DAL CAPITOLO 3

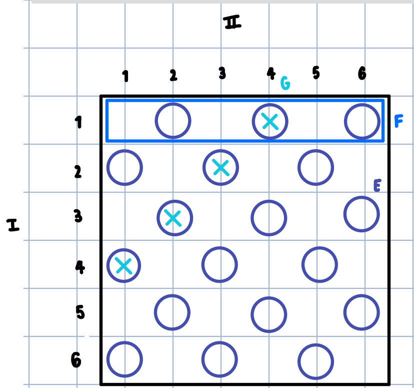
Spazio degli esiti = insieme di tutti gli esiti possibili
Eventi= sottoinsiemi dello spazio degli esiti

L'unione di due eventi E,F $(E \cup F)$ dello stesso spazio degli esiti S, è definita come l'insieme formato dagli esiti che stanno o in E o in F.
L'evento $(E \cup F)$ si verifica se almeno uno tra E e F si verifica.

$E \cap F$ = insieme formato dagli esiti che stanno in E e in F.
L'evento $E \cap F$ rappresenta il verificarsi di entrambi gli eventi.

Se un evento non contiene esiti possibili ($E \cap F = \emptyset$), ovvero se E e F non possono verificarsi entrambi, si dicono mutualmente esclusivi o eventi disgiunti.

4. Si tirano due dadi. Sia E l'evento che la somma dei punteggi sia pari, F che il primo dado realizzi un 1, e G che la somma sia 5. Si descrivano gli eventi
(a) $E \cap F$; **(b)** $E \cup F$; **(c)** $F \cap G$; **(d)** $E \cap F^c$; **(e)** $E \cap F \cap G$.



$S = [6]^2$ $\quad \#S = 3$

$\#E = 18$ $\quad P(E) = \frac{1}{2}$

$\#F = 6$ $\quad P(F) = \frac{1}{6}$

$\#G = 4$ $\quad P(G) = \frac{1}{9}$

a) $E \cap F = \{(1,2), (1,4), (1,6)\}$

b) $E \cup F = \{(1,1), (1,2), (1,3), (1,4), (1,5), (1,6), (2,1), (2,3), (2,5), (3,2), (3,4), (3,6), (4,1), (4,3), (4,5), (5,2), (5,4), (5,6), (6,1), (6,3), (6,5)\}$

c) $F \cap G = \{(1,4)\}$

d) $E \cap F^c = \{(2,1), (2,3), (2,5), (3,2), (3,4), (3,6), (4,1), (4,3), (4,5), (5,2), (5,4), (5,6), (6,1), (6,3), (6,5)\}$

e) $E \cap F \cap G = \{(1,4)\}$

Leggi di De Morgan:

1. $(E \cup F)^c = E^c \cap F^c$

2. $(E \cap F)^c = E^c \cup F^c$

Si associa ad ogni evento E sullo spazio degli esiti S, un numero che si denota con P(E) e che si dice probabilità dell'evento E.
Le probabilità devono rispettare degli assiomi:
1- $0 \leq P(E) \leq 1$
Afferma che ogni probabilità è un numero compreso tra 0 e 1
2- $P(S) = 1$
Stabilisce che l'evento S si verifica con probabilità 1, ovvero che vi è assoluta certezza che si realizzi

un esito contenuto in S OPPURE S contiene tutti necessariamente tutti gli esiti possibili del nostro esperimento

3- $P(\bigcup_{i=1}^{n} E_i) = \sum_{i=1}^{n} P(E_i)$ $\quad n = 1, 2, 3, \dots, \infty$

Preso un insieme finito o numerabile di eventi esclusivi, la probabilità che se ne verifichi almeno uno è uguale alla somma delle loro probabilità

Possiamo notare che E e E^c sono eventi disgiunti, e quindi:

$$1 = P(S) = P(E \cup E^c) = P(E) + P(E^c)$$

PROPOSIZIONE: Per ogni evento $E \subset S$, vale la relazione:

$$P(E^c) = 1 - P(E)$$

La probabilità che un evento non si verifichi è pari a uno meno la possibilità che non si verifichi

PROPOSIZIONE: se E e F sono due eventi qualsiasi, allora:

$$P(E \cup F) = P(E) + P(F) - P(E \cap F)$$

Esempio:

La percentuale di maschi americani che fuma la sigaretta è del 28%, quelli che fumano il sigaro sono il 7%, quelli che fumano entrambi sono il 5%. Qual è la percentuale di chi non fuma né la sigaretta né il sigaro?

E= sigaretta   F= sigaro

$$P(E \cup F) = P(E) + P(F) - P(E \cap F) = 0.07 + 0.28 - 0.05 = 0.3$$

SPAZI DI ESITI EQUIPROBABILI

Equiprobabilità: evento che ha la stessa probabilità di verificarsi di un altro

S= insieme finito $= \{1, 2, 3, \dots, N\}$

$$P(\{1\}) = P(\{2\}) = \dots = P(\{N\}) =: p$$

Dagli assiomi 2 e 3 segue:

$$1 = P(S) = P(\{1\}) + P(\{2\}) + \dots + P(\{N\}) = Np$$

Da qui, si deduce che $P\{(i)\} = p = 1/N$ per tutti gli i=1, 2, …, N

Da qui e dall''assioma 3, si conclude che per ogni evento E,

$$P(E) = \frac{\#E}{N}$$

#E= cardinalità di E

PRINCIPIO DI ENUMERAZIONE:

Se si eseguono r esprimenti, ed è noto che il primo esperimento ammette n_1 esiti possibili, per ognuno dei quali il secondo esperimento ammette n_2 esiti diversi, inoltre se per ogni combinazione di esiti dei primi due esperimenti il terzo ammette n_3 esiti diversi, e così via, allora vi sono un totale di

$n_1 * n_2 * \dots * n_r$ combinazioni degli esiti degli r esperimenti considerati tutti insieme.

Esempio:

Si estraggono a caso due palline da un'urna che ne contiene 6 bianche e 5 nere. Qual è la probabilità che le due estratte siano una bianca e una nera?
Se consideriamo le estrazioni con un ordine, la prima pallina viene scelta tra le 11 presenti nell'urna all'inizio, mentre la seconda tra le 10 che restano dopo la prima estrazione.
Lo spazio degli esiti ha 11*10 = 110 elementi.
Ci sono 6*5 = 30 casi che la prima sia bianca e la seconda sia nera, e 5*6 = 30 che la prima sia nera e la seconda bianca

$$\frac{30+30}{110} = \frac{6}{11}$$

Proviamo a determinare il numero di modi diversi in un si possono ordinare n oggetti, per esempio *a*, *b* e *c*, sono 6. Ciascuno di questi prende il nome di permutazione.
Il primo simbolo può essere scelto in tre modi differenti, il secondo tra i due restanti e il terzo con l'unico che rimane.
Supponiamo di avere n oggetti, allora avremo n! permutazioni degli n oggetti.
RICORDA: 0! = 1

IL COEFFICIENTE BINOMIALE
Poiché il numero diverso di scegliere r oggetti su n tenendo conto dell'ordine è dato da n(n-1) … (n-r+1),e poiché ogni gruppo di lettere fissato viene contato r! volte (una per ogni sua permutazione), il numero di diversi gruppi di r elementi, scelti in un insieme di n oggetti è dato dalla formula

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

Questo valore si dice il numero di combinazioni di n elementi presi r alla volta e prende il nome di coefficiente binomiale
Esempio:
$\binom{8}{2} = \frac{8!}{2!\,6!} = \frac{8\cdot 7}{2\cdot 1} = 28$ Gruppi diversi di due elementi su un insieme di 8.

Inoltre, poiché 0! = 1
$\binom{n}{0} = 1 = \binom{n}{n}$

PROBABILITÀ CONDIZIONATA
Due dadi, lancio un dado ed esce 3, qual é la probabilità che la somma dei due dadi valga 8?
Dato che il primo dado ha dato 3, ci sono solo 6 risultati equiprobabili possibili per l'esperimento.
Ciò significa che, se il primo dado ha dato 3, allora la probabilità condizionata di ciascuno degli esiti possibili è 1/6, mentre la possibilità degli altri 30 elementi di S é 0.
E= somma dei due dati sia 8
F=il primo dato dia 3
Abbiamo calcolato la probabilità condizionata di E dato F

$$P(E|F)$$

In generale:
Se si è verificato l'evento F, affinché si verifichi anche E, il caso favorito un elemento che sta sia in F che in E, ovvero che appartiene a $E \cap F$.

In secondo luogo essendosi verificato F, questo evento diventa il nuovo (ridotto) spazio degli esiti e per questo la probabilità condizionata dell'evento $E \cap F$ sarà pari al rapporto tra la sua probabilità e quella di F.

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

L'equazione ha senso solo se P(F) > 0, in caso contrario P(E | F) non si definisce.

Siano assegnati una quantità finita (o numerabile) di eventi mutualmente esclusivi F_1 , F_2, …, F_n tali che

$$\bigcup_{i=1}^{n} F_i = S$$

Questa proprietà si cita dicendo che gli eventi F ricoprono S e significa che si verifica sempre almeno uno di essi (esattamente uno, se sono disgiunti). Consideriamo un ulteriore evento E, che riscriviamo come

$$E = \bigcup_{i=1} (E \cap F_i)$$

Notando che anche gli eventi $E \cap F_i$ per i= 1, 2, …, n sono mutualmente esclusivi.
Si ottiene:

$$P(E) = \sum_{i=1}^{n} P(E \cap F_i)$$

$$= \sum_{i=1}^{n} P(E|F_i)\, P(F_i)$$

FORMULA DI FATTORIZZAZIONE: è possibile calcolare la probabilità di un evento E condizionando rispetto a quale si verifichi tra un gruppo di eventi accessori mutualmente esclusivi e che ricoprono S.

EVENTI INDIPENDENTI
Nel caso particolare in cui invece P(E|F) e P(E) siano uguali, diciamo che E è indipendente da F.
Quindi E è indipendente da F se la conscenza che F si è avverato non cambia la probabilità di E.
Siccome $P(E|F) = \frac{P(E \cap F)}{P(F)}$, si vede che E è indipendente da F se:

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

Poiché questa equazione è simmetrica in E e F, quando E è indipendente da F, è anche vero il contrario.

DEFINIZIONE: due eventi E d F si dicono indipendenti se vale $P(E \cap F) = P(E)\,P(F)$ altrimenti si dicono dipendenti.
PROPOSIZIONE: se E d F sono indipendenti, lo sono anche E e F^c.

DEFINIZIONE:
Tre eventi E, F, G si dicono indipendenti se valgono tutte e quattro le seguenti equazioni:
1. $P(E \cap F \cap G) = P(E)\,P(F)\,P(G)$

2. $P(E \cap F) = P(E)\,P(F)$

3. $P(E \cap G) = P(E)\,P(G)$

4. $P(F \cap G) = P(F)\,P(G)$

Si noti che se tre eventi E, F e G sono indipendenti, allora ciascuno di essi è indipendente da qualunque evento si possa costruire dagli altri due.

ESERCIZI

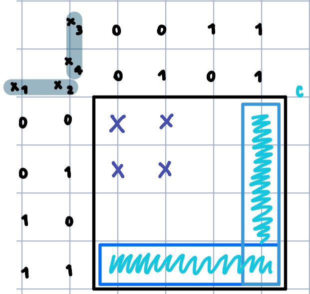
5. Un sistema è composto da 4 componenti, ciascuno dei quali funziona oppure è guasto. Si osserva lo stato dei componenti, ottenendo un vettore (x_1, x_2, x_3, x_4), dove x_i è 1 oppure 0 a seconda che il componente i-esimo funzioni oppure no.

(a) Da quanti elementi è formato lo spazio degli esiti?

(b) Il sistema nel suo insieme funziona fintantoché entrambi i componenti 1 e 2 oppure quelli 3 e 4 funzionano. Specifica tutti gli esiti dell'evento "il sistema funziona".

(c) Sia E l'evento "i componenti 1 e 3 sono guasti". Quanti esiti contiene?

$S = \{0,1\}^4 = \{(0,0,0,0), (0,0,0,1), (0,0,1,0), \cdots\}$ $\quad \# 2^4 = 16$

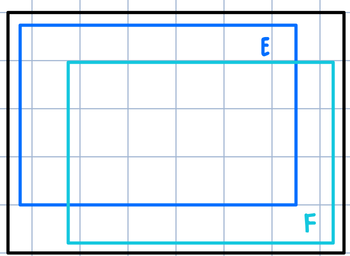


b) il sistema funziona $=: C$ $\quad \# C = 7$

c) $\# E = 4$

12. Dimostra che se $P(E) = 0.9$ e $P(F) = 0.9$, allora $P(E \cap F) \geq 0.8$. Poi dimostra che in generale vale la disuguaglianza seguente

$$P(E \cap F) \geq P(E) + P(F) - 1$$



$$P(E \cap F) \geq \overset{0.9}{P(E)} + \overset{0.9}{P(F)} - 1 = 0.8$$

$$P(E \cap F) \overset{D.M.}{=} P((E^c \cup F^c)^c) \overset{P4}{=} 1 - P(E^c \cup F^c) \overset{P7}{=} 1 - P(E^c) - P(F^c) + P(E^c \cap F^c)$$

$$\overset{P4}{=} \cancel{1} - \cancel{1} + P(E) - 1 + P(F) + \overset{\geq 0 \text{ per A1}}{P(E^c \cap F^c)} \geq P(E) + P(F) - 1$$

RICORDO LE PROPRIETÁ:

P4) probabilità del complementare

$P(A^c) = 1 - P(A)$

P5) $P(\emptyset) = 0$

P6) $A \subseteq B \Rightarrow P(A) \leq P(B)$

P7) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

ESERCIZIO:

Una moneta può essere vista come un d2 marcato con testa e croce.

Si tira 6 volte una moneta, contando quante volte esce testa, sia B questo valore.

Fare P(B= 3), verificare la somma.

$$\binom{6}{3} = \frac{n!}{r!\,(n-r)!} = \frac{6!}{3!\;3!} = \frac{6 \cdot 5 \cdot 4}{3 \cdot 2 \cdot 1} = 20$$

k	$\#\{B : k\}$
0	1
1	6
2	15
3	20
4	15
5	6
6	1
	64

ESERCIZIO: SUPERENALOTTO

Da un’urna contenente i primi 90 numeri interi positivi, se ne estraggono 6 a caso. Giocandone 6 a caso, qual è la probabilità di indovinarne 3? 4? 5? 6?

