On the Mean Absolute Error in Inverse Binomial Sampling

Luis Mendo∗

April 24, 2009

Abstract
A closed-form expression and an upper bound are obtained for the mean absolute error of the unbiased estimator of a probability in inverse binomial sampling. The results given permit the estimation of an arbitrary probability with a prescribed level of the normalized mean absolute error.

Keywords: Inverse binomial sampling, Sequential estimation, Mean-absolute error.

1 Introduction

The estimation of a probability from a set of observations is one of the most basic problems in statistics, with many applications in signal processing and related fields. One commonly used technique is inverse binomial sampling (also termed negative-binomial Monte Carlo). Given a sequence of independent Bernoulli trials with probability of success \( p \), this technique consists in observing the sequence until a given number \( N \) of successes is obtained. The resulting number of trials is denoted as \( n \). The uniformly minimum variance unbiased estimator of \( p \), for \( N \geq 2 \), is \[
\hat{p} = \frac{N - 1}{n - 1}.
\] (1)

For \( N \geq 3 \) the mean square error (MSE) of (1) is known to satisfy \( \mathbb{E}[(\hat{p} - p)^2]/p^2 < 1/(N - 2) \) irrespective of \( p \) [1]. Recent works [2] [3] have shown that, for the modified estimator \( \hat{p} = (N - 1)/n \), the confidence level associated to a relative interval of the form \( [p/\mu_2, p\mu_1] \) also satisfies a lower bound irrespective of \( p \), for \( N \geq 3 \) and a certain range of \( \mu_1, \mu_2 \) values. A similar result also holds for the estimator (1), albeit for a reduced range of \( \mu_1, \mu_2 \) values [4].

This paper analyzes the mean absolute error (MAE) of the estimator (1), for \( N \geq 2 \). Although the MAE is less often used than the MSE, it is a more natural error measure, and has several advantages [5] [6]. It is simpler, it has a clearer meaning, and it is less sensitive to outlying values. Apparently, its lack of use is in large part motivated by the analytical difficulty associated to the absolute value [5] [6] [7].

∗E.T.S. Ingenieros de Telecomunicación, Polytechnic University of Madrid, 28040 Madrid, Spain. E-mail: lmendo@grc.ssr.upm.es.
2 Result

Let \( n_0 = \lfloor (N - 1)/p \rfloor + 1 \) and \( \alpha_N = 2 \exp(-N+1)(N-1)^{N-2}/(N-2)! \). For \( N \geq 2 \), the MAE of the estimator (1) satisfies the following.

\[
\frac{\text{E}(|\hat{p} - p|)}{p} = 2 \left( \frac{n_0 - 1}{N - 1} \right) p^{N-1}(1 - p)^{n_0 - N + 1}
\]

(2)

Furthermore, \( \frac{\text{E}(|\hat{p} - p|)}{p} = \alpha_N \).

(3)

The limit result (3) follows from the Poisson theorem [8].

Let \( i \) at (1) it stems that \( \text{E}(|\hat{p} - p|) = 2(1 - p)b_{n_0 - 1,p}(N - 1) \), which establishes (2).

The corresponding distribution function is denoted as \( F \).

The MAE is computed, using the identity \( f_N(n) = \text{Pr}[\text{I} = n] \), is

\[
f_N(n) = \left( \frac{n - 1}{N - 1} \right) p^{N}(1 - p)^{n - N}, \quad n \geq N.
\]

(5)

The MAE is computed, using the identity \( f_N(n)/(n - 1) = pf_{N-1}(n - 1)/(N - 1) \), as

\[
\begin{align*}
\text{E}(\hat{p} - p) &= \sum_{n=N}^{n_0} f_N(n) \left( \frac{N - 1}{n - 1} - p \right) - \sum_{n=n_0+1}^{\infty} f_N(n) \left( \frac{N - 1}{n - 1} - p \right) \\
&= (N - 1) \left( 2 \sum_{n=N}^{n_0} \frac{f_N(n)}{n - 1} - \sum_{n=n_0+1}^{\infty} \frac{f_N(n)}{n - 1} \right) - p \left( 2 \sum_{n=N}^{n_0} f_N(n) - \sum_{n=n_0+1}^{\infty} f_N(n) \right) \\
&= 2p[F_{N-1}(n_0 - 1) - F_N(n_0)].
\end{align*}
\]

(6)

Let \( b_{n,p}(i) \) denote the binomial probability function with parameters \( n, p \) evaluated at \( i \). Taking into account that \( F_{N-1}(n_0 - 1) = F_N(n_0) + (1 - p)b_{n_0 - 1,p}(N - 1) \), from (6) it stems that \( \text{E}(\hat{p} - p)/p = 2(1 - p)b_{n_0 - 1,p}(N - 1) \), which establishes (2).

