

概率论与数理统计笔记

衍瑜¹

2021 年 12 月 14 日

¹Email: fish233yeah@163.com

目录

1	知识点总结	1
1.1	随机事件及其概率	1
1.1.1	基本知识	1
1.2	计算事件的概率	2
1.2.1	部分公式	2
1.3	随机变量及其分布	2
1.3.1	几种离散型随机变量分布律	2
1.3.2	几种连续型随机变量分布律	3
1.3.3	完整概率分布表	3
1.4	多维随机变量及其分布	5
1.4.1	边缘分布	5
1.4.2	条件分布	5
1.4.3	二维随机变量独立性	5
1.4.4	二维随机变量函数 $Z = g(X, Y)$ 的分布	6
1.4.5	二维均匀分布及二维正态分布的性质	6
1.5	随机变量的数字特征	7
1.5.1	期望定义	7
1.5.2	方差定义	8
1.5.3	期望与方差的性质	10
1.5.4	协方差与相关系数	10
1.5.5	矩与协方差矩阵	11
1.5.6	其他	11
1.6	大数定律与中心极限定理	11
1.6.1	依概率收敛	11
1.6.2	大数定律	12
1.6.3	中心极限定理	12
1.7	样本及抽样分布	13
1.7.1	基本定义	13
1.7.2	Γ 函数	14

1.7.3	抽样分布	14
1.7.4	抽样分布样本均值和方差的分布	18
1.8	参数估计	19
1.8.1	矩估计	19
1.8.2	极大似然估计	20
1.8.3	估计量无偏性、有效性、一致性	21
1.8.4	正态总体参数的区间估计	21
2	解题笔记	25
2.1	多维随机变量及其分布	25
2.2	随机变量的数字特征	25

Chapter 1

知识点总结

1.1 随机事件及其概率

1.1.1 基本知识

1. 样本空间与样本点：对于随机事件 E ，它的所有可能结果组成的集合即为 E 的样本空间，通常记为 S ；而样本空间中的元素（即 E 的每个结果）则称为样本点
2. 摩根率（对偶率）： $\overline{A+B} = \overline{A}\overline{B}$, $\overline{AB} = \overline{A} + \overline{B}$
3. 排列组合公式： $A_n^m = \frac{n!}{(n-m)!}$, $C_n^m = \frac{A_n^m}{m!}$ 。此外我们有： $C_n^m = C_n^{n-m}$ 和 $C_n^m = C_{n-1}^m + C_{n-1}^{m-1}$
4. 条件概率：称 $P(B|A) = \frac{P(AB)}{P(A)}$ 为事件 B 在事件 A 发生的条件下发生的概率。而对于多个事件 B_1, B_2, \dots ，若它们为两两互不相容的事件，则我们有 $P\left(\bigcup_{i=1}^{\infty} B_i|A\right) = \sum_{i=1}^{\infty} P(B_i|A)$
5. 全概率公式：设 E 的样本空间为 S ， A 为 E 的事件， B_1, B_2, \dots, B_n 为 S 的一个划分，且 $P(B_i) > 0$ ，则有：

$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \dots + P(A|B_n)P(B_n)$$

6. 贝叶斯公式：设 E 的样本空间为 S ， A 为 E 的事件， B_1, B_2, \dots, B_n 为 S 的一个划分，且 $P(A) > 0$ ， $P(B_i) > 0$ ，则有：

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{j=1}^n P(A|B_j)P(B_j)}$$

7. 事件独立定义：设 A_1, A_2, \dots, A_n 为 n 个事件，若对于其中任意2个、任意3个、 \dots 、任意 n 个事件的积事件的概率都等于各事件概率之积，则称事件 A_1, A_2, \dots, A_n 相互独立
8. 若非负实数 k_1, k_2, \dots, k_n 满足 $k_1 + k_2 + \dots + k_n = 1$ ，则分布函数 $F_1(x), F_2(x), \dots, F_n(x)$ 的线性组合 $F(x) = k_1F_1(x) + k_2F_2(x) + \dots + k_nF_n(x)$ 仍为某一随机变量的分布函数

1.2 计算事件的概率

1.2.1 部分公式

$$P(A - B) = P(A) - P(AB)$$

$$P(A + B) = P(A) + P(B) - P(AB)$$

$$P(A + B + C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC)$$

$$P(A_1 + A_2 + \cdots + A_n) = \sum_{i=1}^n P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i A_j) + \cdots + (-1)^{n-1} P(A_1 A_2 \cdots A_n)$$

$$P(AB) = P(A)P(B|A) = P(B)P(A|B)$$

$$P(A_1 A_2 \cdots A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1 A_2) \cdots P(A_n|A_1 A_2 \cdots A_{n-1})$$

1.3 随机变量及其分布

1.3.1 几种离散型随机变量分布律

1. 两点分布：又称为伯努利分布、(0-1)分布，随机变量 X 只可能取0和1两个值，分布律为：

$$P\{X = k\} = p^k (1-p)^{1-k}, \quad k = 0, 1$$

2. 二项分布：设试验 E 只有两个可能的结果 A 和 \bar{A} ， $P(A) = p$ ，则称 E 为伯努利试验，我们将试验 E 独立重复 n 次的话则称其为 n 重伯努利试验。现在假设我们用 X 表示 n 重伯努利试验中事件 A 发生的次数，则其符合二项分布，记为 $X \sim b(n, p)$ 或 $X \sim B(n, p)$ ，分布律为：

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

3. 泊松分布：设随机变量 X 所有可能取的值为 $0, 1, 2, \cdots$ ，分布率为：

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

其中 $\lambda > 0$ ，此时我们称随机变量 X 服从参数为 λ 的泊松分布，记为 $X \sim \pi(\lambda)$ 或 $X \sim P(\lambda)$ 。此外，若我们记 $np_n = \lambda$ ，我们还有泊松定理：

$$\lim_{n \rightarrow \infty} C_n^k p_n^k (1-p_n)^{n-k} = \frac{\lambda^k e^{-\lambda}}{k!}$$

4. 几何分布：设在伯努利试验中事件 A 发生的概率为 p ，若记事件 A 首次发生时的经历的试验总次数为 X ，则我们称随机变量 X 服从参数为 p 的几何分布，分布律为：

$$P\{X = k\} = pq^{k-1}, \quad k = 1, 2, \cdots$$

5. 超几何分布：设在 N 个物品中有 N_1 个次品，从中任取 n ($1 \leq n \leq N_1$) 个，设 X 为这 n 个物品中的次品的数量，则随机变量 X 服从超几何分布，分布律为：

$$P\{X = k\} = \frac{C_{N_1}^k C_{N-N_1}^{n-k}}{C_N^n}$$

1.3.2 几种连续型随机变量分布律

1. 均匀分布：记为 $X \sim U(a, b)$

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

$$F(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \leq x < b \\ 1, & x \geq b \end{cases}$$

2. 指数分布：记服从参数 $\lambda > 0$ 的指数分布为 $X \sim E(\lambda)$ 或 $X \sim \text{Exp}(\lambda)$

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

指数分布具有无记忆性，即：

$$P(X > s + t | X > s) = P(X > t)$$

3. 正态分布：又叫做高斯分布，记服从参数 $\mu, \sigma (\sigma > 0)$ 的正态分布为 $X \sim N(\mu, \sigma^2)$

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

$$F(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt, \quad -\infty < x < \infty$$

