

Successive process of statistical inference Associated with an Additive Family of Sufficient Statistics

Kitagawa, Toshio
Kyushu University

<https://doi.org/10.5109/12981>

出版情報 : 統計数理研究. 7 (3/4), pp.92-112, 1957-03. Research Association of Statistical Sciences

バージョン :

権利関係 :



SUCCESSIVE PROCESS OF STATISTICAL INFERENCE ASSOCIATED WITH AN ADDITIVE FAMILY OF SUFFICIENT STATISTICS

By

Tosio KITAGAWA

(Received Feb. 12, 1957)

§ 1. Introduction. The object of this paper is to introduce in § 2 the notion of additive family of sufficient statistics and also that of relative statistics associated with the former family under a certain restricted condition and to show in § 3~§ 7 how far these notions are useful for developing some aspects of two sample or several sample formulations, which seem to be of the fundamental importance for successive processes of statistical inferences developed by the author in a series of his papers [2]~[9]. The approach by means of these notions seems to be at least one of the possible ways by which to have deeper understandings about the statistical theory of R. A. FISHER which has been illustrated by himself in his recent monograph [1] more definitely than he has ever done. According to our standpoints, one of the possible approaches to his fiducial theory, which has been in fact remote from the current theory of mathematical statistics, is to rely upon the two sample formulation whose limiting case will at least suggest the idea of R. A. FISHER more relevantly than other authors have ever tried to understand. It has been also suggested that the fiducial arguments seem to be restricted within some restrictive classes of statistics. Our answer in this paper appealing to the notion of additive family of sufficient statistics and relative statistics does not assume the sole one to the understandings of the fiducial arguments, but may be at least one of its possible approaches.

In conclusion the author expresses his hearty thanks to Prof. R. A. FISHER and Prof. P. C. MAHALANOBIS for their discussions with him about the epistemological aspects of statistical inferences which he could enjoy while he was working as one member of the Reviewing Committee of the National Sample Survey in the Indian Statistical Institute during the period from December 2, 1956 to January 16, 1957, and which gave a strong stimulus for preparing this paper.

§ 2. Additive family of two independent sufficient statistics. Let u_m and u_n be two independent sufficient statistics of the same parameter τ having their probability density functions with respect to a common measure μ over the real axis R such that

$$(2.01) \quad p_m(u_m; \tau) d\mu(u_m) = \exp \{ -\tau u_m + b_m(\tau) + a_m(u_m) \} d\mu(u_m)$$

$$(2.02) \quad p_n(u_n; \tau) d\mu(u_n) = \exp \{ -\tau u_n + b_n(\tau) + a_n(u_n) \} d\mu(u_n)$$

respectively.⁽¹⁾ In what follows, let us consider without loss of generality the case when $d\mu(x) = d\mu(y)$ for all x and y so far as both of $d\mu(x)$ and $d\mu(y)$ do not vanish, that is, x and y belong to the carrier $Car.(\mu)$ of the measure μ . We assume that $x \in Car.(\mu)$ and $y \in Car.(\mu)$ imply $x + y \in Car.(\mu)$.

Lemma 2.1. *Under the assumption of this paragraph, the necessary and sufficient condition that $u_m + u_n$ be a sufficient statistic for τ having the probability density function with respect to the measure μ such that*

$$(2.03) \quad p(u_m + u_n; \tau) d\mu(u_m + u_n) \\ = \exp \{ -\tau(u_m + u_n) + b_m(\tau) + b_n(\tau) + a_{m,n}(u_m + u_n) \} \cdot d\mu(u_m + u_n)$$

is that there is a function $a_{m,n}(\cdot)$ such that

$$(2.04) \quad \exp \{ a_{m,n}(u) \} = \int_R \exp \{ a_m(u-v) \} \exp \{ a_n(v) \} d\mu(v).$$

Proof: The simultaneous probability density function of u_m and u_n is the product of (2.01) and (2.02). Now the change of u_m and u_n to $g = u_m + u_n$, $v = u_n$ gives us

$$(2.05) \quad \exp \{ -\tau g + b_m(\tau) + b_n(\tau) \} d\mu(g) \\ \cdot \exp \{ a_m(g-v) + a_n(v) \} d\mu(v)$$

The integration of the right-hand side of (2.05) gives us the probability density function of g with respect to the measure μ . It is now evident that the necessary and sufficient condition for (2.03) is given by (2.04).

Lemma 2.2. *Under the assumption that the hypothesis to Lemma 2.1 holds for all integers $m, n \geq 1$, the necessary and sufficient condition for $a_{m,n}(u) = a_{m+n}(u)$ for all $m, n \geq 1$ is that there holds the additivity such that*

$$(2.06) \quad \exp \{ a_m(u) \} = \int_R \exp \{ a_{m-1}(u-v) \} \exp \{ a_1(v) \} d\mu(v),$$

$$(2.07) \quad b_m(\tau) = mb_1(\tau) \equiv mb(\tau), \quad \text{say,}$$

for all positive integers $m \geq 1$.

Definition 2.1. We define, for $j \geq 1$ and for all u ,

$$(2.08) \quad A_j(u) = \exp \{ a_j(u) \}.$$

Proof of Lemma 2.2. The convolution relation yields us that $A_{i,j}(u) \equiv A_{i',j'}(u)$ for $i+j=i'+j'$ and $i, j, i', j' = 1, 2, 3, \dots$. Consequently we may

(1) See [10] ~ [13] and [15].

and we shall put $A_{m,n}(u) = A_{m+n}(u)$, which also shows the relation (2.07), as was to be proved. These two Lemmas gives us

Theorem 2.1. *Let $\{u_j\}$ ($j=1, 2, 3, \dots$) be a family of independent sufficient statistics for the same parameter τ having their probability density functions with respect to common measure μ over the real axis R such that (2.01) and (2.02) hold true respectively for positive integers m and n . Then the necessary and sufficient condition that $g = u_m + u_n$ be a sufficient statistic for τ having the probability density function with respect to the measure μ such that*

$$(2.09) \quad \begin{aligned} & p_{m+n}(g; \tau) d\mu(g) \\ &= \exp \{ -\tau g + (m+n)b(\tau) + a_{m+n}(g) \} d\mu(g) \end{aligned}$$

for any pair of positive integers m and n is that

$$(2.10) \quad A_{m+n}(u) = \int_R A_m(u-v) A_n(v) d\mu(v)$$

for any pair of positive integers m and n .

