进一步根据正态分布的性质有 $\bar{X}_1 - \bar{X}_2 \sim \mathcal{N}(0, 1/2)$, 于是可得

$$\Pr(|\bar{X}_1 - \bar{X}_2| > 0.3) = 2 - 2\Phi(0.3/\sqrt{1/2}).$$

假设 $X_1, X_2, \cdots, X_n$ 是来自总体 $X$ 的一个样本, 定义 **最小次序统计量** 和 **最大次序统计量** 分别为:

$$X_{(1)} = \min\{X_1, X_2, \cdots, X_n\} \qquad \text{和} \qquad X_{(n)} = \max\{X_1, X_2, \cdots, X_n\},$$

以及定义 **样本极差** 为

$$R_n = X_{(n)} - X_{(1)}.$$

设总体 $X$ 的分布函数为 $F(x)$, 则有

$$F_{X_{(1)}}(x) = \Pr(X_{(1)} \leqslant x) = 1 - \Pr(X_{(1)} > x) = 1 - (1 - F(x))^n, \qquad F_{X_{(n)}}(x) = F^n(x).$$

**定理 9.1** 设总体 $X$ 的密度函数为 $f(x)$, 分布函数为 $F(x)$, $X_1, X_2, \cdots, X_n$是来自总体 $X$ 的一个样本, 则第 $k$ 次序统计量 $X_{(k)}$ 的分布函数和密度函数分别为

$$\begin{aligned}
F_k(x) &= \sum_{r=k}^{n} \binom{n}{r} [F(x)]^r [1 - F(x)]^{n-r} \\
f_k(x) &= \frac{n!}{(k-1)!(n-k)!} [F(x)]^{k-1} [1 - F(x)]^{n-k} f(x).
\end{aligned}$$

**证明** 根据题意有第 $k$ 次序统计量 $X_{(k)}$ 的分布函数为

$$\begin{aligned}
F_k(x) &= \Pr[X_{(k)} \leqslant x] = \Pr[X_1, X_2, \cdots, X_n \text{ 中至少有 } k \text{ 个随机变量 } \leqslant x] \\
&= \sum_{r=k}^{n} \Pr[X_1, X_2, \cdots, X_n \text{ 中恰有 } r \text{ 个随机变量 } \leqslant x,\ n-r \text{ 个随机变量 } > x] \\
&= \sum_{r=k}^{n} \binom{n}{r} [F(x)]^r [1 - F(x)]^{n-r}.
\end{aligned}$$

利用恒等式

$$\sum_{r=k}^{n} \binom{n}{r} p^r (1-p)^{n-r} = \frac{n!}{(k-1)!(n-k)!} \int_0^p t^{k-1} (1-t)^{n-k} dt \quad (r \in [n],\ p \in [0, 1])$$

由此可知

$$F_k(x) = \frac{n!}{(k-1)!(n-k)!} \int_0^{F(x)} t^{k-1} (1-t)^{n-k} dt,$$

根据积分函数求导完成证明.

## 9.3 Beta 分布、Γ 分布、Dirichlet 分布

首先介绍两积分函数.

**定义 9.3 (Beta-函数)** 对任意给定 $\alpha_1 > 0$ 和 $\alpha_2 > 0$, 定义 Beta 函数为

$$\mathrm{Beta}(\alpha_1, \alpha_2) = \int_0^1 x^{\alpha_1 - 1}(1-x)^{\alpha_2 - 1} dx,$$

有些书简记为 $B(\alpha_1, \alpha_2)$, 被称为第一类欧拉积分函数.