$S = \{ \{1,2,3,4,5,6\}, \{1,2,3,4,5,7\}, \cdots \}$ $\qquad \#S = \binom{90}{6} = \dfrac{90!}{6!\,84!} = \dfrac{90\cdot 89\cdot 88\cdot 87\cdot 86\cdot 85}{6\cdot 5\cdot 4\cdot 3\cdot 2\cdot 1} \approx 600$ mln

E: "ne indovino 3" $\qquad E = \{ V \cup U : V$ sono 3 indovinate e U sono 3 non indovinate $\}$

$\#E = \binom{6}{3}\cdot\binom{84}{3}$ $\qquad P(E) = \dfrac{\binom{6}{3}\binom{84}{3}}{\binom{90}{6}}$

ESERCIZIO:

In un comuni vi sono 5 alberghi. Se 3 persone devono scegliere un albergo in cui pernottare, qual è la probabilità che finiscano tutti in alberghi differenti? Che cosa stiamo assumendo senza dirlo esplicitamente?

$S = \{ (1,1,1), (1,1,2), \cdots, (4,3,5) \}$ $\qquad S = [5]^3 = 125$ $\qquad$ simmetria

I persona albergo 4 $\quad$ II persona albergo 3 $\quad$ III persona albergo 5

A: "tutte in alberghi differenti"

$\#A: 5\cdot 4\cdot 3 = 60$ $\qquad P(A) = \dfrac{60}{125} = \dfrac{12}{25} = 48\%$

Stiamo assumendo che:

- i 5 alberghi siano equivalenti (per tutte e tre le persone)
- Stiamo supponendo che le 3 persone stanno facendo scelte indipendenti

ESERCIZIO:

Una donna ha un mazzo di chiavi con $\overset{4}{n}$ chiavi, una delle quali apre la porta. Se le prova a caso scartando quelle che non aprono, qual è la probabilità che trovi la chiave giusta al $\overset{2}{k}$-esimo tentativo? E se non scartasse le chiavi già provate?

chiavi: $\{1, 2, 3, 4\}$ $\quad$ (la 1 è quella che apre)

$S = \{(1,2,3,4), (1,2,4,3), \cdots\}$ esiti simmetrici $\#S = 4 \cdot 3 \cdot 2 \cdot 1 = 4!$

A_2 : "chiave giusta al secondo tentativo" $= \{(2,1,3,4), (2,1,4,3), \cdots, (4,1,3,2)\}$

$\#A_2 = 3 \cdot 1 \cdot 2 \cdot 1 = 6$

$\#A_k = 1 \cdot 3 \cdot 2 \cdot 1 = 6 \qquad k = 1,2,3,4 \qquad P(A_k) = \frac{1}{4}$

$S = \{(1,2,\cdots,n), (1,2,\cdots,n,n-1), \cdots\}$ PERMUTAZIONI $\#S = n!$

$$\#A_k = (n-1)! \qquad P(A_k) = \frac{(n-1)!}{n!} = \frac{1}{n}$$

Secondo caso:

$S = \{(1), (2,1), (3,1), \cdots, (6,1), (2,2,1), (2,3,1), \cdots, \}$ non equiprobabili

= Lancio ripetutamente un dado a n facce finché non esce 1

$P(1 \text{ tentativo}) = \frac{1}{n}$

$$P(2 \text{ tentativi}) = \frac{n-1}{n} \cdot \frac{1}{n} = \frac{n-1}{n^2}$$

$$P(k \text{ tentativi}) = \frac{(n-1)^{k-1}}{n^k}$$

ESERCIZIO:

Vi è il 60% di probabilità che l'evento A si realizzi. Se ciò non accade, vi è un 10% di probabilità che si realizzi un altro evento B.

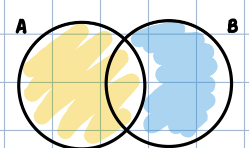
a) qual é la probabilità che si realizzi almeno uno tra A e B?

$S = ?$ non importa $\qquad A, B \in \mathcal{E}$ due eventi

hp:

$P(A) = 0.6 \qquad P(B|A^c) = 0.1 \xrightarrow{\text{recall}} P(B|A^c) = \frac{P(B \cap A^c)}{P(A^c)}$

a) $P(A \cup B) = ?$



$P(A^c) = 1 - P(A) = 0.4$

$P(B \cap A^c) = P(B|A^c)\, P(A) = 0.1 \cdot 0.4 = 0.04$

$P(A \cup B) = P(A) + P(B \cap A^c) = 0.64$

ESERCIZIO:

Hai chiesto ad un vicino di annaffiare una piantina delicata mentre sei in vacanza. Pensi che senza acqua la tua piantina non muoia con probabilità 0.8, mentre se innaffiata questa probabilità si ridurrebbe a 0.15. La tua fiducia che il vicino si ricordi di annaffiarla è del 90%.

A) qual è la probabilità che la pianta sia ancora viva al tuo ritorno? V:

B) se fosse morta, quale sarebbe la probabilità che il vicino si sia dimenticato di annaffiarla? D:

$P(V) = ?$

$P(D|V^c) = ?$

$0.8 = P(V^c|D) \qquad 0.2 = P(V|D)$

$0.15 = P(V^c|D^c) \xrightarrow{P4} 0.85 = P(V|D^c)$

$0.9 = P(D^c) \qquad 0.1 = P(D)$

A) $P(V)$ osserva che P dipende da D

FORMULA delle PROBABILITÀ TOTALI / FATTORIZZAZIONE — media pesata

$= P(V|D)\,P(D) + P(V|D^c)\,P(D^c) = \boxed{0.2 \cdot 0.1 + 0.85 \cdot 0.9} = 0.785$

B) $P(D|V^c)$ = FORMULA di BAYES

$$= P(V^c|D) \cdot \frac{P(D)}{P(V^c)} = 0.8 \; \frac{0.1}{1-(0.785)} \approx 0.372$$

Spesso il denominatore della formula di Bayes non viene dato e quasi sempre serve la formula di fattorizzazione per ricavarlo (al denominatore viene la seconda formula di Bayes)

FORMULE DI BAYES e TEOREMA DI BAYES

serve per calcolare la probabilità condizionata di un evento E rispetto a un altro evento F, a patto di conoscere, oppure di saper calcolare, le probabilità dei due eventi e la probabilità condizionata di F rispetto a E.

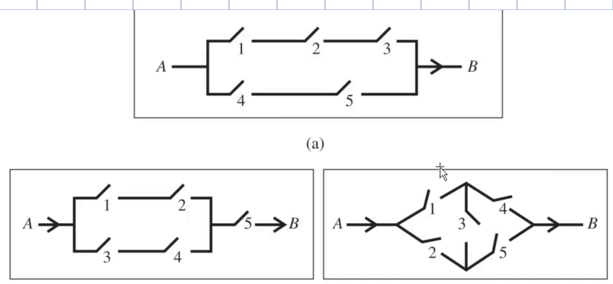
ENUNCIATO: siano E e F due venti con probabilità non nulle. La probabilità condizionata di E rispetto a F è uguale al prodotto tra la probabilità condizionata di F rispetto a E e la probabilità di E, tutto fratto la probabilità di F.

$$P(E|F) = \frac{P(F|E)\,P(E)}{P(F)} \qquad \text{con} \quad P(E),\ P(F) \neq 0$$

SECONDA FORMULA DI BAYES:

Numeratore uguale alla prima, al denominatore: numeratore + numeratore all'altro mondo

$$\frac{P(F|E)\,P(E)}{\boxed{P(F|E)\,P(E) + P(F|E^c)\,P(E^c)}} \rightarrow \text{formula di fattorizzazione}$$



(a)

(b)

(c)

C_1, C_2, C_3, C_4, C_5:

44. Le probabilità di chiusura dei cinque relè in ciascuna delle tre figure qui sopra sono p_1, p_2, p_3, p_4 e p_5. Tutti i relè sono indipendenti. Quali sono le probabilità che passi corrente tra gli estremi A e B dei tre circuiti? A:

a) $P(A) = ?$

$= P(\text{"ramo sopra"} \cup \text{"ramo sotto"}) = P((C_1 \cap C_2 \cap C_3) \cup (C_4 \cap C_5))$

p7 $\hookrightarrow = P(C_1 \cap C_2 \cap C_3) + P(C_4 \cap C_5) - P(C_1 \cap C_2 \cap C_3 \cap C_4 \cap C_5)$

indip. $\hookrightarrow = P(C_1)P(C_2)P(C_3) + P(C_4)P(C_5) - P(C_1)P(C_2)P(C_3)P(C_4)P(C_5)$

$= p_1 p_2 p_3 + p_4 p_5 - p_1 p_2 p_3 p_4 p_5$

b) $P(A) = P(C_5 \cap ((C_1 \cap C_2) \cup (C_3 \cap C_4))) =$

$= p_5 \cdot P(C_1 \cap C_2) \cup (C_3 \cap C_4)) =$

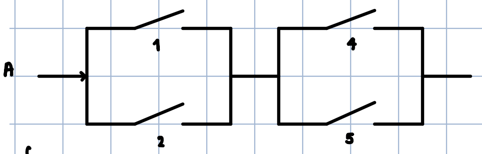
$= p_5 \cdot (p_1 p_2 + p_3 p_4 - p_1 p_2 p_3 p_4)$

fattorizz.

c) $P(A) = P(A|C_3)\,P(C_3) + P(A|C_3^c)\,P(C_3^c)$

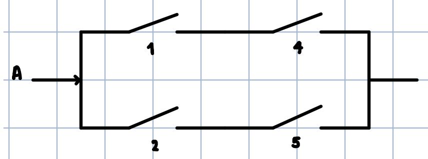
$= P(C_1 \cup C_2) \cap (C_4 \cap C_5))\, p_3 + P(C_1 \cap C_4) \cup P(C_2 \cap C_5) \cdot (1 - p_3)$

C_3



$= P((C_1 \cup C_2) \cap (C_4 \cap C_5))\, p_3$

C_3^c



$= P(C_1 \cap C_4) \cup P(C_2 \cap C_5)(1-p_3)$

ESERCIZI SU VARIABILI ALEATORIE

X + Y

		II dado					
		1	2	3	4	5	6
I dado	1	2	3	4	5	6	7
	2	3	4	5	6	7	8
	3	4	5	6	7	8	9
	4	5	6	7	8	9	10
	5	6	7	8	9	10	11
	6	7	8	9	10	11	12

X+Y

		II dado					
		1	2	3	4	5	6
I dado	1	1	1	1	1	1	1
	2	1	2	2	2	2	2
	3	1	2	3	3	3	3
	4	1	2	3	4	4	4
	5	1	2	3	4	5	5
	6	1	2	3	4	5	6

1) rappresentare V= min (X,Y)
2) P (V ⩾ 3)
3) P (X > Y)

il minimo tra x e y

2) $P(V \geq 3) = P(\{\omega \in S : V(\omega) \geq 3\}) = \dfrac{\#\{\dots\}}{\# S} = \dfrac{16}{36} = \dfrac{4}{9}$

3) $P(X > Y) = P(\{\omega \in S : X(\omega) > Y(\omega)\}) = \dfrac{\#\{\dots\}}{\# S} = \dfrac{15}{36} = \dfrac{5}{12}$

ESERCIZIO:

Si forma la classifica dei punteggi di un gruppo di 10 studenti- 5 studenti maschi e 5 femmine- dopo un esame. Non vi sono ex aequo, e tutte le 10! classifiche diverse hanno pari probabilità. Sia X la migliore posizione ottenuta da una studentessa (ad esempio X= 2 se il primo classificato è maschio e la seconda è femmina). Calcola, per i = 1, 2, …, 10, quanto vale P(X = i)

Chiede $P(x=1), P(x=2), \dots, P(x=10)$ ovvero la pmf

$\varphi_x(k) := P(x=k)$

$k=1$ $\quad P(x=1) = \dfrac{\#\{\text{permutazioni che iniziano con una } f\}}{\# S} = \dfrac{1}{2}$

$k=10$ $\quad P(x=10) = 0$ ci sono 5 femmine : alla peggio è in $x=6$

$k=9$ $\quad P(x=9) = 0$

$k=8$ $\quad P(x=8) = 0$

$k=7$ $\quad P(x=7) = 0$

$k=2$ $\quad P(x=2) = \dfrac{5\cdot 5\cdot \cancel{8\cdot 7\cdot 6\cdot 5\cdot 4\cdot 3\cdot 2\cdot 1}}{10\cdot 9\cdot \cancel{8\cdot 7\cdot 6\cdot 5\cdot 4\cdot 3\cdot 2\cdot 1}} = \dfrac{5}{18}$

$k=3$ $\quad P(x=3) = \dfrac{5\cdot 4\cdot 5\cdot \cancel{7\cdot 6\cdot 5\cdot 4\cdot 3\cdot 2\cdot 1}}{10\cdot 9\cdot 8\cdot \cancel{7\cdot 6\cdot 5\cdot 4\cdot 3\cdot 2\cdot 1}} = \dfrac{5}{36}$

$k=4$ $\quad P(x=4) = \dfrac{5\cdot 4\cdot 3\cdot 5\cdot \cancel{6\cdot 5\cdot 4\cdot 3\cdot 2\cdot 1}}{10\cdot 9\cdot 8\cdot 7\cdot \cancel{6\cdot 5\cdot 4\cdot 3\cdot 2\cdot 1}} = \dfrac{5}{84}$

$k=5$ $\quad P(x=5) = \dfrac{5\cdot 4\cdot 3\cdot 2\cdot 5\cdot \cancel{5\cdot 4\cdot 3\cdot 2\cdot 1}}{10\cdot 9\cdot 8\cdot 7\cdot 6\cdot \cancel{5\cdot 4\cdot 3\cdot 2\cdot 1}} = \dfrac{5}{252}$

$k=6$ $\quad P(x=6) = \dfrac{5\cdot 4\cdot 3\cdot 2\cdot 1\cdot \cancel{5\cdot 4\cdot 3\cdot 2\cdot 1}}{10\cdot 9\cdot 8\cdot 7\cdot 6\cdot \cancel{5\cdot 4\cdot 3\cdot 2\cdot 1}} = \dfrac{1}{252}$

ESERCIZIO

Valori possibili e pmf (con grafico) delle seguenti vv. aa. discrete.

