

1. For continuous case:

we know that

$$P(T=y) = F(y) - \lim_{y' \downarrow y} F(y') = F(y) - F(y) = 0$$

if  $a \geq b$ , then

$$\textcircled{1} \Pr(T \in (a, b)) = 0$$

$$\textcircled{2} \Pr(T \in (a, b)) = 0$$

$$\textcircled{3} \Pr(T \in [a, b]) = 0$$

else, we have

$$\textcircled{1} \Pr(T \in (a, b]) = \Pr(T \leq b) - \Pr(T \leq a) = F(b) - F(a)$$

$$\begin{aligned} \textcircled{2} \Pr(T \in (a, b)) &= \Pr(T \leq b) - \Pr(T \leq a) \\ &= \Pr(T < b) + \Pr(T = b) - \Pr(T \leq a) \\ &= F(b) - F(a) \end{aligned}$$

$$\begin{aligned} \textcircled{3} \Pr(T \in [a, b]) &= \Pr(T \leq b) - \Pr(T < a) \\ &= \Pr(T \leq b) - \Pr(T \leq a) - \Pr(T = a) \\ &= F(b) - F(a) \end{aligned}$$

For discrete case:

we know that

$$P(T=y) = F(y) - \sup_{y'=y} F(y')$$

if  $a > b$ , then

$$\textcircled{1} \quad P(T \in (a, b]) = 0$$

$$\textcircled{2} \quad P(T \in (a, b)) = 0$$

$$\textcircled{3} \quad \text{if } a > b, \quad P(T \in [a, b]) = 0$$

$$\text{if } a = b, \quad P(T \in [a, b]) = P(T = a) = F(a) - \sup_{y'=a} F(y')$$

else, we have

$$\textcircled{1} \quad P(T \in (a, b]) = P(T \leq b) - P(T \leq a) = F(b) - F(a)$$

$$\begin{aligned} \textcircled{2} \quad P(T \in (a, b)) &= P(T = b) - P(T \leq a) \\ &= P(T < b) + P(T = b) - P(T = b) - P(T \leq a) \\ &= F(b) - (F(b) - \sup_{y'=b} F(y')) - F(a) \\ &= \sup_{y'=b} F(y') - F(a) \end{aligned}$$

$$\begin{aligned} \textcircled{3} \quad P(T \in [a, b]) &= P(T \leq b) - P(T < a) \\ &= P(T \leq b) - P(T \leq a) - P(T = a) + P(T = a) \\ &= F(b) - F(a) + (F(a) - \sup_{y'=a} F(y')) \\ &= F(b) - \sup_{y'=a} F(y') \end{aligned}$$

2. (a)  $W = g(Y) = -\log Y$ ,  
 $Y \in (0, 1)$ ,  $Y = e^{-W}$ , then  $g$  is monotonic,  $W \in (-\infty, +\infty)$   
 $p_W(W) = \left| \frac{d e^{-W}}{d W} \right| p_Y(e^{-W}) = e^{-W}$

(b)  $W = g(Y) = 1/Y$   
 $Y \in \mathbb{R}$ ,  $Y = 1/W$ , then  $g$  is not monotonic at  $Y = 0$

if  $Y < 0$ ,  
 $p_W(W) = \left| \frac{d 1/W}{d W} \right| \cdot p_Y(1/W) = \frac{1}{2(W^2+1)}$ ,  $W \in (-\infty, +\infty)$

if  $Y > 0$ ,  
 $p_W(W) = \left| \frac{d 1/W}{d W} \right| \cdot p_Y(1/W) = \frac{1}{2(W^2+1)}$ ,  $W \in (-\infty, 0)$

