

概率论与数理统计

眠云跂石整理

附录

A.1 常用积分

特殊函数 伽马函数与贝塔函数.

伽马函数与递推式: $\Gamma(x) = \int_0^{+\infty} e^{-t} t^{x-1} dt = (x-1)\Gamma(x-1) \quad (x > 0)$

贝塔函数与关系式: $B(x, y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} \quad (x, y > 0)$

勒让德倍量公式: $\Gamma(s)\Gamma\left(s + \frac{1}{2}\right) = \frac{\sqrt{\pi}}{2^{2s-1}}\Gamma(2s) \quad (s > 0)$

余元公式: $B(s, 1-s) = \Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin \pi s} \quad (0 < s < 1)$

$$\begin{cases} \Gamma(n) = (n-1)!, & n \in \mathbb{N}^+, \\ \Gamma\left(\frac{n}{2}\right) = \frac{(n-2)!!}{2^{(n-1)/2}}\sqrt{\pi}, & n \text{ 为正奇数.} \end{cases}$$

$$B(s, s) = \frac{1}{2^{2s-1}}B\left(\frac{1}{2}, s\right) \quad (s > 0)$$

特殊函数的应用

一般的

$$\int_0^{+\infty} \frac{x^c dx}{(1+x^a)^b} = \frac{1}{a} B\left(b - \frac{c+1}{a}, \frac{c+1}{a}\right) \quad (t = (1+x^2)^{-1})$$

$$\int_0^{+\infty} \frac{x^c dx}{(1-x^a)^b} = \frac{1}{a} B\left(1-b, \frac{c+1}{a}\right)$$

$$\int_0^{+\infty} x^n e^{-ax^p} dx = \frac{\Gamma\left(\frac{n+1}{p}\right)}{pa^{\frac{n+1}{p}}}$$

特殊的

$$\int_{-\infty}^{+\infty} \frac{x^m dx}{(1+x^2)^n} = B\left(n - \frac{m+1}{2}, \frac{m+1}{2}\right) \quad (\text{注意积分限})$$

$$\int_0^{+\infty} \frac{x^m dx}{(1+x)^n} = B(n-m-1, m+1)$$

$$\int_0^{+\infty} e^{-ax^p} dx = \frac{\Gamma\left(\frac{1}{p}\right)}{pa^{\frac{1}{p}}}$$

$$\int_0^{+\infty} e^{-x^p} dx = \frac{1}{p}\Gamma\left(\frac{1}{p}\right)$$

$$\int_{-\infty}^{+\infty} e^{-ax^2} dx = \sqrt{\frac{\pi}{a}} \quad (\text{注意积分限})$$

$$\int_0^{+\infty} x^n e^{-ax} dx = \frac{\Gamma(n+1)}{a^{n+1}} = \frac{n!}{a^{n+1}}$$

$$\int_0^{+\infty} x^n e^{-ax^2} dx = \frac{\Gamma\left(\frac{n+1}{2}\right)}{2a^{\frac{n+1}{2}}}$$

$$\int_0^{+\infty} x^{2n} e^{-ax^2} dx = \frac{(2n-1)!!}{2(2a)^n} \sqrt{\frac{\pi}{a}}$$

$$\int_0^{+\infty} x^{2n+1} e^{-ax^2} dx = \frac{(2n)!!}{(2a)^{n+1}}$$

$$\int_0^{+\infty} x^2 e^{-ax^2} dx = \frac{1}{4a} \sqrt{\frac{\pi}{a}}$$

A.2 常用分布

A.2.1 一维离散型

1 二项分布

1.1 基础概念

- $X \sim B(n, p)$.
- 理解: 事件发生的概率为 p , 则重复 n 次试验, 事件发生的次数为 x .
- 概率分布: $P(X = i) = b(i; n, p) = \binom{n}{i} p^i (1-p)^{n-i}$.

1.2 数字特征

- 最可能数: $x = \lfloor (n+1)p \rfloor$.
- 期望: np .
- 方差: $np(1-p)$.