A) somma di due dadi

B) numero di croci prima che esca testa una volta

C) numero di teste lanciando 4 volte una moneta

a) X, Y i valori di due lanci di dado

$X+Y$

I dado \ II dado	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

X+Y

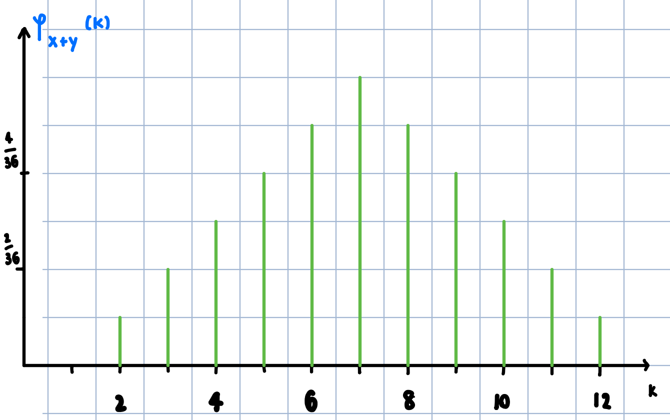
probabilità che la nostra variabile aleatoria assuma valore k

$$\varphi_{X+Y}(k) := P(X+Y=k)$$

$V_{X+Y} :=$

K	Quante volte compare?
2	1/36
3	2/36
4	3/36
5	4/36
6	5/36
7	6/36
8	5/36
9	4/36
10	3/36
11	2/36
12	1/36

$$=: \frac{6 - |7-k|}{36}$$



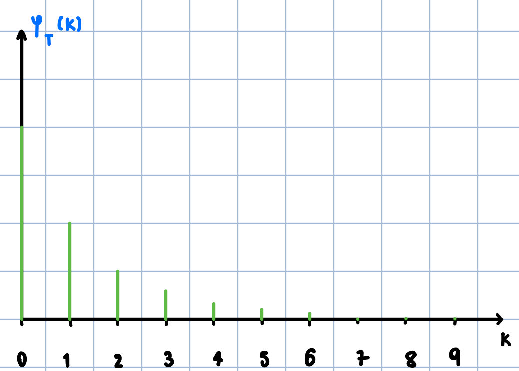
b) T: n° di croci prima che esca testa

$\varphi_T(k) = ?$ $\quad V_T = \{0, 1, \cdots, \} = \mathbb{N}$

$k=0$ $\quad P(T=0) = P(\text{primo lancio, testa}) = \frac{1}{2}$

$k=1$ $\quad P(T=1) = P(\text{croce, poi testa}) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$

k qualsiasi $\quad P(T=k) = P(k \text{ volte croce, una volta testa}) = \left(\frac{1}{2}\right)^k \cdot \left(\frac{1}{2}\right) = \left(\frac{1}{2}\right)^{k+1}$



C) X = n° di volte che esce testa lanciando 4 volte una moneta

$\varphi_X(k)$ $\quad V_X = \{0, 1, 2, 3, 4\}$

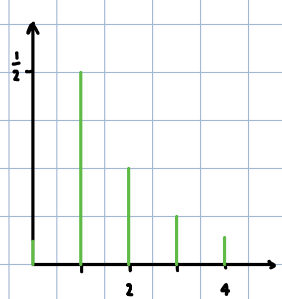
$k=1$ $\quad P(T=1) = \frac{1}{2}$

$k=2$ $\quad P(T=2) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$

$k=3$ $\quad P(T=3) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$

$k=4$ $\quad P(T=4) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{16}$

$k=0$ $\quad P(T=0) = \frac{1}{16}$



ESERCIZIO

Calcolare la probabilità di due eventi che riguardano la variabile aleatoria W.

W: numero di lanci di moneta per vedere la prima testa.

A) P (W > 5)

B) P (W sia pari)

Ricorda: $P(X \in A) = \sum_{k \in A \cap V_x} \varphi_x(k)$ la probabilità che a una v.a. succeda qualcosa è la somma della phi per i valori che fanno succedere quella cosa intersecati a, eventualmente, ai valori possibili

a) $P(W > 5) = \sum_{k=6}^{\infty} \varphi_W(k) = \sum_{k_6}^{\infty} \frac{1}{2^k}$

Recall: per la serie geometrica $\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\dots = 1 \Leftarrow \sum_{k=0}^{\infty} a^k = \frac{1}{1-a} \quad \forall a: |a|<1$

$= \sum_{k=0}^{\infty} \frac{1}{2^k} - \left(\underset{1}{\frac{1}{2^0}} + \underset{\frac{1}{2}}{\frac{1}{2^1}} + \underset{\frac{1}{4}}{\frac{1}{2^2}} + \underset{\frac{1}{8}}{\frac{1}{2^3}} + \underset{\frac{1}{16}}{\frac{1}{2^4}} + \underset{\frac{1}{32}}{\frac{1}{2^5}}\right)$ 1° metodo: da 0 a ∞ e tolgo solo i primi 5

$= 2 - 1 - \left(\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\frac{1}{16}+\frac{1}{32}\right) =$

$= 1 - \left(\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\frac{1}{16}+\frac{1}{32}\right) = 1 - P(W \le 5) = 1 - \sum_{k=1}^{5} \varphi_W(k)$ 2° metodo: passo al complementare [easier]

3° metodo: cambio di variabili

$= (k = i+6$ cambio variabili$)$ $\sum_{i=0}^{\infty} \frac{1}{2^{i+6}} = \frac{1}{2^6} \underbrace{\sum_{i=0}^{\infty} \frac{1}{2^i}}_{=2} = \frac{1}{2^5} = \frac{1}{32}$

B) $P(W \text{ sia pari}) = \sum_{\substack{k \text{ pari} \\ k=2}}^{\infty} \varphi_W(k) = \underset{\varphi_W(2)}{\frac{1}{4}} + \underset{\varphi_W(4)}{\frac{1}{16}} + \underset{\varphi_W(6)}{\frac{1}{64}} + \dots = \boxed{\sum_{k=1}^{\infty} \frac{1}{4^k}} = \overbrace{\frac{1}{1-\frac{1}{4}}}^{\sum_0^{\infty}} - \overbrace{1}^{\frac{1}{4^0}} = \frac{1}{3}$

serie geometrica che non parte da 0, devo aggiungerlo

CdF= funzione di ripartizione

DEFINIZIONE: data una variabile casuale (sia discreta che continua) X, la funzione che fa corrispondere ai valore di x, le probabilità cumulate P ($X \leqslant x$) viene detta funzione di ripartizione, ed è indicata con F_x ed è così definita:

$$F_x : \mathbb{R} \longrightarrow [0,1] \qquad\qquad F_X(x) := P(X \le x)$$

In altre parole: la probabilità che la variabile casuale X assuma uno specifico valore minore o uguale ad un dato valore x.

PROPRIETÁ:

P1) $0 \leqslant F_X(x) \leqslant 1$

P2) $\lim_{x \to -\infty} F_X(x) = 0$ $\boxed{\lim_{x \to +\infty} F_x(x) = 1 \quad \text{Proprietà di normalizzazione}}$

P3) F (x) è monotona non decrescente: $x_1 < x_2 \Rightarrow F_X(x_1) \le F_x(x_2)$

P4) nel caso di una variabile discreta, la funzione di ripartizione è continua a destra:

$\lim_{x \to x_0^+} F_x(x) = F_x(x_0)$

FACTS nel caso di variabili CONTINUE:

F1) $F_x(x) = \int_0^x f_x(s)\,ds$

F2) $f_x(x) = F'_x(x)$

F3) $P(a < X < b) = F_x(b) - F_x(a)$ TEOREMA FONDAMENTALE DEL CALCOLO

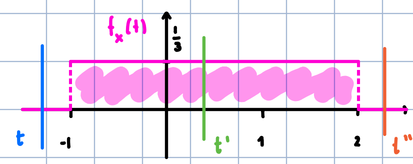
F3.1) $P(X < b) = F_x(b)$

F3.2) $P(X > a) = 1 - F_x(a)$

ESEMPIO GUIDA:

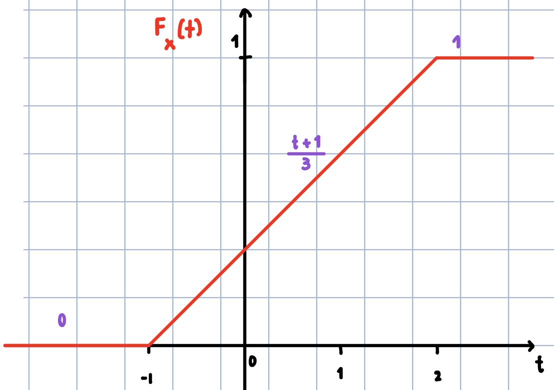
Sia X una v.a. Continua con pdf $f_X(t) = \begin{cases} \frac{1}{3} & -1 < t < 2 \\ 0 & \text{altrimenti} \end{cases}$. Calcolare la CdF e fare il grafico.

1) $F_X(t) = \int_{-\infty}^{t} f_X(u)\, du$ non usare la stessa lettera che fa da estremo d'integrazione



2) Distinguere i casi: un t che cade nella regione con densità $\neq 0$, il prima e il dopo → otterrò una CdF definita a tratti

$$F_X(t) = \int_{-\infty}^{t} f_X(u)\, du = \begin{cases} \int_{-\infty}^{0} 0\, du = 0 \quad \text{perché la densità è 0} & t < -1 \\ \int_{-\infty}^{-1} 0\, du + \int_{-1}^{t} \frac{1}{3}\, du = \frac{t+1}{3} & -1 \le t \le 2 \\ \int_{-\infty}^{-1} 0\, du + \int_{-1}^{2} \frac{1}{3}\, du + \int_{2}^{+\infty} 0\, du = 1 & t > 2 \end{cases}$$



F3.1) $P(X < 1) = F_X(1)$ guardo in che "settore" mi trovo, sostituisco

$= \frac{t+1}{3} = \frac{1+1}{3} = \frac{2}{3}$

ESEMPIO GUIDA 2:

Y una v.a. Continua con pdf parametrica

$f_Y(t) = c(2-t) \qquad 0 \le t \le 2$

$$F_Y(t) = \int_{-\infty}^{t} f_Y(s)\, ds = \begin{cases} 0 & t < 0 \\ \int_{0}^{t} c(2-s)\, ds & 0 \le t \le 2 \\ 1 & t > 2 \end{cases}$$

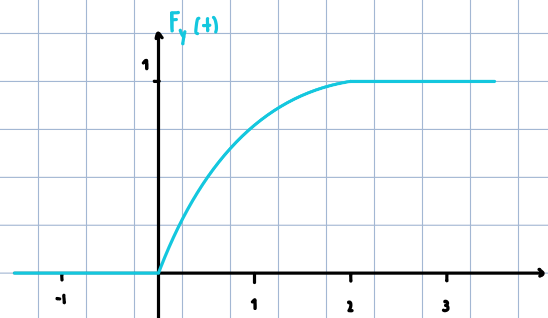
$$= c\int_0^t (2-s)\, ds = c\left[2s - \frac{s^2}{2}\right]_0^t = c\left(2t - \frac{t^2}{2} - 0 + 0\right) = c\left(2t - \frac{t^2}{2}\right)$$

funzione di ripartizione se io conoscessi c

Come trovo c? Voglio che $1 = \int_0^2 f_Y(s)\, ds$ (= che l'integrale della densità tra 0 e 2 valga 1) = $P(Y \le 2) = F_Y(2)$

$F_Y(2) = c\left(2 \cdot 2 - \frac{2^2}{2}\right) = c \cdot 2 = 1 \rightarrow c = \frac{1}{2}$ voglio che faccia 1

$$F_Y(t) = \begin{cases} t - \frac{t^2}{4} & 0 \le t \le 2 \\ 1 & t > 2 \\ 0 & \text{altrimenti} \end{cases}$$ una parabola



$P(Y \le a) = ? \;=\; F_Y(a) = a - \frac{a^2}{4}$

$P\left(\frac{1}{2} \le Y \le \frac{3}{2}\right) = ? \xrightarrow{F3} F_Y\left(\frac{3}{2}\right) - F_Y\left(\frac{1}{2}\right) = \frac{15}{16} - \frac{7}{16} = \frac{1}{2}$

ESERCIZIO:

Sia Z v.a. continua con pdf $f_Z(t) = t \cdot e^{-t}$, $t \ge 0$.

$$F_Z(t) = \int_{-\infty}^{t} f_Z(u)\,du = \int_0^t u \cdot e^{-u}\,du = \begin{cases} \dfrac{-t-1+e^t}{e^t} & t \ge 0 \\ 0 & \text{altrimenti} \end{cases}$$

ESERCIZIO:

Sia V v.a. continua con pdf $f_V(t) = -\log t$, $0 < t < 1$

$$F_V(t) = \int_{-\infty}^{t} f_V(u)\,du = \int_{-\infty}^{t} -\log u\,du = \begin{cases} 0 & t < 0 \\ t(1-\log t) & 0 < t \le 1 \\ 1 & t > 1 \end{cases}$$

ESERCIZIO:

Supponiamo che una variabile aleatoria continua X abbia funzione di densità di probabilità:

$$f(x) = \begin{cases} cx^3 & 0 \le x \le 1 \\ 0 & \text{altrimenti} \end{cases}$$

Determina:

A) il valore di c

B) la probabilità P(0.4 < X < 0.8)

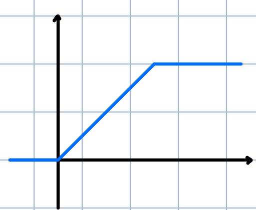
$$F_Y(x) = \int_{-\infty}^{t} f_Y(s)\,ds = \begin{cases} \int_0^t cs^3\,ds = c\,\dfrac{t^4}{4} & 0 \le t \le 1 \\ 0 & t < 0 \\ 1 & t > 1 \end{cases}$$

