文章编号:1000 - 582X(2004)05 - 0093 - 03

关于"上限为三法则"(Rule of Three)的形成条件*

杨　劲

(重庆工商大学 计算机科学与信息工程学院,重庆　400020)

摘　要:"上限为三法则"(Rule of Three)在实验数据统计等带有抽样调查性质的研究上有相当广泛的应用。但是真正了解其由来,及其理论推导的研究工作并不多。人们一般把它当作是一种习惯、一种通则。文中就这一法则分别使用二项分布、普阿松分布、Louis 法和贝氏法四种不同的方法进行推导,并通过实际的数值结果对该法则进行验证比较,得到该法则的四个形成条件结论。

关键词:上限为三法则;概率统计;形成条件

中图分类号:O241　　**文献标识码**:A

当试验或测试对象是属于柏努利随机变量(Bernoulli random variable),而其发生的概率为 p 时,如果欲探讨的问题是:在 n 次相互独立的试验之中,并无所关心的事件发生,而却又想得知事件可能发生的概率,则此时使用"上限为三法则"(Rule of Three) 即可快速地提供笔者发生概率为 p 的事件其 95% 的信赖上限,而其结果等于 $3/n$。

在使用"上限为三的法则" 时有一个先决限制条件,即事件发生的概率 p 值应相当的小。例如某种缺陷发生的几率、某种产品品检不合格的几率等。举例来说,倘若笔者想了解某种缺陷的出现几率,现在观察 n 个实验样本中若该类缺陷,则笔者并不能下结论说这种缺陷的复发率是 0,如此的说法似乎是过于乐观了。笔者应找出一个适当且合理的方法,进而推出一个合适的置信区间。而此时"上限为三法则" 即能立即告知 95% 信赖上限等于 $3/n$。所以,当 $n = 10, p_u = 0.3; n = 100, p_u = 0.03; n = 1\,000, p_u = 0.003$;依此类推。(此处 p_u 为 p 的信赖上限)[1-6]。

1　理论推演

对于"上限为三法则" 的推导,将用三个不同的方法来加以验证。包括二项分布(Binomial distribution)、普阿松分布(Poissondistribution)以及贝氏法(Bayesian approach)。其中加入一段 Louis(1981) 先生对"上限为三法则" 的看法,在数理推导之外,看经由实际调查的结果是否依然支持本定理。

1) 二项分布(Binomial distribution)

倘若随机变量(random variable)X 是一个二项分布的变量,此变量具有两个参数(parameter)n(观察的个数) 和 p(事件发生的概率)。则观察 n 个观测值,而没有事件发生的概率可表示如下:

$$P(X = 0 \mid n,p) = \binom{n}{0} p^0 (1 - p)^n = (1 - p)^n$$

而在显著水准 α 之下,$(1 - \alpha)100\%$ 置信区间的信赖上限,可借助 $(1 - p)^n \geq \alpha$ 的限制条件式,来解得 $p \leq 1 - \alpha^{1/n}$,而我们将 p 的信赖上限 p_u 定义为 $1 - \alpha^{1/n}$,即

$$p_u = 1 - \alpha^{1/n}$$

换句话说,$(0, p_u)$ 这个范围可视为 p 的 $(1 - \alpha)100\%$ 置信区间。经由泰勒展开式(Taylor's expansion) 可得到

$$\alpha^{1/n} = 1 - \ln(\alpha)/n + [\ln(\alpha)]^2/2n^2 + \cdots$$

经移项后,可得

$$1 - \alpha^{1/n} \cong - \ln(\alpha)/n$$

即

* 收稿日期:2004 - 01 - 06
基金项目:教育部优秀青年教师资助计划项目(教人司[2003]40 号)
作者简介:杨劲(1968 -),男,重庆市人,重庆工商大学讲师,主要从事多媒体网络技术研究。

$$p_u \doteq -\ln(\alpha)/n$$

而在显著水准 $\alpha = 0.05$ 的情况之下，$-\ln(\alpha)$ 即 $-\ln(0.05) = 2.996$，因此，p_u 会十分接近 $3/n$。

2）普阿松分布（Poisson distribution）

倘若随机变量（random variable）X 之分布为普阿松分布，在参数 $\lambda = np$ 的情况下

$$P(X = 0 \mid \lambda = np) = \exp(-np)$$

根据限制条件式 $\exp(-np) \geqslant -\alpha$，在等号两边取自然对数（natural log）即可解得 $p \leqslant -\ln(\alpha)/n$，而可得

$$p_u = -\ln(\alpha)/n$$

而在取 $\alpha = 0.05$ 的情况下，p_u 即约等于 $3/n$。

3）Louis（1981）先生的说法[4]

