ECE:226, Probability and Random Processes Notes

Anis Chihoub, Rutgers University

May 2020-July 2020

1 Introduction/Review Notes

- Grading
  - Mini Exams: 60%
    - Usually every week and take place on sakai for 24 hrs
  - Project: 10%
  - HW: 30%
    - 6 sets of HW due on Saturday usually

2 Week 1/Chapter 1

I'm assuming that this is Chapter 1 material, will add more when necessary

2.1 Set Builder Notation

A set is basically a defined set of objects
Example: C = {1,2,3}

Sets can be the subset of another set, showcased by the following symbol: $\subset$
Example: C $\subset$ A

Below are further symbols that might be of use:

- $A^c \Rightarrow$ complement of set A
- A$\emptyset \Rightarrow$ empty set
- $x \in A \cup B \iff x \in A or x \in B$

One part of set theory is Demorgan's law, which basically states that that $(A \cap B)^c == A^c \cup B^c$

Sets are basically used in probability in order describe sample spaces. Take for example rolling a dice and one wants to sort it by odds and evens. The sets are as follows.

1. $S_e = \{0, 1, 3, 5\}$
2. $S_o = \{2, 4, 6\}$

2.2 Axioms for probability

There are several axioms that the book lists here:

1. For any event, A, $P[A] \geq 0$
2. The probability of any element in the sample space is 0
3. The probability of mutually exclusive events is as follows: $P[A_1 \cup A_2 \cup A_3 ...] = P[A_1] \cap P[A_2] \cap P[A_3]...$

2.3 Relevant Theorems

Using the above axioms we can get the following theorems

1. **Theorem 1** $P[A_1 \cup A_2] = P[A_1] + P[A_2]$ *for mutually exclusive events*
2. **Theorem 2** $a \in A, A_i \cap A_j \neq \emptyset, i \neq j$, *then* $P[A] = \sum_{i=1}^{m} P[A_i]$ *Or the sum of all probabilities in a set is the probability of that set.*
3. (a) **Theorem 3** $P[\emptyset] = 0$
   (b) **Theorem 4** $P[A^c] = 1 - P[A]$
   (c) **Theorem 5** $P[A \cup B] = P[A] + P[B] - P[A \cap B]$
   (d) **Theorem 6** $A \subset B$, *then* $P[A] \leq P[B]$

   It is important to note that you can prove all of these using the aforementioned axioms.
4. **Theorem 7** *The probability of an event, B is equal to the sum of the probabilities of the outcomes contained in that event is equal to* $P[B] = \sum_{i=1}^{m} P[s_i]$ *for some set* $B = \{S_1, S_2, S_3\}$

   It is important to note that the book uses $P[A \cap B] = P[A, B], P[AB]$ **as notation for A AND B**
5. For elements of equal probability, use the following notation $P[S] = \frac{1}{n}$
   The following is a proof of the above theorem

(a) Given that the set has equal probability, we know that $\sum_{i=1}^{m} P[B_i] = mp$

(b) Axiom 1 gives that $P[B] = 1$

(c) solve for $p = \frac{1}{m}$

2.4 Conditional Probability

In this part of the book, we are dealing with probability in which we already have a prior probability. This is expressed as $P(A \mid B)$. This can be expressed in the following formula.

$$P(A \mid B) = \frac{P[A \cap B]}{P[B]} \tag{1}$$

There are also several axioms that can be derived from this equation.

1. $P(A \mid B) \geq 0$
2. $P(B \mid B) = 1$
3. $P(A \mid B) = \sum_{i=1}^{m} P(A_i \mid B)$

2.5 Partitions

A partition splits a set into mutually exclusive,collectively exhaustive sets. For example, as follows in the next figure.

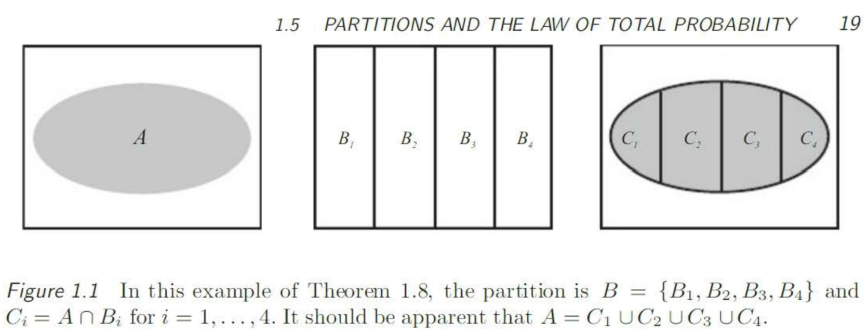

Figure 1: A Partitioned Data Set

An important theorem to note is as follows:

Theorem 8 *Let B be a partition of Set A. Given that $A \cap B_i = C$, then one can say $A = \sum_{i=1}^{m} C$. So the sum of the partition probabilities in their intersection with A will give A*

2.6 Law of Total Probability

Theorem 9 *One way to write P[A], given a partition set B, can be expressed as follows.*

$$P[A] = \sum_{i=1}^{m} P[A \mid B_i]P[B_i] \tag{2}$$

The above theorem allows us to state a total probability in terms of the above quantities. Below is a figure to help explain the theorem a little more.

