

APPUNTI COMPLEMENTARI PER IL CORSO DI STATISTICA

CORSI DI LAUREA IN INGEGNERIA EDILE E TESSILE

Ilia Negri*

24 settembre 2002

1 Schemi di campionamento

Con il termine campionamento si intende l'operazione di estrazione di un certo numero di palline da un'urna. Il risultato dell'estrazione è detto campione. L'insieme di tutti i possibili campioni costituisce lo spazio campionario. Diversi esperimenti casuali possono dare come esito uno stesso tipo di campionamento. In questa sezione ci proponiamo di catalogare l'insieme degli esiti di alcuni diversi tipi di esperimenti casuali. Gli esperimenti consistono nell'estrazione di n palline da un'urna che contiene M palline. Essi si distinguono prima di tutto per il modo in cui possono essere estratte le palline e in secondo luogo per l'importanza o meno data all'ordine in cui sono estratte. L'utilizzo del modello costituito dall'urna e dalle palline è fatto solo per semplicità. Come vedremo nelle applicazioni tale modello può essere visto come l'esemplificazione di modelli assai più complessi. Il problema della costruzione di Ω per questi esperimenti si riduce quindi a cercare di dare una risposta alla domanda: *in quanti modi si possono estrarre n palline da un'urna che ne contiene M distinte?* Come abbiamo già accennato il numero di modi in cui si possono estrarre le palline dipende da due fattori che caratterizzano l'esperimento casuale:

1. Si ripone nell'urna la pallina estratta prima dell'estrazione successiva?

*Dipartimento di Ingegneria – Università degli Studi di Bergamo – Dalmine

2. Ha importanza l'ordine con cui le palline sono estratte?

A seconda delle risposte date alle due domande precedenti si distinguono 4 tipi diversi di schema di campionamento.

Si suppone che le M palline contenute nell'urna siano contraddistinte dai numeri $1, 2, \dots, M$. Un campione di lunghezza n estratto dall'urna può essere indicato come (a_1, a_2, \dots, a_n) , dove ciascun a_i per $i = 1, 2, \dots, n$ può assumere valori nell'insieme $\{1, 2, \dots, M\}$ e rappresenta il valore della i -esima estrazione. Il numero delle n -uple (a_1, a_2, \dots, a_n) diverse che si possono formare dipende dalle risposte che si danno ai due ultimi quesiti. Ad esempio una possibile realizzazione nel gioco del totocalcio può essere vista come un campione ottenuto estraendo 13 palline da un'urna che ne contiene 3, riponendo nell'urna la pallina prima di procedere all'estrazione successiva e dando importanza all'ordine con cui le 13 palline sono estratte. Nel seguito si analizzano i quattro schemi diversi di campionamento.

Primo caso. Estrazione con riposizione e si dà importanza all'ordine. In questo caso due campioni (a_1, a_2, \dots, a_n) e (b_1, b_2, \dots, b_n) sono diversi tra loro o se differiscono per il valore che assumono le a_i e le b_i ovvero se queste sono tutte uguali ma cambia l'ordine con cui compaiono nella n -upla. Per contare quanti campioni diversi si possono formare si ragiona nel modo seguente. Nella prima estrazione la pallina può essere scelta in M modi diversi. Nella seconda estrazione, poiché la pallina estratta viene rimessa nell'urna, questa può essere scelta di nuovo in M modi diversi. Per ognuno dei modi in cui è estratta la prima pallina vi sono M modi possibili di estrarre la seconda pallina, quindi in totale M^2 modi di estrarre due palline. In generale se si effettuano n estrazioni si avranno M^n modi diversi di estrarre le n palline. Il numero di campioni ottenuti è anche detto *disposizioni con ripetizione*. Ad esempio le combinazioni possibili al gioco del totocalcio sono $3^{13} = 1594323$. Si noti che in questo schema di campionamento si può avere $M \leq n$.

Secondo caso. Estrazione senza riposizione e si dà importanza all'ordine. Per contare quanti campioni si possono formare si osserva che la prima estrazione può essere fatta in M modi. La seconda può essere fatta in $M - 1$ modi in quanto la pallina scelta alla prima estrazione non viene rimessa nell'urna. Per ogni scelta

della prima pallina vi sono quindi $M - 1$ scelte della seconda. In totale le prime due scelte possono essere effettuate in $M \cdot (M - 1)$ modi. I casi possibili se si eseguono n estrazioni sono quindi $M \cdot (M - 1) \cdot \dots \cdot (M - n + 1)$. Il numero di campioni ottenuti è detto *disposizioni semplici*. Introduciamo la notazione fattoriale dove con $k!$ per k intero e maggiore di zero si intende il prodotto di tutti gli interi da k fino a 1. Si assume per definizione che $0! = 1$. Allora il numero di casi possibili può essere scritto come

$$\frac{M!}{(M - n)!}.$$

Quanti numeri di sei cifre tutte diverse si possono formare? Si tratta di contare il numero dei campioni diversi che si ottengono facendo sei estrazioni senza riposizione in un'urna che contiene 10 palline. Chiaramente in questo tipo di esperimento ha importanza l'ordine con cui vengono estratte le palline in quanto un numero in una posizione assume un significato ben preciso. I casi possibili sono $10 \cdot 9 \cdot \dots \cdot 5 = 151200$. In quanti modi possibili si possono ordinare 12 persone diverse? Ogni ordine possibile corrisponde ad effettuare 12 estrazioni da un'urna contenente 12 palline in questo schema di campionamento. I casi possibili sono $12! = 479001600$.

