ON TESTING CERTAIN HYPOTHESES

Yukio Suzuki

(Received May 30, 1967)

1. Introduction

Consider the following data generating process: For a non-negative integer r

$$\tilde{x}_i=\alpha_0\tilde{\theta}_i+\alpha_1\tilde{\theta}_{i-1}+\cdots+\alpha_r\tilde{\theta}_{i-r}+\tilde{\varepsilon}_i \qquad i=0,1,\cdots \tag{1.1}$$

where (i) $\tilde{\theta}_{-r},\cdots,\tilde{\theta}_0,\tilde{\theta}_1,\cdots,\tilde{\theta}_n,\cdots$ is a Bernoulli process with a parameter p, i.e. $\tilde{\theta}_i$'s are independently and identically distributed with a binomial distribution $B(1,p)$, (ii) $\alpha_0,\alpha_1,\cdots,\alpha_r$ are real and assumed to be known and (iii) $\{\tilde{\varepsilon}_i\}_{i=0,1,\ldots}$ is a sequence of random variables which are independently and identically distributed with the normal distribution $N(0,\sigma^2)$. Also it is assumed that $\{\tilde{\theta}_i\}$ and $\{\tilde{\varepsilon}_i\}$ are independent. (It is clear from (1.1) that $\tilde{x}_i$ and $\tilde{x}_j$ are independent if $|i-j|>r$.) Further, we assume that $\{\tilde{x}_i\}$ is observable but $\{\tilde{\theta}_i\}$ and $\{\tilde{\varepsilon}_i\}$ are not observable, the realized values of these random variables remaining always unknown. Then, our problem can be stated as follows: On the basis of the observed values $\{x_i\}_{i=0,\ldots,n}$, how can we detect whether the realized value, θ_n, of $\tilde{\theta}_n$ was 0 or 1? This problem is considered as a special case of the general problem which was treated by M. Tainiter [3]. However, it will be worthwhile treating the model (1.1), because the model seems to be prevailing in various fields of application. Also this will serve as a good example of the empirical Bayes approach ([1], [2]).

2. Bayes solution when p and σ^2 are known

From the data generating process (1.1), it is easily seen that the process $(\tilde{x}_n,\tilde{\theta}_n)_{n=0,1,\ldots}$ is strictly stationary. Also the stochastic process $\{\tilde{x}_n\}_{n=0,1,2,\ldots}$ is marginally strictly stationary. The definition of marginal stationarity is given by Tainiter ([3] Definition 2.2).

The conditional joint distribution function of $\{\tilde{x}_{k_i}\}_{i=1,\ldots,s}$ given specified values of all random parameters which appear in the model of each $\tilde{x}_{k_i}$ $i=1,\cdots,s$ is given by

$$\begin{aligned}(2.1)\qquad & F(x_{k_1},\cdots,x_{k_s}\mid\boldsymbol{\theta}_{k_1},\cdots,\boldsymbol{\theta}_{k_s})\\ &=\left(\frac{1}{2\pi\sigma^2}\right)^{s/2}\int_{-\infty}^{x_{k_1}}\cdots\int_{-\infty}^{x_{k_s}}\exp\left[-\frac{\sum_{i=1}^{s}(x_{k_i}-\boldsymbol{\theta}_{k_i}\boldsymbol{\alpha})^2}{2\sigma^2}\right]dx_{k_s}\cdots dx_{k_1}\end{aligned}$$

where we assume $k_1<k_2<\cdots<k_s$ and define

$$(2.2)\qquad \begin{aligned}&\boldsymbol{\theta}_{k_i}=(\theta_{k_i},\theta_{k_i-1},\cdots,\theta_{k_i-r})\qquad i=1,\cdots,s\\ &\boldsymbol{\alpha}=(\alpha_0,\alpha_1,\cdots,\alpha_r)'\,.\end{aligned}$$

