Optimal Weights for Independent Poisson Variables with Background Poisson Noise

Joo-Hwan Kim¹⁾, Ki-Hak Hong²⁾, Ki-Seong Nam³⁾, Chang-Kyoon Son⁴⁾

¹⁾ Dongguk University, Department of Statistics and Information Science,(jhk@mail.dongguk.ac.kr)
 ²⁾ Dongshin University, Department of Computer Science, (khhong@blue.dongshinu.ac.kr)
 ³⁾ Changwon University, Department of Education (kisnam@dreamx.net)
 ⁴⁾ Dongshin University, Department of Computer Science, (ckson85@blue.dongshinu.ac.kr)

Abstract

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

$$\sum_{i} \omega_{i} x$$

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. A comparison is made of the limit values of ω_i for large signal and noise in the three cases.

1. Introduction

Independent Poisson counts in k bins that have different means in each bin are considered. The application is signal processing of return neutron signals from an object irradiated by a neutral particle beam(Kim [3] and [4]). The objective is to discriminate between a re-entry vehicle (RV) and a decoy using the return signal in the presence of background Poisson noise. This discrimination problem is formulated as a test of hypothesis:

 H_0 : object is an RV vs. H_1 : object is a decoy

(The theory of testing hypotheses is given in [5] and [7].) The observations are formed into a vector of neutron counts in several energy bins: $x = (x_1, x_2, \dots, x_k)$, where $x_i = X_i^s + X_i^n$.

The signal X_i^{s} and background noise X_i^{n} are independent Poisson random variables. Let

$$EX^{s} = \frac{(t_{1}, t_{2}, \dots, t_{k}) \text{ under } H_{0},}{(d_{1}, d_{2}, \dots, d_{k}) \text{ under } H_{1},}$$

and

 $EX^n = (n_1, n_2, \dots, n_k)$ under both H_0 and H_1

Then

$$H_0: EX = t + n H_1: EX = d + n \qquad d_i < t_i \text{ for all } i$$

Consider a summary statistic that is a linear combination of $X_i = X_i^s + X_i^n$,

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

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.

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.

2. Signal-to-noise

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

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

under H_0 , but

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

under H_1 . The variance V(Y) is

$$V(Y) = \sum_{i=1}^{k} \omega_i^2(t_i + n_i), \text{ under } H_0$$
$$V(Y) = \sum_{i=1}^{k} \omega_i^2(d_i + n_i), \text{ under } H_1$$

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 $\mu = \mu_1$) is an increasing function of the signal-to-noise ratio

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

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

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

If $d_i = 0$ for all *i*, then S/N is indeed the signal-to-noise ratio. R.E. Graves [2] gives the following proof that (2.2) below is the optimal choice of the ω_i .

By the Cauchy-Schwarz inequality, we have

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

with equality holding if and only if

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

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

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

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

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

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

It is no additional difficulty to allow $d_i > 0$. Therefore, the d_i will remain in the analysis. One may set $d_i \equiv 0$ whenever desired. The detection rate is $1 - \alpha$, where

$$\begin{aligned} \alpha &= P(reject H_0 \text{ when } H_0 \text{ is true}) \\ &= P(Y \le c | E(X) = t + n) \\ &\cong P \Biggl(Z \le \frac{c \cdot \sum_{i=1}^k \omega_i(t_i + n_i)}{\sqrt{\sum_{i=1}^k \omega_i^2(t_i + n_i)}} \Biggr), \end{aligned}$$

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

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

where Φ is the cumulative distribution function of the standard normal random variable Z. The false alarm rate is $\beta = P(accept H_0 when H_1 is true)$

$$= P(Y > c | E(X) = d + n)$$

So

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

or

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

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

$$\Phi^{-1}(1-\beta) = \frac{\sum_{i=1}^{k} \omega_i (t_i - d_i) - \Phi^{-1}(1-\alpha) \sqrt{\sum_{i=1}^{k} \omega_i^2 (t_i + n_i)}}{\sqrt{\sum_{i=1}^{k} \omega_i^2 (d_i + n_i)}}$$
(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 . If H_1 is regarded as

noise plus a small decoy signal, then the first term in (3.3):

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

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 $\overline{\omega}_i = \omega_i / \omega_1$, $i = 2, 3, \dots, k$. Then $h(\{\omega_i\})$ becomes a function of k-1 variables $h(\{\overline{\omega}_i\})$. A sufficient condition (see [7]) that the point $\{\overline{\omega_i}^*\}$ 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

$$h(x) = \frac{x(t_2 - d_2) + (t_1 - d_1) - \Phi^{-1}(1 - \alpha)\sqrt{x^2(t_2 + n_2) + t_1 + n_1}}{\sqrt{x^2(d_2 + n_2) + d_1 + n_1}}$$
(3.4)

Put $\bar{t}_1 = t_1 + n_1$, $\bar{t}_2 = t_2 + n_2$, $\bar{d}_1 = d_1 + n_1$ and $\bar{d}_2 = d_2 + n_2$. Then (3.4) becomes

$$h(x) = \frac{x(\bar{t}_2 - \bar{d}_2) + (\bar{t}_1 - \bar{d}_1) - \Phi^{-1}(1 - \alpha)\sqrt{x^2\bar{t}_2 + \bar{t}_1}}{\sqrt{x^2\bar{d}_2 + \bar{d}_1}}$$
(3.5)

