概率论的基本概念

随机现象及其统计规律 Page 2
- 大量随机现象呈现的内在规律

随机试验 Page 2
- 可重复
- 知道全部可能结果
- 单次结果不能预言

样本点 Page 4
- 一次试验中的结果
- 不可分的结果

样本空间 Page 4
- 全部样本点构成的集合 Ω

事件 Page 3
- 样本点构成的集合 A, A_i
- 必然事件：Ω，全集。
- 不可能事件：ϕ，不包含样本点的集合

事件的关系与运算 Page 5
- 等价于集合的关系与运算
- 研究事件转化成研究集合

包含关系 Page 5
- A 是 B 的子事件
- $\forall x \in A \Rightarrow x \in B$
- $A \subset B$ and $B \subset A \Rightarrow A = B$

和事件 Page 6
- A 和 B 至少有一个事件发生（A 或 B）
- $A \cup B = \{x|x \in A \ or \ x \in B\}$

积事件 Page 6
- A 与 B 同时发生
- $A \cap B = \{x|x \in A \ and \ x \in B\}$

运算规律 Page 7
- Venn 图可以形象描述
- 交换律：$A \cup B = B \cup A, AB = BA$
- 结合律：$A \cup (B \cup C) = (A \cup B) \cup C, (AB)C = A(BC)$
- 分配律：$A(B \cup C) = AB \cup BC, A \cup (BC) = (A \cup B)(A \cup C), A(B - C) = AB - AC$
- 吸收律：$A \subset B \Rightarrow AB = A, A \cup B = B$
- 对偶律：$\overline{A \cup B} = \bar{A} \cap \bar{B}, \overline{A \cap B} = \bar{A} \cup \bar{B}$

互不相容事件 Page 6
- A 与 B 不同时发生
- $A \cap B = \phi$

对立事件 Page 7
- A 和 B 互逆，非此即彼
- $A \cap B = \phi$ and $A \cup B = \Omega$

差事件 Page 7
- A 发生的同时 B 不发生
- $A - B = A\bar{B} = A - AB$

频率 Page 10
- 定义在同一概率模型上
- 可以用来估计概率
- 试验次数增多时，频率具有稳定于一个定值的趋势
- 多次试验结束后的事件发生频率是个常数

概率 Page 9
- 度量事件发生可能性

组合数学 Page No
- 加法原理、乘法原理
- 排列与组合等计数公式与方法

古典概率 Page 11
- 特殊的概率模型（古典概型）
- 仅有有限个基本事件
- 样本点等可能出现
- $P(A) = \frac{|A|}{|\Omega|}$

几何概率 Page 10
- 特殊的概率模型（均匀分布）
- 样本点无穷多
- 样本点等可能出现
- $P(A)$ =长度比、面积比或体积比

概率公理化定义 Page 13
- 事件的函数：$P(A)$
- 非负性：$0 \leq P(A) \leq 1$
- 规范性：$P(\Omega) = 1$
- 可列可加性：
$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

概率性质 Page 13
- $P(\phi) = 0$
- $P\left(\bigcup_{i=1}^{m} A_i\right) = \sum_{i=1}^{m} P(A_i)$
- $P(A) + P(\bar{A}) = 1$
- $A \subset B \Rightarrow P(A) \leq P(B), P(B - A) = P(B) - P(A)$
- $P(B - A) = P(B) - P(AB)$
- $P(A \cup B) = P(A) + P(B) - P(AB)$
- 多除少补原理

Notice:
- 区分互斥、对立与独立
- 区分两两独立和相互独立

乘法公式 Page 18
- $\forall B, P(B) > 0 \Rightarrow P(AB) = P(B)P(A|B)$
- $P(A_1A_2 \ldots A_{n-1}) > 0 \Rightarrow P(A_1A_2 \ldots A_n) = P(A_1)P(A_2|A_1) \ldots P(A_n|A_1A_2 \ldots A_{n-1})$
- 乘法公式用来计算积的概率
- 选择恰当的的事件做条件

条件概率 Page 16
- $P(A|B) = \frac{P(AB)}{P(B)} \quad (P(B) > 0)$
- Ω 上的条件概率
- B 上的概率

样本空间的有限划分 Page 20
- $B_i \cap B_j = \phi, i \neq j$
- $B_1 \cup B_2 \cup \cdots \cup B_n = \Omega$
- 这是一种分类讨论思想
- 适用于复杂问题分析

Byes 公式 Page 22
- $B_1 \cup B_2 \cup \cdots \cup B_n$ 是 Ω 的有限划分
- $P(B_k|A) = \frac{P(B_k)P(A|B_k)}{\sum_{i=1}^{n} P(B_i)P(A|B_i)}$
- 事后概率
- 执果索因

全概率公式 Page 20
- $B_1 \cup B_2 \cup \cdots \cup B_n$ 是 Ω 的有限划分
- $P(A) = \sum_{i=1}^{n} P(B_i)P(A|B_i)$
- 事前概率
- 由因导果

条件概率的性质 Page 17
- 相同的发生条件下，概率对应性质都成立
- 符合概率公理化定义，是概率模型

事件的独立 Page 24
- 源起：$P(A|B) = P(A)$
- 含义：事件 B 的发生不改变事件 A 的概率
- 定义：$P(AB) = P(A)P(B)$

替换性定理 Page 24
- 下列命题等价
- A, B 相互独立
- A, $\bar{B}$ 相互独立
- $\bar{A}$, B 相互独立
- $\bar{A}$, $\bar{B}$ 相互独立

n 个事件的独立性 Page 27
- $\forall s(1 < s \leq n), 1 \leq i_1 < i_2 < \cdots < i_s \leq i_n$
$P(A_{i_1}A_{i_2} \cdots A_{i_s}) = P(A_{i_1})P(A_{i_2}) \cdots P(A_{i_s})$

替换性定理 Page 27
- n 个事件相互独立，则把它们中的任意多个事件换成其对立事件后所得的 n 个事件仍然相互独立。

n 个事件两两独立 Page 26
- $\forall \ 1 \leq i_1 < i_2 <\leq i_n$
$P(A_{i_1}A_{i_2}) = P(A_{i_1})P(A_{i_2})$

随机变量的分布

随机变量 Page 37
- $\omega \to X \in R$
- 样本空间的数值化处理
- 用随机变量的取值规律刻画样本点分布规律

分布函数 Page 38
- $F(x) = P\{X \le x\}$
- 刻画随机变量取值分布的重要方法
- 可以刻画随机变量取任何值的概率

分布函数的性质 Page 40
- 单调不降性：$\forall x_1 < x_2 \Rightarrow F(x_1) \le F(x_2)$
- 有界性：
  $0 \le F(x) \le 1, \quad \lim\limits_{x\to-\infty} F(x) = 0 \quad \lim\limits_{x\to+\infty} F(x) = 1$
- 右连续性：$F(x+0) = F(x)$

分布函数计算概率 Page 40
- $P\{a < X \le b\} = F(b) - F(a)$
- $P\{X = a\} = F(a) - F(a-0)$
- $P\{a < X\} = 1 - F(a)$

其它类型随机变量 Page 38
- 分布函数可描述其概率分布
- 结合第一章知识分析

离散型随机变量 Page 41
- 取值有限或可列（数）的随机变量

分布律（列） Page 41
- $P\{X = x_i\} = p_i, i = 1, 2, \cdots$
- 描述离散型 R.V. 概率分布的方法
- 分布律有两个要素：（1）取值；（2）值出现的概率
- 非负性：$\forall i,\ p_i \ge 0$
- 规范性（归一性）：$\sum\limits_{i=1}^{\infty} p_i = 1$

通过分布律计算概率 Page 41
- $P(X \in A \subset R) = \sum\limits_{x_i \in A} P\{X = x_i\}$
- $P(X \le x) = \sum\limits_{x_i \le x} P\{X = x_i\}$

常见离散型随机变量 Page 42、43、45
- 0–1分布、二项分布、泊松分布
- 要熟悉其物理意义

0–1 分布，$B(1,p)$ Page 42
- Bernoulli 试验
- 只考虑两个对立事件的试验
- $P\{X = k\} = p^k(1-p)^{1-k}, k = 0, 1$
- 意义：一次 Bernoulli 试验中事件出现的次数

Poisson 分布，$P(\lambda)$ Page 45
- $P\{X = k\} = \frac{\lambda^k}{k!}e^{-\lambda}, k = 0, 1, 2, \cdots$
- 稀有事件近似服从 Poisson 分布
- 二项分布（n 大，p 小，$\lambda = np$）的近似分布

二项分布，$B(n,p)$ Page 43
- Bernoulli 概型
- $P\{X = k\} = C_n^k p^k (1-p)^{n-k}, k = 0, 1, \cdots, n$
- Bernoulli 概型中同一事件出现的次数

连续型随机变量 Page 48
- 可由概率密度描述的随机变量

连续型随机变量的性质 Page 49
- $F(x)$ 在 R 上连续
- $\forall x_0, P\{X = x_0\} = 0$
- $P\{a < x \le b\} = F(b) - F(a)$
- 有限个或可列个点处的取值不影响概率
- 概率为 0 的事件可能发生
- 概率为 1 的事件有可能不发生

概率密度函数 Page 48
- 描述连续型 R.V. 概率分布的方法
- $F(x) = \int_{-\infty}^{x} f(u)du$
- 非负性：$f(x) \ge 0$
- 规范性（归一性）：$\int_{-\infty}^{+\infty} f(x)dx = 1$
- $f(x)$ 连续点处，$F'(x) = f(x)$

