

计量经济学的数理统计学基础

一、随机变量的概率分布

1 . 随机变量

随机变量是指取值具有随机性的变量。

随机变量有两种：离散随机变量和连续随机变量。

2 . 离散随机变量的概率分布

(1) 概率函数

通常用一个二维表格直观描述离散随机变量 X 的概率分布：

X	x_1	x_2	$\dots$	x_n
P	p_1	p_2	$\dots$	p_n

其中 $\sum_{i=1}^n p_i = 1$, $0 \leq p_i \leq 1$

(2) 分布函数

累计分布概率：

$$F(x) = P\{X \leq x\} = \sum_{x_i \leq x} p_i$$

3 . 连续随机变量的概率分布

(1) 概率函数

用概率密度函数 $f(x)$ 描述 ,

它满足以下性质 :

$$f(x) \geq 0, \forall x ; \int_{-\infty}^{+\infty} f(x)dx = 1 ;$$

$$P(a < X < b) = \int_a^b f(x)dx$$

(2) 分布函数

累计分布函数 $F(x) = P\{X \leq x\} = \int_{-\infty}^x f(x)dx ;$

另有 :

$$\int_a^b f(x)dx = F(b) - F(a) = P(a < x < b)$$

二、随机变量的数字特征（分布参数）

1 . 数学期望

数学期望 记为 EX 或 μ_x

对于离散变量 , $EX = \sum_{i=1}^n p_i x_i$;

对于连续变量 , $EX = \int f(x) x dx$

性质 :

$$E(C) = C$$

$$E(C \cdot X) = C \cdot E(X)$$

$$E(X_1 \pm X_2) = EX_1 \pm EX_2$$

$$E(X_1 \cdot X_2) = EX_1 \cdot EX_2 \quad X_1 \text{ 和 } X_2 \text{ 互相独立}$$

2 . 方差

方差 记为 DX 或 σ_x^2

$$DX = E(X - EX)^2 = E(X^2) - (EX)^2$$

性质 :

$$D(C) = 0 \quad ; \quad .$$

$$D(CX) = C^2DX$$

$$D(X_1 \pm X_2) = DX_1 \pm DX_2 \quad X_1 \text{和} X_2 \text{互相独立}$$

3 . 标准差 (均方差)

$$\text{标准差} \quad \sigma_X = \sqrt{DX}$$

4 . 矩

矩 $E(X_n)$ 称为变量 X 的 n 阶矩 ,
 $n = 1$ 时就是 X 的期望。

5 . 协方差

协方差用于度量两个变量的**线性**
 相关程度 , 记为 σ_{xy} 或 $\text{cov}(X, Y)$;

$$\text{cov}(X, Y) = E[(X - EX)(Y - EY)]$$

$$= E(XY) - E(X)E(Y)$$

$\sigma_{xy} > 0$ 意味着两个变量同方向变动，称之为正相关；

$\sigma_{xy} < 0$ 称之为负相关；

$\sigma_{xy} = 0$ 称之为不相关。

相关系数 $\rho = \frac{\sigma_{xy}}{\sigma_x \sigma_y} ; |p| < 1$.

如果X, Y 独立，那么 $\sigma_{xy} = 0, p = 0$,

三、样本统计量

1 . 总体和样本

所谓**总体**就是一个随机变量 X , X 的分布函数通常记为 $F(x;\theta)$, 其中 θ 就是待估计的参数。

在进行 n 次重复独立实验后，得到总体 X 的 n 个观察值 $x_1, x_2, \dots, x_n$ ，而在实验之前， $x_1, x_2, \dots, x_n$ 实际上是相互独立均与总体 X 同分布的 n 个随机变量 $X_1, X_2, \dots, X_n$ 。称 $x_1, x_2, \dots, x_n$ 为总体 X 的容量为 n 的简单随机样本，简称样本；称 $x_1, x_2, \dots, x_n$ 为样本的一个观察值，简称样本值。

2 . 常见的样本统计量

- 统计量的概念

设 $x_1, x_2, \dots, x_n$ 是来自总体 X 的一个样本，若随机变量 $x_1, x_2, \dots, x_n$ 的函数 $g(x_1, x_2, \dots, x_n)$ 中不含有任何未知

