ANGEL → SEE RAW SCORES → LAB
8 1 1 13

grade[8,1,1,.. 13, ..,2,2,2]
THRU EXAM 2 (SCORES)

SEC 1 SNAFU F Q2
GIVE YOU Q2 RAW SCORE

Raoul LePage
Professor
STATISTICS AND PROBABILITY
www.stt.msu.edu/~lepage
click on STT200_Sp09

3-18-09

GOLD STANDARD ANSWER

{\$1a, \$1b, \$5c}

Jack draws a bill first

Jill draws second

from the **two** bills then remaining

		Jill		
		a	b	c
	a	elim		*
Jack	b		elim	*
	c			elim

all 6 possibilities

P(Jill \$5)

= 2 / 6 = 1 / 3

same as Jack

MULT RULE $P(AB) = P(A)P(B|A)$

GOLD STANDARD ANSWER

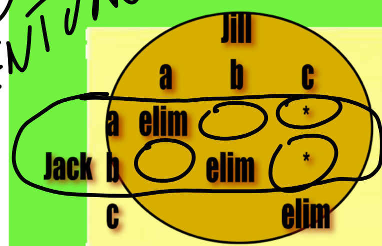
{\$1a, \$1b, \$5c}

Jack draws a bill first

Jill draws second

from the **two** bills then remaining

CONDITIONING EVENT JACK 1



P(Jill \$5)