常见连续型随机变量 Page 50
- 均匀分布、指数分布、正态分布
- 均通过概率密度定义

均匀分布，$U(a,b)$ Page 50
- $f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & otherwise \end{cases}$
- 通过计算区间长度比计算概率
- $\forall C = (c, c+\ell) \subset (a,b), P\{x \in C\} = \frac{\ell}{b-a}$
- 能判断随机变量是否服从均匀分布
- 均匀分布是模拟其它分布的基础

均匀分布的分布函数 Page 50
- $F(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \le x < b \\ 1, & b \le x \end{cases}$

指数分布，$EXP(\lambda)$ Page 52
- 通常用来给“寿命”建模
- $f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & otherwise \end{cases}$
- 具有无后效性，也叫无记忆性。

指数分布的分布函数 Page 52
- $F(x) = \begin{cases} 1 - e^{-\lambda x}, & x > 0 \\ 0, & otherwise \end{cases}$

正态分布，$N(\mu, \sigma^2)$ Page 45
- $\varphi(x;\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\{-\frac{(x-\mu)^2}{2\sigma^2}\}, x \in R$
- $\mu \in R, \sigma > 0$
- $X \sim N(0,1)$ 为标准正态分布
- $\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\{-\frac{x^2}{2}\}, x \in R$

正态分布的性质 Page 53
- $\varphi(x;\mu,\sigma^2)$，且具有各阶连续偏导函数
- $\varphi(x;\mu,\sigma^2)$ 在 $(-\infty,\mu)$ 内单调增加，在 $(\mu,+\infty)$ 内单调减少，
- $\varphi(x;\mu,\sigma^2)$ 在 $x = \mu$ 达到最大值 $\frac{1}{\sigma\sqrt{2\pi}}$
- $\varphi(x;\mu,\sigma^2)$ 关于直线 $x = \mu$ 对称
- μ 不影响 $\varphi(x;\mu,\sigma^2)$ 的形状，但会影响位置
- σ 不影响 $\varphi(x;\mu,\sigma^2)$ 的位置，但会影响形状
- σ 越大，$\varphi(x;\mu,\sigma^2)$ 越平坦

上侧分位数 Page 56
- $P\{X > u_\alpha\} = \int_{u_\alpha}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \alpha$
- $\Phi(u_\alpha) = 1 - \alpha$
- $u_\alpha + u_{1-\alpha} = 0$

正态分布的概率计算 Page 54
- 标准正态分布函数 $\Phi(x), x > 0$ 查表可得
- $\Phi(x) = 1 - \Phi(-x), x < 0$
- 一般正态分布算概率要借助于标准正态分布
- $\Phi(x;\mu,\sigma^2) = \Phi(\frac{x-\mu}{\sigma})$
- 随机变量落入任何两个关于 μ 对称的区间的概率总相等

正态分布概率计算一个公式 Page 41
- $P\{|X - \mu| < x\} = 2\Phi(\frac{x}{\sigma}) - 1$
- 结果与 μ 无关

多维随机变量

多维随机变量 Page 64
- $\omega \to (X, Y) \in R^2$
- 样本空间的样本点要用向量描述
- 用随机向量的取值规律刻画样本点分布规律

联合分布函数的性质 Page 65
- 单调不降性 $\begin{cases} \forall x_1 < x_2, \forall y \Rightarrow F(x_1, y) \le F(x_2, y) \\ \forall y_1 < y_2, \forall x \Rightarrow F(x, y_1) \le F(x, y_2) \end{cases}$
- 右连续 $\begin{cases} \lim\limits_{x \to x_0^+} F(x, y) = F(x_0, y) \\ \lim\limits_{y \to y_0^+} F(x, y) = F(x, y_0) \end{cases}$
- 有界性 $0 \le F(x, y) \le 1$, $\lim\limits_{y \to -\infty} F(x, y) = 0$, $\lim\limits_{x \to -\infty} F(x, y) = 0$, $\lim\limits_{\substack{x \to +\infty \\ y \to +\infty}} F(x, y) = 1$,
- 相容性 $\forall x_1 \le x_2, y_1 \le y_2 \Rightarrow F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1) \ge 0$

n 维联合分布 Page 66
- $F(x_1, x_2, \cdots, x_n) = P\{X_1 \le x_1, X_2 \le x_2, \cdots, X_n \le x_n\}$

k 维边缘分布 Page 66
- 任意 k 个分量的联合分布函数
- $F_{X_{i_1}, X_{i_2}, \cdots, X_{i_k}}(x_{i_1}, x_{i_2}, \cdots, x_{i_k}) = F(+\infty, \cdots, X_{i_1}, +\infty, \cdots, X_{i_2}, +\infty, \cdots, X_{i_k}, +\infty, \cdots)$
- 联合分布可以确定边缘分布

Notice:
- 如果 $F_X(x), F_Y(y)$ 连续且严格单调，则下面两个结论成立:
- $X \sim U(0, 1), Z = F_Y^{-1}(X) \Rightarrow F_Z(y) = F_Y(y)$
- $Y = F_X(X) \Rightarrow Y \sim U(0, 1)$

二维离散型随机变量 Page 66
- (X, Y) 取值为有限个或可列个

二维连续型随机变量 Page 69
- 可以由联合概率密度描述的随机变量

联合分布律 Page 66
- $P\{X = x_i, Y = y_j\} = p_{ij}, i, j = 1, 2, \cdots$
- 描述二维离散型 R.V. 概率分布的方法
- 分布律有两个要素:（1）R^2 内取值点;（2）点出现的概率
- 非负性: $\forall i, j,\ p_{ij} \ge 0$
- 规范性（归一性）: $\sum\limits_{i=1}^{\infty} \sum\limits_{j=1}^{\infty} p_{ij} = 1$
- 概率计算: $\forall G \subset R^2, P\{(X, Y) \in G\} = \sum\limits_{(x_i, y_j) \in G} p_{ij}$
- 求联合分布律可能用到条件概率或独立性等第一章知识
- $F(x, y) = \sum\limits_{x_i \le x} \sum\limits_{y_j \le y} p_{ij}$

联合分布函数 Page 64
- $F(x, y) = P\{X \le x, Y \le y\}$
- $\{X \le x, Y \le y\} = \{X \le x\} \cap \{Y \le y\}$
- 刻画随机向量取值分布的重要方法
- 可以刻画随机向量取任何值的概率

联合概率密度 Page 69
- $F(x, y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f(u, v) du dv$
- 非负性: $f(x, y) \ge 0$
- 规范性（归一性）: $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1$
- 概率计算: $\forall G \subset R^2, P\{(X, Y) \in G\} = \iint\limits_{G} f(x, y) dx dy$
- $f(x, y)$ 的连续点处: $\frac{\partial^2 F(x,y)}{\partial x \partial y} = f(x, y)$

二维均匀分布 Page 71
- $f(x, y) = \begin{cases} \frac{1}{S(G)}, & (x, y) \in G \\ 0, & (x, y) \notin G \end{cases}$
- $\forall D \subset G \Rightarrow P\{(X, Y) \in D\} = \frac{S(D)}{S(G)}$
- 借助于长度比、面积比、体积比来算概率的问题叫几何概率

二维正态分布，$U(a, b)$ Page 50
- $f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\} \quad x \in R, y \in R$
- 二维正态分布的边缘分布为正态分布。
- 若 $(X, Y) \sim N(\mu_1, \sigma_1^2; \mu_2, \sigma_2^2; \rho)$ 则 $X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2)$
- 二维正态的条件分布也是正态分布。
- 正态分布的联合概率密度与 ρ 有关.边缘概率密度与 ρ 无关.
- 边缘分布不能唯一确定联合分布.

正态分布具有可加性 Page 91
- $\begin{cases} X\text{与}Y\text{相互独立} \\ X \sim N(\mu_1, \sigma_1^2) \\ Y \sim N(\mu_2, \sigma_2^2) \end{cases} \Rightarrow X \pm Y \sim N(\mu_1 \pm \mu_2, \sigma_1^2 + \sigma_2^2)$

$X + Y$ 的概率密度 Page 90
- $f_Z(z) = \int_{-\infty}^{+\infty} f(z - y, y) dy$
- $f_Z(z) = \int_{-\infty}^{+\infty} f(x, z - x) dx$
- X, Y 相互独立: $f_Z(z) = \int_{-\infty}^{+\infty} f_X(x) f_Y(z - x) dx$
- X, Y 相互独立: $f_Z(z) = \int_{-\infty}^{+\infty} f_X(z - y) f_Y(y) dy$

边缘概率密度 Page 48
- 联合分布可以确定边缘分布
- $f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy$
- $f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx$
- 通过联合求边缘密度实质是含参变量积分
- 画与积分变量对应轴平行的辅助线，讨论边缘概率密度的自变量

边缘分布律 Page 67
- 联合分布可以确定边缘分布
- $P\{X = x_i\} = \sum\limits_{j=1}^{\infty} p_{ij} = p_{i.} \quad i = 1, 2, \cdots$
- $P\{Y = y_j\} = \sum\limits_{i=1}^{\infty} p_{ij} = p_{.j} \quad j = 1, 2, \cdots$

分量之间的关系 Page 52
- 联合分布可以给出分量之间的关系

边缘分布函数 Page 64
- X 的分布函数 $F_X(x) = \lim\limits_{y \to +\infty} F(x, y)$
- Y 的分布函数 $F_Y(y) = \lim\limits_{x \to +\infty} F(x, y)$
- 边缘分布函数完全由联合分布可以确定，仅有边缘分布确定不了联合分布

