Indice

I materiali per lo studente

In questo capitolo raccogliamo i materiali a supporto delle attività didattiche e indichiamo il ruolo dei video e dei file GeoGebra nello sviluppo del percorso.
Ricordiamo che i video sono stati caricati su YouTube e ad essi si può accedere direttamente cliccando sui link in rosso che riportano il relativo indirizzo URL.
I file GeoGebra sono allegati al presente elaborato e ad essi si può accedere direttamente tramite il link rappresentato, nel testo, dal loro nome (ad esempio *PoissonBinomiale.ggb*, nel paragrafo 4.2).
Le dispense che seguono, invece, si rivolgono direttamente allo studente e sono costruite in modo che i paragrafi siano il più possibile indipendenti[1], per questo talvolta presentano alcune ripetizioni.

[1] Ciò per dare la possibilità al docente di scegliere quali materiali utilizzare e in che ordine proporli.

1 Situazioni motivanti, la questione

1.1 Dalle motivazioni alla questione - video

*Nel video “**Modellizzazione di eventi “rari”**” ti vengono presentate alcune situazioni che possono essere schematizzate efficacemente mediante il modello probabilistico che ci accingiamo a studiare: si tratta del modello di Poisson. Viene inoltre illustrata la questione che ci poniamo e che condurrà alla costruzione di tale modello.*

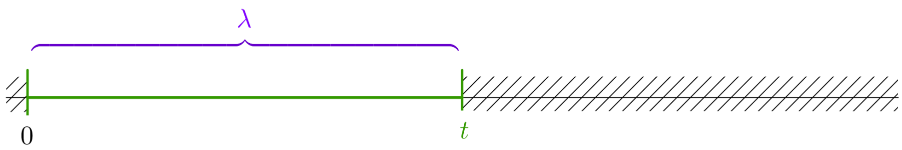
Trovi il video al link seguente: *www.youtube.com/watch?v=kc1WosTkCMA&feature=youtu.be*.

1.2 La questione

Di seguito ti proponiamo la questione che abbiamo illustrato nel video "Modellizzazione di eventi "rari"" che ti consigliamo di esaminare.

Un problema di conteggio[2]

Consideriamo un evento aleatorio "*raro*".
Fissiamo un intervallo di *ampiezza* $\boldsymbol{t}$.
Sia $\boldsymbol{\lambda}$ il *numero medio di realizzazioni* dell'evento in tale intervallo.



La domanda che ci poniamo è la seguente.

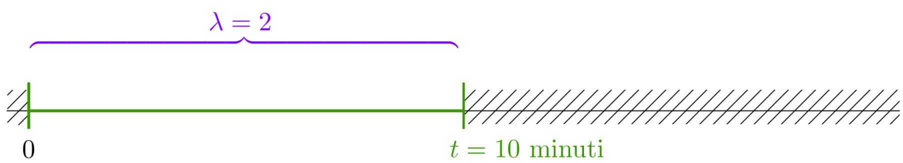
Qual è la probabilità che nell'intervallo fissato *l'evento si realizzi esattamente* $\boldsymbol{k}$ *volte*?

Per capire meglio i termini della questione, esaminiamo un esempio.

Il centralino

Ad un centralino tra le 9 e le 12 arriva, in media, una telefonata ogni 5 minuti.
Qual è la probabilità che in 10 minuti arrivino *esattamente* 3 telefonate?

In questa situazione l'intervallo di riferimento ha ampiezza $\boldsymbol{t = 10}$ minuti.
Per calcolare il numero medio di telefonate nell'intervallo partiamo dall'ipotesi che si abbia, in media, una telefonata ogni 5 minuti. *Decidiamo* che il numero di telefonate sia proporzionale al tempo: quindi avremo mediamente 2 telefonate in 10 minuti. Perciò poniamo $\boldsymbol{\lambda = 2}$.



Infine siamo interessati all'arrivo di 3 telefonate, da cui $\boldsymbol{k = 3}$.

[2] In modo informale diremo "raro" un evento che si verifica con probabilità "piccola": anche se effettuassimo numerose prove otterremmo, molto probabilmente, poche realizzazioni dell'evento stesso.

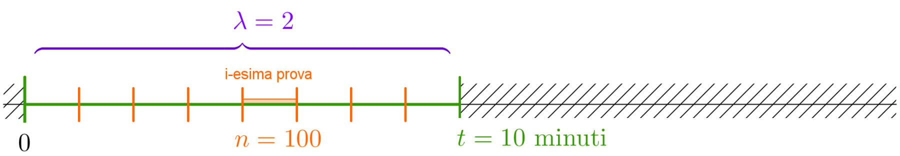
2 Il modello binomiale

2.1 Il centralino - attività

Prima di considerare il caso generale ti proponiamo di affrontare un problema specifico che è già stato introdotto nel video "Modellizzazione di eventi "rari"".

Ad un centralino tra le 9 e le 12 arriva, in media, una telefonata ogni 5 minuti.
Qual è la probabilità che in 10 minuti arrivino *esattamente* 3 telefonate?

Suggerimento
Proviamo a risolvere mediante una opportuna **distribuzione binomiale**.
Allo scopo suddividiamo l'intervallo $[0, t]$ in n intervallini uguali e supponiamo che su ciascuno di essi sia impossibile ricevere più di una chiamata. Scegliamo, ad esempio $n = 100$; dunque ciascun intervallino avrà un'ampiezza di 6 secondi.



A ciascun intervallino corrisponde una prova e il numero di telefonate in arrivo[3] si può descrivere tramite una distribuzione binomiale.

[3] Uno dei due esiti possibili delle prove si indica spesso, convenzionalmente, con il termine "successo"; l'altro esito con "insuccesso".

2.1.1 Risoluzione

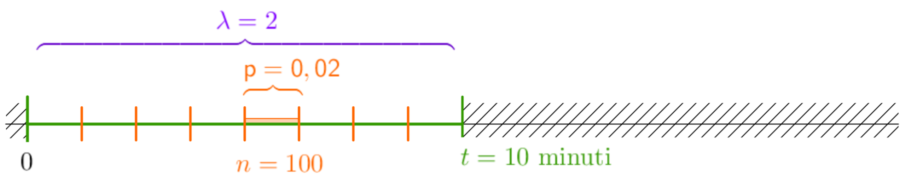
Secondo la costruzione proposta, la situazione si può interpretare in questo modo:

- una **sequenza** di $n = 100$ **prove**, una per ogni intervallino di ampiezza 6 secondi,
- ciascuna prova ha **due soli esiti** possibili: "*arriva una telefonata*", "*non arriva una telefonata*" nell'intervallino ad essa relativo.

Per modellizzare mediante la distribuzione binomiale resta da determinare la **probabilità** p che si realizzi l'evento "*arriva una telefonata nell'intervallino di 6 secondi*".
L'idea è di seguire un approccio frequentista, ossia interpretare la probabilità come frequenza relativa dell'evento in esame sull'insieme delle 100 prove. Cioè[4]:

$$p = \frac{\#\ \textit{realizzazioni evento}}{\#\ \textit{prove}} = \frac{2}{100} = 0{,}02.$$

Nel determinare il numero di realizzazioni dell'evento abbiamo utilizzato il fatto che nell'intervallo $[0, 10]$ vi sono in media $\lambda = 2$ telefonate, come discusso nel paragrafo 1.2.



Osservazione
Stiamo assumendo che p sia **costante** per ogni intervallino, ossia che il numero medio di telefonate sia lo stesso su ciascun intervallino di 6 secondi e non solo su ogni intervallo di 5 minuti come indicato nel testo.

Inoltre assumiamo che l'arrivo di una telefonata sia **indipendente** da quello delle altre.

È opportuno assumere queste due ipotesi affinché abbia senso modellizzare mediante la binomiale.

Abbiamo così tutti gli elementi per concludere che la probabilità richiesta, ossia la probabilità che in 10 minuti arrivino 3 telefonate è:

$$\binom{100}{3} \cdot 0{,}02^3 \cdot 0{,}98^{97} \simeq \mathbf{0{,}1823}.$$

Infatti in generale, secondo il modello binomiale, la probabilità di avere k successi su n prove ciascuna con probabilità p di successo[5] è:

$$\binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}.$$

[4] In generale, dato un insieme E che ha un numero finito di elementi, con il simbolo $\#E$ si indica il numero di elementi di E. Analogamente, nel nostro contesto con il simbolo # indichiamo il "numero di".

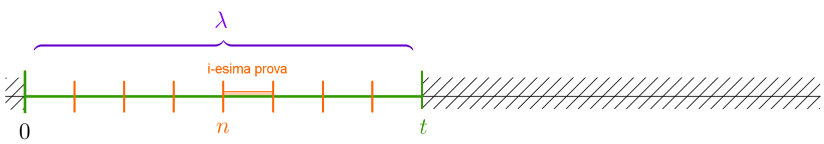
[5] Abbiamo già precisato nella nota (2) cosa si intende per successo in una prova.

2.2 Il caso generale

Torniamo a considerare il problema generale di conteggio descritto nel paragrafo 1.2. Un possibile modello probabilistico per descrivere la situazione è quello binomiale, quindi proviamo a modellizzare la situazione mediante quest'ultima distribuzione, in modo analogo a quanto visto nell'esempio del centralino.

2.2.1 Costruzione del modello

L'idea
Suddividiamo l'intervallo $[0, t]$ in $\boldsymbol{n}$ **intervallini uguali** in modo che in ogni intervallino si abbia "al più" una sola realizzazione dell'evento.



La situazione si può interpretare nel modo seguente.

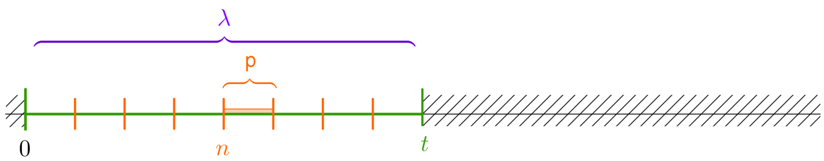
- Consideriamo una **sequenza**[6] di n **prove**.
- Ciascuna prova ha **due soli esiti** possibili: "*l'evento si realizza*", "*l'evento non si realizza*" nell'intervallino ad essa relativo[7].

  Siccome vogliamo al più un solo successo su ogni intervallino, consideriamo $\boldsymbol{n}$ **"grande"** rispetto al numero medio di successi nell'intervallo $[0, t]$.

- A questo punto dobbiamo determinare la **probabilità** p che l'evento si realizzi nell'intervallino. L'idea è di utilizzare un approccio frequentista, ossia considerare la frequenza relativa dell'evento in esame sull'insieme delle n prove.
  Abbiamo così:
$$p = \frac{\#\ \textit{realizzazioni evento}}{\#\ \textit{prove}} = \frac{\lambda}{n}$$

  Assumiamo che $\boldsymbol{p}$ sia **costante** per ogni prova. Osserviamo che ciò equivale ad ipotizzare che il numero medio di realizzazioni sia costante su tutti gli intervallini considerati.

- Inoltre le prove sono **indipendenti**, ovvero nessuna realizzazione è influenzata dalle altre.



Il modello ottenuto, quindi, è la **distribuzione binomiale** con parametri n e p definiti sopra.

[6] Tale sequenza si indica spesso come "schema di Bernoulli" o "schema successo-insuccesso".

[7] Anche in questo caso spesso l'esito che corrisponde alla realizzazione dell'evento si indica come "*successo*"; l'altro come "*insuccesso*". Nell'esempio del centralino, considerato in precedenza, l'evento è l'arrivo di una telefonata.

2.2.2 In sintesi

Abbiamo così risolto la questione di conteggio posta all'inizio, infatti abbiamo fornito un modello: la **distribuzione binomiale** che ha parametri n e p del tipo indicato (in particolare tali che $np = \lambda$).
Di tale modello è nota l'espressione analitica. Precisamente:

* sia S la **variabile aleatoria** che conta il numero di realizzazioni dell'evento nell'intervallo $[0, t]$;
* la **probabilità** che S assuma valore k è

$$\boldsymbol{P(S = k) = \binom{n}{k} \cdot p^n \cdot (1-p)^{n-k}}, \qquad \text{per } k = 0, 1, 2, ...$$

Dove con il simbolo $\binom{n}{k}$ si denota il coefficiente binomiale n su k, la cui espressione esplicita è $\dfrac{n!}{k!(n-k)!}$ [8].

2.2.3 Valutazione del modello

Il modello binomiale così costruito:

- **non** è **efficiente** dal punto di vista computazionale
  visto che la probabilità è espressa in termini dei coefficienti binomiali e il loro calcolo è articolato, infatti, il fattoriale è una funzione che cresce molto rapidamente:

$$10! \simeq 3.6 \cdot 10^6, \qquad 20! \simeq 2,4 \cdot 10^{18}, \qquad 70! > 10^{100}$$

- **non** è molto **espressivo**
  infatti, il parametro λ, che caratterizza la situazione, non compare esplicitamente.

Quindi è utile costruire un nuovo modello per rappresentare la situazione.

[8] Dato un numero naturale m, si indica con $m!$ il prodotto $m \cdot (m-1) \cdot ... \cdot 2 \cdot 1$, se $m \geq 1$; e 1 se $m = 0$.

3 Un nuovo modello

3.1 Il modello

Costruzione del modello [9]

L'idea
Passare al limite per $\boldsymbol{n \to \infty}$ nella distribuzione binomiale, nell'ipotesi che $\boldsymbol{\lambda} = np$ sia **costante**.
Osserviamo che ciò significa passare al limite per $p \to 0$.

3.1.1 Le ipotesi del modello

Iniziamo esplicitando le ipotesi che è opportuno assumere affinché abbia senso schematizzare mediante il "nuovo" modello il problema generale di conteggio proposto nel paragrafo 1.2.

1. Una data realizzazione[a] dell'evento ha la **stessa probabilità** di verificarsi in ogni intervallo di uguale ampiezza.

2. Ogni realizzazione non è influenzata dalle altre, ovvero le realizzazioni sono **indipendenti**.

3. Il numero medio di realizzazioni λ è "piccolo" rispetto al numero delle prove (evento "**raro**").

[a] Nell'esempio del centralino, già analizzato, tale evento è l'arrivo di una telefonata.

Possiamo sintetizzare le ipotesi 1. e 2. dicendo che il processo è "**casuale**" e **uniforme**[10].

3.1.2 La nuova distribuzione

Si può dimostrare[11] che passando al limite nel modo indicato si ottiene una "nuova" variabile aleatoria X che conta, come la variabile aleatoria S, il numero di realizzazioni dell'evento nell'intervallo e una distribuzione di probabilità per X. Precisamente, la **distribuzione di probabilità** di X si può ottenere *per passi* in questo modo:

- si inizia da
$$P(X = 0) = e^{-\lambda}$$

- per ogni naturale $k \geq 1$, il valore $P(X = k)$ si ottiene da $P(X = k - 1)$ moltiplicando per $\frac{\lambda}{k}$, ovvero
$$P(X = k) = \frac{\lambda}{k} \cdot P(X = k - 1)$$

[9] Preciseremo il significato dell'affermazione all'interno del riquadro "L'idea" nel paragrafo 3.2.
[10] Utilizzeremo più volte nel seguito questi due termini e nel farlo intenderemo riferirci proprio alle ipotesi 1. e 2.
[11] Lo faremo, mediante un'attività guidata, nel paragrafo 3.2.

