第一部分 随机变量及其分布

一、离散型随机变量及其分布律

1. (0-1)分布

$$P\{X=k\}=\begin{cases}0, & k=0\\ 1, & k=1\end{cases}$$

2. 二项分布 $X\sim B(n,p)$

$$P\{X=k\}=C_n^k p^k(1-p)^{n-k}=\binom{n}{k}p^k(1-p)^{n-k}\qquad k=0,1,2,\cdots,n$$

3. 泊松分布 $X\sim \pi(\lambda)$

$$P\{X=k\}=\frac{\lambda^k e^{-\lambda}}{k!}\qquad k=0,1,2,\cdots,\quad(\lambda)$$

二、随机变量的分布函数

1. 概念

设X是一个随机变量，x是任意实数，函数$F(x)=P\{X\leq x\}\qquad -\infty<x<+\infty$

2. 基本性质

- 单调性：
  $F(x)$ 是 x 的不减函数。
  对任意实数 x_1,x_2 且 $x_1<x_2$，有：

$$F(x_2)-F(x_1)=P\{x_1<X\leqslant x_2\}\geqslant 0$$

- 有界性：
  $0\leqslant F(x)\leqslant 1$，且满足：

$$F(-\infty)=\lim_{x\to-\infty}F(x)=0,\quad F(\infty)=\lim_{x\to\infty}F(x)=1$$

  几何解释：
  当 $x\to-\infty$ 时，事件"随机点 X 落在点 x 左边"趋于不可能事件，其概率趋于0

当 $x \to \infty$ 时，事件"随机点 X 落在点 x 左边"趋于必然事件，其概率趋于1

- 右连续性：
  $F(x+0)=F(x)$，即 $F(x)$ 是右连续的（证明略）

三、连续随机变量及其概率密度

1. 定义

如果对于随机变量X的分布函数$F(x)$，存在非负可积函数$f(x)$，使对于任意实数x有

$$F(x)=\int_{-\infty}^{x} f(t)dt$$

称X为连续型随机变量，$f(x)$称为X的概率密度函数，简称概率密度

2. 性质

- $f(x) \geq 0$
- $\int_{-\infty}^{+\infty} f(x)=1$
- 对于任意实数$x_1,x_2(x_1 \leq x_2)$

$$P\{x_1<X \leq x_2\}=F(x_2)-F(x_1)=\int_{x_1}^{x_2} f(x)dx$$

- 若$f(x)$在点x处连续，则有$F'(x)=f(x)$

3. 三种重要的连续型随机变量

(1) 均匀分布$X \sim U(a,b)$

$$f(x)=\begin{cases} \dfrac{1}{b-a}, & a<x<b \\ 0, & \text{其他} \end{cases}$$

(2) 指数分布$X \sim exp(\theta)$

$$f(x)=\begin{cases} \dfrac{1}{\theta}e^{-\dfrac{x}{\theta}}, & x>0 \\ 0, & \text{其他} \end{cases}$$

(3) 正态分布(或高斯分布) $X \sim N(\mu, \sigma^2)$

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad -\infty < x < +\infty$$

- 标准正态分布$X \sim N(0,1)$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

对于任意的$X \sim N(\mu, \sigma^2)$，都可以转成$Z = \frac{X-\mu}{\sigma} \sim (0,1)$

第二部分 多维随机变量及其分布

一、二维随机变量

1. 定义

一般，设E是一个随机试验，它的样本空间是$S=\{e\}$，设$X=X(e)$和$Y=Y(e)$是定义在S上的随机变量，由它们构成的一个向量(X,Y),叫做二维随机向量或二维随机变量

2. 联合分布函数

(1) 定义

设(X,Y)是二维随机变量，对于任意实数x, y，二元函数:

$$F(x,y) = P\{X \leq x, Y \leq y\}$$

称为二维随机变量(X,Y)的分布函数，或称为随机变量X和Y的联合分布函数

(2) 性质

- 单调性:
  $F(x,y)$ 是变量 x 和 y 的不减函数:
  - 对任意固定的 y，当 $x_2 > x_1$ 时 $F(x_2,y) \geqslant F(x_1,y)$
  - 对任意固定的 x，当 $y_2 > y_1$ 时 $F(x,y_2) \geqslant F(x,y_1)$

- 有界性：
  $0 \leqslant F(x,y) \leqslant 1$，且满足极限条件：
  - $\lim\limits_{x\to-\infty} F(x,y)=0$（固定 y）
  - $\lim\limits_{y\to-\infty} F(x,y)=0$（固定 x）
  - $\lim\limits_{\substack{x\to-\infty \\ y\to-\infty}} F(x,y)=0$
  - $\lim\limits_{\substack{x\to+\infty \\ y\to+\infty}} F(x,y)=1$

