ANGEL → SEE RAW SCORES → LAB grade [8, 1, 1, 13, 12, 2, 2]

SEC 1 SNAFU F 02 GIVE YOU #2 RAW SCORE

THRU Exam 2 (scores)

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

P( Jill $5 )
= 2 \div 6 = \frac{1}{3}

same as Jack
CONDITIONING TOLD JACK 1.
NEW CONDITIONAL PROB
FOR JILL 5. GIVEN INFO JACK GOT 1.
WRITE P(JILL 5 | JACK 1)

THINK = JILL FACES

INTUIT P(JILL 5 | JACK 1) = \frac{1}{5}

DEF OF COND.

DEFINE THIS COND. PROB AS REGULAR PROB IN REDUCED MODEL $\frac{2}{4} = \frac{1}{2}$

FIND P(JILL 5 | JACK 1) REG PROB $\frac{1}{6} / (\frac{2}{4}) = \frac{P(JILL 5 | JACK 1)}{P(1 | JACK 1)}$

MUT 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

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

all 6 possibilities

DEF OF COND.
EX 1. GIVEN INFO \( P(A) = .1 \), \( P(B) = .3 \)
\[ P(AB) = .06 \]

\( A \subseteq A \)
\( A \cap B \subseteq B \)

So \( P(AB) = .06 \)
\[ P(AB^c) = .04 \]
\[ P(A^cB) = .24 \]
\[ P(A^cB^c) = .66 \]

\[ P(B | A) = .6 \]
\[ AB = \frac{P(AB)}{P(A)} = 0.6/0.1 = .6 \]

DEF \( P(B | A) \)
\[ P(B | A^c) \]
\[ P(B | A^c) = 0.4 \]
\[ AB^c = \frac{P(A^cB)}{P(A^c)} = \frac{0.24}{0.9} \]

\[ P(B | A^c) = \frac{P(A^cB)}{P(A^c)} = \frac{0.24}{0.9} \]
“oil” = oil is present

“+” = a test for oil is positive

“−” = a test for oil is negative

- false negative
- false positive
\[ P(\text{oil}) = 0.3 \]
\[ P(+ | \text{oil}) = 0.9 \]
\[ P(+ | \text{no oil}) = 0.4 \]

\[ P(\text{oil} +) = (0.3)(0.9) = 0.27 \]
TOTAL OF BRANCHES = 1

\[ P(\text{oil}) = 0.3 \]
\[ P(\text{+ | oil}) = 0.9 \]
\[ P(\text{+ | no oil}) = 0.4 \]

sum of unconditional probabilities is one
TOTAL OF CONDITIONAL BRANCHES = 1

\[
P(\text{oil}) = 0.3
\]

\[
\begin{align*}
P(+) \mid \text{oil} &= 0.9 & P(-) \mid \text{oil} &= 0.1 \\
P(+) \mid \text{no oil} &= 0.4
\end{align*}
\]

The sum of conditional probabilities is one.
P(oil) = 0.3
P(+ | oil) = 0.9
P(+ | no oil) = 0.4
VENN DIAGRAM

S

oil

0.03 0.27 0.28 0.42

oil

0.3 0.7

no oil

0.9 0.1 0.4 0.6

0.27 oil+

0.03 oil-

0.28 oil+

0.42 oil-
Oil contributes 0.27 to the total $P(+) = 0.55$. 

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

0.55 = 0.27 + 0.28
Oil contributes 0.27 of the total $P(+) = 0.27 + 0.28$. 

$$P(\text{oil} \mid +) = \frac{P(\text{oil} +)}{P(+)} = \frac{0.27}{0.27 + 0.28} = 0.4909..$$
The test for this infrequent disease seems to be reliable having only 3% false positives and 2% false negatives. **What if we test positive?**
We need to calculate $P(\text{diseased} \mid +)$, the **conditional probability** that we have this disease **GIVEN** we’ve tested positive for it.
CALCULATING OUR CHANCES OF HAVING THE DISEASE IF +

\[ P(+) = 0.0098 + 0.0297 = 0.0395 \]
\[ P(\text{disease} \mid +) = \frac{P(\text{disease+})}{P(+)} = \frac{0.0098}{0.0395} = 0.248. \]
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:
\( P(\text{diseased} | +) \) is a mere \( 3.2\% \) because the huge non-diseased group has completely overwhelmed the test, which no longer has value
FOR MEDICAL PRACTICE: Good diagnostic tests will be of little use if the system is overwhelmed 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.