Lecture 4

David Armstrong

STATS 67 - UCI

Discrete Random Variables

- *Variable*: A quantity that may take different values.

- *Random variable*: A variable that may assume different values with certain probabilities.

  - One way to think of it as a function that assigns a real number to each outcome in the sample space.

  - A *discrete* random variable is one who can only be discrete values (integers).

  - A discrete random variable has countably many outcomes.

  - For now, we will focus on bounded discrete random variables.

- Examples of a discrete random variable:

  - Flip a coin 3 times.
    Then $S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$.

  - Let X be a random variable that is equal to the number of heads in 3 flips of a coin.
  - X can be 0,1,2, or 3.

  Another example is the number of days you do an activity each week.

  - X can be 0,1,2,3,4,5,6 or 7.

Discrete Random Variables

- Denote the *support* of X as $\mathbb{S}_X$.
- The *support* of X is the space of values which X has a positive probability of occurring.
  - Notationally, $X : S \to \mathbb{S}_X$
- An example is flipping a coin once.
- $S = \{T, H\}$.
- If we let X be equal to the number of heads (this is the same as setting heads to equal 1 and tails to equal 0).
- Then $\mathbb{S}_X = \{0, 1\}$.
- In the example on the previous slide with 3 coin flips, $\mathbb{S}_X = \{0, 1, 2, 3\}$.
- In the example on the previous slide with days per week an activity is done, $\mathbb{S}_X = \{0, 1, 2, 3, 4, 5, 6, 7\}$.

Continuous Random Variables

- If $\mathbb{S}_X$ is continuous, then the random variable X is continuous.
  - X can be any number in a given range.
  - There are uncountably many possible values of X.
  - Examples include the the height of a person, the blood pressure of a subject, or the distance a vehicle travels.

- The concept of the support of a random variable is an important one.

- Once the appropriate random variable is specified, we can focus only on the support of it as opposed to the entire sample space.

- Example: If we flip a coin 100 times and interested in the number of heads seen.
  - What does X represent?
  - How many elements are in the sample space?
  - What is the support of the random variable X?

Discrete Random Variables

- Assume we flip a coin 3 times and let X be the number of heads.
  - What does $X = 0$ represent?
  - What is the probability all three of the flips land in tails?
  - What is the probability that $X = 1$
  - What is the probability that $X = 2$
  - What is the probability that $X = 3$
  - Is this a valid distribution?

Discrete Random Variables: Distribution

The *probability distribution* of X assigns a number to all values x in $\mathbb{S}_X$ such that:

- $0 \leq P(X = x) \leq 1$

- $\sum\limits_{x \in S_x} P(X = x) = 1$

Notationally we state $f(x) = P(X = x)$. With discrete random variables, $f(x)$ is termed the *probability mass function (p.m.f.)*.

- From here on out, we will refer to X as the random variable.

- We will denote x as the values that X can be.

- For example flipping a coin 3 times, and setting X to be the number of heads.
  - X is the random variable.
  - X can be set equal to x where $x = 0, 1, 2$, or 3.

- For example the number of days a week an activity is done, and setting X to be the the number of days.
  - X is the random variable. It can equal to x where $x = 0, 1, 2, 3, 4, 5, 6$, or 7.

Discrete Random Variables: Distribution

Returning to the example of flipping a coin 3 times. A distribution table for X, the number of heads, can be constructed below.

	$x = 0$	$x = 1$	$x = 2$	$x = 3$	Total
$P(X = x)$					

Note that functions of the random variable X are also random variables.

- For example X^2.
- In the number of heads in 3 coin flips, $\mathbb{S}_X = \{0, 1, 2, 3\}$
- The support of X^2 will be $\mathbb{S}_{X^2} = \{0, 1, 4, 9\}$

Discrete Random Variables: Distribution

- With the p.m.f. of a discrete random variable, we can compute quantities such as $P(X < a)$ or $P(a \leq X < b)$ for some set constants of a and b.

- The *cumulative distribution function* (cdf) of a random variable at value X is $P(X \leq x)$.

