Alcune questioni sul problema degli accoppiamenti

Supponiamo che N uomini ($N \geqslant 2$) mettano il proprio cappello in una cesta e che poi, ciascuno di loro, ne scelga uno a caso. Ci sono diverse questioni interessanti riguardo questo problema. Vediamone alcune.

1. Qual è la probabilità che nessuno degli uomini scelga il proprio cappello?

E' più semplice calcolare prima la probabilità che almeno un uomo scelga il proprio cappello. Sia E_i ($i = 1, 2, \ldots, N$) l'evento che l' i-esimo uomo scelga il suo cappello. L'evento che almeno un uomo scelga il proprio cappello è dato da

$$\bigcup_{i=1}^{N} E_i$$

e, per il principio di inclusione-esclusione, la sua probabilità è

$$\mathbb{P}\left(\bigcup_{i=1}^{N} E_i\right) = \sum_{i=1}^{N} \mathbb{P}(E_i) - \sum_{i_1<i_2}^{1,N} \mathbb{P}(E_{i_1} \cap E_{i_2}) + \sum_{i_1<i_2<i_3}^{1,N} \mathbb{P}(E_{i_1} \cap E_{i_2} \cap E_{i_3}) + \cdots + (-1)^{N+1}\mathbb{P}(E_1 \cap E_2 \cap \cdots \cap E_N)\,. \tag{1}$$

Per quanto riguarda il primo termine, essendo ovviamente

$$\mathbb{P}(E_i) = \frac{1}{N} \quad (i = 1, 2, \ldots, N),$$

si ha

$$\sum_{i=1}^{N} \mathbb{P}(E_i) = N \cdot \frac{1}{N} = 1\,.$$

Calcoliamoci ora $\mathbb{P}(E_{i_1} \cap E_{i_2})$, ossia la probabilità che (almeno) due uomini abbiano scelto il proprio cappello. Per semplicità supponiamo che i primi due abbiano pescato il proprio cappello. Come possiamo rappresentare questo evento ?

In generale, possiamo rappresentare il risultato delle scelte di tutti gli uomini come un vettore a N componenti intere, nel quale la componente i-sima indica il proprietario del cappello scelto dall'i-simo uomo (ed è quindi un intero compreso tra 1 ed N). Per esempio, il vettore $(1, 2, \ldots, N)$ indica l'evento nel quale tutti hanno scelto il proprio cappello.

L'evento in cui i primi due uomini hanno scelto il proprio cappello si rappresenta quindi come l'unione dei vettori $(1, 2, i_3, \ldots, i_N)$, con gli i_j tutti distinti tra di loro (non si può scegliere un cappello che è stato scelto da un'altro) e tali che $3 \leqslant i_j \leqslant N$, $j = 3, \ldots, N$. Quanti vettori di questo tipo ci sono ? Evidentemente ce ne sono tanti quante sono le permutazioni della $(N-2)$-pla $(3, \ldots, N)$ e dunque ci sono $(N-2)!$ vettori del tipo $(1, 2, i_3, \ldots, i_N)$. Essendoci in tutto $N!$ eventi possibili, abbiamo

$$\mathbb{P}(E_1 \cap E_2) = \frac{(N-2)!}{N!}\,.$$

Visto che ci sono $\binom{n}{2}$ modi diversi di scegliere la coppia di persone che hanno preso il proprio cappello e che $\mathbb{P}(E_{i_1} \cap E_{i_2}) = \mathbb{P}(E_1 \cap E_2)$ per ogni scelta di (i_1, i_2) con $i_1 \neq i_2$, si trova

$$\sum_{i_1<i_2}^{1,N} \mathbb{P}\left(E_{i_1} \cap E_{i_2}\right) = \binom{N}{2} \frac{(N-2)!}{N!} = \frac{\cancel{N!}}{2!\,\cancel{(N-2)!}} \frac{\cancel{(N-2)!}}{\cancel{N!}} = \frac{1}{2}\,.$$

