

1. VARIABILE STATISTICA X
2. MEDIA ARITMETICA SEMPLICE - Ponderata / VARIANZA
3. SCARTO QUADRATICO MEDIO / EVENTO / SPAZIO DEI CAMPIONI
4. COMBINAZIONI SEMPLICI
5. ALGEBRA DI BOOLE / SPAZIO FONDAMENTALE DELLA PROBABILITÀ (DEFINIZIONI)
6. DEFINIZIONE CLASSICA
7. CONCEZIONE FREQUENTISTA
8. LEGGE EMPIRICA DEL CASO / CONCEZIONE SOGGETTIVA
9. CONCEZIONE ASSIOMATICA DI KOLMOGOROV / EVENTI INCOMPATIBILI - INDIPENDENTI
10. PROBABILITÀ CONDIZIONATA
11. FORMULA DI BAYES
13. TEOREMA DELLE CAUSE
14. VARIABILE ALEATORIA / SPAZIO DI PROBABILITÀ INDOTTO / INSIEME DI BOREL
15. VARIABILE ALEATORIA DISCRETA
16. FUNZIONE DI RIPARTIZIONE
17. V. A. CONTINUA
21. MOMENTI DELLA VAR. ALEATORIA - DALLA MEDIA - DALL' ORIGINI / MEDIA DI UNA FUNZIONE DI UNA VARIABILE ALEATORIA
22. FUNZIONE GENERATRICE DEI MOMENTI / V. A. DI BERNOULLI
23. V. A. BINOMIALE
25. V. A. DI POISSON
26. V. A. CONTINUE: X = GAMMA / INTEGRALE EULERIANO
28. χ^2 = CHI QUADRATO
30. TEOREMA DELL'ADDITIVITÀ DEL χ^2 / V. A. NORMALE o GAUSSIANA
31. STANDARDIZZAZIONE
33. TEOREMA DEL LIMITE CENTRALE
34. V. A. T DI STUDENT

36. V.D. F DI FISHER
37. INFERENZA STATISTICA PARAMETRICA
38. POPOLAZIONE: MODELLO STATISTICO PARAMETRICO / SPAZIO DEI PARAMETRI
40. STIMA / FUNZIONE STATISTICA / STIMATORE / STIMA PUNTUALE
41. METODO DEI MOMENTI / M. TEORICI - EMPIRICI
43. METODO DELLA MASSIMA VEROSIMIGLIANZA / LIKELY HOOD FUNCTION
43. PROPRIETÀ DEGLI STIMATORI
50. P. DELLA CORRETTA - CONSISTENZA / DISTORSIONE / ERRORE QUADRATICO MEDIO DI θ
51. P. DELL'EFFICIENZA (ASSOLUTA - RELATIVA) / DISEGUAGLIANZA DI CRAWER-RAO
52. COVARIANZA
53. STIMA PER INTERVALLO (= INTERVALLO DI CONFIDENZA)
54. PERCENTILE / LIVELLO DI FIDUCIA
55. PROVA D'IPOTESI / TEORIA DI NEYMAN - PEARSON
56. ERRORE DI PRIMA - SECONDA SPECIE
58. POTENZA DEL TEST / RIGORE DEL TEST + POTENTE / RAPPORTO DI VEROSIMIGLIANZA GENERALIZZATO
61. REGRESSIONE
63. REGRESSIONE LINEARE SEMPLICE
64. METODO DEI MINIMI QUADRATI
65. DEVIANZA / CODEVIANZA / COEFF. DI REGRESSIONE / INDIPENDENZA IN MEDIA / COEFFICIENTE DI DETERMINAZIONE R^2
69. INTERDIPENDENZA
70. DISEGUAGLIANZA DI SCHWARTZ / COEFFICIENTE DI CORRELAZIONE LINEARE DI BRAVUS - PEARSON

Libro: Giuseppe Cicchitelli - Probabilità e Statistica - Raccardi

H

VARIABILE STATISTICA X

È un insieme di modalità (valore che può assumere X) e frequente (associate alla modalità).

Modalità x_i	Frequente assoluta m_i
x_1	m_1
x_2	m_2
\vdots	\vdots
x_i	m_i
\vdots	\vdots
x_n	m_n
	N

 $N = \text{numero totale}$

$$N = \sum_{i=1}^n m_i$$

H

Approccio pratico:

Se materializzo X (ad ex. $X = \text{altezza}$, $N = 100$)

X	h
170	5
174	35
178	20
180	15
185	25

100 (totale)

La variabile statistica è teorica.

Si può materializzare e si ha una frequenza statistica reale.

 $(x_1, x_2, \dots, x_i, \dots, x_n)$

x	h
x_1	1
x_2	1
\vdots	\vdots
x_n	1
N	

Gli elementi in questo 2° caso hanno frequenza unitaria

Variabile statistica (altezza, peso, ...)

(1)

Mutabile (senza colori occhi) non QUANTITA'

↳ FENOMENO COLLETTIVO espresso già più minore

MUTABILE fenomeno li di attributi

H

MEDIA

$$\mu_x = \frac{\sum_{i=1}^k (x_i) \cdot (n_i)}{N}$$

→ MEDIA ARITMETICA
Ponderata

$$\mu_x = \frac{\sum_{i=1}^N x_i}{N}$$

→ MEDIA ARITMETICA
SEMPLICE

Indice che sintetizza la distribuzione:

SCART $x_i - \mu_x$
 (o SCOSTAMENTO)
 $\begin{cases} \text{POSITIVO} > 0 \\ \text{NEGATIVO} < 0 \end{cases}$
 H

VARIANZA

"La Variabilità è l'attribuzione di un fenomeno ad assumere certe modalità"

Se c'è variabilità ci può essere statistica

La Varianza è un indice x la misura della variabilità

$$\sigma^2 = \sum_{i=1}^k \frac{(x_i - \mu_x)^2 \cdot n_i}{N} \quad (1^{\circ} \text{ caso})$$

$$\sigma^2 = \sum_{i=1}^N \frac{(x_i - \mu_x)^2}{N} \quad (2^{\circ} \text{ caso})$$

(2)

- Se lo scarto è piccolo, la media rappresenta bene la distribuzione.

$$0 \leq \sigma^2 \leq \infty$$

- Se i valori sono tutti uguali la variabilità del fenomeno è 0 (scarti nulli)

- $>$ è la variabilità, $<$ la μ è rappresentativa

H

SCARTO QUADRATICO MEDIO

$$\sigma = \sqrt{\sigma^2} \rightarrow \text{ottengo un indice di variabilità nella stessa unità di misura del fenomeno}$$

Valido x una variabile statistica. Essi infatti sono:

↓

VARIABILE: { STATISTICA: insieme di modalità e frequenze osservate su un collettivo
ALEATORIA: fenomeno casuale

H

ESPERIMENTO CASUALE: ... i cui risultati non si possono prevedere con certezza.

SPAZIO DEI CASUALI: insieme dei possibili risultati. [S]

Ciascun elemento dello S è detto EVENTO ELEMENTARE.

- INCOMPATIBILI (non se ne possono verificare + di 1)

- NECESSARIO (se ne deve verificare almeno 1)

Ex: lancio di 1 dado: $S = \{1, \dots, 6\}$

Se lancio 2 dadi ho $S = \{(1,1), (1,2), \dots, (1,6), (2,1), \dots, (6,6)\}$

↓

La struttura di S è strettamente connessa al tipo di esperimento.

EVENTO: "otto insieme dello spazio dei campioni (S)". ③

$$E_1 = \{2, 4, 6\} \quad E_1 \subset S \quad \text{EVENTO: numero pari}$$

$$E_2 = \{1, 3, 5\} \quad \text{e dispari}$$

$$E_3 = \{1, 2, 3\} \quad \text{numero} \leq 3$$

$E =$ SPAZIO DEGLI EVENTI: insieme di tutti i possibili sottoinsiemi di S . E' formato da 2^5 elementi
 POWER SET \rightarrow numerata dallo spazio dei campioni

In questo caso $E = 2^5$ eventi = 64 eventi

COMBINAZIONI SEMPLICI

"Quanti gruppi di k elementi posso formare in un insieme di n elementi?"

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

n : elementi
 k : n° di elementi nel gruppo

Sottoinsiemi di dimensione:	n dei sottoinsiemi
1	6
2	$\binom{6}{2} = \frac{6!}{2!4!} = \frac{6 \cdot 5 \cdot 4!}{2! 4!} = 15$
3	$\binom{6}{3} = \frac{6!}{3!3!} = \frac{6 \cdot 5 \cdot 4 \cdot 3!}{3! 3!} = 20$
4	$\binom{6}{4} = \frac{6!}{4!2!} = \frac{6 \cdot 5 \cdot 4!}{4! 2!} = 15$
5	$\binom{6}{5} = \frac{6!}{5!1!} = 6$
6	$\binom{6}{6} = 1 \rightarrow$ EVENTO: PR = 1 CERTO: PR = 1
0	1 EVENTO: Insieme vuoto IMPOSSIBILE vuoto

Si considerano diversi 2 gruppi quando differiscono per almeno un elemento.

④ Lo spazio degli eventi e' un'algebra di BOOLE.

Insieme chiuso rispetto alle operazioni di unione e complementazione degli insiemi.

• UNIONE su $2E$:

$$\{1, 2, 3\} \cup \{2, 5\} = \{1, 2, 3, 5\} \text{ questo insieme } \in \text{ già } \text{col } E$$

• COMPLEMENTO su E (\bar{E})



Complemento di $\{1, 2, 3, 4\}$ è $\{5, 6\}$ che già \in col E

• INTERSEZIONE dati $A, B \in E$ e se $\bar{A}, \bar{B} \in E$ allora $(\bar{A} \cup \bar{B}) \in E$,
 $(\bar{A} \cup \bar{B}) \in E$ allora per la legge di DE MORGAN.
 \hookrightarrow (Complemento delle unioni dei complementi)

$$(\bar{A} \cup \bar{B}) = (\bar{A} \cup \bar{B}) = (A \cap B)$$

CARATTERISTICHE DI UN'ALGEBRA DI BOOLE

Dati $A, B \in E$:

1) $A \cup B \in E$

2) $A \in E \Rightarrow \bar{A} \in E$

3) $A \cap B \in E$

Dato S , la complementazione di S è l'insieme vuoto \emptyset .

\bar{A}

SPAZIO FONDAMENTALE DELLA PROBABILITA'

(S, E, P) \rightarrow probabilità di ogni evento

La PROBABILITA' è un numero che esprime il grado di verosimiglianza di un evento

$$\text{IMPOSSIBILE} \leftarrow 0 < P < 1 \rightarrow \text{CERTO}$$

DEFINIZIONI DELLA PROBABILITA':

- 1) Definizione classica o matematica della probabilità
- 2) " Frequenzistica
- 3) " soggettivistica

DEFINIZIONE CLASSICA DI PROBABILITÀ

1) "E" è il rapporto tra n° casi favorevoli all'evento e n° di casi possibili"

$$P\{1\} = \frac{1}{6}$$

$$P(E) = \frac{\text{n° casi favorevoli}}{\text{n° casi possibili}}$$

$$P(E) = \frac{n \text{ casi favorevoli evento}}{n \text{ " " possibili}} ; S = \{1, 2, 3, 4, 5, 6\}$$

$$P(1) = \frac{1}{6} \text{ se } E = \{2, 4, 6\} \rightarrow P(E) = \frac{3}{6} = \frac{1}{2}$$

Campionato a 20 squadre. P che anno dopo si ha stessa classifica. Casi possibili: $20!$ (permutazioni di 20 squadre) \rightarrow stessi oggetti con ordine diverso. Casi favorevoli: le prime 3 nello stesso ordine, le rimanenti 17 in tutti i modi possibili ($17!$)

$$P(E) = \frac{17!}{20!}$$

Se voglii che ordine prime non conta ho $3! / 17!$

Contr. 4 giocatori di poker con 52 carte (13 carte a testa)

P che tutti i giocatori hanno stesse carte mano prelevate. Casi possibili: $52!$

" favorevoli: $13!$ A ogni giocatore, contr. che conten. contemporaneamente deve avvenire, ho $13! / 13! / 13! / 13! = (13!)^4$

Ex: 1 2 (6 carte a 2 giocatori)

A	B	C
A	C	B
B	A	C
B	C	A
C	A	B
C	B	A

D	E	F
E	F	D
D	F	E
E	D	F
F	D	E
F	E	D

$$6 \cdot 6 = 36 \quad [3!]^2$$

Casi poss: 6!

[Concezione classica]

Ex: urna con 10 R e 5 N



P che esca rosso? $10/15$ Nero: $5/15$

CONCEZIONE FREQUENTISTA



So che ho R e N ma non so quante ce ne sono.

(7)

No conc. classica.

Ripeto esperimento n volte; ho $\frac{R}{n}$ e $\frac{N}{n}$
↓
FREQUENZA RELATIVA

Tela STATISTICA:

x_i	m_i
x_1	m_1
x_2	m_2
\vdots	\vdots
x_i	m_i
\vdots	\vdots
x_k	m_k
Tot	N

$$S_2 \quad y_i = \frac{m_i}{N}$$

x_i	y_i
x_1	y_1
x_2	y_2
\vdots	\vdots
x_i	y_i
\vdots	\vdots
x_k	y_k

Perché f, π al posto di quelle assolute? Se volessimo confrontarli.

Es: ho N_1 motorini e m_1 incidenti a Roma. A Viterbo N_2 e m_2 incidenti a Roma o a Viterbo? Ovvero che $m_1 > m_2$, collettivo diverso. $y_1 = \frac{m_1}{N_1}$ e $y_2 = \frac{m_2}{N_2} \Rightarrow$ eliminare influenza N_1 e N_2

(come se ponessi numerata collettivo pari a 1)]

Prendo freq. come PROBABILITÀ dell'evento (è la LEGGE EMPIRICA DEL CASO: "in una Σ di prove fatte sotto le stesse condizioni al \uparrow del n di prove, la f rel. di 1 evento n avvicina alla P dell'evento stesso". (interpret. "asintotica" della \uparrow)
La PROBABILITÀ di un E è il $\lim_{n \rightarrow \infty}$ F.R. (come freq.)

CONCEZIONE SOGGETTIVA

Prevedo una legge, valore, FIDUCIA del sospetto al verificarsi dell'evento e la P misura la

⑧ fiducia.

$P(\Delta)$ è quanto proberei x avere 1€ se si verifica A ,
 $0 \in \mathbb{R}$ B.

Rx: giochi senza memoria (roulette), minimo 2 volte
che ad ex. è uscito R [conc. freq.]

CONCEZIONE ASSIOMATICA DI KOLMOGOROV

$$P: S \rightarrow \mathbb{R}_{[0,1]} \quad \forall E \in S, P(E) \geq 0 \quad P(S) = 1 \text{ (evento Certo)}$$

- se $E_1, E_2 \in S$ sono incompatibili [$E_1 \cap E_2 = \emptyset$], la
 $P(E_1 \cup E_2) = P(E_1) + P(E_2)$

(E_1) (E_2)

- se E_1, E_2, \dots, ∞ , $P(E_1 \cup E_2 \cup \dots \cup E_\infty) = \sum_{i=1}^{\infty} P(E_i)$

- $P(\emptyset) = 0$

(S, E, P) = spazio di probabilità o spazio fondamentale

che varia ad ogni esperimento

1. Eventi compatibili e incompatibili

2. " " dipendenti e indipendenti

- Due σ + E sono INCOMPATIBILI se il verificarsi dell'uno
esclude il verificarsi dell'altro

- Due E sono DIPENDENTI se il verificarsi dell'uno
influenza sulla P dell'altro in una prova successiva.

(ex: Tiro a bersaglio; E_1 : 90 alla 1° estratt.; E_2 : 90 alla 2° estratt.)

$P(E_1)$ [conc. classica] = $1/90$

\downarrow $P(E_2) = \begin{cases} 0 & \text{se } E_1 \text{ si è verificato} \\ 1/89 & \text{" " non si è "} \end{cases}$

Non ho ricomposto la \mathbb{R} iniziale.

Altrimenti ho eventi INDIPENDENTI.

Riepilogo:

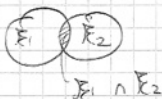
(\rightarrow)

(3)

- E_1 e E_2 INDIP: $P(E_1 \cup E_2) = P(E_1) + P(E_2)$

- " " " COMP: $P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$

nonno 2 volte e l'intersezione



- E_1 e E_2 INDIP: $P(E_1 \cap E_2) = P(E_1) P(E_2)$

- " " " DIP: $P(E_1 \cap E_2) = P(E_1) P(E_2/E_1) = P(E_2) P(E_1/E_2)$

Probabilità CONDIZIONATA

Dato E_2 , qual è la P su E_1

E_2 : urna con 5 palline

$\left. \begin{matrix} 1 \\ 2 \end{matrix} \right\}$ Rose Calcolo $P(I/R)$. Ho 5 risultati possibili. $P(I) = 1/5$. Conoscendo l'evento CONDIZIONATO, lo spazio del campione cambia! Verifichiamo: Posso, campione e solo 1 e 2 $\Rightarrow P(I/R) = 1/2$

PROBABILITÀ CONDIZIONATA

la P cond su un evento: $P(A/B) = \frac{P(A \cap B)}{P(B)}$ da cui

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

$$P(B/A) = \frac{P(A \cap B)}{P(A)} \text{ da cui } P(A \cap B) = P(A) P(B/A)$$



quindi def. eventi DIPENDENTI

Se indipendenti, $P(A/B) = P(A)$ e $P(A \cap B) = P(A)P(B)$

$$P(A/B) = \frac{P(A \cap B)}{P(B)} ; P(B/A) = \frac{P(A \cap B)}{P(A)} \quad 3/10/07$$

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

INDIPENDENZA
STOCASTICA

Se $P(A/B) = P(A)$ [indipendenti] e $P(B/A) = P(B)$

⑩ Quindi in questo caso $P(A \cap B) = P(A) P(B)$

Contr.urna di 10 palline numerate (eventi incompatibili)

O pallina	P(0)	colore	P(colore)	P(0/col)
R { ① ② ③ ④ ⑤ ⑥	1/10 " " " " "	R = {①, ②, ③}	$P(R) = P(① \cup ② \cup ③) = P(①) + P(②) + P(③) = 3/10$	$P(0/R) = 1/3$
V { ⑦ ⑧ ⑨ ⑩	" " " "	V = {④, ⑤, ⑥, ⑦, ⑧, ⑨, ⑩}	$P(V) = 7/10$	$P(0/V) = 1/7$

Supponiamo che siano le prime 3 rosse e le 7 verdi.

Voglio $P(①/R) = \frac{P(① \cap R)}{P(R)} \quad (① \cap R = ①) = \frac{P(①)}{P(R)} = \frac{1/10}{3/10} = \frac{1}{3}$

P. rose. $P(④/R) = \frac{P(④ \cap R)}{P(R)} = \frac{P(\emptyset)}{P(R)} = 0$

Se c'è evento condizionante, matris campione e sottoinsieme di essi che danno luogo a evento condizionante.