= 2 / 6 = 1 / 3
same as Jack

all 6 possibilities

DEF OF COND'L

$$\boxed{P(B|A) = P(AB)/P(A)}$$

CONDITIONING

TOLD JACK 1.
NEW CONDITIONAL
(REVISED) PROBY
FOR JILL 5. GIVEN
INFO JACK GOT 1.
WRITE $P(\text{JILL }5 | \text{JACK }1)$

GIVEN

THINK

JILL FACES

1 5

INTUIT $P(\text{JILL }5 | \text{JA }1) = \frac{1}{2}$

DEFINE THIS COND'L PROBY AS REGULAR PROBY IN REDUCED MODEL $2/4 = \frac{1}{2}$

FROM REG PROBY $2/6 \,/\, (4/6) = P(\text{JA1 JILL5}) / P(\text{JACK1})$

EX 1. GIVEN INFO $P(A)=.1$ $P(B)=.3$

$P(AB)=.06$

$AB \subseteq A$

$AB \subseteq B$

A .1 .9

SO $P(AB)=.06$

$P(AB^c)=.04$

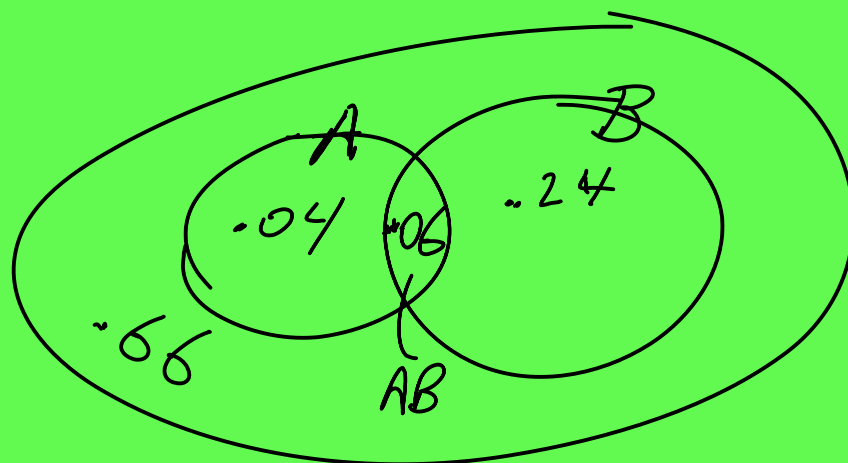
$\boxed{P(A^cB)=.24}$

$P(A^cB^c)=.66$

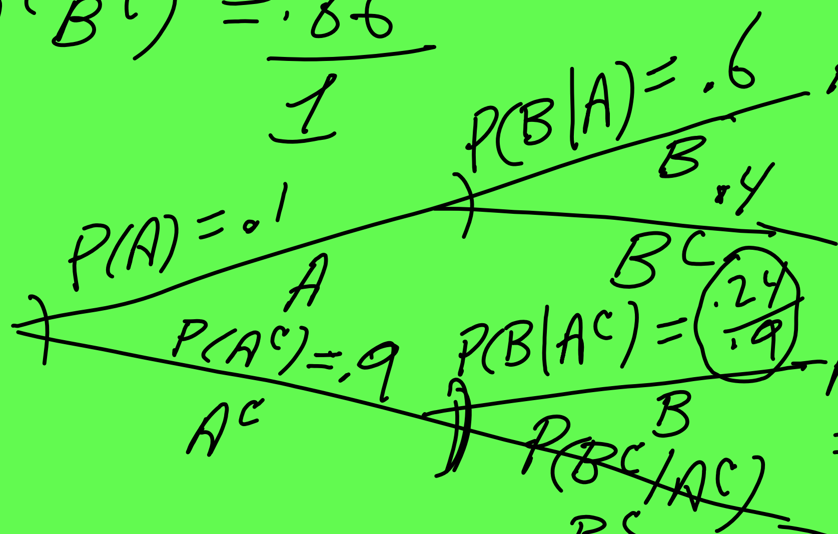
