

法律声明

□ 本课件包括：演示文稿，示例，代码，题库，视频和声音等，小象学院拥有完全知识产权的权利；只限于善意学习者在本课程使用，不得在课程范围外向任何第三方散播。任何其他人或机构不得盗版、复制、仿造其中的创意，我们将保留一切通过法律手段追究违反者的权利。

□ 课程详情请咨询

■ 微信公众号：小象

■ 新浪微博：ChinaHadoop



数理统计

目录

第一章	统计量与抽样分布
第二章	点估计
第三章	区间估计
第四章	假设检验
第五章	分布的检验
第六章	Bayes统计与统计判决理论
第七章	非参数统计
第八章	抽样调查

目 录

7.1 非参数统计

7.2 符号检验

7.3 Wilcoxon符号秩检验

非参数统计

在前几章讨论的统计问题大多属于这样一种情形：样本分布族的数学形式已知（如正态、指数、Poisson 分布等），但其中包含有限个未知的实参数。最典型的情况是关于正态分布族的均值、方差的估计和检验问题，这种统计问题称为参数型的。

然而，在实际问题中也有这样的情况：对样本分布族并未给出其数学形式，而只作了某些一般的假定，如只假定总体分布的连续性或对称性，而对其分布的数学形式则一无所知。类似这种情形的统计问题称为非参数型的。

如果关于总体分布族事先有比较多的信息，那么把统计推断问题放在参数型的范围内讨论是恰当的。如果事先关于总体分布族只有很少信息，此时就需要一种与总体分布族的具体数学形式无关的统计推断方法。这种方法称为非参数方法。

非参数统计

非参数方法具有以下几个特点：① 适用面广而针对性较差. 这是由于非参数方法对总体分布族没有提出太多的假设. 但是, 如果对特定的统计推断问题, 参数模型适合该问题, 且存在一个具有优良性质的参数型统计方法, 则二者相比, 非参数统计方法一般效率较低. ② 大样本理论在非参数统计中起着极其重要的作用. 这是因为对总体分布族知之甚少, 很难指望导出有关统计量的精确分布, 迫使考虑用统计量的极限分布近似代替其精确分布. ③ 只能使用样本中的“一般”信息, 如位置、次序关系等. 这也是由于对总体分布族所知甚少之故. ④ 非参数方法具有稳健性, 即当真实模型与设定的理论模型有一定偏离时, 能维持较好的性质. 这是由于非参数方法对模型限制小, 故自然就具有稳健性 (robustness).

非参数统计

非参数统计可以分为两个范畴，一个是比较经典的基于秩的，以检验为主的非参数统计推断；另一部分是近三四十年来发展的非参数回归、非参数密度估计、自助法以及小波方法等现代非参数统计方法。这两者均不对总体分布做任何假定，但除此之外，这两部分内容在方法上和概念上均没有多少共同点。

基于秩的非参数统计

基于秩的非参数统计

- (1) 符号检验.
- (2) 符号秩和检验.
- (3) 分位数的估计.

现代非参数统计

非参数推断的基本思想是在尽可能少的假定时利用数据对一个未知量做出推断。通常，这意味着利用具有无限维的统计模型。事实上，对非参数推断的另一个更好的名字可能就是无穷维推断。

现代非参数统计的典型问题：

(1)估计分布函数. 给定一个 IID 样本 $X_1, \dots, X_n \sim F$, 估计 CDF $F(x) = P(X \leq x)$.

(2)估计泛函. 给定一个 IID 样本 $X_1, \dots, X_n \sim F$, 估计一个泛函 $T(F)$, 如均值 $T(F) = \int x dF(x)$.

(3)密度估计. 给定一个 IID 样本 $X_1, \dots, X_n \sim F$, 估计密度 $f(x) = F'(x)$.

(4)非参数回归或曲线估计. 给定 $(X_1, Y_1), \dots, (x_n, Y_n)$, 估计回归函数 $r(x) = E(Y|X = x)$.

适应任意分布统计量

定义：适应任意分布统计量. 设随机变量 $X_1, X_2, \dots, X_n$ 是来自总体 $F(x)$ 的样本, 一切可能的 $F(x)$ 组成的分布类 $\mathcal{F}$. 如果统计量 $T(X_1, X_2, \dots, X_n)$ 对任意的 $F \in \mathcal{F}$ 均有相同的分布, 则称 T 关于 $\mathcal{F}$ 是适应任意分布的.

