Statistical Methods
APPM 4570/5570, STAT 4000/5000

Week 2:

Probability: Counting, Sets, Bayes

Random variable

Random variable is a measurable quantity whose outcome is unknown (random) upfront, before an experiment or study is carried out.

Examples:

- Outcome of a coin toss
- A random card selected from a deck
- Commuting time on a particular morning

Experiment -- associated with a carefully controlled laboratory conditions

Study – more general, can be observational (eg. doing surveys)

Sidenote: this distinction is at the core of **causal** vs **non-causal relationships**. Observational studies (and even some experiments) can only give non-causal interpretations.

Sample Space

The set of all possible outcomes of an experiment or study.

Thus, if there are *n* units under study, we have *n* random variables, and *n* possible sets of outcomes.

Example – continuous outcomes:
Study: Measuring the birth weight of babies born in 2000.

Sample space for each baby birth weight: positive real numbers

All positive reals? Well, no. But… we don't know what the upper bound is

In that case, we can say "in theory yes, all real numbers" with the understanding that big numbers are really really unlikely (eg, really chubby babies – say those over 15 pounds - are improbable (1 in a billion or fewer)

Examples of sample spaces

The simplest study is the one where we only have two possible outcomes and a single random variable:

- tossing a single coin once
- examining a single fuse to see whether it is defective

The sample space set for the first example is

$S = \{H(ead), T(ails)\} = \{H, T\}$

The sample space set for the second example can be abbreviated as

$S = \{N, D\}$, where N stands for "not defective", D stands for "defective".

Events

So, the sample space is the **set** of all possible outcomes

An **event** is any collection (**subset**) of outcomes from the sample space

An event is **simple** if it consists of exactly one outcome

An event is **compound** if it consists of more than one outcome

We say that an event (set) "*A*" occurred if the experimental outcome is contained in the set *A.*

Events - example

Study: birth weight and gestational age for a random baby born in 2000

– note this is a 2-dimensional RV (time, weight)

Sample space: **R^+ x R^+ (the upper right quadrant of R^2)**

Event A = small for gestational age (SGA)

This event occurs if our experiment returns a baby whose weight is in the bottom 10% of all baby birth weights for that particular gestational age.

This event is the subset bounded by the line time=0, the line weight=0, and the boundary defined by SGA(t)

Events – example

Consider an exam where there are 3 true/false questions.

So, for each question, you can select T (*true*) or F (*false*)

There are eight possible outcomes that make up the sample space of all exam answers {*TTT*, *TTF*, *TFT*, *FTT*, *TFF*, *FTF*, *FFT*, *FFF*}.

We can make 2 choices at each of the 3 questions => total number of all possible answers is

$2 * 2 * 2 = 2^3$

There are eight simple events, among which are

E_1 = {*TTT*} and E_8 = {*FFF*}.

Set Theory

An event is a set, so set theory can be used to study events and do their probability calculus.

Definition

The complement of an event A, denoted by A'

(pronounced A-prime, or complement of A, or not-A)

is the set of all outcomes in the sample space S not contained in A

Some Relations from Set Theory

Definitions

The union of two events *A* and *B*,

denoted $A \cup B$ and read "*A or B*"

is the event consisting of all outcomes that are **either in A or in B or in both**

Both A and B can occur, or just one of them can occur

The intersection of two events *A* and *B*

denoted $A \cap B$ and read "*A and B*"

is the event consisting of all outcomes that are in **both A and B.**

Copyright Prof. Vanja Dukic, Applied Mathematics, CU-Boulder STAT4000/5000

Set Theory

Sometimes *A* and *B* have no outcomes in common, so that the intersection of *A* and *B* contains no outcomes.

Definition

Let $\varnothing$ denote the *null event* (the event consisting of no outcomes whatsoever).

When $A \cap B = \varnothing$, *A* and *B* are said to be **mutually exclusive** or **disjoint** events.

When one occurs, the other cannot occur. One precludes the other.

Copyright Prof. Vanja Dukic, Applied Mathematics, CU-Boulder STAT4000/5000

Set Theory

The operations of union and intersection can be extended to more than two events.

