Econometrics 120A: Discussion Section

Week 3

Natalia Madrid & Lapo Bini

Department of Economics

Some Stuff

Some stuff

About the office hours

Natalia Madrid

- ⇨ Office hours: Tuesdays, **4:30 pm to 5:30 pm**, SDSC E129
- ⇨ E-mail: namadrid@ucsd.edu
- ⇨ Website: ▸ Natalia Madrid

Lapo Bini

- ⇨ Office hours: Monday, 12:30 pm to 1:30 pm
- ⇨ E-mail: lbini@ucsd.edu
- ⇨ Website: ▸ Lapo Bini

Chapter 3: Probability distribution

Random Variable

A **random variable** is a function that maps outcomes from a **sample space** (the set of all possible outcomes of a random experiment) to real numbers. Formally, a random variable is a function:

$$X : \Omega \to \mathbb{R}$$

where:

- ⇨ Ω is the **sample space** (the set of all possible outcomes of a random process),
- ⇨ $X(\omega)$ is the value assigned by the random variable to an outcome $\omega \in \Omega$,
- ⇨ $\mathbb{R}$ is the set of real numbers.

Random Variable: example

If you flip once a coin, the sample space is $\Omega = \{\text{Heads}, \text{Tails}\}$.

A **random variable** X could be defined as:

- ⇨ X(Heads) = 1
- ⇨ X(Tails) = 0

Here, **X maps the outcomes of the coin flip to numbers.**

The probability of getting heads (H) or tails (T) is:

$$P(H) = P(T) = \frac{1}{2}$$

Random Variables and Probability Distribution

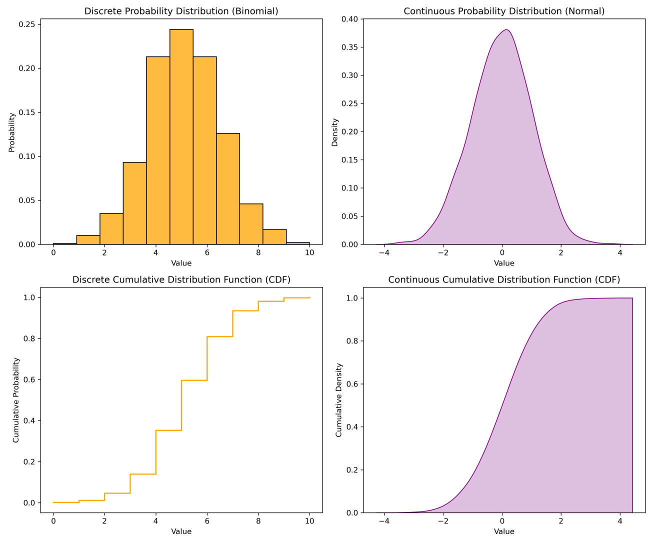
Probability distribution

A **probability distribution** describes how probabilities are assigned to possible outcomes of a random variable.

- ⇨ **Discrete Probability Distribution**: Deals with countable outcomes, like rolling a dice. Each outcome has a specific probability.

- ⇨ **Continuous Probability Distribution**: Deals with infinite possible outcomes within a range, like measuring height. Probabilities are assigned over intervals.

Probability distribution



Discrete Probability Distribution

- **Probability Mass Function (PMF):** Defines the probability that a discrete random variable X equals a specific value.

$$P(X = x) = p(x)$$

The sum of probabilities for all possible values must equal 1:

$$\sum_{x} P(X = x) = 1$$

- **Cumulative Distribution Function (CDF):** Gives the probability that a discrete random variable is less than or equal to a certain value:

$$F(x) = P(X \leq x) = \sum_{x' \leq x} P(X = x')$$

Discrete Probability Distribution

Discrete Distribution

⇨ **Mean:**

$$E[X] = \sum_x x \cdot P(X = x)$$

⇨ **Variance:**

$$\text{Var}(X) = E[(X - E[X])^2] = \sum_x (x - E[X])^2 \cdot P(X = x)$$

⇨ **Standard deviation:**

$$\sigma = \sqrt{\text{Var}(X)}$$

⇨ **Examples:** Bernoulli, Binomial, and Poisson distributions.

