
Statistical Confidence Levels for Estimating Error Probability

Note: This article has been previously published in Lightwave Magazine (April 2000), and in the Maxim Engineering Journal (volume 37).



Statistical Confidence Levels for Estimating Error Probability

I. The Application: Estimation of Bit Error Probability

Many components used in digital communication systems, such as the MAX3675 and MAX3875, are required to meet minimum specifications for probability of bit error, $P(\epsilon)$. $P(\epsilon)$ can be estimated by comparing the output bit pattern of a system with a pre-defined bit pattern applied to the input. Any discrepancies between the input and output bit streams are flagged as errors. The ratio of the number of detected bit errors, ϵ , to the total number of bits transmitted, n , is $P'(\epsilon)$, where the prime (') character signifies that $P'(\epsilon)$ is an estimate of the true $P(\epsilon)$. The quality of the estimate improves as the total number of transmitted bits increases. This can be expressed mathematically as:

$$P'(\epsilon) = \frac{\epsilon}{n} \xrightarrow{n \rightarrow \infty} P(\epsilon) \quad (1)$$

It is important to transmit enough bits through the system to ensure that $P'(\epsilon)$ is an accurate reflection of the true $P(\epsilon)$ (that would be obtained if the test were allowed to proceed for an infinite time period). In the interest of limiting testing to a reasonable length of time, however, it is important to know the minimum number of bits for a statistically valid test.

In many cases, we only need to verify that the $P(\epsilon)$ is at least as good as some pre-defined standard. In other words, it is sufficient to prove that the $P(\epsilon)$ is less than some upper limit. For example, many telecommunication systems require a $P(\epsilon)$ of 10^{-10} or better (an upper limit of 10^{-10}). The statistical idea of associating a *confidence level* with an upper limit can be used to postulate, with quantifiable confidence, that the true $P(\epsilon)$ is less than the specified limit. The primary advantage of this method is that it provides a method to trade-off test time versus measurement accuracy.

II. Definition and Interpretation of Statistical Confidence Level

Statistical confidence level is defined as the probability, based on a set of measurements, that the actual probability of an event is better than some specified level. (For purposes of this definition, actual probability means the probability that would be measured in the limit as the number of trials tends toward infinity.) When applied to $P(\epsilon)$ estimation, the definition of statistical confidence level can be restated as: the probability, based on ϵ detected errors out of n bits transmitted, that the actual $P(\epsilon)$ is better than a specified level, γ (such as 10^{-10}). Mathematically, this can be expressed as

$$CL = P [P(\epsilon) < \gamma | \epsilon, n] \quad (2)$$

where $P[]$ indicates probability and CL is the confidence level. Since confidence level is, by definition, a probability, the range of possible values is 0 – 100%.

Once the confidence level has been computed we may say that we have CL percent confidence that the $P(\epsilon)$ is better than γ . Another interpretation is that, if we were to repeat the bit error test many times and re-compute $P'(\epsilon) = \epsilon/n$ for each test period, we would expect $P'(\epsilon)$ to be better than γ for CL percent of the measurements.

III. Confidence Level Calculation

A. The Binomial Distribution Function

Calculation of confidence levels is based on the binomial distribution function, the details of which are included in many statistics texts^{1,2}. The binomial distribution function is generally written as

$$P_n(k) = \binom{n}{k} p^k q^{n-k}, \quad \text{where } \binom{n}{k} \text{ is defined as } \frac{n!}{k!(n-k)!}. \quad (3)$$

Equation (3) gives the probability that k events (i.e., bit errors) occur in n trials (i.e., n bits transmitted), where p is the probability of event occurrence (i.e., a bit error) in a single trial and q is the probability that the event does not occur (i.e., no bit error) in a single trial. Note that the binomial distribution models events that have two possible outcomes, such as success/failure, heads/tails, error/no error, etc., and therefore $p + q = 1$.

When we are interested in the probability that N or fewer events occur (or, conversely, greater than N events occur) in n trials, then the cumulative binomial distribution function (Equation 4) is useful:

$$P(\epsilon \leq N) = \sum_{k=0}^N P_n(k) = \sum_{k=0}^N \left(\frac{n!}{k!(n-k)!} \right) p^k q^{n-k}$$
$$P(\epsilon > N) = 1 - P(\epsilon \leq N) = \sum_{k=N+1}^n \left(\frac{n!}{k!(n-k)!} \right) p^k q^{n-k} \quad (4)$$

Graphical representations of Equations (3) and (4), along with some of their properties, are summarized in Figure 1.