  > 几何解释：当 $x\to-\infty$ 时，事件概率趋于不可能事件；当 $x,y\to+\infty$ 时，事件概率趋于必然事件

- 右连续性：
  $F(x,y)$ 关于 x 和 y 右连续：

$$F(x+0,y)=F(x,y),\quad F(x,y+0)=F(x,y)$$

- 矩形不等式：
  对任意 (x_1,y_1)，(x_2,y_2) 且 $x_1<x_2$，$y_1<y_2$，有：

$$F(x_2,y_2)-F(x_2,y_1)-F(x_1,y_2)+F(x_1,y_1)\geqslant 0$$

3. 联合概率密度

(1) 定义

存在非负可积函数 $f(x,y)$，使得对于任意实数 x 和 y，分布函数满足：

$$F(x,y)=\int_{-\infty}^{y}\int_{-\infty}^{x} f(u,v)dudv$$

满足上述条件的 (X,Y) 称为二维连续型随机变量，函数 $f(x,y)$ 称为：(X,Y) 的概率密度或X 和 Y 的联合概率密度

(2) 性质

- 非负性：
  概率密度函数 $f(x,y)$ 恒非负：

$$f(x,y)\geqslant 0$$

- 全域归一化：
  $f(x,y)$ 在整个平面上的二重积分值为1：

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = F(+\infty,+\infty) = 1$$

- 区域概率计算：
  若 G 是 xOy 平面上的任意区域，则点 (X,Y) 落在 G 内的概率为：

$$P\{(X,Y)\in G\} = \iint_G f(x,y)dxdy$$

- 微分关系：
  若 $f(x,y)$ 在点 (x,y) 处连续，则分布函数与概率密度满足微分关系：

$$\frac{\partial^2 F(x,y)}{\partial x \partial y} = f(x,y)$$

二、边缘分布

1. 边缘分布函数

- 边缘分布函数定义：
  设 (X,Y) 是二维随机变量，其联合分布函数为 $F(x,y)$。
  X 和 Y 各自的分布函数分别记作：
  - $F_X(x)$：关于 X 的边缘分布函数
  - $F_Y(y)$：关于 Y 的边缘分布函数
- 边缘分布计算公式：

$$\begin{cases} F_X(x) = P\{X \leqslant x\} = F(x,\infty) \\ F_Y(y) = P\{Y \leqslant y\} = F(\infty,y) \end{cases}$$

- 计算原理：
  - $F_X(x) = F(x,\infty) = \lim\limits_{y\to+\infty} F(x,y)$
    (固定 x，令 y 趋向正无穷)
  - $F_Y(y) = F(\infty,y) = \lim\limits_{x\to+\infty} F(x,y)$
    (固定 y，令 x 趋向正无穷)
- 离散型随机变量：
  若 (X,Y) 是离散型，可通过概率求和公式直接计算：

$$F_X(x) = \sum_{x_i \leqslant x} \sum_{y_j} p_{ij}, \quad F_Y(y) = \sum_{y_j \leqslant y} \sum_{x_i} p_{ij}$$

其中 $p_{ij} = P\{X = x_i, Y = y_j\}$

2. 边缘概率密度

- 边缘分布函数:
  对于联合分布函数 $F(x,y)$，X 的边缘分布函数定义为:

$$F_X(x) = F(x, \infty) = \lim_{y \to +\infty} F(x,y)$$

- 边缘概率密度推导:
  X 的边缘概率密度由联合概率密度 $f(x,y)$ 经积分运算得到:

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

- **Y** 的边缘概率密度:
  同理，Y 的边缘概率密度为:

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx$$

四、相互独立的随机变量

二维随机变量独立性

- 定义
  设 $F(x,y)$ 及 $F_X(x)$，$F_Y(y)$ 分别是 (X,Y) 的分布函数及边缘分布函数。
  若对所有 x,y 满足:

$$P\{X \leq x, Y \leq y\} = P\{X \leq x\}P\{Y \leq y\}$$

  等价于:

$$F(x,y) = F_X(x)F_Y(y)$$

  则称 X 和 Y 相互独立
- 连续型随机变量的独立条件:
  当 (X,Y) 是连续型随机变量，概率密度为 $f(x,y)$，边缘概率密度为 $f_X(x)$，$f_Y(y)$ 时，独立条件等价于:

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

该等式在平面上几乎处处成立

- 离散型随机变量的独立条件:
当 (X,Y) 是离散型随机变量时，对所有可能取值 (x_i,y_j) 满足:

$$P\{X=x_i,Y=y_j\}=P\{X=x_i\}P\{Y=y_j\}$$

五、两个随机变量函数的分布

1. $Z=X+Y$的分布

设 (X,Y) 是二维连续型随机变量，它具有概率密度 $f(x,y)$。则 $Z=X+Y$ 仍为连续型随机变量，其概率密度为:

$$f_{X+Y}(z)=\int_{-\infty}^{\infty}f(z-y,y)dy$$

$$f_{X+Y}(z)=\int_{-\infty}^{\infty}f(x,z-x)dx$$

- 证明:

$$\because F_z(z)=P(X+Y\leq z)=\iint_D f(x,y)dxdy,其中D=\{(x,y)|x+y\leq z\}$$
$$\therefore F_z(z)=\int_{-\infty}^{+\infty}dx\int_{-\infty}^{z-x}f(x,y)dy$$
$$\therefore f_z(z)=\frac{d}{dz}[F_z(z)]=\int_{-\infty}^{+\infty}\frac{\partial}{\partial z}[\int_{-\infty}^{z-x}f(x,y)dy]$$
$$=\int_{-\infty}^{+\infty}[f(x,z-x)\cdot(z-x)']dx$$

