Stat 201: Statistics I
Chapter 4

Metropolitan State University

February 5, 2018

Chapter 4
Probability

Section 4.1
Basic Concepts of Probability

Terms

A **trial** is conducted to obtain a result from a process with uncertain outcomes. The book also refers to this as a **procedure**.

An **event** is an outcome of interest from a trial.

A **simple event** is an event that cannot be broken down into simpler parts.

A **sample space** is the set of all possible simple events for a trial.

Terms, example

Example

Consider flipping a coin...

- **Trial:** One flip of a coin
- **Event:** Getting heads (H)
- **Simple event:** This event cannot be broken down, it is a simple event
- **Sample space:** Two possible events: { H, T }

Terms, example cont.

Example

Consider flipping a coin three times...

- **Trial:** Three flips of a coin
- **Event:** Getting two heads and a tail
- **Simple event:** There are several ways this event can occur. It can be broken down into { HHT, HTH, THH }. Getting heads on the first two flips and tails on the third (HHT) is a simple event.
- **Sample space:** Eight possible simple events:
  { HHH, HHT, HTH, HTT, THH, THT, TTH, TTT }

Probability

Probability is a measure of the chance an event will occur in a trial.

- Though sometimes expressed as a percentage, probability is always a number between 0 and 1.
- Probability of 1 means the event is certain to occur and probability of 0 means it is impossible.

Can interpret probability in two ways:

- Probability is the proportion an event will occur over a large number of trials. The **law of large numbers** says this proportion will approach the "true" probability as the number of trials increases.
- Some trials can't be repeated (i.e. the weather tomorrow). Then, probability is the level of confidence that an event will occur in a trial (30% chance of rain tomorrow).

Notation

Events are designated with capital letters:

- $A =$ Get two heads and a tail in three coin flips
- $B =$ Rain tomorrow
- $C =$ A randomly selected person is taller than 78 inches

Probabilities are designated with $P()$.

- $P(A)$ is the probability of event A.

Determining probabilities

Classical Approach: If all simple events are equally likely, then

$$P(A) = \frac{\text{number of simple events satisfying } A}{\text{total number of simple events in sample space}}$$

Relative Frequency Approximation: Given a sample of trials,

$$P(A) = \frac{\text{number of times } A \text{ occured}}{\text{number of trials}}$$

Classical approach, example

Example

Consider flipping a coin three times. Assuming a fair coin, all simple events are equally likely. Recall the sample space,

{ HHH, HHT, HTH, HTT, THH, THT, TTH, TTT }

- $A =$ Get two heads and a tail, { HHT, HTH, THH }

$$P(A) = \frac{3}{8}$$

- $B =$ Get *at least* two heads, { HHH, HHT, HTH, THH }

$$P(B) = \frac{4}{8} = \frac{1}{2}$$

Relative frequency, example

Example

The Youth Risk Behavior Survey (YRBS) for 2015 reports that 9421 out of 15624 teenagers surveyed had driven a car at least once in the previous month. Of the 9421 teen drivers, 3806 had texted or emailed while driving.

What is the probability that a randomly selected teen driver had texted or emailed while driving?

- $A =$ Teen driver has texted or emailed while driving

$$P(A) = \frac{3806}{9421} = 0.404$$

Complements

The **comlement** of event A, denoted as $\bar{A}$, consists of all outcomes in the sample space which are not included in A.

Example

- $A =$ Get a head on one coin flip
  $\bar{A} =$ Get a tail on one coin flip
- $B =$ Get exactly two heads on three coin flips
  $\bar{B} =$ Get zero, one or three heads on three coin flips

Complement rule

Since an event and its complement (A and $\bar{A}$) comprise all possible outcomes, then it is *always* the case that

$$P(A) + P(\bar{A}) = 1$$

Example

- One coin flip: $A = \{$ H $\}$, $\bar{A} = \{$ T $\}$

$$P(A) + P(\bar{A}) = \frac{1}{2} + \frac{1}{2} = 1$$

- Two heads in three coin flips: $B = \{$ HHT, HTH, THH $\}$, $\bar{B} = \{$ TTT, HTT, THT, TTH, HHH $\}$

$$P(B) + P(\bar{B}) = \frac{3}{8} + \frac{5}{8} = 1$$

Using the complement rule

The Youth Risk Behavior Survey (YRBS) for 2015 reports that 9421 out of 15624 teenagers surveyed had driven a car at least once in the previous month. Of the 9421 teen drivers, 3806 had texted or emailed while driving.

