Introduzione alla teoria della probabilità

Teoria della probabilità

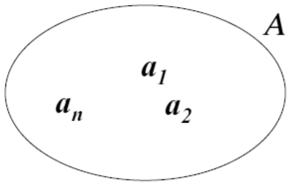
- Primi sviluppi nel XVII secolo (Pascal, Fermat, Bernoulli);
- Nasce nell'ambito dei giochi d'azzardo;
- La prima formalizzazione rigorosa si deve a Laplace (1812).

Teoria della probabilità

- Prima di cominciare a parlare di probabilità, è utile riprendere alcune nozioni sugli **insiemi**;
- La teoria della probabilità, infatti, è tutta basata sugli **insiemi**, che, nel linguaggio probabilistico, si chiameranno **eventi**.

Insieme

- Un insieme è una collezione di oggetti diversi tra loro, e generalmente non ordinati, del tipo:
- $A = \{a_1, a_2, \dots, a_n\}$

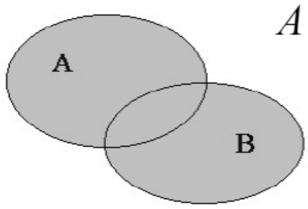


- Un insieme che non contiene elementi si chiama *insieme vuoto*:
- $\varnothing = \{ \}$

Operazioni con gli insiemi

UNIONE

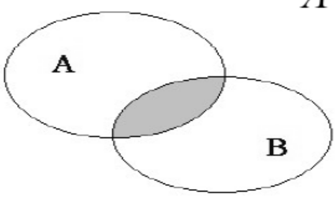
$$A \cup B$$



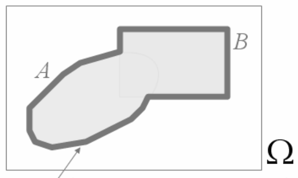
Operazioni con gli insiemi

INTERSEZIONE

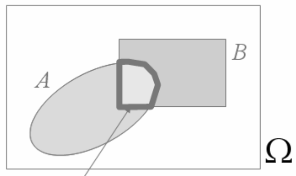
$$A \cap B$$



Unione e intersezione di insiemi



$$A \cup B = \{\omega : \omega \in A \text{ oppure } \omega \in B\}$$

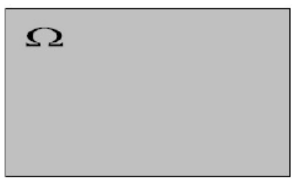


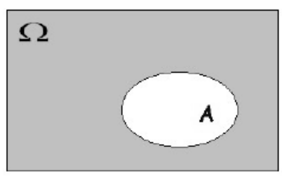
$$A \cap B = \{\omega : \omega \in A \text{ e } \omega \in B\}$$

$$A \cup \Omega = \Omega$$

$$A \cap \Omega = A$$

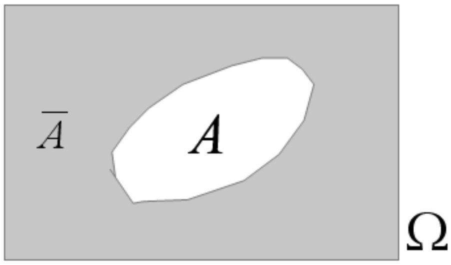
Insieme complementare di A





$$A^C = \Omega - A$$

Insieme complementare di A (detto anche “non A”)



$$\overline{A} = \{\omega : \omega \notin A\}$$

Alcune definizioni

- ***spazio campione (o spazio campionario, o spazio degli eventi)*** Ω = insieme di tutti gli esiti possibili di un esperimento aleatorio;
- ***punti campione*** = elementi di Ω;
- ***evento*** = sottoinsieme dello spazio campione Ω (semplice o complesso), in simboli $\boldsymbol{E} \subseteq \Omega$;
- uno spazio campione è ***discreto*** se contiene un numero finito di elementi o se è un insieme infinito numerabile;
- uno spazio campione è ***continuo*** se contiene una infinità non numerabile di punti.

Un <u>evento</u> può essere:

- ***certo*** quando si presenta con certezza: in questo caso l'evento coincide con lo spazio campione Ω ($\boldsymbol{E} = \Omega$);
- ***impossibile*** quando non può mai verificarsi: in questo caso coincide con l'insieme vuoto $\varnothing$ ($\boldsymbol{E} = \varnothing$);
- ***semplice*** quando non può essere scomposto ulteriormente: coincide con un punto dello spazio campionario Ω;
- ***complesso (o composto)*** quando è dato da un insieme di eventi semplici, cioè da un sottoinsieme E di Ω, $\boldsymbol{E} \subseteq \Omega$

La teoria della probabilità

La teoria della probabilità si occupa del comportamento degli esperimenti aleatori.

