Calcolo delle Probabilità e Statistica 2021/2022

Calcolo combinatorio e spazi di probabilità finiti e uniformi

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

1 Problemi di conteggio

In questo capitolo studiamo nel dettaglio il seguente caso particolare: Ω è *finito* e gli esiti sono *equiprobabili*, ovvero la probabilità $\mathbb{P}$ è *uniforme*. Ricordiamo allora che vale il seguente risultato.

Teorema 1.1. *Consideriamo un esperimento aleatorio descritto da uno spazio campionario* ***finito****:*

$$\Omega \;=\; \{\omega_1, \dots, \omega_N\}$$

con esiti ***equiprobabili****, ovvero con* ***probabilità uniforme****:*

$$\mathbb{P}(\{\omega_1\}) \;=\; \mathbb{P}(\{\omega_2\}) \;=\; \cdots \;=\; \mathbb{P}(\{\omega_N\}).$$

Allora valgono le proprietà seguenti:

1) *dato un qualunque evento elementare $\{\omega_i\}$, vale che*

$$\mathbb{P}(\{\omega_i\}) \;=\; \frac{1}{N};$$

2) *dato un qualunque evento A, vale la* ***formula di Laplace***

$$\mathbb{P}(A) \;=\; \frac{n^o \textit{ di eventi elementari che compongono } A}{N} \;=\; \frac{\textbf{\textit{casi favorevoli}}}{\textbf{\textit{casi possibili}}}.$$

I problemi che rientrano nella situazione qui sopra descritta sono detti **problemi di conteggio**, in quanto il calcolo della probabilità di un evento A si riduce al conteggio del numero di casi favorevoli e del numero di casi possibili. Il *calcolo combinatorio* è lo strumento matematico che permette di svolgere questi calcoli anche quando tali numeri sono particolarmente elevati.

1.1 Cardinalità e corrispondenza biunivoca

Ricordiamo che il simbolo $|\Omega|$ (oppure $\#\Omega$) indica la *cardinalità* di un qualunque insieme Ω, ovvero il numero dei suoi elementi. La formula di Laplace può essere scritta in termini di cardinalità come segue:

$$\mathbb{P}(A) \;=\; \frac{|A|}{|\Omega|}.$$

Corrispondenza biunivoca. Dati due insiemi A e B, si dice che

A è in corrispondenza biunivoca con B

se, per definizione, esiste una funzione $f\colon A \to B$ biettiva, cioè iniettiva e suriettiva.

Ricordiamo dunque il seguente *principio basilare*:

$$|A| \;=\; |B| \quad \textit{se e solo se} \quad A \text{ e } B \text{ sono in corrispondenza biunivoca.}$$

Per determinare la cardinalità di un insieme A spesso si ricorre alla *corrispondenza biunivoca*, ovvero si determina un altro insieme B che si sa essere in corrispondenza biunivoca con A, quindi $|A| = |B|$, e di cui è più facile calcolare la cardinalità.

1.2 Fattoriale e coefficiente binomiale

Ricordiamo il significato del simbolo *fattoriale:*

$$n! \;=\; n(n-1)\cdots 1, \qquad \forall\, n = 1, 2, \ldots$$

Inoltre si pone per convenzione

$$0! \;=\; 1.$$

Il *coefficiente binomiale* è invece dato da

$$\binom{n}{k} \;=\; \frac{n!}{k!(n-k)!}, \qquad \forall\, n, k = 0, 1, 2, \ldots, \text{ con } k \leq n.$$

Dalla definizione segue direttamente che

$$\binom{n}{k} = \binom{n}{n-k}, \qquad \binom{n}{0} = \binom{n}{n} = 1, \qquad \binom{n}{1} = n.$$

Inoltre, per $k, n \in \mathbb{N}$ con $k < n$, vale la *formula di Stifel:*

$$\binom{n}{k} \;=\; \binom{n-1}{k-1} + \binom{n-1}{k}.$$

Infine, vale la *formula del binomio di Newton:*

$$(a+b)^n \;=\; \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}, \qquad a, b \in \mathbb{R}.$$

1.3 Metodo delle scelte successive

In questa sezione illustriamo un metodo, noto come *metodo delle scelte successive* (o *schema delle scelte successive* o anche *principio fondamentale del calcolo combinatorio*), che permette di determinare la cardinalità di un insieme una volta caratterizzati univocamente i suoi elementi tramite un numero finito di scelte successive. Iniziamo con un esempio.

Esempio 1.1. *Quante password di otto caratteri, ognuno dei quali scelto tra trentasei valori alfanumerici, possono essere generate?*
Come cambia la risposta se gli otto caratteri devono essere tra loro distinti?

Soluzione.

1) *Caratteri* ***non necessariamente*** *distinti.* Sia Ω l'insieme di tutte le password di otto caratteri non necessariamente distinti. In tal caso Ω è dato da

$$\Omega \;=\; \{a, b, c, \dots, 8, 9\}^8.$$

Possiamo determinare ogni password di Ω tramite le seguenti otto scelte successive:

- scelta del primo carattere della password: 36 possibilità
- scelta del secondo carattere: 36 possibilità
- $\cdots$
- scelta dell'ottavo carattere: 36 possibilità

Come seguirà dal *metodo delle scelte successive*, la cardinalità di Ω è data dal prodotto di questi otto numeri:

$$|\Omega| \;=\; 36 \cdot 36 \cdot \dots \cdot 36 \;=\; 36^8.$$

Questo risultato è confermato dal fatto che $\Omega = \{a, b, c, \dots, 8, 9\}^8$.

