

符号检验法及其在动物生态学研究中的应用*

宛新荣 钟文勤**

(中国科学院动物研究所 农业虫鼠害综合治理研究国家重点实验室 北京 100080)

摘要: 符号检验法是非参数统计法中的一种重要方法,在动物生态学的实验数据统计分析领域有一定的实用价值,本文简要介绍这一方法的使用。

关键词: 符号检验法; 非参数统计; 实验设计; 动物生态学

中图分类号: Q141 **文献标识码:** A **文章编号:** 0250-3263(2002)06-56-04

The Sign Test and its Application in Fields of Animal Ecology

WAN Xin-Rong ZHONG Wen-Qin

(Institute of Zoology, Chinese Academy of Sciences Beijing 100080, China)

Abstract: The Sign test is an important non-parametric statistical method useful in the analysis of experimental data in the field of animal ecology. This paper provides a brief description of the method and how to use it.

Key words: Experimental design; Sign test; Non-parametric test; Animal ecology

在生物学研究中,对实验数据的分析和处理常常借助于数理统计方法^[1-3]。常用的多为参数检验法,如 *t*-检验等等。对于一些特殊离散型实验数据,需要涉及非参数统计方法,非参数统计方法是对常规统计方法的一个重要补充。由于在野外条件下不可控制因素较多,动物生态学工作者常常会采取特殊的设计方案,这就使实验数据的复杂程度提高。在某些场合下需要借助非参数统计方法来完成。了解非参数统计方法,无疑会大大提高动物生态学工作者的实验设计和统计分析能力。与动物生态学研究领域关系密切的非参数统计方法主要有:符号检验法(sign test)、独立性检验法以及秩和检验法。本文简要介绍符号检验法在动物学研究领域的一些应用场合以及统计运算方法。

1 符号检验法的使用规程

严格意义上说,符号检验法是配对 *t*-检验法的一种简化方式,是比较一系列对等的实验数据大小差值的统计方法,与配对检验法相比,它只比较两组配对实验数据之间的大小,不涉及实验数据的具体差值。为便于读者熟悉符号检验的统计流程,现举二个数据范例。

例1,不同季节 20 个实验动物体重变化情况。A 组表示冬季,B 组表示秋季,统计目的是比较从秋季到冬季动物个体的体重是否增加(表 1 为假定的实验数据)。

表 1 不同季节动物体重的变化

个体序号	冬季	秋季	符号	个体序号	冬季	秋季	符号
1	78	76	+	11	75	78	-
2	73	75	-	12	69	65	+
3	68	65	+	13	86	80	+
4	65	61	+	14	72	70	+
5	77	72	+	15	77	77	0
6	83	74	+	16	81	78	+
7	90	79	+	17	80	76	+
8	74	77	-	18	78	79	-
9	79	79	0	19	76	74	+
10	81	73	+	20	82	80	+

* 中国科学院知识创新重大项目(KSCX1-08), 中国科学院生物特别支持资助项目(STZ-01-06);

** 通讯作者;

第一作者介绍 宛新荣,男,33岁,副研究员;主要从事啮齿动物生态学研究。

收稿日期:2002-02-15,修回日期:2002-07-10

例2,布氏田鼠(*Microtus brandti*)对3种主要贮草成分的选储实验结果^[4]列于表2。如何确定、检验该鼠对3种主要贮草成分的选择嗜好顺序?

上述的2个范例中,例1有具体数值,例2中无具体数值,但只有数值比较结果(优先顺序),相对而言,例1更具普遍性。现根据例1来说明符号检验法的检验步骤。

首先去掉符号为“0”的两个无效数据点,剩下18个

有效数据点计算“+”和“-”的个数。表1中“+”号为14个,“-”号为4个。“-”号较少,因此采用“-”号数目与有效统计样本总数(即“+”号和“-”号的总和)来进行比较。在本例中符号“-”占18个有效统计样本量中的4个,据此计算正常分布下出现这种情形的概率(零假设为:秋季和冬季个体体重无变化;备择假设:冬季体重大于秋季体重,为单侧检验),如果概率很低,则否定原始假设,接受备择假设。

表2 布氏田鼠对主要贮草成分的选储优先次序

记录序号	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
黄蒿根分值	2	2	2	2	1	2	2	2	2	2	2	1	2	2	2
冷蒿分值	1	1	1	1	2	1	1	1	1	1	1	2	1	1	1
黄蒿果枝分值	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

