})}{\Pr(H)} = \frac{\frac{1}{4}}{\frac{2}{4}} = \frac{1}{2}$$
$$\Pr_H(\{(T,T)\}) = \frac{\Pr(\{(T,T)\}\cap H)}{\Pr(H)} = \frac{\Pr(\{(T,T)\})}{\Pr(H)} = \frac{\frac{1}{4}}{\frac{2}{4}} = \frac{1}{2}$$

quindi la probabilità dell'evento desiderato ("due teste") è $\frac{1}{2}$

Esercizio 39. *Se invece l'evento accaduto è* $H = \{$*è uscita almeno una testa*$\}$ *come si modifica lo spazio condizionale?*

Esercizio 40. *Un dado regolare viene lanciato due volte. Nell'ipotesi che si sappia che il punteggio totale dei due lanci è* 6*, qual è la probabilità che il punteggio del primo lancio sia stato* 3*?*

Soluzione 26. *Detto* H *l'evento "punteggio totale uguale a 6" e* A *l'evento "punteggio del primo lancio uguale a 3" si ha*

$$H = \{(1,5),(2,4),(3,3),(4,2),(5,1)\}$$
$$A = \{(3,1),(3,2),(3,3),(3,4),(3,5),(3,6)\}$$

da cui $\Pr(H) = \frac{5}{36}$ *e* $\Pr(A) = \frac{6}{36}$*, inoltre* $A \cap H = \{(3,3)\}$ *e quindi* $\Pr(A \cap H) = \frac{1}{36}$ *ed infine*

$$\Pr(A|H) = \frac{\frac{1}{36}}{\frac{5}{36}} = \frac{1}{5}$$

Esercizio 41 (Ruote di bicicletta)**.** *Nell'assemblaggio delle ruote di bicicletta vengono effettuati due controlli per verificarne la funzionalità. Il primo controllo è effettuato sulla camera d'aria e sul pneumatico (*C_1*) mentre il secondo sulla centratura dei raggi e sulla regolazione del mozzo (*C_2*). Poniamo* A *l'evento "il controllo* C_1 *è superato" e* B *l'evento "il controllo* C_2 *è superato".*

Da un'indagine svolta nel passato si è valutato che $\Pr(A) = 0.8$, $\Pr(B) = 0.9$ *e* $\Pr(A \cap B) = 0.75$.

Sapendo che una ruota è difettosa qual è la probabilità che non abbia superato il controllo C_1*? E che non abbia superato solo il controllo* C_1*?*

Soluzione 27. *L'evento* $H = \{$*ruota difettosa*$\}$ *è dato da* $A^c \cup B^c$ *e la sua probabilità è*

$$\Pr(H) = \Pr((A \cap B)^c) = 1 - \Pr(A \cap B) = 1 - 0.75 = 0.25 \ .$$

Quindi

$$\begin{aligned}
\Pr(A^c|H) = & \frac{\Pr(A^c \cap (A^c \cup B^c))}{\Pr(H)} \\
= & \frac{\Pr(A^c)}{\Pr(H)} \\
= & \frac{1-0.8}{0.25} = \frac{0.2}{0.25} = 0.8
\end{aligned}$$

Per rispondere al secondo quesito dobbiamo invece calcolare

$$\begin{aligned}
\Pr(A^c \cap B|H) = & \frac{\Pr(((B/(A \cap B)) \cap H)}{\Pr(H)} \\
= & \frac{\Pr(B/(A \cap B))}{\Pr(H)} \\
= & \frac{\Pr(B) - \Pr(A \cap B)}{\Pr(H)} \\
= & \frac{0.9 - 0.75}{0.25} = \frac{0.15}{0.25} = 0.6
\end{aligned}$$

Esercizio 42 (Ancora sul lancio di dadi?)**.** *Un dado regolare viene lanciato tre volte. Qual è la probabilità che i puntegi realizzati nei tre lanci siano tutti differenti?*

Soluzione 28. *Se* A_1, A_2, A_3 *rappresentano gli eventi*

- $A_1 = \{$*punteggio qualunque nel primo lancio*$\}$*;*
- $A_2 = \{$*punteggio del secondo lancio differente da quello del primo*$\}$*;*
- $A_3 = \{$*punteggio del terzo lancio differente da quelli del primo e del secondo*$\}$*;*

allora

$$\Pr(A_1 \cap A_2 \cap A_3) = \Pr(A_1) \cdot \Pr(A_2|A_1) \cdot \Pr(A_3|A_1 \cap A_2) = 1 \cdot \frac{5}{6} \cdot \frac{4}{6} = \frac{5}{9} \ .$$

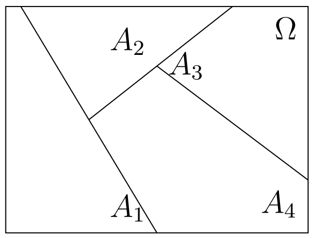
Naturalmente la probabilità richiesta poteva essere calcolata come rapporto tar il numero di allineamenti dei tre punteggi tutti differenti e quello di tutti i possibili allineamenti

$$\Pr(A_1 \cap A_2 \cap A_3) = \frac{6 \cdot 5 \cdot 4}{6^3} = \frac{5}{9} \ .$$

Definizione 29 (Classe completa di eventi (Partizione di un insieme))**.** *Dato un spazio probabilizzabile* $(\Omega, \mathcal{A})$ *la famiglia di eventi* $\{A_n\}_{n=1}^{\infty}$ *è detta Classe Completa se*

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

Esempio 8 (Classe completa di eventi)**.**



Teorema delle Probabilità Totali

Un risultato estremamente utile per il calcolo delle probabilità di eventi mediante l'uso delle probabilità condizionali è quello stabilito dal seguente teorema

Teorema 8 (delle Probabilità Totali)**.** *Sia* $\{A_i\}_{i=1}^{\infty}$ *una famiglia di eventi che costituisce una Classe Completa di* Ω *tale che*

- $\Pr(A_i) > 0$, $i = 1, 2, \cdots$.

Sia B *un qualunque evento. Allora*

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

Dimostrazione 11. *Dalla relazione* $\Omega = \cup_{i=1}^{\infty} A_i$, *intersecando con* B *ambo i membri, si ottiene, per la proprietà distributiva dell'intersezione rispetto all'unione*

$$\Omega \cap B = B = (\cup_{i=1}^{\infty} A_i) \cap B = \cup_{i=1}^{\infty}(A_i \cap B)$$

con gli eventi $A_i \cap B$, $i = 1, 2, \cdots$ *a due a due incompatibili. Segue allora che*

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

Esercizio 43. *Si abbiano due urne* U_1 *e* U_2. *La prima urna* U_1, *contiene* 5 *palline bianche e* 10 *nere; la seconda* U_2, *contiene* 8 *palline bianche e* 10 *nere. Si sceglie a caso una delle due urne e si estrae una pallina. Qual è la probabilità di osservare una pallina bianca?*

Soluzione 29. *Se si sceglie l'urna* U_1 *si ha* $\Pr(B|U_1) = \frac{5}{15} = \frac{1}{3}$; *se si sceglie l'urna* U_2 *si ha* $\Pr(B|U_2) = \frac{8}{18} = \frac{4}{9}$. *Poiché poi si sceglie a caso una delle due urne sarà* $\Pr(U_1) = \Pr(U_2) = \frac{1}{2}$ *e pertanto*

$$\Pr(B) = \Pr(U_1)\Pr(B|U_1) + \Pr(U_2)\Pr(B|U_2) = \frac{1}{2} \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{4}{9} = \frac{7}{18}$$

Esercizio 44. *Esplicitare gli spazi probabilizzati coinvolti.*

Esercizio 45 (Lavoratori e disoccupati). *La forza lavoro di un paese è ripartita in* 5 *regioni* R_i, $i = 1, \cdots, 5$ *nelle percentuali, rispettivamente, del* 10%, 22%, 19%, 30% *e* 19%. *I tassi di disoccupazione nelle stesse regioni sono, nell'ordine,* 5%, 2%, 3%, 1% *e* 8%.

- *Estraendo a caso un individuo qual è la probabilità che egli sia un disoccupato?*
- *Avendo osservato un individuo disoccupato qual è la probabilità che egli provenga dalla regione* R_1?

Soluzione 30. *Si indichi* A_i, $i = 1, \cdots, 5$ *l'evento "l'individuo estratto appartiene alla regione* R_i*" e sia* D *l'evento "l'individuo estratto è disoccupato". Dato che ogni individuo ha la stessa probabilità di essere estratto* $\Pr(A_1) = 0.1$, $\Pr(A_2) = 0.22$, $\Pr(A_3) = 0.19$, $\Pr(A_4) = 0.3$ *e* $\Pr(A_5) = 0.19$ *mentre* $\Pr(D|A_1) = 0.05$, $\Pr(D|A_2) = 0.02$, $\Pr(D|A_3) = 0.03$, $\Pr(D|A_4) = 0.01$ *e* $\Pr(D|A_5) = 0.08$.

Per il Teorema delle Probabilità Totali abbiamo

$$\Pr(D) = \sum_{i=1}^{5} \Pr(A_i) \Pr(D|A_i) \simeq 0.033$$

Per rispondere al secondo quesito si tratta di calcolare

$$\Pr(A_1|D) = \frac{\Pr(A_1 \cap D)}{\Pr(D)} = \frac{\Pr(A_1)\Pr(D|A_1)}{\Pr(D)} = \frac{0.1 \cdot 0.05}{0.033} \simeq 0.15$$

Si tratta di un'applicazione della cosidetta *formula di Bayes.*

Teorema di Bayes

Teorema 9 (di Bayes)**.** *Sia* $\{A_i\}_{i=1}^{\infty}$ *una Classe Completa di eventi tale che*

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

e B *un qualunque evento con* $\Pr(B) > 0$*. Allora*

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

Dimostrazione 12. *La dimostrazione è immediata usando il Teorema delle Probabilità Totali e viene lasciata per esercizio.*

La *formula di Bayes* è una delle più interessanti ed importanti formule della teoria della probabilità. Essa è dovuta all'ecclesiastico Thomas Bayes. La formula parla della *revisione della probabilità* alla luce delle informazioni sperimentali. Nella terminologia statistica A_1, A_2, $\cdots$ sono dette *ipotesi* e la probabilità $\Pr(A_i)$ è detta *probabilità iniziale* (o a priori) di A_i, $\Pr(B|A_i)$ è detta *verosimiglianza* di A_i ed infine, $\Pr(A_i|B)$ è detta *probabilità finale* (o a posteriori) di A_i.

Esercizio 46 (Associazione Sportiva)**.** *Il* 26% *degli iscritti ad un'associazione sportiva è costituito da maschi. Inoltre, il* 45% *di questi praticano la pallacanestro, mentre tra le femmine solo il* 25%*. Determinare la probabilità che, scegliendo a caso un giocatore di pallacanestro, questi sia maschio.*

Soluzione 31. *Possiamo sintetizzare le informazioni in questo modo:*

A*: giocatore di pallacanestro;*

M*: maschio;*

F*: femmina.*

Inoltre:

- $\Pr(M) = 0.26$*;*
- $\Pr(F) = 1 - \Pr(M) = 0.74$*;*
- $\Pr(A|M) = 0.45$*;*
- $\Pr(A|F) = 0.25$.

La probabilità richiesta si determina ricorrendo alla formula di Bayes

$$\begin{aligned} \Pr(M|A) = & \frac{\Pr(M)\Pr(A|M)}{\Pr(M)\Pr(A|M) + \Pr(F)\Pr(A|F)} \\ = & \frac{0.26\ 0.45}{0.26\ 0.45 + 0.74\ 0.25} = 0.387 \end{aligned}$$

1.5 Indipendenza Stocastica

Indipendenza di due eventi (Indipendenza Stocastica)

Definizione 30. *In uno spazio probabilizzato* $(\Omega, \mathcal{A}, P)$ *due eventi* A, B *si dicono tra loro stocasticamente indipendenti se e solo se*

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

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

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

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

Esempio 9. *Si supponga di lanciare un dado a sei facce e, come di consueto, sia* $\Omega = \{1, 2, 3, 4, 5, 6\}$ *l'insieme degli eventi elementari. Se* $\Pr(\{\omega\}) = \frac{1}{6}$ *per* $\omega \in \Omega$, $A = \{2, 4, 6\}$ *e* $B = \{3, 6\}$, *allora*

$$A \cap B = 6, \ \Pr(A) = \frac{1}{2}, \ \Pr(B) = \frac{1}{3} \ e \ \Pr(A \cap B) = \Pr(A) \cdot \Pr(B) = \frac{1}{6} \ .$$

Quindi gli eventi A *e* B *sono indipendenti.*

Esempio 10. *Si noti che la condizione di indipendenza dipende in maniera cruciale dalla funzione di Probabilità definita sulla classe degli eventi. Infatti, consideriamo la seguente continuazione dell'esempio e supponiamo che* $\Pr(\{\omega\}) = \frac{1}{12}$ *per* $\omega \in \{1, 2, 3, 4\}$ *e* $\Pr(\{\omega\}) = \frac{1}{3}$ *per* $\omega \in \{5, 6\}$, *si ha*

$$\Pr(A) = \frac{1}{2}, \ \Pr(B) = \frac{5}{12} \ e \ \frac{1}{3} = \Pr(A \cap B) \neq \Pr(A) \cdot \Pr(B) = \frac{5}{24} \ .$$

Quindi gli eventi A *e* B ***non*** *sono indipendenti.*

Indipendenza di n eventi

La nozione di indipendenza può essere estesa a più di due eventi nel modo seguente

Definizione 31. *In uno spazio probabilizzato* $(\Omega, \mathcal{A}, P)$ n *eventi* $A_1, A_2, \cdots, A_n$, *si dicono stocasticamente indipendenti se e solo se*

$$\Pr(A_{i_1} \cap A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = \Pr(A_{i_1}) \cdot \Pr(A_{i_2}) \cdots \Pr(A_{i_n}) \ .$$

per ogni $k = 2, 3, \cdots, n$ *e per gli allineamenti* $(i_1 < i_2 < \cdots < i_k)$ *degli interi* $1, 2, \cdots, n$.

Se $A_1, A_2, \cdots$ è una famiglia di eventi, allora A_i $i = 1, 2, \cdots$ si dicono mutuamente indipendenti se ogni n–upla di eventi da essa estratta è indipendente.

Esercizio 47 (Quattro famiglie)**.** *Si considerino quattro famiglie con due figli classificate nel modo seguente (M, M), (M, F), (F, M), (F, F) in cui M indica maschio, F indica femmina e l'ordine di ogni coppia indica l'anzianità. Si sceglie a caso una famiglia e si considerino i seguenti eventi*

- $A_1 =$ *"primo figlio maschio";*
- $A_2 =$ *"due figli di sesso differente";*
- $A_3 =$ *"secondo figlio maschio"; Sono i tre eventi indipendenti?*

Soluzione 32. *Si ha*

$$A_1 \cap A_2 = \{(M, F)\}, \; A_1 \cap A_3 = \{(M, M)\}, \; A_2 \cap A_3 = \{(F, M)\},$$

e

$$A_1 \cap A_2 \cap A_3 = \emptyset \; .$$

Poiché

$$\Pr(A_i) = \frac{1}{2}, \qquad i = 1, 2, 3$$

segue che

$$\Pr(A_i \cap A_j) = \frac{1}{4} = \Pr(A_i) \cdot \Pr(A_j) \qquad i, j = 1, 2, 3, i \neq j \; .$$

Ma

$$\Pr(A_1 \cap A_2 \cap A_3) = 0 \neq \frac{1}{8} = \Pr(A_1) \cdot \Pr(A_2) \cdot \Pr(A_3)$$

e quindi i tre eventi ***non*** *sono indipendenti*

Esercizio 48 (Ancora sul lancio di un dado?)**.** *Si lancia un dado regolare e si considerano i tre eventi seguenti:*

- $A_1 =$ *"realizzare un punteggio pari"*$= \{2, 4, 6\}$*;*
- $A_2 =$ *"realizzare un punteggio maggiore di 2"*$= \{3, 4, 5, 6\}$*;*
- $A_3 =$ *"realizzare un punteggio divisibile per 3 o per 5"*$= \{3, 5, 6\}$*;*

Sono i tre eventi indipendenti?

Soluzione 33. *Allora*

$$\Pr(A_1) = \Pr(A_3) = \frac{1}{2}, \ \Pr(A_2) = \frac{2}{3}$$

e

$$\Pr(A_1 \cap A_2 \cap A_3) = \frac{1}{6} = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{2}{3} = \Pr(A_1) \cdot \Pr(A_2) \cdot \Pr(A_3)$$

Ma, ad esempio, $\Pr(A_2 \cap A_3) = \Pr(A_3) = \frac{1}{2} \neq \Pr(A_2)\ \ \Pr(A_3) = \frac{1}{3}$.

Tribù indipendenti

Definizione 32. *Dato uno spazio probabilizzato* $(\Omega, \mathcal{A}, \Pr)$*. Due Tribù contenute in* $\mathcal{A}$ *si dicono tra loro indipendenti se ogni elemento dell'uno è indipendente da ogni elemento dell'altra.*

Esercizio 49 (Pressione alta)**.** *È stato rilevato che il* 5% *delle persone abitanti in una certa zona ha la pressione alta. Inoltre il* 75% *delle persone con pressione alta beve alcolici, mentre solo il* 50% *delle persone senza pressione alta beve alcolici. Qual è la probabilità che una persona che beve alcolici abbia la pressione alta?*

Esercizio 50 (Ancora sull'estrazione da urne)**.** *L'urna A contiene* 2 *palline bianche e* 2 *nere; l'urna B ne contiene* 3 *bianche e* 2 *nere. Si trasferisce una pallina da A a B e poi si estrae da B una pallina che risulta essere bianca. Qual è la probabilità che fosse bianca anche la pallina trasferita da A a B?*

Esercizio 51 (Ancora sul lancio di un dado)**.** *Un dado bilanciato viene lanciato consecutivamente fino a che non esce la faccia con il* 6 *per la prima volta. Dato che il* 6 *non appare al primo lancio, qual è la probabilità che siano necessari più di 4 lanci?*

Esercizio 52 (A scuola)**.** *In una scuola il* 4% *dei maschi e l'*1% *delle femmine hanno una statura superiore a* 185*cm. Inoltre il* 40% *degli studenti sono maschi. Calcolare la probabilità che uno studente scelto a caso con statura superiore a* 185*cm sia femmina.*

Esercizio 53 (Estrazione da un'urna)**.** *Un'urna contiene n palline nere (N) e b palline bianche (B). Si estrae casualmente una pallina dall urna e, dopo averne osservato il colore, la si rimette nell urna con altre* 2 *palline del colore estratto e* 3 *del colore non estratto. Calcolare la probabilità che in* 4 *estrazioni successive e effettuate secondo la regola sopra stabilita, si ottenga la stringa ordinata $BNNB$.*

Esercizio 54 (Estrazione da un'urna)**.** *Un'urna contiene* 5 *palline bianche e* 6 *nere. Si estraggono due palline in blocco dall'urna, se ne registra il colore e si reinseriscono le due palline nell'urna assieme ad altre due palline dello stesso colore di quelle estratte. Viene poi effettuata un'altra estrazione di due palline in blocco. Detto A l'evento che si verifica se le due palline estratte alla seconda estrazione sono entrambe bianche:*

- *calcolare la probabilità dell'evento A;*
- *condizionatamente al fatto che A si è verificato, calcolare la probabilità che le due palline alla prima estrazione siano entrambe nere.*

Esercizio 55 (Tavolo e urne)**.** *Sul tavolo ci sono due urne: la prima contiene* 2 *palline nere e* 5 *bianche e la seconda contiene* 3 *palline nere e* 2 *bianche. Si sceglie a caso un'urna, si estrae una pallina e la si depone nell'altra urna. Da quest'ultima si procede dunque all'estrazione di un altra pallina.*

- *Qual è la probabilità di estrarre due palline bianche?*
- *Qual è la probabilità che la seconda pallina sia nera?*

Esercizio 56 (Partizione)**.** *C_1 e C_2 sono una partizione di Ω e hanno la stessa probabilità. Se l'evento A è tale che $\Pr(A|C_1) = \Pr(A|C_2) = 1/2$, si calcoli la $\Pr(C_1|A)$.*

Esercizio 57 (Tre monete)**.** *Tre monete hanno rispettivamente due facce bianche, una faccia bianca e una nera e due facce nere. Se ne sceglie una a caso e la si lancia. Qual è la probabilità che esca una faccia bianca? Se è uscita una faccia bianca, qual è la probabilità che anche l'altra sia bianca?*

Esercizio 58 (Fornitori di pneumatici)**.** *Due fornitori A e B di pneumatici per una fabbrica di automobili hanno rispettivamente* 0.3% *e* 0.8% *di pezzi difettosi nella loro produzione. Inoltre A fornisce il* 60% *del totale degli pneumatici acquistati dalla fabbrica e B il* 40%*.*

- *Qual è la probabilità che uno pneumatico scelto a caso dal magazzino della fabbrica risulti difettoso?*
- *Avendo trovato un pezzo difettoso, qual è la probabilità che sia stato fornito da A?*

Esercizio 59 (Prigionieri e porte)**.** *Un prigioniero è rinchiuso in una cella con* 3 *porte, A, B e C. La porta A riporta il prigioniero in cella dopo* 2 *giorni di lavori forzati; la porta B lo riporta in cella dopo* 3 *giorni di lavori forzati; infine la porta C ridà al prigioniero la libertà. Il prigioniero sceglie la porta* *da prendere lanciando un dado equilibrato: se il risultato è pari sceglie* A*, se esce il numero* 1 *sceglie* B *e nei rimanenti casi sceglie* C*. Se il prigioniero torna in cella, sceglie in modo equiprobabile fra le due porte non ancora scelte. Qual è la probabilità che il prigioniero impieghi* 3 *oppure* 5 *giorni per uscire?*

1.6 Variabili Aleatorie

Variabili aleatorie

Consideriamo il seguente problema:

I costi da sostenere per produrre un'attrezzatura per una macchina utensile è di 1000*euro*. Ci sono state ordinate 3 attrezzature al prezzo di 2000*euro* cadauna. Nel processo di produzione vi è un'operazione che può fallire con probabilità $\Pr = 0.1$. Quando ciò accade le materie prime devono essere gettate e si deve ricominciare il processo dall'inizio.

- Qual è la probabilità di avere un quadagno?
- Se l'ordine viene ripetuto ogni mese, qual è il quadagno medio che ci aspettiamo?

Attrezzatura per macchine utensili

Per prima cosa dobbiamo considerare quante volte dobbiamo ricominciare il processo per ottenere 3 attrezzature funzionanti.

A questo scopo consideriamo $\Omega_I = \{3, 4, 5, \cdots\}$ dove ogni elemento è il numero di volte che abbiamo dovuto iniziare il procedimento, ad esempio 4 significa che 3 processi sono andati a buon fine (di cui 1 è l'ultimo) e 1 no.

La funzione di probabilità è

- $\Pr_I(3) = (1-p)^3 = 0.729;$
- $\Pr_I(4) = \binom{3}{1}(1-p)^3 \; p = 0.2187;$
- $\Pr_I(5) = \binom{4}{2}(1-p)^3 \; p^2 = 0.04374;$
- $\cdots$

e in generale

$$\begin{aligned} \Pr_I(i) &= \binom{i-1}{i-3}(1-p)^3 \; p^{i-3} \qquad i = 3, 4, 5, \cdots \\ &= \frac{(i-1)(i-2)}{2!}(1-p)^3 \; p^{i-3} \qquad i = 3, 4, 5, \cdots \end{aligned}$$

Che può essere rappresentata attraverso il seguente grafico

È chiaro, che il guadagno (G) è legato al numero di volte che dobbiamo fare il processo (I),

$$g = 6000 - 1000 \cdot i$$

cioè

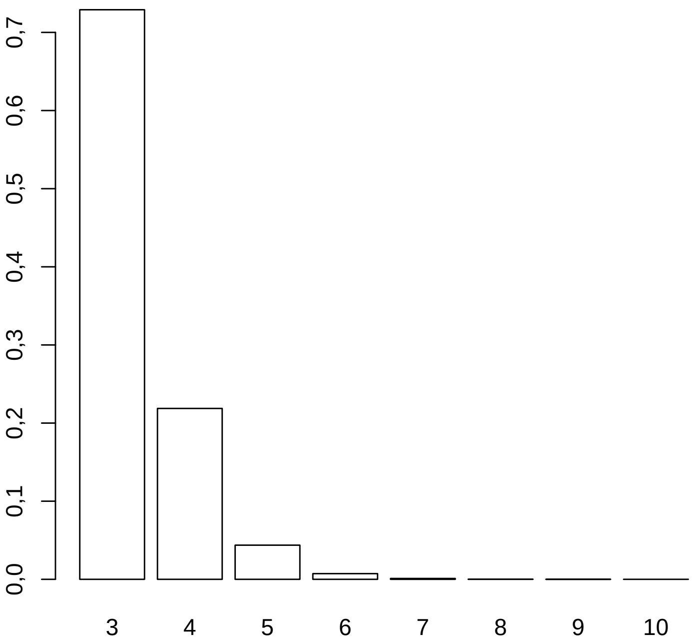


Figura 1.1:

i	3	4	5	6	7	$i > 7$
$\Pr_I(I = i)$	0.729	0.2187	0.04374	0.00729	0.0010935	0.0001765
g	3000	2000	1000	0	-1000	$g < -1000$

da cui abbiamo

$$\Pr_G(G = g) = \Pr_I(I = \frac{6000 - g}{1000})$$

Infine per rispondere al nostro primo quesito

$$\Pr_G(G > 0) = \sum_{i=3}^{5} \Pr_I(i) = 0.99144$$

Variabili aleatorie

Il quadagno G, appartiene allo spazio campionario $\Omega_G = \{3000, 2000, 1000, 0, -1000, \cdots\}$. Il fenomeno che abbiamo studiato non è definito nello spazio $(\Omega_I, \mathcal{A}_I, \Pr_I)$ ma in uno diverso diciamo $(\Omega_G, \mathcal{A}_G, \Pr_G)$ che è stato definito a partire dal primo (abbiamo per ora tralasciato la definizione di $\mathcal{A}_G$).

In molte situazioni, che ricorrono nelle applicazioni, abbiamo che lo spazio campionario in cui avviene l'esperimento aleatorio può essere rappresentato da oggetti concreti (famiglie, individui) di natura **non numerica**. Si pensi, ad esempio, all'esperimento del duplice lancio di una moneta i cui possibili risultati sono $\{(T, T), (T, C), (C, T), (C, C)\}$.

Tuttavia, più che all'esperimento in sé, si è spesso interessati alle *conseguenze*, numericamente valutate.

È quindi molto importante studiare (e formalizzare) come sono legati i due spazi (nel nostro esempio) $(\Omega_I, \mathcal{A}_I, \Pr_I)$ e $(\Omega_G, \mathcal{A}_G, \Pr_G)$.

A tale scopo risulta fondamentale il concetto di **Variabile Aleatoria**.

Definizione 33. *Sia dato lo spazio probabilizzabile* $(\Omega, \mathcal{A})$. *Si dice* ***Variabile aleatoria*** *(v.a.) ogni funzione a valori reali definita in* Ω, $y = X(\omega)$, *tale che*

$$\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{A} \text{ per ogni valore reale } x.$$

- Giova osservare che nella definizione la probabilità non gioca alcun ruolo e che quando $\mathcal{A}$ è la classe di tutti i sottoinsiemi di Ω la condizione nella definizione è sempre soddisfatta.

- Per rendersi conto della necessità di imporre alla funzione $X(\omega)$ la condizione riportata sopra basterà dire che, intendendo assegnare una probabilità agli insiemi $\{\omega \in \Omega : X(\omega) \leq x\}$ per ogni reale x ed avendo probabilizzato la classe $\mathcal{A}$, occorre che tali insiemi appartengano ad $\mathcal{A}$.

Variabili aleatorie e Tribù

Esercizio 60. *Trovare un esempio di funzione* $f : \tilde{\Omega} \to \Omega$ *e una Tribù* $\tilde{\mathcal{A}}$ *su* $\tilde{\Omega}$ *tale che*

$$\mathcal{A} = \{f(A) : A \in \tilde{\mathcal{A}}\}$$

non è una Tribù su Ω.

