LETTERS TO THE EDITOR

Letters to the editor will be confined to discussions of papers that have appeared in *The American Statistician* and to important issues facing the statistical community. Letters discussing papers in *The American Statistician* must be received within two months of publication of the paper; the author of the paper will then be given an opportunity to reply, and the letters and reply will be published together. All letters to the editor will be refereed. Corrections of errors that have been noted in papers published in *The American Statistician* will be listed as corrections at the end of this section.

INMAN, HENRY F. (1994), "KARL PEARSON AND R. A. FISHER ON STATISTICAL TESTS: A 1935 EXCHANGE FROM *NATURE*," *THE AMERICAN STATISTICIAN*, 48, 2–11: COMMENT BY CHRISTENSEN AND REPLY

The exchange on statistical testing between Fisher and Pearson contained in Inman (1994) was very interesting. Surprisingly, I found myself agreeing with both Fisher and Pearson. The purpose of my letter is to explain this phenomenon, and to make some gratuitous comments on prediction, one-sided tests, and confidence intervals. In particular, it seems popular these days to denounce testing, especially "two-sided" testing, in favor of confidence intervals. I disagree with this trend on philosophical, if not practical, grounds.

Consider the simplest standard case: $y_1, \ldots, y_n$ independent $N(\mu, \sigma^2)$ and testing $H_0 : \mu = 0$. There are at least five assumptions being made here: (1) that the observations are independent, (2) that the observations are normally distributed, (3) that the observations all have the same mean, (4) that the observations all have the same variance, and (5) that $\mu = 0$. Together, the assumptions constitute a model for the data. A model allows us to conduct science, that is, make predictions and evaluate the accuracy of those predictions. If we have future observations to which we can apply this model, our point prediction for them would be 0 and a 95% interval prediction contains the values between $0 \pm t(.975, n)\hat{\sigma}$ where $\hat{\sigma}^2 = \sum_{i=1}^n y_i^2/n$.

The essence of Statistics is finding better and more useful models for observable phenomena. In practice, one must agree with Pearson that models are never actually correct. The real question is whether they are useful. Where do useful models come from? Who knows? Often, they result from interactions between statisticians and subject matter specialists.

Testing plays its role in establishing whether models are tenable. You can never prove that models are correct, but they come from somewhere, and to conduct science we must be willing to stand by our models until they are no longer tenable. Different people can have different models, and there will be no way to choose between them until some models are shown to be untenable. As Fisher points out, a statistical test is essentially an attempt at proof by contradiction. We assume the model and ask whether the observed data contradict it. Pearson mentions that absolute contradictions are rare, but his point does not vitiate the force of Fisher's argument. In our normal theory example the model indicates that $(\bar{y} - 0)/(s/\sqrt{n})$ has a $t(n-1)$ distribution. The support of a $t(n-1)$ is the entire real line, so no data can give an absolute contradiction to the model (at least when conducting a t test). On the other hand, some possible values of $(\bar{y} - 0)/(s/\sqrt{n})$ are so unusual for a $t(n-1)$ distribution that if they were observed, the model would be called in question. The P value is simply a measure of how rare a particular value for $(\bar{y} - 0)/(s/\sqrt{n})$ is relative to the $t(n-1)$ distribution. The conclusion of the test is that we have either seen a rare event, as measured by the P value, or the model is incorrect. As the P value gets smaller, it becomes progressively more logical to assume that the model is incorrect.

How do you compute the P value? The $t(n-1)$ distribution has a density. Rare values of $(\bar{y} - 0)/(s/\sqrt{n})$ are those values that give small values for the density. The P value is simply the probability of getting a value of $(\bar{y} - 0)/(s/\sqrt{n})$ with a density that is less than or equal to the density of the observed value of $(\bar{y} - 0)/(s/\sqrt{n})$. Note that the P value thus obtained is identical to that in the usual "two-sided" test.

