Probability and Statistics Notes

Note 2

Confidence and Reliability in an Infinite Population

by
Chris Ashley

7 October 1971

Air Force Weapons Laboratory

Abstract

An approach to the problem of quantifying confidence in conclusions drawn from results of experiments is presented, assuming an infinite number of elements in the set from which the samples are being taken and assuming dichotomy has been imposed. Three different kinds of experimental problems are presented as examples of circumstances to which the approach is applicable. Confidence in reliability of a missile guidance computer is treated in one of these problems. Three general and four special cases are analyzed as computational examples.

Acknowledgements

I am grateful to Jim Flagstead for giving me the problem which led to this note; to him and to Carl Baum for checking the final draft; to Berry Huelsman and Bill Rapp for their ideas for simplifying some of the analysis in one of the examples; and to Elaine Dekker, without whose typing patience and perseverance this note might never have attained a legible form.

PREFACE

Experiments are conducted in order to suggest and justify conclusions and predictions about the subject experimented upon. The greater the amount of supporting experimental data gathered, the higher the confidence one may reasonably have in conclusions drawn from that data. Thus, confidence in a conclusion is a monotone increasing function of the amount of experimental evidence supporting the conclusion. It is possible to quantify this confidence precisely.

In this paper it is shown by reasoning directly from basic principles how to quantify confidence in the case where the population has been dichotomized by the experimenter. No further assumptions are made regarding the distribution of the population.

The analysis begins with intuitively acceptable statements about confidence and proceeds without recourse to devices, such as Bayes' rule, use of which would be second nature for the statistician but which are foreign to many engineers. It is hoped that this avoidance of specialized terminology will make quantification of confidence more accessible and acceptable to the conscientious worker who lacks familiarity with the concepts of probability and statistics.

Some Notes on Confidence and Reliability in an Infinite Population.

1. Consider three experimenters, each with a problem.

2. (Experiment 1.) Suppose the first experimenter is attempting to discover the distribution or pattern with which a collimated stream of protons is deflected after colliding with a target. (He may be performing this experiment either in a laboratory or as a Monte Carlo computer calculation.) As part of his effort to do this he monitors a particular region of solid angle about the target to see what fraction of the protons is deflected in that direction. Suppose that after one hour he has observed that L protons have been fired at the target and, of these, M have been deflected into the region of solid angle which he is presently studying. The experimenter might then reasonably guess that the probability is about M/L that the next proton fired will be deflected into the instrumented region. Suppose the probability is actually some real number P; then, in view of the data, with what confidence may the experimenter assert that P is between, say, (M/L)/2 and (1 + M/L)/2 ?

3. (Experiment 2.) Suppose the second experimenter is attempting to discover the value of a physical constant. He has made L measurements of the constant and, of these, M have fallen inside a real interval (r_1,r_2). That is, the experimenter has found a fraction M/L of his measurements to fall within the interval (r_1,r_2). At this point the experimenter may do any one of several things.

4. He may guess that he knows the distribution of numbers which his particular measuring apparatus will yield, and from his set of measurements make estimates of the parameters of that distribution. Typically he will decree that measurements from his apparatus will be normally

distributed (or at least that their distribution will have moments -- which it won't if it is, say, the Cauchy distribution -- so he can invoke the Central Limit Theorem and have their sample mean becoming normally distributed with increasing sample size). So he will calculate the finite mean and finite variance of his set of measurements. Using these as the mean and variance he would have if he made an infinite number of measurements, he will then announce to the world that the actual value of the physical constant lies inside some interval (r_3, r_4) with "3σ probability" (99.865%) (or 1σ, or 2σ, or whatever). If the experimenter is very conscientious he may even run a goodness-of-fit test to see if his finite set of measurements really could have come from a normal population. Unfortunately, even such a conscientious experimenter sometimes omits to advertise the level of significance of his fit test. Also, he sometimes does not tell the users of his experimental results what other distributions would fit the data as well.