在 Bickel and Doksum（1977, p.180）这本书中曾指出：若将下述的 S_n 视为在将来实验中，相同数目的试验对象（即 n 相同）之下，所关心的事件发生的数目时，则 $\lim\limits_{n\to\infty} S_n = -\ln(\alpha)$。加上之前已知的 $1-\alpha^{1/n} \cong -\ln(\alpha)/n$，则可以得到：

$$1 - \alpha^{1/n} = S_n/n$$

而 Louis 先生表示，在他所观察的试验中，$S_n = 3$ 是一个相当普遍能被接受的数值，而这刚好符合我们之前所求得的 95% 置信区间以上事件发生的概率。Louis 先生也指出说，他所调查的结果和具有"均匀分布事前概率 p"（uniform prior on p）的贝氏法所求得的结果相符合。

4）贝氏法（Bayesian approach）

假设事前概率 p 是属于 Beta(a,b)，则

$$\pi(p) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} p^{a-1}(1-p)^{b-1} \qquad 0 < p < 1$$

所以必须先决定参数 (a,b)，才能充分了解事前概率 p 所表现的特性。先假设事前概率 p 是符合 Beta$(1,b)$，$b \geqslant 1$；因为在 b 愈大于 1 的情况之下，事前概率 p 会愈接近 0；而当 $b = 1$ 时，事前概率 p 则会变成均匀分布 uniform$(0,1)$。而在这个情况之下，事前概率 p 则可说是最保守（conservative）的极限值。至于在参数 a 若不等于 1 时，所生成的事前概率 p 均会较 $a = 1$ 大，所以在没有特殊的情况之下，均取参数 a 为 1。图 1 所示为在 Beta$(1,b)$ 的情况之下，不同的 b 值对于事前概率 p 的影响。

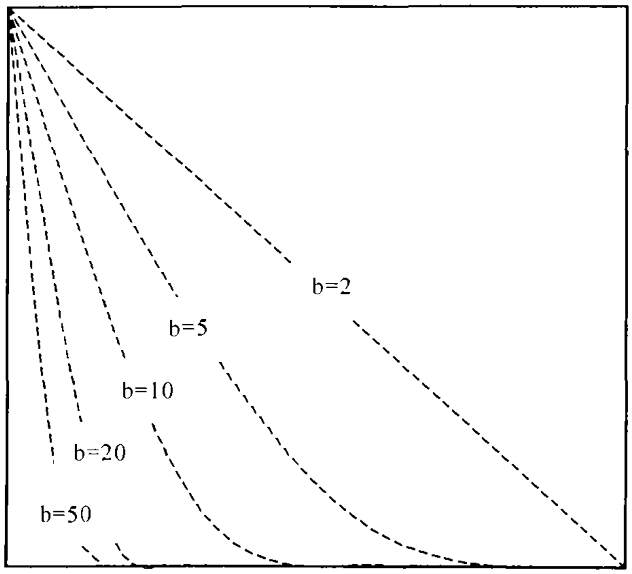

图 1 在 Beta$(1,b)$ 的情况之下，不同的 b 值对于事前概率 P 的影响

在 $\pi(p) \sim$ Beta$(1,b)$ 的情况下，可借助简单的积分，来计算出事后概率 p 的置信区间。（显著水准 α）

$$P(0 < p < p_u \mid X = 0, b, n = 1 - (1 - p_u)^{(n+b)} \geqslant 1 - \alpha$$

则

$$p_u \geqslant 1 - \alpha^{1/(n+b)}$$

再使用泰勒展开式（Taylor's expansion），可得到

$$1 - \alpha^{1/(n+b)} = -\ln(\alpha)/(n+b)$$

即

$$p_u = -\ln(\alpha)/(n+b)$$

所以用贝氏法所求得的 Rule of Three（Bayesian Rule of Three）为

$$3/(n+b)$$

因为 $b \geqslant 1$，故可以很清楚地知道，当 $b = 1$ 时，$3/(n+1)$ 为最大值，特称其为"改进后的上限为三法则"（Improved Rule of Three），而此时的事前概率 p 即属于均匀事前概率（uniform prior；Beta$(1,b)$，$b = 1$）。

2 结果分析与讨论

为了比较上述几种近似法（Rule of Three, Improved Rule of Three 及 Bayesian Rule of Three）何者近似效果最佳，所以将这三种近似法以 $\alpha = 0.05$ 带入，在不同的 n 值下相互比较，结果列于"表 1"。

