标题

Didnelpsun

目录

1　总体与样本

1.1　总体定义

定义：研究对象的全体称为**总体**，组成总体的每一个元素称为**个体**。

1.2　样本

1.2.1　定义

定义：n 个相互独立且域总体 X 有相同概率分布的随机变量 $X_1, X_2, \cdots, X_n$ 所组成的整体 $(X_1, X_2, \cdots, X_n)$ 称为来自总体 X，容量为 n 个一个**简单随机样本**，简称**样本**。一次抽样结果的 n 个具体值 $(x_1, x_2, \cdots, x_n)$ 称为来自样本 $X_1, X_2, \cdots, X_n$ 的一个**观测值**或**样本值**。

在概率论中称为独立同分布，而在数理统计就称为简单随机样本。

1.2.2　分布

对于容量为 n 的样本 $X_1, X_2, \cdots, X_n$ 有如下定理：假设总体 X 的分布函数为 $F(x)$（概率密度为 $f(x)$，或概率分布为 $p_i = P\{X = x_i\}$），则 $(X_1, X_2, \cdots, X_n)$ 的分布函数为 $F(x_1, x_2, \cdots, x_n) = \prod\limits_{i=1}^{n} F(x_i)$。

对于离散型随机变量联合分布：$F(X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n) = \prod\limits_{i=1}^{n} P\{X_i = x_i\}$。

对于连续型随机变量联合概率密度：$f(x_1, x_2, \cdots, x_n) = \prod\limits_{i=1}^{n} f(x_i)$。

2　统计量与分布

2.1　统计量

设 $X_1, X_2, \cdots, X_n$ 来自总体 X 的一个样本，$g(x_1, x_2, \cdots, x_n)$ 为 n 元函数，若 g 中不含有任何未知参数，则称 $g(X_1, X_2, \cdots, X_n)$ 为样本 $X_1, X_2, \cdots, X_n$ 的一个**统计量**。若 $(x_1, x_2, \cdots, x_n)$ 为样本值，则称 $g(x_1, x_2, \cdots, x_n)$ 为 $g(X_1, X_2, \cdots, X_n)$ 的**观测值**。

2.2 常用统计量

- 样本均值：$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$。
- 样本方差：$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$。
- 样本标准差：$S=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2}$。
- 样本 k 阶（原点）矩：$A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k$（$k=1,2,\cdots$）。
- 样本 k 中心矩：$B_k=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^k$（$k=1,2,\cdots$）。

2.3 顺序统计量

2.3.1 概念

将样本 $X_1,X_2,\cdots,X_n$ 的 n 个观测量按其值从小到大的顺序排列，得到 $X_{(1)}\leqslant X_{(2)}\leqslant\cdots\leqslant X_{(n)}$。

随机变量 $X_{(k)}$（$k=1,2,\cdots,n$）称为**第 k 顺序统计量**，其中 $X_{(1)}$ 是最小顺序统计量，而 $X_{(n)}$ 是最大顺序统计量。

$X_{(n)}$ 的分布函数为 $F_{(n)}(x)=[F(x)]^n$，概率密度为 $f_{(n)}(x)=n[F(x)]^{n-1}f(x)$。

证明：$F_{(n)}(x)=P\{X_{(n)}\leqslant x\}=P\{\max\{x_1,\cdots,x_n\}\leqslant x\}=P\{x_1\leqslant x,\cdots,x_n\leqslant x\}=P\{x_1\leqslant x\}\cdots P\{x_n\leqslant x\}=F_{(1)}(x)\cdots F_{(n)}(x)=[F(x)]^n$。

$X_{(1)}$ 的分布函数为 $F_{(1)}(x)=1-[1-F(x)]^n$，概率密度为 $f_{(1)}(x)=n[1-F(x)]^{n-1}f(x)$。

证明:$F_{(1)}(x)=P\{X_{(1)}\leqslant x\}=P\{\min\{x_1,\cdots,x_n\}\leqslant x\}=1-P\{\min\{x_1,\cdots,x_n\}>x\}=1-P\{x_1>x,\cdots,x_n>x\}=1-P\{x_1>x\}\cdots P\{x_n>x\}=1-[1-P\{x_1\leqslant x\}]\cdots[1-P\{x_n\leqslant x\}]=1-[1-F_{(1)}(x)]\cdots[1-F_{(n)}(x)]=1-[1-F(x)]^n$。

2.3.2 性质

设总体 X 的期望 $EX=\mu$，方差 $DX=\delta^2$，样本 $X_1,X_2,\cdots,X_n$ 取自 X，$\overline{X}$ 和 S^2 分别为样本的均值和方差，则：

