Review of Last Lecture

Key concepts and/or techniques:

- bivariate normal distribution and its properties

1. marginal distributions are normal
2. conditional distributions are normal
3. independence $\Longleftrightarrow$ Uncorrelation

- find the distribution of the function of RVs, i.e., determine the pmf or pdf of the functions of RVs

Review of Last Lecture

Definition

Let X and Y be 2 continuous RVs and have the joint pdf

$$f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}\exp[-\frac{1}{2}q(x,y)], x \in \mathbb{R}, y \in \mathbb{R},$$

$$q(x,y) = \frac{1}{1-\rho^2}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right] \geq 0$$

where $\mu_X, \mu_Y \in \mathbb{R}$, $\sigma_X, \sigma_Y > 0$ and $|\rho| < 1$. Then X and Y are said to be bivariate normally distributed.

Key components: Scaled exponential function with a quadratic and negative function as its exponent.

Review of Last Lecture

1. Marginal distributions are normal:

$$X \sim N(\mu_X, \sigma_X^2), Y \sim N(\mu_Y, \sigma_Y^2)$$

2. Conditional distributions are normal:

$$X|Y = y \sim N\left(\mu_X + \frac{\sigma_X}{\sigma_Y}\rho(y - \mu_Y), (1-\rho^2)\sigma_X^2\right)$$

$$Y|X = x \sim N\left(\mu_Y + \frac{\sigma_Y}{\sigma_X}\rho(x - \mu_X), (1-\rho^2)\sigma_Y^2\right)$$

3. Independence $\Longleftrightarrow$ Uncorrelation

Review of Last Lecture

[Function of One Random Variable]

Let X be a RV of either discrete or continuous type with its pmf or pdf denoted by $f(x)$. Consider a function of X, say $Y = u(X)$. Then Y is also a RV and has its pmf or pdf.

How to compute the pmf or pdf of Y?

In what follows, we consider the case where $Y = u(X)$ is a one-to-one mapping.

Review of Last Lecture

1. For discrete RV, when $Y = u(X)$ be a one-to-one mapping with inverse $X = v(Y)$. Then the pmf of Y is

$$g(y) = f[v(y)] \quad \text{for} \quad y \in \overline{S_Y}$$

2. For continuous RV, when $Y = u(X)$ is continuous, strictly decreasing or increasing and has inverse function $X = v(Y)$, whose derivative $\frac{dv(y)}{dy}$ exists, the pdf of Y, denoted by $g(y)$,

$$g(y) = f(v(y))\left|\frac{dv(y)}{dy}\right|$$

STA2001 Probability and Statistics (I)

Lecture 19

Tianshi Chen

The Chinese University of Hong Kong, Shenzhen

Example 2, page 174

Let X have the pdf

$$f(x) = 3(1-x)^2, \quad 0 < x < 1.$$

Consider $Y = (1-X)^3$ and calculate the pdf of Y, $g(y)$.

Example 2, page 174

$$Y = u(X) = (1-X)^3 \quad \longrightarrow \quad \text{continuous, strictly decreasing}$$

$$\text{Inverse function} \quad \longrightarrow \quad X = v(Y) = 1 - Y^{\frac{1}{3}}$$

1. The sample space of Y is $\overline{S_Y} = (0, 1)$, since $0 < x < 1$.
2. $$g(y) = f(v(y)) \left| \frac{dv(y)}{dy} \right| \quad \text{where} \quad \frac{dv(y)}{dy} = -\frac{1}{3} y^{-\frac{2}{3}}$$
   $$= 3(1-(1-y^{\frac{1}{3}}))^2 \left| -\frac{1}{3} y^{-\frac{2}{3}} \right| = 1,$$
   $$0 < y < 1 \quad Y \sim U(0,1)$$

Theorem 5.1-1, page 175

Given a random distribution, it is possible to construct a RV such that this RV has the given random distribution.

Theorem[Random Number Generator]

Let $Y \sim U(0,1)$ and $F(x)$ have the properties of a cdf of a continuous RV with $F(\mathrm{a}) = 0, F(b) = 1$. Moreover, $F(x)$ is strictly increasing such that $F(x) : (a,b) \to [0,1]$, where a could be $-\infty$, b could be ∞. Then $X = F^{-1}(Y)$ is continuous RV with cdf $F(x)$.