In maniera del tutto analoga si trova

$$\mathbb{P}(E_1 \cap \cdots \cap E_m) = \frac{(N-m)!}{N!}$$

e

$$\sum_{i_1<i_2\cdots<i_m}^{1,N} \mathbb{P}\left(E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_m}\right) = \binom{N}{m} \frac{(N-m)!}{N!} = \frac{\cancel{N!}}{m!\,\cancel{(N-m)!}} \frac{\cancel{(N-m)!}}{\cancel{N!}} = \frac{1}{m!}\,.$$

Ricordando la (1)

$$\mathbb{P}\left(\bigcup_{i=1}^{N} E_i\right) = 1 - \frac{1}{2!} + \frac{1}{3!} - \cdots + (-1)^{N+1} \frac{1}{N!}\,.$$

Essendo l'evento che nessuno degli uomini scelga il proprio cappello il complementare di $\bigcup_{i=1}^{N} E_i$, la sua probabilità è data da

$$1 - \left(1 - \frac{1}{2!} + \frac{1}{3!} - \cdots + (-1)^{N+1} \frac{1}{N!}\right) = \frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{(-1)^N}{N!}\,.$$

E' interessante osservare che questa probabilità, per $N \to +\infty$, tende a e^{-1}, dato che

$$e^{-1} = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!}$$

Questo vuol dire che, essendo $e^{-1} \approx 0.36788$, la probabilità che nessun uomo scelga il proprio cappello quando N è grande è approssimativamente uguale a 0.37 (e non 1, come forse ci si potrebbe aspettare).

2. Qual'è la probabilità che esattamente k persone prendano il proprio cappello ?

Per rispondere a questa domanda, cominciamo col fissare k persone, immaginando che loro abbiano scelto il proprio cappello. Per fissare le idee supponiamo che siano i primi k. E' ovvio che se $k = N$ la risposta è semplicemente $1/N!$. Anche se $k = N - 1$ la risposta è ovvia, visto che anche l'ultima persona rimasta sceglierà l'unico cappello rimasto, che è il suo. Non avverrà quindi mai che esattamente $N - 1$ persone scelgano il proprio cappello.

Sia quindi $k \leqslant N - 2$. L'evento in cui i primi k uomini hanno scelto il proprio cappello si rappresenta come l'unione dei vettori $(1, \ldots, k, i_{k+1}, \ldots i_N)$, con gli i_j tutti distinti tra di loro e tali che $k + 1 \leqslant i_j \leqslant N$, $j = k + 1, \ldots, N$. Se poi vogliamo che non ci siano altre persone che hanno scelto il proprio cappello al di fuori delle prime k dobbiamo anche richiedere

$i_j \neq j,\ j = k+1, \dots, N$. Dobbiamo quindi contare quante k-ple abbiamo del tipo seguente, supponendo $k \leqslant N-2$:

(2) $$(1, \dots, k, i_{k+1}, \dots i_N), \quad k+1 \leqslant i_{k+1}, \dots, i_N \leqslant N,\ i_j \neq j,\ j = k+1, \dots, N.$$

In base al risultato precedente, la probabilità che in un gruppo di $(N-k)$ persone nessuna scelga il proprio cappello è data da

(3) $$\frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{(-1)^{N-k}}{(N-k)!}\,.$$

Questo implica che il numero di k-ple del tipo (2) è dato da

$$(N-k)!\left(\frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{(-1)^{N-k}}{(N-k)!}\right),$$

visto che la probabilità (3) è uguale al numero di k-ple del tipo (2) diviso per $(N-k)!$. Si noti che supponendo $k \leqslant N-2$, abbiamo $N-k \geqslant 2$.