$\underline{}$

1

VENN



TREE



$AB = .1(.6)$

$AB^c = .1(.4)$

$A^cB = .9\left(\frac{.24}{.9}\right)$

$= .24$

$= 1 - \frac{.24}{.9}$

DEF $P(B|A)$

$= P(AB)/P(A)$

$= .06/.1 = .6$

$P(B|A^c)$

$= P(A^cB)/P(A^c)$

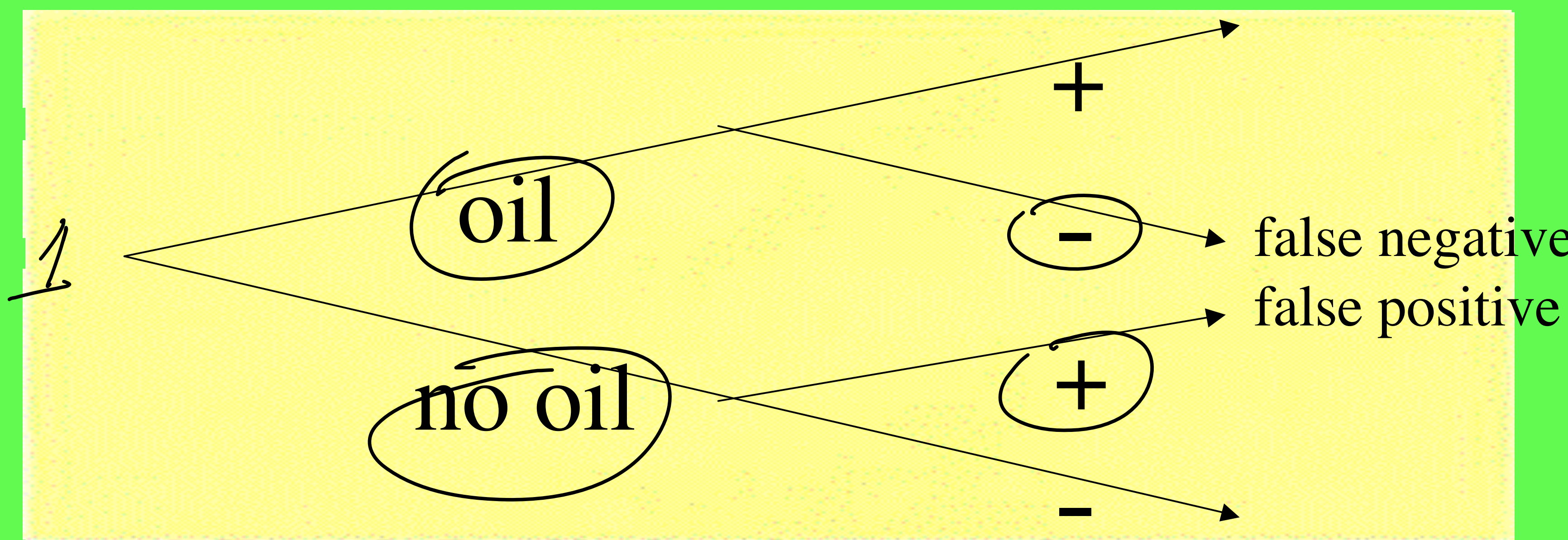
$= .24/.9$

TREE DIAGRAM

"oil" = oil is present

"+" = a test for oil is positive

"–" = a test for oil is negative

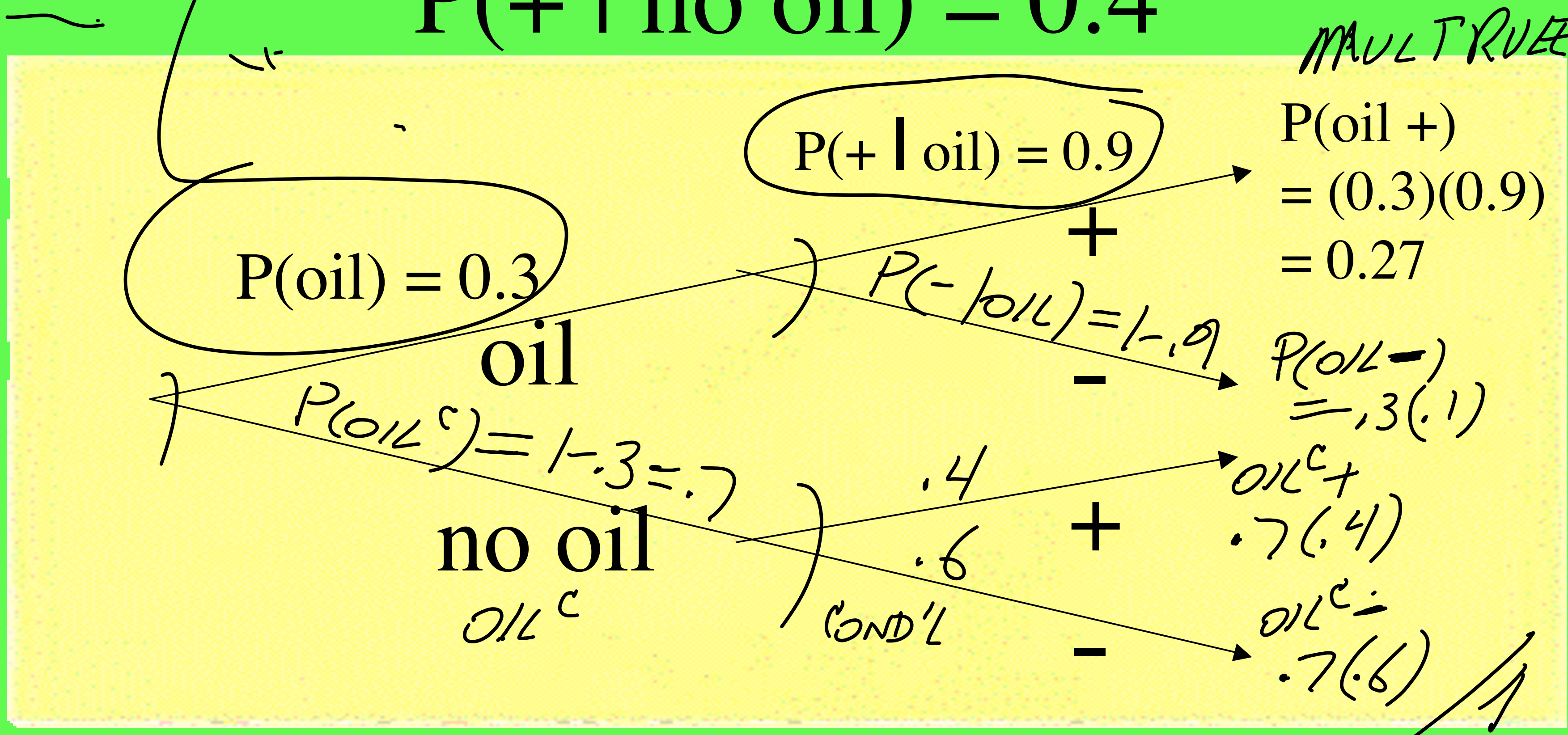


TREE DIAGRAM CONVENTIONS

P(oil) = 0.3

P(+ | oil) = 0.9

P(+ | no oil) = 0.4

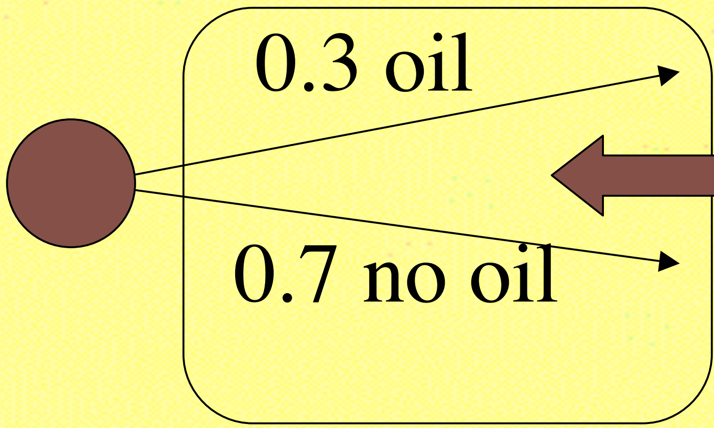


TOTAL OF BRANCHES = 1

$$P(oil) = 0.3$$

$$P(+ \mid oil) = 0.9$$

$$P(+ \mid no\ oil) = 0.4$$



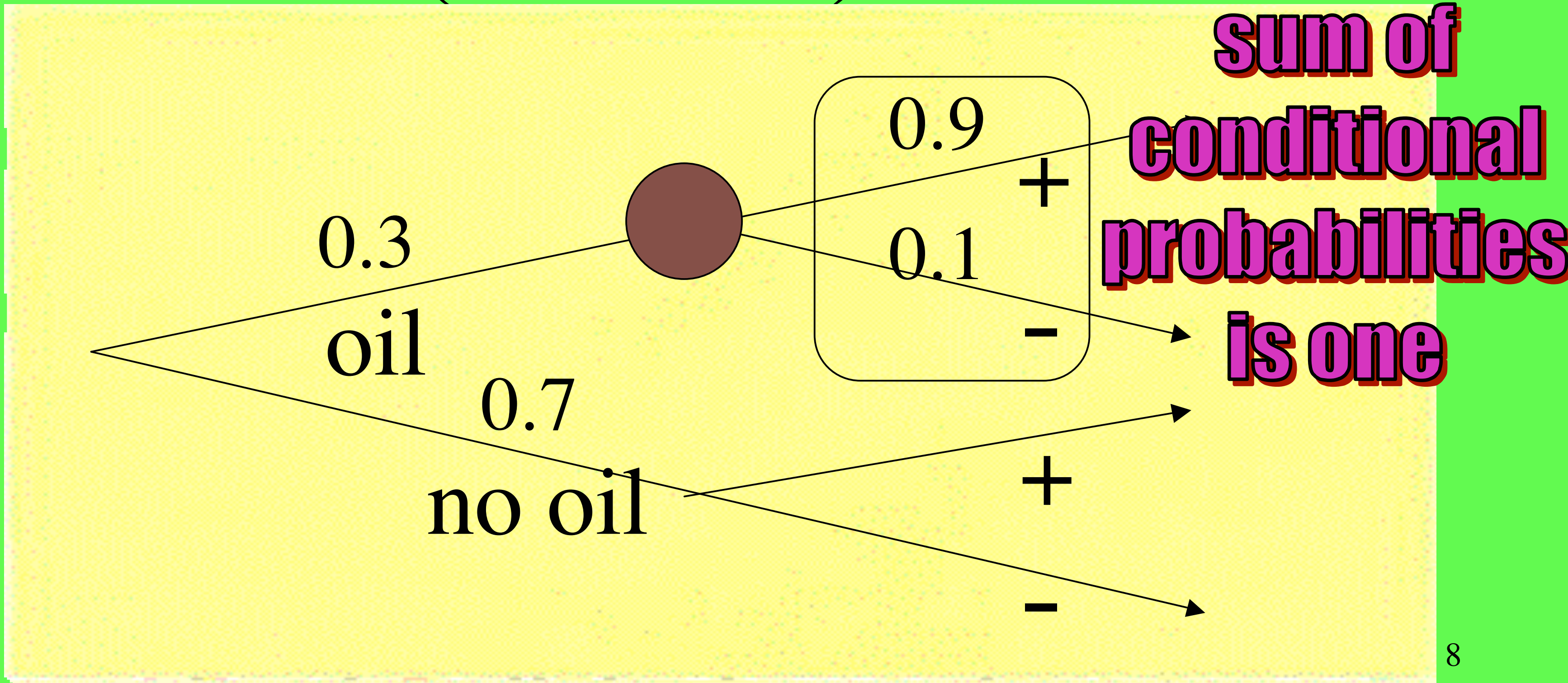
sum of unconditional probabilities is one

