THE UPPER BOUNDS OF THE CONFIDENCE INTERVALS OF BIT ERROR PROBABILITIES BASED ON A MARKOV CHAIN BIT ERROR MODEL

M. MIZUKI
ITT-Federal Electric Corp.
Vandenberg A.F.B., California

Summary Confidence intervals for the bit error probability of an actual PCM telemetry data can be determined based on the analysis of received redundant bits. The procedure usually requires the assumption of independence of bit errors. However, bit errors may occur in clusters under various conditions of multipath, injection of non-thermal noise of long duration, and bit jitters. As a representation of bit errors in clusters, a Markov chain model is introduced. Some results on the confidence interval of bit error probability are obtained as functions of a Markovian parameter, which designates the degree of departure from the binomial model. The computations are quite laborious compared to the case of the binomial model. This paper gives step-by-step instructions for computing the probabilities that r error bits occur among mn received bits which can then be used for the derivation of the confidence interval.

Introduction The bit error rate of an actual PCM telemetry data can be obtained by examining received redundant bits either at the standard detection threshold level or, in the case of slow rate transmissions, using monitoring devices with varied detection threshold levels [1,2]. Under the assumption of the independence of bit errors, the confidence interval for the bit error probability can be determined in an elementary manner, cf. the confidence intervals for binomial or Poisson parameters [3]. It can be shown that the Bayesian posterior distribution of the bit error probability is not markedly different from the prior distribution adding in fact no knowledge due to the observation of redundant bits [4].

Bit errors in clusters may be observed under possible conditions of multipath, injection of non-thermal noise of long durations, bit jitters, etc. One of the probability models suitable for describing bit errors of this nature is the Markov chain model which is investigated in this paper. In order to establish the confidence intervals similar to those obtained under the binomial assumption, it is necessary to introduce a Markovian parameter which designates the degree of departure from the binomial model. This approach must be employed here, because no adequate methods are available for

estimating the transition probabilities under the stipulated conditions of having sync bits separated by information bits.

1. PRELIMINARIES It will be assumed as before, cf. [4], that (i) all the formats of the transmitted sync bits are known at the decommutation station, (ii), the errors of received word sync bits are identified and enumerated over n consecutive words.

The Markov chain model consists of two states C for "correct" and E for "error" of the received sync bits. Let s_k denote the realized state of the k-th received bit, i.e., s_k = C or E depending on either correct or in error. The initial probabilities are denoted by

$$(1.1) \qquad p_c = \text{Prob}\{s_1 = C\}$$
$$p_e = \text{Prob}\{s_1 = E\} = 1 - p_c \quad .$$

The transition probabilities are assumed to be constant for all k, k=l, 2.... and are denoted by

$$(1.2) \qquad p_{cc} = \text{Prob}\{s_{k+1} = C | s_k = C\}$$
$$p_{ce} = \text{Prob}\{s_{k+1} = E | s_k = C\} = 1 - p_{cc}$$
$$p_{ec} = \text{Prob}\{s_{k+1} = C | s_k = E\}$$
$$p_{ee} = \text{Prob}\{s_{k+1} = E | s_k = E\} = 1 - p_{ec} \quad .$$

A Markov chain sequence of length n is designated by an n-tuple, denoted by $(s_1,...,s_n)$, of the realized states s_k's, k=1,...n. From (1.1) and (1.2) the probability that an n-tuple $(s_1,...,s_n)$ is realized is given by the produce of transition probabilities

$$(1.3) \qquad P(s_1,...,s_n) = p_{s_1} \prod_{k=1}^{n-1} p_{s_k\, s_{k+1}} \quad .$$

Denote the ensemble of sequences of states C and E in each of which exactly r E's and (n-r) C's appear by S(r). The probability that exactly r E's appear in a Markov chain sequence of length n, denoted by P(n,r), is then given by the sum of (1.3) over all sequences in the ensemble S(r). For small values of r, P(n,r) are given by

