UNITELMA SAPIENZA

SAPIENZA
UNIVERSITÀ DI ROMA
DIPARTIMENTO DI INFORMATICA

Probabilità

Marco Isopi

8. Estrazioni con e senza reinserimento

Scopo della lezione

In questa lezione discuteremo uno dei modelli paradigmatici del calcolo delle probabilità: l'estrazione di palle di (almeno) due colori diversi da un urna.

Molti problemi elementari possono essere riformulati in termini di estrazione da un'urna. Pertanto è utile familiarizzarsi con alcuni problemi di questo tipo.

Abbiamo visto alcuni esempi di estrazioni da urne nelle lezioni precedenti, ma ora vogliamo dare una panoramica più ampia di questo modello.

Estrazioni da un'urna

La situazione di base vede un'urna contenete b palle bianche e r palle rosse per un totale di $N = b + r$ palle.

Nelle estrazioni *con reinserimento* *(o con rimpiazzo)*, una palla, dopo essere stata estratta, viene reintrodotta nell'urna prima dell'estrazione successiva.

Nelle estrazioni *senza reinserimento* *(o senza rimpiazzo)*, una palla, dopo essere stata estratta, viene messa da parte.

Estrazioni da un’urna

Di solito specificheremo quante estrazione vengono effettuate, oppure daremo una regola di arresto (p.e. ci fermiamo dopo l’estrazione della seconda palla rossa).

Naturalmente nel caso di estrazioni senza reinserimento sono possibile al più $r + b$ estrazioni, mentre un’estrazione con reinserimento può andare avanti quanto si vuole.

Calcoliamo ora la probabilità di alcuni eventi relativi a queste procedure.

Estrazioni con rimpiazzo

Nelle estrazioni con reinserimento, la composizione dell'urna è la stessa ad ogni estrazione, pertanto la probabilità di estrarre una palla bianca (o rossa) non dipenderà da quante estrazioni sono state precedentemente effettuate o dal loro esito.

La probabilità di estrarre una palla bianca in una singola estrazione è

$$\mathbf{P}(B) = \frac{b}{N} =: p$$

Ovviamente è $\frac{r}{N} = 1 - \frac{b}{N}$ la probabilità di estrarre una palla rossa.

Notiamo che se raddoppiamo sia il numero di palle rosse che di palle bianche, questa probabilità non cambia.

Estrazioni con rimpiazzo

UNITELMA SAPIENZA

SAPIENZA
Università di Roma
Dipartimento di Informatica

Se effettuiamo n estrazioni la probabilità di estrarre esattamente k palle bianche (e $n - k$ rosse) in un fissato ordine è

$$\left(\frac{b}{N}\right)^k \left(\frac{r}{N}\right)^{n-k} = p^k(1-p)^{n-k}.$$

Per esempio la probilità di osservare in 8 estrazioni la sequenza $r\,r\,n\,n\,n\,r\,n\,n$ è $p^5(1-p)^3$.

Notiamo che tale probabilità non dipende dall'ordine, ma solo da quante palle bianche e quante palle rosse vi compaiono e *dal fatto che abbiamo fissato un ordine*.

UNITELMA SAPIENZA
SAPIENZA
UNIVERSITÀ DI ROMA
DIPARTIMENTO DI INFORMATICA

Se invece vogliamo trovare la probabilità di estrarre esattamente k palle bianche (e $n-k$ rosse) senza tenere conto dell'ordine di estrazione, dobbiamo moltiplicare $p^k(1-p)^{n-k}$ per il numero di sequenze che contengono k palle bianche e $n-k$ rosse.

Abbiamo $\binom{n}{k}$ modi per scegliere le k posizioni in cui compaiono le palle bianche.

Quindi la probabilità di estrarre k palle bianche e $n-k$ rosse è
$$\binom{n}{k} p^k(1-p)^{n-k}.$$

Estrazioni con rimpiazzo

Osserviamo che

$$\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = \big(p + (1-p)\big)^n = 1.$$

Questo è il cosiddetto *modello binomiale* che incontreremo molte volte nelle lezioni successive.

Estrazioni senza rimpiazzo

Nelle estrazioni senza reinserimento, la composizione dell'urna cambia dopo ogni estrazione, pertanto le probabilità di estrarre una palla bianca (o rossa) non dipende da quante estrazioni sono state precedentemente effettuate o dal loro esito.

Supponiamo di avere come prima un'urna con b palle bianche e r palle rosse per un totale di $N = b + r$ palle.

Le probabilità per la prima estrazione sono naturalmente le $\frac{b}{N}$ e $\frac{r}{N} = 1 - \frac{b}{N}$.

La probabilità di estrarre una palla bianca alla seconda estrazione dipende invece dall'esito della prima:

se abbiamo estratto una palla bianca alla prima estrazione, la probabilità di vederne una anche alla seconda è $\frac{b-1}{N-1}$,
mentre se la prima estrazione è stata di una alla rossa, la probabilità è $\frac{b}{N-1}$.

Quindi la probabilità di vedere due palle bianche nelle prime due estrazioni è $\dfrac{b(b-1)}{N(N-1)}$.

