第七章　统计量及其分布

图文并茂详细学习笔记

依据课程 PPT 课件整理

2026 年 11 月 16 日

摘　要

本章是数理统计（第七至九章）的开篇，实现从概率论到统计推断的过渡。主要内容包括：数理统计的任务与分类；总体、个体、样本与样本值的概念；样本分布（联合分布函数等于边际分布函数之积）；统计量的定义及常用样本矩——样本均值、样本方差、样本 k 阶原点矩与中心矩；样本均值与样本方差的数字特征 $E\overline{X}=EX$、$D\overline{X}=\frac{1}{n}DX$、$ES^2=DX$。考试常考：样本与统计量概念的辨析、样本联合分布的写法、样本均值与样本方差期望方差的计算与证明。

目录

本章知识框架 **2**

1　数理统计概述【了解】 **2**

1.1　数理统计的任务与分类 2

2　总体与样本 **2**

2.1　总体与个体【重点】 2

2.2　样本与样本值【高频考点】 3

2.3　样本分布【重点】 3

3　样本矩和统计量 **4**

3.1　统计量的定义【高频考点】 4

3.2　样本矩（样本的矩统计量）【高频考点】 4

3.3　样本均值与样本方差的数字特征【高频考点】 5

4　本章总结 **6**

4.1　必背公式汇总 6

4.2　重要求解方法与套路 6

4.3　重要题型归纳 7

4.4　易错点清单 7

本章知识框架

数理统计概述（描述统计学 / 推断统计学：参数估计、假设检验、方差分析、回归分析）
→ 总体与样本：总体即随机变量 X → 样本为 n 个相互独立且与 X 同分布的随机变量 → 样本分布为边际分布之积
→ 统计量：不含未知参数的样本函数 → 常用样本矩：样本均值、样本方差、k 阶原点矩、k 阶中心矩
→ 样本矩的数字特征：样本均值的期望与方差、样本方差的期望（等于总体方差）

1 数理统计概述【了解】

1.1 数理统计的任务与分类

数理统计学是一门**以数据为基础**的学科。它的任务就是**如何获得样本，并利用样本对事物的某些未知方面进行分析、推断，然后作出一定的决策**。

数理统计分为两大类：

- **描述统计学**：进行观测或试验，以取得有代表性的观测值。数据的收集（调查数据和实验数据）可根据数据本身的特点有多种不同的方法和设计；
- **推断统计学**：对已取得的观测值进行整理分析，作出推断决策，从而找出所研究对象的规律性。主要包括**参数估计**、**假设检验**、**方差分析**、**回归分析**（本课程第七至九章的核心内容）。

注意

本部分以了解为主：知道数理统计”用样本推断总体”的基本思想即可，历史沿革（凯特勒、费希尔等）一般不考。

2 总体与样本

2.1 总体与个体【重点】

- **总体**——具有一定共同属性的研究对象的全体；
- **个体**——组成总体的每一个元素。

例如：研究某批灯泡的平均寿命（平均耐用时数），全体灯泡的寿命组成了总体，而其中每个灯泡的寿命就是个体。

若研究对象的某一（或某几项）数量指标是一个**随机变量** X（如灯泡的寿命、钢筋的强度），则：

- 总体：随机变量（数量指标）X 的全体取值构成的集合；
- 总体的分布：随机变量 X 的分布。

关键观念：总体就是一个随机变量 X，谈总体就是谈一个分布。

2.2 样本与样本值【高频考点】

从一个总体 X 中随机抽取 n 个个体（有放回的重复抽样）：

- 一次抽样结果：$x_1, x_2, \ldots, x_n$（完全确定的一组数）；
- 再一次抽样结果：$x_1', x_2', \ldots, x_n'$（与上次抽样结果不同）。

抽样的随机性决定了每次抽样结果都会变化。引入随机变量 $X_1, X_2, \ldots, X_n$，把每次抽样结果看成随机变量的取值：

- $X_1, X_2, \ldots, X_n$ 称为来自于总体 X 的 ** 样本容量为 n 的样本 **；
- $x_1, x_2, \ldots, x_n$ 是样本 $X_1, X_2, \ldots, X_n$ 的一组观察值，称为**样本值**。

总结（务必背熟）：

1. 总体就是一个随机变量 X；
2. 样本就是 n 个 ** 相互独立的、与 X 同分布 ** 的随机变量 $X_1, X_2, \ldots, X_n$。

随机抽样：按机会均等的原则，从总体（母体）中选取一些个体进行试验或观察的过程。获得简单随机样本的方法是**简单随机抽样**。

易错提示

- 样本 $X_1, X_2, \ldots, X_n$（大写）是随机变量，样本值 $x_1, x_2, \ldots, x_n$（小写）是确定的一组数，答题时符号大小写不能混；
- 样本的两个本质属性：**相互独立**且**与总体同分布**，后续所有公式（联合分布、期望方差计算）都建立在这两条上。

