Review of Chapters 4-5

Tianshi Chen

The Chinese University of Hong Kong, Shenzhen

Outline

1. Bivariate Random Variable
2. Multivariate Random Variable

Outline

1. Bivariate Random Variable
2. Multivariate Random Variable

Backbone

1. Why to study the bivariate RV?

Backbone

1. Why to study the bivariate RV?
2. What are used to describe the bivariate RV?

Backbone

1. Why to study the bivariate RV?
2. What are used to describe the bivariate RV?
   - full description: pmf or pdf
   - partial description: mathematical expectations

Backbone

1. Why to study the bivariate RV?
2. What are used to describe the bivariate RV?
   - full description: pmf or pdf
   - partial description: mathematical expectations
3. How to derive the pmf or pdf?

Backbone

1. Why to study the bivariate RV?
2. What are used to describe the bivariate RV?
   - full description: pmf or pdf
   - partial description: mathematical expectations
3. How to derive the pmf or pdf?
   - joint ones
   - marginal ones
   - conditional ones

Backbone

1. Why to study the bivariate RV?
2. What are used to describe the bivariate RV?
   - full description: pmf or pdf
   - partial description: mathematical expectations
3. How to derive the pmf or pdf?
   - joint ones
   - marginal ones
   - conditional ones
4. How to calculate the mathematical expectations?

Backbone

1. Why to study the bivariate RV?
2. What are used to describe the bivariate RV?
   - full description: pmf or pdf
   - partial description: mathematical expectations
3. How to derive the pmf or pdf?
   - joint ones
   - marginal ones
   - conditional ones
4. How to calculate the mathematical expectations?
   - by definition
   - by using the marginal and conditional distributions and their mathematical expectations

Backbone

1. Why to study the bivariate RV?
2. What are used to describe the bivariate RV?
   - full description: pmf or pdf
   - partial description: mathematical expectations
3. How to derive the pmf or pdf?
   - joint ones
   - marginal ones
   - conditional ones
4. How to calculate the mathematical expectations?
   - by definition
   - by using the marginal and conditional distributions and their mathematical expectations
5. What typical bivariate random distributions have we learned?

Backbone

1. Why to study the bivariate RV?
2. What are used to describe the bivariate RV?
   - full description: pmf or pdf
   - partial description: mathematical expectations
3. How to derive the pmf or pdf?
   - joint ones
   - marginal ones
   - conditional ones
4. How to calculate the mathematical expectations?
   - by definition
   - by using the marginal and conditional distributions and their mathematical expectations
5. What typical bivariate random distributions have we learned?
   - discrete: trinomial distribution
   - continuous: normal distribution

Why to Study Bivariate RV?

- the outcome of a random experiment is a pair of two scalars.
- to study two random phenomena jointly whose outcome is a scalar, or equivalently, to study two univariate RVs jointly

Full Description: Joint pmf and pdf

Given a pair of discrete or continuous RVs (X,Y) taking values in $\overline{S}$, we define accordingly a pmf or pdf:

1. pmf for discrete RV: $f(x, y) : \overline{S} \to (0, 1]$
   (1) $f(x, y) > 0, \quad (x, y) \in \overline{S}$
   (2) $\sum_{(x,y) \in \overline{S}} f(x) = 1$
   (3) $P((X, Y) \in A) = \sum_{(x,y) \in A} f(x, y), \quad A \subseteq \overline{S}$
2. pdf for continuous RV: $f(x, y) : \overline{S} \to (0, \infty)$
   (1) $f(x, y) > 0, \quad (x, y) \in \overline{S}$
   (2) $\iint_{\overline{S}} f(x, y) dxdy = 1$
   (3) $P((X, Y) \in A) = \iint_A f(x, y) dxdy, \quad A \subseteq \overline{S}$

Partial Description: Mathematical Expectation

$$E[u(X,Y)] = \begin{cases} \sum_{(x,y)\in\overline{S}} u(x,y)f(x,y), & \text{discrete RV} \\ \iint_{\overline{S}} u(x,y)f(x,y)dxdy, & \text{continuous RV} \end{cases}$$

