PROBABILITÀ – SCHEDA N. 8
LA LEGGE DEI GRANDI NUMERI E IL TEOREMA CENTRALE DEL LIMITE
Ovvero:
ci si può fidare dell'esperienza per stimare probabilità?

1. La legge empirica del caso

Abbiamo visto che non esiste un modo univoco per dare una valutazione delle probabilità di eventi.
Nell'assegnazione bisogna che siano però soddisfatte alcune proprietà (la probabilità di tutti i possibili eventi deve essere uguale a 1 e la probabilità che si verifichi l'unione di eventi incompatibili deve coincidere con la somma delle probabilità dei singoli eventi).
Comunque abbiamo osservato che per avere una valutazione ragionevole della probabilità di un evento si può osservare ripetutamente il suo verificarsi o meno, se è possibile l'osservazione ripetuta in condizioni identiche, oppure simularlo col calcolatore. Ci "fidiamo" del fatto che, ripetendo più volte un esperimento, nelle stesse condizioni, la frequenza relativa di un evento che riguarda quell'esperimento è una sensata valutazione della probabilità dell'evento stesso. La valutazione è tanto più affidabile quanto maggiore è il numero di prove.
A questa "legge", la cui validità è rilevabile sperimentalmente, si è attribuito il nome di "**legge empirica del caso**".
Ad esempio, nel lancio di una moneta a due facce non truccata, la frequenza relativa, all'aumentare del numero di lanci tende a 0.5. Questo non vuol dire che il numero t delle teste tende a essere uguale al numero c delle croci, ma è la frequenza relativa dell'evento "esce testa" che tende a essere uguale alla frequenza relativa dell'evento "esce croce".

2. La legge dei grandi numeri

Si legge in alcuni libri che la "legge empirica del caso" è anche nota col nome di "legge dei grandi numeri". Questa affermazione non è corretta, perché, in realtà, si tratta di due enunciati concettualmente molto diversi. Infatti, la "**legge dei grandi numeri**", detta anche "Teorema di Bernoulli", è, appunto, un teorema (e come tale, è una proposizione dimostrabile in una teoria). La legge empirica del caso è invece un'affermazione che sintetizza una regolarità osservabile sperimentalmente.
Per comprendere la legge dei grandi numeri occorrono nozioni avanzate della teoria delle probabilità. Tra l'altro sono note e usate formulazioni differenti della legge dei grandi numeri, che si collocano a diversi gradi di "generalità". Noi ne presentiamo una e lasciamo a eventuali studi futuri il compito di approfondire la trattazione.

Consideriamo una successione di variabili casuali X_1, X_2, ..., X_n indipendenti e con la stessa distribuzione e, in particolare, uguale media μ. Consideriamo quindi la variabile aleatoria $\overline{X}_n$ che si ottiene facendo la media delle variabili date : $\overline{X}_n = \dfrac{X_1 + X_2 + \ldots + X_n}{n}$.
Dalle proprietà della media e della varianza si ha che $\overline{X}_n$ ha valore atteso μ e varianza σ^2/n. [La dimostrazione è riportata in Appendice]

Prima di enunciare la legge dei grandi numeri, facciamo due esempi per capire il significato della variabile aleatoria $\overline{X}_n$.

ESEMPIO 1. Consideriamo l'esperimento del lancio di n dadi identici a sei facce. La funzione che associa, a ogni faccia dell'i-esimo dado ($1 \le i \le n$) un numero dell'insieme {1,2,3,4,5,6} è una variabile aleatoria che indica il punteggio ottenuto nel lancio e che chiamiamo X_i ($1 \le i \le n$). Consideriamo la variabile aleatoria S_n che corrisponde alla somma dei punteggi ottenuti negli n

lanci e che, quindi, assume valori nell'insieme numeri naturali compresi tra n e $6n$. La variabile $\overline{X}_n$ è $\frac{S_n}{n}$ e indica il punteggio medio ottenuti negli n lanci. Se effettuiamo effettivamente l'esperimento otterremo una particolare realizzazione $\overline{x}_n$ di $\overline{X}_n$.

