Università di Udine, Facoltà di Scienze della Formazione
Corso di Laurea in Scienze e Tecnologie Multimediali

Corso di Matematica e Statistica

(Giorgio T. Bagni)

Appunti: elementi di Probabilità

1. LA PROBABILITÀ

1.1. Definizione di probabilità

La *probabilità* di un evento esprime una valutazione numerica del grado di fiducia che viene attribuito all'avverarsi dell'evento stesso: eventi attesi con maggiore fiducia saranno contrassegnati da una maggiore probabilità, mentre eventi considerati più rari saranno contrassegnati da una probabilità minore.

Iniziamo col riportare la seguente definizione classica di *probabilità*.

Definizione 1. Definizione di Laplace. La *probabilità* $P(\mathrm{A})$ di un evento A è il rapporto tra il numero dei casi favorevoli al manifestarsi di A ed il numero di casi possibili, giudicati tutti egualmente possibili.

Esempio 1. Consideriamo lo spazio probabilistico $\Omega = \{\alpha_1; \alpha_2; \alpha_3; \alpha_4; \alpha_5; \alpha_6\}$ dell'esperimento del lancio dei dadi. *Se giudichiamo tutti "egualmente possibili" i sei eventi*, possiamo ipotizzare che ogni sei lanci ogni evento si verifichi una ed una sola volta; risulta dunque:

$$P(\{\alpha_i\}) = \frac{1}{6}$$

per ogni i intero tale che $1 \le i \le 6$.

Osservazione. La precedente definizione, pur essendo spesso utilizzata per la sua semplicità, non risolve il problema di una definizione di probabilità. Che significa, infatti, affermare che i casi considerati devono risultare *tutti egual-*

mente possibili? Saremmo tentati di rispondere che i casi considerati devono avere la stessa *probabilità* di manifestarsi, ma così facendo finiremmo per riferire la definizione di un concetto al concetto stesso. Inoltre, non tutte le situazioni riferibili alla probabilità possono essere caratterizzate da un numero finito di eventi *tutti egualmente possibili*.

1.2. Definizioni frequentista e soggettiva della probabilità

Una diversa definizione di probabilità può essere collegata al concetto di *frequenza*. Ricordiamo che si dice *frequenza relativa* $p \in \mathbf{Q}$ di un determinato accadimento il rapporto tra il numero dei casi in cui si riscontra l'accadimento in questione (frequenza *assoluta* di tale accadimento) e il totale dei casi osservati.

Consideriamo ora un esperimento "ripetibile", ovvero *effettuabile numerose volte nelle stesse condizioni* (ad esempio: il lancio di un dado, oppure il lancio di una moneta): la probabilità da associare all'esito A di tale esperimento può essere riferita alla frequenza relativa con cui A si manifesta.

Una simile impostazione, però, porta alla necessità di considerare un *elevato* numero di prove: un insufficiente numero di prove potrebbe causare notevoli difficoltà e risultati assai diversi da quelli ottenuti applicando la precedente definizione di probabilità. Possiamo giungere così alla definizione seguente.

Definizione 2. Definizione di von Mises. La *probabilità* P(A) dell'esito A di un esperimento ripetibile è il limite a cui tende la sua frequenza relativa, al tendere all'infinito del numero delle prove effettuate.

Osservazione. Alla definizione specificata, detta *definizione frequentista della probabilità*, sono associate alcune difficoltà. Innanzitutto, è necessario supporre l'esistenza del limite; inoltre, il carattere di *ripetibilità* è un'astrazione: è *sempre* possibile ripetere un esperimento più volte, nelle *stesse* condizioni? La probabilità verrebbe definita solo *dopo* la pratica esecuzione della serie di esperimenti; ma un esperimento *ripetuto n volte* può essere realmente considerato *lo stesso esperimento* effettuato n volte, o dovrebbe essere considerato alla stregua di n esperimenti (analoghi) effettuati successivamente?

Accenniamo infine alla *definizione soggettiva* di probabilità.

Definizione 3. Definizione di de Finetti. La *probabilità* P(A) dell'evento A, secondo l'opinione di un individuo coerente, è il prezzo che tale individuo ritiene equo attribuire a un importo unitario esigibile al verificarsi dell'evento A.

Esempio 2. Un individuo coerente stimerà di scommettere $\frac{1}{6}$ di lira sulla vincita di un euro attribuita all'uscita del numero 1 nel lancio di un dado. In base a ciò, diremo che la probabilità dell'evento:

Evento A: *il risultato del lancio del dado è 1*

è: $P(\mathrm{A}) = \frac{1}{6}$.

Osservazione. Il punto di vista della definizione ora presentata è soggettivo, e ciò sembra limitare quel contenuto di oggettività che caratterizza le definizioni precedenti (e tutte le definizioni matematiche). Il lettore noti tuttavia che una valutazione *soggettiva* non deve essere considerata necessariamente *arbitraria*: se l'individuo è *coerente*, la sua valutazione coinciderà, in linea di principio, con quella di *altri* individui coerenti.

1.3. Gli assiomi di Kolmogorov della probabilità

Definizione 4. Dato l'insieme Ω, sia $\wp(\Omega)$ l'insieme delle parti di Ω; si dice *algebra di eventi* su Ω un sottoinsieme E di $\wp(\Omega)$ tale che:

- $\Omega \in \mathrm{E}$;
- $(\mathrm{A} \in \mathrm{E} \wedge \mathrm{B} \in \mathrm{E}) \Rightarrow (\mathrm{A} \cup \mathrm{B} \in \mathrm{E} \wedge \mathrm{A} \cap \mathrm{B} \in \mathrm{E} \wedge \overline{\mathrm{A}} \in \mathrm{E} \wedge \overline{\mathrm{B}} \in \mathrm{E})$.

Esempio 3. Dato l'insieme Ω, un'algebra di eventi E su Ω è l'insieme delle parti di Ω, $\wp(\Omega)$:

$$\mathrm{E} = \wp(\Omega) = \{\mathrm{A}: \mathrm{A} \subseteq \Omega\}$$

Proposizione 1. Dato l'insieme Ω, sia E un'algebra di eventi su Ω. Allora:

$$\varnothing \in \mathrm{E}$$

Dimostrazione. Ricordiamo che: $\varnothing = \overline{\Omega}$. Per la definizione:

$$\Omega \in \mathrm{E} \wedge (\mathrm{A} \in \mathrm{E} \Rightarrow \overline{\mathrm{A}} \in \mathrm{E})$$

Dunque: $\varnothing \in E \;\Rightarrow\; \varnothing = \overline{\Omega} \in E$ ■

Introduciamo ora l'impostazione assiomatica della probabilità facendo riferimento, in una prima fase, a spazi probabilistici *finiti*.

Definizione 5. Definizione assiomatica di Kolmogorov per spazi probabilistici finiti. Sia Ω un insieme finito e sia $E \subseteq \wp(\Omega)$ un'algebra di eventi su Ω. Considerata una funzione P: $E \rightarrow [0; 1]$ che ad ogni $A \in E$ associa il reale $P(A)$, esso si dice *probabilità* dell'evento A quando:

Assioma 1. $\forall A \in E,\ P(A) \geq 0$
Assioma 2. $P(\Omega) = 1$
Assioma 3. $(A_1 \in E \wedge A_2 \in E \wedge A_1 \cap A_2 = \varnothing) \Rightarrow [P(A_1 \cup A_2) = P(A_1) + P(A_2)]$

Osservazione. L'assioma 3 afferma che *la probabilità di due eventi incompatibili è la somma delle probabilità dei singoli eventi*. In questa forma, tale assioma è applicabile soltanto nel caso in cui Ω è un insieme *finito*.

Esempio 4. Sia $\Omega = \{\alpha_1; \alpha_2; \alpha_3; \alpha_4; \alpha_5; \alpha_6\}$ lo *spazio probabilistico* (finito) dell'esperimento del lancio di un dado, come indicato nell'esempio 3. Consideriamo l'algebra di eventi E su Ω data dall'insieme delle parti di Ω:

$$E = \wp(\Omega) = \{A: A \subseteq \Omega\}$$

Indichiamo con il simbolo #(A) la cardinalità di A (ovvero il numero di elementi di A). Definiamo in $E = \wp(\Omega)$ la funzione:

P: $E \rightarrow [0; 1]$ ponendo, per ogni $A \in E$: $P(A) = \frac{1}{6} \cdot \#(A)$

Tale funzione, in base alla definizione precedente, è una *probabilità*. Infatti:

- $A \in E,\ P(A) = \frac{1}{6} \cdot \#(A) \geq 0$
- $P(\Omega) = \frac{1}{6} \cdot \#(\Omega) = \frac{1}{6} \cdot 6 = 1$
- $(A \in E \wedge B \in E \wedge A \cap B = \varnothing) \Rightarrow$
  $\Rightarrow [P(A \cup B) = \frac{1}{6} \cdot \#(A \cup B) = \frac{1}{6} \cdot \#(A) + \frac{1}{6} \cdot \#(B) = P(A) + P(B)]$

Estendiamo ora l'impostazione assiomatica a spazi probabilistici infiniti.

Definizione 6. Definizione assiomatica di Kolmogorov per spazi probabilistici infiniti. Sia Ω un insieme infinito e sia $E \subseteq \wp(\Omega)$ un'algebra di eventi su Ω. Considerata una funzione P: $E \rightarrow [0; 1]$ che ad ogni $A \in E$ associa il reale $P(A)$, esso si dice *probabilità* dell'evento A quando:

Assioma 1. $\forall A \in E, P(A) \geq 0$
Assioma 2. $P(\Omega) = 1$
Assioma 3bis.

$$[\forall k \in \mathbf{N}, A_k \in E \wedge (\forall i \in \mathbf{N} \wedge \forall j \in \mathbf{N} \wedge i \neq j), A_i \cap A_j = \varnothing] \Rightarrow P\left(\bigcup_{k=0}^{+\infty} A_k\right) = \sum_{k=0}^{+\infty} P(A_k)$$

Osservazione. Il lettore noti che l'assioma 3*bis*, ora enunciato, è l'estensione dell'assioma 3 ad una collezione infinita (numerabile) di elementi di E. L'assioma 3*bis* afferma che *la probabilità dell'unione di infiniti (numerabili) eventi incompatibili è la somma delle probabilità dei singoli eventi*.