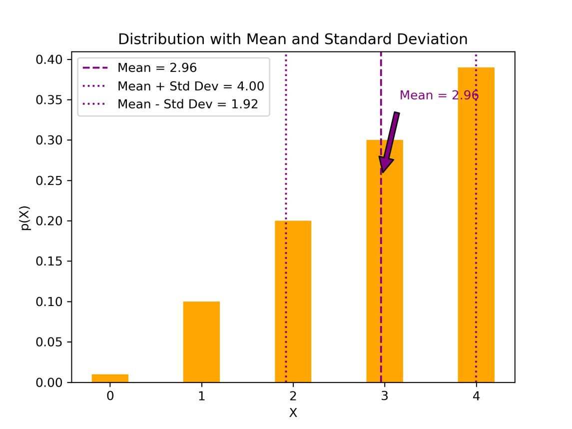
Kind of PS2 Question 2

Consider the following distribution with **mean 2.96** and **standard deviation 1.04**. Draw the graph and show the mean is the balancing point and the standard deviation is a typical deviation.

X	0	1	2	3	4
p(x)	0.01	0.10	0.20	0.30	0.39

Kind of PS2 Question 2

Consider the following distribution with **mean 2.96** and **standard deviation 1.04**. Draw the graph and show the mean is the balancing point and the standard deviation is a typical deviation.



Kind of PS2 Question 3

The students in Econometrics 120A have a bet on whether Natalia will win the award for best TA. She can either win the award, not win it, or share it with Lapo. If she doesn't win, she won't receive any prize, but if she wins alone, she'll receive $\$5000$, and if she shares the award, she'll receive $\$3000$. The probability that she won't win is 20%, that she wins alone is 50%, and that she shares the prize is 30%.

Define a random variable that measures the prize and its distribution.

Random variables and distribution questions

Kind of PS2 Question 3

The students in Econometrics 120A have a bet on whether Natalia will win the award for best TA. She can either win the award, not win it, or share it with Lapo. If she doesn't win, she won't receive any prize, but if she wins alone, she'll receive $5000, and if she shares the award, she'll receive $3000. The probability that she won't win is 20%, that she wins alone is 50%, and that she shares the prize is 30%.

Define a random variable that measures the prize and its distribution.

When we define a **random variable**, we have to include:

- ⇨ the description of the outcomes
- ⇨ the collection of possible outcomes
- ⇨ the probabilities we give to the outcomes

Random variables and distribution questions

Kind of PS2 Question 3

The students in Econometrics 120A have a bet on whether Natalia will win the award for best TA. She can either win the award, not win it, or share it with Lapo. If she doesn't win, she won't receive any prize, but if she wins alone, she'll receive $\$5000$, and if she shares the award, she'll receive $\$3000$. The probability that she won't win is 20%, that she wins alone is 50%, and that she shares the prize is 30%.

Define a random variable that measures the prize and its distribution.

Discrete Probability Distributions

Bernoulli Distribution

The Bernoulli distribution is a discrete probability distribution for a random variable which takes the value 1 with probability p (success) and the value 0 with probability $1 - p$ (failure).

⇨ **Parameter:** p (probability of success, $0 \leq p \leq 1$)

⇨ **Probability Mass Function (PMF):**

$$P(X = x) = p^x(1 - p)^{1-x} \quad \text{for } x \in \{0, 1\}$$

⇨ **Mean (Expected Value):**

$$E[X] = p$$

⇨ **Variance:**

$$\text{Var}(X) = p(1 - p)$$

Binomial Distribution

The Binomial distribution models the number of successes in a fixed number of independent Bernoulli trials, each with the same probability of success.