Example 1.19

A company has three machines B_1, B_2, and B_3 making 1 kΩ resistors. Resistors within 50 Ω of the nominal value are considered acceptable. It has been observed that 80% of the resistors produced by B_1 and 90% of the resistors produced by B_2 are acceptable. The percentage for machine B_3 is 60%. Each hour, machine B_1 produces 3000 resistors, B_2 produces 4000 resistors, and B_3 produces 3000 resistors. All of the resistors are mixed together at random in one bin and packed for shipment. What is the probability that the company ships an acceptable resistor?

Let $A = \{\text{resistor is acceptable}\}$. Using the resistor accuracy information to formulate a probability model, we write

$$\mathrm{P}[A|B_1] = 0.8, \quad \mathrm{P}[A|B_2] = 0.9, \quad \mathrm{P}[A|B_3] = 0.6. \tag{1.27}$$

The production figures state that $3000 + 4000 + 3000 = 10{,}000$ resistors per hour are produced. The fraction from machine B_1 is $\mathrm{P}[B_1] = 3000/10{,}000 = 0.3$. Similarly, $\mathrm{P}[B_2] = 0.4$ and $\mathrm{P}[B_3] = 0.3$. Now it is a simple matter to apply the law of total probability to find the acceptable probability for all resistors shipped by the company:

$$\mathrm{P}[A] = \mathrm{P}[A|B_1]\,\mathrm{P}[B_1] + \mathrm{P}[A|B_2]\,\mathrm{P}[B_2] + \mathrm{P}[A|B_3]\,\mathrm{P}[B_3] \tag{1.28}$$

$$= (0.8)(0.3) + (0.9)(0.4) + (0.6)(0.3) = 0.78. \tag{1.29}$$

For the whole factory, 78% of resistors are within 50 Ω of the nominal value.

Figure 2: Example using Theorem 9

2.7 Independence

The following are important theorems that can be discussed.

Theorem 10 *Two events are Independent if and only if* $P[A \cap B] = P[A]P[B]$

Theorem 11 *If* $P[A], P[B] \neq 0$ *and are independent events, then*

1. $P[A \mid B] = P[B]$
2. $P[B \mid A] = P[A]$

Theorem 12 *Mutual Independence is defined for* A_n *sets iff.*

- *A_1 and A_2 are independent*
- *A_2 and A_3 are independent*
- *A_1 and A_3 are independent*

So on and so forth.

3 Week 2

3.1 Chapter 2

Tree diagrams are basically the law of total probability that can be visualized. More formally, the outcomes of sub-experiments performed

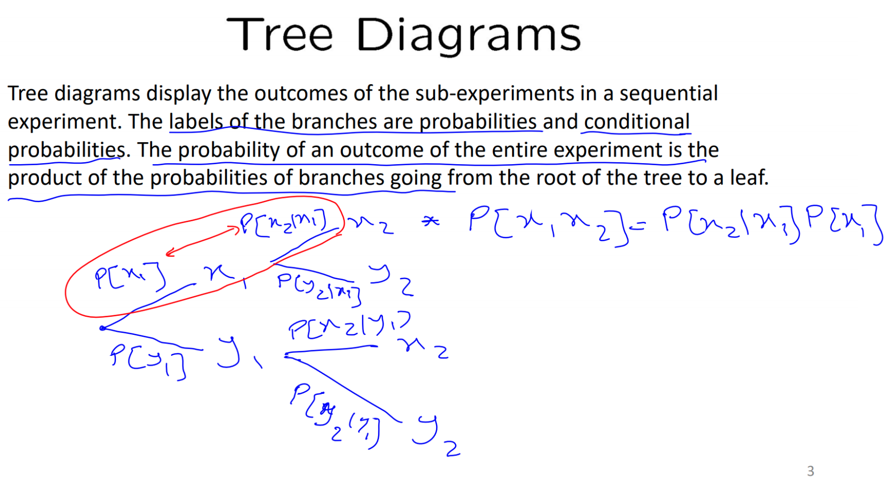

Figure 3: Tree Diagram Probability

By following the events of a subexperiment, one can then deduce some sort of conditional probability. See the figure below for an example.