2) *Caratteri distinti.* Sia Ω l'insieme di tutte le password di otto caratteri tra loro distinti. Procedendo come prima possiamo determinare ogni password di Ω tramite le seguenti scelte successive:

- scelta del primo carattere della password: 36 possibilità
- scelta del secondo carattere: $36 - 1$ possibilità
- $\cdots$
- scelta dell'ottavo carattere: $36 - 7$ possibilità

Come seguirà dal *metodo delle scelte successive*, la cardinalità di Ω è data da

$$|\Omega| \;=\; 36 \cdot (36-1) \cdot \dots \cdot (36-7) \;=\; 36 \cdot 35 \cdot \dots \cdot 29.$$

$\square$

Formuliamo dunque il metodo delle scelte successive, che generalizza quanto appena visto nell'esempio precedente.

Metodo delle scelte successive. *Supponiamo che* ***ciascun*** *elemento di un insieme A possa essere determinato tramite* ***una e una sola*** *sequenza di k scelte successive, in cui ogni scelta viene effettuata tra un numero fissato di possibilità (tale* numero di possibilità, *qui di seguito indicato con $n_1, n_2, \dots, n_k$, non dipende dalle scelte precedenti ma solo da k):*

- *la prima scelta viene effettuata tra n_1 possibilità*

- *la seconda scelta viene effettuata tra* n_2 *possibilità*
- $\cdots$
- *la* k*-esima scelta viene effettuata tra* n_k *possibilità*

Allora la cardinalità di A *è pari a*

$$|A| \ = \ n_1 \cdot n_2 \cdots n_k.$$

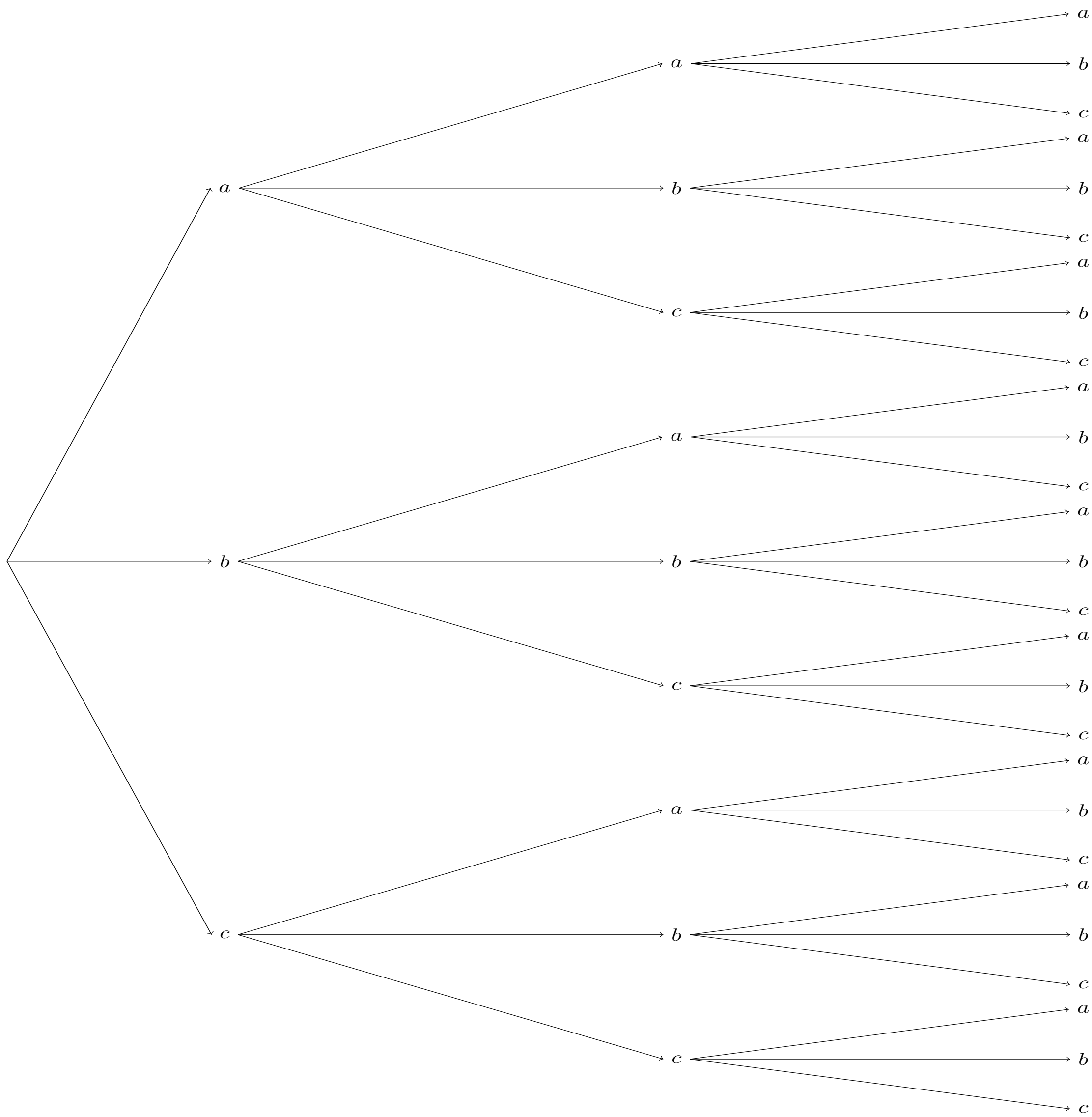
Osservazione 1. *Così enunciato, il metodo delle scelte successive sembra essere un po' vago. Una riformulazione matematica precisa (come teorema) è possibile, tuttavia essa comporta notazioni piuttosto ingombranti e risulta di poco aiuto nelle applicazioni. Per tale ragione nella pratica si fa tipicamente riferimento all'enunciato riportato qui sopra.*

Osservazione 2. *Il metodo delle scelte successive dice essenzialmente che l'insieme* A *è in corrispondenza biunivoca con le sequenze di* k *scelte, il cui numero totale è appunto pari a* $n_1 \cdot n_2 \cdots n_k$.