None of this has anything to do with an alternative hypothesis. Either the data are consistent with the model or are inconsistent with the model. If they are consistent, they yield very little new information because we have attempted a proof by contradiction without getting a contradiction. Consistent data certainly do not prove the model to be correct. If the data are inconsistent with the model, the model is inadequate. (Although sometimes demonstrably inadequate models can still be useful.) In our example the model involves at least five assumptions. The inadequacy of the model could stem from violation of any or all of those assumptions. Inconsistent data certainly do not imply that the particular assumption embodied in $H_0 : \mu = 0$ is the one assumption that was violated. Concluding that $\mu \neq 0$ is only reasonable if the other four assumptions have been validated.

Procedures for testing $H_0 : \mu = 0$ vs. $H_A : \mu \neq 0$ or, say, $H_0 : \mu = 0$ vs. $H_A : \mu > 0$, whether they be Neyman–Pearson or Bayesian, take for granted that the first four assumptions are correct. Given a belief in the first four assumptions, my personal preference is to conduct a Bayesian analysis, but I find the choice between Bayesian or Neyman–Pearson to be of little importance compared to the crucial issue of validating the assumptions. The point here is that it is easy to test whether data are inconsistent with a model, but it is very difficult to conduct a valid test in either a Neyman–Pearson or Bayesian setting.

I have already indicated that I think prediction is fundamental, and my dislike for one-sided alternatives stems from a basic dislike of alternatives. My last point is a question, "What is this creature we call a confidence interval?" We start with a perfectly valid probability statement about future observables. In the example, ignoring H_0, it is

$$1 - \alpha = \Pr\left[\bar{y} - t\left(1 - \frac{\alpha}{2}, n-1\right)\frac{s}{\sqrt{n}} < \mu < \bar{y} + t\left(1 - \frac{\alpha}{2}, n-1\right)\frac{s}{\sqrt{n}}\right].$$

How do we substitute observed values for $\bar{y}$ and s and *magically* turn this into $(1-\alpha)100\%$ "confidence" that μ is between $\bar{y} \pm t(1-\frac{\alpha}{2}, n-1)\frac{s}{\sqrt{n}}$? What in the world could "confidence" possibly mean, and how in the world does the $1 - \alpha$ convert itself from probability into confidence? And please do not give me the long-run frequency interpretation; I know the Law of Large Numbers. It rarely applies, and even when it does, it solves nothing. We all know that, except for those that have been browbeaten into giving the "correct" answer, everyone treats confidence as a synonym for probability. So this approach to confidence seems like a blatant attempt at making Bayesian omelettes without breaking Bayesian eggs or even admitting to making omelettes. At least Fisher's fiducial approach to inverse probability admitted it was making omelettes.

My real point here is simply that frequentists should not bad mouth testing while praising confidence intervals. Until they can give good answers to the questions raised in the previous paragraph, a confidence interval must be simply the collection of parameter values that are consistent with the data as determined by an α level test. Moreover, prediction intervals can be thought of as those future observable values that are consistent with the model based on an α level test. I grant that it is stupid to do a basic α level test because either a P value or a confidence interval is uniformly more informative. However, testing seems to be the foundation of frequentist inference.

Ronald Christensen
Department of Mathematics and Statistics
University of New Mexico
Albuquerque, NM 87131

I am glad that Professor Christensen found Inman (1994) interesting and that it provoked him to examine what a statistical test actually achieves. Because one aim of my article was to demonstrate that confusion among statisticians and scientists regarding the relationship between testing statistical hypotheses and scientific inference is not a recent development, I welcome his effort to grapple with this issue.

However one judges their claims to scientific relevance, the Neyman–Pearson and neo-Bayesian arguments have contributed significantly to the clarity and rigor of the theory of testing hypotheses *as a statistical problem*. In my view this is largely due to the fact that both approaches explicitly incorporate the alternative hypothesis and provide a decision rule that is based on sampling models that apply under the alternative as well as the null hypothesis. As I noted in Inman (1994), Karl Pearson and R. A. Fisher adhered to the tradition of the classical test of significance. Although both men recognized and addressed issues related to sampling procedures in their work, neither saw validation of a sampling model as a necessary requirement for performing a statistical test. The difference between their views lies in how the results of the test should be interpreted; this, in turn, is based on how Pearson and Fisher regarded probability statements and the sorts of scientific investigations they pursued.

