TSUKUBA J. MATH.
Vol. 20 No. 1 (1996) 77-92

SOME PROPERTIES ON TESTS BASED ON THE BAYESIAN CONFIDENCE INTERVAL

By

Michikazu SATO

Abstract. In testing statistical hypotheses, quite generally, if we admit the result of Neyman-Pearson (apart from the interpretation of them) in case that we specify n in advance and admit the likelihood principle, the stopping rule that "continue the experiments until rejecting the null hypothesis" is closed. As a matter of fact, a stronger phenomenon happens, and we shall show it with some examples.

1. Introduction

Let $X_1, X_2, \dots$ be independently and identically distributed (i.i.d.) random variables with a normal distribution $N(\theta, 1)$. We observe them in the order $X_1, X_2, \dots$ Let the prior distribution of θ be the Lebesgue measure (improper in this case). When we observe $X_1, \dots, X_n$, the posterior distribution is $N(\overline{X}_n, 1/n)$, where $\overline{X}_n = \sum_{j=1}^{n} X_j / n$. Let $0 < \alpha < 1$ and define k by $P(|Z| > k) = \alpha$, where Z is a random variable with a normal distribution $N(0,1)$. Then, the Bayesian $100(1-\alpha)$ percent confidence interval is given by $[\overline{X}_n - (k/\sqrt{n}), \overline{X}_n + (k/\sqrt{n})]$. Consider a significance test of a hypothesis $\mathrm{H}_0 : \theta = 0$. The Bayesian test with a significance level α is that we accept H_0 if $\theta = 0$ falls into the confidence interval and that we reject H_0 otherwise. Note that the significance level here is not in the sense of Neyman-Pearson's. In this case, the critical region is $\{|\overline{X}_n| > k/\sqrt{n}\}$, so if we specify n in advance, the result coincides with that of Neyman-Pearson. In application, it usually holds. Note that in the standpoint of usual Bayesian, we do not have to specify n in advance (Akaike [1], however, mentions this respect critically.). Tests like this method are described in Lindley [10], [11] and Shigemasu [15]. In [12], [13], however, Lindley seems to have abandoned this standpoint and have taken the standpoint of Bayesian tests of Jeffreys [7]. In

This research was supported in part by Grant-in-Aid for Science Research, No. 00001777, Ministry of Education, Science, Sports and Culture, Japan.
Received March 4, 1994. Revised May 19, 1994.

Shigemasu [15], both standpoints are described. Tests of Jeffreys are free from the problem as will be mentioned later (Cornfield [5] p. 581).

Now consider the stopping rule that "continue the experiments as long as $|\overline{X}| \le k/\sqrt{n}$ holds and stop when it is violated." Then, since this stopping rule is closed (for the definition, see Section 2), if we take sufficiently small α and make these experiments, we can make the supporters for this test believe that H_0 is not true, with probability 1 irrespective of H_0 being true or not. This fact is described in Robbins [14], Lindley [9] (There is a mistake in this paper. See Bartlett [2].), Cornfield [5], Berger and Wolpert [4], Basu [3]. As a matter of fact, quite generally, if we admit the result of Neyman-Pearson (apart from the interpretation of them) in case that we specify n in advance and admit the likelihood principle, such a phenomenon happens. We call this fact WSC as will be mentioned later. Moreover, a stronger phenomenon happens. We see from the above that the result of the test in the standpoint of Neyman-Pearson and the likelihood principle are quite incompatible. Note that this is not the difference of interpretations on the same result.

2. Main concepts

Let $X_1, X_2, \ldots$ be a sequence of random variables. They are not necessarily real-valued nor i.i.d. Assume that the distribution of $(X_1, X_2, \ldots)$ is defined for each parameter θ and we observe them in the order $X_1, X_2, \ldots$. We denote $X_n^* := (X_1, \ldots X_n)$. A stopping rule σ is said to be closed at $\theta = \theta_0$ if

$$P_{\theta_0}(\text{"the random stopping time based on } \sigma \text{"} < \infty) = 1.$$

When we say only "σ is closed," it means that σ is closed for all θ, and when we say "σ is not closed," it means that σ is not closed for some θ. Similar usage is adopted for the following WSC, SSC and ASC. In the sequel, we denote a null hypothesis by H_0 and an alternative hypothesis by H_1. We assume that, for each n, a critical region R_n is given when we observe X_n^*. When we consider a randomized test, we transform it to a nonrandomized test by introducing random numbers. Then we take the following definitions.

DEFINITION 2.1. A sequence of tests $\{R_n\}$ is said to be weakly sophistically closed (or WSC for short) at $\theta = \theta_0$ if

$$P_{\theta_0}(X_n^* \in R_n \quad \text{for some } n) = 1.$$

DEFINITION 2.2. A sequence of tests $\{R_n\}$ is said to be strongly sophistically closed (or SSC for short) at $\theta = \theta_0$ if

$$P_{\theta_0}(X_n^* \in R_n \quad \text{i.o.})=1,$$

where "i.o." means "infinitely often."

The reader might think that SSC is an empty, abstract and only theoretical concept because we cannot experiment infinitely. But, there is a sophistical meaning in SSC as follows:

Now, I want to insist that H_0 is not true but there have already been a predecessor's experiments. Even if the results are unfavorable for me, I cannot ignore them, but I can make supplementary examinations. So, if H_0 is rejected by the results of the predecessor's, I myself do not experiment, and insist, "H_0 is rejected by the results of a predecessor's experiments". If H_0 is accepted by the results of the predecessor's, I make supplementary examinations. I continue to do them until H_0 is rejected, and insist "H_0 is rejected by the results of my supplementary examinations added to a predecessor's experiments." Then it raises a question whether I succeed (that is, end finitely) or not. As for this question, the following assertion holds.

