STAT 515 Lec 04 slides

Random variables

Karl Gregory

University of South Carolina

$\mathcal{X} = \{1, 2, \ldots, 6\}$ $\qquad$ $\mathcal{X} = [0, \infty)$

These slides are an instructional aid; their sole purpose is to display, during the lecture, definitions, plots, results, etc. which take too much time to write by hand on the blackboard. They are not intended to explain or expound on any material.

A random variable is a numeric encoding of the outcome of an experiment.

Random variable

A *random variable* is a function from a sample space S to the real numbers.

That is, a *random variable* X is a function $X\colon S \to \mathbb{R}$.

"real numbers"

sample space.

Denote by $\mathcal{X}$ the range of X, the set of values X may take.

We often call $\mathcal{X}$ the *support* of X.

X ← random variable

$\mathcal{X}$ ← support of X

Examples:

1. Flip a coin and let $X = 1$ if heads, $X = 0$ otherwise.
2. Flip a coin three times and let $X =$ the number of heads.
3. Count jellyfish washed up on the beach. Let $X = \#$ jellyfish.
4. Let $X =$ time until you drop your new phone. $\mathcal{X} = [0, \infty)$ ← continuous.
5. Let $X =$ number on up-face of rolled die. $S = \{⚀, ⚁, \dots, ⚅\}$

$\mathcal{X} = \{1, \dots, 6\}$. Discrete

(1) $S = \{\text{heads}, \text{tails}\}$

$$X(\text{outcome}) = \begin{cases} 1 & \text{if outcome is heads} \\ 0 & \text{if outcome is tails} \end{cases}$$

$$X = \begin{cases} 1 & \text{if heads} \\ 0 & \text{if tails} \end{cases}$$

support: $\mathcal{X} = \{0, 1\}$ Discrete

2. Flip a coin three times and let $X =$ the number of heads.

$S = \{$ HHH, HHT, HTH, THH, TTH, THT, HTT, TTT $\}$

$\mathcal{X} = \{0, 1, 2, 3\}$ Discrete

$$X = \begin{cases} 0 & \text{if outcome is TTT} \\ 1 & \text{if outcome is TTH, THT, HTT} \\ 2 & \text{if outcome is HHT, HTH, THH} \\ 3 & \text{if outcome is HHH} \end{cases}$$

3. Count jellyfish washed up on the beach. Let $X = \#$ jellyfish.

$S = \{0, 1, 2, \ldots\}$

$X =$ outcome

same

$\mathcal{X} = \{0, 1, 2, \ldots\}$ Discrete

Discrete and continuous random variables

- *Discrete*: Support $\mathcal{X}$ is a list of numbers (finite or countably infinite)
- *Continuous*: Support $\mathcal{X}$ is an interval (or union of intervals).

But what about *categorical data?*

- Record eye color of randomly selected student.
- Rate professor as *miserable*, *mediocre*, *middling*, or *magnificent*.

These we can encode numerically into rvs; rvs are always numbers.

Discuss nominal/ordinal.

$P(A)$

↑

Exercise: Consider some events involving random variables:

1. Flip a coin and let $X = 1$ if heads, $X = 0$ otherwise. Find $P(X = 1)$?
2. Flip a coin three times and let $X = \#$ heads. Find $P(X = 0)$?
3. Let $X = \#$ jellyfish washed up on the beach. Find $P(X > 10)$?
4. Let $X =$ time until you drop your new phone. Find $P(X \leq 1)$?
5. Let $X =$ number on up-face of rolled die. Find $P(X \in \{3, 4\})$?

↑ "in" " $\frac{2}{6} = \frac{1}{3}$.

(1) $P(X=1) = \frac{1}{2}$

(2) $P(X=0) = \frac{1}{8}$

The probability distribution of a random variable tells

1. what values it can take $\mathcal{X}$
2. with what probabilities

rv: random variable

possible values X can take

Probability distribution of a discrete random variable

The *probability distribution* of a discrete rv X with support $\mathcal{X} = \{x_1, x_2, x_3, \dots\}$ is an assignment of probabilities $p_1, p_2, p_3, \dots$ to the values $x_1, x_2, x_3, \dots$ such that

- $p_i \in [0, 1]$ for $i = 1, 2, \dots$ (in)
- $\sum_i p_i = 1$ (Probabilities sum to 1).

Exercise: Tabulate probability distributions of the following discrete rvs:

1. Roll a die and let X = number on up-face of die.
2. Flip two coins and let $X = \#$ heads.

②

specific value

$\mathcal{X}$

x	1	2	3	4	5	6
$P(X=x)$	1/6	1/6	1/6	1/6	1/6	1/6

$P(X=1)$ $P(X=2)$

As-yet unobserved random variable.

little x

The table lists $P(X=x)$ for all $x \in \mathcal{X}$.

Big X Curly $\mathcal{X}$

② Flip Two coins

$$S = \{HH \quad HT \quad TH \quad TT\}$$

$$X = \begin{cases} 0 & \text{if } TT \\ 1 & \text{if } HT \quad TH \\ 2 & \text{if } HH \end{cases}$$

$$\mathcal{X} = \{0, 1, 2\}$$

x	0	1	2
$P(X=x)$	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$

Exercise: Let X = sum of two rolls of a die.

