3.2 Dalla binomiale alla Poisson - attività

Leggi la questione e poi segui il procedimento che ti viene proposto per affrontarla. Per farlo dovrai anche rispondere ad alcune domande: prima di passare alla domanda successiva controlla la tua risposta con la risoluzione che trovi di seguito.

3.2.1 La questione

Consideriamo le distribuzioni binomiali di parametri n e p, con il vincolo che il prodotto $\lambda = np$ sia costante.
Intendiamo determinare il *limite* di tali distribuzioni per $n \to \infty$.

Tale questione va precisata, ma prima sintetizziamo le ragioni che ci hanno condotto ad essa.

Le ragioni

Per risolvere il problema di conteggio posto nel paragrafo 1.2, abbiamo costruito un modello binomiale di parametri n e p opportuni, dove n rappresenta il numero di prove e p la probabilità di realizzazione dell'evento in esame[14] in una (ogni) prova[15]. Tali parametri verificano la condizione $\lambda = np$, dove λ è il numero medio di realizzazioni dell'evento nell'intervallo costituito dall'unione delle n prove.
Ma poi, nel paragrafo 3.1, abbiamo osservato che è opportuno considerare anche una nuova distribuzione.

L'idea
La nuova distribuzione sarà il limite per $n \to \infty$ delle distribuzioni binomiali che verificano la condizione $\lambda = np$ costante.

A priori non possiamo dire che tale limite sia effettivamente una distribuzione di probabilità. Ma ciò si può dimostrare formalmente.

Precisamente

Vediamo di precisare tale richiesta. Innanzitutto indichiamo con S_n la variabile aleatoria binomiale di parametri n, p, nell'ipotesi $\lambda = np$ costante: essa, in quanto variabile aleatoria binomiale, conta il numero di realizzazioni dell'evento su n prove.
Inoltre indichiamo con X la "nuova" variabile aleatoria: essa ha la distribuzione limite ora indicata. Dunque anch'essa conta il numero di realizzazioni dell'evento nell'intervallo costituito dall'unione delle n prove.

Proseguiamo ricordando che, per definizione di distribuzione discreta di probabilità, la distribuzione della variabile aleatoria X è la funzione g, definita da $g(k) := P(X = k)$ per $k = 0, 1, 2...$
Secondo l'idea prima esposta, g è il limite della distribuzione binomiale. Ossia poniamo:

$$\boxed{P(X = k) := \lim_{n\to\infty} P(S_n = k)}$$

[14] Ad esempio, l'arrivo di una telefonata.
[15] La realizzazione dell'evento si indica spesso come "successo".

Questa è l'idea. Ma affinché essa risulti utile, serve ricavare un'espressione analitica per la funzione g. Ci proponiamo di farlo mediante l'attività che segue.

3.2.2 La costruzione

Il primo passo consiste nel determinare $P(X = 0)$, o meglio, nell'esprimere tale probabilità in funzione del solo parametro λ. Per quanto osservato, deve essere

$$P(X = 0) := \lim_{n \to \infty} P(S_n = 0)$$

pertanto dovremo prima esprimere $P(S_n = 0)$ in termini di λ e di n.[16]

1. *Esprimi* $P(S_n = 0)$ *come indicato.*

Ora per ottenere $P(X = 0)$ basta passare al limite per $n \to \infty$ nell'espressione trovata per $P(S_n = 0)$.

2. *Determina* $P(X = 0)$.

   Suggerimento
   Osserva che si tratta di un limite nella forma indeterminata $[1^\infty]$, perciò, come spesso avviene in questa situazione, conviene riscrivere la funzione nella forma:

$$\left(1 - \frac{\lambda}{n}\right)^n = e^{\log(1-\lambda/n)^n} = e^{n \log(1-\lambda/n)} = e^{\frac{\log(1-\lambda/n)}{1/n}}$$

A questo punto dovremo determinare $P(X = k)$ per ogni $k \geq 1$. Però, invece di procedere con il calcolo diretto, preferiamo seguire un approccio diverso:

a) esprimiamo il rapporto $\dfrac{P(X = k)}{P(X = k - 1)}$ in funzione del solo parametro λ e di k

b) poi utilizziamo tale risultato per ricavare $P(X = k)$ a partire dal valore di probabilità $P(X = 0) = e^{-\lambda}$, procedendo per passi successivi.

Nel paragrafo 3.1 viene mostrato come attuare il proposito espresso nel punto b). Resta dunque da determinare il rapporto richiesto nel punto a). Per farlo, seguiamo lo schema già utilizzato nell'affrontare il caso $k = 0$: ossia determiniamo prima il rapporto tra le corrispondenti probabilità ottenute mediante il modello binomiale.

