THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 52, No. 8, October, 1973
Printed in U.S.A.

Statistical Properties of Gilbert's Burst Noise Model

By MIN-TE CHAO

(Manuscript received April 17, 1973)

Simple statistical procedures for analyzing error data, e.g., in digital data transmission systems, are usually based on the assumption of independence. This paper studies the performance and potential utility of such simple statistical procedures in the case of nonindependent error occurrences. The burst noise model is selected for this purpose because of its neatness, its mathematical tractability, its built-in structure of dependence, and its importance in communication theory. We show that statistical procedures designed under the assumption of independence tend to be conservative for the burst noise model. For example, the usual binomial test will reject, on the average, more channels with small error rates than it would if the errors were independent. The case that the sample size n and the error rate ρ *converge in such a way that* $n\rho \rightarrow \mu_0$ *is also studied. It is shown that the error process can be approximated by a compound Poisson process in continuous time t. The statistical implications of this fact are also discussed.*

I. INTRODUCTION

A dilemma long existing in the theory and applications of digital data transmission is the precise determination of the error structure. On the one hand, it is a well-recognized fact that errors do not occur independently; on the other hand, only the assumption of independence offers us a model sufficiently tractable that ordinary statistical procedures can be designed accordingly. A direct consequence is, of course, that we are using statistical methods designed for independent observations to make statistical inferences on dependent data.

The fact is, we do not have much knowledge of the error structure of data transmission channels. Mathematical models have been constructed for fitting observed data streams containing errors, noticeably the burst noise model of Gilbert,[1] the Markov error process and renewal error process of Elliott,[2,3] and the binary regenerative model of McCullough.[4]

One of the most pertinent models with a built-in dependence structure is Gilbert's burst noise model. It is this model that we shall study in this paper. One of the prime concerns of this study is the behavior of various statistical procedures under the burst noise model.

Gilbert[1] constructs a model for burst noise as follows. An input binary signal (0 or 1) is transmitted through a noisy channel with noise z (0 or 1) so that the output is given by

$$\text{output} \equiv \text{input} + z \qquad (\text{mod } 2).$$

The channel can be in either of the two states, good (G) or bad (B). If, at time n, the channel is in G, there is no noise so $z_n = 0$; if the channel is in B, a "coin" with $P[\text{head}] = h$ is tossed and $z_n = 1$ is identified with a tail outcome.

The channel can shift from a good state to a bad state and vice versa. Identify 1 as G and 2 as B and let X_n denote that state of the channel at time n. It is assumed that the process $\{X_n: n \geqq 1\}$ is a two-state Markov chain with stationary transition probabilities

$$T = \begin{bmatrix} 1 - P & P \\ p & 1 - p \end{bmatrix} \tag{1}$$

and initial distribution (π_1, π_2).

Let $Z_n = z_1 + \cdots + z_n$ denote the number of errors through the nth-bit output (0 or 1) digits of the channel where $z_i = 1$ if and only if an error occurs at the ith bit. The statistic Z_n is obviously the quantity that will be used in any statistical procedure concerning the bit error rate. The statistical behavior of Z_n will be studied extensively in this work.

In Section II, we derive most of the exact formulas concerning Z_n, including explicit expressions for its probability-generating function and its first and second moments. The exact form of the probability distribution of Z_n is quite involved in general. For the special case $p + P = 1$, Z_n reduces to the binomial variable. The quantity $\lambda = 1 - p - P$ can thus be used as a measure of dependence; most of the complications in this work are caused by the presence of a nonzero λ. The effect of dependence is discussed in some detail in Section III. Transmission in blocks of digits is considered; one of our major results is that it can be shown in this model that the block error and the bit error have essentially the same covariance structure. Thus, most results concerning bit error rate can be transferred easily to results about block error rate. As a corollary, the variance for Z_n is obtained as a sum of two components, one due to the sum of variances (as if the z's were independent) and the other due to the fact that $\lambda \neq 0$ (the effect of dependence).

Since Z_n is known to be asymptotically normally distributed, the variance formula of Z_n can be used to judge the effect of dependence on the robustness of statistical procedures (i.e., on how well procedures based on the independence assumption perform if this assumption is violated). A general conclusion of Section IV is that statistical procedures designed under the assumption of independence tend to be conservative for the burst noise model. For example, the usual binomial test will reject, on the average, more channels with small error rate then it is supposed to.

It is shown in Section V that if the bit error rate $\rho \to 0$ in such a way that $n\rho \to \mu_0 > 0$, then Z_n converges in distribution to a compound Poisson distribution. The statistical implications of this fact are also discussed. In particular, Z_n is a minimal sufficient statistic for $\mu_0(\rho)$ in some approximate sense. This justifies the use of Z_n in any statistical decision procedures concerning the error rate ρ.

Despite the model's simplicity, the insight we gained in studying this burst noise model enables us to investigate more deeply the structure of error processes. For example, it is possible to treat the underlying Markov chain $\{X_n\}$ as an s-state stationary Markov chain. Details of this and other extensions and their implications will be discussed in a forthcoming report.

II. STATISTICAL PROPERTIES OF Z_n

We shall assume, for simplicity and without loss of too much generality, in the sequel that the initial distribution (π_1, π_2) of the two-state Markov chain $\{X_n\}$ agrees with its absolute stationary distribution $(p/(p + P), P/(p + P))$. Under this assumption, $\{X_n\}$ is strictly stationary.

Let

$$g_n = P[Z_n = 0].$$

Note that the bit error rate ρ is given by

$$\begin{aligned} \rho_1 &= 1 - g_1 \\ &= P[Z_1 \neq 0] \\ &= P[z_1 = 1] \\ &= (1 - h)P/(p + P); \end{aligned} \tag{2}$$

and the block error rate ρ_k, the probability that a block of size k contains at least one error, is

$$\rho_k = 1 - g_k. \tag{3}$$

Thus, $\rho_1 = \rho$.

Since the event $[z_i = 1]$ implies $[X_i = 2]$ and thus signifies a return to a bad state (a recurrent event), it is possible to utilize the renewal equation to derive an exact expression for g_n. The following theorem is essentially due to Gilbert [Ref. 1, eq. (14)].

Theorem 1: For $n \geqq 1$,

$$g_n = \frac{A_1\alpha_1^{n+1}}{1-\alpha_1} + \frac{A_2\alpha_2^{n+1}}{1-\alpha_2}, \tag{4}$$

where

$$\alpha_1 = \tfrac{1}{2}[-(1-h)(1-p) - (p+P-2) + \sqrt{[(1-P)-h(1-p)]^2 + 4pPh}]$$

$$\alpha_2 = \tfrac{1}{2}[-(1-h)(1-p) - (p+P-2) - \sqrt{[(1-P)-h(1-p)]^2 + 4pPh}]$$

$$A_1 = \rho[\alpha_1 + (p+P-1)]/\alpha_1(\alpha_1-\alpha_2)$$

$$A_2 = \rho[\alpha_2 + (p+P-1)]/\alpha_2(\alpha_2-\alpha_1).$$

A proof of Theorem 1 different from that of Gilbert (and the proofs of all other theorems) will be presented in the appendix. We remark here that since a broader view and a more systematic approach is adopted in our new proof, it is possible to extend our method readily to a more general framework than a two-state Markov chain.

