Sta532 Hw1

Xu Chen 1/24/2020

Exercise 1

Based on the definition of CDF, we have

$$\mathsf{Pr}(Y = y) = \mathsf{Pr}(Y \leqslant y) - \mathsf{Pr}(Y < y) = F(y) - \sup_{y'<y} F(y')$$

Therefore,

(a)

$$\mathsf{Pr}(Y \in (a, b]) = F(b) - F(a)$$

(b)

$$\begin{aligned} \mathsf{Pr}(Y \in (a, b)) &= F(b) - F(a) - \left(F(b) - \sup_{y'<b} F(y')\right) \\ &= \sup_{y'<b} F(y') - F(a) \end{aligned}$$

(c)

$$\begin{aligned} \mathsf{Pr}(Y \in [a, b]) &= F(b) - F(a) + \left(F(a) - \sup_{y'<a} F(y')\right) \\ &= F(b) - \sup_{y'<a} F(y') \end{aligned}$$

Exercise 2

(a) As the transformation is monotonic, we have

$$p_W(w) = p_Y(e^{-w}) \left| \frac{\partial e^{-w}}{\partial w} \right| = e^{-w}, \quad w \in \mathbb{R}^+$$

This is an exponential distribution.

(b) This transformation is not monotonic and we derive the pdf of W using definition.

$$\mathsf{Pr}(W \leqslant w) = \mathsf{Pr}(1/Y \leqslant w) = \begin{cases} \mathsf{Pr}(1/w \leqslant Y < 0) & \text{if } w \leqslant 0 \\ \mathsf{Pr}(1/w \leqslant Y) + \mathsf{Pr}(Y \leqslant 0) & \text{otherwise} \end{cases}$$

The CDF of W is given by

$$F_W(w) = \begin{cases} F_Y(0) - F_Y(1/w) & \text{if } w \leqslant 0 \\ 1 - (F_Y(1/w) - F_Y(0)) & \text{otherwise} \end{cases}$$

Therefore, the pdf is given by

$$\begin{aligned}
p_W(w) &= -\frac{\partial}{\partial w} F_Y(1/w) = p_Y(1/w)/w^2 \\
&= \frac{1}{\pi(1+(1/w)^2)} \frac{1}{w^2} = \frac{1}{\pi(1+w^2)}, \; w \in \mathbb{R}
\end{aligned}$$

which is a Cauchy distribution.

(c) As the transformation is monotonic, we have

$$\begin{aligned}
p_W(w) &= p_Y(\log(w)) \left| \frac{\partial \log w}{\partial w} \right| \\
&= \frac{1}{w\sqrt{2\pi}} \exp\{-(\log w)^2/2\}, \quad w \in \mathbb{R}^+
\end{aligned}$$

This is a log normal distribution.

(d) As the transformation is not monotonic, we have

$$\mathsf{Pr}(W \leqslant w) = \mathsf{Pr}(-\sqrt{w} \leqslant Y \leqslant \sqrt{w}) = F_Y(\sqrt{w}) - F_Y(-\sqrt{w})$$

The pdf of W is given by

$$\begin{aligned}
p_W(w) &= \frac{\partial}{\partial w} F_Y(\sqrt{w}) - \frac{\partial}{\partial w} F_Y(-\sqrt{w}) = \frac{1}{2} p_Y(\sqrt{w}) w^{-1/2} + \frac{1}{2} p_Y(-\sqrt{w}) w^{-1/2} \\
&= \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{w}{\nu^2}\right)^{-\frac{\nu+1}{2}} \frac{1}{\sqrt{w}}, \quad w \in \mathbb{R}^+
\end{aligned}$$

This is a F distribution.

Exercise 3

(a) The CDF of U is

$$\begin{aligned}
F_U(u) &= \mathsf{Pr}(U \leqslant u) = \mathsf{Pr}(F_Y(Y) \leqslant u) = \mathsf{Pr}(Y \leqslant F_Y^{-1}(u)) \\
&= F_Y(F_Y^{-1}(u)) = u, \quad u \in [0,1]
\end{aligned}$$

which means that $U \sim \mathsf{Unif}(0,1)$.

(b) The CDF of X is

$$F_X(x) = \mathsf{Pr}(X \leqslant x) = \mathsf{Pr}(F_Y^{-1}(U) \leqslant x) = \mathsf{Pr}(U \leqslant F_Y(x)) = F_Y(x)$$

It means that X and Y follow the same distribution.

(c) The result in the previous part suggests a method of generating random variable which follows distribution F:

(1) Generate random variable $U \sim \mathsf{Unif}(0, 1)$

(2) Let $Y = F^{-1}(U)$

Then we have $Y \sim F$. This strategy is called inverse-CDF method. Therefore, to simulate a normal random variable in R, we could first use runif to generate U and then use qnorm (inverse CDF (quantile) function of normal distribution) to perform the transformation. Code is as follows.

```
set.seed(532)
n <- 1e5
u <- runif(n)
y <- qnorm(u)
```

We generate 10^5 samples using the inverse-CDF method. To verify our claim that these samples (y) follow a standard normal distribution, we plot the approximated density function of these samples (blue), compared against the exact density function of $\mathsf{N}(0, 1)$ (red) in Figure 1.