(1.4)
$$P(n,0) = p_c p_{cc}^{n-1} ,$$
$$P(n,1) = p_c p_{cc}^{n-2} p_{ce} + (n-2) p_c p_{cc}^{n-3} p_{ce} p_{ec} + p_e p_{ec} p_{cc}^{n-2} ,$$
$$P(n,2) = p_c p_{cc}^{n-3} p_{ce} p_{ee} + (n-3) p_c p_{cc}^{n-4} p_{ce}^2 p_{ec}$$
$$+ (n-3) p_c p_{cc}^{n-4} p_{ce} p_{ee} p_{ec}$$
$$+ \left[\binom{n-2}{2} - (n-3)\right] p_c p_{cc}^{n-5} p_{ce}^2 p_{ec}^2$$
$$+ p_e p_{ee} p_{ec} p_{cc}^{n-3} + p_e p_{ec} p_{cc}^{n-3} p_{ce}$$
$$+ (n-3) p_e p_{ec}^2 p_{cc}^{n-4} p_{ce} ,$$

etc. The expressions of P(n,r) become complex for higher values of r. An extensive analysis of P(n,r) is given in the Appendix.

2. STEADY STATE PROBABILITIES TO BE USED IN THE MODEL. Let the initial probabilities and transition probabilities be denoted in matrix notation by

(2.1)
$$p^{(1)} = (p_c^{(1)}, p_e^{(1)}) \quad \text{and} \quad P = \begin{bmatrix} p_{cc} & p_{ce} \\ p_{ec} & p_{ee} \end{bmatrix} .$$

The probability vector of the n-th received bit, denoted by $p^{(n)} = (p_c^{(n)}, p_e^{(n)})$, is then given by

(2.2)
$$p^{(n)} = (p_c^{(n)}, p_e^{(n)}) = p^{(1)} p^{n-1}$$
$$= (p_c^{(1)}, p_e^{(1)}) \begin{bmatrix} p_{cc} & p_{ce} \\ p_{ec} & p_{ee} \end{bmatrix}^{n-1} .$$

The two-state Markov chain with states C and E of interest is aperiodic and irreducible because $0<p_{cc},p_{ee}<1$.

Then the probability vector $p^{(n)}$ converges to a limit vector as n approaches infinity. The limit vector of $p^{(n)}$ is called the steady state probability vector and is denoted by $p^* = (p_c^*, p_e^*)$. The values of p_c^* and p_e^* are obtained by solving the one-step transition

(2.3)
$$(p_c^*, p_e^*) = (p_c^*, p_e^*) \begin{bmatrix} p_{cc} & p_{ce} \\ p_{ec} & p_{ee} \end{bmatrix} .$$

It is easily shown that

(2.4)
$$p_c^* = \frac{p_{ec}}{p_{ce} + p_{ec}}$$

$$p_e^* = \frac{p_{ce}}{p_{ce} + p_{ec}} \quad .$$

Actually $P^{(n)}$ converges fairly rapidly to p* if $p^{(1)}$ is not far from p*. In the present problem of bit error analysis $p_e^{(1)}$ is going to be set equal to the bit error probability of the binomial model, and therefore $P_e^{(1)}$ isexpected to be very close to p_e^* .

3. USE OF WORD SYNC BITS. In the binomial bit error model the bits in the sync words can be used as redundant bits in addition to the word sync bits. In fact, under the assumption of such independent bit errors, any redundant bits of known configurations can be used for the determination of the confidence intervals regardless of the positions of these redundant bits anywhere within each frame. However, the same is not true in the case of the Markov chain bit error model. In order to obtain symmetry and simplify computations involved, only the word sync bits can be used in the subsequent analysis.

Suppose each word consists of h information bits and m word sync bits. Suppose the initial probability vector for the first bit of a train of telemetry data is $p^{(1)} = (q,p)$, where $q=p_c^{(1)}$ and $p=p_e^{(1)}$. Let the transition matrix be written as

$$P = \begin{bmatrix} p_{cc} & p_{ce} \\ p_{ec} & p_{ee} \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

for the ease of notation. Then, the state probabilities of the i-th of m bits ($1 \le i \le m$) of the j-th word sync bit group (j=1,2,...) is given by

$$(q,p) \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{(j-1)(h+m)+h+i-1}$$

However, if the initial probability vector $p^{(1)}$ is set equal to the steady state probability vector, p* = (q*,p*), then, the above reduces to

$$(q^*,p^*) \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{i-1}$$

for $1 \le i \le m$ for all j. This is a great simplification of the algebra required for the determination of the state probabilities. The only requirement is that the initial

probabilities are equal to the steady state probabilities discussed in Section 2. Notice also that j can be any positive integer without any restrictions.

