Topic 02: Transformations

Ethan P. Marzban University of California, Santa Barbara PSTAT 120B

Outline

1. Univariate Transformations
2. Method of Distribution Functions (CDF Method)
3. Method of Transformations (Change of Variable Formula)
4. Method of Moment-Generating Functions (MGF Method)
5. Multivariate Transformations

Leadup

- Recall, from PSTAT 120A, that given an appropriate probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we can think of a **random variable** X as a mapping:

$$X : \Omega \to \mathbb{R}$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Leadup

- Recall, from PSTAT 120A, that given an appropriate probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we can think of a **random variable** X as a mapping:

$$X : \Omega \to \mathbb{R}$$

- Additionally, recall the following fact from precalculus: given a mapping $f_1 : A \to B$ and another mapping $f_2 : B \to C$, then $(f_2 \circ f_1) : A \to C$.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Leadup

- Recall, from PSTAT 120A, that given an appropriate probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we can think of a **random variable** X as a mapping:

$$X : \Omega \to \mathbb{R}$$

- Additionally, recall the following fact from precalculus: given a mapping $f_1 : A \to B$ and another mapping $f_2 : B \to C$, then $(f_2 \circ f_1) : A \to C$.
- This means, given a function $g : \mathbb{R} \to \mathbb{R}$ and a random variable $X : \Omega \to \mathbb{R}$, we have $(g \circ X) : \Omega \to \mathbb{R}$.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Leadup

- In this way, we can think of $(g \circ X)$ as a random variable itself!

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Leadup

- In this way, we can think of $(g \circ X)$ as a random variable itself!
  - For example, given a random variable X, then the quantity X^2 will also be a random variable.

Leadup

- In this way, we can think of $(g \circ X)$ as a random variable itself!
  - For example, given a random variable X, then the quantity X^2 will also be a random variable.
- Another way of saying this: functions of random variables are themselves random variables.

Leadup

- In this way, we can think of $(g \circ X)$ as a random variable itself!
  - For example, given a random variable X, then the quantity X^2 will also be a random variable.
- Another way of saying this: functions of random variables are themselves random variables.
- "Functions of random variables?" That sounds awfully abstract...

Leadup

- In this way, we can think of $(g \circ X)$ as a random variable itself!
    - For example, given a random variable X, then the quantity X^2 will also be a random variable.
- Another way of saying this: functions of random variables are themselves random variables.
- "Functions of random variables?" That sounds awfully abstract...
- But, if we think about it a bit more, this isn't as abstract as it may seem!

Leadup

- For example, let H_I denote the height of a randomly-selected individual as measured in inches, and suppose $H_I \sim \mathcal{N}(70, 2)$.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Leadup

- For example, let H_I denote the height of a randomly-selected individual as measured in inches, and suppose $H_I \sim \mathcal{N}(70, 2)$.
- Let H_C denote the height of a randomly-selected individual as measured in centimeters.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Leadup

- For example, let H_I denote the height of a randomly-selected individual as measured in inches, and suppose $H_I \sim \mathcal{N}(70, 2)$.
- Let H_C denote the height of a randomly-selected individual as measured in <u>centimeters</u>.
- Clearly, the random variables H_I and H_C are related: specifically, $H_C = g(H_I)$ where $g(t) = 2.54 * t$ [since this is the conversion formula between inches and centimeters].

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Leadup

- For example, let H_I denote the height of a randomly-selected individual as measured in inches, and suppose $H_I \sim \mathcal{N}(70, 2)$.
- Let H_C denote the height of a randomly-selected individual as measured in <u>centimeters</u>.
- Clearly, the random variables H_I and H_C are related: specifically, $H_C = g(H_I)$ where $g(t) = 2.54 * t$ [since this is the conversion formula between inches and centimeters].
    - So, **<u>unit conversion</u>** is a fairly simple example of one way transformations (i.e. taking functions of random variables) can be useful.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Leadup

- Transformations can also be used to **summarize** data.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Leadup

- Transformations can also be used to <u>**summarize**</u> data.
- For example, consider a sequence $\{X_i\}_{i=1}^{n} := X_1, \cdots, X_n$ of random variables.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Leadup

- Transformations can also be used to **summarize** data.
- For example, consider a sequence $\{X_i\}_{i=1}^n := X_1, \cdots, X_n$ of random variables.
  - By the way, I'll be using this notation a lot: $\{X_i\}_{i=1}^n$ is a shorthand for $X_1, \cdots, X_n$.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Leadup

- Transformations can also be used to **summarize** data.
- For example, consider a sequence $\{X_i\}_{i=1}^{n} := X_1, \cdots, X_n$ of random variables.
  - By the way, I'll be using this notation a lot: $\{X_i\}_{i=1}^{n}$ is a shorthand for $X_1, \cdots, X_n$.
- The **sample mean** $\overline{X}_n := n^{-1} \sum_{i=1}^{n} X_i$ [which you hopefully saw in PSTAT 120A!] is actually a *function* of the original sequence of random variables, and is hence an example of a transformation.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Leadup

- Now, these two examples indicate that there are perhaps two sub-cases to consider: transformations of *single* random variables, and transformations of *multiple* random variables.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Leadup

- Now, these two examples indicate that there are perhaps two sub-cases to consider: transformations of *single* random variables, and transformations of *multiple* random variables.
  - We often refer to a transformation of a single random variable as a **univariate transformation**, and a transformation of multiple random variables as a **multivariate transformation**.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Leadup

- Now, these two examples indicate that there are perhaps two sub-cases to consider: transformations of *single* random variables, and transformations of *multiple* random variables.
  - We often refer to a transformation of a single random variable as a **univariate transformation**, and a transformation of multiple random variables as a **multivariate transformation**.

- For simplicity's sake, let's start off with univariate transformations.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Leadup

- Now, these two examples indicate that there are perhaps two sub-cases to consider: transformations of *single* random variables, and transformations of *multiple* random variables.
  - We often refer to a transformation of a single random variable as a **univariate transformation**, and a transformation of multiple random variables as a **multivariate transformation**.

- For simplicity's sake, let's start off with univariate transformations.
  - Specifically, given a random variable Y and a function $g : \mathbb{R} \to \mathbb{R}$, we will seek to explore properties of the random variable $U := g(Y)$.

Univariate Transformations

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Goal

Goal

Given a random variable Y and a function $g()$, we seek to describe the random variable $U := g(Y)$.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Goal

Goal

Given a random variable Y and a function $g()$, we seek to describe the random variable $U := g(Y)$.

- What do we mean by "describe" the random variable U?

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Goal

Goal

Given a random variable Y and a function $g()$, we seek to describe the random variable $U := g(Y)$.

- What do we mean by "describe" the random variable U?
- Well, there are a couple of things we could seek to do. First, we could try to compute $\mathbb{E}[U]$.

UC SANTA BARBARA
Department of Statistics and Applied Probability

LOTUS

- It turns out... we've already done that!

LOTUS

- It turns out... we've already done that!
- Specifically, since $U := g(Y)$, we have that $\mathbb{E}[U] = \mathbb{E}[g(Y)]$.

LOTUS

- It turns out... we've already done that!
- Specifically, since $U := g(Y)$, we have that $\mathbb{E}[U] = \mathbb{E}[g(Y)]$.
- The **<u>Law of the Unconscious Statistician</u>** (LOTUS), which we saw in PSTAT 120A, tells us

$$\mathbb{E}[g(Y)] = \int_{\mathbb{R}} g(y) f_Y(y) \, \mathrm{d}y$$

UC SANTA BARBARA
Department of Statistics and Applied Probability

LOTUS

- It turns out... we've already done that!
- Specifically, since $U := g(Y)$, we have that $\mathbb{E}[U] = \mathbb{E}[g(Y)]$.
- The **Law of the Unconscious Statistician** (LOTUS), which we saw in PSTAT 120A, tells us

$$\mathbb{E}[g(Y)] = \int_{\mathbb{R}} g(y) f_Y(y)\ \mathrm{d}y$$

- Similar considerations will allow us to compute $\mathrm{Var}(U)$.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Distributions

- Okay, that's useful! But it's not the whole picture.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Distributions

- Okay, that's useful! But it's not the whole picture.
- Why don't we get a little more ambitious, and seek to find the *distribution* of U?

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Distributions

- Okay, that's useful! But it's not the whole picture.
- Why don't we get a little more ambitious, and seek to find the *distribution* of U?
- First, let me be a little more clear about what I mean by "distribution".

UC SANTA BARBARA
Department of Statistics and Applied Probability

Distributions

- Okay, that's useful! But it's not the whole picture.
- Why don't we get a little more ambitious, and seek to find the *distribution* of U?
- First, let me be a little more clear about what I mean by "distribution".
- Sometimes, we can identify a distribution by name (e.g. "Exponential distribution with parameter $\theta = 0.5$", or "Standard normal distribution").

UC SANTA BARBARA
Department of Statistics and Applied Probability

Distributions

- Okay, that's useful! But it's not the whole picture.
- Why don't we get a little more ambitious, and seek to find the *distribution* of U?
- First, let me be a little more clear about what I mean by "distribution".
- Sometimes, we can identify a distribution by name (e.g. "Exponential distribution with parameter $\theta = 0.5$", or "Standard normal distribution").
- But, a distribution could just as easily have been described by any of the following:

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Distributions

- Okay, that's useful! But it's not the whole picture.
- Why don't we get a little more ambitious, and seek to find the *distribution* of U?
- First, let me be a little more clear about what I mean by "distribution".
- Sometimes, we can identify a distribution by name (e.g. "Exponential distribution with parameter $\theta = 0.5$", or "Standard normal distribution").
- But, a distribution could just as easily have been described by any of the following:
  - Its **distribution function** (i.e. CDF)

UC SANTA BARBARA
Department of Statistics and Applied Probability

Distributions

- Okay, that's useful! But it's not the whole picture.
- Why don't we get a little more ambitious, and seek to find the *distribution* of U?
- First, let me be a little more clear about what I mean by "distribution".
- Sometimes, we can identify a distribution by name (e.g. "Exponential distribution with parameter $\theta = 0.5$", or "Standard normal distribution").
- But, a distribution could just as easily have been described by any of the following:
  - Its **distribution function** (i.e. CDF)
  - Its **density function** (PDF)

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Distributions

- Okay, that's useful! But it's not the whole picture.
- Why don't we get a little more ambitious, and seek to find the *distribution* of *U*?
- First, let me be a little more clear about what I mean by "distribution".
- Sometimes, we can identify a distribution by name (e.g. "Exponential distribution with parameter $\theta = 0.5$", or "Standard normal distribution").
- But, a distribution could just as easily have been described by any of the following:
  - Its **distribution function** (i.e. CDF)
  - Its **density function** (PDF)
  - Its **MGF** (moment-generating function)

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Distributions

- For example, suppose I tell you the random variable W has density function given by
$$f_W(w) = 2e^{-2w} \cdot \mathbb{1}_{\{w \geq 0\}}$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Distributions

- For example, suppose I tell you the random variable W has density function given by
$$f_W(w) = 2e^{-2w} \cdot \mathbb{1}_{\{w \geq 0\}}$$
- You would immediately be able to tell me "oh, W follows the Exponential distribution with parameter $\theta = 1/2$."

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Distributions

- For example, suppose I tell you the random variable W has density function given by

$$f_W(w) = 2e^{-2w} \cdot \mathbb{1}_{\{w \geq 0\}}$$

- You would immediately be able to tell me "oh, W follows the Exponential distribution with parameter $\theta = 1/2$."
- This would, in turn, automatically tell you that W has distribution function

$$F_W(w) = \begin{cases} 1 - e^{-2w} & \text{if } w \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

and MGF

$$M_W(t) = \begin{cases} (1 - t/2)^{-1} & \text{if } t < 1/2 \\ \infty & \text{otherwise} \end{cases}$$

UC SANTA BARBARA
Department of Statistics and Applied Probability

Distributions

- Similarly, if I tell you that the random variable T has MGF given by

$$M_X(t) = \exp\left\{2t + \frac{1}{2}t^2\right\}$$

Distributions

- Similarly, if I tell you that the random variable T has MGF given by

$$M_X(t) = \exp\left\{2t + \frac{1}{2}t^2\right\}$$

you would immediately be able to say

$$f_X(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(x-2)^2\right\}$$

and

$$F_X(x) = \Phi\left(x - 2\right); \qquad \Phi(x) := \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \, \mathrm{d}z$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Distributions

- Now, what if we have a random variable X whose density is given by

$$f_X(x) = \cos(x) \cdot \mathbb{1}_{\{0 \leq x \leq \pi/2\}}$$

Distributions

- Now, what if we have a random variable X whose density is given by

$$f_X(x) = \cos(x) \cdot \mathbb{1}_{\{0 \leq x \leq \pi/2\}}$$

- What is the distribution of X?

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Distributions

- Now, what if we have a random variable X whose density is given by

$$f_X(x) = \cos(x) \cdot \mathbb{1}_{\{0 \leq x \leq \pi/2\}}$$

- What is the distribution of X?
- Well... it's just the density above!

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Distributions

- Now, what if we have a random variable X whose density is given by

$$f_X(x) = \cos(x) \cdot \mathbb{1}_{\{0 \leq x \leq \pi/2\}}$$

- What is the distribution of X?
- Well... it's just the density above!
- What I mean is this - the distribution of X doesn't have a name, like "Exponential" or "Gamma". But it certainly *has* a distribution!

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Distributions

- Now, what if we have a random variable X whose density is given by

$$f_X(x) = \cos(x) \cdot \mathbb{1}_{\{0 \leq x \leq \pi/2\}}$$

- What is the distribution of X?
- Well... it's just the density above!
- What I mean is this - the distribution of X doesn't have a name, like "Exponential" or "Gamma". But it certainly *has* a distribution!
- All of this is to say: I encourage you to get into the habit of thinking about "distributions" fairly broadly, and thinking of a distribution as either a density function, distribution function, or MGF (or all three).

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Back to Transformations

Goal

Given a random variable Y and a function $g()$, we seek to describe the random variable $U := g(Y)$.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Back to Transformations

Goal

Given a random variable Y and a function $g()$, we seek to describe the random variable $U := g(Y)$.

- Now, our discussion on the previous few slides tells us that there are three approaches to achieving our goal above.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Back to Transformations

Goal

Given a random variable Y and a function $g()$, we seek to describe the random variable $U := g(Y)$.

- Now, our discussion on the previous few slides tells us that there are three approaches to achieving our goal above.
- We could go after the density function of U.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Back to Transformations

Goal

Given a random variable Y and a function $g()$, we seek to describe the random variable $U := g(Y)$.

- Now, our discussion on the previous few slides tells us that there are three approaches to achieving our goal above.
- We could go after the density function of U.
- Or we could go after the distribution function of U.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Back to Transformations

Goal

Given a random variable Y and a function $g()$, we seek to describe the random variable $U := g(Y)$.

- Now, our discussion on the previous few slides tells us that there are three approaches to achieving our goal above.
- We could go after the density function of U.
- Or we could go after the distribution function of U.
- Or we could go after the MGF of U.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Back to Transformations

Goal

Given a random variable Y and a function $g()$, we seek to describe the random variable $U := g(Y)$.

