1st Adition

交 大 学 辅

Probability and Mathematical Statistics

概 率 论 与 数 理 统 计

（期末版）

（第一版）

Contributers

仲英书院学业辅导中心

钱学森书院学业辅导中心

文治学辅与发展中心

启德书院朋辈导师团

SEP

17

WHAT WE WANT TO CONVEY

希望这份资料能够对各位的概率论学习有所帮助，解决学习是遇到的问题，巩固课内知识。

祝愿大家在期末取得佳绩！

——2021.2.28

内容简介

本资料就概率论期末考试内容进行了全面详细的总结，并列举了典型例题，结合知识分析题目，以帮助同学们夯实基础，准备考试。

感谢校学生会、仲英学辅、钱院学辅、文治学辅及启德学辅各位工作人员与志愿者的努力工作，使本资料可以顺利完工。由于编者们时间精力限制，难免有错误之处。如果同学们在本资料中发现错误，请联系交小权QQ：3171478351进行反馈，我们将会在下一版进行修改。

资料版次及编者信息：

2021年3月 第一版

编者：

仲英书院学业辅导中心：软件73 田丰瑞 力学理81 叶义晨 金禾91 李凌蕴

钱学森书院学业辅导中心：人工智能91 常瀚之 化生91 干金群

文治书院学业辅导中心：自动化86 张亦斌

排版：

仲英书院学业辅导中心：电气钱92 曾小于 软件73 田丰瑞

启德书院学业辅导中心：经济92 王悦帆

文治书院学业辅导中心：自动化86 张亦斌

材料名称 概率论学业辅导材料（第1版）期末版

编 委 会 西安交通大学学生会学术资讯部

启德书院学业辅导中心、钱学森书院学业辅导中心、文治书院学业辅导中心

仲英书院学业辅导中心

总 页 数 45页

字 数 24000字

Email xjtuxszx@163.com

目录

Part I

随机事件与概率

1 随机事件

1.1 随机现象与随机试验

定义 1.1.1. 随机现象：任何现象分为必然现象、随机现象（具有统计规律性）。

概率论与数理统计是研究和揭示随机现象的统计规律性的学科。

定义 1.1.2. 随机试验（试验），用 E 表示，有以下特征：

1. 可以在相同条件下重复进行
2. 每次试验的可能结果不止一个，且能事先明确试验的所有可能结果
3. 进行一次试验之前不能确定哪一个结果会出现

1.2 样本空间与随机事件

定义 1.2.1. 样本空间：随机试验 E 的所有可能结果组成的集合称为 E 的样本空间，用 Ω 表示。

定义 1.2.2. 基本结果（样本点）：样本空间的元素，即 E 的每个结果称为基本结果（样本点），用 ω 表示。

定义 1.2.3. 随机事件（事件）：在试验中，可能发生、也可能不发生的事情叫做随机事件（事件），用 A、B、C 等表示，当且仅当属于该随机事件的某个样本点在试验中出现，称该事件发生。

任何事件分为三类：

1. 必然事件：用 Ω 表示
2. 不可能事件：用 $\varnothing$ 表示
3. 随机事件：用 A、B、C 等表示

1.3 事件的关系与运算

定义 1.3.1. 事件的包含: 用 $A \subset B$ 表示

定义 1.3.2. 事件的相等:$A = B$ 表示

定义 1.3.3. 事件的和（并）：用 $A \cup B$ 表示

定义 1.3.4. 事件的积（交）：用 AB 表示

定义 1.3.5. 事件的互斥: 用 $AB = \varnothing$ 表示

定义 1.3.6. 事件的对立: 用 $\overline{A}$ 表示

定义 1.3.7. 事件的差: 用 $A - B = A\overline{B}$ 表示

定义 1.3.8. 交换律：$A \cup B = B \cup A$

定义 1.3.9. 结合律：$(A \cup B) \cup C = A \cup (B \cup C)$，$(AB)C = A(BC)$

定义 1.3.10. 分配律：$(A \cup B)C = AC \cup BC$，$(AB) \cup C = (A \cup C)(B \cup C)$

定义 1.3.11. 对偶律：$\overline{A} = \overline{A}\ \overline{B}$，$\overline{AB} = \overline{A} \cup \overline{B}$

运算的优先顺序为：补运算为先，积运算其次，和、差运算最后，若有括号，则括号内的运算优先。

2 概率

2.1 定义

定义 2.1.1. 古典定义，包括古典概型和几何概型，每次试验具备等可能性。

超几何概型：从 a 件正品 b 件次品无放回抽取 n 件，恰好抽出 k 件正品的概率为

$$p_k = \frac{C_a^k C_b^{n-k}}{C_{a+b}^n}$$

其中 $k \le a$，$n - k \le b$。

性质：

1.$P(A) \ge 0$

2.$P(\Omega) = 1$

3. 当 A_1，A_2，$\cdots$ 两两互斥，则 $P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots$。

定义 2.1.2. 统计定义:

频数与频率：相同条件下，重复进行 n 次试验，n 次试验中事件 A 发生的次数 n_A 称为事件 A 发生的频数。n_A/n 称为事件 A 发生的频率，记作 $f_n(A)$，$f_n(A) = n_A/n$。

性质：

1.$f_n(A) \ge 0$

2.$f_n(\Omega) = 1$

3. 当 A_1，A_2，$\cdots$，A_k 两两互斥，则

$$f_n(A_1 \cup A_2 \cup \cdots \cup A_k) = f_n(A_1) + f_n(A_2) + \cdots + f_n(A_k)$$

当试验的重复次数 n 充分大时，若事件 A 发生的频率 $f_n(A)$ 稳定地在某数 p 附近摆动，则称 p 为事件 A 的概率，记为 $P(A) = p$。

统计定义不要求具备等可能性，可以用来估算概率，这样定义的概率也有频率的性质。

定义 2.1.3. 公理化定义:

1. 非负性：$P(A) \geq 0$

2. 规范性：$P(\Omega) = 1$

3. 可列可加性：当 $A_iA_j = \varnothing(i \neq j$；$i$，$j = 1$，$2$，$\cdots)$，有

$P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots$。

$f_n(A_1 \cup A_2 \cup \cdots \cup A_k) = f_n(A_1) + f_n(A_2) + \cdots + f_n(A_k)$

2.2 概率的性质

1. 基本性质：

$P(A) \geq 0$, $P(\Omega) = 1$,

当 A_1，A_2，$\cdots$ 两两互斥，则

$$P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots$$

2.$P(\varnothing) = 0$

3.（概率的加法定理）：当 A_1，A_2，$\cdots$，A_n 两两互斥，则

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n)$$

4. 若 $A \subset B$，$P(B) \geq P(A)$

5.$P(\overline{A}) = 1 - P(A)$

6.$P(A \cup B) = P(A) + P(B) - P(AB)$

推广至 n 个事件，有

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{i=1}^{n} P(A_i) - \sum_{1 \leq i < j \leq n} P(A_iA_j) + \sum_{1 \leq i < j < k \leq n} P(A_iA_jA_k) + \cdots + (-1)^n P(A_1A_2 \cdots A_n)$$

7.（概率的连续性）：若 $A_n \subset A_{n+1}$，$n=1$，2，$\cdots$，令 $A=\bigcup_{i=1}^{\infty} A_i$，则

$$P(A)=\lim_{n\to\infty} P(A_n)$$

8. 推论（概率的连续性）：若 $A_n \supset A_{n+1}$，$n=1$，$2,\cdots$，令 $A=\bigcap_{i=1}^{\infty} A_i$，则

$$P(A)=\lim_{n\to\infty} P(A_n)$$

3 条件概率、全概率公式与贝叶斯公式

3.1 条件概率与乘法公式

定义 3.1.1. 条件概率：

若 A、B 为两个事件，且 $P(B)>0$，称

$$P(A \mid B)=\frac{P(AB)}{P(B)}$$

为事件 B 发生的条件下事件 A 发生的条件概率。

定义 3.1.2. 乘法公式：

若 $P(B)>0$，则

$$P(AB)=P(B)P(A \mid B)$$

推广至 n 个事件 A_1，A_2，$\cdots$，A_n，若 $P(A_1A_2\cdots A_{n-1})>0$，则

$$P(A_1A_2\cdots A_n)=P(A_1)P(A_2 \mid A_1)\cdots P(A_n \mid A_1A_2\cdots A_{n-1})$$

3.2 全概率公式与贝叶斯公式

定义 3.2.1. 互斥完备事件群（互不相容完备事件组）：

设 B_1，B_2，$\cdots$，B_n，$\cdots$ 为有限多个或无限多个事件，若它们满足

1.$B_iB_j=\varnothing(i\neq j$；$i$，$j=1$，$2$，$\cdots)$；

2.$B_1\cup B_2\cup\cdots B_n\cup\cdots=\Omega$；

则称这组事件为互斥完备事件群（互不相容完备事件组）。

定义 3.2.2. 全概率公式：

若事件 B_1，B_2，$\cdots$，B_n，$\cdots$ 构成互斥完备事件群，且 $P(B_i) > 0(i = 1, 2, \cdots)$，则对任一事件 A 有

$$P(A) = \sum_j P(B_j)P(A \mid B_j)$$

定义 3.2.3. 贝叶斯公式：

若事件 B_1，B_2，$\cdots$，B_n，$\cdots$ 构成互斥完备事件群，且 $P(B_i) > 0(i = 1, 2, \cdots)$，则对任一事件 A 有

$$P(B_i|A) = \frac{P(B_i)P(A \mid B_i)}{\sum\limits_j P(B_j)P(A \mid B_j)}$$

其中，称 $P(B_1)$，$P(B_2)$，$\cdots$ 为先验概率，称 $P(B_1 \mid A)$，$P(B_2 \mid A)$，$\cdots$ 为后验概率。

4　随机事件的独立性

4.1　两个事件的独立性

定义 4.1.1. 两个事件的独立性：

设 A，B 为两个事件，若有

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

，则称 A 与 B 相互独立（独立）。

定理 4.1.2. 两个事件的独立性：

若四对事件 $\{A, B\}$，$\{A, \overline{B}\}$，$\{\overline{A}, B\}$，$\{\overline{A}, \overline{B}\}$ 中有一对是相互独立的，则另外三对也是相互独立的。

4.2　多个事件的独立性

定义 4.2.1. 多个事件独立：

设 A_1，A_2，$\cdots$，A_n 为 n 个事件，若对任意 $k(1 < k \leq n)$，任意 $1 \leq i_1 < i_2 < \cdots < i_k \leq n)$，都成立

$$P(A_{i_1}A_{i_2}\cdots A_{i_k}) = P(A_{i_1})P(A_{i_2})\cdots P(A_{i_k})$$

则称 A_1，A_2，$\cdots$，A_n 相互独立。

上式中包含的等式总数为 $2^n - n - 1$。

两两独立为上式中 $k = 2$ 的特殊情况，此时它们中的任意两件事件都相互独立。

两两独立并不能保证相互独立。

事件的相互独立与事件的互斥是两个不同的概念：

1. 当 A_1，A_2，$\cdots$，A_n 两两互斥时，事件的互斥性可简化为和事件的概率计算，即

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n)$$

2. 当 A_1，A_2，$\cdots$，A_n 相互独立时，事件的独立性可简化为积事件的概率计算，即

$$P(A_1A_2\cdots A_n) = P(A_1)P(A_2)\cdots P(A_n)$$

并且，$P(\bigcup\limits_{k=1}^{n} A_k) = 1 - P(\bigcap\limits_{k=1}^{n} \overline{A_k}) = 1 - \prod\limits_{k=1}^{n} P(\overline{A_k}) = 1 - \prod\limits_{k=1}^{n} [1 - P(A_k)]$。

定理 4.2.2. 多个事件的独立性：

若 n 个事件 A_1，A_2，$\cdots$，A_n 相互独立，则把其中任意 $m(1 \le m \le n)$ 个事件相应的换成他们的对立事件，所得的 n 个事件仍为相互独立。

