Settimo modulo: Probabilità e statistica
Obiettivi

1. Saper conteggiare il numero totale di scelte in uno schema ad albero
2. Saper risolvere problemi con permutazioni, disposizioni e combinazioni
3. Saper calcolare media, varianza, frequenze relative ed assolute di un assegnato insieme di dati
4. Saper risolvere semplici problemi relativi alla probabilità di verificarsi di un evento

1 Elementi di analisi combinatoria

Si dice che n oggetti distinti sono disposti in un *allineamento* quando sono collocati in n posti numerati da 1 a n.
Dati n e k numeri naturali, si definisce il *coefficiente binomiale* per $n \geq 0$ e $0 \leq k \leq n$ nel seguente modo

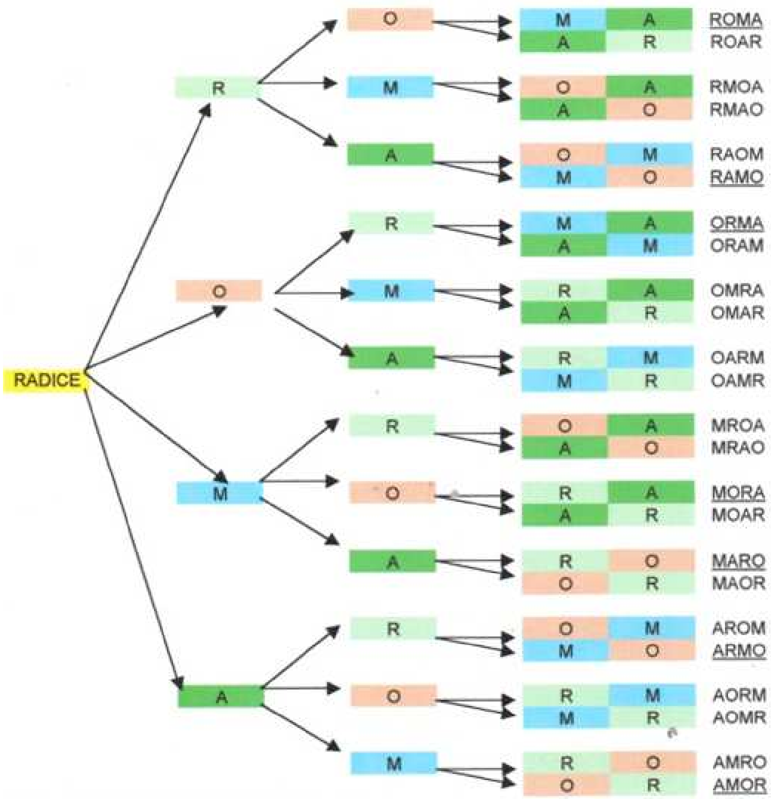
$$\boxed{\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n \cdot (n-1) \ldots (n-k+1)}{k!}}$$

con $n! = n \cdot (n-1) \cdot (n-2) \ldots \cdot 2 \cdot 1$

Le *permutazioni semplici di n oggetti* sono allineamenti di n elementi distinti che si distinguono esclusivamente per l'ordine degli oggetti.
Il numero delle permutazioni di n oggetti distinti è dato da

$$\boxed{P_n = n! = n \cdot (n-1) \cdot \ldots \cdot 3 \cdot 2 \cdot 1}$$

Esempio 1.1 Quanti sono gli anagrammi, anche non significativi, della parola ROMA?
Il diagramma ad albero è il seguente



Gli anagrammi, anche non significativi, sono $P_4 = 4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$.