However, since $\boldsymbol{\theta}_{k_1},\cdots,\boldsymbol{\theta}_{k_s}$ may have common elements, the following expression will be more appropriate. Let us define $\boldsymbol{\theta}_{(k_1,\ldots,k_s)}$ as the vector which is obtained by arranging the elements of $\boldsymbol{\theta}_{k_1},\cdots,\boldsymbol{\theta}_{k_s}$ in the order of magnitude of the subscripts without duplication. Further, let $d(k_1,\cdots,k_s)$ denote the dimension of $\boldsymbol{\theta}_{(k_1,\ldots,k_s)}$. Then, the left-hand side of (2.1) can be written as $F(x_{k_1},\cdots,x_{k_s}\mid\boldsymbol{\theta}_{(k_1,\ldots,k_s)})$. Now, writing $\boldsymbol{\theta}_{(k_1,\ldots,k_s)}=(\theta_1,\theta_2,\cdots,\theta_{d(k_1,\ldots,k_s)})$, the probability function of $\tilde{\boldsymbol{\theta}}_{(k_1,\ldots,k_s)}$ is given by

$$(2.3)\qquad g_d(\boldsymbol{\theta}_{(k_1,\ldots,k_s)})=\begin{cases}p^{\boldsymbol{\theta}_{(k_1,\ldots,k_s)}\mathbf{1}_d}(1-p)^{d-\boldsymbol{\theta}_{(k_1,\ldots,k_s)}\mathbf{1}_d} & \text{for } \boldsymbol{\theta}_{(k_1,\ldots,k_s)}\in S\,,\\ 0 & \text{otherwise}\,,\end{cases}$$

where

$$S=\{(\theta_1,\cdots,\theta_d)\mid\theta_i=0\text{ or }1,\ 1\leqq i\leqq d=d(k_1,\cdots,k_s)\}$$

and

$\mathbf{1}_d=d(k_1,\cdots,k_s)$-dimensional column vector whose components are all 1.

Using the integral representation with respect to a discrete measure, we have the marginal distribution function of $\tilde{x}_{k_1},\cdots,\tilde{x}_{k_s}$ as follows:

$$\begin{aligned}(2.4)\qquad & F(x_{k_1},\cdots,x_{k_s})\\ &=\int_S F(x_{k_1},\cdots,x_{k_s}\mid\boldsymbol{\theta}_{(k_1,\ldots,k_s)})g_d(\boldsymbol{\theta}_{(k_1,\ldots,k_s)})dN(\boldsymbol{\theta}_{(k_1,\ldots,k_s)})\end{aligned}$$

where $N(A)$ is the counting measure of $A\subset S$ which is defined by $N(A)$ $=$number of points of S contained in A.

The problem stated in the previous section is a two-decision problem, the action space being $\{A_0, A_1\}$ where A_0 and A_1 are actions to decide $\theta=0$ and $\theta=1$, respectively. We shall consider the following loss function:

$$(2.5)\qquad L(A_i,\theta)=\begin{cases}0 & \text{if } \theta=i\,,\\ w_1 & \text{if } i=0 \text{ and } \theta=1\,,\\ w_0 & \text{if } i=1 \text{ and } \theta=0\,.\end{cases}$$

A decision procedure for our detection problem is a sequence of decision functions, $T=\{t_k\}_{k=0,1,2,\dots}$, where each decision function t_k takes on values in the action space A. When p and σ^2 are known, it is reasonable to restrict ourselves to decision functions t_k which depend only on $x_{k-r},\cdots,x_k$, since $x_0,\cdots,x_{k-r-1}$ do not supply any information for the detection of θ_k. Then, the Bayes risk of a decision function t_ν for the detection of θ_ν, say $R(t_\nu;p,\sigma^2)$, is, for $\nu\geqq r$,

$$(2.6)\qquad R(t_\nu;p,\sigma^2)\equiv\int_{R_d}\int_{R_{r+1}} L[t_\nu(x_{\nu-r},\cdots,x_\nu),\theta_\nu]\cdot dF(x_{\nu-r},\cdots,x_\nu\mid\boldsymbol{\theta}_{(\nu-r,\dots,\nu)})g_d(\boldsymbol{\theta}_{(\nu-r,\dots,\nu)})dN(\boldsymbol{\theta}_{(\nu-r,\dots,\nu)})$$

where, in fact, $d=d(\nu-r,\cdots,\nu)=2r+1$ and

$$(2.7)\qquad \boldsymbol{\theta}_{(\nu-r,\dots,\nu)}=(\theta_{\nu-2r},\cdots,\theta_\nu)\,.$$

