CS2210 Discrete Structures
Discrete Probability

Fall 2017

Sukumar Ghosh

Sample Space

DEFINITION. The **sample space S** of an experiment is the set of possible outcomes. An **event E** is a subset of the sample space.

What is probability?

- The probability of an event occurring is:

$$p(E) = \frac{|E|}{|S|}$$

  - Where E is the set of desired events (outcomes) and
  - S is the set of all possible events (outcomes)
  - Note that $0 \leq |E| \leq |S|$
    - Thus, the probability will always between 0 and 1
    - An event that will never happen has probability 0
    - An event that will always happen has probability 1

Probability distribution

Consider an experiment where there are n possible outcomes $x_1, x_2, x_3, x_4, \ldots x_n$. Then

1. $0 \leq p(x_i) \leq 1$ $(1 \leq i < n)$
2. $p(x_1) + p(x_2) + p(x_3) + p(x_4) + \ldots + p(x_n) = 1$

You can treat p as a *function* that maps the set of all outcomes to the set of real numbers. This is called the probability distribution function.

Probability of independent events

- When two events E and F are independent, the occurrence of one gives no information about the occurrence of the other.

- The probability of two independent events

$$p(E \cap F) = p(E) \cdot p(F)$$

Example of dice

What is the probability of two 1's on two six-sided dice?

- Probability of getting a 1 on a 6-sided die is 1/6
- Via product rule, probability of getting two 1's is the probability of getting a 1 AND the probability of getting a second 1
- Thus, it's 1/6 * 1/6 = 1/36

What is the probability of getting a 7 by rolling two dice?

- There are six combinations that can yield 7: (1,6), (2,5), (3,4), (4,3), (5,2), (6,1)
- Thus, |E| = 6, |S| = 36, P(E) = 6/36 = 1/6

Example from Card games

There are (13 x 4) = 52 cards in a pack

Poker game: Royal flush

- What is the chance of getting a royal flush?
  - That's the cards 10, J, Q, K, and A of the same suit

- There are only 4 possible royal flushes
- Possibilities for 5 cards: C(52,5) = 2,598,960
- Probability = 4/2,598,960 = 0.0000015
  - Or about 1 in 650,000

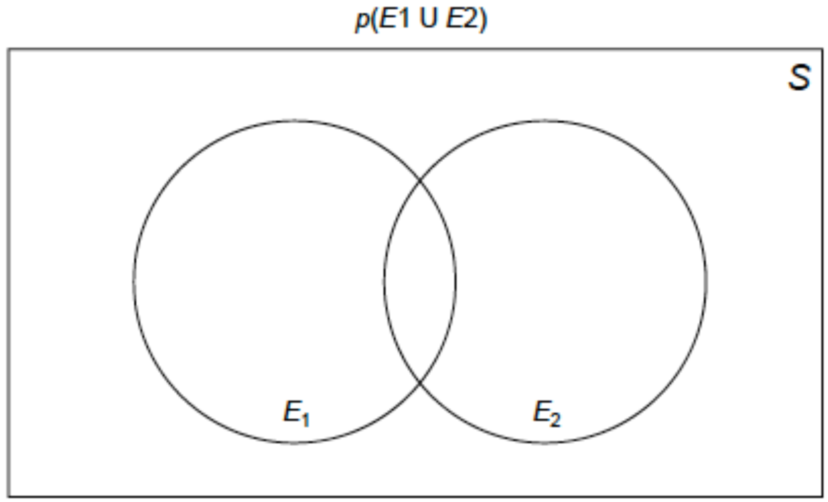
More on probability

- Let E be an event in a sample space S. The probability of the complement of E is:

$$p(\overline{E}) = 1 - p(E)$$

- Recall the probability for getting a royal flush is 0.0000015
  - The probability of *not* getting a royal flush is 1-0.0000015 or 0.9999985

Probability of the union of events



- Let E_1 and E_2 be events in sample space S
- Then $p(E_1 \cup E_2) = p(E_1) + p(E_2) - p(E_1 \cap E_2)$
- Consider a Venn diagram dart-board

Example

- If you choose a number between 1 and 100, what is the probability that it is divisible by 2 or 5 or both?
- Let n be the number chosen
  - $p(2|n) = 50/100$ (all the even numbers)
  - $p(5|n) = 20/100$
  - $p(2|n)$ and $p(5|n) = p(10|n) = 10/100$
  - $p(2|n)$ or $p(5|n) = p(2|n) + p(5|n) - p(10|n)$

    $= 50/100 + 20/100 - 10/100$

    $= 3/5$

When is gambling worth?

