X-560-69-151

PREPRINT

NASA TM X- 63528

ESTIMATION OF BIT PROBABILITY OF ERROR USING SYNC WORD ERROR RATE DATA

DR. CHARLES A. ROHDE

APRIL 1969

GODDARD SPACE FLIGHT CENTER
GREENBELT, MARYLAND

X-560-69-151

ESTIMATION OF BIT PROBABILITY OF ERROR USING SYNC WORD ERROR RATE DATA

Dr. Charles A. Rohde
Information Processing Division

April 1969

GODDARD SPACE FLIGHT CENTER
Greenbelt, Maryland

ESTIMATION OF BIT PROBABILITY OF ERROR USING SYNC WORD ERROR RATE DATA

Dr. Charles A. Rohde
Information Processing Division

SUMMARY

Assuming bit errors are independently distributed with a constant probability of error, p_e, it is shown that a simple estimator is highly efficient for estimation of p_e. The estimator is based on a simple function of the number of sync words containing no bit errors. The estimator is shown to be maximum likelihood, minimum chi-square and modified minimum chi-square when the quality index reported is simply the percent of frames containing zero errors. An approximate confidence interval for p_e is obtained and a determination of the number of sync words to observe in order to obtain an approximate confidence interval of fixed length is indicated. The method of scoring which can be used to obtain a more efficient estimator of p_e is described.

ESTIMATION OF BIT PROBABILITY OF ERROR USING SYNC WORD ERROR RATE DATA

A. Estimators of p_e for grouped sync words

0. Introduction

This report considers the use of sync word error rate data for determining the data quality insofar as it is reflected by the bit probability of error p_e. In the reporting of the number of errors in the sync word the current practice is to group the sync words according to the number of errors. More precisely the number of sync words with 0 errors, 1 error, . . . k errors and more than k errors are reported. Usually $k = 2$ but for generality we let k be any integer between 0 and 27. Such grouping sacrifices some information on p_e and complicates the use of statistical inference procedures.

In order to illustrate the problems and suggest possible techniques of solution we assume that N sync words have been observed and that p_e may be treated as constant over such a time span. If we assume that bit errors are independently distributed then the probability of observing X errors in a sync word follows a binomial distribution, i.e.

$$P[X \text{ errors}] = \binom{27}{X} p_e^X (1 - p_e)^{27-X}$$

where p_e is the probability of an error in a single bit and we have used a 27 bit sync word for definiteness. If for each sync word the number of errors is reported then a good estimator for p_e is

$$\sum_1^N \frac{X_i}{27N}$$

where X_i is the number of bit errors in the i^{th} sync word and N is the total number of words observed. The properties of such an estimator can be readily established including confidence intervals and tests of hypotheses.

When the data are grouped it is not obvious what to use as an estimator of p_e since the total number of errors

$$\left(\sum_{1}^{N} X_i\right)$$

is now unknown. In order to see the difficulties, let $Z_0, Z_1, \cdots, Z_{k+1}$ be defined as follows

Z_i = number of sync words having i errors $i = 0, 1, 2, \cdots, k$

Z_{k+1} = number of sync words having more than k errors.

Then the likelihood or probability of observing $z_0, z_1, \cdots, z_{k+1}$ is

$$f(\underline{z}) = \frac{N!}{\prod_{i=0}^{k+1} z_i!} \prod_{i=0}^{k+1} \left[\gamma_i(p_e)\right]^{z_i} \tag{1}$$

where

$$\gamma_i(p_e) = \binom{27}{i} p_e^i (1-p_e)^{27-i} \qquad i = 0, 1, \cdots, k$$

$$\gamma_{k+1}(p_e) = 1 - \sum_{i=0}^{k} \gamma_i(p_e)$$

It should be clear that estimation of p_e is not a simple matter.

