PARTE 4

INDICE

Statistica descrittiva	pag. 1
Probabilità	pag. 15
Probabilità condizionata	pag. 21
Esercizi relativi	pag. 27
Calcolo combinatorio ed esempi/esercizi	pag. 35
SCHEMA DI RICAPITOLAZIONE DEL CALCOLO COMBINATORIO	pag. 59
APPENDICE: PERCENTUALI (FARE MOLTO BENE GLI ESERCIZI)	pag. 60
APPENDICE: PERCENTUALI - TRATTAZIONE PIÙ APPROFONDITA (LEGGERE)	pag. 65

(Ultima pagina: pag. 82)

PARTE 4

STATISTICA

STATISTICA DESCRITTIVA: Sintesi e propietà principali di tantissimi dati, che si presentano (talvolta) anche in (apparente) disordine.

STATISTICA INDUTTIVA o INFERENZIALE: imparare dall'esperienza e descrivere modelli, e risalire dalle proprietà di un campione alle corrispondenti proprietà dell'intera popolazione.

STATISTICA DESCRITTIVA

Tecniche per la raccolta, l'organizzazione e la descrizione delle osservazioni, sopratutto quando si tratta di raccogliere dati. La moltitudine dei dati, in un primo momento, può sembrare disordinata: allora bisogna raccogliere elementi di sintesi per descrivere il comportamento di tantissimi dati attraverso FONDAMENTALI PROPRIETÀ. Queste proprietà possono essere descritte in diversi modi: ad esempio, tabelle, grafici per punti, spezzate, istogrammi, diagrammi a torta, eccetera ....

Gli elementi rispetto ai quali si fa la sintesi sono in generale dei NUMERI, un elenco di tantissimi numeri.

Supponiamo dunque che, nei nostri esperimenti, si vogliano studiare alcune proprietà dei dati raccolti. Indichiamo con $X = \{x_1, x_2, \ldots, x_n\}$ l'insieme dei dati. In Statistica, si dice anche che l'insieme X è una serie di dati. [N.B.: Qui LE RIPETIZIONI CONTANO.]

Ora diamo delle definizioni e facciamo degli esempi.
Considerato un dato x_i appartenente all'insieme dei dati $X = \{x_1, x_2, ..., x_n\}$, diremo che la frequenza assoluta del dato x_i è il numero di volte in cui compare il dato x_i (perché, nelle nostre osservazioni, i dati si possono ripetere), mentre la frequenza relativa del dato x_i è il rapporto fra la frequenza assoluta del dato x_i e il numero totale dei dati. La frequenza percentuale del dato x_i è il prodotto della frequenza relativa del dato x_i per il numero 100.

Esempio: Il numero dei battiti cardiaci al minuto registrati nell'arco di 10 giorni ad una persona è

73, 72, 73, 74, 76, 70, 71, 72, 72, 74

Dato	Frequenza assoluta	Frequenza relativa	Frequenza percentuale
70	1	1/10	10 %
71	1	1/10	10 %
72	3	3/10	30 %
73	2	2/10 = 1/5	20 %
74	2	2/10 = 1/5	20 %
76	1	1/10	10 %
Somme:	10 (numero dei dati)	1	100 %

Osserviamo che la somma delle varie frequenze assolute è il numero totale dei dati, la somma delle frequenze relative è 1 $\left(= \frac{\text{numero dei dati}}{\text{numero dei dati}}\right)$, e la somma delle frequenze percentuali è 100 $(= 1 \cdot 100)$.

La <u>moda</u> della nostra serie di dati $x_1, x_2, \ldots, x_n$ è il numero che si ripete più spesso, cioè il numero avente <u>maggiore</u> <u>frequenza</u> (assoluta, relativa, o percentuale), purché ~~la~~ la sua frequenza assoluta sia ≥ 2.

Nella nostra serie di dati considerata nella pagina precedente 73, 72, 73, 74, 76, 70, 71, 72, 72, 74 la moda è 72, perché è il dato che si ripete più spesso (3 volte): ha frequenza assoluta uguale a 3, mentre tutti gli altri dati hanno frequenza assoluta uguale a 1 o 2.

N.B.: In una serie di dati ci possono essere anche due o più mode. Per esempio, nella seguente serie di 5 dati 1, 2, 2, 3, 3 ci sono due mode: il 2 e il 3, perché entrambi si ripetono 2 volte, mentre l'altro dato (1) compare una volta sola.

La <u>MEDIANA</u> di una serie di dati $x_1, x_2, \ldots, x_n$, una volta che si dispongono i dati in ordine <u>crescente</u>, è:

<u>il numero centrale</u>, se n è dispari;

<u>la media</u> $\frac{1}{2} \cdot \left(x_{\frac{n}{2}} + x_{\frac{n}{2}+1}\right)$, cioè "la media aritmetica dei due valori centrali", se n è pari.

Nel nostro caso, serie di dati 73, 72, 73, 74, 76, 70, 71, 72, 72, 74, mettendo i dati in ordine crescente si ottiene

70, 71, 72, 72, 72, 73, 73, 74, 74, 76

Siccome i dati sono in numero pari ($n = 10$), allora la <u>mediana</u> è la media aritmetica fra il <u>5°</u> e il <u>6° dato</u>, DOPO AVERE DISPOSTO I DATI IN ORDINE CRESCENTE.

Nel nostro caso, la mediana è $\frac{72+73}{2} = 72,5$.

Se invece la serie di dati fosse stata 1, 1, 2, 3, 3, allora n = numero dei dati = 5, e quindi la mediana sarebbe stato il dato centrale, cioè il terzo dato, ossia 2.

La <u>media</u> della nostra serie di dati si definisce come

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

Se i dati <u>distinti</u> della serie di dati sono $\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_p$ (con $p \leq n$), che compaiono con frequenze assolute rispettive $F_1, F_2, \dots, F_p$, si può scrivere – equivalentemente –

$$\bar{x} = \frac{\tilde{x}_1 F_1 + \tilde{x}_2 F_2 + \dots + \tilde{x}_p F_p}{n} = \frac{1}{n} \sum_{j=1}^{p} F_j \tilde{x}_j =$$

$$= \tilde{x}_1 f_1 + \tilde{x}_2 f_2 + \dots + \tilde{x}_p f_p = \sum_{j=1}^{p} f_j \tilde{x}_j$$, dove le f_j sono le frequenze relative dei dati.

Esempio: Consideriamo sempre la nostra serie di dati. Nel fare la media, è indifferente se li disponiamo in ordine crescente oppure no, perché nell'addizione "cambiando l'ordine degli addendi, il risultato non cambia mai". Disponiamo i dati in ordine crescente, solamente "per vederli meglio"

70, 71, 72, 72, 72, 73, 73, 74, 74, 76

La media è: $\bar{x} = \frac{70+71+72+72+72+73+73+74+74+76}{10} = 72{,}7$

che si esprime anche (con le frequenze assolute)

$$\bar{x} = \frac{70 \cdot 1 + 71 \cdot 1 + 72 \cdot 3 + 73 \cdot 2 + 74 \cdot 2 + 76 \cdot 1}{10}$$

In una serie di dati, spesso è molto importante analizzare non solo la loro media, ma anche la loro "distribuzione" o "dispersione" rispetto alla media, cioè "di quanto i dati si allontanano dalla media" o se "sono molto vicini alla media".

Ad esempio, si può vedere che le due serie di 5 dati
$x_1 = 27$ $x_2 = 28$ $x_3 = 28$ $x_4 = 28$ $x_5 = 29$
$y_1 = 26$ $y_2 = 27$ $y_3 = 28$ $y_4 = 29$ $y_5 = 30$
hanno la stessa media (infatti

$$\frac{27+28+28+28+29}{5} = \frac{26+27+28+29+30}{5} = 28),$$

ma si può osservare – anche intuitivamente – che tra le due serie di dati c'è una grande differenza di "comportamento": infatti i dati x_i sono <u>molto</u> più addensati intorno al 28 che i dati y_i

Data la serie di dati $x_1, \dots, x_n$, si chiama <u>VARIANZA</u> (più precisamente VARIANZA CAMPIONARIA) della serie il numero

$$S^2 = S_X^2 = \frac{(x_1-\bar{x})^2+(x_2-\bar{x})^2+\dots+(x_n-\bar{x})^2}{n-1} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

che valuta la "distanza media al quadrato" dei dati dalla media, cioè la loro "dispersione".
Se i dati assumono p valori distinti $\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_p$ con rispettive frequenze assolute $F_1, F_2, \dots, F_p$, si può scrivere, equivalentemente

$$S^2 = S_X^2 = \frac{F_1(\tilde{x}_1-\bar{x})^2+F_2(\tilde{x}_2-\bar{x})^2+\dots+F_p(\tilde{x}_p-\bar{x})^2}{n-1} = \frac{1}{n-1}\sum_{j=1}^{p}F_j(\tilde{x}_j-\bar{x})^2$$

La quantità $\sqrt{S_X^2} = S_X$ (oppure $\sqrt{S^2} = S$) si chiama <u>SCARTO QUADRATICO MEDIO</u> (CAMPIONARIO) oppure <u>DEVIAZIONE STANDARD</u> (CAMPIONARIA) (è una specie di "distanza", in un certo senso ....
Si chiama <u>DEVIANZA</u> il numero $\sum_{i=1}^{n}(x_i-\bar{x})^2 = \sum_{j=1}^{p}F_j(\tilde{x}_j-x)^2$
(senza dividere per $n-1$; quindi si ha: DEVIANZA $= (n-1)\cdot$ VARIANZA;
VARIANZA $= \frac{\text{DEVIANZA}}{n-1}$, riferendosi alla varianza campionaria)

Noi considereremo varianze CAMPIONARIE

N.B.: Se, nella definizione di varianza, si divide per n invece che per $n-1$, allora si parlerà di VARIANZA DI POPOLAZIONE. Quindi, per questo tipo di varianza, si avrà: DEVIANZA $= n \cdot$ VARIANZA; VARIANZA $= \dfrac{\text{DEVIANZA}}{n}$.

Esempio: $\begin{cases} x_1=27 & x_2=28 & x_3=28 & x_4=28 & x_5=29 & \bar{x}=28 \\ y_1=26 & y_2=27 & y_3=28 & y_4=29 & y_5=30 & \bar{y}=28 \end{cases}$

$$s_X^2=\frac{(27-28)^2+(28-28)^2+(28-28)^2+(28-28)^2+(29-28)^2}{4}=$$

$$=\frac{1\cdot(27-28)^2+3\cdot(28-28)^2+1\cdot(29-28)^2}{4}=\frac{1+0+1}{4}=\frac{2}{4}=\boxed{\frac{1}{2}}$$

$\frac{1}{2}$ è la varianza delle x_i, mentre la devianza delle x_i è

$$1\cdot(27-28)^2+3\cdot(28-28)^2+1\cdot(29-28)^2=1+0+1=\boxed{2}$$

$$s_Y^2=\frac{(26-28)^2+(27-28)^2+(28-28)^2+(29-28)^2+(30-28)^2}{4}=$$

$=\frac{4+1+0+1+4}{4}=\boxed{\frac{10}{4}=\frac{5}{2}=2{,}5}$ è la varianza delle y_i, mentre la devianza delle y_i è

$$(26-28)^2+(27-28)^2+(28-28)^2+(29-28)^2+(30-28)^2=$$

$$=4+1+0+1+4=\boxed{10}$$

Notare che la varianza delle y_i è molto più grande di quella delle x_i.

I rispettivi scarti quadratici medi sono

$$s_X=\sqrt{\frac{1}{2}}=\frac{1}{\sqrt{2}}=\frac{\sqrt{2}}{2}\cong 0{,}7071$$

$$s_Y=\sqrt{2{,}5}\cong 1{,}5811$$

Ora introdurremo i concetti di covarianza, codevianza e il coefficiente di correlazione tra due serie di dati $X=\{x_1,x_2,\dots,x_n\}$ ed $Y=\{y_1,y_2,\dots,y_n\}$, per stabilire che tipo di legame, correlazione, ..., eccetera, ..., c'è tra la prima e la seconda serie di dati. N.B.: n è lo stesso per X e per Y.

Date due serie di dati $x_1, \dots, x_n$; $y_1, \dots, y_n$ aventi lo stesso numero di dati n, si chiama covarianza (più precisamente, covarianza campionaria) il numero

$$Cov(X,Y) = \frac{(x_1-\bar{x})\cdot(y_1-\bar{y})+\dots+(x_n-\bar{x})\cdot(y_n-\bar{y})}{n-1} = \boxed{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})\cdot(y_i-\bar{y})}$$

, mentre la codevianza è data dal numero $(x_1-\bar{x})\cdot(y_1-\bar{y})+\dots+(x_n-\bar{x})\cdot(y_n-\bar{y}) = \sum_{i=1}^{n}(x_i-\bar{x})\cdot(y_i-\bar{y})$ N.B.: Qui NON compaiono i quadrati!

Si ha: CODEVIANZA = (n-1)·COVARIANZA ;

$$\text{COVARIANZA} = \frac{\text{CODEVIANZA}}{n-1}$$

Se, nella definizione di covarianza, si sostituisce $n-1$ con n, allora si parlerà di COVARIANZA DI POPOLAZIONE, e si avrà: CODEVIANZA = n · COVARIANZA;

$$\text{COVARIANZA} = \frac{\text{CODEVIANZA}}{n}$$

Noi considereremo covarianze campionarie.

Esempio: $\begin{cases} x_1=27 & x_2=28 & x_3=28 & x_4=28 & x_5=29 & \bar{x}=28 \\ y_1=26 & y_2=27 & y_3=28 & y_4=29 & y_5=30 & \bar{y}=28 \end{cases}$

$$Cov(X,Y) = \frac{(27-28)\cdot(26-28)+(28-28)\cdot(27-28)+(28-28)\cdot(28-28)+(28-28)\cdot(29-28)+(29-28)\cdot(30-28)}{4} =$$

$$= \frac{(-1)\cdot(-2)+0\cdot(-1)+0\cdot 0+0\cdot 1+1\cdot 2}{4} = \frac{4}{4} = \boxed{1},$$

mentre la codevianza tra X ed Y è data da

$$(27-28)\cdot(26-28)+(28-28)\cdot(27-28)+(28-28)\cdot(28-28)+(28-28)\cdot(29-28)+(29-28)\cdot(30-28) =$$

$$= (-1)\cdot(-2)+0\cdot(-1)+0\cdot 0+0\cdot 1+1\cdot 2 = \boxed{4}.$$

Ora introduciamo una quantità (r) che sarà il COEFFICIENTE DI CORRELAZIONE tra due serie di dati $X=\{x_1, x_2, \dots, x_n\}$, $Y=\{y_1, y_2, \dots, y_n\}$, che esprime che legame c'è tra X ed Y, se sono ben correlate oppure no, e anche quanta correlazione c'è. Si definisce il COEFFICIENTE DI CORRELAZIONE TRA X ED Y il numero

$$r = \frac{Cov(X,Y)}{s_X \cdot s_Y} = \frac{Cov(X,Y)}{\sqrt{s_X^2 \cdot s_Y^2}} = \frac{Cov(X,Y)}{\sqrt{Varianza(X) \cdot Varianza(Y)}} =$$

(Ove la varianza può essere presa indifferentemente campionaria o di popolazione, ed s_X, s_Y indicano i rispettivi scarti quadratici medi di X e di Y)

$$= \frac{Codevianza(X,Y)}{\sqrt{Devianza(X) \cdot Devianza(Y)}}$$

[Si può vedere che queste definizioni sono fra loro equivalenti, ma non facciamo la dimostrazione]

Esempio: $\begin{cases} x_1=27 \quad x_2=28 \quad x_3=28 \quad x_4=28 \quad x_5=29 \quad \bar{x}=28 \\ y_1=26 \quad y_2=27 \quad y_3=28 \quad y_4=29 \quad y_5=30 \quad \bar{y}=28 \end{cases}$

Abbiamo visto che $Varianza(X)=\frac{1}{2}$, $Varianza(Y)=\frac{5}{2}$, $Covarianza(X,Y)=1$, $Devianza(X)=2$, $Devianza(Y)=10$, $Codevianza(X,Y)=4$. Si ha:

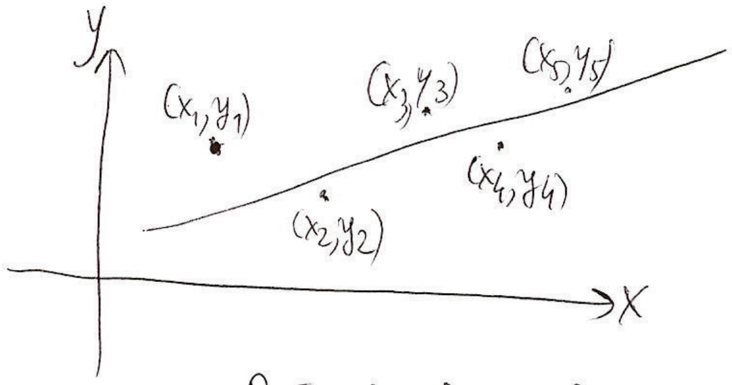
$$r = \frac{Cov(X,Y)}{\sqrt{Varianza(X) \cdot Varianza(Y)}} = \frac{1}{\sqrt{\frac{1}{2} \cdot \frac{5}{2}}} = \frac{1}{\sqrt{\frac{5}{4}}} =$$

$$= \frac{1}{\frac{\sqrt{5}}{2}} = \frac{2}{\sqrt{5}} = \frac{2\sqrt{5}}{5} \simeq 0,894.$$ Si ha anche

$$r = \frac{Codevianza(X,Y)}{\sqrt{Devianza(X) \cdot Devianza(Y)}} = \frac{4}{\sqrt{2 \cdot 10}} = \frac{4}{\sqrt{20}} = \frac{4}{\sqrt{4} \cdot \sqrt{5}} =$$

$$= \frac{\cancel{4}^{2}}{\cancel{2}\sqrt{5}} = \frac{2}{\sqrt{5}} = \frac{2\sqrt{5}}{5}$$ si ritrova lo stesso risultato.

Il fatto che $r \simeq 0{,}894$ sta a indicare che le nostre due serie di dati sono <u>molto</u> correlate fra di loro. Infatti si può vedere che IL COEFFICIENTE DI CORRELAZIONE È UNA QUANTITÀ SEMPRE COMPRESA TRA 1 E −1, E UN VALORE MOLTO VICINO A 1 OPPURE A −1 STA A INDICARE UN'OTTIMA CORRELAZIONE, MENTRE UN VALORE MOLTO VICINO A 0 INDICA UNA CATTIVA CORRELAZIONE.



Facciamo un disegno puramente "indicativo"

RETTA DI REGRESSIONE
FRA DUE SERIE DI DATI
$X = \{x_1, \dots, x_n\}$ ed $Y = \{y_1, \dots, y_n\}$ (n è lo stesso)

Nel piano cartesiano, mettiamo i punti $x_1, x_2, \dots, x_n$ sull'asse delle x e i punti $y_1, y_2, \dots, y_n$ sull'asse delle y: avremo quindi i punti del tipo $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ nel piano cartesiano. La retta che "approssima meglio i punti $(x_1, y_1), \dots (x_n, y_n)$", che "è più vicina ai punti $(x_1, y_1), \dots, (x_n, y_n)$", si chiama (andamento GLOBALE dei punti) RETTA DI REGRESSIONE.