Under this condition we have

$$(2.11) \quad \begin{aligned} & p_m(u_m; \tau) d\mu(u_m) p_n(u_n; \tau) d\mu(u_n) \\ &= \exp \{ -\tau g + (m+n)b(\tau) + a_{m+n}(g) \} d\mu(g) \\ & \cdot \frac{A_m(g-u_n) A_n(u_n)}{A_{m+n}(g)} d\mu(u_n), \end{aligned}$$

where for each assigned value of g we have

$$(2.12) \quad \int_R \frac{A_m(g-u_n) A_n(u_n)}{A_{m+n}(g)} d\mu(u_n) = 1.$$

The proof is immediate. In view of (2.12) we have now the conditional probability density of u_n with respect to the measure μ for each assigned value of g , which does not contain the population parameter τ . Now our statistical inference theory will be particularly interested with the situation under which a suitable choice of variables will make the conditional probability density independent of the value of g . We prepare ourselves with

Theorem 2.2. *Under the hypothesis to Theorem 2.1, the necessary and sufficient condition that there are continuously differentiable functions $f_{m,n}(g, h)$ and a function $D_{m,n}(h)$ with*

$$(2.13) \quad u_m = f_{m,n}(g, h), \quad u_n = g - f_{m,n}(g, h)$$

for which

$$\begin{aligned}
 (2.14) \quad & p(u_m; \tau) p(u_n; \tau) d\mu(u_m) d\mu(u_n) \\
 &= \exp \{ -\tau g + (m+n)b(\tau) + a_{m+n}(g) \} d\mu(g) \\
 &\quad \cdot D_{m,n}(h) d\mu(h)
 \end{aligned}$$

is that we have

$$(2.15) \quad \frac{\exp \{ a_m(f_{m,n}(g, h)) \} \exp \{ a_n(g - f_{m,n}(g, h)) \}}{\exp \{ a_{m+n}(g) \}} \frac{\partial f_{m,n}(g, h)}{\partial h}$$

is independent of g and is equal to $D_{m,n}(h)$.

Example 2.1. For the normal distribution $N(\tau, 1)$ with unknown τ , we are concerned with

$$(2.16) \quad p(u_j; \tau) = (2\pi j)^{-1/2} \exp \left\{ -\tau u_j - \frac{j\tau^2}{2} - \frac{u_j^2}{2j} \right\}$$

for $j = m, n$. Consequently we have

$$(2.17) \quad \exp \{ a_j(u_j) \} = A_j(u_j) = (2\pi j)^{-1/2} \exp \left\{ -\frac{u_j^2}{2j} \right\}$$

$$(2.18) \quad b_j(\tau) = -j\tau^2/2$$

for $j = m, n, m+n$. The direct calculation gives us

$$\begin{aligned}
 (2.19) \quad & \exp \{ a_m(u_m) \} \exp \{ a_n(u_n) \} \exp \{ -a_{m+n}(u_m + u_n) \} \\
 &= \frac{1}{\left(2\pi \left(\frac{1}{m} + \frac{1}{n} \right) \right)^{1/2}} \exp \left\{ -\frac{1}{2 \left(\frac{1}{m} + \frac{1}{n} \right)} \left(\frac{u_m}{m} - \frac{u_n}{n} \right)^2 \right\},
 \end{aligned}$$

which shows that a choice of functions $f_{m,n}$ and $D_{m,n}$ can be given by

$$(2.20) \quad u_m = \frac{m}{m+n}(g + nh), \quad u_n = g - \frac{m}{m+n}(g + nh),$$

$$(2.21) \quad h = \frac{u_m}{m} - \frac{u_n}{n}.$$

Example 2.2. For the gamma distribution with an unknown parameter τ with the form

$$(2.22) \quad p(u_j; \tau) = \tau^j (\Gamma(j))^{-1} u_j^{j-1} e^{-\tau u_j}$$

for $j = m, n$. We have in this case

$$(2.23) \quad \exp \{ a_j(u_j) \} = A_j(u_j) = (\Gamma(j))^{-1} u_j^{j-1},$$

$$(2.24) \quad b_j(\tau) = j \log \tau$$

for $j = m, n, m+n$, which gives us

$$(2.25) \quad \exp \{a_m(u_m)\} \exp \{a_n(u_n)\} \exp \{-a_{m+n}(u_m + u_n)\} \\ = \frac{\Gamma(m+n)}{\Gamma(m)\Gamma(n)} \cdot \frac{u_m^{m-1} u_n^{n-1}}{(u_m + u_n)^{m+n-1}}.$$

But the change of variables u_m and u_n to g and h by the relation

$$(2.26) \quad g = u_m + u_n, \quad u_m = gh$$

gives us

$$(2.27) \quad \frac{\Gamma(m+n)}{\Gamma(m)\Gamma(n)} h^{m-1} (1-h)^{n-1} dh = D_{m,n}(h) dh$$

with $h = u_m(u_m + u_n)^{-1}$. The transformation $F = nh/m(1-h)$ gives us the probability function of the F -distribution with the pair of the degrees of $[2m, 2n]$ such that

$$(2.28) \quad \frac{\Gamma(m+n) m^m n^n}{\Gamma(m)\Gamma(n)} \cdot \frac{F^{m-1}}{(mF+n)^{m+n}} dF.$$

Example 2.3. When x_i 's ($i=1, 2, \dots, m$) are drawn independently from the Poisson distribution with an unknown mean λ , $0 < \lambda < \infty$. Then $u = \sum_{i=1}^m x_i$ is a sufficient statistic for the parameter $\tau = -\log \lambda$ with the probability density function⁽²⁾

$$(2.29) \quad \exp \{-\tau u - m e^{-\tau} + u \log m - \log \Gamma(u+1)\} d\mu(u),$$

where μ is a step-function having jumps of height 1 at every nonnegative integer. In this case we have

$$(2.30) \quad a_m(u) = u \log m - \log \Gamma(u+1),$$

$$(2.31) \quad b_m(\tau) = -m e^{-\tau}.$$

Consequently we have

$$(2.32) \quad \frac{A_m(g-v) A_n(v)}{A_{m+n}(g)} = \frac{n_1^{g-v} n_2^v}{(n_1 + n_2)^g} \cdot \frac{\Gamma(g+1)}{\Gamma(g-v+1) \Gamma(v+1)} \\ = {}_g C_v p^v (1-p)^{g-v},$$

where we have put $p = n_2(n_1 + n_2)^{-1}$.

Example 2.4. When x_i 's ($i=1, 2, \dots, m$) are drawn independently from the binomial distribution with an unknown parameter p . Then $u = \sum_{i=1}^m x_i$ is a sufficient statistic for the parameter $\tau = \log \{p(1-p)^{-1}\}$ with the probability density function⁽³⁾

$$(2.33) \quad \exp \{-\tau u - m \log(1 + e^{-\tau}) + \log B(m+1-u, u+1)\}$$

(2) For this transformation $\tau = -\log \lambda$, see [15] p. 92.

(3) For this transformation $\tau = \log \{p(1-p)^{-1}\}$, see [15] p. 92.

for $0 \leq u \leq m$ and elsewhere zero, with respect to the measure μ , which has the jump of height 1 at every positive integer.