函数的分布 Page 82
- $Y = g(X)$，Y 的分布
- $Z = G(X, Y)$，Z 的分布

特殊函数的分布 Page 92
- $\max, \min, X + Y, X/Y$

min, max 的分布 Page 90
- $\{\max\{X, Y\} \le z\} \Leftrightarrow \{X \le z\} \cap \{Y \le z\}$
- $\{\min\{X, Y\} \le z\} \Leftrightarrow \{X \le z\} \cup \{Y \le z\}$
- $F_{\max}(z) = F(z, z)$
- $F_{\min}(z) = F_X(z) + F_Y(z) - F(z, z)$
- 相互独立或同分布条件下结论可继续推进
- 前两点最重要

X/Y 的概率密度 Page 92
- $f_Z(z) = \int_{-\infty}^{+\infty} |y| f(zy, y) dy$

条件分布 Page 77
- 条件分布律（离散型）: $P\{X = x_i | Y = y_j\} = \frac{p_{ij}}{p_{.j}} \quad i, j = 1, 2, \cdots$
- 条件概率密度（连续型）: $f_{X|Y}(x|y_0) = \frac{f(x, y_0)}{f_Y(y_0)}$
- 条件分布律与条件概率密度都满足（1）非负性;（2）归一性
- 条件分布函数: $F_{X|Y}(x|y) = \lim\limits_{\Delta y \to 0^+} P\{X \le x | y_0 - \Delta y < Y \le y_0\}$

相互独立 Page 74
- $\forall (x, y) \Rightarrow P\{X \le x, Y \le y\} = P\{X \le x\} P\{Y \le y\}$
- $F(x, y) = F_X(x) F_Y(y)$
- 离散型: $\forall, i, j,\ P_{ij} = P_{i.} P_{.j}$
- 连续型: $f(x, y) = f_X(x) f_Y(y)$ 在平面上除去“面积”为 0 的集合外成立

连续型 R.V. 函数的分布 Page 85
- 先求 $F_Y(y), F_Z(z)$，后求导。
- $F_Y(y) = P\{g(X) \le y\} = \int\limits_{\{x | g(x) \le y\}} f_X(x) dx$
- $F_Z(z) = P\{G(X, Y) \le z\} = \iint\limits_{\{(x, y) | G(x, y) \le z\}} f(x, y) dx dy$
- f 的连续点处: $f_Y(y) = F_Y'(y), f_Z(z) = F_Z'(z)$

Notice:
- 对连续的问题而言 $P\{X < a | Y < b\}$ 与 $P\{X < a | Y = b\}$ 计算方法完全不同
- 前者利用条件概率定义，后者利用条件概率密度

n 维随机变量的独立性 Page 76
- $\forall x_1, x_2, \cdots, x_n \Rightarrow F(x_1, x_2, \cdots, x_n) = \prod\limits_{i=1}^{n} F_i(x_i)$
- $f(x_1, x_2, \cdots, x_n) = \prod\limits_{i=1}^{n} f_i(x_i)$ 几乎处处成立
- $P\{X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n\} = \prod\limits_{i=1}^{n} P\{X_i = x_i\}$

离散型 R.V. 函数的分布 Page 82、83
- $P\{Y = y_j\} = P\{g(X) = y_j\} = \sum\limits_{x_i \in S_j} P_{i.} \quad j = 1, 2, \cdots$

  其中: $S_j = \{x_i | g(x_i) = y_j\}$
- $P\{Z = z_k\} = P\{G(X, Y) = z_k\} = \sum\limits_{(x_i, y_j) \in T_k} P_{ij} \quad k = 1, 2, \cdots$

  其中: $T_k = \{(x_i, y_j) | G(x_i, y_j) = z_k\}$
- 函数取值应通过遍历自变量的函数值得到
- 同一函数值出现的概率要将对应的自变量出现的概率叠加起来

函数概率密度的一个计算公式 Page 87
- 条件: $g'(x) > 0$ 或者 $g'(y) < 0$
- 结论: $f_Y(y) = \begin{cases} f_x[h(y)]|h'(y)| & \alpha < y < \beta \\ 0 otherwise \end{cases}$
- $h(y)$ 是 $y = g(x)$ 的反函数
- $\alpha = min(g(-\infty), g(+\infty)), \beta = max(g(-\infty), g(+\infty))$

正态分布的线性变换 Page 91
- $X \sim N(\mu, \sigma^2) \xrightarrow{a \ne 0} aX + b \sim N(a\mu + b, a^2\sigma^2)$
- $X \sim N(\mu, \sigma^2) \longrightarrow \frac{x-\mu}{\sigma} \sim N(0, 1)$

$X_1, X_2, \cdots, X_n$ 相互独立的性质 Page 77
- $X_1, X_2, \cdots, X_n$ 中任意 k 个随机变量 $(2 \le k \le n)$ 也相互独立。
- 随机变量 $g_1(X_1), g_2(X_2), \cdots, g_n(X_n)$ 也相互独立。
- $(X_1, X_2, \cdots, X_m)$ 与 $(X_{m+1}, X_{m+2}, \cdots, X_n)$ 也相互独立。
- $h(X_1, X_2, \cdots, X_m)$ 与 $g(X_{m+1}, X_{m+2}, \cdots, X_n)$ 也相互独立。

函数分布律的一个计算公式 Page 84
- 条件: X, Y 相互独立，其分布律为:

  $P\{X = k\} = p(k), P\{Y = r\} = q(r) \quad k, r = 0, 1, 2, \cdots$
- 结论: $X + Y$ 的分布律为:

  $P\{X + Y = m\} = \sum\limits_{k=0}^{m} p(k) q(m - k) \quad m = 0, 1, 2, \cdots$

离散型 R.V. 的可加性 Page 85
- $\begin{cases} X\text{与}Y\text{相互独立} \\ X \sim B(n_1, p) \\ Y \sim B(n_2, p) \end{cases} \Rightarrow X + Y \sim B(n_1 + n_2, p)$
- $\begin{cases} X\text{与}Y\text{相互独立} \\ X \sim P(\lambda_1) \\ Y \sim P(\lambda_2) \end{cases} \Rightarrow X + Y \sim P(\lambda_1 + \lambda_2)$

随机变量的数字特征

随机变量的数字特征 Page 103

- 数学期望、方差、协方差、相关系数、矩
- 实践需要知道随机变量平均取值、散布程度以及随机变量间的关联程度
- 数字特征是对这些感兴趣的东西的量化描述

X 的矩 Page 120

- k 阶原点矩：$\gamma^k = E(X^k) \quad k = 1, 2, 3, \cdots$
- k 阶绝对原点矩：$\alpha^k = E(|X|^k) \quad k = 1, 2, 3, \cdots$
- k 阶中心矩：$\mu^k = E\{[X - E(X)]^k\} \quad k = 1, 2, 3, \cdots$
- k 阶绝对中心矩：$\beta^k = E[|X - E(X)|^k] \quad k = 1, 2, 3, \cdots$
- 一阶原点矩是期望，二阶中心矩是方差

常见分布的期望与方差表： Page 113～115

分布类型	期望	方差
$P(\lambda)$	λ	λ
$B(n, p)$	np	$np(1-p)$
$N(\mu, \sigma^2)$	μ	σ^2
0-1分布	p	$p(1-p)$
$U(a, b)$	$(a+b)/2$	$(b-a)^2/12$
$Exp(\lambda)$	$1/\lambda$	$1/\lambda^2$

- 按方差计算公式，$E(X^2) = D(X) + E^2(X)$ 可直得

Notice:

- 如果随机变量既不连续也不离散，则本书并未给出其数字特征计算方法，其计算需要全期望公式。详情参见数学专业参考书
- 两个正态和二维正态不是同一个概念。二维正态的两个边缘分布一定是正态分布，但是两个边缘分布是正态分布的二维随机变量不一定构成二维正态。

数学期望 Page 103

- 离散：$E(X) = \sum_{i=1}^{\infty} x_i p_i$
- 连续：$E(X) = \int_{-\infty}^{+\infty} x f(x) dx$
- 上述级数和积分要求绝对收敛，但因实践中通常满足而不予考虑
- 期望刻画了随机变量分布的中心位置，是随机变量概率意义上的平均值

随机变量的函数的数学期望 Page 105

- 一维离散：$E(g(X)) = \sum_{i=1}^{\infty} g(x_i) p_i$
- 一维连续：$E(g(X)) = \int_{-\infty}^{+\infty} g(x) f(x) dx$
- 二维离散：$E(G(X, Y)) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} G(x_i, y_j) p_{ij}$
- 二维连续：$E(G(X, Y)) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} G(x, y) f(x, y) dx$
- 上述级数和积分要求绝对收敛，但因实践中通常满足而不予考虑

随机变量的数学期望的性质 Page 108

- $E(cX + b) = cE(X) + b$
- $E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i)$
- 若 $X_1, X_2, \cdots, X_n$ 相互独立，则 $E\left(\prod_{i=1}^{n} X_i\right) = \prod_{i=1}^{n} E(X_i)$

方差 Page 110

- $D(X) = E\{[X - E(X)]^2\}$
- 离散：$D(X) = \sum_{i=1}^{\infty} [x_i - E(X)]^2 P\{X = x_i\}$
- 连续：$D(X) = \int_{-\infty}^{+\infty} [x - E(X)]^2 f_X(x) dx$
- $D(X) \geq 0$
- 计算公式：$D(X) = E(X^2) - E^2(X)$
- 方差刻画了随机变量相对于分布中心（期望）的散布程度，方差越小，散布程度就越小
- 标准差是方差的算术方根，也叫均方差，均方差的量纲与随机变量的量纲一致