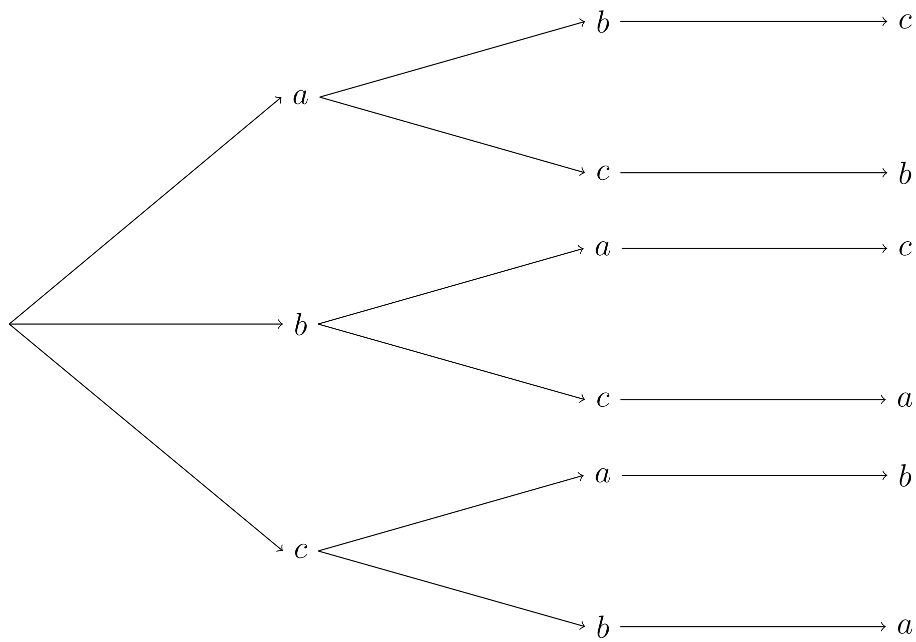
Osservazione 3. *Gli errori più comuni che si commettono nell'utilizzo di tale metodo sono:*

- *non contare tutti gli elementi di* A *(da qui l'importanza del termine "**ciascun**" nell'enunciato del metodo delle scelte successive);*
- *contare più di una volta lo stesso elemento (da qui l'importanza del termine "**una e una sola**" nell'enunciato del metodo delle scelte successive).*

Osservazione 4. *Come abbiamo visto per gli esperimenti aleatori, anche il metodo delle scelte successive può essere descritto graficamente tramite un* diagramma ad albero*, anche se non insisteremo su questo punto per non creare confusione. Il diagramma ad albero relativo al metodo delle scelte successive è costruito in modo tale che ogni ramificazione corrisponda ad una scelta. Consideriamo ad esempio la seguente versione semplificata del problema della password (Esempio 1.1):* quante password di tre caratteri, ognuno dei quali scelto tra le tre lettere a, b, c, possono essere generate?



Ogni cammino, *dalla radice ad una foglia,* ***corrisponde ad uno e un solo elemento dell'insieme*** A, *ovvero ad una e una sola password. Il diagramma ad albero relativo al metodo delle scelte successive ha come peculiarità che ogni ramificazione ha sempre lo stesso numero di rami: ad esempio, la terza ramificazione ha sempre* n_3 *rami. Non è tuttavia necessario, se si considera ad esempio la terza ramificazione, che compaiano sempre le stesse* n_3 *scelte. Ciò può infatti dipendere da quali nodi sono stati scelti nelle precedenti ramificazioni, come accade nell'esempio della password con caratteri distinti:*



Esempio 1.2. *Consideriamo un mazzo di carte da poker da 52 carte, identificate dal* ***seme*** *(cuori* $\heartsuit$*, quadri* $\diamondsuit$*, fiori* $\clubsuit$*, picche* $\spadesuit$*) e dal* ***tipo*** *(un numero da 2 a 10 oppure J, Q, K, A). Calcolare la cardinalità dei seguenti insiemi.*

1) *L'insieme dei* ***full****; un* full *è un sottoinsieme di* 5 *carte costituito dall'unione di un* tris *(un sottoinsieme di* 3 *carte dello stesso* tipo*) e di una* coppia *(un sottoinsieme di* 2 *carte dello stesso* tipo*, necessariamente diverso da quel del tris).*

2) *L'insieme delle* ***doppie coppie****; una* doppia coppia *è un sottoinsieme di* 5 *carte costituito da due* coppie *di* tipi *diversi, più una quinta carta di* tipo *diverso dai* tipi *delle due coppie.*

Soluzione.