A) $1 = \int_0^1 f_Y(x)\,dx = \int_0^1 cx^3\,dx = c\left[\frac{x^4}{4}\right]_0^1 = c\cdot\left(\frac{1}{4} - 0\right) = \frac{c}{4}$ $\qquad \frac{c}{4} = 1 \rightarrow c = 4$

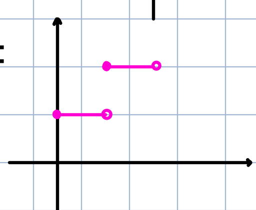
B) $P(a < Y < b) = F_Y(b) - F_Y(a) = F_Y(0.8) - F_Y(0.4) = 0.384$

AD OCCHIO:

Variabile aleatoria continua:

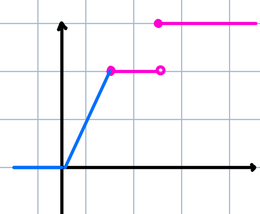


Variabile aleatoria discreta:



Variabile mista: discreta + continua

- non ha ne una densità
- non ha una pmf



FACTS nel caso di variabili DISCRETE:

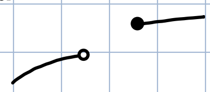
3) $P(a < X \le b) = F_X(b) - F_X(a)$

3.1) $P(X \le b) = F_X(b)$

3.2) $P(X > a) = 1 - F_X(a)$

Altre proprietà della CdF:

P5) CADLAG: continua a destra, limiti a sinistra:



$F(t) = \lim_{s \to t^+} F(s)$ continuo a dx

$\exists \lim_{s \to t^-} F(s) = F(t^-)$ limite a sx

P6) salti sono "atomi": $P(X = t) = F_X(t) - F_X(t^-)$

TRASFORMAZIONI DI VARIABILI ALEATORIE

ESERCIZIO: caso lineare

$f_W(t) = e^{-t}$, $t > 0$, consideriamo $\frac{1}{2}W$

FORMULA per il caso LINEARE:

$Y = a + bX$

$$f_Y(t) = \frac{1}{|b|} f_X\left(\frac{t-a}{b}\right)$$

$\frac{1}{2}W = 0 + \frac{1}{2}W$

densità di w

$$f_{\frac{1}{2}W}(t) = \frac{1}{\frac{1}{2}} \cdot f_W\left(\frac{t-0}{\frac{1}{2}}\right) = 2 \cdot e^{-2t} \qquad t > 0$$

ESERCIZIO: Caso lineare con coefficiente numerico

$f_S(t) = 2t$, $0 \le t \le 1$ $\qquad 3S - 2 = -2 + 3S$

$$f_{3S-2}(t) = \frac{1}{3} f_x\left(\frac{t+2}{3}\right)$$

$= \frac{1}{3} \cdot 2 \frac{t+2}{3}$, $\quad \boxed{0 \le \frac{t+2}{3} \le 1}$ da scrivere sempre

$= \frac{2t+4}{9}$, $\quad -2 \le t \le 1$ tolto il 3 e portato 2 a sx

ESERCIZIO: caso NON LINEARE

$V = e^{-W}$ devo trovare la densità di V

cambiando di segno, si inverte il simbolo

$F_V(t) = P(V \le t) = P(e^{-W} \le t) = P(-W \le \log t) = P(W \ge -\log t)$

applico il log per togliere e, essendo crescente, la disuguaglianza rimane

$= 1 - F_W(-\log t)$

$$f_V(t) = \frac{d}{dt}\left[1 - F_W(-\log t)\right] = 0 - \underbrace{f_W(-\log t) \cdot \left(-\frac{1}{t}\right)}_{\text{derivata di } F_W(-\log t)} = \frac{1}{t} f_W(-\log t)$$

$f_W(t) = e^{-t}$, $t > 0$ $\qquad = \frac{1}{t} \cdot \overbrace{e^{-(-\log t)}}^{= t}$, $\quad -\log t > 0$

$= 1$, $\quad 0 < t < 1$ $\qquad$ quando il log è > 0?

quando il log è negativo, quando $0 < t < 1$

ESERCIZIO: più difficile

$f_R(t) = 1 - |t|$ $\qquad -1 \le t \le 1$ $\qquad A = \sqrt{|R|}$

$$F_A(t) = P(A \le t) = P(\sqrt{|R|} \le t) = \begin{cases} P(|R| \le t^2) & t \ge 0 \\ 0 & t < 0 \end{cases} = P(-t^2 \le R \le t^2)$$

$\overset{3}{=} F_R(t^2) - F_2(-t^2)$, $\quad t > 0$

$$f_A(t) = \frac{d}{dt}\left[F_R(t^2) - F_R(-t^2)\right] = f_R(t^2) \cdot 2t - f_R(-t^2)(-2t) = 2t f_R(t^2) + 2t f_R(-t^2)$$

$= 2t\,\underline{(1 - |t^2|)} + 2t\,\underline{(1 - |-t^2|)} = 4t(1 - t^2)$ $\quad 0 \le t \le 1$ $\qquad$ sanity check: $\int_0^1 4t(1-t^2)\,dt = \int_0^1 (4t - 4t^3)\,dt = 4\frac{1}{2} - 4\frac{1}{4} = 1$ ✓

ESERCIZIO ESAME:

Sia X una variabile aleatoria continua con funzione di densità di probabilità: $f_X(t) = \begin{cases} c(8-t) & 0 \le t \le 8 \\ 0 & \text{altrimenti} \end{cases}$

7 PUNTI) si determini C, si tracci un grafico di fx, ~~si calcolino media, deviazione standard, moda e mediana di X.~~

3 PUNTI) sia Y =⌊X⌋= int (x) la parte intera di x, ovvero il più grande intero minore o uguale a X, determinare la legge di Y ~~e calcolare la sua media~~

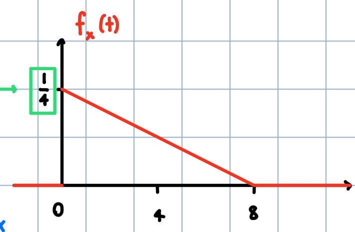
2 PUNTI) SIA Z = {X} = X- Y si determini la parte frazionaria di X, se ne determini la legge ~~e se ne calcoli il valore atteso~~

a) $1 \overset{?}{=} \int_{\mathbb{R}} f_x(t)\,dt = \int_0^8 c(8-t)\,dt = c\left[8t - \frac{t^2}{2}\right]_0^8 = c\,[64-32]$ $\qquad c = \frac{1}{32}$

v.a. continua

$V_x = [0;8]$ valori possibili di x

trovato usando $c = \frac{1}{32}$, $t = 0$ e sostituiti in f_x $\rightarrow$ $\boxed{\frac{1}{4}}$



b) $Y = \lfloor X \rfloor$ $\qquad V_Y = \{0,1,2,3,4,5,6,7\}$ v.a. discreta

Calcolo la funzione di massa:

$$\varphi_Y(k) = P(Y=k) = P(\lfloor X \rfloor = k) = P(k \le X < k+1)$$
$$= F_X(k+1) - F_X(k)$$

$$F_X(t) = \int_{-\infty}^{t} f_X(s)\,ds = \begin{cases} 0 & t < 0 \\ \int_0^t c(8-s)\,ds & 0 \le t \le 8 \\ 1 & t > 8 \end{cases} \qquad = c\left[8t - \frac{t^2}{2}\right] = \frac{1}{4}t - \frac{t^2}{64}$$

$$\varphi_Y(k) = \frac{1}{4}(k+1) - \frac{(k+1)^2}{64} - \frac{1}{4}k - \frac{k^2}{64} = \frac{1}{4} - \frac{2k+1}{64} = \frac{15-2k}{64} \qquad k = 0,1,2,\dots,7$$

c) $Z = \{X\} = X - \lfloor X \rfloor = X - Y$ $\qquad V_Z = [0,1)$ v.a. continua

$$F_Z(t) = P(Z \le t) = P(0 \le X \le t \text{ o } 1 \le X \le t+1 \text{ o } \dots \text{ } 7 \le X \le t+7) \qquad 0 \le t < 1$$

$\{\dots\} \cup \{\dots\} \cup \dots \cup \{\dots\}$ sono eventi disgiunti

$$\overset{A3}{=} P(0 \le X < t) + P(1 \le X \le t+1) + \dots + P(7 \le X \le t+7) = \sum_{k=0}^{7} P(k \le X \le k+t)$$

$$= \sum_{k=0}^{7} \left[F_X(k+t) - F_X(k)\right] = \sum_{k=0}^{7} \left[\frac{1}{4}k+t - \frac{(k+t)^2}{64} - \frac{1}{4}k + \frac{k^2}{64}\right]$$

ESERCIZI SU VETTORI ALEATORI:

Due o più variabili aleatorie (X e Y) che hanno una densità congiunta.

ESERCIZIO:

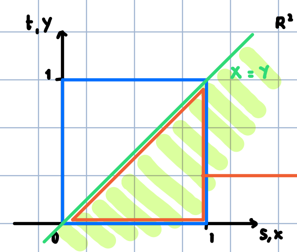
$f_{X,Y}(s,t) = s+t \qquad 0 \le s \le 1,\ 0 \le t \le 1$

a) $P(X > \frac{1}{2})$

b) $P(X > Y)$

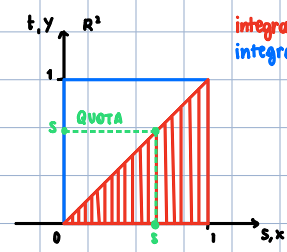
c) $P(X + Y < 1)$

b) $P(X > Y) = \overbrace{P((X,Y) \in R) = \iint_R f_{X,Y}(s,t)\,ds\,dt}^{\text{generico}}$



X > Y (R: semi piano)

integrale qua: T



integrale interno = verticale

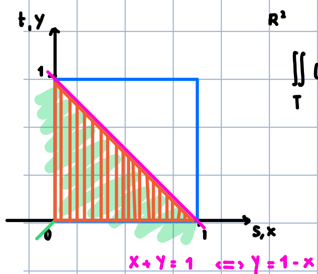
integrale esterno = orizzontale

$$\iint_T (s+t)\,ds\,dt = \int_0^1\int_0^s (s+t)\,dt\,ds$$
$$=\int_0^1 \left(s\cdot s + \frac{s^2}{2}\right) ds$$

l'integrale esterno ha sempre gli estremi fissati

$$= \frac{3}{2}\int_0^1 s^2\,ds = \frac{3}{2}\cdot\left[\frac{s^3}{3}\right]_0^1 = \frac{3}{2}\cdot\frac{1}{3} = \frac{1}{2}$$

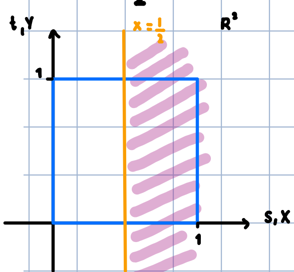
c) $P(X+Y<1) = P((X,Y)\in R) =$



costante · intervallo

$$\iint_T (s+t)\,ds\,dt = \int_0^1\int_0^{1-s}(s+t)\,dt\,ds = \int_0^1 s\cdot(1-s) + \left[\frac{t^2}{2}\right]_0^{1-s} ds$$
$$=\int_0^1 s\cdot(1-s) + \frac{1+s^2-2s}{2}\,ds$$
$$=\int_0^1 s-s^2 + \frac{1+s^2-2s}{2}\,ds$$
$$=\int_0^1 \frac{2s-2s^2+1+s^2-2s}{2}\,ds = \int_0^1 \frac{-s^2+1}{2}\,ds = \frac{1}{2}\int_0^1 -s^2+1\,ds = \frac{1}{2}\left(-\left[\frac{s^3}{3}\right]_0^1+1\right) =$$
$$=\frac{1}{2}\cdot\frac{2}{3} = \frac{1}{3}$$

a) $P(X>\frac{1}{2})$



modo 1) $\int_{\frac12}^1\int_0^1 (s+t)\,dt\,ds \dots$ costante· intervallo

Recall
$$f_X(s) = \int_R f_{X,Y}(s,t)\,dt = \int_0^1 (s+t)\,dt = (s\cdot 1) + \left[\frac{t^2}{2}\right]_0^1 = s+\frac{1}{2} \qquad 0\le s\le 1$$

modo 2) $\int_{\frac12}^1 f_X(s)\,ds = \int_{\frac12}^1 \left(s+\frac{1}{2}\right) ds = \left[\frac{s^2}{2}\right]_{\frac12}^1 + \frac{1}{2}\cdot\frac{1}{2} = \frac{5}{8}$

Ma c'è un altro modo: **[modo 2]**

Se ho eventi che riguardano sia X che Y, mi serve per forza la legge congiunta (la densità congiunta).

Se però ho un evento che riguarda una sola variabile, basterebbe avere la sua densità.

E come la ricavo?

LEGGI MARGINALI:

Ricavare f_X e f_Y (leggi marginali di X, Y) da $f_{X,Y}$ (legge congiunta di X, Y).

Mi basta integrare sull'altra variabile:

$$\boxed{f_X(s) = \int_R f_{X,Y}(s,t)\,dt \quad \forall s\in\mathbb{R} \qquad\qquad f_Y(t) = \int_R f_{X,Y}(s,t)\,ds \quad \forall t\in\mathbb{R}}$$

ESERCIZIO: densità congiunta + indipendenza

La densità di X e Y è data da

$$f_{X,Y}(u,v) = \begin{cases} ue^{-(u+v)} & u>0 \quad v>0 \\ 0 & \text{altrimenti}\end{cases}$$

A) calcola la densità di Y

B) calcola la densità di X

C) le due variabili aleatorie sono indipendenti?