- **Parameters:** n (number of trials, fixed) and p (probability of success)
- **Probability Mass Function (PMF):**

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k} \quad \text{for } k=0,1,2,\dots,n$$

  The binomial coefficient is defined as:

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

- **Mean (Expected Value):**

$$E[X]=np$$

- **Variance:**

$$\mathrm{Var}(X)=np(1-p)$$

Kind of PS2 Question 4

At RIMAC Fitness Gym, the cardio machines are available for use in cycles where each machine is in use for **20 minutes** followed by a **10 minute** break for maintenance. Students often wait to use the machines (which is annoying), especially during peak hours.

(a) What is the chance that a student will arrive at RIMAC Fitness Gym and find a particular cardio machine currently in use?

(b) Suppose that a student visits RIMAC Fitness Gym 8 times during peak hours. What is the expected amount of time the member is likely to spend waiting to use the cardio machines during these visits?

Kind of PS2 Question 4

At RIMAC Fitness Gym, the cardio machines are available for use in cycles where each machine is in use for **20 minutes** followed by a **10 minute** break for maintenance. Students often wait to use the machines (which is annoying), especially during peak hours.

(a) What is the chance that a student will arrive at RIMAC Fitness Gym and find a particular cardio machine currently in use?

Kind of PS2 Question 4

At RIMAC Fitness Gym, the cardio machines are available for use in cycles where each machine is in use for **20 minutes** followed by a **10 minute** break for maintenance. Students often wait to use the machines (which is annoying), especially during peak hours.

(b) Suppose that a student visits RIMAC Fitness Gym 8 times during peak hours. What is the expected amount of time the member is likely to spend waiting to use the cardio machines during these visits?

Kind of PS2 Question 5

The UCSD Economics program has 40 spots available for incoming students. Historically, 65% of applicants are American students, while 35% are international students.

(a) Calculate the probability that 10 or fewer international students are accepted into the Economics program, assuming each applicant has an equal chance of being accepted.

(b) What is the probability that exactly half of the accepted students in the program are international?

Binomial Questions

Kind of PS2 Question 5

The UCSD Economics program has 40 spots available for incoming students. Historically, 65% of applicants are American students, while 35% are international students.

(a) Calculate the probability that 10 or fewer international students are accepted into the Economics program, assuming each applicant has an equal chance of being accepted.

Cumulative Binomial distribution table

n=40

	p													
x	**0.05**	**0.1**	**0.15**	**0.2**	**0.25**	**0.3**	**0.35**	**0.4**	**0.45**	**0.5**	**0.55**	**0.6**	**0.65**	**0.7**
0	0.129	0.015	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1	0.399	0.080	0.012	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	0.677	0.223	0.049	0.008	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3	0.862	0.423	0.130	0.028	0.005	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
4	0.952	0.629	0.263	0.076	0.016	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
5	0.986	0.794	0.433	0.161	0.043	0.009	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
6	0.997	0.900	0.607	0.286	0.096	0.024	0.004	0.001	0.000	0.000	0.000	0.000	0.000	0.000
7	0.999	0.958	0.756	0.437	0.182	0.055	0.012	0.002	0.000	0.000	0.000	0.000	0.000	0.000
8	1.000	0.985	0.865	0.593	0.300	0.111	0.030	0.006	0.001	0.000	0.000	0.000	0.000	0.000
9	1.000	0.995	0.933	0.732	0.440	0.196	0.064	0.016	0.003	0.000	0.000	0.000	0.000	0.000
10	1.000	0.999	0.970	0.839	0.584	0.309	0.121	0.035	0.007	0.001	0.000	0.000	0.000	0.000

Binomial Questions

Kind of PS2 Question 5

The UCSD Economics program has 40 spots available for incoming students. Historically, 65% of applicants are American students, while 35% are international students.

(b) What is the probability that exactly half of the accepted students in the program are international?

Continuous Distribution

A continuous distribution describes the probabilities of possible values of a continuous random variable. Values can take any real number within a given range.

⇨ **Probability Density Function (PDF)**: Represents the likelihood of a random variable taking on a particular value.

