# ST 516: Foundations of Data Analytics

## Discrete Random Variables

# Random Variable Distribution

Recall that a random variable is a value that is assigned to the outcome of a random experiment.

Since the value of a random variable depends on the outcome of some experiment, we can talk about the *probability* of the random variable taking on its different possible values.

## Definition

We refer to these probabilities as the **distribution** of the random variable: The distribution of a random variable describes the probability of the random variable taking on its different possible values.

# More Random Variable Notation

We often use lower case letters like $x, y$, or $z$ to denote particular values that a random variable can take on.

Notation Recap:

- Capital letters like $X, Y$, or $Z$ are used to represent **random variables**. These letters represent a quantity that can take on different values depending on the outcome of an experiment.
- Lower case letters like $x$, $y$, and $z$ are used to represent particular values

# Random Variable Distribution Properties

We are typically interested in quantities like

- $P(X = x)$: The probability that the random variable $X$ takes on the particular value $x$.
- $P(X \leq w)$: The probability that the random variable $X$ is less than or equal to the particular value $w$
- $P(y < X < z)$: The probability that the random variable $X$ is between the values $y$ and $z$.

# Random Variable Distribution Examples

Example:

- Random Trial: Roll a die
- Outcome space (set of all possible outcomes): The faces '1', '2', '3', '4', '5', and '6'.
- Random Variables of Interest:

$$X = \begin{cases} -5 & \text{If the number showing on the die is '1', '2', or '3'} \\ 0 & \text{If the number showing on the die is '4' or '5'} \\ 10 & \text{If the number showing on the die is '6'} \end{cases}$$

What is the distribution of the random variable $X$? We answer this question by describing the probability that $X$ takes on each of its possible values: -5, 0, and 10.

# Random Variable Distribution Examples

Since the die is fair, we know that

$$P(\text{Die shows 1}) = P(\text{Die shows 2}) = P(\text{Die shows 3}) =$$
$$P(\text{Die shows 4}) = P(\text{Die shows 5}) = P(\text{Die shows 6}) = \frac{1}{6}$$

Then we can make the following table:

| Outcome of Die Roll | Value of $X$ ($x$) | Probability ($\mathrm{P}(X = x)$) |
|---|---|---|
| 1 | -5 | $\frac{1}{6}$ |
| 2 | -5 | $\frac{1}{6}$ |
| 3 | -5 | $\frac{1}{6}$ |
| 4 | 0 | $\frac{1}{6}$ |
| 5 | 0 | $\frac{1}{6}$ |
| 6 | 10 | $\frac{1}{6}$ |

# Random Variable Distribution Examples

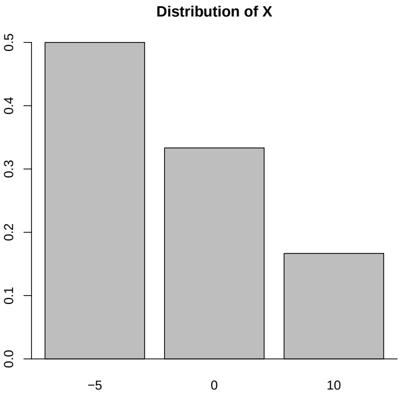
Now we can combine rows that have the same value of the random variable $X$ to summarize the probability distribution of the random variable $X$. For instance, the probability that the random variable $X$ takes on the value $-5$ is the probability of getting a '1', '2', or '3' when we roll the die.

The distribution of $X$ can be described as follows:

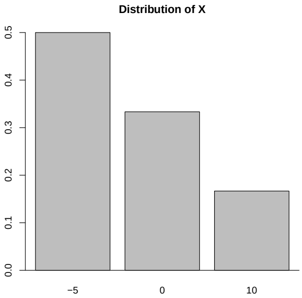
| Value of $X$ ($x$) | Probability ($\mathrm{P}(X = x)$) |
|---|---|
| -5 | $\frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$ |
| 0 | $\frac{1}{6} + \frac{1}{6} = \frac{1}{3}$ |
| 10 | $\frac{1}{6}$ |

# Random Variable Distribution Examples

We can represent the distribution of a discrete random variable using a **bar graph** (or **bar chart**).



