Calcolo Probabilità e Statistica
Prof. Aniello Buonocore

Antonio Garofalo

March 2021

Indice

1 Introduction

L'obiettivo specifico di apprendimento dell'insegnamento è quello dell'acquisizione dei principi teorici riguardanti il calcolo delle probabilità e di alcune metodologie della statistica sia descrittiva che inferenziale. In particolare, si intende rendere consapevole lo studente del fatto che gli assiomi rivestono il ruolo di formalizzare alcune idee-forza, più o meno intuitive e naturali e che strumenti propri dell'analisi matematica, dell'algebra e della logica assumono particolare importanza nella determinazione dei risultati e nella coerenza dell'impianto assiomatico. Un ulteriore obiettivo è quello di fornire un'iniziale indicazione di come i risultati teorici del Calcolo delle Probabilità trovino naturale e piena applicazione in Statistica Matematica.

- Informazioni generali: ?
- Orario di ricevimento: Venerdì ORE: 17.00
- Materiale didattico:
  - Calcolo delle probabilità - Sheldon Ross (Apogeo Editore - Education)[1]
- Modalitià d'esame: Orale + esercizi svolti

2 Lezione 1

2.1 Antico gioco della ZARA

La **zara** è un gioco d'azzardo in uso nel Medioevo.
Si gioca con tre dadi: a turno ogni giocatore chiama un numero da 3 a 18, quindi getta i dadi. Vince chi per primo ottiene il punteggio pari al numero chiamato.

- 3 dadi (onesti).
- danno luogo a 3 punteggi da 1 a 6.
- $S = P_1 + P_2 + P_3$
  - $(1,1,1) = 3$
  - $(6,6,6) = 18$

L'insieme dei possibili valori assumibili (con probabilità non nulla) da S, è l'insieme: {3,4,...,17,18}

La probabiltà che S assume valori >18, <3 è nulla (0%).

$$\mathrm{P}_{(S=10)} > 0 = \frac{\#\text{delle terne con somma 10}}{\#\text{delle terne}} \tag{1}$$

La probabilità è sempre compresa tra 0 e 1 ([0,1]).

2.2 Problema del contare

Principio fondamentale

Se una procedura di scelta si può decomporre in r procedure di scelta allora il numero delle scelte è uguale al prodotto del numero delle scelte di ciascuna sottoprocedura.

$N = N_1 \cdot N_2 \cdot ... \cdot N_r$

Esempio: Negozio di abbigliamento vende camice di una casa produttrice differenti per:

- Colore 5
- Taglia 8
- Foggia 3
- Collo 2

Quante etichette diverse devono essere stampate? Per le semplici proprietà del **Calcolo combinatorio**:
$N = 5 \cdot 8 \cdot 3 \cdot 2$
$A = \{e_1, e_2, ..., e_n\}$ k elementi da A, $k, n \in \mathbf{N}$

2.2.1 Criteri

I criteri imprtanti da tenere in considerazione sono:

- Semplice con ripetizione (o senza).
- Bisogna rispettare l'ordine con il quale sussegono gli elementi (o non bisogna).

Esempi visti a lezione:

- Calcio tris
- Terni al lotto
- Schedina totocalcio

Esempio schedina totocalcio:
$A = \{1, X, 2\}, n = 3, k = 14$
prese 2 schedine, con stesso numero di 1, X e 2 esse sono considerate diverse.

Esempio corsa tris:
$A = \{1, 2, ..., 9\}, n = 9, k = 3$
si consideri la terna (1,3,5), si scommette sulla posizione del podio (risp. primo, secondo e terzo posto) dei cavalli 1,3 e 5. non sono valide:

- $(1, 1, 5)$ - non si può piazzare un cavallo in 2 posizioni diverse.
- $(1, 3, 5) \neq (1, 5, 3)$ - due tris con gli stessi elementi sono considerati diversi (l'ordine è importante).

Esempio terni al lotto:
$A = \{1, 2, ..., 90\}, n = 9, k = 3$
in un terno a lotto la terna (1,1,5) non è valida, ma sono valide le terne (1,54,89)=(54,1,89) ovvero senza contare l'ordine.

2.3 Combinazione e Disposizione

Introduciamo due concetti fondamentali la **Combinazione** che affronteremo meglio nella lezione 2 e **Disposizione**.
Definaimo in inoltre:

- $R, \overline{R}$ rappresenta se si considerano le **R**ipetizione o senza.
- $O, \overline{O}$ rappresenta se si considerano le **O**rdine o senza.

Definiamo

- $D^{(r)}_{n,k}$ - Disposizione (insieme di ordine n con lunghezza della procedura).
- $C^{(r)}_{n,k}$ - Combinazione (insieme di ordine n con lunghezza della procedura).

Ovviamente, senza (r) apice si intende senza ripetizioni (**SEMPLICI**).

2.3.1 Casi di disposizione e risoluzione del problema ZARA

$[R, O]$ **- Disposizione con ripetizione**
Consideriamo quindi di prendere una decomposizione k, considerando ripetizioni e ordine, concluderemo facilmente che il numero di decomposizione è n per ogni componente da 1...k:
$D^{(r)}_{n,k} = n^k$

$[\overline{R}, O]$ **- Disposizione semplice**
Consideriamo quindi di prendere una decomposizione k, senza considerare ripetizioni ma considerando l'ordine, concluderemo che:
$D_{n,k} = n \cdot (n-1) \cdot ... \cdot (n-(k-1))$ è possibile scrivere (per l'ultima parte della formula) $(n-k+1)$.

si potrebbe abbreviare con $\frac{n!}{(n-k)!}$ ovvero il rapporto tra la cardinalità dell' insieme di scelta e la differenza tra tale cardinalità e la lunghezza di scelta (k).

Caso particolare: $n = k$
$D_{n,k} = n!$
In questo caso avremo $P_n = n!$
P = Permutazioni

Tornando al problema della **ZARA**:
#delle terne $= 6 \cdot 6 \cdot 6 = 6^3 = 216$ $(D^{(r)}_{6,3})$
per quanto riguarda il **numeratore** dobbiamo analizza:

- $6 + 3 + 6 \longrightarrow 6 = P_3$
- $6 + 2 + 2 \longrightarrow 3$
- $5 + 4 + 1 \longrightarrow 6 = P_3$
- $5 + 3 + 2 \longrightarrow 6 = P_3$

- $4 + 4 + 2 \longrightarrow 3$
- $4 + 3 + 3 \longrightarrow 3$

Il risultato è 27.

Nota bene si noti che le permutazioni con lo almeno due elementi uguali il totale delle permutazioni è $\frac{P_3}{P_2}$, ciò spiega che una terna con due componenti uguali ha 3 permutazioni in meno.
Il problema fu risolto da Cosimo Minore dei Medici (ft. Galielo Galieli).

3 Lezione 2

Consideriamo il seguente problema:
una parola, quanti anagrammi (anche senza senso logico) possiede?

$$"casa" \gg ca_1sa_2 \gg 4! \tag{1}$$

i pedici alla lettera "a" sono un modo per "Distinguere" le parole che possiamo creare con le permutazioni, se eliminassimo i pedici scopriremo che ci sono coppie di parole uguali.

$$"casata" \gg ca_1sa_2ta_3 \gg \frac{6!}{3! \cdot 1! \cdot 1! \cdot 1!} \tag{2}$$

Nota bene 3! = 6, 6 infatti è il numero totale di anagrammi uguali per ogni anagramma unico (3 infatti è il numero di ripetizioni di a).

Queste sono considerate: **Permutazioni con ripetizione** con $k > n$.

3.1 Permutazioni con ripetizione

Le permutazioni con ripetizione (${}_{k_1,\dots k_m}P_n^{(r)}$) hanno un semplice vincolo:

$$k_1 \dots k_m : k_1 + k_2 + \dots + k_m = n \tag{1}$$

quindi:

$$\sum_{i=1}^{m} k_i = n \tag{2}$$

il risultato di una permutazione con ripetizione sarà:

$$\frac{n!}{k_1! \cdot \dots \cdot k_m!} \tag{3}$$

questo da origine al cosiddetto "**Coefficiene multinomiale**":

$$\binom{n}{k_1 \dots k_m} \tag{4}$$

Esempio con la parola "statistica" nella parola statistica abbiamo 10 lettere, alcune di esse hanno delle ripetizioni (i.e t, s,i e c), se facciamo la somma di tutte le ripetizioni di ogni lettera otteremo 10.

Caso particolare quando gli elementi che si ripetono sono solo 2, la forma dell'equazione 3 può essere descritta nel seguente modo:

$$_{k_1,k_2}P_n^{(r)} =_k P_n^{(r)} \tag{5}$$

ciò è riassumibile in questo modo: $\binom{n}{k,n-k}$.
ragion per cui:

$$\frac{n!}{k! \cdot (n-k)!} \tag{6}$$

questo ci porta a definire che il "**Coefficiente binomiale**" non è altro che un caso particolare di coefficiente multinomiale, definito in questo modo:

$$\binom{n}{k} \tag{7}$$

3.2 Combinazione semplice

Una combinazione semplice in particolare viene definita in questo modo $[\overline{R}, \overline{O}]$ ovvero, senza tener conto delle ripetizioni e senza tener conto (essendo combinazione e non disposizione) dell'ordine.

Definiamo quindi: $C_{n,k}$.

$[\overline{R}, \overline{O}]$ - Combinazione semplice

$$C_{n,k} = \frac{D_{n,k}}{P_k} \tag{1}$$

Al numeratore avremo le disposizioni semplici di un insieme n su una lunghezza k mentre al denominatore le permutazioni di lunghezza k.
quindi:

$$\frac{n!}{k!(n-k)!} \tag{2}$$

Ma la formula 2 corrispone proprio a $\binom{n}{k}$ ovvero al coefficiente binomiale.

Cosa hanno in comune le combinazioni semplici con le permutazioni con ripetizione?

3.2.1 Combinazione o Permutazione?

Consideriamo le prime 6 lettere dell'alfabeto (a,b,c,d,e,f):

- {A,P} - A=Assente e P=Presente.
- n=6.
- k=4.

Parola	Verso	a b c d e f
adef	»	P A A P P P
abef	«	P P A A P P

Tabella 3.1: Schema di esempio

Nella Tabella 3.1 si noti come è possibile attraverso questo schema passare da una prima selezione ad uno schema predisposto (una disposizione) e viceversa.

Concludiamo:

$$ {}_{4,2}P_6^{(r)} = C_{6,4} \tag{3}$$

questa uguaglianza definisce un equipotenza in termini insiemistici dato che:

$$ \binom{6}{4,2} = \binom{6}{4} \tag{4}$$

3.3 Combinazione con ripetizione

La combinazione con ripetizione ($C_{n,k}^{(r)}$) si tratta di una selezione di questo tipo $[R, \overline{O}]$, ovvero consideriamo le ripetizioni ma non l'ordine.

$[R, \overline{O}]$ - **Combinazione con ripetizione**
Riprendiamo lo schema della tabella 3.2.1 possiamo fare una modifca alla tabella aggiungendo una sezione finale che inserisca P aggiuntive per ogni carattere che si ripete ed A per finire/marcare la fine del conteggio delle ripetizioni:

Esempio Trasformazione:

Parola	Verso	a b c d e f	Conteggio aggiuntivo
aabc	»	P P P A A A	P A A A

Tabella 3.2: Esempi - Combinazione con ripetizione

Si noti, nella tabella 3.2 nel campo "**Conteggio Aggiuntivo**" abbiamo, la prima P che indica una seconda ripetizione della lettera "a" e la A successiva che marca la fine del conteggio della lettera "a" (si noti che l'ordine di conteggio è l'ordine che si è scelto di implementare (a b c d e f). Le successive A marcano rispettivamente le lettere "b" e "c" che non hanno ripetizioni.

Esempio Trasformazione (aggiungiamo un campo "Terminazione" con il carattere "A" che indica la fine del conteggio:

Parola	Verso	a b c d e f	Conteggio aggiuntivo	Terminazione
aaaa	»	P A A A A A	P P P	A
bbba	»	Errore	Errore	Errore
abbb	»	P P A A A A	A P P	A
cdde	«	A A P P P A	A P A	A

Tabella 3.3: Esempi - Combinazione con ripetizione

Nota bene l'ultimo carattere di terminazione sarà sempre una "A" (ridondante)
concluderemo con:

$$ {}_{k}P^{(r)}_{n+k-1} = \binom{n+k-1}{k} = C^{(r)}_{n,k} \tag{1} $$

3.4 Schema riassuntivo

Ecco un breve schema riassuntivo

Tipologia	Ripetizione	Ordine	Simbolo	Forumula
Combinazione	Si	No	$C^{(r)}_{n,k}$	$\binom{n+k-1}{k}$
Combinazione	No	No	$C_{n,k}$	$\binom{n}{k}$
Permutazione	No	Si	P_n	$n!$
Permtazione	Si	Si	$_{k_1,\dots k_m}P^{(r)}_n$	$\binom{n}{k_1\dots k_m}$
Disposizione	Si	Si	$D^{(r)}_{n,k}$	n^k
Disposizione	No	Si	$D_{n,k}$	$\frac{n!}{(n-k)!}$

Tabella 3.4: Schema riassuntivo

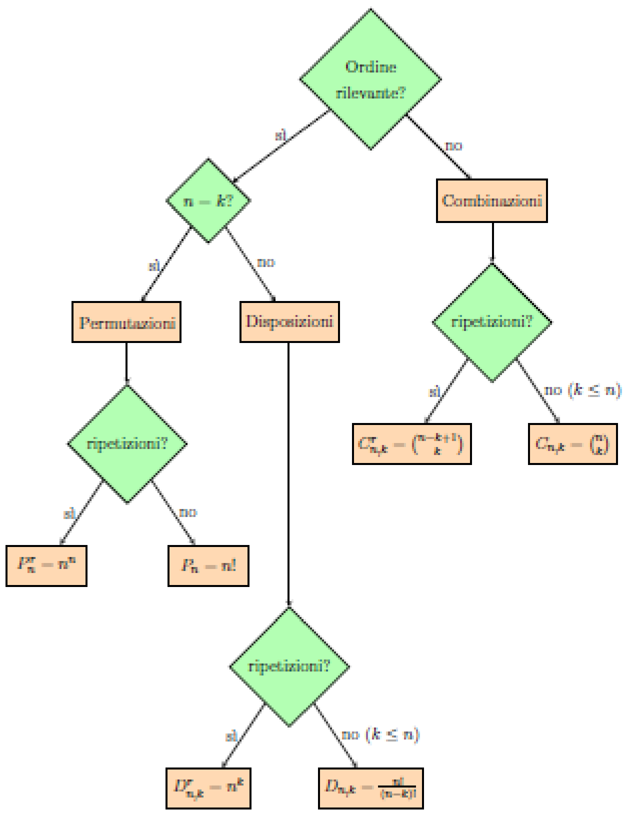


Figura 1: Flow Chart Riassuntivo

Nota bene: Il precedente flow chart è preso dagli appunti (a.a precedenti) del nostro dio **Luigi Starace**.

4 Lezione 3

4.1 Triangolo di Tartaglia

4.1.1 Definizione

In matematica, il **triangolo di Tartaglia** (detto anche triangolo di Pascal o Khayyām o Yang Hui) è una disposizione geometrica dei coefficienti binomiali, ossia dei coefficienti dello sviluppo del binomio $(a + b)$ elevato a una qualsiasi potenza n, a forma di triangolo.

Consideriamo la seguente tabella:

n,k	0	1	2	3	4	5	Somma	Somma a segni alterni
0	1	-	-	-	-	-	$1 = 2^0$	1
1	1	1	-	-	-	-	$2 = 2^1$	$1 - 1 = 0$
2	1	2	1	-	-	-	$4 = 2^2$	$1 - 2 + 1 = 0$
3	1	3	3	1	-	-	$8 = 2^3$	$1 - 3 + 3 - 1 = 0$
4	1	4	6	4	1	-	$16 = 2^4$	$1 - 4 + 6 - 4 + 1 = 0$
5	1	5	10	10	5	1	$32 = 2^5$	$1 - 5 + 10 - 10 + 5 + 1 = 0$

Tabella 4.1: Coefficenti Binomiali (n su k)

4.1.2 Interpretazione insiemistica

Il **Coefficiente Binomiale** ($\binom{n}{k}$) conta il numero di sottoinsiemi di cardinalità k costituiti dagli elemnti di un insieme di cardianlità n:

$$\{e_1, e_2, ..., e_{n-1}, e_n\}$$

4.2 Corrispondenza biunivoca tra insieme e sottoinsiemi

Ci sono tanti sottoindiemi di ordine k di un insieme di ordine N elemeti quante sono le "Combinazioni Semplicit" di lunghezza k di elementi presi in un alfabeto di cardinalità N.

4.3 Alcuni risultati

4.3.1 Risultato 1: $\binom{n}{0} = 1$

Infatti, c'è un solo sottoinsieme di cardinalità 0: l'insieme vuoto.

4.3.2 Risultato 2: $\binom{n}{n} = 1$

Infatti, c'è un unico sottoinsieme di cardinalità n: l'insieme stesso.

4.3.3 Risultato 3: $\binom{n}{k} = \binom{n}{n-k}$

Infatti, ad ogni sottoinsieme di cardinalità di k corrisponde biunivocamente il sottoinsieme complementare.

Esempio $\{e_1, ..., e_5\}$
k=2

$$\{e_1, e_2\} \iff \{e_3, e_4, e_5\}$$
$$\{e_3, e_2\} \iff \{e_2, e_4, e_5\}$$
$$...$$

4.3.4 Risultato 4: $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$

Vincoli: $n \geq 1, 1 \leq k < n$
Infatti, i sottoinsiemi di cardinalità k si possono ottenere con gli elementi (i primi n-1) dell'iniseme di cardinalità n.

Esempio $\{e_1, ..e_{n-1}, e_n\}$
k=3

$$\{e_1, e_2, e_3\}, \{e_2, e_3, e_{n-2}\}$$
$$\{e_{n-3}, e_{n-2}, e_{n-1}\}$$

In più ci sono i sottoinsiemi di cardinalità k che contengono l'ultimo elemento.

$$\{e_1, e_2, e_n\}, \{e_5, e_6, e_n\}$$
$$\{e_{n-2}, e_{n-3}, e_n\}$$

4.3.5 Risultato 5: $|S| = n, |P(S)| = 2^n$

Un sottoinsieme di cardinalità n, ammette 2^n sottoinsiemi:

DIM $\{e_1, e_2, ..., e_n\}, \{A, P\}$ come abbiamo visto per la tabella 3.2 possiamo creare un insieme di ordine 2 (quello contenente solo "A" e "P" per dimostrare tale tesi:

Parola	Verso	$e_1 e_2 e_3 ... e_{n-1} e_n$
$\{e_1, e_3\}$	»	P A P ... A A
$\{e_1, e_2, e_3\}$	«	P P P ... A A
$\{e_n\}$	«	A A A ... A P
$\{S\}$	«	P P P ... P P
$\{\varnothing\}$	«	A A A ... A A

Tabella 4.2: Dimostrazione - Risultato 5

Da questa trasformazione biunivoca si ottiene che i sottoinsiemi $\{e_1, ..., e_n\}$ sono tanti quante le **DISPOSIZIONI** con ripetizione (l'ordine è importante) da A,P di lunghezza n ($D_{2,n}^{(r)} = 2^n$).

4.3.6 Risultato 6: $\sum_{k=0}^{n} \binom{n}{k} = 2^n$

Infatti, $\binom{n}{k}$ è il numero di sottoinsiemi di cardinalità k consiste degli elementi di $e_1, e_2, ..., e_n$ e la somma rappresenta il numero complessivo di sottoinsiemi di $\{e_1, ..., e_n\}$ ovvero 2^n.

4.3.7 Risultato 7: $\sum_{k=0}^{n}(-1)^k \cdot \binom{n}{k} = 0$

Vincolo: $n \geq 1$
Se si prende il triangolo di tartaglia 4.1 nella colonna denominata "Somma a segni alterni" si nota come il risultato (escluso il primo elemento) sia sempre 0.

DIM Il risultato può essere dimostrato tramite il principio di induzione.

Base di induzione: n=1 si sviluppa in questo modo:

$$\sum_{k=0}^{1}(-1)^k \cdot \binom{n}{k} \tag{1}$$

dati i risultati 1 e 2 possiamo concludere che l'espressione 1 si traduce in questo modo:

$$1 + \left((-1)^1 \cdot \binom{1}{1}\right) \tag{2}$$

svolgendo i calcoli otterremmo banalmente come risultato 0 e quindi la base di induzione risulterà valida (si ricordi che per k=0 il risultato è 1 e non -1).

Ipotesi di induzione: poniamo l'ipotesi per n-1:

$$\sum_{k=0}^{n-1}(-1)^k \cdot \binom{n-1}{k} = 1 + \sum_{k=1}^{n-1}(-1)^k \cdot \binom{n-1}{k} \tag{3}$$

dato che per il caso con k=0 il risultato è pari a 1, sostituiamo 1 e per il risultato 4 possiamo riscrivere l'equazione nel seguente modo:

$$1 + \sum_{k=1}^{n-1}(-1)^k \cdot \left(\binom{n-2}{k} + \binom{n-2}{k-1}\right) \tag{4}$$

Proviamo per k=1 e per k=2:

k=1

$$1 - 1 + \binom{n-2}{1} \cdot (-1)^1 \tag{5}$$

k=2

$$\binom{n-2}{2} \cdot (-1)^2 + \binom{n-2}{1} \cdot (-1)^2 \tag{6}$$

se si svolgono i risultati per questi due casi si noteranno come le varie sommatorie e risultati vanno a collidere fino a trovarsi come risultato finale 0 quindi possiamo continuare cosi anche per $n \geq 2$.

Se appunto abbiamo visto che per P(n-1) per ipotesi è vera, allora concluderemo che P(n) è vera.

4.4 Ulteriori proprietà del triangolo di Tartaglia

Il trianngngolo di **Tartaglia** visto nella tabella 4.1 nasconde tanti altri segreti della matematica (i.e Fibonacci, Gauss, numeri triangolari ($t_1 = 1, m \geq 1, t_m = t_{m-1} + m$), etc...)

Altre informazioni qui[1].

[1] https://it.wikipedia.org/wiki/Triangolo_di_Tartaglia

4.5 Gioco della ZARA (analisi)

Esperimento aleatorio un processo che porta ad un risultato incerto ξ (xi).

Ulteriore definizione: Un esperimento casuale è un esperimento per cui non e possibile prevedere a priori il risultato. Ad ogni esperimento casuale è possibile associare una terna (Ω,F,P).

Evento (gli esiti sono risultanti di ξ)
sono alcuni sottoinsiemi di Ω (nel gioco della ZARA in particolare sono tutti).

nascono dalle funzioni nel contesto in cui esso è visto. (nel gioco della ZARA, è S).

SONO CONTROIMMAGINI.

Considerazioni:

- $\Omega = \{(1,1,1), ..., (1,1,6), (1,2,1), (1,2,2), ..., (1,2,6), ..., (6,6,5), (6,6,6)\}$ è l'insieme dei possibili esiti di ξ.
- $|\Omega| = 216 = 6^3 (\omega \in \Omega)$ con ω intendiamo generica tripla (in questo caso) di Ω.
- $S : \Omega \longrightarrow \mathcal{F}$ definita in questo modo $(i,j,k) \longrightarrow S = i + j + k$

Effetto di S su Ω prendiamo in considerazione la seguente tabella:

Insieme	S	$\mathcal{P}$
$\{(1,1,1)\} \subseteq \Omega$	$S = 3$	$\mathcal{P} = \frac{1}{216}$
$\{(1,1,1), (2,1,1), (1,2,1)\} \subseteq \Omega$	$S = 4$	$\mathcal{P} = \frac{3}{216}$
$\{(1,1,3), (3,1,1), (1,3,1), (1,2,2), (2,1,2), (2,2,1)\} \subseteq \Omega$	$S = 5$	$\mathcal{P} = \frac{6}{216}$
...	...	...

Tabella 4.3: Tabella effetti - S su Ω

Nel nostro caso visto nella "Lezione 1" avevamo $S = 10$.

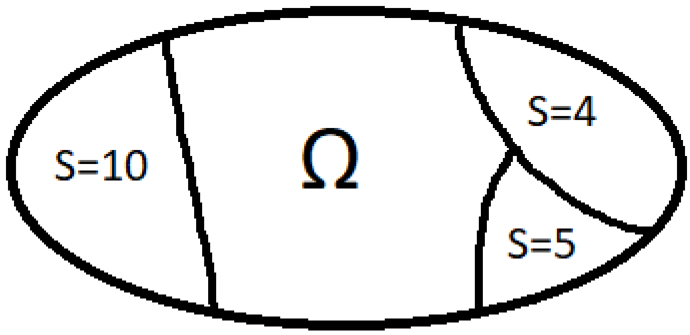


Figura 2: Esempio di partizione di Ω

Sono partizioni (le partizioni sono **eventi**) infatti la probabilità che l'evento per $S = 10$ è di $\mathcal{P}(S = 10) = \frac{27}{216}$.

Ulteriori esempi possono essere i seguenti:

- $S^{-1}(\{3, 18\}) = \{(1,1,1), (6,6,6)\}$.
- $S^{-1}(\{19\}) = \varnothing$.

Conclusione $\mathrm{B} \subseteq \mathrm{R}$ (dove B = Borel), $S^{-1}(B) \subseteq \Omega$.

4.6 Albero delgi esperimenti casuali

$$\begin{array}{ccc} & \xi & \\ \swarrow & \downarrow & \searrow \\ \Omega & \mathcal{F} & \mathcal{P} \end{array}$$

- Ω - insieme degli esiti di ξ.
- $\mathcal{F}$ - famiglia degli elemnti.
- P - la misura delle probabilità degli eventi.

In particolare avremo che la probabilità può essere calcolata tra 0 e 1 dato che:

- $\mathcal{P}(\varnothing) = 0$.
- $\mathcal{P}(\Omega) = 1$.

5 Lezione 4

Esempio introduttivo: si consideri il seguente esperimento:
un solo dado (onesto, anche se dichiararlo tale in questo caso è superficiale) e come insieme di eventi si consideri il lancio del dado e il risultato. Avremo dunque il seguente insieme:

$$\Omega = \{1, 2, 3, 4, 5, 6\} \tag{1}$$

Ω viene detto **Exemple space** (Spazio campione) e fa parte della descrizione di ξ.

Per il momento definiremo $\xi : (\Omega, -, -)$

5.1 L'ordine di Ω ed esperimenti

$$|\Omega| = 6 \Longrightarrow |P(\Omega)| = 2^6 \tag{1}$$

L'implicazione 1 è definita tramite le proprietà insiemistiche per cui l'insieme delle parti di un insieme S di ordine n è definita da P(S) con ordine 2^n.

Suddividiamo Ω in vari sottoinsiemi:

- P (Pari) = {2,4,6};
- D (Dispari) = {1,3,5}
- A (Alti) = {5,6}
- M (Medi) = {3,4}
- B (Bassi) = {1,2}

Da queste suddivisioni (alcune tra le tante possibili) si ipotizza che si scommetta su varie giocate, come risultato si vuole definire la famiglia $\mathcal{F}$ di eventi.

Giocate si suppongono le seguenti giocate:

- Dispari: il banco vince se esce un numero pari ({2,4,6}) (il negato/-complemento di D).
- Bassi: il banco vince se esce un numero alto o medio ($\{3, 4\} \cup \{5, 6\}$) (il negato/complemento di B)

come si vede si possono fare un numero limitato di combinazioni tra gli insiemi (con l'ausilio delle operazioni insiemistiche) per creare un nuovo insieme, uno tra i tanti che compongono $\mathcal{F}$.

5.1.1 Proprietà dell'algebra

Le proprietà che si possono estrarre dalle passate supposizioni sono:

- Se c'è un evento, allora anche il suo complementare è un evento.
- L'unione di eventi (finita) è un evento.
- Ω è un evento (evento certo).

Alcune definizioni le principali sono:

- **Generatori**: sono gli eventi iniziali che si prendono in considerazione.
- **Atomi**: è l'intersezione di tutti gli eventi possibili.

5.2 Esempi: generatori e atomi

Prendiamo in considerazione D e B:

- $D \cap B = \{1\}$
- $D \cap B^c = \{3, 5\}$
- $D^c \cap B = \{2\}$
- $D^c \cap B^c = \{4, 6\}$

Nota bene con A^c ci si riferisce spesso al complemento di A ($\overline{A}$).

La nostra famiglia di eventi ($\mathcal{F}$) sarà composta da tutte le intersezioni a due a due nonchè comprenderà anche il vuoto ($\varnothing$) e l'evento certo ({1,2,3,4,5,6}). Si noterà che l'ordine di $\mathcal{F}$ sarà pari a 16 ovvero:

$|\mathcal{F}| = 16 = 2^4$ dove 4 è proprio il numero di atomi.

Ulteriore esempio si prendano in considerazione D,B ed M e si facciano le varie intersezioni in seguito riportate:

Dispari	Bassi	Medi	Result
D	B	M	$\varnothing$
D	B	M^c	$\{1\}$
D	B^c	M	$\{3\}$
D	B^c	M^c	$\{5\}$
D^c	B	M	$\varnothing$
D^c	B	M^c	$\{2\}$
D^c	B^c	M	$\{4\}$
D^c	B^c	M^c	$\{6\}$

Tabella 5.1: Esempio

Adesso è possibile aggiungere ad ξ la famiglia degli eventi:

$$\xi : (\Omega, \mathcal{F}, -)$$

La famiglia degli eventi possiamo chiamarla "**ALGEBRA**", infatti la famiglia degli eventi "$\mathcal{F}$" ha delle proprietà caratteristiche che la fanno diventare struttura algebrica.

5.3 Unione numerabile di eventi

Consideriamo adesso il seguente esperimento: lancio di un numero **INDEFINITO** di volte di una monetina. Definiamo il suo insieme Ω:

$$\Omega = \{(w_1, w_2, ...) : w_n = T \text{ oppure } w_n = C\} \tag{1}$$

Definiamo

- $(w_1, w_2, ...)$ è un asuccessione di lanci.
- T e C stanno rispettivamente per testa o croce.
- $\{T_n\}_{n \in \mathbf{N}}$: T_n : esce testa nel lancio di ordine n.
- $\{C_n\}_{n \in \mathbf{N}}$: C_n : esce croce nel lancio di ordine n.

Considereremo ora un probabile evento A = "TT esce prima di CC":

$A_1 = TT; A_2 = CTT; A_3 = TCTT; A_4 = CTCTT; ...$

il nostro A totale sarà uguale a:

$$A = A_1 \cup A_2 \cup A_3 \cup ... = \bigcup_{n=1}^{\infty} A_n \in \mathcal{F} \tag{2}$$

Rielenchiamo in questo caso di nuovo le **Proprietà dell'Algebra**:

i Unione **NUMERABILE** di eventi appartiene a $\mathcal{F}$ - $\Omega \in \mathcal{F}$.

ii Complemento di un evento è esso stesso un evento - $A \in \mathcal{F} \Longrightarrow A^C \in \mathcal{F}$.

iii Ω è un evento - $(A_n)_{n \in \mathbf{N}} \subseteq \mathcal{F} \Longrightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

5.3.1 Cosa vuol dire enumerabile?

$\mathbf{N} = \{1, 2, 3, ...\}$ insieme degli interi positivi.
$|\mathbf{N}| = \propto_0$ (alep 0) è rappresenta la cardialinalità di $\mathbf{N}$.

- $|\{a, b, ..., z\}| = 21$
- $|\{a, e, i, o, u\}| = 5$

Definiamo la seguente funzione biettiva (sui numeri dispari, $D \subseteq \mathbf{N}$):

- $n \longrightarrow 2 \cdot n - 1$
- $1 \longrightarrow 1$
- $2 \longrightarrow 3$

- $3 \longrightarrow 5$

N è un insieme infinito e : $|\mathbf{N}| = \propto_{\mathbf{0}}$;
N ha la cardinalità del **numerabile**;
N è numerabile;

da qui in poi si sussegue un discorso sulla definizione di **NUMERABILE** e il come ogni insieme infinito partendo dai numeri naturali si arriva in una successione di "⊆" fino da **R** che si definirà di cardinalita del **CONITNUO** ($c > \propto_0$).

Vedesi anche il paradosso di Hilbert[1].

[1] https://it.wikipedia.org/wiki/Paradosso_del_Grand_Hotel_di_Hilbert

6 Lezione 5

Teorema 1. $(1)\Omega \in \mathcal{F},\ (2)A \in \mathcal{F} \Longrightarrow A^c \in \mathcal{F}$ *tesi* $\varnothing \in \mathcal{F}$

Dimostrazione. $\Omega \in \mathcal{F} \overset{(2)}{\Longrightarrow} \Omega^c \in \mathcal{F} \Longleftrightarrow \varnothing \in \mathcal{F}$ □

possiamo definire il vuoto ($\varnothing$) come **evento impossibile**.