又若 X 和 Y 相互独立，设 (X,Y) 关于 X,Y 的边缘概率密度分别为 $f_X(x)$、$f_Y(y)$，则上两式分别化为:

$$f_{X+Y}(z)=\int_{-\infty}^{\infty}f_X(z-y)f_Y(y)dy$$

$$f_{X+Y}(z)=\int_{-\infty}^{\infty}f_X(x)f_Y(z-x)dx$$

这两个公式称为 f_X 和 f_Y 的卷积公式，记为 f_X*f_Y，即

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

2. $Z = \frac{Y}{X}$的分布、$Z = XY$的分布

- $Z = Y/X$ 的概率密度（商分布）：
  设 (X, Y) 的联合概率密度为 $f(x, y)$，则 $Z = Y/X$ 的概率密度为：

$$f_{Y/X}(z) = \int_{-\infty}^{\infty} |x| f(x, xz) dx$$

  独立情形：若 X 与 Y 独立，且边缘概率密度分别为 $f_X(x)$、$f_Y(y)$，则公式简化为：

$$f_{Y/X}(z) = \int_{-\infty}^{\infty} |x| f_X(x) f_Y(xz) dx$$

- $W = XY$ 的概率密度（积分布）：
  设 (X, Y) 的联合概率密度为 $f(x, y)$，则 $W = XY$ 的概率密度为：

$$f_{XY}(z) = \int_{-\infty}^{\infty} \frac{1}{|x|} f\left(x, \frac{z}{x}\right) dx$$

  独立情形：若 X 与 Y 独立，且边缘概率密度分别为 $f_X(x)$、$f_Y(y)$，则公式简化为：

$$f_{XY}(z) = \int_{-\infty}^{\infty} \frac{1}{|x|} f_X(x) f_Y\left(\frac{z}{x}\right) dx$$

3. $M = \max\{X, Y\}$及$N = \min\{X, Y\}$的分布

最大值与最小值函数的分布（相互独立随机变量）

- 最大值函数 $M = \max\{X, Y\}$ 的分布函数：
  由于 X 和 Y 相互独立：

$$\begin{aligned} F_{\max}(z) &= P\{M \leqslant z\} \\ &= P\{X \leqslant z, Y \leqslant z\} \\ &= P\{X \leqslant z\} \cdot P\{Y \leqslant z\} \\ &= F_X(z) \cdot F_Y(z) \end{aligned}$$

- 最小值函数 $N = \min\{X, Y\}$ 的分布函数：
  由于 X 和 Y 相互独立：

$$
\begin{aligned}
F_{\min}(z) &= P\{N \leqslant z\} \\
&= 1 - P\{N > z\} \\
&= 1 - P\{X > z, Y > z\} \\
&= 1 - [1 - F_X(z)] \cdot [1 - F_Y(z)]
\end{aligned}
$$

推广至 n 个独立随机变量

- n 维最大值函数 $M = \max\{X_1, \ldots, X_n\}$ 的分布函数：
  设 $X_1, \ldots, X_n$ 相互独立，分布函数为 $F_{X_i}(x_i)$：

$$
F_{\max}(z) = \prod_{i=1}^{n} F_{X_i}(z) = F_{X_1}(z) \cdot F_{X_2}(z) \cdots F_{X_n}(z)
$$

- n 维最小值函数 $N = \min\{X_1, \ldots, X_n\}$ 的分布函数：
  设 $X_1, \ldots, X_n$ 相互独立，分布函数为 $F_{X_i}(x_i)$：

$$
F_{\min}(z) = 1 - \prod_{i=1}^{n} [1 - F_{X_i}(z)]
$$

第三部分 随机变量的数字特征

一、数学期望

1. 定义

数学期望定义（离散型）

- 分布律条件：
  设离散型随机变量 X 的分布律为：

$$
P\{X = x_k\} = p_k, \quad k = 1, 2, \cdots
$$

- 数学期望公式：
  若级数 $\sum_{k=1}^{\infty} x_k p_k$ 绝对收敛，则定义 X 的数学期望为：

$$
E(X) = \sum_{k=1}^{\infty} x_k p_k
$$

  关键要求：级数必须绝对收敛（$\sum |x_k| p_k < \infty$）

数学期望定义（连续型）

- 概率密度条件:
  设连续型随机变量 X 的概率密度为 $f(x)$
- 数学期望公式:
  若积分 $\int_{-\infty}^{\infty} xf(x)dx$ 绝对收敛，则定义 X 的数学期望为:

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx$$

关键要求：积分必须绝对收敛（$\int |x|f(x)dx < \infty$）

2. 随机变量函数的期望定理

- 定理核心表述:
  设 $Y = g(X)$ 是随机变量 X 的函数（g 为连续函数）
- 离散型情形:
  若 X 是离散型随机变量，分布律为 $P\{X = x_k\} = p_k$ $(k = 1, 2, \cdots)$，且级数满足绝对收敛条件:

$$\sum_{k=1}^{\infty} |g(x_k)|p_k < \infty$$

则 Y 的数学期望为:

$$E(Y) = E[g(X)] = \sum_{k=1}^{\infty} g(x_k)p_k$$

- 连续型情形:
  若 X 是连续型随机变量，概率密度为 $f(x)$，且积分满足绝对收敛条件:

$$\int_{-\infty}^{\infty} |g(x)|f(x)dx < \infty$$

则 Y 的数学期望为:

$$E(Y) = E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)dx$$

3. 性质

(1) 常数的期望

- 性质：
  若 C 为常数，则

$$E(C) = C$$

(2) 常数倍性质

- 性质：
  设 X 为随机变量，C 为常数，则

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

(3) 可加性（核心性质）

- 性质：
  设 X, Y 为随机变量，则

$$E(X + Y) = E(X) + E(Y)$$

  可推广至 n 个随机变量：

$$E\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} E(X_k)$$

(4) 独立变量的乘积期望

- 性质：
  若 X, Y 相互独立，则

$$E(XY) = E(X) \cdot E(Y)$$

二、方差

1. 方差定义与性质

- 定义式：

$$D(X) = \mathrm{Var}(X) = E\left\{[X - E(X)]^2\right\}$$

- $E(X)$：随机变量 X 的数学期望（均值）
- $[X - E(X)]^2$：偏差的平方
- 方差 $D(X)$ 是偏差平方的期望值

- 标准差：

$$\sigma(X) = \sqrt{D(X)}$$

- 物理意义：度量波动大小的直接指标（与原始量纲一致）

2. 方差计算体系

随机变量类型	计算公式	关键要素说明
离散型	$D(X) = \sum_{k=1}^{\infty} [x_k - E(X)]^2 p_k$	$p_k = P\{X = x_k\}$（分布律）
连续型	$D(X) = \int_{-\infty}^{\infty} [x - E(X)]^2 f(x) dx$	$f(x)$（概率密度函数）

$$D(X) = E(X^2) - [E(X)]^2$$

2. 性质

- 常数方差：
设 C 是常数，则：

$$D(C) = 0$$

- 常数倍与平移性质：
设 X 是随机变量，C 是常数：

$$D(CX) = C^2 D(X)$$
$$D(X + C) = D(X)$$

- 随机变量和的方差：
设 X, Y 是两个随机变量：

$$D(X + Y) = D(X) + D(Y) + 2E\{[X - E(X)][Y - E(Y)]\}$$

独立情形：若 X, Y 相互独立，则简化为：

$$D(X+Y)=D(X)+D(Y)$$

推广：此性质可推广到任意有限多个相互独立的随机变量之和

- 方差为零的充要条件：

$$D(X)=0 \quad \Leftrightarrow \quad P\{X=E(X)\}=1$$

即 X 以概率 1 取常数 $E(X)$

3. 切比雪夫不等式

设随机变量 X 满足：

- 数学期望 $E(X)=\mu$
- 方差 $D(X)=\sigma^2$

则对 任意正数 ε ，有：

$$P\{|X-\mu| \geqslant \varepsilon\} \leqslant \frac{\sigma^2}{\varepsilon^2}$$

证明核心步骤

$$\begin{aligned}
P\{|X-\mu| \geqslant \varepsilon\} &= \int_{|x-\mu| \geqslant \varepsilon} f(x)dx \quad (\text{连续型情形}) \\
&\leqslant \int_{|x-\mu| \geqslant \varepsilon} \left(\frac{x-\mu}{\varepsilon}\right)^2 f(x)dx \quad \left(\text{因} \left(\frac{|x-\mu|}{\varepsilon}\right)^2 \geqslant 1\right) \\
&\leqslant \frac{1}{\varepsilon^2} \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx \\
&= \frac{\sigma^2}{\varepsilon^2}
\end{aligned}$$

三、协方差及相关系数

1. 定义

- 协方差定义：
  设 (X,Y) 是二维随机变量，定义 X 与 Y 的协方差为：

$$\mathrm{Cov}(X,Y)=E\{[X-E(X)][Y-E(Y)]\}$$

- 相关系数定义：
  X 与 Y 的相关系数为：

$$\rho_{XY} = \frac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$$

- 协方差基本性质：

$$\begin{cases} \mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X) \\ \mathrm{Cov}(X,X) = D(X) \end{cases}$$

- 方差与协方差关系：
  对任意两个随机变量 X 和 Y，有：

$$D(X+Y) = D(X) + D(Y) + 2\,\mathrm{Cov}(X,Y)$$

- 协方差计算式：
  协方差可展开为：

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

2. 性质

协方差的核心性质

标量乘法不变性

设 a，b 为常数，则：

$$\mathrm{Cov}(aX,\ bY) = ab\cdot\mathrm{Cov}(X,\ Y)$$

加法分配律

对任意随机变量 X_1，X_2，Y：

$$\mathrm{Cov}(X_1+X_2,\ Y) = \mathrm{Cov}(X_1,\ Y) + \mathrm{Cov}(X_2,\ Y)$$

3. 定理

$$|\rho_{XY}| \le 1 \quad \text{当且仅当 } \exists a,b\in\mathbb{R}, P\{Y=a+bX\}=1\text{时，等号成立}$$

