STA2001 Probability and Statistics (I)

Lecture 13

Tianshi Chen

The Chinese University of Hong Kong, Shenzhen

Review of Last Lecture

Let X and Y be two continuous random variables and (X, Y) be a pair of RVs with their range denoted by $\overline{S} \subseteq R^2$. Then (X, Y) or X and Y is said to be a bivariate continuous RV.

> **[Road map]**
>
> To study the bivariate continuous random variable
>
> | discrete RV $\longrightarrow$ continuous RV | Mathematical expectations |
> |---|---|
> | pmf $\longrightarrow$ pdf | mean |
> | joint pmf $\longrightarrow$ joint pdf | variance |
> | marginal pmf $\longrightarrow$ marginal pdf | covariance |
> | conditional pmf $\longrightarrow$ conditional pdf | correlation coefficient |

Review of Last Lecture

> **[Marginal pdf]**
>
> The marginal pdf of X and Y are:
>
> $$f_X(x) = \int_{\overline{S_Y}(x)} f(x,y) \, dy : \overline{S_X} \to (0, \infty)$$
>
> $$\overline{S_Y}(x) = \{y | (x,y) \in \overline{S}\} \text{ for } x \in \overline{S_X}$$
>
> $$f_Y(y) = \int_{\overline{S_X}(y)} f(x,y) \, dx : \overline{S_Y} \to (0, \infty)$$
>
> $$\overline{S_X}(y) = \{x | (x,y) \in \overline{S}\} \text{ for } y \in \overline{S_Y}$$

Review of Last Lecture

> **[Covariance and Correlation Coefficient]**
>
> Let $g(X, Y)$ be a function of X and Y, whose joint pdf $f(x, y) : \overline{S} \to [0, \infty)$. Then
>
> $$E[g(X, Y)] = \iint_{\overline{S}} g(x, y) f(x, y) dxdy$$
>
> 1. covariance $Cov(X, Y)$
> 2. correlation coefficient $\rho(X, Y)$

> **[Independent Variables]**
>
> Two continuous RVs X and Y are independent if
>
> $$f(x, y) = f_X(x) f_Y(y), \qquad x \in \overline{S_X}, \quad y \in \overline{S_Y}$$

Review of Last Lecture

[Conditional pdf and Conditional Expectation]

The conditional pdf, mean, and variance of Y, given that $X = x$ are

$$h(y|x) = \frac{f(x,y)}{f_X(x)} \quad \text{for} \quad f_X(x) > 0, y \in \overline{S_Y}(x)$$

$$P(Y \in A|X = x) = \int_{y \in A} h(y|x)dy, A \subseteq \overline{S_Y}(x)$$

$$E(Y|X = x) = \int_{\overline{S_Y}(x)} yh(y|x)dy$$

$$Var(Y|X = x) = E\{[Y - E(Y|X = x)]^2|X = x\}$$

$$= \int_{\overline{S_Y}(x)} [y - E(Y|X = x)]^2 h(y|x)dy$$

$$= E[Y^2|X = x] - [E(Y|X = x)]^2$$

Section 4.5 Bivariate Normal distribution

Pdf of Bivariate Normal Distribution

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

Definition

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

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

双变量
正态分布

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

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

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.

Properties of Bivariate Normal Distribution

1. Marginal pdf of X and Y are normal with

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

Properties of Bivariate Normal Distribution

1. Marginal pdf of X and Y are normal with

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

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

2. Conditional pdf of X given that $Y = y$ is normal with mean

$$\mu_X + \frac{\sigma_X}{\sigma_Y}\rho(y - \mu_Y)$$

and variance

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

$$(1 - \rho^2)\sigma_X^2$$

i.e.,

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

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

Moreover,

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

Proof of the Two Properties

First, recall that

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy.$$

Clearly, it is equivalent to prove that

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left(-\frac{1}{2}\left(\frac{x - \mu_X}{\sigma_X}\right)^2\right).$$