ESEMPIO 2. Consideriamo, sempre nel lancio di un dado, l'evento "esce la faccia contrassegnata con 6". I sei possibili esiti a ogni lancio del dado sono descritti da una variabile aleatoria X che assume valore 1 se si verifica l'evento "esce la faccia contrassegnata con 6" e assume il valore 0 altrimenti. La variabile X ha quindi distribuzione di Bernoulli con probabilità p. Se si effettuano n lanci si hanno n variabili casuali indipendenti e identicamente distribuite (assumono sempre il valore 0 o 1 a seconda che non si verifichi o si verifichi l'evento "esce il 6") , X_1, X_2, ..., X_n . Tali variabili aleatorie hanno con la stessa media (visto che hanno la stessa distribuzione).
La variabile aleatoria $\overline{X}_n = \frac{X_1 + X_2 + \ldots + X_n}{n}$ indica la frequenza relativa dell'evento "esce la faccia contrassegnata con 6 in n lanci". Anche in questo caso se effettuiamo effettivamente l'esperimento otterremo una particolare realizzazione $\overline{x}_n$ di $\overline{X}_n$.

La **legge dei grandi numeri** dice che:
Se la numerosità n delle prove tende all'infinito la probabilità che $\overline{X}_n$ *assuma valori al di fuori dell'intervallo* $(\mu - \delta, \mu + \delta)$ *tende a 0, qualunque sia la semiampiezza* δ *dell'intervallo.* In formula:

$$\lim_{n\to+\infty} P\left(\left|\overline{X}_n - \mu\right| > \delta\right) = 0 .$$

Il che equivale a dire che se n tende a infinito la probabilità che la media campionaria $\overline{X}_n$ converga alla media comune μ delle X_i è uguale a 1, ossia che per ogni δ >0,

$$\lim_{n\to+\infty} P(|\,\overline{X}_n - \mu\,| < \delta) = 1$$

Interpretiamo la legge dei grandi numeri nel caso dell'**Esempio 1**.
All'aumentare del numero n dei dadi, cresce la probabilità che la media dei punteggi ottenuti lanciando gli n dadi, sia una buona valutazione del valore atteso μ di ciascuna variabile X_i (osserviamo che, se i dadi non sono truccati, si ha μ = 3.5). La legge dei grandi numeri specifica che se il numero di dadi considerati n tende a infinito, allora è uguale a 1 la probabilità che lo scarto assoluto tra la variabile aleatoria media campionaria e il valore atteso μ sia minore di un qualunque numero reale δ, piccolo a piacere. Il che equivale a dire che, per n tendente a infinito, è uguale a 1 la probabilità che la media campionaria sia uguale al valore atteso μ, cioè questo è un evento praticamente certo.

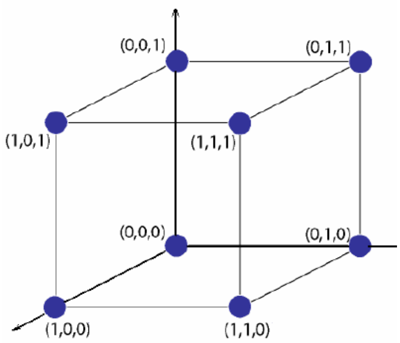
Interpretiamo la legge dei grandi numeri nel caso dell'**Esempio 2**.
In questo caso abbiamo un evento ("esce il sei" nel lancio di un dado) che immaginiamo abbia probabilità sconosciuta p (sconosciuta perché il dado potrebbe essere truccato, o semplicemente difettoso: non possiamo saperlo in anticipo).
Eseguendo n lanci consecutivi otteniamo una valutazione della probabilità di ottenere sei con quel dado, data da

$$\overline{x}_n = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

dove i valori x_1, x_2, ..., x_n indicano l'esito dei lanci e valgono 1 se in quel lancio è uscito il 6, mentre valgono 0 se è uscito un altro numero.
$\overline{x}_n$ è una realizzazione della variabile aleatoria $\overline{X}_n = \frac{X_1 + X_2 + \ldots + X_n}{n}$.

La legge dei grandi numeri afferma che, tante più prove si fanno, tanto più la *valutazione* della probabilità fatta sulla base dell'esperienza sarà vicina, *probabilmente*, alla probabilità reale p dell'evento.
Se la valutazione è molto vicina a 1/6, probabilità teorica che esca il sei per un dado perfettamente equilibrato, potremo essere ragionevolmente certi che il dado in questione non è polarizzato per il sei (per essere ragionevolmente sicuri che il dado non sia truccato in nessun modo dovremmo ripetere il test anche per gli altri cinque numeri). Che cosa significhi *ragionevolmente sicuri* dipende da quanto vogliamo essere precisi nel nostro test: con dieci prove avremmo una valutazione grossolana, con cento ne otterremmo una molto più precisa, con mille ancora di più e così via: il valore di n che siamo disposti ad accettare come sufficiente sarà oggetto di future trattazioni.
Attenzione, però: lanciando un dado "vero" molte volte, questo può deformarsi e non mantenere costante la probabilità di uscita di una faccia. Questo è successo realmente quando un matematico ha voluto effettivamente verificare quello che succedeva con un numero molto alto di lanci!!!

