Tutorial 4 Solutions

STAT 3013/8027

---

1. **Answer:** See the handwritten solutions for SI: 2.13, 2.14.
2. **Answer:** We assume that $U \sim \text{uniform}(0, 1)$.

   a.) Let's consider the first case: Let $Y = -log(U)$. To find the density of Y let's use the cdf method:

$$\begin{aligned}
P(Y \leq y) &= P(-log(U) \leq y) \\
&= P(-log(U) \leq y) = P(log(U) > -y) \\
&= P(U > exp(-y)) = 1 - P(U \leq exp(-y)) \\
&= 1 - exp(-y)
\end{aligned}$$

   So we have $F_Y(y) = 1 - exp(-y)$ and $f_Y(y) = exp(-y)$ which is the density for an exponential distribution with $\beta = 1$ for $0 \leq y \leq \infty$.

   - Now let's consider the next case: $Y = -log(1 - U)$. All we need to show is that $V = 1 - U$ is also a uniform $(0, 1)$ random variable and use the previous result. Let's directly use the 'pdf method' (note: V is monotone for $0 \leq v \leq 1$):

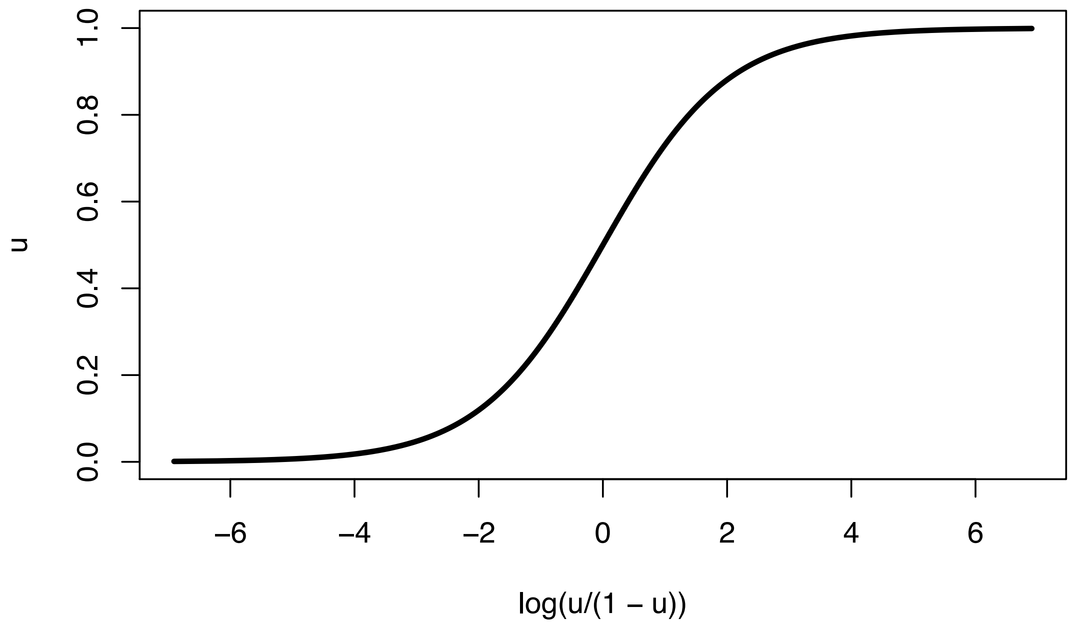
$$\begin{aligned}
V = 1 - U = g(u) \quad &\rightarrow \quad U = 1 - V = g^{-1}(v) \\
&\quad\;\; \frac{d}{dv} g^{-1}(v) = -1
\end{aligned}$$

$$\begin{aligned}
f_V(v) &= f_U\left(g^{-1}(v)\right) \left| \frac{d}{dv} g^{-1}(v) \right| \\
&= 1 \times \Bigg| -1 \Bigg| = 1 \text{ for } 0 \leq v \leq 1
\end{aligned}$$

   - We can see that $V \sim \text{uniform}(0, 1)$, which means $Y \sim \text{exponential}(\beta = 1)$.