We can write $\boldsymbol{\theta}_{(\nu-r,\dots,\nu)}=(\boldsymbol{\theta}_{(\nu-r,\dots,\nu-1)},\theta_\nu)$ and so

$$(2.8)\qquad \begin{aligned} g_d(\boldsymbol{\theta}_{(\nu-r,\dots,\nu)})&=p^{\boldsymbol{\theta}_{(\nu-r,\dots,\nu-1)}\mathbf{1}_{d-1}+\theta_\nu}(1-p)^{d-\boldsymbol{\theta}_{(\nu-r,\dots,\nu-1)}\mathbf{1}_{d-1}-\theta_\nu}\\ &=p^{\theta_\nu}(1-p)^{1-\theta_\nu}g_{d-1}(\boldsymbol{\theta}_{(\nu-r,\dots,\nu-1)})\,.\end{aligned}$$

Let us define, for $i=0$ or 1,

$$(2.9)\qquad F_i(x_{\nu-r},\cdots,x_\nu\mid\boldsymbol{\theta}_{(\nu-r,\dots,\nu-1)})\equiv F(x_{\nu-r},\cdots,x_\nu\mid(\boldsymbol{\theta}_{(\nu-r,\dots,\nu-1)},i))$$

and

$$(2.10)\qquad F_i(x_{\nu-r},\cdots,x_\nu)\equiv\int_{R_{d-1}}F_i(x_{\nu-r},\cdots,x_\nu\mid\boldsymbol{\theta}_{(\nu-r,\dots,\nu-1)})\cdot g_{d-1}(\boldsymbol{\theta}_{(\nu-r,\dots,\nu-1)})dN(\boldsymbol{\theta}_{(\nu-r,\dots,\nu-1)})\,.$$

Since

$$(2.11)\qquad \begin{aligned} R(t_\nu;p,\sigma^2)=(1-p)&\int_{R_{d-1}}\int_{R_{r+1}}L(t_\nu,0)dF_0(x_{\nu-r},\cdots,x_\nu\mid\boldsymbol{\theta}_{(\nu-r,\dots,\nu-1)})\\ &\cdot g_{d-1}(\boldsymbol{\theta}_{(\nu-r,\dots,\nu-1)})dN(\boldsymbol{\theta}_{(\nu-r,\dots,\nu-1)})\\ &+p\int_{R_{d-1}}\int_{R_{r+1}}L(t_\nu,1)dF_1(x_{\nu-r},\cdots,x_\nu\mid\boldsymbol{\theta}_{(\nu-r,\dots,\nu-1)})\\ &\cdot g_{d-1}(\boldsymbol{\theta}_{(\nu-r,\dots,\nu-1)})dN(\boldsymbol{\theta}_{(\nu-r,\dots,\nu-1)})\end{aligned}$$

we have

$$(2.12)\qquad \begin{aligned} R(t_\nu;p,\sigma^2)=(1-p)w_0&\int_{R_{r+1}}t_\nu(x_{\nu-r},\cdots,x_\nu)dF_0(x_{\nu-r},\cdots,x_\nu)\\ &+pw_1\int_{R_{r+1}}[1-t_\nu(x_{\nu-r},\cdots,x_\nu)]dF_1(x_{\nu-r},\cdots,x_\nu)\,.\end{aligned}$$

Now, the random variables $\tilde{x}_{\nu-r},\cdots,\tilde{x}_{\nu-1}$ are independent of the random variable $\tilde{\theta}_\nu$, hence

$$(2.13)\qquad F(x_{\nu-r},\cdots,x_{\nu-1}\,|\,(\boldsymbol{\theta}_{(\nu-r,\dots,\nu-1)},i))=F(x_{\nu-r},\cdots,x_{\nu-1}\,|\,\boldsymbol{\theta}_{(\nu-r,\dots,\nu-1)})\,.$$