Dato che le k persone che hanno scelto il proprio cappello possono essere scelte in $\binom{N}{k}$ modi diversi, troviamo che la probabilità cercata è

$$\frac{1}{N!}\binom{N}{k}(N-k)!\left(\frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{(-1)^{N-k}}{(N-k)!}\right) = \frac{1}{k!}\left(\frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{(-1)^{N-k}}{(N-k)!}\right).$$

Riassumendo, se indichiamo con q_k la probabilità cercata, abbiamo

(4) $$q_k = \begin{cases} \dfrac{1}{k!}\left(\dfrac{1}{2!} - \dfrac{1}{3!} + \cdots + \dfrac{(-1)^{N-k}}{(N-k)!}\right) & \text{se } 0 \leqslant k \leqslant N-2; \\ 0 & \text{se } k = N-1; \\ \dfrac{1}{N!} & \text{se } k = N. \end{cases}$$

Possiamo anche scrivere, più compattamente,

$$q_k = \frac{1}{k!}\sum_{j=0}^{N-k}\frac{(-1)^j}{j!}\,, \qquad k = 0, 1, \dots, N.$$

3. Qual'è il valore atteso del numero di persone che prendono il proprio cappello ?

Indichiamo con X il numero di persone che prendono il proprio cappello. Se definiamo X_i la v. a. uguale a 1 se l'uomo i-simo ha preso il proprio cappello e uguale a 0 se ha preso quello di un altro, avremo ovviamente

$$X = \sum_{j=1}^{N} X_j\,.$$

Essendo la speranza lineare possiamo scrivere

$$E[X] = \sum_{j=1}^{N} E[X_j]\,.$$

Ma essendo X_j una variabile di Bernouilli, risulta

$$E[X_j] = \mathbb{P}(X_j = 1) = \frac{1}{N}\,,$$

visto che la probabilità a priori che una persona prenda il suo cappello è uguale a $1/N$, essendoci in tutto N cappelli [(1)]. Abbiamo dunque

$$E[X] = \sum_{j=1}^{N} E[X_j] = N \cdot \frac{1}{N} = 1.$$

Se questo risultato un po' sorprende, possiamo andare a vedere cosa succede in qualche caso particolare. Consideriamo $N = 3$. Gli eventi possibili, con i relativi valori di X, sono contenuti nella seguente tabella:

cappelli	X
1 2 3	3
1 3 2	1
2 3 1	0
2 1 3	1
3 1 2	0
3 2 1	1

Ricordando che il valore atteso si può calcolare anche con la formula

$$E[X] = \sum_{\omega \in \Omega} X(\omega)\, \mathbb{P}(\omega)\,,$$

dalla tabella deduciamo subito

$$E[X] = \frac{3+1+0+1+0+1}{6} = 1\,. \tag{5}$$

Dalla stessa tabella troviamo anche

$$\mathbb{P}(X=0) = \frac{1}{3}\,, \quad \mathbb{P}(X=1) = \frac{1}{2}\,, \quad \mathbb{P}(X=2) = 0, \quad \mathbb{P}(X=3) = \frac{1}{6}\,, \tag{6}$$

da cui ricaviamo

$$E[X] = 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{2} + 2 \cdot 0 + 3 \cdot \frac{1}{6} = 1\,. \tag{7}$$

Sia la (5) che la (7) illustrano perché la media sia uguale a 1.

Qui sotto abbiamo l'analoga tabella relativa al caso $N = 4$.

[(1)] Notiamo che le variabili X_j non sono indipendenti tra di loro. Basta pensare che se la persona j-sima prende il suo cappello (e quindi certamente non quello dell'i-simo), la probabilità che l'i-simo prenda il suo diventa $1/(N-1)$. In formule: se $i \neq j$, $\mathbb{P}(X_i = 1 \mid X_j = 1) = 1/(N-1)$ e quest'ultima è diversa da $\mathbb{P}(X_i = 1) = 1/N$.

cappelli	X	cappelli	X	cappelli	X	cappelli	X
4 1 2 3	0	1 4 2 3	1	1 2 4 3	2	1 2 3 4	4
4 1 3 2	1	1 4 3 2	2	1 3 4 2	1	1 3 2 4	2
4 2 3 1	2	2 4 3 1	1	2 3 4 1	0	2 3 1 4	1
4 2 1 3	1	2 4 1 3	0	2 1 4 3	0	2 1 3 4	2
4 3 1 2	0	3 4 1 2	0	3 1 4 2	0	3 1 2 4	1
4 3 2 1	0	3 4 2 1	0	3 2 4 1	1	3 2 1 4	2