TOTAL OF CONDITIONAL BRANCHES = 1

$$P(\text{oil}) = 0.3$$

$$\boxed{P(+ \mid \text{oil}) = 0.9 \qquad P(- \mid \text{oil}) = 0.1}$$

$$P(+ \mid \text{no oil}) = 0.4$$

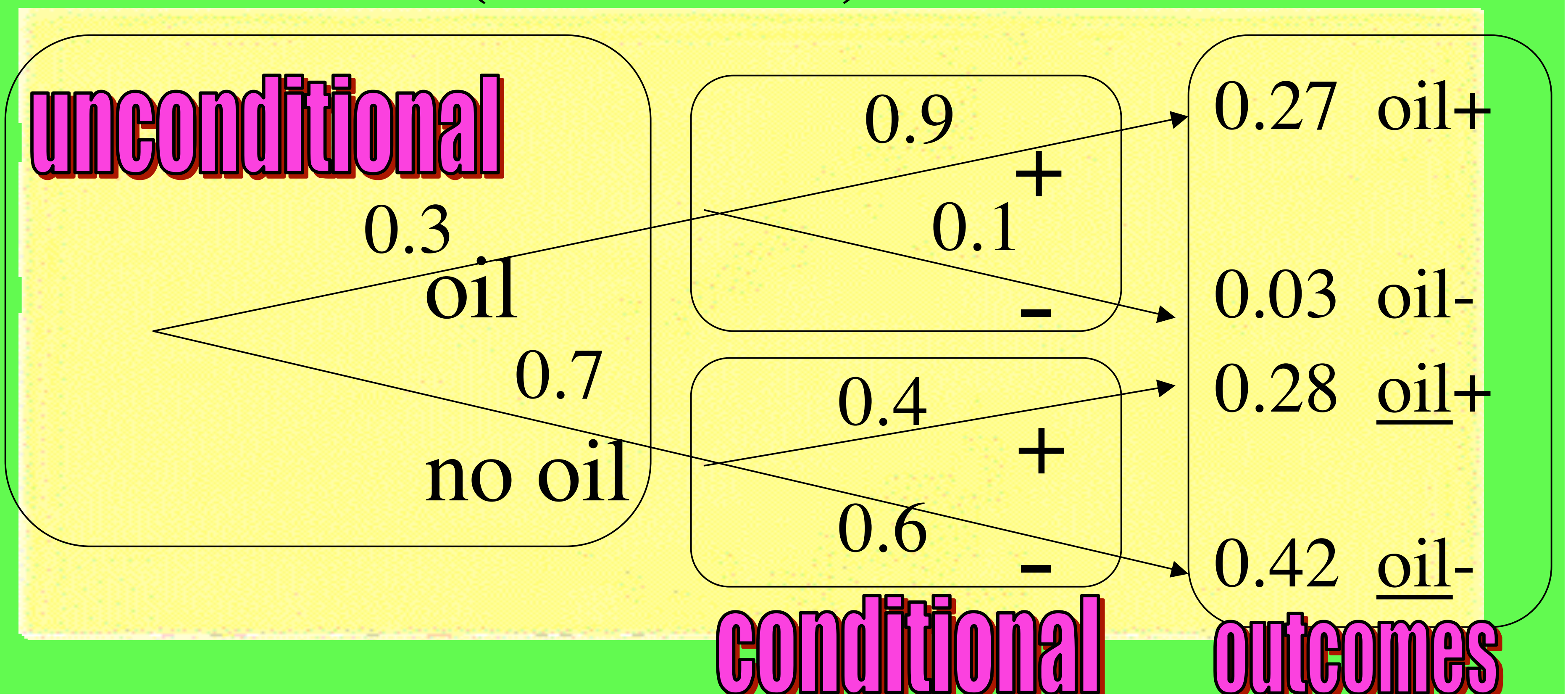


COMPLETE TREE

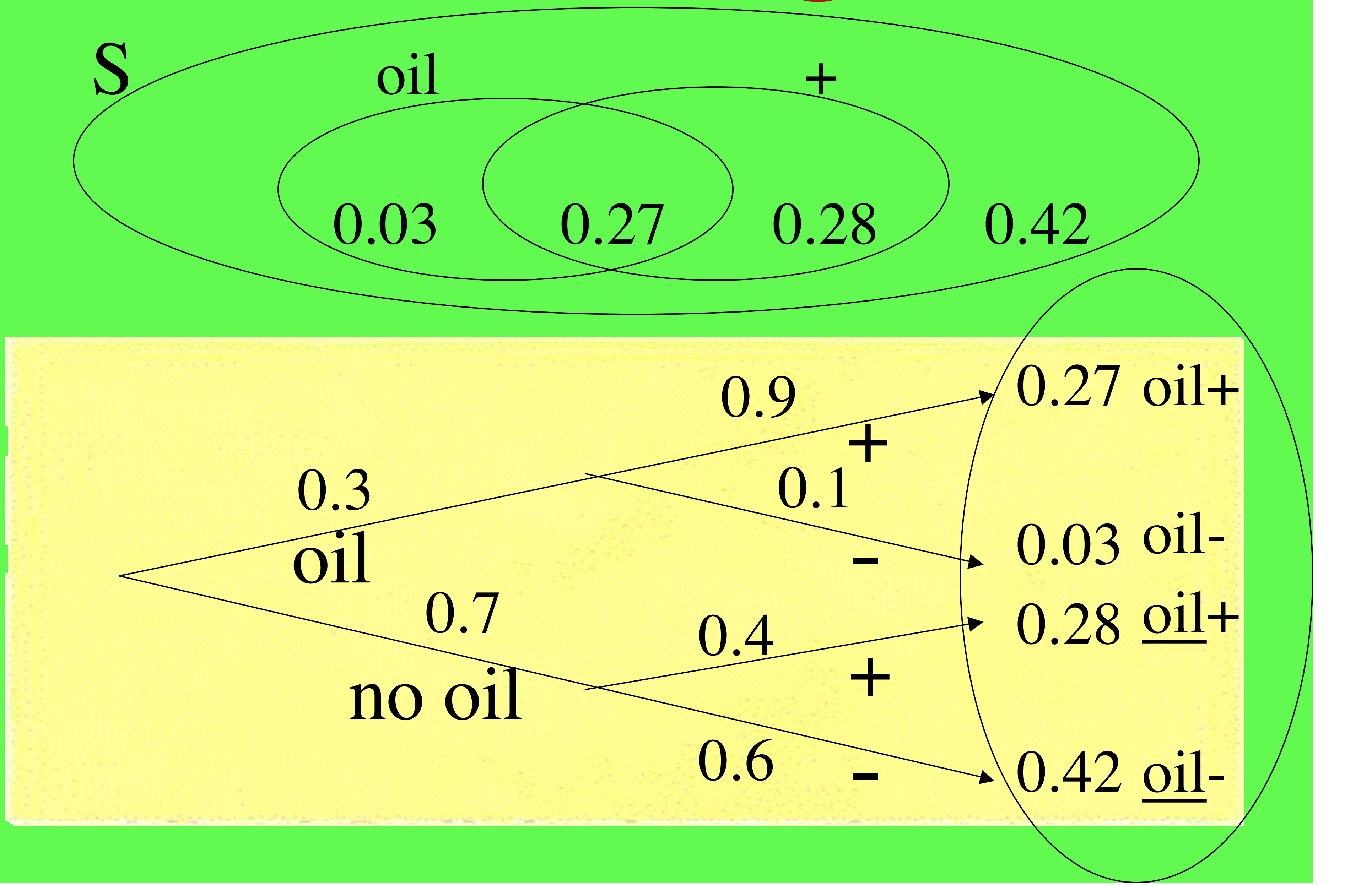
$$P(\text{oil}) = 0.3$$

$$P(+ \mid \text{oil}) = 0.9$$