What is the probability that a randomly selected teen driver had *not* texted or emailed while driving?

- $A =$ Teen driver has texted or emailed while driving
  $\bar{A} =$ Teen driver has not texted or emailed while driving
- $$P(A) = \frac{3806}{9421} = 0.404$$
- $$P(\bar{A}) = 1 - P(A) = 1 - 0.404 = 0.596$$

Unlikely vs. unusual events

An event is **unlikely** if its probability is below some threshold, usually 0.05.

An event is **unusual** if it represents an extreme outcome.

Example

Consider flipping a fair coin 1000 times. Expect heads half of the time, for a total of 500.

- Let A be the event of exactly 523 heads.
  - A is unlikely ($P(A) = 0.00876$), but not unusual.
- Let B be the event of exactly 46 heads.
  - B is very unlikely ($P(B) = 5.929 \times 10^{-222}$) and unusual.

Practice: Cancer screening

Suppose a company is testing a new, cheaper screening test for cancer. They gather a random sample of 1000 people, giving every subject the new test and a doctor visit for definitive diagnosis. These are the results.

	Test Result	
Diagnosis	Positive	Negative
Cancer	74	13
No cancer	26	887

- The number 26 represents **false positives**, positive test results for those with no cancer.
- The number 13 represents **false negatives**, negative test results for those with cancer.

Practice: Cancer screening, cont.

	Test Result	
Diagnosis	Positive	Negative
Cancer	74	13
No cancer	26	887

What is the probability of a randomly selected person having cancer?

- $$P(\text{cancer}) = \frac{74 + 13}{74 + 13 + 26 + 887} = \frac{87}{1000} = 0.087$$

What is the probability of a false negative (person has cancer, but test result is negative)?

- $$P(\text{false negative}) = \frac{13}{1000} = 0.013$$

Section 4.2
Addition Rule and Multiplication Rule

Compound events

A **compound event** is an event which occurs when at least one of two or more simple events occur.

- Denoted as $C = A$ **or** B

Example

- $A =$ Get exactly two heads in three flips
  $A =$ HHT or HTH or THH
- $A =$ Student gets an A on midterm
  $B =$ Student gets a B on midterm
  $C =$ A or B $=$ Student gets an A or a B on midterm
- $A =$ Student gets an A on midterm
  $B =$ Student is female
  $C =$ A or B $=$ Student gets an A on midterm or is female

Disjoint events

Disjoint events are two (or more) events that cannot occur simultaneously. Also called **mutually exclusive**.

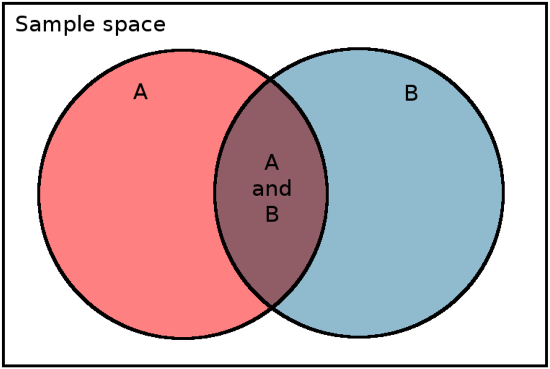
- Complements are always disjoint.

