AN EVALUATION OF TWO METHODS FOR THE DETECTION OF SIGNALS IN NOISE

A THESIS

Presented to

the Faculty of the Graduate Division

Georgia Institute of Technology

by

Murray Rae

June 1957

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Electrical Engineering

Georgia Institute of Technology

June 1957

3$
12k

AN EVALUATION OF TWO METHODS FOR
THE DETECTION OF SIGNALS IN NOISE

Approved: ______________________

____

____

____

____

____

Date approved by Chairman: May 31, 1957

"In presenting the dissertation as a partial fulfillment of the requirements for an advanced degree from the Georgia Institute of Technology, I agree that the Library of the Institution shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to copy from, or to publish from, this dissertation may be granted by the professor under whose direction it was written, or, in his absence, by the dean of the Graduate Division when such copying or publication is solely for scholarly purposes and does not involve potential financial gain. It is understood that any copying from, or publication of, this dissertation which involves potential financial gain will not be allowed without written permission.

ACKNOWLEDGEMENTS

The author wishes to express his sincere appreciation to Professor Immel of the Mathematics Department, Georgia Institute of Technology, for his assistance in the mathematical part of this work and to the members of the reading committee, Professors B. J. Dasher and D. L. Finn.

TABLE OF CONTENTS

ABSTRACT

A problem of considerable importance in communication theory is the detection of a repetitive signal which has been masked by noise. Since the noise is of a random nature the detection process is usually based on probability theory. In a practical case some criterion for detection is considered and, if measurements on the received waveform meet this criterion, then it is assumed that a signal is present and, if they do not, then it is said that no signal is present.

Various means for deriving the best form for this criterion have been suggested but invariably they involve some prior knowledge of the signal's probability distribution. Usually the form of this distribution is assumed to be a constant since nothing is normally known about its true form. In this work it is shown that this assumption is, for large values of noise power per unit bandwidth, liable to lead to quite erroneous results. Various other forms for the a priori distribution function have been substituted into the expression for the probability of detection of the signal and the results show that a better form for the a priori distribution function would be a function which increases as the variable increases. It is suggested that the optimum form could be obtained by means of the Theory of Games, since the other methods depend on obtaining the value of some characteristic which gives a maximum probability of detection which might not be of much utility if there is a considerable spread about this maximum.

In addition to showing how some prior knowledge of the signal

affects the probability of the signal's detection it is shown that there is a probable maximum in the amount of information that may be extracted from an information carrying waveform. A method, based on the probability of existence of the signal, for making a comparison of receivers on an efficiency basis is developed.

An extension of the idea of existence probabilities is used to show that the position of a target may be indicated merely by looking for the presence of a signal in the received waveform, thus suggesting that it might be possible to detect a target without knowing anything about the characteristics of the signal.

LIST OF FIGURES

LIST OF SYMBOLS

x, $x(t)$: Voltage function representing the message.

y, $y(t)$: Received voltage function consisting of the message function and noise.

$p(x)$: Probability that a sample of x lies in the range x to x dx.

$p(y)$: Probability that a sample of y lies in the range y to y dy.

$p(x/y)$: Probability that x is in the range x to x dx when y is in the range y to y dy.

$p(E/y)$: Observer's probability of the existence of a signal when given the waveform y.

$p(N/y)$: Observer's probability of the non-existence of a signal when given the waveform y.

$\overline{p}(A/B)$: Mean probability of A given B averaged over some other parameter.

λ: A priori probability of the presence of a signal.

μ: A priori probability of the absence of a signal.

k: Normalizing factor.

T: Period of observation.

N_o: Mean square power per unit bandwidth.

a,b,R: Functions of the noise waveform.

I_o: Zero order Bessel function with imaginary argument.

T': Pulse period.

n: Number of samples per scan.

CHAPTER I

INTRODUCTION

A significant advance in communication engineering in recent years is the development of a theory of communication in which the methods and techniques of the statistician have augmented those of the communications engineer. The basis for the new development is the concept that the flow of information, which is the primary concern of a communication system, is a statistical phenomenon. In addition to providing effective and practical methods for the solution of a number of problems which have faced considerable difficulty under the classical theory, statistical theory in the present state of development has already indicated the need and the method for recasting certain accepted theories. It has also indicated the possibility of new and more effective systems of transmission, reception, and detection.

In a communication system, varying quantities, such as currents or voltages, distributed in time, are processed during their passage through the system for the purpose of producing some desired result. Thus these functions, which are usually continuous, carry information from the transmitter to the receiver. They may be periodic, aperiodic, or random. However, since a periodic wave does not maintain a continuous flow of information and since aperiodic functions of time are usually associated with transient phenomena, it is the random function which is of interest to the communications engineer since, if information is to be kept in a steady flow, the receiver has to be uncertain of forthcoming events so that what he or the machine receives is a series of selections made by the sender from a finite set of all possible choices. When the receiver has full knowledge of future events then whatever message he continues to receive actually contains no information. Thus it is clear that a function which represents a message should be of a random type and cannot therefore be described exactly for all possible cases but must instead be described by a probability distribution.

The problem of signal detectability is then to find some method of making a decision on whether the source of a time-varying function, which is observed for a prescribed interval of time, is noise or signal plus noise. Breaking down this problem further, it can be said that it is desired to find some criterion such that when a sample from the received waveform meets this criterion then a signal is present and when it does not meet the criterion then no signal is present. However, since probabilities are concerned here, it cannot be said with complete certainty that a signal is or is not present but only that there is a certain likelihood that, if the value of the sample meets the criterion, then a signal is present and if it does not meet it then noise alone is present.

Two probabilities are of particular interest:

(a) The probability of detection, i.e., the probability of saying that a signal is present when a signal is in fact present.

(b) The probability of false alarm, i.e., the probability of saying that a signal is present when in fact no signal is present.

Various criteria for making a decision have been suggested [1] and in every case it has been concerned with the above two probabilities;

in some cases the probability of detection having been maximized, in others the probability of false alarm having been minimized, but in the usual case an optimum value having been chosen. However, in a particular application, the design specification would determine the probabilities of detection and of false alarm which could be admitted and would indicate the criterion to be used.

This criterion may have several forms. In the trivial case it will only be a number which is given to the observer who will take samples of the time-varying function and then make a decision according as the value of the sample is greater than or less than this number. In the more practical case the form of the criterion will form the basis for the design of an instrument which will sample the waveform and render a decision quite objectively.