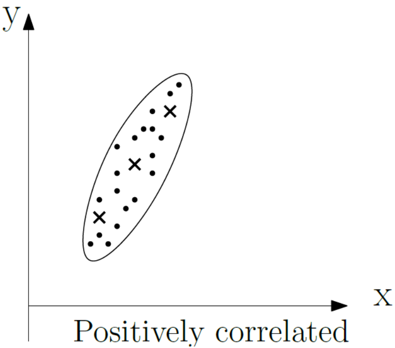
- Covariance and Correlation Coefficient

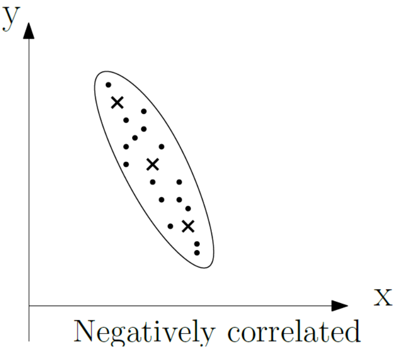
$$Cov(X,Y) = E[(X-E[X])(Y-E[Y])]$$

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(x)}\sqrt{Var(Y)}}, \quad |\rho| \leq 1$$

Covariance and Correlation Coefficient

1. $Cov(X, Y) > 0, (X - E[X])$ and $(Y - E[Y])$ tend to have the same sign; $\rho = 1 \Rightarrow X - E[X] = c(Y - E[Y])$ with $c > 0$
2. $Cov(X, Y) < 0, (X - E[X])$ and $(Y - E[Y])$ tend to have the opposite sign; $\rho = -1 \Rightarrow X - E[X] = c(Y - E[Y])$ with $c < 0$





How to Derive Joint Distributions

1. by definition: go back to the original sample space, calculate the probability, and then derive the joint pmf or pdf

2. independent RV: the joint pmf or pdf is the product of the marginal ones

Independent Random Variables

X and Y are said to be independent if

$$f(x,y) = f_X(x)f_Y(y)$$

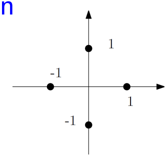
The implication of independent RVs: for any $A \subset \overline{S_X}$ and $B \subset \overline{S_Y}$, the two events $X \in A$ and $Y \in B$ are independent.

A necessary condition for X and Y to be independent

$$\overline{S} = \overline{S_X} \times \overline{S_Y}$$

$$\boxed{\text{Independence} \Rightarrow \text{Uncorrelation}}$$

1. The Converse is not true in general.
2. The Converse is however true for multivariate normal (Gaussian) distribution



How to Derive Marginal Distributions

Given two discrete or continuous RVs (X,Y) taking values in $\overline{S}$ and their joint pmf or joint pdf $f(x,y)$, define accordingly the marginal pmf or marginal pdf to assign the probability of events for RV X:

$$\overline{S_X} = \{\text{all possible values of } X\}, \overline{S_Y}(x) = \left\{y|(x,y) \in \overline{S}\right\}, \text{for } x \in \overline{S_X}$$
$$\overline{S_Y} = \{\text{all possible values of } Y\}, \overline{S_X}(y) = \left\{x|(x,y) \in \overline{S}\right\}, \text{for } y \in \overline{S_Y}$$

1. marginal pmf for discrete RV: $\quad f_X(x) : \overline{S_X} \to (0,1]$

$$f_X(x) = \sum_{y \in \overline{S_Y}(x)} f(x,y)$$

2. marginal pdf for continuous RV: $\quad f_X(x) : \overline{S_X} \to (0,\infty)$

$$f_X(x) = \int_{\overline{S_Y}(x)} f(x,y)dy$$

How to Derive Conditional Distributions

Conditional pmf and pdf

Given two discrete or continuous RVs (X,Y) taking values in $\overline{S}$ and their joint pmf or joint pdf $f(x, y)$, and marginal pmf or marginal pdf $f_X(x)$, define accordingly the conditional pmf or conditional pdf to assign the probability of events for RV Y given that $X = x$:

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

In particular, for $A \subseteq \overline{S_Y}(x)$,

$$P(Y \in A|X = x) = \sum_{y \in A} h(y|x), \quad \text{or} \quad \int_{y \in A} h(y|x) dy$$

Conditional mathematical expectation

$$E[u(Y)|X = x] = \begin{cases} \sum_{y \in \overline{S_Y}(x)} u(y) h(y|x), & \text{discrete RV} \\ \int_{\overline{S_Y}(x)} u(y) h(y|x) dy, & \text{continuous RV} \end{cases}$$

How to Calculate Mathematical Expectations?