Pertanto se si vuole esplicitare l'espressione analitica di $P(X = k)$ basta applicare più volte l'ultima uguaglianza a partire dal caso $k = 0$:

$$\mathrm{P}(X = 0) = e^{-\lambda}$$

$$\boxed{\cdot \frac{\lambda}{1}} \quad \mathrm{P}(X = 1) = \boxed{\frac{\lambda}{1}} \cdot \mathrm{P}(X = 0) = \lambda \cdot e^{-\lambda}$$

$$\boxed{\cdot \frac{\lambda}{2}} \quad \mathrm{P}(X = 2) = \boxed{\frac{\lambda}{2}} \cdot \mathrm{P}(X = 1) = \frac{\lambda}{2} \cdot \lambda \cdot e^{-\lambda} = \frac{\lambda^2}{2} \cdot e^{-\lambda}$$

$$\boxed{\cdot \frac{\lambda}{3}} \quad \mathrm{P}(X = 3) = \boxed{\frac{\lambda}{3}} \cdot \mathrm{P}(X = 2) = \frac{\lambda}{3} \cdot \frac{\lambda^2}{2} \cdot e^{-\lambda} = \frac{\lambda^3}{3!} \cdot e^{-\lambda}$$

$$\vdots$$

In generale, quindi, si ottiene:

$$\boldsymbol{P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}}$$

Si può dimostrare formalmente che questa è una distribuzione di probabilità; essa prende il nome di **distribuzione di Poisson**.

3.1.3 Un esempio di calcolo: ancora il centralino

Possiamo modellizzare mediante la nuova distribuzione il problema del centralino, discusso dal paragrafo 1.2.
Ricordando che, in tale situazione, il parametro che caratterizza la distribuzione è $\lambda = 2$ e che siamo interessati alla realizzazione di $k = 3$ telefonate, abbiamo:

$$\mathrm{P}(X = 0) = e^{-2}$$

$$\boxed{\cdot \frac{2}{1}} \quad \mathrm{P}(X = 1) = \boxed{\frac{2}{1}} \cdot \mathrm{P}(X = 0) = 2\, e^{-2}$$

$$\boxed{\cdot \frac{2}{2}} \quad \mathrm{P}(X = 2) = \boxed{\frac{2}{2}} \cdot \mathrm{P}(X = 1) = 1 \cdot 2\, e^{-2} = 2\, e^{-2}$$

$$\boxed{\cdot \frac{2}{3}} \quad \mathrm{P}(X = 3) = \boxed{\frac{2}{3}} \cdot \mathrm{P}(X = 2) = \frac{2}{3} \cdot 2\, e^{-2} = \frac{4}{3}\, e^{-2} \approx 0{,}180$$

Oppure si giunge allo stesso risultato applicando direttamente la formula che fornisce l'espressione analitica di $P(X = k)$:

$$P(X = 3) = \tfrac{2^3}{3!} e^{-2} \simeq \mathbf{0{,}180}.$$

Concludiamo dunque che la probabilità che si realizzino esattamente 3 telefonate in 10 minuti è quasi del 20%.

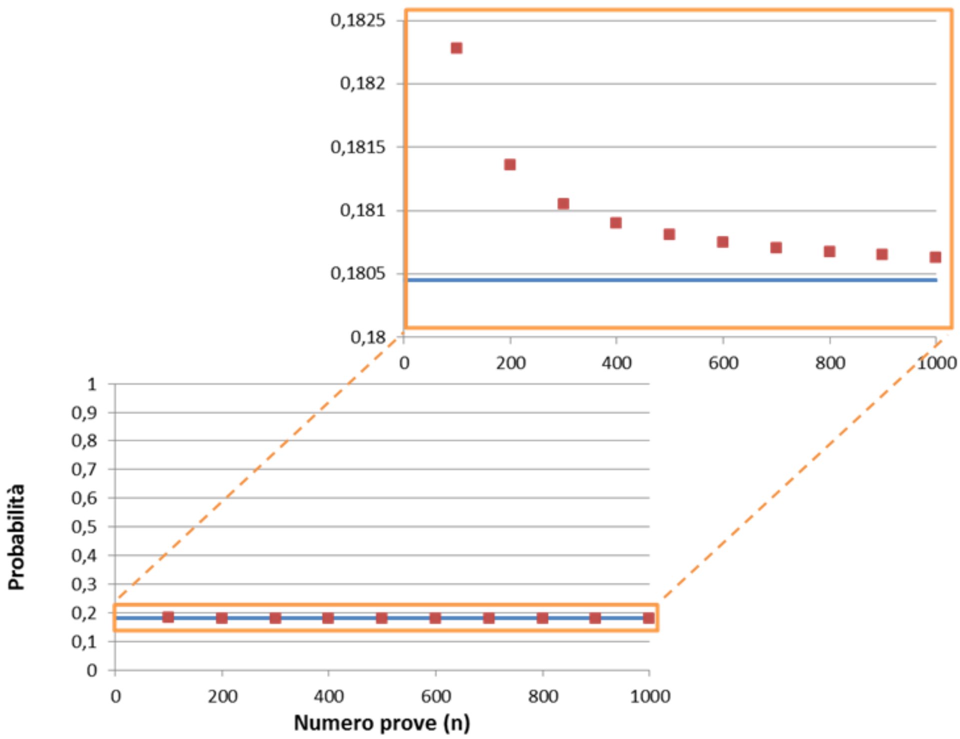
Osservazione
Torniamo a considerare il modello binomiale che avevamo adottato inizialmente[12] per schematizzare la situazione delle telefonate. Abbiamo scelto di suddividere l'intervallo $[0, 10]$ in $n = 100$ intervallini e dunque abbiamo considerato $n = 100$ prove.
Come sarebbe cambiata la probabilità richiesta se avessimo scelto un altro valore per n, ad esempio $n = 200$?
In tal caso il parametro $\lambda = 2$ (numero medio di telefonate nell'intervallo di 10 minuti) rimane lo stesso, mentre il valore di p cambia:

$$p = \frac{\#\ \textit{realizzazioni evento}}{\#\ \textit{prove}} = \frac{2}{200} = 0{,}01.$$

E mediante questo secondo modello binomiale (cioè con $n = 200$ e p =0,01), la probabilità richiesta vale:

$$\binom{200}{3} \cdot 0{,}01^3 \cdot 0{,}99^{197} \simeq \mathbf{0{,}1814}.$$

I seguenti grafici mostrano cosa accade all'aumentare del numero n di prove.
Le ordinate dei punti rossi rappresentano i valori di probabilità ottenuti mediante il modello binomiale, mentre la retta blu ha equazione $y = c$, dove $c \simeq$0,180 è il valore trovato utilizzando il nuovo modello, ossia la distribuzione di Poisson. Possiamo osservare che il valore ottenuto col nuovo modello si può considerare, in un senso opportuno[13], il limite a cui tende il valore ottenuto dal modello binomiale, al crescere di n.



Dal grafico in alto, sembra che i valori di probabilità ottenuti al variare di n siano molto diversi tra loro. Ciò, però, dipende fortemente dalla scala adottata: utilizzando come range, per la probabilità l'intero l'intervallo $[0, 1]$, come nel grafico in basso, i valori risultano praticamente indistinguibili.

[12] Vedi il paragrafo 1.2.
[13] Come vedremo nel paragrafo 3.2.

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*:
   $$\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n = e^{\lim\limits_{n\to\infty}\frac{\log\left(1-\frac{\lambda}{n}\right)}{1/n}} = e^{\lim\limits_{n\to\infty}\frac{\frac{1}{1-\lambda/n}\cdot\frac{\lambda}{n^2}}{-1/n^2}} =$$
   $$= e^{\lim\limits_{n\to\infty}\frac{-\lambda}{1-\lambda/n}} = e^{-\lambda}$$
   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* $\dfrac{P(S_n = k)}{P(S_n = k-1)}$ *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* $\dfrac{P(X = k)}{P(X = k-1)}$ *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:

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

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

4 La distribuzione di Poisson

4.1 Facciamo il punto

Definizione

Consideriamo un numero reale $\lambda > 0$.
Sia X una variabile aleatoria che può assumere i valori $k = 0, 1, 2, ...$
X ha **distribuzione di Poisson** se vale:

$$\boldsymbol{P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}}$$

Interpretazione

Tale *distribuzione* modellizza, in opportune ipotesi (indicate nel paragrafo 3.1), situazioni caratterizzate da eventi "**rari**".

La *variabile aleatoria* X **conta** il numero di realizzazioni dell'evento in un intervallo e il *parametro* λ rappresenta il **numero medio** di **realizzazioni** in tale intervallo.

Inoltre la distribuzione di Poisson **approssima** la distribuzione *binomiale*.

Osservazione

La distribuzione di Poisson approssima la binomiale[20] quando il numero $\boldsymbol{n}$ delle prove è **"grande"** e la probabilità $\boldsymbol{p}$ di realizzazione dell'evento nella prova è **"piccola"**[21].
Infatti, come abbiamo visto[22], la distribuzione di Poisson si può ottenere considerando il limite delle binomiali di parametri n, p, per n che tende all'infinito nell'ipotesi $np =$ costante.

Spesso si considera l'approssimazione accettabile per $np \leq 10$ e $n > 50$.

[20] Nel senso che preciseremo nel paragrafo 4.2.

[21] Questa legge che permette di approssimare la distribuzione binomiale con la Poisson ha carattere empirico e prese il nome di "legge dei piccoli numeri", quando nel 1898 Bortkiewicz la introdusse per studiare il numero di cavalieri dell'esercito prussiano morti per il calcio di un cavallo.

[22] Vedi quanto indicato nel paragrafo 3.2.

4.1.1 Nota storica

La distribuzione che abbiamo ottenuto prende il nome di **distribuzione di Poisson** dal matematico e fisico francese Siméon-Denis Poisson (1781-1840, fig.) che nel 1837 pubblicò "*Recherches sur la probabilité des jugements en matière criminelle et en matière civile*" trattato in cui, per la prima volta, venne introdotta esplicitamente la formula precedente in un contesto matematico. Solo più tardi, nei primi decenni del secolo successivo, ci si rese conto che il modello della distribuzione di Poisson modellizza efficacemente vari processi fisici come il decadimento radiattivo, l'effetto fotoelettrico e quello termoionico.

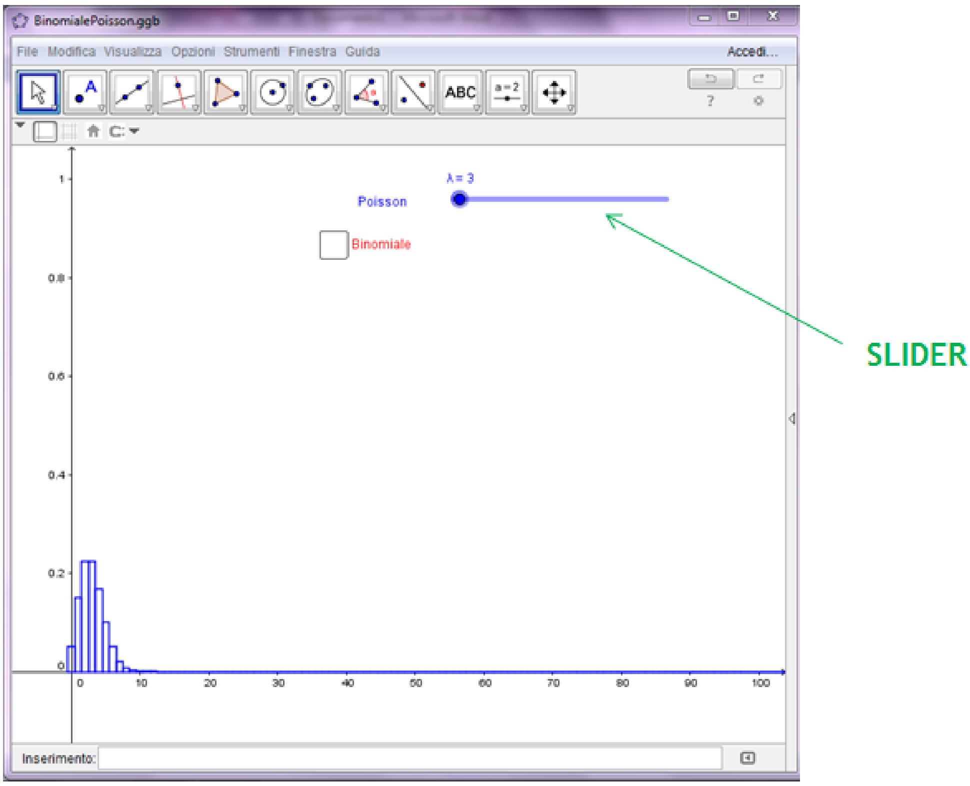
4.2 Significato geometrico del parametro - attività con GeoGebra

Utilizza il file GeoGebra PoissonBinomiale.ggb per tracciare i grafici della distribuzione di Poisson per alcuni valori. Considera valori di λ sia "piccoli" (quali, ad esempio, λ =0,1) sia "grandi" (quali, ad esempio, $\lambda = 100$).

> *Prova a descrivere come varia il grafico della distribuzione al crescere del parametro λ.*

4.2.1 Come utilizzare il file *PoissonBinomiale.ggb*

Il file è provvisto di uno slider (la barra indicata in figura) mediante il quale si può variare il valore del parametro λ.
Per il valore di λ assunto dallo slider viene (in automatico) visualizzato il grafico della distribuzione di Poisson di parametro λ.



4.2.2 Risoluzione e conclusioni

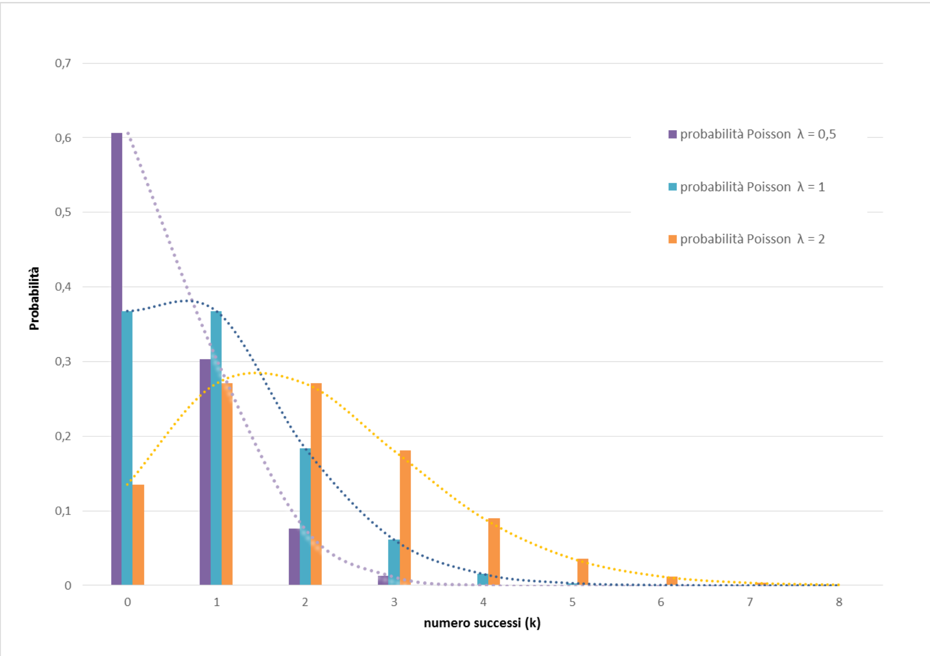
Al crescere di λ:

- il "punto di massimo[a]" della distribuzione cresce;
- l'"apertura" del grafico cresce e il valore massimo diminuisce.