For any three events *A*, *B*, and *C*, the event $A \cup B \cup C$ is the set of outcomes contained in at least one of the three events, whereas $A \cap B \cap C$ is the set of outcomes contained in all three events.

Given events $A_1, A_2, A_3, \ldots$, these events are said to be mutually exclusive (or pairwise disjoint) if no two events have any outcomes in common.

Copyright Prof. Vanja Dukic, Applied Mathematics, CU-Boulder STAT4000/5000

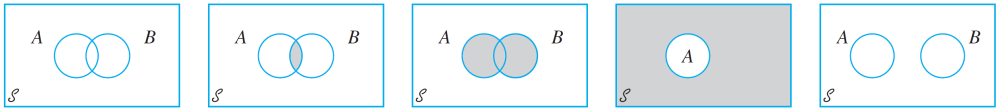
Venn diagrams

We can also use Venn diagrams for sets and events.

To construct a Venn diagram, draw a rectangle whose interior will represent the sample space S

Then any event *A* is represented as the interior of a closed curve (simplest: circle) contained in S

Can shade the desired subset which event is of interest



(a) Venn diagram of events *A* and *B* (b) Shaded region is $A \cap B$ (c) Shaded region is $A \cup B$ (d) Shaded region is A' (e) Mutually exclusive events

Copyright Prof. Vanja Dukic, Applied Mathematics, CU-Boulder STAT4000/5000

Axioms of Probability

For each event A, the number P(A) is called the probability of the event A.

P(A) quantifies how likely is that A will occur.

For something to be a proper probability, we have to have:

Axiom 1: For any event A, $P(A) \geq 0$.

Axiom 2: $P(S) = 1$.

Axiom 3:

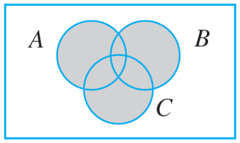
If $A_1, A_2, A_3, \ldots$ is a (possibly infinite) collection of disjoint events:

$$P(A_1 \cup A_2 \cup A_3 \cup \ldots) = \sum_{i=1}^{\infty} P(A_i)$$

More Probability Properties

Propositions

1. For any event A, $P(A) + P(A') = 1$, from which $P(A') = 1 - P(A)$.
2. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
3. $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$



What is Probability?

Probability can be understood as a limit of relative frequency.

Consider an experiment that can be repeatedly performed, in an identical and independent fashion, indefinitely.

An example: repeated coin tossing

The longer we perform the experiment, the more stable the relative frequency of "tails" will get

It will converge on the probability P(tails)

Interpreting Probability

Let's say we perform an experiment n times

The event A will occur (the outcome will be in the set A) in some of the replications. In other replications, A will not occur.

Let $n(A)$ denote the number of replications on which A does occur.

Then the ratio $n(A)/n$ is called the *relative frequency* of occurrence of the event A in the sequence of n replications of the experiment.

Example (Devore)

Let *A* be the event that

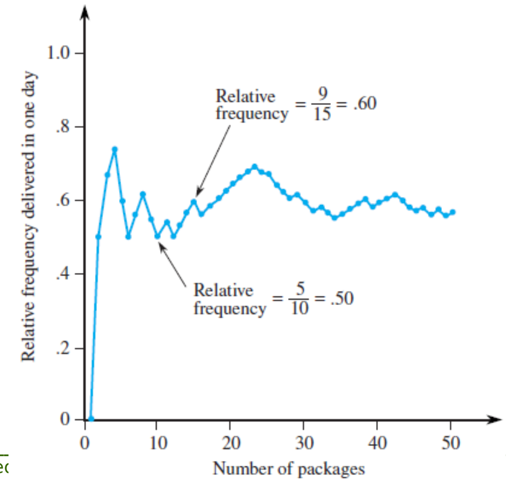
"a package sent within the state of California for 2nd day delivery actually arrives within one day"

The results from sending 10 such packages (the first 10 replications of the experiment) are as follows:

Package #	1	2	3	4	5	6	7	8	9	10
Did *A* occur?	N	Y	Y	Y	N	N	Y	Y	N	N
Relative frequency of *A*	0	.5	.667	.75	.6	.5	.571	.625	.556	.5

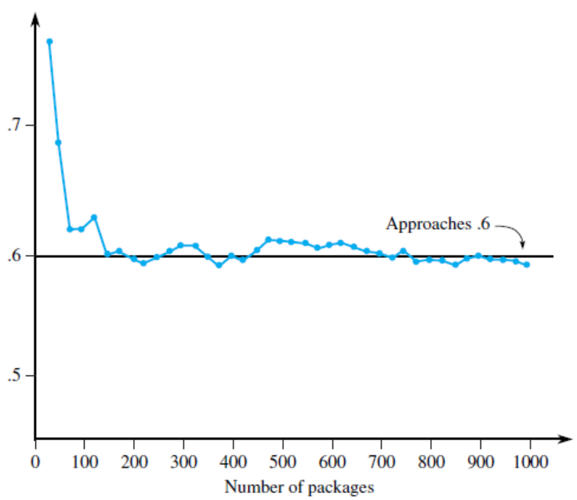
Example (Devore)

The relative frequency $n(A)/n$ fluctuates rather substantially over the course of the first 50 replications.



Example (Devore)

But as the number of replications continues to increase, the relative frequency stabilizes at a limiting value, P(A).



Behavior of relative frequency – Long-run stabilization

Interpreting Probability (Devore)

In that case, a statement such as "the probability of a package being delivered within one day of mailing is .6" means that out of a large number of mailed packages, roughly 60% will arrive within one day (and the other 40% will arrive in 2 or more days).

For coin tossing, we say "fair coin" if

$$P(H) = P(T) = .5$$

and if we were to toss a fair coin 1000 times, the relative frequency of heads will be very close to 50%

Assigning Probabilities to events

There are several ways we can learn about probabilities of events:

1) Empirically – ie, assess probabilities based on experience
2) Analytically, using counting and combinatorial tricks
   - Counting
   - Permutations
   - Combinations

3) Via computer simulations – simulating repeated experiments, and finding out relative frequencies over a large number of simulations

Analytical computation example

A train has five cars. Suppose a commuter is twice as likely to select the middle car (#3) then to select either adjacent car (#2 or #4), and is twice as likely to select any of the adjacent cars then to select any of the end cars (#1 or #5).

Let $p_i = P(\text{car } i \text{ is selected}) = P(E_i)$, for $i = 1 \ldots 5$. All E_i are mutually exclusive. Thus we have

$$p_3 = 2p_2 = 2p_4 \text{ and } p_2 = 2p_1 = 2p_5 = p_4.$$

This gives

$$1 = \Sigma P(E_i) = p_1 + 2p_1 + 4p_1 + 2p_1 + p_1 = 10p_1$$

implying $p_1 = p_5 = .1$, $p_2 = p_4 = .2$, $p_3 = .4$. The probability that one of the three middle cars is selected (a compound event) is then $p_2 + p_3 + p_4 = .8$.

Counting: enumeration for equally likely outcomes

Think of drawing a card at random from a deck of cards.

If there are N equally likely outcomes, the probability for each is $1/N$.

So if we have a compound event, like A = {king}, we have to count the number $N(A)$ of outcomes contained in an event A, and divide by the number of outcomes in a sample space:

$$P(A) = \frac{N(A)}{N}$$

In the case of the above event A = {king}, P(A) = 4/52

Counting: k-tuples

A family requires the services of an obstetrician and a pediatrician. There are two medical clinics, each having 2 obstetricians and 3 pediatricians.

The family wishes to select both doctors from the same clinic. In how many ways can this be done?

Denote the obstetricians by O_1, O_2, O_3, and O_4 and the pediatricians by $P_1, \ldots, P_6$.

Then we wish the number of pairs (O_i, P_j) for which O_i and P_j are associated with the same clinic.

There are 4 obstetricians, and for each there are 3 choices of pediatricians within the same clinic, so there are 12 possibilities.

Counting: k-tuples

If a six-sided die is tossed five times, then the outcome is an ordered collection of five numbers -- a “5-tuple”.

