36 Gli stimatori puntuali

IL PRIMO COMPITO DELLA STATISTICA INFERENZIALE: DISTINGUERE FRA VARIAZIONI SIGNIFICATIVE E NON SIGNIFICATIVE: PER ESEMPIO PER DISTINGUERE
PLAUSIBILI EFFETTI CASUALI
DA
SPERABILI EFFETTI CAUSALI
DEI FARMACI.

36.1 Citazione da un testo importante

BOZZA DRAFT

Il presente testo, nonostante le sue centinaia di pagine, è una sorta di "microbo che fluttua in una piastra di Petri".
Il Lettore potrà dover approfondire determinate tematiche, nella sua eventuale futura professione. Esistono moltissimi testi validi.

Leggiamo([185]) in "Matematica per le scienze della Vita":

> Capire i dati è un processo che richiede diversi passi:
>
> 1. raccogliere i dati;
> 2. riassumere dati;
> 3. analizzare i dati;
> 4. interpretare i risultati e presentarli.
>
> (...) La fase 2 del processo (riassumere i dati) è quella della "statistica descrittiva", in cui l'obiettivo è di astrarre dai dati alcune proprietà per poterli meglio interpretare.
> (...) La fase 3 del processo coinvolge tipicamente l'area della "statistica inferenziale", che consiste nella stima dei parametri e nel testare le ipotesi.

[185] Di Erin N. Bodine, Suzanne Lenhart, Louis J. Gross, a cura di Gabriella Caristi, Maurizio Mozzanica, Giacomo Tommei, 2017, UTET Università, titolo originale Mathematics for the Life Sciences, trad. Patrizia Ferreri

Per quanto riguarda la stima dei parameteri, vedremo gli stimatori, Lezione (36.6), e gli intervalli di fiducia, Lezioni (37), (47). In Lezioni successive vedremo i test statistici.

36.2 Introduzione

Se abbiamo un dado appena comprato da un rivenditore di cui ci fidiamo, lo riteniamo regolare, cioè riteniamo equiprobabili i suoi risultati, e similmente per una moneta nuova di zecca:

$$X := \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix}^{\textit{dado equilibrato}}_{\textit{ovvero regolare}} \qquad Y := \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}^{\textit{moneta equilibrata}}_{\textit{ovvero regolare}}$$

Su queste basi con il Calcolo delle Probabilità possiamo fare moltissime affermazioni certe sui suoi risultati comunque incerti:
– per esempio che conviene scommettere che la somma di 2 lanci del dado sarà 7 piuttosto che 8,
– e conviene scommettere che su un milione di lanci della moneta il numero di teste sarà fra 499 000 e 501 000 piuttosto che no.
Questo è appunto il Calcolo delle Probabilità, e abbiamo visto quante implicazioni ha nella Farmacia.

BOZZA - DRAFT

Tuttavia, gli effetti di un farmaco sono ben lontani da questo tipo di regolarità che hanno il dado e la moneta equilibrati.

Qua entra in campo la Statistica Inferenziale. Mentre nel Calcolo delle Probabilità si parte da una distribuzione e si trovano molte affermazioni conseguenti, come quelle sopra riportate, nella Statistica Inferenziale si suppone noto il tipo di distribuzione, per esempio binomiale o normale o log-normale, ma si ignorano i valori dei parametri della distribuzione, e si vogliono fare affermazioni ragionevoli su essi. Specialmente sulla media e la varianza.

Per esempio potremmo avere un dado di forma irregolare, fatto con un astragalo, o giocare a testa e croce con un frisbee, e ci ritroviamo con distribuzioni di tipo noto ma con parametri incogniti:

$$X := \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ p_1 & p_2 & p_3 & p_4 & p_5 & 1 - p_1 - p_2 - p_3 - p_4 - p_5 \end{pmatrix}$$

$$Y := \begin{pmatrix} 0 & 1 \\ p & 1 - p \end{pmatrix}$$

La statistica inferenziale vuole fare affermazioni ragionevoli sui parametri di distribuzioni per il resto note.