1. Maximum likelihood estimation of p_e.

In the rest of this report we shall discuss and describe several methods for estimating p_e and some other associated inference problems. The method of maximum likelihood chooses $\hat{p}_e$ to maximize $f(\underline{z})$. Taking logs and differentiating leads to the equation

$$\sum_{i=0}^{k+1} z_i \frac{\left[\gamma_i'(\hat{p}_e)\right]}{\gamma_i(\hat{p}_e)} = 0 \tag{2}$$

which must be solved for $\hat{p}_e$. From the form of the γ_i's it is clear that some sort of iterative procedure is needed. In the special case k = 0 we have

$$\gamma_0(p_e) = (1-p_e)^{27}; \qquad \gamma_0'(p_e) = -27(1-p_e)^{26}$$

$$\gamma_1(p_e) = 1-(1-p_e)^{27}; \qquad \gamma_1'(p_e) = 27(1-p_e)^{26}$$

$$z_0 = x; \qquad z_1 = N - x$$

where x is the number of sync words having zero errors. Then Equation 2 becomes

$$\frac{-27x(1-\hat{p}_e)^{26}}{(1-\hat{p}_e)^{27}} + \frac{(N-x)\,27(1-\hat{p}_e)^{26}}{1-(1-\hat{p}_e)^{27}} = 0$$

so that

$$\hat{p}_e = 1 - \left(\frac{x}{N}\right)^{1/27}$$

2. Minimum chi-square estimation of p_e.

Another method of estimation is the method of minimum chi-square which chooses $\hat{p}_e$ to minimize

$$\underset{\sim}{X}^2 = \sum_{i=0}^{k+1} \frac{\left[z_i - N\gamma_i(\hat{p}_e)\right]^2}{N\gamma_i(\hat{p}_e)} = \sum_{i=0}^{k+1} \frac{z_i^2}{N\gamma_i(\hat{p}_e)} - N$$

Differentiating yields the equation

$$\sum_{i=0}^{k+1} \frac{z_i^2\,\gamma_i'(\hat{p}_e)}{N\left[\gamma_i(\hat{p}_e)\right]^2} = 0$$

For $k = 0$ the equation becomes

$$\frac{x^2\left[-27(1-p_e)^{26}\right]}{N\left[(1-p_e)^{27}\right]^2} - \frac{(N-x)^2\, 27(1-p_e)^{26}}{N\left[1-(1-p_e)^{27}\right]^2} = 0$$

or

$$\frac{x}{(1-p_e)^{27}} = \frac{N-x}{\left[1-(1-p_e)^{27}\right]}$$

so that

$$\hat{p}_e = 1 - \left(\frac{x}{N}\right)^{1/27}$$

The general case again requires iteration and <u>need not</u> yield the same estimates as maximum likelihood even though agreement is obtained for $k = 0$.

3. Modified minimum chi-square estimation of p_e.

The method of modified minimum chi-square is often simpler and is based on minimizing

$$(\underset{\sim}{X}')^2 = N \sum_{i=0}^{k+1} \frac{\gamma_i^2(\hat{p}_e)}{z_i} - N$$

which is simply the expression for minimum $\underset{\sim}{X}^2$ with $N\gamma_i(\hat{p}_e)$ in the denominator replaced by z_i. Taking the derivative with respect to p_e yields the equation

$$2N \sum_{i=0}^{k+1} \frac{\gamma_i(\hat{p}_e)\,\gamma_i'(\hat{p}_e)}{z_i} = 0$$

to be solved for $\hat{p}_e$. If $k = 0$ the equation becomes

$$0 = 2N \left\{ \frac{-27(1-p_e)^{27}(1-p_e)^{26}}{x} + \frac{27\left[1-(1-p_e)^{27}\right](1-p_e)^{26}}{N-x} \right\}$$

Hence

$$\hat{p}_e = 1 - \left(\frac{x}{N}\right)^{1/27}$$

as before. Once again the equations require iterative techniques in general and need not yield the same estimates as maximum likelihood or minimum chi-square.

Since $\hat{p}_e$ cannot be solved as an explicit function of the data (the z_i's) except in special cases the variance of $\hat{p}_e$ and confidence interval statements about p_e cannot be easily determined. Asymptotic statements concerning p_e can be made, however, and since N is frequently large in satellite data situations we may expect the results to be reliable.