The proposed analysis will take the following steps:

(1) The state probabilities of the i-th bit, i=l,2,...,m, of each m-bit word sync group is given by the steady state probabilities

$$(q^*, p^*) \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{i-1}$$

(2) The probability of r' bits in error among m received word sync bits is computed using P(m,r') as given in Section 1.

(3) The probabilities of r bits in error over mn word sync bits where n is the number of received words are computed as the sum of the probabilities of all the realizable sequences with r bits in error over mn word sync bits by combining the values of P(m,r').

Since each word sync bit group is separated by h information bits, and since the steady state probabilities are used, the probability P(m,r') of the k-th word sync bit group is treated independently of P(m,r") of the (k+1)st word sync bit group. This results in the symmetry of P(m,r)'s and substantial simplification of the overall computations.

4. PROBABILITIES OF r BITS IN ERROR OVER mn WORD SYNC BITS.

The initial task is to express r as a sum of n non-negative integers each of which is less than or equal to m. By denoting the k-th such integer by i_k, $0 \le i_k \le m$, k=l,...,n, r may be expressed as

(4.1) $$r = i_1 + i_2 \ldots + i_n \ .$$

Suppose there are c_0 zero's, c_1 one's, ..., and c_m m's in this particular sum, then clearly

(4.2) $$r = \sum_{j=0}^{m} j c_j \ .$$

Let $P(m,i_k)$ denote the probability that i_k bits are in error among m bits given the steady state probabilities for the first bit. Then the probability that there are r bits in error over mn word sync bits with the configuration of $(i_1, i_2,...,i_n)$ is given by

(4.3) $$Pr\{i_1,i_2,\ldots,i_n\} = \prod_{k=1}^{n} P(m,i_k) \quad .$$

The probability of r bits in error is given by summing over all distinct configurations of $(i_1,i_2,\ldots,i_n)$, i.e.,

(4.4) $$Pr\{r;mn\} = \sum^{*} \prod_{k=1}^{n} P(m,i_k)$$

where the summation (*) is taken over all distinct $(i_1,i_2,\ldots,i_n)$'ss satisfying $\Sigma i_k = r$. The actual computation of (4.4) can be simplified using the following conventions due to the symmetry of $P(m,i_k)$'s

For the ease of discussion, consider each word sync group consisting of m bits as a box and consider each bit in error as a ball to be placed in these boxes. Then, there are n boxes of interest into which r balls must be placed with the restriction that no boxes should contain more than m balls. The k-th box will contain the i_k balls, and there will be c_j boxes in which j balls are found according to the above formulation.

STEP 1. Decompose r into a sum of non-increasing sequence of positive integers, each of which is less than or equal to m. For instance, for r=6, m=3, n≥6,

$$6 = 3+3 = 3+2+1 = 3+1+1+1$$
$$= 2+2+2 = 2+2+1+1$$
$$= 2+1+1+1+1 = 1+1+1+1+1+1$$

STEP 2. For each sum, enumerate the number of terms of the same integer, $c_m,c_{m-1},\ldots,c_1c_0$, e.g.,

$6 = 3+3$	$c_3 = 2$	$c_2 = 0$	$c_1 = 0$	$c_0 = n-2$
$= 3+2+1$	$c_3 = 1$	$c_2 = 1$	$c_1 = 1$	$c_0 = n-3$
$= 3+1+1+1$	$c_3 = 1$	$c_2 = 0$	$c_1 = 3$	$c_0 = n-4$
$= 2+2+2$	$c_3 = 0$	$c_2 = 3$	$c_1 = 0$	$c_0 = n-3$
$= 2+2+1+1$	$c_3 = 0$	$c_2 = 2$	$c_1 = 2$	$c_0 = n-4$
$= 2+1+1+1+1$	$c_3 = 0$	$c_2 = 1$	$c_1 = 4$	$c_0 = n-5$
$= 1+1+1+1+1+1$	$c_3 = 0$	$c_2 = 0$	$c_1 = 6$	$c_0 = n-6$