1.3 其它性质

- 二项分布和的函数

$$X_1 \sim B(n_1, p), X_2 \sim B(n_2, p) \Rightarrow X_1 + X_2 \sim B(n_1 + n_2, p).$$

- 发生偶数次的概率为 $p_n = \frac{1}{2}[1 + (1 - 2p)^n]$.
- 记 $f(p) = P(X \leq k)$, 则 $f'(p) < 0$, 并且

$$f(p) = \frac{n!}{k!(n-k-1)!} \int_0^{1-p} t^k (1-t)^{n-k-1} dt.$$

1.4 参数估计

- 矩估计: $p = m/n$. (MVU 估计)
- 极大似然估计: $p = m/n$. (MVU 估计)
- 贝叶斯估计

- 同等无知原则: $p = \frac{X+1}{n+2}$.
- 若先验密度 $h(p) = p^{a-1}(1-p)^{b-a-1}$, 则 $\tilde{p} = \frac{X+c}{n+d}$.

- 区间估计

- 大样本法: 近似地取枢轴变量 $(Y_n - np)/\sqrt{np(1-p)} \sim N(0, 1)$, 则

$$\theta_1, \theta_2 = \frac{n}{n + u_{\alpha/2}^2} \left(\hat{p} + \frac{u_{\alpha/2}^2}{2n} \pm u_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n} + \frac{u_{\alpha/2}^2}{4n^2}} \right), \quad \hat{p} = Y_n/n.$$

取 $\hat{p}(1-\hat{p}) = 1/4$, 则区间长度为 $u_{\alpha/2}/\sqrt{n + u_{\alpha/2}^2}$.

当 $\alpha = 0.05$, $n \geq 40$ 时, 有 $\theta_2 - \theta_1 \leq 0.3$.

- p^k ($k \leq n$) 的无偏估计是 $\frac{X^k}{n^k}$. (下降阶乘幂)

2 泊松分布

2.1 基础概念

- $X \sim P(\lambda)$.
- 理解: 单位时间内事件平均发生 λ 次, 则某一段单位时间内发生的次数为 x .
- 概率分布: $P(X = i) = \lim_{n \rightarrow \infty} b(i; n, \frac{\lambda}{n}) = \frac{e^{-\lambda} \lambda^i}{i!}$.
- 当二项分布满足 $n > 50, p < 0.1, np < 5$ 时, 用泊松分布近似效果较好.

2.2 数字特征

- 最可能数: $k = \lfloor \lambda \rfloor$.
- 期望: $E(X) = \lambda$.
- 方差: $\text{Var}(X) = \lambda$.
- 中位数: $m_e = \frac{\ln 2 \lambda}{\lambda}$.
- $E|X - m_e| = m_e$.

2.3 其它性质

- 泊松分布和的函数 (可加性)

$$X_1 \sim P(\lambda_1), X_2 \sim P(\lambda_2) \Rightarrow X_1 + X_2 \sim P(\lambda_1 + \lambda_2).$$

- 记 $f(\lambda) = P(X \leq k)$, 则 $f'(\lambda) < 0$, 并且

$$f(\lambda) = \frac{1}{k!} \int_{\lambda}^{+\infty} t^k e^{-t} dt.$$

- 若 $X \sim P(\lambda)$, $Y \sim B(X, p)$, 则 $Y \sim P(\lambda p)$.

2.4 参数估计

- 矩估计

- $\lambda = m$. (MVU 估计)
- $\lambda = m_2$ 或 S^2 .

- 极大似然估计: $\lambda = \bar{X}$.

- 贝叶斯估计: 见第四章第五题.

- 区间估计

- 大样本法: 近似地取 $(Y_n - n\lambda)/\sqrt{n\lambda} \sim N(0, 1)$, 则

$$A, B = \bar{X} + u_{\alpha/2}^2/(2n) \pm u_{\alpha/2} \sqrt{u_{\alpha/2}^2/(4n^2) + \bar{X}/n}, \quad \bar{X} = Y_n/n.$$

3 超几何分布

3.1 基础概念

- $X \sim H(N, n, M)$.
- 理解: N 件产品中有 M 件次品, 从总体中抽 n 件时次品的数量 m .
- 概率分布: $P(X = m) = \frac{\binom{M}{m} \binom{N-M}{n-m}}{\binom{N}{n}}$.

3.2 数字特征

- 期望: $E(X) = \frac{nM}{N}$.
- 方差: $\text{Var}(X) = \frac{nM(N-n)(N-M)}{N^2(N-1)} = \frac{nM}{N} \frac{N-n}{N-1} \left(1 - \frac{M}{N}\right)$.