四、矩、协方差矩阵

1. 矩的定义体系

设 (X,Y) 是二维随机变量：

- k阶原点矩（X 的矩）：
  若存在，定义为：

$$E(X^k), \quad k=1,2,\cdots$$

- k阶中心矩（X 的偏差矩）：
  若存在，定义为：

$$E\left\{[X-E(X)]^k\right\}, \quad k=2,3,\cdots$$

- $k+l$阶混合矩（X 和 Y 的联合矩）：
  若存在，定义为：

$$E(X^kY^l), \quad k,l=1,2,\cdots$$

- $k+l$阶混合中心矩（联合偏差矩）：
  若存在，定义为：

$$E\left\{[X-E(X)]^k[Y-E(Y)]^l\right\}, \quad k,l=1,2,\cdots$$

矩的特例关系

- $E(X)$ 是 X 的一阶原点矩
- $D(X)$ 是 X 的二阶中心矩
- $\mathrm{Cov}(X,Y)$ 是 X 和 Y 的二阶混合中心矩

第四部分 大数定理及中心极限定理

一、大数定理

1. 弱大数定律（辛钦大数定律）

- 表述一：
  设$X_1,X_2,\cdots$是相互独立，服从同一分布的随即变脸序列，且具有数学期望 $E(X_k)=\mu\ (k=1,2,\cdots)$.作前n个变量的算术平均$\dfrac{1}{n}\sum_{k=1}^{n}X_k$,则对于任意$\varepsilon>0$,有

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \mu\right| < \varepsilon\right\} = 1$$

- 表述二：
  设随机变量$X_1, X_2, \cdots, X_n, \cdots$相互独立，服从同一分布且具有数学期望 $E(X_k) = \mu\ (k = 1, 2, \cdots)$，则序列$\overline{X} = \frac{1}{n}\sum_{k=1}^{n}$依概率收敛于$\mu$，即$\overline{X} \xrightarrow{P} \mu$
  其中：$X \xrightarrow{P} \mu$表示X按概率收敛于μ

2. 伯努利大数定律

设f_A是n次独立重复试验中事件A发生的次数，p是事件A在每次试验中发生的概率，则对

$$\lim_{n\to\infty} P\left\{\left|\frac{f_A}{n} - p\right| < \varepsilon\right\} = 1 \quad 或 \quad \lim_{n\to\infty} P\left\{\left|\frac{f_A}{n} - p\right| \geq \varepsilon\right\} = 0$$

二、中心极限定理

独立同分布的中心极限定理

设随机变量$X_1, X_2, \cdots, X_n, \cdots$相互独立，服从同一分布，且具有数学期望和方差：$E(X_k) = \mu, D(X_k) = \sigma^2 > 0\ (k = 1, 2, \cdots)$，则随机变量之和$\sum_{k=1}^{n} X_k$的标准化变量

$$Y_n = \frac{\sum_{k=1}^{n} X_k - E(\sum_{k=1}^{n} X_k)}{\sqrt{D(\sum_{k=1}^{n} X_k)}} = \frac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma}$$

的分布函数$F_n(x)$对于任意x满足

$$\lim_{n\to\infty} F_n(x) = \lim_{n\to\infty} P\left\{\frac{\sum_{k=1}^{n} X_k - E(\sum_{k=1}^{n} X_k)}{\sqrt{D(\sum_{k=1}^{n} X_k)}}\right\} = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \Phi(x)$$

李雅普诺夫($\mathcal{L}yapunov$) 定理

设随机变量$X_1, X_2, \cdots, X_n, \cdots$相互独立，它们具有数学期望和方差

$$E(X_k) = \mu_k \text{ , } D(X_k) = \sigma_k^2 > 0 \text{ , } k = 1, 2, \cdots$$

$$\text{记} \quad B_n^2 = \sum_{k=1}^{n} \sigma_k^2,$$

若存在正数δ，使得当$n \to \infty$时，

$$\frac{1}{B_n^{2+\delta}} \sum_{k=1}^{n} E\left\{|X_k - \mu_k|^{2+\delta}\right\} \to 0,$$

则随机变量之和$\sum_{k=1}^{n} X_k$的标注化变量

$$Z_n = \frac{\sum_{k=1}^{n} X_k - E(\sum_{k=1}^{n} X_k)}{\sqrt{D(\sum_{k=1}^{n} X_k)}} = \frac{\sum_{k=1}^{n} X_k - \sum_{k=1}^{n} \mu_k}{B_n}$$

的分布函数$F_n(x)$对于任意x，满足

$$\lim_{n \to \infty} F_n(x) = \lim_{n \to \infty} P\left\{\frac{\sum_{k=1}^{n} X_k - \sum_{k=1}^{n} \mu_k}{B_n} \le x\right\} = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \varPhi(x)$$