- by definition

How to Calculate Mathematical Expectations?

- by definition
- by using the marginal and conditional distributions and their mathematical expectations

$$
\begin{aligned}
Cov(X, Y) &= \sum_{(x,y)\in\overline{S}} (x - E[X])(y - E[Y])f(x, y) \\
&= \sum_{x\in\overline{S_X}} (x - E[X]) \sum_{y\in\overline{S_Y}(x)} (y - E[Y])f(x, y) \\
&= \underbrace{\sum_{x\in\overline{S_X}} (x - E[X])f_X(x)}_{\text{expectation of function of } X} \underbrace{\sum_{y\in\overline{S_Y}(x)} (y - E[Y])h(y|x)}_{\text{expectation of function of } Y|X=x}
\end{aligned}
$$

The calculation of $Cov(X, Y)$ is converted to that of mathematical expectation of functions of two univariate RV, which could be easier if the distributions of two univariate RV are known.

Typical Bivariate Distribution: Trinomial Distribution

Description: The random experiment has three mutually exclusive and exhaustive outcomes:

- "perfect",
- "second"
- "defective"

We repeat the experiment n independent times, and moreover, the probabilities

- p_X: the probability of "perfect",
- p_Y: the probability of "second"
- p_Z: the probability of "defective"

remain the same for each repetition. Such n repetitions can be called a trinomial experiment.

For the trinomial experiment, we are interested in the number of perfects, the number of seconds and the number of defectives.

Typical Bivariate Distribution: Trinomial Distribution

If $(X, Y) \sim \text{Trinomial}(n, p_X, p_Y)$, then

1. Joint pmf:

$$f(x,y) = \frac{n!}{x!y!(n-x-y)!} p_X^x p_Y^y (1 - p_X - p_Y)^{n-x-y}, (x,y) \in \overline{S},$$
$$\overline{S} = \{(x,y) | x + y \leq n, x = 0, 1, \cdots, n, y = 0, 1, \cdots, n\}$$

2. Marginal pmf: $X \sim b(n, p_X)$ and $Y \sim b(n, p_Y)$
3. Conditional pmf:

$$Y|X = x \sim b(n - x, \frac{p_Y}{1 - p_X}), \quad X|Y = y \sim b(n - y, \frac{p_X}{1 - p_Y})$$

4. Covariance and correlation coefficient:

$$Cov(X, Y) = -np_X p_Y$$

Typical Bivariate Distribution: Normal Distribution

Definition

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

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

where $|\rho| < 1$ and

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

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.

Typical Bivariate Distribution: Normal Distribution

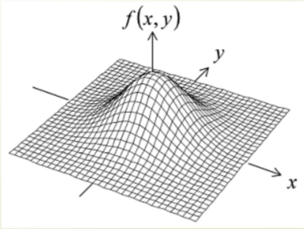
1. Marginal distribution is normal

$$X \sim N(\mu_X, \sigma_X^2) \text{ and } Y \sim N(\mu_Y, \sigma_Y^2)$$

2. Conditional distribution is normal
   given $X = x$, Y is normally distribution with

$$E[Y|X = x] = \mu_Y + \rho\sigma_Y \frac{x - \mu_X}{\sigma_X},$$
$$Var(Y|X = x) = (1 - \rho^2)\sigma_Y^2.$$

3. Uncorrelation is equivalent to independence



Outline

1. Bivariate Random Variable
2. Multivariate Random Variable or Functions of Multivariate Random Variable

Backbone

1. Why to study the multivariate RV?

Backbone

1. Why to study the multivariate RV?
2. What are used to describe the multivariate RV?

Backbone

1. Why to study the multivariate RV?
2. What are used to describe the multivariate RV?
   - full description: pmf or pdf
   - partial description: mathematical expectations

Backbone

1. Why to study the multivariate RV?
2. What are used to describe the multivariate RV?
    - full description: pmf or pdf
    - partial description: mathematical expectations
3. How to derive the pmf or pdf?

