Appunti Probabilità e Statistica

Nicolas Alberti

Anno Accademico 2020/21

Abstract

Questi sono appunti redatti seguendo le video-lezioni.
Dovrebbero contenere tutto, quindi dovrebbero essere pronti per lo studio finale.

Contents

Chapter 1

Lezione 1

1.1 Spazio di Probabilità

1.1.1 Definizione:

Uno spazio di probabilità è una terna (Ω, F, P) dove:

- Ω: è un Insieme
- F: è una famiglia di sottoinsiemi di Ω con una particolare struttura detta σ-algebra
- P: da F $\rightarrow [0, 1]$, è una mappa con misura positiva e normalizzata

Descriviamo ora nel particolare il primo simbolo. Ω è lo **spazio campionario**, ovvero è l'insieme di **tutti gli esiti** dell'esperimento aleatorio che stiamo considerando.

Oss.1 : bisogna scegliere un insieme Ω sufficientemente ricco capace di contenere tutti i risultati dell'esperimento considerato. Quindi la scelta dello spazio campionario **non è** univoca.

Oss.2 :

- Se (Ω) $\leq$ N, si dice che lo spazio campionario è **discreto**.
- Se (Ω) $\geq$ N, si dice che lo spazio campionario è **continuo**.

1.1.2 Esempi:

1. Lancio 1 moneta, osservo la faccia uscita:

$$\Omega = \{T, C\}$$

dove T e C sono Testa e Croce

2. Lancio 1 moneta per 3 volte, vado a contare il numero di Teste uscite:

$$\Omega = \{0, 1, 2, 3\}$$

3. Lancio 1 moneta per 3 volte, ora osservo la sequenza delle facce ottenute:

$$\Omega = \{TTT, TTC, TCT, CTT, TCC, CTC, CCT, CCC\}$$

4. Lancio di 1 moneta fino a che non ottengo Testa e conto il numero di lanci effettuati:

$$\Omega = \mathrm{N}+ = \{1, 2, 3, ...\}$$

5. Considero il tempo di vita di un Hard Disk:

$$\Omega = \mathrm{R}+ = [0, +\infty]$$

1.1.3 Esiti ed Eventi:

Esiti ed Eventi sono degli **elementi speciali** dei sottoinsiemi.

- se $\omega \in \Omega$ (ovvero ω è **elemento** di Ω), allora è detto **esito** (o *evento elementare*
- se E $\subseteq \Omega$ (ovvero E è **sottoinsieme** di Ω, allora è detto **evento**

Se l'esecuzione dell'esperimento dà come risultato $\omega \in \Omega$, diremo che si è verificato ω.

Per ogni E tale che $\omega \in$ E, allora diremo che si è verificato E.

1.1.4 Esempio:

Riprendiamo l'esempio 3 descritto prima. Lo spazio campionario che rappresenta i possibili risultati di 3 lanci di una moneta è :

$$\Omega = \{TTT, TTC, TCT, CTT, TCC, CTC, CCT, CCC\}$$

Gli eventi possibili saranno quindi:

- E_k = "ottengo K croci (in 3 lanci) ", con K = 0, 1, 2, 3 Per ogni K devo esaminarne gli eventi **favorevoli**:

  1. E_0 = TTT
  2. E_1 = TCT, CTT, TTC
  3. E_2 = TCC, CTC, CCT
  4. E_3 = CCC

  Supponiamo di avere ottenuto come esito, dopo i 3 lanci, CTC: si verifica quindi E_2 ma non si verificano E_0, E_1 ed E_3.

1.1.5 Eventi banali ed onnipresenti:

Sono 2 e sono quelli di seguito:

- Ω: evento **certo**
- $\emptyset$: evento **impossibile**

1.2 Operazioni elementari sugli eventi

Le operazioni insiemistiche permettono di combinare più spazi campionari, al fine di ottenere eventi più interessanti.

Siano E, F $\in \Omega$ degli eventi. Abbiamo:

- il **complementare** di E (evento):

  $$E^c = \{(\omega \in \Omega | \omega \ni) E\}$$

  E' l'insieme degli esiti che **non stanno** in E. Probabilisticamente, E^c si verifica quando non si verifica E.

- **Intersezione** di E con F:

  $$E \cap F \ = \ \{\omega \in \Omega | \omega \in E \ e \ \omega \in F\}$$

  Probabilisticamente, $E \cap F$ si verifica solo se si verificano **sia E che F**. Se E ed F sono tali che:

  $$E \cap F \ = \varnothing$$

  si dice che E ed F sono **incompatibili** o **disgiunti**, cioè E ed F non si possono verificare allo stesso momento.

- **Unione** di E con F:

  $$E \cup F \ = \ \{\omega \in \Omega | \omega \in E \ o \ \omega \in F\}$$

  Probabilisticamente, $E \cup F$ si verifica se **almeno uno** tra E ed F si verifica. L'unione corrisponde ad una **"o" inclusiva**.

- **Differenza** tra insiemi. Ne esistono di 2 tipi:
  - $E \backslash F$ *(si legge E meno F)* **non simmetrica**: significa che ($F \backslash E \neq E \backslash F$).

    $$E \backslash F \ = \{\omega \in \Omega \mid \omega \in E \ e \ \omega \notin F\}$$

    Probabilisticamente, si dice che $E \backslash F$ si verifica quando si verifica E **ma non si verifica** F.
  - $E \triangle F$: è la differenza **simmetrica**:

    $$E \triangle F \ = \ \{\omega \in \Omega \mid \omega \in (E \backslash F) \ o \ \omega \in (F \backslash E)\}$$

    Probabilisticamente, si dice che $E \triangle F$ si verifica se **esattamente uno** tra E ed F si verifica. Corrisponde ad una **"o" esclusiva**.

Esempio Vediamo l'esempio dei 3 lanci della moneta visto in precedenza. Abbiamo definito: $E_K = $ "*Ottengo* K *croci*", $K = 0, 1, 2, 3..$ Abbiamo un evento

$$E = \text{"}\textit{esce almeno una testa}\text{"} = E_3^C$$

il quale è il **complementare** di

$$F = \text{"}\textit{escono almeno 2 croci}\text{"} = E_2 \cup E_3$$

Abbiamo anche l'evento

$$G = \text{"}\textit{escono almeno due teste}\text{"} = (E_2 \cup E_3)^C$$

1.3 Proprietà fondamentali di unione ed intersezione

Esistono varie leggi

- Leggi **commutative**

  $$E \cup F = F \cup E$$

  ed anche

  $$E \cap F = F \cap E$$

- Leggi **associative**

  $$(E \cup F) \cup G = E \cup (F \cup G)$$

  ed anche

  $$(E \cap F) \cap G = E \cap (F \cap G)$$

- Leggi **distributive**

  $$E \cup (F \cap G) = (E \cup F) \cap (E \cup G)$$

  ed anche

  $$E \cap (F \cup G) = (E \cap F) \cup (E \cap G)$$

- Leggi di **De Morgan**

  $$(E \cup F)^C = E^C \cap F^C$$

  ed anche

  $$(E \cap F)^C = E^C \cup F^C$$

1.4 Decomposizioni in unioni disgiunte

E' un metodo conveniente quando si tratteranno le probabilità. Dobbiamo innanzitutto **ripartire lo spazio campionario**.

- **Partizione di Ω**: E' una famiglia di eventi $\{E_N\}_{N\geq 1}$. Gli eventi sono **a 2 a 2 disgiunti** e la proprietà è anche detta *mutua disgiunzione* o *mutua incompatibilità* e la loro unione porta ad ottenere lo spazio campionario Ω di partenza. Simbolicamente diciamo:
  - $E_i \cap E_j = \varnothing$ per $i \neq j$
  - $\bigcup_{N\geq 1} E_N = \Omega$

  Significa che fra gli eventi **non deve esserci intersezione** e quando verranno riuniti essi copriranno **tutto lo spazio campionario**.

  Esempi

  1. Se $E \subseteq \Omega$ è un evento, allora la famiglia $\{E, E^C\}$ è una partizione di Ω.
  2. Prendiamo di nuovo l'esempio dei 3 lanci di una moneta. La famiglia che costruiamo con gli eventi $E_K = \{E_0, E_1, E_2, E_3\}$ (che, ricordiamo, sono l'ottenimento di K croci), questa famiglia **è una partizione di Ω**.

- **Decomposizione di un evento rispetto ad una partizione**: Siano F evento e $\{E, E^C\}$ una partizione di Ω. Allora

  $$F = (F \cap E) \cup (F \cap E^C)$$

  Possiamo dire che $(F \cap E)$ e $(F \cap E^C)$ sono eventi **disgiunti**.

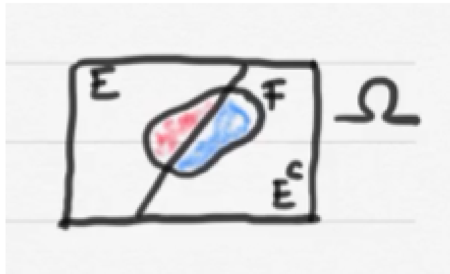

Figure 1.1: Questa è la rappresentazione grafica di ciò che abbiamo appena descritto

In generale, se $\{E_n\}_{n\geq 1}$ è una partizione di Ω, abbiamo

$$F = \bigcup_{n\geq 1} (F \cap E_n)$$

- **Decomposizione dell'unione di due eventi**: serve a riscrivere un unione in modo che gli eventi che vado ad unire per rappresentarla siano **a due a due disgiunti**. Non esiste una regola ed esistono varie decomposizioni. Siano

  $$E, F \subseteq \Omega \ eventi$$

  si può avere il primo caso (1.2)

  $$E \cup F = (E \backslash F) \cup (E \cap F) \cup (F \backslash E)$$

  Altrimenti possiamo avere anche il secondo caso (1.3):

  $$E \cup F = E \cup (F \backslash E)$$

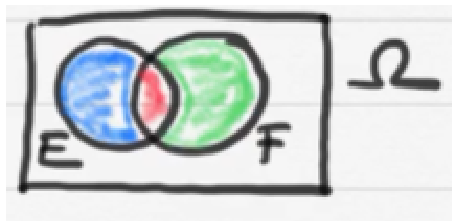

Figure 1.2: Ecco la raffigurazione del primo esempio di decomposizione: la parte blu indica $(E \backslash F)$, la parte rossa indica $(E \cap F)$ mentre la parte verde indica $(F \backslash E)$. Interessante notare come questi tre eventi siano anche **disgiunti** tra loro.

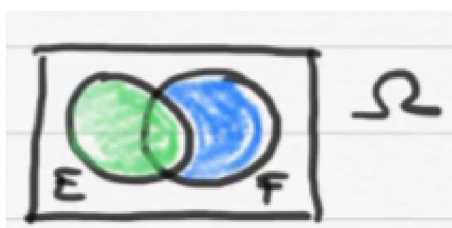

Figure 1.3: Ecco la raffigurazione del secondo esempio. In verde abbiamo E, mentre in blu abbiamo $(F \backslash E)$. Questi due eventi restano **disgiunti**.

Chapter 2

Lezione 2

2.1 La sigma-algebra degli Eventi

Una famiglia F (ricorda F corsivo) di sottoinsiemi di Ω si chiama $\sigma - algebra$ se:

1. F **è non vuota**
2. se $E \in F$, allora $E^C \in F$
3. se $E_n \in F$ per $n \geq 1$, allora $\bigcup_{n\geq 1} E_n \in F$; identifica la chiusura per unione numerabile.

2.1.1 Esempi

Vediamo alcuni esempi:

1. $\sigma - algebra$ **banale**: qualunque sia Ω, la famiglia $\{\Omega, \varnothing\}$ è una $\sigma - algebra$. E' quindi $\sigma - algebra$ lo spazio campionario Ω e lo spazio vuoto $\varnothing$.
2. $\sigma - algebra$ **generata da un evento**: dato Ω e un evento $E \subseteq \Omega$, la famiglia $\{\varnothing, E, E^C, \Omega\}$ è una $\sigma - algebra$.
3. $\sigma - algebra$ **massima**: qualunque sia Ω, la famiglia di tutti i suoi sottoinsiemi (detta **insieme delle parti**)

   $$F = \mathbb{P}(\Omega)$$

   è una $\sigma - algebra$. E' la più grande $\sigma - algebra$ costruibile.