3.3 其它性质

3.4 参数估计

已知 N, n 估计 M .

- 贝叶斯估计: 采用同等无知原则, 则 $M = \frac{N+2}{n+2} (X+1) - 1$.

4 负二项分布

4.1 基础概念

- $X \sim NB(r, p)$.
- 理解: 合格率为 p , 抽取到 r 个合格产品时, 抽到的不合格产品的个数 x .
- 概率分布: $P(X = i) = d(i; r, p) = \binom{i+r-1}{r-1} p^r (1-p)^i$.

4.2 数字特征

- 数学期望: $r(1-p)/p$.

4.3 其它性质

4.4 参数估计

注: $m_e := (X_1 + X_2 + \cdots + X_n)/n$.

- 矩估计: $p = \frac{r}{m_e + r}$.
- 极大似然估计: $p = \frac{r}{m_e + r}$.
- 贝叶斯估计: $p = \frac{nr + 1}{nr + nm_e + 1}$.

5 几何分布

5.1 基础概念

- $X \sim GE(p)$.
- 理解: 合格率为 p , 抽取到第一个合格产品时, 抽到的不合格产品的个数 x .
- 概率分布: $P(X = i) = p(1-p)^i$.

5.2 数字特征

- 数学期望: $(1-p)/p$.

5.3 其它性质

- 几何分布具有无记忆性.
- 若 X_1, X_2, \cdots, X_r 独立同分布 $GE(p)$, 则 $X_1 + X_2 + \cdots + X_r \sim NB(r, p)$.

A.2.2 一维连续型

概率分布函数, 概率密度函数

注: 以下偏度系数定义为 $\beta_1 = \mu_3/\mu_2^{3/2}$, 峰度系数定义为 $\beta_2 = \mu_4/\mu_2^2$.

1 正态分布

1.1 基础概念

- $X \sim N(\mu, \sigma^2)$.
- 概率密度函数: $f(x) = (\sqrt{2\pi}\sigma)^{-1} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$.
- 标准正态分布: $Y = (X - \mu)/\sigma \sim N(0, 1)$.

1.2 数字特征

- 期望: μ .
- 方差: σ^2 .
- k 阶中心矩: $\mu_k = \begin{cases} \sigma^k(k-1)!!, & k \text{ 为偶数,} \\ 0, & k \text{ 为奇数.} \end{cases}$
- 偏度系数: $\beta_1 = 0$.
- 峰度系数: $\beta_2 = 3$.

1.3 其它性质

- 3σ 原则: 0.6826, 0.9544, 0.9974.
- 上 α 分位数: $\Phi(z_\alpha) = 1 - \alpha$.
- 若 $X_1 \sim N(\mu_1, \sigma_1^2)$, $X_2 \sim N(\mu_2, \sigma_2^2)$, 则 $X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.
- 若 X_1, X_2, \dots, X_n 独立同分布 $N(0, 1)$, 则 $Y = X_1^2 + X_2^2 + \dots + X_n^2 \sim \chi_n^2$.
- $\sqrt{n}(\bar{X} - \mu)/\sigma \sim N(0, 1)$.
- $\sqrt{n}(\bar{X} - \mu)/S \sim t_{n-1}$.
- $(n-1)S^2/\sigma^2 \sim \chi_{n-1}^2$.

1.4 参数估计

- 已知 σ^2 , 估计 μ .
 - 矩估计
 - $\mu = m$. (MVU 估计)
 - 区间估计
 - 枢轴变量法
根据 $\sqrt{n}(\bar{X} - \mu)/\sigma \sim N(0, 1)$, 知
$$[\hat{\theta}_1, \hat{\theta}_2] = \left[\bar{X} - \sigma u_{\alpha/2}/\sqrt{n}, \bar{X} + \sigma u_{\alpha/2}/\sqrt{n} \right].$$
- 已知 μ , 估计 σ^2 .
 - 矩估计 $\theta = \sigma^2$.
 - $\hat{\theta} = m_2$. (μ 已知时的 MVU 估计, 且此时均方误差为 $\frac{2}{n}\sigma^4$)
- 估计 μ 和 σ^2 .
 - 矩估计
 - $\mu = m$. (MVU 估计)
 - $\sigma^2 = S^2$. (μ 未知时的 MVU 估计)
 - 极大似然估计: $\mu = m, \sigma^2 = m_2$.
 - 区间估计
 - 枢轴变量法
 - 根据 $\sqrt{n}(\bar{X} - \mu)/S \sim t_{n-1}$, 知一样本 t 区间估计为
$$[\hat{\theta}_1, \hat{\theta}_2] = \left[\bar{X} - St_{n-1}(\alpha/2)/\sqrt{n}, \bar{X} + St_{n-1}(1 - \alpha/2)/\sqrt{n} \right].$$
 - 根据 $(n-1)S^2/\sigma^2 \sim \chi_{n-1}^2$, 知
$$[\hat{\theta}_1, \hat{\theta}_2] = \left[(n-1)S^2/\chi_{n-1}^2(\alpha/2), (n-1)S^2/\chi_{n-1}^2(1 - \alpha/2) \right].$$
 - 无偏估计