1) Sia A l'insieme dei full. Ogni elemento di A può essere determinato tramite quattro scelte successive:

   - scelta del tipo del tris: 13 possibilità
   - scelta del tipo della coppia: 12 possibilità (chiaramente il tipo della coppia deve essere diverso dal tipo del tris perché non esistono cinque carte dello stesso tipo)
   - scelta dei semi delle carte che compaiono nel tris: 4 possibilità
   - scelta dei semi delle carte che compaiono nella coppia: 6 possibilità

   Quindi
$$|A| \;=\; 13 \cdot 12 \cdot 4 \cdot 6 \;=\; 3\,744.$$

2) Sia B l'insieme delle doppie coppie. Per calcolare $|B|$ si potrebbe essere tentati di procedere analogamente al caso dei full, attraverso sei scelte successive:

- scelta del tipo della prima coppia: 13 possibilità
- scelta del tipo della seconda coppia: 12 possibilità
- scelta del tipo della quinta carta: 11 possibilità
- scelta dei semi delle carte che compaiono nella prima coppia: 6 possibilità
- scelta dei semi delle carte che compaiono nella seconda coppia: 6 possibilità
- scelta del seme della quinta carta: 4 possibilità

Si otterrebbe dunque

$$|B| \;=\; 13 \cdot 12 \cdot 11 \cdot 6 \cdot 6 \cdot 4 \;=\; 247\,104.$$

Tuttavia questo risultato è errato. Infatti, ogni doppia coppia non viene determinata da *una e una sola* sequenza di 6 scelte, ma da *esattamente due sequenze distinte.* La ragione è che le prime due scelte sono ambigue, dal momento che non esiste una "prima" e una "seconda" coppia. Per esempio, la doppia coppia $\{5\heartsuit, 5\diamondsuit, 6\heartsuit, 6\clubsuit, 7\spadesuit\}$ viene determinata sia compiendo come prima scelta 5 e come seconda scelta 6, sia viceversa. Per tale ragione il risultato corretto è

$$|B| \;=\; \frac{247104}{2} \;=\; 123\,552.$$

Un modo alternativo di ottenere il risultato corretto è di riunire le prime due scelte nell'unica scelta seguente:

- scelta dei tipi delle due coppie: $\frac{13 \cdot 12}{2} = 78$ possibilità.

Si ottiene

$$|B| \;=\; 78 \cdot 11 \cdot 6 \cdot 6 \cdot 4 \;=\; 123\,552.$$

$\square$

2 Disposizioni e combinazioni

In questa sezione introduciamo le *disposizioni con ripetizione*, le *disposizioni semplici* (o *senza ripetizione*) e le *combinazioni* (*semplici* o *senza ripetizione*). Nel seguito indicheremo con E un insieme di n elementi distinti:

$$E \;=\; \{e_1, e_2, \ldots, e_n\}.$$

E sarà ad esempio l'insieme della carte che compongono un mazzo, oppure l'insieme delle palline contenute in un'urna (in cui l'i-esima pallina ha come etichetta e_i).

2.1 Disposizioni con ripetizione

Definizione 2.1 (Disposizioni con ripetizione). *Siano E un insieme con $|E| = n$ e $k \in \mathbb{N}$. Indichiamo con $\mathbf{DR}_{n,k}$ l'insieme delle* disposizioni con ripetizione di k elementi di E, *ossia l'insieme di tutte le **sequenze ordinate** di k elementi di E, **non necessariamente distinti**:*

$$\mathbf{DR}_{n,k} \;=\; \{(x_1, \dots, x_k) \colon x_i \in E\} \;=\; \underbrace{E \times E \times \cdots \times E}_{k \text{ volte}}.$$

La cardinalità di $\mathbf{DR}_{n,k}$ è pari a

$$|\mathbf{DR}_{n,k}| \;=\; n^k.$$

Esempio 2.1. *Siano $E = \{a, b, c\}$ e $k = 2$. Allora $|\mathbf{DR}_{3,2}| = 3^2$ e precisamente*

$$\mathbf{DR}_{3,2} \;=\; \{(a, a), (a, b), (a, c), (b, a), (b, b), (b, c), (c, a), (c, b), (c, c)\}.$$

Osservazione 1. $\mathbf{DR}_{n,k}$ *esprime i modi in cui possiamo disporre, in maniera* **ordinata** *ed eventualmente* **ripetuta***, un numero k di oggetti scelti da un insieme di n oggetti.*

Osservazione 2. *Si noti che scriviamo $\mathbf{DR}_{n,k}$ senza specificare l'insieme E, dato che ogni volta sarà chiaro dal contesto a quale insieme di oggetti E ci stiamo riferendo.*

Osservazione 3. *La cardinalità di $\mathbf{DR}_{n,k}$ si trova applicando il metodo delle scelte successive, procedendo come nel punto 1) dell'Esempio 1.1.*

Esempio 2.2. *Si consideri un'urna contenente n palline, etichettate con $e_1, e_2, \dots, e_n$, da cui si estraggono **con reimmissione** $k \in \mathbb{N}$ palline. Sia $E = \{e_1, e_2, \dots, e_n\}$. Uno spazio di probabilità $(\Omega, \mathbb{P})$ che descrive tale esperimento è*

$$\Omega \;=\; \mathbf{DR}_{n,k}, \qquad\qquad \mathbb{P} \;\; \textit{probabilità uniforme}.$$

La quantità $|\mathbf{DR}_{n,k}| = n^k$ è dunque pari al numero di "casi possibili" di questo esperimento aleatorio.

