CS145: Lecture 5 Outline

- Discrete random variables
- Expectations of discrete variables
Discrete Random Variables

A random variable assigns values to outcomes of uncertain experiments.

\[ X : \Omega \rightarrow \mathbb{R} \quad \text{where} \quad x = X(\omega) \in \mathbb{R} \quad \text{for} \quad \omega \in \Omega \]

Mathematically: A function from sample space \( \Omega \) to real numbers \( \mathbb{R} \).

May define several random variables on the same sample space, if there are several quantities you would like to measure.

Examples: power consumed, temperature measured, gain/loss of money, number of requests served, etc.
A random variable assigns values to outcomes of uncertain experiments:

\[ X : \Omega \rightarrow \mathbb{R} \quad x = X(\omega) \in \mathbb{R} \text{ for } \omega \in \Omega \]

The range of a random variable is the set of values with positive probability:

\[ X = \{ x \in \mathbb{R} \mid X(\omega) = x \text{ for some } \omega \in \Omega, P(\omega) > 0 \} \]

For a discrete random variable, the range is finite or countably infinite (we can map it to the integers). Coming later: continuous random variables.
A random variable assigns values to outcomes of uncertain experiments

\[ X : \Omega \rightarrow \mathbb{R} \quad \text{for } x = X(\omega) \in \mathbb{R} \quad \text{for } \omega \in \Omega \]

The probability mass function (PMF) or probability distribution of variable:

\[ p_X(x) = P(X = x) = P(\{\omega \in \Omega \mid X(\omega) = x\}) \]

\[ p_X(x) \geq 0, \quad \sum_{x \in \mathcal{X}} p_X(x) = 1. \]

If range is finite, this is a vector of non-negative numbers that sums to one.
Computing a PMF

• Notation:

\[ p_X(x) = P(X = x) = P(\{\omega \in \Omega \text{ s.t. } X(\omega) = x\}) \]

\[ p_X(x) \geq 0 \quad \sum_x p_X(x) = 1 \]

• Example: Two independent rolls of a fair tetrahedral die

\( F: \) outcome of first throw
\( S: \) outcome of second throw
\( X = \min(F, S) \)

4
3
2
1

1 2 3 4

F = First roll
S = Second roll
Computing probabilities of sets of values:

\[ P(X \in S) = \sum_{x \in S} p_X(x) \text{ for any } S \subset \mathbb{R}. \]

The probability mass function or probability distribution of random variable:

\[ p_X(x) = P(X = x) = P\left(\{\omega \in \Omega \mid X(\omega) = x\}\right) \]

\[ p_X(x) \geq 0, \quad \sum_{x \in \mathcal{X}} p_X(x) = 1. \]

If range is finite, this is a vector of non-negative numbers that sums to one.
Functions of Random Variables

- A random variable assigns values to outcomes of uncertain experiments.

\[ X : \Omega \to \mathbb{R} \quad x = X(\omega) \in \mathbb{R} \quad \text{for} \quad \omega \in \Omega \]

\[ p_X(x) = P(X = x) \]

\[ p_X(x) \geq 0, \quad \sum_{x \in \mathcal{X}} p_X(x) = 1. \]

- If we take any non-random (deterministic) function of a random variable, we produce another random variable:

\[ Y = g(X) \quad g : \mathbb{R} \to \mathbb{R} \]

\[ g \circ X : \Omega \to \mathbb{R} \]

- Example: Degrees Celsius \( X \) to degrees Fahrenheit \( Y \):

\[ Y = 1.8X + 32 \]

- Example: Current drawn \( X \) to power consumed \( Y \):

\[ Y = rX^2 \]
A random variable assigns values to outcomes of uncertain experiments

\[ X : \Omega \rightarrow \mathbb{R} \quad \quad x = X(\omega) \in \mathbb{R} \quad \text{for} \quad \omega \in \Omega \]

\[ p_X(x) = P(X = x) \]

\[ p_X(x) \geq 0, \quad \sum_{x \in \mathcal{X}} p_X(x) = 1. \]

If we take any non-random (deterministic) function of a random variable, we produce another random variable:

\[ Y = g(X) \quad g : \mathbb{R} \rightarrow \mathbb{R} \]

\[ g \circ X : \Omega \rightarrow \mathbb{R} \]

