STA257: Probability and Statistics 1

Instructor: Katherine Daignault

Department of Statistical Sciences
University of Toronto

Week 8

Outline

Outline

Calculus Review - Derivatives

- Because we are moving into the case where we are dealing with more than one random variable at a time, you will now need to be able to take partial derivatives.
- When dealing with a function of several variables, $f(x, y, \ldots)$, we can take the derivative with respect to each variable, as

$$\frac{\partial}{\partial x} f(x, y, \ldots) \text{ or } \frac{\partial}{\partial y} f(x, y, \ldots)$$

- We sometimes use a shorthand notation of f'_x or f'_y to denote the partial derivative with respect to x and y respectively.
- When taking partial derivatives, we just treat all other variables as constants, and all normal rules of differentiation apply.

Calculus Review - Multiple Integration

- Just as we need to take partial derivatives for multiple variables, we need to take multiple integrals over all the variables in our function.
- We define the double integral over $f(x, y)$ as

  $$\int_{supp(X)} \int_{supp(Y)} f(x, y) \partial y \partial x$$

  where we use the term $supp(X)$ to denote the support of X, or the values that X is defined over.
- In this case, you (1) take the integral with respect to Y, treating the x's as constants, then (2) take the integral of the result with respect to X.

Calculus Review - Multiple Integration

- Be careful about the order of the integrals!
  - If you want to integrate Y first, then the inner integral must correspond to the integral over y values
  - If you want to integrate X first, then you need to switch the order of the integrals to

$$\int_{supp(Y)} \int_{supp(X)} f(x, y) \partial x \partial y$$

- Since a single integral is the area under a curve, a double integral find the area **between** two curves.
- Often the support of one variable will be defined in terms of the other variable so you will need to be able to manipulate the bounds of integration.

Calculus Review - Multiple Integration

- If we are dealing with a function $f(x, y)$ defined on the region

  $$F = \{(x, y) : a \leq x \leq b, p(x) \leq y \leq q(x)\}$$

  then we must take the integral in the following order

  $$\int \int_F f(x, y)dA = \int_a^b \int_{p(x)}^{q(x)} f(x, y)\partial y \partial x$$

- If I wanted to take the integral in the reverse order, then I would need to determine the functions $r(y)$ and $s(y)$ so that

  $$\int \int_F f(x, y)dA = \int_c^d \int_{r(y)}^{s(y)} f(x, y)\partial x \partial y$$

  and $F = \{(x, y) : c \leq y \leq d, r(y) \leq x \leq s(y)\}$

Outline

Joint Distributions

- We have only been dealing with a single variable and its distribution up until this point.
- In practice, we are often interested in knowing how two or more random quantities (variables) are associated to one another.
  - in clinical trials, response to a drug/treatment is often associated with drug dosage, severity of disease as well as possible other demographic factors like sex or age.
  - in ecology, the total number of each species is affected by the number of predatory species and the number of prey species, as one eats the other.
- So we need to account for the fact that variables may be associated with each other by looking at their **joint distribution**.

Joint Distributions - Definition

- Suppose we are dealing with the clinical trial example, where we think drug response is going to change depending on a subject's age.
- So to determine the probability distribution, I need to consider these two variables (drug response and age) together, jointly.
- I can therefore write out a distribution function (CDF), representing the probability of having a drug response less than x ($X \leq x$) and age less than y ($Y \leq y$) simultaneously:

$$F(x, y) = P(X \leq x, Y \leq y) = P(X \leq x \cap Y \leq y)$$

- We refer to $F(x, y)$ as the joint CDF.

