This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TCOMM.2016.2541140

Bit Error Rate Comparison Statistics and Hypothesis Tests for Inverse Sampling (Negative Binomial) Experiments

Brian A. Mazzeo, *Senior Member, IEEE*, Michael Rice, *Senior Member, IEEE*

***Abstract*—Inverse, or negative binomial, sampling is often used when the observation of interest occurs extremely infrequently. As this is the case in bit error rate (BER) simulations, especially in high signal-to-noise ratio cases, negative binomial sampling can be advantageously employed in a computationally economic fashion to compare bit error rates between different systems. When the results of two negative binomial sampling tests are compared, point estimates and interval estimates quantify the performance relationship between the results of the tests. This paper derives a new, optimal, logarithmically-symmetric confidence interval estimator for the ratio of BER estimates derived from two negative binomial tests. In addition, a three-sided hypothesis test with a single significance level is derived to quantify the confidence of the relationship between the two systems. Low-BER approximations for the confidence interval and decision thresholds are derived based on the F-distribution. The approximation is shown to work with BERs as high as 10^{-2}. An example inspired by bit interleaved coded modulation shows how the technique can be used to reduce simulation time by an order of magnitude and facilitate straightforward interpretation and comparison between different systems. Negative binomial sampling is recommended for comparison experiments where BER is the key metric.**

***Index Terms*—Inverse sampling, negative binomial, confidence interval, bit error rate**

I. Introduction

Because the bit error rate (BER) is perhaps the most important figure of merit in a digital communication system, it is necessary to rigorously analyze the methods through which this statistic is calculated and how this statistic is employed to make quantitative comparisons between competing communications strategies. Computer simulation is often used to estimate the BERs of two systems. The systems are then compared based on these estimates. For complex communication systems, the simulations can take days or even weeks, especially for extremely small BERs.

The problem addressed in this paper is illustrated by the simulation results presented in Fig. 1. Here, the simulated BER performance of two systems, labeled "System 0" and "System 1", and based on binomial sampling involving 100,000,000 bits at each SNR level is shown. The two systems are described in Section VII. The two BER curves are relatively smooth until $E_b/N_0 = 15$ and 16 dB, the points at which a relatively small number of errors occur. At $E_b/N_0 = 15$ dB it appears System 1 is slightly better than System 0 but the situation is reversed at $E_b/N_0 = 16$ dB. To illustrate this point, the ratio of the BER estimates is plotted in the inset. The ratio is approximately 1 for $10 \leq E_b/N_0 \leq 14$ dB, less than 1 for $E_b/N_0 = 15$ dB, and greater than 1 for $E_b/N_0 = 16$ dB. The natural conclusion is that the two systems have equivalent BER performance and that the differences at $E_b/N_0 = 15$ and 16 dB are small enough to be attributed to the variance of the estimates, characteristic of Monte Carlo simulations. (And this would be the correct conclusion.) However, if the variations in the ratio $\hat{p}_1/\hat{p}_0$ could be quantified in a statistical sense, then the confidence associated with making claims such as "System 0 and System 1 have equivalent BERs" or "System 0 has a lower BER than System 1" can be computed and reported with the simulation results. We will also show that the statistical characterization of the random variable $\hat{p}_1/\hat{p}_0$ resulting from a negative binomial test requires much shorter simulations to make such claims than would otherwise be required to satisfy subjective "smoothness" criteria often selected when examining experimental BER curves.

As outlined in our recent paper [1], BER experiments based on negative binomial (NB) sampling, also known as

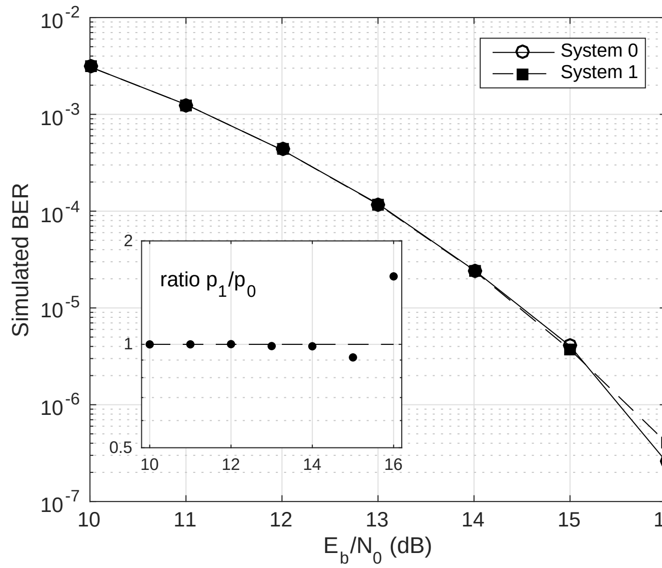


Figure 1. Illustration of BER simulation conditions when it is desirable to have statistical tests to compare performance of two different systems. Due to natural variation inherent in numerical simulations, BER estimates are often not "smooth" curves, particularly under conditions where very few errors are present. The difficult points in this plot at $E_b/N_0 = 15$, 16 dB have BERs that require significant simulation time in order to estimate performance.

B. Mazzeo and M. Rice are with the Dept. Electrical & Computer Engineering, Brigham Young University, Provo, Utah, USA (email: bmazzeo@byu.edu, mdr@byu.edu).

inverse sampling, have the potential to increase the transparency under which different simulations are run. While not increasing simulation time on average, single NB experiments provide exact analyses for confidence intervals (in particular, the logarithmically-centered confidence interval) and improvements for test design and interpretation. Perhaps most importantly, use of NB sampling removes the researcher from the test procedure.

Building significantly on the results for single NB BER experiments, this paper addresses how to statistically compare the results of two NB BER experiments:

- Point estimates and interval estimates *quantify the relationship* and *uncertainty* between two different BER estimates.
- Hypothesis tests provide *decision criteria* to indicate statistically significant differences between two BER tests and to order systems according to performance.

This paper establishes both a computationally economic framework to conduct comparative BER tests and quantitative statistical interpretation of the results. The contributions are as follows. First, we derive a new optimal, logarithmically-symmetric confidence interval estimator for the ratio of two BER estimates obtained using an NB test. Second, we develop a three-sided hypothesis test, with a single significance level, to quantify the confidence of the relationship between the two systems. Third, low-BER approximations for the confidence interval and decision thresholds are derived based on the F-distribution. Simulation results show the approximations to hold for a BER as high as 10^{-2}. Fourth, a rule of thumb is presented to aid the researcher in performing computationally economic NB BER tests.

These contributions apply to *any* communications system whose output errors are well-modeled, or approximately well-modeled, by a Bernoulli process. Care must be taken when attempting to apply these results to situations where the Bernoulli assumption may not hold, such as coded systems where decoder errors occur in bursts or time-varying channels.

II. Performance Considerations for Comparative Tests Using Traditional Binomial Tests and Negative Binomial Tests

For many communications systems, it is assumed that bit errors are independent from other bit errors, resulting in a Bernoulli process (we know that this is often not strictly true, but we often rely on this assumption when we report and discuss performance - models with memory can simulate certain communication strategies with greater fidelity but often result in additional complexity in selection of the memory model [2]).

When comparing the BERs of two different communication systems, we have a choice to either simulate for a fixed number of Bernoulli trials (binomial) or to simulate the experiment until a certain number of errors, r, are reached (NB). In the binomial case, a fixed number of trials, N, is set for the systems to be compared at a particular condition, *e.g.* a signal-to-noise ratio. For a given condition, the two systems will have BERs given by p_0 and p_1. If we desire a certain number of errors on average (though this is not guaranteed in the binomial case), an approximate formula, using Jeruchim's guideline, $r = 10$ [3], might be to set

$$N = \max\left(\frac{r}{p_0},\ \frac{r}{p_1}\right). \tag{1}$$

Thus, for a comparison of two different systems under the same condition and where each system takes unit time to decode one bit, the total binomial simulation time, T_B, is $2N$. If the systems are run in parallel, the total time will be N. In practice we often set N to be some integer power of 10.

If we instead run our simulation until we achieve a certain number of errors, then this procedure is a NB sampling experiment. In this procedure the total number of trials for a single condition will be $X + r$, where X is the NB random variable and r is the specified number of errors, or the stopping criterion for the test. While there are multiple ways of parameterizing a NB random variable, we will use a probability mass function based on the formula

$$\Pr(X = x;\ r,\ p) = \binom{r + x - 1}{x} p^r (1 - p)^x, \tag{2}$$

where p is the underlying BER which must lie on the interval $(0,\ 1)$ [4]. This parametrization is useful because it has support for the random variable X on the interval $[0, \infty)$. For this parameterization of the NB random variable the expected value is $\mathrm{E}\{X + r\} = r/p$, the maximum likelihood (ML) estimator is $\hat{p} = r/(x+r)$, and the minimum variance unbiased (MVU) estimator is $\mathring{p} = (r - 1)/(x + r - 1)$. These estimators are point estimates that can be used to characterize the results of individual NB experiments. To quantify the reliability of these estimates in the context of NB BER simulations, we prefer the use of the MVU point estimator and logarithmic confidence interval, as outlined in our previous work [1].

