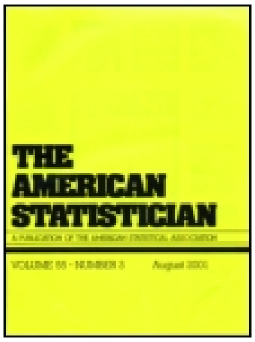

Taylor & Francis
Taylor & Francis Group

The American Statistician

ISSN: 0003-1305 (Print) 1537-2731 (Online) Journal homepage: http://www.tandfonline.com/loi/utas20

A Look at the Rule of Three

B. D. Jovanovic & P. S. Levy

A Look at the Rule of Three

B. D. JOVANOVIC and P. S. LEVY

The Rule of Three states that $3/n$ is an upper 95% confidence bound for binomial probability p when in n independent trials no events occur. We discuss the derivation of this rule, its validity in small samples, and propose some alternatives. The material may serve well as a thought-provoking introduction to a clinical trials, statistical consulting, or a categorical data class, and is interesting on its own merit.

KEY WORDS: Binomial probability; Clinical research; Safety.

1. INTRODUCTION

Consider a Bernoulli random variable with unknown probability p. If in n independent trials no events occur, the Rule of Three provides a quick-and-ready approximation to the upper 95% confidence bound for p, as $3/n$. This rule of seemingly unknown origin was first discussed in medical literature by Hanley and Lippman-Hand (1983), although Louis (1981) provided a short and somewhat technical discussion of it in *The American Statistician.* The appeal of the Rule of Three to a clinician is in its simplicity and usefulness in safety evaluation of clinical procedures or other research, in which it is hoped that no adverse events will occur. Thus the probability p is a priori known to be small, and when in n patients no events occur, this suggests a successful intervention or a safe procedure. The question regarding the upper bound for p then translates to a question very important to a clinician: "What is the worst possible scenario for p given that in n treated patients no adverse events occurred"? The clinician familiar with the Rule of Three will immediately know that the 95% upper confidence bound for p is $3/n$, that is, .3 when $n = 10$, .03 when $n = 100$, .003 when $n = 1000$, etc. Alternatively, a clinician may require a sample size calculation in order to determine the number of patients needed to determine (with 95% confidence) that $p < p_0$ if no events occur. In this case $n \geq \text{int}(3/p_0)$ would be the smallest such sample size.

It is worth noting that in the context of safety evaluation in clinical research, in which a clinician has to "prove" the safety of a new procedure (i.e., show that the probability of an adverse event is lower than some small acceptable probability), the lower bound for p is not of practical interest; one is mainly concerned with the upper bound on p as it represents the "worst case scenario" or the largest probability for placing a patient at risk. Thus the lower bound for p may a priori be taken to equal zero.

B. D. Jovanovic is a Research Assistant Professor and P. S. Levy is a Professor and Director, Division of Epidemiology and Biostatistics, School of Public Health, University of Illinois at Chicago, Chicago, IL 60612.

2. DERIVATION OF THE RULE OF THREE

Derivation of the Rule of Three presents an interesting exercise in elementary statistics. We first paraphrase the argument given by Hanley and Lippman-Hand (1983) prepared for a medically inclined audience. Let the random variable X have a Binomial distribution with parameters n and p. Then $P(X = 0|n, p) = (1-p)^n$, and one can obtain a $(1-\alpha)$ 100% upper bound for p by solving $(1-p)^n \geq \alpha$ for p. This yields $p_u \geq 1 - \alpha^{1/n}$, and by taking

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

for the least upper bound for p, the interval $(0, p_u)$, in some sense provides $(1-\alpha)$ 100% coverage for true p. Now, $3/n$ appears for the following reason. From Taylor expansion $\alpha^{1/n} = 1 + \ln(\alpha)/n + [\ln(\alpha)]^2/2n^2 + \ldots$, by retaining only the linear portion, one obtains

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

For $\alpha = 0.05$, $-\ln(\alpha) = 2.996$, and thus p_u is numerically close to $3/n$. A similar argument using a Poisson random variable with $\lambda = np$ yields $P(X = 0|np) = \exp(-np) \geq \alpha$ that, after taking the natural log of both sides, produces the least upper bound

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

that yields $3/n$ as in the Binomial case. Thus Hanley and Lippman-Hand derive the upper bound for p by assigning probability of at least α to the observed value $X = 0$. Although this in some sense does the job of providing an upper bound for p, no explicit reference to an interval is given in their derivation, and someone less inclined to make a leap of faith will need an additional step. This additional step is rather simple, and consists of observing that the same upper bound may be derived using exact Binomial probabilities. In particular, if $X = x$ is the observed number of events in n trials, the Clopper–Pearson (max-P) upper $(1-\alpha)$ 100% bound may be obtained as a solution to

$$\sum_{t=0}^{x} \binom{n}{t} p^t (1-p)^{n-t} = \alpha$$

(cf. Vollset, 1992). Clearly, when $x = 0$ the expression reduces to $(1-p)^n = \alpha$, and one has a more sound justification of the derived upper bound.