Example

- $A =$ Get exactly two heads in three flips
  $B =$ Get exactly zero, one or three heads in three flips
  A and B are complements, thus they are disjoint.
- $A =$ Student gets an A on midterm
  $B =$ Student gets a B on midterm
  A and B are disjoint, cannot get an A and a B on the midterm
- $A =$ Student gets an A on midterm
  $B =$ Student is female
  A and B are *not* disjoint, possible to get an A and be female

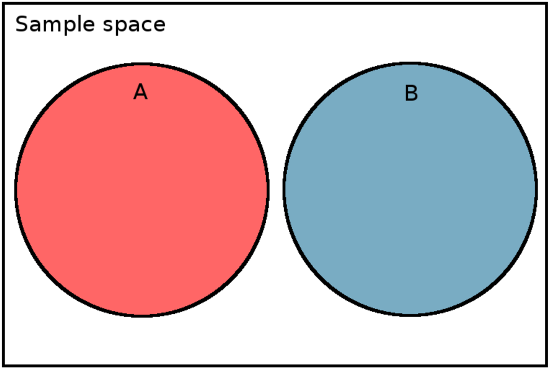
Venn diagrams

Venn diagrams are a good way to visualize events in a sample space.



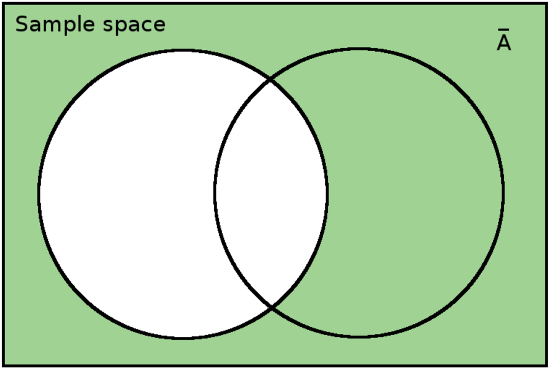
Venn diagrams, disjoint events

Disjoint events are represented non-overlapping circles.



Venn diagrams, complements

Complements are the whole sample space except the event area.



Addition rule

The general rule for calculating probabilities of compound events is to count the outcomes satisfying A and the outcomes satisfying B, making sure to only count each outcome once, and then divide by total number of outcomes.

The formal rule is

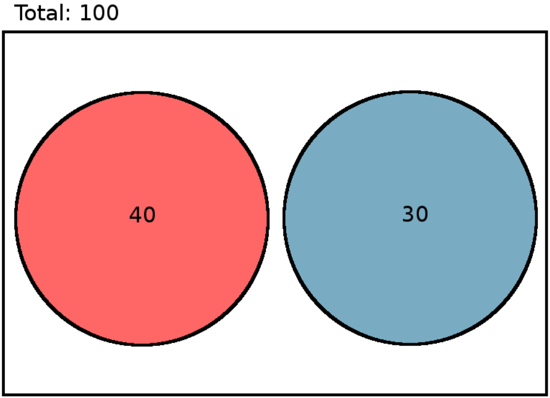
$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

For disjoint events, $P(A \text{ and } B) = 0$, so the rule becomes,

$$P(A \text{ or } B) = P(A) + P(B)$$

Addition rule, example

Suppose a statistics class has 100 students. On the midterm, 40 students got an A (event A) and 30 students got a B (event B).



Addition rule, example cont.

Suppose a statistics class has 100 students. On the midterm, 40 students got an A (event A) and 30 students got a B (event B).

What is the probability that a randomly selected student got an A or a B?

- By the general rule, 40 outcomes for event A and 30 outcomes for event B, and none are counted twice. So,

$$P(A \text{ or } B) = \frac{40+30}{100} = 0.7$$

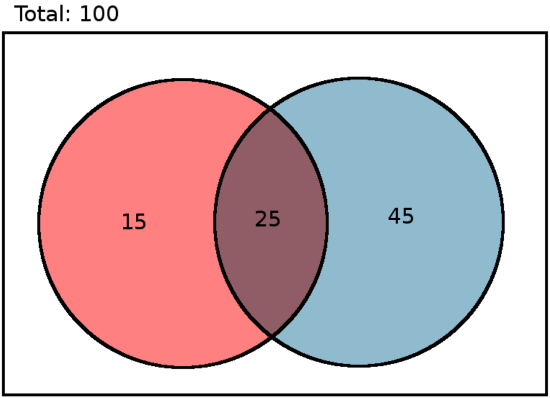
- By the formal rule, $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$.

$$P(A) = 0.4 \qquad P(B) = 0.3 \qquad P(A \text{ and } B) = 0$$

$$P(A \text{ or } B) = 0.4 + 0.3 - 0 = 0.7$$

Addition rule, example 2

Suppose a statistics class has 100 students. On the midterm, 40 students got an A (event A), 70 students are female (event B) and 25 of the female students got an A.