In contrast to the fixed simulation time for the binomial case, the total simulation time for the NB experiment, T_{NB}, is not fixed. However, we can calculate the expected value for the total time, T_{NB} as

$$\mathrm{E}[T_{NB}] = \frac{r}{p_0} + \frac{r}{p_1}. \tag{3}$$

For the case when $p_0 \neq p_1$, $\mathrm{E}[T_{NB}] < 2N$. Under these common experimental conditions, on average it will take less time to simulate the NB version of the comparison experiment. If the two systems are run in parallel the total time can be estimated by $\max\left(r/p_0, r/p_1\right) = N$. This implies that the equivalent binomial test could be run in the same amount of time with known p_0 and p_1, but in a binomial test the number of errors is not guaranteed and often p_0 and p_1 are not known exactly. Thus, in general for unknown BERs, when a certain number of errors are desired, less total simulation computations will be performed in the NB case than in the binomial case.

In addition to the improved performance of the NB against the binomial sampling under these conditions, we want to emphasize that if an experimenter decides that he has not run enough trials in a traditional binomial test and restarts his simulation and combines the previous data with the simulation data from the restart, this type of experiment is neither binomial nor NB - it is a hybrid experiment and would have an alternative probability mass function that includes the probabilistic behavior of the experimenter. From this point of view, NB sampling effectively removes the experimenter from influencing the type of test being conducted if insufficient errors are produced in a binomial-type experiment. Because of this experimental design consideration, and the fact that, on average, the NB test will produce a desired number of errors reliably and with less simulation time than the binomial test, this again recommends the NB experimental procedure over the binomial one. Consequently, the following analysis will restrict itself to the NB sampling scheme for BER simulations and comparisons.

III. Statistical comparisons of negative binomial experimental outcomes

Once the experiment of generating bit errors using a NB method is performed to compare the performance of two competing schemes under the same conditions, how to analyze the outcomes is important. In general, there are two approaches: simple difference and the ratio[1]. The simple difference is straightforward in that it captures the estimated difference between the two BERs; however, it does not scale well as a comparative performance metric over the orders of magnitude common in BER simulations. The estimated ratio of two BERs captures how well one algorithm performs against another algorithm and does so stably over the wide range of BERs; it has the drawback that it loses the absolute nature that the simple difference contains. For both simple difference and ratios, point estimators and interval estimators can be used. The advantage of interval estimators is that by their length they also provide a measure of the confidence one has in their ability to cover the true parameter. In this section, point estimators and interval estimators are derived for cases relevant to BER comparisons.

A. General point estimators and interval estimators for negative binomial comparisons

NB sampling has been used extensively for biological and epidemiological studies [5], [6]. Because of this, there are already a number of point estimators available for the simple difference and ratio cases. The simple difference estimator formulas are generally straightforward as far as combining the NB ML and MVU estimators in a linear fashion [7]. Different forms of the NB ML estimators and MVU estimators are combined to form ML and MVU ratio estimators [8]. These point estimators are found in Table II. There are many ways of defining confidence intervals for the simple difference and ratio with no clear, optimal choice. Some recent papers compare these confidence intervals for simple difference [9] and ratios [10]. Additionally, alternative parametrizations [11] are often used to model highly dispersed data which perform a fitting to both NB parameters.

For the case of BER testing, more information is known about the underlying error probabilities due to the fact that $p \to 0$ for most BER testing scenarios and the number of desired errors for the two conditions can be set a priori so that $r = r_0 = r_1$. These conditions allow for some simplifications and approximations that can yield specialized estimators and, importantly, an optimal set of confidence intervals for the ratio.

[1]The term "relative difference" is often employed in the statistics literature.

B. Asymptotic ratio as F-distributed random variable

A ratio of NB random variables can be successfully approximated by a simple analytical expression when the underlying probabilities are small. Consider two NB random variables, X_0 and X_1, with BER probability parameters $p_0 \to 0$ and $p_1 \to 0$ and stopping conditions r_0 and r_1, respectively. As outlined by Bennett [12] and alternatively shown in Casella, Berger [13]

$$2p_0X_0 \sim \chi^2_{2r_0} \qquad 2p_1X_1 \sim \chi^2_{2r_1}. \tag{4}$$

Because the scaled ratio of two independent chi-squared random variables is an F-distributed random variable, we have

$$Y = \frac{2p_1X_1/(2r_1)}{2p_0X_0/(2r_0)} \sim F(2r_1, 2r_0) \tag{5}$$

where $F(a,b)$ means an F-distributed random variable with parameters a and b. The corresponding probability density function is

$$f(y;a,b) = \frac{\Gamma\left(\frac{a+b}{2}\right) a^{a/2} b^{b/2}}{\Gamma(a/2)\,\Gamma(b/2)} \frac{y^{a/2-1}}{(b+ay)^{(a+b)/2}}. \tag{6}$$

Simplifying (5) shows that

$$Y = \left(\frac{p_1 r_0}{p_0 r_1}\right)\frac{X_1}{X_0}. \tag{7}$$

That is, a scaled version of the the experimental outcome X_1/X_0 has an F-distribution with parameters $2r_1$ and $2r_0$.

C. Logarithmically symmetric confidence interval for ratio

Because of (5), it is immediately apparent that logarithmically symmetric confidence intervals can be constructed when $r_0 = r_1 = r$. This follows because the F-distributed random variable, Y, under this condition can be written as

$$\ln Y = \ln(p_1X_1) - \ln(p_0X_0) \tag{8}$$

This expression is logarithmically symmetric when these chi-squared random variables are equivalently distributed and independent, which must be the case when we choose $r = r_0 = r_1$. This is unlike the simple difference case where the estimators for the terms in the simple difference are themselves dependent on the underlying probabilities. Because of this logarithmic symmetry, if the confidence interval with $1-\alpha$ confidence level is defined with equal probability above and below the desired midpoint of the interval (in a Clopper-Pearson sense [14]), then the bounds of the interval can be obtained as follows. The significance level, α, represents the probability found in the tails outside of the confidence interval. The logarithmically-symmetric confidence interval estimator is given by

$$\left[\left(\frac{p_1}{p_0}\right)_l, \left(\frac{p_1}{p_0}\right)_u\right]. \tag{9}$$

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TCOMM.2016.2541140

Table I
LOWER AND UPPER BOUNDS FOR LOGARITHMICALLY-CENTERED CLOPPER-PEARSON CONFIDENCE INTERVALS FOR RATIO UNDER THE CONDITIONS $r = r_0 = r_1$ AND $p \to 0$

r	$(1-\alpha)=0.90$	$(1-\alpha)=0.95$	$(1-\alpha)=0.99$
2	[0.1565, 6.3882]	[0.1041, 9.6045]	[0.0432, 23.1545]
5	[0.3358, 2.9782]	[0.2690, 3.7168]	[0.1710, 5.8467]
10	[0.4708, 2.1242]	[0.4058, 2.4645]	[0.3014, 3.3178]
11	[0.4883, 2.0478]	[0.4241, 2.3579]	[0.3200, 3.1246]
20	[0.5907, 1.6928]	[0.5333, 1.8752]	[0.4356, 2.2958]
50	[0.7185, 1.3917]	[0.6742, 1.4833]	[0.5949, 1.6809]
100	[0.7920, 1.2626]	[0.7573, 1.3204]	[0.6937, 1.4416]

The lower bound of the confidence interval is obtained by first finding y_l, defined by

$$\frac{\alpha}{2} = \Pr(Y \le y_l) = F(y_l; 2r, 2r) \tag{10}$$

where $F(y; a, b)$ the cumulative distribution function for the F-distribution and is given by integration of the density function (6). The lower bound is thus

$$y_l = F^{-1}\left(\frac{\alpha}{2}; 2r, 2r\right). \tag{11}$$

From (7) we have, for $r_0 = r_1$,

$$\frac{p_1}{p_0} = \frac{X_0}{X_1} Y. \tag{12}$$

Substituting gives the desired result:

$$\left(\frac{p_1}{p_0}\right)_l = \frac{X_0}{X_1} y_l = \frac{X_0}{X_1} F^{-1}\left(\frac{\alpha}{2}; 2r, 2r\right). \tag{13}$$

The upper bound of the confidence interval is obtained in a similar manner. First, we define y_u using

$$1 - \frac{\alpha}{2} = \Pr(Y \le y_u). \tag{14}$$

Solving for y_u and applying the relationship (12), the upper bound may be expressed as

$$\left(\frac{p_1}{p_0}\right)_u = \frac{X_0}{X_1} F^{-1}\left(1 - \frac{\alpha}{2}; 2r, 2r\right) \tag{15}$$

$$= \frac{X_0/X_1}{F^{-1}\left(\frac{\alpha}{2}; 2r, 2r\right)}. \tag{16}$$

This is a particularly clean result for determining the confidence interval. Because we have shown that the logarithmic interval is logarithmically symmetric, this is also the shortest logarithmic confidence interval that can be constructed for a given $1 - \alpha$ confidence level. Figure 2 illustrates these confidence intervals as a function of the number of errors. As the number of errors increases (as can be determined by the experimenter), it is evident that the confidence interval will be significantly shorter. A useful compilation of the logarithmic confidence interval bounds for values of r that would be common in BER comparison experiments is found in Table I.