Definizione 7. Sia Ω uno spazio probabilistico (finito o infinito), E un'algebra di eventi su Ω e P una funzione probabilità definita in E. Si dice *spazio probabilizzato* la terna $\{\Omega; E; P\}$.

Esempio 5. Lo spazio probabilizzato descritto nell'esempio precedente (riferito al lancio di un dado) è $\{\Omega; E; P\}$ con:

$\Omega = \{\alpha_1; \alpha_2; \alpha_3; \alpha_4; \alpha_5; \alpha_6\}$ $\qquad$ $E = \wp(\Omega) = \{A: A \subseteq \Omega\}$

$P: A \rightarrow P(A) = \frac{1}{6} \cdot \#(A)$

1.4. Conseguenze degli assiomi della probabilità

Proposizione 2. Sia $\{\Omega; E; P\}$ uno spazio probabilizzato; sia $A \in E$ e sia $\overline{A}$ l'evento complementare di A. Allora:

$$P(\overline{A}) = 1 - P(A)$$

Dimostrazione. Per definizione, è: $\overline{A} \cup A = \Omega$, Essendo (assioma 2) $P(\Omega) = 1$ ed essendo (assioma 3 o assioma 3*bis*) $P(\overline{A} \cup A) = P(\overline{A}) + P(A)$, risulta:

$$P(\overline{A}) + P(A) = P(\overline{A} \cup A) = P(\Omega) = 1 \quad \Rightarrow \quad P(\overline{A}) = 1 - P(A) \quad \blacksquare$$

Esempio 6. Riferiamoci allo spazio probabilizzato descritto nell'esempio precedente (del lancio di un dado) $\{\Omega;\ E;\ P\}$ con:

$$\Omega = \{\alpha_1;\ \alpha_2;\ \alpha_3;\ \alpha_4;\ \alpha_5;\ \alpha_6\} \qquad E = \wp(\Omega) = \{A\colon A \subseteq \Omega\}$$
$$P\colon A \to P(A) = \frac{1}{6} \cdot \#(A)$$

(dove $\alpha_1, ..., \alpha_6$ sono gli esiti). Si verifica che:

$$P(\{\alpha_1\}) = \frac{1}{6}$$
$$P(\overline{\{\alpha_1\}}) = P(\{\alpha_2;\ \alpha_3;\ \alpha_4;\ \alpha_5;\ \alpha_6\}) = 1 - \frac{1}{6} = \frac{5}{6}$$

Esempio 7. Con riferimento allo spazio probabilizzato $\{\Omega;\ E;\ P\}$ esaminato nell'esempio precedente, per gli eventi:

$$A = \{\alpha_2;\ \alpha_3;\ \alpha_4\} \qquad B = \{\alpha_1;\ \alpha_6\}$$

risulta:

$$P(A) = \frac{3}{6} = \frac{1}{2} \qquad P(B) = \frac{2}{6} = \frac{1}{3}$$

La somma delle precedenti probabilità è:

$$P(A) + P(B) = \frac{1}{2} + \frac{1}{3} = \frac{5}{6} \neq 1$$

Infatti i due eventi $A \in E$ e $B \in E$ sono *incompatibili*, in quanto $A \cap B = \emptyset$, ma non sono *complementari*, in quanto: $A \cup B = \{\alpha_1;\ \alpha_2;\ \alpha_3;\ \alpha_4;\ \alpha_6\} \neq \Omega$.

Proposizione 3. Sia $\{\Omega;\ E;\ P\}$ uno spazio probabilizzato; allora:

$$P(\emptyset) = 0$$

Dimostrazione. Per la proposizione 1: $\emptyset \in E$. Essendo $\emptyset = \overline{\Omega}$ e ricordando che $P(\Omega) = 1$ (assioma 2), risulta:

$$P(\emptyset) = P(\overline{\Omega}) = 1-P(\Omega) = 1-1 = 0 \quad \blacksquare$$

Proposizione 4. Sia $\{\Omega;\ E;\ P\}$ uno spazio probabilizzato; siano $A \in E$ e $B \in E$ tali che: $A \subseteq B$. Allora:

$$P(A) \leq P(B)$$

Dimostrazione. Possiamo scrivere:

$$A \cup (B \setminus A) = B \qquad \text{essendo:} \qquad A \cap (B \setminus A) = \emptyset$$

e dunque (assioma 1 ed assioma 3):

$$P(A) \leq P(A)+P(B \setminus A) = P[A \cup (B \setminus A)] = P(B) \quad \blacksquare$$

Esempio 8. Riferiamoci allo spazio probabilizzato descritto nell'esempio precedente (del lancio di un dado) $\{\Omega;\ E;\ P\}$ con:

$$\Omega = \{\alpha_1;\ \alpha_2;\ \alpha_3;\ \alpha_4;\ \alpha_5;\ \alpha_6\} \qquad E = \wp(\Omega) = \{A: A \subseteq \Omega\}$$
$$P: A \rightarrow P(A) = \frac{1}{6} \cdot \#(A)$$

(dove $\alpha_1, ..., \alpha_6$ sono gli esiti). Per gli eventi:

$$A = \{\alpha_2;\ \alpha_3;\ \alpha_4\} \qquad B = \{\alpha_2;\ \alpha_3;\ \alpha_4;\ \alpha_5\}$$

risulta $A \subseteq B$ ed è: $\quad P(A) = \frac{3}{6} = \frac{1}{2} \leq \frac{2}{3} = \frac{4}{6} = P(B)$

Proposizione 5. Sia $\{\Omega;\ E;\ P\}$ uno spazio probabilizzato; per ogni $A \in E$, è:

$$0 \leq P(A) \leq 1$$

Dimostrazione. La prima disuguaglianza ricalca l'assioma 1. Per quanto riguarda la seconda, possiamo scrivere: $A = \Omega \cap A \subseteq \Omega$ ed infine:

$0 = P(\varnothing) \leq P(\mathrm{A}) \leq P(\Omega) = 1 \quad \blacksquare$

Dimostriamo ora un risultato sulla probabilità dell'unione di due eventi *qualsiasi* (*non* necessariamente incompatibili, come previsto nell'assioma 3).

Proposizione 6. Sia $\{\Omega; \mathrm{E}; P\}$ uno spazio probabilizzato; siano $\mathrm{A}\in\mathrm{E}$ e $\mathrm{B}\in\mathrm{E}$ due eventi *qualsiasi*. Allora:

$$P(\mathrm{A}\cup\mathrm{B}) = P(\mathrm{A})+P(\mathrm{B})-P(\mathrm{A}\cap\mathrm{B})$$

Dimostrazione. Possiamo scrivere:

$$\mathrm{A}\cup\mathrm{B} = \mathrm{A}\cup(\overline{\mathrm{A}}\cap\mathrm{B}) \qquad \text{essendo:} \qquad \mathrm{A}\cap(\overline{\mathrm{A}}\cap\mathrm{B}) = \varnothing$$

e dunque (assioma 3):

$$P(\mathrm{A}\cup\mathrm{B}) = P(\mathrm{A})+P(\overline{\mathrm{A}}\cap\mathrm{B}) \qquad (*)$$

Possiamo inoltre scrivere:

$$(\mathrm{A}\cap\mathrm{B})\cup(\overline{\mathrm{A}}\cap\mathrm{B}) = \mathrm{B} \qquad \text{essendo:} \qquad (\mathrm{A}\cap\mathrm{B})\cap(\overline{\mathrm{A}}\cap\mathrm{B}) = \varnothing$$

e dunque (assioma 3):

$$P(\mathrm{A}\cap\mathrm{B})+P(\overline{\mathrm{A}}\cap\mathrm{B}) = P(\mathrm{B}) \quad \Rightarrow \quad P(\overline{\mathrm{A}}\cap\mathrm{B}) = P(\mathrm{B})-P(\mathrm{A}\cap\mathrm{B})$$

Sostituendo quest'ultima uguaglianza in (*), risulta:

$$P(\mathrm{A}\cup\mathrm{B}) = P(\mathrm{A})+P(\mathrm{B})-P(\mathrm{A}\cap\mathrm{B}) \quad \blacksquare$$

Possiamo dunque affermare che *la probabilità dell'unione di due eventi è non maggiore della somma delle probabilità dei singoli eventi; risulta uguale a tale somma se e solo se gli eventi in questione sono incompatibili*.

Esempio 9. Riferiamoci ancora una volta allo spazio probabilizzato descritto nell'esempio precedente (del lancio di un dado) $\{\Omega; \mathrm{E}; P\}$ con:

$$\Omega = \{\alpha_1; \alpha_2; \alpha_3; \alpha_4; \alpha_5; \alpha_6\} \qquad \mathrm{E} = \wp(\Omega) = \{\mathrm{A}: \mathrm{A}\subseteq\Omega\}$$

$$P: \mathrm{A}\rightarrow P(\mathrm{A}) = \frac{1}{6}\cdot\#(\mathrm{A})$$

(dove α_1, ..., α_6 sono gli esiti). Consideriamo gli eventi:

$A = \{\alpha_2; \alpha_3; \alpha_4\}$ $\qquad$ $B = \{\alpha_1; \alpha_4\}$

$P(A) = \frac{3}{6} = \frac{1}{2}$ $\qquad$ $P(B) = \frac{2}{6} = \frac{1}{3}$

$P(A \cap B) = P(\{\alpha_4\}) = \frac{1}{6}$

Risulta:

$$P(A \cup B) = P(\{\alpha_2; \alpha_3; \alpha_4\} \cup \{\alpha_1; \alpha_4\}) = P(\{\alpha_1; \alpha_2; \alpha_3; \alpha_4\}) = \frac{4}{6} =$$
$$= \frac{2}{3} = \frac{1}{2} + \frac{1}{3} - \frac{1}{6} = P(A)+P(B)-P(A \cap B)$$

Diretto corollario della proposizione precedente è il risultato seguente.