FORMULA DI BAYES

S. si $C_1, C_2, \dots, C_i, C_k$ partizione di S [Partizione insieme se è insieme dei sottoinsieme di S che sono disgiunti e la cui unione è S].

$$C_i \cap C_j = \emptyset; i \neq j; i, j = 1, \dots, k, \quad \bigcup_{i=1}^k C_i = S$$

Ex: $S = \{1, 2, 3, 4, 5, 6\}$; $E_1 = \{1, 3, 5\}$ e $E_2 = \{2, 4, 6\}$

Siano $P(C_i)$ $P(C_k)$ le probabilità di questi eventi.

[prob. A PRIORI delle C_i] (almeno 1 necessario)

Se A evento che si verifica o con C_1 , o con C_2 , ... o con C_i , ... con

$$A = (A \cap C_1) \cup (A \cap C_2) \cup (A \cap C_i) \dots \cup (A \cap C_k)$$

Allora $P(A) = P(A \cap C_1) + P(A \cap C_2) + \dots + P(A \cap C_k) = \sum_{i=1}^k P(A \cap C_i)$
 = (dalle prob. condizionate) = $\sum_{i=1}^k \underbrace{P(C_i)}_{\text{A PRIORI}} \underbrace{P(A/C_i)}_{\text{A POSTERIORI}}$

Voglio $P(C_i) \rightarrow P(C_i/A) \quad i=1 \dots k$
 ↳ prob. A POSTERIORI delle C_i

$$P(C_i/A) = \frac{P(C_i \cap A)}{P(A)} = \frac{P(C_i) P(A/C_i)}{\sum_{i=1}^n P(C_i) P(A/C_i)} \quad i=1, 2, \dots, n$$

le denom. è un fattore di normalizzazione, porta
 [So che $\sum_{i=1}^n P(C_i/A) = 1 = \frac{1}{\sum_{i=1}^n P(C_i) P(A/C_i)} [P(C_1) P(A/C_1) + \dots + P(C_n) P(A/C_n)]$]

le somme delle P pari a 1

Ex: 3 urne con palline B, R, N in diversa composizione.

URNA	Composizione			Totale
	B	R	N	
C ₁	2	3	1	6
C ₂	3	4	2	9
C ₃	4	2	3	9

Scegliere un'urna a caso ed estrarre, in ordine sup.
 casuale [no rimpiazzamento] due palline.

Supponiamo che sia verificato $A = \{(\text{B}, \text{N})\}$

Prob. di scelta di 1 delle 3 urne e $P(C_1) = P(C_2) = P(C_3) = \frac{1}{3}$

Voglio quindi $P(C_i/A)$. Use, Verificatore A, qual è la
 prob. che è l'urna sia la C_i ?

Il calcolo:

$$P(A/C_1) = P((\text{B}, \text{N})/C_1) = P(R \cap N/C_1) = P(R) P(N/A) = \frac{2}{6} \cdot \frac{1}{6} = \frac{1}{18}$$

$$P(A/C_2) = \frac{3}{9} \cdot \frac{2}{9} = \frac{1}{9}$$

$$P(A/C_3) = \frac{2}{9} \cdot \frac{3}{9} = \frac{1}{12}$$

$$\text{Calcolo } P(A) = \sum_{i=1}^3 P(C_i) P(A/C_i) = \frac{1}{3} \left[\frac{1}{18} + \frac{1}{9} + \frac{1}{12} \right] = \frac{1}{3} \cdot \frac{53}{180}$$

$$\text{Quindi } P(C_1/A) = \frac{\frac{1}{3} \cdot \frac{1}{18}}{\frac{1}{3} \cdot \frac{53}{180}} = \frac{180}{530} = \frac{18}{53}$$

$$P(C_2/A) = \frac{\frac{1}{3} \cdot \frac{1}{9}}{\frac{1}{3} \cdot \frac{53}{180}} = \frac{20}{53} \quad ; \quad P(C_3/A) = \frac{\frac{1}{3} \cdot \frac{1}{12}}{\frac{1}{3} \cdot \frac{53}{180}} = \frac{15}{53}$$

per alcune urne la P_A port. e $> P_A$ priori e V. Cervera:

$$\frac{18}{53} = 0,34; \quad \frac{20}{53} = 0,38; \quad \frac{15}{53} = 0,28$$

$\downarrow \qquad \qquad \downarrow \qquad \qquad \downarrow$

$$P((1/\Delta) > P(c_1)) \quad P((2/\Delta) > P(c_2)) \quad P((3/\Delta) < P(c_3))$$

Formula detta anche TEOREMA DELLE CAUSE

2° esempio di applicazione della f. di Bayes.

8/10/2007

Azienda che produce scarpe, ha 2 macchine che producono scarpe. La 1° produce il 60%, la 2° il 40%.

$\pi_1 \rightarrow 60\%$, 10% delle scarpe prodotte sono difettose

$\pi_2 \rightarrow 40\%$, 20% " " " " "

A = scarpe NON difettose

$$P(M_1) = \frac{60}{100} = \frac{6}{10} \quad \text{Probabilità che una scarpa della fabbrica provenga da } \pi_1$$

$$P(M_2) = \frac{4}{10}$$

$$A = (A \cap M_1) \cup (A \cap M_2)$$

$$1^*) \quad P(A) = P(A \cap M_1) + P(A \cap M_2) = P(M_1)P(A/\pi_1) + P(M_2) + P(A/\pi_2)$$

$$P(\pi_i / A) = \frac{P(\pi_i) \cdot P(A/\pi_i)}{\sum_{i=1}^n P(\pi_i) \cdot P(A/\pi_i)} \quad i = 1, 2$$

$$P(A/\pi_1) = \frac{30}{100} = \frac{3}{10}; \quad P(A/\pi_2) = \frac{8}{10}$$

$$P(A) = \frac{6}{10} \cdot \frac{3}{10} + \frac{4}{10} \cdot \frac{8}{10} = 0,86 \quad \text{da } 1^*)$$

$$\left. \begin{aligned} P(\pi_1 / A) &= \frac{P(\pi_1) \cdot P(A/\pi_1)}{0,86} = \frac{0,18}{0,86} = 0,21 \\ P(\pi_2 / A) &= \frac{P(\pi_2) \cdot P(A/\pi_2)}{0,86} = \frac{0,32}{0,86} = 0,37 \end{aligned} \right\} \text{ la somma è pari a 1}$$

Probabilità a priori

Probabilità a posteriori

0,6 $\xrightarrow{\quad\quad\quad}$ 0,63

0,4 $\xrightarrow{\quad\quad\quad}$ 0,37

4

VARIABILITA' ALEATORIA O CASUALE (RANDOM VARIABLE)

$(S, \mathcal{F}, P) =$ Spazio fondamentale

2^6 elementi di \mathcal{F}

Variable Dominio Codominio

$X(\omega)$ $S \rightarrow R$

Con questa variabile posso definire lo SPAZIO DI

PROBABILITA' INDOTTO (R, \mathcal{B}, P_X)

\rightarrow Insieme di BOREL

INSIEME DI BOREL:

"E' la minima classe additiva in cui tutti gli elementi sono intervalli aperti a sinistra e chiusi a destra" $(-\infty, x]$

E' chiusa rispetto alle operazioni di complementazione e di unione.

Es: lancio di 3 monete (testa T o croce C)

$S =$

1	TTT	\rightarrow	1	3
2	CCC	\rightarrow	2	0
3	CTT	\rightarrow	3	2
4	TCT	\rightarrow	4	1
5	TTC	\rightarrow		
6	CCT	\rightarrow		
7	CTC	\rightarrow		
8	TCC	\rightarrow		

$\rightarrow R = \{0, 1, 2, 3\}$

$\forall B \in \mathcal{B}, P_X(B) = P\{1 \in S, X(\omega) = B\} = P(X^{-1}(B))$

(14) con $X^{-1}(B)$: INVERSA INVERSA DI R

$\forall B \in \mathcal{B} \quad X^{-1}(B) \in \mathcal{F}$, e la CONDIZIONE DI MISURABILITÀ

collegamento tra la \mathcal{P} del vecchio S e l'insieme di Borel

la VARIABILE ALEATORIA è una funzione numerabile da S a \mathbb{R} .
la variabile casuale è una funzione dallo spazio
fondamentale ($S \in \mathcal{P}$) allo spazio \mathbb{R} misurabile.
L'immagine inversa degli insiemi $B \in \mathcal{B}$ è allo spazio
degli eventi originari \mathcal{E} .

H

TIPICI DI VARIABILE ALEATORIA (V.a.)

• V.a. DISCRETA: è definita su uno spazio dei campioni
discreto (formato da un numero finito di punti o
un'infinità numerabile di campioni)

X : tutta la variabile

x : i valori che essa assume

$f_x(x) =$ funzione di probabilità o f di probabilità di massa

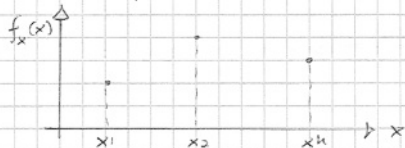
$F_x(x) =$ " " ripartizione

$f_x(x) =$ prob. che X assuma un valore uguale x
 $P\{X=x\}$

$$1) f_x(x) \geq 0 \quad \forall x \in D$$

$$2) \sum_{x \in D} f_x(x) = 1$$

Ad ogni variabile $x_1 \dots x_n$ corrisponde $f_x(x_1) \dots f_x(x_n)$



Fix: lancio di 2 dadi

X = somma delle 2 facciate

Valori di X	2	3	4	5	6	7	8	9	10	11	12
$f_x(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

$$X=2 \quad \{(1,1)\} \quad \Pr\{1 \cap 1\} = P(1) \cdot P(1) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$$

$$X=3 \quad \{(1,2) \cup (2,1)\} \quad \Pr\{X=3\} = P(1 \cap 2) + P(2 \cap 1) =$$

REUNIONE
DISGIUNTA

$$= \frac{1}{36} + \frac{1}{36} = \frac{2}{36}$$

$$X=4 \quad \{(2,2) \cup (1,3) \cup (3,1)\} = \frac{3}{36}$$

$$X=5 \quad \{(1,4) \cup (4,1) \cup (2,3) \cup (3,2)\} = \frac{4}{36}$$

$$X=7 \quad \{(1,6) \cup (6,1) \cup (3,4) \cup (4,3) \cup (5,2) \cup (2,5)\} = \frac{6}{36}$$

Verifico che $\sum f_x(x) = 1$

FUNZIONE DI RIPARTIZIONE

$$F_x(x) = \Pr\{X \leq x\} = \sum_{x \leq x} f_x(x) = \text{SET FUNCTION}$$

$$F_x(2) = \frac{1}{36}; \quad F_x(3) = f_x(2) + f_x(3) = \frac{3}{36}$$

$$F_x(4) = f_x(2) + f_x(3) + f_x(4) = \frac{7}{36} \quad (\text{probabilità che } X \leq 4)$$

X	2	3	4	5	6	7	8	9	10	11	12
$F_x(x)$	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{6}{36}$	$\frac{10}{36}$	$\frac{15}{36}$	$\frac{21}{36}$	$\frac{26}{36}$	$\frac{30}{36}$	$\frac{33}{36}$	$\frac{35}{36}$	1

1) $F_x(x)$ è una fun. non decrescente: $a < b$, $b > a$

16 allora $F_x(b) \geq F_x(a)$

$$2) \lim_{x \rightarrow -\infty} F_X(x) = F_X(-\infty) = 0$$

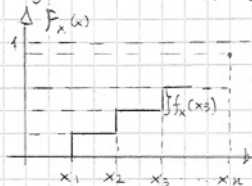
$$3) \lim_{x \rightarrow +\infty} F_X(x) = F_X(+\infty) = 1$$

4) Questa funzione è continua a destra:

$$\lim_{y \rightarrow x^+} F_X(y) = F_X(x) \quad , \quad \lim_{h \rightarrow 0^+} F_X(x+h) = F_X(x)$$

5) Se a, b con $a < b$ allora $\Pr\{a \leq x \leq b\} = F_X(b) - F_X(a)$
 Si presenta come una funzione a gradini:

$$\text{Impatto: } F_X(x_3) - F_X(x_2) = [f_X(x_1) + f_X(x_2) + f_X(x_3)] - [f_X(x_1) + f_X(x_2)] = f_X(x_3)$$



H

• V.a. CONTINUA

Variabile che assume tutti i valori di uno spazio campionario continuo.

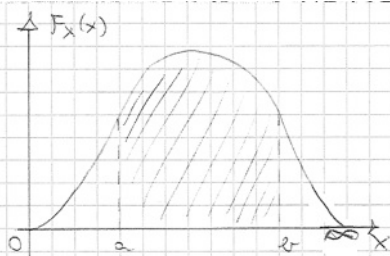
ha v.a. X è continua se \exists una funzione $f_X(x) \geq 0 \quad \forall$ valore della variabile x la probabilità

$$\Pr\{x \leq u\} = F_X(u) \text{ sia pari a:}$$

$$F_X(u) = \int_{-\infty}^u f_X(x) dx = F_X(u) - F_X(-\infty)$$

(nel caso di v.a. continua la soluzione diventa un integrale)

(\rightarrow)

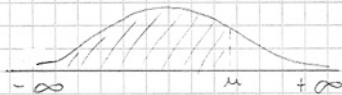


$$\int_0^{\infty} f_X(x) dx = 1$$

$$P_R \{a \leq x \leq b\} = \text{area} \square = \int_a^b f_X(x) dx = \int_0^b f_X(x) dx - \int_0^a f_X(x) dx = F_X(b) - F_X(a)$$

↓

allo stesso modo:



$$\int_{-\infty}^u f_X(x) dx = F_X(u) - F_X(-\infty) = F_X(u)$$

la X è continua se $f_x(x)$ non negativa 9/10/07

$$\text{se } R = \{x \leq u\} = F_x(u) = \int_{-\infty}^u f_x(x) dx = F_x(u) - F_x(-\infty)$$

$$f_x(x) = \Pr\{x=x\} \rightarrow \text{funzione di probabilità}$$

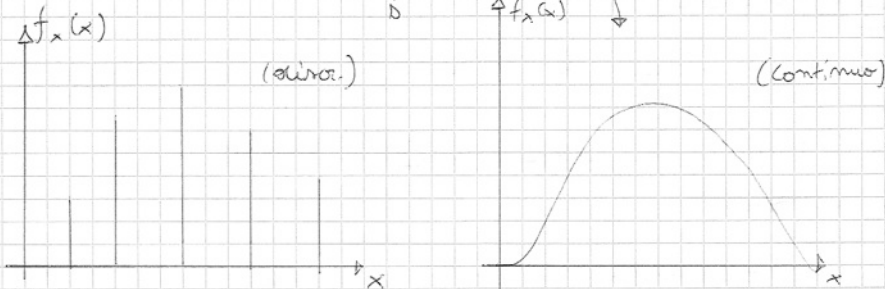
$$x \text{ discreta: } F_x(x) = \Pr\{x \leq x\} \rightarrow \text{funzione di ripartizione}$$

$$x \text{ continua: } \begin{cases} f_x(x) \rightarrow \text{funzione di DENSITA' della PROBABILITA' =} \\ = \Pr\{x \leq X \leq x + \Delta x\} \end{cases}$$

$$F_x(x) = \Pr\{x \leq x\} = \text{funzione di RIPARTIZIONE}$$

$$\Pr\{x=x_1\} = \Pr\{x_1 \leq x \leq x_1\} = \int_{x_1}^{x_1} f_x(x) dx = 0$$

$$f_x(x) \geq 0. \text{ Sia } x \in D, \int_{-\infty}^{+\infty} f_x(x) dx = 1$$



Proprietà funzione di ripartizione:

$$1. \lim_{x \rightarrow -\infty} F_x(x) = F_x(-\infty) = 0; \lim_{x \rightarrow +\infty} F_x(x) = F_x(+\infty) = 1$$

$$2. F_x(x), a, b \text{ con } a \leq b, F_x(b) \geq F_x(a)$$

$$3. \Pr\{a \leq x \leq b\} = F_x(b) - F_x(a)$$

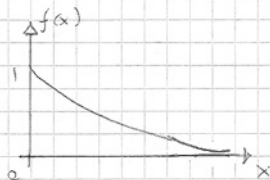
$$\Pr\{a \leq x \leq b\} = \int_a^b f_x(x) dx = F_x(b) - F_x(a)$$

4. F_X è assolutamente continua $\Rightarrow \frac{d}{dx} F_X(x) = f_X(x)$

$$\frac{d}{dx} F_X(x) = \frac{d}{dx} \int_0^x f(t) dt \quad ; \text{dal th. del val. medio si ha}$$

$$= f(x)$$

Es: sia $f_X(x) = \begin{cases} e^{-x} & \text{per } x \geq 0 \\ 0 & \text{altrimenti} \end{cases}$

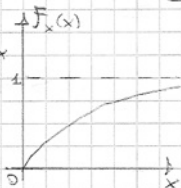


Allora:

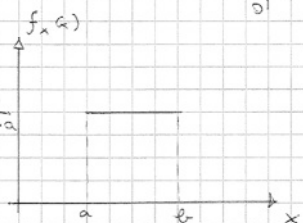
$$F_X(u) = \int_{-\infty}^u f_X(x) dx = \int_0^u e^{-x} dx = - \int_0^u e^{-x} dx = - \left[e^{-x} \right]_0^u = -(-1)e^{-u} = e^{-u}$$

$$= -e^{-x} \Big|_0^u = -e^{-u} + 1 = 1 - e^{-u} \Rightarrow F_X(x) = 1 - e^{-x}$$

$$\frac{d}{dx} F_X(x) = -(-1)e^{-x} = e^{-x}$$



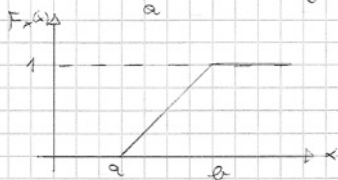
Es: $f_X(x) = \begin{cases} \frac{1}{b-a} & x \in [a, b] \\ 0 & \text{altrimenti} \end{cases}$



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

$$F_X(u) = \int_a^u \frac{1}{b-a} dx = \frac{1}{b-a} \int_a^u dx =$$

$$= \frac{1}{b-a} x \Big|_a^u = \frac{u-a}{b-a}$$



MOMENTI DELLA V.V. X

2) Momenti Centratò (o "dalla media") : $E[x - \mu_x]^2$
 $\mu = E[x - \mu_x]$

1) Momenti non centratò (o "dalla origine") : $E(x^n)$

1) $E(x^n)$: ^{media delle} n -esime dei valori della X

→ "expectation", media variabile aleatoria

$$E(x^n) = \begin{cases} \text{X discreta: } \sum_{x \in D} x^n \cdot f_x(x) \\ \text{X continua: } \int_D x^n f_x(x) dx \end{cases}$$

• Con $n=1$ ho la media ARITHMETICA = $\sum_{x \in D} x f_x(x)$ o $\int_D x f_x(x) dx$

• Con $n=2$ ho il momento SECONDO = $\sum_{x \in D} x^2 f_x(x)$ o $\int_D x^2 f_x(x) dx$

$$2) E(x - \mu)^n = \begin{cases} \sum_{x \in D} (x - \mu)^n f_x(x) \\ \int_D (x - \mu)^n f_x(x) dx \end{cases}$$

• Con $n=2$ ($E(x - \mu)^2$ o $E(x - E(x))^2$) ho la VARIANZA di x =
 $Var(x) = \sigma_x^2 = E(x^2) - [E(x)]^2$

Dim nel caso continuo:

$$\begin{aligned} \sigma_x^2 &= E(x - \mu)^2 = \int_D (x - \mu)^2 f(x) dx = \int_D (x^2 + \mu^2 - 2x\mu) f(x) dx = \\ &= \int_D x^2 f(x) dx + \mu^2 \int_D f(x) dx - 2\mu \int_D x f(x) dx = E(x^2) + \mu^2 + \\ &- 2\mu \mu = E(x^2) - [E(x)]^2 \end{aligned}$$

TEOREMA DI UNA FUNZIONE DI UNA VARIABILE ALEATORIA

Sia X con $f_x(x)$ e sia $g(x)$ una funzione di questa X (2)

Allora:

$$E(g(x)) = \begin{cases} \sum_{x \in D} g(x) f_x(x) \\ \int_D g(x) f(x) dx \\ H \end{cases}$$

FUNZIONE GENERATRICE DEI MOMENTI (F.G.M.)