   Nota: se Ω è discreto, prenderemo sempre $F = \mathbb{P}(\Omega)$

2.1.2 Conseguenze elementari degli assiomi

Vediamo alcune conseguenze degli assiomi appena descritti:

1. $\Omega \in F$ e $\varnothing \in F$
   Come lo vediamo? Esiste un $E \in F$ per la condizione 1. vista prima, ciò implica $\longrightarrow E^C \in F$ per la condizione 2. di prima.
   Ora vediamo $\Omega = \{E \cup E^C\} \in F$, poichè la loro unione determina un numero **finito** di elementi, dato dalla condizione 3. vista prima. In conclusione diciamo anche che $\Omega \in F$ e anche che $\varnothing = \Omega^C \in F$ per la condizione 2.

2. Chiusura per **intersezione numerabile**: dice che se ho $E_n \in F$ per ogni $n \geq 1$, allora anche $\bigcap_{n\geq 1} E_n \in F$ sarà nella $\sigma - algebra$.
   Come la vediamo? Devo riscrivere questa intersezione in un modo che mi permetta di utilizzare tutte le informazioni che ho, che sono solo su **unione e complementazione** di eventi sulla $\sigma - algebra$. Sfrutto le **Leggi di De Morgan**:

   - $(\bigcap_{n\geq 1} E_n)^C \ = \bigcup_{n\geq 1} E_n^C \ \in F.$
     Gli $E_n^C \in F$ per la **condizione 2.**. La loro **unione numerabile** $\in F$ per la **condizione 3.**
   - Chiudiamo dicendo: $\bigcap_{n\geq 1} E_n = ((\bigcap_{n\geq 1} E_n)^C)^C$ dove la **prima intersezione** complementare $\in F$per il primo punto di questo elenco, ed il **suo complementare** $\in F$ per la **condzione 2.** della definizione precedente. Quindi l'oggetto è un elemento della $\sigma - algebra$.

2.2 Misura di Probabilità P

Definizione Una misura di probabilità è una **mappa**

$$P : F \longrightarrow [01]$$

in cui

$$E \longmapsto P(E)$$

dove $P(E)$ è la probabilità dell'evento E, tale che soddisfi 2 proprietà:

1. $P(\Omega) = 1$, è la **normalizzazione**

2. se $\{E_n\}_{n\geq 1}$ famiglia di eventi **mutualmente incompatibili** (ovvero quando l'intersezione fra i,j di una famiglia è vuota quando $i \neq j$), allora

$$P(\bigcup_{n\geq 1} E_n) = \sum_{n\geq 1} P(E_n)$$

La misura, come detto nella prima lezione, deve essere **positiva e normalizzata**.

- La **positività** è implicita nel fatto che la probabilità prende **valori nell'intervallo [0,1]**, quindi certamente positivi
- La condizione di **normalizzazione** invece è la condizione sulla probabilità dell'**intero spazio campionario**, come indicato prima.

2.2.1 Conseguenze elementari degli assiomi

1. $P(E^C) = 1 - P(E)$: come lo vediamo?
   Si ha che $1 = P(\Omega) = P(E \cup E^C) =^{per\ condizione\ 2\ definizione} P(E) + P(E^c)$, dove $E\ ed\ E^c$ sono **disgiunti**.
   Da cui quindi otteniamo
   $$P(E^c) = 1 - P(E)$$
2. $P(\varnothing) = 0$: come lo vediamo?
   Si ha $\varnothing = \Omega^c$, quindi posso usare la prima proprietà appena definita per dire:
   $$P(\varnothing) = P(\Omega^c) =^{(per\ proprieta'\ 1)} 1 - P(\Omega)_{(=1\ per\ condizione\ 1)} = 0$$
3. Se $E \subset F$, allora $P(E) < P(F)$: come lo vediamo?
   Se $E \subset F$, allora $F = E \cup (F \backslash E)$, dove i due eventi che sto unendo sono **disgiunti**. Quindi
   $$P(F) =^{(per\ proprieta'\ 2\ Misura\ delle\ Probabilita')} P(E) + P(F \backslash E)$$
   $P(F \backslash E) \in [0, 1]$, che è **positivo**, da cui segue
   $$P(F) > P(E)$$
4. Formula di **inclusione/esclusione**: ci dà modo di calcolare la probabilità di una **unione di eventi**.
   $$P(E \cup F) = P(E) + P(F) - P(E \cap F)$$
   Come la vedo?
   Si ha $E \cup F = E \cup (F \backslash E)$ (unione disgiunta) e quindi $P(E \cup F) = P(E) + P(F \backslash E)$.
   Inoltre $F = (F \backslash E) \cup (F \cap E)$ (unione disgiunta) e quindi $P(F) = P(F \backslash E) + P(F \cap E)$.
   Ricavo che: $P(F \backslash E) = P(F) - P(F \cap E)$. Lo sostituisco nella probabilità dell'unione scritta inizialmente ed ottengo:
   $$P(E \cup F) = P(E) + P(F) - P(F \cap E)$$

2.2.2 Osservazione:

Dato (Ω, F) per completare la terna, la scelta della misura di probabilità P **non è univocamente determinata**, poichè gli assiomi **non determinano** un' unica misura di probabilità.

Esempio: Consideriamo il lancio di una moneta:

$$\Omega = \{T, C\} \ e \ F = \mathbb{P}(\Omega)$$

Abbiamo **infinite** misure di probabilità compatibili con gli assiomi:

$$(T) = p \in [0, 1] \ , \ P(C) = 1 - p$$

Chapter 3

Lezione 3

3.1 Scelta della Misura di Probabilità P

3.1.1 Scelta di P nel caso di Omega Discreto

Consideriamo (Ω, F, P), con Ω discreto e $F = \mathbb{P}(\Omega)$. Come si assegna P su F?
Se $\Omega = \{\omega_1, \omega_2, ...\}$ è discreta, allora ogni evento $E \subseteq \Omega$ (o, equivalentemente $E \in F$) può essere scritto come **unione numerabile (o finita)** di elementi di Ω, cioè significa che:

$$E = \{\omega_{i1}, \omega_{i2}, \omega_{i3}, ...\} = \{\omega_{i1}\} \cup \{\omega_{i2}\} \cup \{\omega_{i3}\} \cup ...$$

(dove gli indici i sono scelti precisamente) è una **unione disgiunta di singoletti**.
Se P fosse la probabilità a cui siamo interessati, allora avremmo che

$$P(E) = P(\{\omega_{i1}) + P(\{\omega_{i2}) + P(\{\omega_{i3}) + ...\}$$

Se leggiamo dal **punto di vista opposto**: è sufficiente assegnare

$$P(\{\omega_i\}) = p_i$$

con i = 1, 2, 3, ... tali che:

- $p_i \in [0, 1]$, quindi la **positività**
- $\sum_{i \geq 1} p_i = 1$, quindi la **normalizzazione**

Quindi se Ω è discreto, assegno le **proprietà ai singoletti** che poi andrò ad **unire** per ottenere la probabilità dell'evento P.

Esempio

Consideriamo $\Omega = \mathbb{N}$ e definiamo $P(\{i\}) = \frac{1}{2^i}$ per ogni $i \in \mathbb{N}$. Questa è **una misura di probabilità**. Infatti verifico:

- $0 < \frac{1}{2^i} \leq 1$
- $\sum \frac{1}{2^i} = -1 + \sum_{i \geq 0} \frac{1}{2^i} = -1 + \frac{1}{1-\frac{1}{2}} = 1$
  (Piccolo ripasso **serie geometrica**: $\sum_{K \geq 0} x^K$, dove x è detta **ragione**, abbiamo che:
  - **converge** se $0 < x < 1$
  - **diverge** altrimenti

  Se converge $\sum_{K \geq 0} x^K = \frac{1}{1-x}$)

Possiamo utilizzare questa misura di probabilità per calcolare la probabilità dell'insieme dei multipli di 3 . Si ha infatti:

$$P(\{3, 6, 9, 12, ...\}) = P(\{3 * i \ con \ i \in \mathbb{N}\}) =$$

$$= \sum_{i \geq 1} \frac{1}{2^{3i}} = \sum_{i \geq 1} \frac{1}{8^i} = -1 + \sum_{i \geq 0} \frac{1}{8^i} = -1 + \frac{1}{1 - \frac{1}{8}} = \frac{1}{7}$$

3.1.2 Scelta di P nel caso di Omega Finito

Se $\Omega = \{\omega_1, \omega_2, ..., \omega_N\}$ ha **cardinalità finita** $\mathbb{N}$, possiamo scegliere la **misura di probabilità uniforme**, detta uniforme poichè assegna a **tutti gli esiti la stessa probabilità**, ovvero:

$$P(\{\omega_1\}) = P(\{\omega_2\}) = ... = P(\{\omega_N\}) = \frac{1}{N}$$

Questo mi dice che **tutti** gli esiti sono **equiprobabili**.
Lo spazio di probabilità che ottengo scegliendo $\Omega finito, \sigma - algebra \ delle \ parti$ e spazio di probabilità uniforme è detto **spazio di probabilità uniforme**. Come faccio a calcolare la probabilità di un evento E?
Per ogni evento $E \in F$, si ha

$$P(E) = \sum_{\omega_i \in E} P(\{\omega_i\}) = |E| * \frac{1}{N} = \frac{|E|}{|\Omega|}$$

(ovvero cardinalità di E fratto cardinalità di Omega). Questa ultima dicitura indica la famosa descrizione:

$$\frac{Casi \ Favorevoli}{Casi \ Possibili}$$

Esempio

Consideriamo un mazzo di carte da poker (52 carte) e peschiamo una carta.

- La probabilità di estrarre una regina è $\frac{4}{52}$
- La probabilità di estrarre una carta di cuori è $\frac{13}{52}$
- La probabilità di estrarre la regina di cuori è $\frac{1}{52}$

Esercizio

Lanciamo due dadi equilibrati.

1. E_1 = "Qual è la probabilità che la somma dei punteggi sia 8?"
2. E_2 = "Qual è la probabilità che escano due punteggi uguali?"
3. E_3 = "Qual è la probabilità di ottenere 8 con due dadi con lo stesso punteggio?"

Soluzione Lo spazio campionario è

$$\Omega = \{(i, j) : \ i, j \in \{1, 2, ..., 6\}\}$$

cioè l'insieme delle coppie dei punteggi ottenuti sui due dadi. Si ha $|\Omega| = 6 * 6 = 36$

1. Consideriamo l' evento
$$E_1 = \{(2, 6), (3, 5), (4, 4), (5, 3), (6, 2)\}$$
si ha
$$P(E) = \frac{|E|}{|\Omega|} = \frac{5}{36}$$
2. Consideriamo l'evento
$$E_2 = \{(1, 1), (2, 2), (3, 3), (4, 4), (5, 5), (6, 6)\}$$
si ha
$$P(E) = \frac{|E|}{|\Omega|} = \frac{6}{36}$$
3. Consideriamo l'evento
$$E_3 = E_1 \cap E_2 = \{(4, 4)\}$$
si ha
$$P(E) = \frac{|E|}{|\Omega|} = \frac{1}{36}$$

Punto chiave: Il punto chiave è il **contare gli elementi di un insieme**.

3.2 Combinatoria Elementare

3.2.1 Principio Fondamentale del Conteggio (P.F.C)

Sia E un insieme. Supponiamo che ogni elemento di E **si possa determinare univocamente mediante K scelte** successive, tali che:

- La prima scelta ha n_1 esiti possibili
- La seconda scelta ha n_2 esiti possibili
- La **K-esima** scelta ha n_k esiti possibili

allora la cardinalità di E sarà:

$$|E| = n_1 * n_2 * ... * n_k$$

Esempio

Contiamo le carte in un mazzo da poker, usando il P.F.C. Ogni carta del mazzo è individuata da un seme (C, Q, F, P) e da un valore (A, 2, 3, ..., J, Q, K). Per individuare una carta, cosa faccio?

1. Scelgo il **seme**: **4 esiti possibili**
2. Scelgo il **valore**: **13 esiti possibili**

Quindi otteniamo che il numero di carte nel mazzo è:

$$4 * 13 = 52$$

Osservazione Cosa vuol dire "determinare univocamente"? Vuol dire che **sequenze distinte di scelte individuano elementi diversi di E**.
E' importante, perchè **se non è soddisfatta** la cardinalità risulterà **errata!** e il P.F.C non varrà.