B. Application of the Binomial Distribution Function to Confidence Level Calculation

In a typical confidence level measurement, we start by hypothesizing a value for p (the probability of bit error in the transmission of a single bit) and we choose a satisfactory level of confidence. We will use p_h to represent our chosen hypothetical value of p . Generally we choose these values based on an imposed specification limit (e.g., if the specification is $P(\epsilon) \leq 10^{-10}$, we choose $p_h = 10^{-10}$ and choose a confidence level of, say 99%). We can then use Equation (4) to determine the probability, $P(\epsilon > N | p_h)$, that more than N bit errors will occur when n total bits are transmitted, based on p_h . If, during actual testing, less than N bit errors occur (even though $P(\epsilon > N | p_h)$ is high) then there are two possible conclusions: (a) we just got lucky, or (b) the actual value of p is less than p_h . If we repeat the test over and over and continue to measure less than N bit errors, then we become more and more confident in conclusion (b). A measure of our level of confidence in conclusion (b) is defined as $P(\epsilon > N | p_h)$. This is because if $p_h = p$, we would have had a high probability of detecting more bit errors than N . When we measure less than N errors, we conclude that p is probably less than p_h , and we define the probability that our conclusion is correct as the confidence level. In other words, we are $CL\%$ confident that $P(\epsilon)$ (the actual probability of bit error) is less than p_h .

