

概率论与数理统计

习题详解

刘海燕 编

(下册)



西南交通大学出版社

<http://press.swjtu.edu.cn>

概率论与数理统计
习 题 详 解

下 册

刘海燕



西南交通大学出版社

· 成 都 ·

图书在版编目 (C I P) 数据

概率论与数理统计习题详解/刘海燕等编, --成都:
西南交通大学出版社, 2000.9
ISBN 7-81057-449-3

I. 概... II. 刘... III. ①概率论 - 研究生 - 入学考试 - 解题 ②数理统计 - 研究生 - 入学考试 - 解题
IV.021 - 44

中国版本图书馆 CIP 数据核字 (2000) 第 21005 号

概率论与数理统计
习 题 详 解
下 册
刘海燕 编

*

出 版 人 宋绍南
责任编辑 戴本文
封面设计 唐利群

西南交通大学出版社出版发行
(成都交大路 148 号 邮政编码: 610031 发行科电话: 7600564)

<http://press.swjtu.edu.cn>

E-mail: cbs@center2.swjtu.edu.cn

四川森林印务有限责任公司印刷

*

开本: 850mm × 1168mm 1/32 印张: 18.625

字数: 473 千字 印数: 1 ~ 4000 册

2000 年 9 月第 1 版 2000 年 9 月第 1 次印刷

ISBN 7-81057-449-3/O · 039

套价: 27.00 元

前 言

概率论与数理统计是研究随机现象的统计规律性的科学，它不同于确定性科学的学习，学习起来有一定难度。特别是初学者，往往在一些概念的理解上停留很久，或是对自己完成的练习解答感到毫无把握，或是感到困惑不安，这时候他们就特别需要有人给与肯定或否定的回答，帮助他们能进一步学习，否则一些问题就像一座山一样压得喘不过气来。

为了帮助初学者能尽快学好概率论与数理统计，特别是在做练习与实际应用中解决一些困难，我们编写了《概率论与数理统计习题详解》一书，本书不仅系统地介绍了概论率与数理统计各部分内容的基本原理、思想和方法，更侧重展示解题的方法与技巧，有一些范例还是概率论与数理统计内容的扩充和引伸，且为兼顾一些读者报考硕士研究生的需要，本书还收集了历年考研试卷中概率论与数理统计方面的试题，并给出全部的参考解答。希望通过本书的学习，读者能牢固树立概率论与数理统计的基本概念与思想，学会基本的应用方法。

本书分为上、下两册，上册为概率论部分内容，由李裕奇负责编写，下册分为数理统计与历年考研试题解答两部分内容，由刘海燕负责编写。本书的习题基本取自李

裕奇、刘海燕编《应用概率论与数理统计》一书，经过整理与扩充，共收录了300余道概率论方面习题，160余道数理统计方面习题，140余道考研试题，分门别类较详尽地阐述了概率论与数理统计的解题方法与技巧，以帮助初学者提高分析问题与解决问题的能力。

本书在编写过程中得到了西南交通大学应用数学系及概率统计教研室同行的热情帮助与支持，特别是得到了西南交通大学出版社的鼎力协助，才使这本书得以顺利出版，作者在此深表感谢。

由于作者的水平与学识所限，书中难免有错误与不当之处，敬请读者指正。

作 者

2000年9月于成都

目 录

(上 册)

第一部分	概率论习题详解	1
第一章	概率论的基本概念	3
第二章	随机变量及其分布	81
第三章	多维随机变量及其分布	154
第四章	随机变量的数字特征	223
第五章	大数定律与中心极限定理	280

(下 册)

第二部分	数理统计习题详解	297
第六章	样本及抽样分布	299
第七章	参数估计	330
第八章	假设检验	382
第九章	线性统计推断	434
第三部分	历年概率论与数理统计研究生入学 考试试题详解	507
第十章	历年概率论与数理统计研究生入学 考试试题详解	509

第二部分

数理统计习题详解

第六章 样本及抽样分布

一、基本内容

1. 数理统计的基本概念

数理统计中被研究的对象的全体叫做总体；组成总体的每个单元(或元素)叫做个体。

从总体中抽取的一部分个体叫做样本；样本中所包含的个体的数量叫做样本容量。

从总体中随机地、独立地抽取样本叫做简单随机抽样，这样得到的样本叫做简单随机样本，今后提到的抽样及样本都是指简单随机抽样及简单随机样本。

从总体 X 中抽取容量为 n 的样本， n 次抽样的结果应分别看作是 n 个随机变量 $X_1, X_2, \dots, X_n$ ，它们相互独立，并且与总体 X 服从相同的分布。具体得到的是 n 个观测值 $x_1, x_2, \dots, x_n$ ，即在第 i 次抽样时随机变量 X_i 取得观测值 $x_i (i = 1, 2, \dots, n)$ 。

