

1 Quanto devi aspettare per fare tombola

Assumi di essere l'unico giocatore, e che hai una sola cartella. Trova: (a) la probabilità di fare tombola entro le prime k estrazioni, e (b) il valore atteso delle estrazioni necessarie per fare tombola.

(a) è facile: nelle prime k estrazioni ci devono essere i tuoi 15 numeri. Quindi ipergeometrica con due tipi: uno i tuoi numeri (15 elementi), l'altro gli altri (75 elementi). La probabilità cercata è dunque

$$\binom{75}{k-15} / \binom{90}{k}$$

Per il punto successivo conviene avere k solo al numeratore:

$$\binom{75}{k-15} \cdot \binom{90}{k}^{-1} = \frac{75!}{(k-15)!(90-k)!} \cdot \frac{k!(90-k)!}{90!} \cdot \frac{15!}{15!} = \binom{90}{15}^{-1} \binom{k}{15}.$$

Il punto (b) è più impegnativo. Innanzitutto conviene conoscere un modo utile per calcolare il valore atteso di variabili a valori interi non negativi.

Lemma. *Sia X una variabile a valori interi non negativi. Allora*

$$EX = \sum_{k \geq 1} P(X \geq k) = \sum_{k \geq 0} P(X > k).$$

Dimostrazione (incompleta). L'uguaglianza è chiara perché X ha valori interi. L'osservazione furba è che $X(\omega) = \sum_{k \geq 1} \mathbf{1}_{\{X \geq k\}}(\omega)$ - di facile verifica. Sappiamo che $E\mathbf{1}_{\{X \geq k\}} = P(X \geq k)$. Quello che *non* sappiamo, e che dobbiamo accettare, è che in questo caso possiamo portare la E dentro la serie. Così otteniamo

$$EX = E \sum_{k \geq 1} \mathbf{1}_{\{X \geq k\}} = \sum_{k \geq 1} E\mathbf{1}_{\{X \geq k\}} = \sum_{k \geq 1} P(X \geq k),$$

che è la tesi. □

L'altra cosa che conviene ripassare è che

$$\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1} \quad \text{da cui} \quad \binom{n}{k} = \binom{n+1}{k+1} - \binom{n}{k+1}.$$

Dovresti già saperlo, comunque: le combinazioni $\binom{n+1}{k+1}$ si possono dividere in quelle che contengono un certo elemento - che sono $\binom{n}{k}$ - e quelle che non lo contengono - che sono $\binom{n}{k+1}$.

A questo punto indica con T il numero di estrazioni necessarie per fare tombola - che può prendere gli interi da 15 a 89. Nel punto (a) abbiamo visto che $P(T \leq k) = \binom{90}{15}^{-1} \binom{k}{15}$, e usando il lemma di sopra (nella forma $EX = \sum_{k \geq 0} P(X > k)$) otteniamo

$$\begin{aligned} ET &= \sum_{k=0}^{89} P(T > k) = \sum_{k=0}^{89} [1 - P(T \leq k)] \\ &= 90 - \sum_{k=15}^{89} P(T \leq k) = 90 - \binom{90}{15}^{-1} \sum_{k=15}^{89} \binom{k}{15} \\ &= 90 - \binom{90}{15}^{-1} \binom{90}{16} \end{aligned}$$

dove abbiamo usato

$$\begin{aligned} \sum_{k=15}^{89} \binom{k}{15} &= 1 + \left[\binom{17}{16} - \binom{16}{16} \right] + \left[\binom{18}{16} - \binom{17}{16} \right] + \cdots + \left[\binom{90}{16} - \binom{89}{16} \right] \\ &= 1 - \binom{16}{16} + \binom{90}{16} = \binom{90}{16}. \end{aligned}$$

Per curiosità, questo valore atteso è molto alto:

$$90 - \binom{90}{15}^{-1} \binom{90}{16} = 91 \cdot \frac{15}{16} \approx 85.3.$$

Osservazione. Non abbiamo parlato di $(\Omega\mathcal{F}, P)$. Lo facciamo qui per i più curiosi. Qui possiamo prendere Ω uguale all'insieme delle permutazioni degli interi da 1 a 90 - quindi $\omega = (\omega_1, \dots, \omega_{90})$. Poiché Ω è finito $\mathcal{F}$ è l'insieme delle parti di Ω ; e P sia uniforme su Ω .

