Mathematics for Political Science

Lecture 5 – Probability & Looking Ahead

Sarah B. Bouchat

University of Wisconsin–Madison*

August 22, 2016

*Thanks to Dave Ohls and Brad Jones for past years' teaching materials!

Overview

Many important events have a random (stochastic) element to them—outcomes are uncertain, and may take on different values according to chance.

This uncertainty may arise either as a result of...

- a truly random data-generating process
- a combination of many small factors that may be difficult to observe or are theoretically uninteresting
- inherently unobservable events or conditions

The language, techniques, and tools of probability provide a systematic way to discuss these events, how they relate to each other, and the patterns they exhibit.

Probability in Political Science

- Discuss the chances of uncertain political events
- Estimate the likelihood of certain unobservable conditions producing the data and events observed
- Speak about the degree of certainty about the estimated patterns from the data
- Represent the beliefs and perceptions of strategic actors

Agenda

(1) Counting
(2) Set Theory
(3) Probability
(4) Applying Probabilities
(5) Bayes' Rule
(6) Conclusion & Looking Ahead

Counting

To calculate probabilities, it is often necessary to count the outcomes where the event in question occurs (outcomes with a particular property) as well as count all possible outcomes. The probability the event occurs is their ratio.

With k distinct components of a data generating process and n_k possible alternative outcomes for each stage, the number of possible distinct outcomes for the process as a whole will be:

$$\prod_{i}^{k} n_k$$

Two relevant dimensions for problems of selecting objects from a pool of candidates:

- is order of selection is meaningful or not
- are selected individuals replaced (can be selected again) or not

Ordered, Replacement

The number of possible ways to select k objects from a larger pool of n candidates where order matters and replacement occurs:

$$n * n * n * ... * n * n = n^k$$

Intuition: In each draw there are n possible outcomes. Each can be combined with any of the n choices from any (and all) other draws.

Ordered, No Replacement (Permutation)

The number of possible ways to select k objects from a larger pool of n candidates where order matters and replacement does not occur:

$$n * (n-1) * (n-2) * .... * (n-k+1) = \frac{n!}{(n-k)!}$$

Intuition: In each successive draw there is now one fewer possible outcome, so the candidate pool of objects to be selected reduces by one each time. Outcomes in each round are still paired with all others from all other rounds.

Unordered, No Replacement (Combination)

The number of possible ways to select k objects from a larger pool of n candidates where order does not matter and replacement does not occur:

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

Intuition: This reduces the result from "ordered with no replacement" since some are now substantively identical (picking A first and B second vs. picking B first and A second).

Unordered, Replacement

The number of possible ways to select k objects from a larger pool of n candidates where order does not matter and replacement does occur:

$$\frac{(n+k-1)!}{(n-1)!k!} = \binom{n+k-1}{k}$$

Intuition: Not terribly straightforward, but adjusts upwards the "unordered without replacement" result to account for more choices.

Counting

Consider a group of 10 of which 3 are to be selected. How many different possible combinations are there:

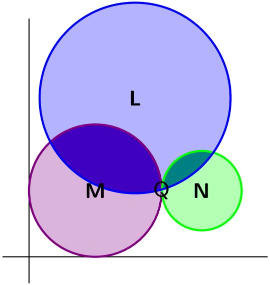
- ordered, with replacement
- ordered, no replacement
- unordered, no replacement
- unordered, with replacement

Sets

A *set* is a collection of elements. This could take the form of a list of units, a range of numbers, or a spatial area in any number of dimensions.

F: {1, 2, 3, 4, 5, 6}
G: {1, 3, 5}
H: {2, 4, 6}

R: [0,1]
S: [0,.7]
T: [.5,1]
U: [.5,1]
V: [.8,1]



P: {Reagan, Bush41, Clinton, Bush43, Obama}
D: {Carter, Mondale, Dukakis, Clinton, Gore, Kerry, Obama}
R: {Reagan, Bush41, Dole, Bush43, McCain}
I: {Perot, Nader}

Everything and Nothing

The *sample space* (denoted S or Ω) is the set that contains all elements in question. This can also be known as the *universal set*.

A *null set* or *empty set* (denoted $\emptyset$) is a set that contains no elements.

A *singleton* is a set that contains exactly one element. A *non-singleton* is a set that contains multiple elements.

Subsets and Supersets

Set A is a *subset* of B (denoted $A \subset B$), and B is a *superset* of A (denoted $B \supset A$) , if all elements of A are included in B.