$$\tilde{\sigma} = \sqrt{\frac{n-1}{2}} \frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} S.$$

- 估计变异系数 σ/μ .
 - 矩估计: $\sqrt{m_2}/m$ 或 S/m .
- 估计 $N(\theta, 1)$ 的 θ .
 - 贝叶斯估计: 先验密度 $h(\theta) \sim N(\mu, \sigma^2)$, 则

$$\tilde{\theta} = \frac{n}{n+1/\sigma^2} \bar{X} + \frac{1/\sigma^2}{n+1/\sigma^2} \mu.$$

- 对于 $N(\mu_1, \sigma^2)$, $N(\mu_2, \sigma^2)$, 已知 σ^2 , 估计 $\mu_1 - \mu_2$.

- 区间估计 (两样本 t 区间估计)

- 枢轴估计法

$$\text{记 } S^2 = \left[\sum_{i=1}^n (X_i - \bar{X})^2 + \sum_{i=1}^m (Y_i - \bar{Y})^2 \right] / \sqrt{n+m-2},$$

$$\text{由 } T = \sqrt{\frac{mn}{m+n}} \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S} \sim t_{n+m-2}, \text{ 知区间估计为}$$

$$[\hat{\theta}_1, \hat{\theta}_2] = \left[(\bar{X} - \bar{Y}) - St_{n+m-2}(\alpha/2) \sqrt{\frac{n+m}{nm}}, (\bar{X} - \bar{Y}) + St_{n+m-2}(\alpha/2) \sqrt{\frac{n+m}{nm}} \right].$$

- 贝伦斯 - 费歇尔问题: 对于 $N(\mu_1, \sigma_1^2)$, $N(\mu_2, \sigma_2^2)$, 所有参数均未知, 估计 $\mu_1 - \mu_2$.

- 区间估计

- 大样本法: 取枢轴变量

$$N(0, 1) \sim \left[(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2) \right] / \sqrt{\sigma_1^2/n + \sigma_2^2/m} \quad (\text{严格的})$$

$$\sim \left[(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2) \right] / \sqrt{S_1^2/n + S_2^2/m} \quad (\text{近似的})$$

- 对于 $N(\mu_1, \sigma_1^2)$, $N(\mu_2, \sigma_2^2)$, 所有参数均未知, 估计 $\lambda = \sigma_1^2/\sigma_2^2$.

- 区间估计

- 枢轴变量法

$$\text{由 } (S_2^2/\sigma_2^2)/(S_1^2/\sigma_1^2) \sim F_{m-1, n-1}, \text{ 知}$$

$$[\hat{\theta}_1, \hat{\theta}_2] = [(S_1^2/S_2^2)F_{m-1, n-1}(1 - \alpha/2), (S_1^2/S_2^2)F_{m-1, n-1}(\alpha/2)].$$

2 指数分布

2.1 基础概念

- $X \sim E(\lambda)$.
- 概率密度函数: $f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0, \\ 0, & x \leq 0. \end{cases}$
- 分布函数: $F(x) = \begin{cases} 0, & x \leq 0, \\ 1 - e^{-\lambda x}, & x > 0. \end{cases}$

2.2 数字特征

- 数学期望: $E[X] = \lambda^{-1}$.
- 方差: $\text{Var}[X] = \lambda^{-2}$.