Estrazioni senza rimpiazzo

Notiamo che potevamo pensare di estrarre due palle in una volta sola ("in blocco").

In questo caso abbiamo $\binom{N}{2}$ coppie di palle di cui $\binom{b}{2}$ sono coppie di palle bianche.

La probabilità diventa

$$\frac{\binom{b}{2}}{\binom{N}{2}} = \frac{b!}{2!(b-2)!} \cdot \frac{2!(N-2)!}{N!} = \frac{b(b-1)}{N(N-1)},$$

ovvero lo stesso risultato di prima.

Estrazioni senza rimpiazzo

Guardando a una palla bianca e una rossa si ragiona nello stesso modo e la probabilità diventa $\dfrac{b \cdot r}{N(N-1)}$.

Vediamo che la probabilità dipende dal numero di palle dei due colori, ma *non dipende* dall'ordine di estrazione e quindi possiamo pensare anche che le palle siano estratte in blocco.

Possiamo ripetere il ragionamento per un numeri qualunque di palle (ma non più grande di N).

Estrazioni senza rimpiazzo

Per trovare la probabilità di estrarre k palle bianche e $n-k$ rosse osserviamo che ci sono $\binom{N}{n}$ modi di scegliere n palle da N.

Poi ci sono $\binom{b}{k}$ modi di scegliere k palle bianche da b e $\binom{r}{n-k}$ modi di scegliere $n-k$ palle bianche da r.

Estrazioni senza rimpiazzo

Quindi la probabilità di vedere k palle bianche e $n - k$ rosse nelle prime n estrazioni è

$$\frac{\binom{b}{k}\binom{r}{n-k}}{\binom{N}{n}}.$$

Questo è il cosiddetto *modello ipergeometrico* che pure incontreremo molte volte nelle lezioni successive.

Generalizzazioni

Trattare sistematicamente generalizzazioni dei modelli di estrazione da urna ci prenderebbe troppo tempo.

Ci limitiamo quindi a qualche accenno. Variazioni sul tema sono presenti in numerosi esercizi.

UNITELMA SAPIENZA
SAPIENZA
UNIVERSITÀ DI ROMA
DIPARTIMENTO DI INFORMATICA

Più di due colori

In un'urna con più di due colori le estrazioni con rimpiazzo sono descritte dal modello *multimomiale*.

L'analisi è analoga e ci limiteremo per semplicità un esempio con tre colori.
Un'urna contiene b palle bianche, r palle rosse e g palle gialle per un totale di $N = b + r + g$ palle.

Se effettuiamo n estrazioni la probabilità di estrarre esattamente k palle bianche, j rosse e $n - k - j$ gialle in un fissato ordine è

$$\left(\frac{b}{N}\right)^k \left(\frac{r}{N}\right)^j \left(\frac{g}{N}\right)^{n-k-j}.$$

Generalizzazioni

Notiamo che, come prima, tale probabilità non dipende dall'ordine, ma solo da quante palle bianche rosse e gialle vi compaiono e *dal fatto che abbiamo fissato un ordine*.

Se l'ordine di estrazione non ci interessa dobbiamo moltiplicare la probabilità di sopra per il numero di sequenze che contengono k palle bianche, j rosse e $n - k - j$ gialle.

Procediamo come abbiamo fatto con gli anagrammi di parole con lettere ripetute.

Generalizzazioni

Ci sono $n!$ modi per ordinare le palle estratte, ma le $k!$ permutazioni fra palle bianche asciano invariata la sequenza. Lo stesso per le $j!$ permutazioni fra palle rosse e le $(n-k-j)!$ permutazioni fra palle gialle.

Quindi la probabilità di estrarre k palle bianche, j rosse e $n-k-j$ gialle è

$$\frac{n!}{k!\,j!\,(n-k-j)!}\left(\frac{b}{N}\right)^{k}\left(\frac{r}{N}\right)^{j}\left(\frac{g}{N}\right)^{n-k-j}.$$

Generalizzazioni

La stessa analisi funziona per un qualunque numero di colori. Per due colori si riduce al modello binomiale visto prima.

Per le estrazoni senza rimpiazzo generalizziamo invece il modello ipergeometrico.

Sempre per un'urna a tre colori come sopra, possiamo ripetere senza difficoltà il ragionamento che ci ha portato a derivare la probabilità ipergeometrica e otteniamo che la probabilità di vedere k palle bianche, j rosse e $n-k-j$ gialle nelle prime n estrazioni è

$$\frac{\binom{b}{k}\binom{r}{j}\binom{g}{n-k-j}}{\binom{N}{n}}.$$

Generalizzazioni

Altre generalizzazioni di cui si trovano esempi negli esercizi sono:

- urne con rinforzo, dove dopo ogni estrazione oltre a rimettere dentro la palla estratta si inseriscono altre palle a seconda del colore di quella appena estratte;
- numero di estrazioni non fissato in anticipo ma che dipende dall'esito delle estrazioni stesse. Per esempio si decide di arrestarsi dopo che sono avvenute un certo numero di estrazioni di palle di un fissato colore.