Joint CDF - Generalized

- Of course, I can look at the joint behaviour of any number of random variables by writing out a generic joint CDF involving n variables:

$$F(x_1, x_2, \ldots, x_n) = P(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n)$$

- Suppose we are dealing with only 2 random variables, $F(x, y)$.
- What does it mean to find the following joint probability?

$$F(a, b) = P(X \leq a, Y \leq b)$$

Probabilities using Joint CDFs

- When considering X and Y jointly, we are looking at a 2 dimensional surface
  - X corresponds to the x-axis
  - Y corresponds to the y-axis
- If $X \leq a$, we are considering all x values below a
  - corresponds to a vertical line at $x = a$ and shading to the left.
- If $Y \leq b$, this considers all y values below b
  - corresponds to a horizontal line at $y = b$ and shading below it.
- So $F(a, b)$ is the area under the curves $X = a$ and $Y = b$

Probabilities using Joint CDFs

- The above implies that if both X and Y are defined down to $(-\infty, -\infty)$, then we are finding the area for a semi-infinite rectangle in the (X, Y) plane
  - i.e. the probability that the point (X, Y) is contained in this rectangle
- Suppose instead we are told that both X and Y are bounded below by 0 (i.e. X and Y takes values from $[0, \infty)$).
- Then our rectangle representing $P(X \leq a, Y \leq b)$ is cut off at $X = 0$ and $Y = 0$
- My probability is now found as the area of the rectangle with corners (0,0), (a, 0), (0, b) and (a, b)

Probabilities using Joint CDFs

- What we have done by stopping our rectangle at $X = 0$ and $Y = 0$ is to draw two more lines, with my rectangle now defined by:
  - vertical lines $X = 0$ and $X = a$
  - horizontal lines $Y = 0$ and $Y = b$
- The area of this new rectangle can now be represented in terms of probabilities for each combination of X and Y lines:
  - $P(X \leq a, Y \leq b)$
  - $P(X \leq a, Y \leq 0)$
  - $P(X \leq 0, Y \leq b)$
  - $P(X \leq 0, Y \leq 0)$

Probabilities using Joint CDFs

- What we get from the above probabilities is the expression for finding the probability that X takes values in $[0, a]$ and Y takes values in $[0, b]$:

$$P(0 \leq X \leq a, 0 \leq Y \leq b)$$

- We can express this in terms of the probabilities on the previous slide as

$$\begin{aligned} P(X \leq a, Y \leq b) &- P(X \leq a, Y \leq 0) \\ &- P(X \leq 0, Y \leq b) + P(X \leq 0, Y \leq 0) \end{aligned}$$

where we add back the last term because it was unshaded twice.

Probabilities using Joint CDFs

- But these probabilities can be written in terms of the CDF, giving us

$$F(a, b) - F(a, 0) - F(0, b) + F(0, 0)$$

- We can actually view this as finding the area under the curve between 0 and a for two different values of Y:

- When $Y \leq b$, we can find the area below Y when $0 \leq X \leq a$: $F(a, b) - F(0, b)$
- When $Y \leq 0$, we can find the area below Y when $0 \leq X \leq a$: $F(a, 0) - F(0, 0)$
- By subtracting these, I get the area of my rectangle.

Outline

Discrete Joint Distributions

- Before we can do much else with these joint CDFs, we need to introduce the corresponding probability mass/density function for jointly distributed random variables.
- We first being with the case of jointly distributed **discrete** random variables.
- Suppose that X and Y are both discrete random variables, defined on the same sample space.
- If X takes on values $x_1, x_2, \ldots$ and Y takes values $y_1, y_2, \ldots$, their joint probability mass function $p(x, y)$ is

$$p(x_i, y_j) = P(X = x_i, Y = y_j)$$

Example: Joint Probability Mass Function

- Suppose we toss a fair coin 3 times.
  - X denotes the number of heads on the first toss
  - Y denotes the total number of heads.
  - These are defined on the same sample space

$$\Omega = \{hhh, hht, hth, htt, thh, tht, tth, ttt\}$$

- We can represent the joint probabilities that $(X = x_i, Y = y_j)$ using a contingency table:

	y			
x	0	1	2	3
0	$1/8$	$2/8$	$1/8$	0
1	0	$1/8$	$2/8$	$1/8$

- We can use this table to find e.g. probability of a total of 2 heads and no head on the first toss, $P(X = 0, Y = 2) = 1/8$

Example: Grocery Store Checkout

A supermarket has 3 checkout counters. Two customers arrive at different times when the counters are not serving any customers. They each pick a counter at random. Let X denote the number of customers who choose counter 1, and Y the number of customers who choose counter 2. What is the joint PMF of X and Y?