Relation (4) can be viewed as a relation between bit error rate and block error rate. If $\lambda = 1 - p - P > 0$, it can be shown that $0 < \alpha_2 < \alpha_1 < 1$ so that $g_n \to 0$ exponentially fast. One effect of dependence in this model is reflected in (4), namely that g_n is the sum of two exponential terms instead of one. In general, if the underlying Markov chain is s-state, g_n will be a sum of s exponential terms.

The right-hand side of (4) is a function of p, P, and h. We shall write $g_n = g_n(p, P, h)$ when we want to emphasize this point. An important connection between g_n and Eu^{Z_n}, the probability-generating function (PGF) of Z_n, is stated in Theorem 2.

Theorem 2: The probability-generating function of Z_n is given by

$$Eu^{Z_n} = g_n(p, P, H), \tag{5}$$

where

$$H = (1-h)u + h.$$

Thus, replacing each h by $(1 - h)u + h$ in (4), we obtain the PGF of Z_n. The exact expressions for $P[Z_n = i]$ are involved unless i is small. Using (4) and the fact that $0 < \alpha_2 < \alpha_1 < 1$, it is possible to express $P[Z_n = i]$ approximately in terms of its leading term as

$$P[Z_n = i] \approx \frac{A_1^{i+1}}{1 - \alpha_1} \binom{n}{i} \alpha_1^{n+1}. \tag{6}$$

Relation (6) can be used to establish the Poisson convergence of Z_n if $\rho = \mu_0/n \to 0$. However, an indirect proof will be presented later.

Moments of Z_n can be obtained by differentiating the right-hand side of (5) and setting $u = 1$. Specifically, we have

$$EZ_n = n\rho, \tag{7}$$

$$\operatorname{Var} Z_n = n\rho(1 - p) + 2C\left[\frac{(n-1)\lambda}{1-\lambda} - \frac{\lambda^2(1-\lambda^{n-1})}{(1-\lambda)^2}\right], \tag{8}$$

where

$$C = (1 - h)^2 \pi_1 \pi_2$$
$$\lambda = 1 - p - P.$$

Relation (8) also can be obtained by other methods which we shall discuss in Section III.

III. MEASURE OF DEPENDENCE AND ITS EFFECT

If the transition matrix of a Markov chain has identical rows, then this Markov chain is merely a sequence of independent and identically distributed (iid) random variables. For the two-state Markov chain $\{X_n\}$ underlying this burst noise model, the matrix T in (1) has identical rows if and only if $p + P = 1$. Letting $\lambda = 1 - p - P$, we see that $|\lambda| \leqq 1$ and that $\lambda = 0$ if and only if the channel is memoryless.

The eigenvalues of the transition matrix play important roles in the theory of Markov chains. The largest (in absolute value) eigenvalue is always 1; in general, it is the second largest eigenvalue that affects all the essential features of a Markov chain. The parameter λ defined earlier is the second largest eigenvalue of the matrix T in (1).

The significance of the parameter λ can be interpreted intuitively. If p and P are small, the underlying Markov chain $\{X_n\}$ tends to stay in a certain state (G or B) once it enters this state; hence, $\lambda > 0$ indicates the tendency of producing bursty errors. If both p and P are large, then $\{X_n\}$ tends to shift between the good state and bad state alternatively. Since the latter case is obviously not very interesting, we shall always assume $\lambda \geqq 0$ in the sequel.

Let Π denote the 2×2 matrix with identical rows

$$\Pi = \begin{bmatrix} \pi_1 & \pi_2 \\ \pi_1 & \pi_2 \end{bmatrix},$$

where (π_1, π_2) is the absolute stationary distribution of $\{X_n\}$. By the definition of the absolute stationary distribution and by some simple calculations, it can be seen that

$$\Pi T = T\Pi = \Pi^2 = \Pi. \tag{9}$$

It follows from (9) and simple induction that, for $n \geqq 1$,

$$\begin{aligned} T^n - \Pi &= (T - \Pi)^n \\ &= \lambda^n \begin{bmatrix} \pi_2 & \pi_2 \\ \pi_1 & \pi_1 \end{bmatrix}. \end{aligned} \tag{10}$$

Relation (10) allows us to calculate the ℓ-step transition probabilities of $\{X_n\}$ accurately. It can also be used to find the covariance of z_i and z_j. We restate eq. (17) of Ref. 1 as follows:

Theorem 3: The covariance of z_i, z_j $(i \neq j)$ is given by

$$\text{Cov}(z_i, z_j) = C\lambda^{|i-j|}, \tag{11}$$

where $C = (1 - h)^2\pi_1\pi_2$.

Corollary:

$$\text{Var}\,(Z_n) = n\rho(1 - \rho) + 2C\left[\frac{(n-1)\lambda}{1-\lambda} - \frac{\lambda^2(1-\lambda^{n-1})}{(1-\lambda)^2}\right]. \tag{12}$$

Define, for $i = 1, 2, \cdots$,

$$\begin{aligned} T_i &= 1 \quad \text{if} \quad z_{(i-1)k+1} + z_{(i-1)k+2} + \cdots + z_{ik} \geqq 1 \\ &= 0 \quad \text{otherwise}; \end{aligned} \tag{13}$$

namely, $T_i = 1$ if and only if the ith block of length k is not error-free. It is possible to extend eq. (11), and therefore (12), to the corresponding equations involving the T's.

Theorem 4: There exists $0 < C_1 < \infty$ *such that*

$$\text{Cov}\,(T_i, T_j) = C_1\lambda^{|i-j|k}. \tag{14}$$

The value of C_1 can be found explicitly. However, we shall be satisfied with a crude estimate $C_1 = C_2\pi_1\pi_2\lambda^{1-k}$ where $|C_2| \leqq \frac{1}{4}$.

Note that $T_i = z_i$ if $k = 1$. In this case, eq. (14) reduces to (11). Theorem 4 not only states that the T's are "less dependent" than the z's but it also tells us, in some sense, how much less dependent the

T's are. Let

$$S_n = T_1 + T_2 + \cdots + T_n.$$

The statistic S_n is the obvious statistical quantity to analyze if digits are transmitted in blocks of size k. For example, in the 1969–70 Connection Survey[5,6] on the Bell System Switched Telecommunications Network conducted by Bell Laboratories, statistics of block errors are presented for both high-speed and low-speed data transmission. Hence, the more important implication of Theorem 4 is that eq. (14) exhibits the same general structure as eq. (11). For example, replacing C by C_1, ρ by ρ_k, and λ by $\lambda^* = \lambda^k$ in (12), we immediately obtain the formula for Var (S_n).

