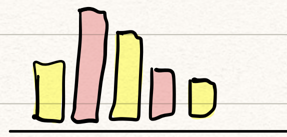
To draw the relative distribution doesn't matter the total count, the only thing that matter are the relative proportions (like relative frequencies)

$$f_i = \frac{n_i}{n}$$



c_1	n_1	f_1
$\vdots$	$\vdots$	$\vdots$
c_m	n_m	f_m

We talk about **PROBABILITY** as a form of abstraction of the concept of relative frequency

So you can look the relative frequency as a sort of manifestation of the abstract idea of probability

A **THEORETICAL POPULATION** can be constituded by even infinite statistical units or entities

BIVARIATE DISTRIBUTION

We can consider each variable separetly and we can compute the univariate distribution.

X	Y
x_1	y_1
$\vdots$	$\vdots$
x_n	y_n

↳ **MARGINAL DISTRIBUTION**

BIVARIATE DISTRIBUTION

CONTINGENCY TABLE

JOINT FREQUENCY

$X \backslash Y$	y_1	...	y_j	...	y_c	SUM
x_1	n_{11}	...	n_{1j}	...	n_{1c}	$n_{1\bullet}$
$\vdots$	$\vdots$		$\vdots$			
x_i	n_{i1}	- -	n_{ij}			
$\vdots$	$\vdots$					
x_r	n_{r1}					$n_{r\bullet}$
SUM	$n_{\bullet 1}$				$n_{\bullet c}$	

frequencies

UNIVARIATE DISTRIBUTION OF X

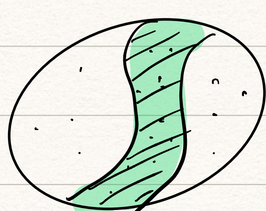
UNIVARIATE DISTRIBUTION OF Y

MARGINAL DISTRIBUTION

The row and the column inside are called

CONDITIONAL DISTRIBUTION

this indicate how many units have both variable contemporary



A distribution conditioned by an attribute

$Freq(X = x_i \mid Y = y_j)$

relative frequency in bivariate distribution

relative frequency in univariate distribution of Y

$$\frac{n_{ij}}{n_{\bullet j}} = \frac{n_{ij}}{n} \frac{n}{n_{\bullet j}} = \frac{n_{ij}}{n} \Big/ \frac{n_{\bullet j}}{n}$$

$$\frac{n_{ij}}{n_{i\bullet}} = \frac{n_{ij}}{n} \Big/ \frac{n_{i\bullet}}{n}$$

relative

$$\mathrm{Freq}(X = X_i \mid Y = Y_J) = \frac{\mathrm{Freq}(X = X_i, Y = Y_J)}{\mathrm{Freq}(Y = Y_J)}$$

BAYES THEOREM

DEF CONDITIONAL PROBABILITY

$$P(X \mid Y) = \frac{P(X, Y)}{P(Y)}$$

$$\mathrm{Freq}(X = X_i) =$$

$$= \sum_J \mathrm{Freq}(X = X_i \text{ AND } Y = Y_J)$$

For example

JOINT FREQUENCY

$$Freq(X = Male) = Freq(X = Male \text{ AND } Y = High) + Freq(X = Male \text{ AND } Y = low)$$

$$\wedge = \text{AND}$$

$$Freq(X = x_i) = \sum_J Freq(x_i \wedge Y = y_J)$$

$$Freq(X|Y) = \frac{Freq(X \wedge Y)}{Freq(Y)}$$

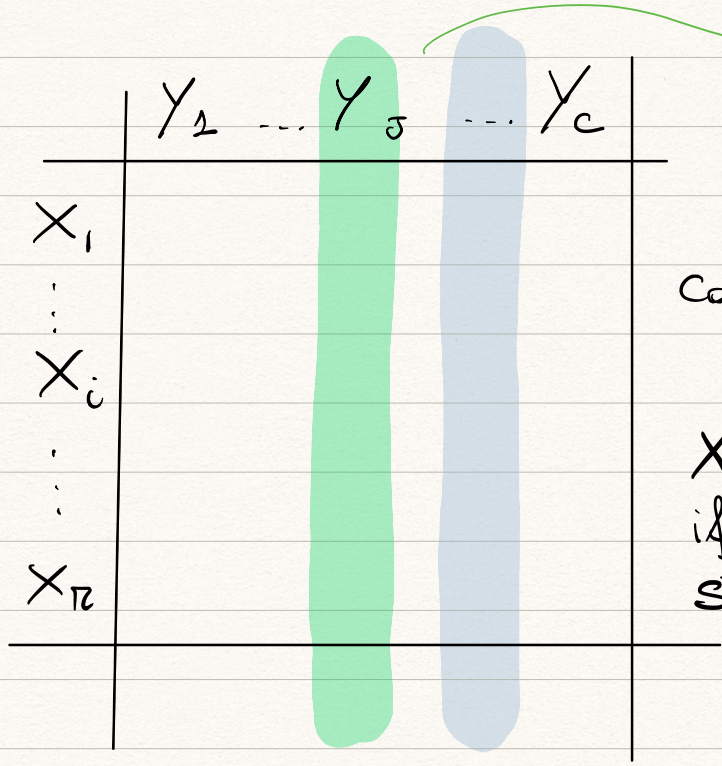
IDENTITY

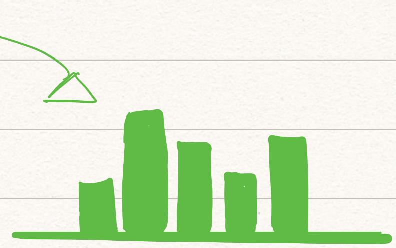
DEFINITION (BAYES THEOREM)

$$P(X|Y) = \frac{P(X \wedge Y)}{P(Y)}$$

INDEPENDENCE

$$\boxed{P(X \wedge Y) = P(X)\,P(Y)}$$





X and Y are independent if they have the same structure

Condition of INDEPENDENCE

For each i $\text{Freq}(X_i | Y_J)$ is equal for all J and therefore also equals to the marginal $[\text{Freq}(X_i)]$

$$\boxed{\frac{n_{iJ}}{n_{\cdot J}} = \frac{n_{i\cdot}}{n}}$$

$$\frac{n_{ij}}{n_{i\cdot}} = \frac{n_{\cdot j}}{n}$$

$$\Updownarrow$$

$$\frac{n_{ij}}{n_{\cdot j}} = \frac{n_{i\cdot}}{n}$$

$$\Rightarrow \boxed{n_{ij}\, n = n_{i\cdot}\, n_{\cdot j}}$$

If we divide by n

$$n_{ij} = \frac{n_{i\cdot}}{n}\,\frac{n_{\cdot j}}{n}$$

$$\Downarrow$$

$$f_{ij} = f_i f_j$$

INDEPENDENCE

$$\mathrm{Freq}(X \cap Y) = \mathrm{Freq}(X)\,\mathrm{Freq}(Y)$$

KOLMOGROV

Probability is defined Axiomatically

Defining properties of probability are the same are the same that hold for the frequencies