根据数学分析可知 $\mathrm{Beta}(\alpha_1, \alpha_2)$ 在定义域 $(0, +\infty) \times (0, +\infty)$ 连续. 利用变量替换 $t = 1 - x$,根据定义有

$$\begin{aligned}
\mathrm{Beta}(\alpha_1, \alpha_2) &= \int_0^1 t^{\alpha_1 - 1}(1-t)^{\alpha_2 - 1} dt = \int_1^0 (1-x)^{\alpha_1 - 1} x^{\alpha_2 - 1} d(1-x) \\
&= \int_0^1 x^{\alpha_2 - 1}(1-x)^{\alpha_1 - 1} dx = \mathrm{Beta}(\alpha_2, \alpha_1),
\end{aligned}$$

由此可知 Beta 函数的对称性: $\mathrm{Beta}(\alpha_1, \alpha_2) = \mathrm{Beta}(\alpha_2, \alpha_1)$.

**定义 9.4 (Γ-函数)** 对任意给定 $\alpha > 0$, 定义 Γ-函数为

$$\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha - 1} e^{-x} dx,$$

又被称为第二类欧拉积分函数.

**性质 9.1** 对 Γ-函数, 有 $\Gamma(1) = 1$ 和 $\Gamma(1/2) = \sqrt{\pi}$, 以及对 $\alpha > 1$ 有 $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$.

**证明** 根据定义有

$$\Gamma(1) = \int_0^{+\infty} e^{-x} dx = 1.$$

利用变量替换 $x = t^{1/2}$ 有

$$\Gamma(1/2) = \int_0^{+\infty} t^{-\frac{1}{2}} e^{-t} dt = \int_0^{+\infty} x^{-1} e^{-x^2} dx^2 = 2\int_0^{+\infty} e^{-x^2} dx = \int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}.$$

进一步有

$$\Gamma(\alpha) = -\int_0^{\infty} x^{\alpha-1} de^{-x} = -[x^{\alpha-1}e^{-x}]_0^{+\infty} + (\alpha - 1)\int_0^{+\infty} x^{\alpha-2} e^{-x} dx = (\alpha - 1)\Gamma(\alpha - 1)$$

对任意正整数 $n$, 根据上面的性质有

$$\Gamma(n) = (n-1)!$$

关于 Beta 函数和 $\Gamma$-函数, 有如下关系:

**定理 9.2** 对任意给定 $\alpha_1 > 0$ 和 $\alpha_2 > 0$, 有

$$\mathrm{Beta}(\alpha_1, \alpha_2) = \frac{\Gamma(\alpha_1)\Gamma(\alpha_2)}{\Gamma(\alpha_1 + \alpha_2)}.$$

**证明** 根据 $\Gamma$-函数的定义有

$$\Gamma(\alpha_1)\Gamma(\alpha_2) = \int_0^{+\infty} t^{\alpha_1 - 1} e^{-t} dt \int_0^{+\infty} s^{\alpha_2 - 1} e^{-s} ds = \int_0^{+\infty}\int_0^{+\infty} e^{-(t+s)} t^{\alpha_1 - 1} s^{\alpha_2 - 1} dt ds.$$

引入变量替换 $x = t + s$ 和 $y = t/(t + s)$, 反解可得 $t = xy$ 和 $s = x - xy$, 计算雅可比行列式有

$$\begin{vmatrix} \frac{\partial t}{\partial x} & \frac{\partial t}{\partial y} \\ \frac{\partial s}{\partial x} & \frac{\partial s}{\partial y} \end{vmatrix} = \begin{vmatrix} y & x \\ 1 - y & -x \end{vmatrix} = -x.$$

同时有 $x \in (0, +\infty)$ 和 $y \in (0, 1)$ 成立, 由此可得

$$\begin{aligned}
\Gamma(\alpha_1)\Gamma(\alpha_2) &= \int_0^1 \int_0^{+\infty} e^{-x} x^{\alpha_1 - 1} y^{\alpha_1 - 1} x^{\alpha_2 - 1} (1 - y)^{\alpha_2 - 1} |x| dx dy \\
&= \int_0^1 \int_0^{+\infty} e^{-x} x^{\alpha_1 + \alpha_2 - 1} y^{\alpha_1 - 1} (1 - y)^{\alpha_2 - 1} dx dy \\
&= \int_0^{+\infty} e^{-x} x^{\alpha_1 + \alpha_2 - 1} dx \int_0^1 y^{\alpha_1 - 1} (1 - y)^{\alpha_2 - 1} dy \\
&= \Gamma(\alpha_1 + \alpha_2)\mathrm{Beta}(\alpha_1, \alpha_2)
\end{aligned}$$

定理得证.