Proposizione 7. Sia $\{\Omega; E; P\}$ uno spazio probabilizzato; siano $A \in E$ e $B \in E$ due eventi *qualsiasi*. Allora:

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

Dimostrazione. Per la proposizione precedente è:

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

e ricordando che (assioma 1): $P(A \cap B) \geq 0$, otteniamo la tesi:

$$P(A \cup B) \leq P(A)+P(B) \quad \blacksquare$$

2. PROBABILITÀ SUBORDINATA E CORRELAZIONE

2.1. La probabilità subordinata

Consideriamo uno spazio probabilizzato $\{\Omega; E; P\}$ ed un evento $A \in E$; la probabilità $P(A)$, come sopra visto, esprime la valutazione della fiducia che appare sensato accordare al verificarsi di A.

Ammettiamo ora di voler quantificare la fiducia che appare sensato accordare *non* al verificarsi dell'evento A, bensì al verificarsi di A *nell'ipotesi che si sia precedentemente verificato* un evento $B \in E$: è ciò che accade quando (a causa, ad esempio, di un incremento delle informazioni in nostro possesso) riteniamo di limitare i casi possibili da Ω (evento certo) al solo $B \subseteq \Omega$. La probabilità da attribuire ad A in questa nuova situazione *non* è più $P(A)$, ma prende il nome di *probabilità subordinata di* A *a* B.

Definizione 8. Sia $\{\Omega; E; P\}$ uno spazio probabilizzato; siano $A \in E$ e $B \in E$ due eventi. Si dice *probabilità subordinata* $P(A|B)$ dell'evento A rispetto all'evento B la probabilità che si verifichi A nell'ipotesi che si sia verificato B.

Osservazione. Nella situazione ora descritta, non appare molto interessante il caso $B = \Omega$: in tale situazione, infatti, la probabilità subordinata di A rispetto all'evento certo Ω altro non è che $P(A)$, ovvero la probabilità stessa dell'evento A. Né può essere considerato particolarmente significativo il caso $B = \varnothing$: non ha infatti molto senso valutare la probabilità che si verifichi l'evento A nell'ipotesi che si sia verificato l'evento impossibile $\varnothing$ (anzi: in molti manuali la scrittura $A|B$ *non viene definita* se non si verifica l'evento B).

Esempio 10. Consideriamo nuovamente lo spazio probabilizzato riferito al lancio di un dado $\{\Omega; E; P\}$, con:

$$\Omega = \{\alpha_1; \alpha_2; \alpha_3; \alpha_4; \alpha_5; \alpha_6\} \qquad E = \wp(\Omega) = \{A: A \subseteq \Omega\}$$
$$P: A \to P(A) = \frac{1}{6} \cdot \#(A)$$

(dove $\alpha_1, ..., \alpha_6$ sono gli esiti). Desideriamo valutare la probabilità che il risultato del lancio di un dado sia 1 nell'ipotesi che tale risultato sia un numero dispari. Dunque, considerati gli eventi:

Evento A': *il risultato del lancio del dado è 1*
Evento B: *il risultato del lancio del dado è un numero dispari*

vogliamo determinare la probabilità subordinata $P(A'|B)$ dell'evento A' rispetto all'evento B.

Procediamo in base alla definizione di Laplace: i casi in cui il risultato è un numero dispari (evento B) sono tre e sono dati da:

Evento A': *il risultato del lancio del dado è 1*
Evento A'': *il risultato del lancio del dado è 3*

Evento A''': *il risultato del lancio del dado è 5*

Pertanto, la probabilità subordinata P(A'|B) dell'evento A' rispetto a B deve essere valutata come il rapporto tra il numero dei casi favorevoli al manifestarsi di A' (un caso) ed il numero di casi possibili (tre casi); risulta dunque:

$$P(\mathrm{A'}|\mathrm{B}) = \frac{1}{3}$$

Mantenendo il riferimento alla definizione di Laplace, enunciamo il risultato seguente, fondamentale nella trattazione di casi di probabilità subordinata.

Proposizione 8. Teorema delle probabilità composte. Sia $\{\Omega; \mathrm{E}; P\}$ uno spazio probabilizzato e siano $\mathrm{A}\in\mathrm{E}$, $\mathrm{B}\in\mathrm{E}$ due eventi. Risulta:

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

formula che se $\mathrm{B} \neq \varnothing$ può scriversi:

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

Esempio 11. Con riferimento all'esempio precedente, risulta:

$$P(\mathrm{A'}\cap\mathrm{B}) = P(\mathrm{A'}) = \frac{1}{6}$$

$$P(\mathrm{B}) = P(\mathrm{A'}\cup\mathrm{A''}\cup\mathrm{A'''}) = \frac{3}{6} = \frac{1}{2}$$

ed è verificato il teorema delle probabilità composte:

$$P(\mathrm{A'}|\mathrm{B}) = \frac{P(\mathrm{A'}\cap\mathrm{B})}{P(\mathrm{B})} = \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{1}{3}$$

È interessante notare che nell'impostazione assiomatica la proposizione 8 *non è un teorema, ma è considerata un assioma*. Ed è proprio in virtù di tale assioma che viene definita (implicitamente) la probabilità subordinata P(A|B) di un evento A rispetto ad un evento B.

2.2. Indipendenza stocastica e correlazione

In base a quanto introdotto nel paragrafo precedente, le probabilità $P(\mathrm{A})$ e $P(\mathrm{A}|\mathrm{B})$ sono entrambe riferite al verificarsi dell'evento A, ma in situazioni diverse: la prima, nell'ipotesi che si verifichi l'evento Ω (ipotesi sempre verificata: Ω è l'evento certo), la seconda nell'ipotesi che si verifichi l'evento $\mathrm{B}\subseteq\Omega$.

Possiamo ora chiederci: le probabilità $P(\mathrm{A})$ e $P(\mathrm{A}|\mathrm{B})$ sono *sempre* diverse? La risposta è *no*, ed a tale proposito introduciamo la definizione seguente.

Definizione 9. Sia $\{\Omega; \mathrm{E}; P\}$ uno spazio probabilizzato; siano $\mathrm{A}\in\mathrm{E}$ e $\mathrm{B}\in\mathrm{E}$ due eventi. Si dice che A è *stocasticamente indipendente* da B se:

$$P(\mathrm{A}) = P(\mathrm{A}|\mathrm{B})$$

ovvero se il verificarsi di B lascia immutata la probabilità che si verifichi A.

Esempio 12. Lanciamo *due volte* una moneta e consideriamo lo spazio probabilizzato: $\{\Omega; \mathrm{E}; P\}$, con:

$$\Omega = \{(\mathrm{T};\mathrm{T}); (\mathrm{T};\mathrm{C}); (\mathrm{C};\mathrm{T}); (\mathrm{C};\mathrm{C})\} \qquad \mathrm{E} = \wp(\Omega) = \{\mathrm{A}: \mathrm{A}\subseteq\Omega\}$$

$$P: \mathrm{A}\rightarrow P(\mathrm{A}) = \frac{1}{4}\cdot\#(\mathrm{A})$$

dove intendiamo:

Esito (T; T): *il risultato del primo lancio è testa ed il risultato del secondo lancio è testa.*
Esito (T; C): *il risultato del primo lancio è testa ed il risultato del secondo lancio è croce.*
Esito (C; T): *il risultato del primo lancio è croce ed il risultato del secondo lancio è testa.*
Esito (C; C): *il risultato del primo lancio è croce ed il risultato del secondo lancio è croce.*

Consideriamo ora gli eventi $\mathrm{A}\in\mathrm{E}$, $\mathrm{B}\in\mathrm{E}$:

Evento A: *il risultato del secondo lancio è testa.*
Evento B: *il risultato del primo lancio è testa.*

$$\mathrm{A} = \{(\mathrm{T};\mathrm{T}); (\mathrm{C};\mathrm{T})\} \qquad \mathrm{B} = \{(\mathrm{T};\mathrm{T}); (\mathrm{T};\mathrm{C})\}$$

Le probabilità associate ad A e a B sono:

$$P(\mathrm{A}) = \frac{2}{4} = \frac{1}{2} \qquad\qquad P(\mathrm{B}) = \frac{2}{4} = \frac{1}{2}$$

Vogliamo determinare $P(\mathrm{A}|\mathrm{B})$. Risulta:

$$\mathrm{A}\cap\mathrm{B} = \{(\mathrm{T};\ \mathrm{T})\} \qquad \Rightarrow \qquad P(\mathrm{A}\cap\mathrm{B}) = \frac{1}{4}$$

e, in base alla proposizione precedente, è:

$$P(\mathrm{A}|\mathrm{B}) = \frac{P(\mathrm{A}\cap\mathrm{B})}{P(\mathrm{B})} = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2}$$

Essendo:

$$P(\mathrm{A}) = \frac{1}{2} = P(\mathrm{A}|\mathrm{B})$$

concludiamo che l'evento A è stocasticamente indipendente dall'evento B. (In altri termini: gettando due volte una moneta, la probabilità di ottenere testa al secondo lancio non è influenzata dall'avere ottenuto testa al primo).

Esempio 13. Consideriamo nuovamente lo spazio probabilizzato riferito al lancio di un dado è $\{\Omega;\ \mathrm{E};\ P\}$, con:

$$\Omega = \{\alpha_1;\ \alpha_2;\ \alpha_3;\ \alpha_4;\ \alpha_5;\ \alpha_6\} \qquad\qquad \mathrm{E} = \wp(\Omega) = \{\mathrm{A}\colon \mathrm{A}\subseteq\Omega\}$$

$$P\colon \mathrm{A}\rightarrow P(\mathrm{A}) = \frac{1}{6}\cdot\#(\mathrm{A})$$

(dove $\alpha_1, ..., \alpha_6$ sono gli esiti). Consideriamo gli eventi:

Evento A': *il risultato del lancio del dado è 1*
Evento B: *il risultato del lancio del dado è un numero dispari*

La probabilità subordinata $P(\mathrm{A}'|\mathrm{B})$ dell'evento A' rispetto all'evento B è:

$$P(\mathrm{A}'|\mathrm{B}) = \frac{P(\mathrm{A}'\cap\mathrm{B})}{P(\mathrm{B})} = \frac{1}{3}$$

ed essendo:

$$P(\mathrm{A}') = \frac{1}{6}$$

concludiamo che A’ *non* è stocasticamente indipendente dall’evento B.