By definition, the probability mass function of \( Y \) equals

\[ p_Y(y) = \sum_{\{x | g(x) = y\}} p_X(x) \quad p_Y(y) \geq 0, \quad \sum_{y \in \mathcal{Y}} p_Y(y) = 1. \]
Example: Absolute Value

\[ g(X) = |X| \]

\[ p_Y(y) = \sum_{x \mid g(x) = y} p_X(x) \]

\[ p_X(x) = \begin{cases} 
1/9 & \text{if } x \text{ is an integer in the range } [-4, 4], \\
0 & \text{otherwise.} 
\end{cases} \]

\[ p_Y(y) = \begin{cases} 
2/9 & \text{if } y = 1, 2, 3, 4, \\
1/9 & \text{if } y = 0, \\
0 & \text{otherwise.} 
\end{cases} \]
Example: Complex Functions

X: current “state” of atmosphere/ocean/hurricane

g: sophisticated climate simulator

Y=g(X): future “state” (in N days) of atmosphere/ocean/hurricane
CS145: Lecture 5 Outline

- Discrete random variables
- Expectations of discrete variables
The expectation or expected value of a discrete random variable is:

\[ E[X] = \sum_{x \in \mathcal{X}} x p_X(x) \]

The expectation is a single number, not a random variable. It encodes the “center of mass” of the probability distribution.

Example: Uniform distribution on \{0, 1, ..., n\}

\[
E[X] = 0 \cdot \frac{1}{n+1} + 1 \cdot \frac{1}{n+1} + \cdots + n \cdot \frac{1}{n+1} = \frac{n(n+1)}{2(n+1)} = \frac{n}{2}
\]
A Bernoulli or indicator random variable $X$ has one parameter $p$:

$$p_X(1) = p, \quad p_X(0) = 1 - p, \quad X = \{0, 1\}$$

For an indicator variable, expected values are probabilities:

$$E[X] = p$$

Examples:

- Flip a possibly biased coin with probability of coming up heads $p$
- A user answers a true/false question in an online survey
- Does it snow or not on some day
Expectation

- The *expectation* or *expected value* of a discrete random variable is:

\[ E[X] = \sum_{x \in \mathcal{X}} xp_X(x) \]

- The expectation is a single number, not a random variable. It encodes the “center of mass” of the probability distribution:

If \( X \) takes two possible values, say \( a \) and \( b \), with probabilities \( P(a) \) and \( P(b) \), then

\[ E(X) = aP(a) + bP(b) \quad P(a) + P(b) = 1 \]
The **expectation** or **expected value** of a discrete random variable $X$ is:

$$E[X] = \sum_{x \in \mathcal{X}} x p_X(x)$$

The expectation is a single number, not a random variable. It encodes the “center of mass” of the probability distribution:

$$x_{\text{min}} \leq E[X] \leq x_{\text{max}}$$

$$x_{\text{min}} = \min\{x \mid x \in \mathcal{X}\}$$

$$x_{\text{max}} = \max\{x \mid x \in \mathcal{X}\}$$

The expectation is an average or interpolation. It is possible that $p_X(E[X]) = 0$ for some random variables $X$. 

---

**Expectation**

- The expectation or expected value of a discrete random variable is:
  $$E[X] = \sum_{x \in \mathcal{X}} x p_X(x)$$
  - The expectation is a single number, not a random variable. It encodes the “center of mass” of the probability distribution:
    $$x_{\text{min}} \leq E[X] \leq x_{\text{max}}$$
    $$x_{\text{min}} = \min\{x \mid x \in \mathcal{X}\}$$
    $$x_{\text{max}} = \max\{x \mid x \in \mathcal{X}\}$$
  - The expectation is an average or interpolation. It is possible that $p_X(E[X]) = 0$ for some random variables $X$. 

A geometric random variable $X$ has parameter $p$, countably infinite range:

$$p_X(k) = (1 - p)^{k-1} p$$

$$X = \{1, 2, 3, \ldots \}$$

Examples:
- Flip a coin with bias $p$, count number of tosses until first heads (success)
- Your laptop hard drive independently fails on each day with (hopefully small) probability $p$. What is the distribution of the number of days until failure?

