

1 Probability, Random Variables, and Probability Distributions

1.1 Basic Probability and Simulations

- Outcomes are governed by chance, but in many repetitions, a pattern emerges.
- Sample space: all of the possible outcomes.
- Event: a specific desired outcome or set of outcomes.
- Notation: probability of Event A $\rightarrow P(A)$.
- Range of probabilities: probability of event is between 0 and 1.
- Sum of probabilities: probability of whole sample is 1.

Theoretical probability is what should happen given the sample space: $\frac{\# \text{ of desired outcomes}}{\text{total } \# \text{ outcomes}}$.

Empirical probability is when you are performing a simulation or experiment, it is what does happen given the trials: $\frac{\# \text{ successes}}{\# \text{ trials}}$.

The law of large numbers states that simulated probabilities tend to get closer to the true probability as the number of trials increases.

The notation for the probability that an event does not occur is $P(A^C)$. The probability an event does not occur is $P(A^C) = 1 - P(A)$.

- The idea of probability is that randomness is predictable in the long run.
- Probability does not allow us to make short run predictions.
- Probability tells us random behavior events out in the long run.
- Future outcomes are not affected by past behavior.

The imitation of chance behavior, based on a model that accurately reflects the situation, is called a simulation. Simulations are usually done with a table of random digits, random number generator, dice, deck of cards, spinner, etc.

Four principles of simulation:

- State: identify the probability calculation.
- Plan: Describe how to use your chance process.
- Do: Perform the simulation
- Conclude: use the results of your simulation to answer the question.

1.2 The Addition Rule

- When two events have no outcomes in common, we refer to them as mutually exclusive events.
- $P(A \text{ and } B) = 0$
- $P(A \text{ or } B) = P(A) + P(B)$

If A and B are any two events resulting from some chance process, then

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

Notice how if A and B are mutually exclusive, $P(A \text{ and } B) = 0$ and $P(A \text{ or } B) = P(A) + P(B)$.

1.3 Venn Diagrams and the Multiplication Rule

- We can represent events with a Venn diagram - a display of potential probabilities.
- The box around the Venn diagram represents the total sample space where $P(S) = 1$.
- The circles themselves represent the probability of each event.

Union: \cup is the probability that either occurs.

Intersection: \cap is the probability both will occur.

- Two events are independent if they do not influence one another.
- The occurrence of one has no effect on the occurrence of the other.
- If A and B are independent events, then the probability that both A and B occur is found using the multiplication rule:

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

- Two events are dependent if they influence one another.
- The occurrence of one affects the occurrence of the other.
- If A and B are dependent events, then the probability that both A and B occur is found using:

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

- Where $P(B|A)$ is the probability B "given" that A has occurred.
- The conditional probability of an event is the probability that one event will happen, if it is known that another event has happened.

1.4 Conditional Probability and Tree Diagrams

- If A and B are dependent events, then the probability that both A and B occur is found using $P(A \cap B) = P(A) \cdot P(B|A)$ where $P(B|A)$ is the probability of B "given" that A has already occurred.
- The conditional probability of an event is the probability that one event will happen, if it is known that another event has happened.
- We can rearrange the multiplication rule a little and get a formula for conditional probability.

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

- If we draw two cards with replacement, we know that those events are independent.
- If we draw two cards without replacement, we know that those events are dependent. Using the following formula we can prove independence:

$$P(A \cap B) = P(A) \cdot P(B) \text{ or } P(A) = P(A|B)$$

1.5 Discrete and Continuous Random Variables

A random variable takes numerical values that describe the outcomes of some chance process. Random variables involve the same probability rules we have already learned, but we will extend those rules to be able to model more events.

A probability distribution tells us the value that our random variable can take and the probability associated with each value.

Every probability distribution must satisfy each of the following requirements.

1. There is a numerical random variable X , and each of the values of X has an associated probability with it.
2. The sum of all the probabilities in the distribution $\sum P(x)$ must equal 1.
3. $0 \leq P(x) \leq 1$ for all probabilities in the distribution.

A discrete random variable takes on a fixed set of possible values with whole number outcomes.

- Random variables are usually capital letters.
- Random variables can be discrete or continuous
- Random variables must be numeric in value

The mean of a discrete random variable X , is the mean outcome for infinitely many trials.

We think of this as the “expected” value because it is the average value we would expect to get if the trials could continue indefinitely.

