Synchronous Online: Start 12pm, Wednesday 2nd December

MAS2905: Probability and Bayesian Methods with R

Newcastle University
lee.fawcett@ncl.ac.uk

Semester 1, 2020/21

The story so far...

- You should now have completed **all of Chapter 4**
- This includes videos for **Home Study 4** and **Home Study 5** and the **Computer Practical** from last week
- You should also have completed **Home Study 6** which covers some of the theory from **Chapter 5**, which we will continue with today
- Don't forget: You should now be making some inroads into the project!

The story so far...

- You should now have completed **all of Chapter 4**
- This includes videos for **Home Study 4** and **Home Study 5** and the **Computer Practical** from last week
- You should also have completed **Home Study 6** which covers some of the theory from **Chapter 5**, which we will continue with today
- Don't forget: You should now be making some inroads into the project!

5. Transformation of random variables

- In Chapter 5 of the course notes we look at how to obtain the distribution of a **transformation** of a random variable
- We can do this two ways:
  - Informally, using **simulation** in R
  - More formally, by deriving the distribution of the transformed random variable through the **Transformation Method** (see proof on page 42)
- We also thought about **why** we might want to explore the distribution of a transformed random variable (e.g. variables that might only be defined over the positive real line)

5. Transformation of random variables

- In Chapter 5 of the course notes we look at how to obtain the distribution of a **transformation** of a random variable
- We can do this two ways:
  - Informally, using **simulation** in R
  - More formally, by deriving the distribution of the transformed random variable through the **Transformation Method** (see proof on page 42)
- We also thought about **why** we might want to explore the distribution of a transformed random variable (e.g. variables that might only be defined over the positive real line)

5. Transformation of random variables

- In Chapter 5 of the course notes we look at how to obtain the distribution of a **transformation** of a random variable
- We can do this two ways:
  - Informally, using **simulation** in R
  - More formally, by deriving the distribution of the transformed random variable through the **Transformation Method** (see proof on page 42)
- We also thought about **why** we might want to explore the distribution of a transformed random variable (e.g. variables that might only be defined over the positive real line)

5. Transformation of random variables

- In Chapter 5 of the course notes we look at how to obtain the distribution of a **transformation** of a random variable
- We can do this two ways:
  - Informally, using **simulation** in R
  - More formally, by deriving the distribution of the transformed random variable through the **Transformation Method** (see proof on page 42)
- We also thought about **why** we might want to explore the distribution of a transformed random variable (e.g. variables that might only be defined over the positive real line)

5. Transformation of random variables

- In Chapter 5 of the course notes we look at how to obtain the distribution of a **transformation** of a random variable
- We can do this two ways:
  - Informally, using **simulation** in R
  - More formally, by deriving the distribution of the transformed random variable through the **Transformation Method** (see proof on page 42)
- We also thought about **why** we might want to explore the distribution of a transformed random variable (e.g. variables that might only be defined over the positive real line)

5.3 The transformation method

Recall (from page 42) that if X has PDF $f_X(x)$ and $h(x)$ is a monotonically increasing or monotonically decreasing function then

- h is invertible and
- the PDF of $Y = h(X)$ is

$$f_Y(y) = f_X(h^{-1}(y)) \times \left| \frac{d\left(h^{-1}(y)\right)}{dy} \right|$$

when y is in the range of $h(x)$ and zero otherwise.

This result allows direct calculation of $f_Y(y)$ and is called the univariate transformation method.