当 $\mu = 0, \sigma = 1$ 时称随机变量 X 服从标准正态分布，其概率密度函数和分布函数为

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad -\infty < x < \infty$$

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt, \quad -\infty < x < \infty$$

正态分布有如下几个性质：

- (a) 正态分布 $N(\mu, \sigma^2)$ 的数学期望为 μ ，方差为 σ^2
- (b) 若 $X \sim N(\mu, \sigma^2)$ ，则 $Y = \frac{X-\mu}{\sigma} \sim N(0, 1)$
- (c) 若 $X \sim N(\mu, \sigma^2)$ ，则 $Y = aX + b \sim N(a\mu + b, (a\sigma)^2)$ ， $a \neq 0$
- (d) 若 $X_i \sim N(\mu_i, \sigma_i^2)$ ， $(i = 1, 2, \dots, n)$ ，且 X_1, X_2, \dots, X_n 相互独立，则其线性组合 $Y = a_1X_1 \pm a_2X_2 \pm \dots \pm a_nX_n$ 仍服从正态分布，且有

$$Y \sim N(a_1\mu_1 \pm a_2\mu_2 \pm \dots \pm a_n\mu_n, a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \dots + a_n^2\sigma_n^2)$$

1.3.3 完整概率分布表

分布	参数	分布律或概率密度函数	期望	方差
(0-1)分布	$0 < p < 1$	$P\{X = k\} = p^k (1-p)^{1-k}, k = 0, 1$	p	$p(1-p)$
二项分布	$n \geq 1$ $0 < p < 1$	$P\{X = k\} = C_n^k p^k (1-p)^{n-k}$ $k = 0, 1, 2, \dots, n$	np	$np(1-p)$
负二项分布 (帕斯卡分布)	$r \geq 1$ $0 < p < 1$	$P\{X = k\} = C_{k-1}^{r-1} p^r (1-p)^{k-r}$ $k = r, r+1, \dots$	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$
几何分布	$0 < p < 1$	$P\{X = k\} = p(1-p)^{k-1}$ $k = 1, 2, \dots$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
超几何分布	N, M, n $(M \leq N)$ $(n \leq N)$	$P\{X = k\} = \frac{C_M^k \cdot C_{N-M}^{n-k}}{C_N^n}$ $\max\{0, n-N+M\} \leq k \leq \min\{n, M\}$	$\frac{nM}{N}$	$\frac{nM}{N} \left(1 - \frac{M}{N}\right) \left(\frac{N-n}{N-1}\right)$
泊松分布	$\lambda > 0$	$P\{X = k\} = \frac{\lambda^k e^{-\lambda}}{k!}$ $k = 0, 1, 2, \dots$	λ	λ
均匀分布	$a < b$	$f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{else} \end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
正态分布	μ $\sigma > 0$	$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2
Γ 分布	$\alpha > 0$ $\beta > 0$	$f(x) = \begin{cases} \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}}, & x > 0 \\ 0, & \text{else} \end{cases}$	$\alpha\beta$	$\alpha\beta^2$
指数分布 (负指数分布)	$\lambda > 0$	$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
χ^2 分布	$n \geq 1$	$f(x) = \begin{cases} \frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, & x > 0 \\ 0, & \text{else} \end{cases}$	n	$2n$
韦布尔分布	$\eta > 0$ $\beta > 0$	$f(x) = \begin{cases} \frac{\beta}{\eta} \left(\frac{x}{\eta}\right)^{\beta-1} e^{-\left(\frac{x}{\eta}\right)^\beta}, & x > 0 \\ 0, & \text{else} \end{cases}$	$\eta \Gamma\left(\frac{1}{\beta} + 1\right)$	$\eta^2 \left\{ \Gamma\left(\frac{2}{\beta} + 1\right) - \left[\Gamma\left(\frac{1}{\beta} + 1\right) \right]^2 \right\}$
瑞利分布	$\sigma > 0$	$f(x) = \begin{cases} \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}, & x > 0 \\ 0, & \text{else} \end{cases}$	$\sqrt{\frac{\pi}{2}} \sigma$	$\frac{4-\pi}{2} \sigma^2$
β 分布	$\alpha > 0$ $\beta > 0$	$f(x) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}, & 0 < x < 1 \\ 0, & \text{else} \end{cases}$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$
对数正态分布	μ $\sigma > 0$	$f(x) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma x} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}, & x > 0 \\ 0, & \text{else} \end{cases}$	$e^{\mu + \frac{\sigma^2}{2}}$	$e^{2\mu + \sigma^2} (e^{\sigma^2} - 1)$
柯西分布	a $\lambda > 0$	$f(x) = \frac{1}{\pi} \cdot \frac{\lambda}{\lambda^2 + (x-a)^2}$	不存在	不存在
t 分布	$n \geq 1$	$f(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi} \cdot \Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}$	$0, n > 1$	$\frac{n}{n-2}, n > 2$
F 分布	n_1, n_2	$f(x) = \begin{cases} \frac{\Gamma\left(\frac{n_1+n_2}{2}\right) \cdot \left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} \cdot x^{\frac{n_1}{2}-1}}{\Gamma\left(\frac{n_1}{2}\right) \cdot \Gamma\left(\frac{n_2}{2}\right) \cdot \left(1 + \frac{n_1}{n_2} x\right)^{\frac{n_1+n_2}{2}}}, & x > 0 \\ 0, & \text{else} \end{cases}$	$\frac{n_2}{n_2-2}$ $n_2 > 2$	$\frac{2n_2^2(n_1+n_2-2)}{n_1(n_2-2)^2(n_2-4)}$ $n_2 > 4$

表 1.1: 概率分布表

1.4 多维随机变量及其分布

1.4.1 边缘分布

1. 离散型随机变量边缘分布:

$$p_{i\cdot} = P\{X = x_i\} = \sum_{j=1}^{\infty} p_{ij}$$

$$p_{\cdot j} = P\{Y = y_j\} = \sum_{i=1}^{\infty} p_{ij}$$

2. 连续型随机变量边缘分布:

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

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

1.4.2 条件分布

1. 离散型随机变量条件分布:

$$P\{X = x_i | Y = y_j\} = \frac{P\{X = x_i, Y = y_j\}}{P\{Y = y_j\}} = \frac{p_{ij}}{p_{\cdot j}}$$

$$P\{Y = y_j | X = x_i\} = \frac{P\{X = x_i, Y = y_j\}}{P\{X = x_i\}} = \frac{p_{ij}}{p_{i\cdot}}$$

2. 连续型随机变量条件分布:

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)}$$

$$f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)}$$

1.4.3 二维随机变量独立性

1. 设 $F(x, y)$ 及 $F_X(x)$, $F_Y(y)$ 分别是二维随机变量 (X, Y) 的分布函数及边缘分布函数, 则随机变量 X 和 Y 相互独立的两个相互等价的条件为:

(a) $F(x, y) = F_X(x) F_Y(y)$ 在平面上恒成立

(b) $f(x, y) = f_X(x) f_Y(y)$ 在平面上几乎处处成立

当随机变量 X 和 Y 相互独立时, 显然我们有:

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)} = \frac{f_X(x) f_Y(y)}{f_Y(y)} = f_X(x)$$

$$f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)} = \frac{f_X(x) f_Y(y)}{f_X(x)} = f_Y(y)$$