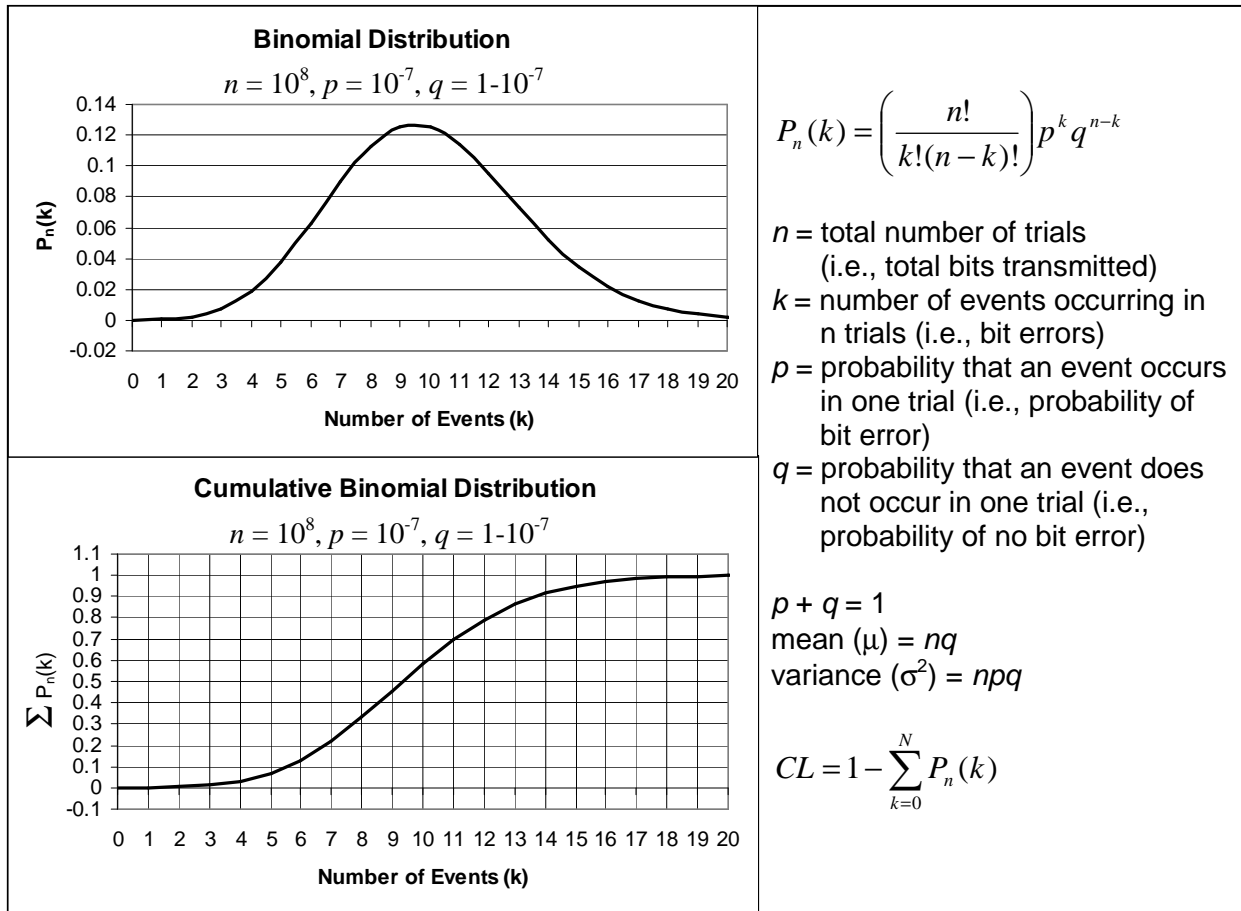


Figure 1. The binomial distribution function

In terms of the cumulative binomial distribution function, the confidence level is defined as

$$CL = P(\varepsilon > N | p_h) = 1 - \sum_{k=0}^N \left(\frac{n!}{k!(n-k)!} \right) (p_h)^k (1 - p_h)^{n-k} \quad (5)$$

where CL is the confidence level in terms of a percent.

As noted above, when using the confidence level method we generally choose a hypothetical value of p (p_h) along with a desired confidence level (CL) and then solve Equation (5) to determine how many total bits, n , we must transmit (with N or less errors) through the system in order to prove our hypothesis. Solving for n and N , can prove difficult unless some approximations are used.

If $np > 1$ (i.e., we transmit at least as many bits as the mathematical inverse of the bit error rate) and k has the same order of magnitude as np , then the Poisson theorem¹ (Equation 6) provides a conservative estimate of the binomial distribution function.

$$P_n(k) = \left(\frac{n!}{k!(n-k)!} \right) p^k q^{n-k} \xrightarrow{n \rightarrow \infty} \frac{(np)^k}{k!} e^{-np} \quad (6)$$

Equation (7) shows how we can use Equation (6) to obtain an approximation for the cumulative binomial distribution as well.

$$\sum_{k=0}^N P_n(k) \approx \sum_{k=0}^N \frac{(np)^k}{k!} e^{-np} \quad (7)$$

Now we can combine Equations 5 and 7, and then solve for n , as follows:

$$\sum_{k=0}^N P_n(k) = 1 - CL \quad (\text{by re-arranging Equation (5)})$$

$$\sum_{k=0}^N \frac{(np)^k}{k!} e^{-np} = 1 - CL \quad (\text{using Equation (7)})$$

$$-np = \ln \left[\frac{1 - CL}{\sum_{k=0}^N \frac{(np)^k}{k!}} \right]$$

$$n = -\frac{\ln(1 - CL)}{p} + \frac{\ln \left(\sum_{k=0}^N \frac{(np)^k}{k!} \right)}{p} \quad (8)$$

Note that the second term in Equation (8) is equal to zero for $N = 0$, and in this case the Equation is simple to solve. For $N > 0$, solutions to Equation (8) are more difficult, but they can be obtained empirically using a computer.

We are now ready to determine the total number of bits that must be transmitted through the system in order to achieve a desired confidence level. Following is an example of this procedure:

- (1) **Select p_n , the hypothetical value of p .** This is the probability of bit error that we would like to verify. For example, if we want to show that $P(\epsilon) \leq 10^{-10}$, then we would set p in Equation (8) equal to $p_h = 10^{-10}$.
- (2) **Select the desired confidence level.** Here we are forced to trade off confidence for test time. Choose the lowest reasonable confidence level for the application in order to minimize test time. The trade-off between test time and confidence level is proportional to $-\ln(1-CL)$. This is illustrated in Figure 2.
- (3) **Solve Equation (8) for n .** In most cases, this is simplified by assuming that no bit errors will occur during the test (i.e., $N=0$).
- (4) **Calculate the test time.** The time required to complete the test is n/R , where R is the data rate.