In this case we have

$$(2.34) \quad b_m(\tau) = -m \log(1 + e^{-\tau}), \quad a_m(u) = \log B(m+1-u, u+1).$$

Consequently we have for $0 \leq v \leq n$

$$(2.35) \quad \frac{A_m(g-v)A_n(v)}{A_{m+n}(g)} = \frac{B(m+1, g-v+1)B(n+1, v+1)}{B(m+n+1, g+1)}$$

and elsewhere vanishes with respect to the measure μ .

Definition 2.2. A family of independent sufficient statistics for which (2.10) holds true is called to be an *additive family* of independent sufficient statistics with respect to the parameter τ .

Definition 2.3. The statistic $h = \varphi_{m,n}(u_m, u_n)$ satisfying the condition enunciated in (2.15) is called to be a *relative statistic* between u_m and u_n associated with the additive family of independent statistics with respect to the parameter τ .

§ 3. Some characterisations of additive families of independent sufficient statistics. First we shall prepare ourselves with the following two functional equations which may have some interests on themselves.

Lemma 3.1. Let $\{\theta_j(u)\}$ ($j=1, 2, 3, \dots$) be a set of functions defined over $-\infty < u < \infty$ and integrable in the sense of Lebesgue in any finite interval of u .

Let us assume that there holds a set of functional equations

$$(3.01) \quad \theta_m(u) + \theta_n(v) = \theta_{m+n}(u+v) + \psi_{m,n}(u + c_{m,n}v)$$

for all u and v in $-\infty < u, v < \infty$ and for all $m, n=1, 2, 3, \dots, c_{m,n}$ being constants which are independent of m and n and are neither equal to 1 nor 0.

Then we have

$$(3.02) \quad \theta_m(u) = \frac{\alpha}{m}(u - \beta m)^2 + \tau_m,$$

$$(3.03) \quad \psi_{m,n}(u + c_{m,n}v) = \frac{\alpha n}{(m+n)m} \left(u - \frac{m}{n}v\right)^2 + \tau_m + \tau_n - \tau_{m+n},$$

where α, β and τ_m are constants, α and β being independent of m and n .

Proof: In view of (3.01), $\psi_{m,n}(u + cv)$ is L -integrable in any finite interval of u or v , and we have

$$(3.04) \quad \int_0^u \theta_m(w)dw + u\theta_n(v) = \int_0^u \theta_{m+n}(w+v)dw + \int_0^u \psi_{m,n}(w + c_{m,n}v)dw.$$

Since now the right-hand side of (3.04) is differentiable with respect to v , it turns out that $\theta_n(v)$ is differentiable for $n=1, 2, 3, \dots$. In view of (3.01) $\psi_{m,n}$ is also differentiable, and hence the right-hand side of (3.04) is twice differentiable. This leads us to the consequence that $\theta_n(v)$ is twice-differentiable for $n=1, 2, 3, \dots$. Similarly we can proceed and we can show that $\theta_n(v)$ is infinitely differentiable.

The twice differentiations of each of two sides of (3.01) first with respect to u and secondly with respect to v give us

$$(3.05) \quad \theta''_{m+n}(u+v) + c_{m,n} \psi''_{m,n}(u + c_{m,n}v) = 0$$

for all u, v in $-\infty < u, v < \infty$. Let us put $u+v=w$.

Then (3.05) and $c_{m,n} \neq 1$ lead us that $\theta''_{m+n}(w) = \text{const}$ independent of w and consequently that

$$(3.06) \quad \theta_{m+n}(w) = a_{m+n}w^2 + b_{m+n}w + c_{m+n}$$

for all $m+n \geq 2$.

Again now (3.05) gives us that $\psi_{m,n}(u + c_{m,n}v)$ is at most quadratic function of $u + c_{m,n}v$ for $m, n \geq 1$. Specially when $m=n=1$ and $u=v$, the functional equation (3.01) shows that we can write

$$(3.07) \quad \theta_1(w) = a_1w^2 + b_1w + c_1.$$

Consequently we have a set of functional relations

$$(3.08) \quad \begin{aligned} & a_m u^2 + b_m u + c_m + a_n v^2 + b_n v + c_n \\ &= a_{m+n}(u+v)^2 + b_{m+n}(u+v) + c_{m+n} \\ &+ p_{m,n}(u + c_{m,n}v)^2 + q_{m,n}(u + c_{m,n}v) + r_{m,n} \end{aligned}$$

for all $m, n \geq 1$. The comparisons of each of the coefficients of u^2, v^2 and uv give us

$$(3.09) \quad a_m = a_{m+n} + p_{m,n},$$

$$(3.10) \quad a_n = a_{m+n} + p_{m,n}c_{m,n}^2,$$

$$(3.11) \quad a_{m+n} + p_{m,n}c_{m,n} = 0.$$

The combinations of these three relations yield us the two relations

$$(3.12) \quad a_n = a_{m+n}(1 - c_{m,n})$$

$$(3.13) \quad c_{m,n}a_m = a_{m+n}(c_{m,n} - 1)$$

for any pair of two positive integers m and n . Since $c_{m,n} \neq 1, 0$, it follows that either all a_m are zero for $m=1, 2, 3, \dots$ or none of a_m are zero and there holds

$$(3.14) \quad a_m^{-1} + a_n^{-1} = a_{m+n}^{-1}$$

for any pair of two positive integers m and n . Both of these two cases give us $a_m = \alpha m^{-1}$ where α is a constant independent of m and may or may not be equal to zero. When $\alpha \neq 0$, then the relations $a_m = \alpha m^{-1}$ gives us in view of (3.12) and (3.11), $c_{m,n} = -mn^{-1}$, $p_{m,n} = \alpha n((m+n)m)^{-1}$, and hence

$$(3.15) \quad b_m = b_{m+n} + q_{m,n},$$

$$(3.16) \quad b_n = b_{m+n} - \frac{m}{n} q_{m,n}$$

and hence the functional equation

$$(3.17) \quad mb_m + nb_n = (m+n)b_{m+n},$$

which leads us $b_m = b = \text{const}$, hence $q_{m,n} = 0$, $r_{m,n} = c_m + c_n - c_{m+n}$. When $\alpha = 0$, we have $p_{m,n} = 0$ and

$$(3.18) \quad b_m = b_{m+n} + q_{m,n},$$

$$(3.19) \quad b_n = b_{m+n} + c_{m,n} q_{m,n},$$

which gives us

$$(3.20) \quad c_{m,n} b_m - b_n = (c_{m,n} - 1) b_{m+n},$$

whose solutions are given by $b_m = b = \text{const}$. Also we have $r_{m,n} = c_m + c_n - c_{m+n}$, q.e.d.

Lemma 3.2. Let $\{\theta_j(u)\}$ ($j=1, 2, \dots$) be a set of functions defined over $-\infty < u < \infty$ and integrable in the sense of Lebesgue in any finite interval of u .