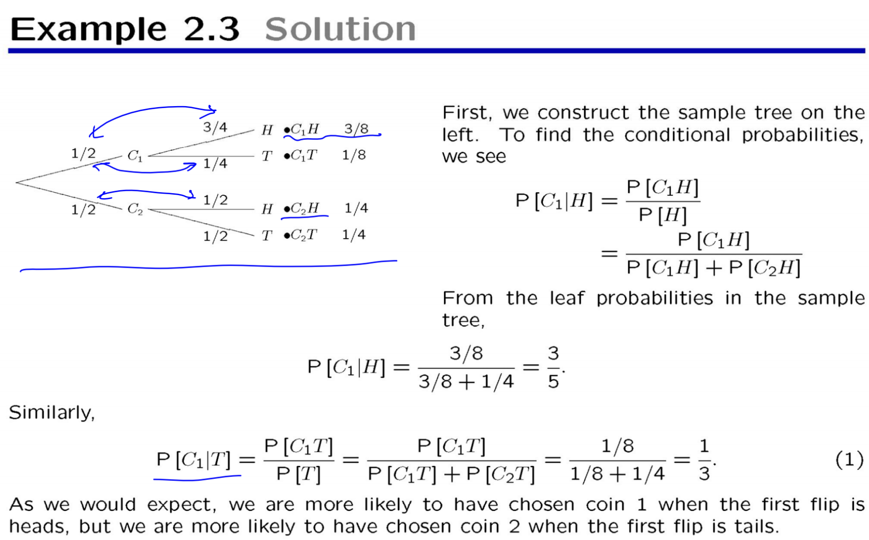

Figure 4: Tree Diagram Probability Example

Now let us begin with counting. Counting essentially shows the maximum number of outcomes of an experiment. Below are the important theorems that the textbook outlines.

Theorem 13 *For two distinct sub experiments, if one has n outcomes and the other k outcomes. The total number of outcomes is nk*

Theorem 14 *The number of k permutations in n objects can be represented by the following equation $\frac{n!}{(n-k)!}$*

Theorem 15 *The number of ways to choose k objects out of n distinguishable objects is $\binom{n}{k} = \frac{n!}{k!(n-k)!}$, otherwise called a **Combination**. This is defined for all $n \geq 0$ otherwise it is just 0*

Now, what if order does matter. In this case there are several theorems that help us do the same.

Theorem 16 *Given m distinguishable objects, there are m^n ways to choose an ordered sample of n objects.*

Theorem 17 *For n repetitions of a subexperiment with a sample space of S with m items there are m^n possible outcomes*

Think of the above two theorems as kinds of inverses of each other. To this end, now we have the ability to know the number of items that one can pick. Applying the above to the following theorems yields the following.

Theorem 18 *The number of observation sequences for n subexperiments with sample space* $S = \{0,1,\ldots n\}$ *can be expressed as follows* $\binom{n}{n_0 \ldots n_m} = \frac{n!}{n_0!n_1!n_2!\ldots n_m!}$ *This is defined for* $n \geq 0$.

For the last part of chapter 2. We discuss Independent Trials. Below, one can write the number of n_0 and n_1 failures and successes as follows.

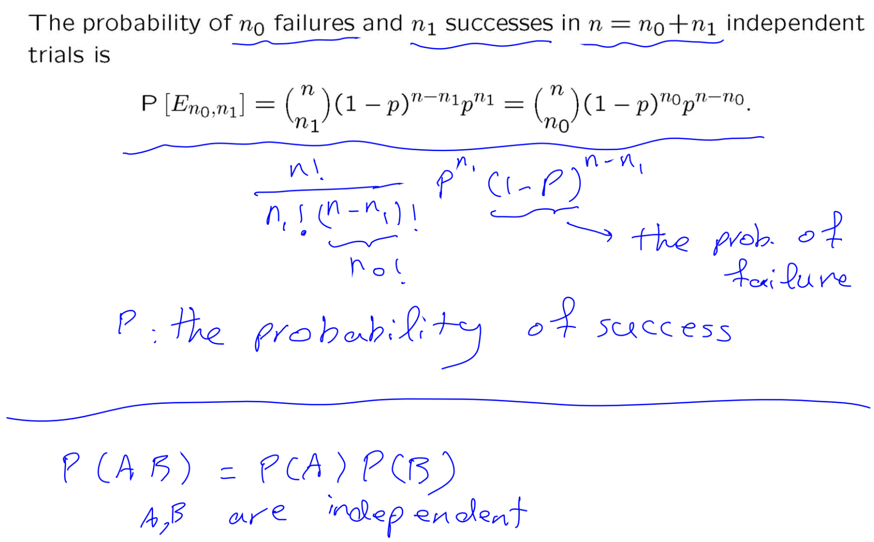

Figure 5: Independent Trials Equation

For independent trials, the probability of occurrences of some event can be expressed as.

$$\binom{n}{n_0 \ldots n_m} = \frac{n!}{n_0! \ldots . n_(m-1)!} \tag{3}$$

3.2 Chapter 3

This chapter begins with a probability mass function. It is defined as follows - the PMF of random variable X expresses the probability model of an experiment as a mathematical function. The function is the probability P[X = x] for every number x.