Backbone

1. Why to study the multivariate RV?
2. What are used to describe the multivariate RV?
   - full description: pmf or pdf
   - partial description: mathematical expectations
3. How to derive the pmf or pdf?
   - joint ones
   - marginal ones
   - conditional ones

Backbone

1. Why to study the multivariate RV?
2. What are used to describe the multivariate RV?
   - full description: pmf or pdf
   - partial description: mathematical expectations
3. How to derive the pmf or pdf?
   - joint ones
   - marginal ones
   - conditional ones
4. How to calculate the mathematical expectations?

Backbone

1. Why to study the multivariate RV?
2. What are used to describe the multivariate RV?
   - full description: pmf or pdf
   - partial description: mathematical expectations
3. How to derive the pmf or pdf?
   - joint ones
   - marginal ones
   - conditional ones
4. How to calculate the mathematical expectations?
   - by definition
   - by using the marginal and conditional distributions and their mathematical expectations

Backbone

1. Why to study the multivariate RV?
2. What are used to describe the multivariate RV?
   - full description: pmf or pdf
   - partial description: mathematical expectations
3. How to derive the pmf or pdf?
   - joint ones
   - marginal ones
   - conditional ones
4. How to calculate the mathematical expectations?
   - by definition
   - by using the marginal and conditional distributions and their mathematical expectations
5. How to derive the distribution of the functions of independent multivariate random variable?

Backbone

1. Why to study the multivariate RV?
2. What are used to describe the multivariate RV?
   - full description: pmf or pdf
   - partial description: mathematical expectations
3. How to derive the pmf or pdf?
   - joint ones
   - marginal ones
   - conditional ones
4. How to calculate the mathematical expectations?
   - by definition
   - by using the marginal and conditional distributions and their mathematical expectations
5. How to derive the distribution of the functions of independent multivariate random variable?
   - by definition
   - by mgf technique

Why to Study Multivariate RV?

- the outcome of a random experiment is a tuple of several scalars.
- to study several random phenomena jointly whose outcome is a scalar, or equivalently, to study several univariate RVs jointly

Full and Partial Description

Joint pmf and pdf:

1. pmf for discrete RV: $f(x_1, \cdots, x_n) : \overline{S} \to (0, 1]$
   (1) $f(x_1, \cdots, x_n) > 0, \quad (x_1, \cdots, x_n) \in \overline{S}$
   (2) $\sum_{(x_1, \cdots, x_n) \in \overline{S}} f(x_1, \cdots, x_n) = 1$
   (3) $P((X_1, \cdots, X_n) \in A) = \sum_{(x_1, \cdots, x_n) \in A} f(x_1, \cdots, x_n), \quad A \subset \overline{S}$
2. pdf for continuous RV: $f(x_1, \cdots, x_n) : \overline{S} \to (0, \infty)$
   (1) $f(x_1, \cdots, x_n) > 0, \quad (x_1, \cdots, x_n) \in \overline{S}$
   (2) $\int_{\overline{S}} f(x_1, \cdots, x_n) dx_1, \cdots, dx_n = 1$
   (3) $P((X_1, \cdots, X_n) \in A) = \int_A f(x_1, \cdots, x_n) dx_1, \cdots, dx_n, \quad A \subset \overline{S}$

Full and Partial Description

Joint pmf and pdf:

1. pmf for discrete RV: $f(x_1, \cdots, x_n) : \overline{S} \to (0, 1]$
   (1) $f(x_1, \cdots, x_n) > 0, \quad (x_1, \cdots, x_n) \in \overline{S}$
   (2) $\sum_{(x_1, \cdots, x_n) \in \overline{S}} f(x_1, \cdots, x_n) = 1$
   (3) $P((X_1, \cdots, X_n) \in A) = \sum_{(x_1, \cdots, x_n) \in A} f(x_1, \cdots, x_n), \quad A \subset \overline{S}$
2. pdf for continuous RV: $f(x_1, \cdots, x_n) : \overline{S} \to (0, \infty)$
   (1) $f(x_1, \cdots, x_n) > 0, \quad (x_1, \cdots, x_n) \in \overline{S}$
   (2) $\int_{\overline{S}} f(x_1, \cdots, x_n) dx_1, \cdots, dx_n = 1$
   (3) $P((X_1, \cdots, X_n) \in A) = \int_A f(x_1, \cdots, x_n) dx_1, \cdots, dx_n, \quad A \subset \overline{S}$