- We start by considering the pairs (i, j), denoting the counter choice i for customer 1 and choice j for customer 2.
- We next determine the sample space for these pairs:

  $$\Omega = \{(1,1),(1,2),(1,3),(2,1),(2,2),(2,3),(3,1),(3,2),(3,3)\}$$

- Since the customers select a counter at random and independently, these outcomes are equiprobable, so each has probability $1/9$
- The possible values for both X and Y are $\{0, 1, 2\}$

Example: Grocery Store Checkout (cont.)

- To determine the joint PMF of X and Y, we need to compute probabilities of the form $P(X = x_i, Y = Y_j)$
  - e.g. the probability 2 customers chose counter 1 and 0 customers chose counter 2 is

$$P(X = 2, Y = 0) = P(\{1, 1\}) = 1/9$$

  - e.g. the probability 1 customer chose counter 1 and 1 chose counter 2 is

$$P(X = 1, Y = 1) = P(\{1, 2\} \cup \{2, 1\}) = 2/9$$

- So the joint PMF would be

	y		
x	0	1	2
0	1/9	2/9	1/9
1	2/9	2/9	0
2	1/9	0	0

Joint Probability Mass Function

- Since the joint PMF is still a probability mass function, it must satisfy the same criteria to be a valid probability function.
  1. $p(x, y) \geq 0$ for all x, y
  2. $\sum_{x,y} p(x, y) = 1$ where the sum is over all values (x, y) that are assigned non-zero probability.
- Just as before, we can write the joint CDF as the sum over values of the joint PMF:

$$F(x, y) = \sum_{x_i \leq x} \sum_{y_j \leq y} p(x_i, y_j)$$

where we need to sum over values below x and below y.

Example: Grocery Store Checkout

Based on the joint PMF below, what is the probability that there were at most 1 customer at counter 1 and at most 1 customer at counter 2?

	y		
x	0	1	2
0	$1/9$	$2/9$	$1/9$
1	$2/9$	$2/9$	0
2	$1/9$	0	0

- This translates to finding $P(X \leq 1, Y \leq 1) = F(1,1)$
- Now just add up the respective PMF probabilities for $X \leq 1$ and $Y \leq 1$:

$$F(1,1) = P(X = 1, Y = 1) + P(X = 1, Y = 0) \\ + P(X = 0, Y = 1) + P(X = 0, Y = 0) \\ = 2/9 + 2/9 + 2/9 + 1/9 = 7/9$$

Exercise - Give it a try!

Suppose we flip 3 fair coins, and denote X as the number of heads of the first flip, and Y as the total number of heads. What is the probability that we have at least 2 total heads when the first flip was a tail?

	y			
x	0	1	2	3
0	$1/8$	$2/8$	$1/8$	0
1	0	$1/8$	$2/8$	$1/8$

Marginal Probability Mass Functions

- Suppose we are now interested in looking only at the probability function for one of our jointly distributed variables.
- We can do this by finding the marginal probability function of a single random variable.
- The marginal PMFs for jointly distributed X and Y are

$$p_X(x) = \sum_i p(x, y_i) \text{ and } p_Y(y) = \sum_i p(x_i, y)$$

- That is, we sum the joint PMF over all values of the random variable for which we do not want the marginal PMF

Example: Grocery Store Checkout

Based on the joint PMF below, find the marginal PMF for both X and Y.

	y		
x	0	1	2
0	$1/9$	$2/9$	$1/9$
1	$2/9$	$2/9$	0
2	$1/9$	0	0

- We want the marginal PMF of X, so for each value of X we must sum across the columns of Y:

$$p_X(0) = \sum_{y=0}^{2} p(0, y) = 1/9 + 2/9 + 1/9 = 4/9$$

$$p_X(1) = \sum_{y=0}^{2} p(1, y) = 2/9 + 2/9 + 0 = 4/9$$

$$p_X(2) = \sum_{y=0}^{2} p(2, y) = 1/9 + 0 + 0 = 1/9$$

Example: Grocery Store Checkout (cont.)