[a]Ossia per ogni λ fissato, il numero naturale k per cui è massimo il valore di probabilità $P(X = k)$, dove X è la variabile aleatoria di Poisson di parametro λ.

Forniremo più avanti[23] una giustificazione intuitiva di tali fatti. Essa, come vedremo, si basa sul fatto che λ è sia il valore atteso sia la varianza della distribuzione di Poisson di parametro λ.

Per ora ci possiamo accontentare di osservare i due fatti nella figura seguente[24]. In essa sono rappresentati i grafici delle distribuzioni di Poisson per alcuni valori di λ.



[23]Nel paragrafo 7.
[24]Nel grafico il tratteggio è stato aggiunto solo per dare l'idea dell'andamento della distribuzione di Poisson che è, comunque, una distribuzione discreta.

4.3 Convergenza della binomiale - attività con GeoGebra

Utilizza il file GeoGebra PoissonBinomiale.ggb per confrontare la distribuzione di Poisson di parametro λ con la binomiale di paremetri n e $p = \lambda/n$. Allo scopo rispondi ai seguenti quesiti.

1. *Considera prima il caso $\lambda = 2$.*

   Trova il più piccolo n per il quale ogni scarto[a] *tra il valore di probabilità ottenuto con il modello binomiale e il corrispondente ottenuto con il modello di Poisson risulta minore di 0,01.*

2. *Fissa poi alcuni valori del parametro λ. Ad esempio $\lambda = 2$; $\lambda = 10$; $\lambda = 0{,}1$; $\lambda = 100$.*

   a) *Per ciascuno di tali valori, traccia i grafici della distribuzione binomiale per alcuni valori del parametro n (p resta determinato dalla condizione $p = \lambda/n$) e confrontali con il grafico della distribuzione binomiale di parametro λ. Per quali valori di n l'approssimazione della binomiale con la Poisson ti sembra "buona"?*

   b) *Per quale dei valori di λ considerati hai indicato valori maggiori per n nel punto a)?*

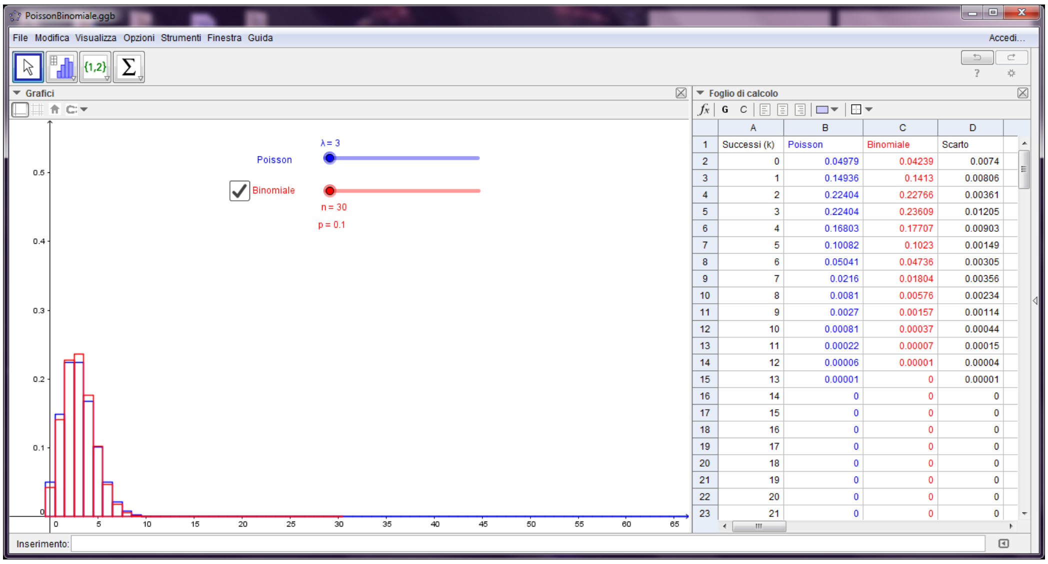
   Attenzione: mantieni la stessa scala per tutti e quattro i casi.

[a] Lo scarto è il valore assoluto della differenza tra i corrispondenti valori di probabilità ottenuti mediante la distribuzione binomiale e la distribuzione di Poisson.

4.3.1 Come utilizzare il file *PoissonBinomiale.ggb*

Oltre allo slider relativo al parametro λ, il file ha anche uno slider relativo al parametro n della binomiale.

Puoi confrontare i grafici delle due distribuzioni (identificati mediante il colore dei rispettivi slider), ma puoi osservare anche gli scarti tra esse (nel foglio di calcolo[25]).



Attenzione: per visualizzare il grafico della distibuzione binomiale, oltre a quello della distribuzione di Poisson, si deve cliccare sull'opzione "binomiale".

[25]Il foglio di calcolo può essere attivato dal menù "visualizza".

4.3.2 Risoluzione

1. Nel caso $\lambda = 2$ il più piccolo n per cui è minore di 0,01 *ogni* scarto tra i corrispondenti valori di probabilità delle due distribuzioni, è $n = 29$.

2. La domanda è qualitativa, quindi il risultato non è univoco, dato che non abbiamo precisato cosa intendiamo per "buona" approssimazione.
   Possiamo dire che si ottiene una "buona" approssimazione[26]:
   - nel caso $\lambda = 2$ per valori di n intorno a 30;
   - nel caso $\lambda = 10$ per valori di n intorno a 75;
   - nel caso $\lambda =$0,1 per valori di n intorno a 2;
   - nel caso $\lambda = 100$ per valori di n intorno a 300.

   Pertanto nei quattro casi considerati, più λ è grande e più n deve essere grande, se si vuole ottenere una "buona" approssimazione.

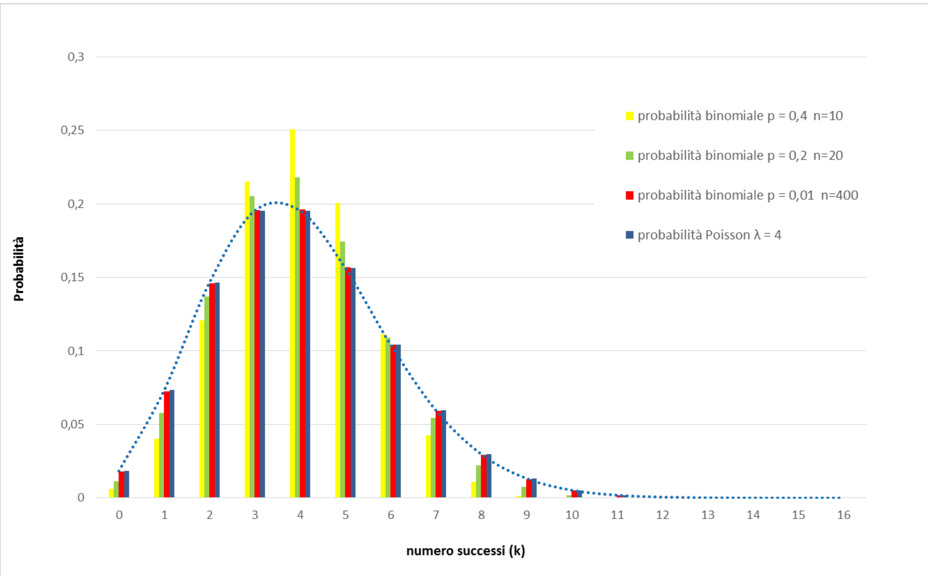
4.3.3 Conclusione

In generale:

> la distribuzione di Poisson di parametro λ approssima la distribuzione binomiale per n "grande" nell'ipotesi $\lambda = np$. E dunque p "piccolo".

Tale risultato[27] è conseguenza del fatto che la Poisson è il limite della distribuzione binomiale per n che tende all'infinito, nell'ipotesi $\lambda = np$ costante.

Con questa attività abbiamo precisato, in alcuni casi, tale approssimazione. Spesso si considera accettabile l'approssimazione per $n > 50$ e $np \leq 10$[28].
In figura[29] vediamo che l'approssimazione con la binomiale è migliore per valori di n più grandi (nell'ipotesi $\lambda = np = 4$).

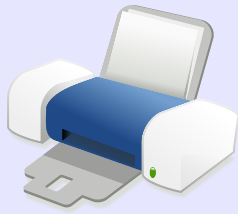


[26] Per indicare i valori abbiamo seguito il seguente criterio: ogni scarto tra i corrispondenti valori di probabilità nei due modelli è minore di 0,01.
[27] Già presentato nel paragrafo 4.
[28] Vedi quanto esposto nel paragrafo 4.
[29] Nel grafico il tratteggio è stato aggiunto solo per dare l'idea dell'andamento della distribuzione di Poisson che è, comunque, una distribuzione discreta.

5 Non solo calcoli

5.1 Errori di stampa

Un compositore tipografico commette, in media, un errore di stampa ogni 3000 parole.
Predispone un libro che contiene 300 parole per pagina.
Costruisci un modello probabilistico che descriva gli errori per pagina.



É sensato modellizzare la situazione mediante:

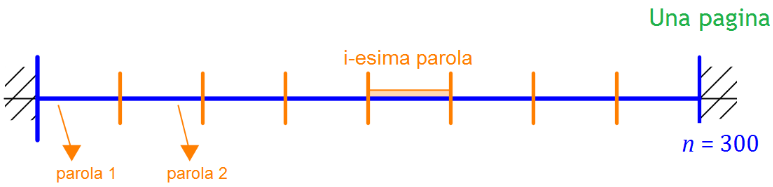
a) il modello binomiale;

b) il modello di Poisson.

Vediamo allora come costruire ciascuno dei due modelli.

a) **Modello binomiale**

Possiamo schematizzare la composizione di una pagina del libro, in modo analogo a quanto visto nell'esempio del centralino[30].
Ossia interpretiamo la situazione in questo modo:



- una sequenza di $n = 300$ prove, una per ogni parola;
- ciascuna prova può avere due soli esiti ("*errore*", "*non errore*"); la probabilità p di errore è $p = \dfrac{1}{3.000}$.

Assumiamo, inoltre, che le prove siano indipendenti, cioè che l'errore nel comporre una parola non influenzi la composizione delle altre.
In queste ipotesi[31] ha senso modellizzare mediante la binomiale.

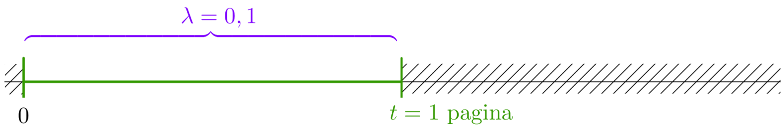
Precisamente:
indichiamo con S la variabile aleatoria che conta il numero di errori per pagina.
Allora, la distribuzione di S è

$$P(S = k) = \binom{\boldsymbol{300}}{\boldsymbol{k}} \left(\frac{\mathbf{1}}{\mathbf{3.000}}\right)^{\boldsymbol{k}} \cdot \left(\frac{\mathbf{2.999}}{\mathbf{3.000}}\right)^{\mathbf{300}-\boldsymbol{k}} \qquad \text{per } k = 0, 1, 2, \ldots$$

[30] Vedi il paragrafo 1.2.
[31] Cioè, in sostanza nell'ipotesi che le prove avvengano nelle stesse condizioni e siano indipendenti.

b) **Modello di Poisson**

Per costruire il modello di Poisson dobbiamo determinare il parametro λ che indica il numero medio di errori per pagina.
Allo scopo decidiamo[32] di seguire uno schema proporzionale: su 3.000 parole si ha, in media, 1 errore, dunque su 300 parole, ossia su 1/10 delle 3.000 parole, ci aspettiamo, in media, 0,1 errori. Quindi poniamo λ =0,1.



Affinché abbia senso modellizzare con la distribuzione di Poisson assumiamo che la distribuzione degli errori sia "casuale" e uniforme[33]; inoltre osserviamo che l'evento "errore nella parola" è "raro" (il numero di errori per pagina è "piccolo" rispetto al numero di parole in essa contenute).

Sia, ora, X la variabile aleatoria che conta il numero di errori per pagina.
Secondo il modello di Poisson, si ha:

$$P(X = k) = \frac{\boldsymbol{0{,}1^k}}{\boldsymbol{k!}} \boldsymbol{e^{-0{,}1}} \qquad \text{per } k = 0, 1, 2, ...$$

Se poi vogliamo determinare il valore di probabilità relativo ad un dato k, ad esempio $k = 3$, possiamo procedere come indicato nel problema del centralino[34].

Confrontiamo le due distribuzioni
Confrontiamo i valori di probabilità che forniscono i due modelli. Otteniamo, arrotondando alla quinta cifra decimale:

# errori (k)	**Binomiale** $P(S = k)$	**Poisson** $P(X = k)$
0	0,90482	0,90484
1	0,09051	0,09048
2	0,00451	0,00452
3	0,00015	0,00015
4	0,00000	0,00000

per $k \geq 4$ i valori di probabilità sono "praticamente nulli".

Il modello di Poisson fornisce valori "vicinissimi" a quelli ottenuti tramite il modello della distribuzione binomiale.
Notiamo infine che vale: $np = 300 \cdot \dfrac{1}{3.000}$ =0,1= λ.

[32] In sostanza stiamo assumendo che l'errore si distribuisca in modo omogeneo nel libro. Ribadiamo però che, per quanto ragionevole, si tratta di una nostra decisione, tra altre possibili.
[33] Nel senso indicato nel paragrafo 3.1.
[34] Come abbiamo visto nel paragrafo 3.1.