Ora diamo l'equazione della retta di regressione (senza dimostrazione) [Per evitare il caso della retta verticale, supporremo che gli x_i non siano tutti quanti uguali a uno stesso numero]. Sarà del tipo $\boxed{y = mx + q}$, dove

$$m = \frac{cov(X, Y)}{S_X^2} \left(\text{cioè } \frac{cov(X, Y)}{\text{Varianza}(X)}\right) = \frac{\text{Codevianza}(X, Y)}{\text{Devianza}(X)}$$

(si può vedere che le due definizioni sono equivalenti, e che è indifferente, nel calcolo di m (coefficiente angolare della retta di regressione), considerare la (co)varianza campionaria (come facciamo) o la (co)varianza di popolazione (senza dimostrazione)), $\boxed{q = \bar{y} - m\bar{x}}$ (perché, cosa IMPORTANTISSIMA DA EVIDENZIARE, LA RETTA DI REGRESSIONE HA LA PROPRIETÀ DI PASSARE SEMPRE PER IL PUNTO $(\bar{x}, \bar{y})$)

Nel nostro esempio $x_1 = 27 \quad x_2 = x_3 = x_4 = 28 \quad x_5 = 29 \quad \bar{x} = 28$
$y_1 = 26 \quad y_2 = 27 \quad y_3 = 28 \quad y_4 = 29 \quad y_5 = 30 \quad \bar{y} = 28$

Varianza $(X) = \frac{1}{2}$, ~~Devianza~~ Devianza $(X) = 2$, Covarianza $(X, Y) = 1$,
Codevianza $(X, Y) = 4$

$$\boxed{m} = \frac{\text{Covarianza}(X, Y)}{\text{Varianza}(X)} = \frac{1}{\frac{1}{2}} = \frac{2}{1} = \boxed{2}$$

oppure

$$\boxed{m} = \frac{\text{Codevianza}(X, Y)}{\text{Devianza}(X)} = \frac{4}{2} = \boxed{2}$$

$$\boxed{q} = \bar{y} - m\bar{x} = 28 - 2 \cdot 28 = 28 - 56 = \boxed{-28}$$

Pertanto l' equazione della retta di regressione è

$$\boxed{y = 2x - 28}.$$

ESERCIZIO: In un gruppo di 5 adulti la somministrazione di dosi di un farmaco ha causato diminuzioni di pressione.

Dose (x_i)	7	12	15	20	22
Diminuzione di pressione (y_i)	10	18	20	25	25

Il farmaco è veramente efficace, o no?
Cioè, il coefficiente di correlazione associato alle x_i ed alle y_i è veramente molto vicino a 1 (o a -1), o no?
Intanto calcoliamo le rispettive medie $\bar{x}, \bar{y}$ delle x_i e delle y_i:

$$\bar{x} = \frac{7 + 12 + 15 + 20 + 22}{5} = \frac{76}{5} = 15,2$$

$$\bar{y} = \frac{10 + 18 + 20 + 25 + 25}{5} = \frac{98}{5} = 19,6$$

Facciamo allora la seguente tabella (che va fatta in TUTTI gli esercizi di statistica di questo tipo):

x_i	$\bar{x}$	y_i	$\bar{y}$	$x_i-\bar{x}$	$(x_i-\bar{x})^2$	$y_i-\bar{y}$	$(y_i-\bar{y})^2$	$(x_i-\bar{x})\cdot(y_i-\bar{y})$
7	15,2	10	19,6	-8,2	67,24	-9,6	92,16	78,72
12	15,2	18	19,6	-3,2	10,24	-1,6	2,56	5,12
15	15,2	20	19,6	-0,2	0,04	0,4	0,16	-0,08
20	15,2	25	19,6	4,8	23,04	5,4	29,16	25,92
22	15,2	25	19,6	6,8	46,24	5,4	29,16	36,72
					$\Sigma=146,8$ (devianza(X))		$\Sigma=153,2$ (devianza(Y))	$\Sigma=146,4$ codevianza(X,Y)

s_X^2 = Varianza$(X) = \frac{146,8}{4} = 36,7$ deviazione standard $=\sqrt{s_X^2} \cong 6,06$

s_Y^2 = varianza$(Y) = \frac{153,2}{4} = 38,3$ deviazione standard $=\sqrt{s_Y^2} \cong 6,19$

$\operatorname{cov}(X,Y)$ = covarianza $= \frac{146,4}{4} = 36,6$

Coefficiente di correlazione $= r = \dfrac{\text{Codevianza}(X,Y)}{\sqrt{\text{Devianza}(X)\cdot\text{Devianza}(Y)}} =$

$= \dfrac{146,4}{\sqrt{146,8\cdot 153,2}} \cong \dfrac{146,4}{149,966} \cong 0,976$ molto vicino a 1. Pertanto c'è un'ottima correlazione, e il farmaco funziona molto bene!

Retta di regressione $y = mx + q$, dove $m = \dfrac{\text{Codevianza}(X,Y)}{\text{Devianza}(X)} =$

$= \dfrac{146,4}{146,8} \cong 0,997$ $q = \bar{y} - m\bar{x} \cong 19,6 - 0,997\cdot 15,2 \cong 4,4456$

Ora, riconsideriamo gli y_i, $y_1=10$, $y_2=18$, $y_3=20$, $y_4=y_5=25$, e facciamo la TABELLA DELLE FREQUENZE

Dato y_j	Frequenza assoluta F_j	Frequenza relativa f_j	Frequenza percentuale
10	1	1/5	20 %
18	1	1/5	20 %
20	1	1/5	20 %
25	2	2/5	40 %
Somme:	5 (numero dei dati)	1	100 %

La moda degli y_i (10 - 18 - 20 - 25 - 25) è 25, perché è il dato che si ripete più spesso, cioè il dato che ha maggiore frequenza (assoluta, relativa o percentuale). Per quanto riguarda la mediana, notiamo che gli y_i sono già disposti in ordine crescente, i dati sono 5 (numero dispari) e quindi la mediana degli y_i è uguale al dato centrale, cioè 20.

Per quanto riguarda gli x_i (7 - 12 - 15 - 20 - 22), notiamo che i valori sono tutti quanti diversi fra di loro, e quindi ciascuno di essi ha frequenza assoluta 1, frequenza relativa 1/5 e frequenza percentuale del 20%. I dati x_i sono disposti in ordine crescente, pertanto la mediana corrisponde al valore "centrale" 15 in quanto il numero dei dati ($n = 5$ in questo caso) è DISPARI. Non c'è una vera e propria moda, perché ciascun dato compare una volta soltanto, e quindi nessun dato si ripete più di una volta.

STATISTICA DESCRITTIVA

RICAPITOLANDO:

Raggruppare (molti) dati e studiarne le principali proprietà significative in modo sintetico.

MEDIA

MODA

MEDIANA

FREQUENZA ASSOLUTA, RELATIVA E %

VARIANZA

SCARTO QUADRATICO MEDIO o DEVIAZIONE STANDARD

una serie di dati

COVARIANZA

COEFFICIENTE DI CORRELAZIONE

RETTA DI REGRESSIONE

due serie di dati

↓

C'è correlazione, oppure è un frutto del caso?

PROBABILITÀ (prevedere il futuro, non predire!!!)

Adesso esaminiamo più da vicino il concetto di frequenza relativa, e introduciamo il concetto di probabilità. (COLLEGAMENTO PROFONDO)

ESEMPIO: Consideriamo la seguente serie di 36 dati che corrisponde al risultato della somma dei punteggi ottenuti con 2 dadi DISTINTI:

2, 3, 3, 4, 4, 4, 5, 5, 5, 5, 6, 6, 6, 6, 6, 7, 7, 7, 7, 7, 7, 8, 8, 8, 8, 8, 9, 9, 9, 9, 10, 10, 10, 11, 11, 12 $\quad n = 36$

e facciamo la tabella delle frequenze assolute e relative.

~~Dato~~ (=punteggio)	Frequenze assolute	Frequenze relative
2 (1+1)	1	1/36
3 (1+2, 2+1)	2	1/18
4 (1+3, 2+2, 3+1)	3	1/12
5 (1+4, 2+3, 3+2, 4+1)	4	1/9
6 (1+5, 2+4, 3+3, 4+2, 5+1)	5	5/36
7 (1+6, 2+5, 3+4, 4+3, 5+2, 6+1)	6	1/6
8 (2+6, 3+5, 4+4, 5+3, 6+2)	5	5/36
9 (3+6, 4+5, 5+4, 6+3)	4	1/9
10 (4+6, 5+5, 6+4)	3	1/12
11 (5+6, 6+5)	2	1/18
12 (6+6)	1	1/36
Somme:	36	1

Ora, lanciamo due dadi distinti, non truccati, e consideriamo l'esito (o evento) E_7 = "la somma dei punteggi dei due dadi è 7".

In letteratura, la PROBABILITÀ di E_7 è il seguente rapporto:

$$P(E_7) = \frac{\text{numero di casi favorevoli ad } E_7}{\text{numero di } \underline{\text{tutti}} \text{ i casi possibili lanciando due dadi}}$$

[Più in generale, si chiama probabilità di un evento E, e la si indica con P(E), il seguente rapporto:

$$P(E) = \frac{\text{numero di casi favorevoli ad } E}{\text{numero di casi possibili}}$$

purché - naturalmente - tutti i casi siano "equamente possibili" (per intenderci, nei dadi, purché i due dadi non siano truccati)]

Ora, nel nostro esempio, i casi possibili sono:

(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), $\boxed{(1, 6)}$

(2, 1), (2, 2), (2, 3), (2, 4), $\boxed{(2, 5)}$, (2, 6)

(3, 1), (3, 2), (3, 3), $\boxed{(3, 4)}$, (3, 5), (3, 6)

(4, 1), (4, 2), $\boxed{(4, 3)}$, (4, 4), (4, 5), (4, 6)

(5, 1), $\boxed{(5, 2)}$, (5, 3), (5, 4), (5, 5), (5, 6)

$\boxed{(6, 1)}$, (6, 2), (6, 3), (6, 4), (6, 5), (6, 6)

quindi $6 \times 6 = 6^2 = 36$ casi possibili

I casi messi in evidenza $\boxed{(1,6)}$, $\boxed{(2,5)}$, $\boxed{(3,4)}$, $\boxed{(4,3)}$, $\boxed{(5,2)}$, $\boxed{(6,1)}$ sono i 6 casi favorevoli ad E_7 perché in questi 6 casi (e solo in questi) la somma è 7. Quindi si ha:

$\boxed{P(E_7)} = \frac{6}{36} = \boxed{\frac{1}{6}}$. Si può vedere che in questo esempio LA PROBABILITÀ È ESATTAMENTE LA FREQUENZA RELATIVA.

Per esempio: abbiamo visto che la frequenza relativa del dato 11 (come somma dei punteggi dei due dadi) è $\frac{1}{18}$. Vediamo che anche la probabilità di E_{11} = "la somma dei punteggi dei due dadi è 11" è $\frac{1}{18}$, perché

$$P(E_{11}) = \frac{\text{numero di casi favorevoli ad } E_{11}}{\text{numero di tutti i casi possibili lanciando 2 dadi}} = \frac{2}{36} = \frac{1}{18};$$

infatti i casi favorevoli ad E_{11} sono solamente due ((5,6) e (6,5)). E così via ....

Altro esempio: in un'urna ci sono 11 palline non truccate, 5 bianche e 6 nere. Si estrae una pallina, e sia B = "pallina bianca", N = "pallina nera". Si ha:

$$P(B) = \frac{\text{numero di casi favorevoli a } B}{\text{numero totale dei casi possibili}} = \frac{5}{11}$$

$$P(N) = \frac{\text{numero di casi favorevoli a } N}{\text{numero totale dei casi possibili}} = \frac{6}{11}$$

Ora, in Medicina e in Farmacia (in particolare) ci si pone spesso il seguente <u>PROBLEMA</u>: Se un individuo risulta positivo ad un test su una malattia qual è la probabilità che sia veramente malato? E quella che sia sano?

A questo scopo introdurremo il concetto di <u>PROBABILITÀ CONDIZIONATA</u>.

ESPERIMENTO: insieme di azioni, il cui risultato non può essere predetto con certezza (p.es. il lancio di 1 o 2 dadi o il tiro a segno)

PROVA: è una singola esecuzione di un esperimento (p.es. un singolo lancio o tiro)

ESITO: Ogni possibile insieme di dati ottenuti da un esperimento, che ne definisce un risultato (p.es. la somma dei punti nel lancio di 2 dadi; oppure, nel tiro a segno, la distanza della freccia dal centro del bersaglio)

SPAZIO CAMPIONE o CAMPIONARIO Ω: l'insieme di tutti gli esiti di un esperimento (p.es. l'insieme di tutti i possibili punteggi nei 2 dadi; oppure l'insieme di tutte le possibili distanze della freccia dal centro del bersaglio)

PUNTO CAMPIONE: Ogni elemento di Ω.

$$\Omega = \{(1,1), (2,1), (1,2), (2,2), \dots (6,6)\}$$

caso discreto (finito o come $\mathbb{N}$) / $\Omega = [0, +\infty[$ caso continuo (intervallo)

EVENTO: Un sottoinsieme di Ω, cioè un insieme di esiti (p.es., nel lancio di 2 dadi, $E=\{(6,5),(5,6)\}$). oppure $E=\{1,3,5\}$ nel lancio di un ~~dado~~

GLI EVENTI SARANNO IDENTIFICATI CON GLI INSIEMI

sono i cosiddetti "atomi"

EVENTO ELEMENTARE: Se $a \in \Omega$, cioè a è un punto campione, allora $\{a\}$ è un evento elementare (p.es. $\{6,6\}$, nel lancio di 2 dadi)

Ω = tutto lo spazio = EVENTO CERTO

ϕ = insieme vuoto = EVENTO IMPOSSIBILE

- Due eventi A e B si dicono INCOMPATIBILI se non si possono verificare contemporaneamente, e si dirà $A \cap B = \phi$ (p.es. $A=\{$numero pari$\}$, $B=\{$numero dispari$\}$, nel lancio di un dado)

Incompatibili: insiemi disgiunti

Con $P(\Omega)$ indichiamo la totalità dei sottoinsiemi di Ω, ϕ ed Ω compresi.

$A \cap B$ si verificano sia A sia B

$A \cup B$ almeno uno tra A e B si verifica in latino, A vel B (non necessariamente aut) (nel senso che A e B non sono "esclusivi").

Una **PROBABILITÀ** è una funzione $P: \mathcal{P}(\Omega) \to [0,1]$ tale che: i) $\boxed{P(\Omega)=1}$ ii) se $A \cap B = \phi$ (cioè se A e B sono INCOMPATIBILI), allora $\boxed{P(A \cup B) = P(A) + P(B)}$

Da i) e ii) si ha: iii) $\boxed{P(\phi)=0}$

iv) se $A \subset B$ (cioè, quando si verifica A, si verifica anche B) allora $\boxed{P(A) \leq P(B)}$. (P è NON DECRESCENTE)

$A=\{3\}$ $B=\{$numero dispari$\}$

Sia Ω uno spazio campione finito. Si dice che Ω è equiprobabile se ogni evento elementare ha la stessa probabilità di verificarsi

Nel dado $P(\{1\}) = P(\{2\}) = 1/6$ perché $P(\Omega) = P(\{1,...,6\}) = 1$

Supposto di essere in uno spazio equiprobabile Ω, si definisce PROBABILITÀ DI UN EVENTO $E \subset \Omega$

$$\boxed{P(E) = \frac{\text{numero di casi favorevoli ad } E}{\text{numero di casi possibili}}}$$

(definizione classica di LAPLACE, 1812)

(Esempio: dado non truccato, monetina non truccata ...)

DEFINIZIONE DI

PROBABILITÀ CONDIZIONATA

Sia $P(B)>0$. Si definisce "probabilità condizionata di A dato B" la quantità

$$\boxed{P(A|B)=\frac{P(A\cap B)}{P(B)}=\frac{P(B\cap A)}{P(B)}}$$

Analogamente, se $P(A)>0$, si definisce

$$P(B|A)=\frac{P(A\cap B)}{P(A)}=\frac{P(B\cap A)}{P(A)}$$

DEFINIZIONE DI EVENTI INDIPENDENTI

Si dice che 2 eventi A, B sono INDIPENDENTI se

$$\boxed{P(A\cap B)=P(A)\cdot P(B)}$$ → (definizione)

[cioè se il verificarsi di uno dei 2 eventi non influisce sull'andamento dell'altro.]

→ risultato che proveremo tra poco

	prodotto	intersezione	DEFINIZIONE	
Indipendenti	$\cdot$	$\cap$	$P(A\cap B)=P(A)\cdot P(B)$	è la DEFINIZIONE di INDIPENDENZA
Incompatibili	$+$	$\cup$	$P(A\cup B)=P(A)+P(B)$	è un RISULTATO NON È LA DEFINIZIONE DI INCOMPATIBILITÀ
	somma	unione	PROPRIETÀ	

Ricordiamo che la definizione di incompatibilità tra A e B è:

$$\boxed{A\cap B=\emptyset}$$

Riprendiamo la <u>definizione</u> di eventi A e B INDIPENDENTI: A e B sono indipendenti se e solo se

$$(\nabla\nabla\nabla) \quad \boxed{P(A \cap B) = P(A) \cdot P(B)}$$

Sia $P(A) \neq 0$, $P(B) \neq 0$ (o equivalentemente $P(A) > 0$, $P(B) > 0$, perché la probabilità è sempre una quantità <u>non</u> <u>negativa</u>). Dall'uguaglianza $(\nabla\nabla\nabla)$, dividendo per $P(B)$ otteniamo

$$\frac{P(A \cap B)}{P(B)} = P(A) \quad , \text{ma } \frac{P(A \cap B)}{P(B)} = P(A|B) \text{ per la}$$

<u>DEFINIZIONE</u> di probabilità condizionata, e quindi si deduce

$$P(A|B) = P(A)$$

Quindi abbiamo provato che, se $P(A \cap B) = P(A) \cdot P(B)$ allora $P(A|B) = P(A)$. Rifacendo all'indietro tutti i passaggi che abbiamo fatto ora, si può vedere che, se $P(A|B) = P(A)$, allora $P(A \cap B) = P(A) \cdot P(B)$. Quindi si ha

$$\boxed{P(A \cap B) = P(A) \cdot P(B) \text{ se e solo se } P(A|B) = P(A)}$$

Analogamente si può vedere che

$$\boxed{P(A \cap B) = P(A) \cdot P(B) \text{ se e solo se } P(B|A) = P(B)}$$

Ma la scrittura $P(A|B) = P(A)$ vuol dire

P di A dato B uguale a P di A

cioè il verificarsi di B non influenza l'andamento di A, e così la scrittura $P(B|A) = P(B)$ sta a indicare che il verificarsi di A non influenza l'andamento di B.

Quindi, tenendo conto che, come già detto, la scrittura $P(A \cap B) = P(A) \cdot P(B)$ vuol dire che gli eventi A e B sono INDIPENDENTI, allora otteniamo il risultato che era

stato precedentemente annunciato:
due eventi A e B sono indipendenti se e solo se il verificarsi di uno dei due eventi non influisce sull' andamento dell' altro.