To find the mean, μ_x , or expected value, $E(x)$, of a discrete random variable X , multiply each possible value by its probability, then add all the products.

$$E(x) = \sum x_i \cdot p_i$$

The standard deviation of a random variable X is a measure of how much the values of the variable typically vary from the expected value.

$$\sigma_x^2 = Var(X) = \sum (x_i - \mu_x)^2 \cdot p_i$$

$$\sigma_x = \sqrt{\sum (x_i - \mu_x)^2 \cdot p_i}$$

- Continuous Random Variables take on all possible values in an interval of numbers. The probability distribution of X is described by a density curve.
- The probability of any event is the area under the density curve and above, below, or between the values x that define the event.
- Area under a density curve is always equal to 1.
- The probability is on the y-axis in a distribution, so finding the area equates to finding the probability of getting an interval of numbers.
- The probability at an event is 0 for a continuous random variable because the area under a point is 0.

Calculator: Mean and Standard Deviation from a Probability Distribution Table

- Enter outcomes into L1 and probabilities in L2
- Run 2nd - Vars - 1VarStats with List:L1 and FreqList:L2
- Mean will be represented by \bar{x}
- Standard deviation will be represented by σ_x

1.6 Combining Random Variables

If X and Y are two independent random variables:

$$E(X + Y) = E(X) + E(Y)$$

$$E(X - Y) = E(X) - E(Y)$$

$$\sigma_{X+Y} = \sqrt{\sigma_x^2 + \sigma_y^2}$$

$$\sigma_{X-Y} = \sqrt{\sigma_x^2 + \sigma_y^2}$$

If X is a random variable and a and b are both constants:

$$E(a + bX) = a + bE(x)$$

$$\sigma_{a+bX} = |b|E(x)$$

1.7 The Binomial Distribution

Requirements for a Bernoulli Trial:

1. Two Possible Outcomes
2. Probability of Success is the Same for Each Trial
3. Trials are Independent

If we take a Bernoulli trial and then we are interested in the number of successes in a specific number of trials, we create a Binomial Probability Distribution.

There are four requirements for a setting to follow the binomial probability distribution. As long as these four things are true, we can use the binomial probability model for calculating the probabilities.

- BINARY: Each trial falls into one of two categories - we call them "success" or "failure". Success does not necessarily mean a positive outcome, but instead, the outcome we are looking for.
- Independent Trials
 - Each trial is independent of the next
 - Reasonable assumption for coins, cards with replacement, spinners, rolling a die
 - What about sampling without replacement, when it can't be avoided? This is where the "10% condition" comes in. The 10% condition says that if you are sampling from a large enough population, you can proceed with sampling without replacement as though the trials were independent.
- Number of Trials is fixed: we are counting the number of successes from a set number of trials (called n)
- Same probability: The probability of successes (called p) is the same for each trial

If all four points are satisfied, you can compute the probability you get a certain number of successes in a specified number of trials:

n is the number of independent trials, p is the probability of successes, and x is the number of successes.

$$P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}$$

where $\binom{n}{x} = nCx = \frac{n!}{x!(n-x)!}$

The mean is $\mu_x = np$ and the standard deviation is $\sigma_x = \sqrt{np(1-p)}$.

Distributions in your calculator often have "pdf" and "cdf".

pdf is the probability distribution function and is used when trying to calculate $P(X = k)$ and cdf is the cumulative distribution function and is used when you are trying to calculate $P(X \leq k)$.

Specifically for Binomial distributions we want

- 2nd - VARS - A: binompdf(trials, probability of success, k) for $P(X = k)$
- 2nd - VARS - B: binomcdf(trials, probability of success, k) for $P(X \leq k)$

1.8 The Geometric Distribution

When we perform many independent trials of the same chance process and are interested in the occurrence of the first success, a geometric setting arises.

If X has a geometric distribution with a probability of success p on each trial and x represents the trial that you get your first success, the probability that X equals x is given by

$$P(X = x) = (1 - p)^{x-1}(p)$$

Specifically for geometric distributions we want...

- 2nd - Vars - E: `geompdf(probability of success, k)` for $P(X = x)$
- 2nd - Vars - F: `geometcdf(probability of success, k)` for $P(X \leq x)$

For this distribution, $\mu_x = \frac{1}{p}$ and $\sigma_x = \frac{\sqrt{1-p}}{p}$