The limit result (3) follows from the Poisson theorem [8].

Let \( I_n = ((N - 1)/n_0, (N - 1)/\lfloor (n_0 - 1)/p \rfloor) \). To prove (4), it is noted that all values \( p \in I_n \) give \( n_0 = n \). From (2),

\[
\frac{\partial \text{E}(\hat{p} - p)/p}{\partial p} = 2 \left( \frac{n_0 - 1}{N - 1} \right) p^{N-2}(1 - p)^{n_0 - N}(N - 1 - n_0p),
\]

which is negative for \( p \in I_n \). Therefore, it is sufficient to consider \( p \) restricted to the set \( S = \{p \in (0,1) \mid (N - 1)/p \in \mathbb{N}\} \); and

\[
n_0 = \frac{N - 1}{p} + 1 \quad \text{for} \ p \in S.
\]

(7)
Defining
\[ x = \frac{1}{p} \ln \lim_{p \to 0} \frac{E(|\hat{p} - p|)/p}{E(|\hat{p} - p|)/p}, \tag{8} \]
the inequality (4) is equivalent to \( x > 0 \). It follows from (2), (3), (7) and (8) that
\[ x = -\frac{1}{p} \sum_{i=1}^{N-2} \ln \left( 1 - \frac{i p}{N - 1} \right) - \frac{1}{p} \left( \frac{N - 1}{p} - N + 2 \right) \ln(1 - p) - \frac{N - 1}{p}. \tag{9} \]
Equation (9) has the same form as [3, eq. (17)]. However, the result in the cited reference which establishes the nonnegativity of \( x \) is not applicable, and a separate analysis is required. The variable \( x \) can be written as
\[ \sum_{j=0}^{\infty} x_j p^j \]
with
\[ x_j = \frac{1}{(j+1)(N-1)^{j+1}} \sum_{i=1}^{N-2} i^{j+1} + \frac{N - 1}{j + 2} - \frac{N - 2}{j + 1}. \tag{10} \]
For \( N = 2 \), \( x_j \) reduces to \( 1/(j + 2) \), and is thus positive. For \( N \geq 3 \), using the inequality \( \sum_{i=1}^{N-2} i^{j+1} > (N - 2)^{j+2}/(j + 2) \) in (10),
\[ (j + 1)(j + 2)x_j > (N - 2) \left( 1 - \frac{1}{N - 1} \right)^{j+1} + j + 3 - N. \tag{11} \]
Let \( y_j \) denote the right-hand side of (11). Computing \( \partial y_j / \partial j \) as if \( j \) were a continuous variable and using the inequality \( \ln(1 + t) > t/(1 + t) \) gives
\[ \frac{\partial y_j}{\partial j} = (N - 2) \left( 1 - \frac{1}{N - 1} \right)^{j+1} \ln \left( 1 - \frac{1}{N - 1} \right) + 1 > - \left( 1 - \frac{1}{N - 1} \right)^{j+1} + 1 > 0. \]
Thus \( y_j > y_0 = 1/(N - 1) > 0 \) for any \( j \geq 1 \), which implies that all the coefficients \( x_j \) are positive. Therefore, \( x > 0 \) for \( N \geq 2 \). This establishes (4).

The positivity of \( x_j \) for all \( j \geq 0 \), \( N \geq 2 \) implies that the values of \( xp \) corresponding to \( p = (N - 1)/n \), \( n = N - 1, N, N + 1, \ldots \) form a decreasing sequence. Together with the negative character of \( \partial E(|\hat{p} - p|)/p|\partial p \) for \( p \in I_n \), this implies that \( E(|\hat{p} - p|)/p \) is a monotonically decreasing function of \( p \).

