Esercitazioni di Matematica e Statistica, Anno Accademico 2008-2009,

Scienze Biologiche e Molecolari C

V.M.Tortorelli

schema XV esercitazione, 3 Marzo 2009

ESERCIZI ED ARGOMENTI SVOLTI

1. Schema della soluzione del "LOTTO FRANCESE": su una tabella a sette per sette sono scritti i primi quarantanove numeri. Presa a caso un'altra tabella simile con che probabilità scegliendo su questa sei caselle nessun numero di quelli scelti occupi la stessa posizione nella prima tabella? E con quali probabilità rispettivamente si trova un numero , due numeri etc, che occupano la stessa posizione nelle due tabelle?

Vengono date una permutazione casuale p sui primi 49 numeri, un'altra permutazione casuale q sui primi 49 numeri, ed un sottoinsieme casuale I con 6 elementi dei primi 49 numeri.

Si sottointende che i tre eventi casuali siano indipendenti e uniformemente distribuiti.

Si chiede di calcolare $p_{49,6,k} =: \mathbf{P}(\#\{i \in I : p_i = q_i\} = k)$, $0 \leq k \leq 6$,
ovvero come è distribuita la variabile aleatoria, definita tramite gli accadimenti casuali in questione:
$N = \#\{i \in I : p_i = q_i\}$ a valori in $\{0, 1, 2, 3, 4, 5, 6\}$.

Prima ancora di studiare lo spazio degli eventi nella prima formalizzazione conviene introdurre qualche semplificazione.

Prima semplificazione: invece di scegliere le due tabelle (permutazioni) e le sei caselle (il sottoinsieme), il problema è equivalente a fare un'estrazione senza rimpiazzo di 6 elementi tra i primi 49 numeri, annotarsi il loro ordine, quindi rimettere tutto dentro e procedere analogamente con un'altra estrazione. La variabile N sarà il numero di elementi che nelle due estrazioni sono stati presi nello stesso ordine.

Le due estrazioni casuali indipendenti possono essere viste come due disposizioni casuali senza ripetizione di $\{1, \ldots 6\}$ in $\{1, \ldots, 49\}$, e distribuite uniformemente.
Chiamiamole d e δ.

Si ha quindi $p_{49,6,k} =: \mathbf{P}(\#\{1 \leq j : d_j = \delta_j\} = k)$

Seconda semplificazione: una delle due disposizioni può essere considerata non casuale ma una arbitraria distibuzione costante, e.g. $\delta_1 = 1, \ldots \delta_6 = 6$.

Infatti date due disposizioni A e B di 6 in 49, l'equazione $A_j = B_j$ è equivalente a $\pi^B_{A_j} = j$ ove

π^B è la permutazione che manda B_1 in 1, ... , B_6 in 6 e numera in maniera crescente partendo da 7 ed arrivando a 49 i rimanenti nell'ordine che hanno. Inoltre al variare della disposizione A la disposizione $j \mapsto \pi^B_{A_j}$ varia tra tutte le disposizioni, e traforma disposizioni A differenti in disposizioni differenti.

In particolare questa trasformazione di disposizioni in disposizioni non cambia la probabilità uniforme sullo spazio avente come eventi elementari le disposizioni indicate con A.

A questo punto (volendo essere formali ma i passagi seguenti possono esser tralasciati) si ha:

$p_{49,6,k} =: \mathbf{P}(\#\{1 \leq j : d_j = \delta_j\} = k) =$

$\sum_{B \ disposizione} \mathbf{P}(\#\{1 \leq j : d_j = \delta_j\} = k|\delta = B)\mathbf{P}(\delta = B) =$

$\sum_{B \ disposizione} \mathbf{P}(\#\{1 \leq j : d_j = B_j\} = k|\delta = B)\mathbf{P}(\delta = B) =$

$\sum_{B \ disposizione} \mathbf{P}(\#\{1 \leq j : \pi^B_{d_j} = j\} = k|\delta = B)\mathbf{P}(\delta = B) =$

qui si usa che si conserva la probabilità in una variabile

$\sum_{B \ disposizione} \mathbf{P}(\#\{1 \leq j : d_j = j\} = k|\delta = B)\mathbf{P}(\delta = B) =$

$\mathbf{P}(\#\{1 \leq j : d_j = j\} = k)$

Quindi il problema è ridotto al calcolo di $p_{49,6,k} = \mathbf{P}(\#\{1 \leq j : d_j = j\} = k)$ ove d è una disposizione casuale uniformemente distribuita.

ORA è utile specificare lo spazio degli eventi elementari: le disposizioni semplici di 6 in 49 con la probabilità uniforme, che vista come rapporto tra numero di casi favorevoli su numero di casi possibili per un insieme di disposizioni D è: $\mathbf{P}(D) = \frac{\#D}{49 \cdot 48 \cdot 47 \cdot 46 \cdot 45 \cdot 44}$.