2.3 其它性质

- 指数分布具有无记忆性, 即 $P(X > m + t | X > m) = P(X > t)$.
- 若有一批元件寿命 $X \sim E(\lambda)$, 让一个元件开始工作, 每当这个元件坏了就用一个新的替换, 则到经历时间 T 后替换的次数 $Y \sim P(\lambda T)$.
- 若 X_1, X_2, \dots, X_n 独立同分布 $E(\lambda)$, 则

$$X = 2\lambda(X_1 + X_2 + \dots + X_n) \sim \chi_{2n}^2.$$

2.4 参数估计

- 矩估计: $1/\lambda = m$. (MVU 估计)
- 极大似然估计: $\lambda = 1/m$.
- 贝叶斯估计: 若先验密度为 $h(\lambda) = \lambda e^{-\lambda} (\lambda > 0)$, 其它值为零, 则 $\lambda = \frac{n+2}{n\bar{X}+1}$.
- 区间估计

◦ 枢轴变量法

- 估计 λ .

由 $2n\lambda\bar{X} \sim \chi_{2n}^2$, 知

$$[\hat{\theta}_1, \hat{\theta}_2] = \left[\chi_{2n}^2(1-\alpha/2)/(2n\bar{X}), \chi_{2n}^2(\alpha/2)/(2n\bar{X}) \right].$$

- 估计 $1/\lambda$.

由 $2n\lambda\bar{X} \sim \chi_{2n}^2$, 知

$$[\hat{\theta}_1, \hat{\theta}_2] = \left[(2n\bar{X})/\chi_{2n}^2(1-\alpha/2), (2n\bar{X})/\chi_{2n}^2(\alpha/2) \right].$$

- 若 X_1, X_2, \dots, X_n 独立同分布 $E(\lambda_1)$, Y_1, Y_2, \dots, Y_m 独立同分布 $E(\lambda_2)$, 估计 λ_2/λ_1 .

◦ 区间估计 (枢轴变量法)

- 利用 $\frac{4\lambda_1 n \bar{X}}{4\lambda_2 m \bar{Y}} \sim \frac{2n\chi_{2n}^2}{2m\chi_{2m}^2} \sim F_{2n, 2m}$.

3 威布尔分布

概率密度函数: $f(x) = \begin{cases} \lambda \alpha x^{\alpha-1} e^{-\lambda x^\alpha}, & x > 0, \\ 0, & x \leq 0. \end{cases}$

分布函数: $F(x) = \begin{cases} 1 - e^{-\lambda x^\alpha}, & x > 0, \\ 0, & x \leq 0. \end{cases}$

4 均匀分布

4.1 基础概念

- $X \sim R(a, b)$.
- 概率密度函数: $f(x) = \begin{cases} 1/(b-a), & a \leq x \leq b, \\ 0, & x < a \text{ 或 } x > b. \end{cases}$
- 分布函数: $F(x) = \begin{cases} 0, & x \leq a, \\ (x-a)/(b-a), & a < x < b, \\ 1, & x \geq b. \end{cases}$

4.2 数字特征

- 数学期望: $\frac{a+b}{2}$.
- 方差: $\frac{(b-a)^2}{12}$.
- k 阶原点矩: $\alpha_k = \frac{1}{k+1} \frac{b^k - a^k}{b-a}$.
- k 阶中心矩: $\mu_k = \begin{cases} \frac{1}{k+1} \left(\frac{b-a}{2}\right)^k, & k \text{ 为偶数,} \\ 0, & k \text{ 为奇数.} \end{cases}$
- 偏度系数: $\beta_1 = 0$.
- 峰度系数: $\beta_2 = \frac{9}{5}$.

4.3 其它性质

4.4 参数估计

- 估计 $R(\theta_1, \theta_2)$ 的参数.
 - 矩估计: $\theta_1 = m - \sqrt{3m_2}, \theta_2 = m + \sqrt{3m_2}$.
 - 极大似然估计: $\theta_1 = \min_i (X_i), \theta_2 = \max_i (X_i)$.
- 估计 $R(0, \theta)$ 的参数.
 - 极大似然估计: $\hat{\theta} = \max_i (X_i)$.
 - 无偏估计
 - $\hat{\theta} = \frac{n+1}{n} \max_i (X_i)$. (MVU 估计)
 - $\hat{\theta} = (n+1) \min_i (X_i)$. (方差很大)
 - $\hat{\theta} = \max_i (X_i) + \min_i (X_i)$.
 - 区间估计
 - 由 $\hat{\theta}_1 := \max_i (X_i) \sim F_{\hat{\theta}_1}(x) = \frac{nx^{n-1}}{\theta^n}$,
[$\max(X_i), (1-\alpha)^{-\frac{1}{n}} \max(X_i)$] 的置信系数为 $1-\alpha$.