- Now, our discussion on the previous few slides tells us that there are three approaches to achieving our goal above.
- We could go after the density function of U.
- Or we could go after the distribution function of U.
- Or we could go after the MGF of U.
- Indeed, each of these three approaches are what our textbook calls different "methods".

UC SANTA BARBARA
Department of Statistics and Applied Probability

Support

- Before we dive into these three methods, let's talk a bit about **support**.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Support

- Before we dive into these three methods, let's talk a bit about **support**.
- Recall that the support (aka "state space") of a random variable Y is the set of all values that Y maps to: i.e. $S_Y := Y(\Omega)$. Equivalently, it's the set of all values y for which the density $f_Y(y)$ is nonzero.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Support

- Before we dive into these three methods, let's talk a bit about **support**.
- Recall that the support (aka "state space") of a random variable Y is the set of all values that Y maps to: i.e. $S_Y := Y(\Omega)$. Equivalently, it's the set of all values y for which the density $f_Y(y)$ is nonzero.
- Then, given a random variable $U := g(Y)$, we have $S_U = g(S_Y)$.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Support

- Before we dive into these three methods, let's talk a bit about **support**.
- Recall that the support (aka "state space") of a random variable Y is the set of all values that Y maps to: i.e. $S_Y := Y(\Omega)$. Equivalently, it's the set of all values y for which the density $f_Y(y)$ is nonzero.
- Then, given a random variable $U := g(Y)$, we have $S_U = g(S_Y)$.
  - That is, the support of a transformed random variable is the image of the original support under the transformation.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Support

- Before we dive into these three methods, let's talk a bit about **support**.
- Recall that the support (aka "state space") of a random variable Y is the set of all values that Y maps to: i.e. $S_Y := Y(\Omega)$. Equivalently, it's the set of all values y for which the density $f_Y(y)$ is nonzero.
- Then, given a random variable $U := g(Y)$, we have $S_U = g(S_Y)$.
    - That is, the support of a transformed random variable is the image of the original support under the transformation.

- Though this formula seems inoccuous enough, finding the support of a transformed random variable can be trickier than it first appears...

UC SANTA BARBARA
Department of Statistics and Applied Probability

Support

- A simple way I like to think about things is to draw a picture.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Support

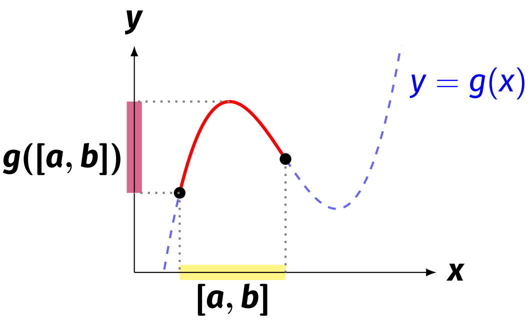
- A simple way I like to think about things is to draw a picture.
- Specifically, let's say we have an interval $[a, b]$ and a transformation $g : \mathbb{R} \to \mathbb{R}$.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Support

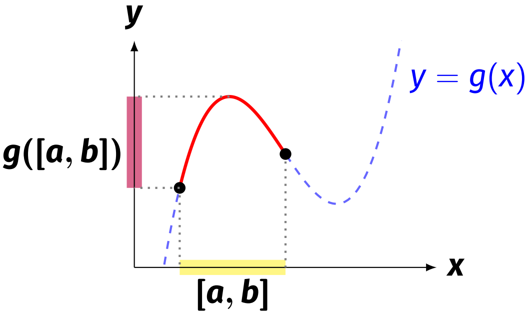
- A simple way I like to think about things is to draw a picture.
- Specifically, let's say we have an interval $[a, b]$ and a transformation $g : \mathbb{R} \to \mathbb{R}$.
- To figure out what $g([a, b])$ looks like, simply graph the function $y = g(x)$, indicate $[a, b]$ on the $x-$axis, and figure out what the corresponding values on the $y-$axis are.

Support



UC SANTA BARBARA
Department of Statistics
and Applied Probability

Support



- Note: in general, $g([a, b]) \neq [g(a), g(b)]$!

Clicker Question!

Clicker Question 1

For $A = [0\ , 6]$ and $g(x) = \cos(\pi x)$, what is the correct expression for $g(A)$?

(A) $[0\ ,\ 1]$ (B) $[0\ ,\ 6]$ (C) $[-1\ , 1]$ (D) $\{0\}$
(E) None of the above

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Clicker Question!

Clicker Question 1

For $A = [0\ , 6]$ and $g(x) = \cos(\pi x)$, what is the correct expression for $g(A)$?

(A) $[0\ ,\ 1]$ (B) $[0\ ,\ 6]$ (C) $[-1\ , 1]$ (D) $\{0\}$
(E) None of the above

Try this On Your Own:

Example

For $A = [-1\ ,\ 1]$ and $g(x) = x^2$, what is the correct expression for $g(A)$?

Method of Distribution Functions (CDF Method)

UC SANTA BARBARA
Department of Statistics
and Applied Probability

CDF Method

- Let's consider the following rephrasing of our goal:

UC SANTA BARBARA
Department of Statistics
and Applied Probability

CDF Method

- Let's consider the following rephrasing of our goal:

Goal

Given a random variable Y and a function $g()$, we seek to derive an expression for $F_U(u) := \mathbb{P}(U \leq u)$, the CDF of U.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

CDF Method

- Let's consider the following rephrasing of our goal:

Goal

Given a random variable Y and a function $g()$, we seek to derive an expression for $F_U(u) := \mathbb{P}(U \leq u)$, the CDF of U.

- As a concrete example, let $Y \sim \text{Exp}(\theta)$ and let $U := cY$ for a positive constant c.

UC SANTA BARBARA
Department of Statistics and Applied Probability

CDF Method

- Let's consider the following rephrasing of our goal:

Goal

Given a random variable Y and a function $g()$, we seek to derive an expression for $F_U(u) := \mathbb{P}(U \leq u)$, the CDF of U.

- As a concrete example, let $Y \sim \text{Exp}(\theta)$ and let $U := cY$ for a positive constant c.
  - If it helps, you can think of this in terms of our inches-to-centimeter conversion example from the start of this lecture: Y can denote the heights in inches and U can denote the heights in cenimeters.

UC SANTA BARBARA
Department of Statistics and Applied Probability

CDF Method

- Now, we know everything we could want to know about Y.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

CDF Method

- Now, we know everything we could want to know about Y.
- Specifically, we have the CDF of Y:

$$F_Y(y) = \begin{cases} 1 - e^{-y/\theta} & \text{if } y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

UC SANTA BARBARA
Department of Statistics and Applied Probability

CDF Method

- Now, we know everything we could want to know about Y.
- Specifically, we have the CDF of Y:

$$F_Y(y) = \begin{cases} 1 - e^{-y/\theta} & \text{if } y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

- So, if we can relate $F_U(u)$ to $F_Y(y)$, we'd be done.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

CDF Method

- Now, we know everything we could want to know about Y.
- Specifically, we have the CDF of Y:

$$F_Y(y) = \begin{cases} 1 - e^{-y/\theta} & \text{if } y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

- So, if we can relate $F_U(u)$ to $F_Y(y)$, we'd be done.
- Note:

$$F_U(u) := \mathbb{P}(U \leq u) = \mathbb{P}(cY \leq u)$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

CDF Method

- Now, we know everything we could want to know about Y.
- Specifically, we have the CDF of Y:

$$F_Y(y) = \begin{cases} 1 - e^{-y/\theta} & \text{if } y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

- So, if we can relate $F_U(u)$ to $F_Y(y)$, we'd be done.
- Note:

$$F_U(u) := \mathbb{P}(U \leq u) = \mathbb{P}(cY \leq u)$$

- Divide through by c:

$$F_U(u) = \mathbb{P}\left(Y \leq \frac{u}{c}\right) = F_Y\left(\frac{u}{c}\right)$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

CDF Method

- So, plugging into our expression for $F_Y(y)$, we have:

$$F_U(u) = F_Y\left(\frac{u}{c}\right)$$

CDF Method

- So, plugging into our expression for $F_Y(y)$, we have:

$$
\begin{aligned}
F_U(u) &= F_Y\left(\frac{u}{c}\right) \\
&= \begin{cases} 1 - e^{(u/c)/\theta} & \text{if } (u/c) \geq 0 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

CDF Method

- So, plugging into our expression for $F_Y(y)$, we have:

$$
\begin{aligned}
F_U(u) &= F_Y\left(\frac{u}{c}\right) \\
&= \begin{cases} 1 - e^{(u/c)/\theta} & \text{if } (u/c) \geq 0 \\ 0 & \text{otherwise} \end{cases} \\
&= \boxed{\begin{cases} 1 - e^{u/(c\theta)} & \text{if } u \geq 0 \\ 0 & \text{otherwise} \end{cases}}
\end{aligned}
$$

UC SANTA BARBARA
Department of Statistics and Applied Probability

CDF Method

- So, plugging into our expression for $F_Y(y)$, we have:

$$
\begin{aligned}
F_U(u) &= F_Y\left(\frac{u}{c}\right) \\
&= \begin{cases} 1 - e^{(u/c)/\theta} & \text{if } (u/c) \geq 0 \\ 0 & \text{otherwise} \end{cases} \\
&= \boxed{\begin{cases} 1 - e^{u/(c\theta)} & \text{if } u \geq 0 \\ 0 & \text{otherwise} \end{cases}}
\end{aligned}
$$

- And we're done! We've accomplished our goal, and found an expression for $F_U(u)$, the CDF of U.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Going Further

- Now, in this particular case, we can take things a step further.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Going Further

- Now, in this particular case, we can take things a step further.
- Specifically, doesn't that CDF look awfully familiar?

UC SANTA BARBARA
Department of Statistics and Applied Probability

Going Further

- Now, in this particular case, we can take things a step further.
- Specifically, doesn't that CDF look awfully familiar?
- Indeed, it is the CDF of the Exp$(c\theta)$ distribution!

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Going Further

- Now, in this particular case, we can take things a step further.
- Specifically, doesn't that CDF look awfully familiar?
- Indeed, it is the CDF of the $\text{Exp}(c\theta)$ distribution!
- So, what we've essentially shown is:

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Going Further

- Now, in this particular case, we can take things a step further.
- Specifically, doesn't that CDF look awfully familiar?
- Indeed, it is the CDF of the $\text{Exp}(c\theta)$ distribution!
- So, what we've essentially shown is:

Theorem (Closure of Exponential Distribution under Multiplication)

Given $Y \sim \text{Exp}(\theta)$ and a positive constant c, then $(cY) \sim \text{Exp}(c\theta)$.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Going Further

- Now, in this particular case, we can take things a step further.
- Specifically, doesn't that CDF look awfully familiar?
- Indeed, it is the CDF of the Exp($c\theta$) distribution!
- So, what we've essentially shown is:

Theorem (Closure of Exponential Distribution under Multiplication)

Given $Y \sim \text{Exp}(\theta)$ and a positive constant c, then $(cY) \sim \text{Exp}(c\theta)$.

- We're going to use this result a **LOT**!

UC SANTA BARBARA
Department of Statistics and Applied Probability

Interpretation

- I know this might seem a little abstract - what does it mean to “multiply the exponential distribution by a constant?”

UC SANTA BARBARA
Department of Statistics and Applied Probability

Interpretation

- I know this might seem a little abstract - what does it mean to "multiply the exponential distribution by a constant?"
- Again, if it helps, you can always think in terms of our inches-to-centimeter problem from the start of these slides.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Interpretation

- I know this might seem a little abstract - what does it mean to "multiply the exponential distribution by a constant?"
- Again, if it helps, you can always think in terms of our inches-to-centimeter problem from the start of these slides.
- If $Y \sim \text{Exp}(\theta)$ denotes the height of a randomly selected person in inches, then the distribution of heights in centimeters will *also* be exponential, this time with mean 2.54θ.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Example

- Let's do another example together.

Example

- Let's do another example together.
- Suppose Y has density function given by

$$f_Y(y) = 2y^2 \cdot \mathbb{1}_{\{0 \le y \le 1\}}$$

and again define $U := cY$ for a positive constant c.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Example

- Let's do another example together.
- Suppose Y has density function given by

$$f_Y(y) = 2y^2 \cdot \mathbb{1}_{\{0 \leq y \leq 1\}}$$

  and again define $U := cY$ for a positive constant c.
- Now, before we got lucky because we immediately knew what the CDF of Y was.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Example

- Let's do another example together.
- Suppose Y has density function given by

$$f_Y(y) = 2y^2 \cdot \mathbb{1}_{\{0 \leq y \leq 1\}}$$

  and again define $U := cY$ for a positive constant c.
- Now, before we got lucky because we immediately knew what the CDF of Y was.
- But, even though we can't *immediately* recognize the CDF of Y in this example, we can still derive it!

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Example

- By definition, for a $y \in [0, 1]$,

$$F_Y(y) = \int_{-\infty}^{y} f_Y(t)\ \mathrm{d}t$$

UC SANTA BARBARA
Department of Statistics and Applied Probability

Example

- By definition, for a $y \in [0, 1]$,

$$
\begin{aligned}
F_Y(y) &= \int_{-\infty}^{y} f_Y(t) \, \mathrm{d}t \\
&= \int_{-\infty}^{y} 2t \cdot \mathbb{1}_{\{0 \leq t \leq 1\}} \, \mathrm{d}t
\end{aligned}
$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Example

- By definition, for a $y \in [0, 1]$,

$$\begin{aligned} F_Y(y) &= \int_{-\infty}^{y} f_Y(t)\, \mathrm{d}t \\ &= \int_{-\infty}^{y} 2t \cdot \mathbb{1}_{\{0 \le t \le 1\}}\, \mathrm{d}t = \int_0^y 2t\, \mathrm{d}t = y^2 \end{aligned}$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Example

- By definition, for a $y \in [0, 1]$,

$$\begin{aligned} F_Y(y) &= \int_{-\infty}^{y} f_Y(t)\,\mathrm{d}t \\ &= \int_{-\infty}^{y} 2t \cdot \mathbb{1}_{\{0 \leq t \leq 1\}}\,\mathrm{d}t = \int_{0}^{y} 2t\,\mathrm{d}t = y^2 \end{aligned}$$

- Clearly, for $y < 0$ we have $F_Y(y) = \mathbb{P}(Y \leq y) = 0$ and for $y > 1$ we have $\mathbb{P}(Y \leq y) = 1$, meaning

$$F_Y(y) = \begin{cases} 0 & \text{if } y < 0 \\ y^2 & \text{if } 0 \leq y < 1 \\ 1 & \text{if } y \geq 1 \end{cases}$$

Example

- And now we're in the same position as before!

$$\mathbb{P}(U \leq u) = \mathbb{P}(cY \leq u) = \mathbb{P}\left(Y \leq \frac{u}{x}\right)$$

Example

- And now we're in the same position as before!

$$\begin{aligned}\mathbb{P}(U \leq u) = \mathbb{P}(cY \leq u) = \mathbb{P}\left(Y \leq \frac{u}{x}\right) \\ = F_Y\left(\frac{u}{c}\right)\end{aligned}$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Example

- And now we're in the same position as before!

$$
\begin{aligned}
\mathbb{P}(U \le u) = \mathbb{P}(cY \le u) &= \mathbb{P}\left(Y \le \frac{u}{X}\right) \\
&= F_Y\left(\frac{u}{c}\right) \\
&= \begin{cases} 0 & \text{if } (u/c) < 0 \\ (u/c)^2 & \text{if } 0 \le (u/c) < 1 \\ 1 & \text{if } (u/c) \ge 1 \end{cases}
\end{aligned}
$$

Example

- And now we're in the same position as before!

$$\mathbb{P}(U \leq u) = \mathbb{P}(cY \leq u) = \mathbb{P}\left(Y \leq \frac{u}{x}\right)$$

$$= F_Y\left(\frac{u}{c}\right)$$

$$= \begin{cases} 0 & \text{if } (u/c) < 0 \\ (u/c)^2 & \text{if } 0 \leq (u/c) < 1 \\ 1 & \text{if } (u/c) \geq 1 \end{cases} = \boxed{\begin{cases} 0 & \text{if } u < 0 \\ u^2/c^2 & \text{if } 0 \leq u < c \\ 1 & \text{if } u \geq c \end{cases}}$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Example

- One more example before we summarize.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Example

- One more example before we summarize.
- Let $Y \sim \mathcal{N}(0, 1)$ and $U := Y^2$.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Example

- One more example before we summarize.
- Let $Y \sim \mathcal{N}(0, 1)$ and $U := Y^2$.
- A quick sketch (see chalkboard) reveals that $S_U = [0, \infty)$. So, $F_U(u) = 0$ whenever $u < 0$.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Example

- One more example before we summarize.
- Let $Y \sim \mathcal{N}(0, 1)$ and $U := Y^2$.
- A quick sketch (see chalkboard) reveals that $S_U = [0, \infty)$. So, $F_U(u) = 0$ whenever $u < 0$.
- Additionally, we (again) have the CDF of Y: $F_Y(y) = \Phi(y)$, where $\Phi(\cdot)$ denotes the standard normal CDF.