Teorema 2. $n \in \mathbf{N},\ A_1, A_2, ..., A_n \in \mathcal{F}$ *tesi* $A_1 \cup A_2 \cup ...A_n = \bigcup_{i=1}^{n} A_i \in \mathcal{F}$

Dimostrazione. Consideriamo 2 casi:

- $i = (1, 2, ..., n),\ B_i = A_i;$
- $i > n,\ B_i = \varnothing$

Allora definiremo:

$$\bigcup_{m=1}^{\infty} B_m \in \mathcal{F} \Longleftrightarrow \bigcup_{m=1}^{n} A_m \cup \left(\bigcup_{m=n+1}^{\infty} B_m\right) \tag{1}$$

$$\left(\bigcup_{m=1}^{n} A_m\right) \cup \varnothing \Longleftrightarrow \bigcup_{m=1}^{n} A_m \in \mathcal{F} \tag{2}$$

□

6.1 Assimomi fondamentali

$$\left.\begin{aligned} (i)\Omega \in \mathcal{F} \\ (ii)A \in \mathcal{F} \Longrightarrow A^c \in \mathcal{F} \\ (iii)(A_n)_{n\in\mathbf{N}} \subseteq \mathcal{F} \Longrightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F} \end{aligned}\right\} \text{Sistema assiomi} \tag{1}$$

se si considerano questi 3 assiomi (chiusura su unioni **numerabili** per via della proprietà iii) $\mathcal{F}$ possiamo definirla una **sigma algebra**.

$$\left.\begin{aligned} (i')\Omega \in a \\ (ii')A \in a \Longrightarrow A^c \in a \\ (iii')(A_1, A_2, ..., A_n) \subseteq a \Longrightarrow \bigcup_{i=1}^{n} A_I \in a \end{aligned}\right\} \text{Sistema assiomi} \tag{2}$$

in questo secondo caso a è definita come **algebra** (chiusa per le successioni finite per via della proprietà iii').

Teorema 3. $(A_n)_{n\in\mathbf{N}} \subseteq \mathcal{F} \Longrightarrow \bigcap_{n\in\mathbf{N}} A_n \in \mathcal{F}$

Dimostrazione. **De Morgan:**

$$\bigcap_{n=1}^{\infty} A_n = (\bigcup_{n=1}^{\infty} A_n^c)^c \tag{1}$$

per avere quel risultato sfruttiamo gli assiomi visti (Schema 1)

- $A_n^c \in \mathcal{F}$ per la proprietà ii;
- $\bigcup_{n=1}^{\infty} A_n^c \in \mathcal{F}$ per la proprietà iii;
- $(\bigcup_{n=1}^{\infty} A_n^c)^c \in \mathcal{F}$ per la proprietà ii (riferito all'apice c);

□

Teorema 4. $A \in \mathcal{F},\ (A^c)^c = A$

6.1.1 Considerazioni

$n \in \mathbf{N}$
$(A_1, A_2, ..., A_n) \in \mathcal{F}$
$(\bigcap_{i=1}^{n} A_i)^c$ = si verifica almeno uno degli eventi tra $A_1...A_n = \bigcup_{i=1}^{n} A_i^c$

dal teorema 4 si edice successivamente che:

$$\Longrightarrow [(\bigcap_{i=1}^{n} A_i)^c]^c = [\bigcup_{i=1}^{n} A_i^c]^c \tag{1}$$

di conseguenza:

$$\Longleftrightarrow \bigcap_{i=1}^{n} A_i = [\bigcup_{i=1}^{n} A_i^c]^c \tag{2}$$

6.2 Proposizioni

- $A, B \in \mathcal{F},\ A \cup B = A \cup (B \cap A^c)$
- $A, B \in \mathcal{F},\ A = (A \cap B) \cup (A \cap B^c)$

6.3 Misura della probabilità

Esperimento: ξ, quando ci riferiamo ad un esperimento prendiamo sempre in considerazione un sottoinsieme di eventi definiti con **g: eventi generatori** e abbiamo costruito una terna di $\xi : (\Omega, \mathcal{F}, -)$.

in particolare $\mathcal{F}$ può essere riscritta in come $\mathcal{F} : \sigma(g)$.

6.4 Introduciamo la probabilità

$\mathcal{P} : \mathcal{F} \to \mathbf{R}$
in particolare il dominio sarà la famiglia degli eventi mentre come codominio l'insieme dei numeri reali.
$A \to \mathcal{P}(A)$
viene chiamata **"Misura di probabilità"**.

6.4.1 Tipologie di probabilità

- **Probabilità classica** - Il metodo classico al calcolo della probabilità è quello di contare il numero di risultati favorevoli, il numero di risultati possibili (si presuppone che i risultati siano contemporaneamente esclusivi ed equiparabili), ed esprimere la probabilità come un rapporto di questi due numeri.

- **Probabilita frequetista** - Nella teoria di frequenza della probabilità, la probabilità è il limite della frequenza relativa con cui un evento si verifica nelle prove ripetute (notare che le prove devono essere indipendenti).

- **Probabilità soggettiva**- Nella teoria soggettiva della probabilità, la probabilità misura il "grado di convinzione" di chi parla che si verifichi l'evento, su una scala da 0% (incredulità completa che l'evento accadrà) al 100% (certezza che l'evento accadrà).

- **Probabilità Bayes (Teoria)** - La teoria di Bayes, invece, assegna alle probabilità ogni dichiarazione sorta, anche se nessun processo casuale è coinvolto. La probabilità, per una teoria di Bayes, è un modo di rappresentare il grado di attendibilità in un'affermazione di un individuo, date le prove.

6.5 Probabilità classica ($\mathcal{P}_c$)

$\exists n \in \mathbf{N}$, $|\Omega| = n$
$\Omega = \{w_1, w_2 ... w_n\}$ e $\mathcal{F} = \sigma(g) = a(g)$ (ovvero $\mathcal{F}$ è un algebra)

ξ =nessuno degli esiti elementari è preferito da ξ esiste una perfetta simmetria tra gli elementi di Ω.

$$\left. A \in \mathcal{F},\ \mathcal{P}(A) := \frac{|A|}{|\Omega|} \right\} \textbf{Probabilità classica}(\mathcal{P}_c) \tag{1}$$

in particolare la probabilità classica di Ω:
$\mathcal{P}_c(\Omega) = \frac{|\Omega|}{|\Omega|} = 1$

Esempio consideriamo ancora l'esempio del lancio di un dado e prendiamo in considerazione i risultati (eventi) medi (M) e bassi (B):

$$\begin{gathered} \mathcal{P}_c(M \cup B) = \mathcal{P}(\{1,2,3,4\}) = \frac{4}{6} \\ \mathcal{P}_c(M) = \frac{2}{6};\ \mathcal{P}_c(B) = \frac{2}{6};\ \text{la loro somma è pari a } \frac{4}{6} \\ \mathcal{P}_c(M \uplus B) = \mathcal{P}_c(M) + \mathcal{P}_c(B) \end{gathered} \tag{2}$$

Se $A \cap B = \varnothing \Longrightarrow \mathcal{P}_c(A \cup B) = \mathcal{P}_c(A) + \mathcal{P}_c(B)$ (proprietà fondamentale).

in generale:

$$\left. \begin{aligned} A \cap B = \varnothing \\ B \cap C = \varnothing \\ A \cap C = \varnothing \end{aligned} \right\} \mathcal{P}_c(A \cup B \cup C) = \mathcal{P}_c(A) + \mathcal{P}_c(B) + \mathcal{P}_c(C) \tag{3}$$

in conclusione: $A \in \mathcal{F} \Longrightarrow \mathcal{P}_c(A) \geq 0$

6.6 Probabilità relativa (o frequentista, $\mathcal{P}_f$)

L'eseperimento è ripetibile nelle medesime condizioni in un numero **indefinito** di volte (ciascuna di esse è una prova).

$$A \in \mathcal{F},\ \mathcal{P}_f(A) = \lim_{n\to\infty} \frac{n_A}{n} \Big\} \textbf{Probabilità frequentista}(\mathcal{P}_f) \tag{1}$$

Nota bene il limite non si intende come limite matematico e andrebbe messo tra apici (') per differenziarlo, per semplicità lo ometto.

- n_A - numero tra le n prove dove A si verifica.
- n - numero delle prove.

6.6.1 $\mathcal{P}_f(\Omega)$

$$\mathcal{P}_f(\Omega) = \lim_{n\to\infty} \frac{n_\Omega}{n} = \lim_{n\to\infty} \frac{n}{n} = 1 \tag{2}$$

6.6.2 $\mathcal{P}_f(A \cap B)$

Si considerino A e B eventi disgiunti ($A \cap B = \emptyset$):

$$\mathcal{P}_f(A \cup B) = \lim_{n\to\infty} \frac{n_{A\cap B}}{n} \tag{3}$$

il che è possibile scomporre nel seguente modo:

$$\lim_{n\to\infty} \frac{n_A + n_B}{n} = \lim_{n\to\infty} \frac{n_A}{n} + \frac{n_B}{n} \tag{4}$$

e con questo concludiamo che:

$$\mathcal{P}_f(A \uplus B) = \mathcal{P}_f(A) + \mathcal{P}_f(B) \tag{5}$$

ciò ci porta a dire che sia $A \in \mathcal{F}$ e sia $\mathcal{P}_f(A) = \lim_{n\to\infty} \frac{n_A}{n} \geq 0$.

7 Lezione 6

7.1 Probabilità soggettiva ($\mathcal{P}_s$)

Non sussiste simmetria perfetta tra gli esisti dell'esperimento e non si possono ripetere in numero indefinito di volte. In particolare si fa uso di due principi fondamentali:

- $(b_1)_s$ principio di equità - è basato sulla scommessa equa che permette lo scambio tra banco e giocatore.
- $(b_2)_s$ principio di coerenza - non è possibile individuare un sistema di eventi che possa assicurare vantaggio/svantaggio all'individuo che pone le probabilità.

$A \in \mathcal{F}$, la probabilità di A è prezzo che un individuo **coerente** ritiene **equo** pagare per ricevere 1 nel caso che l'evento si verifichi.

Valgono le solite proprietà viste delle misure di probabilità classica e frequenziale:

- $\mathcal{P}_s(\Omega) = 1$
- $\mathcal{P}_s(A) \geq 0$
- $A, B : \ A \cap B = \varnothing, \ \mathcal{P}(A \cup B) = \mathcal{P}_s(A) + \mathcal{P}_s(B)$

Secondo **Kolmogorov**, la misura della probabilità degli eventi (elementi di una sigma algebra) è una funzione:

$$\begin{gathered} P : \mathcal{F} \to [0,1](\mathbf{R}) \\ A \to \mathcal{P}(A) \end{gathered}$$

con le seguenti proprietà (in aggiuna alle prime tre di $\mathcal{F}$):

- $A \in \mathcal{F}, \ \mathcal{A} \geq 0$
- $\mathcal{P}(\Omega) = 1$
- $(A_n)_{n \in \mathbf{N}} \subseteq \mathcal{F} : \ A_i \bigcap_{i \neq j} B_i = \varnothing, \ \mathcal{P}(\dot\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mathcal{P}(A_n)$

7.2 Teoremi su Kolnogrov

Teorema 5. $A \in \mathcal{F},\ \mathcal{P}(A) \leq 1$

Dimostrazione. In generale sappiamo che:
$\mathcal{P}(S) = \mathcal{P}(A \cup A^c) = \mathcal{P}(A) + \mathcal{P}(A^c)$

per l'assioma numero 4 che definisce $\mathcal{P}(A^c) \geq 0$ allora possiamo concludere che:
$\mathcal{P}(A) = 1 - \mathcal{P}(A^c) \leq 1$ □

Teorema 6. ***Numerabile additiva***
$\mathcal{P}(\varnothing) = 0$

Dimostrazione. $\varnothing = \varnothing \cup \varnothing \cup ...$
si ponga:

$$\mathcal{P}(\varnothing) = \sum_{n=1}^{\infty} \mathcal{P}(\varnothing) * \tag{1}$$

p.a - sia $\epsilon = \mathcal{P}(0 < \epsilon = \mathcal{P}(\varnothing))$ la serie a secondo membro diverge, per cui (dato l'assurdo) $\mathcal{P}(\varnothing) = 0$ (unica possibiltà per far valere * è porre $\epsilon = 0$):

$$\mathcal{P}(\varnothing) = 0$$

□

Teorema 7. ***Finita additiva***
$n = \mathbf{N},\ (A_1, A_2, ..., A_n) \subseteq \mathcal{F}:\ A_i \bigcap_{i \neq j} A_j$

$$\mathcal{P}(\dot{\bigcup_{i=1}^{n}} A_i) = \sum_{i=1}^{n} \mathcal{P}(A_i) \tag{1}$$

Dimostrazione. Conideriamo: $i = (1, 2, ..., n),\ B_i = A_i$;
$i > n,\ B_i = \varnothing$;

$$\begin{aligned} \mathcal{P}(\dot{\bigcup_{n=1}^{\infty}} B_n) &= \sum_{n=1}^{\infty} \mathcal{P}(B_n) \\ = \sum_{i=1}^{n} \mathcal{P}(B_i) &+ \sum_{i=n+1}^{n} \mathcal{P}(B_i) \\ = \sum_{i=1}^{n} \mathcal{P}(A_i) &+ \sum_{i=n+1}^{n} \mathcal{P}(\varnothing) \end{aligned} \tag{2}$$

D'altra parte:

$$\begin{aligned}\mathcal{P}(\dot{\bigcup_{i=1}^{\infty}} B_n) &= \mathcal{P}[(\dot{\bigcup_{i=1}^{n}}) \cup (\dot{\bigcup_{i=n+1}^{\infty}} \varnothing)] \\ &= \mathcal{P}[(\dot{\bigcup_{i=1}^{n}} A_i) \cup \varnothing] \\ &= \mathcal{P}(\dot{\bigcup_{i=1}^{n}} A_i)\end{aligned} \tag{3}$$

□

Teorema 8. $A \subseteq B,\ \mathcal{P}(A) \le \mathcal{P}(B)$

Dimostrazione.

$$\begin{aligned} B = A \dot{\bigcap} (B \cap A^c) &\Longrightarrow \\ \mathcal{P}(B) = \mathcal{P}(A) + \mathcal{P}(B \cap A^c) &\Longrightarrow \\ \mathcal{P}(B) \ge \mathcal{P}(A) & \end{aligned} \tag{1}$$

in quanto $\mathcal{P}(B \cap A^c) \ge 0$ per l'assioma 4. □

7.3 Probabilità indotta da una funzione S

Nel gioco della Zara si prende un elemento (un tripla) della σ−algebra (insieme di partenza) e poi esegue la somma della compenti della tripla.

$\mathcal{P}$ è la misura di probalità classica (indotta da funzione S).

A partire dela misura di probabilità $\mathcal{P}$.

$$(\Omega, \mathcal{F} = \mathcal{P}(\Omega), \mathcal{P}_c)$$

creiamo una funzione:
$S : \Omega \to \mathbf{R}$
$\omega = (\omega_1, \omega_2, \omega_3) \to \omega_1 + \omega_2 + \omega_3\ (S(\underline{\omega}))$

con questo avremo la nuova terna:

$$(\mathbf{R}, B, \mathcal{P}_s)$$

con **B** ci riferiamo all'algebra di Borel, mentre con $\mathcal{P}_s(B)$:

$$\mathcal{P}(S \in B) = \mathcal{P}[S{-}1(B)] - \textbf{Contro immagine} \tag{1}$$

7.4 Esperimento di tipo Bernulli

Consideriamo l'esperimento del lancio di una moneta un numero indefinito di volta e si supponga che essa sia onesta o truccata e che nel secondo caso si conosca tale trucco per il calcolo delle probabilità:

$$\Omega = \{\underline{\omega} = (\omega_1, \omega_2, ...) : \omega_n \in \{T, C\}\}$$

$$(g_n)_{n\in\mathbf{N}} = (T_n)_{n\in\mathbf{N}} \Longrightarrow \mathcal{F} = \sigma((T_n)_{n\in\mathbf{N}})$$

se la moneta è equa:

$n \in \mathbf{N},\ \mathcal{P}(T_n) = \frac{1}{2}$

se la moneta non è equa:

$n \in \mathbf{N},\ \mathcal{P}(T_n) = p$ con $p \in \{0, 1\}$
$\mathcal{P}(\mathcal{C}) = 1 - p$

definiamo i valori ottenibili dalla funzione (S):

$$n \in \mathbf{N},\ \Omega \to \mathbf{R}$$
$$\underline{\omega} \to \begin{cases} 1, & se\ \omega_n = T \\ 0, & se\ \omega_n = C \end{cases} \tag{1}$$

Ovvero se si verifica T_n la funzione X_n vale 1 altrimenti vale 0.

Esempi vari esempi mostrati:

1) $\mathcal{P}(X_n = 0) = \mathcal{P}(\{0\}) = \mathcal{P}(C_n) = 1 - p$

2) $\mathcal{P}(X_n = 2) = \mathcal{P}(\{2\}) = 0$

3) $\mathcal{P}(]\frac{1}{2}, 2[) = \mathcal{P}(\{1\}) = \mathcal{P}(X_n = 1) = \mathcal{P}(T_n) = p$

4) $\mathcal{P}(]-\infty, -3]) = \mathcal{P}(X_n \in]-\infty, -3]) = \mathcal{P}(X_n \leq -3) = \mathcal{P}(\varnothing) = 0$

Si noti per l'esempio 2 e 4 che la probabilità cade allo 0 (ovvero racchiude eventi impossibili).

7.5 Ulteriore esempio e cosa significa X "misurabile"

Esempio:

$$\mathcal{P}_{x_n}(]-\infty,2]) = \mathcal{P}(x_n \in]-\infty,2]) \tag{1}$$

scomponiamo l'intervallo creando un insieme con i valori accettati (nel nostro caso entrambi i valori 0 e 1 si trovano all'interno dell'intervallo):

$$\mathcal{P}(x_n \in \{0,1\}) \tag{2}$$

ora scriviamo in termini insiemistici:

$$\mathcal{P}[\{x_n = 0\} \cup \{x_n = 1\}] = \mathcal{P}(C_n \cup T_n) \tag{3}$$

per le proprietà della probabilità classica:

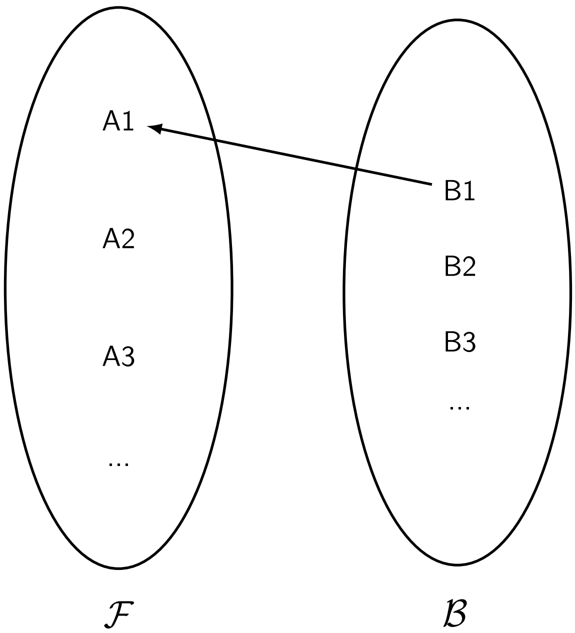
$$\mathcal{P}(x_n = 0) + \mathcal{P}(x_n = 1) = \mathcal{P}(C_n) + \mathcal{P}(T_n) \tag{4}$$

il risultato sarà quindi:

$$1 - p + p = 1 \tag{5}$$

Concluso questo esempio ragioniamo sull'insieme B (Borel):

$B \in \mathcal{B},\ \mathcal{P}_{x_n}(B) = \mathcal{P}(x_n \in B) = \mathcal{P}(X^{-1}(B))$
a patto che $\{X_n \in B\}$ sia un evento.



Dal grafico precendente si immagini di avere tanti eventi in $\mathcal{F}$ (A1,A2,...) la freccia che va da B1 ad A1 ($B1 \to A1$ sta ad indicare la trasformazione dell'antimmagine di X_n ovvero X_n^{-1}.

In generale $X_\Omega \to \mathbf{R}$ deve avere questa proprietà:
$B \in \mathcal{B},\ X^{-1}(B) \in \mathcal{F}$
e diremo che **X è misurabile**.

Funzione misurabile definizione:
In analisi matematica, una funzione misurabile è una funzione tra due spazi misurabili compatibile con la loro struttura di σ-algebra.

La richiesta di misurabilità di una funzione è in genere un'ipotesi di regolarità minima, ed è molto spesso richiesta per l'applicazione dei teoremi e dei metodi dell'analisi matematica e della teoria della misura.

8 Lezione 7

Nella lezione precedente abbiamo considerato le seguenti 3 terne di elemeniti:

- $(\Omega, \mathcal{F} = \sigma(g), \mathcal{P})$
- $(\mathbf{R}, \mathcal{B} = \sigma(\mathcal{F}), \mathcal{P}_x)$

$$X: \ \Omega \xrightarrow{misurabile} \mathbf{R}$$
$$\omega \longrightarrow X_{(\omega)}$$

specifichiamo:

- $\mathcal{P}_x$, misura la probalità $\mathcal{P}$ con la funzione misurabile X
- $\mathcal{B}$=Borel, ha insiemi g di generatori (intervalli di $\mathbf{R}$)

per la quale:

$$\forall B \in \mathcal{B}, \ \{x \in \mathcal{B}\} \in \mathcal{F}$$
$$= \{\omega \in \Omega : \ X(\omega) \in \mathcal{B}\} \subseteq \Omega$$

in particolare se è rispettata la proprietà ($\subseteq \mathcal{F}$) allora è possibile dire che X è misurabile.

questa verifica è equivalente:

$$\forall B \in \mathbf{R}, \ \{X \leq x\} \in \mathcal{F}$$
$$= \{\omega \in \Omega : \ X(\omega) \leq x\} \subseteq \Omega$$

Stiamo considerando: $B =]-\infty, X]$
se la proprietà è valida per tutte le semirette sinistre, X è misurabile.

Dopo di ciò:

$$\mathcal{B} \longrightarrow \mathbf{R}$$
$$B \longrightarrow \mathcal{P}_x(B) := \mathcal{P}(X \in \mathcal{B})$$

per definizione X è un evento.

Una funzione X tra Ω e $\mathbf{R}$ che sia misurabile è **VARIABILE ALEATORIA** (numero) possiamo definire di "Aleatorio" ω, il risultato dell'esperimento.

8.1 Variabili aleatorie

Dato uno spazio di probabilità una variabile aleatoria (chiamata anche variabile casuale o variabile stocastica) X è una funzione.

Nello schema di Bernulli (prove ripetute):
$\Omega = \{T, X\}^N$ - prodotto cartesiano T e C un numerabile di volte ($\mathbf{N}$)
$g = \{(T_n)_{n \in \mathbf{N}}\}, \ \mathcal{F} = \sigma(g)$

$$A \in \mathcal{F}, \ \mathcal{P}(A) n \in \mathbf{N}, \ X_n : \ \Omega \longrightarrow \mathbf{R}$$
$$\underline{\omega} \longrightarrow X_n(\omega) \begin{cases} 1, & se\ T\ in\ posizione\ n \\ 2, & se\ C\ in\ posizione\ n \end{cases}$$

Analizziamo i 3 casi principali:$]-\infty, x] \subseteq \mathbf{R}, \ \{X_n \leq x\} \subseteq \Omega$

- x>1 - $\{\omega \in \Omega : \ X_n < x\} = \Omega \in \mathcal{F}$
- $0 \leq x < 1$ - $\{\omega \in \Omega : \ X_n < x\} = (C_n = T_n^c) \in \mathcal{F}$
- x<0 - $\{\omega \in \Omega : \ X_n < x\} = \varnothing \in \mathcal{F}$

nello schema di Bernulli, X_n rispetta le proprietà di variabile aleatoria.

8.2 Somma di variabili aleatorie e spettro

Lo spettro (o supporto) di una variabile aleatoria X è l'insieme S_x di valori (o stati) che questa può assumere.

La somma di variabili aleatorie è una variabile aleatoria:

$$n \in \mathbf{N}, \ S_n = X_1 + X_2 + ... + X_n \sim B(n, p)$$

qui definiamo con n il numero di ripetizioni e con B **(VARIABILE BINOMIALE)**.

Tornando all'esperimento Bernilli $n \in \mathbf{N}, \ X_n \sim B(1, p), \ p = \mathcal{P}(T) \in (0, 1)$ ovviamente consideriamo anche il complemento ovverò C definito come $1 - p = \mathcal{P}(C) \in (0, 1)$

Spettro $S_{S_n} = \{0, 1, 2...n\}$ è l'insieme dei possibili valori.

$$k \in S_{S_n},\ \mathcal{P}_{S_n}(\{k\} = \mathcal{P}(S_n = k)) \tag{1}$$

che come risultato definisce:

$$= \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \tag{2}$$

la precedente equazione viene chiamata **Legge di B(n,p)**

Considerando una sequenza: $T_1 T_2 ... T_K C_{k+1} C_{k+2} ... C_n$
anche se facessimo una permutazione dell'ordine con cui appaino le teste e croci

i.e $C_1 T_2 ... T_K C_{k+1} C_{k+2} ... T_n$ il risultato dell'equazione 2 non cambia e infatti possiamo dire che esistono $\binom{n}{k}$ permutazioni di tale sequenza.

8.3 Problema del Cavaliere De Méré

4 lanci di un dado onesto "almeno un 6"
$n = 4,\ p = \frac{1}{6}$
dovremmo definire la somma S_4 e sarebbe troppo complesso, quindi definiamo il complemento (si onsideri il caso in cui 1 se il dado mostra 6 e 0 se mostra una faccia da 1 a 5):

$$\{S_4 \geq 1\} = \{S_4 = 0\}^c$$

quindi calcoliamo in termini di misura della probabilità $S_4 = 0$:

$$\mathcal{P}(S_4 = 0) = \binom{4}{0} \cdot p^0 \cdot (1-p)^4 = (\frac{5}{6})^4 \tag{1}$$

adesso complementiamo e noteremo che il Cavaliere De Mèrè aveva un agio a suo favore:

$$\mathcal{P}(S_4 \geq 1) = 1 - (\frac{5}{6})^4 \longrightarrow > 0.5 \tag{2}$$

il cavaliere ci riprova con un secondo esperimento, 24 lanci del doppio dato onesto "almeno un doppio 6", facciamo valere le condizioni del precedente esperimento: $n = 24,\ p = \frac{1}{36}$

$$\{S_{24} \geq 1\} = \{S_{24} = 0\}^c$$

quindi calcoliamo in termini di misura della probabilità $S_{24} = 0$:

$$\mathcal{P}(S_{24} = 0) = \binom{24}{0} \cdot p^0 \cdot (1-p)^{24} = (\frac{35}{36})^4 \tag{3}$$

adesso complementiamo e noteremo che il Cavaliere De Mèrè aveva un agio a suo favore:

$$\mathcal{P}(S_24 \geq 1) = 1 - (\frac{35}{36})^4 \longrightarrow < 0.5 \tag{4}$$

Nota bene facciamo il complemento come accennato prima per evitare la sommatoria:
$\mathcal{P}(S_{24} \geq 1) = \mathcal{P}(S_{24} = 1) + ... + \mathcal{P}(S_{24} = 24) = p - \mathcal{P}(S_{24} = 0)$

Consideriamo $\mathcal{P}(\Omega)$:

$$= \mathcal{P}(\{S_n = 0\} \uplus ... \uplus \{S_n = n\} = \tag{5}$$

Teorema 9. ***Binomio di Newton***

$$\sum_{k=0}^{n} \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} = [p \cdot (1-p)]^n = 1 \tag{1}$$

Dimostrazione.

$$\mathcal{P}(S_n = 0) + ... + \mathcal{P}(S_n = n)) \sum_{k=0}^{n} \mathcal{P}(S_n = k) = \sum_{k=0}^{\infty} \binom{24}{0} p^k (1-p)^{n-k} = 1 \tag{2}$$

infatti $[p \cdot (1-p)]^n$ viene detto **Quadrato di Newton** □

8.4 Ulteriore variabile aleatoria (geometrica)

Consideriamo adesso un altra variabile aleatoria nello schema di Bernulli: $T_1 = inf\{n \in \mathbf{N} : \; X_n = 1\} \sim G(P)$ dove la variabile G è detta **geometrica**.

C C C C C T C C T C C C T T T...

la precedente sequenza vede la nostra T_1 (posizione dove appare per la prima volta T) con valore pari a 6, vediamo lo spettro di tale funzione:

$$S_{T_1} = \{1, 2, ...\} = \mathbf{N} \text{ ovver il numerabile}$$

$$k \in S_{T_1}, \; \mathcal{P}_{T_1}(\{k\}) = \mathcal{P}(T_1 = k) = p \cdot (1 - p)^{k-1} = p \cdot q^{k-1} \tag{1}$$

Nota bene $(p \in (0, 1) \Longrightarrow Q \in (0, 1)$ infatti:

$$1 = \mathcal{P}(\Omega) = \sum_{k=1}^{\infty} p \cdot q^{k-1} = p \sum_{k=1}^{\infty} q^{k-1} = p \sum_{m=0}^{\infty} q^m = 1 \cdot \frac{1}{1-q} = \frac{p}{1-q} = \frac{p}{p} = 1 \tag{2}$$

nella prcedente equazione definiremo "la serie geometrica di ragione g".

9 Lezione 8

Nella lezione precedente abbiamo introdotto alcune variabili aleatorie nell'esperimento ripetuto (Bernulli:

- X_n - prova di ordine n (guarda a tempo n, l'istante della prova).
- S_n - somma dele prime n prove (guarda a tempo n e al passato, conta i successi).
- T_1 - numero delle prove per osservare il primo successo (guarda al futuro)

9.1 Variabile aleatoria (Generalizzazione geometrica)

Domanda determinare il numero delle prove necessarie per assumere il k-esimo successo per la prima volta. Quindi consideriamo:

$k \in \mathbf{N}$ - numero delle prove con successo.

Risposta entra in gioco la seguente variabile aleatoria:
W_k - numero delle prove con il quale si sono visti k successi.

Nota bene, vale almeno k (almeno k prove per k successi).

9.1.1 Formula

$S_{W_k} = \{k, k+1, k+2, ..\}$ l'ordine è pari al numerabile ($\mathbf{N}$) infatti se si ricorda nelle lezioni precedenti $S_{T_1} = \mathbf{N}$ (insieme dei numeri interi).

Ricodiamo T_1: $S_{T_1} = \mathbf{N} = \{1, 2, ...\}$

Prendiamo questa sequenza di esperimenti:

$$C_1...Cn-1|T_n$$

primo successo a tempo n

avremo che:
$n \in \mathbf{N},\ \mathcal{P}(T_1 = n) = p \cdot (1-p)^{n-1}$

quindi sapendo che ciò che è stato cisto precedentemente vale per T_1 vediamo come valper W_k (generalizzando per k elementi).

Per T_1 vale va il seguente insieme: $\{S_{n-1} = 0\} \cap \{X_n = 1\}$

Ciò che noi vogliama che accada è il seguente insieme:
$\{S_{n-1} = k - 1\} \cap \{X_n = 1\}$

Probabiltà per W_k è la seguente:

$$\begin{aligned}\mathcal{P}(W_k) &= \mathcal{P}(\{S_{n-1} = k-1\} \cap \{X_n = 1\}) \\ &= \mathcal{P}(S_{n-1} = k-1) \cdot \mathcal{P}(X_n = 1) \\ &= \binom{n-1}{k-1} \cdot p^{k-1} \cdot (1-p)^{n-k} \cdot p \\ &= \binom{n-1}{k-1} \cdot p^{k} \cdot (1-p)^{n-k}\end{aligned}$$

alla fine possiamo riassumere anche W_k "somigliante" al Pascal (come anche T_1):

- $W_k \sim Pascal(k, p)$
- $T_1 \sim Pascal(1, p) \sim G(p)$

9.1.2 Vari esempi

Consideriamo la seguente sequenza:

C C C T | C C C C C T

Se consideriamo T_1 arriveremo al conteggio fino alla 4 posizione, se consideriamo il successivo T_1 resettiamo il contatore e avremo 6.

Definizione W_k è la somma di k numeri aleatori geometrici **INDIPENDENTI** e **SOMIGLIANTI**.

Esempio Consideriamo queste due sequenze:

k=4, T C C T C T C C C T C C T...
k=3, C C C C T T C C C C C C T C T C C T...

Trasformiamo in termini probabilistici il risultato di W_k delle due sequenze:

- $10 = 1 + 3 + 2 + 4 = t_1^{(1)} + t_1^{(2)} + t_1^{(3)} + t_1^{(4)}$
- $13 = 5 + 1 + 7 = t_1^{(1)} + t_1^{(2)} + t_1^{(3)}$

Un ulteriore esempio è il seguente:
$W_k \sim Pascal(5, \frac{1}{4})$

- $\mathcal{P}(W = 3) = \varnothing$ - non è possibile avere 5 successi in 3 prove.
- $\mathcal{P}(W = \pi) = \varnothing$ - non è possibile se il numero di prove non è intero.
- $\mathcal{P}(W = 7) = \binom{6}{4} \cdot p^5 \cdot (1-p)^2$
- $\mathcal{P}(W = 5) = \binom{4}{4} \cdot p^5 \cdot (1-p)^0$

9.2 Risultati

$A, B \in \mathcal{F},\ A \cap B = \varnothing \Longrightarrow \mathcal{P}(A \uplus B) = \mathcal{P}(A) + \mathcal{P}(B)$
questa viene detta finita additività.