Disclaimer. *This is a statistical analysis, not a moral or ethical discussion*

- What if you gamble $1, and have a ½ probability to win $10?
  - If you play 100 times, you will win (on average) 50 of those times
    - Each play costs $1, each win yields $10
    - For $100 spent, you win (on average) $500
  - Average win is $5 (or $10 * ½) per play for every $1 spent
- What if you gamble $1 and have a 1/100 probability to win $10?
  - If you play 100 times, you will win (on average) 1 of those times
    - Each play costs $1, each win yields $10
    - For $100 spent, you win (on average) $10
  - Average win is $0.10 (or $10 * 1/100) for every $1 spent

Powerball lottery

Disclaimer. This is a statistical analysis, not a moral or ethical discussion

- Modern powerball lottery is a bit different
  - Source: http://en.wikipedia.org/wiki/Powerball
- You pick 5 numbers from 1-55
  - Total possibilities: C(55,5) = 3,478,761
- You then pick one number from 1-42 (the powerball)
  - Total possibilities: C(42,1) = 42
- By the product rule, you need to do both
  - So the total possibilities is 3,478,761* 42 = 146,107,962
- While there are many "sub" prizes, the probability for the jackpot is about 1 in 146 million
  - You will "break even" if the jackpot is $146M
  - Thus, one should only play if the jackpot is greater than $146M

Conditional Probability

You are flipping a coin 3 times. The first flip is a tail. Given this, what is the probability that the 3 flips produce an odd number of tails?

Deals with the probability of an event E when another event F has *already occurred*. The occurrence of F actually shrinks the sample space.

Given F, the probability of E is
$p(E|F) = p(E \cap F) / p(F)$

Conditional Probability

Sample space S = {TTT, THH, THT, TTH, HTT, HHH, HHT, HTH}
F = {TTT, THH, THT, TTH}   (the reduced sample space)
E = {TTT, THH}   {the target event set)

$p(E \cap F) = 2/8$,
$p(F) = 4/8$.

So $p(E|F) = p(E \cap F) / p(F) = 1/2$

Example of Conditional Probability

What is the probability that a family with two children has two boys, given that *they have at least one boy*?