Addition rule, example 2 cont.

Suppose a statistics class has 100 students. On the midterm, 40 students got an A (event A), 70 students are female (event B) and 25 of the female students got an A.

What is the probability that a randomly selected student got an A or is female?

- By the general rule, if outcomes were counted as in the previous example,

$$P(A \text{ or } B) \neq \frac{40+70}{100} = 1.1$$

- The females who got A's were counted twice. Instead, count distinct outcomes in the circles of the Venn diagram.

$$P(A \text{ or } B) = \frac{15+25+45}{100} = 0.85$$

Addition rule, example 2 cont.

Suppose a statistics class has 100 students. On the midterm, 40 students got an A (event A), 70 students are female (event B) and 25 of the female students got an A.

What is the probability that a randomly selected student got an A or is female?

- By the formal rule, $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$.

$$P(A) = 0.4 \qquad P(B) = 0.7 \qquad P(A \text{ and } B) = 0.25$$

$$P(A \text{ or } B) = 0.4 + 0.7 - 0.25 = 0.85$$

Addition rule, example 2 table

Suppose a statistics class has 100 students. On the midterm, 40 students got an A (event A), 70 students are female (event B) and 25 of the female students got an A.

Construct a table for this situation.

	Midterm grade		
Gender	A	Not A	Total
Female	25	45	70
Male	15	15	30
Total	40	60	100

Complements, revisited

Remember, an event and its complement comprise the whole sample space. Whatever the result of a trial, it satisfies either the event or its complement.

$$P(A \text{ or } \bar{A}) = 1$$

Also, complements are disjoint, so the addition rule is

$$P(A \text{ or } \bar{A}) = P(A) + P(\bar{A})$$

Then,

$$P(A) + P(\bar{A}) = 1$$

$$P(A) = 1 - P(\bar{A})$$

$$P(\bar{A}) = 1 - P(A)$$

Practice: Cancer screening

	Test Result		
Diagnosis	Positive	Negative	Total
Cancer	74	13	87
No cancer	26	887	913
Total	100	900	1000

What is the probability of a randomly selected person having cancer (A) or not having cancer (B)?

- Having cancer and not having cancer are complements,

$$P(A \text{ or } B) = 1$$

Practice: Cancer screening, cont.

	Test Result		
Diagnosis	Positive	Negative	Total
Cancer	74 (0.074)	13 (0.013)	87 (0.087)
No cancer	26 (0.026)	887 (0.887)	913 (0.913)
Total	100 (0.1)	900 (0.9)	1000 (1)

What is the probability of a randomly selected person having cancer (A) or getting a positive test result (B)?

- Using the addition rule,

$$P(A \text{ or } B) = 0.087 + 0.1 - 0.074 = 0.113$$

Practice: Cancer screening, cont.

	Test Result		
Diagnosis	Positive	Negative	Total
Cancer	74 (0.074)	13 (0.013)	87 (0.087)
No cancer	26 (0.026)	887 (0.887)	913 (0.913)
Total	100 (0.1)	900 (0.9)	1000 (1)

What is the probability of a test being wrong? That is, what is the probability of getting a false positive (A) or a false negative (B)?

- The events are disjoint. Using simplified addition rule,

$$P(A \text{ or } B) = 0.026 + 0.013 = 0.039$$

Independent events

Two events are said to be **independent** if the probability of one is unaffected by the occurrence of the other.

Example

- Let $A =$ get a head on the first flip
  and $B =$ get a tail on the second flip.
  $P(B) = \frac{1}{2}$ regardless of what happens on the first flip.

Dependent events

If two events are not independent, then they are **dependent**. That is, the probability of one changes depending on the outcome of the other.

Example

Consider an urn with 2 red balls and 3 blue balls. A trial consists of randomly selecting a ball from the urn and not replacing it.