Per esempio, vogliamo affermare ragionevolmente se p è $\frac{1}{2}$ oppure no, che potrebbe corrispondere al farmaco che non fa nè bene nè male, oppure no, qualcosa fa.
Solo ragionevolmente parlando.
La certezza sfugge inesorabilmente alla Statistica Inferenziale.

Tali affermazioni ragionevoli verranno fatte non sulla base della simmetria com'è per dadi e monete, che fa ritenere l'equiprobabilità dei casi possibili, bensì sulla base di molti valori sperimentali ovvero empirici (determinazioni) della variabile aleatoria oggetto di indagine – cioè sulla base di una determinazione (storica, realmente avvenuta nel tempo) di un campione aleatorio. Cioè, lanciamo molte volte l'astragalo o il frisbee, o diamo il farmaco a molti soggetti, e conteggiamo i risultati.

BOZZA - DRAFT

Così spiega Wikipedia, l'enciclopedia libera, alla voce *inferenza statistica*:

> Si considereranno principalmente campioni casuali semplici di dimensione $n > 1$, che possono venire interpretati come n realizzazioni indipendenti di un esperimento di base, nelle medesime condizioni. Dal momento che si considera un esperimento casuale, si coinvolge il calcolo delle probabilità. Nell'inferenza statistica c'è, in un certo senso, un rovesciamento di punto di vista rispetto al calcolo delle probabilità. Nell'ambito di quest'ultimo, noto il processo di generazione dei dati sperimentali (modello probabilistico) siamo in grado di valutare la probabilità dei diversi possibili risultati di un esperimento. Nella statistica il processo di generazione dei dati sperimentali non è noto in modo completo (il processo in questione è, in definitiva, l'oggetto di indagine) e le tecniche statistiche si prefiggono di indurre le caratteristiche di tale processo sulla base dell'osservazione dei dati sperimentali da esso generati.

Dovremo però cercare di ragionevolmente distinguere il plausibile effetto casuale, dallo sperabile effetto causale.
Per esempio è ben possibile che una moneta o frisbee pur regolare abbia fatto 3 teste su 4 lanci, ma se dà testa 300 volte su 400, sarà ragionevole ritenere che no, non sia regolare.
Fra il plausibile effetto casuale e lo sperabile effetto causale una certa linea di confine viene demarcata da formule basate sulle funzioni del chi quadrato, della t di student, della gaussiana, e altre.
La linea di confine resta sottile e problematica.

Così ancora spiega Wikipedia, l'enciclopedia libera, alla stessa voce *inferenza statistica*:

> Nell'ambito dell'inferenza statistica, si distinguono due scuole di pensiero, legate a diverse concezioni, o interpretazioni, del significato della probabilità:
>
> - Inferenza classica, o frequentista;
> - Inferenza bayesiana.
>
> La prima è legata agli storici contributi di R. Fisher, K. Pearson, e rappresenta la posizione maggioritaria. La seconda, allo stato attuale (2005) ancora minoritaria, ma in crescita, è fondata sull'uso del risultato del teorema di Bayes ai fini dell'inferenza statistica.

BOZZA - DRAFT

Questa trattazione elementare segue il primo modello.

36.3 Lo scarto di una radice quadrata dalla metà

Il seguente risultato avanzato di Calcolo delle Probabilità ci traghetterà nella Statistica Inferenziale.

Teorema. C'è circa il 5% di probabilità che lanciando una moneta regolare il numero di teste disti dalla metà di un *grande* numero dei lanci più della radice quadrata del numero dei lanci, ovvero, che stia entro quello scarto, circa 95%:

$$\boxed{P\left(\frac{n}{2}-\sqrt{n} \leq \text{ numero di teste } \leq \frac{n}{2}+\sqrt{n}\right) \approx 95\% \qquad n \gg} \tag{140}$$

e più precisamente di poco maggiore di 95.4% per n molto grande.

Insomma, con 10 000 lanci, fra 4900 e 5100: così è chiarissimo.