F = {BB, BG, GB}
E = {BB}

If the four events {BB, BG, GB, GG} are equally likely, then

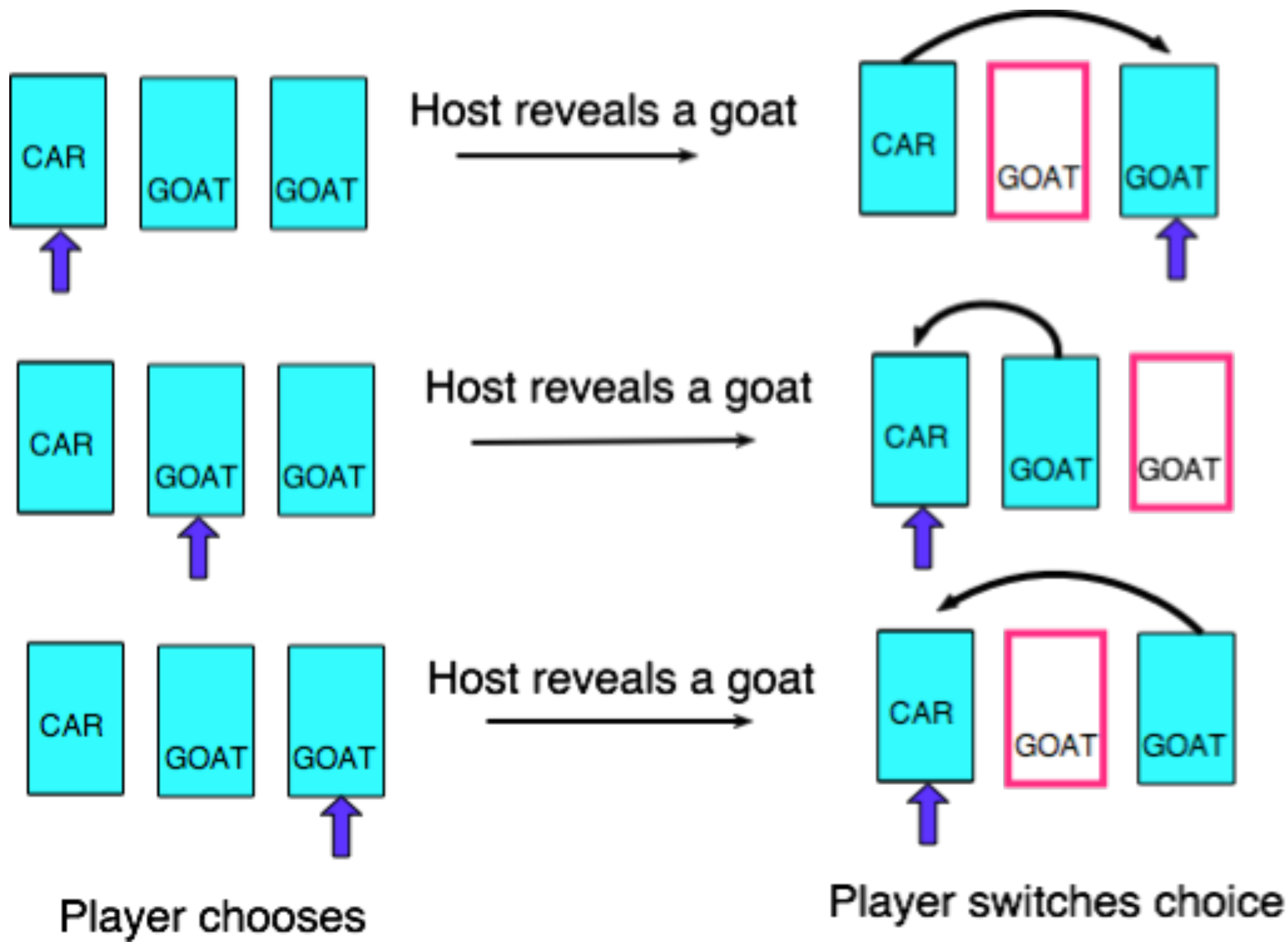
$$p(F) = 3/4, \text{ and } p(E \cap F) = 1/4$$

So the answer is ¼ divided by ¾ = 1/3

Monty Hall 3-door Puzzle

- The Monty Hall problem paradox
  - Consider a game show where a prize (a car) is behind one of three doors
  - The other two doors do not have prizes (goats instead)
  - After picking one of the doors, the host (Monty Hall) opens a different door to show you that the door he opened is not the prize
  - Do you change your decision?
- Your initial probability to win (i.e. pick the right door) is 1/3
- What is your chance of winning if you change your choice after Monty opens a wrong door?
- After Monty opens a wrong door, if you change your choice, your chance of ***winning*** is 2/3
  - Thus, your chance of winning ***doubles*** if you change
  - Huh?

What is behind the doors?



Warm up

Problem 1. A sequence of 10 bits is randomly generated. What is the probability that at least one of these bits is 0?

Answer.

Probability that none of these bits is 0 is $1/2^{10}$

So, the probability that at least one of these bits is 0 is

$(1-1/2^{10}) = 1023/1024$

Warm up

Problem 2. Find the probability of selecting none of the correct six integers in a lottery, (where the order in which these integers are selected does not matter) from the positive integers 1-40?

Answer. The number of ways of selecting *all wrong numbers* is the number of ways of selecting six numbers from the 34 *incorrect numbers*. There are C(34,6) ways to do this. Since there are C(40,6) ways to choose numbers in total, the probability of selecting none of the correct six integers is

$$C(34,6)/C(40,6)$$

Bernoulli trials

An experiment with only two outcomes (like 0, 1 or T, F) is called a Bernoulli trial . Many problems need to compute the probability of exactly k successes when an experiment consists of n independent Bernoulli trials.