Che cos'è un esperimento aleatorio?

Esperimento aleatorio

Un esperimento aleatorio è un esperimento il cui esito non può essere predetto con certezza prima di essere eseguito.

Esperimento aleatorio vs esperimento deterministico

In questo, un esperimento aleatorio si differenzia rispetto agli **esperimenti deterministici**, come ad esempio gli esperimenti della fisica classica (es. caduta di un grave, ecc.), il cui esito può essere predetto con certezza praticamente assoluta.

Esperimento aleatorio vs esperimento deterministico

- Lanciare una moneta e **osservare** l'esito che produce (testa o croce), oppure

-Lasciare cadere un grave e **misurare** il tempo che impiega a raggiungere il suolo

Sono entrambi **esperimenti**.

Esperimento deterministico

Se l'esperimento consiste nella misurazione del tempo impiegato da un grave a raggiungere il suolo, se conosciamo le condizioni iniziali del sistema e risolviamo le equazioni del moto, è possibile **predire** con esattezza lo stato del sistema, in ogni suo istante.

Esperimento aleatorio

Se l'esperimento consiste nel lancio di una moneta, **<u>non è possibile stabilire con certezza</u>** se l'esito di un singolo lancio sarà testa o croce.

Lo spazio campione (o spazio campionario)

Come già accennato, l'insieme costituito da tutti gli eventi elementari che costituiscono i risultati possibili di un esperimento aleatorio si dice **spazio campione (o spazio campionario, o spazio degli eventi)** e ciascun evento semplice si dice **punto campione**.

Esempio: lancio di un dado

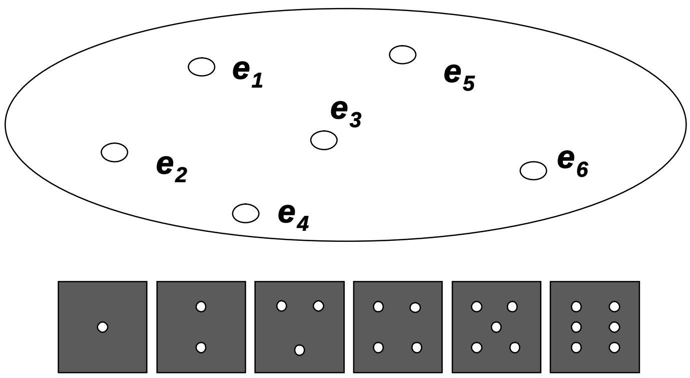
Nel caso dell'esperimento costituito dal **lancio di un dado**, lo **spazio campione** Ω è l'insieme dei punti campione corrispondenti ai sei eventi elementari, cioè alle sei facce del dado e_i, con i = 1, 2, ..., 6:

$$\Omega = \{e_1, e_2, e_3, e_4, e_5, e_6\}$$

Ciascuno degli esiti possibili che costituiscono lo spazio campione Ω si chiama **punto campione**

Esempio

$$\Omega = \{e_1, e_2, e_3, e_4, e_5, e_6\}$$



Esempio: estrarre una carta

Nel caso dell'esperimento "estrarre una carta da un mazzo", lo **spazio campione** è costituito dalle 52 carte del mazzo.

Ω = {1, 2, ..., 13} X {*Cuori*, *Quadri*, *Fiori*, *Picche*}

Spazi campione finiti e infiniti

Si possono distinguere tre tipi di spazio campione:

- spazio campione *finito*
- spazio campione *infinito numerabile*
- spazio campione *infinito non numerabile*

Spazio campione finito

Lo **spazio campione finito** è costituito da un numero finito di elementi.

Esempio: il lotto

Un'urna contiene 90 palline numerate. Per l'esperimento "estrazione di una pallina dall'urna", lo spazio campione sarà finito e sarà costituito da 90 punti campione, ciascuno dei quali corrisponde a una delle palline.

Spazio campione infinito numerabile

Lo spazio campione **infinito numerabile** è uno spazio campione infinito nel quale a ogni punto campione può essere associato un numero naturale.

Spazio campione infinito non numerabile

Si dice **infinito non numerabile** uno spazio campione i cui eventi semplici sono tali per cui, fissati due di essi, è sempre possibile determinarne almeno un terzo intermedio.
Ad esempio, l'insieme dei numeri reali R, o un suo sottoinsieme, è uno spazio campione non numerabile.