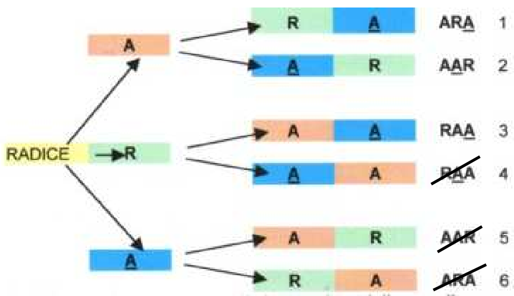
Osservazione 1.1 Si possono avere anche *permutazioni con ripetizione*, ovvero permutazioni di n oggetti non necessariamente distinti.
Se gli n oggetti sono divisi in due gruppi, uno di k elementi e l'altro di $(n-k)$ elementi, il numero delle permutazioni con ripetizione è dato da

$$\boxed{P^*_{k,k-1} = \frac{n!}{k!(n-k)!}}$$

Osservazione 1.2 Se gli n oggetti sono suddivisi in più gruppi k_1, k_2, ... k_h, il numero delle permutazioni con ripetizione è dato da

$$\boxed{P^*_{k_1 k_2 \ldots k_h} = \frac{n!}{k_1! k_2! \ldots k_h!}}$$

Esempio 1.2 Quanti sono gli anagrammi, anche non significativi, della parola ARA?
Il diagramma ad albero è il seguente

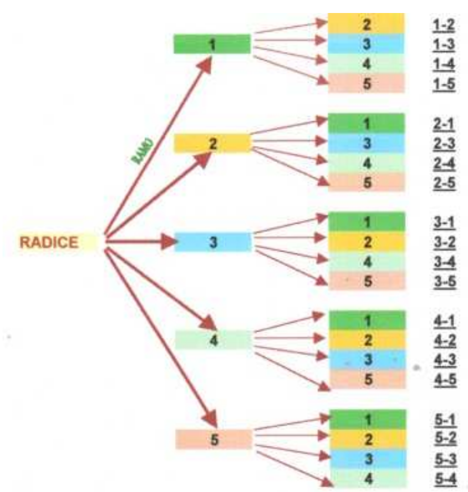


Gli anagrammi, anche non significativi, sono $P^*_{2,1} = \frac{3!}{2!1!} = \frac{3 \cdot 2 \cdot 1}{2 \cdot 1 \cdot 1} = 3$.

Le *disposizioni semplici di n oggetti di classe k* sono allineamenti di k oggetti distinti tali che ogni raggruppamento differisce dagli altri o per l'ordine o per gli oggetti contenuti.
Il numero di disposizioni semplici di n elementi di classe k sono

$$\boxed{D_{n,k} = n \cdot (n-1) \cdot \ldots (n-k+1)}$$

Esempio 1.3 Assegnati i cinque numeri $\{1, 2, 3, 4, 5\}$, quanti numeri di due cifre si possono formare?
Il diagramma ad albero è il seguente

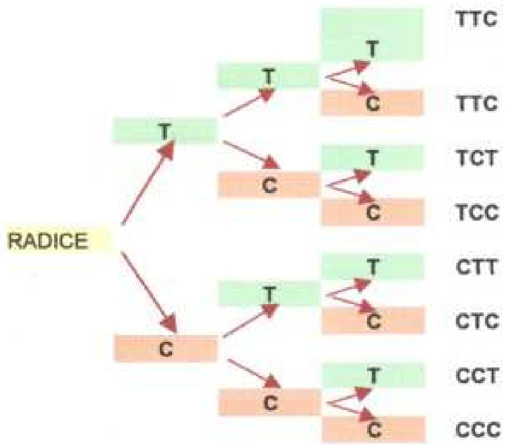


Le coppie di numeri che si possono formare (differiscono per numero e ordine) sono $D_{5,2} = 5 \cdot 4 = 20$.

Osservazione 1.3 Si hanno *disposizioni con ripetizione* di n oggetti di classe k quando si formano raggruppamenti di k elementi che differiscono tra loro o per l'ordine o per gli oggetti contenuti o per la ripetizione.
Il numero di disposizioni con ripetizione di n elementi di classe k sono

$$\boxed{D^*_{n,k} = n^k}$$

Esempio 1.4 Lanciando tre volte una moneta (T, C), quali sono i risultati possibili?
Il diagramma ad albero è il seguente