In a radar detection system the information is contained in the position of the signal whose presence is determined by one of its characteristics, such as frequency, phase, or amplitude. The noise is assumed to be stationary, band-limited with a uniform power spectrum over the complete band and with its amplitude peaks having a Gaussian distribution.

This paper will deal, in particular, with two different methods of obtaining detection criteria and will show how one of them, which is used in practice can, due to an assumption which is made without foundation, lead to quite erroneous results. The other method, which is purely theoretical and is based on existence probabilities, will indicate how a new method of detection might be developed. This latter method will also show that there is a probability maximum in the amount of information that may be extracted from an information-carrying waveform.

CHAPTER II

THEORY

A Method of Detection Based on Inverse Probability

Assume that a message function x, which is independent of time, is mixed with a fluctuating disturbance, which is independent of x, and denote the resulting waveform by y. Then the problem is to operate on y so as to extract as much of x as is possible. It is quite clear that the problem is not one of maximizing the information that may be obtained from y but that of conserving the information contained in x and eliminating the unwanted information in y.

Since x contains information, the various values that it may be expected to take may be described by a probability distribution function, say $p(x)$. Similarly, $p(y)$ may be used to describe the distribution of the various values that samples of the waveform y might have. Thus the problem may be written mathematically since all that is required is to find the value of the probability of detecting x when given y, i.e., $p(x/y)$. The ideal receiver may then be defined as something which, when given y at the input, will supply $p(x/y)$ at the output.

By the product law of probabilities:

$$p(x,y) = p(x)p(y/x) = p(y)p(x/y) \tag{1}$$

$$p(x/y) = \frac{p(x)p(y/x)}{p(y)}$$

$$= k\, p(x)\, p(y/x) \tag{2}$$

since the values of the samples of y are presumed to be known and therefore need not be described by a probability distribution.

Let the random noise voltage waveform be denoted by n(t). Then it may be shown [2] that this voltage has a Gaussian distribution and may be written:

$$G(n) = k \exp \left[- \frac{1}{N_o} \int_0^T \left[n(t) \right]^2 dt \right] \tag{3}$$

where N_o is the mean noise power per unit bandwidth and T is the time of observation.

Since:

$$y(t) = x + n(t) \tag{4}$$

where x and n(t) are assumed independent:

$$G(n) = k \exp \left[- \frac{1}{N_o} \int_0^T (y - x)^2 \, dt \right] \tag{5}$$

and, since the possible values of y when x is present are dependent on the noise:

$$G(n) = G(y - x) = p(y/x) \tag{6}$$

Hence, the a posteriori distribution of the message function is:

$$p(x/y) = k \, p(x) \exp \left[- \frac{1}{N_o} \int_0^T (y - x)^2 \, dt \right] \tag{7}$$

It is now seen that Equation (7) gives the probable amount of the x information which may be obtained from the waveform y. It has, however, greater significance since it specifies the conditions for the design of the optimum receiver. Thus all that is required is to maximize p(x/y) and design a receiver with a transfer function $p(x/y)_{max.}$ Then, when a waveform y is applied to the input of the instrument, the output will be $p(x/y)_{max.}$ times the information represented by x.

The outstanding obstacle to the use of this method is in the presence of the a priori distribution p(x). Since, in a radar system, it is a target, which may or may not be present, which determines the value of p(x), the designer of the receiver has no idea of the form to assign to p(x).

Faced with this difficulty, it may be assumed that all states of x are equally probable, but this is merely a mathematical way of expressing ignorance of what really happens. In addition, since there is no real basis for such an assumption, this is merely guessing and Woodward [3] has shown that guesswork destroys information.

However, making some assumption about the form of p(x) (and p(x)= a constant is the most obvious one to make) enables one to maximize p(x/y) and, for p(x) constant, this is equivalent to maximizing p(y/x).

Now, for white Gaussian niose, from Equation (7):

$$p(y/x) = k \exp\left[-\frac{1}{N_o}\int_0^T (y-x)^2 dt\right] \tag{8}$$

Since all signals are assumed equally likely and will thus have equal energies:

$$\exp\left[-\frac{1}{N_o}\int_0^T x^2 dt\right] = A \tag{9}$$

Similarly, since the noise power is assumed uniform over the whole spectrum, it follows that:

$$\exp\left[-\frac{1}{N_o}\int_0^T y^2 dt\right] = B \tag{10}$$

and A and B may be absorbed into the normalizing constant.

$$p(y/x) = k' \exp\left[\frac{2}{N_o}\int_0^T yx \, dt\right] \tag{11}$$

Since p(y/x) in Equation (9) is a single valued function all that is required is to evaluate $\int_0^T yx \, dt$ in order to derive the posterior distribution p(x/y). This function, $\int_0^T yx \, dt$, is called the cross-correlation between y and x and it is seen from Equation (11) that the most probable message state is the one which yields the largest positive cross-correlation.

This idea has been used at Massachusetts Institute of Technology [4] to develop a cross-correlation receiver which gives very satisfactory results, but, so far as the author knows, the instrument has not so far been used in a radar detection set.

A Method of Detection based on Existence Probability*

An alternative form for the detection probability may be developed by means of Bayes' Theorem of Inverse Probability [5] which states that the probability that a signal is present after a waveform has been received is proportional to the product of the a priori probability and the probability that the waveform would have occurred had a signal been present. In this method we consider the probability of the existence of a signal, i.e., the likelihood that a signal is present.

Since the characteristics of the signal and noise are known, it is possible to calculate the probability distributions for y when a signal is present and when it is absent, namely $p(y/S)$ and $p(y/0)$ respectively. Then if the a priori probabilities of the presence and absence of the signal are λ and μ $(=1-\lambda)$ respectively, it may be written by the Theorem of Inverse Probability (Bayes' Theorem), that the probability that a signal is present when some characteristic, y, of a waveform occurs is:

$$p(S/y) = \frac{\lambda\, p(y/S)}{\lambda\, p(y/S) + \mu p(y/0)} \tag{12}$$

---

*This is the author's interpretation of the application of Bayes' Theorem to the detection of signals in noise as given by Davies.[6] Many workers, especially at M.I.T., disagree strongly with this use of Bayes' Theorem on the grounds that it is fitting artificial constraints to a practical situation. In this work an attempt is made to remove some of the obscurities from Davies' paper in order that his method may be more easily understood by the non-mathematician. It should be noted that this is only one of several possible ways of interpreting the paper.

