DSA 8010 - Probability foundations

What is probability?

Probability is a branch of mathematics that deals with uncertain outcomes.

- Probability is the foundation of inference, the framework we use to draw conclusions using data that are limited and variable.
- Some of the skills required for probability calculations include logic, counting, and measuring.

Combinatorics

Permutations

There are four chairs in an office and four people plan to sit down. How many unique seating arrangements are possible?

Factorials

Factorial function. For any integer $n \geq 0$, $n!$ is defined as

$$n! = \begin{cases} n \cdot (n-1) \cdot \ldots \cdot 1 & n = 1, 2, \ldots, . \\ 1 & n = 0 \end{cases}$$

There are $k!$ ways to rearrange k objects.

Permutations

There are four chairs in an office and six people need to sit down. How many unique seating arrangements are possible?

Permutations

Counting permutations. There are $k!$ ways to rearrange k objects.

Permutation formula. The number of ways to rearrange k items from a total of n items is

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

Combinations

There are four chairs in an office and six people need to sit down. How many unique combinations of people sitting are possible, regardless of their arrangement?

Permutations and combinations

Counting permutations. There are $k!$ ways to rearrange k objects.

Permutation formula. The number of ways to rearrange k items from a total of n items is

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

Combination formula (Binomial coefficient). The number of ways to select k items from a total of n items, ignoring their order, is

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

often denoted as

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

Permutations with replacement

There are 20 flavors of ice cream at Jeni's. I plan to get three scoops and am willing to repeat flavors. How many ordered combinations are possible?

Permutations and combinations

Counting permutations. There are $k!$ ways to rearrange k objects.

Permutation formula. The number of ways to rearrange k items from a total of n items is

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

Combination formula (Binomial coefficient). The number of ways to select k items from a total of n items, ignoring their order, is

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

Permutations "with replacement." The number of unique (ordered) ways to select n items out of k possibilities, with replacement, is k^n.

Permutations and combinations

Counting permutations. There are $k!$ ways to rearrange k objects.

Permutation formula. The number of ways to rearrange k items from a total of n items is ${}_nP_r = \frac{n!}{(n-k)!}$.

Combination formula (Binomial coefficient). The number of ways to select k items from a total of n items, ignoring their order, is ${}_nC_r = \binom{n}{k} = \frac{n!}{(n-k)!k!}$.

Permutations "with replacement." The number of unique (ordered) ways to select n items out of k possibilities, with replacement, is k^n.

Warning: this list is not exhaustive.

Permutations and combinations

Counting permutations. There are $k!$ ways to rearrange k objects.

Permutation formula. The number of ways to rearrange k items from a total of n items is ${}_nP_r = \frac{n!}{(n-k)!}$.

Combination formula (Binomial coefficient). The number of ways to select k items from a total of n items, ignoring their order, is ${}_nC_r = \binom{n}{k} = \frac{n!}{(n-k)!k!}$.

Permutations "with replacement." The number of unique (ordered) ways to select n items out of k possibilities, with replacement, is k^n.

Warning: the verbal descriptions of these (and other combinatorical) scenarios can be difficult to parse. Logic will often serve you better than searching for the correct formula.

Probability experiments

Probability experiments: vocabulary

Experiment. A process from which an outcome is observed.

Examples:

1. Flip a fair coin twice.
2. Roll a fair die once.

Outcome. A measurable result (i.e. a thing that can happen and be observed).

Examples:

1. Coin experiment - HH
2. Die experiment - 5

Probability experiments: vocabulary

Sample space. Set of all outcomes in an experiment.

Examples:

1. Coins - $\mathcal{S} = \{HH, HT, TH, TT\}$
2. Die - $\mathcal{S} = \{1, 2, 3, 4, 5, 6\}$

Probability experiments: vocabulary

Sample space. Set of all outcomes in an experiment.

Examples:

1. Coins - $\mathcal{S} = \{HH, HT, TH, TT\}$
2. Die - $\mathcal{S} = \{1, 2, 3, 4, 5, 6\}$

Event. A subset of the sample space.

Examples:

1. Coins - $E =$ I get at least one tails.
   $E = \{HT, TH, TT\}$.
2. Die - $A =$ I roll a number less than 5.
   $A = \{1, 2, 3, 4\}$.