Let Σ be a family of closed randomized stopping rules which satisfies that there exists a sequence $\{\sigma_j\}_{j=1}^{\infty} \subset \Sigma$ such that if we denote the random stopping time based on σ_j by M_j (generally à random variable), by appropriately determining a conditional joint probability distribution of stopping the experiments for any given observed value,

$$\lim_{j\to\infty} M_j = \infty \quad P_{\theta_0} - \text{a.e.} \tag{2.1}$$

holds. Then, the following (1) and (2) are equivalent.

(1) $\{R_n\}$ is SSC at $\theta = \theta_0$.

(2) For any stopping rule of a predecessor's in Σ , I succeed in the above with P_{θ_0} probability 1.

In particular, let $n_1 < n_2 < \ldots, n_j \in N$ and let σ_j be the stopping rule corresponding to experimenting exactly n_j times. Put $\Sigma = \{\sigma_j\}_{j=1}^{\infty}$, then (2.1) holds. Note that the assertion above does not only clarify the sophistical meaning of SSC, but also is used in order to show SSC.

For $n_1 < n_2 < \ldots, n_j \in N$, let

$$R_n^* = \begin{cases} R_n & (n = n_j) \\ \phi & (n \neq n_j \text{ for all } j). \end{cases}$$

Let us denote $\{R_n^*\}_{n=1}^{\infty}$ by $\{R_{n_j}^*\}_{j=1}^{\infty}$ and call it a subsequence of $\{R_n\}$. We can also regard it as a sequence of tests based on $Y_1, Y_2, \ldots$ where $Y_j = (X_{n_{j-1}+1}, X_{n_{j-1}+2}, \ldots, X_{n_j})$, $n_0 = 0$.

DEFINITION 2.3. A sequence of tests $\{R_n\}$ is said to be all-subsequentially sophistically closed (or ASC for short) at $\theta = \theta_0$ if all subsequences of $\{R_n\}$ are SSC at $\theta = \theta_0$.

It is easily derived from the assertion above that this is equivalent to that all subsequences of $\{R_n\}$ are WSC at $\theta = \theta_0$.

By definition, ASC implies SSC, and SSC implies WSC.

In the subsequent discussion we shall not explicitly distinguish between a test and a sequence of tests unless there is a possibility of misunderstanding.

3. An exact test and an asymptotic test

In the following discussion, we assume $0 < \alpha < 1$.

DEFINITION 3.1. Fix n. A test that satisfies the following assumptions (3a) and (3b) is called a left-sided exact test based on T_n with Neyman-Pearson significance level α.

(3a) $T_n = g_n(X_1, \ldots X_n)$ is a real-valued random variable, and the distribution of T_n does not depend on θ under H_0.

(3b) There exists t_n such that,

(i) We reject H_0 if $T_n < t_n$, and accept H_0 if $T_n > t_n$.

(ii) We reject H_0 with a constant conditional probability independent of observed values if $T_n = t_n$.

(iii) For θ under H_0, P_θ (reject H_0)$=\alpha$.

Then t_n is called a critical point.

DEFINITION 3.2. Fix n. A test that satisfies (3a) and (i), (iii) of (3b) is called a left-sided exact test based on T_n except on a critical point with Neyman-Pearson significance level α.

DEFINITION 3.3. A test that satisfies (3a) for each n and the following (3c) and (3d) is called a left-sided asymptotic test with an open critical region based on T_n with Neyman-Pearson significance level α.

(3c) The distribution of T_n under H_0 converges weakly to a probability distribution λ whose distribution function is continuous.

(3d) There exists t_∞ such that,

(iv) We reject H_0 if $T_n < t_\infty$, and accept H_0 if $T_n \geq t_\infty$.

(v) $\lambda((-\infty, t_\infty)) = \alpha$.

Then t_∞ is called a critical point.

DEFINITION 3.4. In Definition 3.3, when we replace (iv) by the following (iv)′, the test is called a left-sided asymptotic test with a closed critical region based on T_n with Neyman-Pearson significance level α.

(iv)′ We reject H_0 if $T_n \leq t_\infty$, and accept H_0 if $T_n > t_\infty$.

Usually, so far as SSC or ASC is concerned, the tests of Definition 3.1 to 3.4 are equivalent, but there are delicate problems. We shall now clarify them.

THEOREM 3.1. *Under the assumption* (3a), *the following* (1) *and* (2) *are equivalent.*

(1) *For any* α, *any left-sided exact test based on* T_n *with Neyman-Pearson significance level* α *is SSC at* $\theta = \theta_0$.

(2) *For any* α, *there exists a left-sided exact test based on* T_n *with Neyman-Pearson significance level* α *that is SSC at* $\theta = \theta_0$.

PROOF. It is obvious that (1) implies (2). To prove the converse, we have only to compare a critical region of level α with that of level $\alpha/2$ of (2).

It is noted that the similar results to Theorem 3.1 hold for ASC and WSC.