Formally, we can write the probability mass function as follows:

$$P_X(x) = P[X = x] \tag{4}$$

That is for each x in the function, it must sum to 1 somehow.

The same probability theorems discussed in Chapter 1 can be expressed for discrete random variables.

Theorem 19 *a For any x* $P_X(x) \geq 0$

b $\sum_{x \epsilon S_x} P_X(x) = 1$

c All events inside $P_X(x)$ *must sum to zero.*

Often, discrete random variables are expressed as a family. For example, if you want to combine multiple variables, one must write a probability mass function. See the below figure for an example.

Example 3.6

Consider the following experiments:

- Flip a coin and let it land on a table. Observe whether the side facing up is heads or tails. Let X be the number of heads observed. $X=1$ $X=0$
- Select a student at random and find out her telephone number. Let $X = 0$ if the last digit is even. Otherwise, let $X = 1$.
- Observe one bit transmitted by a modem that is downloading a file from the Internet. Let X be the value of the bit (0 or 1).

All three experiments lead to the probability mass function

$$P_X(x) = \begin{cases} 1/2 & x = 0, \\ 1/2 & x = 1, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Figure 6: Family of Discrete Variables Equations

Now lets discuss a Bernoulli discrete variable. They often take the following form.

$$P_X(x) = \begin{cases} 1-p & \text{x} = 0 \\ p & \text{x} = 1 \\ 0 & otherwise \end{cases}$$

Another important variable can be described as a geometric variable.

$$P_X(x) = \begin{cases} p(1-p)^{x-1} & x = 1, 2... \\ 0 & otherwise \end{cases}$$

The last two variables are binomial and pascal variables. Examples will not be included but I will reference the book later. (5) is a binomial function while (6) is the pascal variable.

$$P_X(x) = \binom{n}{x} p^x (1-p)^{n-x} \tag{5}$$

$$P_X(x) = \binom{x-1}{k-1} p^x (1-p)^{k-x} \tag{6}$$

The above equations must satisfy the following conditions:

1. $0 < p < 1$
2. $k \geq 1$

With multiple equations being covered, when should each one be used? In that sense, the following can be surmised.

1. Geometric: No set number of trials
2. Binomial: Set number of trials, no more than
3. Pascal: Exact number

Now, the book introduced another type of variable, discrete uniform,used for equiprobable outcomes:

$$P_X(x) = \begin{cases} \frac{1}{l-k+1} & k = k, k+1, k+2... \\ 0 & otherwise \end{cases}$$

Where k and l are parameters described and $k < l$

To find the average number of something, we use the poisson variable.

$$P_X(x) = \begin{cases} \frac{\alpha^x * (-e^\alpha)}{x!} & x = 0, 1, 2... \\ 0 & otherwise \end{cases}$$

4 Week 3

4.1 CMFs

The book begins with a discussion of the cumulative distribution function. Cumulative distribution functions are functions that display probability as a function of all previous probabilities. Below are some relevant properties, given a discrete random variable X with range $S_X = \{x_1, x_2, ...\}$ and $x_1 \leq x_2 \leq ..$:

1. $F_X(-\infty) = 0$
2. $F_X(\infty) = 1$
3. For all $x' \geq x$, $F_X(x') \geq F_X(x)$, so the CDF is always increasing
4. For all $x_i \in S_X, P_X(x_i)$ is given as follows

$$F_X(x_i) - F_X(x_i - \epsilon) = P(x_i) \tag{7}$$

   Given that ϵ is an extremely small positive number. So this means that the difference between the two values is the probabilty

5. $F_X(x) = F_X(x_i), \forall x_i \leq x \leq x_{i+1}$, meaning there is a vertical discontinuity between jumps

Now, there is an important theorem to consider.

Theorem 20 *For all $b \geq a$, it follows that if $F_X(x)$ is the probability function, then:* $F_X(b) - F_X(a) = P[a < x \leq b]$

Below is an image of a CDF being sketched.