Corollary:

$$\text{Var}\,(S_n) = n\rho_k(1 - \rho_k) + 2C_1\left[\frac{(n-1)\lambda^*}{1-\lambda^*} - \frac{\lambda^{*2}(1-\lambda^{*n-1})}{(1-\lambda^*)^2}\right]. \quad (15)$$

Consequently, statistical procedures using S_n and concerning the inferences on the block error rate ρ_k should have essentially the same behavior as those procedures using Z_n and concerning the bit error rate ρ. The above reasoning implies that, at least as far as the large sample properties are concerned, it is sufficient to consider inference on ρ only.

Both the law of large numbers and the central limit theorem hold true for the sum of Markovian random variables; see, for example, Ref. 7. Hence,

$$\frac{S_n}{n} \to \rho_k \quad (16)$$

with probability 1; and

$$P\left[\frac{S_n - n\rho_k}{\sqrt{\text{Var}\,S_n}} \leqq v\right] \to \Phi(v) \quad (17)$$

for each $-\infty < v < \infty$, where $\Phi(v)$ denotes the cumulative distribution function of an $N(0, 1)$ random variable. Relations (16) and (17) will be used in Section IV to discuss the robustness of some statistical procedures concerning inferences on ρ_k.

IV. STATISTICAL INFERENCES ON ρ_k

For simplicity, we shall consider the special case $k = 1$ and concentrate our discussion on problems of statistical estimation and hypothesis testing of $\rho = \rho_1$. As remarked earlier in Section III, the restriction $k = 1$ can easily be extended to the general case.

Since $\{X_n\}$ is assumed to be stationary, so is $\{z_n\}$; we have seen that

$$E[Z_n] = n\rho, \tag{18}$$

so that the obvious estimator $\hat{\rho}_n = Z_n/n$ of ρ is unbiased. Relation (16), specialized for the case $k = 1$, states that $\hat{\rho}_n$ is a strongly consistent estimator of ρ.

Very few (optimal) small sample properties of $\hat{\rho}_n$ can be stated, however. For $n \geqq 3$, it can be shown that no uniformly minimum variance estimator of ρ exists. Nevertheless, it is intuitively obvious that $\hat{\rho}_n$ is about the best we can do if the z's are the only observables. From (12),

$$\begin{aligned} n \operatorname{Var}(\hat{\rho}_n) &= \rho(1-\rho) + 2C\frac{\lambda}{1-\lambda} + o(1) \\ &\triangleq \sigma^2 + A, \end{aligned} \tag{19}$$

where

$$\begin{aligned} \sigma^2 &= \rho(1-\rho) \\ A &= 2(1-h)^2\pi_1\pi_2\lambda/(1-\lambda). \end{aligned}$$

Note that the term $\rho(1-\rho)$ in eq. (19) corresponds to $n \operatorname{Var}(\hat{\rho}_n)$ if the z's were independent. Since we have assumed that $\lambda \geqq 0$, it follows that $A \geqq 0$ and $n \operatorname{Var}(\hat{\rho}_n) \geqq \sigma^2$. Thus, the presence of a positive λ actually causes loss of efficiency in estimating ρ. Writing $A = \tau^2$, we see that if the parameters h, p, and P (hence σ^2 and τ^2) can be estimated from the data, the loss of efficiency due to dependence can be estimated as the ratio $\hat{\tau}/\hat{\sigma}$, where $\hat{\tau}$ and $\hat{\sigma}$ denote the estimates of τ and σ from the sample. Hence, if control or confidence limits are used to evaluate the channel performance, the actual 3 standard deviation or 2 standard deviation limits should be wider by $100(\tau/\sigma)$ percent.

We may also consider the loss of power for statistical tests for H_0: $\rho \leqq \rho_0$ of the form

$$\text{reject } H_0 \quad \text{if} \quad Z_n \geqq C^*.$$

Based on the assumption of independence, the power function is approximately

$$\beta_I = 1 - \Phi\left(\frac{C^* - n\rho}{\sqrt{n\sigma^2}}\right); \tag{20}$$

whereas for our model, the power is approximately

$$\beta_D = 1 - \Phi\left(\frac{C^* - n\rho}{\sqrt{n(\sigma^2 + \tau^2)}}\right). \tag{21}$$

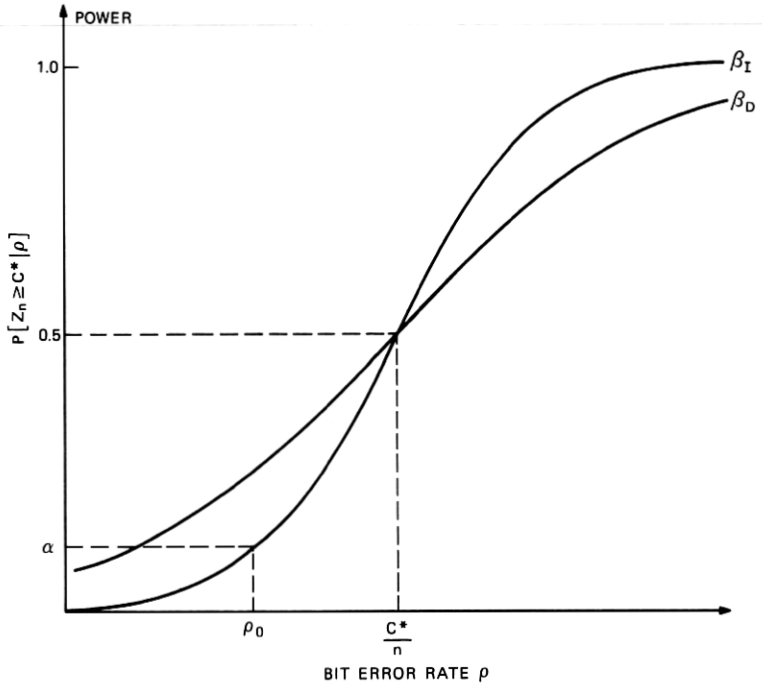

Fig. 1—Comparison of power functions.

If the first-type error $\alpha \leqq \frac{1}{2}$, we see from (20) that $C^* - n\rho_0 \geqq 0$. We see that $\beta_I \leqq \beta_D$ if $C^* - n\rho \geqq 0$ and $\beta_I \geqq \beta_D$ otherwise. This means that it might be possible to design more powerful tests for H_0 based on the knowledge that the dependent model obtains. On the other hand, the test is conservative in the sense that it may reject more channels than expected if the bit error rate ρ is close to the service objective ρ_0 and if the dependent model obtains. The rules of the game shift in the other direction if $C^* - n\rho < 0$. However, it is the smaller values of ρ that we are really concerned with and we may claim that the test based on the assumption of independence gives a pessimistic estimate of channel reliability (see Fig. 1).