即二维随机变量 X, Y 相互独立的充要条件为 X, Y 的所有条件分布都与其边缘分布相同

2. 设多维随机变量 (X_1, X_2, \dots, X_m) 和 (Y_1, Y_2, \dots, Y_n) 相互独立, 则 X_i 和 Y_j 相互独立。又若 h, g 是连续函数, 则 $h(X_1, X_2, \dots, X_m)$ 和 $g(Y_1, Y_2, \dots, Y_n)$ 相互独立

1.4.4 二维随机变量函数 $Z = g(X, Y)$ 的分布

1. 离散型随机变量函数概率分布计算:

$$P(Z = z_k) = P(g(X, Y) = z_k) = \sum_{g(x_i, y_j) = z_k} p_{ij}$$

2. 连续型随机变量函数概率分布计算:

$$P(Z = z) = P(g(X, Y) \leq z) = P((X, Y) \in G) = \iint_G f(x, y) dx dy$$

其中 $G = \{(x, y) | g(x, y) \leq z\}$

1.4.5 二维均匀分布及二维正态分布的性质

1. 二维均匀分布

- (a) 定义:

设 G 为平面上的有界区域, 面积为 $S(G)$, 若二维随机变量 (X, Y) 在 G 上服从均匀分布, 则

$$f(x, y) = \begin{cases} \frac{1}{S(G)}, & (x, y) \in G \\ 0, & else \end{cases}$$

记为 $(X, Y) \sim U(G)$

- (b) 性质:

- i. 设矩形区域为 $G = \{(x, y) | a \leq x \leq b, c \leq y \leq d\}$, 则二维随机变量 (X, Y) 在矩形区域 G 上服从均匀分布的充分必要条件为 X, Y 分别在 $(a, b), (c, d)$ 上分别服从均匀分布

2. 二维正态分布

- (a) 定义:

二维随机变量 (X, Y) 在平面上服从二维正态分布, 则

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right]}$$

记为 $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$

- (b) 性质:

- i. 对于二维正态随机变量 (X, Y) , X 和 Y 相互独立的充分必要条件为 $\rho = 0$
- ii. 二维正态随机变量 $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ 的边缘密度函数 $f_X(x)$ 与 $f_Y(y)$ 与 ρ 的取值无关, 恒为

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}$$

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}$$

即若 (X, Y) 服从二维正态分布, 则 X 和 Y 都服从一维正态分布

- iii. 若 X 和 Y 都服从一维正态分布, 且 X 和 Y 相互独立, 则 (X, Y) 服从二维正态分布
- iv. (X, Y) 服从二维正态分布的充分必要条件为 X 和 Y 的任意线性组合 $l_1X + l_2Y$ (l_1, l_2 不全为零) 均服从一维正态分布
- v. (正态变量线性变换不变性) 若 (X, Y) 服从二维正态分布, 记随机变量

$$U = aX + bY, V = cX + dY, \text{ 且 } \begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0$$

则 (U, V) 也服从二维正态分布

1.5 随机变量的数字特征

1.5.1 期望定义

1. 定义:

- (a) 对于离散型随机变量 X , 设分布律为 $P(X = x_k) = p_k, k = 1, 2, \dots$, 若级数 $\sum_{k=1}^{\infty} x_k p_k$ 绝对收敛, 则 X 的期望定义为

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

- (b) 对于连续型随机变量 X , 设概率密度函数为 $f(x)$, 若积分 $\int_{-\infty}^{+\infty} x f(x) dx$ 绝对收敛, 则 X 的期望定义为

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

2. 一维随机变量函数期望:

设函数 $g(x)$ 连续, 则随机变量 $Y = g(X)$ 的期望如下

- (a) 对于离散型随机变量 X , 设分布律为 $P(X = x_k) = p_k, k = 1, 2, \dots$, 若级数 $\sum_{k=1}^{\infty} g(x_k) p_k$ 绝对收敛, 则有

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

- (b) 对于连续型随机变量 X , 设概率密度函数为 $f(x)$, 若积分 $\int_{-\infty}^{+\infty} g(x) f(x) dx$ 绝对收敛, 则有

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

3. 二维随机变量函数期望:

设函数 $g(x, y)$ 连续, 则随机变量 $Z = g(X, Y)$ 的期望如下

- (a) 对于二维离散型随机变量 (X, Y) , 设分布律为 $P(X = x_i, Y = y_j) = p_{ij}$, $i, j = 1, 2, \dots$, 若级数 $\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} g(x_i, y_j) p_{ij}$ 绝对收敛, 则有

$$E(Z) = E[g(X, Y)] = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} g(x_i, y_j) p_{ij}$$

特别地, 我们有

$$E(X) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} x_i p_{ij} = \sum_{i=1}^{\infty} x_i p_{i\cdot}$$

$$E(Y) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} y_j p_{ij} = \sum_{j=1}^{\infty} y_j p_{\cdot j}$$

- (b) 对于连续型随机变量 (X, Y) , 设概率密度函数为 $f(x, y)$, 若积分 $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) f(x, y) dx dy$ 绝对收敛, 则有

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

特别地, 我们有

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

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

1.5.2 方差定义

1. 定义: 随机变量 X 的期望存在时若 $E[X - E(X)]^2$ 也存在, 则有

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

而标准差定义为

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

- (a) 对于离散型随机变量 X , 设分布律为 $P(X = x_k) = p_k$, $k = 1, 2, \dots$, 若其期望 $E(X)$ 存在, $\sum_{k=1}^{\infty} [x_k - E(X)]^2 p_k$ 绝对收敛, 则 X 的方差定义为

$$D(X) = \sum_{k=1}^{\infty} [x_k - E(X)]^2 p_k$$

(b) 对于连续型随机变量 X , 设概率密度函数为 $f(x)$, 若其期望 $E(X)$ 存在,

$\int_{-\infty}^{+\infty} [x - E(X)]^2 f(x) dx$ 绝对收敛, 则 X 的方差定义为

$$D(X) = \int_{-\infty}^{+\infty} [x - E(X)]^2 f(x) dx$$

2. 一维随机变量函数方差:

设函数 $g(x)$ 连续, 则随机变量 $Y = g(X)$ 的方差为

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

(a) 对于离散型随机变量 X , 设分布律为 $P(X = x_k) = p_k, k = 1, 2, \dots$, 若其期望 $E(X)$ 存在,

在, $\sum_{k=1}^{\infty} [g(x_k) - E(X)]^2 p_k$ 绝对收敛, 则 X 的方差定义为

$$D(X) = \sum_{k=1}^{\infty} [g(x_k) - E(X)]^2 p_k$$

(b) 对于连续型随机变量 X , 设概率密度函数为 $f(x)$, 若其期望 $E(X)$ 存在,

$\int_{-\infty}^{+\infty} [g(x) - E(X)]^2 f(x) dx$ 绝对收敛, 则 X 的方差定义为

$$D(X) = \int_{-\infty}^{+\infty} [g(x) - E(X)]^2 f(x) dx$$

3. 二维随机变量函数方差:

设函数 $g(x, y)$ 连续, 则随机变量 $Z = g(X, Y)$ 的方差为

(a) 对于二维离散型随机变量 (X, Y) , 设分布律为 $P(X = x_i, Y = y_j) = p_{ij}, i, j = 1, 2, \dots$,

若级数 $\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \{g(x_i, y_j) - E[g(X, Y)]\}^2 p_{ij}$ 绝对收敛, 则有

$$D(Z) = D[g(X, Y)] = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \{g(x_i, y_j) - E[g(X, Y)]\}^2 p_{ij}$$