Spazi campione discreti e continui

Lo spazio campione si dice **discreto** se è uno spazio finito o infinito numerabile.

Lo spazio campione si dice **continuo** se è uno spazio infinito non numerabile.

Evento

Si dice **evento** l'insieme costituito da uno o più dei possibili risultati di un esperimento aleatorio.

Eventi

- Eventi elementari
- Eventi complessi

Eventi elementari

Gli eventi elementari sono costituiti da uno solo dei possibili risultati di un esperimento aleatorio.

Eventi complessi (o composti)

Gli **eventi complessi** sono costituiti da **più di uno** dei possibili risultati di un esperimento aleatorio.

Un evento complesso può sempre essere scomposto in eventi elementari. Se un evento non risulta ulteriormente scomponibile, allora è un evento elementare.

Eventi complessi

Nello spazio campione associato a un esperimento aleatorio, un evento composto corrisponde ad un insieme che contiene più di un punto campione.

Un evento è un **sottoinsieme** dello spazio campionario.

Esempio

Dato l’esperimento costituito dal lancio di un dado, viene definito l'evento A:
“**si osserva un numero dispari**”.

A={1,3,5}

A è un **evento complesso**.

Esempio

Nel caso dell’esperimento “lancio di un dado”, viene definito l'evento E_2:
“**si osserva la faccia 2**”.

A={2}

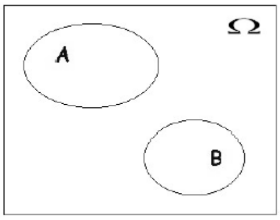
Dato che tale evento corrisponde a **un solo punto dello spazio campione**, E_2 è un **evento elementare**.

Ancora sugli eventi

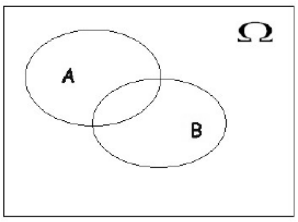
- Dati due eventi A e B, essi si dicono:
- ***Incompatibili*** o ***disgiunti*** quando non hanno alcun elemento in comune:
- $A \cap B = \varnothing$
- ***Compatibili*** quando hanno almeno un punto campione in comune:
- $A \cap B \neq \varnothing$
- ***Complementari*** quando sono disgiunti e la loro unione dà luogo allo spazio campione:
- $A \cup B = \Omega$
- In questo caso B si dice **complementare di A** e si indica con A^C

Eventi

INCOMPATIBILI



COMPATIBILI



Spazio probabilizzato

Una volta descritto lo spazio campione di cui disponiamo, il passo che dobbiamo fare adesso consiste nell'associare delle "misure", che chiameremo "probabilità", a ciascuno degli eventi del nostro spazio campione.

Spazio probabilizzato

Uno spazio probabilizzato si costruisce assegnando una numero reale *P*, chiamato probabilità, **a ciascun evento** dello spazio campione Ω.

Spazio probabilizzato

Arrivati a questo punto, dobbiamo quindi capire che cosa si intende con la nozione di “probabilità”.

Probabilità: tre definizioni

- Storicamente, la teoria della probabilità non è stata in grado di produrre una definizione univoca di probabilità, accettata da tutti.
- Le definizioni di probabilità sono tre:
- 1) la definizione classica (Pascal);
- 2) la definizione frequentista (Von Mises);
- 3) la definizione soggettivista o bayesiana (De Finetti)

Definizione classica (Pascal)

Se un evento si può verificare in **N** modi mutuamente esclusivi e **<u>tutti ugualmente probabili</u>**, se **m** di questi possiedono una caratteristica E, la probabilità di E è il **rapporto tra il numero di casi favorevoli e il numero totale dei casi possibili**

P(E) =
= casi favorevoli / casi possibili =
= m / N

Esempio: lancio di un dado.
<u>Se il dado non è truccato</u>, ognuna delle 6 facce ha la stessa probabilità di uscire: 1/6,
perché i casi possibili sono 6, e non c'è una faccia che sia "più probabile" delle altre.

Pro e contro della definizione classica

Problemi della definizione classica:

1) non sempre posso dire che gli eventi sono equiprobabili (*esempio: il dado è truccato*);
2) il numero di casi deve essere <u>finito.</u>

Aspetti positivi:

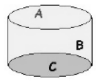
- è una definizione immediatamente operativa:
- Infatti, posso determinare la probabilità facendo uso del calcolo combinatorio.