36.4 Primo esempio completo di Statistica Inferenziale

Consideriamo il contatore X_n di teste in n lanci di moneta regolare. Useremo il risultato del paragrafo precedente, che adesso scriviamo così:

$$P\left(\left|X_n - \frac{n}{2}\right| \leq \sqrt{n}\right) > 95.4\% \qquad \forall\, n > 4. \tag{141}$$

Questo ci permette di fare una prima puntatina nella statistica della Farmacia.

BOZZA - DRAFT

Supponiamo di avere una popolazione di monete regolari, e consideriamo malattia la loro regolarità, cioè il fare testa con probabilità $\frac{1}{2}$.
Ne prendiamo un bel campione per fare un test terapeutico, diciamo 10 000. (Il numero di soggetti ragionevole da considerare è questione statistica non banale; qua scegliamo 10 000).
Le curiamo con una goccia di farmaco (per esempio una goccia di colla al centro dalla parte della testa) e ci chiediamo se la cura ha funzionato, cioè se adesso sono meno regolari. Cioè ci chiediamo se le abbiamo curate dalla loro regolarità.
Lanciamo le 10 000 monete 1 volta ciascuna ottenendo 10 000 risultati indipendenti e diciamo $X_{10\,000}$ il numero di teste.
Noi vorremmo che questo numero sia molto diverso 5 000, cioè la metà di 10 000.
Supponiamo di ottenere 5 100: possiamo essere soddisfatti?
(5100 è proprio $n + \sqrt{n}$).
Sì perché se le monete erano ancora regolari c'era meno del 4.6% di probabilità (100% − 95.4%) che venisse un risultato così estremo, lontano dalla metà. (Siamo soddisfatti ma di poco, se veniva 5 200 o magari 6 000 era molto meglio).
La soglia ordinaria di *significatività* della Statistica è il 5%.

Se ne otteniamo 5 050 non siamo soddisfatti: una tale deviazione è ben compatibile con la regolarità.
Questa è solo un'idea iniziale: nella pratica scientifica seria, bisogna fissare la soglia prima di fare l'esperimento.

36.5 Secondo esempio di Statistica Inferenziale

Durante una guerra viene costruito per 20mila prigionieri un campo di detenzione che ha due ali con ugual numero di detenuti: in una i detenuti bevono l'acqua di un pozzo, e nell'altra ala l'acqua di un altro pozzo. Non c'è nessun problema etico e nessuna malevolenza nel dare l'acqua dei due diversi pozzi nelle due ali; questo è quello che si chiama un *esperimento naturale*, che si forma da sè. Ad un certo punto sono ormai morti metà dei prigionieri, e nello stesso giorno gli altri vengono liberati e un tecnico del personale analizza l'acqua dei due pozzi e scopre che in uno c'è molto nickel, o qualunque altra sostanza, ma diciamo nickel per fissare le idee, e nell'altro non c'è (in nessuna misura rilevabile). A questo punto il medico del campo ipotizza che il nickel possa influire significativamente sulla sopravvivenza ovvero la mortalità, seppure non sa se favorevolmente o sfavorevolmente. Scopre che 5050 sono morti in un'ala e 4950 nell'altra, e uno potrebbe dire che sì, c'è un forte indizio che il nickel abbia influenzato la sopravvivenza ovvero la mortalità. Questa sarebbe stata la logica conclusione di uno studioso dei secoli passati. Invece la moderna Statistica Inferenziale dice che no, il risultato non è *statisticamente significativo*. Lo scarto di 50 da 5000 è troppo piccolo. Cè una probabilità circa del 5% che il valore cada fuori di $[4900, 5100]$, e ben maggiore del 5% che cada fuori di $4951, 5049$, com'è successo, anche se è $\frac{1}{2}$ la probabilità che ogni singolo deceduto provenisse da una delle due ali del campo.
La moderna Statistica Medica viene fatta in generale proprio così, “al 95%" (di *confidenza*, come si dirà).

