Optimal Weights for a Vector of Independent Poisson Random Variables[1]

Joo-Hwan Kim[2]

Abstract

Suppose one is given a vector X of a finite set of quantities X_i which are independent Poisson random variables. A null hypothesis H_0 about $E(X)$ is to be tested against an alternative hypothesis H_1. A quantity

$$\sum_i \omega_i x_i$$

is to be computed and used for the test. The optimal values of ω_i are calculated for three cases: (1) signal to noise ratio is used in the test, (2) normal approximations with unequal variances to the Poisson distributions are used in the test, and (3) the Poisson distribution itself is used. The above three cases are considered to the situations that are without background noise and with background noise. A comparison is made of the optimal values of ω_i in the three cases for both situations.

Keywords : Poisson signal, Hypothesis testing, Background noise, Power, Optimal weights,

1. Introduction

A beam of neutral particles can be used to estimate the density or mass of an object (Feller (1970)). A method of discrimination here is to use a neutral particle beam(NPB) aimed at the object, and a small number of neutron signals are counted at the detector. Beyer and Qualls (1987) showed that the number of return neutron particles from an object interrogation for a given dwell time follow Poisson distributon. Kim([4], [5], [6] and [7]) studied an application of signal processing of return neutron signals from an object irradiated by a neutral particle beam in the case of one probe-one object-one detector.

To extend the previous studies, we consider the case of one probe-one object- k detectors. Independent Poisson counts in k bins that have different means in each bin are considered. The objective is also to discriminate between a re-entry vehicle (RV) and a decoy using the

---

return signal. Figure 1 shows the situation we consider in this study. This discrimination problem is formulated as a test of hypothesis:

$$H_0 : \text{object is an RV} \quad \text{vs.} \quad H_1 : \text{object is a decoy}$$

(The theory of testing hypotheses is given in [8].) The observations without background noise are formed into a vector of neutron counts in several energy bins: $x = (x_1, x_2, \cdots, x_k)$.

The signal X_i are independent Poisson random variables. Let

$$EX = (t_1, t_2, \cdots, t_k) \text{ under } H_0$$

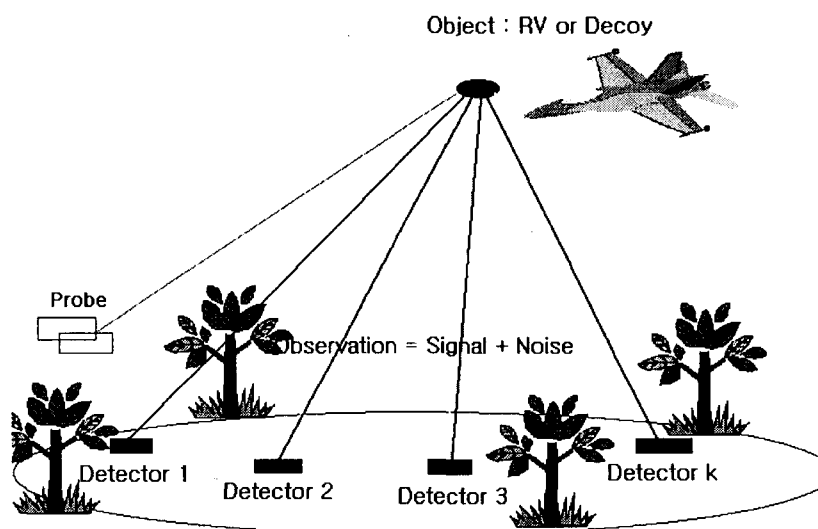$$EX = (d_1, d_2, \cdots, d_k) \text{ under } H_1$$

Then

$$H_0 : EX = t$$

$$H_1 : EX = d \qquad d_i < t_i \text{ for all } i$$

To discriminate an object using k signals from k detectors, we consider a summary statistic that is a linear combination of X_i

$$Y = \sum_{i=0}^{k} \omega_i X_i$$

where the weight ω_i are to be chosen later. Since the ω_i are to be chosen positive (see (2.2), (3.3), and (4.1) below), we reject H_0 if $Y \leq$ some critical value c. Three methods of choosing the ω_i are given and compared.