V. POISSON APPROXIMATIONS

The bit error rates of high-speed digital channels are usually very small, say 10^{-5} to 10^{-8}; therefore, the normal approximation and the statistical theory discussed earlier may not be too helpful in practice unless n is large. In this section, we prove that Z_n converges in distribution to a Poisson distribution if $n\rho \rightarrow \mu_0$ in a suitable way. Using this result, we construct a Poisson process in continuous time t that approximates the process $\{Z_n(t):t > 0\}$ where n denotes the number of transmitted digits per unit time.

We have shown earlier in (2) that the error rate ρ is given by

$$\rho = (1 - h)P/(p + P). \tag{22}$$

If $\rho \rightarrow 0$ in such a way that $n\rho \rightarrow \mu_0 > 0$, what do we expect to be the asymptotic distribution for $Z_n = z_1 + \cdots + z_n$, the number of errors in the first n digits? Note that we have quite a few choices for the convergence $n\rho \rightarrow \mu_0$. For example, keeping p fixed and letting $P = (\mu/n)^{\epsilon_1}$, $1 - h = (\mu/n)^{\epsilon_2}$, $\epsilon_1 + \epsilon_2 = 1$, we have, by (8),

$$\begin{aligned} \text{Var}\,(Z_n) &\approx \mu_0(1 - \rho) + 2Cn\cdot\frac{\lambda}{1 - \lambda} \\ &\approx \mu_0 + \frac{2}{n^{\epsilon_2}}\cdot\mu^{1+\epsilon_2}\frac{1 - p}{p^2}, \end{aligned} \tag{23}$$

where $\mu_0 = \mu/p$. Also,

$$\begin{aligned} EZ_n &= n\rho \\ &= \mu_0. \end{aligned}$$

Hence, if $\epsilon_2 = 0$ is selected, we see that for large n, Var $Z_n \neq EZ_n$ so that the limiting distribution of Z_n cannot be Poisson.

In order that $\rho = (1 - h)P/(p + P) \approx \mu_0/n$, the most general choice of h and P would be

$$\begin{aligned} 1 - h &= a_1x + a_2x^2 + a_3x^3 + \cdots, \\ P &= b_1y + b_2y^2 + b_3y^3 + \cdots, \end{aligned} \tag{24}$$

where $x = n^{-\epsilon_2}$, $y = n^{-\epsilon_1}$, $\epsilon_1 + \epsilon_2 = 1$, $\epsilon_1 \geqq 0$, $\epsilon_2 > 0$, and $a_1b_1/p = \mu_0$ (the case $\epsilon_2 = 0$ is of particular interest and will be considered separately later). We state the main theorem of this section as follows:

Theorem 5: If $\rho \rightarrow 0$ in such a way that (24) *holds, then*

$$P[Z_n = i] \rightarrow \frac{1}{i!}\mu_0^i e^{-\mu_0}$$

as $n \rightarrow \infty$, where $\mu_0 = a_1b_1/p$. Furthermore, the convergence is uniform in $i = 0, 1, 2, \cdots$.

By using the result of Theorem 5, we may construct a Poisson process in continuous time t as an approximation to the process of partial sums $\{Z_n\colon n \geqq 1\}$. Suppose the underlying channel can transmit n digits per unit time. Let $Z_n(t)$ denote the number of errors in $(0, t)$. Theorem 5 states that, for $i = 0, 1, 2, \cdots$,

$$P[Z_n(t) = i] \rightarrow \frac{1}{i!} (\mu t)^i e^{-\mu t};$$

here μ denotes the limiting error rate per unit time. Let $Z(t)$ denote the number of errors in $(0, t)$ in the limiting case. The fact that $Z(t)$ is a process with independent increments, namely that $Z(t)$ is indeed a Poisson process, is easy to prove and we shall omit it.

Theorem 5 implies that Z_n is asymptotically a minimum sufficient statistic for the bit error rate ρ if (24) can be justified; this provides theoretical support for the use of Z_n in any statistical inferences concerning ρ. We remark here that, by replacing ρ by ρ_k and Z_n by S_n, the same comment applies for block error rates. Another consequence of Theorem 5 is that

$$\text{Var}\,(Z_n) \rightarrow \mu_0 = a_1 b_1/p. \tag{25}$$

Note that if $\lambda = 1 - p - P = 0$ (the independent case), (24) implies $P \rightarrow 0$ and this in turn implies $p \rightarrow 1$. From (25), we see that Var (Z_n) is minimized in the independent case. The increase of variance due to dependence is therefore $100(1/p - 1)$ percent. Hence, in the dependent case, the confidence interval for ρ should be wider than we thought in the independent case.

The null hypothesis $H_0: \rho \leqq \rho_0$ becomes $H_0': \mu_0 \leqq \mu_0^*$ in the limiting case. The uniformly most powerful test for H_0' exists and is given by the rule:

$$\text{reject } H_0' \quad \text{if} \quad Z_n \geqq C^*.$$

Based on the approximation that Z_n is Poisson, we may compute the power functions as

$$\begin{aligned}\beta_D(\mu_0) &= P[Z_n \geqq C^* \mid \mu_0] \\ &= \sum_{i=C^*}^{\infty} \frac{1}{i!} \mu_0^i e^{-\mu_0} \\ &= \int_0^{\mu_0} \frac{1}{(C^* - 1)!} e^{-x} x^{C^*-1} dx\end{aligned}$$

and

$$\beta_I = \int_0^{a_1 b_1} \frac{1}{(C^* - 1)!} e^{-x} x^{C^*-1} dx.$$

It follows that $\beta_I \leqq \beta_D$ so that a test for H_0 based on the assumption of independence and used when dependence is present rejects more channels than it should. In other words, tests designed for independent observations protect customers in the sense that channels they are using may have better quality than inferred.

The effect of dependence reported for both the binomial and the Poisson cases has an intuitive explanation. By using Z_n or S_n, we are actually abandoning some of the information contained in the sequence $z_1, z_2, \cdots$, so that statistical inferences based on Z_n or S_n tend to be more conservative in the sense that channel reliability is estimated pessimistically.

We now return to (24) and consider the special case $\epsilon_2 = 0$. This case cannot be ignored because previous papers, for example Ref. 1, indicate that sometimes $h \approx 0.5$ (rather than 0.999) is a reasonable value. The fact that Poisson processes do not describe certain error processes well has also been reported in the literature.

If $\epsilon_2 = 0$, eq. (24) reduces to

$$P = [b_1 + o(1)]/n \qquad b_1 > 0. \tag{26}$$

We have

Theorem 6: *If* (26) *holds, then*

$$Eu^{Z_n} \to \exp\left[-\frac{b_1(1-H)}{1-H+pH}\right], \tag{27}$$

where $H = (1-h)u + h$.

