第五章 基本统计推断

本章重点介绍统计推断的主要方法，包括参数估计和假设检验性检验。其中参数估计分点估计（矩估计与最大似然估计）和区间估计。最大似然估计与一致估计具有不变性，而无偏估计和估计不具有不变性。

5.1 参数估计

一、估计量的优良标准

1.无偏性

2.有效性

3.充分性

4.相容性

二、矩估

矩法求得总体参数θ 的估计称为矩估计，记作 $\hat{\theta}_M$

例5.1设与古典概型的随机试验相应的总体X，以等概率$\dfrac{1}{\theta}$ 取值 $1,2,\cdots,\theta$，试求总体大小θ 的矩估计。

解：现在只有一个参数θ 要估计，因为

$$EX = 1\times\frac{1}{\theta}+2\times\frac{1}{\theta}+\cdots+\theta\times\frac{1}{\theta}=\frac{\theta+1}{2}$$

而对应的样本矩是$\overline{X}$，为了保证样本与总体矩之间的对应，θ 的估计$\hat{\theta}$ 应满足

$$\overline{X}=\frac{\theta+1}{2}$$

解得$\hat{\theta}_M = 2\overline{X}-1$是 θ 的矩估计。

- **例5.2** 设总体$X\sim U(a,b)$, a,b是未知参数，求 a,b 的矩估计$\hat{a},\hat{b}$ 。
- **解**：总体有两个未知参数，须建立两个方程。由于

$$EX=\frac{a+b}{2},DX=\frac{(b-a)^2}{12}$$

- 按对应关系，a,b 的估计应满足

$$\begin{cases}\overline{X}=(\hat{a}+\hat{b})/2\\ S^2=\sum_{k=1}^{n}(X_k-\overline{X})^2/n=(\hat{b}-\hat{a})^2/12\end{cases}$$

- 第二个方程开方得到 $\sqrt{3}S=(\hat{b}-\hat{a})/2$ 最后求出

$$\begin{cases}\hat{a}_M=\overline{X}-\sqrt{3}S\\ \hat{b}_M=\overline{X}+\sqrt{3}S\end{cases}$$

- **例5.3** 设总体 $X \sim (0-1)$ 分布，即

$$P(X=x)=p^{x}(1-p)^{1-x}, x=0,1$$

- 于是对于给定的样本样本的联合概率函数

$$L(p)=\prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i}=p^{\sum x_i}(1-p)^{n-\sum x_i}$$

- 求驻点，解得最大似然估计

$$\hat{p}=\frac{1}{n}\sum x_i=\bar{x}$$

- 例5.5 设总体 $X \sim \begin{pmatrix} 1 & 2 & 3 \\ 1-3\theta & 2\theta & \theta \end{pmatrix}$，1，1，2，3是取自总体的样本，(1)求 θ 的矩估计量和矩估计值；
- (2) 对于给定的样本,求样本的联合概率函数及 θ 的最大似然估计值。
- 解：(1) $\mathrm{E}X = 1\times(1-3\theta)+2\times 2\theta+3\times\theta = 1+4\theta = \overline{X}$

  得到 $\hat{\theta} = \dfrac{\bar{x}-1}{4} = \dfrac{3}{16}$
- (2) 对于给定的样本,样本的联合概率函数为

$$L(\theta) = (1-3\theta)^2 \times 2\theta \times \theta = 2\theta^2(1-3\theta)^2$$

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\mathrm{L}(\theta) = 4\theta(1-3\theta)(1-6\theta) = 0$$

- 故 θ 的最大似然估计值为 1/6 。
- 注意：理解样本1，1，2，3的作用

三、最大似然估计

似然函数是样本观测值和未知参数 θ 的函数。对于离散型它是参数取θ 时，$x_1,\cdots,x_n$ 出现（或发生）的概率。对于连续型，它虽不是概率，但也刻画了取 $x_1,\cdots,x_n$ 这些值可能性的大小。

当样本值固定时，它是参数θ 的函数，对应不同的θ ，$L(\theta)$ 是参数为θ 时的联合概率密度，并随θ 的不同而不同，最大似然估计的基本想法是：选择使 $x_1,\cdots,x_n$ 出现的可能性最大 $\hat{\theta}$ 的作为 θ 的估计。

$1-\alpha(0<\alpha<1)$

三、区间估计

设θ 为总体X的一个参数，$x_1,x_2,\cdots,x_n$ 为从此总体抽取的一个样本，如果一个随机区间$(T_1(\boldsymbol{X}),T_2(\boldsymbol{X}))$能满足

$$P(T_1(\boldsymbol{X})\le\theta\le T_2(\boldsymbol{X}))=1-\alpha \quad ,$$

其中$\boldsymbol{X}=(x_1,x_2,\cdots,x_n)$ 则称此区间为 θ 的置信度为 $1-\alpha(0<\alpha<1)$的双侧置信区间（confidence interval）。