2.3 样本分布【重点】

若总体 X 具有分布函数 $F(x)$，设 $X_1, X_2, \ldots, X_n$ 为来自于总体 X 的样本，则 $X_1, X_2, \ldots, X_n$ 相互独立，且 X_i 的分布函数：

$$F_{X_i}(x_i) = P\{X_i \le x_i\} = P\{X \le x_i\} = F(x_i)$$

$(X_1, X_2, \ldots, X_n)$ 的分布函数称为**样本分布**，即

$$F(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} F(x_i)$$

例 1 设总体 X 服从参数为 λ 的指数分布，即概率密度为 $f(x) = \lambda e^{-\lambda x}$ $(x > 0)$，$x \le 0$ 时 $f(x) = 0$。$X_1, X_2, \ldots, X_n$ 为来自总体 X 的样本，写出样本的联合概率密度。

解由样本的独立同分布性，X_i 与总体同分布，密度同为 $f(x_i)$，再由独立性，联合密度等于各边际密度之积：

$$f(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} f(x_i) = \lambda^n e^{-\lambda \sum_{i=1}^{n} x_i}, \quad x_i > 0,\ i = 1, 2, \ldots, n$$

其余情形联合密度为 0。

3 样本矩和统计量

数理统计所要解决的问题是：借助总体 X 的样本 $X_1, X_2, \ldots, X_n$ 对总体的性质进行推断，这类问题统称为**统计推断问题**。利用样本推断总体的分布，还需要样本的函数，即由样本构造的统计量。

3.1 统计量的定义【高频考点】

定义设 $X_1, X_2, \ldots, X_n$ 为总体 X 的一个样本，$g(x_1, x_2, \ldots, x_n)$ 为一个**不含总体未知参数**的连续函数，则称 $g(X_1, X_2, \ldots, X_n)$ 为样本的一个**统计量**。

例 2 设 $X_1, X_2, \ldots, X_n$ 为来自总体 $N(\mu, \sigma^2)$ 的样本，其中 μ 已知、σ^2 未知。判断下列各式是否为统计量：

$$T_1 = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad T_2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2, \qquad T_3 = \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \mu)^2$$

解按定义，统计量只能依赖于样本和已知量，**不能含有未知参数**。

- T_1 只是样本的函数，是统计量；
- T_2 含 μ，但 μ 已知，是统计量；
- T_3 含未知参数 σ^2，**不是**统计量。

考场提醒

判断统计量只看一点：表达式中是否出现**未知参数**。出现未知参数就不是统计量；含已知参数或只含样本的都是统计量。

3.2 样本矩（样本的矩统计量）【高频考点】

设 $X_1, X_2, \ldots, X_n$ 为来自于总体 X 的一个样本，称：

样本均值：

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

样本方差：

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

样本 k 阶矩（或 k 阶原点矩）：

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

样本 k 阶中心矩：

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

显然，上述样本矩都是**随机变量**。

如果 $x_1, x_2, \ldots, x_n$ 是样本 $X_1, X_2, \ldots, X_n$ 的一组观察值，则

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \qquad s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$$

$$a_k = \frac{1}{n}\sum_{i=1}^{n} x_i^k, \qquad b_k = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^k$$

分别是 $\overline{X}$、S^2、A_k、B_k 的观察值。

易错提示

- 样本方差 S^2 的分母是 $n-1$（不是 n），而样本二阶中心矩 B_2 的分母是 n，两者关系为 $B_2 = \frac{n-1}{n}S^2$，极易混淆；
- 样本均值、样本方差都是随机变量，可以讨论其期望、方差与分布；代入观察值后才是一个数。

3.3 样本均值与样本方差的数字特征【高频考点】

例 3 设 $X_1, X_2, \ldots, X_n$ 为来自总体 X 的一个样本，$\overline{X}$ 和 S^2 分别为样本均值和样本方差，求 $E(\overline{X})$、$D(\overline{X})$、$E(S^2)$。

解 $X_1, X_2, \ldots, X_n$ 相互独立、同分布。

** 第 1 步：求 $E(\overline{X})$。** 由期望的线性性，

$$E\overline{X} = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} EX_i = \frac{1}{n}\sum_{i=1}^{n} EX = EX$$

** 第 2 步：求 $D(\overline{X})$。** 由独立性，方差可加，

$$D\overline{X} = D\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} DX_i = \frac{1}{n^2}\sum_{i=1}^{n} DX = \frac{1}{n}DX$$

** 第 3 步：求 $E(S^2)$。** 先化简离差平方和，插入 EX_i 与 $E\overline{X}$ 配平方：

$$\sum_{i=1}^{n}(X_i - \overline{X})^2 = \sum_{i=1}^{n}\left[(X_i - EX_i) - (\overline{X} - E\overline{X})\right]^2$$

$$= \sum_{i=1}^{n}\left[(X_i - EX_i)^2 - 2(\overline{X} - E\overline{X})(X_i - EX_i) + (\overline{X} - E\overline{X})^2\right]$$

$$= \sum_{i=1}^{n}(X_i - EX_i)^2 - 2(\overline{X} - E\overline{X})\sum_{i=1}^{n}(X_i - EX_i) + n(\overline{X} - E\overline{X})^2$$