参数，则称 $g(X_1, X_2, \dots, X_n)$ 为一个统计量。

注意：统计量本身是一个随机变量；其值可由样本值计算出来。

- 最常见的统计量有：

样本均值 $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ ；

样本方差 $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ ；

样本标准差 $S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}$ ；

样本 k 阶原点矩 $M_k = \frac{1}{n} \sum_{i=1}^n X_i^k$ ；

样本 k 阶中心矩 $M'_k = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^k$ 。

假设 (X_i, Y_i) , $i = 1, \dots, n$, 是某个 X 和 Y 联合分布的样本，那么

样本协方差 $S_{XY} = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})$

样本相关系数 $\rho_{XY} = \frac{S_{XY}}{S_X S_Y}$

四、抽样分布

1 . 几个常用分布

●正态分布

定义：如果随机变量 X 的密度函数为

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}(x - \mu)^2\right]$$

则称 X 服从参数为 μ 、 σ 的**正态分布**，
通常记为 $X \sim N(\mu, \sigma^2)$ 。

令 $z = (x - \mu) / \sigma$ ，

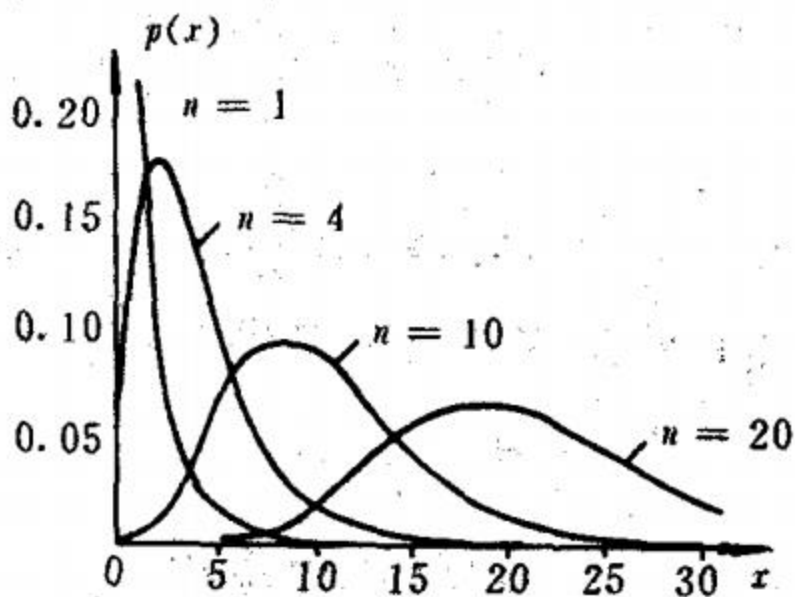
那么 z 服从**标准正态分布** $N(0, 1)$ ，

$$f(z) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}z^2\right]$$

●卡方分布

假设 n 维向量 $X \sim N(0, I_n)$ ，那么

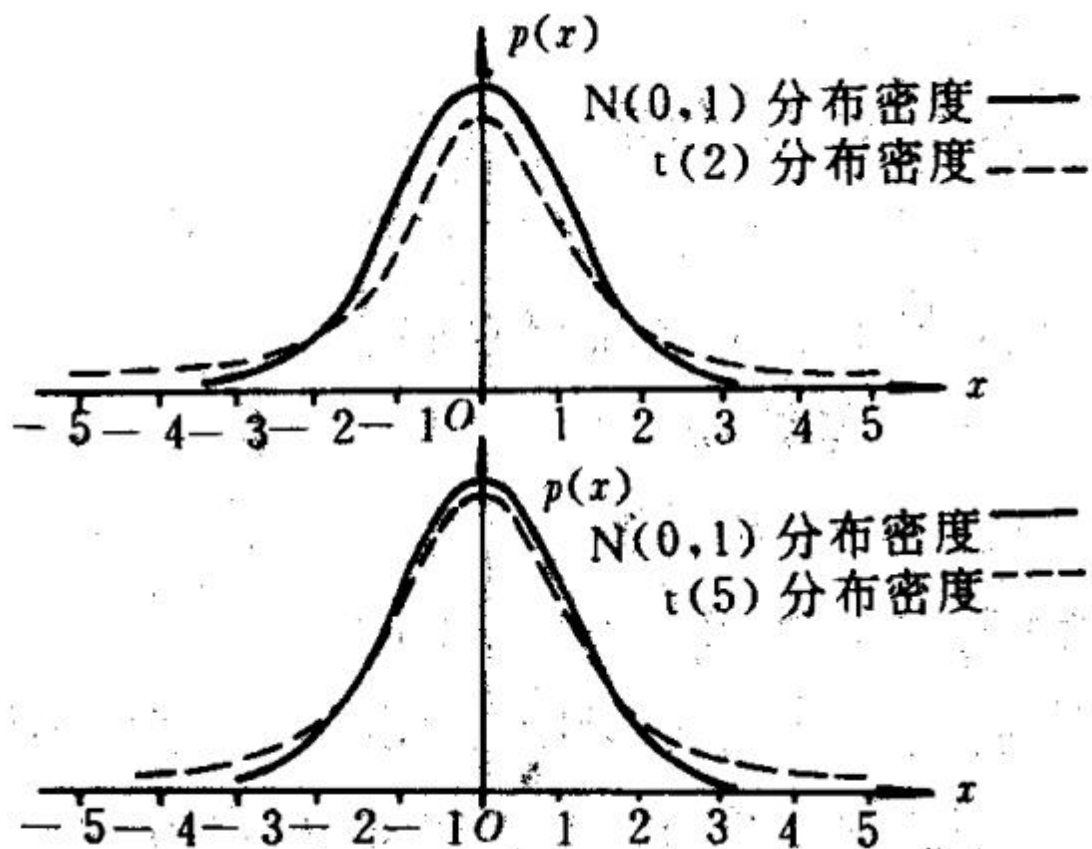
$$X'X = (X_1^2 + X_2^2 + \dots + X_n^2) \sim \chi^2(n) ;$$