5. Alternatively, the experimenter may admit that he does not really know precisely how his apparatus distributes measurements. He does know that he can partition all possible measurements into two classes, those within the interval (r_1, r_2) and those outside that interval. He knows that, whatever the distribution of the measurements, there exists some probability P that a single measurement will fall within the interval. His best approximation of P, in view of his L measurements so far, is M/L. With what confidence can the experimenter then declare that P lies within the interval (p_1, p_2), where p_1 and p_2 are bounds picked as interesting by the experimenter?

6. (Experiment 3.) Consider the third experimenter. Suppose that a missile guidance computer is being subjected to electromagnetic or particle radiation in a simulator. Suppose that sometimes a pulse of radiation causes the computer to malfunction non-destructively but

that at other times a pulse is not followed by a malfunction. The experimenter has observed that in L pulses no malfunctions have occurred on M occasions. He might reasonably guess that the probability is about M/L that the next pulse (which might occur when the missile is in use) will not cause the computer to malfunction. Suppose the probability of no malfunction is actually some real number P; then, in view of the data, with what confidence may the experimenter assert that P is, say, greater than .5?

7. The reason for discussing these three experiments is to exhibit a small sample of the wide range of situations in which the theory of this note is applicable. They have in common that the experimenters do not really know, even after the experiment, how the experimental results are distributed. Let f denote the unknown probability density function which gives the distribution of measurements of the variable being measured. Then $\int_{-\infty}^{b} f(x)\,dx$ is the probability that a single measurement will be less than or equal to b, and $\int_{a}^{b} f(x)\,dx$ is the probability that a single measurement will fall between a and b. If the measurement technique does not contain any systematic error, the distribution mean

$$\mu = \int_{-\infty}^{\infty} x\, f(x)\, dx$$

(if it exists, i.e., if $x*f(x)$ is integrable from $-\infty$ to ∞) will be both the expected value of a measurement and the "true" value of the variable being measured.

8. In the first of the three experiments at the beginning of this paper, then, the unknown probability P would be

$$P = \int_{\substack{\text{over the} \\ \text{solid angle} \\ \text{of interest}}} f(A)\, dA \quad .$$

In the second experiment,

$$P = \int_{r_1}^{r_2} f(x)\, dx \quad .$$

In the third experiment,

$$P = \int_{.5}^{1} f(x)\, dx \quad .$$

In this paper we shall assume (as we did in an earlier note* on this subject: "Some Notes on Confidence and Reliability in a Finite Population", dated 18 February 1971) that the distribution of f is unknown and that the only information the evaluator of the data can get about it is from a finite number L of measurements.

9. (Assumptions.) We shall also assume the evaluator dichotomizes the experimental results, so that each is either inside or outside the region of interest to him. Let M be the number which fall inside the region or interval of interest; then L-M is the number outside the area of interest. Since L is not bounded by any finite number N for the kind of experiments being discussed in this paper, i.e., since the experimenter can always take more measurements and thereby gain more information (unlike the situation in the earlier note, in which it was possible for the experimenter to get all the information possible,

* Probability and Statistics Notes, Note 1.

and thus achieve certainty, in only N, say 1000, experiments), therefore P may be any real number between 0 and 1 (instead of being restricted to proper rational numbers with denominator N).

10. (Confidence requirement.) We shall also define confidence so that it satisfies the following requirement (as in the earlier note). We want to discover if our results would be more likely if $P = p_1$ than if $P = p_2$. Let $g(p,L,M)$ be the probability of getting exactly M measurement values inside the interval of interest in L samples if the infinite population, or set of possible measurement values, actually has a fraction p in that interval. Then if $g(p_1,L,M) = 2g(p_2,L,M)$, we would be twice as confident that $P = p_1$ as we would that $P = p_2$. (Read that sentence again.) That is, our confidence in p_1 is twice our confidence in p_2. We write $C(p_1,L,M) = 2C(p_2,L,M)$. One motivation for levying this requirement is the fact that $g(p_1,L,M) = 2g(p_2,L,M)$ means there are "twice as many ways" we could have obtained our results from a population for which $P = p_1$ as from a population for which $P = p_2$. So if we imagine that we are getting our results from the union of two populations, one of each kind, then before experimenting it would be twice as likely that from the union we will get a set of experimental results of the kind we have (L total, M interesting) from the p_1 population as from the p_2 population. Looked at another way, if both of these kinds of populations will occur equally often in our future experience (the maximum ignorance assumption), then in a large number of sets drawn from randomly selected populations of these two kinds in the future we will obtain sets of this kind (L total, M interesting) twice as often from p_1 populations as from p_2 populations. Hence, if we bet strictly according to confidence defined subject to this requirement, then over bets placed on many randomly encountered populations we will "break even". In general, we require of our definition of