We will call an ordered collection of k objects a ***k-tuple***

There are

$$n_1 n_2 \cdots n_k$$

possible k-tuples.

When k=2, like on the previous slide, it is simply called a “pair”.

Copyright Prof. Vanja Dukic, Applied Mathematics, CU-Boulder

Permutations and Combinations

A **subset** where order matters is called a **permutation.**

The number of permutations of size k that can be formed from n objects will be denoted by $P_{k,n}$.

An unordered subset is called a **combination.**

One way to denote the number of combinations is $C_{k,n}$ or:

$$\binom{n}{k}$$

(pronounced: “n choose k”).

Copyright Prof. Vanja Dukic, Applied Mathematics, CU-Boulder

Permutations (k-tuples from the same set, without replacement)

Example: A college of engineering has 7 departments. Each has one representative on the student council.

From these 7 representatives, one is to be chosen chair, another vice-chair, and one secretary.

How many ways are there to select these three officers?

That is, how many permutations (order matters!) of size 3 can be formed from the 7 representatives?

Copyright Prof. Vanja Dukic, Applied Mathematics, CU-Boulder

Permutations

The chair can be selected from any department, i.e., in
$n_1 = 7$ ways.

Once we select the chair, there are $n_2 = 6$ departments left – so 6 ways to select the vice-chair, and hence $7 \times 6 = 42$ (chair, vice-chair) pairs.

Then, after having selected the chair and vice chair, there remain 5 ways to select the secretary.

This gives

$$P_{3,7} = (7)(6)(5) = 210$$

This is the number of permutations of size 3 that can be formed from 7 distinct individuals.

Copyright Prof. Vanja Dukic, Applied Mathematics, CU-Boulder

Permutations – formula

Recall, for any positive integer n,

$$n! = n(n-1)(n-2)\cdot\cdot\cdot\cdot(2)(1)$$

(0! = 1 by definition).

Then, it follows: $P_{3,7} = (7)(6)(\frac{(7)(6)(5)(4!)}{(4!)} = \frac{7!}{4!}$

More generally,

$$P_{k,n} = n(n-1)(n-2)\cdot\cdot\cdot\cdot(n-(k-2))(n-(k-1))$$

or

$$P_{k,n} = \frac{n!}{(n-k)!}$$

Combinations

Combinations are subsets where order does not matter

Again refer to the student council scenario, and suppose that 3 of the 7 representatives are to be selected to attend a convention.

The order of selection is not important; all that matters is which three get selected.

We can simply take the result we got for $P_{3,7}$ and divide it by the number of ways you can arrange 3 elements

…. 3! is the number of ways we can order a set of 3

$$P_{3,7} = (3!)\cdot\binom{7}{3} \Rightarrow \binom{7}{3} = \frac{P_{3,7}}{3!} = \frac{7!}{(3!)(4!)} = \frac{(7)(6)(5)}{(3)(2)(1)} = 35$$

Combinations

Generalizing ….

$$\binom{n}{k} = \frac{P_{k,n}}{k!} = \frac{n!}{k!(n-k)!}$$

Notice that $\binom{n}{n} = 1$ and $\binom{n}{0} = 1$ since there is only one way to choose a set of (all) n elements or a set of no elements.

Also, since there are n subsets of size 1, we have:

$$\binom{n}{1} = n$$

Examples

A particular iPod playlist contains 100 songs, 10 of which are by the Beatles.

Suppose the shuffle feature is used to play the songs in random order, without repetition.

What is the probability that the 1st Beatles song heard is the 5th song played?

In order for this event to occur, it must be the case that the first 4 songs played are not Beatles' songs (NBs) and that the 5th song is by the Beatles (B).

Using permutations and combinations in computing probabilities cont'd

The number of ways to select the first five songs is 100(99)(98)(97)(96).

The number of ways to select these five songs so that the first four are NBs and the next is a B is 90(89)(88)(87)(10).