In the usual case, if the relation, $p(S/y) \geq \epsilon$ (where ϵ is some predetermined constant) holds, a signal is present and, if $p(S/y) < \epsilon$ no signal is present.

However, in a practical situation, λ and μ are not known exactly and estimates, $\hat{\lambda}$ and $\hat{\mu}$ must be made of them. Since those estimates are probability estimates they must contain some margin of error, and hence, the observer cannot compute $p(S/y)$ exactly. He can, however, make an estimate, $p(E/y)$, of $p(S/y)$ and if $p(E/y) \geq \epsilon$, he can state that a signal has a certain degree of likelihood of being present.

Thus, from Equation (10):

$$p(E/y) = \frac{\hat{\lambda}\, p(y/S)}{\hat{\lambda}\, p(y/S) + \hat{\mu} p(y/0)} \qquad (13)$$

Since this gives the probability of the existence of a signal when a particular y is considered, it is not of much use in estimating the performance of a system. However, if this expression is weighted by $p(y/S)$ and averaged over all possible values of y, the following mean existence probabilities are obtained:

$$\begin{aligned} \bar{p}(E/S) &= \int p(E/y)p(y/S)dy \\ &= \int \frac{\hat{\lambda}\left[p(y/S)\right]^2}{\hat{\lambda}\, p(y/S) + \hat{\mu}\, p(y/0)}\, dy \qquad (14) \end{aligned}$$

$$\begin{aligned} \bar{p}(E/0) &= \int p(E/y)p(y/0)dy \\ &= \int \frac{\hat{\lambda}\, p(y/S)p(y/0)}{\hat{\lambda}\, p(y/S) + \hat{\mu}\, p(y/0)}\, dy \qquad (15) \end{aligned}$$

Similarly we have that:

$$\bar{p}(N/S) = \int \frac{\hat{\mu}\, p(y/0)p(y/S)}{\hat{\lambda}\, p(y/S) + \hat{\mu}\, p(y/0)}\, dy \qquad (16)$$

$$\bar{p}(N/O) = \int \frac{\hat{\mu}\,[p(y/O)]^2}{\hat{\lambda}\, p(y/S) + \hat{\mu}\, p(y/O)}\, dy \tag{17}$$

$$\text{where } p(N/y) = \frac{\hat{\mu}\, p(y/O)}{\hat{\lambda}\, p(y/S) + \hat{\mu}\, p(y/O)}$$

$$= 1 - p(E/y)$$

From a consideration of fundamental probability relationships and Bayes' Theorem, the following relationships may be derived:

$$\bar{p}(E/S) + \bar{p}(N/S) = 1 \tag{18-a}$$

$$\bar{p}(E/O) + \bar{p}(N/O) = 1 \tag{18-b}$$

$$\hat{\lambda}\, \bar{p}(E/S) + \hat{\mu}\, \bar{p}(E/O) = \hat{\lambda} \tag{19-a}$$

$$\hat{\lambda}\, \bar{p}(N/S) + \hat{\mu}\, \bar{p}(N/O) = \hat{\mu} \tag{19-b}$$

$$\hat{\lambda}\, \bar{p}(N/S) = \hat{\mu}\, \bar{p}(E/O) \tag{20}$$

If it is assumed that the noise is Gaussian and has a uniform power spectrum over a band which is wider than that of the signal and if some characteristic, x(t), of the signal is completely known, it may be written as previously that:

$$p(y/S,x) = k \exp\left[-\frac{1}{N_o}\int_0^T (y-x)^2\, dt\right] \tag{8}$$

Then, substituting in Equation (10):

$$\bar{p}(E/y) = \frac{\hat{\lambda}\, k \exp\left[-\frac{1}{N_o}\int_0^T (y-x)^2\, dt\right]}{\hat{\lambda}\, k \exp\left[-\frac{1}{N_o}\int_0^T (y-x)^2\, dt\right] + \hat{\mu}\, k \exp\left[-\frac{1}{N_o}\int_0^T y^2 dt\right]} \tag{21}$$

$$= \frac{\hat{\lambda} \exp\left[\frac{2}{N_o} \int_0^T yx \, dt\right]}{\hat{\lambda} \exp\left[\frac{2}{N_o} \int_0^T yx \, dt\right] + \hat{\mu} \exp\left[\frac{\rho^2}{2}\right]} \tag{22}$$

where ρ^2 is the ratio of the received signal energy during the observation time, T, to the mean noise power per unit bandwidth.

This expression for $p(E/y)$ may be simplified by letting:

$$a = \frac{2}{N_o} \int_0^T yx \, dt \tag{23}$$

Thus:

$$p(E/y) = \frac{\hat{\lambda} \exp[a]}{\hat{\lambda} \exp[a] + \hat{\mu} \exp\left[\frac{\rho^2}{2}\right]} \tag{24}$$

and since 'a' may be shown to have a Gaussian distribution of zero mean and mean square value ρ^2, we may substitute this value for $p(E/y)$ in Equation (13) and get:

$$\bar{p}(E/0) = \int_{-\infty}^{\infty} \frac{\hat{\lambda} \exp[a]}{\hat{\lambda} \exp[a] + \hat{\mu} \exp\left[\frac{\rho^2}{2}\right]} \, \frac{\exp\left[\frac{-a^2}{2\rho^2}\right]}{\rho\sqrt{2\pi}} \, da \tag{25}$$

By means of relations (16), (17), and (18) we may obtain similar expressions for $\bar{p}(E/S)$, $\bar{p}(N/O)$, and $\bar{p}(N/S)$.

Davies [6] has carried this approach further by applying it to a modulated carrier of the form:

$$x(t) = v(t)\cos(wt + \phi) \tag{26}$$

He assumes that v(t) varies slowly compared to the carrier frequency and also that, except for the carrier phase ϕ, the form of x(t) is known precisely.