Based on the joint PMF below, find the marginal PMF for both X and Y.

	y		
x	0	1	2
0	$1/9$	$2/9$	$1/9$
1	$2/9$	$2/9$	0
2	$1/9$	0	0

- For the marginal PMF of Y we must sum across the rows of X for each value of Y:

$$p_Y(0) = \sum_{x=0}^{2} p(x,0) = 1/9 + 2/9 + 1/9 = 4/9$$

$$p_Y(1) = \sum_{x=0}^{2} p(x,1) = 2/9 + 2/9 + 0 = 4/9$$

$$p_Y(2) = \sum_{x=0}^{2} p(x,2) = 1/9 + 0 + 0 = 1/9$$

Generalized Marginal PMF

- Of course, we are able to extend the concept of the joint PMF to the case where we have more than two jointly distributed discrete variables.

$$p(x_1, \ldots, x_m) = P(X_1 = x_1, \ldots, X_m = x_m)$$

  as long as $X_1, \ldots, X_m$ are defined on the same sample space.

- Then we can also define the marginal PMF for any one of these m random variables as

$$p_{X_i}(x_i) = \sum_{j: x_j \neq x_i} p(x_1, \ldots, x_m)$$

Exercise - Give it a try!

Suppose we flip 3 fair coins, and denote X as the number of heads of the first flip, and Y as the total number of heads. Find the marginal distributions of X and Y.

	y			
x	0	1	2	3
0	$1/8$	$2/8$	$1/8$	0
1	0	$1/8$	$2/8$	$1/8$

Outline

Continuous Joint Distributions

- Just as when we were dealing with a single continuous random variable, we can really only talk about the joint PDF in terms of how it relates to the joint CDF.
- Suppose that X and Y are continuous random variables with a joint CDF $F(x, y)$
- Then their joint PDF is a piecewise continuous function of X and Y, $f(x, y)$, which
  - must be non-negative, and
  - $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)dydx = 1$
- So we can define the joint CDF of X and Y as

$$F(x, y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(u, v)dvdu$$

Continuous Joint Distributions

- So like the univariate case, we can express the joint CDF as the area under the curve up to the values $(X, Y) = (x, y)$
- Again, we have that anywhere the discrete case used a summation, the continuous case uses an integral.
- From the fundamental theorem of multivariable calculus, we have the same relationship between the CDF and PDF as in the univariate case:

$$f(x, y) = \frac{\partial^2}{\partial x \partial y} F(x, y)$$

  where we now have a second order partial derivative to represent the two random variables

Probabilities Using Joint PDFs

- Finding probabilities involving the joint PDF can be tricky sometimes.
- Because these are joint probabilities, often the question is asking about how one variable can relate to another.
- Therefore it is very important to know what values X and Y are defined on, and how X and Y relate to each other.
- Further, it is often very helpful to sketch out exactly what are we are integrating over (area of integration).
  - This will help you decide what the bound of your integrals should be in order to arrive at the correct area.

Example 1: Radioactive Particle

Suppose a radioactive particle is randomly located in a square with sides of length 1. Let X and Y denote the coordinates of the particle's location in this square. A reasonable model for the location of the particle is

$$f(x,y) = \begin{cases} 1, & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

Find $F(0.2, 0.4)$.

- First we should sketch the area of integration for this problem.

Example 1: Radioactive Particle (cont.)