Fisher argued that the logical force of a statistical test stemmed from the rejection of the null hypothesis tested, but much of Fisher's work involved the analysis of designed experiments. Although claiming to avoid formal statistical tests altogether, Pearson noted the scientific value of adopting, at least provisionally, a model fit to data which (in most of his work) derived from what we would now classify as observational investigations. Pearson, it seems to me, saw statistical methodology within a framework consistent with Bayesian inference, but because Pearson rejected the truth of scientific models he did not recognize the Bayesian goal of attaching probabilities to hypotheses *as measures of truth*. Instead, Pearson interpreted probability statements in terms of hypothetical future observations generated by the phenomenon under study. For Pearson an adequate statistical model, fit to scientific data (however obtained), incorporated within it the stochastic basic for probability calculations.

Finally, I sadly must take this opportunity to note the death of Arthur J. Lee, who retired in 1980 as Director of Fisheries Research at the Ministry of Agriculture, Fisheries and Food's Lowestoft Laboratory. Mr. Lee kindly communicated to me his own recollections of H. J. Buchanan–Wollaston; provided me the address of Geoffrey Wollaston, Buchanan–Wollaston's son; and guided me to relevant information in Lee (1992). Without his help Inman (1994) would have been far less interesting to write and (I believe) to read. Unfortunately Mr. Lee died shortly before the article appeared in *The American Statistician*; thus he was unable to see the result of the assistance he generously gave me.

Henry F. Inman
2016 A Park Avenue
Richmond, VA 23220

REFERENCE

Lee, A. J. (1992), *The Directorate of Fisheries Research: Its Origins and Development*, Lowestoft, England: Ministry of Agriculture, Fisheries and Food, Directorate of Fisheries Research for England and Wales.

KELLER, J. B. (1995), "A CHARACTERIZATION OF THE POISSON DISTRIBUTION AND THE PROBABILITY OF WINNING A GAME," *THE AMERICAN STATISTICIAN*, 48, 294–298: COMMENTS BY MENG AND SEN AND REPLY

I read this article with joy—it is fun and intriguing. However, its Appendix, and thus the extension to continuous variables, is built upon an elementary error.

The Appendix started with: "Let X and Y be real-valued random variables with densities $p(x, \lambda)$ and $q(y)$, respectively, where λ is the mean of X." It then arrived at

$$\Pr[X = Y] = \int p(y, \lambda) q(y)\, dy. \tag{A.2}$$

This is obviously wrong, even assuming the independence between X and Y, an assumption I deduced from the integrand as well as from the treatment of the discrete cases in the article. In fact, when the distribution of $X - Y$ is continuous, $\Pr[X = Y] = 0$. Another simple way of rejecting (A.2) is to note that its right side may exceed 1. For instance, suppose $X, Y \sim {}_{\text{iid}}N(0, \sigma^2)$; then the right side is $(2\sqrt{\pi}\sigma)^{-1}$, which exceeds 1 for $\sigma < (2\sqrt{\pi})^{-1}$.

All of these, of course, are well known. I surmise that the error was a simple "mindo" (that is, a "typo" of the mind), as the article's main focus was on discrete cases, and (A.2) was thought to be a natural extension of the discrete setting. I also surmise that it would be more appropriate to publish such an extension in *Teacher's Corner*, for it provides a good classroom illustration of the need for caution when generalizing from the discrete case to the continuous one.