Notiamo che nel caso in cui l’evento A sia stocasticamente indipendente da B, essendo $P(\mathrm{A}) = P(\mathrm{A}|\mathrm{B})$, la proposizione precedente assume la forma seguente.

Proposizione 9. Teorema delle probabilità composte in condizioni di indipendenza stocastica. Sia $\{\Omega; \mathrm{E}; P\}$ uno spazio probabilizzato e sia $\mathrm{A}\in\mathrm{E}$ un evento stocasticamente indipendente dall’evento $\mathrm{B}\in\mathrm{E}$. Risulta:

$$P(\mathrm{A}\cap\mathrm{B}) = P(\mathrm{A})\cdot P(\mathrm{B})$$

Consideriamo ora i due eventi A, B tali che A *non* sia stocasticamente indipendente da B e fissiamo la definizione seguente.

Definizione 10. Sia $\{\Omega; \mathrm{E}; P\}$ uno spazio probabilizzato e sia $\mathrm{A}\in\mathrm{E}$ un evento stocasticamente indipendente dall’evento $\mathrm{B}\in\mathrm{E}$; se:

$$P(\mathrm{A}) < P(\mathrm{A}|\mathrm{B})$$

l’evento A si dice *correlato positivamente* con l’evento B; se invece:

$$P(\mathrm{A}) > P(\mathrm{A}|\mathrm{B})$$

l’evento A si dice *correlato negativamente* con l’evento B.

Esempio 14. Consideriamo la situazione esaminata nell’esempio precedente. Lo spazio probabilizzato riferito al lancio di un dado è $\{\Omega; \mathrm{E}; P\}$, con:

$$\Omega = \{\alpha_1; \alpha_2; \alpha_3; \alpha_4; \alpha_5; \alpha_6\} \qquad \mathrm{E} = \wp(\Omega) = \{\mathrm{A}: \mathrm{A}\subseteq\Omega\}$$
$$P: \mathrm{A}\to P(\mathrm{A}) = \frac{1}{6}\cdot\#(\mathrm{A})$$

(dove $\alpha_1, ..., \alpha_6$ sono gli esiti). Dati gli eventi:

Evento A’: *il risultato del lancio del dado è 1*
Evento B: *il risultato del lancio del dado è un numero dispari*

si nota che A *non* è stocasticamente indipendente da B, in quanto:

$$P(\text{A'}) = \frac{1}{6} < \frac{1}{3} = P(\text{A'}|\text{B})$$

Possiamo affermare che l'evento A' è correlato positivamente con l'evento B.

2.3. Formula di Bayes

Un risultato classico del calcolo delle probabilità è il teorema di Bayes, che trova una sua prima semplice espressione nella proposizione seguente.

Proposizione 10. Teorema di Bayes (in forma elementare). Sia {Ω; E; *P*} uno spazio probabilizzato e siano $\text{A}\in\text{E}$, $\text{B}\in\text{E}$ due eventi; risulta:

$$P(\text{A}|\text{B})\cdot P(\text{B}) = P(\text{B}|\text{A})\cdot P(\text{A})$$

formula che se $\text{A} \neq \varnothing$ e $\text{B} \neq \varnothing$ può scriversi:

$$\frac{P(\text{A}|\text{B})}{P(\text{A})} = \frac{P(\text{B}|\text{A})}{P(\text{B})} \qquad \text{oppure:} \qquad P(\text{A}|\text{B}) = \frac{P(\text{B}|\text{A})\cdot P(\text{A})}{P(\text{B})}$$

Dimostrazione. Applicando due volte il teorema delle probabilità composte (proposizione 8), otteniamo:

$$P(\text{A}\cap\text{B}) = P(\text{A}|\text{B})\cdot P(\text{B}) \qquad P(\text{A}\cap\text{B}) = P(\text{B}|\text{A})\cdot P(\text{A})$$

$$P(\text{A}|\text{B})\cdot P(\text{B}) = P(\text{B}|\text{A})\cdot P(\text{A})$$

e se $\text{A} \neq \varnothing$ e $\text{B} \neq \varnothing$ possiamo scrivere: $\dfrac{P(\text{A}|\text{B})}{P(\text{A})} = \dfrac{P(\text{B}|\text{A})}{P(\text{B})}$ ■

Diretti corollari del teorema di Bayes sono i risultati seguenti (le dimostrazioni sono immediate).

Proposizione 11. Sia {Ω; E; *P*} uno spazio probabilizzato e siano $\text{A}\in\text{E}$, $\text{B}\in\text{E}$; se l'evento A è stocasticamente indipendente dall'evento B, allora anche B è stocasticamente indipendente da A.

Pertanto, invece di dire: *A è stocasticamente indipendente da B* potremo dire: *A e B sono stocasticamente indipendenti* (talvolta: *A e B sono indipendenti*).

Proposizione 12. Sia $\{\Omega; E; P\}$ uno spazio probabilizzato e siano $A \in E$, $B \in E$; se l'evento A è correlato positivamente con l'evento B, allora anche B è correlato positivamente con A; se A è correlato negativamente con B, allora anche B è correlato negativamente con A.

Pertanto, invece di dire: *A è correlato positivamente* (o: *negativamente*) *con B* potremo dire: *A e B sono correlati positivamente* (o: *negativamente*).

Esempio 15. Riprendiamo la situazione esaminata nell'esempio precedente, dove lo spazio probabilizzato riferito al lancio di un dado è $\{\Omega; E; P\}$, con:

$$\Omega = \{\alpha_1; \alpha_2; \alpha_3; \alpha_4; \alpha_5; \alpha_6\} \qquad E = \wp(\Omega) = \{A: A \subseteq \Omega\}$$
$$P: A \to P(A) = \frac{1}{6} \cdot \#(A)$$

(dove $\alpha_1, ..., \alpha_6$ sono gli esiti). Dati gli eventi:

Evento A': *il risultato del lancio del dado è 1*
Evento B: *il risultato del lancio del dado è un numero dispari*

affermiamo che l'evento A' è correlato positivamente con B, in quanto:

$$P(A') = \frac{1}{6} < \frac{1}{3} = P(A'|B)$$

La probabilità subordinata $P(B|A')$ di B rispetto ad A' è:

$$P(B \cap A') = P(A') = \frac{1}{6} \quad \Rightarrow \quad P(B|A') = \frac{P(B \cap A')}{P(A')} = \frac{\frac{1}{6}}{\frac{1}{6}} = 1$$

Essendo:

$$P(B) = P(\{\alpha_1; \alpha_3; \alpha_5\}) = \frac{3}{6} = \frac{1}{2} < 1 = P(B|A')$$

concludiamo che anche l'evento B è correlato positivamente con l'evento A'.

Inoltre il teorema di Bayes è verificato, essendo:

$$P(A'|B) \cdot P(B) = \frac{1}{3} \cdot \frac{1}{2} = 1 \cdot \frac{1}{6} = P(B|A') \cdot P(A')$$

Il teorema di Bayes, espresso in una forma particolarmente semplice nella proposizione 10, può essere esteso e riformulato con riferimento non ad un singolo evento B, ma ad un numero (finito) n di eventi B_1, B_2, B_3, ..., B_n; si noti che tali eventi B_i devono costituire una *partizione*, ovvero l'evento A deve potersi verificare soltanto al verificarsi di uno (ed uno solo) degli eventi B_i.

Proposizione 13. Teorema di Bayes. Sia $\{\Omega; E; P\}$ uno spazio probabilizzato e si considerino l'evento $A \in E$ e la partizione $B_1 \in E$, $B_2 \in E$, ..., $B_n \in E$; risulta:

$$P(B_h|A) = P(B_h) \cdot \frac{P(A|B_h)}{\sum_{i=1}^{n} P(B_i)P(A|B_i)}$$

per ogni valore dell'indice h tale che: $1 \le h \le n$.

Osservazione. Il risultato ora enunciato è talvolta denominato *teorema della probabilità delle cause*. Esso infatti fornisce le varie $P(B_h|A)$, probabilità subordinate di B_h (spesso dette *cause*) all'evento A, le quali risultano uguali alle probabilità $P(B_h)$, moltiplicate per le rispettive $P(A|B_h)$ (dette *verosimiglianze di* B_h *per* A) e divise per il denominatore $\sum_{i=1}^{n} P(B_i)P(A|B_i)$, che *non dipende* dalla particolare $P(B_h|A)$ cercata. Si dimostra infatti che $\sum_{i=1}^{n} P(B_i)P(A|B_i) = P(A)$.

Esempio 16. I responsabili di una fabbrica acquistano un componente, necessario per la produzione, da tre fornitori, F_1, F_2, F_3. Alcuni componenti, però, sono di qualità insufficiente e non possono utilizzati. È noto che:

- il 10% dei componenti prodotti da F_1 sono di qualità insufficiente;
- il 20% dei componenti prodotti da F_2 sono di qualità insufficiente;
- il 15% dei componenti prodotti da F_3 sono di qualità insufficiente;
- la fabbrica acquista il 45% del totale dei componenti da F_1;
- la fabbrica acquista il 25% del totale dei componenti da F_2;
- la fabbrica acquista il 30% del totale dei componenti da F_3.

Vogliamo determinare le probabilità che un componente che si rivela di qualità insufficiente sia stato acquistato rispettivamente da F_1, da F_2 o da F_3.

Prendiamo in considerazione gli eventi:

Evento A: *il componente acquistato è di qualità insufficiente.*
Evento B_1: *il componente è stato acquistato da* F_1.