Definition of probability

Let E be some event.

1. Classical definition.

$$P(E) = \frac{\#\text{ outcomes in E}}{\#\text{ outcomes in } \mathcal{S}}$$

2. Relative frequency interpretation (empirical approach)

$$P(E) = \frac{\#\text{ times } E \text{ occurred}}{\#\text{ number of possibilities for } E \text{ to occur}}$$

3. Subjective probability

Example: cards

Consider a probability experiment in which you draw one card from a deck of 52 cards.

Define two events:
$A =$ the card is red.
$B =$ the card is a face card.

Find $P(B)$ and $P(A)$.

Example: cards

Use H, S, D, and C to denote hearts, spade, diamonds, and clubs, respectively.

$$\mathcal{S} = \begin{Bmatrix} 1H & 2H & 3H & 4H & \dots & 10H & JH & QH & KH \\ 1S & 2S & 3S & 4S & \dots & 10S & JS & QS & KS \\ 1D & 2D & 3D & 4D & \dots & 10D & JD & QD & KD \\ 1C & 2C & 3C & 4C & \dots & 10C & JC & QC & KC \end{Bmatrix}$$

$P(A) =$

$P(B) =$

Basic probability rules

- $0 \leq P(E) \leq 1$
- If $E = \emptyset$, then $P(E) = 0$. (E cannot occur.)
- If $E = \mathcal{S}$, then $P(E) = 1$. (E always occurs.)

Complements, unions, and intersections

Complement. The complement of E is the set of outcomes not included in E.

Notation: $\bar{E}$ or E'

Union. The union of A and B is the set of outcomes included in either A or B (including those outcomes in both A and B).

Notation: $A \cup B$

Intersection. The intersection of A and B is the set of outcomes included in both A and B.

Notation: $A \cap B$

Disjoint. Two events are *disjoint* if $A \cap B = \emptyset$, where $\emptyset$ is the empty set containing no outcomes. More simply, disjoint events never occur together.

Venn diagrams

Example: cards

Consider a probability experiment in which you draw one card from a deck of 52 cards.

Define two events:
$A =$ the card is red.
$B =$ the card is a face card.

Find $P(B')$, $P(A \cup B)$, and $P(A \cap B)$.

Example: cards

$A =$ the card is red; $B =$ the card is a face card.

Use H, S, D, and C to denote hearts, spade, diamonds, and clubs, respectively.

$$\mathcal{S} = \left\{ \begin{array}{ccccccccc} AH & 2H & 3H & 4H & \ldots & 10H & JH & QH & KH \\ AS & 2S & 3S & 4S & \ldots & 10S & JS & QS & KS \\ AD & 2D & 3D & 4D & \ldots & 10D & JD & QD & KD \\ AC & 2C & 3C & 4C & \ldots & 10C & JC & QC & KC \end{array} \right\}$$

$P(B') =$

$P(A \cup B) =$

$P(A \cap B) =$

Example: cards

Probabilities associated with two events can also be represented using a two-way table.

	face card	number card	Total
red	$6/52$	$20/52$	$26/52$
black	$6/52$	$20/52$	$26/52$
Total	$12/52$	$40/52$	$52/52 = 1$

More probability rules

Complement rule.

$$P(E') = 1 - P(E)$$

Probabilities of unions.

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

Probabilities of intersections.

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

Example: cards

Define the events A, B, and C as follows:
$A =$ card is black.
$B =$ card is a Queen.
$C =$ card is hearts.

$P(A \cup B) =$

$P(B') =$

$P(A \cap C) =$

Conditional probability

Conditional probability

The conditional probability of "A given B" is the probability of A based on the knowledge that B has occurred.

Notation: $P(A|B)$

Examples:

1. Roll one fair die. Let $A =$ roll a 4 and $B =$ roll an even number.

Conditional probability

The conditional probability of "A given B" is the probability of A based on the knowledge that B has occurred.

Notation: $P(A|B)$

Examples:

2. Select a student at random from a university population. Let $A =$ the student is above 68 inches. $B =$the student is female.

