ECE 2521: Analysis of Stochastic Processes

Lecture 2

Department of Electrical and Computer Engineering
University of Pittsburgh

September, 8^{th} 2021

Azime

Counting Methods

- Useful for calculating probabilities of events in finite samples spaces with equally likely outcomes
- Consider a process consisting of
  - r stages
  - n_i choices in stage i

  $$\text{Total number of outcomes} = \prod_{i=1}^{r} n_i = n_1 n_2 \cdots n_r$$

- Note: If $n_1 = n_2 = \ldots = n_r = n$, then $\prod_{i=1}^{r} n_i = n^k$

Examples

- Example 1
  - Find the number of license plates with 3 letters and 4 digits.
  - What if repetition is prohibited?

- Example 2
  - What is the total number of subsets of an n-element set including the empty set and itself?

Permutations (Ordering)

- Number of possible ordered sequences of k out of n distinct elements (k permutations of n objects)
- Sampling **without replacement** and **with ordering**

$$P_k^n = n(n-1)(n-2)\cdots(n-(k-1)) = \frac{n!}{(n-k)!}$$

Permutations (Ordering)

- Number of possible ordered sequences of k out of n distinct elements (k permutations of n objects)
- Sampling **without replacement** and **with ordering**

$$P_k^n = n(n-1)(n-2)\cdots(n-(k-1)) = \frac{n!}{(n-k)!}$$

- There are n choices for the first position, $(n-1)$ choices for the second position, ..., and $(n-(k-1))$ choices for the kth position

Permutations (Ordering)

- Number of possible ordered sequences of k out of n distinct elements (k permutations of n objects)
- Sampling **without replacement** and **with ordering**

$$P_k^n = n(n-1)(n-2)\cdots(n-(k-1)) = \frac{n!}{(n-k)!}$$

- There are n choices for the first position, $(n-1)$ choices for the second position, ..., and $(n-(k-1))$ choices for the kth position

Note 1: If $k = n$, then the total of n permutations of the n distinct elements is $P_n^n = n!$ (n *factorial*)

Permutations (Ordering)

- Number of possible ordered sequences of k out of n distinct elements (k permutations of n objects)
- Sampling **without replacement** and **with ordering**

$$P_k^n = n(n-1)(n-2)\cdots(n-(k-1)) = \frac{n!}{(n-k)!}$$

- There are n choices for the first position, $(n-1)$ choices for the second position, ..., and $(n-(k-1))$ choices for the kth position

Note 1: If $k = n$, then the total of n permutations of the n distinct elements is $P_n^n = n!$ (n *factorial*)

Note 2: Correction for indistinguishable elements: If q of the n elements are identical, then the total number of distinguishable permutations of the n elements is $n!/q!$

Examples

a) How many distinct 2 letter permutations can you form from the word *BASE*?

b) How many distinct 2 letter permutations can you form from the word *BABY*?

c) Throw three six-sided dice. Are the outcomes 11 and 12 equally likely?

Combinations (No Ordering)

- Sampling **without replacement** and **without ordering**.
- Let $C_k^n = \binom{n}{k}$ be the number of k-element subsets of a given n-element set (n choose k), then $\binom{n}{k} = \frac{n!}{k!(n-k)!}$.

Combinations (No Ordering)

- Sampling **without replacement** and **without ordering**.
- Let $C_k^n = \binom{n}{k}$ be the number of k-element subsets of a given n-element set (n choose k), then $\binom{n}{k} = \frac{n!}{k!(n-k)!}$.

Proof Consider ways of constructing permutations of k distinct items from an n-element set. This is given by P_k^n. We can also do this by first choosing k items from the n-element set and then ordering them:

$$\binom{n}{k} k! = P_k^n = \frac{n!}{(n-k)!} \quad \Rightarrow \quad \binom{n}{k} = \frac{n!}{(n-k)!k!}$$

- The $\binom{n}{k}$ terms are also known as *Binomial Coefficients*.

Application of Combinations to Binomial Probabilities

- The expansion of the n-ordered polynomial in p and q is given by the **binomial theorem**

$$
\begin{aligned}
(p+q)^n &= \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} \\
&= \binom{n}{0} p^0 q^n + \binom{n}{1} p q^{(n-1)} + \binom{n}{2} p^2 q^{(n-2)} + \cdots + \binom{n}{n} p^n q^0 \\
&= q^n + npq^{n-1} + \frac{n(n-1)}{2} p^2 q^{(n-2)} + \cdots + p^n
\end{aligned}
$$

- The expansion on the right provides the **probabilities of an n order binomial (Bernoulli) experiment**, where each experiment leads to only two possible outcomes with probabilities p and $q = 1 - p$
- If $(p+q) = 1$, then the right-hand side sums to 1

Application of Combinations to Binomial Probabilities

For example, for an experiment comprising of n tosses of a coin.