To this goal, we arrange $f(x, y)$ as follows:

$$f(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1 - \rho^2}} \times$$

$$\exp\left(-\frac{1}{2(1 - \rho^2)}\left[\left(\frac{y - \mu_Y}{\sigma_Y}\right) - \rho\left(\frac{x - \mu_X}{\sigma_X}\right)\right]^2 - \frac{1}{2}\left(\frac{x - \mu_X}{\sigma_X}\right)^2\right)$$

$$= \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left(-\frac{1}{2}\left(\frac{x - \mu_X}{\sigma_X}\right)^2\right) \times h(y|x)$$

Proof of the Two Properties

$$h(y|x) = \frac{1}{\sqrt{2\pi}\sigma_Y\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left[\left(\frac{y-\mu_Y}{\sigma_Y}\right) - \rho\left(\frac{x-\mu_X}{\sigma_X}\right)\right]^2\right)$$

$$= \frac{1}{\sqrt{2\pi}\sqrt{(1-\rho^2)\sigma_Y^2}} \exp\left(-\frac{1}{2}\left[\frac{y-[\mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X)]}{\sqrt{(1-\rho^2)\sigma_Y^2}}\right]^2\right)$$

Clearly, to complete the proof, we only need to show that

$$\int_{-\infty}^{\infty} h(y|x)\,dy = 1$$

which is true because $h(y|x)$ is the probability density function of $N(\mu_Y + \frac{\sigma_Y}{\sigma_X}\rho(x-\mu_X), (1-\rho^2)\sigma_Y^2)$.

Actually, $h(y|x)$ is the conditional probability density function of Y given $X = x$ because

$$h(y|x) = \frac{f(x,y)}{f_X(x)}.$$

Properties of Bivariate Normal Distribution

Inde $\Longleftarrow$ uncorrelation

3. Independence $\Longleftrightarrow$ Uncorrelation

▶ prove that ρ is indeed the correlation coefficient of X and Y
    ▶ $Cov(X, Y) = \rho\sqrt{Var(X)}\sqrt{Var(Y)} = \rho\sigma_X\sigma_Y$
    ▶ $f(x, y) = h(y|x)f_X(x)$
▶ Uncorrelation means $\rho = 0 \implies$

$$f(x, y) = f_X(x)f_Y(y) \quad \text{means independence}$$

The proof is left as a question in the assignment.

Example 1, page 165

Observe a group of college students. Let X and Y denote their grades in high school and in the 1st year in college, respectively, have a bivariate normal distribution with

$$\mu_X = 2.9, \quad \mu_Y = 2.4, \quad \sigma_X = 0.4, \quad \sigma_Y = 0.5 \quad \text{and} \quad \rho = 0.8$$

Find $P(2.1 < Y < 3.3)$ and $P(2.1 < Y < 3.3 | X = 3.2)$.

> **Question**
>
> Observe a group of college students. Let X and Y denote their grades in high school and in the 1st year in college, respectively, have a bivariate normal distribution with
>
> $$\mu_X = 2.9, \quad \mu_Y = 2.4, \quad \sigma_X = 0.4, \quad \sigma_Y = 0.5 \quad \text{and} \quad \rho = 0.8$$
>
> Find $P(2.1 < Y < 3.3)$ and $P(2.1 < Y < 3.3 | X = 3.2)$.

Since $Y \sim N(\mu_Y, \sigma_Y^2) = N(2.4, 0.5^2)$, then

$$P(2.1 < Y < 3.3) = P\left(\frac{2.1 - 2.4}{0.5} < \frac{Y - 2.4}{0.5} < \frac{3.3 - 2.4}{0.5} \right)$$