Per finire:

- si calcola prima la probabilità di alemno una coincidenza contando le disposzioni per cui $\#\{1 \leq j : d_j = j\} \geq 1$ con il seguente ragionamento: ho 6 scelte per l'indice j che deve stare fisso. Gli altri 5 li posso "disporre " nei rimanenti 48 posti, quindi:

$$\mathbf{P}(\#\{1 \leq j : d_j = j\} \geq 1)\frac{6\frac{48!}{43!}}{\frac{49!}{43!}} = \frac{6}{49}$$

- ma allora la probabilità di non azzeccarne manco una è $1 - \frac{6}{49} = \frac{43}{49}$

- ciò posto si osserva che ragionando in modo analogo si ottiene la probabilità desiderata.

- per prima cosa: il numero di disposizioni con almeno un punto fisso di $6 - k$ numeri tra i primi 6 nei primi 49 numeri tranne i k tra i primi 6 che si son tolti, sono $(6 - k)$ (il numero delle scelte di quello che si vuole fisso) per il numero delle disposizioni semplici di $6 - k - 1$(uno è già stato assegnato a se stesso) in $49 - k - 1$(il posto asegnato al primo). Totale: $(6 - k)\frac{(48-k)!}{(43-k)!}$

-quindi si scelgono k indici su 6 che devon rimanere fissi, per ognuna di tali scelte che sono $\binom{6}{k}$ i rimanenti $6-k$ indici devono essere disposti senza andare in se stessi nei rimanenti $49-k$ posti. Cioè si devon contare le disposizioni semplici di $6-k$ in $49-k$ che non lascianmo fisso nulla, ma queste sono il numero di tutte le disposizioni semplici di $6-k$ in $49-k$ meno quello di quelle con almeno un fisso sopra calcolate.

- concludendo:

$$p_{49,6,k} = \binom{6}{k}\frac{(49-k)! - (6-k)(48-k)!}{49!} = \binom{6}{k}\frac{(48-k)!}{49!}43 = \binom{6}{k}\frac{43}{49 \cdot \ldots \cdot (49-k)}$$

2.Segregazione sessuale, indipendenza dal sesso, bigenitorialità sessuata portano allo stesso risultato di frequenze genotipiche di una riproduzione ermafrodita non esclusivamente bigenitoriale.

3.Tra n amici ci sono solo un Paolo e solo un Pietro. Si siedono ad una tavola (rotonda!). Qual'è la media del minimo numero di amici seduti tra Paolo e Pietro ? (cfr. esercizio corrispondente del XIV schema).

4 Media e varianza del numero di successi in n tentativi.

4.1 Media e varianza di una variabile aleatoria di "guadagno pedita" da uno schema di successo insuccesso: $\mathbf{P}(Y = a) = p$, $\mathbf{P}(Y = b) = 1 - p$.

4.2 Media (e varianza) del numero di successi in n tentativi usando le derivate.

4.3 Media e varianza delle percentuali di successo insuccesso su 100 tentativi

4.4 Stima dellla probabilità che tali percentuali siano al più del 10% usando la diseguaglianza di Tchebyschev.

4.5 Calcolo della varianza del numero di successi usando l'indipendenza e confrontando con il precedente risultato.

- Digressione sulla media come baricentro.
- Somma dei primi H numeri naturali, somma dei loro quadrati.
- Somma di numeri di successi indipendenti rispettivamente su n ed m tentativi indipendenti e senza rimpiazzo con egual probabilità p è ancora un numero di successi su $n+m$ tentativi indipendenti e probabilità p.

5 Media del numero N di successi in n estrazioni senza rimpiazzo

5.1 N è somma di n variabili aleatorie $Y_1 \dots Y_n$ *dipendenti* a valori o 0 o 1 (insuccesso successo). Anche se si possono pensare in prima istanza come estrazioni successive senza rimpiazzo da un'urna ciò nasconde il fatto che hanno la stessa probabilità cioè:

$\mathbf{P}(Y_i = a) = \mathbf{P}(Y_1 = a)$

(verifica diretta con Bayes che ciò è vero per l aprima e la seconda estrazione).

anzi

$\mathbf{P}(Y_i = a, Y_j = b) = \mathbf{P}(Y_1 = a, Y_2 = b)$ se $i \neq j$

Dimostrazione del primo fatto considerandole come disposizioni e considerando la trasformazione che scambia il posto i con il primo posto che lascia inalterata la probabilità uniforme sullo spazio delle disposizioni.

5.2 Alla luce di ciò calcolo della media di N per linearità La media comune delle Y_i è sorpendentemente la stessa del caso di estrazioni con rimpiazzo. La media di N sarà quindi quella del numero di successi in n lanci indipendenti!