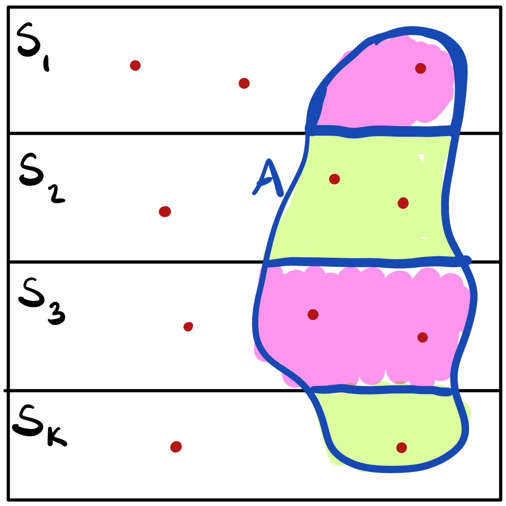
§7.6 Bayes' Theorem

Law of Total Probability.



Sample space S broken into k events

$$S_1, S_2, \dots, S_k$$

such that all mutually exclusive, &

$$S_1 \cup S_2 \cup \dots \cup S_k = S.$$

i.e. every possible outcome is contained in 1 & only 1 event S_i.

Then given any event A,

$$n(A) = n(A \cap S_1) + n(A \cap S_2) + \dots + n(A \cap S_k)$$

$$\frac{n(A)}{n(S)} = \frac{n(A \cap S_1)}{n(S)} + \frac{n(A \cap S_2)}{n(S)} + \dots + \frac{n(A \cap S_k)}{n(S)}$$

$$P(A) = P(A \cap S_1) + P(A \cap S_2) + \dots + P(A \cap S_k)$$

$$\boxed{P(A) = P(S_1)P(A|S_1) + P(S_2)P(A|S_2) + \dots + P(S_k)P(A|S_k)}$$

Law of Total Probability.

Removing conditions!

ex. Suppose a marketing company gives a survey to its client's customers and learns that

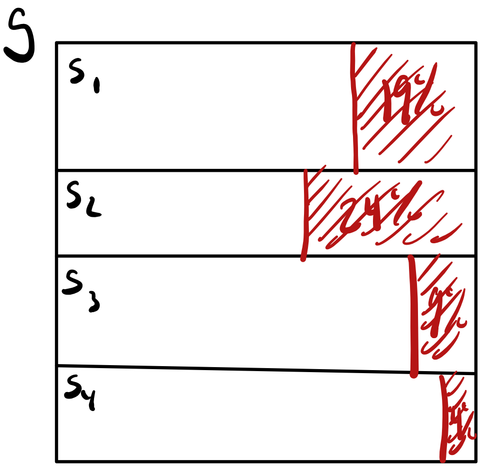
30% of its customers are under age 18,
25% are 18-30,
25% 31-50,
20% are over 50.

BREAKING POPULATION OF CUSTOMERS INTO SUBPOPULATIONS
DEMOGRAPHIC GROUPS
MUTUALLY EXCLUSIVE CATEGORIES

Of the customers less than 18, 19% follow the client on IG.
Of the customers 18-30, 24% follow the client on IG.
Of the customers 31-50, 9% follow the client on IG.
Of the customers over 50, 4% follow the client on IG.

$P(F \mid S_1) = .19$
$P(F \mid S_2) = .24$
$P(F \mid S_3) = .09$
$P(F \mid S_4) = .04$

What percent of the client's customers follow them on IG?



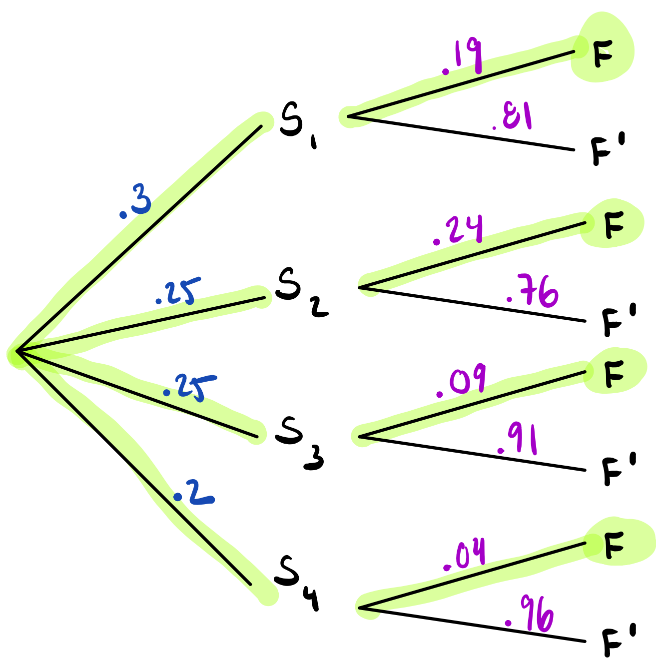
S_1 : AGE < 18 $\qquad P(S_1) = .3$

S_2 : 18 ≤ AGE ≤ 30 $\qquad P(S_2) = .25$

S_3 : 31 ≤ AGE ≤ 50 $\qquad P(S_3) = .25$

S_4 : 51 ≤ AGE $\qquad P(S_4) = .2$

LoTP



$$P(F) = P(S_1)P(F \mid S_1) + P(S_2)P(F \mid S_2) + P(S_3)P(F \mid S_3) + P(S_4)P(F \mid S_4)$$