Sia X con $f(x)$. Allora $m_x(t) = E(e^{tx}) = \begin{cases} \sum_{x \in D} e^{tx} f_x(x) \\ \int_D e^{tx} f_x(x) dx \end{cases}$

Nota: F.G.M. $\exists \forall t \in (-h, h)$

Ove \exists e' derivabile con continuita' intorno all' origine.
Se voglio m^n , calcolo la $\frac{d^n m_x(t)}{dt^n}$ e la calcolo in $t=0$.

Es: $E(x) = \mu$

$\frac{d}{dt} m_x(t) \Big|_{t=0} \quad m_x(t) = \int_D e^{tx} f(x) dx$

$m'_x(t) = \int_D x e^{tx} f(x) dx$. In $t=0$ ho $\int_D x f(x) dx (= \mu)$

$E(x^2)$: $m''_x(t) = \int_D x \cdot x e^{tx} f(x) dx$. In $t=0$ ho $\int_D x^2 f(x) dx$

H

3 tipi di variabile aleatorie: (discrete)

1. BERNOLLIANA

2. BINOMIALE

3. POISSON

1) $f_x(x) = p^x (1-p)^{1-x}$ con $x=0,1$ e $p \in [0,1]$

Quindi $P\{x=0\} = f_x(0) = 1-p$

(22) $P\{x=1\} = f_x(1) = p$

Vogliamo μ e σ_x^2 .

$$\bullet \underline{E(x)} = \sum_{x=0}^{\infty} x f_x(x) = \sum_{x=0}^{\infty} x P^x (1-P)^{1-x} = \underline{P}$$

$$\bullet \underline{Var(x)} = E[x - E(x)]^2 = E[x - P]^2 = \sum_{x=0}^{\infty} (x-P)^2 f_x(x) = \\ = \sum_{x=0}^{\infty} (x-P)^2 P^x (1-P)^{1-x} = P^2(1-P) + (1-P)^2 P = P(1-P)[P + (1-P)] = \underline{P(1-P)}$$

Confrontiamo i risultati ottenuti con la F.G.M.:

$$\underline{m_x(t)} = E[e^{tx}] = \sum_{x=0}^{\infty} e^{tx} f_x(x) = \sum_{x=0}^{\infty} e^{tx} P^x (1-P)^{1-x} = \\ = 1-P + e^t P ; \frac{d}{dt} m_x(t) = e^t P. \text{ In } t=0 \text{ ho } \underline{P}$$

So che $\sigma_x^2 = E(x^2) - [E(x)]^2$. Calcolo $E(x^2)$ con F.G.M.:

$$\frac{d^2}{dt^2} m_x(t) = e^t P. \text{ In } t=0 \text{ ho } P. \underline{Var(x)} = P - P^2 = \underline{P(1-P)}$$

2) Considero il problema delle prove ripetute: ho le alternative A e B. Estraggo e rimetto dentro per n volte. Il campione con N elementi e ci sono N_1 di A e N_2 di B. Ho $P(A) = N_1/N$, $P(B) = N_2/N$.
 $\hookrightarrow P$ $\hookrightarrow Q = 1-P$

$$\underline{f_x(x) = \binom{n}{x} P^x (1-P)^{n-x}} \text{ con } x=1, 2, \dots, n$$

Vogliamo μ e σ_x^2 :

$$\bullet E(x) = \sum_{x=0}^n x f_x(x) = \sum_{x=0}^n x \binom{n}{x} P^x (1-P)^{n-x} = \\ = \sum_{x=1}^n x \binom{n}{x} P^x (1-P)^{n-x} = \sum_{x=1}^n x \frac{n!}{x!(n-x)!} P^x (1-P)^{n-x} =$$

$$= \sum_{x=1}^n \cancel{x} \frac{n(n-1)!}{\cancel{x}(x-1)!} \cdot \frac{1}{(n-x)!} \cdot P \cdot P^{x-1} (1-P)^{n-x} = nP \sum_{x=1}^n \binom{n-1}{x-1} P^{x-1} (1-P)^{n-x}$$

Variable aleatoria binomiale $f_x(x) = \binom{n}{x} P^x Q^{n-x} \quad x=0, \dots, n \quad P \in [0,1]$ 10/10/07

$$\sum_{x=0}^n \binom{n}{x} P^x Q^{n-x} = (P+Q)^n; \text{ pongo } t=x-1 \begin{cases} t=0; x=1 \\ t=n-1; x=n \end{cases}$$

$$\text{Quindi ho } nP \sum_{x=0}^{n-1} \binom{n-1}{t} P^t Q^{n-t-1} = nP (P+Q)^{n-1}$$

Con $E(x) = nP$

$$\sigma_x^2 = \text{Var}(x) = E(x^2) - [E(x)]^2 = E(x^2) - n^2 P^2$$

$$E(x^2) = E[x(x-1) + x] = E[x(x-1)] + E(x) = E[x(x-1)] + nP$$

$$\begin{aligned} \sigma_x^2 &= E[x(x-1)] + nP - n^2 P^2; \quad E[x(x-1)] = \sum_{x=0}^n x(x-1) f_x(x) = \\ &= \sum_{x=0}^n x(x-1) \binom{n}{x} P^x Q^{n-x} = \sum_{x=2}^n \cancel{x}(\cancel{x}-1) \frac{n(n-1)(n-2)!}{\cancel{x}(\cancel{x}-1)(x-2)!} \cdot \frac{1}{(n-x)!} \cdot \\ &\quad \underline{P \cdot P^{x-2} \cdot Q^{n-x}} = n(n-1) P^2 \sum_{x=2}^n \binom{n-2}{x-2} P^{x-2} Q^{n-x} \end{aligned}$$

Pongo $t=x-2 \begin{cases} t=0; x=2 \\ t=n-2; x=n \end{cases}$ Quindi ho:

$$n(n-1) P^2 \sum_{t=0}^{n-2} \binom{n-2}{t} P^t Q^{n-t-2} = n(n-1) P^2 (P+Q)^{n-2} =$$

$$= n(n-1) P^2$$

$$\begin{aligned} \text{Quindi } \sigma_x^2 &= n(n-1) P^2 + nP - n^2 P^2 = nP [(n-1)P + 1 - nP] = \\ &= nP(1-P) = \underline{nPQ} \end{aligned}$$

Approccio con la F.G.M.

$$\textcircled{24} \quad m_x(t) = E[e^{tx}] = \sum_{x=0}^n e^{tx} f_x(x) = \sum_{x=0}^n e^{tx} \binom{n}{x} P^x Q^{n-x} =$$

$$\sum_{x=0}^n \binom{n}{x} (e^t P + Q)^x Q^{n-x} = (e^t P + Q)^n$$

$$\frac{d m_x(t)}{dt} = m(e^t P + Q)^{n-1} e^t P ; \quad \frac{d^2}{dt^2} m_x(t) = m(n-1)(e^t P + Q)^{n-2} e^t P \cdot e^t P +$$

$$+ m(e^t P + Q)^{n-1} e^t P.$$

$$\cdot \underline{E(x)} = \frac{d}{dt} m_x(t) \Big|_{t=0} = m \underbrace{(P+Q)}_{I+1}^{n-1} P = \underline{mP} \quad \text{CVD}$$

$$E(x^2) = \frac{d^2}{dt^2} m_x(t) \Big|_{t=0} = m(n-1)(P+Q)^{n-2} \cdot P^2 + mP =$$

$$= m(n-1)P^2 + mP.$$

$$\cdot \underline{\text{Var}(x)} = m(n-1)P^2 + mP - m^2 P^2 = \underline{mPQ} \quad \text{CVD}$$

3) Discrete: $f_x(x) = \frac{e^{-\lambda} \lambda^x}{x!}$ com $x=0, 1, 2, \dots; \lambda \geq 0$

Qu'on $E(x) = \sigma_x^2 = \lambda$

$$\cdot \underline{E(x)} = \sum_{x=0}^{\infty} x f_x(x) = \sum_{x=0}^{\infty} x \frac{e^{-\lambda} \lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} x \frac{\lambda \cdot \lambda^{x-1}}{x!} =$$

$$= e^{-\lambda} \cdot \lambda \sum_{x=0}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} = \left(\text{développement en série de } e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} \right)$$

$$= \lambda e^{-\lambda} e^{\lambda} = \underline{\lambda}$$

$$\cdot \sigma_x^2 = E(x^2) - [E(x)]^2 = E[x(x-1) + x] - \lambda^2 = E[x(x-1)] + \lambda - \lambda^2$$

$$E[x(x-1)] = \sum_{x=0}^{\infty} x(x-1) \frac{e^{-\lambda} \lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} x(x-1) \frac{\lambda^2 \lambda^{x-2}}{x(x-1)(x-2)!} =$$

$$= e^{-\lambda} \cdot \lambda^2 \sum_{x=0}^{\infty} \frac{\lambda^{x-2}}{(x-2)!} = e^{-\lambda} \cdot \lambda^2 \cdot e^{\lambda} = \underline{\lambda^2}$$

$$\underline{\sigma_x^2} = \lambda^2 + \lambda - \lambda^2 = \underline{\lambda} \quad \text{CVD}$$

Applichiamo la FGR:

$$m_x(t) = E[e^{tx}] = \sum_{x=0}^{\infty} e^{tx} \cdot e^{-\lambda} \cdot \frac{\lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} (\lambda e^t)^x / x! =$$

$$= e^{-\lambda} e^{\lambda e^t}$$

$$\frac{d}{dt} m_x(t) = e^{-\lambda} \cdot \lambda e^t e^{\lambda e^t}; \quad \frac{d^2}{dt^2} m_x(t) = e^{-\lambda} \left[\lambda e^t e^{\lambda e^t} + \lambda e^t \lambda e^t e^{\lambda e^t} \right] = e^{-\lambda} \lambda e^t e^{\lambda e^t} [1 + \lambda e^t]$$

$$\bullet \underline{E(x)} = \left. \frac{d}{dt} m_x(t) \right|_{t=0} = e^{-\lambda} \lambda e^{\lambda} = \underline{\lambda} \quad \text{CVD}$$

$$E(x^2) = \left. \frac{d^2}{dt^2} m_x(t) \right|_{t=0} = e^{-\lambda} \lambda e^{\lambda} [1 + \lambda] = \lambda(1 + \lambda)$$

$$\bullet \underline{\sigma_x^2} = \lambda(1 + \lambda) - \lambda^2 = \underline{\lambda} \quad \text{CVD}$$

H

VARIABILI ALGEBRICHE CONTINUE

• X = GAMMA

$$f_x(x, \alpha, \beta) = \frac{1}{(\Gamma(\alpha)\beta^\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} \quad \text{con } x \geq 0 \text{ e } \alpha, \beta \geq 0$$

$\Gamma(\alpha) = (\alpha-1)!$. Ricorda integrale euleriano:

$$\alpha! = \int_0^{\infty} x^\alpha e^{-x} dx = \Gamma(\alpha+1)$$

Voglio media e Varianza.

con generico

norm. p.e. m.m.

$$\textcircled{26} \quad E(x^n) = \int_0^{\infty} x^n f_x(x) dx = \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^{\infty} x^n \cdot x^{\alpha-1} e^{-\frac{x}{\beta}} dx =$$

$$= \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty x^{\alpha+\pi-1} e^{-\frac{x}{\beta}} dx = \text{Ponzo } y = \frac{x}{\beta} \quad \begin{matrix} x = y\beta \\ dx = \beta dy \end{matrix}$$

$$= \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty (y\beta)^{\alpha+\pi-1} e^{-y} \beta dy = \frac{\beta^{\alpha+\pi}}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty y^{\alpha+\pi-1} e^{-y} dy =$$

$$= \frac{\beta^\pi}{\Gamma(\alpha)} \Gamma(\alpha+\pi)$$

$$\text{Quindi } E(x^\pi) = \frac{\beta^\pi}{\Gamma(\alpha)} \Gamma(\alpha+\pi)$$

• $E(x) = \frac{\beta}{\Gamma(\alpha)} \Gamma(\alpha+1)$; Poiché $\Gamma(\alpha+1) = \alpha \Gamma(\alpha)$ (è un fattore!)
 $= \alpha(\alpha-1)\Gamma(\alpha-1) \dots$ ho:

$$E(x) = \frac{\beta}{\Gamma(\alpha)} \alpha \Gamma(\alpha) = \alpha \beta$$

$$E(x^2) = \frac{\beta^2}{\Gamma(\alpha)} \Gamma(\alpha+2) = \frac{\beta^2}{\Gamma(\alpha)} (\alpha+1) \alpha \Gamma(\alpha) = \alpha(\alpha+1) \beta^2$$

$$\sigma_x^2 = \alpha(\alpha+1) \beta^2 - \alpha^2 \beta^2 = \alpha \beta^2 [(\alpha+1) - \alpha] = \alpha \beta^2$$

Approccio FGR:

$$m_x(t) = E[e^{tx}] = \int_0^\infty e^{tx} f(x) dx = \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty e^{tx} x^{\alpha-1} e^{-\frac{x}{\beta}} dx =$$

$$= \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty x^{\alpha-1} e^{-x \left[\frac{1}{\beta} - t \right]} dx \xrightarrow{\frac{1-t\beta}{\beta}} \text{Ponzo } y = x \left(\frac{1-t\beta}{\beta} \right)$$

quindi $x = y \frac{\beta}{(1-t\beta)}$ e $dx = \left(\frac{\beta}{1-t\beta} \right) dy$

$$m_x(t) = \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty \left(\frac{y}{1-t\beta} \right)^{\alpha-1} e^{-y} \cdot \frac{\beta}{1-t\beta} dy =$$

$$= \frac{1}{\Gamma(\alpha)\beta^\alpha} \frac{\beta^\alpha}{(1-t\beta)^\alpha} \int_0^\infty y^{\alpha-1} e^{-y} dy = \frac{\Gamma(\alpha)\beta^\alpha}{\Gamma(\alpha)\beta^\alpha} \frac{1}{(1-t\beta)^\alpha} = (1-t\beta)^{-\alpha}$$

$$E(X) = \frac{d}{dt} m_X(t) \Big|_{t=0} = \alpha \beta (1 - t\beta)^{-(\alpha+1)} \Big|_{t=0} = \underline{\alpha \beta} \quad (V)$$

$$E(X^2) = \frac{d^2}{dt^2} m_X(t) \Big|_{t=0} = (\alpha+1) \alpha \beta^2 (1 - t\beta)^{-(\alpha+2)} \Big|_{t=0} = \underline{\alpha \beta^2 (\alpha+1)} \quad (V)$$

Caso particolare Gamma: $X^2 = (H)$ QUADRATO

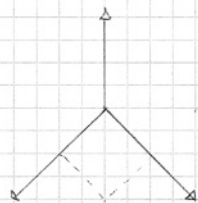
$\alpha = \frac{V}{2}$ e $\beta = 2$. Pongo $u = X^2$, quindi:

$$f_u(u) = \frac{1}{\Gamma(\frac{V}{2}) 2^{\frac{V}{2}}} \cdot u^{\frac{V}{2}-1} \cdot e^{-\frac{u}{2}} \quad \text{con } V = \text{gradi di libert\`a}$$

Con 3 numeri e normaliamo: $X_1 + X_2 + X_3$, 3 gradi

// // // / la somma ≤ 100 , ≥ 2 gradi

// // // e mediamo: $\mu = \frac{X_1 + X_2 + X_3}{3} \rightarrow$ Piano di termine
nota 3 μ



Il piano sventura il vincolo.

