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

Bayes' formula

Let E and F be two events such that $P(E) > 0$ and $P(F) > 0$. Then

$$\boxed{P(E|F) = \frac{P(E \cap F)}{P(F)}} \qquad (1)$$

and

$$P(F|E) = \frac{P(F \cap E)}{P(E)}$$

but $F \cap E = E \cap F$

so $P(F \cap E) = P(E \cap F)$

$$P(F|E) = \frac{P(E \cap F)}{P(E)} \qquad (2)$$

From (2) we get that

multiply both sides by $P(E)$

$$\boxed{P(F|E)\,P(E) = P(E \cap F)} \qquad (3)$$

Substituting (3) into (1) we get

$$\boxed{P(E|F) = \frac{P(F|E)\,P(E)}{P(F)}} \qquad \textbf{Bayes' formula} \quad (4)$$

version 1

In general $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

if $A \cap B = \emptyset$ i.e. if A and B are disjoint

$P(A \cap B) = 0$

$P(A \cup B) = P(A) + P(B)$

Bayes' theorem

Bayes' formula says that $\boxed{P(E \mid F) = \dfrac{P(F \mid E)\,P(E)}{P(F)}}$. From set theory we have that, given any two sets A and B,

$$A = \underbrace{(A \cap B)}_{} \cup \underbrace{(A \cap \neg B)}_{} \tag{5}$$

disjoint

and the two sets $(A \cap B)$ and $(A \cap \neg B)$ are disjoint. Thus $P(A) = P(A \cap B) + P(A \cap \neg B)$.

Hence in the denominator of Bayes' formula we can replace $P(F)$ with

$F = (F \cap E) \cup (F \cap \neg E)$

Then, using conditional probability we get that $P(F \cap E) =$ and $P(F \cap \neg E) =$

$P(E \cap F) = P(F \mid E)\,P(E)$

$P(F \cap \neg E) = P(F \mid \neg E)\,P(\neg E)$

Thus $\boxed{P(F) = P(F \mid E)\,P(E) + P(F \mid \neg E)\,P(\neg E)}$.

Replacing this in Bayes' formula we get

$$\boxed{P(E \mid F) = \frac{P(F \mid E)\,P(E)}{\underbrace{P(F \mid E)\,P(E) + P(F \mid \neg E)\,P(\neg E)}_{P(F)}}} \tag{6}$$

Bayes' theorem (6)

version 2

version 2

$$\boxed{P(E \mid F) = \frac{P(F \mid E)\,P(E)}{\underbrace{P(F \mid E)P(E) + P(F \mid \neg E)P(\neg E)}_{= P(F)}}}$$

version 1

$$P(E \mid Pass) = \frac{P(Pass \mid E)\,P(E)}{P(Pass)}$$

EXAMPLE.

Enrollment in a class is as follows: 60% econ majors (E), 40% other majors ($\neg E$). In the past, 80% of the econ majors passed and 65% of the other majors passed. A student tells you that she passed the class. What is the probability that she is an econ major? Let P_{ass} stand for "Pass the class".

$P(E)=60\%=\frac{60}{100}$

$P(\neg E)=40\%=\frac{40}{100}$

$P(Pass \mid E)=80\%=\frac{80}{100}$

$\rightarrow P(Pass \mid \neg E)=65\%=\frac{65}{100}$

$$P(E \mid P_{ass}) = \frac{P(Pass \mid E)\ P(E)}{\underbrace{P(Pass \mid E)\,P(E) + P(Pass \mid \neg E)\,P(\neg E)}_{P(Pass)}}$$

$$= \frac{\frac{80}{100} \times \frac{60}{100}}{\frac{80}{100} \times \frac{60}{100} + \frac{65}{100} \times \frac{40}{100}} = 64.86\%$$

Back to previous examples

EXAMPLE 1. Testing for a disease

Base rate of a disease: percentage of the population that has the disease

Sensitivity of a test: percentage of those who have the disease that tests positive

Specificity of a test: percentage of those who **do not** have the disease that tests negative

Suppose:

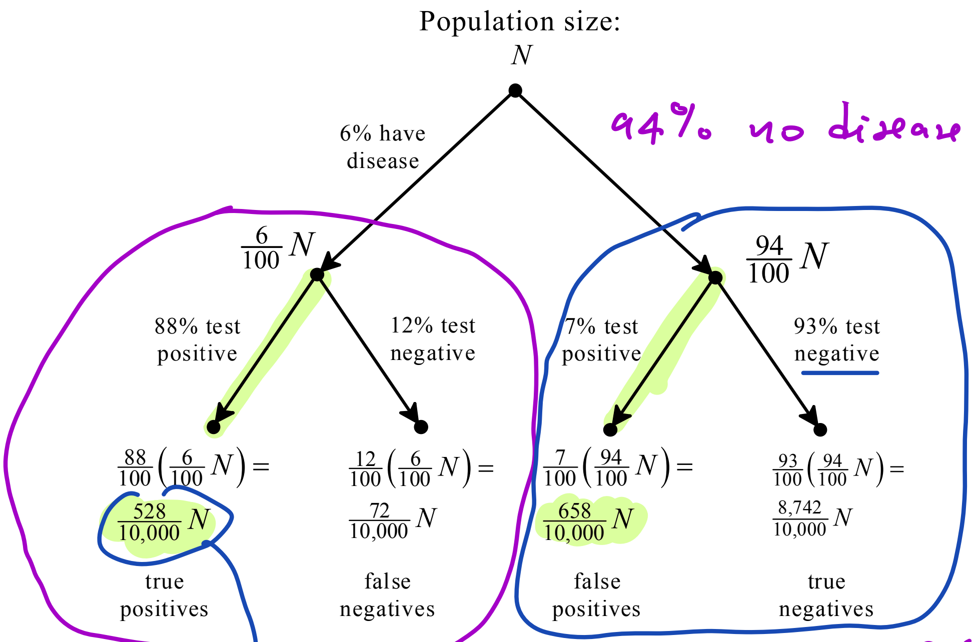
Base rate = 6%

Sensitivity = 88% true positives

Specificity = 93% true negatives

Suppose you test positive. What is the probability that you have the disease?

Previous analysis:



The probability of having the disease, conditional on testing positive is: $P(D|+) = ?$

$$P(D|+) = \frac{528\frac{N}{10,000}}{528\frac{N}{10,000}+658\frac{N}{10,000}} = \frac{528\frac{N}{10,000}}{1,186\frac{N}{10,000}} = \frac{528}{1,186} = 0.4452 = 44.52\%$$

$$P(+|D) = \frac{\frac{528}{10,000}N}{\frac{6}{100}N}$$

$$\frac{P(+|D)\,P(D)}{P(+|D)\,P(D) + \underbrace{P(+|\neg D)\,P(\neg D)}_{=\frac{658}{10,000}N}}$$

$$P(D) = \frac{\frac{6}{100}N}{N}$$

$$= \frac{\frac{94}{100}N}{N}$$

$$\frac{94}{100}N$$

D = have the disease $\neg D =$ do not have disease

$+$ = test positive $-$ = test negative

Base rate = 6% $P(D) = 6\%$ $P(\neg D) = 94\%$

Sensitivity = 88% $P(+|D) = 88\%$ $P(-|D) = 12\%$

Specificity = 93% $P(+|\neg D) = 7\%$ $P(-|\neg D) = 93\%$

By Bayes' rule:

$$P(D|+) = \frac{P(+|D)\,P(D)}{P(+|D)P(D) + P(+|\neg D)\,P(\neg D)}$$

$$= \frac{\frac{88}{100} \times \frac{6}{100}}{\frac{88}{100} \times \frac{6}{100} + \frac{7}{100} \times \frac{94}{100}} = 44.52\%$$

EXAMPLE 2. More than two categories

Base rates of seniors who graduated within the past 6 months:

CS	*Econ*	*Hist*	*Math*
28%	40%	12%	20%

"base rates"