Esempio

Sia dato un cilindro di plastica infrangibile, le cui facce vengono indicate con A, B, C. Vogliamo calcolare la probabilità che, se lo lanciamo in aria, esso cada con la faccia C rivolta verso l'alto.

Lo spazio campione sarà:

$$\Omega = \{A, B, C\}$$



Possiamo dire che

$$p(C) = \frac{1}{3} \quad \textbf{?}$$

NO: gli eventi non sono EQUIPROBABILI

Definizione frequentista (o a posteriori) – Von Mises

Si ripete un esperimento N volte e se un evento con una certa caratteristica E si verifica m volte, la frequenza relativa di successo sarà data da:

f(E) = frequenza assoluta di E / numero di ripetizioni =
= m / E

Solo apparentemente la formula è uguale a quella
Della definizione classica; in realtà, in questo caso, la frequenza relativa di E è soltanto una **stima** della "vera" probabilità P(E)

Definizione frequentista (o a posteriori)

Nella concezione frequentista della probabilità, per ***probabilità*** di un evento si intende il limite a cui tende la frequenza relativa delle prove in cui l'evento si verifica (frequenza relativa dei "casi favorevoli"), quando il numero di prove, tutte effettuate nelle stesse condizioni, tende all'infinito:

$$P(E) = \lim_{N \to \infty} f(E) = \lim_{N \to \infty} \frac{m}{N} = p$$

Problemi della definizione frequentista

- In situazioni concrete, il "passaggio al limite" su cui si basa la definizione è di difficile o impossibile attuazione: le prove sono troppe!
- Non è sempre possibile assicurare l'invarianza delle condizioni in cui si realizzano le prove.

Esempio

- Vogliamo calcolare la probabilità che un giocatore di basket faccia canestro effettuando un tiro libero.
- La frequenza relativa NON può essere assunta come probabilità se i vari tiri non sono stati realizzati NELLE STESSE CONDIZIONI !!
- E, come sappiamo, molto difficilmente un giocatore è "nelle stesse condizioni" durante tutta la partita!

Definizione soggettivista (De Finetti)

- Secondo la concezione soggettivista della probabilità, la probabilità è il "grado di fiducia", espresso soggettivamente, rispetto al verificarsi dell'evento;
- Tale grado dipende da "quanto si è disposti a scommettere" sull'evento (prezzo equo per la scommessa);
- A differenza delle altre concezioni probabilistiche, quella soggettiva **non crede che la probabilità sia una proprietà degli oggetti osservati** (esperimenti), ma crede invece che sia una proprietà insita nell'osservatore, che giudica sempre sulla base delle sue conoscenze:
- **Probabilità "ontologica"**
- **VS**
- **Probabilità "epistemica"**
- La differenza tra le due concezioni è profonda:
- **Un evento non è più o meno probabile: sono io che giudico più o meno probabile un evento.**

Problemi della definizione soggettivista

- I detrattori della concezione soggettivista ritengono, ovviamente, che la formulazione di una probabilità "soggettiva" non sia sufficientemente "obiettiva".

Come è stato risolto il "problema" delle tre definizioni?

- Le differenti concezioni della probabilità non hanno mai portato a un "accordo";
- Tuttavia, il lavoro di Andrei Nikolaevich Kolmogorov, portò nel 1933 alla formulazione assiomatica della teoria della probabilità.

La formulazione assiomatica

- Come era avvenuto per altre discipline (es. l'aritmetica, con gli assiomi di Peano), Kolmogorov decise di non occuparsi dei "fondamenti" della probabilità ("che cosa è la probabilità") ma di occuparsi invece di fondare una solida teoria che "funzionasse" comunque, a prescindere da quale fosse la concezione probabilistica.

La formulazione assiomatica

- Per fare ciò, Kolmogorov fissò TRE assiomi, cioè tre nozioni "intuitive", non dimostrabili, che fossero accettate da tutti, senza considerare minimamente questa o quella concezione probabilistica.
- Tutte e tre le concezioni probabilistiche, infatti, soddisfano i tre assiomi di Kolmogorov.

La formulazione assiomatica

- Ancora oggi, infatti, si studia una unica “teoria della probabilità”, che vale sempre, qualunque sia la concezione probabilistica;
- Le varie concezioni possono dare luogo a determinati “casi particolari” e “arricchimenti” della teoria; ma la teoria di fondo, comunque, rimane sempre la stessa.