$V \equiv n$ Variabili libere $= n$ Variabili - Vincoli

$$(X_1 - \mu)^2 + (X_2 - \mu)^2 + (X_3 - \mu)^2 = 3\sigma^2$$

Sfera di raggio $\sqrt{3}\sigma$. Se metto a sistema col piano,

$V = 1$

$$X_V^2 = u$$

$$E(U^R) = \frac{1}{\Gamma(\frac{V}{2}) 2^{\frac{V}{2}}} \int_0^\infty U^{\frac{V}{2}-1} e^{-\frac{U}{2}} dU = \frac{1}{\Gamma(\frac{V}{2}) 2^{\frac{V}{2}}} \int_0^\infty U^{\frac{V}{2}-1} e^{-\frac{U}{2}} dU$$

$$e^{-\frac{u}{2}} du; \text{ Pongo } y = \frac{u}{2} \rightarrow u = 2y \rightarrow du = 2 dy$$

$$E(V^n) = \frac{1}{\Gamma(\frac{\nu}{2}) 2^{\frac{\nu}{2}}} \int_0^{\infty} (2y)^{\frac{\nu}{2} + n - 1} \cdot e^{-y} \cdot 2 dy = \frac{2^{\frac{\nu}{2} + n}}{\Gamma(\frac{\nu}{2}) 2^{\frac{\nu}{2}}}$$

$$\int_0^{\infty} y^{\frac{\nu}{2} + n - 1} \cdot e^{-y} dy = \frac{2^n}{\Gamma(\frac{\nu}{2})} \Gamma\left(\frac{\nu}{2} + n\right)$$

$$\Gamma\left(\frac{\nu}{2} + n\right) = \left(\frac{\nu}{2} + n - 1\right) \Gamma\left(\frac{\nu}{2} + n - 1\right) = \left(\frac{\nu}{2} + n - 1\right) \left(\frac{\nu}{2} + n - 2\right) \dots$$

$$\dots \frac{\nu}{2} \Gamma\left(\frac{\nu}{2}\right)$$

$$E(V^n) = 2^n \left(\frac{\nu}{2} + n - 1\right) \left(\frac{\nu}{2} + n - 2\right) \dots \frac{\nu}{2} = 2^n \cdot \prod_{i=1}^n \left(\frac{\nu + 2n - 2i}{2}\right)$$

$$= 2^n \cdot \prod_{i=1}^n \left(\frac{\nu + 2n - 2i}{2}\right) \Rightarrow E[X_V^2]^n = \prod_{i=1}^n (\nu + 2n - 2i)$$

$$\text{Se } n=1, \quad \underline{E(X_V^2) = \nu}$$

$$\text{Se } n=2, \quad E[X_V^2]^2 = \prod_{i=1}^2 (\nu + 4 - 2i) = (\nu + 2)^{\nu}$$

$$\sigma_x^2 = (\nu + 2)^{\nu} - \nu^2 = 2\nu$$

$$m_u(t) = E[e^{tu}] = \frac{1}{\Gamma(\frac{\nu}{2}) 2^{\frac{\nu}{2}}} \int_0^{\infty} e^{tu} \cdot e^{-\frac{y}{2}} \cdot u^{\frac{\nu}{2} - 1} dy =$$

$$= \frac{1}{\Gamma(\frac{\nu}{2}) 2^{\frac{\nu}{2}}} \int_0^{\infty} u^{\frac{\nu}{2} - 1} \cdot e^{-u(\frac{1}{2} - t)} du = \frac{1}{\Gamma(\frac{\nu}{2}) 2^{\frac{\nu}{2}}} \int_0^{\infty} u^{\frac{\nu}{2} - 1} \cdot e^{-u(\frac{1-t}{2})} du$$

$$\text{Ponho } y = u \left(\frac{1-t}{2}\right) \rightarrow u = y \left(\frac{1-t}{2}\right)^{-1} \rightarrow du = \left(\frac{1-t}{2}\right)^{-1} dy$$

$$m_u(t) = \frac{1}{\Gamma(\frac{\nu}{2}) 2^{\frac{\nu}{2}}} \int_0^{\infty} y^{\frac{\nu}{2} - 1} \cdot e^{-y} \cdot \frac{1}{\left(\frac{1-t}{2}\right)} dy =$$

$$= \frac{2^{\frac{\nu}{2}}}{\Gamma(\frac{\nu}{2}) \cdot 2^{\frac{\nu}{2}} \cdot (1-2t)^{\frac{\nu}{2}}} \cdot \int_0^{\infty} y^{\frac{\nu}{2}-1} \cdot e^{-y} \cdot dy = \frac{\cancel{\Gamma(\frac{\nu}{2})}}{\Gamma(\frac{\nu}{2})} \cdot \frac{1}{(1-2t)^{\frac{\nu}{2}}} = \frac{1}{(1-2t)^{\frac{\nu}{2}}}$$

$$\frac{d}{dt} m X_{\nu}^2(t) = -\frac{\nu}{2} (1-2t)^{-\left(\frac{\nu}{2}+1\right)} (-2) = \nu (1-2t)^{-\left(\frac{\nu}{2}+1\right)}$$

$$\frac{d^2}{dt^2} m X_{\nu}^2(t) = \left(\frac{\nu}{2}+1\right)(-\nu) \cdot (1-2t)^{-\left(\frac{\nu}{2}+1\right)} (-2) =$$

$$= 2\nu \left(\frac{\nu+2}{2}\right) (1-2t)^{-\left(\frac{\nu}{2}+1\right)}$$

$$-E(X_{\nu}^2) = \frac{d}{dt} m(X_{\nu}^2(t)) = \nu$$

$$-E[(X_{\nu}^2)^2] = \frac{d^2}{dt^2} m(X_{\nu}^2(t)) = \nu(\nu+2)$$

TEOREMA DELL'ADDITIVITA' DEL X_{ν}^2

Data $X_{\nu_1}^2$ con i_1 gradi di libertà e $X_{\nu_2}^2$ " i_2 " " " " con $X_{\nu_1}^2$ e $X_{\nu_2}^2$ indep.

$$f(X_{\nu_1}^2, X_{\nu_2}^2) = f(X_{\nu_1}^2) \cdot f(X_{\nu_2}^2) \text{ allora}$$

$$\underline{X_{\nu_1}^2 + X_{\nu_2}^2 = X_{\nu_1+\nu_2}^2} \quad \text{i cui g.d.l. sono } i_1 + i_2$$

Data n variabili $X_{\nu_1}, \dots, X_{\nu_n}$ allora $\sum_{i=1}^n X_{\nu_i}^2 = X_{\sum_{i=1}^n \nu_i}^2$

VARIABILE ALEATORIA NORMALE O GAUSSIANA

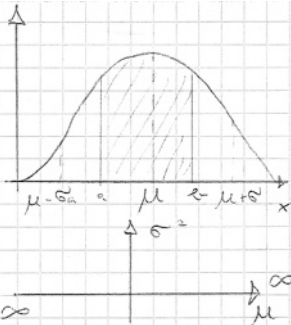
$$\textcircled{30} \quad X \sim N(\mu, \sigma_x^2) = f_x(x) = \frac{1}{\sigma_x \cdot \sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma_x^2} (x-\mu)^2\right\}$$

$$x, \mu \in \mathbb{R}; \sigma_x^2 \in \mathbb{R}^+$$

$$N(\mu, \sigma_x^2)$$

$$P_n \{a \leq x \leq b\} = F_x(b) - F_x(a) = f(-x) = f(x)$$

$$= \frac{1}{\sigma_x \sqrt{2\pi}} \left[\int_{-\infty}^b e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma_x}\right)^2} dx - \int_{-\infty}^a e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma_x}\right)^2} dx \right]$$



$$\{N(\mu, \sigma_x^2) / \mu \in \mathbb{R}, \sigma_x^2 \in \mathbb{R}^+\}$$

↳ Insieme di tutti i punti dell'insieme

STANDARDIZATION

$$X \sim N(\mu, \sigma_x^2) \quad Z = \frac{X - \mu}{\sigma_x} \sim N(0, 1) \Rightarrow Z \text{ è detta NORMALE STANDARDIZATA}$$

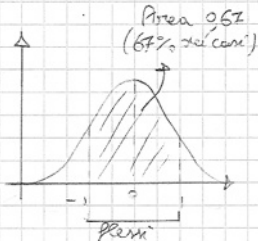
$$E(Z) = E\left[\frac{X - \mu}{\sigma_x}\right] = \frac{E(X) - \mu}{\sigma_x} = \frac{\mu - \mu}{\sigma_x} = 0$$

$$\sigma_z^2 = E[Z - E(Z)]^2 = E(Z)^2 = E\left[\frac{X - \mu}{\sigma_x}\right]^2 = \frac{E(X - \mu)^2}{\sigma_x^2} = \frac{\sigma_x^2}{\sigma_x^2} = 1$$

$$Z = \frac{X - \mu}{\sigma_x} \Rightarrow X = Z \cdot \sigma_x + \mu \rightarrow dx = \sigma_x dz$$

$$f_Z(z) dz = \frac{1}{\sigma_x \sqrt{2\pi}} \exp\left\{-\frac{1}{2} z^2\right\} \sigma_x dz =$$

$$P_n \{a \leq x \leq b\} = \frac{1}{\sqrt{2\pi}} \left[\int_{-\infty}^{\frac{b-\mu}{\sigma_x}} e^{-\frac{1}{2} z^2} dz - \int_{-\infty}^{\frac{a-\mu}{\sigma_x}} e^{-\frac{1}{2} z^2} dz \right] = F_Z\left(\frac{b-\mu}{\sigma_x}\right) - F_Z\left(\frac{a-\mu}{\sigma_x}\right)$$



$$m_T(t) = E(e^{tZ}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tZ} \cdot e^{-\frac{1}{2} Z^2} dZ = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tZ - \frac{1}{2} Z^2} dZ$$

$$tZ - \frac{1}{2} Z^2 + \frac{1}{2} t^2 - \frac{1}{2} t^2 = -\frac{1}{2} (t - Z)^2 + \frac{1}{2} t^2$$

$$\frac{e^{\frac{1}{2} t^2}}{\sqrt{2\pi}} \int_0^{\infty} e^{-\frac{1}{2} (t - Z)^2} dZ \quad \text{Poniamo } y = Z - t, Z = y + t \rightarrow dZ = dy$$

$$e^{\frac{t^2}{2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dy = e^{\frac{t^2}{2}}; m_z(t) = e^{\frac{t^2}{2}} \rightarrow m_x(t)$$

Poiché la Moments Generating Function è $Z = \frac{x-\mu}{\sigma_x} \Rightarrow x = Z \cdot \sigma_x + \mu \sim N(\mu, \sigma_x)$

Ogni combinazione lineare di variabili normali è normale:

$$E(x) = E(Z) \sigma_x + \mu = \mu$$

$$E(x) = E[x - E(x)]^2 = E[Z \sigma_x + \mu - \mu]^2 = E(Z^2) \cdot \sigma_x^2 = \sigma_x^2$$

Tr: "Data una var. aleatoria $y = \cos t$ si ha che $m_y(t) = e^{\frac{\sigma^2}{2}} m_x(t)$ "

$$m_x(t) = e^{\frac{\mu t}{\sigma_x}} \cdot m_z\left(\frac{t}{\sigma_x}\right) \quad m_x(t) = e^{\frac{\mu t}{\sigma_x}} \cdot e^{\frac{1}{2} \sigma_x^2 \frac{t^2}{\sigma_x^2}} = e^{\frac{\mu t}{\sigma_x}} \cdot e^{\frac{1}{2} t^2} = e^{\frac{\mu t}{\sigma_x} + \frac{1}{2} t^2}$$

$$\frac{d}{dt} m_x(t) = \frac{d}{dt} e^{\frac{\mu t}{\sigma_x} + \frac{1}{2} t^2} = \left(\frac{\mu}{\sigma_x} + t \right) e^{\frac{\mu t}{\sigma_x} + \frac{1}{2} t^2}$$

$$\frac{d^2}{dt^2} m_x(t) = \left(\frac{\mu}{\sigma_x} + t \right)^2 e^{\frac{\mu t}{\sigma_x} + \frac{1}{2} t^2} + e^{\frac{\mu t}{\sigma_x} + \frac{1}{2} t^2} = \left(\frac{\mu}{\sigma_x} + t \right)^2 + 1 \cdot e^{\frac{\mu t}{\sigma_x} + \frac{1}{2} t^2}$$

$$E(x) = \frac{d}{dt} m_x(t) \Big|_{t=0} = \mu; \quad E(x)^2 = \frac{d^2}{dt^2} m_x(t) \Big|_{t=0} = \mu^2 + \sigma_x^2$$

$$\text{Var } x = \mu^2 + \sigma_x^2 - \mu^2 = \sigma_x^2$$

Quadrato della variabile. Moments Generating Function: $Z^2 = X^2$

$$X \sim N(\mu, \sigma_x^2); \quad Z = \frac{X - \mu}{\sigma_x} \sim N(0, 1). \quad Z^2 = U$$

$$F_U(u) = P\{U \leq u\} = P\{Z^2 \leq u\} = P\{-\sqrt{u} \leq Z \leq \sqrt{u}\} = F_Z(\sqrt{u}) - F_Z(-\sqrt{u})$$

$$f_U(u) = \frac{d}{du} F_U(u) = F_Z'(\sqrt{u}) \cdot \frac{1}{2\sqrt{u}} - F_Z'(-\sqrt{u}) \cdot \left(-\frac{1}{2\sqrt{u}}\right)$$

$$f_z(z) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}z^2\right\} \quad z \in \mathbb{R}$$

17/10/2007

$$X_2^2 = U; \quad f_U(u) = \frac{1}{\Gamma(\frac{U}{2}) 2^{\frac{U}{2}}} \cdot e^{-\frac{U}{2}} \cdot u^{\frac{U}{2}-1} \quad u \geq 0$$

$$Z^2 = X_1^2; \quad X_1^2 \sim \frac{1}{\Gamma(\frac{1}{2}) 2^{\frac{1}{2}}} \cdot u^{\frac{1}{2}-1} \cdot e^{-\frac{u}{2}}$$

$$\text{Poniamo } U = Z^2. \quad F_U(u) = \Pr\{U < u\} = \Pr\{Z^2 \leq u\} = \\ = \Pr\{-\sqrt{u} \leq Z \leq \sqrt{u}\} = F_Z(\sqrt{u}) - F_Z(-\sqrt{u})$$

Applico la var. media:

$$f_u(u) = \frac{d}{du} F_U(u) = F_Z'(\sqrt{u}) (\sqrt{u})' - F_Z'(-\sqrt{u}) (-\sqrt{u})' = \\ = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u} \cdot \frac{1}{2\sqrt{u}} + \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u} \cdot \left(+\frac{1}{2\sqrt{u}}\right) = \\ = \frac{2}{2\sqrt{u}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u} = \frac{e^{-\frac{1}{2}u} u^{\frac{1}{2}-1}}{2^{\frac{1}{2}} \underbrace{\left(\sqrt{u}\right)}_{\substack{\text{rimuovendo la} \\ \text{parte } \Gamma(\frac{1}{2})}}} \quad \text{C.V.D.}$$

TEOREMA DEL LIMITE CENTRALE (Nella defn. che cont.)

Siano x_1, x_2, \dots, x_n n var. aleatorie indipendenti e con la stessa $f_X(x) \sim X$. [si dice "i.i.d."]

Quindi $E(x_i) = \mu \quad \forall i$ e $\sigma^2 = \text{Var}(x_i) \quad \forall i$.

Sia S_n la somma $\sum_{i=1}^n x_i$ (combin. lineare delle x_i).

$$\text{Allora } E(S_n) = E(x_1) + E(x_2) + \dots + E(x_i) + E(x_n) = n\mu,$$

$$\text{e } \text{Var}(S_n) = E[S_n - E(S_n)]^2 = E[x_1 + x_2 + \dots + x_n - n\mu]^2 =$$

$$= E[(x_1 - \mu) + (x_2 - \mu) + \dots + (x_n - \mu)]^2 = n E(x - \mu)^2 = \textcircled{33}$$

$n \sigma_x^2$ (solo in questo caso)

le th. afferma che:

$$\lim_{n \rightarrow +\infty} \Pr \left\{ \frac{S_n - n\mu}{\sqrt{n} \sigma_x} \leq t \right\} = \int_{-\infty}^t \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz$$

di X i.i.d.
ha come l.i.mite $N(0,1)$ Rate per $n \rightarrow \infty$ va verso una
normale standardizzata. ($N(0,1)$)

//

V.A. $T_D = t$ DI STUDENT CON D G.D.L.

Sia $Z \sim N(0,1)$. Sia $U \sim \chi_D^2$ e Z indipendenti.

Allora: (come nelle 2 var.)

$$T_D = \frac{Z}{\sqrt{\frac{U}{D}}}$$

Voglio det. la $f_{T_D}(t)$.

$$\text{Patto sia } f(z, u) = f_z(z) f_u(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \cdot \frac{1}{\Gamma(\frac{D}{2}) 2^{\frac{D}{2}}} u^{\frac{D}{2}-1} e^{-\frac{1}{2}u} = \frac{1}{\sqrt{2\pi} \Gamma(\frac{D}{2}) 2^{\frac{D}{2}}} u^{\frac{D}{2}-1} e^{-\frac{1}{2}(u+z^2)}$$

Trovo le variabili. Pongo T_D , quindi $z = T_D \sqrt{\frac{u}{D}}$ e
 $dz = \sqrt{\frac{u}{D}} dT_D$

$$\text{Ho } f(T_D, u) dT_D du = K \cdot u^{\frac{D}{2}-1} e^{-\frac{u}{2}(1+T_D^2/D)}$$

$$\sqrt{\frac{u}{D}} dT_D du$$

Voglio $f(T_D)$ ["marginale"]. Se in generale ho
 $f(x, y)$ con $x \in D$ e $y \in \mathbb{R}$, se voglio $f(x)$ ho
(34) (\rightarrow)

$$f(x) = \int_{\mathbb{R}} f(x, y) dy \quad e \quad f_y(y) = \int_{\mathbb{D}} f(x, y) dx$$

$$f_{T_V}(t) = \int_0^{\infty} f(T_V, u) du = \frac{k}{\sqrt{v}} \int_0^{\infty} e^{-\frac{v}{2}(1+\frac{t^2}{v})} u^{\frac{v}{2}+\frac{1}{2}-1} du$$

$[u \geq 0, T_V \in \mathbb{R}]$

Mi ricordo l'int. euleriano. Pongo $y = \frac{v}{2}(1+\frac{t^2}{v}) \Rightarrow$

$$u = \frac{2y}{1+\frac{t^2}{v}} \quad e \quad du = \frac{2}{1+\frac{t^2}{v}} dy \quad e \quad ho:$$

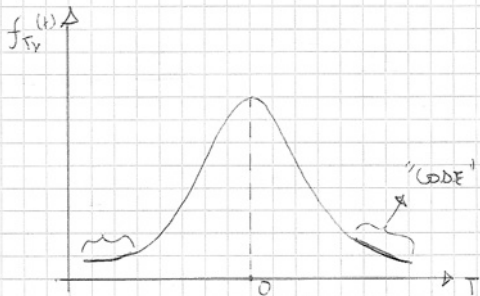
$$\begin{aligned} f_{T_V}(t) &= k \frac{1}{\sqrt{v}} \int_0^{\infty} \left(\frac{2u}{1+\frac{t^2}{v}} \right)^{\frac{v+1}{2}-1} \cdot e^{-y} \cdot \frac{2}{1+\frac{t^2}{v}} dy = \\ &= \frac{2^{\frac{v+1}{2}}}{\sqrt{2\pi v} \cdot \Gamma(\frac{v}{2})} \cdot \left(\frac{1}{1+\frac{t^2}{v}} \right)^{\frac{v+1}{2}} \cdot \int_0^{\infty} y^{\frac{v+1}{2}-1} \cdot e^{-y} dy = \end{aligned}$$

$$= \frac{\Gamma(\frac{v+1}{2})}{\sqrt{2\pi v} \cdot \Gamma(\frac{v}{2})} \cdot \left(1 + \frac{t^2}{v} \right)^{-\frac{v+1}{2}}$$

con $t \in \mathbb{R}$

$$E(T_V) = 0$$

$$\sigma^2 = \frac{v}{v-2}$$



"code" sono + "doppi", \rightarrow + lentamente
a 0 delle normale.

H

V.A. F DI FISHER

Sia $U \sim \chi^2_{n_1}$. Sia $V \sim \chi^2_{n_2}$ e U e V indipendenti.