$$= \Phi(1.8) - \Phi(-0.6) = 0.69$$

$Y \sim N(2.4, 0.5^2)$ $\qquad \frac{2.1-2.4}{0.5} < z < \frac{3.3-2.4}{0.5}$

Example 1, page 165

Note that

$$Y|X = x \sim N\left(\mu_Y + \frac{\sigma_Y}{\sigma_X}\rho(x - \mu_X), (1 - \rho^2)\sigma_Y^2\right)$$
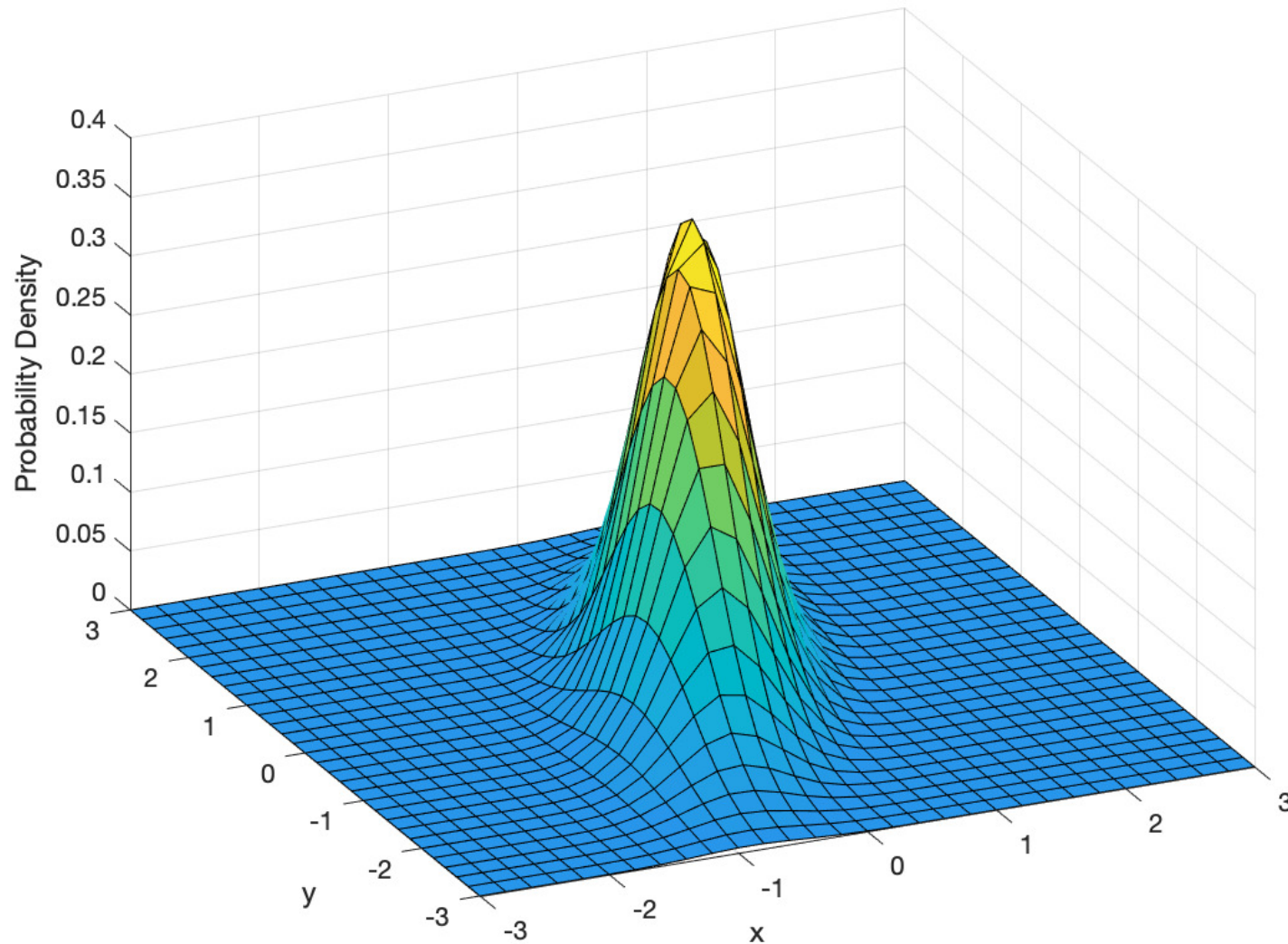
Let $X = 3.2$. Then $Y|X = 3.2 \sim N(2.7, 0.3^2)$ 代入.