Bernoulli trials

Example. A coin is *biased* so that the probability of *heads* is 2/3. What is the probability that exactly four heads come up when the coin is flipped exactly seven times?

Bernoulli trials

The number of ways 4-out-of-7 flips can be heads is C(7,4).

H H H H T T T

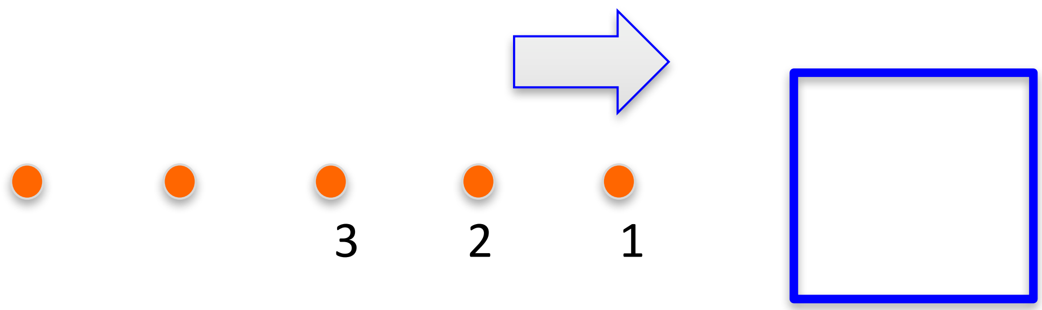
T H H T H H T

T T T H H H H

Each flip is an independent flips. For each such pattern, the probability of 4 heads (and 3 tails) = $(2/3)^4$. $(1/3)^{3.}$ So, in all, the probability of exactly 4 heads is C(7,4). $(2/3)^4$. $(1/3)^3$ = 560/2187

The Birthday Problem

The problem. What is the smallest number of people who should be in a room so that the probability that at least two of them have the same birthday is greater than ½?

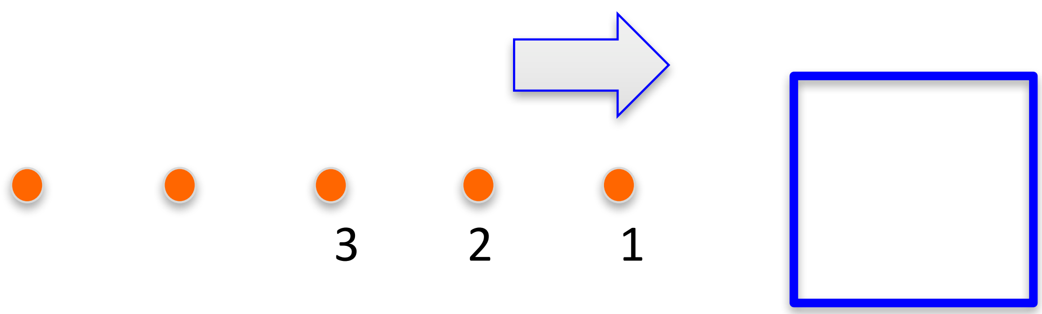


Consider people entering the room one after another. Assuming birthdays are randomly assigned dates, the probability that the second person has the same birthday as the first one is 1 - 365/366

Probability that third person has the same birthday as any one of the previous persons is 1 – 364/366 x 365/366

The Birthday Problem

Continuing like this, probability that the n^{th} person has the same birthday as one of the previous persons is 1 – 365/366 x 364/366 x … x (367 –n)/366



Solve the equation so that for the n^{th} person, this probability exceeds ½. You will get n = 23

Also sometimes known as the **birthday paradox**.

Random variables

DEFINITION. A random variable is a function from the ***sample space*** of an experiment to the set of *real numbers*

Note. A random variable is a function, not a variable ☺

Example. A coin is flipped three times. Let X(t) be the random variable that equals the number of heads that appear when the outcome is t. Then

X(HHH) = 3

X(HHT) = X(HTH) = X(THH) = 2

X(TTH) = X(THT) = X(HTT) = 1

X(TTT) = 0

Expected Value

Informally, the ***expected value*** of a random variable is its average value. Like, “what is the average value of a Die?”