Allora:

$F \in \mathbb{R}^+$

$$F = \frac{U}{V} \cdot \frac{V_2}{V_1} \sim F_{V_1, V_2}$$

$$e \quad F = \frac{V}{U} \cdot \frac{V_1}{V_2} \sim F_{V_2, V_1}$$

$$f(u, v) = f_{X_{V_1}}^2(u) \cdot f_{X_{V_2}}^2(v) = \frac{U^{\frac{V_1}{2}-1} \cdot e^{-\frac{U}{2}}}{\Gamma(\frac{V_1}{2}) \cdot 2^{\frac{V_1}{2}}} \cdot \frac{V^{\frac{V_2}{2}-1} \cdot e^{-\frac{V}{2}}}{\Gamma(\frac{V_2}{2}) \cdot 2^{\frac{V_2}{2}}} =$$

$$= \frac{U^{\frac{V_1}{2}-1} \cdot V^{\frac{V_2}{2}-1} \cdot e^{-\frac{U+V}{2}}}{\Gamma(\frac{V_1}{2}) \cdot \Gamma(\frac{V_2}{2}) \cdot 2^{\frac{V_1+V_2}{2}}} \xrightarrow{\text{Pongo } F \rightarrow U=VF \rightarrow dU=V \cdot dF}$$

$$f(F, V) dV dF = \frac{1}{K} \cdot V^{\frac{V_2}{2}-1} \cdot \left(V \cdot F \cdot \frac{V_1}{V_2} \right)^{\frac{V_1}{2}-1} \cdot e^{-\frac{V}{2} \left(1 + F \cdot \frac{V_1}{V_2} \right)} \cdot V^{\frac{V_1}{2}} dV dF$$

$$= \frac{1}{K} \cdot \left(F \cdot \frac{V_1}{V_2} \right)^{\frac{V_1}{2}-1} \cdot V^{\frac{V_1+V_2}{2}-1} \cdot e^{-\frac{V}{2} \left(1 + F \cdot \frac{V_1}{V_2} \right)} \cdot \frac{V_1}{V_2} dV dF$$

$$f(F) = \int_0^\infty f(F, V) dV = \frac{1}{K} \cdot F^{\frac{V_1}{2}-1} \cdot \left(\frac{V_1}{V_2} \right)^{\frac{V_1}{2}} \cdot \int_0^\infty V^{\frac{V_1+V_2}{2}-1} \cdot e^{-\frac{V}{2} \left(1 + F \cdot \frac{V_1}{V_2} \right)} dV =$$

$$\text{Pongo } y = \frac{V}{2} \left(1 + F \cdot \frac{V_1}{V_2} \right) \rightarrow V = \frac{2y}{1 + F \cdot \frac{V_1}{V_2}} \rightarrow dV = \frac{2}{1 + F \cdot \frac{V_1}{V_2}} dy$$

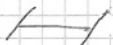
$$\text{Quindi } f(F) = \frac{1}{K} \cdot F^{\frac{V_1}{2}-1} \cdot \left(\frac{V_1}{V_2} \right)^{\frac{V_1}{2}} \cdot \int_0^\infty \left(\frac{2y}{1 + F \cdot \frac{V_1}{V_2}} \right)^{\frac{V_1+V_2}{2}-1} \cdot e^{-y} \cdot \frac{2}{1 + F \cdot \frac{V_1}{V_2}} dy =$$

$$= \frac{1}{K} \cdot F^{\frac{V_1}{2}-1} \cdot 2^{\frac{V_1+V_2}{2}} \cdot \left(1 + F \cdot \frac{V_1}{V_2} \right)^{-\frac{V_1+V_2}{2}} \cdot \int_0^\infty y^{\frac{V_1+V_2}{2}-1} \cdot e^{-y} dy = \Gamma\left(\frac{V_1+V_2}{2}\right)$$

$$= \frac{2^{\frac{V_1+V_2}{2}}}{\Gamma(\frac{V_1}{2}) \cdot \Gamma(\frac{V_2}{2}) \cdot 2^{\frac{V_1+V_2}{2}}} \cdot F^{\frac{V_1}{2}-1} \cdot \left(1 + F \cdot \frac{V_1}{V_2} \right)^{-\frac{V_1+V_2}{2}} \cdot \Gamma\left(\frac{V_1+V_2}{2}\right) \text{ Quindi}$$

$$f(F) = \frac{\Gamma\left(\frac{V_1+V_2}{2}\right)}{\Gamma(\frac{V_1}{2}) \cdot \Gamma(\frac{V_2}{2})} \cdot \left(\frac{V_1}{V_2} \right)^{\frac{V_1}{2}} \cdot F^{\frac{V_1}{2}-1} \cdot \left(1 + F \cdot \frac{V_1}{V_2} \right)^{-\frac{V_1+V_2}{2}} \quad F \in \mathbb{R}^+$$

(36)



BINOMIALE: $P_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$

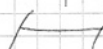
$n \rightarrow \infty \quad P_X(x) \rightarrow N(np, npq)$

la normale è il lim a cui tende la binom. al ↑ della prob.

$X_V^2 \sim \frac{(X_V^2)^{\frac{V}{2}-1} \cdot e^{-\frac{X_V^2}{2}}}{\Gamma(\frac{V}{2}) \cdot 2^{\frac{V}{2}}}$ Per $V \rightarrow \infty \quad X_V^2 \rightarrow N(V, 2V)$

$T_V = \frac{Z}{\sqrt{1/V}}$ Per $V \rightarrow \infty \quad T_V \rightarrow N(0, \frac{V}{V-2})$

la N è il limite di questa distrib. di prob.

 (fine calc. prob.)

INFERENZA STATISTICA PARAMETRICA

Dato un collettivo, se analizzo fenomeno nel coll. faccio un censimento. Pongo però esaminare un sottoinsieme, le campione e poi i risultati li ESTENDO alla pop. con un grado di INCERTEZZA espresso dalla probabilità.



Ex: Voglio indagare sugli \bar{E} l'attesa di studente 22/10/02

in aula. Potrei fare censimento, ma voglio campione casuale estruendo con sostituzione. Numero ogni stud. estraggo, rimetto dentro ed estraggo $\times 4$ volte. ($n=4$)
Prima dell'extr. ho $\underline{X} = (x_1, x_2, x_3, x_4)$, evento aleatorio, extr. casuale. Pongo a $\underline{x} = (x_1, x_2, x_3, x_4)$ [vettore di numeri, mentre prima era vett. var. aleatorie]. Variabile che sono i.i.d (ipotesi a monte). Convi. quindi estrazione Bernoulliana.

\underline{X} = campione.

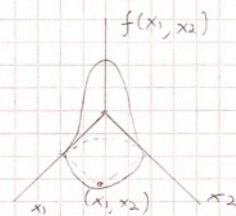
$$f(x_1, \dots, x_m) = f(x_1) f(x_2) \dots f(x_m) = \prod_{i=1}^m f(x_i, \theta)$$

↓
D.M.S. congiunta
di m var. al.

↳ f marginale della variabile

Il campione è un punto $\in \mathbb{R}^m$
e f può essere "campione"

Spazio dei campioni $\mathcal{X} = \{(x_1, x_2, \dots, x_m)\}$
e su questo si estende $f(x_1, x_2, \dots, x_m)$



[Prendendo le x_i i.i.d. allora fattorizziamo la f]
La forma della f è la stessa (non il valore)

Qual è la popolazione?

La pop. si identifica con il modello statistico parametrico:

$$= \{f(x, \theta) / \theta \in \bar{\Theta}\} \rightarrow \text{spazio dei parametri}$$

Ex: Poisson $f(x, \lambda) = \frac{e^{-\lambda} \lambda^x}{x!}$ $\lambda \in \mathbb{R}^+$ e $\bar{\Theta} = \mathbb{R}^+$

Binomiale: $f(x, \theta) = \binom{n}{x} \theta^x (1-\theta)^{n-x}$ $\bar{\Theta} = [0, 1]$

Normale: $f(x, \mu, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left\{-\frac{1}{2} \left(\frac{x-\mu}{\sigma}\right)^2\right\}$ $\bar{\Theta} = \mathbb{R} \times \mathbb{R}^+$
(μ) (σ^2)

Problema è individuare il PARAMETRO. Variab. x che estraggo con certa $f(x)$ la μ varia.

Ex: ditta che produce $N=100,000$ lampadine con probab. che il % di pezzi difettosi (ex: fulgurano entro 50 ore)
↳ incognita.

Estraggo $M=50$ lampadine $x = (0, 1, \dots, 0)$ [0: difettosa,

38 1: buona]. $\mathcal{X} = (x_1, x_2, \dots, x_{50})$

Con estrazione Bernoulliana.

Supponiamo che $y = 4$ difettore.

$P \in [0, \frac{1}{100000}, \frac{2}{100000}, \frac{3}{100000}, \dots, 1]$, ma incognita.

Allora $\hat{p} = \frac{4}{50}$ (stimata). Come diventa questa la popolazione?

$y = (\text{n. lamp. difettore}) = 0, 1, 2, \dots, 50$ (n) n configura come var. al. binomiale $if(x, n) = \binom{n}{x} P^x (1-P)^{n-x}$

$$\rightarrow f(y, P) = \binom{50}{y} P^y (1-P)^{50-y}$$

Nel campione $y = 4 \rightarrow f(4, P) = \binom{50}{4} P^4 (1-P)^{50-4}$

Allora $f(4, P)$ è la f. di prob. osservata calcolabile solo se conosco P . Quindi popolazione diventa

$$f(4, P) = \binom{50}{4} P^4 (1-P)^4; P \in [0, 1]$$

Problema è P delle 100.000 lamp., fatta l'estrazione non passate a var. al. y con distr. prob. $f(y, P)$.

Con campionamento ho il prob. matematico di determinare la P di $f(y, P)$ che diventa il modello statistico parametrico, cioè la popolazione.

L'estrazione casuale del campione genera la var. aleatoria con mea $f(x, \theta)$; inferenza parametrica.

Come impostare prob. di inf. parametrica?

PRESUPPOSTO:

"Dato un modello stat. parametrico $\mathcal{F} = \{f(x, \theta) | \theta \in \Theta\}$,
date n X_i iid $\sim f(x, \theta)$. Voglio le θ vero
che ha generato le vett. di variabili aleatorie X_i ."

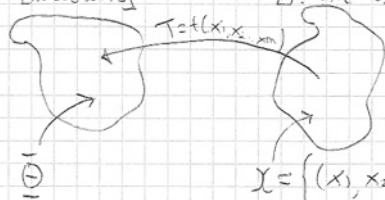
2 modi:

- STILY

- PROVA DI IPOTESI

[CUCCHIAIA]

[D. CHE OSSERVO]



$X = \{(x_1, x_2, \dots, x_m)\} =$ spazio dei campioni
(a m dimensioni)

STIMA: "ci" deve trovare $\hat{T} = t(x_1, x_2, \dots, x_m)$ [funz. del campione] scegliibile x rappresentare il param.
 T è solo del campione e NON è Θ , e' chiamata FUNZIONE STATISTICA.

\exists + T sufficiente x stessa stima:

$$T_1 = t_1(x_1, x_2, \dots, x_m)$$

$$T_2 = t_2(x_1, x_2, \dots, x_m)$$

$$T_k = t_k(x_1, x_2, \dots, x_m)$$

Tra le varie T_k scelgo lo STIMATORE: $\hat{\Theta} = t(X_1, x_2, \dots, x_m)$

$\hat{\Theta} = t(x_1, x_2, \dots, x_m) =$ stima (valore che assume) (vett. var. aleatorie)
 lo stimatore nel campione

È applicazione da X a Θ , operazione di stima.

Da prima: $\hat{p} = \frac{4}{50} = \bar{x} = \frac{\sum x_i}{n}$; dal punto $(1, 1, 1, \dots, 0, 0)$ in \underline{X}

X mi ha preso $\frac{4}{50}$ in $\underline{\Theta}$.

Come trovare la forma dello STIMATORE?

$\Theta = \hat{\Theta}$ e $f(x; \Theta) = f(x; \hat{\Theta})$ [operaz. di stima] \rightarrow

STIMA PUNTUALE

\exists \Leftrightarrow metodi di stima:

- METODO DEI MOMENTI

- " DELLA MASSIMA VEROSIMILITUDINE

METODO DEI MOMENTI

Cerca $\theta = h(x_1, x_2, \dots, x_n) \in \mathbb{R}$

MOMENTI TEORICI

da distribuzione probab.

$$\mu_r = \int x^r f_x(x, \theta) dx \quad \text{con } r=1, 2, \dots$$

dal modello.

MOMENTI EMPIRICI

Calcolati nel campione

$$\bar{x}_r = \sum_{i=1}^n \frac{x_i^r}{n}$$

Sempre stessa forma

il modello statistico

P.D.F. "converte" nel particolare l'uguaglianza tra
n mom. teorici e n mom. empirici

$$\begin{cases} \mu_1(\theta) = \bar{x}_1 \\ \mu_2(\theta) = \bar{x}_2 \\ \mu_n(\theta) = \bar{x}_n \end{cases}$$

k = n elementi che costituiscono
il vettore dei param.

se ho $\tilde{T} = \{N(\mu, \sigma_x^2) : \mu \in \mathbb{R}, \sigma_x^2 \in \mathbb{R}^+\}$

postulo $\begin{cases} \mu_1(\mu, \sigma_x^2) = \bar{x}_1 \\ \mu_2(\mu, \sigma_x^2) = \bar{x}_2 \end{cases}$

$$\begin{cases} \mu_1 = \bar{x}_1 \\ \mu_2 = \bar{x}_2 \end{cases} = \begin{cases} \sum_{i=1}^n \frac{x_i}{n} = \mu \\ \sum_{i=1}^n \frac{x_i^2}{n} = \mu_2 \end{cases}$$

Allora dato che $\mu_1 = \mu = \bar{x}$
e $\sigma_x^2 = \mu_2 - \mu^2$ proprio:

$$\mu_1 = \mu = \bar{x}$$

$$\sigma_x^2 + \mu^2 = \bar{x}_2 \rightarrow \sigma_x^2 = \bar{x}_2 - \mu^2 = \bar{x}_2 - \bar{x}^2$$

Stima della Varianza e' pari a $\hat{\sigma}_x^2 = \sum_{i=1}^n \frac{x_i^2}{n} - \left(\sum_{i=1}^n \frac{x_i}{n} \right)^2 =$

$$= \left(\frac{\sum x_i}{n} \right)^2 = \sum_{i=1}^n \frac{x_i^2}{n} - \frac{m \bar{x}^2}{n} = \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n} = \sum_{i=1}^n (x_i^2 + \bar{x}^2 - 2\bar{x}x_i) =$$

$$= \sum_{i=1}^n x_i^2 + m \bar{x}^2 - 2\bar{x} \left(\sum_{i=1}^n x_i \right) = \sum_{i=1}^n x_i^2 + m \bar{x}^2 - 2m \bar{x}^2 =$$

$$= \sum_{i=1}^n x_i^2 - m \bar{x}^2$$

Quinnoli:

$$\hat{\mu} = \bar{x}$$

$$\hat{\sigma}_x^2 = \sum_{i=1}^n (x_i - \bar{x})^2 / n$$

Conv.: $f_x(x, \alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} \cdot e^{-\frac{x}{\beta}} \quad x \geq 0, \alpha, \beta \geq 0$

$$\begin{cases} \mu_1(\alpha, \beta) = \bar{x}_1 \\ \mu_2(\alpha, \beta) = \bar{x}_2 \end{cases} \quad \left(\begin{aligned} \mu_n &= \int_0^\infty x^n \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} \cdot e^{-\frac{x}{\beta}} dx \\ &= \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty x^{\alpha+n-1} e^{-\frac{x}{\beta}} dx \end{aligned} \right)$$

Pongo $y = \frac{x}{\beta} \rightarrow x = y\beta \rightarrow dx = \beta dy$, quindi:

$$= \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty (y\beta)^{\alpha-1} \cdot e^{-y} \cdot \beta dy$$

$$\mu_n = \frac{\beta^{\alpha+n}}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty y^{\alpha+n-1} \cdot e^{-y} \cdot dy = \frac{\Gamma(\alpha+n)}{\Gamma(\alpha)} \beta^n = \prod_{i=1}^n (\alpha+i-1) \beta^n$$

$$\mathcal{F} = \{f_x(x, \alpha, \beta) = G(\alpha, \beta) \mid \alpha, \beta \in \mathbb{R}^+\}$$

23/10/02

$$E(x^n) = \mu_n(\alpha, \beta) = \beta^n \prod_{i=1}^n (\alpha+i-1), n \geq 1 \rightarrow \mu_1(\alpha, \beta) = \beta \alpha$$

$$n=2 \rightarrow \mu_2(\alpha, \beta) = (\alpha+1)\alpha$$

$$\begin{cases} \alpha\beta = \bar{x}_1 = \bar{x} \\ \alpha(\alpha+1)\beta^2 = \bar{x}_2 \end{cases} \quad \begin{array}{l} \text{Risolvere} \\ \text{rim. } \alpha \\ \text{da } \beta \end{array} \quad \begin{cases} \beta = \frac{\bar{x}}{\alpha} \\ (\alpha(\alpha+1)) \frac{\bar{x}^2}{\alpha^2} = \bar{x}_2 \rightarrow (\alpha+1)\bar{x}^2 = \alpha\bar{x}_2 \end{cases}$$

$$\textcircled{42} \quad \alpha\bar{x}^2 + \bar{x}^2 = \alpha\bar{x}_2 \rightarrow \alpha(\bar{x}_2 - \bar{x}^2) = \bar{x}^2 \rightarrow \alpha = \frac{\bar{x}^2}{\bar{x}_2 - \bar{x}^2} = h(x_1, x_2)$$

$$\hat{\alpha} = \frac{\sum_{i=1}^n x_i^2}{\sum_{i=1}^n x_i} \quad \text{Minima della Varianza} = \hat{\alpha}_x^2 = \sum_{i=1}^n \frac{x_i^2}{n} \quad \bar{x}^2 = \sum_{i=1}^n \frac{x_i^2 - \bar{x}^2}{n} =$$

$$= \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n} \quad \text{Risolto risp. a } \beta:$$

$$\hat{\beta} = S_x^2 / \bar{x}$$

107

METODO DELLA MASSIMA VEROSIMILIANZA

(Intersezione)

"Otteniamo lo stimatore di max. verosim."

È lo stimatore che massimizza la funzione di verosimiglianza detta LIKELIHOOD FUNCTION $= L(\theta; x_1, x_2, \dots, x_n) = L(\theta | x)$ (L di likelihood il mio campione)

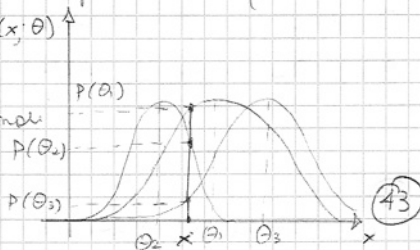
$$f_x(x_1, x_2, \dots, x_n; \theta) = \prod_{i=1}^n f_x(x_i; \theta) \quad \text{definita}$$

nello spazio dei campioni $\mathcal{X} = \{(x_1, x_2, \dots, x_i, \dots, x_n)\}$

$$L(\theta, x_1, x_2, \dots, x_i, \dots, x_n) = \prod_{i=1}^n f_x(x_i, \theta) \rightarrow \hat{\theta}$$

(ridefinire quella di prima). È definita invece nel SPAZIO DEI PARAMETRI, e prima ma var. è θ Opposto a defniz. di prob. Come varia la prob. del campione al variare di θ . Non è fun. di probabilità (θ non è una var. aleatoria)

Prendere x , al variare di θ varia la curva e quindi la P del campione



Ex: Bernoulli

$$f_x(x; \theta) = \theta^x (1-\theta)^{1-x} \quad x=0,1$$

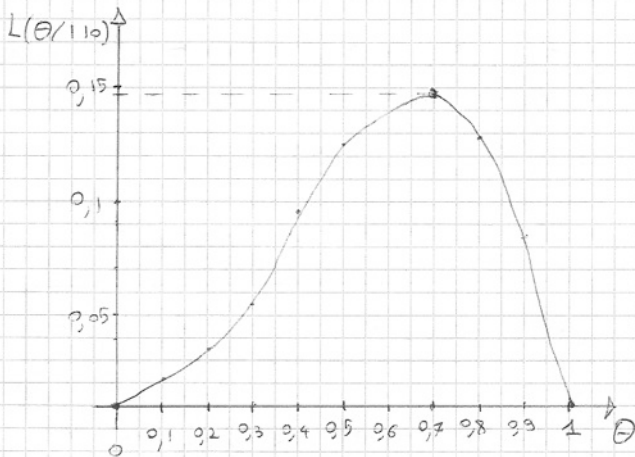
$$\mathcal{F} = \{f_x(x; \theta) = \theta^x (1-\theta) / \theta \in [0,1]\}$$

Con $n=3$, $\underline{x} = (x_1, x_2, x_3)$. Estimatione e ho $\underline{x} = (1, 1, 0)$

Ho che $f_x(1; \theta) = \theta$ e $f_x(0; \theta) = 1 - \theta$. Facio variare θ .