●t-分布

假设两个独立的随机变量
 $Z \sim N(0,1)$, $Y \sim \chi^2(n)$, 那么

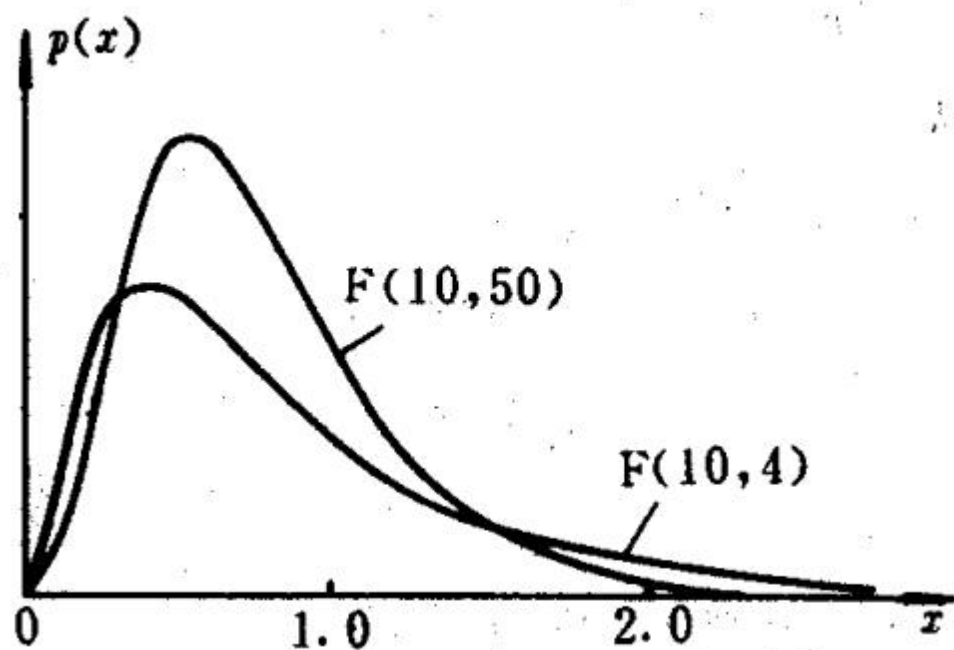
$$\frac{Z}{\sqrt{\frac{Y}{n}}} \sim t(n)$$



●F-分布

假设 $Y_1 \sim \chi^2(n_1)$ 和 $Y_2 \sim \chi^2(n_2)$ 是两个独立

的卡方分布，那么 $\frac{Y_1/n_1}{Y_2/n_2} \sim F(n_1, n_2)$



2 . 样本均值的分布

总体 $X \sim N(\mu, \sigma^2)$

样本 $X_1, X_2, \dots, X_n \sim N(\mu, \sigma^2)$

则： $\bar{X} \sim N(\mu, \sigma^2/n)$

3 . 样本方差的分布

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。
。如要下载或阅读全文，请访问：

<https://d.book118.com/337001041144010014>