Adesso vediamo, per quanto riguarda la PROBABILITÀ CONDIZIONATA, che cosa giustifica la definizione $\boxed{P(A|B) = \frac{P(A\cap B)}{P(B)}}$? (con $P(B) \neq 0$)

Facciamo un esempio.
Lanciamo un dado non truccato (il nostro spazio campionario è $\Omega = \{1, 2, 3, 4, 5, 6\}$). Sia A = "esce un numero non più grande di 3", cioè $A = \{1, 2, 3\}$. Si ha:

$$P(A) = \frac{\text{numero dei casi favorevoli ad } A}{\text{numero di tutti i casi possibili}} = \frac{3}{6} = \frac{1}{2}.$$

Adesso ho una nuova informazione: so che si è verificato l' evento B = "è uscito un numero dispari", cioè $B = \{1, 3, 5\}$.
Adesso, senza pensare alla formula ma pensando solo alla definizione di probabilità, chi sarà la probabilità condizionata $P(A|B)$, P di A dato B, cioè la probabilità dell' evento A ma questta volta SAPENDO CHE SI È VERIFICATO L' EVENTO B? Se so questo, allora tutti i casi possibili sono 1, 3 e 5 (non sono più 1, 2, 3, 4, 5 e 6 perché ho la nuova informazione!). Allora chi saranno i casi favorevoli ad A? $A = \{1, 2, 3\}$ ma bisogna tenere conto che i casi possibili sono 1, 3 e 5 e quindi 2 è escluso. Quindi, in questo nuovo contesto, i casi favorevoli ad A sono 1 e 3,

e pertanto la probabilità condizionata $P(A|B)$, per la DEFINIZIONE DI PROBABILITÀ, sarà data da

$$\frac{\text{numero dei "nuovi" casi favorevoli ad } A \text{, visto che è uscito } B}{\text{numero dei "nuovi" casi possibili, cioè quelli di } B} =$$

$= \dfrac{2}{3}$ (Torniamo ora al "vecchio" spazio campionario $\Omega = \{1, 2, 3, 4, 5, 6\}$) (TRUCCO) $= \boxed{\dfrac{2}{6} \cdot \dfrac{6}{3}} =$

(TRUCCO) $= \dfrac{\frac{2}{6}}{\frac{3}{6}}$ (perché ora vogliamo esprimere il tutto in termini delle "originali" probabilità, dove il numero dei casi possibili è 6, perché è il numero di <u>tutte</u> le facce del dado di partenza)

Ma 3 è il numero dei casi favorevoli a B, quindi

$$P(B) = \frac{\text{numero dei casi favorevoli a } B}{\text{numero di tutti i casi possibili}} = \frac{3}{6}$$

Invece, 2 è il numero dei casi favorevoli ad A visto che si è verificato l'evento B. Questi casi favorevoli ad A - abbiamo visto - sono 1 e 3, quindi non sono <u>tutti</u> gli elementi dell'insieme A, ma sono esattamente gli elementi di A che <u>appartengono anche all'insieme B</u>, perché <u>si è verificato B</u> e quindi dobbiamo stare dentro l'insieme B (!).

Ma gli elementi di A **che** appartengono anche a B sono gli elementi che stanno contemporaneamente sia in A sia in B, cioè gli elementi di $(A \cap B)$. Allora si ha:

$$\frac{2}{6} = \frac{\text{numero dei casi favorevoli ad } A \cap B}{\text{numero di tutti i casi possibili}} = P(A \cap B),$$ per <u>DEFINIZIONE di prob.</u>

Allora $\boxed{P(A|B)} = \dfrac{\frac{2}{6}}{\frac{3}{6}} = \boxed{\dfrac{P(A \cap B)}{P(B)}}$ <u>ritroviamo la FORMULA della probabilità condizionata</u> (!)

APPLICAZIONE DELLA PROBABILITÀ CONDIZIONATA

ALBERO ... DI NATALE !!

ESTRAZIONI DI PALLINE

L'albero si fa quando non c'è reimbussolamento (Cioè le palline, una volta estratte, NON VENGONO RIMESSE nell'urna), e in questo caso entra in gioco la PROBABILITÀ CONDIZIONATA.

SENZA REIMBUSSOLAMENTO

Esempio: 7 palline: 3 bianche, 4 nere. B_1 = "bianca alla 1ª estrazione", N_1 = "nera alla 1ª estrazione", B_2 = "bianca alla 2ª estrazione", N_2 = "nera alla 2ª estrazione". Alla 1ª estrazione ci sono due possibilità: la 1ª pallina è bianca (B_1) oppure nera (N_1). Alla 2ª estrazione, bisogna distinguere 2 sottocasi: nel primo, la pallina uscita alla 1ª estrazione è bianca; nell'altro, la pallina uscita alla 1ª estrazione è nera.

1ª estrazione

$P(B_1) = \frac{3}{7}$ $\qquad$ $P(N_1) = \frac{4}{7}$

2ª estrazione

$P(B_2|B_1) = \frac{2}{6} = \frac{1}{3}$ $\qquad$ $P(N_2|B_1) = \frac{4}{6} = \frac{2}{3}$ $\qquad$ $P(B_2|N_1) = \frac{3}{6} = \frac{1}{2}$ $\qquad$ $P(N_2|N_1) = \frac{3}{6} = \frac{1}{2}$

Nella diramazione che parte da $P(B_1)$ si estrae la 2ª pallina: può essere bianca oppure nera. Ma l'uscita della pallina bianca alla 1ª estrazione influisce sulla 2ª estrazione, perché la 1ª pallina viene posta sul tavolo e non viene messa dentro l'urna. Quindi si deve parlare non di $P(B_2)$ e di $P(N_2)$ in assoluto, ma di $P(B_2|B_1)$ e di $P(N_2|B_1)$. Così pure, se alla 1ª estrazione esce una pallina nera, si deve parlare di $P(B_2|N_1)$ e di $P(N_2|N_1)$. Allora, quando si valuta la probabilità dell'evento intersezione $P(B_2 \cap B_1)$ (tutte e due le palline bianche), bisogna applicare la definizione di probabilità condizionata, ottenendo

$$P(B_2|B_1) = \frac{P(B_2 \cap B_1)}{P(B_1)}$$

, da cui

$$P(B_2 \cap B_1) = P(B_2|B_1) \cdot P(B_1) = \frac{1}{3} \cdot \frac{3}{7} = \frac{1}{7}$$

Similmente, si ha: $P(N_2|B_1) = \frac{P(N_2 \cap B_1)}{P(B_1)}$, da cui

$$P(N_2 \cap B_1) = P(N_2|B_1) \cdot P(B_1) = \frac{2}{3} \cdot \frac{3}{7} = \frac{2}{7}.$$

Inoltre, $P(B_2|N_1) = \frac{P(B_2 \cap N_1)}{P(N_1)}$, da cui $P(B_2 \cap N_1) = P(B_2|N_1) \cdot P(N_1) = \frac{1}{2} \cdot \frac{4}{7} = \frac{2}{7}$, mentre $P(N_2|N_1) = \frac{P(N_2 \cap N_1)}{P(N_1)}$, da cui

$$P(N_2 \cap N_1) = P(N_2|N_1) \cdot P(N_1) = \frac{1}{2} \cdot \frac{4}{7} = \frac{2}{7}.$$

Le due estrazioni sono <u>dipendenti</u>, perché l'esito della prima <u>influisce</u> sull'esito della seconda. Lo possiamo vedere anche da un punto di vista matematico, in quanto si ha, ad esempio, $P(B_2 \cap B_1) \neq P(B_2) \cdot P(B_1)$: infatti $P(B_2)$, presa "da sola", è uguale a $\frac{3}{7}$ (3 palline bianche su 7 in totale), e quindi $P(B_2 \cap B_1) = \frac{1}{7}$, mentre $P(B_2) \cdot P(B_1) = \frac{3}{7} \cdot \frac{3}{7} = \frac{9}{49}$.

Se invece <u>c'è</u> reimbussolamento, cioè la pallina estratta <u>viene</u> rimessa nell'urna, allora le due estrazioni sono INDIPENDENTI, e non si fa l'albero, ma si calcola la probabilità dell'evento intersezione come prodotto delle 2 probabilità, applicando direttamente la definizione di eventi indipendenti. È:

$$P(B_2 \cap B_1) = P(B_2) \cdot P(B_1) = \frac{3}{7} \cdot \frac{3}{7} = \frac{9}{49}$$

("Da sola", $P(B_2) = P(B_1) = \frac{3}{7}$, $P(N_2) = P(N_1) = \frac{4}{7}$)

$$P(N_2 \cap B_1) = P(N_2) \cdot P(B_1) = \frac{4}{7} \cdot \frac{3}{7} = \frac{12}{49}$$

$$P(B_2 \cap N_1) = P(B_2) \cdot P(N_1) = \frac{3}{7} \cdot \frac{4}{7} = \frac{12}{49}$$

$$P(N_2 \cap N_1) = P(N_2) \cdot P(N_1) = \frac{4}{7} \cdot \frac{4}{7} = \frac{16}{49}.$$

ESERCIZIO PROB. CONDIZIONATA

Ci sono 17 palline, di cui 8 bianche e 9 nere. Si fanno due estrazioni, e si considera sia il caso senza reimbussolamento sia il caso con reimbussolamento. Indichiamo con: B_1 = "pallina bianca alla 1ª estrazione"; N_1 = "pallina nera alla 1ª estrazione"; B_2 = "pallina bianca alla 2ª estrazione"; N_2 = "pallina nera alla 2ª estrazione". Le palline e le estrazioni non sono "truccate".

a) Calcolare $P(B_1)$

b) Calcolare $P(B_2 \cap B_1)$, sia senza reimbussolamento sia con reimbussolamento

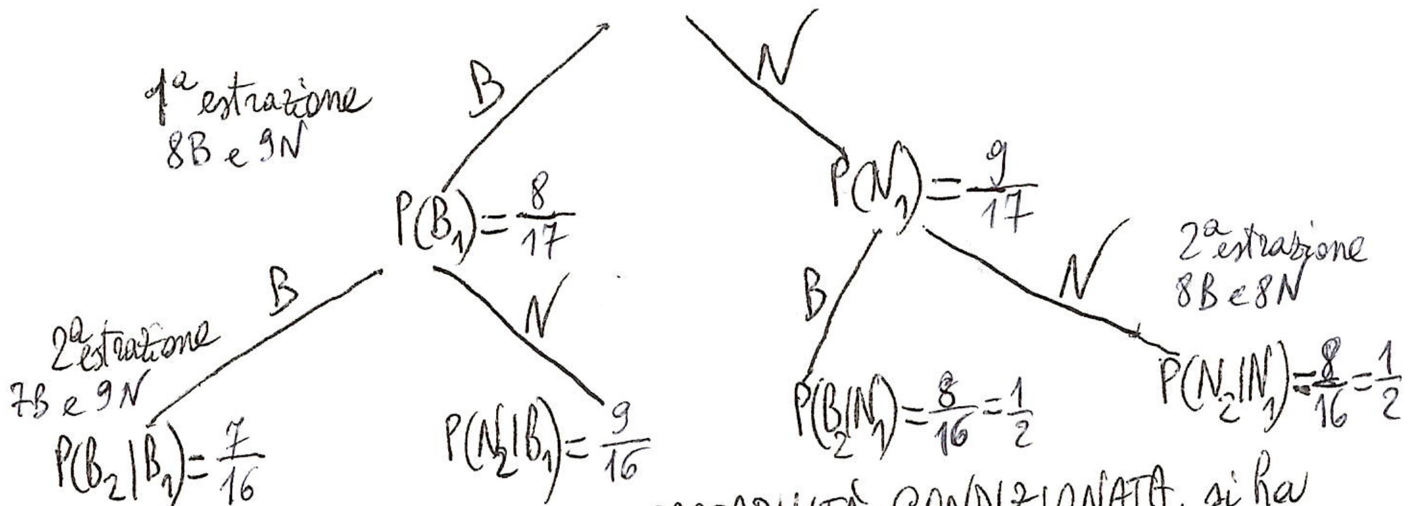
a) Considerando la definizione "classica" di probabilità

$$P(E) = \frac{\text{numero dei casi favorevoli ad } E}{\text{numero di tutti i casi possibili}}$$

, allora si ha

$$P(B_1) = \frac{\text{numero delle palline bianche}}{\text{numero totale delle palline}} = \frac{8}{17}$$

b) Consideriamo dapprima il caso senza reimbussolamento, e allora facciamo l'<u>albero</u>. (la prima pallina, una volta estratta, <u>NON</u> viene rimessa nell'urna)



Per la DEFINIZIONE DI PROBABILITÀ CONDIZIONATA, si ha $P(B_2|B_1) = \frac{P(B_2 \cap B_1)}{P(B_1)}$. Moltiplicando entrambi i membri di questa uguaglianza per $P(B_1)$, si ottiene

$$P(B_2 \cap B_1) = P(B_2|B_1) \cdot P(B_1) = \frac{7}{\cancel{16}_2} \cdot \frac{\cancel{8}}{17} = \frac{7}{34}$$

Ora vediamo il caso con reimbussolamento (cioè la prima pallina, una volta estratta, VIENE rimessa nell'urna). In questo caso le estrazioni sono indipendenti, e pertanto gli eventi B_2 e B_1 sono indipendenti. Quindi si ha, per la DEFINIZIONE di eventi indipendenti:

$P(B_2 \cap B_1) = P(B_2) \cdot P(B_1) = P(B_1) \cdot P(B_1)$ (le estrazioni sono indipendenti, e in questo caso alla 2ª estrazione si "ricomincia da capo", e quindi – in termini di probabilità – la 2ª estrazione "si comporta" come la 1ª) $= \frac{8}{17} \cdot \frac{8}{17} = \frac{64}{289}$

ESERCIZIO PROB. CONDIZIONATA

Ci sono 14 palline, di cui 5 bianche e 9 nere. Si fanno due estrazioni, e si considera sia il caso senza reimbussolamento sia il caso con reimbussolamento. Indichiamo con: B_1 = "pallina bianca alla 1ª estrazione"; N_1 = "pallina nera alla 1ª estrazione"; B_2 = "pallina bianca alla 2ª estrazione"; N_2 = "pallina nera alla 2ª estrazione". Le palline e le estrazioni non sono "truccate".

a) Calcolare $P(N_1)$

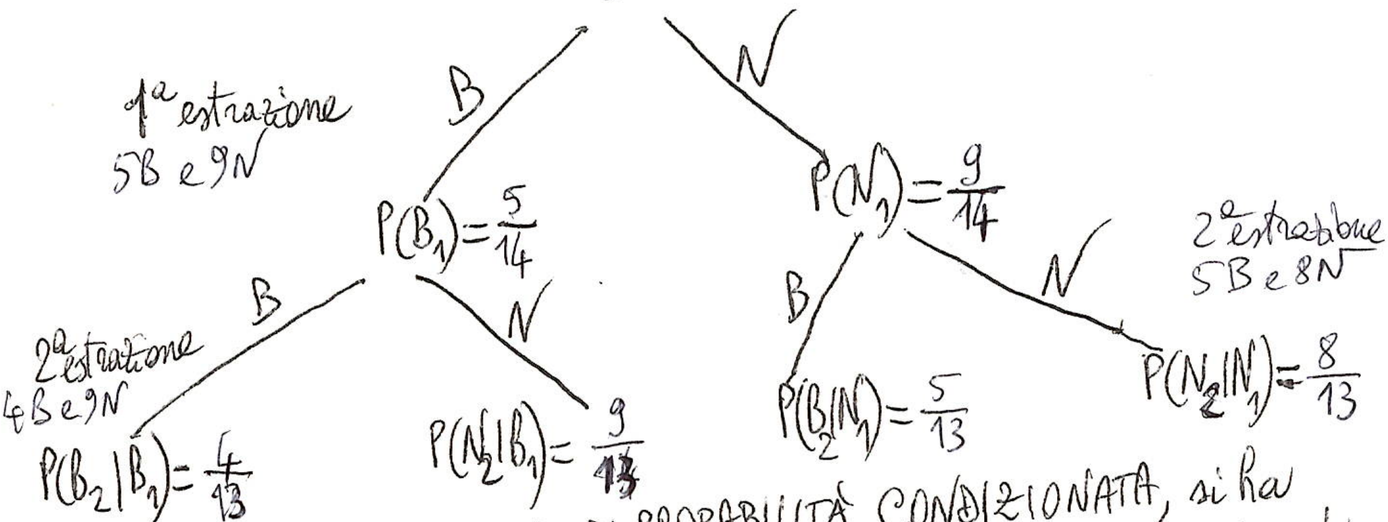
b) Calcolare $P(N_2 \cap N_1)$, sia senza reimbussolamento sia con reimbussolamento

a) Considerando la definizione "classica" di probabilità

$$P(E) = \frac{\text{numero dei casi favorevoli ad } E}{\text{numero di tutti i casi possibili}}$$, allora si ha

$$P(N_1) = \frac{\text{numero delle palline nere}}{\text{numero totale delle palline}} = \frac{9}{14}$$

b) Consideriamo dapprima il caso senza reimbussolamento, e allora facciamo l' <u>albero</u>. (la prima pallina, una volta estratta, <u>NON</u> viene rimessa nell'urna)



Per la DEFINIZIONE DI PROBABILITÀ CONDIZIONATA, si ha
$P(N_2|N_1) = \frac{P(N_2 \cap N_1)}{P(N_1)}$. Moltiplicando entrambi i membri di questa uguaglianza per $P(N_1)$, si ottiene

$$P(N_2 \cap N_1) = P(N_2|N_1) \cdot P(N_1) = \frac{\cancel{8}\,4}{13} \cdot \frac{9}{\cancel{14}\,7} = \frac{36}{91}$$

Ora vediamo il caso con reimbussolamento (cioè la prima pallina, una volta estratta, VIENE rimessa nell'urna). In questo caso le estrazioni sono indipendenti, e pertanto gli eventi N_2 e N_1 sono indipendenti. Quindi si ha, per la DEFINIZIONE di eventi indipendenti:

$P(N_2 \cap N_1) = P(N_2) \cdot P(N_1) = P(N_1) \cdot P(N_1)$ (le estrazioni sono indipendenti, e in questo caso alla 2ª estrazione si "ricomincia da capo", e quindi – in termini di probabilità – la 2ª estrazione "si comporta" come la 1ª) $= \frac{9}{14} \cdot \frac{9}{14} = \frac{81}{196}$

ESERCIZIO PROB. CONDIZIONATA

Ci sono 23 palline, di cui 12 bianche e 11 nere. Si fanno due estrazioni, e si considera sia il caso senza reimbussolamento sia il caso con reimbussolamento. Indichiamo con: B_1 = "pallina bianca alla 1ª estrazione"; N_1 = "pallina nera alla 1ª estrazione"; B_2 = "pallina bianca alla 2ª estrazione"; N_2 = "pallina nera alla 2ª estrazione". Le palline e le estrazioni non sono "truccate".

a) Calcolare $P(B_1)$

b) Calcolare $P(N_2 \cap B_1)$, sia senza reimbussolamento sia con reimbussolamento

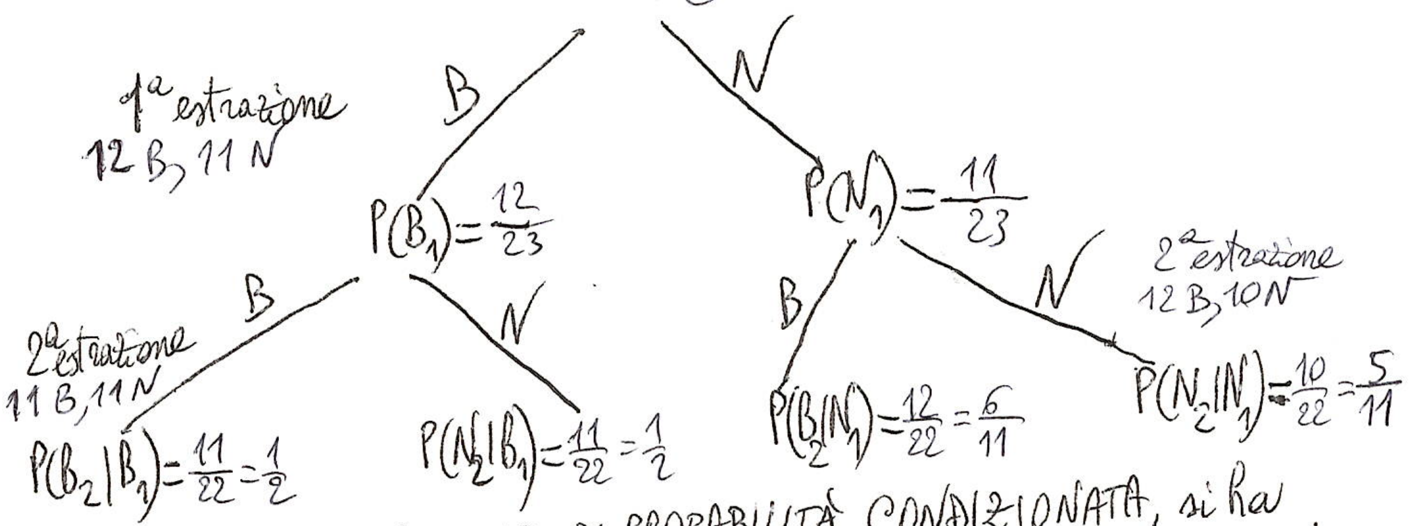
a) Considerando la definizione "classica" di probabilità

$$P(E) = \frac{\text{numero dei casi favorevoli ad } E}{\text{numero di tutti i casi possibili}}$$