Let us assume that there holds a set of functional equations

$$(3.21) \quad \theta_m(xy) + \theta_n(x(1-y)) + \log x = \theta_{m+n}(x) + \psi_{m,n}(y)$$

for all x in $0 < x < \infty$ and all y in $0 < y < 1$ and for all pairs of positive integers m and n .

Then there are constants $\alpha \neq 0$, β and $\{\gamma_j\}$ such that

$$(3.22) \quad \theta_m(x) = (\alpha m - 1) \log x + \beta x + \gamma_m,$$

$$(3.23) \quad \psi_{m,n}(x) = (\alpha m - 1) \log x + (\alpha n - 1) \log(1-x) \\ + \gamma_m + \gamma_n - \gamma_{m+n}.$$

Proof: Putting $x=1$ in (3.21) we have

$$(3.24) \quad \psi_{m,n}(y) = \theta_m(y) + \theta_n(1-y) - \theta_{m+n}(1),$$

which gives in combination of (3.21)

$$(3.25) \quad \theta_m(xy) + \theta_n(x(1-y)) + \log x \\ = \theta_{m+n}(x) + \theta_m(y) + \theta_n(1-y) - \theta_{m+n}(1).$$

The argument similar to the proof of Lemma 3.1 shows that $\theta_m(x)$ are differentiable for all x and m . The differentiation of both sides of (3.25) with respect to y gives us

$$(3.26) \quad x\theta'_m(xy) - x\theta'_n(x(1-y)) = \theta'_m(y) - \theta'_n(1-y)$$

and specially for $y=1/2$ we have

$$(3.27) \quad x\theta'_m\left(\frac{x}{2}\right) - x\theta'_n\left(\frac{x}{2}\right) = \theta'_n\left(\frac{1}{2}\right) - \theta'_n\left(\frac{1}{2}\right)$$

for all x in $0 < x < \infty$.

Let us now choose $n=1$ and consider x instead of $x/2$. Then the integrations of (3.27) give us

$$(3.28) \quad \theta_m(x) = \theta_1(x) + (a_m - a_1) \log x + (b_m - b_1),$$

where a_m, b_m ($m=1, 2, \dots$) are constants independent of x and will be determined below.

Let $m=n=1$ in (3.25), which then gives us in view of (3.28),

$$(3.29) \quad \begin{aligned} & \theta_1(xy) + \theta_1(x(1-y)) + \log x \\ &= \theta_1(x) + (a_2 - a_1) \log x \\ & \quad + \theta_1(y) + \theta_1(1-y) - \theta_1(1). \end{aligned}$$

The differentiation of both sides of (3.29) with respect to x gives us

$$(3.30) \quad y\theta'_1(xy) + (1-y)\theta'_1(x(1-y)) = \theta'_1(x) + (a_2 - a_1 - 1)x^{-1}.$$

Now let us multiple both sides of (3.30) by x , and then put $xy=u$, $x(1-y)=v$, $x\theta'_1(x) \equiv \varphi(x)$. Then we have the functional equation

$$(3.31) \quad \varphi(u) + \varphi(v) = \varphi(u+v) + (a_2 - a_1 - 1),$$

which holds true for all $u, v < 0$. The solution of the functional equation gives us

$$(3.32) \quad u\theta'_1(u) \equiv \varphi(u) = \beta u + (a_2 - a_1 - 1)$$

and hence we can write, in view of (3.28)

$$(3.33) \quad \theta_m(u) = \beta u + A_m \log u + r_m$$

with constants β, A_m and r_m . Putting the formula (3.33) into (3.25), there must hold the relations $A_{m+n} = A_m + A_n + 1$, for all positive integers m and n which give $A_m = \alpha m - 1$. q.e.d.

In view of the two Lemmas 3.1 and 3.2, we shall observe the following two Theorems which characterise normal distribution and gamma distributions respectively.

Theorem 3.1. *Under the hypothesis to Theorem 2.2, let us consider the case when $\mu(x)$ becomes the linear Lebesgue measure, that is, $d\mu(x)$*

$=dx$ in $(-\infty < x < \infty)$. Then the necessary and sufficient condition that the function $f_{m,n}(g, h)$ has the special form

$$(3.34) \quad g = u_m + u_n, \quad h = u_m + c_{m,n} u_n$$

with constant $c_{m,n} \neq 0, 1$ for $m, n = 1, 2, 3, \dots$, and for which (2.15) holds true is that there are constants $\sigma^2 > 0$ and β such that

$$(3.35) \quad p(u_m; \tau) = (2\pi m \sigma^2)^{-1/2} \exp \left\{ -\frac{(u_m + \sigma m(\tau - \beta))^2}{2\sigma^2 m} \right\},$$

which can be reduced to the normalised form

$$(3.36) \quad p(u_m^*; \tau^*) = \exp \left\{ -\tau^* u_m^* + m \tau^{*2} + \frac{u_m^{*2}}{2m} - \frac{1}{2} \log 2\pi m \right\}$$

by the transformation

$$(3.37) \quad u_m^* = \sigma^{-1} u_m, \quad \tau^* = \sigma(\tau - \beta).$$

Proof: The sufficiency is immediate. The necessity can be observed from Lemma 3.1.

Theorem 3.2. Under the hypothesis to Theorem 2.2, let us consider the case when $\mu(x)$ becomes the linear Lebesgue measure, that is, $d\mu(x) = dx$ in $-\infty < x < \infty$. Then the necessary and sufficient condition that the function $f_{m,n}(g, h)$ has the special form for which

$$(3.38) \quad u_m = gh, \quad u_n = g(1-h)$$

and for which (2.10) holds true is that

$$(3.39) \quad p(u_m; \tau) = (\Gamma(\alpha m))^{-1} \tau^{\alpha m} \exp \{ -\tau u \} u^{\alpha m - 1}$$

with a positive constant α .

Proof: The sufficiency is evident. The necessity follows from the fact that the relations

$$(3.40) \quad A_m(gh) A_n(g(1-h)) g = A_{m+n}(g) B_{m,n}(h)$$

are equivalent to the functional equations (3.21), provided that $\log A_m(\cdot) = \theta_m(\cdot)$, $\log B_{m,n}(\cdot) = \psi_{m,n}(\cdot)$ and $g = x$, $h = y$.

§ 4. Two sample formulation based on the decomposition of two independent additive statistics. The two Theorems 2.1 and 2.2 will be now shown to be of the fundamental importance in developing two sample theory in a general formulation which we discussed in the previous paper of ours [2] rather with reference to the particular forms of distribution-functions. We are not trying to give any mathematical refinements beyond the two Theorems, but we are rather concerned with the interpretations

and the uses and of the simple mathematical formula. Let us begin with the simplest cases among the two possible approaches, that is, the case when there exists a family of relative statistics $\varphi_{m,n}(u_m, u_n)$ defined in Definition 2.2.