- Now I need to integrate over this region of the joint PDF given on the previous slide:

Example 1: Radioactive Particle (cont.)

- Now suppose instead I want $P(0.1 \leq X \leq 0.3, 0 \leq Y \leq 0.5)$
- I can sketch the new area of integration for this probability:

Example 2: A Messier Joint Problem

Consider the joint PDF

$$f(x,y) = \frac{12}{7}(x^2 + xy),\ 0 \leq x \leq 1, 0 \leq y \leq 1.$$

Find the probability $P(X > Y)$.

- Here our joint PDF is not as simple as before, and we are trying to find the area under the joint PDF where $X > Y$.
- Let's again start by sketching the area of integration, where we now have the added information that we want $0 \leq y \leq x \leq 1$

Example 2: A Messier Joint Problem (cont.)

- In the other example, it didn't really matter whether we first integrate X or Y.
- Here, we can make our job easier by first integrating Y, followed by X.
  - We know that $0 \leq y \leq x \leq 1$, based on the probability we are looking for.
  - So y is between $(0, x)$ and x is between $(y, 1)$.
  - I end up with a simpler integral if I start by integrating y between $(0, x)$,

Example 2: A Messier Joint Problem (cont.)

- So I can now write out the integral corresponding to this area:

$$P(X > Y) = \frac{12}{7}\int_0^1 \int_0^x (x^2 + xy)dydx$$

- Start by integrating out y:

$$P(X > Y) = \frac{12}{7}\int_0^1 \left[x^2y + \frac{xy^2}{2}\right]_0^x dx = \frac{12}{7}\int_0^1 (x^3 + \frac{x^3}{2})dx$$

- Now finish by integrating the x's

$$P(X > Y) = \frac{12}{7}\left[\frac{x^4}{4} + \frac{x^4}{8}\right]_0^1 = \frac{9}{14}$$

Exercise - Give it a try!

A gas station receives a delivery of gas every week. Let X be the proportion of gasoline in the holding tank after delivery. Let Y be the proportion of gasoline in the holding tank that has been sold that week. We thus have that $Y \leq X$. If the joint PDF is $f(x, y) = 3x$, if $0 \leq y \leq x \leq 1$, find the probability that in a week we started with less than $1/2$ a tank and sell more than $1/4$.

Marginal PDFs

- Analogous to the discrete case, we can find the marginal PDF of either X or Y by integrating over the values of the random variable we are not interested in.
- Therefore the marginal PDF of X would be

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy$$

and the marginal PDF of Y would be

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx$$

Example 2 revisited

- Recall that the joint PDF was

$$f(x,y) = \frac{12}{7}(x^2 + xy),\ 0 \le x \le 1, 0 \le y \le 1.$$

- To find the marginal PDF of X, we just integrate out Y from the joint PDF:

$$f_X(x) = \frac{12}{7}\int_0^1 (x^2 + xy)dy = \left[x^2 y + \frac{xy^2}{2}\right]_0^1 = \frac{12}{7}\left(x^2 + \frac{x}{2}\right)$$

- To find the marginal ODF of Y, integrate out X:

$$f_Y(y) = \frac{12}{7}\int_0^1 (x^2 + xy)dx = \left[\frac{x^3}{3} + \frac{x^2 y}{2}\right]_0^1 = \frac{12}{7}\left(\frac{1}{3} + \frac{y}{2}\right)$$

Example: Exercise revisited

- Recall that the gasoline tank exercise had joint PDF

$$f(x, y) = 3x, \text{ if } 0 \leq y \leq x \leq 1$$

- We can still find the marginal PDFs of X and Y but here we need to be careful of how X and Y are defined relative to each other.
  - Here $0 \leq y \leq x$ and $y \leq x \leq 1$ so this must be reflected in our integrals.
  - This means, to integrate over x values, we use the bounds y and 1
  - and to integrate over y values, we use the bounds 0 and x

Example: Exercise revisited (cont.)