THEOREM 3.2. *Under the assumptions* (3a) *and* (3c), (1) *and* (2) *in Theorem* 3.1 *and the following* (3) *to* (8) *are equivalent.*

(3) *For any* α, *any left-sided exact test based on* T_n *except on a critical point with Neyman-Pearson significance level* α *is SSC at* $\theta = \theta_0$.

(4) *For any* α, *there exists a left-sided exact test based on* T_n *except on a critical point with Neyman-Pearson significance level* α *that is SSC at* $\theta = \theta_0$.

(5) *For any* α, *any left-sided asymptotic test with an open critical region based on* T_n *with Neyman-Pearson significance level* α *is SSC at* $\theta = \theta_0$.

(6) *For any* α, *there exists a left-sided asymptotic test with an open critical region based on* T_n *with Neyman-Pearson significance level* α *that is SSC at* $\theta = \theta_0$.

(7) (5) *where "open" is replaced by "closed" holds.*

(8) (6) *where "open" is replaced by "closed" holds.*

PROOF. The proof that (5) and (6), (7) and (8) are equivalent, respectively, is similar to Theorem 3.1. It is obvious that (6) implies (8), (3) implies (1), and that (2) implies (4). If we show that (4) implies (5) and (8) implies (3), the proof is completed. First, we shall prove that (4) implies (5). We take a critical point $t_n(\alpha)$ and $t_\infty(\alpha)$ of (4) and (5), respectively. For θ under H_0,

$$\lim_{n\to\infty} P_\theta(T_n < t_\infty(\alpha/2)) = \alpha/2 \tag{3.1}$$

holds. Hence, there exists υ such that

$$\alpha/3 < P_\theta(T_n < t_\infty(\alpha/2)) < \alpha \quad \text{for } n \geq \upsilon.$$

Hence,

$$t_\infty(\alpha/3) \leq t_\infty(\alpha/2) < t_\infty(\alpha) \quad \text{for } n \geq \upsilon \tag{3.2}$$

and we get (5). We get that (8) implies (3) by noting

$$t_n(\alpha/3) < t_\infty(\alpha/2) < t_n(\alpha) \quad \text{for } n \geq \upsilon.$$

The similar result to Theorem 3.1 holds for ASC.

REMARK 1. In order to prove non-WSC, we cannot disregard critical regions of finite n's. Hence Theorem 3.2 does not hold for WSC.

REMARK 2. Under only (3a), (4) does not imply (3). For a counter-example, let $X_1, X_2, \ldots$ be i.i.d. random variables with a uniform distribution $U(0,1)$ under H_0 and $T_n \equiv 0$. Then, the critical point $t_n(\alpha)$ is equal to 0. Both $R_n := \{X_n < \alpha\}$ and $\hat{R}_n := \{X_1 < \alpha\}$ are left-sided exact tests based on T_n except on a critical point with Neyman-Pearson significance level α. But $\{R_n\}$ is ASC at H_0 and $\{\hat{R}_n\}$ is non-WSC at H_0.

REMARK 3. In Definition 3.3, if we exclude the assumption in (3c) that the distribution function of λ is continuous and exclude the sign of equality in (iv) and replace (v) by "A test function φ based on T_n satisfies $\int \varphi(t)\lambda(dt) = \alpha$.", then, we get into trouble as follows: In the assumptions of Section 1, denote the distribution function of $N(0,1)$ by Φ. Under H_0, $\Phi(\sqrt{n}\overline{X}_n)$ is distributed as $U(0,1)$, hence $T_n := \{1 + \Phi(\sqrt{n}\overline{X}_n)\}/n$ is distributed as $U(1/n, 2/n)$ which converges weakly to the Dirac measure on 0. Hence in the definition above, which is milder than Definition 3.3, we always accept H_0 in a left-sided asymptotic test based on T_n, while on the other hand we always reject H_0 in a left-sided asymptotic test based on $-T_n$. Note that, in the proof of Theorem 3.2, the assumption that the distribution function of λ is continuous is used in (3.1) and the strict inequality in (3.2). Also note that when we say weak convergence, generally, the limit is not necessarily a probability distribution, but in such cases, (3.1) does not necessarily hold and we get into trouble, hence we exclude such cases.

REMARK 4. In Definition 3.3, the distribution function of T_n is not necessarily continuous under H_0, hence the sign of equality in (iv) is not generally nonessential. For example, under H_0, let $X_1, X_2, \dots$ be i.i.d. random variables with $U(0,1)$ and

$$T_n := \begin{cases} X_1 & \text{if } X_1 < 1/2, \\ 1/2 & \text{if } 1/2 \le X_1 \text{ and } X_n < 1/n, \\ (nX_n + n - 2)/2(n-1) & \text{if } 1/2 \le X_1 \text{ and } 1/n \le X_n. \end{cases}$$

Then, the limiting distribution of T_n is $U(0,1)$ under H_0. Let Neyman-Pearson significance level α be 1/2. Then, the critical point t_∞ of the asymptotic test is equal to 1/2 and if we take the open critical region, which coincides with the exact test in this case, the test is non-WSC at H_0, but if we take the closed critical region, the test is SSC at H_0 because

$$\begin{aligned} P_{H_0}(T_n \le 1/2 \text{ i.o.}) &= P_{H_0}(X_1 < 1/2) + P_{H_0}(X_1 \ge 1/2) P_{H_0}(X_n < 1/n \text{ i.o.}) \\ &= 1, \end{aligned}$$

where the last equality follows from the Borel-Cantelli lemma.