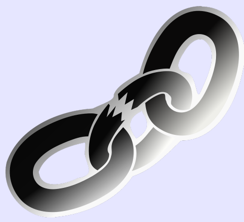
5.2 Dall'Esame di Stato del Liceo Scientifico

5.2.1 Sessione suppletiva 2015, Quesito 7

Una fabbrica produce il 3% di prodotti difettosi. Determinare la probabilità che in un campione di 100 prodotti ve ne siano 2 difettosi, usando:

a) *la distribuzione binomiale;*

b) *la distribuzione di Poisson.*



a) **Modello binomiale**

Interpretiamo la situazione in questo modo:

- una sequenza[35] di $\boldsymbol{n = 100}$ **prove**, una per ogni prodotto del campione,
- ciascuna prova ha due soli esiti possibili: "il prodotto è difettoso", "il prodotto non è difettoso"; la probabilità p che il pezzo sia difettoso è $\boldsymbol{p = 0{,}03}$.

Inoltre dobbiamo assumere che tutti i controlli siano tra loro *indipendenti* e vengano effettuati nelle stesse condizioni.

Dunque la richiesta si può così riformulare:

nel modello ora descritto, qual è la probabilità di avere esattamente[36] *2 pezzi difettosi?*

Quindi secondo tale modello binomiale si ha che la probabilità richiesta è[37]:

$$\binom{100}{2} \cdot 0{,}03^2 \cdot 0{,}97^{98} \simeq \mathbf{0{,}225}.$$

b) **Modello della distribuzione di Poisson**

Affinché abbia senso modellizzare con la distribuzione di Poisson assumiamo che la produzione del pezzo difettoso avvenga in modo "casuale" e uniforme[38]. Per costruire il modello dobbiamo determinare il parametro λ: esso è il numero medio di prodotti difettosi, nel campione considerato.
Dato che il campione è composto da 100 prodotti, si ha

$$\boldsymbol{\lambda = 3}$$

[35] Tale sequenza si indica spesso come "schema di Bernoulli" o "schema successo-insuccesso".

[36] Interpretiamo in questo modo la richiesta "2 pezzi difettosi" formulata nel testo del quesito. Un'altra possibile interpretazione poteva essere "**almeno** 2 pezzi difettosi".

[37] Ricordiamo che la distribuzione di una variabile aleatoria binomiale S è data da

$$P(S = k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}.$$

Discuteremo, alla fine della risoluzione, un modo per ricostruire tale formula, volta per volta.

[38] Nel senso indicato nel paragrafo 3.1.

In alternativa si poteva determinare λ anche mediante l'uguaglianza

$$\lambda = np.$$

In ogni caso, per determinare la probabilità richiesta (ossia la probabilità che nel campione vi siano *esattamente* 2 prodotti difettosi) seguiamo l'approccio iterativo illustrato nel paragrafo 3.1.
Precisamente, detta X la variabile aleatoria che conta i pezzi difettosi tra i 100 pezzi del campione, si ha:

$$\mathrm{P}(X=0) = e^{-3}$$

$$\boxed{\cdot\frac{3}{1}} \qquad \mathrm{P}(X=1) = \boxed{\frac{3}{1}} \cdot \mathrm{P}(X=0) = 3\,e^{-3}$$

$$\boxed{\cdot\frac{3}{2}} \qquad \mathrm{P}(X=2) = \boxed{\frac{3}{2}} \cdot \mathrm{P}(X=1) = \frac{3}{2} \cdot 3\,e^{-2} = \frac{9}{2}\,e^{-2} \approx 0{,}224$$

In alternativa usando l'espressione analitica della distribuzione di Poisson[39] si ottiene direttamente che la probabilità richiesta è:

$$P(X=2) = \frac{3^2}{2!} \cdot e^{-3} = \frac{9}{2} \cdot e^{-3} \simeq \mathbf{0{,}224}.$$

Osservazioni
- *I modelli a confronto*
I valori ottenuti mediante le due distribuzioni considerate sono "vicini".
Ci si poteva aspettare tale conclusione dato che la distribuzione di Poisson approssima la distribuzione binomiale che ha parametri n, p rispettivamente "grande" e "piccolo". Spesso l'approssimazione è accettabile per $np \leq 10$ e $n > 50$, condizioni che sono appunto verificate nella situazione proposta dal quesito.

- *La formula della binomiale*
Senza ricorrere a formule ritenute a memoria, il valore di probabilità si può ricavare volta per volta in due passi, determinando:

1. la probabilità di una sequenza di 2 pezzi difettosi (D) e 98 non difettosi (N), quale

   $$DDNN...N$$

   per la legge della moltiplicazione (eventi indipendenti) la probabilità di tale sequenza è

   $$0{,}03^2 \cdot (1 - 0{,}03)^{98}$$

2. il numero di tutte le sequenze di 2 D e 98 N
   esso è il numero di sottoinsiemi di 2 elementi (posizioni per D), contenuti in un insieme di 100 elementi (posizioni possibili nella sequenza); dunque è

   $$\binom{100}{2}$$

[39] La distribuzione della variabile aleatoria di Poisson X è data da

$$P(X=k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda} \quad \text{per } k = 0, 1, 2, ...$$

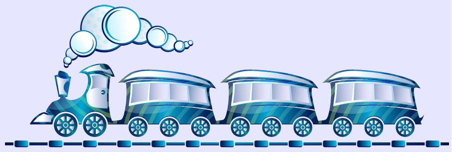
5.2.2 Seconda simulazione ministeriale 2015, Quesito 10

In una stazione ferroviaria, fra le 8 e le 10 del mattino arrivano in media ogni 20 minuti due treni. Determina la probabilità che in 20 minuti :

a) non arrivi nessun treno;

b) ne arrivi uno solo;

c) al massimo ne arrivino 4.



Decidiamo di utilizzare il modello di Poisson per modellizzare la situazione. Affinché ciò abbia senso assumiamo[40] che l'arrivo di un treno sia un evento "casuale" e uniforme[41]. Osserviamo, inoltre, che esso si verifica "raramente", visto che nei 20 minuti considerati arrivano solo 2 treni.

Pertanto proseguiamo indicando con X la variabile aleatoria che conta i treni che arrivano nei 20 minuti in esame.
Il prossimo passo consiste nello stabilire il valore del parametro λ, ossia del numero medio di realizzazioni dell'evento (in questo caso l'arrivo del treno) nell'intervallo considerato (in questo caso 20 minuti). Il testo afferma proprio quanto ci serve: in media, arrivano 2 treni in 20 minuti. Quindi stabiliamo che il parametro λ che caratterizza le distribuzione sia 2.

a) b) Possiamo così calcolare le due probabilità richieste:

$$P(X=0) = e^{-2} \approx 0{,}135$$

$$\boxed{\cdot \frac{2}{1}}$$

$$P(X=1) = \boxed{\frac{2}{1}} \cdot P(X=0) = 2\,e^{-2} \approx 0{,}271$$

c) Conviene interpretare l'evento "*arrivano al massimo 4 treni*" come "*non arriva alcun treno, **oppure** arriva 1 treno, **oppure** arrivano 2 treni, **oppure** arrivano 3 treni, **oppure** arrivano 4 treni*".
Pertanto la probabilità richiesta è:

$$P(X \leq 4) = P(X=0) + P(X=1) + P(X=2) + P(X=3) + P(X=4).$$

[40] Nel testo del quesito tali ipotesi non erano precisate e non veniva indicata, nemmeno, quale distribuzione usare, pertanto la nostra è una soluzione tra le tante possibili. Ad esempio poteva avere senso modellizzare mediante un'opportuna distribuzione binomiale.

[41] Nel senso indicato nel paragrafo 3.1.

E si può calcolare come segue:

$$P(X = 0) = e^{-2}$$

$$P(X = 1) = \boxed{\frac{2}{1}} \cdot P(X = 0) = 2\,e^{-2}$$

$$P(X = 2) = \boxed{\frac{2}{2}} \cdot P(X = 1) = 1 \cdot 2\,e^{-2} = 2\,e^{-2}$$

$$P(X = 3) = \boxed{\frac{2}{3}} \cdot P(X = 2) = \frac{2}{3} \cdot 2\,e^{-2} = \frac{4}{3}\,e^{-2}$$

$$P(X = 4) = \boxed{\frac{2}{4}} \cdot P(X = 3) = \frac{1}{2} \cdot \frac{4}{3}\,e^{-2} = \frac{2}{3}\,e^{-2}$$

da cui:

$$P(X \leq 4) = e^{-2} + 2e^{-2} + 2e^{-2} + \frac{4}{3}e^{-2} + \frac{2}{3}e^{-2} = 7e^{-2} \simeq \mathbf{0{,}947}.$$

5.3 Nascite all'ospedale "Alto Vicentino"

[42]

Su un quotidiano locale leggiamo che nel corso del 2015 all'ospedale "Alto Vicentino" di Santorso (VI), sono nati 1.637 bambini.

1. *È verosimile che in qualche giorno siano nati 10 bambini?*
   E che ne siano nati 20? E più di 20?

Suggerimento
Prova a modellizzare la situazione con la distribuzione di Poisson.
Quale valore di probabilità ottieni per l'evento "*nascono esattamente 10 bambini*"? E per l'evento "*nascono esattamente 20 bambini*"? E per l'evento "*nascono più di 20 bambini*"?

Per sondare la "bontà" del modello probabilistico così costruito, abbiamo richiesto all'azienda "Ulss 4" del Veneto il numero dei nati in tale ospedale in ogni giorno dell'anno.
L'azienda fornisce le tabelle che trovi di seguito.

2. *Il modello costruito ti sembra coerente con i dati osservati?*

Suggerimento
Confronta i tre valori di probabilità che fornisce il modello con le frequenze relative osservate.

Approfondimento
Recupera la distribuzione del numero di nati al giorno nell'ospedale della tua zona; indica con λ il numero medio di nati al giorno e confronta la probabilità dell'evento "nascono più di $2 \cdot \lambda$ bambini nel giorno" con i dati osservati.

[42] I dati numerici dell'esercizio sono stati forniti, per scopi didattici, dall'Ulss 4 del Veneto.

REGIONE VENETO ULSS4 ALTO VICENTINO

Numero di nati al giorno, anno 2015

Gennaio

DATA	# nati
01/01/15	7
02/01/15	9
03/01/15	3
04/01/15	5
05/01/15	5
06/01/15	7
07/01/15	8
08/01/15	5
09/01/15	6
10/01/15	2
11/01/15	9
12/01/15	8
13/01/15	4
14/01/15	2
15/01/15	5
16/01/15	5
17/01/15	1
18/01/15	5
19/01/15	5
20/01/15	3
21/01/15	7
22/01/15	2
23/01/15	7
24/01/15	4
25/01/15	1
26/01/15	6
27/01/15	2
28/01/15	5
29/01/15	8
30/01/15	6
31/01/15	7

Febbraio

DATA	# nati
01/02/15	6
02/02/15	1
03/02/15	6
04/02/15	8
05/02/15	3
06/02/15	4
07/02/15	6
08/02/15	3
09/02/15	4
10/02/15	3
11/02/15	5
12/02/15	4
13/02/15	3
14/02/15	4
15/02/15	9
16/02/15	2
17/02/15	5
18/02/15	6
19/02/15	5
20/02/15	5
21/02/15	2
22/02/15	9
23/02/15	3
24/02/15	3
25/02/15	3
26/02/15	3
27/02/15	3
28/02/15	6

Marzo

DATA	# nati
01/03/15	3
02/03/15	4
03/03/15	2
04/03/15	4
05/03/15	7
06/03/15	6
07/03/15	2
08/03/15	8
09/03/15	5
10/03/15	4
11/03/15	8
12/03/15	8
13/03/15	7
14/03/15	4
15/03/15	2
16/03/15	6
17/03/15	5
18/03/15	7
19/03/15	4
20/03/15	4
21/03/15	7
22/03/15	5
23/03/15	2
24/03/15	3
25/03/15	7
26/03/15	3
27/03/15	4
28/03/15	5
29/03/15	2
30/03/15	7
31/03/15	2

Aprile

DATA	# nati
01/04/15	6
02/04/15	4
03/04/15	7
04/04/15	4
05/04/15	5
06/04/15	4
07/04/15	4
08/04/15	9
09/04/15	3
10/04/15	4
11/04/15	2
12/04/15	4
13/04/15	7
14/04/15	1
15/04/15	4
16/04/15	3
17/04/15	3
18/04/15	2
19/04/15	4
20/04/15	3
21/04/15	6
22/04/15	4
23/04/15	2
24/04/15	3
25/04/15	4
26/04/15	3
27/04/15	4
28/04/15	5
29/04/15	6
30/04/15	3

Maggio

DATA	# nati
01/05/15	3
02/05/15	2
03/05/15	5
04/05/15	3
05/05/15	8
06/05/15	10
07/05/15	2
08/05/15	7
09/05/15	8
10/05/15	5
11/05/15	9
12/05/15	7
13/05/15	3
14/05/15	3
15/05/15	6
16/05/15	0
17/05/15	4
18/05/15	6
19/05/15	2
20/05/15	2
21/05/15	2
22/05/15	4
23/05/15	6
24/05/15	4
25/05/15	7
26/05/15	6
27/05/15	8
28/05/15	1
29/05/15	3
30/05/15	4
31/05/15	2

Giugno

DATA	# nati
01/06/15	2
02/06/15	5
03/06/15	6
04/06/15	1
05/06/15	3
06/06/15	7
07/06/15	4
08/06/15	3
09/06/15	6
10/06/15	4
11/06/15	7
12/06/15	4
13/06/15	4
14/06/15	6
15/06/15	5
16/06/15	4
17/06/15	3
18/06/15	7
19/06/15	4
20/06/15	3
21/06/15	2
22/06/15	7
23/06/15	6
24/06/15	0
25/06/15	5
26/06/15	5
27/06/15	6
28/06/15	4
29/06/15	2
30/06/15	5