We remark here that the limiting value in eq. (27) is the PGF of a compound Poisson process. More specifically, let N be a Poisson variable with mean $b_1/(1-p)$, and let $W_1, W_2, \cdots$ be iid random variables with the geometric distribution

$$P[W_1 = i] = \left[\frac{p}{1-(1-p)h}\right]\left[\frac{(1-p)(1-h)}{1-(1-p)h}\right]^i, \tag{28}$$
$$i = 0, 1, 2, \cdots.$$

If the W's are independent of N, then the left-hand side of (27) is simply $Eu^{W_1+W_2+\cdots+W_N}$. It is of course possible to introduce a continuous time parameter t and consider the following random mechanism which describes the bursty nature of this error process vividly. The bursts are generated by a Poisson process; given that a burst occurs, the errors are generated by a geometric distribution.

From the right-hand side of (27), it is possible to compute the moments of the limiting distribution of Z_n. We have

$$E(W_1 + W_2 + \cdots + W_N) = \frac{b_1(1-h)}{p} \tag{29}$$

$$\text{Var}\,(W_1 + W_2 + \cdots + W_N) = \frac{b_1(1-h)}{p} + \frac{2b_1(1-h)^2(1-p)}{p^2}. \tag{30}$$

Note that the variance is always larger than the mean in this case. Note also that, as h approaches 1, the second term on the right-hand side of eq. (30) is of higher order and vanishes in the limiting case. Another interesting thing is that it is possible to show that the right-hand side of (30) is minimized at $p = 1$, and as p approaches 1, the limiting distribution is Poisson.

Branching renewal processes have been suggested in the literature[8] as a model for series of events. The basic structure for branching renewal processes can be described in terms of our problem as follows: The series of primary events (bursts) are generated by a Poisson process. Each of these primary events generates a subsidiary series of events (bit errors), separated by the waiting time $Y_1, Y_2, \cdots, Y_S$, where S is random. If we assume that these subsidiary series of events take no time, then the branching renewal process reduces to the compound Poisson process.

VI. CONCLUSIONS AND FURTHER EXTENSIONS

(*i*) The burst noise model of Gilbert discussed in this paper provides a vehicle for studying the robustness of some fixed sample size statistical procedures. The general result is that the presence of dependence increases the variance of the random variable Z_n, for the case where the bit error rate ρ is fixed and the case in which $\rho = [\mu_0 + o(1)]/n$. Thus, use of statistical tests based on the assumption of independence increases the power at the cost of rejecting more satisfactory channels than would be rejected if dependence were absent. The use of blocks does reduce the covariances among errors compared with bits or smaller blocks. However, the covariance structure among the blocks is essentially the same as that among the bits.

(*ii*) Although the dependence structure of the Gilbert's burst noise model is a simple one, it is by no means a trivial one. In fact, from the insight gained through this study, many results obtained in this paper have generalizations in error processes defined over an s-state Markov chain as well. A unified treatment on channels with Markov type of memory will be reported elsewhere.

(*iii*) The second largest eigenvalue (in absolute value) of the $(s \times s)$ transition matrix of the underlying Markov chain is a parameter which should not be overlooked. It can be viewed as a measure of dependence of a Markovian model. The effect of this parameter $(=\lambda$ in this work) is visible in many important formulas, for example, in (14).

(*iv*) Another important question to ask is what kind of stochastic process can be used to approximate the error process of a binary channel with memory. If the bit error rate is small, we can extend the proof of Theorem 6 (in a nontrivial way) to find an important conclusion: the compound Poisson process can serve the purpose.

(*v*) The by-products of this work are also fruitful. For example, the variance formula of Z_n can be generalized to find the variance of $T_n = f(X_1) + \cdots + f(X_n)$ where $\{X_i\}$ is an s-state Markov chain, $s \leqq \infty$, and f is an arbitrary function. Since many continuous sampling plans, such as CSP1, CSP2, CSP3, can be described as random walks of the form T_n (see Refs. 9 and 10), the application of this formula to quality assurance is evident.

(*vi*) Mathematically speaking, there is an essential difference between Gilbert's original treatment and our generalizations to the s-state Markov chain. More specifically, Gilbert viewed his problem as one of the renewal type whereas the s-state Markov case should be handled by the semigroup property (of taboo probabilities). We remark here that many results of the theory of recurrent events (see, for example, Ref. 11) can be applied to Gilbert's model. We also remark that the renewal process is a one-state semi-Markov process. A general question can be raised at this point: What is the behavior of an s-state semi-Markov channel? Since it is known that distributions other than the exponential (for example, the Pareto distribution, see Ref. 12) describe the waiting time distribution well, the question raised is a realistic one and should not be merely considered as an attempt at mathematical generality.

VII. ACKNOWLEDGMENTS

The author wishes to thank J. A. Tischendorf for his critical reading. He also wants to thank L. B. Boza for many helpful discussions.

APPENDIX

A.1 *Proof of Theorem 1*

Consider $Y_n = (X_n, z_n)$ as a three-state Markov chain with transition matrix

$$\begin{array}{c} \\ (\mathrm{G},0) \\ (\mathrm{B},0) \\ (\mathrm{B},1) \end{array} \begin{array}{c} \begin{array}{ccc} (\mathrm{G},0) & (\mathrm{B},0) & (\mathrm{B},1) \end{array} \\ \begin{bmatrix} 1-P & hP & (1-h)P \\ p & h(1-p) & (1-h)(1-p) \\ p & h(1-p) & (1-h)(1-p) \end{bmatrix} \end{array} = Q = (q_{ij}), \tag{31}$$

say. We have

$$Eu^{Z_n} = Eu^{z_1+z_2+\cdots+z_n}$$
$$= \textstyle\sum_{y_0} \sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} u^{z_1+\cdots+z_n} q_{y_{n-1}y_n} \cdot q_{y_{n-2}y_{n-1}} \cdots q_{y_0y_1} \cdot \lambda_{y_0}, \quad (32)$$

where $\lambda_{y_0} = P[Y_0 = y_0]$. Note that the value of z_ℓ is completely determined by the value $Y_\ell = y_\ell$. Let

$$r_{y_{\ell-1},y_\ell} = u^{z_\ell} q_{y_{\ell-1},y_\ell},$$

and let

$$R = (r_{ij}).$$

Relation (32) can then be written as

$$Eu^{Z_n} = \boldsymbol{\lambda}' R^n \mathbf{1}, \quad (33)$$

where

$$\boldsymbol{\lambda}' = (\lambda_{(G,0)}, \lambda_{(B,0)}, \lambda_{(B,1)}),$$
$$\mathbf{1}' = (1, 1, 1),$$
$$R = (r_{ij})$$
$$= Q\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & u \end{bmatrix}.$$

We remark here that eq. (33) can be extended to the case of an s-state Markov chain easily. Letting $u = 0$ in eq. (33), the PGF of Z_n, we have

$$P[Z_n = 0] = \boldsymbol{\lambda}'\left[Q\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}\right]^n \mathbf{1}. \quad (34)$$

The last column of the 3×3 matrix in eq. (34) is always a zero vector for every $n \geqq 1$. Hence, the right-hand side of (34) is essentially the nth power of a 2×2 matrix. The explicit formula for g_n in eq. (4) follows from (34) by straightforward calculations.