特别地, 我们有

$$D(X) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} [x_i - E(X)]^2 p_{ij} = \sum_{i=1}^{\infty} [x_i - E(X)]^2 p_{i.}$$

$$D(Y) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} [y_j - E(Y)]^2 p_{ij} = \sum_{j=1}^{\infty} [y_j - E(Y)]^2 p_{.j}$$

(b) 对于连续型随机变量 (X, Y) , 设概率密度函数为 $f(x, y)$, 若积分 $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) f(x, y) dx dy$ 绝对收敛, 则有

$$D(Z) = D[g(X, Y)] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \{g(x, y) - E[g(X, Y)]\}^2 f(x, y) dx dy$$

特别地, 我们有

$$D(X) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} [x - E(X)]^2 f(x, y) dx dy = \int_{-\infty}^{+\infty} [x - E(X)]^2 f_X(x) dx$$

$$D(Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} [y - E(Y)]^2 f(x, y) dx dy = \int_{-\infty}^{+\infty} [y - E(Y)]^2 f_Y(y) dy$$

1.5.3 期望与方差的性质

1. $E(kX \pm c) = kE(X) \pm c$, $D(kX \pm c) = k^2 D(X)$
2. $E(X_1 \pm X_2) = E(X_1) \pm E(X_2)$, $D(X_1 \pm X_2) = D(X_1) + D(X_2) \pm \text{Cov}(X_1, X_2)$
3. 若 X 与 Y 相互独立, 则有

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

$$D(XY) = D(X)D(Y) + [E(X)]^2 D(Y) + [E(Y)]^2 D(X)$$

4. 若 X 与 Y 相互独立, $g(x)$ 和 $h(x)$ 是两个连续函数, 则 $U = g(X)$ 和 $V = h(Y)$ 也都是随机变量, 且 U 和 V 相互独立

1.5.4 协方差与相关系数

1. 定义:

(a) 协方差 $\text{Cov}(X, Y)$ 定义为:

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

- i. 二维离散型随机变量协方差:

$$\text{Cov}(X, Y) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} [x_i - E(X)][y_j - E(Y)] p_{ij}$$

- ii. 二维连续型随机变量协方差:

$$\text{Cov}(X, Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} [x - E(X)][y - E(Y)] f(x, y) dx dy$$

(b) 相关系数 ρ_{XY} 定义为:

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

2. 协方差性质:

- (a) $\text{Cov}(X, Y) = \text{Cov}(Y, X)$
- (b) $\text{Cov}(X, k) = 0$, $\text{Cov}(k, Y) = 0$
- (c) $\text{Cov}(aX \pm b, cY \pm d) = ac \text{Cov}(X, Y)$
- (d) $\text{Cov}(X_1 \pm X_2, Y) = \text{Cov}(X_1, Y) \pm \text{Cov}(X_2, Y)$

3. 相关系数性质:

- (a) $|\rho_{XY}| \leq 1$
- (b) $|\rho_{XY}| = 1$ 的充要条件是, 存在常数 a, b 使

$$P\{Y = aX + b\} = 1$$

成立 (即相关系数描述的是随机变量的线性相关性, $\rho_{XY} = 0$ 时即表示 X, Y 不相关)

1.5.5 矩与协方差矩阵

1. 矩定义:

(a) k 阶原点矩: $E(X^k)$

(b) k 阶中心矩: $E\{[X - E(X)]^k\}$

(c) $k + l$ 阶混合矩: $E(X^k Y^l)$

(d) $k + l$ 阶混合中心矩: $E\{[X - E(X)]^k [Y - E(Y)]^l\}$

2. 协方差矩阵定义: 设 n 维随机变量 (X_1, X_2, \dots, X_n) 的协方差

$$c_{ij} = c_{ji} = \text{Cov}(X_i, X_j)$$

均存在, 则称矩阵

$$C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix}$$

为 n 维随机变量 (X_1, X_2, \dots, X_n) 的协方差矩阵, 显然协方差矩阵为一个对称矩阵

1.5.6 其他

1. 切比雪夫不等式: 设随机变量 X 具有数学期望 $E(X) = \mu$, 方差 $D(X) = \sigma^2$, 则对于任意正数 ε , 我们有下述不等式成立:

$$P\{|X - \mu| \geq \varepsilon\} \leq \frac{\sigma^2}{\varepsilon^2}$$

1.6 大数定律与中心极限定理

1.6.1 依概率收敛

1. 定义: 设 X_1, X_2, \dots 是一个随机变量序列, a 是一个常数, 若对于任意 $\varepsilon > 0$, 有

$$\lim_{n \rightarrow \infty} P\{|X_n - a| < \varepsilon\} = 1$$

则称序列 X_1, X_2, \dots 依概率收敛于 a , 记为

$$X_n \xrightarrow{P} a$$

2. 性质: 设 $X_n \xrightarrow{P} a$, $Y_n \xrightarrow{P} b$, 又函数 $g(x, y)$ 在点 (a, b) 处连续, 则

$$g(X_n, Y_n) \xrightarrow{P} g(a, b)$$

1.6.2 大数定律

1. 大数定律：设 X_1, X_2, \dots 为一随机变量序列，若对于任意 $\varepsilon > 0$ ，有

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{1}{n} \sum_{k=1}^n X_k - E \left(\frac{1}{n} \sum_{k=1}^n X_k \right) \right| < \varepsilon \right\} = 1$$

则称随机变量序列 X_1, X_2, \dots 服从大数定律

2. 切比雪夫大数定律：设 X_1, X_2, \dots 是相互独立的随机变量序列，且方差一致有上界，则对于任意 $\varepsilon > 0$ ，有

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{1}{n} \sum_{k=1}^n X_k - \frac{1}{n} \sum_{k=1}^n E(X_k) \right| < \varepsilon \right\} = 1$$

3. 辛钦大数定律（弱大数定律）：设 X_1, X_2, \dots 是相互独立且服从同一分布的随机变量序列，且具有数学期望 $E(X_k) = \mu$ ，则对于任意 $\varepsilon > 0$ ，有

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

即 $\bar{X} = \frac{1}{n} \sum_{k=1}^n X_k$ 依概率收敛于 μ ， $\bar{X} \xrightarrow{P} \mu$

4. 伯努利大数定律：设 n_A 是 n 次独立试验中事件 A 发生的次数， p 是事件 A 在每次试验中发生的概率，则对于任意 $\varepsilon > 0$ ，有

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{n_A}{n} - p \right| < \varepsilon \right\} = 1$$

即事件 A 发生的频率依概率收敛于事件 A 发生的概率， $\frac{n_A}{n} \xrightarrow{P} p$

1.6.3 中心极限定理

1. 李雅普诺夫中心极限定理：设随机变量序列 X_1, X_2, \dots 相互独立，且具有期望和方差

$$E(X_k) = \mu_k, D(X_k) = \sigma_k^2 > 0$$

则随机变量

$$Z_n = \frac{\sum_{k=1}^n X_k - \sum_{k=1}^n \mu_k}{\sqrt{\sum_{k=1}^n \sigma_k^2}}$$

的分布函数 $F_n(x)$ 对任意实数 x 满足

$$\lim_{n \rightarrow \infty} F_n(x) = \lim_{n \rightarrow \infty} P \left\{ \frac{\sum_{k=1}^n X_k - \sum_{k=1}^n \mu_k}{\sqrt{\sum_{k=1}^n \sigma_k^2}} \leq x \right\} = \Phi(x)$$