The methods of maximum likelihood, minimum chi-square and modified minimum chi-square applied to estimation of p_e all have the same asymptotic properties. In particular the asymptotic variance of $\hat{p}_e$ is the reciprocal of

$$E\left[\frac{\partial \log f(\underline{z})}{\partial p_e}\right]^2$$

Since

$$\log f(\underline{z}) = \text{constant} + \sum_{i=0}^{k+1} z_i \log \gamma_i (p_e)$$

we have

$$\frac{\partial \log f(\underline{z})}{\partial p_e} = \sum_{i=0}^{k+1} \left[\frac{\gamma_i'(p_e)}{\gamma_i(p_e)}\right] z_i$$

Since the z_i are multinomial random variables we have

$$E[z_i \; z_j] = \begin{cases} \text{Var } z_i + [Ez_i]^2 & i = j \\ \text{cov}(z_i, z_j) + [Ez_i][Ez_j] & i \neq j \end{cases}$$

$$= \begin{cases} N\gamma_i(p_e)[1-\gamma_i(p_e)] + [N\gamma_i(p_e)]^2 & i = j \\ -N\gamma_i(p_e)\gamma_j(p_e) + N^2\gamma_i(p_e)\gamma_j(p_e) & i \neq j \end{cases}$$

Hence

$$E\left[\frac{\partial \log f(\underline{z})}{\partial p_e}\right]^2 = \sum_{i=0}^{k+1}\sum_{j=0}^{k+1} \frac{\gamma_i'(p_e)}{\gamma_i(p_e)} \frac{\gamma_j'(p_e)}{\gamma_j(p_e)} Ez_i \; z_j$$

$$= \sum_{i=0}^{k+1}\left[\frac{\gamma_i'(p_e)}{\gamma_i(p_e)}\right]^2 \left[N\gamma_i(p_e)[1-\gamma_i(p_e)] + N^2\gamma_i(p_e)^2\right]$$

$$+ \sum_{\substack{i=0 \\ i\neq j}}^{k+1}\sum_{j=0}^{k+1}\left[\frac{\gamma_i'(p_e)\gamma_j'(p_e)}{\gamma_i(p_e)\gamma_j(p_e)}\right]\left[-N\gamma_i(p_e)\gamma_j(p_e) + N^2\gamma_i(p_e)\gamma_j(p_e)\right]$$

$$= \left\{\sum_{i=0}^{k+1}[\gamma_i'(p_e)]^2\right\} N(N-1) + N\sum_{i=0}^{k+1}\frac{[\gamma_i'(p_e)]^2}{\gamma_i(p_e)}$$

$$+ \left\{\sum_{\substack{i=0 \\ i\neq j}}^{k+1}\sum_{j=0}^{k+1}\gamma_i'(p_e)\gamma_j'(p_e)\right\} N(N-1)$$

$$= N(N-1)\left[\sum_{i=0}^{k+1}\gamma_i'(p_e)\right]^2 + N\sum_{i=0}^{k+1}\frac{[\gamma_i'(p_e)]^2}{\gamma_i(p_e)}$$

$$= N \sum_{i=0}^{k+1} \frac{[\gamma_i'(p_e)]^2}{\gamma_i(p_e)} \qquad \left(\text{since } \sum_{i=0}^{k+1} \gamma_i'(p_e) = 0 \right)$$

so that the asymptotic variance of $\hat{p}_e$ is

$$\frac{1}{N \sum_{i=0}^{k+1} \left\{ \frac{[\gamma_i'(p_e)]^2}{\gamma_i(p_e)} \right\}} \tag{3}$$

Since both γ_i' and γ_i are polynomials in p_e the above expression can easily be tabulated for values of p_e.

Most satellite work yields values of p_e in the range $0 \leq p_e \leq .01$ so tabulation in this range should suffice.

B. **Estimators of p_e for ungrouped sync words**

It is of some interest to compare the asymptotic variance of $\hat{p}_e$ above with that obtained if the data were not grouped. This gives some insight into the information lost by grouping. If the data were not grouped and Y represents the total number of sync word errors then Y has a binomial distribution with parameters 27N and p_e, i.e.

$$P[Y = y] = \binom{27N}{y} p_e^y (1 - p_e)^{27N-y}$$

so that $\hat{p}_e = y/27N$ and the variance of $\hat{p}_e$ is

$$\frac{p_e(1 - p_e)}{27N}$$

Thus the ratio of asymptotic variances is

$$\frac{\dfrac{p_e(1-p_e)}{27N}}{\dfrac{1}{N \sum_{i=0}^{k+1} \left\{ \frac{[\gamma_i'(p_e)]^2}{\gamma_i(p_e)} \right\}}} = \left\{ \sum_{i=0}^{k+1} \frac{[\gamma_i'(p_e)]^2}{\gamma_i(p_e)} \right\} \frac{p_e(1 - p_e)}{27}$$

Since γ_i' and γ_i are polynomials in p_e the above ratio could be easily tabulated for values of p_e in the ranges customarily found in telemetry data.

