$$S_y^2 = \frac{\sum(y_i-\bar y)^2}{n-1} = \frac{\sum(a+bx_i-a-b\bar x)^2}{n-1} = b^2\frac{\sum(x_i-\bar x)^2}{n-1} = b^2 S_x^2$$

$$S_y = \sqrt{b^2 S_x^2} = |b|\,S_x$$

$\bar x + S_x$: 68%
$\bar x + 2S_x$: 95%
$\bar x + 3S_x$: 99.7%

BOX PLOT: un'alternativa più semplice all'istogramma

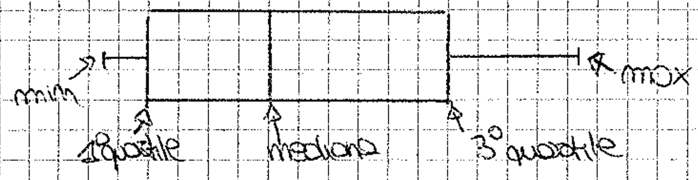


DIAGRAMMA A SCATOLA E BAFFI (o BOXPLOT)

PROBABILITÀ

28-09-09

Esperimento casuale: i suoi esiti (risultati) non si possono prevedere con certezza → non deterministico

S = spazio degli esiti dell'esperimento

1) Lancio di un dado — $S = \{1,2,3,4,5,6\}$
2) Lancio di 2 monete — $S = \{(T,T),(T,C),(C,T),(C,C)\}$
3) Corsa di 7 cavalli — $S = \{\ldots\} = \{$tutte le possibili $7!$ classifiche$\}$
4) Tempo in cui si brucia una lampadina — $S = \mathbb{R}^+$

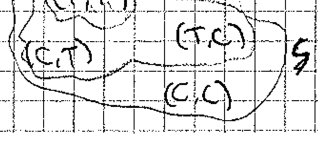
Un evento è un sottoinsieme di S (a meno di complicazioni matematiche del tipo dei paradossi di teoria degli insiemi)
↳ unione $\cup$ cup, intersezione $\cap$ cap, negazione A^c, complemento
= operazioni logiche definite sugli eventi

1) E = "esce un numero pari" $= \{2,4,6\} \subseteq S$
2) E = "esce almeno una testa" $= \{(T,T),(C,T),(T,C)\}$ — E_2 "non esce croce" $= \{(T,T)\}$
3) E = "il 5 si piazza" $= \{(5,i_1,i_2\ldots i_6),(i_1,i_2,5,i_3\ldots i_6),(i_1,5,i_2,\ldots i_6)\}$, $i_1,i_2\ldots i_6 \in \{1,2,3,4,5,6\}$
4) E_1 = "la lampadina dura meno di 3 ore" $= (0,3) \subseteq \mathbb{R}^+$
E_2 = "la lampadina dura almeno mezz'ora" $= [0.5, \infty)$

Operazioni logiche:

2) $E_1 \cap E_2 = E_2 = \{(T,T)\}$
4) $E_1 \cap E_2 = [0.5, 3)$
$E_1 \cup E_2 = \mathbb{R}^+ = S$
$(E_1 \cup E_2)^c = \emptyset$

Diagramma di Venn

2)


$E_1^c = \{(C,C)\}$

PROBABILITÀ = funzione d'insieme che gode di 3 proprietà
1) $0 \leq P(A) \leq 1$
2) $P(S) = 1$
3) $P(A \cup B) = P(A) + P(B)$ se $A \cap B = \emptyset$

Note: • P è una funzione (d'insieme) il cui dominio è l'insieme degli eventi (dei sottoinsiemi) di S
• l'immagine di P è $[0, 1]$
• P è normalizzata a 1 $\rightarrow P(S) = 1$

PROPRIETÀ:
• $P(A^c) = 1 - P(A)$
• $P(\emptyset) = 0 = 1 - P(S) = 1 - 1$
insieme vuoto = evento impossibile

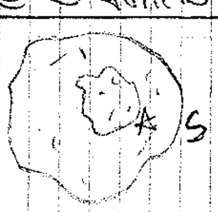
In generale, se A e B qualsiasi, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ REGOLA DELL'ADDIZ.

SPAZI UNIFORMI = ESITI EQUIPROBABILI (G.G)
: 1) lo spazio S è finito, $\#S < \infty$
cardinalità
2) tutti gli esiti hanno la stessa probabilità

2) Se le monete sono eque (non truccate) e se vengono lanciate a caso, S è uniforme
$P(\{(T,T)\}) = P(\text{"T,T"}) = P(\text{"C,T"}) = P(\text{"T,C"}) = P(\text{"C,C"}) = \frac{1}{4}$
Se S è uniforme, allora ciascun esito ha la stessa probabilità
$p = P(\text{"esito"}) = \frac{1}{\#S}$ in quanto $\sum_{i=1}^{\#S} p = 1$
cioè $p \#S = 1 \Rightarrow p = \frac{1}{\#S}$

L'esempio dei cavalli non funziona, poiché i cavalli hanno diversa forza e non tutte le classifiche sono equiprobabili

Se S finito uniforme:



$$P(A) = \#A \cdot \frac{1}{\#S} = \frac{\#A}{\#S} = \frac{\#\text{Casi favorevoli}}{\#\text{Casi possibili}}$$

2) $P(E_1) = 3 \cdot \frac{1}{4} = \frac{3}{4}$

Sono importanti le tecniche di CONTEGGIO (4.7)

TECNICHE DI CONTEGGIO - MATEMATICA COMBINATORIA ELEMENTARE

$\#A \times B = \#A \cdot \#B$ n° di possibili coppie di A e B
• 10 donne, 4 uomini $\#$ coppie $= 4 \cdot 10 = 40$
• 10 uomini, 4 donne, 5 bambini $\#$ coppie $= 10 \cdot 4 \cdot 5$

DISPOSIZIONE (SEMPLICE) SENZA RIPETIZIONE (con ordine) di n elementi presi r per volta
è una r-pla degli n elementi senza ripetizione
$n(n-1) \cdot \ldots \cdot (n-r+1)$