Xiao-Li Meng
Department of Statistics
University of Chicago
Chicago, IL 60637
meng@galton.uchicago.edu

Keller (1994) provides an interesting interpretation of the Poisson distribution in terms of probabilities of winning and drawing a game. I would like to open this letter by congratulating the author on an innovative and stimulating piece of work. The intent of this letter, however, is to point out a substantial simplification of the treatment presented in the appendix of the aforementioned article, and subsequently establish a general result which is apparently overlooked in the appendix.

Specifically, the author starts with the defining relation (A.3), namely,

$$\frac{\partial}{\partial \lambda} \Pr[X > Y] = \Pr[X = Y] \tag{1}$$

for every continuous random variable Y independent of X where X is continuous with finite mean λ. From this, the author, via a differential equation approach, derives a characterizing Equation (A.8) for the distribution of X, and consequently establishes that no nonnegative random variable X satisfies (1). The treatment simplifies considerably by noting that the right-hand side of (1) actually equals zero for independent, continuous random variables. The assumption of independence (also assumed by the author) is crucial because dependence between X and Y may lead to a positive value of $\Pr[X = Y]$ (e.g., in the case of Marshall and Olkin's (1967) bivariate exponential distribution). In the following I will establish that, in fact, there is no continuous random variable X for which (1) holds, a stronger contention than that arrived at by the author.

Note that in view of our observation, (1) entails that $\Pr[X > Y]$ is free of λ for every continuous random variable Y, independent of X. In particular, choosing Y to be a *Uniform*$(0, \theta)$ random variable, we get $\int_0^\theta \Pr[X > x]\, dx$ free of λ for every $\theta > 0$, and so

$$\int_0^\infty \Pr[X > x]\, dx = \lim_{\theta \uparrow \infty} \int_0^\theta \Pr[X > x]\, dx \tag{2}$$

is free of λ. But for $X \geq 0$, the left-hand side of (2) equals $E(X) = \lambda$, a contradiction, establishing that (1) fails to hold for any nonnegative continuous X. For a general continuous X the expression for the mean equals

$$\lambda = E(X) = \int_0^\infty \Pr[X > x]\, dx - \int_{-\infty}^0 \Pr[X \leq x]\, dx. \tag{3}$$

An argument very similar to the above enables us to conclude that both terms on the right of (3) are free of λ, thus reaching the desired contradiction in the general case.

As a concluding note I would like to mention that a similar characterization for geometric random variables exists in connection with probabilities of wins and ties. Specifically, we need to replace the author's condition (2.1) by

$$\Pr[P > Q] \propto \Pr[P = Q]. \tag{4}$$

It is easily seen that (4) holds for all nonnegative, integer-valued random variables Q independent of P (also nonnegative integer-valued) if and only if P is geometric with mean equaling the constant of proportionality. In this case, as in the Poisson case, the crux of the proof rests on the fact that the class Q is rich enough to include all degenerate distributions which enable us to extract the exact structure of the distribution of P.

Ananda Sen
Department of Mathematical Sciences
Oakland University
Rochester, MI 48309

REFERENCE

Marshall, A.W., and Olkin, I. (1967), "A Multivariate Exponential Distribution," *Journal of the American Statistical Association*, 62, 30–44.

Xiao-Li Meng has pointed out a blunder in Equation (A.2) in the appendix to my paper, which invalidates the "proof" of the conclusion stated there. Fortunately, Ananda Sen has presented a correct proof of an even stronger conclusion. He has also given a characterization of geometric random variables, analogous to my characterization of Poisson variables. I apologize for the blunder and thank the authors for pointing it out and giving a correct proof, respectively.

Joseph B. KELLER
Department of Mathematics and Mechanical Engineering
Stanford University
Stanford, CA 94305

COYLE, C. A., AND WANG, C. (1993), "WANNA BET? ON GAMBLING STRATEGIES THAT MAY OR MAY NOT WORK IN A CASINO," *THE AMERICAN STATISTICIAN*, 47, 108–111: COMMENT BY DONEGAN

The article by Coyle and Wang (1993) presents some interesting comparisons among various gambling strategies. The first comparison involves a game (called Game 1) based on tossing a "fair" coin, and another game (Game 2) based on an "almost fair" roulette wheel with 38 possible outcomes. The first result in the article states that with the games as described, the gambler is more likely to reach his or her target in game 1 rather than game 2. This is, of course, true. But, if game 2 were to be based on a French roulette wheel (with only a single zero, i.e., 37 possible outcomes), then the situation would be reversed.