confidence that $P = p$, denoted by $C(p,L,M)$, that, given experimental results L and M, then for any real number r

$$g(p_1,L,M) = r * g(p_2,L,M) \Rightarrow C(p_1,L,M) = r * C(p_2,L,M).$$

We might rewrite this expression more generally to cover the case in which we want to quantify our confidence that P is contained either within interval (p_1,p_2) or within interval (p_3,p_4). Then the reasoning at the beginning of this paragraph would require that

$$\int_{p_1}^{p_2} g(p,L,M)\, dp = r * \int_{p_3}^{p_4} g(p,L,M)\, dp \Rightarrow$$

$$\Rightarrow C(p_1,p_2,L,M) = r * C(p_3,p_4,L,M) \qquad , (1)$$

where $C(p_i,p_j,L,M)$ denotes our confidence that P is within the interval (p_i,p_j) given L samples among which exactly M fall within the region of interest. (The link between the notation in which C has three arguments and that in which it has four is provided by $C(p,L,M) =$
$= C(p,p,L,M)$.)

11. (Confidence axioms.) Also, we would like to feel that confidence is similar to probability. That is, given the experimental results L and M, we would like to be able to say that the measurement apparatus and the subject being experimented upon are "probably" of such and such character (although, since they are given, the probability that they actually are is either 0 or 1). So it is natural to require confidence to satisy the three basic axioms of probability theory. Applied to disjoint intervals (p_1,p_2), (p_3,p_4),..., these axioms, for confidence, are:

I. $0 \le C(p_1,p_2,L,M)$ for $p_1 \le p_2$

II. $C(0,1,L,M) = 1$

III. $C((p_1,p_2) \text{ or } (p_3,p_4) \text{ or } \ldots,L,M) =$

$$= C(p_1,p_2,L,M) + C(p_3,p_4,L,M) + \ldots \quad .$$

Axiom I states that confidence may be zero, but never negative. Axiom II states that confidence is 1 that the unknown probability P is <u>some</u> real number between 0 and 1; i.e., we are completely certain that there is such a probability P. Axiom III states that the confidence that P is contained in one or another of several nonoverlapping intervals is just the sum of the confidences that P is in each one alone.

12. (Existence.) Expression (1) tempts us to try defining $C(p,L,M) = g(p,L,M)$ for fixed experimental results L and M. However, the fact that g gives the probability of getting M values inside the interval of interest given L and p, so that g is a discrete probability density function, assures us that $\sum_{M=0}^{L} g(p,L,M) = 1$ but not that $\int_0^1 g(p,L,M)\, dp = 1$. Consequently such a trial definition would leave Axiom II unsatisfied. If instead we define confidence that P is contained in the interval (p_1,p_2) by

$$C(p_1,p_2,L,M) = \frac{\int_{p_1}^{p_2} g(p,L,M)\, dp}{\int_0^1 g(p,L,M)\, dp} \quad , (2)$$

then all the requirements of the foregoing paragraphs are satisfied. That g is Riemann integrable from 0 to 1 will be clear when we derive its form in paragraph 14, below. Plugging this definition of C into expression (1) and using the linearity of the Riemann integral shows immediately that expression (1) is satisfied. The fact that g is itself a probability density function, and so satisfies the probability axioms, and so is non-negative, is sufficient to assure that Axiom I is satisfied. That Axiom II is satisfied is immediate. Axiom III is the natural definition of our confidence that P is in either (p_1,p_2) or (p_3,p_4) or ....