PERMUTAZIONI SENZA RIPETIZIONE: (senza ordine) disposizioni di n elementi n alla volta

COMBINAZIONE di n elementi r alla volta è un sottoinsieme degli n elementi lungo r

$$\frac{\#\text{ disposizioni di } n \text{ r alla volta}}{\#\text{ permutazioni di } r \text{ elementi}} = \frac{n(n-1)\dots(n-r+1)}{r!} = \frac{n!}{r!(n-r)!} = \binom{n}{r}$$ coefficiente binomiale "n su r"

30-09-09

RIASSUNTO

S = spazio degli eventi

$E \subseteq S$ evento

$P(E)$ è la sua probabilità

1) $0 \le P(E) \le 1$

2) $P(S) = 1$

3) se $A \cap B = \emptyset$ (A e B disgiunti)

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

3*) se $A_1, A_2 \dots$ sono disgiunti $(A_i \cap A_j) = \emptyset$ per ogni coppia (i,j) allora

$P(\underbrace{\bigcup_{i=1}^{\infty} A_i}_{A_1 \cup A_2 \cup \dots \cup A_i}) = \sum_{i=1}^{\infty} P(A_i)$ ← è una sottigliezza matematica

Se gli esiti sono equiprobabili (x simmetria, ad es. nei giochi d'azzardo), allora

$$P(E) = \frac{\#E}{\#S} = \frac{\#CF}{\#CP}$$

- Le disposizioni di n elementi r alla volta sono:

$$n(n-1) \dots (n-r+1)$$

- Le combinazioni di n elementi r alla volta sono

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

- es. il n° di possibili "prime file" lunghe 9 da una classe di 100 è

$$100 \cdot 99 \cdot 98 \cdot 97 \cdot 96 \cdot 95 \cdot 94 \cdot 93 \cdot 92$$

il n° di possibili commissioni di 9 studenti della stessa classe è

$$\binom{100}{9} = \frac{100!}{9!\,91!} = \frac{100 \cdot 99 \cdot 98 \cdot 97 \cdot 96 \cdot 95 \cdot 94 \cdot 93 \cdot 92 \cdot 91!}{9! \cdot 91!}$$

- esempio: esperimento casuale → lancio di due dadi equi a caso

$\#S = 6 \times 6 = 36$ esiti equiprobabili

A = "somma è ≥ 11" $\qquad \#A = \{(5,6), (6,5), (6,6)\} = 3$

B = "il primo dado è 6" $\qquad \#B = \{(6,1), (6,2), (6,3), (6,4), (6,5), (6,6)\} = 6$

$P(A) = \frac{\#CF}{\#CP} = \frac{3}{36} = \frac{1}{12}$ $\qquad P(B) = \frac{\#CF}{\#CP} = \frac{6}{36} = \frac{1}{6}$

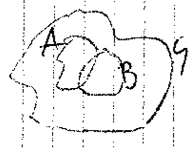
PROBABILITÀ CONDIZIONATA

PROBABILITÀ CONDIZIONATA = probabilità valutata in presenza dell'informazione che un altro evento si sia verificato.

$P(A)$ = probabilità dell'evento A

$P(A|B)$ = probabilità di A dato B $\qquad P(A|B) = \frac{P(A\cap B)}{P(B)}$

Esempio: nel lancio di due dadi, qual è $P(A|B)$?



la $P(A|B)$ è $P(A|B) = \frac{P(A\cap B)}{P(B)} = \frac{2/36}{6/36} = \frac{1}{3} > P(A) = \frac{1}{2}$

se $P(B) > 0$

$P(A\cap B) = P(\text{"}(6,5) \text{ o } (6,6)\text{"}) = \frac{2}{36}$

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

Se il conoscere B non cambia la probabilità di A, allora A e B si dicono INDIPENDENTI e $P(A|B) = P(A)$ ovvero $P(A\cap B) = P(A)P(B)$

- Esempio: test diagnostici
  - SENSITIVITÀ: P("test positivo" / "malatto")
  - SPECIFICITÀ: P("test negativo" / "no malatto")

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

$= P(A|B)P(B) + P(A|B^c)P(B^c)$ ← formula delle probabilità totali

Es. 17 pag 170

- Esempio: esperimento screening, cioè selezione a caso di un individuo da una popolazione seguito da somministrazione di un test diagnostico

eventi: M = "l'individuo scelto è malato"

$+$ = "il teste è positivo"

dati: $P(M) = 0.005$ = prevalenza = proporzione di malati nella popolazione

sensitività $P(+|M) = 0.99$

specificità $P(-|M^c) = 0.98$

per le probabilità totali:

$P(+) = P(+/M)P(M) + P(+|M^c)P(M^c) =$

$= 0.99 \times 0.005 + 0.02(1-0.005) = 0.025$

in realtà siamo interessati a

$P(M/+) = \frac{P(M\cap +)}{P(+)} = \frac{P(+|M)P(M)}{P(+)} = 0.199 \simeq 20\%$

TEOREMA DI BAYES: $P(B|A) = \frac{P(A|B)P(B)}{P(A)}$

05-10-09

RIASSUNTO

S = insieme esiti

PROBABILITÀ: se A evento ($A \subseteq S$)
1) $0 \le P(A) \le 1$
2) $P(S) = 1$
3) se $A \cap B = \emptyset$ (A, B disgiunti) $\Rightarrow P(A \cup B) = P(A) + P(B)$

A e B sono indipendenti se $P(A \cap B) = P(A) \cdot P(B)$

PROBABILITÀ CONDIZIONATA di A dato B, A e B eventi e $P(B) > 0$

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

TEOREMA di BAYES: $$P(H|E) = \frac{P(E|H)\,\overset{\text{ipotesi}}{P(H)}}{\underbrace{P(E|H)P(H) + P(E|H^c)P(H^c)}_{=P(E)}}$$