即当 n 很大时，随机变量 Z_n 近似地服从标准正态分布 $N(0, 1)$

2. 列维-林德伯格中心极限定理（独立同分布的中心极限定理）：设随机变量序列 X_1, X_2, \dots 相互独立且同分布，且具有期望和方差

$$E(X_k) = \mu, D(X_k) = \sigma^2 > 0$$

则随机变量

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

的分布函数 $F_n(x)$ 对任意实数 x 满足

$$\lim_{n \rightarrow \infty} F_n(x) = \lim_{n \rightarrow \infty} P \left\{ \frac{\sum_{k=1}^n X_k - n\mu}{\sqrt{n}\sigma} \leq x \right\} = \Phi(x)$$

即当 n 很大时，随机变量 Y_n 近似地服从标准正态分布 $N(0, 1)$ ，或者说，当 n 很大时随机变量序列的算术平均 $\bar{X} = \frac{1}{n} \sum_{k=1}^n X_k$ 服从均值为 μ 方差为 $\frac{\sigma^2}{n}$ 的正态分布

3. 棣莫弗-拉普拉斯中心极限定理：本定理是列维-林德伯格中心极限定理的特殊情况。设伯努利试验中试验 A 出现的概率为 p ，记 n 重伯努利试验中 A 出现的次数为 η_n （即 η_n 服从参数为 n, p 的二项分布），则对任意实数 x 有

$$\lim_{n \rightarrow \infty} P \left\{ \frac{\eta_n - np}{\sqrt{np(1-p)}} \leq x \right\} = \Phi(x)$$

1.7 样本及抽样分布

1.7.1 基本定义

1. 样本：设 X 是具有分布函数 F 的随机变量，若 X_1, X_2, \dots, X_n 相互独立且均有同一分布函数 F ，则称 X_1, X_2, \dots, X_n 为从总体 X 得到的容量为 n 的简单随机样本，简称样本，它们的观察值 x_1, x_2, \dots, x_n 称为样本值，又称为 X 的 n 个独立的观察值
2. 统计量：设 X_1, X_2, \dots, X_n 是来自总体 X 的一个样本， $g(X_1, X_2, \dots, X_n)$ 是 X_1, X_2, \dots, X_n 的函数，若 g 中不含未知参数，则称 $g(X_1, X_2, \dots, X_n)$ 是一统计量，统计量的分布就称为抽样分布

常见统计量有：

(a) 样本均值：

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

(b) 样本方差：

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

(c) 样本标准差:

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

1.7.2 Γ 函数

1. Γ 函数的定义为:

$$\Gamma(s) = \int_0^{+\infty} e^{-x} x^{s-1} dx \quad (s > 0)$$

2. 计算性质:

(a) $\Gamma(1) = 1, \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

(b) $\Gamma(s+1) = s\Gamma(s)$ (即我们有, 对任意正整数 n , $\Gamma(n+1) = n!$)

(c) $\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin \pi s}$

(d) $\lim_{s \rightarrow 0^+} \Gamma(s) = +\infty$

1.7.3 抽样分布

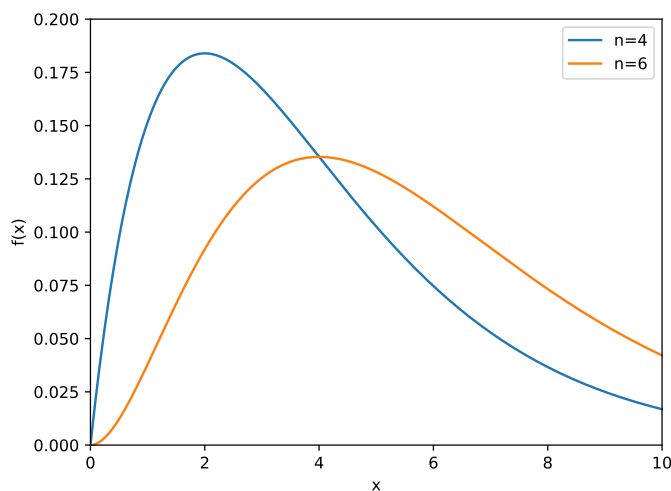
1. χ^2 分布 (卡方分布): 设 X_1, X_2, \dots, X_n 是来自总体 $N(0, 1)$ 的样本, 则称统计量

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

服从自由度为 n 的 χ^2 分布, 记为 $\chi^2 \sim \chi^2(n)$, 概率密度函数为

$$f(x) = \begin{cases} \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, & x > 0 \\ 0, & \text{else} \end{cases}$$

函数图像为:



χ^2 分布性质:

(a) 设 $\chi_1^2 \sim \chi^2(n_1)$, $\chi_2^2 \sim \chi^2(n_2)$, 且 χ_1^2 和 χ_2^2 相互独立, 则有

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

(b) 设 $\chi^2 \sim \chi^2(n)$, 则有

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

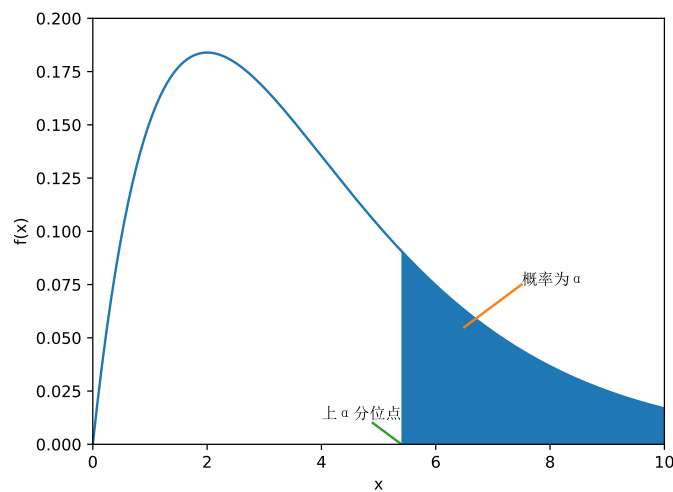
(c) 设总体 $X \sim N(\mu, \sigma^2)$, X_1, X_2, \dots, X_n 是来自总体 X 的样本, 则

$$\frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma^2} \sim \chi^2(n)$$

(d) 设 $\chi^2 \sim \chi^2(n)$, 对于给定的正数 $\alpha \in (0, 1)$, 则满足

$$P\{\chi^2 > \chi_\alpha^2(n)\} = \alpha$$

的点 $x = \chi_\alpha^2(n)$ 称为 χ^2 的上 α 分位点



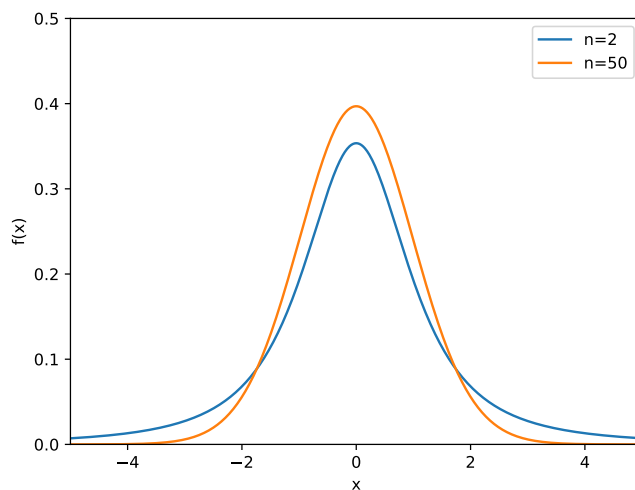
2. t 分布: 设 $X \sim N(0, 1)$, $Y \sim \chi^2(n)$, 且 X, Y 相互独立, 则称随机变量