4. Criteria for SSC and ASC

In this section, we shall consider criteria for SSC and ASC. Note that the law of iterated logarithm (Feller [6]) is useful to judge SSC, but useless to judge ASC.

THEOREM 4.1. *For a sequence of tests* $\{R_n\}$, *the following assertions hold.*
(1) *If there exist* $\varepsilon > 0$ *and* $m_o \in N_0 := N \cup \{0\}$ *such that for any* $m \ge m_0$ *and* $X_m^* = x_m^*$, *there exists* $n (> m)$ *satisfying*

$$P_{\theta_0}(X_n^* \in R_n | X_m^* = x_m^*) \ge \varepsilon,$$

then $\{R_n\}$ *is SSC at* $\theta = \theta_0$.

(2) *There exist* $\varepsilon > 0$ *and* $m_o \in N_0$ *such that for any* $m \ge m_o$ *and* $X_m^* = x_m^*$,

$$\liminf_{n\to\infty} P_{\theta_0}(X_m^* \in R_n \mid X_m^* = x_m^*) \ge \varepsilon ,$$

then $\{R_n\}$ *is SSC at* $\theta = \theta_0$.

In this theorem, the conditional probability for $m = 0$ means the unconditional probability.

PROOF. Denote $P = P_{\theta_0}$ and denote the sample space of X_n by $(\mathscr{X}_n, \mathscr{A}_n)$.

(1) Step 1. For $m_0 = 0,$, we shall prove WSC at $\theta = \theta_0$. For $m \in N_0$ and x_m^*, the least n that satisfies the assumption is denoted by $N_m(x_m^*)$. We easily get measurability of N_m. We can assume that $\mathscr{X}_n$'s are mutually disjoint by giving a registration number to each element if necessary. Furthermore, by adding a symbol $*_n$ which denotes a value that X_n never takes, to $\mathscr{X}_n$, we identify a finite sequence $(x_i, x_{i+1}, \ldots, x_j) \in \prod_{k=i}^{j} \mathscr{X}_k$ with an infinite sequence $(*_1, *_2, \ldots, *_{i-1}, x_i, x_{i+1}, \ldots, x_j, *_{j+1}, *_{j+2}, \ldots) \in \prod_{k=1}^{\infty} \mathscr{X}_k$. In the subsequent discussion, we shall not explicitly distinguish between a random variable and its value. Let

$$N_m = N_m(X_m^*) \quad \text{for} \quad m \in N_0$$

$$M_0 := 0, \quad M_n := N_{M_{n-1}} \quad \text{for} \quad n \in N.$$

We easily get $\prod_{n=1}^{\infty} \mathscr{A}_n$-measurability of M_n. Define a sequence of random variables $\{Y_n\}$ by

$$Y_n := (X_{M_{n-1}+1}, X_{M_{n-1}+2}, \ldots, X_{M_n}),$$

where Y_n takes values of $\prod_{n=1}^{\infty} \mathscr{X}_n$. Note that M_{n-1} depends on $X_1, \ldots, X_{M_{n-1}}$. We easily get that Y_n is $\prod_{n=1}^{\infty} \mathscr{A}_n \to \prod_{n=1}^{\infty} \mathscr{A}_n$ – measurable, that is, $Y_n^{-1}(\prod_{n=1}^{\infty} \mathscr{A}_n) \subset \prod_{n=1}^{\infty} \mathscr{A}_n$. Denote $Y_n^* := (Y_1, \ldots, Y_n)$. Then $Y_n^* = X_{M_n}^*$. Regard $\{R_{M_n}\}_{n=1}^{\infty}$ as a sequence of tests based on Y_n^*. Then,

$$P(X_n^* \in R_n \quad \text{for some } n) \ge P(Y_n^* \in R_{M_n} \quad \text{for some } n).$$

Hence, we need only prove

$$P(Y_n^* \in R_{M_n} \quad \text{for some n}) = 1.$$

Generally, for $\{A_n\}_{n=1}^{\infty}$, if

$$P(A_n \mid A_1^c \cap \cdots \cap A_{n-1}^c) \ge \varepsilon > 0 \quad \text{for all } n,$$

then $P(\bigcup_{n=1}^{\infty} A_n) = 1$, where the left-hand side in case of $n = 1$ means unconditional probability and if $P(B) = 0$, define $P(A|B)$ arbitrarily. We easily get it by considering the complementary event. Fix n and denote

$$B := \{Y_1^* \notin R_1, \ldots, Y_n^* \notin R_{M_{n-1}}\},$$

then we need only prove

$$P(Y_n^* \in R_{M_n} | B) \geq \varepsilon. \tag{4.1}$$

We can assume $P(B) > 0$. Since M_1 is constant by the definition, we denote it by m_1. Let

$$B_{m_2,\ldots,m_n} := \{M_2 = m_2, \ldots, M_n = m_n, Y_1^* \notin R_{M_1}, \ldots, Y_n^* \notin R_{M_n}\}.$$

Then $B = \sum_{m_2,\ldots,m_n} B_{m_2,\ldots,m_n}$, where $\sum$ means the direct sum, and the summation is taken over all possible values of $(M_2, \ldots, M_n)$. Generally,

$$P(A | \sum_k B_k) = \sum_k P(A|B_k) \Big/ \sum_k B_k$$

holds if $\sum_k P(B_k) > 0$. Hence, in order to prove (4.1), we need only prove

$$P(Y_n^* \in R_{M_n} | B_{m_2,\ldots,m_n}) \geq \varepsilon. \tag{4.2}$$

We get

$$\begin{aligned} P(Y_n^* \in R_{M_n} | B_{m_2,\ldots,m_n}) &= P(X_{m_n}^* \in R_{m_n} | B_{m_2,\ldots,m_n}) \\ &= E\left[P(X_{m_n}^* \in R_{m_n} | X_{m_{n-1}}^*) \,|\, B_{m_2,\ldots,m_n}\right]. \end{aligned} \tag{4.3}$$