He then shows that:

$$p(E/y) = \frac{\dfrac{\hat{\lambda}}{2\pi}\displaystyle\int_0^{2\pi} \exp\left[\frac{2}{N_o}\int_0^T y(t)v(t)\cos(wt+\phi)\,dt\right]d\phi}{\dfrac{\hat{\lambda}}{2\pi}\displaystyle\int_0^{2\pi} \exp\left[\frac{2}{N_o}\int_0^T y(t)x(t)dt\right]d\phi + \hat{\mu}\exp\left[\frac{\rho^2}{2}\right]} \tag{27}$$

and by expanding the cosine term, integrating with respect to ϕ and then averaging over all values of y he obtains:

$$\bar{p}(E/0) = \int_0^{\infty} \frac{\hat{\lambda}\, I_o(R)}{\hat{\lambda}\, I_o(R) + \hat{\mu}\exp\left[\frac{\rho^2}{2}\right]} \; \frac{R\exp\left[-\frac{R^2}{2\rho^2}\right]}{\rho^2}\, dR \tag{28}$$

where $R = \sqrt{(a^2 + b^2)}$

$$a = \frac{2}{N_o} \int_0^T y(t)v(t)\cos wt \, dt$$

$$b = \frac{2}{N_o} \int_0^T y(t)v(t)\sin wt \, dt$$

$I_o =$ zero order Bessel function with imaginary argument

It is then possible to plot values of $\bar{p}(E/S)$ and $\bar{p}(N/O)$ against the ratio of received signal energy to mean noise power per unit bandwidth, i.e., E/N_o, for different values of $\hat{\lambda}$ and $\hat{\mu}$ (Figures 1 and 2) and also the receiver operating characteristics (Figure 3).

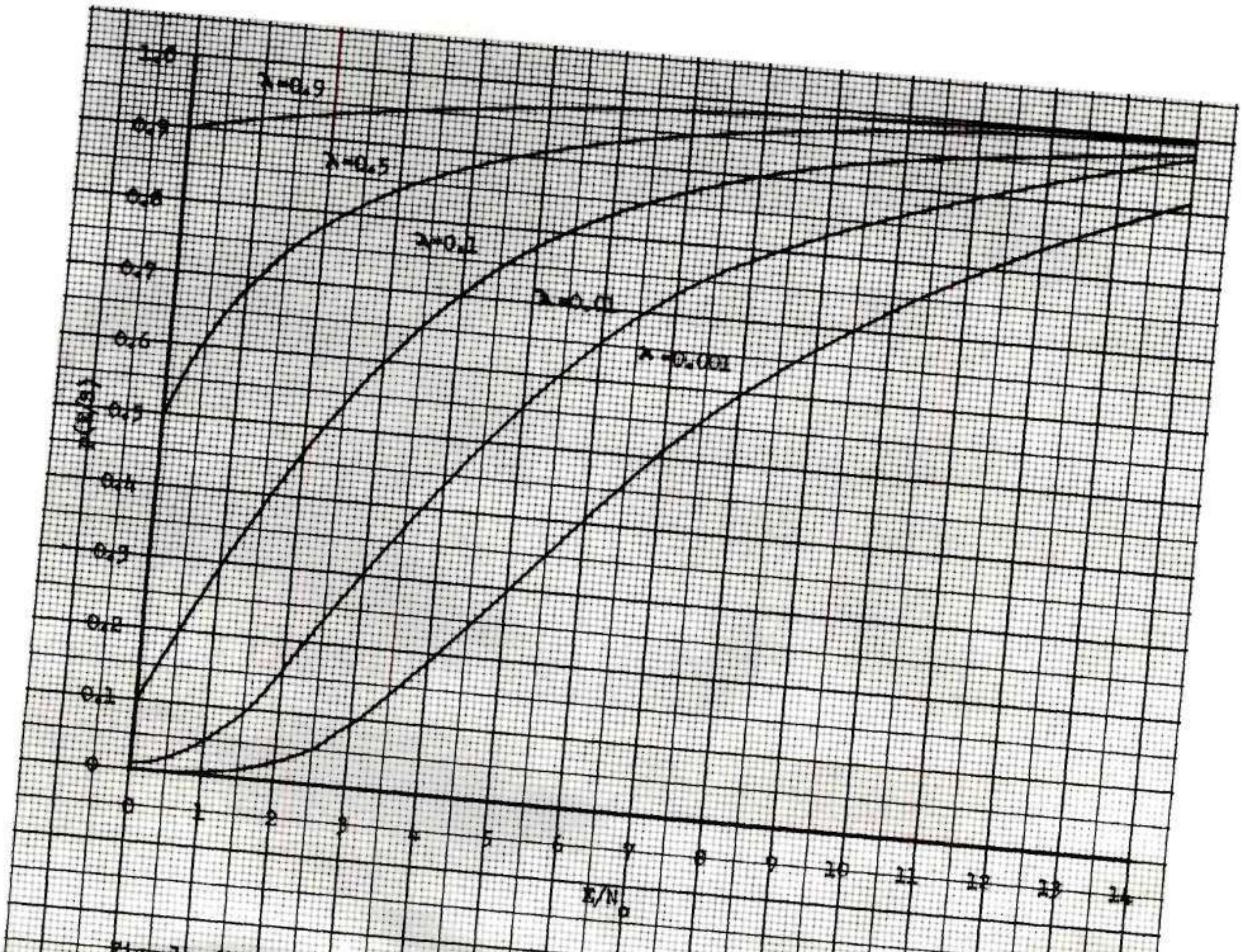


Fig. 1: Mean Existence Probability When Signal is Present.