计数统计量

定义: 计数统计量. 设 X 是一个随机变量, 对于一个给定的实数 θ_0 , 定义随机变量

$$\Psi = \psi(X - \theta_0)$$

其中

$$\psi(t) = \begin{cases} 1, & \text{当 } t > 0 \text{ 时} \end{cases} \quad (5)$$

$$\begin{cases} 0, & \text{当 } t \leq 0 \text{ 时} \end{cases} \quad (6)$$

随机变量 Ψ 称为 X 按 θ_0 分段的计数统计量.

秩统计量

定义：秩统计量. 设 $Z_1, Z_2, \dots, Z_N$ 是来自连续分布 $F(z)$ 的简单随机样本, $Z_{(1)} \leq Z_{(2)} \leq \dots \leq Z_{(N)}$ 为其次序统计量, 定义随机变量

$$R_i = r, \quad Z_i = Z_{(r)}, \quad i = 1, 2, \dots, N$$

当 R_i 是唯一确定时, 称样本观测值 Z_i 有秩 $R_i (i = 1, 2, \dots, N)$. (由于 $F(z)$ 连续, 因而 R_i 不唯一的概率为 0)

秩统计量

定理:记 $R = (R_1, R_2, \dots, R_N)$, 集合

$$\mathcal{R} = (r_1, r_2, \dots, r_N) | (r_1, r_2, \dots, r_N) (1, 2, \dots, N)$$

则 R 在 $\mathcal{R}$ 上均匀分布.

秩统计量

定理: $R = (R_1, R_2, \dots, R_N)$ 的边缘分布也是均匀分布. 特别的, 一维边缘分布有

$$P(R_i = r) = \begin{cases} \frac{1}{N}, & \text{当 } r = 1, 2, \dots, N \text{ 时} \\ 0, & \text{else} \end{cases} \quad (7)$$

(8)

二维边缘分布当 $i \neq j$ 时, 有

$$P(R_i = r, R_j = s) = \begin{cases} \frac{1}{N(N-1)}, & \text{当 } r \neq s, r, s = 1, 2, \dots, N \text{ 时} \\ 0, & \text{else} \end{cases} \quad (9)$$

(10)

秩统计量

定理:对秩统计量 $R = (R_1, R_2, \dots, R_N)$, 有

$$E(R_i) = \frac{N+1}{2}, i = 1, 2, \dots, N$$

$$Var(R_i) = \frac{(N+1)(N-1)}{12}, i = 1, 2, \dots, N$$

$$Cov(R_i, R_j) = -\frac{(N+1)}{12}, i \neq j, i, j = 1, \dots, N$$

目 录

7.1 非参数统计

7.2 符号检验

7.3 Wilcoxon符号秩检验

符号检验

设随机变量 $X_1, X_2, \dots, X_n$ 相互独立同分布, 分布函数为 $F(x)$, $F(x)$ 在 $x = 0$ 处连续. 假设检验问题

$$H_0 : F(0) = 0.5 \longleftrightarrow H_1 : F(0) \neq 0.5$$

检验统计量可取 $B = \sum_{i=1}^n \Psi_i = \sum_{i=1}^n \psi(X_i)$, 即为 $X_1, X_2, \dots, X_n$ 中取正号的个数. 在 H_0 成立的条件下, B 的分布为参数为 n 和 $\frac{1}{2}$ 的二项分布 $b(n, \frac{1}{2})$. 可确定 c_1, c_2 , 使得

$$P(B \leq c_1) \leq \frac{\alpha}{2}, P(B \geq c_2) \leq \frac{\alpha}{2}$$

否定域为 $\{0, 1, \dots, c_1\} \cup \{c_2, c_2 + 1, \dots, n\}$.

符号检验

1. 检验一个样本是否来自某个总体

假设某个总体的中位数为 M_0 , 如果样本中位数 $M = M_0$, 我们就接受样本来自某个总体的假设. 其具体的检验方法是这样的. 首先从每个样本观察值中减去总体中位数 M_0 , 得出的正、负差额用正 (+)、负 (-) 号加以表示. 如果总体中位数等于样本中位数, 即 $M = M_0$, 那么, 样本观察值在中位数上、下的数目应各占一半, 因现时出现正号或负号的概率应各占 $1/2$. 设样本容量为 n , 就可以用二项分布 $B(n, 1/2)$ 来计算出现负号 (或正号) 个数的概率, 从而根据一定的显著性水平 α , 作出是否接受原假设 $H_0: M = M_0$ 的判定.

