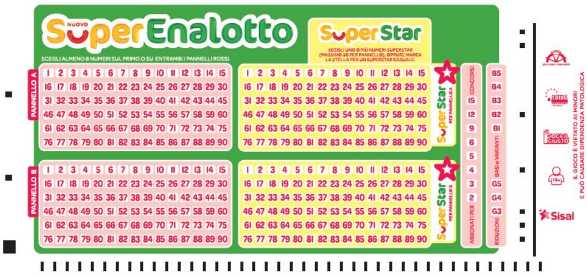
6.3 Prevedere il numero di vincitori al SuperEnalotto - attività

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	$\boxed{4.471.671,00€}$
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 $k = 0, 1, 2, \dots$) 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

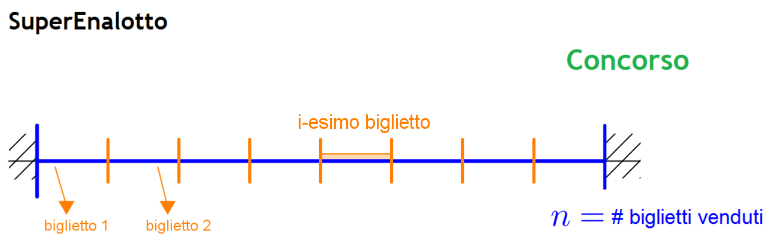
Considera più concorsi (per esempio quelli di un mese specifico), calcola i parametri λ relativi alla vincita nelle tre categorie considerate.
(Attenzione: il valore del parametro λ cambia da concorso a concorso!)
Confronta i valori di probabilità calcolati con il numero effettivo di vincitori. Tale 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 $p = \dfrac{1}{11.907}$.
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.