   b.) Let $X = \log\left(\frac{U}{1-U}\right)$. Let's visually check that X is monotone on $0 \leq u \leq 1$.

```
u <- seq(0, 1, by=0.001)
plot( log(u/(1-u)), u, type="l", lwd=3)
```



$$x = \log\left(\frac{u}{1-u}\right) = g(u) \quad \rightarrow \quad u = \frac{exp(x)}{1+exp(x)} = \frac{1}{1+exp(-x)} = g^{-1}(x)$$

$$\frac{d}{dx} g^{-1}(x) = \frac{exp(-x)}{(1+exp(-x))^2}$$

$$\begin{aligned} f_X(x) &= f_U\left(g^{-1}(x)\right)\left|\frac{d}{dx} g^{-1}(x)\right| \\ &= 1 \times \left|\frac{exp(-x)}{(1+exp(-x))^2}\right| \\ &= = \frac{exp(-x)}{(1+exp(-x))^2} \text{ for } -\infty \le x \le \infty \end{aligned}$$

- We can see that X has the density of a logistic distribution with $\mu = 0$ and $\beta = 1$.

c.) Now let's generate from $Y \sim \text{logistic}(\mu = 3, \beta = 2)$.

- We know how to generate $X \sim \text{logistic}(\mu = 0, \beta = 1)$
- Now we want to generate Y which has a pdf:

$$\begin{aligned} f_Y(y) &= \frac{1}{\beta} \frac{exp\left(-\frac{(y-\mu)}{\beta}\right)}{\left[1 + exp\left(-\frac{(y-\mu)}{\beta}\right)\right]^2} \\ &= \frac{1}{\beta} f_X\left(\frac{(y-\mu)}{\beta}\right) \end{aligned}$$

This suggests that the right transformation would be:

$$Y = \beta X + \mu$$

$$Y = \beta X + \mu = g(x) \quad \rightarrow \quad X = \frac{(Y-\mu)}{\beta} = g^{-1}(y)$$
$$\frac{d}{dy} g^{-1}(y) = 1/\beta$$

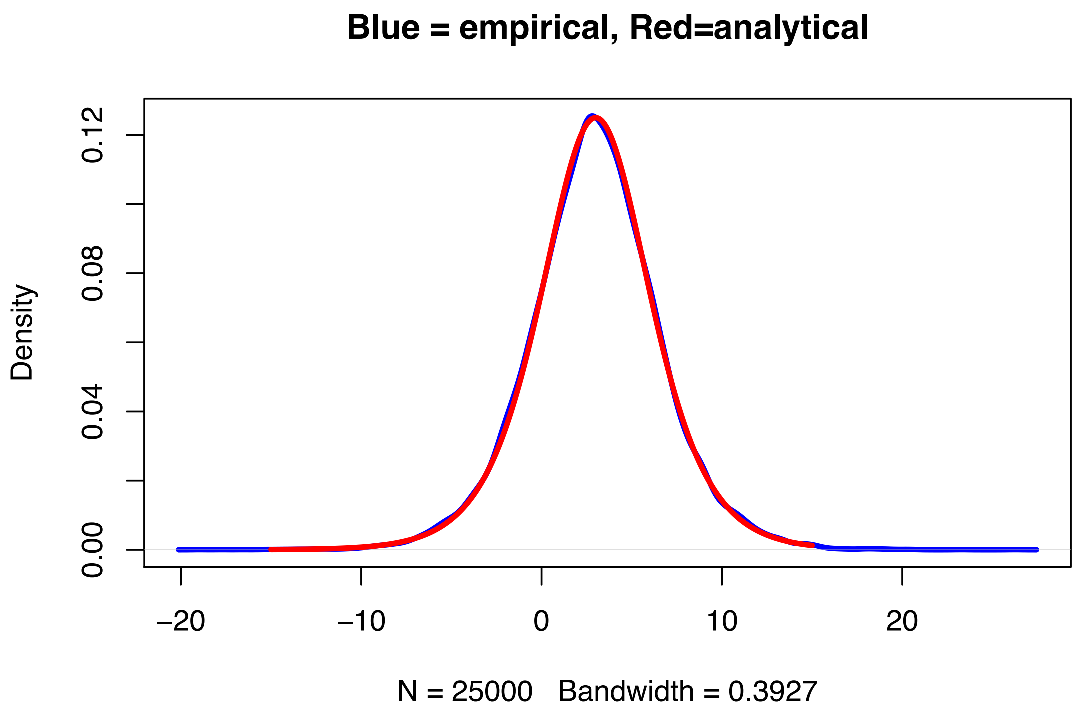
$$\begin{aligned} f_Y(y) &= f_X\left(g^{-1}(y)\right)\left|\frac{d}{dy} g^{-1}(y)\right| \\ &= \frac{exp\left(-\frac{(x-\mu)}{\beta}\right)}{\left[1 + exp\left(-\frac{(x-\mu)}{\beta}\right)\right]^2}\frac{1}{\beta} \end{aligned}$$