方差的性质 Page 111

- $D(cX + b) = c^2 D(X)$
- $D\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} D(X_i) + 2 \sum_{1 \leq i < j \leq n} cov(X_i, X_j)$
- 若 $X_1, X_2, \cdots, X_n$ 相互独立，则 $D\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} D(X_i)$
- $D(X) = 0 \iff P\{X = E(X)\} = 1$ 利用切比雪夫不等式

标准化随机变量 Page 112

- $X^* = \frac{X - E(X)}{\sqrt{D(X)}}$ 被称为 X 的标准化随机变量
- 它是一个没有量纲的量

n 维正态随机变量 Page 121

- $\phi(\mathrm{x}_1, \mathrm{x}_2, \cdots, \mathrm{x}_n) = \frac{1}{(2\pi)^{n/2}|\mathrm{C}|^{1/2}} \exp\left\{-\frac{1}{2}(\mathrm{X} - \mu)' \mathrm{C}^{-1} (\mathrm{X} - \mu)\right\}$
- $X = (x_1, x_2, \cdots x_n)' \quad \mu = (\mu_1, \mu_2, \cdots \mu_n)'$
- detC$=0$ 时，无法写出概率密度，分布被称为退化正态或奇异正态
- n 维正态记为：$(X_1, X_2, \cdots, X_n)$ 服从 $N(\mu, C)$

n 维正态随机变量的性质 Page 122

- n维正态分布随机变量的任一 $m(m \leq n)$ 维子向量服从m维正态分布，特别地，$\mathrm{X_i}$ 均为一维正态分布
- n 维随机变量 $(\mathrm{X}_1, \mathrm{X}_2, \cdots .\mathrm{X}_n)$ 相互独立的充要条件是其协方差矩阵为对角阵：

$$C = \begin{bmatrix} C_{11} & & & \\ & C_{22} & & \\ & & \ddots & \\ & & & C_{nn} \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_n^2 \end{bmatrix}$$

- 相互独立的正态随机变量的有限非零线性组合仍服从正态分布.(正态分布具有可加性)
- n维随机变量$(\mathrm{X}_1, \mathrm{X}_2, \cdots .\mathrm{X}_n)$服从正态分布的充分必要条件是：$\mathrm{X}_1, \mathrm{X}_2, \cdots .\mathrm{X}_n$的任意非零线性组合 $\ell_1 \mathrm{X}_1 + \ell_2 \mathrm{X}_2 + \cdots + \ell_n \mathrm{X}_n$ 服从正态分布. $(\ell_1, \ell_2, \cdots ., \ell_n$不全为0)
- 设有矩阵 $B = (b_{ij})_{m \times n}$，随机变量 $(X_1, X_2, \cdots .X_n)$ 的线性变换

$$\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \sim N(B\mu, BCB^T)$$

B 满秩时是非退化的。

协方差、相关系数、矩 Page 115

- 协方差：$cov(X, Y) = E\{[X - E(X)][Y - E(Y)]$
- 常用计算公式：$cov(X, Y) = E(XY) - E(X)E(Y)$
- 协方差用来描述 X, Y 之间的某种关联的程度，但跟随机变量所使用的单位有关

协方差矩阵 Page 116

- $C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix}$
- $c_{ij} = cov(X_i, X_j)$
- $c_{ii} = D(X_i) \quad i = 1, 2, \cdots, n$
- $c_{ij} = c_{ji} \quad i, j = 1, 2, \cdots, n$

协方差的性质 Page 116

- $cov(X, Y) = cov(Y, X)$
- $cov(aX, bY) = abcov(X, Y)$
- $cov(X_1 + X_2, Y) = cov(X_1, Y) + cov(X_2, Y)$
- $cov(X, X) = D(X)$

相关系数 Page 116

- $\rho_{XY} = \frac{cov(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$
- ρ_{XY} 是一无量纲的量
- $\rho_{XY} = E\left[\frac{X - E(X)}{\sqrt{D(X)}} \cdot \frac{Y - E(Y)}{\sqrt{D(Y)}}\right]$
  $= E[X^* \cdot Y^*] = cov(X^*, Y^*) = \rho_{X^*Y^*}$
- ρ_{XY} 是 $cov(X, Y)$ 的标准化处理

相关系数的性质 Page 117

- $|\rho| \leq 1$ 证明用到 $cov(X, Y) = \rho_{X^*Y^*}$
- $|\rho| = 1 \iff X$ 与 Y 依概率为1线性相关。即 $\exists \alpha, \beta(\alpha \neq 0) \quad s.t. \ P\{Y = \alpha X + \beta\} = 1$ 充分性证明用到 $Y, \alpha X + \beta$ 同分布
- 若 $\xi = a_1 X + b_1, \eta = a_2 Y + b_2$ 则 $\rho_{\xi\eta} = \frac{a_1 a_2}{|a_1 a_2|} \rho_{XY}$

相关系数的意义与相关概念 Page 118

- 相关系数是衡量两个随机变量之间线性相关程度的数字特征。
- $\rho_{XY} = 1$ 称 X, Y 正相关
- $\rho_{XY} = -1$ 称 X, Y 负相关
- $\rho_{XY} = 0$ 称 X, Y 不相关，不相关仅指无线性关联
- 相互独立则一定不相关，反之未必成立
- 对于二维正态来讲，不相关$\Leftrightarrow$相互独立$\Leftrightarrow$ $\rho = 0$

大数定律和中心极限定理

Notice:

- 本章的知识点涉及到的数字特征均假定存在

Chebyshev 不等式 Page 130

- $P\{|X-E(X)|\geq\varepsilon\}\leq\frac{D(X)}{\varepsilon^2}$
- $P\{|X-E(X)|<\varepsilon\}>1-\frac{D(X)}{\varepsilon^2}$
- 用数字特征估计概率
- 其估计是很粗略的
- 理论推导

随机变量序列的收敛性 Page 129、130

- 依概率收敛（$X_n\xrightarrow{P}X$）：$\lim\limits_{n\to\infty}P\{|X_n-X|\geq\varepsilon\}=0$
- 依概率收敛是指 n 充分大时，X_n 几乎与 X 等同，但并非必然。
- 依分布收敛（$X_n\xrightarrow{L}X$）：$\lim\limits_{n\to\infty}F_n(x)=F(x)$
- 依分布收敛是指 n 充分大时 $F_n(x)\approx F(x)$，即 $P\{X_n\leq x\}\approx P\{X\leq x\}$。因此可以用 $F(x)$ 近似计算 X_n 有关的概率问题

Notice:

- 大数定律讨论的是序列前 n 项算术均值的取值问题，大数定律成立时，我们几乎确定 $\frac{1}{n}\sum\limits_{k=1}^{n}X_k$ 取什么值
- 中心极限定理讨论的是序列前 n 项和的分布问题，中心极限定律成立时，我们确定 $\sum\limits_{k=1}^{n}X_k$ 具有什么近似分布
- 由于定理成立条件的重合，我们可以同时清楚地知道这两方面的信息

大数定律 Page 130

- 序列 $\{X_n\}$ 服从大数定律是指序列具有性质：
  $$\lim_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i-\frac{1}{n}\sum_{i=1}^{n}E(X_i)\right|<\varepsilon\right\}=1$$
- $\frac{1}{n}\sum\limits_{i=1}^{n}X_i-\frac{1}{n}\sum\limits_{i=1}^{n}E(X_i)\xrightarrow{P}0$
- n 充分大时，（固定 n）序列前 n 项的算术均值几乎就是一个常量

中心极限定理 Page 135

- $X\sim N(0,1)$，$\{X_k\}$，$k=1,2,\cdots$ 相互独立，且数学期望和方差都存在，若 $\dfrac{\sum\limits_{k=1}^{n}X_k-\sum\limits_{k=1}^{n}E(X_k)}{\sqrt{\sum\limits_{k=1}^{n}D(X_k)}}\xrightarrow{L}X$，则称随机变量序列 $\{X_k\}$，$k=1,2,\cdots$ 服从中心极限定理
- $\sum\limits_{k=1}^{n}X_k$ 的标准化随机变量的极限分布为标准正态分布
- 当 n 足够大时，$\dfrac{\sum\limits_{k=1}^{n}X_k-\sum\limits_{k=1}^{n}E(X_k)}{\sqrt{\sum\limits_{k=1}^{n}D(X_k)}}\sim N(0,1)$ 近似成立
- 当 n 足够大时 $\sum\limits_{k=1}^{n}X_k\sim N\left(\sum\limits_{k=1}^{n}E(X_k),\sum\limits_{k=1}^{n}D(X_k)\right)$ 近似成立。
- 理论上给出了现实中大量随机变量服从或近似服从正态分布的原因

切比雪夫大数定律 Page 131

- 条件：$\{X_k\}$ 是相互独立的随机变量序列，其数学期望和方差都存在，且存在一个常数 C，使得 $D(X_k)<C$，$k=1,2,\cdots$
- 结论：$\lim\limits_{n\to\infty}P\left\{\left|\frac{1}{n}\sum\limits_{i=1}^{n}X_i-\frac{1}{n}\sum\limits_{i=1}^{n}E(X_i)\right|<\varepsilon\right\}=1$

辛钦大数定律 Page 131

- 条件：$\{X_k\}$ 是相互独立且同分布的随机变量序列，其 $E(X_k)=\mu$，$k=1,2,\cdots$
- 结论：$\lim\limits_{n\to\infty}P\left\{\left|\frac{1}{n}\sum\limits_{k=1}^{n}X_k-\mu\right|<\varepsilon\right\}=1$
- 成立所需条件较少，适用范围更广