- Let $P(\{H\}) = p$, $P(\{T\}) = q = 1 - p$
- Then, the probability of obtaining k heads in n tosses is

$$P(k \text{ heads in } n \text{ tosses}) = \binom{n}{k} p^k (1-p)^{n-k} = \frac{n!}{(n-k)!k!} p^k (1-p)^{n-k}$$

Application of Combinations to Binomial Probabilities

For example, for an experiment comprising of n tosses of a coin.

- Let $P(\{H\}) = p$, $P(\{T\}) = q = 1 - p$
- Then, the probability of obtaining k heads in n tosses is

$$P(k \text{ heads in } n \text{ tosses}) = \binom{n}{k} p^k (1-p)^{n-k} = \frac{n!}{(n-k)!k!} p^k (1-p)^{n-k}$$

- The binomial coefficients are given by Pascal's triangle:

n=1: 1 1
n=2: 1 2 1
n=3: 1 3 3 1
n=4: 1 4 6 4 1
n=5: etc.

- The binomial coefficients sum to $\sum_{k=0}^{n} \binom{n}{k} = \binom{n}{0} + \cdot + \binom{n}{n} = 2^n$

Example 1

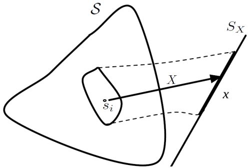
Consider three tosses of a biased coin with the probability of obtaining a head in each toss given by p. Find the probability of obtaining

a) three heads in the three tosses,

b) exactly two heads in the three tosses.

Random Variables

- Random variable X assigns a number $X(s)$ for every outcome s in the sample space S (*domain* of X)



- A random variable is a number or measurement associated with the outcome of an experiment
- Since the outcomes are random, the results of the measurements are also *random*
- The set of possible values of a random variable X is the *range* of X denoted by S_X

Random Variables

- A random variable is a function (rather than a variable) that maps points of the sample space to real numbers

$$\begin{aligned} X(\cdot):\quad S &\rightarrow R \\ s &\rightarrow X(s) \end{aligned}$$

  so that $S_X = \{x : x = X(s), s \in S\}$
- Random variables can be *discrete* (S_X is a discrete set) or *continuous* (S_X is a continuous set)

Note The uppercase alphabet X is used to denote a random variable, and the corresponding lowercase alphabet x denotes the values taken by the random variable.

Experiment: Toss two six-sided dice

Define Random Variables:

X is the minimum of the two dice. X assigns each outcome from sample space S a number from the following set: $S_X = \{1, 2, 3, 4, 5, 6\}$

Y is the sum of the two dice: $S_Y = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$

- Both X and Y are discrete random variables

Experiment: Symphony Hall

Define Random Variables:

N is the number of attendees on a given day:
$S_N = \{0, 1, 2, \cdots, 2262\}$
N is a discrete random variable

W is the weight of an individual attendee in pounds:
$S_W = \{w | 20 < w < 250\}$
W is a continuous random variable

Probability Mass Function (PMF)

- A random variable that can take finite or countably infinite number of values is characterized by a PMF
- The PMF - $p_X(x)$ - provides the probability of each numerical value a discrete random variable can take:

$$p_X(x) = \text{Probability of event}\{X = x\} = P[X = x] = P[\{\xi | X(\xi) = x\}]$$

Probability Mass Function (PMF)

- A random variable that can take finite or countably infinite number of values is characterized by a PMF
- The PMF - $p_X(x)$ - provides the probability of each numerical value a discrete random variable can take:

$$p_X(x) = \text{Probability of event}\{X = x\} = P[X = x] = P[\{\xi | X(\xi) = x\}]$$

- Procedure for calculating PMF, $p_X(x)$:
    - For each value x the RV X takes, collect all outcomes that give rise to $X = x$, and add their probabilities

Probability Mass Function (PMF)

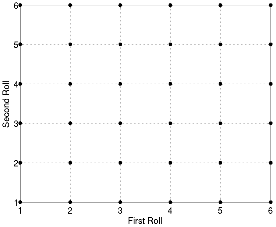
- A random variable that can take finite or countably infinite number of values is characterized by a PMF
- The PMF - $p_X(x)$ - provides the probability of each numerical value a discrete random variable can take:

$$p_X(x) = \text{Probability of event}\{X = x\} = P[X = x] = P[\{\xi | X(\xi) = x\}]$$