根据上述定理可知

**推论 9.1** 对任意 $\alpha_1 > 1$ 和 $\alpha_2 > 0$, 有

$$Beta(\alpha_1, \alpha_2) = \frac{\alpha_1 - 1}{\alpha_1 + \alpha_2 - 1} Beta(\alpha_1 - 1, \alpha_2).$$

**证明** 根据前面的定理有

$$\mathrm{Beta}(\alpha_1, \alpha_2) = \frac{\Gamma(\alpha_1)\Gamma(\alpha_2)}{\Gamma(\alpha_1 + \alpha_2)} = \frac{(\alpha_1 - 1)\Gamma(\alpha_1 - 1)\Gamma(\alpha_2)}{(\alpha_1 + \alpha_2 - 1)\Gamma(\alpha_1 + \alpha_2 - 1)} = \frac{\alpha_1 - 1}{\alpha_1 + \alpha_2 - 1}\mathrm{Beta}(\alpha_1 - 1, \alpha_2).$$

**定义 9.5** 对任意 $\alpha_1, \alpha_2, \cdots, \alpha_k > 0$, 定义多维 Beta 函数为

$$\mathrm{Beta}(\alpha_1, \alpha_2, \cdots, \alpha_k) = \frac{\Gamma(\alpha_1)\Gamma(\alpha_2)\cdots\Gamma(\alpha_k)}{\Gamma(\alpha_1 + \alpha_2 + \cdots + \alpha_k)}.$$

下面介绍三种分布:

**定义 9.6 (Beta 分布)** 给定 $\alpha_1 > 0$ 和 $\alpha_2 > 0$, 若随机变量 $X$ 的概率密度为

$$f(x) = \begin{cases} \frac{x^{\alpha_1-1}(1-x)^{\alpha_2-1}}{B(\alpha_1,\alpha_2)} & x \in (0,1) \\ 0 & \text{其它}. \end{cases}$$

称 $X$ 服从参数为 $\alpha_1$ 和 $\alpha_2$ 的 Beta 分布，记 $X \sim B(\alpha_1, \alpha_2)$.

**定理 9.3** 若随机变量 $X \sim B(\alpha_1, \alpha_2)$, 则有

$$E[X] = \frac{\alpha_1}{\alpha_1+\alpha_2} \quad \text{和} \quad \mathrm{Var}(X) = \frac{\alpha_1\alpha_2}{(\alpha_1+\alpha_2)^2(\alpha_1+\alpha_2+1)}.$$

**证明** 根据期望的定义有

$$\begin{aligned} E[X] &= \frac{1}{B(\alpha_1,\alpha_2)}\int_0^1 x\cdot x^{\alpha_1-1}(1-x)^{\alpha_2-1}dx = \frac{B(\alpha_1+1,\alpha_2)}{B(\alpha_1,\alpha_2)} = \frac{\alpha_1}{\alpha_1+\alpha_2}, \\ E[X^2] &= \frac{1}{B(\alpha_1,\alpha_2)}\int_0^1 x^{\alpha_1+1}(1-x)^{\alpha_2-1}dx = \frac{B(\alpha_1+2,\alpha_2)}{B(\alpha_1,\alpha_2)} = \frac{\alpha_1+1}{\alpha_1+\alpha_2+1}\frac{\alpha_1}{\alpha_1+\alpha_2}, \end{aligned}$$

由此可得

$$\mathrm{Var}(X) = E[X^2] - E[X]^2 = \frac{\alpha_1(1+\alpha_1)}{(\alpha_1+\alpha_2)(\alpha_1+\alpha_2+1)} - (\frac{\alpha_1}{\alpha_1+\alpha_2})^2 = \frac{\alpha_1\alpha_2}{(\alpha_1+\alpha_2)^2(\alpha_1+\alpha_2+1)}.$$