Theorem 4.1. *Under the hypothesis to Theorem 2.2, let us assume that the condition (2.15) is satisfied.*

Let us assume furthermore that for an assigned value of α in $0 < \alpha < 1$ we can find out a domain $R_{m,n}(\alpha)$ of the real line of h such that

$$(4.01) \quad \int_{R_{m,n}(\alpha)} D_{m,n}(h) d\mu(h) = \alpha.$$

Let us write $h = \varphi_{m,n}(u_m, u_n)$. Then we have

$$(4.02) \quad Pr. \{ \varphi_{m,n}(u_m, u_n) \in R_{m,n}(\alpha) \} = \alpha.$$

The interpretations of the simple formula (4.02) can be done in several ways. The definition of $R_{m,n}(\alpha)$ and the value of α may be different according to their uses and their interpretations. Among others there are three fundamental interpretations.

[1] *The prediction.* Let u_m be a sufficient statistic which we have already obtained, and let u_n be another independent statistic to be obtained in another occasion, where u_m and u_n are distributed according to (2.01) and (2.02) respectively. Under this condition the formula (4.02) gives the domain of the values of u_n to be expected to belong for an assigned value of α .

[2] *The test of significance.* Let u_m be a sufficient statistic distributed according to (2.01), and let u_n be another independent statistic distributed according to the distribution of the form (2.02) however with parameter τ^* which is not necessarily equal to τ .

The null hypothesis $H_0: \tau = \tau^*$ can be tested by the formula (4.02) by means of u_m and u_n . That is to say, for an assigned significance level of α , we shall reject the null hypothesis H_0 when $\varphi_{m,n}(u_m, u_n)$ does belong to the domain $R_{m,n}(\alpha)$.

[3] *The fiducial argument.* After our understandings to the theory of fiducial argument due to R. A. FISHER, there are two conditions for the application of the fiducial argument. The first is the absence of an information a priori. The second is that it is concerned solely with sufficient statistics.

Our formulation discussed in §2 satisfies both of these two conditions. Now the author prefers here the argument of two sample formulation, because it contains the prediction, the test of significance as we have just referred and moreover it will give some explanation in the limiting case for the idea of the fiducial argument concerning the unknown parameter.

This argument is based upon the argument to the effect that, when the size of sample n tends to the infinity, the sequence of the functions of u_n will tend to the population parameter τ . We have already mentioned the interpretation of this sort in our previous paper [1]. This interpretation seems to the author to be one of the possible approaches to the fiducial argument about an unknown parameter τ which R. A. FISHER can agree with, because he points out in his recent monograph [1] p. 114 the equivalence of fiducial probability statements to the predictions in the form of probability statements about the future observations.

Now let us turn to the more general situation when the formula (2.15) does not necessarily hold but merely the formula (2.04), (2.06) and (2.07) hold true. Under this situation we have to deal with the conditional probability density function of v for each assigned value of $g = u + v$ such that

$$(4.03) \quad \frac{A_m(v) A_n(g-v)}{A_{m+n}(g)} d\mu(v).$$

Because of the fact that the sum of u and v has an assigned value, some of the possibilities of statistical inferences are not so straight as we have just shown for the case when (2.10) holds true. Indeed it is at least very difficult to establish the prediction [1] and the fiducial argument [3] for the general case (4.03), because of the lack of the statistic $h = \varphi_{m,n}(u, v)$ which is independent of $g = u + v$.

However the test of significance [2] can be established generally under our present situation.

Theorem 4.2. *Under the hypothesis to Theorem 2.1, let us assume that for an assigned value of α in $0 < \alpha < 1$ we can find out a domain $R_{m,n}(g; \alpha)$ of the real line v such that*

$$(4.04) \quad \int_{R_{m,n}(g; \alpha)} \frac{A_m(v) A_n(g-v)}{A_{m+n}(g)} d\mu(v) = \alpha.$$

Then the probability that, under the condition that the sum of the two independent sufficient statistics $u_m + u_n$ has the assigned value g , that is, $u_m + u_n = g$, the probability of u_n belonging to the domain $R_{m,n}(g; \alpha)$ is equal to α , namely

$$(4.05) \quad Pr.\{u_n \in R_{m,n}(g; \alpha); u_m + u_n = g\} = \alpha.$$

This Theorem can be applied when we have already obtained the values of u_m and u_n . We have then a certain value of the sum $g = u_m + u_n$. We may restrict ourselves within the reference set for which $u_m + u_n$ is equal to an assigned value of $g = u_m + u_n$. It is to be noted that (4.05) has no unknown parameter τ , and that under our situation the statistic $u_m + u_n$ exhausts all of the relevant informations for estimating the unknown parameter τ .

The uses and the implications of the two Theorems will be explained by the following examples.

Example 4.1. For Example 2.1 let us define the function k_α of α such that, for each assigned value of α in $0 < \alpha < 1$, we have

$$(4.06) \quad \Phi(k_\alpha) - \Phi(-k_\alpha) \equiv \frac{1}{\sqrt{2\pi}} \int_{-k_\alpha}^{k_\alpha} \exp\left\{-\frac{u^2}{2}\right\} d\mu = \alpha.$$

Then we can define the following formula:

(a) The domain of rejection for an assigned level of significance α

$$(4.07) \quad R_{m,n}(\alpha) = \left[h; |h| \geq k_{1-\alpha} \left(\frac{1}{m} + \frac{1}{n} \right)^{1/2} \right].$$

(b) The prediction formula for u_n by means of u_m with the confidence probability α

$$(4.08) \quad n \left(\frac{u_m}{m} - k_\alpha \left(\frac{1}{m} + \frac{1}{n} \right)^{1/2} \right) < u_n < n \left(\frac{u_m}{m} + k_\alpha \left(\frac{1}{m} + \frac{1}{n} \right)^{1/2} \right),$$

which has been derived from the relation

$$(4.09) \quad |h| \leq k_\alpha \left(\frac{1}{m} + \frac{1}{n} \right)^{1/2}.$$

(c) The fiducial argument for the unknown parameter τ by means of the statistic u_m . This can be shown as the limiting case of the probability distribution function of $h = u_m/m - u_n/n$

$$(4.10) \quad \left\{ 2\pi \left(\frac{1}{m} + \frac{1}{n} \right) \right\}^{-1/2} \exp \left\{ -\frac{1}{2} \left(\frac{1}{m} + \frac{1}{n} \right)^{-1} \left(\frac{u_m}{m} - \frac{u_n}{n} \right)^2 \right\} d \left(\frac{u_n}{n} - \frac{u_m}{m} \right)$$

when n tends to infinity, which gives us

$$(4.11) \quad (2\pi)^{-1/2} m^{1/2} \exp \left\{ -\frac{m}{2} \left(\tau - \frac{u_m}{m} \right)^2 \right\} d \left(\tau - \frac{u_m}{m} \right).$$

Exactly speaking what we are concerned with is the limiting asymptotic probability distribution of u_n/n as n tends to infinity. However at least symbolically it is simpler to appeal to (4.11), and there may not be any confusion so far as we are concerned with the application of Theorem 4.1.