- Procedure for calculating PMF, $p_X(x)$:
  - For each value x the RV X takes, collect all outcomes that give rise to $X = x$, and add their probabilities
- The PMF satisfies the axioms of probability:
  - *Non-negativity*: $p_X(x) \geq 0$ for any x
  - *Normalization*: $\sum_{x \in S_X} p_X(x) = 1$
  - $P[X \text{ in } B] = \sum_{x \in B} p_x(x)$, where $B \subset S_X$

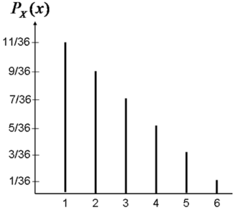
Experiment: Toss two six-sided dice

Consider an experiment consisting of tossing two six-sided dice. Define the random variable X to be the minimum of the two dice. Find and plot the probability mass function (PMF) of X.



Sample Space

$$p_X(x) = \begin{cases} 11/36 & \text{if } x = 1, \\ 9/36 & \text{if } x = 2, \\ 7/36 & \text{if } x = 3, \\ 5/36 & \text{if } x = 4, \\ 3/36 & \text{if } x = 5, \\ 1/36 & \text{if } x = 6, \\ 0 & \text{otherwise.} \end{cases}$$



Probability Mass Function

Note: PMF sums to 1 (normalization axiom): $\sum_X p_X(x) = 1$

Functions of a Random Variable

- Let X be a random variable with probability mass function $p_X(x)$
- Let random variable $Y = g(X)$ be a function of X
- The probability mass function of Y can be calculated from the probability mass function of X:

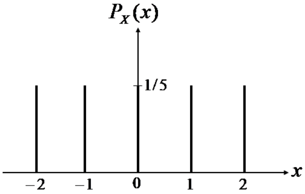
$$p_Y(y) = \text{Prob}(\{g(x) = y\}) = \sum_{x|g(x)=y} p_X(x)$$

Example 1

The discrete random variable X has the following probability mass function (PMF):

$$p_X(x) = \begin{cases} 1/5 & \text{if } x \in \{-2, -1, 0, 1, 2\}, \\ 0 & \text{otherwise.} \end{cases}$$

Find the PMF of random variables $Y = 2|X| + 3$, and $Z = X^2$.

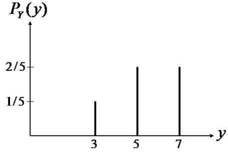


Example 1

The PMF of random variables Y and Z:

$$P_Y(y) = \begin{cases} 1/5 & \text{if } y = 3, \\ 2/5 & \text{if } y = 5, \\ 2/5 & \text{if } y = 7, \\ 0 & \text{otherwise.} \end{cases}$$

$$P_Z(z) = \begin{cases} 1/5 & \text{if } z = 0, \\ 2/5 & \text{if } z = 1, \\ 2/5 & \text{if } z = 4, \\ 0 & \text{otherwise.} \end{cases}$$



Expected Value (Mean) of a Random Variable

- The **expected value** or **mean** of a random variable X is defined as

$$\mu_X = E[X] = \sum_x x p_X(x)$$

- The mean is the weighted average of the possible values of X, weighted by its PMF
- Can also be considered as the center of gravity of the PMF
- Describes a typical value of a random variable

Moments of a Random Variable

- The n^{th} **moment** of a random variable X is

$$\mu_X = E[X^n] = \sum_x x^n p_X(x)$$

Note: The mean is the first moment of X.

- The n^{th} **central moment** of a random variable X is

$$E[(X - \mu_X)^n] = \sum_x (x - \mu_X)^n p_X(x)$$

Variance and Standard Deviation

- The **variance** of a random variable X:

$$\mathrm{Var}[X] = E[(X - E[X])^2] = E[(X - \mu_X)^2]$$

Note: The variance is the the 2^{nd} central moment of X.

- The **standard deviation** of a random variable X:

$$\sigma_X = \sqrt{\mathrm{Var}[X]}$$

- Provides a measure of the dispersion of X about its mean: the larger the standard deviation, the more dispersed are the values of X
- Tells us how uncertain the random variable is

Note: The variance and standard deviation are non-negative!