, allora si ha

$$P(B_1) = \frac{\text{numero delle palline bianche}}{\text{numero totale delle palline}} = \frac{12}{23}$$

b) Consideriamo dapprima il caso senza reimbussolamento, e allora facciamo l'albero. (la prima pallina, una volta estratta, NON viene rimessa nell'urna)



Per la DEFINIZIONE DI PROBABILITÀ CONDIZIONATA, si ha $P(N_2|B_1) = \frac{P(N_2 \cap B_1)}{P(B_1)}$. Moltiplicando entrambi i membri di questa uguaglianza per $P(B_1)$, si ottiene

$$P(N_2 \cap B_1) = P(N_2|B_1) \cdot P(B_1) = \frac{1}{\cancel{2}} \cdot \frac{\overset{6}{\cancel{12}}}{23} = \frac{6}{23}$$

Ora vediamo il caso con reimbussolamento (cioè la prima pallina, una volta estratta, VIENE rimessa nell'urna). In questo caso le estrazioni sono indipendenti, e pertanto gli eventi N_2 e B_1 sono indipendenti. Quindi si ha, per la DEFINIZIONE di eventi indipendenti:

$P(N_2 \cap B_1) = P(N_2) \cdot P(B_1) = P(N_1) \cdot P(B_1)$ (le estrazioni sono indipendenti, e in questo caso alla 2ª estrazione si "ricomincia da capo", e quindi – in termini di probabilità – la 2ª estrazione "si comporta" come la 1ª) $= \frac{11}{23} \cdot \frac{12}{23} = \frac{132}{529}$

ESERCIZIO (APPLICAZIONE DELLA PROBABILITÀ CONDIZIONATA ALLA MEDICINA E ALLA FARMACIA)

Una certa malattia colpisce il 5% di una popolazione. Un test individua la malattia (cioè risulta essere positivo sugli individui malati) nel 90% dei casi, e il test risulta positivo su soggetti sani nel 5% dei casi. Qual è la probabilità che l'individuo risulti SANO, pur essendo POSITIVO al test? (Verrà 0,5135, quindi avremo ancora speranze...)

IMPOSTAZIONE DEL PROBLEMA IN TERMINI DI PROBABILITÀ (CONDIZIONATA)

$$P(\text{malato}) = \frac{5}{100} = \frac{1}{20}, \text{ e quindi}$$

(Se il 5% sono malati, allora il 95% sono sani)

$$P(\text{sano}) = 1 - P(\text{malato}) = 1 - \frac{1}{20} = \frac{19}{20} = \frac{95}{100}$$

$$P(\text{positivo} \mid \text{malato}) = \frac{90}{100} = \frac{9}{10}$$

$$P(\text{positivo} \mid \text{sano}) = \frac{5}{100} = \frac{1}{20}$$

Da determinare: $P(\text{sano} \mid \text{positivo}) = \dfrac{P(\text{positivo} \cap \text{sano})}{P(\text{positivo})}$.

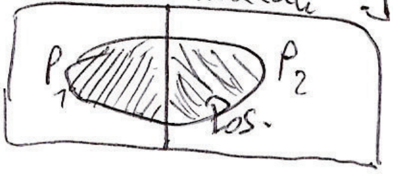
Intanto, calcoliamo $P(\text{positivo} \cap \text{sano}) = P(\text{sano} \cap \text{positivo})$

N.B (TRUCCO): Noi conosciamo $P(\text{positivo} \mid \text{sano}) \underset{\text{prob. condiz.}}{\overset{\text{definiz.}}{=}} \dfrac{P(\text{positivo} \cap \text{sano})}{P(\text{sano})}$

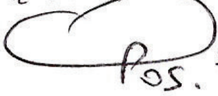
Quindi $\dfrac{1}{20} = \dfrac{P(\text{positivo} \cap \text{sano})}{\frac{19}{20}}$, da cui $P(\text{positivo} \cap \text{sano}) = \dfrac{1}{20} \cdot \dfrac{19}{20} = \dfrac{19}{400}$

Ora calcoleremo $P(\text{positivo})$

Per calcolare P(positivo), ci vuole un TRUCCO (che è LO STESSO PER TUTTI GLI ESERCIZI DI QUESTO TIPO): PENSARE A UNA "PARTIZIONE"!



Ω = l'intero spazio, l'intera popolazione



Pos. = individui positivi

P_1 //// : individui sani e positivi — positivo $\cap$ sano

P_2 |||| : individui malati — positivo $\cap$ malato

Si ha: Ω = "sano" $\cup$ "malato" (l'unione è disgiunta, perché un individuo o è sano o è malato), quindi "positivo" = "positivo $\cap$ sano" $\cup$ "positivo $\cap$ malato", ove l'unione è disgiunta, e quindi i due eventi "positivo $\cap$ sano" e "positivo $\cap$ malato" sono INCOMPATIBILI. Allora, per le proprietà degli eventi incompatibili che abbiamo visto precedentemente, si ha

$$P(\text{"positivo"}) = P(\text{"positivo} \cap \text{sano"}) + P(\text{"positivo} \cap \text{malato"})$$

$P(\text{"positivo} \cap \text{sano"}) = \frac{19}{400}$ $P(\text{"positivo} \cap \text{malato"})$ la calcoliamo con la stessa tecnica con cui si è trovato $P(\text{"positivo} \cap \text{sano"})$.

Sappiamo che $P(\text{positivo}|\text{malato}) \underset{\text{condiz.}}{\overset{\text{prob.}}{=}} \frac{P(\text{"positivo} \cap \text{malato"})}{P(\text{malato})}$. Quindi

$\frac{9}{10} = \frac{P(\text{positivo} \cap \text{malato})}{\frac{1}{20}}$, perché sappiamo che $P(\text{positivo}|\text{malato}) = \frac{9}{10}$ e che $P(\text{malato}) = \frac{1}{20}$. Allora si ha

$P(\text{positivo} \cap \text{malato}) = \frac{1}{20} \cdot \frac{9}{10} = \frac{9}{200} = \frac{18}{400}$; $P(\text{positivo}) =$
$= P(\text{positivo} \cap \text{sano}) + P(\text{positivo} \cap \text{malato}) = \frac{19}{400} + \frac{18}{400} = \frac{37}{400}$;

$$P(\text{sano}|\text{positivo}) = \frac{P(\text{positivo} \cap \text{sano})}{P(\text{positivo})} = \frac{\frac{19}{400}}{\frac{37}{400}} = \frac{19}{37} \cong \boxed{0{,}5135}$$

CALCOLO COMBINATORIO

Nello studio dell'analisi di dati, e in tantissimi campi della Matematica e delle varie Scienze, ci si occupa spesso dello studio dei metodi, dei modi per raggruppare un numero finito di elementi.

Il Calcolo Combinatorio è lo studio dei diversi metodi con i quali si può raggruppare un numero finito di elementi, e del modo di contare il numero di tutti i possibili raggruppamenti che si possono fare per ciascun metodo (nel senso che ogni metodo ha le sue ben precise regole: se si cambia metodo, le regole cambieranno)

Siano n e k due numeri interi positivi, e consideriamo un insieme di n elementi distinti (per esempio $A_n = \{a_1, a_2, \dots, a_n\}$)

Da questo insieme, consideriamo i raggruppamenti di k elementi che si possono fare.

Per esempio, se $n = 5$, $A_5 = \{a_1, a_2, a_3, a_4, a_5\}$ e $k = 3$, allora $a_1 a_2 a_4$ è un raggruppamento di 3 elementi, e $a_2 a_3 a_5$ è un altro raggruppamento di 3 elementi.

Si dicono <u>DISPOSIZIONI</u> DI <u>n</u> OGGETTI a <u>k</u> a <u>k</u> i raggruppamenti che si possono formare a partire dagli n oggetti, in modo tale che

- in ogni raggruppamento ci sono esattamente k elementi;
- ciascun raggruppamento differisca dagli altri raggruppamenti per almeno un elemento oppure per l'<u>ordine</u> dei vari elementi.

[Insomma: nelle DISPOSIZIONI, L'ORDINE CONTA!]

Se nei vari raggruppamenti non sono ammesse ripetizioni, ossia se i k elementi devono essere <u>tutti</u> quanti <u>DIVERSI</u> fra di loro, allora le disposizioni si chiamano DISPOSIZIONI SEMPLICI; se invece sono ammesse ripetizioni, parleremo di DISPOSIZIONI CON RIPETIZIONI.

<u>Esempio</u>: In quanti modi 2 sedie <u>distinte</u> (per esempio: una rossa e l'altra verde) possono essere occupate da 5 ragazzi (Alberto, Ernesto, Ivan, Omar, Umberto)?

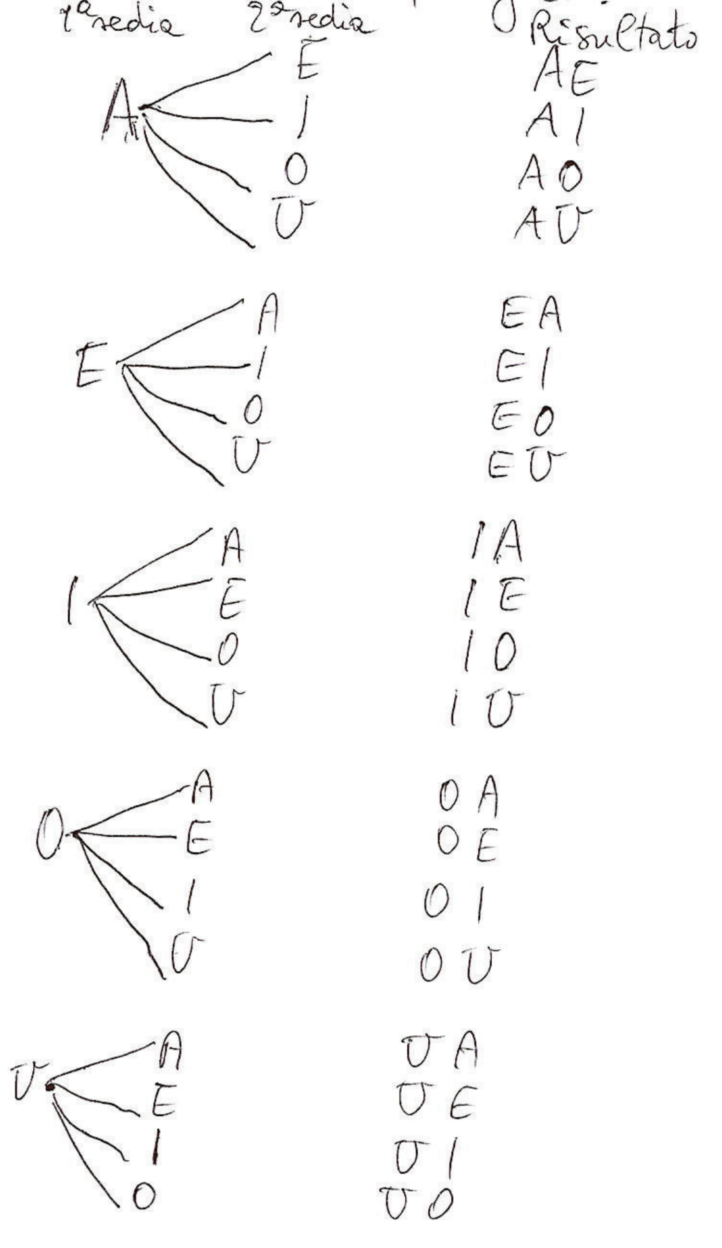
1ª sedia (rossa)	2ª sedia (verde)

Si tratta quindi di contare i possibili raggruppamenti di 2 elementi (perché le sedie sono due) presi dall'insieme {Alberto, Ernesto, Ivan, Omar, Umberto}, che denoteremo per semplicità con {A, E, I, O, U} con le seguenti regole:

1) l'ordine CONTA (perché le 2 sedie, rossa e verde, sono <u>distinte</u>: quindi dire AE significa che Alberto si siede sulla 1ª sedia (rossa) ed Ernesto sulla 2ª sedia (verde), mentre EA vuol dire che Ernesto occupa la 1ª sedia (rossa) e Alberto la 2ª sedia (verde): pertanto AE ed EA NON SONO LA STESSA COSA).
2) <u>non</u> ci sono ripetizioni, perché una persona NON PUÒ occupare due sedie diverse.

Si tratta quindi di $\boxed{\text{DISPOSIZIONI SEMPLICI.}}$

Adesso, <u>contiamo</u> le disposizioni semplici: <u>QUANTE</u> sono le disposizioni semplici di 5 oggetti a 2 a 2?
La prima sedia può essere occupata in 5 modi diversi (A, E, I, O oppure U): una volta occupata la prima sedia da un ragazzo, la seconda sedia può essere occupata solamente da uno dei rimanenti 4 ragazzi.



Il numero totale delle disposizioni semplici, come si può vedere da questo elenco, è $5 \cdot 4 = \boxed{20}$
[Qui svolge un ruolo fondamentale l'operazione di PRODOTTO]

$n = 5 \quad k = 2$

$5 \cdot 4 = n \cdot (n-1)$

(lo teniamo presente)

Più in generale, procedendo analogamente: quante sono le <u>disposizioni semplici</u> di n oggetti a k a k (con $k \leq n$: perché se fosse $k > n$, dato che <u>NON</u> sono ammesse <u>ripetizioni</u>, allora gli n oggetti non basterebbero per "coprire" i k posti)?

Dunque: supponiamo di avere k caselle DISTINTE (come le sedie, che sono <u>distinte</u>), e di doverle riempire con k elementi <u>diversi</u> (perché NON CI SONO RIPETIZIONI) scelti tra gli n elementi (diversi) dell'insieme di partenza. 1) L'elemento da collocare nella 1^a casella può essere scelto in n modi diversi. 2) In corrispondenza ad ognuna delle n scelte di cui al punto 1), ci sono $\boxed{n-1}$ possibilità di scelta per l'elemento da inserire nella 2^a casella, perché l'elemento messo nella 1^a casella NON PUÒ essere ripetuto. A questo punto si hanno $\boxed{n \cdot (n-1)}$ modi di riempire le prime due caselle. Allora, rimangono $n-2$ elementi per riempire la terza casella e, procedendo analogamente, si hanno $\boxed{n \cdot (n-1) \cdot (n-2)}$ modi di riempire le prime tre caselle. Quindi, procedendo ancora, quando si arriva alla k-esima (ed ultima) casella, si hanno $n-k+1$ elementi per riempire il k-simo posto.
$n-(k-1)$

$\boxed{1^o}$ elemento	: $\boxed{n}$ scelte possibili	$1+n$
$\boxed{2^o}$ elemento	: $\boxed{n-1}$ scelte possibili	$2+n-1=1+n$
$\boxed{3^o}$ elemento	: $\boxed{n-2}$ scelte possibili	$3+n-2=1+n$
$\boxed{k^o}$ elemento	: $\boxed{n-k+1}$ scelte possibili	$k+n-k+1=1+n$

Per ricordarsi il numero delle scelte possibili, teniamo conto che qui la somma deve fare sempre $1+n$ (!)

Quindi (usando sempre il PRODOTTO) si ha che il numero $D_{n,k}$ delle disposizioni semplici di n oggetti a k a k ($k \leq n$) è dato da

$$\boxed{D_{n,k} = n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot (n-k+1)}$$

Nel calcolo di $D_{n,k}$, nel fare l'operazione del prodotto, abbiamo tenuto conto del fatto che vale il seguente

PRINCIPIO FONDAMENTALE (senza dimostrazione)

Se una prima procedura può essere realizzata in n_1 modi diversi, e se successivamente una seconda procedura può essere realizzata in n_2 modi diversi, e una terza in n_3, e così via, ..., allora il numero di modi in cui la procedura totale può essere realizzata è $n_1 \cdot n_2 \cdot n_3 \dots$

Ora, il prossimo passo sarà di esprimere il numero $D_{n,k}$ con i FATTORIALI

ORA DEFINIAMO IL FATTORIALE $n!$, si legge "enne fattoriale"

Si definisce, per convenzione: $0! = 1$, $1! = 1$.

Se n è un intero, $n \geq 2$, si definisce FATTORIALE di n il PRODOTTO DEI primi n numeri interi positivi, cioè $2! = 1 \cdot 2 = 2$ $3! = 1 \cdot 2 \cdot 3 = 6$, $\boxed{4!} = 1 \cdot 2 \cdot 3 \cdot 4 = (1 \cdot 2 \cdot 3) \cdot 4 = \underbrace{6}_{3!} \cdot 4 = \boxed{24}$, $5! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 = (1 \cdot 2 \cdot 3 \cdot 4) \cdot 5 = 24 \cdot 5 = 120$ $(= 4! \cdot 5)$, ...,

$$\boxed{n!} = \underbrace{1 \cdot 2 \cdot 3 \cdot 4 \cdot \dots (n-1)}_{\text{prodotto dei primi } n-1 \text{ numeri interi positivi}} \cdot n = \boxed{(n-1)! \cdot n}$$

$$\boxed{(n+1)!} = \underbrace{1 \cdot 2 \cdot 3 \cdot 4 \cdot \dots \cdot n}_{\text{prodotto dei primi } n \text{ numeri interi positivi}} \cdot (n+1) = \boxed{n! \cdot (n+1)}$$

Sussistono quindi le importantissime FORMULE DI RICORRENZA $\boxed{n! = (n-1)! \cdot n}$ $\boxed{(n+1)! = n! \cdot (n+1)}$

usando le quali è possibile calcolare $n!$ conoscendo il valore di $(n-1)!$ e calcolare $(n+1)!$ conoscendo il valore di $n!$ Questo, in realtà, è quello che facciamo di fatto quando scriviamo $5! = (1 \cdot 2 \cdot 3 \cdot 4) \cdot 5 = 4! \cdot 5 = 24 \cdot 5 = 120$, perché partiamo dal fatto che già conosciamo il valore di $4!$ (che è 24), e poi ci ricaviamo il valore di $5!$

Abbiamo detto che $D_{n,k} = n\cdot(n-1)\cdot(n-2)\cdot\ldots\cdot(n-k+1)$
Da ciò otteniamo che

$$D_{n,k} = \frac{n\cdot(n-1)\cdot(n-2)\cdot\ldots\cdot(n-k+1)\cdot(n-k)\cdot(n-k-1)\cdot\ldots\cdot 3\cdot 2\cdot 1}{(n-k)\cdot(n-k-1)\cdot\ldots\cdot 3\cdot 2\cdot 1} =$$

$$= \frac{n!}{(n-k)!}$$

$$\boxed{D_{n,k} = \frac{n!}{(n-k)!}}$$

Veniamo ora alle DISPOSIZIONI CON RIPETIZIONE.
Questa volta k può essere un intero positivo qualunque.