Come è possibile notare, la legge dei grandi numeri ha un enunciato più complesso e meno vago di quello della legge empirica del caso; soprattutto, è un teorema, ossia può essere dimostrato in una teoria (quella del calcolo delle probabilità).
La sua portata conoscitiva è quindi molto superiore di quella della legge empirica del caso e ci si può aspettare che conduca a sviluppi più significativi rispetto alle applicazioni consentite dalla legge empirica del caso.
Il fatto di accontentarsi della legge empirica del caso o, invece, di apprezzare la prospettiva che dà la legge dei grandi numeri è un interessante indice della nostra minore o maggiore propensione a una visione matematica della realtà. "Matematizzare" vuol dire descrivere fenomeni e situazioni utilizzando un linguaggio particolare, fatto di simboli e formule, che hanno sicuramente qualche referente concreto, ma che spesso consentono di vedere il concreto con nuovi occhi, da diverse prospettive e che per questo, spesso, offrono interpretazioni inattese e aprono nuove strade alla conoscenza. Si pensi, per esempio, al linguaggio delle variabili aleatorie.
Ad esempio, quando eseguiamo 3 lanci di una moneta non truccata, potremmo accontentarci di considerare solamente gli esiti, ossia una lista di tre numeri.
Invece, i matematici, per ogni lancio, considerano una variabile aleatoria, ossia a una funzione che associa, per esempio, all'uscita di testa il numero 1 e all'uscita di croce il numero 0. La lista dei 3 numeri che indicano gli esiti dei tre lanci è allora una realizzazione fra le 2^3 possibili realizzazioni, ossia è un elemento del prodotto cartesiano {0,1}x{0,1}x{0,1} (qui a fianco visualizzato).



Tutto ciò significa solo rendere inutilmente complicate cose semplici? Vuol dire ingegnarsi a oscurare i significati? Oppure vuol dire offrire nuovi occhi per osservare i fenomeni, per potenziare la nostra capacità di osservare gli oggetti di studio? La risposta può costituire un indice significativo per valutare la nostra eventuale vicinanza o lontananza dalla matematica e dal modo con cui, attraverso essa, possiamo vedere i fenomeni. Nel paragrafo seguente e nelle schede di inferenza statistica verrà evidenziato proprio come la matematica teorica abbia permesso di "valutare" osservazioni statistiche sperimentali.

3. Il teorema del limite centrale (o teorema centrale del limite)

Con questo nome vengono presentati diversi enunciati che riguardano la convergenza alla distribuzione normale di medie o somme di variabili aleatorie indipendenti.

Sia la legge dei grandi numeri che il teorema del limite centrale riguardano la variabile aleatoria $\overline{X}_n$, media di n variabili aleatorie indipendenti e con la stessa distribuzione:

$$\overline{X}_n = \frac{X_1 + X_2 + \ldots + X_n}{n}.$$

La legge dei grandi numeri precisa che la variabile aleatoria $\overline{X}_n$ permette di valutare il valore atteso μ di X. Il teorema del limite centrale dà informazioni sulla distribuzione di $\overline{X}_n$.

Questo teorema fu enunciato per la prima volta da Laplace nel 1812, ma solo intorno agli anni trenta del Novecento fu definitivamente sistemato da Bernstein e poi da Feller.
Il teorema, nelle diverse formulazioni e generalizzazioni ha un ruolo importante, sia dal punto di vista teorico, sia, ancora di più, nelle applicazioni. Molti dei principali metodi utilizzati nell'inferenza statistica si basano su questo teorema (che è detto *centrale* proprio per tale motivo).

Il teorema afferma che se si hanno n variabili aleatorie indipendenti X_1, X_2, ..., X_n, con stessa media μ e stessa varianza σ^2, allora la successione di variabili aleatorie

$$Z_n = \frac{\overline{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}}$$ (cioè le variabile aleatorie $\overline{X}_n$ standardizzate)

"tende" ad avere una distribuzione **normale** di media 0 e varianza 1; dove "tende" significa che, all'aumentare di n, la distribuzione di Z_n assomiglia sempre più a quella di una variabile aleatoria con distribuzione N(0,1).

Una formulazione più precisa del teorema è la seguente.

