



工程概率统计

Probability and Statistics for Engineering

第三章 联合分布

Chapter 3 Joint Distributions

Chapter 3 Joint Distributions

- 3.1 Random Vector and Joint Distribution
- 3.2 Relationship between Two Random Variables
- 3.3 Function of Multiple Random Variables
- 3.4 Multivariate Normal Distribution



3.2 Relationship Between Two Random Variables

- In the end of Section 3.1, we mentioned the concept of independence between random variables.
- Recall the independence between random events, the independence between random variables can be defined similarly: the value of Y does not affect the distribution of X or vice versa.
- For example, for the continuous case:

$$f_{X|Y}(x|y) = f_X(x) \Rightarrow f(x, y) = f_{X|Y}(x|y)f_Y(y) = f_X(x)f_Y(y),$$

or

$$f_{Y|X}(y|x) = f_Y(y) \Rightarrow f(x, y) = f_{Y|X}(y|x)f_X(x) = f_X(x)f_Y(y).$$

By $f(x, y) = f_X(x)f_Y(y)$, we have

$$F(x, y) = \int_{-\infty}^y \int_{-\infty}^x f(u, v) du dv = \int_{-\infty}^x f_X(u) du \int_{-\infty}^y f_Y(v) dv = F_X(x)F_Y(y).$$



3.2 Relationship Between Two Random Variables

Independence of Random Variables

- Let $F(x_1, x_2, \dots, x_n)$ be the joint CDF of (X_1, X_2, \dots, X_n) , $F_{X_i}(x_i)$ be the marginal CDF of X_i , then if for $\forall x_1, x_2, \dots, x_n \in \mathbb{R}$ we have

$$F(x_1, x_2, \dots, x_n) = F_{X_1}(x_1)F_{X_2}(x_2) \cdots F_{X_n}(x_n),$$

then we say that random variables X_1, X_2, \dots, X_n are **(mutually) independent** (相互独立).

- For discrete random variables X_1, X_2, \dots, X_n , if they are independent, then the PMF satisfies

$$p(x_1, x_2, \dots, x_n) = P(X_1 = x_1, \dots, X_n = x_n) = p_{X_1}(x_1)p_{X_2}(x_2) \cdots p_{X_n}(x_n).$$

- For continuous random variables X_1, X_2, \dots, X_n , if they are independent, then the PDF satisfies

$$f(x_1, x_2, \dots, x_n) = f_{X_1}(x_1)f_{X_2}(x_2) \cdots f_{X_n}(x_n).$$

- Question:** the definition of independence between three or more random events require multiple equations, why the independence between three or more random variables only require one?



3.2 Relationship Between Two Random Variables

Example 3.6

- The PDF of a standard normal random variable Z is $f(z) = ce^{-z^2/2}$, $-\infty < z < \infty$.
- We already know that $c = 1/\sqrt{2\pi}$, however, how is this value obtained?
- Surprisingly, the easiest way to determine c is to define two independent standard normal random variables and use the fact that their joint PDF must integrate to 1.



3.2 Relationship Between Two Random Variables

Solution



3.2 Relationship Between Two Random Variables



Example 3.7

- Suppose that the number of people who enter a shopping mall on a randomly selected weekday follows a Poisson distribution with parameter λ .
- If each person who enters the shopping mall is a male with probability 0.2 and a female with probability 0.8.
- Show that the number of males and females entering the shopping mall are independent Poisson random variables with parameters 0.2λ and 0.8λ , respectively.

Solution



3.2 Relationship Between Two Random Variables

Solution



3.2 Relationship Between Two Random Variables

- If X and Y are independent random variables, then, for any functions g and h , we have

$$E[g(X)h(Y)] = E[g(X)]E[h(Y)],$$

$$\text{Var}[g(X) \pm h(Y)] = \text{Var}[g(X)] + \text{Var}[h(Y)].$$

Proof: Without loss of generality, show the case for the continuous case.

Suppose that X and Y have joint density $f(x, y)$, then

$$\begin{aligned} E[g(X)h(Y)] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x)h(y)f(x, y)dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x)h(y)f_X(x)f_Y(y)dx dy \\ &= \int_{-\infty}^{\infty} g(x)f_X(x)dx \int_{-\infty}^{\infty} h(y)f_Y(y)dy = E[g(X)]E[h(Y)]. \end{aligned}$$

Let $E[g(X)] = a$ and $E[h(Y)] = b$, then

$$\begin{aligned} \text{Var}[g(X) + h(Y)] &= E[(g(X) + h(Y) - a - b)^2] \\ &= E[(g(X) - a)^2] + E[(h(Y) - b)^2] + 2E[(g(X) - a)(h(Y) - b)] \\ &= \text{Var}[g(X)] + \text{Var}[h(Y)]. \end{aligned}$$

$$= E[g(X) - a]E[h(Y) - b] = 0$$



3.2 Relationship Between Two Random Variables

- Special case: $E(XY) = E(X)E(Y)$ and $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$ if X and Y are independent.
- What if X and Y are not independent?
- Think about the difference between $\text{Var}(X + Y)$ and $\text{Var}(X) + \text{Var}(Y)$:
$$\text{Var}(X + Y) - \text{Var}(X) - \text{Var}(Y) = 2E[(X - E(X))(Y - E(Y))] = 2[E(XY) - E(X)E(Y)].$$
- So, if $E[(X - E(X))(Y - E(Y))] \neq 0$, X and Y cannot be independent.
- Therefore, $E[(X - E(X))(Y - E(Y))]$ can be used to measure the relationship between X and Y .

Covariance

- The **covariance (协方差)** between X and Y , denoted by $\text{Cov}(X, Y)$, is defined by
$$\text{Cov}(X, Y) \triangleq E[(X - E(X))(Y - E(Y))] = E(XY) - E(X)E(Y)$$
- If X and Y are independent, then $\text{Cov}(X, Y) = 0$. However, if $\text{Cov}(X, Y) = 0$, X and Y may not be independent, we can only say that X and Y are **uncorrelated (不相关的)**.



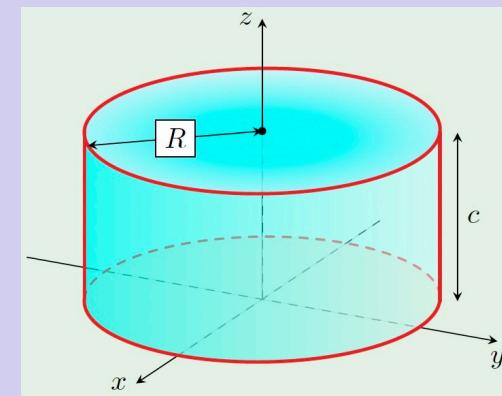
3.2 Relationship Between Two Random Variables

Example 3.4 (Continued)

- We already know that X and Y are not independent, however, show that $\text{Cov}(X, Y) = 0$.