Si dicono DISPOSIZIONI CON RIPETIZIONE DI n OGGETTI a k a k i raggruppamenti che si possono formare a partire dagli n oggetti, in modo tale che
- in ogni raggruppamento ci sono esattamente k elementi, ma non necessariamente distinti
- uno stesso elemento può comparire fino a k volte nello stesso gruppo (cioè, questa volta, SONO AMMESSE RIPETIZIONI)
- ciascun raggruppamento differisca dagli altri raggruppamenti per almeno un elemento oppure per l'ordine dei vari elementi (cioè, L'ORDINE CONTA)

Esempio: Quanti numeri di 2 cifre si possono avere usando le cifre 1, 2, 3? Si tratta di contare i possibili raggruppamenti di 2 elementi (perché le cifre da riempire, da "incasellare", sono due) presi dall'insieme $\{1,2,3\}$

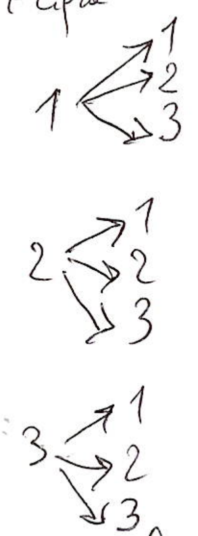
1a cifra	2a cifra
1a posizione	2a posizione

con le seguenti regole:
1) l'ordine conta (per esempio, 21 e 12 non sono lo stesso numero)
2) sono ammesse ripetizioni (perché esistono i numeri 11, 22 e 33)

Si tratta quindi di DISPOSIZIONI CON RIPETIZIONE ($D^{(r)}_{n,k}$)

Adesso, <u>contiamo</u> le disposizioni con ripetizione: QUANTE sono le disposizioni con ripetizione di 3 oggetti a 2 a 2? La prima cifra può essere occupata in 3 modi diversi; una volta scritta la prima cifra, questa volta - essendo ammesse ripetizioni - anche la seconda cifra può essere occupata in 3 modi diversi. Si ha: $\boxed{n=3;\ k=2}$

1ª cifra 2ª cifra



Numero
11
12
13

21
22
23

31
3~~2~~2
33

Il numero totale delle disposizioni con ripetizioni è $3 \cdot 3 = 3^2 = \boxed{9}$

$3 \cdot 3 = n \cdot n = n^2 = \boxed{n^k}$

[anche qui il prodotto svolge un ruolo fondamentale e si applica il PRINCIPIO FONDAMENTALE]

Più in <u>generale</u>, <u>procedendo analogamente</u>: quante sono le <u>disposizioni con ripetizione</u> di n oggetti a k a k?

Supponiamo di avere k caselle <u>distinte</u> (che corrisponde alla cifra di posizione da riempire: quindi l'ordine CONTA) e di doverle riempire con k elementi (tenendo conto che questa volta sono ammesse ripetizioni) scelti tra gli n elementi dell'insieme di partenza.

1) L'elemento da collocare nella 1ª casella può essere scelto in n modi diversi. 2) Essendoci la possibilità di avere ripetizioni, anche l'elemento da collocare nella 2ª casella può essere scelto in n modi diversi, e quindi si hanno $n \cdot n = n^2$ modi per riempire le prime 2 caselle. Continuando in questo modo, anche l'elemento da collocare nella 3ª, nella 4ª, ..., nella k-sima casella può essere scelto in n modi diversi, e quindi (tenendo conto anche del "principio fondamentale") abbiamo $n \cdot n \cdot n = n^3$ modi per riempire le prime 3 caselle, e così via, ... in totale abbiamo $\boxed{n^k}$ modi per riempire

tutte le k caselle, ottenendo

$$\boxed{D_{n,k}^{(r)} = n^k}$$

Ora veniamo alle PERMUTAZIONI SEMPLICI di n oggetti (che si indicano con P_n)

<u>Definizione</u>: Le permutazioni semplici di n oggetti sono le disposizioni semplici di n oggetti a n a n (N.B.: come detto anche per le disposizioni, questi n oggetti devono essere <u>distinti</u>)

$$\boxed{P_n = D_{n,n}}$$

Quante sono le permutazioni semplici? Siccome $D_{n,k} = \frac{n!}{(n-k)!}$, allora, sostituendo k con n, si ottiene

$$\boxed{P_n} = D_{n,n} = \frac{n!}{(n-n)!} = \frac{n!}{0!} = \frac{n!}{1} = \boxed{n!}$$

Dunque, $\boxed{P_n = n!}$

<u>Esempio</u>: Quanti sono gli anagrammi, anche senza senso, della parola ROMA?

Dobbiamo riempire 4 caselle e formare un'altra parola, per esempio AMOR, ORMA, RAMO...

Per la 1ª casella, abbiamo 4 scelte. Per la 2ª casella: una volta che è stata "sistemata" la 1ª lettera nella 1ª casella, restano 3 lettere. Quindi per la 2ª casella, restano 3 scelte. Sistemate le prime 2 lettere nelle prime 2 caselle, per la 3ª casella restano 2 lettere, cioè 2 scelte. Riempita anche la 3ª casella, resterà una scelta sola per la 4ª e ultima casella. Tenendo conto del "principio fondamentale", il numero globale di scelte per tutte e 4 le caselle è dato dal PRODOTTO delle scelte che riguarda ogni singola casella, cioè $4 \cdot 3 \cdot 2 \cdot 1 = 4! = 24$.

N.B. Notiamo che qui, naturalmente, l'ordine conta. Sono <u>disposizioni</u> semplici di 4 oggetti a 4 a 4, cioè <u>permutazioni</u> <u>di</u> 4 oggetti.

Più in generale, procedendo analogamente come nel calcolo del numero degli anagrammi della parola ROMA, che il numero degli anagrammi di una parola di n lettere distinte è $\boxed{n!}$, cioè il numero P_n delle permutazioni semplici di n oggetti, ossia il numero $D_{n,n}$ delle disposizioni semplici di n oggetti ad n ad n.

Quindi il numero degli anagrammi della parola NAPOLI è $\boxed{6!} = (1\cdot2\cdot3\cdot4\cdot5)\cdot6 = 120\cdot6 = \boxed{720}$, mentre il numero degli anagrammi della parola PERUGIA è $\boxed{7!} = (1\cdot2\cdot3\cdot4\cdot5\cdot6)\cdot7 = 6!\cdot7 = 720\cdot7 = \boxed{5040}$.

Adesso invece calcoliamo il numero di anagrammi di parole dove ci sono ripetizioni di lettere.

Facciamo subito un esempio, Calcoliamo il numero di anagrammi della parola $\boxed{\text{CATANIA}}$

Intanto, scriviamo sempre il numero totale delle lettere e scriviamo anche le ripetizioni, cioè quali lettere si ripetono e, per ciascuna di queste lettere, quante volte si ripete.

Nel nostro caso abbiamo: numero di lettere totali $= 7$; poi c'è la A che si ripete 3 volte. Non ci sono altre ripetizioni. Come procediamo? Osserviamo che, se le lettere fossero tutte quante diverse, per esempio $CA_1TA_2NIA_3$, allora il numero degli anagrammi sarebbe $7! = 5040$. Invece c'è una sola possibilità A A A (A si ripete 3 volte) al posto di tutti i possibili anagrammi di $A_1A_2A_3$, che, siccome abbiamo supposto che $A_1A_2A_3$ siano 3 lettere diverse, sono $3!$, come abbiamo detto all'inizio di questa pagina. Allora il numero totale degli anagrammi della parola CATANIA è $\frac{7!}{3!} = \frac{\cancel{1\cdot2\cdot3}\cdot4\cdot5\cdot6\cdot7}{\cancel{1\cdot2\cdot3}} = 840$.

Calcoliamo ora il numero degli anagrammi della parola FERRARA.

Numero totale delle lettere: 7

Ripetizioni: la A 2 volte e la R 3 volte

Se le lettere fossero diverse $FER_1R_2A_1R_3A_2$ allora il numero totale degli anagrammi sarebbe $7!$ Ma bisogna considerare le ripetizioni della A e della R.

Abbiamo solo una possibilità ($A\ A$: la lettera A si ripete 2 volte) al posto di tutti i possibili anagrammi della parola A_1A_2 che, essendo A_1, A_2 per ipotesi diverse, sono $2!$

Quindi intanto dobbiamo dividere per $2!=2$ ma otterremo ancora gli anagrammi non di FERRARA, ma di $FER_1R_2AR_3A$ (la R ancora non l'abbiamo considerata): quindi (siccome abbiamo una sola scelta $R\,R\,R$ al posto di tutti i possibili anagrammi della parola $R_1R_2R_3$, che sono $3!$, in quanto abbiamo supposto che R_1, R_2 ed R_3 siano distinte tra loro) dobbiamo ancora dividere per $3!$. Pertanto il numero totale degli anagrammi della parola FERRARA sarà $\frac{7!}{2!\,3!}=$

$$= \frac{\cancel{1\cdot 2}\cdot 3\cdot 4\cdot 5\cdot \cancel{6}\cdot 7}{\cancel{1\cdot 2}\cdot \cancel{1\cdot 2\cdot 3}} = 420$$

Quindi il numero degli anagrammi di una parola costituita da n lettere, ma in cui ci sono h lettere tali che la prima lettera si ripete n_1 volte, la 2^a lettera si ripete n_2 volte, ..., la h-sima lettera si ripete n_h volte, è dato da

$$\boxed{\frac{n!}{n_1!\cdot n_2!\cdot \ldots \cdot n_h!}}$$

Questa formula si chiama FORMULA DELLE <u>PERMUTAZIONI CON RIPETIZIONI</u>

<u>Esempio</u>: Usando direttamente questa formula, calcolare gli anagrammi della parola ECCELLENTE.
Il numero totale delle lettere è 10.
La E si ripete 4 volte; la C e la L si ripetono 2 volte. Quindi il numero degli anagrammi sarà

$$\frac{10!}{4!\cdot 2!\cdot 2!} = \frac{\cancel{1\cdot 2\cdot 3\cdot 4}\cdot 5\cdot 6\cdot 7\cdot \overset{2}{\cancel{8}}\cdot 9\cdot 10}{\cancel{1\cdot 2\cdot 3\cdot 4\cdot 2\cdot 2}} = 37800$$

Passiamo adesso alle COMBINAZIONI DI n OGGETTI a k a k. Finora abbiamo visto e studiato situazioni in cui "l'ordine conta"; adesso vediamo che cosa succede quando "l'ordine non conta" (per esempio, se si deve fare solo una scelta di 3 vincitori in un concorso, senza necessariamente fare una classifica o graduatoria, allora la terna, diciamo, "Alberto-Bruno-Carlo" è equivalente alla terna "Carlo-Alberto-Bruno": queste 2 terne sono, per così dire, "la stessa cosa", in quanto, se si deve fare solo una scelta dei vincitori, è importante solamente che Alberto, Bruno e Carlo abbiano vinto, e non chi, fra i 3, si sia classificato primo, o secondo, o terzo).

Si dicono COMBINAZIONI SEMPLICI (O SENZA RIPETIZIONE) DI n OGGETTI a k a k i raggruppamenti che si possono formare a partire dagli n oggetti, in modo tale che

- in ogni raggruppamento ci sono esattamente k elementi DISTINTI TRA DI LORO (e quindi senza ripetizione)
- ciascun raggruppamento differisca dagli altri raggruppamenti per almeno un elemento, ma non per l'ordine.

Quindi, nelle combinazioni, "l'ordine non conta", vale a dire due o più raggruppamenti fatti con gli stessi elementi sono considerati la stessa combinazione, anche se l'ordine con cui si presentano è diverso. Per esempio, nel caso dei tre vincitori Alberto, Bruno e Carlo, che chiameremo per semplicità A, B e C, i seguenti raggruppamenti danno luogo alla stessa combinazione:

ABC , ACB, BAC, BCA, CAB, CBA

Questi 6 raggruppamenti sono tutte le possibili "sottoclassifiche" relative ai 3 vincitori Alberto, Bruno, Carlo (notazione: ABC vuol dire 1° Alberto, 2° Bruno e 3° Carlo, ACB significa 1° Alberto, 2° Carlo e 3° Bruno, e così via): sono $3! = 6$, perché si vede che questi raggruppamenti possono essere "identificati" con la totalità degli anagrammi della parola ABC di 3 lettere fra loro distinte, che sono per l'appunto $3! = 6$. Questo, teniamolo presente!

Adesso, prima di calcolare in generale il numero delle combinazioni semplici di n oggetti a k a k, stabiliamo un collegamento PROFONDO con il numero delle disposizioni semplici di n oggetti a k a k. Per fare ciò, iniziamo con un esempio/esercizio.

ESEMPIO / ESERCIZIO : Ad un concorso si presentano 5 concorrenti (A = Alberto, B = Bruno, C = Carlo, D = Domenico, E = Ernesto), e vincono i primi 3.

a) Determinare tutte le possibili classifiche (o graduatorie) dei 3 vincitori (p.es. 1° Alberto, 2° Bruno, ecc....)

b) Determinare tutte le possibili scelte (o selezioni) dei 3 vincitori (non è importante che un concorrente si sia classificato 1° o 2° o 3°, ma solo che si sia piazzato fra i primi 3 e che abbia vinto!)

Cominciamo dal punto a). Intanto: qui l'ordine conta: [1° Alberto e 2° Bruno] è diverso da [1° Bruno e 2° Alberto], e non ci sono ripetizioni, perché naturalmente un concorrente non può essersi classificato in 2 o più posizioni diverse... Quindi in questo caso abbiamo DISPOSIZIONI SEMPLICI di 5 oggetti a 3 a 3, che sono in numero di $\frac{5!}{(5-3)!} = \frac{5!}{2!} = \frac{\cancel{1 \cdot 2} \cdot 3 \cdot 4 \cdot 5}{\cancel{1 \cdot 2}} = 60$. Vediamolo più da vicino.

Il 1° posto può essere assegnato in 5 modi diversi; una volta "sistemato" il primo classificato, il 2° posto può essere assegnato in 4 modi diversi, tra i 4 concorrenti che rimangono. Una volta "sistemato" anche il secondo classificato, il 3° posto può essere assegnato in 3 modi diversi. Complessivamente, applicando il "principio fondamentale", i 3 posti relativi alla <u>classifica</u>, <u>graduatoria</u> dei 3 vincitori possono essere assegnati in $5 \cdot 4 \cdot 3 = \boxed{60}$ modi diversi.

primi due elementi

primo elemento

A → AB: ABC, ABD, ABE
A → AC: ACB, ACD, ACE
A → AD: ADB, ADC, ADE
A → AE: AEB, AEC, AED

B → BA: BAC, BAD, BAE
B → BC: BCA, BCD, BCE
B → BD: BDA, BDC, BDE
B → BE: BEA, BEC, BED

C → CA: CAB, CAD, CAE
C → CB: CBA, CBD, CBE
C → CD: CDA, CDB, CDE
C → CE: CEA, CEB, CED

D → DA: DAB, DAC, DAE
D → DB: DBA, DBC, DBE
D → DC: DCA, DCB, DCE
D → DE: DEA, DEB, DEC

E → EA: EAB, EAC, EAD
E → EB: EBA, EBC, EBD
E → EC: ECA, ECB, ECD
E → ED: EDA, EDB, EDC

Tutto questo se si prendono le DISPOSIZIONI semplici. E con le combinazioni?

Nelle combinazioni, gli elementi del tipo

ABC, ACB, BAC, BCA, CAB, CBA

di cui parlavamo prima, vanno presi "una volta sola", perché mentre come DISPOSIZIONI sono 6 elementi <u>distinti</u>, invece come COMBINAZIONI sono lo stesso elemento.

E così anche per gli elementi del tipo

CDE, CED, DCE, DEC, ECD, EDC, ..., e via dicendo.

Praticamente, quando si passa dal computo del numero (cioè dall'atto di contare il numero) delle combinazioni semplici al computo delle corrispondenti disposizioni semplici, "si moltiplica per 6", perché, per esempio nel caso di Alberto, Bruno e Carlo, si devono considerare TUTTI gli anagrammi della parola ABC (che sono 6, e che invece – come combinazioni, sono un solo elemento). Quindi, nel passare dalle disposizioni semplici alle corrispondenti combinazioni semplici, "si divide per 6". Quindi si ha che: il <u>numero</u> delle <u>combinazioni</u> <u>semplici</u> di 5 elementi (oppure oggetti) a 3 a 3 è $\frac{1}{6}$ del numero delle disposizioni semplici di 5 oggetti a 3 a 3, cioè $\boxed{10}$.

Verifichiamolo. Sono veramente 10? Si tratta di "collocare" 3 elementi dell'insieme {A, B, C, D, E} in 3 caselle, rispettando però le seguenti 2 regole:

a) non ci devono essere ripetizioni;

b) "l'ordine non conta", nel senso che, se ho sistemato il raggruppamento ABC, non devo sistemare gli altri raggruppamenti ACB, BAC, BCA, CAB, CBA.

A questo punto, intanto, facciamo una cosa che ci può fare comodo. Tra i 5 elementi A, B, C, D, E, inseriamo un "ordine", oppure "ordinamento": per esempio, diremo che

$A<B<C<D<E$

(Questo non vuol dire, se si scrive $A<B$, che A è più piccolo di B, ma vuol dire semplicemente "A viene prima di B" ed è in senso astratto (non è riferito a nessuna graduatoria, anche perché stiamo studiando il problema delle scelte e non quello delle graduatorie!!))

Adesso elenchiamo i nostri raggruppamenti che costituiscono le nostre combinazioni

1) ABC ora, teniamo fissi A e B e facciamo variare il terzo elemento 2) ABD 3) ABE non ci sono altri raggruppamenti che hanno A come primo elemento e B come 2°

Allora teniamo sempre A come primo elemento, ma prendiamo C come secondo elemento: otteniamo

4) ACD 5) ACE e basta, perché adesso <u>non</u> <u>possiamo</u> considerare ACB, perché abbiamo già preso ABC e, a livello di combinazioni, ACB e ABC sono "la stessa cosa"

Ci può essere d'aiuto il fatto che, nelle nostre combinazioni, le lettere le prendiamo sempre in ordine "crescente", rispetto all'ordinamento $A<B<C<D<E$ di cui sopra, anche per evitare di sbagliarci e per avere maggiore chiarezza.

Ora prendiamo A come primo elemento e D come secondo elemento: resta 6) ADE (non possiamo prendere ADB e ADC, perché abbiamo preso già ABD e ACD). Ora tutti i casi in cui A è il primo elemento sono finiti:

non possiamo prendere AEB, AEC, AED perché abbiamo preso ABE, ACE, ADE. Insomma, osserviamo che, quando si prendono le combinazioni, gli elementi vanno presi in ordine "crescente" rispetto all'ordinamento $A<B<C<D<E$:

si può "andare avanti" ma non "tornare indietro". Restano quindi: 7) BCD 8) BCE 9) BDE 10) CDE, e basta. Quindi sono <u>veramente</u> 10 elementi.

Quindi, nel passare dalle disposizioni semplici alle corrispondenti combinazioni semplici, si ha:

$$C_{5,3} = \frac{D_{5,3}}{3!} \quad \text{ove } 3! \text{ corrisponde a } k!$$

Più in generale si può vedere che $C_{n,k} = \frac{D_{n,k}}{k!}$

ove $k!$ è il numero delle permutazioni di k oggetti.

Quindi $C_{n,K} = \frac{1}{k!} \cdot \frac{n!}{(n-k)!} = \frac{n!}{k! \cdot (n-k)!}$

La quantità $\frac{n!}{k!(n-k)!}$ la si indica anche col simbolo $\binom{n}{k}$ che si legge "n su k".