Evento B_2: *il componente è stato acquistato da* F_2.
Evento B_3: *il componente è stato acquistato da* F_3.

$$P(A|B_1) = \frac{10}{100} = \frac{1}{10} \qquad P(B_1) = \frac{45}{100} = \frac{9}{20}$$
$$P(A|B_2) = \frac{20}{100} = \frac{1}{5} \qquad P(B_2) = \frac{25}{100} = \frac{1}{4}$$
$$P(A|B_3) = \frac{15}{100} = \frac{3}{20} \qquad P(B_3) = \frac{30}{100} = \frac{3}{10}$$

Le probabilità richieste che un componente acquistato che si rivela di qualità insufficiente sia stato acquistato da F_1, da F_2 o da F_3 si indicano rispettivamente con $P(B_1|A)$, $P(B_2|A)$, $P(B_3|A)$.

Applicando il teorema di Bayes possiamo calcolare tali probabilità; risulta:

$$P(B_1|A) = P(B_1) \cdot \frac{P(A|B_1)}{\sum_{i=1}^{n} P(B_i)P(A|B_i)} = \frac{9}{20} \cdot \frac{\frac{1}{10}}{\frac{9}{20}\cdot\frac{1}{10}+\frac{1}{4}\cdot\frac{1}{5}+\frac{3}{10}\cdot\frac{3}{20}} = \frac{9}{28}$$

$$P(B_2|A) = P(B_2) \cdot \frac{P(A|B_2)}{\sum_{i=1}^{n} P(B_i)P(A|B_i)} = \frac{1}{4} \cdot \frac{\frac{1}{5}}{\frac{9}{20}\cdot\frac{1}{10}+\frac{1}{4}\cdot\frac{1}{5}+\frac{3}{10}\cdot\frac{3}{20}} = \frac{5}{14}$$

$$P(B_3|A) = P(B_3) \cdot \frac{P(A|B_3)}{\sum_{i=1}^{n} P(B_i)P(A|B_i)} = \frac{3}{10} \cdot \frac{\frac{3}{20}}{\frac{9}{20}\cdot\frac{1}{10}+\frac{1}{4}\cdot\frac{1}{5}+\frac{3}{10}\cdot\frac{3}{20}} = \frac{9}{28}$$

----

3. VARIABILI CASUALI

3.1. Variabili casuali e distribuzioni di probabilità

Molti sono gli esempi di *numero aleatorio* (o *variabile casuale*): è aleatorio il risultato del lancio di un dado che il lettore si impegnerà ad effettuare domani mattina, è aleatorio il numero di battiti del mio cuore nei prossimi quaranta minuti, è aleatorio il numero di bambini che nasceranno in Canada nel mese prossimo. Possiamo immaginare già determinati tutti questi numeri; diremo che essi non sono a noi noti (per ora) per un'insufficiente informazione.

Definizione 11. Sia $\{\Omega; \mathrm{E}; P\}$ uno spazio probabilizzato; si dice *variabile casuale* (o *variabile aleatoria*, o *numero aleatorio*) una funzione X: $\Omega \rightarrow \mathbf{R}$ in modo tale che se:

$$X: \omega \rightarrow x$$

con $\omega \in \Omega$, per ogni valore x assunto da X sia definita la probabilità che proprio x sia il valore assunto dalla variabile casuale.

Osservazione. Come anticipato, il termine *variabile*, che in ossequio ad una consolidata tradizione non possiamo non adottare, è improprio: infatti una *variabile casuale* dovrebbe essere detta, ben più precisamente, *funzione casuale*.

La definizione afferma dunque che per ogni valore x assunto dalla variabile casuale X deve essere definita la probabilità che x sia il valore assunto da X.

Definizione 12. Sia $\{\Omega; \mathrm{E}; P\}$ uno spazio probabilizzato e sia definita la variabile casuale X: $\Omega \rightarrow \mathbf{R}$; si dice *distribuzione di probabilità* della variabile casuale X una relazione che ad ogni valore x assunto da X fa corrispondere la sua probabilità.

Se una variabile casuale può assumere valori soltanto in un insieme finito o numerabile di reali (ad esempio: nell'insieme $\{0; 1\}$, o nell'insieme dei reali interi), essa si dice *discreta*. In altri casi, una variabile casuale può assumere valori in tutto **R**, o in intervallo di **R** (ad esempio, in: $\{x \in \mathbf{R}: -1 \leq x \leq 5\}$), avente, come **R**, potenza del continuo; allora tale variabile casuale si dice *continua*.

Esempio 17. Consideriamo lo spazio probabilizzato riferito al lancio di un dado $\{\Omega; \mathrm{E}; P\}$ con:

$$\Omega = \{\mathrm{A}_1; \mathrm{A}_2; \mathrm{A}_3; \mathrm{A}_4; \mathrm{A}_5; \mathrm{A}_6\} \qquad \mathrm{E} = \wp(\Omega) = \{\mathrm{A}: \mathrm{A} \subseteq \Omega\}$$

$$P: \mathrm{A} \rightarrow P(\mathrm{A}) = \frac{1}{6} \cdot \#(\mathrm{A})$$

dove $\mathrm{A}_1, ..., \mathrm{A}_6$ sono gli eventi:

A_1	l'esito (il risultato del lancio) è:	1
A_2	l'esito (il risultato del lancio) è:	2
A_3	l'esito (il risultato del lancio) è:	3
A_4	l'esito (il risultato del lancio) è:	4

A_5	l'esito (il risultato del lancio) è:	5
A_6	l'esito (il risultato del lancio) è:	6

Possiamo definire molte variabili casuali X, Y, Z... (con le relative distribuzioni di probabilità) su di esso. Ad esempio:

X: "il risultato del lancio è un numero dispari" $\rightarrow x_1$
"il risultato del lancio è un numero pari" $\rightarrow x_2$

Cioè: $A_1 \rightarrow x_1 \quad \wedge \quad A_3 \rightarrow x_1 \quad \wedge \quad A_5 \rightarrow x_1$
$A_2 \rightarrow x_2 \quad \wedge \quad A_4 \rightarrow x_2 \quad \wedge \quad A_6 \rightarrow x_2$

Risulta: $P(x_1) = \frac{1}{2} \qquad P(x_2) = \frac{1}{2}$

Si noti che risulta: $P(x_1)+P(x_2) = 1$

Esempio 18. Nello spazio probabilizzato $\{\Omega;\ E;\ P\}$ dell'esempio precedente consideriamo la variabile casuale:

Y: "il risultato del lancio è il numero 1" $\rightarrow y_1$
"il risultato del lancio è un numero diverso da 1" $\rightarrow y_2$

Cioè: $A_1 \rightarrow y_1 \quad \wedge \quad A_2 \rightarrow y_2 \quad \wedge \quad A_3 \rightarrow y_2$
$A_4 \rightarrow y_2 \quad \wedge \quad A_5 \rightarrow y_2 \quad \wedge \quad A_6 \rightarrow y_2$

Risulta: $P(y_1) = \frac{1}{6} \qquad P(y_2) = \frac{5}{6}$

Si noti che risulta: $P(y_1)+P(y_2) = 1$

Esempio 19. Nello spazio probabilizzato $\{\Omega;\ E;\ P\}$ degli esempi precedenti consideriamo la variabile casuale:

Z: "il risultato del lancio è un numero pari maggiore di 3" $\rightarrow z_1$
"il risultato del lancio è un numero pari minore di 3" $\rightarrow z_2$
"il risultato del lancio è un numero dispari maggiore di 4" $\rightarrow z_3$
"il risultato del lancio è un numero dispari minore di 4" $\rightarrow z_4$

Cioè: $A_1 \rightarrow z_4 \quad \wedge \quad A_2 \rightarrow z_2 \quad \wedge \quad A_3 \rightarrow z_4$
$A_4 \rightarrow z_1 \quad \wedge \quad A_5 \rightarrow z_3 \quad \wedge \quad A_6 \rightarrow z_1$

Risulta: $P(z_1) = \frac{1}{3} \qquad P(z_2) = \frac{1}{6}$

$$P(z_3) = \frac{1}{6} \qquad\qquad P(z_4) = \frac{1}{3}$$

Si noti che risulta: $P(z_1)+P(z_2)+P(z_3)+P(z_4) = 1$

3.2. Funzione di ripartizione di una variabile casuale

Definizione 13. Sia $\{\Omega; \mathrm{E}; P\}$ uno spazio probabilizzato e sia definita la variabile casuale X: $\Omega\to\mathbf{R}$; si dice *funzione di ripartizione F* di tale variabile casuale la funzione F: $\mathbf{R}\to[0; 1]$ che ad ogni valore x assunto dalla variabile casuale associa la probabilità che X sia non maggiore di x:

$$F: x\to P(\{X \le x\})$$

Proposizione 14. Sia $\{\Omega; \mathrm{E}; P\}$ uno spazio probabilizzato, sia definita la variabile casuale X: $\Omega\to\mathbf{R}$ e sia F la *funzione di ripartizione* di X: F: $x\to P(\{X \le x\})$. Allora:

- F è limitata: $0 \le F(x) \le 1$;
- F è monotona non decrescente;
- $\lim_{x\to-\infty} F(x) = 0$;
- $\lim_{x\to+\infty} F(x) = 1$ (*proprietà di normalizzazione*).

3.3. Media e varianza di una variabile casuale discreta

Considereremo ora variabili casuali *discrete* e introdurremo i concetti di *media* e di *varianza* per una tale variabile casuale. Una variabile casuale discreta può assumere valori in un insieme finito o numerabile di reali (se i valori assunti dalla variabile casuale considerata costituiscono un insieme infinito, la validità delle definizioni seguenti richiederà alcune condizioni di convergenza).