Subset and Superset

$$A \subset B \equiv B \supset A \equiv \forall\, X \in A, X \in B$$

Recall sets:

- F: {1, 2, 3, 4, 5, 6}
- G: {1, 3, 5}
- H: {2, 4, 6}

G ⊂ F and H ⊂ F

- (G and H are subsets of F)

F ⊃ G and F ⊃ H

- (F is a superset of G and H)

Equal Sets and Disjoint Sets

Set A is *equal* to set B (denoted $A = B$) if they contain exactly the same elements. Set A and set B are *disjoint* (or *mutually exclusive*) if they share no common elements.

Equality

$$A = B \equiv \forall\ X \in A, X \in B \text{ and } \forall\ Y \in B, Y \in A$$

Disjointness

$$A \text{ and } B \text{ are disjoint } \equiv \forall\ X \in A, X \notin B \text{ and } \forall\ Y \in B, Y \notin A$$

Recall sets:
R: [0,1] , S: [0,.7], T: [.5,1],
U: [.5,1], V: [.8,1]

T = U: T and U contain exactly the same elements
S and V are disjoint: they contain no elements in common

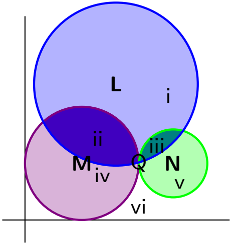
Intersections

The *intersection* of sets A and B (denoted $A \cap B$) is the set of all elements they have in common - all elements that are in both sets.

Intersection

$$A \cap B \equiv X \text{ such that } X \in A \text{ and } X \in B$$

Recall sets:



$L \cap M = \{ii\}$

- (the intersection of L and M is region ii)

$L \cap N = \{iii\}$

- (the intersection of L and N is region iii)

$M \cap N = \emptyset$

- (the intersection of M and N is the null set - they are disjoint)

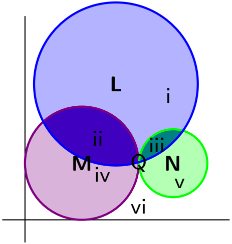
Unions

The *union* of sets A and B (denoted $A \cup B$) is the set of all elements that are in either set.

Union

$$A \cup B \equiv X \text{ such that } X \in A \text{ or } X \in B$$

Recall sets:



$L \cup M = \{i, ii, iii, iv\}$

- (the union of L and M is regions i, ii, iii, and iv)

$L \cup N = \{i, ii, iii, v\}$

- (the union of L and N is regions i, ii, iii, and v)

$M \cup N = \{ii, iii, iv, v\}$

- (the union of M and N is regions ii, iii, iv, and v)

Complements

The *complement* of set A (denoted A^C) is the set of all elements in the sample space (universal set) that are not in set A.

Complement

$$A^C \equiv \text{X such that X} \notin A$$

Recall sets:
R: [0,1] , S: [0,.7],
T: [.5,1], U: [.5,1],
V: [.8,1]

$R^C = \emptyset$

- the complement of R is the null set - R includes the all elements

$S^C = (.7, 1]$

- the complement of S is the region bounded by .7 (open) on the left and 1 (closed) on the right

$T^C = [0, .5)$

- the complement of T is the region bounded by 0 (closed) on the left and .5 (open) on the right

Set Theory

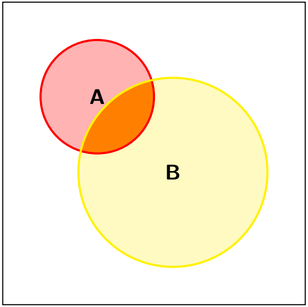
Recall sets:

- R: [0,1]
- S: [0,.7]
- T: [.5,1]
- U: [.5,1]
- V: [.8,1]

- Which sets are subsets of T?
- Which sets are supersets of V?
- What is the set $S \cap U$?
- What is the set $S \cup V$?
- What is the set $(S \cap T) \cup (U \cap V)$?
- What is the complement of V?

Probability

The *probability* of A is the likelihood event A occurs, denoted $p(A)$.



This can be conceptualized as an area of probability space representing probability 1 (100% of the likelihood), where the event A covers some sub-set of that probability space, according to how likely it is.

Probability

Playing cards:

- 4 suits: Hearts, Diamonds, Spades, Clubs
- 13 card values: Ace, 2, 3, 4, 5, 6, 7, 8, 9, 10, Jack, Queen, King
  - 3 face cards: Jack, Queen, King

	A	2	3	4	5	6	7	8	9	10	J	Q	K
♡H													
♢D													
♠S													
♣C													

Total probability area: 1

Probability of drawing any particular card: $\frac{1}{52}$

Conceptualizing Probability

Two main ways to think about probability:

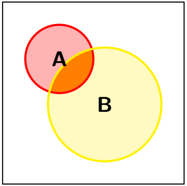
- Over the long run of a very large number of repeated, controlled trials, how often (in what fraction of trials, with what frequency) would some event happen?
- Given a situation and all available evidence about the process, what is the most rational belief about how likely an event is to happen?