Example

- So, let's try and proceed like we did before! For a fixed $u \geq 0$,

$$F_U(u) := \mathbb{P}(U \leq u) = \mathbb{P}(Y^2 \leq u)$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Example

- So, let's try and proceed like we did before! For a fixed $u \geq 0$,

$$F_U(u) := \mathbb{P}(U \leq u) = \mathbb{P}(Y^2 \leq u)$$

- Now, it's tempting to continue this as

$$F_U(u) = \mathbb{P}(Y^2 \leq u) = \mathbb{P}(Y \leq \sqrt{u})$$

This is, however, <u>INCORRECT</u>.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Example

- So, let's try and proceed like we did before! For a fixed $u \geq 0$,

$$F_U(u) := \mathbb{P}(U \leq u) = \mathbb{P}(Y^2 \leq u)$$

- Now, it's tempting to continue this as

$$F_U(u) = \mathbb{P}(Y^2 \leq u) = \mathbb{P}(Y \leq \sqrt{u})$$

  This is, however, <u>INCORRECT</u>.
- Let's understand why.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Example

- There are a couple of ways to understand why the above is incorrect.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Example

- There are a couple of ways to understand why the above is incorrect.
- One is to recall a fact from algebra/precalculus that you might have forgotten: $\sqrt{\cdot}$ means the *principal* square root, and so, for any real number x, we have $\sqrt{x^2} = |x|$.

Example

- There are a couple of ways to understand why the above is incorrect.
- One is to recall a fact from algebra/precalculus that you might have forgotten: $\sqrt{\cdot}$ means the *principal* square root, and so, for any real number x, we have $\sqrt{x^2} = |x|$.
  - Remember, both -3 and 3 have squares equal to 9! But, when we write $\sqrt{9}$, we implicitly mean the principal square root which is why we write $\sqrt{9} = 3$.

Example

- There are a couple of ways to understand why the above is incorrect.
- One is to recall a fact from algebra/precalculus that you might have forgotten: $\sqrt{\cdot}$ means the *principal* square root, and so, for any real number x, we have $\sqrt{x^2} = |x|$.
  - Remember, both -3 and 3 have squares equal to 9! But, when we write $\sqrt{9}$, we implicitly mean the principal square root which is why we write $\sqrt{9} = 3$.
- So, what we really have is:

$$F_U(u) := \mathbb{P}(U \leq u) = \mathbb{P}(Y^2 \leq u)$$

Example

- There are a couple of ways to understand why the above is incorrect.
- One is to recall a fact from algebra/precalculus that you might have forgotten: $\sqrt{\cdot}$ means the *principal* square root, and so, for any real number x, we have $\sqrt{x^2} = |x|$.
  - Remember, both -3 and 3 have squares equal to 9! But, when we write $\sqrt{9}$, we implicitly mean the principal square root which is why we write $\sqrt{9} = 3$.
- So, what we really have is:

$$F_U(u) := \mathbb{P}(U \leq u) = \mathbb{P}(Y^2 \leq u) = \mathbb{P}(|Y| \leq \sqrt{u}) = \mathbb{P}(-\sqrt{u} \leq Y \leq \sqrt{u})$$

Example (cont'd)

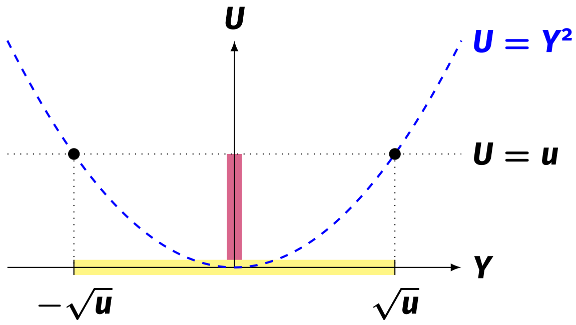
- Now, there's another way to see how to get from $\mathbb{P}(Y^2 \leq u)$ to $\mathbb{P}(-\sqrt{u} \leq Y \leq \sqrt{u})$; one that doens't require us to dig into our memory banks and dredge up something from algebra/precalculus, and instead uses pictures.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Video

https://www.youtube.com/watch?v=HtzqjHfoRbw

Static Image



UC SANTA BARBARA
Department of Statistics
and Applied Probability

Example (cont'd)

- So, let's finish up our example!

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Example (cont'd)

- So, let's finish up our example!

$$F_U(u) = \cdots = \mathbb{P}(-\sqrt{u} \leq Y \leq \sqrt{u})$$

Example (cont'd)

- So, let's finish up our example!

$$
\begin{aligned}
F_U(u) = \cdots &= \mathbb{P}(-\sqrt{u} \leq Y \leq \sqrt{u}) \\
&= F_Y(\sqrt{u}) - F_Y(-\sqrt{u})
\end{aligned}
$$

UC SANTA BARBARA
Department of Statistics and Applied Probability

Example (cont'd)

- So, let's finish up our example!

$$\begin{aligned} F_U(u) &= \cdots = \mathbb{P}(-\sqrt{u} \le Y \le \sqrt{u}) \\ &= F_Y(\sqrt{u}) - F_Y(-\sqrt{u}) = \Phi(\sqrt{u}) - \Phi(-\sqrt{u}) \end{aligned}$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Example (cont'd)

- So, let's finish up our example!

$$
\begin{aligned}
F_U(u) &= \cdots = \mathbb{P}(-\sqrt{u} \leq Y \leq \sqrt{u}) \\
&= F_Y(\sqrt{u}) - F_Y(-\sqrt{u}) = \Phi(\sqrt{u}) - \Phi(-\sqrt{u}) \\
&= \Phi(\sqrt{u}) - [1 - \Phi(\sqrt{u})]
\end{aligned}
$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Example (cont'd)

- So, let's finish up our example!

$$
\begin{aligned}
F_U(u) &= \cdots = \mathbb{P}(-\sqrt{u} \leq Y \leq \sqrt{u}) \\
&= F_Y(\sqrt{u}) - F_Y(-\sqrt{u}) = \Phi(\sqrt{u}) - \Phi(-\sqrt{u}) \\
&= \Phi(\sqrt{u}) - [1 - \Phi(\sqrt{u})] = \boxed{2\Phi(\sqrt{u}) - 1}
\end{aligned}
$$

UC SANTA BARBARA
Department of Statistics and Applied Probability

Example (cont'd)

- So, let's finish up our example!

$$
\begin{aligned}
F_U(u) = \cdots = \mathbb{P}(-\sqrt{u} \leq Y \leq \sqrt{u}) \\
&= F_Y(\sqrt{u}) - F_Y(-\sqrt{u}) = \Phi(\sqrt{u}) - \Phi(-\sqrt{u}) \\
&= \Phi(\sqrt{u}) - [1 - \Phi(\sqrt{u})] = 2\Phi(\sqrt{u}) - 1
\end{aligned}
$$

- That's a bit anticlimactic... Let's differentiate wrt. u and obtain the PDF of U:

UC SANTA BARBARA
Department of Statistics and Applied Probability

Example (cont'd)

$$f_U(u) = \frac{\mathrm{d}}{\mathrm{d}u} F_U(u)$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Example (cont'd)

$$\begin{aligned} f_U(u) &= \frac{\mathrm{d}}{\mathrm{d}u} F_U(u) \\ &= \frac{\mathrm{d}}{\mathrm{d}u} [2\Phi(\sqrt{u}) - 1] \end{aligned}$$

Example (cont'd)

$$\begin{aligned} f_U(u) &= \frac{\mathrm{d}}{\mathrm{d}u} F_U(u) \\ &= \frac{\mathrm{d}}{\mathrm{d}u} [2\Phi(\sqrt{u}) - 1] \\ &= 2 \cdot \frac{1}{2\sqrt{u}} \cdot \phi(\sqrt{u}) = \frac{1}{\sqrt{u}} \phi(\sqrt{u}) \end{aligned}$$

Example (cont'd)

$$\begin{aligned} f_U(u) &= \frac{\mathrm{d}}{\mathrm{d}u} F_U(u) \\ &= \frac{\mathrm{d}}{\mathrm{d}u}[2\Phi(\sqrt{u}) - 1] \\ &= 2 \cdot \frac{1}{2\sqrt{u}} \cdot \phi(\sqrt{u}) = \frac{1}{\sqrt{u}} \phi(\sqrt{u}) \end{aligned}$$

- Let's incorporate the support of U, and simplify:

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Example (cont'd)

$$f_U(u) = \frac{1}{\sqrt{u}} \phi(\sqrt{u}) \cdot \mathbb{1}_{\{u \geq 0\}}$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Example (cont'd)

$$\begin{aligned} f_U(u) &= \frac{1}{\sqrt{u}} \phi(\sqrt{u}) \cdot \mathbb{1}_{\{u \geq 0\}} \\ &= \frac{1}{\sqrt{u}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\sqrt{u})^2} \cdot \mathbb{1}_{\{u \geq 0\}} \end{aligned}$$

Example (cont'd)

$$
\begin{aligned}
f_U(u) &= \frac{1}{\sqrt{u}} \phi(\sqrt{u}) \cdot \mathbb{1}_{\{u \geq 0\}} \\
&= \frac{1}{\sqrt{u}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\sqrt{u})^2} \cdot \mathbb{1}_{\{u \geq 0\}} \\
&= \frac{1}{\sqrt{\pi} \cdot 2^{1/2}} \cdot u^{1/2-1} \cdot e^{-u/2} \cdot \mathbb{1}_{\{u \geq 0\}}
\end{aligned}
$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Example (cont'd)

$$
\begin{aligned}
f_U(u) &= \frac{1}{\sqrt{u}} \phi(\sqrt{u}) \cdot \mathbb{1}_{\{u \geq 0\}} \\
&= \frac{1}{\sqrt{u}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\sqrt{u})^2} \cdot \mathbb{1}_{\{u \geq 0\}} \\
&= \frac{1}{\sqrt{\pi} \cdot 2^{1/2}} \cdot u^{1/2 - 1} \cdot e^{-u/2} \cdot \mathbb{1}_{\{u \geq 0\}}
\end{aligned}
$$

- One useful fact: $\Gamma(1/2) = \sqrt{\pi}$. Hence:

$$
f_U(u) = \frac{1}{\Gamma(1/2) \cdot 2^{1/2}} \cdot u^{1/2 - 1} \cdot e^{-u/2} \cdot \mathbb{1}_{\{u \geq 0\}}
$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Example (cont'd)

$$\begin{aligned} f_U(u) &= \frac{1}{\sqrt{u}} \phi(\sqrt{u}) \cdot \mathbb{1}_{\{u \geq 0\}} \\ &= \frac{1}{\sqrt{u}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\sqrt{u})^2} \cdot \mathbb{1}_{\{u \geq 0\}} \\ &= \frac{1}{\sqrt{\pi} \cdot 2^{1/2}} \cdot u^{1/2 - 1} \cdot e^{-u/2} \cdot \mathbb{1}_{\{u \geq 0\}} \end{aligned}$$

- One useful fact: $\Gamma(1/2) = \sqrt{\pi}$. Hence:

$$f_U(u) = \frac{1}{\Gamma(1/2) \cdot 2^{1/2}} \cdot u^{1/2-1} \cdot e^{-u/2} \cdot \mathbb{1}_{\{u \geq 0\}}$$

- Indeed, $U \sim \text{Gamma}(1/2\ ,\ 2) \stackrel{\text{d}}{=} \chi^2_1$!

UC SANTA BARBARA
Department of Statistics and Applied Probability

Theorem

- This is an ***extremely*** important result which we will use repeatedly throughout this course. Let's make it more formal by rephrasing it as a theorem:

Theorem (Square of Standard Normal)

If $Y \sim \mathcal{N}(0, 1)$ and $U := Y^2$, then $U \sim \chi^2_1$.

- The proof of this theorem is exactly the work we did on the previous slides.

Recap

- Whew- that was a lot of work! Let's recap.

