論　文

最適檢出器의 Threshold를 求하는 有用한 方法

正會員　韓　榮　烈*　　準會員　朴　文　榮**

An Useful Method for Evaluating the Threshold of the Optimum Detector

Young Yeul HAN*, *Regular Member*　Moon Young PARK**, *Associate Member*

要　約　充分統計量(sufficient statistics)이 존재할 때 最適檢出器의 threshold를 찾는 有用한 方法을 提示하고 예를 들어 說明하였다. 充分統計量이 존재하지 않을 때에도 적절한 檢定統計量(test statistics)을 선택함으로써 suboptimun 檢出器의 設計가 可能함을 보였다. 일반적으로 주어진 檢定統計量에 대한 threshold를 求하는 方法을 提示하고 그 結果가 이미 알려진 理論과 一致됨을 알 수 있다. 結論的으로 광범위한 應用을 전제하고 있다.

ABSTRACT　An useful method for evaluating the threshold of the optimum detector can be used if sufficient statistics exists. This was done by giving examples. The design of the suboptimum detector can be carried out by finding the threshold of the appropriate test statistics. The results are conformed with the existing theory and the method given above is applicable in extensive area of designing the detectors.

1. 序　論

充分統計量(sufficient statistics)은 確率分布函數의 모수(parameter)를 推定하는 모든 情報를 가지고 있으며 標本 $X_1, X_2, X_3, \cdots, X_n$의 函數로써 주어진다는 것은 잘 알려진 事實이다. 이러한 充分統計量은 推定問題뿐만 아니라 統計的推理 및 統計的假說問題에서도 널리 使用된다[1]. 充分統計量, $S=s(X_1, X_2, \cdots X_n)$은 確率變數(random variable)이므로 n개의 確率變數인 $X_1, X_2, \cdots X_n$을 하나의 確率變數로 축소함으로써 最適檢出器(optimum detector)의 誤率(error probability)을 計算하는데 有用하게 使用할 수 있다. 즉 S의 確率分布函數(probability density function)가 存在하면 最適檢出器의 threshold를 알므로써 誤率을 求할 수 있다. S의 確率分布가 해석적으로 求하기 난이한 때에도 우리가 檢定하려는 모수의 充分統計量이 存在하면 threshold를 얻을 수 있으며 이 threshold가 決定됨으로써 檢出器의 設計는 이미 이루어졌다고 볼 수 있다. 充分統計量은 주어진 確率分布의 모수에 대하여 항상 存在하는 것이 아니며 또한 充分統計量의 存在性을 數學的으로 입증하기가 곤란할 때도 있다.

本論文에서는 2進符號디지탈通信(binary code digital communication)에서 하나의 모수를 檢定할 경우 充分統計量을 使用하여 threshold를 求하는 有用한 方法을 提示하고 例를 들어 說明하였다. 充分統計量이 存在하지 않더라도 檢定(test)하려는 確率分布의 特性을 알므로써 이미 알려진 充分統計量을 利用하여 suboptimum detector를 設計할 수 있음을 보였다.

2. 充分統計量의 諸性質

充分統計量은 統計量중 特別한 種類의 統計量

* **漢陽大學校 工科大學 電子通信工學科
Dept. of Electronic Communication Engineering, Hanyang University, Seoul, 133 Korea
論文番號：82-08　(接受 1982.5.31)

이라고 볼 수 있다. 즉 充分統計量은 모수 θ 의 모든 情報를 標本値로서 얻을 수 있다는 뜻이 된다. $X_1, X_2, \cdots, X_n$이 確率分布 $f(\bullet;\theta)$에서 임의추출표본 (random sample)이면 充分統計量 S는 $X_1, X_2, \cdots X_n$의 함수 즉 $S=s(X_1, X_2, \cdots X_n)$으로 이루어지고 充分統計量을 求하는 方法은 다음과 같다[1]. $X_1, X_2, \cdots, X_n$이 確率分布 $f(\bullet;\theta)$에서의 任意抽出標本이고 $X_1, X_2, \cdots, X_n$의 多次元確率分布 (joint probability density)를 $f(X_1, X_2, \cdots, X_n;\theta)$라 하면 $f(X_1, X_2, \cdots, X_n;\theta)$가 다음과 같이 분류될 수 있고 $g(s(x_1, x_2, \cdots, x_n);\theta)$이 θ를 포함하고 $s(X_1, X_2, \cdots, X_n)$이 $X_1, X_2, \cdots, X_n$의 함수이면 S는 充分統計量이다[1]. 즉

$$f(x_1, x_2, \cdots, x_n;\theta)=g(s(x_1, x_2, \cdots, x_n);\theta)\times h(x_1, x_2, \cdots, x_n)=g(S;\theta)h(x_1, x_2, \cdots, x_n) \tag{1}$$