<u>Esempio</u>: In una gelateria ci sono 6 gusti (fragola, limone, vaniglia, cioccolato, caffè, pistacchio). Quanti sono i gelati che si possono fare con 2 gusti diversi?

Innanzi tutto l'ordine non conta (dire "fragola e limone" è la stessa cosa che dire "limone e fragola"). Inoltre è specificato che i 2 gusti sono <u>diversi</u>, e quindi non sono ammesse ripetizioni. Si tratta quindi di COMBINAZIONI SEMPLICI DI 6 OGGETTI A 2 a 2, che sono $(n=6, k=2)$

$$\binom{6}{2} = \frac{6!}{2!\,4!} = \frac{\cancel{1 \cdot 2 \cdot 3 \cdot 4} \cdot 5 \cdot \overset{3}{\cancel{6}}}{1 \cdot \cancel{2} \cdot \cancel{1 \cdot 2 \cdot 3 \cdot 4}} = 15.$$

Passiamo ora alle COMBINAZIONI CON RIPETIZIONE

Siano n e k due numeri interi positivi (Qui, non necessariamente $k \leq n$).

Si dicono <u>COMBINAZIONI CON RIPETIZIONE</u> di n OGGETTI a k a k i raggruppamenti che si possono formare a partire dagli n oggetti, in modo tale che

- ogni raggruppamento contenga esattamente k elementi (non necessariamente distinti tra loro, perché sono ammesse ripetizioni);
- ogni elemento possa essere ripetuto fino a k volte nello stesso raggruppamento;
- ciascun raggruppamento differisca dagli altri raggruppamenti per almeno un elemento, e non per l'ordine

Esempio: In una gelateria ci sono 6 gusti (gusto 1 = fragola; 2 = limone; 3 = vaniglia; 4 = cioccolato; 5 = caffè; 6 = pistacchio). Quanti sono i gelati che si possono fare con 2 palline di gelato, tenendo conto che le 2 palline non sono numerate e ci possono essere ripetizioni di uno stesso gusto?

Si tratta di COMBINAZIONI CON RIPETIZIONE DI 6 OGGETTI A 2 a 2

Quante sono? Cerchiamo di elencarle, tenendo conto che: 11 (uno-uno) (limone + limone), ecc., è ammesso; 12 (uno-due) (fragola e limone) e 21 (due-uno) (limone e fragola) sono "la stessa cosa". Quindi, analogamente al caso delle combinazioni semplici, conviene considerare i raggruppamenti dove gli elementi 1, 2, 3, 4, 5, 6 sono disposti in ordine non-decrescente. Diciamo "non-decrescente" anziché "crescente" (che vorrebbe dire "strettamente crescente") perché sono ammesse ripetizioni, quindi sono ammessi i raggruppamenti 11, 22 (due-due), eccetera, ..., ed inoltre, visto che – per esempio – 12 e 21 sono "la stessa cosa", il gelato al limone e fragola – come raggruppamento – sarà indicato con 12 e non con 21 (tanto, il discorso è: dobbiamo scegliere tra 12 e 21, e allora prendiamo quel "raggruppamento" in cui le cifre si presentano in ordine non-decrescente, cioè 12). Quindi, tenuto conto di queste "accortezze", presentiamo le nostre combinazioni con ripetizione.

1ª pallina	2ª pallina
1	1
1	2
1	3
1	4
1	5
1	6
2	2
2	3
2	4
2	5
2	6

1ª pallina	2ª pallina
3	3
3	4
3	5
3	6
4	4
4	5
4	6
5	5
5	6
6	6

Sono dunque 21 (ventuno) raggruppamenti!

$\boxed{\text{Quindi } C^{(r)}_{6,2} = 21,}$

Anche $C_{7,2} = \binom{7}{2} = \frac{7!}{2! \cdot 5!} = \frac{1 \cdot 2 \cdot \cancel{3 \cdot 4 \cdot 5} \cdot 6 \cdot 7}{1 \cdot 2 \cdot \cancel{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5}} = \boxed{21}$

Non è un caso che $C^{(r)}_{6,2} = C_{7,2} = 21$

In generale, si può vedere che la formula del numero di combinazioni con ripetizioni di n elementi a k a k è

$$\boxed{C^{(r)}_{n,k} =} C_{n+k-1,k} = \binom{n+k-1}{k} = \frac{(n+k-1)!}{k!\,(n+\cancel{k}-1-\cancel{k})!} =$$

$$= \boxed{\frac{(n+k-1)!}{k!\,(n-1)!}}$$

Di questa formula non facciamo la dimostrazione nel caso generale, ma mostriamo l'uguaglianza $C^{(r)}_{n,k} = C_{n+k-1,k}$ nel caso $n=6, k=2$, cioè $C^{(r)}_{6,2} = C_{7,2}\ (=21)$. Nelle prime due colonne ci sono i raggruppamenti che costituiscono le combinazioni con ripetizioni di 6 oggetti a 2 a 2 (cioè 11, 12, 13, 14, 15, 16, 22, 23, ...) (11 = uno-uno, ...). I numeri della 3ª colonna sono ottenuti aggiungendo 1 ai corrispondenti numeri della 2ª colonna. Osserviamo che (ed è questa la "mate-magica"!) nella 1ª e 3ª colonna ci sono i raggruppamenti che COSTITUISCONO LE COMBINAZIONI SEMPLICI DI 7 OGGETTI A 2 a 2 (!!!!) (infatti abbiamo: 12 (uno-due), 13, 14, 15, 16, 17, 23, 24, 25, 26, 27, 34, 35, 36, 37, 45, 46, 47, 56, 57, 67) (sono sempre 21 (ventuno) elementi!)

1ª pallina	2ª pallina	2ª colonna +1
1	1	2
1	2	3
1	3	4
1	4	5
1	5	6
1	6	7
2	2	3
2	3	4
2	4	5
2	5	6
2	6	7
3	3	4
3	4	5
3	5	6
3	6	7
4	4	5
4	5	6
4	6	7
5	5	6
5	6	7
6	6	7

Ora facciamo un <u>COLLEGAMENTO PROFONDO</u> fra le disposizioni, combinazioni di n oggetti a k a k e le <u>funzioni</u> del tipo $f:\{1,2,\dots,k\} \to \{1,2,\dots,n\}$.

Una <u>funzione</u> (di questo tipo) è una legge che ad ogni elemento dell'insieme $\{1,2,\dots,k\}$ fa corrispondere uno e un solo elemento dell'insieme $\{1,2,\dots,n\}$.

Ci domandiamo: a quali tipi di funzioni $f:\{1,2,\dots,k\} \to \{1,2,\dots,n\}$ corrispondono le disposizioni semplici, le disposizioni con ripetizione, le combinazioni semplici, le combinazioni con ripetizione di n oggetti a k a k? Quando abbiamo parlato di combinazioni semplici, abbiamo parlato di "ordine crescente", mentre quando abbiamo studiato le combinazioni con ripetizione abbiamo parlato di "ordine non-decrescente". Allora l'idea (e il risultato) è che <u>alle combinazioni semplici</u> corrispondono le funzioni $f:\{1,2,\dots,k\} \to \{1,2,\dots,n\}$ STRETTAMENTE CRESCENTI, mentre alle combinazioni <u>con ripetizione</u> corrispondono le funzioni NON-DECRESCENTI.

<u>Definizioni</u>: Una funzione $f:\{1,2,\dots,k\} \to \{1,2,\dots,n\}$ si dice: <u>STRETTAMENTE CRESCENTE</u>, se, comunque presi $i,j \in \{1,2,\dots,k\}$ con $i<j$, si ha $\boxed{f(i)<f(j)}$; <u>NON DECRESCENTE</u>, se, comunque presi $i,j \in \{1,2,\dots,k\}$ con $i<j$, si ha $f(i) \leq f(j)$.

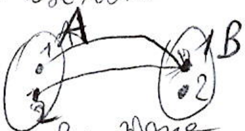
CALCOLO COMBINATORIO E INIETTIVITÀ — 55 — NEW NEW parte 4 testo adottato

Le definizioni di funzione strettamente crescente e non decrescente si formulano più in generale per funzioni del tipo $f: A \to B$ definite in un insieme non vuoto qualsiasi A e che assumono valore in un insieme non vuoto qualsiasi B. (n.b.: Una funzione è una legge che ad ogni elemento $a \in A$ associa uno e un solo elemento di B, che chiameremo $f(a)$, vedi anche Parte 2)

[Da ogni elemento di A parte UNA e UNA SOLA FRECCIA]

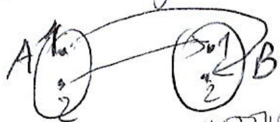
INIETTIVITÀ E SURIETTIVITÀ

Una funzione $f: A \to B$ si dice iniettiva se non esistono due elementi $i, j \in A$ tali che $i \neq j$ ed $f(i) = f(j)$, cioè, equivalentemente, se comunque presi 2 elementi $i, j \in A$ con $i \neq j$ si ha $f(i) \neq f(j)$. Una funzione $f: A \to B$ si dice non-iniettiva se esistono almeno due elementi $i, j \in A$ tali che $i \neq j$ ed $f(i) = f(j)$.



Una funzione $f: A \to B$ si dice suriettiva se per ogni $b \in B$ esiste almeno un $a \in A$ con $f(a) = b$.
Una funzione $f: A \to B$ si dice biiettiva se f è sia iniettiva sia suriettiva.

funzione NON INIETTIVA: almeno un elemento di B è raggiunto da 2 frecce; funzione NON SURIETTIVA: almeno un elemento di B non è raggiunto da nessuna freccia



funzione INIETTIVA: ogni elemento di B raggiunto da AL PIÙ una freccia; funzione SURIETTIVA: ogni elemento di B è raggiunto da almeno una freccia

N.B.: RISULTATO: Ogni funzione STRETTAMENTE CRESCENTE è INIETTIVA (ma non viceversa)

N.B.: Come si mettono in relazione i raggruppamenti (di n oggetti a k a k) con le funzioni $f: A \to B$, ove $A = \{1, 2, .., k\}$, $B = \{1, 2, .., n\}$? Per esempio: $n = 5$: un possibile raggruppamento di $k = 3$ oggetti presi dai 5 oggetti 1, 2, 3, 4, 5 è: $\boxed{2\ 4\ 5}$. Che vuol dire? $\boxed{2\ 4\ 5}$ è quella funzione $f: \{1,2,3\} \to \{1, 2, .., 5\}$ tale che $f(1) = 2$, $f(2) = 4$, $f(3) = 5$ (vedi figura). Invece il raggruppamento $\boxed{4\ 5\ 2}$ corrisponde a quell'(altra) funzione $h: \{1,2,3\} \to \{1,2,3,4,5\}$ tale che $h(1) = 4$, $h(2) = 5$, $h(3) = 2$.

f(1)	f(2)	f(3)
2	4	5

Adesso facciamo vedere i tipi di funzioni che corrispondono a disposizioni semplici, disposizioni con ripetizione, combinazioni semplici, combinazioni con ripetizione: daremo l'enunciato nel caso generale, ma lo facciamo vedere, per semplicità, con un esempio ($n = 4$, $k = 2$).

Funzioni $f: \{1, 2, \ldots, k\} \to \{1, 2, \ldots, n\}$ $n=4$ $k=2$

Disposizioni con ripetizione di 4 oggetti $(1,2,3,4)$ a 2 a 2

a 2 a 2 vuol dire che bisogna "sistemare", "riempire" 2 caselle.

Già un aspetto del COLLEGAMENTO PROFONDO con le funzioni sta in ciò:

ALLA PRIMA CASELLA CORRISPONDE $f(1)$,

ALLA SECONDA CASELLA CORRISPONDE $f(2)$.

Prima casella	Seconda casella	$f(1)$	$f(2)$
1 →	1	1	1
	2	1	2
	3	1	3
	4	1	4
2 →	1	2	1
	2	2	2
	3	2	3
	4	2	4
3 →	1	3	1
	2	3	2
	3	3	3
	4	3	4
4 →	1	4	1
	2	4	2
	3	4	3
	4	4	4

Le disposizioni con ripetizione sono <u>TUTTE QUANTE</u> le funzioni $f: \{1,2\} \to \{1,2,3,4\}$

← Qui ci sono TUTTE le possibilità. Corrisponde anche a <u>TUTTI</u> i possibili punteggi ottenuti lanciando 2 tetraedri regolari distinti (uno rosso e l'altro blu) con le 4 facce contrassegnate ciascuna da un punteggio che va da 1 a 4, (e "non truccati") (come il dado da 1 a 6): $f(1)$ corrisponde al punteggio del 1° tetraedro, mentre $f(2)$ corrisponde al punteggio del 2° tetraedro.

Sono $n^k = 4^2 = \boxed{16}$

1) Disposizioni semplici (cioè senza ripetizione) di 4 oggetti a 2 a 2.

1ª casella 2ª casella

$1 \to 2, 3, 4$

$2 \to 1, 3, 4$

$3 \to 1, 2, 4$

$4 \to 1, 2, 3$

Sono $\frac{n!}{(n-k)!} = \frac{4!}{2!} = \frac{1 \cdot 2 \cdot 3 \cdot 4}{1 \cdot 2} = 12$

$f(1)$	$f(2)$
1	2
1	3
1	4
2	1
2	3
2	4
3	1
3	2
3	4
4	1
4	2
4	3

Qui non sono TUTTE le funzioni $f: \{1, \dots, k\} \to \{1, \dots, n\}$ perché, per esempio, manca la funzione definita ponendo $f(1) = 1$, $f(2) = 1$. E questo non è un caso. Perché si tratta delle disposizioni senza ripetizione. Senza ripetizione vuol dire che i valori di $f(1)$ e di $f(2)$ devono essere distinti. [Più in generale, nel caso k non necessariamente uguale a 2, i valori di $f(1), f(2), \dots, f(k)$ devono essere distinti].

Questo vuol dire che, comunque si prendano i, j con $i \neq j$, si deve avere necessariamente $f(i) \neq f(j)$. Quindi f deve essere INIETTIVA. Inoltre, siccome l'ordine conta, bisogna prendere tutte quante le funzioni iniettive, perché l'unico vincolo che c'è è che nel raggruppamento "non ci siano ripetizioni", e abbiamo visto che, nel linguaggio matematico, questo si traduce "non ci sono punti $i \neq j$ con $f(i) = f(j)$".

QUINDI ALLE DISPOSIZIONI SEMPLICI, CIOÈ SENZA RIPETIZIONI, DI n OGGETTI A k A k CORRISPONDONO TUTTE E SOLE LE FUNZIONI INIETTIVE $f: \{1, 2, \dots, k\} \to \{1, 2, \dots, n\}$. Similmente, si può vedere che ALLE PERMUTAZIONI SEMPLICI DI n OGGETTI CORRISPONDONO TUTTE E SOLE LE FUNZIONI BIIETTIVE $f: \{1, 2, \dots, n\} \to \{1, 2, \dots, n\}$.

Come accennato, alle combinazioni semplici corrispondono le funzioni STRETTAMENTE CRESCENTI $f:\{1,2,\dots,k\}\to\{1,2,\dots,n\}$

Combinazioni semplici (cioè senza ripetizione) di 4 oggetti a 2 a 2

f(1)	f(2)
1	2
1	3
1	4
2	3
2	4
3	4

Sono: $\binom{n}{k} = \dfrac{n!}{k!(n-k)!} \underset{(k=2)}{\overset{(n=4)}{=}}$

$= \dfrac{4!}{2!2!} = \dfrac{\cancel{1}\cdot 2\cdot 3\cdot \cancel{4}^{2}}{\cancel{1}\cdot\cancel{2}\cdot 1\cdot \cancel{2}} = \boxed{6}.$

... e alle combinazioni con ripetizione corrispondono le funzioni NON DECRESCENTI $f:\{1,2,\dots,k\}\to\{1,2,\dots,n\}$

Combinazioni con ripetizione di 4 oggetti a 2 a 2 $(n=4, k=2)$

f(1)	f(2)	2ª colonna +1
1	1	2
1	2	3
1	3	4
1	4	5
2	2	3
2	3	4
2	4	5
3	3	4
3	4	5
4	4	5

$\binom{n+k-1}{k} = \dfrac{(n+k-1)!}{k!\cdot(n-1)!} =$

$= \dfrac{5!}{2!3!} = \dfrac{\cancel{1\cdot 2\cdot 3}\cdot \cancel{4}^{2}\cdot 5}{\cancel{1\cdot 2}\cdot \cancel{1\cdot 2\cdot 3}} = 10$

Alla formula $\dfrac{5!}{2!3!}$ si arriva anche considerando la 1ª e la 3ª colonna da sole, cioè le COMBINAZIONI SEMPLICI (o SENZA RIPETIZIONE) di 5 OGGETTI a 2 a 2 $\binom{5}{2} = \dfrac{5!}{2!3!}$.

SCHEMA DI RICAPITOLAZIONE DEL CALCOLO COMBINATORIO

Raggruppamenti (n oggetti a k a k)	Quantità	Funzioni $f:\{1,\dots,k\}\to\{1,\dots,n\}$
$k\le n$ oppure $k>n$ Disposizioni con ripetizione $D^{(r)}_{n,k}$	n^k	TUTTE LE FUNZIONI l'ordine conta
$k\le n$ Disposizioni semplici (senza ripetizione) $D_{n,k}$	$\frac{n!}{(n-k)!}=n(n-1)\cdots(n-k+1)$	Funzioni INIETTIVE l'ordine conta
$k\le n$ Combinazioni semplici (senza ripetizione) $C_{n,k}$	$\binom{n}{k}=\frac{n!}{k!(n-k)!}$	Funzioni STRETTAMENTE CRESCENTI l'ordine non conta
$k\le n$ (oppure anche $k>n$) Combinazioni con ripetizione $C^{(r)}_{n,k}$	$\binom{n+k-1}{k}=\frac{(n+k-1)!}{k!(n-1)!}$	Funzioni NON DECRESCENTI l'ordine non conta
$k=n$ Permutazioni di n oggetti tra loro (permutazioni semplici) P_n	$n!$	Funzioni BIIETTIVE $f:\{1,\dots,n\}\to\{1,\dots,n\}$ l'ordine conta

APPENDICE: PERCENTUALI

Il concetto di percentuale è molto usato in moltissime situazioni della vita quotidiana (non solo in Matematica!) Se vogliamo dare una <u>definizione</u> <u>di</u> <u>percentuale</u>: LA PERCENTUALE È UNA FRAZIONE AVENTE COME DENOMINATORE 100, E, SU UN DATO TOTALE, INDICA QUANTE UNITÀ SU 100 SODDISFANO UNA CERTA CONDIZIONE.

<u>Dunque: la (quasi unica!) cosa da "ricordare" è</u>

$$\boxed{\text{"kappa per cento"} = k\% = \frac{k}{100}}$$

<u>Il</u> numero k si chiama TASSO PERCENTUALE

Analogamente, "kappa per mille" $= \boxed{k‰ = \frac{k}{1000}}$.

Diamo <u>subito</u> esercizi.

<u>ESERC.</u>: Il poeta Eugenio Montale visse anagraficamente 85 anni. In una sua poesia, lui afferma di essere vissuto "al 5%" ("non aumentate la dose"). Quanti sarebbero gli anni "veramente vissuti", secondo questo suo (pessimistico) pensiero?

Il 5% di 85 è uguale a $85 \cdot \frac{5}{100} = 4{,}25$, cioè 4 anni e $\frac{1}{4}$, ossia <u>4 anni</u> e <u>3 mesi</u> (1 anno = 12 mesi). E non di più, per rispetto nei suoi confronti.

