Econ 103: Probability and Statistics

Manu Navjeevan

UCLA

June 28, 2021

Content Outline

Single Random Variable

- Discrete and Continuous random variables
- Mean, variance, and expectations

Multiple Random Variables

- Conditional probabilities and conditional means
- Covariance and independence

The Normal Distribution

- Properties and computing probabilities

The Law of Large Numbers and the Central Limit Theorem

- The sample mean as a random variable

Random Variable

Question: What is a random variable?

Intuitively, we can think of a random variable as a model of an outcome that is uncertain.

- Example: Flipping a coin, traffic in the morning, etc.

Random Variable

Question: What is a random variable?

Intuitively, we can think of a random variable as a model of an outcome that is uncertain.

- Example: Flipping a coin, traffic in the morning, etc.

While the outcome is random, the random variable does have a *distribution*. For any subset of the outcome space, the distribution describes the probability that the random variable takes a value in that subset.

- Example: We know that our flipped coin has a 50% probability of taking a value in the set $\{H\}$, a 50% probability of taking a value in the set $\{T\}$, and a 100% probability of taking a value in the set $\{H, T\}$.

Random Variables

Question: Why do we care about random variables? What does this have to do with econometrics?

Consider the population of California. Suppose we want to know about the education levels of people in the population.

- The education level of a randomly selected person in the population is not deterministic. Different people have different levels of education.

Random Variables

Question: Why do we care about random variables? What does this have to do with econometrics?

Consider the population of California. Suppose we want to know about the education levels of people in the population.

- The education level of a randomly selected person in the population is not deterministic. Different people have different levels of education.
- We can think about the education level as a *random variable* with a distribution that describes the probability that a randomly selected person has an education level in certain range.
  - i.e 30% of people have high school diplomas, 40% of people have college degrees, etc.

Random Variables

Question: Why do we care about random variables? What does this have to do with econometrics?

Consider the population of California. Suppose we want to know about the education levels of people in the population.

- The education level of a randomly selected person in the population is not deterministic. Different people have different levels of education.
- We can think about the education level as a *random variable* with a distribution that describes the probability that a randomly selected person has an education level in certain range.
  - i.e 30% of people have high school diplomas, 40% of people have college degrees, etc.
- In general however, we may not know the exact distribution of the random variable. Econometrics is about using a random sample of data to make inferences about the underlying distribution of the random variable.

Single Random Variables: Outcome Spaces

Let's formalize the discussion above. Let X denote a random variable. All random variables come equipped with an *outcome space* that contains all the values that the random variable can take up.

.

Let's formalize the discussion above. Let X denote a random variable. All random variables come equipped with an *outcome space* that contains all the values that the random variable can take up.

- If X represents the flipping a coin, the outcome space is $\{H, T\}$.

.

Single Random Variables: Outcome Spaces

Let's formalize the discussion above. Let X denote a random variable. All random variables come equipped with an *outcome space* that contains all the values that the random variable can take up.

- If X represents the flipping a coin, the outcome space is $\{H, T\}$.
- If X represents rolling a die, the outcome space is $\{1, 2, 3, 4, 5, 6\}$.

.

Single Random Variables: Outcome Spaces

Let's formalize the discussion above. Let X denote a random variable. All random variables come equipped with an *outcome space* that contains all the values that the random variable can take up.

- If X represents the flipping a coin, the outcome space is $\{H, T\}$.
- If X represents rolling a die, the outcome space is $\{1, 2, 3, 4, 5, 6\}$.
- If X represents the 100m sprint time (in seconds) of an Olympic athlete, the outcome space may be $[9.5, 10.5]$.

.

Single Random Variables: Outcome Spaces

Let's formalize the discussion above. Let X denote a random variable. All random variables come equipped with an *outcome space* that contains all the values that the random variable can take up.

- If X represents the flipping a coin, the outcome space is $\{H, T\}$.
- If X represents rolling a die, the outcome space is $\{1, 2, 3, 4, 5, 6\}$.
- If X represents the 100m sprint time (in seconds) of an Olympic athlete, the outcome space may be $[9.5, 10.5]$.

If the outcome space of X is *countable* (think finite), then we say that X is a discrete random variable. If the outcome space of X is *uncountable* (think infinite), we say that X is a continuous random variable.

- Flipping a coin and rolling a die would be discrete random variables
- The 100m sprint time of an Olympic athlete would be a continuous random variable.

Single Random Variables: Outcome Spaces and Probability

In general in this class, we will notate the outcome space of a random variable X as $\mathcal{O}_X$. Let $2^{\mathcal{O}_X}$ denote all the subsets of $\mathcal{O}_X$.