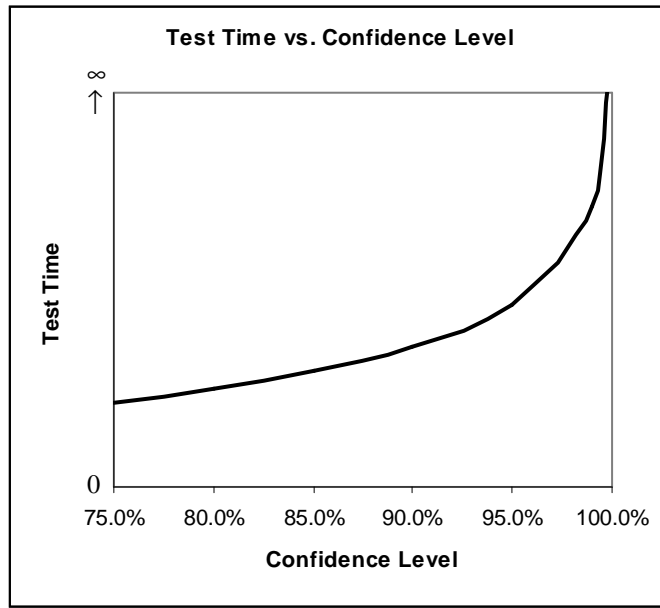


Figure 2. Test Time Versus Confidence Level

C. Example: Application of Confidence Level Concepts to Estimation of Bit Error Probability

Many telecommunication systems specify a $P(\epsilon)$ of 10^{-10} or better. For example, assume we are to test the MAX3675 (at 622Mbps) and MAX3875 (at 2.5Gbps) clock/data recovery chips to verify compliance with this specification. In this case we set $p_h = 10^{-10}$. We would like to design a test that would result in 100% confidence in the desired specification, but, since we do not have infinite test time, we will settle for a 99% confidence level. Next we solve Equation (8) for n using various values of N (e.g., 0,1,2,3,4). The results are shown in TABLE I below.

TABLE I
Example: Estimation of Bit Error Probability ($CL = 99\%$ and $p_h = 10^{-10}$)

Bit Errors $\leq N$ $N =$	Required Number of Bits to Transmit (n)	Test Time for Bit Rate of 622 Mbps (seconds)	Test Time for Bit Rate of 2.5 Gbps (seconds)
0	4.61×10^{10}	74.1	18.5
1	6.64×10^{10}	106	26.7
2	8.40×10^{10}	135	33.7
3	1.00×10^{11}	161	40.2
4	1.16×10^{11}	186	46.6

From TABLE I we see that if no bit errors are detected for 18.5 seconds (in a 2.5Gbps system) then we have a 99% confidence level that $P(\epsilon) \leq 10^{-10}$. If one bit error occurs in 26.7 seconds of testing, or two bit errors in 33.7 seconds, and so on, the result is the same (i.e., 99% confidence level that $P(\epsilon) \leq 10^{-10}$).

In order to develop a standard $P(\epsilon)$ test for the MAX3675 and MAX3875, we might select the test time corresponding to $N = 3$ from TABLE I. Using a bit error rate tester (BERT) we transmit 10^{11} bits through each of the two chips. The test time for 10^{11} bits is 2 min., 41 sec. at 622Mbps and 40.2 sec. at 2.5 Gbps. At the end of the test time, we check the number of detected bit errors, ϵ . If $\epsilon \leq 3$ the device has passed and we are 99% confident that $P(\epsilon) \leq 10^{-10}$.

IV. Stressing the System to Reduce Test Time

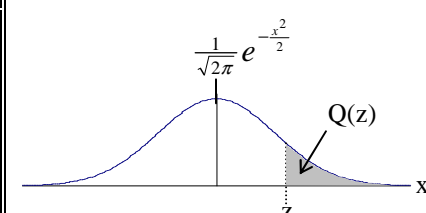
Dan Wolaver has documented a method for reducing test time by stressing the system³. It is based on the assumption that thermal (Gaussian) noise at the input of the receiver is the dominant cause of bit errors. (Note that this assumption excludes other potential error causing effects, such as jitter, etc.) For systems where this assumption is valid, the signal-to-noise ratio (SNR) can be reduced by a known quantity through inserting a fixed attenuation in the transmission path (i.e., the attenuation applies only to the signal and not the dominant noise source). In our previous example of the MAX3675 and MAX3875, it was determined that jitter effects and the non-linear gain of the input limiting amplifier violated the key assumption of this method, so it was not employed.