Esempio 2

Si deve formare un comitato di 2 persone da scegliersi dal gruppo $\{U, D_1, D_2\}$ (dove U = Uomo e D = Donna). I possibili comitati sono 3:

1. $\{U, D_1\}$
2. $\{U, D_2\}$
3. $\{D_1, D_2\}$

Ora ragioniamo come segue. Scegliere un comitato equivale a fare le seguenti scelte:

1. Scelgo una delle due donne: ho 2 esiti possibili
2. Scelgo la persona che manca tra i rimanenti: ho 2 esiti possibili.

Applico il P.F.C ed ottengo che il numero dei comitati è $2*2 = 4$, ma **in realtà** i comitati **sono solo 3**. Dove sta l'errore?
Sta nel fatto che ci sono **2 distinte sequenze di scelte** che mi danno come risultato il comitato $\{D_1, D_2\}$ e sono:

- $\{D_1, D_2\}$
- $\{D_2, D_1\}$

Chapter 4

Lezione 4

4.1 Disposizioni con Ripetizione

Scopo: Contare le K-uple **ordinate** che posso creare scegliendo ogni entrata da n oggetti con la possibilità di ripetizione.
Ogni entrata può essere scelta tra **n alternative** (ci può essere **ripetizione**) e faccio **K scelte** totali, una per ogni entrata del vettore che voglio costruire. Quindi, per sapere il numero totale di disposizioni con ripetizione è

$$n^K$$

Esempi

1. Una cassaforte con codice a 7 cifre ha 10^7 codici possibili.

2. Metodo per generare sottoinsiemi di un dato insieme: se E è un insieme **finito**, con cardinalità **K**, allora i possibili sottoinsiemi di E sono

   $$2^K$$

   Infatti, possiamo specificare un sottoinsieme F di E nel seguente modo: ad ogni elemento di F possiamo assegnare:

   - Il valore **1** se **sta in F**
   - Il valore **0** se **non sta in F**

   Ogni stringa di bit $\{0, 1\}$ di lunghezza K **codifica un sottoinsieme di E**.

   Esempio: E = $\{+, -, *, :\}$, $|E| = 4$ con stringa di bit: 0101 che genera il sottoinsieme:

   $$\{-, :\}$$

poichè i valori corrispondono esattamente agli elementi in E ($" + " = 0$ *quindi no*, $" - " = 1$ *quindi si, eccetera.*) Per contare i sottoinsiemi devo quindi **contare le stringhe**. Quante stringhe di lunghezza K posso formare scegliendo le entrate in $\{0, 1\}$?
Sono 2^K.

4.2 Disposizioni senza Ripetizione

Scopo: contare le K-uple **ordinate** (poichè sono disposizioni quindi l'ordine conta) che posso creare scegliendo ciascuna entrata da n oggetti senza possibilità di ripetizione.

- La prima entrata può essere scelta tra **n alternative**
- La seconda tra **n-1 alternative**
- Fino alla **K-esima** che può essere scelta tra **n-k+1** alternative

Il numero delle disposizioni senza ripetizione sarà $n*(n-1)*(n-2)*...*(n-k+1)$ con K fattori. Se **K = n** ottengo le **permutazioni** che contano tutti i modi possibili per ordinare gli n oggetti. In particolare si ottiene che il numero di possibili permutazioni è

$$n!$$

Esempio

Cinque amici fanno una gara di nuoto. Le possibili classifiche sono

$$5! = 120$$

ovvero tutti i possibili ordinamenti dei nomi.
I possibili podii sono

$$5 * 4 * 3 = 60$$

rispettivamente 5 per l'oro, 4 per l'argento e 3 per il bronzo.

4.3 Esercizi

1. Un'urna contiene palline numerate da 1 a 50. Si estraggono contemporaneamente 2 palline. Calcolare la probabilità di ottenere:
   (a) Due numeri dispari
   (b) Un numero divisibile per 5 e uno non divisibile per 5
   (c) Due numeri la cui somma è 50.

Soluzione Considero Ω l'insieme di combinazioni di 2 palline scelte tra le 50 nell'urna, senza tenere conto dell'ordine. Si ha $|\Omega| = \binom{50}{2} = 1225$
Il coefficiente binomiale $\binom{n}{k}$ è

$$\frac{k!}{(n-k)!}$$

(a) Evento A: Ottengo

$$P(A) = \frac{\binom{25}{2}}{1225} = \frac{12}{49}$$

dove 25 sono le palline dispari (50-25 = 25)

(b) Evento B: Ottengo

$$P(B) = \frac{10 * 40}{1225} = \frac{16}{49}$$

dove 10 sono i numeri **divisibili** per 5 e 40 sono i numeri **non divisibili** per 5

(c) Evento C: Ottengo

$$P(C) = \{1, 49\}, \{2, 48\}, ..., \{24, 26\} = 24 \xrightarrow{porta\ al\ risultato} \frac{24}{1225}$$

Soluzione Alternativa Si può risolvere tenendo conto dell' ordine, considerando $\overline{\Omega}$ come l'insieme delle disposizioni di 2 palline scelte dalle 50 dall'urna. Si ha:

$$|\overline{\Omega}| = 50 * 49$$

(a) Evento A: Ottengo

$$P(A) = \frac{25 * 24}{50 * 49} = \frac{12}{49}$$

(b) Evento B: Questo evento diventa l'**unione disgiunta** dei due eventi
   - B_1 = "il primo numero è divisibile per 5 e il secondo no"
   - B_2 = "il primo numero non è divisibile per 5 e il secondo si"

   Ottengo

$$P(B) = P(B_1 \cup B_2) = P(B_1) + P(B_2) = \frac{10 * 40}{50 * 49} + \frac{40 * 10}{50 * 49} = \frac{16}{49}$$

(c) Evento C: Le coppie ordinate che sommano a 50 sono (1,49), (49,1), ... (24,26), (26,24)

$$P(C) = \frac{48}{50 * 49} = \frac{24}{1225}$$

2. Lanciamo 3 dadi equilibrati. Calcolare la probabilità di ottenere:
   (a) Tre numeri dispari
   (b) Due numeri pari e uno dispari
   (c) Tre numeri la cui somma è 5
   (d) Almeno due 1.

Soluzione: Considero

$$\Omega = \{(i, j, k) : i, j, k \in \{1, ..., 6\}\}$$

Questo spazio campionario **tiene conto dell'ordine**. Si ha $|\Omega| = 6^3 = 216 esiti possibili.$

(a) Evento A: si ha

$$P(A) = \frac{3^3}{216} = \frac{1}{8}$$

(b) Evento B: si ha

$$P(B) = \frac{3^4}{216} = \frac{3}{8}$$

Si ha 3^4 perchè si hanno 3 posizioni diverse per i dadi che vanno incluse nel calcolo della probabilità totale, oltre alle scelte da fare sui dadi.

(c) Evento C: si ha

$$P(C) = \frac{2 * 3}{216} = \frac{1}{36}$$

Questo perchè si hanno solo 2 terne che rispondono alla richiesta (le terne "113" e "221") ed hanno 3 ordinamenti diversi.

(d) Evento D: questo evento è l'unione disgiunta dei due seguenti eventi:

- D_1 = "Ottengo esattamente due 1"
- D_2 = "Ottengo tre 1"

si ha quindi

$$P(D) = (D_1 \cup D_2) = \frac{3 * 5 * 1}{216} + \frac{1}{216} = \frac{2}{27}$$

questo perchè: in D_1 vi è un caso in cui non deve uscire "1" $(\frac{3*5}{6})$ mentre gli altri due casi sono $\frac{1}{6} * \frac{1}{6}$;
in D_2 invece ho solo un caso in cui ottengo esattamente tre "1", quindi ho $\frac{1}{6} * \frac{1}{6} * \frac{1}{6} = \frac{1}{216}$.

4.4 Birthday Problem

E' anche detto *paradosso dei compleanni*, in informatica esiste un attacco di tipo brute force chiamato "Birthday Attack".
Questo problema consiste nel calcolare la probabilità dell'evento
E_n = "in una classe di n bambini almeno 2 hanno lo stesso compleanno" con $n \leq 365$.
L'anno è composto da 365 giorni (no anno bisestile), che identifico con i numeri da 1 (1 Gennaio) a 365 (31 Dicembre). Lo spazio campionario sarà:

$$\Omega_n = \{1, ..., 365\}^n$$

e la misura di probabilità è **uniforme**. Dobbiamo calcolare

$$P(E_n) = \frac{|E_n|}{|\Omega_n|} = \frac{|E_n|}{365^n}$$

Il calcolo di $|E_n|$ è complicato, quindi passo all'evento complementare, ossia $E_n^c =$ "in una classe di n bambini tutti i compleanni avvengono in giorni distinti". Si ha

$$P(E_n) = 1 - P(E_n^c) = 1 - \frac{365 * (365 - n + 1)}{365^n}$$

Quello che ci interessa è calcolare n_* definito come il primo n per cui $P(E_n) \geq \frac{1}{2}$, cioè

$$n_* = min\{n \mid P(E_n) \geq \frac{1}{2}\}$$

Ciò lo si vede con un grafico fatto con un calcolatore e si trova che $n_* = 23$, quindi basta un numero piccolissimo di persone per avere una alta probabilità di persone con lo stesso compleanno all'interno del gruppo.

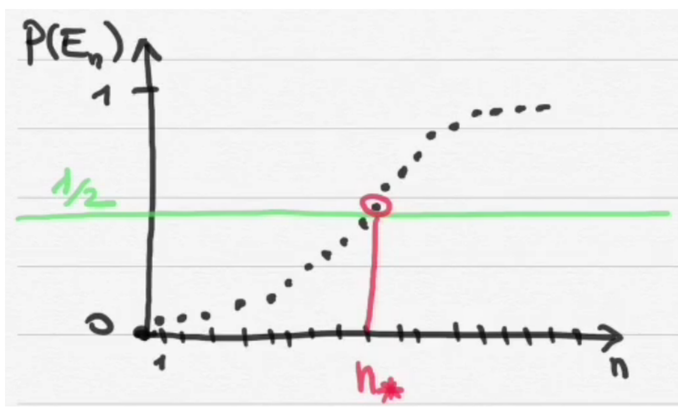

Figure 4.1: Questo è il grafico che permette di vedere che quale è il minimo n_* da prendere in considerazione

Chapter 5

Lezione 5

5.1 Probabilità Condizionata

Esempio Urna: 7 palline nere e 3 palline rosse. Estraggo 2 palline. Qual è la probabilità che la seconda pallina estratta sia rossa?
La risposta dipende da **quale pallina ho estratto alla prima estrazione**: la prima estrazione può avere 2 esiti diversi

- Pallina nera: resteranno **6 nere** e **3 rosse**, la probabilità che la seconda pallina sia rossa è $\frac{1}{3}$
- Pallina rossa: resteranno **7 nere** e **2 rosse**, la probabilità che la seconda pallina sia rossa è $\frac{2}{9}$

quindi la probabilità **cambia a seconda della scelta** che faccio.
Con questo concetto rispondiamo alle affermazioni del tipo: "la probabilità di E sapendo che si è verificato l'evento F".

5.2 Definizione Probabilità Condizionata

Siano (Ω, F, P) uno spazio di probabilità e $F \in F$ un evento tale che $P(F) > 0$. Allora, per ogni altro evento $E \in F$ è ben definita la quantità

$$P(E|F) = \frac{P(E \cap F)}{P(F)}$$

Questa è la probabilità condizionata di **E dato F**.

Osservazione: Se P(E) e P(F) sono entrambe strettamente positive, allora

$$P(E \cap F) = P(E|F) * P(F) = P(F|E) * P(E)$$

Questa formula si chiama **formula di moltiplicazione**, è un trucco per calcolare l'intersezione.

Giustificazione dell'osservazione: Mettiamoci su uno spazio di probabilità **uniforme** e supponiamo di avere un'urna fatta in questo modo:

- Ci sono 8 palline numerate da 1 a 8
- Le palline "3", "5" e "8" sono nere
- Le restanti palline sono rosse

Consideriamo gli eventi:

- E = "estraggo una pallina con numero pari"
- F = "estraggo una pallina rossa"

$P(E) = \frac{1}{2}$. Come cambia la situazione se voglio calcolare $P(E|F)$?

$$P(E|F) = \frac{P(E \cap F)}{P(F)} = \frac{|E \cap F|}{|\Omega|} * \frac{|\Omega|}{|F|} = \frac{|E \cap F|}{|F|} = \frac{3}{5}$$

Questa ultima espressione risponde all'affermazione "casi favorevoli su casi possibili", ma in uno **spazio di probabilità ridotto**. Quindi si può dire che **condizionare la scelta riduce lo spazio degli esiti possibili**.