Gli assiomi di Kolmogorov

Nel caso di uno spazio campione finito Ω, un numero reale $P(A)$, chiamato *probabilità* di A, può essere assegnato a ciascun evento A (dove A è un sottoinsieme di Ω) se i seguenti assiomi vengono rispettati:

Assioma 1

La probabilità è un numero reale maggiore o uguale a zero.

$$P(A) \geq 0$$

Assioma 2

La probabilità dell'evento certo è pari a 1.

$$P(\Omega) = 1$$

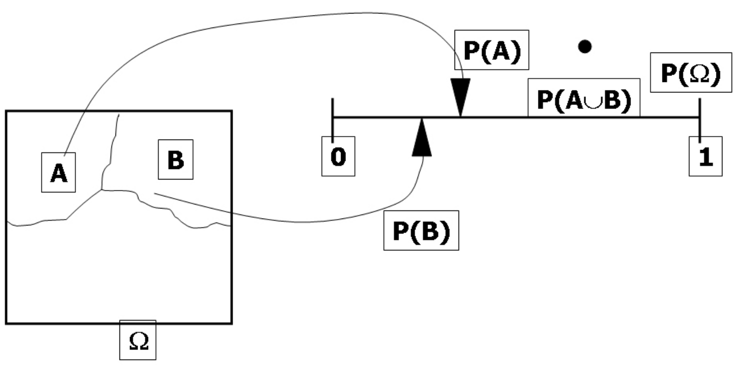
Assioma 3: additività

Se $A_1, A_2, \ldots, A_m$ sono eventi incompatibili in Ω,
cioè se: $A_i \cap A_j = \varnothing \quad \forall i \neq j$,

allora la probabilità dell'evento unione è pari alla somma selle singole probabilità.

$$P(A_1 \cup A_2 \cup \ldots \cup A_m) = \sum_{i=1}^{m} P(A_i)$$

Assioma 3: additività



Spazio probabilizzato (o spazio probabilistico)

- Torniamo adesso alla nozione di spazio probabilizzato;
- In sostanza, uno spazio probabilizzato è dato dalla **terna**:
- **(Ω, B_Ω, P)**
- Dove:
- Ω è lo spazio campione;
- B_Ω è l'insieme delle parti di Ω, cioè l'insieme di tutti i possibili sottoinsiemi di Ω.
- P è la probabilità assegnata a ogni possibile evento compreso in Ω.

Spazio probabilizzato

- Unica condizione:
- La probabilità P deve rispettare gli assiomi di Kolmogorov.

Regole probabilistiche

- Dai tre assiomi di Kolmogorov discendono alcuni importanti teoremi sulla probabilità, che sono poi diventate le regole (o leggi) probabilistiche.

Legge della somma

- Il terzo assioma di Kolmogorov ci dice già come comportarci quando gli eventi sono disgiunti (incompatibili);
- Cosa accade, invece, nel caso più generale, in cui invece gli eventi sono compatibili (cioè la loro intersezione non è vuota)?

Legge della somma

- Se $A \cap B \neq \emptyset$

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

Esempio

- Vogliamo calcolare la probabilità che il risultato del lancio di un dado sia un multiplo di 3 oppure un numero maggiore di 4. I due eventi che compongono l'evento unione E sono compatibili:

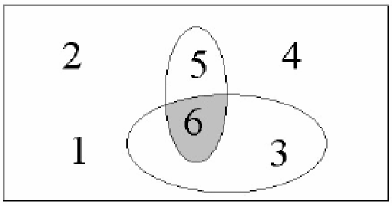
$$E_1 = \{3, 6\}$$
$$E_2 = \{5, 6\}$$

l'elemento 6 è presente in entrambi gli insiemi

$$E = E_1 \cup E_2 = \{3, 5, 6\}$$

Vogliamo calcolare la probabilità dell'evento unione

Graficamente



$$p(E) = \frac{3}{6}$$

Se consideriamo separatamente i due eventi, abbiamo che:

$$p(E_1) = \frac{2}{6} \qquad p(E_2) = \frac{2}{6}$$

e quindi

$$p(E) < p(E_1) + p(E_2)$$

Infine:

Si ha

$$p(E_1 \cup E_2) < p(E_1) + p(E_2)$$

perché i casi favorevoli nel secondo membro vengono contati due volte; per ristabilire l'equilibrio bisogna sottrarre al secondo membro la probabilità dell'intersezione.