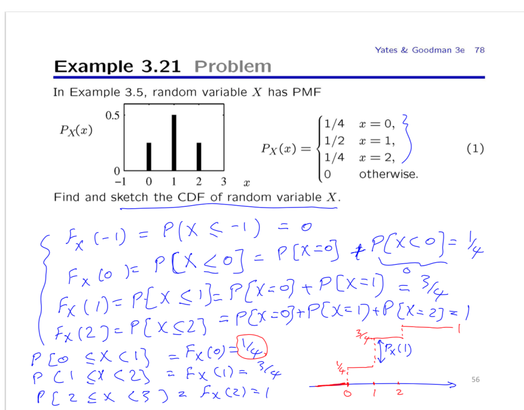

Figure 7: CDF Example

4.2 Expected Value

Now, we will discuss expected values in probability. The expected value of a variable, say X is given as follows:

$$\mu_X = \sum_{x \in S_X} x P_X(x) \tag{8}$$

That is, the average value is the value at x multiplied the probability at x

Now, given the variables discussed before in chapter 3, below is a list of their expected values:

- Bernoulli:p
- Geometric: $\frac{1}{p}$
- Poisson: α,i.e. the rate multiplied by the time
- Binomial: np
- Pascal: k/p
- Discrete Uniform: $\frac{k+l}{2}$

4.3 Functions of Random Variables

A Derived Random Variable is expressed as a function of another random variable. Thus, the pmf of Y is:

$$P_Y(y) = \sum_{x:g(x)=y} P_X(x) \tag{9}$$

The expected value is:

$$\mu_Y = \sum_{x \epsilon S_X} g(x) P_X(x) \tag{10}$$

The expected value of any point and its residual is 0, so we can then write.

When dealing with constants in front of the random variable, we can use the following theorem:

Theorem 21 $E[aX + b] = aE[X + b]$

The variance of a random variable is defined as follows:

$$\sigma_X = E[(X - \mu_x)^2] \tag{11}$$

The standard deviation is:

$$\sqrt{\sigma_X} \tag{12}$$

The variance can also be rewritten as follows:

$$Var[X] = E[X^2] - (E[X])^2 \tag{13}$$

Manipulating the exponent on each of the above produced a moment. This is denoted in numerical order. It is also important to note that should you remove constant coefficient, it becomes a^2 instead of a

4.4 Probability Density Function

The PDF is simply the derivative with respect to X of the PMF. Shown as follows:

$$f_X(x) = \frac{dF_x}{dx} \tag{14}$$

Below are some important properties:

1. The function is strictly increasing
2. $F_X(x) = \int_{-\infty}^{x} f_X(x)$
3. $1 = \int_{-\infty}^{\infty} f_X(x)$

Therefore, taking the appropriate integral of the PDF yields the probability in question.

4.5 Expected Values of CDFs

Similar to the PDF, the expected value of a CDF can be written as follows:

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx \tag{15}$$

Likewise, for a random variable defined by a function:

$$E[g(x)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx \tag{16}$$

Now, we shall list some relevant theorems relating to random variables as review

1. $E[X - \mu_X] = 0$
2. $E[aX + b] = aE[X] + b$
3. $Var[X] = E[X^2] - \mu_X^2$
4. $Var[aX + b] = a^2 Var[X]$

Below is an example that ties the above all together

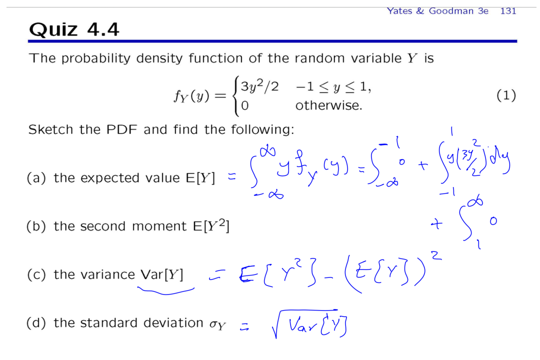

Figure 8: PDF Example

4.6 Derived Random Variables

From (16), we know the expected value of a Derived Random Variable. The same theorems for standard random variables also apply. They will be listed for convenience

1. $E[X - \mu_X] = 0$
2. $E[aX + b] = aE[X] + b$
3. $Var[X] = E[X^2] - \mu_X^2$
4. $Var[aX + b] = a^2 Var[X]$

4.7 Uniform Random Variables

A Uniform Random Variable is Defined as Follows if its PDF is as follows:

$$F_X(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{otherwise} \end{cases} \tag{17}$$

Some relevant properties are listed below:

1. The CDF of a Uniform Random Variable is
$$f_X(x) = \begin{cases} 0 & x \leq a \\ \frac{x-a}{b-a} & a \leq x \leq b \\ 1 & x > b \end{cases} \tag{18}$$
2. The expected value is
$$E[X] = \frac{b+a}{2} \tag{19}$$
3. The variance is
$$Var[X] = \frac{(b-a)^2}{12} \tag{20}$$

4.8 Exponential Random Variables

A variable is an exponential random variable if its PDF is as follows:

$$f_X(x) = \begin{cases} \lambda e^{\lambda x} & x \geq 0 \\ 0 & otherwise \end{cases} \tag{21}$$

The relevant properties are listed as follows:

1. The CDF of a Exponential Random Variable is

$$f_X(x) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & otherwise \end{cases} \tag{22}$$

2. The expected value is

$$E[X] = \frac{1}{\lambda} \tag{23}$$

3. The variance is

$$Var[X] = \frac{1}{\lambda^2} \tag{24}$$

4.9 Erland Random Variables

The final variable discussed is the Erland variable. If the PDF is as follows, then it is an Erland variable.

$$F_X(x) = \begin{cases} \frac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!} & x \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{25}$$

Given that $\lambda > 0$ and $n \geq 1$

The following is the properties of the Erlang variable.