Example 5.4

Example

When $X \sim U(0,1)$, use the transformation method to find the PDF of $Y = X^3$.

$$h(x) = x^3, \quad h^{-1}(y) = y^{1/3}.$$

$$\begin{aligned} f_Y(y) &= 1 \times \frac{d}{dy} y^{1/3} \\ &= \frac{1}{3} y^{-2/3}, \quad 0 \le y \le 1. \end{aligned}$$

Example 5.4

Example

When $X \sim U(0,1)$, use the transformation method to find the PDF of $Y = X^3$.

$$h(x) = x^3, \quad h^{-1}(y) = y^{1/3}.$$

$$\begin{aligned} f_Y(y) &= 1 \times \frac{d}{dy} y^{1/3} \\ &= \frac{1}{3} y^{-2/3}, \quad 0 \le y \le 1. \end{aligned}$$

Example 5.4

Example

When $X \sim U(0,1)$, use the transformation method to find the PDF of $Y = X^3$.

$$h(x) = x^3, \quad h^{-1}(y) = y^{1/3}.$$

$$\begin{aligned} f_Y(y) &= 1 \times \frac{d}{dy} y^{1/3} \\ &= \frac{1}{3} y^{-2/3}, \quad 0 \le y \le 1. \end{aligned}$$

Example 5.4

Example

When $X \sim U(0,1)$, use the transformation method to find the PDF of $Y = X^3$.

$$h(x) = x^3, \quad h^{-1}(y) = y^{1/3}.$$

$$\begin{aligned} f_Y(y) &= 1 \times \frac{d}{dy} y^{1/3} \\ &= \frac{1}{3} y^{-2/3}, \quad 0 \le y \le 1. \end{aligned}$$

Example 5.5

Example

When $X \sim U(-1, 1)$ and $Y = \exp(X)$, use the transformation method to find the PDF of Y.

$$h(x) = e^x, \quad h^{-1}(y) = \log(y).$$

$$\begin{aligned} f_Y(y) &= \frac{1}{2} \times \left| \frac{d}{dy} \log(y) \right| \\ &= \frac{1}{2|y|} = \frac{1}{2y}, \quad e^{-1} \leq y \leq e^1, \text{ zero otherwise.} \end{aligned}$$

Example 5.5

Example

When $X \sim U(-1, 1)$ and $Y = \exp(X)$, use the transformation method to find the PDF of Y.

$$h(x) = e^x, \quad h^{-1}(y) = \log(y).$$

$$\begin{aligned} f_Y(y) &= \frac{1}{2} \times \left| \frac{d}{dy} \log(y) \right| \\ &= \frac{1}{2|y|} = \frac{1}{2y}, \quad e^{-1} \leq y \leq e^1, \text{ zero otherwise.} \end{aligned}$$

Example 5.5

Example

When $X \sim U(-1, 1)$ and $Y = \exp(X)$, use the transformation method to find the PDF of Y.

$$h(x) = e^x, \quad h^{-1}(y) = \log(y).$$

$$\begin{aligned} f_Y(y) &= \frac{1}{2} \times \left| \frac{d}{dy} \log(y) \right| \\ &= \frac{1}{2|y|} = \frac{1}{2y}, \quad e^{-1} \leq y \leq e^1, \text{ zero otherwise.} \end{aligned}$$

Example 5.5

Example

When $X \sim U(-1, 1)$ and $Y = \exp(X)$, use the transformation method to find the PDF of Y.

$$h(x) = e^x, \quad h^{-1}(y) = \log(y).$$

$$\begin{aligned} f_Y(y) &= \frac{1}{2} \times \left| \frac{d}{dy} \log(y) \right| \\ &= \frac{1}{2|y|} = \frac{1}{2y}, \quad e^{-1} \leq y \leq e^1, \text{ zero otherwise.} \end{aligned}$$

Example 5.6

Example

Use the transformation method to find the PDF of the log-normal distribution.

$$h(x) = e^x, \quad h^{-1}(y) = \log(y).$$

$$\begin{aligned} f_Y(y) &= f_X(\log(y)) \times \frac{1}{y} \qquad \text{for } y > 0 \\ &= \frac{1}{y} \times \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\log(y)-\mu}{\sigma}\right)^2\right]. \end{aligned}$$

Example 5.6

Example

Use the transformation method to find the PDF of the log-normal distribution.

$$h(x) = e^x, \quad h^{-1}(y) = \log(y).$$

$$\begin{aligned} f_Y(y) &= f_X(\log(y)) \times \frac{1}{y} \qquad \text{for } y > 0 \\ &= \frac{1}{y} \times \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\log(y)-\mu}{\sigma}\right)^2\right]. \end{aligned}$$