Avremo pertanto:

$$p(E) = \frac{2}{6} + \frac{2}{6} - \frac{1}{6} = \frac{3}{6}$$

Probabilità condizionata e indipendenza stocastica

- **Esempio: un'urna contiene 15 palline rosse e 5 nere.**
- **Vogliamo calcolare la probabilità di ottenere in 2 estrazioni consecutive senza reinserimento una pallina rossa e poi una nera:**
- **A= estraggo una pallina rossa p(A)=15/20=3/4**
-
- **B= estraggo una pallina nera**
-
- **La probabilità di estrarre una nera dopo aver estratto una rossa è 5/19, perché i casi possibili adesso sono diventati 19.**
- **Cioè: la conoscenza dell'evento A ha ridotto la dimensione dello spazio campionario.**

Probabilità condizionata

Dati due eventi A e B, si dice <u>probabilità di B condizionata ad A</u> , e si scrive: P(B|A), la probabilità di B calcolata nel caso in cui si sia verificato A.

Naturalmente, si presume che l'evento A sia in grado di condizionare B, cioè che NON sia stocasticamente indipendente da A.

La probabilità condizionata

- L'osservazione "empirica" che abbiamo fatto nell'esempio appena visto ci permette di introdurre la formula della probabilità condizionata:

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

La probabilità composta

- Dalla formula della probabilità condizionata, possiamo ricavare direttamente una regola per calcolare, in generale, la probabilità di un **evento composto**.
- In generale, se vogliamo ottenere la probabilità del verificarsi composto di due eventi, si deve applicare la "regola del prodotto":

$$P(A \cap B) = P(A)P(B \mid A)$$

La probabilità composta

- Nel nostro esempio:

$P(A \cap B) = P(A) * P(B|A) = 3/4 * 5/19 = 15/76$

Un altro esempio

Abbiamo un sacchetto contenente due gettoni rossi ed uno nero. Estraiamo due gettoni in sequenza (senza reinserire il primo estratto nel sacchetto); vogliamo determinare la probabilità dell'evento

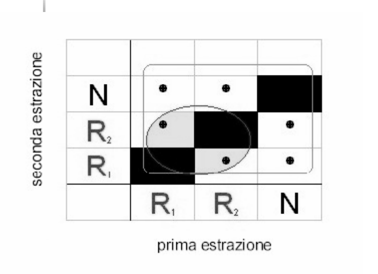
$E = \{\text{i gettoni estratti sono entrambi rossi}\}$

L'evento E risulta composto dai due eventi elementari:

$E_1 = \{R'\}$
$E_2 = \{R''\}$
$E = E_1 \cap E_2$

Si noti però che, mentre $p(E_1) = 2/3$, non così vale per $p(E_2)$ che <u>può variare a seconda del risultato della prima estrazione</u>.

La probabilità composta



Gli eventi possibili sono 6, quelli favorevoli al verificarsi di E sono due; pertanto la probabilità cercata sarà

$$p(E) = \frac{2}{6}$$

Come si può ottenere questo valore?

La probabilità di estrarre come primo un gettone rosso (E_1) è data da

$$p(E_1) = \frac{2}{3}$$

la probabilità di estrarre un rosso anche alla seconda estrazione (E_2) dipende dal risultato della prima; supponendo di avere estratto già un gettone rosso tale probabilità sarà

$$p(E_2 \mid E_1) = \frac{1}{2}$$

Possiamo quindi effettuare il calcolo con la formula:

$$P(E_2 \cap E_1) = P(E_1)P(E_2 \mid E_1) = \frac{2}{3} \times \frac{1}{2} = \frac{2}{6}$$

Indipendenza stocastica

- Dalla nozione di probabilità condizionata deriva anche la nozione di **indipendenza stocastica**;
- Se, infatti, si ha:
- P(A|B) = P(A)
- Si dice che i due eventi, A e B, sono tra loro stocasticamente indipendenti.

Regola del prodotto nel caso di indipendenza stocastica

- Se e solo se A e B sono stocasticamente indipendenti, dalla "regola del prodotto" generale, si ricava la relazione:
- **$P(A \cap B) = P(A)P(B)$**
- Che è anche una definizione alternativa di indipendenza stocastica.

Una misura della dipendenza tra A e B

- Dati due eventi, A e B, non stocasticamente indipendenti, la quantità:
- $\mathrm{Dip}(A,B) = | P(A \cap B) - P(A)P(B) |$
- Si dice **dipendenza** tra A e B.
- Ovviamente, se i due eventi sono stocasticamente indipendenti, si avrà:
- $\mathrm{Dip}(A,B) = 0$