Object : RV or Decoy

Probe

Observation = Signal + Noise

Detector 1   Detector 2   Detector 3   Detector k

<Fig. 1> One probe-one object- k detectors case

The first method is based on a signal-to-noise ratio S/N. Maximizing S/N usually is

intended to maximize the power of the statistical test, defined by the criterion for rejecting H_0. In the present situation the variance of Y (the test statistic) is not the same under H_0 and H_1; consequently, maximizing S/N and maximizing power are not equivalent. The second method of choosing weight is based on maximizing power for normal distribution approximations with unequal variances under H_0 and H_1. The third method maximizes power for the exact Poisson distributions. Note that we only consider the optimal weighting of the linear combinations of Poisson counts, but it has not been verified here as being able to discriminate the hypothesis at reasonable risks α and β.

2. Signal-to-noise

Since the X_i are independent Poisson random variables, the mean is

$$E(Y) = \sum_{i=1}^{k} \omega_i t_i$$

under H_0, but

$$E(Y) = \sum_{i=1}^{k} \omega_i d_i$$

under H_1. The variance $V(Y)$ is

$$V(Y) = \sum_{i=1}^{k} \omega_i^2 t_i, \quad \text{under} \quad H_0$$

$$V(Y) = \sum_{i=1}^{k} \omega_i^2 d_i, \quad \text{under} \quad H_1$$

In the theory of testing hypotheses concerning means μ_0 and μ_1 with common variance σ^2, the power (probability of rejecting $H_0: \mu = \mu_0$ when $H_1: \mu = \mu_1$) is an increasing function of the signal-to-noise ratio(i.e. mean difference between H_0 and H_1 over common standard deviation)

$$\frac{|\mu_0 - \mu_1|}{\sigma}.$$

This suggests that in the present case of the hypothesis testing, one might choose ω_i to maximize

$$\frac{S}{N} = \frac{\sum_{i=1}^{k} \omega_i (t_i - d_i)}{\sqrt{\sum_{i=1}^{k} \omega_i^2 d_i}} \tag{2.1}$$

where the denominator of (2.1) was chosen to the standard deviation of Y under H_1. Since we assume the standard deviation under H_0 is greater than or equal to the one under H_1, the maximum of signal-to-noise is obtained when we use the standard deviation of Y under

H_1 as denominator.

By the Cauchy-Schwarz inequality(Rohatgi [12] p.165), we have

$$\sum_{i=1}^{k}(\omega_i\sqrt{d_i})\frac{t_i-d_i}{\sqrt{d_i}} \leq \sqrt{\left(\sum_{i=1}^{k}\omega_i^2 d_i\right)}\sqrt{\left(\sum_{i=1}^{k}\frac{(t_i-d_i)^2}{d_i}\right)}$$

with equality holding if and only if

$$\omega_i\sqrt{d_i} = K\frac{t_i-d_i}{\sqrt{d_i}}, \quad i=1,2,\cdots,k,$$

for some constant K. In other words, the signal-to-noise ratio is maximized by the choice

$$\omega_i = \frac{t_i-d_i}{d_i} \tag{2.2}$$

Since any constant multiple of the ω_i also maximizes S/N, the ω_i of (2.2) can be rescaled so that $\sum_{i=1}^{k}\omega_i = 1$.

3. Normal Approximation with Unequal Variances

Assume that the independent Poisson distributions $P(\lambda_i)$ of the X_i can be approximated by normal distributions. Then, approximately,

$$X_i \sim N(\lambda_i, \sigma_i^2)$$

where $\lambda_i = \sigma_i^2 = t_i$ or $\lambda_i = \sigma_i^2 = d_i$ according to whether H_0 or H_1 is true. Also, approximately,

$$Y \sim N(\sum_{i=1}^{k}\omega_i\lambda_i, \sum_{i=1}^{k}\omega_i^2\lambda_i).$$

