

Summary (Unit 3B - Random Variables)

We have almost reached the end of our discussion of probability. We were introduced to the important concept of **random variables**, which are quantitative variables whose value is determined by the outcome of a random experiment.

We discussed discrete and continuous random variables.

We saw that all the information about a **discrete random variable** is packed into its probability distribution. Using that, we can answer probability questions about the random variable and find its **mean and standard deviation**. We ended the part on discrete random variables by presenting a special class of discrete random variables – **binomial random variables**.

As we dove into **continuous random variables**, we saw how calculations can get complicated very quickly, when probabilities associated with a continuous random variable are found by calculating **areas under its density curve**.

As an example for a continuous random variable, we presented the **normal random variable**, and discussed it at length. The normal distribution is extremely important, not just because many variables in real life follow the normal distribution, but mainly because of the important role it plays in statistical inference, our ultimate goal of this course.

We learned how we can avoid calculus by using the **standard normal calculator or table** to find probabilities associated with the normal distribution, and learned how it can be used as an **approximation to the binomial** distribution under certain conditions.

Random Variables

A random variable is a variable whose values are numerical results of a random experiment.

- A **discrete random variable** is summarized by its probability distribution — a list of its possible values and their corresponding probabilities.

The sum of the probabilities of all possible values must be 1.

The probability distribution can be represented by a table, histogram, or sometimes a formula.

- The **probability distribution** of a random variable can be supplemented with numerical measures of the center and spread of the random variable.

Center: The center of a random variable is measured by its mean (which is sometimes also referred to as the **expected value**).

The mean of a random variable can be interpreted as its long run average.

The mean is a weighted average of the possible values of the random variable weighted by their corresponding probabilities.

Spread: The spread of a random variable is measured by its variance, or more typically by its standard deviation (the square root of the variance).

The standard deviation of a random variable can be interpreted as the typical (or long-run average) distance between the value that the random variable assumes and the mean of X .

Binomial Random Variables

- The binomial random variable is a type of discrete random variable that is quite common.
- The binomial random variable is defined in a random experiment that consists of n independent trials, each having two possible outcomes (called “success” and “failure”), and each having the same probability of success: p . Such a random experiment is called the binomial random experiment.
- The binomial random variable represents the number of successes (out of n) in a binomial experiment. It can therefore have values as low as 0 (if none of the n trials was a success) and as high as n (if all n trials were successes).
- There are “many” binomial random variables, depending on the number of trials (n) and the probability of success (p).
- The probability distribution of the binomial random variable is given in the form of a formula and can be used to find probabilities. Technology can be used as well.
- The mean and standard deviation of a binomial random variable can be easily found using short-cut formulas.

Continuous Random Variables

The probability distribution of a continuous random variable is represented by a probability density curve. The probability that the random variable takes a value in any interval of interest is the area above this interval and below the density curve.

An important example of a continuous random variable is the **normal random variable**, whose probability density curve is symmetric (bell-shaped), bulging in the middle and tapering at the ends.

- There are “many” normal random variables, each determined by its mean μ (mu) (which determines where the density curve is centered) and standard deviation σ (sigma) (which determines how spread out (wide) the normal density curve is).
- Any normal random variable follows the Standard Deviation Rule, which can help us find probabilities associated with the normal random variable.
- Another way to find probabilities associated with the normal random variable is using the standard normal table. This process involves finding the z-score of values, which tells us how many standard deviations below or above the mean the value is.
- An important application of the normal random variable is that it can be used as an approximation of the binomial random variable (under certain conditions). A continuity correction can improve this approximation.

Summary (Unit 3B - Random Variables) is shared under a [CC BY-NC-SA 4.0](https://creativecommons.org/licenses/by-nc-sa/4.0/) license and was authored, remixed, and/or curated by LibreTexts.