In systems where the assumption is valid, the probability of bit error can generally be mathematically calculated as^{4,5}

$$P(\epsilon) = Q\left(\frac{\sqrt{SNR_{electrical}}}{2}\right) = Q\left(\frac{SNR_{optical}}{2}\right) \quad (7)$$

where $Q(x)$ is the complementary error (or ‘Q’) function included in many communications textbooks⁶ (a variety of other sources for this data are available, including the NORMDIST function in MicrosoftTM Excel). Key values for the complementary error function are listed in TABLE II.

TABLE II
Tabulated Values for the Complementary Error (‘Q’) Function

$z = \frac{x-\mu}{\sigma}$	$Q(z) = \frac{1}{\sqrt{2\pi}} \int_z^{\infty} e^{-\frac{x^2}{2}} dx$	
3.71	10^{-4}	
4.26	10^{-5}	
4.75	10^{-6}	
5.19	10^{-7}	
5.61	10^{-8}	
5.99	10^{-9}	
6.36	10^{-10}	
6.70	10^{-11}	
7.03	10^{-12}	

Equation (7) shows that the probability of bit error increases as the SNR decreases. If a fixed attenuation, α , is inserted in the transmission path, then the signal power, P_S , is reduced by a factor of α , while the noise power, P_N , is unchanged. The SNR is therefore reduced from $SNR = P_S/P_N$ to $SNR = P_S/\alpha P_N$. The corresponding $P(\epsilon)$ is increased by a factor that can be calculated using Equation (7) and TABLE II. We can now repeat the test method outlined previously using a modified value for p_h . These calculations can then be extrapolated to any other SNR by using Equation (7). The result is the same, but the test time may be significantly shorter.

TABLE III is an example of the reduction in test time achievable (as compared to TABLE I) by using an optical signal attenuation factor of $\alpha = 1.132$, which results in a reduction of SNR from 12.7 to 11.2. In this example, p_h is increased from 10^{-10} to approximately 10^{-8} (calculated using Equation (7)). This factor of 100 increase in p_h results in a corresponding decrease in test time. Extrapolations of the results in TABLE III to other SNRs retain the same confidence level (99%)

TABLE III
 Example: Estimation of Bit Error Probability ($CL = 99\%$ and $p_h = 10^{-8}$)

Bit Errors $\leq N$ $N =$	Required Number of Bits to Transmit (n)	Test Time for Bit Rate of 622 Mbps (seconds)	Test Time for Bit Rate of 2.5 Gbps (seconds)
0	4.61×10^8	.741	.185
1	6.64×10^8	1.06	.267
2	8.40×10^8	1.35	.337
3	1.00×10^9	1.61	.402
4	1.16×10^9	1.86	.466

The disadvantage of stressing the system is that measurements and calculations must be carried out with more precision. Errors (due to round off, measurement tolerances, etc.) will be multiplied when the results are extrapolated to their non-stressed levels.

V. Conclusion

The concept of statistical confidence levels can be used to establish the quality of an estimate. Application of this idea allows a trade off of test time versus the level of confidence we desire to have in the test results. Test time may be further reduced (at the expense of required precision) by introducing an artificial stress into the system, making the measurements, and extrapolating the results to their non-stressed levels.

-
- ¹ A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. New York: McGraw-Hill, 1984.
- ² K.S. Shanmugan and A.M. Breipohl, *Random Signals: Detection, Estimation, and Data Analysis*. New York: John Wiley and Sons, 1988.
- ³ D.H. Wolaver, "Measure Error Rates Quickly and Accurately," *Electronic Design*, pp. 89-98, May 30, 1995.
- ⁴ J.G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1995.
- ⁵ J.M. Senior, *Optical Fiber Communications: Principles and Practice (second edition)*. Englewood Cliffs, New Jersey: Prentice Hall, 1992
- ⁶ B. Sklar, *Digital Communications: Fundamentals and Applications*. Englewood Cliffs, New Jersey: Prentice Hall, 1988.