Example 5.6

Example

Use the transformation method to find the PDF of the log-normal distribution.

$$h(x) = e^x, \quad h^{-1}(y) = \log(y).$$

$$\begin{aligned} f_Y(y) &= f_X(\log(y)) \times \frac{1}{y} \qquad \text{for } y > 0 \\ &= \frac{1}{y} \times \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\log(y)-\mu}{\sigma}\right)^2\right]. \end{aligned}$$

Example 5.6

Example

Use the transformation method to find the PDF of the log-normal distribution.

$$h(x) = e^x, \quad h^{-1}(y) = \log(y).$$

$$\begin{aligned} f_Y(y) &= f_X(\log(y)) \times \frac{1}{y} \qquad \text{for } y > 0 \\ &= \frac{1}{y} \times \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\log(y)-\mu}{\sigma}\right)^2\right]. \end{aligned}$$

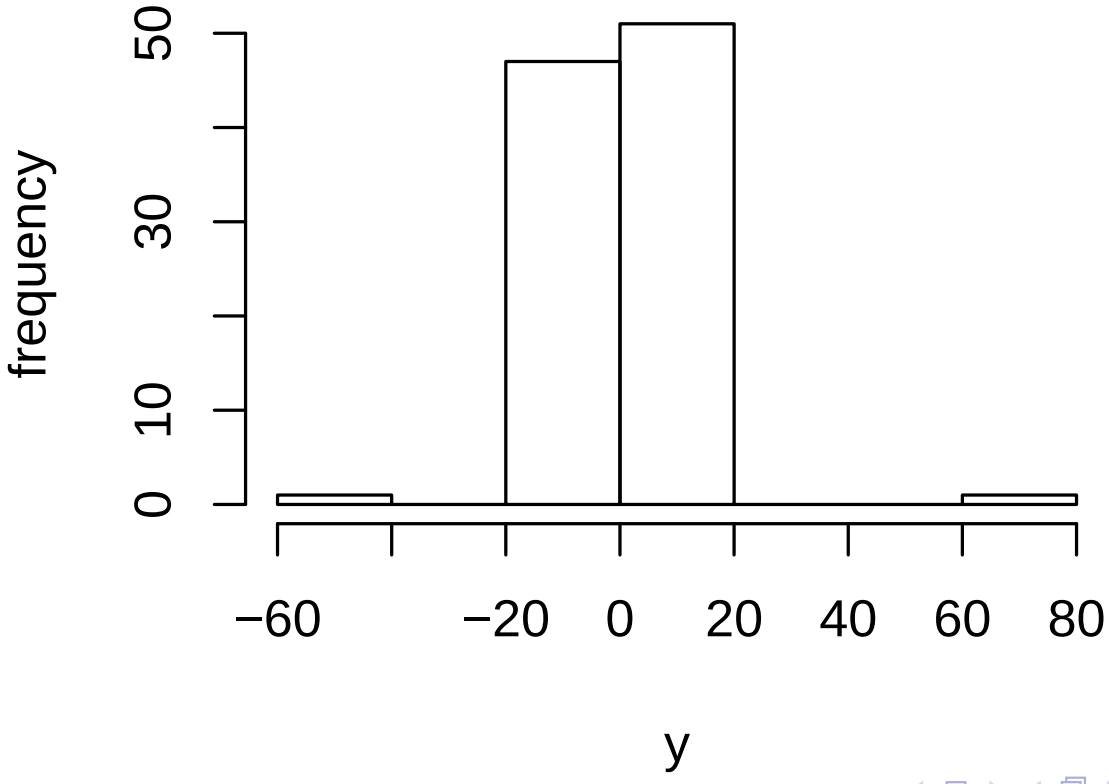
Example 5.7

Example

Suppose $X \sim U(-\frac{\pi}{2}, \frac{\pi}{2})$ and $Y = \tan(X)$.
Use simulation to study the distribution of Y and explicitly compute the PDF.

