Che cosa è la statistica? È la disciplina che si occupa dell'elaborazione dei risultati derivanti dall'osservazione di uno o più caratteri posseduti dagli elementi di un insieme, con l'intento di esprimere un giudizio sul "fenomeno collettivo". - Studio di fenomeni collettivi con variabilità.

FENOMENI DETERMINISTICI: data la legge di evoluzione del sistema e le sue C.I., posso prevedere il comportamento futuro con precisione

FENOMENI ALEATORI: → "random" che non possono essere previsti con precisione, ma possono essere stimati (stima probabilistica). Questi fenomeni vengono descritti attraverso variabili aleatorie

Un sistema dinamico presenta n var. in ingresso, m in uscita, e k variabili che ne descrivono lo stato interno. 3 TIPI:

- SIST. DIN. DETERMINISTICO: è immediato valutare le V. in USCITA date quelle in IN
- SIST. DIN. STOCASTICO: *
- SIST. DIN. CAOTICO: caratterizzato da leggi non lineari (anche se semplici) che con piccole variazioni degli ingressi determina una grande variazione delle uscite

* non è individuabile una corrispondenza biunivoca tra ingresso e uscita, ed è quindi necessaria una descrizione probabilistica

Scopo della statistica: studiare un fenomeno che abbia una certa variabilità nei dati.

STATISTICA DESCRITTIVA = contiene tutti metodi utili a sintetizzare con poche grandezze una mole di dati (fornisce le formule)

INDICI:

MASTER COPY
Tel. 050 8312126
Cell. 388 9837745

MISURE DI POSIZIONE: o tendenza centrale

- MEDIA: $\bar{x} = \frac{\sum_i x_i}{N}$
- MODA: punto max della distribuzione (valore più frequente)
- MEDIANA: Valore sotto al quale cadono la metà dei valori campionari; dati i dati in ordine crescente, si prende quello nella posizione centrale (se campione con dati dispari) o la media dei 2 centrali (se campione con dati pari)

MISURE DI DISPERSIONE: valutano la distribuzione nelle code

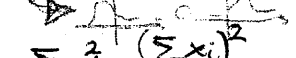
- VARIANZA $s^2 = \frac{\sum_i (x_i - \bar{x})^2}{N-1}$
- DEVIAZIONE STANDARD $s = \sqrt{s^2}$
- RANGE $R = X_{MAX} - X_{MIN}$ ha il difetto di basarsi solo sul 1° e ultimo elemento che sono i più soggetti ad errori di misura o variabilità del campionamento

(SCARTO QUADRATICO MEDIO) vedi dopo

- SCARTO INTERQUARTILE
- DEVIANZA $D = \sum (x_i - \bar{x})^2$ *

MISURE DI FORMA

- SKEWNESS (coeff. di asimmetria) $\frac{\sum_i \left(\frac{x_i - \bar{x}}{\sigma}\right)^3}{N}$ $\begin{cases} >0 \text{ coda a dx} \\ <0 \text{ " " sx} \\ =0 \text{ simmetria} \end{cases}$
- CURTOSI (quanto è appuntita la dist.) $\frac{\sum_i \left(\frac{x_i - \bar{x}}{\sigma}\right)^4}{N}$ $\begin{cases} <3 \text{ molto appuntita} \\ =3 \text{ per la gaussiana} \\ >3 \text{ poco appuntita} \end{cases}$



* $D = \sum x_i^2 - \frac{(\sum x_i)^2}{n}$

Si dice CAMPIONE un aggregato di unità statistiche appartenenti ad una determinata popolazione, selezionate attraverso un esperimento di campionamento.

(se la pop. virtualmente è ∞, spesso la pop ~~desiderata~~ campionata è ≠ da quella obiettivo)

"fare inferenza" ≡ estendere le info derivate da un campione su tutta la pop. (in ultima → "sperimentare", "descrivere")

⇓

STATISTICA INFERENZIALE (questa estensione non è però certa)

↳ ha alla base la statistica descrittiva; che in primis (sintetizza i dati)

→ ha come scopo finale il calcolo della probabilità, in quanto il processo di inferenza non è certo ma ha sempre un grado di probabilità.