$$P(+ \mid \text{no oil}) = 0.4$$



VENN DIAGRAM



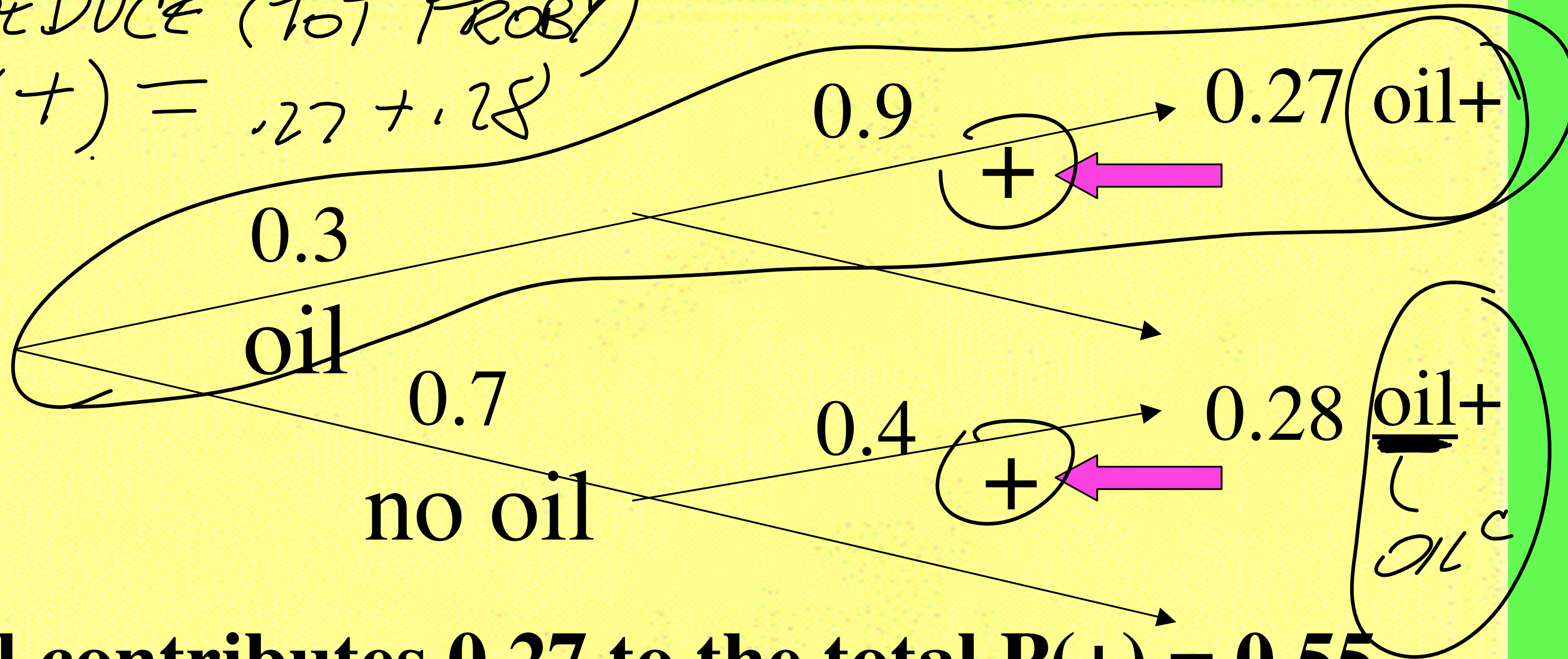
TOTAL PROBABILITY

$$P(+) = P(\text{oil}+) + P(\text{no oil}+)$$

$$0.55 = 0.27 + 0.28$$

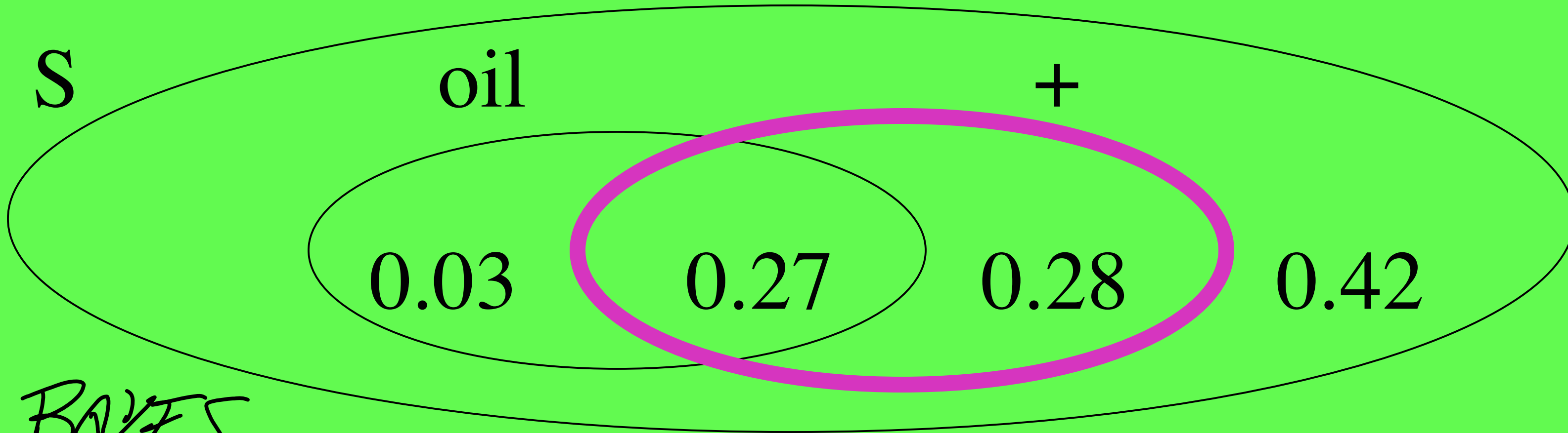
DEDUCE (TOT PROBY)

$P(+) = .27 + .28$



Oil contributes 0.27 to the total P(+) = 0.55.