Proof

$$\begin{aligned} E(X) &= E(Y) = 0 \\ E(XY) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f(x,y) dx dy \\ &= \iint_{x^2+y^2 \leq R^2} cxy dx dy \end{aligned}$$



$\text{Cov}(X, Y) = 0$: uncorrelated 不相关
 $\text{Cov}(X, Y) > 0$: positively correlated 正相关
 $\text{Cov}(X, Y) < 0$: negatively correlated 负相关

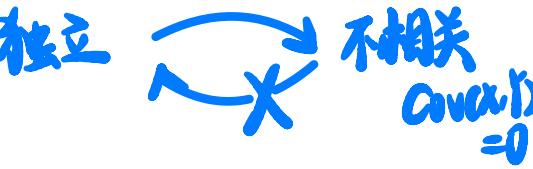
3.2 Relationship Between Two Random Variables

- Note that $\text{Cov}(X, Y)$ is positive when X and Y tend to vary in the same direction and negative when they tend to vary in the opposite direction.

- The covariance has the following properties: (a, b, c are constants)

- Covariance-variance relationship: $\text{Cov}(X, X) = \text{Var}(X)$.
- Symmetry: $\text{Cov}(X, Y) = \text{Cov}(Y, X)$.
- Constants cannot covary: $\text{Cov}(X, c) = 0$.
- Pulling out constants: $\text{Cov}(aX, bY) = ab\text{Cov}(X, Y)$.
- Distributive property: $\text{Cov}(X_1 + X_2, Y) = \text{Cov}(X_1, Y) + \text{Cov}(X_2, Y)$.
- Bilinear property:

$$\text{Cov}(a_1X_1 + a_2X_2 + \cdots + a_nX_n, b_1Y_1 + b_2Y_2 + \cdots + b_mY_m) = \sum_{i=1}^n \sum_{j=1}^m a_i b_j \text{Cov}(X_i, Y_j).$$



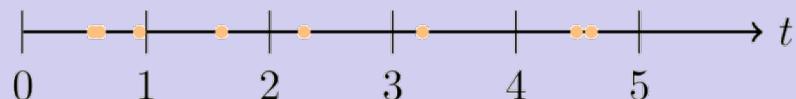
 独立
 不相关
 $\text{Cov}(X, Y) = 0$



3.2 Relationship Between Two Random Variables

Example 3.8

- A Geiger counter (盖革计数器) is a device used for detecting and measuring ionizing radiation (电离辐射).
- Each time it detects a radioactive particle, it makes a clicking sound. E.g., the orange points below indicates the times at which the Geiger counter detect a particle.



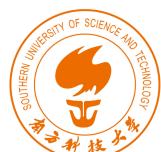
- Suppose that in a city, radioactive particles reach a Geiger counter according to a Poisson process at a rate of $\lambda = 0.8$ particles per second.
- The time that the first particle is detected and the time that the second particle is detected are denoted by X and Y , respectively.
- Calculate the covariance between X and Y .



3.2 Relationship Between Two Random Variables

Solution

Let $Z = Y - X$ 第一个粒子与第二个粒子间隔时间
 $X \sim \text{Exp}(\lambda)$, $Z \sim \text{Exp}(\lambda)$
 $E(X) = E(Z) = \frac{1}{\lambda}$ $\text{Cov}(X, Z) = 0$
 $\text{Cov}(X, Y) = \text{Cov}(X, X + Z) = \text{Cov}(X, X) + \text{Cov}(X, Z)$
 $= \text{Var}(X) = \frac{1}{\lambda^2} = 1.5625$
 $\text{Cov}(X, Y) > 0$ 同样增加



3.2 Relationship Between Two Random Variables

- While the covariance measures the relationship between two random variables, its value depends on the unit/scale on which we measure the random variables.
 - E.g., let X (in m) and Y be the height and weight (in kg) of a randomly selected person, and $\tilde{X} = 100X$ (i.e., \tilde{X} is the height measured in cm) then $\text{Cov}(\tilde{X}, Y) = 100\text{Cov}(X, Y)$. **会受到尺度影响**
 - Therefore, a larger covariance does not necessarily suggest a stronger relationship.
- To make the measure comparable, we need to remove the impact of unit/scale.

Correlation Coefficient

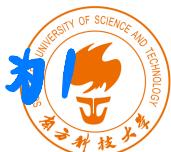
- The **correlation coefficient** (相关系数) between X and Y , denoted by $\text{Cor}(X, Y)$ or ρ_{XY} , is defined by

$$\rho_{XY} = \text{Cor}(X, Y) \triangleq E \left[\frac{(X - E(X))}{\text{SD}(X)} \cdot \frac{(Y - E(Y))}{\text{SD}(Y)} \right] = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}.$$

标准差 消除单位、尺度的影响

- ρ_{XY} is a normalized version of the covariance, which is a dimensionless quantity (无量纲数值).

- Question:** what kind of relationship is ρ_{XX} measuring?
$$\rho_{XX} = \frac{\text{Cov}(X, X)}{\sqrt{\text{Var}(X)\text{Var}(X)}} = \frac{\text{Var}(X)}{\text{Var}(X)} = 1$$
与自身相关系数为1



3.2 Relationship Between Two Random Variables

Example 3.9

- Calculate ρ_{XY} if the joint PDF of X and Y is ($0 < c \leq 1$)

$$f(x, y) = \frac{1}{2\pi c} \exp \left\{ -\frac{x^2 - 2\sqrt{1-c^2}xy + y^2}{2c^2} \right\}, -\infty < x, y < \infty.$$

Solution

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f(x, y) dy = \frac{1}{2\pi c} \int_{-\infty}^{\infty} \exp \left\{ -\frac{(y - \sqrt{1-c^2}x)^2 + cx^2}{2c^2} \right\} dy \\ &= \frac{1}{2\pi c} \int_{-\infty}^{\infty} \exp \left\{ -\frac{(y - \sqrt{1-c^2}x)^2}{2c^2} - \frac{x^2}{2} \right\} dy \\ \Rightarrow N(\sqrt{1-c^2}x, c^2) &\quad \text{This step indicates that the marginal distribution of } X \text{ is a normal distribution with mean } \sqrt{1-c^2}x \text{ and variance } c^2. \\ \Rightarrow \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi} c} \exp \left(-\frac{(y - \sqrt{1-c^2}x)^2}{2c^2} \right) dy &= \frac{1}{2\pi c} \times \sqrt{2\pi} C \times \exp \left(-\frac{x^2}{2} \right) \\ &= \frac{1}{\sqrt{2\pi}} \times \exp \left(-\frac{x^2}{2} \right) \\ &= \frac{1}{\sqrt{2\pi}} \times \exp \left(-\frac{x^2}{2} \right) \end{aligned}$$

$X \sim N(0, 1)$



3.2 Relationship Between Two Random Variables

同样 $Y \sim N(0,1) \Rightarrow E(X)=E(Y)=0$

Solution

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$$

$$\text{Var}(X)=\text{Var}(Y)=1$$

$$\text{let } v = \frac{y - \sqrt{1-c^2}x}{c} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f(x,y) dx dy$$

$$v = y \quad | = \frac{1}{2\pi c} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy \exp \left\{ -\frac{(y - \sqrt{1-c^2}x)^2}{2c^2} - \frac{x^2}{2} \right\} dx dy$$

$$\left| \begin{array}{l} \frac{\partial x}{\partial u} \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} \frac{\partial y}{\partial v} \end{array} \right| = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (cu + \sqrt{1-c^2}v^2) e^{-u^2+v^2/2} du dv$$

$$\left| \begin{array}{l} \frac{\partial x}{\partial u} \frac{\partial y}{\partial v} \\ \frac{\partial y}{\partial u} \frac{\partial y}{\partial v} \end{array} \right| = \frac{c}{2\pi} \int_{-\infty}^{\infty} ue^{-u^2/2} du \int_{-\infty}^{\infty} ve^{-v^2/2} dv + \frac{\sqrt{1-c^2}}{2\pi} \int_{-\infty}^{\infty} e^{-u^2/2} du \int_{-\infty}^{\infty} v^2 e^{-v^2/2} dv$$

奇函数

正态分布



3.2 Relationship Between Two Random Variables

- ρ_{XY} actually measure the direction and strength of **the linear relationship** between X and Y .