（一） 总体均值μ 的区间估计

1. σ 已知的情况下的μ 置信区间

今假设总体X是正态的$N(\mu,\sigma^2)$，并且σ 已知的，要寻求μ 的置信区间。

对于抽取的样本$x_1,x_2,\cdots,x_n$ ，我们知道样本均值$\overline{X}$ 是μ 的自然的估计量，并且由（2.7）我们可以构造 $\overline{X}$ 与μ 的函数

$$z=G(\overline{X},\mu)=\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$$

因此由标准正态分布的分位数，

$$P(-z_{\alpha/2}\le\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\le z_{\alpha/2})=1-\alpha \qquad （5.6）$$

等价的可以转换成下式

$$P(\overline{X}-z_{\alpha/2}\frac{\sigma}{\sqrt{n}}<\mu<\overline{X}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}})=1-\alpha \qquad （5.7）$$

而（5.7）说的正是置信区间

$$(\overline{X}-z_{\alpha/2}\frac{\sigma}{\sqrt{n}},\overline{X}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}})$$

例5.7 设总体为X运输车队投入运营的卡车数目，已知其服从方差 $\sigma^2=225$ 的正态分布，今抽取 $n=25$ 的样本，由其算出 $\bar{x}=107$ ，求 μ 的80%置信区间。

解：这个题相当于 $\alpha=0.2$ 时求正态总体均值 μ 的置信区间问题。根据题意，

$$\frac{\sigma}{\sqrt{n}}=3$$

首先查表得：$z_{\alpha/2}=z_{0.1}=1.28$

利用置信区间公式，即可得出80%置信区间为：

$$107-1.28\times3\le\mu\le107+1.28\times3$$

$$103.16\le\mu\le110.84$$

也就是说在100次中有80次区间[103.16,110.84]将含有 μ

2. σ 未知的情况下 μ 的置信区间

在假定总体为正态分布的前提下，有

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

为了决定 $100(1-\alpha)\%$ 置信区间，要查t-分布表，查出自由度为n-1的t-分布上侧概率为 $\alpha/2$ 的点 $t_{\alpha/2}(n-1)$, 即

$$P(-t_{\alpha/2}(n-1) \le \frac{\overline{X}-\mu}{\tilde{S}/\sqrt{n}} \le t_{\alpha/2}(n-1)) = 1-\alpha$$

解此不等式便得到 σ 未知情况下 μ 的 $100(1-\alpha)\%$ 置信区间公式为

$$\overline{X} - t_{\alpha/2}(n-1)\frac{\tilde{S}}{\sqrt{n}} \le \mu \le \overline{X} + t_{\alpha/2}(n-1)\frac{\tilde{S}}{\sqrt{n}} \quad （5.8）$$

例5.8 某地区为制定高中学生的体育锻炼的成绩标准，在该地区随机抽取36名男生，测验100米短跑的成绩，假定短跑成绩服从正态分布，根据这36名男生的测试成绩计算得到样本均值为$\bar{x}=13.8$ 秒，样本标准差$s=1.5$ 秒，试求该地区高中学生100米跑的平均成绩的区间估计。

解：因为方差 σ^2 未知，用 $\tilde{S}^2$ 替换，置信区间为

$$\bar{X} \pm t_{\alpha/2}(n-1)\frac{\tilde{S}}{\sqrt{n}}$$

代入 $\bar{X}$ 的估计值及 $t_{0.025}(35)=2.03$ 得到

$$13.8 \pm 2.03 \times \frac{1.5}{6} = 13.8 \pm 0.51$$

即μ 的95%置信区间为 $[13.29, 14.31]$

3. 总体正态性条件的释放

根据中心极限定理，不论总体X服从均值为μ，方差σ^2的任何分布，$\overline{X}$的分布随着样本容量n的增加都逼近于平均值μ和方差$\frac{\sigma^2}{n}$的正态分布。这样，只要样本容量n足够大，随机变量

$$z=\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$$

仍可认为是标准正态分布的，从而公式仍可使用。至于t-分布的公式也同样使用。

实用中，通常以n>30作为足够大的标准。

例5.10 某厂负责人想估计某种材料的平均重量，随机抽取100包组成一个样本，样本的均值和修正标准差分别为37公斤和6公斤，试求总体均值 μ 的置信度为95%的置信区间。