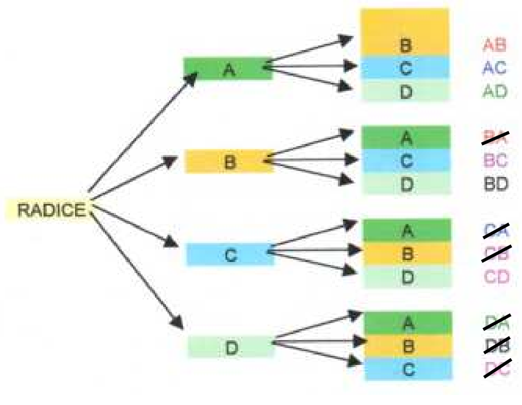


I risultati sono superiori al numero delle disposizioni semplici perchè nel risultato sono ripetuti alcuni elementi.
Il numero delle disposizioni con ripetizione sono $D^*_{2,3} = 2^3 = 8$.

Le *combinazioni semplici di n oggetti di classe k* sono raggruppamenti di k oggetti comunque scelti fra gli n in modo tale che ogni raggruppamenti differisca dagli altri per almeno un oggetto contenuto (attenzione: non si tiene conto dell'ordine!).
Il numero di combinazioni semplici di n elementi di classe k sono

$$\boxed{C_{n,k} = \left(\begin{array}{c} n \\ k \end{array}\right) = \frac{n \cdot (n-1) \ldots (n-k+1)}{k!}}$$

Esempio 1.5 In quanti modi una squadra di due persone (doppio di tennis), viene scelta tra quattro persone (giocatori)?
Il diagramma ad albero è il seguente

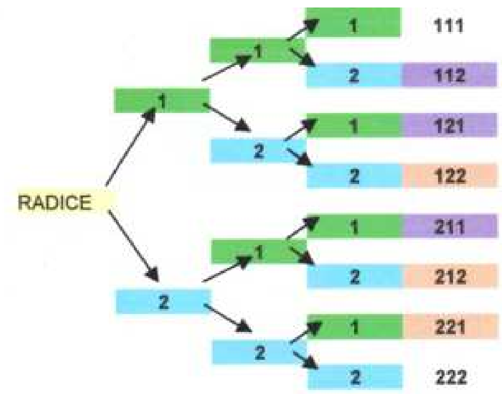


Il numero di coppie che si possono formare sono $C_{4,2} = \left(\begin{array}{c} 4 \\ 2 \end{array}\right) = \frac{4 \cdot 3}{2!} = \frac{12}{2} = 6$.

Osservazione 1.4 Si hanno *combinazioni con ripetizione* di n oggetti di classe k quando si formano raggruppamenti di k elementi che differiscono tra loro o per l'ordine o per la ripetizione.
Il numero di combinazioni con ripetizione di n elementi di classe k sono

$$\boxed{C^*_{n,k} = \binom{n+k-1}{k} = \frac{n\cdot(n+1)\ldots(n+k-1)}{k!}}$$

Esempio 1.6 Dati due numeri (1,2), quanti numeri di 3 cifre si possono formare?
Il diagramma ad albero è il seguente



Attenzione: non dobbiamo considerare l'ordine!
Quindi le combinazioni (viola) 112, 121 e 211 sono identiche tra loro; altrettanto le combinazioni (rosa) 122, 221 e 212.
Il numero combinazioni con ripetizioni che si possono formare sono
$C^*_{2,3} = \binom{4}{3} = \frac{2\cdot3\cdot4}{3!} = \frac{24}{6} = 4.$

2 Elementi di statistica descrittiva

Indici di posizione centrale

Una media è un valore opportunamente scelto e compreso fra il minimo e il massimo dei dati. In tutti i casi, la *media è un numero che ne sintetizza molti e consente di averne una visione unitaria*, ovviamente nascondendo la molteplicità dei dati da cui è ottenuta. Ad esempio il reddito medio delle famiglie italiane è un valore unico, utile per fare confronti con altre nazioni o con periodi passati, ma non evidenzia che i redditi sono molto diversi e molte famiglie sono al di sotto della soglia della sopravvivenza, mentre altre possiedono beni in grande quantità.