**例 9.2** 设独立同分布随机变量 $X_1, X_2, \cdots, X_n$ 服从均匀分布 $\mathcal{U}(0,1)$, 记 $X_{(k)}$ 为其顺序统计量, 则

$$X_{(k)} \sim B(k, n-k+1).$$

**证明** 若随机变量 $X_i \sim U(0,1)$ $(i \in [n])$, 则当 $x \in (0,1)$ 时其分布函数 $F(x) = x$. 由此可得到第 $k$ 个统计量 $X_{(k)}$ 的概率密度函数

$$\begin{aligned} f(x) &= \frac{n!}{(k-1)!(n-k)!}(F(x))^{k-1}(1-F(x))^{n-k} = \frac{n!}{(k-1)!(n-k)!}x^{k-1}(1-x)^{n-k} \\ &= \frac{1}{B(k,n-k+1)}x^{k-1}(1-x)^{n-k}. \end{aligned}$$

下面定义 Γ 分布:

**定义 9.7** 如果随机变量 $X$ 的概率密度

$$f(x) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x} & x > 0 \\ 0 & x \leqslant 0 \end{cases}$$

其中 $\alpha > 0$ 和 $\lambda > 0$, 则称随机变量 $X$ 服从参数为 $\alpha$ 和 $\lambda$ 的 $\Gamma$ 分布, 记为$X \sim \Gamma(\alpha, \lambda)$.

**定理 9.4** 若随机变量 $X \sim \Gamma(\alpha, \lambda)$, 则有 $E(X) = \alpha/\lambda$ 和 $\mathrm{Var}(X) = \alpha/\lambda^2$.

**证明** 根据期望的定义有

$$E[X] = \int_0^\infty \frac{\lambda^\alpha}{\Gamma(\alpha)} x^\alpha e^{-\lambda x} dx = \frac{\Gamma(\alpha+1)}{\lambda\Gamma(\alpha)} \int_0^\infty \frac{\lambda^{\alpha+1}}{\Gamma(\alpha+1)} x^\alpha e^{-\lambda x} dx = \alpha/\lambda.$$

以及

$$E[X^2] = \int_0^\infty \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha+1} e^{-\lambda x} dx = \frac{\Gamma(\alpha+2)}{\lambda^2\Gamma(\alpha)} \int_0^\infty \frac{\lambda^{\alpha+2}}{\Gamma(\alpha+2)} x^{\alpha+1} e^{-\lambda x} dx = \alpha(\alpha+1)/\lambda^2,$$

由此可得

$$\mathrm{Var}(X) = E[X^2] - (E[X])^2 = \alpha(\alpha+1)/\lambda^2 - \alpha^2/\lambda^2 = \alpha/\lambda^2.$$

我们有 $\Gamma$ 分布的可加性:

**定理 9.5** 若随机变量 $X \sim \Gamma(\alpha_1, \lambda)$ 和 $Y \sim \Gamma(\alpha_2, \lambda)$, 且 $X$ 与 $Y$ 相互独立, 则 $X + Y \sim \Gamma(\alpha_1 + \alpha_2, \lambda)$.

**证明** 设随机变量 $Z = X + Y$, 根据独立同分布随机变量和函数的分布有随机变量 $Z$ 的概率密度为

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^\infty f_X(x) f_Y(z-x) dx = \int_0^z \frac{\lambda^{\alpha_1}}{\Gamma(\alpha_1)} x^{\alpha_1 - 1} e^{-\lambda x} \frac{\lambda^{\alpha_2}}{\Gamma(\alpha_2)} (z-x)^{\alpha_2 - 1} e^{-\lambda(z-x)} dx \\ &= \frac{\lambda^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} e^{-\lambda z} \int_0^z x^{\alpha_1 - 1} (z-x)^{\alpha_2 - 1} dx \end{aligned}$$