<u>Esercizio:</u> C'è una tassa del 20%. Un uomo, durante una giornata, fa un bel guadagno di 450 Euro lordi. Quanto guadagna al netto?

$$450 - \frac{45\cancel{0} \cdot 2\cancel{0}}{1\cancel{0}\cancel{0}} = 450\left(1 - \frac{20}{100}\right) = 450 \cdot \frac{80}{100}$$

$$= 450 - 90 = 360 \text{ Euro}$$

Resta l'80%.

<u>Altro Esercizio:</u> Un uomo gioca 50 Euro alla Roulette. Al primo passo la sua quantità aumenta del 10%. Al secondo passo la sua quantità diminuisce del 10%. Quanto gli rimane? 50 Euro, di più o di meno?

(N.B.: $10\% = \frac{10}{100}$)

1° passo: $50+50\cdot\frac{10}{100}=50\left(1+\frac{10}{100}\right)=\frac{5\cancel{0}\cdot 11\cancel{0}}{1\cancel{00}}=55$

2° passo: $55-55\cdot\frac{10}{100}=55\left(1-\frac{10}{100}\right)=55\cdot\frac{9\cancel{0}}{10\cancel{0}}=49{,}5$.

Quindi la sua quantità è diminuita.

Esercizio: Una merce costa 27 Euro, IVA compresa. Ammettendo che l'IVA sia al 22%, quanto sarebbe costata la merce senza l'IVA? E quant'è l'ammontare dell'IVA?

Indichiamo con x il costo della merce senza IVA. Si ha:

$x\left(1+\frac{22}{100}\right)=27$, cioè $x\cdot\frac{\cancel{122}^{61}}{\cancel{100}_{50}}=27$, da cui

$x=\frac{27\cdot 50}{61}\cong 22{,}1311$ L'ammontare dell'IVA è

$27-22{,}1311=4{,}8689$.

Esercizio: Ogni settimana la popolazione mondiale aumenta di 1,4 milioni di persone. In un certo anno fissato (non importa quale), la popolazione era di 3,8 miliardi di persone.

Calcolare l'incremento percentuale annuo.

In un anno ci sono (circa) 52 settimane ($52\cdot 7=364$).

Quindi in un anno la popolazione aumenta di $1{,}4\cdot 52$ milioni di persone, cioè 72,8 milioni di persone. Ma all'inizio erano $3{,}8\cdot 10^9=3{,}8\cdot 10^3\cdot 10^6=3800\cdot 10^6=3800$ milioni di persone. Quindi l'incremento è

$$\frac{72,8}{3800} \simeq 0,0191 = 1,91\%$$ (infatti $1,91\% = \frac{1,91}{100} = 0,0191$).

<u>Esercizio</u>: In un'aula magna, ci sono 1139 europei e 461 extraeuropei. Qual è la percentuale degli europei e degli extraeuropei?

$$\text{Europei} = \frac{1139}{1139+461} = \frac{1139}{1600} \cong 0,712 = 71,2\%$$

(perché $71,2\% = \frac{71,2}{100} = 0,712$).

$$\text{Extraeuropei} = \frac{461}{1600} \cong 0,288 = 28,8\%$$

(perché $28,8\% = \frac{28,8}{100} = 0,288$)

<u>Esercizio</u>: Il 3,5% di una somma S è 70 Euro. A quanto è uguale S?

Si ha: $\frac{3,5}{100} \cdot S = 70$ (N.B.: $k\% = \frac{k}{100}$),

e quindi

$$S = \frac{70 \cdot 100}{3,5} = \frac{7000}{3,5} = \frac{70000}{35} = 2000 \text{ Euro}.$$

Esercizio:

Il 15% dei membri di una popolazione si ammalarono di coronavirus, e l'8% delle persone ammalate morirono. Calcolare la mortalità del coronavirus rispetto all'intera popolazione.

Su 100 individui, 15 presero il coronavirus.

L'8% di 15 è: $\cancel{15}^{3} \cdot \frac{\cancel{8}^{2}}{\cancel{100}\,\cancel{25}_{5}} = \frac{6}{5} = 1{,}2$

Quindi la mortalità è stata dell'1,2%.

Esercizio:

Il 4% del 20% di un numero è 1. Qual è il numero?

Teniamo conto che "$k\% = \frac{k}{100}$". Indichiamo con x il numero da trovare. Si ha:

$$\frac{\cancel{4}^{1}}{\cancel{100}_{25}} \cdot \frac{\cancel{20}^{1}}{\cancel{100}_{5}} \cdot x = 1, \text{ cioè } \frac{1}{125} \cdot x = 1, \text{ e quindi}$$

$$\boxed{x = 125}.$$

Esercizio:

In un campione estratto da una popolazione ci sono 380 persone con la malattia A, 219 con la malattia B e 3050 con la malattia C. Trovare le percentuali di ogni gruppo

Il numero totale delle persone del campione è 3649.

Malattia A: $\frac{380}{3649} \simeq 0,10414 = 10,414\%$

$k\% = \frac{k}{100}$: quindi, quando si ha un numero (come nel nostro caso 0.10414) e lo si deve scrivere in percentuale, lo si moltiplica per 100, perché si deve fare l'operazione inversa della divisione per 100 (Esempio: $0,25 = 25\%$: infatti $0,25 \cdot 100 = 25$, e $0,25 = \frac{25}{100}$). Analogamente si ha:

Malattia B: $\frac{219}{3649} \simeq 0,06002 = 6,002\%$

Malattia C: $\frac{3050}{3649} \simeq 0,83584 = 83,584\%$

PERCENTUALI – TRATTAZIONE PIÙ APPROFONDITA (LEGGERE)

Si dice <u>percentuale</u> una frazione decimale in cui il denominatore è uguale a 100, e, su un dato totale, indica quante unità su 100 soddisfano una certa condizione.

Dunque, la cosa fondamentale è

$$\boxed{\text{“kappa per cento”} = k\% = \frac{k}{100}}$$

Il numero k si chiama TASSO PERCENTUALE.

Ovviamente, $0\% = 0$, perché $\frac{0}{100} = 0$.

Ora, come si trasformano le frazioni in percentuali?

Per esempio, $\frac{1}{2}$ di una certa quantità, quale percentuale è di quella quantità?

A questa domanda rispondiamo così.

Deve essere $\frac{1}{2} = \frac{k}{100}$. Allora, moltiplicando per 100 la frazione $\frac{1}{2}$, si ottiene k, cioè $k = 100 \cdot \frac{1}{2} = 100 : 2 = 50$.

Quindi $\frac{1}{2}$ corrisponde, equivale al 50%.

Quindi, data una frazione $\frac{p}{q}$ con $p, q \in \mathbb{N}$, $q \neq 0$, per “trasformarla” in percentuale $k\% = \frac{k}{100}$, si pone

$\frac{p}{q} = \frac{k}{100}$ e si trova

$$\boxed{k = 100 \cdot \frac{p}{q}}$$

(cioè si moltiplica la frazione stessa $\frac{p}{q}$ per 100)

$\frac{1}{4}$ corrisponde a $100 \cdot \frac{1}{4}$ %, cioè 25%

Notiamo che $100 = 10^2 = (2 \times 5)^2 = 2^2 \times 5^2 = 4 \times 25$, quindi $\frac{100}{4} = 25$ e $\frac{100}{25} = 4$

$\frac{1}{25}$ corrisponde a $100 \cdot \frac{1}{25}$ % = 4%

N.B.: Noi terremo conto, quando serve, che $\boxed{100 = 2^2 \times 5^2 = 4 \times 25}$

$\frac{1}{5}$ corrisponde a $100 \cdot \frac{1}{5} = \frac{2 \times 2 \times 5 \times 5}{1 \times 5} =$ (proprietà invariantiva) $= \frac{2 \times 2 \times 5}{1} = 20$%. Quindi

(Non diremo più esplicitamente che si usa la proprietà invariantiva e anche le proprietà associativa, commutativa....)

$\frac{1}{5}$ corrisponde a 20%

$\frac{1}{20}$ corrisponde a $100 \cdot \frac{1}{20} = \frac{2 \times 2 \times 5 \times 5}{2 \times 2 \times 5 \times 1} = \frac{5}{1} = 5$%. Quindi

$\frac{1}{20}$ corrisponde a 5%

$\frac{4}{5}$ corrisponde a $100 \cdot \frac{4}{5} = \frac{2 \times 2 \times 5 \times 5 \times 4}{5} = \frac{2 \times 2 \times 4 \times 5 \times 5}{5} =$

$= 2 \times 2 \times 4 \times 5 = (2 \times 5) \times (2 \times 4) = 10 \times 8 = 80$%

$\frac{3}{4}$ corrisponde a $100 \cdot \frac{3}{4} = \frac{4 \times 25 \times 3}{4} = 25 \times 3 = 75$%

da u
$^{1}2\ 5 \times$
3
$7\ ^{1}5$

$\frac{1}{6}$ corrisponde a $100 \cdot \frac{1}{6} = \frac{100 : 2}{6 : 2} = \frac{50}{3} = 10 \times \frac{5}{3}$%...

Abbiamo visto poche pagine indietro che $\frac{5}{3} = 1,\overline{6} = 1,6666666...$

quindi $10 \times \frac{5}{6} = 16,6666666... = 16,\overline{6}$. Quindi,

$\frac{1}{6}$ corrisponde a $16,\overline{6}$% $= 16,666666...$%

$\frac{1}{3}$ corrisponde a $100 \cdot \frac{1}{3}$% Avevamo visto in questo materiale didattico che $\frac{1}{3} = 0,333333... = 0,\overline{3}$. Quindi, moltiplicando per **100**, cioè spostando la virgola a destra di 2 posti, si ha: $100 \times \frac{1}{3} = 33,33333.... = 33,\overline{3}$. Pertanto,

$\frac{1}{3}$ corrisponde a $33,\overline{3}$% $= 33,333333....$%

$\frac{5}{6}$ corrisponde a $100 \cdot \frac{5}{6}$% Avevamo visto in queste note che $\frac{5}{6} = 0,8\overline{3} = 0,8333333...$, e pertanto

$100 \cdot \frac{5}{6} = 100 \cdot 0{,}8333333\ldots = 83{,}333333\ldots\% = 83{,}\bar{3}\%$

$\frac{5}{6}$ corrisponde a $83{,}\bar{3}\% = 83{,}333333\ldots\%$

$\frac{1}{9}$ corrisponde a $100 \cdot \frac{1}{9}\%$ Avevamo visto che $\frac{1}{9} = 0{,}\bar{1} =$
$= 0{,}11111111\ldots$, e pertanto $100 \cdot \frac{1}{9} = 11{,}111111\ldots = 11{,}\bar{1}$. Si ha:

$\frac{1}{9}$ corrisponde a $11{,}\bar{1}\% = 11{,}111111\ldots\%$

$\boxed{\frac{1}{10}}$ corrisponde a $100 \cdot \frac{1}{10}\% = \boxed{10\%}$

$\frac{2}{3}$ corrisponde a $100 \cdot \frac{2}{3} = \frac{200}{3}\%$

```
200,0 0   | 3
18        |---------
---       | 66,66....
 20
 18
---
  20
```

Qui inseriamo la virgola sia al dividendo sia al quoziente

Il processo si ripete all'infinito, quindi $\frac{200}{3} = 66{,}66666\ldots = 66{,}\bar{6}$

```
 18
---
  20
  18
---
   2....
```

Si ha:

$\frac{2}{3}$ corrisponde a $66{,}\bar{6}\% = 66{,}6666666\ldots\%$

$\frac{8}{9}$ corrisponde a $100 \cdot \frac{8}{9}$ Avevamo visto che $\frac{8}{9} = 0{,}88888888\ldots = 0{,}\bar{8}$
e quindi $100 \cdot \frac{8}{9} = 88{,}888888\ldots = 88{,}\bar{8}$. Quindi

$\frac{8}{9}$ corrisponde a $88{,}\bar{8}\% = 88{,}88888\ldots\%$ (= circa 89%)

NB: Si può vedere che la probabilità che la prova del nove fallisca equivale alla probabilità che, dato un numero intero relativo, questo sia multiplo di 9, quindi $\frac{1}{9}$ (circa 11,1%, o circa 11%), e che la proprietà che la prova del nove sia veritiera equivale alla probabilità che un numero non sia multiplo di 9, quindi $\frac{8}{9}$ = circa 88,9% o circa 89%: ECCO DA DOVE "SBUCA" la quantità $\boxed{89\%}$ (!!!)

Abbiamo dunque la seguente tabella:

(N.B.)

FRAZIONE	PERCENTUALE	NUMERO DECIMALE
$\frac{1}{25}$	4%	0,04
$\frac{1}{20}$	5%	0,05
$\frac{1}{10}$	10%	0,1
$\frac{1}{9}$	11,1111...% $=11,\overline{1}\%$	$0,11111... = 0,\overline{1}$
$\frac{1}{6}$	16,6666...% $=16,\overline{6}\%$	$0,166666... = 0,1\overline{6}$
$\frac{1}{5}$	20%	0,2
$\frac{1}{4}$	25%	0,25
$\frac{1}{3}$	33,333...% $=33,\overline{3}\%$	$0,33333... = 0,\overline{3}$
$\frac{1}{2}$	50%	0,5
$\frac{2}{3}$	66,6666...% $=66,\overline{6}$	$0,666666... = 0,\overline{6}$
$\frac{3}{4}$	75%	0,75
$\frac{4}{5}$	80%	0,8
$\frac{5}{6}$	83,3333...% $=83,\overline{3}$	$0,83333... = 0,8\overline{3}$
$\frac{8}{9}$	88,8888...% $=88,\overline{8}\%$	$0,88888... = 0,\overline{8}$

Per passare dalla 2^a alla 3^a colonna, si divide per 100 ($k\% = \frac{k}{100}$), cioè "si sposta la virgola di 2 posti verso sinistra", mentre naturalmente per passare dalla 3^a alla 2^a colonna, si moltiplica per 100, cioè "si sposta la virgola di 2 posti verso destra".

E così continuando (avevamo visto che $\frac{1}{8} = 0,125$ e $\frac{3}{8} = 0,375$...)

$\frac{1}{8}$	12,5%	0,125
$\frac{3}{8}$	37,5%	0,375

e così via.....

In Statistica, ci sono la <u>frequenza relativa</u> e la <u>frequenza percentuale</u> di dati, che corrispondono rispettivamente alla 1^a o 3^a colonna e alla 2^a colonna.
(frequenza relativa di un dato $= \frac{\text{numero di volte in cui c'è quel dato}}{\text{numero totale degli esperimenti}}$)

CALCOLI VELOCI CON LE PERCENTUALI...

1% DI 300 → 3

10% DI 300 → 30

1%: si divide per 100

10%: si divide per 10

$\left(10\% = \frac{10}{100} = \frac{1}{10}\right)$

ESEGUI

1% DI 400 = 4

1% DI 600 = 6

1% DI 800 = 8

1% DI 640 = 6,4

10% DI 600 = 60

10% DI 480 = 48

10% DI 620 = 62

10% DI 40 = 4

10% DI 300 = 30

1% DI 1000 = 10

1% DI 540 = 5,4

10% DI 500 = 50

10% DI 500 = 50

Più in generale: quant'è il k% di un numero $n \in \mathbb{N}$? È

$$\boxed{\frac{k}{100} \cdot n}$$

Vale a dire: k%, come abbiamo visto, si "traduce" con $\frac{k}{100}$, mentre la parola "di" si "traduce" con il simbolo della moltiplicazione.

Esempi: 1) A che cosa è uguale il 20% di 70 (ℓ_1)?

È uguale a: $\ell_1 = \frac{20}{100} \cdot 70 = \frac{2 \times 10 \times 7 \times 10}{10 \times 10} = \frac{2 \times 7 \times 10 \times 10}{1 \times 10 \times 10} =$

$= \frac{14}{1} = 14$ (abbiamo applicato la proprietà INVARIANTIVA e le proprietà commutativa e associativa ......) (in seguito lo faremo anche se non lo diremo esplicitamente)

2) A quanto è uguale il 5% di 40 (ℓ_2)?

È uguale a $\boxed{\ell_2} = \frac{5}{100} \times 40 = \frac{5 \times 4 \times 10}{10 \times 10} = \frac{20 \times 10}{10 \times 10} = \frac{2 \times 10 \times 10}{1 \times 10 \times 10} = \frac{2}{1} = \boxed{2}$.

3) Calcolare ℓ_3 = il 2% di 5.000

Si ha: $\boxed{\ell_3} = \frac{2}{100} \times 5.000 = \frac{2 \times 5 \times 10 \times 10 \times 10}{1 \times 10 \times 10} = \frac{2 \times 5 \times 10}{1} = 10 \times 10 = \boxed{100}$.

4) Calcolare il 10% di 120 (ℓ_4)

Si ha: $\boxed{\ell_4} = \frac{10}{100} \times 120 = \frac{12 \times 10 \times 10}{10 \times 10} = \boxed{12}$

5) Calcolare il 5% di 340 (ℓ_5)

Si ha: $\boxed{\ell_5} = \frac{5}{100} \times 340 = \frac{5 \times 34 \times 10}{10 \times 10} = \frac{5 \times 34}{10} = \frac{34 \times 5}{2 \times 5} = \frac{34}{2} = \boxed{17}$

$$\begin{array}{r|l} 34 & 2 \\ \hline 2 & 17 \\ \hline 14 & \\ 14 & \\ \hline = & \end{array}$$

$\boxed{6)}$ Calcolare il 2% di 800 (ℓ_6)

Si ha: $\boxed{\ell_6} = \frac{2}{100} \times 800 = \frac{2 \times 8 \times 10 \times 10}{10 \times 10} = \boxed{16}$

ESERCIZI SULLE PERCENTUALI

Le percentuali sono spesso usate nei problemi in cui intervengono i concetti di

SCONTO, **AUMENTO**, **INTERESSE**.

In breve,

1) LO SCONTO È IL VALORE PERCENTUALE CHE VIENE TOLTO DALLA CIFRA DA PAGARE.
2) L'AUMENTO È IL VALORE PERCENTUALE CHE VIENE AGGIUNTO AL PREZZO INIZIALE.
3) L'INTERESSE È IL VALORE PERCENTUALE CHE VIENE AGGIUNTO A UNA SOMMA DI DENARO PRESTATA O DEPOSITATA.

1) Una merce costa 70 €. Poi iniziano i saldi, e viene applicato uno sconto del 20%. Quanto costa quella merce nel periodo dei saldi?

Nella pagina precedente abbiamo detto che il 20% di 70 è 14. Nel periodo dei saldi, quindi, la merce costa 70 € − 14 € = = (70 − 14) € = 56 €. Infatti, sulla colonna u 0 − 4 non si può fare; allora ci facciamo prestare 1 da = 10 u dalla colonna da, e quindi le unità (del minuendo), che erano 0, diventano 10, mentre le decine (del minuendo), che erano 7, diventano 6. Quindi si ha: sulla colonna u, $10 - 4 = \boxed{6}$; sulla colonna da, $6 - 1 = \boxed{5}$, quindi $\boxed{5}$ da + $\boxed{6}$ u, cioè 56.

$$\begin{array}{cc} da & u \\ \cancel{7}^{6} & \cancel{0}^{10} & - \\ 1 & 4 & = \\ \hline 5 & 6 \end{array}$$

2) Una merce l'anno scorso costava 40 €; adesso è aumentata del 5%. Quanto costa ora?

Nella pagina precedente abbiamo visto che il 5% di 40 è 2. Quindi la merce è aumentata di 2 €, e ora costa 40 + 2 = 42 €.

3) Una somma di 5'000 € viene depositata in banca, e dopo un anno frutta un interesse del 2%. Quant'è, dopo un anno, l'ammontare della somma?
Abbiamo visto che il 2% di 5'000 è 100. Allora l'ammontare della somma dopo un anno sarà di (5'000+100)€, cioè 5'100 €.