Therefore, we can write

$$(2.14)\qquad \begin{aligned} F_i(x_{\nu-r},\cdots,x_\nu)&=\int_{R_{d-1}}\int_{R_r}F_i(x_\nu\,|\,x_{\nu-r},\cdots,x_{\nu-1})\\ &\quad\cdot dF(x_{\nu-r},\cdots,x_{\nu-1}\,|\,\boldsymbol{\theta}_{(\nu-r,\dots,\nu-1)})\\ &\quad\cdot g_{d-1}(\boldsymbol{\theta}_{(\nu-r,\dots,\nu-1)})dN(\boldsymbol{\theta}_{(\nu-r,\dots,\nu-1)})\\ &=\int_{R_r}F_i(x_\nu\,|\,x_{\nu-r},\cdots,x_{\nu-1})dF(x_{\nu-r},\cdots,x_{\nu-1}) \end{aligned}$$

or

$$(2.15)\qquad dF_i(x_{\nu-r},\cdots,x_\nu)=dF_i(x_\nu\,|\,x_{\nu-r},\cdots,x_{\nu-1})dF(x_{\nu-r},\cdots,x_{\nu-1})\,.$$

From (2.12) and (2.15), we obtain

$$(2.16)\qquad \begin{aligned} R(t_\nu\,;p,\sigma^2)=pw_1+\int_{R_{r+1}}&[(1-p)w_0f_0(x_\nu\,|\,x_{\nu-r},\cdots,x_{\nu-1})\\ &-pw_1f_1(x_\nu\,|\,x_{\nu-r},\cdots,x_{\nu-1})]t_\nu(x_{\nu-r},\cdots,x_\nu)dx_\nu\\ &\cdot f(x_{\nu-r},\cdots,x_{\nu-1})\,dx_{\nu-r}\cdots dx_{\nu-1} \end{aligned}$$

where $f_0(x_\nu\,|\,x_{\nu-r},\cdots,x_{\nu-1})$, $f_1(x_\nu\,|\,x_{\nu-r},\cdots,x_{\nu-1})$ and $f(x_{\nu-r},\cdots,x_{\nu-1})$ are probability density functions of $F_0(x_\nu\,|\,x_{\nu-r},\cdots,x_{\nu-1})$, $F_1(x_\nu\,|\,x_{\nu-r},\cdots,x_{\nu-1})$ and $F(x_{\nu-r},\cdots,x_{\nu-1})$, respectively.

The Bayes decision rule which minimizes $R(t_\nu\,;p,\sigma^2)$ is defind by

$$(2.17)\qquad t(x_{\nu-r},\cdots,x_\nu)=\begin{cases} 0 & \text{if } (1-p)w_0f_0(x_\nu\,|\,x_{\nu-r},\cdots,x_{\nu-1})\\ & \qquad -pw_1f_1(x_\nu\,|\,x_{\nu-r},\cdots,x_{\nu-1})>0\,,\\ 1 & \text{if } (1-p)w_0f_0(x_\nu\,|\,x_{\nu-r},\cdots,x_{\nu-1})\\ & \qquad -pw_1f_1(x_\nu\,|\,x_{\nu-r},\cdots,x_{\nu-1})\leqq 0\,. \end{cases}$$

Since $f_i(x_\nu\,|\,x_{\nu-r},\cdots,x_{\nu-1})=f_i(x_{\nu-r},\cdots,x_\nu)/f_i(x_{\nu-r},\cdots,x_{\nu-1})$, $i=0,1$, where $f_i(\cdot)$ is pdf of the corresponding cdf $F_i(\cdot)$, and $f_0(x_{\nu-r},\cdots,x_{\nu-1})=f_1(x_{\nu-r},\cdots,x_{\nu-1})=f(x_{\nu-r},\cdots,x_{\nu-1})$, the decision rule (2.17) can be written as follows:

$$(2.18)\qquad t(x_{\nu-r},\cdots,x_\nu)=\begin{cases} 0 & \text{if } (1-p)w_0f_0(x_{\nu-r},\cdots,x_\nu)\\ & \qquad -pw_1f_1(x_{\nu-r},\cdots,x_\nu)>0\,,\\ 1 & \text{if } (1-p)w_0f_0(x_{\nu-r},\cdots,x_\nu)\\ & \qquad -pw_1f_1(x_{\nu-r},\cdots,x_\nu)\leqq 0\,. \end{cases}$$