Variance and Standard Deviation

- The variance can be calculated using either one of the two equivalent formulas below

$$\mathrm{Var}[X] = \sum_x (x - \mu_X)^2 p_X(x)$$

$$\mathrm{Var}[X] = E[X^2] - \mu_X^2$$

Example

The discrete random variable X has the following probability mass function (PMF):

$$p_X(x) = \begin{cases} 1/5 & \text{if } x \in \{-2, -1, 0, 1, 2\}, \\ 0 & \text{otherwise.} \end{cases}$$

Find the mean and the variance of random variable $Y = 2|X| + 3$.

Example - Solution

- The mean of Y is

$$\begin{aligned}\mu_Y = E[Y] &= \sum_y y P_Y(y) \\ &= \frac{1}{5}(3) + \frac{2}{5}(5+7) \\ &= \frac{27}{5} = 5.4\end{aligned}$$

- The variance of Y is:

$$\begin{aligned}\mathrm{Var}[Y] &= \sum_y (y-\mu_Y)^2 P_Y(y) \\ &= \frac{1}{5}(3-5.4)^2 + \frac{2}{5}((5-5.4)^2 + (7-5.4)^2) \\ &= 56/25 = 2.24\end{aligned}$$

Example - Solution (continued)

- The variance can also be obtained by first calculating:

$$\begin{aligned} E[Y^2] &= \sum_y y^2 P_Y(y) \\ &= \frac{1}{5}(3^2) + \frac{2}{5}(5^2 + 7^2) \\ &= \frac{157}{5} \end{aligned}$$

- Then the variance is obtained from:

$$\mathrm{Var}[Y] = E[Y^2] - (E[Y])^2 = \frac{157}{5} - (\frac{27}{5})^2 = \frac{56}{25} = 2.24$$

- The standard deviation of Y is:

$$\sigma_Y = \sqrt{\mathrm{Var}[Y]} = \sqrt{\frac{56}{25}} = 1.5$$

Calculating Expectations of a Derived Random Variable

- Given a random variable X with probability mass function $p_X(x)$ and a random variable $Y = g(X)$ which is a function of X. We can compute the expected value of Y without deriving its probability mass function:

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

- The formula is called the **expected value rule**:

$$\begin{aligned} E[Y] &= \sum_y y \cdot p_Y(y) = \sum_y y \cdot \sum_{\{x|g(x)=y\}} p_X(x) \\ &= \sum_x g(x) p_X(x) \end{aligned}$$

Mean and Variance of a Linear Function

- Given a random variable X with probability mass function $p_X(x)$, and a random variable $Y = aX + b$ is a linear function of X (a and b are constants) then

$$E[Y] = E[aX + b] = aE[X] + b$$

$$\text{Var}[Y] = \text{Var}[aX + b] = a^2\text{Var}[X]$$

Note: The constant a is a scale factor while the constant b is a translation factor in the linear transformation from random variable X to Y. The variance is affected by scaling but unaffected by translation.

Note: In general $E[g(X)]$ is not equal to $g(E[X])$ unless $g(X)$ is a linear function of X.

Mean and Variance of a Linear Function

Proof:

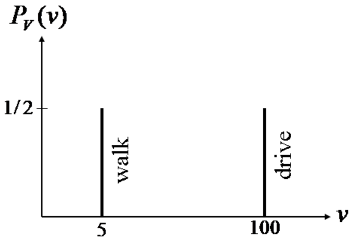
$$\begin{aligned} E[Y] = E[aX+b] &= \sum_x (ax+b)p_X(x) \\ &= a\sum_x xp_X(x) + b\sum_x p_X(x) \\ &= aE[X]+b \end{aligned}$$

$$\begin{aligned} \mathrm{Var}[Y] = \mathrm{Var}[aX+b] &= \sum_x (ax+b-E[aX+b])^2 p_X(x) \\ &= \sum_x (ax+b-aE[X]-b)^2 p_X(x) \\ &= a^2\sum_x (x-E[X])^2 p_X(x) = a^2\mathrm{Var}[X] \end{aligned}$$

Example

Consider the following probability mass function describing the travel speed V of a person's commute to work in units of miles per hour:

$$p_V(v) = \begin{cases} 1/2 & \text{if } v = 5, \\ 1/2 & \text{if } v = 100, \\ 0 & \text{otherwise.} \end{cases}$$



Example - Solution

- The average speed in miles per hour is

$$
\begin{aligned}
E[V] &= \sum_v v p_V(v) \\
&= \frac{1}{2}(5 + 100) = 52.5
\end{aligned}
$$

- The mean square speed in (miles/hour)^2 is

$$
\begin{aligned}
E[V^2] &= \sum_v v^2 p_V(v) \\
&= \frac{1}{2}(5^2 + 100^2) = 5012.5
\end{aligned}
$$

Note: $(E[V])^2 = 52.5^2 = 2756.2 \neq E[V^2]$.