- Notationally this is $F(x) = P(X \leq x) = \sum_{\tilde{x} \leq x} P(X = \tilde{x})$

  - It is the sum of all probabilities which have $X \leq x$.

- A useful rule to remember when calculating cdf's is the law of the complement, $P(X \leq a) = 1 - P(X > a)$.

Discrete Random Variables: Distribution

Assume we roll a die once. Let X equal to the number that shown, then we can create the following distribution table.

	$x=0$	$x=1$	$x=2$	$x=3$	$x=4$	$x=5$	$x=6$
$P(X=x)$							

- Calculate $f(3)$

- Calculate $F(3)$.

- What is the probability of rolling a number greater than 1 but less than 5.

- Calculate $F(5)$

Discrete Random Variables: Distribution

Properties of discrete random variables

- Let a be a value of x. $P(X < a) = P(X \leq a - 1)$.

- $P(X > a) = P(X \geq a + 1)$

- Let a and b be constants specified by you. Something that looks counter intuitive, but holds true for discrete distributions.

  - $P(a \leq X \leq b) = P(X \leq b) - P(X < a)$.
  - $P(a < X \leq b) = P(X \leq b) - P(X \leq a)$.

Discrete Random Variables: Distribution

The p.m.f. completely determines the probability distribution of a discrete random variable.

- The *expectation* of X can be viewed as the mean or average of X.
- Within a frequentist framework, it can be seen as the average of X across many trials of the experiment.
- The *expectation* of X is denoted as $\mathrm{E}(X)$.
- $\mathrm{E}(X) = \sum_{x \in \mathbb{S}_X} xP(X = x) = \sum_{x \in \mathbb{S}_X} xf(x)$.
- Can be viewed as averaging over all possible X values while weighting each possible value by its probability.

Discrete Random Variables: Expectation

Returning to the example of flipping a coin 3 times. A distribution table for X, the number of heads, can be constructed as follows.

	$x = 0$	$x = 1$	$x = 2$	$x = 3$
$P(X = x)$	$\left(\frac{1}{2}\right)^3$	$3\left(\frac{1}{2}\right)^3$	$3\left(\frac{1}{2}\right)^3$	$\left(\frac{1}{2}\right)^3$

What is the expectation of X?

- $\mathrm{E}(X) = \sum_{x \in \mathbb{S}_X} x f(x)$

- The expectation of X can be viewed as the average number of heads seen when we repeatedly (a large number of times) flip a coin 3 times.

Discrete Random Variables: Expectation

- Remember that functions of X are also random variables.
- We can set $h(X)$ and $g(X)$to be a function of X.
  - Example: $h(X) = X^2$.
  - Example: $g(X) = \frac{1}{X}$.
- We can take expectation of these functions without having to first find the distribution of $h(X)$ first.
  - $\mathrm{E}(h(X)) = \sum_{x \in \mathbb{S}_X} h(x)P(X = x) = \sum_{x \in \mathbb{S}_X} h(x)f(x)$.
- Just like with X, $\mathrm{E}(h(X))$ can be viewed as averaging over all possible $h(X)$ values while weighting each possible value by the probability of X.

Properties of expectations.

- If a and b are constants and X is a random variables, then $\mathrm{E}(a + bX) = a + b\mathrm{E}(X)$.
- If X and Y are random variables, then $\mathrm{E}(X + Y) =\mathrm{E}(X)+\mathrm{E}(Y)$.
  - As a result if a and b are constants, then $\mathrm{E}(aX + bY) =a\mathrm{E}(X)+b\mathrm{E}(Y)$
  - As a further result, let X_i be random variables and a_i's be constants.
  - $\mathrm{E}\left(\sum_i a_i X_i\right) = \sum_i a_i \mathrm{E}(X_i)$

Discrete Random Variables: Expectation

Example: Say X is the number of days a student is registered to take classes at University of California.

x	1	2	3	4	5
$f(x)$	0.10	0.30	0.25	0.25	0.10

- What is the expectation of X?

- What is the expectation of $5 + 3X$?

- Let $h(X) = X^2$. What is the expectation of $h(x)$?

- let $g(X) = \frac{1}{X}$. What is the expectation of $g(X)$