Example 4.2. For Example, let us define the function $F_{\nu_2}^{\nu_1}(\alpha)$ of such that, for each assigned value of α in $0 < \alpha < 1$, we have

$$(4.12) \quad \frac{\nu_1^{\nu_1} \nu_2^{\nu_2}}{B \left(\frac{\nu_1}{2}, \frac{\nu_2}{2} \right)} \int_{F_{\nu_2}^{\nu_1}(\alpha)}^{\infty} \frac{F^{\nu_1/2-1}}{(\nu_1 F + \nu_2)^{(\nu_1+\nu_2)/2}} dF = \alpha.$$

Then we can define the following formula:

(a) The domain of rejection for an assigned level of significance α .

$$(4.13) \quad \begin{aligned} R_{m,n}(\alpha) &= [F; F \geq F_{2n}^{2m}(\alpha)] \\ &= [u_m/u_n; u_m \geq n^{-1}u_n m F_{2n}^{2m}(\alpha)]. \end{aligned}$$

(b) The prediction formula for u_m by means of u_n with the confidence probability α . There are a lot of possibilities. One of them is given by

$$(4.14) \quad mn^{-1}u_n F_{2n}^{2m}\left(1 - \frac{\alpha}{2}\right) < u_m < mn^{-1}u_n F_{2n}^{2m}\left(\frac{\alpha}{2}\right).$$

(c) The fiducial argument. It can be readily seen that we shall have for any assigned positive number G_1 and G_2 such that $G_1 < G_2$

$$(4.15) \quad \lim_{n \rightarrow \infty} Pr. \left\{ G_1 < \frac{m u_m}{u_n} < G_2 \right\} = \int_{G_1}^{G_2} \frac{h^{m-1} e^{-h}}{\Gamma(m)} dh,$$

which gives

$$(4.16) \quad \lim_{n \rightarrow \infty} Pr. \left\{ \frac{G_1}{u_m} < \frac{n}{u_n} < \frac{G_2}{u_m} \right\} = \int_{G_1}^{G_2} \frac{h^{m-1} e^{-h}}{\Gamma(m)} dh.$$

Now the law of large number assures us that n/u_n tends to the population parameter τ with probability one. Hence with probability one we have

$$(4.17) \quad Pr. \left\{ \frac{G_1}{u_m} < \tau < \frac{G_2}{u_m} \right\} = \int_{G_1}^{G_2} \frac{h^{m-1} e^{-h}}{\Gamma(m)} dh.$$

Example 4.3. Let u and v be two independent statistics having their probability density functions of (2.09) for u and

$$(4.18) \quad \exp \{ -\tau^* v - n e^{-\tau^*} + v \log n - \log \Gamma(v+1) \} d\mu(v)$$

for v as in Example 2.3. The hypothesis $H_0: \tau = \tau^*$ can be tested by defining the domain of rejection $R_p(g; \alpha)$ such that

$$(4.19) \quad \sum_{v \in R_p(g; \alpha)} {}_0C_v p^v (1-p)^{g-v} = \alpha,$$

where $p = n_2(n_1 + n_2)^{-1}$ and $g = u + v$.

Example 4.4. For the case of the binomial distribution, the argument similar to Example 4.3 holds. Here we appeal to the hypergeometric distribution of the type (2.35) for which we choose $R_{m,n}(g; \alpha)$ such that

$$(4.20) \quad \sum_{v \in R_{m,n}(g; \alpha)} \frac{B(m+1, g-v+1) B(n+1, v+1)}{B(m+n+1, g+1)} = \alpha.$$

It is to be noted that we have not entered the details how to choose the domain $R_{m,n}(\alpha)$ and $R_{m,n}(g; \alpha)$ among possible choices of them and how to manage the difficulties associated with discontinuous distributions of the types (4.19) and (4.20). We have rather concerned with the broad aspects of the uses of additive families of sufficient statistics.

§ 5. **Several sample formulations and limiting procedures associated with them.** We have been so far concerned with the two sample formulation and some of statistical inferences associated with them. Our notions of an additive family of independent sufficient statistics and a family of relative statistics associated with the former make it possible to give several sample formulations. In fact our additive family of independent sufficient statistics gives us

$$(5.01) \quad \prod_{j=1}^k \exp \left\{ -\tau u_{m_j} + m_j b(\tau) + a_{m_j}(u_{m_j}) \right\} \\ = \exp \left\{ -\tau \sum_{j=1}^k u_{m_j} + \sum_{j=1}^k m_j b(\tau) + a_{M_k} \left(\sum_{j=1}^k u_{m_j} \right) \right\} \\ \cdot \left[A_m \left(\sum_{j=1}^k u_{m_j} \right) \right]^{-1} \prod_{j=1}^k A_{m_j}(u_{m_j}),$$

where we have put $M_k = \sum_{j=1}^k m_j$. The cases when there exists a family of relative statistics can be discussed more in details, and there holds an associative law concerning the family of relative statistics such as

$$(5.02) \quad D_{m_1, m_2}(\varphi_{m_1, m_2}(u_{m_1}, u_{m_2})) D_{m_1+m_2, m_3}(\varphi_{m_1+m_2, m_3}(u_{m_1}+u_{m_2}, u_{m_3})) \\ d\varphi_{m_1, m_2}(u_{m_1}, u_{m_2}) d\varphi_{m_1+m_2, m_3}(u_{m_1}+u_{m_2}, u_{m_3}) \\ \equiv d\varphi_{m_1, m_2, m_3}(u_{m_1}, u_{m_2}, u_{m_3})$$

is invariant under the permutation of (m_1, m_2, m_3) to $(m_{\alpha_1}, m_{\alpha_2}, m_{\alpha_3})$ when $(\alpha_1, \alpha_2, \alpha_3)$ is any permutation of $(1, 2, 3)$. These are rather simple observations, but will serve to some sorts of statistical inferences discussed in our previous papers [2] ~ [9]. Let us begin with the two sample formulation.

In § 4 we have already referred to the limiting procedure to be applied to our two sample formulation by means of which to show that the fiducial argument of R. A. FISHER can be considered as the limiting case of two sample observation where the size of one of the two sample tends to infinity. This can be proved under a fairly general condition, although it will be given in somewhat complicated expression. Indeed we shall observe.

Theorem 5.1. *Let us assume that under the hypothesis to Theorem 2.1 the mean and the variance of the statistic u_1 exists and is finite.*

Let x, x_1, x_2, v_1 and v_2 be any assigned real numbers such that $x_1 < x_2$ and $v_1 < v_2$. Let us write $g = u_{m+n} = u_m + u_n$.