La *moda* è il valore cui corrisponde la massima frequenza.
Esempio 2.1 La sequenza di numeri 5, 6, 8, 8, 8, 12, 12, 14 ha moda 8.
Esempio 2.2 In un cantiere lo stipendio medio mensile dei quattro apprendisti è 600 euro, dei venti operai è 1.000 euro, del capocantiere 2.000 euro. Quindi, 600 euro ha frequenza 4, 1.000 euro ha frequenza 20 e 2.000 euro ha frequenza 1. La moda è 1.000 euro e sintetizza efficacemente la paga media dei dipendenti.

La *mediana* è il valore che occupa il posto di mezzo, quando i dati sono disposti in ordine crescente.
Esempio 2.3 I voti di Pierino sono, in ordine crescente: 4, 5, 5, 6, 7, 8, 9. Il voto che occupa il posto di mezzo, la mediana, è 6.

Dati n valori $x_1, x_2, ...x_n$ si dice *media aritmetica* (o semplicemente media) il valore che si ottiene dividendo la loro somma per il loro numero n; indicando con $\overline{M}$ la media aritmetica, si ha:

$$\boxed{\overline{M} = \frac{x_1+x_2+\ldots+x_n}{n}}$$

Esempio 2.4 la media aritmetica degli stipendi in euro riportati nell'esempio 2.2 è
$\overline{M} = \frac{4 \cdot 600 + 20 \cdot 1000 + 1 \cdot 2000}{25} = \frac{24400}{25} = 976$

Dati n valori $x_1, x_2, ...x_n$ ed n pesi $p_1, p_2, ...p_n$ si dice *media ponderata* il valore che si ottiene attribuendo a ciascun valore il suo peso prima di farne la media; indicando con M_p la media ponderata, si ha:

$$\boxed{M_p = \frac{x_1 \cdot p_1 + x_2 \cdot p_2 + ... + x_n \cdot p_n}{p_1 + p_2 + ... + p_n}}$$

Esempio 2.5 Per superare un esame uno studente deve sostenere una prova pratica, una prova scritta e una prova orale e ottenere una media superiore a 60. La prova pratica è meno importante di quella scritta, la quale, a sua volta, è meno importante di quella orale; esse hanno pesi 1, 2 e 3. Se uno studente merita 78 nella prova pratica, 44 nella scritta e 66 nella prova orale, la sua media ponderata è: $M_p = \frac{78 \cdot 1 + 44 \cdot 2 + 66 \cdot 3}{1+2+3} = \frac{364}{6} = 60,67$, per cui ha superato l'esame.

Dati n valori $x_1, x_2, ...x_n$ si dice *media quadratica* la radice quadrata della media aritmetica dei loro quadrati; indicando con M_q la media quadratica, si ha:

$$\boxed{M_q = \sqrt{\frac{x_1^2 + x_2^2 + ... + x_n^2}{n}}}$$

Esempio 2.6 Si vogliono sostituire tre tubi di raggio rispettivamente 2 cm, 3 cm e 4 cm con tre tubi di uguale raggio in modo che la portata complessiva resti inalterata. Quale deve essere il loro raggio?
Detta x la misura in cm del raggio incognito, deve essere:
$3\pi x^2 = \pi 2^2 + \pi 3^2 + \pi 4^2$ e quindi $x = \sqrt{\frac{2^2+3^2+4^2}{3}} \cong 3,11$.

Dati n valori $x_1, x_2, ...x_n$, si dice *media geometrica* la radice n-esima del loro prodotto; indicando con M_g la media geometrica, si ha:

$$\boxed{M_g = \sqrt[n]{x_1 \cdot x_2 \cdot ... \cdot x_n}}$$

Esempio 2.7 Un trasformatore rende l'81%, un altro il 64%. Se si applicano in serie il rendimento complessivo è pari al prodotto dei due rendimenti. La loro media geometrica: $\sqrt{0,81 \cdot 0,64} = 0,72 = 72\%$ ci dice quale rendimento uguale dovrebbero avere i due trasformatori per lasciare inalterato il loro rendimento complessivo.