1. Tabulate the probability distribution of X.
2. Give $P(X \leq 7)$.
3. Give $P(X > 10)$.

$$S = \left\{ \begin{matrix} (1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\ (5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\ (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) \end{matrix} \right\}.$$

$x = \{2, \ldots, 12\}$

x	2	3	4	5	6	7	8	9	10	11	12
$P(X=x)$	$\frac{1}{36}$	$\frac{2}{36}$	$3/36$	$4/36$	$5/36$	$6/36$	$5/36$	$4/36$	$3/36$	$2/36$	$\frac{1}{36}$

② $P(X \leq 7) = \frac{21}{36}$

③ $P(X > 10) = \frac{3}{36}$

$\mathcal{X} = \{0, 1, 2, \ldots\}$

When $\mathcal{X}$ is countably infinite, we cannot write down the entire table:

Exercise: If $X = \#$ jellyfish washed up on the beach, we might have

x	0	1	2	3	4	5	$P(X>5)$
$P(X = x)$	0.050	0.149	0.224	0.224	0.168	0.101	

1 minus the sum of the other numbers.

How can we find $P(X > 5)$?

Of interest later on: These are Poisson probabilities with $\lambda = 3$.

Expected value (mean/average value)

Expected value of a discrete rv

For X a discrete rv which takes the values $x_1, x_2, x_3, \ldots$ with the probabilities $p_1, p_2, p_3, \ldots$, the *expected value* of X is given by

$\mathbb{E}X$

$$\mathbb{E}X = p_1 x_1 + p_2 x_2 + p_3 x_3 + \ldots$$

Expected value, call $\mathbb{E}$ the expectation operator.

- The average of many realizations of X should be close to $\mathbb{E}X$.
- $\mathbb{E}X$ is the "balancing point" of probability distribution.
- We often use μ to denote $\mathbb{E}X$.
- We often call $\mathbb{E}X$ the *mean* of X

μ = "mu"

$$X = \begin{cases} 1 & \text{heads} \\ 0 & \text{tails} \end{cases}$$

x	0	1
$P(X=x)$	1/2	1/2

Exercise: Flip a (fair) coin and let $X = 1$ if heads, $X = 0$ otherwise.

1. Find $\mathbb{E}X$.
2. Discuss...

$$\mathbb{E}X = \frac{1}{2} \cdot 0 + \frac{1}{2} \cdot 1 = \frac{1}{2}$$

Repeat experiment, observe many realizations of X:

1 1 0 1 0 0 1 0 0 1 1 0 1 1 0 1 0 1 1 1 0 0 1 0 0 ....

$$\text{average} = \frac{\#\{1\text{'s (heads)}\}}{\#\text{ times I flipped coin}}$$ should be close to 1/2.

Exercise: Let X = money won from playing this game:

Roll a die and draw one bill from a bag...

Roll ⚀, ⚁, ⚂	$\longrightarrow$	draw from bag 1:	3 \$5 bills and 1 \$20 bill
Roll ⚃, ⚄	$\longrightarrow$	draw from bag 2:	2 \$5 bills and 2 \$20 bills
Roll ⚅	$\longrightarrow$	draw from bag 3:	1 \$5 bill and 3 \$20 bills

1. Give $\mathcal{X}$. $\mathcal{X} = \{5, 20\}$
2. Tabulate the probability distribution of X.
3. Give $\mathbb{E}X$.
4. If the game costs 7 dollars to play, do you recommend playing it?

(2)

x	5	20
$P(X=x)$	$\frac{7}{12}$	$5/12$

(3)

$$\mathbb{E}X = \frac{7}{12} \cdot 5 + \frac{5}{12} \cdot 20$$

$$= \frac{35 + 100}{12}$$

$$= \frac{135}{12}$$

$$= 11 \frac{3}{12}$$

Exercise: Consider a 10-sided die with sides displaying 1, 2, 3, and 4 as:

side of die	1	2	3	4	5	6	7	8	9	10
number displayed	1	1	1	1	2	2	2	3	3	4

Let $X =$ the number on the up-face of the die when it is rolled.

1. Tabulate the probability distribution of X.
2. Add to the table the *cumulative probabilities* $P(X \leq x)$ for all $x \in \mathcal{X}$.
3. Find $P(X > 3)$.
4. Find $\mathbb{E}X$.