STEP 3. Compute the probability of r bits in error with the first c_m boxes with m balls, next c_{m-1} boxes with (m-1) balls, ..., next c_2 boxes with two balls, next c_1 boxes with one ball each, and the remaining $n-c_m,c_{m-1}-\ldots-c_1 = c_0$ boxes with no balls as

(4.5) $$\Pr\{c_m, c_{m-1}, \ldots, c_1, c_0\} = \prod_{j=0}^{m} P(m,j)^{c_j}$$

STEP 4. Compute the number of different ways of choosing $(n-c_0)$ boxes in such a way that c_j boxes contain j balls, j=l,...,m; i.e., this is given by

$$\binom{n}{n-c_0} \frac{(n-c_0)!}{\prod_{j=1}^{m} c_j!} .$$

STEP 5. The probability that r balls are placed in various combination of j balls in c_j boxes is given by

(4.6) $$\Pr\{r; c_j, j=1,\ldots,m\} = \binom{n}{n-c_0} \frac{(n-c_0)!}{\prod_{j=1}^{m} c_j!} \prod_{j=0}^{m} P(m,j)^{c_j} .$$

STEP 6. The probability that r balls are placed in n boxes is then given as the sum of (4.6) for all ossible combinations of c_j's,

(4.7) $$\Pr\{r; mn\} = \sum_{**} \Pr\{r; c_j, j=1,\ldots,m\}$$

where the swmation (**) is taken over all combinations of $(c_0,\ldots,c_m)$ enumerated in Step 2.

5. THE UPPER BOUNDS OF CONFIDENCE INTERVALS. The confidence intervals must be constructed for the Markov chain model in some meaningful manner that can be shown as a natural logical extension of the binomial model. For this purpose, the transition matrix P is designated by

(5.1) $$P = \begin{bmatrix} 1-\pi & \pi \\ 1-\pi-\epsilon & \pi+\epsilon \end{bmatrix}$$

where $\epsilon(>0)$ is a parameter which reflects the departure of P from the binomial case. The Markov chain model of (5.1) reduces to the binomial model when $\epsilon = 0$. The upper bounds of the confidence intervals for π can then be determined for given values of ϵ. As stated earlier in the introduction, no adequate methods based on the observable transition numbers are available for the estimation of true transition probabilities of P due to the fact that no more than m consecutive redundant bits can be observed over data

bits except for the frame sync words. Moreover, the frame sync words are not used in the present investigation for the sake of simplifying computations.

From (5.1) and (2.4), the steady state probabilities are given by

(5.2) $$q^* = \frac{1-\pi-\epsilon}{(1-\pi-\epsilon) + \pi} = \frac{1-\pi-\epsilon}{1-\epsilon}$$

$$p^* = \frac{\pi}{1-\epsilon}$$

Notice now that p*, q*, and P are all functions of π and ϵ. Denote the probability Pr{r;mn} for a given value of ϵ by Pr{r;mn,ϵ}, which is now a function of π alone. Then, the upper bound of the confidence interval of level 1-α is given by π satisfying

(5.3) $$\sum_{i=0}^{r} Pr\{i;mn,\epsilon\} \leq \alpha/2$$

if r error bits are found among mn received bits given for a fixed value of ð ϵE. In particular if ϵ=0, the above (5.3) is equal to the equation used for the determination of the upper bounds of confidence intervals for the binomial bit error model. Since the value of π treated in bit error models are usually very small, one-sided confidence intervals of the form $(0,\pi_0)$ with confidence 1-α may be considered. Under such alternative formulation (5.3) may be replaced by

(5.3)' $$\sum_{i=0}^{r} Pr\{i;mn,\epsilon\} \leq \alpha \quad .$$

6. EXAMPLES. Case r=0: Suppose m=3 and n is arbitrary. Consider the simplest case of r=0. Then, $c_0 = n$, $c_1 = c_2 = c_3 = 0$. Since

(6.1) $$P(3,0) = q^*a^2 = \frac{1-\pi-\epsilon}{1-\epsilon} (1-\pi)^2 \quad .$$

Using (4.6) and (4.7)

(6.2) $$P\{0;3n,\epsilon\} = \left[\frac{(1-\pi-\epsilon)(1-\pi)^2}{1-\epsilon}\right]^n$$