H, H^c → statistica Bayesiana

VARIABILI ALEATORIE

- o casuali (random variables)
- o stocastiche

<u>Una variabile aleatoria X è un aspetto numerico del nostro esperimento</u>

- Lanciamo 2 monete a caso (cioè in maniera indipendente)

Il primo è equo: $P(\text{"primo} = T\text{"}) = \frac{1}{2}$

Il secondo non è equo: $P(\text{"secondo} = T\text{"}) = \frac{1}{3}$

Sia X = n° totale di teste: i valori possibili per X sono $x = 0, 1, 2$

L'evento "$X=0$" = "(C,C)" $\subseteq S$ ha probabilità

$$P(X=0) = P(\text{"C,C"}) = \underbrace{P(\text{"primo C"})\,P(\text{"secondo C"})}_{\text{per indipendenza}} = \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3} = \frac{2}{6}$$

L'evento "$X=1$" $= \{(T,C), (C,T)\}$

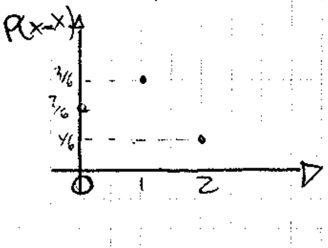
$$P(X=1) = P(\{(T,C),(C,T)\}) = P(\text{"T,C"} \cup \text{"C,T"}) =$$
$$= P(\text{"T,C"}) + P(\text{"C,T"}) = P(\text{"primo T"})P(\text{"secondo C"}) + P(\text{"primo C"})P(\text{"secon}$$
$$= \frac{1}{2} \cdot \frac{2}{3} + \frac{1}{2} \cdot \frac{1}{3} = \frac{3}{6} = \frac{1}{2}$$

L'evento "$X=2$" $= \{(T,T)\}$

$$P(X=2) = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$$

Matematicamente, X è una funzione da S a $\mathbb{R}$

↳ abbiamo implicitamente trasportato delle probabilità su $\mathbb{R}$



Una variabile aleatoria DISCRETA può assumere (logicamente) un numero finito o un'infinità numerabile di valori - (X nell'esempio è discreta)

nell'esempio, anche la v.a. $F = \frac{X}{2}$, cioè la frequenza relativa di teste, è discreta perché può assumere solo i valori $0, \frac{1}{2}, \frac{2}{2}$

• Esempio 3 : campioniamo un cofano di auto

$X =$ n° totale di imperfezioni sul cofano

i possibili valori di X sono $x = 0, 1, 2$

X è una v.a. discreta suscettibile di assumere infiniti valori

Tramite la definizione di X veniamo definito su $\mathbb{R}$ la retta reale, una distribuzione di probabilità descritta dalla funzione (chiamata funzione di massa di probabilità) $x \to P(X = x)$: $\sum P(X = x_i) = 1$

verifico nell'esempio 1: $\frac{2}{6} + \frac{1}{6} + \frac{3}{6} = 1$

VALORE ATTESO

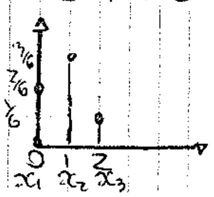
Sia X una variabile aleatoria che può assumere i valori $x_1, x_2, x_3 \dots$ con rispettive probabilità $P(X = x_1), P(X = x_2) \dots$

Si definisce VALORE ATTESO (expectation) di X :

$$E(X) = \sum_i x_i P(X = x_i)$$

↑ "baricentro" della distribuzione di probabilità

$\sum_i$ sarà una somma finita o infinita secondo che il numero di valori $x_1, x_2 \dots$ sia finito o infinito



$$E(X) = \sum_{i=1}^{3} x_i P(X = x_i) =$$
$$= 0 \cdot \frac{2}{6} + 1 \cdot \frac{3}{6} + 2 \cdot \frac{1}{6} = \frac{5}{6}$$

Analogia formale tra distribuzione statistica di frequenze (x_i, f_i) con f_i frequenza relativa, per esempio tabulata, e $(x_i, P(X = x_i))$

Analogia tra $\bar{x} = \sum x_i f_i$ e $E(X) = \sum x_i P(X = x_i)$

$X_1 =$ possibile risultato dell'investimento 1

$X_2 =$ " " " " 2

VARIANZA

Se X è una variabile aleatoria con valore atteso $\mu = E(X)$, allora $X - \mu$ è la variabile aleatoria scarto e $(X - \mu)^2$ è la v.a. "quadrato dello scarto". Il valore atteso di $(X - \mu)^2$ si dice VARIANZA

$$E((X - \mu)^2) = \sum (x_i - \mu)^2 P(X = x_i) \qquad VAR(X)$$

$$Var(X) = \left(0 - \frac{5}{6}\right)^2 \cdot \frac{2}{6} + \left(1 - \frac{5}{6}\right)^2 \cdot \frac{3}{6} + \left(2 - \frac{5}{6}\right)^2 \cdot \frac{1}{6}$$

PROPRIETÀ DI $E(X)$ e $Var(X)$

• $E(X)$ è lineare, cioè:

1) se $Y = cX$ c costante, allora $E(Y) = E(cX) = c\,E(X)$

2) se $Y = X_1 + X_2$ X_1 e X_2 v.a. allora $E(Y) = E(X_1) + E(X_2)$

in generale $E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i)$

- se $Y = g(x)$ è una funzione di x, $E(Y) = \sum g(x_i)\, P(X = x_i)$
- $Var(X) = E((X-\mu)^2) = E(X^2 + \mu^2 - 2\mu X) =$

$$= E(X^2) + \mu^2 - 2\mu \underbrace{E(X)}_{\mu} =$$

$$= E(X^2) - \mu^2$$

- La radice della varianza si chiama

$$\text{DEVIAZIONE STANDARD} = SD(X) = \sqrt{Var(X)}$$

07-10-08

<u>Due variabili aleatorie X e Y si dicono indipendenti se $P(X = x_i \cap Y = y_j) = P(X = x_i)\,P(Y = y_j)$</u>
<u>per ogni scelta x_i, y_j</u>

