7 Il TLC

7.1 Esplorazioni: i parametri -attività-

Cosa. Ora disponiamo di vari strumenti relativi alla variabile aleatoria normale, dunque possiamo comprendere più a fondo il risultato di approssimazione introdotto all'inizio del percorso. In particolare, investighiamo sui valori da attribuire ai parametri della distribuzione normale affinché essa approssimi una certa distribuzione binomiale.

Come. Ha senso effettuare tale lavoro con un approccio operativo-sperimentale: si tratta di un'attività di scoperta che lo studente può svolgere tramite il file Geogebra proposto. Va osservato che le richieste sono di tipo volutamente qualitativo e dunque non ammettono una risposta univoca. Anche per questo, come del resto in varie altre parti del percorso, quanto scrive lo studente non deve necessariamente coincidere con quanto si trova sui materiali e ciò va esplicitato alla classe. A tal proposito è significativa l'osservazione di R. Zan [23, pag. 27-28]:

> [...] implicita nella preoccupazione di evitare domande "troppo difficili" c'è spesso la valorizzazione della correttezza dei prodotti, che viene considerata più importante dell'attivazione di processi di pensiero significativi [...]. È il compromesso delle risposte corrette, denunciato dallo psicologo americano Howard Gardner: un comodo patto fra insegnanti e studenti, che si accordano sul fingere che la risposta corretta garantisca comprensione.

Uso. L'attività si presta ad esser svolta individualmente, anche a casa, e si avvale del file Geogebra NormaleBinomiale1.ggb. Al docente spetterà poi raccogliere le considerazioni fatte e discutere le conclusioni con gli studenti.

7.2 Esplorazioni: il numero di prove -attività-

Cosa. Fissato il legame tra i parametri della distribuzione normale e quelli della binomiale, è interessante sondare la convergenza, ossia studiare cosa succede al crescere del numero di prove dello schema di Bernoulli.

Come. Le scelte didattiche sottese e le competenze sviluppate sono analoghe a quelle dell'attività precedente.

Uso. Anche se abbiamo preferito presentarlo nella forma di lezione partecipata, il materiale si presta ad essere proposto come lavoro individuale. In ogni caso, è opportuno che lo studente si immerga nelle prove in prima persona, mediante il file Geogebra NormaleBinomiale2.ggb.

7.3 La sostanza del teorema

Aspetti didattici

Cosa. Grazie alle attività esplorative condotte come indicato nei paragrafi precedenti, gli studenti possono apprezzare l'enunciato semplificato del teorema limite centrale. È la nostra seconda formulazione, dopo quella ancor più ridotta presentata all'inizio del percorso.

Certo, il risultato non compare esplicitamente nelle Indicazioni nazionali, ma è assai utile per semplificare il calcolo, come nel nostro problema guida relativo ai sondaggi, ma anche concettualmente, dato che che stabilisce un legame tra due distribuzioni che sono oggetto di studio per lo studente della classe quinta.

Come. Il docente (e anche lo studente) *deve* esser consapevole del fatto che l'enunciato proposto non è preciso, ma d'altra parte è proprio in tale versione che il teorema trova concretamente applicazione in molti problemi di calcolo.
Del resto non ha nemmeno tanto senso appellarsi ad una presunta mancanza di rigore per non presentarlo in classe, dato che il concetto di rigore è cambiato nel tempo, come afferma in ambito didattico Colette Laborde, citata da D'Amore e Sbaragli [13].

> La nozione di "correttezza" non è assoluta e si riferisce sempre ad un dato sapere; il sapere di riferimento può anche evolversi. I criteri di rigore in Matematica sono cambiati considerevolmente nel tempo. Ogni concezione ha un suo dominio di validità e funziona per quel preciso dominio.

Uso. Questo materiale è stato pensato per una lezione in modalità partecipata.