5 对数正态分布

5.1 基础概念

- $\ln X \sim N(\mu, \sigma^2)$.
- 概率密度函数: $f(x, \mu, \sigma) = \begin{cases} (x\sqrt{2\pi}\sigma) \exp\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right], & x > 0, \\ 0, & x \leq 0. \end{cases}$

5.2 数字特征

- 期望: $E(X) = e^{\mu+\sigma^2/2}$.
- 方差: $\text{Var}(X) = (e^{\sigma^2} - 1)e^{2\mu+\sigma^2}$.
- k 阶原点矩: $\alpha_k = e^{\mu k + k^2 \sigma^2/2}$.

6 柯西分布

$$X \sim C(\gamma, x_0).$$

- 概率密度函数: $f(x; x_0, \gamma) = \frac{1}{\pi} \cdot \frac{\gamma}{(x - x_0)^2 + \gamma^2} \quad (-\infty < x < +\infty).$
- 参数估计: 使用样本中位数 \tilde{m} 估计.

7 拉普拉斯分布

7.1 基础概念

- $X \sim \text{La}(\mu, b).$
- 概率密度函数: $f(x) = \frac{1}{2\lambda} e^{-\frac{|x-\mu|}{\lambda}}.$

7.2 数字特征

7.3 其它性质

7.4 参数估计

- 估计 b .
 - 矩估计: $\tilde{b} = m.$
 - 极大似然估计: $\tilde{b} = m_e.$

8 卡方分布

8.1 基础概念

- 自由度为 n 的皮尔逊卡方密度与卡方分布 $X \sim \chi_n^2.$
- 概率密度函数

$$k_n(x) = \begin{cases} \frac{e^{-x/2} x^{(n-2)/2}}{\Gamma\left(\frac{n}{2}\right) 2^{n/2}}, & x > 0, \\ 0, & x \leq 0. \end{cases}$$

8.2 数字特征

- $E(\chi_n^2) = n.$
- $E(\chi_n^2)^{-1} = \frac{1}{n-2}.$
- $E(\chi_n^2)^k = \frac{2^k \Gamma\left(\frac{n}{2} + k\right)}{\Gamma\left(\frac{n}{2}\right)} \quad (k \in \mathbb{Z}).$
- $\text{Var}(\chi_n^2) = 2n.$

注意到方差是均值的两倍, 可以以此检验是否为卡方分布.

8.3 其它性质

- 若 X_1, X_2, \dots, X_n 独立同分布 $N(0, 1)$, 则

$$Y = X_1^2 + X_2^2 + \dots + X_n^2 \sim \chi_n^2.$$

- 若 $X_1 \sim \chi_m^2$ 与 $X_2 \sim \chi_n^2$ 独立, 则

$$X_1 + X_2 \sim \chi_{m+n}^2.$$

- 若 X_1, X_2, \dots, X_n 独立同分布 $E(\lambda)$, 则

$$X = 2\lambda(X_1 + X_2 + \dots + X_n) \sim \chi_{2n}^2.$$

- 若 X_1, X_2, \dots, X_n 独立同分布 $N(\mu, \sigma^2)$, 则

$$(n-1)S^2/\sigma^2 \sim \chi_{n-1}^2.$$

9 t 分布

9.1 基础概念

- 自由度为 n 的 t 分布 $X \sim t_n$.
- 概率密度函数

$$t_n(y) = \frac{\Gamma((n+1)/2)}{\sqrt{n\pi}\Gamma(n/2)} \left(1 + \frac{y^2}{n}\right)^{-\frac{n+1}{2}}.$$

9.2 数字特征

- $E(t_n) = 0$ ($n > 1$).
- $\text{Var}(t_n) = \frac{n}{n-2}$ ($n > 2$).