- Let's first find the marginal PDF of Y, where x takes values on $(y, 1)$:

$$f_Y(y) = \int_y^1 3x dx = \left[\frac{3x^2}{2}\right]_y^1 = \frac{3}{2} - \frac{3}{2}y^2, \text{ if } 0 \leq y \leq 1$$

- To get the marginal PDF of X, integrate out y which takes values on $(0, x)$:

$$f_X(x) = \int_0^x 3x dy = [3xy]_0^x = 3x^2, \text{ if } 0 \leq x \leq 1$$

Exercise - Give it a try!

Suppose we have the following joint PDF

$$f(x, y) = \begin{cases} \lambda^2 e^{-\lambda y}, & 0 \le x \le y \\ 0, & \text{otherwise} \end{cases}$$

where $\lambda > 0$. Find the marginal PDF for X and Y.

Outline

Independent Random Variables

- We have talked about independence in a number of different cases:
  - we saw that events A and B are independent (week 2) when
    $$P(A \cap B) = P(A)P(B)$$
  - we saw that random variables X and Y are independent (week 3) when
    $$P(X = x \cap Y = y) = P(X = x)P(Y = y)$$
- We can now show how we can determine if two jointly distributed random variables X and Y are independent using their joint PMF/PDF.

Independent Random Variables

Definition

Random variables $X_1, X_2, \ldots, X_n$ are said to be independent if their joint CDF factors into the product of their marginal CDFs:

$$F(x_1, x_2, \ldots, x_n) = F_{X_1}(x_1)F_{X_2}(x_2)\cdots F_{X_n}(x_n)$$

for all $x_1, x_2, \ldots, x_n$.

- This holds for both discrete and continuous random variables.
- It is also equivalent to say variables are independent when the joint PDF/PMF factors into the produce of the marginal PDFs/PMFs

Independent Random Variables

- We can see this easily by considering two jointly continuous RVs.
- If they are independent then $F(x, y) = F(x)F(y)$
- In order to move from CDF to PDF, I take the second mixed partial derivative of $F(x, y)$:

$$\frac{\partial^2}{\partial x \partial y}\left[F(x)F(y)\right]$$

- But since this is the same as taking the derivative w.r.t. y and then the derivative w.r.t x, I can rewrite this as

$$\frac{\partial}{\partial x}\left[\frac{\partial}{\partial y}F(x)F(y)\right] = \frac{\partial}{\partial x}\left[F(x)f(y)\right] = f(x)f(y)$$

Independent Random Variables

- We can also see that the CDF factors if the PDF factors.
- Consider the joint CDF of two independent random variables:

$$F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_X(u) f_Y(v) dv du$$

- If I do the first integration, I see that

$$F(x, y) = \int_{-\infty}^{x} f_X(u) \left[\int_{-\infty}^{y} f_Y(v) dv \right] du = \int_{-\infty}^{x} f_X(u) F_Y(y) du$$

- And then doing the second integration gives me

$$F(x, y) = F_Y(y) \int_{-\infty}^{x} f_X(u) du = F_X(x) F_Y(y)$$

- So you can check the factorization of either the CDF or PMF/PDF to check independence.

Example 2 revisited... again

- We had the following joint PDF for X and Y:

$$f(x,y) = \frac{12}{7}(x^2 + xy), \; 0 \leq x \leq 1, 0 \leq y \leq 1.$$

- We also found the marginal PDFs for both X and Y as

$$f_X(x) = \frac{12}{7}\left(x^2 + \frac{x}{2}\right), \text{ if } 0 \leq x \leq 1$$

and

$$f_Y(y) = \frac{12}{7}\left(\frac{1}{3} + \frac{y}{2}\right), \text{ if } 0 \leq y \leq 1$$

- Are X and Y independent?