The detection rate is $1-\alpha$, where

$$\alpha = P(reject\ H_0\ when\ H_0\ is\ true)$$

$$= P(Y \leq c \mid E(X) = t)$$

$$\cong P\left(Z \leq \frac{c-\sum_{i=1}^{k}\omega_i t_i}{\sqrt{\sum_{i=1}^{k}\omega_i^2 t_i}}\right)$$

and $Z \sim N(0,1)$. This is equivalent to

$$\frac{c-\sum_{i=1}^{k}\omega_i t_i}{\sqrt{\sum_{i=1}^{k}\omega_i^2 t_i}} = \Phi^{-1}(\alpha) = -\Phi^{-1}(1-\alpha), \tag{3.1}$$

where Φ is the cumulative distribution function of the standard normal random variable Z.
   The false alarm rate is

$$\beta = P(\text{accept } H_0 \text{ when } H_1 \text{ is true})$$

$$= P(Y > c \mid E(X) = d)$$

We assume $d_i > 0$. Therefore, the d_i will remain in the analysis. So

$$1 - \beta \cong P\left(Z \leq \frac{c - \sum_{i=1}^{k} \omega_i d_i}{\sqrt{\sum_{i=1}^{k} \omega_i^2 d_i}}\right)$$

or

$$\frac{c - \sum_{i=1}^{k} \omega_i d_i}{\sqrt{\sum_{i=1}^{k} \omega_i^2 d_i}} = \Phi^{-1}(1 - \beta) \tag{3.2}$$

The combination of (3.1) and (3.2), for fixed α, shows that satisfies

$$\Phi^{-1}(1 - \beta) = \frac{\sum_{i=1}^{k} \omega_i (t_i - d_i) - \Phi^{-1}(1 - \alpha)\sqrt{\sum_{i=1}^{k} \omega_i^2 t_i}}{\sqrt{\sum_{i=1}^{k} \omega_i^2 d_i}} \tag{3.3}$$

The quantity $\sum_{i=1}^{k} \omega_i (t_i - d_i)$ is the "excess" signal under hypothesis H_0 over that under H_1. Note that the first term in (3.3):

$$\frac{\sum_{i=1}^{k} \omega_i (t_i - d_i)}{\sqrt{\sum_{i=1}^{k} \omega_i^2 d_i}}$$

is the signal-to-noise ratio and is just (2.1).

It is required to find $\{\omega_i\}$ that maximize the function $h(\{\omega_i\})$ defined by the right-hand side of (3.3). One can assume that not all $\omega_i = 0$, since otherwise $\alpha = 1$. One first assumes that $\omega_1 = 0$ and investigates all maxima of the resulting function of $k-1$ variables. Then assume $\omega_1 \neq 0$ and divide numerator and denominator of (3.3) by ω_1 and put $\overline{\omega}_i = \omega_i / \omega_1$, $i = 2, 3, \cdots, k$. Then $h(\{\omega_i\})$ becomes a function $k-1$ of variables: $h(\{\overline{\omega}_i\})$. A sufficient condition (see Luenberger [9]) that the point $h(\{\overline{\omega}_i^*\})$ be a strict local maximum for h is that $\nabla h(\{\overline{\omega}_i^*\}) = 0$ and that the matrix $\nabla^2 h(\{\overline{\omega}_i^*\})$ be negative definite. This condition can be verified by checking that the eigenvalues of the matrix $\nabla^2 h(\{\overline{\omega}_i^*\})$ are negative. Finally, having checked all strict local maxima, it is necessary to insure that the function does not become elsewhere greater than its value at one of the strict local maxima. There are other possibilities that must be checked such as nonstrict maxima.

This procedure will be illustrated in only one case: $k=2$. Assume $\omega_1 \neq 0$. Put $x = \overline{\omega_2}$ $= \omega_2/\omega_1$. Then

$$h(x) = \frac{x(t_2 - d_2) + (t_1 - d_1) - \Phi^{-1}(1-a)\sqrt{x^2 t_2 + t_1}}{\sqrt{x^2 d_2 + d_1}} \tag{3.4}$$

After calculating $h'(x) = 0$, transposing a square root and squaring both sides, a quadratic equation in x is obtained:

$$[d_2(t_2 - d_1)x - d_1(t_2 - d_2)]^2(t_2 x^2 + t_1) - (d_1 t_2 - t_1 d_2)^2 \Phi^{-1}(1-a)x^2 = 0. \tag{3.5}$$