注意 $\sum_{i=1}^{n}(X_i - EX_i) = n(\overline{X} - E\overline{X})$，代入合并得

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

两边取期望：

$$ES^2 = E\left[\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2\right] = \frac{1}{n-1}E\left[\sum_{i=1}^{n}(X_i - EX_i)^2 - n(\overline{X} - E\overline{X})^2\right]$$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n}E(X_i - EX_i)^2 - nE(\overline{X} - E\overline{X})^2\right] = \frac{1}{n-1}\left[\sum_{i=1}^{n}DX_i - nD\overline{X}\right]$$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n}DX - n\cdot\frac{DX}{n}\right] = DX$$

结论（必背）：

$$E\overline{X} = EX, \qquad D\overline{X} = \frac{1}{n}DX, \qquad ES^2 = DX$$

意义：$ES^2 = DX$ 说明样本方差 S^2 是总体方差 DX 的无偏估计（第八章），这正是 S^2 分母取 $n-1$ 而不是 n 的原因。

4 本章总结

4.1 必背公式汇总

1. 样本分布（联合分布函数）：$F(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} F(x_i)$；
2. 统计量：样本 $X_1, X_2, \ldots, X_n$ 的不含未知参数的（连续）函数 $g(X_1, X_2, \ldots, X_n)$；
3. 样本均值：$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$；
4. 样本方差：$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$；
5. 样本 k 阶（原点）矩：$A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$；
6. 样本 k 阶中心矩：$B_k = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^k$；
7. 样本均值的期望与方差：$E\overline{X} = EX$，$D\overline{X} = \frac{1}{n}DX$；
8. 样本方差的期望：$ES^2 = DX$。

4.2 重要求解方法与套路

写出样本联合分布：第 1 步，写出总体的分布律或密度 $f(x)$；第 2 步，利用独立性把联合密度（分布律）写成 $\prod_{i=1}^{n} f(x_i)$；第 3 步，整理成紧凑形式并注明各 x_i 的取值范围。

判断一个量是否为统计量：第 1 步，检查表达式中出现了哪些总体参数；第 2 步，确认这些参数是否全部已知；第 3 步，含未知参数则不是统计量，否则是统计量。

求样本矩的期望与方差：第 1 步，写出统计量的定义式；第 2 步，利用样本独立同分布，用期望线性性、方差可加性化归到 EX、DX；第 3 步，对 S^2 类含 $\overline{X}$ 的统计量，先作配平方分解 $\sum_{i=1}^{n}(X_i-\overline{X})^2=\sum_{i=1}^{n}(X_i-EX_i)^2-n(\overline{X}-E\overline{X})^2$，再逐项取期望。

4.3 重要题型归纳

题型	识别特征	解题套路/所用公式
概念辨析	问总体、样本、样本值、统计量的含义	总体即随机变量 X；样本独立同分布；样本值是确定数值
写样本联合分布	给出总体分布，求样本联合密度/分布律	独立性：$F=\prod_{i=1}^{n}F(x_i)$，密度同理
统计量判断	给出若干样本函数，问哪些是统计量	看是否含未知参数，含则不是
求样本矩数字特征	求 $E\overline{X}$、$D\overline{X}$、ES^2 等	独立同分布 + 期望方差性质；S^2 用配平方分解
与第六章综合	大样本下样本均值的近似分布	$E\overline{X}=\mu$、$D\overline{X}=\frac{\sigma^2}{n}$ + 中心极限定理：$\overline{X}$ 近似服从 $N\left(\mu,\frac{\sigma^2}{n}\right)$

4.4 易错点清单

- 样本 $X_1,\ldots,X_n$ 是随机变量、样本值 $x_1,\ldots,x_n$ 是常数，符号与性质不能混用；
- 样本的两大属性” 相互独立、与总体同分布” 缺一不可，无放回抽样在小总体下不满足独立性；
- 样本方差分母是 $n-1$，样本二阶中心矩分母是 n，两者差一个因子 $\frac{n-1}{n}$；
- 统计量中一旦出现未知参数（如 σ、μ 未知时的 μ），立即不是统计量；
- $D\overline{X}=\frac{1}{n}DX$，常错写成 DX 或 $\frac{1}{n^2}DX$；
- $ES^2=DX$ 依赖分母 $n-1$，若用 $\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$ 则期望为 $\frac{n-1}{n}DX$，不是无偏的；
- 求 $E(S^2)$ 时不能直接对 $(X_i-\overline{X})^2$ 用方差公式，因为 $\overline{X}$ 也是随机的，必须先配平方分解。