θ	$L(\theta/110)$
0	0
1/10	0,009
2/10	0,032
3/10	0,063
4/10	0,096
5/10	0,125
6/10	0,144
7/10	0,147
8/10	0,128
9/10	0,081
1	0

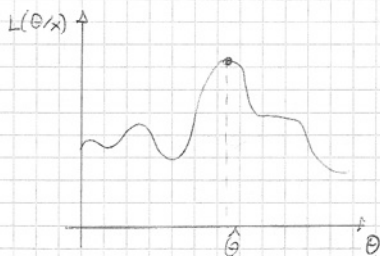
$$L(\theta/110) = f_x(1; \theta) \cdot f_x(1; \theta) \cdot f_x(0; \theta) = \theta^2 (1-\theta)$$



$$\hat{\theta} = \underset{\theta \in [0,1]}{\arg \max} L(\theta/x)$$

Qual è il θ + verosimile?

Dalle ho $\max L \rightarrow$



Con θ scalare. Per trovare $\hat{\theta}$, dato $L(\theta; x_1, x_2, \dots, x_m)$
 devo fare $\frac{d}{d\theta} L(\theta; x_1, x_2, \dots, x_m) = 0$. $\} \Leftrightarrow$ punti di
 min locale, max locale, flessi. Allora devo verificare
 che $\frac{d^2}{d\theta^2} L(\theta; x_1, x_2, \dots, x_m) \Big|_{\theta = \hat{\theta}} = 0$ (max globale)

Con θ vettoriale. $\theta = [\theta_1, \theta_2, \dots, \theta_i, \dots, \theta_n]$. Dev.

$$\begin{cases} \frac{\partial}{\partial \theta_1} L(\theta/x) = 0 \\ \frac{\partial}{\partial \theta_2} L(\theta/x) = 0 \\ \vdots \\ \frac{\partial}{\partial \theta_n} L(\theta/x) = 0 \end{cases} \Rightarrow \hat{\theta} = [\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_n], \text{ punto} \\ \text{stationario. Dobbiamo calcolare} \\ \text{e' Hessiana.}$$

- Poisson:

$$f_x(x, \lambda) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}; \quad \mathcal{F} = \left\{ f_x(x, \lambda) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}, \lambda \in \mathbb{R}^+ \right\}$$

Con m x_i i.i.d. $\sim f_x(x, \lambda)$. Soluz:

$$\begin{aligned} L(\lambda, x_1, x_2, \dots, x_m) &= \frac{e^{-\lambda} \lambda^{x_1}}{x_1!} \cdot \frac{e^{-\lambda} \lambda^{x_2}}{x_2!} \dots \frac{e^{-\lambda} \lambda^{x_m}}{x_m!} = \\ &= \frac{e^{-m\lambda} \lambda^{\sum_{i=1}^m x_i}}{\prod_{i=1}^m x_i!} \end{aligned}$$

Lineare rispetto con la trasformazione
 logaritmica (che e' monotona) essendo
 una f. exp.

$$\log L(\lambda, x_1, x_2, \dots, x_m) = -m\lambda + \sum_{i=1}^m x_i \log \lambda - \log \prod_{i=1}^m x_i!$$

$$\left\{ \frac{d}{d\lambda} \log L(\lambda/\underline{x}) = -n + \sum_{i=1}^n \frac{x_i}{\lambda} = 0 \rightarrow \text{La 1' eq. si chiama sempre} \right.$$

Equazione di verosimiglianza

$$\left\{ \frac{d^2}{d\lambda^2} \log L(\lambda/\underline{x}) = - \sum_{i=1}^n \frac{x_i}{\lambda^2} = -n\bar{x} \right.$$

$$-n\lambda + \sum_{i=1}^n x_i = 0 \Rightarrow \boxed{\hat{\lambda} = \sum_{i=1}^n \frac{x_i}{n} = \bar{x}}$$

$$\left. \frac{d^2}{d\lambda^2} \log L(\lambda/\underline{x}) \right|_{\lambda=\bar{x}} = - \frac{n\bar{x}}{\bar{x}^2} = - \frac{n}{\bar{x}} < 0 \quad \left(\begin{array}{l} \text{La Poisson ha} \\ x=0,1,2,\dots,\infty \end{array} \right)$$

- Normale

$$f_i(x; \mu) = \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2}(x_i - \mu)^2 \right\} \sim N(\mu, \sigma^2)$$

\downarrow

$$\mathcal{F} = \left\{ f_x(x, \mu) = N(\mu, 1) / \mu \in \mathbb{R} \right\}$$

Varianza nota
ma non nota
di generalità rimane
per tutti i valori

$$L(\mu/\underline{x}) = \left(\frac{1}{\sqrt{2\pi}} \right)^n \left[e^{-\frac{1}{2}(x_1 - \mu)^2} \cdot e^{-\frac{1}{2}(x_2 - \mu)^2} \cdot \dots \cdot e^{-\frac{1}{2}(x_n - \mu)^2} \right] =$$

$$= (2\pi)^{-\frac{n}{2}} \cdot e^{-\frac{1}{2} \sum_{i=1}^n (x_i - \mu)^2}$$

Anche qui si massimizza il log:

$$\log L(\mu/\underline{x}) = -\frac{nn}{2} \log 2\pi - \frac{1}{2} \sum_{i=1}^n (x_i - \mu)^2$$

$$\left\{ \frac{d}{d\mu} \log L(\mu/\underline{x}) = 0 \right.$$

$$\left. \frac{d^2}{d\mu^2} \log L(\mu/\underline{x}) \right|_{\mu=\hat{\mu}} < 0$$

$$\frac{d}{d\mu} \log L(\mu/\mathcal{X}) = \frac{1}{n} \sum_{i=1}^n (x_i - \mu) \cdot (-1) = 0 = \sum_{i=1}^n x_i - n\mu = 0$$

$$\boxed{\hat{\mu} = \sum_{i=1}^n x_i / n = \bar{x}} \quad ; \quad \text{essendo } \frac{d^2}{d\mu^2} < 0 \Rightarrow -n \text{ e' ok.}$$

Problema nota, σ_x^2 incognita:

$$f_x(x, \sigma_x^2) = \frac{1}{\sigma_x \sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma_x^2} x^2\right\} \sim N(0, \sigma_x^2)$$

$$\mathcal{F} = \left\{ f_x(x, \sigma_x^2) = N(0, \sigma_x^2) / \sigma_x^2 \in \mathbb{R}^+ \right\} \quad \begin{array}{l} \text{+ } \mu=0 \text{ ma} \\ \text{non perche} \\ \text{di simmetria} \end{array}$$

$$L(\sigma_x^2 / \mathcal{X}) = (2\pi \sigma_x^2)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma_x^2} \sum_{i=1}^n x_i^2\right\}$$

$$\log L(\sigma_x^2 / \mathcal{X}) = \left(-\frac{n}{2}\right) \log 2\pi + \frac{n}{2} \log \sigma_x^2 - \frac{1}{2\sigma_x^2} \sum_{i=1}^n x_i^2$$

$$\frac{d}{d(\sigma_x^2)} \log L(\sigma_x^2 / \mathcal{X}) = -\frac{n}{2\sigma_x^2} + \frac{\sum_{i=1}^n x_i^2}{2\sigma_x^4} = 0$$

$$= \left[\frac{n}{2} (\sigma_x^2)^{-1} + \frac{\sum_{i=1}^n x_i^2}{2} (\sigma_x^2)^{-2} \right] \rightarrow \text{derivata ancora!}$$

$$= \frac{n}{2} (\sigma_x^2)^{-2} - \sum_{i=1}^n x_i^2 (\sigma_x^2)^{-3} = \frac{n}{2\sigma_x^4} - \frac{\sum_{i=1}^n x_i^2}{\sigma_x^6}$$

$$\left\{ -\frac{n}{\sigma_x^4} + \frac{\sum_{i=1}^n x_i^2}{\sigma_x^6} = 0 \Rightarrow \boxed{\hat{\sigma}_x^2 = \frac{\sum_{i=1}^n x_i^2}{n}} \right.$$

$$\left. \left\{ \frac{d^2}{d(\sigma_x^2)^2} \log L(\sigma_x^2 / \mathcal{X}) \right\} \right|_{\hat{\sigma}_x^2 = \frac{\sum_{i=1}^n x_i^2}{n}} = \frac{n}{2\sigma_x^6} - \frac{n}{\sigma_x^6} = \frac{n-2n}{2\sigma_x^6} < 0 \Rightarrow \text{ok!}$$

Conv. @ vettoriale:

$$\mathcal{F} = \left\{ f_x(x, \mu, \sigma_x^2) = N(\mu, \sigma_x^2) / (\mu, \sigma_x^2) \in \mathbb{R} \times \mathbb{R}^+ \right\}$$

$$n \text{ i.i.d. } \sim N(\mu, \sigma_x^2)$$

$$L(\mu, \sigma_x^2 / \underline{x}) = \frac{1}{\sigma_x \sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma_x^2}(x_1 - \mu)^2\right\} \cdots \frac{1}{\sigma_x \sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma_x^2}(x_n - \mu)^2\right\} = (2\pi \sigma_x^2)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma_x^2} \sum_{i=1}^n (x_i - \mu)^2\right\}$$

$$\log L(\mu, \sigma_x^2 / \underline{x}) = -\frac{n}{2} \log 2\pi - \frac{n}{2} \log \sigma_x^2 - \frac{1}{2\sigma_x^2} \sum_{i=1}^n (x_i - \mu)^2$$

$$\begin{cases} \frac{\partial}{\partial \mu} \log L(\mu, \sigma_x^2 / \underline{x}) = 0 \\ \frac{\partial}{\partial \sigma_x^2} \log L(\mu, \sigma_x^2 / \underline{x}) = 0 \end{cases} = \begin{cases} \frac{\sum_{i=1}^n (x_i - \mu)}{\sigma_x^2} = 0 \\ -\frac{n}{2\sigma_x^2} + \frac{\sum_{i=1}^n (x_i - \mu)^2}{2\sigma_x^4} = 0 \end{cases}$$

$$\sum_{i=1}^n x_i - n\mu = 0 \rightarrow \boxed{\hat{\mu} = \frac{\sum_{i=1}^n x_i}{n} = \bar{x}}$$

$$-n\sigma_x^2 + \frac{\sum_{i=1}^n (x_i - \mu)^2}{\sigma_x^4} = 0 \rightarrow \hat{\sigma}_x^2 = \boxed{s_x^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}$$

Punto Massimo. Vero e' l'alternativa (det. ∂^2 e ∂^2 mista):

$$H(\mu, \sigma_x^2) = \begin{vmatrix} \frac{\partial^2}{\partial \mu^2} \log L & \frac{\partial^2}{\partial \mu \partial \sigma_x^2} \log L \\ \frac{\partial^2}{\partial \sigma_x^2 \partial \mu} \log L & \frac{\partial^2}{(\sigma_x^2)^2} \log L \end{vmatrix}_{\bar{x}, \hat{s}_x}$$

- Se il det. > 0 , abbiamo 1 max e min relativi
- Se il det. < 0 , " punto di sella
- Se il det. $= 0$, e' indefinito, non possiamo dire nulla.

(48) Se $H_{1,1} > 0$ e' min relativo, se < 0 e' max relativo

$$\frac{\partial^2}{\partial \mu^2} \log L = - \frac{n}{\sigma_x^2}$$

$$\frac{\partial^2}{\partial \sigma_x^2} \log L = \frac{n}{2\sigma_x^4} - \frac{\sum_{i=1}^n (x_i - \mu)^2}{\sigma_x^6}$$

$$\begin{aligned} \frac{\partial^2}{\partial \mu \partial \sigma_x^2} \log L &= \frac{\partial}{\partial \mu} \cdot \frac{\partial}{\partial \sigma_x^2} \log L = \frac{\partial}{\partial \mu} \left[- \frac{n}{2\sigma_x^2} + \frac{\sum_{i=1}^n (x_i - \mu)^2}{2\sigma_x^4} \right] = \\ &= - \frac{\sum_{i=1}^n (x_i - \mu)}{\sigma_x^4} \end{aligned}$$

Le calcolo in \bar{x}, s_x^2 :

$$H(\mu, \sigma_x^2) = \begin{vmatrix} -\frac{n}{\sigma_x^2} & -\frac{\sum_{i=1}^n (x_i - \bar{x})}{\sigma_x^4} = -\frac{n\bar{x} - n\bar{x}}{\sigma_x^4} = 0 \\ 0 & \frac{n}{2\sigma_x^4} - \frac{n s_x^2}{\sigma_x^6} = -\frac{n}{2\sigma_x^4} \end{vmatrix} =$$

$$= \frac{n^2}{2 \cdot \sigma_x^6} > 0; \text{essendo } -\frac{n}{\sigma_x^2} < 0, (\bar{x}, s_x^2) \text{ è max relativo,}$$

ma essendo l'unica sol. è anche ASSOLUTO \Rightarrow trovato gli stimevoli

Trovato $\theta = \hat{\theta}$, calcolo $f_x(x, \theta) = f_x(x, \hat{\theta})$

Quale metodo applichiamo, a valle PROPRIETÀ degli stimevoli:

1. P. della CORRETTA

2. P. // CONSISTENZA

3. P. // EFFICIENZA

Dato sempre il modello \mathcal{F} :

1) $\hat{\theta}$ è corretto, se $E(\hat{\theta}) = \theta \quad \forall \theta \in \bar{\Theta}$
(o "centrato")

$\hat{\theta}$ è var. aleatoria con $\hat{\theta} \sim g(\hat{\theta}, \theta)$

Se $\hat{\theta} \in \mathbb{R}$, $f(\hat{\theta}) = \int_{\mathbb{R}} \hat{\theta} g(\hat{\theta}, \theta) d\hat{\theta} \quad (\text{cont.})$

Se $\hat{\theta} \in D$, $f(\hat{\theta}) = \sum_{\theta \in D} \hat{\theta} g(\hat{\theta}, \theta) \quad (\text{discret.})$

Se $E(\hat{\theta}) \neq \theta$ si chiamerà distorto e la quantità

$$b(\hat{\theta}) = E(\hat{\theta}) - \theta = \underline{\text{distorsione}}$$

↳ Bias

Tanto $\hat{\theta}$ lo conosco, θ non lo so (altrimenti non farei inferenza).

2) Corr. è prop. asintotica, e' prop. di una sequenza di stimatori:

$$\hat{\theta}_1 = h(x_1), \quad \hat{\theta}_2 = h(x_1, x_2), \quad \hat{\theta}_n = h(x_1, x_2, \dots, x_n)$$

[↑ la dimensione del campione]

ERRORE QUADRATICO MEDIO di $\hat{\theta}$

$$E[\hat{\theta} - \theta]^2 \quad \left\{ \begin{array}{l} \text{Se } \hat{\theta} \text{ è corretto} = \underline{\text{Var}(\hat{\theta})} = E[\hat{\theta} - E(\hat{\theta})]^2 \\ \text{Se } \hat{\theta} \text{ è distorto} = \underline{\text{Var}(\hat{\theta})} + b(\hat{\theta})^2 \end{array} \right.$$

$$\begin{aligned} E[\hat{\theta} - \theta]^2 &= E[\hat{\theta} - E(\hat{\theta}) + E(\hat{\theta}) - \theta]^2 = E[\hat{\theta} - E(\hat{\theta})]^2 + \\ &+ E[E(\hat{\theta}) - \theta]^2 + 2E[(\hat{\theta} - E(\hat{\theta}))][E(\hat{\theta}) - \theta] = \\ &= \text{Var}(\hat{\theta}) + b(\hat{\theta})^2 \end{aligned}$$

Consistenza $\begin{cases} \text{in "senso debole"} & a) \\ \text{in "media quadratica"} & b) \end{cases}$

b):
$$\lim_{n \rightarrow \infty} E[\hat{\theta}_n - \theta]^2 = 0$$

a):
$$\lim_{n \rightarrow \infty} P_n\left\{|\hat{\theta}_n - \theta| < \varepsilon\right\} = 1$$

ε piccolo a piacere
($\hat{\theta}_n \rightarrow \theta$ con $P_n \rightarrow 1$)

[b) \Rightarrow a)]

3) Efficienza $\begin{cases} \text{"assoluta"} & a) \\ \text{"relativa"} & b) \end{cases}$

a) si rif. alla σ_n^2

\rightarrow del $\hat{\theta}$. $< \sigma_n^2$,

$<$ oscillazione intorno

alla media

a) si riferisce solo alla classe dei $\hat{\theta}$ corretti
CRAUER-RAO hanno una. un limite inf. di $VAR(\hat{\theta})$
DISUGUAGLIANZA DI CRAUER-RAO

$$\sigma_{\hat{\theta}}^2 \geq \frac{1}{n E \left[\frac{\partial}{\partial \theta} \log f_n(x, \theta) \right]^2} \quad \mathbb{P} \left\{ f(x, \theta) / \theta \in \Theta \right\}$$

Veroe & Kramers

Se uguale, $\hat{\theta}$ è efficiente in senso assoluto

$E(x)$

Suppongo normale!

$$\mathcal{F} = \{f_x(x, \theta) = N(\mu, 1) \mid \mu \in \mathbb{R}\}$$

$n \times$ i.i.d. $N(\mu, 1)$

$$\hat{\mu} = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{n} (x_1 + x_2 + \dots + x_n)$$

(da \bar{x} e' anche normale) $\bar{x} \sim N(\mu, \frac{1}{n})$

$$\mathbb{E}(\bar{x}) = \frac{1}{n} \mathbb{E}\left[\sum_{i=1}^n x_i\right] = \frac{n\mu}{n} = \mu$$

$$\text{Var}(\bar{x}) = \mathbb{E}(\bar{x} - \mu)^2 = \mathbb{E}\left[\frac{x_1 + x_2 + \dots + x_n}{n} - \mu\right]^2 =$$

$$= \mathbb{E}\left[\frac{\sum x_i - n\mu}{n}\right]^2 = \frac{1}{n^2} \mathbb{E}\left[\sum_{i=1}^n (x_i - \mu)\right]^2 = \frac{1}{n^2} \mathbb{E}\left[\sum_{i=1}^n (x_i - \mu)^2 +$$

$$\sum_{i \neq j} \sum_{j=1}^n (x_i - \mu)(x_j - \mu)\right] = \frac{1}{n^2} \left[\sum_{i=1}^n \mathbb{E}(x_i - \mu)^2 +$$

$$+ \sum_{i=1}^n \sum_{j \neq i} \mathbb{E}\{(x_i - \mu)(x_j - \mu)\} \right] = \frac{1}{n^2} [n \sigma_x^2] = \frac{\sigma_x^2}{n}$$

$\rightarrow x$ con stessa varianza
 \rightarrow generico

\hookrightarrow COVARIANZA

Prendendo $n \times$ i.i.d. allora la COVAR(x) e' annulla

\bar{x} e' corretto. Quindi posso verif. efficienza:

$$f(x, \mu, \sigma_x^2) = \frac{1}{\sigma_x \sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma_x^2} (x - \mu)^2\right\}$$

$$\log f(x, \mu, \sigma_x^2) = -\frac{1}{2} \log 2\pi \sigma_x^2 - \frac{1}{2\sigma_x^2} (x - \mu)^2 \quad [\theta \text{ e' } \mu]$$

$$\mathbb{E}\left[\frac{\partial}{\partial \mu} \log f\right]^2 = \mathbb{E}\left[\frac{x - \mu}{\sigma_x^2}\right]^2 = \frac{1}{\sigma_x^4} \mathbb{E}(x - \mu)^2 = \frac{1}{\sigma_x^2}$$

$$\sigma_u^2 \geq \frac{1}{\sigma_{\hat{x}}^2} = \frac{\sigma_x^2}{n} \Rightarrow \bar{x} \text{ è efficiente in senso assoluto.}$$

b) Confronta gli errori quad. medi dei due $\hat{\theta}$.
Se ho $\hat{\theta}_1$ e $\hat{\theta}_2$ stimatori di θ ,

Se $\frac{E(\hat{\theta}_1 - \theta)^2}{E(\hat{\theta}_2 - \theta)^2} < 1$, $\hat{\theta}_1$ è più efficiente di $\hat{\theta}_2$

Eff. relativa non corr. se corretto o no. Confronta 2 stimatori.