B) $$f_Y(v) = \int_R f_{X,Y}(u,v)\,dv = \begin{cases} \int_0^\infty ue^{-(u+v)} = \int_0^\infty u\cdot e^{-u-v}\,dv = u\cdot e^{-u}\int_0^\infty e^{-v}\,dv = u\cdot e^{-u}\cdot[-e^{-v}]_0^\infty = u\cdot e^{-u}\cdot 1 & u>0 \\ 0 & u\le 0\end{cases}$$

A) $f_X(u) = \int_{\mathbb{R}} f_{X,Y}(u,v)\, du = \begin{cases} \int_0^\infty u \cdot e^{-u-v}\, du & \\ 0 & u \le 0 \end{cases} = e^{-v} \int_0^\infty u \cdot e^{-u} = e^{-v} \cdot 1 = e^{-v}$ $\quad v > 0$

c)

DEFINIZIONE:

Due variabili aleatorie sono indipendenti se la densità congiunta è il prodotto delle marginali:

$$f_{X,Y}(s,t) = f_X(s) \cdot f_Y(t)$$

in questo caso: $f_{X,Y}(u,v) = f_X(u) \cdot f_Y(v)$ ✓ $\quad u > 0,\ v > 0$

$u \cdot e^{-(u+v)} = u \cdot e^{-u} \cdot e^{-v}$

LEGGE DEL MINIMO E DEL MASSIMO PER VVAA INDIPENDENTI

Funzione con n > 2 qualsiasi

$X_1, X_2, \dots, X_n$ vvaa indipendenti

R: $\min_i X_i$

T: $\max_i X_i$

MINIMO

$$1 - F_R(t) = \prod_i (1 - F_{X_i}(t))$$

MASSIMO

$$F_T(t) = \prod_i F_{X_i}(t)$$

ESERCIZIO:

X, Y vvaa continue indipendenti

$f_X(t) = 2e^{-2t}, \quad t > 0$ $\qquad f_Y(t) = 3e^{-3t}, \quad t > 0$

Determinare la pdf di S= min (X, Y)

$1 - F_S(t) = (1 - F_X(t)) \cdot (1 - F_Y(t))$

da ricavare

$F_X(t) = \int_{-\infty}^t f_X(u)\, du = \begin{cases} \int_0^t 2e^{-2u}\, du & \\ 0 & t \le 0 \end{cases} = 2 \cdot \left[-\frac{1}{2} \cdot e^{-2u}\right]_0^t = -e^{-2t} + 1 \quad t > 0$

$F_Y(t) = \int_{-\infty}^{0} f_Y(v)\, dv = \begin{cases} \int_0^t 3e^{-3v}\, dv & \\ 0 & t \le 0 \end{cases} = 3\left[-\frac{1}{3} \cdot e^{-3u}\right]_0^t = -e^{-3t} + 1 \quad t > 0$

$1 - F_S(t) = [1 - \overbrace{1 + e^{-2t}}^{0\ t \le 0}] \cdot [1 - \overbrace{1 + e^{-3t}}^{0\ t \le 0}]$

$= \begin{cases} e^{-2t} \cdot e^{-3t} & t > 0 \\ 1 & t \le 0 \end{cases}$

min

$F_S(t) = \begin{cases} 1 - e^{-5t} & t > 0 \\ 1 - 1 & t \le 0 \end{cases}$

pdf del min

$f_S(t) = F'_S(t) = 0 - (-5e^{-5t}) = 5e^{-5t}, \ t > 0$

ESERCIZIO: nel caso DISCRETO + distribuzioni condizionali (non congiunte)

Tiro una moneta finche non esce testa: sia X il numero di lanci fatti; costruisco un dado con X facce e lo lancio: sia Y il risultato.

Determinare $\varphi_X, \varphi_{Y|X}, \varphi_{X,Y}, \varphi_Y, \varphi_{X|Y}$ (in questo ordine)

a) $\varphi_x(k) := P(X=k)$ pmf: funzione di massa di probabilità

$V_x = \{1, 2, \cdots, n\}$ valori possibili

$P(X=1) = \frac{1}{2}$

$P(X=2) = \frac{1}{2} \cdot \frac{1}{2} \longrightarrow \varphi_x(k) = P(X=k) = \frac{1}{2^k} \quad k \in V_x$

$P(X=3) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2}$

caso discreto

b) $\varphi_{Y|X}(j|k)$ legge di Y condizionata a X $:= P(Y=j \mid X=k)$ pmf condizionata

$= \begin{cases} \frac{1}{k} & j \in \{1, 2, \cdots, k\}, \ k \in V_x \\ 0 & \text{altrimenti} \end{cases} \quad = \frac{1}{k} \quad 1 \le j \le k$ interi

c) $\varphi_{X,Y}$ legge congiunta

RICORDO LA FORMULA: $\varphi_{Y|X}(j|k) = \dfrac{\varphi_{Y,X}(j,k)}{\varphi_X(k)}$

INVERSA: $\varphi_{Y,X}(j,k) = \varphi_{Y|X}(j|k) \cdot \varphi_X(k)$

$= \frac{1}{k} \cdot \frac{1}{2^k} \quad 1 \le j \le k \in \mathbb{N}^+$

!! non è finita, non è ciò che l'es mi chiede

$\varphi_{X,Y}(k,j) = \varphi_{Y,X}(j,k) = \frac{1}{k 2^k} \quad 1 \le j \le k$

$\varphi_{X,Y}(j,k) = \frac{1}{j 2^j}$ ho cambiato variabile $1 \le k \le j$

d) $\varphi_Y(k)$ legge marginale

Parto dalla legge congiunta e ricavo la marginale

essendo nel caso discreto: Σ e non $\int$

$= \sum_j \varphi_{X,Y}(j,k) \quad = \sum_{j=k}^{\infty} \frac{1}{j \cdot 2^j}$

faccio sparire l'altro argomento

VALORE ATTESO E VARIANZA

Il valore atteso, che viene chiamato anche media, della distribuzione di una variabile casuale. È un indice di posizione. Il valore atteso di una variabile casuale rappresenta il valore previsto che si potrà ottenere in un gran numero di prove.

DEFINIZIONE: il valore atteso di una variabile casuale discreta, se la distribuzione è finita, è un numero reale dato dalla somma dei o prodotti di ogni valore della variabile casuale per la rispettiva probabilità, cioè:

$$E(x) = x_1 p_1 + x_2 p_2 + \cdots + x_n p_n = \sum_{i=1}^{n} x_i p_i$$

Il valore atteso è una somma pesata dei valori che la variabile casuale assume con pesi le probabilità associate. Può essere sia negativo che positivo.

- se la variabile casuale discreta, assume un'infinità numerabile di valori:

$$E(x) = \sum_{i=1}^{\infty} x_i P(X = x_i)$$

Con la condizione che, nel caso in cui i valori x , x , …, x non siano tutti dello stesso segno (tranne al più un numero finito di essi), la serie $\sum_{i=1}^{\infty} x_i P(X = x_i)$ sia assolutamente convergente, altrimenti si dice che X non ammette valore atteso.

- se la variabile casuale è continua, abbiamo:

$$E(x) = \int_{-\infty}^{\infty} x\, f(x)\, dx$$

Con la condizione che, nel caso in cui i valori x_1, x_2, …, x_n non siano tutti dello stesso segno (tranne al più un numero finito di essi), è necessario che: $\int_{-\infty}^{\infty} |x|\, f(x)\, dx < +\infty$

Altrimenti si dice che X non ammette valore atteso. Dove f è la funzione di densità della variabile casuale continua.

PROPRIETÀ:

1) se a è una costante reale, allora:

$$E(aX) = a \cdot E(x)$$

2) se a è una costante reale, allora:

$E(X+a) = E(x) + a$ in particolare $E(a) = a$

3) se X e Y sono due variabili casuali, si ha una nuova variabile Z= X +Y i cui valori sono dati dalla somma di tutti i possibili valori di x con tutti i possibili valori di y di Y, con probabilità $P(X = x_i, Y = y_k) = p_{ik}$

$$E(z) = E(X+Y) = E(X) + E(Y)$$

Questa proprietà si può generalizzare alla somma di n variabili.

4) se A e B sono costanti reali, allora:

$$E(aX + bY) = a\,E(X) + b\,E(Y)$$

Questa proprietà si può estendere anche al caso di combinazioni lineari di più di due variabili casuali.

5) se a è una costante reale, allora:

$$E((aX)^2) = a^2\, E(X^2)$$

ESEMPIO: anche di VARIANZA + DEVIAZIONE STANDARD LINEARI

a) $f_W(t) = e^{-t}$, $t > 0$

$$E(W) = \int_R t \cdot f_W(t)\, dt = \int_0^{\infty} t \cdot e^{-t}\, dt = \left[-t \cdot e^{-t}\right]_0^{\infty} + \int_0^{\infty} e^{-t}\, dt = 0 - 0 + \left[-e^{t}\right]_0^{\infty} = 1$$

La varianza di una variabile aleatoria ha due formule che possiamo usare:

1) $E[(x - E(x))^2]$ scomoda

2) $E(x^2) - E(x)^2$

Uso la 2)

$$E(W^2) = \int_R t^2 \cdot f_W(t)\, dt = \int_0^{\infty} t^2 \cdot e^{-t}\, dt = \underbrace{\left[-t^2 e^{-t}\right]_0^{\infty}}_{= 0 - 0} + 2\underbrace{\int_0^{\infty} t \cdot e^{-t}\, dt}_{\text{fatto prima } = 1} = 2$$

$$Var(W) = E(W^2) - E(W)^2 = 2 - 1 = 1$$

DEVIAZIONE STANDARD: radice quadrata della varianza

$$\sqrt{Var(W)} = 1$$

ESEMPIO: VARIANZA + DEV STANDARD NON LINEARI

b) $W'' = e^{-W}$ $\qquad R'' = \sqrt{|R|}$

$E(w') = E(e^{-w})$

TEOREMA:

Se devo fare il valore atteso di una cosa complicata, di una funzione di una variabile aleatoria , posso integrare quella funzione li contro la densità

$$= \int_R e^{-t} \cdot f_w(t)\, dt = \int_0^\infty e^{-t} \cdot e^{-t}\, dt = \int_0^\infty e^{-2t}\, dt = \left[-\frac{1}{2} e^{-2t}\right]_0^\infty = \frac{1}{2}$$

$$E(w''^2) = E((e^{-w})^2) = E(e^{-2w}) = \int_0^\infty e^{-2t} \cdot e^{-t}\, dt = \int_0^\infty e^{-3t}\, dt = \left[-\frac{1}{3} e^{-3t}\right]_0^\infty = \frac{1}{3}$$

$$Var(w') = \frac{1}{3} - \frac{1}{4} = \frac{1}{12} \qquad \sqrt{Var(w')} = \frac{1}{2\sqrt{3}}$$

PROBLEMI STRANI DAL CAP. 4 DEL ROSS:

La funzione di densità X è data da

$$f(x) = \begin{cases} a + bx^2 & \text{se } 0 \le x \le 1 \\ 0 & \text{altrimenti} \end{cases}$$

Determina il valore di a e b, sapendo che $E[X] = \frac{3}{5}$

$$E(x) = \int_R t \cdot f_x(t)\, dt = \int_0^1 t \cdot (a + bt^2)\, dt = \left[a\frac{t^2}{2} + b\frac{t^4}{4}\right]_0^1 = \frac{a}{2} + \frac{b}{4} - 0 = \frac{3}{5}$$

$$1 = \int_R f_x(t)\, dt = \int_0^1 a + bt^2\, dt = a + \frac{b}{3}$$

$$\begin{array}{l} 3 \\ 4 \end{array}\begin{cases} a + \frac{b}{3} = 1 \\ \frac{a}{2} + \frac{b}{4} = \frac{3}{5} \end{cases} \qquad \begin{cases} 3a + b = 3 \\ 2a + b = \frac{12}{5} \end{cases} \qquad \begin{cases} b = -3a + 3 \\ 2a - 3a + 3 = \frac{12}{5} \end{cases} \qquad \begin{cases} \text{———} \\ a = \frac{3}{5} \end{cases} \qquad \begin{cases} b = \frac{6}{5} \\ a = \frac{3}{5} \end{cases}$$

ESERCIZIO:

Siano $X_1, X_2, \dots, X_n$ variabili aleatorie indipendenti, tutte con densità

$$f(x) = \begin{cases} 1 & \text{se } 0 < x < 1 \\ 0 & \text{altrimenti} \end{cases}$$

Calcola E [max($X_1, \dots, X_n$) e E [min ($X_1, \dots, X_n$)]

$$\underbrace{\max(X_1, \dots, X_n)}_{S} \qquad \underbrace{\min(X_1, \dots, X_n)}_{T}$$

$$\boxed{\begin{array}{l} F_S(t) = \prod_{i=1}^n F_{x_i}(t) \\ 1 - F_T(t) = \prod_{i=1}^n (1 - F_{x_i}(t)) \end{array}}$$

nel caso di v.a. continue

$$F_{x_i}(t) := P(x_i \le t) = \int_{-\infty}^t f_{x_i}(s)\, ds = \begin{cases} \int_0^t f_{x_i}(s)\, ds = \int_0^t 1\, ds = t & 0 < t < 1 \\ 0 & t \le 0 \\ 1 & t \ge 1 \end{cases}$$

$$\max: F_S(t) = \prod_1^n \begin{cases} t \\ 0 \\ 1 \end{cases} = \begin{cases} t^n & 0 < t < 1 \\ 0^n & 0 \le t \\ 1^n & t \le 1 \end{cases}$$

$$E(s) = \int_R t \cdot f_s(t)\, dt \qquad f_s = F'_s$$

OPPURE

Formula alternativa per la media di vvaa positive

$$E(x) = \int_0^\infty [1 - F_x(t)]\, dt$$

$$E(S) = \int_0^1 [1 - t^n]\, dt = 1 - \left[\frac{t^{n+1}}{n+1}\right]_0^1 = \frac{n}{n+1}$$

min: $$E(T) = \int_R (1 - F_T(t))\, dt = \int_R (1 - F_{x_i}(t))^n\, dt = \int_0^1 (1-t)^n\, dt = \left[-\frac{(1-t)^{n+1}}{n+1}\right]_0^1 = \frac{1}{n+1}$$

ESERCIZIO:

Sapendo che E[X] = 2 e E[X^2]= 8, calcola

a) $E[(2+4x)^2]$

b) $E[x^2 + (x+1)^2]$

a) $$E[(2+4x)^2] = E[(4 + 16x^2 + 16x)] \overset{\text{linearità}}{=} 4 + 16\underbrace{E(x)}_{=2} + 16\underbrace{E(x^2)}_{=8} = 164$$

b) $$E[x^2 + (x+1)^2] = E[x^2 + x^2 + 2x + 1] = E[2x^2 + 2x + 1] = 2E(x^2) + 2E(x) + 1 = 16 + 4 + 1 = 21$$