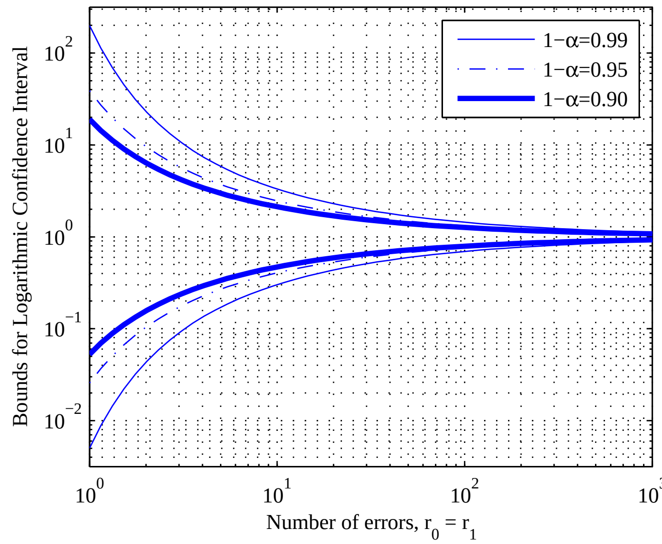


Figure 2. Upper and lower confidence interval bounds for logarithmically-centered ratio confidence intervals using the F-distribution approximation for different values of $r_0 = r_1$.

D. Median estimator

The unique nature of the log-symmetric ratio also introduces the possibility of using the upper and lower bound expressions as an additional estimator. We define a median estimator in the following fashion:

$$\left(\frac{p_1}{p_0}\right)_{\text{median}} = \frac{X_0}{X_1} F^{-1}\left(\frac{1}{2}; 2r, 2r\right) = \frac{X_0}{X_1} \tag{17}$$

This is definitely a biased estimator but has the attractive property that the median of this estimator will indeed converge to the desired ratio. Because of the heavy right tail of the F-distribution, however, we expect this in general to be biased above the actual value of the ratio. An additional use of this estimator for plotting is that it will be found at the center of the logarithmically-centered confidence interval described in the previous section.

E. Summary of comparative estimators

In this section a number of estimators have been described in the literature and logarithmic confidence interval bounds and a median estimator have been derived in this section based on the chi-square approximation to the NB when $p \to 0$. These point estimators and confidence interval bounds are shown in Table II.

IV. ANSWERS TO THE QUESTION: WHICH SYSTEM IS BETTER?

With the point and interval estimators for the ratio p_1/p_0 established, we now move to the question of ordering. Based on the estimate of the ratio p_1/p_0, a researcher often wants to know if the simulation results establish the superiority ($p_1 < p_0$), equivalence ($p_0 = p_1$), or inferiority ($p_1 > p_0$) of the two

Table II
COMPARATIVE NEGATIVE BINOMIAL POINT ESTIMATORS AND THEIR ASSOCIATED FORMULAS WHEN $r = r_0 = r_1$.

Point Estimate	Formula
ML simple difference [7]	$\hat{\Delta} = \hat{p}_0 - \hat{p}_1 = \frac{r}{x_0+r} - \frac{r}{x_1+r}$
MVU simple difference [7]	$\mathring{\Delta} = \mathring{p}_0 - \mathring{p}_1 = \frac{r-1}{x_0+r-1} - \frac{r-1}{x_1+r-1}$
ML ratio [8]	$\left(\frac{p_1}{p_0}\right)_{\text{ML}} = \frac{r+x_0}{r+x_1}$
MVU ratio [8]	$\left(\frac{p_1}{p_0}\right)_{\text{MVU}} = \frac{(r-1)(x_0+r)}{r(r+x_1-1)}$
Median ratio	$\left(\frac{p_1}{p_0}\right)_{\text{median}} = \frac{x_0}{x_1}$
Log-confidence interval lower bound	$\left(\frac{p_1}{p_0}\right)_l = \frac{x_0}{x_1} F^{-1}\left(\frac{\alpha}{2}; 2r, 2r\right)$
Log-confidence interval upper bound	$\left(\frac{p_1}{p_0}\right)_u = \frac{x_0}{x_1} F^{-1}\left(1-\frac{\alpha}{2}; 2r, 2r\right)$

systems. The question is best formulated as a hypothesis test involving the following three hypotheses:

$$\begin{aligned} H_+ &: p_1 < p_0 \\ H_0 &: p_1 = p_0 \\ H_- &: p_1 > p_0 \end{aligned} \tag{18}$$

Because there are three hypotheses, there are a number of ways to formulate a hypothesis test to answer the question [15], [16]. The approach outlined by Goeman, Solari and Stijnen [16] seems best for our application. Here Finner's and Strassburger's partition principle [17] is used to define a composite test for the three hypotheses: a one-sided test is used for H_+ (the "null" hypothesis is $H_0 \cup H_-$) and for H_- (the "null" hypothesis is $H_0 \cup H_+$) whereas a two-sided test is used for H_0. By the partition principle, the significance level (the probability of an incorrect rejection) of the composite test is guaranteed to be α if the significance level of of each of the constituent tests is set to α. Doing so defines the following test:

$$\begin{aligned} &\text{Reject } H_+ \text{ if } X_1/X_0 \le t_{1-\alpha} \\ &\text{Reject } H_- \text{ if } X_1/X_0 \ge t_{\alpha} \\ &\text{Reject } H_0 \text{ if } X_1/X_0 > t_{1-\alpha/2} \text{ or } X_1/X_0 < t_{\alpha/2} \end{aligned} \tag{19}$$

where t_q is the q-th quartile of the distribution of X_1/X_0 under the null hypothesis. The q-th quartile is defined by

$$q = \Pr\left(X_1/X_0 \le t_q\right) = F(t_q; 2r, 2r) \tag{20}$$

where $F(t_q; 2r, 2r)$ is the cumulative distribution function for the F-distribution. The rejection test for H_0 is equivalent to "inverting" the confidence interval estimator described in Section III-C.

With three hypotheses, we use somewhat more precise language than that often used with two hypotheses. Because a statistical test such as this one can either reject or fail to reject H_0 (but never "prove" H_0 true), the language of the test is based on the rejection criteria.

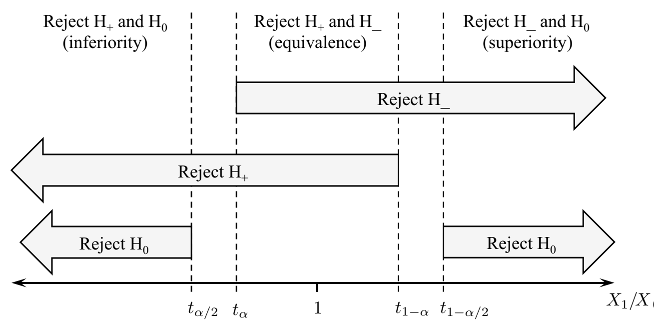


Figure 3. A graphical representation of the test (19).

A graphical representation of the test is illustrated in Figure 3. The graphical representation makes clear how the test may be used to answer the question. This answer may be summarized as follows:

$$\text{Decision} = \begin{cases} p_1 < p_0 & \text{if } X_1/X_0 > t_{1-\alpha} \\ p_1 = p_0 & \text{if } t_{\alpha/2} \le X_1/X_0 \le t_{1-\alpha/2} \\ p_1 > p_0 & \text{if } X_1/X_0 < t_{\alpha}. \end{cases} \tag{21}$$

Note that there are two regions where an answer is not given. The first is $t_{\alpha/2} \le X_1/X_0 < t_{\alpha}$, where only H_+ is rejected. In this case the answer is "non-superiority" in the sense that the hypothesis $p_1 < p_0$ is rejected, but the hypotheses $p_0 = p_1$ and $p_1 > p_0$ are not rejected. The second case, $t_{1-\alpha} < X_1/X_0 \le t_{1-\alpha/2}$, is similar: the answer is "non-inferiority" and means that the hypothesis $p_1 > p_0$ is rejected, but the hypotheses $p_1 = p_0$ and $p_1 < p_0$ are not rejected.