Redefine Game 2.

Game 2A: The gambler will bet \$1 on either red or black on an *American* roulette wheel (the probability of winning is 18/38 = .474).

Game 2F: The gambler will bet \$1 on either red or black on a *French* roulette wheel (the probability of winning is 18/37 = .486).

Readers who are familiar with casinos outside North America will be aware that, despite the differing probabilities of a winning bet, the odds paid are identical. It is obvious that game 2F is closer to being "fair" than game 2A.

Following the notation of Coyle and Wang, and keeping other parameters unchanged, the games can be summarized as follows:

	Capital		Goal	
Game 1	\$900	→	\$1,000,000	(p = 50%)
Game 2A	\$900	→	\$1,000	(p = 47.4%)
Game 2F	\$900	→	\$1,000	(p = 48.6%).

The probabilities of reaching the goal can be shown to be

Game 1	$h(\$900) = .09\%$
Game 2A	$h(\$900) = .0027\%$
Game 2F	$h(\$900) = .4486\%$.

Note that the probability of success in game 2F is considerably higher than in Game 1.

Introducing Game 2F into the analysis involving the Dubins–Savage strategy yields the following comparisons which do not show such dramatic differences.

The probabilities of reaching the goal and expected winnings are as follows:

Game 2A	$h(9) = 87.94\%$	E(Winnings) = −\$20.53
Game 2F	$h(9) = 88.98\%$	E(Winnings) = −\$10.24.

Thus employing the Dubins–Savage strategy in French roulette, a gambler is about 200 times more likely to reach goal than using a \$1 bet.

Comparisons of the Dubins–Savage strategy with variations of the original (constant bet) strategy are given in the following tables.

Game 2A

Bet	*h(9)*	*E(loss)*	*m(9) = expected number of hands to completion*	*Expected loss using D–S over same time*
\$1	.000027	\$899.97	17,099	\$168,690
\$5	.12158	\$778	2,958	\$29,181
\$20	.58837	\$312	296	\$2,920
D–S	.87942	\$21	2.0856	

Game 2F

Bet	*h(9)*	*E(loss)*	*m(9) = expected number of hands to completion*	*Expected loss from using D–S over same time*
\$1	.00448	\$895.51	33,134	\$166,089
\$5	.33913	\$561	4,150	\$20,805
\$20	.74612	\$154	285	\$1,427
D–S	.88976	\$10	2.0436	

Now consider the case of the Japanese whale. We are not informed as to which game Mr. Kashiwagi played, so we shall consider three cases. The table below shows results based on American roulette (A), French roulette (F), and the "best" case as suggested by Coyle and Wang (p = 49.375%).

	Game 2A	*Game 2F*	*Best*
p	.473684	.486486	.49375
$1-p$	.526315	.513513	
K	60	60	60
G	120	120	120
$h(60)$	.001793	.037541	.182413
Expected loss	\$11,942,580	\$10,798,640	\$6,162,740
Expected number of hands	1,135.910	2,053.314	3,048.826

The expected number of hands reported by Coyle and Wang (≈5,000) should be 3,049, while it is considerably less on either American or French roulette. In this light it is surprising that Mr. Kashiwagi lasted as long as he did. In this case even crossing the Atlantic would not have been any advantage to him.