Soluzione 34. *Si consideri ad esempio* $\tilde{\Omega} = \{1, 2, 3\}$ *e* $\Omega = \{1, 2\}$, *la Tribù*

$$\tilde{\mathcal{A}} = \{\emptyset, \{1\}, \{2, 3\}, \{1, 2, 3\}\}$$

e la funzione f tale che

$$f(1) = f(2) = 1 \qquad f(3) = 2$$

Allora

$$\mathcal{A} = \{f(A) : A \in \tilde{\mathcal{A}}\} = \{\emptyset, \{1\}, \{1, 2\}\}$$

che non è nemmeno un'algebra perché non ha 2^n elementi

Variabili aleatorie e Tribù

Teorema 10. *Siano $\tilde{\Omega}$ e Ω due insiemi arbitrari e sia $X : \tilde{\Omega} \to \Omega$ una funzione. Se $\mathcal{A}$ è una Tribù su Ω allora*

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

è una Tribù su $\tilde{\Omega}$.

Dimostrazione 13. *Infatti*

- $X^{-1}(\Omega) = \tilde{\Omega}$;
- $X^{-1}(\Omega/A) = \tilde{\Omega}/X^{-1}(A)$;
- $X^{-1}(\cup_{i=1}^{\infty} A_i) = \cup_{i=1}^{\infty} X^{-1}(A_i)$;

Teorema 11. *Siano $\tilde{\Omega}$ e Ω due insiemi arbitrari e sia $X : \tilde{\Omega} \to \Omega$ una funzione. Se $\tilde{\mathcal{A}}$ è una Tribù su $\tilde{\Omega}$ allora*

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

è una Tribù su Ω.

Dimostrazione 14. *Infatti*

- *Siccome $\tilde{\Omega} \in \tilde{\mathcal{A}}$ e $\tilde{\Omega} = X^{-1}(\Omega)$ allora $\Omega \in \mathcal{A}$;*
- *Supponiamo che $A \in \mathcal{A}$ allora $X^{-1}(A) \in \tilde{\mathcal{A}}$ che implica $X^{-1}(\Omega/A) = \tilde{\Omega}/X^{-1}(A) \in \tilde{\mathcal{A}}$ e quindi $\Omega/A \in \mathcal{A}$;*
- *Supponiammo che $A_1, A_2, \cdots \in \mathcal{A}$. Questo significa che $X^{-1}(A_1), X^{-1}(A_2), \cdots \in \tilde{\mathcal{A}}$ così*

  $$X^{-1}(\cup_{i=1}^{\infty} A_i) = \cup_{i=1}^{\infty} X^{-1}(A_i) \in \tilde{\mathcal{A}}$$

  e quindi $\cup_{i=1}^{\infty} A_i \in \mathcal{A}$

Appurare che una data funzione è una variabile aleatoria non è sempre agevole, tuttavia nell'importante caso in cui $\Omega = \mathbb{R}$ vale il seguente teorema

Teorema 12. *Ogni funzione continua oppure monotòna crescente o decrescente $f : (\mathbb{R}, \mathcal{B}(\mathbb{R})) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ è una variabile aleatoria.*

Variabili aleatorie e funzioni di probabilità

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

$$\text{Probabilità di } (a < X \leq b) = \Pr(X \in (a, b]), \ -\infty \leq a < b < +\infty \ .$$

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

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

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

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

La funzione di probabilità P_X, definita sulla classe di Borel di $\mathbb{R}$, è nota col nome di *distribuzione* della v.a. X e mediante essa sarà possibile determinare $\Pr_X(B) = \Pr(X \in B)$ per ogni $B \in \mathcal{B}(\mathbb{R})$.

Variabili aleatorie discrete

Definizione 34. *Una v.a. X definita su $(\Omega, \mathcal{A})$ è detta discreta se i valori distinti dell'insieme $\cup_{\omega \in \Omega}\{X(\omega)\}$ costituiscono un insieme R_X finito o numerabile*

Esempio 11. *L'esperimento consiste nel registrare il numero di biciclette che transitano per un certo tratto di strada nell'arco di un periodo temporale. I risultati sperimentali saranno $\Omega = \{0, 1, 2, \cdots\}$. La funzione $y = X(\omega) = \omega$, $\omega \in \Omega$ è una v.a. con $R_X = \Omega$*

Per scrivere la distribuzione di una v.a. discreta risulta comoda la *funzione di probabilità* (o *densità discreta*) di cui alla seguente

Definizione 35 (funzione di probabilità (o densità discreta))**.** *Se X è una v.a. discreta con $R_X = \{x_1, x_2, \cdots\}$, allora la funzione, definita in $\mathbb{R}$, data da*

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

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

Teorema 13. *Se X è una v.a. discreta con $R_X = \{x_1, x_2, \cdots\}$ allora*

$$\mathrm{p}(x) \geq 0 \textit{ per ogni } x \textit{ reale e } \sum_{x \in R_X} \mathrm{p}(x) = 1$$

Questo significa che qualsiasi funzione che abbia le proprietà descritte dal teorema precedente è una funzione di probabilità e può servire alla descrizione probabilistica di una v.a. discreta. In particolare ci sono delle funzioni che ricorrono spesso nelle applicazioni per cui vale la pena di studiarle in dettaglio. Nel seguito ne vedremo alcune.

Distribuzione Binomiale

Questa distribuzione di probabilità regola il numero dei successi (o risultati favorevoli) conseguito in una successione (finita) di prove indipendenti.

Si supponga che un certo esperimento venga replicato $N \geq 1$ volte e l'esito di ognuno di essi possa essere *favorevole* (evento A) oppure *non favorevole* (evento A^c). Ad ogni prova dell'esperimento associamo una v.a. X_i, $i = 1, 2, \cdots, N$ che ne rappresenta l'esito: $X_i = 1$ se si verifica A e $X_i = 0$ se si verifica A^c. Si supponga che le v.a. X_i, $i = 1, 2, \cdots, N$ siano indipendenti e che $\Pr(X_i = 1) = p$, $0 \leq p \leq 1$. Qual è la distribuzione di probabilità del numero *totale* di successi nelle N prove? Cioè qual è la distribuzione di probabilità della v.a.

$$X = X_1 + X_2 + X_3 + \cdots + X_N \ ?$$

Definizione 36 (Distribuzione Binomiale)**.** *Si dice che una v.a. X si distribuisce secondo la distribuzione di probabilità (o legge) binomiale di parametri $N \geq 1$ (intero) e $0 \leq p \leq 1$, se*

$$\Pr(X = x) = \begin{cases} \binom{N}{x} p^x (1-p)^{N-x} & x = 0, 1, \cdots, N \\ 0 & \textit{altrove} \ . \end{cases}$$

e scriveremo $X \sim \mathrm{Bi}(N, p)$.

Nel caso speciale in cui $N = 1$ la v.a. è chiamata anche v.a. di Bernoulli.

Teorema 14. *La funzione di probabilità del numero totale di successi ottenuti in N prove indipendenti con probabilità di successo pari a p ad ogni prova è dato dalla distribuzione di probabilità di una v.a. binomiale*

Dimostrazione 15. *Si consideri una particolare realizzazione per cui il numero di successi è x, ad esempio l'allineamento*

$$B = \{\underbrace{A, A, \cdots, A}_{x \textit{ volte}} \underbrace{A^c, A^c, \cdots, A^c}_{N-x \textit{ volte}}\}$$

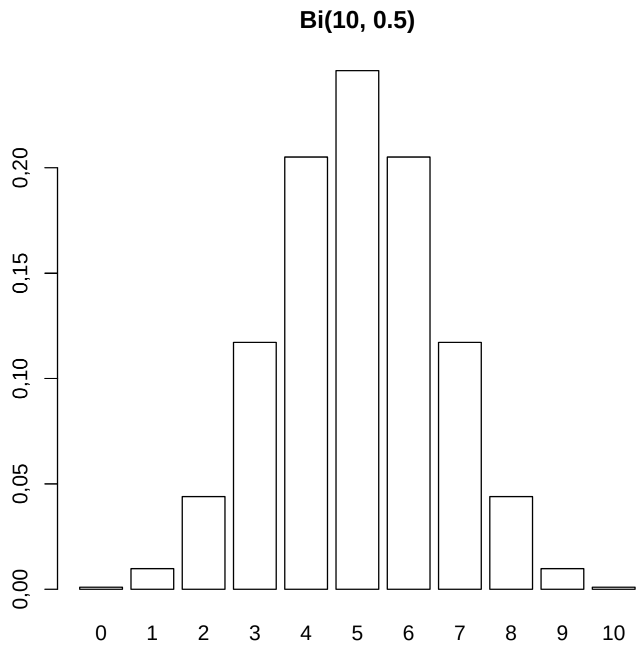


Figura 1.2: Binomiale $n = 10$, $p = 0.5$, funzione di probabilità.

Poiché le prove sono indipendenti la probabilità di questo allineamento è

$$\Pr(B) = p^x \ (1-p)^{N-x}$$

e questa è la probabilità di un qualsiasi allineamento con x successi che sono

$$\binom{N}{x}$$

e quindi

$$\Pr(X = x) = \binom{N}{x} p^x \ (1-p)^{N-x}$$

Teorema 15. *Siano $X_1, X_2, \cdots, X_N$, N v.a. Bernoulliane $Bi(1,p)$ stocasticamente indipendenti allora la v.a. X così definita*

$$X = \sum_{i=1}^{n} X_i$$

è tale che $X \sim Bi(N,p)$

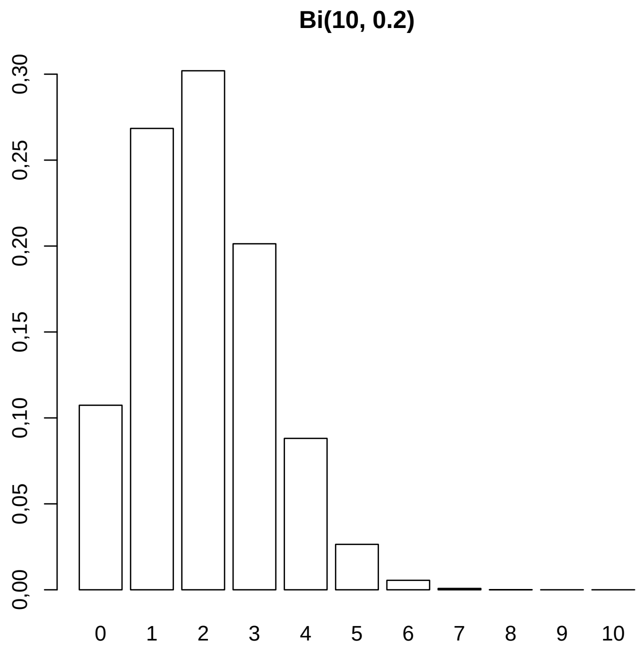


Figura 1.3: Binomiale $n = 10$, $p = 0.2$, funzione di probabilità.

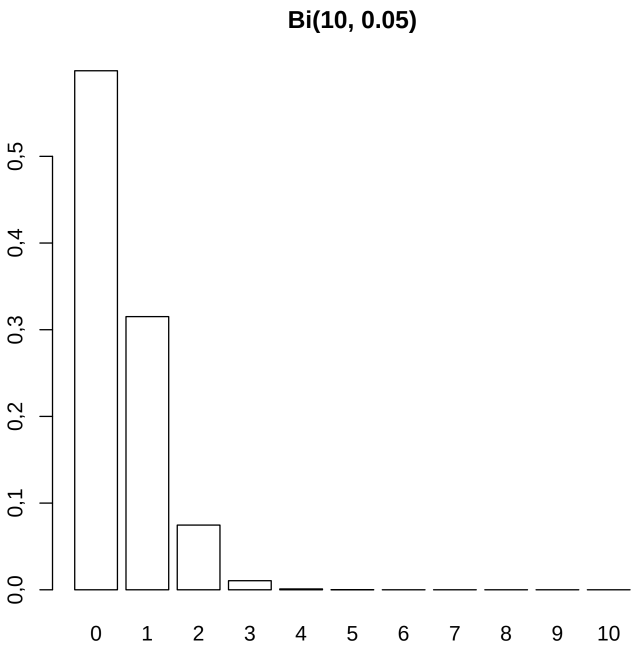


Figura 1.4: Binomiale $n = 10$, $p = 0.0.05$, funzione di probabilità.

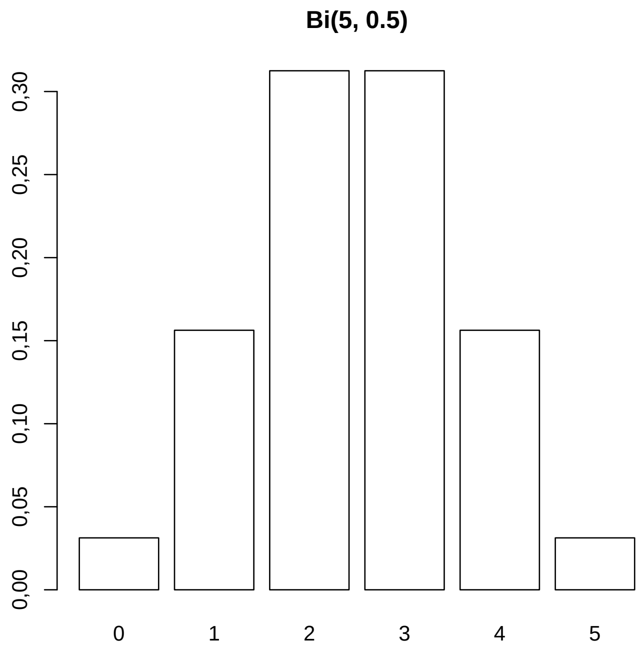


Figura 1.5: Binomiale $n = 5$, $p = 0.5$, funzione di probabilità.

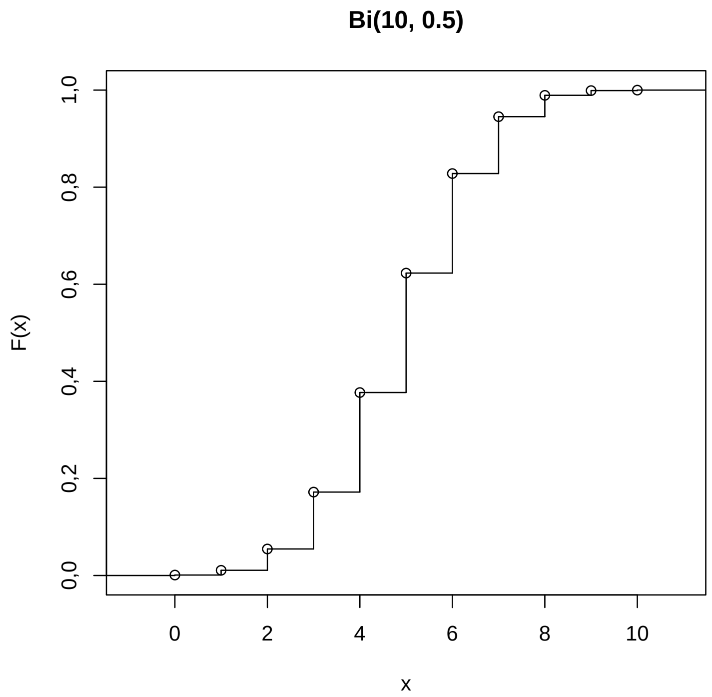


Figura 1.6: Binomiale $n = 10$, $p = 0.5$, funzione di ripartizione.

Funzione di Ripartizione

Definizione 37. *Sia X una v.a.. Si dice funzione di ripartizione della v.a. X la funzione $y = F(x)$, definita per ogni x reale, data da*

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

Si noti la somiglianza della funzione di ripartizione con la funzione che abbiamo utilizzato per probabilizzare lo spazio $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Distribuzione Binomiale

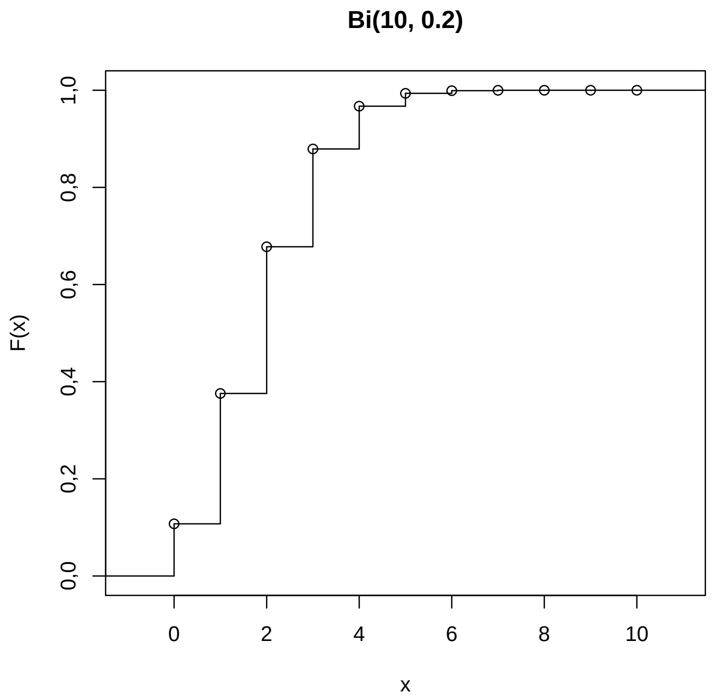


Figura 1.7: Binomiale $n = 10$, $p = 0.2$, funzione di ripartizione.

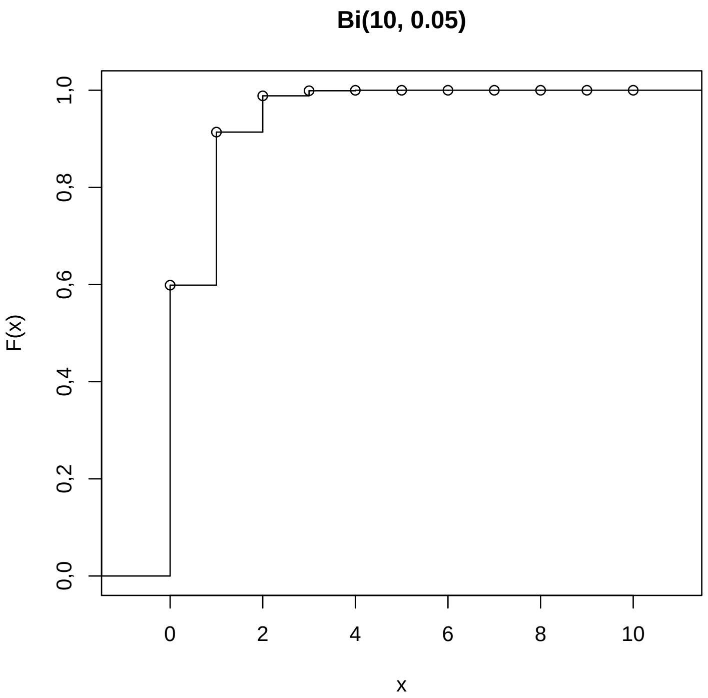


Figura 1.8: Binomiale $n = 10$, $p = 0.05$, funzione di ripartizione.

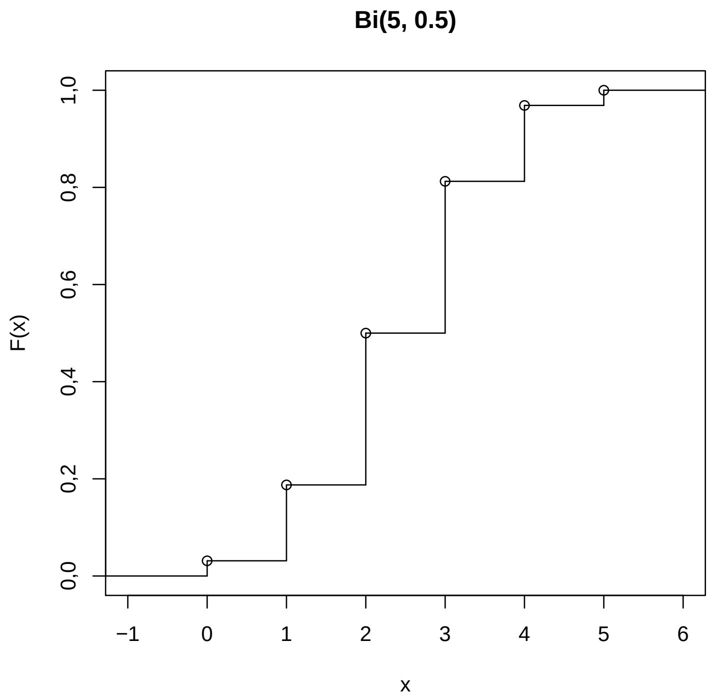


Figura 1.9: Binomiale $n = 5$, $p = 0.5$, funzione di ripartizione.

Funzione di ripartizione e funzione di probabilità

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

$$\mathrm{p}(X = x) = F(x) - F(x^-)$$

Esercizio 61. *Si calcoli la funzione di ripartizione di una v.a.* $Bi(10, 0.5)$ *e si controlli che vale la relazione sopra riportata.*

Distribuzione Geometrica

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

Definizione 38 (Distribuzione Geometrica)**.** *Si dice che una v.a.* X *si distribuisce secondo una distribuzione geometrica di parametro* $0 \leq p \leq 1$ *se la sua funzione di probabilità è*

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

e scriverremo $X \sim Ge(p)$*.*

La funzione di ripartizione di una v.a. Geometrica è data da

$$\begin{aligned} F(x) = & \quad \Pr(X \leq x) = \sum_{t=1}^{[x]} \Pr(X = t) = p \ \sum_{t=1}^{[x]} (1-p)^{t-1} \\ = & \quad \begin{cases} p\frac{1-(1-p)^{[x]}}{1-(1-p)} = 1 - (1-p)^{[x]} & x \geq 1 \\ 0 & x < 1 \end{cases} \end{aligned}$$

dove $[\cdot]$ indica la parte intera.

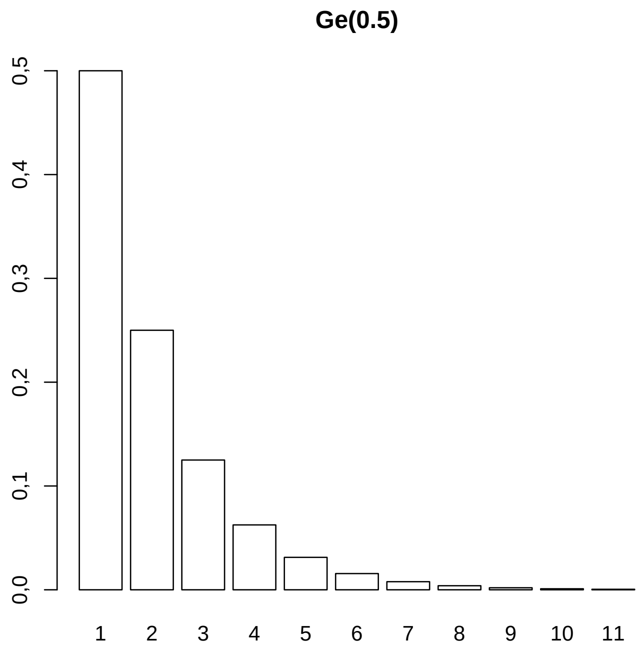


Figura 1.10: Geometrica $p = 0.5$, funzione di probabilità.

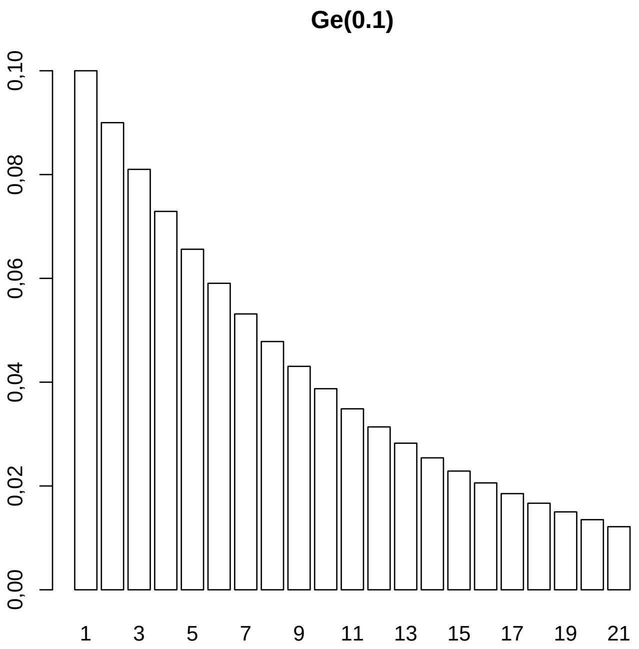


Figura 1.11: Geometrica $p = 0.1$, funzione di probabilità.

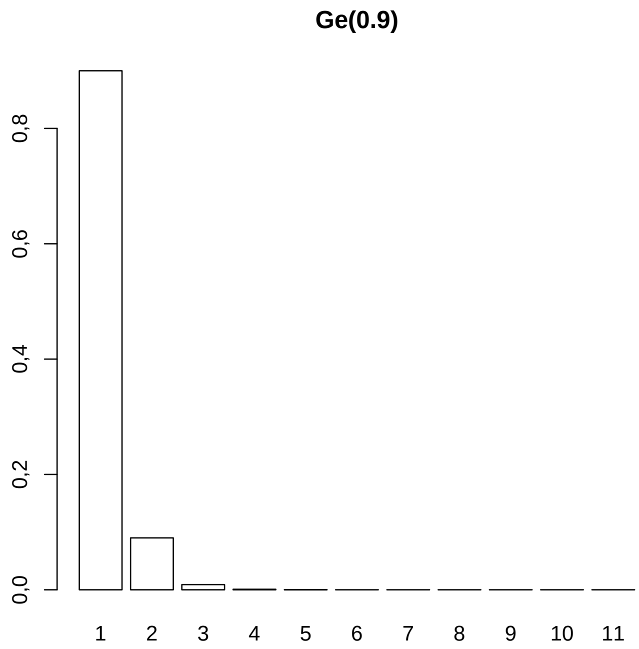


Figura 1.12: Geometrica $p = 0.9$, funzione di probabilità.

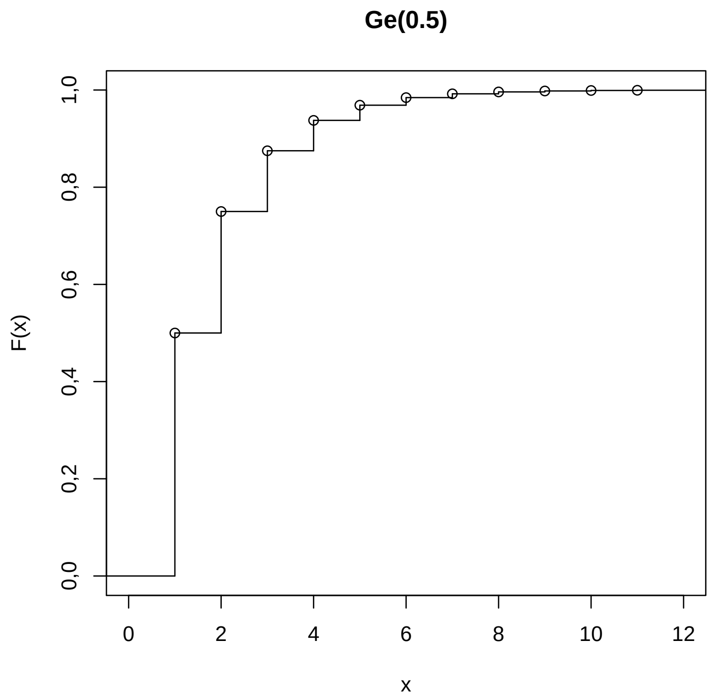


Figura 1.13: Geometrica $p = 0.5$, funzione di ripartizione.

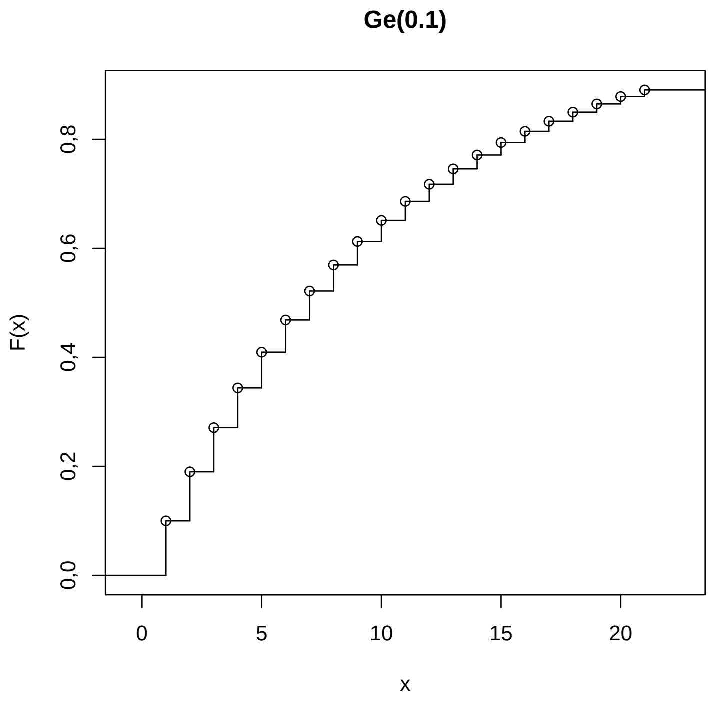


Figura 1.14: Geometrica $p = 0.1$, funzione di ripartizione.