- Let $A =$ get a red ball the first trial and $B =$ get a blue ball the second trial.
    - If A occurs, then $P(B) = 3/4$
    - If A does not occur, then $P(B) = 2/4$

Example

The probability of a randomly selected student getting an A on the final is probably different depending on whether they got an A on the midterm.

Dependent events as independent

When dealing with large populations and small sample sizes, events that are technically dependent can be treated as independent. The rule of thumb the book uses is sample sizes less than 5% of population can be treated as independent.

Example

Suppose the urn has 2000 red balls and 3000 blue balls. The probability of selecting a blue ball is approximately $3/5$, regardless of whether a red ball was previously selected or not.

Conditional probability

The **conditional probability** of an event is the probability assuming another event occurred.

- It is denoted $P(B|A)$ and read as "probability of B given A"
- For independent events, $P(B|A) = P(B)$

Example

There is an urn with 2 red balls and three blue balls. A trial consists of randomly selecting a ball from the urn and not replacing it.
Let $A =$ get a red ball the first trial
and $B =$ get a blue ball the second trial.

- $P(B|A) = 3/4$
- $P(B|\bar{A}) = 2/4$

Multiplication rule

To find the probability of all events in a series of trials, multiply the probability of the first by the probability of the second given the first occurred, etc.

Formally,

$$P(A \text{ and } B) = P(A) \times P(B|A)$$

For independent events,

$$P(A \text{ and } B) = P(A) \times P(B)$$

Multiplication rule, example

Example

There is an urn with 2 red balls and 3 blue balls. A trial consists of randomly selecting a ball from the urn and not replacing it.

What is the probability of selecting a red ball and then selecting a blue ball?

- Let $A =$ get a red ball the first trial
  and $B =$ get a blue ball the second trial.
- B is dependent on A.

$$P(A \text{ and } B) = P(A) \times P(B|A)$$

- $P(A) = 2/5 \qquad P(B|A) = 3/4$
- $$P(A \text{ and } B) = \frac{2}{5} \times \frac{3}{4} = \frac{6}{20} = \frac{3}{10} = 0.3$$

Multiplication rule, example

Example

Consider flipping a coin three times.

What is the probability of get heads on the first two flips and a tail on the third (HHT)?

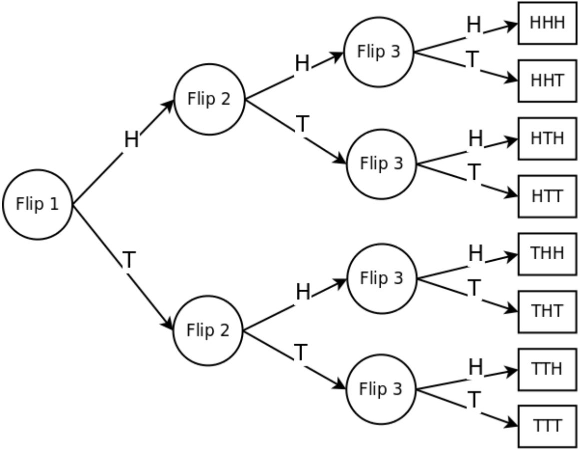
- Let $A =$ get a head the first flip,
  $B =$ get a head on the second flip,
  and $C =$ get a tail on the third flip.
- A, B and C are independent events.

$$P(A \text{ and } B \text{ and } C) = P(A) \times P(B) \times P(C)$$

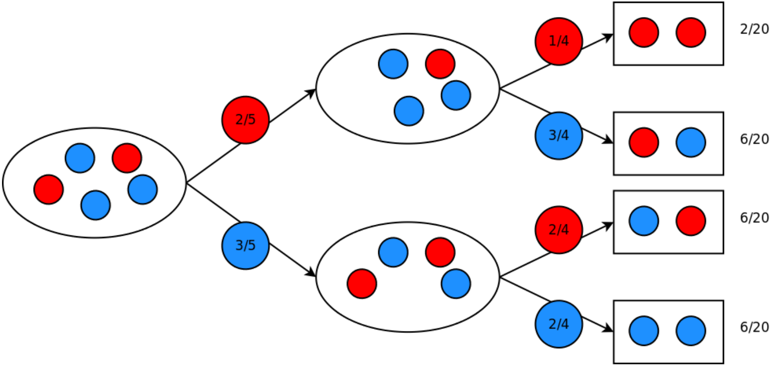
- $P(A) = P(B) = P(C) = 1/2$
- $$P(A \text{ and } B \text{ and } C) = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$$

Tree diagrams

Tree diagrams are a good way to visualize events in a series of trials.