解：我们不知道总体是否服从正态分布，方差 σ^2 也未知，但由于是大样本，可用 $\tilde{S}^2$ 替换 σ^2 ，置信区间为

$$\overline{X} \pm z_{\alpha/2}\frac{\tilde{S}}{\sqrt{n}}$$

代入 $\overline{X}$ 的值及 $Z_{0.025}=1.96$，得到

$$37\pm1.96\times\frac{6}{10}=37\pm1.176$$

即 μ 的95%置信区间为 $[35.824, 37.176]$

四、n的确定

1. σ 已知的情形

置信区间的半径 $z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ 即相当于最大允许误差：

$$\Delta = z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \qquad （5.11）$$

在α 给定，σ 已知时，这式子给出了 Δ 与n的关系。给定n时可用以求 Δ，反过来在指定 Δ 时也可用以求n：

估计样本均值所需要的最小样本容量

$$n = \frac{z_{\alpha/2}^2\sigma^2}{\Delta^2} \qquad （5.12）$$

n再大还能缩小 Δ ,但为达到最小的n由上式所确定。

2. σ 未知的情况

用序贯抽样通过迭代求最小容量n的方法，不介绍。

（二） 总体方差 σ^2 的置信区间

根据抽样分布的结果，在总体正态的假定之下

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

可得到 σ^2 的置信度为 $100(1-\alpha)\%$ 的置信区间为

$$\frac{(n-1)\tilde{S}^2}{\chi^2_{\alpha/2}(n-1)} \le \sigma^2 \le \frac{(n-1)\tilde{S}^2}{\chi^2_{1-\alpha/2}(n-1)} \qquad （5.14）$$

5.2参数假设检验

一、假设检验基本概念

对于总体的某种性质或是它的某个参数的值事先给出的某种猜想就称为统计假设或简称为**假设**（hypothesis），决定一个统计假设是否成立，就称为**假设检验**（hypothesis testing）。

原假设或**零假设**（null hypothesis），用H_0表示。H_0 的对立面，用 H_1（注：有的教材备择假设用下标*a*）表示，称为**备选假设**（alternative hypothesis）。两个假设 H_0 和 H_1 是互为余集的，因此两者非此即彼，没有别的可能。

二、两种类型的错误

第Ⅰ类错误指的是H_0客观上真实但被检验所拒绝，这种错误也称为**弃真错误**。**第Ⅱ类错误**指的是H_0客观上不真实但被检验所接受，这种错误也称为**取伪错误**。

用α表示第Ⅰ种类型错误的概率，用β表示第Ⅱ种类型错误概率。α也称为**显著性水平**（level of significance），$1-\alpha$称为**置信水平**（confidence level），它代表检验识别真实H_0的能力。$1-\beta$称为势（power），它代表检验识别不真H_0的能力。

三、假设检验步骤

1.根据实际问题提出零假设和备选假设

2.选适当的检验统计量

3.确定拒绝域

4.计算检验统计量的值

5.做出检验结论并给出解释

四、总体均值检验

关于总体均值μ 的检验都使用样本均值 $\overline{X}$ 来构造检验统计量，下面总假定总体是正态的或n>30以保证 $\overline{X}$ 是近似于正态。

1. 检验 $H_0:\mu=\mu_0$ ，$H_1:\mu\neq\mu_0$ （方差σ^2已知）

这种检验已在前面讨论过，检验统计量为

$$Z=\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} \qquad (5.15)$$

而拒绝域为

$$\{Z<-Z_{\frac{\alpha}{2}}\}\cup\{Z>Z_{\frac{\alpha}{2}}\} \qquad (5.16)$$

2．检验 $H_0: \mu = \mu_0$ ，$H_1: \mu \neq \mu_0$（方差 σ^2 未知）

这种情形只要用样本方差 S^2 代替σ^2 而使用自由度为 $\nu = n-1$ 的 $t-$ 分布即可。检验统计量为

$$t = \frac{\bar{X} - \mu_0}{\tilde{S} / \sqrt{n}} \qquad (5.17)$$

拒绝域为

$$\{t < -t_{\frac{\alpha}{2},\nu}\} \quad , \quad \{t > t_{\frac{\alpha}{2},\nu}\} \qquad (5.18)$$

3. 单边检验 $H_0:\mu\geq\mu_0$ ，$H_1:\mu<\mu_0$,或 $H_0:\mu\leq\mu_0$ ，$H_1:\mu>\mu_0$