Notice:

- 事件的概率小到什么程度，就能够从实践上认为它是不可能事件呢？这取决于所讨论的事件的重要程度。例如，0.01 是一个不大的数。如果我们有一批炮弹，0.01 是炮弹落地时不爆炸的概率，那么这意味着，大约 1% 的发射是无效的。这是可以容忍的。假如我们有一批降落伞，而 0.01 是跳伞时降落伞不张开的概率。这显然是不能容忍的，因为这意味着，在每一百次跳伞中，就会损失一条生命。这两个例子告诉我们，在考虑一个具体问题时，要根据实际情况，来确定事件的概率如何小，我们就能够于事无损地认为它不可能。

独立同分布大数定律 Page 132

- 条件：$\{X_k\}$ 是相互独立且同分布的随机变量序列，其 $E(X_k)=\mu$，$D(X_k)=\sigma^2$，$k=1,2,\cdots$
- 结论：$\lim\limits_{n\to\infty}P\left\{\left|\frac{1}{n}\sum\limits_{k=1}^{n}X_k-\mu\right|<\varepsilon\right\}=1$
- 定理成立时，n 个随机变量的算术平均，当 n 无限增加时将几乎变成一个常数。

Bernulli 大数定律 Page 133

- 条件：$\frac{m}{n}$ 是 n 次重复独立试验中事件 A 发生的频率，p 是事件 A 在每次试验中发生的概率
- 结论：$\lim\limits_{n\to\infty}P\left\{\left|\frac{m}{n}-p\right|<\varepsilon\right\}=1$
- 此定理证明了随着试验次数的增加，频率有趋于稳定的趋势
- 可得小概率事件原理：概率很小的事件，在一次试验中几乎是不可能发生的，从而在实际中可看成不可能事件

独立同分布中心极限定理 Page 135

- 成立条件：设 $\{X_k\},k=1,2,\cdots$ 是相互独立且同分布的随机变量序列，且 $E(X_k)=\mu$，$D(X_k)=\sigma^2$
- 结论：$\lim\limits_{n\to\infty}P\left\{\dfrac{\sum\limits_{k=1}^{n}X_k-n\mu}{\sqrt{n}\sigma}\leq x\right\}=\Phi(x)$

独立同分布中心极限定理的应用 Page 136

- 求随机变量之和 $S_n=\sum\limits_{k=1}^{n}X_k$ 取值的概率
- 已知 $S_n=\sum\limits_{k=1}^{n}X_k$ 取值的概率，反求 n
- 数理统计中大样本推断的理论基础：设 $\{X_k\},k=1,2,\cdots,n$ 为样本，$E(x)=\mu$，$D(x)=\sigma^2$，则对充分大的 n：$\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$。

棣莫弗–拉普拉斯中心极限定理 Page 137

- 成立条件：随机变量序列 $\{Y_n\},Y_n\sim B(n,p),n=1,2,\cdots$
- 结论：$\lim\limits_{n\to\infty}P\left\{\dfrac{Y_n-np}{\sqrt{np(1-p)}}\leq x\right\}=\Phi(x)$
- 对于足够大的 n，二项分布可被近似成正态分布处理

棣莫弗—拉普拉斯中心极限定理的应用 Page 137

- 计算在大 n 情况下二项分布概率的近似值。
- 已知在大 n 情况下二项分布 R.V. 在某范围内取值的概率，求该范围。
- 近似计算与用频率估计概率的有关问题。

数理统计的基本概念

背景知识 Page 142

- 描述统计学：利用研究对象有关的全部资料，得出全部研究对象某些特征的描述。
- 数理统计学（抽样统计学）：研究怎样用有效的方法去搜集、整理、分析带有随机性影响的数据，并在此基础上对所讨论的问题给出统计性的推断
- 概率作用：度量统计性推断中结论的不确定性。如果总体是知道的，概率论可以用来计算观察到特定样本的概率。如果总体是未知的，但可以得到总体的样本，利用概率论，通过样本信息，可以推断总体构成信息

Notice:

- 数理统计中，有已知信息就尽量用已知信息
- 数理统计中，能用低阶矩就尽量用低阶矩

Notice:

- Gamma函数：$\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1}e^{-x}dx,\ \alpha > 0$
- $\Gamma(1) = 1, \Gamma(\frac{1}{2}) = \sqrt{\pi}, \Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$

总体与个体 Page 142

- 研究对象的全体称为 **总体** ，总体中的每个元素称为 **个体**
- 通常也把研究对象的某个数量指标 X 全体称为 **总体** ，每个数量指标 X_i 称为 **个体**
- 总体被看作是具有一定分布规律的一个随机变量

样本与抽样 Page 142

- **样本** ：数理统计中，用来推断总体信息的具有代表性的一组个体
- 样本是一个 **指标集** ： $X_1, X_2, \cdots, X_n$
- **样本容量** ：样本中个体数量 n
- **抽样** ：从总体中提取样本的过程
- **样本观测值** ：抽样完成后，对样本的观测结果 $x_1, x_2, \cdots, x_n$

简单随机样本 Page 143

- **代表性**：样本中每个个体与总体同分布
- **独立性** 样本中的个体是相互独立的随机变量

统计量 Page 144

- **统计量** ：获取样本后，对得到的信息进行加工和整理的结果，通常表现为样本的函数。
- 对总体的分布规律或数字特征进行推断的基础

常用的统计量 Page 144

- **样本均值** ：$\overline{X} = \frac{1}{n}\sum\limits_{k=1}^{n} X_k$，描述样本分布的中心。与总体均值有密切关联
- **样本方差** ：$S^2 = \frac{1}{n-1}\sum\limits_{k=1}^{n}(X_k - \overline{X})^2$，描述样本的散布程度。与总体方差有密切关联
- S 叫样本标准差

样本矩 Page 144

- 样本 k 阶原点矩：$A_k = \frac{1}{n}\sum\limits_{i=1}^{n} X_i^k$
- 样本 k 阶中心矩：$M_k = \frac{1}{n}\sum\limits_{i=1}^{n}(X_i - \overline{X})^k$
- $A_1 = \overline{X}$
- $M_2 = \frac{n-1}{n}S^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}^2 = A_2 - A_1^2$

χ^2 分布 Page 145

- 构造方法：$\begin{cases} X_1, X_2, \cdots, X_n \text{相互独立} \\ X_i \sim N(0,1),\ i = 1, 2, \cdots, n \end{cases} \Rightarrow \chi^2 = \sum\limits_{i=1}^{n} X_i^2 \sim \chi^2(n)$
- 性质1： $E(\chi^2) = n, D(\chi^2) = 2n$
- 性质2：$\begin{cases} Y_1, Y_2 \text{ 相互独立} \\ Y_1 \sim \chi^2(n_1), Y_2 \sim \chi^2(n_2) \end{cases} \Rightarrow Y_1 + Y_2 \sim \chi^2(n_1 + n_2)$
- 性质3：当 n 足够大(如 $n > 45$)时，有 $\chi^2_\alpha(n) \approx n + u_\alpha\sqrt{2n}$，其中 u_α 满足 $\Phi(u_\alpha) = 1 - \alpha$

t 分布 Page 147

- 构造方法：$\begin{cases} X, Y \text{相互独立} \\ X \sim N(0,1) \\ Y \sim \chi^2(n) \end{cases} \Rightarrow T = \frac{X}{\sqrt{Y/n}} \sim t(n)$
- 性质1：$t_{1-\alpha}(n) = -t_\alpha(n)$
- 性质2：n 较大时($n > 45$)，$t_\alpha(n) \approx u_\alpha$

F 分布 Page 148

- 构造方法：$\begin{cases} X, Y \text{相互独立} \\ X \sim \chi^2(n_1) \\ Y \sim \chi^2(n_2) \end{cases} \Rightarrow F = \frac{X/n_1}{Y/n_2} \sim F(n_1, n_2)$
- F 服从第一自由度为 n_1，第二自由度为 n_2 的 F 分布
- $F \sim F(n_1, n_2) \Leftrightarrow \frac{1}{F} \sim F(n_2, n_1)$
- $F \sim F(n_1, n_2) \Rightarrow F_{1-\alpha}(n_1, n_2) = \frac{1}{F_\alpha(n_2, n_1)}$

抽样分布 Page 145

- 统计量的分布称为抽样分布

单正态总体抽样分布定理 Page 150

- $\overline{X}$与S^2相互独立
- $\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1) \to \sigma$ 已知，推断 μ 时使用
- $\frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t(n-1) \to \sigma$ 未知，推断 μ 时使用
- $\sum\limits_{i=1}^{n}\left(\frac{X_i - \mu}{\sigma}\right)^2 \sim \chi^2(n) \to \mu$ 已知，推断 σ^2 时使用
- $\frac{n-1}{\sigma^2}S^2 \sim \chi^2(n-1) \to \mu$ 未知，推断 σ^2 时使用