Recap

- Whew- that was a lot of work! Let's recap.
- Given a random variable Y, and $U := g(Y)$ for some function $g : \mathbb{R} \to \mathbb{R}$, we can use the **method of distribution functions** (aka the **CDF**) method to find the distribution of U.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Recap

- Whew- that was a lot of work! Let's recap.
- Given a random variable Y, and $U := g(Y)$ for some function $g : \mathbb{R} \to \mathbb{R}$, we can use the **method of distribution functions** (aka the **CDF**) method to find the distribution of U.
- Specifically, this entails:

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Recap

- Whew- that was a lot of work! Let's recap.
- Given a random variable Y, and $U := g(Y)$ for some function $g : \mathbb{R} \to \mathbb{R}$, we can use the **method of distribution functions** (aka the **CDF**) method to find the distribution of U.
- Specifically, this entails:
  (1) Writing $F_U(u)$, the CDF of U, in terms of $F_Y(y)$, the CDF of Y, by basically finding an equivalent formulation for the event $A_U := \{U \leq u\}$ that is in terms of Y

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Recap

- Whew- that was a lot of work! Let's recap.
- Given a random variable Y, and $U := g(Y)$ for some function $g : \mathbb{R} \to \mathbb{R}$, we can use the **method of distribution functions** (aka the **CDF**) method to find the distribution of U.
- Specifically, this entails:
  (1) Writing $F_U(u)$, the CDF of U, in terms of $F_Y(y)$, the CDF of Y, by basically finding an equivalent formulation for the event $A_U := \{U \leq u\}$ that is in terms of Y
  (2) Plugging into the CDF of Y, and simplifying as necessary.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Method of Transformations (Change of Variable Formula)

UC SANTA BARBARA
Department of Statistics and Applied Probability

Leadup

- Let's, for a moment, return the example where we squared the standard normal distribution.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Leadup

- Let's, for a moment, return the example where we squared the standard normal distribution.
- Specifically, we had $Y \sim \mathcal{N}(0, 1)$ and $U := Y^2$.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Leadup

- Let's, for a moment, return the example where we squared the standard normal distribution.
- Specifically, we had $Y \sim \mathcal{N}(0, 1)$ and $U := Y^2$.
- After some work, we found that $F_U(u) = 2\Phi(\sqrt{u}) - 1$.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Leadup

- Let's, for a moment, return the example where we squared the standard normal distribution.
- Specifically, we had $Y \sim \mathcal{N}(0, 1)$ and $U := Y^2$.
- After some work, we found that $F_U(u) = 2\Phi(\sqrt{u}) - 1$.
- Then, we differentiated wrt. u to obtain a formula for $f_U(u)$.

Leadup

- Let's, for a moment, return the example where we squared the standard normal distribution.
- Specifically, we had $Y \sim \mathcal{N}(0, 1)$ and $U := Y^2$.
- After some work, we found that $F_U(u) = 2\Phi(\sqrt{u}) - 1$.
- Then, we differentiated wrt. u to obtain a formula for $f_U(u)$.
- This begs the question - can we perhaps "extend" the CDF method to give us a formula for the *PDF* of U directly?

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Leadup

- Let's, for a moment, return the example where we squared the standard normal distribution.
- Specifically, we had $Y \sim \mathcal{N}(0, 1)$ and $U := Y^2$.
- After some work, we found that $F_U(u) = 2\Phi(\sqrt{u}) - 1$.
- Then, we differentiated wrt. u to obtain a formula for $f_U(u)$.
- This begs the question - can we perhaps "extend" the CDF method to give us a formula for the *PDF* of U directly?
- The answer turns out to be "yes, under some conditions."

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Leadup

Goal

Given a random variable Y and a function $g()$, we seek to describe the random variable $U := g(Y)$.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Leadup

Goal

Given a random variable Y and a function $g()$, we seek to describe the random variable $U := g(Y)$.

- Let's see what happens if we try to apply the CDF method:

$$F_U(u) := \mathbb{P}(U \leq u) = \mathbb{P}(g(Y) \leq u)$$

UC SANTA BARBARA
Department of Statistics and Applied Probability

Leadup

Goal

Given a random variable Y and a function $g()$, we seek to describe the random variable $U := g(Y)$.

- Let's see what happens if we try to apply the CDF method:

$$F_U(u) := \mathbb{P}(U \leq u) = \mathbb{P}(g(Y) \leq u)$$

- Isn't it tempting to apply $g^{-1}(\cdot)$ to both sides of the inequality?

UC SANTA BARBARA
Department of Statistics and Applied Probability

Leadup

Goal

Given a random variable Y and a function $g()$, we seek to describe the random variable $U := g(Y)$.

- Let's see what happens if we try to apply the CDF method:

$$F_U(u) := \mathbb{P}(U \leq u) = \mathbb{P}(g(Y) \leq u)$$

- Isn't it tempting to apply $g^{-1}(\cdot)$ to both sides of the inequality?
- It is! <u>But we need to be careful</u>. First, remember that we don't have any guarantee that $g^{-1}(\cdot)$ even exists!

UC SANTA BARBARA
Department of Statistics and Applied Probability

Leadup

- Alright, then - let's add some assumption about our function $g(\cdot)$.

Leadup

- Alright, then - let's add some assumption about our function $g(\cdot)$.

Goal

Given a random variable Y and a strictly increasing function $g()$, we seek to find $f_U(u)$, the PDF of U.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Leadup

- Alright, then - let's add some assumption about our function $g(\cdot)$.

Goal

Given a random variable Y and a strictly increasing function $g()$, we seek to find $f_U(u)$, the PDF of U.

- *Now* we are guaranteed the existence of $g^{-1}(\cdot)$.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Leadup

- Alright, then - let's add some assumption about our function $g(\cdot)$.

Goal

Given a random variable Y and a strictly increasing function $g()$, we seek to find $f_U(u)$, the PDF of U.

- *Now* we are guaranteed the existence of $g^{-1}(\cdot)$.
- Furthermore, since we assumed $g(\cdot)$ itself to be strictly *increasing*, $g^{-1}(\cdot)$ will also be strictly increasing.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Leadup

- Alright, then - let's add some assumption about our function $g(\cdot)$.

Goal

Given a random variable Y and a strictly increasing function $g()$, we seek to find $f_U(u)$, the PDF of U.

- *Now* we are guaranteed the existence of $g^{-1}(\cdot)$.
- Furthermore, since we assumed $g(\cdot)$ itself to be strictly *increasing*, $g^{-1}(\cdot)$ will also be strictly increasing.
- Hence, we "preserve the direction of inequality" when applying $g^{-1}(\cdot)$ to both sides of an inequality.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Leadup

- Then:

$$F_{(}u) := \mathbb{P}(U \leq u) = \mathbb{P}(g(Y) \leq u) = \mathbb{P}(Y \leq g^{-1}(u)) = F_Y(g^{-1}(u))$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Leadup

- Then:

$$F_{(}u) := \mathbb{P}(U \leq u) = \mathbb{P}(g(Y) \leq u) = \mathbb{P}(Y \leq g^{-1}(u)) = F_Y(g^{-1}(u))$$

- We can now differentiate wrt. U and apply the chain rule (from calculus; we can discuss this further on the chalkboard):

$$\begin{aligned} f_U(u) &:= \frac{\mathrm{d}}{\mathrm{d}u} F_U(u) \\ &= \frac{\mathrm{d}}{\mathrm{d}u} F_Y(g^{-1}(u)) \\ &= f_Y(g^{-1}(u)) \cdot \frac{\mathrm{d}}{\mathrm{d}u} g^{-1}(u) \end{aligned}$$

Leadup

- If we instead assume that $g(\cdot)$ is strictly decreasing, a similar computation (which I'll be asking you to complete on your homework) yields

$$f_U(u) = f_Y(g^{-1}(u)) \cdot \left[-\frac{\mathrm{d}}{\mathrm{d}u} g^{-1}(u) \right]$$

Leadup

- If we instead assume that $g(\cdot)$ is strictly decreasing, a similar computation (which I'll be asking you to complete on your homework) yields

$$f_U(u) = f_Y(g^{-1}(u)) \cdot \left[-\frac{\mathrm{d}}{\mathrm{d}u} g^{-1}(u) \right]$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Leadup

- If we instead assume that $g(\cdot)$ is strictly decreasing, a similar computation (which I'll be asking you to complete on your homework) yields

$$f_U(u) = f_Y(g^{-1}(u)) \cdot \left[-\frac{\mathrm{d}}{\mathrm{d}u} g^{-1}(u) \right]$$

- So, if we instead simply assume that $g(\cdot)$ is strictly monotonic, we can summarize our findings as:

$$f_U(u) = \begin{cases} f_Y(g^{-1}(u)) \cdot \left[\frac{\mathrm{d}}{\mathrm{d}u} g^{-1}(u) \right] & \text{if } g(\cdot) \text{ is increasing} \\ f_Y(g^{-1}(u)) \cdot \left[-\frac{\mathrm{d}}{\mathrm{d}u} g^{-1}(u) \right] & \text{if } g(\cdot) \text{ is decreasing} \end{cases}$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Change of Variable Formula

- A bit of simplification (and recollections of how derivatives of increasing/decreasing functions behaves) allows us to rewrite our result above as:

UC SANTA BARBARA
Department of Statistics and Applied Probability

Change of Variable Formula

- A bit of simplification (and recollections of how derivatives of increasing/decreasing functions behaves) allows us to rewrite our result above as:

Theorem (Change of Variable Formula)

Given a random variable $Y \sim f_Y$ and a function $g(\cdot)$ that is strictly monotonic over the support of Y, then the random variable $U := g(Y)$ has density given by

$$f_U(u) = f_Y[g^{-1}(u)] \cdot \left| \frac{\mathrm{d}}{\mathrm{d}u} g^{-1}(u) \right|$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Example

- As an example, let's re-derive the closure under multiplication property of the Exponential distribution, this time using the Change of Variable formula.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Example

- As an example, let's re-derive the closure under multiplication property of the Exponential distribution, this time using the Change of Variable formula.
- That is: let $Y \sim \text{Exp}(\theta)$, and set $U := cY$ for some positive constant $c > 0$.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Example

- As an example, let's re-derive the closure under multiplication property of the Exponential distribution, this time using the Change of Variable formula.
- That is: let $Y \sim \text{Exp}(\theta)$, and set $U := cY$ for some positive constant $c > 0$.
- Since the transformation $g(y) = cy$ is strictly monotonic (specifically, it's strictly increasing) it's inverse exists and is calculable as $g^{-1}(u) = u/c$. Hence:

$$\left| \frac{\mathrm{d}}{\mathrm{d}u} g^{-1}(u) \right| = \left| \frac{\mathrm{d}}{\mathrm{d}u} \left(\frac{u}{c} \right) \right| = \left| \frac{1}{c} \right| = \frac{1}{c}$$

  where we have dropped the absolute values in the last step since we are assuming $c > 0$.

Example

- Additionally, since $Y \sim \text{Exp}(\theta)$ we know that

$$f_Y(y) = \frac{1}{\theta} \exp\left\{-\frac{y}{\theta}\right\} \cdot \mathbb{1}_{\{y \geq 0\}}$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Example

- Additionally, since $Y \sim \text{Exp}(\theta)$ we know that

$$f_Y(y) = \frac{1}{\theta} \exp\left\{-\frac{y}{\theta}\right\} \cdot \mathbb{1}_{\{y \geq 0\}}$$

- Therefore, plugging into the change of variable formula, we have

$$\begin{aligned} f_U(u) &= f_Y[g^{-1}(u)] \cdot \left| \frac{\mathrm{d}}{\mathrm{d}u} g^{-1}(u) \right| \\ &= \frac{1}{\theta} \exp\left\{ -\frac{\left(\frac{u}{c}\right)}{\theta} \right\} \cdot \mathbb{1}_{\{\frac{u}{c} \geq 0\}} \cdot \frac{1}{c} \\ &= \frac{1}{c\theta} \exp\left\{-\frac{u}{c\theta}\right\} \cdot \mathbb{1}_{\{u \geq 0\}} \end{aligned}$$

UC SANTA BARBARA
Department of Statistics and Applied Probability

Clicker Question!

Clicker Question 1

Given $Y \sim \text{Unif}[1, 2]$ and $U := 2X + 3$, does U also follow a Uniform Distribution?

(A) Yes; (B) No

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Change of Variable Formula

- Now, note that the only assumption we need to make about $g(\cdot)$ in order for the Change of Variable formula to hold is that it is strictly monotone *over the support of* Y.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Change of Variable Formula

- Now, note that the only assumption we need to make about $g(\cdot)$ in order for the Change of Variable formula to hold is that it is strictly monotone *over the support of Y*.
- For example, suppose $Y \sim \text{Unif}[-1, 0]$ and take $U := Y^2$.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Change of Variable Formula

- Now, note that the only assumption we need to make about $g(\cdot)$ in order for the Change of Variable formula to hold is that it is strictly monotone *over the support of Y*.
- For example, suppose $Y \sim \text{Unif}[-1, 0]$ and take $U := Y^2$.
- Though the function $g(y) = y^2$ is not strictly monotone over $\mathbb{R}$, it *is* strictly monotone over $S_Y := [-1, 0]$ (i.e. the support of Y), and hence its inverse exists and is given by $g^{-1}(u) = -\sqrt{u}$.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Change of Variable Formula

- Now, note that the only assumption we need to make about $g(\cdot)$ in order for the Change of Variable formula to hold is that it is strictly monotone *over the support of* Y.
- For example, suppose $Y \sim \text{Unif}[-1, 0]$ and take $U := Y^2$.
- Though the function $g(y) = y^2$ is not strictly monotone over $\mathbb{R}$, it *is* strictly monotone over $S_Y := [-1, 0]$ (i.e. the support of Y), and hence its inverse exists and is given by $g^{-1}(u) = -\sqrt{u}$.
- The Change of Variable formula can therefore safely be applied.

Change of Variable Formula

- In general, however, the Change of Variable formula does not work when we are dealing with transformations that are not strictly monotone.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Change of Variable Formula

- In general, however, the Change of Variable formula does <u>not</u> work when we are dealing with transformations that are not strictly monotone.
- For example, given $Y \sim \mathcal{N}(0, 1)$ and $U := Y^2$, we *cannot* directly apply the Change of Variable formula.

Change of Variable Formula

- In general, however, the Change of Variable formula does <u>not</u> work when we are dealing with transformations that are not strictly monotone.
- For example, given $Y \sim \mathcal{N}(0, 1)$ and $U := Y^2$, we *cannot* directly apply the Change of Variable formula.
  - Admittedly, there does exist a way to generalize the Change of Variable formula to work in a situation like this, but we won't cover that in PSTAT 120B. If you're curious, I'm happy to walk you through the general outline during Office Hours.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Method of Moment-Generating Functions (MGF Method)

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Leadup

Goal

Given a random variable Y and a function $g()$, we seek to describe the random variable $U := g(Y)$.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Leadup

Goal

Given a random variable Y and a function $g()$, we seek to describe the random variable $U := g(Y)$.

- So far, we've talked about "describing" the distribution of U by both its CDF (using the CDF method) and its PDF (using the Change of Variable formula).

UC SANTA BARBARA
Department of Statistics and Applied Probability

Leadup

Goal

Given a random variable Y and a function $g()$, we seek to describe the random variable $U := g(Y)$.

- So far, we've talked about "describing" the distribution of U by both its CDF (using the CDF method) and its PDF (using the Change of Variable formula).
- We know that there is a third way of classifying distributions - **moment-generating functions** (MGFs).

UC SANTA BARBARA
Department of Statistics and Applied Probability

MGFs

Definition (MGF)

The MGF of a random variable X, notated $M_X(t)$, is defined as

$$M_X(t) := \mathbb{E}[e^{tX}]$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

MGFs

Definition (MGF)

The MGF of a random variable X, notated $M_X(t)$, is defined as

$$M_X(t) := \mathbb{E}[e^{tX}]$$

- Recall that this expectation is computed as a sum if X is discrete and as an integral if X is continuous.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Useful Result

Theorem

Given two random variables X and Y with MGFs $M_X(t)$ and $M_Y(t)$, respectively, that are both continuous in a small neighborhood of the origin, then $M_X(t) = M_Y(t)$ implies that X and Y have the same distribution.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Useful Result

Theorem

Given two random variables X and Y with MGFs $M_X(t)$ and $M_Y(t)$, respectively, that are both continuous in a small neighborhood of the origin, then $M_X(t) = M_Y(t)$ implies that X and Y have the same distribution.