1. $E[X] = \frac{n}{x}$
2. $Var[X] = \frac{n}{\lambda^2}$

Finally, some examples to consider (see the attatched pdf)

5 Week 4

5.1 Gaussian Variables

A Gaussian Random Variable is a variable who's PDF is as follows:

$$f_X(x) = \frac{-e^{\frac{-(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} \tag{26}$$

Its expected value and Variance are as follows:

- $E[X] = \mu$
- $Var[X] = \sigma^2$

For any derived random variable, we can write the gaussian with the following parameters $(\mu, \sigma) \longrightarrow (a\mu + b, a\sigma)$ for some $Y = aX + b$ is X is a Gaussian Variable.

If μ is 0 and σ is 1, then the distribution is **normal**.

The CDF of a standard normal variable is expressed as follows:

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{\frac{-u^2}{2}} du \tag{27}$$

Should that variable happen to be normal, it can be written as follows for $P[a < X \leq b]$

$$\phi(\frac{b-\mu}{\sigma}) - \phi(\frac{a-\mu}{\sigma}) \tag{28}$$

As usual, $\phi(-z) = 1 - \phi(z)$

5.2 Multiple Random Variables Introduction

Basically, this section is combining two random variables together and analyzing the result. Below are some important properties:

1. $0 \leq F_{X,Y} \leq 1$
2. $F_{X,Y}(-\infty, \infty) = 1$
3. $F_X(x) = F_{X,Y}(X, \infty)$
4. $F_Y(y) = F_{X,Y}(\infty, Y)$
5. $F_{X,Y}(x, -\infty) = 0$
6. $F_{X,Y}(-\infty, Y) = 0$
7. If $x \leq x_1$ and $y \leq y_2$ we can say the function is strictly increasing

To find the respective PDFs of multiple random variables, this will require the use of partial derivatives for each variable. Likewise, to find the respective probability of an event given the CDF, multiple integrals will need to be used.

5.3 Marginal PMFs

Now, we introduce another theorem:

1. $$P_X(x) = \sum_{y \in S_Y} P_{X,Y}(x, y) \tag{29}$$

2. $$P_Y(y) = \sum_{x \in S_X} P_{X,Y}(x, y) \tag{30}$$

The problems in this section deal exactly with table like problems and the application of principles from this section.

5.4 Joint Probability Density Function

In order to express the joint PDF, one needs to use a double integral as follows:

$$F_{X,Y}(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(u, v) dv du \tag{31}$$

Likewise, given the CDF, the PDF can be expressed as follows:

$$f_{X,Y}(x, y) = \frac{\partial^2 F_{X,Y}(x, y)}{\partial x \partial y} \tag{32}$$

Some of the theorems in the previous section may also apply here as well. They are listed below for conveniece:

1. $f_{X,Y}(x, y) \geq 0$
2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx dy = 1$

Again, given a probability as follows $P[a < x \leq b, c < y \leq d]$. It is important to set up the double integral to evaluate. Therefore, if we are given the region that $P[A]$ exists, the simplest thing to do is just employ the double integral.

6 Week 5

6.1 Marginal PDF

When given a joint pdf of $f_{X,Y}(x,y)$, we can find the marginal pdfs of x and y as follows:

$$f(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy \tag{33}$$

$$f(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx \tag{34}$$

Using these equations, we can then determine independence:

6.2 Independence and Expected Values

We can say that two variables are independent if and only if the following is true:

Theorem 22 *Given a PMF:* $P_X(x) * P_Y(y) = P_{X,Y}(x,y)$

Theorem 23 *Given a PDF:* $F_X(x) * F_Y(y) = F_{X,Y}(x,y)$

Likewise, for any expected values, we can write:

Theorem 24 *Given a PMF:* $\sum_{x \in S} \sum_{y \in S} g(x,y)P(x,y)$

Theorem 25 *Given a PDF:* $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y)f(x,y)$

As always, for multiple variables in a set, just sum them together.
Now, we discuss the variance. For two random variables, we can write:

$$E[X+Y] = E[X] + E[Y] \tag{35}$$

$$Var[X+Y] = Var[X] + Var[Y] - 2E[(X-\mu_X)(Y-\mu_Y)] \tag{36}$$

The above can also be expressed as follows given that one variable, say Y is expressed as follows: $Y = X + Z$:

$$Var[X+Y] = (\sigma_X)^2 + (\sigma_Z)^2 - 2E[XZ] \tag{37}$$

6.3 Covarriance, Correlation

We can express the covarriance of two variables as follows:

$$Cov[X,Y] = E[(X-\mu_X)(Y-\mu_Y)] \tag{38}$$

Now, the co variance represents the mean value of the product of the deviations of two variates from their respective means. We can use the covariance to measure correlation:

$$\rho_{X,Y} = \frac{Cov[X,Y]}{\sqrt{Var[X]Var[Y]}} \tag{39}$$

Correlation ranges between -1 and 1. The closer to 1, there is a positive correlation. The closer to -1, the more there is a negative correlation.