A.2 *Proof of Theorem 2*

The z's are conditionally independent if the values of the X's are given. Hence,

$$P[Z_n = 0] = E\{P[z_1 = z_2 = \cdots = z_n = 0 \mid X_1, X_2, \cdots, X_n]\}$$
$$= E\left\{\prod_{i=1}^{n} P[z_i = 0 \mid X_i]\right\}$$
$$= E \prod_{i=1}^{n} h^{X_i - 1}$$
$$= Eh^{X_1+X_2+\cdots+X_n-n}. \quad (35)$$

Similarly,

$$\begin{aligned} Eu^{Z_n} &= E\{E[u^{z_1+z_2+\cdots+z_n}|X_1, X_2, \cdots, X_n]\} \\ &= E\left\{\prod_{i=1}^{n} E[u^{z_i}|X_i]\right\} \\ &= E\left\{\prod_{i=1}^{n} [h + (1-h)u]^{X_i-1}\right\} \\ &= EH^{X_1+X_2+\cdots+X_n-n}, \end{aligned} \tag{36}$$

where $H = h + (1 - h)u$. By comparing eqs. (35) and (36), Theorem 2 follows.

A.3 *Proof of Theorem 3*

By eq. (10),

$$P[X_n = 2|X_0 = 2] = \pi_2 + \lambda^n \pi_1.$$

Hence,

$$\begin{aligned} \text{Cov}\ (z_i, z_j) &= P[z_i = z_j = 1] - \rho^2 \\ &= (1-h)^2 P[X_i = X_j = 2] - \rho^2 \\ &= (1-h)^2 \pi_2(\pi_2 + \pi_1 \lambda^{|i-j|}) - \rho^2 \\ &= \pi_1\pi_2(1-h)^2\lambda^{|i-j|}. \quad \text{QED.} \end{aligned}$$

A.4 *Proof of Theorem 4*

Let us compute a special case first. Consider $P[T_1 = 0, T_n = 0]$. A typical path of the underlying Markov chain $\{X_1, X_2, \cdots, X_{kn}\}$ may be of the following form:

$$b(x_k, x_{k+1}, x_{(n-1)k}, x_{(n-1)k+1}) = (x_1x_2\cdots x_kx_{k+1}\cdots x_{(n-1)k}x_{(n-1)k+1}x_{(n-1)k+2}\cdots x_{nk}). \tag{37}$$

|← first block of size k →| |← last block of size k →|

The rest of the x's in b (and in W_1, W_2 later) are omitted for typographical reasons. Note that the values of $x_{k+2}, \cdots, x_{(n-1)k-1}$ are deliberately unspecified; also, $n > 2$ is assumed *pro tem.*

For fixed first block and last block, there are four different kinds of paths, according to the values of x_{k+1}, $x_{(n-1)k}$. Let m denote the number of 2's in the first and last blocks together. We have

$$P[T_1 = 0, T_n = 0|b(x_k, x_{k+1}, x_{(n-1)k}, x_{(n-1)k+1})] = h^m \tag{38}$$

and

$$P[b(x_k, x_{k+1}, x_{(n-1)k}, x_{(n-1)k+1})] = W_1(x_k)p_{x_kx_{k+1}}p^{[(n-2)k-1]}_{x_{k+1}x_{(n-1)k}}p_{x_{(n-1)k}x_{(n-1)k+1}}W_2(x_{(n-1)k+1}), \tag{39}$$

where

$$W_1(x_k) = P[X_1 = x_1, X_2 = x_2, \cdots, X_{k-1} = x_{k-1}, X_k = x_k]$$

and

$$W_2(x_{(n-1)k+1}) = P[X_{(n-1)k+2} = x_{(n-1)k+2}, \cdots, X_{nk} = x_{nk} \mid X_{(n-1)k+1} = x_{(n-1)k+1}].$$

We may also find $P[T_1 = 0]\cdot P[T_n = 0]$ by considering their conditional probabilities over the first and the nth blocks. It is not difficult to see that

$$P[T_1 = 0]P[T_n = 0] = \sum h^m W_1(x_k) W_2(x_{(n-1)k+1}) \cdot \pi_{x_{(n-1)k+1}}, \quad (40)$$

where the summation ranges over all 2^{2k} possible blocks. The expression for $P[T_1 = 0, T_n = 0]$ can be obtained by taking the product of (38) and (39) and summing over all 2^{2k+2} possibilities. The 2^{2k+2} terms in this form of $P[T_1 = 0, T_n = 0]$ outnumbered the terms in (40) by a margin of 4 to 1, and there is an obvious 4:1 correspondence between these terms. Consider

$$\begin{aligned} \mathrm{Cov}\,(T_1, T_n) &= \mathrm{Cov}\,(1 - T_1, 1 - T_n) \\ &= P[T_1 = 0, T_n = 0] - P[T_1 = 0]P[T_n = 0]. \end{aligned} \quad (41)$$

For fixed first and last blocks, a typical difference between the (4:1) correspondent terms is

$$\begin{aligned} h^m[W_1(x_k)W_2(x_{(n-1)k+1})][&p_{x_k1}p_{11}^{[(n-2)k-1]}p_{1x_{(n-1)k+1}} \\ &+ p_{x_k1}p_{12}^{[(n-2)k-1]}p_{2x_{(n-1)k+1}} + p_{x_k2}p_{21}^{[(n-2)k-1]}p_{1x_{(n-1)k+1}} \\ &+ p_{x_k2}p_{22}^{[(n-2)k-1]}p_{2x_{(n-1)k+1}} - \pi_{x_{(n-1)k-1}}]. \end{aligned} \quad (42)$$

By (10), it can be shown that the third factor of (42) becomes

$$(p_{x_k1}p_{1x_{(n-1)k+1}} + p_{x_k1}p_{2x_{(n-1)k+1}} + p_{x_k2}p_{1x_{(n-1)k+1}} + p_{x_k2}p_{2x_{(h-1)k+1}})\lambda^{(n-2)k-2}$$

$$= \begin{cases} \dfrac{P}{p+P}\lambda^{(n-2)k+1} & \text{if} \quad (x_k, x_{(n-k)+1}) = (1,1) \\ -\dfrac{P}{p+P}\lambda^{(n-2)k+1} & \qquad\qquad = (1,2) \\ -\dfrac{p}{p+P}\lambda^{(n-2)k+1} & \qquad\qquad = (2,1) \\ \dfrac{p}{p+P}\lambda^{(n-2)k+1} & \qquad\qquad = (2,2). \end{cases} \quad (43)$$

Note that in all terms we have a common factor $\lambda^{(n-2)k+1}$. By factoring out this common factor, Theorem 4 follows immediately.