(c)  $W = g(Y) = e^Y$   
 $Y \sim N(0, 1)$ ,  $Y = \log W$ , then  $g$  is monotonic,  $W \in (0, +\infty)$

$$p_W(W) = \left| \frac{d \log W}{d W} \right| p_Y(\log W) = \frac{1}{W} \frac{\exp(-\frac{1}{2}(\log W)^2)}{\sqrt{2\pi}}$$

$$(d) \quad W = g(Y) = Y^2$$

$Y \sim t_\nu$ ,  $g$  is not monotonic,  $W \in [0, +\infty)$

$$F_W(w) = \Pr(W \leq w) = \Pr(Y^2 \leq w) = \Pr(-\sqrt{w} \leq Y \leq \sqrt{w})$$

$$= F_Y(\sqrt{w}) - F_Y(-\sqrt{w})$$

$$p_W(w) = \frac{\partial F_W(w)}{\partial w} = \frac{\partial F_Y(\sqrt{w})}{\partial w} - \frac{\partial F_Y(-\sqrt{w})}{\partial w}$$

$$= p_Y(\sqrt{w}) \left| \frac{\partial \sqrt{w}}{\partial w} \right| - p_Y(-\sqrt{w}) \left| \frac{\partial (-\sqrt{w})}{\partial w} \right|$$

$$= \frac{(1 + \frac{\nu}{w})^{-\frac{\nu+1}{2}} \Gamma(\frac{\nu+1}{2})}{w^{-\frac{1}{2}} \sqrt{\nu\pi} \Gamma(\frac{\nu}{2})}$$

$$\begin{aligned}
 3. (a) \quad F_U(u) &= \Pr(U \leq u) = \Pr(F_Y(Y) \leq u) \\
 &= \Pr(Y \leq F_Y^{-1}(u)) = F_Y(F_Y^{-1}(u)) \\
 &= u
 \end{aligned}$$

$$\begin{aligned}
 (b) \quad F_X(x) &= \Pr(X \leq x) = \Pr(F_Y^{-1}(U) \leq x) \\
 &= \Pr(U \leq F_Y(x)) = F_U(F_Y(x)) \\
 &= F_Y(x)
 \end{aligned}$$

$$(c) \quad \text{qnorm}(\text{runif}(n))$$

$$P_U(u) = 1$$

From the derivation above we can see that a continuous strictly increasing CDF follows a uniform distribution.

We can first use `runif()` to generate values from uniform distribution, then use `qnorm()` to simulate a normal distribution.

$$\begin{aligned}
7. & \int \Pr(X \in A \mid Y=y) P_Y(y) dy \\
&= \int \frac{\Pr(X \in A, Y=y)}{P_Y(y)} P_Y(y) dy \\
&= \int \Pr(X \in A, Y=y) dy \\
&= \iint_A P_{X,Y}(x,y) dx dy \\
&= \int_A P_X(x) dx \\
&= \Pr(X \in A)
\end{aligned}$$

for all possible values of  $y \in Y$

We sum up all the probabilities,  $\wedge$  of both  $Y=y$  and  $x \in A$  under the circumstance of  $Y=y$  happening together, we can get the probability of  $X \in A$  happening. This is an analogy to the discrete case.

For the continuous case, we can divide the value domain of  $Y$  to very small divisions, then sum up all the divisions and get the probability of  $X \in A$ .

$$\begin{aligned}
& \int_A \lim_{\varepsilon \rightarrow 0} \Pr(X \in A \mid Y \in B_\varepsilon) \\
&= \lim_{\varepsilon \rightarrow 0} \frac{\Pr(X \in A, Y \in B_\varepsilon)}{\Pr(Y \in B_\varepsilon)} \\
&= \lim_{\varepsilon \rightarrow 0} \frac{\int_A \int_{y-\varepsilon}^y p_{X,Y}(x, t) dx dt}{\int_{y-\varepsilon}^y p_Y(t) dt}
\end{aligned}$$