REGIONE VENETO ULSS4 ALTO VICENTINO

Numero di nati al giorno, anno 2015

Luglio		Agosto		Settembre		Ottobre		Novembre		Dicembre	
DATA	# nati	DATA	# nati	DATA	# nati	DATA	# nati	DATA	# nati	DATA	# nati
01/07/15	4	01/08/15	4	01/09/15	3	01/10/15	5	01/11/15	5	01/12/15	3
02/07/15	5	02/08/15	4	02/09/15	9	02/10/15	11	02/11/15	2	02/12/15	2
03/07/15	4	03/08/15	4	03/09/15	4	03/10/15	4	03/11/15	7	03/12/15	4
04/07/15	8	04/08/15	3	04/09/15	6	04/10/15	3	04/11/15	5	04/12/15	2
05/07/15	4	05/08/15	5	05/09/15	4	05/10/15	6	05/11/15	1	05/12/15	2
06/07/15	3	06/08/15	2	06/09/15	4	06/10/15	6	06/11/15	2	06/12/15	4
07/07/15	7	07/08/15	10	07/09/15	7	07/10/15	4	07/11/15	2	07/12/15	7
08/07/15	5	08/08/15	4	08/09/15	4	08/10/15	4	08/11/15	4	08/12/15	5
09/07/15	8	09/08/15	2	09/09/15	1	09/10/15	8	09/11/15	4	09/12/15	8
10/07/15	6	10/08/15	7	10/09/15	3	10/10/15	2	10/11/15	4	10/12/15	5
11/07/15	3	11/08/15	3	11/09/15	4	11/10/15	5	11/11/15	4	11/12/15	4
12/07/15	4	12/08/15	6	12/09/15	5	12/10/15	6	12/11/15	4	12/12/15	6
13/07/15	3	13/08/15	4	13/09/15	3	13/10/15	4	13/11/15	4	13/12/15	2
14/07/15	5	14/08/15	5	14/09/15	2	14/10/15	2	14/11/15	2	14/12/15	3
15/07/15	7	15/08/15	8	15/09/15	3	15/10/15	4	15/11/15	6	15/12/15	4
16/07/15	8	16/08/15	5	16/09/15	5	16/10/15	4	16/11/15	7	16/12/15	6
17/07/15	1	17/08/15	3	17/09/15	7	17/10/15	3	17/11/15	7	17/12/15	3
18/07/15	3	18/08/15	6	18/09/15	6	18/10/15	4	18/11/15	3	18/12/15	5
19/07/15	3	19/08/15	8	19/09/15	9	19/10/15	4	19/11/15	4	19/12/15	4
20/07/15	2	20/08/15	7	20/09/15	9	20/10/15	1	20/11/15	4	20/12/15	4
21/07/15	4	21/08/15	10	21/09/15	8	21/10/15	2	21/11/15	6	21/12/15	3
22/07/15	4	22/08/15	1	22/09/15	8	22/10/15	1	22/11/15	4	22/12/15	3
23/07/15	8	23/08/15	5	23/09/15	1	23/10/15	2	23/11/15	2	23/12/15	6
24/07/15	6	24/08/15	4	24/09/15	5	24/10/15	3	24/11/15	3	24/12/15	5
25/07/15	3	25/08/15	0	25/09/15	2	25/10/15	4	25/11/15	3	25/12/15	4
26/07/15	5	26/08/15	4	26/09/15	6	26/10/15	4	26/11/15	4	26/12/15	6
27/07/15	5	27/08/15	8	27/09/15	5	27/10/15	5	27/11/15	3	27/12/15	2
28/07/15	2	28/08/15	6	28/09/15	2	28/10/15	3	28/11/15	7	28/12/15	4
29/07/15	11	29/08/15	5	29/09/15	3	29/10/15	2	29/11/15	3	29/12/15	8
30/07/15	5	30/08/15	5	30/09/15	2	30/10/15	1	30/11/15	5	30/12/15	5
31/07/15	2	31/08/15	5			31/10/15	0			31/12/15	4

1. Per costruire un modello di Poisson opportuno dobbiamo determinare il parametro λ. Esso è il numero medio di nascite al giorno, quindi lo possiamo calcolare dividendo il numero totale di nati per il numero totale di giorni (del 2015). Si ha:

$$\lambda = \frac{1.637}{365} \simeq 4,48$$

A questo punto, utilizzando il modello di Poisson, possiamo calcolare la probabilità che in un giorno si verifichino esattamente 10 nascite e quella che se ne verifichino 20.

Sia X la variabile aleatoria che conta il numero di nascite nel giorno, otteniamo:

$$P(X = 10) = \frac{\lambda^{10}}{10!}e^{-\lambda} \simeq \mathbf{0{,}0102}$$

$$P(X = 20) = \frac{\lambda^{20}}{20!}e^{-\lambda} \simeq \mathbf{5 \cdot 10^{-8}}$$

Per calcolare $P(X > 20)$ conviene considerare la probabilità dell'evento complementare, ossia "*nascono 0 oppure 1 oppure 2 oppure ... oppure 20 bambini*".
Pertanto:

$$P(X > 20) = 1 - (P(X = 0) + P(X = 1) + ...P(X = 20)) \simeq \mathbf{10^{-8}}$$

La probabilità che vi siano 10 nascite nel giorno è circa dell'1%: "piccola", ma non trascurabile. Quindi è ragionevole aspettarsi che tale evento accada effettivamente qualche volta in un anno.
Invece la probabilità che vi siano 20 nascite è minore di 10^{-7}. Dunque l'evento è "estremamente raro".
Analoghe considerazioni valgono per la probabilità che vi siano più di 20 nascite.

Osservazioni
Affinché abbia senso modellizzare con la distribuzione di Poisson è opportuno assumere che la distribuzione delle nascite sia "casuale" e uniforme[43].
Ma non è così scontato lo sia...

2. Dalle tabelle riportate, fornite dalla Ulss 4, possiamo ricavare le frequenze (assolute) osservate e possiamo calcolare[44] le frequenze relative sul periodo di un giorno. Allora abbiamo:

# nati	Frequenza osservata	Frequenze relative
10	3	0,0082
20	0	0
più di 20	0	0

Confrontando i valori delle frequenze relative con quelli delle probabilità calcolate in precedenza possiamo vedere che, relativamente ai tre casi considerati, essi sono "vicini".

[43] Nel senso indicato nel paragrafo 3.1.
[44] Dividendo le frequenze assolute per il numero totale di giorni dell'anno, ossia per 365.

In alternativa[45], per comprendere meglio, possiamo confrontare le frequenze assolute osservate con le frequenze teoriche previste dal modello.
Per frequenza teorica dell'evento intendiamo il prodotto della probabilità dell'evento per il numero di prove, ossia il numero di giorni dell'anno (365) e poi arrotondiamo all'intero.
Otteniamo così:

# nati	Frequenza osservata	Frequenze teorica
10	3	4
20	0	0
più di 20	0	0

Dunque anche le frequenze osservate sono "vicine" ai valori previsti dal modello, nei tre casi analizzati.

Siccome le previsioni del modello sembrano "vicine" alle osservazioni, si può concludere che il nostro modello probabilistico sembra coerente con i dati osservati, almeno per quanto riguarda i tre eventi in esame.

Cosa possiamo dire invece relativamente agli altri numeri di nati al giorno? Sono anch'essi coerenti con il modello?
Ebbene, seguendo un procedimento analogo anche per gli altri numeri di nascite, si ottiene:

# nati	Frequenze osservate	Frequenze relative	Probabilità
0	4	0,0110	0,0113
1	14	0,0384	0,0508
2	48	0,1315	0,1137
3	58	0,1589	0,1698
4	82	0,2247	0,1902
5	53	0,1452	0,1704
6	39	0,1068	0,1273
7	31	0,0849	0,0814
8	22	0,0603	0,0456
9	9	0,0247	0,0227
10	3	0,0082	0,0102
11	2	0,0055	0,0041
12	0	0	0,0016
...	...	...	...
20	0	0	$5 \cdot 10^{-8}$

Pertanto, l'intera distribuzione delle nascite prevista dal modello sembra "vicina" ai dati osservati.

[45] Adotteremo questo approccio anche nell'attività "Un esperimento storico: il decadimento radioattivo" discussa nel paragrafo 6.2.

5.4 Aspetti algebrici

Sappiamo che, per la variabile aleatoria di Poisson X di parametro λ, vale:

$$P(X = k) = \frac{\lambda^k}{k!}e^{-\lambda} \qquad \text{per } k = 0, 1, 2, ...$$

Utilizza tale formula per esprimere $P(X = m)$ in funzione di $P(X = m - 1)$, dove m è un intero maggiore di 1.

Un modo di affrontare la questione è manipolare l'espressione $P(X = m)$ in modo da poter individuare la sua relazione con l'espressione $P(X = m - 1)$.

Ma come conviene manipolare l'espressione $P(X = m)$?
Per comprenderlo iniziamo con l'esplicitare l'espressione $P(X = m - 1)$.

$$P(X = m - 1) = \frac{\lambda^{m-1}}{(m-1)!}e^{-\lambda}$$

A questo punto possiamo intravvedere l'approccio da seguire:

- per ricondursi a $(m - 1)!$ conviene esprimere $m!$ nella forma

$$m! = m(m - 1)!$$

- per ricondursi a λ^{m-1} conviene esprimere λ^m nella forma

$$\lambda^m = \lambda \cdot \lambda^{m-1}.$$

Pertanto concludiamo che vale:

$$P(X = m) = \frac{\lambda^m}{m!} \cdot e^{-\lambda} = \frac{\lambda}{m} \cdot \frac{\lambda^{m-1}}{(m-1)!}e^{-\lambda} = \frac{\lambda}{m}P(X = m - 1)$$

Osservazioni

Abbiamo ottenuto la stessa uguaglianza discussa nel paragrafo 3.1 e dimostrata, per altra via, nel paragrafo 3.2.

Inoltre quello proposto è solo uno dei modi per risolvere la questione.
Ad esempio, dopo le considerazioni iniziali, si poteva pensare di considerare direttamente il rapporto

$$\frac{P(X = m)}{P(X = m - 1)} = \frac{\dfrac{\lambda^m}{m!} \cdot e^{-\lambda}}{\dfrac{\lambda^{m-1}}{(m-1)!} \cdot e^{-\lambda}} = \frac{\dfrac{\lambda}{m} \cdot \dfrac{\lambda^{m-1}}{(m-1)!}e^{-\lambda}}{\dfrac{\lambda^{m-1}}{(m-1)!} \cdot e^{-\lambda}} = \frac{\lambda}{m}$$

Infine, questo ultimo approccio è simile a quello utilizzato nella dimostrazione presentata al paragrafo 3.2, pertanto potrebbe essere utile affrontare questo esercizio prima di analizzare la dimostrazione.

6 Applicazioni

6.1 Bombe su Londra - video

*Nel video "**Bombe su Londra**" viene proposto un episodio storico, accaduto durante la Seconda Guerra Mondiale.*[46] *Potrai saggiare l'efficacia del modello di Poisson nell'affrontare la questione che si ponevano gli inglesi: il bombardamento era mirato?*
Lo schema di ragionamento che seguiamo costituisce un valido riferimento per altre attività del nostro percorso.

Trovi il video al link seguente: *www.youtube.com/watch?v=gfAreJzx5O0&feature=youtu.be*.

[46] Un riferimento sintetico, in forma scritta, dei contenuti affrontati nel video si può trovare anche in seguito, nel paragrafo 6.4.

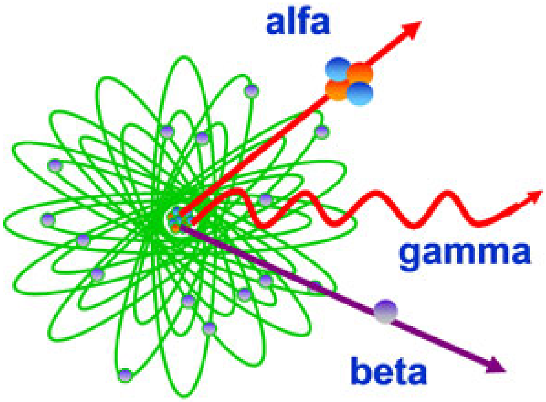
6.2 Un esperimento storico: il decadimento radioattivo - attività

L'intento dell'attività è quello di investigare il fenomeno del decadimento radioattivo mediante il modello di Poisson. Lo faremo seguendo, essenzialmente, l'approccio seguito da Rutherford all'inizio del Novecento.

Leggi il testo proposto e poi rispondi ai quesiti che seguono aiutandoti con un foglio elettronico. Può esserti d'aiuto esaminare prima il video "Bombe su Londra".

I fisici Ernest Rutherford (1871-1937) e Johannes Wilhelm Geiger (1882-1945) assieme ai loro collaboratori James Chadwick (1891-1974) e Charles Drummond Ellis (1895-1980) nei laboratori di Cambridge, studiarono i processi di emissione di particelle da sostanze radioattive.

Essi osservarono che alcune sostanze, come per esempio il radio e l'uranio, cambiano lentamente natura nel tempo. Accade, infatti, che gli atomi si disintegrino ed emettano di conseguenza particelle radioattive che, in seguito, furono classificate in particelle alfa (α), beta (β) e raggi gamma (γ). Ad esempio l'uranio, con un lungo procedimento di decadimento che dura miliardi di anni, si trasforma da uranio-238 (radioattivo) in piombo-206 (non radioattivo).



[47]

Fig.1

Le emissioni di radiazioni, quindi, sono utili segnali dell'avvenuta disintegrazione.
I fisici presentarono i risultati del loro lavoro nel libro "Radiations from Radioactive Substances" (1930). Nel testo viene presentato un esperimento del 1910 in cui essi studiarono le emissioni di un campione di polonio-210. Per farlo si servirono di uno strumento che emetteva un segnale quando captava una particella α proveniente dal campione. Lo strumento, in seguito, prese il nome di contatore di Geiger e viene usato ancora oggi per conteggiare le emissioni radioattive. Gli scienziati divisero il tempo di osservazione totale, pari a circa 5 ore, in 2.608 intervalli di 7,5 secondi ciascuno e contarono il numero di particelle α emesse in ogni intervallo: nel complesso registrarono 10.097 emissioni. Poi elaborarono tali dati, contando il numero di intervalli in cui nessun atomo si era disintegrato, il numero di intervalli in cui era avvenuta una sola disintegrazione, il numero di intervalli in cui ne erano avvenute due, poi tre, quattro e così via.
La distribuzione delle emissioni è descritta dalla seguente tabella.

Numero particelle emesse	Frequenze osservate
0	57
1	203
2	383
3	525
4	532
5	408
6	273

[47] Fonte dell'immagine: www.passionescienza.it/la-radioattivita/.

7	139
8	45
9	27
10 o più	16

6.2.1 Quesiti

1. *Il matematico Harry Bateman propose di modellizzare il fenomeno mediante la distribuzione di Poisson. Stabilisci il valore del parametro della distribuzione e calcola mediante tale distribuzione le probabilità di avere* $k = 0, 1, 2, \dots, 10$ *o più emissioni nell'intervallo di ampiezza 7,5 secondi. Approssima alla quarta cifra decimale.*

2. *Confronta i valori di probabilità forniti dal modello di Poisson con i dati osservati. Allo scopo realizza una tabella e un grafico opportuno.*

3. *Quali ipotesi sul fenomeno è opportuno assumere affinché abbia senso modellizzarlo mediante la distribuzione di Poisson?*

4. *Tramite questo esperimento, Rutherford e Geiger, intendevano sondare l'ipotesi che le emissioni fossero processi "casuali" e uniformi. Quali conclusioni a tale proposito permette di trarre l'esperimento? Argomenta. (Suggerimento: quali assunzioni sono alla base del modello utilizzato? Ripensa al ragionamento seguito per esaminare il bombardamento di Londra durante la Seconda Guerra Mondiale*[a]*.)*

5. *Il modello proposto è analogo a quello utilizzato per rappresentare il bombardamento su Londra nella Seconda Guerra Mondiale. Indica quali elementi dei due modelli sono in corrispondenza.*

[a] Qui si fa riferimento a quanto presentato nel video "Bombe su Londra".