Osservazione $P(E|F)$ è **minore, maggiore o uguale** a $P(E)$.

5.2.1 Teorema

Sia $F \in F$ un evento con $P(F) > 0$, allora la mappa

$$P(\cdot|F) : F \to \mathbb{R}$$

$$E \longmapsto P(E|F)$$

è una misura di probabilità. Da ciò seguono due conseguenze importanti:

1. $P(E^c|F) = 1 - P(E|F)$
2. $P(E \cup G|F = P(E|F) + P(G|F) - P(E \cap G|F)$, è la formula di **inclusione/esclusione**

Attenzione: Fissato $E \in F$, la mappa di

$$P(E|\cdot) : F \to \mathbb{R}$$

$$F \longmapsto P(E|F)$$

non è una misura di probabilità e quindi

$$P(E|F^c) \neq 1 - P(E|F)$$

5.3 Formula delle probabilità totali

5.3.1 Teorema

Siano (Ω, F, P) uno spazio di probabilità e $F \in F$ un evento tale che $0 < P(F) < 1$, quindi un evento non impossibile e non certo, allora, per ogni $E \in F$, vale la formula delle probabilità totali:

$$P(E) = P(E|F) \cdot P(F) + P(E|F^c) \cdot P(F^c)$$

In generale, se ho $\{F_k\}_{k=1}^n$ che è una partizione di Ω con $P(F_k) > 0$ per ogni k, vale la formula delle probabilità totali:

$$P(E) = \sum_{k=1}^{n} P(E|F_k) \cdot P(F_k)$$

Dimostrazione: Dimostriamo il caso generale. Scomponiamo l'evento E sulla partizione $\{F_k\}_{k=1}^n$ ed otteniamo:

$$E = \bigcup_{k=1}^{n} (E \cap F_k)$$

ovvero una unione disgiunta. Quindi calcoliamo:

$$P(E) = P(\bigcup_{k=1}^{n} (E \cap F_k)) = \sum_{k=1}^{n} P(E \cap F_k)$$

che ci porta a

$$\sum_{k=1}^{n} P(E|F_k) \cdot P(F_k)$$

c.v.d.

Esempio

Urna del primo esempio: 7 palline nere e 3 rosse. Estraggo 2 palline e devo determinare quale è la probabilità che la seconda pallina sia rossa.
Definisco gli eventi: R_i = "l'i-esima pallina estratta è rossa", con i = 1,2 ed ottengo:

$$P(R_2) = P(R_2|R_1) \cdot P(R_1) + P(R_2|R_1^c) \cdot P(R_1^c) = \frac{2}{9} \cdot \frac{3}{10} + \frac{1}{3} \cdot \frac{7}{10} = \frac{3}{10}$$

dove le frazioni indicano, rispettivamente:

- La probabilità della seconda pallina estratta dopo aver estratto una rossa ($\frac{2}{9}$)
- La probabilità di estrarre una pallina rossa alla prima estrazione ($\frac{3}{10}$)
- La probabilità di avere una pallina rossa dopo una pallina nera ($\frac{1}{3}$)
- La probabilità di estrarre una pallina nera alla prima estrazione ($\frac{7}{10}$)

5.3.2 Esercizio: Tre Urne - Prima parte

Ho tre urne.

- Prima urna: 3 palline bianche e 2 nere
- Seconda urna: 3 palline bianche e 3 nere
- Terza urna: 4 palline bianche e 1 nera

Lancio un dado equilibrato:

- Se esce **6** estraggo una pallina dalla **terza urna**
- Se esce **4 o 5** estraggo dalla **seconda urna**
- Negli altri casi estraggo dalla **prima urna**

Quale è la probabilità di estrarre una pallina bianca?

Soluzione Consideriamo gli eventi U_i = "peschiamo dall'urna i", con i = 1,2,3 e l'evento B = "peschiamo una pallina bianca". Abbiamo:

- $P(U_1) = P(\textit{punteggio dado} \in \{1,2,3\}) = \frac{1}{2}$
- $P(U_2) = P(\textit{punteggio dado} \in \{4,5\}) = \frac{1}{3}$
- $P(U_3) = P(\textit{punteggio dado} = 6) = \frac{1}{6}$

Dal testo ci possiamo anche ricavare le probabilità condizionate:

- $P(B|U_1) = \frac{3}{5}$
- $P(B|U_2) = \frac{1}{2}$
- $P(B|U_3) = \frac{4}{5}$

Quindi, per la formula delle probabilità totali otteniamo:

$$P(B) = P(B|U_1) \cdot P(U_1) + P(B|U_2) \cdot P(U_2) + P(B|U_3) \cdot P(U_3)$$

che equivale a

$$\frac{3}{5} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{3} + \frac{4}{5} \cdot \frac{1}{6} = \frac{18}{30} = \frac{3}{5}$$

5.3.3 Formula di Bayes

Questa formula serve per stimare le probabilità a posteriori, quindi a "rovesciare" il condizionamento. Questo avviene da

$$P(F_k|E) = \frac{P(F_k \cap E)}{P(E)} \cdot \frac{P(F_k)}{P(F_k)}$$

Dati questi elementi posso capire che

$$\frac{P(F_k \cap E)}{P(F_k)} = \frac{P(E|F_k) \cdot P(F_k)}{P(E)}$$

ed è questa ultima parte (dopo l'uguale) la formula di Bayes.

Esercizio: Tre Urne - Seconda Parte

Tenendo le informazioni della sezione 5.3.2, sapendo che è stata estratta una pallina bianca, quale è la probabilità che sia stata estratta dalla terza urna?

Soluzione Dobbiamo calcolare $P(U_3|B)$. Da quanto visto prima sappiamo che

$$P(U_3) = \frac{1}{6},\ P(B|U_3) = \frac{4}{5}\ e\ P(B) = \frac{3}{5}$$

Per la formula di Bayes, si ha

$$P(U_3|B) = \frac{P(B|U_3) \cdot P(U_3)}{P(B)} = \frac{\frac{4}{5} \cdot \frac{1}{6}}{\frac{3}{5}} = \frac{2}{9}$$

5.3.4 Esercizi

1. Lungo un canale di trasmissione sono spediti dati in formato 0,1. Il canale è disturbato e a volte capita che si trasmetta uno 0 e si riceva un 1 e vicenversa. Si stima che:

   - Se si trasmette uno 0, viene ricevuto correttamente con probabilità 0.95
   - Se si trasmette un 1, viene ricevuto correttamente con probabilità 0.75
   - Si sa inoltre che il 70% dei segnali trasmessi sono 0.

   Viene spedito un segnale lungo il canale, calcolare la probabilità che:

   a) Venga ricevuto uno 0
   b) Si verifichi un errore di trasmissione
   c) Il segnale spedito fosse uno 0, sapendo che è stato ricevuto uno 0.

   Soluzione Definiamo gli eventi:

   - T_i = "Trasmetto il segnale i", con $i \in \{0, 1\}$
   - R_i = "Ricevo il segnale i", con $i \in \{0, 1\}$

   Dal testo ricaviamo che:

   $$P(R_0|T_0) = 0.95,\ P(R_1|T_1) = 0.75\ e\ P(T_0) = 0.7$$

   a) Dobbiamo calcolare la probabilità di $P(R_0)$. Uso la formula delle probabilità totali:

   $$P(R_0) = P(R_0|T_0) \cdot P(T_0) + P(R_0|T_1) \cdot P(T_1)$$

   Si ha:

- $P(R_0|T_1) = P(R_1^c|T_1) = 1 - P(R_1|T_1) = 1 - 0.75 = 0.25$
- $P(T_1) = P(T_0^c) = 1 - P(T_0) = 1 - 0.7 = 0.3$

Sostituisco ed ottengo:

$$(0.95 \cdot 0.7) + (0.25 \cdot 0.3) = 0.74$$

b) Sia E l'evento "Si verifica un errore di trasmissione".

$$E = (T_0 \cap R_1) \cup (T_1 \cap R_0)$$

Questa è una unione disgiunta, quindi ottengo:

$$P(E) = P(T_0 \cap R_1) + P(T_1 \cap R_0) = P(R_1|T_0) \cdot P(T_0) + P(R_0|T_1) \cdot P(T_1)$$

Si ha :

- $P(R_1|T_0 = P(R_0^c|T_0) = 1 - P(R_0|T_0) = 1 - 0.95 = 0.05$
- $P(R_0|T_1 = P(R_1^c|T_1) = 1 - P(R_1|T_1) = 1 - 0.75 = 0.25$

Infine ottengo che

$$P(E) = (0.05) \cdot (0.7) + (0.25) \cdot (0.3) = 0.11$$

c) Dobbiamo calcolare $P(T_0|R_0)$. E' una probabilità **a posteriori**, quindi si usa la formula di Bayes e otteniamo:

$$P(T_0|R_0) = \frac{P(R_0|T_0) \cdot P(T_0)}{P(R_0)}$$

$$= \frac{(0.95) \cdot (0.7)}{0.74} = 0.9$$

2. Siano A,B eventi tali che P(A) = $\frac{1}{4}$ e $P(A|B) = P(B|A) = \frac{1}{2}$. Si dica, giustificando, se è vero o falso che:

a) A e B sono incompatibili

b) $P(A \cap B) = P(A) \cdot P(B)$

c) $P(A \cap B) > P(A) \cdot P(B)$

Soluzione

a) $P(A \cap B) = P(B|A) \cdot P(A) = \frac{1}{8}$. Se A,B fossero incompatibili, si avrebbe $A \cap B = \varnothing$ e quindi $P(A \cap B) = 0$. **FALSO**.

b) $P(A \cap B) = P(A) \cdot P(B)$. Si ha che:

$$P(A \cap B) = P(A|B) \cdot P(B)$$

. Quindi posso semplificare per P(B) ed otterrei che

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

il che è **ASSURDO**. Quindi b) è **FALSO**.

c) Dato $P(A \cap B) > P(A) \cdot P(B)$, abbiamo, dopo la moltiplicazione per P(B)

$$P(A|B) \cdot P(B) > P(A) \cdot P(B) = P(A \cap B) > P(A) \cdot P(B)$$

e poichè se P(B) non fosse strettamente positivo, $P(A|B)$ non esisterebbe nemmeno, concludiamo che c) è **VERO**.

Chapter 6

Lezione 6

6.1 Eventi Indipendenti

6.1.1 Indipendenza di due eventi

Sia (Ω, F, P) spazio di probabilità. Gli eventi $E, F \in F$ si dicono indipendenti se

$$P(E \cap F) = P(E) \cdot P(F)$$

e si scrive

$$E \perp\!\!\!\perp F \text{ } o \text{ } \{E, F\} \text{ } indipendente$$

Attenzione Indipendenza è **diverso da** incompatibilità.

- Indipendenza è una nozione **probabilistica**, dipende da E,F e dalla misura P.
- Incompatibilità è una nozione **insiemistica**.

In particolare, se E,F $\in F$ di probabilità strettamente positiva sono incompatibili, allora **non possono essere indipendenti**. Infatti, essendo incompatibili, significa che se uno si verifica l'altro **non può certamente verificarsi**, quindi c'è una **forte dipendenza** tra i due eventi.

Esempio (banale) Qualunque sia $E \in F$, allora $E \perp\!\!\!\perp \varnothing$ *e* $E \perp\!\!\!\perp \Omega$. Infatti, si ha

- $P(E \cap \varnothing) = P(\varnothing) = 0 = P(\varnothing) \cdot P(E)$
- $P(E \cap \Omega) = P(E) = P(E) \cdot P(\Omega)$ dove $P(\Omega) = 1$

Esempi

1. Lanciamo una moneta e un dado. Lo spazio campionario naturale è

$$\Omega = \{(a, i) : a \in \{T, C\}, i \in \{1, ..., 6\}\}$$

con la misura uniforme.
Osserviamo che $|\Omega| = 2 \cdot 6 = 12$. Gli eventi:

    - E = "Esce testa"
    - F = "Esce 4"

    sono indipendenti. Infatti, se calcoliamo

$$P(E) = P(\{(T, i) : i \in \{1, ..., 6\}\}) = \frac{1}{2}$$

$$P(F) = P(\{(a, 4) : a \in \{T, C\}\}) = \frac{1}{6}$$

$$P(E \cap F) = P(\{(T, 4)\}) = \frac{1}{12}$$

    otteniamo che $\frac{1}{12} = P(E \cap F) = P(E) \cdot P(F) = \frac{1}{2} \cdot \frac{1}{6}$ Questo ci permette di dire che gli eventi sono indipendenti.