- This theorem is essentially just a more formal way of saying "MGFs uniquely determine random variables." For example,

$$M_X(t) = \exp\left\{2t + \frac{1}{2}t^2\right\} \iff X \sim \mathcal{N}(2, 1)$$

UC SANTA BARBARA
Department of Statistics and Applied Probability

Useful Result

Theorem

Given a random variable Y with MGF $M_Y(t)$, and $U := aY + b$ for constants $a, b \in \mathbb{R}$,

$$M_U(t) = e^{bt} M_Y(at)$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Proof.

$$M_U(t) := \mathbb{E}[e^{tU}]$$ [Definition of MGF]

$\square$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Proof.

$$
\begin{aligned}
M_U(t) &:= \mathbb{E}[e^{tU}] && \text{[Definition of MGF]} \\
&:= \mathbb{E}[e^{t(aY+b)}] && \text{[Definition of } U\text{]}
\end{aligned}
$$

$\square$

Proof.

$$
\begin{aligned}
M_U(t) &:= \mathbb{E}[e^{tU}] && \text{[Definition of MGF]} \\
&:= \mathbb{E}[e^{t(aY+b)}] && \text{[Definition of } U\text{]} \\
&:= \mathbb{E}[e^{(at)Y+bt}] && \text{[Algebra]}
\end{aligned}
$$

$\square$

UC SANTA BARBARA
Department of Statistics and Applied Probability

Proof.

$$
\begin{aligned}
M_U(t) &:= \mathbb{E}[e^{tU}] && \text{[Definition of MGF]} \\
&:= \mathbb{E}[e^{t(aY+b)}] && \text{[Definition of } U\text{]} \\
&:= \mathbb{E}[e^{(at)Y+bt}] && \text{[Algebra]} \\
&:= \mathbb{E}[e^{(at)Y} \cdot e^{bt}] && \text{[Algebra]}
\end{aligned}
$$

$\square$

UC SANTA BARBARA
Department of Statistics and Applied Probability

Proof.

$$
\begin{aligned}
M_U(t) &:= \mathbb{E}[e^{tU}] && \text{[Definition of MGF]} \\
&:= \mathbb{E}[e^{t(aY+b)}] && \text{[Definition of } U\text{]} \\
&:= \mathbb{E}[e^{(at)Y+bt}] && \text{[Algebra]} \\
&:= \mathbb{E}[e^{(at)Y} \cdot e^{bt}] && \text{[Algebra]} \\
&:= e^{bt}\mathbb{E}[e^{(at)Y}] && \text{[Linearity of } \mathbb{E}\text{]}
\end{aligned}
$$

$\square$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Proof.

$$
\begin{aligned}
M_U(t) &:= \mathbb{E}[e^{tU}] && \text{[Definition of MGF]} \\
&:= \mathbb{E}[e^{t(aY+b)}] && \text{[Definition of } U\text{]} \\
&:= \mathbb{E}[e^{(at)Y+bt}] && \text{[Algebra]} \\
&:= \mathbb{E}[e^{(at)Y} \cdot e^{bt}] && \text{[Algebra]} \\
&:= e^{bt}\mathbb{E}[e^{(at)Y}] && \text{[Linearity of } \mathbb{E}\text{]} \\
&:= e^{bt}M_Y(at) && \text{[Definition of MGF]}
\end{aligned}
$$

$\square$

UC SANTA BARBARA
Department of Statistics and Applied Probability

Proof.

$$
\begin{aligned}
M_U(t) &:= \mathbb{E}[e^{tU}] && \text{[Definition of MGF]} \\
&:= \mathbb{E}[e^{t(aY+b)}] && \text{[Definition of } U\text{]} \\
&:= \mathbb{E}[e^{(at)Y+bt}] && \text{[Algebra]} \\
&:= \mathbb{E}[e^{(at)Y} \cdot e^{bt}] && \text{[Algebra]} \\
&:= e^{bt}\mathbb{E}[e^{(at)Y}] && \text{[Linearity of } \mathbb{E}\text{]} \\
&:= e^{bt}M_Y(at) && \text{[Definition of MGF]}
\end{aligned}
$$

□

- It turns out, we can use this theorem to (again) prove the closure of the exponential distribution under multiplication!

Example

- Once again, let $Y \sim \text{Exp}(\theta)$, and let $U = cY$ for a positive constant c.

Example

- Once again, let $Y \sim \text{Exp}(\theta)$, and let $U = cY$ for a positive constant c.
- First recall that the MGF of the exponential distribution is

$$M_Y(t) = \begin{cases} (1 - \theta t)^{-1} & \text{if } t < 1/\theta \\ \infty & \text{otherwise} \end{cases}$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Example

- Once again, let $Y \sim \text{Exp}(\theta)$, and let $U = cY$ for a positive constant c.
- First recall that the MGF of the exponential distribution is

$$M_Y(t) = \begin{cases} (1 - \theta t)^{-1} & \text{if } t < 1/\theta \\ \infty & \text{otherwise} \end{cases}$$

- Hence, by the previous theorem:

$$M_U(t) = e^{0 \cdot t} \cdot M_Y(ct)$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Example

- Once again, let $Y \sim \text{Exp}(\theta)$, and let $U = cY$ for a positive constant c.
- First recall that the MGF of the exponential distribution is

$$M_Y(t) = \begin{cases} (1-\theta t)^{-1} & \text{if } t < 1/\theta \\ \infty & \text{otherwise} \end{cases}$$

- Hence, by the previous theorem:

$$M_U(t) = e^{0 \cdot t} \cdot M_Y(ct) = 1 \cdot \begin{cases} (1-\theta(ct))^{-1} & \text{if } (ct) < 1/\theta \\ \infty & \text{otherwise} \end{cases}$$

Example

- Once again, let $Y \sim \text{Exp}(\theta)$, and let $U = cY$ for a positive constant c.
- First recall that the MGF of the exponential distribution is

$$M_Y(t) = \begin{cases} (1-\theta t)^{-1} & \text{if } t < 1/\theta \\ \infty & \text{otherwise} \end{cases}$$

- Hence, by the previous theorem:

$$M_U(t) = e^{0 \cdot t} \cdot M_Y(ct) = 1 \cdot \begin{cases} (1-\theta(ct))^{-1} & \text{if } (ct) < 1/\theta \\ \infty & \text{otherwise} \end{cases}$$

$$= \begin{cases} (1-(c\theta)t))^{-1} & \text{if } t < 1/(c\theta) \\ \infty & \text{otherwise} \end{cases}$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Example

- Once again, let $Y \sim \text{Exp}(\theta)$, and let $U = cY$ for a positive constant c.
- First recall that the MGF of the exponential distribution is

$$M_Y(t) = \begin{cases} (1-\theta t)^{-1} & \text{if } t < 1/\theta \\ \infty & \text{otherwise} \end{cases}$$

- Hence, by the previous theorem:

$$M_U(t) = e^{0 \cdot t} \cdot M_Y(ct) = 1 \cdot \begin{cases} (1-\theta(ct))^{-1} & \text{if } (ct) < 1/\theta \\ \infty & \text{otherwise} \end{cases}$$

$$= \begin{cases} (1-(c\theta)t))^{-1} & \text{if } t < 1/(c\theta) \\ \infty & \text{otherwise} \end{cases}$$

which is, as expected, the MGF of the $\text{Exp}(c\theta)$ distribution.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Clicker Question!

Clicker Question 2

If $Y \sim \text{Pois}(\lambda)$ and $U := cY$ for some positive constant c, what is the distribution of U?

(A) $\text{Pois}(c\lambda)$

(B) $\text{Pois}(c/\lambda)$

(C) $\text{Pois}(\lambda/c)$

(D) None of the above

UC SANTA BARBARA
Department of Statistics and Applied Probability

Leadup

- Now, it may seem a little pointless to use the method of MGFs to identify distributions of transformed random variables.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Leadup

- Now, it may seem a little pointless to use the method of MGFs to identify distributions of transformed random variables.
- For one thing, we really only seem to be able to use it when we are dealing with linear transformations.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Leadup

- Now, it may seem a little pointless to use the method of MGFs to identify distributions of transformed random variables.
- For one thing, we really only seem to be able to use it when we are dealing with linear transformations.
- For another, we are required to simply recognize the MGF of the resulting distribution at the end of the problem.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Leadup

- Now, it may seem a little pointless to use the method of MGFs to identify distributions of transformed random variables.
- For one thing, we really only seem to be able to use it when we are dealing with linear transformations.
- For another, we are required to simply recognize the MGF of the resulting distribution at the end of the problem.
- The last point is definitely a legitimate criticism of the method - the method of MGFs will not give you the PDF of the transformed random variable unless you're able to recognize the MGF.

Leadup

- Now, it may seem a little pointless to use the method of MGFs to identify distributions of transformed random variables.
- For one thing, we really only seem to be able to use it when we are dealing with linear transformations.
- For another, we are required to simply recognize the MGF of the resulting distribution at the end of the problem.
- The last point is definitely a legitimate criticism of the method - the method of MGFs will not give you the PDF of the transformed random variable unless you're able to recognize the MGF.
- However, we will see that the method of MGFs is incredibly useful in the next section of our lecture... so stay tuned!

Summary: Univariate Transformations

Goal

Given a random variable Y and a function $g()$, we seek to describe the random variable $U := g(Y)$.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Summary: Univariate Transformations

Goal

Given a random variable Y and a function $g()$, we seek to describe the random variable $U := g(Y)$.

- So far, we've accomplished this goal in three different ways:

UC SANTA BARBARA
Department of Statistics and Applied Probability

Summary: Univariate Transformations

Goal

Given a random variable Y and a function $g()$, we seek to describe the random variable $U := g(Y)$.

- So far, we've accomplished this goal in three different ways:
  - The CDF Method (Method of Distribution Functions)

UC SANTA BARBARA
Department of Statistics and Applied Probability

Summary: Univariate Transformations

Goal

Given a random variable Y and a function $g()$, we seek to describe the random variable $U := g(Y)$.

- So far, we've accomplished this goal in three different ways:
  - The CDF Method (Method of Distribution Functions)
  - The Change of Variable formula (Method of Transformations)

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Summary: Univariate Transformations

Goal

Given a random variable Y and a function $g()$, we seek to describe the random variable $U := g(Y)$.

- So far, we've accomplished this goal in three different ways:
  - The CDF Method (Method of Distribution Functions)
  - The Change of Variable formula (Method of Transformations)
  - The MGF Method.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

CDF Method

CDF Method

(1) Write $F_U(u) := \mathbb{P}(A_U)$, where $A_U := \{U \leq u\}$.

CDF Method

(1) Write $F_U(u) := \mathbb{P}(A_U)$, where $A_U := \{U \leq u\}$.

(2) Find an equivalent expression for A_U in terms of Y; call this A_Y.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

CDF Method

(1) Write $F_U(u) := \mathbb{P}(A_U)$, where $A_U := \{U \leq u\}$.

(2) Find an equivalent expression for A_U in terms of Y; call this A_Y.

(3) Compute $\mathbb{P}(A_Y)$ using the distribution of Y (which is known), to then find the CDF of U.

CDF Method

(1) Write $F_U(u) := \mathbb{P}(A_U)$, where $A_U := \{U \leq u\}$.

(2) Find an equivalent expression for A_U in terms of Y; call this A_Y.

(3) Compute $\mathbb{P}(A_Y)$ using the distribution of Y (which is known), to then find the CDF of U.

- Remember: when carrying out step 3, drawing a picture can be incredibly helpful!

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Change of Variable Formula

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Change of Variable Formula

(1) Compute $g^{-1}(u)$ [remember that this can be done by solving the equation $u = g(y)$ for u in terms of y].

Change of Variable Formula

(1) Compute $g^{-1}(u)$ [remember that this can be done by solving the equation $u = g(y)$ for u in terms of y].

(2) Plug into the Change of Variable formula:

$$f_U(u) = f_Y[g^{-1}(y)] \cdot \left| \frac{\mathrm{d}}{\mathrm{d}u} g^{-1}(u) \right|$$

UC SANTA BARBARA
Department of Statistics and Applied Probability

Change of Variable Formula

(1) Compute $g^{-1}(u)$ [remember that this can be done by solving the equation $u = g(y)$ for u in terms of y].
(2) Plug into the Change of Variable formula:

$$f_U(u) = f_Y[g^{-1}(y)] \cdot \left| \frac{\mathrm{d}}{\mathrm{d}u} g^{-1}(u) \right|$$

- Remember: this method only works when the transformation $g(\cdot)$ is strictly monotonic over the support of Y!

UC SANTA BARBARA
Department of Statistics and Applied Probability

Change of Variable Formula

(1) Compute $g^{-1}(u)$ [remember that this can be done by solving the equation $u = g(y)$ for u in terms of y].
(2) Plug into the Change of Variable formula:

$$f_U(u) = f_Y[g^{-1}(y)] \cdot \left| \frac{\mathrm{d}}{\mathrm{d}u} g^{-1}(u) \right|$$

- Remember: this method only works when the transformation $g(\cdot)$ is strictly monotonic over the support of Y!
- Also, a side note: so long as you are careful, the change of variable formula will give you the support of U. But, in some cases, it might be easier to find the support first (by drawing a picture), and then incorporating that into your answer later.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

MGF Method

MGF Method

(1) Compute the MGF $M_U(t)$ of U by writing it in terms of the MGF $M_Y(t)$ of Y, and then recognize the resulting MGF as belonging to a particular distribution.

UC SANTA BARBARA
Department of Statistics and Applied Probability

MGF Method

(1) Compute the MGF $M_U(t)$ of U by writing it in terms of the MGF $M_Y(t)$ of Y, and then recognize the resulting MGF as belonging to a particular distribution.

- This works well for linear transformations, but not too well for nonlinear transformations.

MGF Method

(1) Compute the MGF $M_U(t)$ of U by writing it in terms of the MGF $M_Y(t)$ of Y, and then recognize the resulting MGF as belonging to a particular distribution.

- This works well for linear transformations, but not too well for nonlinear transformations.
- Also, the MGF method won't (typically) give you a PDF/CDF, so if you really want the PDF/CDF of U you should use a different method.

MGF Method

(1) Compute the MGF $M_U(t)$ of U by writing it in terms of the MGF $M_Y(t)$ of Y, and then recognize the resulting MGF as belonging to a particular distribution.

- This works well for linear transformations, but not too well for nonlinear transformations.
- Also, the MGF method won't (typically) give you a PDF/CDF, so if you really want the PDF/CDF of U you should use a different method.
- And, again, we'll see that the MGF method really shines in a slightly different context...