> *La funzione di distribuzione cumulata di Z_n, calcolata in ogni punto t, cioè $F_{Z_n}(t)$, quando n tende all'infinito, tende alla funzione di distribuzione cumulata di Z calcolata in t, cioè $F_Z(t)$, con Z variabile aleatoria normale standardizzata.*
> In formula:
> $$\forall t,\ \lim_{n\to\infty}\left(F_{Z_n}(t)\right) = F_Z(t)$$

Il teorema centrale, come già detto, è un risultato teorico e contiene, nel suo enunciato un limite per n che tende a infinito; ha però un risvolto operativo che si può sintetizzare in questo modo.

> *Qualunque sia la distribuzione di una variabile aleatoria X, la funzione di distribuzione cumulata di $\overline{X}_n$ si può approssimare con quella di una variabile aleatoria di distribuzione normale (con stessa media e varianza di $\overline{X}_n$) e tale approssimazione è tanto migliore quanto più grande è n.*

Vediamo due esempi allo scopo di chiarire quanto è stato detto.

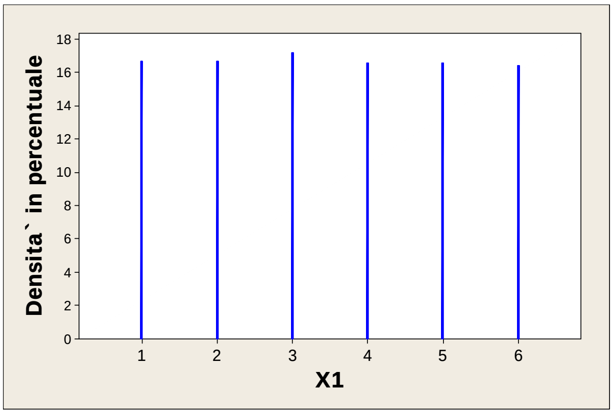
ESEMPIO 3. Consideriamo n dadi non truccati. I possibili esiti di ciascuno di essi sono modellizzati da una variabile aleatoria X che assume valori sull'insieme {1,2,3,4,5,6} e che ha una distribuzione uniforme (P(1)=P(2)=P(3)=P(4)=P(5)=P(6) = 1/6), quindi ben lontana da quella normale.
Ritorniamo al significato del teorema del limite centrale: esso assicura che la forma della distribuzione di $\overline{X}_1$, $\overline{X}_2$, $\overline{X}_4$, ..., $\overline{X}_n$ tende ad assumere sempre più la forma di una gaussiana man mano che si procede nei termini della successione (ossia man mano che aumenta n) basta che le variabili X_i siano indipendenti, con stessa varianza e stessa media. Questo anche nel caso in cui ciascuna di essa (come nel caso dei lanci di un dado) ha una funzione di densità di probabilità molto diversa dalla curva "a campana" tipica della funzione di densità di probabilità normale.
Sappiamo che esiste una dimostrazione di tale teorema, ma non è per nulla semplice ed è al di fuori della portata di studenti di una scuola secondaria. Possiamo però dare una visualizzazione del teorema in un contesto in cui siamo in grado di fare direttamente i "conti". Consideriamo il lancio di 1, 2, 3, 4, ... , n dadi. Potremmo calcolare (con molta pazienza) la distribuzione teorica delle variabili aleatorie $\overline{X}_1$, $\overline{X}_2$, $\overline{X}_4$, ..., $\overline{X}_n$ (come visto nella scheda 5) con

$$\overline{X}_1 = X_1, \quad \overline{X}_2 = \frac{X_1 + X_2}{2}, \quad \overline{X}_3 = \frac{X_1 + X_2 + X_3}{3} \ ...., \quad \overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i\ .$$

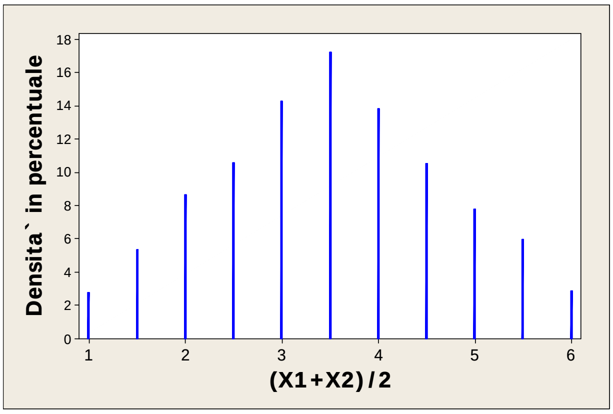
Invece di fare ciò, simuliamo la situazione.
Sappiamo che la distribuzione di $\overline{X}_1$, cioè di X_1, è uniforme su {1,2,3,4,5,6}, ma simuliamo anche questa.
Generiamo 10 000 numeri casuali da una variabile uniforme Y in [0,1] e calcoliamo

$$X_1 = \mathrm{int}(Y \times 6 + 1)$$

dove la funzione "int" indica la parte intera di un numero.
Alcuni software hanno direttamente la distribuzione uniforme intera.
A fianco è riportato il grafico della densità di X_1 simulata.