On the other hand, from (2.1) and (2.14) we have

$$
\begin{aligned}
f_0(x_{\nu-r},\cdots,x_\nu)&=\left(\frac{1}{2\pi\sigma^2}\right)^{(r+1)/2}\int\exp\left[-\frac{1}{2\sigma^2}\left\{\sum_{j=1}^{r}(x_{\nu-j}-\boldsymbol{\theta}_{\nu-j}\boldsymbol{\alpha})^2\right.\right.\\
&\qquad\left.\left.+(x_\nu-\boldsymbol{\theta}_\nu^*\boldsymbol{\alpha}^*)^2\right\}\right]g_{d-1}(\boldsymbol{\theta}_{(\nu-r,\cdots,\nu-1)})dN(\boldsymbol{\theta}_{(\nu-r,\cdots,\nu-1)})
\end{aligned}
\tag{2.19}
$$

$$
\begin{aligned}
f_1(x_{\nu-r},\cdots,x_\nu)&=\left(\frac{1}{2\pi\sigma^2}\right)^{(r+1)/2}\int\exp\left[-\frac{1}{2\sigma^2}\left\{\sum_{j=1}^{r}(x_{\nu-j}-\boldsymbol{\theta}_{\nu-j}\boldsymbol{\alpha})^2\right.\right.\\
&\qquad\left.\left.+(x_\nu-\alpha_0-\boldsymbol{\theta}_\nu^*\boldsymbol{\alpha}^*)^2\right\}\right]g_{d-1}(\boldsymbol{\theta}_{(\nu-r,\cdots,\nu-1)})dN(\boldsymbol{\theta}_{(\nu-r,\cdots,\nu-1)})
\end{aligned}
$$

where $\boldsymbol{\alpha}^*=(\alpha_r,\cdots,\alpha_1)'$, $\boldsymbol{\theta}_\nu^*=(\theta_{\nu-r},\cdots,\theta_{\nu-1})$.

Using the definition of $g_{d-1}(\boldsymbol{\theta}_{(\nu-r,\cdots,\nu-1)})$, we have the decision rule (2.18) in the following form:

$$
t_\nu(x_{\nu-r},\cdots,x_\nu)=\begin{cases}0 & \text{if } \phi(x_{\nu-r},\cdots,x_\nu;p,\sigma^2)>0\\ 1 & \text{if } \phi(x_{\nu-r},\cdots,x_\nu;p,\sigma^2)\leqq 0\end{cases}
\tag{2.20}
$$

where

$$
\begin{aligned}
&\phi(x_{\nu-r},\cdots,x_\nu;p,\sigma^2)\\
&=\sum_{\boldsymbol{\theta}_{(\nu-r,\cdots,\nu-1)}}\left\{(1-p)w_0-pw_1\cdot\exp\left(-\frac{1}{2\sigma^2}[\alpha_0^2-2\alpha_0(x_\nu-\boldsymbol{\theta}_\nu^*\boldsymbol{\alpha}^*)]\right)\right\}\\
&\quad\cdot p^{\boldsymbol{\theta}_{(\nu-r,\cdots,\nu-1)}\mathbf{1}_{d-1}}(1-p)^{d-1-\boldsymbol{\theta}_{(\nu-r,\cdots,\nu-1)}\mathbf{1}_{d-1}}\\
&\quad\cdot\exp\left(-\frac{1}{2\sigma^2}\left[(x_\nu-\boldsymbol{\theta}_\nu^*\boldsymbol{\alpha}^*)^2+\sum_{j=1}^{r}(x_{\nu-j}-\boldsymbol{\theta}_{\nu-j}\boldsymbol{\alpha})^2\right]\right).
\end{aligned}
\tag{2.21}
$$