Properties of Probabilities

Key properties of all probabilities:

- A probability must be strictly bounded on the closed interval between 0 and 1. Either it is *impossible* for A to occur (probability 0), or it is *certain* that A will occur (probability 1), or somewhere *in between*.
  - $p(A) \in [0, 1]$
- Given a set of exhaustive and mutually exclusive outcomes, their probabilities must sum to 1. *Some outcome* must happen.
  - $\sum p(A_i) = 1$

Probability of Complements



$$p(A^C) = 1 - p(A)$$

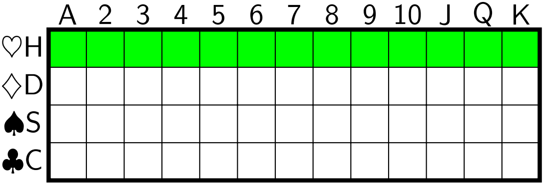
The probability of the complement of an event is...

- ... 1 minus the probability of that event.

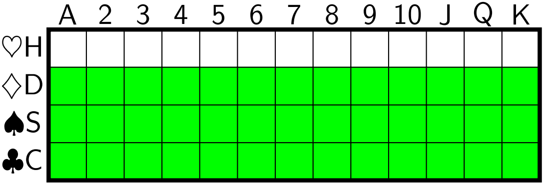
Intuition: Since *something* must happen with probability 1, either an event or its complement will occur (by definition).

Probability of Complements

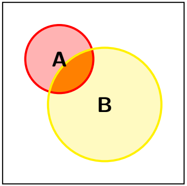
$p(H) = \frac{1}{4}$:



$p(H^C) = 1 - p(H) = \frac{3}{4}$:



Probability of Unions



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

The probability of the union of two events (either or both of them occurring) is...

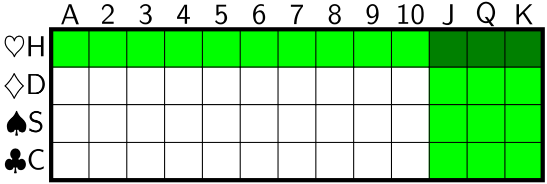
- ... the sum of the probabilities of each event...
- ... minus the probability of their intersection.

Intuition: Collect all the probability area covered by both events, and subtract that area which is counted twice.

Probability of Unions

Probability of drawing a card that is *either* a heart *or* a face card:

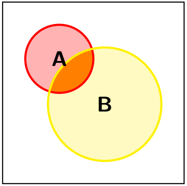
$$p(H \cup F) = p(H) + p(F) - p(H \cap F)$$



$p(H) = \frac{1}{4}$ $\quad p(F) = \frac{3}{13}$ $\quad p(H \cap F) = \frac{3}{52}$

$$\begin{aligned} p(H \cup F) &= \frac{1}{4} + \frac{3}{13} - \frac{3}{52} \\ &= \frac{13}{52} + \frac{12}{52} - \frac{3}{52} \\ &= \frac{22}{52} \end{aligned}$$

Probability of Intersections



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

The probability of the intersection of two events (both of them occurring) is...

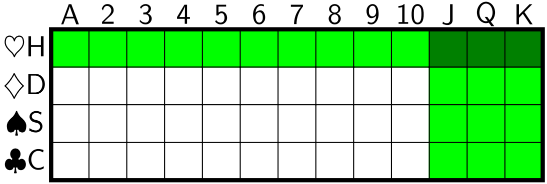
- ... the sum of the probabilities of each event...
- ... minus the probability of their union.

Intuition: Collect all the area they cover, and subtract all of it once, leaving *only* that area which was initially counted twice.

Probability of Intersections

Probability of drawing a card that is *both* a heart *and* a face card:

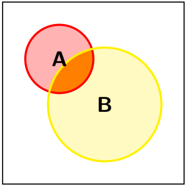
$$p(H \cap F) = p(H) + p(F) - p(H \cup F)$$



$$p(H) = \tfrac{1}{4} \qquad p(F) = \tfrac{3}{13} \qquad p(H \cup F) = \tfrac{22}{52}$$

$$\begin{aligned} p(H \cup F) &= \frac{1}{4} + \frac{3}{13} - \frac{22}{52} \\ &= \frac{13}{52} + \frac{12}{52} - \frac{22}{52} \\ &= \frac{3}{52} \end{aligned}$$

Conditional Probability



$$p(A|B) = \frac{p(A \cap B)}{p(B)}$$

The probability of one event given that another has occurred is...