棣莫弗 - 拉普拉斯($\mathcal{D}e\ \mathcal{M}oivre - \mathcal{L}aplace$) 定理

设随机变量$\eta_n \ (n = 1, 2, \cdots)$服从参数为$n, p \ (0 < p < 1)$的二项分布，则对于任意$x$，有

$$\lim_{n \to \infty} P\left\{\frac{\eta_n - np}{\sqrt{np(1-p)}} \le x\right\} = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \varPhi(x)$$

第五部分 抽样分布

一、基础定义

- 样本均值（期望） $E(X)=\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$
- 样本方差 $D(X)=S^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2=\frac{1}{n-1}\left(\sum_{i=1}^{n}X_i^2-n\overline{X}\right)$
- 样本标准差 $S=\sqrt{S^2}$
- 样本k阶（原点）矩 $A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k$, $k=1,2,\cdots$
- 样本k阶中心矩 $B_k=\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^k$, $k=2,3,\cdots$
- 经验分布函数
  设$x_1,x_2,\cdots,x_n$是来自分布函数为$F(x)$的总体X的样本观察值.X的经验分布函数，记为$F_n(x)$，定义为样本观察值$x_1,x_2,\cdots,x_n$中小于或等于指定值x所占的比率，即

$$F_n(x)=\frac{\#(x_i\leq x)}{n}\qquad -\infty<x<+\infty$$

  其中$\#(x_i\leq x)$表示$x_1,x_2,\cdots,x_n$中小于等于x的个数。
  - 具有分布函数的条件
    1. $F_n(x)$是x的不减函数
    2. $0\leq F_n(x)\leq 1$，且$F(\infty)=0,F(+\infty)=1$
    3. $F(x)$是一个右连续函数
  - 格里汶科定理
    设$X_1,X_2,\cdots,X_n$的函数是来自以$F_(x)$为分布函数的总体X的样本，$F(x)$是经验分布函数，则有

$$P\left\{\lim_{n\to\infty}\sup_{-\infty\to+\infty}|F_n(x)-F(x)|=0\right\}=1$$

二、抽样分布

1. χ^2分布

- 定义
  设$X_1,X_2,\cdots,X_n$是来自总体$N(0,1)$的样本，则称统计量

$$\chi = X_1^2 + X_2^2 + \cdots + X_n^2$$

服从自由度为n的χ^2分布，记为$\chi^2 \sim \chi^2(n)$

- χ^2分布的概率密度

$$f(y) = \begin{cases} \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} y^{\frac{n}{2}-1} e^{-\frac{y}{2}}, & y > 0 \\ 0, & 其他 \end{cases}$$

$$\chi^2 = \sum_{i=1}^{n} X_i^2 \sim \Gamma(\frac{n}{2}, 2)$$

- χ^2分布的可加性
  设$\chi_1^2 \sim \chi^2(n_1), \chi_2^2 \sim \chi^2(n_2)$，并且$\chi_1^2, \chi_2^2$相互独立，则有

$$\chi_1^2 + \chi_2^2 \sim \chi^2(n_1 + n_2)$$

- χ^2分布的期望和方差
  若$\chi^2 \sim \chi^2(n)$，则有

$$E(\chi^2) = n \qquad D(\chi^2) = 2n$$

- χ^2分布的上分位数
  对于给定的正数α，$0 < \alpha < 1$，满足条件

$$P\{\chi^2 > \chi_\alpha^2(n)\} = \int_{\chi_\alpha^2(n)}^{+\infty} f(y)dy = \alpha$$

- 近似计算

$$\chi_\alpha^2(n) \approx \frac{1}{2}(z_\alpha + \sqrt{2n-1})^2$$

2. t分布（学生氏$(Student)$分布）

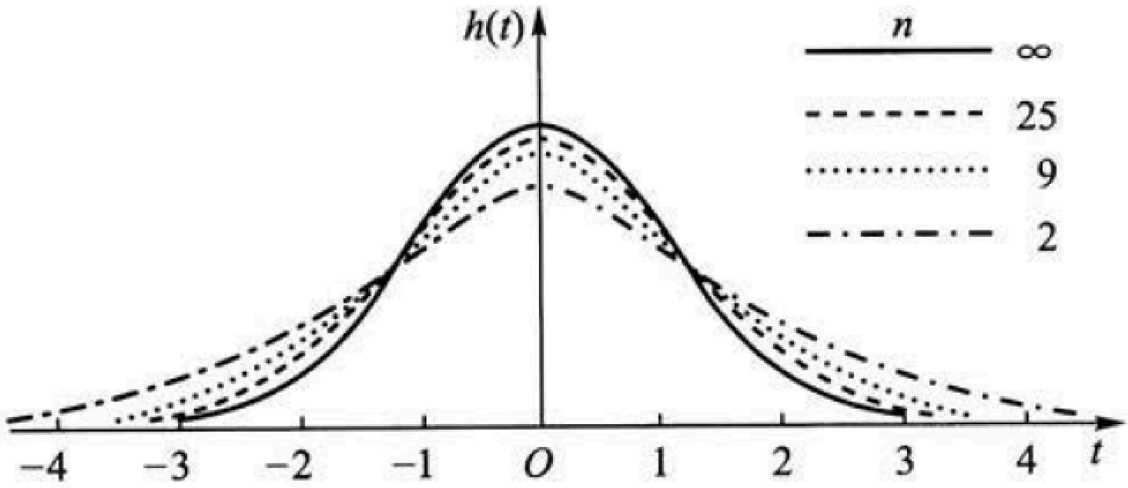
- 定义
  设$X \sim N(0,1), Y \sim \chi^2(n)$，且$X, Y$相互独立，则称变量