We will typically be interested in the probability that X takes values in some $A \in 2^{\mathcal{O}_X}$ (that is $A \subseteq \mathcal{O}_X$). This probability is a number between 0 and 1 and will be notated as $\mathbb{P}_X(A)$. We will require the probability $\mathbb{P}_X(\cdot)$ to satisfy certain properties:

Single Random Variables: Outcome Spaces and Probability

In general in this class, we will notate the outcome space of a random variable X as $\mathcal{O}_X$. Let $2^{\mathcal{O}_X}$ denote all the subsets of $\mathcal{O}_X$.

We will typically be interested in the probability that X takes values in some $A \in 2^{\mathcal{O}_X}$ (that is $A \subseteq \mathcal{O}_X$). This probability is a number between 0 and 1 and will be notated as $\mathbb{P}_X(A)$. We will require the probability $\mathbb{P}_X(\cdot)$ to satisfy certain properties:

- $\mathbb{P}_X(\mathcal{O}_X) = 1$ and $\mathbb{P}_X(\emptyset) = 0$.

In general in this class, we will notate the outcome space of a random variable X as $\mathcal{O}_X$. Let $2^{\mathcal{O}_X}$ denote all the subsets of $\mathcal{O}_X$.

We will typically be interested in the probability that X takes values in some $A \in 2^{\mathcal{O}_X}$ (that is $A \subseteq \mathcal{O}_X$). This probability is a number between 0 and 1 and will be notated as $\mathbb{P}_X(A)$. We will require the probability $\mathbb{P}_X(\cdot)$ to satisfy certain properties:

- $\mathbb{P}_X(\mathcal{O}_X) = 1$ and $\mathbb{P}_X(\emptyset) = 0$.
- $0 \leq \mathbb{P}_X(A) \leq 1$ for any $A \in 2^{\mathcal{O}_X}$.

In general in this class, we will notate the outcome space of a random variable X as $\mathcal{O}_X$. Let $2^{\mathcal{O}_X}$ denote all the subsets of $\mathcal{O}_X$.

We will typically be interested in the probability that X takes values in some $A \in 2^{\mathcal{O}_X}$ (that is $A \subseteq \mathcal{O}_X$). This probability is a number between 0 and 1 and will be notated as $\mathbb{P}_X(A)$. We will require the probability $\mathbb{P}_X(\cdot)$ to satisfy certain properties:

- $\mathbb{P}_X(\mathcal{O}_X) = 1$ and $\mathbb{P}_X(\emptyset) = 0$.
- $0 \leq \mathbb{P}_X(A) \leq 1$ for any $A \in 2^{\mathcal{O}_X}$.
- If $A_1, A_2, \ldots$ are pairwise disjoint, then $\mathbb{P}_X(\cup_i A_i) = \sum_i \mathbb{P}_X(A_i)$.

Single Random Variables: Outcome Spaces and Probability

In general in this class, we will notate the outcome space of a random variable X as $\mathcal{O}_X$. Let $2^{\mathcal{O}_X}$ denote all the subsets of $\mathcal{O}_X$.

We will typically be interested in the probability that X takes values in some $A \in 2^{\mathcal{O}_X}$ (that is $A \subseteq \mathcal{O}_X$). This probability is a number between 0 and 1 and will be notated as $\mathbb{P}_X(A)$. We will require the probability $\mathbb{P}_X(\cdot)$ to satisfy certain properties:

- $\mathbb{P}_X(\mathcal{O}_X) = 1$ and $\mathbb{P}_X(\emptyset) = 0$.
- $0 \leq \mathbb{P}_X(A) \leq 1$ for any $A \in 2^{\mathcal{O}_X}$.
- If $A_1, A_2, \ldots$ are pairwise disjoint, then $\mathbb{P}_X(\cup_i A_i) = \sum_i \mathbb{P}_X(A_i)$.

When we say we are interested in the *distribution* of the random variable X, we really mean we are interested in $\mathbb{P}_X(\cdot)$ as viewed as a map from $2^{\mathcal{O}_X}$ onto $[0, 1]$.

Single Random Varibles: Discrete Random Variables

If X is a discrete random variable the distribution or probability function $\mathbb{P}_X$ can be described by the *probability mass function* or *pmf*, $p_X(\cdot) : \mathcal{O}_X \to [0, 1]$.

- Recall that a discrete random variable has a countable (think finite) outcome space

.

If X is a discrete random variable the distribution or probability function $\mathbb{P}_X$ can be described by the *probability mass function* or *pmf*, $p_X(\cdot) : \mathcal{O}_X \to [0, 1]$.

- Recall that a discrete random variable has a countable (think finite) outcome space

For each element a of the outcome space ($a \in \mathcal{O}_X$), the probability mass function evaluated at a, $p_X(a)$, describes the probability that X takes value a. That is $p_X(a) = \mathbb{P}_X(\{a\})$.