符号检验

例 5.18 联合国人员在世界上 66 个大城市的生活花费指数 (以纽约市 1996 年 12 月为 100) 按自小至大的次序排列如下 (这里北京的指数为 99)：

66	75	78	80	81	81	82	83	83	83	83
84	85	85	86	86	86	86	87	87	88	88
88	88	88	89	89	89	89	90	90	91	91
91	91	92	93	93	96	96	96	97	99	100
101	102	103	103	104	104	104	105	106	109	109
110	110	110	111	113	115	116	117	118	155	192

假设这个样本是从世界许多大城市中随机抽样得到的。试用符号检验分析，北京是在中位数之上，还是在中位数之下。

符号检验

```
> X<-scan()  
1: 66 75 78 80 81 81 82 83 83 83 83  
12: 84 85 85 86 86 86 86 87 87 88 88  
23: 88 88 88 89 89 89 89 90 90 91 91  
34: 91 91 92 93 93 96 96 96 97 99 100  
45: 101 102 103 103 104 104 104 105 106 109 109  
56: 110 110 110 111 113 115 116 117 118 155 192  
67:  
Read 66 items  
> binom.test(sum(X>99), length(X), al="l")
```

Exact binomial test

```
data: sum(X > 99) and length(X)  
number of successes = 23, number of trials = 66, p-value = 0.009329  
alternative hypothesis: true probability of success is less than 0.5  
95 percent confidence interval:  
 0.0000000 0.4563087  
sample estimates:  
probability of success  
      0.3484848
```

符号检验

2. 用成对样本来检验两个总体间是否存在显著差异

符号检验法也可用于以成对随机样本观察值来检验两个总体之间是否存在显著差异. 如果两个总体无显著差异, 则两个成对随机样本观察值正、负差额的个数应大体相等. 假定 $x_i - y_i > 0$ 用正号表示, $x_i - y_i < 0$ 用负号表示, 则如果两个总体无显著差异, 那么出现正号和负号的概率各占 $1/2$. 和上面检验样本是否来自某个总体一样, 可用二项分布 $B(n, 1/2)$, 根据一定的显著性水平和正号 (或负号) 的个数, 作出接受或拒绝两个总体无显著差异的判断.

符号检验

例 5.19 用两种不同的饲料养猪，其增重情况如表 5.10 所示。 试分析两种饲料

表 5.10: 不同饲料养猪的增重情况

对编号	1	2	3	4	5	6	7	8	9	10	11	12	13	14
饲料 X	25	30	28	23	27	35	30	28	32	29	30	30	31	16
饲料 Y	19	32	21	19	25	31	31	26	30	25	28	31	25	25

养猪有无显著差异.

符号检验

```
> x<-scan()  
1: 25 30 28 23 27 35 30 28 32 29 30 30 31 16  
15:  
Read 14 items  
> y<-scan()  
1: 19 32 21 19 25 31 31 26 30 25 28 31 25 25  
15:  
Read 14 items  
> binom.test(sum(x<y), length(x))
```

Exact binomial test

```
data: sum(x < y) and length(x)  
number of successes = 4, number of trials = 14, p-value = 0.1796  
alternative hypothesis: true probability of success is not equal to 0.5  
95 percent confidence interval:  
 0.08388932 0.58103526  
sample estimates:  
probability of success  
 0.2857143
```

目 录

7.1 非参数统计

7.2 符号检验

7.3 Wilcoxon符号秩检验

Wilcoxon符号秩检验

1. 对来自一个总体样本的检验

符号检验利用了观测值和原假设的中心位置之差的符号来进行检验，但是它并没有利用这些差的大小（体现于差的绝对值的大小）所包含的信息，不同的符号代表了中心位置的哪一边，而差的绝对值的秩的大小代表距中心位置的远近。如果将两者结合起来，自然比仅仅利用符号更有效。这也是下面要介绍的Wilcoxon(威尔科克逊)符号秩检验 (Wilcoxon signed-rank test) 的宗旨。

Wilcoxon符号秩检验

为了弥补符号检验法之不足, 在这里将介绍一种在一定程度上考虑到样本观察值与总体中位数之间的差额, 即 $|x_i - M_0|$ (其中 $i = 1, 2, \dots, n$) 的大小的检验方法. 在这里假定: (1) 总体分布是连续的; (2) 总体对其中位数是对称的. 这样, 将以上 $|x_i - M_0|$ 得到的差额, 按递增次序排列, 并据差额的次序给出相应的秩次 R_i , 如差额绝对值最小者给以秩次 1, 次小者给以秩次 2, $\dots, \dots$, 最大值给以秩次 n . 再按 $x_i - M_0 > 0$ 为正秩次, $x_i - M_0 < 0$ 为负秩次. 然后按照正秩次和进行检验, 这就是顺序和检验. 这种方法首先由 Wilcoxon 提出的, 所以称为 Wilcoxon 符号秩检验.