Teorema 10. ***(inclusione-esclusione a 2 eventi)***
$A, B \in \mathcal{F} \Longrightarrow \mathcal{P}(A \cup B) = \mathcal{P}(A) + \mathcal{P}(B) - \mathcal{P}(A \cap B)$

Dimostrazione. Per dimostrare questo teorema abbiamo bisogno di 2 proprietà:

1. $B = (B \cap \Omega) = B \cap (B \cup A^c) = (B \cap A) \uplus (B \cap A^c)$
   che in termini di probabilità:
   $\mathcal{P}(B) = \mathcal{P}(B \cap A) + \mathcal{P}(B \cap A^c)$

2. $A \cup B = A \cup (B \cap A^c)$

adesso è possibile calcolare la probabilità:

$$\begin{aligned}\mathcal{P}(A \cup B) &\overset{(2)}{=} \mathcal{P}[A \uplus (B \cap A^c)] \\ &= \mathcal{P}(A) + \mathcal{P}(B \cap A^c) \\ &\overset{(1)}{=} \mathcal{P}(A) + \mathcal{P}(B) - \mathcal{P}(B \cap A)\end{aligned}$$

□

Teorema 11. ***(inclusione-esclusione a 3 eventi)***
$A, B, C \in \mathcal{F} \Longrightarrow \mathcal{P}(A \cap B \cap C) = \mathcal{P}(A) + \mathcal{P}(B) + \mathcal{P}(C) - [\mathcal{P}(A \cap B) + \mathcal{P}(A \cap C) + \mathcal{P}(B \cap C)] + \mathcal{P}(A \cap B \cap C)$

Dimostrazione. Visibile su Wikipedia[1] □

Teorema 12. ***(inclusione-esclusione a n eventi)***
$(A_1, ..., A_n) \subseteq \mathcal{F}$
$A_1, ..., A_n \in \mathcal{F}$

$$\begin{aligned}\mathcal{P}(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mathcal{P}(A_i) - ... \\ - \sum_{i=1}^{n} \sum_{j=i+1}^{n} \mathcal{P}(A_i A_j) + ... \\ + (-1)^{n-1} \mathcal{P}(A_1, ..., A_n)\end{aligned}$$

è possibile tradurre doppia sommatoria (o più) in un unica sommatoria a parte che per ogni indice i da 1 a n, $i_k < i_{k+1}$

[1] `https://it.wikipedia.org/wiki/Principio_di_inclusione-esclusione`

10 Lezione 9

Iniziamo la lezione tenendo presente la **Formula inclusione-esclusione** vista nella lezione precedente.

10.1 Problema: CONCORDANZE

Ci sono n cartonicini numerati sul fronte che vengono mischiati. Poi il mazzetto viene posto sul tavolo e una persona ne prende uno che si trova più in alto. Poi una persona prende tale cartoncino pronunciando "uno" e lo presenta con il fronte sul tavolo accato al mazzetto.

Se un cartoncino punta il numero "uno" allora si dice che si è presentata "**una concordanza**". Il procedimento si ripete con la persona che pronuncia "due" e rovescia la carta attualmente più in alto del mazzetto. C'è concorrenza se il cartoncino presenta il "due" e così via...

Se $k \in \{1, 2, ..., n\}$ determinare la probabilità dell'evento:

$E_{k,n}$: **"esattamente" k concordanze nelle n chiamate**

10.2 Soluzione: CONCORDANZE

Sia M_n il numero aleatorio che rappresenta il numero delle concordanze su n-chiamate:
$k = 0, 1, ..., n$
$E_{k,n} = \{M_n = k\}$

risolviamo il problema con il caso 0 (nessuna concordanza) e con il caso k (k concordanze).

10.2.1 Caso 1: k=0

Definiamo:
$i = 1...n,\ A_i$ = **"Concordanza nell'i-esima chiamata"**
si vuole studiare quindi (attraverso la formula inclusione-esclusione) la seguente equazione:

$$C = \bigcup_{i=1}^{n} A_i \tag{1}$$

calcoliamo la probabilità:

$$\begin{aligned}\mathcal{P}(C) &= \mathcal{P}(A_i)\\ &= [\sum_{i=1}^{n} \mathcal{P}(A_i)] -\\ &- [\sum_{i<j}^{n} \mathcal{P}(A_i \cap A_j)] +\\ &+ [\sum_{i<j<k}^{n} \mathcal{P}(A_i \cap A_j \cap A_k] + \ldots +\\ &+ (-1)^{n-1} \cdot [\mathcal{P}(\bigcap_{i=1}^{n} A_i)]\end{aligned}$$

Vediamo:

(i) $\mathcal{P}(A_n) = \frac{(n-1)!}{n!}$

(ii) $\mathcal{P}(A_i A_j) = \frac{(n-2)!}{n!}$

(iii) $\mathcal{P}(A_i A_j A_k) = \frac{(n-3)!}{n!}$

analizziamo ogni caso uno per uno.

Caso: (i) sostituiamo alla prima sommatoria la prima soluzione(i):

$$\begin{aligned}
\sum_{i=1}^{n} \mathcal{P}(A_i) &= \\
&= \sum_{i=1}^{n} \frac{(n-1)!}{n!} = \\
&= \frac{(n-1)!}{n!} \sum_{i=1}^{n} 1 = \\
&= \frac{n(n-1)!}{n!} = \frac{n!}{n!} = 1
\end{aligned}$$

risolviamo quindi che la prima sommatoria da come risultato 1.

Caso: (ii) sostituiamo alla seconda sommatoria la seconda soluzione(ii):

$$\begin{aligned}
\sum_{i<j=1}^{n} \mathcal{P}(A_i A_j) &= \\
&= \sum_{i<j=1}^{n} \frac{(n-2)!}{n!} = \\
&= \frac{(n-2)!}{n!} \sum_{i<j=1}^{n} 1 = \\
&= \frac{n(n-1)(n-2)!}{2n!} = \frac{1}{2}
\end{aligned}$$

risolviamo quindi che la seconda sommatoria da come risultato $\frac{1}{2}$.

Esempio
$n = 6,\ Totale = 15$

$$\frac{n(n-1)}{2}$$

Caso: (iii) sostituiamo alla terza sommatoria la terza soluzione(iii):

$$\begin{aligned}
\sum_{i<j<k=1}^{n} \mathcal{P}(A_i A_j A_k) &= \\
&= \sum_{i<j<k=1}^{n} \frac{(n-3)!}{n!} = \\
&= \frac{(n-3)!}{n!} \sum_{i<j<k=1}^{n} 1 = \\
&= \binom{n}{3} \cdot \frac{(n-3)!}{n!} \\
&= \frac{\cancel{n}! \cdot (n \cancel{-} 3)!}{3! \cdot (n \cancel{-} 3)! \cdot \cancel{n}!} = \frac{1}{3}
\end{aligned}$$

risolviamo quindi che la terza sommatoria da come risultato $\frac{1}{3}$.

Tips
$\binom{n}{1} = 1 = \sum_{i=1}^{n} 1$
$\binom{n}{2} = \frac{n!}{2!(n-2)!} = \frac{n \cdot (n-1) \cdot \cancel{(n-2)!}}{2 \cdot \cancel{(n-2)!}}$

quindi concludiamo prendendo in considerazione la **FORMULA INCLUSIONE-ESCLUSIONE** e sostituiamo i termini che abbiamo trovato dai precedenti 3 risultati.

$$\mathcal{P}(C) = \mathcal{P}(\bigcup_{i=1}^{n} A_i) = 1 - \frac{1}{2} + \frac{1}{3!} + \dots + (-1)^{n-1} \cdot \frac{1}{n!} \tag{2}$$

possiamo riassumere il tutto:

$$\mathcal{P}(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} (-1)^{i-1} \cdot \frac{1}{i!} \tag{3}$$

calcoliamo la negata (ovvero il caso che abbiamo preso in considerazione:

$$(\bigcup_{i=1}^{n} A_i)^c = \{M_n = 0\} \tag{4}$$

la probabilità:

$$\mathcal{P}(M_n = 0) = 1 - \mathcal{P}(\bigcup_{i=1}^{n} A_i) = \tag{5}$$

$$= 1 - \sum_{i=1}^{n} (-1)^{i-1} \cdot \frac{1}{i!} = \tag{6}$$

$$= \sum_{i=2}^{n} (-1)^{i} \cdot \frac{1}{i!} \tag{7}$$

e se n tende all'infinito? consideriamo:

$$\lim_{n \to \infty} \mathcal{P}(M_n = 0) = \frac{1}{e} = e^{-1}$$

infatti:

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots \Longrightarrow e^{-1} = 1 + \frac{(-1)}{1!} + \frac{1}{2!} - \frac{1}{3!} + \dots \tag{8}$$

10.2.2 Caso 2: k>0

Continuiamo la dimostrazione della precedente lezione, abbiamo avuto dalle passate considerazioni il seguente risultato:

$$\mathcal{P}(E_{0,n}) = \frac{\#\text{mischiate con 0 concordanze nelle n chiamate}}{n!}$$

$n \in \mathbf{N}$, il numeratore di tale frazione è uguale a $n! \cdot \mathcal{P}(E_{0,n})$

ora si consideri il caso generale:
$k = (0...n)$, $\mathcal{P}(E_{k,n}) = \mathcal{P}(M_n = k)$ e prendiamo in considerazione la seguente successione (C-Concordanze e D-Discordanze):

$$\text{C ... C} \mid \text{D ... D} \Longrightarrow A_{k,n} \text{ (evento)}$$

in questo caso abbiamo per i primi k posti delle concordanze e da k+1 a n delle discordanze: $\{M_{m-k} = 0\} = E_{0,m-k}$.

$$\mathcal{P}(A_{k,n}) = \frac{\#\text{misciate con concidenze in n-k chiamate}}{n!} = \frac{(n-k)! \cdot \mathcal{P}(E_{0,n-k})}{n!}$$

$$\begin{aligned}
\mathcal{P}(M_n = k) = \mathcal{P}(E_{k,n}) &= \binom{n}{k} \cdot \frac{(n-k)! \cdot \mathcal{P}(E_{0,n-k})}{n!} \\
&= \frac{\cancel{n}!}{k! \cdot ((n \cancel{-} k)!} \cdot \frac{(n \cancel{-} k)!}{\cancel{n}!} \\
&= \frac{1}{k!} \cdot \mathcal{P}(M_{n-k} = 0)
\end{aligned}$$

M_n : è un numero aleatorio.
$S_{M_n} = \{0, ..., n\}$
$k \in S_{M_n}$, $\mathcal{P}(M_n = k) = \frac{\mathcal{P}(M_{n-k}=0)}{k!}$

così dimostriamo il caso generale della precedente lezione.

11 Lezione 10

11.1 Problema della ripartizione della posta in gioco

Nelle lezioni precedenti abbiamo definito W_k è il numero delle prove per osservare il k-esimo successo per la prima volta.

Il problema della ripartizione della posta è un quesito antico che viene brillantemente risolto dal famoso matematico francese del Seicento Blaise Pascal. Fu proprio grazie a questa ricerca che Pascal giunse alla teoria del calcolo delle probabilità. Ora, dalle teorie delle probabilità ai giochi online abbiamo già espresso chiaramente la nostra opinione. Però, se supponiamo due giocatori impegnati in una sfida di più partite. Si dà per scontato che ogni partita possa essere vinta da ciascuno dei due sfidanti con le stesse probabilità. Se a un certo punto i due decidono di interrompere il gioco, qual'è il modo più equo per ripartire la posta in gioco?

Considerazioni prima di vedere giusto un paio di casi ricordiamo $S_{W_k} = \{k, k+1, ...\}$ quindi la seguente formula:

$$n \in S_{W_k},\ \mathcal{P}(W_k = n) = \binom{n-1}{k-1} \cdot p^{k-1} \cdot (1-p)^{n-k} \cdot p = \\ = \binom{n-1}{k-1} \cdot p^k \cdot (1-p)^{n-k}$$

Caso 1 Partita interrotta = 5-3 su 64 monete in palio.

$\{W_3 = 3\}$ e $\mathcal{P}(W_3 = 3) = \binom{2}{2} \cdot p^3 \cdot (1-p)^0 = p^3 = \frac{1}{8}$

Caso 2 Partita interrotta = 4-2 su 64 monete in palio.
considerando sempre chi è in svantaggio possiamo considerare le successioni di vincita sui prossimi lanci:

C C C C - $\Longrightarrow \{W_4 = 4\}$
C C C T C $\Longrightarrow \{W_4 = 5\}$
C C T C C $\Longrightarrow \{W_4 = 5\}$
C T C C C $\Longrightarrow \{W_4 = 5\}$
T C C C C $\Longrightarrow \{W_4 = 5\}$

quindi: $\mathcal{P}(W_4 = 4) + \mathcal{P}(W_4 = 5)$, sviluppiamo:

$$\begin{aligned}
&= \binom{3}{3} \cdot p^4 \cdot (1-p)^0 + \binom{4}{3} \cdot p^4 \cdot (1-p)^1 = \\
&= p^4 + 4 \cdot p^4 \cdot (1-p) = \\
&= \frac{1}{16} + {}^2\cancel{4} \cdot \frac{1}{16} \cdot \frac{1}{\cancel{2}_1} = \frac{3}{16} \text{ di monete.}
\end{aligned}$$

11.2 Legge di probabilità di Poission

In teoria delle probabilità la distribuzione di Poisson (o poissoniana) è una distribuzione di probabilità discreta che esprime le probabilità per il numero di eventi che si verificano successivamente ed indipendentemente in un dato intervallo di tempo, sapendo che mediamente se ne verifica un numero λ.

Consideriamo una successione di schemi binomiali ognuno con probabilità "p" e un numero positivo λ.

in generale conosciamo:

$$\mathcal{P}(S_n = k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

consideriamo adesso $n \to \infty$:

$$\lim_{n\to\infty} \mathcal{P}(S_n = k) = \lim_{n\to\infty} \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

considerando $k \in \{0...\infty\},\ k \in \mathbf{N_0}$, aggiungiamo:

- $p \to 0$
- $n \cdot p \to \lambda$

quindi:

$$\lim_{n\to\infty} n \cdot p = \lambda$$

per ogni schema bionomiale consideriamo $\frac{\lambda}{k}$ ma il rapporto è >1 (per numeri $<$ k).

la tendenza è pari a π (3.14).

Per considerare ∞ probabilità:

$$\lim_{n\to\infty} \binom{n}{k} \cdot p_m^k \cdot (1-p_m)^{n-k}$$

al posto di utilizzare il limite per $n \cdot \lambda$:
$n \in \mathbf{N},\ n \cdot p_n = \lambda \iff p_n = \frac{\lambda}{n}$
quindi in conclusione:

$$\lim_{n\to\infty} \binom{n}{k} \cdot (\frac{\lambda}{n})^k \cdot (1-\frac{\lambda}{n})^{n-k}$$

11.2.1 Calcoliamo

$$\lim_{n\to\infty}\left[\frac{n!}{k!(n-k)!}\cdot\frac{\lambda^k}{n^k}\cdot\frac{(1-\frac{\lambda}{n})^n}{(1-\frac{\lambda}{n})^k}\right] \tag{1}$$

fuori dal segno di limite:

$$\frac{\lambda^k}{k!}\cdot\lim_{n\to\infty}\left[\frac{(n-0)(n-1)...(n-k+1)\ \cancel{(n-k)}}{\cancel{(n-k)}}\cdot\frac{1}{n^k}\cdot\frac{(1-\frac{\lambda}{n})^n}{(1-\frac{\lambda}{n})^k}\right] \tag{2}$$

consideriamo i limiti:

$$\frac{\lambda^k}{k!}\cdot\lim_{n\to\infty}\left[\frac{(n-0)(n-1)...(n-k+1)}{n\cdot\dots\cdot n}\cdot\frac{\lim_{n\to\infty}(1-\frac{\lambda}{n})^n}{\lim_{n\to\infty}(1-\frac{\lambda}{n})^k}\right] \tag{3}$$

il primo limite e la frazione di limiti danno come risultato 1, quindi procedendo con vari calcoli (si veda la fine di queso paragrafo):

$$\frac{\lambda^k}{k!}\cdot\lim_{n\to\infty}\left[(1-\frac{\lambda}{n})^{-\frac{n}{\lambda}}\right]^{-\lambda} \tag{4}$$

$$\frac{\lambda^k}{k!}\cdot\lim_{n\to\infty}\left[(1+\frac{\lambda}{n})^{-\frac{n}{\lambda}}\right]^{-\lambda} \tag{5}$$

$$\lim_{n\to\infty}(1+\frac{1}{n})^n = e \textbf{ (numero di Nepero)} \tag{6}$$

Nota bene

- $(a^b)^c = a^{bc} \Longrightarrow (-\frac{n}{\lambda})^{-\lambda} = -\frac{n}{\lambda}^{(-\lambda)} = n$
- $\lim_{n\to\infty} a_n^{\lambda}$=$[\lim_{n\to\infty} a_n]^{\lambda}$

11.2.2 Conclusioni

Si dimostra che tale successione è:

- Crescente
- Limitata

quindi **Convergente** pari a: $\frac{\lambda^k}{k!}\cdot e^{-\lambda}$ (*).

Un numero aleatorio **X** che ha spettro $\mathbf{N_0}$ e relative probabilità nella (*) su dice avere legge di Possion con parametro λ e si scrive:

$$X \sim \Pi(\lambda)$$
$$k \in \mathbf{N_0},\ \mathcal{P}(X = k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!},\ \lambda > 0$$

La legge di Possion è detta "legge dei piccoli numeri" o "degli eventi rari". Ciò è dovuto alla rapida decrescita di $e^{-\lambda} \cdot \frac{\lambda^k}{k!}$ al crescere di k. Ciò, ovviamente, è dato dal fattoriale di k che compare nel denominatore.

12 Lezione 11

12.1 Differenza tra un numero deterministico ad esempio $\sqrt{2}$, e un numero aleatorio ad esempio S_{10}

Vediamo un breve riassunto di ciò che abbiamo appreso da precedenti lezioni:

- Numero deterministico: $\sqrt{2} = 1.414...$;
- Numero non deterministico: S_{10} = Somma successi in 10 eventi;

per specificare S_{10} dobbiamo definire l'insieme dei possibili valori:

$\{0, 1, ..., 9, 10\}$ (con probabilità non nulla)

con quale probabilità ognuno di loro può essere assunto:

$k \in \{0, 1, ..., 9, 10\}, \ \mathcal{P}(S_n = k) = \binom{10}{k} \cdot p^k \cdot (1-p)^{10-k} > 0$

queste due cose sono la legge di S_{10}.

12.1.1 Ricapitoliamo le leggi

Legge binomiale $\mathbf{B(n, p)}$ essa può anche essere definita come $\mathbf{G(p)}$

Legge $\mathbf{X_n}$ $X_n, \ S_{X_n} = \{0, 1\}$

- $\mathcal{P}(X_n = 0) = 1 - p$
- $\mathcal{P}(X_n = 1) = p$

Legge $\mathbf{T_1}$ $T_1, \ S_{T_1} = \{1, 2, ...\} \ n \in S_{T_1}$

Probabilità:
$\mathcal{P}(T_1 = n) = p \cdot (1-p)^{n-1}$

Legge Pascal ($\mathbf{W_k}$) $Pascal(k, p), \ W_k$
$S_{W_k} = \{k, k+1, ...\}$
$n \in S_{W_k}$

Probabilità:
$\mathcal{P}(W_k = n) = p \cdot \binom{n-1}{k-1} \cdot p^{k-1} \cdot (1-p)^{n-k}$

Legge di Poisson $\lambda > \mathbf{0} \quad X \sim \Pi(\lambda)$
$S_X = \{0, 1, ...\infty\} = \mathbf{N_0},\ n \in S_X$

Probabilità:
$\mathcal{P}(X = n) = e^{-\lambda} \cdot \frac{\lambda^n}{n!}$ per una successione di $\frac{\lambda}{n} < 1$

Legge delle concordanze ($\mathbf{E_{n,k}}$)

- M_n =non deterministico
- $S_{M_n} = \{0, 1, 2, ...n\}$
- $k \in S_{\{}M_n\}$

Probabilità:
$\mathcal{P}(M_n = k) = \frac{1}{k!} \cdot \sum_{i=2}^{n-k} (-1)^i \cdot \frac{1}{i!}$

è possibile far partire la sommatoria da 0 (ma andrebbero le prime 2 somme a collidere[1]):

$$\sum_{i=0}^{n-k} (-1)^i \frac{1}{i!} = (-1)^0 + (-1)^1 \frac{1}{1!} + \sum i = 2^{n-k} (-1)^i \frac{1}{i!}$$

[1]Visione del file excel proposto dal prof.

12.2 Eventi quasi certi/impossibili

12.2.1 Eventi "quasi" certi

Consideriamo un evento A:TT prima di CC (probabilià p onesta pari a $\frac{1}{2}$):

- $A_1 = T_1T_2$
- $A_2 = C_1T_2T_3$
- $A_3 = T_1C_2T_3T_4$
- $\ldots$
- A_n = Alternanza T ed C $\ldots T_nT_{n+1}$

calcoliamo la probabilità: $\mathcal{P}(A) = \mathcal{P}(A_1)\ldots$ quindi:

$$\begin{aligned}
\mathcal{P}(A) &= \mathcal{P}(T_1T_2) + \mathcal{CTT} + \ldots \\
&= \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \ldots \\
&= \frac{1}{4}(1 + \frac{1}{2} + \frac{1}{4} + \ldots \\
&= \frac{1}{4}\sum_{i=0}^{\infty}(\frac{1}{2}^{i}) \\
&= \frac{1}{4} \cdot \frac{1}{1-\frac{1}{2}} = \frac{1}{4} \cdot \frac{1}{\frac{1}{2}} = \frac{1}{4} \cdot 2 = \frac{1}{2}
\end{aligned}$$

questa viene detta serie di ragione $\frac{1}{2}$ (converge $0 < \frac{1}{2} < 1$)

Consideriamo l'evento opposto (CC prima di TT) e chiamiamolo B:

$$\mathcal{P}(B) = \frac{1}{2}$$

$$\mathcal{P}(A \uplus B) = \mathcal{P}(A) + \mathcal{P}(B) = \frac{1}{2} + \frac{1}{2} = 1$$

notiamo che:

$$A \cup B \neq \{T, C\}^{\mathbf{N}} = \Omega$$

concludiamo che $A \uplus B$ è un evento "quasi" certo (ha la "stessa" possibilità di un evento certo) solo in caso di eventi in cui Ω è numerabile.

12.2.2 Eventi "quasi" impossibili

Iniziamo da un esempio, ovvero una successione (infinita) di alternanza T e C nei due modi diversi:

- U ={(C T C T C...)}
- V ={(T C T C T...)}

si noti che $U \cap V = \varnothing$, $U \uplus V = (A \uplus B)^c$,
la probabilità $\mathcal{P}[(A \cup B)^c] = 1 - \mathcal{P}(A \cup B) = 1 - 1 = 0$

$\mathbf{U} \uplus \mathbf{V} \neq \varnothing$ questa unione disgiunta ha probabilità 0 come visto in precedenza ma non è l'evento impossibile, ovvero non è $\varnothing$, tale evento viene detto evento "quasi" impossibile. Possiamo notare anche i seguenti due eventi:

- $\mathcal{P}(U) = 0$
- $\mathcal{P}(V) = 0$

questi sono altri elementi "quasi" impossibili.

In generale esistono $\mathbf{N_o}$ eventi "quasi" impossibili, esempio, consideriamo la seguente probabilità:

$$\mathcal{P}\{(TCTTCCTTTCCCTTTTCCCC...)\} = 0$$

$$\begin{aligned} U \subseteq U \uplus V \Longrightarrow \mathcal{P}(U) &\leq \mathcal{P}(U \uplus V) \\ \mathcal{P}(U) &\leq 0 \\ \mathcal{P}(V) &\leq 0 \\ \mathcal{P}(U) &= 0 = \mathcal{P}(V) \end{aligned}$$

A seconda di dove ci fermiamo abbiamo una successione decrescente:

$$\frac{1}{2} > \frac{1}{4} > \cdots > \frac{1}{2^{10}} > \frac{1}{2^n}$$

consideriamo:

- $\mathcal{P}(\ldots|) = \frac{1}{2^n} \to 0 -$ (limite esterno $\lim \mathcal{P}(\ldots|)$)
- $\mathcal{P}(\ldots) =$ (limite interno $\mathcal{P}(\lim \ldots)$)

Esempi $E_1, E_2, ... E_n, ...$

$$\lim_n \mathcal{P}(E_n) = \mathcal{P}(\lim_n E_n)$$

più in generale, tutti i singoletti di $\{T, C\}^{\mathbf{N}}$, ovvero tutti i sottoinsiemi dello spazio campione contenenti una sola successione di T e C, sono eventi "quasi" impossibili.

Una maniera alternativa per verificare che U e V sono quasi impossibili è la seguente:

$$0 = \mathcal{P}(U \cap V) = \mathcal{P}(U) + \mathcal{P}(D) \Longrightarrow \mathcal{P}(E) = \mathcal{P}(D) = 0$$

Nota bene un evento "quasi" impossibile è anche un evento "quasi" certo.

13 Lezione 12

La lezione inizia con il riprendere un discorso generale sugli eventi "quasi" certi ed eventi "quasi" impossibili. Si consiglia di rivedere la precedente lezione per capire.

13.1 Eventi indipendenti

Consideriamo la definizione di **eventi incompatibili o disgiunti**:

$$A \uplus B \Longleftrightarrow A \cup B = \varnothing$$

la definizione invece di **indipendeza** tra eventi è definita in questo modo:

Due eventi, A e B, si dicono ***indipendenti*** *se e solo se:*

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

la probabilità dell'evento congiunto si fattorizza nel prodotto delle probabilità dei singoli eventi.

Esempio consideriamo T_1 e T_2 essendo lanci di moneta indipendenti possiamo concludere velocemente che essi sono eventi indipendenti:

$$\mathcal{P}(T_1 \cap T_2) = \mathcal{P}(T_1) \cdot \mathcal{P}(T_2)$$

per la precisione essendo probabilità "oneste" (pari a $\frac{1}{2}$) la definizione combaia.

Se si può determinare un insieme di due sottoesperimenti che sono ***fisicamente*** *indipendenti allora tutti gli eventi che sono generati dal primo di esso sono indipendenti da ciascun evento generato dal secondo sottoesperimento.*

13.2 Teoremi

Teorema 13. *Siano* $A, I \in \mathcal{F}$ *con* $\mathcal{P}(I) = 0$ *allora* :
$\mathcal{P}(A \cap I) = \mathcal{P}(A) \cdot \mathcal{P}(I)$

Dimostrazione.

$$\begin{aligned} A \cap I \subseteq I &\Longrightarrow \mathcal{P}(A \cap I) \leq \mathcal{P}(I) = 0 \\ &\Longrightarrow \mathcal{P}(A \cap I) = 0 \\ &\mathcal{P}(A \cap I) = 0 \\ &\mathcal{P}(A) \cdot \mathcal{P}(I) = 0 \end{aligned}$$

□

Teorema 14. *Siano* $A, I \in \mathcal{F}$ *con* $\mathcal{P}(C) = 1$ *allora* :
$\mathcal{P}(A \cap C) = \mathcal{P}(A) \cdot \mathcal{P}(C)$

Dimostrazione. Consideriamo prima le seguenti proprietà:

- $A = A \cap \Omega = A \cap (C \cup C^c)$ per distributività $(A \cap C) \cup (A \cap C^c?$
- $\mathcal{P}(C^c) = 0$

fatte le precedenti premesse, dimostriamo:

$$\begin{aligned} \mathcal{P}(A) &= \mathcal{P}(A \cap C) + \mathcal{P}(A \cap C^C) \\ &= \mathcal{P}(A \cap C) + 0 \end{aligned}$$

quindi

$$\begin{aligned} \mathcal{P}(A) = \mathcal{P}(A \cap C) &\Longleftrightarrow \mathcal{P}(A \cap C) = \mathcal{P}(A) \cdot 1 \\ &\Longleftrightarrow \mathcal{P}(A \cap C) = \mathcal{P}(A) \cdot \mathcal{P}(C) \end{aligned}$$

□

13.3 Esempi di indipendenza

13.3.1 Lancio di un dado onesto (6 facce) - 2 eventi

Semplicemente si prendano in considerazione i seguenti eventi:

- $A = \{2, 4, 6\}$
- $B = \{1, 2\}$

calcoliamo le probabilità singole degli eventi:

- $\mathcal{P}(A) = \frac{3}{6}$
- $\mathcal{P}(B) = \frac{2}{6}$

per dire che A e B sono indipendenti calcoliamo anche la probabilità dell'intersezione e vediamo se vale la fattorizzazione:

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

effettivamente se moltiplichiamo tra loro le probabilità (A e B) otteremo $\frac{1}{6}$.

13.3.2 Lancio di un dado onesto (4 facce) - 3 eventi

$\Omega = \{1, 2, 3, 4\}$
$g = \{\{1\}\{2\}\{3\}\{4\}\}$
$\mathcal{P} = A \in g = \frac{1}{4}$

consideriamo:

- $A = \{1, 2\}$
- $B = \{1, 3\}$
- $C = \{1, 4\}$

le probabilità dei singoli eventi sono:
$\mathcal{P}(A) = \frac{2}{4} = \mathcal{P}(B) = \mathcal{P}(C)$

consideriamo le probabilità degli insiemi a 2 a 2 in intersezione:

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

a sua volta la fattorizzazione sarà pari a:

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

la stessa logica va applicata anche agli altri casi e ci troveremo che tutti hanno lo stesso risultato.

Infine calcoliamo la probabilità dell'intersezione tra tutti e 3 gli insiemi e controlliamo la fattorizzazione:

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

la fattorizzazione sarà:

$$\mathcal{P}(A) \cdot \mathcal{P}(B) \cdot \mathcal{P}(C) = \frac{2}{4} \cdot \frac{2}{4} \cdot \frac{2}{4} = \frac{1}{8}$$

non sono **collettivamente** indipendenti.

13.4 Collettivamente indipendenti

Visti i precedenti esempi possiamo definire la collettiva indipendenza:

Per 3 eventi:
Quindi 3 eventi sono indipendenti se si fattorizza la probabilità dell'intersezione di 2 qualsiasi di essi e la probabilità dell'intersezione dei tre eventi

Per n eventi:
Quindi n eventi sono indipendenti se si fattorizza la probabilità di k degli eventi con l = $(2, 3, ..., n)$ *e comunque si scelgano i k eventi.*

13.5 Conclusione e introduzione alla prossima lezione

Consideriamo l'esempio di un urna con 20 palline, tra cui 10 bianche e 10 nere. L'esperimento viene visto in due funzioni differenti, la prima consiste in 2 pescaggi dove però la prima pallina va rimessa all'interno dell'urna, il secondo invece dopo il primo pescaggio la pallina non va rimessa nell'urna.

Caso 1 Consideriamo la seguente probabilità:
$\mathcal{P}(B_1 \cap N_1) = \mathcal{P}(B_1) \cdot \mathcal{P}(N_2)$
la fattorizzazione secondo l'esperimento sarà pari a:
$\frac{10}{20} \cdot \frac{10}{20}$

Caso 2 Consideriamo la seguente probabilità:
$\mathcal{P}(B_1 \cap N_1) = \mathcal{P}(B_1) \cdot \mathcal{P}(N_2)$
si può intuire che eliminando una pallina si dice che "l'urna ha memoria" e quindi la probabilità (a seconda di che colore è la pallina pescata per prima) cambia e cio ci fa definire che:
$\mathcal{P}(B_1 \cap N_1) \neq \mathcal{P}(B_1) \cdot \mathcal{P}(N_2)$

14 Lezione 13

14.1 Introduzione

In generale: **Proposizione**:

Siano A e B eventi indipendenti, valgono le seguenti proprietà:

(i) AeB^c sono indipendenti.

(ii) A^ceB sono indipententi.

(iii) A^ceB^c sono indipendeti.

Dimostrazione. La dimostrazione va fatta per il punto (i) e di conseguenza applicata alle coppie del punto (ii) e (iii):

Formula di decomposizione evento certo :
$\mathcal{P}(A) = \mathcal{P}(A \cap B) + \mathcal{P}(A \cap B^c)$

$$\begin{aligned}\mathcal{P}(A \cap B^c) &= \mathcal{P}(A) - \mathcal{P}(A \cap B)\\ &= \mathcal{P}(A) - \mathcal{P}(A) \cdot \mathcal{P}(B)\\ &= \mathcal{P}(A)[1 - \mathcal{P}(B)]\\ &= \mathcal{P}(A) \cdot \mathcal{P}(B^c)\end{aligned}$$

infatti:

$$\begin{aligned}\mathcal{P}(A) = \mathcal{P}(A \cap \Omega) &= \mathcal{P}[A \cap (B \cup B^c)]\\ &= \mathcal{P}((A \cap B) \cup (A \cap B^c))\\ &= \mathcal{P}(A \cap B) + \mathcal{P}(A \cap B^c)\end{aligned}$$

□

Generalizziamo Siano $A_1, ..., A_n$ eventi indipendenti//
(comunque si sceglie k=(2,3,...,n) e comunque si scelgano k di essi, la probabilità della loro intersezione si fattorizza nel prodotto delle rispettive probabilità)

Comunque si sceglie k=(1,2,...,n) e comunque si scelgano k di essi, sostituendo a cessi i loro rispettivi complementi si ottiene un altra n-upla di eventi indipententi.

Esempio se $A_1A_2A_3A_4$ sono indipendeti allora (scegliendo k da 1 a 4) e sostituendo ai rispettivi eventi i complementi, avremo sempre l'indipendenza (i.e k=1 -> $A_1^cA_2A_3A_4$.

14.2 Probabilità condizionata

14.3 Esperimento

Consideriamo il seguente esperimento:
Lancio di un dado nella stanza T_1 e lancio di un altro dado nella stanza T_2 e si consideri S la somma dei due punteggi (i dadi sono onesit).