The behavior of the decision regions is illustrated by the numerical examples in Figures 4 and 5. In Figure 4, the significance level is fixed at $\alpha = 0.05$. The decision thresholds

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TCOMM.2016.2541140

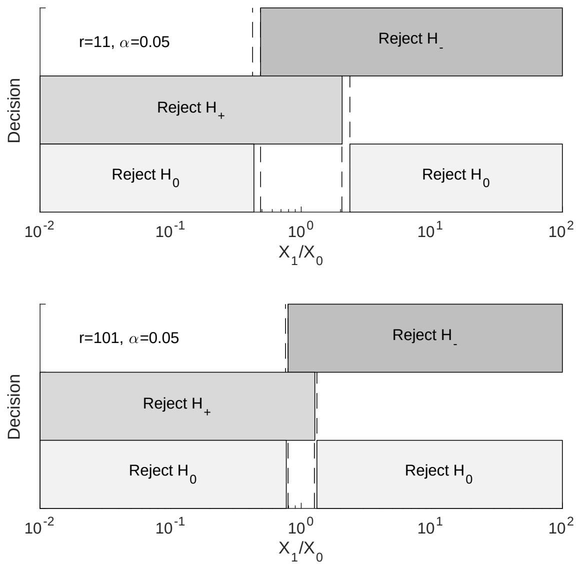


Figure 4. A graphical representation of the test (19) for $\alpha = 0.05$ and $r = 11$ (top), $r = 101$ (bottom).

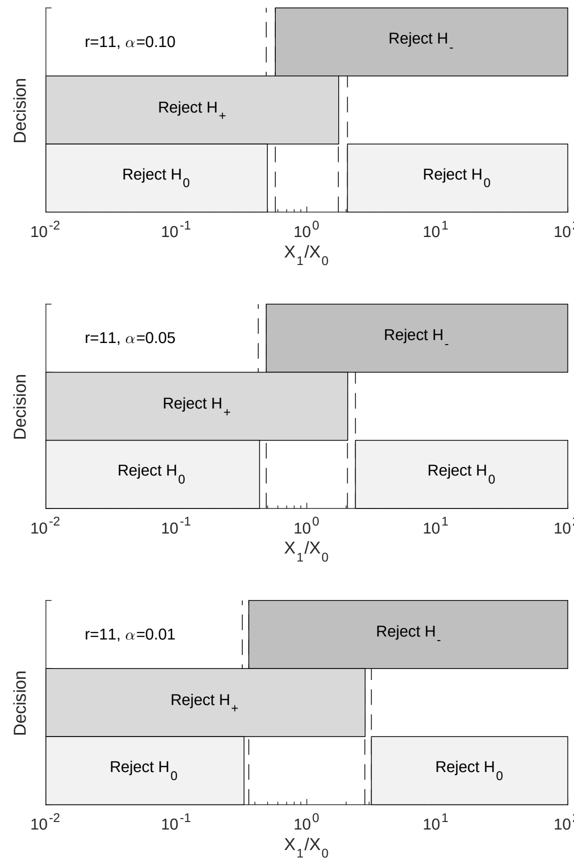


Figure 5. A graphical representation of the test (19) for $r = 11$ and $\alpha = 0.10$ (top), $\alpha = 0.05$ (middle), $\alpha = 0.01$ (bottom).

are computed for two values of the stopping criterion: $r = 11$ and $r = 101$. The plots show for fixed α the width of the "non-inferiority" and "non-superiority" decision regions shrink. This makes intuitive sense because decisions about superiority, equivalence, and inferiority become more certain as the number of observed errors increases.

In Figure 5, the stopping criteria r is fixed at 11 and the decision thresholds are computed for three significance levels $\alpha = 0.10, 0.05, 0.01$. The plots show that the rejection regions for H_- and H_+ get longer with decreasing α whereas the rejection regions for H_0 get shorter for with decreasing α. This behavior is expected because t_α decreases and $t_{1-\alpha}$ increases as α decreases. The same behaviors are observed with $t_{\alpha/2}$ and $t_{1-\alpha/2}$, although these two terms change with α at a slightly different rate with decreasing α, hence the slight decreases in the widths of the "non-inferiority" and "non-superiority" decision regions.

A list of the decision thresholds used by (21) for various values of stopping criterion r and confidence level $1 - \alpha$ is given in Table III. It is noted that because of the "inverse" relationship between the confidence interval and the two-sided hypothesis the bounds of the confidence intervals in Table I are the same values for the thresholds $t_{\alpha/2}$ and $t_{1-\alpha/2}$ in Table III. These decision thresholds are logarithmically symmetric because they originate from the ratio of the chi-squared random variables shown in (5).

V. Simulation

To test the effectiveness of the theoretical approximations outlined above, the different point estimators and interval estimators were simulated on a computer for a variety of situations of interest in BER comparison studies.

The performance of the ML, median, and MVU ratio point esimators was assessed by comparing scenarios where $p_0 > p_1$ and $p_0 < p_1$. The stopping criteria r was varied through approximately logarithmically-distributed integer values from 1 to 1000. For each r, two comparative tests were performed:

1) One NB random variable was simulated with $p_0 = 10^{-7}$ and a second NB random variable was simulated with $p_1 = 10^{-6}$.
2) One NB random variable was simulated with $p_0 = 10^{-7}$ and a second NB random variable was simulated with $p_1 = 10^{-8}$.

For each r, 100,000 trials were conducted and the mean, median, and standard deviation of the different ratio point estimators were calculated. The results for the ML point estimator are shown in Fig. 6. The results for the median and MVU point estimators are shown in Figs. 7 and 8.

To test the performance of the asymptotic logarithmically-centric confidence interval (and equivalently the performance of the two-sided hypothesis test) an additional set of simulations were employed. From the preceding analysis, we understand from the asymptotic analysis that these ratio test relationships will certainly be accurate at very low BERs when $p_i \to 0$. How these confidence interval estimators perform at high BERs can also be important.

To assess this, NB variables at varying BERs were simulated and confidence intervals were calculated so the estimated coverage probabilities could be compared quantitatively with the theoretically predicted coverage probabilities. This was performed for three different stopping criteria, $r = 1$, 11, and

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TCOMM.2016.2541140

Table III
DECISION THRESHOLDS FOR THE TEST (21) FOR VARIOUS VALUES OF STOPPING CRITERION r AND CONFIDENCE LEVEL $1-\alpha$.

$(1-\alpha)=0.90$				
r	$t_{\alpha/2}$	t_{α}	$t_{1-\alpha}$	$t_{1-\alpha/2}$
2	0.1565	0.2435	4.1072	6.3882
5	0.3358	0.4306	2.3226	2.9782
10	0.4708	0.5575	1.7938	2.1242
11	0.4883	0.5734	1.7440	2.0478
20	0.5907	0.6642	1.5056	1.6928
50	0.7185	0.7731	1.2934	1.3917
100	0.7920	0.8339	1.1991	1.2626
$(1-\alpha)=0.95$				
r	$t_{\alpha/2}$	t_{α}	$t_{1-\alpha}$	$t_{1-\alpha/2}$
2	0.1041	0.1565	6.3882	9.6045
5	0.2690	0.3358	2.9782	3.7168
10	0.4058	0.4708	2.1242	2.4645
11	0.4241	0.4883	2.0478	2.3579
20	0.5333	0.5907	1.6928	1.8752
50	0.6742	0.7185	1.3917	1.4833
100	0.7573	0.7920	1.2626	1.3204
$(1-\alpha)=0.99$				
r	$t_{\alpha/2}$	t_{α}	$t_{1-\alpha}$	$t_{1-\alpha/2}$
2	0.0432	0.0626	15.9770	23.1545
5	0.1710	0.2062	4.8491	5.8467
10	0.3014	0.3404	2.9377	3.3178
11	0.3200	0.3591	2.7849	3.1246
20	0.4356	0.4730	2.1142	2.2958
50	0.5949	0.6259	1.5977	1.6809
100	0.6937	0.7188	1.3912	1.4416

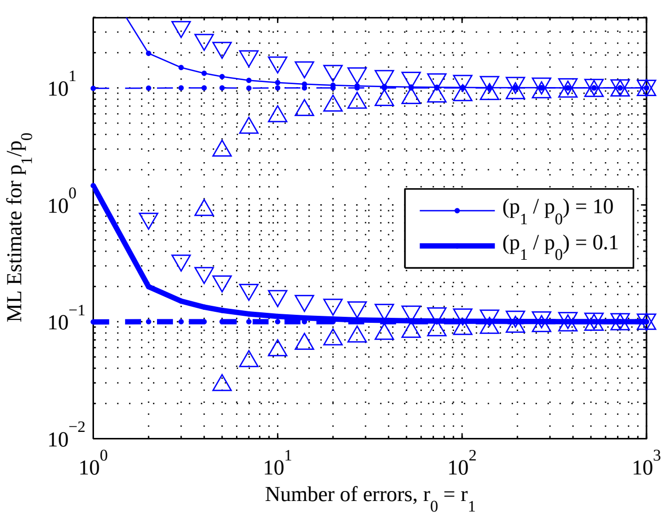


Figure 6. Simulated performance of ML ratio estimator for $p_0 = 10^{-7}$. Solid lines represent the calculated means for the estimators. Dashed lines represent the calculated medians for the estimators. The triangle symbols represent points $\pm$ 1 standard deviation from the means.