Proof: Consider to use a linear function of X to approximate Y , i.e., $\hat{Y} = a + bX$.

Then, the mean squared error (MSE, 均方误差) of the approximation is

$$\begin{aligned}\text{MSE} &= E[(Y - \hat{Y})^2] = E[(Y - a - bX)^2] \\ &= E(Y^2) + b^2 E(X^2) + a^2 - 2bE(XY) + 2abE(X) - 2aE(Y).\end{aligned}$$

Next, we would like to minimize the MSE w.r.t. a and b .

$$\begin{cases} \frac{\partial \text{MSE}}{\partial a} = 2a + 2bE(X) - 2E(Y) = 0 \\ \frac{\partial \text{MSE}}{\partial b} = 2bE(X^2) - 2E(XY) + 2aE(X) = 0 \end{cases} \Rightarrow \begin{cases} b_0 = \frac{E(XY) - E(X)E(Y)}{E(X^2) - [E(X)]^2} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} \\ a_0 = E(Y) - b_0E(X) = E(Y) - E(X) \frac{\text{Cov}(X, Y)}{\text{Var}(X)} \end{cases}$$

Therefore,

$$\begin{aligned}\min_{a,b} \text{MSE} &= E[(Y - a_0 - b_0X)^2] = E[(Y - E(Y) + b_0E(X) - b_0X)^2] \\ &= \text{Var}(Y) + b_0^2 \text{Var}(X) - 2b_0 \text{Cov}(X, Y) = \text{Var}(Y) \left[1 - \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X)\text{Var}(Y)} \right] = \text{Var}(Y)(1 - \rho_{XY}^2).\end{aligned}$$

$$\Rightarrow -1 \leq \rho_{XY} \leq 1$$



3.2 Relationship Between Two Random Variables

- Since $\min_{a,b} \text{MSE} = \text{Var}(Y)(1 - \rho_{XY}^2) \geq 0$, so $\rho_{XY}^2 \leq 1 \Rightarrow -1 \leq \rho_{XY} \leq 1$.

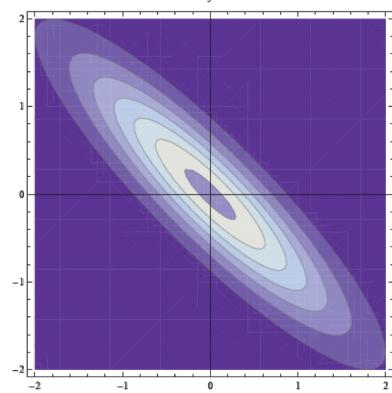
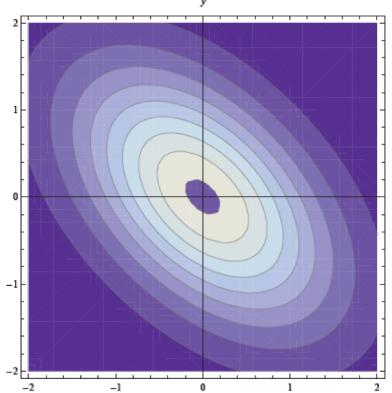
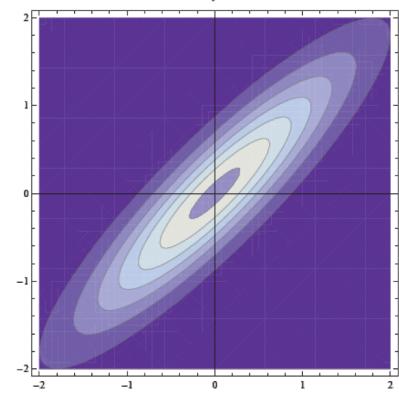
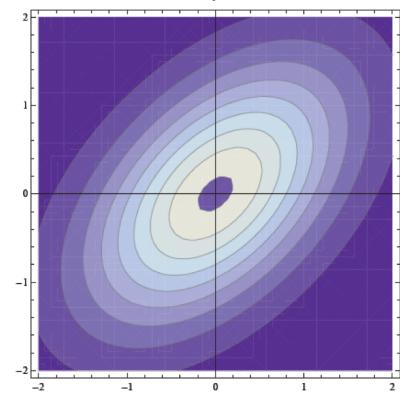
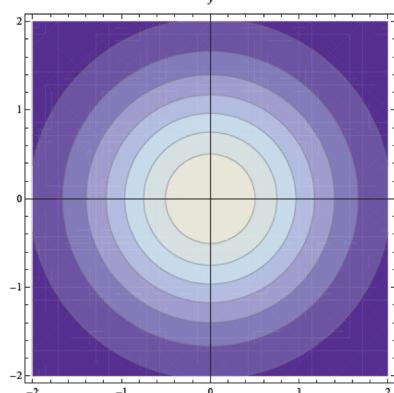
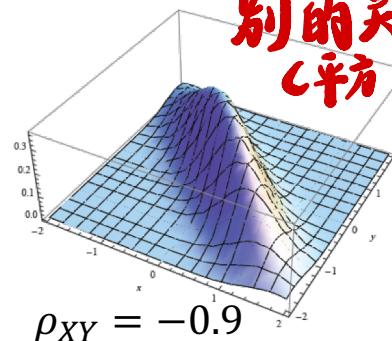
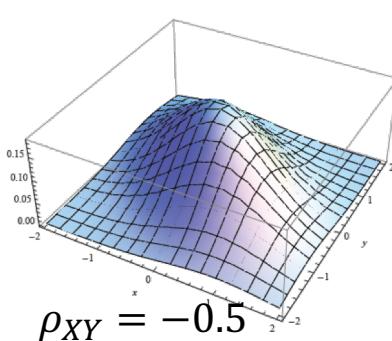
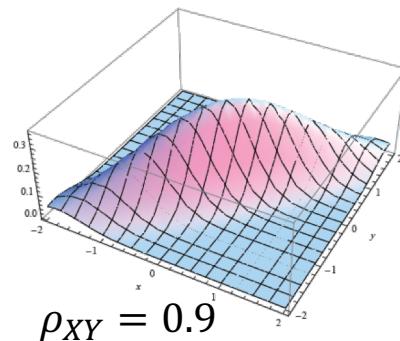
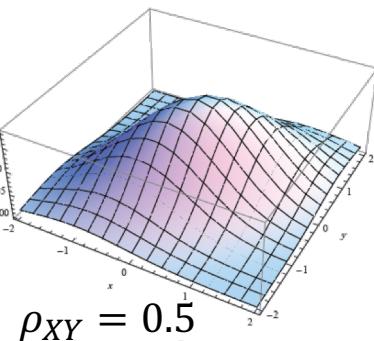
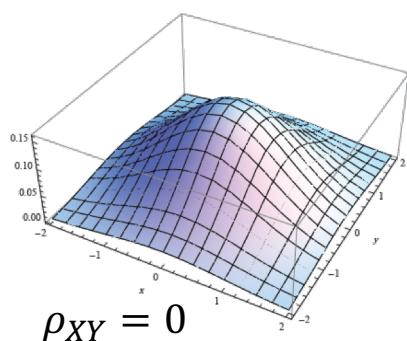
Not necessarily independent!