In the special case that $d_1 = 1$, $d_2 = 2$, $t_1 = 3$, $t_2 = 4$, $\Phi^{-1}(1-a) = 4$, equation (3.5) reduces to $16x^4 - 16x^3 - 12x + 3 = 0$ which has two real roots, only one of which gives a value zero to h' : $x = 1.3998\cdots$ and $h''(1.3998\cdots) < 0$. The condition $\omega_1 + \omega_2 = 1$ finally gives $\omega_1 = .4274\cdots$, $\omega_2 = .5726\cdots$. Since $h(0) < 0$ and $h(\infty) = -\sqrt{2}$, there is no other maximum value of $h(x)$ to the special case.

For the case $k=3$, h in (3.4) becomes bivariate function of $\overline{\omega_2^*}$ and $\overline{\omega_3^*}$ and we may find the maximum point for some specified values of the parameters. For general k, h becomes $k-1$ variate function of $\overline{\omega_2^*}$, $\overline{\omega_3^*}$, $\cdots$, $\overline{\omega_k^*}$.

4. Optima Weights for Poisson Distributions

For Poisson counts in one energy bin, Beyer and Qualls [1] give a rather complete analytical analysis. For two energy bins, the discrimination surface, analogous in the discrimination curve in Kim [6] is a mapping of R^2 to R^2. For this analysis, one needs to develop one or more test statistics. In this section we give an optimal test statistic based on the observation that to minimize β is to maximize the power $1 - \beta$. We seek the most powerful (MP) test of the hypothesis H_0. The Neyman-Pearson lemma [8, p.74] computes the MP test in terms of a

$$rejection\ region = \left\{ x \mid \frac{p(x,d)}{p(x,t)} > K \right\},$$

where

$$\frac{p(x;d)}{p(x;t)} = \frac{\prod_{i=1}^{k} d_i^{x_i} e^{-d_i}/x_i!}{\prod_{i=1}^{k} t_i^{x_i} e^{-t_i}/x_i!} = \exp\left\{ -\sum_{i=1}^{k} x_i \log\frac{t_i}{d_i} \right\} \prod_{i=1}^{k} e^{(t_i - d_i)}$$

is the likelihood ratio. The MP test has the form

$$Reject\ H_0\ if\ y = \sum_{i=1}^{k} \omega_i x_i \leq c,$$

where

$$\omega_i = \log \frac{t_i}{d_i}$$
(4.1)

Again, we can rescale the ω_i so that $\sum_{i=1}^{k} \omega_i = 1$.

5. Optimal Weights with Background Poisson Noise

We have considered the situation in section 1 to section 4 without background Poisson noise. In this section we consider the case with background Poisson noise. The observations are formed into a vector of neutron signals and noise counts in several energy bins:

$$x = (x_1^s + x_1^n, \ x_2^s + x_2^n, \ \cdots, \ x_k^s + x_k^n).$$

where the signal X_i^s and background noise X_i^n are independent Poisson random variables. Let

$$EX^s = \begin{cases} (t_1, t_2, \cdots, t_k) & \text{under } H_0 \\ (d_1, d_2, \cdots, d_k) & \text{under } H_1 \end{cases}$$

$$EX^n = (n_1, n_2, \cdots, n_k) \quad \text{under both } H_0 \text{ and } H_1$$

Then

$$H_0 : \ EX = EX^s + EX^n = t + n$$

$$H_1 : \ EX = EX^s + EX^n = d + n \ , \qquad d_i < t_i \ \text{for all } i$$

For a test statistic $Y = \sum_{i=1}^{k} \omega_i X_i$, the mean $E(Y)$ is $\sum_{i=1}^{k} \omega_i (t_i + n_i)$ and $\sum_{i=1}^{k} \omega_i (d_i + n_i)$ under H_0 and H_1, respectively. The variance $V(Y)$ is $\sum_{i=1}^{k} \omega_i^2 (t_i + n_i)$ under H_0, and $\sum_{i=1}^{k} \omega_i^2 (d_i + n_i)$ under H_1.