Definizione 14. Sia $\{\Omega; \mathrm{E}; P\}$ uno spazio probabilizzato e sia definita la variabile casuale discreta X: $\Omega\to\mathbf{R}$; ai valori x_1, x_2, x_3, ... assunti da X corrispondono rispettivamente le probabilità p_1, p_2, p_3, ...; si dice *media* di X la media aritmetica ponderata di tali valori (assumendo le probabilità come pesi), ovvero il numero reale $M(X)$:

$$M(X) = \sum_{\forall i} x_i p_i = x_1p_1+x_2p_2+x_3p_3+...$$

dove la sommatoria è estesa a tutti gli indici i a cui corrisponde un valore x_i assunto dalla variabile casuale X.

Se i valori assunti dalla variabile casuale X costituiscono un insieme infinito, la media $M(X)$ è definita solo se la serie è convergente.

Definizione 15. Sia {Ω; E; P} uno spazio probabilizzato e sia definita la variabile casuale discreta X: $\Omega\to\mathbf{R}$; ai valori x_1, x_2, x_3, ... assunti da X corrispondono rispettivamente le probabilità p_1, p_2, p_3, ...; si dice *varianza* di X la media aritmetica ponderata dei quadrati delle differenze tra tali valori e $M(X)$ (assumendo le probabilità come pesi), ovvero il numero reale $\sigma^2(X)$:

$$\sigma^2(X) = \sum_{\forall i}\left[x_i - M(X)\right]^2 p_i =$$
$$= \left[x_1 - M(X)\right]^2 p_1 + \left[x_2 - M(X)\right]^2 p_2 + \left[x_3 - M(X)\right]^2 p_3 + \ldots$$

dove la sommatoria è estesa a tutti gli indici i a cui corrisponde un valore x_i assunto dalla variabile casuale X.

Se i valori assunti dalla variabile casuale X costituiscono un insieme infinito, la varianza $\sigma^2(X)$ è definita solo se la serie è convergente.

Ricordiamo che con il termine *scarto quadratico medio* $\sigma(X)$ di una variabile casuale si indica la radice quadrata della varianza:

$$\sigma(X) = \sqrt{\sum_{\forall i}\left[x_i - M(X)\right]^2 p_i} =$$
$$= \sqrt{\left[x_1 - M(X)\right]^2 p_1 + \left[x_2 - M(X)\right]^2 p_2 + \left[x_3 - M(X)\right]^2 p_3 + \ldots}$$

Non approfondiamo le condizioni di convergenza per le definizioni.

Esempio 20. Calcoliamo media, varianza e scarto quadratico medio della variabile casuale Z considerata in uno spazio probabilizzato {Ω; E; P} tale che:

$$P(z_1) = p_1 = \frac{1}{3} \qquad P(z_2) = p_2 = \frac{1}{6} \qquad P(z_3) = p_3 = \frac{1}{6} \qquad P(z_4) = p_4 = \frac{1}{3}$$

Assegniamo (ad esempio) a z_1, z_2, z_3, z_4 i valori:

$$z_1 = 20 \qquad z_2 = 40 \qquad z_3 = 60 \qquad z_4 = 80$$

Per la media, in tale caso, risulta:

$$M(Z) = \sum_{i=1}^{4} z_i p_i = z_1p_1+z_2p_2+z_3p_3+z_4p_4 = 20\cdot\frac{1}{3}+40\cdot\frac{1}{6}+60\cdot\frac{1}{6}+80\cdot\frac{1}{3} = 50$$

Per la varianza:

$$\sigma^2(Z) = \sum_{i=1}^{4}\left[z_i - M(Z)\right]^2 p_i =$$
$$= \left[z_1 - M(Z)\right]^2 p_1 + \left[z_2 - M(Z)\right]^2 p_2 + \left[z_3 - M(Z)\right]^2 p_3 + \left[z_4 - M(Z)\right]^2 p_4$$
$$= (20-50)^2\cdot\frac{1}{3}+(40-50)^2\cdot\frac{1}{6}+(60-50)^2\cdot\frac{1}{6}+(80-50)^2\cdot\frac{1}{3} = \frac{1900}{3} = 633{,}3$$

Per lo scarto quadratico medio, risulta:

$$\sigma(Z) = \sqrt{\sigma^2(Z)} = \sqrt{\frac{1900}{3}} = \frac{10}{3}\cdot\sqrt{57} = 25{,}2$$

3.4. Densità di una variabile casuale continua

Si può dimostrare la proposizione seguente.

Proposizione 15. Sia $\{\Omega; E; P\}$ uno spazio probabilizzato, sia definita la variabile casuale continua X: $\Omega\rightarrow\mathbf{R}$ e sia F la *funzione di ripartizione* di X: F: $x\rightarrow P(\{X \le x\})$. Allora F è continua in **R** ed esiste una funzione f non negativa tale che:

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

e risulta (*proprietà di normalizzazione*):

$$\int_{-\infty}^{+\infty} f(x)dx = \lim_{x\to+\infty} F(x) = 1$$

La funzione f ora introdotta è denominata *funzione densità*.

Definizione 16. Sia $\{\Omega; E; P\}$ uno spazio probabilizzato, sia definita la variabile casuale continua X: $\Omega\rightarrow\mathbf{R}$ e sia F la *funzione di ripartizione* di X: F: $x\rightarrow P(\{X \le x\})$. La funzione f non negativa tale che:

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

si dice *funzione densità*.

Esempio 21. Consideriamo la funzione di ripartizione:

$$F(x) = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$$

Ad essa corrisponde la funzione densità:

$$f(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}$$

Si verifica che:

$$\frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx = \lim_{x \to +\infty} F(x) = 1$$

3.5. Media e varianza di una variabile casuale continua

Definizione 17. Sia {Ω; E; P} uno spazio probabilizzato e sia definita la variabile casuale continua X: Ω→**R**, essendo F la funzione di ripartizione e f la funzione densità; si dice *media* di X il numero reale $M(X)$:

$$M(X) = \int_{-\infty}^{+\infty} xf(x)dx = \lim_{a \to -\infty} \lim_{b \to +\infty} \int_{a}^{b} xf(x)dx$$

se l'integrale improprio esiste finito.

Definizione 18. Sia {Ω; E; P} uno spazio probabilizzato e sia definita la variabile casuale continua X: Ω→**R**, essendo F la funzione di ripartizione e f la funzione densità; si dice *varianza* di X il numero reale $\sigma^2(X)$:

$$\sigma^2(X) = \int_{-\infty}^{+\infty} [x - M(X)]^2 f(x)dx = \lim_{a \to -\infty} \lim_{b \to +\infty} \int_{a}^{b} [x - M(X)]^2 f(x)dx$$

se l'integrale improprio esiste finito.

Esempio 22. Consideriamo la variabile casuale continua caratterizzata dalla funzione di ripartizione:

$$F(x) = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$$

e dalla funzione densità:

$$f(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}$$

(presentata nell'esempio precedente). Si prova che risulta:

$$M(X) = \int_{-\infty}^{+\infty} x f(x) dx = \int_{-\infty}^{+\infty} x \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}} dx = 0$$

$$\sigma^2(X) = \int_{-\infty}^{+\infty} [x - M(X)]^2 f(x) dx = \int_{-\infty}^{+\infty} x^2 \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}} dx = 1$$

3.6. Variabili casuali e inferenza statistica

Completiamo questa sezione con alcuni richiami di statistica. In particolare, accenneremo ai *procedimenti statistici inferenziali*.

In molte situazioni applicative, si è portati a considerare variabili casuali caratterizzate da distribuzioni di probabilità *non* conosciute *a priori*. In situazioni del genere è necessario introdurre alcune ipotesi sulle variabili casuali in questione (ad esempio sulla loro media e sulla loro varianza); per fare ciò si ricorre all'analisi dei dati a disposizione o, se possibile, si procede all'assunzione di nuove informazioni sulla variabile casuale in questione.

In questo paragrafo accenneremo a come tali dati vengono ricercati ed ottenuti, secondo quella che viene detta *inferenza statistica*; ovvero, introdurremo i criteri che portano a considerare, di una *collettività statistica* (*universo statistico*), un ben determinato *campione di individui* (di *unità statistiche*) che sia ben rappresentativo della totalità.

L'inferenza statistica è quel complesso di procedimenti che consente di dedurre informazioni sull'universo statistico sulla base delle informazioni ottenute esaminando un più ristretto campione di individui (*rilevamento statistico*). In particolare, le fasi più importanti e significative dell'inferenza statistica sono:

- il *problema della stima dei parametri ignoti*: analizzando il campione vengono attribuiti dei valori ai parametri ignoti, ad esempio alla media ed alla varianza di una variabile casuale non nota a priori;
- il *problema della verifica di ipotesi*: analizzando il campione vengono confermate oppure smentite e quindi corrette alcune ipotesi precedentemente formulate sulla variabile casuale non nota a priori.

La scelta del campione, dunque, è un momento delicato nell'inferenza statistica: spesso è la scelta del campione che determina il successo o l'insuccesso del procedimento inferenziale.

L'esempio seguente è liberamente ispirato a un esempio del manuale L. Daboni, *Calcolo delle probabilità ed elementi di statistica* (UTET, Torino 1980).

Esempio 23. Consideriamo un universo statistico costituito da n unità nel quale ciascuna unità può possedere o non possedere una caratteristica. Ad esempio, possiamo considerare un'urna contenente n palline; di esse, alcune sono bianche (assenza della caratteristica "colore") ed alcune sono colorate (presenza della caratteristica "colore"). Pur conoscendo il numero totale (finito) n di palline, ammettiamo di *non* conoscere anticipatamente la composizione dell'urna (il numero delle palline bianche e delle palline colorate in essa contenute).

Indichiamo mediante la variabile casuale X, che può assumere valori 0 o 1, la situazione delle singole unità. In particolare, diremo:

$X = 0$ riferendoci ad ogni pallina bianca
(assenza della caratteristica "colore");

$X = 1$ riferendoci ad ogni pallina colorata
(presenza della caratteristica "colore").

In questa semplice situazione, per conoscere completamente la distribuzione di X è sufficiente conoscere:

la percentuale α delle unità (palline) colorate (tali che $X = 1$);
la percentuale $1-\alpha$ delle unità (palline) bianche (tali che $X = 0$).