2. Lanciamo un dado due volte. Lo spazio campionario naturale sarà

$$\Omega = \{(i, j) : i, j \in \{1, ..., 6\}\}$$

con la misura di probabilità **uniforme**, quindi tutti gli esiti sono equiprobabili.
Osserviamo che $|\Omega| = 6 \cdot 6 = 36$. Gli eventi:

    - E = "La prima faccia è 4"
    - F = "La somma dei due punteggi è 9"

    non sono indipendenti.

$$P(E) = P(\{(4, j) : j \in \{1, ..., 6\}\}) = \frac{1}{6}$$

$$P(F) = P(\{(3, 6), (4, 5), (5, 4), (6, 3)\}) = \frac{1}{9}$$

$$P(E \cap F) = P(\{(4, 5)\}) = \frac{1}{36}$$

    poichè è solo una la coppia che rispetta la condizione.
    Otteniamo $\frac{1}{36} = P(E \cap F) \neq P(E) \cdot P(F) = \frac{1}{6} \cdot \frac{1}{9} = \frac{1}{54}$, quindi gli eventi **non sono indipendenti**.
    Se invece prendessimo l'evento G = "la somma dei punteggi è 7 " e sostituissimo F con G, gli eventi E e G sarebbero indipendenti.

6.2 Conseguenze elementari dell'indipendenza

Proposizione 1 Siano E,F $\in F$ con $P(E) > 0$ e $P(F) > 0$. Allora le seguenti affermazioni:

(i) $E \perp\!\!\!\perp F$

(ii) $P(E|F) = P(E)$

(iii) $P(F|E) = P(F)$

sono **equivalenti**.

6.2.1 Dimostrazione

(i) $\rightarrow (ii)$ $E \perp\!\!\!\perp F \xrightarrow{implica} P(E \cap F) = P(E) \cdot P(F)$

$$P(E \cap F) = P(E|F) \cdot P(F)$$

quindi posso semplificare i $P(F)$ ed ottengo ciò che volevo dimostrare, ossia

$$P(E|F) = P(E)$$

(ii) $\rightarrow (iii)$ $P(E|F) = P(E) \xrightarrow{(Bayes)} \frac{P(F|E) \cdot P(E)}{P(F)} = P(E)$ semplifico per P(E) e moltiplico per P(F), ottenendo

$$P(F|E) = P(F)$$

(iii) $\rightarrow (i)$ $P(F|E) = P(F) \xrightarrow{implica} \frac{P(E \cap F)}{P(E)} = P(F)$ moltiplico per P(E) ed ottengo

$$P(E \cap F) = P(F) \cdot P(E) \xrightarrow{ci\ porta\ a} E \perp\!\!\!\perp F$$

c.v.d

6.2.2 Proposizione 2

Abbiamo **equivalenza** tra le seguenti istanze:

(i) $E \perp\!\!\!\perp F$

(ii) $E^c \perp\!\!\!\perp F$

(iii) $E^c \perp\!\!\!\perp F^c$

(iv) $E \perp\!\!\!\perp F^c$

Dimostrazione Abbiamo anche qui una catena di implicazioni:

(i) $\rightarrow$ (ii) Dobbiamo mostrare che $P(E^c \cap F) = P(E^c) \cdot P(F)$.
Si ha $E^c \cap F = F \backslash (E \cap F)$ ed inoltre si ha che $E \cap F \subseteq F$.
Quindi $P(E^c \cap F) = P(F) - P(E \cap F)$. Usiamo l'ipotesi (i), quindi otteniamo:

$$P(E^c \cap F) = P(F) - P(E) \cdot P(F) = P(F) \cdot (1 - P(E))$$

ed $(1 - P(E)) = P(E^c)$, quindi

$$P(E^c \cap F) = P(E^c) \cdot P(F)$$

Per concludere si deve dimostrare che $(ii) \rightarrow (iii) \rightarrow (iv) \rightarrow (i)$.

6.3 Indipendenza di Tre Eventi

Gli eventi E_1, E_2, E_3 sono indipendenti se le seguenti due condizioni sono **entrambe** soddisfatte:

1. $E_1 \perp\!\!\!\perp E_2, E_1 \perp\!\!\!\perp E_3, E_2 \perp\!\!\!\perp E_3$
2. $P(E_1 \cap E_2 \cap E_3) = P(E_1) \cdot P(E_2) \cdot P(E_3)$

Osservazione Le due condizioni **non sono rindondanti**, sono infatti logicamente **indipendenti** e necessarie.

6.3.1 Esempi

1. Lancio due volte una moneta. Considero gli eventi:

   - A = "al primo lancio ottengo testa"
   - B = "al secondo lancio ottengo testa"
   - C = "nei due lanci ottengo due facce uguali"

   Mostriamo che gli eventi sono a 2 a 2 indipendenti, ovvero che **vale la prima condizione** ma la seconda non è soddisfatta:

   Soluzione Lo spazio campionario $\Omega = \{(l_1, l_2) : l_1, l_2 \in \{T, C\}\}$ con $|\Omega| = 2 \cdot 2 = 4$

   - Evento A: $P(A) = P(\{(T, l_2) : l_2 \in \{T, C\}\}) = \frac{1}{2}$
   - Evento B: $P(B) = P(\{(l_1, T) : l_1 \in \{T, C\}\}) = \frac{1}{2}$
   - Evento C: $P(C) = P(\{(T, T), (C, C)\}) = \frac{1}{2}$

Inoltre si ha:

$$P(A \cap B) = P(A \cap C) = P(B \cap C) = P(\{(T,T)\} = \frac{1}{4}$$

Si vede facilmente che le coppie {A,B}, {A,C} e {B,C} sono indipendenti. Per esempio prendiamo:

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

le altre sono analoghe. Mostriamo ora che

$$P(A \cap B \cap C) \neq P(A) \cdot P(B) \cdot P(C)$$

- $P(A) \cdot P(B) \cdot P(C) = \frac{1}{8}$
- $P(A \cap B \cap C) = P(\{(T,T)\}) = \frac{1}{4}$

2. Lancio due volte un dado. Considero gli eventi:
   - A = "Il punteggio del primo lancio $\in \{1, 2, 3\}$"
   - B = "Il punteggio del primo lancio $\in \{3, 4, 5\}$
   - C = "La somma dei due punteggi è uguale a 9"

Soluzione Mostriamo che:

- $P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$
- Gli eventi **non sono indipendenti** a 2 a 2.

Lo spazio campionario è

$$\Omega = \{(d_1, d_2) : d_1, d_2 \in \{1, ..., 6\}\} \to |\Omega| = 6 \cdot 6 = 36$$

Calcoliamo:

- $P(A) = P(\{(d_1, d_2) : d_1 \in \{1, 2, 3\}, d_2 \in \{1, ..., 6\}\}) = \frac{3 \cdot 6}{36} = \frac{1}{2}$
- $P(A) = P(\{(d_1, d_2) : d_1 \in \{3, 4, 5\}, d_2 \in \{1, ..., 6\}\}) = \frac{3 \cdot 6}{36} = \frac{1}{2}$
- $P(C) = P(\{(3, 6), (4, 5), (5, 4), (6, 3)\}) = \frac{1}{9}$

Inoltre:

$$P(A \cap B \cap C) = P(\{(3, 6)\} = \frac{1}{36}$$

Quindi

$$P(A \cap B \cap C) = \frac{1}{36} = P(A) \cdot P(B) \cdot P(C) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{9} = \frac{1}{36}$$

Consideriamo ora {A,C} e mostriamo che non sono indipendenti. Sappiamo già che:

$$P(A) \cdot P(C) = \frac{1}{2} \cdot \frac{1}{9} = \frac{1}{18}$$

Calcoliamo

$$P(A \cap C) = P(\{3,6\}) = \frac{1}{36}$$

Si può facilmente vedere che

$$\frac{1}{18} \neq \frac{1}{36}$$

quindi gli eventi **non sono a due a due indipendenti**. Si può procedere in modo analogo per le altre coppie, ma il risultato è lo stesso: basta che solo una coppia non soddisfi la condizione e anche le altre non la soddisferanno.

6.3.2 In generale

La famiglia di $\{E_1, ..., E_n\}$ di eventi è indipendente se **per ogni r** $(2 \leq r \leq n)$ e per **ogni possibile scelta di r eventi distinti** degli n eventi della famiglia, la probabilità dell'intersezione degli r eventi scelti è **pari al prodotto delle loro probabilità**.
Devo quindi dimostrare la **proprietà di fattorizzazione** per ogni coppia, terna, quaterna, ecc. di eventi. Devo considerare **tutte le possibili sottofamiglie**.

Osservazione Se $\{E_1, ..., E_n\}$ è una famiglia di eventi indipendenti, quando sostituisco qualche E_i (non importa quanti e quali) con E_i^c ottengo ancora una famiglia di eventi **indipendenti**.
La famiglia **numerabile** di eventi $\{E_1, E_2, ...\}$ è indipendente se ogni sottofamiglia finita lo è.

Chapter 7

Lezione 7

7.1 Variabili Aleatorie Discrete

Sia $(\Omega, \mathbb{P}(\Omega), P)$ uno spazio di probabilità discreto. Ogni mappa:

$$X : \Omega \to \mathbb{R}$$

è detta **variabile aleatoria** (o casuale) discreta su $\mathbb{R}$.
Nella mappa X, **ad ogni esito viene associato un numero**. Abbiamo che

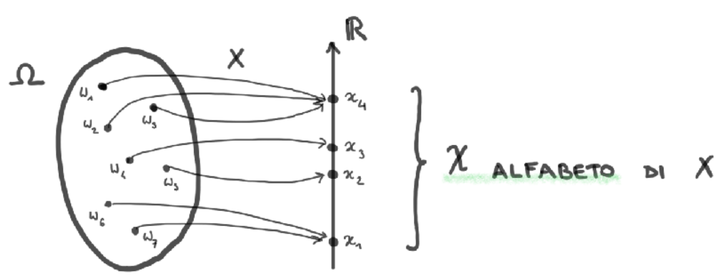

Figure 7.1: Vediamo qui l'Alfabeto di X, che è l'immagine di X

$$\mathcal{X} = X(\Omega) = \{x \in \mathbb{R} : \; X(\omega) = x \; per \; qualche \; \omega \in \Omega\}$$

Importante! $\mathcal{X}$ è **discreto** ed ho che

$$|\mathcal{X}| \leq |\Omega| \leq |\mathbb{N}|$$

7.1.1 Esempi

1. Sia $(\Omega, \mathbb{P}(\Omega), P)$ uno spazio di probabilità discreto. Sia E $\in \mathbb{P}(\Omega)$ un qualsiasi evento. Possiamo definire la variabile aleatoria $\mathbb{1}_E(\omega) : \Omega \to \mathbb{R}$

definita come:

$$\mathbb{1}_E(\omega) = \begin{cases} 1, & \text{se } \omega \in E \\ 0, & \text{altrimenti} \end{cases}$$

Questa variabile si chiama **indicatrice** ed ha alfabeto {0,1}.

2. Sia ora $\Omega = \{(d_1, d_2) : d_1, d_2 \in \{1, ..., 6\}\}$ lo spazio campionario relativo al lancio di una coppia di dadi, dove d_i indica il punteggio dell'i-esimo dado, ed i = 1,2.
   Vediamo qualche variabile aleatoria che si può creare.
   - $\omega = (d_1, d_2) \longmapsto X_1(\omega) = d_1$
     Mappo gli eventi verso la variabile aleatoria X_1 e questo mi dice di considerare solo il lancio del valore d_1. L'alfabeto sarà $X_1 = \{1, ..., 6\}$.
     Il discorso sulla mappatura vale anche per i prossimi esempi.
   - $\omega = (d_1, d_2) \longmapsto X_2(\omega) = d_2$ con alfabeto $X_2 = \{1, ..., 6\}$
   - $\omega = (d_1, d_2) \longmapsto X_3(\omega) = min\{d_1, d_2\}$ con alfabeto $X_3 = \{1, ..., 6\}$
   - $\omega = (d_1, d_2) \longmapsto X_4(\omega) = max\{d_1, d_2\}$ con alfabeto $X_4 = \{1, ..., 6\}$
   - $\omega = (d_1, d_2) \longmapsto X_5(\omega) = d_1 + d_2$ con alfabeto $X_5 = \{2, 3, ..., 12\}$

Osservazione Se vengono definite sullo stesso spazio di probabilità, diverse variabili aleatorie possono essere **combinate** attraverso **operazioni**, come ad esempio

$$X + Y, X - Y, X \cdot Y, X/Y, min\{X, Y\}, max\{X, Y\}$$

queste sono ancora variabili aleatorie, come nell'esempio 2, in cui avevamo X_3, X_4 e X_5.