Therefore the desired probability is the ratio of the number of outcomes for which the event of interest occurs to the number of possible outcomes:

$$P(1^{\text{st}}\text{ B is the }5^{\text{th}}\text{ song played}) = \frac{90\cdot 89\cdot 88\cdot 87\cdot 10}{100\cdot 99\cdot 98\cdot 97\cdot 96} = \frac{P_{4,\,90}\cdot(10)}{P_{5,\,100}} = .0679$$

Conditional Probability

Can the information "event *B* has occurred" affect the probability of event *A?*

For example,
A = having a particular disease in the presence of certain symptoms.

B = blood test result is negative

The updated (post-test) probability of disease will be different than pre-test (if the test is at all valuable)

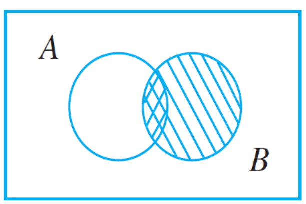
We will use the notation $P(A\,|\,B)$ to represent the **conditional probability of *A* given that *B* has occurred.** *B* is called the "conditioning event."

Conditional Probability

The conditional probability is expressed as a ratio of unconditional probabilities --

probability of the intersection of the two events
----
probability of the conditioning event *B*

Given that *B* has occurred, **the relevant sample space is no longer *S* but it boils down to only the outcomes in *B*.**



Example – conditional probability

Parts are assembled in two different assembly lines, *A* and *A'*
Line *A* uses older equipment than *A'*, so it is somewhat less reliable.

Suppose on a given day line *A* has assembled 8 parts, whereas *A'* has produced 10.

From the 8 parts from *A*,
2 were defective (B) and
6 as nondefective (B').

From the 10 parts from *A'*,
1 was defective (B) and
9 nondefective (B').

		Condition	
		B	***B'***
Line	***A***	2	6
	A'	1	9

Example

cont'd

The sales manager randomly selects 1 of these 18 parts for a test.

Then before the test:

P(line A part selected) = $P(A) = \frac{N(A)}{N}$ = 8/18 = .44

If the test revealed that the part is defective – event B={*defective part*} has occurred. Then the selected part must have been one of the 3 total defective parts made that day.

What is the chance that it was made by the line A?

$$P(A \mid B) = \frac{2}{3} = \frac{2/18}{3/18} = \frac{P(A \cap B)}{P(B)}$$

Copyright Prof. Vanja Dukic, Applied Mathematics, CU-Boulder

The Multiplication Rule for $P(A \cap B)$

The definition of conditional probability yields the following result:

The Multiplication Rule

$$P(A \cap B) = P(A \mid B) \cdot P(B)$$

Copyright Prof. Vanja Dukic, Applied Mathematics, CU-Boulder

Independence

The definition of conditional probability enables us to revise the probability $P(A)$ originally assigned to A when we are subsequently informed that another event B has occurred; the new probability of A is $P(A \mid B)$.

In our examples, it was frequently the case that $P(A \mid B)$ differed from the unconditional probability $P(A)$, indicating that the information "B has occurred" resulted in a change in the chance of A occurring.

Often the chance that A will occur or has occurred is not affected by knowledge that B has occurred, so that $P(A \mid B) = P(A)$.

Copyright Prof. Vanja Dukic, Applied Mathematics, CU-Boulder

Independence

It is then natural to regard A and B as independent events, meaning that the occurrence or nonoccurrence of one event has no bearing on the chance that the other will occur.

Definition

Two events A and B are **independent** if $P(A \mid B) = P(A)$ and are **dependent** otherwise.

The definition of independence might seem "unsymmetric" because we do not also demand that $P(B \mid A) = P(B)$.

Example: A = winning a lottery, and B = raining today.

Copyright Prof. Vanja Dukic, Applied Mathematics, CU-Boulder

Independence

However, using the definition of conditional probability and the multiplication rule,

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A \mid B)P(B)}{P(A)}$$

The right-hand side is $P(B)$ if and only if $P(A \mid B) = P(A)$ (independence), so the equality in the definition implies the other equality (and vice versa).