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

$$[\overline{d}_{2}(\overline{t}_{2}-\overline{d}_{1})x-\overline{d}_{1}(\overline{t}_{2}-\overline{d}_{2})]^{2}(\overline{t}_{2}x^{2}+\overline{t}_{1})-(\overline{d}_{1}\overline{t}_{2}-\overline{t}_{1}\overline{d}_{2})^{2}\Phi^{-1}(1-\alpha)x^{2}=0.$$
(3.6)

In the special case that $\overline{d_1} = 1$, $\overline{d_2} = 2$, $\overline{t_1} = 3$, $\overline{t_2} = 4$, $\Phi^{-1}(1-\alpha) = 4$, (3.6) reduces to $16x^4 - 16x^3 - 12x + 3 = 0$ which has two real roots, only one of which gives a value zero to h': $x = 1.3398\cdots$ and $h''(1.3398\cdots) < 0$. 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).

4. Optimal Weights for Poisson Distributions

For Poisson counts in one energy bin defined by a threshold, Kim [3] give a rather complete analytical analysis. For two energy bins, the discrimination surface, analogous in the discrimination curve in [4] 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 [5, p.74] computes the MP test in terms of a

rejection region =
$$\{x | \frac{p(x; d+n)}{p(x; t+n)} > K\}$$

where

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

is the likelihood ratio. The MP test has the form Reject H_0 if $y = \sum_{i=1}^{k} \omega_i x_i \le c$, where

$$\omega_i = \log \frac{t_i + n_i}{d_i + n_i} \tag{4.1}$$

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

5. Comparison of Limit Optimal Weights

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 three methods as the amount of signal and noise becomes infinite.

Comparison is made in the limit for large signal and noise. Let $d_i = 0$ for all i, and $t_i = \lambda_i n_i$ with $\lambda_i \ge 0$ for all i. Then consider the limit of the ω_i for the three methods of this paper as the noise becomes infinite, i.e. as $\min(n_i) \to \infty$.

For the S/N method, we obtain

$$\omega_i = \frac{t_i}{n_i} = \lambda_i \to \lambda_i, \quad as \min(n_i) \to \infty$$
(5.1)

For the Poisson MP test method, we obtain

$$\omega_i = \log(\frac{t_i + n_i}{n_i}) = \log(1 + \lambda_i) \to \log(1 + \lambda_i) \quad \text{as} \quad \min(n_i) \to \infty .$$
(5.2)

For the normal approximation method, we further specialize the limiting process so that the noise $n_i = v n_i^0$ for $n_i^0 \ge 0$ for all *i* with $v \to \infty$. Expression (3.3) becomes

$$\sqrt{\nu} \frac{\sum \omega_{i} \lambda_{i} n_{i}^{0}}{\sqrt{\sum \omega_{i}^{2} n_{i}^{0}}} - \Phi^{-1} (1 - \alpha) \frac{\sqrt{\sum \omega_{i}^{2} (1 + \lambda_{i}) n_{i}^{0}}}{\sqrt{\sum \omega_{i}^{2} n_{i}^{0}}}$$
(5.3)

The second term of (5.3) for various choices of the ω_i is bounded above (and below), since the $\min(n_i^0) > 0$. Consequently, as $\nu \to \infty$, the first term dominates and the ω_i that maximize (5.3) converge to the S/N weights; i.e. $\omega_i \to \lambda_i$ as $\nu \to \infty$

It is interesting that the normal approximations to the Poisson distributions implicitly included in (3.3) and (5.3) become better as t_i and n_i become large $(\nu \to \infty)$ but that the ω_i that maximize (5.1) do not converge to the

Poisson MP test weights.

6. Remark

Note that the optimal weighting have not been verified here as being able to discriminate the hypothesis at reasonable risks α and β .

References

- [1] W. A. Beyer, and C. R. Qualls, Discrimination with Neutral Particle Beams and Neutrons, Los Alamos National Laboratory, LA-87-3140, 1987.
- [2] R. E. Graves, Signal-to-background noise ratio optimization for an NPB neutron detector with energy measurement capability, Los Alamos National Lab., LA-UR-86-3101, Sept. 1986.
- [3] J.H. Kim, Discrimination of Objects with Poisson Signals : Theory and Application, University of New Mexico, Ph.D. Dissertation, 1991
- [4] J.H. Kim, Error Rate for the Limiting Poisson-power Function Distribution, The Korean Communications in Statistics, Vol. 3, No. 1, pp. 243-255, 1996.
- [5] E. L. Lehmann, Testing Statistical Hypotheses, Second Edition, John Wiley and Sons, 1986.
- [6]. D. G. Luenberger, Introduction to Linear and Nonlinear Programming, Addison-Wesley Publishing Co., 1973.
- [7]. J. Neyman and E. S. Pearson, On the problem of the most efficient tests of statistical hypotheses, Phil. Trans. Roy. Soc. A 231, 287-337, 1933.