- $0 < \rho_{XY} \leq 1$: positively correlated; $-1 \leq \rho_{XY} < 0$: negatively correlated; $\rho_{XY} = 0$: uncorrelated.

不存在线性关系

- When $|\rho_{XY}|$ is closer to 1, the mean squared error is smaller, i.e., the relationship between X and Y is closer to linear. Specifically, if $\rho_{XY} = \pm 1$, X and Y have an almost perfect linear relationship.

但可能有
别的关系
(平方...)



3.2 Relationship Between Two Random Variables



Example 3.10

- We would like to invest \$10,000 into shares of companies XX and YY.
- Shares of XX cost \$20 per share and the market analysis shows that the expected return is \$1 per share, with a standard deviation of \$0.5.
- Shares of YY cost \$50 per share, with an expected return of \$2.5 and a SD of \$1.
- What is the optimal portfolio (资产组合) consisting of shares of XX and YY, given their correlation coefficient ρ ? (Note: number of shares can be any non-negative real value)

Solution

投资 10000 dollars, XX: C YY: 10000-C , the resulting return is R_C

let r.v.s X, Y 为每股收益

$$E(R_C) = E\left(X \cdot \frac{C}{20} + Y \cdot \frac{10000-C}{50}\right) = 1 \times \frac{C}{20} + 2.5 \times \frac{10000-C}{50} = 500$$

收益期望不变

找风险最小 (方差)



3.2 Relationship Between Two Random Variables

Solution

$$\text{Var}(R_C) = \text{Var}(X + \frac{C}{20} + Y \times \frac{10000-C}{50})$$

$$\text{Var}(X \pm Y) = \text{Var}(X) + \text{Var}(Y) \pm 2\text{Cov}(X, Y)$$

$$= \left(\frac{C}{20}\right)^2 \text{Var}(X) + \left(\frac{10000-C}{50}\right)^2 \text{Var}(Y) + 2 \frac{C}{20} \frac{10000-C}{50} \text{Cov}(X, Y)$$

$$= \left(\frac{C}{20}\right)^2 \times 0.5^2 + \left(\frac{10000-C}{50}\right)^2 \times 1^2 + 2 \frac{C}{20} \left(\frac{10000-C}{50}\right) \rho \times 0.5 \times 1$$

$$= \left(\frac{41-40\rho}{40000}\right) C^2 - (8-10\rho)C + 10000$$

$$\rho \geq 0.8 \quad \text{二次函数对称轴} \Rightarrow \text{取 } C=0 \quad \text{Var}(R_C) = 10000$$

$$\rho < 0.8 \quad C = \frac{40000(4-5\rho)}{(41-40\rho)}$$

$$\text{Var}(R_C) = 10000 - 40000 \times \frac{(4-5\rho)^2}{41-40\rho}$$

当 $\rho = -1$ 时 $\text{Var}(R_C) = 0$
无风险



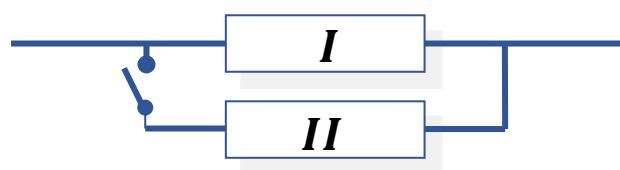
Chapter 3 Joint Distributions

- 3.1 Random Vector and Joint Distribution
- 3.2 Relationship between Two Random Variables
- 3.3 Function of Multiple Random Variables
- 3.4 Multivariate Normal Distribution



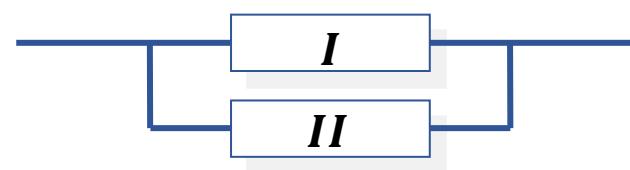
3.3 Function of Multiple Random Variables

- In Section 2.4, we talked about how to determine the distribution of some function of a r.v.
 - Similarly, we may sometimes know the joint distribution of a random vector, e.g., (X, Y) , and would like to derive the distribution of some function of it, e.g., $Z = g(X, Y)$.
 - E.g., we want to derive the distribution of the lifespan of a system consists of two components.



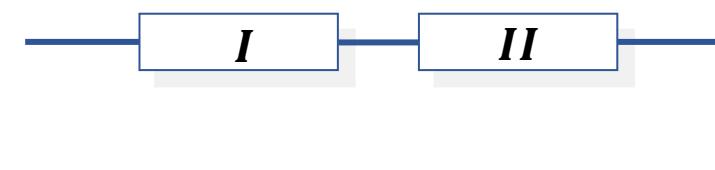
Switch to II if I is broken

System lifespan = $X + Y$



I and II are connected
parallelly

System lifespan = $\max\{X, Y\}$



I and II are connected
in series

$$\text{System lifespan} = \min\{X, Y\}$$



3.3 Function of Multiple Random Variables

- Consider the continuous first.
- The most general solution is to derive the CDF of $Z = g(X, Y)$ starting from the definition of CDF:

$$F_Z(z) = P(Z \leq z) = P(g(X, Y) \leq z) = \iint_{g(x,y) \leq z} f(x, y) dx dy = \dots = \int_{-\infty}^z f_z(u) du.$$

The PDF of Z

The PDF of $Z = X + Y$ – Continuous Case

- Let $f(x, y)$ be the PDF of random vector (X, Y) , $f_X(x)$ and $f_Y(y)$ be the marginal PDF of X and Y , respectively. Then, the PDF of $Z = X + Y$ is

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

- Specifically, if X and Y are **independent**, then

$$f_Z(z) = f_X * f_Y \triangleq \int_{-\infty}^{\infty} f_X(z - y) f_Y(y) dy = \int_{-\infty}^{\infty} f_X(x) f_Y(z - x) dx,$$

- where the two integrals are called the **convolution (卷积)** of f_X and f_Y , denoted as $f_X * f_Y$.