Approfondimenti teorici

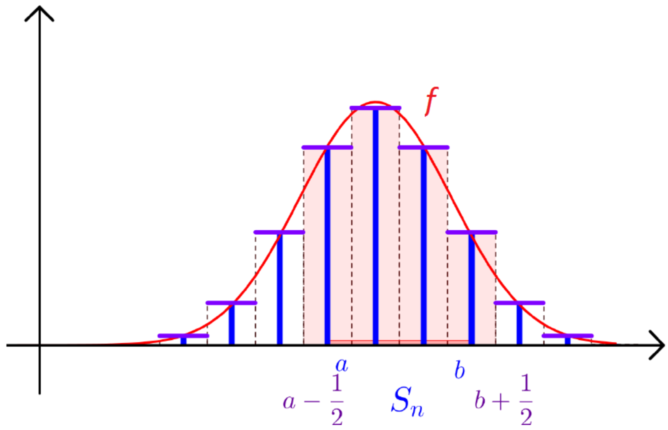
- **Confronto tra distribuzioni**
  Nel trattare l'approssimazione in questione, abbiamo confrontato i grafici di due oggetti matematici diversi: la distribuzione di probabilità $p(s)$ della variabile aleatoria binomiale S_n e la densità di probabilità $f(x)$ della v.a. normale X.
  Mentre nel caso della v.a. binomiale la distribuzione è una funzione i cui valori *sono* valori di probabilità, nel caso della v.a. normale questi sono invece forniti dagli integrali della densità.
  Allora cosa possiamo affermare sul legame tra le probabilità delle due distribuzioni?
  Per rispondere, consideriamo la distribuzione binomiale e valutiamo la probabilità che la v.a. binomiale S_n sia compresa tra due numeri interi a, b, dove $a \leq b$, ovvero

  $$P(a \leq S_n \leq b) = \sum_{j=a}^{b} p(j) \quad \text{con } j \text{ numero intero}$$

  I valori $p(j)$ sono rappresentati dalla lunghezza delle linee in figura e ciascuna di queste può essere sostituita con un rettangolo che ha come base il segmento di estremi $j - 1/2,\ j + 1/2$ e come altezza $p(j)$. Così facendo il valore numerico dell'altezza è uguale all'area del rettangolo corrispondente. Pertanto i valori della distribuzione binomiale si possono interpretare come somma di lunghezze o equivalentemente come area del sottografico della funzione costante a tratti ivi definita, ossia la densità della distribuzione.

  L'interpretazione in termini di aree permette di confrontare due densità, cioè oggetti matematici *dello stesso tipo*. È sulla base di questa interpretazione che ha senso l'approssimazione espressa nel TLC.

- **Dall'approssimazione all'uguaglianza**
  Riprendiamo l'enunciato semplificato del teorema limite centrale.



Ossia, sia S_n la variabile aleatoria binomiale di parametri n e p, dove n rappresenta il numero di prove e p la probabilità di successo nella prova. Sia poi X la v.a. normale di densità f, che ha media μ e varianza σ^2 uguali rispettivamente alla media e alla varianza di S_n, quindi $\mu = np$ e $\sigma^2 = np(1-p)$. Per n "grande" vale

$$P(a \leq S_n \leq b) \simeq \int_a^b f(x)\,\mathrm{d}x$$

Passando al limite per $n \to +\infty$ con a e b fissati, la probabilità relativa alla variabile aleatoria binomiale va a zero, quindi bisogna trovare un'altra strada. L'idea è di operare il confronto con la variabile aleatoria normale standard, ma allora occorre trasformare la v.a. S_n in una v.a. che abbia anch'essa media 0 e varianza 1. Ciò si può ottenere negli usuali due passi:

1. si sottrae il valore atteso np; la nuova v.a. $S_n - np$ ha media 0
2. si divide $S_n - np$ per la deviazione standard $\sqrt{np(1-p)}$; così la nuova v.a. ha varianza 1

Quindi la standardizzata di S_n è

$$S_n^* = \frac{S_n - np}{\sqrt{np(1-p)}}$$

Allora il TLC si può formulare come segue[10]:

TEOREMA LIMITE CENTRALE -TLC

Sia S_n la v.a. binomiale di parametri n e p, dove n rappresenta il numero di prove e p la probabilità di successo nella prova. Sia S_n^* la v.a. standardizzata di S_n.
Vale

$$\lim_{n\to+\infty} P(a \leq S_n^* \leq b) = \int_a^b g(x)\,\mathrm{d}x$$

dove g è la densità normale standard e a e b sono due numeri reali tali che $a \leq b$.