Calculating conditional probabilities

If you have $\mathcal{S}$ and can count outcomes:

$$P(A|B) = \frac{\text{no. outcomes in } A \cap B}{\text{no. outcomes in } B}$$

More generally:

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

Conditional probability examples

1. Die experiment. Let $A =$ roll a 4 and $B =$ roll an even number. Find $P(A|B)$.

Conditional probability examples

2. Select a student at random from a university population. Let $A =$ the student is above 68 inches. $B =$the student is female. Find $P(A|B)$. Use the following information about the university population.

	female	male	total
below 68"	0.46	0.26	0.72
above 68"	0.11	0.17	0.28
Total	0.57	0.43	1

Conditional probability examples

2. Select a student at random from a university population. Let $A =$ the student is above 68 inches. $B =$the student is female. Find $P(B|A)$. Use the following information about the university population.

	female	male	total
below 68"	0.46	0.26	0.72
above 68"	0.11	0.17	0.28
Total	0.57	0.43	1

Multiplication rule

Probabilities of intersections can be found using conditional probabilities by applying the multiplication rule.

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

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

Multiplication rule

Warning!

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

$$P(A \cap B) \neq P(A|B)P(A)$$
$$P(A \cap B) \neq P(B|A)P(B)$$

Independence

Independence

Definition of independence

If events A and B are independent, then the following hold:

1. $P(A|B) = P(A)$
2. $P(A \cap B) = P(A)P(B)$

- Independence is the absence of association between events or variables. If two events are independent, then knowing the that one event occurred does not affect what I know about the probability of the other event.

Example: cards

$A =$ the card is red.
$B =$ the card is a face card.

Verify that A and B are independent by showing that they satisfy the following properties.

1. $P(A|B) = P(A)$
2. $P(A \cap B) = P(A)P(B)$

Independence

Independence might be known/assumed because of the nature of our experiment. Then properties (1) and (2) make some probability calculations easier. Examples:

- consecutive coin flips are independent.
- "independent samples:" in a sample that contains n observations of a random variable (X), the outcome for each sampled unit is independent of the outcome for any other sampled unit. In other words, knowing the value of X for one unit does not tell me any information about likely values for any other units.

Probabilities of independent events

Multiplication rule for independent processes

If the events A_1, A_2, $\ldots, A_k$ are independent, then the probability of all events occurring is

$$P(A_1)P(A_2)\ldots P(A_k).$$

Example: suppose 70% of voters support a proposed tax bill. In a random sample of 5 voters, what is the probability that all of them support the bill?

Sampling with replacement

An urn contains 17 green balls and 6 yellow balls. A ball is drawn at random, its color noted, and is replaced in the urn. This is repeated 5 times (sampling with replacement). What is the probability that all five draws are green?

Sampling without replacement

An urn contains 17 green balls and 6 yellow balls. Fives balls are drawn at random, without replacing the selected balls in the urn. What is the probability that all five draws are green?

Independent samples

Often, it is reasonable to assume independent (or at least approximately independent) samples if data are collected using a simple random sample or other well-defined, reasonable sampling scheme.

- Convenience and volunteer samples are less likely to produce independent samples.
- Measurements are often *dependent* over time and/or space.

  Examples: measure the air humidity on a grid of locations throughout South Carolina, record the height of a child once a month for 2 years.

Bayes' theorem

Bayes' theorem

For events A and B such that $P(A) \neq 0$,

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

Bayes' theorem

For events A and B such that $P(A) \neq 0$,

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

- Bayes' theorem is the foundation for Bayesian statistics, a popular framework for statistical inference.
- In foundational probability calculations, the theorem is useful when you want to "reverse" a conditional probability.

Bayes' theorem

Here is a more general, but equivalent statement. Let $B_1, \ldots, B_K$ be disjoint events such that $\mathcal{S} = \{B_1 \cup B_2 \cup B_k\}$. (The sets $B_1, \ldots, B_K$ form a partition of the sample space.)

Then for any i from 1 to K,

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{k=1}^{K} P(A|B_k)P(B_k)}.$$

Example: diagnostic testing

A diagnostic test is developed for a disease that is present in 10% of the patient population. The test produces a positive result in 95% of patients who have the disease. It also incorrectly produces a positive result in 15% of patients who do not have it.

If a patient tests positive, what is the probability that they do indeed have the disease?