Apart from the slight inaccuracies in the computations there is a typographic error in the article. On p. 109, right column, paragraph 4, the equation should be:

$$h(9) = p + \frac{pq + p * q^2}{1 - (pq)^2} = h(\$900).$$

Kevin DONEGAN
Faculty of Business and Technology
University of Western Sydney—Macarthur
Campbelltown, NSW 2560
Australia

TIETJEN, G. (1993), "RECURSIVE SCHEMES FOR CALCULATING CUMULATIVE BINOMIAL AND POISSON PROBABILITIES," *THE AMERICAN STATISTICIAN*, 48, 136–137: COMMENT BY NEMENYI

For the calculation of Binomial individual, right tail, and cumulative (= left tail) probabilities, Tietjen points out computational advantages of using one of the recursions

$$\Pr_n(r) = p \cdot \Pr_{n-1}(r-1) + q \cdot \Pr_{n-1}(r) \qquad (1a)$$

beginning with 0 q p 0 in a sea of zeros,

$$\mathrm{Tail}_n(r) = p \cdot \mathrm{Tail}_{n-1}(r-1) + q \cdot \mathrm{Tail}_{n-1}(r) \tag{1b}$$

beginning with 1 1 p 0 0,

$$\mathrm{Cum}_n(r) = q \cdot \mathrm{Cum}_{n-1}(r) + p \cdot \mathrm{Cum}_{n-1}(r-1) \tag{1c}$$

starting with 0 0 q 1 1, or

$$\mathrm{Pr}_n(r) = (n-r+1)p \cdot \mathrm{Pr}_n(r-1)/(qr) \tag{2}$$

beginning with $\mathrm{Pr}_n(0) = q^n$—their advantages over a closed formula like

$$\mathrm{Tail}_n(r) = \sum_{j=r}^{n} \frac{n!}{j!(n-j)!} \cdot p^j(1-p)^{n-j}. \tag{3}$$

(The notation here is a bit different from Tietjen's, but self-explanatory.)

It should be noted that (1) is free of underflow and overflow problems. If d = how many decimals matter for your purposes (plus a safety margin of 3 to allow for accumulation of rounding errors, if you plan to take n pretty high), then immediately round off any calculated probability $< .5 \cdot 10^{-d}$ to 0. (Modify this if you measure accuracy in significant digits.)

To me the most important use for (1) and then (2) is in education, to let Elementary Statistics be elementary. For a beginning student with little math, (3) is intimidating and likely to produce paralysis before statistical inference even begins. The simple recursions (1) follow directly from verbal definitions of Pr, Tail, and Cum and rudimentary rules of probability, and are easy to see happening numerically in a table. In the symmetric case $p = .5$, it means you simply keep averaging two adjacent entries to get the one below. In the general case you have weighted averages ("p of those plus q of these").

A nonparalytic introduction to statistics can begin with Arbuthnot's (1710) proof of Divine Providence based on 82 available years of London birth records showing more baby boys than girls born every year. Students have no trouble seeing that each successive year like that halves the probability of chance occurrence, resulting in a chance probability of $1/2^{82}$, smaller than 10^{-24}. Eisenhart and Alan Birnbaum (1967) exhumed this early sign test. $1/2^{82}$ is also the probability that the Minimum of a random sample of 82 items from a quantitative (continuous) population exceeds the population median (because it means that all 82 items do), and so you have a lower confidence limit for the population median with confidence probability $1 - 1/2^{82}$.

The 24 zeros, or nines, ask us: Would 81 years of excess male births out of 82 have been enough evidence to establish Divine Providence? In other words, is not the second lowest value out of a sample of 82 safe enough as a lower limit for a population median? The answer depends upon Pr(No exceptions) + Pr(1 exception), (or Pr(82) + Pr(81)).