By equating this to α and taking the n-th root,

(6.3) $$(1-\pi-\epsilon)(1-\pi)^2 = \alpha^{\frac{1}{n}} (1-\epsilon) \quad .$$

This can be rewritten as

(6.4)
$$1-\varepsilon-(3-2\varepsilon)\pi + (3-\varepsilon)\pi^2 - \pi^3 = \alpha^{\frac{1}{n}} (1-\varepsilon) \quad .$$

In particular, if π is expected to be very small, the higher order terms of π w can be neglected. Using the linear term only, the upper bound of the one-sided confidence interval of level $1-\alpha$ is given by

(6.5)
$$\pi_0 = \frac{(1-\alpha^{\frac{1}{n}})(1-\varepsilon)}{3-2\varepsilon}$$

Case $r \neq 0$: For m=3, it is only necessary to compute P(3,1), P 3,2 , and P(3,3). From the equations given in (1.4),

(6.6)
$$\begin{aligned} P(3,1) &= q^*ab + q^*bc + p^*ca \\ &= \frac{1-\pi-\varepsilon}{1-\varepsilon}\{(1-\pi)\pi + \pi(1-\pi-\varepsilon)\} + \frac{\pi}{1-\varepsilon}(1-\pi-\varepsilon)(1-\pi) \\ &= \frac{(1-\pi-\varepsilon)}{1-\varepsilon}\{3(1-\pi)-\varepsilon\} \quad , \\ P(3,2) &= q^*bd + p^*cb + p^*dc \\ &= \frac{1-\pi-\varepsilon}{1-\varepsilon}\pi(\pi+\varepsilon) + \frac{\pi}{1-\varepsilon}\{(1-\pi-\varepsilon)\pi + (\pi+\varepsilon)(1-\pi-\varepsilon)\} \\ &= \frac{\pi(1-\pi-\varepsilon)}{1-\varepsilon}\{3\pi+2\varepsilon\} \quad , \\ P(3,3) &= p^*d^2 \\ &= \frac{\pi}{1-\varepsilon}(\pi+\varepsilon)^2 \end{aligned}$$

respectively.

For r=1, c_0=n-1, and c_1=1, and therefore

(6.7)
$$\begin{aligned} Pr\{1;3n,\varepsilon\} &= \binom{n}{1} P(3,1)P(3,0)^{n-1} \\ &= n\left[\frac{1-\pi-\varepsilon}{1-\varepsilon}\right]^n \pi(1-\pi)^{2n-2}\{3(1-\pi)-\varepsilon\} \end{aligned}$$

For r=2, there are two cases of c_0=n-1, c_1=0, c_2=1 and c_0=n-2, c_1=2, c_2 =0 to be analzed. In this case

(6.8)
$$\begin{aligned}\Pr\{2;3n,\varepsilon\} &= \binom{n}{1}P(3,2)P(3,0)^{n-1} \\ &+ \binom{n}{2}P(3,1)P(3,0)^{n-2} \\ &= n\left(\frac{1-\pi-\varepsilon}{1-\varepsilon}\right)^{n}\pi(1-\pi)^{2n-2}(3\pi+2\varepsilon) \\ &+ \binom{n}{2}\left(\frac{1-\pi-\varepsilon}{1-\varepsilon}\right)^{n}\pi^{2}(1-\pi)^{2n-4}\{3(1-\pi)-\varepsilon\}^{2} .\end{aligned}$$

For r=3, there are three cases of

$c_0 = n-1, \quad c_1 = 0, \quad c_2 = 0, \quad c_3 = 1$

$c_0 = n-2, \quad c_1 = 1, \quad c_2 = 1, \quad c_3 = 0$

$c_0 = n-3, \quad c_1 = 3, \quad c_2 = 0, \quad c_3 = 0$.