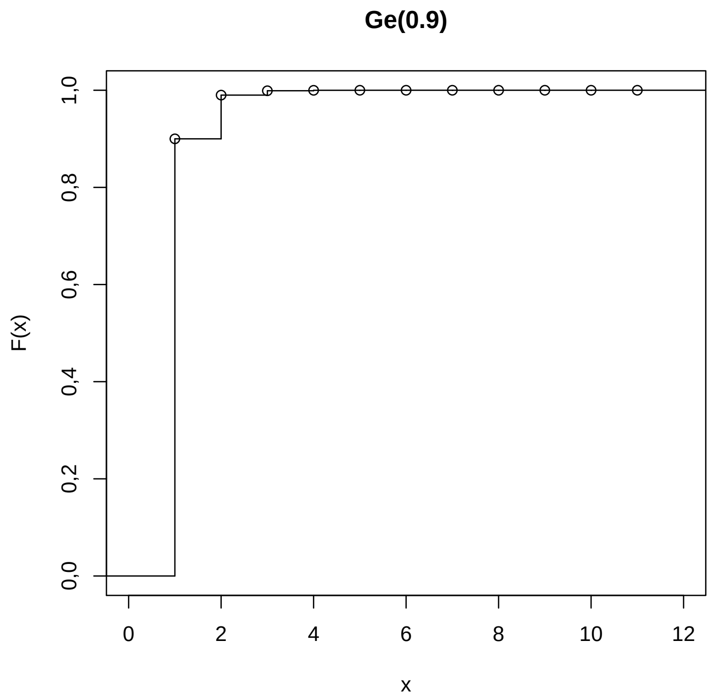


Figura 1.15: Geometrica $p = 0.9$, funzione di ripartizione.

Esempio 12 (Distribuzione Geometrica)**.** *Un lotto di lampadine contiene N lampadine di cui $0 < M < N$ difettose. Estraendo con reimmissione, qual è la probabilità che una lampadina difettosa appaia per la prima volta alla x–esima estrazione?*

Soluzione 35. *Questo problema può essere risolto considerando la v.a. Geometrica $X \sim Ge(p = \frac{M}{N})$ e quindi*

$$\Pr(X = x) = \frac{M}{N}\left(1 - \frac{M}{N}\right)^{(x-1)}$$

Esempio 13 (Distribuzione Geometrica)**.** *Una proprietà interessante della distribuzione Geometrica è quella dell'assenza di memoria. Infatti continuando l'esempio precedente*

Si supponga di sapere che nelle prime y estrazioni non è comparsa una lampadina difettosa. Qual è la probabilità che essa si presenti per la prima volta dopo ulteriori $r > 0$ estrazioni?

Soluzione 36. *Dobbiamo valutare*

$$\Pr(X = r + y | X > y) = \frac{\Pr(\{X = r + y\} \cap \{X > y\})}{\Pr(\{X > y\})}$$

ma poiché

$$\Pr(\{X > y\}) = 1 - \Pr(\{X \leq y\}) = (1 - p)^y > 0$$

e $\{X = r + y\} \subseteq \{X > y\}$ quindi

$$\begin{aligned} \Pr(X = r + y | X > y) = & \quad \frac{\Pr(\{X = r + y\})}{(1 - p)^y} = \frac{p(1 - p)^{y+r-1}}{(1 - p)^y} \\ = & \quad p(1 - p)^{r-1} = \Pr(X = r)\ . \end{aligned}$$

Distribuzione Binomiale negativa

Definizione 39 (Distribuzione Binomiale negativa (o di Pascal))**.** *Si dice che una v.a. X si distribuisce secondo la distribuzione binomiale negativa di parametri $0 < p \leq 1$ e $r \geq 1$ (intero) se la sua funzione di probabilità è data da*

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

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

Si noti che per $r = 1$ si ottiene la distribuzione Geometrica.

Relazione tra la distribuzione Binomiale e la distribuzione Binomiale negativa

Riportiamo questo teorema senza dimostrazione (si veda ad esempio, Cifarelli (1998), Introduzione al Calcolo delle Probabilità, pag. 262, teorema n. 3).

Teorema 16 (Distribuzione Binomiale negativa (o di Pascal))**.** *Sia* $X \sim BiNe(r, p)$ *e* $Z \sim Bi(N, p)$ *allora*

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

Distribuzione Binomiale negativa

Teorema 17. *Siano* $X_1, X_2, \cdots, X_r$*,* r *v.a. Geometriche* $Ge(p)$ *stocasticamente indipendenti, dove* X_1 *indica il numero di prove per avere il primo successo,* X_2 *indica l'ulteriore numero di prove per avere il secondo successo, allora la v.a.* X *così definita*

$$X = \sum_{i=1}^{n} X_i$$

è tale che $X \sim BiNe(r, p)$

Variabili aleatorie continue

Definizione 40. *Una v.a.* X *definita su* $(\Omega, \mathcal{A})$ *è detta continua se la sua funzione di ripartizione è continua.*

Densità

Definizione 41 (Densità)**.** *Si dice che la v.a.* X *è dotata di densità se la probabilità con cui* X *assume valori nell'intervallo* $(a, b]$ *è data mediante la formula*

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

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

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

Variabili aleatorie assolutamente continue

Definizione 42. *Una v.a. X definita su $(\Omega, \mathcal{A})$ è detta assolutamente continua se la sua funzione di ripartizione è continua e la v.a. X ammette densità.*

Densità e funzione di Ripartizione

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

$$\begin{aligned}
\Pr(X \in (a, b]) &= \int_a^b f(x)\ dx \\
&= \int_{-\infty}^b f(x)\ dx - \int_{-\infty}^a f(x)\ dx \\
&= \Pr(X \le b) - \Pr(X \le a) \\
&= F(b) - F(a)
\end{aligned}$$

Infine si noti che quando $F(x)$ è derivabile allora (dal teorema fondamentale del calcolo integrale)

$$f(x) = F'(x) = \frac{\partial}{\partial x} F(x)$$

e

$$F(x) = \int_{-\infty}^x f(t)\ dt$$

Variabili aleatorie dotate di Densità

È una caratteristica delle v.a. dotate di densità che per ogni $x \in \mathbb{R}$ is abbia $\Pr(X = x) = 0$. Infatti, dalla proprietà di continuità di una funzione di probabilità, se a_n è una qualsiasi successione crescrente con $a_n \to x$ per $n \to \infty$, allora

$$\begin{aligned}
\Pr(X = x) &= \Pr(\lim_{n\to\infty} (a_n, x]) \\
&= \lim_{n\to\infty} \Pr((a_n, x]) \\
&= \lim_{n\to\infty} \int_{a_n}^x f(x)\ dx = 0
\end{aligned}$$

Esempio 14 (v.a. continue dotate di densità)**.** *Sia* $\Omega = [0,1]$ *e* $\mathcal{A} = \mathcal{B}([0,1])$ *la Tribù associata. Sia* Pr *la funzione di probabilità su* $\mathcal{A}$ *definita dalla funzione* $F(x)$

$$F(x) = \begin{cases} 0 & x < 0 \\ x & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$$

Soluzione 37. *Consideriamo la v.a.* $X(\omega) = \omega$ *e quindi*

$$\begin{aligned} \Pr(a < X \leq b) &= \Pr(\{\omega \in [0,1] : a < X(\omega) \leq b\}) \\ &= \Pr(\{\omega \in [0,1] : a < \omega \leq b\}) \\ &= \Pr((a,b]) \\ &= \begin{cases} b - a = \int_a^b 1 \; dx & 0 \leq a < b \leq 1 \\ b & a < 0 < b \\ 0 & a < b < 0 \end{cases} \end{aligned}$$

La v.a. X *è dotata di densità*

$$f(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \textit{altrove} \end{cases}$$

Consideriamo sullo stesso spazio la v.a. $Y(\omega) = \omega^2$ *allora abbiamo*

$$\begin{aligned} \Pr(a < Y \leq b) &= \Pr(\{\omega \in [0,1] : a < Y(\omega) \leq b\}) \\ &= \Pr(\{\omega \in [0,1] : a < \omega^2 \leq b\}) \\ &= \Pr(\{\omega \in [0,1] : \sqrt{a} < \omega \leq \sqrt{b}\}) \\ &= \Pr((\sqrt{a}, \sqrt{b}]) \qquad = \begin{cases} \sqrt{a} - \sqrt{b} = \int_a^b \frac{1}{2\sqrt{x}} \; dx & 0 \leq a < b \leq 1 \\ \sqrt{b} & a < 0 < b \\ 0 & a < b < 0 \end{cases} \end{aligned}$$

La v.a. Y *è dotata di densità*

$$f(x) = \begin{cases} \frac{1}{2\sqrt{x}} & 0 \leq x \leq 1 \\ 0 & \textit{altrove} \end{cases}$$

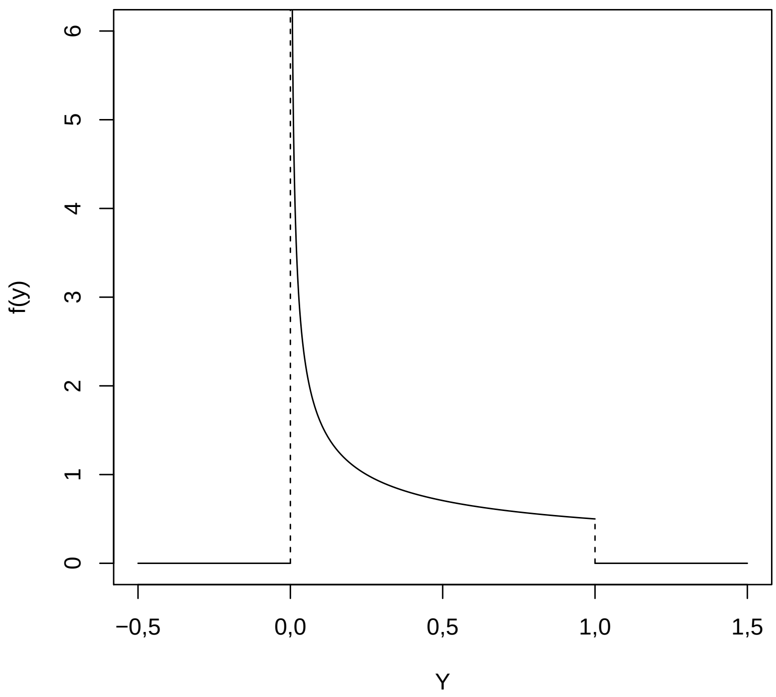


Figura 1.16: .

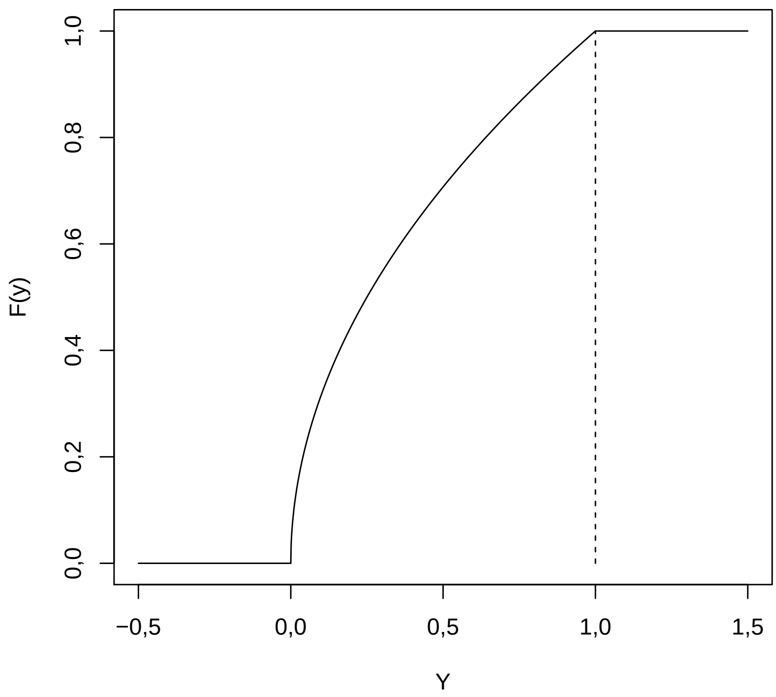


Figura 1.17: .

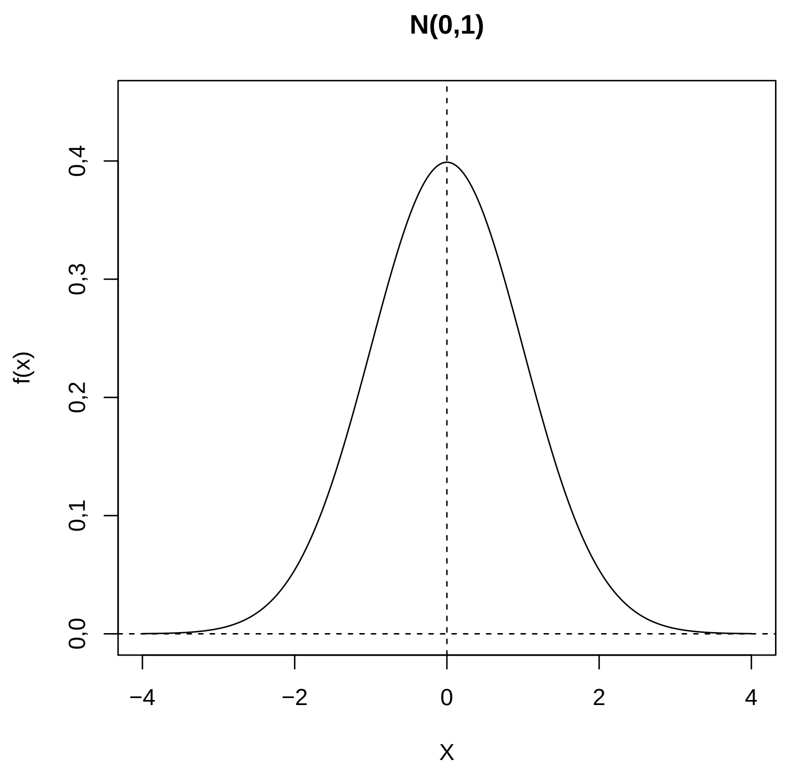


Figura 1.18: Densità della v.a. Normale Standard ($\mu = 0$, $\sigma^2 = 1$).

Distribuzione Normale (o di Gauss)

Definizione 43 (Distribuzione Normale)**.** *Si dice che una v.a.* X *si distribuisce con legge di probabilià Normale (o Gaussiana) di parametri* $-\infty < \mu < +\infty$ *e* $0 < \sigma < +\infty$ *se possiede la seguente densità*

$$f(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right) \qquad -\infty < x < +\infty$$

e la indichiamo con $X \sim \mathrm{N}(\mu, \sigma^2)$

La distribuzione Normale è quella che più di ogni altra trova applicazione nella metodologia statistica. Essa può essere utilizzata per approssimare la distribuzione di molti fenomeni presenti in natura e gioca un ruolo fondamentale nei risultati teorici.

La v.a. $X \sim \mathrm{N}(0, 1)$ è chiamata *Normale standard*.

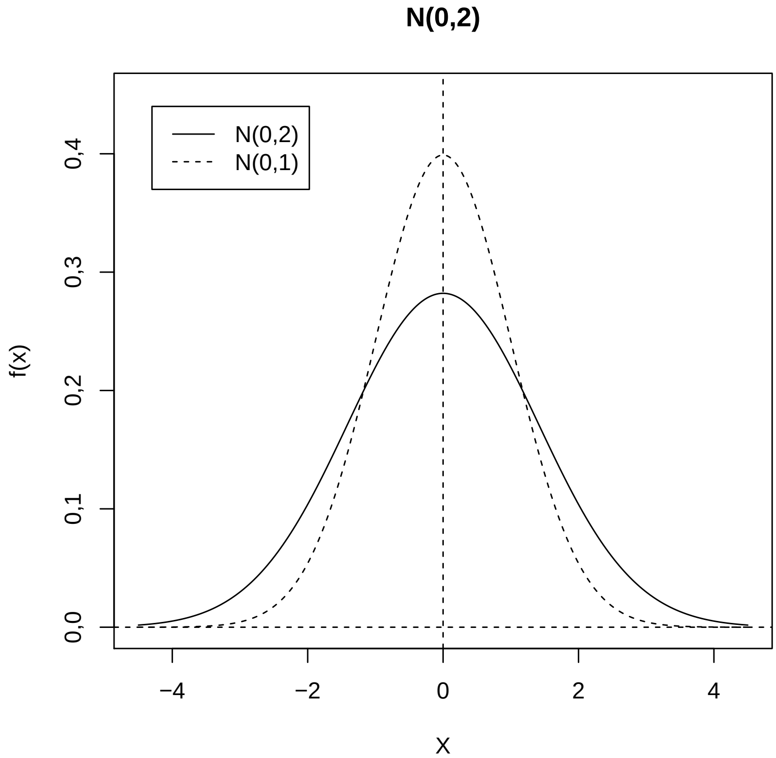


Figura 1.19: Densità della v.a. Normale $\mu = 0$, $\sigma^2 = 2$.

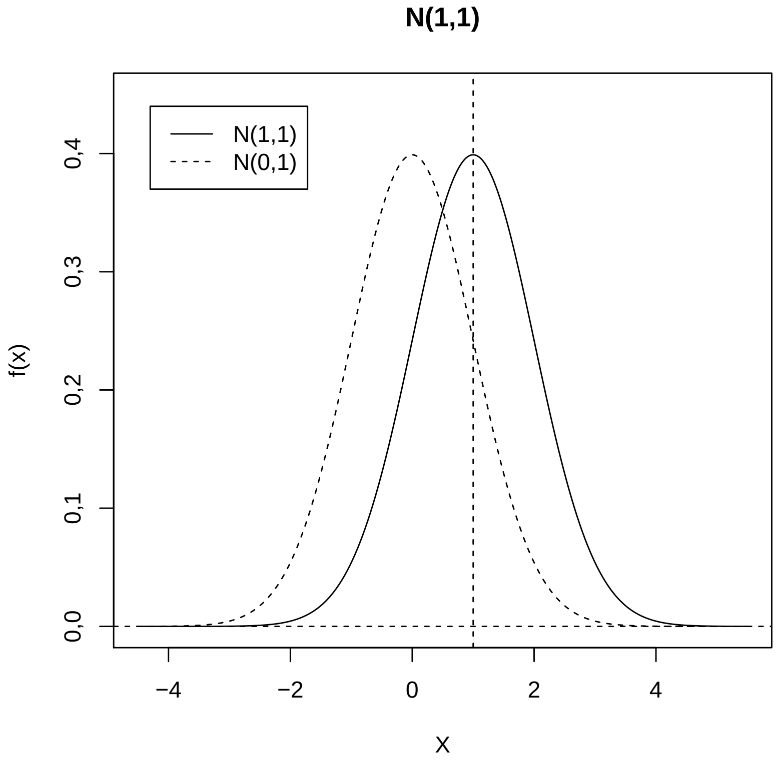


Figura 1.20: Densità della v.a. Normale $\mu = 1$, $\sigma^2 = 1$.

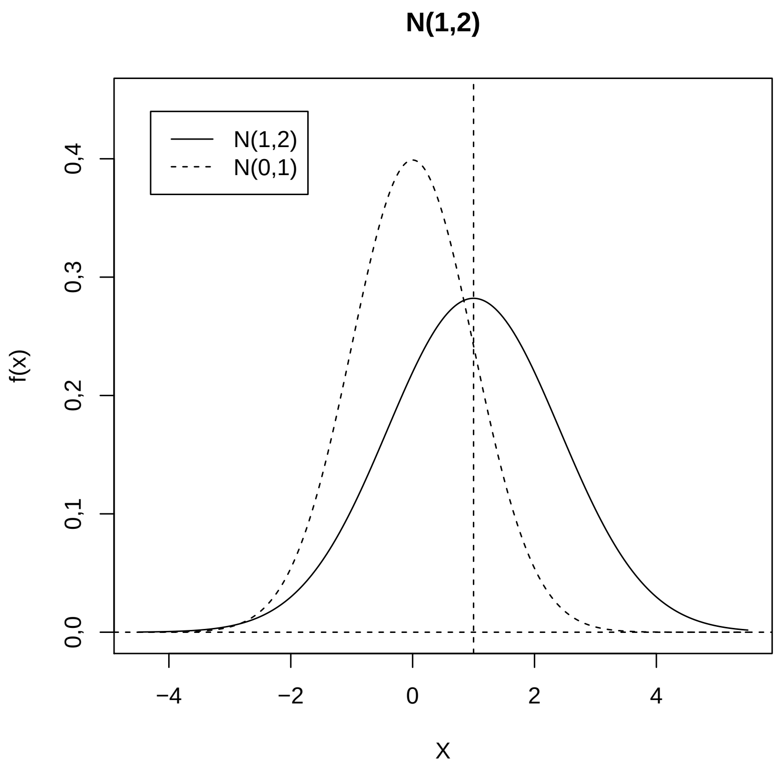


Figura 1.21: Densità della v.a. Normale $\mu = 1$, $\sigma^2 = 2$.

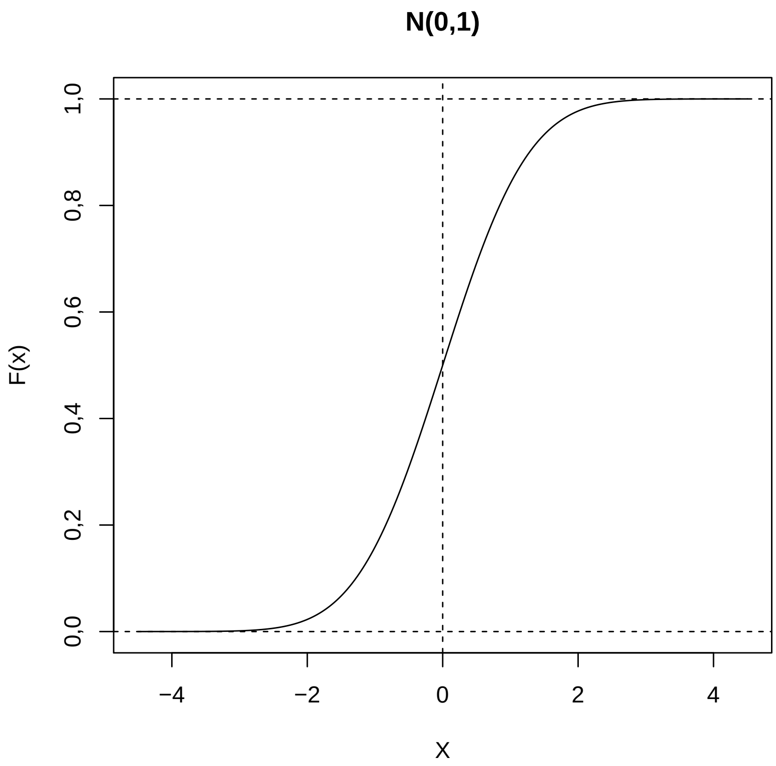


Figura 1.22: Funzione di Ripartizione della v.a. Normale Standard ($\mu = 0$, $\sigma^2 = 1$).

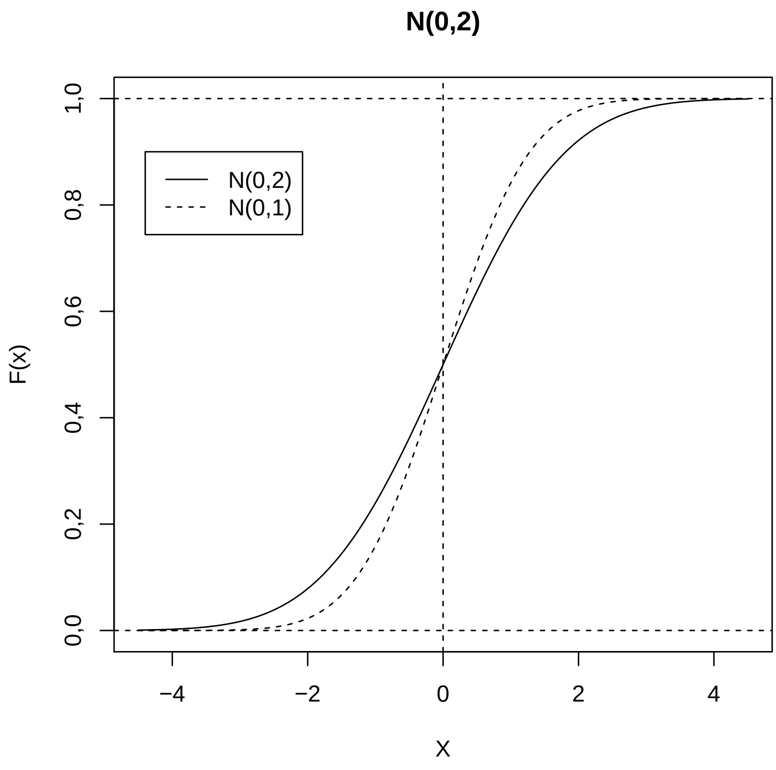


Figura 1.23: Funzione di Ripartizione della v.a. Normale $\mu = 0$, $\sigma^2 = 2$).

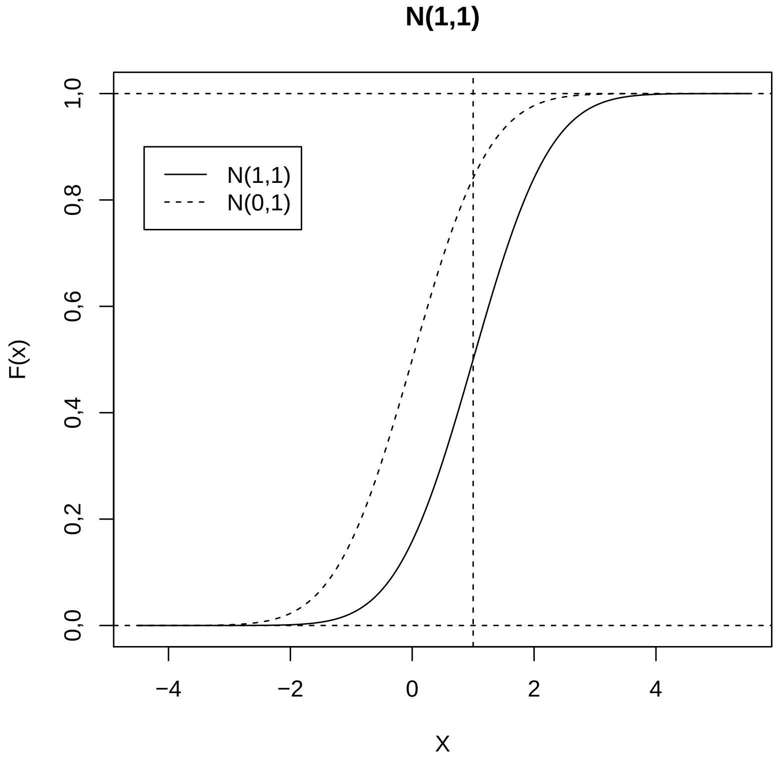


Figura 1.24: Funzione di Ripartizione della v.a. Normale $\mu = 1$, $\sigma^2 = 1$).