Then the following assertions hold

(1°) *We have*

$$(5.03) \quad \lim_{m \rightarrow \infty} \text{Pr.} \left\{ x_1 < \frac{u_m + u_n - (m+n)\bar{\xi}}{\sigma_1' m + n} \leq x_2, \quad v_1 < u_n < v_2 \right\} \\ = (\Phi(x_2) - \Phi(x_1)) \int_{v_1}^{v_2} \hat{p}_n(v; \tau) d\mu(v).$$

(2°) We have

$$(5.04) \quad \lim_{m \rightarrow \infty} \text{Pr.} \{ (m+n)\xi + \sigma x_1 \sqrt{m+n} < g < (m+n)\xi + \sigma x_2 \sqrt{m+n} \} \\ = \Phi(x_2) - \Phi(x_1).$$

(3°) We have

$$(5.05) \quad \lim_{m \rightarrow \infty} \int_{v_1}^{v_2} e^{-\tau v} A_n(v) d\mu(v) \cdot \frac{\int_{g_1}^{g_2} e^{-\tau(g-v)} A_m(g-v) d\mu(g-v)}{\int_{g_1}^{g_2} e^{-\tau g} A_{m+n}(g) d\mu(g)} \\ = \int_{v_1}^{v_2} p_n(v; \tau) d\mu(v),$$

where $g_i = (m+n)\xi + \sigma x_i \sqrt{m+n}$ ($i = 1, 2$).

Proof :⁽⁴⁾ Our statistic u_j belonging to our additive family of independent sufficient statistics can be written as the sum of j independent stochastic variables each of whose summands has the same probability density function with respect to the measure μ , and hence the statistic u_j has the mean $j\xi$ and the variance $j\sigma^2$.

Consequently the central limit theorem can be applied to the probability event enciated in the left-hand side (5.03), which gives us

$$(5.06) \quad \lim_{m \rightarrow \infty} \text{Pr.} \left\{ x_1 < \frac{u_m + u_n - (m+n)\xi}{\sigma \sqrt{m+n}} \leq x_2, \quad v_1 < u_n \leq v_2 \right\} \\ = \lim_{m \rightarrow \infty} \int_{v_1}^{v_2} \text{Pr.} \left\{ x_1 < \frac{u_m + v - (m+n)\xi}{\sigma \sqrt{m+n}} \leq x_2 \mid u_n = v \right\} f_n(v) d\mu(v) \\ = \int_{v_1}^{v_2} \lim_{m \rightarrow \infty} \text{Pr.} \left\{ x_1 < \frac{u_m + v - (m+n)\xi}{\sigma \sqrt{m+n}} \leq x_2 \mid u_n = v \right\} f_n(v) d\mu(v) \\ = \int_{v_1}^{v_2} \left\{ \int_{x_1}^{x_2} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx \right\} p_n(v; \tau) d\mu(v) \\ = \frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} \exp\left\{-\frac{x^2}{2}\right\} dx \int_{v_1}^{v_2} p_n(v; \tau) d\mu(v),$$

as was to be proved in (5.03).

Assertion (2°) follows immediately from (1°), by letting $v_1 = -\infty$, $v_2 = +\infty$, and the combination of (1°) and (2°) give (3°).

Remark 1. (5.05) is an exact enunciation to the suggestion that $A_m(g-v)A_n(v)/A_{m+n}(g)$ will tend to $p_n(v; \tau)$ as m tends to infinity.

In virtue of (2.06) which can be written as

$$(5.07) \quad A_{m+n}(u) = \int_{-\infty}^{\infty} A_m(u-v) A_n(v) d\mu(v),$$

(4) This is an improved proof after the remark of Mr. Nobuyuki Ikeda.

we have, for the family of the probability density functions $f_m(u) \equiv \exp\{-\tau u + mb(\tau)\} A_m(u)$,

$$(5.08) \quad f_{m+n}(u) = \int_{-\infty}^{\infty} f_m(u-v) f_n(v) d\mu(v)$$

and hence

$$(5.09) \quad R_{m,n}(g, v) \equiv \frac{A_m(g-v) A_n(v)}{A_{m+n}(g)} = \frac{p_m(g-v; \tau) p_n(v; \tau)}{p_{m+n}(g; \tau)}.$$

Remark 2. In order to establish the more direct result to the effect that

$$(5.10) \quad \lim_{m \rightarrow \infty} Pr. \{v_1 < u_n \leq v_2 \mid u_m + u_n = (m + \xi) \xi + \sigma x \sqrt{m+n}\} \\ = \int_{v_1}^{v_2} p_n(v; \tau) d\mu(v),$$

it will be sufficient either to have established the central limit theorem concerning the probability density function not concerning the probability distribution or to have assumed that

$$(5.11) \quad \lim_{m \rightarrow \infty} \frac{p_{m+n}((m+n)\xi + \sigma x \sqrt{m+n}; \tau)}{p_m((m+n)\xi + \sigma x \sqrt{m+n-v}; \tau)} = 1$$

for each assigned τ and uniform for x and v in any finite domain of x and v .

The similar argument as to Theorem 5.1 can be applied to the limiting procedures in our several sample formulation. For the sake of brevity, we shall enunciate the three sample one which gives us.

Theorem 5.2. *Under the same hypothesis to Theorem 5.1, let us consider the statistic u_i , n_m and u_n . Let us write $g = u_i + u_m + u_n$. Let $x, x_1, x_2, v_1, v_2, w_1$ and w_2 , be any assigned real numbers such that $x_1 < x_2$; $v_1 < v_2$; $w_1 < w_2$. Let us put $g_i = (l + m + n)\xi + \sigma x_i \sqrt{l + m + n}$ ($i = 1, 2$). Then the following assertions hold:*

(1°) *We have*

$$(5.12) \quad \lim_{m \rightarrow \infty} Pr. \left\{ x_1 < \frac{u_l + u_m + u_n - (m + n + l)\xi}{\sigma \sqrt{l + m + n}} \leq x_2, \right. \\ \left. v_1 < u_l \leq v_2, w_1 < u_n < w_2 \right\} \\ = (\Phi(x_2) - \Phi(x_1)) \int_{v_1}^{v_2} \exp\{-\tau v + lb(\tau) + a_l(v)\} d\mu(v) \\ \cdot \int_{w_1}^{w_2} \exp\{-\tau w + nb(\tau) + a_n(w)\} d\mu(w).$$

(2°) *We have*

$$(5.13) \quad \lim_{m \rightarrow \infty} Pr. \{g_1 < g < g_2\} = \Phi(x_2) - \Phi(x_1).$$

(3°) *We have*

$$(5.14) \quad \lim_{m \rightarrow \infty} \int_{v_1}^{v_2} A_i^*(u) d\mu(u) \int_{w_1}^{w_2} A_n^*(w) d\mu(w) \frac{\int_{g_1}^{g_2} A_m^*(g-u-w) d\mu(g-u-w)}{\int_{g_1}^{g_2} A_{l+m+n}^*(g) d\mu(g)} \\ = \int_{v_1}^{v_2} p_i(u; \tau) d\mu(u) \int_{w_1}^{w_2} p_n(w; \tau) d\mu(w),$$

where we have put $A_j^*(x) = e^{-\tau x} A_j(x)$.