(1)

$\mathcal{X} = \{1, 2, 3, 4\}$

(2)

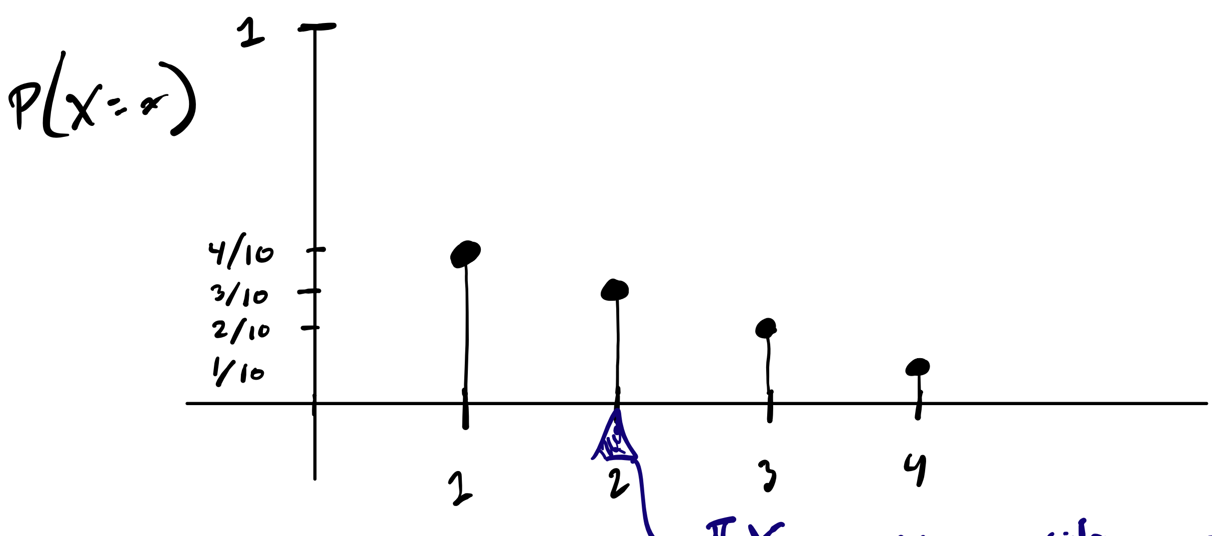
x	1	2	3	4
$P(X=x)$	$\frac{4}{10}$	$\frac{3}{10}$	$\frac{2}{10}$	$\frac{1}{10}$
$P(X\leq x)$	$\frac{4}{10}$	$\frac{7}{10}$	$\frac{9}{10}$	1

(3) $P(X>3) = \frac{1}{10}$

$$P(X>3) = 1 - P(X\leq 3)$$
$$= 1 - \frac{9}{10}$$
$$= \frac{1}{10}$$

(4) $\mathbb{E}X = \frac{4}{10}\cdot 1 + \frac{3}{10}\cdot 2 + \frac{2}{10}\cdot 3 + \frac{1}{10}\cdot 4$

$$= \frac{4+6+6+4}{10}$$
$$= 2.$$



$\mathbb{E}X$ gives position of fulcrum for balancing the prob. dist.

"expected value"

Variance of a random variable

The *variance* of a random variable X with mean μ is defined as

$$\mathrm{Var}\,X = \mathbb{E}(X-\mu)^2.$$

- $\mathrm{Var}\,X$ is the expected squared deviation of X from μ.
- Measure of "spread" for the distribution of X.
- Often use σ^2 to denote $\mathrm{Var}\,X$.
- Use σ to denote $\sqrt{\mathrm{Var}\,X}$, which is called the *standard deviation* of X.

σ : "sigma"

Σ

Variance for discrete rvs

If X has mean μ and takes the values $x_1, x_2, x_3, \ldots$ w/probs $p_1, p_2, p_3, \ldots$, then

$$\text{Var}\, X = p_1(x_1 - \mu)^2 + p_2(x_2 - \mu)^2 + p_3(x_3 - \mu)^2 + \ldots$$

Exercise: Get the variance of the following random variables

1. Let $X = 1$ if coin flip "heads", $X = 0$ if "tails."
2. Let $X =$ number on the up-face of a 6-sided die when it is rolled.

(1)

x	0	1
$P(X=x)$	$1/2$	$1/2$

$$\mu = \mathbb{E}X = \tfrac{1}{2}\cdot 0 + \tfrac{1}{2}\cdot 1 = \tfrac{1}{2}$$

$$\sigma^2 = \text{Var } X = \tfrac{1}{2}(0-\mu)^2 + \tfrac{1}{2}(1-\mu)^2 \qquad (\mu = \tfrac{1}{2})$$

$$= \tfrac{1}{2}\left(0 - \tfrac{1}{2}\right)^2 + \tfrac{1}{2}\left(1 - \tfrac{1}{2}\right)^2$$

$$= \tfrac{1}{2}\,\tfrac{1}{4} + \tfrac{1}{2}\,\tfrac{1}{4}$$

$$= \tfrac{1}{4}$$

Standard deviation: $\sigma = \sqrt{\text{Var } X} = \sqrt{\tfrac{1}{4}} = \tfrac{1}{2}$.

2. Let X = number on the up-face of a 6-sided die when it is rolled.

x	1	2	3	4	5	6
$P(X=x)$	1/6	1/6	1/6	1/6	1/6	1/6

$$\mu = \mathbb{E}X = \frac{1+2+3+4+5+6}{6} = \frac{21}{6} = 3\tfrac{3}{6} = 3.5$$

$$\sigma^2 = \text{Var } X = \tfrac{1}{6}(1-3.5)^2 + \tfrac{1}{6}(2-3.5)^2 + \cdots + \tfrac{1}{6}(6-3.5)^2$$

$$= \ldots$$