The last equality holds because $B_{m_2,\ldots,m_n}$ is a set determined only by $X_{m_{n-1}}^*$. It is essentially the definition of the conditional probability. From the assumption and the definition of M_n, we get

$$E\left[P(X_m^* \in R_m \,|\, X_{m_{n-1}}^*) \,\middle|\, B_{m_2,\ldots,m_n}\right] \geq E\left[\varepsilon \,|\, B_{m_2,\ldots,m_n}\right] = \varepsilon.$$

From (4.3) and the above, we get (4.2), hence $\{R_n\}$ is SSC at $\theta = \theta_0$.

Step 2. For arbitrary m_0, we shall prove that $\{R_n\}$ is SSC at $\theta = \theta_0$. We need only prove that, for any fixed $\ell (> m_0)$, if a predecessor experiments exactly ℓ times, I can reject H_0 by making supplementary examinations (P_{θ_0} − a.e.). In order to show the above, we need only prove that there exists $\tilde{m}_0 > \ell$ such that $\{R_{\tilde{m}_0+n-1}\}_{n=1}^{\infty}$ is WSC at $\theta = \theta_0$. In the notation of Step 1, the possible values of N_l are $\ell+1, \ell+2, \ldots$, hence we can take $\tilde{m}_0 > \ell (> m_0)$ such that

$$\delta := P(N_l = \tilde{m}_0) > 0.$$

Denote $\tilde{\varepsilon} := \varepsilon\delta(>0)$. Consider $\{\tilde{R}_n\} := \{R_{\tilde{m}_0+n-1}\}_{n=1}^{\infty}$, which is a subsequence of $\{R_n\}$. We can regard it as a sequence of tests based on $Z_1 := X^*_{\tilde{m}_0}, Z_2 := X_{\tilde{m}_0+1}, Z_3 := X_{\tilde{m}_0+2}, \ldots$. Denote the indicator of A by I_A. Then, we get

$$\begin{aligned}
P(Z_1 \in \tilde{R}_1) = P(X^*_{\tilde{m}_0} \in R_{\tilde{m}_0}) &= E\left[P(X^*_{\tilde{m}_0} \in R_{\tilde{m}_0} | X^*_l)\right] \\
&\geq E\left[I_{\{N_l = \tilde{m}_0\}} P(X^*_{\tilde{m}_0} \in R_{\tilde{m}_0} | X^*_l)\right] \\
&\geq E\left[I_{\{N_l = \tilde{m}_0\}} \varepsilon\right] = \varepsilon P(N_l = \tilde{m}_0) = \varepsilon\delta = \tilde{\varepsilon},
\end{aligned}$$

where the second inequality follows from the assumption and the definition of N_l. Hence, for $m = 0$, if we let $n = 1$, $P(Z_1 \in \tilde{R}_1) \geq \tilde{\varepsilon}$ holds. For $m = 1,2,\ldots$, for any given value of $Z^*_m := (Z_1, \ldots, Z_m) = X^*_{\tilde{m}_0+m-1}$, if we let $n = N_{\tilde{m}_0+m-1} - \tilde{m}_0 + 1$, we get

$$\begin{aligned}
P(Z^*_n \in \tilde{R}_n | Z^*_m) &= P(X^*_{N_{\tilde{m}_0}+m-1} \in R_{N_{\tilde{m}_0}+m-1} | X^*_{\tilde{m}_0+m-1}) \\
&\geq \varepsilon \geq \tilde{\varepsilon}.
\end{aligned}$$

Hence, for a sequence of tests $\{\tilde{R}_n\}$ based on $\{Z_n\}$, the assumption holds for $\tilde{\varepsilon}$ instead of ε and $m_0 = 0$. Hence, from Step 1, we get $\{\tilde{R}_n\}$ is WSC at $\theta = \theta_0$. This completes the proof of (1).

(2) For any subsequence of $\{R_n\}$, the assertion of (1) for $\varepsilon/2$ instead of ε is satisfied. Hence we get (2).

THEOREM 4.2. *Let $X_1, X_2, \ldots$ be indepedent random variables and let $T_n = g_n(X^*_n)$ be real-valued and $R_n \supset \{T_n < t_0\}$, where t_0 is independent of n. Assume the following* (4a) *and* (4b).

(4a) *There exists $t_1 < t_0$ such that*

$$\liminf_{n\to\infty} P_{\theta_0}(T_n < t_1) > 0$$

(4b) *There exists $m_0 \in \mathbf{N}_0$ such that for any $m \geq m_0$ and $x_1, x_2, \ldots, x_m$ (x_j is a value that x_j can take),*

$$g_n(x_1, \ldots, x_m, X_{m+1}, \ldots, X_n) - T_n \to 0 \quad (P_{\theta_0}) \quad (n \to \infty)$$

holds in the sense of convergence in probability.

Then, $\{R_n\}$ is ASC at $\theta = \theta_0$.