Nel caso generale defniremo:

- $\Omega = \{(i,j)\}$ con i,j=1...6, $g = (\{i,j\})$.
- $\sigma(g) = P(\Omega)$.
- $\mathcal{P}$ è quella classica.

consideriamo A: "S è multiplo di 4"

risulta:

$\mathcal{P}(A) = \{$unione disgiunta delle coppie che danno somma 4, 8 o 12$\} = \mathcal{P}(\{(1,3\} + \cdots + \mathcal{P}(\{(6,6)\})$
9 addendi uguali $\frac{1}{36} = \frac{9}{36}$

Conclusione:

- $(\Omega, \mathcal{F}, \mathcal{P})$
- $\mathcal{P}_\Omega(A) = \frac{\mathcal{P}(A)}{\mathcal{P}(\Omega} = \mathcal{P}(A)$

Si supponga che prima della valutazione di A si venga a conoscenza che il dado nella stanza T_2 abbia dato esito "3" e quindi definire $\{3\} = B$ (indica lo spazio campione delle coppie con seconda coordinata 3).

Allora $\mathcal{P}_B(A) = \mathcal{P}(1,3) + \mathcal{P}(5,3)$

con l'informazione cambia la terna: $(B, \mathcal{F}_B, \mathcal{P}_B$

$$\mathcal{P}_B(A) = \frac{\mathcal{P}(\{1,3\}) \cup \mathcal{P}(\{(5,3)\})}{\mathcal{P}(B)} = \frac{\mathcal{P}(1,3) + \mathcal{P}(5,3)}{\mathcal{P}(B)} = \frac{\frac{2}{36}}{\frac{6}{36}} = \frac{12}{36}$$

14.3.1 Considerazioni

Per le considerazioni precedenti:

- Con le informazioni: $\mathcal{P}_B(A) = \frac{\mathcal{P}(A \cap B}{\mathcal{P}_B(A)}$
- Senza le informazoni: $\mathcal{P}_\Omega(A) = \frac{\mathcal{P}(A \cap B)}{\mathcal{P}(\Omega)} = \mathcal{P}(A)$

Nota bene la probabilità condizionata possono essere <, = o > della probabilità del caso base (in questo caso A), **NON** esiste una relazione. Essa inoltre è valida solo se B non è un evento impossinile o "quasi" dato che aritmeticamente il denominatore non può essere 0 ($\mathcal{P}(\varnothing) = 0$.

14.3.2 Definizione

$A, B \in \mathcal{F} e B : \mathcal{P}(B) > 0$. Si definisce:

$$\mathcal{P}_B(A) := \frac{\mathcal{P}(A \cap B}{\mathcal{P}(B)}$$

viceversa se $A : \mathcal{P}(A) > 0$. Si definisce:

$$\mathcal{P}_A(B) := \frac{\mathcal{P}(A \cap B)}{\mathcal{P}(A)}$$

Osservazione dalle considerazioni ed esempi precedenti mostra che appunto non esiste relazione d'ordine tra $\mathcal{P}(A)$ e $\mathcal{P}_B(A)$.

14.4 La legge della probabilità congiunta

- Formula 1 - $\mathcal{P}_\Omega(A \cap B) = \mathcal{P}(B) \cdot \mathcal{P}_B(A)$
- Formula 2 - $\mathcal{P}_\Omega(A \cap B) = \mathcal{P}(A) \cdot \mathcal{P}_A(B)$

quindi concludiamo che:

- A e B ind. $\mathcal{P}(B) \cdot \mathcal{P}(A) \longrightarrow \mathcal{P}_B(A) = \mathcal{P}(A)$
- A e B ind. $\mathcal{P}(A) \cdot \mathcal{P}(B) \longrightarrow \mathcal{P}_A(B) = \mathcal{P}(B)$

14.4.1 Esempio palline (lezione precedente)

$B_1 = Bianca(posto1)$
$N_2 = Nera(posto2)$

$$\begin{aligned}\mathcal{P}(B_1 N_2) &= \mathcal{P}(B_1) \cdot \mathcal{P}_{B_1}(N_2)\\ &= \frac{10}{20} \cdot \frac{10}{19} = \frac{5}{19}\end{aligned}$$

Esempio dei giocatori Siano $G_1 ... G_5$ 5 giocatori e sia presente un'unra contenente 4 palline **Bianche** e una pallina **Nera**. I giocatori (in ordine di numerazione) pescano dall'urna, chi pesca la pallina nera vince.

Calcoliamo le probabilità dei vari giocatori:

Giocatore 1:

$$\mathcal{P}(G_1) = \mathcal{P}(N_1) = \frac{1}{5}$$

Giocatore 2:

$$\begin{aligned}\mathcal{P}(G_2) = \mathcal{P}(B_1 \cap N_2) &= \mathcal{P}(B_1) \cdot \mathcal{P}_{B_1}(N_2)\\ &= \frac{4}{5} \cdot \frac{1}{4} = \frac{1}{5}\end{aligned}$$

Giocatore 3:

$$\begin{aligned}\mathcal{P}(G_3) &= \mathcal{P}(B_1 \cap B_2 \cap N_3)\\ &= \mathcal{P}(B_1) \cdot \mathcal{P}_{B_1}(B_2) \cdot ?\\ &= \ldots \mathcal{P}_{B_1 \cap B_2}(N_3)\\ &= \frac{4}{5} \cdot \frac{3}{4} \cdot \frac{1}{3}\end{aligned}$$

14.4.2 Proposizione

$A, B, C \ : \ \mathcal{P}(B \cap C) > 0$

$(\mathcal{P}_B)_C(A) = \mathcal{P}_{B\cap C}(A)$

14.5 Dimostrazione

Dimostrazione. Poniamo $Q = \mathcal{P}_B$:

$$
\begin{aligned}
(\mathcal{P}_B)_C(A) = Q_C(A) &= \frac{Q(A \cap C}{Q(C)} && (1)\\
&= \frac{\mathcal{P}_B(A \cap C)}{\mathcal{P}_B(C)} && (2)\\
&= \frac{\frac{\mathcal{P}(A\cap C\cap B)}{\mathcal{P}(B)}}{\frac{\mathcal{P}(B\cap C)}{\mathcal{P}(B)}} && (3)\\
&= \frac{\mathcal{P}(A \cap B \cap C)}{\mathcal{P}(B \cap C)} && (4)\\
&= \mathcal{P}_{B\cap C}(A) && (5)
\end{aligned}
$$

□

14.5.1 Conclusione

Tornando all'esempio dei giocatori e avedo fatto la precedente dimostrazione, conclufiamo i calcoli:
$G_1 \uplus ... \uplus G_5 = \Omega$

Giocatore 4:

$$
\begin{aligned}
\mathcal{P}(G_4) &= \mathcal{P}(B_1 B_2 B_3 N_4)\\
&= \mathcal{P}(B_1) \cdot \mathcal{P}_{B_1}(B_2) \cdot \mathcal{P}_{B_1\cap B_2}(B3) \cdot \mathcal{P}_{B_1\cap B_2\cap B_3}(N_4)\\
&= \frac{4}{5} \cdot \frac{3}{4} \cdot \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{2}
\end{aligned}
$$

Giocatore 5:

$$
\begin{aligned}
\mathcal{P}(G_5) = \mathcal{P}(B_1 B_2 B_3 B_4 N_5) &= \mathcal{P}(G_1^c G_2^c G_3^c G_4^c)\\
&= \mathcal{P}((G_1 \uplus G_2 \uplus G_3 \uplus G_4))^c\\
&= 1 - \mathcal{P}(G_1 \uplus G_2 \uplus G_3 \uplus G_4)\\
&= 1 - \frac{4}{5} = \frac{1}{5}
\end{aligned}
$$

Osservazione finale Un interpretazione **"frequentista"** del risultato ottenuto è la seguente:

su ogni 100 partite, G1 le gioca tutte e "mediamente", ne vince 20; G2 allora, mediamente, ne gioca 80 e ne vince 20; G3, mediamente, ne gica 60 e ne vince 20; G4, mediamente ne gioca 40 e ne vince 20; G5, mediamente ne gioca 20 e ne vince 20.

Infatti, l'urna di G1 gli conente di vincere con probabilità $\frac{1}{5}$ quella di G2 con $\frac{1}{4}$, G3 con $\frac{1}{3}$, G con $\frac{1}{2}$ e infine per G la probabilità è 1.

15 Lezione 14

15.1 Esempio e generalizzazione

Introduciamo la lezione consiferando un caso generale dell'esercizio visto in precedenza dei giocatori con un urna composta da k palline bianche e una pallina nera, quindi:

- Giocatori: 2
- Palline bianche: k
- Palline nere: 1

Si vuole calcolare la probabilità che giocatore 1 vinca (da notare che il pescaggio (in caso di pallina bianca pescata) si alterna tra i dude giocatori, quindi il giocatore 1 farà sempre pescate dispari, mentre il giocatore 2 sempre quelle pari:

$G_{1,1}$ la probabilità è la seguente:
$\mathcal{P}(G_{1,1}) = \frac{1}{k+1}$ (k+1 sono il numero di palline totali).

$G_{1,3}$ la probabilità è la seguente:

$$\begin{aligned}\mathcal{P}(G_{1,3}) &= \mathcal{P}(B_1B_2N_3)\\ &= \mathcal{P}(B_1)\mathcal{P}_{B_1}(B_2)\mathcal{P}_{B_1B_2}(N_3)\\ &= \frac{k}{k+1}\cdot\frac{k-1}{k}\cdot\frac{1}{k-1} = \frac{1}{k+1}\end{aligned}$$

$G_{1,5}$ la probabilità è la seguente:

$$\begin{aligned}\mathcal{P}(G_{1,5}) &= \mathcal{P}(B_1B_2B_3B_4N_5)\\ &= \mathcal{P}(B_1)\mathcal{P}_{B_1}(B_2)\mathcal{P}_{B_1B_2}(B_3)\mathcal{P}_{B_1B_2B_3}(N_4)\mathcal{P}_{B_1B_2B_3B_4}(N_5)\\ &= \frac{k}{k+1}\cdot\frac{k-1}{k}\cdot\frac{k-2}{k-1}\cdot\frac{k-3}{k-2}\cdot\frac{1}{k-3} = \frac{1}{k+1}\end{aligned}$$

e così via...

quindi:

- $G_1 = G_{1,1} \uplus ... \uplus G_1$ se k è pari (k+1 tiri)
- $G_1 = G_{1,1} \uplus ... \uplus G_1$ se k è dispari (k tiri)

$$\#G_1 = \begin{cases} \frac{k}{2}+1 & \text{if pari} \\ \frac{k+1}{2} & \text{if dispari} \end{cases}$$

suddividiamo i due casi allora:

k dispari

$$\mathcal{P}(G_1) = \mathcal{P}(G_{1,1}) + ... + \mathcal{P}(G_{1,k}$$

in totale avremo (per considerazioni precedenti) $\frac{k+1}{2}$ addendi:

$$= \frac{k+1}{2} \cdot \mathcal{P}(G_{1,1}) \tag{1}$$

$$= \frac{k+1}{2} \cdot \frac{1}{k+1} = \frac{1}{2} \tag{2}$$

k pari

$$\mathcal{P}(G_1) = \mathcal{P}(G_{1,1}) + ... + \mathcal{P}(G_{1,k+1}$$

in totale avremo (per considerazioni precedenti) $\frac{k}{2}+1$ addendi:

$$= \frac{k}{2} + 1 \cdot \mathcal{P}(G_{1,1}) \tag{3}$$

$$= \frac{k}{2} + 1 \cdot \frac{1}{k+1} = \frac{k+2}{2(k+1)} > \frac{1}{2} \tag{4}$$

Infatti se k è pari il giocatore 1 riceve un agio:
$\frac{1}{2} \cdot \frac{k+2}{k+1} > \frac{1}{2}$
$\lim_{k\to\infty} \mathcal{P}(G_1) = \frac{1}{2} \lim_{k\to\infty} \frac{k+2}{k+1} = \frac{1}{2}$

Se ne conclude che per ilmitare il "vantaggio" del giocatore 1, il giocatore 2 chiede di riempire l'urna con il numero massimo di palline bianche disponibili. Vedesi il file excel messo a disposizione dal prof.

15.2 Formula di Bayes

$A, B \in \mathcal{F} : \mathcal{P}(A \cap B) > 0$

- Formula 1 - $\mathcal{P}(A \cap B) = \mathcal{P}(A) \cdot \mathcal{P}_A(B)$
- Formula 2 - $\mathcal{P}(A \cap B) = \mathcal{P}(B) \cdot \mathcal{P}_B(A)$

per definire la formula finale indichiamo:

- $\mathcal{P}_B(A)$ - a posteriori;
- $\mathcal{P}_A(B)$ - verosomiglianza;
- $\mathcal{P}(A)$ - a priori;

a finale:

$$\mathcal{P}_B(A) = \frac{\mathcal{P}_A(B) \cdot \mathcal{P}(A)}{\mathcal{P}(B)}$$

15.2.1 Esempio SERD

S: Portatore di una particolare dipendenza lieve. T: Portatore di una particolare dipendenza grave.

- $\mathcal{P}_S(T)$ - a lungo termine (guarda al futuro);
- $\mathcal{P}_T(S)$ - studiabile (guarda al passato);

$$\begin{aligned} \mathcal{P}_S(T) = \frac{\mathcal{P}_A(B)\mathcal{P}(A)}{\mathcal{P}(B)} &= \frac{\mathcal{P}_T(S)\mathcal{P}(T)}{\mathcal{P}(S)} \\ &= \mathcal{P}(T) \cdot \frac{\mathcal{P}_T(S)}{\mathcal{P}(S)} \end{aligned}$$

La formula precedente si riassume in questo modo:

- $\mathcal{P}_S(T)$ si ottiene "ora" in quanto si chiede ad un certo numero di individui T se nel passato era un individuo con S.
- $\mathcal{P}_S(T)$ si otterrà "in futuro" in quanto individuato un certo numero di individui S sarà necessario aspettare un periodo di tempo per chiedere ad essi se ne frattempo sono diventati individui con T.
- Quindi, se si è in grado di valitare anche il rapporto $\frac{\mathcal{P}_T(S)}{\mathcal{P}(S)}$ la formula di Bayes consente di procedere con la rilevazione di dati v

Ritornando alla formula di Bayes, in generale bisogna valutare $\mathcal{P}(B)$.

Informazioni aggiuntive sul SERD[1]

[1] `https://en.wikipedia.org/wiki/Selective_estrogen_receptor_degrader`

16 Lezione 15

16.1 Formula delle alternative

Un insieme di eventi, $H_1 ... H_n$ si dice costruire "sistema" completo di alternative se fode di 3 proprietà:

- i=(1,...n): $\mathcal{P}(H_1) > 0$
- i,j=(1,...,n): $i \neq j$, $H_i \cap H_j = \varnothing$
- $\dot\bigcup_{i=1}^{n} H_i = \Omega$ (traducibile ance come $\mathcal{P}(\dot\bigcup_{i=1}^{n} H_i) = \sum_{i=1}^{n} \mathcal{P}(H_i)$)

Formula:

$$\mathcal{P}(B) = \sum_{i=1}^{n} \mathcal{P}_{H_i}(B) \cdot \mathcal{P}(H_i)$$

16.2 Dimostrazione

Dimostrazione.

$$\begin{aligned} B &= B \cap \Omega & (1) \\ &\overset{(3)}{=} B \cap (\dot\bigcup_{i=1}^{n} H_i) & (2) \\ &= \dot\bigcup_{i=1}^{n} (B \cap H_i) & (3) \end{aligned}$$

calcoliamo adesso la probabilità di B, sfruttando la **finita additività** e la **legge delle prob. congiunte**:

$$\begin{aligned} \mathcal{P}(B) &= \mathcal{P}[\dot\bigcup_{i=1}^{n} (B \cap H_i)] & (4) \\ &= \sum_{i=1}^{n} \mathcal{P}(B \cap H_i) & (5) \\ &= \sum_{i=1}^{n} \mathcal{P}(H_i)\mathcal{P}(B) & (6) \end{aligned}$$

□

16.3 Esempio articolato

Valutiamo il seguente esperimento:

- Un sacchetto contenente 11 dischi numerati da 0 a 10.
- Un armadietto con 10 palline bianche e 10 nere.
- Un'urna inizialmente vuota.

si pesca dal sacchetto un disco targato (un certo i=(0...10)), si riempie l'urna con i palline bianche e 10-i palline nere.

16.3.1 Svolgimento

i=(0...10), H_i ="esce il dischetto con etichetta i"

Verifichiamo le tre proprietà della formula delle alternative:

- $i \neq j,\ H_i \cap H_j = \varnothing$ - Verificata
- $\dot{\bigcup}_{i=0}^{n} H_i = \Omega$ - Verificata
- $i = (0...10), \mathcal{P}(H_i) = \frac{1}{11} > 0$ - Verificata

In pratica:

$$
\begin{aligned}
\mathcal{P}(B) &= \sum_{i=0}^{10} \mathcal{P}(H_i) \cdot \mathcal{P}_{H_i}(B) \\
&= \sum_{i=0}^{10} \frac{1}{11} \cdot \mathcal{P}_{H_i}(B) = \frac{1}{11} \sum_{i=0}^{10} \mathcal{P}(B) \\
&= \frac{1}{11} \sum_{i=0}^{10} \frac{1}{10} = \frac{1}{110} \sum : i = 0i \\
&= \frac{1}{2}
\end{aligned}
$$

se il banco vuole costruire un **agio** a suo favore può eliminare il dischetto con valore 0:

$$
\begin{aligned}
\sum : i = 1^{10} \frac{1}{10} \mathcal{P}_{H_i}(B) &= \frac{1}{10} \sum : i = 1^{10} \frac{1}{10} \\
&= \frac{1}{100} \cdot \frac{11 \cdot 10}{2} (\frac{1}{10} \sum : i = 1^{10} i) \\
&= \frac{1}{2} \cdot \frac{11}{10} > \frac{1}{2}
\end{aligned}
$$

16.4 Ulteriore esempio

- Si lanciano le monete e si conta il numero X di T ($X \sim B(n,p) \sim S_n$.
- Si escludono le monete che hanno presentato C e si rilanciano le rimanenti. Si conta il numero y di T.

Determinare la legge di y.

16.4.1 Svolgimento

a) $S_y = \{0...n\}$.

b) $y \in Y,\ \mathcal{P}(Y = y) =?$

svolgimento:
$H_0 = \{X = 0\}...H_n = \{X = n\}$

$$\begin{aligned}
\mathbf{P}(Y = y) &= \sum : i = 0^n \mathcal{P}(H_i) \cdot \mathcal{P}_{H_i}(Y = y) \\
&= \sum : i = 0^n \mathcal{P}(X = i) \cdot \mathcal{P}_{H_i}(Y = y) \\
&= \sum : i = 0^n \binom{n}{i} p^i (1-p)^{n-i} \mathcal{P}(Y = y) \\
&= \sum : i = y^n \binom{n}{i} p^i (1-p)^{n-i} \\
&= \sum : i = y^n \binom{n}{i} p^i (1-p)^{n-i} \binom{n}{y} p^y (1-p) i - y
\end{aligned}$$

Risultato finale:

$$\binom{n}{y} (p^2)^y (1-p^2)^{n-y}$$

da cui:
$Y \sim B(n, p^2)$

$$\begin{aligned}
\mathcal{P}(T_1 \cap T_2) &= \mathcal{P}(T_1) \cdot \mathcal{P}(T_2) \\
&= p \cdot p = p^2
\end{aligned}$$

Nota bene il precedente esercizio rispetta le 3 proprietà:

- $\binom{n}{k} p^i (1-p)^{n-k} = 1;$
- $k = (0...n),\ \binom{n}{i} p^i (1-p)^{n-i} > 0;$
- $\Omega = \{X = 0\} \uplus ... \uplus \{X = n\}$

16.5 Legge discreta uniforme

$S = \{e_1 ... e_n\}$
$\mathcal{P}(S = e_i) = \frac{1}{m}$
$i = 1...m$

16.5.1 Esempi

- casuale.tra(1;5) - funzione excel.
- P: numero aleatorio che rappresenta il punteggio nel lancio di un dado onesto - $\mathcal{P}(S_p = 1) = \frac{1}{6},\ S_p = \{1...6\}(m = 6)$.

17 Lezione 16

17.1 Teorema di Bayes

Teorema 15. *Consideriamo ($H_1 ... H_n$) un s.c.a (sistema completo di alternative)*
i=(1,...n), $\mathcal{P}(H_i) > 0$ ed è nota, allora:

$B \in \mathcal{F}$, B si è verificato,
i=(1,...,n), $\mathcal{P}_B(H_i) =$?

Dimostrazione.

$$\begin{aligned} &= \mathcal{P}(B \cap H_i) = \mathcal{P}_B(H_i) \cdot \mathcal{P}(B) \\ &= \mathcal{P}(B \cap H_i) = \mathcal{P}_{H_i}(B) \cdot \mathcal{P}(H_i) \end{aligned}$$

a noi interessa $\mathcal{P}_B(H_i)$:

$$\begin{aligned} \mathcal{P}_B(H_i) = \mathcal{P}(H_i) &= \frac{\mathcal{P}_{H_i}(B) \cdot \mathcal{P}(H_i)}{\mathcal{P}(B)} \\ &= \frac{\mathcal{P}_{H_i}(B) \cdot \mathcal{P}(H_i)}{\sum_{j=1}^{n} \mathcal{P}_{H_j}(B) \cdot \mathcal{P}(H_j)} \end{aligned}$$

□

17.2 Esempio 1 - Contraffazione

Vediamo il seguente esempio, 3 tipografie falsificano le banconote e differiscono tra di loro in dimensioni e quantità di banconote false prodotte:

- $T_1, \ \mathcal{P}(T_1) = 0.05;$
- $T_2, \ \mathcal{P}(T_2) = 0.20;$
- $T_3, \ \mathcal{P}(T_3) = 0.75;$

Evento B: "la banconota falsa è nella mia tasca", verifichiamo quindi le probabilità condizionate da B:

Calcoliamo per T1 risultato:

$$\begin{aligned}\mathcal{P}_B(T_1) &= \frac{\mathcal{P}_{T_1}(B) \cdot \mathcal{P}(T_1)}{\mathcal{P}_{T_1}(B) \cdot \mathcal{P}(T_1) + \mathcal{P}_{T_3}(B) \cdot \mathcal{P}(T_2) + \mathcal{P}_{T_3}(B) \cdot \mathcal{P}(T_3)} \\ &= \frac{0.90 \cdot 0.05}{0.90 \cdot 0.05 + 0.40 \cdot 0.20 + 0.30 \cdot 0.75} = \frac{0.045}{0.35} = 0.128\end{aligned}$$

va da se che possiamo calcolare anche le probabilità delle altre tipografie:

Calcoliamo per T3 risultato:

$$\mathcal{P}_B(T_3) = \frac{0.30 \cdot 0.75}{0.35} = \frac{0.225}{0.35} = 0.64$$

Calcoliamo per T2 risultato:

$$\mathcal{P}_B(T_2) = 1 - 0.128 - 0.64 = 0.23$$

17.3 Esempio 2 - Tiro con l'arco

Calcolare la probabilità che essendo stato colpito ik bersaglio una sola volta, l'autore del detto centro è il tiratore con probabilità maggiore. Si consideri quindi il tiro con lparco e ogni giocatore un tiro e una propria probabilità di fare centro.

Svolgimento Evento: "uno dei due tiratori ha fatto centro"
analizziamo i dati:

- Giocatore 1 - $\mathcal{P}(G1) = 0.9$
- Giocatore 2 - $\mathcal{P}(G2) = 0.5$

alternative:

- H1: "due colpi non fanno centro"
- H2: "due colpi fanno centro"
- H3: "G1 centra e G2 non centra"
- H4: "G1 non centrra e G2 centra"

calcoliamo le probabilità delle alternative:

- $\mathcal{P}(H1) = 0.1 \cdot 0.5 = 0.05$
- $\mathcal{P}(H2) = 0.9 \cdot 0.5 = 0.45$
- $\mathcal{P}(H3) = 0.9 \cdot 0.5 = 0.45$
- $\mathcal{P}(H4) = 0.1 \cdot 0.5 = 0.05$

si verifica B:

- $\mathcal{P}_B(H1) = 0$
- $\mathcal{P}_B(H2) = 0$
- $\mathcal{P}_B(H3) = 0.90$
- $\mathcal{P}_B(H4) = 0.10$

ciò è dimostrabile calcolando tramite la formula di Bayes le probabilità delle alternative sullpevento B (si faccia il lungo calcolo e l'applicazione della formula vista nel precedente eserizio).

17.4 Medie delle variabili aleatorie

Consideriamo una rilevazione di dati di un carattere quantitativo:

- Rilevazione di dati: $y_1, y_2, ..., y_n$;
- n = taglia della rilevazione;
- y = dato grezzo;

Definiamo la **MEDIA ARITMETICA**:

$$\overline{y} = \frac{y_1 + ... + y_n}{n}$$

Alcuni termini:

- Carattere quantitativo: Y;
- Modalità: $x_1, ..., x_k$ con $k \leq n$;
- $(x_1, n_1), ..., (x_n, n_k)$ - si presenta x_i per n_i volte;
- $n_1 + n_2 + ...n_k = n$

Infatti:

$$\begin{aligned} \overline{y} &= \frac{n_1 x_1 + n_2 x_2 + ... + n_k x_k}{n} = \\ &= \frac{n_1}{n} x_1 + \frac{n_2}{n} x_2 + ... + \frac{n_k}{n} x_k \end{aligned}$$

si noti come tutti i rapporti sono **probabilità relative/frequentiste**, per l'appunto definisce secondo varie rappresentazioni (i.e Grafico a barre) la **MODA** di y:

$$\simeq \mathcal{P}(Y = x_1) + ... + \mathcal{P}(Y = x_k)$$

Y è **un numero aleatorio discreto**, infatti definiamo una legge:

- $S_Y = \{y_1, ..., y_n\}$
- $y \in S_Y, \; \mathcal{P}(Y = y)$

Si definisce **media** di numero aleatorio Y:

$$E(Y) = \sum_{i=1}^{k} y_i \cdot \mathcal{P}(Y = y_i)$$

Sia $r \in \mathbf{N}$, si definisce la media di Y^r:

$$E(Y^r) = \sum_{i=1}^{k} y_i^r \cdot \mathcal{P}(Y = y_i)$$

per r che va da 1 a 2.

17.4.1 Varie rappresentazioni

Le rappresentazioni possono essere fatte sotto forma di grafici oppure anche in forma tabellare:

Modalità	Frequenza assoluta	relativa	Intensità
x_1	n_1	$f_1 = n_1/n$	$i_1 = x_1 f_1$
x_2	n_2	$f_2 = n_2/n$	$i_2 = x_2 f_2$
...	...	...	...
x_k	n_k	$f_k = n_k/n$	$i_k = x_k f_k$
$\sum_i^k$	n (Taglia)	1	$\overline{y}$ (Media aritmentica)

Esempio si consideri una raccolta dati:[1]
$\underline{y} = (2, 5, 3, 2, 1, 3, 5, 2, 2, 3, 1, 4, 1, 1, 3, 2, 2, 3, 4, 4)$

Quindi $n = 20, x_1 = 1,\ x_2 = 2,\ x_3 = 3,\ x_4 = 4,\ x_5 = 5$

se rappresentiamo quesi dati nella tabella che abbiamo visto sopra concludiamo che:

- Taglia $=$ 20 (somma delle occorrenze);
- Frequenza $(n_k/n) = 1$;
- Media aritmetica $= \frac{53}{20} = 2.65$

[1]Vedesi le slide prodotte dal professore in merito a questa lezione e si consideri una ripasso riguardante la probabilita relativa (frequentista) vista nelle prime lezioni.

18 Lezione 17

18.1 Visione del problema: Salomone e le due meritrici

Un giorno il saggio re Salomone[1] dovette giudicare una causa in cui c'erano contrapposte due donne che si contendevano lo stesso bambino.

Ognuna di loro sosteneva che il figlio era suo. Vista l'impossibilità di poter stabilire chi delle due contendenti era la vera madre, al Sovrano venne un'idea. Egli propose di tagliare il bimbo in due parti uguali e di offrirne metà a ciascuna di esse.

La prima fu d'accordo con il Monarca e con tono glaciale disse: "se non lo posso tenere io che sono la vera madre, allora è meglio tagliarlo a metà, così che nemmeno lei lo possa avere per intero". La seconda donna, invece, pregò il Re di non fare del male al bambino. "Datelo a lei, preferisco perderlo anziché vederlo morire", disse cadendo in ginocchio davanti ai suoi piedi.

Con questo stratagemma il re Salomone riuscì a capire immediatamente che quella era la madre legittima del bambino. "Alzati - le disse con dono deciso - e riprenditi tuo figlio!"[2].

18.2 Definizione probabilità frequentista

Ricapitolando quindi lo studio di una raccolta di dati y=(y1...yn) arriviamo ad una conclusione:

- $x_1, ..., x_k$
- $n_1, ..., n_K$

definiamo:

$$\frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{n}(n_1x_1, ..., n_kx_k)$$

il tutto formalizzato:

$$x_1 \cdot \frac{n_1}{n} + ... + x_k \cdot \frac{n_k}{n}$$

quindi definiamo la probabilità frequentista (vista nelle prime lezioni):

$$\mathcal{P}_f(A) = \lim_n \tfrac{n_a}{n}$$
$$\mathcal{P}_f(A) \approx \tfrac{n_a}{n}$$

X è una v.a con spettro S_x e con $x \in \S_x$, $\mathcal{P}(X = x_i) = p_i$ (conosciamo la legge di X).

$$E(X) = x_1 \cdot p_1 + ... + x_k \cdot p_k$$
$$\textstyle\sum_{i=1}^{k} x_i \cdot p_i$$

[1] https://it.wikipedia.org/wiki/Salomone

[2] Visione di alcune slide e focalizzazione sulla formula di Bayes

18.2.1 Esempio 1

Legge:

- $X \sim B(1, p)$
- $S_x = \{0, 1\}$
- $p_1 = \mathcal{P}(X = 0) = 1 - p$
- $p_2 = \mathcal{P}(X = 1) = p$

Media aritmetica di tale legge:

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

(p è il parametro di probabilità della legge)

18.2.2 Esempio 2

Legge:

- $X \sim B(n, p)$
- $S_x = \{0, ..., n\}$
- $j = (0, ..., n),\ \mathcal{P}(X = j) = \binom{n}{j} \cdot p^j \cdot (1 - p)^{n-j}$

Media aritmetica di tale legge:

$$E(X) = n \cdot p$$

Svolgimento

$$\begin{aligned}
E(X) &= \sum_{j=0}^{n} j \binom{n}{j} \cdot p^j \cdot (1 - p)^{n-j} && (1)\\
&= \sum_{j=1}^{n} j \left(\frac{n!}{j!(n - j)!}\right) \cdot p^j \cdot (1 - p)^{n-j} && (2)\\
&= \sum_{j=1}^{n} j \left(\frac{\cancel{j}\, n!}{\cancel{j}\,(j - 1)!(n - j)!}\right) \cdot p^j \cdot (1 - p)^{n-j} && (3)\\
&= n \sum_{j=1}^{n} j \left(\frac{(n - 1)!}{(j - 1)!(n - j)!}\right) \cdot p^j \cdot (1 - p)^{n-j} && (4)\\
&= n \sum_{j=1}^{n} \binom{n - 1}{j - 1} \cdot p^{j-1} \cdot (1 - p)^{n-j} \cdot p && (5)\\
&= n \cdot p \sum_{j=1}^{n} \binom{n - 1}{j - 1} \cdot p^{j-1} \cdot (1 - p)^{n-j} && (6)\\
&= \text{si consideri l=j-1} && (7)\\
&= n \cdot p \sum_{l=0}^{n-1} \binom{n - 1}{l} \cdot p^l \cdot (1 - p)^{(n-l)-1} && (8)
\end{aligned}$$

Calcoli:

- $l = j - 1 \Longrightarrow j = l + 1$
- $n - j = n - (l + 1) = (n - 1) - l$

nel risultato finale:

$$= n \cdot p \sum_{l=0}^{n-1} \binom{n-1}{l} \cdot p^l \cdot (1-p)^{(n-1)-l} \tag{9}$$

$$= n \cdot p[p + (1-p)]^{n-1} \tag{10}$$

$$= n \cdot p \cdot p^{n-1} = n \cdot p \tag{11}$$

Nota Bene la sommatoria si riferisce ad una legge ben precisa:
$B((n-1), p)$

18.2.3 Conclusioni

$S_n = x_1 + x_2 + ... + x_n$ con $x_i \sim B(1, p)$
$E(S_n) = E(X) = E(x_1) + ... + E(x_n)$ oppure $E(x_1 + ... + x_n)$

la media di una somma è ugale alla somma delle medie e nel nostro caso:

$$= p + p + p...$$
$$= n \cdot p$$

18.3 Proprietà

Le proprietà che possiamo evidenziare sono le seguenti:

- $E(X+Y) = E(X) + E(Y)$
- $b \in \mathbf{R}(b \neq 0),\ E(bX) = B \cdot E(X)$
- $a, b \in \mathbf{R},\ E(aX+bY) = aE(X) + bE(Y)$

Esempio 1:

$$\begin{aligned} a \in \mathbf{R},\ E(X+a) &= E(X) + E(a) \\ &= E(X) + a \end{aligned}$$

Esempio 2:

$$\begin{aligned} b \in \mathbf{R}(b \neq 0),\ E(bX) &= \sum_{j=1}^{k} bx_j \mathcal{P}(X = x_j) \\ &= b \sum_{j=1}^{k} x_j \mathcal{P}(X = x_j) \\ &= b \cdot E(X) \end{aligned}$$

Esempio 3:

$$\begin{aligned} a, b \in \mathbf{R},\ E(aX+bY) &= E(aX) + E(bY) \\ &= aE(X) + bE(Y) \end{aligned}$$

18.4 Generalizziamo le proprietà

$$\left.\begin{array}{r} a_1 ... a_r \in \mathbf{R} \\ x_1 ... x_r v.a \end{array}\right\} E(\sum_{n=1}^{r} a_m X_m) = \sum : m = 1^r a_m \cdot E(X_m) \quad (1)$$

Esempio definiamo la seguente legge:

- $T_1 \sim X \sim G(p)$
- $S_x = \{1, ...\} = \mathbf{N}$
- $j \in \mathbf{N},\ \mathcal{P}(X = j)(1 - p)^{n-1}$.

la media:

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

Svolgimento:

$$\begin{aligned} E(T) = E(X) &= \sum_{j=1}^{\infty} j \cdot (1-p)^{j-1} \cdot p \\ &= p \cdot \sum_{j=1}^{\infty} j \cdot (1-p)^{j-1} \\ &= \frac{1}{p} \end{aligned}$$

18.5 Finale

Legge:

- $X \sim \Pi(\lambda)$
- $S_x = \{0, 1, ...\} = \mathbf{N_0}$
- $j \in \mathbf{N_0}$, $\mathcal{P}(X = j) = e^{-\lambda} \cdot \frac{\lambda^j}{j!}$

la media:

$$E = \lambda$$

(anche questa volta λ è risultato della media ed è il parametro della legge)

Svolgimento:

$$\begin{aligned}
E(X) &= \sum_{j=0}^{\infty} j \cdot e^{-\lambda} \cdot \frac{\lambda^j}{j!} \\
&= e^{-\lambda} \sum_{j=0}^{\infty} j \cdot \frac{\lambda^j}{j!} \\
&= e^{-\lambda} \sum_{j=1}^{\infty} j \cdot \frac{\lambda^j}{j!} \\
&= e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^{j-1}}{(j-1)!} \cdot \lambda \\
&= (l = j - 1) \\
&= e^{-\lambda} \cdot \lambda \sum_{l=0}^{\infty} \frac{\lambda^l}{l!} = e^{-1} \cdot \lambda \cdot e^{\lambda} \\
&= \lambda \sum_{l=0}^{\infty} e^{-\lambda} \cdot \frac{\lambda^l}{l!} = \lambda
\end{aligned}$$

la sommatoria, come nei casi precedenti si riferisce ad una legge, in questo caso alla legge $\Pi(\lambda)$.