If just one year out of 82 shows an excess of baby girls, this means either the last year shows more baby girls and none of the first 81 did, or the last showed more boys and so did all but one of the first 81, so that

$$\mathrm{Pr}_{82}(\text{All but 1}) = .5 \cdot \mathrm{Pr}_{81}(\text{All 81}) + .5 \cdot \mathrm{Pr}_{81}(\text{All but 1}) \quad \text{(based on chance)}$$

or

$$q \cdot \mathrm{Pr}_{81}(\text{All 81}) + p \cdot \mathrm{Pr}_{81}(\text{All but 1}) \quad \text{(in general)}.$$

The result: Pr(No exception) + Pr(1 exception) still begins with 22 zeros, and the corresponding confidence probability begins with 22 nines, calling for the calculation of another term:

$$\mathrm{Pr}_{82}(\text{2 exceptions}) = q \cdot \mathrm{Pr}_{81}(\text{1 exception}) + p \cdot \mathrm{Pr}_{81}(\text{2 exceptions}),$$

and so on until an accumulated tail probability rises into the realm of possible chance occurrence in the students' judgement. Or switch to an example with smaller n before going on to the general formulation (1a).

In an experimental Math 111 at the University of Maryland in 1974, we gave the students an A-table of Pr calculated by (1a) and a B-table of tails obtained by accumulation, each for $p = .5$ and $p = .2$ (and with some gaps left to fill in). A student, Damon Sui, came in a little confused because he thought the recursion formula applied to the Tail probabilities. So we checked a few examples to straighten him out—and saw that he was right, and then why: (1b) of course follows from (1a) by the distributive law of arithmetic, and also directly from first principles by the same reasoning as above with inequalities in place of some "equals." The course text was simplified accordingly. (Nemenyi and Harry Bushar (1975), *Motivated Math*, privately printed for the course.)

When two-sided confidence intervals for a median are introduced, the table of confidence probabilities is generated by the same averaging formula (starting with 1, .5, 0 for $n = 2$) because

$$\begin{aligned}\mathrm{Conf}_n(r) &= \Pr((x_{(r)}, x^{(r)}) \text{ of the first } n-1 \text{ sample values already brackets the median}) \\ &\quad + \Pr(\text{it misses by one and the } n\text{th value falls on the deficient side}) \\ &= \mathrm{Conf}_{n-1}(r) + 0.5((\mathrm{Conf}_{n-1}(r-1) - \mathrm{Conf}_{n-1}(r)) \\ &= (\mathrm{Conf}_{n-1}(r-1) + \mathrm{Conf}_{n-1}(r))/2. \end{aligned} \tag{1d}$$

The recursion formula (2) can be obtained from (1a) by a few short inductive steps, and then used to derive the *law of large numbers* directly, without having to learn means, variances, or Tschebychev first. The *normal approximation* can also be derived from (2) by showing that the ratio of successive probabilities becomes more and more indistinguishable from the ratio of the corresponding normal approximations as n gets big. For students with some calculus this can be done more formally, showing that the limit of (2), centered, is the differential equation of a normal density function.

The closed formula (3) also follows from (2), of course. Really, there is no need for the big formula at all, but it is a good idea to show it to the students at some point, if only because they will see it all over the literature in later life.

Peter Nemenyi
8211/2 Burch Avenue
Durham, NC 27701

REFERENCES

Arbuthnot, J. (1710), "An Argument for Divine Providence, Taken from the Constant Regularity in the Births of Both Sexes," *Philosophical Transactions of the Royal Society of London*, 186–190.

Eisenhart, C., and Birnbaum, A. (1967), "Anniversaries in 1966–67 of Interest to Statisticians," *The American Statistician*, 22, 22–29.

CORRECTION

JEFFREY F. BROMAGHIN (1993), "SAMPLE SIZE DETERMINATION FOR INTERVAL ESTIMATION OF MULTINOMIAL PROBABILITIES," *THE AMERICAN STATISTICIAN*, 47, 203–206

The quantity α_i in Equation (3) should be replaced by $1 - \alpha_i$. In the expression of π_i^+ in Equation (4), the $-$ following $\hat{\pi}_i$ should be replaced by a +. In the numerator of the expression of π_i^+ in Equation (12), the $-$ following $2n_i$ should be replaced by a +.

Jeffrey Bromaghin expresses his appreciation to G. Li and R. C. Tiwari and bringing two of these errors to his attention.