Wilcoxon符号秩检验

Wilcoxon 检验不仅考虑到每个观察值比总体中位数 M_0 大还是小, 而且在一定程度上也考虑了大多少, 小多少. 在进行检验时, 如果观察值与总体中位数的差额的绝对值相等时, 就要用平均秩次来代替. 例如, $|x_i - M_0| = |x_j - M_0| = |x_k - M_0|$, 首先, 给以相应的秩次为 4、5、6, 其平均值为 5 (R 软件以平均值定义相同值的秩次, 三个数据的秩次均是 5). 此外, 如果 $x_i - M_0 = 0$, 就将 x_i 从观察数据中去掉.

Wilcoxon符号秩检验

如果原观察值的数目为 n' , 减去差额为 0 的观察数据后, 其样本数为 n . 用 $R_i^{(+)}$ 表示正秩次, W 表示正秩次的和, 则 Wilcoxon 统计量为

$$W = \sum_{i=1}^n R_i^{(+)}. \quad (5.26)$$

因为 n 个整数 $1, 2, \dots, n$ 的总和用 $\frac{n(n+1)}{2}$ 计算, 而正秩次总和可以在区间 $\left(0, \frac{n(n+1)}{2}\right)$ 内变动, 如果观察值来自中位数为 M_0 的某个总体的假设为真, 那么 Wilcoxon 检验统计量的取值将是秩次和的平均数, 即 $\mu_W = \frac{n(n+1)}{4}$ 的左右变动. 如果该假设不成立, 则 W 的取值将向秩次和的两头的数值靠近. 这样, 在一定的显著性水平, 便可进行检验了.

Wilcoxon符号秩检验

例 5.24 假定某电池厂宣称该厂生产的某种型号电池寿命的中位数为 140 安培小时。为了检验该厂生产的电池是否符合其规定的标准，现从新近生产的一批电池中抽取 20 个随机样本，并对这 20 个电池的寿命进行了测试，其结果如下 (单位：安培小时)：

137.0 140.0 138.3 139.0 144.3 139.1 141.7 137.3 133.5 138.2
141.1 139.2 136.5 136.5 135.6 138.0 140.9 140.6 136.3 134.1

试用 *Wilcoxon* 符号秩检验分析该厂生产的电池是否符合其标准。

Wilcoxon符号秩检验

```
> X<-scan()  
1: 137.0 140.0 138.3 139.0 144.3 139.1 141.7 137.3 133.5 138.2 141.1 139.2 136.5 136.5 135.6  
.9 140.6 136.3 134.1  
21:  
Read 20 items  
> wilcox.test(X, mu=140, alternative="less", exact=FALSE, correct=FALSE, conf.int=TRUE)
```

wilcoxon signed rank test

```
data: X  
V = 34, p-value = 0.007034  
alternative hypothesis: true location is less than 140  
95 percent confidence interval:  
-Inf 139.2  
sample estimates:  
(pseudo)median  
138.2
```

Wilcoxon符号秩检验

上面介绍了用 Wilcoxon 符号秩检验方法检验一个样本是否来自某个总体的内容. 同样, 这个方法也可用于成对样本的检验, 从而说明两个总体是否存在显著差异.

例 5.25 为了检验一种新的复合肥和原来使用的肥料相比是否显著地提高了小麦的产量, 在一个农场中选择了 10 块田地, 每块等分为两部分, 其中任指定一部分使用新的复合肥料, 另一部分使用原肥料. 小麦成熟后称得各部分小麦产量如表 5.14 所示. 试用 Wilcoxon 符号检验法检验新复合肥是否会显著提高小麦的

表 5.14: 使用不同肥料情况下小麦的产量 (单位: 千克)

田 块	1	2	3	4	5	6	7	8	9	10
新复合肥	459	367	303	392	310	342	421	446	430	412
原肥料	414	306	321	443	281	301	353	391	405	390

产量, 并与符号检验作比较 ($\alpha = 0.05$).