3.3 Function of Multiple Random Variables

- Here we provide the derivation of the PDF of $Z = X + Y$.

Proof:

$$\begin{aligned} F_Z(z) &= P(Z \leq z) = P(X + Y \leq z) = \iint_{x+y \leq z} f(x, y) dx dy \\ &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{z-y} f(x, y) dx \right) dy \quad \text{Let } x = u - y \quad \int_{-\infty}^{\infty} \int_{-\infty}^z f(u - y, y) du dy \\ &= \int_{-\infty}^z \left[\int_{-\infty}^{\infty} f(u - y, y) dy \right] du \Rightarrow f_Z(z) = \int_{-\infty}^{\infty} f(z - y, y) dy. \end{aligned}$$

Similarly, we can show $f_Z(z) = \int_{-\infty}^{\infty} f(x, z - x) dx$.



3.3 Function of Multiple Random Variables



Example 3.8 (Continued)

- Let T_1 be the duration from 0 to the arrival of the first particle, T_2 be the duration from the arrival of the first particle to the arrival of the second particle, ...
- Derive the distribution of time until the r th particle arrives.

Solution

独立同分布 i.i.d

when $r=2$ i.e. $Z = T_1 + T_2$

$$\begin{aligned}\Rightarrow f_Z(z) &= \int_{-\infty}^{\infty} f_{T_1}(t) f_{T_2}(z-t) dt \\ &= \int_0^z \lambda e^{-\lambda t} \lambda e^{-\lambda(z-t)} dt \\ &= \lambda^2 e^{-\lambda z} \int_0^z 1 dt = \lambda^2 z e^{-\lambda z}\end{aligned}$$



3.3 Function of Multiple Random Variables

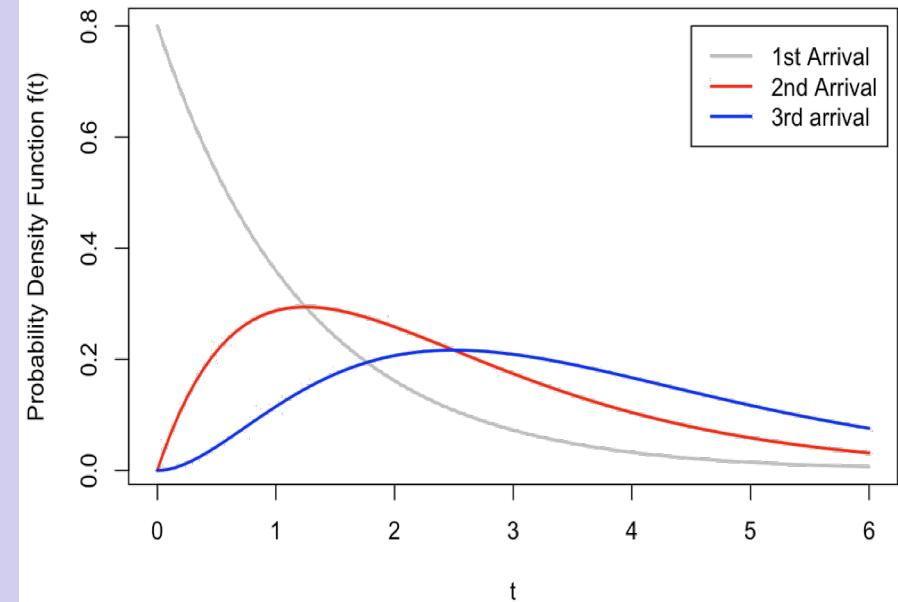
Solution

$$\begin{aligned}f_Z(z) &= \int_{-\infty}^{\infty} f_{T_1+T_2}(t) f_{T_3}(z-t) dt \\&= \int_0^z \lambda^2 t e^{-\lambda t} \lambda e^{-\lambda(z-t)} dt \\&= \lambda^3 e^{-\lambda z} \int_0^z t dt = \frac{\lambda^3 z^2 e^{-\lambda z}}{2}\end{aligned}$$

$$Z = T_1 + T_2 + \dots + T_r$$

$$f_Z(z) = \frac{\lambda^r}{(r-1)!} z^{r-1} e^{-\lambda z}$$

伽马分布



3.3 Function of Multiple Random Variables

Example 3.11

- Let X and Y be independent standard normal random variables, $T = X + Y$.
- You should quickly be able to determine $E(T)$ and $\text{Var}(T)$, but what's the distribution of T ?

Solution

$$\begin{aligned}f_T(t) &= \int_{-\infty}^{\infty} f_X(x) f_Y(t-x) dx \\&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{(t-x)^2}{2}} dx \\&= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2} - \frac{t^2 - 2tx + x^2}{2}} dx \quad \text{经正态上乘} \rightarrow \text{指数为平方项} \\&= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\frac{2x^2 - 2tx + t^2}{2}} dx \\&= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\frac{(x - \frac{t}{2})^2 - \frac{t^2}{4}}{2}} dx \quad \hookrightarrow N(\frac{t}{2}, \frac{1}{2}) \\&= \frac{1}{2\pi} \sqrt{\frac{1}{2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{\frac{1}{2}}} e^{-\frac{(x - \frac{t}{2})^2}{2}} \cdot e^{-\frac{t^2}{4}} dx \\&= \frac{1}{2\sqrt{\pi}} e^{-\frac{t^2}{4}} = \frac{1}{\sqrt{\pi}} e^{-\frac{t^2}{4}} \quad \text{积分1(正态)} \\&\Rightarrow t \sim N(0, 2)\end{aligned}$$



3.3 Function of Multiple Random Variables

General Results about the Sum of Independent Normal Random Variables

- Let X and Y be two independent random variables, $X \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_2, \sigma_2^2)$. Then
只要独立, 相加之和仍为正态 $X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.
- More generally, if random variables X_1, X_2, \dots, X_n are independent and $X_i \sim N(\mu_i, \sigma_i^2)$ ($i = 1, 2, \dots, n$). Then for constants a_1, a_2, \dots, a_n ,
($a_1 \sim a_n$ 不全为0)

$$a_1X_1 + a_2X_2 + \dots + a_nX_n \sim N\left(\sum_{i=1}^n a_i\mu_i, \sum_{i=1}^n a_i^2\sigma_i^2\right).$$

一定先相加

- In summary, a linear combination of independent normal random variables still follows a normal distribution.



3.3 Function of Multiple Random Variables

- The discrete case is similar.

The PDF of $Z = X + Y$ – Discrete Case

- Let X and Y be two discrete random variables, for simplicity, assume that the support of X and Y are both $\{0, 1, 2, \dots\}$, then the PMF of $Z = X + Y$ is: ($k = 0, 1, 2, \dots$)

$$P(Z = k) = \sum_{i=0}^k P(X = i, Y = k - i) = \sum_{j=0}^k P(X = k - j, Y = j).$$

- Specifically, if X and Y are **independent**, then

$$P(Z = k) = \sum_{i=0}^k P(X = i) \cdot P(Y = k - i) = \sum_{j=0}^k P(X = k - j) \cdot P(Y = j),$$

- which is the **discrete convolution formula** (离散卷积公式).