We may even push the computations further to find an exact expression for the constant C_1 in Theorem 4. Note that we have four types of combinations of blocks, according to the values of x_k and $x_{(n-k)+1}$. The quantity in (42) becomes

$$
\begin{aligned}
&(1, 1) \Rightarrow h^m W_1(1) W_2(1) \frac{P}{p+P} \lambda^{(n-2)k+1},\\
&(1, 2) \Rightarrow -\; h^m W_1(1) W_2(2) \frac{P}{p+P} \lambda^{(n-2)k+1},\\
&(2, 1) \Rightarrow -\; h^m W_1(2) W_2(1) \frac{p}{p+P} \lambda^{(n-2)k+1},\\
&(2, 2) \Rightarrow h^m W_1(2) W_2(2) \frac{p}{p+P} \lambda^{(n-2)k+1},
\end{aligned}
\tag{44}
$$

and Cov (T_1, T_n) is the sum of all 2^{2k} terms in (44).

Let

$$T_1' = z_1 + z_2 + \cdots + z_{k-1}, \qquad T_2' = z_{k+2} + \cdots + z_{2k}.$$

(We should use $T_n' = z_{(n-1)k+2} + \cdots + z_{nk}$; however, the distribution of T_n' is independent of n so we may take $n = 2$.) Then

$$
\begin{aligned}
P[T_1' &= 0 | X_k = 1] P[T_2' = 0 | X_{k+1} = 1]\\
&= \sum_{2^{2(k-1)}} h^m P[X_1 = x_1, \cdots, X_{k-1} = x_{k-1} | X_k = 1]\\
&\qquad\qquad \cdot P[X_{k+2} = x_{k+2}, \cdots, X_{2k} = x_{2k} | X_{k+1} = 1]\\
&= \frac{1}{\pi_1} \sum_{2^{2(k-1)}} h^m W_1(1) W_2(1),
\end{aligned}
$$

where m denotes the number of 2's in the sequence $x_1, x_2, \cdots, x_{k-1}, x_{k+2}, \cdots, x_{2k}$, which is equal to the number of 2's in the sequence $x_1, x_2, \cdots, x_{2k}$ in the case $x_k = x_{k+1} = 1$. Thus, the sum of terms of the type (1, 1) in (44) is simply

$$\pi_1 P[T_1' = 0 | X_k = 1] P[T_2' = 0 | X_{k+1} = 1] \cdot \pi_2 \lambda^{(n-1)k+1}.$$

Similarly, we may find the sums of other types of terms in (44). We have

$$
\begin{aligned}
&(1, 2) \Rightarrow -\; \pi_1 h P[T_1' = 0 | X_k = 1] P[T_2' = 0 | X_{k+1} = 2] \pi_2 \lambda^{(n-2)k+1}\\
&(2, 1) \Rightarrow -\; \pi_2 h P[T_1' = 0 | X_k = 2] P[T_2' = 0 | X_{k+1} = 1] \cdot \pi_1 \lambda^{(n-2)k+1}\\
&(2, 2) \Rightarrow \pi_2 h^2 P[T_1' = 0 | X_k = 2] P[T_2' = 0 | X_{k+1} = 2] \pi_1 \lambda^{(n-2)k+1}.
\end{aligned}
$$

Thus, if $n > 2$, $i \geqq 1$,

$$\begin{aligned}\text{Cov}\,(T_1, T_n) &= \text{Cov}\,(T_i, T_{i+n-1})\\ &= \pi_1\pi_2\lambda^{(n-2)k+1}\{P[T_1' = 0 \mid X_k = 1]P[T_2' = 0 \mid X_{k+1} = 1]\\ &\qquad - hP[T_1' = 0 \mid X_k = 1]P[T_2' = 0 \mid X_{k+1} = 2]\\ &\qquad - hP[T_1' = 0 \mid X_k = 2]P[T_2' = 0 \mid X_{k+1} = 1]\\ &\qquad + h^2P[T_1' = 0 \mid X_k = 2]P[T_2' = 0 \mid X_{k+1} = 2]\}. \qquad (45)\end{aligned}$$

The case $n = 2$ should be considered separately; this is because $(n - 2)k - 1 < 0$ if $n = 2$ so that (39) simplifies to

$$P[b(x_k, x_{k+1})] = W_1(x_k)p_{x_kx_{k+1}}W_2(x_{k+1}). \qquad (46)$$

In this case, the number of terms in $P[T_1 = 0, T_2 = 0]$ equals the number of terms in the product $P[T_1 = 0]P[T_2 = 0]$ and there is an obvious one-to-one correspondence between the terms. Consider the difference $P[T_1 = 0, T_2 = 0] - P[T_1 = 0]\ P[T_2 = 0]$. For fixed first and last (second) blocks, the term-wise difference is

$$h^mW_1(x_k)W_2(x_{k+1})[p_{x_kx_{k+1}} - \pi_{x_{k+1}}]. \qquad (47)$$

The last factor in (47) can be computed. We have

$$\begin{aligned}p_{x_kx_{k+1}} - \pi_{x_{k+1}} &= \frac{P}{p+P}\lambda \quad \text{if} \quad (x_k, x_{k+1}) = (1, 1)\\ &= -\frac{P}{p+P}\lambda \qquad\qquad\qquad = (1, 2)\\ &= -\frac{p}{p+P}\lambda \qquad\qquad\qquad = (2, 1)\\ &= \frac{p}{p+P}\lambda \qquad\qquad\qquad\ \ = (2, 2). \qquad (48)\end{aligned}$$

Note that (43) reduces to (48) if $n = 2$; hence, all arguments leading to (45) hold true even if $n = 2$.

Let C_2 be the quantity in the large square bracket of (45); we may write (45) as

$$\text{Cov}\,(T_i, T_j) = C_2\pi_1\pi_2\lambda^{(|i-j|-1)k+1} \qquad (49)$$

if $i \leqq j$.

It is possible to find the value of C_2 through an argument similar to that of finding g_n in Theorem 1. However, we shall be satisfied with a crude estimate

$$|\pi_1\pi_2C_2| \leqq 1,$$

which follows from $\pi_1\pi_2 = \pi_1(1 - \pi_1) \leqq \frac{1}{4}$ trivially.