Tree diagram, urn example



Practice: Cancer screening

	Test Result		
Diagnosis	Positive	Negative	Total
Cancer	74	13	87
No cancer	26	887	913
Total	100	900	1000

What is the probability of two randomly selected people both having positive test results?

- Sample size of 2 is less than 5% of population of 1000, so can treat events as independent.
- $P(\text{positive}) = \frac{100}{1000} = 0.1$
- $P(\text{both positive}) = 0.1 \times 0.1 = 0.01$

Practice: Statistics club

The Metro State Statistics Club has 10 members, 6 men and 4 women. They need to select a president, a vice-president and a treasurer. They decide to choose members randomly for the officers positions, in order.

What is the probability that women are selected for president and vice-president, and a man for treasurer?

- $A =$ Woman selected president
  $B =$ Woman selected vice-president
  $C =$ Man selected treasurer
- $P(A \text{ and } B \text{ and } C) = P(A) \times P(B|A) \times P(C|A \text{ and } B)$
- $P(A) = \frac{4}{10} = 0.4$
  $P(B|A) = \frac{3}{9} = 0.33$
  $P(C|A \text{ and } B) = \frac{6}{8} = .75$
- $P(A \text{ and } B \text{ and } C) = 0.4 \times 0.33 \times 0.75 = 0.099$

Recap of probability rules

- To calculate the probability of at least one of two events occurring (A **or** B), use the addition rule. Be aware of whether the events are disjoint or not.
- To calculate the probability of all of a sequence of two, or more, events occurring (A **and** B), use the multiplication rule. Be aware of whether the events are independent or dependent.

Testing for independence

It is sometimes difficult to tell if events are independent. The rule for independent events, that $P(B|A) = P(B)$, can be used to test for independence.

Testing for independence, example

Example

Consider rolling two fair six-sided dice. Let $A =$ total of the dice is 5 and $B =$ at least one of the dice is a 3.

Are A and B independent events?

- $P(B) = P(\text{first die is 3 or second die is 3})$
  $P(B) = P(\text{first die is 3}) + P(\text{second die is 3}) - P(\text{both are 3})$
  $P(B) = \frac{1}{6} + \frac{1}{6} - \frac{1}{36} = \frac{11}{36}$
- If A occurred, then the possible dice values are
  $$\{(1,4),(2,3),(3,2),(4,1)\}$$
- $P(B|A) = \frac{2}{4} = \frac{1}{2}$
- A and B are not independent.

Section 4.3
Complements, Conditional Probability and Bayes' Theorem

Complex events

A **complex event** is a compound event in the form of "at least one" of some simpler event occurs.

- This is a shortcut for "exactly one **or** exactly two **or**..."

The complement of a complex event is none of the simpler events occur, or that the simpler event didn't occur all the time.

Example

Let $A =$ Get at least one head in three flips.

- $A =$ exactly one head **or** exactly two heads **or** exactly three heads
- $\bar{A} =$ Get no heads on three flips $=$ Get tails every flip

Using complements for complex events

As sample sizes increases, calculating probabilities for complex events becomes very difficult. It is often much easier to use complements for such calculations.

- Recall the complement rule,

$$P(A) = 1 - P(\bar{A})$$

Using complements for complex events, example

Example

A manufacturer of insulin test strips has a defect rate of 0.7% ($P(\text{strip is defective}) = 0.007$). In a box a 50 strips, what is the probability of getting at least one defective strip?

- $A =$ At least one strip is defective
  $\bar{A} =$ No strips are defective $=$ All strips are good
- $P(\text{strip is good}) = 1 - P(\text{strip is defective}) = 1 - 0.007 = 0.993$
- $$P(\bar{A}) = 0.993 \times 0.993 \times \cdots \times 0.993$$
  $$P(\bar{A}) = (0.993)^{50} = 0.704$$
- $$P(A) = 1 - P(\bar{A}) = 1 - 0.704 = 0.296$$

Other kinds of complex events

A complex event could be in the form of “at least twice” or ”at least three times”, etc.