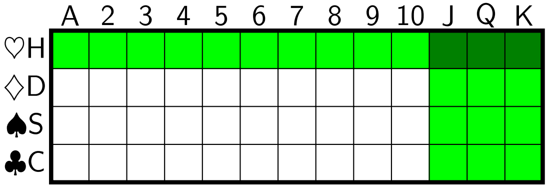
- ... the probabilities of their intersection...
- ... divided by the probability of the event known to have occurred.

Intuition: The given event is known to have occurred, so only that subset of probability space is relevant. Within that, the probability of the other event (the conditional probability) is the area where they overlap.

Conditional Probability

Probability of drawing a card that is a face card *given* that a heart has been drawn:

$$p(F|H) = \frac{p(F \cap H)}{p(H)}$$



$$p(F \cap H) = \frac{3}{52} \qquad p(H) = \frac{1}{4} = \frac{13}{52}$$

$$p(F|H) = \frac{\frac{3}{52}}{\frac{13}{52}} = \frac{3}{13}$$

Probabilities

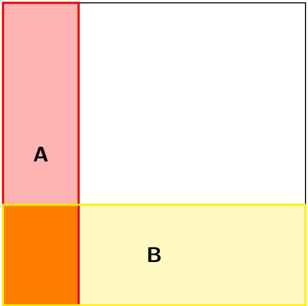
	A	2	3	4	5	6	7	8	9	10	J	Q	K
♡H													
♢D													
♠S													
♣C													

Calculate the following probabilities:

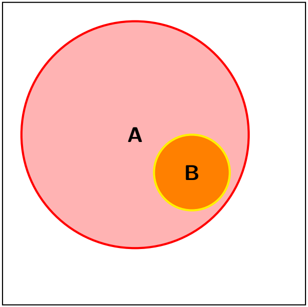
- $p(\{8, 9, 10\})$
- $p(\{5, 6, 7, 8\} \cup \{8, 9, 10\})$
- $p(\{5, 6, 7, 8\} \cap \{D\})$
- $p((\{2, 3\} \cap \{C, S\})^C)$
- $p(\{5\} | \{5, 6, 7, 8\})$

Independence of Events

Two events are *independent* if knowing the outcome of one event provides no information on outcome of the other.



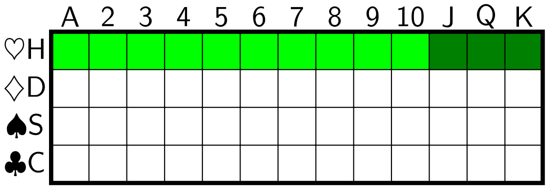
independent



not independent

Independence of Events

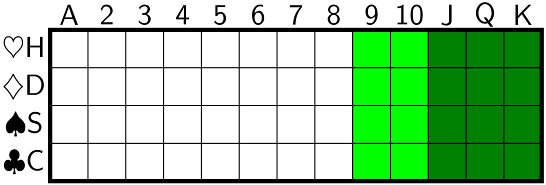
Drawing a heart and drawing a face card are independent events:



$$p(F|H) = \frac{3}{13} = p(F)$$

Independence of Events

Drawing a card greater than 8 and drawing a face card are *not* independent events:



$$p(F| > 8) = \frac{12}{20} = \frac{3}{5} \neq p(F)$$

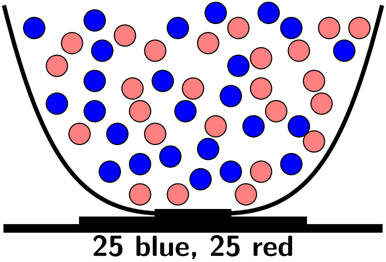
Independent Events

Explain why the following events are or are not independent:

- $p(\{2,3,4,5,6\}$ and $p(D)$
- $p(\{2,3,4,5,6\}$ and $p(\{7,8\})$

Applying Probability

To understand probability and its manipulations, it is useful to think in terms of straightforward, concrete examples.

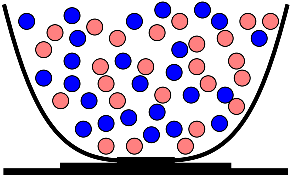


Consider an urn containing a large number of balls of different colors, which may be drawn at random.

An 'event' is drawing a ball of a certain color, or a certain sequence of colors, or a certain pair of colors from different urns.

Probability

Consider drawing one ball from the urn:



25 blue, 25 red

The probability of any given colored ball is the ratio of balls of that color to the total number of balls:

$$p(\text{blue}) = \frac{25}{50} = .5$$
$$p(\text{red}) = \frac{25}{50} = .5$$

Because there are the same number of each color, the probabilities are equal.