Example 2 revisited... again

- We can check by seeing of the product of the marginal PDFs gives us the joint PDF:

$$f_Y(y)f_X(x) = \frac{12}{7}\left(\frac{1}{3} + \frac{y}{2}\right) \times \frac{12}{7}\left(x^2 + \frac{x}{2}\right)$$

- Through some simplification, we get

$$f_Y(y)f_X(x) = \frac{12}{49}x(1+2x)(2+3y) \neq \frac{12}{7}(x^2 + xy)$$

- So we find that X and Y are not independent.

Exercise - Give it a try!

Suppose a radioactive particle is randomly located in a square with sides of length 1. Let X and Y denote the coordinates of the particle's location in this square. The joint PDF is

$$f(x, y) = \begin{cases} 1, & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

Are X and Y independent?

Outline

Recall: Moment Generating Functions

Definition of Moment Generating Function

The MGF of a random variable X is $M(t) = E\left(e^{tX}\right)$ if the expectation is defined. In the discrete case,

$$M(t) = E\left(e^{tX}\right) = \sum_x e^{tx} p(x)$$

and in the continuous case,

$$M(t) = E\left(e^{tX}\right) = \int_{\infty}^{\infty} e^{tx} f(x) dx$$

- We can use the same idea to talk about the MGF for independent variables X and Y

MGFs of Independent Random Variables

MGFs of Independent RVs

If X and Y are independent random variables with MGFs M_X and M_Y, and $Z = X + Y$, then $M_Z(t) = M_X(t)M_Y(t)$ on the common interval where both MGFs exist.

Proof:

$\square$

Example: Gamma Distribution

Suppose that $X_1 \sim Gamma(\alpha_1, \lambda)$ and $X_2 \sim Gamma(\alpha_2, \lambda)$ are independent random variables. Show that the MGF of $X_1 + X_2$ is also a Gamma MGF.

- We can use the previous result to get the MGF for the sum of 2 Gamma random variables
- The MGF for X_i, $i = 1, 2$ is $M_{X_i}(t) = \left(\frac{\lambda}{\lambda - t}\right)^{\alpha_i}$, when $t < \lambda$
- Since X_1 and X_2 are independent, we can write

$$M_{X_1+X_2}(t) = \left(\frac{\lambda}{\lambda - t}\right)^{\alpha_1}\left(\frac{\lambda}{\lambda - t}\right)^{\alpha_2} = \left(\frac{\lambda}{\lambda - t}\right)^{\alpha_1+\alpha_2}$$

- We can recognize this as the MGF of a $Gamma(\alpha_1 + \alpha_2, \lambda)$ because MGFs are unique

Example: Gamma Distribution (cont.)

Now suppose that $Y_1 \sim Exp(\lambda)$ and $Y_2 \sim Exp(\lambda)$ are independent random variables. Show that the MGF of $Y_1 + Y_2$ is equivalent to the MGF of a Gamma.

- It can be shown that the MGF for an Exponential random variable is $M_{Y_i}(t) = \frac{\lambda}{\lambda - t}$ when $t < \lambda$
- Again we use the previous result to find the MGF of $Y_1 + Y_2$:

$$M_{Y_1+Y_2}(t) = \frac{\lambda}{\lambda - t} \times \frac{\lambda}{\lambda - t} = \left(\frac{\lambda}{\lambda - t}\right)^2$$

- If we compare this with the result on the previous slide, we see that if $\alpha_1 = \alpha_2 = 1$, then $M_{X_1+X_2}(t) = M_{Y_1+Y_2}(t)$
- Thus we have shown that, when α is an integer, the Gamma distribution is just a sum of independent and identically distributed Exponentials.

Joint Moment Generating Functions

- Random variables are often not independent.
- We can also define the MGF for two jointly distributed variables X and Y.
- If X and Y have a joint distribution then

$$M_{XY}(s,t) = E\left(e^{sX+tY}\right)$$

- We can work with this MGF as we have done before:
  - Take the first partial derivative and set $t = s = 0$ to get $E(XY)$
- Can also get the marginal MGFs by e.g. $M_X(s) = M_{XY}(s,0)$