The Bayes risk of the above decision rule $t_\nu(\cdot)$ is, for $\nu\geqq r$,

$$
R(t_\nu;p,\sigma^2)=pw_1+\int_{R_r}\left[\int_K\phi(x_{\nu-r},\cdots,x_\nu;p,\sigma^2)dx_\nu\right]dx_{\nu-1}\cdots dx_{\nu-r}
\tag{2.22}
$$

where K denotes the set defined for each $(x_{\nu-r},\cdots,x_{\nu-1})$ as

$$
K(x_{\nu-r},\cdots,x_{\nu-1})=\{x_\nu\,|\,\phi(x_{\nu-r},\cdots,x_\nu;p,\sigma^2)<0\}\,.
$$

It should be noted that $R(t_\nu;p,\sigma^2)$ does not depend on ν at all if $\nu\geqq r$.

3. Empirical Bayes solution when p and σ^2 are unknown

In this section we will treat the same problem as in the previous sections, when p and σ^2 are unknown; we shall construct an empirical Bayes solution for this problem. For this purpose we need consistent estimates of p and σ^2. As such estimates we consider $\hat{p}$ and $\hat{\sigma}^2$ which are given in (3.1) and (3.9). Let us define, for $\nu=0,1,2,\cdots$

$$a_0\hat{p}_\nu=\frac{1}{\nu+1}\sum_{i=0}^{\nu}x_i \tag{3.1}$$

and

$$S_\nu^2=\frac{1}{\nu}\sum_{i=0}^{\nu}(x_i-a_0\hat{p}_\nu)^2 \tag{3.2}$$

for the data $\{x_i, i=0, 1,\cdots,\nu\}$ from the data generating process (1.1), where $a_0=\sum_{i=0}^{r}\alpha_i$ and we assume $a_0\neq 0$. Further, define

$$a_0\tilde{\hat{p}}_\nu=\frac{1}{\nu+1}\sum_{i=0}^{\nu}\tilde{x}_i \tag{3.3}$$

and

$$\tilde{S}_\nu^2=\frac{1}{\nu}\sum_{i=0}^{\nu}(\tilde{x}_i-a_0\tilde{\hat{p}}_\nu)^2\,. \tag{3.4}$$

As is easily seen we have

$$\mathrm{E}\{\tilde{\hat{p}}_\nu\}=p\,, \tag{3.5}$$

and for $\nu\geqq r$

$$\mathrm{E}\{\tilde{S}_\nu^2\}=\frac{\nu+1}{\nu}p(1-p)\sum_{i=0}^{r}\alpha_i^2+\sigma^2-\frac{1}{\nu(\nu+1)}p(1-p)A_\nu\,, \tag{3.6}$$

where $a_1=\sum_{i=0}^{r-1}\alpha_i,\ \cdots,\ a_j=\sum_{i=0}^{r-j}\alpha_i,\ \cdots,\ a_r=\alpha_0,\ a_{-1}=\sum_{i=1}^{r}\alpha_i,\ \cdots,\ a_{-j}=\sum_{i=j}^{r}\alpha_i,\ \cdots,$ $a_{-r}=\alpha_r$ and $A_\nu=\sum_{j=-r}^{r}a_j^2+(\nu-r)a_0^2$ for $\nu\geqq r$. Since the observable random variables are generated from independent random variables $\{\tilde{\theta}_i\}$ and $\{\tilde{\varepsilon}_i\}$ as in the model (1.1), we can apply the zero-one law to the proofs of the following statements: When p and σ^2 are the true parameters we have

$$\mathrm{P}\{\lim_{\nu\to\infty}\tilde{\hat{p}}_\nu=p\}=1 \tag{3.7}$$

and

$$\mathrm{P}\left\{\lim \tilde{S}_\nu^2=p(1-p)\sum_{i=0}^{r}\alpha_i^2+\sigma^2\right\}=1\,. \tag{3.8}$$