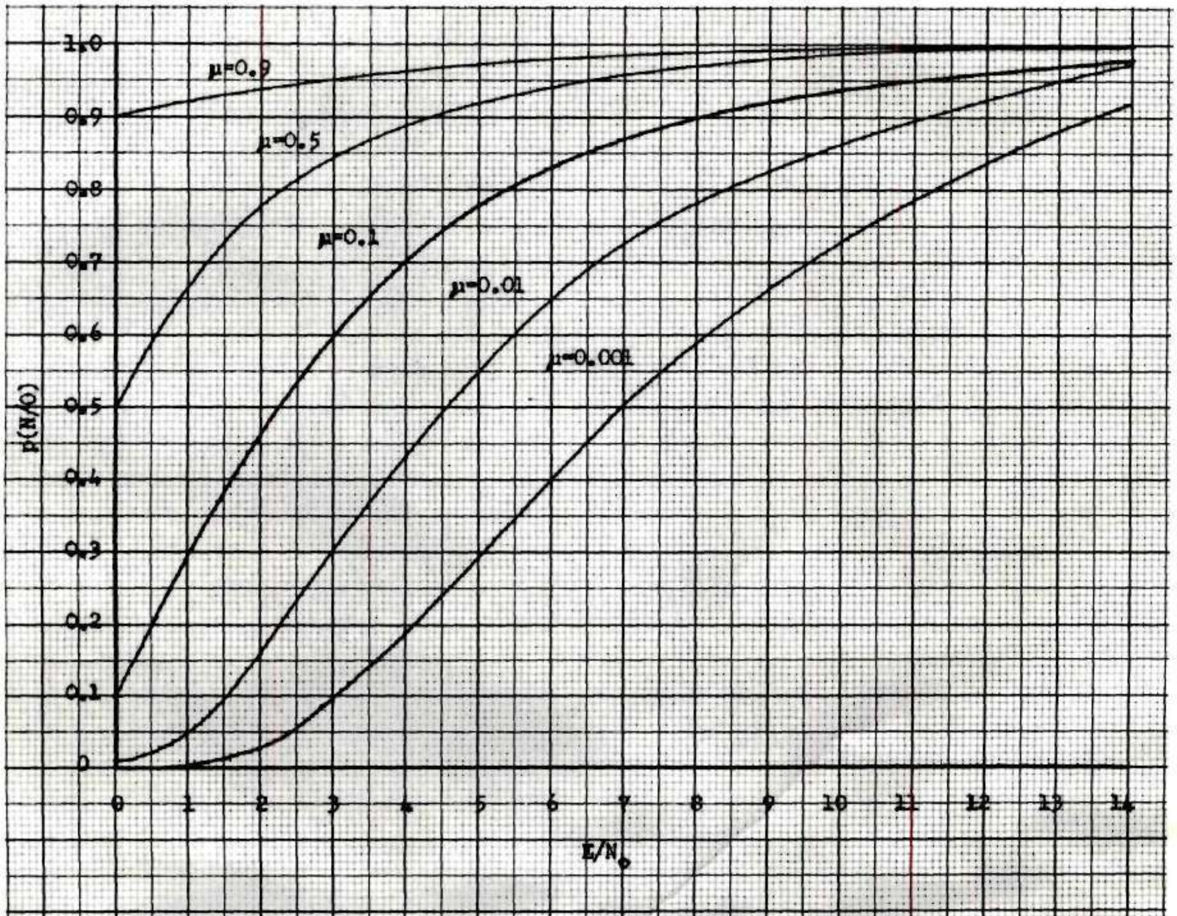


Fig. 2: Mean Non-existence Probability when No Signal is Present.

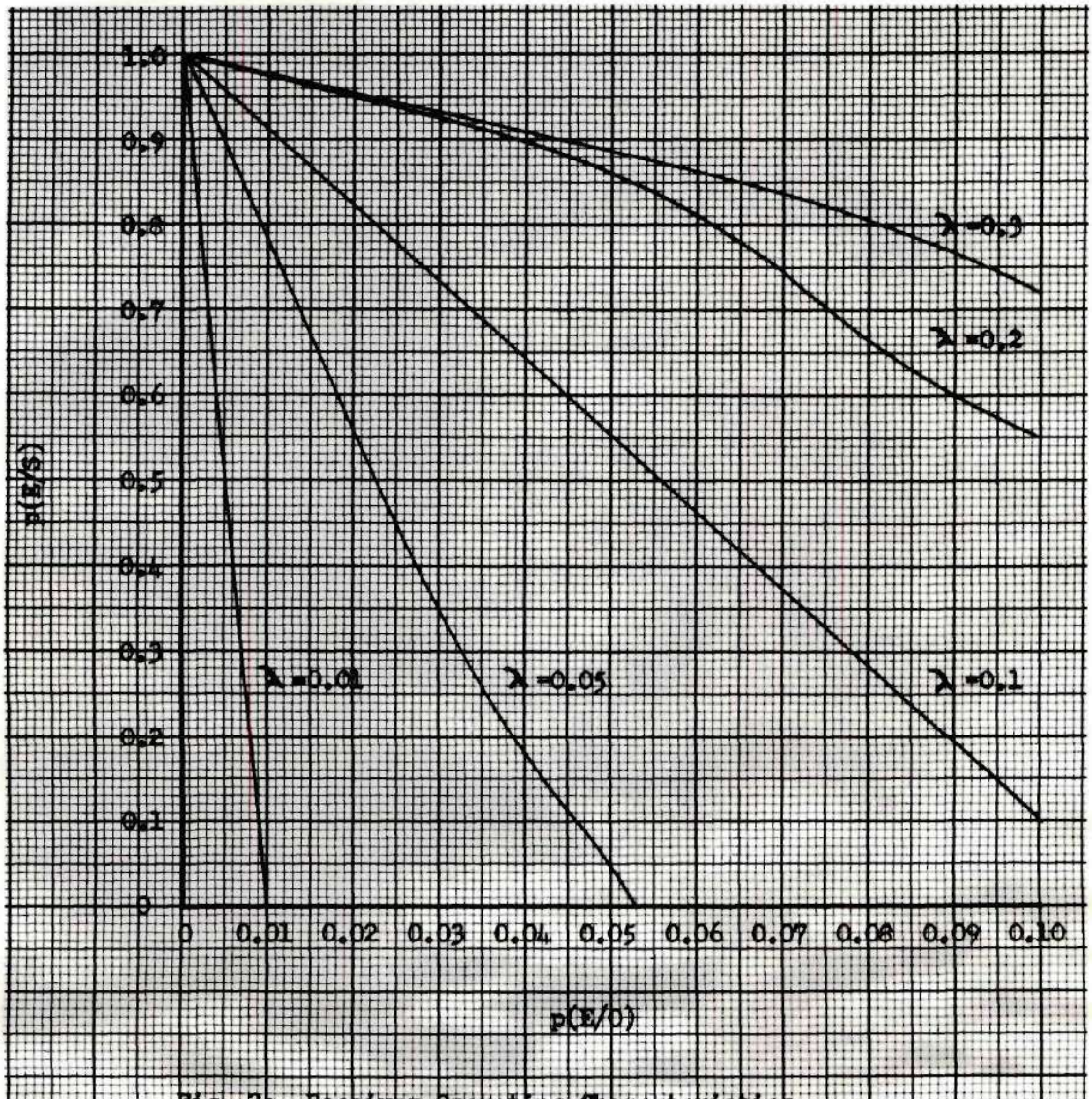


Fig. 3: Receiver Operating Characteristics.