BAYES FORMULA



BAYES

$$P(oil \mid +) = P(oil+) \,/\, P(+)$$

$$= 0.27 \,/\, (0.27 + 0.28)$$

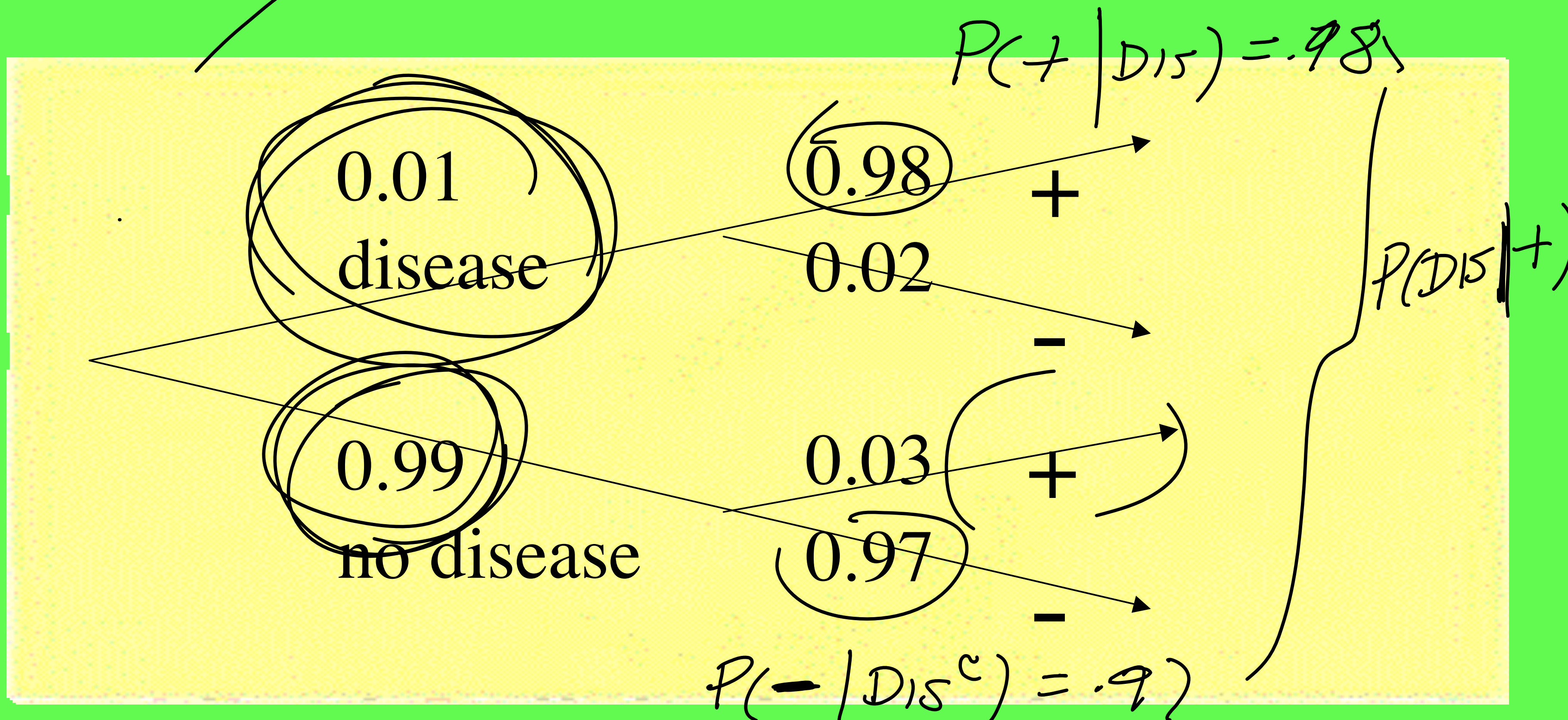
$$= 0.4909..$$

0.27 oil+

0.28 $\underline{oil}$+

Oil contributes 0.27 of the total P(+) = 0.27+0.28.