Dati n valori $x_1, x_2, ...x_n$, si dice *media armonica* l'inverso della media aritmetica dei loro inversi; indicando con M_a la media armonica, si ha:

$$\boxed{M_a = \frac{1}{\frac{\frac{1}{x_1} + \frac{1}{x_2} + ... + \frac{1}{x_n}}{n}} = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + ... + \frac{1}{x_n}}}$$

Esempio 2.8 Percorro 21 Km alla velocità di 30 Km/h e altri 21 Km alla velocità di 70 Km/h. Qual è la velocità media?
La velocità media, in Km/h, risulta $v_m = \frac{2s}{t} = \frac{2s}{t_1+t_2} = \frac{2s}{\frac{s}{v_1} + \frac{s}{v_2}} = \frac{2}{\frac{1}{v_1} + \frac{1}{v_2}} = \frac{2}{\frac{1}{30} + \frac{1}{70}} = 42$

Indici di dispersione

Le medie riassumono in un unico valore il fenomeno studiato, ma non forniscono alcuna informazione sulla sua variabilità.
Ai fini di una descrizione sintetica ma significativa, è necessario definire dei parametri che indichino la dispersione dei dati ovvero la loro maggiore o minore concentrazione attorno a un valore medio.

La più immediata misura della variabilità è il *campo di variazione*, cioè la differenza fra il massimo e il minimo dei valori osservati.
Esempio 2.9 Dati i valori 161, 163, 170, 170, 173, 175, 178 rappresentanti le altezze in cm di sette

persone, il loro campo di variazione è di 17 cm.

In genere però il campo di variazione, che tiene conto soltanto dei due valori estremi e non è influenzato in alcun modo da quelli intermedi, costituisce una misura troppo rozza della variabilità.

Si definisce *scarto semplice medio* da una media M la media aritmetica dei valori assoluti degli scarti da M, ovvero

$$\boxed{s.s.m. = \frac{|x_1-M|+|x_2-M|+...+|x_n-M|}{n}}$$

Esempio 2.10 Per i valori dell'esempio 2.9, la media aritmetica delle altezze è 170 cm e lo scarto semplice medio è $s.s.m. = \frac{|161-170|+|163-170|+|170-170|+|170-170|+173-170|+|175-170|+|178-170|}{7} = \frac{32}{7} \cong 4,57$.

Si dice *scarto quadratico medio* la radice quadrata della media aritmetica degli scarti di ciascuna modalità dalla media aritmetica elevati al quadrato, ovvero

$$\boxed{s.q.m. = \sqrt{\frac{(x_1-M)^2+(x_2-M)^2+...+(x_n-M)^2}{n}}}$$

Esempio 2.11 Per i valori dell'esempio 2.9, lo scarto quadratico medio rispetto alla media aritmetica (170 cm) è
$s.q.m. = \sqrt{\frac{(161-170)^2+(163-170)^2+(170-170)^2+(170-170)^2+(173-170)^2+(175-170)^2+(178-170)^2}{7}} = \sqrt{\frac{228}{7}} \cong 5,71$.

Si dice *varianza* la media aritmetica degli scarti di ciascuna modalità dalla media aritmetica elevati al quadrato (lo scarto quadratico medio al quadrato), ovvero

$$\boxed{\sigma = \frac{(x_1-M)^2+(x_2-M)^2+...+(x_n-M)^2}{n}}$$

Esempio 2.12 Per i valori dell'esempio 2.9, la varianza rispetto alla media aritmetica (170 cm) è
$\sigma = \frac{(161-170)^2+(163-170)^2+(170-170)^2+(170-170)^2+(173-170)^2+(175-170)^2+(178-170)^2}{7} = \frac{228}{7} \cong 32,57$.

3 Elementi di probabilità

La *probabilità* che l'evento E si verifichi è uguale al quoziente fra il numero di casi in cui l'evento E accade e il numero di casi che sono possibili:

p(E) = numero di casi favorevoli al verificarsi di E / numero dei casi possibili

Esempio 3.1 Qual è la probabilità di estrarre un re di denari da un mazzo di 40 carte?
Casi favorevoli: 1
Casi possibili: 40

$$p(\text{esca un re di denari})=\frac{1}{40}$$

Esempio 3.2 Qual è la probabilità di estrarre una carta di picche da un mazzo di 52 carte?
Casi favorevoli: 13
Casi possibili: 52

$$p(\text{esca una carta di picche})=\frac{13}{52} = \frac{1}{4}$$

Esempio 3.3 Qual è la probabilità che esca 7 se si lancia un dado (cubico, con le facce numerate da 1 a 6)?
Casi favorevoli: 0
Casi possibili: 6

$$p(\text{esca una carta di picche})=\frac{0}{6}=0$$

Osservazione 3.1 Le situazioni illustrate riguardano eventi che hanno la stesa probabilità di accadere, ovvero sono equiprobabili (ad esempio il dado non deve essere truccato).