- For $X \sim \text{Poisson}(\lambda_1)$, $Y \sim \text{Poisson}(\lambda_2)$, and X, Y are independent, then $X + Y \sim \text{Poisson}(\lambda_1 + \lambda_2)$ can be shown accordingly. Actually, [Example 3.7](#) serves as an example.
- Therefore, sum of independent Poisson random variables still follows a Poisson distribution.

但不是任何线性组合都服从 poisson $X \sim \text{Poisson}(\lambda)$ $Y = cX$ $E(Y) = c\lambda$ $\text{Var}(Y) = c^2\lambda$

同样：均匀分布也不成立



3.3 Function of Multiple Random Variables

- The sum of random variables $S_n = X_1 + X_2 + \dots + X_n$ appear in many real-life problems, however, determining the exact distribution of S_n is not an easy task generally.
- Each time we add one more random variable, we have to calculate a convolution. What if we work with the sum of hundreds of random variables? Calculating many convolutions is impractical.
- It would be great if there is an approximated distribution of S_n that is accurate and easy to use.
- The **Central Limit Theorem (CLT, 中心极限定理)** provides such an approximation.

Central Limit Theorem for i.i.d. Random Variables

- X_1, X_2, \dots is a sequence of i.i.d. random variables with $\mu \triangleq E(X_i)$ and $\sigma^2 \triangleq \text{Var}(X_i)$. Let $S_n = X_1 + X_2 + \dots + X_n$ and $\bar{X}_n = S_n/n$, consider the standardized version of S_n :

标准化

$$Z_n = \frac{S_n - E(S_n)}{\sqrt{\text{Var}(S_n)}} = \frac{S_n - n\mu}{\sqrt{n}\sigma} = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma}$$

$\xrightarrow{n \rightarrow \infty}$ would be = if X_1, \dots, X_n are normal r.v.s, even for small n .

- As $n \rightarrow \infty$, Z_n converges in distribution (依分布收敛) to a standard normal random variable, that is:

$$F_{Z_n}(z) = P(Z_n \leq z) \xrightarrow{n \rightarrow \infty} \Phi(z) \text{ for all } z.$$

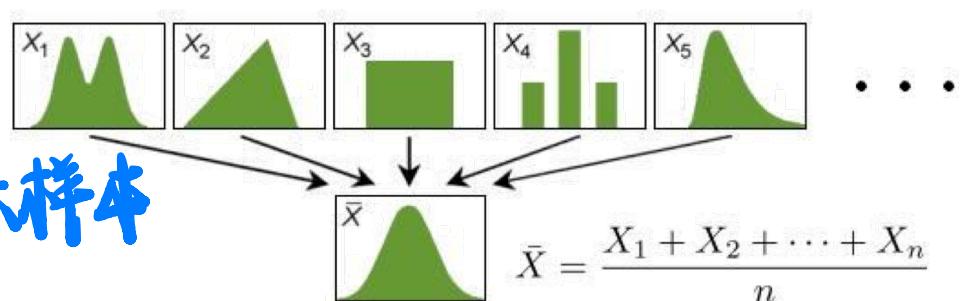
$$Z_n \xrightarrow{d} N(0, 1)$$



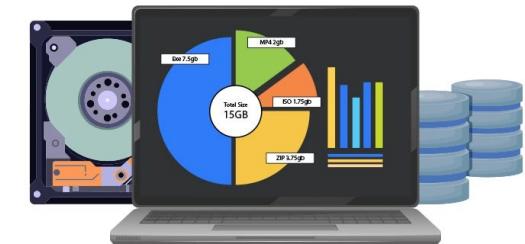
3.3 Function of Multiple Random Variables

只要独立且同分布

- The CLT does not require X_1, X_2, \dots, X_n to follow any specific distribution, so it is a universal behavior across different probability distributions with finite expectation and variance.
- The further research findings are surprising: even if X_1, X_2, \dots, X_n are not independent and do not follow the same distribution, the CLT still holds. Details are omitted here.
- All the complexity and chaos are dissolved under the mysterious curve of the normal distribution.
- Initially, mathematicians refer to this theorem as the Limit Theorem. However, due to its importance in probability theory, the word “central” was added.
- The CLT explains why many measures in reality are normally distributed: they are typically the combined effect of multiple factors. 量会与很多因素相关，综合起来服从正态
- How large n should be to apply the CLT? The rule of thumb (经验法则) is $n \geq 30$.



3.3 Function of Multiple Random Variables



Example 3.12

- A disk has free space of 330 megabytes. Is it likely to be sufficient for 300 independent images, if each image has expected size of 1Mb with a standard deviation of 0.5Mb?

Solution

$$n=300 \quad \mu=1, \sigma=0.5$$

CLT applies to total size S_n

$$\begin{aligned} P(S_n \leq 330) &= P\left(\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \leq \frac{330 - 300 \times 1}{\sqrt{300 \times 0.5^2}}\right) \\ &= \Phi(3.46) = 0.9997 \end{aligned}$$

sufficient



3.3 Function of Multiple Random Variables

伯努利分布加起来是二项分布

- The binomial variable represent a special case of $S_n = X_1 + \dots + X_n$, where $X_i \stackrel{\text{i.i.d.}}{\sim} \text{Bernoulli}(p)$.
- In this case, the exact distribution of S_n is $\text{Binomial}(n, p)$, and consider the approximated distribution of S_n applying the CLT:

对二项分布
标准化
$$\frac{S_n - E(S_n)}{\sqrt{\text{Var}(S_n)}} = \frac{S_n - np}{\sqrt{np(1-p)}} \underset{\text{approx.}}{\sim} N(0, 1) \Rightarrow S_n \underset{\text{approx.}}{\sim} N(np, np(1-p)).$$

- This suggests that the binomial distribution $\text{Binomial}(n, p)$ can be approximated by the normal distribution $N(\mu, \sigma^2)$, where $\mu = np, \sigma = \sqrt{np(1-p)}$ 高频钢板
- This is called the **normal approximation to binomial distribution** (二项分布的正态近似).
- Recall that we talked about the Poisson theorem, which is about the Poisson approximation to binomial distribution (see the PPT of Chapter 2, Page 28-29).
- **Question:** what's the relationship between these two approximations?



3.3 Function of Multiple Random Variables

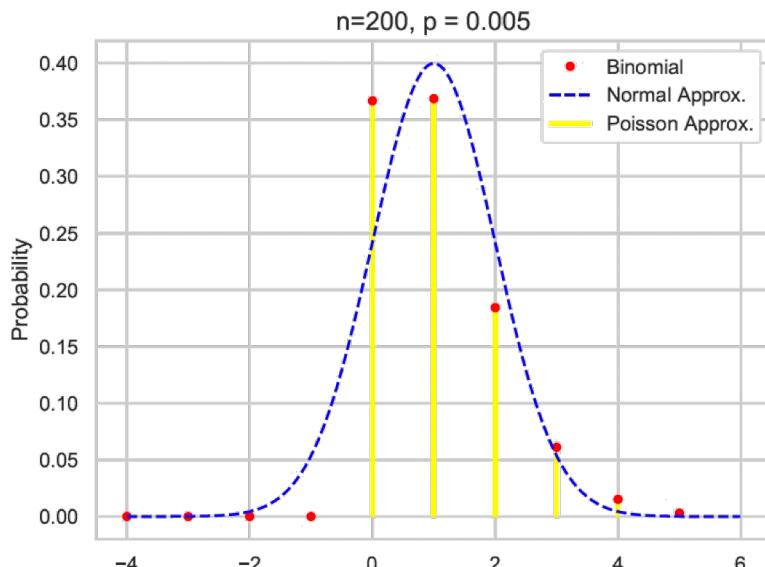
适用场景不同

Poisson approximation

- Binomial(n, p) can be approximated by Poisson(np).
- The approximation works well when n is large and p is small, e.g., $n > 100$ and $p < 0.05$.