$h(x_1, x_2, \cdots, x_n)$는 θ를 포함하지 않고 임의추출표본 $x_1, x_2, \cdots, x_n$만의 함수이다. 상기의 경우는 單一充分統計量이 존재할 경우이고 充分統量이 다수 존재할 경우는

$$f(x_1, x_2, \cdots, x_n;\theta)\ g(s_1(x_1, x_2, \cdots, x_n), \cdots, s_i(x_1, x_2, \cdots, x_n);\theta)\times h(x_1, x_2, \cdots, x_n)$$
$$=g(s_1, s_2, \cdots s_i;\theta)\ h(x_1, x_2, \cdots, x_n) \tag{2}$$

$g(s_1, s_2, \cdots, s_i;\theta)$가 θ를 포함하고 $x_1, x_2 \cdots, x_n$의 함수로 이루어지면 $s_1, s_2, \cdots, s_i$는 充分統計量들이다[1]. 本論文에서는 2進符號디지탈通信에서 하나의 모수에 관심이 있으므로 앞으로 單一充分統計量에 대하여 論하기로 한다.

3. 充分統計量과 Threshold

단순통계적가설의 경우 檢定될 두 가설을 $H_0;\theta=\theta_0$, $H_1;\theta=\theta_1$이라 하면 베이어스 코스트 (Bayes cost)를 最低로 하는 最適檢定 (optimum test)은

$$\frac{\prod_{i=1}^{n} f(x_i|H_1)}{\prod_{i=1}^{n} f(x_i|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{3}$$

라 놓을 수 있으며[2] threshold η는 事前確率 (a priori probability), $p(H_1)$ 및 $p(H_0)$가 각각 50%이고 統計的가설 H_j가 진실일 때 H_i를 선택하는 코스트 (cost)를 c_{ij}라 하면 2進符號 디지탈通信에서는 $c_{00}=c_{11}=0$, $c_{10}=c_{01}=1$이라 놓을 수 있고 η는 1이 된다[2]. 즉 다음과 같이 쓸 수 있다.

$$\frac{\prod_{i=1}^{n} f(x_i|H_1)}{\prod_{i=1}^{n} f(x_i|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} 1 \tag{4}$$

檢定하려는 모수가 다른 確率分布 $f(x_i|H_1)$, $f(x_i|H_0)$가 同一한 充分統計量을 가지고 있고 $x_1, x_2, \cdots, x_n$이 相互獨立的이면 다음과 같이 充分統計量 S를 쓸 수 있다.

$$\frac{\prod_{i=1}^{n} f(x_i|H_1)}{\prod_{i=1}^{n} f(x_i|H_0)}=g(s(x_1, x_2, \cdots, X_n;\theta)\times h(x_1, x_2, \cdots, x_n)) \tag{5}$$

그러므로 檢定統計量 (test statistics)는 充分統計量으로 表示될 수 있음을 알 수 있다.

(4)의 式을 양변이 같다고 놓고 양변에 logarithm을 取하면

$$\sum_{i=1}^{n} \ln\frac{f(x_i|H_1)}{f(x_i|H_0)}=0$$

이 된다. 좌우변이 같기 위해서는 각 항이 0이 되어야 한다. 즉

$$\frac{\ln f(x_1|H_1)}{\ln f(x_1|H_0)}=0,\ \frac{\ln f(x_2|H_1)}{\ln f(x_2|H_0)}=0, \cdots,$$
$$\frac{\ln f(x_n|H_1)}{\ln f(x_n|H_0)}=0 \tag{6}$$

이 된다. 0이 되는 x_1의 값을 a_1, x_2의 값을 $a_2, \cdots$, x_n의 값을 a_n이라 하면 檢定統計量은 充分統計量으로 되어 있으므로 다음의 式이 成立된다.

$$s(x_1, x_2, \cdots, x_n) \underset{H_0}{\overset{H_1}{\gtrless}} s(a_1, a_0, \cdots a_n)=\lambda \tag{7}$$

$s(x_1, x_2, \cdots, x_n)$은 檢定하려는 모수의 充分統計量이며 同時에 檢定統計量이고 $s(a_1, a_2, \cdots, a_n)$는 函數 $s(x_1, x_2, \cdots, x_n)$에 $x_1=a_1$, $x_2=a_2, \cdots$, $x_n=a_n$을 代入하여 얻은 檢定統計量의 threshold λ의 값이다. 만일 $\eta\neq 1$인 경우는

$$\sum_{i=1}^{n} \ln\frac{f(x_i|H_1)}{f(x_i|H_0)}=\ln\eta \tag{8}$$

이 되며 좌우변이 등식이 되도록 x_i, $i=1, 2, 3, \cdots$, n의 값을 찾아 a_i, $i=1, 2, 3, \cdots, n$으로 놓고 (7)式에서 threshold를 찾으면 된다.