Theorem 5.1-1, page 175

Proof: Idea — we need to show $P(X \leq x) = F(x)$

$$P(X \leq x) = P(F^{-1}(Y) \leq x) = P(Y \leq F(x))$$

$$\text{since} \quad \{y | F^{-1}(y) \leq x\} = \{y | y \leq F(x)\}.$$

Note that

$$Y \sim U(0,1) \Longrightarrow P(Y \leq y) \overset{0<y<1}{=\!=\!=\!=} \int_0^y 1 dz = y$$

Therefore,

$$P(X \leq x) = P(Y \leq F(x)) = F(x)$$

Remarks

Theorem 5.1-1 can be used to construct a random number generator for distributions with strictly increasing cdf based on the random generator for a uniform distribution.

Random number generator

1. generator a random number y from $U(0, 1)$

2. Take $x = F^{-1}(y)$

Then x is a random number generated from the distribution or RV with cdf $F(x)$.

Example 3

Question

Assume that we know how to generate a random number from $Y \sim U(0, 1)$.

- Can we generate a random number from the exponential distribution with parameter θ, whose pdf is given by

$$f(x) = \frac{1}{\theta} e^{-\frac{x}{\theta}}, \quad x \geq 0$$

Example 3

Question

Assume that we know how to generate a random number from $Y \sim U(0,1)$.

- Can we generate a random number from the exponential distribution with parameter θ, whose pdf is given by
$$f(x) = \frac{1}{\theta} e^{-\frac{x}{\theta}}, \quad x \geq 0$$
- Can we run computer simulation to collect the waiting time until the first customer that arrives at a cinema or a hospital?

Example 3

The cdf of X is

$$F(x) = P(X \leq x) = \int_0^x \frac{1}{\theta} e^{-\frac{t}{\theta}} dt = 1 - e^{-\frac{x}{\theta}}, \quad x \geq 0,$$

which is strictly increasing.

1. If we know how to generate a random number from $Y \sim U(0,1)$, say the random number generated is y.
2. $x = F^{-1}(y)$ is the random number generated from the exponential distribution with θ, where

$$x = F^{-1}(y) = -\theta \ln(1-y), \quad y \in (0,1)$$

Histogram for continuous distribution

The simplest form of a histogram is constructed as follows

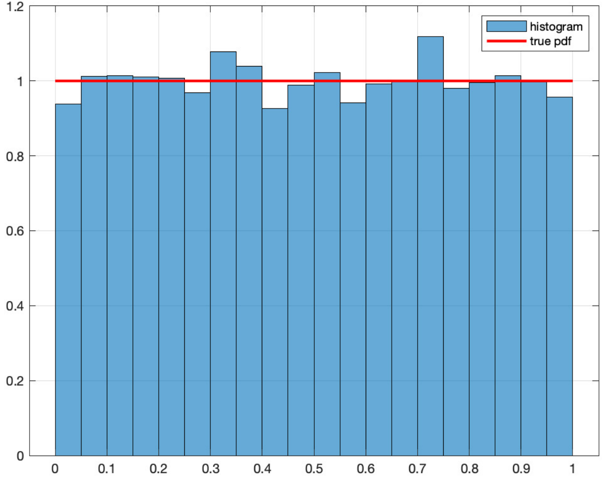
1. divide (or "bin") the sample space of the distribution into a sequence of adjacent, non-overlapping and equally spaced subintervals.
2. treat each subinterval as an event, then count how many observed numerical outcomes fall into each subinterval and calculate the relative frequency
3. draw a rectangle erected over the bin with height equal to the relative frequency divided by the width of each subinterval.

Remark:

- Note that the area of the histogram is equal to 1, thus histogram gives an approximation of the probability density function of the underlying random variable.

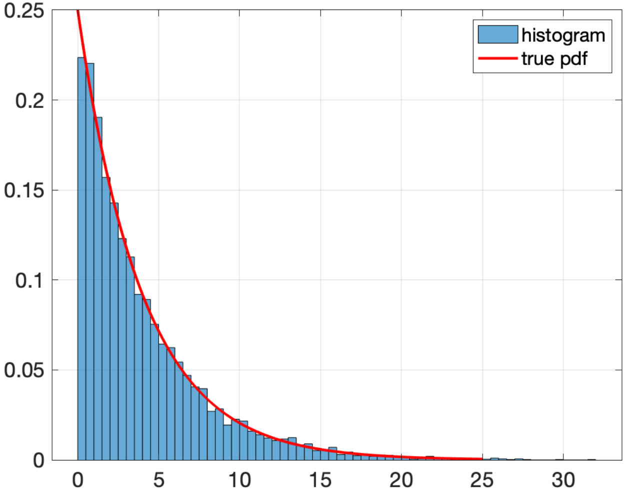
Example 3

The histogram of 10000 random numbers y generated from $U(0, 1)$



Example 3

The histogram of 10000 random numbers x for exponential distribution with $\theta = 4$