表 1　实际数据对照($\alpha = 0.05, x = 0$)

n	$-\ln(\alpha)/n$	$1-\alpha^{1/n}$	$3/n$	$1-\alpha^{1/(n+1)}$	$3/(n+1)$	$3/(n+20)$
3	0.998 58	0.631 60	1.00 000	0.527 13	0.75 000	0.130 43
4	0.748 93	0.527 13	0.75 000	0.45 072	0.600 00	0.125 00
5	0.599 15	0.450 72	0.600 00	0.393 04	0.500 00	0.120 00
6	0.499 29	0.393 04	0.500 00	0.348 16	0.428 57	0.115 38
7	0.427 96	0.348 16	0.428 57	0.312 34	0.375 00	0.111 11
8	0.374 77	0.312 34	0.375 00	0.283 13	0.333 33	0.107 14
9	0.332 86	0.283 13	0.333 33	0.258 77	0.300 00	0.103 45
10	0.299 57	0.258 77	0.300 00	0.238 40	0.272 73	0.100 00
20	0.149 79	0.139 11	0.150 00	0.132 95	0.142 86	0.075 00
50	0.059 91	0.058 16	0.060 00	0.057 05	0.058 82	0.042 85
100	0.029 96	0.029 51	0.030 00	0.029 23	0.029 70	0.025 00

在"表1"中共列有7个字段:第1列"n"是指观察试验的数目;而第2到第7列皆为在没有事件发生的情况下($X=0$),95% 置信区间以上的概率;第 2 列"$-\ln(\alpha)/n$"是指在普阿松分布下;第3列"$1-\alpha^{1/n}$"是指在二项分布下;第 4 列"$3/n$"是使用"上限为三法则"求得的近似值;第5列"$1-\alpha^{1/(n+1)}$"是指在事前概率p属于均匀分布时(Beta$(1,b)$,$b=1$),利用贝氏法来求值;第 6 列"$3/(n+1)$"则是使用"改进后的上限为三法则"Improved Rule of Three 来求值;第 7 列"$3/(n+20)$"亦是使用贝氏法来求值,但是取 $b=20$。在这里取 $b=20$ 的意义为之前我们随机取19个观测值,而并无我们所关心的事件发生。

3　结论

根据上述表格中的数据,可以得到以下几点结论:

1) 当n越大时,$1-\alpha^{1/n}$、$1-\alpha^{1/(n+1)}$ 以及它们所对应的近似值$3/n$和$3/(n+1)$即越相近。事实上六种方法所求得的值也是随着 n 越大而越相近,倘若 $n\to\infty$ 时,Rule of Three($3/n$) 应具有相当的代表性。

2) 再进一步的观察我们可发现,在本表当中($n\leqslant 100$) 由"改进后的上限为三法则""$3/(n+1)$"所求得的近似值,似乎较"上限为三法则"($3/n$) 求得的近似值更接近 $1-\alpha^{1/n}$。

3) 倘若对事前概率p有所了解的话,则可决定一个合适的"b"值, 进而采用 Bayesian"上限为三法则""$3/n+b$"的方法来求值,如此应较为准确。

4) 在 $n\geqslant 20$ 的情况下,我们可说"上限为三法则""$3/n$"有相当不错的近似效果,故我认为$n\geqslant 20$可说是"上限为三法则"使用的限制条件。

参考文献

[1] HOGG R V , TANIS E A . Probability and Statistical Inference[J]. Journal of the American Mathematical Association,1997(7):38 - 42.

[2] HANLEY J, LIPPMAN A. Is Everything All right? [J]. Journal of the American Mathematical Association, 1983 (13):143 - 145.

[3] JOVANOVIC B D, LEVY P S. A Look at the Rule of Three [J]. The American Statistician, 1998(51):137 - 139.

[4] LOUIS T A. Confidence Intervals for a Binomial Parameter after Observing no Successes [J]. The American Statistician, 1981,(35):154.

[5] PAUL S. Bayesian Statistics: Principles, Models and Applications[M]. New York: New York Press,1989. 564 - 632.

[6] VOLLSET S E. Confidence Intervals for a Binomial Proportion[J], Statistics in Medicine, 1993,(12):809 - 824.

Preconditions About Rule of Three

YANG Jin

(Computer College, Chongqing Technology and Business University, Chongqing 400020, China)

Abstract: In those researches of sample test, especially in the works of experimental data counting and statistics, Rule of Three has been applied widely. But few people understand its origin really , and there are little research work have been down on the theory's reasoning. It has be regard as a general rule. Four different ways. Binomial distribution, Poisson distribution, Louis distribution and Bayesian approach are used, Which are. Then the rule is verified and companed through real investigation data, and four tenable conditions of the rule are given.

Key words: Rule of Three; probability statistics; probability distribution

(编辑　张小强)