PROBLEMA DA ESAME:

Sia X una variabile aleatoria continua con funzione di densità di probabilità

$$f_x(t) = \begin{cases} c(e^{-2t} - e^{-3t}) & t \geq 0 \\ 0 & \text{altrimenti} \end{cases}$$

7 punti) si determini il valore di c, si tracci un grafico di f , si calcolino media, moda e deviazione standard di X

2 punti) Se determini la funzione di ripartizione di X e si trovi dove cade la mediana rispetto alla moda e alla media. (Non è richiesto il valore esatto, ma solo di determinare l'rdine relativo di queste tre grandezze)

3 punti) Determinare se possibile una distribuzione F tale che X~max(U, V), dove U~V, V~ Expo(1), U e V sono indipendenti e il simbolo vuol dire "distribuito come"

MODA: valore di ascissa in cui c'è il massimo := si fa la derivata, si vede dove si annulla

MEDIANA: linea che taglia l'integrale metà a destra e metà a sinistra

$$1 = \int_R f_x(t)\, dt = c\int_0^\infty (e^{-2t} - e^{-3t})\, dt$$

$$= c\left[-\frac{1}{2}\cdot e^{-2t} - \left(\frac{1}{3}\cdot e^{-3t}\right)\right]_0^\infty = c\left[\frac{1}{2} - \frac{1}{3}\right] = c\,\frac{1}{6} \iff c = 6$$

$$E(x) = \int_0^\infty c\,t\,(e^{-2t} - e^{-3t})\, dt$$

$$= \int_0^\infty 6te^{-2t}\,dt + \int_0^\infty 6te^{-3t}\,dt = \left[-6t\cdot\frac{1}{2}e^{-2t}\right]_0^\infty + \left[-6t\cdot\frac{1}{3}e^{-3t}\right] + \int_0^\infty 6\cdot\frac{1}{2}e^{-2t}\,dt + \int_0^\infty 6\cdot\frac{1}{3}e^{-3t}\,dt$$

$$= 3\int_0^\infty e^{-2t}\,dt + 2\int_0^\infty e^{-3t}\,dt = 3\cdot\left[-\frac{1}{2}e^{-2t}\right]_0^\infty + 2\cdot\left[-\frac{1}{3}e^{-3t}\right]_0^\infty = \frac{3}{2} - \frac{2}{3} = \frac{5}{6}$$

moda: ascissa con f_x massima

$$0 \overset{!}{=} f_x'(t) = c(-2e^{-2t} + 3e^{-3t}) = \overset{6}{c}\cdot \overset{>0}{e^{-2t}}(3e^{-t} - 1) \iff 3e^{-t} - 1 = 0 \iff e^{-t} = \frac{1}{3} \iff e^{t} = 3 \iff t = \log 3$$

Se non dovessi avere la legge cosa devo usare?

DISUGUAGLIANZA DI MARKOV

Funziona su variabili aleatorie non negative

$$P(X > t) \leq \frac{1}{t}\cdot E(x) \quad \forall t > 0$$

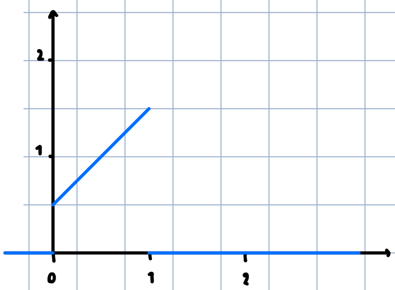
OPPURE

DISUGUAGLIANZA DI CHEBYSHEV

$$P(|x - E(x)| \geq t) \leq \frac{1}{t^2} \; Var(x) \qquad \forall t > 0$$

ESERCIZIO: media, moda, mediana + grafico

$f_x(t) = \frac{1}{2} + t \; , \; 0 < t < 1$



moda: $E(x) = \int_R t \cdot f_x(t)\, dt = \int_0^1 t \cdot (\frac{1}{2} + t)\, dt = \left[\frac{t^2}{4} + \frac{t^3}{3}\right]_0^1 = \frac{1}{4} + \frac{1}{3} = \frac{7}{12}$

mediana: Calcolo la $F_x(t)$ e trovo dove vale $\frac{1}{2}$

$F_x(t) := P(x \leq t) = \int_{-\infty}^{t} f_x(s)\, ds = \int_0^t \frac{1}{2} + s \; ds = \frac{t}{2} + \frac{t^2}{2}$

$\frac{1}{2} = F_x(M_x)$

$\frac{1}{2} = \frac{M_x}{2} + \frac{M_x^2}{2} \longrightarrow M_x^2 + M_x - 1 = 0 \qquad x_{1,2} = \frac{-1 \pm \sqrt{5}}{2} \longrightarrow \frac{\sqrt{5}-1}{2} \approx 0.618$

uso solo quella nell'intervallo di $f_x(t)$

la moda è 1 (dal grafico)

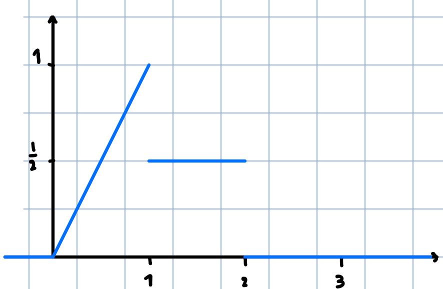
ESERCIZIO: esame (tipologia 2)

Sia X una va continua con funzione di densità di probabilità:

$$f_x(t) = \begin{cases} t & 0 \leq t \leq 1 \\ \frac{1}{2} & 1 < t \leq 2 \\ 0 & \text{altrimenti} \end{cases}$$

7 punti) si verifichi che fx sia una funzione di densità valida, se ne tracci un grafico, si calcolino media, moda, mediana di X. Si calcoli P (X > $\frac{1}{2}$).

2 punti) si determini la dev standard di X e si calcoli P ($|X - 1| > \frac{1}{2}$)



moda: 1, dal grafico

verifico densità:

1) $f_x \geq 0$ ✓ [a occhio]

2) $\int_R f_x(t)\, dt = 1$ ✓ [a occhio]

mediana: $M_x = 1$, perche $P(x \leq 1) = \frac{1}{2}$ dal grafico

$E(x) = \int_R t\, f_x(t)\, dt = \int_0^1 t \cdot f_x(t)\, dt + \int_1^2 t \cdot f_x(t)\, dt = \left[\frac{t^3}{3}\right]_0^1 + \frac{1}{2}\left[\frac{t^2}{2}\right]_1^2 = \frac{1}{3} + 1 - \frac{1}{4} = \frac{13}{12}$

$P(x > \frac{1}{2}) = 1 - P(x \leq \frac{1}{2}) = 1 - \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{7}{8}$

$\sqrt{Var(x)} = ? \qquad Var(x) = E(x^2) - E(x)^2$ ✓

$E(x^2) = \int_0^2 t^2 \cdot f_x(t) = \int_0^1 t^2 \cdot f_x(t)\, dt + \int_1^2 t^2 \cdot f_x(t)\, dt = \int_0^1 t^3\, dt + \int_1^2 \frac{t^2}{2}\, dt = \left[\frac{t^4}{4}\right]_0^1 + \frac{1}{2}\left[\frac{t^3}{3}\right]_1^2 = \frac{1}{4} + \frac{8}{6} - \frac{1}{6} = \frac{3+14}{12} = \frac{17}{12}$

$Var(x) = \frac{17}{12} + \left(\frac{13}{12}\right)^2 = \frac{35}{144} \qquad \sqrt{Var(x)} = \frac{\sqrt{35}}{12} \approx 0.5$

$|x - 1| > \frac{1}{2} \longrightarrow \begin{cases} x - 1 > 0 & x - 1 > \frac{1}{2} & x > \frac{3}{2} \\ x - 1 < 0 & 1 - x < \frac{1}{2} & x < \frac{1}{2} \end{cases}$

$P(|x-1| > \frac{1}{2}) = P(x < \frac{1}{2}) + P(x > \frac{3}{2}) = \frac{1}{8} + \frac{1}{4} = \frac{3}{8}$ ✓

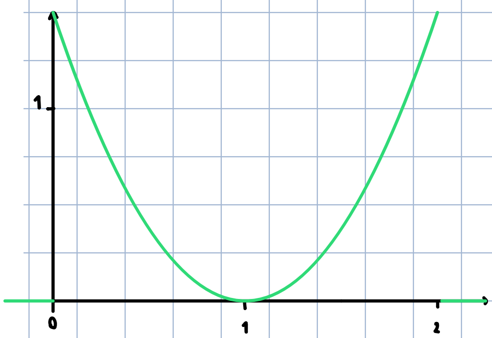
ESERCIZIO: esame (tipologia 2)

Sia X una variabile aleatoria continua con funzione di densità di probabilità

$$f_x(t) = \begin{cases} c(t-1)^2 & 0 \leq t \leq 2 \\ 0 & \text{altrimenti} \end{cases}$$

7 punti) si determini il valore di c, si tracci un grafico, si calcolino moda, media e dev standard
2 punti) si determini la funzione di ripartizione di X e si trovi la mediana
2 punti) si trovino due reali a e b tali che $(X - a)^b$ abbia distribuzione uniforme con estremi qualsiasi

$$1=\int_0^2 f_x(t)\,dt = \int_0^2 c(t-1)^2 = c\int_0^2 t^2-2t+1\,dt = c\left[\frac{t^3}{3}-\cancel{2}\frac{t^2}{\cancel{2}}+t\right]_0^2 = c\left[\frac{8}{3}-4+2\right] = c\,\frac{8-12+6}{3} = \frac{2}{3}c \rightarrow c=\frac{3}{2}$$



a) moda: bimodale $\{0,2\}$

media: 1 } per simmetria

mediana: 1

$\sqrt{Var(x)}$ $\qquad Var(x) = E(x^2) + E(x)^2$

$$E(x^2) = \int_R t^2\cdot f_x(t)\,dt = \int_0^2 t^2\cdot f_x(t)\,dt = \int_0^2 t^2\cdot\frac{3}{2}(t-1)^2\,dt = \frac{3}{2}\int_0^2 (t^4-2t^3+t^2)\,dt = \frac{3}{2}\left[\frac{t^5}{5}-\frac{t^4}{2}+\frac{t^3}{3}\right]_0^2 = \frac{3}{2}\left[\frac{32}{5}-\frac{16}{2}+\frac{8}{3}\right]_0^2 = \frac{8}{25}$$

$$Var(x) = E(x^2) - E(x)^2 = \frac{8}{5} - 1 = \frac{3}{5} \qquad \sqrt{Var(x)} = \frac{\sqrt{15}}{5} \approx 0.77$$

b) $F_x(t) = \int_{-\infty}^t f_x(t)\,dt = \int_0^t \frac{3}{2}(u-1)^2\,du = \frac{3}{2}\left[\frac{u^3}{3}-u^2+u\right]_0^t = \frac{3}{2}\left(\frac{t^3}{3}-t^2+t\right) = \frac{1}{2}$

Per la mediana (se non me ne accorgo prima), devo risolvere:

moltiplico per 2

$t^3-3t^2+3t-1=0 \longrightarrow (t-1)^3$

c) Voglio che $Z := (X-a)^b$ sia uniforme

TALMENTE DIFFICILE CHE ANCHE IL DOCENTE FA FATICA!

ESERCIZIO: esame (tipologia 1)

Giulia mangia cioccolatini abbastanza regolarmente. Ogni giorno indipendentemente dai precedenti, ne mangia un numero casuale pari a 0, 1, 2 o 5, con uguale probabilità che 1/4 ciascuno.

7 PUNTI) determinare la media e la dev standard del numero di cioccolatini mangiati da Giulia in un 1 giorno e in 7 giorni. ~~Determinare approssimativamente la probabilità che una scatola di 12 cioccolatini non sia sufficiente per 7 giorni.~~

3 ~~PUNTI~~) Giulia ha ora due scatole nuove (24 cioccolatini in tutto). Per quanti giorni le basteranno come minimo di 80% di probabilità? In altre parole, detto T il numero di giorni prima di finire i cioccolatini, trovare L intero tale che P(T > L) 80%.

2 PUNTI) determinare la probabilità esatta che in 3 giorni Giulia mangi 6 cioccolatini.

Giorni $i = 1, 2, \dots, n$

X_i = n° cioc. mangiati nel giorno 1

$x_1, x_2, \dots$ indip e equiprobabili

discreta

a) $E(X_i) = \sum_k k\cdot\varphi_{x_i}(k)$

$$= 0\cdot\frac{1}{4} + 1\cdot\frac{1}{4} + 2\cdot\frac{1}{4} + 5\cdot\frac{1}{4} = 2$$

$E(x_i^2) = \sum_{K} K^2 \cdot \varphi_{x_i}(K)$

$= 0^2 \cdot \frac{1}{4} + 1^2 \cdot \frac{1}{4} + 2^2 \cdot \frac{1}{4} + 5^2 \cdot \frac{1}{4} = \frac{1^2+2^2+5^2}{4} = \frac{15}{2}$

$Var(x) = E(x_i^2) - E(x_i)^2 = \frac{15}{2} - 2 = \frac{7}{2}$ $\sqrt{Var(x)} = \frac{\sqrt{14}}{2} \approx 1.87$

S_7 : # di cioc. mangiati in 7 giorni

$E(S_7) = E(x_1 + \cdots x_7) = E(x_1) + \cdots E(x_7) = 7 \cdot 2 = 14$

$\sqrt{Var(S_7)} \overset{\text{indip.}}{=} \sqrt{7} \cdot \sqrt{Var(x)} = \sqrt{7} \cdot \frac{\sqrt{14}}{2} \approx 4.9$

$\sqrt{Var\left(\sum_i^7 x_i\right)} = \sqrt{7}\sqrt{Var(x_i)}$

c) $P(S_3 = 6) = P((x_1, x_2, x_3) \in \{\underbrace{(2,2,2)}_{\frac{1}{4^3}}, (5,1,0), (5,0,1), (1,5,0), (0,5,1), (1,0,5), (0,1,5)\}$

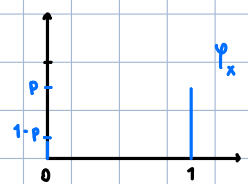
$= 7 \cdot \frac{1}{4^3} = \frac{7}{64} \approx 10.9\%$

DISTRIBUZIONE DI BERNOULLI, discreta

Esperimento con due esiti possibili

- ha un parametro: p, la prob che esca 1
- Valori possibili: $V_x = \{0, 1\}$
- Legge (pmf):

$$\varphi_x(K) = \begin{cases} 1-p & K=0 \\ p & K=1 \end{cases}$$



- si denota con: "bin (1, p)", es $X \sim bin(1, p)$ X ha distribuzione di Bernoulli di parametro p
- momenti: $E(x) = p$ $Var(X) = p(1-p)$

DISTRIBUZIONE BINOMIALE, discreta

Numero di successi su un certo numero di prove iid ripetute

- ha due parametri: n = n° prove $n \geq 1$ p = prob. di succ. $p \in [0, 1]$
- Valori possibili: $V_x = \{0, 1, 2, \cdots, n\}$
- Legge (pmf):

$$\varphi_x(K) = \binom{n}{K} p^K (1-p)^{n-K} \qquad K = 0, 1, \cdots, n$$

- si denota con: "bin" $X \sim bin(n, p)$ X ha la distribuzione binomiale di parametri di n e p
- momenti: $E(x) = np$ $Var(X) = np(1-p)$
- riproducibilità: $X \sim bin(m, p)$ $Y \sim bin(n, p)$ indipendenti stesso p

  Allora $X + Y \sim bin(m+n, p)$

ESERCIZIO:

Se un votante scelto a caso è favorevole a una certa riforma con probabilità 0.7, qual è la probabilità che su 10 votanti, esattamente 7 siano favorevoli?