Single Random Varibles: Discrete Random Variables

If X is a discrete random variable the distribution or probability function $\mathbb{P}_X$ can be described by the *probability mass function* or *pmf*, $p_X(\cdot) : \mathcal{O}_X \to [0, 1]$.

- Recall that a discrete random variable has a countable (think finite) outcome space

For each element a of the outcome space ($a \in \mathcal{O}_X$), the probability mass function evaluated at a, $p_X(a)$, describes the probability that X takes value a. That is $p_X(a) = \mathbb{P}_X(\{a\})$.

By the last property of probability measures, the pmf can be used to recover the probability that X takes values in any subset A of the outcomes space $\mathcal{O}_X$

$$\mathbb{P}_X(A) = \sum_{a \in A} \mathbb{P}_X(\{a\}) = \sum_{a \in A} p_X(a).$$

Single Random Variables: Discrete Random Variables

Let's see an example of this. Let X denote the outcome of a fair dice roll. We can describe the distribution of X via the probability mass function

$$p_X(a) = \begin{cases} \frac{1}{6} & \text{if } a \in \{1,2,3,4,5,6\} \\ 0 & \text{for any other value of } a \end{cases}$$

Let's see an example of this. Let X denote the outcome of a fair dice roll. We can describe the distribution of X via the probability mass function

$$p_X(a) = \begin{cases} \frac{1}{6} & \text{if } a \in \{1,2,3,4,5,6\} \\ 0 & \text{for any other value of } a \end{cases}$$

Let's use the pmf to compute $\mathbb{P}_X(A)$ for $A = \{2,4,6\}$, that is use the pmf to compute the probability that X takes on an even value.

Single Random Variables: Discrete Random Variables

Let's see an example of this. Let X denote the outcome of a fair dice roll. We can describe the distribution of X via the probability mass function

$$p_X(a) = \begin{cases} \frac{1}{6} & \text{if } a \in \{1,2,3,4,5,6\} \\ 0 & \text{for any other value of } a \end{cases}$$

Let's use the pmf to compute $\mathbb{P}_X(A)$ for $A = \{2,4,6\}$, that is use the pmf to compute the probability that X takes on an even value.

$$\mathbb{P}_X(\{2,4,6\}) = \sum_{a \in \{2,4,6\}} \mathbb{P}_X(\{a\})$$

Single Random Variables: Discrete Random Variables

Let's see an example of this. Let X denote the outcome of a fair dice roll. We can describe the distribution of X via the probability mass function

$$p_X(a) = \begin{cases} \frac{1}{6} & \text{if } a \in \{1,2,3,4,5,6\} \\ 0 & \text{for any other value of } a \end{cases}$$

Let's use the pmf to compute $\mathbb{P}_X(A)$ for $A = \{2,4,6\}$, that is use the pmf to compute the probability that X takes on an even value.

$$\begin{aligned} \mathbb{P}_X(\{2,4,6\}) &= \sum_{a \in \{2,4,6\}} \mathbb{P}_X(\{a\}) \\ &= \sum_{a \in \{2,4,6\}} p_X(a) \end{aligned}$$

Single Random Variables: Discrete Random Variables

Let's see an example of this. Let X denote the outcome of a fair dice roll. We can describe the distribution of X via the probability mass function

$$p_X(a) = \begin{cases} \frac{1}{6} & \text{if } a \in \{1,2,3,4,5,6\} \\ 0 & \text{for any other value of } a \end{cases}$$

Let's use the pmf to compute $\mathbb{P}_X(A)$ for $A = \{2,4,6\}$, that is use the pmf to compute the probability that X takes on an even value.

$$\begin{aligned} \mathbb{P}_X(\{2,4,6\}) &= \sum_{a \in \{2,4,6\}} \mathbb{P}_X(\{a\}) \\ &= \sum_{a \in \{2,4,6\}} p_X(a) \\ &= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} \end{aligned}$$

Single Random Variables: Discrete Random Variables

Let's see an example of this. Let X denote the outcome of a fair dice roll. We can describe the distribution of X via the probability mass function

$$p_X(a) = \begin{cases} \frac{1}{6} & \text{if } a \in \{1,2,3,4,5,6\} \\ 0 & \text{for any other value of } a \end{cases}$$

Let's use the pmf to compute $\mathbb{P}_X(A)$ for $A = \{2,4,6\}$, that is use the pmf to compute the probability that X takes on an even value.

$$\begin{aligned} \mathbb{P}_X(\{2,4,6\}) &= \sum_{a \in \{2,4,6\}} \mathbb{P}_X(\{a\}) \\ &= \sum_{a \in \{2,4,6\}} p_X(a) \\ &= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} \\ &= 1/2 \end{aligned}$$