13. (Uniqueness.) A question which arises naturally at this point is, does equation (2) define confidence uniquely, i.e., is the function C as defined by equation (2) the only function which satisfies expression (1) and the three axioms listed in paragraph 11, within our assumptions, above? To answer this, we begin by rewriting expression (1) as:

$$\frac{\int_{p_1}^{p_2} g(p,L,M)\,dp}{\int_{p_3}^{p_4} g(p,L,M)\,dp} = r = \frac{C(p_1,p_2,L,M)}{C(p_3,p_4,L,M)} .$$

Next, let $\overline{C}$ be any function which satisfies all our four requirements for confidence, viz., expression (1) and the three axioms. Then by Axiom II we know that $\overline{C}(0,1,L,M) = 1$. For general p_1 and p_2 we may therefore write:

$$\overline{C}(p_1,p_2,L,M) = \frac{\overline{C}(p_1,p_2,L,M)}{\overline{C}(0,1,L,M)} =$$

$$= \frac{\int_{p_1}^{p_2} g(p,L,M)\, dp}{\int_0^1 g(p,L,M)\, dp} =$$

$$= \frac{C(p_1,p_2,L,M)}{C(0,1,L,M)} =$$

$$= C(p_1,p_2,L,M) \quad .$$

So we know that $\overline{C}$ is the same as C for all permissible sets of values of the parameters. Therefore $\overline{C}$ is the same function as C. Hence equation (2) gives a definition of confidence which satisifies our requirements uniquely.

14. (g(p,L,M).) What is the form of g? The answer to this question can be had directly from the verbal definition of g in paragraph 10, above. If the subject being studied and the experimental apparatus are such that the probability of getting a measurement value within the interval of interest is p, then the probability of getting M independently taken measurement values within this interval is just $p*p*...*p = p^M$. Similarly if the probability of getting a measurement value outside the interval of interest is 1-p, then the probability of getting L-M independently taken measurement values outside this interval is just $(1-p)(1-p)...(1-p) = (1-p)^{L-M}$. So the probability of getting M measurement values inside the interval of interest and L-M values outside that interval, for a total of L values, is just

$$g(p,L,M) = \binom{L}{M} p^M (1-p)^{L-M} \qquad (3)$$

where the binomial coefficient $\binom{L}{M}$ takes account of the fact that the M interesting measurements could have shown up in the total of L measurements in any of $\binom{L}{M}$ different sequences. (Equation (3) is the defining equation for the binomial distribution. The binomial distribution can be shown to become similar to the hypergeometric distribution as the size of a finite population becomes large, so it can be anticipated that under that condition some of the formulas developed in this note may in practice be used in place of corresponding formulas in the earlier note on confidence in finite populations, cited above. This fact arises again in paragraphs 24 and 25, below.)

15. (Confidence for infinite populations.) Plugging (3) into (2) yields

$$C(p_1,p_2,L,M) = \frac{\int_{p_1}^{p_2} p^M (1-p)^{L-M}\, dp}{\int_0^1 p^M (1-p)^{L-M}\, dp} \quad . \quad (4)$$

(This expression for infinite populations is analogous to equation (4) in the earlier note, cited above, for finite populations.)

16. (Symmetry Theorem.) From this general expression for confidence in an infinite population we can immediately establish two other facts which can sometimes be of aid. Bear in mind that $C(x_1,x_2,x_3,x_4)$ is our symbol for confidence that the actual probability P is contained in the interval (x_1,x_2) given that x_3 samples have been taken and, of these, x_4 were discovered to be in the region of interest. Using these four parameters, in that order, we then prove the

<u>Symmetry Theorem</u>: $C(0,Q,L,M) = C(1-Q,1,L,L-M)$. (5)

Proof: $$C(0,Q,L,M) = \frac{\int_0^Q p^L (1-p)^{L-M} \, dp}{\int_0^1 p^L (1-p)^{L-M} \, dp}$$