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

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

$$f(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi} \cdot \Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}$$

函数图像为:



t 分布性质:

(a) 对于 t 分布的密度函数 $f(x)$ 我们有

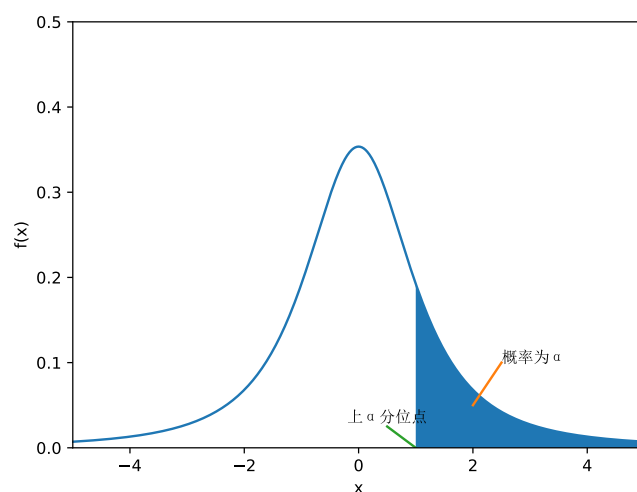
$$\lim_{n \rightarrow \infty} f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

即当 n 足够大时 t 分布近似于标准正态分布 $N(0, 1)$

(b) 设 $t \sim t(n)$, 对于给定的正数 $\alpha \in (0, 1)$, 则满足

$$P\{t > t_\alpha(n)\} = \alpha$$

的点 $x = t_\alpha(n)$ 称为 t 的上 α 分位点



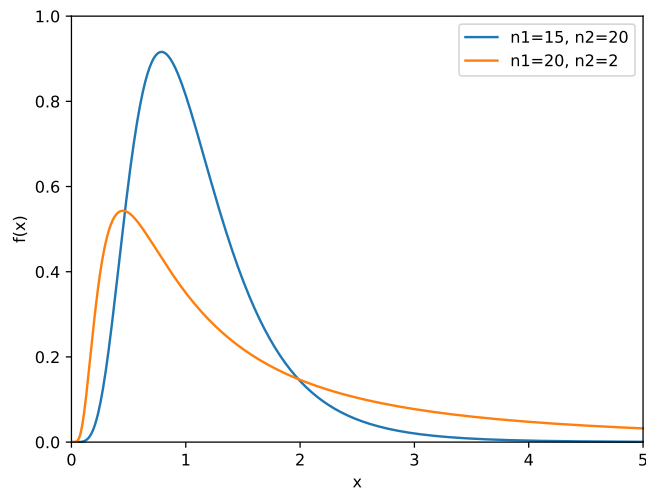
3. F 分布: 设 $U \sim \chi^2(n_1)$, $V \sim \chi^2(n_2)$, 且 U, V 相互独立, 则称随机变量

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

服从自由度为 (n_1, n_2) 的 F 分布, 记为 $F \sim F(n_1, n_2)$, 概率密度函数为

$$f(x) = \begin{cases} \frac{\Gamma\left(\frac{n_1+n_2}{2}\right) \cdot \left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} \cdot x^{\frac{n_1}{2}-1}}{\Gamma\left(\frac{n_1}{2}\right) \cdot \Gamma\left(\frac{n_2}{2}\right) \cdot \left(1 + \frac{n_1}{n_2}x\right)^{\frac{n_1+n_2}{2}}}, & x > 0 \\ 0, & \text{else} \end{cases}$$

函数图像为:



F 分布性质:

(a) 若 $F \sim F(n_1, n_2)$, 则有

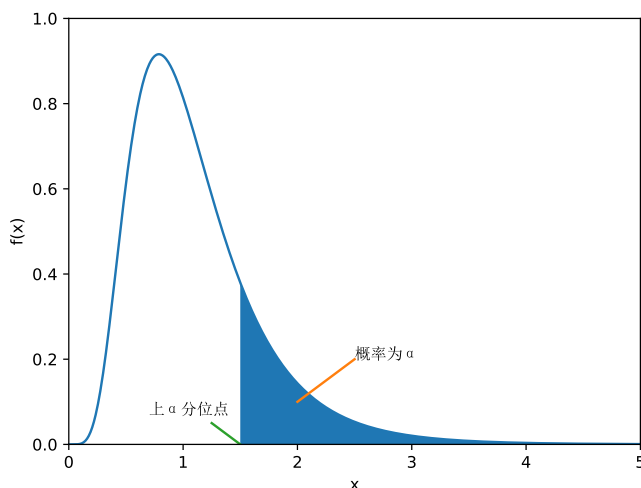
$$\frac{1}{F} \sim F(n_2, n_1)$$

(b) 若 $X \sim t(n)$, 则 $X^2 \sim F(1, n)$

(c) 设 $F \sim F(n_1, n_2)$, 对于给定的正数 $\alpha \in (0, 1)$, 则满足

$$P\{F > F_\alpha(n_1, n_2)\} = \alpha$$

的点 $x = F_\alpha(n_1, n_2)$ 称为 F 的上 α 分位点



1.7.4 抽样分布样本均值和方差的分布

1. 设总体 X 具有均值 μ 和方差 σ^2 , X_1, X_2, \dots, X_n 是来自 X 的一个样本, \bar{X} 和 S^2 分别是样本均值和样本方差, 则我们有

$$E(\bar{X}) = \mu, D(\bar{X}) = \frac{\sigma^2}{n}, E(S^2) = \sigma^2$$

2. 单个正态总体样本均值分布: 设 X_1, X_2, \dots, X_n 是来自正态总体 $N(\mu, \sigma^2)$ 的一个样本, \bar{X} 是样本均值, 则有

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

3. 两个正态总体样本均值分布: 设 X_1, X_2, \dots, X_m 与 Y_1, Y_2, \dots, Y_n 分别是来自正态总体 $N(\mu_1, \sigma_1^2)$ 和 $N(\mu_2, \sigma_2^2)$ 的样本, 且这两个正态总体 $N(\mu_1, \sigma_1^2)$ 和 $N(\mu_2, \sigma_2^2)$ 相互独立, \bar{X}, \bar{Y} 分别是其样本均值, 则有

$$\bar{X} \pm \bar{Y} \sim N\left(\mu_1 \pm \mu_2, \frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}\right)$$

4. 非正态总体样本均值近似分布: 设 X_1, X_2, \dots, X_n 是来自任意总体 X 的一个样本, 总体的期望为 $E(X) = \mu$, 方差为 $D(X) = \sigma^2$, 则当 n 较大时, 有

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

5. 设 X_1, X_2, \dots, X_n 是来自正态总体 $N(\mu, \sigma^2)$ 的一个样本, \bar{X} 和 S^2 分别是样本均值和样本方差, 则 \bar{X} 和 S^2 相互独立, 且有

$$\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2} = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

6. 设 X_1, X_2, \dots, X_n 是来自正态总体 $N(\mu, \sigma^2)$ 的一个样本, \bar{X} 和 S^2 分别是样本均值和样本方差, 则有