Joint Probability

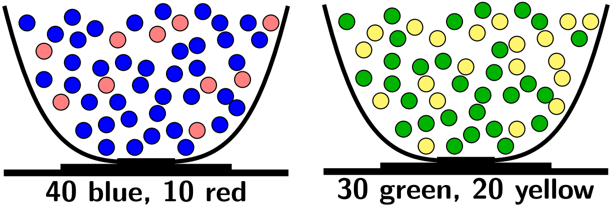
The *joint probability* (denoted $p(A, B)$) of multiple events is the probability that all occur.

This is identical to the probability of the intersection of multiple events in set notation.

For independent events, the joint probability is the product of their individual probabilities.

Joint Probability

Consider randomly drawing one ball from each of two urns:



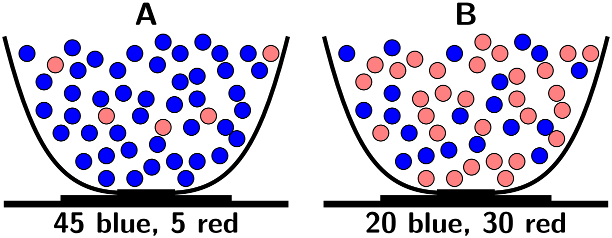
The joint probability of...

$$
\begin{aligned}
\text{p(blue, green)} &= (\tfrac{40}{50})(\tfrac{30}{50}) = (.8)(.6) = .48 \\
\text{p(blue, yellow)} &= (\tfrac{40}{50})(\tfrac{20}{50}) = (.8)(.4) = .32 \\
\text{p(red, green)} &= (\tfrac{10}{50})(\tfrac{30}{50}) = (.2)(.6) = .12 \\
\text{p(red, yellow)} &= (\tfrac{10}{50})(\tfrac{20}{50}) = (.2)(.4) = .08
\end{aligned}
$$

Note that since these are exhaustive and mutually exclusive outcomes, their probabilities sum to 1.

Joint Probability

Now consider randomly flipping a fair coin to choose between two urns, and then randomly drawing a ball from the one chosen:



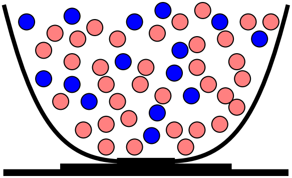
There are two ways of choosing a blue ball.

- Choosing A, and drawing a blue ball:
  - $p(A, \text{blue}) = (\frac{1}{2})(\frac{45}{50}) = (.5)(.9) = .45$
- Choosing B, and drawing a blue ball:
  - $p(B, \text{blue}) = (\frac{1}{2})(\frac{20}{50}) = (.5)(.4) = .20$

The total probability of drawing a blue ball is the sum of these joint probabilities:

Joint Probability

Consider instead drawing randomly 5 balls from a single urn, replacing each time:



The probability of a *specific sequence* blue, red, blue, blue, red:

$$(.3)(.7)(.3)(.3)(.7) = .01323$$

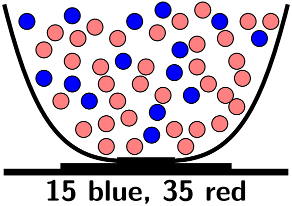
The probability of drawing 3 blue balls and 2 red balls in any order:

$$\binom{5}{3}(.3)^3(.7)^2 = (10)(.01323) = .1323$$

- The $\binom{5}{3}$ term counts the number of unique places where the 3 blue balls could be drawn (since we don't care about order, any of these generate the outcome of interest).

Joint Probability

Consider instead drawing randomly 5 balls from a single urn with *no replacement*:



The probability of a *specific sequence* blue, red, blue, blue, red:

$$(\frac{15}{50})(\frac{35}{49})(\frac{14}{48})(\frac{13}{47})(\frac{34}{46}) = \frac{3,248,700}{254,251,200} = \frac{32,487}{2,542,512} \approx .052$$

Conditional Probability

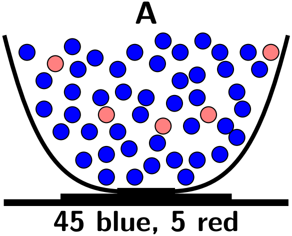
The *conditional probability* (denoted $p(A|B)$) of two events is the probability that the first occurs, given that the second has occurred.

This is identical to the conditional probability of events in set notation.

For independent events, the conditional probability $p(A|B)$ is equal to the probability $p(A)$.