Chalkboard Example

Example

The **kinetic energy** of a particle with mass m traveling at a velocity V is given by

$$E = \frac{1}{2} m V^2$$

Consider a particle selected at random, whose velocity is a random variable V with density

$$f_V(v) = 2v^3 e^{-v^2} \cdot \mathbb{1}_{\{v > 0\}}$$

Find the distribution of the kinetic energy of this particle once using the CDF method and once using the Change of Variable formula.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Multivariate Transformations

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Leadup

- Recall that our goal up until now has been:

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Leadup

- Recall that our goal up until now has been:

Goal

Given a random variable Y and a function $g()$, we seek to describe the random variable $U := g(Y)$.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Leadup

- Recall that our goal up until now has been:

Goal

Given a random variable Y and a function $g()$, we seek to describe the random variable $U := g(Y)$.

- As mentioned previously, this falls under the umbrella of **univariate transformations**, as we are investigating transformations of a *single* random variable Y.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Leadup

- Recall that our goal up until now has been:

Goal

Given a random variable Y and a function $g()$, we seek to describe the random variable $U := g(Y)$.

- As mentioned previously, this falls under the umbrella of **univariate transformations**, as we are investigating transformations of a *single* random variable Y.
- In many situations, it's desired to investigate the transformation of *two or more* random variables, leading us to the realm of **multivariate transformations**.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Bivariate Transformations

- For now, let's restrict ourselves to considering only **bivariate transformations**; i.e. transformations of two random variables.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Bivariate Transformations

- For now, let's restrict ourselves to considering only **bivariate transformations**; i.e. transformations of two random variables.

Goal

Given a bivariate random vector $(Y_1, Y_2) \sim f_{Y_1,Y_2}$ and a function $g : \mathbb{R}^2 \to \mathbb{R}$, we seek to describe the random variable $U := g(Y_1, Y_2)$.

Bivariate Transformations

- For now, let's restrict ourselves to considering only **bivariate transformations**; i.e. transformations of two random variables.

Goal

Given a bivariate random vector $(Y_1, Y_2) \sim f_{Y_1,Y_2}$ and a function $g : \mathbb{R}^2 \to \mathbb{R}$, we seek to describe the random variable $U := g(Y_1, Y_2)$.

- First, note that a "bivariate random vector" is just a fancy way of saying "pair of random variables."

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Bivariate Transformations

- For now, let's restrict ourselves to considering only **bivariate transformations**; i.e. transformations of two random variables.

Goal

Given a bivariate random vector $(Y_1, Y_2) \sim f_{Y_1,Y_2}$ and a function $g : \mathbb{R}^2 \to \mathbb{R}$, we seek to describe the random variable $U := g(Y_1, Y_2)$.

- First, note that a "bivariate random vector" is just a fancy way of saying "pair of random variables."
  - The notation $(Y_1, Y_2) \sim f_{Y_1,Y_2}$ is just a shorthand for saying that the bivariate random vector (Y_1, Y_2) has joint density given by $f_{Y_1,Y_2}(y_1, y_2)$.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Examples

- Some examples of bivariate transformations include:

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Examples

- Some examples of bivariate transformations include:
  - **Sums:** $U := Y_1 + Y_2$

Examples

- Some examples of bivariate transformations include:
  - **Sums:** $U := Y_1 + Y_2$
  - **Differences:** $U := Y_2 - Y_1$

Examples

- Some examples of bivariate transformations include:
  - **Sums:** $U := Y_1 + Y_2$
  - **Differences:** $U := Y_2 - Y_1$
  - **Averages:** $U := (Y_1 + Y_2)/2$

UC SANTA BARBARA
Department of Statistics and Applied Probability

Examples

- Some examples of bivariate transformations include:
  - **Sums:** $U := Y_1 + Y_2$
  - **Differences:** $U := Y_2 - Y_1$
  - **Averages:** $U := (Y_1 + Y_2)/2$
  - More involved transformations, like $U := 2Y_1^2 - \sqrt{Y_2}$

Examples

- Some examples of bivariate transformations include:
  - **Sums:** $U := Y_1 + Y_2$
  - **Differences:** $U := Y_2 - Y_1$
  - **Averages:** $U := (Y_1 + Y_2)/2$
  - More involved transformations, like $U := 2Y_1^2 - \sqrt{Y_2}$

- Notice that in each of these cases our transformation $g(\cdot)$ is a function that takes *two* inputs and returns only *one* output.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Reference Example

- To ground ourselves even further, let's consider the following reference example: let $Y_1, Y_2 \overset{\text{i.i.d.}}{\sim} \text{Exp}(\theta)$, and set $U := Y_1 + Y_2$. Our goal is to find the distribution of U.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Reference Example

- To ground ourselves even further, let's consider the following reference example: let $Y_1, Y_2 \overset{\text{i.i.d.}}{\sim} \text{Exp}(\theta)$, and set $U := Y_1 + Y_2$. Our goal is to find the distribution of U.
- Well... nothing's stopping us from trying to start where we started our univariate considerations, and trying to find an expression for the CDF of U!

Reference Example

- To ground ourselves even further, let's consider the following reference example: let $Y_1, Y_2 \overset{\text{i.i.d.}}{\sim} \text{Exp}(\theta)$, and set $U := Y_1 + Y_2$. Our goal is to find the distribution of U.
- Well... nothing's stopping us from trying to start where we started our univariate considerations, and trying to find an expression for the CDF of U!
- Specifically,

$$F_U(u) := \mathbb{P}(U \leq u) = \mathbb{P}(Y_1 + Y_2 \leq u)$$

UC SANTA BARBARA
Department of Statistics and Applied Probability

Reference Example

- To ground ourselves even further, let's consider the following reference example: let $Y_1, Y_2 \overset{\text{i.i.d.}}{\sim} \text{Exp}(\theta)$, and set $U := Y_1 + Y_2$. Our goal is to find the distribution of U.
- Well... nothing's stopping us from trying to start where we started our univariate considerations, and trying to find an expression for the CDF of U!
- Specifically,
$$F_U(u) := \mathbb{P}(U \leq u) = \mathbb{P}(Y_1 + Y_2 \leq u)$$
- Now, this should look **VERY** familiar to you... from PSTAT 120A!

Reference Example

- Specifically, the probability on the RHS is something we can find by double integrating the joint density $f_{Y_1,Y_2}(y_1, y_2)$ over the region

$$\mathcal{R} := \{(y_1, y_2) \in S_{Y_1,Y_2} : y_1 + y_2 \leq u\}$$

where S_{Y_1,Y_2} denotes the joint support of (Y_1, Y_2).

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Reference Example

- Specifically, the probability on the RHS is something we can find by double integrating the joint density $f_{Y_1,Y_2}(y_1, y_2)$ over the region

$$\mathcal{R} := \{(y_1, y_2) \in S_{Y_1,Y_2} : y_1 + y_2 \leq u\}$$

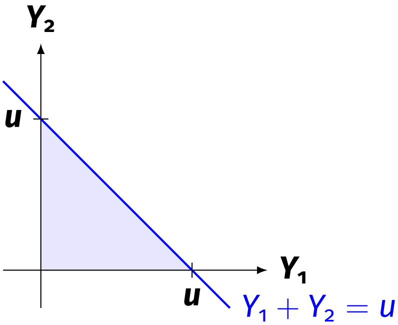
where S_{Y_1,Y_2} denotes the joint support of (Y_1, Y_2).

- That is,

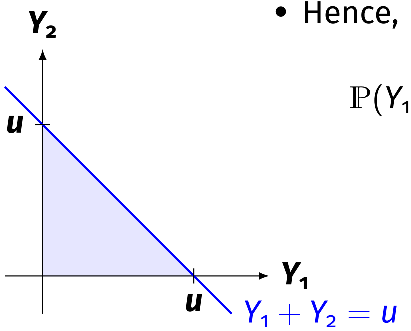
$$\mathbb{P}(Y_1 + Y_2 \leq u) = \iint_{\mathcal{R}} f_{Y_1,Y_2}(y_1, y_2)\, \mathrm{d}A$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Region of Integration



Region of Integration



- Hence,

$$\begin{aligned}
\mathbb{P}(Y_1 + Y_2 \leq u) &= \int_0^u \int_0^{u-y_1} f_{Y_1,Y_2}(y_1, y_2)\ \mathrm{d}y_2\ \mathrm{d}y_1 \\
&= \int_0^u \int_0^{u-y_2} f_{Y_1,Y_2}(y_1, y_2)\ \mathrm{d}y_1\ \mathrm{d}y_2
\end{aligned}$$

Joint Density

- In this case, neither order of integration will be particularly easier than the other - as such, let's (somewhat arbitrarily) use the order $dy_2\, dy_1$.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Joint Density

- In this case, neither order of integration will be particularly easier than the other - as such, let's (somewhat arbitrarily) use the order $dy_2 \, dy_1$.

- Additionally, note that since $Y_1, Y_2 \overset{\text{i.i.d.}}{\sim} \text{Exp}(\theta)$:

$$f_{Y_1,Y_2}(y_1, y_2) = f_{Y_1}(y_1) \cdot f_{Y_2}(y_2) = \frac{1}{\theta^2} e^{-(y_1+y_2)/\theta} \cdot \mathbb{1}_{\{y_1 \geq 0, \; y_2 \geq 0\}}$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Integration

- So:

$$\mathbb{P}(Y_1 + Y_2 \leq u) = \int_0^u \int_0^{u-y_1} \frac{1}{\theta^2} e^{-(y_1+y_2)/\theta} \, \mathrm{d}y_2 \, \mathrm{d}y_1$$
$$= \frac{1}{\theta} \cdot \int_0^u e^{-y_1/\theta} \int_0^{u-y_1} \left[\frac{1}{\theta} e^{-y_2/\theta} \right] \mathrm{d}y_2 \, \mathrm{d}y_1$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Integration

- So:

$$\mathbb{P}(Y_1 + Y_2 \leq u) = \int_0^u \int_0^{u-y_1} \frac{1}{\theta^2} e^{-(y_1+y_2)/\theta} \, dy_2 \, dy_1$$
$$= \frac{1}{\theta} \cdot \int_0^u e^{-y_1/\theta} \int_0^{u-y_1} \left[\frac{1}{\theta} e^{-y_2/\theta} \right] dy_2 \, dy_1$$

- Note that the inner integral (which I've highlighted in blue) is just the CDF of the Exp(θ) distribution evaluated at $u - y_1$, and is therefore equal to $1 - e^{-(u-y_1)/\theta}$.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Integration

- Therefore:

$$\begin{aligned}
\mathbb{P}(Y_1 + Y_2 \leq u) &= \frac{1}{\theta} \cdot \int_0^u e^{-y_1/\theta} \int_0^{u-y_1} \left[\frac{1}{\theta} e^{-y_2/\theta} \right] \, dy_2 \, dy_1 \\
&= \frac{1}{\theta} \cdot \int_0^u e^{-y_1/\theta} \left(1 - e^{-(u-y_1)/\theta}\right) \, dy_1 \\
&= \frac{1}{\theta} \cdot \int_0^u \left(e^{-y_1/\theta} - e^{-u/\theta}\right) \, dy_1 \\
&= \int_0^u \frac{1}{\theta} e^{-y_1/\theta} \, dy_1 - \frac{u}{\theta} e^{-u/\theta}
\end{aligned}$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Integration

- Therefore:

$$
\begin{aligned}
\mathbb{P}(Y_1 + Y_2 \leq u) &= \frac{1}{\theta} \cdot \int_0^u e^{-y_1/\theta} \int_0^{u-y_1} \left[\frac{1}{\theta} e^{-y_2/\theta} \right] \, \mathrm{d}y_2 \, \mathrm{d}y_1 \\
&= \frac{1}{\theta} \cdot \int_0^u e^{-y_1/\theta} \left(1 - e^{-(u-y_1)/\theta} \right) \, \mathrm{d}y_1 \\
&= \frac{1}{\theta} \cdot \int_0^u \left(e^{-y_1/\theta} - e^{-u/\theta} \right) \, \mathrm{d}y_1 \\
&= \int_0^u \frac{1}{\theta} e^{-y_1/\theta} \, \mathrm{d}y_1 - \frac{u}{\theta} e^{-u/\theta}
\end{aligned}
$$

- The blue integral is, once again, the CDF of the Exp(θ) distribution, this time evaluated at u.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Integration

- Therefore:

$$\begin{aligned}\mathbb{P}(Y_1 + Y_2 \leq u) &= \int_0^u \frac{1}{\theta} e^{-y_1/\theta}\, \mathrm{d}y_1 - \frac{u}{\theta} e^{-u/\theta} \\ &= 1 - e^{-u/\theta} - \frac{u}{\theta} e^{-u/\theta}\end{aligned}$$

Integration

- Therefore:

$$\mathbb{P}(Y_1 + Y_2 \le u) = \int_0^u \frac{1}{\theta} e^{-y_1/\theta}\, dy_1 - \frac{u}{\theta} e^{-u/\theta}$$
$$= 1 - e^{-u/\theta} - \frac{u}{\theta} e^{-u/\theta}$$

- And this is the CDF of U! If we want the density, we can differentiate wrt. u:

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Differentiation

$$\begin{aligned} f_U(u) &= \frac{\mathrm{d}}{\mathrm{d}u} F_U(u) \cdot \mathbb{1}_{\{u \geq 0\}} \\ &= \frac{\mathrm{d}}{\mathrm{d}u}\left[1 - e^{-u/\theta} - \frac{u}{\theta} e^{-u/\theta}\right] \cdot \mathbb{1}_{\{u \geq 0\}} \\ &= \frac{1}{\theta} e^{-u/\theta} - \frac{1}{\theta}\left[e^{-u/\theta} - \frac{u}{\theta} e^{-u/\theta}\right] \cdot \mathbb{1}_{\{u \geq 0\}} \\ &= \boxed{\frac{u}{\theta^2} e^{-u/\theta} \cdot \mathbb{1}_{\{u \geq 0\}}} \end{aligned}$$

UC SANTA BARBARA
Department of Statistics and Applied Probability

Differentiation

$$
\begin{aligned}
f_U(u) &= \frac{\mathrm{d}}{\mathrm{d}u} F_U(u) \cdot \mathbb{1}_{\{u \geq 0\}} \\
&= \frac{\mathrm{d}}{\mathrm{d}u} \left[1 - e^{-u/\theta} - \frac{u}{\theta} e^{-u/\theta} \right] \cdot \mathbb{1}_{\{u \geq 0\}} \\
&= \frac{1}{\theta} e^{-u/\theta} - \frac{1}{\theta} \left[e^{-u/\theta} - \frac{u}{\theta} e^{-u/\theta} \right] \cdot \mathbb{1}_{\{u \geq 0\}} \\
&= \boxed{\frac{u}{\theta^2} e^{-u/\theta} \cdot \mathbb{1}_{\{u \geq 0\}}}
\end{aligned}
$$

- This is, in fact, the density of a distribution we've encountered before - specifically, it is the density of the Gamma$(2, \theta)$ distribution!

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Result

Theorem (Sum of IID Exponential Distributions)

Given $Y_1, Y_2 \overset{\text{i.i.d.}}{\sim} \text{Exp}(\theta)$, we have $U := (Y_1 + Y_2) \sim \text{Gamma}(2, \theta)$.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Result

Theorem (Sum of IID Exponential Distributions)

Given $Y_1, Y_2 \overset{\text{i.i.d.}}{\sim} \text{Exp}(\theta)$, we have $U := (Y_1 + Y_2) \sim \text{Gamma}(2, \theta)$.