E' possibile calcolare la probabilità che si verifichi un evento legato in qualche modo al verificarsi di due eventi elementari E ed E'.

La *probabilità dell'evento complementare* di un evento dato vale:

$$p(non\ E)=1\text{-}p(E)$$

dove *non E* indica l'evento che corrisponde al fatto che l'evento E non si realizzi ed esclude che gli eventi E e *non E* si verifichino contemporaneamente.

Esempio 3.4 Qual è la probabilità che due persone che sono nate nello stesso anno bisestile compiano gli anni in giorni diversi?
La probabilità che, essendo la prima nata in un certo giorno ($p(E)=\frac{1}{366}$),la seconda sia nata in un altro è $p(non\ E)=1-\frac{1}{366}=\frac{365}{366}$.

La *probabilità* che si verifichino *due eventi equiprobabili e indipendenti uno dall'altro* corrisponde al *prodotto logico* $E \cap E'$, ovvero al prodotto delle probabilità dei singoli eventi.

Esempio 3.5 Qual è la probabilità di ottenere 12 tirando contemporaneamente due dadi?
L'unico modo per ottenere 12 è che tanto il primo dado quanto il secondo diano 6. Sono due eventi indipendenti l'uno dall'altro. La probabilità del primo è $\frac{1}{6}$, come quella del secondo. Quindi la probabilità che esca 12 è $p(12)=\frac{1}{6}\cdot\frac{1}{6}=\frac{1}{36}$.

Esempio 3.6 Qual è la probabilità, se si estraggono successivamente due carte da un mazzo di 40 carte e si rimette la prima al suo posto, di estrarre prima il 7 di spade e poi una carta di bastoni?
Sono due eventi indipendenti l'uno dall'altro poichè la carta estratta per prima viene rimessa nel mazzo. La probabilità di estrarre il 7 di spade è $\frac{1}{40}$, mentre quella di estrarre una carta di bastoni è $\frac{10}{40}$. La probabilità che i due eventi accadano insieme è $p(E \cap E')=\frac{1}{40}\cdot\frac{10}{40}=\frac{10}{1600}=\frac{1}{160}$.

Osservazione 3.2 Cosa sarebbe successo se la prima carta estratta non fosse stata rimessa nel mazzo? L'insieme delle carte da cui estrarre la seconda carta sarebbe stato formato solo da 39 carte e la probabilità di estrarre una carta di bastoni sarebbe stata condizionata da quanto successo nella prima estrazione: *i due eventi non sarebbero stati indipendenti.*
Il calcolo delle probabilità che escano il 7 di spade e una carta di bastoni vale il prodotto della probabilità di $E=\{$*esce il 7 di spade* $\}$ per la probabilità di $E'=\{$*esce una carta di bastoni* $\}$ *condizionata dal verificarsi di E.*
Quindi $p(E \cap E')=p(E)\cdot p(E'/E)=\frac{1}{40}\cdot\frac{10}{39}=\frac{10}{1560}=\frac{1}{156}$.

La *probabilità* che si verifichi *almeno uno degli eventi elementari coinvolti* corrisponde alla *somma logica* $E \cup E'$ dei due eventi, ovvero la loro probabilità è la somma delle singole probabilità.

Esempio 3.7 Se due dadi vengono lanciati contemporaneamente, qual è la probabilità che esca 11?
La probabilità che esca almeno una delle coppie (5,6) e (6,5) è la somma delle singole probabilità e vale $p(11)=\frac{1}{36}+\frac{1}{36}=\frac{1}{18}$.