Esempio 2.3. *Determinare un possibile spazio campionario per i seguenti esperimenti aleatori[a]:*

i) *Si sceglie a caso una parola (anche senza senso) composta da 8 lettere dell'alfabeto italiano.*

[a] Soluzioni: i) $\Omega = \mathbf{DR}_{21,8}$, quindi $|\Omega| = 21^8$; ii) $\Omega = \mathbf{DR}_{3,13}$, quindi $|\Omega| = 3^{13}$; iii) $\Omega = \mathbf{DR}_{6,10}$, quindi $|\Omega| = 6^{10}$.

ii) *Si gioca una schedina al totocalcio, in cui per ognuna delle* 13 *partite si può scegliere tra* 1*,* 2 *o X.*

iii) *Si lancia* 10 *volte un dado (non truccato) a sei facce.*

2.2 Disposizioni semplici

Definizione 2.2 (Disposizioni semplici). *Siano E un insieme con $|E| = n$ e $k \leq n$. Indichiamo con $\mathbf{D}_{n,k}$ l'insieme delle* disposizioni semplici *(o* senza ripetizione*)* di k elementi di E*, ossia l'insieme di tutte le* ***sequenze ordinate*** *di k elementi* ***distinti*** *di E:*

$$\mathbf{D}_{n,k} \; = \; \{(x_1, \dots, x_k) \colon x_i \in E \text{ distinti}\}.$$

La cardinalità di $\mathbf{D}_{n,k}$ è pari a

$$|\mathbf{D}_{n,k}| \; = \; n(n-1)\cdots(n-k+1) = \frac{n!}{(n-k)!}.$$

Esempio 2.4. *Siano $E = \{a, b, c\}$ e $k = 2$. Allora $|\mathbf{D}_{3,2}| = 6$ e precisamente*

$$\mathbf{D}_{3,2} \; = \; \{(a,b), (a,c), (b,a), (b,c), (c,a), (c,b)\}.$$

Osservazione 1. $\mathbf{D}_{n,k}$ *esprime i modi in cui possiamo disporre, in maniera* **ordinata** *e* **non ripetuta***, un numero k di oggetti scelti da un insieme di n oggetti.*

Osservazione 2. *La cardinalità di $\mathbf{D}_{n,k}$ si trova applicando il metodo delle scelte successive come nel punto 2) della soluzione dell'Esempio 1.1.*

Esempio 2.5. *Si consideri un'urna contenente n palline, etichettate con $e_1, e_2, \dots, e_n$, da cui si estraggono* ***senza reimmissione*** *$k \leq n$ palline. Sia $E = \{e_1, e_2, \dots, e_n\}$. Uno spazio di probabilità $(\Omega, \mathbb{P})$ che descrive tale esperimento è*

$$\Omega \; = \; \mathbf{D}_{n,k}, \qquad\qquad \mathbb{P} \text{ probabilità uniforme.}$$

La quantità $|\mathbf{D}_{n,k}| = n(n-1)\cdots(n-k+1)$ è dunque pari al numero di "casi possibili" di questo esperimento aleatorio.

Esempio 2.6. *Supponiamo di giocare un'unica cinquina (ad esempio la sequenza ordinata* 13, 5, 45, 21, 34*) al gioco del lotto, in cui si estraggono* senza reimmissione *cinque numeri dai primi novanta naturali.*

1) *Qual è la probabilità di fare una* cinquina secca *(per cui* conta *l'ordine di estrazione)?*

2) *Qual è la probabilità di fare una* cinquina semplice *(per cui* non conta *l'ordine di estrazione)?*

Soluzione. Come suggerito dall'Esempio 2.5, consideriamo lo spazio di probabilità $(\Omega, \mathbb{P})$ dato da

$$\Omega \;=\; \mathbf{D}_{90,5}, \qquad\qquad \mathbb{P} \text{ probabilità uniforme.}$$

1) La probabilità di fare una *cinquina secca* è semplicemente

$$\frac{1}{|\mathbf{D}_{90,5}|} \;\approx\; 1.89 \cdot 10^{-10}.$$

2) Se invece si considera una *cinquina semplice*, dobbiamo innanzitutto contare in quanti modi differenti si possono ordinare 5 numeri. Tale numero è pari a $|\mathbf{D}_{5,5}| = 5!$
Allora la probabilità di una cinquina semplice dopo 5 estrazioni è

$$\frac{|\mathbf{D}_{5,5}|}{|\mathbf{D}_{90,5}|} \;\approx\; 2.27 \cdot 10^{-8}.$$

$\square$

Introduciamo infine le *permutazioni*, che sono un caso particolare di disposizione semplice.

Definizione 2.3 (Permutazioni). *Sia E un insieme con $|E| = n$. Indichiamo con $\mathbf{P}_n$ l'insieme delle* permutazioni degli n elementi di E, *ossia l'insieme $\mathbf{D}_{n,n}$.*
La cardinalità di $\mathbf{P}_n$ è dunque pari a

$$|\mathbf{P}_n| \;=\; |\mathbf{D}_{n,n}| \;=\; n!$$

Osservazione. $\mathbf{P}_n$ *esprime i modi in cui possiamo disporre, in maniera* **ordinata** *e* **non ripetuta**, *n oggetti.*

2.3 Combinazioni

Definizione 2.4 (Combinazioni). *Siano E un insieme con $|E| = n$ e $k \leq n$. Indichiamo con $\mathbf{C}_{n,k}$ l'insieme delle* combinazioni *(*semplici *o* senza ripetizione*)* di k elementi di E, *ossia la famiglia dei sottoinsiemi di E di cardinalità k:*

$$\mathbf{C}_{n,k} \;=\; \{A \subset E \colon |A| = k\}.$$

La cardinalità di $\mathbf{C}_{n,k}$ è pari a

$$|\mathbf{C}_{n,k}| \;=\; \binom{n}{k}. \tag{2.1}$$

Esempio 2.7. *Siano* $E = \{a, b, c\}$ *e* $k = 2$*. Allora* $|\mathbf{C}_{3,2}| = 3$ *e precisamente*

$$\mathbf{C}_{3,2} = \big\{\{a, b\}, \{a, c\}, \{b, c\}\big\}.$$

OSSERVAZIONE 1. $\mathbf{C}_{n,k}$ *è l'insieme di tutti i gruppi di* k *oggetti scelti da un insieme di* n *oggetti, in maniera* **non ordinata** *e* **non ripetuta**.