Since

$$\sum_{i=0}^{k+1} \frac{\left[\gamma_i'(p_e)\right]^2}{\gamma_i(p_e)}$$

is monotone increasing in k it is clear that the most information is lost when $k = 0$ (this is also intuitively obvious). For the case $k = 0$ we have (see (3))

$$\gamma_0'(p_e) = -27(1-p_e)^{26} \qquad \gamma_0(p_e) = (1-p_e)^{27}$$

$$\gamma_1'(p_e) = 27(1-p_e)^{26} \qquad \gamma_1(p_e) = 1-(1-p_e)^{27}$$

so that

$$\sum_{i=0}^{1} \frac{\left[\gamma_i'(p_e)\right]^2}{\gamma_i(p_e)} = \frac{27^2\left[(1-p_e)^{26}\right]^2}{(1-p_e)^{27}} + \frac{27^2\left[(1-p_e)^{26}\right]^2}{1-(1-p_e)^{27}}$$

$$= 27^2\left[(1-p_e)^{26}\right]^2 \left[\frac{1}{1-(1-p_e)^{27}} + \frac{1}{(1-p_e)^{27}}\right]$$

$$= \frac{27^2(1-p_e)^{26}}{(1-p_e)\left[1-(1-p_e)^{27}\right]}$$

Thus the ratio of asymptotic variances is

$$\left\{\frac{27^2(1-p_e)^{26}}{(1-p_e)\left[1-(1-p_e)^{27}\right]}\right\} \frac{p_e(1-p_e)}{27} = \frac{27(1-p_e)^{26}\,p_e}{\left[1-(1-p_e)^{27}\right]}$$

If p_e is very small (say $\leq .01$) then

$$(1-p_e)^{26} \cong 1 - 26\,p_e \qquad (1-p_e)^{27} \cong 1-27\,p_e$$

Hence the ratio of asymptotic variances is approximately

$$\frac{27\left(1-26\,p_e\right)p_e}{1-\left(1-27\,p_e\right)} \cong 1-26\,p_e$$

which for $p_e = .01$ is .74, for $p_e = .001$ is .974.

C. Estimation of p_e when p_e is small

1. The estimator and a large sample confidence interval.

A recommendation for practical usage is clear. If p_e is expected to be small (say $\lesssim .01$) little efficiency is lost by reporting only the number X of sync word frames with zero errors in N observed sync words. The estimator for p_e is then

$$\hat{p}_e = 1-\left(\frac{X}{N}\right)^{1/27} \tag{4}$$

and $\hat{p}_e$ may be treated as approximately normal with mean p_e and variance

$$\frac{\left(1-p_e\right)\left[1-\left(1-p_e\right)^{27}\right]}{N27^2\left(1-p_e\right)^{26}} \cong \frac{\left(1-p_e\right)27\,p_e}{27^2\,N\left[1-26\,p_e\right]}$$

$$= \frac{p_e\left(1-p_e\right)}{27N\left[1-26\,p_e\right]}$$

It follows that an approximate $1-\alpha$ level confidence interval for p_e is

$$\hat{p}_e \pm z_{\alpha/2}\left\{\frac{\hat{p}_e\left(1-\hat{p}_e\right)}{27N\left[1-26\,\hat{p}_e\right]}\right\}^{1/2}$$

where

$$\int_{z_{\alpha/2}}^{\infty}\frac{e^{-u^2/2}}{\sqrt{2\pi}}\,du = 1-\alpha$$

can be obtained from tables of the standard normal distribution. Since $\left[\hat{p}_e\left(1-\hat{p}_e\right)/1-26\,\hat{p}_e\right]$ is a decreasing function of p_e for small p_e we have that