Mathematical expectation of a function $u(X_1, X_2, \cdots, X_n)$:

$$E[u(X_1, X_2, \cdots, X_n)] = \begin{cases} \displaystyle\sum_{(x_1, \cdots, x_n) \in \overline{S}} u(x_1, \cdots, x_n) \cdot f(x_1, \cdots, x_n) & \text{“discrete RVs”} \\ \int_{\overline{S}} u(x_1, \cdots, x_n) f(x_1, \cdots, x_n) dx_1, \cdots, dx_n & \text{“continuous RV”} \end{cases}$$

Note: Mathematical Expectation is a linear operator.

How to Derive Joint Distributions

1. by definition: go back to the original sample space, calculate the probability, and then derive the joint pmf or pdf, which is in general very complicated

2. independent RV: the joint pmf or pdf is the product of the marginal ones and it is thus in general assumed that the multivariate RVs $X_1, \cdots, X_n$ are independent

As a result of the independence of $X_1, \cdots, X_n$, it is easier to derive

- marginal and conditional distributions, which are same
- mathematical expectation, which still involves multiple summation or integral

Independent Random Variables

RVs $X_1, X_2, \cdots, X_n$ are said to be independent if

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

A necessary condition for $X_1, X_2, \cdots, X_n$ to be independent

$$\overline{S} = \overline{S}_{X_1} \times \cdots \overline{S}_{X_n}$$

Random sample of size n from a common distribution: n independent and identically distributed, i.i.d., RVs $X_1, X_2, \cdots, X_n$.

Functions of Independent Multivariate RVs

- by definition: for functions of one RV

Assume $Y = u(X)$ has an inverse function $X = v(Y)$ and for continuous case, further assume $Y = u(X)$ is strictly increasing or decreasing and $v'(y)$ exists. Then the pmf or pdf of Y is

$$g(y) = \begin{cases} f(v(y)), & y \in u(\overline{S}) \quad \text{discrete RV} \\ f(v(y))|v'(y)|, & y \in u(\overline{S}) \quad \text{continuous RV} \end{cases}$$

For continuous RV:

1. cdf of Y: $G(y) = P(Y \leq y), \qquad y \in u(\overline{S})$
2. pdf of Y: $g(y) = G'(y), \qquad y \in u(\overline{S})$

Note: You need to check whether the assumptions hold before you apply the above formulas!

Functions of Independent Multivariate RVs

- by definition: for functions of one RV
- by mgf technique: for more general cases

Theorem 5.3-1

Assume $X_1, X_2, \cdots, X_n$ are independent and $Y = u_1(X_1)$ $u_2(X_2) \cdots u_n(X_n)$. If $E[u_i(X_i)]$, $i = 1, 2, \cdots, n$, exist, then

$$E(Y) = E[u_1(X_1) \cdots u_n(X_n)] = E[u_1(X_1)] \cdots E[u_n(X_n)]$$

Functions of Independent Multivariate RVs

Theorem 5.3-2

If $X_1, X_2, \cdots, X_n$ are independent random variables with respective means $\mu_1, \mu_2, \cdots \mu_n$ and variances $\sigma_1^2, \sigma_2^2, \cdots, \sigma_n^2$, then the mean and the variance of

$$Y = \sum_{i=1}^{n} a_i X_i,$$

where $a_1, a_2, \cdots, a_n$ are constants, are, respectively,

$$E(Y) = \sum_{i=1}^{n} a_i \mu_i \quad \text{and} \quad Var(Y) = \sum_{i=1}^{n} a_i^2 \sigma_i^2.$$

Sample Mean and Sample Variance

Definition

Let $X_1, X_2, \cdots, X_n$ be independent and identically distributed with mean μ and variance σ^2. Then

- the sample mean is defined as

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

  and an <u>estimator</u> of mean μ, because $E(\overline{X}) = \mu$.

- the sample variance is defined as

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2,$$

  and an <u>estimator</u> of the variance σ^2, because $E(S^2) = \sigma^2$.