双正态总体抽样分布定理 Page 151

- $F = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F(n_1 - 1, n_2 - 1) \to \mu_1, \mu_2$ 未知，推断 σ_2^2/σ_1^2 时使用
- $F = \frac{\sum\limits_{i=1}^{n_1}\left(\frac{X_i-\mu_1}{\sigma_1}\right)^2 \Big/ n_1}{\sum\limits_{j=1}^{n_2}\left(\frac{Y_j-\mu_2}{\sigma_2}\right)^2 \Big/ n_2} = \frac{\sigma_2^2}{\sigma_1^2}\frac{\frac{1}{n_1}\sum\limits_{i=1}^{n_1}(X_i-\mu_1)^2}{\frac{1}{n_2}\sum\limits_{j=1}^{n_2}(Y_j-\mu_2)^2} \sim F(n_1, n_2) \to \mu_1, \mu_2$ 已知，推断 σ_2^2/σ_1^2 时使用
- $U = \frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \sim N(0,1) \to \sigma_1^2, \sigma_2^2$ 已知，推断 $\mu_1 - \mu_2$ 时使用
- 当 $\sigma_1^2 = \sigma_2^2$ 时 $T = \frac{(\overline{X} - \overline{Y}) - (\mu_1 - \mu_2)}{S_w\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2)$ 其中，$S_w = \sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1 + n_2 - 2}} \to \sigma_1^2 = \sigma_2^2$ 未知，推断 $\mu_1 - \mu_2$ 时使用

参数估计

背景知识 Page 156

- 实践中，经常遇到总体分布类型已知，需要推断参数未知的问题
- 分为点估计和区间估计两种方法

参数的点估计 $\hat{\theta}$ Page 156

- 构造统计量 $\hat{\theta}_i(X_1, X_2, \cdots, X_n)$ 估计 θ_i，$\hat{\theta}_i(x_1, x_2, \cdots, x_n)$ 称为 θ 的估计值
- 方法 1：矩估计
- 方法 2：最大似然估计

矩估计 Page 156

- 思想来源：A_k 是 γ_k 的良好估计
- 原理，依大数定律：$A_k \xrightarrow{P} \gamma_k$，即样本容量足够大时，样本矩几乎等于总体矩
- 缺点：总统体矩不存在时失效

矩估计法的一般步骤 Page 156

- 1. 求总体数字特征，一般来讲它们是待估参数的函数
  2. 求解估计方程组：将未知参数用总体数字特征来表示
  3. 将总体数字特征的矩估计代入即可得到未知参数的矩估计
- *注意：一般有几个未知参数就需要建立几个估计方程
- 能用低阶矩就尽量用低阶矩

最大似然估计 Page 158

- Fisher 提出
- 按照最大可能性原则进行推断，即待估参数的取值应该使得样本观测值出现的概率达到最大
- 优点：由于利用了分布信息，估计结果（与矩估计不一样时）一般比矩估计更好

似然函数 Page 159

- 用来描述样本观测值 $x_1, x_2, \cdots, x_n$ 出现的概率
- 离散型随机变量的似然函数是样本联合分布律：
  $L(x_1, x_2, \cdots, x_n; \theta) = \prod_{i=1}^{n} p(x_i; \theta)$
- 连续型随机变量的似然函数是样本联合概率密度：
  $L(x_1, x_2, \cdots, x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta)$

最大似然估计值、量 Page 159

- 估计值：$\hat{\theta} = \arg\max\limits_{\theta} L(x_1, x_2, \cdots, x_n; \theta)$
- 估计量：$\hat{\theta}(X_1, X_2, \cdots, X_n)$

求最大似然估计的一般步骤 Page 160

- 写出似然函数：
  $$L(x_1, x_2, \cdots, x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta_1, \theta_2, \cdots, \theta_m)$$
- 对似然函数取对数：$\ln L = \sum\limits_{i=1}^{n} \ln f(x_i; \theta_1, \theta_2, \cdots, \theta_m)$
- 分别求 $\ln L$ 对 $\theta_j (j = 1, \cdots, m)$ 的偏导，并令其为 0 得似然方程(组)：$\frac{\partial \ln L}{\partial \theta_j} = 0, \quad (j = 1, 2, \cdots, m)$
- 解似然函数方程组得 $\hat{\theta}_1, \hat{\theta}_2, \cdots, \hat{\theta}_m$ 即为所求。

Notice:

- 若不能通过似然函数找驻点的方法求得最大似然估计值，比如似然函数不可导、驻点不存在等情形，就得回归到最大似然估计的定义求解

估计的优良性准则 Page 162

- 估计量不唯一时，需要比较哪个估计量更好
- 三个标准：无偏性、有效性、相合性

无偏性准则 Page 162

- $E(\hat{\theta}) = \theta \Rightarrow \hat{\theta}$ 是 θ 的无偏估计量
- 不是无偏的就是有偏估计
- 无偏估计意味着无系统误差，即 $E(\hat{\theta} - \theta) = 0$

有效性准则 Page 163

- $\begin{cases} E(\hat{\theta}_1) = E(\hat{\theta}_2) = \theta \\ D(\hat{\theta}_1) \le D(\hat{\theta}_2) \end{cases} \Rightarrow \hat{\theta}_1$ 比 $\hat{\theta}_1$ 有效
- $\begin{cases} E(\hat{\theta}_0) = E(\hat{\theta}) = \theta,\ \forall \hat{\theta} \\ D(\hat{\theta}_0) \le D(\hat{\theta}) \end{cases} \Rightarrow \hat{\theta}_0$ 是最小方差无偏估计
- 无偏估计的方差越小，其偏离真值的可能性就越小

相合性准则 Page 164

- $\hat{\theta} \xrightarrow{P} \theta \Rightarrow \hat{\theta}$ 是 θ 的相合估计量
- 相合性保证了：样本容量越大，估计量越精确

参数的区间估计 $[\hat{\theta}_1, \hat{\theta}_2]$ Page 166

- 点估计无法给出估计值的可靠程度
- 区间估计以指明的可靠程度包含估计值
- 精度：区间长度($\hat{\theta}_2 - \hat{\theta}_1$)
- 可靠度：区间包含未知参数的概率($P\{\hat{\theta}_1 \le \theta \le \hat{\theta}_2\} = 1 - \alpha$)
- $[\hat{\theta}_1, \hat{\theta}_2]$ 是 θ 的置信度为 $1 - \alpha$ 的置信区间

枢轴变量法 Page 167

- 点估计是待估参数的估计值，从参数的优良点估计出发构造置信区间
- 优良点估计的变换函数应满足（1）包含待估参数；（2）不包含其它未知参数；（3）变换函数的分布参数完全已知，则变换函数为枢轴变量
- Newman 区间估计原则：（1）先照顾可靠度；（2）然后使精度尽可能高
- 按照 Neyman 提出的原则，利用枢轴变量以及指定的置信度构造置信区间

非正态总体区间估计 Page 173

- 大样本方法构造置信区间依然是根据优良性准则，从点估计出发，利用中心极限定理化成近似正态分布枢轴变量

单侧置信区间 Page 175

- $P\{\hat{\theta}_1(X_1, X_2, \cdots, X_n) \le \theta\} = 1 - \alpha \Rightarrow$
  $[\hat{\theta}_1, +\infty)$ 为参数 θ 的置信度为 $1 - \alpha$ 的单侧置信区间
- $\hat{\theta}_1$ 叫单侧置信下限
- $P\{\theta \le \hat{\theta}_2(X_1, X_2, \cdots, X_n)\} = 1 - \alpha \Rightarrow$
  $(-\infty, \hat{\theta}_2]$ 为参数 θ 的置信度为 $1 - \alpha$ 的单侧置信区间
- $\hat{\theta}_2$ 叫单侧置信上限

单正态总体 μ 的区间估计 Page 168

- σ^2 已知，枢轴变量：$U = \dfrac{\overline{X} - \mu}{\sigma_0/\sqrt{n}} \sim N(0, 1)$

  μ 的置信度为 $1 - \alpha$ 的双侧置信区间：$\left[\overline{X} - \frac{\sigma_0}{\sqrt{n}} u_{\frac{\alpha}{2}}, \overline{X} + \frac{\sigma_0}{\sqrt{n}} u_{\frac{\alpha}{2}}\right]$

  单侧置信区间：$\left[\overline{X} - \frac{\sigma_0}{\sqrt{n}} u_\alpha, +\infty\right)$ 或者 $\left(-\infty, \overline{X} + \frac{\sigma_0}{\sqrt{n}} u_\alpha\right]$
- σ^2 未知，枢轴变量：$T = \dfrac{\overline{X} - \mu}{S/\sqrt{n}} \sim t(n-1)$

  μ 的置信度为 $1 - \alpha$ 的双侧置信区间：$\left[\overline{X} - \frac{S}{\sqrt{n}} t_{\frac{\alpha}{2}}(n-1), \overline{X} + \frac{S}{\sqrt{n}} t_{\frac{\alpha}{2}}(n-1)\right]$

  单侧置信区间：$\left[\overline{X} - \frac{S}{\sqrt{n}} u_\alpha, +\infty\right)$ 或者 $\left(-\infty, \overline{X} + \frac{S}{\sqrt{n}} u_\alpha\right]$

单正态总体 σ^2 的区间估计 Page 168

- μ已知，枢轴变量：$\chi^2 = \sum\limits_{i=1}^{n} \left(\dfrac{X_i - \mu}{\sigma}\right)^2 \sim \chi^2(n)$

  σ^2 的置信度为 $1 - \alpha$ 的双侧置信区间：$\left[\dfrac{\sum_{i=1}^{n}(X_i - \mu)^2}{\chi^2_{\alpha/2}(n)}, \dfrac{\sum_{i=1}^{n}(X_i - \mu)^2}{\chi^2_{1-\alpha/2}(n)}\right]$

  单侧置信区间：$\left[\dfrac{\sum_{i=1}^{n}(X_i - \mu)^2}{\chi^2_{\alpha}(n)}, +\infty\right)$ 或者 $\left(-\infty, \dfrac{\sum_{i=1}^{n}(X_i - \mu)^2}{\chi^2_{1-\alpha}(n)}\right]$
- μ 未知，枢轴变量：$\chi^2 = \dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$