DEFINITION. The ***expected value*** of a random variable X(s) is equal to $\sum_{s \in S} p(s)X(s)$

EXAMPLE 1. *Expected value of a Die*

Each number 1, 2, 3, 4, 5, 6 occurs with a probability 1/6. So the expected value is $1/6\ (1+2+3+4+5+6) = 21/6 = 7/2$

Expected Value (continued)

EXAMPLE 2. *A fair coin is flipped three times. Let X be the random variable that assigns to an outcome the number of heads that is the outcome. What is the expected value of X?*

There are eight possible outcomes when a fair coin is flipped three times. These are HHH, HHT, HTH, HTT, THH, THT, TTH, TTT. Each occurs with a probability of 1/8. So,

$E(X) = 1/8(3+2+2+2+1+1+1+0) = 12/8 = 3/2$

Geometric distribution

Consider this:

You flip a coin and the probability of a tail is p. This coin is repeatedly flipped until it comes up tails.

What is the *expected number of flips* until you see a tail?

Geometric distribution

The sample space {T, HT, HHT, HHHT …} is infinite.

The probability of a tail (T) is p.
Probability of a head (H) is (1-p)
The probability of (HT) is (1-p)p
The probability of (HHT) is $(1-p)^2p$ etc. Why?

Let X be the random variable that counts the number of flips to see a tail. Then $p(X=j) = (1-p)^{j-1} \cdot p$

This is known as geometric distribution.

Expectation in a Geometric distribution

X = the random variable that counts the number of flips to see a tail.

So, $X(T) = 1, X(HT) = 2, X(HHT) = 3$ and so on

$$E(X) = \sum_{1}^{\infty} j.p(X = j)$$

$$= 1.p + 2.(1-p).p + 3.(1-p)^2.p + 4.(1-p)^3.p + \ldots$$

This infinite series can be simplified to 1/p.

Thus, if p = 0.2 then the expected number of flips after which you see a tail is 1/0.2 = 5

Explanation

Probability	**Value**
0.2	30
0.3	40
0.5	20

What is the average value?

0.2 x 30 + 0.3 x 40 + 0.5 x 20 = 28

Linearity of Expectation

Theorem

If X_i, i = 1,2,...,n with n a positive integer, are random variables on S, and if a and b are real numbers, then

(i) $E(X_1 + X_2 + \cdots + X_n) = E(X_1) + E(X_2) + \cdots + E(X_n)$

(ii) $E(aX+b) = aE(X)+b$.

Example 1. What is the expected value of **the sum of the numbers** that appear when a pair of fair dice is rolled?

Let X_1 and X_2 be the random variables so that X_1 appears in the first die and X_2 appears on the second die. $E(X_1 + X_2) = E(X_1) + E(X_2) = 7/2 + 7/2 = 7$.

Useful Formulas

$$p(\bar{E}) = 1 - p(E)$$

$$p(E \cap F) = p(E).p(F)$$ (E and F are mutually independent)

$$p(E \cup F) = p(E) + p(F)$$ (E and F are mutually independent)

$$p(E \cup F) = p(E) + p(F) - p(E \cap F)$$ (E and F are not independent: Inclusion- Exclusion)

$$p(E \mid F) = \frac{p(E \cap F)}{p(F)}$$ (Conditional probability: given F, the probability of E)

Monte Carlo Algorithms

A class of probabilistic algorithms that make a random choice at one or more steps.

Example. Has this batch of n chips *not* been tested by the chip maker?

Randomly pick a chip and test it.
  If it is bad, then the answer is true (i.e. it has not been tested).
  If the chip is good then the answer is "don't know." Then randomly pick another.

After the answer is "don't know" for K different random picks, with you certify the batch to be good.

What is the probability of a wrong conclusion?

Monte Carlo Algorithms

Assume that in previously untested batches, the probability that a particular chip is bad has been observed to be 0.1. So the probability of a chip being good from an untested batch is (1-0.1) = 0.9.

Each test is independent. So the probability that all K steps produce the result "don't know" is 0.9^{k}. By making K large enough, one can make the probability as small as possible. Thus, if K=66, then $0.9^{66} < 0.001$

The fact that so many chips are OK tells that the probability that the batch has not been tested is very small. So we certify the batch. Usually K is a constant. Each test takes a *constant time* – so we can certify (or discard) a batch in constant time.