For a dervied variable, say, $\hat{X} = aX + b$ and $\hat{Y} = cY + d$, we can write the following:

1. $\rho_{\hat{X},\hat{Y}} = \rho_{X,Y}$
2. $acCov[\hat{X},\hat{Y}] = Cov[aX + b, cY + d]$

Correlation can also be expressed as $r = E[XY]$. Some theorems about correlation can be seen as follows:

1. $Cov[X,Y] = r_{X,Y} - \mu_X\mu_Y$
2. $Var[X+Y] = Var[X] + Var[Y] + 2Cov[X,Y]$
3. If $X = Y$, then $COV[X] = COV[Y]$ and $R_{X,Y} = E[X^2] = E[Y^2]$

6.4 Special Cases

The following are some special cases:

- If X and Y are orthogonal, $r_{X,Y} = 0$
- If X and Y are uncorrelated, $Cov[X,Y] = 0$

If we are dealing with Independent variables, then the following applies:

1. $E[g(x)h(y)] = E[g(x)]E[h(y)]$
2. $r_{X,Y} = E[X]E[Y]$
3. $Cov[X,Y] = 0$
4. $Var[X+Y] = Var[X] + Var[Y]$

6.5 Bivariate Gaussian Random Variables

Please refer to the slides in this part since the equations in question are way too complicated to input.

7 Week 6

8 Week 7

9 Week 8

10 Relevant Figures

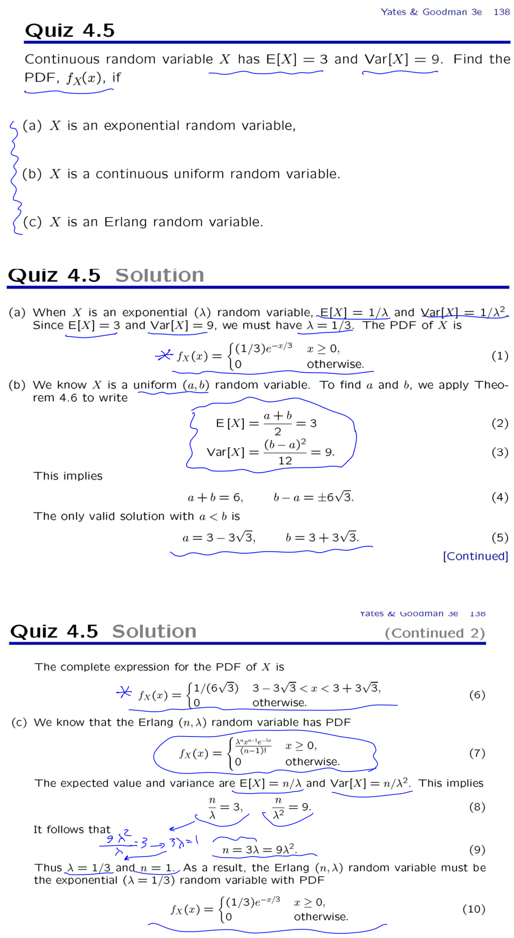

Figure 9: Variable Identification Example

11 Any Extraneous Things I Found Useful

Damn LaTex is fucking useful My personal ratings of chapters $0 - 10$ scale

1. Chapter 1: 0 brain dead easy
2. Chapter 2: 6 counting is kinda hard ngl
3. Chapter 3: 2 memorize formulae

4. Chapter 4: 0 pretty ex
5. Chapter 5: 4 double integrals, that's it.

12 Review Sheet

Set Theory

$$P[A \cap B] = P[AB]$$

$$P[A] \geq 0$$

$$P[\emptyset] = 0$$

$$P[A \cup B] = P[A] + P[B] - P[A \cap B]$$

$$P[A^c] = 1 - P[A]$$

$$P[A|B] = \frac{P[AB]}{P[B]}$$

$$P[B|B] = 1$$

$$\text{Bayes Theorem: } P[B|A] = \frac{P[A|B]P[B]}{P[A]}$$

$$\text{n choose k} = \big(n\big) = \frac{n!}{k!(n-k)!}$$

Table of Discrete Random Variables —Name of Random Variable— Description of variable — PMF Function — Expected Value-

$$E[X]$$

— Variance -

$$Var[X]$$

— ————————————————————-————————— —Bernoulli—number successes in 1 trial —

$$P_X(x) = \begin{cases} 1-p, & x = 0\ p, \\ x = 1\ 0, & \text{otherwise} \end{cases}$$