Nota sui due esempi. Si noti però che in generale nella Far-

BOZZA - DRAFT

macia in generale con un farmaco si vuole sbilanciare un parametro in una fissata direzione, (si pensi alla glicemia troppo alta o alla sideremia troppo bassa), non semplicemente smuoverlo in una direzione qualunque.
Questo complicherà le cose. Per intanto abbiamo capito essenzialmente cosa vuol dire distinguere variazioni significative da variazioni non significative: queste ultime *potrebbero ben essere* (qui sta il 95%) *fluttuazioni casuali.*

Con l'approssimazione normale si può dimostrare che, purché n sia sufficientemente grande, lo sbilanciamento di teste rispetto alla metà in una direzione o nell'altra

$$\left|X_n - \frac{n}{2}\right|$$

con probabilità $\approx 68.3\%$ è $< \frac{1}{2}\sqrt{n}$,
con probabilità $\approx 27\%$ sta fra $\frac{1}{2}\sqrt{n}$ e $\sqrt{n}$,
con probabilità $\approx 4.6\%$ è $> \sqrt{n}$

Riscriviamo l'ultima:

$$P\left(\left|X_n - \frac{n}{2}\right| > \sqrt{n}\right) \approx 4.6\% \qquad \forall\, n\, molto\, grande \tag{142}$$

strettamente imparentata alla (141) perché 4.6 è 100-95.4.
Nota. In realtà dovremmo fare un *test statistico* un po' diverso, e più complicato, perchè vogliamo dimostrare lo sbilanciamento in una direzione specifica. Tuttavia, nella pratica Biomedica spesso si fa comunque il *test bilatero* anche se il *test unilatero* sarebbe più appropriato, perchè quello bilatero è più "standard".

Nota. Di questa Lezione ci si dovrebbe aspettare che lo studente conosca a memoria la formula (142), oltre a tutto l'inquadramento teorico della questione.

BOZZA - DRAFT

36.6 Stimatori e stimatori non distorti

36.7 Parametri e stimatori; stimatori non distorti

Supponiamo che fra poco avremo le altezze di $n := 100$ persone, prese a caso dai 60milioni di italiani, e allora l'altezza di un italiano a caso è una variabile aleatoria X, e gli n numeri che stiamo per avere sono essi stessi variabili aleatorie X_1,...,X_n, e dopo che li avremo saranno numeri $x_1, ..., x_n$, detti *determinazioni* della variabile aleatoria. Chiameremo anche *campione aleatorio* le variabili aleatorie X_1,...,X_n e *campione* i numeri $x_1, ..., x_n$, ma non ci formalizzeremo su questa distinzione.

Ottenere da n numeri una *stima* di un *parametro* incognito di una v.a. ovvero della sua legge, è un *problema di stima.* Ogni funzione

$$h(X_1, ..., X_n)$$

(per esempio $\max(X_1, .., X_n)$) di un campione aleatorio si dice *stimatore*, in generale stimatore di un parametro incognito u della densità della variabile aleatoria, variabile aleatoria che sappiamo esistere ma non conosceremo mai in forma esatta (legge ovvero distribuzione, cioè densità o funzione di ripartizione), ma di cui avremo n determinazioni. Scriveremo

$$\hat{v} := h(X_1, ..., X_n)$$

e diremo che $\hat{v} = h(X_1, ..., X_n)$ (variabile aleatoria) è uno stimatore di v e $\hat{v} = h(x_1, ..., x_n)$ (numero) è una stima di v.

Rimarchiamo che lo stimatore è una variabile aleatoria, funzione di un numero indeterminato n di variabili aleatorie. Certo, dopo che delle variabili aleatorie del campione avremo una determinazione, immediatamente avremo una determinazione dello stimatore, che a quel punto sarà un numero.

BOZZA - DRAFT

Esempio duplice di stimatore. Per esempio per una v.a. $\mathbb{U}[0,a]$ o $\mathbb{U}\{0,a\}$ (uniforme fra 0 e a, continua o rispettivamente discreta)

$$\hat{a} := 2\,\frac{X_1 + ... + X_n}{n} \qquad = 2\,\bar{X}_n \tag{143}$$

è lo *stimatore dei momenti* del parametro a. (Il doppio della media aritmetica).