3 Discussion
The result above allows the estimation of a probability \( p \) with a prescribed value of the normalized MAE, \( E(|\hat{p} - p|)/p \). This value is guaranteed irrespective of the unknown \( p \). For example, if a normalized MAE not exceeding 10% is desired, \( N = 65 \) suffices, according to (4).

The behaviour of the normalized MAE as a function of \( p \) is depicted in Figure 1 with solid lines. The discontinuity of the derivative at the points \( p = (N - 1)/n \), \( n \in \mathbb{N} \) (see proof of theorem) can be clearly observed, specially for low \( N \) and large \( p \). Figure 2 shows the bound (4) as a function of \( N \). The bound for the normalized root mean square error (RMSE), \( (E[|\hat{p} - p|^2])/p \), is also shown for comparison.
Figure 1: Normalized MAE as a function of $p$

Figure 2: Bounds on normalized MAE and RMSE as a function of $N$
Both error measures are seen to have the same type of behaviour, with MAE lower than RMSE.

It is interesting to compare the bound (2) with the normalized MAE resulting from a fixed sample size \( n \). This is obtained from \([9, \text{eq. (1.1)}]\) as

\[
\frac{E[|\hat{p} - p|]}{p} = 2 \left( \frac{n - 1}{N_0 - 1} \right) p^{N_0 - 1} \left( 1 - p \right)^{n - N_0 + 1}
\]

with \( N_0 = \lfloor np \rfloor + 1 \). Since the average sample size in inverse binomial sampling is \( N/p \), the comparison is restricted to probabilities \( p \) such that \( N/p \) is an integer value, and the sample size \( n \) in the fixed case is taken equal to this value. The resulting fixed-size normalized MAE is shown in Figure 1 with dashed lines. Dividing (2) by (12) with \( n = N/p \), it is easily seen that, for \( p \to 0 \), the MAE with inverse binomial sampling is asymptotically \((1 + 1/(N-1))^{-N+1}e\) times larger than the MAE with fixed size. This value is close to 1 except for very small values of \( N \). This is observed in Figure 1 which also shows that the MAE ratio is approximately maintained for all values of \( p \). It is thus concluded that, in order to guarantee a given normalized MAE, inverse binomial sampling gives an average sample size that is only slightly larger than the sample size that would be necessary in the fixed case (which is a function of the unknown \( p \)).

4 Conclusion

The MAE in inverse binomial sampling has been analyzed. It has been shown that the estimator guarantees a certain value of the normalized MAE. This result is analogous to that known for the MSE, and allows to select a value of \( N \) that meets a prescribed error level irrespective of \( p \).

References

[1] G. Prasad and A. Sahai, “Sharper variance upper bound for unbiased estimation in inverse sampling,” *Biometrika*, vol. 69, no. 1, p. 286, 1982.

[2] L. Mendo and J. M. Hernando, “A simple sequential stopping rule for Monte Carlo simulation,” *IEEE Transactions on Communications*, vol. 54, no. 2, pp. 231–241, February 2006.

[3] ———, “Improved sequential stopping rule for Monte Carlo simulation,” *IEEE Transactions on Communications*, vol. 56, no. 11, pp. 1761–1764, November 2008.

[4] ———, “Unbiased monte carlo estimator with guaranteed confidence,” in *IEEE International Workshop on Signal Processing Advances in Wireless Communications*. IEEE, 2008.

[5] C. R. Blyth, “Expected absolute error of the usual estimator of the binomial parameter,” *The American Statistician*, vol. 34, no. 3, pp. 155–157, August 1980.
[6] S. Gorard, “Revisiting a 90-year old debate: The advantages of the mean deviation,” *British Journal of Educational Studies*, vol. 53, no. 4, pp. 417–430, 2005.

[7] S. K. Bar-Lev, B. Boukai, and P. Enis, “On the mean squared error, the mean absolute error and the like,” *Communications in Statistics - Theory and Methods*, vol. 28, no. 8, pp. 1813–1822, 1999.

[8] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. McGraw-Hill, 1991.

[9] P. Diaconis and S. Zabell, “Closed form summation for classical distributions: Variations on a theme of De Moivre,” *Statistical Science*, vol. 6, no. 3, pp. 284–302, 1991.