<u>ESEMPIO</u>: Lancio a caso 2 monete una equa e la seconda tale che $P(\text{"seconda = testa"}) = \frac{1}{3}$

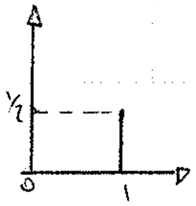
Sia $X_1 =$ n° di teste della prima moneta
$X_2 =$ n° di teste della seconda moneta
$X = X_1 + X_2$ X_1 e X_2 sono indipendenti : $P(X_1 = 0 \cap X_2 = 1) = P(\text{"primo = croce"}) \cdot P(\text{"seconda = testa"})$

X_1 e X non sono indipendenti:

$$P(X_1 = 0 \cap X = 0) = P(\text{"primo C"} \cap \text{"secondo C"}) = \frac{1}{2} \cdot \frac{2}{3} \neq \underset{P(X_1=0)}{\frac{1}{2}} \cdot \underset{P(X=0)}{\frac{2}{6}}$$

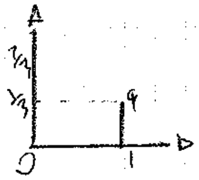
PROPRIETÀ DI $E(X)$ e $Var(X)$

- $E(X+Y) = E(X) + E(Y)$ per qualsiasi X, Y
- $Var(X+Y) = Var(X) + Var(Y)$ se X e Y indipendenti



$$E(X_1) = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{1}{2}$$

$$E(X_2) = 0 \cdot \frac{2}{3} + 1 \cdot \frac{1}{3} = \frac{1}{3}$$



- $E((X - E(X))^2) = Var(X) = E(X^2) - (E(X))^2 = \sum x_i^2\, P(X = x_i) - \mu^2$

$$Var(X_1) = \sum x_i^2\, P(X = x_i) - E(X_1)^2 =$$

$$= 0^2 \cdot \frac{1}{2} + 1^2 \cdot \frac{1}{2} - \left(\frac{1}{2}\right)^2 = \frac{1}{2} - \left(\frac{1}{2}\right)^2 = \frac{1}{4}$$

$$Var(X_2) = \frac{2}{9} = \frac{1}{3} \cdot \frac{2}{3}$$

$X = X_1 + X_2$

$$E(X) = \ldots = 0 \cdot \frac{2}{6} + 1 \cdot \frac{3}{6} + 2 \cdot \frac{1}{6} = \frac{5}{6}$$

$$Var(X) = 0^2 \cdot \frac{2}{6} + 1^2 \cdot \frac{3}{6} + 2^2 \cdot \frac{1}{6} - \left(\frac{5}{6}\right)^2 = \frac{3}{6} + \frac{4}{6} - \frac{25}{36} = \frac{17}{36}$$

$$Var(X) = \frac{17}{36} \overset{?}{=} Var(X_1) + Var(X_2) = \frac{1}{4} + \frac{2}{9} = \frac{17}{36}$$ ← SÌ, PERCHÉ X_1 e X_2 SONO INDIPENDENTI

<u>Una variabile si dice</u> <u>BERNOULLIANA (o BINARIA)</u> <u>se assume i valori 0 e 1</u> $X \sim \text{Bernoulli}(p)$

Se poniamo $p = P(X = 1)$, allora si dice che X è una bernoulliana con parametro p

Abbiamo $E(X) = p$ $Var(X) = p(1-p)$

Nell'esempio X_1 è Bernoulli$(\frac{1}{2})$ e X_2 è Bernoulli$(\frac{1}{3})$

Se considero n prove binarie indipendenti, ciascuno con lo stesso distribuzione, è interessante considerare il risultato cumulativo.

Formalmente considero $x_1, x_2, \dots, x_n$ indipendenti ciascuno Bernoulli (p)
dove $p = P(x_i = \underset{\text{successo}}{1})$ = probabilità di successo di ogni singola prova $i = 1, 2 \dots n$

e definisco $x = x_1 + x_2 + \dots + x_n$ cioè il numero di successi sulle n prove

- Moneta truccata $p = \frac{1}{3}$

$n = 4$

x = n° di teste (successi) su 4 lanci casuali

$P(x=0) = P(\text{"}\underset{I}{C}\underset{I}{C}\underset{I}{C}\underset{I}{C}\text{"}) = \frac{2}{3}\frac{2}{3}\frac{2}{3}\frac{2}{3} = \left(\frac{2}{3}\right)^4$

$P(x=1) = P(\text{"}\underset{\text{successo}}{S}\underset{\text{insuccesso}}{III} \cup ISII \cup IISI \cup IIIS\text{"}) = \frac{1}{3}\cdot\frac{2}{3}\cdot\frac{2}{3}\cdot\frac{2}{3} + \frac{2}{3}\frac{1}{3}\frac{2}{3}\frac{2}{3} + \dots = 4\left(\frac{1}{3}\right)^1\left(\frac{2}{3}\right)^3$

$P(x=3) = 4\left(\frac{2}{3}\right)^1\left(\frac{1}{3}\right)^3$

$P(x=4) = \left(\frac{1}{3}\right)^4$

$P(x=2) = \binom{4}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^2$

Si può dimostrare che $\boxed{P(x=i) = \binom{n}{i} p^i (1-p)^{n-i}}$

DIMOSTRAZIONE: la probabilità di una singola stringa con i successi e $n-i$ insuccessi è $p^i(1-p)^{n-i}$
di tali stringhe equiprobabili ce ne sono $\binom{n}{i}$, il numero di modi di scegliere i posti su n

Una variabile aleatoria si dice BINOMIALE con parametri n e p se

$$P(x=i) = \binom{n}{i} p^i (1-p)^{n-i} \qquad i = 0, 1, 2, \dots n$$

Verifichiamo: $\sum_{i=0}^{n} P(x=i) = \sum_{i=0}^{n} \binom{n}{i} \underset{a}{p^i} \underset{b}{(1-p)^{n-i}}$