Example 3

Matlab script for the random number generator:

```
figure(1);
y=rand(10000,1); % generate 10000 random number y from U(0,1)
histogram(y,'normalization','pdf') % draw the histogram of
10000 y
hold on;
plot(0:0.01:1,ones(1,length(0:0.01:1)),'r',
'linewidth',2) % plot the true pdf of U(0,1)
grid on;
legend('histogram', 'true pdf')
figure(2);
theta=4;
x=-theta*log(1-y); % generate 10000 random number x
histogram(x,'normalization','pdf') % draw the histogram of
10000 x
hold on;
plot(0:0.01:25,exp(-(0:0.01:25)/theta)/theta,'r',
'linewidth',2) % plot the true pdf of exponential distribution with
theta=4
grid on;
```

Theorem 5.1-2, page 176

[Theorem]

Suppose that X is a continuous RV with $\overline{S_X} = (a, b)$, and moreover, its cdf $F(x)$ is strictly increasing. Then the RV Y, defined by $Y = F(X)$, has a uniform distribution, that is, $Y \sim U(0, 1)$.

Theorem 5.1-2, page 176

Proof:

Since $F(\mathrm{a}) = 0$, $F(b) = 1$ and $F(x)$ is strictly increasing,

$Y = F(X)$ with the range $\overline{S_Y} = (0, 1)$.

Consider the cdf of Y:

$$P(Y \leq y) = P(F(X) \leq y), \quad y \in (0, 1)$$

Theorem 5.1-2, page 176

Since $F(x)$ is strictly increasing, $\{F(X) \leq y\} \iff \{X \leq F^{-1}(y)\}$.

$$P(Y \leq y) = P(F(X) \leq y) = P(X \leq F^{-1}(y)), \quad 0 < y < 1.$$

Since $P(X \leq x) = F(x)$, we have

$$P(Y \leq y) = F(F^{-1}(y)) = y, \quad 0 < y < 1 \longrightarrow \quad \text{cdf of } U(0,1)$$

Example 2, page 174, continued

Let X have the pdf

$$f(x) = 3(1-x)^2, \quad 0 < x < 1$$

Consider $Y = (1-X)^3$ and then $Y \sim U(0,1)$.

The result can be obtained from Theorem 5.1-2. Since

$$F(x) = 1 - (1-x)^3$$

is strictly increasing, then

$$F(X) = 1 - (1-X)^3 \sim U(0,1)$$

which implies $(1-X)^3 = [1 - F(X)] \sim U(0,1)$.

The case: $Y = u(X)$ not one-to-one

[Function of One Random Variable]

Let X be a RV of either discrete or continuous type with its pmf or pdf denoted by $f(x)$. Consider a function of X, say $Y = u(X)$. Then Y is also a RV and has its pmf or pdf.

How to compute the pmf or pdf of Y?

When $Y = u(X)$ is not one-to-one, there is no general result.

Example 4

Question

Assume that X is a continuous RV with pdf

$$f(x) = \frac{1}{\pi(1+x^2)} \quad x \in (-\infty, \infty)$$

Let $Y = X^2$. Find the pdf of Y.

Clearly, $\overline{S_Y} = [0, \infty)$

Let the cdf of Y be $G(y)$. Then

$$G(y) = P(Y \leq y), \quad y \in [0, \infty)$$

$$= P(X^2 \leq y) = P(-\sqrt{y} \leq X \leq \sqrt{y}) = \int_{-\sqrt{y}}^{\sqrt{y}} f(x)dx$$

Example 4

$$
\begin{aligned}
G(y) &= P(Y \leq y), \quad y \in [0, \infty) \\
&= P(X^2 \leq y) = P(-\sqrt{y} \leq X \leq \sqrt{y}) = \int_{-\sqrt{y}}^{\sqrt{y}} f(x)dx \\
&= \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{\pi(1+x^2)} dx = 2\int_0^{\sqrt{y}} \frac{1}{\pi(1+x^2)} dx
\end{aligned}
$$

$$\Rightarrow g(y) = G'(y) = \tfrac{2}{\pi(1+y)} \times \tfrac{1}{2} \times \tfrac{1}{\sqrt{y}} = \tfrac{1}{\pi(1+y)\sqrt{y}}$$