PROBABILITÀ $P[i] = \lim_{N\to\infty} \frac{m_i}{N}$ ← casi favorevoli dell'evento i / ← casi totali

Vero ⇐⇒ tutti i casi sono equiprobabili

DEFINIZIONE: è costituita da una terna di elementi:

- Ω è un insieme ≡ SPAZIO CAMPIONE = insieme di tutte le possibili uscite delle v. osservate e le loro combinazioni
- A è una σ-algebra di Ω, cioè contiene tutti i possibili sottoinsiemi di Ω compreso quello nullo e quello totale
- P è la probabilità se A → $\cup$ elem. di A → $\in$ A; $\cap$ " " A → $\in$ A

Allora la probabilità è un insieme di funzioni matematiche $P: A \to \mathbb{R}^+ = [0;1]$, con delle assunzioni, cioè $P(\Omega)=1$, $P(\omega_i)=p_i$ e

Se i dati sono disgiunti (cioè non si influenzano a vicenda) → eventi mutuamente esclusivi → p. 3

↳ $P\left(\bigcup_{m=1}^{\infty} A_m\right) = \sum_{m=1}^{\infty} P(A_m)$

ASSIOMI:

- la probabilità è sempre ≥ 0
- $P(\Omega)=1$
- $P(A \cup B) = P(A) + P(B)$ se $A \cap B = \emptyset$ (dati disgiunti)
  - → se $A \cap B \neq 0$, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ (altrimenti conterei 2 volte l'intersezione)

PROB. CONDIZIONATA: prob di A, conoscendo già B

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

ANALISI ↔ SEMPLICE (1 o più campioni) / DATI ACCOPPIATI / MULTIVARIATA

Viene utilizzato il metodo scientifico (ipotetico-deduttivo)

osservazione → ipotesi → predizione del fenomeno basata sull'ipotesi → test della predizione ↔ esito positivo → ipotesi confermata / esito negativo → ipotesi rigettata.

H_0 (ipotesi nulla) le differenze osservate sono frutto del caso

H_1 (" alternativa) le differenze osservate sono tali da dire che i campioni ∈ a 2 pop. diverse

MASTER COPY
Tel. 050 8312126
Cell. 388 9837745

Ho vi è in ogni caso un certo grado di incertezza (grado di probabilità)

VARIABILI → QUALITATIVE → nominali (non c'è un ordine)
ordinali (c'è un ordine)
→ QUANTITATIVE → DISCRETE
CONTINUE.

2

PROPRIETÀ DI MUTUA ESCLUSIVITÀ se $A \cap B = \phi$

u

DI ESAUSTIVITÀ COLLETTIVA se l'unione degli eventi da lo spazio campione

CONSEGUENZE DEGLI ASSIOMI:

- se ho $P(A)$, $P(A^c) = 1 - P(A)$ per es. slide 16
- se $A \cap B \neq 0$, $P(A \cup B) = P(A) + P(B) - P(A \cap B) \Rightarrow P(A \cap B) = P(A) + P(B) - P(A \cup B)$

→ in generale $P(E \cup F \cup G) = P(E) + P(F) + P(G) - P(EF) - P(EG) - P(FG) - P(EFG)$

→ se 2 eventi sono indipendenti (uno non influenza l'altro) $\Rightarrow P(A \cap B) = P(A) \cdot P(B)$

→ se non indipendenti $P(A \cap B) = P(B|A) \cdot P(A) = P(A|B) \cdot P(B)$

TEOREMA DELLA PROBABILITÀ TOTALE: Dato un insieme di eventi ($B_i, \dots B_n$) mutuamente esclusivi e collettivamente esaustivi, la probabilità di un evento A è: $P(A) = \sum_i P(A \cap B_i) = \sum_i P(A|B_i) P(B_i)$

il t. delle p. totali ⇒ la rappresentazione ad albero

TEOREMA DI BAYES Combinando $P(A \cap B) = P(B|A) \cdot P(A)$ e $P(A \cap B) = P(A|B) \cdot P(B)$ si ottiene

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

$P(A|B)$: conoscenza a posteriori riguardo la verità di un ipotesi, date le osservazioni

$P(B|A) \cdot P(A)$: conoscenza a priori riguardante l'ipotesi

$B = \{\text{osservazioni}\}$

$A = \{\text{ipotesi}\}$

$P(B)$: costante di normalizzazione che non dipende dalle ipotesi

$P(B|A)$: osservazioni che ottengo sotto l'ipotesi

Si riesce quindi a determinare la veridicità di una ipotesi partendo dall'esperienza → approccio Bayesiano.