BINOMIO DI NEWTON $= (a+b)^n = (p+1-p)^n = 1$

Calcoliamo valore atteso e varianza di x

$\underline{\underline{E(x)}} = \sum_{i=0}^{n} i \binom{n}{i} p^i (1-p)^{n-i}$

$E(x_1 + x_2 + \dots + x_n) = \sum_{i=1}^{n} E(x_i) = p + p + p \dots = \underline{\underline{np}}$

$\underline{Var(x)} = Var(x_1 + x_2 + \dots x_n) = Var(\sum x_i) =$ x_i indipendenti

$= \sum Var(x_i) = \underline{np(1-p)}$

* Esempio: un'urna contiene 4 palle bianche e 5 nere.
ben mescolato, estraggo una palla, ne registro il colore poi la reintroduco nell'urna. Ripeto per 10 volte

Se indichiamo successo = bianco

Abbiamo $n = 10$ prove indipendenti (causa la reintroduzione e rimescolamento), con lo stesso probabilità di successo $p = \frac{4}{9}$

Quindi il numero totale di bianche estratte è $x \sim$ binomiale $(10, \frac{4}{9})$

* Controesempio: estrazione senza reintroduzione, $n = 3$ volte

$P(x=2) = P(\text{"2 bianche su 3 estratte"}) = P(\text{"}BBN \cup BNB \cup NBB\text{"}) =$

$$= \frac{4\cdot 3\cdot 5}{8\cdot 8\cdot 7} + \frac{4\cdot 5\cdot 3}{8\cdot 8\cdot 7} + \frac{5\cdot 4\cdot 3}{8\cdot 8\cdot 7}$$

In generale, l'estrazione senza reintroduzione di m da un'urna con B bianche e N nere dà origine ad un numero totale X di bianche sulle m estratte che ha probabilità

$$P(X=i) = \frac{\binom{B}{i}\binom{N}{m-i}}{\binom{B+N}{m}}$$

X si dice variabile aleatoria IPERGEOMETRICA

Nell'esempio $m=3$, $B=4$, $N=5$, $i=2$

$$\Rightarrow P(X=2) = \frac{\binom{4}{2}\binom{5}{1}}{\binom{8}{3}} = \frac{\frac{4\cdot 3}{2}\cdot 5}{\frac{8\cdot 8\cdot 7}{3\cdot 2}} = 3\,\frac{3\cdot 4\cdot 5}{8\cdot 8\cdot 7}$$

Una variabile aleatoria X si dice di POISSON con parametro λ se

$$P(X=i) = \frac{e^{-\lambda}}{i!}\lambda^i \qquad i=0,1,2\ldots$$

È un esempio di variabile aleatoria discreta che può assumere un numero infinito di valori

12-10-09

Una v.a. discreta X assume diversi valori $x_1, x_2 \ldots$ con diverse probabilità $P(X=x_1)$, $P(X=x_2)$ ....

V.A. NOTEVOLI

1) $X \sim$ Bernoulli (p)
se $P(X=1) = p = 1 - P(X=0)$ parametro p: $0 \le p \le 1$

2) $X \sim$ Binomiale (n, p)
se $P(X=i) = \binom{n}{i} p^i (1-p)^{n-i}$ $\quad i = 0, 1 \ldots n \quad$ n intero positivo

X rappresenta il numero di successi su n prove binarie, indipendenti, ciascuna con la stessa probabilità di successo p

$E(X) = n\cdot p$
$Var(X) = n\cdot p(1-p)$

3) $X \sim$ Poisson (λ)
se $P(X=i) = \frac{e^{-\lambda}\lambda^i}{i!}$ $\quad i=0,1,2\ldots$

Si dimostra che

SERIE ESPONENZIALE

3.1) $$\sum_{i=0}^{\infty} P(X=i) = \sum_{i=0}^{\infty} \frac{e^{-\lambda}}{i!}\lambda^i = e^{-\lambda}\sum_{i=0}^{\infty}\frac{\lambda^i}{i!} = e^{-\lambda}e^{\lambda} = e^0 = 1$$

3.2) $$E(X) = \sum_{i=0}^{\infty} i\,\frac{e^{-\lambda}}{i!}\lambda^i = \ldots = \lambda$$

3.3) $$Var(X) = E(X^2) - \lambda^2 = \ldots = \lambda$$

ESEMPIO 5.23

$X =$ n° di incidenti per settimana

$\lambda = 1.2$ n° medio di incidenti, $E(X)$

P("almeno un incidente nella settimana") $= P(x>0) = 1 - P(x=0) =$

$= 1 - \frac{e^{-\lambda}}{0!}\lambda^0 = 1 - e^{-1.2} \approx 0.70$

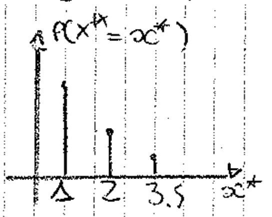
VARIABILI ALEATORIE CONTINUE

Aspetti quantitativi di un esperimento aleatorio di tipo continuo (spazio, tempo, pressi...

ESEMPIO: x = tempo di reazione in un certo esperimento chimico

Potremmo approssimare x con x^*

x^*	
≈ 1	0.6
≈ 2	0.3
≈ 3.5	0.1



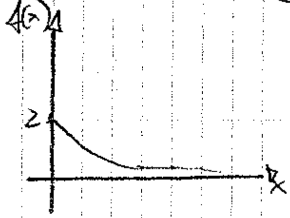
In realtà x^* (o qualunque approssimazione discreta) non è adatta

Concettualmente occorre pensare a una FUNZIONE DI DENSITÀ DI PROBABILITÀ $f(x)$

tale che $P(a \le X \le b) = \int_a^b f(x)\,dx$

Nell'esempio tempo di reazione potrebbe essere:

$$f(x) = \begin{cases} 0 & x<0 \\ 2e^{-2x} & x \ge 0 \end{cases}$$



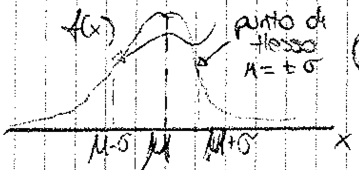
$P(0 \le x \le 2) = \int_0^2 2e^{-2x}dx = -\frac{1}{2}[e^{-2x}]_0^2 = 1 - e^{-4}$