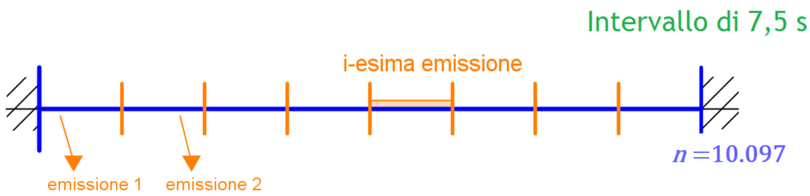
6.2.2 Risoluzione

1. *Il matematico Harry Bateman propose di modellizzare il fenomeno mediante la distribuzione di Poisson. Stabilisci il valore del parametro della distribuzione e calcola mediante tale distribuzione le probabilità di avere* $k = 0, 1, 2, \ldots, 10$ *o più emissioni nell'intervallo di ampiezza 7,5 secondi. Approssima alla quarta cifra decimale.*

   Vogliamo stabilire il parametro λ della distribuzione di Poisson: esso rappresenta il numero medio di emissioni (successi) dell'intero campione di polonio nell'intervallo di 7,5 s. Possiamo calcolarlo, allora, dividendo il numero totale di emissioni osservate per il numero di intervalli temporali:

$$\lambda = \frac{10.097}{2.608} \simeq 3,87.$$

Osservazione
Fissato un intervallo di durata 7,5 s, possiamo schematizzare la situazione come segue:



- si considera una sequenza di $n = 10.097$ prove, una per ogni emissione;
- ciascuna prova può avere due soli esiti: nel fissato intervallo l'emissione si verifica (successo), l'emissione non si verifica in esso;
- la probabilità che una data emissione avvenga nell'intervallo fissato è[48] $p = \dfrac{1}{2.608}$ dato che gli intervalli sono 2.608.

Anche considerando lo schema grafico proposto, osserviamo che si può esprimere λ nella forma:

$$\lambda = np = 10.097 \cdot \frac{1}{2.608} = \frac{10.097}{2.608}.$$

Sia ora X la variabile aleatoria che conta il numero di emissioni nell'intervallo di 7,5 s. Secondo il modello di Poisson la probabilità di avere k emissioni nell'intervallo è[49] $P(X = k) = \dfrac{\lambda^k}{k!} \cdot e^{-\lambda}$ dove $\lambda \simeq 3,87$.

[48] Per determinare tale valore stiamo assumendo l'ipotesi che l'emissione sia "casuale" e uniforme, ma di questo ci occuperemo in dettaglio nel quesito 3.

[49] Per effettuare i calcoli può essere utile utilizzare un foglio Excel. Per calcolare $P(X = k)$ si utilizza la funzione di Excel "=Poisson(k;λ,0)" dove $\lambda \simeq 3,87$.

Allora, al variare di k, otteniamo:

k	P(X=k)
0	0,0209
1	0,0807
2	0,1562
3	0,2015
4	0,1949
5	0,1509
6	0,0973
7	0,0538
8	0,0260
9	0,0112
10 o più	0,0065

2. *Confronta i valori di probabilità forniti dal modello di Poisson con i dati osservati. Allo scopo realizza una tabella e un grafico opportuno.*

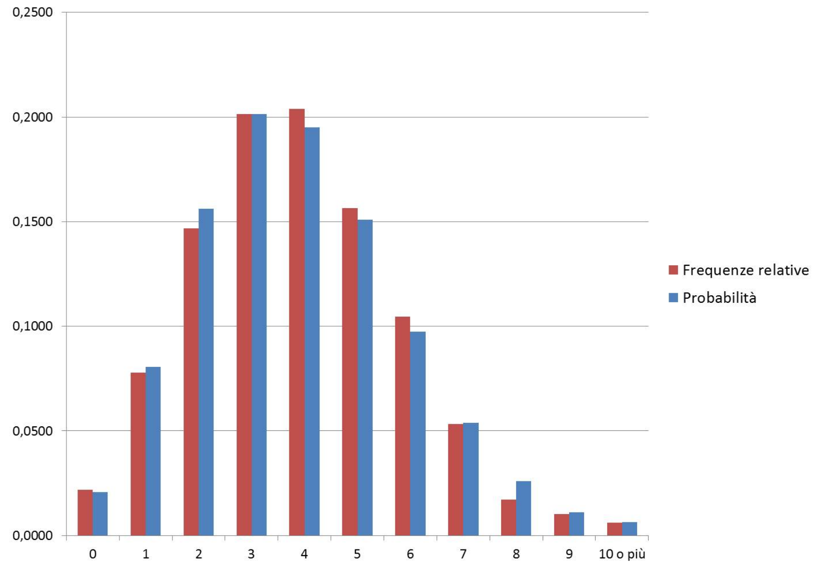
   Vogliamo confrontare le previsioni del modello con i dati osservati. Possiamo procedere in uno dei modi seguenti.

   a) Calcolare le frequenze relative osservate (dividendo ogni frequenza per il numero totale degli intervalli) e confrontarle con le probabilità calcolate.
   b) Confrontare le frequenze osservate con le frequenze teoriche (che possiamo ottenere moltiplicando le probabilità calcolate per il numero totale di intervalli e poi approssimando all'intero più vicino).

   Ecco i risultati delle due elaborazioni.

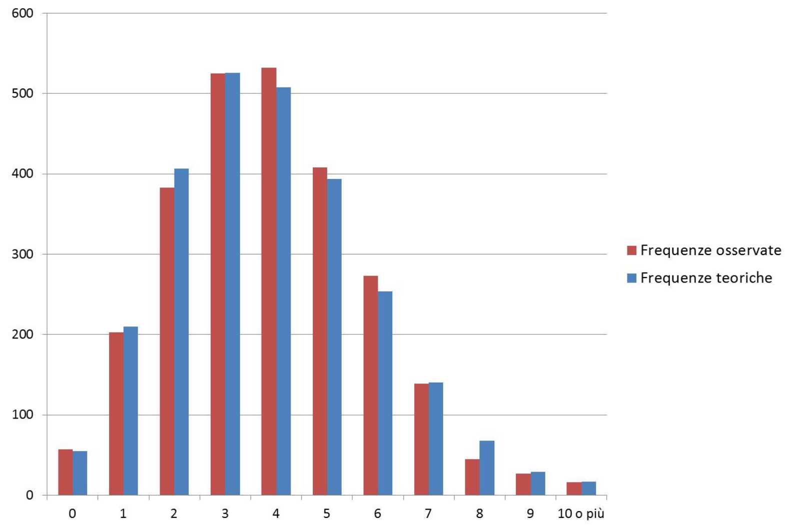
   a) Confronto tra frequenze relative osservate e probabilità.

k	**Frequenze osservate**	**Frequenze relative**	**Probabilità**
0	57	0,0219	0,0209
1	203	0,0778	0,0807
2	383	0,1469	0,1562
3	525	0,2013	0,2015
4	532	0,2040	0,1949
5	408	0,1564	0,1509
6	273	0,1047	0,0973
7	139	0,0533	0,0538
8	45	0,0173	0,0260
9	27	0,0104	0,0112
10 o più	16	0,0061	0,0065



b) Confronto tra frequenze osservate e teoriche.

k	Frequenze osservate	Probabilità	Frequenze teoriche
0	57	0,0209	55
1	203	0,0807	210
2	383	0,1562	407
3	525	0,2015	526
4	532	0,1949	508
5	408	0,1509	394
6	273	0,0973	254
7	139	0,0538	140
8	45	0,0260	68
9	27	0,0112	29
10 o più	16	0,0065	17



Possiamo quindi osservare che le frequenze osservate sono "vicine" a quelle previste dal modello di Poisson. Pertanto il modello di Poisson sembra rappresentare in modo adeguato il fenomeno in esame[50].

3. *Quali ipotesi sul fenomeno è opportuno assumere affinché abbia senso modellizzarlo mediante la distribuzione di Poisson?*

   Le ipotesi che è opportuno assumere per utilizzare il modello di Poisson sono:

   (a) una (ogni) data emissione ha la stessa probabilità di verificarsi in ogni intervallo di 7,5 s;

   (b) le emissioni sono indipendenti, ossia ogni emissione non è influenzata dalle altre[51];

   (c) l'evento è "raro", cioè λ, il numero medio di emissioni nell'intervallo di 7,5 s è "piccolo" rispetto al numero totale di prove.

4. *Tramite questo esperimento, Rutherford e Geiger, intendevano sondare l'ipotesi che le emissioni fossero processi "casuali" e uniformi. Quali conclusioni a tale proposito permette di trarre l'esperimento? Argomenta.*
   (Suggerimento: quali assunzioni sono alla base del modello utilizzato? Ripensa al ragionamento seguito per esaminare il bombardamento di Londra durante la Seconda Guerra Mondiale.)

   Cerchiamo di riassumere ciò che abbiamo fatto finora in questa attività e di riorganizzarlo secondo l'ordine logico che ci è utile.
   In sostanza ciò che abbiamo seguito è lo schema utilizzato nel secolo scorso da Rutherford e Geiger. Ossia:

   (a) abbiamo assunto l'ipotesi che le emissioni siano fenomeni "casuali" e uniformi,

   (b) abbiamo costruito un modello di Poisson per descrivere la situazione il processo,

   (c) abbiamo calcolato le probabilità previste dal modello,

   (d) abbiamo confrontato i dati osservati con quelli forniti dal modello.

   A questo punto, siccome lo scostamento tra i dati osservati e quelli forniti dal modello è "piccolo"[52], possiamo affermare che i dati sono coerenti col modello. Pertanto, come pensavano Rutherford e Giger, è ***plausibile*** che l'emissione radioattiva sia un processo "casuale" e uniforme[53]. Tale conclusione è stata poi confermata da numerosi esperimenti analoghi.

[50] Queste ultime considerazioni sulla "bontà del modello" si possono precisare mediante opportuni strumenti matematici, ma ciò esula dai nostri scopi.

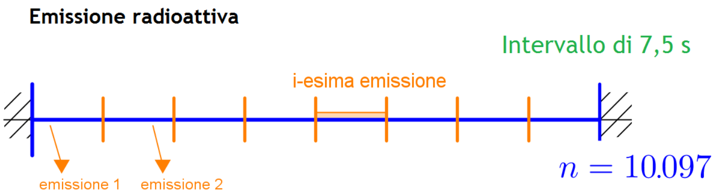
[51] Possiamo esprimere sinteticamente le ipotesi (a) e (b) parlando di emissione "causale" e uniforme, come indicato nel paragrafo 3.1.

[52] Precisamente lo abbiamo osservato, in modo qualitativo, nella domanda 2 tramite i grafici.

[53] Utilizzando il linguaggio statistico, possiamo concludere che tale ipotesi "non viene rifiutata". In altre parole, non si è così dimostrato che il fenomeno del decadimento radioattivo è "causale" e uniforme, ma si è giunti ad una conclusione più debole: cioè che tale ipotesi sia plausibile.

5. *Il modello proposto è analogo a quello utilizzato per rappresentare il bombardamento su Londra nella Seconda Guerra Mondiale. Indica quali elementi dei due modelli sono in corrispondenza.*

La corrispondenza tra i due modelli è suggerita dal confronto tra gli espressivi schemi grafici che abbiamo proposto in precedenza.



Precisamente, la corrispondenza tra i due modelli di Poisson si può così sintetizzare.

	Bombe su Londra	**Emissione radioattiva**
Suddivisione	settore della città (576) bomba (537)	intervallo di tempo (2.608) particella α (10.097)
Evento conteggiato	esplosione bomba sul settore	emissione particella α nell'intervallo

Inoltre:

# medio realizzazioni (λ)	0,93	3,87
# prove (n)	537	10.097
probabilità realizzazione evento (nella prova) (p)	$\dfrac{1}{576}$	$\dfrac{1}{2.608}$

Osserviamo infine che alla base di entrambi i modelli si ha l'ipotesi di "casualità" e uniformità del bombardamento oppure dell'emissione.

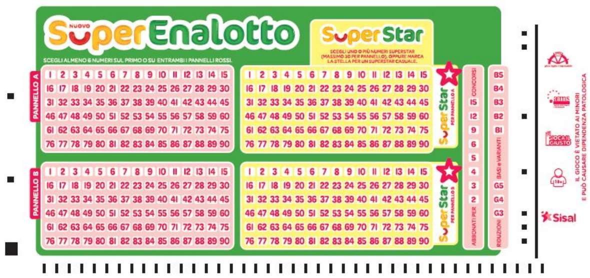
6.3 Prevedere il numero di vincitori al SuperEnalotto - attività

L'intento dell'attività è quello di costruire dei modelli probabilistici che permettano di affermare qualcosa sul numero di vincitori ad un dato concorso del SuperEnalotto.

Leggi il testo proposto e poi rispondi ai quesiti che seguono aiutandoti con un foglio elettronico. Può esserti d'aiuto esaminare prima il video "Bombe su Londra".

Per giocare al SuperEnalotto si devono scegliere 6 numeri interi compresi tra 1 a 90. Si vince se una delle combinazioni giocate contiene 2, 3, 4, 5, o 6 numeri uguali ai 6 numeri estratti[54].

Il gioco è gestito dalla società Sisal sul sito della quale si possono [55] trovare tutte le informazioni utili al riguardo.



In particolare sono indicati i modi (detti categorie) in cui si può vincere: indovinare tutti e sei i numeri estratti (categoria 6), indovinare 5 numeri più il jolly (categoria 5+1)...

Categorie	Probabilità di Vincita	Vincite
6 PIÙ RICCO	1 su 622.614.630	Jackpot milionario
5+1	1 su 103.769.105	311.000€
5	1 su 1.250.230	32.000€
4	1 su 11.907	300€
3 PIÙ RICCO	1 su 327	25€
2 NOVITÀ	1 su 22	5€

Inoltre, per ogni categoria, sono riportate le **probabilità** di vincita e la cifra che viene pagata in caso di vincita.

Il sito non fornisce il numero n di combinazioni giocate. Però esso si può ricavare in modo indiretto osservando che:

- il montepremi di ogni concorso è il 60% delle quote giocate[56];
- il costo di ogni combinazione giocata è 1 euro.

Pertanto n si può ricavare dall'uguaglianza:

$$\text{montepremi concorso} = \frac{60}{100} \cdot n.$$

[54] In realtà ci sono altre combinazioni che portano alla vittoria, per esempio quelle che utilizzano il numero jolly o il superstar, ma per semplicità non ce ne occuperemo in questa attività.

[55] Il sito a cui ci si riferisce è *http : //www.sisal.it/superenalotto* e da esso sono state tratte le immagini di questa scheda.

[56] Dal regolamento del SuperEnalotto (Decreto 109175 dell'Agenzia delle Dogane e dei Monopoli del 15/11/2015) in vigore dal 31 gennaio 2016.

Il montepremi di ogni concorso è riportato sul sito.
Ad esempio, quello relativo al "Concorso 22" del 20/02/2016 è indicato nel rettangolo evidenziato in figura.