Conditional Probability

Consider again randomly flipping a fair coin to choose between two urns, and then randomly drawing a ball from the one chosen:

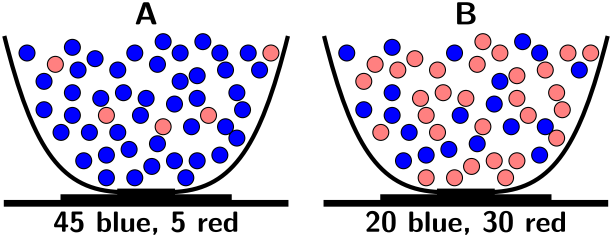


The probabilities of choosing a blue ball *given* a certain urn is chosen are:

$$\mathrm{p}(\mathrm{blue}|\mathrm{A}) = \frac{45}{50} = .90 \qquad \mathrm{p}(\mathrm{blue}|\mathrm{B}) = \frac{20}{50} = .40$$

Conditional Probability

Suppose instead a die were used to choose the urn, with a $\{1,2,3,4,5\}$ result selecting urn A and a $\{6\}$ result selecting urn B:



The conditional probabilities of choosing a blue ball *given* a certain urn is chosen are unchanged:

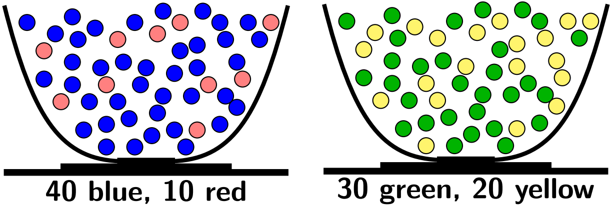
$$p(\text{blue}|A) = \tfrac{45}{50} = .90 \qquad p(\text{blue}|B) = \tfrac{20}{50} = .40$$

But the total probability that a blue ball will be chosen is now higher:

$$\begin{aligned} p(\text{blue}) &= p(A, \text{blue}) + p(B, \text{blue}) \\ &= \left(\frac{5}{6}\right)\left(\frac{45}{50}\right) + \left(\frac{1}{6}\right)\left(\frac{20}{50}\right) = \frac{49}{60} \approx .81667 \end{aligned}$$

Conditional Probability

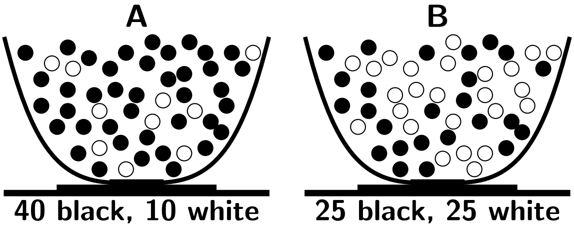
Consider again drawing one ball from each of two urns:



Since the draws are independent, whether a blue or red ball is chosen from the first has no impact on the probability of a green or yellow ball being chosen from the second.

$$
\begin{aligned}
\text{p(green|blue)} &= \text{p(green)} = \left(\tfrac{30}{50}\right) = .60 \\
\text{p(green|red)} &= \text{p(green)} = \left(\tfrac{30}{50}\right) = .60 \\
\text{p(yellow|blue)} &= \text{p(yellow)} = \left(\tfrac{20}{50}\right) = .40 \\
\text{p(yellow|red)} &= \text{p(yellow)} = \left(\tfrac{20}{50}\right) = .40
\end{aligned}
$$

Applying Probabilities



Find the following joint or conditional probabilities:

- Drawing one ball from each urn and both are black.
- Flipping a fair coin to choose an urn and drawing a white ball from that urn.
- Mixing all balls together and drawing a black ball.
- The probability of drawing a black ball conditional on a coin flip having selected A.
- Drawing a black ball from B conditional on having drawn a black ball from A.

Bayes' Rule

Bayes' Rule (or *Bayes' Theorem* or *Bayes' Law*) relates conditional probabilities to their inverse (the reverse conditional probability).

This is a very powerful tool to infer the probability of unobservable events or circumstances, using information on the outcome of some observable event or signal that is (even imperfectly) related.

Bayes Rule

Formally, Bayes' Rule says that the conditional probability of some unobservable event given some signal is...

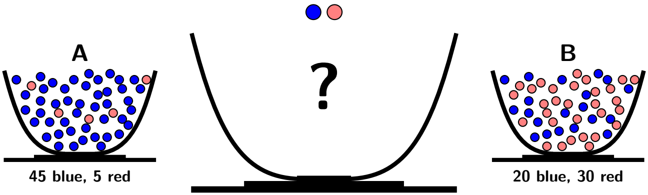
- ... the joint probability of that event producing that signal...
- ... divided by the sum of the joint probabilities of all possible events producing that signal.

Bayes Rule

$$p(E|s) = \frac{p(E)p(s|E)}{p(E)p(s|E) + p(E^C)p(s|E^C)} = \frac{p(E)p(s|E)}{\sum p(x)p(s|x)}$$

Inverse Conditional Probability

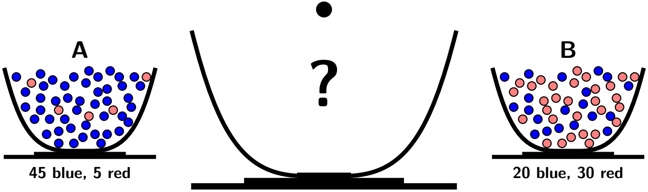
Consider once again drawing a ball from one of two urns:



But now suppose the urn is hidden, and consider the reverse conditional probabilities: the probabilities it came from one urn or the other, given a certain color of ball.