OSSERVAZIONE 2 (CALCOLO DI $|\mathbf{C}_{n,k}|$). *A differenza del calcolo di* $|\mathbf{DR}_{n,k}|$ *e* $|\mathbf{D}_{n,k}|$*, non è possibile scomporre il calcolo di* $|\mathbf{C}_{n,k}|$ *in una sequenza di scelte successive. Tuttavia, dimostrare la* (2.1) *equivale a dimostrare la seguente uguaglianza:*

$$\frac{n!}{(n-k)!} \;=\; |\mathbf{C}_{n,k}|\, k!$$

ossia

$$|\mathbf{D}_{n,k}| \;=\; |\mathbf{C}_{n,k}|\, |\mathbf{P}_k|. \tag{2.2}$$

Dimostriamo la (2.2) *applicando il metodo delle scelte successive all'insieme* $\mathbf{D}_{n,k}$*:*

- *scelta dei* k *elementi di* E *da ordinare, ovvero di un sottoinsieme di* E *di cardinalità* k*:* $|\mathbf{C}_{n,k}|$ *possibilità (per definizione di* $\mathbf{C}_{n,k}$*)*
- *scelta dell'ordine in cui disporre i* k *elementi, ossia permutazione dei* k *elementi:* $|\mathbf{P}_k|$ *possibilità*

Dal metodo delle scelte successive si ottiene (2.2) *e dunque* (2.1).

Esempio 2.8. *Si consideri un'urna contenente* n *palline, etichettate con* $e_1, e_2, \dots, e_n$*, da cui si estraggono* ***simultaneamente*** $k \leq n$ *palline. Sia* $E = \{e_1, e_2, \dots, e_n\}$*. Uno spazio di probabilità* $(\Omega, \mathbb{P})$ *che descrive tale esperimento è*

$$\Omega \;=\; \mathbf{C}_{n,k}, \qquad\qquad \mathbb{P} \;\; \textit{probabilità uniforme.}$$

La quantità $|\mathbf{C}_{n,k}| = \binom{n}{k}$ *è dunque pari al numero di "casi possibili" di questo esperimento aleatorio.*

Esempio 2.9. *Si consideri la* formula di Stifel:

$$\binom{n}{k} \;=\; \binom{n-1}{k-1} + \binom{n-1}{k}.$$

Possiamo interpretare questa formula in termini di combinazioni come segue:

$$\underbrace{\binom{n}{k}}_{\substack{n^o \text{ di combinazioni} \\ \text{di } k \text{ elementi di } E}} \quad = \quad \underbrace{\binom{n-1}{k-1}}_{\substack{n^o \text{ di combinazioni} \\ \text{di } k \text{ elementi di } E \text{ in cui} \\ \text{è presente l'elemento } \bar{e}}} \quad + \quad \underbrace{\binom{n-1}{k}}_{\substack{n^o \text{ di combinazioni} \\ \text{di } k \text{ elementi di } E \\ \text{senza l'elemento } \bar{e}}}$$

dove $\bar{e}$ *è un elemento fissato di* E *scelto in modo arbitrario.*

Esempio 2.10. *Si consideri la* formula del binomio di Newton:

$$(a+b)^n \;=\; \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}, \qquad a, b \in \mathbb{R}.$$

Una dimostrazione di carattere combinatorio della formula di Newton è la seguente. Il prodotto $(a+b)(a+b)\cdots(a+b)$ di n fattori si sviluppa in una somma di monomi di grado n del tipo $a^{n-k}b^k$ con $0 \le k \le n$. Resta dunque da determinare il coefficiente di ciascun monomio $a^{n-k}b^k$, ossia calcolare quante volte compare facendo il prodotto $(a+b)(a+b)\cdots(a+b)$. Tale monomio si ottiene scegliendo il valore b da k degli n fattori disponibili e, quindi, scegliendo a dai rimanenti $n-k$, ovvero in $\binom{n}{k}$ modi.

3 Tre esperimenti aleatori di riferimento

Estrazioni da un'urna

Quando si utilizza il calcolo combinatorio per lo studio di un esperimento aleatorio, la scelta dello spazio campionario Ω è importante perché può semplificare il conteggio dei casi possibili e dei casi favorevoli. La scelta più conveniente dipende dall'esperimento aleatorio in questione. Tuttavia nella maggior parte dei casi è possibile scegliere come spazio campionario uno dei tre spazi che ora introduciamo. Questo perché è possibile ripensare l'esperimento aleatorio che si sta studiando come uno dei tre esperimenti aleatori di riferimento riportati qui di seguito (che corrispondono agli Esempi 2.2, 2.5, 2.8).

Si consideri un'urna contenente n palline, etichettate con $e_1, e_2, \dots, e_n$. Si estraggono k palline dall'urna in uno dei tre modi seguenti:

1) **estrazione *con reimmissione***[1], con $k \in \mathbb{N}$, in cui, per l'estrazione successiva, la pallina estratta viene reinserita nell'urna (Esempio 2.2);

2) **estrazione *senza reimmissione***, con $k \le n$, in cui la pallina estratta non viene reinserita nell'urna (Esempio 2.5);

3) **estrazione *simultanea***, con $k \le n$, in cui le k palline vengono estratte simultaneamente (Esempio 2.8).