Then

(6.9)
$$\begin{aligned}\Pr\{3;3n,\varepsilon\} &= \binom{n}{1}P(3,3)P(3,0)^{n-1} \\ &+ 2\binom{n}{2}P(3,2)P(3,1)P(3,0)^{n-2} \\ &+ \binom{n}{3}P(3,1)^{3}P(3,0)^{n-3} \\ &= n\left(\frac{1-\pi-\varepsilon}{1-\varepsilon}\right)^{n-1}\frac{\pi(\pi+\varepsilon)^{2}}{1-\varepsilon}(1-\pi)^{2n-2} \\ &+ 2\binom{n}{2}\left(\frac{1-\pi-\varepsilon}{1-\varepsilon}\right)^{n}\pi^{2}(1-\pi)^{2n-4}(3\pi+2\varepsilon)\{3(1-\pi)-\varepsilon\} \\ &+ \binom{n}{3}\left(\frac{1-\pi-\varepsilon}{1-\varepsilon}\right)^{n}\pi^{3}(1-\pi)^{2n-6}\{3(1-\pi)-\varepsilon\}^{3}\end{aligned}$$

In the case r=4,, there are four cases of

$c_0 = n-2, \quad c_1 = 1, \quad c_2 = 0, \quad c_3 = 1$

$c_0 = n-2, \quad c_1 = 0, \quad c_2 = 2, \quad c_3 = 0$

$c_0 = n-3, \quad c_1 = 2, \quad c_2 = 1, \quad c_3 = 0$

$c_0 = n-4, \quad c_1 = 4, \quad c_2 = 0, \quad c_3 = 0$

Therefore

(6.10)
$$\begin{aligned}
\Pr\{4;3n,\varepsilon\} &= 2\binom{n}{2}P(3,3)P(3,1)P(3,0)^{n-2} \\
&+ \binom{n}{2}P(3,2)^2 P(3,0)^{n-2} \\
&+ 3\binom{n}{3}P(3,2)P(3,1)^2 P(3,0)^{n-3} \\
&+ \binom{n}{4}P(3,1)^4 P(3,0)^{n-4} \\
&= 2\binom{n}{2}\left\{\frac{1-\pi-\varepsilon}{1-\varepsilon}\right\}^{n-1}\frac{\pi^2(1-\pi)^{2n-4}}{1-\varepsilon}(\pi+\varepsilon)^2\{3(1-\pi)-\varepsilon\} \\
&+ \binom{n}{2}\left\{\frac{1-\pi-\varepsilon}{1-\varepsilon}\right\}^{n}\pi^2(1-\pi)^{2n-4}(3\pi+2\varepsilon)^2 \\
&+ 3\binom{n}{3}\left\{\frac{1-\pi-\varepsilon}{1-\varepsilon}\right\}^{n}\pi^3(1-\pi)^{2n-6}\{3(1-\pi)-\varepsilon\}^2(3\pi+2\varepsilon) \\
&+ \binom{n}{4}\left\{\frac{1-\pi-\varepsilon}{1-\varepsilon}\right\}^{n}\pi^4(1-\pi)^{2n-8}\{3(1-\pi)-\varepsilon\}^4 .
\end{aligned}$$

As seen easily from these derivations, the expressions of $\Pr\{r;3n,\epsilon\}$ become very complex as r increases. The solutions of (5.3) or (5.3)' for such complex polynomials of π are generally hard to obtain.

REFERENCES

[1] J. C. Ashlock, and E. C. Posner, "*Application of the statistical theory of extreme values to spacecraft receivers*," Technical Report No. 32-737, Jet Propulsion Lab., Calif. Inst. of Tech., Pasadena, Calif.; May 1965.

[2] D. J. Gooding, "*Performance monitor techniques for digital receivers based on extrapolation of error rate*," pp. 380-387, IEEE transactions on Comm. Tech., Vol. COM-16; June 1968.

[3] M. G. Kendall, and A. Stuart, "*The Advanced Theory of Statistics*," Vol. 2, Griffin, London; 1961.

[4] M. Mizuki, "*Confidence intervals for bit error probabilities derived from observed redundant bits,*" Report 4000-70-05, SPAD, FEC/ITT, Vandenberg Air Force Base, Calif.; April 1970.

APPENDIX

BASIC PROPERTIES OF P(n,r) OF THE MARKOV CHAIN MODEL

1. PRELIMINARIES. Let s_k denote the realized k-th state of a Markov chain with two states C and E. Let the initial probabilities be denoted by p_c and p_e , and let the transition probabilities be denoted by p_{cc}, p_{ce}, P_{ec}, and P_{ee} respectively.