- This is a very useful result that we will frequently use throughout this course!

UC SANTA BARBARA
Department of Statistics and Applied Probability

Result

Theorem (Sum of IID Exponential Distributions)

Given $Y_1, Y_2 \overset{\text{i.i.d.}}{\sim} \text{Exp}(\theta)$, we have $U := (Y_1 + Y_2) \sim \text{Gamma}(2, \theta)$.

- This is a very useful result that we will frequently use throughout this course!
- Additionally, its proof (which is just the work we did over the past several slides) highlights how we can use the CDF method to find distributions of bivariate transformations.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Chalkboard Exercise

Example

Given $Y_1, Y_2 \overset{\text{i.i.d.}}{\sim} \text{Unif}[0, 1]$, find the densities of $U_1 := Y_1 Y_2$ and $U_2 := Y_1 + Y_2$.

Leadup

- Now, there is actually another way we could have proved the "Sum of IID Exponential Distributions" result.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Leadup

- Now, there is actually another way we could have proved the "Sum of IID Exponential Distributions" result.
- Specifically, note that if $Y_1 \perp Y_2$, then

$$M_{Y_1+Y_2}(t) := \mathbb{E}[e^{t(Y_1+Y_2)}] = \mathbb{E}[e^{tY_1}] \cdot \mathbb{E}[e^{tY_2}] = M_{Y_1}(t) \cdot M_{Y_2}(t)$$

Leadup

- Now, there is actually another way we could have proved the "Sum of IID Exponential Distributions" result.
- Specifically, note that if $Y_1 \perp Y_2$, then

$$M_{Y_1+Y_2}(t) := \mathbb{E}[e^{t(Y_1+Y_2)}] = \mathbb{E}[e^{tY_1}] \cdot \mathbb{E}[e^{tY_2}] = M_{Y_1}(t) \cdot M_{Y_2}(t)$$

- Hence, given $Y_1, Y_2 \overset{\text{i.i.d.}}{\sim} \text{Exp}(\theta)$,

$$\begin{aligned} M_{Y_1+Y_2}(t) &= M_{Y_1}(t) \cdot M_{Y_2}(t) = [M_{Y_1}(t)]^2 \\ &= \left(\begin{cases} (1-\theta t)^{-1} & \text{if } t < 1/\theta \\ \infty & \text{otherwise} \end{cases} \right)^2 = \begin{cases} (1-\theta t)^{-2} & \text{if } t < 1/\theta \\ \infty & \text{otherwise} \end{cases} \end{aligned}$$

which we recognize as the MGF of the Gamma$(2, \theta)$ distribution. Done!

UC SANTA BARBARA
Department of Statistics and Applied Probability

Result

Theorem (Closure of Gamma Distribution under Sums)

Given an independent sequence $\{Y_i\}_{i=1}^n$ of random variables with $Y_i \sim \text{Gamma}(\alpha_i, \beta)$, then

$$U := \left(\sum_{i=1}^n Y_i \right) \sim \text{Gamma} \left(\sum_{i=1}^n \alpha_i \ , \ \beta \right)$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Proof.

$$M_U(t) := \mathbb{E}[e^{tU}] = \mathbb{E}\left[e^{t\sum_{i=1}^{n} Y_i}\right] \qquad \text{[Definition of MGF and } U\text{]}$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Proof.

$$M_U(t) := \mathbb{E}[e^{tU}] = \mathbb{E}\left[e^{t\sum_{i=1}^n Y_i}\right]$$ [Definition of MGF and U]

$$= \mathbb{E}\left[\prod_{i=1}^n e^{tY_i}\right]$$ [Exponential of Sum is Product of Exponentials]

Proof.

$$M_U(t) := \mathbb{E}[e^{tU}] = \mathbb{E}\left[e^{t\sum_{i=1}^{n} Y_i}\right]$$ [Definition of MGF and *U*]

$$= \mathbb{E}\left[\prod_{i=1}^{n} e^{tY_i}\right]$$ [Exponential of Sum is Product of Exponentials]

$$= \prod_{i=1}^{n} \mathbb{E}[e^{tY_i}]$$ [Independence]

UC SANTA BARBARA
Department of Statistics and Applied Probability

Proof.

$$M_U(t) := \mathbb{E}[e^{tU}] = \mathbb{E}\left[e^{t\sum_{i=1}^n Y_i}\right] \qquad \text{[Definition of MGF and } U\text{]}$$

$$= \mathbb{E}\left[\prod_{i=1}^n e^{tY_i}\right] \qquad \text{[Exponential of Sum is Product of Exponentials]}$$

$$= \prod_{i=1}^n \mathbb{E}[e^{tY_i}] \qquad \text{[Independence]}$$

$$= \prod_{i=1}^n M_{Y_i}(t) \qquad \text{[Definition of MGF]}$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Proof.

$$M_U(t) := \mathbb{E}[e^{tU}] = \mathbb{E}\left[e^{t\sum_{i=1}^n Y_i}\right] \qquad \text{[Definition of MGF and } U\text{]}$$

$$= \mathbb{E}\left[\prod_{i=1}^n e^{tY_i}\right] \qquad \text{[Exponential of Sum is Product of Exponentials]}$$

$$= \prod_{i=1}^n \mathbb{E}[e^{tY_i}] \qquad \text{[Independence]}$$

$$= \prod_{i=1}^n M_{Y_i}(t) \qquad \text{[Definition of MGF]}$$

$$= \prod_{i=1}^n \left(\begin{cases} (1-\beta t)^{-\alpha_i} & \text{if } t < 1/\beta \\ \infty & \text{otherwise} \end{cases} \right) \qquad \text{[MGF of } Y_i\text{]}$$

UC SANTA BARBARA
Department of Statistics and Applied Probability

Proof.

$$M_U(t) = \prod_{i=1}^{n} \left(\begin{cases} (1 - \beta t)^{-\alpha_i} & \text{if } t < 1/\beta \\ \infty & \text{otherwise} \end{cases} \right) \qquad \text{[MGF of } Y_i\text{]}$$

$\square$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Proof.

$$M_U(t) = \prod_{i=1}^{n} \left(\begin{cases} (1-\beta t)^{-\alpha_i} & \text{if } t < 1/\beta \\ \infty & \text{otherwise} \end{cases} \right) \qquad \text{[MGF of } Y_i\text{]}$$

$$= \left(\begin{cases} (1-\beta t)^{-\sum_{i=1}^{n} \alpha_i} & \text{if } t < 1/\beta \\ \infty & \text{otherwise} \end{cases} \right) \qquad \text{[Algebra]}$$

$\square$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Proof.

$$M_U(t) = \prod_{i=1}^{n} \left(\begin{cases} (1-\beta t)^{-\alpha_i} & \text{if } t < 1/\beta \\ \infty & \text{otherwise} \end{cases} \right) \qquad \text{[MGF of } Y_i\text{]}$$

$$= \left(\begin{cases} (1-\beta t)^{-\sum_{i=1}^{n} \alpha_i} & \text{if } t < 1/\beta \\ \infty & \text{otherwise} \end{cases} \right) \qquad \text{[Algebra]}$$

- The final step of the proof is to note that this is precisely the MGF of a Gamma($\sum_{i=1}^{n} \alpha_i \ , \ \beta$) distribution.

□

MGF Method, Revisited

- So, perhaps we can now see where the MGF method really shines - in the context of sums and linear combinations of random variables!

UC SANTA BARBARA
Department of Statistics and Applied Probability

MGF Method, Revisited

- So, perhaps we can now see where the MGF method really shines - in the context of sums and linear combinations of random variables!
- As an exercise, try proving the following result:

UC SANTA BARBARA
Department of Statistics
and Applied Probability

MGF Method, Revisited

- So, perhaps we can now see where the MGF method really shines - in the context of sums and linear combinations of random variables!
- As an exercise, try proving the following result:

Theorem (MGF of Linear Combinations)

Consider an independent sequence $\{Y_i\}_{i=1}^n$ of random variables with MGFs $M_{Y_i}(t)$. Given $U := \sum_{i=1}^n a_i Y_i$ for constants $\{a_i\}_{i=1}^n$, we have

$$M_U(t) = \prod_{i=1}^n M_{Y_i}(a_i t)$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Closure of Normal Distribution

- After proving this result, you can verify that the following result also follows:

UC SANTA BARBARA
Department of Statistics and Applied Probability

Closure of Normal Distribution

- After proving this result, you can verify that the following result also follows:

Theorem (Closure of Normal Distribution under Linear Combinations)

Given an independent sequence $\{Y_i\}_{i=1}^n$ of random variables with $Y_i \sim \mathcal{N}(\mu_i\ , \sigma_i^2)$, we have

$$\left(\sum_{i=1}^n \alpha_i Y_i \right) \sim \mathcal{N}\left(\sum_{i=1}^n \alpha_i \mu_i\ , \ \sum_{i=1}^n (\alpha_i \sigma_i)^2 \right)$$

UC SANTA BARBARA
Department of Statistics and Applied Probability

Leadup

- So, we've seen how the CDF method can be extended to allow us to derive distributions of multivariate transformations.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Leadup

- So, we’ve seen how the CDF method can be extended to allow us to derive distributions of multivariate transformations.
- We’ve now also seen how the MGF method can be used to help us explore the distributions of multivariate transformations.

Leadup

- So, we've seen how the CDF method can be extended to allow us to derive distributions of multivariate transformations.
- We've now also seen how the MGF method can be used to help us explore the distributions of multivariate transformations.
- A natural question arises: is there an analog of the Change of Variable formula?

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Leadup

- So, we've seen how the CDF method can be extended to allow us to derive distributions of multivariate transformations.
- We've now also seen how the MGF method can be used to help us explore the distributions of multivariate transformations.
- A natural question arises: is there an analog of the Change of Variable formula?
- The answer is, "yes - for bivariate transformations."

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Leadup

- So, we've seen how the CDF method can be extended to allow us to derive distributions of multivariate transformations.
- We've now also seen how the MGF method can be used to help us explore the distributions of multivariate transformations.
- A natural question arises: is there an analog of the Change of Variable formula?
- The answer is, "yes - for bivariate transformations."
- This method is often referred to as the **Jacobian Method**.

Leadup

- I admit that the Jacobian Method can seem a bit daunting at first.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Leadup

- I admit that the Jacobian Method can seem a bit daunting at first.
- So, allow me to simply state the formula for you, make a few comments, and then walk through an example (which I hope should help demystify things a bit).

UC SANTA BARBARA
Department of Statistics and Applied Probability

Jacobian Method

Theorem (The Jacobian Method)

Consider $Y_1, Y_2 \sim f_{Y_1,Y_2}$, and random variables

$$U_1 := h_1(Y_1, Y_2); \qquad U_2 := h_2(Y_1, Y_2)$$

with inverse transformations

$$Y_1 = h_1^{-1}(U_1, U_2); \qquad Y_2 = h_2^{-1}(U_1, U_2)$$

The joint density of U_1, U_2 is given by

$$f_{U_1,U_2}(u_1, u_2) = f_{Y_1,Y_2}\left(h_1^{-1}(u_1, u_2)\ ,\ h_2^{-1}(u_1, u_2)\right) \cdot |J|; \quad J := \begin{vmatrix} \frac{\partial h_1^{-1}}{\partial u_1} & \frac{\partial h_1^{-1}}{\partial u_2} \\ \frac{\partial h_2^{-1}}{\partial u_1} & \frac{\partial h_2^{-1}}{\partial u_2} \end{vmatrix}$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Some Comments

- First, the quantity J is called a **Jacobian** (hence why this method is called the Jacobian Method)

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Some Comments

- First, the quantity J is called a **Jacobian** (hence why this method is called the Jacobian Method)
- Additionally, recall that

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \quad \text{is the } \underline{\textbf{determinant}} \text{ of a } 2 \times 2 \text{ matrix.}$$

Some Comments

- First, the quantity J is called a **Jacobian** (hence why this method is called the Jacobian Method)
- Additionally, recall that

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \quad \text{is the } \textbf{determinant} \text{ of a } 2 \times 2 \text{ matrix.}$$

- Hence,

$$J := \begin{vmatrix} \frac{\partial h_1^{-1}}{\partial u_1} & \frac{\partial h_1^{-1}}{\partial u_2} \\ \frac{\partial h_2^{-1}}{\partial u_1} & \frac{\partial h_2^{-1}}{\partial u_2} \end{vmatrix} = \left(\frac{\partial h_1^{-1}}{\partial u_1} \right) \left(\frac{\partial h_2^{-1}}{\partial u_2} \right) - \left(\frac{\partial h_1^{-1}}{\partial u_2} \right) \left(\frac{\partial h_2^{-1}}{\partial u_1} \right)$$

Some Comments

- First, the quantity J is called a **Jacobian** (hence why this method is called the Jacobian Method)
- Additionally, recall that

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \quad \text{is the } \textbf{determinant} \text{ of a } 2 \times 2 \text{ matrix.}$$

- Hence,

$$J := \begin{vmatrix} \frac{\partial h_1^{-1}}{\partial u_1} & \frac{\partial h_1^{-1}}{\partial u_2} \\ \frac{\partial h_2^{-1}}{\partial u_1} & \frac{\partial h_2^{-1}}{\partial u_2} \end{vmatrix} = \left(\frac{\partial h_1^{-1}}{\partial u_1} \right) \left(\frac{\partial h_2^{-1}}{\partial u_2} \right) - \left(\frac{\partial h_1^{-1}}{\partial u_2} \right) \left(\frac{\partial h_2^{-1}}{\partial u_1} \right)$$

  - It's just a bit more succinct to write this as a determinant, as opposed to writing it in its expanded form (like we did above).

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Example

Example

Given $Y_1, Y_2 \overset{\text{i.i.d.}}{\sim} \text{Exp}(\theta)$, find the joint density of (U_1, U_2) where

$$U_1 := \frac{Y_1}{Y_1 + Y_2}; \qquad U_2 := Y_1 + Y_2$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Example

Example

Given $Y_1, Y_2 \overset{\text{i.i.d.}}{\sim} \text{Exp}(\theta)$, find the joint density of (U_1, U_2) where

$$U_1 := \frac{Y_1}{Y_1 + Y_2}; \qquad U_2 := Y_1 + Y_2$$

- When applying the method of Jacobian transformations, a good first step is always to find the inverse transformations. This amounts to solving the system of equations

$$\begin{cases} u_1 = h_1(y_1, y_2) \\ u_2 = h_2(y_1, y_2) \end{cases} \qquad \text{for } y_1, y_2 \text{ in terms of } u_1, u_2$$

Example

- In this example, $h_1(y_1, y_2) = y_1/(y_1 + y_2)$ and $h_2(y_1, y_2) = y_1 + y_2$. Hence, we seek to solve:

$$\begin{cases} u_1 = \frac{y_1}{y_1+y_2} \\ u_2 = y_1 + y_2 \end{cases}$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Example

- In this example, $h_1(y_1, y_2) = y_1/(y_1 + y_2)$ and $h_2(y_1, y_2) = y_1 + y_2$. Hence, we seek to solve:

$$\begin{cases} u_1 = \frac{y_1}{y_1+y_2} \\ u_2 = y_1 + y_2 \end{cases}$$

- Substituting the second equation into the first yields

$$u_1 = \frac{y_1}{u_2} \quad \Longrightarrow \quad y_1 = u_1 u_2 =: h_1^{-1}(u_1, u_2)$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Example

- In this example, $h_1(y_1, y_2) = y_1/(y_1 + y_2)$ and $h_2(y_1, y_2) = y_1 + y_2$. Hence, we seek to solve:

$$\begin{cases} u_1 = \frac{y_1}{y_1+y_2} \\ u_2 = y_1 + y_2 \end{cases}$$

- Substituting the second equation into the first yields

$$u_1 = \frac{y_1}{u_2} \implies y_1 = u_1 u_2 =: h_1^{-1}(u_1, u_2)$$

- Then, the second equation yields

$$y_2 = u_2 - y_1 = u_2 - u_1 u_2 = u_2(1 - u_1) =: h_2^{-1}(u_1, u_2)$$

UC SANTA BARBARA
Department of Statistics and Applied Probability

Example

- So, we have $h_1^{-1}(u_1, u_2) = u_1 u_2$ and $h_2^{-1}(u_1, u_2) = u_2(1 - u_1)$.