* $P(B) = P(A) P(B|A) + P(A^c) P(B|A^c)$

ESEMPI SLIDE 35-40

CAMPIONAMENTO (n elem. da insieme di M elem.)

- CAMP. SENZA RIPETIZIONE e CON ORDINE: $\frac{M!}{(M-n)!}$ disposizioni semplici
- CAMP. SENZA RIPETIZIONE e SENZA ORDINE $\binom{M}{n} = \frac{M!}{n!(M-n)!}$
- CAMP CON RIPETIZIONE e CON ORDINE M^n

reimbussolamento

$$P(S) = \frac{n!(M-n)!}{M!}$$

[3] VARIABILI ALEATORIE = è una entità matematica i cui valori sono governati dal caso, cioè seguono le leggi della teoria delle probabilità

Allora la probabilità è quanto frequentemente otteniamo valori diversi riguardanti la stess. v.a se si ripete più volte l'esperimento.

V. CONTINUE: hanno un numero ∞ di valori (es concentrazione colesterolo nel sangue)

V. DISCRETE: hanno un numero finito di valori (es. dado)

PER LE V.A. CONTINUE:

- $\int_{-\infty}^{+\infty} f(x)dx = 1$ $f(x)$ è la funzione densità di probabilità (PDF), (non negativa)
  - ↳ corrisponde all'area sottesa la curva
  - → [se $\int_{-\infty}^{+\infty} f(x)dx = 1$, allora la probabilità dell'osservazione di quella variabile tra a e b è $\int_a^b f(x)dx = P(a \le X \le b)$
  - → $[P(a \le X \le a) = P(x=a) = \int_a^a f(t)dt = 0]$
  - → $[F(x) = \int_{-\infty}^{a} f(x)dx = P(X \le a)$ è la f. densità di probabilità cumulativa
  - → $f(x) = \frac{dF(x)}{dx}$

ASPETTAZIONE e MOMENTI → per momenti di ordine superiore $E[X^m] = \int_{-\infty}^{\infty} x^m f_x(x)dx$

MOMENTI CENTRALI: momento della v.a a cui è sottratto il valor medio

V.A. CONTINUE

$E(X) = \int_{\infty x} x_i \underbrace{p(x_i)}_{\text{prob.}} dx$

V.A DISCRETE

$E(X) = \sum_{\infty x} x_i p(x_i)$

> in generale $E[g(x)] = \int_{-\infty}^{+\infty} g(x) f(x) dx$

- MEDIA $\mu = \int_{-\infty}^{+\infty} x f(x)dx$
- VARIANZA (*) $\sigma^2 = \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)dx = E([X-E(X)]^2)$ (*) p. 10
- SKEWNESS $\gamma = \frac{1}{\sigma^3}\int_{-\infty}^{+\infty}(x-\mu)^3 f(x)dx = \frac{E[[X-E(X)]^3]}{\sigma^3}$

(*) $\sigma^2 = \int_{\infty x}(x_i-\mu)^2 p(x_i)dx = E[X^2] - (E[X])^2$ dice quanto è stretta o larga $f(x)$ rispetto al valor medio

- DEV. STANDARD $\sigma = \sqrt{\sigma^2}$

Proprietà di linearità dell'operatore aspettazione:

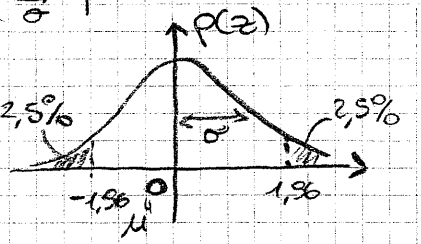
- → $E(X+Y) = E(X) + E(Y)$
- → $E(aX+bY) = aE(X) + bE(Y)$
- !!! → $Var(aX) = a^2 Var(X)$ dimo su (5) slide 33
  - → se $Y = X+b \Rightarrow Var(Y) = Var(X)$

in generale vale: (SOMMA DI V.A) $Var(X+Y) = Var(X) + Var(Y) + \overbrace{2E([X-E(X)][Y-E(Y)])}^{2Cov(X,Y)}$