$P(2.1 < y < 3.3|X = 3.2)$

$$= P\left(\frac{2.1 - 2.7}{0.3} < \frac{Y - 2.7}{0.3} < \frac{3.3 - 2.7}{0.3}\Big| X = 3.2\right)$$
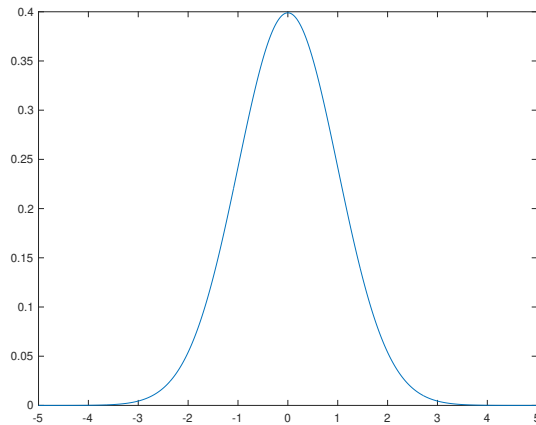
$$= \Phi(2) - \Phi(-2) = 0.95$$

Geometrical interpretation [P.166]

Let $z = f(x, y)$ and draw it in $x - y - z$ 3-dimensional space

Geometrical interpretation [P.166]

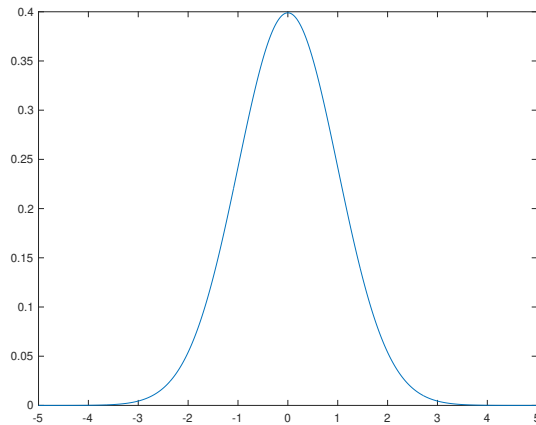Let $z = f(x, y)$ and draw it in $x - y - z$ 3-dimensional space

1. Consider $z = f(x_0, y) = f_X(x_0) h(y|x_0)$

   ▶ the intersection of the surface $z = f(x, y)$ with the plane $x = x_0$, which is parallel to the yz-plane

   ▶ a bell-shaped curve and has the shape of a normal pdf

Geometrical interpretation [P.166]

Let $z = f(x, y)$ and draw it in $x - y - z$ 3-dimensional space

2. Consider $z = f(x, y_0) = f_Y(y_0)g(x|y_0)$

► the intersection of the surface $z = f(x, y)$ with the plane $y = y_0$, which is parallel to the xz-plane

► a bell-shaped curve and has the shape of a normal pdf

Geometrical interpretation [P.166]

3. Consider $z_0 = f(x, y)$ with $0 < z_0 < \dfrac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}$

▶ the intersection of the surface $z = f(x, y)$ with the plane $z = z_0$, which is parallel to the xy-plane

▶ an ellipse

$$\exp[-\frac{1}{2}q(x, y)] = z_0 \cdot 2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}$$