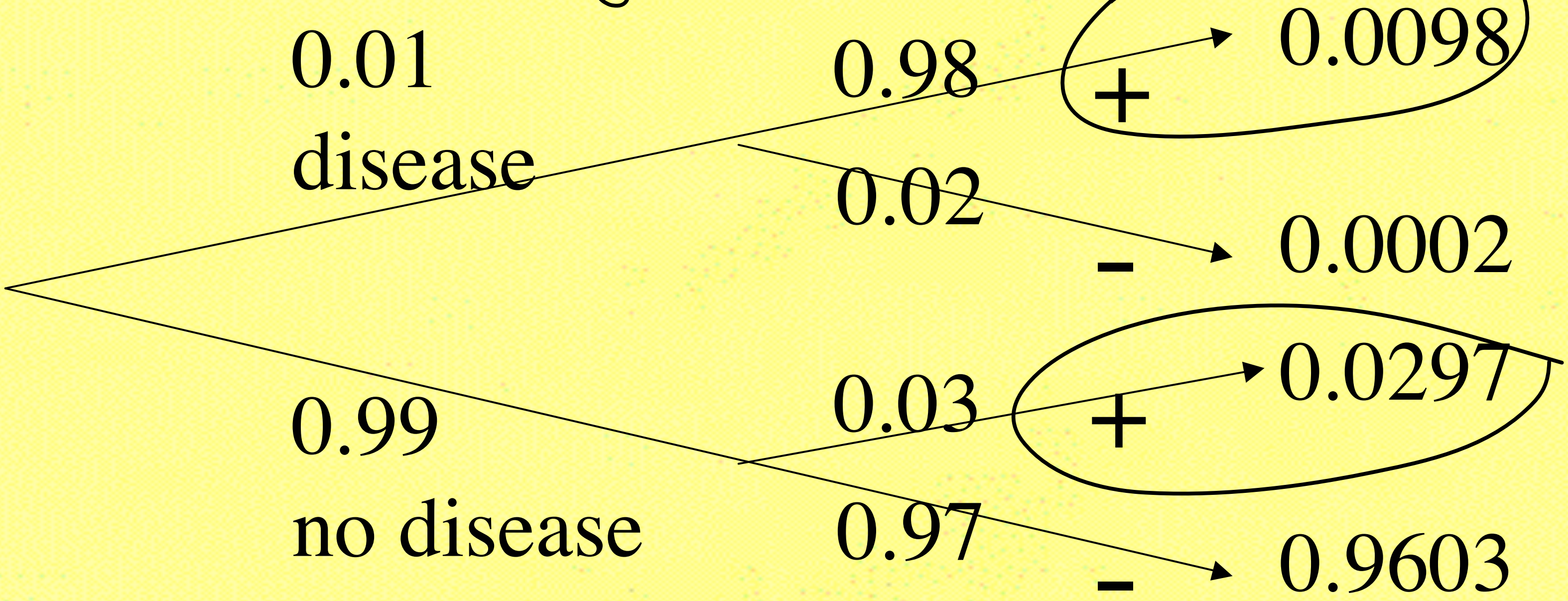
MEDICAL TEST



The test for this infrequent disease seems to be reliable having only 3% false positives and 2% false negatives. **What if we test positive?**

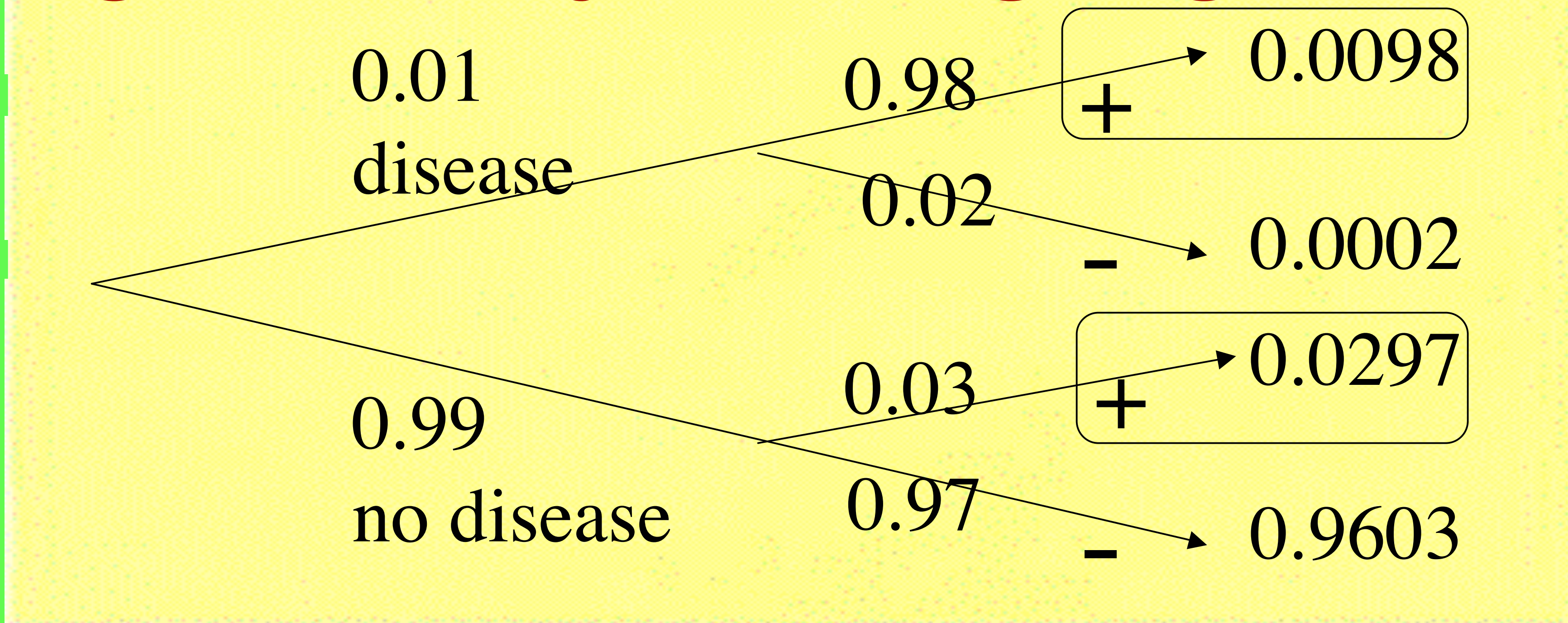
MEDICAL TEST

$$P(DIS\,|+) = \frac{.0098}{.0098 + .0297} \approx$$



We need to calculate P(diseased | +), the **conditional probability** that we have this disease **GIVEN** we've tested positive for it.