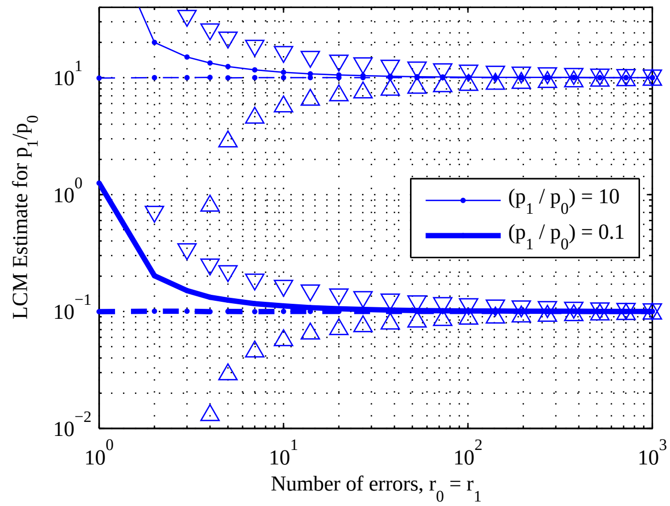


Figure 7. Simulated performance of median ratio estimator for $p_0 = 10^{-7}$. Solid lines represent the calculated means for the estimators. Dashed lines represent the calculated medians for the estimators. The triangle symbols represent points $\pm$ 1 standard deviation from the means.

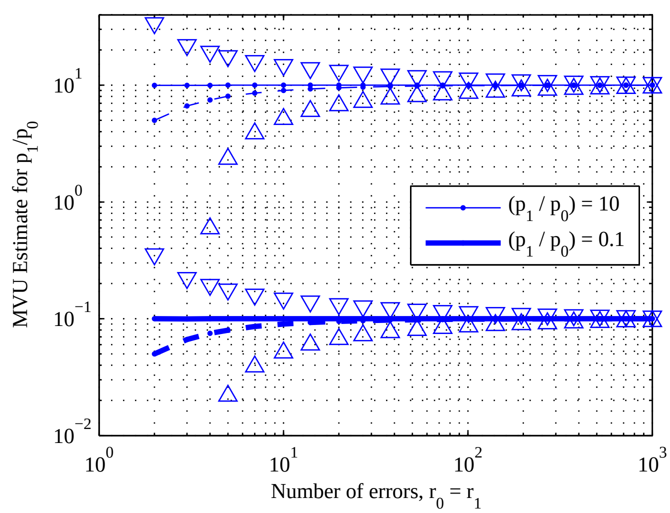


Figure 8. Simulated performance of MVU ratio estimator for $p_0 = 10^{-7}$. Solid lines represent the calculated means for the estimators. Dashed lines represent the calculated medians for the estimators. The triangle symbols represent points $\pm$ 1 standard deviation from the means.

100. The condition $r = 11$ was chosen to match criteria for the single parameter estimation case [1]. The BERs for two NB variables were set equal, $p_0 = p_1$, and were logarithmically spaced from 10^{-5} to 3×10^{-2}. At each stopping condition and BER, 10^6 trials of the two NB variables were conducted. Lower and upper bound criteria for the confidence intervals at three confidence levels $1-\alpha = 0.90$, 0.95, and 0.99 were used to determine if the calculated confidence interval from the two NB variable outcomes contained the true BER ratio $p_1/p_0 = 1$. The coverage probabilities were then calculated for each pair of stopping condition and BER and plotted as shown in Fig. 9.

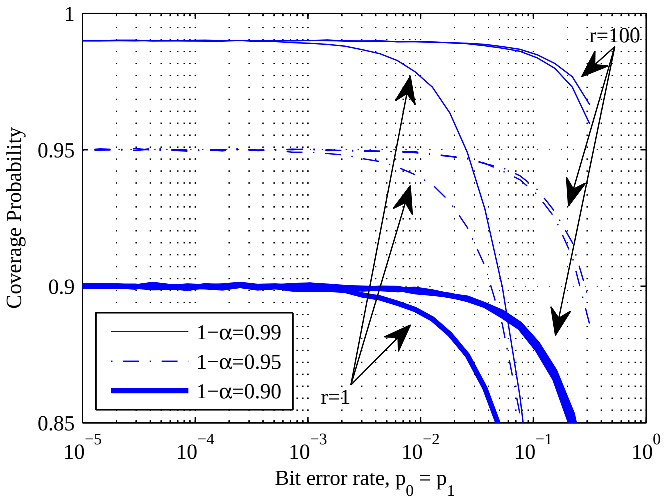


Figure 9. Simulated coverage probability versus BER using logarithmically-centered confidence intervals for $r = 1,\ 11,\ 100$.

VI. Discussion

A. Comparison of ratio point estimators

The ratio point estimator simulations reveal the nature of each of the different types of estimators. For the ML estimator, the mean is consistently biased above the actual ratio. As r increases, the ML estimator significantly improves. The ML estimator also performs very well as a median estimator under conditions relevant for BER testing. This is because of the approximations that can be made in the small probability case:

$$\begin{aligned} \Pr\left(\frac{X_0+r}{X_1+r}<\frac{p_1}{p_0}\right) &\approx \Pr\left(\frac{X_0}{X_1}<\frac{p_1}{p_0}\right) \\ &\approx \Pr\left(\frac{\chi_0^2 p_1}{\chi_1^2 p_0}<\frac{p_1}{p_0}\right) \\ &= \Pr\left(\chi_0^2<\chi_1^2\right) \\ &= \frac{1}{2}. \end{aligned} \tag{22}$$

The chi-squared random variables have the same number of degrees of freedom and so are equivalent - one variable is equally likely to be greater or smaller than the other so the probability is split evenly between the estimate for p_1/p_0 being greater or less than the true value.

Similar to the ML estimator, but derived directly from the F-distribution approximation to the ratio of the NB random variables, the median estimator clearly has its median at the a priori ratio. The performance of the ML estimator and the median estimator are almost identical because r is negligible in the ML formula for the large X_0 and X_1 expected when $p \to 0$.

The MVU estimator is clearly the best choice if a mean of the ratio estimates is used. It works as designed and its mean clearly estimates correctly the a priori ratio. Its median is below the mean, but that is expected because of the skew inherent in the ratio distribution. The MVU estimator would be the preferred statistic to report for the estimated ratio in a bit error test because it is unbiased.

B. Performance of logarithmically-centered confidence interval estimator

From the simulation data comparing logarithmically-centered confidence intervals for a fixed number of errors, we see that the coverage probabilities achieve their targeted values using the F-distribution ratio approximation for $p \to 0$ as expected. For the condition $r = 1$ the asymptotic confidence intervals are already almost accurate to their desired coverage probabilities below the 10^{-2} BER level. The confidence intervals increase in accuracy as the confidence level increases. Additionally, while the approximation is accurate at the $r = 1$ level (when the distributions are actually geometric random variables), the accuracy of the approximation increases as the number of errors, r increases, which will be typical of real comparisons. Because the two-sided hypothesis test uses the same bounds as the confidence intervals due to their "inverse" relationship, this also validates our proposed hypothesis testing under these conditions. For almost all digital communication systems, the BERs of interest are those that are below the level 10^{-3}. This provides the rational justification for the use of the F-distribution approximation for most of the digital communication systems of practical interest.

When comparing the estimated BERs using NB sampling, logarithmically-centered confidence intervals for the BER ratios are relatively easy to construct and to visualize on logarithmic plots and thus can be used in comparisons of estimated BER probabilities. Based on the confidence interval and three-sided hypothesis test presented in this work, this naturally results in the following rule of thumb:

- If $r_0 = r_1 = 11$, a ratio of BERs of approximately 2:1 is required to be statistically significant at 90% confidence.

While this paper has dealt with the case when $r_0 = r_1$, which is the most common case when the researcher has control of the simulation parameters, the F-distribution approximation analysis can be useful for the case when $r_0 \neq r_1$. However, under those changed conditions, the distribution of Y will no longer be logarithmically symmetric and so other criteria would be needed to establish additional parameters to select a confidence interval. This demonstrates that the analysis for BER simulations where control of r_0 and r_1 is possible (the focus of this work) is particularly clean and yields useful estimators and interval analysis that can be readily applied to comparison tests where the metric of interest is the BER.