PROOF.

$$\begin{aligned}
&P_{\theta_0}(T_n < t_0 | X_1 = x_1, \ldots, X_m = x_m) = P_{\theta_0}(g_n(x_1, \ldots, x_m, X_{m+1}, \ldots, X_n) < t_0) \\
&\geq P_{\theta_0}(T_n < t_1) - P_{\theta_0}(g_n(x_1, \ldots, x_m, X_{m+1}, \ldots, X_n) \geq t_0 - t_1)
\end{aligned}$$

holds. Take $\liminf_{n\to\infty}$ and we can reduce it to Theorem 4.1.

REMARK. We must not replace (4a) by

(4a)′
$$\liminf_{n\to\infty} P_{\theta_0}(T_n < t_0) > 0$$

For a counter-example, $X_1, X_2, \ldots$ be independent random variables and let $P_{\theta_0}(X_1 \in A) = \alpha \in (0,1)$ and

$$T_n := \begin{cases} -1/n & \text{if } X_1 \in A, \\ 0 & \text{otherwise,} \end{cases}$$

$$t_0 := 0, \quad R_n := \{T_n < 0\}.$$

Then, (4a)′ and (4b) hold but it is non-WSC at $\theta = \theta_0$ because

$$P_{\theta_0}(T_n < 0 \quad \text{for some} \quad n) = \alpha.$$

EXAMPLE 4.1. (1) $T_n = a_n \sum_{j=1}^n X_j$ and $\lim_{n\to\infty} a_n = 0$ implies that the assumption (4b) holds.

(2) Let $X_1, X_2, \ldots$ be i.i.d. and $R_n \supset \{\sqrt{n}\overline{X}_n < t_0\}$ and assume that one of the following (4c) to (4e) holds at $\theta = \theta_0$.

(4c)
$$E_{\theta_0} X_1 \le 0 < \mathrm{Var}_{\theta_0} X_1 < \infty.$$

(4d)
$$-\infty \le E_{\theta_0} X_1 < 0.$$

(4e) X_1 is distributed as a Cauchy distribution.

Then, $\{R_n\}$ is ASC at $\theta = \theta_0$.

(3) The example in Section 1 is ASC.

(4) Let $X_1, X_2, \ldots$ be i.i.d. and $R_n = \{\sqrt{n}\overline{X}_n < t_0\}$ and assume $0 < E_{\theta_0} X_1 \le \infty$. Then $\{R_n\}$ is non-SSC at $\theta = \theta_0$. For further details, this example satisfies

$$P_{\theta_0}(\sqrt{n}\overline{X}_n < t_0 \quad \text{i.o.}) = 0.$$

(5) In the assumptions of (4), if X_1 is distributed as a normal distribution at $\theta = \theta_0$ and $E_{\theta_0} X_1 > t_0$, then $\{R_n\}$ is non-WSC at $\theta = \theta_0$.

Indeed, we can easily get (1) to (4) by using the central limit theorem, the strong law of large numbers, and the reproducibility of a Cauchy distribution. We shall prove (5). Let X_1 be distributed as $N(\xi_0, \sigma^2)$ at $\theta = \theta_0$. We may assume $\sigma = 1$ and $t_0 > 0$. Let $Z_j := X_j - \xi_0, \xi_1 := \xi_0 - t_0 (> 0), Y_j := Z_j - \xi_1$ Then, we get $R_n = \{\sqrt{n}\overline{Z}_n < t_0 - \sqrt{n}\xi_0\} \subset \{\sqrt{n}\overline{Y}_n < 0\}$. Since Y_j's are i.i.d. random variables with $N(\theta, 1)$, we need only prove that, if X_j's are i.i.d. random variables with

$N(\theta,1)$, and

$$\mathrm{H}_0:\theta=\theta_0,\quad \mathrm{H}_1:\theta=-\theta_0\quad(\theta_0>0),$$

then $\tilde{R}_n=\{\sqrt{n}\overline{X}_n<0\}$ is non-WSC at H_0. Let the prior distribution be

$$P(\theta=\theta_0)=P(\theta=-\theta_0)=1/2,$$

then

$$\tilde{R}_n=\left\{P(\theta=\theta_0\,|\,\overline{X}_n)<\frac{1}{2}\right\}$$

holds. Hence, from Cornfield [5] p. 581, it is non-WSC at H_0.

EXAMPLE 4.2. In multinomial trials, if we specify Neyman-Pearson significance level α, the usual χ^2-test is ASC. We can get it from Theorem 4.2 or Koike [8].

THEOREM 4.3. *Let* $X_1,X_2,\ldots$ *be i.i.d. and*

$$\mathrm{H}_0: X_1 \text{ is distributed as } \nu_0,\quad \mathrm{H}_1: X_1 \text{ is distributed as } \nu_1,$$

where $\nu_0\neq\nu_1,\nu_0$ *and* ν_1 *are mutually absolutely continuous and assume*

$$\int d\nu_0(\log d\nu_1/d\nu_0)^2<\infty. \tag{4f}$$

Then, if we specify Neyman-Pearson significance level α *and* R_n *be one of the most powerful tests, then* $\{R_n\}$ *is ASC.*

PROOF. Let $Y_n:=\log(d\nu_0/d\nu_1)(X_n)$. Then, the most powerful test is a likelihood ratio test, hence it is a left-sided exact test based on $\overline{Y}_n$ except on the critical point. By using Theorem 3.2, we can easily reduce it to Example 4.1 (2).