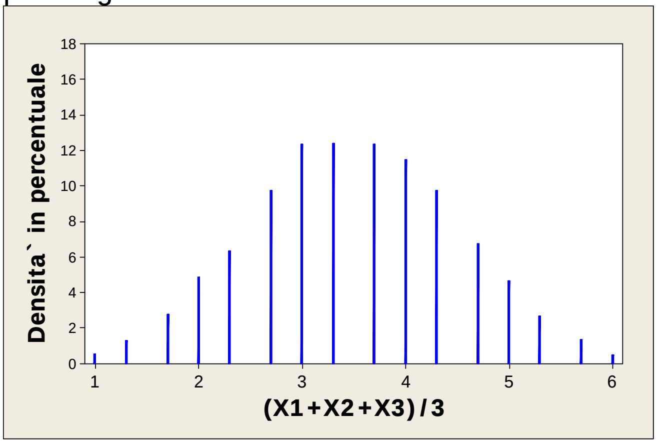


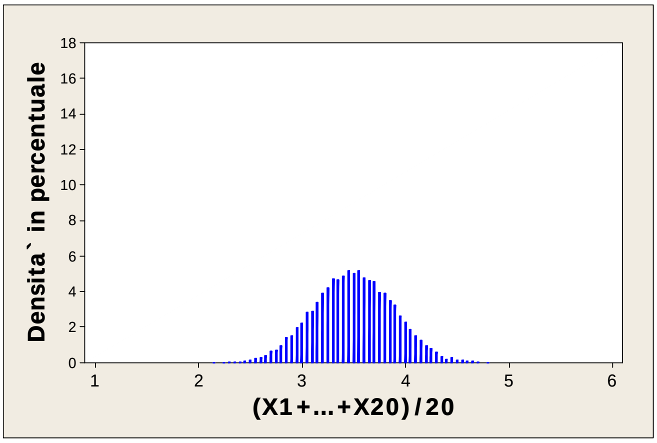
La densità di probabilità della media di due uniformi indipendenti ha forma "triangolare". Il grafico a fianco riporta la densità della media di due variabili aleatorie ottenute per simulazione.



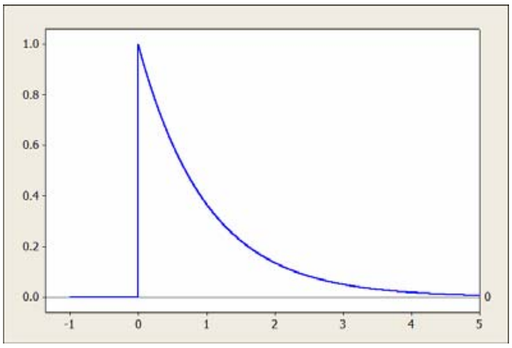
All'aumentare di n, la **media** di n variabili aleatorie uniformi assume sempre più una forma "a campana", tipica della distribuzione normale.

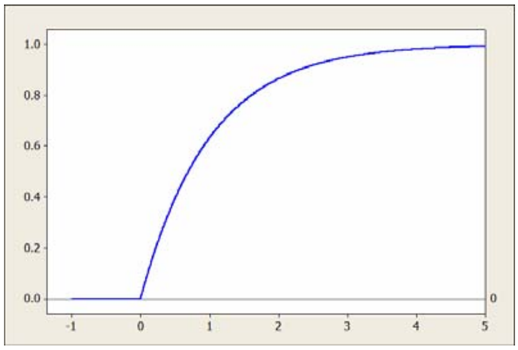
Riportiamo di seguito i grafici delle densità della somma dei punteggi ottenuti nel lancio di n dadi, per n uguale a 3 e a 20.