2. 统计量

(1) 统计量的概念

设样本 $x_1, x_2, \dots, x_n$ 的函数 $g(x_1, x_2, \dots, x_n)$ 中不含有任何未知参数，则样本函数 $g(x_1, x_2, \dots, x_n)$ 叫做统计量。

(2) 数理统计中最常用的统计量

1° 样本平均值(简称样本均值)

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

2° 样本方差

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

3° 样本标准差

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

4° 样本 k 所原点矩

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

特别样本一阶原点矩就是样本均值，即

$$\nu_1 = \bar{X}.$$

5° 样本 k 阶中心矩

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

特别的，样本一阶中心矩恒等于零，即

$$\mu_1 = 0.$$

样本二阶中心矩记作 $\hat{\sigma}^2$ ，即

$$\hat{\sigma}^2 = \mu_2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2.$$

3. 数理统计中常用的分布

(1) χ^2 分布

设随机变量 $X_1, X_2, \dots, X_n$ 相互独立，并且 $X_i \sim N(0, 1)$ ， $i = 1, 2, \dots, n$ ，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，记作 $\chi^2 \sim \chi^2(n)$ ；其概率密度为

$$f_{\chi^2}(x) = \begin{cases} \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, & x > 0, \\ 0, & x \leq 0. \end{cases}$$

对不同的自由度 n 及不同的数 α ($0 < \alpha < 1$), $\chi^2_{\alpha}(n)$ 为满足等式

$$P(\chi^2 \geq \chi^2_{\alpha}(n)) = \int_{\chi^2_{\alpha}(n)}^{+\infty} f_{\chi^2}(x) dx = \alpha$$

的一个数值, 此数值称为上 α 百分位点.

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

(2) t 分布

设随机变量 X 与 Y 相互独立, 并且 $X \sim N(0, 1)$, $Y \sim \chi^2(n)$, 则随机变量

$$T = \frac{X}{\sqrt{Y/n}}$$

服从自由度为 n 的 t 分布, 记作 $t \sim t(n)$; 其概率密度为

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

对不同的自由度 n 及不同的数 α ($0 < \alpha < 1$), $t_{\alpha}(n)$ 为满足等式

$$P(t \geq t_{\alpha}(n)) = \int_{t_{\alpha}(n)}^{+\infty} f_t(x) dx = \alpha$$

的数值, 此数值 $t_{\alpha}(n)$ 称为上 α 百分位点.

因为, 如 $n \rightarrow \infty$ 时, t 分布趋于标准正态分布 $N(0, 1)$, 一般当 $n > 50$ 时, $t_{\alpha}(n) \approx z_{\alpha}$, 其中 z_{α} 为

$$P(U \geq z_{\alpha}) = \frac{1}{\sqrt{2\pi}} \int_{z_{\alpha}}^{+\infty} e^{-\frac{x^2}{2}} dx = \alpha,$$

注意 $U \sim N(0, 1)$, $EU(n) = 0$.

(3) F 分布

设随机变量 X 与 Y 独立, 并且 $X \sim \chi^2(n_1)$, $Y \sim \chi^2(n_2)$, 则随机变量

$$F = \frac{X/n_1}{Y/n_2}$$

服从自由度为 (n_1, n_2) 的 F 一分布, 记作 $F \sim F(n_1, n_2)$. 其概率密度为

$$f_F(x) = \begin{cases} \frac{\Gamma\left(\frac{n_1+n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)} n_1^{\frac{n_1}{2}} n_2^{\frac{n_2}{2}} \frac{x^{\frac{n_1}{2}-1}}{(n_1x+n_2)^{\frac{n_1+n_2}{2}}}, & x > 0, \\ 0, & x \leq 0, \end{cases}$$

$F_\alpha(n_1, n_2)$ 是满足等式

$$P(F \geq F_\alpha) = \int_{F_\alpha}^{+\infty} f_F(x) dx = \alpha$$

的一个数值, $F_\alpha(n_1, n_2)$ 称为 F 分布的上百分位点. 关于数值 $F_\alpha(n_1, n_2)$, 有下面的等式:

$$F_{1-\alpha}(n_1, n_2) = \frac{1}{F_\alpha(n_2, n_1)}.$$

4. 正态总体统计量的分布

(1) 单个正态总体统计量的分布

设总体 $X \sim N(\mu, \sigma^2)$, 从总体 X 中抽取样本 $X_1, X_2, \dots, X_n$, 则有:

1° 样本均值 $\bar{X}$ 服从正态分布 $N\left(\mu, \frac{\sigma^2}{n}\right)$, 即

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

2° 统计量 $\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$ 服从标准正态分布, 即

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

3° 统计量 $\frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \mu)^2$ 服从自由度为 n 的 χ^2 分布, 即

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

4° 统计量 $\frac{(n-1)S^2}{\sigma^2} = \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{X})^2$ 服从自由度为 $n-1$ 的 χ^2 分布, 即

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

5° 样本均值 $\bar{X}$ 与样本方差 S^2 相互独立.

6° 统计量 $\frac{\bar{X} - \mu}{S / \sqrt{n}}$ 服从自由度为 $n-1$ 的 t -分布, 即

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

(2) 两个正态总体统计量的分布

设总体 $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$, 从总体 X 中抽取样本 $X_1, X_2, \dots, X_{n_1}$, 从总体 Y 中抽取样本 $Y_1, Y_2, \dots, Y_{n_2}$, 样本均值分别记作

$$\bar{x} = \frac{1}{n_1} \sum_{i=1}^{n_1} x_i, \quad \bar{y} = \frac{1}{n_2} \sum_{j=1}^{n_2} y_j,$$

样本方差分别记作

$$S_1^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (X_i - \bar{X})^2, \quad S_2^2 = \frac{1}{n_2 - 1} \sum_{j=1}^{n_2} (Y_j - \bar{Y})^2,$$

则有:

1° 统计量 $\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$, 服从标准正态分布,

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

2° 当 $\sigma_1 = \sigma_2$ 时, 统计量 $\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} S_w}$ 服从自由度为

$n_1 + n_2 - 2$ 的 t 分布, 即

$$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} S_w} \sim t(n_1 + n_2 - 2).$$

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

3° 统计量 $\frac{\sum_{i=1}^{n_1} (X_i - \mu_1)^2 / n_1 \sigma_1^2}{\sum_{i=1}^{n_2} (Y_i - \mu_2)^2 / n_2 \sigma_2^2}$ 服从自由度为 (n_1, n_2) 的 F

分布, 即

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

4° 统计量 $\frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2}$ 服从自由度为 $(n_1 - 1, n_2 - 1)$ 的 F 分布,

即

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

二、基本要求

1. 理解总体、个体、样本和统计量的概念、掌握样本均值、样本方差的计算.
2. 理解样本分布函数的实际意义, 会计算样本分布函数, 会画直方图.
3. 了解 χ^2 分布、 t 分布、 F 分布的定义, 并会查表计算.
4. 了解正态总体的某些常用统计量的分布.

三、习题详解

6.1 在一小时内观测电话用户对总机的呼唤次数, 按每一分钟统计, 得到数据如下:

0	0	1	2	1	2	2	3	4	1	1	2
0	2	3	1	3	1	4	1	0	1	2	5
3	1	2	2	2	4	1	2	0	2	3	4
2	5	0	2	2	4	3	1	1	3	2	3
4	6	1	1	0	3	2	1	3	1	2	0

计算样本均值、样本标准差及样本方差.

解 为了计算方便, 首先对上述 60 个数据进行整理, 写出一分钟内电话呼唤次数 X 的频数分布表:

一分钟内的呼唤次数 x_i	0	1	2	3	4	5	6	总计
频 数 m_i	8	16	17	10	6	2	1	60

然后利用具有统计计算功能的计算器进行计算.

把计算器设置于统计计算状态, 按上述频数分布表, 把全部数据分 7 次存入计算器中,

- 按 $\bar{x}$ 键, 得样本均值 $\bar{x} = 2$,
- 按 S 键, 得样本标准差 $s \approx 1.402$,
- 按 S^2 键, 得样本方差 $s^2 \approx 1.966$.