It is also straightforward to show that if A and B are independent, then so are the following pairs of events:
(1) A' and B,

(2) A and B'

(3) A' and B'

The Multiplication Rule for $P(A \cap B)$

A and B are independent if and only if (iff)

$$P(A \cap B) = P(A)\ P(B)$$

The verification of this multiplication rule is as follows:

$$P(A \cap B) = P(A \mid B)\ P(B) = P(A)\ P(B)$$

where the second equality is valid iff A and B are independent.

The Multiplication Rule for $P(A \cap B)$

In summary: independence means that:

$P(A|B) = P(A)$

$P(B|A) = P(B)$

$P(A \cap B) = P(A)P(B)$

$P(A \cup B) = P(A) + P(B) - P(A \cap B) = P(A) + P(B) - P(A)P(B)$

Independence of More Than Two Events

Definition
Events $A_1, \ldots, A_n$ are **mutually independent** if for every k ($k = 2, 3, \ldots, n$) and every subset of indices $i_1, i_2, \ldots, i_k$,

$$P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdot \cdots \cdot P(A_{i_k})$$

Bayes' Theorem

The computation of a posterior probability $P(B|A_j)$ from given prior probabilities $P(A_i)$ and conditional probabilities $P(A_j|B)$

The general rule is just a simple application of the multiplication rule, goes back to Reverend Thomas Bayes, who lived in the eighteenth century.

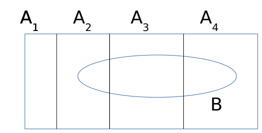
Recall that events $A_1, \ldots, A_k$ are mutually exclusive if no two have any common outcomes. The events are *exhaustive* if one A_i must occur, so that $A_1 \cup \ldots \cup A_k = S$

Copyright Prof. Vanja Dukic, Applied Mathematics, CU-Boulder

The Law of Total Probability

Let $A_1, \ldots, A_k$ be mutually exclusive and exhaustive events. Then for any other event B we have that B can be broken up into pieces belonging to each of the A's:

$B = (B \cap A_1) \cup (B \cap A_2) \ldots \cup (B \cap A_k)$



Thus:

$$P(B) = P(B|A_1)P(A_1) + \cdots + P(B|A_k)P(A_k)$$

$$= \sum_{i=1}^{k} P(B|A_i)P(A_i)$$

Copyright Prof. Vanja Dukic, Applied Mathematics, CU-Boulder

Example

An individual has 3 different email accounts:

70% of messages come into account #1,

20% come into account #2

10% into account #3.

Of the messages in account #1, only 1% are spam, whereas the corresponding percentages for accounts #2 and #3 are 2% and 5%, respectively.

What is the probability that a randomly selected message is spam?

Copyright Prof. Vanja Dukic, Applied Mathematics, CU-Boulder

Example

cont'd

To answer this question, let's first establish some notation:

A_i = {message is from account **#** i} for i = 1, 2, 3,

B = {message is spam}

Then the given percentages imply that

$P(A_1) = .70$, $P(A_2) = .20$, $P(A_3) = .10$

$$P(B|A_1) = .01,\ P(B|A_2) = .02,\ P(B|A_3) = .05$$

Copyright Prof. Vanja Dukic, Applied Mathematics, CU-Boulder

Example

cont'd

Now it is simply a matter of substituting into the equation for the law of total probability:

$P(B) = (.01)(.70) + (.02)(.20) + (.05)(.10) = .016$

In the long run, 1.6% of this individual's messages will be spam.

This is like a weighted average of the spam probabilities, weighted by the probability of each account.

Example

cont'd

$P(B) = (.01)(.70) + (.02)(.20) + (.05)(.10) = .016$

Now, say we randomly selected a message and it was indeed spam. What is the probability that it came from account #1?

In other words, we are looking for $P(A_1 \mid B)$

We can find that by $P(A_1 \mid B) = P(A_1 \cap B) / P(B)$

where $P(A_1 \cap B) = P(B \mid A_1) P(A_1)$, and P(B) was found above.