$$\frac{\hat{p}_e\left(1-\hat{p}_e\right)}{1-26\,\hat{p}_e} \leq \frac{.01(.99)}{.74} = .013$$

so that the interval

$$\hat{p}_e \pm z_{\alpha/2}\left[\frac{.013}{27N}\right]^{1/2}$$

is at least a $1-\alpha$ level confidence interval for p_e. If we desire a confidence interval of length 2d then we can set

$$d = z_{\alpha/2}\left[\frac{.013}{27N}\right]^{1/2}$$

so that

$$d^2 = z_{\alpha/2}^2\,\frac{.013}{27N}$$

Hence

$$N \geq \frac{z_{\alpha/2}^2\;.013}{d^2\,27}$$

will yield an approximate $1-\alpha$ confidence interval of length 2d.

2. Asymptotic variances of more accuracy.

If asymptotic variances are desired to more accuracy we may simply observe that

$$\gamma_i\left(p_e\right) = \binom{27}{i} p_e^{\,i}\left(1-p_e\right)^{27-i} \qquad i = 0, 1, 2, \cdots, k$$

so

$$\gamma_i'(p_e) = \binom{27}{i}\left\{ i p_e^{i-1}(1-p_e)^{27-i} + (27-i)\, p_e^{i}(1-p_e)^{27-i-1}\right\}$$

$$= \binom{27}{i} p_e^{i-1}(1-p_e)^{27-i-1}\left[i(1-p_e) - (27-i)\, p_e\right]$$

$$= \frac{\binom{27}{i} p_e^{i}(1-p_e)^{27-i}}{p_e(1-p_e)}\,(i - 27\, p_e)$$

for $i = 1, 2, \cdots, k$. For $k+1$

$$\gamma_{k+1}(p_e) = 1 - \sum_{i=0}^{k}\binom{27}{i} p_e^{i}(1-p_e)^{27-i}$$

so

$$\gamma_{k+1}'(p_e) = -\sum_{i=0}^{k}\frac{\binom{27}{i} p_e^{i}(1-p_e)^{27-i}(i-27\, p_e)}{p_e(1-p_e)}$$

Now define

$$P(p_e;\, k) = \sum_{i=0}^{k}\binom{27}{i} p_e^{i}(1-p_e)^{27-i}$$

$$M(p_e;\, k) = \sum_{i=0}^{k}\binom{27}{i} p_e^{i}(1-p_e)^{27-i}(i-27\, p_e)$$

$$V(p_e;\, k) = \sum_{i=0}^{k}\binom{27}{i} p_e^{i}(1-p_e)^{27-i}(i-27\, p_e)^2$$

Then for $i = 0, 1, 2, \cdots, k$ we have

$$\frac{\left[\gamma_i'(p_e)\right]^2}{\gamma_i(p_e)} = \frac{\binom{27}{i}^2 \left[p_e^i (1-p_e)^{27-i}\right]^2 \left[(i - 27p_e)\right]^2 / p_e^2}{\binom{27}{i} p_e^i (1-p_e)^{27-i}}$$

$$= \frac{\binom{27}{i} p_e^i (1-p_e)^{27-i}}{p_e^2 (1-p_e)^2} (i - 27 p_e)^2$$

so

$$\sum_{i=0}^{k} \frac{\left[\gamma_i'(p_e)\right]^2}{\gamma_i(p_e)} = \frac{1}{p_e^2 (1-p_e)^2} V(p_e; k)$$

Also

$$\frac{\left[\gamma_{k+1}'(p_e)\right]^2}{\gamma_{k+1}(p_e)} = \frac{\left[M(p_e; k)\right]^2 / p_e^2 (1-p_e)^2}{1 - P(p_e; k)}$$

Hence

$$\sum_{i=0}^{k+1} \frac{\left[\gamma_i'(p_e)\right]^2}{\gamma_i(p_e)} = \frac{1}{p_e^2 (1-p_e)^2} \left[V(p_e; k) + \frac{\left[M(p_e; k)\right]^2}{1 - P(p_e; k)}\right]$$

$$= \frac{1}{p_e^2 (1-p_e)^2} \frac{\left\{V(p_e; k)\left[1 - P(p_e; k)\right] + \left[M(p_e; k)\right]^2\right\}}{1 - P(p_e; k)}$$