Example 5.7: Solution

```
y = tan(runif(100, -pi/2, pi/2))
hist(y, xlab = 'y', ylab = 'frequency', main = '')
```



Solution

This distribution looks odd! Calculate the PDF:

$$h(x) = \tan(x), \quad h^{-1}(y) = \tan^{-1}(y).$$

$$\begin{aligned} f_Y(y) &= \frac{1}{\pi}\frac{d}{dy}\left(\tan^{-1} y\right) \quad -\infty < y < \infty, \\ &= \frac{1}{\pi(1+y^2)} \quad -\infty < y < \infty. \end{aligned}$$

This is the **Cauchy distribution**. It has *very* heavy tails.

Solution

This distribution looks odd! Calculate the PDF:

$$h(x) = \tan(x), \quad h^{-1}(y) = \tan^{-1}(y).$$

$$\begin{aligned} f_Y(y) &= \frac{1}{\pi}\frac{d}{dy}\left(\tan^{-1} y\right) \quad -\infty < y < \infty, \\ &= \frac{1}{\pi(1+y^2)} \quad -\infty < y < \infty. \end{aligned}$$

This is the **Cauchy distribution**. It has *very* heavy tails.

Solution

This distribution looks odd! Calculate the PDF:

$$h(x) = \tan(x), \quad h^{-1}(y) = \tan^{-1}(y).$$

$$\begin{aligned} f_Y(y) &= \frac{1}{\pi}\frac{d}{dy}\left(\tan^{-1} y\right) \quad -\infty < y < \infty, \\ &= \frac{1}{\pi(1+y^2)} \quad -\infty < y < \infty. \end{aligned}$$

This is the **Cauchy distribution**. It has *very* heavy tails.

Solution

This distribution looks odd! Calculate the PDF:

$$h(x) = \tan(x), \quad h^{-1}(y) = \tan^{-1}(y).$$

$$\begin{aligned} f_Y(y) &= \frac{1}{\pi}\frac{d}{dy}\left(\tan^{-1} y\right) \quad -\infty < y < \infty, \\ &= \frac{1}{\pi(1+y^2)} \quad -\infty < y < \infty. \end{aligned}$$

This is the **Cauchy distribution**. It has *very* heavy tails.

Solution

This distribution looks odd! Calculate the PDF:

$$h(x) = \tan(x), \quad h^{-1}(y) = \tan^{-1}(y).$$

$$\begin{aligned} f_Y(y) &= \frac{1}{\pi}\frac{d}{dy}\left(\tan^{-1} y\right) \quad -\infty < y < \infty, \\ &= \frac{1}{\pi(1+y^2)} \quad -\infty < y < \infty. \end{aligned}$$

This is the **Cauchy distribution**. It has *very* heavy tails.

Bonus example

Hourly wind speed maxima (X miles per hour), observed on the roof of Claremont Tower as part of the University's *Urban Observatory*, are assumed exponentially distributed with rate 0.1.

Unfortunately, the data recorded are the logged wind speeds.

Use the transformation method to find the PDF of $Y = \log X$.

Bonus example: Solution

Recall that if $X \sim Exp(\lambda)$, then it has PDF

$$f_X(x) = \lambda e^{-\lambda x}, \quad x > 0, \quad \lambda > 0.$$

Now here, we have

$$h(x) = \log x, \quad \text{and} \quad h^{-1}(y) = e^y.$$

Therefore

$$\begin{aligned} f_Y(y) &= f_X(h^{-1}(y)) \times \left| \frac{d\left(h^{-1}(y)\right)}{dy} \right| \\ &= \lambda \exp\left\{-\lambda \exp\{y\}\right\} \times \exp\{y\} \\ &= \lambda \exp\left\{y - \lambda e^y\right\}, \end{aligned}$$

with $\lambda = 0.1$.