The probability that an n-tuple (s_1, ...,s_n) of states is realized is given by the product of transition probabilities,

$$(1) \qquad P(s_1,\ldots,s_n) = p_{s_1} \prod_{k=1}^{n-1} p_{s_k s_{k+1}} \quad .$$

Denote the ensemble of sequences of C and E in each of which exactly r E's and (n-r) C's appear by S(r). The probability that exactly r E's appear in a Markov chain sequence of length r, denoted by P(n,r), is then given by the sun of products of (1) over all sequences in this ensemble S (r) .

In order to facilitate the evaluation of P(n,r), recursive formulas are obtained for the subsets of S(r). Let $P(n,r;s_1=a,s_n=b)$ denote the probability that a Markov chain sequence of length n containing r E's and (n-r) C's satisfies that $s_1=a$ and $s_n =b$. Then

$$(2) \qquad P(n,r;s_1=a,s_n=b) = \sum_{*} p_{s_1} \prod_{k=1}^{n-1} p_{s_k\, s_{k+1}}$$

where $s_1=a$, $s_n=b$, and the summation * is taken over all n-tuples, $(a,s_2,\ldots,s_{n-1},b)$, in which E's appear exactly r times.

2. RECURSIVE AND RELATED FORMULAS FOR $P(n,r;s_1=a,s_n =b)$. For a given n and r, by switching the initial and terminal states, one readily obtains that

$$(3) \qquad P(n,r;s_1=E,s_n=C)=P(n,r;s_1=C,s_n=E)p_e p_{ec}/p_c p_{ce} \quad .$$

The following recursive formulas are obtained by considering the adjunction of an added state either in the front of or at the end of (n-1)-tuples.

(4) $$P(n,r;s_1=E,s_n=C)=P(n-1,r-1;s_1=E,s_{n-1}=C)p_{ee}$$
$$+ P(n-1,r-1;s_1=C,s_{n-1}=C)p_e p_{ec}/p_c$$
$$= P(n-1,r;s_1=E,s_{n-1}=C)p_{cc}$$
$$+ P(n-1,r;s_1=E,s_{n-1}=E)p_{ec} ,$$

(5) $$P(n,r;s_1=C,s_n=E) = P(n-1,r;s_1=C,s_{n-1}=E)p_{cc}$$
$$+ P(n-1,r;s_1=E,s_{n-1}=E)p_c p_{ce}/p_e$$
$$= P(n-1,r-1;s_1=C,s_{n-1}=E)p_{ee}$$
$$+ P(n-1,r-1;s_1=C,s_{n-1}=C)p_{ce}$$

(6) $$P(n,r;s_1=E,s_n=E) = P(n-1,r-1;s_1=E,s_{n-1}=E)p_{ee}$$
$$+ P(n-1,r-1;s_1=C,s_{n-1}=E)p_e p_{ec}/p_c$$
$$= P(n-1,r-1;s_1=E,s_{n-1}=E)p_{ee}$$
$$+ P(n-1,r-1;s_1=E,s_{n-1}=C)p_{ce} .$$

There exist no recursive expressions for $P(n,r;s_1=C,s_n=C)$ given in terms of $P(n-1,r-1;s_1=a,s_{n-1}=b)$. However, it is trivially true that

(7) $$P(n,r;s_1=C,s_n=C) = P(n-1,r;s_1=C,s_{n-1}=C)p_{cc} .$$

By imposing the initial conditions that

(8) $$P(n,0;s_1=E,s_n=C) = 0,$$
$$P(n,0;s_1=C,s_n=E) = 0,$$
$$P(n,0;s_1=E,s_n=E) = 0, \text{ and}$$
$$P(n,0;s_1=C,s_n=C)=p_c p_{cc}^{n-1}$$

for any n, the systems of recursive equations can be solved if $P(n,r;s_1=C, s_n=C)$ are known for all r.

3. DERIVATION OF $P(n,r;s_1=C,s_n=C)$. The following derivation is based on some simple combinatorial results. Because of this, the elementary lemmas are proved using the analogy of picking balls out of an array of balls.