Entropy as a Function of Probabilities

Since communication has to deal with the measurement of information received by an observer, there must be some means of measuring the observer's state of knowledge. Before reception each message X_i will have the probability $p(X_i)$ of occurring and afterwards one particular message X will have been singled out in the mind of the observer and the uncertainty described by its initial probability, $P(X)$, will be removed and information will have been gained. Mathematically this means that $P(X)$ increases to unity and the probabilities of all the other states diminish to zero. Thus the extent of the change may be measured in terms of $P(X)$. The prior probabilities of the states which failed to occur need not be considered individually but can be grouped together as having the probability $1 - P(X)$. It can now be postulated that the gain in information when two independent messages X and Y are received is the sum of the independent gains. For this case the joint probability $P(X,Y)$ is equal to $P(X)P(Y)$ and for the above assumption J must be chosen such that:

$$J\left[P(X)P(Y)\right] = J\left[P(X)\right] + J\left[P(Y)\right] \tag{29}$$

One form of J which would satisfy this identity is the logarithm and in order to make the gain in information positive, the following form is chosen:

$$J(P) = -\log P \tag{30}$$

However, in a practical communication system, the effects of random interference will make it impossible for the receiver to identify the message with complete certainty. Thus, in general, the probability of having received a particular message state X after some signal has been received is not unity but is, say, P'(X) and thus the received information is:

$$- \log P(X) - \left[- \log P'(X) \right] = \log \frac{P'(X)}{P(X)} \qquad (31)$$

This expression is the information gain of the system.

Instead of discrete message states, let a continuous function of possible messages, x, be considered and also the received information as a function, y. Then from Equations (1) and (29), the gain in information is:

$$\begin{aligned} I_{x,y} &= \log \frac{p(x/y)}{p(x)} \\ &= \log \frac{p(x,y)}{p(x)\ p(y)} \end{aligned} \qquad (32)$$

If 2 is taken as the logarithmic base, the unit of information is called a "bit" and Equation (32) is the basic expression for the quantity of information which is implicit in Shannon's theory. [7]

Equation (32) enables one to calculate the transfer of information for specific values of x and y but, in general, the observer will not know which x has caused y to occur and thus the observer's gain in information will be found by averaging $I_{x,y}$ over all the possible values of x which could have caused y to occur weighted with the relative probabilities of their occurrence.

Thus, for each y, the observer's gain in information is:

$$I_y = \int p(x/y) \log \frac{p(x,y)}{p(x)p(y)} \, dx \tag{33}$$

and by averaging over all possible values of y the average overall gain in information may be found. This is:

$$I = \int p(y) \int p(x/y) \log \frac{p(x,y)}{p(x)p(y)} \, dx \, dy \tag{34}$$

which may be written as:

$$\begin{aligned} I &= \int p(y) \int p(x/y) \log p(x/y) \, dx \\ &\quad - \iint p(x,y) \log p(x) \, dx \, dy \\ &= H(x) - H(x/y) \end{aligned} \tag{35}$$

$$\begin{aligned} \text{where } H(x) &= - \int p(x) \log p(x) \, dx \\ H(x/y) &= - \int p(y) \, dy \int p(x/y) \log p(x/y) \, dx \end{aligned}$$

These expressions H are the expressions for entropy as defined by Shannon [7] and from which he obtains the expression for the maximum amount of information which may be passed through a system of bandwidth W in a time T when the signal power is P and the noise power is N. This expression is:

$$I_{max} = W T \log \left[\frac{P+N}{N} \right] \tag{36}$$

CHAPTER III

DISCUSSION

In the first section of Chapter II an expression (Equation 7) was developed a basis for the design of an optimum receiver. This expression included the a priori probability of some characteristic of the signal which is, in general, unknown. As stated earlier, the form of this probability function is normally taken as a constant. This is done for two reasons; firstly, it indicates that nothing special happens anywhere in the distribution and this is the same as expressing complete ignorance of what happens, and, secondly, such a distribution function simplifies the maximizing of the a posteriori probability. The question now is, "Is this the most likely form of the distribution in an actual case?"

In the Appendix it is shown that this form implies comparatively good conditions for detection and this would indicate that the assumption that $p(x)$ is a constant is not the best one that could be made. It would in fact be better to utilize the Theory of Games in order to determine what the form of the enemy's (or nature's) worst strategy could be and to design the receiver on this basis.

The assumption that the a priori probability is a constant does not include the effects of the noise power on which, as shown in the Appendix, the most probable value of the signal characteristic depends.

Since the nearer the target is to the receiver, the more likely it is to be detected and the more pronounced are the effects produced in the receiver, it would appear that a more likely form of the distribution would be an increasing function of the characteristic which is being used for detection.

The theory as given in the second section of Chapter II provides better opportunities for the study of the design of detection systems than the method of the first section. As in the previous method, the a priori probabilities are included in the expression for the existence probabilities and their form would have to be assumed if it were desired to use this as a basis for the design of a receiver. The real usefulness of this form is in its use as a measure of the efficiency of a receiver. Thus the expression for $\overline{p}(E/S)$ gives the mean probability that the presence of a signal is indicated when a signal is in fact present. However it can also be thought of as being the probable fraction of the information in the received waveform that may be extracted. In the third section of Chapter II it was shown that there is a maximum amount of information that could be carried over a communication system. This relation was obtained by consideration of the probability distributions and the most probable capacity of the system was given by Equation (36). Thus, if a receiver is designed on the basis of maximizing the mean existence probability $\overline{p}(E/S)$ it will have a capacity depending on the bandwidth, the time of observation, and the signal-to-noise ratio as given in Equation (36), but the amount of information that could, on the average, be expected to be taken from the output would be only:

$$\overline{p}(E/S)\; W\; T \log \left[\frac{N + P}{N} \right]$$

Thus this method could be used to determine the output which could be expected when a known amount of information was put into a receiver. It could also be used as a basis for the comparison of different receivers.