- $\mathrm{p}(A|\mathrm{blue})$
- $\mathrm{p}(A|\mathrm{red})$

Inverse Conditional Probability



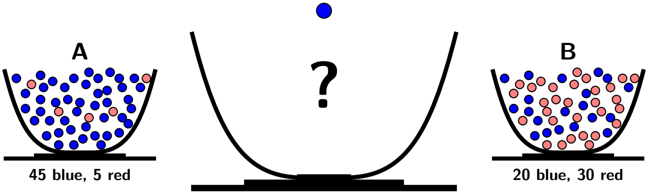
Before observing the color of the ball, the only information about which urn it came from is the prior probability.

Because the coin is fair, the probability for the urns must be even.

$$p(A) = p(B) = .5$$

Bayesian Updating

Suppose the ball is observed to be blue. This provides information that makes it possible to update the probability estimate.

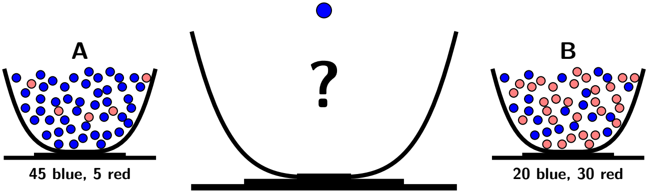


Recall the conditional probabilities of a blue ball for each urn, based on the proportions of colors of balls in the urn:

- $p(\text{blue}|A) = .90$
- $p(\text{blue}|B) = .40$

This suggests intuitively that the updated probability of the ball having come from urn A should increase (from the prior .5), since it is more likely to produce a blue ball.

Bayesian Updating



Using Bayes' Rule, the probability the ball came from urn A or B must be:

$$
\begin{aligned}
p(A|\text{blue}) &= \frac{p(A)p(\text{blue}|A)}{p(A)p(\text{blue}|A) + p(B)p(\text{blue}|B)} \\
&= \frac{(.5)(.9)}{(.5)(.9) + (.5)(.4)} \\
&\approx .69
\end{aligned}
$$

$$
\begin{aligned}
p(B|\text{blue}) &= \frac{p(B)p(\text{blue}|B)}{p(B)p(\text{blue}|B) + p(A)\text{p(blu}} \\
&= \frac{(.5)(.4)}{(.5)(.4) + (.5)(.9)} \\
&\approx .31
\end{aligned}
$$

- The probabilities sum to 1.
- The probability the ball came from urn A has increased.

Bayesian Updating

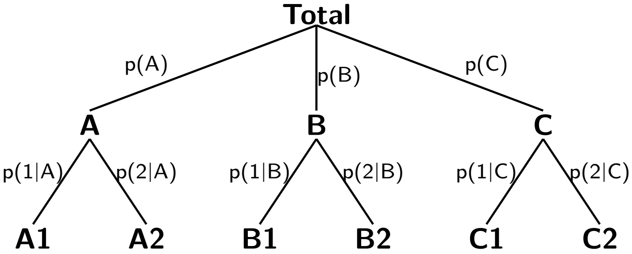
Consider a rare condition that occurs in only 1% of some population, which has a test that is 95% accurate: 95% of individuals with the condition will test positive, and 95% of individuals without the condition will test negative.

If an individual tests positive, the probability it has the condition is:

$$
\begin{aligned}
p(C|pos) &= \frac{p(C)p(pos|C)}{p(C)p(pos|C) + p(\sim C)p(pos|\sim C)} \\
&= \frac{(.01)(.95)}{(.01)(.95) + (.99)(.05)} \\
&= \frac{.0095}{.0590} \\
&\approx .16
\end{aligned}
$$

Intuition of Bayes

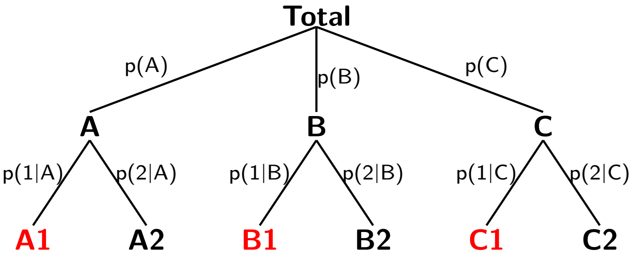
Consider dividing a population according to two uncertain events:



- The first (unobservable) may result in outcomes A, B, or C
  - The proportions in each group are the prior probabilities
  - p(A), p(B), p(C)
- The second (observable) may result in outcomes sending signal 1 or 2
  - The proportions in each sub group are the conditional probabilities
  - p(1|A), p(2|A), p(1|B), p(2|B), p(1|C), p(2|C)

Intuition of Bayes

Suppose signal 1 is observed as the outcome of the second event:



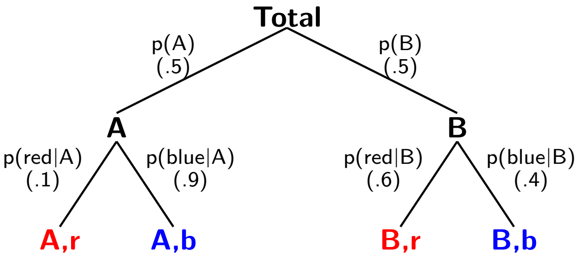
The probability it came from the joint outcome involving event A is:

$$\begin{aligned} p(A|1) &= \frac{p(A,1)}{p(A,1)+p(B,1)+p(C,1)} \\ &= \frac{p(A)p(1|A)}{p(A)p(1|A)+p(B)p(1|B)+p(C)p(1|C)} \\ &= \frac{p(A)p(1|A)}{\sum p(x)p(1|x)} \end{aligned}$$

This is exactly Bayes' Rule.

Applying Bayesian Intuition

Consider again drawing a ball from one of two urns:

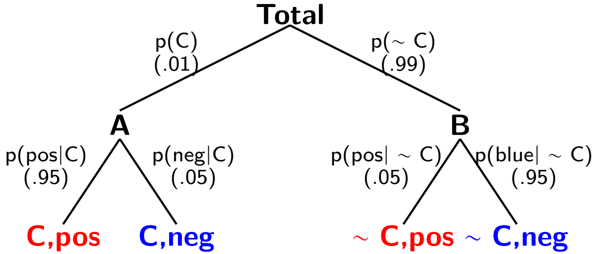


The probability it was urn A if a blue ball is drawn is:

$$\begin{aligned}
p(A|blue) &= \frac{p(A, blue)}{p(A, blue) + p(B, blue)} \\
&= \frac{p(A)p(blue|A)}{p(A)p(blue|A) + p(B)p(blue|B)} \\
&= \frac{(.5)(.9)}{(.5)(.9) + (.5)(.4)} \\
&\approx .69
\end{aligned}$$

Applying Bayesian Intuition

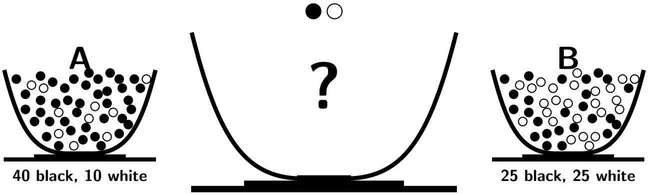
Consider again the rare condition



The probability of the condition given a positive test is:

$$
\begin{aligned}
p(C|pos) &= \frac{p(C)p(pos|C)}{p(C)p(pos|C) + p(\sim C)p(pos| \sim C)} \\
&= \frac{(.01)(.95)}{(.01)(.95) + (.99)(.05)} \\
&= \frac{.0095}{.0590}
\end{aligned}
$$

Bayesian Updating



Consider an unobserved die roll to choose an urn, with a 1 or 2 choosing the left urn and a 3, 4, 5, or 6 choosing the right urn. Find the conditional probabilities of each urn having been selected given the color of the ball drawn.

- $p(A|black)$
- $p(A|white)$
- $p(B|black)$
- $p(B|white)$

Courses & Training Opportunities

Foundations

- 812* & 813*, 817, 835*, 553*

Going further...

- Formal: 836, 837
- Empirical: 818, [Bayes], 919, 948* PED, Ethnography
- 999

Other departments

- Formal: ECON 711, 713
- Empirical: AAE, Sociology, Statistics/Biostat, Anthropology
- ITV

Summer/non-semester opportunities

- ICPSR
- EITM
- IPSA short courses and conferences

Technical Skills

- Languages
- Computing: R, STATA, Stan, C++/Java, Python
- LaTeX

Ways to Move Forward

- Take the foundations courses, then select what interests you
- Minor in Methodology
  - 3 courses, specialize in some field of methodology
  - 2 "substantive" fields as primary focus
- Second field Methodology
  - Foundational courses plus courses across 2 subfields (e.g., empirical, formal)
  - 2 paper prelim: one empirical application, one review
  - Not signaling "good at methods;" signaling "I want to teach methodology courses and conduct research in methdology"
- First field Methodology
  - Still need "substantive" interests, but want to focus on methodological questions
  - Small community
  - Sat exam prelim in subfields of specialization