Formula:

$$P(a \leq X \leq b) = \int_a^b p(x)\, dx$$

⇨ **Cumulative Distribution Function (CDF)**: Represents the probability that a random variable is less than or equal to a certain value.

Formula:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} p(t)\, dt$$

Continuous Distribution

Mean (Expected Value):

$$\mu = E[X] = \int_{-\infty}^{\infty} xp(x)\,dx$$

Variance:

$$\sigma^2 = Var(X) = E[(X-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 p(x)\,dx$$

Standard Deviation:

$$\sigma = \sqrt{Var(X)}$$

Kind of PS2 Question 7

Suppose we have the distribution

$$p(x) = \begin{cases} \frac{1}{8} & 0 \leq x \leq 8 \\ 0 & \text{otherwise} \end{cases}$$

(a) Draw the distribution. Where does its mean (balancing point) seem to be?

(b) Show that the probabilities 'add up' to 1 (i.e., the area under the curve is equal to one).

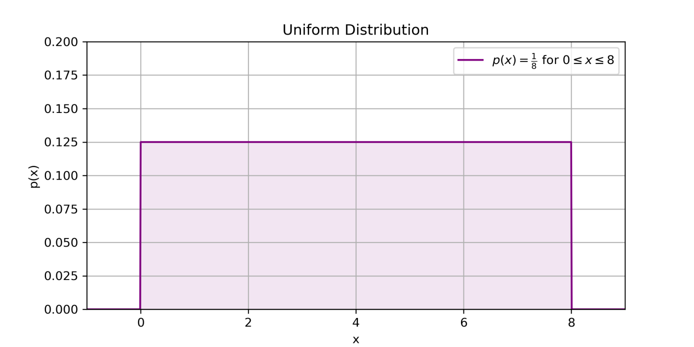
(c) Derive $E[X]$ using the formula from class. Does it agree with your answer in (a)?

Kind of PS2 Question 7

Suppose we have the distribution

$$p(x) = \begin{cases} \frac{1}{8} & 0 \leq x \leq 8 \\ 0 & \text{otherwise} \end{cases}$$

(a) Draw the distribution. Where does its mean (balancing point) seem to be?



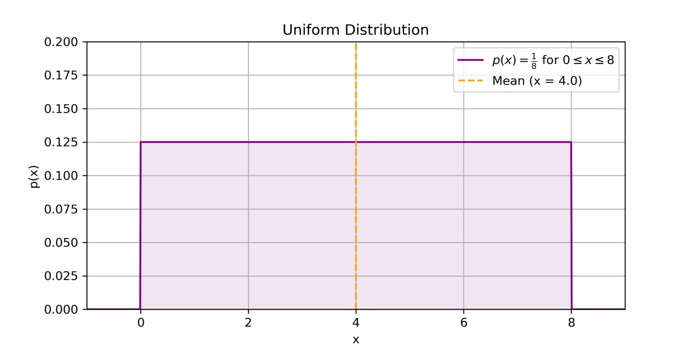
Continuous Distribution questions

Kind of PS2 Question 7

Suppose we have the distribution

$$p(x) = \begin{cases} \frac{1}{8} & 0 \leq x \leq 8 \\ 0 & \text{otherwise} \end{cases}$$

(a) Draw the distribution. Where does its mean (balancing point) seem to be?



Continuous Distribution questions

Kind of PS2 Question 7

Suppose we have the distribution

$$p(x) = \begin{cases} \frac{1}{8} & 0 \leq x \leq 8 \\ 0 & \text{otherwise} \end{cases}$$

(b) Show that the probabilities 'add up' to 1 (i.e., the area under the curve is equal to one).

Continuous Distribution questions

Kind of PS2 Question 7

Suppose we have the distribution

$$p(x) = \begin{cases} \frac{1}{8} & 0 \leq x \leq 8 \\ 0 & \text{otherwise} \end{cases}$$

(c) Derive $E[X]$ using the formula from class. Does it agree with your answer in (a)?