Bonus example: Solution

Recall that if $X \sim Exp(\lambda)$, then it has PDF

$$f_X(x) = \lambda e^{-\lambda x}, \quad x > 0, \quad \lambda > 0.$$

Now here, we have

$$h(x) = \log x, \quad \text{and} \quad h^{-1}(y) = e^y.$$

Therefore

$$\begin{aligned} f_Y(y) &= f_X(h^{-1}(y)) \times \left| \frac{d\left(h^{-1}(y)\right)}{dy} \right| \\ &= \lambda \exp\left\{-\lambda \exp\{y\}\right\} \times \exp\{y\} \\ &= \lambda \exp\left\{y - \lambda e^y\right\}, \end{aligned}$$

with $\lambda = 0.1$.

Bonus example: Solution

Recall that if $X \sim Exp(\lambda)$, then it has PDF

$$f_X(x) = \lambda e^{-\lambda x}, \quad x > 0, \quad \lambda > 0.$$

Now here, we have

$$h(x) = \log x, \quad \text{and} \quad h^{-1}(y) = e^y.$$

Therefore

$$\begin{aligned} f_Y(y) &= f_X(h^{-1}(y)) \times \left| \frac{d\left(h^{-1}(y)\right)}{dy} \right| \\ &= \lambda \exp\left\{-\lambda \exp\{y\}\right\} \times \exp\{y\} \\ &= \lambda \exp\left\{y - \lambda e^y\right\}, \end{aligned}$$

with $\lambda = 0.1$.

Bonus example: Solution

Recall that if $X \sim Exp(\lambda)$, then it has PDF

$$f_X(x) = \lambda e^{-\lambda x}, \quad x > 0, \quad \lambda > 0.$$

Now here, we have

$$h(x) = \log x, \quad \text{and} \quad h^{-1}(y) = e^y.$$

Therefore

$$\begin{aligned} f_Y(y) &= f_X(h^{-1}(y)) \times \left| \frac{d\left(h^{-1}(y)\right)}{dy} \right| \\ &= \lambda \exp\left\{-\lambda \exp\left\{y\right\}\right\} \times \exp\left\{y\right\} \\ &= \lambda \exp\left\{y - \lambda e^y\right\}, \end{aligned}$$

with $\lambda = 0.1$.

Bonus example: Solution

Recall that if $X \sim Exp(\lambda)$, then it has PDF

$$f_X(x) = \lambda e^{-\lambda x}, \quad x > 0, \quad \lambda > 0.$$

Now here, we have

$$h(x) = \log x, \quad \text{and} \quad h^{-1}(y) = e^y.$$

Therefore

$$\begin{aligned} f_Y(y) &= f_X(h^{-1}(y)) \times \left| \frac{d\left(h^{-1}(y)\right)}{dy} \right| \\ &= \lambda \exp\left\{-\lambda \exp\{y\}\right\} \times \exp\{y\} \\ &= \lambda \exp\left\{y - \lambda e^y\right\}, \end{aligned}$$

with $\lambda = 0.1$.

Bonus example: Solution

Recall that if $X \sim Exp(\lambda)$, then it has PDF

$$f_X(x) = \lambda e^{-\lambda x}, \quad x > 0, \quad \lambda > 0.$$

Now here, we have

$$h(x) = \log x, \quad \text{and} \quad h^{-1}(y) = e^y.$$

Therefore

$$\begin{aligned} f_Y(y) &= f_X(h^{-1}(y)) \times \left| \frac{d\left(h^{-1}(y)\right)}{dy} \right| \\ &= \lambda \exp\left\{-\lambda \exp\left\{y\right\}\right\} \times \exp\left\{y\right\} \\ &= \lambda \exp\left\{y - \lambda e^y\right\}, \end{aligned}$$

with $\lambda = 0.1$.

Bonus example: Solution

We can check this in R:

```
X = rexp(1000,  0.1)
Y = log(X)
hist(Y, prob=TRUE, main="Y=log(Wind Speeds)", xlab="
    Miles per hour")
lambda = 0.1
curve(lambda*exp(x-lambda*exp(x)), add=TRUE, col="red",
     lty=2, lwd=2)
```