Since the probability of indicating the presence of a signal when a signal is present will normally be greater than the probability of false alarm, we need only look for the existence of a signal in order to know whether or not it is present. However, instead of taking samples at intervals along each scan and averaging over these samples in order to find out if a signal occurred during that scan, only one sample need be taken from each of several scans and, if it has been arranged that each scan has the duration of the pulse period and each sample is taken at the same time interval after the pulse has been transmitted, one can now average over these samples and look for the existence of a signal. This would have the effect of taking time samples at a particular range. In order to cover the whole range, n samples could be taken over each pulse period, each sample being of T'/n seconds duration, where T' is the pulse period and T'/n is of the order of the pulse width. For greater accuracy, overlapping of the samples could be allowed, one series of samples being made at 0, 1, 2, ... and another series being made at 1/2. 3/2, 5/2, ... The theory given in the second section of Chapter II could then be used to find whether a signal existed in the i th sample, say. If the existence of a signal was indicated in this sample, then the target's position would be known since this sample would indicate the range.

Thus this theory may be used to develop the idea that it is not necessary to know anything about the characteristics of a signal in order to detect it. Instead all that need be known is that a signal exists and with this information the source may be located.

CHAPTER IV

CONCLUSIONS

It has been indicated that the assumption that all possible values of the signal characteristic are equally probable is not a good one to make when designing a receiver since it implies the existence of comparatively good conditions for detection. In practice a receiver should be designed on a basis of the worst conditions to be expected and thus $p(x)$ should be chosen such that when there is a large value of noise power per unit bandwidth it is indicated that, for best detection, a large value of x is required.

However, since no statistics exist for estimating the best function to use, and since all functions will, in many cases, lead to quite erroneous results, the decision on the form of the function to be used should depend on how it will affect the complexity of the receiver design.

In the previous work for the message function x has been assumed to be independent of time which really makes the case considered a trivial one. Such an assumption was made in order to illustrate the theory and yet keep the mathematics simple. An attempt was made to maximize $p(x/y)$, with x a function of time, by means of the Calculus of Variations, but since there were too many unknown parameters present, the mathematics became extremely complex and since finding a maximum of $p(x/y)$ does not indicate anything about the spread of values about the maximum, and is therefore of doubtful utility, the work was not carried through to a stage such that any conclusions could be drawn.

Although the relations for the existence of a signal also involve the a priori distributions, the ideas expressed in the theory indicate that a theoretical comparison of receivers could be made on an efficiency basis. The theory also suggests that targets could be detected on the basis of the mere presence or absence of a signal and not on some characteristic such as amplitude or phase.

APPENDIX

To Show the Effect of Assuming Different Forms of

the a priori Distribution Function, $p(x)$

The a posteriori probability distribution function is:

$$F = p(x/y) = k\, p(x) \quad \exp\left[-\frac{1}{N_o}\int_0^T (y - x)^2 dt\right]$$

Assume that, due to physical effects, the greatest value of x possible is M and the smallest is -M.

(1) $$p(x) = A; \quad A = \frac{1}{2M}$$

In this case,

$$F = kA \exp\left[-\frac{1}{N_o}\int_0^T (y - x)^2\, dt\right]$$

To find the most probable value of x take the logarithm of F, differentiate with respect to x and equate to zero.

$$\therefore \qquad \text{Log } F = \text{Log } kA - \frac{1}{N_o}\int_0^T (y - x)^2 dt$$

$$\therefore \qquad \frac{d \operatorname{Log} F}{dx} = \frac{2}{N_o} \int_0^T (y - x)\, dt$$

$$= 0 \quad \text{for a turning value.}$$

$$\therefore \qquad \hat{x} = \frac{1}{T} \int_0^T y\, dt$$

Since the second derivative is negative F is a maximum for this value of $\hat{x}$.

(2) $$p(x) = Ax + B; \quad B = \frac{1}{2M}, \quad o < A \leq \frac{1}{2M^2}$$

$$F = k(Ax + B) \exp\left[-\frac{1}{N_o} \int_0^T (y - x)^2 dt \right]$$

$$\operatorname{Log} F = \operatorname{Log} k + \operatorname{Log}(Ax + B) - \frac{1}{N_o} \int_0^T (y - x)^2 dt$$

$$\frac{d \operatorname{Log} F}{dx} = \frac{A}{Ax + B} + \frac{2}{N_o} \int_0^T (y - x) dt$$

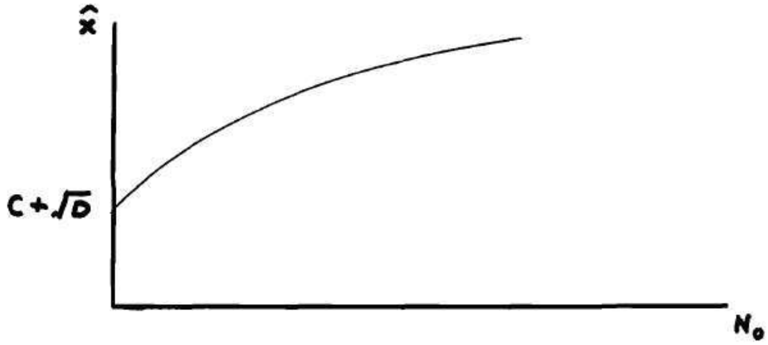
$$= 0 \text{ for a turning value.}$$

$$ATx^2 + (BT - A \int_0^T y\, dt)x - B \int_0^T y\, dt - \frac{AN_o}{2} = 0$$

$$\therefore\ \hat{x} = \frac{1}{2T}\int_0^T y\,dt - \frac{B}{2A} \pm \sqrt{\left[\frac{1}{2T}\int_0^T y\,dy - \frac{B}{2A}\right] \quad \frac{N_o}{2T} \quad \frac{B}{2A}\int_0^T y\,dt}$$

$$= C + \sqrt{D + \frac{N_o}{2T}}$$

By plotting values of F against the corresponding values of $\hat{x}$ we find that a maximum occurs at the upper value of $\hat{x}$.



Form of curve of $\hat{x}$ plotted against N_o

It is noted that, for this form of p(x), the most probable value of $\hat{x}$ depends on the noise power and increases as the noise power increases.