Example

- So, we have $h_1^{-1}(u_1, u_2) = u_1 u_2$ and $h_2^{-1}(u_1, u_2) = u_2(1 - u_1)$.
- Now, let's take some partial derivatives:

$$\frac{\partial}{\partial u_1} h_1^{-1} = \frac{\partial}{\partial u_1}[u_1 u_2] = u_2$$
$$\frac{\partial}{\partial u_2} h_1^{-1} = \frac{\partial}{\partial u_2}[u_1 u_2] = u_1$$
$$\frac{\partial}{\partial u_1} h_2^{-1} = \frac{\partial}{\partial u_1}[u_2(1 - u_1)] = -u_2$$
$$\frac{\partial}{\partial u_2} h_2^{-1} = \frac{\partial}{\partial u_2}[u_2(1 - u_1)] = 1 - u_1$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Example

- The Jacobian is therefore:

$$J := \begin{vmatrix} \dfrac{\partial h_1^{-1}}{\partial u_1} & \dfrac{\partial h_1^{-1}}{\partial u_2} \\ \\ \dfrac{\partial h_2^{-1}}{\partial u_1} & \dfrac{\partial h_2^{-1}}{\partial u_2} \end{vmatrix}$$

$$= \begin{vmatrix} u_2 & u_1 \\ \\ -u_2 & 1 - u_1 \end{vmatrix} = u_2(1 - u_1) + u_1 u_2 = u_2$$

Example

- Hence, by the formula provided in the theorem above:

$$
\begin{aligned}
f_{U_1,U_2}(u_1,u_2) &= f_{Y_1,Y_2}\left(h_1^{-1}(u_1,u_2)\ ,\ h_2^{-1}(u_1,u_2)\right)\cdot|J| \\
&= f_{Y_1}(u_1u_2)\cdot f_{Y_2}(u_2(1-u_1))\cdot|u_2| \\
&= \left[\frac{1}{\theta}e^{-u_1u_2/\theta}\cdot \mathbb{1}_{\{u_1u_2\geq 0\}}\right]\cdot\left[\frac{1}{\theta}e^{-u_2(1-u_1)/\theta}\cdot \mathbb{1}_{\{u_2(1-u_1)\geq 0\}}\right]\cdot u_2 \\
&= \boxed{\frac{u_2}{\theta^2}e^{-u_2/\theta}\cdot \mathbb{1}_{\{u_2\geq 0\}}\cdot \mathbb{1}_{\{0\leq u_1\leq 1\}}}
\end{aligned}
$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Example

- Again, what we have shown is that

$$f_{U_1,U_2}(u_1, u_2) = \frac{u_2}{\theta^2} e^{-u_2/\theta} \cdot \mathbb{1}_{\{u_2 \geq 0\}} \cdot \mathbb{1}_{\{0 \leq u_1 \leq 1\}}$$

Example

- Again, what we have shown is that

$$f_{U_1,U_2}(u_1, u_2) = \frac{u_2}{\theta^2} e^{-u_2/\theta} \cdot \mathbb{1}_{\{u_2 \geq 0\}} \cdot \mathbb{1}_{\{0 \leq u_1 \leq 1\}}$$

- Notice that this is of the form (something involving only u_2) times (something involving only u_1). This allows us to conclude that $U_1 \perp U_2$!

Example

- Again, what we have shown is that

$$f_{U_1,U_2}(u_1, u_2) = \frac{u_2}{\theta^2} e^{-u_2/\theta} \cdot \mathbb{1}_{\{u_2 \geq 0\}} \cdot \mathbb{1}_{\{0 \leq u_1 \leq 1\}}$$

- Notice that this is of the form (something involving only u_2) times (something involving only u_1). This allows us to conclude that $U_1 \perp U_2$!
- Not only that, we can simply read off the marginal densities from our joint densities: perhaps unsurprisingly, $U_2 \sim \text{Gamma}(2, \theta)$ but perhaps surprisingly

$$U_1 := \left(\frac{Y_1}{Y_1 + Y_2} \right) \sim \text{Unif}[0, 1]$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Marginals

- This example, in addition to revealing a very useful result, also illustrates how the Jacobian Method can be used to find distributions of multivariate transformaions.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Marginals

- This example, in addition to revealing a very useful result, also illustrates how the Jacobian Method can be used to find distributions of multivariate transformaions.
- Specifically, we found the distribution of $U_1 := Y_1/(Y_2 + Y_1)$ by first finding the joint density of (U_1, U_2) and then effectively integrating out u_2 to find the marginal density of U_1.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Marginals

- This example, in addition to revealing a very useful result, also illustrates how the Jacobian Method can be used to find distributions of multivariate transformaions.
- Specifically, we found the distribution of $U_1 := Y_1/(Y_2 + Y_1)$ by first finding the joint density of (U_1, U_2) and then effectively integrating out u_2 to find the marginal density of U_1.
- Let's see this in action again:

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Convolution

Theorem (Convolution Formula)

Given $(Y_1, Y_2) \sim f_{Y_1,Y_2}$, the density of $S := Y_1 + Y_2$ is given by

$$f_S(s) = \int_{-\infty}^{\infty} f_{Y_1,Y_2}(s - t\ ,\ t)\ \mathrm{d}t$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Convolution

Theorem (Convolution Formula)

Given $(Y_1, Y_2) \sim f_{Y_1,Y_2}$, the density of $S := Y_1 + Y_2$ is given by

$$f_S(s) = \int_{-\infty}^{\infty} f_{Y_1,Y_2}(s - t\ ,\ t)\ \mathrm{d}t$$

- Note that we could have proved this using the CDF method (and, indeed, you will do so on a homework problem).

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Convolution

Theorem (Convolution Formula)

Given $(Y_1, Y_2) \sim f_{Y_1,Y_2}$, the density of $S := Y_1 + Y_2$ is given by

$$f_S(s) = \int_{-\infty}^{\infty} f_{Y_1,Y_2}(s - t\ ,\ t)\ \mathrm{d}t$$

- Note that we could have proved this using the CDF method (and, indeed, you will do so on a homework problem).
- But, for practice, let's use the Jacobian Method.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Proof

- When using the Jacobian Method to find the distribution of a multivariate transformation S, it's customary to introduce a **auxilliary** random variable T so that we can use the Jacobian Method to first find the joint density $f_{S,T}(s, t)$ of (S, T) and then integrate out t.

Proof

- When using the Jacobian Method to find the distribution of a multivariate transformation S, it's customary to introduce a **auxilliary** random variable T so that we can use the Jacobian Method to first find the joint density $f_{S,T}(s,t)$ of (S,T) and then integrate out t.
- There is no unique choice for T! For now, let's simply take $T = Y_2$.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Proof

- When using the Jacobian Method to find the distribution of a multivariate transformation S, it's customary to introduce a **auxilliary** random variable T so that we can use the Jacobian Method to first find the joint density $f_{S,T}(s,t)$ of (S,T) and then integrate out t.
- There is no unique choice for T! For now, let's simply take $T = Y_2$.
- Now, we find the inverse transformations:

$$\begin{cases} s = y_1 + y_2 & \implies y_1 = s - y_2 = s - t =: h_1^{-1}(s,t) \\ t = y_2 & \implies y_2 = t =: h_2^{-1}(s,t) \end{cases}$$

Proof

- When using the Jacobian Method to find the distribution of a multivariate transformation S, it's customary to introduce a **auxilliary** random variable T so that we can use the Jacobian Method to first find the joint density $f_{S,T}(s, t)$ of (S, T) and then integrate out t.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Proof

- When using the Jacobian Method to find the distribution of a multivariate transformation S, it's customary to introduce a **auxilliary** random variable T so that we can use the Jacobian Method to first find the joint density $f_{S,T}(s, t)$ of (S, T) and then integrate out t.
- There is no unique choice for T! For now, let's simply take $T = Y_2$.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Proof

- When using the Jacobian Method to find the distribution of a multivariate transformation S, it's customary to introduce a **auxilliary** random variable T so that we can use the Jacobian Method to first find the joint density $f_{S,T}(s, t)$ of (S, T) and then integrate out t.
- There is no unique choice for T! For now, let's simply take $T = Y_2$.
- Now, we find the inverse transformations:

$$\begin{cases} s = y_1 + y_2 & \implies y_1 = s - y_2 = s - t =: h_1^{-1}(s, t) \\ t = y_2 & \implies y_2 = t =: h_2^{-1}(s, t) \end{cases}$$

Proof

- So, we have $h_1^{-1}(s, t) = s - t$ and $h_2^{-1}(s, t) = t$.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Proof

- So, we have $h_1^{-1}(s,t) = s - t$ and $h_2^{-1}(s,t) = t$.
- Now, let's take some partial derivatives:

$$\begin{aligned}
\frac{\partial}{\partial s} h_1^{-1}(s,t) &= \frac{\partial}{\partial s}[s - t] = 1 \\
\frac{\partial}{\partial t} h_1^{-1}(s,t) &= \frac{\partial}{\partial t}[s - t] = -1 \\
\frac{\partial}{\partial s} h_2^{-1}(s,t) &= \frac{\partial}{\partial s}[t] = 0 \\
\frac{\partial}{\partial t} h_2^{-1}(s,t) &= \frac{\partial}{\partial t}[t] = 1
\end{aligned}$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Proof

- So, the Jacobian is

$$J := \begin{vmatrix} \dfrac{\partial h_1^{-1}}{\partial s} & \dfrac{\partial h_1^{-1}}{\partial t} \\ \\ \dfrac{\partial h_2^{-1}}{\partial s} & \dfrac{\partial h_2^{-1}}{\partial t} \end{vmatrix} = \begin{vmatrix} 1 & -1 \\ \\ 0 & 1 \end{vmatrix} = 1$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Proof

- So, the Jacobian is

$$J := \begin{vmatrix} \dfrac{\partial h_1^{-1}}{\partial s} & \dfrac{\partial h_1^{-1}}{\partial t} \\ & \\ \dfrac{\partial h_2^{-1}}{\partial s} & \dfrac{\partial h_2^{-1}}{\partial t} \end{vmatrix} = \begin{vmatrix} 1 & -1 \\ & \\ 0 & 1 \end{vmatrix} = 1$$

- Therefore, the joint density of (S, T) is:

$$\begin{aligned} f_{S,T}(s,t) &= f_{Y_1,Y_2}\left(h_1^{-1}(s,t)\ ,\ h_2^{-1}(s,t)\right) \cdot |J| \\ &= f_{Y_1,Y_2}(s - t\ ,\ t) \end{aligned}$$

UC SANTA BARBARA
Department of Statistics and Applied Probability

Proof

- Finally, we obtain the marginal density of S by integrating out t:

$$
\begin{aligned}
f_S(s) &= \int_{-\infty}^{\infty} f_{S,T}(s, t) \, \mathrm{d}t \\
&= \int_{-\infty}^{\infty} f_{Y_1,Y_2}(s - t, t) \, \mathrm{d}t
\end{aligned}
$$

UC SANTA BARBARA
Department of Statistics
and Applied Probability

Proof

- Finally, we obtain the marginal density of S by integrating out t:

$$
\begin{aligned}
f_S(s) &= \int_{-\infty}^{\infty} f_{S,T}(s, t) \, \mathrm{d}t \\
&= \int_{-\infty}^{\infty} f_{Y_1,Y_2}(s - t, t) \, \mathrm{d}t
\end{aligned}
$$

- This completes the proof!

UC SANTA BARBARA
Department of Statistics and Applied Probability

Proof

- Finally, we obtain the marginal density of S by integrating out t:

$$
\begin{aligned}
f_S(s) &= \int_{-\infty}^{\infty} f_{S,T}(s, t) \, \mathrm{d}t \\
&= \int_{-\infty}^{\infty} f_{Y_1,Y_2}(s - t, t) \, \mathrm{d}t
\end{aligned}
$$

- This completes the proof!
- A quick side note: often times, when using this formula, the bounds of integration will depend on s. There's an example of this on one of the Discussion Worksheets.

A Word of Caution

- One thing I should mention (that I did not mention before) is that, much like the Change of Variable formula, the Jacobian Method really only works when we are dealing with strictly monotone functions.

UC SANTA BARBARA
Department of Statistics
and Applied Probability

A Word of Caution

- One thing I should mention (that I did not mention before) is that, much like the Change of Variable formula, the Jacobian Method really only works when we are dealing with strictly monotone functions.
- Exactly what that means for multivariate transformations is a bit complicated. Suffice it to say: when solving the system

$$\begin{cases} u_1 & = h_1(y_1, y_2) \\ h_2 & = h_2(y_1, y_2) \end{cases}$$

  check to ensure that that only *one* solution [for (y_1, y_2) in terms of (u_1, u_2)] exists.

A Word of Caution

- One thing I should mention (that I did not mention before) is that, much like the Change of Variable formula, the Jacobian Method really only works when we are dealing with strictly monotone functions.
- Exactly what that means for multivariate transformations is a bit complicated. Suffice it to say: when solving the system

$$\begin{cases} u_1 & = h_1(y_1, y_2) \\ h_2 & = h_2(y_1, y_2) \end{cases}$$

  check to ensure that that only *one* solution [for (y_1, y_2) in terms of (u_1, u_2)] exists.
    - If more than one solution exists, there *is* a way to modify the Jacobian Method to work, but for the purposes of this class you should simply resort to another method.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Summary

- So, to quickly recap what we've discussed on Multivariate Transformations.

UC SANTA BARBARA
Department of Statistics and Applied Probability

Summary

- So, to quickly recap what we've discussed on Multivariate Transformations.
- For Bivariate Transformations (i.e. transformations of exactly two random variables), we can use either the CDF method or the Jacobian method.

Summary

- So, to quickly recap what we've discussed on Multivariate Transformations.
- For Bivariate Transformations (i.e. transformations of exactly two random variables), we can use either the CDF method or the Jacobian method.
- For linear combinations of any number of random variables, the MGF method is best-suited.