REMARK. It is not clear whether (4f) is necessary or not, but we must not omit the assumption that ν_0 and ν_1 are mutually absolutely continuous. For a counter-example, let $\mathrm{H}_0:\theta=1,\mathrm{H}_1:\theta\neq 1$. Then, the most powerful test with Neyman-Pearson significance level α is essentially unique and it is

$$R_n=\{T_n<\sqrt[n]{\alpha}\quad\text{or}\quad 1<T_n\}$$

where $T_n:=\max_{1\leq j\leq n}X_j$. (Note that if H_1 is composed of only one point, it is not generally essentially unique.) If the prior distribution is $d\theta/\theta$ and take the shortest Bayesian confidence interval, the test based on it coincides with R_n. Since

$$P_{\theta=1}(X_n^* \in R_n \quad \text{for some } n) = \alpha\left\{1+\sum_{n=2}^{\infty}(1-\alpha^{1/n(n-1)})^{n-1}\right\}$$
$$< \alpha\sum_{n=1}^{\infty}(1-\alpha)^{n-1} = 1,$$

$\{R_n\}$ is non-WSC at H_0. Also we easily get that it is ASC at H_1.

The following theorem assures us that in order to judge SSC or ASC when there is a nuisance parameter, we can reduce the problem to the case that the nuisance parameter is known under some regularity conditions.

THEOREM 4.4. *Denote the parameter space by* $\Omega = \Omega^1 \times \Omega^2$ *and a parameter by* $\theta = (\xi, \eta)$. *Regard* η *as a nuisance parameter and consider the test*

$$\mathrm{H}_0 : \xi \in \Omega_0^1, \quad \mathrm{H}_1 : \xi \in \Omega_1^1.$$

Fix ξ_0 *and* η_0.

(I) *Assume that the following* (4g) *and* (4h) *hold.*

(4g) $\Omega^2 \subset \boldsymbol{R}^k$ and $\hat{\eta}_n = \hat{\eta}_n(X_n^*)$ *is a strongly consistent estimator of* η at $\eta = \eta_0$. Let $\hat{\eta}_n \in \Omega^{2*}, \Omega^2 \subset \Omega^{2*} \subset \boldsymbol{R}^k$.

(4h) $T_n = g_n(X_n^*)$ *is independent of* ξ, η *and real-valued, and if we fix* η, *the distribution of* T_n *is independent of* $\xi \in \Omega_0^1$ *and its asymptotic distribution exists as a probability distribution whose distribution function is continuous and we can take* $t_\infty : (0,1) \times \Omega^{2*} \to \boldsymbol{R}$ *such that*

$$\lambda_\eta((-\infty, t_\infty(\alpha, \eta))) = \alpha \qquad \text{for all} \quad \eta \in \Omega^2$$

and $\eta \mapsto t_\infty(\alpha, \eta)$ *is continuous on* Ω^{2*}.

Then, the following (1) *and* (2) *are equivalent.*

(1) *For any* α, *when* $\eta = \eta_0$ *is known, a left-sided test based on* T_n *with Neyman-Pearson significance level* α *is SSC at* $\xi = \xi_0$.

(2) *For any* α, *the test* $R_n^{(\alpha)} := \{T_n < t_\infty(\alpha, \hat{\eta}_n)\}$ *is SSC at* $\xi = \xi_0, \eta = \eta_0$.

(II) *In* (I), *if "a strongly consistent estimator", "SSC" are replaced by "a (weakly) consistent estimator", "ASC", respectively, then the similar result holds.*

PROOF. We shall prove that (1) implies (2). Fix α, then,

$$\lim_{n\to\infty} t_\infty(\alpha, \hat{\eta}_n) = t_\infty(\alpha, \eta), \quad P_{\xi_0, \eta_0} - \text{a.e.}$$

and

$$t_\infty(\alpha/2, \eta_0) < t_\infty(\alpha, \eta_0)$$

hold. Hence,

$$\lim_{\nu\to\infty} P_{\xi_0,\eta_0}(\text{ for all } n\geq \nu, t_\infty(\alpha/2,\eta_0)<t_\infty(\alpha,\hat{\eta}_n))=1.$$

And we get

$$\begin{aligned}
&P_{\xi_0,\eta_0}(\,T_n<t_\infty(\alpha,\hat{\eta}_0)\ \text{ i.o.})\\
&\geq P_{\xi_0,\eta_0}(\,T_n<t_\infty(\alpha/2,\eta_0)\ \text{ i.o.})\\
&\quad - P_{\xi_0,\eta_0}(\text{ not for all } n\geq\nu,\ \ t_\infty(\alpha/2,\eta_0)<t_\infty(\alpha,\hat{\eta}_n))\\
&= P_{\xi_0,\eta_0}(\text{ for all } n\geq\nu,\ \ t_\infty(\alpha/2,\eta_0)<t_\infty(\alpha,\hat{\eta}_n))\\
&\to 1 \ \ as \ \ \ \nu\to\infty.
\end{aligned}$$

We can similarly prove that (2) implies (1).

(II) Fix $\{n_j\}\subset \boldsymbol{N}$, $n_1<n_2<\cdots$. Then, there exists a subsequence $\{n_{j_k}\}$ such that

$$\lim_{k\to\infty}\hat{\eta}_{n_{j_k}}\to\eta_0,\qquad P_{\xi_0,\eta_0}-\text{a.e.},$$

and we can reduce it to (I).