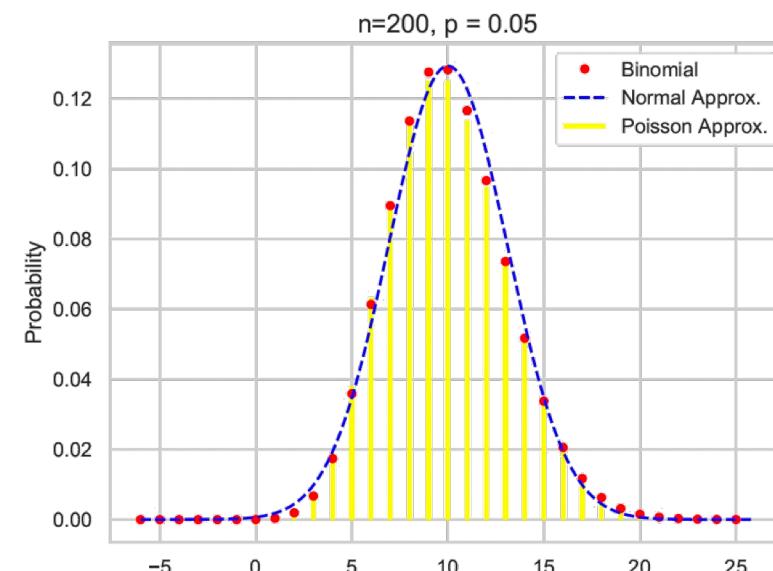


- The best way to understand this is to visualize the three distributions in Python.

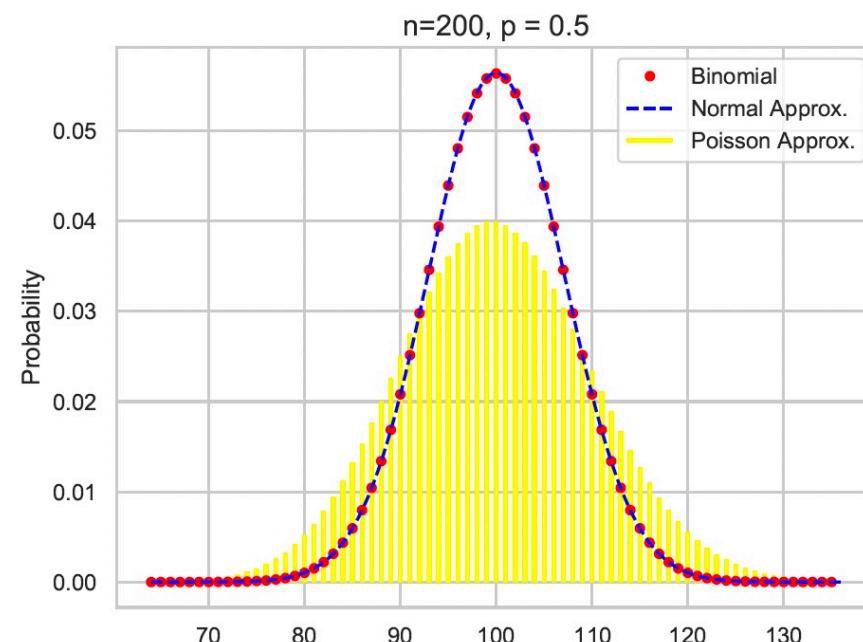
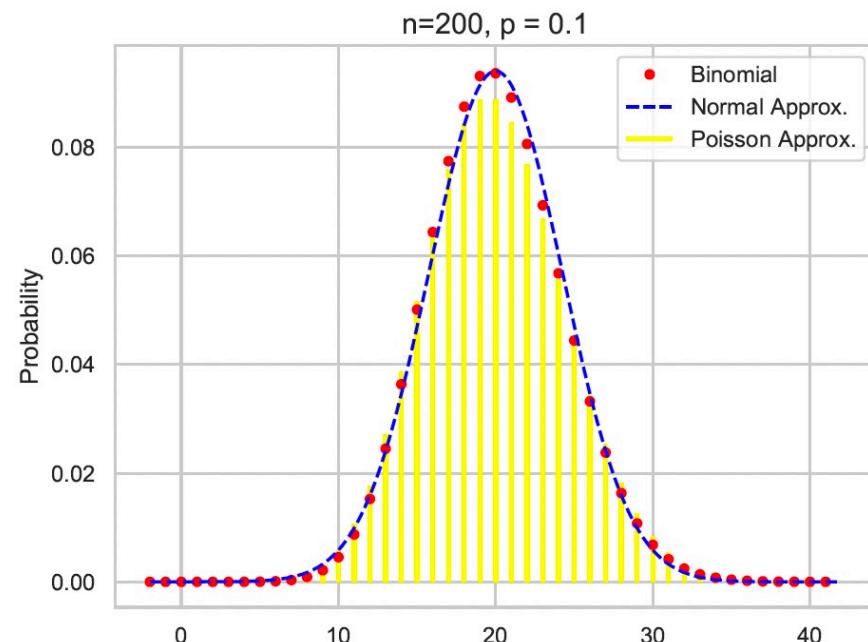


Normal approximation

- Binomial(n, p) can be approximated by $N(np, np(1 - p))$.
- The approximation works well when $np \geq 5$ and $n(1 - p) \geq 5$.



3.3 Function of Multiple Random Variables



- We see that when p is small, the Poisson approximation is better (for a given small value of p , we need larger n for the normal approximation), while the normal approximation is better for large p .
- It is not surprising that the Poisson approximation works poorly for large p if we consider the variance of $\text{Binomial}(n, p)$ and $\text{Poisson}(np)$.