PAGINA 233
n° S 22
errato tipo 3
almeno coinvolge
di più

VARIABILI ALEATORIE NORMALI

Variabile aleatoria x descrivere valori continui (lunghezza, luminosità → GAUSS)

Una variabile aleatoria normale (o gaussiana) ha densità



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

$x \in \mathbb{R}$

Si scrive $x \sim N(\mu, \sigma^2)$

Si dice: x ha distribuzione normale con parametri μ e σ^2

(equivalentemente μ e $\sigma = \sqrt{\sigma^2}$)

Caso particolare: densità normale standard, quando $\mu = 0$, e $\sigma^2 = 1$ (cioè $\sigma = 1$)

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

Le probabilità si ottengono per integrazione della densità:

$P(a \le X \le b) = \int_a^b f(x)dx$

se $f(x)$ è (*) questo integrale non è esprimibile tramite funzioni elementari

2 possibilità < tavole (Nel testo pag 053)
computer (R)

La tavola D.1 contiene per il caso normale standard

$$P(Z \le z) = P(-\infty \le Z \le z) \quad \text{per tanti valori di } z \text{ (x nel testo)}$$

Per l'additività dell'integrale questo ci consente di calcolare qualunque probabilità per la normale standard:

$$P(a \le X \le b) = P(X \le b) - P(X \le a)$$

Nota: se X è una v.a. continua

$$P(X \le a) = P(X < a) + \cancel{P(X = a)} = P(X < a)$$

14-10-09

RIPASSO:

Una variabile aleatoria continua X è dotata di funzione di densità $f(x)$

tale che $P(a \le x \le b) = \int_a^b f(x)dx$ $\qquad \int_{-\infty}^{+\infty} f(x)dx = 1$

Inoltre $E(g(x)) = \int_{-\infty}^{+\infty} g(x) f(x) dx$

Per esempio, se $g(x) = x$ (identità)

$$\mu = E(x) = \int x f(x) dx$$

se $g(x) = (x-\mu)^2$

$$\sigma^2 = Var(x) = \int_{-\infty}^{+\infty} (x-\mu)^2 f(x) dx \quad \text{non sul testo}$$

Una famiglia di v.a. importante è la normale, con densità

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

Si dimostra che se X è normale con parametri μ e σ^2, è proprio

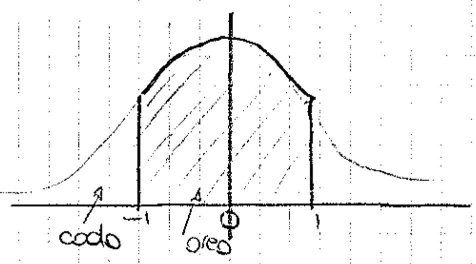
$$E(x) = \mu, \quad Var(x) = \sigma^2 \qquad SD(x) = \sqrt{\sigma^2} = \sigma$$

Esempio:

se $Z \sim N(\overset{\mu}{0}, \overset{\sigma^2}{1})$

$$P(-1 \le x \le 1) = \int_{-1}^{+1} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} = 1 - 2\,\text{coda} =$$

$$= 1 - 2(1 - \text{coda} + \text{area}) \simeq 68\%$$



si fa questo perché le tavole danno le probabilità a sinistra $P(Z \le x)$ per x variabile

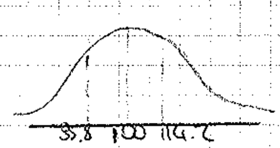
E per una normale non standard? Standardizzo in virtù di un:

TEOREMA: (standardizzazione)

$$\text{se } X \sim N(\mu, \sigma^2) \Rightarrow Z = \frac{X-\mu}{\sigma} \sim N(0,1)$$

Esempio 6.6a:

X = QI di studente delle medie



d) Qual è la probabilità che uno studente scelto a caso abbia punteggio maggiore di 130?

$$P(X>130) = P\left(\frac{X-100}{14.2} > \frac{130-100}{14.2}\right) = P(Z > 2.1127) = 1 - P(Z<2.1127) = 1-0.9826 \simeq 0.017$$

Il teorema si generalizza in quanto trasformazioni lineari di normali sono normali

Per esempio, la somma di due o più normali è normale

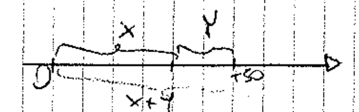
se $X \sim N(\mu_X, \sigma_X^2)$
$\quad Y \sim N(\mu_Y, \sigma_Y^2)$ allora $X+Y \sim N(\mu_X+\mu_Y, ?)$

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

In generale non conosciamo la varianza, ma nel caso particolare in cui X e Y sono indipendenti, allora

$Var(X+Y) = \sigma_X^2 + \sigma_Y^2$ e quindi $\underline{X+Y \sim N(\mu_X+\mu_Y, \sigma_X^2+\sigma_Y^2)}$

Esempio 6.8



$P(X+Y>750)$ = standardizzo

$$= P\left(Z > \frac{750-(\mu_X+\mu_Y)}{\sqrt{\sigma_X^2+\sigma_Y^2}}\right) =$$

$$= P\left(Z > -\frac{50}{\sqrt{3200}}\right) = P(Z > -0.884) \simeq 0.81$$

Per esempio, $z_{0.10}$ taglia una coda a destra uguale a 0.10, e la parte restante a sinistra è 0.90.