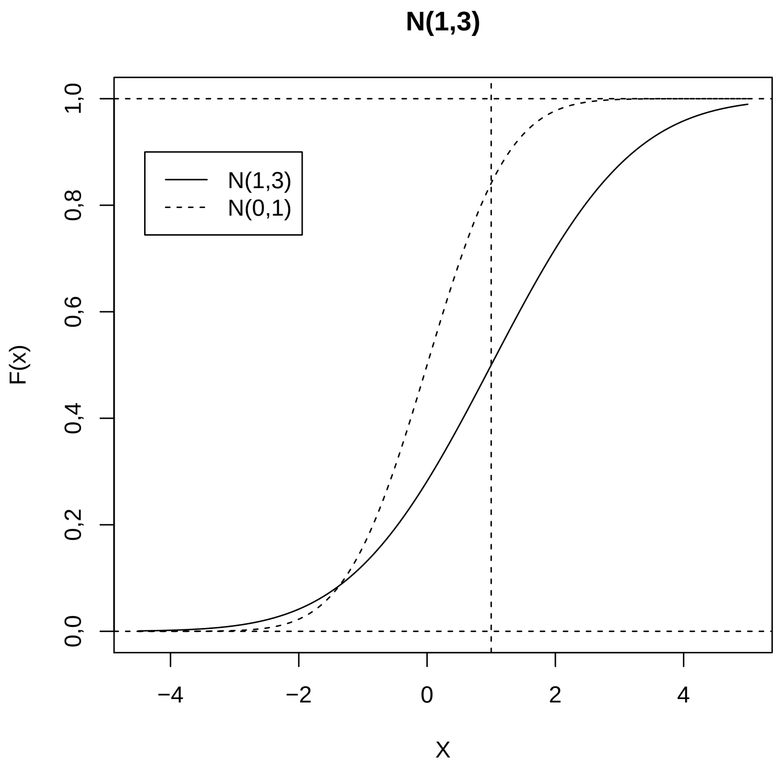


Figura 1.25: Funzione di Ripartizione della v.a. Normale $\mu = 1$, $\sigma^2 = 3$).

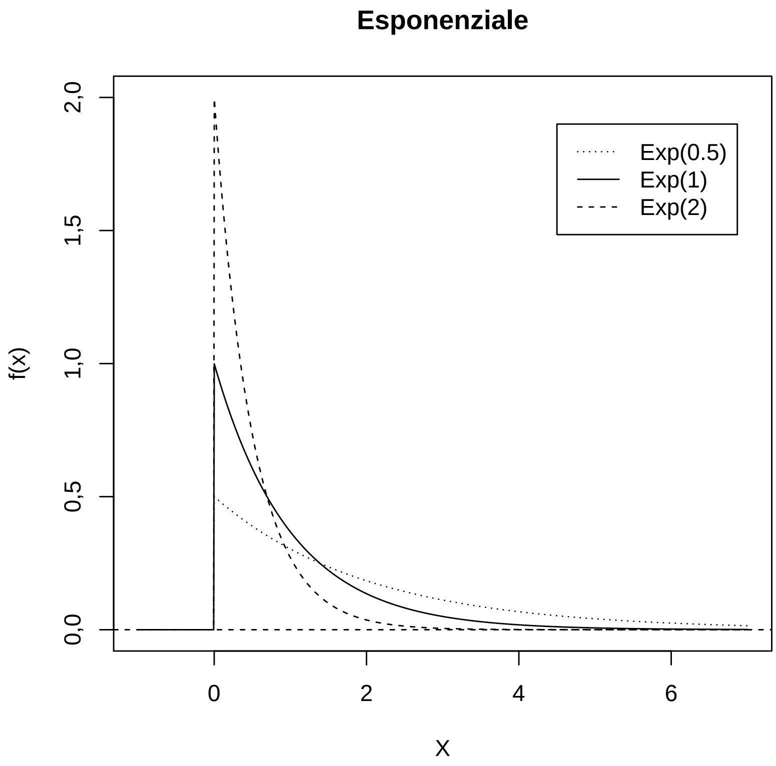


Figura 1.26: Funzione di Densità della v.a. Esponenziale $\lambda = 0.5, 1, 2$.

Distribuzione Esponenziale

Definizione 44 (Distribuzione Esponenziale)**.** *Si dice che una v.a. X ha legge Esponenziale con parametro $\lambda > 0$ se la sua funzione di densità à*

$$f(x;\lambda) = \begin{cases} \lambda \exp(-\lambda x) & x > 0 \\ 0 & \textit{altrove} \end{cases}$$

e la indichiamo nel seguente modo $X \sim \mathrm{Exp}(\lambda)$.

In alcuni casi la densità è espressa in funzione di $\mu = \frac{1}{\lambda}$ (cioè attraverso un'altra parametrizzazione).

Esercizio 62. *Si scriva la densità in funzione di μ.*

La funzione di ripartizione della v.a. Esponenziale è

$$F(x) = \begin{cases} \int_0^x \lambda e^{-\lambda t}\, dt = 1 - e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

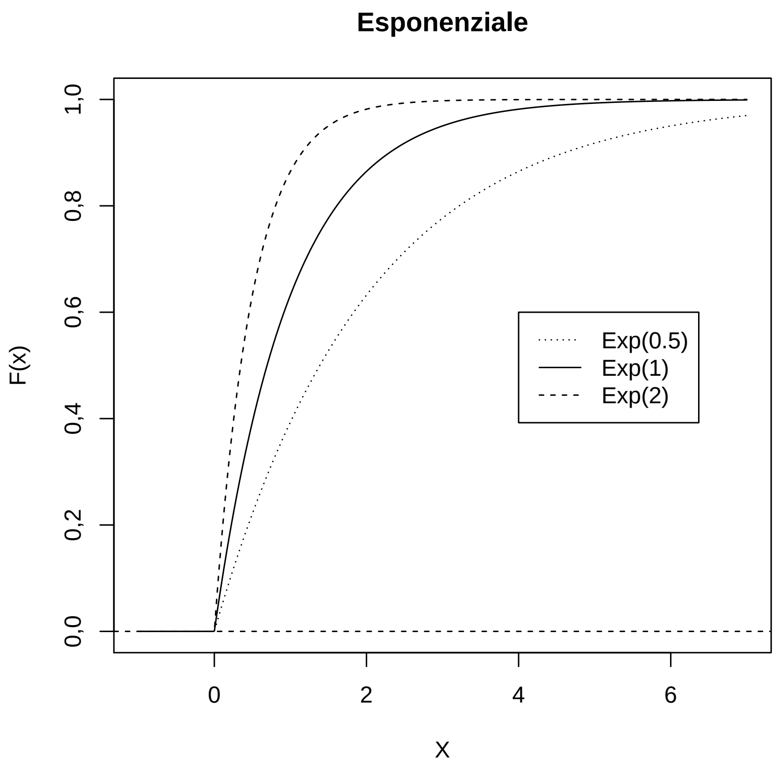


Figura 1.27: Funzione di Ripartizione della v.a. Esponenziale $\lambda = 0.5, 1, 2$.

Al pari della legge Geometrica, anche la legge Esponenziale possiede il requisito di essere *senza memoria* nel senso che se X è una v.a. Esponenziale allora, dati due reali positivi y e x, abbiamo

$$\begin{aligned}\Pr(y < X \leq y + x | X > y) &= \frac{\Pr(\{y < X \leq y + x\} \cap \{X > y\})}{\Pr(\{X > y\})} \\ &= \frac{\Pr(\{y < X \leq y + x\})}{\Pr(\{X > y\})} \\ &= \frac{e^{-\lambda y} - e^{-\lambda(y+x)}}{e^{-\lambda y}} \\ &= 1 - e^{-\lambda x} = \Pr(X \leq x)\end{aligned}$$

Trasformazione di varibili aleatorie

Si immagini un'esperimento definito sullo spazio $(\Omega, \mathcal{A}, \Pr)$. Sia $X(\omega)$ una v.a. il cui valore è determinato dall'esito $\omega \in \Omega$.

Consideriamo una funzione $Y(\omega) = g(X(\omega))$ dove $g : \mathbb{R} \to \mathbb{R}$. Se g è tale che

$$\{x \in \mathbb{R} : g(x) \leq z\} \in \mathbb{B}(\mathbb{R}) \text{ per ogni } z \in \mathbb{R}$$

allora $Y(\omega)$ è una variabile aleatoria.

Tale condizione risulta soddisfatta se g è una funzione con almeno una delle seguenti caratteristiche

- continua
- monotona (crescente o descrescente)

Quindi, ad esempio, se $X(\omega)$ è una v.a. allora anche

- $|X(\omega)|$
- $X^m(\omega)$ (m intero naturale)
- $aX(\omega)$ (a numero reale)
- $\exp(X(\omega))$

Posto

$$A(y) = \{x \in \mathbb{R} : g(x) \leq y\}$$

allora abbiamo

$$F_Y(y) = \Pr(Y(\omega) \leq y) = \Pr(X(\omega) \in A(y))$$

Nel caso discreto abbiamo:

$$\Pr(Y \leq y) = \sum_{\{x: g(x) \leq y\}} p_X(x)$$

mentre nel caso di una v.a. dotata di densità

$$\Pr(Y \leq y) = \int_{\{x: g(x) \leq y\}} f_X(x)\ dx$$

Teorema 18. *Sia X una v.a. con densità $f_X(x)$ con supporto l'intervallo (a, b), eventualmente non limitato. Sia $g(x)$ una funzione strettamente monotona con derivata in (a, b).*

Allora la v.a. $Y = g(X)$ è dotata di densità

$$f_Y(y) = \begin{cases} f_X(g^{-1}(y)) \left| \frac{\partial}{\partial y} g^{-1}(y) \right| & \alpha < y < \beta \\ 0 & \textit{altrove} \end{cases}$$

dove $\alpha = \min(g(a), g(b))$ e $\beta = \max(g(a), g(b))$.

Esempio 15. *Sia X una v.a. esponenziale di parametro λ. Sia $g(x) = e^x$. Allora applicando il teorema precedente abbiamo*

- $\alpha = g(a) = g(0) = 1$
- $\beta = g(b) = g(+\infty) = +\infty$
- $g^{-1}(y) = \log(y)$ *per* $1 < y < +\infty$
- $\frac{\partial}{\partial y} g^{-1}(y) = \frac{1}{y}$

allora

$$f_Y(y) = \lambda \exp\left(-\lambda \log(y)\right) \frac{1}{y} = \frac{\lambda}{y^{\lambda+1}} \qquad 1 < y < +\infty$$

e 0 altrove.

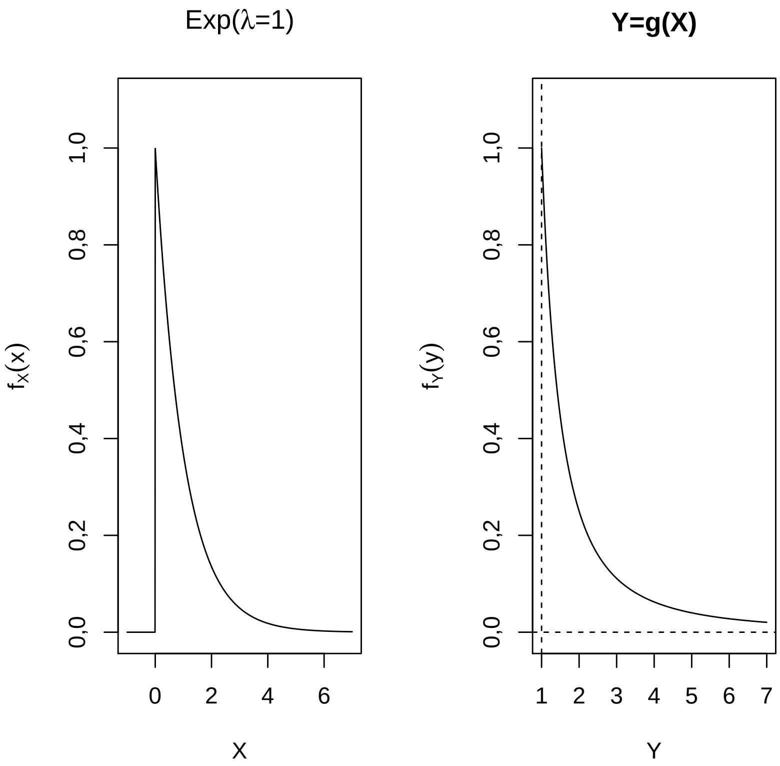


Figura 1.28: .

Caratteristiche, scale di misura e v.a.

Le v.a. rappresentato una misura di un particolare aspetto del fenomeno che viene osservato attraverso l'esperimento aleatorio. Gli aspetti che vengono misurati sono spesso classificati in una scala nel seguente modo

- qualitativi: se le modalità sono espresse in forma verbale
  - nominali (o categoriali)
  - ordinali
- quantitativi: se le modalità sono espresse in forma numerica
  - ordinali
  - intervallo
  - rapporto

Scala nominale (o categoriale)

Esempi di v.a. che sono su scala nominale sono

- il colore degli occhi: verdi, grigi, azzurri, castani, neri, $\cdots$
- i modelli di un determinato prodotto
- il sesso di un individuo: maschio o femmina
- la religione professata da un individuo

Ogni valore assunto dalla variabile è chiamato *modalità* e spesso queste variabili sono chiamate *mutabili.*

Le variabili su scala nominale sono variabili in cui non è possibile (o non ha senso) stabilire una relazione d'ordine tra le modalità, infatti posto ad esempio $\Omega = \{\text{“}verdi\text{''}, \text{“}grigi\text{''}, \text{“}azzurri\text{''}, \text{“}castani\text{''}, \text{“}neri\text{''}\}$ e definiamo una v.a. $X(\omega)$ tale che $X(\text{“}verdi\text{''}) = 1$, $X(\text{“}grigi\text{''}) = 2$, $X(\text{“}azzurri\text{''}) = 3$, $X(\text{“}castani\text{''}) = 4$, $X(\text{“}neri\text{''}) = 5$. È chiaro che non ha senso in questo contesto la scrittura

$$X(\omega) < 3 \qquad .$$

Scala ordinale (qualitative e quantitative)

Quando tra le modalità di una v.a. è possibile (ha senso) instaurare una relazione d'ordine allora la variabile è su scala ordinale. Per tali variabili non ha però senso parlare di *quanto* una modalità è più grande di un'altra. Ad esempio, consideriamo la variabile "quanto zuccherata è la marmellata" con le seguenti modalità "poco", "abbastanza", "molto" mentre è chiaro che "poco" indica una quantità di zucchero inferiore a "abbastanza" e a "molto".

Non è vero, in generale, invece che per passare da una marmellata "poco" zuccherata a una "abbastanza" zuccherata ci voglia la stessa quantità di zucchero che per passare da una marmellata "abbstanza" zuccherata a una "molto" zuccherata.

- il livello di istruzione (qualitativa)
- l'ordine di arriva dei partecipanti ad una gara (quantitativa)
- la classe di reddito annuo (quantitativa) con le seguenti modalità $0 \leq R \leq 5000$, $5000 < R \leq 10000$, $10000 < R \leq 20000$, $20000 < R \leq 40000$ e $R > 40000$.
- gli scaglioni di reddito per determinare le aliquote irpef (quantitativa)

Scala intervallare

Una v.a. è misurata su scala intervallo quando oltre alla possibilità di stabilire una relazione d'ordine ha senso anche confrontare le loro differenze, ad esempio ha senso dire che $|x_1 - x_2| = |x_2 - x_3$. Mentre per queste variabili ha senso parlare di una loro trasformazione lineare, non ha senso considerare il loro rapporto. Il problema nasce dalla mancanza di un zero naturale.

Un esempio di v.a. su scala intervallo è una variabile che misuri il tempo. Mentre il tempo trascorso tra due istanti rimane inalterato al cambiare dell'istante in cui noi fissiamo uno zero il rapporto invece si può alterare.

Infatti, consideriamo i due calendari, Gregoriano e Mussulmano.

Il primo fissa come anno zero la nascita di Cristo, il secondo invece l'anno 622 d.C. (rispetto a quello Gregoriano) così per il calendario Gregoriano questo è l'anno 2004 mentre per quello Mussulmano è il 1382. Consideriamo la variabile numero di anni necessari per ottenere una laurea=anno di iscrizione−anno di laurea+1. Questa variabile assume lo stesso valore anche se misurata attraverso i due calendari.

Mentre la variabile

$$\frac{\text{anno di iscrizione}}{\text{anno di laurea} + 1}$$

dipende dal calendario che stiamo adoperando.

Un esempio in cui non risulta immediato stabilire se la v.a. è ordinale o su scala intervallo è la variabile

"voto all'esame di Probabilità e Statistica"

Scala rapporto

Una v.a. è su scala rapporto quando eiste uno zero naturale. Si consideri ad esempio la temperatura misurata in **gradi Kelvin**. Per questa variabile esiste uno zero naturale che è lo *zero assoluto*.

Valori di sintesi di una v.a.

Abbiamo visto che una v.a. può essere definita, in maniera equivalente, attraverso la funzione di probabilità (di densità) o attraverso la funzione di ripartizione. In molti casi risulta utile poter sintetizzare la legge di una v.a. attraverso la determinazione di un certo numero di caratteristiche sintetiche che pongono in luce particolari aspetti dell'intera distribuzione.

Le caratteristiche che in genere si considerano sono

- moda
- quantili
- momenti (centrati e non centrati)

Scale di misura e valori di sintesi

A seconda della scala di misura della variabile considerata vi sono diversi valori di sintesi che possono essere considerati

	valore di sintesi		
scala	moda	quantili	momenti
nominale	$\times$		
ordinale	$\times$	$\times$	
intervallo	$\times$	$\times$	$\times$
rapporto	$\times$	$\times$	$\times$

Moda

Definizione 45. *Si dice moda, indicata con* $\mathrm{Mo}(X)$ *di una v.a.* X *la modalità che si presenta con la densità (discreta nel caso di v.a. discreta) più elevata.*

Quantili

Definizione 46. *Si dice quantile di ordine* $0 \leq \alpha \leq 1$ *di una v.a.* X *la modalità* $Q_\alpha(X) = x_\alpha$ *tale che*

$$\Pr(X \leq x_\alpha) \geq \alpha \text{ e } \Pr(X \geq x_\alpha) \geq 1 - \alpha$$

cioè

$$F(x_\alpha) \geq \alpha \text{ e } 1 - F(x_\alpha) \geq 1 - \alpha$$

In particolare abbiamo:

- **centili**: $x_{0.01}, x_{0.02}, \cdots, x_1$;
- **decili**: $x_{0.1}, x_{0.2}, \cdots, x_1$;
- **quartili**:
  - 1^o quartile: $q_1 = \mathrm{Q}_{0.25}(X) = x_{0.25}$,
  - mediana: $\mathrm{Me}(X) = q_2 = \mathrm{Q}_{0.5}(X) = x_{0.5}$,
  - 3^o quartile: $q_3 = \mathrm{Q}_{0.75}(X) = x_{0.75}$
  - $q_4 = \mathrm{Q}_1(X) = x_1$.

Quartili

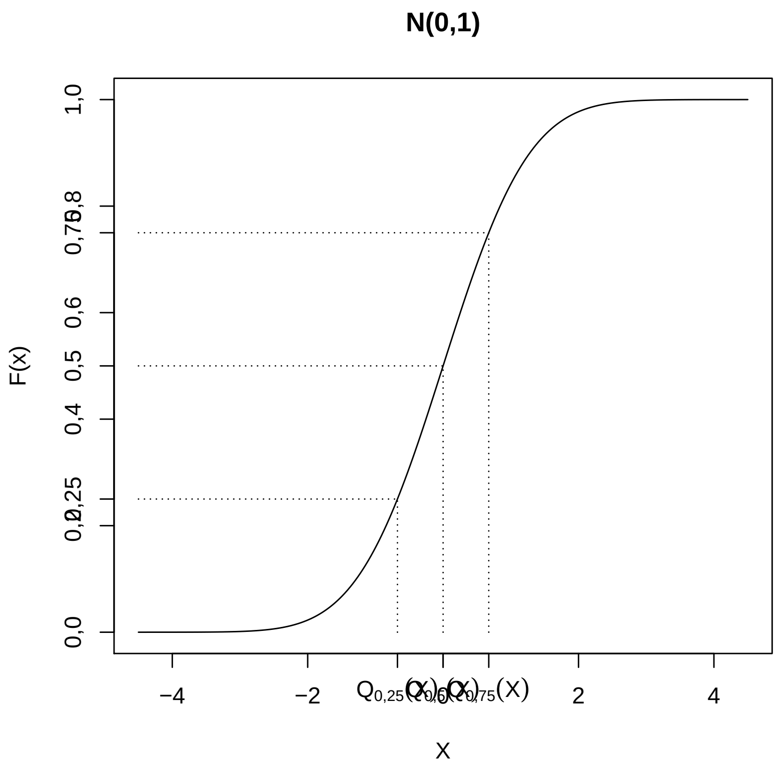


Figura 1.29: .

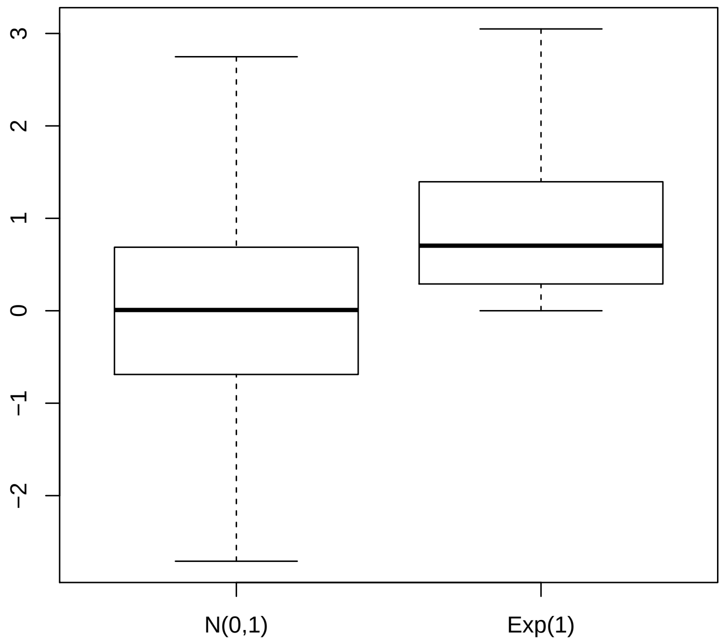


Figura 1.30: .

Quartili e Diagramma con scatola e baffi (*Box and whiskers plot*)

Una rappresentazione basata interamente sui quantili è il Diagramma con scatola e baffi. La scatola (un rettangolo tagliato da una linea) rappresenta i tre quartili q_i, $i = 1, 2, 3$ mentre i baffi sono costruiti nella seguente maniera:

- Si fissa una costante, in genere: $c = 1.5$
- Si calcola $\mathrm{SI}(X) = \mathrm{Q}_{0.75}(X) - \mathrm{Q}_{0.25}(X)$ e $m = \mathrm{Q}_0(X)$, $M = \mathrm{Q}_1(X)$.
- $\mathrm{Ba}_{inf}(X) = \max\{m, \mathrm{Q}_{0.25}(X) - c \cdot \mathrm{SI}(X)\}$
- $\mathrm{Ba}_{sup}(X) = \min\{M, \mathrm{Q}_{0.75}(X) + c \cdot \mathrm{SI}(X)\}$

$\mathrm{Ba}_{inf}(X)$ e $\mathrm{Ba}_{sup}(X)$ determinano la posizione dei baffi.

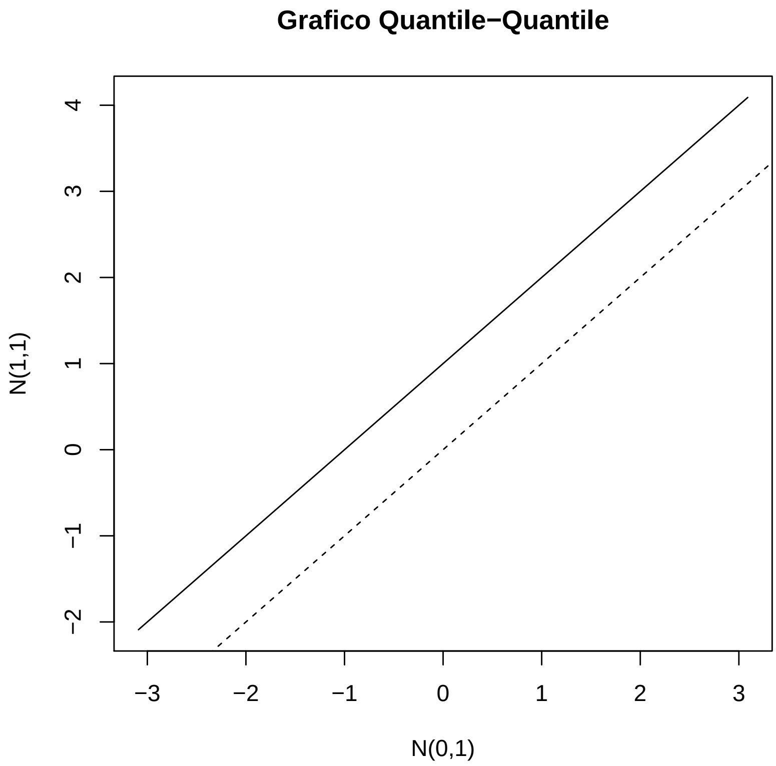


Figura 1.31: .

Grafico Quantile-Quantile

Il Grafico Quantile-Quantile è un utile strumento per confrontare la distribuzione di due v.a. X, Y. È prodotto ponendo in ascissa i valori $\mathrm{Q}_\alpha(X)$ e in ordinata $\mathrm{Q}_\alpha(Y)$, cioè si rappresentano sul piano cartesiano le coppie $(\mathrm{Q}_\alpha(X), \mathrm{Q}_\alpha(Y))$ per $\alpha \in [0,1]$.

Quando le due v.a. hanno la stessa distribuzione allora il grafico è la bisettrice del primo e terzo quadrante.

Speranza matematica o valore atteso per v.a. discrete

Definizione 47. *Sia X una v.a. discreta con funzione di probabilità $p_X(x)$. Allora, si chiama speranza matematica di X la quantità (finita)*

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

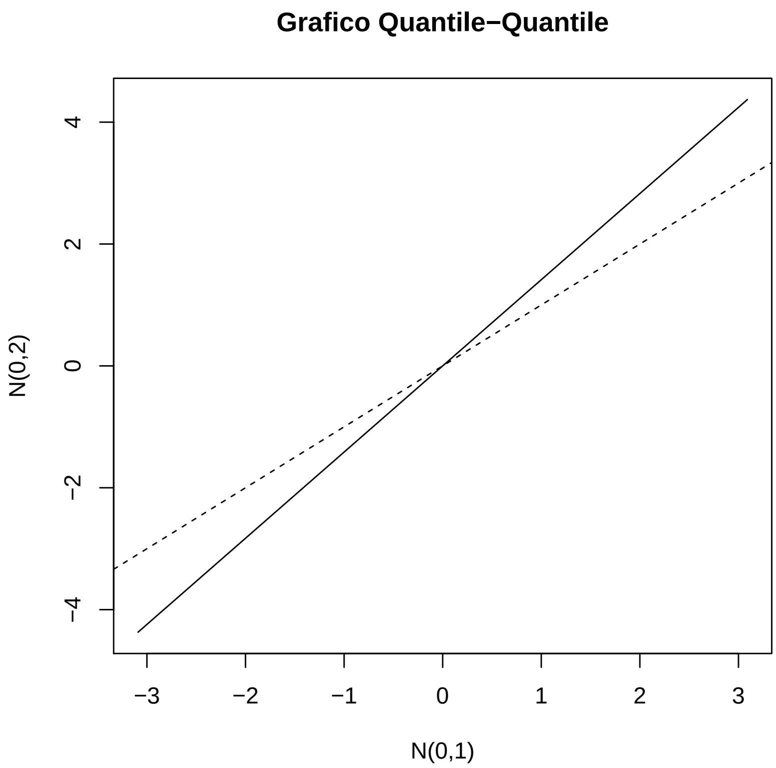


Figura 1.32: .

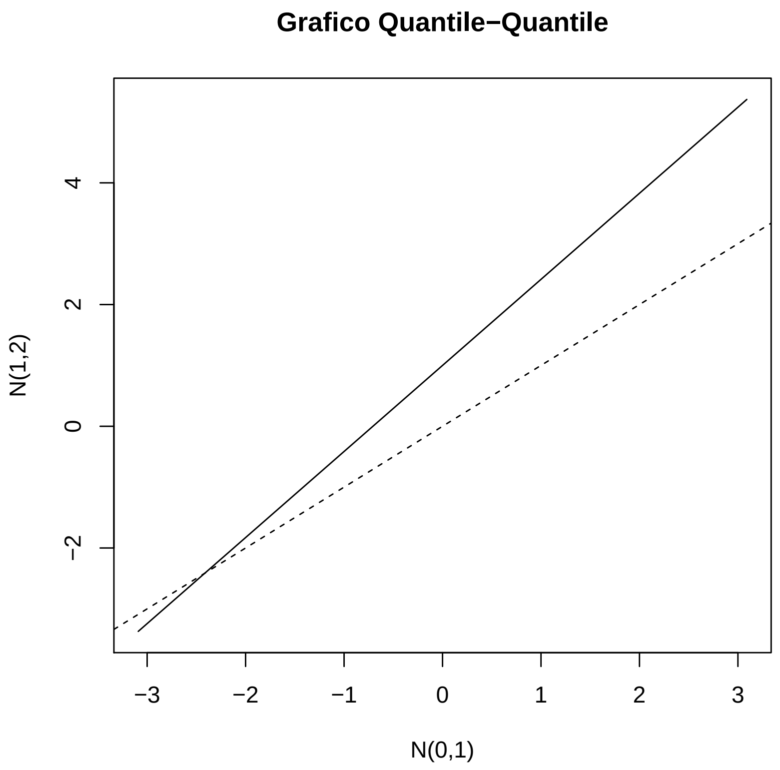


Figura 1.33: .