Osserviamo che

$$\mathbb{P}(X=0)=\frac{3}{8}\,,\ \mathbb{P}(X=1)=\frac{1}{3}\,,\ \mathbb{P}(X=2)=\frac{1}{4},\ \mathbb{P}(X=3)=0\,,\ \mathbb{P}(X=4)=\frac{1}{24}\,.$$

Ovviamente questi valori, così come quelli in (6), coincidono con quelli dati dalla formula generale (4) per $N=3,4$.

Ripetendo i ragionamenti fatti per $N=3$, troviamo $E[X]=1$ anche in questo caso.

4. Qual'è la varianza della v.a. X introdotta nel numero precedente ?

Con le notazioni introdotte nel punto precedente, possiamo scrivere

$$\text{Var}\,(X)=\sum_{j=1}^{N}\text{Var}\,(X_j)+2\sum_{i<j}^{1,N}\text{Cov}\,(X_i,X_j).$$

Si noti che, pur essendo somma di v.a. di Bernouilli, la X non risulta essere una binomiale, dato che le X_j non sono indipendenti tra di loro

Essendo X_j delle v.a. di Bernouilli, abbiamo

$$\text{Var}\,(X_j)=\frac{1}{N}\left(1-\frac{1}{N}\right)=\frac{N-1}{N^2}\,.$$

Per quanto riguarda le covarianze, intanto osserviamo che

$$\text{Cov}\,(X_i,X_j)=E[X_iX_j]-E[X_i]\,E[X_j].$$

Poiché X_iX_j - assumendo solo i valori 0 ed 1 - risulta essere una Bernoulli, possiamo scrivere

$$E[X_iX_j]=\mathbb{P}(X_iX_j=1).$$

Ovviamente $X_iX_j=1$ se e solo se $X_i=X_j=1$ e dunque

$$\mathbb{P}(X_iX_j=1)=\mathbb{P}(X_i=1,X_j=1).$$

Quest'ultima probabilità (per $i\neq j$) è data da

$$\mathbb{P}(X_i=1\mid X_j=1)\,\mathbb{P}(X_j=1)=\frac{1}{N-1}\,\frac{1}{N}$$

(cfr. la nota (1)), da cui ricaviamo

$$\text{Cov}\,(X_i,X_j)=\frac{1}{N(N-1)}-\frac{1}{N^2}=\frac{1}{N}\left(\frac{1}{N-1}-\frac{1}{N}\right)=\frac{1}{N}\,\frac{\cancel{N}-\cancel{N}+1}{N(N-1)}=\frac{1}{N^2(N-1)}.$$

In definitiva, tenendo presente che ci sono $\binom{N}{2}$ coppie (i,j) con $1 \leqslant i < j \leqslant N$, otteniamo

$$\mathrm{Var}\,(X) = N\,\frac{N-1}{N^2} + 2\binom{N}{2}\frac{1}{N^2(N-1)} = \frac{N-1}{N} + \cancel{2}\,\frac{N\cancel{(N-1)}}{\cancel{2}}\,\frac{1}{N^2\cancel{(N-1)}} = \frac{N-1}{N} + \frac{1}{N} = 1\,.$$

Una giustificazione "visiva" del perché la varianza sia uguale a 1 si può avere guardando il seguente grafico, nel quale è rappresentata la densità di probabilità della X per $N = 10$. Come si vede, i valori più probabili si hanno per $k = 0, 1, 2$ mentre la probabilità di quelli successivi decresce molto rapidamente. Considerando che la media è 1, si capisce perché la varianza sia uguale a 1: il grafico mostra, infatti, che è improbabile che ci discostiamo a una distanza maggiore di 1 dalla media .