18.5.1 Svolgimento in serie di e^{λ}

$$\sum_{l=0}^{\infty} \frac{\lambda^l}{e!} = \frac{1}{0!} + \frac{\lambda}{1!} + \frac{\lambda}{2!} + ... = e^{\lambda} \tag{1}$$

In particolare:

$$e^{-1} = 1 - 1 + \frac{1}{2!} - \frac{1}{3!} + ... \tag{2}$$

$$e^2 = 1 + 2 + \frac{2^2}{2!} + \frac{2^3}{3!} + ... \tag{3}$$

In generale per ottenere la somma A della serie di termine generale a_i:

(a) si sommano i primi n addendi per ottenere la somma parziale:

$$A_n = \sum_{i=1}^{n} a_i \tag{4}$$

(b) si determina il limite della successione $(A_n)n \in \mathbf{N}$:

$$\lim_{n\to\infty} A_n \tag{5}$$

(c) se il limite in (b) appartiene a **R** allora:

$$A = \lim_{n\to\infty} A_n \tag{6}$$

(d) se:

$$\lim_n A_n = +\infty \text{ oppure } \lim_n A_n = -\infty \tag{7}$$

allora si dice che la serie è divergente.

(e) se:

$$\lim_{n\to\infty} A_n \tag{8}$$

non esiste allora si dice che la serie è non regolare.

18.6 Lezione 18

18.7 Esempio per applicare il Teorema di Bayes

Un aereo risulta disperso, la ricerca si focalizza su 3 zone: R_1, R_2, R_3, si consideri una probabilità per queste tre zone/regioni pari ad $\frac{1}{3}$.

Consideriamo l'evento E: la ricerca nella zona 1 produce il ritrovamento dell'aereo., quindi calcoliamo le probabilità applicando la formula di Bayes:

Consideriamo:

$$\mathcal{P}_{R_1}(E) = 1 - \alpha_1$$

con $\alpha \in (0,1)$, allora procediamo conoscendo le varie probabilità con la formula:

$$\mathcal{P}_{E^c}(R_1) = \frac{\mathcal{P}_{R_1}(E^c) \cdot \mathcal{P}(R_1)}{\mathcal{P}_{R_1}(E^c) \cdot \mathcal{P}(R_1) + \mathcal{P}_{R_2}(E^c) \cdot \mathcal{P}(R_2) + \mathcal{P}_{R_3}(E^c) \cdot \mathcal{P}(R_3)} \quad (1)$$

$$= \frac{\mathcal{P}(R_1)}{\mathcal{P}(R_1)} \cdot \frac{\mathcal{P}_{R_1}(E^c)}{\mathcal{P}_{R_1}(E^c) + \mathcal{P}_{R_2}(E^c) + \mathcal{P}_{R_3}(E^c)} \quad (2)$$

$$= \frac{\alpha_1}{\alpha_1 + 1 + 1} = \frac{\alpha_1}{\alpha_1 + 2} \quad (3)$$

Verifichiamo anche le restati probabilità:

$$\mathcal{P}_{E^c}(R_2) = \frac{\mathcal{P}_{R_2}(E^c)}{\mathcal{P}_{R_1}(E^c) + \mathcal{P}_{R_2}(E^c) + \mathcal{P}_{R_3}(E^c)} \quad (4)$$

$$= \frac{1}{\alpha_1 + 2} = \mathcal{P}_{E^c}(R_3) \quad (5)$$

Conclusioni verifichiamo quindi i risultati ottenuti:

$$\frac{\alpha_1}{\alpha_1+2} \leq \frac{1}{\alpha_1+2}$$

Infatti, se l'aereo non è caduto nella regione 1 di sicuro non avverrà il ritrovamento in quella data regione ma la probabilità che il ritrovamento dell'aereo avviene nella regione 2 o 3 (quindi non nella regione 1) è pari a 1.

18.8 Statistica descrittiva

Un carattere di una **popolazione** è una caratteristica che si può ritrovare in qualsiasi elemento della popolazione.

Analogie

- Caratteristica (in statistica) = v.a (in probabilità);
- Popolazione (in statistica) = Spazio Campione (in probabilità);

Il modo in cui si manifesta un carattere, si chiamano **modalità**.
I caratteri si possono classificare in questo modo:

- Qualitativi:
  - Nominali
  - Ordinali
- Quantitativi
  - Discreti
  - Continui

18.9 Varie definizioni

18.9.1 Moda

La moda M_o di una rilevazione dati è la modalita alal quale compete la frequenza maggiore:

$$M_o := \text{argmax } \{n_1, n_2, ..., n_k\},\ x \in \{x_1, x_2, ..., x_k\}$$

Osservazione Si può determinare la moda per una rilevazione dati da qualsiasi tipo di carattere.

18.9.2 Mediana

La mediana M_e di una rilevazione dati è il dato che si trova posizione centrale nella sequenza ordinata dei dati. Se i dato sono organizzati in una distribuzine di frequenze la mediana è ka nodalità più piccola tra quelle che hanno frequenza relativa cumulata maggiore di $\frac{1}{2}$.

Osservazione Non è possibile determinare la mediana di una rilevazione dati relativa a un carattere di tipo qualitativo nominale.

Osservazione Se la taglia n è pari allora nella sequenza dati ordinati:

$$\underline{y}_o = y_{(1)} \leq y_{(2)} \leq \dots \leq y_{(n)}$$

ci sono due posti centrali, in tal caso, se:

- il carattere è quantitativo la mediana è la semisomma dei due dati di posto centrale nella sequenza $\underline{y}_o$;
- il carattere è qualitativo ordinale la mediana non esiste se i dati di posto centrale nella sequenza $\underline{y}_o$ sono diversi;
- il carattere è qualitativo ordinale la mediana e i dati di posto centrale nella sequenza $\underline{y}_o$ sono uguali allora la mediana coincide con ciascuno di essi.

Ai fini dimostrativi la lezione si è spostata su un file excel e teso[3].

[3]Visione di un file excel di esempio - Salto in lungo, che comprende esempi e spiegazioni su vari punti introduttivi di statistica descrittiva

19 Lezione 19

19.1 Proprietà di continuità

Definizione Definiamo una successione di eventi $(C_n)_{n \in \mathbf{N}} \subseteq \mathcal{F}$ si dice crescente se risulta che:

$$n \geq 1,\ C_n \subseteq C_{n+1}$$

Osservazione

$$C_1 \subseteq C_2 \subseteq ... \subseteq C_n \subseteq ...$$

Definizione *Si definisce limite di una successione di eventi crescente con la posizione:*

$$C \equiv \lim_{n \to \infty} C_n := \bigcup_{n=1}^{\infty} C_n$$

Ovviamente: $n \in \mathbf{N},\ C \supseteq C_n$

Definizione Definiamo una successione di eventi $(D_n)_{n \in \mathbf{N}} \supseteq \mathcal{F}$ si dice crescente se risulta che:

$$n \geq 1,\ D_n \supseteq D_{n+1}$$

Osservazione

$$D_1 \supseteq D_2 \supseteq ... \supseteq D_n \supseteq ...$$

Definizione *Si definisce limite di una successione di eventi decrescenti con la posizione:*

$$D \equiv \lim_{n \to \infty} D_n := \bigcap_{n=1}^{\infty} D_n$$

Ovviamente: $n \in \mathbf{N},\ D \subseteq D_n$

19.2 Proprietà di continuità di P

Teorema 16. *Assegnata una successione crescente* $(C_n)_{n \in \mathbf{B}})$ *du eventi si ha:*

$$\lim_n \mathcal{P}(C_i) = \mathcal{P}(\lim_n C_n)$$

Dimostrazione. F = successione ausiliaria così definita:

- $F_1 = C_1$;
- $F_2 = C_1^c C_2$;
- $F_2 = C_1^c C_2^c C_3$;
- $F_n = C_1^c C_2^c ... C_{n-1}^c C_n$

Ovviamente si considerino queste due proprietà aggiuntive:

a) $n \in \mathbf{N},\ \dot\cup_{i=1}^n F_i = C_n = \cup_{i=1}^n C_i$

b) $\lim_n \dot\cup_{i=1}^n F_i = \lim_n \cup_{i=1}^n C_i \Longleftrightarrow \dot\cup_{i=1}^\infty F_i = \cup_{i=1}^\infty C_i$

Allora,

$$\mathcal{P}(\bigcup_{I=1}^{\infty} C_i) = \mathcal{P}(\bigcup_{I=1}^{\infty} F_i) \overset{num.add.}{=} \sum_{i=1}^{\infty} \mathcal{P}(F_i) \tag{1}$$

$$= \lim_{n\to\infty} \sum_{i=1}^{n} \mathcal{P}(F_i) \tag{2}$$

$$\overset{fin.add.}{=} \lim_{n\to\infty} \mathcal{P}(\dot\bigcup_{i=1}^{n} F_i) \tag{3}$$

$$\overset{a)}{=} \lim_n \mathcal{P}(\bigcup_{=} \mathcal{C}) = \mathcal{C} \tag{4}$$

In altri termini:

$$\lim_n \mathcal{P}(C_n) = \mathcal{P}(\bigcup_{n=1}^{\infty} C_n) = \mathcal{P}(\lim_n C_n)$$

□

Teorema 17. *Assegnata una successione decrescente $(D_n)_{n\in\mathbf{N}}$ di eventi si ha:*

$$\lim_n \mathcal{D} = \mathcal{P}(\lim_n D_n) = \mathcal{P}(D)$$

Dimostrazione.

$$D_n \downarrow D \Longrightarrow D_n^c \uparrow D^c \Longrightarrow$$
$$\mathcal{P}(\bigcup_N^\infty D_n^c) = \mathcal{P}(\lim_n D_n^c) \overset{Teorema1}{=} \lim_n \mathcal{P}(D_n^c)$$

Per De Morgan vale la seguente cosa:

$$\bigcup_{n=1}^\infty D_n^c = (\bigcap_{n=1}^\infty D_n)^c$$

quindi continuando la dimostrazione:

$$\Longleftrightarrow \mathcal{P}[(\bigcap_{n=1}^\infty D_n)^c] = \lim_n \mathcal{P}(D_n^c) \tag{1}$$
$$\Longleftrightarrow 1 - \mathcal{P}(\bigcap_{n=1}^\infty D_n) = \lim_n [1 - \mathcal{P}(D_n)] = 1 - \lim_n \mathcal{P}(D_n) \tag{2}$$
$$\Longleftrightarrow \mathcal{P}(\bigcap_{n=1}^\infty D_n) = \lim_n \mathcal{P}(D_n) \tag{3}$$

In altri termini:

$$\lim_n \mathbf{P}(D_n) = \mathcal{P}(\bigcap_{n=1}^\infty D_n) = \mathcal{P}(\lim_n D_n)$$

Caso particolare

$$D_n \downarrow \varnothing \Longrightarrow \lim_n \mathcal{P}(D_n) = 0$$
$$\Longleftrightarrow D_n \downarrow \varnothing \Longrightarrow \mathcal{P}(D_n) \downarrow 0$$

□

19.2.1 Esempio

$(T_n)_{n\in\mathbf{N}}$, T_n :"esce T al lancio n"

- $A_1 = T_1\omega\omega\omega$ - $\mathcal{P}(A_1 = \frac{1}{2} = p$
- $A_2 = T_1T_2\omega\omega$ - $\mathcal{P}(A_2 = \frac{1}{4} = p^2$
- $A_3 = T_1T_2T_3\omega$ - $\mathcal{P}(A_3 = \frac{1}{8} = p^3$
- ...
- $A_n = T_1...T_n\omega$ - - $\mathcal{P}(A_1 = \frac{1}{2^n} = p^n$

$A_n \downarrow (T_1, T_2, ...T_n, ...)$

Per il teorema:

$$\lim_{n\to\infty} \mathcal{P}(A_n) = \lim_{n\to\infty} \frac{1}{2^n} = 0 = \mathcal{P}(\lim_{n\to\infty} A_n$$
$$= \mathcal{P}(T_1, T_2, ..., T_n, ..)$$

19.3 Teorema di continuità

Se $(A_n)_{n\in\mathbf{N}} \subseteq \mathcal{F}$ crescente oppure decrescente allora:

$$\lim_n \mathcal{P}(A_n) = \mathcal{P}(\lim_n A_n)$$

Il teorema di continuità è molto importate in quanto, aknebi oer le successioni crescenti e decrescenti, consente lo scambio di segno di limite e il simbolo della misura di probabilià. Il teorema di continuità continua anche a valere nel caso delle succesioni non monotone (crescenti o decrescenti) ma dotate di limite:

$$A_n \longrightarrow A \Longrightarrow \lim_n \mathcal{P}(A_n) = \mathcal{P}(\lim_n A_n) = \mathcal{P}$$

Ciò, facilita in molti casi il calcolo delle probabilità di un evento e , in più, consente di ottenere le proprietà della funzione di distribuzione.

Spettro	Probabilità	Probabilità accumulata
1	1/6	1/6
2	1/6	2/6
3	1/6	3/6
4	1/6	4/6
5	1/6	5/6
6	1/6	6/6

20 Lezione 20

20.1 Esemi introduttivi

20.1.1 Esempio 1

$X \sim U_d(1,2,3,4,5,6)$, lancio di un dado onesto:

(si faccia un grafico con ascisse e ordinate, e si studi la funzione F_X, si noteranno le proprietà che vedremo alla fine).

Proviamo a calcolare:

$$\begin{aligned}
\mathcal{P}(2 \leq X \leq 4) &= \mathcal{P}(X \in \{2,3,4\}) && (1)\\
&= \mathcal{P}(X = 2) + \mathcal{P}(X = 3) + \mathcal{P}(X = 4) && (2)\\
&= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} && (3)\\
&= \frac{3}{6} && (4)
\end{aligned}$$

adesso proviamo:

$$F_4 - F_2 = \frac{4}{6} - \frac{2}{6} = \frac{2}{6}\text{sbagliato}$$
$$F_4 - F_1 = \frac{4}{6} - \frac{1}{6} = \frac{3}{6}\text{corretto}$$

Se consideriamo il grafico precedente possiamo rilevare alcune probabilità ponendo:

- $x_1 = -1,\ \mathcal{P}(X \leq x_1) = 0$
- $x_2 = 0,\ \mathcal{P}(X \leq x_2) = 0$
- $\mathcal{P}(X \leq 1) = \frac{1}{6}$
- $x_3 = 1.5,\ \mathcal{P}(X \leq x_3) = \frac{1}{6}$
- ...
- $\mathcal{P}(X \leq 7) = 1$

è possibile attraverso le proprietà aggevolare il calcolo per calcolare (per differenza) tutte le proprietà:

$F_X(4) - F_x(2^-) = \frac{3}{6}$
$F(2^-) = \lim_{\epsilon \to 0^+} F_x(2-\epsilon) = \frac{4}{6} - \frac{1}{6} = \frac{1}{6}$

Ulteriori calcoli:

$$\begin{aligned}\frac{1}{6} = \mathcal{P}(X=4) &= F_X(4^+) - F_X(4^-)\\ &= F_X(4) - F_X(4^-)\\ &= \frac{4}{6} - \frac{3}{6} = \frac{1}{6}\end{aligned}$$

- $F_X(2.5) = 1 \cdot \frac{1}{6} + 1 \cdot \frac{1}{6} + 0 \cdot \frac{1}{6} + \ldots$
- $F_X(3.4) = 1 \cdot \frac{1}{6} + 1 \cdot \frac{1}{6} + 1 \cdot \frac{1}{6} + 0 \cdot \frac{1}{6} + \ldots$
- $F_X(1) = 1 \cdot \frac{1}{6} + 0 \cdot \frac{1}{6} + \ldots$

20.1.2 Esempio 2

$X \sim B(1, \frac{1}{3})$, legge di bernulli di parametro $p = \frac{1}{3}$
$S_x = \{0, 1\}$
$\mathcal{P}(X=0) = q = 1 - p = \frac{2}{3}$
$\mathcal{P}(X=1) = p = \frac{1}{2}$

20.1.3 Esempio 3

$X \sim B(3, \frac{1}{3})$, legge di bernulli di parametro $p = \frac{1}{3}$
l'esempio è riportato sulle slide del prof.

20.2 Funzione di distribuzione

"Non c'è nulla che colpisca più di questo fatto: via via che la Matematica si elevava e appartava nelle regioni più alte del pensiero astratto, tornava poi a terra con uno strumento sempre più importante per l'analisi dei fatti concreti"

20.2.1 Definizione

Sia $(\Omega, \mathcal{P})$ uno spazio di probabilità e sia:

$$X : \Omega \xrightarrow{\text{misurabile}} \mathbf{R}$$

un numero aleatorio.
Si definisce la **funzione di distribuzione** di ponendo:

$$F_X : \mathbf{R} \longrightarrow [0,1]$$
$$x \longrightarrow F_X(x) = \mathcal{P}(X \leq x)$$

$\{X \leq x\} = \{\omega \in \Omega : x(\omega) \leq x\} \subseteq \Omega$
$\{X \leq x\} \in \mathcal{F}$, in quanto:
$\{X \leq x\} = X^{-1}(] - \infty, x]$ - algebra di Borel (è misurabile)

20.2.2 Proprietà

a) F_X è non decrescente.

b) $\lim_{\epsilon \to 0^+} F_X(x - \epsilon) = F_X(x)$

c) $\lim_{x \to -\infty} F_X(x) = 0 \;\; \lim_{x \to +\infty} F_X(x) = 1$

20.2.3 Dimostrazione delle proprietà

Proprietà 1: la non decrescenza.

Dimostrazione. Si considerino due punti $(x_1,\ x_2)$ con $x_1 < x_2$ bisogna dimostrare che:

$F_X(x_2) \geq F_X(x_1)$ si procede nel seguente modo:

$$\begin{aligned}
F_X(x_2) = \mathcal{P}(x \leq x_2) &= \mathcal{P}[\{X \leq x_1\} \cup \{x_1 < X \leq x_2\}] \\
&= \mathcal{P}(x \leq x_1) + \mathcal{P}(x_1 < x \leq x_2) \\
&\geq \mathcal{P}(X \leq x_1) = F_X(x_1)
\end{aligned}$$

□

Proprietà 2: $\lim_{\epsilon \to 0^+} F_X(x - \epsilon) = F_X(x)$

Dimostrazione.

$$\begin{aligned}
\lim_n F_X(x + \frac{1}{n}) = \\
&= \lim_n \mathcal{P}(x \leq x + \frac{1}{n}) \\
&= \lim_n \mathcal{P}(\{X \leq x\} \cup \{x < X \leq x + \frac{1}{n}\} \\
&= \lim_n [\mathcal{P}(x \leq X) + \mathcal{P}(x < X \leq x + \frac{1}{n})] \\
&= \lim_n \mathcal{P}(x \leq X) + \lim_n \mathcal{P}(x < X \leq x + \frac{1}{n}) \\
&= F_X + \lim \mathcal{P}(x < X \leq x + \frac{1}{n}) \\
&= F_X + \mathcal{P}(\lim_n E_n) \\
&= F_X + \mathcal{P}(\varnothing) = F_X
\end{aligned}$$

Questo fatto insieme alla proprietà a garantisce la dimostrazione. □

Proprietà 3: $\lim_{x\to-\infty} F_X(x) = 0$ $\lim_{x\to+\infty} F_X(x) = 1$

Dimostrazione. Iniziamo sostituendo al posto di x, n. Ciò significa che $\lim_n \mathcal{P}(x \le n)$

$$\begin{aligned}(C_n)_{n\in\mathbf{N}},\ C_n &= \{\omega \in \Omega : X(\omega) \le n\} \\ &\equiv \{X \le n\}\end{aligned}$$

è una successione crescebte e $C_n \uparrow \Omega$ che per le proprietà di continuità:

$$\begin{aligned}&= \mathcal{P}(\lim_n \{X \le n\}) \\ &= \mathcal{P}(\Omega) = 1\end{aligned}$$

questa è la dimostrazione per la successione crescente, $\lim_n F_X(x) = 1$, adesso consideriamo:

$$\begin{aligned}(D_n)_{n\in\mathbf{N}},\ D_n &= \{\omega \in \Omega : X(\omega) \le -n\} \\ &\equiv \{X \le -n\}\end{aligned}$$

è una successione crescebte e $D_n \downarrow \varnothing$ che per le proprietà di continuità:

$$\begin{aligned}&= \mathcal{P}(\lim_n \{X \le -n\}) \\ &= \mathcal{P}(\varnothing) = 0\end{aligned}$$

questa è la dimostrazione per la successione decrescente, $\lim_n F_X(x) = 0$, questo iniseme anche alla proprietà a) garantiscono la dimostrazione. □

Vedesi footnote.[1]

[1] Vedere da libri/appunti di Analisi 1 il "Teorema Ponte".

21 Lezione 21

21.1 Esempio

La lezione 21 del 29/04/2021 è stata introdotta con un esempio in particolare per poi passare ad una serie di esempi tratti dal libro.

- $X \sim G(p),\ p \in (0,1);\ q = (1-p) \in (0,1);$
- $S_x = \mathbf{N}$

Consideriamo:

$$
\begin{aligned}
&k \in S_x,\ \mathcal{P}(X = k) = p \cdot q^{k-1} \\
&k \in S_x, F_X(k) = \mathcal{X} \leq\ = \sum_{i=0}^{k} \mathcal{X} = \\
&= p \sum_{i=1}^{k} q^{i-1} = p(1 + q + \dots + q^{k-1}) \\
&= p \cdot \frac{1-q^k}{1-q} = p \cdot \frac{1-q^k}{p} = 1 - q^k
\end{aligned}
$$

Progressione geometrica di ragione q.

$$
\begin{aligned}
k \in S_x, \mathcal{P}(X > k) &= 1 - \mathcal{P}(X \leq k) = 1 - F_X(k) \\
&= 1 - (1 - q^k) = q^k
\end{aligned}
$$

Il complemento a q della funzione di distribuzione di una qualsiasi variabile aleatoria X si disegna con simbolo $\overline{F}_X$ e si chiama:

Funzione di sopravvivenza
$X \sim G(p)$
$k \in S_x, \overline{F}_X(k) = 1 - F_X(k) = q^k = (1-p)^k$

21.2 Proposizione

$X \sim B(n,p),\ p \in (0,1), q = (1-p) \in (0,1)$
$n \in \mathbf{N}$
$S_X = \{0, ..., n\}$
$k \in S_X,\ \mathcal{P}(X = k) = \binom{n}{k} \cdot p^k \cdot q^{n-k}$ In più distinguiamo dei casi:

- $k = 0$
- $k = (1, ..., n)$

Nel primo caso abbiamo:

$$\mathcal{P}(X = 0) = \binom{n}{0} \cdot p^0 \cdot q^{n-0} = q^n$$

invece nel secondo caso (casi generali da 1 ad n):

$$\begin{aligned}\frac{\mathcal{P}(X = k)}{\mathcal{P}(X = k-1)} &= \frac{\binom{n}{k} \cdot p^k \cdot q^{n-k}}{\binom{n}{k-1} \cdot p^{k-1} \cdot q^{n-k+1}} \\ &= \frac{n!}{k! \cdot (n-k)!} \cdot \frac{(k-1)! \cdot (n-k+1)!}{n!} \cdot \frac{p}{q} = \frac{(n-k+1)}{k} \cdot \frac{p}{q}\end{aligned}$$

Quindi:

$$\mathcal{P}(X = k) = \mathcal{P}(X = k-1) \cdot \frac{p}{q} \cdot \frac{n-k+1}{k}$$

Esempi

- $\mathcal{P}(X = 0) = q^n$
- $\mathcal{P}(X = 1) = q^n \cdot \frac{p}{q} \cdot n = n \cdot q^{n-1} \cdot p$
- $\mathcal{P}(X = 2) = n \cdot q^{n-1} \cdot p \cdot \frac{n-1}{2}$

21.3 Riferimenti

Vari esempi sono stati trattati a lezione presi dal libro[1] di riferimento.

22 Lezione 22

22.1 Momenti di un numero aleatorio

$y = (y_1, ..., y_n)$ dati quantitativi.
$\overline{y} = \frac{1}{n} \cdot \sum_{i=0}^{n} y_i$
$s^2 = \frac{1}{n-1} \cdot \sum_{i=1}^{n} (y_1 - \overline{y})^2 \geq 0$

$$\overline{y} \pm \sqrt{s^2}$$

Esempi

- n=2, [-1,1] $\overline{y} = 0$ $s^2 = \frac{1}{2-1} \cdot [(-1-0)^2 + (1-0)^2] = 2$;
- n=2, [-10,10] $\overline{y} = 0$ $s^2 = \frac{1}{2-1} \cdot [(-10-0)^2 + (10-0)^2] = 200$

22.2 Definizione

Se esiste $E(X^m)$ (ossia se è un numero reale) allora esso si dice **momento** di ordine m. Nel caso m=1, $\mu_1' = E(X)$ è la media della variabile aleatoria. Nel caso m=2, $\mu_2' = E(X^2)$ è la media del secondo ordine (**Media quadratica**). L'operazione per ottenere μ_m' è la seguente (per le variabili discrete):

$$\mu_m' = \sum_{x_i \in S_X} x_i^m \mathcal{P}(X = x_i)$$

22.3 Esempio - Poisson

$X \sim \Pi(\lambda),\ \lambda > 0$
Conosciamo la legge:

- $S_X = \{0, 1, ...\}$
- $\mathcal{P}(X = k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}$

Svolgimento

$$\begin{aligned}
\mu_1' = E(X) &= \sum_{k=0}^{\infty} k \cdot e^{-\lambda} \cdot \frac{\lambda^k}{k!} = \sum_{k-1}^{\infty} k \cdot e^{-\lambda} \cdot \frac{\lambda^k}{k!} \\
&= \sum_{k=1}^{\infty} e^{-\lambda} \cdot \frac{\lambda^k}{(k-1)!} \\
&\text{m=j-1 / k=m+1} \\
&= \sum_{m=0}^{\infty} e^{-\lambda} \cdot \frac{\lambda^{m+1}}{m!} \\
&= \lambda \cdot 1 = \lambda
\end{aligned}$$

La prima conclusione afferma che la media di una legge di Poission è uguale al parametro dato in input (λ), calcoliamo adesso la media quadratica:

$$\mu_2' = E(X^2) = \sum_{k=0}^{\infty} k^2 \cdot e^{-\lambda} \cdot \frac{\lambda^k}{k!} = \sum_{k-1}^{\infty} k^2 \cdot e^{-\lambda} \cdot \frac{\lambda^k}{k!} \tag{1}$$

$$= \sum_{k=1}^{\infty} e^{-\lambda} \cdot k \cdot \frac{\lambda^k}{(k-1)!} \tag{2}$$

$$= \text{m=j-1 / k=m+1} \tag{3}$$

$$= \sum_{m=0}^{\infty} (m-1) \cdot e^{-\lambda} \cdot \frac{\lambda^{m+1}}{m!} \tag{4}$$

$$= \lambda \cdot \sum_{m=0}^{\infty} m \cdot e^{-\lambda} \cdot \frac{\lambda^{m+1}}{m!} + \lambda \cdot \sum_{m=0}^{\infty} e^{-\lambda} \cdot \frac{\lambda^{m+1}}{m!} \tag{5}$$

$$= \lambda^2 + \lambda \tag{6}$$

Per la definizione che seguirà concluderemo che in Poission λ è sia media che varianza e in particolare:

$$D^2(X) = E(X^2) - E^2(X) = \lambda^2 + \lambda - \lambda^2 = \lambda$$

22.3.1 Definizione per Poission

Se X è una variabile aleatoria per la quake $E(X^2) < +\infty$, si definisce la **varianza di X** la quanttà:

$$D^2(X) := E[(x - \mu_x)^2]$$

(la differenza è chiamata **scarto**) dove $\mu_x \equiv \mu_1' = E(X)$. In più, la radice quadrata della varianza:

$$D(X) := \sqrt{D^2(X)}$$

si chiama **"deviazione standard"**.

22.3.2 Proposizione

Se $E(X^2) < +\infty$, allora:

$$D^2(X) = E(X^2) - E^2(X)$$

ovvero

$$\sigma_x^2 = \mu_2' -_x^2 \mu'$$

22.3.3 Dimostrazione

Dimostrazione.

$$
\begin{aligned}
D^2(X) &= E[(X - \mu_x)^2] && (7)\\
&= E(X^2 - 2\mu_x \cdot X + \mu_x^2) && (8)\\
&= E(X^2) + E(-2\mu_x \cdot X) + E(\mu_x^2) && (9)\\
&= E(X^2) + (-2\mu_x) \cdot E(X) + E(\mu_x^2) && (10)\\
&= \mu_2' - 2\mu_x^2 + \mu_x^2 && (11)\\
&= \mu_2' - \mu_x' + \mu_x^2 && (12)
\end{aligned}
$$

$\square$

22.4 Esempio - Binomiale

$X \sim B(n,p),\ n \in \mathbf{N},\ p \in (0,1)\ q = (1-p) \in (0,1)$
Conosciamo la legge:

- $S_X = \{0, ..., n\}$
- $\mathcal{P}(X = k) = \binom{n}{k} \cdot p^k \cdot q^{n-k}$

Svolgimento

$$
\begin{aligned}
E(X) &= \sum_{k=0}^{n} k \cdot \binom{n}{k} \cdot p^k \cdot q^{n-k} && (1)\\
&= \sum_{k=1}^{n} k \cdot \binom{n}{k} \cdot p^k \cdot q^{n-k} && (2)\\
&= \sum_{k=1}^{n} k \cdot \frac{n!}{k!(n-k)!} \cdot p^k \cdot q^{n-k} = \sum_{k=1}^{n} \frac{n(n-1)!}{(k-1)!(n-k)} \cdot p^k \cdot q^{n-k} && (3)\\
&= n \sum_{k=1}^{n} \frac{n(n-1)!}{(k-1)!(n-k)} \cdot p \cdot p^{k-1} \cdot q^{n-k} && (4)\\
&= n \cdot p \sum_{k=1}^{n} k \cdot \binom{n-1}{k-1} \cdot p^{k-1} \cdot q^{n-k} && (5)\\
&= \text{m=j-1 / k=m+1} && (6)\\
&= n \cdot p \sum_{m=0}^{n-1} \binom{n-1}{m} \cdot p^m \cdot q^{n-(m-1)} && (7)\\
&= n \cdot p \cdot 1 = n \cdot p && (8)
\end{aligned}
$$

Nota bene il risultato della somamtoria è 1 dato che è la somma delle probabilità della legge Bionmiale con n-1 e p come parametri (pari a 1).