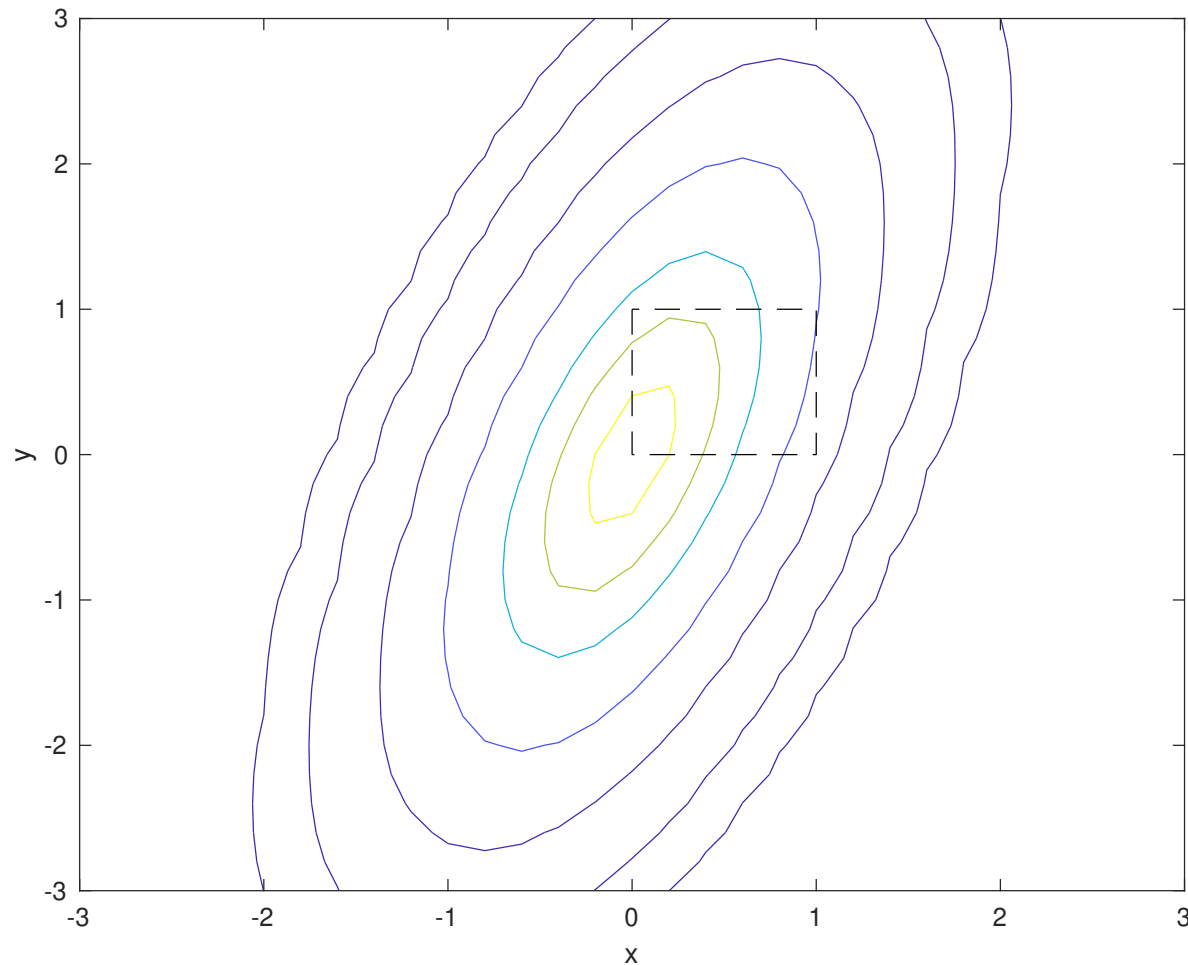Taking logarithm yields 取对数.

$$\left(\frac{x - \mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x - \mu_X}{\sigma_X}\right)\left(\frac{y - \mu_Y}{\sigma_Y}\right) + \left(\frac{Y - \mu_Y}{\sigma_Y}\right)^2$$
$$= -2(1 - \rho^2)\ln(z_0 2\pi\sigma_X\sigma_Y\sqrt{1 - \rho^2})$$

which corresponds to an ellipse.

Geometrical interpretation [P.166]

We can draw these ellipses for different z_0 on the xy-plane, which are called level curves or contours.

Chapter 5. Distribution of Functions of Random Variables

Section 5.1 Function of one random variable

Function of One Random Variable

> **Question**
>
> Let X be a RV of either discrete or continuous type. 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?

已知 X，若 $Y = Y(X)$.
则如何求 Y 分布

的 PMF / PDF

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

Discrete case

Let X be a discrete RV with pmf $f(x) : \overline{S_X} \to (0, 1]$, and $Y = u(X)$ be a one-to-one mapping with inverse $X = v(Y)$.

Then the pmf of Y, denoted by $g(y) : \overline{S_Y} \to (0, 1]$ is

▶ $\overline{S_Y} = \{y | y = u(x), x \in \overline{S_X}\}$

▶ for any $y \in \overline{S_Y}$,

$$g(y) = P(Y = y) = P(u(X) = y) = P(X = v(y)),$$

Since

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

Example 1, page 177

Question

Let X have a Poisson distribution with $\lambda = 4$ and its pmf takes the form of

$$f(x) = \frac{4^x e^{-4}}{x!}, \quad x = 0, 1, 2, \cdots.$$

If $Y = \sqrt{X}$, what is the pmf $g(y)$ of Y?