Moment Generating Function Technique

Mgf, if exists, uniquely determines the distribution of the RV. So distribution of a random variable can be found through its mgf, e.g., normal or Gaussian $N(\mu, \sigma^2)$: $M(t) = \exp(\mu t + \frac{1}{2}\sigma^2 t^2)$

Theorem 5.4-1

If $X_1, X_2, \cdots, X_n$ are independent random variables with respective moment generating functions $M_{X_i}(t), i = 1, 2, 3, \cdots, n$, where $-h_i < t < h_i, i = 1, 2, \cdots, n$, for positive numbers $h_i, i = 1, 2, \cdots, n$, then the moment-generating function of $Y = \sum_{i=1}^{n} a_i X_i$ is

$$M_Y(t) = \prod_{i=1}^{n} M_{X_i}(a_i t), \text{where} -h_i < a_i t < h_i, i = 1, 2, \cdots, n.$$

Functions of Independent Multivariate χ^2 RVs

Theorem 5.4-2

Let $X_1, X_2, \cdots, X_n$ be independent chi-square random variables with $r_1, r_2, \cdots, r_n$ degrees of freedom, respectively. Then $Y = X_1 + X_2 + \cdots + X_n$ is $\chi^2(r_1 + r_2 + \cdots + r_n)$.

Corollary 5.4-2

Let $Z_1, Z_2, \cdots, Z_n$ have standard normal distributions, $N(0, 1)$. If these random variables are independent, then $W = Z_1^2 + Z_2^2 + \cdots + Z_n^2$ has a distribution that is $\chi^2(n)$.

Functions of Independent Multivariate Normal RVs(1/2)

Theorem 5.5-1

If $X_1, X_2, \cdots, X_n$ are n independent normal variables with means $\mu_1, \mu_2, \cdots, \mu_n$ and variances σ_1^2, σ_2^2, $\cdots$, σ_n^2 , respectively, then

$$Y = \sum_{i=1}^{n} a_i X_i$$

has the normal distribution

$$Y \sim N\left(\sum_{i=1}^{n} a_i \mu_i, \sum_{i=1}^{n} a_i^2 \sigma_i^2\right)$$

Functions of Independent Multivariate Normal RVs(1/2)

Theorem 5.5-2

Let $X_1, X_2, \cdots, X_n$ be a random sample of size n from the normal distribution $N(\mu, \sigma^2)$. Then the sample mean $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ and the sample variance $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$ are independent, and

$$\frac{(n-1)S^2}{\sigma^2} = \sum_{i=1}^{n}\left(\frac{X_i - \overline{X}}{\sigma}\right)^2 \sim \chi^2(n-1)$$

$$\sum_{i=1}^{n}\left(\frac{X_i - \overline{X}}{\sigma}\right)^2 \sim \chi^2(n-1)$$

$$\sum_{i=1}^{n}\left(\frac{X_i - \mu}{\sigma}\right)^2 \sim \chi^2(n)$$

Functions of Independent Multivariate Normal RVs(2/2)

Theorem 5.5-3 (Student's t distribution)

Let

$$T = \frac{Z}{\sqrt{U/r}}$$

where $Z \sim N(0,1), U \sim \chi^2(r)$, and Z and U are independent. Then T has a student's t distribution

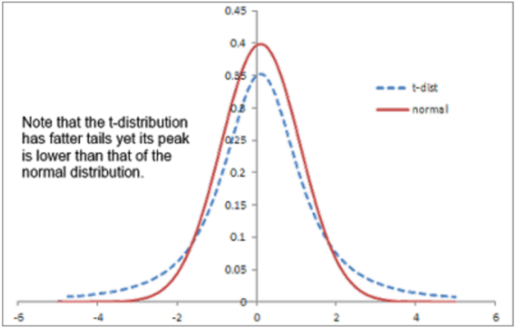
$$f(t) = \frac{\Gamma(\frac{r+1}{2})}{\sqrt{\pi r}\Gamma(\frac{r}{2})} \frac{1}{(1+\frac{t^2}{r})^{\frac{r+1}{2}}}, \quad -\infty < t < \infty$$

Student's t distribution is a heavy tailed distribution in contrast with the normal distribution.

Functions of Independent Multivariate Normal RVs(2/2)

$$T = \frac{\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{(n-1)S^2}{\sigma^2}/(n-1)}} = \frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$$

$$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$$



Heavy-tailed distribution

Central Limit Theorem

Definition

A sequence of random variables Z_1, Z_2, ... is said to converge in distribution to a random variable Z, denoted by $Z_n \xrightarrow{d} Z$, if

$$\lim_{n\to\infty} F_n(z) = F(z),$$

for every number $z \in R$ at which $F(z)$ is continuous, where $F_n(z)$ and $F(z)$ are the cdfs of random variables Z_n and Z, respectively.