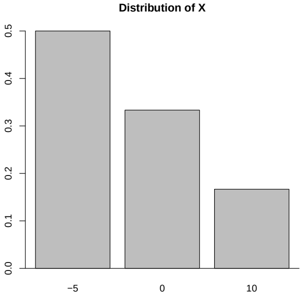
# Random Variable Distribution: Understanding and Interpreting




The height of the bar above each number on the $x$-axis represents probability that $X$ takes on that value.

For instance, the height of the bar above the number '-5' is 0.5, meaning that $P(X = -5) = 0.5$.

# Random Variable Distribution: Understanding and Interpreting




Similarly, we see that the height of the bar above the number '0' is $0.33 = \frac{1}{3}$, meaning that $P(X = 0) = 0.33$ and the height of the bar above the number '10' is $0.167 = \frac{1}{6}$ meaning that $P(X = 10) = 0.167$.

Important Properties of random variable distribution:

- Note that if we add up the probabilities of each of the possible values of $X$, we get 1:

$$\frac{1}{2} + \frac{1}{3} + \frac{1}{6} = 1$$

  This is always the case: for *any* random variable, if we add up the probabilities across all of its possible values we get 1.
- Equivalently, if we add up the heights of the bars in a barplot of the random variable's distribution, we must get 1.

Important Properties of random variable distribution:

- The different values are *mutually exclusive*. For instance, if on a particular roll of the die we get $X = -5$, then that means that we *did not* get $X = 0$ or $X = 10$.
    - Therefore, if we are interested in the probability of a certain set of values, we just add up the probabilities corresponding to each value in that set. For instance, in our example, what is $P(X < 3)$?

      There are two possible values of $X$ that are less than 3: -5 and 0. So we have

      $$P(X < 3) = P(X = -5) + P(X = 0) = \frac{1}{2} + \frac{1}{3} = \frac{5}{6}$$

# Common Discrete Random Variable Distributions

The distribution of a random variable tells us everything we can know about how it behaves.

There are certain distributions that we have given names because we encounter and use them so frequently. You may have heard of some of them:

- Bernoulli
- Binomial
- Poisson
- (Discrete) Uniform

We will briefly describe these distributions. Remember that there are many, many other possible distributions for a random variable–just most of them aren't named.

# Common Discrete Random Variable Distributions: Bernoulli

The first discrete distribution that most people encounter is called the **Bernoulli** distribution. This is the distribution of a random variable that can take on only two possible values: 0 or 1.

For instance, suppose you flip a coin. We can define $X$ to be the random variable that takes on the value 1 if you get heads when you flip the coin once, and 0 if you get tails.

The probabilities of these two possible values are determined by a **parameter** $\pi$. A *parameter* is a constant value that determines the shape (probabilities) of a distribution. Later you will encounter distributions with several parameters, but the Bernoulli distribution is completely determined by the single parameter $\pi$, which is a value between 0 and 1.

# Common Discrete Random Variable Distributions: Bernoulli

The probabilities for a random variable $X$ that has the Bernoulli distribution (in terms of the parameter $\pi$) are

| Value $x$ | $P(X = x)$ |
|---|---|
| 0 | $P(X = 0) = 1 - \pi$ |
| 1 | $P(X = 1) = \pi$ |

For example, as before suppose you flip a coin and $X$ is 1 if you get heads, 0 if you get tails. If the coin has probability $\dfrac{3}{4}$ of coming up heads, and $X =$ Then $X$ has the Bernoulli distribution with $\pi = \frac{3}{4} = 0.75$.

We sometimes use the notation $X \sim \text{Bernoulli}(\pi)$ as a short-hand way of writing "$X$ is distributed according to the Bernoulli distribution with parameter $\pi$".

# Common Discrete Random Variable Distributions: Binomial

The **binomial** distribution is often described as the distribution of the number of 'successes' in $n$ independent trials, each of which is a 'success' with probability $\pi$ or a 'failure' with probability $1-\pi$.

The values $n$ and $\pi$ are *parameters* of the binomial distribution. We generally know $n$ (how many trials were done), but we do not know $\pi$. We typically want to use the number of successes (which is our random variable $Y$) to learn about the value of $\pi$.