Estrazione del 20/2/2016
CONCORSO NUMERO 22

Combinazione vincente: 29 32 33 49 54 85

Jolly: 16

SuperStar: 44

Video estrazione

Montepremi

del concorso	4.471.671,00€
Riporto Jackpot concorso precedente	47.181.750,01€
Attribuzione da D.D. 2011/49938/Giochi/Ena del 16/12/11 art. 2 comma 2	4.058,48€
Montepremi totale del concorso	51.657.479,49€

Quote SuperEnalotto

Categoria	N. Vincite	Euro
Punti 6	0	0,00€
Punti 5+1	0	0,00€
Punti 5	5	37.562,04€
Punti 4	472	402,20€
Punti 3	18854	30,46€
Punti 2	301445	5,93€

Se a questo punto ti stai chiedendo quali siano stati gli esiti del concorso 22, basta consultare il sito (figura a lato).

6.3.1 Quesiti

1. *Consideriamo un concorso del SuperEnalotto. Indichiamo con Y_6 e Y_5 le variabili aleatorie che contano rispettivamente il numero di vincitori nelle categorie "6 punti" e "5 punti". Qual è una distribuzione di probabilità adeguata per tali variabili aleatorie? Spiega le ragioni della tua scelta.*

2. *Considera ora il concorso n.22 del 20/02/2016.*

   a. *Determina il valore dei parametri delle due distribuzioni in esame. (Suggerimento: trova prima il numero n di combinazioni giocate e la probabilità p di vincita qualora si giochi una sola schedina.)*

   b. *Al variare del numero di vincitori k (dove $k = 0, 1, 2, \ldots$) qual è la probabilità di avere k vincitori nella categioria "6"? E nella categoria "5"? Utilizza un foglio elettronico.*

3. *Confronta il modello probabilistico che hai realizzato con il numero effettivo di vincitori al concorso n.22 riportato nella tabella precedente. Il modello ti sembra coerente con le osservazioni?*

4. *Analizziamo, ora, la categoria "4" del concorso e proviamo a modellizzare la situazione mediante la distribuzione di Poisson. Mediante il foglio elettronico determina per quanti valori di k la probabilità che vi siano k vincitori è maggiore di 0,01. E maggiore di 0,02?*

5. *Confronta il modello probabilistico che hai realizzato con il numero effettivo di vincitori al concorso n.22. Cosa osservi?*

6. *Sulla base dei valori di probabilità nelle tabelle, su quale numero di vincitori al prossimo concorso scommetteresti per ciascuna delle tre categorie analizzate? Ritieni che tutti e tre i modelli siano utili per prevedere il numero di vincitori ad un dato concorso del SuperEnalotto? Giustifica.*

Approfondimento

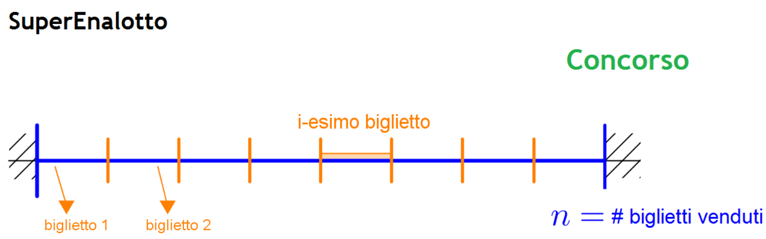
Considera più concorsi (per esempio quelli di un mese specifico), calcola i parametri λ relativi alla vincita nelle tre categorie considerate.
(Attenzione: il valore del parametro λ cambia da concorso a concorso!)
Confronta i valori di probabilità calcolati con il numero effettivo di vincitori. Tale confronto ti suggerisce conclusioni analoghe a quelle a cui eri giunto relativamente alla domanda 6? Ossia tutti e tre i modelli sono utili per la previsione del numero di vincitori in un dato concorso? [57]

[57] Nella risoluzione riportiamo, a titolo di esempio, i dati relativi a tutti i concorsi del febbraio 2016.

6.3.2 Risoluzione

1. *Consideriamo un concorso del SuperEnalotto. Indichiamo con Y_6 e Y_5 le variabili aleatorie che contano rispettivamente il numero di vincitori nelle categorie "6 punti" e "5 punti". Qual è una distribuzione di probabilità adeguata per tali variabili aleatorie? Spiega le ragioni della tua scelta.*

   La distribuzione di probabilità di Poisson è un modello sensato per descrivere la situazione in esame, visto che, almeno in prima approssimazione, possiamo assumere che l'estrazione sia "casuale" e uniforme[58]. E inoltre per un giocatore che punta su una sola combinazione, vincere nelle categorie "5" punti o "6" punti è un evento ragionevolmente "raro"[59].
   La situazione si può schematizzare nel modo seguente.



   Osserviamo che il fenomeno considerato si può modellizzare anche con la distribuzione binomiale. In tal caso però i calcoli sono più articolati[60].

2. *Considera ora il concorso n.22 del 20/02/2016.*

   a. *Determina il valore dei parametri delle due distribuzioni in esame. (Suggerimento: trova prima il numero n di combinazioni giocate e la probabilità p di vincita qualora si giochi una sola schedina.)*

   Categoria "6"
   Per determinare il valore del parametro λ, conviene trovare:
   - il numero n di combinazioni giocate;
   - la probabilità p di vincita, qualora si giochi una sola combinazione.

   Potremmo così ricavare il valore di λ dato che vale $\lambda = np$.

   Ora, nella tabella "categorie/probabilità di vincita" che si può trovare nelle pagine precedenti, leggiamo direttamente

$$p = \frac{1}{622.614.630}$$

   per trovare il numero di giocate, invece, basta ricorrere all'uguaglianza illustrata in precedenza:

$$\text{montepremi} = \frac{60}{100} \cdot n$$

[58] Nel senso indicato nel paragrafo 3.1.
[59] Come risulta dal numero di vincitori nella categoria "5 punti" del concorso n.22 e da quello dei vincitori degli altri concorsi nella stessa categoria.
[60] Come osservato, in generale, nel paragrafo 2.2.

Mentre il valore del montepremi si deduce dalla tabella "montepremi" riportata nelle pagine precedenti da cui si ricava:

$$n = 7.452.785.$$

Concludiamo così che $\lambda = np \simeq$0,0120.
Tale valore rappresenta il numero medio di vincitori della categoria "6".

Categoria "5"
In questo caso la tabella fornisce il valore $p = \dfrac{1}{1.250.230}$.
Visto che il concorso è lo stesso della categoria precedente abbiamo ancora $n = 7.452.785$.
Otteniamo, quindi, che il numero medio di vincitori della categoria "5" è $\lambda = np \simeq 5{,}9611$.

b. *Al variare del numero di vincitori k (dove $k = 0, 1, 2, ...$) qual è la probabilità di avere k vincitori nella categioria "6"? E nella categoria "5"? Utilizza un foglio elettronico.*

Categoria "6"
Utilizzando la distribuzione di Poisson per $\lambda \simeq 0{,}0120$ abbiamo:

k	$P(Y_6=k)$
0	0,9881
1	0,0118
2	0,0001
3	0,0000
4	0,0000
5	0,0000

I valori che seguono nella tabella sono praticamente nulli. La probabilità di avere più di un vincitore è molto piccola.

Categoria "5"
Utilizzando la distribuzione di Poisson per $\lambda \simeq 5{,}9611$ abbiamo:

k	$P(Y_5=k)$
0	0,0026
1	0,0154
2	0,0458
3	0,0910
4	0,1356

k	$P(Y_5=k)$
5	0,1617
6	0,1606
7	0,1368
8	0,1019
9	0,0675

k	$P(Y_5=k)$
10	0,0402
11	0,0218
12	0,0108
13	0,0050
14	0,0021

k	$P(Y_5=k)$
15	0,0008
16	0,0003
17	0,0001
18	0,0000
19	0,0000

3. *Confronta il modello probabilistico che hai realizzato con il numero effettivo di vincitori al concorso n.22 riportato nella tabella precedente. Il modello ti sembra coerente con le osservazioni?*

Nel concorso n.22 vi sono 0 vincitori per la categoria “6” e 5 vincitori per la categoria “5”. Per quanto calcolato nel quesito 2 la probabilità di non avere vincitori nella categoria “6” è circa del 98,8%, quindi il risultato è coerente con la previsione. E la probabilità di avere 5 vincitori nella categoria “5” è del 16,2% ed è il valore di probabilità maggiore tra tutti quelli calcolati. I dati osservati, pertanto, sembrano essere coerenti con quelli forniti dal modello.

4. *Analizziamo, ora, la categoria “4” del concorso e proviamo a modellizzare la situazione mediante la distribuzione di Poisson. Mediante il foglio elettronico determina per quanti valori di k la probabilità che vi siano k vincitori è maggiore di 0,01. E maggiore di 0,02?*

Nella tabella “categoria/vincitori,” leggiamo che per la categoria “4” la probabilità p di vincita, giocando un solo biglietto, è $p = \frac{1}{11.907}$.
Visto che il concorso è lo stesso delle categorie precedenti abbiamo ancora $n = 7.452.785$.
Otteniamo quindi $\lambda = np \simeq 626$. Esso è il numero medio di vincitori della categoria “4”.
Utilizzando la distribuzione di Poisson abbiamo allora[61]

k	$\mathbf{P(Y_4{=}k)}$
0	$1,4724 \cdot 10^{-272}$
...	...
10	$3,7447 \cdot 10^{-251}$
...	...
100	$7,0736 \cdot 10^{-151}$
...	...
500	$2,1864 \cdot 10^{-8}$
...	...
600	$9,4510 \cdot 10^{-3}$
601	$9,8428 \cdot 10^{-3}$
602	$1,0234 \cdot 10^{-2}$
603	$1,0623 \cdot 10^{-2}$
604	$1,1008 \cdot 10^{-2}$
605	$1,1389 \cdot 10^{-2}$
...	...
646	$1,1410 \cdot 10^{-2}$
647	$1,1038 \cdot 10^{-2}$
648	$1,0662 \cdot 10^{-2}$
649	$1,0283 \cdot 10^{-2}$
650	$9,9016 \cdot 10^{-3}$
...	...

Quindi vale $P(Y_4 = k) > 0{,}01$ per tutti i valori di k nell'intervallo $\forall k \in [602, 649]$, ovvero in ben 48 casi.
Non ci sono, invece, valori di k tali che probabilità $P(Y_4 = k) > 0{,}02$.

[61] Indichiamo con Y_4 la variabile aleatoria che conta il numero di vincitori nella categoria “4”.

Possiamo dunque sintetizzare la situazione dicendo che, a differenza delle due categorie “5” e “6”, per la categoria “4” non vi sono valori di probabilità “grandi”, ma ve ne sono “molti” non troppo “piccoli” (come i 48 compresi tra 0,01 e 0,02).

5. *Confronta il modello probabilistico che hai realizzato con il numero effettivo di vincitori al concorso n.22. Cosa osservi?*

   Nel concorso n.22 ci sono 472 vincitori nella categoria “4”, mentre la probabilità teorica di avere $k = 472$ vincitori è $1,88 \cdot 10^{-11}$. Un valore decisamente piccolo. D’altronde, come visto relativamente alla domanda 4, la distribuzione di Y_4 non ha valori di probabilità “grandi”; ad esempio, non vi sono valori di probabilità maggiori di 0,02.

6. *Sulla base dei valori di probabilità nelle tabelle, su quale numero di vincitori al prossimo concorso scommetteresti per ciascuna delle tre categorie analizzate? Ritieni che tutti e tre i modelli siano utili per prevedere il numero di vincitori ad un dato concorso del SuperEnalotto? Giustifica.*

   In base ai valori di probabilità calcolati, ci aspettiamo ragionevolmente che nel prossimo concorso vi siano 0 o 1 vincitori nella categoria “6”, da 3 a 9 (non troppi di meno, non troppi di più) nella categoria “5”. Invece, per quanto osservato nella risposta al quesito 4, non possiamo fare previsioni così precise per la categoria “4”, dato che vi sono “molti” valori di probabilità “non troppo” piccoli.

   Quindi per quanto riguarda le prime due categorie il modello di Poisson è utile per prevedere il numero di vincitori; infatti in questi casi vi sono pochi valori di k che hanno probabilità “non troppo piccola”. Nella categoria “4”, invece, questo non succede; pertanto non è utile modellizzarla mediante la distribuzione di Poisson.

Approfondimento

Considera più concorsi (per esempio quelli di un mese specifico), calcola i parametri λ relativi alla vincita nelle tre categorie considerate. (Attenzione: il valore del parametro λ cambia da concorso a concorso!)
Confronta i valori di probabilità calcolati con il numero effettivo di vincitori. Tale confronto ti suggerisce conclusioni analoghe a quelle a cui eri giunto relativamente alla domanda 6? Ossia tutti e tre i modelli sono utili per la previsione del numero di vincitori in un dato concorso?

Analizziamo, ad esempio, i concorsi del mese di febbraio 2016.

Concorso n.	$\boldsymbol{\lambda_6}$	**Vincite "6"**	$\boldsymbol{\lambda_5}$	**Vincite "5"**	$\boldsymbol{\lambda_4}$	**Vincite "4"**
14	0,0101	0	5,0272	7	527,8591	463
15	0,0097	0	4,8125	2	505,3131	552
16	0,0123	0	6,1259	9	643,2146	515
17	0,0098	0	4,8621	2	510,5178	454
18	0,0097	0	4,8176	6	505,8482	629
19	0,0120	0	5,9942	0	629,3922	456
20	0,0097	0	4,8450	8	508,7186	552
21	0,0096	0	4,7730	3	501,1604	445
22	0,0120	0	5,9611	5	625,9163	472
23	0,0097	0	4,8254	1	506,6631	420
24	0,0095	0	4,7399	4	497,6911	385
25	0,0120	0	5,9545	3	625,2194	589

A questo punto si può ripetere, per ogni singolo concorso, l'intera analisi effettuata per in concorso n.22 secondo lo schema delle domande precedenti.

6.4 Simulazione di eventi "rari" - attività con GeoGebra

Vogliamo investigare ulteriormente la situazione, riguardante il bombardamento di Londra durante la Seconda Guerra Mondiale, discussa nel video "Bombe su Londra".
Ti consigliamo di esaminarlo.
Leggi il testo seguente e utilizza il file GeoGebra Simulazione.ggb per rispondere ai quesiti che seguono. L'uso del file è spiegato in dettaglio nel video "Simulazione bombardamento".