Per esempio se 3 persone si sono iscritte indipendentemente a una corsa con molti partecipanti e hanno avuto i numeri progressivi n_1, n_2, n_3, il doppio della media – in assenza di altre informazioni – è una stima del numero di iscritti. (Approssimativamente: in effetti i 3 numeri si possono ritenere determinazioni di una v.a. $\mathbb{U}\{1,a\}$ e non $\mathbb{U}\{0,a\}$, e inoltre non c'è una perfetta indipendenza perchè noi sappiamo che $n_2 \neq n_1$, però le 2 imperfezioni non pesano molto perchè ci sono molti partecipanti).
Per esempio per

Quello della formula data non è l'unico stimatore per a, ne esistono infiniti, ma ben pochi sensati. Un altro considerato molto sensato per $\mathbb{U}[0,a]$ è il massimo dei valori del campione, $\hat{a} := \max(X_1, ..., X_n)$, di cui non ci occuperemo.

BOZZA - DRAFT

ESERCIZIO$_\mu$ $\approx$
Dato questo campione aleatorio

$$0.5288, 0.0344, 0.6112, 0.072, 0.4584, 0.5104, 0.7296$$

di una variabile aleatoria Z uniformemente distribuita sull'intervallo $[0,a]$, stimare a con lo stimatore dei momenti, con 4 cifre decimali.

SVOLGIMENTO
Viene usato lo standard del punto decimale.

Lo stimatore dei momenti è il doppio della media aritmetica.
(Ovvero in formule, per una v.a. Z, $\hat{a} := 2\,\frac{Z_1+...+Z_n}{n} = 2\,\bar{Z}_n$).

La media degli $n = 7$ numeri è

$$\bar{Z}_7 = \frac{0.5288 + 0.0344 + 0.6112 + 0.072 + 0.4584 + 0.5104 + 0.7296}{7} \approx 0.420686$$

e moltiplicando per 2, si ha, con 4 cifre decimali,

$$\hat{a} = 2\,\bar{Z}_7 =$$

$$\boxed{\hat{a} \approx 0.8414}$$

Nota. (Il campione era stato ottenuto con $a = 0.8$ con l'istruzione

```
v = Table[0.8 RandomInteger[1000]/1000., i, 1, 7]
```

col software Mathematica$^{(R)}$).

BOZZA - DRAFT

Ulteriore esempio. Per una v.a. uniforme $\mathbb{U}[-u, u]$ lo *stimatore dei momenti* di u è

$$\hat{u} := \sqrt{\frac{3\,(X_1^2 + ... + X_n^2)}{n}}$$

36.8 Stimatori non distorti

Esistono molti criteri di bontà di uno stimatore, e noi ne consideremo uno: diremo che lo stimatore $\hat{u}$ del parametro incognito u è *non distorto* (o *corretto*; in inglese *unbiased*) se la speranza matematica di $\hat{u}$ è u, cioè $E(\hat{u}) = u$.

36.9 Stimatori non distorti di media e varianza

La differenza basale fra la *statistica descrittiva* e la *statistica inferenziale* è che la prima opera (in generale) su numeri e la seconda su variabili aleatorie, con il che la prima ricade in quella che abbiamo chiamato *matematica della certezza* e la seconda nella *matematica dell'incertezza*.
Se abbiamo n numeri

$$x_1, x_2, ..., x_n$$

nella statistica descrittiva queste sono la media e la varianza degli n numeri:

$$\bar{x} := \frac{1}{n}\sum_{k=1}^{n} x_k \qquad Var(x_1, x_2, ..., x_n) := \frac{1}{n}\sum_{k=1}^{n}(x_k - \bar{x})^2 \,.$$