Variance $np(1 - p)$

Variance np

$np(1 - p) \approx np$
only when p is small



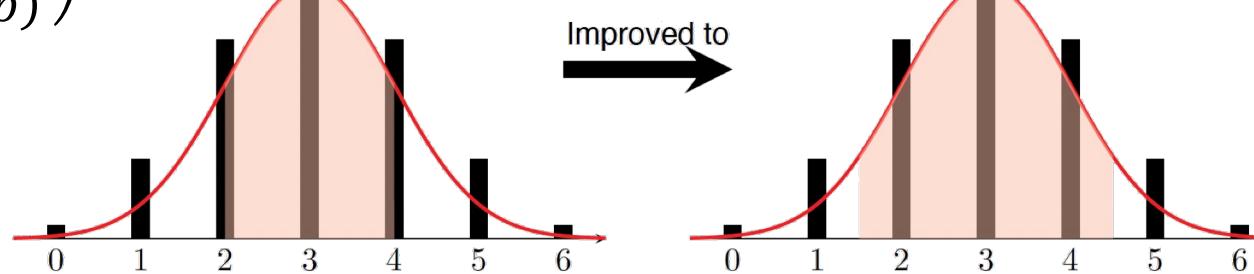
3.3 Function of Multiple Random Variables

- If $X \sim \text{Binomial}(n, p)$ and we want to calculate $P(k \leq X \leq l)$ (k, l are integers) with the normal approximation, note that a **continuity correction (连续性修正)** needs to be applied.
- This correction is needed when we approximate a discrete distribution by a continuous one.
- The essential reason why a correction is needed is that $P(X = x)$ may be positive if X is discrete, whereas it is always 0 for continuous X .

p_{x=k} = P(k - 0.5 ≤ X ≤ k + 0.5) 不改变概率值，只有整数

The continuity correction is to expand the interval by 0.5 in each direction:

$$P(k \leq X \leq l) = P(k - 0.5 \leq X \leq l + 0.5) = P\left(\frac{k - 0.5 - np}{\sqrt{np(1-p)}} \leq \frac{X - np}{\sqrt{np(1-p)}} \leq \frac{l + 0.5 - np}{\sqrt{np(1-p)}}\right)$$
$$\approx \Phi\left(\frac{l + 0.5 - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{k - 0.5 - np}{\sqrt{np(1-p)}}\right).$$



3.3 Function of Multiple Random Variables

Example 3.13

- A new computer virus attacks a folder consisting of 200 files.
- Each file gets damaged with probability 0.2 independently of other files.
- What is the probability that fewer than 50 files get damaged?

$$<50 \Rightarrow \leq 49$$



Solution

$$X \sim BC(200, 0.2)$$

$$X \sim N(200 \times 0.2, 200 \times 0.2 \times 0.8)$$

$$\begin{aligned} P(X < 50) &= P(\underline{X \leq 49}) = P(\underline{X \leq 49.5}) \\ &= P\left(\frac{X - 200 \times 0.2}{\sqrt{200 \times 0.2 \times 0.8}} \leq \frac{49.5 - 200 \times 0.2}{\sqrt{200 \times 0.2 \times 0.8}}\right) \\ &= \Phi(-1.68) = 0.9535 \end{aligned}$$

先等下再expand



3.3 Function of Multiple Random Variables

- Up to this point, we have been talking about the sum of multiple random variables.
- In the following, we consider how to determine the distribution of the maximum/minimum of two random variables, the result can be generalized to multiple random variables.

The CDF of $\max(X, Y)$ and $\min(X, Y)$ - Continuous Case

- Let $f(x, y)$ be the PDF of random vector (X, Y) . Then, the CDFs of $\max(X, Y)$ and $\min(X, Y)$ are

$$F_{\max}(z) = \int_{-\infty}^z \int_{-\infty}^z f(x, y) dx dy, \quad F_{\min}(z) = 1 - \int_z^{\infty} \int_z^{\infty} f(x, y) dx dy.$$

- Let $F_X(x)$ and $F_Y(y)$ be the marginal CDF of X and Y , then if X and Y are **independent**, we have

$$F_{\max}(z) = F_X(z)F_Y(z), \quad F_{\min}(z) = 1 - [1 - F_X(z)][1 - F_Y(z)].$$

- Specifically, if $X_1, X_2, \dots, X_n \stackrel{\text{i.i.d.}}{\sim} F(x)$ with PDF $f(x)$, then the CDFs and PDFs of $\max(X_1, X_2, \dots, X_n)$ and $\min(X_1, X_2, \dots, X_n)$ are

$$F_{\max}(z) = [F(z)]^n, \quad f_{\max}(z) = nf(z)[F(z)]^{n-1},$$

$$F_{\min}(z) = 1 - [1 - F(z)]^n, \quad f_{\min}(z) = nf(z)[1 - F(z)]^{n-1}.$$



3.3 Function of Multiple Random Variables

- Here we provide the derivation of the CDF of $\max(X, Y)$ and $\min(X, Y)$.

Proof:

$$F_{\max}(z) = P(\max(X, Y) \leq z) = P(X \leq z, Y \leq z) = \int_{-\infty}^z \int_{-\infty}^z f(x, y) dx dy.$$

$$\begin{aligned} F_{\min}(z) &= P(\min(X, Y) \leq z) = 1 - P(\min(X, Y) > z) \\ &= 1 - P(X > z, Y > z) = 1 - \int_z^{\infty} \int_z^{\infty} f(x, y) dx dy \end{aligned}$$

When X and Y are independent,

$$F_{\max}(z) = P(X \leq z) \cdot P(Y \leq z) = F_X(z)F_Y(z),$$

$$F_{\min}(z) = 1 - P(X > z) \cdot P(Y > z) = 1 - [1 - F_X(z)][1 - F_Y(z)].$$

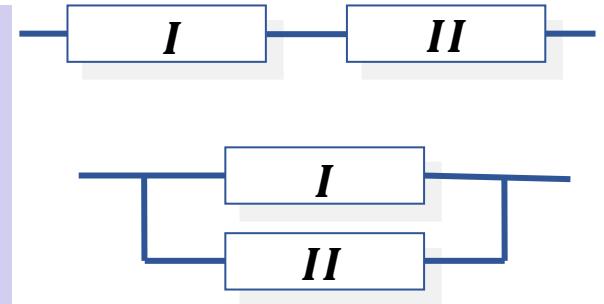
- Suggestion:** don't just memorize the resulting formulas, try to understand the process of derivation.



3.3 Function of Multiple Random Variables

Example 3.14

- A system is made up of two independent components I and II, with lifespan $X_1 \sim \text{Exp}(\lambda_1)$ and $X_2 \sim \text{Exp}(\lambda_2)$, respectively.
- Calculate the expected lifespan of the system in these two scenarios:
(1) I and II are connected in series; (2) I and II are connected parallelly.



Solution



3.3 Function of Multiple Random Variables

Solution



3.3 Function of Multiple Random Variables

- Besides the sum, maximum, minimum functions, the distribution of other functions of multiple random variables can also be derived starting from the definition of CDF and do the integration.

Example 3.15

- X and Y are independent random variables and both follow the distribution $\text{Exp}(1)$.
- Derive the PDF of $Z = X/Y$.

Solution



3.3 Function of Multiple Random Variables

Example 3.16

- A store sells a certain product, where the weekly stock (进货量) and customer demand are independent random variables, both uniformly distributed over the interval (10, 20).
- The store earns a profit of \$1,000 for each unit of the product sold.
- However, if the demand exceeds the stock, the store can order the product from other stores, earning a profit of \$500 per unit in such cases.
- Please calculate the store's expected weekly profit from selling this product.



3.3 Function of Multiple Random Variables

Solution