In both these integrals let us change variables by letting $q = 1-p$. Then $p = 1-q$; $dp = -dq$; $p = Q \Rightarrow q = 1-p = 1-Q$; $p = 1 \Rightarrow q = 1-p = 1-1 = 0$; and $p = 0 \Rightarrow q = 1-p = 1-0 = 1$. Then

$$C(0,Q,L,M) = \frac{\int_1^{1-Q} (1-q)^L q^{L-M} (-dq)}{\int_1^0 (1-q)^L q^{L-M} (-dq)} =$$

$$= \frac{\int_{1-Q}^1 (1-q)^L q^{L-M} \, dq}{\int_0^1 (1-q)^L q^{L-M} \, dq} =$$

$$= \frac{\int_{1-Q}^1 (1-q)^{L-(L-M)} q^{L-M} \, dq}{\int_0^1 (1-q)^{L-(L-M)} q^{L-M} \, dq} =$$

$$= C(1-Q,Q,L,L-M)$$, q.e.d.

17. (Symmetry Corollary.) From this Symmetry Theorem we have immediately the

<u>Symmetry Corollary</u>: $C(0,\frac{1}{2},L,\frac{L}{2}) = \frac{1}{2}$. (6)

Proof: By the Symmetry Theorem,

$$C(0,\tfrac{1}{2},L,\tfrac{L}{2}) = C(1-\tfrac{1}{2},1,L,L-\tfrac{L}{2}) = C(\tfrac{1}{2},1,L,\tfrac{L}{2}) \quad . \qquad (7)$$

But, by Axioms II and III,

$$C(0,x_2,x_3,x_4) + C(x_2,1,x_3,x_4) = 1 \quad . \qquad (8)$$

Therefore

$$C(0,\tfrac{1}{2},L,\tfrac{L}{2}) + C(\tfrac{1}{2},1,L,\tfrac{L}{2}) = 1 \quad .$$

Combining this with equation (7) we have

$$2\ C(0,\tfrac{1}{2},L,\tfrac{L}{2}) = 1 \Rightarrow C(0,\tfrac{1}{2},L,\tfrac{L}{2}) = \tfrac{1}{2} \quad , \qquad \text{q.e.d.}$$

18. (L = L and M = 0.) Let us now apply these results to three general and four special cases of recurrent interest. For the first general example consider the case in which the samples come out all of a kind, and since none stand out we'll say none are of interest. Then L = L and M = 0 and we have:

$$C(0,Q,L,0) = \frac{\int_0^Q p^0\ (1-p)^{L-0}\ dp}{\int_0^1 p^0\ (1-p)^{L-0}\ dp} = \frac{\int_0^Q (1-p)^L\ dp}{\int_0^1 (1-p)^L\ dp} =$$

$$= \frac{\int_{1}^{1-Q} q^{L} \, (-dq)}{\int_{1}^{0} q^{L} \, (-dq)} = \frac{\left. \frac{q^{L+1}}{L+1} \right|_{1-Q}^{1}}{\left. \frac{q^{L+1}}{L+1} \right|_{0}^{1}} =$$

$$= \frac{\frac{1}{L+1} - \frac{(1-Q)^{L+1}}{L+1}}{\frac{1}{L+1}} =$$

$$= 1 - (1-Q)^{L+1} \quad . \qquad (9)$$

This is, then, the confidence we should have that in the long run the fraction of samples which we will get which will be interesting is less than or equal to Q.

19. (L = L and M = L.) Application of the Symmetry Theorem to this result permits us to write immediately, as our second general example,

$$C(1-F,1,L,L) = 1 - (1-F)^{L+1} \quad .$$

This example is especially interesting from the viewpoint of the reliability tester. For if F is the maximum tolerable failure ratio, or perhaps the greatest tolerable chance of computer malfunction as a result of pulsing being administered by the third experimenter (cf. paragraph 6, above), and if all tests so far have resulted in successes (non-failures), then this expression gives the confidence one should have that the "true" success rate is 1-F or better, i.e., that the true failure rate is F or less. If we let R, for reliability, be 1-F, then the equation of interest when test results have been uniformly successful is therefore

$$C(R,1,L,L) = 1 - R^{L+1} \quad . \tag{10}$$

(This important fact may of course be deduced directly from equation (4).)