根据零假设,表1中出现正负号的概率应该是均等的(均为1/2),18个数值中至多出现4个负号的情形有5种:0、1、2、3、4。不妨假定上述5种情形的概率值依次为 $p(0)$ 、 $p(1)$ 、 $p(2)$ 、 $p(3)$ 、 $p(4)$ 。利用二项展开式,可以分别计算5种情形下的概率值^[1,2]:

$$p(0) = C_{18}^0 \times (1/2)^{18}$$

$$p(1) = C_{18}^1 \times (1/2)^{18}$$

$$p(2) = C_{18}^2 \times (1/2)^{18}$$

$$p(3) = C_{18}^3 \times (1/2)^{18}$$

$$p(4) = C_{18}^4 \times (1/2)^{18}$$

$$p(S \leq 4) = p(0) + p(1) + p(2) + p(3) + p(4)$$

$$= (C_{18}^0 + C_{18}^1 + C_{18}^2 + C_{18}^3 + C_{18}^4) \times (1/2)^{18}$$

$$p(S \leq 4) = 0.015$$

由于 $p(S \leq 4) < 0.05$,拒绝零假设,因而该组数据表明动物冬季体重要明显高于秋季体重。此外,表1中的数据可以用配对 t -检验法检验,检验结果也同样表明动物冬季体重要明显高于秋季体重($P < 0.05$),这与符号检验法得出的结论一致。

一般情况下,如果在 n 种情形中出现 m ($m \leq n/2$)个符号的概率计算步骤如下:

$$p(S \leq m) = p(0) + p(1) + p(2) + \dots + p(m)$$

$$= (C_n^0 + C_n^1 + C_n^2 + \dots + C_n^m) \times (1/2)^n$$

如果样本量很大,上述计算公式会很麻烦。在实际应用中,可通过查阅符号检验表值法解决这个问题,一般的数理统计书籍或统计手册均附有符号检验表^[1],可大大简化统计过程。实际上,当 $n \rightarrow \infty$ 时, S 渐近正态分布 $N(\mu, \sigma^2)$ 。其中:

$$\mu = np = n/2$$

$$\sigma = \sqrt{np(1-p)} = \sqrt{n/4} = \sqrt{n}/2$$

u 检验的统计估算公式为:

$$u = \frac{S - \mu}{\sigma} = \frac{S - n/2}{\sqrt{n}/2} = \frac{2S - n}{\sqrt{n}}$$

此外,有条件的还可以借助常用统计软件如SPSS^[5]完成符号检验的统计检验问题。

如果采用查表方法确定布氏田鼠对3种主要贮草成分的选储嗜好(表2),在15次实验中,黄蒿根的选储存优先级别有13次高于冷蒿,2次低于冷蒿^[4],查阅符号检验表中样本总量为15中出现2个的概率,结果是 $P < 0.05$ 。由此可知,布氏田鼠更喜储黄蒿根;而黄蒿根对黄蒿果枝以及冷蒿对黄蒿果枝均为15:0,查表得: $P < 0.01$,这表明,黄蒿果枝的优先级别最低。因此可以得出结论,布氏田鼠在秋季贮草季节对贮草的选储优先次序为:黄蒿根→冷蒿→黄蒿果枝。

2 符号检验法与配对 t -检验法的关系

从上面的统计步骤可以看到,符号检验法是配对 t -检验法的一种简化。从应用范围来看,符号检验法要比配对 t -检验法范围更广,凡是能应用配对 t -检验的场合,均可以使用符号检验;反之,使用符号检验的场合不一定都能使用配对 t -检验法。符号检验法最大的缺陷是在统计过程中省略一些量化的信息。所以一般情况下,能用到配对 t -检验的场合,通常不采用符号检验法。因此,在可以同时使用二种检验方法的情形中,配对 t -检验法更有效,因而优先使用配对 t -检验法;在个别的场合下,两种检验方法可能有不同的检验结论,这时应该采用配对 t -检验法的结论。这可能是符号检验法为多数动物生态学工作者所忽视的主要原因。

由于配对 t -检验法要求统计对象必须具有量化的实验数据, 在一些特殊的场合, 实验数据只有对立的两种结果(最多包含第三种中性结果, 如平局等)大小之分, 即只有定性的数据而没有定量的数据时, 配对 t -检验法是无能为力的, 这时才真正显示符号检验法的价值。

除此之外, 符号检验法与配对 t -检验法还有一个细微的差别: 配对 t -检验法要求实验数据符合正态分布, 而符号检验法对实验数据则无此项要求。不过, 大多数生物学实验数据都符合正态分布^[1], 因而这种差别在动物生态学实验数据的分析和处理过程中实际上意义并不大。

3 符号检验法的一些应用场合举例

例如在确定动物的社群等级序位方法上, 通过动物相遇后打斗胜负结果, 累计一定的次数之后就可以统计个体间实力的强弱对比或者是等级序位的情况。打斗胜负只是定性的数值: 胜、负、平, 而没有量化的指标, 因此, 对这类数据的检验法只能采用符号检验法。即对胜、负、平三种结果依次编号为“+”、“-”和“0”。假如某次的实验结果是胜 52 次, 负 18 次, 平 6 次(平局为无效统计样本), 则检验有效样本总数 70(52 + 18) 次中负场至多出现 18 次的概率。

例如, 按照年龄、性别将布氏田鼠分成 5 个组: 越冬雄鼠、越冬雌鼠、当年成体雄鼠、当年成体雌鼠以及亚成体鼠或者幼鼠, 利用打斗胜负记录结果分析田鼠的社群等级序位(表 3)。