Louis (1981) provided a slightly different and conceptually more involved derivation of the Rule of Three. He quoted Bickel and Doksum (1977, p. 180) pointing out that

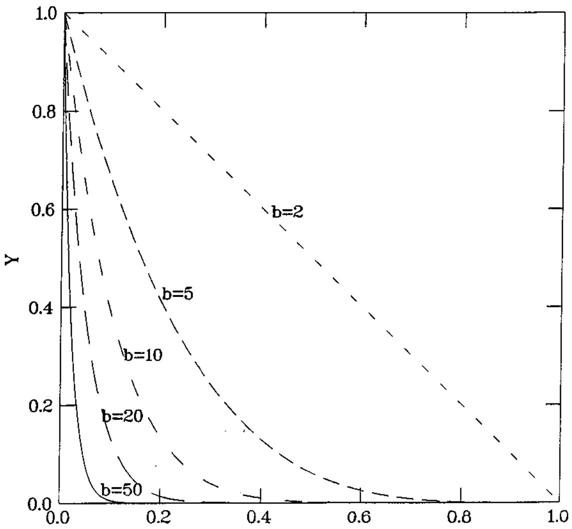

Figure 1. Beta(1, b) Prior Kernels for Various Values of b.

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

where S_n can be viewed as the number of events in a future experiment of the same size, with $\lim_{n\to\infty} S_n = -\ln(\alpha)$. He then argued, based on a survey of students and clinicians, that $S_n = 3$ is an absolute upper bound acceptable to all of the subjects, and that this corresponds to a 95% upper confidence bound for p. He also pointed out that this corresponds to the upper Bayesian credibility bound for a uniform prior on p.

In fact, a Bayesian approach to the problem seems quite instructive and is discussed next. Assume a Beta(a, b) prior on p, that is, $\pi(p) = p^{a-1}(1-p)^{b-1}/B(a, b)$. We need to choose a and b in such a way as to ascertain agreement with the a priori knowledge about p. We know that p is close to zero, without a reasonable lower bound other than zero. Taking a Beta$(1, b), b \geq 1$ prior on $p \in (0, 1)$ seems a good choice because for $b > 1$ the prior favors values of p close to zero, although in the case of $b = 1$ it produces a uniform prior on (0, 1), which may be considered a limiting and a most conservative prior in this context. In addition, values for a other than $a = 1$ provide the prior with a local maximum away from zero that, without additional information, cannot be justified. Figure 1 presents a sequence of kernels of Beta$(1, b)$ priors for various values of b.

With $\pi(p) \sim$ Beta$(1, b)$ simple integration yields a posterior credibility interval for p:

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

which simplifies to

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

The right-hand side can be approximated via Taylor expansion by

$$-\ln(\alpha)/(n + b).$$

This gives us a Bayesian Rule of Three as

$$3/(n + b)$$

with $b \geq 1$. Obviously, for $b = 1$, $3/(n + 1)$ is the largest such upper bound, corresponding to the uniform prior.

We next give a scenario describing how one might come up with a prior value of b possibly acceptable to a clinician. Suppose there is a completed safety evaluation study with n_1 patients and $x_1 = 0$ events. If we use the Rule of Three, this study will yield an upper bound on p as $3/(n_1 + 1)$, which corresponds to uniform prior, that is, Beta(1, 1). The standard result (Press, p. 41) is that if Beta(a, b) is the prior distribution on p and $X = x$ is observed in n trials, the posterior distribution for p is Beta$(a + x, b + n - x)$. Consequently, for prior Beta(1, 1) and no observed events in n_1 patients we have the posterior Beta$(1, 1 + n_1)$. Next, $n_1 + 1$ may be taken as a plausible value for b in the subsequent safety evaluation study with n_2 patients. If no events occur, the Rule of Three gives $3/(n_2 + n_1 + 1)$. Clearly, a sequence of k studies with no observed events will produce $3/(n_1 + n_2 + \cdots + n_k + 1)$ as the upper 95% confidence bound. We believe this line of reasoning could be easily accepted by a clinician. Alternatively, a clinician may be willing to base an initial estimate of b on a "conceptual experiment," which involves prior beliefs about the underlying rate p, about its moments or its range; details may be found, for example, in Novick and Jackson (1974, p. 157).