—

$$p$$

—

$$p(1-p)$$

— —Geometric—number of trials until 1st success—

$$P_X(x) = \begin{cases} p(1-p)^{x-1}, & x = 0, \\ \text{otherwise} \end{cases}$$

—

$$\frac{1}{p}$$

—

$$\frac{1-p}{p^2}$$

— —Binomial —number of successes in n trials —

$$P_X(x) = \left(n\ x\right) p^x (1-p)^{n-x}$$

—

$$np$$

—

$$np(1-p)$$

— —Pascal —number of trials until k successes—

$$P_X(x) = \left(x-1\ k-1\right) p^k (1-p)^{x-k}$$

—

$$\frac{k}{p}$$

—

$$\frac{k(1-p)}{p^2}$$

— — Poisson — Probability of x arrivals in T seconds —

$$P_X(x) = \begin{cases} \frac{\alpha^x e^{-\alpha}}{x!}, x = 0,1,2\dots\ 0, \text{otherwise} \end{cases}$$

—

$$\alpha$$

—

$$\alpha$$

— —Discrete Uniform — Probability of any value between k and l —

$$P_X(x) = \begin{cases} \frac{1}{l-k+1},\ x = k, k+1, k+2\dots\ 0, \text{otherwise} \end{cases}$$

—

$$\frac{k+l}{2}$$

—

$$\frac{(l-k)(l-k+2)}{12}$$

—

Table of Continuous Random Variables — Name — Form — PDF — CDF

—

$$E[X]$$

—

$$Var[X]$$

— — Uniform—

$$(a,b)$$

—

$$\begin{cases} \frac{1}{b-a} & a \leq x < b \ 0 \\ \text{otherwise} \end{cases}$$

—

$$\begin{cases} 0 & x \leq a \ \frac{x-a}{b-a} \\ a < x \leq b \ 1 & x > b \end{cases}$$

—

$$\frac{b+a}{2}$$

—

$$\frac{(b-a)^2}{12}$$

— — Exponential —

$$(\lambda)$$

—

$$\begin{cases} \lambda e^{-\lambda x} & x \geq 0 \ 0 \\ \text{otherwise} \end{cases}$$

—

$$\begin{cases} 1 - e^{-\lambda x} & x \geq 0 \ 0 \\ \text{otherwise} \end{cases}$$

—

$$\frac{1}{\lambda}$$

—

$$\frac{1}{\lambda^2}$$

— —Erlang—

$$(n, \lambda)$$

—

$$\begin{cases} \frac{\lambda x^{n-1} e^{-\lambda x}}{(n-1)!} & x > 0 \ 0 \\ \text{otherwise} \end{cases}$$

— —

$$\frac{n}{\lambda}$$

—

$$\frac{n}{\lambda^2}$$

— — Gaussian —

$$(\mu, \sigma)$$

—

$$\frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$

— *

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-u^2/2} du$$

—

$$\mu$$

—

$$\sigma^2$$

—

* Note that this is the standard normal CDF of the Gaussian random variable. To adjust for

$$\mu$$

and

$$\sigma$$

use

$$\Phi(\frac{x-\mu}{\sigma})$$

List of Essential Equations

$$P[x_1 < x < x_2] = F_X(x_2) - F_X(x_1)$$

$$\int_{-\infty}^{\infty} f_X(x)dx = 1$$

Expected Value

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

$$E[X - \mu_x] = 0$$

$$E[aX + b] = aE[X] + b = 0$$

$$E[X + Y] = E[X] + E[Y]$$

Variance

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

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

$$Var[aX + b] = a^2Var[X]$$

$$Var[X + Y] = Var[X] + Var[Y] + 2Cov[X, Y]$$

Covariance and Correleational Coefficients

$$\rho_{X,Y} = \frac{Cov[X, Y]}{\sqrt{Var[X]Var[Y]}}$$

$$-1 \leq \rho_{X,Y} \leq 1$$
$$r_{X,Y} = E[XY]$$
$$Cov[X,Y] = r_{X,Y} - \mu_X \mu_Y$$
$$Cov[X,X] = Var[X] = E[X^2] - (E[X])^2$$

Independent Variables Given two independent random variables:

$$f_{X,Y}(x,y) = f_X(x) f_Y(y)$$
$$r_{X,Y} = E[XY] = E[X]E[Y]$$
$$Cov[X,Y] = \rho_{X,Y} = 0$$
$$Var[X+Y] = Var[X] + Var[Y]$$