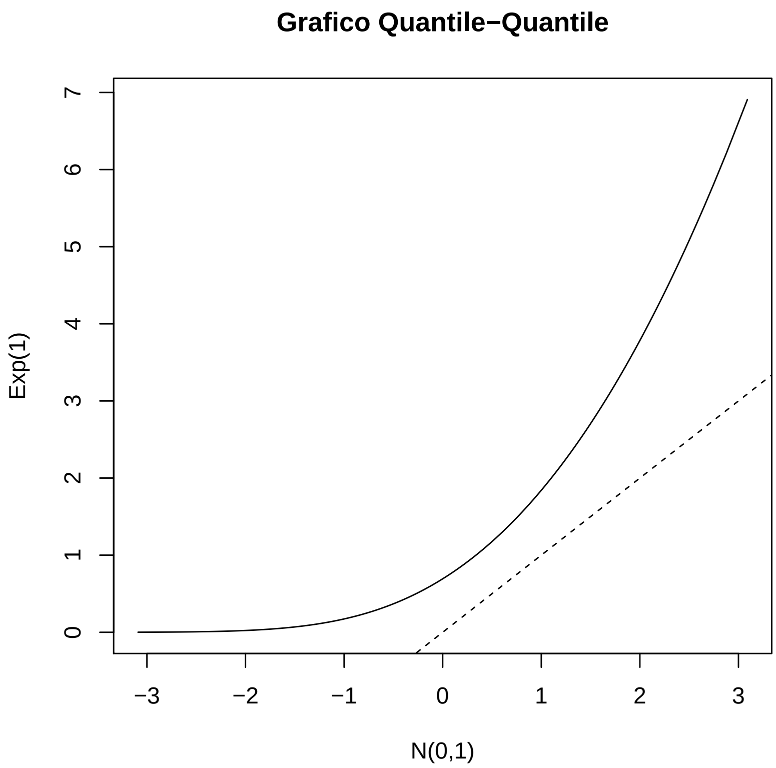


Figura 1.34: .

se

$$\mathbb{E}(|X|) = \sum_{x \in R_X} |x| \; p_X(x) < \infty$$

Definizione 48. *Sia X una v.a. dotata di densità $f_X(x)$ e funzione di ripartizione $F_X(x)$. Si chiama speranza matematica di X la quantità (finita)*

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

se

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

Consideriamo le v.a. Y e X tali che $Y = g(X)$ per una qualche funzione g. Allora

$$\mathbb{E}(Y) = \mathbb{E}(g(X)) = \begin{cases} \sum_{y \in R_Y} y p_Y(y) & = \sum_{x \in R_X} g(x) p_X(x) \\ & \text{se } Y \text{ è una v.a. discreta} \\ \int_{R_Y} y \; dF_Y(y) & = \int_{R_X} g(x) \; dF_X(x) \\ & \text{se } Y \text{ è una v.a dotata di densità} \end{cases}$$

Momenti

Definizione 49. *Data la v.a. X si dice momento non centrato di ordine r (intero positivo) il valore*

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

e si dice momento centrato dalla media di ordine r

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

valori di sintesi basati sui momenti

- Media: $\mu = \mu_1 = \mathbb{E}(X)$
- Varianza: $\mathrm{var}(X) = \sigma^2 = \bar{\mu}_2 = \mathbb{E}((X - \mu_1)^2)$
- Deviazione standard: $\sigma = \sqrt{\sigma^2}$
- Coefficiente di variazione: $CV = \frac{\sigma}{\mu}$
- Indice di Asimmetria: $\rho_3 = \frac{\bar{\mu}_3}{\sigma^3}$
- Indice di Curtosi: $\rho_4 = \frac{\bar{\mu}_4}{\sigma^4} - 3$

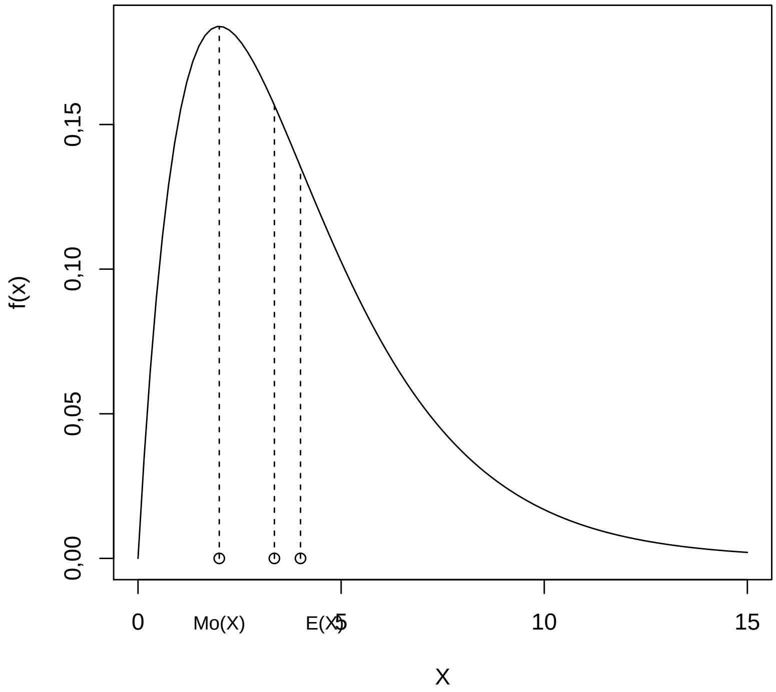


Figura 1.35: .

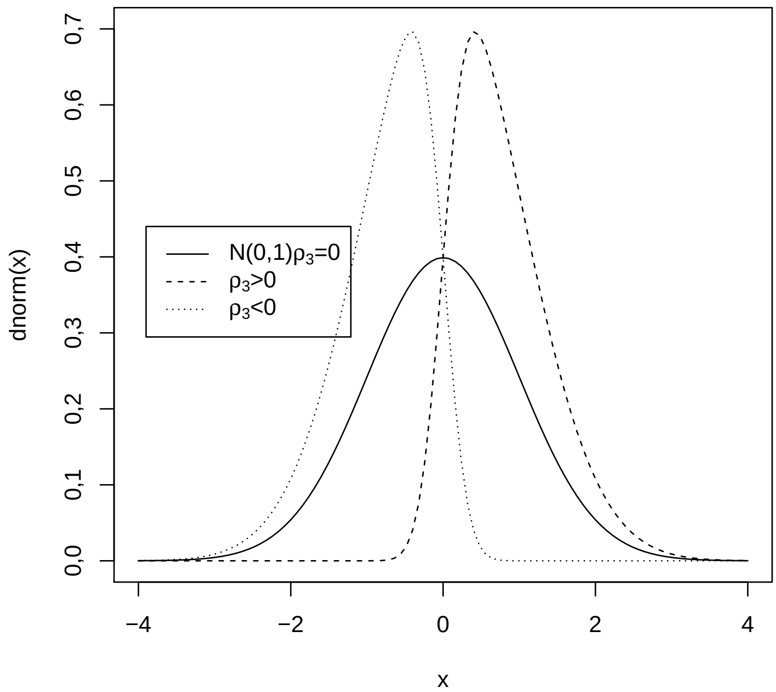


Figura 1.36: .

Media, Mediana, Moda

Asimmetria

Curtosi

Proprietà dell'operatore Speranza Matematica

- Sia $X(\omega) = c$ la v.a. constante che assume con probabilità 1 il valore c allora $\mathbb{E}(X) = \mathbb{E}(c) = c$
- Sia $Y = c\ g(X)$ una v.a. allora $\mathbb{E}(Y) = c\mathbb{E}(g(X))$
- Siano $g(X)$ e $h(Y)$ due v.a. e a e b due constanti allora la v.a. $Z = ag(X) + bh(Y)$ è tale che

$$\mathbb{E}(Z) = a\mathbb{E}(g(X)) + b\mathbb{E}(h(Y))$$

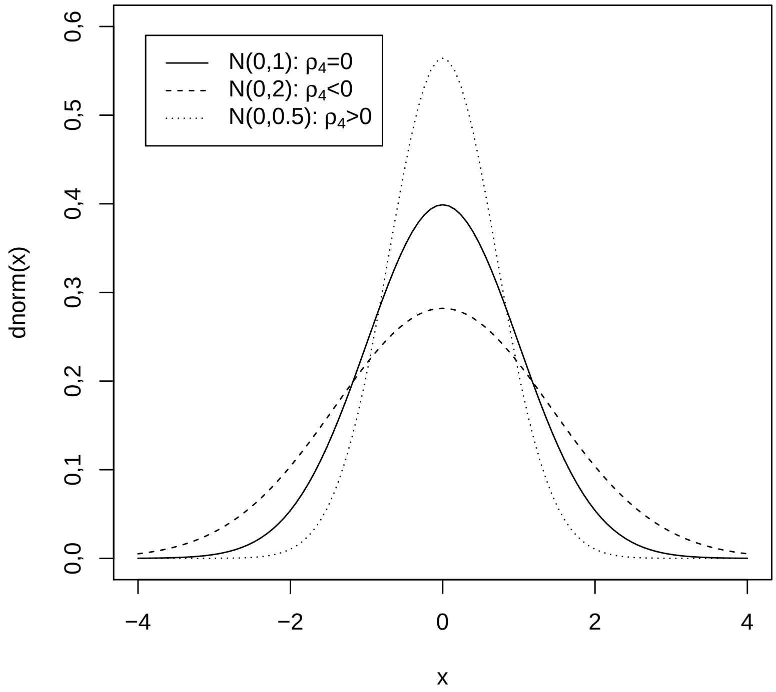


Figura 1.37: .

- Siano X_i, $i = 1, \cdots, n$ v.a., mentre a_i, $i = 1, \cdots, n$ constanti. Poniamo $Z = \sum_{i=1}^{n} a_i X_i$ allora

$$\mathbb{E}(Z) = \sum_{i=1}^{n} a_i \mathbb{E}(X_i)$$

- Siano X e Y due v.a. indipendenti allora

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

Proprietà dell'operatore Varianza

- $\mathrm{var}(X) = \sigma^2 = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \mu_2 - \mu_1^2$

- Sia c una costante e X una v.a. allora posto $Y = X + c$ si ha $\mathrm{var}(Y) = \mathrm{var}(X + c) = \mathrm{var}(X)$

- Posto $Y = c \cdot X$ si ha $\mathrm{var}(Y) = \mathrm{var}(c \cdot X) = c^2 \mathrm{var}(X)$

- Date due costanti a e b e posto $Y = aX + b$ allora

$$\mathrm{var}(Y) = \mathrm{var}(aX + b) = a^2 \mathrm{var}(X)$$

- Se X e Y sono due v.a. indipendenti allora

$$\mathrm{var}(X + Y) = \mathrm{var}(X) + \mathrm{var}(Y)$$

- Più in generale date X_i, $i = 1, \cdots, n$ v.a. a due a due indipendenti allora

$$\mathrm{var}(\sum_{i=1}^{n} a_i X_i) = \sum_{i=1}^{n} a_i^2 \mathrm{var}(X_i)$$

Media della v.a. Binomiale

$$
\begin{aligned}
\mathbb{E}(X) &= \sum_{x=0}^{N} x \binom{N}{x} p^x (1-p)^{N-x} \\
&\downarrow \text{il primo termine della somma è zero} \\
&= \sum_{x=1}^{N} x \binom{N}{x} p^x (1-p)^{N-x} \\
&\downarrow \text{poniamo } s = x-1 \text{ otteniamo} \\
&= \sum_{s=0}^{N-1} (s+1) \frac{N!}{(s+1)!(N-s-1)!} p^{s+1} (1-p)^{N-s-1} \qquad \downarrow \text{portiamo fuori } N \text{ e } p \\
&= Np \sum_{s=0}^{N-1} (s+1) \frac{(N-1)!}{(s+1)!(N-s-1)!} p^{s} (1-p)^{N-s-1} \\
&\downarrow \text{semplifichiamo } (s+1) \text{ con il primo fattore al denominatore} \\
&= Np \sum_{s=0}^{N-1} \frac{(N-1)!}{s!(N-s-1)!} p^{s} (1-p)^{N-s-1} \\
&\downarrow {\scriptstyle \text{La sommatoria è la somma della funzione di probabilità di una Bi}(N-1,p)} \\
&= Np
\end{aligned}
$$

In maniera più semplice consideriamo una v.a. di Bernoulli ($X \sim \text{Bi}(1,p)$), allora dalla definizione di speranza matematica abbiamo

$$\mathbb{E}(X) = \sum_{x=0}^{1} x p_X(x) = 0 \cdot (1-p) + 1 \cdot p = p$$

Infine, consideriamo $X_1, X_2, \cdots, X_N$, N v.a. Bernoulliane indipendenti di parametro p e $X = \sum_{i=1}^{N} X_i \sim \text{Bi}(N,p)$ allora

$$\mathbb{E}(X) = E\left(\sum_{i=1}^{N} X_i\right) = \sum_{i=1}^{N} \mathbb{E}(X_i) = Np$$

Varianza della v.a. Binomiale

Per il calcolo della varianza utilizziamo la relazione

$$\text{var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2$$

e quindi ci rimane da calcolare $\mu_2 = \mathbb{E}(X^2)$.

$$
\begin{aligned}
\mathbb{E}(X^2) &= \sum_{x=0}^{N} x^2 \binom{N}{x} p^x (1-p)^{N-x} \\
&= \sum_{x=1}^{N} x^2 \binom{N}{x} p^x (1-p)^{N-x} \\
&= \sum_{x=1}^{N} x \frac{N!}{(x-1)!(N-x)!} p^x (1-p)^{N-x} \\
&= N \sum_{x=1}^{N} x \binom{N-1}{x-1} p^x (1-p)^{N-x} \\
&= N \sum_{s=0}^{N-1} (s+1) \binom{N-1}{s} p^{s+1} (1-p)^{N-s-1} \\
&= N \left\{ \sum_{s=0}^{N-1} s \binom{N-1}{s} p^{s+1} (1-p)^{N-s-1} \right. \\
&\left. + \sum_{s=0}^{N-1} \binom{N-1}{s} p^{s+1} (1-p)^{N-s-1} \right\} \\
&= N\{p \cdot (N-1) \cdot p + p\} = N(p^2(N-1) + p)
\end{aligned}
$$

Da cui la varianza è

$$
\begin{aligned}
\mathrm{var}(X) &= \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 \\
&= N(p^2(N-1) + p) - (Np)^2 \\
&= N(N-1)p^2 + Np - N^2p^2 \\
&= N^2p^2 - Np^2 + Np - N^2p^2 \\
&= Np(1-p)
\end{aligned}
$$

In maniera più semplice consideriamo una v.a. di Bernoulli ($X \sim \mathrm{Bi}(1, p)$), allora abbiamo

$$
\mathbb{E}(X^2) = \sum_{x=0}^{1} x^2 p_X(x) = 0^2 \cdot (1-p) + 1^2 \cdot p = p
$$

$$
\mathrm{var}(X) = p - p^2 = p \cdot (1-p)
$$

Infine, consideriamo $X_1, X_2, \cdots, X_N$, N v.a. Bernoulliane indipendenti di parametro p e $X = \sum_{i=1}^{N} X_i \sim \mathrm{Bi}(N, p)$ allora

$$
\mathrm{var}(X) = V\left(\sum_{i=1}^{N} X_i\right) = \sum_{i=1}^{N} \mathrm{var}(X_i) = Np(1-p)
$$

Media della v.a. Geometrica

$$
\begin{aligned}
\mathbb{E}(X) &= \sum_{x=1}^{\infty} x \cdot p(1-p)^{x-1} = p \sum_{x=1}^{\infty} x \cdot (1-p)^{x-1} \\
&= -p \sum_{x=1}^{\infty} \frac{\partial}{\partial p}(1-p)^{x} \\
&\downarrow \qquad \text{siccome è una serie assulutamente convergente, possiam} \\
&= -p \frac{\partial}{\partial p} \sum_{x=1}^{\infty} (1-p)^{x} \\
&\downarrow \ \text{la serie è una serie geometrica di ragione } 1-p \text{ dove manca il primo termine} \\
&= -p \frac{\partial}{\partial p} \left(\frac{1}{1-(1-p)} - 1 \right) \\
&= -p \frac{\partial}{\partial p} \left(\frac{1-p}{1-(1-p)} \right) \\
&= -p \left(\frac{-p-(1-p)}{p^2} \right) = \frac{1}{p}
\end{aligned}
$$

Momento secondo della v.a. Geometrica

$$
\begin{aligned}
\mathbb{E}(X^2) &= \sum_{x=1}^{\infty} x^2 p(1-p)^{x-1} \\
&= p\left\{\sum_{x=1}^{\infty} x(x-1)(1-p)^{x-1}\right. \\
&\left.+\sum_{x=1}^{\infty} x(1-p)^{x-1}\right\} \\
&= p\left\{\sum_{x=1}^{\infty} x(x-1)(1-p)^{x-1} + \frac{1}{p^2}\right\} \\
&= p\left\{(1-p)\sum_{x=1}^{\infty} x(x-1)(1-p)^{x-2} + \frac{1}{p^2}\right\} \\
&= p\left\{(1-p)\sum_{x=1}^{\infty} \frac{\partial^2}{\partial p^2}(1-p)^{x} + \frac{1}{p^2}\right\} \\
&= p(1-p)\frac{\partial^2}{\partial p^2}\sum_{x=1}^{\infty}(1-p)^{x} + \frac{1}{p} \\
&= p(1-p)\frac{\partial^2}{\partial p^2}\frac{1-p}{p} + \frac{1}{p} \\
&= p(1-p)\frac{2}{p^3} + \frac{1}{p} = \frac{2-p}{p^2}
\end{aligned}
$$

$$
\mathrm{var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}
$$

Media della v.a. Normale

$$
\begin{aligned}
\mathbb{E}(X) = & \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] dx \\
& \text{poniamo } x = \sigma z + \mu, \text{ cioè } z = \tfrac{x-\mu}{\sigma} \\
& \downarrow \text{ da cui } dx = \sigma dz \\
& = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} (\sigma z + \mu) \exp\left[-\frac{1}{2} z^2\right] \sigma \; dz \\
& = \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z \exp\left[-\frac{1}{2} z^2\right] dz \\
& + \mu \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2} z^2\right] dz
\end{aligned}
$$

- Il primo integrale è nullo in quanto z è una funzione dispari e $\exp\left[-\frac{1}{2}z^2\right]$ è una funzione pari.
- Il secondo integrale è pari a 1 perché integrale di una densità.

Quindi

$$\mathbb{E}(X) = \mu$$

Momente secondo e Varianza della v.a. Normale

Seguendo lo stesso procedimento fatto per la media e utilizzando l'integrazione per parti otteniamo

$$\mathbb{E}(X^2) = \sigma^2 + \mu^2$$

da cui

$$\operatorname{var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \sigma^2 + \mu^2 - \mu^2 = \sigma^2$$

Su trasformazioni lineari di una Normale

Teorema 19. *Sia $X \sim \mathrm{N}(\mu, \sigma^2)$ e siano $a \neq 0$ e b due costanti. Allora la v.a.*

$$Y = aX + b \sim \mathrm{N}(a\mu + b, a^2\sigma^2)$$

Per la dimostrazione si usi il teorema che abbiamo enunciato sulle trasformazioni di v.a.

In particolare si ha

$$\mathbb{E}(Y) = \mathbb{E}(aX + b) = a\mu + b \qquad \mathrm{var}(Y) = \mathrm{var}(aX + b) = a^2\mathrm{var}(X) = a^2\sigma^2$$

Standardizzazione di una Normale

Come risultato del precedente teorema abbiamo che data una $X \sim \mathrm{N}(\mu, \sigma^2)$, allora

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

questa operazione viene chiamata *standardizzazione*.

L'operazione di standardizzazione risulta molto utile per calcolare la probabilità di eventi nella forma $X \leq c$ (con c costante) per una $X \sim \mathrm{N}(\mu, \sigma^2)$. Infatti tale evento è equivalente a

$$\left\{ \frac{X - \mu}{\sigma} \leq \frac{c - \mu}{\sigma} \right\}$$

da cui

$$\Pr(\{X \leq c\}) = \Pr\left(\left\{ \frac{X - \mu}{\sigma} \leq \frac{c - \mu}{\sigma} \right\} \right)$$

ma $\frac{X-\mu}{\sigma} \sim \mathrm{N}(0, 1)$ le cui probabilità sono tabulate.

Nella stessa maniera possiamo calcolare la probabilità di eventi nella forma $a \leq X \leq b$ (con a e b costanti) per una $X \sim \mathrm{N}(\mu, \sigma^2)$. Infatti tale evento è equivalente a

$$\left\{ \frac{a - \mu}{\sigma} \leq \frac{X - \mu}{\sigma} \leq \frac{b - \mu}{\sigma} \right\}$$

da cui

$$\begin{aligned}
\Pr(\{a \leq X \leq b\}) &= \Pr\left(\left\{ \frac{a - \mu}{\sigma} \leq \frac{X - \mu}{\sigma} \leq \frac{b - \mu}{\sigma} \right\} \right) \\
&= \Pr\left(\left\{ \frac{X - \mu}{\sigma} \leq \frac{b - \mu}{\sigma} \right\} \right) \\
&- \Pr\left(\left\{ \frac{X - \mu}{\sigma} \leq \frac{a - \mu}{\sigma} \right\} \right) \\
&= \Phi\left(\frac{b - \mu}{\sigma} \right) - \Phi\left(\frac{a - \mu}{\sigma} \right)
\end{aligned}$$

Media e Varianza di una Esponenziale

$$\mathbb{E}(X) = \int_0^\infty x\lambda \exp\left[-\lambda x\right]\, dx = \frac{1}{\lambda}$$

e

$$\mathrm{var}(X) = \int_0^\infty x^2\lambda \exp\left[-\lambda x\right]\, dx - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

come si verificare integrando per parti.

Esercizio 63 (Direttore). *Il direttore di una società ha il suo domicilio in una data località A. Egli parte dal suo domicilio in macchina alle 8.45 per arrivare alla società che apre alle ore 9.00. Il tempo di percorso è, in media, di 13 minuti, con una deviazione standard di 3 minuti.*

Assumendo che il tempo di percorso sia distribuito normalmente, determinare in quale percentuale di casi il direttore arriva alla società in ritardo.

Esercizio 64 (Differenza in due lanci di un dado). *Si lanciano due dadi bilanciati. Si costruisca la variabile casuale X che rappresenta la differenza tra il punteggio del primo e del secondo dado. Si determini*

- *la distribuzione di probabilità e la funzione di ripartizione di X e se ne disegnino i grafici.*
- $\Pr(-2 < X \leq 0)$*;*
- *media e varianza di X;*
- *distribuzione di probabilità, media e varianza della variabile casuale $Z = X^2$.*

Esercizio 65 (Quattro sfere). *Un'urna contiene quattro sfere delle quali due sono contrassegnate dal numero* 1*, una dal* 3 *e una dal* 4*. Si estraggono senza reinserimento due sfere e sia X la variabile aleatoria che indica la somma dei numeri corrispondenti alle sfere estratte. Si determini*

- *la funzione di probabilità di X, con relativa rappresentazione grafica;*
- *la funzione di ripartizione di X, con relativa rappresentazione grafica;*
- *la media della distribuzione di X;*
- *la varianza di X;*
- $\Pr(X \geq 7)$ *e* $\Pr(3 < X \leq 5)$.

Esercizio 66 (Stereo)**.** *Una partita di 6 stereo ne contiene 2 difettosi. Un locale acquista 3 di questi stereo a caso. Se X conta il numero di stereo difettosi, trovarne la funzione di probabilità e la funzione di ripartizione con relativi grafici. Calcolarne*

- *media e varianza di X;*
- $\Pr(X = 1)$ *e* $\Pr(0 < X \leq 2)$ *a partire dalla funzione di ripartizione.*

Esercizio 67 (Urna con tre palline)**.** *Un'urna contiene tre palline numerate da* 1 *a* 3*. Si estraggono con reinserimento due palline e sia X la variabile aleatoria che indica la differenza in modulo dei numeri estratti. Si determini*

- *la funzione di probabilità di X, con relativa rappresentazione grafica;*
- *la funzione di ripartizione di X, con relativa rappresentazione grafica;*
- *la media della distribuzione di X;*
- *la varianza di X;*
- $\Pr(X \leq 2)$ *e* $\Pr(2 \leq X < 5)$.

Esercizio 68 (Contenitori e palline)**.** *Si distribuiscono casualmente 2 palline in 4 contenitori e sia $X =$"numero di palline nel primo contenitore".*

- *Trovare e disegnare la funzione di probabilità di X;*
- *Trovare e disegnare la funzione di ripartizione di X;*
- *Calcolare media e varianza di X.*

Esercizio 69 (Casello Autostradale)**.** *Ad un casello autostradale arriva ogni ora un numero di automobili che segue una distribuzione di Poisson di parametro $\lambda = 20$. Qual è la probabilità che in un'ora arrivino non più di 7 automobili? E che il numero di macchine sia compreso fra 6 e 12 (estremi inclusi)?*

Esercizio 70 (Piove?)**.** *In una località balneare la probabilità che piova in un qualunque giorno del mese di agosto è 0.05. Assumendo che vi sia indipendenza tra i vari giorni del mese, qual è la probabilità che la prima pioggia del mese si osservi il 15 agosto? E prima del 15 agosto? Dato che fino al 10 agosto non ha piovuto, qual è la probabilità che non piova fino al 25?*

Esercizio 71 (Lancio di una moneta non bilanciata)**.** *Si lancia una moneta che presenta testa con probabilità 0.6. Se il risultato è testa, si estraggono 4 palline con reinserimento da un'urna che contiene 6 palline bianche e 4 palline nere. Se esce croce, si estraggono dalla stessa urna 3 palline senza reinserimento. Trovare funzione di probabilità, funzione di ripartizione e valore atteso della variabile che conta il numero di palline bianche estratte nell'esperimento.*

Esercizio 72 (Alberi)**.** *La probabilità che una albero di un certo frutteto non dia frutto è 0.04. Qual è*

- *la probabilità che su 200 alberi esattamente 7 non diano frutti?*
- *la probabilità che meno di 2 piante non diano frutti?*
- *la probabilità che almeno una pianta dia frutto?*

Esercizio 73 (Prigioniero e porte)**.** *Un prigioniero è rinchiuso in una cella con 3 porte, A, B e C. La porta A riporta il prigioniero in cella dopo 2 giorni di lavori forzati; la porta B lo riporta in cella dopo 3 giorni di lavori forzati; infine la porta C ridà al prigioniero la libertà. Il prigioniero sceglie la porta da prendere lanciando un dado equilibrato: se il risultato è pari sceglie A, se esce il numero 1 sceglie B e nei rimanenti casi sceglie C. Se il prigioniero torna in cella, sceglie in modo equiprobabile fra le due porte non ancora scelte. Si determini*

- *la variabile casuale X che indica il numero di giorni impiegati dal prigioniero per uscire;*
- *la funzione di ripartizione di X e se ne faccia il grafico;*
- *media e varianza di X.*

1.7 Variabili Aleatorie Doppie

Variabili aleatorie doppie

Definizione 50. *Sia* $(\Omega, \mathcal{A}, \Pr)$ *uno spazio probabilizzato. Siano* $X(\omega)$ *e* $Y(\omega)$ *due v.a. definite su* Ω *in modo che*

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

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

È facile vedere come può essere definita una v.a. tripla e così via.