Example 1, page 177

> ## Question
>
> Let X have a Poisson distribution with $\lambda = 4$ and its pmf takes the form of
>
> $$f(x) = \frac{4^x e^{-4}}{x!}, \quad x = 0, 1, 2, \cdots$$
>
> If $Y = \sqrt{X}$, what is the pmf $g(y)$ of Y?

First, $S_Y = \{0, 1, \sqrt{2}, \sqrt{3}, \cdots\}$, and then

$$Y = u(X) = \sqrt{X} \Longrightarrow X = v(Y) = Y^2$$

$$g(y) = P(Y = y) = P(\sqrt{X} = y) = P(X = y^2)$$

$$= f(y^2) = \frac{4^{y^2} e^{-4}}{(y^2)!}, \quad y = 0, 1, \sqrt{2}, \sqrt{3}, \cdots$$

Continuous Case: the idea

> **Question**
>
> Let X be a continuous RV with pdf $f(x) : [c_1, c_2] \to [0, \infty)$ and $Y = u(X)$ is a function of X.
>
> Our goal is to calculate the pdf of Y, say $g(y)$.

For a continuous RV X, recall the relation between a pdf $f(x)$ and its corresponding cdf $F(x)$:

$$F(x) = \int_0^x f(t)dt \Rightarrow F'(x) = \frac{dF(x)}{dx} = f(x),$$

where $F(x)$ is differentiable.
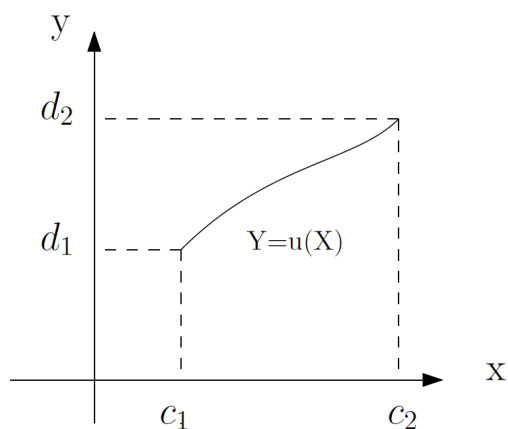
$$F(x) = \int_0^x f(t)dt \qquad F'(x) = \frac{dF(x)}{dx}$$

Continuous Case: Case 1

$Y = u(X)$, CTS, ↗ 反函数.

▶ Case 1: $Y = u(X)$ is continuous, strictly increasing, has inverse function $X = v(Y)$, whose derivative $\frac{dv(y)}{dy}$ exists.

1. Determine the cdf of Y, $G(y)$, $y \in \overline{S_Y}$.
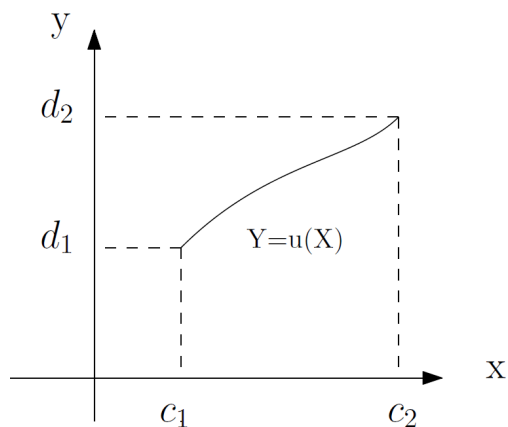
$X = v(Y)$ 且 differentiable

We first find the sample space of Y, $\overline{S_Y}$. Since $Y = u(X)$ is continuous and increasing, $\overline{S_Y} = [d_1,\ d_2]$ with $d_1 = u(c_1)$ and $d_2 = u(c_2)$.