VII. Examples

As an example of the forgoing analysis, we consider using computer simulation to assess the bit error performance for different bit-to-constellation point mappings for non-binary modulations. Bit decisions are made based on bit-level log-likelihood ratios (LLRs) as computed for bit interleaved coded modulation (BICM) [18], [19]. Normally BICM employs an error correction code operating on the bit level LLRs. Unfortunately, decoder errors tend to occur in bursts and therefore do not conform to the statistical model assumed in this paper. For example the decoder for an LDPC code with block size n outputs the correct codeword (a block of n bits with zero errors) or a block of n bits with usually more than $n/10$ errors when the decoder reaches the maximum number of allowed iterations. Here we omit the code (but retain the bit-level interleaver) to illustrate the application of the principles.

a) Example 1: For the first example, we use 16-APSK modulation, defined in the DVB-S2 standard [20], and a length-100,000 S-random interleaver [21]. System "0" is defined by the bit-to-constellation point mappings defined in [20]. The main feature of this mapping is its Gray code property. For the first set of simulations, system "1" is defined by a different Gray code for the bit-to-constellation point mapping. We expect the BER performance of the two systems to be identical because the Gray code property is preserved. To begin, we performed the "traditional" binomial test using $N = 100{,}000{,}000$ bits for each value of E_b/N_0. The MVU point estimates are shown in Fig. 1 and discussed in the introduction. The points of interest are $E_b/N_0 = 15$ and 16 dB, where small variations suggest somewhat conflicting conclusions. Furthermore, the confidence intervals associated with these point estimates have different lengths, an undesirable property [1].

Next, the NB test with $r = 11$ was used to estimate the BER. The MVU point estimates for p are plotted in Fig. 10. With a relatively small number of errors, it is hard to draw any firm conclusions from a plot such as this one. However, a plot of the BER ratio estimate tends to reveal more. The log-confidence intervals for the ratio for this simulation are plotted in Fig. 11. For each value of E_b/N_0, three logarithmically-centered confidence intervals (for increasing confidence levels) are shown. The confidence interval covers the true ratio with probability $1 - \alpha$. Each of the intervals inclues $x_0/x_1 = 1$ and this *suggests* the BERs of the two systems are equivalent. [This will be made more precise below in the discussion about the decisions from the hypothesis test (21).] Also included is the plot is the median point estimate (see Table II). By design, the confidence intervals are logarithmically-centered about the median point estimates. From Table I, we see that the 90% confidence interval extends from approximately one-half the point estimate to twice the point estimate. In other words, if the 90% confidence interval covers unity, then we are 90% confident that the two systems are equivalent. This observation may be "inverted" by placing two horizontal lines, one at 2 and the other at $1/2$ as shown in the figure, and drawing the same qualified conclusion for equivalency if the median point estimate lies between the two lines. This illustrates a simple rule-of-thumb with the median estimator. Even though the BER curves of Fig. 10 are not smooth, the relatively quick NB experiment for $r = 11$ (the simulation time was 4% of that required for binomial test) and the rule-of-thumb for the point estimate of the ratio allows the researcher to reach a quick initial conclusion regarding equivalence.

More formal decisions about equivalence may be drawn from the decision rule (21) and its illustration in Fig. 5. Using the thresholds corresponding to $r = 11$ in Table III, the decision regarding the relationship between the two systems is "equivalence" with significance levels 90%, 95% and 99%.

For comparison, the NB test with $r = 101$ was also run for the two systems. The MVU point estimates for p are plotted in Fig. 12. Here the BER curves are quite a bit smoother than

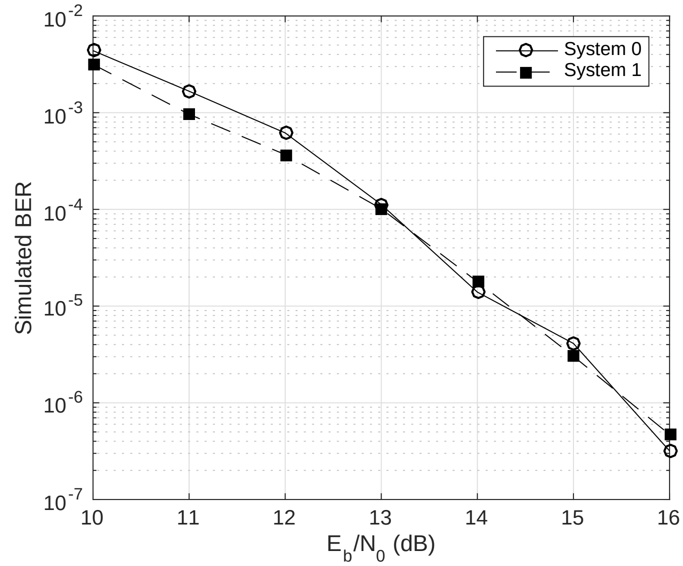


Figure 10. The MVU estimates of p resulting from the negative binomial test with $r = 11$ as found in Example 1.

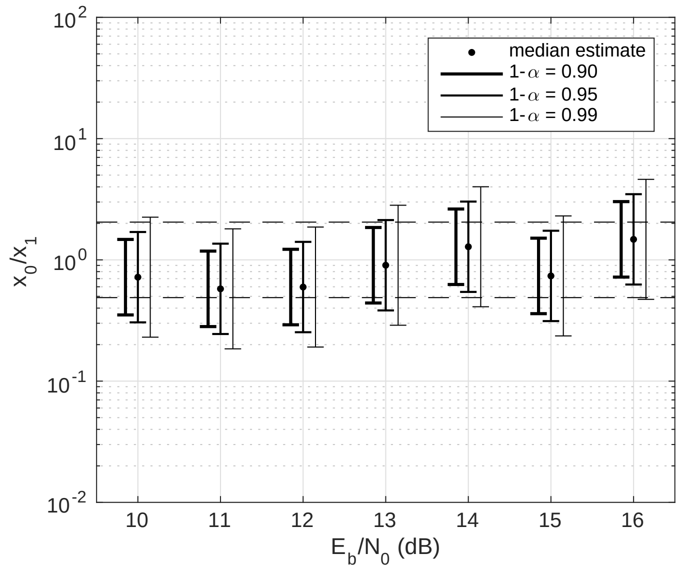


Figure 11. The log-confidence intervals for negative binomial test with $r = 11$ for Example 1 for $1 - \alpha = 0.90$, 0.95, and 0.99. The dashed lines define the quick rule-of-thumb described in the text.

the curves in Fig. 10 and provide a more visually pleasing case for equivalence. The confidence intervals for the BER estimate ratios are plotted in Fig. 13. The confidence intervals are much shorter than their counterparts in Fig. 11. The confidence intervals all cover $x_0/x_1 = 1$, *suggesting* equivalence. But application of the decision rule(21) produces the results shown in Table IV. The decisions are as expected except for $E_b/N_0 = 12 - 14$ dB for 90% confidence. That this might be the case could be inferred from the 90% confidence intervals in Fig. 13: they barely cover $x_0/x_1 = 1$ for $E_b/N_0 = 12 - 14$ dB. In any event, the simulation time for this experiment was 41% of that required for the binomial test.

b) Example 2: As a second example, a random permutation of the bit-to-constellation point mapping in the DVB-S2

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TCOMM.2016.2541140

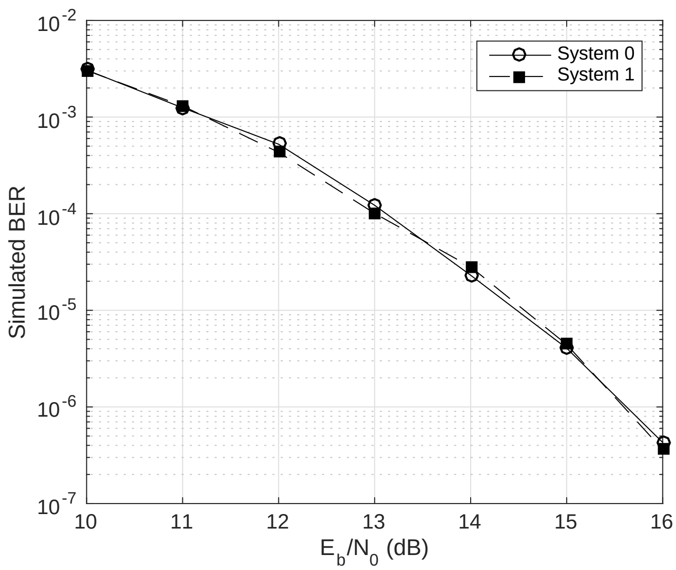


Figure 12. The MVU estimates of p resulting from the negative binomial test with $r = 101$ as found in Example 1.