A.5 *Proof of Theorem 5*

Let $H = (1 - h)u + h$ and let $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{A}_1$, $\hat{A}_2$, $\hat{\rho}$ be the quantities obtained from α_1, α_2, A_1, A_2, ρ by replacing each h with H respectively. As $n \to \infty$, $H \to 1$. Hence,

$$\begin{aligned} \hat{\alpha}_1 &\to 1 \\ \hat{\alpha}_2 &\to 1 - p < 1 \\ \hat{A}_1 &\to 0 \\ \hat{A}_2 &\to 0 \\ \hat{\rho} &\to 0. \end{aligned}$$

It follows from Theorem 2 that

$$\lim_{n\to\infty} Eu^{Z_n} = \lim_{n\to\infty} \frac{\hat{A}_1 \hat{\alpha}_1{}^{n+1}}{1 - \hat{\alpha}_1}. \tag{50}$$

Let $\Delta^2 = [1 - P - H(1 - p)]^2 + 4pPH$ in the expression of $\hat{\alpha}_1$. An important step in our argument is to find the value of Δ. By substituting (24) into the expression for Δ^2, we have

$$\Delta^2 = p^2 + \sum_{n=1}^{\infty} \gamma_n x^n + \sum_{n=1}^{\infty} \delta_n y^n + \epsilon xy + o(xy), \tag{51}$$

where

$$\begin{aligned} \gamma_{2n} &= (1 - p)^2(1 - u)^2(a_n^2 + 2a_1a_{2n-1} + 2a_2a_{2n-2} + \cdots \\ &\qquad + 2a_{n-1}a_{n+1}) + 2p(1 - p)(1 - u)a_{2n} \\ \gamma_{2n+1} &= (1 - p)^2(1 - u)^2(2a_1a_{2n} + 2a_2a_{2n-1} + \cdots + 2a_na_{n+1}) \\ \delta_{2n} &= b_n^2 + 2b_1b_{2n-1} + 2b_2b_{2n-2} + \cdots + 2b_{n-1}b_{n+1} \\ \delta_{2n+1} &= 2b_1b_{2n} + 2b_2b_{2n-1} + \cdots + 2b_nb_{n+1} \\ \epsilon &= -2(1 - p)(1 - u)a_1b_1 - 4p(1 - u)a_1b_1. \end{aligned}$$

Let

$$\Delta = p + \sum_{n=1}^{\infty} (d_nx^n + e_ny^n) + fxy + o(xy). \tag{52}$$

By comparing the Δ^2 in (52) with the same quantity in (51), it is not difficult to see that

$$\begin{aligned} d_k &= (1 - p)(1 - u)a_k \\ e_k &= b_k \end{aligned} \tag{53}$$

for $k = 1, 2, \cdots$. Also, it is easy to find that

$$f = -\frac{2}{p}(1 - u)a_1b_1. \tag{54}$$

Using (53), it can be seen that in the expression of $\hat{\alpha}_1$, the coefficients of x^k, y^k are zero for all k. Hence, recall that $xy = 1/n$,

$$\hat{\alpha}_1 = 1 - \frac{(1-u)a_1b_1}{p} xy + o(xy)$$
$$= 1 - \frac{(1-u)a_1b_1}{np} + o\left(\frac{1}{n}\right). \tag{55}$$

By (55) and (24), it can be seen that

$$\hat{A}_1 = \frac{a_1b_1(1-u)}{pn} + o\left(\frac{1}{n}\right). \tag{56}$$

By (50), (55), and (56), we have

$$\lim_{n\to\infty} Eu^{Z_n} = \exp\left(\frac{a_1b_1(u-1)}{p}\right),$$

which is the PGF of the Poisson distribution with mean equal to a_1b_1/p.

A.6 *Proof of Theorem 6*

Using (26), we may express $\hat{\alpha}_1$, $\hat{\alpha}_1$ in terms of powers of $1/n$ as

$$\hat{\alpha}_1 = 1 - \frac{\alpha}{n} + o\left(\frac{1}{n}\right), \tag{57}$$
$$\hat{A}_1 = \frac{\alpha}{n} + o\left(\frac{1}{n}\right),$$

where

$$\alpha = \frac{b_1(1-H)}{1-H+pH} \tag{58}$$
$$H = (1-h)u + h.$$

By eqs. (50), (57), and (58), we have

$$\lim_{n\to\infty} Eu^{Z_n} = \lim_{n\to\infty} \frac{\hat{A}_1}{1-\hat{\alpha}_1} \lim_{n\to\infty} \hat{\alpha}_1{}^{n+1}$$
$$= \exp\left[-\alpha\right]$$
$$= \exp\left[-\frac{b_1(1-H)}{1-H+pH}\right]. \tag{59}$$

This proves Theorem 6.

REFERENCES

1. Gilbert, E. N., "Capacity of a Burst-Noise Channel," B.S.T.J., *39*, No. 5 (September 1960), pp. 1253–1266.

2. Elliott, E. O., "Estimates of Error Rates for Codes on Burst-Noise Channels," B.S.T.J., *42*, No. 5 (September 1963), pp. 1977–1998.
3. Elliott, E. O., "A Model of the Switched Telephone Network for Data Communications," B.S.T.J., *44*, No. 1 (January 1965), pp. 89–109.
4. McCullough, R. H., "The Binary Regenerative Channel," B.S.T.J., *47*, No. 8 (October 1968), pp. 1713–1735.
5. Balkovic, M. D., Klancer, H. W., Klare, S. W., and McGruther, W. G., "1969–70 Connection Survey: High-Speed Voiceband Data Transmission Performance on the Switched Telecommunications Network," B.S.T.J., *50*, No. 4 (April 1971), pp. 1349–1384.
6. Fleming, H. C., and Hutchinson, R. M., Jr., "1969–70 Connection Survey: Low-Speed Data Transmission Performance on the Switched Telecommunications Network," B.S.T.J., *50*, No. 4 (April 1971), pp. 1385–1405.
7. Chung, K. L., *Markov Chains with Stationary Transition Probabilities*, Part I, Sections 15–16, New York: Springer-Verlag, 1967.
8. Cox, D. R., and Lewis, P. A. W., *The Statistical Analysis of Series of Events*, Chapter 7, London: Methuen & Co. Ltd., 1966.
9. Roberts, S. W., "States of Markov Chains for Evaluating Continuous Sampling Plans," Trans. 17th Annual All-Day Conference on Quality Control, Metropolitan Section ASQC and Rutgers University, New Brunswick, New Jersey (1965), pp. 106–111.
10. Lieberman, G. J., and Solomon, H., "Multilevel Continuous Sampling Plans," Ann. Math. Stat., *26* (1955), pp. 686–704.
11. Feller, W., "Fluctuation Theory of Recurrent Events," Trans. Amer. Math. Soc., *67* (1949), pp. 98–119.
12. Sussman, S. M., "Analysis of the Pareto Model for Error Statistics on Telephone Circuits," IEEE Trans. Commun. Syst., *11*, No. 2 (June 1963), pp. 213–221.