Then, $P(A_1 \mid B) = (.01)(.70) / 0.016 = 43.75\%$

Bayes' Theorem

In general, let $A_1, A_2, \ldots, A_k$ be a collection of k mutually exclusive and exhaustive events with ***prior*** probabilities $P(A_i)$

Then for any other event B for which $P(B) > 0$, the ***posterior*** probability of A_j given that B has occurred is

$$P(A_j \mid B) = \frac{P(A_j \cap B)}{P(B)} = \frac{P(B \mid A_j)P(A_j)}{\sum_{i=1}^{k} P(B \mid A_i) \cdot P(A_i)} \quad j = 1, \ldots, k$$

Example:

A_1 = person has the flu (probability of flu = $P(A_1)$ = 5%)

A_2 = he/she doesn't have the flu (probability $P(A_2)$ = 100-5=% = 95%)

B = flu test came back positive

What is the **updated probability** that the person has the flu, **$P(A_1 \mid B)$** = ?

Bayes' Theorem: diagnostic test example

A_1 = person has the flu (probability of flu = "flu prevalence" at the time = $P(A_1)$ = 5%)

A_2 = he/she doesn't have the flu (probability $P(A_2)$ = 100-5=% = 95%)

B = flu test came back positive

What is the **updated probability** that the person has the flu, **$P(A_1 \mid B)$** = ?

$$P(A_j \mid B) = \frac{P(A_j \cap B)}{P(B)} = \frac{P(B \mid A_j)P(A_j)}{\sum_{i=1}^{k} P(B \mid A_i) \cdot P(A_i)} \quad j = 1, \ldots, k$$

We first have to find the numerator, $P(B|A_1) P(A_1) = P(B|A_1) * 0.05$

The term $P(B|A_1)$ is the probability of a positive test given that the person actually has the flu. It is called "true positive" or "sensitivity" in diagnostic testing and it is the ability of the test to correctly detect the presence of the disease.

This probability is determined by the test manufacturers and FDA, who usually do validation. For flu, let's assume sensitivity is 0.98.

Then, the numerator is 0.98*0.05 = 0.049

Bayes' Theorem: diagnostic test example

A_1 = person has the flu (probability of flu = "flu prevalence" at the time = P(A_1) = 5%)

A_2 = he/she doesn't have the flu (probability P(A_2) = 100-5=% = 95%)

B = flu test came back positive

What is the **updated probability** that the person has the flu, **P(A_1 | B)** = ?

$$P(A_j|B) = \frac{P(A_j \cap B)}{P(B)} = \frac{P(B|A_j)P(A_j)}{\sum_{i=1}^{k} P(B|A_i)\cdot P(A_i)} \quad j = 1,\ldots,k$$

Next, we have to find the denominator.

The denominator is the total probability of B, and can be found by summing over all the ways we can get B (positive test) to happen: the test can be **positive correctly (true positive),** or **positive incorrectly (false positive)**.

False positive is based on specificity, or probability that the test is negative for a disease-free person (true negative). In fact, false positive = 1-specificity. Let's assume that specificity is 0.96 for the flu test. Then:

P(B) = **P(B|A_1)** P(A_1) + **P(B|A_2)** P(A_2) = **0.98*0.05** + (**1-0.96)*0.95** = 0.087, or 8.7%

Bayes' Theorem: diagnostic test example

A_1 = person has the flu (probability of flu = "flu prevalence" at the time = P(A_1) = 5%)

A_2 = he/she doesn't have the flu (probability P(A_2) = 100-5=% = 95%)

B = flu test came back positive

What is the **updated probability** that the person has the flu, **P(A_1 | B)** = ?

Sensitivity of the test = P(B|A_1) = 0.98 (true positive rate)

Specificity of the test = P(B'|A_2) = 0.96 (true negative rate)

Given all the above, using Bayes theorem we find that the probability of having the flu, given that the test came back positive, is

P(B|A_1) P(A_1) / [**P(B|A_1)** P(A_1) + **P(B|A_2)** P(A_2)] =

0.98*0.05 / [0.98*0.05 + (**1-0.96)*0.95]** = 0.049/0.087 = 0.563 = 56.3%

This number is close to a coin flip! It does go up with increased reliability of the test – i.e. higher test specificity and test sensitivity.