Consider the problem of separating a linear array of m balls into r groups without changing the ordering of the balls. The separation can be made, for instance, by inserting group boundary lines between balls. Let the number of different ways of grouping be denoted by G(m,r). Then,

Lemma 1 $$G(m,r) = \binom{m-1}{r-1} .$$

Proof: There are (m-1) spaces between balls. The m balls are divided into r groups if (r-1) boundary lines are drawn among the (m-1) spaces. q.e.d.

For the same linear array of m balls, let N(m.,r) denote the number of different ways that, when r balls are picked, these r balls are not adjacent to each other. Then,

Lemma 2 $$N(m,r) = \binom{m}{r} + \sum_{h=1}^{r-1} (-1)^h \binom{m-h}{h} \binom{r-1}{h-1}$$

$$= \sum_{h-0}^{r-1} (-1)^h \binom{m-h}{r-h} \binom{r-1}{h-1} ,$$

where $\binom{-1}{-1} = 1$ by definition

Proof: Suppose the r chosen balls are colored red, while the remaining (m-r) balls are white. Suppose the balls are picked in order from the head of the array and placed in boxes which are also arranged in an array following the rule given below. The boxes are filled from the head of the array. Each box should contain at least one or more balls. Neither more than one white ball nor mixed color ball should be in each box. Rule I: A set of red balls in an uninterrupted sub-array may be placed in one or several adjacent boxes in groups. Rule II: A set of red balls in an uninterrupted sub-array is placed in a single box as a group. In either case, if h boxes of red balls ($1 \leq h, \leq r$) exist, there are altogether (m-r+h) boxes.

Let E(m,r,h) denote the event that r red balls out of m balls are placed in h boxes, and let A(u,h) denote the event that no two of these h boxes of red balls out of u boxes (u=m-r+h) are adjacent to each other. Then, obviously

(9) $$E(m,r,h) = A(m,h) \cup A(m-1,h-1) \cup \ldots \cup A(m-h+1,1) \quad .$$

By definition, E(m,r,h) designates the event of placing r red balls according to the Rule I, and A(u,h) designates the same according to the Rule II. Let n(.) denote the number of different ways of placing r red balls into h boxes in a given event "." . Then,

(10) $$n(E(m,r,h)) = \binom{r-1}{h-1}\binom{m-r+h}{h}$$

because there are $\binom{m-r+h}{h}$ different ways of placing h boxes allowing them to be adjacent to each other, and $\binom{r-1}{h-1}$ different ways of placing r balls into h boxes by Lemma 1. In particular,

$$n(E(m,r,r)) = \binom{m}{r} \quad .$$

It is easily seen that N(m,r) = h(A(m,r)). However, from (9),

(11) $$n(A(m,r)) = n(E(m,r,r)) + \sum_{j=1}^{r-1} (-1)^j n(E(m,r,r-j)) \quad .$$

Therefore,

(12) $$N(m,r) = \binom{m}{r} + \sum_{h=1}^{r-1} (-1)^h \binom{m-r+h}{h} \binom{r-1}{h-1} \quad . \quad \text{q.e.d.}$$

<u>Theorem</u> Using Lemma 2,

$$P(n,r;s_1=C,s_n=C)$$

$$= \sum_{j=0}^{r-1} \sum_{i=0}^{r-j-1} (-1)^i \binom{r-1}{j} \binom{n-j-i-2}{r-j-i} \binom{r-j-1}{r-j-i-1} p_c p_{cc}^{n-2r+j-1} (p_{ce}p_{ec})^{r-j} p_{ee}^j \quad .$$

Proof: The E's and C's of the original Markov chain sequences of length n correspond to white and red balls of the Lemmas 1 and 2. There can exist r different classes of sequences of boxes in which (r-j) boxes of E's (j=1,...,r-1) appear according to the Rule II of Lemma 2. For a fixed j, the coefficient is given by

(13) $$N(n-2-j,r-j) = \sum_{i=0}^{r-j-1} (-1)^i \binom{n-j-i-2}{r-j-1} \binom{r-j-1}{r-j-i-1} \quad .$$

The r E's of the Markov chain sequence can be grouped into (r-j) boxes in $\binom{r-1}{j}$ $\binom{r-1}{r-j-1}$ different ways. The powers of the transition probabilities are so defined as to account for those r different classes of sequences of boxes. q.e.d.