$$\begin{aligned}
& \lim_{\varepsilon \rightarrow 0} \int_A \int_{y-\varepsilon}^y p_{X,Y}(x, t) dx dt \\
&= \varepsilon \cdot \int_A \lim_{\varepsilon \rightarrow 0} \frac{\int_{y-\varepsilon}^y p_{X,Y}(x, t) dt}{\varepsilon} dx \\
&= \varepsilon \cdot \int_A \lim_{\varepsilon \rightarrow 0} \frac{F_{X,Y}(x, y) - F_{X,Y}(x, y-\varepsilon)}{\varepsilon} dx \\
&= \varepsilon \cdot \int_A p_{X,Y}(x, y) dx
\end{aligned}$$

$$\begin{aligned}
& \lim_{\varepsilon \rightarrow 0} \int_{y-\varepsilon}^y p_Y(t) dt \\
&= \varepsilon \cdot \lim_{\varepsilon \rightarrow 0} \frac{\int_{y-\varepsilon}^y p_Y(t) dt}{\varepsilon} \\
&= \varepsilon \cdot \lim_{\varepsilon \rightarrow 0} \frac{F_Y(y) - F_Y(y-\varepsilon)}{\varepsilon} \\
&= \varepsilon \cdot p_Y(y)
\end{aligned}$$

$$\begin{aligned}
\Rightarrow \lim_{\varepsilon \rightarrow 0} \Pr(X \in A \mid Y \in B_\varepsilon) &= \frac{\varepsilon \int_A p_{X,Y}(x, y) dx}{\varepsilon p_Y(y)} = \int_A \frac{p_{X,Y}(x, y)}{p_Y(y)} dx \\
&= \int_A p_{X|Y}(x, y) dx
\end{aligned}$$

$$6. Y|X \sim \text{Gamma}(c, x)$$

$$P_{Y|X}(y|x) = x^c y^{c-1} e^{-xy} / \Gamma(c)$$

$$X \sim \text{Gamma}(a, b)$$

$$P_X(x) = b^a x^{a-1} e^{-bx} / \Gamma(a)$$

$$\begin{aligned} \Rightarrow P_{X,Y}(x,y) &= P_{Y|X}(x,y) \times P_X(x) \\ &= x^{a+c-1} e^{-xy-bx} y^{c-1} \frac{b^a}{\Gamma(a)\Gamma(c)} \end{aligned}$$

$$\Rightarrow P_Y(y) = \int_0^{+\infty} P_{X,Y}(x,y) dx$$

$$= \int_0^{+\infty} x^{a+c-1} e^{-xy-bx} y^{c-1} \frac{b^a}{\Gamma(a)\Gamma(c)} dx$$

$$= y^{c-1} \frac{b^a}{\Gamma(a)\Gamma(c)} \int_0^{+\infty} \underbrace{x^{a+c-1} e^{-xy-bx}}_{\text{kernel of Gamma}(a+c, y+b)} dx$$

$$\text{we have } \int_0^{+\infty} x^{a+c-1} e^{-xy-bx} (y+b)^{a+c} / \Gamma(a+c) dx = 1$$

$$\Rightarrow \int_0^{+\infty} x^{a+c-1} e^{-xy-bx} dx = \frac{\Gamma(a+c)}{(y+b)^{a+c}}$$

$$P_Y(y) = y^{c-1} \frac{b^a}{\Gamma(a)\Gamma(c)} \frac{\Gamma(a+c)}{(y+b)^{a+c}}$$

$$\Rightarrow P_{X|Y}(x|y) = \frac{P_{X,Y}(x,y)}{P_Y(y)}$$

$$= \frac{x^{a+c-1} e^{-xy-bx} \cancel{y^{c-1} b^a}}{\Gamma(a)\Gamma(c)} \times \frac{\Gamma(a)\Gamma(c)(y+b)^{a+c}}{\cancel{y^{c-1} b^a} \Gamma(a+c)}$$

$$= \frac{x^{a+c-1} e^{-xy-bx} (y+b)^{a+c}}{\Gamma(a+c)}$$