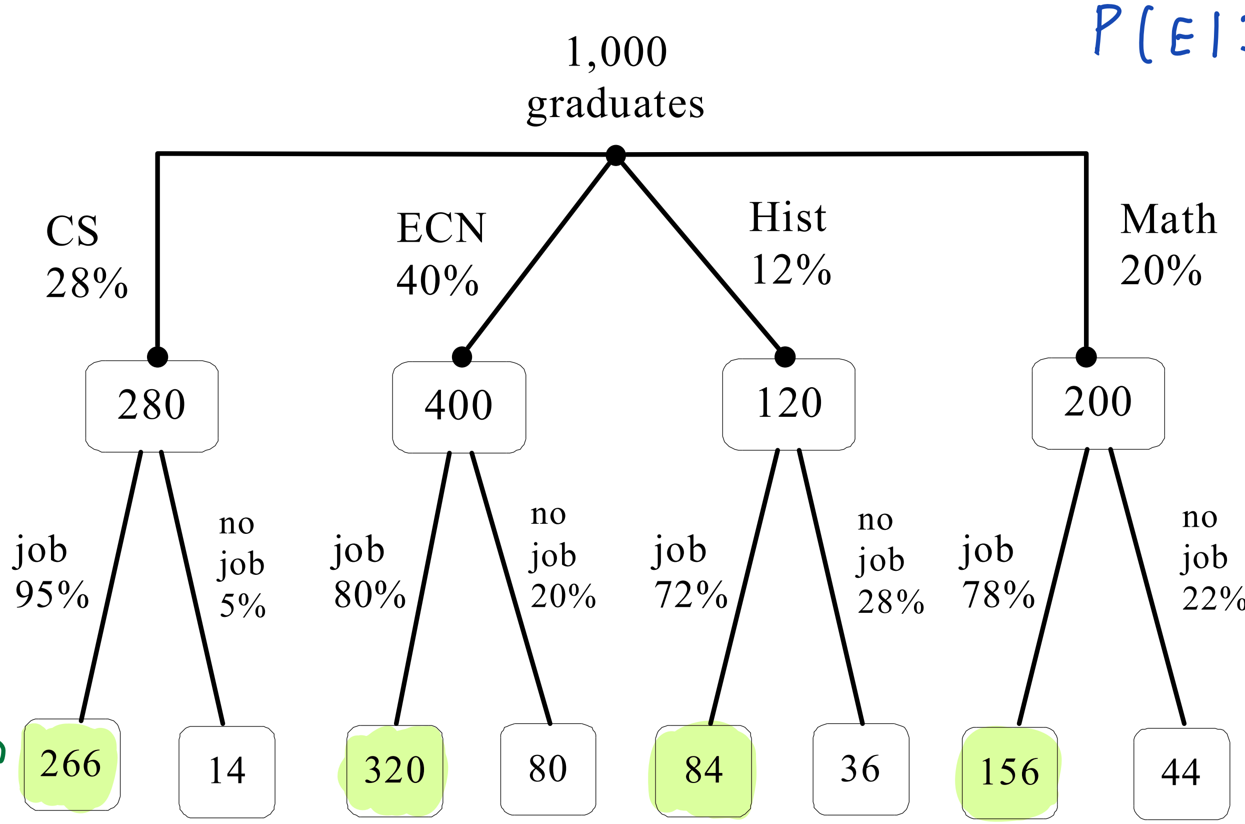
Percentages of those who found a job within 6 months of graduation by major:

Major:	*CS*	*Econ*	*Hist*	*Math*
% who found a job:	95%	80%	~~70%~~ 72%	78%

Conditional probabilities

You learn that Ann graduated 6 months ago and has already found a job. What is the probability that Ann is an Econ major?

P(E | J)



base rates

conditional prob.

The **probability of Ann being and Econ major, given that she found a job is thus**:

$$P(E \mid J) = \frac{320}{266+320+84+156} = \frac{320}{826} = 0.3874 = 38.74\%$$

↑ found a job

CS	Econ	Hist	Math
28%	40%	12%	20%

$P(E) = 40\%$

$P(J|E) = 80\%$

Major:	CS	Econ	Hist	Math
% who found a job:	95%	80%	~~70%~~ 72%	78%

$$P(E|J) = \frac{P(J|E)\cdot P(E)}{\underbrace{P(J)}}$$

$= P(J|CS)\,P(CS) + P(J|E)\,P(E) + P(J|H)\,P(H) + P(J|M)\,P(M)$

Now we need a version of Bayes' rule that allows for more than two conditioning events.

Let S be the set of states and $\{E_1, E_2, ..., E_m\}$ be a partition of S, that is,

- $E_1 \cup E_2 \cup ... \cup E_m = S$
- For all $i, j \in \{1, 2, ...m\}$ with $i \neq j$, $E_i \cap E_j = \varnothing$

$A = (A \cap B) \cup (A \cap \neg B)$

Let $F \subseteq S$ be an arbitrary event. Then

$A = (A \cap E_1) \cup (A \cap E_2) \cup .. \cup (A \cap E_m)$

$F = (F \cap E_1) \cup (F \cap E_2) \cup ... \cup (F \cap E_m)$, all disjoint events. Thus

$$\boxed{P(F) = \underbrace{P(F \cap E_1)} + \underbrace{P(F \cap E_2)} + ... + \underbrace{P(F \cap E_m)}}$$

generalizes $P(F) = P(F \cap E) + P(F \cap \neg E)$

Hence, $P(E_i \mid F) =$

$$P(E|J) = \frac{\frac{80}{100} \times \frac{40}{100}}{\frac{95}{100} \times \frac{28}{100} + \frac{80}{100} \times \frac{40}{100} + \frac{72}{100} \times \frac{12}{100} + \frac{78}{100} \times \frac{20}{100}} = 38.74\%$$

$$\underbrace{S = \{a, b, c, d, e, f, g\}}_{\text{shrinks to}}$$

$\Rightarrow$ information $\overbrace{E = \{a, c, e, g\}}$

interested in prob. of $A = \{a, e, f\}$

$P(A|E)$