令变量替换 $x = zt$ 有

$$\int_0^z x^{\alpha_1 - 1}(z-x)^{\alpha_2 - 1} dx = z^{\alpha_1+\alpha_2-1} \int_0^1 t^{\alpha_1 - 1}(1-t)^{\alpha_2 - 1} dt = z^{\alpha_1+\alpha_2-1} \mathcal{B}(\alpha_1, \alpha_2)$$

在利用 Beta 函数的性质

$$\mathcal{B}(\alpha_1, \alpha_2) = \int_0^1 t^{\alpha_1 - 1}(1-t)^{\alpha_2 - 1} dt = \frac{\Gamma(\alpha_1)\Gamma(\alpha_2)}{\Gamma(\alpha_1 + \alpha_2)}$$

代入完成证明.

特别地, 若随机变量 $X \sim \Gamma(1/2, 1/2)$, 则其密度函数为

$$f(x) = \begin{cases} \frac{1}{\sqrt{2\pi}} x^{-\frac{1}{2}} e^{-\frac{1}{2}x} & x > 0 \\ 0 & x \leqslant 0 \end{cases}$$

**例 9.3** 若随机变量 $X \sim \mathcal{N}(0,1)$, 则有 $X^2 \sim \Gamma(1/2, 1/2)$.

**解** 首先求解随机变量函数 $Y = X^2$ 的分布函数. 当 $y \leqslant 0$ 时有 $F_Y(y) = 0$; 当 $y > 0$ 时有

$$F_Y(y) = \Pr(X^2 \leqslant y) = \Pr(-\sqrt{y} \leqslant X \leqslant \sqrt{y}) = \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx,$$

由此得到概率密度为 $f_Y(y) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{y}} e^{-\frac{y}{2}}$. 从而得到 $X^2 \sim \Gamma(1/2, 1/2)$.

下面介绍 Dirichlet 分布:

**定义 9.8** 给定 $\alpha_1, \alpha_2, \cdots, \alpha_k \in (0, +\infty)$, 若多元随机向量 $X = (X_1, X_2, \cdots, X_k)$ 的密度函数为

$$f(x_1, x_2, \cdots, x_k) = \begin{cases} \frac{x_1^{\alpha_1-1} x_2^{\alpha_2-1} \cdots x_k^{\alpha_k-1}}{\mathrm{Beta}(\alpha_1, \alpha_2, \cdots, \alpha_k)} & \sum_{i=1}^k x_i = 1,\ x_i > 0\ (i \in [k]), \\ 0 & \text{其它} \end{cases}$$

则称 $X$ 服从参数为 $\alpha_1, \alpha_2, \cdots, \alpha_k$ 的 Dirichlet 分布, 记 $X \sim \mathrm{Dir}(\alpha_1, \alpha_2, \cdots, \alpha_k)$.

Dirichlet 分布是 Beta 分布的一种推广, 当 $k = 2$ 时 Dirichlet 分布退化为 Beta 分布.

**定理 9.6** 若随机向量 $X = (X_1, X_2, \cdots, X_k) \sim \mathrm{Dir}(\alpha_1, \alpha_2, \cdots, \alpha_k)$, 设 $\tilde{\alpha} = \alpha_1 + \alpha_2 + \cdots + \alpha_k$ 和 $\tilde{\alpha}_i = \alpha_i / \tilde{\alpha}$, 则

$$E[X_i] = \tilde{\alpha}_i \qquad \text{和} \qquad Cov(X_i, X_j) = \begin{cases} \frac{\tilde{\alpha}_i(1-\tilde{\alpha}_i)}{\tilde{\alpha}+1} & i = j, \\ -\frac{\tilde{\alpha}_i \tilde{\alpha}_j}{\tilde{\alpha}+1} & i \neq j. \end{cases}$$