$$= \underbrace{(.3)(.19) + (.25)(.24) + (.25)(.09) + (.2)(.04)}_{\text{WEIGHTED AVERAGE!}}$$

$$= .1475$$

Q: WHAT IS THE PROBABILITY THAT A CUSTOMER THAT FOLLOWS YOUR CLIENT ON IG IS UNDER 18?

6 $P(S_1 \mid F) = \dfrac{P(S_1)P(F \mid S_1)}{P(F)} = \dfrac{(.3)(.19)}{.1475} = .3864$

LoTP

Note: $P(A \mid B) = \underbrace{\dfrac{P(A \cap B)}{P(B)}}_{\text{Definition of Cond. Prob.}} = \overbrace{\dfrac{P(A)P(B \mid A)}{P(B)}}^{\text{Product Rule for Probability}}$ ✓

$\dfrac{\cancel{P(B)}P(A \mid B)}{\cancel{P(B)}}$ (True but useless)

$$\boxed{P(A \mid B) = \dfrac{P(A)P(B \mid A)}{P(B)}}$$

BAYES' RULE

Flipped the condition!
(opposite conditional probabilities)

Economics *Data from the Bureau of Labor Statistics indicates that for the year 2012, 40.8% of the labor force had a high school diploma or fewer years of education, 27.7% had some college or an associate's degree, and 31.5% had a bachelor's degree or more education. Of those with a high school diploma or fewer years of education, 17.0% earned \$75,000 or more annually. Of those with some college or an associate's degree, 31.4% earned \$75,000 or more, and of those with a bachelor's degree or more education, 55.0% earned \$75,000 or more. Find the following probabilities that a randomly chosen labor force participant has the given characteristics.*

11. Some college or an associate's degree, given that he or she earns \$75,000 or more annually

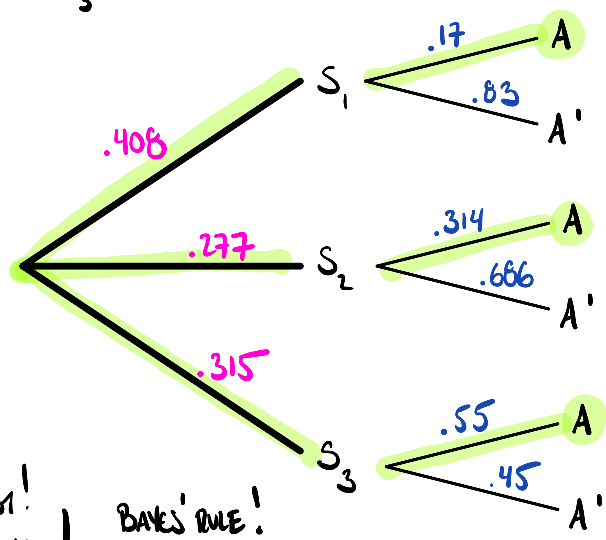
12. A bachelor's degree or more education, given that he or she earns \$75,000 or more annually

13. A high school diploma or less education, given that he or she earns *less* than \$75,000 annually

14. A bachelor's degree or more education, given that he or she earns *less* than \$75,000 annually

S_1 : High School
S_2 : Some College
S_3 : Bach. Degree

A = Earn ≥ \$75,000



$P(A \mid S_1) = .17$

$P(A \mid S_2) = .314$

$P(A \mid S_3) = .55$

Different! But related!

Bayes' Rule!

$P(● \mid ●) = \dfrac{P(●)P(● \mid ●)}{P(●)}$

11. $P(S_2 \mid A) = \dfrac{P(S_2)P(A \mid S_2)}{P(A)}$ ✓ ✓

LoTP

$$P(S_2 \mid A) = \frac{P(S_2)P(A \mid S_2)}{P(S_1)P(A \mid S_1) + P(S_2)P(A \mid S_2) + P(S_3)P(A \mid S_3)}$$

$$= \frac{(.277)(.314)}{(.408)(.17) + (.277)(.314) + (.315)(.55)} = \boxed{.2639}$$

12. $$P(S_3 \mid A) = \frac{P(S_3)P(A \mid S_3)}{P(A)}$$

$$P({\color{red}\bullet} \mid {\color{blue}\bullet}) = \frac{P({\color{red}\bullet})P({\color{blue}\bullet} \mid {\color{red}\bullet})}{P({\color{blue}\bullet})}$$

$$= \frac{(.315)(.55)}{(.408)(.17) + (.277)(.314) + (.315)(.55)}$$

$$= .5257$$

$P(S_3) = .315$

$P(S_3 \mid A) = .5257$