↳ se sono indipendenti $Var(X+Y) = Var(X) + Var(Y)$ (perché $Cov = \emptyset$)

MASTER COPY
Tel. 050 8312126
Cell. 388 9837745

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



- $P(\mu-\sigma \le X \le \mu+\sigma) \simeq 0{,}63$
- $P(\mu-2\sigma \le X \le \mu+2\sigma) = 0{,}954$
- $P(\mu-3\sigma \le X \le \mu+3\sigma) = 0{,}997$

GAUSSIANA STANDARDIZZATA: media nulla e dev. stand. unitaria

$f(x) = p(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \Rightarrow \left(z = \frac{X-\mu}{\sigma}\right) \equiv$ STANDARDIZZARE UNA VARIABILE

[$z_{0,05} = 1{,}96$ cioè in corrispondenza di 1,96 a dx vi è il 2,5% della tabella e a sx di -1,96 vi è il 2,5% → in totale 5%]

Stiamo che lascia nelle 2 code una probabilità del 5% (2,5% a dx e 2,5% a sx)

$z_{0,02} = 2{,}33$; $z_{0,01} = 2{,}58$

Per la v.a continue:

- DISTR CONTINUA: tutti i valori della variabile sono associati alla stessa prob.
- DISTR. ESPONENZIALE $f(t) = \lambda e^{-\lambda t}$ per $t \ge 0$; $\mu = \frac{1}{\lambda}$ e $\sigma^2 = \mu^2 = \frac{1}{\lambda^2}$
- DISTR LOG-NORMALE data la produttoria di v.a. i.i.d. $Y = W_1 \cdot W_2 \cdot W_3 \cdot \ldots \cdot W_n$, e considero $X = \ln Y$, X è una distr. normale e Y è una distr. lognormale
- DISTR. BETA. utilizzata quando la va assume un range di valori. $f_y(Y) \propto y^a (1-y)^b$; se $a=b=0 \Rightarrow$ diventa una distr. uniforme
- DISTR. GAMMA $\Gamma(\alpha) = \int_0^\alpha x^{\alpha-1} e^{-x} dx$

si trasforma in una distr. esponenziale quando $\alpha = 1$ o in una distr. χ^2 quando $\alpha = \frac{V}{2}$ e $\beta = 2$ $\left(f(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}}\right)$

- DISTR. CHI-QUADRO (ha solo 1 coda) Siano $z, z_1, z_2, \ldots z_m$ v.a. i.i.d. normali standardizzate (z), allora $\chi^2_m = \sum_{i=1}^{m} z_i^2$ è la d. χ^2 con $n = gdl = \mu$ (i gdl corrispondono alla media) = numero di variabili

→ la somma di 2 v.a. chi-quadro a n_1 e n_2 gdl è una v.a. chi-quadro con n_1+n_2 gdl

⊕ per $n > 30$, $\chi^2_\alpha(n) = \frac{\left(z_\alpha + \sqrt{2n-1}\right)^2}{2}$ con z_α percentile di ordine α di Z (v. norm. stand.)

- DISTR T-STUDENT (ha 2 code) Siano $z, z_1, z_2, \ldots z_m$ v.a. i.i.d. normali standardizzate, allora $t_n = \frac{Z}{\left(\frac{1}{n}\sum_{i=1}^{m} z_i^2\right)^{1/2}}$ è la dist. t con n gdl

→ $E[t_n] = 0$ e $Var(t_n) = \begin{cases} \frac{n}{n-2} & \text{se } n > 2 \\ 0 & \text{se } n \le 2 \end{cases}$ SIMMETRICA E CENTRATA IN 0

- all'aumentare di n la distribuzione t tende ad una gaussiana standardizzata (le code sono più marcate)

⊕ per $n > 30$ si approssima con una distribuzione normale

MASTER COPY Tel. 050 8312126 Cell. 388 9837745