7.2 Misura di probabilità indotta

Poichè X è un insieme discreto, basta assegnare una probabilità P^X ad **ogni singoletto** di X, compatibile con la probabilità P di $(\Omega, \mathbb{P}(\Omega), P)$ dove X è definita.

7.2.1 Definizione Probabilità Indotta

La definizione è questa:

$$P^X(x_k) = P(X^{-1}(x_k))\forall x_k \in X$$

che si può anche scrivere come:

$$P^X(x_k) = P(X = x_k)\forall x_k \in X$$

Cosa significa? Significa che la probabilità totale dei valori che fanno verificare x_k è **uguale** alla probabilità che si verifichi x_k dell'alfabeto.
Ciascuno dei valori dell'alfabeto si osserverà se **qualcuno degli esiti mappati in questo valore si verifica**.

7.2.2 Lemma e Dimostrazione

P^X è una misura di probabilità.

Dimostrazione Poichè X è discreto, basta verificare:

1. $P^X(x_k) \geq 0 \; \forall x \in X$
2. $\sum_{x_k \in X} P^X(x_k) = 1$

Verifico

1. $P^X(x_k) = P(X^{-1}(x_k))$ ma P è misura di probabilità e quindi positiva
2. Osserviamo il fatto che X sia una funzione: ciò implica che la funzione delle anti-immagini (ovvero gli esiti che poi vengono mappati nell'alfabeto)

   $$\{X^{-1}(X_k)\}_{x_k \in X}$$

   sia una **partizione** di Ω, cioè:

   - $X^{-1}(x_k) \cap X^{-1}(x_j) = \varnothing, \; \forall k \neq j$
   - $\bigcup_{x_k \in X} X^{-1}(x_k) = \Omega$

   Quindi otteniamo che:

   $$\begin{aligned} \sum_{x_k \in X} P^X(x_k) &= \sum_{x_k \in X} P(X^{-1}(x_k)) \\ &= P(\bigcup_{x_k \in X} X^{-1}(x_k)) \\ &= P(\Omega) = 1 \end{aligned}$$

   come volevasi dimostrare.

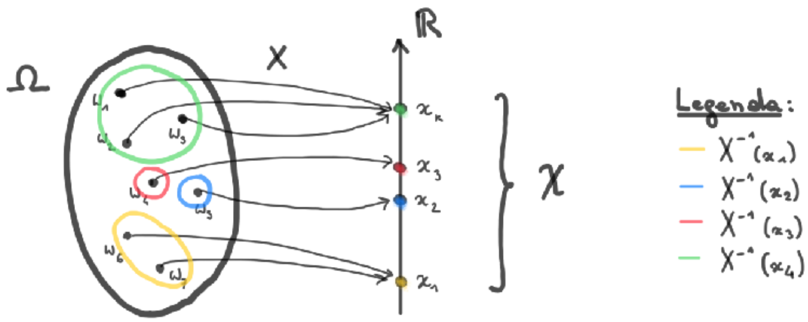

Figure 7.2: In questa figura vediamo chiaramente che le nuvole di probabilità sono distinte e disgiunte e si può vedere chiaramente ciò che è stato appena dimostrato.

7.3 Esempi e descrizione probabilistica

1. Ritorniamo all'esempio del lancio di 2 dadi, e consideriamo la variabile aleatoria $X_1 : \Omega \longmapsto \mathbb{R}$ con $(d_1, d_2) \mapsto d_1$
   Come assegniamo una probabilità sull'alfabeto $X_1 = \{1, ..., 6\}$? Vediamolo con una figura: La probabilità delle anti-immagini è:

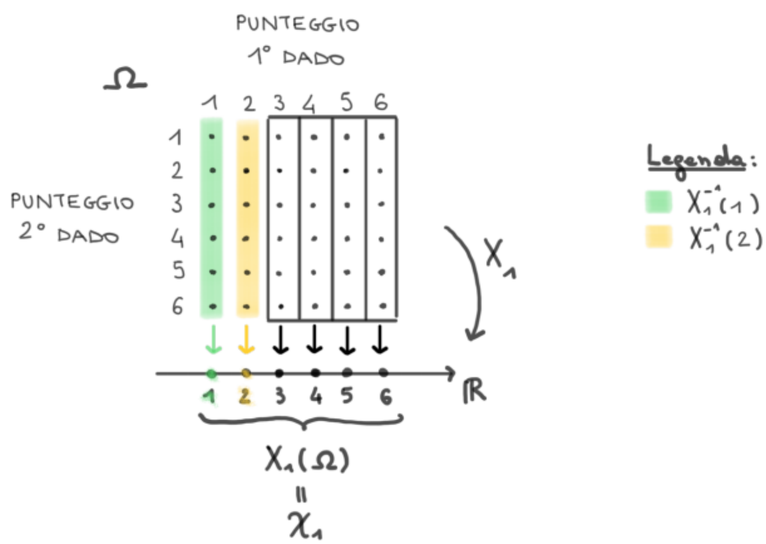

Figure 7.3: Vediamo qui che ogni lancio del primo dado, come lo abbiamo definito prima, porta alla rappresentazione del numero reale, e l'insieme dei valori dati da d_1 mi dà l'alfabeto X_1

$$P^{X_1}(1 \in X_1) = P(X = 1) = P(X_1^{-1}(1))$$
$$= P(\{(1,1), (1,2), ..., (1,6)\}) = \frac{1}{6}$$

$$P^{X_1}(2 \in X_1) = P(X = 2) = P(X_1^{-1}(2))$$
$$= P(\{(2,1), (2,2), ..., (2,6)\}) = \frac{1}{6}$$

Analogamente possiamo trovare

$$P^{X_1}(3) = P^{X_1}(4) = P^{X_1}(5) = P^{X_1}(6) = \frac{1}{6}$$

2. Sempre dall'esempio del lancio dei 2 dadi: consideriamo

$$X_3 : \Omega \to \mathbb{R}$$
$$(d_1, d_2) \mapsto min\{d_1, d_2\}$$

Come assegnamo una probabilità su $X_3 = \{1, ..., 6\}$?

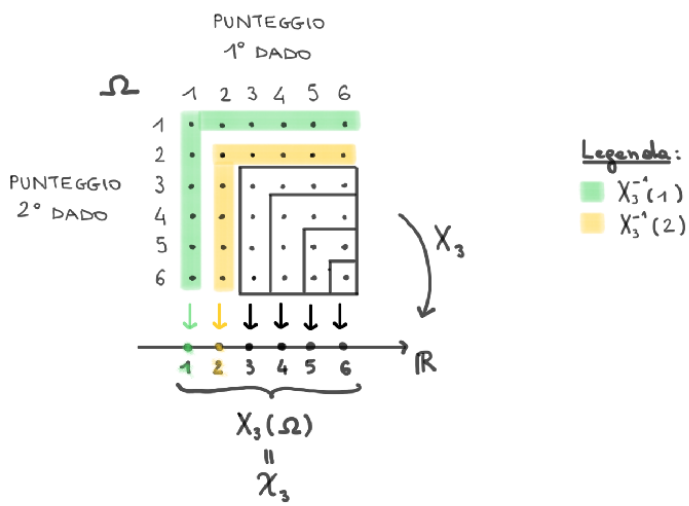

Figure 7.4:

$$P^{X_3}(1) = P(X_3^{-1}(1)) = P(X = 1) = P(\text{ parte verde }) = \frac{11}{36}$$
$$P^{X_3}(2) = P(X_3^{-1}(2)) = P(2) = P(\text{ parte gialla }) = \frac{1}{4}$$

e così via per tutti gli elementi dell'alfabeto X_3

7.3.1 Descrizione Probabilistica della variabile aleatoria

Se X è una variabile aleatoria discreta, per definirla probabilisticamente utilizziamo questa descrizione:

- Abbiamo X alfabeto **unito** a
- $p_x : X \to [0, 1]$
  $x_k \longmapsto p_x(x_k) = P(X = x_k)$ dove esistono due proprietà:
  1. $p_x(x_k) \geq 0, \forall x_k \in X$
  2. $\sum_{x_k \in X} p_x(x_k) = 1$

Identifichiamo p_x come **densità discreta di probabilità** (o **legge**) della variabile aleatoria X.

Esempio

Sia X una var. aleatoria discreta con alfabeto $X = \{0, \sqrt{2}, \pi\}$ e densità discreta

- $p_x(0) = \frac{2}{3}$
- $p_x(\sqrt{2}) = p_x(\pi) = \frac{1}{6}$

Calcoliamo $P(1 < X < 4)$. Abbiamo:

$$\begin{aligned}
P(1 < X < 4) &= P(X \in \{\sqrt{2}, \pi\}) && (7.1)\\
&= P(\{X = \sqrt{2}\} \cup \{X = \pi\}) && (7.2)\\
&= P(X = \sqrt{2} + P(X = \pi) && (7.3)\\
&= \frac{1}{3} && (7.4)
\end{aligned}$$

Chapter 8

Lezione 8

Ricapitolando possiamo definire una variabile aleatoria in due modi:

1. Come funzione $X : \Omega \to \mathbb{R}$ di cui si determinano alfabeto X e misura di probabilità indotta $P^x(x_k) = P(X^{-1}(x_k))$, per ogni $x_k \in X$, da cui si deduce la densità $p_x(x_k) = P^x(x_k)$, per ogni $x_k \in X$, che soddisfa positività e normalizzazione.

2. Si danno direttamente alfabeto X e densità $p_x(\cdot)$ che soddisfa positività e normalizzazione.

8.1 Funzione di una variabile aleatoria

$$\left.\begin{array}{ll} X : \Omega \to \mathbb{R}, & \text{variabile aleatoria} \\ g : \mathbb{R} \to \mathbb{R}, & \text{funzione} \end{array}\right\} \text{Y} = \text{g}oX : \Omega \to \mathbb{R} \text{ variabile aleatoria}$$

Caratterizziamo Y. L'alfabeto è $Y = \text{g}(X)$ e la densità discreta:

$$p_Y(y_l) = P(Y = y_l) = \sum_{x:g(x_k)=y_l} p_x(x_k), \text{ per ogni } y_l \in Y$$

Esempio

Sia X una v.al. con alfabeto $X = \{$-1,1,3$\}$ e densità discreta $p_x(-1) = p_x(3) = \frac{1}{4}$ e $p_x(1) = \frac{1}{2}$.
Caratterizziamo la v.al. Y=X^2: se

- g: $\mathbb{R} \to \mathbb{R}$
- g: x $\longmapsto x^2$

abbiamo Y = g(X).
Alfabeto: $Y = \text{g}(X) = \{$1,9$\}$

Densità:

$$p_y(1) = P(\{X = -1\} \cup \{X = 1\}) \tag{8.1}$$
$$= p_x(-1) + p_x(1) = \frac{1}{4} + \frac{1}{2} = \frac{3}{4} \tag{8.2}$$
$$P_y(9) = P(X = 3) = p_x(3) = \frac{1}{4} \tag{8.3}$$

8.1.1 Esercizi

1. Lancio una moneta e un dado. Se i risultati che ottengo hanno la stessa iniziale vinco 1 euro, altrimenti perdo 50 centesimi.
   Sia X la variabile aleatoria che descrive la vincita/perdita. Determinare alfabeto e densità discreta di X.

   - **Soluzione**: Alfabeto: $X = \{\frac{-1}{2}, 1\}$, corrispondono a perdita e vincita.

   - Densità discreta: Uno spazio campionario per il nostro gioco è

   $$\Omega = \{(a, i) : a \in \{T, C\}, i \in \{1, ..., 6\}\}$$

   Si ha $|\Omega| = 2 \cdot 6 = 12$. Calcoliamo la probabilità indotta, che prendiamo come densità di X.

   $$p_x(1) = P(X = 1) = P(\text{i risultati hanno la stessa iniziale}) \tag{8.4}$$
   $$= P(\{(T, 3), (C, 5)\} = \frac{1}{6} \tag{8.5}$$
   $$p_x(-\frac{1}{2} = 1 - p_x(1) = \frac{5}{6} \tag{8.6}$$

2. Estraggo tre carte da un mazzo di carte da poker. Vinco 1 euro per ogni carta di picche estratta.
   Sia X la variabile aleatori che descrive la vincita. Determinare alfabeto e densità discreta di X.