CALCULATING OUR CHANCES OF HAVING THE DISEASE IF +



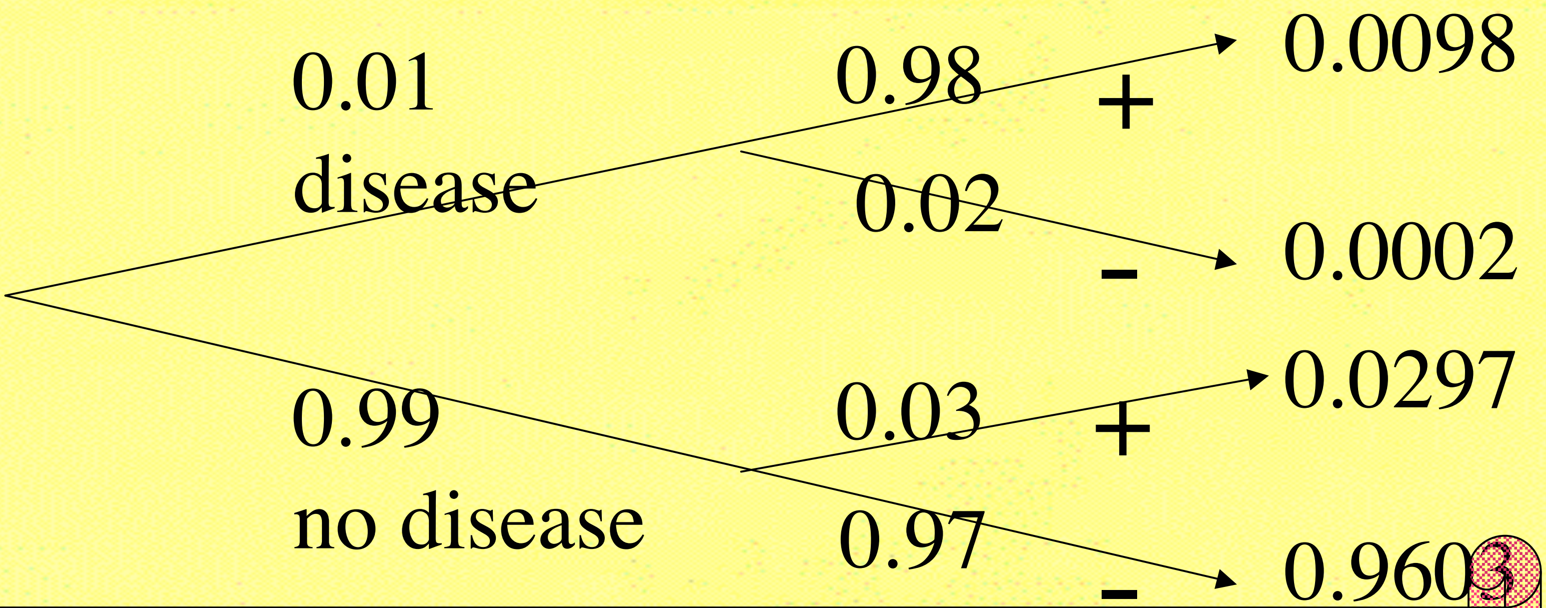
P(+) = 0.0098 + 0.0297 = 0.0395

P(disease | +) = P(disease+) / P(+)

= .0098 / 0.0395 = 0.248. only 25% !

FALSE POSITIVE PARADOX

one may overwhelm a good test by failing to screen

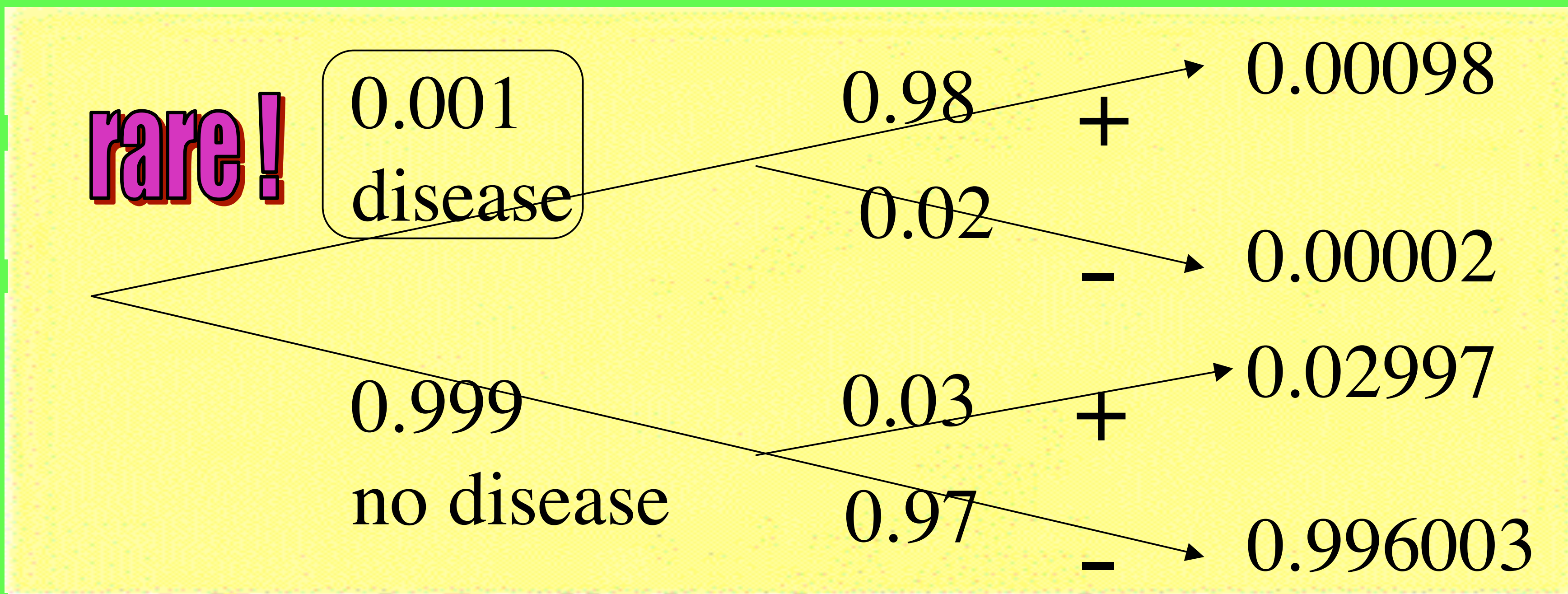


EVEN FOR THIS ACCURATE TEST:

P(diseased | +) is only around 25% because the non-diseased group is **so predominant** that most positives come from it.

FALSE POSITIVE PARADOX

one may overwhelm a good test by failing to screen



WHEN THE DISEASE IS TRULY RARE: P(diseased | +) is a mere **3.2%** because the huge non-diseased group **has completely over-whelmed** the test, which no longer has value

IMPLICATIONS OF THE PARADOX

FOR MEDICAL PRACTICE: Good diagnostic tests will be of little use if the system is over-whelmed by lots of healthy people taking the test. **Screen patients** first.

FOR BUSINESS: Good sales people capably focus their efforts on likely buyers, leading to increased sales. They can be rendered ineffective by feeding them too many false leads, as with massive **un-targeted sales promotions**.