Example: diagnostic testing

A diagnostic test is developed for a disease that is present in 10% of the patient population. The test produces a positive result in 95% of patients who have the disease. It also incorrectly produces a positive result in 15% of patients who do not have it.

If a patient tests positive, what is the probability that they do indeed have the disease?

Example: book store

A book store classifies customers as heavy, medium, or light purchasers, and separate mailings are prepared for each of these groups. Overall, 20% of purchasers are heavy, 30% are medium, and 50% are light. A member is classified 36 months after the first purchase, but a test is made of the feasibility of using the first 6 months' purchases to classify members. The following percentages are obtained from existing records of individuals classified into the purchasing groups.

	Group		
First 6 month's purchases	*Heavy*	*Medium*	*Light*
0	0.10	0.15	0.75
1	0.20	0.70	0.15
2+	0.70	0.15	0.10

What do you notice about this table of probabilties?

An Introduction to Statistical Methods and Data Analysis, Ott & Longnecker

Example: book store (multiplication rule)

A book store classifies customers as heavy, medium, or light purchasers, and separate mailings are prepared for each of these groups. Overall, 20% of purchasers are heavy, 30% are medium, and 50% are light. A member is classified 36 months after the first purchase, but a test is made of the feasibility of using the first 6 months' purchases to classify members. The following percentages are obtained from existing records of individuals classified into the purchasing groups.

	Group		
First 6 month's purchases	*Heavy*	*Medium*	*Light*
0	0.10	0.15	0.75
1	0.20	0.70	0.15
2+	0.70	0.15	0.10

Find the probability that a randomly selected selected customer is a Light purchaser and made 1 purchase in the first 6 months.

Example: book store (Bayes' rule)

	Group		
First 6 month's purchases	*Heavy*	*Medium*	*Light*
0	0.10	0.15	0.75
1	0.20	0.70	0.15
2+	0.70	0.15	0.10

If a customer made 0 purchases in the first 6 months, what is the probability that they are a "Light" purchaser?.

Probabilities using combinatorics

Example 1 (PIN)

A 4 digit PIN is selected at random. What is the probability that there are no repeated digits?

Example 2a (lottery)

In a certain state's lottery, 48 balls numbered 1 through 48 are placed in a machine and six of them are drawn at random. If the six numbers drawn match the numbers that a player had chosen, the player wins $1,000,000. In this lottery, the order the numbers are drawn in doesn't matter. Compute the probability that you win the million-dollar prize if you purchase a single lottery ticket.

Example 2b (lottery)

In a certain state's lottery, 48 balls numbered 1 through 48 are placed in a machine and six of them are drawn at random. If five of the six numbers drawn match the numbers that a player has chosen, the player wins a second prize of $1,000. In this lottery, the order the numbers are drawn in doesn't matter. Compute the probability that you win the second prize if you purchase a single lottery ticket.

Probabilities using simulation

Recap: frequency interpretation of probability

Let E be some event.

1. Classical definition.

$$P(E) = \frac{\#\text{ outcomes in E}}{\#\text{ outcomes in }\mathcal{S}}$$

2. Relative frequency interpretation (empirical approach)

$$P(E) = \frac{\#\text{ times } E \text{ occurred}}{\#\text{ number of possibilities for } E \text{ to occur}}$$

3. Subjective probability

Finding probabilities with simulation

Consider a probability experiment and some event E that is a subset of the sample space.

- Write a program that performs the probability experiment using random numbers.
- Perform the probability experiment and note whether the event E has occurred.
- Repeat the experiment for a large number of iterations.
- The approximate probability of E is

$$\frac{\text{no. of times E occurred}}{\text{no. simulations}}.$$

Example (fair die)

```
> # roll a fair die one time
> sample(1:6,1,replace=TRUE, prob=rep(1/6,6))
[1] 5
>
> # find the probability that a 5 is rolled
> nsims <- 10000
> die.results <- rep(NA,nsims)
>
> for( i in 1:nsims)
+ {
+ die.results[i] <-sample(1:6,1,replace=TRUE, prob=rep(1/6,6))
+ }
>
> table(die.results==5)

FALSE  TRUE
 8349  1651
> prop.table(table(die.results==5))

 FALSE   TRUE
0.8349 0.1651
```

Example (PIN revisited)

A 4 digit PIN is selected at random. What is the probability that there are no repeated digits?