Se invece supponiamo che $x_1, x_2, ..., x_n$ siano *determinazioni* di una variabile aleatoria X con una certa media μ e una certa varianza σ^2 *incognite*, dalle precedenti formule possiamo immediatamente definire questi (ragionevoli) stimatori – che sono variabili aleatorie – di μ e σ^2:

$$\hat{\mu} := \bar{X} := \bar{X}_n := \frac{1}{n}\sum_{k=1}^{n} X_k \tag{144}$$

$$W := \frac{1}{n}\sum_{k=1}^{n}(X_k - \bar{X}_n)^2 \,.$$

BOZZA - DRAFT

Si dimostra che il primo è stimatore *non distorto* della media μ, cioè $E(\bar{X}_n) = \mu$, ma il secondo non è stimatore non distorto di σ^2, cioè la sua speranza matematica è diversa dal parametro che si vuol stimare.
Si dimostra invece che (e si noti la differenza nel denominatore)

$$S^2 := \frac{1}{n-1}\sum_{k=1}^{n}(X_k - \bar{X})^2 \ \left[:= \hat{\sigma}^2 \ _{\leftarrow \textit{scrittura rara, scriveremo } S^2} \right] \tag{145}$$

è lo *stimatore non distorto della varianza*: $E(S^2) = \sigma^2$.

In Farmacia e Medicina. Per esempio dalle altezze degli studenti in aula si potrebbe ottenere una stima dell'altezza media di popolazioni più ampie – o con la sideremia, o quant'altro. Abbastanza bene, per esempio, dell'altezza delle ventenni triestine, o perfino italiane, meno bene dei ventenni italiani (maschi e femmine), ancora meno bene di tutti i ventenni del mondo, o di tutti gli esseri umani.
Ma di quale popolazione un campione è bene rappresentativo è problema drammatico e non ben risolubile di tutta la Farmacia e la Medicina. Le statistiche mediche spesso si fanno presso 1 solo

ospedale, o 2-3 ospedali, con campioni che non bene rappresentano tutta l'umanità, per vari motivi: socio-economici, etnici, eccetera. Poi, certo, gli Autori degli articoli scientifici se la giustamente cavano con frasi standard come "further studies are required" (20 milioni di risultati di ricerca Google nel settembre 2024), e magari "with larger samples".

36.10 Cenno agli stimatori di massima verosimiglianza

Lo stimatore $\hat{a}$ *di massima verosimiglianza* "cerca di essere" quello che renderebbe massimamente probabile l'uscita di $x_1,...,x_n$, se a valesse $\hat{a}$. Il metodo per trovarlo si basa sulle derivate e non è semplicissimo.
Ci limitiamo a dare la formula dello stimatore di massima verosimiglianza del parametro λ della legge esponenziale:

$$\hat{\lambda} = \frac{n}{X_1 + ... + X_n} = \frac{1}{\bar{X}_n} \,. \tag{146}$$

(Che non è non distorto).

BOZZA - DRAFT

ESERCIZIO$_{\mu}$ ≈
Stimare il parametro λ di una densità esponenziale da questo campione:

16.62, 3.810, 35.97, 4.322, 2.725, 11.44, 0.6671, 14.85, 3.816, 12.54.

Si dia il risultato con 2 cifre significative.

SVOLGIMENTO
Viene usato lo standard del punto decimale.
La media aritmetica degli $n = 10$ valori è $\bar{X}_n = 10.676$.
Lo stimatore di massima verosimiglianza è il reciproco della media

$$\hat{\lambda} = \frac{1}{\bar{X}_n} \approx$$

con 2 cifre significative

$$\boxed{\approx 0.094.}$$

BOZZA - DRAFT

Nota. Salvo approssimare i valori a 4 cifre significative, il campione era stato ottenuto su WolframAlpha con parametro 0.1, abbastanza vicino allo 0.094 trovato con lo stimatore, con l'istruzione

```
10 random numbers exponential distribution lambda=0.1
```

che naturalmente, se richiamata da qua, in generale darà altri 10 numeri, sempre provenienti da quella variabile aleatoria esponenziale simulata, con vero parametro $\lambda = 0.1$. Di volta in volta lo stimatore darebbe diversi valori per $\hat{\lambda}$, plausibilmente "vicini" al vero $\lambda = 0.1$.