Vogliamo calcolare la media del second'ordine (sostituendo a $k = k^2$) e k in più che "sopravvive", il risultato è:

$$
\begin{aligned}
&= n \cdot p \cdot (n-p) \cdot p + n \cdot p = && (9)\\
&= n^2 \cdot p^2 - n \cdot p^2 + n \cdot p = && (10)\\
&= n^2 \cdot p^2 + n \cdot p \cdot (1-p) = n^2 \cdot p^2 + n \cdot p \cdot q && (11)
\end{aligned}
$$

Che in definitiva, $E(X) = n \cdot p,\ E(X^2) = n^2 \cdot p^2 + n \cdot p \cdot q$, quindi:

$$
\begin{aligned}
D^2(X) &= E(X^2) - E^2(X)\\
&= n^2 \cdot p^2 + n \cdot p \cdot q - n^2 \cdot p^2\\
&= n \cdot p \cdot q = n \cdot p \cdot (1-p)
\end{aligned}
$$

22.5 Esempio - Geometrica

$X \sim G(p),\ p \in (0,1)\ q = (1-p) \in (0,1)$
Conosciamo la legge:

- $S_X = \{0, 1, ...\}$
- $\mathcal{P}(X = k) = p \cdot q^{k-1}$

Svogimento

$$
\begin{aligned}
E(X) &= \sum_{k-1}^{\infty} k \cdot p \cdot q^{k-1} = p \sum_{k=1}^{\infty} k \cdot q^{k-1} && (1)\\
&= p \cdot (1 + 2q + 3q^2 + 4q^3 + ...) && (2)\\
&= p \cdot [\sum_{k=0}^{\infty} q^k + q \cdot \sum_{k=0}^{\infty} q^k + q^2 \sum_{k=0}^{\infty} q^k + ...] && (3)\\
&= p \cdot \sum_{k=0}^{\infty} q^k (1 + q + q^2 + ...) && (4)\\
&= p \cdot \sum_{k=0}^{\infty} q^k \cdot \sum_{k=0}^{\infty} q^k = p \cdot (\frac{1}{(1-q)^2} && (5)\\
&= p \cdot \frac{1}{p^2} = \frac{1}{p} && (6)
\end{aligned}
$$

22.6 Tabella riepilogativa

La tabella sottostante (ancora incompleta) verrà riempita successivamente:

	Media	Varianza	Parametri
Geometrica	1/p	?	$p \in (0,1)$
Binomiale	$n \cdot p$	$n \cdot p \cdot (1-p)$	$n \in \mathbf{N},\ p \in (0,1)$
Poission	λ	λ	$\lambda > 0$

22.7 Vari esercizi e riferimenti

Verso la fine della lezione vengono svolti l'eservizio del Capitolo 4 del nostro libro di riferimento[1] per la precisione esercizio 7.a e 7.b, inoltre il prof consiglia di prendere visione di tali argomenti alla pagina 162 del testo con annesso esercizio 7.c.

Viene fatto anche riferimento al **"Paradigma di Poission"**qui[1].

[1] `https://it.wikipedia.org/wiki/Distribuzione_di_Poisson`

23 Lezione 23

23.1 Utilizzo della funzione di distribuzione

In generale abbiamo considerato la semiretta sinstra come cinervallo e la probabilità che X (variabile aleatoria) assumesse un valore in tale intervallo.

$x \in \mathbf{R},\ F_X(x) := \mathcal{P}(X \leq x)$

nell'intervallo:

$]-\infty, x]$

23.2 Altri intervalli

23.2.1 Intervallo: $]x_1, x_2]$

$$\begin{aligned} F(X_2) = \mathcal{P}(X \leq x_2) &= \mathcal{P}(\{X \leq x_1\} \uplus \{x_1 < X \leq x_2\}) \\ &= \mathcal{P}(X \leq x_1) + \mathcal{P}(x_1 < X \leq x_2) \\ &= \mathcal{P}(x_1 < X \leq x_2) = F_X(x_2) - F_X(x_1) \end{aligned}$$

23.2.2 Intervallo: $\{x\}$

$$\begin{aligned} \epsilon > 0,\ \mathcal{P}(x - \epsilon < X \leq x) &\overset{(1)}{=} F_X(x) - F_X(x-\epsilon) \\ &\mathcal{P}(\lim_n x - \epsilon < X \leq x) = F_X(x) - F(x^-) \\ &\mathcal{P}(X = x) = F(x) - F_X(x^-) \\ &= F_X(x) - \lim_{\epsilon \to 0^+} F_X(x-\epsilon) \end{aligned}$$

Caso particolare nel caso che F_x è continua in x, allora:
$\mathcal{P}(X = x) = F_X(x) - F_X(x) = 0$

23.2.3 Intervallo: $[x_1, x_2]$

$$[x_{1,2}] =]x_1, x_2] \uplus \{x_1\}$$

$$\begin{aligned} \mathcal{P}(x_1 \leq X \leq x_2) &= \mathcal{P}(x_1 < X \leq x_2) + \mathcal{P}(X = x_1) \\ &\overset{(1)}{=} F_X(x_2) - F_X(x_1) + \mathcal{P}(X = x_1) \\ &\overset{(2)}{=} F_X(x_2) - F_X(x_1) + F_X(x_1) - F_(x_1^-) \\ &= F_X(x_2) - F_X(x_1^-) \end{aligned}$$

23.2.4 Intervallo: $]x_1, x_2[$

$$]x_{1,2}] =]x_1, x_2[\cup \{x_2\}$$

$$\begin{aligned}
\mathcal{P}(x_1 < X \le x_2) &= \mathcal{P}(x_1 < X < x_2) + \mathcal{P}(X = x_2) \\
\mathcal{P}(x_1 < X < x_2) &= \mathcal{P}(x_1 < X \le x_2) - \mathcal{P}(X = x_2) \\
&\overset{(1)}{=} F_X(x_2) - F_X(x_1) - \mathcal{P}(X = x_2) \\
&\overset{(2)}{=} F_X(x_2) - F_X(x_1) - [F_X(x_2) - F_X(x_2^-)] \\
&= F_X(x_2^-) - F(x_1)
\end{aligned}$$

23.2.5 Intervallo: $[x_1, x_2[$

$$[x_{1,2}] = [x_1, x_2[\cup \{x_2\}$$

$$\begin{aligned}
\mathcal{P}(x_1 \le X \le x_2) &= \mathcal{P}(x_1 \le X < x_2) + \mathcal{P}(X = x_2) \\
\mathcal{P}(x_1 \le X < x_2) &= \mathcal{P}(x_1 \le X \le x_2) - \mathcal{P}(X = x_2) \\
&\overset{(1)(3)}{=} F_X(x_2) - F_X(x_1^-) - [F(x_2) - F_x(x_2^-)] \\
&= F_X(x_2) - F_X(x_1^-) - [F_X(x_2) + F_X(x_2^-)] \\
&= F_X(x_2^-) - F(x_1^-)
\end{aligned}$$

23.2.6 Intervallo: $]x, +\infty[$

$$]x, +\infty[= (] - \infty, x])^c$$

$$\begin{aligned}
\mathcal{P}(x \le X \le +\infty) &= 1 - \mathcal{P}(-\infty < X \le x) \\
\mathcal{P}(X > x) &= 1 - \mathcal{P}(X \le x) \\
&= 1 - F_X(x)
\end{aligned}$$

23.2.7 Intervallo: $] - \infty, x[$

$$] - \infty, x] =] - \infty, x[\cup \{X = x\}$$

$$\begin{aligned}
\mathcal{P}(x \le x) &= \mathcal{P}(X < x) + \mathcal{P}(X = x) \\
\mathcal{P}(X < x) &= \mathcal{P}(X \le x) - \mathcal{P}(X = x) \\
&= F_X(x) - [F_X(x) - [F(x^-)] \\
&= F_X(x_2) - F_X(x_1^-) - [F_X(x_2) + F_X(x_2^-)] \\
&= F_X(x) - F_X(x) + F_X(x^-) = F_X(x^-)
\end{aligned}$$

23.2.8 Intervallo: $[x, +\infty[$

$$[x, +\infty[= (]-\infty, x[)^c$$

$$\begin{aligned}\mathcal{P}(X \geq x) &= 1 - \mathcal{P}(X < x) \\ &\overset{(7)}{=} 1 - F_X(x^-)\end{aligned}$$

23.3 Funzione di densità

In statistica descrittiva i dati che vengono rilevati medianti uno strumento di misura si riferiscea un carattere "quantitativo continuo".

In calcolo delle probabilità l'analogo del carattere "quantitativo continuo" è rappresentato da una variabile aleatoria con legge (assolutamente) continua.

Per le variabile discreate abbiamo:

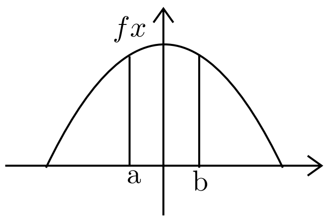
- $S_X = \{x_1, ..., x_n, ...\}$
- $x_i \in S_X, \; \mathcal{P}(X = x_i)$
- $F_X(x) = \sum_i 1_{x_i \leq x} \cdot \mathcal{P}(X = x_i)$

Per le variabili con legge di probabilità di tipo continuo è necessario introdurreintrodurre un'altra funzione: **funzione di densità di probabilità** che si design con il dimbolo $\mathbf{f_x}$ e che integratada $-\infty$ a un reale x qualsiasi restituisce il valore della funzione di distribuzione:

$$\int_{-\infty}^{x} f_x(t)\, dt = F_X(x)$$

è la funzione integrale f.d.p: $f_X(x) = F'_X(x)$

23.3.1 Esempio



Svolgimento

$$\begin{aligned}
\mathcal{P}(a \leq X \leq b) &= F_X(b) - F_X(a^-) \\
&= \int_{-\infty}^{b} f_x(t)\, dt - \int_{-\infty}^{a} f_x(t)\, dt \\
&= \int_{a}^{b} f_x(t)\, dt
\end{aligned}$$

X è un numero aleatorio con legge di probabilità continua: pertato essa è dotata du f.d.p, sia essa f_X, allora, $\mathcal{P}(x < X < x + dt)$:

$$\begin{aligned} \mathcal{P}(X = x) &= \mathbf{P}(x \le X \le x) \\ &= \int_x^x f_x(t)\, dt = 0 \end{aligned}$$

quindi la f.d è continua in $\mathbf{R}$.

23.3.2 Proprietà

- $x \in \mathbf{R},\ f_x(x) \ge 0$
- $\lim_{x \to \infty} F_X(x) = 1$

$$\begin{aligned} 1 &= \lim_{x \to \infty} \int_{-\infty}^{x} f_x(t)\, dt = \int_{-\infty}^{+\infty} f_x(t)\, dt \\ &= \mathcal{P}(-\infty < X < +\infty) \end{aligned}$$

23.3.3 Vari esempi e riferimenti

Visto l'esercizio 5.1.a del libri di testo.[1]

24 Lezione 24

La lezione si è incentrata su alcuni eservizi guidati presi dal nostro libro di testo consigliato[1] sul capitolo 5.

24.1 Esempi

Esercizi - parte 1

- Esempio 5.1.b - Calcolo funzione di distribuzione (e verifica f.d.p).
- Esempio 5.1.c - Calcolo funzione di distribuzione (e verifica f.d.p).
- Esempio 5.1.d - Determinare f.d.p di una v.a (data F.D e f.d.p di un altra variabile aleatoria).

Esercizi - parte 2

- Esempio - 5.2.a - Calcolo della media.
- Esempio - 5.2.b - Calcolo della media.

Esercizi - parte 3 Questi sono gli eserizi consigliati da svolgere e prendere in considerazione sul libro di testo:

- Esempio - 5.2.c (pg. 219)
- Esempio - 5.2.d (pg. 220)

24.2 Teoria

24.2.1 Media e Varianza

Visti i vari esempi precedenti:

Forumal 1:

$$m \in \mathbf{N},\ \mu'_n \equiv E(X^m) = \int_{-\infty}^{+\infty} x^m \cdot f_X(x)\, dx$$

Formula 2:

$$\begin{aligned} D^2(X) &= E(X^2) - E^2(X) \\ &= \mu'_2 - (\mu'_1)^2 \end{aligned}$$

25 Lezione 25

La lezione inizia con lo svolgimento dell'esercizio assegnato nella lezione 24 (5.2.c) e svolto il continuo dell'esercizio 5.2.a (ovvero il 5.2.e).

25.1 Legge Uniforme

25.1.1 Preposizione 1

$\mu_{aX+b} \equiv E(aX + b) = a \cdot E(X) + b$ per linearità

25.1.2 Proposizione 2

$D^2(aX + b) = D^2(aX) = a^2 \cdot D^2(X)$

Dimostrazione.

$$\begin{aligned} D^2(aX + b) &= E\{[(aX + b) - \mu_{aX+b}]^2\} \\ &= E\{[(aX + b) - (a\mu_X + b]^2\} \\ &= E[(aX + b - a\mu_X - b)^2] \\ &= E[(aX - a\mu_x)^2] \\ &= E[a^2(X - \mu_x)^2] \\ &= a^2 \cdot E[(X - \mu_X)^2] = a^2 \cdot D^2(X) \end{aligned}$$

□

25.1.3 Definizione - legge uniforme

Una v.a continua la cui legge ha la seguente f.d.p:

$$f_X(x) = \begin{cases} 1, & 0 < x < 1 \\ 0, & altrimenti \end{cases}$$

si dice "**UNIFORME**" nell'intervallo (0,1) e si scrive:

$$X \sim U(0, 1)$$

Risulta:

- $\mu_X = E(X) = \int_{-\infty}^{+\infty} X \cdot F_x(C)\, dx = \int_0^1 x\, dx = \frac{x^2}{2}\Big|_0^1 = \frac{1}{2}$
- $\mu_2 = E(X^2) = \int_0^1 x^2 \cdot F_x(C)\, dx = \frac{x^3}{3}\Big|_0^1 = \frac{1}{3}$

Allora: $D^2(X) = \mu_2' - \mu_x^2 = \frac{1}{3} - \frac{1}{2} = \frac{1}{12}$

25.1.4 Preposizione 3

$X \sim U(0,1)$
$Y = a + (b-a)X$

La v.a Y si dice che è distribuita uniformemente nell'intervallo (a,b):

$$f_Y(y) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & altrimenti \end{cases} \tag{1}$$

Dimostrazione.

$$x \in \mathbf{R},\ F_X(x) = \int_{-\infty}^{x} 1\,dx \begin{cases} 0, & x \leq 0 \\ x, & 0 < x < 1 \\ 1, & x \geq 1 \end{cases}$$

$x \in \mathbf{R}$,

$$\begin{aligned} F_Y(y) &= \mathcal{P}(Y \leq y) = \mathcal{P}[a + (b-a)X \leq y] \\ &= \mathcal{P}[(b-a)X \leq (y-a)] \\ &= \mathcal{P}(x \leq \frac{y-a}{b-a}) \\ &= F_X(\frac{y-a}{b-a}) \end{aligned}$$

Per derivazione:

$$\begin{aligned} f_Y(y) &= f_X(\frac{y-a}{b-a}) \cdot D_y(\frac{y-a}{b-a} \\ &= f_X(\frac{y-a}{b-a}) \cdot \frac{1}{b-a}) \end{aligned}$$

Se:

- $y < a \Longleftrightarrow \frac{y-a}{b-a} < 0 \Longrightarrow f_Y(y) = 0$
- $y > b \Longleftrightarrow \frac{y-a}{b-a} > 1 \Longrightarrow f_Y(y) = 0$
- $a < y < b \Longleftrightarrow 0 < \Longrightarrow f_Y(y) = \frac{1}{b-a}$

□

25.1.5 Proposizione 4

Se $X \sim U(a,b)$ allora:

- $E(X) = \frac{a+b}{2}$
- $D^2(X) = \frac{(b-a)^2}{12}$

Dimostrazione.

$$\begin{aligned} E(Y) &= E[a + (b-a)X] = a + (b-a) \cdot E(X) \\ &= a + \frac{b-a}{2} = \frac{2a+b-a}{2} = \frac{a+b}{2} \end{aligned}$$

$$\begin{aligned} D^2(Y) &= D^2[(b-a)X] = D^2[(b-a)X] \\ &= (b-a)^2 D^2(X) = \frac{(b-a)^2}{12} \end{aligned}$$

□

25.2 Vari esempi e riferimenti

Alla fine della lezione si sono visti due esempi di applicazione della legge uniforme presenti nel libro di testo[1], per la precisone è stato visto l'esercizio 5.3.c e 5.3.b.

26 Lezione 26

26.1 Leggi Normali (o Gaussiane)

Una v.a X si dice che ha la legge Gaussiana di parametri $\mu \in \mathbf{R}$ e $\sigma^2 > 0$ se è caratterizzata dalla seguente f.d.p:

$$f_X(x; \mu, \sigma^2) = f_X(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

dove $\sigma = \sqrt{\sigma^2} > 0$. In simboli, si scrive:

$$X \sim N(\mu, \sigma^2)$$

Osservazione si dimostra che non si può scrivere la seguente funzione primitiva come composizione di funzioni elementari:

$$x \in \mathbf{R},\ F_X(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{(y-\mu)^2}{2\sigma^2}}\, dy$$

Allora possiamo porre la funzione Φ in cui $\mu =$ t e $\sigma = 1$ per ottenere una primitiva più semplice da calcolare:

$$z \in \mathbf{R},\ \Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{t^2}{2}}\, dt = F_Z(z)$$

Il grafico della f.d.p è simmetrico all'asse verticale di $x = \mu$; presenta due punti di flesso $\mu \pm \sigma$; assume il massimo assoluto per $x = \mu$; decresce a 0 nell'intorno di $-\infty$ e $+\infty$.

Nota bene il grafico il particolare descrive una funzione Gaussiana visibile qui.[1]

[1] https://it.wikipedia.org/wiki/Funzione_gaussiana

26.1.1 Proposizione 1

$$\int_{-\infty}^{+\infty} f_X(x;\mu,\sigma^2)\,dx = 1$$

Dimostrazione.

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

Trasformiamo:

- $y = \frac{x-\mu}{\sigma}$
  - $x \to -\infty \Longrightarrow y \to -\infty$
  - $x \to +\infty \Longrightarrow y \to +\infty$
- $x = \mu + \sigma \cdot y$

Riprendiamo il calcolo[2]:

$$\begin{aligned}
&= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\cdot\sigma}\cdot e^{-\frac{y^2}{2}}\cdot\sigma\,dy \\
&= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{y^2}{2}}\,dy \equiv I
\end{aligned}$$

Da qui in poi continuiamo con I:

$$\begin{aligned}
I^2 &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{y^2}{2}}\,dy\cdot\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}}\cdot e^{\frac{z^2}{2}}\,dz \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \frac{1}{2\pi}\cdot e^{\frac{1}{2}(y^2+z^2)}\,dy\ dz \\
&= \int_{0}^{+\infty}\left[\int_{0}^{2\pi} \frac{1}{2\pi}\cdot e^{-\frac{1}{2}r^2}\cdot r\,dr\right]d\theta \\
&= \int_{0}^{+\infty} r\cdot e^{-\frac{1}{2}r^2}\cdot\left[\int_{0}^{2\pi}\frac{1}{2\pi}\,d\theta\right]dr \\
&= \int_{0}^{+\infty} r\cdot e^{-\frac{1}{2}r}\,dr
\end{aligned}$$

[2]Durante la lezione il prof si è spostato su argomenti della geometria analitica, si consiglia di fare riferimento alla lezione o ai propri appunti del corso.

Consideriamo adesso la seguente trasformazione:

- $t = -\frac{1}{2}r^2$
  - $r \to 0 \Longrightarrow t \to 0$
  - $r \to +\infty \Longrightarrow t \to -\infty$
- $\frac{dt}{dr} = -\frac{1}{2} \cdot r$
  - dt=$-r \cdot dr$
  - dr=$-\frac{dt}{r}$

Contiuniamo la dimostrazione:

$$\begin{aligned} &= \int_0^{+\infty} \cancel{r} \cdot e^t \frac{dt}{\cancel{r}} \\ &= \int_{-\infty}^{0} e^t \, dt = e^t|_{-\infty}^{0} = 1 - 0 = 1 \end{aligned}$$

Allora

$$I^2 = 1 \Longrightarrow I = 1 \Longleftrightarrow I = 1$$

□

26.1.2 Proposizione 2

Per ogni coppia di numeri reali (α, β) tali che $\alpha \neq 0$ si ha:

$$X \sim N(\mu, \sigma^2) \longrightarrow Y = \alpha x + \beta \sim N(\alpha\mu + \beta, \alpha^2\sigma^2)$$

Dimostrazione. $\alpha > 0$

$$\begin{aligned} y \in \mathbf{R},\ F_Y(y) = \mathcal{P}(Y \leq y) &= \mathcal{P}(\alpha X + \beta \leq y) \\ &= \mathcal{P}(\alpha X + \beta \leq Y - \beta) \\ &= \mathcal{P}(X \leq \frac{y-\beta}{\alpha}) \\ &= F_X(\frac{y-\beta}{\alpha}) \end{aligned}$$

Adesso possiamo derivare il risultato ottenuto:
$y \in \mathbf{R},\ f_Y(y) = f_X(\frac{y-\beta}{\alpha}) \cdot \frac{1}{\alpha}$

Ricordando che (D_y) in questo caso va calcolata così:

$$
\begin{aligned}
D_y\left(\frac{y-\beta}{\alpha}\right) &= \frac{1}{\alpha}\cdot D_y(y-\beta) = \frac{1}{\alpha} \\
&= \frac{1}{\sqrt{2\pi}\alpha\sigma}\cdot e^{-\frac{\left(\frac{y-\beta}{\alpha}-\mu\right)^2}{2\sigma^2}} \\
&= \frac{1}{\sqrt{2\pi}\alpha\sigma}\cdot e^{-\frac{(y-\beta-\alpha\mu)^2}{2\alpha^2\sigma^2}} \\
&= \frac{1}{\sqrt{2\pi}\alpha\sigma}\cdot e^{-\frac{[y-(\beta+\alpha\mu)]^2}{2(\alpha\sigma)^2}}
\end{aligned}
$$

il che comporta che:

$$Y \sim N(\beta+\alpha\mu, \alpha\sigma)$$

La dimostrazione per $\alpha < 0$ è analogo e si lascia al lettore. $\square$

26.1.3 Teorema

Teorema 18.

$$X \sim N(\mu, \sigma^2) \Longrightarrow E(X) = \mu \ e \ D^2(X) = \sigma^2$$

Data una legge normale di X, conosciamo sia la sua varianza sia la sua media.

Dimostrazione. $Z \sim N(0,1)$, Z è la v.a con legge normale standard.

Calcoliamo $E(Z)$:

$$\begin{aligned}
E(Z) &= \int_{-\infty}^{+\infty} z \cdot f_Z(z)\, dz = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{z^2}{2}} \cdot z\, dz \\
&= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{z^2}{2}}\, d\frac{z^2}{2} \\
&= \frac{1}{2\sqrt{\pi}} \cdot \int_{-\infty}^{+\infty} e^{-\frac{z^2}{2}}\, d\frac{z^2}{2} \\
&= -\frac{1}{2\pi} \cdot e^{-\frac{z^2}{2}} \Big|_{-\infty}^{+\infty} = 0
\end{aligned}$$

Calcoliamo $E(Z^2)$:

$$\begin{aligned}
E(Z^2) &= \int_{-\infty}^{+\infty} z^2 \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{z^2}{2}}\, dz \\
&= \frac{2}{\sqrt{2\pi}} \int_0^{+\infty} z^2 \cdot e^{-\frac{z^2}{2}}\, dz \\
&= \int_0^{+\infty} z \cdot z \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{z^2}{2}}\, dz \\
&= \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} z \cdot z \cdot e^{-\frac{z^2}{2}}\, dz
\end{aligned}$$

Riscriviamo il risultato in questo modo:

$$E(Z \cdot e^{-\frac{z^2}{2}}) = -e^{-\frac{z^2}{2}} + c$$

si può quindi applicare la definizione dell'integrazione per parti (in partiolare si noti $D(Z) = 1$):

$$\begin{aligned}
&= \frac{2}{\sqrt{2\pi}} \left[-z \cdot e^{-\frac{z^2}{2}} \Big|_0^{+\infty} + \int_0^{+\infty} e^{-\frac{z^2}{2}}\, dt \right] \\
&= \frac{2}{\sqrt{2\pi}} \int_0^{+\infty} e^{\frac{z^2}{2}} = 2 \int_0^{+\infty} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{z^2}{2}}\, dz = 1
\end{aligned}$$

Quindi: $E(Z) = 0$ e $D^2(Z) = E(Z^2) - E^2(Z) = 1$

Consideriamo:
$Y = \mu + \sigma,\ Z \sim N(\mu, \sigma^2)$ da cui si evince che $Y \sim X$

Allora:

- $E(X) = E(Y) = E(\mu + \sigma z) = \mu + \sigma \cdot E(Z) = \mu$
- $D^2(X) = D^2(Y) = D^2(\mu + \sigma z) = D^2(\sigma z) = \sigma^2 \cdot D^2(Z) = \sigma^2$

□

26.2 Osservazioni

La trasformazione che ci consente di passare da una v.a $X \sim N(\mu, \sigma^2)$ ad una v.a $Z \sim B(0, 1)$ è la seguente:

$$\frac{X - \mu}{\sigma} \sim Z \sim N(0, 1)$$

Essa è nota come la "standardizzazione" di Y.

Osservazione 1 Ora è chiaro come utulizzare la funzione Φ per ottenere i valori assunti dalla funzione di distribuzione F_X nel caso di $X \sim N(\mu, \sigma^2)$:

$$\begin{aligned} x \in \mathbf{R},\ F_Z(x; \mu, \sigma^2) &= F_Z(\frac{x - \mu}{\sigma}; 0, 1) \\ &= \Phi(\frac{x - \mu}{\sigma}) \end{aligned}$$

In altri termini l'area sotto la f.d.p di X prima di x è uguale all'area sottoo la f.d.p di Z prima di $\frac{x-\mu}{\sigma}$.

Osservazione 2 Sono state redatte varie tabelle di valori assunti da Φ. Quella più utilizzata compare all'interno del libro di testo[1]. Essa riporta valori di Φ per x=0 fino X=3.49 con incremento pari a $\Delta x = 0.01$.

Definizione Se $X \sim N(\mu, \sigma^2)$ il valore $\frac{x-\mu}{\sigma}$ si dice "il punto standardizzato del punto x, con $x \in \mathbf{R}$".

26.3 Vari esempi e riferimenti

La lezione si è conclusa con metà dello svolgimento dell'esercizio del libro[1] e inoltre stata presentata una tabella di standardizzazione.

27 Lezione 27

La lezione continua sviluppando l'esercizio 5.4.b della lezione 26 e introduce una Proposizione prima di un teorema fondamentale.

27.1 Proposizione

$$x \in \mathbf{R},\ \Phi(-x) + \Phi(x) = 1$$

Dimostrazione. Sia $Z \sim N(0,1)$

$$\begin{aligned} x \geq 0,\ \Phi(-x) &= \mathcal{P}(Z \leq -x) = \mathcal{P}(Z \geq x) \\ &= \mathcal{P}(Z > x) = 1 - \mathcal{P}(Z \leq x) \\ &= 1 - \Phi(x) \end{aligned}$$

In conclusione: $\Phi(-x) + \Phi(x) = 1$

□

Questo risultato è utile nel calcolo di un punto standardizzato negativo.

27.2 Teorema di De Moivre - Laplace

Teorema 19. *Introduciamo questa definizione, sia* $B \in \mathcal{F}$ $(\mathcal{P}(B)) > 0$*:*

- $A \in \mathcal{F},\ \mathcal{P}_B(A) = \frac{\mathcal{P}(A \cap B)}{\mathcal{P}(B)}$
- $A \in \mathcal{F},\ \mathcal{P}(A|B) = \frac{\mathcal{P}(A \cap B)}{\mathcal{P}(B)}$

Enunciato: $n \in \mathbf{N},\ S_n \sim B(n,p),\ p \in (0,1)$.

- $E(S_n) = n \cdot p$
- $D^2(S_n) = n \cdot p \cdot (1-p)$

Standardizzazione:

$$\frac{S_n - n \cdot p}{\sqrt{n \cdot p \cdot (1-p)}}$$

Formula: *sia* $a<b$

$$\mathcal{P}\left[a < \frac{S_n - n \cdot p}{\sqrt{n \cdot p \cdot (1-p)}} < b\right] \to_{n\to\infty} \Phi(b) - \Phi(a)$$

In generale conta n:
$\mathcal{P}(a < Z < b) = \Phi(b) - \Phi(a)$
e per questo che la formula di De Moivre $\approx \mathcal{P}(a < Z < b)$.

27.2.1 Regola Empirica

L'approssimazione è accettabile per: $n \cdot p \cdot (1-p) \geq 10$.

27.2.2 Utilizzo e applicazione del teorema

$X \sim B(40, \frac{1}{2}),\ S_X(\{0, ..., 40\}$

Calcolo classico della probabilita per X=20:

$$\begin{aligned}\mathcal{P}(X=20) &= \binom{40}{20} \cdot (\frac{1}{2})^2 \cdot (1-\frac{1}{2})^{20} \\ &= \frac{40!}{20!20!} \cdot (\frac{1}{2})^{40} \approx 0,1254 \text{ (Calcolo preciso)}\end{aligned}$$

Calcolo con il teorema della probabilita per X=20:

$$\begin{aligned}\mathcal{P}(X=20) &= \mathcal{P}(19.5 < X < 20.5) \\ &= \mathcal{P}(-0.16 < \frac{x-20}{\sqrt{10}} < 0.16) \\ &\approx \Phi(0.16) - \Phi(-0.16) \approx 0.1272\end{aligned}$$

Tornando alla regola empirica, n=40, p=1/2 e 1-p=1/2 allora:

$$n \cdot p \cdot (1-p) = \frac{40}{2 \cdot 2} = 10\ (\geq 10)$$

è quindi possibile applicarla.

27.3 Vari comandi EXCEL

In Excel ci sono varie funzioni che calcolano il risultato della funzione normalizzata (standard e non):

- distrib.nomr.standard(...)
- distrib.norm(...)

La funzione "distrib.norm" calcola il valore della funzione di distribuzione normale se il quarto argomento è "vero" oppure calcola il valore della funzione di densità di probabilità normale se il suo quarto argomento è "falso". Inoltre, per entambi i casi il suo primo argomento è la variabile indipendente, il suo secondo argomento è la media della variabile aleatoria e il suo terzo argomento è la derivazione standard della variabile aleatoria.

28 Lezione 28

La lezione inizia con 3 esempi che vanno a consolidare quello che abbiamo appreso dalla lezione 27 riguardo il **Teorema De Moivre-Laplace**, con i seguenti esercizi:

- Esercizio 5.4.d
- Esercizio 5.4.i
- Esercizio 5.4.j (contrastante al teorema)[1] dove si dimostra che $X \sim$ *ipergeometrica* (N » n).

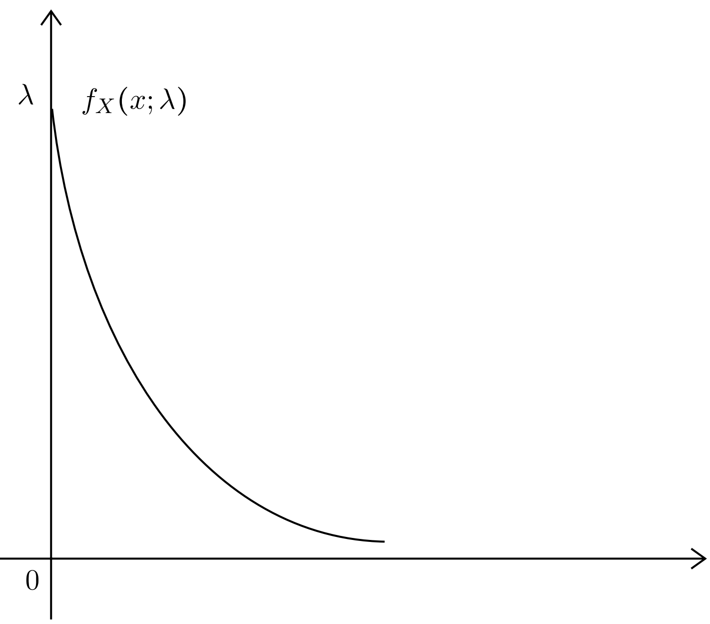
28.1 Variabile aleatoria : Esponenziale - Esp(λ)

Sia $\lambda > 0$, una v.a X si dice avere legge esponenziale di parametro λ e si scrive $X \sim Esp(\lambda)$ se la sua funzione di distribuzione (F.D.) si ottiene per integrazione dalla seguente f.d.p:

$$f_X(x) \equiv f_X(x;\lambda) = \begin{cases} \lambda \cdot e^{-\lambda \cdot x}, \ x > 0 \\ 0, \ x \leq 0 \end{cases}$$

Il grafico della f.d.p è il seguente:

$$\lim_{x \to 0^+} f_X(x;\lambda) = \lambda$$



[1]Vedesi estrazione sequenziale con/senza rimpiazzamento (estrazione a blocchi)

28.2 Verifica delle proprietà

Bisogna verificare le proprietà di f.d.p.