The possible values for a random variable $Y$ with a binomial distribution are $0, 1, 2, \ldots, n$. The probability that $Y$ takes on a particular one of these values, represented by $k$, is described by the formula

$$P(Y=k) = \binom{n}{k} \pi^k (1-\pi)^{n-k}$$

where $\binom{n}{k}$ means '$n$ choose $k$': how many different ways you can choose $k$ objects from $n$ total objects.

# Common Discrete Random Variable Distributions: Poisson

The **Poisson** distribution is often a good approximation to random variables that are counting occurrences of some event. A random variable $X$ that has the Poisson distribution can take on any non-negative integer value: $0, 1, 2, 3, \ldots$.

The shape of the Poisson distribution is determined by a parameter $\lambda$: larger values of the parameter $\lambda$ mean that the random variable $X$ is more likely to take on larger values.

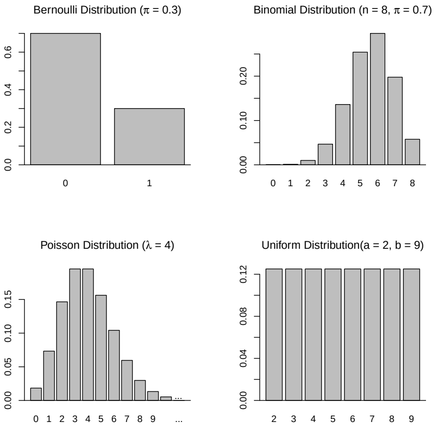
# Common Discrete Random Variable Distributions: Discrete Uniform

The discrete uniform distribution is the distribution of a random variable that is equally likely to be any one of a range of integer values. A random variable $Y$ that has the discrete uniform distribution can take on any integer value between $a$ and $b$.

The numbers $a$ and $b$ are the parameters of the discrete uniform distribution: they determine the set of possible values for the random variable $Y: a, a+1, a+2, \ldots, b$.

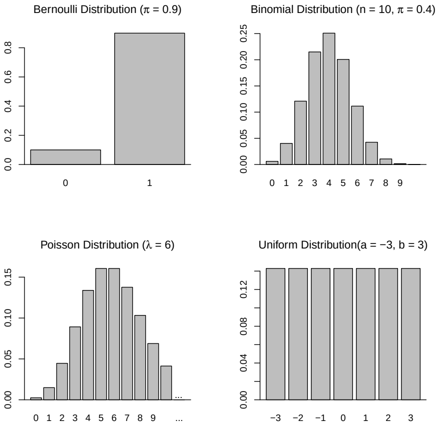
# Common Discrete Random Variable Distributions: Example Barplots

Here are some barplots of the discrete distributions we have just discussed:



# Common Discrete Random Variable Distributions: Example Barplots

And here are some more barplots of these same distribution families, but with different parameters:



# Common Discrete Random Variable Distributions: Example Barplots

Using the plots on the previous two slides, see if you can answer the following questions:

- If $X \sim \text{Binomial}(10, 0.4)$, what is the probability that $X = 6$?

- If $Y \sim \text{Poisson}(\lambda = 4)$, what is the probability that $Y = 8$?

# Common Discrete Random Variable Distributions: Example Barplots

Using the plots on the previous two slides, see if you can answer the following questions:

- If $X \sim$ Binomial$(10, 0.4)$, what is the probability that $X = 6$?

  From the second set of barplots, we look at the Binomial(10, 0.4) distribution and examine the bar over the value 6. It has height approximately equal to $0.11$, so we conclude that if $X \sim$ Binomial$(n = 10, \pi = 0.4)$, then $P(X = 6) \approx 0.11$ (in fact, we can calculate this probability exactly and find that it is 0.1115).

- If $Y \sim$ Poisson$(\lambda = 4)$, what is the probability that $Y = 8$?

  From the first set of barplots, we look at the Poisson(4) distribution and examine the bar over the value 8. It has height approximately equal to $0.03$, so we conclude that if $Y \sim$ Poisson$(\lambda = 4)$, then $P(X = 8) \approx 0.03$.