$$t = \frac{X}{\sqrt{\frac{Y}{N}}}$$

  服从自由度为n的t分布，记为$t \sim t(n)$
- 概率密度函数

$$h(t)=\frac{\Gamma(\frac{n+1}{2})}{\sqrt{\pi n}\,\Gamma(\frac{n}{2})}(1+\frac{t^2}{n})^{-\frac{n+1}{2}}\qquad -\infty<t<+\infty$$

$$\lim_{n\to\infty}h(t)=\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}$$



- t分布的上分位数
  对于给定的$\alpha, 0<\alpha<1$，满足条件

$$P\left\{t>t_\alpha(n)\right\}=\int_{t_\alpha(n)}^{+\infty}h(t)dt=\alpha$$

- 对称性

$$t_{1-n}=-t_\alpha(n)$$

- 近似计算

$$t_\alpha(n)\approx z_n\qquad n>45$$

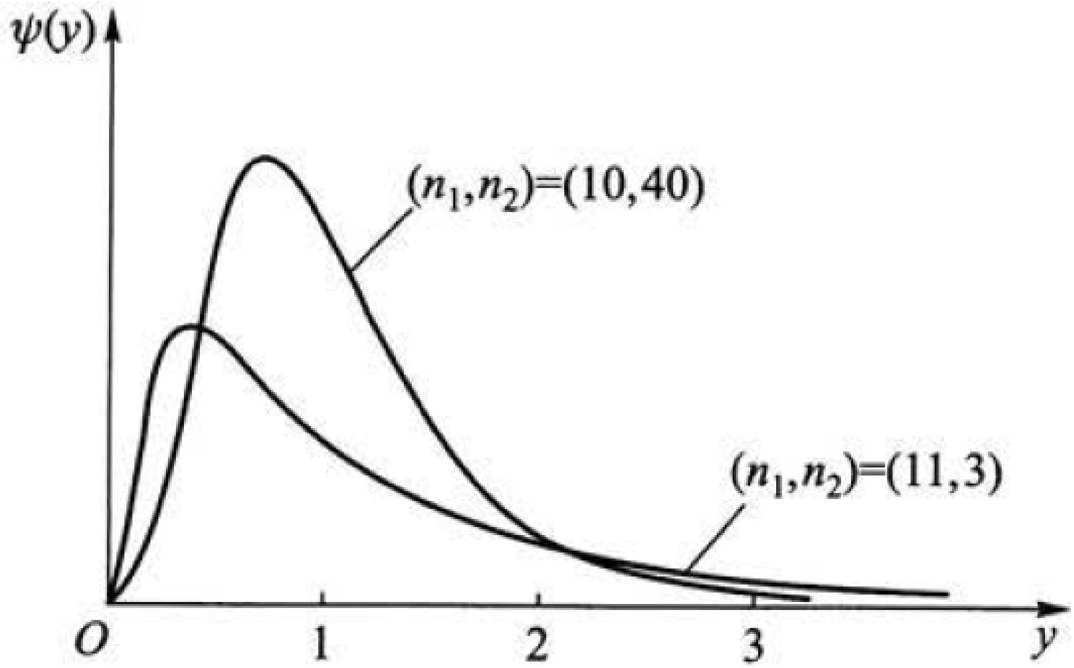
3. F分布

- 定义
  设$U\sim\chi^2(n_1),V\sim\chi^2(n_2)$，且$U,V$相互独立，则称随机变量

$$F=\frac{\frac{U}{n_1}}{\frac{V}{n_2}}$$

  服从自由度为(n_1,n_2)的F分布，记为$F\sim F(n_1,n_2)$.
- 分布概率

$$\psi(y)=\begin{cases}\dfrac{\Gamma(\frac{n_1+n_2}{2})(\frac{n_1}{n_2})^{\frac{n_1}{2}}y^{\frac{n_1}{2}-1}}{\Gamma(\frac{n_1}{2})\Gamma(\frac{n_2}{2})(1+\frac{n_1 y}{n_2})^{\frac{n_1+n_2}{2}}}, & y>0\\ 0, & \text{其他}\end{cases}$$



- 特性

$$若F\sim F(n_1,n_2),则\frac{1}{F}\sim F(n_2,n_1)$$

- 上分位数
  对于给定的$\alpha,0<\alpha<1$，满足条件

$$P\{F>F_\alpha(n_1,n_2)\}=\int_{F_\alpha(n_1,n_2)}^{+\infty}\psi(y)dy=\alpha$$

- 对称性

$$F_{1-\alpha}(n_1,n_2)=\frac{1}{F_\alpha(n_2,n_1)}$$

4. 正态总体的样本均值于样本方差的分布

(1) 单总体样本均值和方差性质

设总体 X 均值为 μ，方差为 σ^2（分布不限），样本 $X_1,\cdots,X_n$，$\overline{X}$ 和 S^2 为样本均值和方差：

$$E(\overline{X}) = \mu$$
$$D(\overline{X}) = \frac{\sigma^2}{n}$$
$$E(S^2) = \sigma^2 \quad \text{(无偏性)}$$