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

È necessario considerare la seguente funzione di ripartizione (detta *congiunta*)

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

Definizione 51. *Per due v.a. discrete* X *e* Y*, la v.a. doppia* $Z = (X, Y)$ *(che è discreta) ha funzione di probabilità (congiunta)*

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

cioè

$$\begin{aligned} \Pr(x_1 < X \leq x_2, y_1 < Y \leq y_2) = & \quad F_{X,Y}(x_2, y_2) - F_{X,Y}(x_1, y_2) \\ - & \quad F_{X,Y}(x_2, y_1) + F_{X,Y}(x_1, y_1) \end{aligned}$$

$$F_{X,Y}(x, y) = \sum_{\{(u,v):u \leq x, v \leq y\}} p_{X,Y}(u, v)$$

Variabili aleatorie doppie discrete

Esercizio 74. *Si consideri il lancio di un dado e le due v.a. così definite*

- $X(\omega) = \omega$
-
$$Y(\omega) = \begin{cases} 0 & \textit{“esce un numero dispari”} \\ 1 & \textit{“esce un numero pari minore di 4”} \\ 2 & \textit{“esce un numero pari maggiore o uguale di 4”} \end{cases}$$

Calcorare la funzione di probabilità e di ripartizione della v.a. doppia $Z = (X, Y)$

Soluzione 38. *È facile determinare la funzione di probabilità e di ripartizione di X e Y infatti abbiamo*

$$p_X(X = x) = \frac{1}{6} \qquad F_X(x) = \begin{cases} 0 & x < 1 \\ \frac{[x]}{6} & 1 \leq x \leq 6 \\ 1 & x > 6 \end{cases}$$

dove $[\cdot]$ indica la parte intera.

Mentre

$$p_Y(Y = y) = \begin{cases} \frac{3}{6} & Y = 0 \quad \text{“esce un numero dispari”} \\ \frac{1}{6} & Y = 1 \quad \text{“esce un numero pari minore di 4”} \\ \frac{2}{6} & Y = 2 \quad \text{“esce un numero pari maggiore o uguale di 4”} \end{cases}$$

e

$$F_Y(y) = \begin{cases} 0 & y < 0 \\ \frac{3}{6} & 0 \leq y < 1 \\ \frac{4}{6} & 1 \leq y < 2 \\ 1 & y \geq 2 \end{cases}$$

invece per determinare $p_{X,Y}(x, y)$ e $F_{X,Y}(x, y)$ abbiamo

$X(\omega)$	$Y(\omega)$ 0	1	2	$p_X(X = x)$
1	$\frac{1}{6}$	0	0	$\frac{1}{6}$
2	0	$\frac{1}{6}$	0	$\frac{1}{6}$
3	$\frac{1}{6}$	0	0	$\frac{1}{6}$
4	0	0	$\frac{1}{6}$	$\frac{1}{6}$
5	$\frac{1}{6}$	0	0	$\frac{1}{6}$
6	0	0	$\frac{1}{6}$	$\frac{1}{6}$
$p_Y(Y = y)$	$\frac{3}{6}$	$\frac{1}{6}$	$\frac{2}{6}$	1

e quindi

$$p_{X,Y}(x, y) = \begin{cases} \frac{1}{6} & (x, y) \in \{(1, 0), (3, 0), (5, 0), (2, 1), (4, 2), (6, 2)\} \subset R_{X,Y} \\ 0 & (x, y) \ \text{altrimenti} \end{cases}$$

e la $F_{X,Y}(x, y)$ è

	$Y(\omega)$			
$X(\omega)$	*0*	*1*	*2*	$F_X(x)$
1	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$
2	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{2}{6}$	$\frac{2}{6}$
3	$\frac{2}{6}$	$\frac{3}{6}$	$\frac{3}{6}$	$\frac{3}{6}$
4	$\frac{2}{6}$	$\frac{3}{6}$	$\frac{4}{6}$	$\frac{4}{6}$
5	$\frac{3}{6}$	$\frac{4}{6}$	$\frac{5}{6}$	$\frac{5}{6}$
6	$\frac{3}{6}$	$\frac{4}{6}$	*1*	*1*
$F_Y(y)$	$\frac{3}{6}$	$\frac{4}{6}$	*1*	

Infine, se è nota la funzione di probabilità congiunta $p_{X,Y}(x,y)$ allora è facile ottenere le funzioni di probabilità (marginali) delle v.a. X e Y

$$p_X(X=x) = \sum_{y \in R_Y} p_{X,Y}(x,y) \qquad p_Y(Y=y) = \sum_{x \in R_X} p_{X,Y}(x,y)$$

e quindi

$$F_X(x) = \sum_{u \le x} \sum_{y \in R_Y} p_{X,Y}(u,y) \qquad F_Y(y) = \sum_{v \le y} \sum_{x \in R_X} p_{X,Y}(x,v)$$

Variabili aleatorie doppie dotate di densità

Definizione 52. *La v.a. doppia $Z = (X,Y)$ si dirà dotata di densità se esiste una funzione $f_{X,Y}(x,y)$ tale che*

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

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\ dx\ dy = 1$$

-

$$\Pr(a < x \leq b, c < y \leq d) = \int_a^b \int_c^d f_{X,Y}(x,y)\ dx\ dy$$

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

Da cui abbiamo

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

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

e $f_X(x)$ e $f_Y(y)$ sono chiamate densità marginali delle v.a. X e Y rispettivamente.

Dalle formule precedenti si ottiene anche

$$F_X(x) = \lim_{y\to+\infty} F_{X,Y}(x,y) \qquad F_Y(y) = \lim_{x\to+\infty} F_{X,Y}(x,y)$$

Esercizio 75. *Sia (X,Y) la v.a. doppia dotata della seguente densità*

$$f_{X,Y}(x,y) = \begin{cases} \exp[-x-y] & x>0, y>0 \\ 0 & \textit{altrove} \end{cases}$$

- *Calcolare le densità marginali*
- *Calcolare la funzione di ripartizione congiunta e le marginali*
- *Calcolare la probabilità dell'evento $C = \{(x,y)|x+y>1\}$*

Soluzione 39. *Le densità marginali si ottengono nel modo seguente*

$$\begin{aligned} f_X(x) = & \int_0^{+\infty} f_{X,Y}(x,v)\ dv \\ = & \int_0^{+\infty} \exp[-x-v]\ dv \\ = & \exp[-x]\int_0^{+\infty} \exp[-v]\ dv \\ & \textit{l'integrale è pari a uno perché integriamo la densità} \\ \downarrow & \textit{ di una v.a. esponenziale } (\lambda=1) \textit{ su tutto il suo supporto} \\ = & \exp[-x] \qquad (x>0) \end{aligned}$$

*perciò $f_X(x)$ è la densità di una v.a. esponenziale con $\lambda = 1$. Lo stesso vale per $f_Y(y)$. Si noti che in questo caso **particolare***

$$f_{X,Y}(x,y) = f_X(x)\cdot f_Y(y)$$

La funzione di ripartizione congiunta si ottiene

$$\begin{aligned} F_{X,Y}(x,y) = & \int_0^x \int_0^y \exp[-u-v]\ du\ dv \\ = & \int_0^x \exp[-u]\ du \int_0^y \exp[-v]\ dv \\ = & (1-\exp[-x])(1-\exp[-y]) \\ = & F_X(x)\cdot F_Y(y) \end{aligned}$$

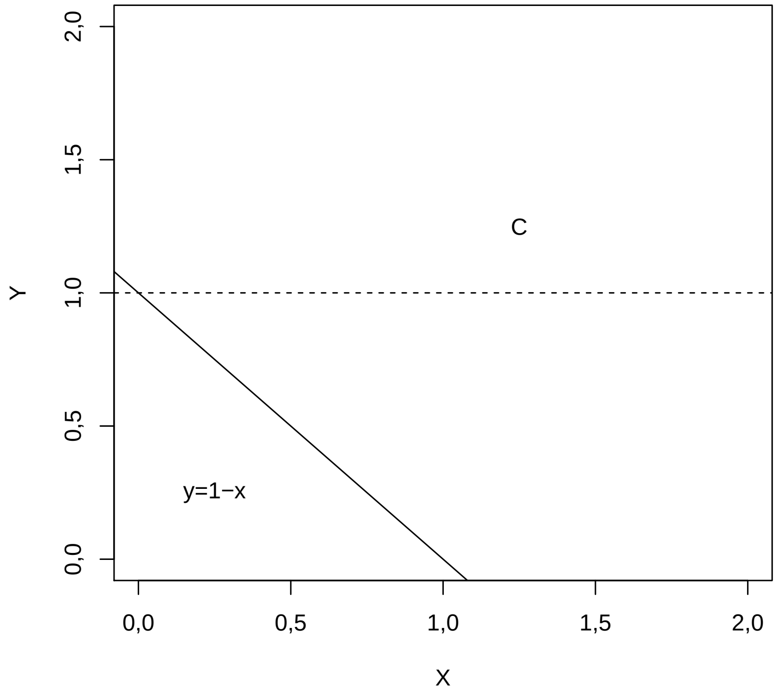


Figura 1.38: .

Per rispondere all'ultimo quesito dobbiamo calcolare $\Pr((X,Y) \in C)$. *Cioè*

Per rispondere all'ultimo quesito dobbiamo calcolare $\Pr((X,Y) \in C)$. *Cioè*

$$\begin{aligned}
\Pr((X,Y) \in C) &= \int\int_C f_{X,Y}(u,v)\; du\; dv = \Pr(X+Y>1) \\
&= \int_0^1 \left(\int_{1-u}^{\infty} \exp[-u] \exp[-v]\; dv \right)\; du \\
&+ \int_1^{\infty} \left(\int_0^{\infty} \exp[-u] \exp[-v]\; dv \right)\; du \\
&= \int_0^1 \exp[-u] \left(\int_{1-u}^{\infty} \exp[-v]\; dv \right)\; du \\
&+ \int_1^{\infty} \exp[-u]\; du \int_0^{\infty} \exp[-v]\; dv \\
&= \int_0^1 \exp[-u] \exp[-(1-u)]\; du + \int_1^{\infty} \exp[-u]\; du \cdot 1 \\
&= \frac{1}{e} \int_0^1 1\; du + \frac{1}{e} = \frac{2}{e}
\end{aligned}$$

Variabili aleatorie doppie

Naturalmente esistono v.a. doppie che non sono né discrete né continue. Un esempio è quando X è una v.a. discreta mentre Y è una v.a. dotata di densità. Noi non tratteremo questo tipo di variabili doppie.

Distribuzioni condizionali per v.a.

Definizione 53 (Probabilità condizionale di $X|Y=y$)**.** *Sia* (X,Y) *una v.a. doppia discreta con funzione di probabilità*

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

allora in accordo con la definizione di probabilità condizionale

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

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

Esercizio 76. *Riprendendo l'ultimo esempio, si calcoli la probabilità condizionale* $p_{X|Y}(X=x|Y=y)$.

	$Y(\omega)$		
$X(\omega)$	0	1	2
1	$\frac{2}{6}$	0	0
2	0	1	0
3	$\frac{2}{6}$	0	0
4	0	0	$\frac{3}{6}$
5	$\frac{2}{6}$	0	0
6	0	0	$\frac{3}{6}$
$p_Y(Y=y)$	$\frac{3}{6}$	$\frac{1}{6}$	$\frac{2}{6}$

Distribuzioni condizionali e indipendenza per v.a.

Definizione 54. *Due v.a. discrete X e Y sono stocasticamente indipendenti se (e solo se)*

$$p_{X,Y}(x,y) = p_X(x) \cdot p_Y(y) \qquad \forall (x,y) \in R_{X,Y} = R_X \times R_Y$$

questo implica che

$$p_{X|Y}(x|y) = p_X(x), p_{Y|X}(y|x) = p_Y(y) \qquad \forall (x,y) \in R_{X,Y}$$

Gli stessi risultati e definizioni valgono anche per le v.a. dotate di densità in particolare abbiamo

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} \qquad \forall y : f_Y(y) > 0$$

e due v.a. X, Y sono stocasticamente indipendenti se e solo se

$$f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y)$$

il che implica (nel caso di indipendenza)

$$f_{X|Y}(x|y) = f_X(x) \qquad \forall (x,y) \in R_{X,Y}$$

Funzioni di ripartizioni condizionali

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

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

e

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

Variabili aleatorie condizionali e speranza matematica

Data la v.a. doppia (X, Y) con funzione di probabilità $p_{X,Y}(x, y)$ allora la funzione $X|Y = y$ $(y \in R_Y)$ è una v.a. con funzione di probabilità $p_{X|Y}(x|y)$.

Quindi dalla definizione di speranza matematica e di Varianza abbiamo

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

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

Esercizio 77. *Continuando l'esercizio si calcoli ad esempio* $\mathbb{E}(X|Y = y)$

$$\begin{aligned}
\mathbb{E}(X|Y = 0) &= 1 \cdot \frac{1}{3} + 3 \cdot \frac{1}{3} + 5 \cdot \frac{1}{3} = 3 \\
\mathbb{E}(X|Y = 1) &= 2 \cdot 1 = 2 \\
\mathbb{E}(X|Y = 2) &= 4 \cdot \frac{1}{2} + 6 \cdot \frac{1}{2} = 5
\end{aligned}$$

Esercizio 78.

$$\begin{aligned}
\operatorname{var}(X|Y = 0) &= (1 - 3)^2 \cdot \frac{1}{3} + 0 + (5 - 3)^2 \cdot \frac{1}{3} = 4\frac{2}{3} = \frac{8}{3} \\
\operatorname{var}(X|Y = 1) &= 0 \\
\operatorname{var}(X|Y = 2) &= (4 - 5)^2 \cdot \frac{1}{2} + (6 - 5)^2 \cdot \frac{1}{2} = 2\frac{1}{2} = 1
\end{aligned}$$

Speranza matematica della speranza matematica condizionale

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

$$\begin{aligned}
\mathbb{E}(\mathbb{E}(X|Y)) &= \sum_{y \in R_Y} \left[\sum_{x \in R_x} x p_{X|Y}(x|y) \right] p_Y(y) \\
&= \sum_{y \in R_Y} \sum_{x \in R_x} x p_{X|Y}(x|y) p_Y(y) \\
&= \sum_{y \in R_Y} \sum_{x \in R_x} x p_{X,Y}(x, y) \\
&= \sum_{x \in R_x} x \sum_{y \in R_Y} p_{X,Y}(x, y) \\
&= \sum_{x \in R_x} x p_X(x) = \mathbb{E}(X)
\end{aligned}$$

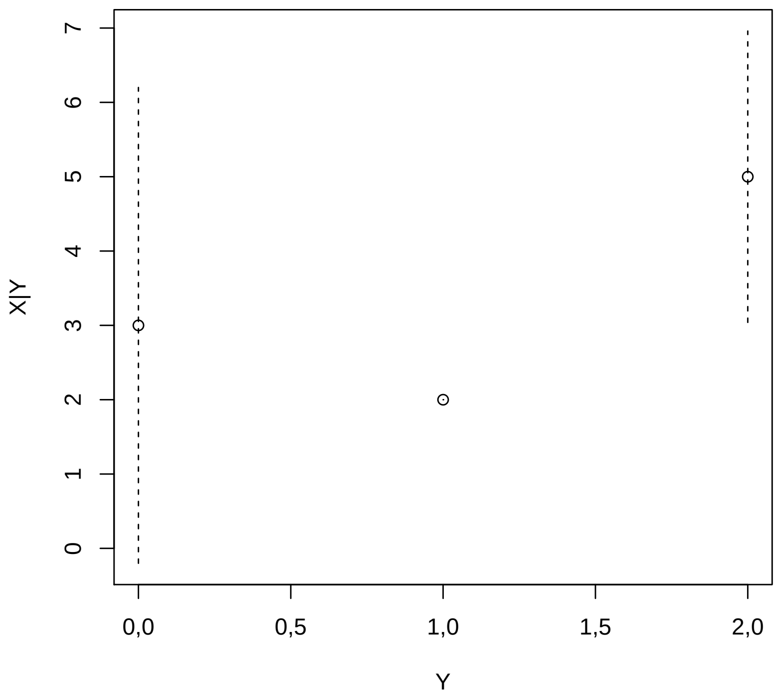


Figura 1.39: .

Varianza e varianza condizionale

Esiste una relazione assai importante tra la varianza di una v.a. e le varianze condizionali di questa v.a. condizionata rispetto alle modalità di un'altra v.a.

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

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

Questa formula viene chiamata **scomposizione della varianza** e assume un ruolo fondamentale in statistica. Essa è il concetto fondamentale utilizzato nel ramo di studi chiamati "Analisi della Varianza" (si veda ad esempio il libro di Scheffe, Analysis of Variance).

Nel caso di v.a. discrete abbiamo

$$\begin{aligned}
\mathrm{var}(X) &= \sum_{x\in R_X} (x-\mathbb{E}(X))^2 p_X(x) \\
&= \sum_{x\in R_X} (x-\mathbb{E}(X))^2 \sum_{y\in R_Y} p_{X,Y}(x,y) \\
&= \sum_{x\in R_X}\sum_{y\in R_Y} (x-\mathbb{E}(X|Y=y)+\mathbb{E}(X|Y=y)-\mathbb{E}(X))^2 p_{X,Y}(x,y) \quad = \sum_{x\in R_X}\sum_{y\in R_Y} (x-\mathbb{E} \\
&+ \sum_{x\in R_X}\sum_{y\in R_Y} (\mathbb{E}(X|Y=y)-\mathbb{E}(X))^2 p_{X,Y}(x,y) \\
&+ 2\sum_{x\in R_X}\sum_{y\in R_Y} (x-\mathbb{E}(X|Y=y))(\mathbb{E}(X|Y=y)-\mathbb{E}(X))p_{X,Y}(x,y) \\
&= \sum_{y\in R_Y}\left[\sum_{x\in R_X}(x-\mathbb{E}(X|Y=y)^2 p_{X|Y}(x|y))\right] p_Y(y) \\
&+ \sum_{y\in R_Y}\left[\sum_{x\in R_X}(\mathbb{E}(X|Y=y)-\mathbb{E}(\mathbb{E}(X|Y)))^2 p_{X|Y}(x|y)\right] p_Y(y) \\
&+ 0 \\
&= \mathbb{E}(\mathrm{var}(X|Y)) + \mathrm{var}(\mathbb{E}(X|Y))
\end{aligned}$$

Rimane da mostrare

$$2\sum_{x\in R_X}\sum_{y\in R_Y} (x-\mathbb{E}(X|Y=y))(\mathbb{E}(X|Y=y)-\mathbb{E}(X))p_{X,Y}(x,y) = 0$$

infatti

$$
\begin{aligned}
&= \sum_{x\in R_X}\sum_{y\in R_Y}(x-\mathbb{E}(X|Y=y))(\mathbb{E}(X|Y=y)-\mathbb{E}(X))p_{X,Y}(x,y)\\
&= \sum_{y\in R_Y}(\mathbb{E}(X|Y=y)-\mathbb{E}(X))\left[\sum_{x\in R_X}(x-\mathbb{E}(X|Y=y))p_{X|Y}(x|y)\right]p_Y(y)\\
&= \sum_{y\in R_Y}(\mathbb{E}(X|Y=y)-\mathbb{E}(X))\left[\sum_{x\in R_X}xp_{X|Y}(x|y)-\mathbb{E}(X|Y=y)\right]p_Y(y)\\
&= \sum_{y\in R_Y}(\mathbb{E}(X|Y=y)-\mathbb{E}(X))\left[\mathbb{E}(X|Y=y)-\mathbb{E}(X|Y=y)\right]p_Y(y)
\end{aligned}
$$

Esercizio 79. *Riprendendo l'esercizio precedente abbiamo*

$$\mathbb{E}(\mathrm{var}(X|Y)) = \left[\frac{8}{3}\frac{3}{6}+0\frac{1}{6}+1\frac{2}{6}\right] = \frac{8}{6}+\frac{2}{6}=\frac{10}{6}$$

$$
\begin{aligned}
\mathrm{var}(\mathbb{E}(X|Y)) &= \left[(3-3.5)^2\frac{3}{6}+(2-3.5)^2\frac{1}{6}+(5-3.5)^2\frac{2}{6}\right]\\
&= \frac{3}{4\cdot 6}+\frac{9}{4\cdot 6}+\frac{9}{2\cdot 6}\\
&= \frac{3+9+9\cdot 2}{4\cdot 6}=\frac{30}{4\cdot 6}=\frac{7.5}{6}
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{var}(X) = &\left[(1-3.5)^2+(2-3.5)^2+(3-3.5)^2+(4-3.5)^2\right.\\
&\left.+(5-3.5)^2+(6-3.5)^2\right]\frac{1}{6}=\frac{17.5}{6}
\end{aligned}
$$

Misure di dipendenza tra v.a.

Perché due v.a. siano stocasticamente dipendenti è sufficiente che la densità (discreta) congiunta non possa essere espressa come prodotto delle densità (discrete) marginali per una qualche coppia $(x,y)\in R_{X,Y}$. È chiaro quindi che vi sono diverse "intensità" e "forme" di dipendenza. Diviene importante allora dare una valutazione numerica all'intensità e alla forma della dipendenza.

Nel seguito presentiamo alcuni modi per misurare il grado di dipendenza tra due v.a.. Ci soffermeremo sulle v.a. discrete, per quelle dotate di densità l'estensione è immediata.

Dipendenza in media

Definizione 55 (Dipendenza in media)**.** *La v.a. X si dice indipendente in media da Y se*

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

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

Indice di dipendenza in media η^2

Definizione 56 (Rapporto di correlazione)**.** *Sia (X, Y) una v.a. doppia discreta, si chiama rapporto di correlazione di X dato Y*

$$\eta^2_{X|Y} = \frac{\mathrm{var}(\mathbb{E}(X|Y))}{\mathrm{var}(X)} = 1 - \frac{\mathbb{E}(\mathrm{var}(X|Y))}{\mathrm{var}(X)} \qquad \mathrm{var}(X) > 0$$

e in modo analogo si definisce $\eta^2_{Y|X}$.

Dalla formula della scomposizione della varianza è facile vedere che

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

inoltre

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

Covarianza e correlazione

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

$$\mathrm{cov}(X, Y) = \mathbb{E}(X \cdot Y) - \mathbb{E}(X) \cdot \mathbb{E}(Y)$$

mentre

$$\rho(X, Y) = \frac{\mathrm{cov}(X, Y)}{\sqrt{\mathrm{var}(X) \cdot \mathrm{var}(Y)}}$$

Varianza di una combinazione lineare di v.a.

Teorema 20. *Sia (X, Y) una v.a. doppia e a e b due costanti. Allora*

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

χ^2 di Pearson

Un indice di dipendenza (stocastica) è il χ^2 di Pearson. Sia (X, Y) una v.a. doppia con funzione di probabilità congiunta $p_{X,Y}(x, y)$ e $p_X(x)$, $p_Y(y)$ le distribuzioni marginali corrispondenti.

Nel caso di indipendenza stocastica ci aspettiamo che

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

Poniamo quindi $\pi_{X,Y}(x, y) = p_X(x)p_Y(y)$ e confrontiamo opportunamente $\pi_{X,Y}(x, y)$ con $p_{X,Y}(x, y)$:

$$\chi^2(X, Y) = \sum_{x \in R_X} \sum_{y \in R_Y} \frac{(p_{X,Y}(x, y) - \pi_{X,Y}(x, y))^2}{\pi_{X,Y}(x, y)}$$

È facile vedere che

$$0 \leq \chi^2(X, Y) < +\infty$$

e che si ha $\chi^2(X, Y) = 0$ se e solo se X e Y sono stocasticamente indipendenti.

1.8 Teoremi limite della Probabilità

Teoremi limite della probabilità

"Me so magna' i schei col 53"
(Anonimo veneziano)

- Legge dei grandi numeri
- Teorema del limite centrale

"Per la legge dei grandi numeri ..."
C'è stata in Italia, alcuni anni or sono, quasi una psicosi (con effetti anche tragici) intorno all'uscita del numero 53 sulla ruota di Venezia nel gioco del lotto. Molti giocatori hanno puntato forti somme su questo numero in quanto la sua estrazione ritardava da molto tempo. Tali giocate erano motivate dall'affermazione che *"prima o poi per la legge dei grandi numeri il numero 53 doveva essere estratto"*.

Indichiamo con Y_i il risultato dell'estrazione i-esima (di 5 numeri in blocco da un'urna che ne contiene 90) sulla ruota di Venezia. NA

Se l'estrazione non è truccata (palline indistinguibili, urna ben mescolata, ...) ogni Y_i ha dunque una distribuzione di Bernoulli con parametro

$$p = \mathbb{E}(Y_i) = 1 \cdot \Pr(Y_i = 1) + 0 \cdot \Pr(Y_i = 0) = \Pr(Y_i = 1) = \frac{5}{90}$$

(perché?).

Inoltre possiamo assumere che ogni Y_i abbia la stessa distribuzione e che i risultati delle estrazioni (ovvero le Y_i) siano tra loro indipendenti.

La variabile $S_n = \sum_{i=1}^n Y_i$ conta il numero di successi (è uscito il 53) in n estrazioni indipendenti e si distribuisce dunque come una v.c. binomiale $\mathcal{B}(n, p)$.

Intuitivamente, siamo portati a dire che per un numero n 'grande' di prove , $n \gg 0$, la frequenza relativa S_n/n si avvicina a p:

$$\left| \frac{S_n}{n} - p \right| < \varepsilon,$$

cioè

$$\frac{S_n}{n} \approx p$$

NA

Ma cosa significa esattamente "si avvicina"?

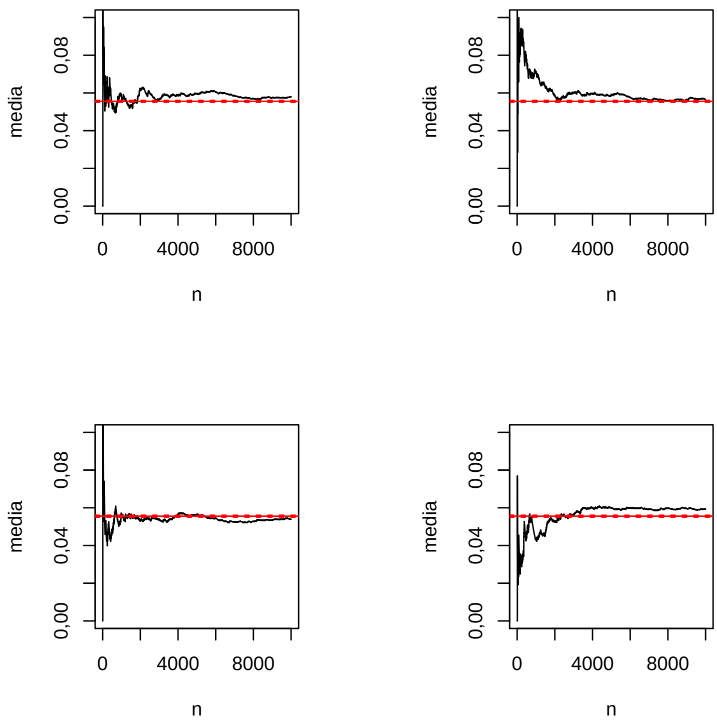

media
0,00
0,04
0,08
0
4000
8000
n
media
0,00
0,04
0,08
0
4000
8000
n
media
0,00
0,04
0,08
0
4000
8000
n
media
0,00
0,04
0,08
0
4000
8000
n

Figura 1.40:

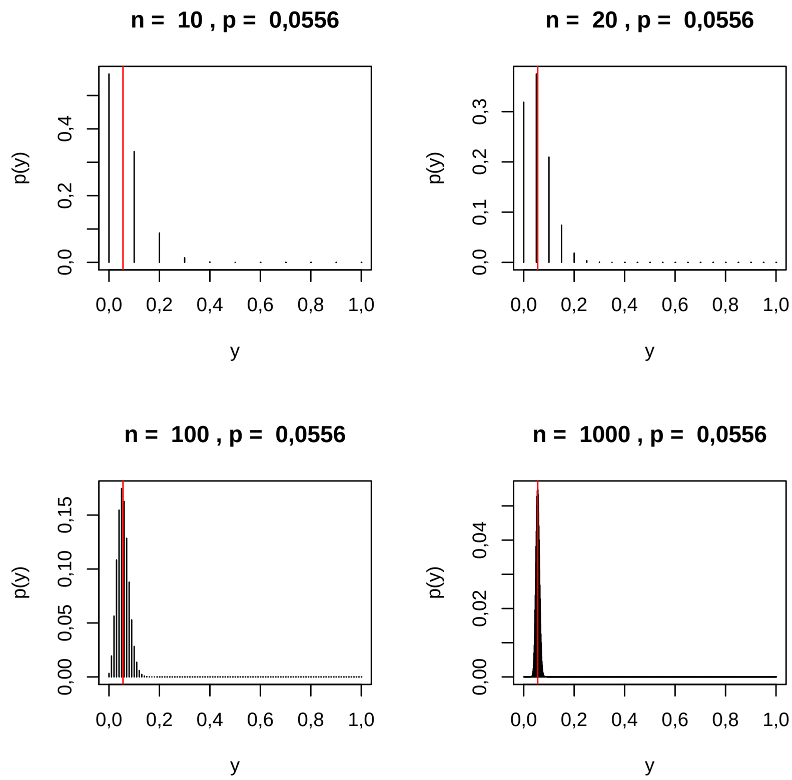


Figura 1.41:

Nel grafico rappresentiamo il risultato di 4 esperimenti con 10.000 estrazioni. Possiamo notare che S_n/n si avvicina a p (la linea rossa), in maniera erratica e in maniera differente per ogni esperimento.

Vediamo cosa accade alla distribuzione di S_n/n quando n aumenta (nei grafici, $p = 5/90$):

La distribuzione tende a concentrarsi intorno al valore $p = 5/90$.

Convergenza in probabilità (o debole)

Ci sono diversi modi per esprimere il fatto che S_n/n si avvicina a p.

Potremmo ad esempio scrivere che, per n grande e per ε piccolo a piacere

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

o equivalentemente

$$\lim_{n\to\infty} \Pr\{|S_n/n - p| \geq \varepsilon\} = 0.$$

In simboli questo tipo di convergenza si denota con

$$S_n/n \xrightarrow{P} p$$

e $\xrightarrow{P}$ si legge "converge in probabilità (o in senso debole)".

Più in generale diremo che una successione $Y_1, Y_2 \dots$ **converge in probabilità (o in senso debole)** ad una v.c. Y se, per ogni $\varepsilon > 0$,

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

ovvero

$$\lim_{n\to\infty} \Pr(|Y_n - Y| < \varepsilon) = 1.$$

Convergenza in media quadratica

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

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

o equivalentemente

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

In simboli questo tipo di convergenza si denota con

$$S_n/n \xrightarrow{m.q.} p$$

e $\xrightarrow{m.q.}$ si legge "converge in media quadratica".

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

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

Proposizione 2. *La convergenza in media quadratica implica la convergenza in probabilità:*

$$Y_n \xrightarrow{m.q.} Y \Longrightarrow Y_n \xrightarrow{P} Y$$

Per dimostrare questa proposizione ci servono alcuni risultati preliminari.

Proposizione 3. *(Disuguaglianza di Markov) Sia Y una v.c. che assume valori non negativi, allora per ogni numero reale $a > 0$*

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

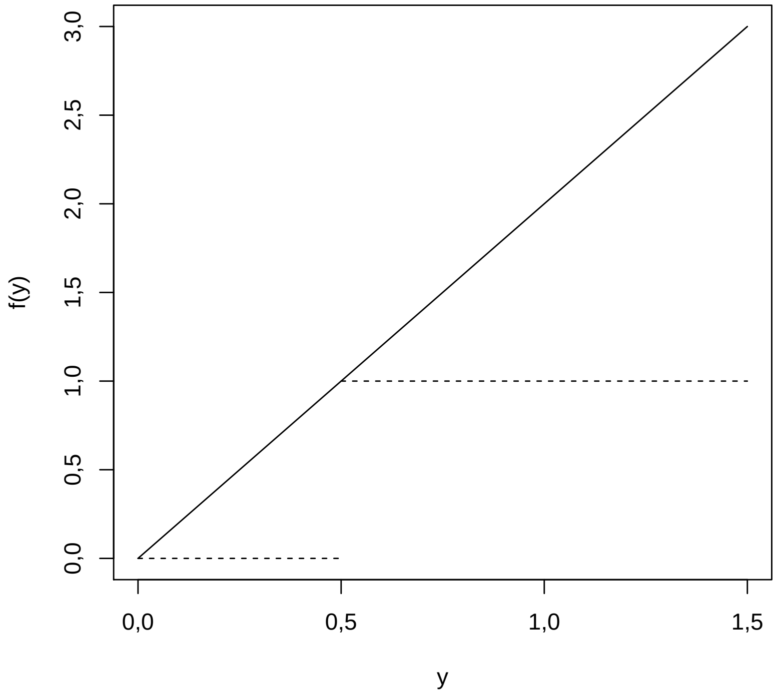


Figura 1.42:

Dimostrazione 16. *Sia*

$$X = \begin{cases} 1 & se\ Y \geq a \\ 0 & altrimenti \end{cases}$$

*Si noti che, poiché $Y \geq 0$ abbiamo $X \leq Y/a$ e quindi, per la proposizione [**??**], $\mathbb{E}(X) \leq \mathbb{E}(Y)/a$, con $\mathbb{E}(X) = 1 \cdot \Pr(X = 1) + 0 \cdot \Pr(X = 0) = \Pr(X = 1) = \Pr(Y \geq a)$.*

Siamo ora in grado di dimostrare la Proposizione [2].

Dimostrazione 17. *Sia $\varepsilon > 0$. Grazie alla disuguaglianza di Markov applicata a $(Y_n - Y)^2$, si ha che*

$$\begin{aligned} \lim_{n\to\infty} \Pr(|Y_n - Y| \geq \varepsilon) &= \lim_{n\to\infty} \Pr((Y_n - Y)^2 \geq \varepsilon^2) \\ &\leq \lim_{n\to\infty} \frac{\mathbb{E}[(Y_n - Y)^2]}{\varepsilon^2}. \end{aligned}$$

Inoltre, per ipotesi, $\mathbb{E}[(Y_n - Y)^2] \longrightarrow 0$ e perciò $\lim_{n\to\infty} \Pr(|Y_n - Y| \geq \varepsilon) = 0$.

Dalla disuguaglianza di Markov discende un altro importante risultato:

Proposizione 4. *(Disuguaglianza di Chebychev) Sia Y una v.c. con valore atteso $\mathbb{E}(Y) = \mu$ e varianza $\mathbb{V}ar(Y) = \sigma^2$. Allora*

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

Dimostrazione 18. *Essendo $(Y - \mu)^2$ una v.c. non negativa, basta applicare la disuguaglianza di Markov con $a = \varepsilon^2$:*

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

Ma $(Y - \mu)^2 \geq \varepsilon^2$ se e solo se $|Y - \mu| \geq \varepsilon$ e quindi

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

L'importanza delle due disuguaglianze di Markov e Chebychev è che ci danno dei limiti superiori per delle probabilità senza alcun riferimento alla distribuzione di probabilità della v.c. Y. Basta conoscerne il valore atteso o il valore atteso e la varianza.

Ad esempio, si supponga che il numero di clienti in un giorno ad uno sportello sia una v.c. Y con $\mathbb{E}(Y) = 130$. Cosa possiamo dire della probabilità che possano arrivare più di 250 clienti? Grazie alla disuguaglianza di Markov possiamo dire che $\Pr(Y \geq 250) \leq 130/250 = 0.52$.

Ovviamente questi limiti possono essere molto grossolani. Se $Y \sim \mathcal{N}(0, 1)$ sappiamo che $\Pr(|Y| \geq 1.96) = 0.05$, mentre la disuguaglianza di Chebychev afferma che $\Pr(|Y| \geq 1.96) \leq 1/(1.96^2) \approx 0.260$.

Richiami sulle somme variabili casuali

Proposizione 5. *Siano* $Y_1, \ldots, Y_n$ *v.c. con valore atteso rispettivamente* $\mu_1, \ldots, \mu_n$. *Allora*

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

Proposizione 6. *Siano* $Y_1, \ldots, Y_n$ *v.c. indipendenti con varianza* $\sigma_1^2, \ldots, \sigma_n^2$ *rispettivamente. Allora*

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

Una conseguenza importante di queste due proposizioni è la seguente:

Proposizione 7. *Siano* $Y_1, \ldots, Y_n$ *v.c. indipendenti, tutte con valore atteso* μ *e varianza* σ^2 *e sia* $\overline{Y}_n = \sum_{i=1}^n Y_i/n$. *Allora*

$$\mathbb{E}(\overline{Y}_n) = \sum_{i=1}^{n} \frac{\mathbb{E}(Y_i)}{n} = n\frac{\mu}{n} = \mu,$$

$$\mathbb{V}ar(\overline{Y}_n) = \sum_{i=1}^{n} \frac{\mathbb{V}ar(Y_i)}{n^2} = n\frac{\sigma^2}{n^2} = \frac{\sigma^2}{n}.$$

La legge debole dei grandi numeri

Proposizione 8 (Legge debole dei grandi numeri)**.** *Sia* $Y_1, Y_2, \ldots$ *una successione di v.c. indipendenti, ciascuna con valore atteso* μ *e varianza* σ^2. *Allora, per ogni* $\varepsilon > 0$,

$$\lim_{n\to\infty} \Pr\{|\overline{Y}_n - \mu| \geq \varepsilon\} = 0,$$

ovvero $\overline{Y}_n \xrightarrow{P} \mu$.

Dimostrazione 19. *Per la Proposizione [7], quando* $n \longrightarrow \infty$,

$$\mathbb{E}[(\overline{Y}_n - \mu)^2] = \mathbb{V}ar(\overline{Y}_n) = \frac{\sigma^2}{n} \longrightarrow 0,$$

che significa che $\overline{Y}_n \xrightarrow{m.q.} \mu$. *La convergenza debole segue allora dalla Proposizione [2].*

"Per la legge dei grandi numeri ...", un'amara conclusione

Tornando al gioco del lotto, la legge dei grandi numeri garantisce che $S_n/n \xrightarrow{P} p$ e quindi che la frequenza relativa del 53 tende a 5/90 quando il numero di estrazioni cresce.

Nella nostra discussione abbiamo supposto l'indipendenza e l'identica distribuzione delle v.c. Y_i. Nel gioco del lotto questo vuol dire che le estrazioni sono indipendenti e che la probabilità di estrarre il 53 è la stessa nelle varie estrazioni.

Anche se l'indipendenza non è un'ipotesi fondamentale per provare delle leggi dei grandi numeri (ma questo richiede nozioni più avanzate del calcolo delle probabilità), lo è nel gioco del lotto. E quindi anche se la frequenza dei successi tende in probabilità alla probabilità di successo, ogni volta che avviene una nuova estrazione il 53 ha **sempre la medesima probabilità** di uscire (sic!).

Il teorema del limite centrale

Ci occupiamo qui di un diverso tipo di convergenza. Abbiamo visto che, sotto determinate ipotesi, S_n/n converge debolmente alla probabilità di successo p. Consideriamo ora invece di S_n/n una sua espressione standardizzata

$$Z_n = \sqrt{n}\frac{S_n/n - p}{\sqrt{p(1-p)}}.$$

Nei grafici abbiamo riportato la funzione di ripartizione della variabile Z_n (con $p = 2$) e, con la linea continua, la funzione di ripartizione di una v.c. $\mathcal{N}(0,1)$.

Proposizione 9 (Teorema del limite centrale)**.** *Sia* $Y_1, Y_2, \ldots$ *una successione di v.c. ciascuna con* $\mathbb{E}(Y_i) = \mu$ *e* $\mathbb{V}ar(Y_i) = \sigma^2$*. Allora, posto* $Z_n = \sqrt{n}(\overline{Y}_n - \mu)/\sigma$*, per ogni* z

$$\lim_{n\to\infty} \Pr(Z_n \leq z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2} dt\,.$$

In simboli il teorema del limite centrale si denota scrivendo

$$Z_n = \sqrt{n}\frac{(\overline{Y}_n - \mu)}{\sigma} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1),$$

dove $\xrightarrow{\mathcal{D}}$ si legge "converge in distribuzione".

Una lettura grezza (ma pratica) del teorema del limite centrale NA

$$\overline{Y}_n \overset{a}{\sim} \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right),$$

dove $\overset{a}{\sim}$ significa "distribuita approssimatamente".

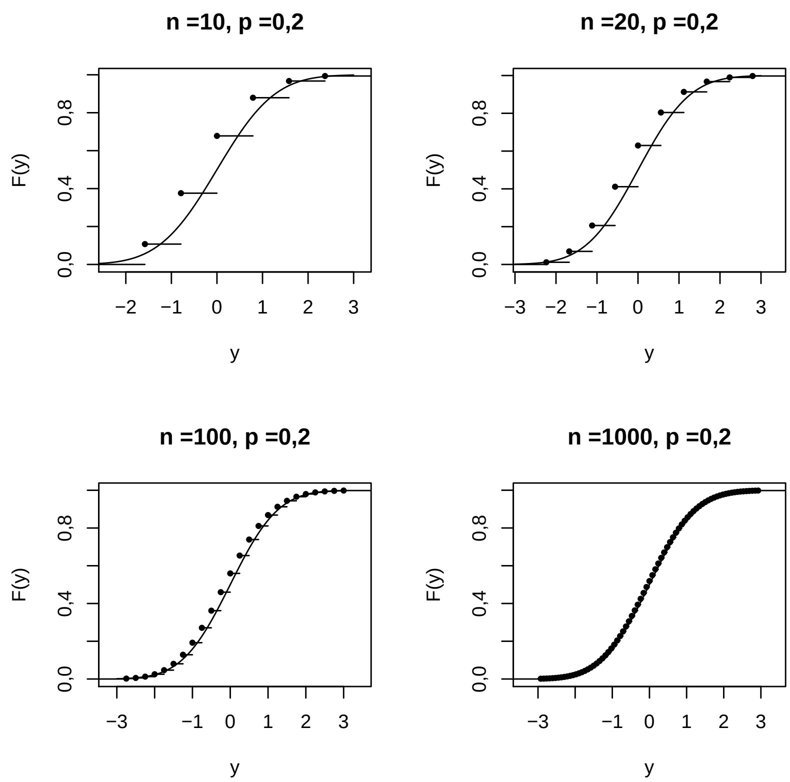


Figura 1.43:

Somme di v.c. normali

Il teorema del limite centrale è un risultato valido solo asintoticamente, cioè quando $n \to \infty$.

Se le v.c. sono normali, la distribuzione esatta di $\overline{Y}_n$ può essere ottenuta dalla proposizione seguente.

Proposizione 10. *Se $Y_1, \ldots, Y_n$ sono variabili casuali normali* indipendenti *tra loro e se $a_0, \ldots, a_n$ sono delle costanti reali qualsiasi, allora*

$$a_0 + \sum_{i=1}^{n} a_i Y_i \sim \mathcal{N}\left(a_0 + \sum_{i=1}^{n} a_i \mu_i, \sum_{i=1}^{n} a_i^2 \sigma_i^2\right)$$

dove μ_i e σ_i^2 indicano rispettivamente la media e la varianza di Y_i.

Esempio 16. *Nel caso della media $\overline{Y}_n$ di n v.c. normali Y_i, indipendenti e con medesime media μ e varianza σ^2, si ha $a_0 = 0$ e $a_i = 1/n$, $i = 1, \ldots, n$. Quindi:*

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

Approssimare la binomiale con la normale

All'aumentare del numero n di prove, la distribuzione di una variabile casuale binomiale di parametri n e p si "avvicina" sempre di più a quella di una normale con parametri $\mu = np$ e $\sigma^2 = np(1-p)$.

Alcune considerazioni:

- L'approssimazione è valida quando n è abbastanza grande e p e $1-p$ non sono vicini a zero.
- Regola pratica
  1. si calcola l'intervallo di estremi $np \pm 3\sqrt{np(1-p)}$
  2. se esso è contenuto nell'intervallo $[0, n]$ allora l'approssimazione può ritenersi valida.

Esempio 17. *Vogliamo calcolare* $\Pr(B \leq 25)$ *dove $B \sim \mathcal{B}(30, 0.7)$. L'intervallo* $[13.47, 28.53]$ *è contenuto nell'intervallo* $[0, 30]$ *quindi procediamo con l'approssimazione. Calcoliamo il valore standardizzato corretto corrispondente a 25:*

$$z_0 = \frac{(25 + \mathbf{0.5}) - 30 \cdot 0.7}{\sqrt{30 \cdot 0.7(1 - 0.7)}} \simeq 1.79$$

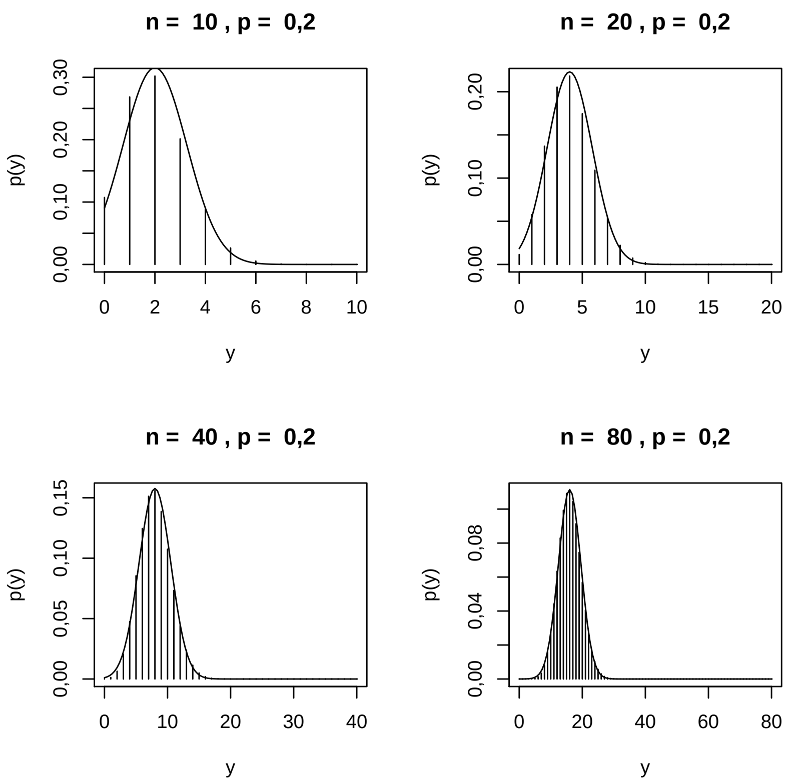

n = 10 , p = 0,2
p(y)
0,00
0,10
0,20
0,30
0
2
4
6
8
10
y
n = 20 , p = 0,2
p(y)
0,00
0,10
0,20
0
5
10
15
20
y
n = 40 , p = 0,2
p(y)
0,00
0,05
0,10
0,15
0
10
20
30
40
y
n = 80 , p = 0,2
p(y)
0,00
0,04
0,08
0
20
40
60
80
y

Figura 1.44:

e utilizziamo le tavole della distribuzione normale:

$$\Pr(B \leq 25) \simeq \Pr(Z \leq 1.79) = 0.9633.$$

Si confronti questa probabilità con la probabilità esatta 0.9699.

Si osservi che nel calcolo di z_0 *abbiamo aggiunto 0.5 al valore dato 25. Si tratta della cosiddetta* ***correzione per continuità*** *che migliora l'approssimazione di una v.c. discreta con una continua.*

Bibliografia

P. Baldi. *Calcolo delle probabilità e statistica.* McGraw-Hill, 1998.

P. Baldi, R. Giuliano, and L. Ladelli. *Laboratorio di probabilità e statistica. Problemi svolti.* McGraw-Hill, 1998.

G. Cicchitelli. *Probabilità e Statistica.* Maggioli editore, Rimini, seconda edition, 2001. ISBN 88–387–1672.2.

D.M. Cifarelli. *Introduzione al calcolo delle probabilità.* McGraw-Hill, 1998.

L. Pace and A. Salvan. *Introduzione alla Statistica - I. Statistica descrittiva.* Cedam, Padova, 1996. ISBN 88–13–19939–2.

L. Pace and A. Salvan. *Introduzione alla Statistica - II. Inferenza, verosimiglianza, modelli.* Cedam, Padova, 2001.

F. Parpinel and C. Provasi. *Elementi di probabilità e statistica per le scienze economiche.* Giappichelli, Torino, 2004.

S. Ross. *Probabilità e statistica per l'ingegneria e le scienze.* Apogeo, 2008.

Appendice A

Lettere dell'alfabeto greco antico

Lettere dell'alfabeto greco antico e loro pronuncia internazionale (e in italiano)

A	α	alpha
B	β	beta
Γ	γ	gamma
Δ	δ	delta
E	ϵ, ε	epsilon
Z	ζ	zeta
H	η	eta
Θ	θ, ϑ	theta
I	ι	iota
K	κ	kappa
Λ	λ	lambda
M	μ	mu (my, mi)
N	ν	nu (ny, ni)
Ξ	ξ	xi
O	o	omicron
Π	π, ϖ	pi
P	ρ, ϱ	rho
Σ	σ, ς	sigma
T	τ	tau
Υ	υ	upsilon (ypsilon)
Φ	ϕ, φ	phi
X	χ	chi
Ψ	ψ	psi
Ω	ω	omega

Appendice B

Strumenti di base

Sommatoria

Siano a_i e b_i, $i = 1, \ldots, n$, due serie di numeri e c una costante qualsiasi. Si definisce la **sommatoria** nel modo seguente:

$$\sum_{i=1}^{n} a_i = a_1 + a_2 + \ldots + a_n \,.$$

Alcune proprietà

1. $\sum_{i=1}^{n} c = c + c + \ldots + c = n\,c$.

2. $\sum_{i=1}^{n} c\,a_i = c\,a_1 + c\,a_2 + \ldots + c\,a_n = c\,(a_1 + a_2 + \ldots + a_n) = c\,\sum_{i=1}^{n} a_i$.

3. $\sum_{i=1}^{n}(a_i + b_i) = \sum_{i=1}^{n} a_i + \sum_{i=1}^{n} b_i$.

4. $\left(\sum_{i=1}^{n} a_i\right)^c \neq \sum_{i=1}^{n} a_i^c$. Ad esempio, con $n = 2$ e $c = 2$, si ha $\left(\sum_{i=1}^{n} a_i\right)^2 = (a_1 + a_2)^2 = a_1^2 + 2\,a_1\,a_2 + a_2^2 \neq a_1^2 + a_2^2 = \sum_{i=1}^{n} a_i^2$.

5. $\sum_{i=1}^{n}(a_i\,b_i) \neq \sum_{i=1}^{n} a_i\,\sum_{i=1}^{n} b_i$. Ad esempio, con $n = 2$, si ha $\sum_{i=1}^{n}(a_i\,b_i) = a_1\,b_1 + a_2\,b_2 \neq (a_1 + a_2)\,(b_1 + b_2) = a_1\,b_1 + a_1\,b_2 + a_2\,b_1 + a_2\,b_2 = \sum_{i=1}^{n} a_i\,\sum_{i=1}^{n} b_i$.

La somma dei primi n interi[1]

Un caso famoso è quello della somma dei primi n numeri

$$\sum_{i=1}^{n} i = 1 + 2 + \cdots + n = \frac{n(n+1)}{2} \,.$$

[1] Si veda ad esempio Polymath

La dimostrazione più rapida è quella di Gauss, che secondo un aneddoto determinò in poco tempo, durante le elementari, la somma dei primi 100 numeri naturali, rispondendo alla richiesta del suo maestro che lo voleva tenere impegnato per un po' di tempo per potersi dedicare anche agli altri bambini, osservando che le somme del primo e dell'ultimo, del secondo e del penultimo, e così via, sono uguali. La dimostrazione per induzione della formula di Gauss, chiamata $P(n)$, è la seguente:

- (base) Per $n = 1$ si ha $1 = (1/2)1(1+1)$ e il secondo membro si riduce a 1.
- (passo induttivo) Ammesso per ipotesi induttiva $P(n)$ cioè che

$$P(n) = \sum_{i=1}^{n} i = \frac{n(n+1)}{2} \ .$$

si ha

$$\sum_{i=1}^{n+1} i = \sum_{i=1}^{n} i + n + 1 = \frac{n(n+1)}{2} + n + 1$$

e il secondo membro con facili calcoli si riduce a $(n+1)(n+2)/2$, e quindi si ha

$$P(n+1) = \sum_{i=1}^{n+1} i = \frac{(n+2)(n+1)}{2} \ .$$

Produttoria

Siano a_i e b_i, $i = 1, \ldots, n$, due serie di numeri e c una costante qualsiasi. Si definisce la **produttoria** nel modo seguente:

$$\prod_{i=1}^{n} a_i = a_1\, a_2\, \ldots a_n \, .$$

Alcune proprietà

1. $\prod_{i=1}^{n} c = c\,c\,\ldots c = c^n$.
2. $\prod_{i=1}^{n} c\,a_i = c\,a_1\,c\,a_2 \ldots c\,a_n = c^n\,(a_1\,a_2 \ldots a_n) = c^n \prod_{i=1}^{n} a_i$.
3. $\prod_{i=1}^{n}(a_i\,b_i) = \prod_{i=1}^{n} a_i \prod_{i=1}^{n} b_i$.
4. $\left(\prod_{i=1}^{n} a_i\right)^c = \prod_{i=1}^{n} a_i^c$.
5. $\prod_{i=1}^{n}(a_i + b_i) \neq \prod_{i=1}^{n} a_i + \prod_{i=1}^{n} b_i$.

Logaritmo

Il logaritmo in base b del numero reale positivo x, $l = \log_b x$, si definisce come l'esponente l da dare a b per ottenere x, ossia: $x = b^l$. Le basi più comunemente utilizzate sono quella *naturale*, ossia il numero neperiano $e = 2.7183\ldots$, e la base 10.

Alcune proprietà

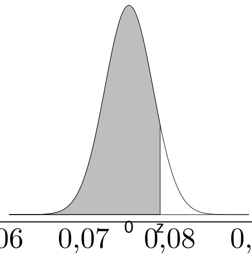
1. $\log_b 1 = 0\,.$
2. $\log_b b = 1\,.$
3. $\lim_{x\to 0^+} \log_b x = -\infty$; $\lim_{x\to +\infty} \log_b x = +\infty$
4. $\log_b(x\,y) = \log_b x + \log_b y$
   (più in generale, $\log_b\left(\prod_{i=1}^n x_i\right) = \sum_{i=1}^n \log_b x_i$).
5. $\log_b(x^c) = c\log_b x$.
6. $\log_b(x/y) = \log_b(x\,y^{-1}) = \log_b x + \log_b y^{-1} = \log_b x - \log_b y$.
7. Cambio di base: $\log_b x = \log_b c \log_c x$.