-- Certification via random witnesses
-- Monte Carlo algorithm for testing prime numbers

Bayes' theorem

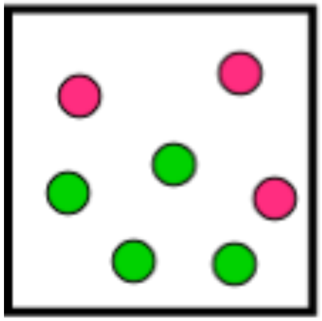
This is related to conditional probability. We can make a realistic estimate when some extra information is available.

Problem 1.
There are two boxes.
Bob first chooses one of the two boxes at random.
He then selects one of the balls in this box at random.

If Bob **has selected a red ball**, what is the probability that ***he selected a ball from the first box?***

(See page 469 of your textbook)

Box 1



Box 2

Bayes' theorem

Let E= Bob chose a red ball. So E' = Bob chose a green ball
F= Bob chose from Box 1. So F' = Bob chose from Box 2
We have to compute p(F|E)


Box 1

p(E|F) = 7/9, p(E|F') = 3/7

We have to find $p(F \mid E) = \frac{p(F \cap E)}{p(E)}$

p(F) = p(F') = 1/2


Box 2

$$p(E \cap F) = p(E \mid F).p(F) = (7/9).(1/2) = 7/18$$

$$p(E \cap F') = p(E \mid F').p(F') = (3/7).(1/2) = 3/14$$

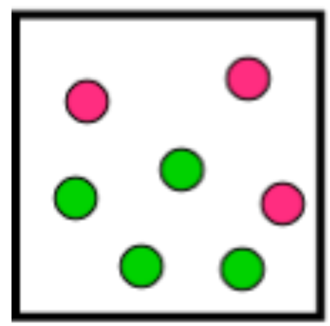
Bayes' theorem

$$p(E) = p(E \cap F) + p(E \cap F^{'}) == 7/18 + 3/14 = 38/63$$

$$p(F \mid E) = \frac{p(F \cap E)}{p(E)} = \frac{7/18}{38/63} = \frac{49}{76}$$

Box 1

This is the probability that
Bob chose the ball from Box 1



Box 2

Bayes' theorem

Let E and F be events from a sample space S such that p(E) ≠ 0 and p(F) ≠ 0. Then

given

$$p(F \mid E) = \frac{p(E \mid F).p(F)}{p(E \mid F)p(F) + p(E \mid \bar{F}).p(\bar{F})}$$

Compute this

Bayes' theorem

Problem 2

1. Suppose that one person in 100,000 has a particular rare disease for which there is a fairly accurate diagnostic test.

2. This test is correct 99.0% of the time when given to a person selected at random who has the disease;

3. The test is correct 99.5% of the time when given to a person selected at random who does not have the disease.

Find the probability that a person who tests positive for the disease really has the disease. (See page 471 of your textbook)

Bayes' theorem

- ✓ 1 in 100,000 has the rare disease (1)
- ✓ This test is 99.0% correct if actually infected; (2)
- ✓ The test is 99.5% correct if not infected (3)

Let F = event that a randomly chosen person has the disease
and E = event that a randomly chosen person tests positive

So, $p(F) = 0.00001$, $p(F') = 0.99999$ {from (1)}
Also, $p(E|F) = 0.99$, and $p(E'|F) = 1 - 0.99 = 0.01$ {from (2)}
Also $p(E'|F') = 0.995$, and $p(E|F') = 1 - 0.995 = 0.005$ {from (3)}

Now, plug into Bayes' theorem.

Bayes' theorem

$$p(F \mid E) = \frac{p(E \mid F).p(F)}{p(E \mid F)p(F) + p(E \mid \bar{F}).p(\bar{F})}$$

$$= \frac{0.99 \times 0.00001}{0.99 \times 0.00001 + 0.005 \times 0.99999} \simeq 0.002$$

So, the probability that a person "who tests positive for the disease" really has the disease is only 0.2%