[16] Abbiamo precisato che vogliamo esprimere $P(X = 0)$ in termini del solo parametro λ. Ma, dato che esso si ottiene al limite per $n \to \infty$, puntiamo ad esprimere $P(S_n = 0)$ in termini anche di n.

3. *Esprimi* $\dfrac{P(S_n = k)}{P(S_n = k-1)}$ *in funzione dei parametri* n, p *e di* k.

Suggerimento
Secondo il modello binomiale, le probabilità di ottenere rispettivamente k e $k-1$ realizzazioni dell'evento su n prove sono

$$P(S_n = k) = \binom{n}{k} p^k (1-p)^{n-k}$$
$$P(S_n = k-1) = \binom{n}{k-1} p^{k-1} (1-p)^{n-k+1}$$

Inoltre, tieni presente che l'espressione esplicita del coefficiente binomiale è

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

Prima di passare al limite e ricavare così il rapporto tra le corrispondenti probabilità nella variabile X, è opportuno compiere un ulteriore passo. Ovvero esprimere il rapporto $\dfrac{P(S_n = k)}{P(S_n = k-1)}$ in funzione del parametro λ, di k e di n.

4. *Esprimi il rapporto* $\dfrac{P(S_n = k)}{P(S_n = k-1)}$ *come indicato.*

Suggerimento
Basta sfruttare l'ipotesi $\lambda = np$.

Disponiamo così di tutti gli elementi per attuare il nostro proposito, ossia determinare il rapporto $\dfrac{P(X = k)}{P(X = k-1)}$ in funzione dei valori noti. Basta passare al limite per $n \to \infty$ nell'espressione appena ricavata per le probabilità binomiali.

5. *Esprimi il rapporto* $\dfrac{P(X = k)}{P(X = k-1)}$ *in funzione del parametro* λ *e di* k.

Suggerimento
Tieni presente che λ e k sono costanti, al variare del parametro n della binomiale.

3.2.3 Risoluzione

Determiniamo $P(X=0) := \lim_{n\to\infty} P(S_n = 0)$ in funzione del solo parametro λ.

1. *Esprimi* $P(S_n = 0)$ *come indicato, ossia in funzione del solo parametro* λ *e di* n.
   $P(S_n = 0)$ è la probabilità di avere 0 realizzazioni dell'evento (successi) su n prove, ossia la probabilità che l'evento non si realizzi (insuccesso) in tutte le n prove. Ora, la probabilità che l'evento non si realizzi in una data prova è $1-p$; pertanto, per la legge della moltiplicazione, applicata nell'ipotesi di indipendenza delle prove, si ha:
   $$P(S_n = 0) = (1-p)^n$$
   ovvero, sfruttando l'ipotesi $\lambda = np$,
   $$\boxed{\boldsymbol{P(S_n = 0) = \left(1 - \frac{\lambda}{n}\right)^n}} \tag{1}$$

2. *Determina* $P(X = 0)$.
   Scriviamo la funzione nella forma indicata dal suggerimento. Poi, per determinarne il limite, possiamo ricorrere essenzialmente[17] al teorema di *de l'Hopital*:
   $$\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n = e^{\lim_{n\to\infty} \frac{\log\left(1-\frac{\lambda}{n}\right)}{1/n}} = e^{\lim_{n\to\infty} \frac{\frac{1}{1-\lambda/n}\cdot\frac{\lambda}{n^2}}{-1/n^2}} =$$
   $$= e^{\lim_{n\to\infty} \frac{-\lambda}{1-\lambda/n}} = e^{-\lambda}$$
   Riassumendo, abbiamo così ricavato:
   $$\boxed{\boldsymbol{P(X = 0) = e^{-\lambda}}} \tag{2}$$
   Osservazione
   In alternativa si può ricondursi ad utilizzare il limite notevole[18]
   $$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = e$$

3. *Esprimi* $\dfrac{P(S_n = k)}{P(S_n = k-1)}$ *in funzione dei parametri* n, p *e di* k.
   Per quanto osservato nel suggerimento, il rapporto in esame si può esprimere nella forma seguente:
   $$\frac{P(S_n = k)}{P(S_n = k-1)} = \frac{\dfrac{n!}{k!(n-k)!}}{\dfrac{n!}{(k-1)!(n-k+1)!}} \cdot \frac{p^k}{p^{k-1}} \cdot \frac{(1-p)^{n-k}}{(1-p)^{n-k+1}} =$$

[17] Per la precisione, il teorema di de l'Hopital si applica al rapporto di due funzioni definite su un intervallo reale. Qui applichiamo il teorema trattando n "come" una variabile reale. Comunque tale modo di operare si può precisare formalmente e dimostrare rigorosamente.