Thus we obtain a consistent estimate of σ^2:

$$\hat{\sigma}_\nu^2=S_\nu^2-\hat{p}_\nu(1-\hat{p}_\nu)\sum_{i=0}^{r}\alpha_i^2\,. \tag{3.9}$$

An empirical Bayes decision rule $\{t_\nu^*\}$ is thus defined by

$$(3.10)\qquad t_\nu^*(x_0,\cdots,x_\nu)=\begin{cases}0 & \text{if } \phi(x_{\nu-r},\cdots,x_\nu;\hat{p}_\nu,\hat{\sigma}_\nu^2)>0\,,\\ 1 & \text{otherwise}\,,\end{cases}$$

where ϕ is the function defined in (2.21). Now, to express explicitly that the density functions in the previous section depend on the unknown parameters p and σ^2, we write $f_i(x_\nu\,|\,x_{\nu-r},\cdots,x_{\nu-1};p,\sigma^2)$, $f(x_0,\cdots,x_{\nu-r-1}\,|\,x_{\nu-r},\cdots,x_{\nu-1};p,\sigma^2)$, $f(x_{\nu-r},\cdots,x_{\nu-1};p,\sigma^2)$ instead of the corresponding density functions.

In the remainder of this section we wish to prove that the sequence of risks of the decision rules $\{t_\nu^*\}$, $\{R_\nu^*(t_\nu^*;p,\sigma^2)\}$, coincides asymptotically with the Bayes risk of the decision rule $\{t_\nu\}$ given by (2.22). Since $\hat{p}_\nu$ and $\hat{\sigma}_\nu^2$ are consistent estimate of p and σ^2, it is quite plausible that for any p $(0\leqq p\leqq 1)$ and σ^2 (>0)

$$(3.11)\qquad \lim_{\nu\to\infty} R_\nu^*(t_\nu^*;p,\sigma^2)=R(t_\nu;p,\sigma^2)\,.$$

It should be noted that $R(t_\nu;p,\sigma^2)$ is independent of ν if $\nu\geqq r$. We shall consider only the case when $\nu\geqq r$.

PROOF OF (3.11).

$$\begin{aligned}(3.12)\qquad R_\nu^*(t_\nu^*;p,\sigma^2)&=pw_1\,\mathrm{E}\{1-t_\nu^*(\tilde{x}_0,\cdots,\tilde{x}_\nu)\,|\,\tilde{\theta}_\nu=1\}\\ &\quad+(1-p)w_0\,\mathrm{E}\{t_\nu^*(\tilde{x}_0,\cdots,\tilde{x}_\nu)\,|\,\tilde{\theta}_\nu=0\}\\ &=pw_1+\int_{R^{\nu-r}}\int_{R^r}\int_{R^1} t_\nu^*\phi(x_{\nu-r},\cdots,x_\nu;p,\sigma^2)dx_\nu\\ &\quad\cdot f(x_{\nu-r},\cdots,x_{\nu-1}\,|\,x_0,\cdots,x_{\nu-r-1};p,\sigma^2)dx_{\nu-r}\cdots dx_{\nu-1}\\ &\quad\cdot f(x_0,\cdots,x_{\nu-r-1}\,|\,p,\sigma^2)dx_0\cdots dx_{\nu-r-1}\end{aligned}$$

where

$$\begin{aligned}(3.13)\qquad \phi(x_{\nu-r},\cdots,x_\nu;p,\sigma^2)&=(1-p)w_0f_0(x_\nu\,|\,x_{\nu-r},\cdots,x_{\nu-1};p,\sigma^2)\\ &\quad-pw_1f_1(x_\nu\,|\,x_{\nu-r},\cdots,x_{\nu-1};p,\sigma^2)\,.\end{aligned}$$