Thus the ratio of asymptotic variances is

$$\frac{\left\{V(p_e; k)\left[1 - P(p_e; k)\right] + \left[M(p_e; k)\right]^2\right\}}{27 p_e (1-p_e)\left[1 - P(p_e; k)\right]}$$

3. Recursive estimation of p_e.

Let

$$S(\underline{Z}, p_e) = \sum_{i=0}^{k+1} \frac{\gamma_i'(p_e)}{\gamma_i(p_e)} Z_i$$

Then we can write

$$S(\underline{Z}, p_e) = \sum_{i=0}^{k} \frac{Z_i\,(i-27)}{p_e\,(1-p_e)} + Z_{k+1} \frac{M(p_e; k)}{p_e\,(1-p_e)\left[1-P(p_e; k)\right]}$$

or

$$S(\underline{Z}, p_e) = \frac{1}{p_e\,(1-p_e)} \left\{ \sum_{i=0}^{k} Z_i\,(i-27p_e) + Z_{k+1} \frac{M(p_e; k)}{p_e\,(1-p_e)\; 1-P(p_e; k)} \right\}$$

Now

$$\sum_{i=0}^{k} i\,Z_i = \text{total number of errors in the first k groups of sync words} = Z_1^*$$

$$\sum_{i=0}^{k} Z_i = \text{total number of sync words with less than k + 1 errors} = Z_2^*$$

Hence

$$S(\underline{Z}, p_e) = \frac{1}{p_e\,(1-p_e)} \left\{ Z_1^* - 27p_e\, Z_2^* + Z_{k+1} \frac{M(p_e; k)}{\left[1-P(p_e; k)\right]} \right\}$$

Also

$$NI(p_e) = N \sum_{i=1}^{k+1} \frac{[\gamma_i'(p_e)]^2}{\gamma_i(p_e)} = \frac{N}{p_e^2(1-p_e)^2} \frac{\{V(p_e; k)[1-P(p_e; k)] + [M(p_e; k)]^2\}}{[1-P(p_e; k)]}$$

If p_e^0 is a trial value of p_e then a better approximation to $\hat{p}_e$ is $p_e' = \hat{p}_e + \delta p_e^0$ where

$$\delta p_e^0 = S(\underline{Z}; p_e^0)/NI(p_e^0)$$

The process may now be repeated using p_e' to get

$$\hat{p}_e^2 = p_e' + \delta p_e'$$

The process may be stopped whenever the correction term is sufficiently small. The practical steps may now be indicated as follows:

(1) Compute

N = total number of sync words observed

Z_1^* = # of errors in sync words with 0, 1, 2, $\cdots$, k errors

Z_2^* = # of sync words with less than k + 1 errors

$Z_{k+1} = N - Z_2^*$

(2) Select a trial value; say $\hat{p}_e^0 = 1 - (X/N)^{1/27}$ where X is the number of sync words with zero errors.

(3) Compute $1/\hat{p}_e^0(1-\hat{p}_e^0)$, $M(\hat{p}_e^0; k)/1 - P(\hat{p}_e^0; k)$ and

$$S(\hat{p}_e^0, \underline{Z}) = \frac{1}{\hat{p}_e^0(1-\hat{p}_e^0)} \left\{ Z_1^* - 27\hat{p}_e^0 Z_2^* + Z_{k+1} \frac{M(\hat{p}_e^0; k)}{[1 - P(\hat{p}_e^0, k)]} \right\}$$

(4) Compute $1/I(\hat{p}_e^{\,0})$ and

$$\delta\hat{p}_e^{\,0} = \frac{S(\hat{p}_e ; \underline{Z})}{NI(\hat{p}_e^{\,0})}$$

(5) Compute $\hat{p}_e^{\,0} + \delta\hat{p}_e^{\,0} = \hat{p}_e{}'$ and return to (2), replacing $\hat{p}_e^{\,0}$ by $\hat{p}_e{}'$. Stop when $\delta\hat{p}_e^{\,i}$ is sufficiently small.

The above iterative procedure is easily programmed for routine estimation problems of the type discussed in this report.

<u>Reference:</u>

C. R. Rao - Linear Statistical Inference and its Applications. John Wiley.