- Generate U from uniform(0,1).
- Generate Y from $\beta \log\left(\frac{U}{1-U}\right) + \mu$

```
set.seed(1001)
S <- 25000
mu <- 3
beta <- 2


## empirical
u <- runif(S, 0, 1)
y <- beta* log( u/(1-u) ) + mu
plot(density(y), type="l", col="blue", lwd=3,
     main="Blue = empirical, Red=analytical")

## analytical
y.an <- seq(-15, 15 , by=0.01)
f.y.an <- (1/beta)*( exp( - (y.an - mu)/beta )/ ( 1 + exp( - (y.an - mu)/beta ))^2 )
lines(y.an, f.y.an, col="red", lwd=3)
```


Blue = empirical, Red=analytical
Density
0.00
0.04
0.08
0.12
−20
−10
0
10
20
N = 25000   Bandwidth = 0.3927

GJJ Q 2.13) $x_1, \dots, x_n \overset{iid}{\sim} f_x(x)$

→ Find minimal sufficient statistics

a.) $X \sim \text{Unif}(\theta - \frac{1}{2}, \theta + \frac{1}{2})$

If $Z \sim \text{Unif}(a, b)$

$f_Z(z) = \frac{1}{b-a}$

$a \leq z \leq b$

$$f(x) = \frac{1}{(\theta + \frac{1}{2}) - (\theta - \frac{1}{2})} = \frac{1}{\frac{1}{2} + \frac{1}{2}}$$

$$= 1\, I_{(\theta - \frac{1}{2} \leq x \leq \theta + \frac{1}{2})}$$

$$L(\theta; \underset{\sim}{x}) = \prod_{i=1}^{n} 1\, I_{(\theta - \frac{1}{2} \leq x_i \leq \theta + \frac{1}{2})}$$

$$= 1^n\, I_{(\theta - \frac{1}{2} \leq x_1, \dots, x_n \leq \theta + \frac{1}{2})}$$

order statistics

$$= 1\, I_{(\underbrace{\theta - \frac{1}{2} \leq \overset{"x_{(1)}}{\min(x_i)}}_{\theta \leq x_{(1)} + \frac{1}{2}} \leq \underbrace{\overset{"x_{(n)}}{\max(x_i)} \leq \theta + \frac{1}{2}}_{\theta \geq x_{(n)} - \frac{1}{2}})}$$

$$= 1\, I_{(x_{(n)} - \frac{1}{2} \leq \theta \leq x_{(1)} + \frac{1}{2})}$$

- To solve for the minimal sufficient statistics we don't want the following ratio to depend on θ

$$R(\theta) = \frac{L(\theta; \underset{\sim}{x})}{L(\theta; \underset{\sim}{y})} = \frac{1\, I_{(x_{(n)} - \frac{1}{2} \leq \theta \leq x_{(1)} + \frac{1}{2})}}{1\, I_{(y_{(n)} - \frac{1}{2} \leq \theta \leq y_{(1)} + \frac{1}{2})}}$$