$X \sim bin(n, p)$ $n = 10$ $p = 0.7$

$$P(X=7) = \varphi_{bin(n,p)}(7) = \binom{n}{k} \cdot p^k (1-p)^{n-k}$$
$$= \binom{n}{7} \cdot p^7 (1-p)^{n-7}$$
$$= \binom{10}{7} \cdot p^7 (1-p)^{10-7}$$
$$= \frac{10!}{7!\,3!} \cdot 0.7 \, (1-0.7)^3 \approx 0.2668$$

ESERCIZIO:
Uno dei sistemi installati su un satellite è costituito da 4 componenti, e erise a funzionare correttamente se almeno 2 di essi sono efficienti. Se ciascuno dei componenti, indipendentemente dagli altri, funziona bene con probabilità do 0.6, qual è la probabilità che l'intero sistema funzioni?

P(sistema funzioni) = $P(X \geq 2)$

X : # di componenti efficienti   mi serve la legge di X $X \sim bin(4, 0.6)$

$$\varphi_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$= \binom{4}{k} 0.6^k (1-0.6)^{4-k} \qquad k = 0,1,2,3,4$$

oppure

$$P(X \geq 2) = \sum_{k=2}^{4} \varphi_X(k) = 1 - P(X \leq 1) = 1 - \varphi_X(0) - \varphi_X(1) = 0.8208$$

RICORDA:
1) Se nell'es2 mi da la f. Di ripartizione, per trovare la legge/pmf devo derivare.
Se devo trovare la moda, studio il segno e quindi derivo la pmf, trovo le soluzione e guardo l'intervallo della funzione che mi interessa.
2) c'è un modo per trovare la media usando direttamente la f. Di ripartizione?
Si, se la variabile in questione non è negativa

$$\int_0^{\infty} (1 - F_X(t))\, dt$$

DISTRIBUZIONE DI POISSON, discreta

Numero di successi in una situazione tipo binomiale con p piccolo (e n, p incerti)
= la legge di Poisson approssima la binomiale quando p è piccola

- ha un parametro: ν, numero medio di successi $\nu \geq 0$
- Valori possibili: $V_X = \{0, 1, 2, \dots\} = \mathbb{N}$
- Legge (pmf):

$$\varphi_X(k) = \frac{\nu^k}{k!} e^{-\nu} \qquad k = 0, 1, \dots$$

- si denota con: $X \sim Pois(\nu)$ X ha distribuzione di Poisson di media ν
- momenti: $E(X) = \nu$ $Var(X) = \nu$

-> bin (n, p) ha media np e varianza np(1- p) che sono quasi uguali con p piccola

- riproducibilità:

$X \sim Pois(\nu_1)$ $Y \sim Pois(\nu_2)$ indip.

Allora $X + Y \sim Pois(\nu_1 + \nu_2)$

ESERCIZIO
Il numero medio di errori tipografici per pagina di una certa rivista è di 0.2. Qual è la probabilità che la

pagina che ti accingi a leggere contenga

A) nessun refuso

B) 2 o più refusi

P(0 refusi)   P(2 o più refusi)

X: n° refusi di una pag. $P(X=0)$, $P(X \geq 2)$

servirebbe la legge di X

$E(X) = 0.2$

Valori nascosti : $V_X = \{0, 1, 2, 3, \dots\}$ ⟶ $X \sim Pois(0.2)$

$$P(X=k) = \varphi_X(k) = \varphi_{Pois}(k) := \frac{\nu^k}{k!} e^{-\nu} \qquad k = 0, 1, 2, \dots$$

$$= \frac{0.2^k}{k!} e^{-0.2}$$

a) $P(X=0) = \frac{0.2^k}{k!} e^{-0.2} = e^{-\nu} \approx 0.819$

b) $P(X \geq 2) = 1 - [P(X=0) + P(X=1)] =$

$= 1 - [0.819 + 0.164] \approx 0.017$

DISTRIBUZIONE UNIFORME, continua:

Sceglie un punto “a caso” su un intervallo reale

- ha due parametri: a, b gli estremi dell'intervallo $a < b$
- Valori possibili: $V_X = [a, b] \subseteq \mathbb{R}$ OPPURE (a, b)
- Legge:

pdf:

$$f(t) = \frac{1}{b-a}, \quad a \leq t \leq b$$

CdF:

$$F(t) = \begin{cases} 0 & t \leq a \\ \frac{t-a}{b-a} & a < t \leq b \\ 1 & t > b \end{cases}$$

- si denota con: $X \sim unif(a, b)$ X ha distribuzione uniforme su [a, b]
- momenti: $E(X) = \frac{a+b}{2}$ $Var(X) = \frac{(b-a)^2}{12}$

  simmetria $\sqrt{Var(X)} \propto (b-a)$
- Trasformazioni lineari: unif → unif

  $U \sim unif(a, b)$ $V = mU + q$ $m \neq 0$

  $V \sim unif(ma + q, mb + q)$

DISTRIBUZIONE ESPONENZIALE, continua:

Modella un tempo casuale per un evento che avverrà all’improvviso, senza avvisaglie

- ha un parametro: λ, il "rate" o il tasso di accadimento
- Valori possibili: $V_X = [0, +\infty)$
- Legge:

pdf:

$$f(t) = \lambda e^{-\lambda t}, \quad t \geq 0$$

CdF:

$$F(t) = 1 - e^{-\lambda t}, \quad t \geq 0$$

- si denota con: $X \sim \text{expo}(\lambda)$
- momenti: $E(x) = \frac{1}{\lambda}$ $\quad Var(x) = \frac{1}{\lambda^2}$
- anotetie: expo $\rightarrow$ expo $\quad T \sim \text{expo}(\lambda) \quad S = \alpha \cdot T \quad \alpha > 0$

allora $S \sim \text{expo}\left(\frac{\lambda}{\alpha}\right)$

- Assenza di memoria:

t_0 $\quad T > t_0$ $\quad T - t_0$

"se anche è passato del tempo, e l'evento non è accaduto, il tempo che rimane da aspettare ha la stessa distribuzione che all'inizio"

$$P(T - t_0 > t \mid T > t_0) = P(T > t) \qquad \forall t, t_0 \geq 0$$

- Minimo di esponenziali indipendenti:

$T_1 \sim \text{expo}(\lambda_1) \quad T_2 \sim \text{expo}(\lambda_2), \dots, T_n \sim \text{expo}(\lambda_n)$ indip. $\quad n \geq 2$

$S := \min_{k=1,\dots,n} T_k$ $\quad$ Allora $\quad S \sim \text{expo}(\lambda_1 + \dots \lambda_n)$

DISTRIBUZIONE GAUSSIANA O NORMALE, continua:

Numero casuale effetto di tanti errori, con distribuzione a campana

- ha due parametri: μ, σ media e dev. standard $\quad \mu \in \mathbb{R} \quad \sigma > 0$
- Valori possibili: $V_x = \mathbb{R}$ $\quad$ [In pratica: $(\mu - 6\sigma, \mu + 6\sigma)$
- Legge:

pdf:

$$f(t) = \frac{1}{\sqrt{2\pi}\,\sigma} \cdot e^{-\frac{(t-\mu)^2}{2\sigma^2}}$$

CdF:

$$F(t) = \Phi\left(\frac{t-\mu}{\sigma}\right)$$

- si denota con: $X \sim \mathcal{N}(\mu, \sigma^2)$
- momenti: $E(x) = \mu$ $\quad Var(x) = \sigma^2$
- Trasformazioni lineari: $\mathcal{N} \rightarrow \mathcal{N}$

$Z \sim \mathcal{N}(\mu, \sigma^2) \qquad X = \alpha + \beta Z \sim \mathcal{N}(?, ?)$

$\Rightarrow X \sim \mathcal{N}(\alpha + \beta\mu, \beta^2 \sigma^2)$

- simmetria: $\Phi(-x) = 1 - \Phi(x)$

$$\Phi^{-1}(\alpha) = -\Phi^{-1}(1-\alpha)$$

- riproducibilità: $X \sim \mathcal{N}(\mu_x, \sigma_x^2)$ $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$ indipendenti

Allora

$$X+Y \sim \mathcal{N}(\mu_x+\mu_Y ; \sigma_Y^2+\sigma_Y^2)$$

TEOREMA DEL LIMITE CENTRALE:

Quando si sommano tanti effetti casuali indipendenti, il risultato è approssimativamente Gaussiano

ENUNCIATO:

X_1, X_2, … successione di vvaa iid con E (X_i)= a e Var (X_i)= b^2

$S_n := X_1 + \dots + X_n$ $\quad E(S_n) = na$ $\quad Var(S_n) = nb^2$

$Z_n := \frac{S_n - na}{\sqrt{n} \cdot b}$ Trasform. lineare standard

Allora $Z_n \xrightarrow[n\to\infty]{} \mathcal{N}(0,1)$ nel senso che $P(Z_n \le t) \xrightarrow[n\to\infty]{} \Phi(t)$ $\forall t \in \mathbb{R}$

ESERCIZIO:

La larghezza di una scanalatura in un trafilato di duralluminio è (in pollici) una variabile aleatoria normale con μ = 0.900 e σ = 0.0030. Le specifiche di fabbricazione assegnate impongono il limite 0.900 ± 0.0050.

A) che percentuale dei trafilati sarà difettosa?

B) qual è il più alto valore di σ accettabile, per avere una percentuale di difettosi non superiore all' 1%?

X: larghezza scanalatura

$X \sim \mathcal{N}(\mu, \sigma^2)$ $\to$ $X \sim \mathcal{N}(0.900, 9 \cdot 10^{-6})$

a) P(X sia fuori dal limite di tolleranza) = $1 - (0.9 - 0.005 < P < 0.9 + 0.005)$

CdF → $= 1 - (F_x(b) - F_x(a)) =$

$= 1 - (F_x(0.905) - F_x(0.895)) =$

ho solo Φ → $= 1 - \Phi\left(\frac{t-\mu}{\sigma}\right)$

$= 1 - \Phi\left(\frac{0.905-\mu}{\sigma}\right) - \Phi\left(\frac{0.895-\mu}{\sigma}\right)$

$= 1 - \Phi\left(\frac{0.005}{0.003}\right) - \Phi\left(\frac{0.005}{0.003}\right) = 1 - 0.9522 + 1 - 0.9522 \approx 9.56\% \approx 10\%$

se ho un negativo $= 1 - \Phi$

sono andata al contrario

b) $1\% \approx 2\left(1 - \Phi\left(\frac{0.005}{\sigma}\right)\right)$ $\quad \frac{0.01}{2} \approx \frac{2}{2} \cdot 1 - \Phi\left(\frac{0.05}{\sigma}\right)$

$-0.005 + 1 \approx + \Phi\left(\frac{0.05}{\sigma}\right)$

$\Phi^{-1} \cdot \Phi\left(\frac{0.005}{\sigma}\right) \approx 0.995 \cdot \Phi^{-1}$

$\frac{0.005}{\sigma} \approx \underbrace{\Phi^{-1}(0.995)}_{2.5758}$

$\sigma \approx \frac{0.005}{2.5758} \approx 0.0019$

ESERCIZIO:

Un test per il QI produce punteggi con distribuzione normale media 100 e dev standard 142. Che intervallo di punteggi raggiunge l'1% della popolazione formato dalle persone più intelligenti?

t

$X \sim \mathcal{N}(\mu, \sigma^2)$ $\mu = 100$ $\sigma = 14.2$

cerco t: $P(X > t) = 1\%$

$1\% = P(X > t) = 1 - F_X(t)$

$0.99 = F_X(t) = \Phi\left(\frac{t-\mu}{\sigma}\right)$

$\Phi^{-1}(0.99) = \frac{t - 100}{14.2} \longrightarrow t = 100 + 14.2\left[\Phi^{-1}(0.99)\right]$

$t = 100 + (14.2 \cdot 2.3263)$

$t \approx 133$

DISTRIBUZIONE DI CHI-QUADRO, continua:

Artificiale: somma di quadrati di $\mathcal{N}(0,1)$ indip.

- ha un parametro: K = n° di Gaussiane da sommare
- valori possibili: $V_X = (0, +\infty)$
- Legge:

CdF: tabulata

pdf:

$f(t) = c_K \cdot t^{\frac{K}{2}-1} \cdot e^{-\frac{t}{2}}$ $t > 0$

- si denota con: $X \sim \chi^2$
- momenti: $E(X) = K$ $Var(X) = 2K$
- definizione: $Z_1, Z_2, \dots, Z_K \sim \mathcal{N}(0,1)$ indip

$W := Z_1^2 + Z_2^2 + \dots + Z_K^2 \sim \chi^2(K)$

- Riproducibilità: $W_1 \sim \chi^2(K_1)$ $W_2 \sim \chi^2(K_2)$ indip

$W_1 + W_2 \sim \chi^2(K_1 + K_2)$

DISTRIBUZIONE T DI STUDENT, continua:

Artificiale: rapporto tra Z e Y indip. $Z \sim \mathcal{N}(0,1)$ $Y = \sqrt{\frac{W}{K}}$ dove $W \sim \chi^2(K)$

è molto simile a $\mathcal{N}(0,1)$ soprattutto per k grande

- ha un parametro: K = n° di gdl della chi-quadro coinvolta
- Valori possibili: $V_X = \mathbb{R}$
- Legge: CdF tabulata
- Si denota con: $X \sim t(K)$
- Approssimazione normale: $t(K) \xrightarrow[K \to \infty]{} \mathcal{N}(0,1)$ $K > 30$ possiamo approssimare $t(K) \dot{\sim} \mathcal{N}(0,1)$

_ valgono le stesse simmetrie

$F_{t(K)}(-x) = 1 - F_{t(K)}(x)$ $F^{-1}_{t(K)}(\alpha) = -F^{-1}_{t(K)}(1-\alpha)$

- Definizione:

$Z \sim \mathcal{N}(0,1)$ $W \sim \chi^2(K)$ indip

$Y = \sqrt{\frac{W}{K}}$ $\frac{Z}{Y} \sim t(K)$ il rapporto è una t di Student

ESERCIZI ESAME, tipologia 1:

Un fornitore produce componenti ciascuno dei quali ha una probabilità di esser difettoso pari a p=7.5%. I pezzi che non sono difettosi si dicono conformi. Consideriamo un lotto di 5600 pezzi di questo fornitore.

7 PUNTI) Quanto vale approssimativamente la probabilità che un numero di pezzi difettosi del lotto sia compreso tra 400 e 600? Cambia la risposta se a seconda di includono o escludono gli estremi dell'intervallo?

2 PUNTI) Determinare un limite inferiore L tale che il numero di pezzi conformi sia maggiore o uguale a L con probabilità dell'80% almeno.

3 PUNTI) per questo quesito non si suppone più che p = 7.5%, ovvero p va considerata come un parametro. Supponiamo di testare un campione di m pezzi e che siano tutti conformi. Qual è il massimo valore di p per cui questo evento abbia probabilità almeno del 10%?

X : n° pezzi difettosi

$P(400 \leq X \leq 600)$

$X \sim bin(n, p)$ $\quad n = 5600 \quad p = 7{,}5\% = 0.075$

$np = E(X) = 420 \quad \sqrt{np(1-p)} = \sqrt{Var(X)} = \sqrt{420 - 31.5} \approx 19.71$

$$P(400 \leq X \leq 600) \underset{\text{TLC}}{\approx} F_{\mathcal{N}}(600) - F_{\mathcal{N}}(400) = \Phi\left(\frac{600-420}{19.71}\right) - \Phi\left(\frac{400-420}{19.71}\right)$$

TLC $: np(1-p) \approx 390$

$$= \underbrace{\Phi(9.13)}_{1} - \underbrace{\Phi(1.014)}_{0.8449} \approx 15.5\% \approx 16\%$$

La c.c. qui cambia poco, e anche includere o escludere gli estremi

b) $L: 80\% = P(X \geq L)$

$0.2 = P(X < L) \approx \Phi\left(\frac{L - 420}{19.71}\right)$

$\left[\Phi^{-1}(0.2) \cdot 19.71\right] + 420 \approx L \longrightarrow \left[-\Phi^{-1}(0.8) \cdot 19.71\right] + 420 = L$

$\Phi^{-1}(\alpha) = -\Phi^{-1}(1-\alpha)$

$-0.8416 \cdot 19.71 + 420 = L$

$L \approx 403$

c) Livello too hard e I'm too tired

ESERCIZIO ESAME, tipologia 1:

Una linea di imbottigliamento alimentare produce n= 400 lotti di bevande all'anno. Durante la produzione di ciascun lotto, sto verificando casualmente degli arresti, in numero casuale, con legge che assume essere di Poisson di media *v*= 0.27 arresti per lotto.

7 PUNTI) determinare la media e la dev standard del numero di arresti in un anno di produzione. Determinare approssimativamente la probabilità che il numero totale di arresti sia superiore a 120. Quali assunzioni stiamo facendo implicitamente?

$n = 400 \quad \nu = 0.27$

$X_1, X_2, \dots, X_n \quad X_i \sim Pois(\nu)$

$S = X_1 + \dots + X_n \quad S \sim Pois(n \cdot \nu) \quad n\nu = 108$

$P(S > 120) \approx 1 - F_{\mathcal{N}}(120.5)$

$$\approx 1 - \Phi\left(\frac{120.5 - 108}{\sqrt{108}}\right) \approx 1 - \Phi(1{,}2) \approx 1 - 0.8849 \approx 11.5\% \approx 12\%$$

😬STATISTICA😬

ESERCIZI ESAME, tipologia 3:
Un'azienda testa l'affidabilità di un modello di matita meccanica. Un campione di 25 prodotti viene testato fino alla rottura e per ciascuno si annota il tempo di vita in mesi, trovando una media campionaria di 7.48 e una dev standard campionaria di 3.79. Si assume che la popolazione abbaio distribuzione Gaussiana di parametri μ e σ .
6 PUNTI) si stimino al 90% di confidenza sia μ (Intervallo bilaterale), sia σ (Intervallo unilaterale destro, tipo $\mu \leqslant$ U)
2 PUNTI) nel caso fosse proprio μ =7.48 e σ =3.79, che frazione delle matite durerebbe meno di due mesi? Qual è il minimo valore di μ per cui questa frazione sia inferiore al 10%?

$X_1, X_2, \dots, X_n$ campione statistico

$n = 25$ $\quad$ $X_i \sim \mathcal{N}(\mu, \sigma^2)$

$\bar{X} = 7.48$ $\quad$ $S_x = 3.79$

media campionaria

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

dev. standard campionaria

$$S_x = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{X})^2}$$

a)

i. Stimare μ al 90% di conf: bilaterale

livello di confidenza : $1 - \alpha = 0.9$

STIMA PUNTUALE: $\mu \approx \bar{x}$ $\qquad$ $\mu \approx 7.48$

formula :

$$\mu \in \bar{X} \pm \boxed{q \cdot \frac{S_x}{\sqrt{n}}} \xrightarrow{} r \qquad \text{dove} \quad q = F^{-1}_{t(n-1)}\left(1 - \frac{\alpha}{2}\right)$$

$\alpha = 0.1$ $\qquad$ $1 - \frac{\alpha}{2} = 1 - 0.05 = 0.95$ $\qquad$ $q = F^{-1}_{t(24)}(0.95) \approx 1.7109$

$r = 1.7109 \cdot \frac{3.79}{\sqrt{25}} \approx 1.3$ $\qquad$ $\mu \in 7.48 \pm 1.3 = [6.18 ; 8.78]$ al 90% di confidenza $\qquad$ rivedi lezione th29-30

ii.

$\alpha = 0.1$ $\qquad$ $1 - \alpha = 0.9$ $\qquad$ $b = F^{-1}_{\chi^2(n-1)}(1-\alpha) \approx 33.20$

$$1 - \alpha = P(\chi^2(n-1) \le b)$$
$$= P\left(\frac{S_x^2}{\sigma^2}(n-1) \le b\right)$$
$$= P\left(\frac{\sigma^2}{S_x^2 (n-1)} \ge \frac{1}{b}\right) = P\left(\sigma^2 \ge S_x^2 \cdot \frac{n-1}{b}\right) = P\left(\sigma \ge S_x \cdot \sqrt{\frac{n-1}{b}}\right)$$

volevo $\sigma^2 \le \dots$ e ho il contrario

$$a = F^{-1}_{\chi^2(n-1)}(\alpha) \approx 15.66$$

$$1 - \alpha = P(\chi^2(n-1) \ge a)$$
$$= P\left(\frac{S_x^2}{\sigma^2}(n-1) \ge a\right)$$
$$= P\left(\frac{\sigma^2}{S_x^2(n-1)} \le \frac{1}{a}\right) = P\left(\sigma^2 \le S_x^2 \cdot \frac{n-1}{a}\right) = P\left(\sigma \le S_x \cdot \sqrt{\frac{n-1}{a}}\right)$$

con il 90% di confidenza $\qquad$ $\sigma \le 3.79 \cdot \sqrt{\frac{24}{15.66}} \approx 4.69$

b) nuova hp: $\mu = 7.48$ $\quad$ $\sigma = 3.79$ $\qquad$ quale frazione ha $x < 2$ mesi?

$$P(x \le 2) = F_{\mathcal{N}(\mu, \sigma^2)}(2) = \Phi\left(\frac{2 - \mu}{\sigma}\right) \approx \Phi(-1{,}446)$$

$= 1 - \Phi(1.446) \approx 1 - 0.9258 \approx 7.42\% \approx 7\%$

$10\% = P(x \leq 2) = \Phi\left(\frac{2-\mu}{\sigma}\right) \iff \frac{2-\mu}{\sigma} = \Phi^{-1}(0.10) = \Phi^{-1}(0.90)$
≈ -1.2816

$\mu = 2 - \sigma \cdot (-1.2816) \approx 6.86$

B2) supponendo $\sigma = 3.79$ (non μ), stimare al 90% di confidenza unilaterale destro ($\leq C$), la frazione di matite con x < 2 mesi

devo stimare $\pi := P(x<2) = \Phi\left(\frac{2-\mu}{\sigma}\right) \approx \Phi\left(\frac{2-\mu}{3.79}\right) = \pi(\mu)$

non conosco π, ma conosco una formula che la lega a μ

* Stima puntuale $\hat{\pi} : \Phi\left(\frac{2-\bar{x}}{3.79}\right) \approx 7.42\%$

nei casi bilaterali $= 1 - \frac{\alpha}{2}$
nei unilaterali $=$ o α o $1-\alpha$

se io voglio $\pi \leq C \iff \mu \geq L$

intervallo di conf per μ unilat dx $\quad \mu \geq \bar{x} - q\frac{\sigma}{\sqrt{n}} \quad q = \Phi^{-1}(1-\alpha) \approx 1.2816$

$\mu \geq 7.48 - 1.2816 \cdot \frac{3.79}{5} \approx 6.51$

$\pi \leq C = \Phi\left(\frac{2-6.51}{3.79}\right) \approx \Phi(-1.19) \approx 1 - 0.883 \approx 11.7\% \approx 12\%$

al 90% di confidenza $\pi \leq 12\%$

MEDIA CAMPIONARIA: $\bar{X} := \frac{1}{n}\sum_{i=1}^{n} x_i$

STIMATORE BERNOULLIANO:

$X := x_1 + \dots + x_n \qquad \hat{p} := \frac{X}{n} \approx p$

DEVIAZIONE STANDARD CAMPIONARIA E VARIANZA CAMPIONARIA

Supponiamo $\mu = E(X_i)$ e $\sigma = \sqrt{Var(X_i)}$

I. $S_x := \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})}$ utilizzabile sempre

Recall: $\sigma = \sqrt{Var(x)} = \sqrt{E[(x-\mu)^2]}$

II. $\tilde{S}_x := \sqrt{\frac{1}{\mu}\sum_{i=1}^{n}(x_i - \mu)}$ solo quando si conosce μ

III. S_x^2 e $\tilde{S}_x^2$ sono le varianze campionarie

Le varianze sono stimatori consistenti

Formula per S_x^2

$S_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \qquad S_n^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right]$

DISTRIBUZIONE DEGLI STIMATORI

- media campionaria con dati Gaussiani $\bar{X} \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$

  $\bar{X} = \frac{1}{n}(X_1 + \dots + X_n)$
- media campionaria con dati non Gaussiani $\bar{X} \dot\sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$

⭐ spesso usiamo la versione standardizzata: (media 0 e var 1)

$$\frac{\bar{X}-\mu}{\sigma} \cdot \sqrt{n} \sim \mathcal{N}(0,1) \qquad \frac{\bar{X}-\mu}{\sigma} \cdot \sqrt{n} \dot\sim \mathcal{N}(0,1)$$

- stimatore p con dati Bernoulliani

⭐ versione std: $\dfrac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}} \dot\sim \mathcal{N}(0,1)$

- stimatore di con dati esponenziali:

⭐ $2n\lambda\bar{X} \sim \chi^2(2n)$

- varianza campionaria con dati Gaussiani

$$\frac{S_x^2}{\sigma^2}(n-1) \sim \chi^2(n-1)$$

- caso Gaussiano, μ e σ incognite, media campionaria e uso della t di Student

⭐ $\dfrac{\bar{X}-\mu}{S_x}\sqrt{n} \sim t(n-1)$

- Caso generale:

  $X_1, \dots, X_n$ qualsiasi $\mu = E(x_i)$ $\sigma = \sqrt{Var(x_i)}$

  $\mu \approx \bar{X}$ $\sigma^2 \approx S^2_x$

FUNZIONI ANCILLARI

DEFINIZIONE: una funzione ancillare per un parametro θ incognito è una statistica dipendete da θ E non altri parametri incogniti, con distribuzione nota.

COSA MANCA?

- μ:
  - Ma io ho σ ?
    - Si: $\dfrac{\bar{X}-\mu}{\sigma} \cdot \sqrt{n} \sim \mathcal{N}(0,1)$
    - No: $\dfrac{\bar{X}-\mu}{S_x} \sim t(n-1)$

- σ:
  - Ma io ho μ?
    - Si:

      $$\frac{\tilde{S}^2_x}{\sigma^2} \cdot (n) \sim \chi^2(n)$$

    - No:

      $$\frac{\tilde{S}^2_x}{\sigma^2} \cdot (n-1) \sim \chi^2(n-1)$$

- p: $$\frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \dot{\sim} \mathcal{N}(0,1)$$

- λ: $2n\lambda\bar{X} \sim \chi^2(2n)$