  σ^2 的置信度为 $1 - \alpha$ 的双侧置信区间：$\left[\dfrac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{\chi^2_{\alpha/2}(n-1)}, \dfrac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{\chi^2_{1-\alpha/2}(n-1)}\right]$

  单侧置信区间：$\left[\dfrac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{\chi^2_{\alpha}(n-1)}, +\infty\right)$ 或者 $\left(-\infty, \dfrac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{\chi^2_{1-\alpha}(n-1)}\right]$

双正态总体区间估计 Page 172

- σ_1^2 和 σ_2^2 已知:，枢轴变量：$U = \dfrac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \sim N(0, 1)$

  $\mu_1 - \mu_2$ 的置信度为 $1 - \alpha$ 的双侧置信区间：$\left[\overline{X} - \overline{Y} - u_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}, \overline{X} - \overline{Y} + u_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right]$
- $\sigma_1^2 = \sigma_2^2 = \sigma^2$ 未知，枢轴变量：

  $T = \dfrac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{S_W\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t(n_1 + n_2 - 2), S_W = \sqrt{\dfrac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}$

  $\mu_1 - \mu_2$ 的置信度为 $1 - \alpha$ 的双侧置信区间：

  $\left[\overline{X} - \overline{Y} - t_{\frac{\alpha}{2}}(n_1 + n_2 - 2)S_W\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}, \overline{X} - \overline{Y} + t_{\frac{\alpha}{2}}(n_1 + n_2 - 2)S_W\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right]$
- μ_1, μ_2 已知，枢轴变量：$F = \dfrac{\frac{\sum_{i=1}^{n_1}\left(\frac{X_i-\mu_1}{\sigma_1}\right)^2}{n_1}}{\frac{\sum_{j=1}^{n_2}\left(\frac{Y_j-\mu_2}{\sigma_2}\right)^2}{n_2}} = \dfrac{\sigma_2^2}{\sigma_1^2} \dfrac{\frac{1}{n_1}\sum_{i=1}^{n_1}(X_i-\mu_1)^2}{\frac{1}{n_2}\sum_{j=1}^{n_2}(Y_j-\mu_2)^2} \sim F(n_1, n_2)$

  $\frac{\sigma_2^2}{\sigma_1^2}$ 的置信度为 $1 - \alpha$ 的双侧置信区间：

  $\left[\dfrac{\frac{1}{n_2}\sum_{j=1}^{n_2}(Y_j-\mu_2)^2}{\frac{1}{n_1}\sum_{i=1}^{n_1}(X_i-\mu_1)^2} F_{1-\frac{\alpha}{2}}(n_1, n_2), \dfrac{\frac{1}{n_2}\sum_{j=1}^{n_2}(Y_j-\mu_2)^2}{\frac{1}{n_1}\sum_{i=1}^{n_1}(X_i-\mu_1)^2} F_{\frac{\alpha}{2}}(n_1, n_2)\right]$
- μ_1, μ_2 未知，枢轴变量：$F = \dfrac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} = \dfrac{\sigma_2^2}{\sigma_1^2} \cdot \dfrac{S_1^2}{S_2^2} \sim F(n_1 - 1, n_2 - 1)$

  $\frac{\sigma_2^2}{\sigma_1^2}$ 的置信度为 $1 - \alpha$ 的双侧置信区间：$\left[\frac{S_2^2}{S_1^2} F_{1-\frac{\alpha}{2}}(n_1 - 1, n_2 - 1), \frac{S_2^2}{S_1^2} F_{\frac{\alpha}{2}}(n_1 - 1, n_2 - 1)\right]$

假设检验

背景知识 Page 182

- 仅涉及总体分布的未知参数的统计假设称为 参数假设 ，判断统计假设是否成立的方法称为 假设检验
- 对分布参数 θ 提出假设 H_0，根据样本提供的信息，做出接受还是拒绝 H_0 的决定。
- 第一个假设称为 原假设 或 零假设 ，记为 H_0
- 第二个假设称为 对立假设 或 备择假设 ，记为 H_1
- 检验分为单侧假设检验和双侧假设检验两种形式，检验的具体形式通常由对立假设的形式判断
- 假设检验有分布参数的信息，而区间估计一般没有这类信息

大样本检验法 Page 173

- 利用中心极限定理，近似求得检验统计量的分布
- 根据对立假设形式，利用近似分布确立拒绝域
- 根据样本观测值做出推断

单正态总体均值 μ 的检验 Page 185

- 拒绝域的形式应根据对立假设的形式选择
- U 检验法：σ^2 已知，检验统计量：$U = \dfrac{\overline{X}-\mu}{\sigma_0/\sqrt{n}} \sim N(0,1)$
  - (1)检验 $H_0:\mu=\mu_0$，$H_1:\mu\neq\mu_0$ 的拒绝域 $|u|>u_{\frac{\alpha}{2}}$
  - (2)检验 $H_0:\mu\leq\mu_0$，$H_1:\mu>\mu_0$ 的拒绝域 $u>u_\alpha$
  - (3)检验 $H_0:\mu\geq\mu_0$，$H_1:\mu<\mu_0$ 的拒绝域 $u<u_\alpha$
- T 检验法：σ^2 未知，检验统计量：$T = \dfrac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$
  - (1)检验 $H_0:\mu=\mu_0$，$H_1:\mu\neq\mu_0$ 的拒绝域 $|t|>t_{\frac{\alpha}{2}}(n-1)$
  - (2)检验 $H_0:\mu\leq\mu_0$，$H_1:\mu>\mu_0$ 的拒绝域 $t>t_\alpha(n-1)$
  - (3)检验 $H_0:\mu\geq\mu_0$，$H_1:\mu<\mu_0$ 的拒绝域 $t<t_\alpha(n-1)$

假设检验的基本步骤 Page 184

- 提出原假设 ：根据实际问题提出原假设 H_0 和备择假设 H_1 ；
- 建立检验统计量 ：寻找一个参数的良好估计量，据此建立一个不带任何未知参数的统计量 W 作为检验统计量，并在 H_0 成立的条件下，确定 W 的分布(或近似分布)；
- 确定 H_0 的拒绝域 ：根据实际问题选定显著性水平 α ，依据检验统计量的分布和的 H_0 内容，确定 H_0 的拒绝域；
- 对 H_0 作判断 ：根据样本值算出检验统计量的统计值 w ，判断 w 是否落在拒绝域，以确定拒绝或接受 H_0 。

双正态总体均值差 $\mu_1-\mu_2$ 的检验 Page 187

- U 检验法：σ_1^2 和 σ_2^2 已知，检验统计量：$U = \dfrac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \sim N(0,1)$
  - (1)检验 $H_0:\mu_1=\mu_2$，$H_1:\mu_1\neq\mu_2$ 的拒绝域 $|u|>u_{\frac{\alpha}{2}}$
  - (2)检验 $H_0:\mu_1\leq\mu_2$，$H_1:\mu_1>\mu_2$ 的拒绝域 $u>u_\alpha$
  - (3)检验 $H_0:\mu_1\geq\mu_2$，$H_1:\mu_1<\mu_2$ 的拒绝域 $u<u_\alpha$
- T 检验法：σ_1^2 和 σ_2^2 未知，检验统计量：

$$T = \frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{S_W\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t(n_1+n_2-2),\ S_W=\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}$$

  - (1)检验 $H_0:\mu_1=\mu_2$，$H_1:\mu_1\neq\mu_2$ 的拒绝域 $|t|>t_{\frac{\alpha}{2}}(n_1+n_2-2)$
  - (2)检验 $H_0:\mu_1\leq\mu_2$，$H_1:\mu_1>\mu_2$ 的拒绝域 $t>t_\alpha(n_1+n_2-2)$
  - (3)检验 $H_0:\mu_1\geq\mu_2$，$H_1:\mu_1<\mu_2$ 的拒绝域 $t<t_\alpha(n_1+n_2-2)$

单正态总体方差 σ^2 的检验 Page 185

- 拒绝域的形式应根据对立假设的形式选择
- χ^2 检验法：μ 已知，检验统计量：$\chi^2 = \sum\limits_{i=1}^{n}\left(\dfrac{X_i-\mu}{\sigma}\right)^2 \sim \chi^2(n)$
  - (1)检验 $H_0:\sigma^2=\sigma_0^2$，$H_1:\sigma^2\neq\sigma_0^2$ 的拒绝域 $\chi^2<\chi^2_{1-\frac{\alpha}{2}}(n)$ or $\chi^2>\chi^2_{\frac{\alpha}{2}}(n)$
  - (2)检验 $H_0:\sigma^2\leq\sigma_0^2$，$H_1:\sigma^2>\sigma_0^2$ 的拒绝域 $\chi^2>\chi^2_\alpha(n)$
  - (3)检验 $H_0:\sigma^2\geq\sigma_0^2$，$H_1:\sigma^2<\sigma_0^2$ 的拒绝域 $\chi^2<\chi^2_{1-\alpha}(n)$
- χ^2 检验法：μ 未知，检验统计量：$\chi^2 = \dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$
  - (1)检验 $H_0:\sigma^2=\sigma_0^2$，$H_1:\sigma^2\neq\sigma_0^2$ 的拒绝域 $\chi^2<\chi^2_{1-\frac{\alpha}{2}}(n-1)$ or $\chi^2>\chi^2_{\frac{\alpha}{2}}(n-1)$
  - (2)检验 $H_0:\sigma^2\leq\sigma_0^2$，$H_1:\sigma^2>\sigma_0^2$ 的拒绝域 $\chi^2>\chi^2_\alpha(n-1)$
  - (3)检验 $H_0:\sigma^2\geq\sigma_0^2$，$H_1:\sigma^2<\sigma_0^2$ 的拒绝域 $\chi^2<\chi^2_{1-\alpha}(n-1)$