[18] Da cui si ha: $\lim_{n\to+\infty}\left(1-\frac{1}{n}\right)^n = e^{-1}$.

$$= \frac{(k-1)!(n-k+1)!}{k!(n-k)!} \cdot \frac{p}{1-p}$$

Per semplificare la frazione che contiene i fattoriali, basta osservare che $k! = k(k-1)!$ e, analogamente, $(n-k+1)! = (n-k+1)(n-k)!$
Ciò permette di esprimere il rapporto tra le probabilità nella forma voluta:

$$\boxed{\frac{P(S_n = k)}{P(S_n = k-1)} = \frac{n-k+1}{k} \cdot \frac{p}{1-p}} \tag{3}$$

4. *Esprimi il rapporto* $\frac{P(S_n = k)}{P(S_n = k-1)}$ *come indicato, ossia in funzione di* λ*, di* k *e di* n*.*
   Sostituiamo $p = \lambda/n$ nell'espressione (3) e poi semplifichiamo (opportunamente!) i denominatori n che compaiono nella seconda frazione:

$$\frac{P(S_n = k)}{P(S_n = k-1)} = \frac{n-k+1}{k} \cdot \frac{p}{1-p} = \frac{n-k+1}{k} \cdot \frac{\lambda/n}{1-\lambda/n} =$$

$$= \frac{n-k+1}{k} \cdot \frac{\lambda}{n-\lambda}$$

   Pertanto:

$$\boxed{\frac{P(S_n = k)}{P(S_n = k-1)} = \frac{\lambda(n-k+1)}{k(n-\lambda)}} \tag{4}$$

5. *Esprimi il rapporto* $\frac{P(X = k)}{P(X = k-1)}$ *in funzione del parametro* λ *e di* k*.*
   Vale:

$$\frac{P(X = k)}{P(X = k-1)} = \lim_{n\to\infty} \frac{P(S_n = k)}{P(S_n = k-1)} = \lim_{n\to\infty} \frac{\lambda(n-k+1)}{k(n-\lambda)} =$$

$$= \lim_{n\to\infty} \frac{\lambda n}{kn} = \frac{\lambda}{k}$$

   Nell'ultimo passaggio abbiamo utilizzato il fatto che i termini costanti ($-\lambda k + \lambda$ a numeratore, e $-k$ a denominatore) non influiscono sul valore del limite in esame[19],

[19] Precisiamo il significato di tale affermazione dimostrandola formalmente.
Iniziamo raccogliendo il termine n sia a numeratore che a denominatore:

$$\frac{\lambda(n-k+1)}{k(n-\lambda)} = \frac{\lambda n\left(1 - \frac{k}{n} + \frac{1}{n}\right)}{kn\left(1 - \frac{\lambda}{n}\right)} = \frac{\lambda\left(1 - \frac{k}{n} + \frac{1}{n}\right)}{k\left(1 - \frac{\lambda}{n}\right)}$$

Poi, passando al limite per $n \to \infty$, otteniamo $\lim_{n\to\infty} \frac{k}{n} = 0$, $\lim_{n\to\infty} \frac{1}{n} = 0$, $\lim_{n\to\infty} \frac{\lambda}{n} = 0$, in quanto sia k che λ sono costanti.
Pertanto, concludiamo che vale

$$\lim_{n\to\infty} \frac{\lambda(n-k+1)}{k(n-\lambda)} = \lim_{n\to\infty} \frac{\lambda\left(1 - \frac{k}{n} + \frac{1}{n}\right)}{k\left(1 - \frac{\lambda}{n}\right)} = \frac{\lambda}{k}$$

ciò si può esprimere dicendo che essi hanno un ordine di infinito minore rispetto ad n. Possiamo così concludere che vale:

$$\boxed{\frac{P(X=k)}{P(X=k-1)} = \frac{\lambda}{k}} \tag{5}$$

3.2.4 Conclusione

Dunque, mediante la nostra costruzione abbiamo ottenuto la seguente espressione analitica per la distribuzione della variabile aleatoria X:

- $P(X=0) = e^{-\lambda}$,
- $P(X=k) = \dfrac{\lambda}{k} \cdot P(X=k-1), \qquad \text{per } k = 1, 2, 3, \ldots$

Si tratta proprio della formula che avevamo anticipato nel paragrafo 3.1.