Central Limit Theorem

Definition

A sequence of random variables Z_1, Z_2, ... is said to converge in distribution to a random variable Z, denoted by $Z_n \xrightarrow{d} Z$, if

$$\lim_{n \to \infty} F_n(z) = F(z),$$

for every number $z \in R$ at which $F(z)$ is continuous, where $F_n(z)$ and $F(z)$ are the cdfs of random variables Z_n and Z, respectively.

CLT

Let $\overline{X}$ be the sample mean of the random sample of size n, $X_1, X_2, \cdots, X_n$ from a distribution with a finite mean μ and a finite nonzero variance σ^2, then as $n \to \infty$, the sequence of random variables $\frac{\overline{X} - \mu}{\sigma / \sqrt{n}} \xrightarrow{d} N(0, 1)$.

Practical use of CLT

Practical use of CLT: for large n,

- $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$ can be approximated by $N(0,1)$.
- $\overline{X}$ can be approximated by $N(\mu, \frac{\sigma^2}{n})$.
- $\sum_{i=1}^{n} X_i$ can be approximated by $N(n\mu, n\sigma^2)$.

Practical use of CLT

Practical use of CLT: for large n,

- $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$ can be approximated by $N(0,1)$.
- $\overline{X}$ can be approximated by $N(\mu, \frac{\sigma^2}{n})$.
- $\sum_{i=1}^{n} X_i$ can be approximated by $N(n\mu, n\sigma^2)$.

For large n, the probabilities of events of $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$, $\overline{X}$ and $\sum_{i=1}^{n} X_i$ can be calculated approximately by treating them as if they are $N(0,1)$, $N(\mu, \frac{\sigma^2}{n})$, and $N(n\mu, n\sigma^2)$, respectively, and by looking up tables of normal distributions.

Practical use of CLT

Practical use of CLT: for large n,

- $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$ can be approximated by $N(0,1)$.
- $\overline{X}$ can be approximated by $N(\mu, \frac{\sigma^2}{n})$.
- $\sum_{i=1}^{n} X_i$ can be approximated by $N(n\mu, n\sigma^2)$.

For large n, the probabilities of events of $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$, $\overline{X}$ and $\sum_{i=1}^{n} X_i$ can be calculated approximately by treating them as if they are $N(0,1)$, $N(\mu, \frac{\sigma^2}{n})$, and $N(n\mu, n\sigma^2)$, respectively, and by looking up tables of normal distributions.

Recall that if $Y \sim N(\mu, \sigma^2)$

$$
\begin{aligned}
P(a \leq Y \leq b) &= P\left(\frac{a-\mu}{\sigma} \leq \frac{Y-\mu}{\sigma} \leq \frac{b-\mu}{\sigma}\right) \\
&= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)
\end{aligned}
$$

where $\Phi(\cdot)$ is the cdf of $N(0,1)$

Approximation for discrete distribution

To find a continuous distribution whose pdf is "close" to the histogram of the discrete distribution.

Let $Y = \sum_{i=1}^{n} X_i$, where $X_1, \cdots, X_n$ are i.i.d. random sample drawn from discrete distributions with mean μ and variance σ^2, then

$P(Y = k)$	$\approx$	$P(k - \frac{1}{2} < Y < k + \frac{1}{2})$
discrete RV		approximate by continuous RV
pmf $f(y)$		by CLT for large n, Y can be approximated by $N(n\mu, n\sigma^2)$ in the sense that the pdf of the normal distribution is close to the histogram of Y
hard to calculate		easy to calculate

Chebyshev's inequality

Chebyshev's Inequality

If the random variable X has a finite mean μ and finite nonzero variance σ^2, then for every $k \geq 1$,

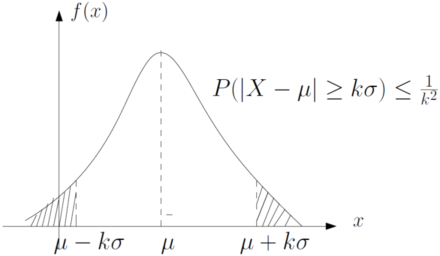
$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$