现在与前面不同的是对H_0不利的方向只有一边，因此拒绝域也只有一边，而且Ⅰ型错误概率α 也由这一边承担。几种情形下的拒绝域由表5.3列出。

本教程单边检验的H_0都以复合零假的形式给出。由于备选假设一旦给定，H_0 也就唯一确定，并且H_0成立时的分布，考虑成参数为 μ_0

表5.3: 几种单边检验拒绝域

方差 σ^2	H_1	拒绝域
已知	$\mu>\mu_0$	$\{Z>Z_\alpha\}$
已知	$\mu<\mu_0$	$\{Z<-Z_\alpha\}$
未知	$\mu>\mu_0$	$\{t>t_\alpha(\nu)\}$
未知	$\mu<\mu_0$	$\{t<-t_\alpha(\nu)\}$

例5.19 健康成年男子脉搏平均为72次/分，高考体检时，某校参加体检的26名男生的脉搏平均为74.2次/分，修正标准差为6.2次/分，问此26名男生每分钟脉搏次数与一般成年男子有无显著差异？(α=0.05)

解：分析 ,如果男大学生每分钟脉搏次数与一般成年男子无显著差异，意味着这26名男生也是来自μo=72的正态总体。换言之，男大学生每分钟脉搏次数比一般成年男子多得多或少得多才认为两者差异显著。因此可以提出假设：

$H_0: \mu = 72$ ， $H_1: \mu \neq 72$ （次/分）

由于总体方差未知，只能用t检验。计算t值：

$$t = \frac{\bar{x} - \mu_0}{\tilde{s}/\sqrt{n}} = \frac{74.2 - 72}{6.2/\sqrt{26}} = 1.874$$

由 $\alpha = 0.05$ 临界值 $t_{0.025} = 2.06$ 判断：由 |1.874|＜2.06，没有充分理由拒绝原假设，故接受Ho认为男大学生每分钟脉搏次数与一般成年男子无显著差异。

4．两总体均值之差的检验

双边检验的基本形式是：

$$H_0 : \mu_1 - \mu_2 = 0(\text{即}\mu_1 = \mu_2) \quad , H_1 : \mu_1 - \mu_2 \neq 0(\text{即}\mu_1 \neq \mu_2)$$

可以通过正态分布或分布求出检验的拒绝域(见P144)

四、总体方差检验

1．单一总体方差σ^2的检验

这种检验零假设的基本形式是 $H_0: \sigma^2 = \sigma_0^2$ (指定值)，所用的检验 χ^2-统计量是如下的统计量，因为在正态总体的条件能够证明它服从自由度 $\nu = n-1$ 的 χ^2-分布：

$$\chi^2 = \frac{\nu \tilde{S}^2}{\sigma^2} \sim \chi_\nu^2 \qquad (5.21)$$

这里是S^2样本方差。利用这一结果就可以构造拒绝域。

5.3 非参数假设检验

一、分布拟合优度检验

拟合优度（goodness of fit）泛指两个事物相似或接近好坏的程度。

拟合优度 χ^2 –统计量如下：

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} = \sum_{i=1}^{k} \frac{O_i^{\,2}}{E_i} - n \sim \chi^2(k - \mathrm{r} - 1) \qquad （5.23）$$

每个 $O_i - E$ 是观测频数与理论频数间的“误差”,*r* 是欲估计的未知参数个数，对正态分布，*r* = 2。

- 例5.22 某公司上半年共售出汽车辆，按售出时间分组统计6个月的月销售量各为

27，18，15，24，36，30

试问能否认为月销售量是均衡的？

- 解：所说的均衡相当于每辆车售出时间服从均匀分布。如果均衡，6个月中每个月应该售出150/6=25辆，因此我们取零假设与备选假设为

$$H_0: E_1 = E_2 = E_3 = E_4 = E_5 = E_6 = 25$$

$H_1: E_1, E_2, \cdots, E_6$ 不全相等

计算统计量，得到它的计算值为

$$\chi_c^2=\frac{(27-25)^2}{25}+\frac{(18-25)^2}{25}+\frac{(15-25)^2}{25}$$
$$+\frac{(24-25)^2}{25}+\frac{(36-25)^2}{25}+\frac{(30-25)^2}{25}=\frac{300}{12}=12$$

今 E_i 已知，$r=0$，$v=k-1=6-1=5$。取 $\alpha=0.025$，查表得临界值 $\chi^2_{0.025}(5)=11.1$，现在是值大时不利于 H_0 的单边检验，拒绝域应为 $\{\chi^2>11.1\}$。因 χ_c^2 落在拒绝域内，故拒绝均衡的零假设，而判定销售量不均衡（受季节影响）。

二、正态性的 χ^2 检验

- 1．样本数据的标准化
- 2．分组

  ① 按区间长度等分

  ② 按概率等分
- 3．频数的计算

略！