   - **Soluzione**: Alfabeto: $X = \{0,1,2,3\}$
   - Densità discreta:

   $$p_x(0) = P(X = 0) = P(\text{non pesco carte di picche}) \tag{8.7}$$
   $$= \frac{\binom{39}{3}}{\binom{52}{3}} = \frac{703}{1700} \tag{8.8}$$

   $$p_x(1) = P(1) = P(\text{pesco 1 carta di picche}) \tag{8.9}$$
   $$= \frac{13 \cdot \binom{39}{2}}{\binom{52}{3}} = \frac{741}{1700} \tag{8.10}$$

$$p_x(2) = P(2) = P(\text{pesco 2 carte di picche}) \tag{8.11}$$

$$= \frac{\binom{13}{2} \cdot 39}{\binom{52}{3}} = \frac{234}{1700} \tag{8.12}$$

$$p_x(3) = 1 - p_x(0) - p_x(1) - p_x(2) = \frac{11}{850}$$

8.2 Valor Medio (o Atteso)

Sia X v.al. con alfabeto X e densità discreta p_x. Il valor medio di X è il numero reale:

$$E(x) = \sum_{x_k \in X} x_k \cdot p_x(x_k)$$

Osservazione Se $|X| < \infty$, allora il v.medio è una somma **finita** ed è **sempre** ben definito. Se $|X| = +\infty$, allora il v.medio è una serie e **può non esistere** se la serie **non converge**.

Esempi

1. Sia X una v.al. con alfabeto $X = \{-7, 0, \pi, 4\}$ e densità $p_x(-7) = \frac{1}{2}$ e $p_x(0) = p_x(\pi) = p_x(4) = \frac{1}{6}$. Otteniamo:

$$E(X) = (-7) \cdot \frac{1}{2} + (4 + \pi) \cdot \frac{1}{6} \approx -2.31$$

2. Consideriamo uno spazio di probabilità $(\Omega, \mathbb{P}(\Omega), P)$ ed un evento $E \in \mathbb{P}(\Omega)$. Definiamo la v.al.

$$X(\omega) = \mathbb{1}_E(\omega) = \begin{cases} 1, & \text{se } \omega \in E \\ 0, & \text{se } \omega \notin E \end{cases}$$

Allora si ha:

$$E(X) = P(X = 1) = P(\{\omega : X(\omega) = 1\}) = P(E)$$

8.2.1 Significato del Valor Medio

Indice di centralità $\Longleftrightarrow$ Baricentro della distribuzione.
abbiamo dei punti $\{x_1, ..., x_n\}$ con masse $p_x(x_1), ..., p_x(x_n)$ e cerchiamo il baricentro $a \in \mathbb{R}$ della nuvola di punti. Cerchiamo il punto dove la risultante dei

momenti è nulla (**equilibrio**):

$$\sum_{x_k}(x_k - a) \cdot p_x(x_k) = 0 \tag{8.13}$$

$$\iff \sum_{x_k} x_k \cdot p_x(x_k) = a \cdot \underbrace{\sum_{x_k} p_x(x_k)}_{=1} \tag{8.14}$$

$$\iff a = E(x) \tag{8.15}$$

8.2.2 Teorema Fondamentale del valor medio

Sia $X : \Omega \to \mathbb{R}$ una v.al. discreta, allora si ha:

$$E(X) = \sum_{\omega \in \Omega} X(\omega) \cdot P(\{\omega\})$$

Dimostrazione

Calcoliamo:

$$E(X) = \sum_{x_k \in X} x_k \cdot p_x(x_k) = \sum_{x_k \in X} x_k \cdot P(X^{-1}(x_k))$$

$$\text{(definizione di antiimmagine)} = \sum_{x_k \in X} x_k \cdot P(\underbrace{\{\omega : X(\omega) = x_k\})}_{\substack{\textbf{Insieme discreto}\text{:ottengo la sua} \\ \text{probabilità} \\ \text{sommando le probabilità degli esiti} \\ \text{che ci appartengono}}} \tag{8.16}$$

$$= \sum_{x_k \in X} x_k \cdot \sum_{\omega : X(\omega) = x_k} \cdot P(\{\omega\}) \tag{8.17}$$

$$= \underbrace{\sum_{x_k \in X} x_k \cdot \sum_{\omega : X(\omega) = x_k}}_{\text{vedi (1) sotto}} \cdot X(\omega) \cdot P(\{\omega\}) \tag{8.18}$$

$$= \sum_{\omega \in \Omega} X(\omega) \cdot P(\{\omega\}) \tag{8.19}$$

(1): Le due somme descrivono l'insieme:

$$\bigcup_{x_k \in X} \{\omega : X(\omega) = x_k\} = \Omega$$

$$\text{dove } X(\omega) \longleftrightarrow \{X^{-1}(x_k)\} \text{ (è una partizione)}$$

Chapter 9

Lezione 9

9.1 Proprietà del Valor Medio

9.1.1 Linearità

Descriviamo la linearità attraverso due osservazioni:

1. Sia X una v.al., $a \in \mathbb{R}$, allora $E(aX) = aE(X)$. Come lo vedo?
   aX è la v.al. $\omega \mapsto aX(\omega)$, allora per il teorema fondamentale otteniamo:

$$E(aX) = \sum_{\omega \in \Omega} aX(\omega)P(\{\omega\}) \tag{9.1}$$

$$= a \sum_{\omega \in \Omega} X(\omega)P(\{\omega\}) = aE(X) \tag{9.2}$$

2. Siano X,Y v.al. definite sullo stesso spazio campionario Ω. Allora E(X+Y) = E(X) + E(Y). Come lo vedo?
   X + Y è la v.al. $\omega \mapsto X(\omega) + Y(\omega)$, allora per il teorema fondamentale otteniamo:

$$E(X+Y) = \sum_{\omega \in \Omega} (X(\omega) + Y(\omega))P(\{\omega\}) \tag{9.3}$$

$$= \sum_{\omega \in \Omega} X(\omega) \cdot P(\{\omega\}) + Y(\omega) \cdot P(\{\omega\}) \tag{9.4}$$

$$= E(X) + E(Y) \tag{9.5}$$

Potevo definire la proprietà 2. con la definizione del valore medio?

$$E(X+Y) = \sum_{x_k \in X} \sum_{y_l \in Y} (x_k + y_l) \cdot p_{X+Y}(x_k + y_l)$$

La scrittura p_{X+Y} avrà senso dalla settimana 8.
Mettendo insieme le proprietà 1. e 2. otteniamo:

$$E(aX + bY) = aE(X) + bE(Y)$$

9.1.2 Positività

Sia X una v.al. positiva (cioè $X \in \mathbb{R}_+$), allora $E(X) \geq 0$

9.1.3 Monotonia

Siano X, Y: $\Omega \to \mathbb{R}$ v.al. tali che $X \geq Y$. Allora si può dire che

$$E(X) \geq E(Y)$$

Come lo vedo?
$X \geq Y \iff X - Y \geq 0$.
Per la positività si ha che $E(X - Y) \geq 0$.
Per la linearità invece si ottiene $E(X) - E(Y) \geq 0$ da cui concludo la dimostrazione che $E(X) \geq E(Y)$.

9.1.4 Limiti Inferiore e Superiore

Sia X v.al. con alfabeto X e siano $\underline{x} = inf X$ e $\overline{x} = supX$. Allora si ha:

$$\underline{x} \leq E(X) \leq \overline{x}$$

Osservazione: Se $b \in \mathbb{R}$ allora E(b) = b.
Possiamo vedere $b \in \mathbb{R}$ come una v.al. **costante**, cioè:
X v.al. con alfabeto X = {b} e densità discreta $p_x(b) = 1$ e$p_x(i) = 0 \; \forall i \neq b$.
Segue dalla definizione di valor medio che E(X) = b.

9.1.5 Teorema (valor medio di funzioni di v.al.)

Siano $X : \Omega \to \mathbb{R}$ v.al. e $g : \mathbb{R} \to \mathbb{R}$ funzione. Si ha:

$$E(g(X)) = \sum_{x_k \in X} g(x_k) p_x(x_k)$$

Esempio

Sia X v.al. con alfabeto X={-1,1,3} e densità discreta $p_x(3) = p_x(-1) = \frac{1}{4}$ e $p_x(1) = \frac{1}{2}$. Sia inoltre $Y = X^2$. Calcoliamo E(Y).

- **Metodo 1:** Caratterizzo Y (alfabeto + densità) e uso la definizione di valor medio.
  Abbiamo Y = {1,9} e $p_y(1) = \frac{3}{4}$ e $p_y(3) = \frac{1}{4}$.
  Quindi
  $$E(Y) = 1 \cdot \frac{3}{4} + 9 \cdot \frac{1}{4} = 3$$
- **Metodo 2:** Applico direttamente il teorema:
  $$E(Y) = (-1)^2 \cdot \frac{1}{4} + 1^2 \cdot \frac{1}{2} + 3^2 \cdot \frac{1}{4} = 3$$

9.1.6 Esercizi

1. Un'urna contiene 8 palline nere e 6 bianche. Si fanno due estrazioni senza reinserimento. Per ogni pallina nera estratta si vince 1 euro. Per ogni pallina bianca estratta si perde 1 euro. Sia X la variabile aleatoria che descrive la vincita/perdita del gioco. Calcolare E(X).

   Soluzione: Alfabeto: $X = \{$-2, 0, 2$\}$ (corrisponde a: estraggo 2 bianche, estraggo 1 nera e 1 bianca, estraggo 2 nere).
   Densità:

$$p_x(-2) = \text{P(estraggo 2 bianche)} \tag{9.6}$$

$$= \frac{\binom{6}{2}}{\binom{14}{2}} = \frac{15}{91} \tag{9.7}$$

$$p_x(0) = \frac{8 \cdot 6}{\binom{14}{2}} = \frac{48}{91} \tag{9.8}$$

$$p_x(2) = 1 - p_x(0) - p_x(-2) = \frac{4}{13} \tag{9.9}$$

   Media: $E(X) = (-2) \cdot \frac{15}{91} + 2 \cdot \frac{4}{13} = \frac{2}{7} \approx 0.286$

2. Due dadi sono truccati in modo che la probabilità di ottenere 6 sia il doppio di quella di ottenere ogni altro punteggio. Qual è la media del punteggio ottenuto lanciando i due dadi?

   Soluzione: Siano X_1 e X_2 le v.al. che corrispondono al punteggio del primo e del secondo dado. Le 2 v.al. hanno lo stesso alfabeto X=$\{$1,...,6$\}$ e la stessa densità discreta che ora determiniamo:
   si ha $p_{xi}(6) = 2p$
   $p_{xi}(j) = p$ per ogni j = 1,...,5
   da cui ricavo $2p + 5p = 1 \iff p = \frac{1}{7}$.
   Quindi la densità è

$$p_{xi}(1) = ... = p_{xi}(5) = \frac{1}{7} \text{ e } p_{xi}(6) = \frac{2}{7}$$

   Ora sia $Y = X_1 + X_2$ la v.al. che corrisponde alla somma dei due punteggi. Abbiamo:

$$E(Y) = E(X_1) + E(X_2) \tag{9.10}$$

$$= (1 + 2 + 3 + 4 + 5) \cdot \frac{1}{7} + 6 \cdot \frac{2}{7} + \frac{15}{7} + \frac{12}{7} \tag{9.11}$$

$$= \frac{54}{7} \approx 7.7 \tag{9.12}$$

   Se i dadi non fossero stati truccati si avrebbe E(Y) = 7.

9.1.7 Varianza

Sia X v.al. con alfabeto X e densità discreta p_x. La varianza è il **numero reale positivo**.

$$Var(X) = \sum_{x_k \in X} (x_k - E(X))^2 p_x(x_k)$$

Esempio

Consideriamo $X_1 \in \{-1, \frac{1}{4}, \frac{3}{4}\}$ con densità discreta $p_{x_1}(-1) = p_{x_1}(\frac{1}{4}) = p_{x_1}(\frac{3}{4}) = \frac{1}{3}$ e $X_2 \in \{-10, 10\}$ con densità discreta $p_{x_2}(-10) = p_{x_2}(10) = \frac{1}{2}$.
Abbiamo $E(X_1) = E(X_2) = 0$, ma

$$Var(X_1) = \frac{1}{3} \cdot [(-1-0)^2 + (\frac{1}{4} - 0)^2 + (\frac{3}{4} - 0)^2] \approx 0.524$$

$$Var(X_2) = \frac{1}{2} \cdot [(-10-0)^2 + (10-0)^2] = 100$$

I valori di X_2 sono tanto dispersi, cioè **lontani dalla media**.