**证明** 根据期望的定义有

$$\begin{aligned} E[X_i] &= \frac{\int\int_{\sum_i x_i = 1, x_i \geqslant 0} x_1^{\alpha_1-1} x_2^{\alpha_2-1} \cdots x_k^{\alpha_k-1} \cdot x_i dx_1 \cdots dx_k}{\mathrm{Beta}(\alpha_1, \alpha_2, \cdots, \alpha_k)} \\ &= \frac{\mathrm{Beta}(\alpha_1, \cdots, \alpha_i + 1, \cdots, \alpha_k)}{\mathrm{Beta}(\alpha_1, \cdots, \alpha_i, \cdots, \alpha_k)} = \frac{\alpha_i}{\alpha_1 + \alpha_2 + \cdots + \alpha_k} = \tilde{\alpha}_i. \end{aligned}$$

若 $i = j$, 则有

$$\mathrm{Cov}(X_i, X_i) = E[X_i^2] - (E[X_i])^2 = \frac{Beta(\alpha_1, \cdots, \alpha_i + 2, \cdots, \alpha_k)}{Beta(\alpha_1, \cdots, \alpha_i, \cdots, \alpha_k)} - (\tilde{\alpha}_i)^2 = \frac{\tilde{\alpha}_i(1 - \tilde{\alpha}_i)}{\tilde{\alpha} + 1}.$$

若 $i \neq j$, 则有

$$\begin{aligned} Cov(X_i, X_j) &= E[X_i X_j] - E[X_i]E[X_j] = \frac{\mathrm{Beta}(\alpha_1, \cdots, \alpha_i + 1, \cdots, \alpha_j + 1, \cdots, \alpha_k)}{\mathrm{Beta}(\alpha_1, \cdots, \alpha_i, \cdots, \alpha_j, \cdots, \alpha_k)} - \tilde{\alpha}_i \tilde{\alpha}_j \\ &= \frac{\alpha_i \alpha_j}{\tilde{\alpha}(\tilde{\alpha} + 1)} - \tilde{\alpha}_i \tilde{\alpha}_j = -\frac{\tilde{\alpha}_i \tilde{\alpha}_j}{\tilde{\alpha} + 1}. \end{aligned}$$

## 9.4 正态总体抽样分布定理

### 9.4.1 $\chi^2$ 分布

**定义 9.9** 若 $X_1, X_2, \cdots, X_n$ 是来自总体 $X \sim \mathcal{N}(0,1)$ 的一个样本, 称 $Y = X_1^2 + X_2^2 + \cdots + X_n^2$ 为服从自由度为 $n$ 的 $\chi^2$ 分布, 记 $Y \sim \chi^2(n)$.

根据 $X_1^2 \sim \Gamma(1/2, 1/2)$ 和 $\Gamma$函数的可加性可得 $Y \sim \Gamma(n/2, 1/2)$. 于是有随机变量 $Y$ 的概率密度为

$$f_Y(y) = \begin{cases} \frac{(\frac{1}{2})^{\frac{n}{2}}}{\Gamma(n/2)} y^{\frac{n}{2}-1} e^{-\frac{y}{2}} & y > 0 \\ 0 & y \leqslant 0 \end{cases}$$

下面研究 $\chi^2$ 分布的性质:

**定理 9.7** 若随机变量 $X \sim \chi^2(n)$, 则 $E(X) = n$ 和 $\mathrm{Var}(X) = 2n$; 若随机变量 $X \sim \chi^2(m)$ 和 $Y \sim \chi^2(n)$ 相互独立, 则 $X + Y \sim \chi^2(m+n)$;

**证明** 若随机变量 $X \sim \chi^2(n)$, 则有 $X = X_1^2 + X_2^2 + \cdots + X_n^2$, 其中 $X_1, X_2, \cdots, X_n$ 是总体为 $X' \sim \mathcal{N}(0,1)$ 的一个样本. 我们有

$$\begin{aligned} E[X] &= E[X_1^2 + X_2^2 + \cdots + X_n^2] = nE[X_1^2] = n, \\ \mathrm{Var}(X) &= n\mathrm{Var}(X_1^2) = n[E(X_1^4) - (E(X_1^2))^2] = n(E(X_1^4) - 1). \end{aligned}$$

计算

$$E(X_1^4) = \int_{-\infty}^{+\infty} \frac{x^4}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = -\int_{-\infty}^{+\infty} \frac{x^3}{\sqrt{2\pi}} de^{-\frac{x^2}{2}} = 3\int_{-\infty}^{+\infty} \frac{x^2}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 3$$

可得 $\mathrm{Var}(X) = 2n$.

若随机变量 $X \sim \mathcal{N}(0,1)$, 则

$$E(X^k) = \begin{cases} (k-1)!! & k \text{ 为偶数} \\ 0 & k \text{ 为奇数} \end{cases}$$

其中 $(2k)!! = 2k \cdot (2k-2) \cdot \cdots \cdot 2$ 和 $(2k+1)!! = (2k+1) \cdot (2k-1) \cdot \cdots \cdot 1$.

**例 9.4** 设 $X_1, X_2, X_3, X_4$ 是来自于总体 $\mathcal{N}(0,4)$ 的样本, 以及 $Y = a(X_1 - 2X_2)^2 + b(3X_3 - 4X_4)^2$. 求 $a, b$ 取何值时, $Y$ 服从 $\chi^2$ 分布, 并求其自由度.

**解** 根据正态分布的性质有 $X_1 - 2X_2 \sim \mathcal{N}(0,20)$ 和 $3X_3 - 4X_4 \sim \mathcal{N}(0,100)$, 因此

$$\frac{X_1 - 2X_2}{2\sqrt{5}} \sim \mathcal{N}(0,1), \qquad \frac{3X_3 - 4X_4}{10} \sim \mathcal{N}(0,1),$$

所以当 $a = 1/20, b = 1/100$ 时有 $Y \sim \chi^2(2)$ 成立.

分布可加性:

- 如果 $X \sim \mathcal{N}(\mu_1, a_1^2)$ 和 $Y \sim \mathcal{N}(\mu_2, a_2^2)$, 且 $X$ 与 $Y$ 独立, 那么 $X \pm Y \sim \mathcal{N}(\mu_1 \pm \mu_2, a_1^2 + a_2^2)$;
- 如果 $X \sim B(n_1, p)$ 和 $Y \sim B(n_2, p)$, 且 $X$ 与 $Y$ 独立, 那么 $X + Y \sim B(n_1 + n_2, p)$;
- 如果 $X \sim P(\lambda_1)$ 和 $Y \sim P(\lambda_2)$, 且 $X$ 与 $Y$ 独立, 那么 $X + Y \sim P(\lambda_1 + \lambda_1)$;
- 如果 $X \sim \Gamma(\alpha_1, \lambda)$ 和 $Y \sim \Gamma(\alpha_2, \lambda)$, 且 $X$ 与 $Y$ 独立, 那么 $X + Y \sim \Gamma(\alpha_1 + \alpha_2, \lambda)$.
- 如果 $X \sim \chi(m)$ 和 $Y \sim \chi(n)$, 且 $X$ 与 $Y$ 独立, 那么 $X + Y \sim \chi(m + n)$.

### 9.4.2 $t$ 分布 (student distribution)

**定义 9.10** 随机变量 $X \sim \mathcal{N}(0, 1)$ 和 $Y \sim \chi^2(n)$ 相互独立, 则随机变量

$$T = \frac{X}{\sqrt{Y/n}}$$

服从自由度为 $n$ 的 $t$-分布, 记 $T \sim t(n)$.