Non ho attendibilità stima, allora procedo 25/10/67
con lei.

STIMA PER INTERVALLO (o INTERVALLO DI CONFIDENZA)

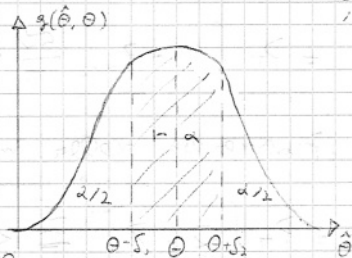
Trovare 2 statistiche $T_1 = t_1(x_1, x_2, \dots, x_n)$ e $T_2 = t_2(x_1, x_2, \dots, x_n)$
funz. del campione con $T_1 < T_2$ / (T_1, T_2) include θ con
certa probabilità (attento, θ non code in nessun intervallo,
è fisso)

$\hat{\theta} \sim g(\hat{\theta}, \theta)$ dove errore noto g .
Corr. dens. di probabilità

$$1 - \alpha = \Pr \{ \theta - s_1 \leq \hat{\theta} \leq \theta + s_2 \}$$

È interv. fisso ma non è quello
di confidenza.

Inverte gli estremi: $\Pr \{ \hat{\theta} - s_2 \leq \theta \leq \hat{\theta} + s_1 \}$
nono parato a int. variabile, oggi estremi lo



lo stimatore che è var. aleatoria. Int. varia a seconda del campione. $1-\alpha$ è quindi la prob. che l'int. include θ .

$$\text{Fix: } \mathcal{F} = \{N(\mu, \sigma^2) / \mu \in R\}$$

$\hat{\mu} = \bar{x}$ + stima puntuale che si distribuisce secondo:

$$\bar{x} \sim N(\mu, \frac{\sigma_x^2}{n}) = g(\bar{x}, \mu)$$

Standardizzazione:

$$Z = \frac{(\bar{x} - \mu)\sqrt{n}}{\sigma_x} \sim N(0, 1)$$

Devo cercare l'int. + piccolo
essendo simmetrica e intorno

a 0. $Z_{\alpha/2}$ e $-Z_{\alpha/2}$: PERCENTILI della normale.

$$\frac{\alpha}{2} = \int_{-\infty}^{-Z_{\alpha/2}} f(z) dz = \int_{Z_{\alpha/2}}^{\infty} f(z) dz$$

$$1 - \alpha = \Pr \left\{ -Z_{\alpha/2} \leq Z \leq Z_{\alpha/2} \right\} = \Pr \left\{ -Z_{\alpha/2} \leq \frac{(\bar{x} - \mu)\sqrt{n}}{\sigma_x} \leq Z_{\alpha/2} \right\} =$$

$$= \Pr \left\{ \underbrace{\mu - Z_{\alpha/2} \frac{\sigma_x}{\sqrt{n}}}_{\delta_1} \leq \bar{x} \leq \mu + \underbrace{Z_{\alpha/2} \frac{\sigma_x}{\sqrt{n}}}_{\delta_2} \right\} \quad \begin{array}{l} \text{Essendo sym,} \\ \delta_1 = \delta_2 \end{array}$$

1) Probabilità della media del campione. Inverso

$$= \Pr \left\{ \underbrace{\bar{x} - Z_{\alpha/2} \frac{\sigma_x}{\sqrt{n}}}_{\text{infer}} \leq \bar{x} \leq \bar{x} + \underbrace{Z_{\alpha/2} \frac{\sigma_x}{\sqrt{n}}}_{\text{super}} \right\} \rightarrow 2) \text{ prob. che}$$

intervallo include μ . $1-\alpha$ con 2 forme diverse

$1-\alpha$ è il LIVELLO di FIDUCIA.

PROVA DI IPOTESI

$$T = \{f_{\theta}(x, \theta) \mid \theta \in \bar{\Theta}\} \quad n \times i.i.d \sim f_{\theta}(x, \theta) \quad \forall i=1, \dots, n$$



$$X = \{(x_1, x_2, \dots, x_n)\}$$

"La p.d.f. è l'ipotesi che un certo $\theta = \theta_0$ oppure una certa regione di apparten. di θ [$\theta \in \theta_0 \subset \bar{\Theta}$] e verificare ^{con TEST} se compie o meno l'ipotesi."

Ipotesi: Congettura che faccia nel valore del parametro θ nulla θ_0 o di θ di un sottoinsieme di $\bar{\Theta}$

$H: \theta = \theta_0$ è ip. SEMPLICE

$H: \theta \in \theta_0 \subset \bar{\Theta}$ è "complessa"

Concludo un TEST: è una regola decisionale stabilita a priori

Posso PARTITIONARE lo spazio dei campioni, determinando una frontiera, in 2 regioni, 1 è consistente con H (ipotesi), l'altra no.



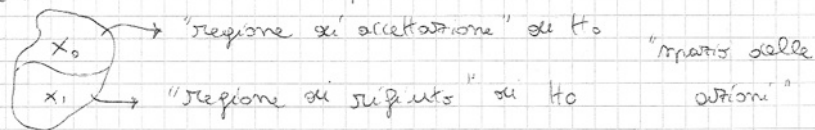
TEORIA DI NEYMAN-PEARSON

Parte da un sistema di ipotesi, compatto da:

$$\begin{cases} H_0: \theta = \theta_0: \text{ip. nulla [quella da provare]} \\ H_1: \theta = \theta_1: \text{" alternativa} \end{cases}$$

$$\text{Amble N.M.} \quad \left\{ \begin{array}{l} H_0: \theta \in \Theta_0, \quad \Theta_0 \cap \Theta_1 = \emptyset \\ \text{Completo} \quad H_1: \theta \in \Theta_1 \quad \Theta_0 \cup \Theta_1 = \bar{\Theta} \end{array} \right.$$

(probi nella semp. prova $H_1: \theta \neq \theta_0$ ed o' completo)



$$X_0 \cap X_1 = \emptyset; \quad X_0 \cup X_1 = X$$

Ogni test è individuato da partizione. \exists no partiz. della partiz!

Secondo la T.R., nel concludere il test posso commettere:

- 1) ERRORE PRIMA SPECIE
- 2) " SECONDA "

- 1) Rifiuto H_0 quando è VERA. Ex. cose in X_1 e quindi rifiuto H_0 .
- 2) Le punto cose in X_0 ma accetto un'ipotesi che è FALSA

Tabella:

IPOTESI VERE	IPOTESI ACCETTATE	
	H_0	H_1
H_0	✓	ERR I
H_1	ERR II	✓

$$\underline{X} = (X_1, X_2, \dots, X_n)$$

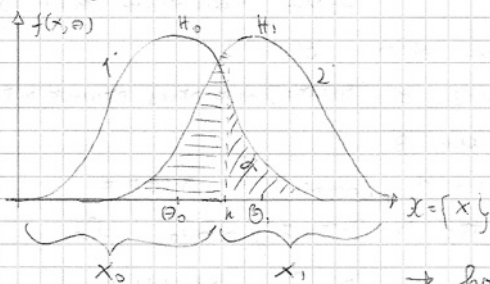
Assunto all'errore c'è una mia $P(\text{err})$

$$\alpha = \Pr \{ X \in X_1 / H_0 \}$$

$$\beta = \Pr \{ Z \in X_0 / H_1 \} \quad (\text{L'errore})$$

Devo scegliere il test migliore, devo cercare di avere $\alpha = \beta = 0$; è molto difficile. Cerco di minimizzarli ma è difficile farlo simultaneamente per entrambe.

$$\begin{cases} H_0: \theta = \theta_0 \\ H_1: \theta = \theta_1 \end{cases} \quad \theta_0 < \theta_1 \quad f_x(x, \theta_0) ; f_x(x, \theta_1)$$



La frontiera è il valore critico k .

$$X_0 = \{x / x \leq k\}$$

$$X_1 = \{x / x > k\}$$

→ ho det. un test.

Se vado in X_1 rispetto H_0 , quindi punto è 2° curva ma può essere anche alla 1°!

$$\alpha = \int_k^{\infty} f_x(x, \theta_0) dx$$

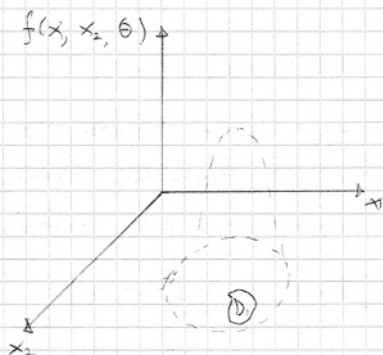
Analogamente se accetto H_0 e rispetto H_1 , ho:

$$\beta = \int_{-\infty}^k f_x(x, \theta_1) dx$$

Se voglio min α , porto $k \rightarrow \infty$ ma $\uparrow \beta$!

Allora:

Stabilisco a priori α , lo controllo. Poi, fra tutti i test con quel dato α , scelgo quello con β min.



$$\iint_{D_1} f(x_1, x_2, \theta_0) dx_1 dx_2 = \alpha$$

il complemento e' β :

$$\iint_{\bar{D}_1} f(x_1, x_2, \theta_1) dx_1 dx_2 = \beta$$

Posso individuare tanti D_1

\Leftrightarrow tanti test diversi, di dim. α .

[α = "LIVELLO DI SIGNIFICATIVITA'" e "DIFFUSIONE DELLA REGIONE DI RISCHIO"]

A ogni D_1 corrisponde un β .

$\alpha + \beta$ non e' detto che ≤ 1 , possono essere DIVERSE

Test con β min e' detto "TEST PIU' POTENTE", dove la

potenza del test e' data da $Pr\{x \in X_1 / H_1\} = 1 - \beta$

[rispetto H_0 quando e' vera, H_1 agisce correttamente]

$$Pr\{x \in X_1 / H_1\} + Pr\{x \in X_0 / H_1\} = 1$$

$$\min \beta = \max 1 - \beta.$$

la ricerca del test e' \Leftrightarrow se con. i p. sempre e' completa.

$$\text{LEMMA DI NEMHOUS-PEARSON} \begin{cases} H_0: \theta = \theta_0 \\ H_1: \theta = \theta_1 \end{cases}$$

RICERCA DEL TEST + POTENTE

$$\text{ROPPERO DI VERDUGO-LIARNA GENERALIZZATO} \begin{cases} H_0: \theta \in \Theta_0 \\ H_1: \theta \in \Theta_1 \end{cases}$$

Se ho campione di dim = 1, se $f_x(x, \theta_0)$ e' pramale
la 1 e + prob.

Scego le valore + probabile. Valor in n campione

$$\frac{L(\theta_0, x_1, x_2 \dots x_n)}{L(\theta_1, x_1, x_2 \dots x_n)} = \frac{\prod_{i=1}^n f_x(x, \theta_0)}{\prod_{i=1}^n f_x(x, \theta_1)} = \lambda$$

(n data
in max
verosimigl.)

Il lemma dice che X_1 [regione critica che max potenza
del test] e' $X_1 = \{ (x_1, x_2 \dots x_n) / \lambda = \frac{L(\theta_0(x))}{L(\theta_1(x))} \leq \lambda \alpha \}$

Dato $F = \{ f(x, \theta) / \theta \in \bar{\Theta} \}$ e nat. ip. semplice: 29/10/07

$$\begin{cases} H_0 / \theta = \theta_0, & \text{de } X_1 = \{ (x_1, x_2, \dots, x_n) / \lambda = \frac{L(\theta_0, x_1, x_2, \dots, x_n)}{L(\theta_1, x_1, x_2, \dots, x_n)} = \frac{\prod_{i=1}^n f(x, \theta_0)}{\prod_{i=1}^n f(x, \theta_1)} \leq \lambda \alpha \} \\ H_1 / \theta = \theta_1 \end{cases}$$

Test sta natio dei campione n-dim a quello
di 1 dim-dim. Le var. aleatoria con $f_1(\lambda)$

$\alpha = \int_0^{\lambda \alpha} g(\lambda, \theta_0) d\lambda$ x scet. e' estremo solo conoscere
 $g(\lambda, \theta_0)$ dietro la sint. di
prob. di il che non e' sempre facile \Rightarrow uso
di artificia. Si trova

$$Pn \{ x \in X_1 / H_1 \} = 1 - \beta = \max.$$

Exi. \uparrow
 $F = \{ N(\mu, 1) / \mu \in \mathbb{R} \}$. Estraggo n X_i iid $\sim N(\mu, 1)$

$$H_0: \mu=0 \quad L(\mu/x) = \left(\frac{1}{2\pi}\right)^{\frac{m}{2}} \cdot e^{-\frac{1}{2} \sum_{i=1}^m (x_i - \mu)^2}$$

$$H_1: \mu=2$$

$$\text{Calcolo } \lambda = \frac{L(0/x)}{L(2/x)} = \frac{e^{-\frac{1}{2} \sum_{i=1}^m x_i^2}}{e^{-\frac{1}{2} \sum_{i=1}^m (x_i - 2)^2}} = \frac{e^{-\frac{1}{2} \sum_{i=1}^m x_i^2}}{e^{-\frac{1}{2} \left[\sum_{i=1}^m x_i^2 + 4m - 4 \sum_{i=1}^m x_i \right]}} = e^{2m - 2m\bar{x}}$$

$$= \exp \left\{ -\frac{1}{2} \left[\sum_{i=1}^m x_i^2 - \sum_{i=1}^m x_i^2 - 4m + 4m\bar{x} \right] \right\} = e^{2m - 2m\bar{x}}$$

la regione critica è formata da:

$$K_1 = \{ (x_1, x_2, \dots, x_m) / \lambda = e^{2m - 2m\bar{x}} \leq \lambda \alpha \}$$

→ lo suppongo noto

Passo al log:

$$K_1 = \{ (x_1, x_2, \dots, x_m) / 2m - 2m\bar{x} \leq \log \lambda \alpha \} =$$

$$= \{ (x_1, x_2, \dots, x_m) / 2m\bar{x} \geq 2m - \log \lambda \alpha \} =$$

$$= \{ (x_1, x_2, \dots, x_m) / \bar{x} \geq \underbrace{1 - \frac{1}{2m} \log \lambda \alpha}_C \} =$$

C, costante che suppongo nota

$$= \{ (x_1, x_2, \dots, x_m) / \bar{x} \geq c \}$$

(Se $x \in K_1 / \bar{x} \geq c$ allora $\mu=2$)

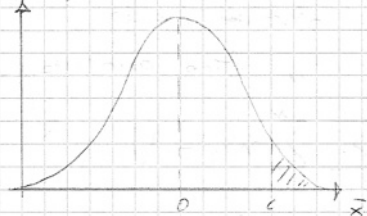
$$\alpha = P_{\mu} \{ \bar{x} \geq c / \mu=0 \} \quad [\alpha = P_{\mu} \{ x \in K_1 / H_0 \}]$$

Con modello normale $\bar{x} \sim (\mu, \frac{\sigma_x^2}{m})$ ($\mu=0, \text{var}(\bar{x}) = \frac{1}{m}$)

$$\bar{x} \sim N(\mu, \frac{1}{m}), \quad \bar{x} \sim N(0, \frac{1}{m}) = \frac{\sqrt{m}}{\sqrt{2\pi}} \exp \left\{ -\frac{m}{2} \bar{x}^2 \right\}$$

$$\textcircled{60} \quad \alpha = \int_c^{\infty} \frac{\sqrt{m}}{\sqrt{2\pi}} \exp \left\{ -\frac{m}{2} \bar{x}^2 \right\} d\bar{x}$$

$$N(0, \frac{1}{n})$$



Voglio usare le tabelle.

La porto a normale standard.

$$Z = (\bar{x} - 0) \sqrt{n} = \bar{z} \sqrt{n} \sim N(0, 1)$$

Quindi diventa:

$$\alpha = \int_{\frac{1}{\sqrt{n}}}^{\infty} e^{-\frac{1}{2}z^2} dz$$

(V_n)

Dalle tabelle conosco

$C\sqrt{n} = Z_{\alpha}$, percentile della normale

$$\text{Allora } X = \left\{ (x_1, x_2, \dots, x_n) / \bar{x} \geq \frac{Z_{\alpha}}{\sqrt{n}} \right\}$$

H (prime prova di ipotesi)

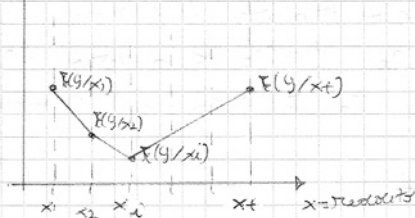
REGRESSIONE e INTERDIPENDENZA

Osserviamo un campione coppia variabili (x, y), var. al. coppia. Vedrò 2 tipi di relazioni, regre. e interdip.