Proprietà della Varianza

1. Var(X) = $E[(X-E(X))^2]$ è la media di una funzione della v.al. X; usiamo g(x) = $(x - E(x))^2$

2. Var(X) ≥ 0 e, in particolare, Var(X) = 0 $\iff X \equiv$ costante
   Come lo vedo?

$$Var(X) = 0 \iff \sum_{x_k \in X} (x_k - E(X))^2 \cdot \underbrace{p_x(x_k)}_{>0} = 0 \quad (9.13)$$

$$\iff x_k - E(X) = 0, \forall x_k \in X \quad (9.14)$$

$$\iff x_k = E(X), \forall x_k \in X \quad (9.15)$$

3. Var(aX) = a^2 Var(X), a $\in \mathbb{R}$
   Come lo vedo?

$$Var(aX) = E[(aX - aE(x))^2] \quad (9.16)$$

$$\text{per linearità valor medio} = E[(aX - aE(x))^2] \quad (9.17)$$

$$\text{per linearità valor medio} = a^2 E[(X - E(X))^2] = a^2 Var(X) \quad (9.18)$$

4. Var(X+c) = Var(X), $c \in \mathbb{R}$

Osservazione Importante: 3 e 4 mostrano che la varianza **NON E' LINEARE!**

5. Var(X) = $E(X^2) - E(X)^2$
   Come lo vedo?

$$Var(X) = E[(X - E(X))^2] \tag{9.19}$$
$$= E[X^2 - 2XE(X) + E(X)^2] \tag{9.20}$$
$$(1) = E(X^2) - E(X)^2 \tag{9.21}$$

   (1) Indica che tale risoluzione è data da $E(X) \in \mathbb{R}$ + linearità del valor medio.

Esercizio

L'urna 1 ha composizione: 1 pallina dorata, 4 palline verdi, 15 palline bianche. L'urna 2 invece contiene: 4 palline verdi e 25 palline bianche. Una pallina a caso viene spostata dall'urna 1 all'urna 2, quindi mi viene chiesto di estrarre una pallina dall'urna 2.
Se estraggo la pallina dorata vinco 50 euro, se estraggo una pallina verde perdo 1 euro, altrimenti non vinco e non perdo.
Sia X la v.al. che corrisponde alla vincita/perdita. Si calcoli la varianza di X.

Soluzione: Alfabeto X = {-1, 0, 50}
Densità discreta. Considero gli eventi $T_i =$ "trasferisco una pallina i" con $i = O, V, B$, dove O sta per Oro, V per Verde e B per bianca,
ed $E_i =$ "estraggo una pallina i" con $i = O, V, B$.
Otteniamo:

$$P(X = -1) = P(E_V) = P(E_V|T_B) \cdot P(T_B) + P(E_V|T_V) \cdot P(T_V) \tag{9.22}$$
$$+ P(E_V|T_O) \cdot P(T_O) \tag{9.23}$$
$$= \frac{4}{30} \cdot \frac{15}{20} + \frac{5}{30} \cdot \frac{4}{20} + \frac{4}{30} \cdot \frac{1}{20} = \frac{7}{50}$$

$$P(X = 50) = P(E_O) = P(T_O \cap E_O) = P(E_O|T_O) \cdot P(T_O) \tag{9.24}$$
$$= \frac{1}{30} \cdot \frac{1}{20} = \frac{1}{600} \tag{9.25}$$

$$P(X = 0) = 1 - P(X = -1) - P(X = 50) = \frac{103}{120} \tag{9.26}$$

Varianza:

$$Var(X) = E(X^2) - E(X)^2 \tag{9.27}$$
$$= (-1)^2 \cdot \frac{7}{50} + (50)^2 \cdot \frac{1}{600} - [(-1) \cdot \frac{7}{50} + 50 \cdot \frac{1}{600}]^2 \tag{9.28}$$
$$\approx 4.3 \tag{9.29}$$

Chapter 10

Lezione 11

La lezione 10 è stata usata per esercizi di ripasso.

10.1 Variabili Aleatorie Discrete Notevoli

La densità discreta contiene la descrizione probabilistica di una variabile aleatoria. Quindi due v.al. X e Y con la stessa densità sono **probabilisticamente indistinguibili** (o equidistribuite o identicamente distribuite), nel senso che $P(X \in B) = P(Y \in B)$ per ogni B sottoinsieme dell'alfabeto.

Attenzione: ciò **non significa che** X = Y.
Assegnata una densita *p(.)* si può allora associare la famiglia delle v.al. X che hanno densità $p_x(.) = p(.)$

10.1.1 Caso di Alfabeto Finito

Variabili Aleatorie di Bernoulli

$$X \sim B_e(p) \text{ con } p \in [0,1]$$

$$\text{se ha: } \begin{cases} \text{alfabeto,} & X = \{0,1\} \\ \text{densità discreta,} & p_x(1) = p, p_x(0) = 1 - p \end{cases}$$

Inoltre, *E(X) = p* e *Var(X) = p(1-p)*

Esempi

1. X v.al. che assume valore 1 se lanciando una moneta ottengo testa e assume valore 0 altrimenti. Si ha $X \sim B_e(\frac{1}{2})$
2. Y v.al. che assume valore 1 se lanciando un dado ottengo un numero pari e assume valore 0 altrimenti. Si ha $Y \sim B_e(\frac{1}{2})$
   In questi due casi X e Y sono probabilisticamente indistinguibili ma **X è diversa da Y**

3. Sia (Ω, F, P) uno spazio di probabilità e $E \in F$ un evento. Definiamo la v.al.: $\mathbb{1}_E(\omega) = \begin{cases} 1, & \text{se } \omega \in E \\ 0, & \text{altrimenti} \end{cases}$

   Si ha $\mathbb{1}_E \sim B_e(P(E))$. Infatti otteniamo:

$$p = P(\mathbb{1}_E = 1) = P(\{\omega \in \Omega : \omega \in E\} = P(E)$$

10.1.2 Schema di Bernoulli (o schema a Prove Indipendenti)

Si divide in due pari: quella del **contesto sperimentale** e quella del **modello probabilistico**, che sono identiche tranne che nella forma in cui sono espresse.

Contesto Sperimentale

- Un certo numero $n \geq 1$ di prove **identiche** effettuate in sequenza
- Ogni prova ha **2 esiti** possibili codificati con **0 e 1**
- Il risultato di ciascuna prova **non influenza** il risultato delle altre

Modello probabilistico

- Consideriamo $X_1, ..., X_n$ v.al. **identicamente distribuite**
- $X_i \sim B_e(p)$ (con i = 1,...,n) e p è la probabilità di ottenere 1
- Gli eventi $\{X_1 = 1\}, ..., \{X_n = 1\}$ sono tra loro **indipendenti**

10.1.3 Esempi

1. Una rete è composta da 150 terminali connessi ad un server. Controllo quali terminali sono pronti per trasmettere un lavoro. Per i = 1,...,150 si ha:

$$X_i = \begin{cases} 1, & \text{se l'i-esimo terminale è pronto per trasmettere un lavoro} \\ 0, & \text{altrimenti} \end{cases}$$

2. Controllo di qualità in una linea di produzione di chip. Ogni giorno ne vengono testati 1000. Per i = 1,...,1000 si ha:

$$X_i = \begin{cases} 1, & \text{se l'i-esimo chip è difettoso} \\ 0, & \text{altrimenti} \end{cases}$$

3. Errori in una trasmissione digitale di 1200000 bit. Per i = 1,...,1200000 si ha:

$$X_i = \begin{cases} 1, & \text{se l'i-esimo bit è stato trasmesso errato} \\ 0, & \text{altrimenti} \end{cases}$$

10.2 Variabili Aleatorie Binomiali

$$X \sim Bin(n,p) \text{ con } n \in \mathbb{N} = \{1,2,...\} \text{ e } p \in [0,1]$$

$$\text{se ha: } \begin{cases} \text{alfabeto,} & X = \{0,1,...,n\} \\ \text{dens. discr.,} & p_x(k) = \binom{n}{k} \cdot p^k(1-p)^{n-k}, k \in X \end{cases}$$

Inoltre, *E(X) = np* e *Var(X) = np(1-p)*.
La variabile aleatoria binomiale conta il numero di successi (cioè di numeri "1") in uno schema di Bernoulli con n prove indipendenti, dove la probabilità di successo è p.

Contesto: le v.al. $X_i \sim B_e(p)$ (i = 1,...,n) sono gli esiti delle n prove e la loro somma

$$X = \sum_{i=1}^{n} X_i \sim Bin(n,p)$$

è il numero di successi ottenuti nelle n prove.
Cruciale: indipendenza degli eventi $\{X_1 = 1\}, ..., \{X_n = 1\}$

Interpretazione densità

$$\begin{aligned} p_x(k) &= P(\text{ottenere k successi in n prove}) && (10.1) \\ &= P(\text{ottenere stringa binaria di n cifre con k cifre 1} && (10.2) \\ &= \underbrace{\binom{n}{k}}_{\text{num. stringhe di lunghezza n con k cifre 1}} \cdot \underbrace{p^k}_{\text{prob. di avere k cifre 1}} \cdot \underbrace{(1-p)^{n-k}}_{\text{prob. di avere n-k cifre 0}} && (10.3) \end{aligned}$$

$p^k \cdot (1-p)^{n-k}$ indicano anche **equidistribuzione + indipendenza**.

10.2.1 Esempio

Supponiamo di fissare n = 5, k = 1 e di voler calcolare la probabilità della stringa '00010'. Se $X_i \sim B_e(p)$ si ottiene:

$$\begin{aligned} P(00010) &= P(X_1 = X_2 = X_3 = 0, X_4 = 1, X_5 = 0) && (10.4) \\ \xrightarrow{indip.} &= P(X_1 = 0)P(X_2 = 0)P(X_3 = 0)P(X_4 = 1)P(X_5 = 0) && (10.5) \\ \xrightarrow{equidistr.} &= (1-p)^4 \cdot p && (10.6) \end{aligned}$$

Analogamente, possiamo calcolare:

$$\begin{aligned} P(01000) &= P(X_1 = 0)P(X_2 = 1)P(X_3 = 0)P(X_4 = 0)P(X_5 = 0) && (10.7) \\ &= (1-p)^4 \cdot p && (10.8) \end{aligned}$$

Per indipendenza ed equidistribuzione, tutte le stringhe di lunghezza 5 con una sola cifra 1 sono equiprobabili.

10.2.2 Esercizio

E' più facile ottenere almeno un 6 lanciando 4 volte un dado oppure ottenere almeno un doppio 6 lanciando 24 volte una coppia di dadi?

Soluzione

- **Lancio singolo:** Costruiamo uno schema di Bernoulli. Per ogni i = 1,...,4 definiamo le v.al.

  $$X_i = \begin{cases} 1, & \text{se all'i-esimo lancio ottengo 6} \\ 0, & \text{altrimenti} \end{cases}$$

  Si ha $X_i \sim B_e(\frac{1}{6})$. Inoltre gli eventi $\{X_1 = 1\}, ..., \{X_4 = 1\}$ sono indipendenti. Allora il numero di punteggi 6 ottenuti nei 4 lanci è:

  $$X = \sum_{i=1}^{4} X_i \sim Bin(4, \frac{1}{6})$$

  Calcoliamo:

  $$P(X \geq 1) = 1 - P(X = 0) = 1 - (\frac{5}{6})^4 \approx 0.518$$

- **Lancio doppio:** Costruiamo uno schema di Bernoulli. Per i = 1,...,24 definiamo le v.al.

  $$Y_i = \begin{cases} 1, & \text{se all'i-esimo lancio ottengo una coppia di 6} \\ 0, & \text{altrimenti} \end{cases}$$

  Si ha $Y_i \sim B_e(\frac{1}{36})$. Inoltre gli eventi $\{Y_1 = 1\}, ..., \{Y_{24} = 1\}$ sono indipendenti. Allora il numero di volte che ottengo un doppio 6 nei 24 lanci è:

  $$Y = \sum_{i=1}^{24} Y_i \sim Bin(24. \frac{1}{36})$$

  Calcoliamo:

  $$P(Y \geq 1) = 1 - P(Y = 0) = 1 - (\frac{35}{36})^{24} \approx 0.49$$

  Quinidi è più facile ottenere almeno un 6 lanciando un dado.