Appendice C

Tavole di alcune variabili casuali

C.1 Funzione di ripartizione della distribuzione normale standard

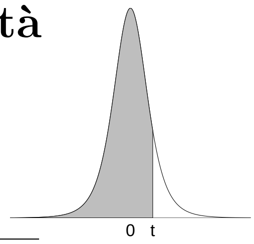
$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \, dx$$



	0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998

C.2 Alcuni quantili della distribuzione t di Student con r gradi di libertà

$$F_r(t) = \int_{-\infty}^{t} \frac{\Gamma[(r+1)/2]}{\sqrt{\pi r}\Gamma[r/2](1+x^2/r)^{(r+1)/2}}\ dx$$



	0,6	0,75	0,9	0,95	0,975	0,99	0,995
1	0,3249	1,0000	3,0777	6,3138	12,7062	31,8205	63,6567
2	0,2887	0,8165	1,8856	2,9200	4,3027	6,9646	9,9248
3	0,2767	0,7649	1,6377	2,3534	3,1824	4,5407	5,8409
4	0,2707	0,7407	1,5332	2,1318	2,7764	3,7469	4,6041
5	0,2672	0,7267	1,4759	2,0150	2,5706	3,3649	4,0321
6	0,2648	0,7176	1,4398	1,9432	2,4469	3,1427	3,7074
7	0,2632	0,7111	1,4149	1,8946	2,3646	2,9980	3,4995
8	0,2619	0,7064	1,3968	1,8595	2,3060	2,8965	3,3554
9	0,2610	0,7027	1,3830	1,8331	2,2622	2,8214	3,2498
10	0,2602	0,6998	1,3722	1,8125	2,2281	2,7638	3,1693
11	0,2596	0,6974	1,3634	1,7959	2,2010	2,7181	3,1058
12	0,2590	0,6955	1,3562	1,7823	2,1788	2,6810	3,0545
13	0,2586	0,6938	1,3502	1,7709	2,1604	2,6503	3,0123
14	0,2582	0,6924	1,3450	1,7613	2,1448	2,6245	2,9768
15	0,2579	0,6912	1,3406	1,7531	2,1314	2,6025	2,9467
16	0,2576	0,6901	1,3368	1,7459	2,1199	2,5835	2,9208
17	0,2573	0,6892	1,3334	1,7396	2,1098	2,5669	2,8982
18	0,2571	0,6884	1,3304	1,7341	2,1009	2,5524	2,8784
19	0,2569	0,6876	1,3277	1,7291	2,0930	2,5395	2,8609
20	0,2567	0,6870	1,3253	1,7247	2,0860	2,5280	2,8453
21	0,2566	0,6864	1,3232	1,7207	2,0796	2,5176	2,8314
22	0,2564	0,6858	1,3212	1,7171	2,0739	2,5083	2,8188
23	0,2563	0,6853	1,3195	1,7139	2,0687	2,4999	2,8073
24	0,2562	0,6848	1,3178	1,7109	2,0639	2,4922	2,7969
25	0,2561	0,6844	1,3163	1,7081	2,0595	2,4851	2,7874
26	0,2560	0,6840	1,3150	1,7056	2,0555	2,4786	2,7787
27	0,2559	0,6837	1,3137	1,7033	2,0518	2,4727	2,7707
28	0,2558	0,6834	1,3125	1,7011	2,0484	2,4671	2,7633
29	0,2557	0,6830	1,3114	1,6991	2,0452	2,4620	2,7564
30	0,2556	0,6828	1,3104	1,6973	2,0423	2,4573	2,7500
50	0,2547	0,6794	1,2987	1,6759	2,0086	2,4033	2,6778
75	0,2542	0,6778	1,2929	1,6654	1,9921	2,3771	2,6430
100	0,2540	0,6770	1,2901	1,6602	1,9840	2,3642	2,6259
∞	0,2533	0,6745	1,2816	1,6449	1,9600	2,3263	2,5758

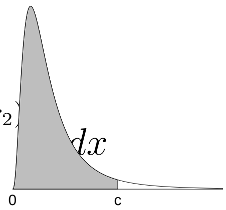
C.3 Alcuni quantili della distribuzione χ^2 con r gradi di libertà

$$F_r(c) = \int_0^c \frac{1}{\Gamma[r/2]2^{r/2}} x^{r/2-1} \exp\left[-\frac{x}{2}\right] dx$$

	0,005	0,01	0,025	0,05	0,1	0,9	0,95	0,975	0,99	0,995
1	0,000	0,000	0,001	0,004	0,016	2,706	3,841	5,024	6,635	7,879
2	0,010	0,020	0,051	0,103	0,211	4,605	5,991	7,378	9,210	10,597
3	0,072	0,115	0,216	0,352	0,584	6,251	7,815	9,348	11,345	12,838
4	0,207	0,297	0,484	0,711	1,064	7,779	9,488	11,143	13,277	14,860
5	0,412	0,554	0,831	1,145	1,610	9,236	11,070	12,833	15,086	16,750
6	0,676	0,872	1,237	1,635	2,204	10,645	12,592	14,449	16,812	18,548
7	0,989	1,239	1,690	2,167	2,833	12,017	14,067	16,013	18,475	20,278
8	1,344	1,646	2,180	2,733	3,490	13,362	15,507	17,535	20,090	21,955
9	1,735	2,088	2,700	3,325	4,168	14,684	16,919	19,023	21,666	23,589
10	2,156	2,558	3,247	3,940	4,865	15,987	18,307	20,483	23,209	25,188
11	2,603	3,053	3,816	4,575	5,578	17,275	19,675	21,920	24,725	26,757
12	3,074	3,571	4,404	5,226	6,304	18,549	21,026	23,337	26,217	28,300
13	3,565	4,107	5,009	5,892	7,042	19,812	22,362	24,736	27,688	29,819
14	4,075	4,660	5,629	6,571	7,790	21,064	23,685	26,119	29,141	31,319
15	4,601	5,229	6,262	7,261	8,547	22,307	24,996	27,488	30,578	32,801
16	5,142	5,812	6,908	7,962	9,312	23,542	26,296	28,845	32,000	34,267
17	5,697	6,408	7,564	8,672	10,085	24,769	27,587	30,191	33,409	35,718
18	6,265	7,015	8,231	9,390	10,865	25,989	28,869	31,526	34,805	37,156
19	6,844	7,633	8,907	10,117	11,651	27,204	30,144	32,852	36,191	38,582
20	7,434	8,260	9,591	10,851	12,443	28,412	31,410	34,170	37,566	39,997
21	8,034	8,897	10,283	11,591	13,240	29,615	32,671	35,479	38,932	41,401
22	8,643	9,542	10,982	12,338	14,041	30,813	33,924	36,781	40,289	42,796
23	9,260	10,196	11,689	13,091	14,848	32,007	35,172	38,076	41,638	44,181
24	9,886	10,856	12,401	13,848	15,659	33,196	36,415	39,364	42,980	45,559
25	10,520	11,524	13,120	14,611	16,473	34,382	37,652	40,646	44,314	46,928
26	11,160	12,198	13,844	15,379	17,292	35,563	38,885	41,923	45,642	48,290
27	11,808	12,879	14,573	16,151	18,114	36,741	40,113	43,195	46,963	49,645
28	12,461	13,565	15,308	16,928	18,939	37,916	41,337	44,461	48,278	50,993
29	13,121	14,256	16,047	17,708	19,768	39,087	42,557	45,722	49,588	52,336
30	13,787	14,953	16,791	18,493	20,599	40,256	43,773	46,979	50,892	53,672
40	20,707	22,164	24,433	26,509	29,051	51,805	55,758	59,342	63,691	66,766
50	27,991	29,707	32,357	34,764	37,689	63,167	67,505	71,420	76,154	79,490
60	35,534	37,485	40,482	43,188	46,459	74,397	79,082	83,298	88,379	91,952
70	43,275	45,442	48,758	51,739	55,329	85,527	90,531	95,023	100,425	104,215
80	51,172	53,540	57,153	60,391	64,278	96,578	101,879	106,629	112,329	116,321

C.4 Alcuni quantili della distribuzione F con r_1 e r_2 gradi di libertà

$$F_{r_1,r_2}(c) = \int_0^c \frac{\Gamma((r_1+r_2)/2)}{\Gamma(r_1/2)\Gamma(r_2/2)} \left(\frac{r_1}{r_2}\right)^{r_1/2} x^{(r_1/2-1)} \left(1+\frac{r_1}{r_2}x\right)^{-(r_1+r_2)/2} dx$$



Per ogni coppia di r_1 (colonna) e r_2 (riga), la tavola fornisce il quantile f_α di ordine α corrispondente. I quantili inferiori della distribuzione F di Fisher–Snedecor si possono determinare tramite la relazione $f_{1-\alpha}(r_1, r_2) = 1/f_\alpha(r_2, r_1)$.

Tavole per l'ordine quantilico α= 0,95

	1	2	3	4	5	6	7	8	9
1	161,448	199,500	215,707	224,583	230,162	233,986	236,768	238,883	240,543
2	18,513	19,000	19,164	19,247	19,296	19,330	19,353	19,371	19,385
3	10,128	9,552	9,277	9,117	9,013	8,941	8,887	8,845	8,812
4	7,709	6,944	6,591	6,388	6,256	6,163	6,094	6,041	5,999
5	6,608	5,786	5,409	5,192	5,050	4,950	4,876	4,818	4,772
6	5,987	5,143	4,757	4,534	4,387	4,284	4,207	4,147	4,099
7	5,591	4,737	4,347	4,120	3,972	3,866	3,787	3,726	3,677
8	5,318	4,459	4,066	3,838	3,687	3,581	3,500	3,438	3,388
9	5,117	4,256	3,863	3,633	3,482	3,374	3,293	3,230	3,179
10	4,965	4,103	3,708	3,478	3,326	3,217	3,135	3,072	3,020
11	4,844	3,982	3,587	3,357	3,204	3,095	3,012	2,948	2,896
12	4,747	3,885	3,490	3,259	3,106	2,996	2,913	2,849	2,796
13	4,667	3,806	3,411	3,179	3,025	2,915	2,832	2,767	2,714
14	4,600	3,739	3,344	3,112	2,958	2,848	2,764	2,699	2,646
15	4,543	3,682	3,287	3,056	2,901	2,790	2,707	2,641	2,588
16	4,494	3,634	3,239	3,007	2,852	2,741	2,657	2,591	2,538
17	4,451	3,592	3,197	2,965	2,810	2,699	2,614	2,548	2,494
18	4,414	3,555	3,160	2,928	2,773	2,661	2,577	2,510	2,456
19	4,381	3,522	3,127	2,895	2,740	2,628	2,544	2,477	2,423
20	4,351	3,493	3,098	2,866	2,711	2,599	2,514	2,447	2,393
25	4,242	3,385	2,991	2,759	2,603	2,490	2,405	2,337	2,282
30	4,171	3,316	2,922	2,690	2,534	2,421	2,334	2,266	2,211
40	4,085	3,232	2,839	2,606	2,449	2,336	2,249	2,180	2,124
50	4,034	3,183	2,790	2,557	2,400	2,286	2,199	2,130	2,073
60	4,001	3,150	2,758	2,525	2,368	2,254	2,167	2,097	2,040
120	3,920	3,072	2,680	2,447	2,290	2,175	2,087	2,016	1,959
∞	3,841	2,996	2,605	2,372	2,214	2,099	2,010	1,938	1,880

Tavole per l'ordine quantilico α= 0,95

	10	15	20	30	40	50	60	120	∞
1	241,882	245,950	248,013	250,095	251,143	251,774	252,196	253,253	254,314
2	19,396	19,429	19,446	19,462	19,471	19,476	19,479	19,487	19,496
3	8,786	8,703	8,660	8,617	8,594	8,581	8,572	8,549	8,526
4	5,964	5,858	5,803	5,746	5,717	5,699	5,688	5,658	5,628
5	4,735	4,619	4,558	4,496	4,464	4,444	4,431	4,398	4,365
6	4,060	3,938	3,874	3,808	3,774	3,754	3,740	3,705	3,669
7	3,637	3,511	3,445	3,376	3,340	3,319	3,304	3,267	3,230
8	3,347	3,218	3,150	3,079	3,043	3,020	3,005	2,967	2,928
9	3,137	3,006	2,936	2,864	2,826	2,803	2,787	2,748	2,707
10	2,978	2,845	2,774	2,700	2,661	2,637	2,621	2,580	2,538
11	2,854	2,719	2,646	2,570	2,531	2,507	2,490	2,448	2,404
12	2,753	2,617	2,544	2,466	2,426	2,401	2,384	2,341	2,296
13	2,671	2,533	2,459	2,380	2,339	2,314	2,297	2,252	2,206
14	2,602	2,463	2,388	2,308	2,266	2,241	2,223	2,178	2,131
15	2,544	2,403	2,328	2,247	2,204	2,178	2,160	2,114	2,066
16	2,494	2,352	2,276	2,194	2,151	2,124	2,106	2,059	2,010
17	2,450	2,308	2,230	2,148	2,104	2,077	2,058	2,011	1,960
18	2,412	2,269	2,191	2,107	2,063	2,035	2,017	1,968	1,917
19	2,378	2,234	2,155	2,071	2,026	1,999	1,980	1,930	1,878
20	2,348	2,203	2,124	2,039	1,994	1,966	1,946	1,896	1,843
25	2,236	2,089	2,007	1,919	1,872	1,842	1,822	1,768	1,711
30	2,165	2,015	1,932	1,841	1,792	1,761	1,740	1,683	1,622
40	2,077	1,924	1,839	1,744	1,693	1,660	1,637	1,577	1,509
50	2,026	1,871	1,784	1,687	1,634	1,599	1,576	1,511	1,438
60	1,993	1,836	1,748	1,649	1,594	1,559	1,534	1,467	1,389
120	1,910	1,750	1,659	1,554	1,495	1,457	1,429	1,352	1,254
∞	1,831	1,666	1,571	1,459	1,394	1,350	1,318	1,221	1,000

Tavole per l'ordine quantilico α= 0,975

	1	2	3	4	5	6	7	8	9
1	647,789	799,500	864,163	899,583	921,848	937,111	948,217	956,656	963,285
2	38,506	39,000	39,165	39,248	39,298	39,331	39,355	39,373	39,387
3	17,443	16,044	15,439	15,101	14,885	14,735	14,624	14,540	14,473
4	12,218	10,649	9,979	9,605	9,364	9,197	9,074	8,980	8,905
5	10,007	8,434	7,764	7,388	7,146	6,978	6,853	6,757	6,681
6	8,813	7,260	6,599	6,227	5,988	5,820	5,695	5,600	5,523
7	8,073	6,542	5,890	5,523	5,285	5,119	4,995	4,899	4,823
8	7,571	6,059	5,416	5,053	4,817	4,652	4,529	4,433	4,357
9	7,209	5,715	5,078	4,718	4,484	4,320	4,197	4,102	4,026
10	6,937	5,456	4,826	4,468	4,236	4,072	3,950	3,855	3,779
11	6,724	5,256	4,630	4,275	4,044	3,881	3,759	3,664	3,588
12	6,554	5,096	4,474	4,121	3,891	3,728	3,607	3,512	3,436
13	6,414	4,965	4,347	3,996	3,767	3,604	3,483	3,388	3,312
14	6,298	4,857	4,242	3,892	3,663	3,501	3,380	3,285	3,209
15	6,200	4,765	4,153	3,804	3,576	3,415	3,293	3,199	3,123
16	6,115	4,687	4,077	3,729	3,502	3,341	3,219	3,125	3,049
17	6,042	4,619	4,011	3,665	3,438	3,277	3,156	3,061	2,985
18	5,978	4,560	3,954	3,608	3,382	3,221	3,100	3,005	2,929
19	5,922	4,508	3,903	3,559	3,333	3,172	3,051	2,956	2,880
20	5,871	4,461	3,859	3,515	3,289	3,128	3,007	2,913	2,837
25	5,686	4,291	3,694	3,353	3,129	2,969	2,848	2,753	2,677
30	5,568	4,182	3,589	3,250	3,026	2,867	2,746	2,651	2,575
40	5,424	4,051	3,463	3,126	2,904	2,744	2,624	2,529	2,452
50	5,340	3,975	3,390	3,054	2,833	2,674	2,553	2,458	2,381
60	5,286	3,925	3,343	3,008	2,786	2,627	2,507	2,412	2,334
120	5,152	3,805	3,227	2,894	2,674	2,515	2,395	2,299	2,222
∞	5,024	3,689	3,116	2,786	2,567	2,408	2,288	2,192	2,114

Tavole per l'ordine quantilico α= 0,975

	10	15	20	30	40	50	60	120	∞
1	968,627	984,867	993,103	1001,414	1005,598	1008,117	1009,800	1014,020	1018,258
2	39,398	39,431	39,448	39,465	39,473	39,478	39,481	39,490	39,498
3	14,419	14,253	14,167	14,081	14,037	14,010	13,992	13,947	13,902
4	8,844	8,657	8,560	8,461	8,411	8,381	8,360	8,309	8,257
5	6,619	6,428	6,329	6,227	6,175	6,144	6,123	6,069	6,015
6	5,461	5,269	5,168	5,065	5,012	4,980	4,959	4,904	4,849
7	4,761	4,568	4,467	4,362	4,309	4,276	4,254	4,199	4,142
8	4,295	4,101	3,999	3,894	3,840	3,807	3,784	3,728	3,670
9	3,964	3,769	3,667	3,560	3,505	3,472	3,449	3,392	3,333
10	3,717	3,522	3,419	3,311	3,255	3,221	3,198	3,140	3,080
11	3,526	3,330	3,226	3,118	3,061	3,027	3,004	2,944	2,883
12	3,374	3,177	3,073	2,963	2,906	2,871	2,848	2,787	2,725
13	3,250	3,053	2,948	2,837	2,780	2,744	2,720	2,659	2,595
14	3,147	2,949	2,844	2,732	2,674	2,638	2,614	2,552	2,487
15	3,060	2,862	2,756	2,644	2,585	2,549	2,524	2,461	2,395
16	2,986	2,788	2,681	2,568	2,509	2,472	2,447	2,383	2,316
17	2,922	2,723	2,616	2,502	2,442	2,405	2,380	2,315	2,247
18	2,866	2,667	2,559	2,445	2,384	2,347	2,321	2,256	2,187
19	2,817	2,617	2,509	2,394	2,333	2,295	2,270	2,203	2,133
20	2,774	2,573	2,464	2,349	2,287	2,249	2,223	2,156	2,085
25	2,613	2,411	2,300	2,182	2,118	2,079	2,052	1,981	1,906
30	2,511	2,307	2,195	2,074	2,009	1,968	1,940	1,866	1,787
40	2,388	2,182	2,068	1,943	1,875	1,832	1,803	1,724	1,637
50	2,317	2,109	1,993	1,866	1,796	1,752	1,721	1,639	1,545
60	2,270	2,061	1,944	1,815	1,744	1,699	1,667	1,581	1,482
120	2,157	1,945	1,825	1,690	1,614	1,565	1,530	1,433	1,310
∞	2,048	1,833	1,708	1,566	1,484	1,428	1,388	1,268	1,000

Tavole per l'ordine quantilico α= 0,99

	1	2	3	4	5	6	7	8	9
1	4052,181	4999,500	5403,352	5624,583	5763,650	5858,986	5928,356	5981,070	6022,473
2	98,503	99,000	99,166	99,249	99,299	99,333	99,356	99,374	99,388
3	34,116	30,817	29,457	28,710	28,237	27,911	27,672	27,489	27,345
4	21,198	18,000	16,694	15,977	15,522	15,207	14,976	14,799	14,659
5	16,258	13,274	12,060	11,392	10,967	10,672	10,456	10,289	10,158
6	13,745	10,925	9,780	9,148	8,746	8,466	8,260	8,102	7,976
7	12,246	9,547	8,451	7,847	7,460	7,191	6,993	6,840	6,719
8	11,259	8,649	7,591	7,006	6,632	6,371	6,178	6,029	5,911
9	10,561	8,022	6,992	6,422	6,057	5,802	5,613	5,467	5,351
10	10,044	7,559	6,552	5,994	5,636	5,386	5,200	5,057	4,942
11	9,646	7,206	6,217	5,668	5,316	5,069	4,886	4,744	4,632
12	9,330	6,927	5,953	5,412	5,064	4,821	4,640	4,499	4,388
13	9,074	6,701	5,739	5,205	4,862	4,620	4,441	4,302	4,191
14	8,862	6,515	5,564	5,035	4,695	4,456	4,278	4,140	4,030
15	8,683	6,359	5,417	4,893	4,556	4,318	4,142	4,004	3,895
16	8,531	6,226	5,292	4,773	4,437	4,202	4,026	3,890	3,780
17	8,400	6,112	5,185	4,669	4,336	4,102	3,927	3,791	3,682
18	8,285	6,013	5,092	4,579	4,248	4,015	3,841	3,705	3,597
19	8,185	5,926	5,010	4,500	4,171	3,939	3,765	3,631	3,523
20	8,096	5,849	4,938	4,431	4,103	3,871	3,699	3,564	3,457
25	7,770	5,568	4,675	4,177	3,855	3,627	3,457	3,324	3,217
30	7,562	5,390	4,510	4,018	3,699	3,473	3,304	3,173	3,067
40	7,314	5,179	4,313	3,828	3,514	3,291	3,124	2,993	2,888
50	7,171	5,057	4,199	3,720	3,408	3,186	3,020	2,890	2,785
60	7,077	4,977	4,126	3,649	3,339	3,119	2,953	2,823	2,718
120	6,851	4,787	3,949	3,480	3,174	2,956	2,792	2,663	2,559
∞	6,635	4,605	3,782	3,319	3,017	2,802	2,639	2,511	2,407

Tavole per l'ordine quantilico α= 0,99

	10	15	20	30	40	50	60	120	∞
1	6055,847	6157,285	6208,730	6260,649	6286,782	6302,517	6313,030	6339,391	6365,864
2	99,399	99,433	99,449	99,466	99,474	99,479	99,482	99,491	99,499
3	27,229	26,872	26,690	26,505	26,411	26,354	26,316	26,221	26,125
4	14,546	14,198	14,020	13,838	13,745	13,690	13,652	13,558	13,463
5	10,051	9,722	9,553	9,379	9,291	9,238	9,202	9,112	9,020
6	7,874	7,559	7,396	7,229	7,143	7,091	7,057	6,969	6,880
7	6,620	6,314	6,155	5,992	5,908	5,858	5,824	5,737	5,650
8	5,814	5,515	5,359	5,198	5,116	5,065	5,032	4,946	4,859
9	5,257	4,962	4,808	4,649	4,567	4,517	4,483	4,398	4,311
10	4,849	4,558	4,405	4,247	4,165	4,115	4,082	3,996	3,909
11	4,539	4,251	4,099	3,941	3,860	3,810	3,776	3,690	3,602
12	4,296	4,010	3,858	3,701	3,619	3,569	3,535	3,449	3,361
13	4,100	3,815	3,665	3,507	3,425	3,375	3,341	3,255	3,165
14	3,939	3,656	3,505	3,348	3,266	3,215	3,181	3,094	3,004
15	3,805	3,522	3,372	3,214	3,132	3,081	3,047	2,959	2,868
16	3,691	3,409	3,259	3,101	3,018	2,967	2,933	2,845	2,753
17	3,593	3,312	3,162	3,003	2,920	2,869	2,835	2,746	2,653
18	3,508	3,227	3,077	2,919	2,835	2,784	2,749	2,660	2,566
19	3,434	3,153	3,003	2,844	2,761	2,709	2,674	2,584	2,489
20	3,368	3,088	2,938	2,778	2,695	2,643	2,608	2,517	2,421
25	3,129	2,850	2,699	2,538	2,453	2,400	2,364	2,270	2,169
30	2,979	2,700	2,549	2,386	2,299	2,245	2,208	2,111	2,006
40	2,801	2,522	2,369	2,203	2,114	2,058	2,019	1,917	1,805
50	2,698	2,419	2,265	2,098	2,007	1,949	1,909	1,803	1,683
60	2,632	2,352	2,198	2,028	1,936	1,877	1,836	1,726	1,601
120	2,472	2,192	2,035	1,860	1,763	1,700	1,656	1,533	1,381
∞	2,321	2,039	1,878	1,696	1,592	1,523	1,473	1,325	1,000

Appendice D

Storia

Versione 0.1-1 14 Settembre 2004, Copyright © 2004 Claudio Agostinelli
Prima stesura del materiale in beamer e Sweave.
Versione 0.1-2 10 Febbraio 2016, Copyright © 2016 Claudio Agostinelli

Appendice E

Versione software

Questo materiale è stato preparato con LaTeX, classe `beamer` e `book` e pacchetto `Sweave` in R . Sono stati generati in R ver. 2.9.2 nel sistema operativo (OS) linux-gnu e usando i seguenti pacchetti oltre a quelli standard:

Pacchetto	Versione
`e1071`	1.5-19
`labstatR`	1.0.5
`sn`	0.4-12
`xtable`	1.5-6

Appendice F

GNU Free Documentation License

Preamble

The purpose of this License is to make a manual, textbook, or other functional and useful document "free" in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or noncommercially. Secondarily, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of "copyleft", which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License in order to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

1. APPLICABILITY AND DEFINITIONS

This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a world-wide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The **"Document"**, below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as **"you"**. You accept the license if you copy, modify or distribute the work in a way requiring permission under copyright law.

A **"Modified Version"** of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

A **"Secondary Section"** is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document's overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.

The **"Invariant Sections"** are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License. If a section does not fit the above definition of Secondary then it is not allowed to be designated as Invariant. The Document may contain zero Invariant Sections. If the Document does not identify any Invariant Sections then there are none.

The **"Cover Texts"** are certain short passages of text that are listed, as Front-Cover Texts or Back-Cover Texts, in the notice that says that the Document is released under this License. A Front-Cover Text may be at most 5 words, and a Back-Cover Text may be at most 25 words.

A **"Transparent"** copy of the Document means a machine-readable copy, represented in a format whose specification is available to the general public, that is suitable for revising the document straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input to text formatters or for automatic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent file format whose markup, or absence of markup, has been arranged to thwart or discourage subsequent modification by readers is not Transparent. An image format is not Transparent if used for any substantial amount of text. A copy that is not "Transparent" is called **"Opaque"**.

Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, LaTeX input format, SGML or XML using a publicly available DTD, and standard-conforming simple HTML, PostScript or PDF designed for human modification. Examples of transparent image formats include PNG, XCF and JPG. Opaque formats include proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools are not generally available, and the machine-generated HTML, PostScript or PDF produced by some word processors for output purposes only.

The **"Title Page"** means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material this License requires to appear in the title page. For works in formats which do not have any title page as such, "Title Page" means the text near the most prominent appearance of the work's title, preceding the beginning of the body of the text.

A section **"Entitled XYZ"** means a named subunit of the Document whose title either is precisely XYZ or contains XYZ in parentheses following text that translates XYZ in another language. (Here XYZ stands for a specific section name mentioned below, such as **"Acknowledgements"**, **"Dedications"**, **"Endorsements"**, or **"History"**.) To **"Preserve the Title"** of such a section when you modify the Document means that it remains a section "Entitled XYZ" according to this definition.

The Document may include Warranty Disclaimers next to the notice which states that this License applies to the Document. These Warranty Disclaimers are considered to be included by reference in this License, but only as regards disclaiming warranties: any other implication that these Warranty Disclaimers may have is void and has no effect on the meaning of this License.

2. VERBATIM COPYING

You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

3. COPYING IN QUANTITY

If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document's license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

4. MODIFICATIONS

You may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

A. Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if there were any, be listed in the History section of the Document). You may use the same title as a previous version if the original publisher of that version gives permission.

B. List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the Document (all of its principal authors, if it has fewer than five), unless they release you from this requirement.

C. State on the Title page the name of the publisher of the Modified Version, as the publisher.

D. Preserve all the copyright notices of the Document.

E. Add an appropriate copyright notice for your modifications adjacent to the other copyright notices.

F. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below.

G. Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice.

H. Include an unaltered copy of this License.

I. Preserve the section Entitled "History", Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled "History" in the Document, create one stating the title, year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence.

J. Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the "History" section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission.

K. For any section Entitled "Acknowledgements" or "Dedications", Preserve the Title of the section, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein.

L. Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles.

M. Delete any section Entitled "Endorsements". Such a section may not be included in the Modified Version.

N. Do not retitle any existing section to be Entitled "Endorsements" or to conflict in title with any Invariant Section.

O. Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all of these sections as invariant. To do this, add their titles to the list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section Entitled "Endorsements", provided it contains nothing but endorsements of your Modified Version by various parties–for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or by arrangement made by the same entity you are acting on behalf of, you may not add another; but you may replace the old one, on explicit permission from the previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.

5. COMBINING DOCUMENTS

You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number. Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections Entitled "History" in the various original documents, forming one section Entitled "History"; likewise combine any sections Entitled "Acknowledgements", and any sections Entitled "Dedications". You must delete all sections Entitled "Endorsements".

6. COLLECTIONS OF DOCUMENTS

You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License, provided you insert a copy of this License into the extracted document, and follow this License in all other respects regarding verbatim copying of that document.

7. AGGREGATION WITH INDEPENDENT WORKS

A compilation of the Document or its derivatives with other separate and independent documents or works, in or on a volume of a storage or distribution medium, is called an "aggregate" if the copyright resulting from the compilation is not used to limit the legal rights of the compilation's users beyond what the individual works permit. When the Document is included in an aggregate, this License does not apply to the other works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document, then if the Document is less than one half of the entire aggregate, the Document's Cover Texts may be placed on covers that bracket the Document within the aggregate, or the electronic equivalent of covers if the Document is in electronic form. Otherwise they must appear on printed covers that bracket the whole aggregate.

8. TRANSLATION

Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections with translations requires special permission from their copyright holders, but you may include translations of some or all Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License, and all the license notices in the Document, and any Warranty Disclaimers, provided that you also include the original English version of this License and the original versions of those notices and disclaimers. In case of a disagreement between the translation and the original version of this License or a notice or disclaimer, the original version will prevail.

If a section in the Document is Entitled "Acknowledgements", "Dedications", or "History", the requirement (section 4) to Preserve its Title (section 1) will typically require changing the actual title.

9. TERMINATION

You may not copy, modify, sublicense, or distribute the Document except as expressly provided for under this License. Any other attempt to copy, modify, sublicense or distribute the Document is void, and will automatically terminate your rights under this License. However, parties who have received copies, or rights, from you under this License will not have their licenses terminated so long as such parties remain in full compliance.

10. FUTURE REVISIONS OF THIS LICENSE

The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See http://www.gnu.org/copyleft/.

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License "or any later version" applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation.

ADDENDUM: How to use this License for your documents

To use this License in a document you have written, include a copy of the License in the document and put the following copyright and license notices just after the title page:

> Copyright ©YEAR YOUR NAME. Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.2 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license is included in the section entitled "GNU Free Documentation License".

If you have Invariant Sections, Front-Cover Texts and Back-Cover Texts, replace the "with...Texts." line with this:

> with the Invariant Sections being LIST THEIR TITLES, with the Front-Cover Texts being LIST, and with the Back-Cover Texts being LIST.

If you have Invariant Sections without Cover Texts, or some other combination of the three, merge those two alternatives to suit the situation.

If your document contains nontrivial examples of program code, we recommend releasing these examples in parallel under your choice of free software license, such as the GNU General Public License, to permit their use in free software.