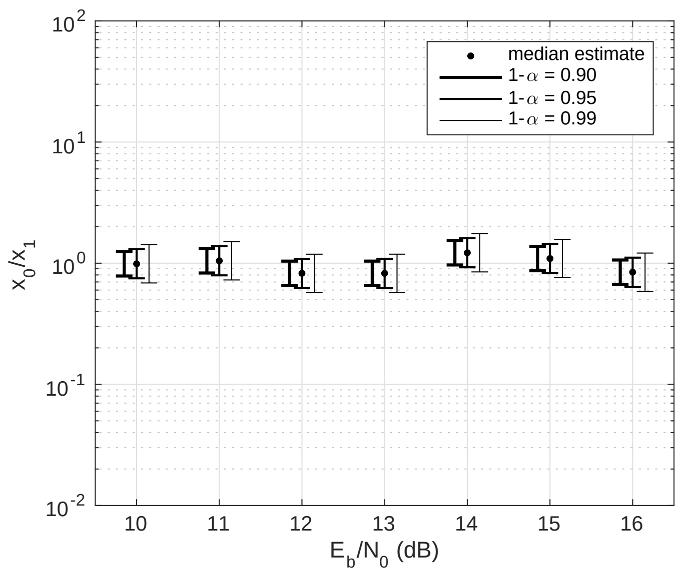


Figure 13. The log-confidence intervals for negative binomial test with $r = 101$ for Example 1 for $1 - \alpha = 0.90$, 0.95, and 0.99.

Table IV
DECISIONS, IN THE SENSE OF (21) AND FIGURE 3, FOR EXAMPLE 1 WITH $r = 101$ FOR THREE DIFFERENT CONFIDENCE LEVELS.

E_b/N_0 (dB)	Decision $(1-\alpha) = 0.90$	Decision $(1-\alpha) = 0.95$	Decision $(1-\alpha) = 0.99$
10	equivalence	equivalence	equivalence
11	equivalence	equivalence	equivalence
12	non-inferiority	equivalence	equivalence
13	non-inferiority	equivalence	equivalence
14	non-inferiority	equivalence	equivalence
15	equivalence	equivalence	equivalence
16	equivalence	equivalence	equivalence

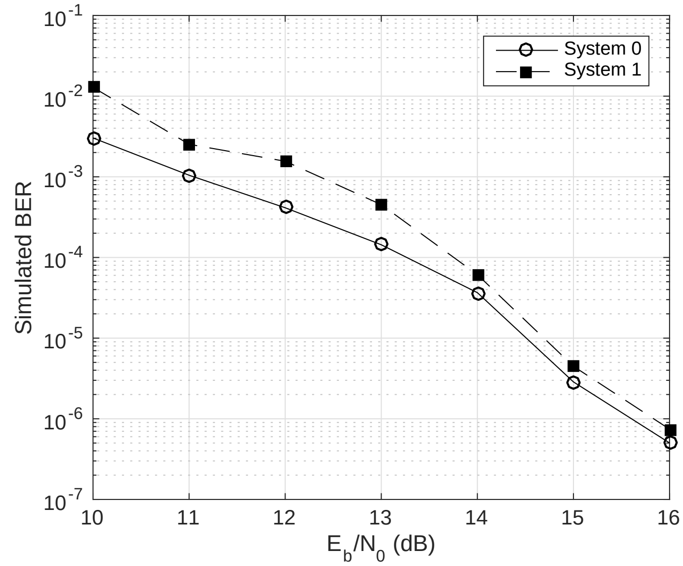


Figure 14. The MVU estimates of p resulting from the negative binomial test with $r = 11$ as found in Example 2.

standard was used to define System 1. One expects the BER for System 1 to be higher (worse) than that of System 0 because the Gray code property of the mapping no longer exists. A plot of the simulated BER using Monte Carlo simulations based on the negative binomial test using $r = 11$ is shown in Figure 14. This plot seems to confirm the belief that System 1 is inferior to System 0, but it is difficult to be too certain. As before, the situation becomes a little more clear when examining the ratio x_0/x_1. Confidence intervals and the median point estimates are plotted in Fig. 15. The horizontal lines defining the rule-of-thumb for equivalence are also shown. The rule of thumb suggests that with 90% confidence, equivalence is rejected for $E_b/N_0 = 10 - 13$ dB but not rejected for $E_b/N_0 = 14 - 16$. With reference to the decision thresholds illustrated in Figure 5 and using the thresholds corresponding to $r = 11$ in Table III, the decisions for each value of E_b/N_0 and for three different confidence levels are summarized in Table V. For the 90% and 95% confidence levels, these results show that System 1 is inferior to System 0 for $E_b/N_0 = 10{-}13$ dB, but equivalent to System 0 for $E_b/N_0 = 14{-}16$. For the higher confidence level of 99%, the conclusion is the same except for $E_b/N_0 = 11$ dB. Given our experienced intuition about how these systems behave, this is probably a good indicator that the simulation for $E_b/N_0 = 11$ dB needs to be examined in greater detail.

In any event, it is important to note that the simulation time required to produce the data in Figures 14 and 15 and in Table V was only 3% of the simulation time required to produce the "smooth" curve (corresponding to the binomial test) in Figure 1. This is less than the 4% reported in the first simulation of Example 1 because in this example the performance of System 1 has a slightly higher BER, information that may not be known a priori in many experiments.

The NB test for $r = 101$ for this example was also run. A plot of the MVU point estimates for the BERs is given in Fig. 16 and the usual information about the ratio x_0/x_1 is shown in Fig. 17. The BER curve in Fig. 16 is much

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TCOMM.2016.2541140

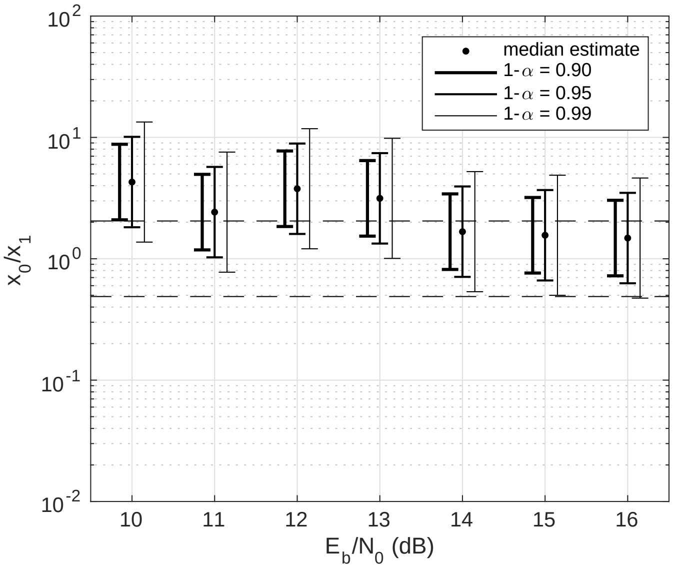


Figure 15. The log-confidence intervals for negative binomial test with $r = 11$ for Example 2 for $1 - \alpha = 0.90$, 0.95, and 0.99.

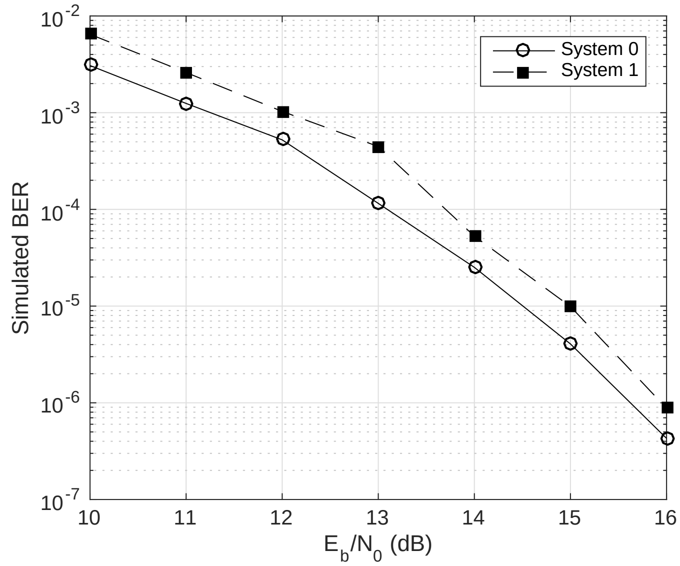


Figure 16. The MVU estimates of p resulting from the negative binomial test with $r = 101$ as found in Example 2.

Table V
DECISIONS, IN THE SENSE OF (21) AND FIGURE 3, FOR EXAMPLE 2 FOR THREE DIFFERENT CONFIDENCE LEVELS.

E_b/N_0 (dB)	Decision $(1-\alpha) = 0.90$	Decision $(1-\alpha) = 0.95$	Decision $(1-\alpha) = 0.99$
10	inferiority	inferiority	inferiority
11	inferiority	inferiority	equivalence
12	inferiority	inferiority	inferiority
13	inferiority	inferiority	inferiority
14	equivalence	equivalence	equivalence
15	equivalence	equivalence	equivalence
16	equivalence	equivalence	equivalence

more smooth than its counterpart in Fig. 14 and makes it more clear that System 1 is inferior to System 0. Both the interval and point estimates for x_0/x_1 in Fig. 17 confirm this. Application of the decision rule of (21) decides System 1 is inferior to System 0 with confidence levels if 90%, 95% and 99%. These relatively certain results were achieved in a simulation run requiring only 28% of the computer time required by the binomial test. Again, this is less computer time required than the 41% reported in the second simulation of Example 1 because of the increased BER of System 1.