- $EX_i=\mu$。
- $DX_i=\delta^2$。

- $E\overline{X} = EX = \mu$。
- $D\overline{X} = D\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = \frac{1}{n^2}n\delta^2 = \frac{1}{n}DX = \frac{\delta^2}{n}$。
- $E(S^2) = E\left(\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2\right) = E\left(\frac{1}{n-1}\sum_{i=1}^{n}(x_i^2 - 2x_i\overline{x} + \overline{x}^2)\right) = E\left(\frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - 2\overline{x}\cdot\sum_{i=1}^{n} x_i + n\overline{x}^2\right)\right) = E\left(\frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n\overline{x}^2\right)\right) = \frac{1}{n-1}E\left(\sum_{i=1}^{n} x_i^2 - n\overline{x}^2\right) = \frac{1}{n-1}\left(\sum_{i=1}^{n} Ex_i^2 - nE\overline{x}^2\right) = \frac{n}{n-1}[(Ex_i)^2 + Dx_i - (E\overline{x})^2 - D\overline{x}] = \frac{n}{n-1}\left(\mu^2 + \delta^2 - \mu^2 - \frac{\delta^2}{n}\right) = DX = \delta^2$。

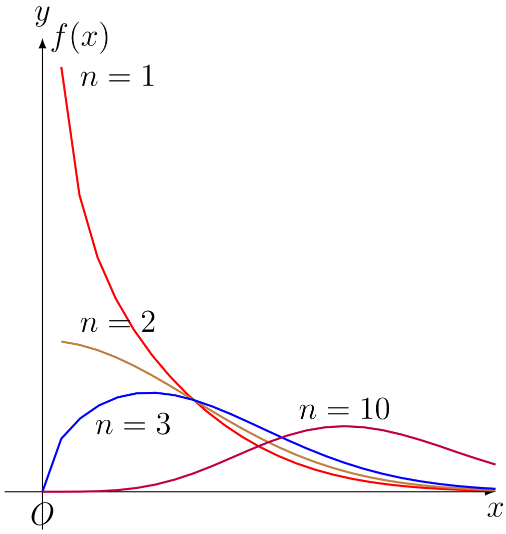
2.4 三大分布

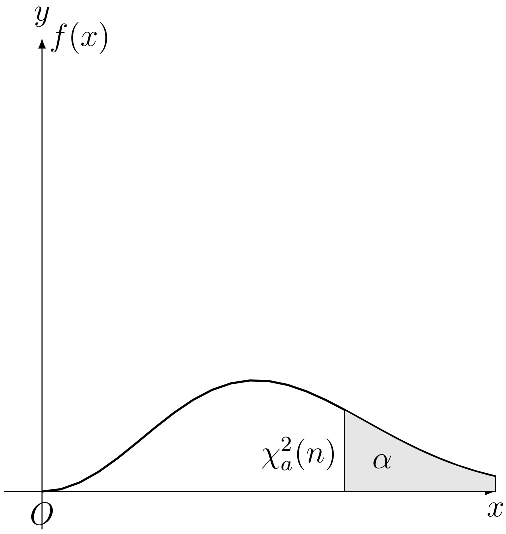
2.4.1 χ^2 分布

2.4.1.1 概念

定义：若随机变量 $X_1, X_2, \cdots, X_n$ 相互独立，且都服从标准正态分布，则随机变量 $X = \sum_{i=1}^{n} X_i^2$ 服从自由度为 n 的 χ^2 分布，记为 $X \sim \chi^2(n)$，特别地 $X_i^2 \sim \chi^2(1)$。

对给定的 α（$0 < \alpha < 1$）称满足 $P\{\chi^2 > \chi_\alpha^2(n)\} = \int_{\chi_\alpha^2(n)}^{+\infty} f(x)\,\mathrm{d}x = \alpha$ 的 $\chi_\alpha^2(n)$ 为 $\chi^2(n)$ 分布的**上 α 分位点**。





2.4.1.2 性质

- 若 $X_1 \sim \chi^2(n_1)$，$X_2 \sim \chi^2(n_2)$，X_1X_2 相互独立，则 $X_1 + X_2 \sim \chi^2(n_1 + n_2)$。一般，若 $X_i \sim \chi^2(n_i)$（$i = 1, 2, \cdots, m$），$X_1, X_2, \cdots, X_m$ 相互独立，则

$$\sum_{i=1}^{m} X_i \sim \chi^2\left(\sum_{i=1}^{m} n_i\right)。$$

- 若 $X \sim \chi^2(n)$，则 $EX = n$，$DX = 2n$。

2.4.2 t 分布

2.4.2.1 概念

也称为学生分布。

若随机变量 $X \sim N(0,1)$，$Y \sim \chi^2(n)$，XY 相互独立，则随机变量 $t = \dfrac{X}{\sqrt{Y/n}}$ 服从自由度为 n 的 t 分布，记为 $t \sim t(n)$。

当 $t \to \infty$ 时，t 分布就是正态分布。其是偶函数，所以 $Et = 0$。

2.4.3 F 分布

若随机变量 $X_1, X_2, \cdots, X_n$

2.5 正态总体下结论

3 参数点估计

3.1 概念

3.2 方法

3.2.1 矩估计法

3.2.2 最大似然估计

3.3 估计量平均标准

3.3.1 无偏性

3.3.2 有效性

最小方差性。

3.3.3 一致性

相合性。

4 参数区间估计与假设检验

4.1 区间估计

4.1.1 概念

4.1.2 正态总体均值的置信空间

4.2 检设检验

4.2.1 思想

4.2.2 正态总体下的六大检验与拒绝域

4.3 两类错误