Proprietà 1 la non negatività:

$$x \leq 0 \Longrightarrow f_X(x;\lambda) = 0$$
$$x > 0 \Longrightarrow f_X(x;\lambda) \text{ è il prodotto di due fattori positivi.}$$

Proprietà 2 verifica dell'area uguale a 1:

$$\begin{aligned}
\int_{-\infty}^{+\infty} f_X(x;\lambda)\,dx &= \int_{0}^{+\infty} \lambda \cdot e^{-\lambda x}\,dx \\
&\textit{Pongo } t = \lambda \cdot x \textit{ quindi abbiamo } (x = \frac{t}{\lambda} \Longrightarrow dx = \frac{dt}{\lambda}) \\
&= \int_{0}^{+\infty} \lambda \cdot e^{-t}\,\frac{dt}{\lambda} \\
&= \int_{0}^{+\infty} e^{-t}\,dt = -e^{-t}\Big|_{0}^{+\infty} = e^{-t}\Big|_{+\infty}^{0} = 1 - 0 = 1
\end{aligned}$$

28.3 Calcolo della funzione di distribuzione

Adesso possiamo calcolare la F.D.

Caso 1

$$x < 0,\ F_X(x;\lambda) = \int_{-\infty}^{x} f_X(y;\lambda)\,dy = \int_{-\infty}^{0} 0\,dy = 0$$

Caso 2

$$\begin{aligned} x > 0,\ F_X(x;\lambda) &= \int_{-\infty}^{x} f_X(y;\lambda)\,dy = \int_{-\infty}^{0} 0\,dy + \int_{0}^{x} \lambda \cdot e^{-\lambda y}\,dy \\ &= e^{-\lambda x}\big|_0^x = e^{-\lambda x}\big|_x^0 = 1 - e^{-\lambda x} \end{aligned}$$

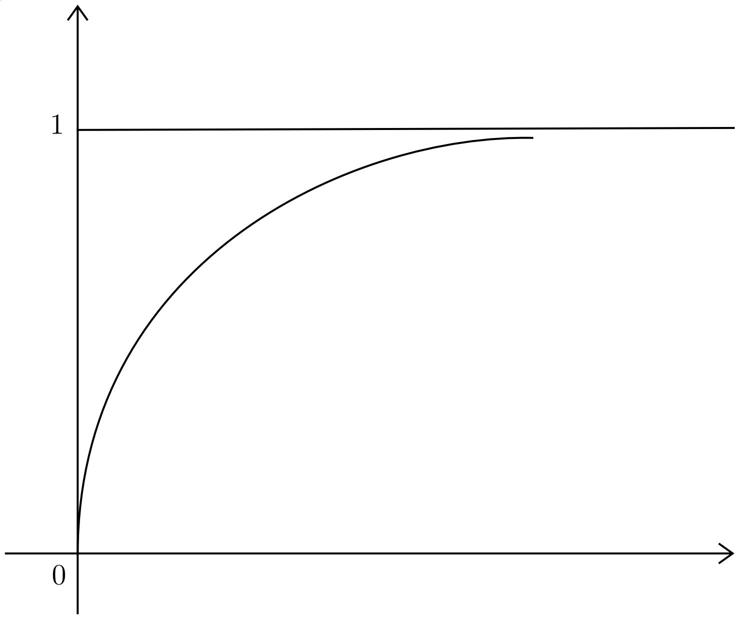
In definitiva:

$$F_X(x;\lambda) = \begin{cases} 1 - e^{-\lambda \cdot x},\ x > 0 \\ 0,\ x \leq 0 \end{cases}$$

Studiamo il comportamento per x che tende ad infinito

$$\lim_{x\to\infty} F_X(x;\lambda) = \lim_{x\to\infty} 1 - e^{1-e^{\lambda x}} = 1 - \lim_{x\to\infty} e^{-\lambda x}$$

Il grafico:



28.4 Proposizione

Sia $X \sim Esp(\lambda)$ con $\lambda > 0$ e sia $m \in \mathbf{N}$

Risulta:

$$\mu'_m = E(X^m) \int_0^{+\infty} x^m \cdot \lambda \cdot e^{-\lambda x}\, dx = \frac{m!}{\lambda^m}$$

Caso base , sviluppo per induzione:
m=1, $\mu'_1 \equiv \mu_x = E(X) = \int_0^{+\infty} x \cdot \lambda \cdot e^{-\lambda x}\, dx$ Integro per parti

$$\begin{aligned} D_X(X) = 1; \int_0^{+\infty} \lambda \cdot e^{-\lambda X}\, dx &= -e^{-\lambda x} + c \\ &= -x \cdot e^{-\lambda x}|_0^{+\infty} + \int_0^{+\infty} 1 \cdot e^{-\lambda x}\, dx \\ &= 0 + \frac{1}{\lambda} \int_0^{+\infty} \lambda e^{-\lambda x}\, dx = \frac{1}{\lambda} = \frac{1!}{\lambda^2} \end{aligned}$$

Caso generale , sviluppo per induzione:
Sia vera la tesi per m > 1, bisogna vedere che essa è vera per m+1.

$$\begin{aligned} E(X^{m+1}) &= \int_0^{+\infty} x^{m+1} \cdot \lambda \cdot e^{-\lambda x}\, dx \\ &= -x^{m+1} \cdot e^{-\lambda x}|_0^{+\infty} + (m+1) \cdot \int_0^{+\infty} x^m e^{-\lambda x}\, dx \\ &= (m+1) \cdot \frac{1}{\lambda} \cdot \int_0^{+\infty} x^m \cdot \lambda \cdot e^{-\lambda x}\, dx \text{ (Per la tesi, l'integrale è pari a } m!/\lambda^m) \\ &= \frac{m!(m+1)}{\lambda \cdot \lambda^m} = \frac{(m+1)!}{\lambda^{m+1}} \end{aligned}$$

28.5 Corollari

Corollario 1 Sia $X \sim Esp(\lambda)$, si ha:

$$D^2(X) = \frac{1}{\lambda^2}$$

Dimostrazione.

$$\begin{aligned} D^2(X) &= E(X^2) - E^2(X) \\ &= \frac{2!}{\lambda^2} - (\frac{1}{\lambda})^2 \\ &= \frac{2!}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2} \end{aligned}$$

□

Corollario 2

$$\textit{Coefficiente di variazione } CV(x) = \frac{\sqrt{D^2(X)}}{|E(X)|}$$

Dimostrazione.

$$\frac{\sqrt{D^2(X)}}{|E(X)|} = \frac{\sqrt{\frac{1}{\lambda^2}}}{\frac{1}{\lambda}} = \frac{\frac{1}{\lambda}}{\frac{1}{\lambda}} = 1$$

□

29 Lezione 29

La lezione inizia con lo svolgimento di due esercizi che ci porteranno poi alla teoria vera e propria (ripercorrendo la lezione precedente riguardante le v.a Esp(λ)):

- Esercizio 5.5.b
- Esercizio 5.5.c
- Esercizio 5.5.d (lasciato allo studente)

29.1 L'assenza di usura (v.a. continue)

Introduciamo questa proprietà considerando due intervalli di tempo s,t>0
X è una v.a. con legge esponenziale di parametro λ.

$$\begin{aligned}
\mathcal{P}(X > s + t | X > t) &= \frac{\mathcal{P}(X > s + t, X > t)}{\mathcal{P}(X > t)} \\
&= \frac{\mathcal{P}(X > t + s)}{\overline{F}_X(t)} \\
&= \frac{\overline{F}_X(t + s)}{\overline{F}_X(t)} \\
&= \frac{e^{-\lambda(t+s)}}{e^{-\lambda t}} = \frac{e^{-\lambda t} \cdot e^{-\lambda t}}{e^{-\lambda s}} = e^{-\lambda s} \\
&= \overline{F}_X(s) = \mathcal{P}(X > s)
\end{aligned}$$

In altri termini:

$$\overline{F}_X(s + t) = \overline{F}(s) \cdot \overline{F}(t)$$

Detto ciò l'esercizio 5.5.c può essere facilmente svolto andando a concludere che pur non avendo il parametro λ il risultato delle probabilità è pari a $\frac{1}{2}$.

29.2 L'assenza di memoria (v.a. discrete)

Sia X una v.a con legge geometrica di parametro $p \in (0,1)$
$m \in \mathbf{N},\ \mathcal{P}(X \leq m) = \mathcal{P}(X = 1) + \dots + \mathcal{P}(X = m)$, q=1-p;

$$\begin{aligned} &= p + pq + pq^2 + \dots + pq^{m.1} \\ &= p(1 + q + q^{+} \dots + q^{m-1}) \\ &= p \cdot \frac{1 - q^m}{1 - q} = 1 - q^m \end{aligned}$$

$m \in \mathbf{N},\ F_X(n) = 1 - q^m,\ \overline{F}_X = q^m$

Si condideri adesso:

$$\begin{aligned} m, n \in \mathbf{N}, \mathcal{P}(X > n + m | X > n) &= \frac{\mathcal{P}(X > n + m)}{\mathcal{P}(X > n)} = \frac{\overline{F}(n + m)}{\overline{F}_X(n)} \\ &= \frac{q^n \cdot q^m}{q^n} = q^m = \overline{F}_X = q^m \end{aligned}$$

29.2.1 Esempio (relazione tra legge esponenziale e geometrica)

Formula per ottenere una legge geometrica a partire da una legge esponenziale
$X \sim Esp(\lambda),\ \lambda > 0$

$$Y = \lfloor X \rfloor + 1$$

Esempio

- $y = 3.4 \longleftarrow y = 4$
- $y = impossibile \longleftarrow y = 0$
- $y = 0.67 \longleftarrow y = 1$
- $y = 0.342 \longleftarrow y = 1$
- $y = m, 561 \longleftarrow y = m + 1$
- $y = (m - 1), 123 \longleftarrow y = m$

Dimostrazione. $S_Y = \{1, 2, ..., n, ...\}$

$$\begin{aligned}
\mathcal{P}(Y = n) &= \mathcal{P}(\lfloor x \rfloor + 1 = n) \\
&= \mathcal{P}(\lfloor x \rfloor = n - 1) \\
&= \mathcal{P}(n - 1 \le X < n) \\
&= 1 - e^{-\lambda n} - [1 - e^{-\lambda(n-1)}] \\
&= e^{-\lambda(n-1)} - e^{-\lambda n} \\
&= e^{-\lambda(n-1)} \cdot (1 - e^{-\lambda}) \\
&= (1 - e^{-\lambda}) \cdot (e^{-\lambda})^{n-1} = p \cdot q^{n-1}
\end{aligned}$$

□

Otteniamo:

$$Y \sim G(1 - e^{-\lambda})$$

Se viene fissato p, quale valore di λ bisogna utilizzare?

$$\begin{aligned}
1 - e^{-\lambda} = p &\Longleftrightarrow 1 - p = e^{-\lambda} \\
&\Longleftrightarrow -\lambda = ln(1 - p) \\
&= \Longleftrightarrow \lambda = ln(\frac{1}{1 - p})
\end{aligned}$$

29.3 Centri (dati quantitativi)

Il concetto di centro riesce a mettere in relazione i concetti di moda, mediana e media.
Voglio trovarmi una funzione che mi rappresenta la distanza di x da tutta la rilevazione dati

Sia $\underline{y}=(y,...,y_n)$ una raccolta dati e sia $x \in \mathbf{R}$ ed $r \in \mathbf{N_0}$ definiamo:

$$d(x,\underline{y}) = \begin{cases} \sqrt[r]{\frac{1}{n} \cdot \sum_{i=1}^{n} |x-y_i|^r}, & r > 0 \\ \frac{1}{n} \cdot \sum_{i=1}^{n} |x-y_i|^0, & r = 0 \end{cases}$$

i=(1,2,...,n), $r \geq 1$, $|x-y_i|^2$ rappresenta la discrepanza di x dal dato i-esimo; l'argomento della radice r-esima è quindi la media aritmetica di tutta la discrepanza e la funzione $d_r(x;\underline{y})$ ha distanza (di ordine r) di x da $\underline{y}$. Lo stesso si può dire per r=0.

29.3.1 Definizione

Il centro di ordine r è il numero reale che rende minima la funzione $d_r(c,\underline{y})$

$$\xi_r := argmin_{x \in \mathbf{R}} d_r(x,\underline{y})$$

Teorema 20. *Il centro di ordine 0 è la moda di $\underline{y}$.*

Dimostrazione. Elechiamo le varie proprietà:

- $x \notin \underline{y},\ d_0(x,\underline{y}) = \frac{1}{n} \cdot \sum_{i=1}^{n} |x-y_i|^0$
- $x \in \underline{y},\ d_0(x;\underline{y}) = \frac{1}{n}(n-n_x)$ *n_x è il numero di volte che x è presente nella rilevazione*
- $y_i = x = y_j,\ |x-y_i|^0 = |0|^0 = 0$

□

Per avere il minimo valore di $d_0(x,\underline{y})$ il valore x deve essere uguale al dato che si è presentato il maggior numero di volte, ovvero la moda.

Teorema 21. *Il centro di ordine 1 è la mediana di $\underline{y}$.*

Dimostrazione. NON DIMOSTRATO □

30 Lezione 30

Le lezioni precedenti ci hanno dato i seguenti risultati:

- $\xi_0(\underline{y})$ è la moda di $\underline{\mathrm{y}}$;
- $\xi_1(\underline{y})$ è la mediana di $\underline{\mathrm{y}}$;

30.1 Altri teoremi sul CENTRO

Teorema 22. *$\xi_2(\underline{y})$ è la media di $\underline{y}$. ovvero:*

$$\xi_2(\underline{y}) = \overline{y} = \frac{1}{N} \cdot \sum_{i=1}^{N} y_i$$

Dimostrazione. Consideriamo:

$$f(x) := d_2(\underline{x}; \underline{y})$$

allora:

$$\xi_2(\underline{y}) = argmin_{x \in \mathbf{R}} f(x)$$

Ma,

$$f(x) = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^{N} (x - y_i)^2}$$

che è infinitamente grande nell'intorno $+\infty$ e $-\infty$ e inoltre è derivabile in tutto $\mathbf{R}$ (ciò significa che il limite per $+\infty$ e $-\infty$ è infinito).

Allora la ricerca del minimo assoluto è da eseguire tra i punti di minimo relativo.

D'altra parte, posto:

$$x \in \mathbf{R},\ f(X) := \frac{1}{N} \cdot \sum_{i=0}^{N} (x - y_i)^2$$

Calcoliamo la derivata prima:

$$f'(x) = D_x[\frac{1}{N} \cdot \sum_{i=1}^{N} (x - y_i)^2] = \frac{2}{N} \cdot \sum_{i=1}^{N} (x - y_i)$$

Calcoliamo la derivata seconda:

$$f''(x) = D_x[\frac{2}{N} \cdot \sum_{i=1}^{N} (x - y_i)] = \frac{2}{N} \cdot \sum_{i=1}^{N} 1 = \frac{2N}{N} = 2 > 0 \tag{1}$$

$$\tag{2}$$

Punti stazionari:

$$f'(x) = 0 \iff \frac{2}{N} \cdot \sum_{i=1}^{N} (x - y_i) = 0 \tag{3}$$

$$\iff \sum_{i=1}^{N} x = \sum_{i=1}^{N} y_i \tag{4}$$

$$\iff N \cdot x = \sum_{i=1}^{N} y_i \tag{5}$$

$$\iff x = \frac{1}{N} \cdot \sum_{i=1}^{N} y_i = \overline{y} \tag{6}$$

Ricerca dei punti di minimo e massimo relativo

$$f''(\overline{y}) > 0 \implies \overline{y} \tag{7}$$

è di minimo relativo e quindi per quanto osservato in precedenza $\overline{y}$ è il punto di minimo assoluto.

In definitiva: $\xi_2(\underline{y}) = argmin_{x \in \mathbf{R}} f(x) = \overline{y}$ $\square$

Teorema 23. *Si ponga:*

$$\xi_\infty(\underline{y}) := \lim_{x\to\infty} \xi_r(\underline{y}) \tag{1}$$

Risulta:

$$\xi_\infty(\underline{y}) = \frac{y_{(1)} + y_N}{2} \tag{2}$$

dove:

- $y_{(1)} = min(1, ..., y_N)$
- $y_{(2)} = max(1, ..., y_N)$

La quantità $\frac{y_{(1)}+y_N}{2}$ *si dice "valore centrale di* $\underline{y}$*"*
No Dimostrazione.

30.2 Quartile

La mediana di una rilevazione dati è il quartile di ordine 2: Q_2. La mediana suddivide $\underline{\text{y}}$ in due parti: i dati "minori" di essa e i dati "maggiori" di essa.

La mediana dei dati "minori" di Q_2 si indica con il simbolo Q_1: il quartile di ordine 1. La mediana dei dati "maggiori" di Q_2 si indica con il simbolo Q_3: il quartile di ordine 3.

Si completa chiamando $y(1)$ quartile di ordine 0:

$$y_{(1)} = Q_0$$

e chiamando $y_{(N)}$ quartile di ordine 4:

$$y_{(N)} = Q_4$$

Nota bene un quartile corrisponde a 1/4 della quantità di dati.

30.3 Indici di Dispersione

Nel caso di caratteri qualitativi quello che viene maggiormente utilizzato è "**l'indice di ricchezza**", ovvero il numero delle modalità con le quali essi si manifesta. (Ad esempio i gruppi sanguigni hanno indice di ricchezza 4 dato che 4 è il numero di modalità che un gruppo sanguigno può avere)

L'indice di ricchezza è un esempio della famiglia degli indici di diversità.

Per i caratteri quantitativi ce ne sono vari:

- Campo o intervallo di variazione (riferito al valore centrale)
  $\Gamma := y_{(N)} - y_{(1)} = Q_4 - Q_0$.
- Differenza tra il terzo e il primo quartile ("differenza interquartilica")
  $\gamma := Q_3 - Q_1$ (riferito alla mediana).
- "Scarto mediano assoluto" (riferito alla mediana)
  $S_{Q_2} := \frac{1}{N} \cdot \sum_{i=1}^{N} |y_i - Q_2|$
- "Scarto media assoluto" (riferito alla media)
  $S_{\overline{y}} := \frac{1}{N} \cdot \sum_{i=1}^{N} |y_i - \overline{y}|$
- Varianza (riferito alla media)
  $S_{\overline{y}}^2 := \frac{1}{N} \cdot \sum_{i=1}^{N} (y_i - \overline{y})^2$
- Varianza (altra forma)
  $S_{\overline{y}} := \sqrt{\frac{1}{N} \cdot \sum_{i=1}^{N} (y_i - \overline{y})^2}$

30.4 Diagramma scatola con baffi

Il diagramma scatola con baffi è possibile visionarlo sia in verticale che in orizzontale ed è facilmente consultabile qui[1].

Elementi:

- VAS - Valore Adiacente Superiore.
- VAI - Valore Adiacente Inferiore.
- Quartili.
- Altezza pari ad una delle misure di dispersione.
- Segmenti $1.5 \cdot \Phi$ (dove gamma equivale ad una misura di dispersione)

[1] https://it.wikipedia.org/wiki/Diagramma_a_scatola_e_baffi

31 Lezione 31

31.1 Medie Analitiche

Sia $\underline{y} = (y_1, y_2, ..., y_n)$, "qual è la media dei dati?" è una domanda non formulata completamente. Bisogna fornire anche un criterdio di mediazione. Il **criterio di mediazione** è detto "funzione di circostanza".

La funzione di circostanza ha come argomento la rilevazione dati. essa si designa, di solito, con:

$$C(\underline{y})$$

Sia y la media da calcolare e sia $\underline{y}*$ (una rilevazione dati falsa):

$$\underline{y}* = (y, y, ..., \underline{y})$$

allora y è soluzione dell' **"equazione di circostanza"**:

$$C(\underline{y}) = C(\underline{y}*)$$

Ogni funzione dei dati che è soluzione di una funzione di circostanza è una **"media analitica"**.

31.1.1 Esempio 1

Asseganta $\underline{y} = (y_1, ..., y_N)$ è assegnato il criterio di mediazione:

$$C(\underline{y}) = \sum_{i=1}^{N} y_i$$

la media analitica corrispondente è la media aritmetica.
Infatti:

$$\begin{aligned} C(\underline{y}) = C(\underline{y}*) \Longrightarrow \sum_{i=1}^{N} y_i &= \sum_{i=1}^{N} y \\ &= \sum_{i=1}^{N} y_i = N \cdot y \end{aligned}$$

Totale:

$$y = \frac{1}{N} \cdot \sum_{i=1}^{N} y_i \equiv \overline{y}$$

che coincide con la media aritmetica.

31.1.2 Esempio 2

Asseganta $\underline{y} = (y_1, ..., y_N)$ è assegnato il criterio di mediazione:

$$C(\underline{y}) = \prod_{i=1}^{N} y_i$$

la media analitica corrispondente è la media geometrica.
Infatti:

$$\begin{aligned} C(\underline{y}) = C(\underline{y}*) \Longrightarrow \prod_{i=1}^{N} y_i &= \prod_{i=1}^{N} y \\ &= \prod_{i=1}^{N} y_i = y^N \end{aligned}$$

Totale:

$$\underline{y} = \sqrt[N]{\prod_{i=1}^{N} y_i} \equiv Mg \text{ media geometrica}$$

che coincide con la media geometrica, unico **vincolo** i dati sono tutti positivi.

31.1.3 Esempio 3

Asseganta $\underline{y} = (y_1, ..., y_N)$ è assegnato il criterio di mediazione:

$$C(\underline{y}) = \sum_{i=1}^{N} \frac{1}{y_i}$$

la media analitica corrispondente è la media armonica, ovvero, il reciproco della media aritmetica dei reciproci.
Infatti:

$$\begin{aligned} C(\underline{y}) = C(\underline{y}*) \Longrightarrow \sum_{i=1}^{N} \frac{1}{y_i} &= \sum_{i=1}^{N} \frac{1}{y} \\ &= \sum_{i=1}^{N} \frac{1}{y_i} = N \cdot \frac{1}{y} \\ &\Longleftrightarrow \frac{1}{y} = \frac{1}{N} \cdot \sum_{i=1}^{N} \frac{1}{y_i} \end{aligned}$$

Totale:

$$y = (\frac{1}{N} \cdot \sum_{i=1}^{N} \frac{1}{y_i})^{-1} \equiv Ma \text{ media armonica}$$

che coincide con la media armonica.

31.1.4 Altri esempi di funzioni di circostanza

- la somma dei quadrati dei dati - (quadratica);
- la somma dei quadrati dei dati rapportata alla somma dei dati - (anti-armonica);
- la somma dei dati trasformati tramite le funzioni esponenziali - $(e^{y_1}, ..., e^{y_N})$;

31.2 Statistica Inferenziale

Supponiamo che esistono finiti valori medi considerati nel seguito, si ricordi che:

- $E(aX+b) = a \cdot E(X) + b$;
- $E(X_1 + X_2) = E(X_1) + E(X_2)$;
- $n \in \mathbf{N}, E(X_1 + ... + X_n) = E(X_1) + \cdots + E(X_n)$ ovvero:

$$E(\textstyle\sum_{i=1}^{N} x_i) = \sum_{i=1}^{N} E(X_i)$$

Sappiamo poi che:

$$D^2(aX+b) = a^2 \cdot D^2(X)$$

Definizione La quantità:

$$Cav(X_1, X_2) = E\{[(X_1 - \mu_{X_1}) \cdot (X_2 - \mu_{X_2})]\}$$

Viene chiamata **COVARIANZA**, ovviamente se X_1 e X_2 sono uguali il risultato sarà la varianza stessa:

$$Cov(X_1, X_1) = E[(X_1 - \mu_{X_1})] = D^2(X_1)$$

Teorema 24. *Citando la ddefinizione precedente risulta che:*

$$D^2(X_1, X_2) = D^2(X_1) + D^2(X_2) + 2Cov(X_1, X_2) \tag{1}$$

Dimostrazione.

$$D^2(X_1 + X_2) = E\{[(X_1 + X_2) - (\mu_{X_1} - \mu_{X_2})]^2\} \tag{2}$$
$$= E\{[(X_1 - \mu_{X_1}) + (X_2 - \mu_{X_2})]\} \tag{3}$$
$$= E[(X_1 - \mu_{X_1})^2 + (X_2 - \mu_{X_2})^2 + 2(X_1 - \mu_{X_1}) \cdot (X_2 - \mu_{X_2})] \tag{4}$$
$$= E[(X_1 - \mu_{X_1})^2] + E[(X_2 - \mu_{X_2})^2] + 2E[(X_1 - \mu_{X_1}) \cdot (X_2 - \mu_{X_2})] \tag{5}$$
$$= D^2(X_1) + D^2(X_2) + 2Cov(X_1, X_2) \tag{6}$$

□

31.2.1 Definizioni e Proposizioni

Definizione X_1 e X_2 sono indipendenti se tutti gli eventi generati da X_1 sono indipendenti da ogni evento generato da X_2.

Proposizione X_1 e X_2 sono indipendeti se e solo se:

$$X_1, X_2 \in \mathbf{R},\ \mathcal{P}(X_1 \le x_1, X_2 \le x_2) = \mathcal{P}(X_1 \le x_1) \cdot \mathcal{P}(X_2 \le x_2)$$

Proposizione Se X_1 e X_2 sono indipendenti:

$$Cov(X_1, X_2) = 0$$

Corollario Se X_1 e X_2 sono indipendenti allora:

$$D^2(X_1 + X_2) = D^2(X_1) + D^2(X_2)$$

31.3 Descrittiva vs Inferenziale

La statistica descrittiva ruota attorno due concetti:

- Ω - Popolazione.
- Y - Carattere.
- $\underline{y} = (y_1, ..., y_N)$.

La statistica inferenziale parametrica mira a determinare delle formule per ottenere **"stime"** del parametro o più in generale di una funzione del parametro.

- Ω - Insieme degli eventi elementari.
- X - Genitrice.
- $(x_1, ..., x_n)$ n-pla di variabili somiglianti a X.

Infatti: $(\Omega, \mathcal{F} = \sigma(g), \mathcal{P}_\theta)$ dove:

$\theta \in \Theta$ e si chiama **"regione parametrica"**

31.3.1 Esempio

Si sa che una certa moneta è truccata con trucco incognito., allora X (la genitrice) ha legge B(1,p):

$$p \in (0, 1)$$

La regione parametrica è l'intervallo (0,1):

$$\Theta = (0, 1)$$

Facciamo n lanci della moneta per ottenere n repliche di X:

$$X_1, X_2, - - -, X_{10} : X_1 \sim X_2 \sim ... \sim X_{10} \sim X \\ \sim B(1, p)$$

- $t(x_1, ..., x_n)$ è una v.a;
- $p \approx t(x_1, ..., x_{10})$ è un numero;

che approssima la probabilità di T per la moneta considerata.

31.4 Vari reference

Il professore ha consigliato un nuovo libro di testo per la parte di statistica consultabile qui[2].

32 Lezione 32

La lezione inizia con uno schema approfondito della differenza tra la statistica descrittiva e inferenziale:

Statistica descrittiva

- Ω : popolazione;
- Y: carattere;
- $\underline{y}=(y_1, ..., y_N)$ - rilevazione dati;
- N: taglia di Ω;

Statistica inferenziale parametrica

- Ω : spazio campione;
- X: genitrice;
- $\underline{X} = (X_1, ..., X_m)$ - Campione casuale semplice (c.c.s) oppure osservazioni, sono indipendenti e somiglianti a X;
- $\underline{x}=(x_1, ..., x_n)$ - realizzazione del campione;
- n = taglia del campione;
- $T=t(\underline{x})$ - statistica se viene usata a scopo di stima è chiamata stimatore;

32.0.1 Definizione di statistica

La disciplina è denominata "STATISTICA" in quanto traduzione del vocabolo inglese "STATISTICS", ovvero è disciplina che "tratta" le funzioni di un campione casuale.

32.0.2 Esempio

Per illustrare la differenza tra punto campione, campionatura e realizzazione del campione si immagini di avere la seguente legge:

$$X \sim B(1, p), \ n = 10$$

si considera il lancio di una moneta ripetuta, ogni risultato oltre ad avere un valore indipendente dal precedente e dal successivo considera che ogni X (da 1 a 10) è simile a quello precedente e successivo e sono tutti indipendenti.

32.0.3 Esempi di Statstiche

$T = X_1 + X_2 + \cdots + X_n$ la somma delle osservazioni.

$$\overline{x} = \frac{T}{n} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i \textit{ (media campionaria)}$$

quindi:

$$k \in \mathbf{N},\ \overline{X}^{(k)} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i^k \textit{ (media empirica)}$$

viene detto "momento empirico di ordine k"
ovviamente $\overline{X}^{(1)} = \overline{X}$ e $\overline{X}^{(1)}$ è l'alter ego "campionario" di:

$$\mu_X = \mu_1' = E(X)$$

$k \in \mathbf{N}$, $\overline{X}^{(k)}$ è l'alter ego "statistico" di:

$$\mu_k' = E(X^k) \textit{ (momento teorico)}$$

viene detto momento teorico di ordine k.

32.0.4 Risultato

Si ha che:

$$k \in \mathbf{N},\ \overline{X}^{(k)} \Longrightarrow_{n\to\infty} \mu_k' = E(X^k)$$

In altri termini, al crescere della taglia n il momento campionario di ordine tende (in un senso che sarà poi specificato nelle prossime lezioni) a un valore deterministico che coincide con il momento teorico di ordine k (se esistente e finito).

Questo risultato p la base del "metodo dei momenti" che è utile per la costruzione degli stimatori.

32.1 Metodo dei momenti

Si preferisce far precedere l'esposizione del metodo da alcuni esempi:

32.1.1 Esempio 1

Definizione

$$X \sim B(1, p),\ p \in (0,1)\ \textit{incognito.}$$

Fatto teorico

$$\mu_1' = E(X) = p$$

Fatto teorico (sperimentale)

$$(X_1, X_2, ..., X_n)$$

questa è la c.c.s e $X_1 \sim X$.

Metodo sostituire a μ_1' con $\overline{X}^{(1)} = \overline{X}$, l'equazione per tanto diventa:

$$\overline{X} \approx p$$

da cui scaturisce la seguende definizione:

$$\hat{P}_{MM} := \overline{X} = \frac{X_1 + X_1 + ... + X_m}{n} \ - \text{ media campionaria}$$

Nota bene $\hat{p}_{MM} \neq p$.

Conclusione Lo stimatore del parametro di una v.a di Bernulli individuato con il metodo dei momenti p la media campionaria.

Quando si otterrà la realizzazione del campione $\underline{x} = (x_1, x_2, ..., x_n)$ allora:

$$\hat{p}_{MM} = \overline{x} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i \ - \text{ media aritmetica}$$

che fornisce la sima di p.

32.1.2 Esempio 2

Definizione

$$X \sim G(p),\ p \in (0,1)\ \textit{incognito.}$$

Fatto teorico

$$\mu_1' = E(X) = \frac{1}{p}$$

Fatto teorico (sperimentale)

$$(X_1, X_2, ..., X_n)$$

questa è la c.c.s e $X_1 \sim X$.

Metodo sostituire μ_1' con $\overline{X}^{(1)} = \overline{X}$, l'equazione diventa:

$$\overline{X} \approx \frac{1}{p}$$

da cui scaturisce la seguente definizione:

$$\hat{p}_{MM} := \frac{1}{\overline{X}} = (\overline{X})^{-1} \cdot (\frac{X_1 + X_2 + ... + X_n}{n})^{-1} \text{ – reciproco della media campionaria}$$

Nota bene $\hat{p}_{MM} \neq p$.

Conclusione Lo stimatore del parametro di una v.a geometrica individuato con il metodo dei momenti è il reciproco della media campionaria.

Quando si otterrà $\underline{x} = (x_1, x_2, ---, x_n)$ allora:

$$\hat{p}_{MM} = \frac{1}{\overline{x}} = (\frac{1}{n} \cdot \sum_{i=1}^{n} x_i)^{-1} \text{ – media aritmetica}$$

fornisce la stima di p.

32.1.3 Esempio 3

Definizione

$$X \sim \Pi(\lambda), \ \lambda > 0 \ \textit{incognito}$$

Fatto teorico

$$\mu_1' = \lambda$$

Fatto teorico (sperimentale)

$$\underline{X} = (X_1, X_2, ..., X_n)$$

sostituire μ_1' con $\overline{X}$:

$$\overline{X} \approx \lambda$$

Metodo

$$\hat{\Lambda}_{MM} := \overline{X}, \ (\hat{\lambda}_{MM} \neq \lambda)$$

lo stimatore sarà $\hat{\lambda}_{MM} = \overline{x}$

Conclusione Lo stimatore del parametro di una v.a di Poission individuato con il metodo dei momenti è la media campionaria.

32.1.4 Esempio 4

Definizione

$$X \sim N(\mu, \sigma^2),\ \mu \in \mathbf{R},\ \sigma^2 > 0\ \textit{incogniti}$$

Fatto teorico

- $E(X) = \mu$
- $D^2(X) = \sigma^2 \Longleftrightarrow E(X^2) - E^2(X) = \sigma^2$

  definaimo un sistema tendendo in considerazione che $E(X^2) = \sigma^2 + \mu^2$.

$$\begin{cases} E(X) = \mu \\ E(X^2) = \sigma^2 + \mu^2 \end{cases} \Longleftrightarrow \begin{cases} \mu = E(X) \\ \sigma^2 = E(X^2) - \mu^2 \end{cases} \Longleftrightarrow \begin{cases} \mu = E(X) \\ \sigma^2 = E(X^2) - E^2(X) \end{cases}$$

Fatto teorico (sperimentale)

$$\underline{X} = (X_1, ..., X_n),\ X_1 \sim X$$

questa è la c.c.s e $X_1 \sim X$.