Possiamo descrivere sinteticamente tali esperimenti tramite la seguente tabella, evidenziando i due aspetti caratterizzanti che sono l'*ordine* e la *ripetizione:*

[1] Invece di "reimmissione" si utilizzano anche i termini "reimbussolamento", "reinserimento", "reintroduzione", "restituzione", "rimpiazzo".

Ripetizione \ Ordine	***Senza* ripetizione**	***Con* ripetizione**
Si tiene conto dell'ordine	Estrazione *senza reimmissione*	Estrazione *con reimmissione*
***Non* si tiene conto dell'ordine**	Estrazione *simultanea*	—

Negli Esempi 2.2, 2.5, 2.8 abbiamo introdotto gli spazi di probabilità che descrivono questi tre esperimenti aleatori. Possiamo dunque completare la precedente tabella, riportando anche gli spazi campionari e le loro cardinalità (ossia il numero di "casi possibili").

Ripetizione \ Ordine	***Senza* ripetizione**	***Con* ripetizione**
Si tiene conto dell'ordine	Estrazione *senza reimmissione* $\Omega = \mathbf{D}_{n,k}$ $\lvert\Omega\rvert = \frac{n!}{(n-k)!}$	Estrazione *con reimmissione* $\Omega = \mathbf{DR}_{n,k}$ $\lvert\Omega\rvert = n^k$
***Non* si tiene conto dell'ordine**	Estrazione *simultanea* $\Omega = \mathbf{C}_{n,k}$ $\lvert\Omega\rvert = \frac{\lvert\mathbf{D}_{n,k}\rvert}{k!} = \binom{n}{k}$	—

Osservazione 1. *Nonostante gli esperimenti aleatori introdotti siano tre, in realtà sarebbe sufficiente considerare solamente i primi due: l'estrazione senza reimmissione e l'estrazione con reimmissione. Infatti l'estrazione simultanea può essere vista come un'estrazione senza reimmissione in cui* non si tiene conto dell'ordine*, ossia come un caso particolare dell'estrazione senza reimmissione. Questo significa che, in alternativa a $\mathbf{C}_{n,k}$, è possibile utilizzare $\mathbf{D}_{n,k}$ come spazio campionario. Ciò segue da $|\mathbf{D}_{n,k}| = k!|\mathbf{C}_{n,k}|$ e, più in generale, dal fatto che ad ogni elemento di $\mathbf{C}_{n,k}$ corrispondono $k!$ elementi di $\mathbf{D}_{n,k}$. Più precisamente, vale la catena di uguaglianze:*

$$\frac{\textit{casi favorevoli in } \mathbf{C}_{n,k}}{\textit{casi possibili in } \mathbf{C}_{n,k}} = \frac{k!\,(\textit{casi favorevoli in } \mathbf{C}_{n,k})}{k!\,(\textit{casi possibili in } \mathbf{C}_{n,k})} = \frac{\textit{casi favorevoli in } \mathbf{D}_{n,k}}{\textit{casi possibili in } \mathbf{D}_{n,k}}. \tag{3.1}$$

Osservazione 2. *La casella vuota nella tabella sopra riportata corrisponde all'insieme delle cosiddette* combinazioni con ripetizione*, ossia all'insieme di tutti i gruppi,* non ordinati *ed eventualmente* ripetuti*, di k oggetti da un insieme di n oggetti. A tal proposito, consideriamo il seguente esempio.*

Si lancia tre volte una moneta. Qual è la probabilità che esca due volte testa?

Dato che possono verificarsi delle ripetizioni, consideriamo come spazio campionario l'insieme delle disposizioni con ripetizione $\mathbf{DR}_{2,3}$ *con* $\mathbb{P}$ *probabilità uniforme. L'insieme* $\mathbf{DR}_{2,3}$ *è costituito dalle nove sequenze ordinate:*

$$(T,T,T),\ \ (T,T,C),\ \ (T,C,T),\ \ (C,T,T),\ \ (T,C,C),\ \ (C,T,C),\ \ (C,C,T),\ \ (C,C,C).$$

Notiamo tuttavia che l'evento

$$A \ = \ \text{“esce due volte testa”}$$

non tiene conto dell'ordine. In tal caso, un altro possibile spazio campionario consiste appunto nell'insieme delle combinazioni con ripetizione, che indichiamo con $\mathbf{CR}_{2,3}$, *che è costituito dalle quattro sequenze*[2] *non ordinate:*

$$[T,T,T], \qquad [T,T,C], \qquad [T,C,C], \qquad [C,C,C].$$

Su $\mathbf{DR}_{2,3}$ *la probabilità è quella uniforme. Quale probabilità va invece utilizzata su* $\mathbf{CR}_{2,3}$*? La risposta è facile se si confrontano le sequenze riportate sopra, infatti l'unica scelta coerente con quella di prendere su* $\mathbf{DR}_{2,3}$ *la probabilità uniforme è la seguente:*

$$\begin{aligned}
\mathbb{P}([T,T,T]) \ &= \ \mathbb{P}((T,T,T)) \ = \ \frac{1}{6}, \\
\mathbb{P}([T,T,C]) \ &= \ \mathbb{P}((T,T,C)) + \mathbb{P}((T,C,T)) + \mathbb{P}((C,T,T)) \ = \ \frac{1}{2}, \\
\mathbb{P}([T,C,C]) \ &= \ \mathbb{P}((T,C,C)) + \mathbb{P}((C,T,C)) + \mathbb{P}((C,C,T)) \ = \ \frac{1}{2}, \\
\mathbb{P}([C,C,C]) \ &= \ \mathbb{P}((C,C,C)) \ = \ \frac{1}{6}.
\end{aligned}$$