Wilcoxon符号秩检验

```
> x<-c(459, 367, 303, 392, 310, 342, 421, 446, 430, 412)
> y<-c(414, 306, 321, 443, 281, 301, 353, 391, 405, 390)
> wilcox.test(x, y, alternative = "greater", paired = TRUE)
```

wilcoxon signed rank test

```
data: x and y
V = 47, p-value = 0.02441
alternative hypothesis: true location shift is greater than 0
```

```
> binom.test(sum(x>y), length(x), alternative = "greater")
```

Exact binomial test

```
data: sum(x > y) and length(x)
number of successes = 8, number of trials = 10, p-value = 0.05469
alternative hypothesis: true probability of success is greater than 0.5
95 percent confidence interval:
 0.4930987 1.0000000
sample estimates:
probability of success
              0.8
```

Wilcoxon符号秩检验

2. 非成对样本的秩次和检验

假定两个非成对样本的观察值为 $X_1, X_2, \dots, X_{n_1}$ 和 $Y_1, Y_2, \dots, Y_{n_2}$, 其样本容量分别为 n_1 和 n_2 . 现要检验两个随机样本来自两个总体的中位数是否相等 (如果中位数相等, 则认为两个总体无差异).

将样本的观察值排在一起, $X_1, X_2, \dots, X_{n_1}, Y_1, Y_2, \dots, Y_{n_2}$, 仍设 $r_1, r_2, \dots, r_{n_1}$ 为由 $X_1, X_2, \dots, X_{n_1}$ 产生的秩统计量, $R_1, R_2, \dots, R_{n_2}$ 为由 $Y_1, Y_2, \dots, Y_{n_2}$ 产生的秩统计量, 则 Wilcoxon-Mann-Whitney 统计量定义为

$$U = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - \sum_{i=1}^{n_2} R_i. \quad (5.27)$$

类似单一总体的 Wilcoxon 符号检验一样, 可以通过统计量 U 进行检验, 该检验称为 Wilcoxon 秩和检验.

Wilcoxon符号秩检验

例 5.26 今测得 10 名非铅作业工人和 7 名铅作业工人的血铅值，如表 5.15 所示。试用 *Wilcoxon* 秩和检验分析两组工人血铅值有无差异。

表 5.15: 两组工人的血铅值 (单位: 10^{-6}mmol/L)

非铅作业组	24	26	29	34	43	58	63	72	87	101
铅作业组	82	87	97	121	164	208	213			

Wilcoxon符号秩检验

```
> x<-c(24, 26, 29, 34, 43, 58, 63, 72, 87, 101)
> y<-c(82, 87, 97, 121, 164, 208, 213)
> #不采取连续修正
> wilcox.test(x,y,alternative="less",exact=FALSE,correct=FALSE)
```

wilcoxon rank sum test

```
data: x and y
W = 4.5, p-value = 0.001449
alternative hypothesis: true location shift is less than 0
```

```
> #采取连续修正
> wilcox.test(x,y,alternative="less",exact=FALSE)
```

wilcoxon rank sum test with continuity correction

```
data: x and y
W = 4.5, p-value = 0.001698
alternative hypothesis: true location shift is less than 0
```

Wilcoxon符号秩检验

例 5.28 某医院用某种药物治疗两型慢性支气管炎患者共 216 例, 疗效由表 5.18 所示. 试分析该药物对两型慢性支气管炎的治疗是否相同.

表 5.18: 某种药物治疗两型慢性支气管炎疗效结果

疗效	控制	显效	进步	无效
单纯型	62	41	14	11
喘息型	20	37	16	15

Wilcoxon符号秩检验

```
> x<-rep(1:4, c(62, 41, 14,11))  
> y<-rep(1:4, c(20, 37, 16, 15))  
> wilcox.test(x, y, exact=FALSE)
```

wilcoxon rank sum test with continuity correction

data: x and y

W = 3994, p-value = 0.0001242

alternative hypothesis: true location shift is not equal to 0

我们在这里

□ 问题答疑：<http://www.xxwenda.com/>

■ 可邀请老师或者其他人回复问题

联系我们

小象学院：互联网新技术在线教育领航者

- 微信公众号：小象
- 新浪微博：ChinaHadoop