Recall the geometric series:

$$\sum_{k=0}^{\infty} q^k = \frac{1}{1 - q}, \quad 0 < q < 1.$$
Geometric Probability Distribution

- Recall the geometric series: 
  \[ \sum_{k=0}^{\infty} q^k = \frac{1}{1-q}, \quad 0 < q < 1. \]

- A **geometric** random variable \( X \) has parameter \( p \), countably infinite range:
  \[
  p_X(k) = (1 - p)^{k-1} p \quad \mathcal{X} = \{1, 2, 3, \ldots\}
  \]

- The expected value equals:
  \[
  \mathbb{E}[X] = \sum_{k=1}^{\infty} k(1 - p)^{k-1} p = \frac{1}{p}
  \]

- Prove by taking derivative of series:
  \[
  \frac{d}{dq} \left( \frac{1}{1-q} \right) = \frac{d}{dq} \left( \sum_{k=1}^{\infty} (1 - p)^k \right) = \sum_{k=1}^{\infty} k(1 - p)^{k-1} = \frac{1}{(1-q)^2}
  \]
Consider a non-random (deterministic) function of a random variable:

\[ Y = g(X) \]

\[ p_X(x) = P(X = x) \quad \rightarrow \quad p_Y(y) = \sum_{\{x | g(x) = y\}} p_X(x) \]

What is the expected value of random variable \( Y \)?

\[ E[Y] = E[g(X)] \]

Correct approach #1:

\[ E[Y] = \sum_y y p_Y(y) \]

Correct approach #2:

\[ E[Y] = E[g(X)] = \sum_x g(x) p_X(x) \]

Incorrect approach:

\[ g(E[X]) \neq E[g(X)] \quad \text{(except in special cases)} \]
Example: Absolute Value

\[ g(X) = |X| \]

\[ p_Y(y) = \sum_{\{x \mid g(x) = y\}} p_X(x) \]

\[ p_Y(y) = \begin{cases} 
2/9 & \text{if } y = 1, 2, 3, 4, \\
1/9 & \text{if } y = 0, \\
0 & \text{otherwise.} 
\end{cases} \]

\[ E[X] = 0 \]
\[ g(E[X]) = g(0) = 0 \]
\[ E[Y] = \frac{1}{9}(0) + \frac{2}{9}(1 + 2 + 3 + 4) = \frac{20}{9} \approx 2.22 \]
Travel at a Random Speed

- You want to travel 200 miles to New York
- With 50% probability, the new high-speed train runs at a constant velocity of 200 mph
- With 50% probability, the train engine overheats and it runs at a constant velocity of 1 mph

- time in hours $= T = t(V) =$
- $E[T] = E[t(V)] = \sum_v t(v) p_V(v) =$
- $E[TV] = 200 \neq E[T] \cdot E[V]$
Linearity of Expectation

- Consider a linear function: \( Y = g(X) = aX + b \)

- Example: Change of units (temperature, length, mass, currency, ...)

- In this special case, mean of \( Y \) is the linear function applied to \( E[X] \):

\[
E[Y] = g(E[X]) = aE[X] + b
\]

\[
E[Y] = \sum_x (ax + b)p_X(x) = a \sum_x xp_X(x) + b \sum_x p_X(x) = aE[X] + b.
\]

Example: You went on vacation to Europe, and want to find the average amount you spent on lodging per day. The following are equivalent (assuming a fixed exchange rate from Euros to US dollars):

- \( E[g(X)] = \) convert each receipt from Euros to US dollars, average result

- \( g(E[X]) = \) average receipts in Euros, convert result to US dollars
Consider a **linear function**: \( Y = g(X) = aX + b \)

**Example**: Change of units (temperature, length, mass, currency, …)

In this special case, mean of \( Y \) is the linear function applied to \( E[X] \):

\[
E[Y] = g(E[X]) = aE[X] + b
\]

**Example**: I offer you to let you play a game where you pay a $20 entrance fee, and then I let you roll a fair 6-sided die, and pay you the rolled value times $5. What is your expected change in money?

\[
Y = 5X - 20 \quad \text{(change in money } Y \text{ for dice outcome } X)\
\]

\[
E[X] = 3.5
\]

\[
E[Y] = 5E[X] - 20 = -2.5
\]