4.3 随机试验的相互独立性

定义 4.3.1. 相互独立：

设有试验 E_1，E_2，若试验 E_1 的任一结果（事件）与试验 E_2 的任一结果（事件）都是相互独立的，则称这两个试验相互独立。

推广至 n 个试验，若试验 E_1 的任一结果，试验 E_2 的任一结果，$\cdots$，试验 E_n 的任一结果都是相互独立的，则称这 n 个试验相互独立。

定义 4.3.2. n 重独立试验。

当这 n 个试验相互独立且相同时，称其为 n 重独立试验。

定义 4.3.3. n 重伯努利试验：

如果在 n 重独立试验中，每次试验的结果为两个，即 A 或 $\overline{A}$，称这种试验为 n 重伯努利试验。

在 n 重伯努利试验中，若事件 A 在每次试验中发生的概率均为 $P(A) = p(0 < p < 1)$，记 B_{nk} 表示 n 重伯努利试验中事件 A 发生的 k 次这一事件，则 $P(B_{nk}) = C_n^k p^k (1-p)^{n-k}$，$k = 0$，$1$，$\cdots$，$n$。

（注：本章的典型例题有 13 页的例 1.2.6，例 1.2.7，例 1.3.2，例 1.4.3，以及习题 1 中的 5，7，9，13，19，30，35。）

Part II

随机变量及其分布

5 一维随机变量

5.1 随机变量

定义 5.1.1. 设 E 为一随机试验，Ω 为其样本空间，若 $X = X(\omega), \omega \in \Omega$ 为单值实函数，且对于任意实数 x，集合 $\{\omega|X(\omega) \le x\}$ 都是随机事件，则称 X 为**随机变量**。随机变量经常用 X、Y、Z 等表示。

5.2 离散型随机变量及其分布律

定义 5.2.1. 随机变量：

设 E 为一随机试验，Ω 为其样本空间，若 $X = X(\omega), \omega \in \Omega$ 为单值实函数，随机变量且对于任意实数 x，集合 $\{\omega|X(\omega) \le x\}$ 都是随机事件，则称 X 为随机变量。随机变量经常用 X、Y、Z 等表示。

定义 5.2.2. 离散型随机变量：

随机变量的全部可能的取值是有限个或可列无穷多个，这种随机变量称为离散型随机变量。

设离散型随机变量 X 的所有可能取的值为 $x_k(k = 1, 2, \cdots)$，X 取各个可能值的概率，即事件 $\{X = x_k\}$ 的概率，为

$$P\{X = x_k\} = p_k, \quad k = 1, 2, \cdots.$$

由概率的定义，p_k 满足如下两个条件：

1.$p_k \ge 0, k = 1, 2, \cdots$

2.$\sum_{k=1}^{\infty} p_k = 1$

5.2.1 (2) 几种重要的离散型随机变量

定义 5.2.3. 0-1 分布：

设随机变量 X 服从 0-1 分布，则其只可能取 0 与 1 两个值，它的分布律是

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

定义 5.2.4. 二项分布 $B(n, p)$：

设试验 E 只有两个可能结果：A 及 $\bar{A}$，则称 E 为伯努利 (Bernoulli) 试验。设 $P(A)=p(0<p<1)$，此时 $P(\bar{A})=q=1-p$。将 E 重复地进行 n 次，则称这一串重复的独立试验为 n 重伯努利试验。

n 次伯努利试验中事件 A 发生的次数为

$$P\{X=k\}=\binom{n}{k}p^kq^{n-k},\quad k=0,1,2,\cdots,n.$$

则称 X 服从二项分布，记为 $X\sim B(n,p)$。

定义 5.2.5. 泊松分布 $P(\lambda)$：

体积相对小的物质在较大空间的稀疏分布，都可以看作泊松分布，参数 λ 可以用观测值的平均值求出。

设随机变量 X 所有可能取的值为 $0,1,2,\cdots$，而取各个值的概率为

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

其中 $\lambda>0$ 是常数。则称 X 服从参数为 λ 的泊松 (Poisson) 分布，记为 $X\sim\pi(\lambda)$。

定理 5.2.6. 泊松定理：

设 $\lambda>0$ 是一个常数，n 是任意正整数，设 $np_n\to\lambda$，则对于任一固定的非负整数 k，有

$$\lim_{n\to\infty}\binom{n}{k}p_n^k(1-p_n)^{n-k}=\frac{\lambda^k\mathrm{e}^{-\lambda}}{k!}.$$

定义 5.2.7. * 几何分布 $G(p)$：

设 X 服从参数为 p 的几何分布，也即 $X\sim G(p)$，则有

$$P\{X=k\}=p(1-p)^{k-1},\quad k=1,2,3\cdots.$$

定义 5.2.8. * 帕斯卡分布 $P(r,p)$：

设 X 服从参数为 r 和 p 的帕斯卡 (Pascal) 分布 $X\sim P(r,p)$，则有

$$P\{X=k\}=\binom{k-1}{r-1}p^r(1-p)^{k-r},\quad k=r,r+1,\cdots.$$

定义 5.2.9. * 超几何分布 $H(n,N,M)$：

设 X 服从参数为 n、N 和 M 的超几何分布 $X \sim H(n, N, M)$，则有

$$P\{X = k\} = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}, \quad , k = 0, 1, 2, \cdots, \min\{n, M\}.$$

5.3 随机变量的分布函数

对于非离散型随机变量 X，其可能的取值不能一一列举出来，因而不能像离散型随机变量那样用分布律描述。非离散型随机变量取任意一指定实数值的概率等于 0。

定义 5.3.1. 分布函数:

设 X 是一个随机变量，x 是任意实数，函数

$$F(x) = P\{X \le x\}, \quad -\infty < x < \infty$$

称为 X 的分布函数。

对于任意实数 $x_1, x_2 (x_1 < x_2)$，有

$$P\{x_1 < X \le x_2\} = P\{X \le x_2\} - P\{X \le x_1\} = F(x_2) - F(x_1).$$

基本性质:

1.$F(x)$ 是一个不减函数。

事实上，由上式，对于任意实数 $x_1, x_2 (x_1 < x_2)$，有

$$F(x_2) - F(x_1) = P\{x_1 < X \le x_2\} \ge 0.$$

2.$0 \le F(x) \le 1$，且

$$F(-\infty) = \lim_{x \to -\infty} F(x) = 0, F(\infty) = \lim_{x \to \infty} F(x) = 1.$$

3.$F(x + 0) = F(x)$，即 $F(x)$ 是右连续的。

注：如果一个函数满足上述性质，那么该函数一定是某个随机变量的分布函数。

5.4 连续型随机变量极其概率密度

定义 5.4.1. 连续型随机变量:

如果对于随机变量 X 的分布函数 $F(x)$，存在非负函数 $f(x)$，使对于任意实数 x，有

$$F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t,$$

则称 X 为连续性随机变量，其中函数 $f(x)$ 称为 X 的概率密度函数，简称概率密度。

性质:

1.$f(x) \geq 0$

2. $\int_{-\infty}^{+\infty} f(x)\mathrm{d}x = 1$

3. 对于任意实数 $x_1, x_2(x_1 < x_2)$，

$$P\{x_1 < X \leq x_2\} = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(x)\mathrm{d}x;$$

4. 若 $f(x)$ 在点 x 处连续，则有 $F'(x) = f(x)$

5.4.1 重要的连续型随机变量

定义 5.4.2. 均匀分布 $U(a,b)$:

若连续型随机变量 X 具有概率密度

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

则称 X 在区间 (a,b) 上服从均匀分布，记为 $X \sim U(a,b)$。

四舍五入精确到第 k 位小数，所产生随机误差服从均匀分布 $U(-\frac{1}{2}10^{-k}, \frac{1}{2}10^{-k})$。

定义 5.4.3. 指数分布 $\exp(\theta)$:

若连续型随机变量 X 具有概率密度

$$f(x) = \begin{cases} \dfrac{1}{\theta}\mathrm{e}^{-\frac{x}{\theta}}, & x > 0, \\ 0, & \text{其他}, \end{cases}$$

其中 $\theta > 0$ 为常数，则称 X 服从参数为 θ 的指数分布。

指数分布具有无记忆性，也即 $P\{X > s + t|X > s\} = P\{X > t\}$。

定义 5.4.4. 正态分布（高斯分布）$N(\mu, \sigma)$:

若连续型随机变量 X 的概率密度为

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

其中 $\mu,\sigma(\sigma>0)$ 为常数，则称 X 服从参数为 μ,σ 的正态分布或高斯 (Gauss) 分布，记为 $X\sim N(\mu,\sigma^2)$。

性质:

曲线关于 $x=\mu$ 对称。这表明对于任意 $h>0$ 有

$$P\{\mu-h<x\leq\mu\}=P\{\mu<x\leq\mu+h\}.$$

当 $x=\mu$ 时取到最大值 $f(\mu)=\dfrac{1}{\sqrt{2\pi}\sigma}$。$x$ 离 μ 越远，$f(x)$ 的值越小。这表明对于同样长度的区间，当区间离 μ 越远，X 落在这个区间上的概率越小。

在 $x=\mu\pm\sigma$ 处曲线有拐点。曲线以 Ox 轴为渐近线。

定义 5.4.5. 标准正态分布:

特别，当 $\mu=0,\sigma=1$ 时称随机变量 X 服从标准正态分布，其概率密度和分布函数分别用 $\varphi(x),\Phi(x)$ 表示，即有

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}},\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t.$$

易知 $\Phi(-x)=1-\Phi(x)$。

定理 5.4.6. 正态分布:

若 $X\sim N(\mu,\sigma^2)$，则 $Z=\dfrac{X_\mu}{\sigma}\sim N(0,1)$。

于是，若 $X\sim N(\mu,\sigma^2)$，则它的分布函数 $F(x)$ 可写成

$$F(x)=P\{X\leq x\}=P\left\{\frac{X-\mu}{\sigma}\leq\frac{x-\mu}{\sigma}\right\}=\Phi\left(\frac{x-\mu}{\sigma}\right).$$

对于任意区间 $(x_1,x_2]$，有

$$P\{x_1<X\leq x_2\}=P\left\{\frac{x_1-\mu}{\sigma}<\frac{X_\mu}{\sigma}\leq\frac{x_2-\mu}{\sigma}\right\}=\Phi\left(\frac{x_2-\mu}{\sigma}\right)-\Phi\left(\frac{x_1-\mu}{\sigma}\right).$$

6 多维随机变量及其应用

6.1 n 维随机变量（向量）

定义 6.1.1. n 维随机变量:

设 E 是随机试验，$\Omega=\{\omega\}$ 是 E 的样本空间，而 $X_1(\omega),X_2(\omega),\cdots,X_n(\omega)$ 是定义在 Ω 上的 n 个随机变量，则 n 维向量 $(X_1(\omega),X_2(\omega),\cdots,X_n(\omega))$ 称为 n 维随机变量或 n 维随机向量，通常简记为 $(X_1,X_2,\cdots,X_n)$。

对于任意 n 个实数 $x_1,x_2,\cdots,x_n$，n 元函数

$$F(x_1,x_2,\cdots,x_n)=P\{X_1\le x_1,X_2\le x_2,\cdots,X_n\le x_n\}$$

称为 n 维随机变量 $(X_1,X_2,\cdots,X_n)$ 的分布函数或随机变量 $X_1,X_2,\cdots,X_n$ 的联合分布函数。

6.2　二维随机变量（向量）

定义 6.2.1. 联合分布函数：

设 (X,Y) 是二维随机变量，对于任意实数 x,y，二元函数

$$F(x,y)=P\{(X\le x)\cap(Y\le y)\}=P\{X\le x,Y\le y\}$$

称为二维随机变量的分布函数，或称为随机变量 X 和 Y 的联合分布函数。

容易算出随机点 (X,Y) 落在矩形域 $\{(x,y)|x_1<x\le x_2,y_1<y\le y_2\}$ 的概率为

$$P\{x_1<X\le x_2,y_1<Y\le y_2\}=F(x_2,y_2)-F(x_2,y_1)+F(x_1,y_1)-F(x_1,y_2).$$

性质：

1.$F(x,y)$ 是变量 x 和 y 的不减函数，即对于任意固定的 y，当 $x_2>x_1$ 时 $F(x_2,y)>F(x_1,y)$；对于任意固定的 x，当 $y_2>y_1$ 时 $F(x,y_2)>F(x,y_1)$

2.$0\le F(x,y)\le 1$，且对于任意固定的 y，$F(-\infty,y)=0$；对于任意固定的 x，$F(x,-\infty)=0$；$F(-\infty,-\infty)=0$，$F(\infty,\infty)=1$