Signal-to-noise ratio with background Poisson noise becomes

$$\frac{S}{N} = \frac{\sum_{i=1}^{k} \omega_i (t_i - d_i)}{\sum_{i=1}^{k} \omega_i^2 (d_i + n_i)}$$
(5.1)

and the optimal choice that maximize the signal-to-noise is

$$\omega_i = \frac{t_i - d_i}{d_i + n_i}.$$
(5.2)

For a test with normal approximation, similar to the procedure in section 3, we have

$$\frac{c - \sum_{i=1}^{k} \omega_i (t_i + n_i)}{\sqrt{\sum_{i=1}^{k} \omega_i^2 (t_i + n_i)}} = \Phi^{-1}(\alpha) = -\Phi^{-1}(1 - \alpha),$$
(5.3)

and

$$\frac{c - \sum_{i=1}^{k} \omega_i (d_i + n_i)}{\sqrt{\sum_{i=1}^{k} \omega_i^2 (d_i + n_i)}} = \Phi^{-1}(1 - \beta) \tag{5.4}$$

The combination of (5.3) and (5.4), for fixed α, shows that satisfies

$$\Phi^{-1}(1 - \beta) = \frac{\sum_{i=1}^{k} \omega_i (t_i - d_i) - \Phi^{-1}(1 - \alpha)\sqrt{\sum_{i=1}^{k} \omega_i^2 (t_i + n_i)}}{\sqrt{\sum_{i=1}^{k} \omega_i^2 (d_i + n_i)}} \tag{5.5}$$

For the case $k = 2$, as we had demonstrated in section 3, we have

$$h(x) = \frac{x(t_2 - d_2) + (t_1 - d_1) - \Phi^{-1}(1 - \alpha)\sqrt{x^2(t_2 + n_2) + t_1 + n_1}}{\sqrt{x^2(d_2 + n_2) + d_1 + n_1}} \tag{5.6}$$

Put $\overline{t_1} = t_1 + n_1$, $\overline{t_2} = t_2 + n_2$, $\overline{d_1} = d_1 + n_1$ and $\overline{d_2} = d_2 + n_2$. Then (5.6) becomes

$$h(x) = \frac{x(\overline{t_2} - \overline{d_2}) + (\overline{t_1} - \overline{d_1}) - \Phi^{-1}(1 - \alpha)\sqrt{x^2 \overline{t_2} + \overline{t_1}}}{\sqrt{x^2 \overline{d_2} + \overline{d_1}}} \tag{5.7}$$

and equation (5.7) is same form to (3.4).

For a test with exact Poisson distribution, similar to the procedure in section 4, the MP test has the form

$$Reject \ H_0 \ \text{if} \ y = \sum_{i=1}^{k} \omega_i x_i \leq c,$$

where

$$\omega_i = \log \frac{t_i + n_i}{d_i + n_i} \tag{5.8}$$

6. Comparison and Remark

We now have three methods of calculating weights: (1) S/N, (2) normal approximations, and (3) Poisson MP test. In this section we compare the optimal weights for three methods.

First of all, comparison is made in the limit for large Poisson counts without background noise. Let $t_i = p_i d_i$ with $p_i > 1$ for all i. Then consider the limit of the ω_i for the three methods of this paper as the decoy counts become infinite, i.e. as $\min(d_i) \to \infty$.

For the S/N method, we obtain

$$\omega_i = \frac{t_i - d_i}{d_i} = \frac{t_i}{d_i} - 1 = p_i - 1 \ \to \ p_i - 1 \quad as \ min(d_i) \to \infty \tag{6.1}$$

For the Poisson MP test method, we obtain

$$\omega_i = \log\left(\frac{t_i}{d_i}\right) = \log(p_i) \;\rightarrow\; \log(p_i) \quad as \; min(d_i) \rightarrow \infty. \tag{6.2}$$