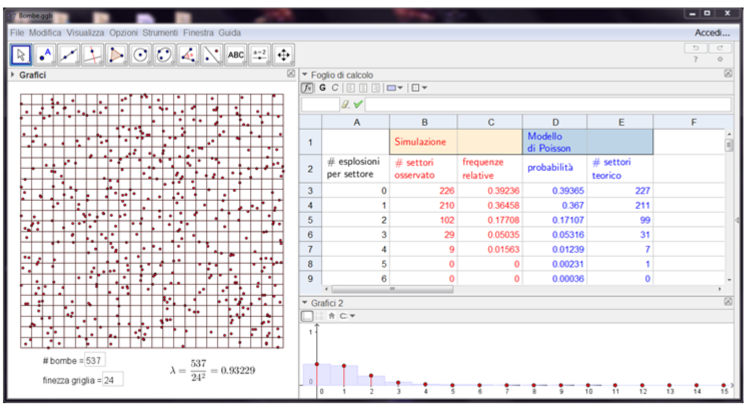
Il modello e la sua storia
Dal giugno del 1944 al marzo del 1945 i tedeschi bombardarono l'Inghilterra con 9.251 bombe volanti V-1 di cui 2.419 raggiunsero Londra. Tra queste 537 caddero nella zona sud di Londra.
Gli analisti inglesi suddivisero in settori l'intera zona interessata, che ha un'estensione di 144 km^2, utilizzando una griglia 24 $\times$ 24 (ossia con 576 quadratini uguali, ciascuno di area pari a 0,25km^2). Poi contarono le bombe esplose in ciascun settore e osservarono che alcuni settori erano stati colpiti anche 4 volte, mentre molti non era mai stati colpiti.
I dati raccolti sono riportati in tabella.

# esplosioni per settore	0	1	2	3	4	5	6	7
frequenza osservata	229	211	93	35	7	0	0	1

Così gli inglesi iniziarono a chiedersi se la tecnologia tedesca fosse tanto avanzata da poter colpire con precisione obiettivi specifici.
L'attuario R.D. Clarke nel 1946 pubblicò[62] i risultati dello studio volto a stabilire se il bombardamento fosse mirato o se le bombe fossero cadute "casualmente".
Utilizzò, in particolare, la distribuzione di Poisson per prevedere le frequenze teoriche di caduta delle bombe e ottenne i dati seguenti.

# esplosioni per settore	0	1	2	3	4	5	6	7
frequenza teorica	227	211	99	31	7	1	0	0

Come utilizzare il file *Simulazione.ggb*
L'utilizzo del file è spiegato in dettaglio nel video "Simulazione bombardamento", che puoi trovare al link www.youtube.com/watch?v=7cLcKSQkMDM&feature=youtu.be.
Ne proponiamo qui una sintesi.



[62] R.D.Clarke, An Application of the Poisson Distribution, Journal of the Institute of Actuaries, vol. 72 (1946), p.481.

Il file *Simulazione.ggb* è stato costruito per permetterti di simulare un bombardamento e confrontare i risultati ottenuti con il modello teorico della distribuzione di Poisson.
In particolare possiamo modificare il numero di bombe e la finezza della griglia.
Per esempio, se vogliamo simulare il lancio di 500 bombe su una griglia 10×10 (100 quadratini) dovremo inserire nelle relative celle i valori dei parametri: "#bombe" = 500; "finezza griglia" = 10. Gli esiti del bombardamento sono rappresentati nel riquadro "Grafici".
Se, invece, vogliamo analizzare la situazione dal punto di vista numerico, allora, possiamo far riferimento al "Foglio di calcolo": in rosso sono riportati i dati riguardanti la simulazione effettuata, mentre nelle adiacenti colonne in blu vengono calcolati i corrispondenti valori teorici previsti dal modello probabilistico di Poisson.
Inoltre la visualizzazione "Grafici2" confronta le frequenze relative osservate nel bombardamento simulato (rappresentate come segmenti rossi) e le frequenze teoriche calcolate col modello (rappresentate come rettangoli blu).

Quesiti

1. *Simula, utilizzando il file "Simulazione.ggb", il bombardamento di Londra del 1944-45 con le V-1: fissa i valori dei parametri come indicato nel testo precedente ed effettua più simulazioni. Confronta i valori osservati nelle simulazioni con quelli previsti dal modello mediante le finestre "Foglio di calcolo" e "Grafico 2" del file. Ti sembra che siano "vicini"?*

2. *Basandoti sugli esiti delle tue prove, spiega perché il bombardamento tedesco poteva essere "casuale" e uniforme.*
   (Suggerimento: segui lo schema di ragionamento adottato nel video "Bombe su Londra" e nell'attività "Un esperimento storico: il decadimento radioattivo", quesito 4.)

Traccia risolutiva

1. La domanda è di tipo qualitativo, pertanto non ammette una risposta univoca. Ci aspettiamo che le simulazioni forniscano dati "vicini" a quelli dei dati previsti dal modello di Poisson.

2. Basta ripercorre il ragionamento seguito nel video *Bombe su Londra.* Esso è riportato in dettaglio anche nelle risposte ai quesiti 4 e 5 dell'attività "Un esperimento storico: il decadimento radioattivo" del paragrafo 6.2.

7 Indici della distribuzione - attività

Leggi la questione e poi segui il procedimento che ti viene proposto per affrontarla. Per farlo dovrai anche rispondere ad una domanda.

7.0.1 La questione

Consideriamo la variabile aleatoria di Poisson X che ha parametro λ. Vogliamo esprimere valore atteso e varianza di X in termini del parametro della distribuzione.

In realtà, dall'intrepretazione di λ come numero medio di realizzazioni dell'evento nell'intervallo, ci aspettiamo che la media della distribuzione sia proprio λ. Proviamo, però, a dedurre il risultato per un'altra via.

7.0.2 Costruzione

L'idea è di ricondursi alla *distribuzione binomiale*.

Infatti:

- conosciamo le espressioni degli indici di tale variabile aleatorie in termini dei parametri[63] n, p;
- la distribuzione di X è il limite delle binomiali per $n \to \infty$, nell'ipotesi np costante[64].

Pertanto ci *aspettiamo* che anche gli indici di X siano il limite per $n \to \infty$ dei corrispondenti indici delle variabili aleatorie binomiali approssimanti.

Valore atteso

Il valore atteso della variabile aleatoria binomiale è np, dunque esso è costantemente uguale a λ per ogni variabile aleatoria binomiale approssimante.
Pertanto il suo limite per $n \to \infty$ è proprio λ.
Dunque ci aspettiamo che anche il valore atteso di X sia $\boldsymbol{\lambda}$[65].

Varianza

Prova ad esprimere la varianza di X in termini del parametro λ.
<u>Suggerimento</u>
Segui il procedimento illustrato per il valore atteso e osserva che la varianza della variabile aleatoria binomiale è $np(1-p)$.

[63] Precisamente essi rappresentano rispettivamente il numero di prove del processo e la probabilità di realizzazione dell'evento (spesso indicato come "successo") nella singola prova.
[64] Nel senso indicato nel paragrafo 3.1.
[65] Si può dimostrare formalmente che la nostra congettura è vera.

7.0.3 Risoluzione

1. Il primo passo consiste nell'esprimere la varianza della variabile aleatoria binomiale approssimante in termini di λ e di n.
   Per farlo sfruttiamo l'ipotesi $\lambda = np$; si ottiene così

$$\boxed{np(1-p) = \boldsymbol{\lambda}\left(1 - \frac{\boldsymbol{\lambda}}{n}\right)} \tag{6}$$

2. Ci aspettiamo che la varianza di X sia il limite per $n \to \infty$ dell'espressione appena trovata. Per determinare tale limite è utile tenere presente che esso va effettuato nell'ipotesi λ costante. In definitiva:

$$\boxed{\lim_{n\to\infty} \boldsymbol{\lambda}\left(1 - \frac{\boldsymbol{\lambda}}{n}\right) = \lambda(1-0) = \boldsymbol{\lambda}} \tag{7}$$

7.0.4 Conclusione

Abbiamo così congetturato che:

> la variabile aleatoria di Poisson di parametro λ ha **valore atteso $\boldsymbol{\lambda}$** e **varianza $\boldsymbol{\lambda}$**.

Si può dimostrare che tale congettura è vera.

Osservazione

In generale il valore atteso di una variabile aleatoria si può interpretare come "centro della distribuzione" e la varianza come "dispersione della distribuzione".

Per la distribuzione di Poisson, tali indici valgono λ, pertanto l'interpretazione geometrica del valore atteso e della varianza ci fornisce anche il significato geometrico del parametro λ: al crescere del parametro λ aumentano anche il punto di massimo e l'apertura del grafico della distribuzione.

E con questo abbiamo giustificato intuitivamente quanto già discusso nel paragrafo 4.2; ciò è visualizzato in modo espressivo nella figura a pagina 21.

Riferimenti bibliografici

Riferimenti relativi alla didattica

Siti

[CProb] *Corso di formazione in didattica della probabilità – DiCoMat Lab in collaborazione con IPRASE*
URL: http://www.iprase.tn.it/formazione/formazione-docenti-e-dirigenti/corsi/didattica-della-probabilita-per-la-secondaria-di-secondo-grado/

[DCMLab] *Laboratorio di Didattica e Comunicazione della Matematica, Dipartimento di Matematica dell'Università degli Studi di Trento*
URL: http://r.unitn.it/it/maths/dicomatlab
Canale YouTube del DiCoMat Lab
URL: https://www.youtube.com/channel/UCcUoK3gxJDNEcUhDrzUuHgw

[IndNaz1] *Indicazioni nazionali per il curricolo del primo ciclo di istruzione*, 2012.
URL: http://www.indicazioninazionali.it/J/

[IndNaz2] *Indicazioni nazionali per il secondo ciclo di istruzione*, 2010.
URL: http://archivio.pubblica.istruzione.it/riforma_superiori/nuovesuperiori/index.html

[UMI-CIIM] UMI-CIIM, *Proposta di un Syllabus di matematica per i Licei Scientifici (nuovo ordinamento)*, 2014.
URL: http://www.umi-ciim.it/materiali-umi-ciim/secondo-ciclo/

[Zan10] *L'errore in matematica: alcune riflessioni*, 2010.
URL: http://www.scuolavalore.indire.it/nuove_risorse/lerrore-in-matematica-alcune-riflessioni/

Testi

[Anz] ANZELLOTTI G. (2011). "*Valutazione e sviluppo delle competenze matematiche di base dall'obbligo scolastico all'ingresso dell'università*", RicercAzione, vol. 3, n. 1.

[AnzCapInn] ANZELLOTTI G., CAPPELLO L., INNOCENTI S. (2005). "*Matematica: obiettivi, itinerari, interpretazioni*", Nuova Secondaria, vol. 23, n. 1, pp. 91-97.

[Bar] BARRA M. (2016). "*Parliamo di probabilità e del suo insegnamento*", Progetto Alice, vol. 17, n. 49, p. 5 e seguenti.

[ConCl] CONSIGLIO DI CLASSE 5BSA (2016). "*Documento del Consiglio di Classe 5BSA*", Liceo Ginnasio Statale "B. G. Brocchi" di Bassano del Grappa (VI).

[DAm] D'AMORE B. (1999). "*Elementi di didattica della matematica*", Pitagora.

[Fre] FREUDENTHAL H. (1994). "*Ripensando l'educazione matematica*", Edizioni La Scuola.

[Lang] LANG S. (1991). "*La bellezza della matematica*", Bollati Boringhieri.

[ParEu] PARLAMENTO EUROPEO (2006). "*Raccomandazione del parlamento europeo del 18 dicembre 2006*".

[Pel] PELLEREY M. (2015). "*Le competenze cosa sono*", L'insegnamento della matematica e delle scienze integrate, vol. 38, n. 5.

[Sfa] SFARD A. (2009). "*Psicologia del pensiero matematico*", Edizioni Erickson.

[Zan07] ZAN R. (2007). "*Difficoltà in matematica. Osservare, interpretare, intervenire*", Edizioni Springer.

Libri di testo in adozione

[BerTriBar] BERGAMINI M., TRIFONE A., BAROZZI G. (2011). "*Matematica.blu.2.0*", Volume 5, Zanichelli.

[BarManFra] BARONCINI P., MANFREDI R., FRAGNI I. (2012). "*Lineamenti.Math.blu*", Volume 5, Ghisetti e Corvi.

[MarPal] MARASCHINI W., PALMA M. (2002). "*MultiForMat*", Probabilità e inferenza statistica, Paravia.

[Sas] SASSO L. (2016). "*La matematica a colori*", Edizione blu per il quinto anno, Petrini.

Riferimenti relativi al calcolo delle probabilità o affini

Siti

[DMar] De Martini G. (2007). "*Un problema di... bombardamento*", Progetto Lauree Scientifiche-Matematica.
URL: http://www.dma.unina.it/laureescientifiche/materiale/II%20anno/modelli%20II%20anno/ModMatSocLez14e22-03-07.pdf

[GriSne] Grinstead C.M., Snell J.L. (1997). "*Introduction to probability*", AMS.
URL: http://www.dartmouth.edu/ chance/teaching_aids/books_articles/probability_book/pdf.html

[LeSci] Le Scienze (2015). "*CNR: Adolescenti d'azzardo: più prevenzione, meno giocatori*", Comunicato Stampa 13 marzo 2015.
URL: http://www.lescienze.it/lanci/2015/03/13/news/cnr_adolescenti_d_azzardo_piu_prevenzione_meno_giocatori-2525314/

Testi

[Bal] Baldi P. (2012). "*Introduzione alla probabilità*", McGraw-Hill, Seconda Edizione.

[Bon] Bonaccorsi S. (2004). "*Appunti di probabilità*", dispensa del corso "Calcolo delle probabilità".

[Bre] Brezis H. (1995). "*Ananlisi Funzionale*", Appendice sull'integrazione astratta di Sbordone C., Liguori Editore.

[CerTom] Cerasoli M., Tomassetti G. (1989). "*La matematica di oggi per domani, Elementi di statistica*", Zanichelli.

[Cla] Clarke R.D. (1946). "*An application of the poisson distribution*", Journal of the Institute of Actuaries, vol. 72, p. 48.

[Cic] Cicchitelli G. (1990). "*Probabilità e statistica*", Maggioli, Prima Edizione.

[DAgl] Dall'Aglio G. (1987). "*Calcolo delle probabilità*", Zanichelli, Prima Edizione.

[Fel] Feller W. (1968). "*An introduction to probability theory and its applications*", Volume 1, Wiley, Terza Edizione.

[Isr] Israel G. (2002). "*Modelli matematici, introduzione alla matematica applicata*", Franco Muzzio Editore.

[PagSal] Pagani C.D., Salsa S. (1990). "*Ananlisi Matematica 1*", Zanichelli, Prima Edizione.

[Pia] Piazza R. (2009). "*I capricci del caso*", Springer.

[Pro] Prodi G. (1992). "*Metodi matematici e statistici*", McGraw-Hill.

[Ros] Rossi C. (1999). "*La matematica dell'incertezza*", Zanichelli.

[RutChaEll] Rutherford E., Chadwick J., Ellis C.D., (2010). "*Radiations from Radioactive Substances*", Cambridge University Press, Seconda Edizione.

[Wol] Wolfson M.M. (2008). "*Everyday Probability and Statistics: Health, Elections, Gambling and War*", World Scientific Publishing Company.

Fonti per le immagini

Siti

[1] URL: https://pixabay.com/it/

[2] URL: https://commons.wikimedia.org/w/

[3] URL: https://it.wikipedia.org/w/

[4] URL: http://www.passionescienza.it/la-radioattivita/

[5] URL: http://www.sisal.it/superenalotto