3.$F(x+0,y)=F(x,y)$，$F(x,y+0)=F(x,y)$，即 $F(x,y)$ 关于 x 右连续，关于 y 也右连续

4. 对于任意 $(x_1,y_1),(x_2,y_2)$，$x_1<x_2$，$y_1<y_2$，下列不等式成立：

$$F(x_2,y_2)-F(x_2,y_1)+F(x_1,y_1)-F(x_1,y_2)\ge 0.$$

定义 6.2.2. 边缘分布函数：

$F_X(x)=F(x,+\infty)$ 和 $F_Y(y)=F(+\infty,y)$ 分别成为二维随机变量 (X,Y) 关于 X，关于 Y 的边缘分布函数。

$F_X(x)$ 和 $F_Y(y)$ 可由 $F(x,y)$ 惟一地确定，但是，反过来并不一定成立。

6.3 二维离散型随机变量

定义 6.3.1. 二维离散型随机变量:

如果二维随机变量 (X,Y) 全部可能取到的值是有限对或可数无穷多对，则称 (X,Y) 是离散型的随机变量。

设二维离散型随机变量 (X,Y) 所有可能取的值为 $(x_i,y_j),i,j=1,2,\cdots$，记

$$P\{X=x_i,Y=y_j\}=p_{ij},\quad i,j=1,2,\cdots,$$

则由概率的定义有

$$p_{ij}\geq 0,\quad \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}p_{ij}=1.$$

已知联合分布律可以求出其联合分布函数，由分布函数也可以求出其联合分布律。

6.4 二维连续型随机变量

定义 6.4.1. 二维连续型随机变量:

与一维随机变量相似，对于二维随机变量 (X,Y) 的分布函数 $F(x,y)$，如果存在非负的函数 $f(x,y)$，使对于任意 x,y，有

$$F(x,y)=\int_{-\infty}^{y}\int_{-\infty}^{x}f(u,v)\mathrm{d}u\mathrm{d}v,$$

则称 (X,Y) 是连续型的二维随机变量，函数 $f(x,y)$ 称为二维随机变量 (X,Y) 的概率密度，或称为随机变量 X 和 Y 的联合密度函数。

性质:

1.$f(x,y)\geq 0$

2. $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f(x,y)\mathrm{d}x\mathrm{d}y=F(\infty,\infty)=1$

3. 设 G 是 xOy 平面上的区域，点 (X,Y) 落在 G 内的概率为

$$P\{(X,Y)\in G\}=\iint_G f(x,y)\mathrm{d}x\mathrm{d}y;$$

4. 若 $f(x,y)$ 在点 (x,y) 连续，则有

$$\frac{\partial^2F(x,y)}{\partial x\partial y}=f(x,y).$$

定义 6.4.2. 二维均匀分布 $U(G)$:

设 (X,Y) 为二维随机变量，G 是平面上的一个有界区域，其面积为 $A(A\neq 0)$，若二维随机

变量的概率密度为

$$f(x,y)=\begin{cases}\dfrac{1}{A}, & (x,y)\in G,\\ 0, & (x,y)\notin G,\end{cases}$$

则称 (X,Y) 在区域 G 上服从二维均匀分布。

定义 6.4.3. 二维正态分布 $N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$:

若二维随机变量的概率密度为

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

其中 $-\infty<\mu_1,\mu_2<\infty$，$\sigma_1,\sigma_2>0$，$|\rho|<1$，则称 (X,Y) 服从二维正态分布 $N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$。

注意：$N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$ 在某些教科书里也记作 $N(\mu_1,\sigma_1^2;\mu_2,\sigma_2^2;\rho)$。

重要结论：

如果二位随机变量 (X,Y) 服从二维正态分布 $N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，则

1. 它们相互独立的充要条件是 $\rho=0$

2.X 服从正态分布 $N(\mu_1,\sigma_1^2)$，Y 服从正态分布 $N(\mu_2,\sigma_2^2)$。

7 条件分布

定义 7.0.1. 二维条件分布:

对于二维随机变量 (X,Y)，考虑当它的一个分量的取值给定时，另一个分量的概率分布，这种分布就是二维条件分布。

7.1 二维离散型随机变量的条件分布

定义 7.1.1. 二维离散型随机变量的条件分布:

设 (X,Y) 为二维离散型随机变量，其联合分布率以及 (X,Y) 关于 X、关于 Y 的边缘分布率分别为 $p_{ij},p_{i\cdot},p_{\cdot j},i,j=1,2,\cdots$，对于固定的 j，若 $p_{\cdot j}=P\{Y=y_j\}>0$，则

$$P\{X=x_i|Y=y_j\}=\frac{p_{ij}}{p_{\cdot j}},\quad i=1,2,\cdots$$

为在 $Y=y_j$ 的调剂下随机变量 X 的条件分布率。

同理，对于固定的 i，若 $p_{i\cdot}=P\{X=x_i\}>0$，则称

$$P\{Y=y_j|X=x_i\}=\frac{p_{ij}}{p_{i\cdot}},\quad j=1,2,\cdots$$

为在 $X = x_i$ 条件下的随机变量 Y 的条件分布率。

7.2 二维连续型随机变量的条件分布

定义 7.2.1. 二维连续型随机变量的条件分布:

设 (X, Y) 为二维连续型随机变量，其联合概率密度维 $f(x, y)$，(X, Y) 关于 X, Y 的边缘概率密度为 $f_X(x), f_Y(y)$。若对于固定的 y，$f_Y(y) > 0$，则在给定 $Y = y$ 时 X 的条件分布函数为

$$F_{X|Y}(x|y) = P\{X \leq x|Y = y\} = \int_{-\infty}^{x} f_{x|y}(u|y)\mathrm{d}u,$$

其中 $f_{X|Y}(x|y) = \dfrac{f(x, y)}{f_Y(y)}$ 称为在 $Y = y$ 的条件下 X 的条件概率密度。

类似地，若对于固定的 x，$f_X(x) > 0$，则在给定 $X = x$ 时 Y 的条件分布函数为

$$F_{Y|X}(y|x) = P\{Y \leq y|X = x\} = \int_{-\infty}^{y} f_{y|x}(v|x)\mathrm{d}v,$$

其中 $f_{Y|X}(y|x) = \dfrac{f(x, y)}{f_X(x)}$ 称为在 $X = x$ 条件下 Y 的条件概率密度。

有乘法公式 $f(x, y) = f_X(x)f_{Y|X}(y|x) = f_Y(y)f_{X|Y}(x|y)$。

8 随机变量的相互独立性

定义 8.0.1. 二维变量的相互独立性:

二维随机变量 (X, Y), 对于任意的 x, y, 有

$$F(x, y) = F_X(x)F_Y(y)$$

定义 8.0.2. 离散变量的相互独立性:

X, Y 相互独立, 当且仅当对于任意可能取值 (x_i, y_j), 有

$$p_i j = p_i p_j$$

定义 8.0.3. 连续变量的相互独立性:

X, Y 相互独立，当且仅当对于实数域内的任意 x, y, 下式几乎处处成立（除去平面上面积为 0 的单点）

$$f(x, y) = f_X(x)f_Y(y)$$

8.1 多维随机变量相互独立

设有 n 维随机变量 $(X_1, ..., X_n)$, 和 m 维随机变量量 $(Y_1, ..., Y_n)$

定义 8.1.1. n 维随机变量的联合分布函数：

$$F(x_1, ..., x_n) = PX_1 \leq x_1, ..., X_n \leq x_n$$

定义 8.1.2. n 维随机变量相互独立：

$$F(x_1, ..., x_n) = F_{X_1}(x_1)F_{X_2}(x_2)...F_{X_n}(x_n)$$

定义 8.1.3. 随机变量序列相互独立：

设随机变量序列 $X_1, ..., X_n, ...$, 若对于任意的 n>0, 有 $X_1, ..., X_n$ 相互独立，那么此随机变量序列相互独立。

定义 8.1.4. $(X_1, ..., X_n)$ 与 $(Y_1, ..., Y_m)$ 相互独立：

设 $(X_1, ..., X_n)$ 的分布函数 $F_1(x_1, ..., x_n)$,$(Y_1, ..., Y_m)$ 的分布函数 $F_2(y_1, ..., y_m)$, 则当 $F(x_1, ..., x_n, y_1, ..., y_m) = F_1(x_1, ..., x_n)F_2(y_1, ..., y_m)$ 称 $(X_1, ..., X_n)$ 与 $(Y_1, ..., Y_m)$ 相互独立。

8.2 相互独立的相关定理

定理 8.2.1. 相互独立的相关定理 1:

设 n 维随机变量 $(X_1, ...X_n)$ 和 m 维随机变量 $(Y_1, ...Y_m)$ 相互独立，则多元连续函数 $h(X_1, ..., X_n)$ 和 $g(Y_1, ..., Y_m)$ 也相互独立。

定理 8.2.2. 相互独立的相关定理 2:

若随机变量 $X_1, ...X_n$，把他们分为不相交的 k 个组，每个组中的所有变量由一个连续函数复合而成一个新的随机变量，则这 k 个新的随机变量仍相互独立。

9 随机变量函数的概率分布

定义 9.0.1. 随机变量函数:

设随机变量 X 取值范围为 D, 函数 $g(x)$ 为定义在 D 上的连续函数或分段单调函数的实函数，如果对于 D 内的任一 x，随机变量 Y 取值 $y = g(x)$，称 Y 是 X 的函数。

9.1 一维随机变量函数的概率分布

定义 9.1.1. 一维离散型随机变量的分布律：

将所有满足 $y = g(x)$ 中 y 的不同取值找出，再对每个 y 取值中所有对应 x 的概率求和。

定义 9.1.2. 一维连续型随机变量函数的概率密度函数：

求取 Y 的分布函数 $F_Y(y)$ 后再求导，此时

$$F_Y(y) = P\{Y \leq y\} = P\{g(X) \leq y\} = \int_{g(x)\leq y} f_X(x)dx$$

Tips: 如何确定积分区间非常重要，需要根据积分上下限的不同情况进行分类讨论。

定理 9.1.3. 重要定理 1：

如果是严格单调的可导函数，则有：

$$f_Y(y) = \begin{cases} f_X(g^{-1}(y))|\dfrac{dg^{-1}(y)}{dy}|, \alpha < y < \beta \\ 0, \end{cases}$$

其中 $\alpha = \min\limits_{-\infty<x<+\infty} g(x), \beta = \max\limits_{-\infty<x<+\infty} g(x)$

定理 9.1.4. 重要定理 2：

如果 $g(x)$ 是可导的偶函数，且当 $x > 0$ 时严格单调，则有：

$$f_Y(y) = \begin{cases} 2f_X(g^{-1}(y))|\dfrac{dg^{-1}(y)}{dy}|, \alpha < y < \beta \\ 0, \end{cases}$$

其中 $\alpha = \min\limits_{-\infty<x<+\infty} g(x), \beta = \max\limits_{-\infty<x<+\infty} g(x)$

9.2 多维随机变量函数的概率分布

定义 9.2.1. 二维离散型随机变量的分布律：

将所有满足 $z = g(x, y)$ 中所有可能的 (x, y) 的不同取值找出，再对每个 z 取值中所有对应 (x, y) 的概率求和。

定义 9.2.2. 二维连续型随机变量函数的概率密度函数：

求取 Z 的分布函数 $F_Z(z)$ 后再求导，此时

$$F_Z(z) = P\{Z \leq z\} = P\{g(X, Y) \leq z\} = \iint_{g(x,y)\leq z} f(x, y)dxdy$$

Tips: 这里对积分区域的讨论要比一维随机变量函数要更加复杂，注意考虑所有可能情况。

常见多维随机变量的概率分布/密度函数:

定义 9.2.3. 随机变量和的概率密度函数:

设 $Z = X + Y$，则有:

$$f_Z(z) = \int_{-\infty}^{+\infty} f(x, z - x)dx = \int_{-\infty}^{+\infty} f(z - y, y)dy$$

如果 X, Y 相互独立，则有:

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x)f_Y(z - x)dx = \int_{-\infty}^{+\infty} f_X(z - y)f_Y(y)dy$$

相互独立的正态分布随机变量的和:

若 $X_i\ N(\mu_i, {\sigma_i}^2)(i = 1, ..., n)$，它们相互独立，设 $Z = \sum\limits_{i=1}^{n} X_i$，则有 $Z\ N(\sum\limits_{i=1}^{n} \mu_i, \sum\limits_{i=1}^{n} {\sigma_i}^2)$

定义 9.2.4. 随机变量商的概率密度函数:

设 Z=$\dfrac{X}{Y}f_Z(z) = \int_{-\infty}^{+\infty} |y|f(yz, y)dy$X、Y$f_Z(z) = \int_{-\infty}^{+\infty} |y|f_X(yz)f_Y(y)dy$

定义 9.2.5. 随机变量最大最小值的概率分布函数:

设 $X_1, X_2, ..., X_n$ 相互独立,$M = max(X_1, X_2, ..., X_n), N = min(X_1, X_2, ..., X_n)$

$$F_M(z) = F_{X_1}(z)F_{X_2}(z)...F_{X_n}(z)$$

$$F_N(z) = 1 - (1 - F_{X_1}(z))(1 - F_{X_2}(z))...(1 - F_{X_n}(z))$$

对其求导即可得到概率密度函数。

10 典型例题

例题 1. 设二维随机变量 (X, Y) 的概率密度

$$f(x, y) = \begin{cases} ae^x, 0 < x < 1, x < y < 1 \\ 0, \end{cases}$$

求：(1)a (2)f$_X(x)f_Y(y)$ (3)XY (4)$P\{X + Y \le 0.5\}$

【解】(1)

$\int_0^1 dx \int_x^1 ae^x dy = a((2 - x)e^x)|_0^1 = a(e - 2) = 1$，因此 $a = 1/(e - 2)$。

（注：此题主要考察概率密度函数积分为 1 的性质）(2)

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)dy=\begin{cases}\int_x^1 ae^x dy=\frac{1}{e-2}(1-x)e^x, 0<x<1\\ 0,\end{cases}$$

$$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)dx=\begin{cases}\int_0^y ae^x dy=\frac{e^y-1}{e-2}, 0<y<1\\ 0,\end{cases}$$

（注：此题主要考察边缘概率密度的计算方法）

(3) 由于在 $0<x<1, x<y<1$ 内, $f(x,y)_X(x)f_Y(y)$, 所以它们不独立。

(4)（首先画图确定积分区间，这里我们先对 y 积分再对 x 积分，由 $X+Y<0.50<x<1x<y<1$ 组成的区域，可得 y 的积分区域为 $(x,0.5-x)$，x 的积分区域为 $(0,0.25)$）

$$P(X+Y<0.5)=\int_0^0 .25dx\int_x^{0.5-x}\frac{e^x}{e-2}dy=\frac{(5-4x)e^x}{2(e-2)}|_0^0 .25=\frac{4e^0 .25-5}{2(e-2)}$$

例题 2. 设二维随机变量 (X,Y) 相互独立且都在 $[0,a](a>0)$ 上均匀分布，求 $Z=X-Y$ 的概率密度。

【解】先画图确定积分区间，由 $X-Y\le Z$、$0\le x\le a$、$0\le y\le a$ 确定积分区间，可以看出积分的上下限根据 z 的取值需要分 4 种情况讨论：$(-\infty,-a)[-a,0][0,a](a,+\infty)$

$$F_Z(z)=\begin{cases}0, z<-a;\\ \int_0^{a+z}dx\int_{x-z}^{a}\frac{1}{a^2}dy=\frac{(a+z)^2}{2a^2}, -a\le z\le 0;\\ 1-\int_z^a dx\int_0^{x-z}\frac{1}{a^2}dy=1-\frac{(a-z)^2}{2a^2}, 0<z\le a;\\ 1,\end{cases}$$

$$f_Z(z)=\frac{dF_Z(z)}{dz}\begin{cases}\frac{a+z}{a^2}, -a\le x\le 0;\\ \frac{a-z}{a^2}, 0<x\le a^2;\\ 0,\end{cases}$$

Part III

随机变量的数字特征

11 数学期望与方差

11.1 数学期望

注意：以下定义的前提均要求所对应的级数、反常积分绝对收敛。

定义 11.1.1. 离散型随机变量的数学期望:

$$E(X) = \sum_{i=1}^{\infty} x_i p_i$$

定义 11.1.2. 连续型随机变量的数学期望:

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

定义 11.1.3. 一维离散型随机变量的函数的数学期望:

设 $Y = g(X)$ 为随机变量函数

$$E(Y) = E(g(Y)) = \sum_{i=1}^{\infty} g(x_i) p_i$$

定义 11.1.4. 一维连续型随机变量的函数的数学期望:

设 $Y = g(X)$ 为随机变量函数

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

定义 11.1.5. 二维离散型随机变量的函数的数学期望:

设 $Z = g(X, Y)$ 为随机变量函数

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

定义 11.1.6. 二维连续型随机变量的函数的数学期望:

设 $Z=g(X,Y)$ 为随机变量函数

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

数学期望的重要性质:

1.$E(C)=c$

2.$E(aX+bY)=aE(X)+bE(Y)$

3. 若 XY 相互独立，则 $E(XY)=E(X)E(Y)$

11.2 方差

定义 11.2.1. 随机变量的方差:

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

方差的另一种求法:

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

方差的重要性质:

1.$D(C)=0$

2.$D(CX)=C^2D(X)$

3.$D(X\pm Y)=D(X)\pm 2E((X-E(X))(Y-E(Y)))+D(Y)$

4. 若 $X_1,...X_n$ 相互独立，则 $D(\sum\limits_{i=1}^{n}C_iX_i)=\sum\limits_{i=1}^{n}{C_i}^2D(X_i)$

5.$D(X)=0\Leftrightarrow P(X=E(X))=1$

定义 11.2.2. 随机变量的标准差:

$$\Sigma(X)=\sqrt{D(X)}=\sqrt{E(X-E(X))^2}$$

定义 11.2.3. Markov 不等式:

若 $E(|X|^r)<\infty(r>0)$，则任意的 $\epsilon>0P\{|X|\geq\epsilon\}\leq\dfrac{E(|X|^r)}{\epsilon^r}$

11.3 常见分布的数学期望与方差

X	$E(X)$	$D(X)$
$B(n,p)$	np	$np(1-p)$
$P(\lambda)$	λ	λ
$N(\mu,\sigma^2)$	μ	σ^2
$U[a,b]$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
$exp(\lambda)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
$G(p)$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
$P\{X=a\}=1$	a	0

12 协方差、相关系数和矩

12.1 协方差

定义 12.1.1. 二维随机变量的协方差:

$$Cov(X,Y)=E\{(X-E(X))(Y-E(Y))\}$$

协方差的另一种求法:

$$Cov(X,Y)=E(XY)-E(X)E(Y)$$

协方差与方差的关系：

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

$$D(X\pm Y)=D(X)+D(Y)\pm 2Cov(X,Y)$$

协方差的重要性质:

1.$Cov(X,Y)=Cov(Y,X)$

2.$Cov(X_1+X_2,Y)=Cov(X_1,Y)+Cov(X_2,Y)$

3.$Cov(aX,bY)=abCov(XY)$

4. 若 XY 相互独立, 则 $Cov(X,Y)=0$。（注：反之不一定成立）

定义 12.1.2. Cauchy-Schwarz 不等式:

$$[E(XY)]^2\leq E(X^2)E(Y^2)$$

特别的:

$$[Cov(X,Y)]^2 \le D(X)D(Y)$$

12.2 相关系数

定义 12.2.1. 二维随机变量的相关系数:

$$rho_X Y = \frac{Cov(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$$

若 $rho_X Y = 0$，则称 XY 不相关

相关系数的重要性质:

1. 若 XY 相互独立，则 $rho_X Y = 0$（注：反之不一定成立）

2.$|\ rho_X Y| \le 1, |\ \mathrm{rho}_X Y| = 1 \Leftrightarrow \exists a, b$, 使得 $P\{Y = a + bX\} = 1$

3. 若 (X,Y) 服从二维正态分布，则 X、Y 相互独立充要条件为 $rho_X Y = 0$

12.3 矩

定义 12.3.1. 随机变量的 k 阶原点矩:

$$\alpha_n = E(X^n)$$

特别地，$\alpha_1 = E(X)$，即数学期望是一阶原点矩。

定义 12.3.2. 随机变量的 k 阶中心矩:

$$\mu_n = E[(X - E(X))^n]$$

特别地，$\mu_2 = D(X)$，即即方差是二阶中心矩。。

定义 12.3.3. 二维随机变量的 k+l 阶混合原点矩:

$$\alpha_k l = E(X^k Y^l)$$

定义 12.3.4. 二维随机变量的 k+l 阶混合中心矩:

$$\mu_k l = E[(X - E(X))^k (Y - E(Y))^l]$$

特别地，$\mu_1 1 = Cov(X,Y)$，即协方差是二阶混合中心矩。

13 典型例题

例题 3. 设二维随机变量 (X, Y) 概率密度为

$$f(x,y)=\begin{cases}\frac{1}{3}(x+y), 0\leq x\leq 2, 0\leq y\leq 1\\ 0,\end{cases}$$

求：(1)$E(X)$　　(2)$E(X^2+Y^2)$

【解】

$$f_X(x)=\begin{cases}\int_0^1\frac{1}{3}(x+y)dy=\frac{1}{6}(2x+1), 0\leq x\leq 2\\ 0,\end{cases}$$

$$E(X)=\int_{-\infty}^{+\infty}xf_X(x)dx=\int_{\frac{1}{6}}x(2x+1)dx=\frac{11}{9}$$

$$E(X^2+Y^2)=\int_0^2dx\int_0^1\frac{1}{3}(x^2+y^2)(x+y)dy=\frac{13}{6}$$

例题 4. 设二维随机变量 (X, Y) 有 $E(X)=0E(Y)=0E(X^2)=1E(XY)=-1\ rho_XY=-0.5$，求 $D(X-Y)$

【解】

$$D(X)=E(X^2)-(E(X))^2=1$$

$$rho_XY=\frac{Cov(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}=\frac{E(XY)-E(X)E(Y)}{\sqrt{D(X)}\sqrt{D(Y)}}=-\frac{1}{\sqrt{D(Y)}}=-0.5$$

得 $D(Y)=4$。

$$D(X-Y)=D(X)+D(Y)-2Cov(X,Y)=D(X)+D(Y)-2E(XY)+2E(X)E(Y)=7$$

Part IV

大数定律与中心极限定理

14 大数定律

14.1 大数定律

定理 14.1.1. 大数定律:

是一种描述当试验次数很大时随机事件所呈现的概率性质的定律。以明确的数学形式表达了频率的稳定性和平均结果的稳定性，并讨论了它们成立的条件。

14.2 切比雪夫不等式

定理 14.2.1. 切比雪夫不等式:

设随机变量 X 的数学期望 $E(X)$ 与方差 $D(X)$ 都存在，则对任意的常数 $\varepsilon > 0$，有

$$P\{| X - E(X) |\geq \varepsilon\} \leq \frac{D(X)}{\varepsilon^2}$$

或

$$P\{| X - E(X) |< \varepsilon\} \geq 1 - \frac{D(X)}{\varepsilon^2}$$

可见方差是反映随机变量取值集中在 $E(X)$ 附近的程度的数量指标。

14.3 伯努利大数定律

定理 14.3.1. 伯努利大数定律:

设 η_n 是 n 重伯努利试验中事件 A 发生的次数，p 是事件 A 发生的概率，则对任意的 $\varepsilon > 0$，有

$$\lim_{n\to\infty} P\{| \frac{\eta_n}{n} - p |< \varepsilon\} = 1$$

记作

$$\frac{\eta_n}{n} \xrightarrow{P} p$$

伯努利大数定律表明频率稳定于概率的含义。

14.4 实际推断原理

定理 14.4.1. 实际推断原理:

在实际的生产活动中，通常认为小概率事件在一次随机试验中是不可能发生的。

14.5 依概率收敛

定理 14.5.1. 依概率收敛:

设 X_1，X_1，$\cdots$ ，X_n ，$\cdots$ 为随机变量序列，a 为常数，如果对任意的 $\varepsilon > 0$，有

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

或

$$\lim_{n\to\infty} P\{| X_n - a |\geq \varepsilon\} = 0$$

则称 $\{X_n\}$ 依概率收敛于 a，记作 $X_n \xrightarrow{P} a$。

性质：

如果 $X_n \xrightarrow{P} a$，$Y_n \xrightarrow{P} b$，设二元函数 $g(x, y)$ 在点 (a, b) 处连续，则 $g(X_n, Y_n) \xrightarrow{P} g(a, b)$。