6.2 从一个总体中抽取了容量为 5 的一个样本，具体观察值为 $(-2.8, -1, 1.5, 2.1, 3.4)$

(1) 求极差；(2) 求经验分布函数，并作出图形。

解 (1) 从样本观察值得顺序统计量，极差统计量

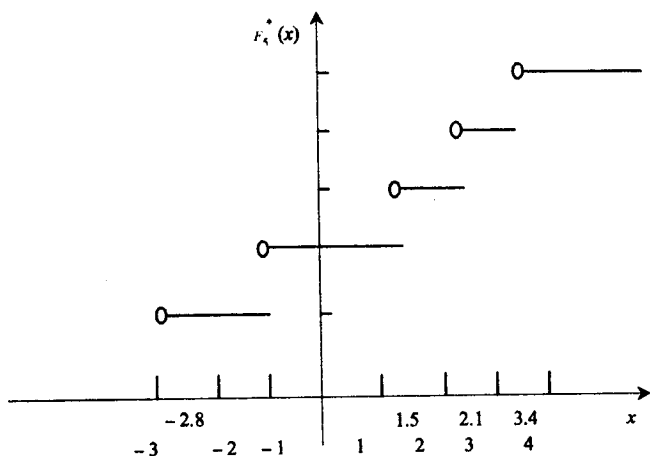
$$x_1^*, x_2^*, x_3^*, x_4^*, x_5^*, D_n^* = x_5^* - x_1^*$$

$$-2.8, -1, -1.5, 2.1, 3.4, D_n^* = 3.4 - (-2.8) = 6.2.$$

(2) 经验分布函数

$$F_5^*(x) = \begin{cases} 0, & x \leq -2.8, \\ \frac{1}{5}, & -2.8 < x \leq -1, \\ \frac{2}{5}, & -1 < x \leq 1.5, \\ \frac{3}{5}, & 1.5 < x \leq 2.1, \\ \frac{4}{5}, & 2.1 < x \leq 3.4, \\ 1, & x > 3.4. \end{cases}$$

图形



6.3 设总体 X 服从贝努里分布 $B(1, p)$, $P(X=1)=p$, 其中 p 是未知参数 $X=(X_1, X_2, \dots, X_5)$ 是从中抽取的一简单随机子样.

(1) 写出 X 的概率分布;

(2) 指出 $X_1 + X_2$, $\min_{1 \leq i \leq 5} X_i$, $X_5 + 2p$ 之中, $(X_5 - X_1)^2$ 之中, 哪些是统计量, 哪些不是统计量, 为什么?

(3) 如果 X 的一个观察值是 $(0, 1, 0, 1, 1)$ 计算它的子样平均值和子样方差.

解 (1) X 的概率分布即为 $(X_1, X_2, \dots, X_5)$ 联合分布

$$\begin{aligned} & P\{X_1 = x_1, X_2 = x_2, X_3 = x_3, X_4 = x_4, X_5 = x_5\} \\ &= p\{X_1 = x_1\}p\{X_2 = x_2\} \cdots p\{X_5 = x_5\} \\ &= p^{x_1}(1-p)^{1-x_1} \cdot p^{x_2}(1-p)^{1-x_2} \cdots p^{x_5}(1-p)^{1-x_5} \\ &= p^{\sum_{i=1}^5 x_i} (1-p)^{5-\sum_{i=1}^5 x_i}, \end{aligned}$$

其中 x_i 只取 0, 1 两个数值.

(2) $X_1 + X_2$, $\min_{1 \leq i \leq 5} X_i$, $(X_5 - X_1)^2$ 是统计量, 因不含未知参数; $X_5 + 2p$ 含有未知参数 p , 故 $X_5 + 2p$ 不是统计量.

(3) 如果 X 的一个观察值是 $(0, 1, 0, 1, 1)$, 计算它的子样平均值和子样方差.

$$\text{解 } \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{3}{5}, \quad s^2 = \frac{1}{n-1} \sum_{i=1}^5 (x_i - \bar{x})^2 = 0.3.$$

6.4 试证: 当 $c = \bar{x}$ 时, $\sum_{i=1}^n (x_i - c)^2$ 达到最小.

$$\begin{aligned} \text{证 因为 } \sum_{i=1}^n (x_i - c)^2 &= \sum_{i=1}^n [(x_i - \bar{x}) + (\bar{x} - c)]^2 \\ &= \sum_{i=1}^n (x_i - \bar{x})^2 + 2 \sum_{i=1}^n (x_i - \bar{x})(\bar{x} - c) + \sum_{i=1}^n (\bar{x} - c)^2, \end{aligned}$$

$$\text{而 } \sum_{i=1}^n (x_i - \bar{x})(\bar{x} - c) = (\bar{x} - c) \sum_{i=1}^n (x_i - \bar{x}) = 0,$$