3. NUMERICAL RESULTS

Following arguments in Section 2 it is tempting to con-

Table 1. Upper Bounds on p when x = 0: Poisson (2), Exact Binomial (3), Rule of Three (4), Bayesian Upper Bound for Uniform (Beta(1, b),b = 1) Prior (5), Improved Rule of Three (6), Bayesian Rule of Three for b = 20 (7)

(1) n	(2) $-\ln(\alpha)/n$	(3) $1 - \alpha^{1/n}$	(4) $3/n$	(5) $1 - \alpha^{1/(n+1)}$	(6) $3/(n+1)$	(7) $3/(n+20)$
3	.99858	.63160	1.00000	.52713	.75000	.13043
4	.74893	.52713	.75000	.45072	.60000	.12500
5	.59915	.45072	.60000	.39304	.50000	.12000
6	.49929	.39304	.50000	.34816	.42857	.11538
7	.42796	.34816	.42857	.31234	.37500	.11111
8	.37477	.31234	.37500	.28313	.33333	.10714
9	.33286	.28313	.33333	.25877	.30000	.10345
10	.29957	.25877	.30000	.23840	.27273	.10000
20	.14979	.13911	.15000	.13295	.14286	.07500
50	.05991	.05816	.06000	.05705	.05882	.04285
100	.02996	.02951	.03000	.02923	.02970	.02500

sider $3/(n+1)$ as an improvement to the Rule of Three. For large n the bounds $1-\alpha^{1/n}$, $1-\alpha^{1/(n+1)}$, and their respective approximations $3/n$ and $3/(n+1)$ are close to each other, but for $n \leq 20$ the discrepancy seems worth noting. In fact, Table 1 shows that $3/(n+1)$ provides a uniformly better approximation to $1-\alpha^{1/n}$ than $3/n$ does, and could be considered as an alternative on that merit only. (Such better approximation is only an artifact of using the linearized portion of the Taylor expansion for the Rule of Three.) If prior knowledge regarding b is available, the Bayesian Rule of Three $3/(n+b)$ could be considered. We illustrate these points in Table 1.

Column (4) in Table 1 shows high accuracy of the Rule of Three for $n \geq 20$. Column (6) shows that $3/(n+1)$ approximates $1-\alpha^{1/n}$ uniformly better than $3/n$ does. Column (7) shows the Bayesian Rule of Three with $b = 20$, which would correspond to a prior study with 19 patients and no events, with the current study size given in column (1). For $n \geq 20$ the difference between various upper bounds (2)–(6), for most practical purposes, diminishes.

4. CONCLUSION

The Rule of Three, although very useful in clinical safety evaluation and screening studies, does not seem to be widely known either among clinicians or among statisticians. At the same time this rule seems to be a part of the "folklore" associated with clinical research. In our consulting practice on several occasions we have met with researchers who had vague knowledge of such a rule, and requested clarification from us. This prompted us to take a closer look at the Rule of Three. This proved to be a rewarding exercise. In the classroom the Rule of Three caused debates among our graduate students. For homework we assign a closer look at the 90%, 99%, and 99.9% upper bounds on p, and ask the question: "Any more rules?"

[Received April 1995. Revised May 1996.]

REFERENCES

Bickel, P. J., and Doksum, K. A. (1980), *Mathematical Statistics*, San Francisco: Holden-Day.

Hanley, J. A., and Lippman-Hand, A. (1983), "If Nothing Goes Wrong, Is Everything Allright? Interpreting Zero Numerators," *Journal of the American Medical Association*, 249(13), 1743–1745.

Kerns, J. R., Shaub, T. B., and Fontanarosa, P. B. (1993), "Emergency Cardiac Testing in the Evaluation of Emergency Department Patients with Atypical Chest Pain," *Annals of Emergency Medicine*, 22, 794–798.

Louis, T. A. (1981), "Confidence Intervals for a Binomial Parameter after Observing no Successes," *The American Statistician*, 35(3), 154.

Novick, M. R., and Jackson, P. H. (1974), *Statistical Methods for Educational and Psychological Research*, New York: McGraw-Hill.

Press, S. J. (1989), *Bayesian Statistics: Principles, Models and Applications*, New York: Wiley.

Vollset, S. E. (1993), "Confidence Intervals for a Binomial Proportion," *Statistics in Medicine*, 12, 809–824.