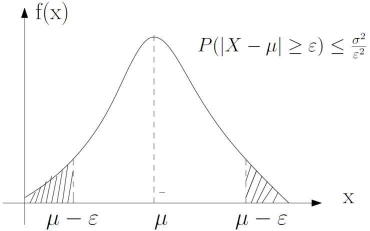
Or equivalently, if $\epsilon = k\sigma$, then

$$P(|X - \mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2}$$

Chebyshev's inequality

Graphical interpretation:





This links to the interpretation of σ^2: a measure of dispersion of the values that X can take with respect to its mean μ.

Convergence in Probability and Law of large numbers

Convergence in Probability

A sequence of RVs $Z_1, Z_2, \cdots$, is said to converge in probability to a RV Z, denoted by, $Z_n \xrightarrow{p} Z$, if for any $\varepsilon > 0$,

$$\lim_{n \to \infty} P(|Z_n - Z| \geq \varepsilon) = 0.$$

Convergence in Probability and Law of large numbers

Convergence in Probability

A sequence of RVs $Z_1, Z_2, \cdots$, is said to converge in probability to a RV Z, denoted by, $Z_n \xrightarrow{p} Z$, if for any $\varepsilon > 0$,

$$\lim_{n\to\infty} P(|Z_n - Z| \geq \varepsilon) = 0.$$

Law of Large Numbers

Let $X_1, X_2, \cdots, X_n$ be a random sample of size n drawn from a distribution with finite mean μ and finite nonzero variance, and let $\overline{X}$ be the sample mean. Then $\overline{X} \xrightarrow{p} \mu$, i.e., for any $\varepsilon > 0$,

$$\lim_{n\to\infty} P\left(\left|\overline{X} - \mu\right| \geq \varepsilon\right) = 0$$

Limiting mgf technique

Limiting mgf technique

Let $\{M_n(t)\}_{n=1}^{\infty}$ be a sequence of mgfs for t in an open interval around $t = 0$. If $\lim_{n\to\infty} M_n(t) = M(t)$, for t in the open interval around $t = 0$. Then the sequence of RVs

$$Z_n \xrightarrow{d} Z,$$

where $M_n(t)$ and $M(t)$ are mgfs of Z_n and Z, respectively.

Limiting mgf technique

Limiting mgf technique

Let $\{M_n(t)\}_{n=1}^{\infty}$ be a sequence of mgfs for t in an open interval around $t = 0$. If $\lim_{n\to\infty} M_n(t) = M(t)$, for t in the open interval around $t = 0$. Then the sequence of RVs

$$Z_n \xrightarrow{d} Z,$$

where $M_n(t)$ and $M(t)$ are mgfs of Z_n and Z, respectively.

Convergence of $b(n, p)$

1. as $n \to \infty$ with $\lambda = np$ being a constant, and let $Z_n \sim b(n, p)$ and $Z \sim \text{Poisson}(\lambda)$, then $Z_n \xrightarrow{d} Z$.
2. as $n \to \infty$ with p being a constant, and let $Z_n \sim b(n, p)$ and $Z \sim N(0, 1)$, then
$$\frac{Z_n/n - p}{\sqrt{p(1-p)/n}} \xrightarrow{d} Z$$

Final Exam

- The exam will cover Chapters 1-5, excluding Sections 3.4, 5.2
- The final exam contains 10 regular questions, which are in the similar format of the ones in the assignments. Among the 10 questions, 5 of them are taken from the assignments and created by merging a couple of questions.
- You are allowed to bring ONE sheet of A4 paper, on which you can write/draw anything you want, but all notes on the paper must be hand-written by yourself (you can print out on a blank A4 paper your own notes hand-written on ipad or the similar, but it is NOT allowed to print out other's notes).
- Tables of probabilities for distributions will be provided.
- You can bring a non-electronic dictionary, but no mathematical formula, notations, etc should be found in your dictionary.
- Please bring your student ID card to the exam for verification of identity and attendance taking.
- Those who arrive more than 30 minutes late shall NOT be permitted to take the exam.
- Calculators are allowed in the exam.

CTE

- web: `https://octe.cuhk.edu.cn/a/login`
- QR code: