证明 第 k 次序统计量 $X_{\langle k\rangle}$ 的分布函数为

$$\begin{aligned}F_{\langle k\rangle}(x) &= P\left(X_{\langle k\rangle} \leqslant x\right) = P\left(X_1, X_2, \cdots, X_n \text{ 中至少有 } k \text{ 个随机变量 } \leqslant x\right)\\ &= \sum_{r=k}^{n} P\left(X_1, X_2, \cdots, X_n \text{ 中恰有 } r \text{ 个随机变量 } \leqslant x, \text{而其余 } n-r \text{ 个随机变量 } > x\right)\\ &= \sum_{r=k}^{n} \binom{n}{r}\left(F(x)\right)^r\left(1-F(x)\right)^{n-r}.\end{aligned}$$

针对 $X_{\langle k\rangle}$ 的密度函数, 对任意 $p\in[0,1]$) 有

$$\sum_{r=k}^{n}\binom{n}{r}p^r(1-p)^{n-r} = \frac{n!}{(k-1)!(n-k)!}\int_0^p t^{k-1}(1-t)^{n-k}dt$$

由此可知

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

在上式两边分别对 x 求导数得到密度函数.

8.3 Beta 分布、Γ 分布、Dirichlet 分布*

8.3.1 两类积分函数

定义 8.3 (Beta-函数) 对任意给定 $\alpha_1>0$ 和 $\alpha_2>0$, 定义 Beta 函数为

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

又称为第一类欧拉积分函数.

容易知道 $\mathrm{Beta}(\alpha_1,\alpha_2)$ 在定义域 $(0,+\infty)\times(0,+\infty)$ 连续. 利用变量替换 $t=1-x$ 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 8.4 (Γ-函数) 对任意给定 $\alpha>0$, 定义 Γ-函数为

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

又称为第二类欧拉积分函数.

性质 8.4 对 Γ-函数有 $\Gamma(1)=1$ 和 $\Gamma(1/2)=\sqrt{\pi}$, 以及当 $\alpha>1$ 时有 $\Gamma(\alpha)=(\alpha-1)\Gamma(\alpha-1)$.

证明 根据定义有

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

利用变量替换 $x=t^{1/2}$ 和正态分布的性质有

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

利用分部积分进一步有

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

根据上述性质, 对任意正整数 n 有

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

由此可知 Γ-函数可以看作 $n!$ 的一种插值函数. 关于 Beta 函数和 Γ-函数, 有如下关系:

性质 8.5 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 Γ-函数的定义有

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

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

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

根据上述性质容易得到

性质 8.6 对任意 $\alpha_1 > 1$ 和 $\alpha_2 > 0$, 有

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

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

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

8.3.2 两类积分函数相关的分布

定义 8.6 (Beta 分布) 给定 $\alpha_1 > 0$ 和 $\alpha_2 > 0$, 若随机变量 X 的概率密度为

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

称 X 服从参数为 α_1 和 α_2 的 Beta 分布，记 $X \sim \mathcal{B}(\alpha_1, \alpha_2)$.

性质 8.7 若随机变量 $X \sim \mathcal{B}(\alpha_1, \alpha_2)$, 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

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

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 8.2 若独立同分布随机变量 $X_1, X_2, \cdots, X_n$ 服从均匀分布 $\mathcal{U}(0, 1)$, 则第 k 次序统计量

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

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

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

定义 8.7 (Γ 分布) 若随机变量 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 α 和 λ 的 Γ 分布, 记为$X \sim \Gamma(\alpha,\lambda)$.

很容易检验密度函数 $f(x)$ 构成一个概率密度函数, 且具有如下性质.

性质 8.8 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

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

定理 8.2 若相互独立的随机变量 $X \sim \Gamma(\alpha_1,\lambda)$ 和 $Y \sim \Gamma(\alpha_2,\lambda)$, 则有 $X+Y \sim \Gamma(\alpha_1+\alpha_2,\lambda)$.

证明 设随机变量 $Z = X+Y$, 根据独立同分布假设和 Γ 分布的概率密度为

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

令变量替换 $x = zt$ 有

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

在利用 Beta 函数的性质

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

代入完成证明.

根据 Γ 分布的定义可知随机变量 $X \sim \Gamma(1/2, 1/2)$ 的密度函数为

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

定理 8.3 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

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

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 8.8 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

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

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 8.4 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

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

证明 根据期望的定义有

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

若 $i=j$, 则有

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

若 $i\neq j$, 则有

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

8.4 重要的统计分布

本节介绍在数理统计中经常用到的三大统计分布: χ^2 分布, t 分布和 F 分布.

8.4.1 χ^2 分布

定义 8.9 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 $Y=X_1^2+X_2^2+\cdots+X_n^2$ 服从自由度为 n 的 χ^2 分布, 记 $Y\sim\chi^2(n)$.

根据 $X_1^2\sim\Gamma(1/2,1/2)$ 和 Γ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 χ^2 分布的性质:

性质 8.9 若随机变量 $X\sim\chi^2(n)$, 则 $E[X]=n$ 和 $\mathrm{Var}(X)=2n$; 若相互独立的随机变量 $X\sim\chi^2(n)$ 和 $Y\sim\chi^2(m)$, 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 $X_1,X_2,\cdots,X_n$ 是来自总体 $\mathcal{N}(0,1)$ 的一个样本. 于是有

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

根据正态分布的密度函数有

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

可得 $\mathrm{Var}(X)=2n$. 也可以根据 Γ 分布的性质 8.8 直接写出期望和方差.

若随机变量 $X\sim\mathcal{N}(0,1)$, 则有

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

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

例 8.3 设 X_1,X_2,X_3,X_4 是来自于总体 $\mathcal{N}(0,4)$ 的一个样本, 当 a 和 b 取何值时 $Y=a(X_1-2X_2)^2+b(3X_3-4X_4)^2$ 服从 χ^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.

8.4.2 t 分布 (student distribution)

定义 8.10 若随机变量 $X\sim\mathcal{N}(0,1)$ 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.

8.4.3 F 分布

定义 8.11 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 $1/F \sim F(n,m)$.

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.

8.5 正态分布的抽样定理

定理 8.5 设 $X_1, X_2, \cdots, X_n$ 是来自总体 $\mathcal{N}(\mu, \sigma^2)$ 的样本, 则有

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

定理 8.6 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

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