Definisci $Y_k(\omega) = \{\omega_i : 1 \leq i \leq k\}$, che indica i primi k numeri estratti. Indicando con C l'insieme dei numeri nella tua cartella, il numero degli elementi di C usciti nelle prime k estrazioni è $Z_k(\omega) = \# \{C \cap Y_k(\omega)\}$. Osserva che nel punto (a) abbiamo trovato $P(Z_k = 15) = \binom{90}{15}^{-1} \binom{k}{15}$.

Il numero di estrazioni necessario per fare tombola è allora

$$T(\omega) = \min\{n : Z_n(\omega) = 15\}.$$

Per concludere osserva che $\{T \leq k\} = \{Z_k = 15\}$.

2 Quanto ci vuole per riempire un album di figurine

Il problema lo formuliamo così: in un'urna ci sono palline di m colori diversi, in proporzioni uguali. Dunque gli m colori sono le figurine che compongono l'album, e nota che tutte le figurine hanno la stessa probabilità $1/m$ (non ci sono figurine “difficili” come era nella realtà quando io ero piccolo). E non c'è un mercato delle figurine (invece noi ragazzini passavamo pomeriggi interi a scambiare figurine): ogni giorno estrai una pallina/figurina dall'urna; se ce l'hai la rimetti nell'urna; se non ce l'hai la metti al suo posto vuoto.¹

La domanda è quante estrazioni ci vogliono in media per completare l'album (se ogni estrazione costa 1 la domanda equivale a quanti soldi ci vogliono).

Per rispondere immagina di fare un punto ogni volta che trovi una figurina che non hai. Devi arrivare ad m punti, e definiremo m variabili $(U_k)_{1 \leq k \leq m}$ dove U_k rappresenta il numero di estrazioni necessarie per passare da $k-1$ a k punti (cioè per collocare la k -esima figurina). Ponendo $T_k = U_1 + \dots + U_k$, il problema è trovare $E(T_m)$.

Ovviamente $U_1 = 1$ - perché l'album è vuoto. Dopo che hai messo la prima figurina, diciamo $M_1 \in \{1, \dots, m\}$ ricominci a estrarre finché non ne trovi una diversa da M_1 . Quante estrazioni ci vogliono? La probabilità di trovarla è in ogni estrazione $(m-1)/m$ (basta che non la trovi del colore M_1), e le estrazioni sono indipendenti; quindi il numero di estrazioni fino al successo - cioè U_2 - è una v.a. con distribuzione geometrica² di parametro $p = (m-1)/m$. E U_3 ? Il processo delle estrazioni successive è analogo al precedente, solo che la probabilità di successo a questo punto è $(m-2)/m$. È chiaro a questo punto che U_{k+1} è una geometrica di parametro $p = (m-k)/m$. E sappiamo che la sua media è $1/p = m/(m-k)$. In conclusione la risposta è

$$E(U_1 + \dots + U_m) = 1 + \frac{m}{m-1} + \dots + \frac{m}{m-(m-1)} = m \sum_{k=1}^m \frac{1}{k}.$$

Per curiosità, le somme dei reciproci degli interi positivi si chiamano “Harmonic numbers”, e per m grande si può usare l'approssimazione $\sum_{k=1}^m \frac{1}{k} \approx 0.58 + \ln m$. Quindi per esempio per un album di 500 figurine ci vogliono in media $500 \cdot (0.58 + \ln 500) \approx 3400$ estrazioni.

¹Immaginando un numero grande di figurine esistenti nella fabbrica dell'album, rimetterla nell'urna equivale a buttarla.

²Nella terminologia di Baldi “geometrica modificata”.