随机变量 $T \sim t(n)$ 的概率密度为

$$f(x) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sqrt{n\pi}} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}} \qquad x \in (-\infty, +\infty).$$

由此可知 $t$-分布的密度函数 $f(x)$ 是偶函数. 当 $n > 1$ 为偶数时有

$$\frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sqrt{n\pi}} = \frac{(n-1)(n-3)\cdots 5 \cdot 3}{2\sqrt{n}(n-2)(n-4)\cdots 4 \cdot 2};$$

当 $n > 1$ 为奇数时有

$$\frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sqrt{n\pi}} = \frac{(n-1)(n-3)\cdots 4 \cdot 2}{\pi\sqrt{n}(n-2)(n-4)\cdots 5 \cdot 3}.$$

当 $n \to \infty$ 时, 随机变量 $T \sim t(n)$ 的概率密度

$$f(x) \to \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

因此当 $n$ 足够大时, $f(x)$ 可被近似为 $\mathcal{N}(0, 1)$ 的密度函数.

### 9.4.3 $F$ 分布

**定义 9.11** 设随机变量 $X \sim \chi^2(m)$ 和 $Y \sim \chi^2(n)$ 相互独立, 称随机变量

$$F = \frac{X/m}{Y/n}$$

服从自由度为 $(m,n)$ 的 $F$-分布, 记 $F\sim F(m,n)$.

随机变量 $F\sim F(m,n)$ 的概率密度为

$$f(x)=\begin{cases}\frac{\Gamma(\frac{m+n}{2})(\frac{m}{n})^{\frac{m}{2}}x^{\frac{m}{2}-1}}{\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})(1+\frac{mx}{n})^{\frac{m+n}{2}}} & x>0\\ 0 & x\leqslant 0\end{cases}$$

若随机变量 $F\sim F(m,n)$, 则 $\frac{1}{F}=F(n,m)$.

课题练习:

- 独立同分布随机变量 $X_1,X_2,\cdots,X_n$ 满足 $X_i\sim\mathcal{N}(\mu_i,\sigma_i^2)$, 求 $\sum_{i=1}^n(X_i-\mu_i)^2/\sigma_i^2$ 的分布.
- 设 $X_1,X_2,\cdots,X_9$ 和 $Y_1,Y_2,\cdots,Y_9$ 是分别来自总体 $\mathcal{N}(0,9)$ 的两个独立样本, 求 $(X_1+X_2+\cdots+X_9)/\sqrt{Y_1^2+Y_1^2+\cdots+Y_9^2}$ 的分布.
- 设 $X_1,X_2,\ldots,X_{2n}$ 来自总体 $\mathcal{N}(0,\sigma_2)$ 的样本, 求 $(X_1^2+X_3^2+\cdots+X_{2n-1}^2)/(X_2^2+X_4^2+\cdots+X_{2n}^2)$ 的分布.

### 9.4.4 正态分布的抽样分布定理

**定理 9.8** 设 $X_1,X_2,\cdots,X_n$ 是来自总体 $\mathcal{N}(\mu,\sigma^2)$ 的样本, 则有

$$\bar{X}=\frac{1}{n}\sum_{i=1}^n X_i\sim\mathcal{N}(\mu,\frac{\sigma^2}{n}),\qquad \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\sim\mathcal{N}(0,1).$$

**定理 9.9** 设 $X_1,X_2,\cdots,X_n$ 是来自总体 $\mathcal{N}(\mu,\sigma^2)$ 的样本, 其样本均值和修正样本方差分别为

$$\bar{X}=\sum_{i=1}^n X_i/n \quad 和 \quad S^2=\frac{1}{n-1}\sum_{i=1}^n(X_i-\bar{X})^2,$$

则有 $\bar{X}$ 和 $S^2$ 相互独立, 且

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

此定理证明参考书的附件.

**定理 9.10** 设 $X_1,X_2,\cdots,X_n$ 是来自总体 $\mathcal{N}(\mu,\sigma^2)$ 的样本, 其样本均值和修正样本方差分别为

$$\bar{X}=\sum_{i=1}^n X_i/n \quad 和 \quad S^2=\frac{1}{n-1}\sum_{i=1}^n(X_i-\bar{X})^2,$$

则有

$$\frac{\bar{X}-\mu}{S/\sqrt{n}}\sim t(n-1).$$