또는 다음과 같이 놓으면 a_i, $i=1, 2, \cdots, n$을 쉽게 찾을 수가 있다.

$$\sum_{i=1}^{n} \ln\frac{f(x_i|H_1)}{f(x_i|H_0)}=\ln\eta=\ln\eta+\ln 1+\cdots+\ln 1 \tag{9}$$

$\ln 1=0$이고 $(n-1)$개의 項으로 이루어져 있다

$$\ln\frac{f(x_1|H_1)}{f(x_1|H_0)}=\ln\eta,\ \ln\frac{f(x_2|H_1)}{f(x_2|H_0)}=\ln 1=0,$$

$$\cdots, \ln \frac{f(x_n | H_1)}{f(x_n | H_0)} = 0 \qquad (10)$$

라 놓고 a_i, $i=1, 2, \cdots, n$을 求할 수 있다. 또 각 항의 동일한 a_i의 값을 求하기 위하여서는 다음과 같이 놓을 수 있다.

$$\sum_{i=1}^{n} \ln \frac{f(x_i | H_1)}{f(x_i | H_0)} = n\left(\frac{\ln \eta}{n}\right) \qquad (11)$$

즉,

$$\ln \frac{f(x_i | H_1)}{f(x_i | H_0)} = \frac{\ln \eta}{n}, \quad i=1, 2, \cdots, n \qquad (12)$$

으로 동일한 a_i의 값을 求할 수 있다.

例로서 分散 σ^2을 갖는 正規分布函數 (normal density function)일 경우 平均値 θ를 檢定하는 2개의 가설을 다음과 같다고 하면,

$$H_1;\ \theta=\theta_0$$

$$H_0;\ \theta=0 \qquad (13)$$

平均値 θ에 대한 充分統計量은 $\sum_{i=1}^{n} x_i$가 되며 最適檢出器의 式은 다음과 같다[2]

$$\frac{\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp[-(x_i-\theta)^2/2\sigma^2]}{\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp[-x_i^2/2\sigma^2]} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \qquad (14)$$

$\eta=1$이라 놓고 a_i를 求하면 $a_i=\theta/2$, $i=1, 2, \cdots, n$으로 되며 $\eta \neq 1$일 경우 동일한 a_i의 값을 求하면

$$a_i = \frac{\sigma^2 \ln \eta}{\theta n} + \frac{\theta}{2}, \quad i=1, 2, \cdots, n \qquad (15)$$

$$\text{즉} \quad \sum_{i=1}^{n} x_i \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\sigma^2 \ln \eta}{\theta} + \frac{n\theta}{2} \qquad (16)$$

가 된다.

다음 H_1; $\sigma^2=\sigma_1^2$, H_0; $\sigma^2=\sigma_0^2$을 檢定할 경우 분산 σ^2의 充分統計量은 $\sum_{i=1}^{n} x_i^2$으로 주어진다. 상기 기술한 方法으로 $\eta=1$일 경우 a_i를 求하여 보면 $\sigma_0^2 > \sigma_1^2$일 때

$$a_i = \frac{2\sigma_0^2 \sigma_1^2}{\sigma_0^2 - \sigma_1^2}\left(\ln \frac{\sigma_0}{\sigma_1}\right), \quad i=1, 2, \cdots, n \qquad (17)$$

$$\sum_{i=1}^{n} x_i^2 \underset{H_0}{\overset{H_1}{\gtrless}} \eta \frac{2\sigma_0^2 \sigma_1^2}{\sigma_0^2 - \sigma_1^2}\left(\ln \frac{\sigma_0}{\sigma_1}\right) = \lambda \qquad (18)$$

가 된다. $\eta \neq 1$일 경우도 동일한 방법으로 求할 수 있다. $\sum_{i=1}^{n} x_i$와 $\sum_{i=1}^{n} x_i^2$의 確率分布는 求할 수 있고[3] 상기의 分布函數를 $f(S)$라 하고 threshold를 λ라고 하면 false alarm probability와 detection probability는 다음과 같다[2].

$$P_F = \int_{\lambda}^{\infty} f(S | H_0)\, dS \qquad (19)$$

$$P_D = \int_{\lambda}^{\infty} f(S | H_1)\, dS \qquad (20)$$

지금까지 充分統計量을 가지고 있는 正規分布를 예를 들어 說明하였다. 그러나 전에 말한 바와 같이 모든 分布가 檢定하려는 모수에 대하여 充分統計量을 보유하고 있지는 않다. 다음은 充分統計量을 찾을 수 없으나 確率分布의 特性을 감안하여 檢定統計量을 설정하여 suboptimum detector를 設計할 수 있음을 예를 들어 보이고 있다.