$z_{0.10}$ è il 90° percentile

In generale, z_α è il $100(1-\alpha)$-esimo percentile

Uso inverso della tavola D.1: per esempio il 90° percentile della normale standard è $z_{0.10} \approx 1.28$

Se invece $X \sim N(\mu, \sigma^2)$, cerchiamo il suo $100(1-\alpha)$-esimo percentile x_α, cioè $P(X > x_\alpha) = \alpha$

Per ottenere x_α standardizzo: $P\left(\underbrace{\frac{X-\mu}{\sigma}}_{Z} > \underbrace{\frac{x_\alpha-\mu}{\sigma}}_{z_\alpha}\right) = \alpha$

quindi ottengo $\frac{x_\alpha-\mu}{\sigma} = z_\alpha$

cioè $x_\alpha = \mu + z_\alpha\sigma$

ESEMPIO: Il ministro dell'educazione di un certo paese vuole premiare il 10% più bravo degli studenti tramite un test il cui valore atteso è $\mu=800$ e la cui DS è $\sigma=40$

La soglia che il ministro deve usare è

$$x_{0.10}=\mu+z_{0.10}\,\sigma=800+1.28\cdot 40$$

02-11-09

CAMPIONAMENTO (cap7)

- **Es 1** (Controllo di qualità)

Campioniamo $n=10$ pistoni da un processo produttivo e ne misuriamo lo scarto X tra lunghezza effettiva e specifica di progetto $T=50\,mm$

Osservazioni: $x_1,\dots,x_n$

x_i = scarto del pistone i-esimo, proveniente da una distribuzione, per esempio normale $(0,\sigma^2)$

So che la media è 0 perché, per esperienza passata, so che il processo è "centrato" su T

Formalmente, $X_1,\dots,X_n$ sono v.a. indipendenti e con la stessa distribuzione

[v.a. i.i.d. = variabili aleatorie indipendenti, identicamente distribuite]

oppure CAMPIONE CASUALE da una certa distribuzione ($\times$ es. normale $(0,\sigma^2)$)

- **ES 2.** (sondaggio)

Campiono n elettori (in maniera indipendente) e a ciascuno di essi chiedo cosa votano $\begin{pmatrix} X=1 & SI \\ X=0 & NO \end{pmatrix}$

percentuale dei sì nella popolazione

Ottengo $X_1,\dots,X_n$ v.a. i.i.d. Bernoulli (p)

Una STATISTICA è una funzione di $X_1,\dots,X_n$

1) Potrei calcolarne la media: $\bar{X}=\frac{\sum X_i}{n}$ è la STATISTICA MEDIA CAMPIONARIA

2) $S=\sum X_i$ = numero dei sì nel campione

Poiché in un campione casuale $X_1,\dots X_n$ sono variabili aleatorie, anche una statistica è una v.a.

La sua distribuzione è chiamata DISTRIBUZIONE CAMPIONARIA

Consideriamo la statistica $\bar{X}=\frac{\sum X_i}{n}$

Il suo valore atteso è, in generale: $E(\bar{X})=E\left(\frac{\sum X_i}{n}\right)=\frac{\sum E(X_i)}{n}$

ma le $X_1,\dots,X_n$ hanno la stessa distribuzione, quindi lo stesso valore atteso

$$\mu=E(X_1)=E(X_2)=\dots=E(X_n)$$

e quindi $E(\bar{X})=\frac{\sum E(X_i)}{n}=\frac{\sum \mu}{n}=\frac{n\mu}{n}=\mu$

La VARIANZA di $\bar{x}$ è invece

$$Var(\bar{x}) = Var\left(\frac{\sum x_i}{n}\right) = \frac{1}{n^2} \, var(\sum x_i) \overset{\text{x indipendenza}}{=} \frac{1}{n^2} \sum var(x_i) = \frac{1}{n^2} \cdot n\sigma^2 = \frac{\sigma^2}{n}$$

dove $\sigma^2 = var(x_1) = \ldots = var(x_n)$ perché le x_i hanno lo stesso distribuzio

Se $x_1 \ldots x_n$ iid, ciascuna con valore atteso μ e varianza σ^2

$$E(\bar{x}) = \mu$$

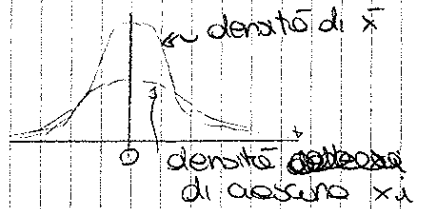
$$Var(\bar{x}) = \frac{\sigma^2}{n} \qquad SD(\bar{x}) = \frac{\sigma}{\sqrt{n}}$$

formule del tutto generali (non dipendono dalla distribuzione delle x_i)

Nel caso normale, cioè se $x_1, \ldots x_n$ iid $N(\mu, \sigma^2)$, allora conosciamo l'intera distribuzione di $\bar{x}$, perché $\bar{x}$ è una combinazione lineare di normali, quindi normale esso stesso. $\bar{x} \sim N(\mu, \frac{\sigma^2}{n})$

1) $x_1, \ldots x_n$ iid $N(0, \sigma^2)$

$$\bar{x} \sim N(0, \frac{\sigma^2}{n})$$



Cosa succede se $x_1, \ldots, x_n$ non sono normali?

Anche se x_i non sono normali, per n grande, approssimativamente

TLC1 $\sum x_i \sim N(n\mu, n\sigma^2)$

TLC2 e quindi $\frac{\sum x_i}{n} \sim N(\mu, \frac{\sigma^2}{n})$

ovvero, approssimativamente,

TLC3 $$\frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \sim N(0, 1)$$

• Esempio 7.3

$\bar{x} = 202$

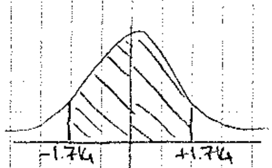
$\sigma = 14$

a) 36 lavoratori

$$P(198 < \bar{x} < 206) = P\left(\frac{198-202}{14/\sqrt{36}} < \underbrace{\frac{\bar{x} - \mu}{\sigma/\sqrt{n}}}_{z \text{ normale}} < \frac{206-202}{14/\sqrt{36}}\right) = P(-1.714 < z < +1.714)$$