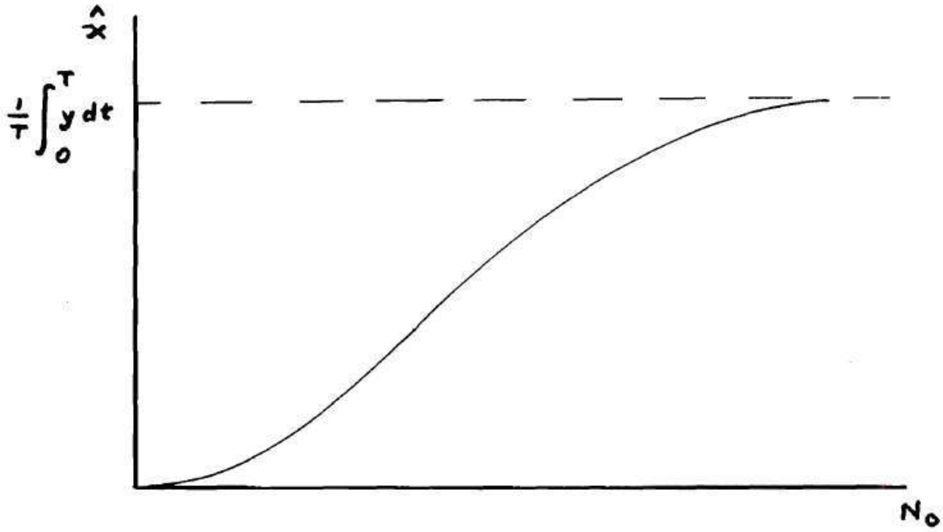
(3) $\qquad p(x) = C \exp\left[-x^2/2\right]$

$$F = C \exp\left[-\frac{x^2}{2} - \frac{1}{N_o}\int_0^T (y - x)^2\,dt\right]$$

$$\therefore \quad \frac{d \operatorname{Log} F}{dx} = -x + \frac{2}{N_o}\int_0^T (y - x)\,dt$$

$$= 0 \text{ for a turning value.}$$

$$\therefore \qquad \hat{x} = \frac{1}{\frac{2}{N_o} + T} \int_0^T y \, dt$$



Form of curve of $\hat{x}$ plotted against N_o.

As in the previous case the most probable value of $\hat{x}$ increases with an increase in the noise power per unit bandwidth. It is noted that, as the value of N_o becomes very large, the probable value of $\hat{x}$ tends to the value obtained when p(x) is assumed constant.

(4) $\qquad p(x) = A \, x^{\alpha} \exp\left[-x\right]$

$$F = k \, A \, x^{\alpha} \exp\left[-x - \frac{1}{N_o}\int_0^T (y - x)^2 dt\right]$$

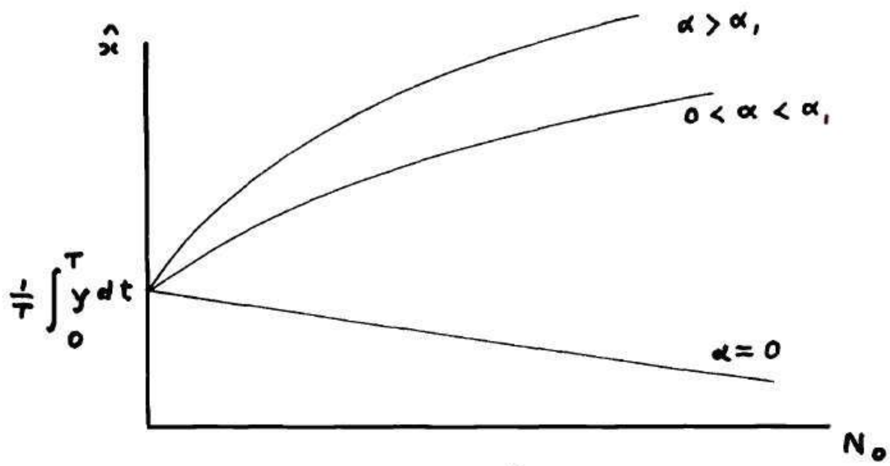
$$\therefore \qquad \frac{d \operatorname{Log} F}{dx} = \frac{\alpha}{x} - 1 + \frac{2}{N_o}\int_0^T (y - x) \, dt$$

$= 0$ for a turning value.

$$\therefore \quad \frac{2T}{N_0}x^2 + 1 - \frac{2}{N_0}\int_0^T y\,dt - \alpha = 0$$

$$\therefore \quad \hat{x} = \frac{-N_0}{4T} + \frac{1}{2T}\int_0^T y\,dt \pm \sqrt{\left[\frac{N_0}{4T} - \frac{1}{2T}\int_0^T y\,dt\right]^2 + \frac{N_0}{2T}\alpha}$$

By plotting values of F against the corresponding values of $\hat{x}$ we see that the maximum occurs for the upper value of $\hat{x}$.



Form of the curve of $\hat{x}$ plotted against N_0.

As in the previous two cases, the most probable value of $\hat{x}$ depends on N_0 when p(x) is assumed to have the form $Ax^{\alpha}\exp\left[-x\right]$ and the value of $\hat{x}$ increases as N_0 is increased.

Cases (2), (3), and (4) all indicate what was intuitively obvious, that is, as the noise power per unit bandwidth is increased then the most probable value of $\hat{x}$ must also increase. However, in case (1), the most probable value of $\hat{x}$ is constant for all values of N_0 and thus, if p(x) is assumed constant, too conservative an estimate is made of the value of $\hat{x}$ for large values of N_o and a receiver designed on this basis would be designed for comparatively good conditions. Thus, since a receiver should be designed for the worst possible conditions, the form of $p(x)$ should not be $p(x) = A$, but should be some other form such that the expected value of x increases as N_o increases.

BIBLIOGRAPHY

1. Peterson, W. W., Birdsall, T. G., Fox, W. G., "The Theory of Signal Detectability." Proceedings of the Institute of Radio Engineers, 1954, Part III, P. G. I. T.

2. Goldman, S., Information Theory, New York: Prentice Hall, Inc., 1953, Chapter III.

3. Woodward, P. M., Probability and Information Theory, New York: McGraw Hill, Inc., Chapter III.

4. Lee, Y. W., M.I.T. Research Laboratory Technical Report, Cambridge, Mass., No. 181, 1950.

5. Jeffreys, H., Theory of Probability, Oxford: University Press, 1948.

6. Davies, I. L., "On Determining the Presence of Signals in Noise." Proceedings of the Institution of Electrical Engineers, 99, (1952) Part 3.

7. Shannon, C. E., "A Mathematical Theory of Communication." Bell System Technical Journal, Vol. 27, (1948).