Nota. Vale proprio la pena trarre più volte un campione di 10 elementi con questa distribuzione, esponenziale di parametro 0.1, su WolframAlpha:

```
10 random numbers exponential distribution lambda=0.1
```

e osservare la loro natura. Li si pensi come intertempi, in minuti, fra ingressi in Farmacia, o al Pronto Soccorso. Alle volte si aspetta molto, alle volte pochissimo. (E si pensi alle possibili conseguenze in Pronto Soccorso). I Lettori che hanno fatto barista hanno esperienza pratica di questi fatti.

36.11 Riassumiamo: gli stimatori puntuali

Prima di tutto, in questo paragrafo chiameremo stimatori puntuali quelli che finora abbiamo chiamato semplicemente stimatori; esistono poi gli stimatori intervallari, che sono coppie di variabili aleatorie da intendere come estremi di un intervallo, che vedremo.

Per una v.a. discreta o continua X, supponiamo di sapere che:
1) ha distribuzione di un certo tipo noto dipendente da un parametro incognito, sia esso a, per esempio $\mathbb{U}[-b,b]$ o $N(\mu,1)$ o $N(0,\sigma^2)$ o $\Gamma(\alpha,1)$ (dove a è rispettivamente b, μ, σ^2 e α);
2) avremo n determinazioni indipendenti x_1,...,x_n di X, che per ora, “prima" di averle, sono n variabili aleatorie X_1,...X_n.

Non sapremo mai quanto vale a ma vogliamo stimarlo con uno *stimatore puntuale* $\hat{a}$ mediante i dati x_1,...,x_n, ovvero (prima di averli) X_1,...,X_n.

Concretamente lo stimatore ci appare come una funzione di un numero arbitrario n di variabili, come $2\,\frac{X_1+...+X_n}{n}$ ovvero $2\,\frac{x_1+...+x_n}{n}$.

Ai fini delle Scienze Applicate, queste formule si trovano semplicemente elencate, sui vari testi, come questo, mentre produrle non è facilissimo. I matematici e in particolare gli statistici le producono, dimostrando le loro varie qualità positive, per esempio l'essere non distorti e altre.

A un livello superiore si considerano parametri 2-dimensionali, come $a := (\mu,\sigma^2)$.

La teoria degli stimatori dei momenti è ricca e interessante e qualcosa si trova nei Complementi a questa Lezione ma nella nostra trattazione elementare ci limitiamo molto.

BOZZA - DRAFT

36.12 Nota dolente finale e prospettive future

La probabilità che lo stimatore dia il valore esatto del parametro può, genericamente e imprecisamente parlando, ritenersi bassa per variabili aleatorie discrete, e nulla per variabili aleatorie continue.
Eppure, il valore che danno è molto meglio che niente, e si pensi semplicemente alla media. Questo è dovuto essenzialmente al fatto che anche se non danno il valore esatto del parametro da stimare, è "grande" la probabilità che l'errore sia "piccolo", seppure qua non approfondiamo la questione.
Resta il fatto che quel singolo numero non contiene in sè l'informazione sulla sua plausibile precisione. Questo limite verrà superato dai molto più moderni stimatori intervallari: invece di dare un singolo valore, daranno un intervallo all'interno del quale "sperabilmente" cade il valore da stimare; e ci sarà una sorta di misura per quel livello di sperabilità, il *livello di confidenza*.

Nota. Di questa Lezione ci si dovrebbe aspettare che lo studente conosca a memoria le formule (143), (144), (145), (146), oltre a tutto l'inquadramento teorico della questione.

Naturalmente bisogna conoscere anche le formule di media e varianza, riportate in questa Lezione, che sono state già illustrate in Statistica Descrittiva.