(standardizzo)

tabelle



$$= 1 - 2(1 - 0.957) = 0.913$$

b) 64 lavoratori

$$P(198 < \bar{x} < 206) = 0.978$$

2) $x_1 \ldots x_n$ iid. Bernoulli (p)

$S = \sum x_i$

In realtà conosciamo l'intera distribuzione di S

$S \sim Binomiale(n, p)$ cap 5.5

09-11-09

Ripasso:

$\underbrace{x_1 \dots x_n}_{\text{campione casuale}}$ iid da $f(x)$

Statistica = qualsiasi funzione di $x_1 \dots x_n$

es.1 $x_1 \dots x_n$ iid. $N(0, \sigma^2)$ $\bar{x} = \sum x_i / n$

es.2 $x_1 \dots x_n$ iid. Bernoulli (p) $S = \sum x_i$

$\bar{X}$ = media campionaria

TEOREMA DEL LIMITE CENTRALE:

Se $x_1 \dots x_n$ campione casuale con $\mu = E(x_i)$, $\sigma^2 = Var(x_i)$

allora, per n grande, approssimativamente $\sum x_i \sim N(n\mu, n\sigma^2)$

ovvero $\bar{x} \sim N(\mu, \frac{\sigma^2}{n})$ ovvero $\frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$ $\Rightarrow \frac{\frac{\sum x_i}{n} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{\sum x_i - n\mu}{\sigma\sqrt{n}}$

$S = x_1 + \dots + x_n$ è Binomiale(n, p) cioè $P(S = i) = \binom{n}{i} p^i (1-p)^{n-i}$ $i = 0, 1, n$ (risultato esatto)

Risultato approssimato:

se $x_1 \dots x_n$ iid. Bernoulli (p) e n grande, possiamo usare il Teorema del limite centrale (TLC) anche in questo caso

$\mu = E(x_i) = 0 \times (1-p) + 1 \times p = p$

$\sigma^2 = Var(x_i) = p(1-p)$

quindi, approssimativamente,

$S \sim N(np, np(1-p))$

Altri esempi di statistiche

- $S^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$ = funzione del campione $x_1 \dots x_n$
- $\max(x_1, \dots x_n)$ " " "
- Potremmo poi avere 2 campioni, per es. $x_1 \dots x_n$ iid $N(\mu_x, \sigma_x^2)$
  $y_1 \dots y_m$ iid $N(\mu_y, \sigma_y^2)$
- allora statistiche di interesse potrebbero essere

$D = \bar{x} - \bar{y}$ differenza delle due medie campionarie (di interesse nel confronto tra 2 piani)

R_i = numero di y_i maggiori o uguali a x_i (STATISTICA DEI RANGHI

se $n = m$ un'altra statistica è $\underset{\text{correlazione}}{corr(x, y)} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$

Ognuna di queste statistiche avrà la sua distribuzione campionaria

Il TLC copre casi in cui la statistica è una somma o quasi

Per esempio, nel caso di 2 campioni:

$$D = \bar{X} - \bar{Y} = \frac{\sum X_i}{n} - \frac{\sum Y_j}{m}$$

Per il TLC, se n e m sono grandi

$$\frac{\sum X_i}{n} \sim N\left(\mu_X, \frac{\sigma_X^2}{n}\right)$$

$$\frac{\sum Y_j}{m} \sim N\left(\mu_Y, \frac{\sigma_Y^2}{m}\right)$$

Se, in aggiunta, il campione X e il campione Y sono indipendenti, allora, approssimativamente,

$$D \sim N(\,?\,,\,?\,) = N\left(\mu_X - \mu_Y, \frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}\right)$$

$$E(\bar{X} - \bar{Y}) = E(\bar{X}) - E(\bar{Y}) = \mu_X - \mu_Y$$

$$Var(\bar{X} - \bar{Y}) = Var(\bar{X} + (-\bar{Y})) = Var(\bar{X}) + Var(-\bar{Y}) = Var(\bar{X}) + (-1)^2 Var(\bar{Y}) = \frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}$$

- $S^2 = \frac{\sum (X_i - \bar{X})^2}{n-1}$ VARIANZA CAMPIONARIA

Il TLC non ci aiuta direttamente, in quanto abbiamo un quadrato e la presenza di $\bar{X}$ dentro parentesi

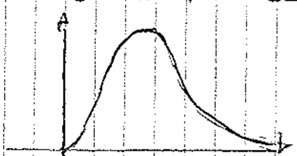
Sez 7.6

Se $X_1, \dots X_n$ sono i.i.d. $N(\mu, \sigma^2)$ la distribuzione della statistica

$$S^2 = \frac{\sum (X_i - \bar{X})^2}{n-1}$$ VARIANZA CAMPIONARIA

si può dedurre dalla distribuzione $\frac{(n-1)S^2}{\sigma^2} = \frac{(n-1)\frac{\sum (X_i - \bar{X})^2}{n-1}}{\sigma^2} \sim \chi^2_{n-1}$

dove χ^2_{n-1} ("chi-quadrato con parametro $n-1$, ovvero con $n-1$ gradi di libertà") è una particolare distribuzione continua con densità di forma



i cui percentili sono tabulati in **D.3** (appendice)

Campionamento da popolazioni finite

Ritorniamo sul campionamento binario $X_1 \dots X_n$ i.i.d. Bernoulli(p)

In realtà, questo campione casuale non può a rigore rappresentare il risultato di un sondaggio da una popolazione finita in quanto non è naturale in un sondaggio prevedere la reintervista della stessa persona. Cioè il sondaggio è "campionamento senza reintroduzione nell'urna".

Tecnicamente le X_i non sono indipendenti

TUTTAVIA, se supponiamo che l'urna (cioè la popolazione) sia composta da N elementi, con $N >> n$ (N molto + grande del campione di n elementi)

Quindi lo schema ipergeometrico può essere tranquillamente sostituito con lo schema binomiale $X_1 \dots X_n$ i.i.d. Bernoulli(p)

$\sum X_i \sim$ Binomiale(n, p)

(se invece N non è molto più grande di n, necessitiamo di fattori di correzione)