IE 605: Engineering Statistics

Solutions of tutorial 7

Solution 1

Given: $U \sim Uniform(0,1)$

- To show: Both $-logU$ and $-log(1-U)$ are exponential random variables.

  Solution:

  Since $U \sim Uniform(0,1)$, the pdf is given as

  $$f_U(u) = \begin{cases} 1, & 0 < u < 1 \\ 0, & \text{otherwise} \end{cases} \quad \text{and} \quad F_U(u) = u, \qquad 0 < u < 1$$

  We have Y = g(U) = -log(U)

  $$\text{Since } \frac{d}{dU}g(U) = \frac{d}{dU}(-log(U)) = \frac{-1}{U} < 0, \qquad \text{for } 0 < u < 1$$
  $$\implies g(U) \text{ is a decreasing function.}$$

  As U ranges between 0 and 1, $-log(U)$ ranges between 0 and $\infty \implies Y \in (0,\infty)$.

  Now for $y > 0$,$y = -log(u) \implies u = e^{-y}$, so $g^{-1}(y) = e^{-y}$.

  Therefore, for $y > 0$,

  $$\begin{aligned} F_Y(y) &= \mathbb{P}\{Y \leq y\} = \mathbb{P}\left\{X \geq g^{-1}(y)\right\} \\ &= 1 - F_X(g^{-1}(y)) \\ &= 1 - F_X(e^{-y}) = 1 - e^{-y}. \end{aligned}$$
  $$\text{and } F_Y(y) = 0, \qquad \text{for } y \leq 0.$$
  $$\implies Y \sim Exponential(1)$$

  Now we have Y = g(U) = -log(1-U)

  $$\text{Since } \frac{d}{dU}g(U) = \frac{d}{dU}(-log(1-U)) = \frac{1}{1-U} > 0, \qquad \text{for } 0 < u < 1$$
  $$\implies g(U) \text{ is a increasing function.}$$

As U ranges between 0 and 1, $-log(1-U)$ ranges between 0 and $\infty \implies Y \in (0, \infty)$.

Now for $y > 0$,$y = -log(1-u) \implies u = 1 - e^{-y}$, so $g^{-1}(y) = 1 - e^{-y}$.

Therefore, for $y > 0$,

$$\begin{aligned} F_Y(y) &= \mathbb{P}\{Y \leq y\} = \mathbb{P}\left\{X \leq g^{-1}(y)\right\} \\ &= F_X(g^{-1}(y)) \\ &= F_X(1-e^{-y}) = 1 - e^{-y}. \end{aligned}$$

and $F_Y(y) = 0, \qquad$ for $y \leq 0$.

$\implies Y \sim Exponential(1)$

- To show: $X = log\left(\frac{U}{1-U}\right)$ is a logistic(0,1) random variable.

  Solution:

$$X = g(U) = -log\frac{1-U}{U}.$$

Then $g^{-1}(x) = \dfrac{1}{1+e^{-y}}$.

Thus $f_X(x) = 1 \times \mid \dfrac{e^{-y}}{(1+e^{-y})^2} \mid$

$$= \frac{e^{-y}}{(1+e^{-y})^2} \qquad -\infty < y < \infty,$$

$\implies X \sim logistic(0,1)$ random variable.

Solution 2

Given: The Box-Muller method for generating normal pseudo-random variables is based on the transformation

$$X_1 = cos(2\pi U_1)\sqrt{-2log(U_2)}, \qquad X_2 = sin(2\pi U_1)\sqrt{-2log(U_2)}$$

where U_1 and U_2 are iid $Uniform(0,1)$.

To prove: X_1 and X_2 are independent $Normal(0,1)$ random variables.

Solution:

Since U_1 and U_2 are iid $Uniform(0,1)$, the joint pdf is given as

$$f_{U_1,U_2}(u_1,u_2) = 1, \qquad 0 \leq u_1, u_2 \leq 1.$$

We are given

$$X_1 = cos(2\pi U_1)\sqrt{-2log(U_2)}, \qquad \text{and } X_2 = sin(2\pi U_1)\sqrt{-2log(U_2)}$$
$$\implies U_1 = \frac{1}{2\pi}tan^{-1}\left(\frac{X_2}{X_1}\right), \qquad \text{and } U_2 = e^{-\frac{1}{2}(X_1^2+X_2^2)}$$
$$\text{Jacobian:}J = \begin{vmatrix} \frac{\delta U_1}{\delta X_1} & \frac{\delta U_1}{\delta X_2} \\ \frac{\delta U_2}{\delta X_1} & \frac{\delta U_2}{\delta X_2} \end{vmatrix}$$
$$= \begin{vmatrix} \frac{-X_2}{2\pi(X_1^2+X_2^2)} & \frac{X_1}{2\pi(X_1^2+X_2^2)} \\ (-X_1)e^{-\frac{1}{2}(X_1^2+X_2^2)} & (-X_2)e^{-\frac{1}{2}(X_1^2+X_2^2)} \end{vmatrix} = \frac{1}{2\pi}e^{-\frac{1}{2}(X_1^2+X_2^2)}$$