ESEMPIO 4.
Consideriamo una variabile aleatoria X esponenziale di parametro 1, con funzione di densità di probabilità e funzione di distribuzione cumulata rappresentate nei grafici sotto, molto "lontane" dalla forma delle corrispondenti funzioni di una variabile aleatoria con distribuzione normale (in questo caso non vi sono nemmeno simmetrie, come invece si avevano nel caso dell'esempio precedente). Abbiamo già detto che un esempio concreto modellabile con una variabile aleatoria esponenziale è la durata di una lampadina. La variabile aleatoria $\overline{X}_n$ modella quindi la durata di vita media di n lampadine.



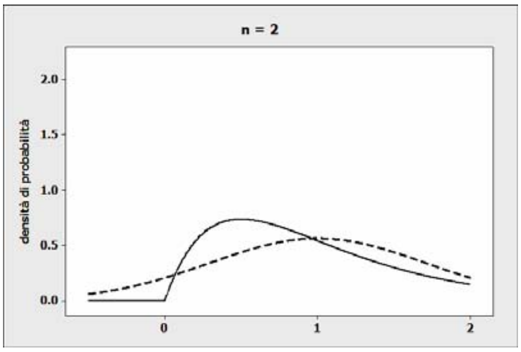


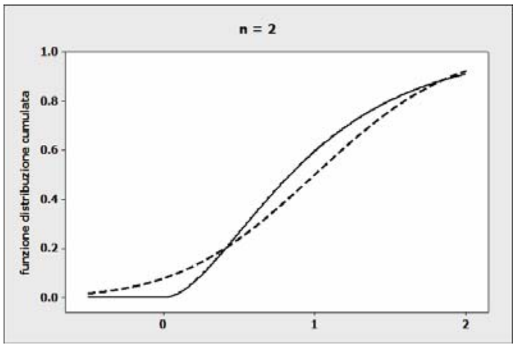
La variabile aleatoria $\overline{X}_n$ modella quindi la durata di vita media di n lampadine.

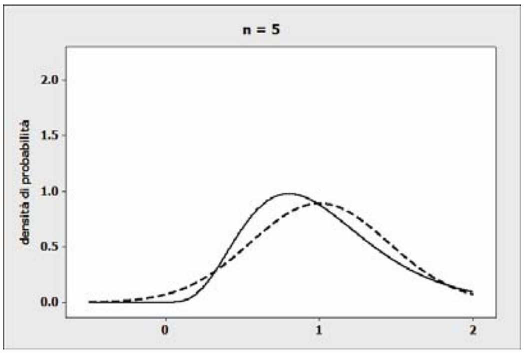
In questo secondo esempio, invece di simulare $\overline{X}_n$, utilizziamo (senza riportarle) le formule delle funzioni di densità di probabilità e di distribuzione cumulata *esatte* per $\overline{X}_n$. Si possono quindi confrontare tali funzioni esatte con quelle di una normale con stessa media e stessa varianza.

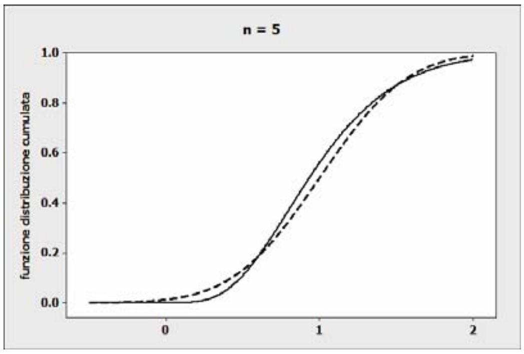
I grafici a linea continua si riferiscono alla distribuzione della variabile aleatoria esponenziale e quelli a linea tratteggiata alla distribuzione normale corrispondente.

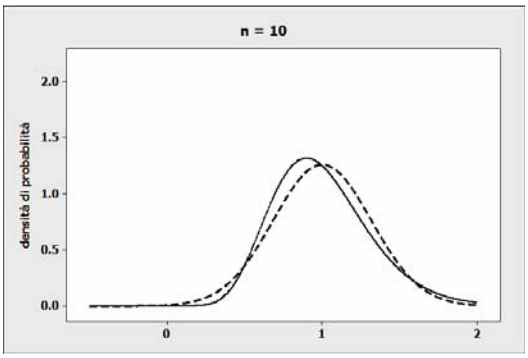
La “tendenza alla normale” di $\overline{X}_n$, si ha già per valori di n relativamente piccoli.