Metodo sostituire a μ_1' con $\overline{X}$ E $\mu_2' = E(X^2)$ con $\overline{X}^{(2)}$:

$$\begin{cases} \hat{\mu} \approx \overline{X} \\ \hat{\sigma}^2 \approx \overline{X}^{(2)} - (\overline{X})^2 \end{cases}$$

risulati:

- stimatore per $\mu \Longrightarrow \hat{\mu}_{MM} := \overline{X}$
- stimatore per $\hat{\sigma}^2_{MM} := \overline{X}^{(2)} - (\overline{X})^2$

Conclusione Lo stimatore del parametro bidimensionale di una v.a di Gauss individuato con il metodo dei momenti p la coppia costituita dalle statistiche mediie compionarie dele secondo ordine e il quadrato della media campionaria.

Se si desidera ottenere uno stimatore per σ basta osservare che:

$$\hat{\sigma}_{MM} = \sqrt{\hat{\sigma}^2_{MM}} = f(\sigma^2)$$

Nota bene la radice quadrata è funzione contina in tutto il suo dominio.

e applicare la stessa funzione allo stimatore di σ^2 individuato dal metodo dei momenti:

$$\begin{aligned}\hat{\sigma}_{MM} = f(\sigma^2_{MM}) &= \sqrt{\hat{\sigma^2}_{MM}} \\ &= \sqrt{\overline{X}^{(2)} - (\overline{X})^2} = g(\underline{X}) \ \text{ – è una funzione del campione}\end{aligned}$$

è uno stimatore "buono" per la deviazione standard: aumentando la taglia del campione "aumenta" la precisione della stima.

33 Lezione 33

33.1 Metodo dei momenti

Sia X una v.a di cui conosciamo la legge a meno di una parte della sua configurazione parametrica:

$$(\theta_1, ..., \theta_n) \ \ incognita$$

Esempio $X \sim N(0.2; \sigma^2)$ e consideriamo $\theta_1 = \sigma^2$ (non è completamente specificata, dato che conosciamo solo il valore di $\mu = 0.2$ e non di σ^2).

Tra i parametri incogniti e i momenti teorici esiste una relazione teorica; si supponga che esistono finiti i primi q e si consideri che $q \geq r$ (per poter risolvere il sistema di q equazioni in r incognite) allora si presenta:

$$\begin{cases} \mu_1' = E(X) = f_1(\theta_1, ..., \theta_r) \\ \mu_2' = E(X^2) = f_2(\theta_1, ..., \theta_r) \\ ... \\ \mu_q' = E(X^q) = f_q(\theta_1, ..., \theta_r) \end{cases}$$

Risolvendo il sistema avremo che:

$$\begin{cases} \theta_1 = g_1(\mu_1', ..., \mu_q') \\ \theta_2 = g_2(\mu_1', ..., \mu_q') \\ ... \\ \theta_r = g_r(\mu_1', ..., \mu_q') \end{cases}$$

Allora il metodo dei momenti richiede la sostituzione dei momenti teorici con quelli empirici/campionari:

$$\begin{cases} \Theta_1 \approx g_1(\overline{x}^{(1)}, ..., \overline{x}^{(q)}) \\ \Theta_2 \approx g_2(\overline{x}^{(1)}, ..., \overline{x}^{(q)}) \\ ... \\ \Theta_r \approx g_r(\overline{x}^{(1)}, ..., \overline{x}^{(q)}) \end{cases}$$

Ciò è giustificato dal fatto che se ho J=(1,2,...q), $\overline{X}^{(j)} \Longrightarrow_{n\to\infty} \mu_j' = E(X^j)$
Se la taglia è grande, è vero che il momento di ordine J è una variabile aleatoria: più aumenta la taglia, più la v.a. diventa un numero reale.

Dopo ciò, si pone:

$$\begin{cases} \hat{\Theta}_{1,MM} := g_1(\overline{x}^{(1)}, ..., \overline{x}^{(q)}) \\ \hat{\Theta}_{2,MM} := g_2(\overline{x}^{(1)}, ..., \overline{x}^{(q)}) \\ ...\hat{\Theta}_{r,MM} := g_r(\overline{x}^{(1)}, ..., \overline{x}^{(q)}) \end{cases}$$

e la r-pla:

$$(\hat{\Theta}_{1,MM}, ..., \hat{\Theta}_{r,MM})$$

è lo stimatore per r-pla dei parametri incognit:

$$(\theta_1, ..., \theta_2)$$

33.2 Riferimenti

La lezione da questo momento si è spostata su degli esempi presi dai libri di testo (che si trovano a fine documento), non vengono riportati dato che non sono completi data la difficoltà dei calcoli ma si evincono varie piccolezze che vanno comunque trattate, per tanto si consiglia di prendere visione dei documenti pdf del professore e delle "lavagne".

33.3 Metodo della massima verosomiglianza

33.3.1 Esempio introduttivo

Consideriamo la v.a $X \sim B(1,p)$ dove p è incognito pero è noto allo sperimentatore che:

$$p = \{\frac{1}{1000}, \frac{999}{1000}\}$$

Lo sperimentatore decide di effettuare 100 lanci ottenendo 100 C, inoltre $p = \mathcal{P}(C)$. Viene più naturale pensare che il valore p sia verosimilmente $\frac{999}{1000}$. Calcoliamo adesso:

$$\begin{aligned}\mathcal{P}_p(X_1 = 1, ..., X_{100} = 1) &= \prod_{i=1}^{100} \mathcal{P}(X_i = 1)\\ &= \prod_{i=1}^{100} p = p^{100}\end{aligned}$$

e quindi:

- $\mathcal{P}_{\frac{1}{1000}}(X_1 = 1, ..., X_{100} = 1) = (\frac{1}{1000})^{100} = 10^{-300}$;
- $\mathcal{P}_{\frac{999}{1000}}(X_1 = 1, ..., X_{100} = 1) = (\frac{999}{1000})^{100}$;

33.3.2 Formalizzando il tutto nel caso di una genitrice discreta

X ha legge con nota almeno una parte dei suoi parametri, allora:

$$\underline{X} = (X_1, ..., X_n)\ con\ X_1 \sim X\ -\ c.c.s$$

quindi:

$$\underline{x} = (x_1, ..., x_n)$$

definiamo quindi:

$$\mathcal{P}(X_1 = x_1, ..., X_n = x_n; \underline{\theta}) = \prod_{i=1}^{n} \mathcal{P}(X_i = x_i; \underline{\theta}) \tag{1}$$

in conclusione otteniamo la funzione di verosimiglianza

$$L(\underline{\theta}; \underline{X}) := \prod_{i=1}^{n} \mathcal{P}(X_i = x_i; \underline{\theta})$$

Ciò implica (riferendosi all'esempio precedente):

- verosomiglianza p=1/1000 di (1,1,...,1) = $\mathcal{P}$(1/1000)...
- verosomiglianza p=999/1000 di (1,1,...,1) = $\mathcal{P}$(999/1000)...

finamo quindi con le definizioni:

- $\underline{\hat{\Theta}}_{MV} = argmax_{\theta \in \Theta} L(\underline{\theta; \underline{x}})$ - stima di massima verosomiglianza.
- $\underline{\hat{\Theta}}_{MV} = argmax_{\theta \in \Theta} L(\underline{\theta; \underline{X}})$ - stimatore di massima verosomiglianza.

33.3.3 Conclusione ed esempio

Consideriamo $X \sim \Pi(\lambda),\ \lambda > 0 \Longrightarrow \Theta = (0, +\infty)$ – *incognito.*

$$L(\lambda; \underline{X}) = \prod_{i=1}^{n} \mathcal{P}(X = x_i; \lambda) = \prod_{i=1}^{n} e^{-\lambda} \cdot \frac{\lambda^n}{n!} \Longrightarrow \prod_{i=1}^{n} e^{-\lambda} \cdot \frac{\lambda^{x_i}}{x_i!}$$

$$A = x_1! \cdots x_n = \frac{e^{-n\lambda}}{A} \cdot (\lambda^{x_1} \ldots \lambda^{x_n})$$

$$t = x_1 + \ldots + x_n = e^{-m\lambda} \cdot \frac{\lambda^n}{A}$$

in conclusione:

$$X > 0,\ L(\lambda; \underline{X}) = \frac{1}{A} \cdot e^{-n\lambda} \cdot \lambda^t$$

Punti stazionari

$$\begin{aligned} L'(\lambda; \underline{x}) = 0 &\Longleftrightarrow \frac{1}{A}(-ne^{-n\lambda} \cdot \lambda^t + e^{-n\lambda} \cdot t\lambda^{t-1} = 0 \\ &\Longleftrightarrow e^{-n\lambda} \cdot \lambda^{t-1}(-n\lambda + t) = 0 \\ &\Longleftrightarrow -n\lambda + t = 0 \\ &\Longleftrightarrow \lambda = \frac{t}{n} = \frac{1}{n} \cdot (x_1 + \ldots + x_n) = \overline{x} \end{aligned}$$

Frontera

$$\begin{aligned} \lim_{\lambda \to 0^+} L(\lambda; \underline{x}) &= \lim_{\lambda \to 0^+} \frac{1}{A} \cdot e^{-n\lambda} \cdot \lambda^t \\ &= \frac{1}{A} \cdot \lim_{\lambda \to 0^+} e^{-n\lambda} \cdot \lim_{\lambda \to 0^+} \lambda^t = 0 \end{aligned}$$

$$\begin{aligned} \lim_{\lambda \to \infty} L(\lambda; \underline{x}) &= \lim_{\lambda \to \infty} \frac{1}{A} \cdot e^{-n\lambda} \cdot \lambda^t \\ &= \frac{1}{A} \cdot \lim_{\lambda \to \infty} e^{-n\lambda} \cdot \lim_{\lambda \to \infty} \lambda^t = 0 \end{aligned}$$

Massimo e minimo relativo

$$\begin{aligned} L''(\lambda; \underline{x}) &= D_\lambda[\frac{1}{A} \cdot e^{-n\lambda} \cdot \lambda^{t-1} \cdot (-n\lambda + t)] \\ &= \frac{1}{A}[e^{-n\lambda}(-n)\lambda^{t-1}(-n\lambda + t) + e^{-n\lambda}(t-1)\lambda^{t-2}(-n\lambda + t) + e^{-n\lambda}\lambda^{t-1}(-n)] \end{aligned}$$

Notiamo che: $n\overline{x} = n \cdot \frac{t}{n} = n \Longrightarrow -n\overline{x} + t = 0$, quindi:

$$\begin{aligned} L''(\overline{x}, \underline{x}) &= \frac{1}{A}[0 + 0 + e^{-t} \cdot (\overline{x})^{t-1} \cdot (-n)] \\ &= -\frac{n}{A} \cdot (\overline{x})^{t-1} \cdot e^{-t} < 0 \end{aligned}$$

Quindi $\overline{x}$ è un punto di massimo relativo per λ.

Conclusione La funzione di verosomiglianza è positiva in $\Theta = (0, +\infty)$, si annulla sulle frontiere, è dotata di derivata prima in Θ, ha un unico punto di massimo relativo in $\lambda = \overline{x}$ e quindi:

$$argmax_{\lambda>0}\ L(\lambda; \underline{x}) = \overline{x}$$

Di conseguenza

- $\hat{\lambda}_{MV} = \overline{x}$ - stima;
- $\hat{\Lambda}_{MV} = \overline{X}$ - stimatore;

34 Lezione 34

La lezione si è incentrata su una osservazione iniziale ed esempi di applicazione sulle v.a (genitrici) che abbiamo studiato durante il corso. Il testo si limita a trarre eventuali conlusioni ma si chiede di far riferimento alle lavagne revisionate del prof.

34.0.1 Osservazioni

Centro di ordine 2:

$$d_2(x, \underline{y}) = \sqrt{\frac{1}{N} \cdot \sum (x - y_i)^2}$$

quindi definiamo: $\Xi_2 = argmin_{x \in \mathbf{R}} D_2(X, \underline{y})$.

Chiamiamo:

$$f(x) = \frac{1}{N} \cdot \sum (x - y_i)^2$$

allora:

$$\Xi_2 = argmin_{x \in \mathbf{R}} D_2(X, \underline{y}) = argmin_{x \in \mathbf{R}} f(x)$$

dato che la funzione radice è funzione strettamente crescente.

Nota bene La funzione $\ln x$ strettamente crescente, per cui posto:

$$l(\underline{\theta}; \underline{x}) := \ln L(\underline{\theta}; \underline{x})$$

risulta:

$$argmax_{\theta \in \Theta} L(\underline{\theta}; \underline{x}) = argmax_{\theta \in \Theta} l(\underline{\theta}; \underline{x})$$

34.0.2 Come si procede nel caso di una genitrice assolutamente continua?

Si assuma che, se X è dotata di funzione di densità $f_x(x)$, risulta:

$$x \in \mathbf{R}, \mathcal{P}[X \in (x, x + dx)] = f_X(x),\ dx$$

allora:

$$\mathbf{P}[X \in (a, b)]\ \int_a^b f_X(x),\ dx$$

e allora si definisce $L(\underline{\theta}; \underline{x}) = \prod f_x(x_i; \underline{\theta})$

34.1 Vari esempi

Vedere le lavagne del professore per prendere visione degli esempi, infatti le osservazioni precedenti vengono utilizzate su ttute le v.a viste durante il corso, si consiglia di prendere visione anche del libro di statistica di riferimento[2].

35 Lezione 35

La lezione riprende un argomento lasciato nell'ultimo esempio della lezione 34, si consiglia di prendere visione delle lavagne del professore, inoltre conclude la parentesi del momento verosomiglianza anche sulla variabile aleatoria Gaussiana (alternando i parametri noti).

Infine introduce e conclude la lezione con delle proprietà.

35.1 Proprietà degli stimatori

$X \sim (0, b)$:

- $\hat{B}_{MM} = 2\overline{X}$;
- $\hat{B}_{MV} = X_{(n)}$

quale scegliere?

35.1.1 Proprietà di correttezza

$\underline{\mathrm{T}} = \mathrm{g}(\underline{\mathrm{X}})$ sia uno stimatore per la funzione Ψ del parametro θ. Si dice che $\underline{\mathrm{T}}$ è corretto per $\Psi(\underline{\theta})$ se e solos e:

$$\underline{\theta} \in \Theta,\ E(\underline{T}) = \Psi(\underline{\theta})$$

35.1.2 Preposizione

Sia X una genitrice dotata di media finita: $E(X) = \mu \in \mathbf{R}$ ricaviamo:

$$\mu \in \mathbf{R}, E(\overline{X}) = \mu$$

Dimostrazione.

$$\begin{aligned} E(\overline{X}) &= E(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) \\ &= \frac{1}{n}\sum_{i=1}^{n} E(X) = \frac{n \cdot E(x)}{n} = \mu \end{aligned}$$

$\square$

35.1.3 Esempio

Sia $X \sim U(o, b) \implies E(X) = \frac{b}{2}$ allora:

$$E(\hat{B}_{MM}) = E(2\overline{X}))2E(\overline{X}) = 2 \cdot \frac{b}{2} = b$$

Di conseguenza:

$$b > 0,\ E(\hat{B}_{MM}) = b$$

ovvero $\hat{B}_{MM}$ è uno stimatore corretto.

D'altra parte, si dimostra che:

$$b > 0,\ E(\hat{B}_{MV}))E(X_{(n)}) = \frac{n}{n+1} \cdot b$$

e $\hat{B}_{MV}$ è uno stimatore "distorto" di b.
Comunque:

$$\begin{aligned} d &= \frac{n}{n+1} \cdot b - b = b \cdot (\frac{n}{n+1} - 1) \\ &= \frac{n-n-1}{n+1} \cdot b = -\frac{1}{n+1} \cdot b \end{aligned}$$

e quindi:

$$\lim_{n \to \infty} d_{(b)} = 0$$

$X_{(n)}$ non è corretto però esso è asintoticamente corretto. Ma non solo:

$$\begin{aligned} E(x_{(n)}) = \frac{n}{n+1} \cdot b \Longrightarrow E(\frac{n+1}{n} X_{(n)}) &= \frac{n+1}{n} \cdot E(X_{(n)}) \\ &= \frac{n+1}{n} \cdot \frac{n}{n+1} \cdot b = b \end{aligned}$$

concludiamo che:

$$T = \frac{n+1}{n} \cdot X_{(n)}$$

è uno stimatore corretto per b.

Una volta ottenuto uno stimatore, in generale, si può considerare una successione di stimatori:

$$n \in \mathbf{N},\ \overline{X}_n = \frac{1}{n} \sum : i = 1^n X_i$$

quindi avremo:

- n=1, $\overline{X}_1 = X_1$
- n=2, $\overline{X}_2 = \frac{X_1 + X_2}{2}$
- ...

35.2 Definizione

Sia, per ogni $n \in \mathbf{N}$, $T_n = g(\underline{X}_n)$ uno stimatore per una funzione Ψ del parametro θ. Si dice che $(T_n)_{n \in \mathbf{N}}$ è "asintoticamente corretto" per $\Psi(\theta)$ se e solo se:

$$\theta \in \Theta, \ \lim_{n \to \infty} E(T_n) = \Psi(\theta)$$

In altri termini, $(T_n)_{n \in \mathbf{N}}$ è asintoticamente corretto se una successione infinitesima per ogni θ appartentente a Θ:

$$\lim_{n \to \infty} [E(T_n) - \Psi(\theta)] = \lim_{n \to \infty} d_n(\theta) = 0$$

36 Lezione 36

36.1 Esempio

36.1.1 Proposizione

Sia x una genitrice per la quale:
$E(X^2) < +\infty$ essa ammette media μ e varianza σ^2 finita. Se $\underline{X}$ è un campione casuale semplice con taglia n, si ha che:

$$D^2(\overline{X}) = \frac{D^2(X)}{n} = \frac{\sigma^2}{n}$$

Dimostrazione. Si ha:

$$\begin{aligned} D^2(\overline{X}) &= D^2(\frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{n^2} D^2(\sum_{i=1}^{n} X_i) \\ &= \frac{1}{n^2}\sum_{i=1}^{n} D^2(X_i) = \frac{1}{n^2}\sum_{i=1}^{n} D^2(X) \\ &= \frac{1}{n^2}\cdot n \cdot D^2(X) = \frac{1}{n}\cdot D^2(X) = \frac{\sigma^2}{n} \end{aligned}$$

□

Osservazione 1 Il fatto che una v.a sia dotata d varianza "piccola" la rende un buon predittore della sua media. Nel caso limite di una v.a dotata di varianza nulla, posto $\mu = E(X)$, risulta:

$$\mathcal{P}(X = \mu) = 1$$

Se X è dotata di f.d.p il risultato precedente segue:

$$0 = D^2(X) = \int_{\mathbf{R}} (X-\mu)^2 f_x(x),\ dx \Longrightarrow x - \mu = 0 \Longleftrightarrow x = \mu$$

per $x \in \mathbf{R}$ tranne, al più, per $x \in B$, essendo B un insieme di Borel di $\mathcal{P}_X(B) = 0$. Per le v.a discrete si provede in maniera analoga.

Osservazione 1 La varianza della media campionaria può essere resa arbitrariamente piccola a patto di incrementare la taglia del campione casuale. Torniamo all'esempio e consideriamo un c.c.s di taglia b da una genitrice X: $E(X^2) < +\infty$.

Posto,

$$T_1 = \frac{1}{10}\sum_{i=1}^{1} 0(X_i)$$

si ha che:

- $E(T_1) = \frac{1}{10} \cdot \sum_{i=1}^{10} E(X_i) = \frac{1}{10} \cdot 10 \cdot E(X) = E(X)$
- $D^2(T_1) = \frac{1}{100} \cdot 10 \cdot D^2(X) = \frac{D^2(X)}{10}$

questo è uno dei 3 esempi visti a lezione, se si considerassero anche gli ultimi 10 elementi, il risultato sarebbe identico mentre se si considerassero solo 2 elementi invece sarebbe molto poco attendibile.

36.2 Definizione

Uno stimatore T ottenuto da un c.c.s da una genitrice X e di taglia n, si dice **"CONSISTENTE"** (coerente) se e solo se:

$$\epsilon > 0, \lim_{n \to \infty} \mathcal{P}(|T_n - \Psi(\theta)| \leq \epsilon) = 1$$

Si scirive anche più simmetricamente:

$$T_n \xrightarrow{\mathcal{P}}_{n \to \infty} \Psi(\theta)$$

e si legge "T_n converge in probabilità a $\Psi(\theta)$".

36.3 Proposizione

Se la genitrice del campione dotata della media del secondo ordine finita e $\underline{X}$ è un campione (e T_n è uno stimatore per $\Psi(\theta)$) allora esso p consistente se:

- T_n è asintoticamente corretto;
- La sua varianza è infinitesima;

36.4 Esempio

Se $\mu_k = E(X^k)$ esiste finita, allora:

$$\overline{X}_n^{(k)} = \frac{1}{n}\sum_{i=1}^{n} X_i^k \xrightarrow{\mathcal{P}}_{n\to\infty} \mu_k'$$

36.5 Esempio

Sappiamo che:

$$E(\overline{X}) = E(X)$$

Se esiste finita $D^2(X)$ e se essa è nota allora:

$$\frac{1}{n}\cdot\sum_{i=1}^{n}[X_1-\mu]^2$$

è uno stimatore corretto per σ^2, infatti:

$$\begin{aligned}\sigma^2>0,\ E(T) &= \frac{1}{n}\sigma^2_{i=1}E[(X_i-\mu)^2]\\ &= \frac{1}{n}\sum_{i=1}^{n}E[(X-\mu)^2] = \frac{1}{n}\sum_{i=1}^{n}D^2(X)\\ &= \frac{n}{n}D^2(X) = \sigma^2\end{aligned}$$

Se invece non conosciamo μ, su può considerare:

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$$

Si può dimostrare che:

$$\sigma^2>0,\ E(S_c^2) = \sigma^2$$

Ne scaturisce che:

$$S_d^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$$

non è corretta. Comunque:

$$S_c^2 = \frac{n}{n-1}\cdot S_d^2$$

36.6 Confronto in base al rischio quadratico medio

36.6.1 Proposizione

X è una v.a e c è in numero reale (X dptata del momento di secondo ordine finito), si ha:

$$E[(X-c)^2] = D^2(X) + E(X) - [c - E(X)]^2$$

Dimostrazione.

$$\begin{aligned} E[(X-c)^2] &= E\{[(x-\mu)+(\mu-c)]\} \\ &= E\{[(x-\mu]^2 + (\mu-c)^2 + (\mu-c)\cdot(x-\mu)]\} \\ &= E[(x-\mu)^2] + (\mu-c)^2 + 2(\mu-c)\cdot E(x-\mu) \\ &= D^2(X) + (\mu-c)^2 + 2(\mu-c)\cdot(\mu-\mu) \\ &= D^2(X) + (\mu-c)^2 \end{aligned}$$

□

36.6.2 Definizione

Si definisce rischio quadratico medio di uno stimatore T tratto da una genitrice dodata del momento di secondo ordine finito, la quantità:

$$R_T(\theta) = E\{[T - \Psi(\theta)]^2\},\ con\ \theta \in \Theta$$

Osservazione Il rischio $R_T(\theta)$ è minimo quando lo stimatore è corretto:

$$R_T(\theta) = D^2(T)$$

Siano S e T due stimatori per $\Psi(\theta)$

- S è "**Preferibile**" a T, e si può scrivere:

  $$\mathrm{S} \leq \mathrm{T}$$

  se e solo se:

  $$\theta \in \Theta,\ R_S(\theta) \leq R_T(\theta)$$

- S è "**Strettamente preferibile**" a T, e si può scrivere:

  $$\mathrm{S} \prec \mathrm{T}$$

  se S ≤ T e se esiste un $\theta_0 \in \Theta$ per il quale:

  $$R_S(\theta_0) < R_T(\theta_0)$$

- Uno stimatore si dice "**Ammissibile**" se non esiste un altro stimatore a lui strettamente preferibile.

37 Lezione 37

37.1 Proposizione 1

Se la genitrice X di un campione casuale è dotata di momento del secondo ordine finito, allora la media campionaria $\overline{X}$ è uno stimatore consistente per la media μ di X.

Dimostrazione. Dal momenti che:

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

si ha:

$$\begin{aligned} \mu \in \mathbf{R},\ E(\overline{X}) &= \frac{1}{n} \cdot \sum_{i=1}^{n} E(X_i) \\ &= \frac{1}{n} \cdot \sum_{i=1}^{n} E(X) \qquad = \frac{n}{n} \cdot E(X) = \mu \end{aligned}$$

Ovvero $\overline{X}$ è uno stimatore corretto per μ (e quindi, a maggior ragione, anche asintoticamente corretto).

D'altra parte, sappiamo che:

$$D^2(\overline{X}) = \frac{D^2(X)}{n}$$

e quindi:

$$\lim_{n\to\infty} D^2(\overline{X}) = \lim_{n\to\infty} \frac{1}{n} \cdot D^2(X) = 0$$

Ne discende che la ... è sufficiente per X_n (è uno stimatore sufficiente). □

37.2 Proposizione 2

Se la genitrice X di un campione casuale è dotata del momento di quarto ordine finito, allora il momento empirico del secondo ordine ($\overline{X}^{(2)}$) è uno stimatore consistente per il momento teorico del secondo ordine μ_2'.

$\mu_k' = E(X)$

Dimostrazione. Dal momenti che:

$$n \in \mathbf{N},\ \mu_2' = E(X^2),\ \overline{X}_n^{(2)} = \frac{1}{n} \cdot \sum_{i=1}^{n} X_i^2$$

si ha:

$$\mu \in \mathbf{R},\ E(\overline{X}^{(2)}) = \frac{1}{n} \cdot \sum_{i=1}^{n} E(X_i^2)$$
$$= \frac{1}{n} \cdot \sum_{i=1}^{n} E(X^2) \qquad = \frac{n}{n} \cdot E(X^2) = \mu_2'$$

Ovvero $\overline{X}_n^{(2)}$ è uno stimatore corretto per μ_2', per ogni n (e quindi, a maggior ragione, anche asintoticamente corretto).

D'altra parte, sappiamo che:

$$D^2(\overline{X}_n^{(2)}) = \frac{\mu_{2n}' - (\mu_2')^2}{n}$$

e quindi:

$$\lim_{n\to\infty} D^2(\overline{X}_n^{(2)}) = \frac{\mu_4' - (\mu_2')^2}{n} = 0$$

Ovvero la varianza di $\overline{X}_n^{(2)}$ è infinitesima. L atesi segue che esso è uno stimatore sufficiente. □

37.3 Proposizione 3

Se la genitrice X di un campione casuale è dotata del momento di ordine 2k finito, allora il momento empirico di ordine k ($\overline{X}^{(k)}$) è uno stimatore consistente per il momento teorico di ordine k.

$\mu_k' = E(X)$

Dimostrazione. Tutto resta inalterato, considerando che si può dimostrare che:

$$n \in \mathbf{R},\quad D^2(\overline{X}_n^{2K}) = \frac{\mu_{2k}' - (\mu_k')^2}{n}\ (n \to \infty)$$

□

37.4 Esempio

37.4.1 Domanda

Ci possono essere dure o più stimatori ammissibili per $\Psi(\theta)$?

37.4.2 Risposta

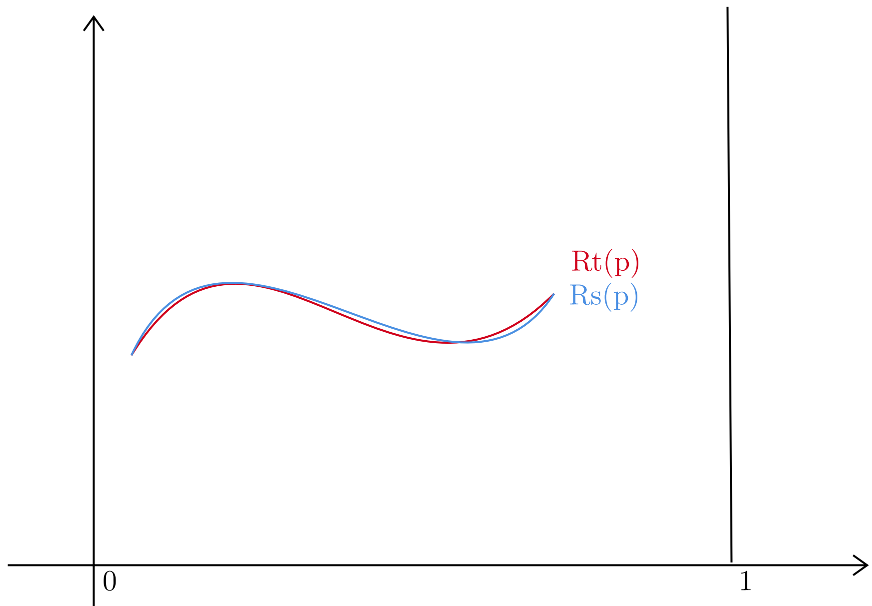
La risposta è affermativa. $\leq$ è una relazione d'ordine che è diversa dalla relazione d'ordine $\leq$ tra due numeri reali che è totale.

T e S sono due stimatori per $\Psi(\theta)$ non è detto che si può affermare:

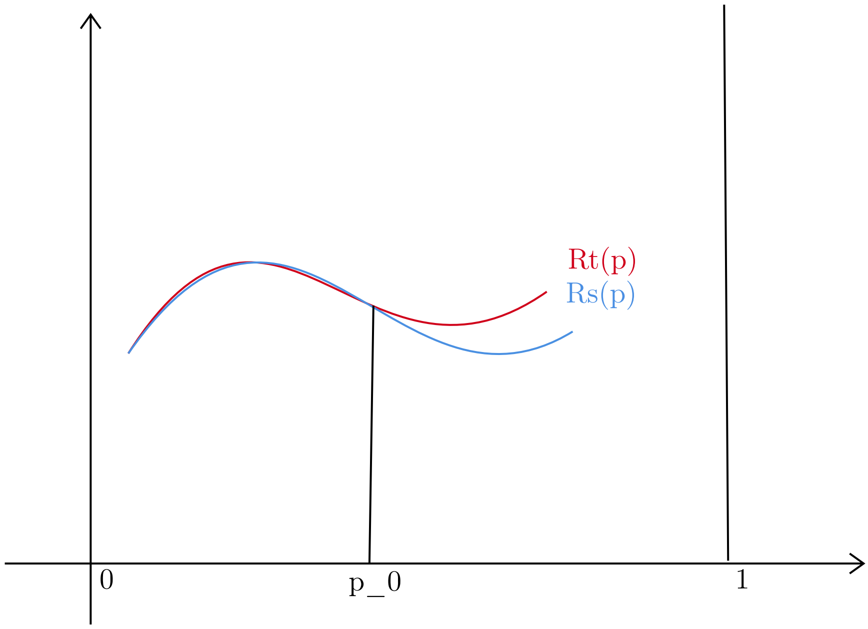
$$S \leq T \text{ } oppure \text{ } T \leq S$$

37.4.3 Dimostrazione

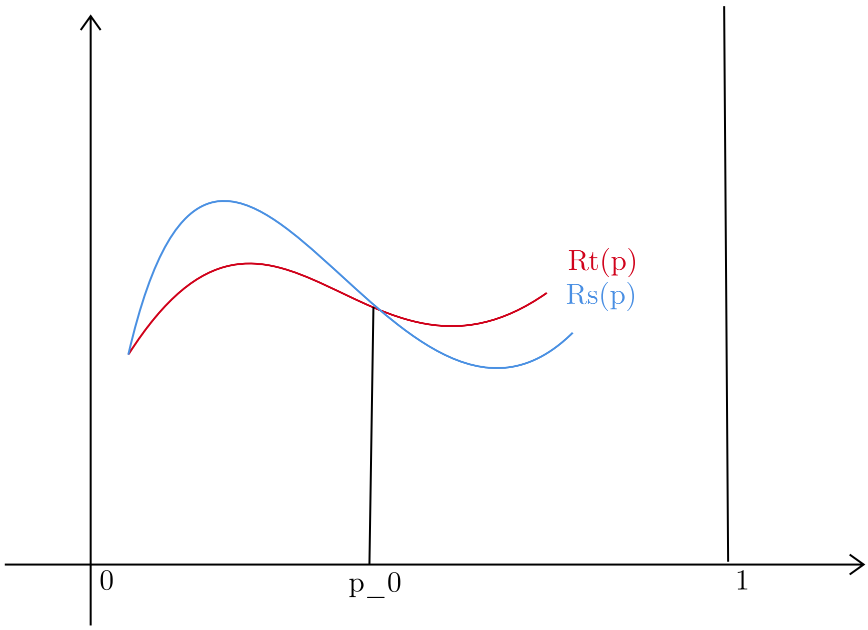
Sia X una genitrice con legge B(1,p).
Allora $\Theta = (0,1)$. Siano S e T due stimatori con le seguenti funzioni di rischio:



Esempio 1
Definizione: $S \leq T$
Infatti, $R_S(p) = R_T(p)$ con $p \in (0,1)$.



Esempio 2
Definizione: $S < T$
Infatti, $R_S(p) \leq R_T(p)$ con $p \in (0,1)$, consideriamo un punto p_0, per $p > p_0$, $R_S(p) < R_T(p)$.



Esempio 3
Definizione: Non confrontabili
Infatti, per $0 < p < p_0$ e $p_0 < p < 1$.

Riferimenti bibliografici

[1] S. Ross, C. Mariconda, and M. Ferrante, *Calcolo delle probabilità.* Idee & strumenti, Apogeo, 2007.

[2] R. Giuliano, *Argomenti di probabilità e statistica.* Springer Milan, 2012.