The joint pdf of X_1 and X_2 is

$$\begin{aligned} f_{X_1,X_2}(x_1,x_2) &= \frac{1}{2\pi}e^{-\frac{1}{2}(X_1^2+X_2^2)}, \qquad -\infty \le x_1, x_2 \le \infty \\ &= \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(X_1^2)}.\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(X_2^2)}, \qquad -\infty \le x_1, x_2 \le \infty \end{aligned}$$
$$\implies X_1 \text{ and } X_2 \text{ are independent } Normal(0,1) \text{ random variables}$$

Solution 3

Given: Park et.al. (1996) describe a method for generating correlated binary variables based on the following scheme:

Let X_1, X_2, X_3 be independent Poisson random variables with mean $\lambda_1, \lambda_2, \lambda_3$ respectively, and create the random variables

$$Y_1 = X_1 + X_3 \qquad \text{and} \qquad Y_2 = X_2 + X_3.$$

1. To show: $Cov(Y_1, Y_2) = \lambda_3$.

   Solution:

$$\begin{aligned} Cov(Y_1, Y_2) &= Cov(X_1 + X_3, X_2 + X_3) \\ &= Cov(X_3, X_3) = \lambda_3 \qquad \text{since } X_1, X_2 \text{ and } X_3 \text{ are independent.} \end{aligned}$$

2. Define $Z_i = \mathbb{I}(Y_i = 0)$ and $p_i = e^{-(\lambda_i+\lambda_3)}$.

   To show: Z_i are $Bernoulli(p_i)$ with

$$Corr(Z_1, Z_2) = \frac{p_1p_2(e^{\lambda_3} - 1)}{\sqrt{p_1(1-p_1)}\sqrt{p_2(1-p_2)}}.$$

   Solution:

$$Z_i = \begin{cases} 1, & \text{if} X_i = X_3 = 0 \\ 0, & \text{otherwise} \end{cases}$$

$$p_i = P(Z_i = 0) = P(Y_i = 0) = P(X_i = 0, X_3 = 0) = e^{-(\lambda_i + \lambda_3)}.$$

Therefore Z_i are $Bernoulli(p_i)$ with $\mathbb{E}[Z_i] = p_i, Var(Z_i) = p_i(1 - p_i)$ and

$$\begin{aligned}
\mathbb{E}[Z_1 Z_2] &= P(Z_1 = 1, Z_2 = 1) \\
&= P(Y_1 = 0, Y_2 = 0) \\
&= P(X_1 + X_3 = 0, X_2 + X_3 = 0) \\
&= P(X_1 = 0).P(X_2 = 0).P(X_3 = 0) \\
&= e^{-\lambda_1} e^{-\lambda_2} e^{-\lambda_3}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
Cov(Z_1, Z_2) &= \mathbb{E}[Z_1 Z_2] - \mathbb{E}[Z_1]\,\mathbb{E}[Z_2] \\
&= e^{-\lambda_1} e^{-\lambda_2} e^{-\lambda_3} - e^{-(\lambda_1+\lambda_3)} e^{-(\lambda_2+\lambda_3)} \\
&= e^{-(\lambda_1+\lambda_3)} e^{-(\lambda_2+\lambda_3)} (e^3 - 1) \\
&= p_1 p_2 (e^{\lambda_3} - 1).
\end{aligned}$$

Thus $Corr(Z_1, Z_2) = \dfrac{p_1 p_2 (e^{\lambda_3} - 1)}{\sqrt{p_1(1 - p1)}.\sqrt{p_2(1 - p_2)}}.$

Solution 4

The given algorithm is the Acceptance-Rejection method (indirect method) of random variable generation for some desired choice of distribution f.

$$\begin{aligned}
P(Y \leq y) &= P(V \leq y \mid U < \frac{1}{c} f_Y(V)) \\
&= \frac{P(V \leq y, U < \frac{1}{c} f_Y(V))}{P(U < \frac{1}{c} f_Y(V))} \\
&= \frac{\int_0^y \int_0^{\frac{1}{c} f_Y(v)} du\, dv}{\frac{1}{c}} \\
&= \frac{\frac{1}{c} \int_0^y f_Y(v) dv}{\frac{1}{c}} \\
&= \int_0^y f_Y(v) dv
\end{aligned}$$

Thus, we may say that $Y \sim f$, and in this case, $f \equiv Beta(m, n)$
You may refer to Sec. 8.2.4 from the book "*Simulation Modeling and Analysis*" by Averill M. Law for a better understanding, in addition to the references already suggested in class.

Solution 5

1. $M = \sup_y \frac{\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} y^{a-1}(1-y)^{b-1}}{\frac{\Gamma([a]+b)}{\Gamma([a])\Gamma([b])} y^{[a]-1}(1-y)^{[b]-1}} < \infty$, since $a - [a] > 0$ and $b - [b] > 0$ and $y \in (0, 1)$

2. $M = \sup_y \frac{\frac{a^b}{\Gamma(b)} e^{-ay} y^{b-1}}{\frac{[a]^b}{\Gamma(b)} e^{-[a]y} y^{b-1}} < \infty$, since $a - [a] > 0$ and $y > 0$.

3. $M = \sup_y \frac{\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} y^{a-1}(1-y)^{b-1}}{\frac{\Gamma([a]+1+\beta)}{\Gamma([a]+1)\Gamma(b')} y^{[a]+1-1}(1-y)^{b'-1}} < \infty$, since $a - [a] - 1 < 0$ and $y \in (0, 1)$. $b - b' > 0$ when $b' = [b]$ and will be equal to zero when $b' = b$, thus it does not affect the result.