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

7. 设 X_1, X_2, \dots, X_{n_1} 与 Y_1, Y_2, \dots, Y_{n_2} 分别是来自正态总体 $N(\mu_1, \sigma_1^2)$ 和 $N(\mu_2, \sigma_2^2)$ 的样本, 且这两个正态总体 $N(\mu_1, \sigma_1^2)$ 和 $N(\mu_2, \sigma_2^2)$ 相互独立, \bar{X}, \bar{Y} 和 S_1^2, S_2^2 分别是样本均值和样本方差, 则有

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

而当 $\sigma_1^2 = \sigma_2^2 = \sigma^2$ 时, 还有

$$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2)$$

1.8 参数估计

1.8.1 矩估计

1. 样本原点矩与中心矩: 对于总体 X , 设 X_1, X_2, \dots, X_n 为其样本, 则样本原点矩与中心矩定义为:

(a) 样本 k 阶原点矩:

$$A_k = \frac{1}{n} \sum_{i=1}^n X_i^k$$

(b) 样本 k 阶中心矩:

$$B_k = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^k$$

2. 矩估计定义: 表示总体原点矩或中心矩与未知参数关系的函数方程(组)

$$\begin{cases} \theta_1 = \theta_1(\mu_1, \mu_2, \dots, \mu_k) \\ \theta_2 = \theta_2(\mu_1, \mu_2, \dots, \mu_k) \\ \vdots \\ \theta_k = \theta_k(\mu_1, \mu_2, \dots, \mu_k) \end{cases}$$

称为矩估计方程(组), 用对应的样本矩替代方程组中的总体矩后得到的

$$\hat{\theta}_i = \theta_i(A_1, A_2, \dots, A_k)$$

分别称为 θ_i 的矩估计量, 矩估计量的观察值称为矩估计值

3. 矩估计性质:

- (a) 若 $\hat{\theta}$ 为参数 θ 的矩估计量, $g(x)$ 为 x 的连续函数, 则 $g(\hat{\theta})$ 是 $g(\theta)$ 的矩估计量
- (b) 若 $\hat{\theta}_1, \hat{\theta}_2$ 为参数 θ_1, θ_2 的矩估计量, $g(x, y)$ 为 (x, y) 的连续函数, 则 $g(\hat{\theta}_1, \hat{\theta}_2)$ 是 $g(\theta_1, \theta_2)$ 的矩估计量

1.8.2 极大似然估计

1. 单参数似然函数: 对于总体 X , 设 $f(x; \theta)$ 为总体 X 的概率密度函数, 其中 θ 是未知参数, X_1, X_2, \dots, X_n 为总体 X 的一个样本, 则自变量为 θ , 定义域为 Θ 的非负函数

$$L(x_1, x_2, \dots, x_n; \theta) = \prod_{i=1}^n f(x_i; \theta), \theta \in \Theta$$

称为容量为 n 的样本似然函数, 或称为样本值 x_1, x_2, \dots, x_n 的似然函数

2. 单参数极大似然估计: 对于总体 X , 设 X_1, X_2, \dots, X_n 为其样本, 若存在 $\hat{\theta} = \hat{\theta}(x_1, x_2, \dots, x_n)$, 使

$$L(\hat{\theta}) = \max_{\theta \in \Theta} L(\theta)$$

则称似然函数的最大值点 $\hat{\theta}(x_1, x_2, \dots, x_n)$ 为参数 θ 的极大似然估计值, 而称 $\hat{\theta}(X_1, X_2, \dots, X_n)$ 为参数 θ 的极大似然估计量

3. 多参数似然函数: 对于总体 X , 设 $f(x; \theta_1, \theta_2, \dots, \theta_m)$ 为总体 X 的概率密度函数, 其中 θ_i 是未知参数, X_1, X_2, \dots, X_n 为总体 X 的一个样本, 则自变量为 $\theta_1, \theta_2, \dots, \theta_m$, 定义域为 Θ^m 的非负函数

$$L(x_1, x_2, \dots, x_n; \theta_1, \theta_2, \dots, \theta_m) = \prod_{i=1}^n f(x_i; \theta_1, \theta_2, \dots, \theta_m), (\theta_1, \theta_2, \dots, \theta_m) \in \Theta^m$$

称为容量为 n 的样本似然函数, 或称为样本值 x_1, x_2, \dots, x_n 的似然函数

4. 多参数极大似然估计: 对于总体 X , 设 X_1, X_2, \dots, X_n 为其样本, 若存在一组 $\hat{\theta}_i = \hat{\theta}_i(x_1, x_2, \dots, x_n)$, 使

$$L(\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_m) = \max_{(\theta_1, \theta_2, \dots, \theta_m) \in \Theta^m} L(\theta_1, \theta_2, \dots, \theta_m)$$

则称似然函数的最大值点 $\hat{\theta}_i(x_1, x_2, \dots, x_n)$ 为参数 θ_i 的极大似然估计值, 而称 $\hat{\theta}_i(X_1, X_2, \dots, X_n)$ 为参数 θ_i 的极大似然估计量

5. 极大似然估计性质:

- (a) 若 $\hat{\theta}$ 为参数 θ 的极大似然估计量, $g(x)$ 为 x 的连续函数, 则 $g(\hat{\theta})$ 是 $g(\theta)$ 的极大似然估计量

- (b) 若 $\hat{\theta}_1, \hat{\theta}_2$ 为参数 θ_1, θ_2 的极大似然估计量, $g(x, y)$ 为 (x, y) 的连续函数, 则 $g(\hat{\theta}_1, \hat{\theta}_2)$ 是 $g(\theta_1, \theta_2)$ 的极大似然估计量
- (c) 若任意总体 X 的数学期望和方差均存在, 则其数学期望的矩估计量和极大似然估计量相同, 均为 $\frac{1}{n} \sum_{i=1}^n X_i$; 其方差的矩估计量和极大似然估计量也相同, 均为 $\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$

1.8.3 估计量无偏性、有效性、一致性

1. 无偏性: 设 $\hat{\theta}$ 是未知参数 θ 的估计量, 若 $E(\hat{\theta}) = \theta$, 则称 $\hat{\theta}$ 是 θ 的无偏估计量
2. 有效性: 设 $\hat{\theta}_1(X_1, X_2, \dots, X_n)$ 与 $\hat{\theta}_2(X_1, X_2, \dots, X_n)$ 均为 θ 的无偏估计量, 若有

$$D(\hat{\theta}_1) < D(\hat{\theta}_2)$$

则称 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 更有效

3. 一致性 (相合性): 设 $\hat{\theta}_n(X_1, X_2, \dots, X_n)$ 为参数 θ 的估计量, 若对任意 $\theta \in \Theta$, 当样本容量 $n \rightarrow \infty$ 时, $\hat{\theta}_n(X_1, X_2, \dots, X_n)$ 依概率收敛于 θ , 即对任意 $\varepsilon > 0$, 有

$$\lim_{n \rightarrow \infty} P(|\hat{\theta}_n - \theta| \geq \varepsilon) = 0$$

则称 $\hat{\theta}_n$ 为 θ 一致估计量, 也叫相合估计量

4. 性质:

- (a) 设 $\hat{\theta}_n$ 为 θ 无偏估计量, 且 $\lim_{n \rightarrow \infty} D(\hat{\theta}_n) = 0$, 则 $\hat{\theta}_n$ 为 θ 一致估计量
- (b) 设 $\hat{\theta}_n$ 为 θ 的一致估计量, 若函数 $g(x)$ 连续, 则 $g(\hat{\theta}_n)$ 为 $g(\theta)$ 的一致估计量
- (c) 任意样本原点矩是总体原点矩的一致估计量。进一步地, 若待估参数 $\theta = g(\mu_1, \mu_2, \dots, \mu_k)$, 而函数 g 连续, 则矩估计量 $\hat{\theta} = g(A_1, A_2, \dots, A_k)$ 为 θ 的一致估计量

1.8.4 正态总体参数的区间估计

1. 一个正态总体

- (a) 待估参数为 μ , 且 σ^2 已知时:
用作估计的随机变量及其分布为:

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$$

置信度为 $1 - \alpha$ 的置信区间为:

$$\left(\bar{X} \pm \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}} \right)$$

单侧置信限为:

$$\begin{aligned}\underline{\mu} &= \bar{X} - \frac{\sigma}{\sqrt{n}} z_{\alpha} \\ \bar{\mu} &= \bar{X} + \frac{\sigma}{\sqrt{n}} z_{\alpha}\end{aligned}$$

(b) 待估参数为 μ , 且 σ^2 未知时:

用作估计的随机变量及其分布为:

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

置信度为 $1 - \alpha$ 的置信区间为:

$$\left(\bar{X} \pm \frac{S}{\sqrt{n}} t_{\frac{\alpha}{2}}(n-1) \right)$$

单侧置信限为:

$$\begin{aligned}\underline{\mu} &= \bar{X} - \frac{S}{\sqrt{n}} t_{\alpha}(n-1) \\ \bar{\mu} &= \bar{X} + \frac{S}{\sqrt{n}} t_{\alpha}(n-1)\end{aligned}$$

(c) 待估参数为 σ^2 , 且 μ 未知时:

用作估计的随机变量及其分布为:

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

置信度为 $1 - \alpha$ 的置信区间为:

$$\left(\frac{(n-1)S^2}{\chi_{\frac{\alpha}{2}}^2(n-1)}, \frac{(n-1)S^2}{\chi_{1-\frac{\alpha}{2}}^2(n-1)} \right)$$

单侧置信限为:

$$\begin{aligned}\underline{\sigma^2} &= \frac{(n-1)S^2}{\chi_{\alpha}^2(n-1)} \\ \bar{\sigma^2} &= \frac{(n-1)S^2}{\chi_{1-\alpha}^2(n-1)}\end{aligned}$$

2. 两个正态总体

(a) 待估参数为 $\mu_1 - \mu_2$, 且 σ_1^2, σ_2^2 已知时:

用作估计的随机变量及其分布为:

$$Z = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0, 1)$$

置信度为 $1 - \alpha$ 的置信区间为:

$$\left((\bar{X} - \bar{Y}) \pm \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \cdot z_{\frac{\alpha}{2}} \right)$$

单侧置信限为:

$$\underline{\mu_1 - \mu_2} = (\bar{X} - \bar{Y}) - \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \cdot z_{\alpha}$$

$$\overline{\mu_1 - \mu_2} = (\bar{X} - \bar{Y}) + \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \cdot z_{\alpha}$$

(b) 待估参数为 $\mu_1 - \mu_2$, 且 $\sigma_1^2 = \sigma_2^2 = \sigma^2$ 未知时:

用作估计的随机变量及其分布为:

$$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2)$$

置信度为 $1 - \alpha$ 的置信区间为:

$$\left((\bar{X} - \bar{Y}) \pm \sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \cdot t_{\frac{\alpha}{2}}(n_1 + n_2 - 2) \right)$$

单侧置信限为:

$$\underline{\mu_1 - \mu_2} = (\bar{X} - \bar{Y}) - \sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \cdot t_{\alpha}(n_1 + n_2 - 2)$$

$$\overline{\mu_1 - \mu_2} = (\bar{X} - \bar{Y}) + \sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \cdot t_{\alpha}(n_1 + n_2 - 2)$$

(c) 待估参数为 $\frac{\sigma_1^2}{\sigma_2^2}$, 且 μ_1, μ_2 未知时:

用作估计的随机变量及其分布为:

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

置信度为 $1 - \alpha$ 的置信区间为:

$$\left(\frac{S_1^2}{S_2^2} \cdot \frac{1}{F_{\frac{\alpha}{2}}(n_1 - 1, n_2 - 1)}, \frac{S_1^2}{S_2^2} \cdot \frac{1}{F_{1-\frac{\alpha}{2}}(n_1 - 1, n_2 - 1)} \right)$$

单侧置信限为:

$$\left(\frac{\sigma_1^2}{\sigma_2^2} \right) = \frac{S_1^2}{S_2^2} \cdot \frac{1}{F_{\alpha}(n_1 - 1, n_2 - 1)}$$

$$\overline{\left(\frac{\sigma_1^2}{\sigma_2^2} \right)} = \frac{S_1^2}{S_2^2} \cdot \frac{1}{F_{1-\alpha}(n_1 - 1, n_2 - 1)}$$

3. 一些性质

- (a) 若 $[\underline{\theta}, \bar{\theta}]$ 为 θ 的置信度为 $1-\alpha$ 的置信区间, $g(x)$ 为单调函数, 则 $[g(\underline{\theta}), g(\bar{\theta})]$ (或 $[g(\bar{\theta}), g(\underline{\theta})]$) 为 $g(\theta)$ 的置信度为 $1-\alpha$ 的置信区间

Chapter 2

解题笔记

2.1 多维随机变量及其分布

1. 离散型随机变量和连续型随机变量混合

例：设随机变量 X 与 Y 独立，其中 X 的概率分布为

$$X \sim \begin{bmatrix} 1 & 2 \\ 0.3 & 0.7 \end{bmatrix}$$

而 Y 的概率密度为 $f(Y)$ ，求随机变量 $U = X + Y$ 的概率密度 $g(u)$

解：设 $F(y)$ 是 Y 的分布函数，则由全概公式以及 X 和 Y 相互独立知 $U = X + Y$ 的分布函数为

$$\begin{aligned} G(u) &= P(X + Y \leq u) \\ &= 0.3P(X + Y \leq u|X = 1) + 0.7P(X + Y \leq u|X = 2) \\ &= 0.3P(Y \leq u - 1|X = 1) + 0.7P(Y \leq u - 2|X = 2) \\ &= 0.3P(Y \leq u - 1) + 0.7P(Y \leq u - 2) \\ &= 0.3F(u - 1) + 0.7F(u - 2) \end{aligned}$$

因此 $U = X + Y$ 的概率密度为

$$g(u) = G'(u) = 0.3f(u - 1) + 0.7f(u - 2)$$

2.2 随机变量的数字特征

1. 利用随机变量分解法和数字特征运算规律求解数字特征

例：将 n 个球（ $1 \sim n$ 号）随机放进 n 个盒子（ $1 \sim n$ 号）中去，一个盒子装一个球，将一个球装入与球同号的盒子中称为一个配对，记 X 为总的配对数，求 $E(X)$

解：记随机变量 X_i 为 i 号球的配对数，显然每个球的配对数只有1和0两种可能，而尽管

随机变量 X_i 互相并不独立，但我们有 $P(X_i = 1) = \frac{1}{n}$ ，而又因为 $X = \sum_{i=1}^n X_i$ ，故我们有

$$E(X) = \sum_{i=1}^n E(X_i) = n \cdot \frac{1}{n} = 1$$