14.6 切比雪夫大数定律

定理 14.6.1. 切比雪夫大数定律:

设 X_1，X_2，$\cdots$，X_n，$\cdots$ 是两两不相关的随机变量序列，且方差一致有界，即存在某一常数 C，使得 $D(X_i) \leq C(i = 1，2，\cdots)$，则对任意的 $\varepsilon > 0$，都有

$$\lim_{n\to\infty} P\{| \frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} E(X_i) |\geq \varepsilon\} = 0$$

或

$$\lim_{n\to\infty} P\{| \frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} E(X_i) |< \varepsilon\} = 1$$

14.7 独立同分布随机变量的大数定律

定理 14.7.1. 独立同分布随机变量的大数定律:

设 X_1，X_2，$\cdots$，X_n，$\cdots$ 是独立同分布的随机变量序列，且存在数学期望 $E(X_i) = \mu$ 与方差 $D(X_i) = \sigma^2$，$i = 1$，2，$\cdots$，则对任意的 $\varepsilon > 0$，都有

$$\lim_{n\to\infty} P\{| \frac{1}{n}\sum_{i=1}^{n} X_i - \mu |\geq \varepsilon\} = 0$$

或

$$\lim_{n\to\infty} P\{| \frac{1}{n}\sum_{i=1}^{n} X_i - \mu |< \varepsilon\} = 1$$

即 $\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} \mu$

伯努利大数定律、独立同分布随机变量的大数定律可以看作切比雪夫大数定律的特例。

15 中心极限定理

15.1 独立同分布中心极限定理

定理 15.1.1. 独立同分布中心极限定理（林德贝格-勒维定理）：

设 X_1，X_2，$\cdots$，X_n 是独立同分布的随机变量序列，且存在数学期望 $E(X_i)=\mu$ 与方差 $D(X_i)=\sigma^2$，$i=1$，2，$\cdots$，则对任意的实数 $x\in R$，都有

$$\lim_{n\to\infty} P\{\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} \le x\} = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \Phi(x)$$

该定理表明，满足该定理的条件的 n 个随机变量的和 $\sum_{i=1}^{n} X_i$ 的极限分布是正态分布 $N(n\mu, n\sigma^2)$。

15.2 德莫佛-拉普拉斯中心极限定理

定理 15.2.1. 德莫佛-拉普拉斯中心极限定理:

设 η_n 是 n 重伯努利试验中事件 A 发生的次数，p 是事件 A 在每次试验中发生的概率，则对任意的实数 $x\in R$，有

$$\lim_{n\to\infty} P\{\frac{\eta_n - np}{\sqrt{np(1-p)}} \le x\} = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \Phi(x)$$

该定理表明，正态分布是二项分布的极限分布。

德莫佛-拉普拉斯中心极限定理可以看作独立同分布中心极限定理的特例。

如果一个随机变量所描述的随机现象是由许多相互独立的因素叠加而成，而且每个因素对该现象的影响都很微小，那么这个随机变量就可以认为近似服从正态分布。

（注：本章的典型例题有例 4.2.4，例 4.2.8，以及习题 4 中的 7，11，18，19。）

Part V

数理统计的基本概念

16 总体与样本

16.1 基本概念

定义 16.1.1. 总体:

研究对象全体元素组成的集合。指所研究的对象的某个或某些数量指标的全体，它是一个随机变量（或多维随机变量），记为 X。X 的分布函数和数字特征称为总体的分布函数和数字特征。

定义 16.1.2. 个体:

组成总体的每一个元素。总体的每个数量指标，可看作随机变量 X 的某个取值，用 X_i 表示。

定义 16.1.3. 样本:

从总体中抽取的部分个体，用 $(X_1, X_2, ..., X_n)$ 表示，n 为样本容量，称 $(x_1, x_2, ..., x_n)$ 为总体 X 的一个容量为 n 的样本观测值，或称样本的一个实现。

定义 16.1.4. 样本空间:

样本所有可能取值的集合

定义 16.1.5. 简单随机样本:

通过简单随机抽样得到的样本，即每个样本 $X_i (i = 1, 2, ..., n)$ 必须与总体具有相同的分布，并且样本 $(X_1, X_2, ..., X_n)$ 是相互独立的。

16.2 样本的联合分布函数

定义 16.2.1. 样本的联合分布函数:

设总体 X 的分布函数为 $F(x)$，则样本 $(X_1, X_2, ..., X_n)$ 的联合分布函数为:

$$P\{X_1 \leq x_1, X_2 \leq x_2, ..., X_n \leq x_n\} = \prod_{i=1}^{n} F(x_i)$$

定义 16.2.2. 样本的联合分布律:

总体 X 是离散型随机变量且其分布律为 $P\{X = x_i\} = p(x_i)$ 时，若样本 $(X_1, X_2, ..., X_n)$ 的观测值为 $(x_1, x_2, ..., x_n)$，则联合分布律为:

$$P\{X_1=x_1,X_2=x_2,...,X_n=x_n\}=\prod_{i=1}^{n}p(x_i)$$

定义 16.2.3. 样本的联合概率密度:

总体 X 是连续性随机变量且其概率密度为 $f(x)$ 时，若 $(X_1,X_2,...,X_n)$ 的联合概率密度为:

$$P(x_1,x_2,...,x_n)=\prod_{i=1}^{n}f(x_i)$$

17 样本分布

总体 X 是随机变量，其概率分布是客观存在的，实际应用中常常用样本分布作为总体分布的近似，常用方法有：频数分布与频率分布（对离散型总体），频率直方图（对连续型总体），经验分布函数（对各类型总体）。

17.1 经验分布函数

定义 17.1.1. 经验分布函数:

设 $(x_1,x_2,...,x_n)$ 是来自于总体 X 的样本值，将这些值按由小到大的顺序进行排列为 $x_{(1)}\le x_{(2)}\le...\le x_{(n)}$，对于任意的实数 x，定义函数 $F_n(x)$ 为总体 X 的经验分布函数:

$$F_nx=\begin{cases}0, x<x_{(1)}\\ \dfrac{k}{n}, x_{(k)}\le x<x_{(k+1)}\\ 1, x\ge x_{(n)}\end{cases}$$

17.2 格里纹科定理

定理 17.2.1. 格里纹科定理:

当样本容量 n 充分大时，经验分布函数和总体分布函数最大的绝对偏差可以足够小，即当 n 充分大时，经验分布函数是总体分布函数的一个很好的近似, 即:

$$P\{sup|F_n(x)-F_n(x)\to 0|\}=1$$

18 统计量及其数字特征

18.1 常用统计量

设 $(X_1,X_2,...,X_n)$ 是来自总体 X 的容量为 n 的样本，有常用统计量如下:

定义 18.1.1. 样本均值:

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

定义 18.1.2. 样本方差：

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

定义 18.1.3. 样本标准差：

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

定义 18.1.4. 样本的 k 阶原点矩：

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

定义 18.1.5. 样本的 k 阶中心矩：

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

定义 18.1.6. 顺序统计量，样本极值与极差：

设 $(X_1, X_2, ..., X_n)$ 为样本，$(x_1, x_2, ..., x_n)$ 为样本值，将观测值从小到大顺序排列，记为 $x_{(1)} \leq x_{(2)} \leq ... \leq x_{(n)}$，定义 $X_{(k)} = x_{(k)}$，则有：

顺序统计量：$X_{(1)}, X_{(2)}, ..., X_{(n)}$

样本极值：$X_{(n)} = \max\{X_k\}, X_{(1)} = \min\{X_k\}$

样本极差：$D_n = X_{(n)} - X_{(1)}$

定义 18.1.7. 样本中位数：

$$M = \begin{cases} X_{(n+1)/2}, odd \\ 0.5X_{n/2} + 0.5X_{(n+2)/2}, even \end{cases}$$

定义 18.1.8. 样本 P 分位数：

$$M_p = X_{(np+p)}, 0 < p < 1$$

18.2 重要定理

定理 18.2.1. 重要定理 1：

设总体 X 的数学期望 $E(X) = \mu$ 和方差 $D(X) = \sigma^2$ 存在，则有：

（1）$E(\overline{X}) = \mu, D(\overline{X}) = \dfrac{\sigma^2}{n}$，且当 $n \to \infty$ 时，$\overline{X}$ 依概率收敛于 μ

（2）当 X 的四阶矩存在，$E(S^2) = \sigma^2$，S^2 依概率收敛于 σ^2

定理 18.2.2. 重要定理 2：

设总体 X 的 k 阶矩 $\alpha_k = E(X^k)$ 存在，则有：

（1）$E(A_k) = \alpha_k$

（2）当 X 的 $2k$ 阶矩存在，A_k 依概率收敛于 α_k

定理 18.2.3. 重要定理 3:

样本的方差 S^2 与二阶中心矩 S_n^2 满足:

（1） $S_n^2 = \frac{n-1}{n} S^2$

（2） $E(S_n^2) = \frac{n-1}{n}\sigma^2, E(S^2) = \sigma^2$

19 抽样分布

统计量的分布称为抽样分布，正态总体的抽样分布与 χ^2 分布，t 分布，F 分布三大分布有着重要关系。

19.1 分位数

定义 19.1.1. 上侧分位数:

设随机变量 X 的分布函数为 $F(x) = P\{X \le x\}$，对于给定的 $\alpha(0 < \alpha < 1)$，若存在使 $P\{X > x_\alpha\} = 1 - F(x_\alpha) = \alpha$，则称 α 为 X 分布的上侧 α 分位数或上侧临界值。

定义 19.1.2. 双侧分位数:

若存在数 λ_1 和 λ_2，使得 $P\{X \ge \lambda_1\} = P\{X \le \lambda_2\} = \frac{\alpha}{2}$，则称 x_α 为 X 分布的双侧 α 分位数或双侧临界值，其中 $\lambda_1 = x_{\alpha/2}, \lambda_2 = x_{1-\alpha/2}$

当 X 的分布关于 y 轴对称时，若存在 $x_{\alpha/2}$，使得 $P\{|X| \ge x_{\alpha/2}\} = \alpha$，则称 $x_{\alpha/2}$ 为 X 分布的双侧 α 分位数或双侧临界值

19.2 χ^2 分布

定义 19.2.1. χ^2 分布:

设总体 $X \sim N(0,1)$，$(X_1, X_2, ..., X_n))$ 是 X 的一个样本，则称统计量 $\chi^2 = X_1^2 + X_2^2 + ... + X_n^2$ 服从自由度为 n 的 χ^2 分布，记作 $\chi^2 \sim \chi^2(n)$。$\chi^2(n)$ 分布的密度函数为:

$$\chi^2(x;n) = \begin{cases} \frac{1}{2^{n/2}\Gamma(n/2)} x^{n/2-1} e^{-x/2}, x \ge 0 \\ 0, x \le 0 \end{cases}$$

其基本性质为：1.$\chi^2(1)$ 为标准正态分布，$\chi^2(2)$ 为参数为 0.5 的指数分布

2.$\chi^2(n)$ 分布的上侧 α 分位数 $P\left\{\chi^2(n) \ge \chi_\alpha^2(n)\right\} = \alpha$

3.$\chi^2(n)$ 分布的双侧 α 分位数 $P\left\{\chi^2(n) \le \chi_{1-\alpha/2}^2(n)\right\} = P\left\{\chi^2(n) \ge \chi_{\alpha/2}^2(n)\right\} = \frac{\alpha}{2}$

4. 若 $Z \sim \chi^2(n)$，则 $E(Z) = n, D(Z) = 2n$

5. 若 $Z_1 \sim \chi^2(n_1)$，$Z_2 \sim \chi^2(n_2)$，且 Z_1, Z_2 相互独立，则 $Z_1 + Z_2 \sim \chi^2(n_1 + n_2)$，且可以推广到多个随机变量的情形

6. 设随机变量 $X_1, X_2, ..., X_n$ 相互独立，且都服从标准正态分布 $N(0,1)$，则随机变量 $\chi^2 = \sum_{i=1}^{n} X_i^2$，服从自由度为 n 的 χ^2 分布，即 $\chi^2 \sim \chi^2(n)$

7. 设 $X_1, X_2, ..., X_n$ 为取自正态总体 $X \sim N(\mu, \sigma^2)$ 的样本，则 $\dfrac{\sum_{i=1}^{n}(X_i - \mu)^2}{\sigma^2} \sim \chi^2(n)$

8. 设 $X_1, X_2, ..., X_n$ 为取自正态总体 $X \sim N()\mu, \sigma^2)$ 的样本，则 $\overline{X}$ 与 S^2 相互独立，且 $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$

19.3 *t* 分布

定义 19.3.1. t 分布:

设随机变量 $X \sim N(0,1), Y \sim \chi^2(n)$，且 X 与 Y 相互独立，则称统计量 $T = \dfrac{X}{\sqrt{Y/n}}$ 服从自由度为 n 的 t 分布或学生式分布，记作 $T \sim t(n)$。其密度函数为:

$$t^2(x;n) = \frac{\Gamma\left(\dfrac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\dfrac{n}{2}\right)}(1+\frac{x^2}{n})^{-\dfrac{n+1}{2}}, -\infty < x < +\infty$$

其基本性质为：1.t 分布密度函数形状类似于标准正态分布，n 比较大时，t 分布近似于标准正态分布，但对较小的 n 值，t 分布与标准正态分布之间有较大差异，且 $P\{|T| \geq t_0\} = P\{|X| \geq t_0\}$

2.$\chi^2(n)$ 分布的上侧 α 分位数 $P\{T > T_\alpha(n)\} = \alpha$

3.$\chi^2(n)$ 分布的双侧 α 分位数 $P\{|T| > t_{a/2}(n)\} = \alpha$

4. 若 $T \sim t(n)$，则 $E(T) = 0, D(T) = \dfrac{n}{n-2}(n > 2)$

5. 设 $X_1, X_2, ..., X_n$ 为取自正态总体 $X \sim N(\mu, \sigma^2)$ 的样本，则有 $T = \dfrac{\overline{X} - \mu}{S/\sqrt{n}} \sim t(n-1)$

6. 设 $X_1, X_2, ..., X_n$ 和 $Y_1, Y_2, ..., Y_n$ 为取自正态总体 $N(\mu_1], \sigma^2)$，$N(\mu_2], \sigma^2)$ 的样本，且相互独立，则有 $T = \dfrac{(\overline{X} - \overline{Y}) - (\mu_1 - \mu_2)}{S_n\sqrt{1/n_1 + 1/n_2}} \sim t(n_1 + n_2 - 2)$，且 $S_n = \sqrt{\dfrac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1 + n_2 - 1}}$

19.4 *F* 分布

定义 19.4.1. F 分布:

设随机变量 $X \sim \chi^2(n_1), Y \sim \chi^2(n_2),$，且 X 与 Y 相互独立，则称统计量 $F = \dfrac{X/n_1}{Y/n_2}$ 服从第一自由度为 n_1，第二自由度为 n_2 的 F 分布。记作 $F \sim F(n_1, n_2)$。其密度函数见书本。

其基本性质为：1. 若 $X \sim F(n_1, n_2)$，则 $\dfrac{1}{X} = F(n_2, n_1)$，进一步，有 $F_\alpha(n_1, n_2) = \dfrac{1}{F_{1-\alpha}(n_2, n_1)}$

2.$\chi^2(n)$ 分布的上侧 α 分位数 $P\{F(n_1, n_2) > F_\alpha(n_1, n_2)\} = \alpha$

3.$P\{F \leq F_{1-\alpha/2}(n_1, n_2)\} = P\{F \geq F_{\alpha/2}(n_1, n_2)\} = \dfrac{\alpha}{2}$

4. 设 n_1, S_1^2 为正态总体 $N(\mu_1, \sigma_1^2)$ 的样本容量和样本方差，n_2, S_2^2 为正态总体 $N(\mu_2, \sigma_2^2)$ 的样本容量和样本方差，且两个样本相互独立，则有：$\dfrac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2} \sim F(n_1-1, n_2-1)$

19.5 抽样分布定理

定理 19.5.1. 抽样分布定理 1：

设 $(X_1, X_2, ..., X_n)$ 是来自正态总体 $N(\mu, \sigma^2)$ 的样本，$\overline{X}$ 为样本均值，S^2 为样本方差，则有：

（1）$\overline{X} \sim N(\mu, \dfrac{\sigma^2}{n})$

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

（3）$\overline{X}$ 与 S^2 相互独立

定理 19.5.2. 抽样分布定理 2：

设 $(X_1, X_2, ..., X_n)$ 是来自正态总体 $N(\mu, \sigma^2)$ 的样本，$\overline{X}$ 为样本均值，S^2 为样本方差，则有：

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

定理 19.5.3. 抽样分布定理 3：

设 $(X_1, X_2, ..., X_{n_1})$，$(Y_1, Y_2, ..., Y_{n_2})$ 是分别来自总体 $N(\mu_1, \sigma^2), N(\mu_2, \sigma^2)$ 的样本，且两组样本相互独立，$\overline{X} = \dfrac{1}{n_1} \sum_{i=1}^{n_1} X_i$，$\overline{Y} = \dfrac{1}{n_2} \sum_{i=1}^{n_2} Y_i$,$S_{1n_1}^2 = \dfrac{1}{n_1 - 1} \sum_{i=1}^{n1}(X_i - \overline{X})^2$，$S_{2n_2}^2 = \dfrac{1}{n_2 - 1} \sum_{i=1}^{n2}(Y_i - \overline{Y})^2$，则有：

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

其中

$$S_W^2 = \frac{(n_1 - 1)S_{1n_1}^2 + (n_2 - 1)S_{2n_2}^2}{n_1 + n_2 - 1}$$

定理 19.5.4. 抽样分布定理 4：

设 $(X_1, X_2, ..., X_{n_1})$，$(Y_1, Y_2, ..., Y_{n_2})$ 是分别来自总体 $N(\mu_1, \sigma^2), N(\mu_2, \sigma^2)$ 的样本，且两组样本相互独立，则

$$F = \frac{\sigma_2^2 S_{1n_1}^2}{\sigma_2^1 S_{2n_2}^2} \sim F(n_1 - 1, n_2 - 1)$$

20 典型例题

例题 5. 设 $X_1, X_2, ..., X_n$ 是来自总体 $X \sim N(0,4)$ 的样本，$Y = a(X_1 - 2X_2)^2 + b(3X_3 - 4X_4)^2$，则 a 和 b 取何值时，统计量 Y 服从 χ^2 分布？

解：令 $Y_1 = X_1 - 2X_2, Y_2 = 3X_3 - 4X_4$，则 Y_1, Y_2 仍服从正态分布且有：

$$E(Y_1) = 0, D(Y_1) = 20, E(Y_2) = 0, D(Y_2) = 100$$

即 $Y_1 \sim N(0,20), Y_2 \sim N(0,100)$，进一步有：

$$\frac{(Y_1 - 0)^2}{20} \sim \chi^2(1), \frac{(Y_2 - 0)^2}{100} \sim \chi^2(1)$$

则有：

$$a = \frac{1}{20}, b = \frac{1}{100}, Y \sim \chi^2(2)$$

【评析】此题较为基本，关键在于掌握正态分布的基本性质，并灵活运用 χ^2 分布的相关结论

例题 6. 设 $X \sim F(n,n)$，证明 $P\{X > 1\} = 0.5$

解：由 $X \sim F(n,n)$，可得 $\frac{1}{X} \sim F(n,n)$，故

$$P\{X > 1\} = P\left\{\frac{1}{X} \geq 1\right\} = 0.5(P\{X > 1\} + P\{X \leq 1\}) = 0.5$$

【评析】此题为 F 分布的性质运用，其关键在于掌握 F 分布的相关性质

例题 7. 设 $(X_1, X_2, ..., X_n, X_{n+1}, ..., X_{n+m})$ 是来自正态总体 $N(0, \sigma^2)$ 的样本，试求下列统计量的概率分布：

$$(1) Y_1 = \frac{\sqrt{m}\sum_{i=1}^{n} X_i}{\sqrt{n}\sqrt{\sum_{i=n+1}^{n+m} X_i^2}} \quad (2) Y_2 = \frac{m\sum_{i=1}^{n} X_i^2}{n\sum_{i=n+1}^{n+m} X_i^2}$$

解：（1）令 $Z_1 = \sum_{i=1}^{n} X_i$，$Z_2 = \sum_{i=n+1}^{n+m} \frac{X_i^2}{\sigma^2}$，则有：

$$Z_1 \sim N(0, \sigma^2), \frac{Z_1}{\sqrt{n}\sigma} \sim N(0,1), Z_2 \sim \chi^2(m)$$

故有：

$$Y_1 = \frac{\sqrt{m}\sum_{i=1}^{n} X_i}{\sqrt{n}\sqrt{\sum_{i=n+1}^{n+m} X_i^2}} = \frac{Z_1/\sqrt{n}\sigma}{\sqrt{Z_2/m}} \sim t(m)$$

（2）令 $X_1 = \sum_{i=1}^{n} \frac{X_i^2}{\sigma^2} \sim \chi^2(n), X_2 = \sum_{i=n+1}^{n+m} \frac{X_i^2}{\sigma^2} \sim \chi^2(m)$，则有：

$$Y_2 = \frac{m\sum_{i=1}^{n} X_i^2}{n\sum_{i=n+1}^{n+m} X_i^2} = \frac{mX_1}{nX_2} = \frac{\frac{X_1}{n}}{\frac{X_2}{m}} \sim F(n.m)$$

【评析】此题是典型的证明统计量服从某分布的例题，对于这种题目，关键在于将题给统计量进行适当化简，并熟记三大分布相关形式，以便解题

Part VI

参数估计

纲要：参数估计分为两部分：点估计和区间估计，区级估计为假设检验提供工具。

21 参数的点估计——估计未知参数的值

参数是刻画总体某概率特性的数量，当此数量未知时，从总体抽出一系列样本，用某种方法对这个未知参数进行估计就是参数估计。

定义 21.0.1. 点估计：

设总体 X 的分布函数形式已知，但含有一个或多个未知参数：$\theta_1, \theta_2, \ldots, \theta_k$，$X_1, X_2, \ldots, X_n$ 设为总体的一个样本，构造 k 个统计量：$\theta_1(X_1, X_2, \ldots, X_n), \ldots, \theta_k(X_1, X_2, \ldots, X_n)$（估计量），代入样本观测值得到 k 个数（估计值）：$\hat{\theta_1}(X_1, X_2, \ldots, X_n), \ldots, \hat{\theta}_k(X_1, X_2, \ldots, X_n)$ 。

简而言之：估计量：根据样本选择出的统计量，如：均值，方差等; 估计值：根据所给样本的数据计算出估计量的数值。

21.1 频率替换法:

利用事件 A 在 n 次试验中发生的频率作为事件 A 发生概率 p 的估计量。

21.2 矩估计法:

主要思想方法：用样本矩的函数代替总体矩的函数，对总体的未知函数做出估计。基本解题思路：通过积分计算均值，方差等与样本特征有关的值，确定矩估计。

1. 用样本的 k 阶原点矩作为总体的 k 阶矩的估计量，建立含有待估参数的方程（组），从而解出待估参数。

定理 21.2.1. 定理：设总体 X 的 k 阶矩 $\alpha_k = E(X^k)$ 存在，则有：$E(A_k) = \alpha_k$；当 X 的 $2k$ 阶矩存在，A_k 依概率收敛于 α_k。

2. 设待估计参数为：$\theta_1, \theta_2, \ldots, \theta_k$，设总体的 r 阶原点矩存在，记为 $E(X^r) = \alpha_r(\theta_1, \theta_2, \ldots, \theta_k)$，样本 $X_1, X_2, \ldots, X_n$ 的 r 阶原点矩为 $A_r = \frac{1}{n}\sum\limits_{i=1}^{n} X_i^r$，令 $\alpha_r(\theta_1, \theta_2, \ldots, \theta_k) = \frac{1}{n}\sum\limits_{i=1}^{n} X_i^r$，得到含未知参数的方程（组），解得 k 个统计量（未知参数的矩估计量），代入样本观测值得 k 个数：$\hat{\theta}_1(X_1, X_2, \ldots, X_n), \ldots, \hat{\theta}_k(X_1, X_2, \ldots, X_n)$（未知参数的矩估计值）。

3. 一般地，不论总体服从什么分布，总体期望 μ 与方差 σ^2 存在时，则有 $\hat{\mu} = \bar{X}, \hat{\sigma^2} = A_2 - \hat{\mu}^2 = \frac{1}{n}\sum\limits_{i=1}^{n} X_i^2 - \bar{X}^2 = \frac{1}{n}\sum\limits_{i=1}^{n}(X_i - \bar{X})^2$。

21.3 极大似然估计法:

1. 思想方法：一次试验就出现的事件有较大的概率，即求概率函数最大值（利用单调性或求导计算极值）。

2. 一般地，设 X 为离散型随机变量，其分布律为 $P(X = x) = f(x, \theta)$，则样本 $X_1, X_2, \ldots, X_n$ 的联合概率分布为 $P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) = f(x_1, \theta)f(x_2, \theta)\ldots f(x_n, \theta)$，记作 $L(x_1, x_2, \ldots, x_n, \theta)$ 或 $L(\theta)$，称 $L(\theta)$ 为样本的似然函数。

   - 若 X 连续，取 $f(x_i, \theta)$ 为 X_i 的密度函数，则样本的似然函数为 $L(\theta_1, \theta_2, \ldots, \theta_k) = \prod\limits_{i=1}^{n} f(x_i, \theta_1, \theta_2, \ldots, \theta_k)$
   - 若有多个未知参数 $\theta_1, \theta_2, \ldots, \theta_k$，则可设 X 的密度（或分布）为 $f(x, \theta_1, \theta_2, \ldots, \theta_k)$，则样本似然函数为 $L(\theta_1, \theta_2, \ldots, \theta_k) = \prod\limits_{i=1}^{n} f(x_i, \theta_1, \theta_2, \ldots, \theta_k)$。

3. 极大似然法的思想：选择适当的 $\theta = \hat{\theta}$ 使 $L(\theta)$ 取最大值（令导数或取对数后导数为零得似然方程组），即 $L(x_1, x_2, \ldots, x_n, \hat{\theta}) = max\{f(x_1, \theta)f(x_2, \theta)\ldots f(x_n, \theta)\}$，称这样得到的为参数的极大似然估计值，称统计量参数的极大似然估计量。

4. 极大似然法估计参数的步骤

   i. 写出似然函数（联合分布函数）：
   $L(x_1, x_2, \ldots, x_n; \theta_1, \theta_2, \ldots, \theta_k) = \prod\limits_{i=1}^{n} f(x_i, \theta_1, \theta_2, \theta_k)$

   ii. 求出 $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k)$，使似然函数极大，即得到参数估计值 $\hat{\theta}_r = g(x_1, x_2, \ldots, x_n)$

iii. 可进一步得到估计量 $\hat{\theta}_r = g(X_1, X_2, \ldots, X_n)$。

iv. 注意：若有不同变量，应对不同的变量进行单独，整体求导（如将 σ^2 看作一个整体求导）

5. 极大似然估计的不变性设 $\hat{\theta}$ 是 θ 的极大似然估计值，$u(\theta)$ 是 θ 的函数，且有单值反函数（单调），则 $\hat{u}(\theta)$ 是 $\hat{u}(\theta)$ 的极大似然估计值。

22 估计量的评价标准

对于同一个未知参数，不同方法得到的估计量可能不同，则需要一定的标准来衡量某个估计量的优劣：无偏性（有无系统性误差）、有效性（波动性的大小，即方差大小）、一致性（当样本容量增大时估计值是否越来越精确）。

定义 22.0.1. 无偏性

1. 设 $\hat{\theta}(X_1, X_2, \ldots, X_n)$ 是未知参数 θ 的估计量，若 $E(\theta) = \theta$（估计值的期望与真值相等），则称 $\hat{\theta}$ 是 θ 的无偏估计量；若 $E(\theta) \neq \theta$，但 $\lim\limits_{n\to\infty} E(\theta) = \theta$，则称 $\hat{\theta}$ 是 θ 的渐近无偏估计量。

2. 对任意总体，样本均值是总体期望的无偏估计量，样本方差是总体方差的无偏估计量，样本 k 阶原点矩是总体 k 阶原点矩的无偏估计量。

3. 对任意总体，样本均值平方是总体期望平方的渐近无偏估计量，样本二阶中心矩是总体方差的渐近无偏估计量。

定义 22.0.2. 有效性

1. 设 $\hat{\theta}_1(X_1, X_2, \ldots, X_n)$ 和 $\hat{\theta}_2(X_1, X_2, \ldots, X_n)$ 都是对总体中某一未知参数的无偏估计量，若 $D(\hat{\theta}_1(X_1, X_2, \ldots, X_n)) < D(\hat{\theta}_1(X_1, X_2, \ldots, X_n))$，则称估计量 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 更有效。

2. 设 $\hat{\theta} = \hat{\theta}(X_1, X_2, \ldots, X_n)$ 是来自总体 X 的样本，其中总体均值 $E(X) = \mu$ 与方差 $D(X) = \sigma^2$ 都存在，则样本均值 $\bar{X}$ 在 μ 的所有线性无偏估计量中是最有效的（称为最小方差线性无偏估计量）。

定义 22.0.3. 一致性（相合性）

设 $\hat{\theta} = \hat{\theta}(X_1, X_2, \ldots, X_n)$ 是总体参数的估计量，若对于任意的 $\theta \in \Theta$，当 $n \to \infty$ 时，$\hat{\theta}$ 依概率收敛于 θ，则称 $\hat{\theta}$ 为 θ 的一致（相合）估计量。若当 $n \to \infty$ 时，$\hat{\theta}$ 均方收敛于 θ，则称为的均方相合估计量。

- 样本均值和样本方差分别是总体期望和总体方差的一致估计量，样本 k 阶原点矩既是总体 k 阶原点矩的一致估计量，又是其均方相合估计量。
- 样本方差和样本二阶中心矩都是总体方差的相合估计。
- 设是的无偏估计量，且，则是的一致估计量。
- 矩法估计得到的估计量一般为一致估计量；在一定条件下，极大似然估计可具有一致性。

23　参数的区间估计——估计未知参数的取值范围

23.1　双侧区间估计

由不同样本求得的参数估计值不同，希望根据所给样本确定一个随机区间，使其包含未知参数真值的概率达到指定要求。

定义 23.1.1. 置信区间

设 θ 为待估参数，$\alpha(0 < \alpha < 1)$ 是一个给定的数，若能找到两个统计量 T_1、T_2，使得 $P(T_1 \leq \theta \leq T_2) = 1 - \alpha$，则称 $[T_1, T_2]$ 为置信水平为 1_α 的置信区间。

- 置信区间的长度 $T_2 - T_1$ 反映了估计精度，$T_2 - T_1$ 越小，估计精度越高。
- α 反映估计的可靠度，α 越小，即 $1 - \alpha$ 越大，估计的可靠度越高，但同时 $T_2 - T_1$ 往往增大，即估计精度降低。
- α 确定以后，置信区间的选取方法不唯一，常选择区间长度最小的一个。
- T_2 称为置信上限，T_1 为置信下限。

求置信区间的步骤:

(1) 寻找一个样本函数 $g(\theta, X_1, X_2, \ldots, X_n)$ 称为枢轴量)，使其含有待估参数，不含其他未知参数，且分布已知，且分布不依赖于待估参数（常由 θ 的点估计出发考虑)。

(2) 给定置信度 $1 - \alpha$，定出常数 a, b 使得 $P(a < g(\theta, X_1, X_2, \ldots, X_n) < b) = 1 - \alpha$ 。

(3) 由 $a < g(\theta, X_1, X_2, \ldots, X_n) < b$, 解得置信区间 T_1, T_2。

(4) 代入样本观测值，得到具体的置信区间。

23.2　一个正态总体参数的区间估计

(1) 均值 μ 的置信区间

- σ^2 已知：$U = \dfrac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \sim N(0,1)$
- σ^2 未知：$T = \dfrac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$

(2) 方差 σ^2 的置信区间

- μ 已知：$\chi^2 = \dfrac{1}{\sigma^2}\sum\limits_{i=1}^{n}(X_i-\mu)^2 \sim \chi^2(n)$
- μ 未知：$\chi^2 = \dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$

23.3　两个正态总体参数的区间估计

(1) 均值差 $\mu_1-\mu_2$ 的置信区间

- σ_1^2,σ_2^2 已知：$U = \dfrac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{\sigma_1^2/n_1^2}+\sqrt{\sigma_2^2/n_2{}^2}} \sim N(0,1)$
- $\sigma_1^2=\sigma_2^2=\sigma^2$ 未知：$T = \dfrac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}} \sim t(n_1+n_2-2)$,

  其中，$S_w = \sqrt{\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}$
- σ_1^2,σ_2^2 未知, 且 $\sigma_1^2 \neq \sigma_2^2$：

  $U = \dfrac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{S_{1n_1}^2/n_1^2}+\sqrt{S_{2n_2}^2/n_2{}^2}} \sim N(0,1)$
- σ_1^2,σ_2^2 未知, 但 $n_1=n_2$：

  令 $Z_i = X_i - Y_i$, 则 $Z \sim N(\mu_1-\mu_2,\sigma_1^2+\sigma_2^2)$

  $S_Z = \dfrac{1}{n-1}\sum\limits_{i=1}^{n}[(X_i-Y_i)-(\bar{X}-\bar{Y})]^2$,

  取 $T = \dfrac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{S_Z/\sqrt{n}} \sim t(n-1)$

(2) 方差比 $\dfrac{\sigma_1^2}{\sigma_2^2}$ 的置信区间

- μ_1, μ_2 已知，则

$$F = \frac{\dfrac{1}{n_1}\dfrac{\sum\limits_{i=1}^{n_1}(X_i-\mu_1)^2}{\sigma_1^2}}{\dfrac{1}{n_2}\dfrac{\sum\limits_{i=1}^{n_2}(Y_i-\mu_2)^2}{\sigma_2^2}} = \frac{\dfrac{n_2}{n_1}\dfrac{\sum\limits_{i=1}^{n_1}(X_i-\mu_1)^2}{\sum\limits_{i=1}^{n_2}(Y_i-\mu_2)^2}}{\dfrac{\sigma_1^2}{\sigma_2^2}} \sim F(n_1, n_2)$$

- μ_1, μ_2 未知，则

$$F = \frac{S_{1n_1}^2 / S_{2n_2}^2}{\sigma_1^2/\sigma_2^2} \sim F(n_1-1, n_2-1)$$

23.4 单侧置信上下限:

1. 与双侧置信区间类似，对于给定的置信度，满足 $P(\theta > \hat{\theta}_1) = 1-\alpha$，则称随机区间 $(\hat{\theta}_1, +\infty)$，为 θ 的置信度为 $1-\alpha$ 的单侧置信区间，$\hat{\theta}_1$ 为 θ 置信度 $1-\alpha$ 的单侧置信下限。单侧置信上限类似。

2. 单侧置信区间的计算步骤法与双侧置信区间相同。

本章精选习题:

求矩估计值以及极大似然估计值：习题 3（离散型随机变量），习题 4（连续型随机变量），习题 6（利用极大似然不变性）

估计量的评价：习题 8（无偏性和有效性），习题 14（相合性），习题 9（利用无偏性计算估计量）

区间估计：习题 18（计算区间长度），习题 24（计算标准差之比的置信区间），习题 26（计算单侧置信下限），习题 22（两样本标准差相同）

Part VII

假设检验

心得总结：

根据往年的出题经验，概率论第七章部分一般只会出一个大题，但也一定会有一个大题。所以这章属于必考内容，考察重点在于各类名词的含义理解与假设检验的流程。此处题型非常固定，因而是重点并不是难点，希望大家不要在此处失分。

24 假设检验的概念

1. 在总体的分布函数完全未知或只知其形式、但不知其参数的情况下，为了推断总体的某些性质，对总体提出某项统计假设（假设），假设是否正确需要利用样本（采样数据）对假设的真假进行判断，即假设检验。