The complements in these cases, “the event occurs zero or one times”, is still probably simpler to calculate.

Practice: cancer screening

	Positive	Negative	Total
Cancer	74 (0.074)	13 (0.013)	87 (0.087)
No cancer	26 (0.026)	887 (0.887)	913 (0.913)
Total	100 (0.1)	900 (0.9)	1000 (1)

If ten subjects are selected randomly, what is the probability at least one of them test positive?

- A = At least one test positive
  $\bar{A}$ = Zero test positive = All test negative
- $$P(\bar{A}) = (0.9)^{10} = 0.349$$
- $$P(A) = 1 - 0.349 = 0.651$$

Formal definition of conditional probability

Recall the multiplication rule,

$$P(A \text{ and } B) = P(A) \times P(B|A)$$

From this, the formal definition of conditional probability is

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

Intuitive approach to conditional probability

An intuitive approach to $P(B|A)$ is to assume A has occurred, then count count instances of B. A is, in a sense, the new sample space.

$$P(B|A) = \frac{\text{number of } B \text{ and } A}{\text{number of } A}$$

Practice: Cancer screening

	Positive	Negative	Total
Cancer	74 (0.074)	13 (0.013)	87 (0.087)
No cancer	26 (0.026)	887 (0.887)	913 (0.913)

What is the probability of a positive test result if the subject has cancer?

- A = Subject has cancer
  B = Positive test result
- Formally,

$$P(B|A) = \frac{P(A \text{ and } B)}{P(A)} = \frac{0.074}{0.087} = 0.851$$

- Intuitive approach,

$$P(B|A) = \frac{\text{number of } B \text{ and } A}{\text{number of } A} = \frac{74}{87} = 0.851$$

Practice: Cancer screening, cont.

	Positive	Negative	Total
Cancer	74 (0.074)	13 (0.013)	87 (0.087)
No cancer	26 (0.026)	887 (0.887)	913 (0.913)

What is the probability of a negative test result if the subject does not have cancer?

- A = Subject does not have cancer
  B = Negative test result
- Formally,

$$P(B|A) = \frac{P(A \text{ and } B)}{P(A)} = \frac{0.887}{0.913} = 0.972$$

- Intuitive approach,

$$P(B|A) = \frac{\text{number of } B \text{ and } A}{\text{number of } A} = \frac{887}{913} = 0.972$$

Sensitivity and specificity

The proceeding examples have specific terms when used with diagnostic tests.

- **Sensitivity** is the probability of a positive test result for a subject which has the conditions, $P(\text{positive}|\text{cancer})$.
- **Specificty** is the probability of a negative test result for a subject which does not have the conditions, $P(\text{negative}|\text{no cancer})$.

Many diagnostic tests work by measuring the level of a certain chemical and returning a positive result if it is above a designated threshold. Adjusting this threshold to increase sensitivity will decrease specificity, and vice versa. There is always a trade-off.

Screening tests for rare events

Example

Suppose there is a screening test for a rare disease which has a prevalence of 0.3%. The screening test has 99% sensitivity and 99% specificity. 100,000 people are screened.

	Positive	Negative	Total
Disease	297	3	300
No disease	997	98703	99700
Total	1294	98706	100,000

What is the probability that someone who tested positive does not have the disease?

- $$P(\text{no disease}|\text{positive}) = \frac{997}{1294} = 0.770$$
- The complement, $P(\text{disease}|\text{positive}) = 0.23$, is known as the **precision** or the **positive predictive value (PPV)** of the test.

Screening tests for rare events, cont.

- This does not mean screening tests are not useful. Often they are a first step before tests that are more accurate, but also more expensive and/or more invasive.
  - Cancer screening, followed by biopsy for confirmation
- Sometimes tests like these can have profound consequences for peoples lives.
  - Drug screening for jobs
  - Vetting for refugees or immigrants
  - etc.
- It is important to remember that no test is perfect and there are often trade-offs (sensitivity / specificity).