La percentuale α (inizialmente a noi non nota) *può* dunque essere:

$\alpha = 0$ se nell'urna si trovano 0 palline colorate e n palline bianche;
$\alpha = \frac{1}{n}$ se nell'urna si trovano 1 pallina colorata e $n-1$ palline bianche;

$\alpha = \frac{2}{n}$ se nell'urna si trovano 2 palline colorate e $n-2$ palline bianche;

$\alpha = \frac{3}{n}$ se nell'urna si trovano 3 palline colorate e $n-3$ palline bianche;

... ...

$\alpha = \frac{n-1}{n}$ se nell'urna si trovano $n-1$ palline colorate e 1 pallina bianca;

$\alpha = 1$ se nell'urna si trovano n palline colorate e 0 palline bianche.

Per acquisire informazioni sulla variabile casuale in questione, ovvero per conoscere la percentuale α di palline colorate (e dunque quella $1-\alpha$ di palline bianche), possiamo effettuare alcune osservazioni su campioni casuali, ovvero estrarre alcune palline dall'urna e stabilire, per ogni pallina estratta, il valore di X. Riporremo la pallina estratta nell'urna dopo ogni estrazione (opereremo dunque il "reimbussolamento"), al fine di mantenere costanti le condizioni in cui avviene l'osservazione.

Sia *m* il numero delle osservazioni effettuate; possiamo tabulare i risultati di tali osservazioni nella *m*-pla:

$$X_1, \quad X_2, \quad X_3, \quad ..., \quad X_m$$

in cui ciascuno dei precedenti X_i assume valore 0 (se l'estrazione ha riguardato una pallina bianca) o 1 (se l'estrazione ha riguardato una pallina colorata).

L'esame di tale *m*-pla consentirà di stimare la percentuale α cercata. È necessario stabilire opportunamente il numero *m* di prove in modo che siano assicurate soddisfacenti condizioni di affidabilità.

Il metodo di campionamento più semplice è quello del *sorteggio* (detto *campionamento casuale semplice*): da un universo statistico costituito da *n* individui sono estratti casualmente *q* individui (costituenti il campione, $0<q<n$), in modo che:

- ciascun individuo appartenente all'universo statistico abbia la stessa probabilità di essere estratto di ogni altro;
- la scelta di un individuo come appartenente al campione non influenzi le probabilità che gli altri individui hanno di poter essere scelti.

Esempio 24. Si consideri un universo statistico costituito da *n* individui e si voglia effettuare un campionamento casuale semplice.

Si può procedere nel modo seguente:

- si predispone l'elenco completo delle unità costituenti l'universo statistico e si identifica ciascuna unità con un numero da 1 a n; nel caso di un campionamento casuale semplice non devono essere fissati altri parametri di riconoscimento per le singole unità statistiche;
- si eseguono q estrazioni di un numero casuale compreso tra 1 e n (ad esempio, è possibile adottare la procedura dell'estrazione casuale dall'urna di una pallina contrassegnata da un numero, dopo avere predisposto n palline numerate da 1 a n), avendo cura di "reimbussolare" sempre il numero estratto;
- si inseriscono nel campione le $q<n$ unità statistiche corrispondenti ai numeri estratti.

Si noti che il citato "reimbussolamento" è indispensabile per ottenere un campionamento casuale semplice: omettere tale fase significherebbe infatti accettare che la scelta di un individuo come appartenente al campione possa influenzare le probabilità che gli altri (rimanenti) individui hanno di essere scelti.

3.7. Inferenza statistica: media e varianza

Consideriamo un universo statistico costituito da n unità; in esso sia definita una variabile aleatoria X (a noi non nota a priori) avente media $M(X)$ e varianza $\sigma^2(X)$. Individuiamo un campione casuale C_h da tale universo statistico estraendo a caso da essa n unità e valutiamo la media μ_h e la varianza $\sigma^2{}_h$ della variabile aleatoria considerata su tale campione.

Ripetiamo quindi successivamente *più* volte la valutazione precedente, sulla base di altri campionamenti effettuati con lo stesso metodo (nel nostro caso, mediante l'estrazione casuale del campione), che indicheremo con:

$C_1, C_2, C_3, \dots$

ottenendo: come medie, i valori: $\mu_1, \mu_2, \mu_3, \dots$
come varianze, i valori: $\sigma^2{}_1, \sigma^2{}_2, \sigma^2{}_3, \dots$

In generale, si può dimostrare che la media dei valori $\mu_1, \mu_2, \mu_3, \dots \mu_m$ calcolata su *tutti i possibili m campioni diversi* coincide con la media $M(X)$ della variabile aleatoria X (non nota):

$$\frac{\mu_1 + \mu_2 + \mu_3 + \dots \mu_m}{m} = M(X)$$

Tale fatto si esprime talvolta dicendo che la media del campione è una *stima corretta* della media della variabile casuale considerata.

Per la varianza, invece, si può dimostrare che risulta (sempre considerando tutti gli m possibili campioni):

$$\frac{\sigma^2{}_1 + \sigma^2{}_2 + \sigma^2{}_3 + \ldots \sigma^2{}_m}{m} = \frac{m-1}{m}\sigma^2(X)$$

Per ottenere una valutazione corretta della varianza $\sigma^2(X)$ sarebbe necessario moltiplicare il suo valore medio su tutti i possibili campioni per il fattore:

$$\frac{m}{m-1} > 1.$$

Si noti che al crescere del numero m dei possibili campioni il fattore $\frac{m}{m-1}$ si approssima al valore 1; in particolare si ha che:

$$\lim_{m \to +\infty} \frac{m}{m-1} = 1$$

Ciò porta ad affermare che la valutazione della varianza sopra descritta è *asintoticamente corretta*.

Talvolta particolari condizioni applicative suggeriscono di impiegare metodi di campionamento ben più sofisticati del campionamento casuale semplice, sopra esaminato. Ad esempio, nel *campionamento casuale stratificato*, l'universo statistico viene suddiviso in alcuni settori (*strati*) in base ad alcune particolari caratteristiche degli individui, caratteristiche che vengono ritenute di rilevante importanza per l'analisi statistica in corso. In ognuno di questi settori si effettuerà quindi il sorteggio di un campione casuale (parziale): l'unione dei campioni (parziali) così ottenuti verrà a costituire il campione sul quale impostare il procedimento inferenziale.

Molti altri metodi di campionamento possono essere utilmente applicati: nel *campionamento sistematico* le unità statistiche vengono individuate scegliendo un'unità ogni k unità nell'elenco degli elementi dell'universo statistico; nel campionamento *a grappoli* l'universo statistico viene diviso in un elevato numero di raggruppamenti e viene quindi scelto un individuo per rappresentare ciascun raggruppamento.

Molte sono le tecniche per valutare la correttezza di stime ottenute grazie a particolari procedimenti di campionamento, ma il loro esame dettagliato esula dagli scopi di questi appunti. Ricordiamo infine che i procedimenti di inferenza statistica sono generalmente suddivisi in *frequentisti* ed in *bayesiani*: nei metodi frequentisti si considerano esclusivamente le informazioni direttamente ricavate dai dati sperimentali; nei metodi bayesiani, le informazioni sperimentali possono essere integrate da altre informazioni utili a interpretare il fenomeno in esame, non direttamente desumibili dall'esecuzione dell'esperimento.

4. ESEMPI DI DISTRIBUZIONI DI PROBABILITÀ

4.1. Variabili casuali discrete

Daremo una sintetica presentazione delle più importanti distribuzioni di probabilità relative a variabili casuali discrete. Nelle definizioni seguenti indichiamo anche la media e la varianza (senza tuttavia riportare le dimostrazioni).

Definizione 19. Caratteristiche della *distribuzione uniforme.*

Possibili valori della X:	$0 \le x \le n \ \wedge \ x \in \mathbf{N}$
Probabilità:	$P(h) = \dfrac{1}{n+1}$
Media:	$\dfrac{n}{2}$
Varianza:	$\dfrac{n^2}{12} + \dfrac{n}{6}$

Esempio 25. Una variabile casuale con distribuzione uniforme di probabilità può essere definita nello spazio probabilizzato $\{\Omega; \mathrm{E}; P\}$ riferito al lancio di un dado, con:

$$\Omega = \{\mathrm{A}_1; \mathrm{A}_2; \mathrm{A}_3; \mathrm{A}_4; \mathrm{A}_5; \mathrm{A}_6\} \qquad \mathrm{E} = \wp(\Omega) = \{\mathrm{A}: \mathrm{A} \subseteq \Omega\}$$

$$P: \mathrm{A} \to P(\mathrm{A}) = \frac{1}{6} \cdot \#(\mathrm{A})$$

essendo A_i gli eventi ($1 \le i \le 6$):

A_i l'esito (il risultato del lancio) è: i

Sia X la variabile casuale:

$$X: \mathrm{A}_i \to x_i \qquad \text{con:} \quad x_i = i - 1$$

La distribuzione di probabilità è uniforme:

- Possibili valori della X: $0 \le x \le 5 \;\wedge\; x \in \mathbf{N}$
- Probabilità: $P(h) = \dfrac{1}{n+1} = \dfrac{1}{6}$
- Media: $\dfrac{n}{2} = \dfrac{5}{2}$
- Varianza: $\dfrac{n^2}{12} + \dfrac{n}{6} = \dfrac{25}{12} + \dfrac{5}{6} = \dfrac{35}{12}$

Definizione 20. Caratteristiche della *distribuzione binomiale* (o *di Bernoulli*).

Possibili valori della X: $0 \le x \le n \;\wedge\; x \in \mathbf{N}$

Probabilità: $P(h) = \dbinom{n}{h} p^h (1-p)^{n-h}$ con $0 < p < 1$

Media: np

Varianza: $np(1-p)$

Ricordiamo che è: $\dbinom{n}{h} = \dfrac{n!}{h!(n-h)!}$

L'esempio seguente è particolarmente importante per la comprensione delle possibilità applicative della distribuzione binomiale (o di Bernoulli).