2. 当总体的分布类型已知，对总体分布中未知参数的检验称为参数假设检验，当总体分布类型未知，对总体分布类型或分布性质的检验称为非参数假设检验。

3. 参数假设检验：对总体分布的某个参数（通常是总体的某个数字特征）提出假设，用总体的样本值检验该假设是否成立。

24.1 假设检验的基本原理

要给出检验假设 H_0 的一个检验法，主要是在 H_0 成立的前提下找到一个适当的小概率事件，按照实际推断原理，在一次具体采样后如果这个小概率事件发生了，就拒绝原假设 H_0，如果这个小概率事件没有发生，就接受原假设 H_0。根据实际问题事先给定一个显著性水平 α，事件发生的概率不超过 α 时，就认为是一个小概率事件。

24.2 两类错误

第一类错误（拒真错误）：原假设 H_0 为真，但拒绝了 H_0。

第二类错误（存伪错误）：原假设 H_0 不真，但接受了 H_0。

24.3 原假设和备择假设的不平等性

H_0 和 H_1 地位不平等，H_0 受特殊保护，因此接受了 H_0，并不一定说明 H_0 一定为真。因为有较大概率实际不真也误认为是真的。但是若在很小的 下仍拒绝 H_0，则有充分理由否定 H_0。接受 H_0 只说明目前样本提供的信息还不足以否定 H_0。

24.4 假设检验的一般步骤

1. 恰当地提出原假设 H_0 和备择假设 H_1。

2. 构造检验统计量 Z，并求出在 H_0 成立的前提下确定 Z 的概率分布，要求 Z 的分布不依赖于任何未知参数（即相当于根据已知样本可以计算出的具体量）。

3. 确定拒绝域。在 H_0 成立时，以不利于 H_0（亦即有利于 H_1）的方式设定拒绝域的形式 W，再根据给定的水平 α 和 Z 的分布，由 $P\{W|H_0$ 为真 $\}$ 确定拒绝域 W。

4. 进行一次抽样，根据得到的样本值与上面确定的拒绝域 W，对 H_0 作出拒绝或接受的判断。

例题 8. 在某年级学生中抽测 9 名跳远成绩, 得样本均值 $\bar{x} = 4.38$, 设跳远成绩 X 服从正态分布且 $\sigma = 0.3$，问是否可以认为该年级学生跳远平均成绩为 $\mu = 4.40$ 米?($\alpha = 0.10$)

解：根据题意提出假设 H_0：$\mu = 4.40$ H_1：$\mu \neq 4.40$

因为 σ 已知，(因而应采用方差已知条件下的均值检验)，构造枢轴量: $U = \dfrac{\bar{X} - 4.40}{\sigma / \sqrt{n}}$

从而拒绝域为：$|u| \geq u_{\alpha/2}$;

查标准正态分布表得：$u_{\alpha/2} = u_{0.05} = 1.645$；

计算可知，统计量 $|U| = |\dfrac{\bar{x} - 4.40}{0.3/3}| = 0.2 < u_{0.05} = 1.645$，并不在拒绝域范围内所以接受 H_0，认为 $\mu = 4.40$ 米，即在 $\alpha = 0.10$ 的置信度下可以认为该年级学生跳远平均成绩为 4.40 米.

25 正态总体参数的假设检验

综述：此处需要牢记各类情况下构造的枢轴量的形式以及相应情况下的拒绝域形式（不等号方向等）

25.1 单个正态总体参数的假设检验

1. 正态总体 $N(\mu, \sigma^2)$ 中均值 μ 的检验

   (1) σ^2 已知——u 检验：$U = \dfrac{\sqrt{n}(\bar{X} - \mu_0)}{\sigma} \sim N(0, 1)$

   (2) σ^2 未知——t 检验：$T = \dfrac{\bar{X} - \mu_0}{S / \sqrt{n}} \sim t(n-1)$

2. 正态总体 $N(\mu, \sigma^2)$ 中方差 σ^2 的检验 χ^2——检验

   (1) μ 已知：$\chi^2 = \dfrac{1}{\sigma_0^2} \sum_{i=1}^{n} (X_i - \mu)^2 \sim \chi^2(n)$

   (2) μ 未知：$\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2} \sim \chi^2(n-1)$

25.2 两个正态总体参数的假设检验

1. 1. 两总体均值差的检验——t 检验

   (1) σ_1^2, σ_2^2 已知：$U = \dfrac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n_1^2} + \sqrt{\sigma_2^2/n_2{}^2}} \sim N(0, 1)$

(2) $\sigma_1^2=\sigma_2^2=\sigma^2$ 未知：$T=\dfrac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\sim t(n_1+n_2-2)$,

其中，$S_w=\sqrt{\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}$

(3) σ_1^2,σ_2^2 未知，且 $\sigma_1^2\neq\sigma_2^2$：

$U=\dfrac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{S_{1n_1}^2/n_1^2}+\sqrt{S_{2n_2}^2/n_2{}^2}}\sim N(0,1)$

(4) σ_1^2,σ_2^2 未知，但 $n_1=n_2$：

令 $Z_i=X_i-Y_i$, 则 $Z\sim N(\mu_1-\mu_2,\sigma_1^2+\sigma_2^2)$

$S_Z=\dfrac{1}{n-1}\sum\limits_{i=1}^{n}[(X_i-Y_i)-(\bar{X}-\bar{Y})]^2$,

取 $T=\dfrac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{S_Z/\sqrt{n}}\sim t(n-1)$

2. 两总体方差比 $\dfrac{\sigma_1^2}{\sigma_2^2}$ 的检验——F 检验

(1) μ_1,μ_2 已知，则

$$F=\frac{\dfrac{1}{n_1}\dfrac{\sum\limits_{i=1}^{n_1}(X_i-\mu_1)^2}{\sigma_1^2}}{\dfrac{1}{n_2}\dfrac{\sum\limits_{i=1}^{n_2}(Y_i-\mu_2)^2}{\sigma_2^2}}=\frac{\dfrac{n_2}{n_1}\dfrac{\sum\limits_{i=1}^{n_1}(X_i-\mu_1)^2}{\sum\limits_{i=1}^{n_2}(Y_i-\mu_2)^2}}{\dfrac{\sigma_1^2}{\sigma_2^2}}\sim F(n_1,n_2)$$

(2) μ_1,μ_2 未知，则

$$F=\frac{S_{1n_1}^2/S_{2n_2}^2}{\sigma_1^2/\sigma_2^2}\sim F(n_1-1,n_2-1)$$

3. 双边假设检验与单边假设检验

(1) 双边假设检验

- $H_0:\theta=\theta_0$、$H_0:\theta\neq\theta_0$

(2) 单边假设检验

- $H_0:\theta=\theta_0$、$H_0:\theta>\theta_0$
- $H_0:\theta=\theta_0$、$H_0:\theta<\theta_0$
- $H_0:\theta\leq\theta_0$、$H_0:\theta>\theta_0$
- $H_0:\theta=\geq\theta_0$、$H_0:\theta<\theta_0$

例题 9. 单样本例题分析：

某厂生产的某种型号的电池，其寿命长期以来服从方差为 $\sigma^2 = 5000$(小时的平方) 的正态分布, 现有一批这种电池, 从它的生产情况来看, 寿命的波动性有所改变. 现随机取 26 只电池, 测出寿命的样本方差 $s^2 = 9200$(小时 ^{2}). 问根据这一数据能否推断这批电池的寿命的波动性较以往的有显著的变化 (取 $\alpha = 0.02$)？

解：根据题意提出假设：$H_0 : \sigma^2 = \sigma_0^2 = 5000, H_1 : \sigma^2 \neq \sigma_0^2$

检验统计量:（检验单样本方差是否发生变化，使用卡方检验法）

$\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$ ，其中：$n = 26, \sigma_0^2 = 5000$

拒绝域为:$\chi^2 \leq \chi^2_{n-1}(1-\alpha/2)$ 或 $\chi^2 \geq \chi^2_{n-1}(\alpha/2)$。

查表得:$\chi^2_{n-1}(1-\alpha/2) = \chi^2_{25}(0.99) = 11.523$ $\chi^2_{n-1}(\alpha/2) = \chi^2_{25}(0.01) = 44.313$

由样本观察值算得 $\chi^2 = \dfrac{(n-1)s^2}{\sigma_0^2} = 46 > 44.313,$

所以拒绝 H_0，即认为在显著水平 $\alpha 0.02$ 下，可以认为这批电池的寿命的波动性较以往有显著的变化。

点评：单样本问题格式较为固定，熟记各类枢轴量构造方法即可。

例题 10. 双样本例题分析：

根据资料用某种旧安眠药时，平均睡眠时间为 20.8 h，标准差为 1.6 h。有一种新安眠药，据说在一定剂量下，能比旧安眠药平均增加睡眠时间 3 h。为了检验这个说法是否正确，收集到一组使用新安眠药的睡眠时间 (单位：h) 为：26.7，22.0，24.1，21.0，27.2，25.0，23.4。试问：从这组数据能否说明新安眠药已达到新的疗效?(假定睡眠时间服从正态分布，$\alpha = 0.05$)

解：由题：两样本的总体方差未知，因而选取假设检验中的 t 检验

第一步：提出原假设和备择假设 $H_0 : \mu = 23.8, H_1 : \mu \neq 23.8$

第二步：构造检验统计量为 t:

$$t = \frac{\bar{x} - 23.8}{s/\sqrt{7}} \sim t(5)$$

拒绝域为双边：

$$W : |t| > T_{1-\alpha/2}(5)$$

第三、四步：查表得

$$t_{0.975}(5) = 2.571 \qquad W: t > 2.571$$

第五步：计算统计量 t，有

$$t = \frac{24.2 - 23.8}{2.295/\sqrt{7}}$$

计算可得，$t = 0.46 < 2.571$ (t 未落入拒绝域)

因而接受原假设, 即此新安眠药已达到新的疗效。

点评：这是一道典型的双样本例题，题型仍然十分固定，考试时很有可能就是这类题型，希望各位读者做到烂熟于心，方可取胜千里之外。

25.3 大样本总体参数的假设检验

综述：此块内容并不作为重点考察。由于大样本的特殊性，往往只需要知道对于大样本可以将未知分布归结为正态分布（中心极限定理），此块内容只需掌握书上有关内容即可，并不作为重点备考对象。

样本抽样分布：多次二项分布抽样 $np > 5$ 时 (p 为小于 1 的数，np 大于 5 表示 n 的值比较大，表示这是一个大样本)，该样本抽样分布满足正态分布。

例题 11. 补充例题：检验一个假设，超过 30% 美国家庭拥有互联网接入，显著性水平 5%。已知采集了 150 个家庭作为样本，结果 57 家拥有接入。

解：零假设：美国家庭网络接入 <=30%；备择假设: 美国家庭网络接入 >30%

分析：我们要根据零假设得到一个总体中的占比值，在这个假设下，看 150 户中有 57 湖接入网络的概率是多少。如果该概率小于 5%，我们就拒绝零假设，承认备择假设。

样本比例的均值：$57/150 = 0.38$

样本标准差：$0.38 * 0.62 = 0.2356$

零假设成立时，总体均值为 0.3，总体标准差为 $\sqrt{0.3 \times 0.7} = \sqrt{0.21}$

因为零假设下，$np = 150 \times 0.3 > 5$, 我们认为零假设下的抽样分布满足正态分布。所以抽样分布均值 =0.3; 标准差 = 总体标准差/ $\sqrt{150}$ = $\sqrt{0.21}/\sqrt{50} = 0.037$

求样本均值与抽样分布均值之间的标准差数，即：$(0.38 - 0.3)/0.037 = 2.14$ 个

（注：这里就是构造枢轴量的另一种说法）

查询 Z 分布表，5% 的概率为 1.65 个标准差。而 2.41 > 1.65. 即零假设下，样本均值距总体均值的距离大于 5% 的概率下的标准差距离，也就是样本均值落入小于 5% 概率下的均值分布，因而应当拒绝零假设。

最后，祝同学们在概率论考试中取得好成绩！

启德书院朋辈导师团

文治学辅与发展中心

钱学森书院学业辅导中心

仲英学业辅导中心及薪火工作室

1st Adition

Contributers

仲英书院学业辅导中心

钱学森书院学业辅导中心

文治学辅与发展中心

启德书院朋辈导师团

关于本材料内容的任何建议及错误指正，欢迎向交小权QQ反映，我们将在下一版对相关内容进行订正、修改。听说有效反馈还会收到精美小权礼物一份哟~