$$G(y) = P(Y \leq y) = P(u(X) \leq y) = P(X \leq v(y)) = \int_{c_1}^{v(y)} f(x)dx$$

$$P(X \leq v(y)) = \int_{c_1}^{v(y)} f(x)\,dx$$

Continuous Case: Case 1

2. Determine the pdf of Y, $g(y)$, $y \in \overline{S_Y}$.

$$g(y) = G'(y) = \frac{dG(y)}{dy} = f(v(y))v'(y)$$

$$= f(v(y))\frac{dv(y)}{dy}$$

$$= f(v(y))\left|\frac{dv(y)}{dy}\right|, \quad y \in \overline{S_Y}$$

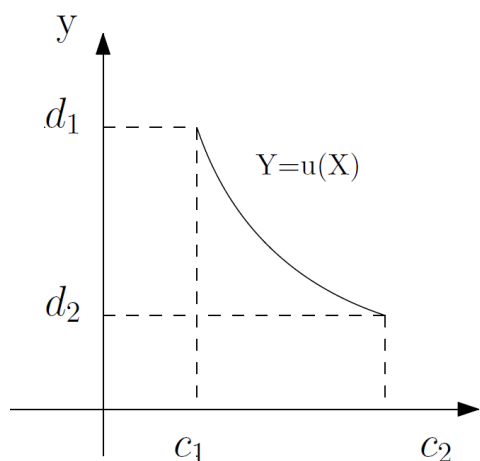$$f(v(y))\left|\frac{dv(y)}{dy}\right|$$

Continuous Case: Case 2

$Y = u(x)$, cts, $\downarrow$, 反函数 $X = v(Y)$, 可导

▶ Case 2: $Y = u(X)$ is continuous, strictly decreasing and has inverse function $X = v(Y)$, whose derivative $\frac{dv(y)}{dy}$ exists.

1. Determine the cdf of Y, $G(y)$, $y \in \overline{S_Y}$.

   We first find the sample space of Y, $\overline{S_Y}$. Since $Y = u(X)$ is continuous and strictly decreasing $\overline{S_Y} = [d_2, \ d_1]$ with $d_1 = u(c_1)$ and $d_2 = u(c_2)$.

$P(Y \le y) = P(u(x) \le y) = P(X \geq v(y))$

$$G(y) = P(Y \le y) = P(u(X) \le y) = P(X \ge v(y))$$

$$= 1 - P(X \le v(y)) = 1 - \int_{c_1}^{v(y)} f(x)dx$$

$\longrightarrow \ 1 - P(X \le v(y)) = 1 - \int_{c_1}^{v(y)} f(x)dx$

Continuous Case: Case 2

2. Determine the pdf of Y, $g(y)$, $y \in \overline{S_Y}$.

$$g(y) = G'(y) = \frac{dG(y)}{dy} = -f(v(y))v'(y)$$

$$= -f(v(y))\frac{dv(y)}{dy}$$

$$= f(v(y))\left|\frac{dv(y)}{dy}\right| \quad y \in \overline{S_Y}$$

Summary: for both strictly increasing and decreasing cases,

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

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

$$Y = (1-X)^3, \quad X = 1 - \sqrt[3]{Y}$$

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

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

1. The sample space of Y is $\overline{S_Y} = (0, 1)$, since $0 < x < 1$.

$$f(v(y)) = 3\left[1 - (1 - \sqrt[3]{y})\right]^2 \cdot \left| -\tfrac{1}{3} y^{-\frac{2}{3}} \right| = 1$$

2.

$$g(y) = f(v(y)) \left| \frac{dv(y)}{dy} \right| \quad \text{where} \quad \frac{dv(y)}{dy} = -\frac{1}{3} y^{-\frac{2}{3}}$$

$$= 3(1 - (1 - y^{\frac{1}{3}}))^2 \left| -\frac{1}{3} y^{-\frac{2}{3}} \right| = 1,$$

$$0 < y < 1 \quad Y \sim U(0, 1)$$

$$0 < y < 1, \quad Y \sim U(0, 1)$$

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 $F(a) = 0, F(b) = 1$. 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 $F(x)$.

$$Y \sim U(0,1) \quad F(x) \uparrow$$

$$F(a) = 0$$
$$F(b) = 1$$

$$X = F^{-1}(Y)$$
反函数

Theorem 5.1-1, page 175

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

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

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

Note that

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

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

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.

$$\int_0^x \frac{1}{\theta} e^{-\frac{t}{\theta}} \, dt = 1 - e^{-\frac{x}{\theta}} = Y$$

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 $\quad 1 - Y = e^{-\frac{x}{\theta}}$

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

$$x = F^{-1}(y) = -\theta \ln(1 - Y)$$