9.3 其它性质

- 设 X_1, X_2 独立, $X_1 \sim \chi_n^2$, $X_2 \sim N(0, 1)$, 则

$$\frac{X_2}{\sqrt{X_1/n}} \sim t_n$$

- 若 X_1, X_2, \dots, X_n 独立同分布 $N(\mu, \sigma^2)$, 则

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

- 设 X_1, X_2, \dots, X_n 独立同分布 $N(\mu_1, \sigma^2)$, Y_1, Y_2, \dots, Y_m 独立同分布 $N(\mu_2, \sigma^2)$, 且 X_i, Y_j 独立, 则

$$\frac{\sqrt{\frac{nm(n+m-2)}{n+m}} [(\bar{X} + \bar{Y}) - (\mu_1 - \mu_2)]}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 + \sum_{j=1}^m (Y_j - \bar{Y})^2}} \sim t_{n+m-2}.$$

10 F 分布

10.1 基础概念

- 自由度为 (m, n) 的 F 分布 $X \sim F_{m,n}$.
- 概率密度函数

$$f_{m,n}(y) = m^{m/2} n^{n/2} \frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} y^{m/2-1} (my+n)^{-(m+n)/2} \quad (y > 0).$$

10.2 数字特征

- $E(f_{m,n}) = \frac{n}{n-2} \quad (n > 2).$
- $\text{Var}(f_{m,n}) = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}.$

10.3 其它性质

- 设 X_1, X_2 独立, $X_1 \sim \chi_n^2, X_2 \sim \chi_m^2$, 则

$$\frac{X_2}{m} \bigg/ \frac{X_1}{n} \sim F_{m,n}.$$

- 设 X_1, X_2, \dots, X_n 独立同分布 $N(\mu_1, \sigma_1^2), Y_1, Y_2, \dots, Y_m$ 独立同分布 $N(\mu_2, \sigma_2^2)$, 且 X_i, Y_j 独立, 则

$$\frac{S_Y}{\sigma_2^2} \bigg/ \frac{S_X}{\sigma_1^2} \sim F_{m-1, n-1}.$$

- $\forall k, n, \in \mathbb{N}, a \in (0, 1) : kF_{k,n}(a) \geq F_{1,n}(a).$

A.2.3 多维离散型

1 多项分布:

$$X = (X_1, \dots, X_n) \sim M(N; p_1, \dots, p_n).$$

$$P(X_1 = k_1, X_2 = k_2, \dots, X_n = k_n) = \frac{N!}{k_1! k_2! \dots k_n!} p_1^{k_1} p_2^{k_2} \dots p_n^{k_n}.$$

多项分布的边缘分布是二项分布.

$$(X_1, X_2, \dots, X_n) \sim M(N; p_1, p_2, \dots, p_n) \Rightarrow X_1 + X_2 \sim B(N; p_1 + p_2).$$

A.2.4 多维连续型

1 矩形均匀分布

2 二维正态分布

$$X = (X_1, X_2) \sim N(a, b, \sigma_1^2, \sigma_2^2, \rho).$$

$$f(x_1, x_2) = (2\pi\sigma_1\sigma_2\sqrt{1-\rho^2})^{-1} \exp \left[-\frac{1}{2(1-\rho^2)} \left(\frac{(x_1-a)^2}{\sigma_1^2} - \frac{2\rho(x_1-a)(x_2-b)}{\sigma_1\sigma_2} + \frac{(x_2-b)^2}{\sigma_2^2} \right) \right].$$

当且仅当 $\rho = 0$ 时, X_1 和 X_2 独立.

其它性质

- 二维正态分布的边缘分布是正态分布.
- 二维正态分布的条件分布是正态分布.

若 $(X, Y) \sim N(a, b, \sigma_1^2, \sigma_2^2, \rho)$, 则给定 $X = x$ 时 Y 的条件分布为

$$N(b + \rho\sigma_2\sigma_1^{-1}(x-a), \sigma_2^2(1-\rho^2)).$$

- 二维正态分布的边缘分布的和仍为正态分布

若 $(X_1, X_2) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, 则 $Y = X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2).$

- 正态分布的联合分布不一定是二维正态分布.
- 相互独立的正态分布的和仍为正态分布

若 $X_i \sim N(\mu_i, \sigma_i^2)$, 则 $X_1 + \cdots + X_n \sim N(\mu_1 + \cdots + \mu_n, \sigma_1^2 + \cdots + \sigma_n^2)$.

- 若 $Y = X_1 + X_2$ 服从正态分布, X_1, X_2 独立, 则 X_1, X_2 也是正态分布.