Terzo caso. Estrazione senza riposizione e non ha importanza l'ordine. Per contare i campioni possibili in questo caso facciamo la seguente osservazione. Due campioni dello schema precedente che differiscono solo per l'ordine delle palline estratte ma non per le palline estratte sono lo stesso campione in questo schema. Per ogni estrazione di n palline diverse vi sono $n!$ modi di ordinare queste n palline. I campioni ottenuti in questo modo rappresentano un'unico campione per questo schema. I casi possibili sono dunque

$$\frac{M \cdot (M - 1) \cdot \dots \cdot (M - n + 1)}{n!}.$$

In pratica si applica quella che viene detta regola del pastore: per sapere quante pecore vi sono nel gregge si contano prima le zampe e poi si divide per quattro. Il numero di campioni ottenuto in questo schema prende il nome di *combinazioni semplici*. Per indicare il numero di campioni ottenuto con questo schema si utilizza

la seguente scrittura:

$$\binom{M}{n} = \frac{M \cdot (M-1) \cdot \dots \cdot (M-n+1)}{n!} = \frac{M!}{(M-n)!n!}.$$

La quantità $\binom{M}{n}$ viene anche detta *coefficiente binomiale*. Ad esempio si supponga di voler sapere quante quintine si possono formare nel gioco del Lotto. In questo caso si devono contare i campioni possibili ottenuti facendo 5 estrazioni senza riposizione da urna che contiene 90 palline. I casi possibili sono $\binom{90}{5} = 43949268$.

Quarto caso.¹ Estrazione con riposizione e non ha importanza l'ordine. Si supponga di identificare le M palline contenute nell'urna con $M-1$ barre che dividono M celle (la prima e l'ultima cella non hanno la parete a sinistra e rispettivamente a destra). L'estrazione di n palline nello schema dell'urna con riposizione e senza dare importanza all'ordine, corrisponde all'assegnazione di n oggetti indistinguibili (chiamiamoli asterischi) alle M celle senza esclusione (cioè ogni cella può contenere più asterischi). Se pensiamo agli $M-1$ bastoncini e agli n asterischi come ad $M-1+n$ oggetti diversi, ogni configurazione corrisponde ad una permutazione di questi oggetti. Ci sono quindi $(M-1+n)!$ configurazioni possibili. Due permutazioni di questo tipo corrispondono allo stesso campione nello schema che si sta considerando, quando, fissate le posizioni e i valori degli asterischi, si permutano gli $M-1$ bastoncini. In modo analogo due permutazioni in cui sono fissate le posizioni dei bastoncini rappresentano lo stesso campione quando si permutano gli asterischi. In definitiva i casi possibili sono

$$\frac{(M-1+n)!}{(M-1)!n!} = \binom{M-1+n}{n}.$$

Il numero di tali campioni è detto *combinazioni con ripetizione*. Ad esempio quante tessere diverse del domino si possono formare? Tante quante il numero di estrazioni diverse che si possono effettuare estraendo due palline con riposizione da un'urna che ne contiene sei. I casi possibili sono $\binom{6-1+2}{2} = 21$.

Per i quattro schemi considerati abbiamo costruito lo spazio Ω degli eventi elementari. Il numero di elementi che appartengono ad Ω nei quattro schemi è riassunto nella tabella 1. Se si vuole calcolare la probabilità di un particolare evento in uno di

¹Questo caso è presentato solo per completezza ma non fa parte del programma d'esame

	ORDINATI	NON ORDINATI
CON RIPOSIZIONE	M^n	$\binom{M-1+n}{n}$
SENZA RIPOSIZIONE	$\frac{M!}{(M-n)!}$	$\binom{M}{n}$

Tabella 1: Numero di campioni elementari nei differenti schemi di campionamento.

questi schemi di campionamento si può supporre che ciascun evento elementare sia equiprobabile. Con l'ipotesi fatta la probabilità degli eventi si può calcolare come numero di casi favorevoli all'evento fratto numero dei casi possibili.

Esempio 1.1. *Calcolare la probabilità di fare ambo avendo giocato 2 numeri su una ruota del Lotto.*

In questo caso i casi possibili sono $\binom{90}{5} = 43949268$. Mentre i casi favorevoli sono $\binom{88}{3} = 109736$. Infatti le cinque favorevoli sono quelle che contengono i due numeri giocati e altri tre numeri qualunque scelti tra gli 88 rimasti. La probabilità richiesta è pertanto $p = 0.0025$.

Quindi per considerare equo il gioco del Lotto una vincita ottenuta con l'ambo dovrebbe essere pagata circa 500 volte non le 50 attuali!