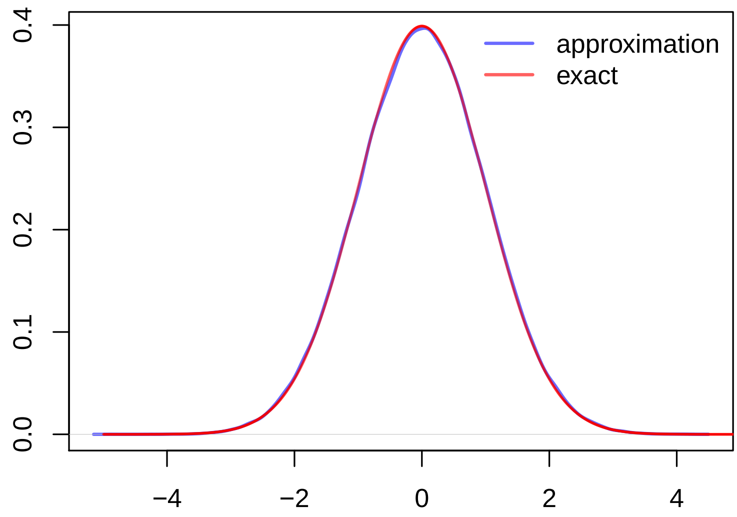

Figure 1: Approximated density and exact density.

Exercise 4

As $\mathsf{Pr}(X \in A \mid Y = y) = \int_A p_{X|Y}(x \mid y)\mathrm{d}x$, we have

$$\begin{aligned}\int \mathsf{Pr}(X \in A \mid Y = y)p_Y(y)\mathrm{d}y &= \int \int_A p_{X|Y}(x \mid y)p_Y(y)\mathrm{d}x\mathrm{d}y \\ &= \int \int_A p_{X,Y}(x, y)\mathrm{d}x\mathrm{d}y = \int_A \int p_{X,Y}(x, y)\mathrm{d}y\mathrm{d}x \\ &= \int_A p_X(x)\mathrm{d}x = \mathsf{Pr}(X \in A)\end{aligned}$$

This can be seen as a continuous version of the law of total probability. In discrete case, $\mathsf{Pr}(A) = \sum_{i=1}^n \mathsf{Pr}(A \cap B_i) = \sum_{i=1}^n \mathsf{Pr}(A \mid B_i)\mathsf{Pr}(B_i)$ where $\{B_i\}_{i=1}^n$ is a partition of the sample space. In the continuous case, we replace the summation and pmf by integral and pdf.

Exercise 5

$$\begin{aligned}
\lim_{\varepsilon\to 0} \mathsf{Pr}(X \in A \mid Y \in B_\varepsilon) &= \lim_{\varepsilon\to 0} \frac{\mathsf{Pr}(X \in A, Y \in B_\varepsilon)}{\mathsf{Pr}(Y \in B_\varepsilon)} \\
&= \lim_{\varepsilon\to 0} \frac{\mathsf{Pr}(X \in A, Y \in B_\varepsilon)}{\mathsf{Pr}(Y \in B_\varepsilon)} \\
&= \lim_{\varepsilon\to 0} \frac{\int_{y-\varepsilon}^{y} \int_A p_{X,Y}(x,u)\mathrm{d}x\mathrm{d}u}{\int_{y-\varepsilon}^{y} p_Y(u)\mathrm{d}u}
\end{aligned}$$

Both numerator and denominator are approaching 0 as ε goes to 0. Therefore, using L'Hopital rule and Leibniz rule, we have

$$\begin{aligned}
\lim_{\varepsilon\to 0} \mathsf{Pr}(X \in A \mid Y \in B_\varepsilon) &= \lim_{\varepsilon\to 0} \frac{\frac{\partial}{\partial\varepsilon}\int_{y-\varepsilon}^{y} \int_A p_{X,Y}(x,u)\mathrm{d}x\mathrm{d}u}{\frac{\partial}{\partial\varepsilon}\int_{y-\varepsilon}^{y} p_Y(u)\mathrm{d}u} \\
&= \lim_{\varepsilon\to 0} \frac{\int_A p_{X,Y}(x, y-\varepsilon)\mathrm{d}x}{p_Y(y-\varepsilon)} \\
&= \lim_{\varepsilon\to 0} \int_A p_{X|Y}(x \mid y-\varepsilon)\mathrm{d}x \\
&= \int_A \lim_{\varepsilon\to 0} p_{X|Y}(x \mid y-\varepsilon)\mathrm{d}x \qquad (1) \\
&= \int_A p_{X|Y}(x \mid y)\mathrm{d}x
\end{aligned}$$

We can exchange the integral and limit because the integration is with respect to x instead of y.

Exercise 6

As $Y \mid X \sim \mathsf{Ga}(c, X)$, we have

$$p_{X,Y}(x,y) = p_{Y|X}(y)p_X(x) = x^c y^{c-1} e^{-xy}/\Gamma(c) \times b^a x^{a-1} e^{-bx}/\Gamma(a)$$

Therefore, the marginal density of Y is

$$\begin{aligned}
p_Y(y) = \int p_{X,Y}(x,y)\mathrm{d}x &= y^{c-1}/\Gamma(c)b^a/\Gamma(a) \int x^{c+a-1} e^{-(b+y)x}\mathrm{d}x \\
&= y^{c-1}/\Gamma(c)b^a/\Gamma(a) \times \Gamma(c+a)/(b+y)^{c+a} \int \underbrace{(b+y)^{c+a} x^{c+a-1} e^{-(b+y)x}/\Gamma(c+a)}_{\text{pdf of } \mathsf{Ga}(a+c,\, b+y)} \mathrm{d}x \\
&= \frac{y^{c-1}b^a}{(b+y)^{a+c}} \frac{\Gamma(a+c)}{\Gamma(a)\Gamma(c)}
\end{aligned}$$

The conditional density of X given Y is

$$
\begin{aligned}
p_{X|Y}(x) &= p_{X,Y}(x,y)/p_Y(y) \\
&= (b+y)^{a+c} x^{a+c-1} e^{-(b+y)x}/\Gamma(a+c)
\end{aligned}
$$

Therefore, $X \mid Y = y \sim \mathsf{Ga}(a+c, b+y)$.