Example - Solution (continued)

- Furthermore:

$$\begin{aligned} E[\frac{1}{V}] &= \sum_v \frac{1}{v} p_V(v) \\ &= \frac{1}{2}(\frac{1}{5} + \frac{1}{100}) = 21/200 = 0.105 \end{aligned}$$

Note: $\frac{1}{E[V]} = 1/52.5 = 0.019 \neq E[\frac{1}{V}]$.

Exercise What is the average time taken for a 200 mile distance travel?

Example - Solution (continued)

- Furthermore:

$$
\begin{aligned}
E[\frac{1}{V}] &= \sum_v \frac{1}{v} p_V(v) \\
&= \frac{1}{2}(\frac{1}{5} + \frac{1}{100}) = 21/200 = 0.105
\end{aligned}
$$

Note: $\frac{1}{E[V]} = 1/52.5 = 0.019 \neq E[\frac{1}{V}]$.

Exercise What is the average time taken for a 200 mile distance travel?

Define a new random variable for the travel time: $T = 200/V$.

The average travel time for the 200 mile journey in hours is:
$E[T] = E[\frac{200}{V}] = 200E[\frac{1}{V}] = (200)(0.105) = 21$ hours.

Compare this value to the time taken when travelling at the average speed of $E[V] = 52.5$ miles per hour: $200/E[V] = 3.8$ hours.

Conditioning a Discrete Random Variable

- For a discrete random variable X and conditioning event B, with $\text{Prob}(B) > 0$, the conditional probability mass function (PMF) of X is

$$p_{X|B}(x) = P(X = x|B)$$

Note: If the value $X = x$ is contained in the conditioning event B, then the conditional PMF is non-zero for that value of x. Otherwise, if it is not contained in event B, the conditional PMF is zero for that value of x.

$$p_{X|B}(x) = \begin{cases} \frac{p_X(x)}{P(B)} & \text{if } x \in B \\ 0 & \text{otherwise.} \end{cases}$$

- The conditional PMF satisfies the axioms of probability
  - *Non-negativity*: For any $x \in B$: $p_{X|B}(x) \geq 0$
  - *Normalization*: $\sum_{x \in B} p_{X|B}(x) = 1$

Example

The random variable X has PMF:

$$p_X(x) = \begin{cases} 1/10 & \text{if } x = 2, 3, 4, 5 \\ 3/10 & \text{if } x = 6, 7 \\ 0 & \text{otherwise.} \end{cases}$$

Define event $B = \{X \geq 4\}$. Find the conditional PMF $p_{X|B}(x)$.

Example

The random variable X has PMF:

$$p_X(x) = \begin{cases} 1/10 & \text{if } x = 2, 3, 4, 5 \\ 3/10 & \text{if } x = 6, 7 \\ 0 & \text{otherwise.} \end{cases}$$

Define event $B = \{X \geq 4\}$. Find the conditional PMF $p_{X|B}(x)$.

Solution

The probability of event B is $P(B) = P(X \geq 4) = 8/10 = 4/5$. Therefore the desired conditional PMF is

$$p_{X|B}(x) = \begin{cases} 1/8 & \text{if } x = 4, 5 \\ 3/8 & \text{if } x = 6, 7 \\ 0 & \text{otherwise.} \end{cases}$$

Conditional Expected Value for a Random Variable

- The **conditional expected value of random variable** X given condition B is:

$$E[X|B] = \mu_{x|B} = \sum_{x \in B} x p_{X|B}(x)$$

- The **conditional expected value of a function of a random variable** $Y = g(X)$ given condition B is

$$E[Y|B] = E[g(X)|B] = \sum_{x \in B} g(x) p_{X|B}(x)$$

- The **conditional variance of random variable** X given condition B can be calculated using

$$\text{Var}[X|B] = E[(X - \mu_{x|B})^2] = E[X^2|B] - \mu_{x|B}^2$$

Total Probability Theorem for Conditional PMF

- Given events B_1, B_2, ..., B_m form a partition of the sample space, we can obtain the **unconditional PMF** for X from the conditional PMFs

$$p_X(x) = \sum_{i=1}^{m} p_{X|B_i}(x) P(B_i)$$

- The **unconditional expectation** of X from the conditional expectations can be obtained using the **total expectation theorem**:

$$E[X] = \sum_{i=1}^{m} E[X|B_i] P(B_i)$$