La probabilità da adottare su $\mathbf{CR}_{2,3}$ *non è dunque uniforme. In conclusione, quando si studia un esperimento aleatorio con ripetizione, sia che si tenga conto dell'ordine sia che non se ne tenga conto, conviene considerare come spazio campionario* $\Omega = \mathbf{DR}_{n,k}$, *in modo tale che* $\mathbb{P}$ *sia la probabilità uniforme (così da poter utilizzare il Calcolo combinatorio per calcolare le probabilità degli eventi). Al contrario, scegliendo* $\Omega = \mathbf{C}_{n,k}$, *è necessario considerare una probabilità* non uniforme. *Ciò è dovuto al fatto che non vale più la catena di uguaglianze* (3.1) *nel caso in cui ci siano ripetizioni. In altri termini, ad ogni combinazione con ripetizione non corrisponde sempre lo stesso numero di elementi di* $\mathbf{DR}_{n,k}$ *(infatti dipende da quante ripetizioni ci sono all'interno della combinazione).*

Esempio 3.1 (**Probabilità binomiale**). *Si consideri un'urna che contiene* b *palline bianche ed* r *palline rosse. Si effettuano* n *estrazioni* con reimmissione. *Calcolare la probabilità dell'evento*

$$A_k \ = \ \text{“si estraggono } k \text{ palline bianche ed } n-k \text{ palline rosse”}$$

con $0 \leq k \leq n$.

[2] Si noti che, ad esempio, dire che si è verificata la combinazione $[T,T,C]$ equivale a dire che si è verificata una delle tre disposizioni (T,T,C), (T,C,T), (C,T,T) senza specificare quale.

Soluzione. Etichettiamo le b palline bianche con $bianca_1, bianca_2, \dots, bianca_b$; analogamente, le r palline rosse con $rossa_1, rossa_2, \dots, rossa_r$. Sia dunque

$$E \; = \; \{bianca_1, bianca_2, \dots, bianca_b, rossa_1, rossa_2, \dots, rossa_r\}.$$

Si noti che $|E| = b + r$.
Come spazio di probabilità $(\Omega, \mathbb{P})$ è naturale considerare $\Omega = \mathbf{DR}_{b+r,n}$ (insieme delle disposizioni con ripetizione di n elementi di E) e $\mathbb{P}$ probabilità uniforme.

Determiniamo la cardinalità di A_k tramite le tre seguenti scelte successive:

- scelta della sequenza (ordinata e con eventuali ripetizioni) delle k palline bianche estratte: $|\mathbf{DR}_{b,k}|$ possibilità;
- scelta della sequenza (ordinata e con eventuali ripetizioni) delle $n-k$ palline rosse estratte: $|\mathbf{DR}_{r,n-k}|$ possibilità;
- scelta delle k estrazioni in cui sono uscite le palline bianche: $|\mathbf{C}_{n,k}|$ possibilità[3].

In definitiva

$$\mathbb{P}(A_k) \; = \; \frac{|\mathbf{DR}_{b,k}||\mathbf{DR}_{r,n-k}||\mathbf{C}_{n,k}|}{|\mathbf{DR}_{b+r,n}|} = \binom{n}{k} \frac{b^k r^{n-k}}{(b+r)^n},$$

o, equivalentemente,

$$\mathbb{P}(A_k) \; = \; \binom{n}{k} p^k (1-p)^{n-k}, \qquad k = 0, 1, \dots, n,$$

dove $p = \frac{b}{b+r}$ è la probabilità di estrarre una pallina bianca in una singola estrazione. $\square$

OSSERVAZIONE. Consideriamo lo spazio di probabilità $(\Omega, \mathbb{P})$ con $\Omega = \{0, 1, 2, \dots, n\}$ e $\mathbb{P}$ data da

$$\mathbb{P}(\{k\}) \; = \; \binom{n}{k} p^k (1-p)^{n-k}, \qquad k = 0, 1, \dots, n.$$

Si noti che $\mathbb{P}$ è effettivamente una probabilità. Infatti, per la formula del binomio di Newton vale che

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

$\mathbb{P}$ non è una probabilità uniforme. $\mathbb{P}$ si chiama ***probabilità binomiale***.

[3] La terza scelta consiste nello scegliere in quali estrazioni/posizioni sono uscite le bianche: ci sono in tutto n estrazioni/posizioni a disposizione, dobbiamo sceglierne k. Quando nel metodo delle scelte successive una scelta riguarda la *posizione*, si procede in questo modo. Si considera l'insieme $I_n = \{1, 2, \dots, n\}$ delle n posizioni e l'insieme $\mathbf{C}_{n,k}$ delle combinazioni di k elementi di I_n. Allora ogni combinazione, ossia ogni sottoinsieme di cardinalità k di I_n, identifica k posizioni, e viceversa. Ad esempio, consideriamo il caso $n = 4$ e $k = 2$. Sia dunque $I_4 = \{1, 2, 3, 4\}$ l'insieme delle 4 posizioni a disposizione, così che $\mathbf{C}_{4,2}$ è l'insieme delle sei combinazioni:

$$\{1,2\}, \quad \{1,3\}, \quad \{1,4\}, \quad \{2,3\}, \quad \{2,4\}, \quad \{3,4\}.$$

Notiamo allora che ogni combinazione individua due posizioni, e viceversa.