For the normal approximation method, we further specialize the limiting process so that the decoy counts $d_i = \nu d_i^0$ for $d_i^0 \geq 0$ for all i with $\nu \rightarrow \infty$. Expression (3.3) becomes

$$\sqrt{\nu}\,\frac{\sum \omega_i p_i d_i^0}{\sqrt{\sum \omega_i^2 d_i^0}} - \Phi^{-1}(1-\alpha)\frac{\sqrt{\sum \omega_i^2 p_i d_i^0}}{\sqrt{\sum \omega_i^2 d_i^0}} \tag{6.3}$$

The second term of (6.3) for various choices of the ω_i is bounded above (and below), since the $min(d_i^0) > 0$. Consequently, as $\nu \rightarrow \infty$, the first term dominates and the ω_i that maximize (6.3) converge to the S/N weights; i.e. $\omega_i \rightarrow (p_i - 1) \; as \; \nu \rightarrow \infty$

Secondly, we compare the optimal weights with background noise as the amount of signal and noise becomes infinite. Let $d_i = 0$ and $n_i > 0$ for all i, and $t_i = p_i n_i$ with $p_i \geq 0$ for all i. Then, similar to the above setup, we obtain the limit for the S/N method,

$$\omega_i = \frac{t_i}{n_i} = p_i \;\rightarrow\; p_i \quad as \; min(n_i) \rightarrow \infty \tag{6.4}$$

For the Poisson MP test method, we obtain

$$\omega_i = \log\left(\frac{t_i + n_i}{d_i + n_i}\right) = \log(1 + p_i) \;\rightarrow\; \log(1 + p_i) \quad as \; min(n_i) \rightarrow \infty \tag{6.5}$$

For the normal approximation method, the ω_i that maximize

$$\sqrt{\nu}\,\frac{\sum \omega_i p_i n_i^0}{\sqrt{\sum \omega_i^2 n_i^0}} - \Phi^{-1}(1-\alpha)\frac{\sqrt{\sum \omega_i^2 (1 + p_i) n_i^0}}{\sqrt{\sum \omega_i^2 n_i^0}} \tag{6.6}$$

converge to the S/N weights; i.e. $\omega_i \rightarrow p_i \; as \; \nu \rightarrow \infty$, where $n_i = \nu n_i^0$ for all i.

It is interesting that the normal approximations to the Poisson distributions implicitly included in (3.3) and (6.3) become better as t_i and d_i become large $(\nu \rightarrow \infty)$ but the ω_i that maximize S/N do not converge to the Poisson MP test weights. Same properties can be found to the case with background Poisson noise.

References

[1] Beyer, W. A. and Qualls, C. R. (1987). Discrimination with Neutral Particle Beams and Neutrons, *Los Alamos National Laboratory, LA-87-3140.*

[2] Feller, W. (1970). *An Introduction to Probability Theory and Its Applications: Volume II,* Third Ed., John Wiley & Sons.

[3] Graves, R. E. (1986). Signal-to-background noise ratio optimization for an NPB neutron detector with energy measurement capability, *Los Alamos National Lab., LA-UR-86-3101.*

[4] Kim, Joo-Hwan (1995). Properties of the Poisson-power Function Distribution, *The Korean Communications in Statistics*, Vol. 2, No. 2, pp. 166-175

[5] Kim, Joo-Hwan (1996). Error Rate for the Limiting Poisson-power Function Distribution, *The Korean Communications in Statistics*, Vol. 3, No. 1, pp. 243-255, 1996.

[6] Kim, Joo-Hwan (1997). The Minimum Dwell Time Algorithm for the Poisson Distribution and the Poisson-power Function Distribution, *The Korean Communications in Statistics*, Vol. 4, No. 1. pp. 229-241.

[7] Kim, Joo-Hwan(1998). Monotone Likelihood Ratio Property of the Poisson Signal Distribution with Three Sources of Errors in the Parameter, *The Korean Communications in Statistics*, Vol. 5, No. 2, pp. 503-515.

[8] Lehmann, E. L. (1986). *Testing Statistical Hypotheses*, 2nd Edition, John Wiley & Sons.

[9] Luenberger, D. G. (1973). *Introduction to Linear and Nonlinear Programming*, Addison-Wesley Publishing Company, 1973.

[10] Lohatgi, V. K. (1976). An Introduction to Probability Theory and Mathematical Statistics, John Wiley & Sons.