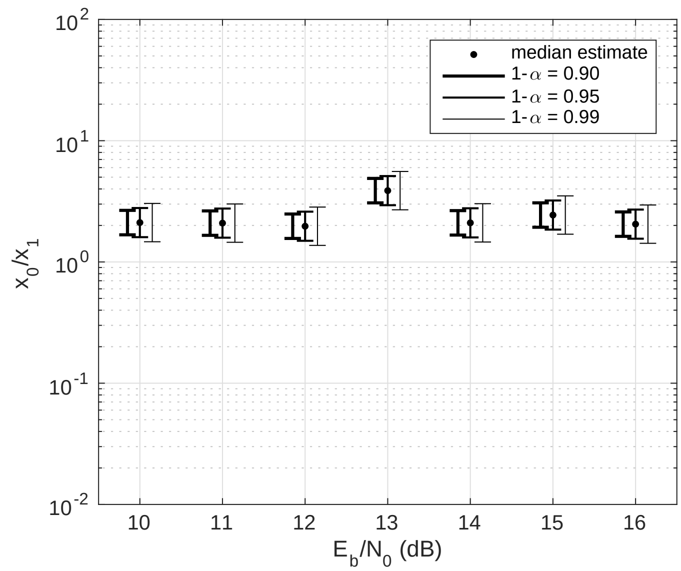


Figure 17. The log-confidence intervals for negative binomial test with $r = 101$ for Example 2 for $1 - \alpha = 0.90$, 0.95, and 0.99.

VIII. CONCLUSION

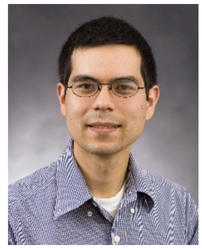
NB sampling is a powerful technique for estimating the probability of rare events without prior knowledge of the order of magnitude of the probability of occurrence. It is thus ideally suited for BER tests, especially when paired with point estimates and logarithmically-centered interval estimates of the bit error probabilities [1]. It effectively removes the researcher from making a priori judgments about the underlying probabilities in their experiments.

Given that a NB experiment can be performed to compare two competing communication schemes, in this paper important point estimate statistics were outlined to compare the performance of the schemes. Their performance was simulated and explained. For most cases, the MVU ratio estimate is preferred because it is an unbiased point estimator (though in situations were the median is used, other statistics may be preferable). Additionally, through the F-distribution approximation using the chi-squared approximation of the NB, a new logarithmically-centered ratio confidence interval was presented with optimal length properties. This confidence interval can and should be used when presenting data about comparison experiments between NB tests. A simple rule of thumb was introduced that can guide researchers when making initial judgments about the results of their experiments.

This paper expands the knowledge base necessary for performing comparative statistical tests of BER simulations using NB sampling. Because of the significant advantages in simulation time, experiment design and interpretation, and simple construction of confidence intervals, it is hoped that the NB approach will be more generally adopted to increase the clarity and rigor of the communications literature.

REFERENCES

[1] M. Rice and B. Mazzeo, "On the superiority of the negative binomial test over the binomial test for estimating the bit error rate," *IEEE Transactions on Communications*, vol. 60, no. 10, pp. 2971–2981, 2012.
[2] P. Sadeghi, R. A. Kennedy, P. B. Rapajic, and R. Shams, "Finite-state markov modeling of fading channels-a survey of principles and applications," *Signal Processing Magazine, IEEE*, vol. 25, no. 5, pp. 57–80, 2008.
[3] M. Jeruchim, "Techniques for estimating the bit error rate in the simulation of digital communication systems," *IEEE J. Sel. Areas Commun.*, vol. 2, no. 1, pp. 153–170, 1984.
[4] N. L. Johnson, A. W. Kemp, and S. Kotz, *Univariate Discrete Distributions*. Hoboken, New Jersey: Wiley, 3rd ed., 2005.
[5] J. B. S. Haldane, "On a method of estimating frequencies," *Biometrika*, vol. 33, no. 3, pp. 222–225, 1945.
[6] K.-J. Lui, *Statistical Estimation of Epidemiological Risk*. Chichester: John Wiley & Sons, Ltd., 2004.
[7] K.-J. Lui, "Interval estimation of simple difference under independent negative binomial sampling," *Biometrical Journal*, vol. 41, no. 1, pp. 83–92, 1999.
[8] K.-J. Lui, "Point estimation on relative risk under inverse sampling," *Biometrical Journal*, vol. 38, no. 6, pp. 669–680, 1996.
[9] M.-L. Tang and M. Tian, "Asymptotic confidence interval construction for risk difference under inverse sampling," *Computational Statistics and Data Analysis*, vol. 53, no. 3, pp. 621–631, 2009.
[10] M. Tian, M. L. Tang, H. K. Ng, and P. S. Chan, "Confidence intervals for the risk ratio under inverse sampling," *Stat Med*, vol. 27, no. 17, pp. 3301–24, 2008.
[11] D. Shilane, S. N. Evans, and A. E. Hubbard, "Confidence intervals for negative binomial random variables of high dispersion," *Int J Biostat*, vol. 6, no. 1, p. Article 10, 2010.
[12] B. M. Bennett, "On the use of the negative binomial in epidemiology," *Biometrical Journal*, vol. 23, no. 1, pp. 69–72, 1981.
[13] G. Casella and R. L. Berger, *Statistical Inference*. Duxbury Press, 2nd ed., 2002.
[14] C. J. Clopper and E. S. Pearson, "The use of confidence or fiducial limits illustrated in the case of the binomial," *Biometrika*, vol. 26, no. 4, pp. 404–413, 1934.
[15] L. Leventhal and C.-L. Huynh, "Directional decisions for two-tailed tests: Power, error rates, and sample size," *Psychological Methods*, vol. 1, no. 3, pp. 278–292, 1996.
[16] J. Goeman, A. Solari, and T. Stijnen, "Three-sided hypothesis testing: Simultaneous testing of superiority, equivalence and inferiority," *Statistics in Medicine*, vol. 29, pp. 2117–2125, 2010.
[17] H. Finner and K. Strassburger, "The partitioning principle: A powerful tool in multiple decision theory," *The Annals of Statistics*, vol. 30, pp. 1194–1213, 2002.
[18] E. Zehavi, "8-PSK trellis codes for a Rayleigh channel," *IEEE Transactions on Communications*, vol. 40, no. 5, pp. 873–884, May 1992.
[19] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Transactions on Information Theory*, vol. 44, no. 3, pp. 947–946, May 1998.
[20] *Digital Video Broadcasting-Satellite-Second Generation (DVB-S2)*, ESTI EN 302 307-2 v1.1.1, 2015.
[21] S. Dolinar and D. Divsilar, "Weight distribution for turbo codes using random and nonrandom permutations," in TDA Progress Rep. 42-122, Aug 1995.



Brian Mazzeo (M'09 SM'14) received the BS degree from the Massachusetts Institute of Technology, Cambridge, MA, USA in 2005 and the Ph.D. degree from Cambridge University, United Kingdom, in 2008 as a Marshall Scholar. He joined the faculty at Brigham Young University in 2008 where he is currently an Associate Professor in the Department of Electrical & Computer Engineering.

Professor Mazzeo's research interests are in the area of nondestructive measurement of material properties and in the area of probabilistic systems and statistical signal processing.

Michael Rice (M'82 SM'98) received a BSEE from Louisiana Tech University in 1987 and his PhD from Georgia Tech in 1991. Dr. Rice was with Digital Transmission Systems, Inc. in Atlanta and joined the faculty at Brigham Young University in 1991 where he is currently a Professor in the Department of Electrical & Computer Engineering. Professor Rice was a NASA/ASEE Summer Faculty Fellow at the Jet Propulsion Laboratory during 1994 and 1995 where he worked on land mobile satellite systems. During the 1999-2000 academic year, Professor Rice was a visiting scholar at the Communication Systems and Signal Processing Institute at San Diego State University.

Professor Rice's research interests are in the area of digital communication theory and statistical signal processing with a special emphasis on applications to aeronautical telemetering and software radio design. He has been a consultant to both government and industry on telemetry related issues and currently serves as an associate member of the Telemetry Group of the Range Commander's Council. He is a member of the IEEE Communications Society and is past chair of the Communication Theory Technical Committee. He is also a member of the Aerospace and Electronic Systems Society and is the Technical Editor for Command, Control and Communication Systems and an Associate Editor-in-Chief for IEEE Transactions on Aerospace and Electronic Systems. He serves on the Board of Governors for the Aerospace and Electronic Systems Society.