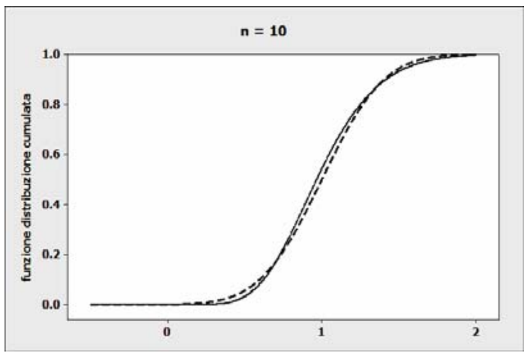


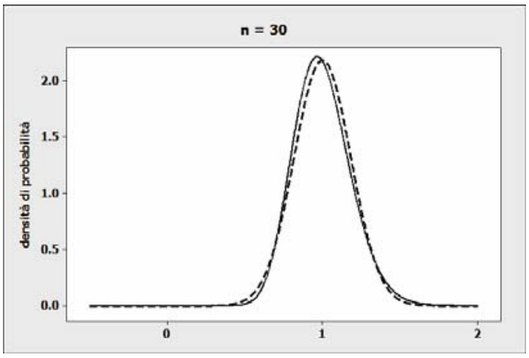


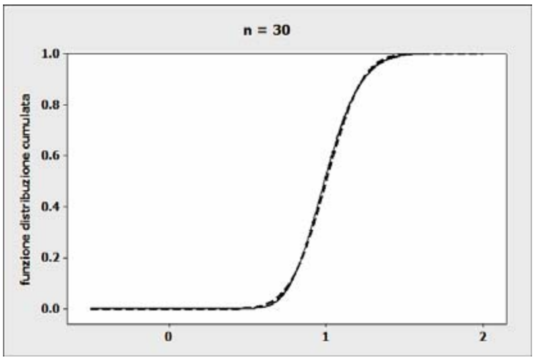












Il teorema centrale giustifica in parte la popolarità del modello "normale" nella descrizione di molti fenomeni. Quando la variabilità di un fenomeno può essere pensata come la somma di varie cause indipendenti con stessa media e stessa varianza, allora si può applicare il teorema del limite centrale e affermare che la media campionaria ha una distribuzione normale (anche nel caso in cui le variabili casuali abbiano diversa distribuzione).

Quello che si osserva è che, anche partendo da distribuzioni molto lontane dalla normale (come abbiamo visto negli esempi precedenti) e anche in condizioni meno restrittive di quelle da noi enunciate nel teorema del limite centrale, la convergenza di Z_n alla normale standard è abbastanza rapida. In genere con $n > 30$ si può essere abbastanza tranquilli nell'approssimare la distribuzione di una somma o una media di variabili aleatorie con quella di una variabile aleatoria normale.

ESEMPIO 5. Si deve verificare il buon funzionamento di una macchina che produce tubi di acciaio. In condizioni di regolarità della produzione il diametro è 30 mm e la deviazione standard è 4 mm. Per verificare se la macchina non si sia "sregolata" si esamina un campione di 100 pezzi e se la media dei diametri del lotto è maggiore di 31 mm si attua una manutenzione straordinaria della macchina. È ragionevole questa scelta? Per rispondere dobbiamo sapere qual è la probabilità di ottenere una media dei diametri maggiore di 31 mm, nel caso in cui la macchina funzioni in modo regolare.
Modelliamo la misura del diametro del tubo con una variabile aleatoria X di valore atteso 30 e di deviazione standard 4 e consideriamo la variabile aleatoria media campionaria $\overline{X}_{100}$.
Dobbiamo valutare se $P\left(\overline{X}_{100} > 31\right)$ è "alta" o "bassa".
Anche se non conosciamo la distribuzione di X, essendo la numerosità campionaria alta, possiamo applicare il teorema del limite centrale e quindi approssimare la distribuzione di $\overline{X}_{100}$ con quella di una normale di media 30 e deviazione standard 0.4. Infatti:

$$E\left(\overline{X}_{100}\right) = E(X) = 30 \quad \text{std}\left(\overline{X}_{100}\right) = \frac{\text{std}(X)}{\sqrt{n}} = \frac{4}{10}.$$

Possiamo quindi dire che:

$$P\left(\overline{X}_{100} > 31\right) = P\left(\frac{\overline{X}_{100} - 30}{0.4} > \frac{31-30}{0.4}\right) \cong P\left(Z > 2.5\right) = 0.0062 \quad \text{con } Z \sim N(0,1).$$

La probabilità di ottenere un valore medio maggiore di 31 mm in condizioni di funzionamento regolare è molto bassa. Quindi si può supporre che la macchina abbia bisogno di una manutenzione straordinaria, in quanto vi sono buone ragioni per ritenere che non stia più funzionando regolarmente.