假设检验的推导依据 Page 184

- 小概率事件的实际推断原理（概率反证法）
- 实际推断原理：实际不可能事件在一次随机试验中是不会出现的
- 实际不可能事件：人为规定一个上界 α，若 $P(A)\leq\alpha$，则实际中认为 A 是不可能事件

假设检验的两类错误 Page 184

- 第一类错误（弃真）：$P\{$拒绝$H_0|H_0$为真$\}=\alpha$
- 第二类错误（纳伪）：$P\{$接受$H_0|H_0$为假$\}$。它跟参数真实值有关
- 实践中，若要同时减少两类错误的概率，则必须增加样本容量，这意味着工作量的加大和经济上的损失
- 样本容量一定时，犯两类错误的概率不能同时降低。假设检验的通常做法是按照 Neyman–Pearson 提出的原则：先控制犯第一类错误的概率 α ，然后再使犯第二类错误的概率尽可能地小

双正态总体方差比 $\frac{\sigma_1^2}{\sigma_2^2}$ 的检验 Page 187

- F 检验法：μ_1,μ_2 已知，检验统计量：

$$F = \frac{\sum\limits_{i=1}^{n_1}\left(\frac{X_i-\mu_1}{\sigma_1}\right)^2 \Big/ n_1}{\sum\limits_{j=1}^{n_2}\left(\frac{Y_j-\mu_2}{\sigma_2}\right)^2 \Big/ n_2} = \frac{\sigma_2^2}{\sigma_1^2}\cdot\frac{\frac{1}{n_1}\sum\limits_{i=1}^{n_1}(X_i-\mu_1)^2}{\frac{1}{n_2}\sum\limits_{j=1}^{n_2}(Y_j-\mu_2)^2} \sim F(n_1,n_2)$$

  - (1)检验 $H_0:\sigma_1^2=\sigma_2^2$，$H_1:\sigma_1^2\neq\sigma_2^2$ 的拒绝域 $F<F_{1-\frac{\alpha}{2}}(n_1,n_2)$ or $F>F_{\frac{\alpha}{2}}(n_1,n_2)$
  - (2)检验 $H_0:\sigma_1^2\leq\sigma_2^2$，$H_1:\sigma_1^2>\sigma_2^2$ 的拒绝域 $F>F_\alpha(n_1,n_2)$
  - (3)检验 $H_0:\sigma_1^2\geq\sigma_2^2$，$H_1:\sigma_1^2<\sigma_2^2$ 的拒绝域 $F<F_{1-\alpha}(n_1,n_2)$
- F 检验法：μ_1,μ_2 未知，检验统计量：$F = \dfrac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} = \dfrac{\sigma_2^2}{\sigma_1^2}\cdot\dfrac{S_1^2}{S_2^2} \sim F(n_1-1,n_2-1)$
  - (1)检验 $H_0:\sigma_1^2=\sigma_2^2$，$H_1:\sigma_1^2\neq\sigma_2^2$ 的拒绝域 $F<F_{1-\frac{\alpha}{2}}(n_1-1,n_2-1)$ or $F>F_{\frac{\alpha}{2}}(n_1-1,n_2-1)$
  - (2)检验 $H_0:\sigma_1^2\leq\sigma_2^2$，$H_1:\sigma_1^2>\sigma_2^2$ 的拒绝域 $F>F_\alpha(n_1-1,n_2-1)$
  - (3)检验 $H_0:\sigma_1^2\geq\sigma_2^2$，$H_1:\sigma_1^2<\sigma_2^2$ 的拒绝域 $F<F_{1-\alpha}(n_1-1,n_2-1)$

回归分析

背景知识 Page 205

- 事物之间的关系 $\xrightarrow{\text{量化}}$ 变量之间的关系
- 相关关系：变量之间存在关联，但没有达到互相确定的程度
- 回归分析：找出 相关关系 中变量之间的近似关系
- 在回归分析中：因变量被看作随机变量,自变量则是可控制的!

回归模型 Page 205

- $Y = \mu(x_1, x_2, \cdots, x_k) + \varepsilon, \varepsilon \sim N(0, \sigma^2)$
  $\mu(x_1, x_2, \cdots, x_k)$ 被称为 Y 对 $(X_1, X_2, \cdots, X_k)$的回归函数
  $y = \mu(x_1, x_2, \cdots, x_k)$ 被称为回归方程
- 一元回归模型：$Y = \mu(x) + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2)$
  一元回归方程为：$y = \mu(x)$

回归分析涉及的主要问题： Page 206

- 建立模型(找出自变量与因变量)
- 确定回归函数 μ 的类型
- 估计参数
- 检验模型的合理性

一元正态线性回归模型 Page 207

- $Y = a + bx + \varepsilon, \varepsilon \sim N(0, \sigma^2)$
- 其中，a 被称为回归常数(又称截距)，b 被称为回归系数(又称斜率)，ε 被称为随机扰动项
- 对于任意一组样本，有 $Y_i = a + bx_i + \varepsilon_i, \quad i = 1, \cdots, n$

  1. 各次试验相互独立
  2. $E(\varepsilon_i) = 0$
  3. $D(\varepsilon_i) = \sigma^2$
  4. $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$相互独立

  从而，$Y_i \sim N(a + bx_i, \sigma^2)$，且相互独立

一元正态线性回归模型参数估计 Page 208

- $\hat{y}_i = \hat{a} + \hat{b}x_i$ 为 Y_i 的估计值，y_i 为 Y_i 的观测值，偏离量 $y_i - \hat{y}_i$ 称为在 $X = x_i$ 处的残差
- 将使得残差平方和 $\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{n}(y_i - \hat{a} - \hat{b}x_i)^2$ 取最小值的 $\hat{a}$、$\hat{b}$ 做为 a、b 的估计值
- 使得残差平方和达到最小来估计未知参数的方法称为最小二乘法，估计出的参数叫参数的最小二乘估计值
- $\hat{a}$ 被称为经验回归常数，$\hat{b}$ 被称为经验回归系数

一元正态线性回归模型参数 a、b 和 σ^2 的最小二乘估计 Page 209

- $\hat{b} = \dfrac{\ell_{xy}}{\ell_{xx}}$
- $\hat{a} = \overline{y} - \hat{b}\overline{x}$
- $\hat{\sigma}^2 = \dfrac{1}{n-2}\sum_{i=1}^{n}\varepsilon_i^2 = \dfrac{1}{n-2}(\ell_{yy} - \hat{b}^2\ell_{xx})$
- 记号：$\ell_{xy} = \sum(x_i - \overline{x})(y_i - \overline{y}) = \sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}$

  $\ell_{xx} = \sum(x_i - \overline{x})^2 = \sum_{i=1}^{n} x_i^2 - n\bar{x}^2$

  $\ell_{yy} = \sum(y_i - \overline{y})^2 = \sum_{i=1}^{n} y_i^2 - n\bar{y}^2$

非线性回归问题的线性化处理 Page 223

- 有些非线性问题可以通过变量替换转化成线性问题
- 检验非线性的显著性可以转化成检验替换后变量间关系的线性显著性

常见的几种可以转化成线性问题的非线性问题 Page 225

- 双曲线函数：$\frac{1}{y} = a + \frac{b}{x}$
- 幂函数：$y = ax^b \quad (x > 0)$
- 对数函数：$y = a + b\ln x \quad (x > 0)$
- 指数函数：$y = ae^{bx}$
- 倒数指数函数：$y = ae^{\frac{b}{x}} \quad (x > 0)$
- S 型曲线：$y = \frac{1}{a+be^{-x}}$

一元线性回归的显著性检验：相关系数检验法 Page 214

- 相关系数检验法：是基于试验数据，检验变量间线性相关关系是否显著的一种方法
- 相关系数

$$\rho_{XY} = \frac{E\{[X - E(X)][Y - E(Y)]\}}{\sqrt{D(X)}\sqrt{D(Y)}}$$

  是表征随机变量 Y 与 X 的线性相关程度的数字特征。
- 样本相关系数：$\hat{\rho}_{XY} = R = \dfrac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\bar{y})^2}} = \dfrac{\ell_{xy}}{\sqrt{\ell_{xx}}\sqrt{\ell_{yy}}}$
- $|R|$ 越接近于 1，X 与 Y 间的线性相关关系越显著；
- $|R|$ 越靠近于 0，X 与 Y 间的线性相关关系越不显著。

一元线性回归的显著性检验：F 检验法 Page 213

- 原假设：$H_0 : b = 0$，如果拒绝 H_0 则自变量 X 和因变量 Y 之间线性相关关系显著；否则，认为两者之间无明显的线性相关关系
- 当原假设成立时，检验统计量 $F = \frac{E_R}{E_E/(n-2)} \sim F(1, n-2)$
- 拒绝域：$f > F_\alpha(1, n-2)$
- 残差平方和：$Q_E = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2, Q_R = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$，它反映了随机因素对 Y 的观测值造成的离散程度
- 回归离差平方和：$Q_R = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$，它反映了观测值 $\hat{y}_1, \hat{y}_2, \cdots, \hat{y}_n$ 的离散程度
- 总离差平方和：$Q_T = \sum_{i=1}^{n}(y_i - \bar{y})^2 = Q_E + Q_R$，它反映了观测值 $y_1, y_2, \cdots, y_n$ 的总的离散程度

Notice:

- 从 $|R| = 1$ 可严格推出观测到的散点精确分布在一条直线上