Single Random Variables: Discrete Random Variables

Let's see an example of this. Let X denote the outcome of a fair dice roll. We can describe the distribution of X via the probability mass function

$$p_X(a) = \begin{cases} \frac{1}{6} & \text{if } a \in \{1,2,3,4,5,6\} \\ 0 & \text{for any other value of } a \end{cases}$$

Let's use the pmf to compute $\mathbb{P}_X(A)$ for $A = \{2,4,6\}$, that is use the pmf to compute the probability that X takes on an even value.

$$\begin{aligned} \mathbb{P}_X(\{2,4,6\}) &= \sum_{a \in \{2,4,6\}} \mathbb{P}_X(\{a\}) \\ &= \sum_{a \in \{2,4,6\}} p_X(a) \\ &= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} \\ &= 1/2 \end{aligned}$$

Of course, this result is a bit obvious. However, if the die was not fair, we would follow the same procedure to compute this probability.

Single Random Variables: Continuous Random Variables

If X is a continuous random variable we cannot use a probability mass function to describe its distribution.

Single Random Variables: Continuous Random Variables

If X is a continuous random variable we cannot use a probability mass function to describe its distribution.

- Recall that if X is a continuous random variable then the outcome space $\mathcal{O}_X$ is (uncountably) infinite.
- What would happen if $\mathbb{P}_X(\{a\}) > 0$ for each $a \in X$?

Single Random Variables: Continuous Random Variables

If X is a continuous random variable we cannot use a probability mass function to describe its distribution.

- Recall that if X is a continuous random variable then the outcome space $\mathcal{O}_X$ is (uncountably) infinite.
- What would happen if $\mathbb{P}_X(\{a\}) > 0$ for each $a \in X$?
  - Then for any set A

$$\mathbb{P}_X(A) = \sum_{a \in A} \mathbb{P}_X(\{a\}) = \infty.$$

  - So we must have $\mathbb{P}_X(\{a\}) = 0$ for all $a \in \mathcal{O}_X$.

Single Random Variables: Continuous Random Variables

If X is a continuous random variable we cannot use a probability mass function to describe its distribution.

- Recall that if X is a continuous random variable then the outcome space $\mathcal{O}_X$ is (uncountably) infinite.
- What would happen if $\mathbb{P}_X(\{a\}) > 0$ for each $a \in X$?
  - Then for any set A

$$\mathbb{P}_X(A) = \sum_{a \in A} \mathbb{P}_X(\{a\}) = \infty.$$

  - So we must have $\mathbb{P}_X(\{a\}) = 0$ for all $a \in \mathcal{O}_X$.
- Intuitively, what is the probability that an Olympic sprinters runs the 100m dash in exactly 9.8412312. . . seconds?
  - Basically zero.

Single Random Variables: Continuous Random Variables

If X is a continuous random variable we cannot use a probability mass function to describe its distribution.

- Recall that if X is a continuous random variable then the outcome space $\mathcal{O}_X$ is (uncountably) infinite.
- What would happen if $\mathbb{P}_X(\{a\}) > 0$ for each $a \in X$?
    - Then for any set A

$$\mathbb{P}_X(A) = \sum_{a \in A} \mathbb{P}_X(\{a\}) = \infty.$$

    - So we must have $\mathbb{P}_X(\{a\}) = 0$ for all $a \in \mathcal{O}_X$.
- Intuitively, what is the probability that an Olympic sprinters runs the 100m dash in exactly 9.8412312... seconds?
    - Basically zero.

This rules out being able to use a pmf to describe the distribution of a continuous random variable.

If X is a continuous random variable we cannot use a probability mass function to describe its distribution.

Single Random Variables: Continuous Random Variables.

If X is a continuous random variable we cannot use a probability mass function to describe its distribution.

Instead we will use the probability density function (pdf), $f_X(\cdot)$ to describe the distribution of X. The pdf is related to the probability measure $\mathbb{P}_X$ via the following equation:

$$\mathbb{P}_X(a \leq X \leq b) = \int_a^b f_X(x)\, dx.$$

If X is a continuous random variable we cannot use a probability mass function to describe its distribution.

Instead we will use the probability density function (pdf), $f_X(\cdot)$ to describe the distribution of X. The pdf is related to the probability measure $\mathbb{P}_X$ via the following equation:

$$\mathbb{P}_X(a \leq X \leq b) = \int_a^b f_X(x)\, dx.$$

The identity above as well as the rules for the probability measure $\mathbb{P}_X$ can be used to calculate $\mathbb{P}_X(A)$ for any set $A \subseteq \mathcal{O}_X$.