$E(S^2) = \sigma^2$ 的证明:

$$\begin{aligned} E(S^2) &= E\left[\frac{1}{n-1}\left(\sum_{i=1}^{n} X_i^2 - n\overline{X}^2\right)\right] \\ &= \frac{1}{n-1}\left[\sum_{i=1}^{n} E(X_i^2) - nE(\overline{X}^2)\right] \\ &= \frac{1}{n-1}\left[\sum_{i=1}^{n}(\sigma^2+\mu^2) - n\left(\frac{\sigma^2}{n}+\mu^2\right)\right] \\ &= \sigma^2 \end{aligned}$$

---

(2) 正态总体分布定理

设 $X \sim N(\mu, \sigma^2)$，则:

- 定理2 $\overline{X} \sim N\left(\mu, \dfrac{\sigma^2}{n}\right)$
- 定理3 $\begin{cases} \dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1) \\ \overline{X} \text{ 与 } S^2 \text{ 相互独立} \end{cases}$
- 定理4 $\dfrac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$

定理4证明思路:

- 由定理2: $\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$
- 由定理3: $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$
- 由t分布定义:

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

(3) 双正态总体分布定理

设两独立样本：

- $X_i \sim N(\mu_1, \sigma_1^2)$（容量 n_1）
- $Y_j \sim N(\mu_2, \sigma_2^2)$（容量 n_2）

则：

$$1^\circ \ \frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}} \sim F(n_1 - 1, n_2 - 1)$$

$$2^\circ \text{ 当 } \sigma_1^2 = \sigma_2^2 = \sigma^2 \text{ 时：}$$
$$\frac{(\overline{X} - \overline{Y}) - (\mu_1 - \mu_2)}{S_W\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2)$$

其中合并方差：

$$S_W^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}, \quad S_W = \sqrt{S_W^2}$$

(4) 应用场景速查表

检验类型	使用统计量	分布	条件
单样本均值检验	$\frac{\overline{X} - \mu}{\frac{S}{\sqrt{n}}}$	$t(n-1)$	正态总体
两样本方差齐性检验	$\frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}}$	$F(n_1 - 1, n_2 - 1)$	独立正态样本
两样本均值差检验	$\frac{(\overline{X} - \overline{Y}) - \delta}{S_W\sqrt{1/n_1 + 1/n_2}}$	$t(\nu)$	方差相等 ($\sigma_1^2 = \sigma_2^2$)

> 注:
> $\nu = n_1 + n_2 - 2$
> $\delta = \mu_1 - \mu_2$

第六部分 参数估计

一、点估计

1. 矩估计法

(1) 基本概念
设总体 X 为连续型随机变量（概率密度 $f(x;\theta_1,\cdots,\theta_k)$）或离散型随机变量（分布律 $P\{X=x\}=p(x;\theta_1,\cdots,\theta_k)$），其中 $\theta_1,\cdots,\theta_k$ 为待估参数。$X_1,\cdots,X_n$ 为样本。

(2) 矩的定义

- 连续型总体:

$$\mu_l = \int_{-\infty}^{\infty} x^l f(x;\theta_1,\cdots,\theta_k)dx$$

- 离散型总体:

$$\mu_l = \sum_{x\in R_X} x^l p(x;\theta_1,\cdots,\theta_k)$$

（R_X 为 X 可能取值范围）

(3) 估计原理

- 样本矩:

$$A_l = \frac{1}{n}\sum_{i=1}^{n} X_i^l$$

- 建立方程组:

$$
\begin{cases}
\mu_1 = \mu_1(\theta_1, \cdots, \theta_k) \\
\vdots \\
\mu_k = \mu_k(\theta_1, \cdots, \theta_k)
\end{cases}
$$

- 解方程组得矩估计量：

$$
\hat{\theta}_i = \theta_i(A_1, \cdots, A_k) \quad (i = 1, \cdots, k)
$$

2. 最大似然估计法

(1) 似然函数定义

- 离散型总体：

$$
L(\theta) = \prod_{i=1}^{n} p(x_i; \theta)
$$

- 连续型总体：

$$
L(\theta) = \prod_{i=1}^{n} f(x_i; \theta)
$$

  ($x_1, \cdots, x_n$ 为样本观测值)

(2) 估计原理
寻找 $\hat{\theta}$ 使似然函数最大化：

$$
L(x_1, \cdots, x_n; \hat{\theta}) = \max_{\theta \in \Theta} L(x_1, \cdots, x_n; \theta)
$$

(3) 求解方法

- 似然方程：

$$
\frac{d}{d\theta} L(\theta) = 0
$$

- 对数似然方程（优先使用）：

$$
\frac{d}{d\theta} \ln L(\theta) = 0
$$

- 估计量：$\hat{\theta}(X_1, \cdots, X_n)$，估计值：$\hat{\theta}(x_1, \cdots, x_n)$