"la REGRESSIONE è una rel. tra x e y dove una è indipendente come antecedente e l'altra conseguente, dipendente quella prima" (ex: reddito e spesa).

C'è DIPENDENZA MUTUA, vedo come variano le medie della vari. dipendente. Variano al variare di quella indep.

Es: $\Delta y = \text{media}$



Con il livello di reddito e il livello di spesa

si trova la spesa (funzione algebrica), $x < c$ distrib. di y

x	y_1	y_2	...	y_5	...	y_j	TOT.
x_1	m_{11}	m_{12}	...	m_{15}	...	m_{1j}	m_{10}
x_2	m_{21}	m_{22}	...	m_{25}	...	m_{2j}	m_{20}
...
x_i	m_{i1}	m_{i2}	...	m_{i5}	...	m_{ij}	m_{i0}
...
x_t	m_{t1}	m_{t2}	...	m_{t5}	...	m_{tj}	m_{t0}
Totale	m_{01}	m_{02}	...	m_{05}	...	m_{0j}	N

$N = \text{collettivo totale}$

intervallo (x_j, y_j)

m_{ij} : freq. assoluta

$$\sum_{i=1}^t \sum_{j=1}^r m_{ij} = N$$

$$\sum_{j=1}^r m_{ij} = m_{i0} \quad \text{e}$$

$$\sum_{i=1}^t m_{ij} = m_{0j}$$

la prima e l'ultima colonna danno variabile statistica X (DISTRIBUZIONE MARGINALE della X)

la 1. e ult. riga danno distrib. marginale della Y .
Generica riga e Y con X fissa: DISTRIB. della Y dato x_i , condizionata al valore x_i della X .

Generica colonna: e' DISTRIB. della X condizionata della Y .

$$\bullet E(Y/X=x_i) = \sum_{j=1}^r (y_j m_{ij}) / m_{i0}$$

$$\bullet E(X/Y=y_j) = \sum_{i=1}^t (x_i m_{ij}) / m_{0j}$$

Si suppone che esista i punti di $E(Y/X=x_i)$ e
" di REGRESSIONE. Ma, non alla rinfusa, la rappresentata X PREVEDERE. Delo trovare funzione che esprime legge che lega la Y alla X , ovvero il MODELLO secondo la quale la $E(Y/X=x_i)$ variando al variare della X .

⑥2 Caso + semplice,

Regressione lineare semplice (data dalla retta di regressione)

Prendiamo i dati e li riportiamo nel piano cartesiano.

$y \uparrow$  [SCATTER DIAGRAM]

$y = \alpha x + \beta$ determinato,
'brutto' (dati dovrebbero
comportarsi su retta)? Aggiungo
variabile spiega di disturbo

$$y = \alpha + \beta x + U = E(y/x) + U$$

Con U accidentale.

* Tutti quelli che frequentano x
hanno spesa con parte comune
(media) + oscillazione (U) che
considera altre variabili oltre

al risultato (ex. proprietario del appartamento). U var. aleatoria.
Della braccio retta delle medie [$y = \alpha x + \beta$] che
passa tra quei punti. Dello premere quella
migliore retta sotto ad ottimizzazione (ex. MINUS II).

$$y_i = E(y/x) + U_i \rightarrow \text{var. aleatoria}$$

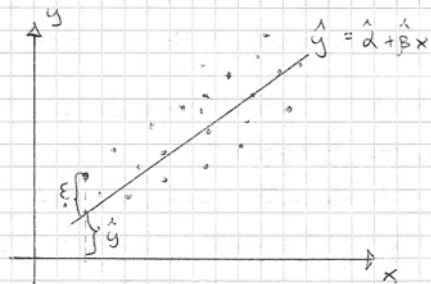
30/10/07

Modello di Reg. lin. semplice

IPOTESI:

$$\left\{ \begin{array}{l} U_i \sim N(0, \sigma_u^2) \\ X \text{ è una variabile matematica} \\ \text{Cov}(U_i, U_j) = E[(U_i - 0)(U_j - 0)] = E(U_i; U_j) = 0 \rightarrow \text{indipend.} \\ E(U_i) = 0 \\ E(U_i^2) = \sigma_u^2 \text{ (ovvero } E(U_i - 0)^2) \end{array} \right.$$

Entrare un campione di coppie (x_i, y_i) $i = 1 \dots n$



Stimo la retta migliore
con il metodo dei minimi quadrati
 E : residuo

$$y = \underbrace{\hat{y}}_{\text{teoria, media stimata}} + \epsilon = \underbrace{\hat{\alpha} + \hat{\beta}x}_{\text{teoria, media stimata}} + \epsilon$$

teoria, media stimata

$$y = \alpha + \beta x + U \quad [\text{Vera, ignota}]$$

U = var. ale. errore, $y = \hat{y} + E$ nel campione [E var. aleat. residua]

Il MQA dice che tra le ∞ rette che passano tra i punti, la migliore è quella x cui mi ha!

$$\varphi(\alpha, \beta) = \min \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n \epsilon_i^2$$

Quindi devo risolvere:

$$\begin{cases} \frac{\partial}{\partial \alpha} \varphi(\alpha, \beta) = 0 = -2 \sum_{i=1}^n (y_i - \alpha - \beta x_i) = 0 \\ \frac{\partial}{\partial \beta} \varphi(\alpha, \beta) = 0 = -2 \sum_{i=1}^n (y_i - \alpha - \beta x_i) x_i = 0 \end{cases}$$

$$\begin{cases} \sum_{i=1}^n y_i = n\alpha + \beta \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i y_i = \alpha \sum_{i=1}^n x_i + \beta \sum_{i=1}^n x_i^2 \end{cases}$$

Sistema delle
eq. normali
della retta

Divido per n :

$$\bar{y} = \alpha + \beta \bar{x} \Rightarrow \underline{\hat{\alpha} = \bar{y} - \hat{\beta} \bar{x}}$$

β è il coeff. ang. della retta. α è l'intercetta.

64 Se cambio origine, β è invariante e semplifico.

Pongo: (\rightarrow)

$$x_i' = x_i - \bar{x} \quad \text{e} \quad \text{ho} \quad \sum_{i=1}^n x_i' y_i = \beta \sum_{i=1}^n x_i'^2$$

$$y_i' = y_i - \bar{y}$$

$$\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \alpha \sum_{i=1}^n \cancel{(x_i - \bar{x})} + \beta \sum_{i=1}^n (x_i - \bar{x})^2$$

$$\left[\sum_{i=1}^n (x_i - \bar{x}) = \sum x_i - n\bar{x} = n\bar{x} - n\bar{x} = 0 \right]$$

$$\hat{\beta} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sum x_i y_i}{\sum x_i'^2} = \frac{\text{COV}(x, y)}{\text{DEV}(x)} \rightarrow \text{DEV}(y)$$

↑ CODERIVAZIONE

$$\text{Ho determinato } \hat{y} = \hat{\alpha} + \hat{\beta}x$$

$\hat{\beta}$ è il COEFFICIENTE DI REGRESSIONE: mi mostra di quanto varia la y al variare di 1 unità della x .

Ex: $\beta = 0.5$ reddito / mese, se ↑ redd. di 1€, spendo 0.5€ in più.

Stima mi serve a la previsione. Ammettendo che sia buono, $\forall x$ posso det. y .

Se $\beta = 0$ non c'è regressione, c'è INDIPENDENZA in CODA.

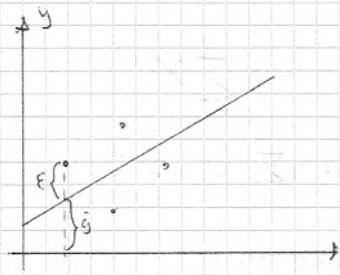
la scelta migliore la trovo sempre, ma non so se la bene.

Uno il COEFFICIENTE DI DETERMINAZIONE

si riferisce alla VARIABILITÀ della y .

$$\begin{aligned} \text{DEV}(y) &= \sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (y_i - \hat{y}_i + \hat{y}_i - \bar{y})^2 = \\ &= \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + 2 \sum_{i=1}^n (y_i - \hat{y}_i)(\hat{y}_i - \bar{y}) \end{aligned}$$

$$\text{Dim. che } \hat{y}_i = \bar{y}$$



$$\hat{y}_i = \hat{\alpha} + \hat{\beta} \bar{x} \quad \text{con} \quad \hat{\alpha} = \bar{y} - \hat{\beta} \bar{x}$$

$$\frac{\sum \hat{y}_i}{n} = \frac{\sum \hat{\alpha}}{n} + \frac{\sum \hat{\beta} \bar{x}}{n} = \bar{\hat{y}}_i = \hat{\alpha} + \hat{\beta} \bar{x} =$$

$$\bar{y} - \hat{\beta} \bar{x} + \hat{\beta} \bar{x} = \bar{y} \quad \text{CVD}$$

si ha quindi:

$$2 \sum_{i=1}^n (y_i - \hat{y}_i)(\hat{y}_i - \bar{y}) = 2 \sum_{i=1}^n \epsilon_i \hat{y}_i \quad \text{con} \quad y_i = y_i - \bar{y}$$

$$\hat{y}_i - \bar{y} = \hat{\beta}(x_i - \bar{x}) + \hat{y}_i = \hat{\beta} x_i \quad \rightarrow \quad = \hat{\beta} \sum_{i=1}^n \epsilon_i x_i$$

$$DEV(y) = DEV(TOTALE) = \sum_{i=1}^n \epsilon_i^2 + \sum_{i=1}^n \hat{y}_i^2 = DEV(E) + DEV(R)$$

\downarrow \downarrow
 accidentale di regressione

le coeff. di determ. e:

$$\boxed{R^2 = \frac{DEV(R)}{DEV(TOT)} = 1 - \frac{DEV(E)}{DEV(TOT)}} = \frac{\sum_{i=1}^n \hat{y}_i^2}{\sum_{i=1}^n y_i^2} = 1 - \frac{\sum_{i=1}^n \epsilon_i^2}{\sum_{i=1}^n y_i^2}$$

Indice $0 \leq R^2 \leq 1$

Se $R^2 = 0$, $DEV(R) = 0 \Rightarrow DEV(TOT) = DEV(E)$

" $R^2 = 1$, $DEV(TOT) = DEV(R)$, i punti sono tutti sulla retta. Se $R^2 \rightarrow 1$, retta + n. avvicina ai dati.

Es: $R^2 = 78\%$ buono; in generale meglio superiore il 50%.

Dal modello base. $y = \alpha x + \beta x + U$, y e x vari. aleat.
 $y = \alpha + \beta x + U \sim N(\alpha + \beta x, \sigma_u^2)$

[la cond. lin. si varia N e N]

$E(y) = \alpha + \beta x + E(U) = 0$ per ipotesi $= \alpha + \beta x$ (media

66) n di disp. su una retta)

$$\text{Var}(y) = E[(y - E(y))^2] = E[\alpha + \beta x + U - \alpha - \beta x]^2 = E(U^2) = \sigma_u^2$$

Stima della Varianza: $\hat{\sigma}_u^2 = \frac{\sum_{i=1}^n E_i^2}{n-2} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-2}$

$$\hat{\beta} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sum_{i=1}^n (x_i - \bar{x}) y_i - \bar{y} \sum_{i=1}^n (x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad \begin{matrix} \text{(visto prima)} \\ \text{Allora} \end{matrix}$$

$$\hat{\beta} = \frac{\sum_{i=1}^n (x_i - \bar{x}) y_i}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sum_{i=1}^n \underbrace{(x_i - \bar{x})}_{w_i} y_i}{\sum_{i=1}^n x_i^2} = \sum_{i=1}^n w_i y_i$$

$\hat{\beta}$ è comb. lineare di y_i normali $\Rightarrow \hat{\beta} \sim N$

$$E(\hat{\beta}) = E\left(\sum w_i y_i\right) = E\left[\sum w_i (\alpha + \beta x_i + U_i)\right] = E\left[\alpha \sum w_i + \beta \sum w_i x_i + \sum w_i U_i\right]$$

$$\sum w_i = \frac{\sum x_i}{\sum x_i^2} = \frac{\sum (x_i - \bar{x})}{\sum x_i^2} = 0$$

$$\begin{aligned} \sum w_i x_i &= \frac{\sum x_i x_i}{\sum x_i^2} = \frac{\sum (x_i - \bar{x}) x_i}{\sum (x_i - \bar{x})^2} = 1, \quad \sum (x_i - x)^2 = \sum (x_i - \bar{x})^2 \\ &= \frac{\sum (x_i - \bar{x})^2}{\sum (x_i - \bar{x})^2} = 1 \quad \text{Quindi} \end{aligned}$$

$$E[\beta + \sum w_i U_i] = \beta + \sum w_i E(U_i) = \beta \quad (\text{e' stim corretto})$$

$\Gamma y = \alpha + \beta x + U ; U \sim N(0, \sigma_u^2)$ 31/10/07

\hat{x} e Var. matematica: $E(U) = 0, E(U^2) = \sigma_u^2, E(U_i U_j) = 0 \text{ i} \neq j$

$$\hat{y} = \hat{\alpha} + \beta \hat{x} ; \hat{\alpha} = \bar{y} - \hat{\beta} \bar{x} ; \hat{\beta} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \textcircled{6x}$$

$$= \sum_{i=1}^N w_i y_i \quad \text{con } w_i = \frac{x_i}{\sum x_i^2}; \quad y \sim N(\alpha + \beta x, \sigma_u^2)$$

$$\hat{\beta} = \sum w_i y_i = \sum w_i (\alpha + \beta x_i + u_i) = \alpha \sum w_i + \beta \sum w_i x_i + \sum w_i u_i$$

$$E(\hat{\beta}) = \beta + \sum w_i E(u_i) = \beta$$

Ora calcola la varianza.

$$\begin{aligned} \text{Var}(\hat{\beta}) &= E[\hat{\beta} - \beta]^2 = E\left[\sum w_i u_i\right]^2 = E\left[\sum_{i=1}^N w_i^2 u_i^2 + \sum_{i \neq j} \sum_{j=1}^N w_i w_j u_i u_j\right] \\ &= \sum_{i=1}^N w_i^2 E(u_i^2) + \sum_{i \neq j} \sum_{j=1}^N w_i w_j E(u_i u_j) \end{aligned}$$

$$\text{Var}(\hat{\beta}) = \sigma_u^2 \left(\sum_{i=1}^N w_i^2 \right) = \sum \left[\frac{x_i}{\sum x_i^2} \right]^2 = \frac{\sum x_i^2}{\left(\sum x_i^2 \right)^2} = \frac{1}{\sum x_i^2}$$

$$= \frac{\sigma_u^2}{\sum_{i=1}^N x_i^2} \quad \text{Quindi la var. di } \hat{\beta} \text{ e' } \sim N\left(\beta, \frac{\sigma_u^2}{\sum_{i=1}^N x_i^2}\right)$$

(da π^2).

$$\sum_{i=1}^N (y_i - \bar{y})^2 = \sum_{i=1}^N (\hat{y}_i - y_i)^2 + \sum_{i=1}^N (\hat{y}_i - \bar{y})^2 + 2 \sum_{i=1}^N (\hat{y}_i - y_i) \cdot (\hat{y}_i - \bar{y})$$

$$, (\hat{y}_i - \bar{y}) \quad ; \quad \text{e' lo che } \sum_{i=1}^N (\hat{y}_i - y_i) (\hat{y}_i - \bar{y}) =$$

$$\textcircled{68} = \sum_{i=1}^N \epsilon_i \hat{y}_i, \quad \hat{y}_i = \hat{\alpha} + \hat{\beta} x_i$$

↳ quindi

$$\begin{aligned} [\hat{y}_i = \hat{\beta} x_i, \text{minimale}] &= \hat{\beta} \sum_{i=1}^N \epsilon_i x_i = \hat{\beta} \sum_{i=1}^N (\hat{y}_i - y_i) x_i = \\ &= \hat{\beta} \sum_{i=1}^N (\hat{y}_i - y_i) x_i \quad \left(\begin{array}{l} \text{aggiungo e} \\ \text{sottr. } \bar{y} \end{array} \right) \end{aligned}$$

$$\begin{aligned} [\hat{y}_i - \bar{y} - (y_i - \bar{y})] &= \hat{\beta} \left[\sum \hat{y}_i x_i - \sum y_i x_i \right] = \\ [\hat{y}_i] & \quad [y_i] \quad \left[\sum x_i y_i = \sum (x_i - \bar{x})(y_i - \bar{y}) = \right. \\ &= \hat{\beta} \left[\hat{\beta} \sum x_i^2 - \sum x_i^2 \right] = 0 \quad \left. = \sum (x_i - \bar{x}) y_i - \bar{y} \sum (x_i - \bar{x}) \right] \end{aligned}$$

$$R^2 = \frac{DEV(R)}{DEV(Y)} = 1 - \frac{DEV(RR)}{DEV(Y)} = \frac{\sum_{i=1}^N \hat{y}_i^2}{\sum_{i=1}^N y_i^2} \quad \text{con } 0 \leq R^2 \leq 1$$

Indice x dice quanta variabilità della y è ascrivibile alla relaz. di regressione.

H

INTERDIPENDENZA tra (x, y)

Prenzo nel di rimbrebui: Ex novo con Diam. e Lung.
non so che è antecedente, sono correlate. $\odot \perp$

Prima recalcito influen. la spora, ora in impl. a vicenda
Concorrenza e Discordanza

(x, y) conc. quando a Valori ^[grandi] piccoli dell'una corr.
Valori ^[grandi] piccoli dell'altra.

Opposto quando discordanti.

la COEFFICIENTE (x, y) misura questa proporzionalità.

$$\text{COB}(x, y) = \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}) \in (-\infty, +\infty)$$

- Se $\text{cov} < 0$, le 2 var. sono DISCORDANTI:

Valori piccoli $<$ media (parto negativo).

Valori grandi $>$ " (" positivo) $\Rightarrow [-] \cdot [+] = [-]$

- Se $\text{cov} > 0$, sono CONCORDANTI:

$(\delta -) \cdot (\delta -) [+]$ oppure $(\delta +) \cdot (\delta +) [+]$

\exists la DISUGUGLIANZA DI SCHWARTZ

$$\left[\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}) \right]^2 \leq \sum_{i=1}^N (x_i - \bar{x})^2 \cdot \sum_{i=1}^N (y_i - \bar{y})^2$$

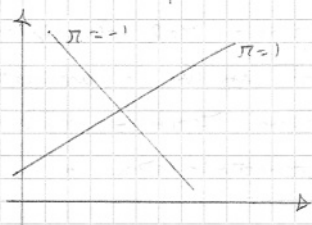
Si ricava il COEFFICIENTE DI CORRELAZIONE LINEARE DI BRAUNIS-PEARSON:

$$r = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^N (y_i - \bar{y})^2}} = \frac{\text{cov}(x, y)}{\sqrt{\text{DEV}(x) \cdot \text{DEV}(y)}} \quad -1 \leq r \leq 1$$

$r = -1$ DISCORD. MAX

$r = 1$ CONCORD. "

↓
r misura l'intensità della



(FINE TEORIA: 31/10/07)