다음 雜音모델은 포아슨分布 (Poisson density)로서 解析할 수 있는 충격잡음과 가우스雜音 (Gaussian noize)를 합한 것으로 여러 人工器機에서 發射되는 전자파 잡음에 적합한 모델로서 간섭잡음이 受信機의 대역폭보다 작은 간섭에 적용되고 있다[4]. 이 分布函數는 다음의 式으로 주어진다[4],[5]

$$f(x) = \sum_{m=0}^{\infty} e^{-A} \frac{A^m}{m!\sqrt{2\pi\sigma_m^2}} e^{-x^2/2\sigma_m^2} \qquad (21)$$

$$\sigma_m^2 = \frac{m/A + B}{1+B} \qquad (22)$$

여기서 A와 B는 2개의 상수로서 모델 설정에 따라 다른 수치를 갖는다. 이 分布의 特徵은 A가 적을 때 충격잡음이 많은 모델에 적합하며 A가 크면 가우스分布가 된다. (13)式과 같이 平均値 θ에 대한 檢定을 하려 해도 充分統計量이 存在하지 않으므로 이 때 A가 크면 가우스 分布가 된다는 것을 감안하여 다음과 같이 sub-optimum detector의 檢定式으로 놓을 수가 있다.

$$\sum_{i=1}^{n} x_i \underset{H_0}{\overset{H_1}{\gtrless}} \frac{n\theta}{2} \qquad (23)$$

이 식은 A가 큰 값이어서 分布函數가 가우스分布가 되면 正確한 檢定式이다. 이 식은 解析的으로는 도저히 不可能한 것을 이 分布函數의 特性과 正規分布函數의 平均値의 充分統計量을 使用하여 상기 기술한 방법으로 平均値를 檢定하는 근사식을 얻을 수 있었다. 本論文에서 論한 方法으로 threshold를 얻을 수 있는 sub-optimum detector의 설계가 가능함을 보여주고 있다. 이 때는 檢定統計量設定에 더욱 주의를 기울여야 하겠다.

4. 結 論

本論文에서는 檢定하려는 모수의 單一充分統

計量이 존재할 경우 充分統計量을 檢定統計量으로 하여 threshold를 구하는 有用한 方法을 유도 제시하고 예를 들어 타당성을 검토하였다. 檢定統計量의 threshold를 구한다는 것은 檢定統計量의 分布函數를 알면 그 檢出器의 모든 特性을 아는 것이므로 threshold를 구하는 方法에 중점을 두었다. 또한 本硏究는 充分統計量이 존재치 않는 난이한 檢出器設計에 많은 도움이 되리라 믿는다.

謝 辭

本硏究는 1982年度大宇文化福祉財團 硏究費支援의 일부로 이루어진 것입니다.

参 考 文 献

(1) A. M. Mood, F. A. Graybil and D. C. Boes "Introduction to the theory of statistics" New York, McGraw-Hill 1974.

(2) H. L. Van Trees, "Detection, estimation, and modulation theory, Part I" New York; John Wiley and Sons, 1968.

(3) H. Cramer, "Mathematical methods of statistics" Princeton New York; Princeton University Press, 1951.

(4) D. Middleton, "Statistical physical models of man made radio noise, Part I", First order probability models of the instanteneous amplitude" Office of telecommunications, Report 74－36, April 1974.

(5) A. D. Spaulding, and D. Middleton "Optimum reception in an impulsive interference environment Part I: Coherent detection" IEEE Trans on Commu., vol COM－25, no. 9, pp. 910－923, Sept. 1977.

韓 榮 烈 (Young Yeul HAN) 正會員
1938年 6月10日生
1960年 2月：서울工大電子工學科卒業
1976年 8月：Missouri州立大學大學院 通信專攻(工學碩士)
1979年 8月：Missouri州立大學大學院 通信專攻(工學博士)
1961年 8月～1964年 8月：西獨Ziemens會社(株)에서 電子分野硏修
1964年 8月～1969年11月：한영공업(株)근무.
1969年11～1970年10月：韓國科學技術硏究所勤務
1980年 8月～現在：漢陽大學校電子通信工學科副敎授
美國Sigma Xi 및 IEEE正會員
本學會理事

朴 文 榮 (Moon Young PARK) 準會員
1957年 6月25日生
1981年 2月：漢陽大學校電子通信工學科 卒業(B. S)
1981年 3月～現在：漢陽大學校大學院碩士過程 在學中
IEEE學生會員