[10]Per la dimostrazione del teorema e approfondimenti, si vedano [33, da pag.325 a pag.329] e [31,

Il TLC fornisce anche una *giustificazione* dell'uso dell'approssimazione, come mostriamo di seguito.
Per n "grande" fissati α e β, dal TLC segue che

$$P(\alpha \leq S_n^* \leq \beta) \simeq \int_\alpha^\beta g(x)\,\mathrm{d}x$$

Ricordando la definizione di S_n viene

$$P(\sqrt{np(1-p)}\alpha + np \leq S_n \leq \sqrt{np(1-p)}\beta + np) \simeq \int_\alpha^\beta g(x)\,\mathrm{d}x$$

Se $a = \sqrt{np(1-p)}\alpha + np$ e $b = \sqrt{np(1-p)}\beta + np$, tale uguaglianza diventa

$$P(a \leq S_n \leq b) \simeq \int_{\frac{a-np}{\sqrt{np(1-p)}}}^{\frac{b-np}{\sqrt{np(1-p)}}} g(x)\,\mathrm{d}x$$

Infine, detti $a^* = \frac{a-np}{\sqrt{np(1-p)}}$ e $b^* = \frac{b-np}{\sqrt{np(1-p)}}$, arriviamo all'approssimazione

$$P(a \leq S_n \leq b) \simeq \int_{a^*}^{b^*} g(x)\,\mathrm{d}x$$

che abbiamo preso come versione semplificata dell'enunciato formale del teorema limite centrale.

7.4 Applicazioni: finalmente i sondaggi

Cosa. A questo punto del percorso, disponiamo di tutti gli elementi per risolvere il problema relativo ai sondaggi in modo meno dispendioso computazionalmente. Possiamo così concludere la risoluzione del problema guida, che ci ha impegnato fin dall'inizio e che ha motivato l'introduzione degli strumenti matematici di cui ci siamo occupati nel percorso.

Come. Le scelte didattiche sottese alla risoluzione seguono ormai da quelle operate in precedenza, proprio allo scopo di introdurre il TLC. In effetti, il procedimento che conduce alla risposta è una sintesi dell'intero percorso.
Il problema è già stato modellizzato all'inizio del percorso, pertanto si tratta di calcolare effettivamente la probabilità relativa al numero di individui che voteranno uno dei due candidati.
In questa fase, ci sembra significativo curare la sintassi, non in modo fine a se stesso, ma per fissare esplicitamente le relazioni esistenti tra i vari oggetti matematici introdotti via via nel percorso. In quest'ottica distinguiamo tra il simbolo di approssimazione e quello di uguaglianza, riservando il primo all'applicazione del TLC, all'arrotondamento di un numero o alla lettura dei valori dalle tavole di probabilità; invece ricorriamo al simbolo di uguaglianza solo quando i due termini sono proprio uguali.
Un altro aspetto è il calcolo della probabilità relativa alla v.a. standard. A tale proposito ci sembra significativo sfruttare la simmetria dell'intervallo di integrazione, anche se, naturalmente, lo studente non è vincolato a seguire questo approccio; anzi, è opportuno che egli si confronti con più modalità risolutive.

paragrafo 3.2].

Uso. Il materiale può essere utilizzato a supporto di una lezione partecipata, che andrebbe condotta riservando importanza all'argomento, visto che è il nostro problema guida. I ragazzi comunque dovrebbero ormai disporre di tutti gli strumenti per poterlo svolgere individualmente.

7.5 Applicazioni: dall'esame di Stato

Cosa. Questo quesito è stato proposto ben tre volte dal MIUR: in due prove del 2015 (sessione straordinaria italiana di settembre 2015 e sessione ordinaria 2015, scuole italiane all'estero, Americhe) e nella simulazione del 29 aprile 2016, sempre per il Liceo scientifico. Sfruttando il modello binomiale si può risolvere il quesito, ma la seconda richiesta comporta calcoli dispendiosi (ed effettivamente non eseguibili senza un calcolatore), anche se non esplicitamente richiesti nel testo. È qui allora che l'uso del TLC acquista valore, permettendo di risolvere l'esercizio in modo più efficace.

Uso. Gli studenti dovrebbero essere in grado di risolvere tale quesito in modo autonomo. Ad esso vanno poi aggiunti altri esercizi di consolidamento, desunti magari dal libro di testo e/o da altre prove dell'esame di Stato.