表 3 利用打斗胜负记录结果分析田鼠的社群等级序位

6 月	越冬雌鼠	当年雄鼠	当年雌鼠	亚成体和幼体
越冬雄鼠	(36, 5)**	(45, 0)**	(52, 4)**	(16, 0)**
越冬雌鼠		(52, 2)**	(43, 4)**	(22, 0)**
当年雄鼠			(18, 36)*	(34, 0)**
当年雌鼠				(37, 0)**

(36, 5) 表示该行的个体与该列的个体在 41 次(36 + 5) 相遇中, 36 次击退该列的对手, 5 次负于(或回避)该列的对手。
** 表示极显著($P < 0.01$), * 表示显著($P < 0.05$)

从表 3 中逐项检验结果(查符号检验表法)为: 越冬雄鼠 \Rightarrow 越冬雌鼠; 越冬雌鼠 \Rightarrow 当年雌鼠; 当年雌鼠 \rightarrow 当年雄鼠; 当年雄鼠 \Rightarrow 亚成体鼠和幼鼠, 其中“ \Rightarrow ”表示社群等级优势极显著($P < 0.01$); “ \rightarrow ”表示社群等级优势显著($P < 0.05$)。其它检验(如越冬雄鼠 \Rightarrow 当年雌鼠)

从略。根据上述结果将确定 1996 年 6 月布氏田鼠社群等级序位如下:

越冬雄鼠 \Rightarrow 越冬雌鼠 \Rightarrow 当年雌鼠 \rightarrow 当年雄鼠 \Rightarrow 亚成体鼠和幼鼠。

实验结论是: 越冬雄鼠保持着最高的社群等级, 其次是越冬雌鼠, 当年雌鼠的等级序位略高于当年雄鼠, 亚成体和幼鼠的等级序位最低。

有关动物的食性选择、栖息地选择等实验也是动物生态学中很常见的问题^[6-10], 多数实验数据可按照类似例 2 的步骤来处理。例如, 如果在确定动物对 2 种食物 A、B 的优先选择次序的实验中, 先吃完(或者先取食完事先规定的比例, 如先吃完 A 记录为“+”, 那么先吃完 B 记为“-”, 同时吃完记为“0”(此为无效统计样本)。重复若干次实验后, 即可采用符号检验法确定动物对两种食物类型 A、B 的优先选择次序。

此外, 符号检验还可以分析动物是否具有扩散或留居倾向。例如一次实验中, 50 例实验中动物个体扩散了 35 例, 15 例没有扩散。如果规定扩散为“+”, 留居的为“-”, 也可以应用符号检验法确定动物的扩散倾向是否明显超出留居倾向。

除了上面列举的一些情形外, 符号检验法还可以在其它场合起到关键性的作用, 这需要动物学研究人员在不同研究领域的摸索和积累经验。鉴于在某些实验设计中符号检验法可能是惟一能被应用的方法, 熟悉这一统计法, 有助于提高动物生态学工作者对野外实验的综合设计能力。当然, 应用符号检验也要争取应用符号检验表或者统计软件以简化计算过程并节省时间。

参 考 文 献

- [1] 杜荣骞. 生物统计学. 北京: 高等教育出版社, 1987.
- [2] Underwood A J. Experiments in Ecology. London: Cambridge University Press, 1997. 1 ~ 212.
- [3] 孙山泽. 非参数统计讲义. 北京: 北京大学出版社, 2000.
- [4] 宛新荣, 王梦军, 王广和等. 布氏田鼠对主要贮草种类的选嗜嗜好. 生态学杂志, 2002, 21(1): 65 ~ 66.
- [5] 刘先勇, 袁长迎, 段宝福等. SPSS 10.0 统计分析软件与应用. 北京: 国防工业出版社, 2001.
- [6] 钟文勤, 周庆强, 孙崇瀚. 布氏田鼠的生境选择与植被条件. 见: 中国科学院内蒙古草原生态系统定位研究站编. 草原生态系统研究第 1 集. 北京: 科学出版社, 1985. 147 ~ 153.
- [7] 周庆强, 钟文勤, 孙崇瀚. 达乌尔黄鼠的食物和食量. 见: 中国科学院内蒙古草原生态系统定位研究站编. 草原生态

- 态系统研究第 1 集,北京:科学出版社,1985.139 ~ 146.
- [8] 王桂明,周庆强,钟文勤等. 自由生活下达乌尔黄鼠的食物选择. 见:中国动物学会编. 中国动物学会成立 60 周年纪念论文集. 北京:中国科学技术出版社,1994.
- [9] 王梦军,钟文勤,宛新荣等. 达乌尔鼠兔扩散过程中的生境选择. 动物学报,1998,44(4):398 ~ 405.
- [10] Wang G M, Zhou Q Q, Zhong W Q *et al.* Comparative food preference of *Microtus brandti* and *Ochotona daurica* in grasslands of Inner Mongolia, China. *Mammalian Biology*, 2001, 66:312 ~ 316