20. (L = 1 and M = 0.) A special case of equation (9) arises when it is possible to take only one sample, perhaps because occurrences which qualify as samples are difficult to come by, and that sample turns out not to lie in the region of interest. Then L = 1 and M = 0 and equation (9) reduces to

$$\begin{aligned} C(0,Q,1,0) &= 1 - (1-Q)^{1+1} = 1 - (1-Q)^2 = \\ &= 1 - (1-2Q + Q^2) = \\ &= 2Q - Q^2 \quad . \end{aligned} \tag{11}$$

21. (L = 1 and M = 1.) A second special case of interest which comes to mind immediately is that in which only one sample can be taken and that sample turns out to be of interest. Applying the Symmetry Theorem to equation (11) yields at once that $C(1-Q,1,1,1) = 2Q - Q^2$. By equation (8) we have that $C(0,1-Q,1,1) = 1 - C(1-Q,1,1,1)$. Changing variables by letting $T = 1-Q$, therefore, we obtain

$$\begin{aligned} C(0,T,1,1) &= 1 - (2Q - Q^2) = \\ &= 1 - 2Q + Q^2 = \\ &= (1 - Q)^2 = \\ &= T^2 \quad . \end{aligned}$$

As with equation (10), this result may of course be had immediately from equation (4):

$$C(0,Q,1,1) = \frac{\int_0^Q p^1 \, (1-p)^{1-0} \, dp}{\int_0^1 p^1 \, (1-p)^{1-0} \, dp} =$$

$$= \frac{\int_0^Q p \, dp}{\int_0^1 p \, dp} = \frac{\frac{Q^2}{2}}{\frac{1^2}{2}} =$$

$$= Q^2$$

This simple result is worth bearing in mind for those frequent occasions in everyday life when one is tempted "to generalize from one data point". One can draw conclusions in such instances, and such results as this enable one to assess how sure he should be of such conclusions. For example, perhaps you are thinking of moving to a new town, but someone has just told you there was signficant earthquake activity in that area last year. If this is all the information you have on the occurrence of earthquakes in that area, then you may still draw some quantitative conclusions about the percentage of years in which significant earthquake activity occurs in that area over a period of time long enough to get a reasonable average but short enough so basic geological conditions relevant to earthquakes in that area do not change appreciably (e.g., so the town does not get replaced by an ocean). In particular, you can conclude on the basis of this fact about one year that the fraction of years in which significant earthquake activity occurs in the area is less than or equal to Q (say $\frac{1}{2}$) with confidence Q^2 (25% if $Q = \frac{1}{2}$). Thus you can be 25% confident that the probability is not more than 50% that one or more significant earthquakes will occur in the area of the town next year (while you are there!). (Note that we assumed you had only the one fact at your disposal, and therefore that you were ignorant of ideas of earthquake periodicity, relaxation's preventing immediate recurrence, etc.)

22. (L = L and M = L-1.) For the third general example consider again the case in which this theory is being applied to reliability testing. Usually in reliability testing one expects, at least at the beginning of the testing, that "most" of the samples will pass the test, or, in the case of our third experimenter, that the computer will not malfunction on most of the pulses. Perhaps failures occur just often enough so we know they are possible, perhaps just once in the test series. Then if we let the successes be the occurrences of interest, i.e., M = L-1, we may calculate

$$C(R,1,L,L-1) = \frac{\int_R^1 p^{L-1}\,(1-p)^{L-(L-1)}\,dp}{\int_0^1 p^{L-1}\,(1-p)^{L-(L-1)}\,dp} = \frac{\int_R^1 p^{L-1}\,(1-p)\,dp}{\int_0^1 p^{L-1}\,(1-p)\,dp} =$$