4) Una merce l'anno scorso costava 120 €, ma adesso il prezzo è aumentato del 10%. Quanto costa adesso?
Abbiamo visto che il 10% di 120 è 12. Pertanto adesso quella merce costa (120 + 12) € = 132 €.

5) Un cellulare l'anno scorso costava € 340, ma ora il prezzo è aumentato del 5%. Quanto costa adesso?
Abbiamo visto che il 5% di 340 è 17. Pertanto il prezzo attuale del cellulare, in Euro, è 340+17=357.

6) Un computer costava 800 Euro l'anno scorso, ma ora il prezzo è aumentato del 2%. Quanto costa adesso?
Abbiamo visto che il 2% di 800 è 16. Allora il computer costa ora Euro 800+16 = 816 -

Esercizio / Problema

In un'urna sono state messe tante palline: 60 gialle, 120 verdi, 240 rosse e 80 blu. Qual è, in percentuale, la probabilità che sia estratta una pallina gialla/verde/rossa/blu? (N.B.: si suppone che l'estrazione sia "onesta" e che le palline non siano "truccate")

Risoluzione

Innanzi tutto, calcoliamo il numero totale di palline presenti nell'urna. Facciamo l'addizione in colonna.

$$\begin{array}{r} \text{da}\ \text{u} \\ ①12\,+ \\ 8\,= \\ \hline 2^{①}0 \end{array} \qquad \begin{array}{rrrl} \text{h} & \text{da} & \text{u} & \\ & 6 & 0 & + \\ ②1 & 2 & 0 & + \\ 2 & 4 & 0 & + \\ & 8 & 0 & = \\ \hline 5 & ^{②}0 & 0 & \end{array}$$

Sulla colonna u, $\boxed{0}$. Sulla colonna da: $6+2+4+8 = 8+4+8 = 12+8 = 2\boxed{0}$: scriviamo $\boxed{0}$ sulla colonna da e riportiamo ② sulla colonna h, dove si ha: $1+2+$ ② (il riporto) $= \boxed{5}$ h, che insieme a $\boxed{0}$ da e $\boxed{0}$ u fornisce il risultato di 500: quindi il numero TOTALE di palline è $\boxed{500}$.

Quindi $\frac{60}{500}$ delle palline sono gialle, $\frac{120}{500}$ sono verdi, $\frac{240}{500}$ sono rosse, $\frac{80}{500}$ sono blu.

Per avere le rispettive percentuali, si moltiplicano le rispettive frazioni per 100. Quindi la percentuale delle palline gialle è: $\frac{60}{500} \times 100 = \frac{2 \times 3 \times 10 \times 100}{5 \times 100} =$

$= \left(\text{proprietà invariantiva}\right) \frac{2 \times 3 \times 10}{5} = \frac{2 \times 3 \times 2 \times 5}{1 \times 5} = \frac{2 \times 3 \times 2}{1} = 6 \times 2 = \boxed{12}$.

Quindi le palline gialle sono il 12% del totale.

La percentuale delle palline verdi è:

$$\frac{120}{500}\times 100=\frac{120\times 100}{5\times 100}=\frac{120}{5}=\frac{12\times 10}{5}=$$

$$=\frac{12\times 2\times 5}{1\times 5}=\frac{12\times 2}{1}=24.$$ Quindi il 24% delle palline sono verdi.

(N.B.: abbiamo applicato la proprietà invariantiva, non lo diremo più esplicitamente)

La percentuale delle palline rosse è:

$$\frac{240}{500}\times 100=\frac{240\times 100}{5\times 100}=\frac{240}{5}=\frac{24\times 10}{5}=\frac{24\times 2\times 5}{1\times 5}=$$

$$=\frac{24\times 2}{1}=48.$$ Quindi il 48% delle palline sono rosse.

La percentuale delle palline blu è $\frac{80}{500}\times 100=$

$$=\frac{80\times 100}{5\times 100}=\frac{80}{5}=\frac{8\times 2\times 5}{1\times 5}=\frac{8\times 2}{1}=16\%.$$

Le palline blu, dunque, costituiscono il 16% della totalità delle palline.

Risposta

La probabilità che sia estratta una pallina gialla/verde/rossa/blu è rispettivamente del 12%/24%/48%/16%,

<u>Osservazioni</u>: N.B.: In Statistica, i numeri 60, 120, 240, 80 sono le <u>frequenze assolute</u> (cioè il numero) delle palline gialli/verdi/rosse/blu, mentre i numeri $\frac{60}{500}$, $\frac{120}{500}$, $\frac{240}{500}$, $\frac{80}{500}$ sono le rispettive <u>frequenze relative</u>, e 12%, 24%, 48%, 16% sono le rispettive <u>frequenze percentuali</u>.

Le frequenze relative si possono ricavare anche dalle frequenze percentuali dividendo per 100, ovvero "spostando la virgola di due posti verso sinistra", ottenendo rispettivamente 0,12; 0,24; 0,48; 0,16.

Adesso si possono fare tante cose, se vogliamo "confrontare" questi numeri con $\frac{60}{500}$, $\frac{120}{500}$, $\frac{240}{500}$, $\frac{80}{500}$.

Riduciamo le 4 frazioni <u>ai minimi termini</u>, e teniamo conto (come fatto in precedenza) che $100 = 2^2 \times 5^2$.

Si ha: $$\frac{60}{500} = \frac{2 \cdot 3 \cdot 10}{5 \cdot 2^2 \cdot 5^2} = \frac{2 \cdot 3 \cdot 2 \cdot 5}{5 \cdot 2 \cdot 2 \cdot 5 \cdot 5} = \frac{5 \cdot 2 \cdot 2 \cdot 3}{5 \cdot 2 \cdot 2 \cdot 5 \cdot 5} = \frac{2 \cdot 3}{5 \cdot 5} = \frac{3}{5^2} = \boxed{\frac{3}{25}}$$

$$\frac{120}{500} = \frac{4 \cdot 3 \cdot 10}{5 \cdot 2^2 \cdot 5^2} = \frac{2 \cdot 2 \cdot 3 \cdot 2 \cdot 5}{5 \cdot 2 \cdot 2 \cdot 5 \cdot 5} = \frac{5 \cdot 2 \cdot 2 \cdot 3 \cdot 2}{5 \cdot 2 \cdot 2 \cdot 5 \cdot 5} = \frac{2 \cdot 3}{5 \cdot 5} = \boxed{\frac{6}{25}}$$

$$\frac{240}{500} = \frac{8 \cdot 3 \cdot 10}{5 \cdot 2^2 \cdot 5^2} = \frac{2 \cdot 2 \cdot 2 \cdot 3 \cdot 2 \cdot 5}{5 \cdot 2 \cdot 2 \cdot 5 \cdot 5} = \frac{5 \cdot 2 \cdot 2 \cdot 3 \cdot 2 \cdot 2}{5 \cdot 2 \cdot 2 \cdot 5 \cdot 5} = \frac{3 \cdot 2 \cdot 2}{5 \cdot 5} = \boxed{\frac{12}{25}}$$

$$\frac{80}{500} = \frac{8 \cdot 10}{5 \cdot 2^2 \cdot 5^2} = \frac{2 \cdot 2 \cdot 2 \cdot 2 \cdot 5}{5 \cdot 2 \cdot 2 \cdot 5 \cdot 5} = \frac{2 \cdot 2 \cdot 2 \cdot 2 \cdot 5}{5 \cdot 5 \cdot 2 \cdot 2 \cdot 5} = \frac{2 \cdot 2}{5 \cdot 5} = \boxed{\frac{4}{25}}$$

Le rispettive divisioni $\frac{3}{25}$, $\frac{6}{25}$, $\frac{12}{25}$, $\frac{4}{25}$ daranno luogo (come già sappiamo) a 4 numeri decimali <u>limitati</u>, perché nel 25 c'è solo il fattore primo 5 (alla 2^a potenza)).

Ma, piuttosto che fare le divisioni in colonna (partendo dai numeri $\frac{3}{25}$, $\frac{6}{25}$, $\frac{12}{25}$, $\frac{4}{25}$), è più rapido moltiplicare numeratore e denominatore per 4, ottenendo frazioni con denominatore 100, e poi "spostare la virgola di 2 posti indietro, cioè verso sinistra", cioè:

$$\frac{3}{25} = \frac{3 \times 4}{25 \times 4} = \frac{12}{100} = 0{,}12 \qquad \frac{6}{25} = \frac{6 \times 4}{25 \times 4} = \frac{24}{100} = 0{,}24$$

$$\frac{12}{25} = \frac{12 \times 4}{25 \times 4} = \frac{6 \times 2 \times 4}{25 \times 4} = \frac{6 \times 8}{25 \times 4} = \frac{48}{100} = 0{,}48$$

$$\frac{4}{25} = \frac{4 \times 4}{25 \times 4} = \frac{16}{100} = 0{,}16$$

Esercizio / Problema

(è una specie di "viceversa" rispetto a quello precedente)

Ci sono 1'200 palline: il 15% di colore giallo, il 35% di colore verde, il 45% di colore rosso, il 5% di colore blu. Quante sono le palline gialle/verdi/rosse/blu?

Risoluzione.

Le palline gialle sono il 15% di 1'200, cioè

$$\frac{15}{100} \times 1'200 = \frac{3 \times 5 \times 4 \times 3 \times 100}{1 \times 100} = 3 \times 5 \times 4 \times 3 =$$

$$= 3 \times \underbrace{5 \times 2}_{10} \times 2 \times 3 = 10 \times 3 \times 2 \times 3 = 10 \times 6 \times 3 = 10 \times 18 = \boxed{180}.$$

Le palline verdi sono il 35% di 1'200, ossia

$$\frac{35}{100} \times 1'200 = \frac{7 \times 5 \times 4 \times 3 \times 100}{1 \times 100} = 7 \times \underbrace{5 \times 2}_{10} \times 2 \times 3 =$$

$$= 10 \times 7 \times 6 = 10 \times 42 = \boxed{420}.$$

Le palline rosse sono il 45% di 1'200, cioè $\frac{45}{100} \times 1'200 =$

$$= \frac{9 \times 5 \times 4 \times 3 \times 100}{1 \times 100} = 9 \times 5 \times 4 \times 3 = 9 \times \underbrace{5 \times 2}_{10} \times 2 \times 3 =$$

$$= 9 \times 10 \times 6 = 10 \times 6 \times 9 = 10 \times 54 = \boxed{540}.$$

Infine, le palline blu sono il 5% di 1'200, cioè $\frac{5}{100} \times 1'200 =$

$$= \frac{5 \times 4 \times 3 \times 100}{1 \times 100} = 5 \times 4 \times 3 = 20 \times 3 = \boxed{60}.$$

Risposta

Le palline gialle/verdi/rosse/blu sono rispettivamente 180/420/540/60.

Problema: In un~~o~~ parco ci sono 120 animali. Di questi, il 60% sono cervi, il 20% sono orsi e il resto sono leoni.

Quanti sono i cervi? E gli orsi? E i leoni?

Risoluzione.

Innanzi tutto, ricaviamoci la percentuale dei leoni. Il totale deve fare il 100%, quindi la percentuale dei leoni è $100-(60+20)=100-80=20$.

Quindi i leoni costituiscono, come gli orsi, il 20% degli animali presenti nel parco.

Allora il numero totale dei cervi è $120 \cdot \frac{60}{100} =$ (ricordiamo che $\boxed{k\% = \frac{k}{100}}$) $= \frac{3 \times 4 \times 10 \times 2 \times 3 \times 10}{10 \times 10} =$

$$= \frac{3 \times 3 \times 4 \times 2 \times 10 \times 10}{1 \times 10 \times 10} = \begin{pmatrix}\text{propietà} \\ \text{invariantiva}\end{pmatrix} = 3 \times 3 \times 4 \times 2 = 9 \times 8 = \boxed{72}$$

Il numero totale degli orsi, come pure il numero totale dei leoni, è $120 \cdot \frac{20}{100} = \frac{4 \times 3 \times 10 \times 2 \times 10}{10 \times 10} = \frac{4 \times 2 \times 3 \times 10 \times 10}{1 \times 10 \times 10} =$

$$= 4 \times 2 \times 3 = 8 \times 3 = \boxed{24}.$$

Risposta

Nel parco ci sono 72 cervi, 24 orsi e 24 leoni.

Problema

a) Una merce costa 19 Euro, ma viene scontata del 10%. Quant'è il prezzo scontato?

b) Dopo un certo periodo di tempo, al costo di 19 Euro viene applicato un aumento del 10%. Quant'è il prezzo aumentato?

Risoluzione

a) Il $\boxed{10\% \text{ di } 19}$ è $19 \cdot \frac{10}{100} = 19 \cdot \frac{1}{10} = 19 : 10 = \boxed{1{,}9}$ (perché quando si divide per 10, si sposta la virgola di un posto verso sinistra).

(Bisogna arrivare ai centesimi) $1u + 9d = 1u + 9d + \boxed{0}c$

Quindi il prezzo scontato, in Euro, è $19 - 1{,}9 = 19 - 1{,}90$.

Calcoliamolo facendo la sottrazione in colonna.

$$\begin{array}{cccc} da & u & d & c \\ 1 & {}^{8}\cancel{9}, & {}^{10}\cancel{0} & 0\ - \\ & 1, & 9 & 0\ = \\ \hline 1 & 7, & 1 & 0 \end{array}$$

Sulla colonna c (centesimi), $0 - 0 = 0$ (ovvio). Sulla colonna d (decimi), $0 - 9$ non si può fare: allora ci facciamo prestare $1u = 10d$: i decimi (del minuendo) erano 0 e diventano 10, le unità (sempre del minuendo) erano 9 e diventano 8. Quindi si ha: sulla colonna d, $10 - 9 = \boxed{1}$; sulla colonna u, $8 - 1 = \boxed{7}$; sulla colonna da, $1 - 0 = \boxed{1}$. Cioè, $\boxed{1}da + \boxed{7}u + \boxed{1}d + \boxed{0}c$, vale a dire $\boxed{17{,}10}$, che è il prezzo scontato.

b) Abbiamo visto al punto a) che il 10% di 19 è 1,90. Allora il prezzo aumentato, in Euro, è $19 + 1{,}90$. Sulla colonna d, $0 + 9 = 9$. Sulla colonna u, $9 + 1 = 10$: scrivo $\boxed{0}$ e riporto ① sulla colonna da, dove si fa $1 + ①$ (il riporto) $= \boxed{2}$.

$$\begin{array}{cccc} da & u & d & c \\ {}^{1}1 & 9, & 0 & 0\ + \\ & 1, & 9 & 0\ = \\ \hline 2 & {}^{1}0, & 9 & 0 \end{array}$$

Quindi: $\boxed{2}da + \boxed{0}u + \boxed{9}d + \boxed{0}c$, che danno luogo al numero $\boxed{20{,}90}$, che è il prezzo aumentato.

Risposte

a) Il prezzo scontato è di 17 Euro e 10 centesimi.

b) Il prezzo aumentato è di 20 Euro e 90 centesimi.

N.B.: Nel caso di 1% e 10%, è spesso conveniente vedere il tutto come NUMERI DECIMALI e con lo SPOSTAMENTO DELLA VIRGOLA, e soprattutto se si tratta di EURO, dove vengono richiesti numeri decimali con i centesimi!!

Esercizio

Qual è il 4% del 20% di 125?

Risoluzione: 4% e 20% vogliono dire, rispettivamente, $\frac{4}{100}$ e $\frac{20}{100}$; inoltre, 20% di 125 vuol dire $\frac{20}{100} \cdot 125$, e ancora, 4% del 20% vuol dire $\frac{4}{100} \cdot \frac{20}{100}$. Quindi, le parole "del" e "di" si traducono con il simbolo della moltiplicazione. Pertanto si ha:

$$\frac{4}{100} \cdot \frac{20}{100} \cdot 125 = \frac{4}{100} \cdot \frac{10}{100} \cdot 2 \cdot 125 =$$

$$= \frac{4}{100} \cdot \frac{10}{100} \cdot 250 = \frac{4}{100} \cdot \frac{10}{100} \cdot 10 \cdot 25 = \frac{4}{100} \cdot \frac{100}{100} \cdot 25 =$$

$$= \frac{4}{100} \cdot 25 = \frac{4 \cdot 25}{100} = \frac{100}{100} = \boxed{1}$$ (Ricordiamo che $100 = 2^2 \cdot 5^2 = 4 \cdot 25$)

Risposta

Il 4% del 20% di 125 è 1.

Il poeta Eugenio Montale visse anagraficamente 85 anni.
In una sua poesia (dal titolo "PER FINIRE"), lui afferma di essere "vissuto al 5%"

[...] Vissi al cinque per cento, non aumentate
la dose. Troppo spesso invece piove
sul bagnato.

Quanti sarebbero gli anni "veramente vissuti", secondo questo suo (pessimistico) pensiero?

Il 5% di 85 è uguale a $\frac{5}{100} \cdot 85 = \frac{5}{2^2 \cdot 5^2} \cdot 85 =$

$= \frac{5 \cdot 85}{5 \cdot 5 \cdot 4} = \frac{85}{5 \cdot 4} = \frac{85}{20} = \frac{85}{10 \cdot 2} = \frac{1}{10} \cdot \frac{85}{2} = \ldots$

Eseguiamo la divisione in colonna 85:2

$$\begin{array}{r|l} 85{,}0 & 2 \\ \hline 8 & 42{,}5 \\ \hline =5 & \\ 4 & \\ \hline 10 & \\ 10 & \\ \hline = & \end{array}$$

qui inseriamo la virgola al dividendo e al quoziente e abbassiamo uno 0

$\ldots = \frac{1}{10} \cdot 42{,}5 = \boxed{4{,}25}$

Quindi gli anni "veramente vissuti" sarebbero 4,25 anni.
Cerchiamo di esprimere questa quantità, se possibile, in termini di un numero <u>intero</u> di anni e di un numero <u>intero</u> di mesi. Intanto, 4,25 anni = 4 anni + 0,25 anni; quindi come numero intero di anni prenderemo $\boxed{4}$ e trasformiamo 0,25 anni in mesi. Come si fa? Si ha: 1 anno = 12 mesi, e pertanto, tenendo conto che $100 = 2^2 \times 5^2 = 4 \times 25$, si ha:

0,25 anni $= 0{,}25 \times 12$ mesi $= \frac{25}{100} \times 12$ mesi $= \frac{1 \times 25}{4 \times 25} \times 12$ mesi $= \frac{1}{4} \times 12$ mesi $= \boxed{3}$ mesi. Pertanto Montale sarebbe «veramente vissuto» solamente $\boxed{\text{4 ANNI E 3 MESI}}$.

Esercizio. Viceversa, se il Signor X.Y. ha 60 anni e dice: "Di questi 60 anni, se faccio la somma di tutti i momenti complessivi, avrò goduto la vita solamente per 3 anni". "In quale percentuale" è "vissuto" il Signor X.Y.?

Si ha: $\frac{3}{60} = k\% = \frac{k}{100}$, quindi $(100 = 2^2 \cdot 5^2)$

$$k = 100 \cdot \frac{3}{60} = \frac{2 \cdot 2 \cdot 5 \cdot 5 \cdot 3}{2 \cdot 3 \cdot 2 \cdot 5} = \frac{2 \cdot 3 \cdot 2 \cdot 5 \cdot 5}{2 \cdot 3 \cdot 2 \cdot 5 \cdot 1} = \frac{5}{1} = 5,$$

$(6 \cdot 10)$

SEMPRE AL 5% (!!!)

Naturalmente, questo esercizio e quello su Montale sono stati scelti con tanta, tanta sottile ironia.

W L'OTTIMISMO !!!