- Suppose $y_{(n)} - \frac{1}{2} < \theta < x_{(n)} - \frac{1}{2}$

$$\Rightarrow R(\theta) = \frac{0}{1} = 0$$

$\therefore$ Set $x_{(n)} = y_{(n)}$; $x_{(1)} = y_{(1)}$

$$R(\theta) = \frac{I(x_{(n)} - 1/2 \leq \theta \leq x_{(1)} - 1/2)}{I(x_{(n)} - 1/2 \leq \theta \leq x_{(1)} - 1/2)} = 1$$

b.) $x \sim \text{Unif}(-\theta, \theta)$; $f(x) = \frac{1}{\theta - (-\theta)} = \frac{1}{2\theta} I_{(-\theta \leq x \leq \theta)}$

$$L(\theta; \underset{\sim}{x}) = \prod_{i=1}^{n} \frac{1}{2\theta} I(-\theta \leq x_i \leq \theta)$$

$$= \left(\frac{1}{2\theta}\right)^n I_{(-\theta \leq x_1, x_2, \ldots, x_n \leq \theta)}$$

$$= \left(\frac{1}{2\theta}\right)^n I_{(\underbrace{-\theta \leq x_{(1)}}_{\theta \geq -x_{(1)}} \leq \underbrace{x_{(n)} \leq \theta}_{\theta \geq x_{(n)}})}$$

$\theta \geq \max(-x_{(1)}, x_{(n)})$

$\theta \geq \max(|x_{(1)}|, |x_{(n)}|)$

$\theta \geq \max(|x_i|)$

$$= \left(\frac{1}{2\theta}\right)^n I_{(\theta \geq \max(|x_i|))}$$

$$R(\theta) = \frac{(1/2\theta)^n I_{(\theta \geq \max(|x_i|))}}{(1/2\theta)^n I_{(\theta \geq \max(|y_i|))}}$$

Set $\max(|x_i|) = \max(|y_i|)$

$\Rightarrow$ R does not depend on θ.

$\therefore$ $\max(|x_i|)$ is a sufficient statistic

GJJ 2.14) $x_1, x_2, \ldots, x_n \overset{iid}{\sim} \text{beta}(a, b)$

$$f(x) = \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1} \; ; \; 0 < x < 1$$

the beta function

gamma function

$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

$$\therefore \; f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}$$

$$L(a,b) = \prod_{i=1}^{n} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x_i^{a-1}(1-x_i)^{b-1}$$

$$= \left(\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\right)^n \left[\prod_{i=1}^{n} x_i^{a-1}\right]\left[\prod_{i=1}^{n}(1-x_i)^{b-1}\right]$$

- For a minimal sufficient statistic we want the ratio to not depend on $\theta = (a, b)$:

$$R = \frac{\left[\prod x_i^{a-1}\right]\left[\prod (1-x_i)^{b-1}\right]}{\left[\prod y_i^{a-1}\right]\left[\prod (1-x_i)^{b-1}\right]}$$

We can see that is $\prod x_i = \prod y_i$ and $\prod(1-x_i) = \prod(1-y_i)$

then $R = 1$. $\therefore$ The minimal sufficient statistics are $\left(\prod_{i=1}^{n} x_i, \; \prod(1-x_i)\right)$.

Note: The order statistics are also minimal sufficient.

Let's check this: Sps $x = (0.2, 0.3) = (x_{(1)}, x_{(2)})$

$y = (0.1, 0.6) = (y_{(1)}, y_{(2)})$

$$\left.\begin{aligned} \prod x_i &= (0.2)(0.3) = 0.6 \\ \prod y_i &= (0.1)(0.6) = 0.6 \end{aligned}\right\}\ \text{these are the same}$$

$$\prod (1-x_i) = (0.8)(0.7) = 0.56$$

$$\prod (1-y_i) = (0.9)(0.4) = 0.36$$

$\therefore$ For R not to depend on θ $x_{(1)} = y_{(1)}$, $x_{(2)} = y_{(2)}$

$\therefore$ Minimal sufficient statistics are $(x_{(1)}, x_{(2)}, \ldots, x_{(n)})$.