Esempio 26. Le probabilità che un evento di probabilità $p = 0{,}4$ si verifichi un certo numero di volte in $n = 2$ prove indipendenti sono descritte da una variabile casuale X che assume valori:

$x_0 = 0 \qquad x_1 = 1 \qquad x_2 = 2$

Si può dimostrare che ad essi sono associate le probabilità:

$$P(0) = 1 \cdot (0{,}4)^0 (1-0{,}4)^{2-0} = 0{,}36$$
$$P(1) = 2 \cdot (0{,}4)^1 (1-0{,}4)^{2-1} = 0{,}48$$
$$P(2) = 1 \cdot (0{,}4)^2 (1-0{,}4)^{2-2} = 0{,}16$$

La distribuzione di probabilità della X è binomiale con $p = 0{,}4$:

- Possibili valori della X: $0 \le x \le 2 \;\wedge\; x \in \mathbf{N}$

- Probabilità: $P(h) = \binom{2}{h}(0,4)^h(1-0,4)^{2-h}$
- Media: $np = 2\cdot 0,4 = 0,8$
- Varianza: $np(1-p) = 2\cdot 0,4\cdot(1-0,4) = 0,48$

Definizione 21. Caratteristiche della *distribuzione di Poisson.*

Possibili valori della X:	$x\in\mathbf{N}$	
Probabilità:	$P(h) = \dfrac{\lambda^h}{h!}e^{-\lambda}$	con $\lambda > 0$
Media:	λ	
Varianza:	λ	

Esempio 27. I valori assunti da una variabile casuale X sono $x\in\mathbf{N}$, con probabilità:

$$P(0) = \frac{(0,2)^0}{0!}e^{-0,2} = 0,8187...$$

$$P(1) = \frac{(0,2)^1}{1!}e^{-0,2} = 0,1637...$$

$$P(2) = \frac{(0,2)^2}{2!}e^{-0,2} = 0,0163...$$

...

$$P(h) = \frac{(0,2)^h}{h!}e^{-0,2}$$

La distribuzione di probabilità è di Poisson con $\lambda = 0,2$:

- Possibili valori della X: $x\in\mathbf{N}$
- Probabilità: $P(h) = \dfrac{(0,2)^h}{h!}e^{-0,2}$
- Media: $\lambda = 0,2$
- Varianza: $\lambda = 0,2$

Definizione 22. Caratteristiche della *distribuzione geometrica.*

Possibili valori della X: $x \in \mathbf{N}$

Probabilità: $P(h) = p^h(1-p)$ con $0 < p < 1$

Media: $\dfrac{p}{1-p}$

Varianza: $\dfrac{p}{(1-p)^2}$

Esempio 28. I valori assunti da una variabile casuale X sono $x \in \mathbf{N}$, con probabilità:

$P(0) = (0,3)^0(1-0,3) = 0,7$ $P(1) = (0,3)^1(1-0,3) = 0,21$

$P(2) = (0,3)^2(1-0,3) = 0,063$ ... $P(h) = (0,3)^h(1-0,3)$

La distribuzione di probabilità è geometrica con $p = 0,3$:

- Possibili valori della X: $x \in \mathbf{N}$
- Probabilità: $P(h) = (0,3)^h(1-0,3)$
- Media: $\dfrac{p}{1-p} = \dfrac{0,3}{1-0,3} = 0,4285...$
- Varianza: $\dfrac{p}{(1-p)^2} = \dfrac{0,3}{(1-0,3)^2} = 0,6122...$

4.2. Variabili casuali continue

In questo paragrafo daremo una sintetica presentazione delle più importanti distribuzioni di probabilità relative a variabili casuali continue.

Nelle definizioni seguenti indichiamo anche la media e la varianza delle variabili casuali in questione (senza riportare le dimostrazioni).

Definizione 23. Caratteristiche della *distribuzione rettangolare.*

Possibili valori della X: $x \in \mathbf{R}$

Densità di probabilità: $f(x) = \frac{1}{a} \cdot \chi_{[0;\, a]} \Leftrightarrow \begin{cases} f(x) = \frac{1}{a} \\ 0 \le x \le a \end{cases} \cup \begin{cases} f(x) = 0 \\ x < 0 \vee x > a \end{cases}$

con $a > 0$

Media: $\frac{a}{2}$

Varianza: $\frac{a^2}{12}$

Esempio 29. I possibili valori di una variabile casuale X sono $x \in \mathbf{R}$ con densità di probabilità:

$$f(x) = \frac{1}{4} \cdot \chi_{[0;\, 4]} \Leftrightarrow \begin{cases} f(x) = \frac{1}{4} \\ 0 \le x \le 4 \end{cases} \cup \begin{cases} f(x) = 0 \\ x < 0 \vee x > 4 \end{cases}$$

La distribuzione di probabilità è rettangolare con $a = 4$:

- Possibili valori della X: $x \in \mathbf{R}$
- Densità di probabilità: $f(x) = \frac{1}{4} \cdot \chi_{[0;\, 4]} \Leftrightarrow$
  $\begin{cases} f(x) = \frac{1}{4} \\ 0 \le x \le 4 \end{cases} \cup \begin{cases} f(x) = 0 \\ x < 0 \vee x > 4 \end{cases}$
- Media: $\frac{a}{2} = \frac{4}{2} = 2$
- Varianza: $\frac{a^2}{12} = \frac{4^2}{12} = 1{,}3333...$

Definizione 24. Caratteristiche della *distribuzione normale.*

Possibili valori della X: $x \in \mathbf{R}$

Densità di probabilità: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-m)^2}{2\sigma^2}}$ con $\sigma > 0$

Media: m

Varianza: σ^2

Esempio 30. I possibili valori di una variabile casuale X sono $x \in \mathbf{R}$ con densità di probabilità:

$$f(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(x-2)^2}{2}}$$

La distribuzione di probabilità è normale:

- Possibili valori della X: $x \in \mathbf{R}$
- Densità di probabilità: $f(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(x-2)^2}{2}}$
- Media: $m = 2$
- Varianza: $\sigma^2 = 1$

Definizione 25. Caratteristiche della *distribuzione esponenziale.*

Possibili valori della X: $x \geq 0 \ \wedge \ x \in \mathbf{R}$

Densità di probabilità: $f(x) = ae^{a-x}$ con $a > 0$

Media: $\frac{1}{a}$

Varianza: $\frac{1}{a^2}$

Esempio 31. I possibili valori di una variabile casuale X sono $x \geq 0 \wedge x \in \mathbf{R}$ con densità di probabilità:

$$f(x) = 2e^{2-x}$$

La distribuzione di probabilità è esponenziale con $a = 2$:

- Possibili valori della X: $x \geq 0 \ \wedge \ x \in \mathbf{R}$
- Densità di probabilità: $f(x) = 2e^{2-x}$
- Media: $\frac{1}{a} = \frac{1}{2}$
- Varianza: $\frac{1}{a^2} = \frac{1}{2^2} = \frac{1}{4}$

Indichiamo con $\Gamma(n)$ la *funzione gamma*, definita per i numeri reali $n > 0$ da:

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

Definizione 26. Caratteristiche della *distribuzione gamma*.		
Possibili valori della X:	$x \geq 0 \quad \wedge \quad x \in \mathbf{R}$	
Densità di probabilità:	$f(x) = \dfrac{b^{a+1}}{\Gamma(a+1)} x^a e^{-bx}$	con $a > -1$, $b > 0$
Media:	$\dfrac{a+1}{b}$	
Varianza:	$\dfrac{a+1}{b^2}$	

Esempio 32. I possibili valori di una variabile casuale X sono $x \geq 0 \wedge x \in \mathbf{R}$ con densità di probabilità:

$$f(x) = \frac{b^{2+1}}{\Gamma(2+1)} x^2 e^{-3x} = \frac{27}{2} x^2 e^{-3x}$$

essendo $\Gamma(3) = 2$. La distribuzione di probabilità è gamma con $a = 2$ e $b = 3$:

- Possibili valori della X: $x \geq 0 \quad \wedge \quad x \in \mathbf{R}$
- Densità di probabilità: $f(x) = \dfrac{b^{a+1}}{\Gamma(a+1)} x^a e^{-bx} = \dfrac{27}{2} x^2 e^{-3x}$
- Media: $\dfrac{a+1}{b} = \dfrac{2+1}{3} = 1$
- Varianza: $\dfrac{a+1}{b^2} = \dfrac{2+1}{3^2} = \dfrac{1}{3}$

Definizione 27. Caratteristiche della *distribuzione beta*.		
Possibili valori della X:	$0 \leq x \leq 1 \quad \wedge \quad x \in \mathbf{R}$	
Densità di probabilità:	$f(x) = \dfrac{(a+b+1)!}{a!b!} x^a (1-x)^b$	con $a > -1$, $b > -1$
Media:	$\dfrac{a+1}{a+b+2}$	

Varianza:	$\left(\dfrac{a+1}{a+b+2}\right)\left(\dfrac{a+2}{a+b+3}\right)$

Esempio 33. I possibili valori di una variabile casuale X sono $0 \leq x \leq 1 \wedge x \in \mathbf{R}$ con densità di probabilità:

$$f(x) = \frac{(3+2+1)!}{3!2!} x^3 (1-x)^2 = 60x^3(1-x)^2$$

La distribuzione di probabilità è beta con $a = 3$, $b = 2$:

- Possibili valori della X: $0 \leq x \leq 1 \quad \wedge \quad x \in \mathbf{R}$
- Densità di probabilità: $f(x) = \dfrac{(a+b+1)!}{a!b!} x^a (1-x)^b =$
  $= 60x^3(1-x)^2$
- Media: $\dfrac{a+1}{a+b+2} = \dfrac{3+1}{3+2+2} = \dfrac{4}{7}$
- Varianza: $\left(\dfrac{a+1}{a+b+2}\right)\left(\dfrac{a+2}{a+b+3}\right) =$
  $= \left(\dfrac{3+1}{3+2+2}\right)\left(\dfrac{3+2}{3+2+3}\right) = \dfrac{4}{7} \cdot \dfrac{5}{8} = \dfrac{5}{14}$