ESEMPIO 6. Riprendiamo l'Esempio 2 della scheda 6, che riportiamo qui integralmente:
"Il rapporto dei sessi nella specie umana alla nascita è di 105 femmine su 100 maschi. Qual è la probabilità che in 6 nascite singole almeno la metà dei neonati siano di sesso femminile? Come varierebbe la probabilità se si considerassero 60 nascite singole? Sarebbe maggiore o minore della precedente?"
Concentriamoci sulla risposta all'ultima domanda.

Indichiamo con S_{60} il numero di neonati di sesso femminile nelle 60 nascite prese in esame. Vogliamo calcolare la probabilità che su 60 nascite almeno la metà siano di sesso femminile, quindi P($S_{60} \geq 30$). Se utilizzassimo il fatto che S_{60} è una binomiale B(60, 0.51), dovremmo svolgere

$$P(S_{60} \geq 30) = P(S_{60}=30) + P(S_{60}=31) + \ldots P(S_{60}=60),$$

Questo calcolo richiede non solo un'addizione di 30 addendi, ma, quel che è peggio, il calcolo del valore di ciascuno degli addendi è piuttosto complesso.

Abbiamo però a disposizione il Teorema del limite centrale, il quale afferma che S_{60} (che si può scrivere come somma di 60 variabili aleatorie indipendenti di media 0.51 e varianza 60*0.51*0.49 = 15) si può approssimare con una normale.

Abbiamo quindi

$$P(S_{60} \geq 30) = P\left(\frac{S_{60} - 60\ 0.51}{\sqrt{15}} > \frac{30 - 60\ 0.51}{\sqrt{15}}\right) \cong P(Z > -0.15) = 1\text{-}0.44 = 0.56$$

Facendo i calcoli direttamente avremmo ottenuto 0.51.

4. Alcune riflessioni sul "fidarsi" dell'esperienza

Negli esempi 3 e 4 abbiamo utilizzato la simulazione o le conoscenze teoriche sulle funzioni di distribuzione e di ripartizione delle variabili casuali considerate per *verificare* l'enunciato del teorema centrale, e anche la sua efficacia, osservando la velocità di convergenza alla normale. In questi esempi, la situazione reale da modellare, cioè i risultati del lancio di un dado o la durata di una lampadina, ci è servita per comprendere meglio, in un certo senso visualizzandolo, l'enunciato di un teorema. In un certo senso lo abbiam "visualizzato". Naturalmente non abbiamo dimostrato niente: la dimostrazione del teorema del limite centrale, come abbiamo già detto è al di fuori della portata di uno studente di scuola secondaria. Sappiamo però che esiste e, quindi, possiamo contare sulla sua correttezza e adeguatezza per applicarne le coseguenza nella risoluzione di problemi applicativi.

Ma che cosa accade quando il risultato teorico è inaccessibile per problemi di complessità di calcolo oppure perché non è proprio noto? In questo caso è possibile fidarsi della simulazione per risolvere un problema, per esempio per risolvere il problema di stimare una probabilità?

La risposta sta, in fondo, nella fiducia che attribuiamo alla legge empirica del caso, ma questa non è un risultato della teoria. Nella forma in cui viene espressa è irriducibile alla teoria. Tutto sembra ritornare daccapo; in realtà – come abbiamo già osservato – la legge dei grandi numeri e il teorema centrale del limite (due risultati teorici) pongono condizioni ben precise, consentendo di essere più consapevoli dei limiti e delle potenzialità della legge empirica del caso.

Appendice. Dimostrazione del valore atteso e della varianza di $\overline{X_n}$

$$E\left(\overline{X}_n\right) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E\left(X_i\right) = \frac{1}{n}\sum_{i=1}^{n}\mu = \frac{1}{n}n\mu = \mu$$

$$\text{Var}\left(\overline{X_n}\right) = \text{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\text{Var}\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n}\text{Var}\left(X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}$$

$$\text{std}\left(\overline{X_n}\right) = \frac{\sigma}{\sqrt{n}}$$

Possiamo standardizzare la variabile aleatoria $\overline{X}_n$ e ottenere la variabile $Z_n = \dfrac{\overline{X}_n - \mu}{\sigma / \sqrt{n}}$ con valore atteso 0 e varianza 1.

Riferimenti bibliografici:

Rossi, C. (1999), *La matematica dell'incertezza*, Zanichelli, Bologna

Sitografia
Wikipedia, http://it.wikipedia.org/wiki/Legge_dei_grandi_numeri e
http://it.wikipedia.org/wiki/Teorema_del_limite_centrale