REMARK. In (I), we must not replace "a strongly consistent estimator" by "a consistent estimator". For a counter-example, let $X_1, X_2, \ldots$ be i.i.d. random variables with $N(\xi,\sigma^2)$ where $\xi\leq 0, \sigma\geq 1$ and $H_0:\xi=0$, $H_1:\xi<0$. Let Φ be the distribution function of $N(0,1)$ and

$$T_n:=\begin{cases}\Phi(X_n) & \text{if } \ \Phi(X_n)<1/n,\\ \Phi(X_1) & \text{otherwise.}\end{cases}$$

Then, in the notation of Theorem 4.4, where $\eta=\sigma$,

$$\lambda_n((-\infty,t))=\Phi(\Phi^{-1}(t)/\sigma)\quad for\ \ 0<t<1,$$

and

$$t_\infty(\alpha,\sigma)=\Phi(\sigma\Phi^{-1}(\alpha))$$

hold. Fix ξ,σ. Then, for a sufficiently large n,

$$R_n^{(\alpha,\sigma)}\supset\{(X_n-\xi)/\sigma<\Phi^{-1}(1/n)\}$$

holds. Hence, from the Borel-Cantelli lemma, $\{R_n^{(\alpha,\sigma)}\}$ is SSC. On the other hand, let $\hat{\sigma}_n$ be a strongly consistent estimator of σ and

$$\hat{\sigma}_n^* :=\begin{cases} n & \text{if } \ \Phi(X_n)<1/n,\\ \hat{\sigma}_n & \text{otherwise.}\end{cases}$$

Then, $\hat{\sigma}_n^*$ is a consistent estimator of σ and $R_n^{(\alpha)} := \{T_n < t_\infty(\alpha, \hat{\sigma}_n^*)\}$ is non-SSC at $\xi = 0$ and any σ if $0 < \alpha < 1/2$ [By Chebyshev's inequality and the Borel-Cantelli lemma, we need only consider the case $\Phi(X_n) \geq 1/n.$].

EXAMPLE 4.3. Let $X_1, X_2, \ldots$ be i.i.d. random variables with $N(\xi, \sigma^2)$. Then, if we specify Neyman-Pearson significance level α, Student's tests

$$\text{(1)} \qquad \mathrm{H}_0 : \xi = 0, \quad \mathrm{H}_1 : \xi \neq 0 \quad \text{(the two-sided test)}$$

$$\text{(2)} \qquad \mathrm{H}_0 : \xi = 0, \quad \mathrm{H}_1 : \xi < 0 \quad \text{(the left-sided test)}$$

are ASC. Note that the interpretation of the left-sided test is more realistic to consider

$$\text{(3)} \qquad \mathrm{H}_0 : \xi \geq 0, \quad \mathrm{H}_1 : \xi < 0$$

than (2). This is ASC at $\xi \leq 0$, non-SSC at all $\xi > 0, \sigma > 0$. We can prove it by reducing it to the case that σ is known (Example 4.1) by using Theorem 3.2 and Theorem 4.4.

References

[1] Akaike, H., Selection of the prior distribution and its application, Y. Suzuki and N. Kunitomo. (eds.), Bayesian Statistics and its Applications (in Japanese), Univ. Tokyo Press, Tokyo , 1989.

[2] Bartlett, M. S., A comment on D. V. Lindley's statistical paradox, Biometrika, **44** (1957), 533–534.

[3] Basu, D., Statistical Information and Likelihood, J. K. Ghosh (ed.), Springer-Verlag, New York, Tokyo, 1988.

[4] Berger, J. O. and Wolpert, R. L., The Likelihood Principle, Institute of Math. Statistics Hayward, California, 1984.

[5] Cornfield, J., A Bayesian test of some classical hypotheses – with application to sequential clinical trials, J. Amer. Statist. Assoc. **61** (1966), 577–594.

[6] Feller, W., The general form of the so-called law of iterated logarithm, Trans. Amer. Math. Soc. **54** (1943) 373–402.

[7] Jeffreys, H., Theory of Probability, Oxford: Claredon Press, Oxford, New York, Tokyo *et al.*, 1961.

[8] Koike, K., Unbiased estimation for sequential multinomial sampling plans, Sequential Anal. **12** (1993), 253–259.

[9] Lindley, D. V., A statistical paradox, Biometrika, **44** (1957), 187–192.

[10] Lindley, D. V., Introduction of Probability and Statistics Part 2 Inference, Cambridge Univ. Press, London, New York, 1965.

[11] Lindley, D. V., Bayesian Statistics, A Review, Society for Industrial and Applied Math., Philadelphia, 1972.

[12] Lindley, D. V., Statistical inference concerning Hardy-Weinberg equilibrium. M. Bernardo *et al.* (eds.), Bayesian Statistics 3, Oxford Science Publications, Oxford, New York, Tokyo *et al.*, 1988.

[13] Lindley, D. V., Is our view of Bayesian statistics too narrow? M. Bernardo *et al.* (eds.), Bayesian Statistics 4, Oxford Science Publications, Oxford, New York, Tokyo *et al.*, 1992.
[14] Robbins, H., Some aspects of the sequential design of experiments, Bull. Amer. Math. Soc. **58** (1952), 527–535.
[15] Shigemasu, K., An Introduction to Bayesian Statistics (in Japanese), Univ. Tokyo Press, Tokyo 1985.

JSPS Research Fellow
Institute of Mathematics
University of Tsukuba
Ibaraki 305, Japan