BOZZA - DRAFT

* Complementi *

36.13 Complementi – gli stimatori da un libro classico

Diamo qualche dettaglio sugli stimatori seguendo abbastanza da vicino un libro che si distingue nella didattica del Calcolo delle Probabilità e della Statistica: Calcolo delle probabilità, di Paolo Baldi, ed. McGraw-Hill.

BOZZA - DRAFT

Stimatore dei momenti col momento primo.
Lo stimatore $\hat{a}$ *dei momenti, col momento primo*, attribuisce al parametro a quel valore che farebbe avere alla densità $f_a(x)$ come media, ovvero speranza matematica ovvero momento primo, il valore che è proprio la media di $X_1,...X_n$.

Esempio 1. Per una densità $N(\mu, \sigma^2)$ con σ^2 noto (ora $a := \mu$)

$$\hat{\mu} := \frac{1}{n}(X_1 + ... + X_n).$$

Esempio 2. Per una densità esponenziale di parametro λ (ora $a := \lambda$) è

$$E(X) = \frac{1}{\lambda} =^{EQ} \bar{X}_n$$

$$\Rightarrow \hat{\lambda} := \frac{1}{\bar{X}_n} = \frac{n}{X_1 + ... + X_n}.$$

Esempio 3. Per una densità $\Gamma(\alpha, \lambda)$ con α noto (ora $a := \lambda$) è

$$E(X) = \frac{\alpha}{\lambda} =^{EQ} \bar{X}_n$$

$$\Rightarrow \hat{\lambda} := \frac{\alpha}{\bar{X}_n} = \frac{n\,\alpha}{X_1 + ... + X_n}.$$

Esempio 4. Per una densità uniforme $\mathbb{U}[-u, u]$ è

$$E(X) = \frac{u - u}{2} = 0$$

che non dipende dal parametro incognito $a = u$: il metodo dei momenti col solo momento primo non si può applicare.

Metodo di calcolo dello stimatore dei momenti 1) Si calcola la speranza matematica di $f_a(x)$ come funzione di a;
2) se dipende da a la si uguaglia a $\bar{X}_n$, media di X_1,...,X_n;
3) si risolve (se possibile) l'equazione in a;
Quella soluzione trovata, in termini di X_1,...,X_n è lo stimatore dei momenti, inteso come funzione (ed è una variabile aleatoria), per esempio

$$\hat{a} := \frac{n}{X_1 + \ldots + X_n}$$

e in termini di x_1,...,x_n è lo stimatore dei momenti inteso come valore dello stimatore, nel caso specifico di un campione x_1,...,x_n considerato, per esempio

$$\hat{a} := \frac{n}{x_1 + \ldots + x_n}$$

e quando a x_1,...,x_n si sostituiscano i loro valori, è un numero, per esempio

$$\hat{a} := \frac{3}{8+7+5} = \frac{3}{20} = 0.15.$$

4) Se $E(X)$ non dipende dal parametro incognito a, si usi allora il momento secondo uguagliandolo alla media quadratica.

Esempio 4 bis. Per una densità uniforme $\mathbb{U}[-u, u]$ è

$$E(X) = \frac{u-u}{2} = 0$$

che non dipende dal parametro incognito $a = u$. Con il momento secondo $\int_{-\infty}^{+\infty} x^2 f_a(x)\,dx$, ora che la densità è $\frac{1}{2u}$ fra $-u$ e u,

$$E(X^2) = \int_{-u}^{u} \frac{x^2}{2u}\,dx = \left[\frac{x^3}{6u}\right]_{-u}^{u} = \frac{u^3}{6u} - \frac{-u^3}{6u} = \frac{u^3}{3u} =$$

$$= \frac{u^2}{3} =^{EQ} \frac{1}{n}(X_1^2 + \ldots + X_n^2) \quad \leftarrow \textit{media quadratica}$$

BOZZA - DRAFT

$$\rightarrow \quad \hat{u} := \sqrt{\frac{3\,(X_1^2 + \ldots + X_n^2)}{n}}\,.$$

BOZZA - DRAFT