Interchanging the order of integration and replacing the variables $x_{\nu-r},\cdots,x_\nu$ by $y_0,y_1,\cdots,y_r$, respectively, we have

$$\begin{aligned}(3.14)\qquad R_\nu^*(t_\nu^*;p,\sigma^2)&=pw_1+\int_{R^r}\left\{\int_{R^1}\left(\int_{R^{\nu-r}} t_\nu^*(x_0,\cdots,x_{\nu-r-1},y_0,\cdots,y_r)\right.\right.\\ &\quad\left.\cdot f(x_0,\cdots,x_{\nu-r-1}\,|\,y_0,\cdots,y_{\nu-1};p,\sigma^2)dx_0\cdots dx_{\nu-r-1}\right)\\ &\quad\left.\cdot\,\phi(y_0,\cdots,y_r;p,\sigma^2)dy_r\right\}f(y_0,\cdots,y_{r-1};p,\sigma^2)\\ &\quad\cdot\,dy_0\cdots dy_{r-1}\,.\end{aligned}$$

Now

$$(3.15)\qquad \int_{R^{\nu-r}} t_\nu^*(x_0,\cdots,x_{\nu-r-1},y_0,\cdots,y_r) \cdot f(x_0,\cdots,x_{\nu-r-1}\mid y_0,\cdots,y_r;p,\sigma^2)dx_0\cdots dx_{\nu-r-1}$$
$$=\mathrm{P}\{t_\nu^*(\tilde{x}_0,\cdots,\tilde{x}_{\nu-r-1},y_0,\cdots,y_r)=1\mid p,\sigma^2\}$$
$$=\mathrm{P}\{\phi(y_0,\cdots,y_r;\tilde{\pi}_\nu,\tilde{\tau}_\nu)<0\mid p,\sigma^2\}$$

where

$$(3.16)\qquad \begin{aligned}\tilde{\pi}_\nu&=\hat{p}_\nu(\tilde{x}_0,\cdots,\tilde{x}_{\nu-r-1},y_0,\cdots,y_r)\\ \tilde{\tau}_\nu&=\hat{\sigma}_\nu^2(\tilde{x}_0,\cdots,\tilde{x}_{\nu-r-1},y_0,\cdots,y_r)\,.\end{aligned}$$

From (3.7), (3.8) and (3.9) we have

$$(3.17)\qquad \begin{aligned}&\mathrm{P}\{\lim_{\nu\to\infty}\tilde{\pi}_\nu=p\mid p,\sigma^2\}=1\\ &\mathrm{P}\{\lim_{\nu\to\infty}\tilde{\tau}_\nu=\sigma^2\mid p,\sigma^2\}=1\,.\end{aligned}$$

From (3.17) and the continuity of $\phi(y_0,\cdots,y_r;p,\sigma^2)$ in p and σ^2, we obtain

$$(3.18)\qquad \lim_{\nu\to\infty}\mathrm{P}\{\phi(y_0,\cdots,y_r;\tilde{\pi}_\nu,\tilde{\tau}_\nu)<0\mid p,\sigma^2\} =\begin{cases}0 & \text{if } \phi(y_0,\cdots,y_r;p,\sigma^2)>0\\ 1 & \text{if } \phi(y_0,\cdots,y_r;p,\sigma^2)<0\,.\end{cases}$$

By the Lebesgue dominated convergence theorem, we have, from (3.14), (3.15) and (3.18),

$$(3.19)\qquad \lim_{\nu\to\infty}R_\nu^*(t_\nu^*;p,\sigma^2)=pw_1+\int_{R^r}\int_K \phi(y_0,\cdots,y_r;p,\sigma^2) \cdot dy_r f(y_0,\cdots,y_{r-1};p,\sigma^2)dy_0\cdots dy_{r-1}$$

where K is the set defined by (2.23). Thus the proof has been completed.

REFERENCES

[1] H. Robbins, "The empirical Bayes approach to statistical decision problems," *Ann. Math. Statist.*, 35 (1964), 1-20.
[2] E. Samuel, "An empirical Bayes approach to the testing of certain parametric hypotheses," *Ann. Math. Statist.*, 34 (1963), 1370-1385.
[3] M. Tainiter, "Sequential hypothesis tests for the r-dependent marginally stationary processes," *Ann. Math. Statist.*, 37 (1966), 90-97.