$$= \frac{\int_R^1 (p^{L-1}-p^L)\,dp}{\int_0^1 (p^{L-1}-p^L)\,dp} = \frac{\left(\frac{p^L}{L} - \frac{p^{L+1}}{L+1}\right)\Big|_R^1}{\left(\frac{p^L}{L} - \frac{p^{L+1}}{L+1}\right)\Big|_0^1} =$$

$$= \frac{\frac{1-R^L}{L} - \frac{1-R^{L+1}}{L+1}}{\frac{1}{L} - \frac{1}{L+1}} = \frac{(L+1)(1-R^L)-L(1-R^{L+1})}{L+1-L} =$$

$$= (L+1)(1-R^L) - L(1-R^{L+1}) = \qquad (12)$$

$$= LR^{L+1} - (L+1)R^L + 1 \quad . \qquad (13)$$

This gives our confidence that the actual fraction of successes in an infinitely large random sample would be greater than or equal to R. Looked at another way, this is the confidence we should have that the probability is R or better that our next independent random sample will be a success (i.e., non-malfunction, non-failure, within the region of interest). Thus, this is the third experimenter's confidence in the reliability R of his computer after one malfunction in L tests, where by reliability we mean simply the minimum actual success ratio. (Cf. paragraph 19, above, for another case of interest in reliability testing.)

23. (L = 2 and M = 1.) A special case of this general example (our third special case so far) arises when two samples have been taken and the "ambiguous" results have been observed that "one sample has gone each way". Then L = 2 and M = 1. Applying the Symmetry Theorem to equation (12), and letting Q = 1-R, so that R = 1-Q, we have

$$C(0,Q,L,1) = (L+1)Q^L - LQ^{L+1} \quad .$$

Therefore

$$C(0,Q,2,1) = (2+1)Q^2 - 2Q^{2+1} =$$

$$= 3Q^2 - 2Q^3 \quad .$$

(As usual, of course, this result could have been had directly from equation (4).) In particular, for $Q = \frac{1}{2}$ we would have $C(0,\frac{1}{2},2,1) =$ $= 3(\frac{1}{2})^2 - 2(\frac{1}{2})^3 = \frac{3}{4} - \frac{2}{8} = \frac{1}{2}$. This is the result we would expect, either from intuition or from the Symmetry Corollary (cf. equation (6)).

24. (L = 5 and M = 4.) To give a somewhat more realistic specific numerical example (our fourth special case), let us calculate the confidence he would have if the third experimenter had performed five tests in which he had observed a malfunction only one time, and if he were interested in a reliability of 65%. Then he would have L = 5, M = 4, and R = .65. Applying equation (13),

$$
\begin{aligned}
C(.65,1,5,4) &= 5(.65)^{5+1} - (5+1).65^5 + 1. = \\
&= 5(.65)^6 - 6(.65)^5 + 1. = \\
&= 5(.07541889062) - 6(.1160290625) + 1. = \\
&= .3770944531 - .6961743750 + 1. = \\
&= 1. - .3190799219 = \\
&= .6809200781 \quad . \qquad (14)
\end{aligned}
$$

That is, after one malfunction in five tests the experimenter would be justified in having a confidence of 68.092% in computer reliability of 65%. (The reader may be interested in comparing this result with that of a very similar example given in paragraph 9 of the earlier note, cited above, on confidence in finite sample spaces.)

25. Numerical integration digital computer programs to calculate

$C(p_1,p_2,L,M)$ directly from equation (4) may easily be written for arbitrary values of the parameters p_1, p_2, L, and M. More importantly, the reader should note that equation (4) may be reduced to such forms as equation (13) "on the back of an envelope" and then, for modest values of L, say 10 or 15 or so, may be evaluated in a matter of moments with the aid of a standard desk top electronic calculator. This is how evaluation (14) was done. Thus if the population size is very large, even though not infinite, it may be worth while to use equation (4) in this note as an approximation of equation (4) in the earlier note, cited above, simply because the calculation may then be done with tools which are more readily available.