The proof of Theorem 5.2 is similar to that of Theorem 5.1. It is to be noted that we have the relation

$$(5.15) \quad \frac{A_i^*(u) A_m^*(g-u-w) A_n^*(w)}{A_{l+m+n}^*(g)} \\ = \frac{A_i^*(v) A_m^*(g-v-w)}{A_{l+m}^*(g-w)} \cdot \frac{A_{l+m}^*(g-w) A_n^*(w)}{A_{l+m+n}^*(g)} \\ = \frac{p_i(v; \tau) p_m(g-v-w; \tau)}{p_{l+m}(g-w; \tau)} \cdot \frac{p_{l+m}(g-w; \tau) p_n(w; \tau)}{p_{l+m+n}(g; \tau)}.$$

§ 6. The estimation after preliminary test of significance. The estimations after preliminary test of significance belong to a class of the elaborated statistical procedures discussed in some of our previous papers [2], [3] and [9]. Now we can generalise the main aspects of our procedures to the more general case associated with two additive sufficient statistics.

Let u be a sufficient statistic having the probability density function of the form (3.01) with respect to an unknown parameter τ_1 . Let v be a sufficient statistic having the probability density function of the form (3.02) with respect to another unknown parameter τ_2 . Whether the values of τ_1 and τ_2 may or may not be equal to each other is not unknown to us. Thus the distinction between two parameter values may be regarded as hypothetical. Let us suppose that we may pool u and v , that is to form $g = u + v$, in case when testing the hypothesis that $\tau_1 = \tau_2$ shows that the hypothesis cannot be rejected.

Our rule of procedure is as follows:

(i) Let the statistic $h = \varphi_{m,n}(u, v)$ is defined according to the formula (2.14) and (2.15).

(ii) Let us define the estimate $\hat{\tau}_1$ of the unknown parameter τ_1 defined in the following manner:

(a) If the statistic $h = \varphi_{m,n}(u, v)$ defined in (2.15) does not belong to the domain $R_{m,n}(\alpha)$, then

$$(6.01) \quad \hat{\tau}_1 = a'_{m+n}(u+v)$$

(b) If the statistic $h = \varphi_{m,n}(u, v)$ defined in (2.15) does belong to the domain then

$$(6.02) \quad \hat{\tau}_1 = a'_m(u),$$

where $R_{m,n}(\alpha)$ is defined by (4.02).

For each assigned value of real number x , let us denote by $D_m(x)$ the domain defined by

$$(6.03) \quad D_m^w(x) = Pr.\{w; a'_m(w) < x\}.$$

Now we observe

Theorem 6.1. *The distribution of $\hat{\tau}_1$ is given by*

$$(6.04) \quad \begin{aligned} & Pr.\{\hat{\tau}_1 < x\} \\ &= \int_{D_{m+n}^g(x)} \exp\{-\tau_1 g + (m+n)b(\tau_1) + a_{m+n}(g)\} d\mu(g) \\ & \cdot \int_{h \in \Omega - R_{m,n}(\alpha)} \exp\{-(\tau_2 - \tau_1)v(g, h) + n(b(\tau_2) - b(\tau_1))\} d_{m,n}(h) d\mu(h) \\ &+ \int_{D_m^u(x)} \exp\{-\tau_1 u + mb(\tau_1) + a_m(u)\} d\mu(u) \\ & \cdot \int_{\varphi_{m,n}(u, v) \in R_{m,n}(\alpha)} \exp\{-\tau_2 v + nb(\tau_2) + a_n(v)\} d\mu(v) \end{aligned}$$

§ 7 Sequential probability ratio test concerning an additive family of sufficient statistics. The sequential probability ratio test developed by A. WALD [14] can be applied to a fairly general class of probability functions. It is to be noted that his theory has been applied to some restrictive class of probability functions in order to secure the concrete constructions of various fundamental functions such as the operating characteristic functions, the average sample number function and so on. Most of his examples discussed in his monograph [10] are concerned with the familiar probability distributions such as normal, binomial, chi-square distributions and so on. These belong to the examples of additive family of independent sufficient statistics. In fact we shall show that there is a common feature which is valid for any additive family of independent sufficient statistics and which will make the sequential analysis to be a most fruitful one.

Let us consider an additive family of independent sufficient statistics $\{u_m\}$ concerning an unknown parameter τ . Let the null hypothesis be $H_1: \tau = \tau_1$, and let its alternative one be $H_2: \tau = \tau_2$.

Then the probability ratio is simply

$$(7.01) \quad \begin{aligned} & \frac{\exp\{-\tau_1 u_m + mb(\tau_1) + a_m(u_m)\}}{\exp\{-\tau_2 u_m + mb(\tau_2) + a_m(u_m)\}} \\ &= \exp\{-(\tau_1 - \tau_2)u_m + m(b(\tau_1) - b(\tau_2))\}. \end{aligned}$$

According to the general approach in the sequential probability ratio test by A. WALD [10], there are two constants A and B to decide one of the following three alternatives.

(1°) When we have

$$(7.02) \quad (\tau_2 - \tau_1)u_m + m(b(\tau_1) - b(\tau_1)) \geq \log A,$$

then we stop the experimentation and accept the hypothesis H_1 .

(2°) When we have

$$(7.03) \quad \log B < (\tau_2 - \tau_1)u_m + m(b(\tau_1) - b(\tau_2)) < \log A,$$

then we make another independent observation which gives us u_{m+1} .

(3°) When we have

$$(7.04) \quad (\tau_2 - \tau_1)u_m + m(b(\tau_1) - b(\tau_2)) \leq \log B,$$

then we stop the experimentation and accept the hypothesis H_2 .

The two parallel straight lines giving the boundaries for a continuation of experimentations are given by

$$(7.05) \quad u_m = \frac{b(\tau_2) - b(\tau_1)}{\tau_2 - \tau_1} m + \frac{1}{\tau_2 - \tau_1} \log r,$$

where $r = A$ or B .

In virtue of the additivity of our independent sufficient statistics we may and we shall assume that there is a sequence of independent stochastic variables $\{v_j\}$ ($j = 1, 2, \dots$) such that each v_j has the probability density function with respect to a common measure such that

$$(7.06) \quad \exp\{-\tau v_j + b(\tau) + a_1(v_j)\} d\mu(v_j)$$

and that $u_m = v_1 + v_2 + \dots + v_m$. The essential instrument for constructing the fundamental aspects as well as the practical calculations concerning sequential analysis by A. WALD [10] is to find out a function $h(\tau)$ such that

$$(7.07) \quad E_\tau \left[\left(\frac{p(v; \tau_1)}{p(v; \tau_2)} \right)^{h(\tau)} \right] = 1$$

for all τ . For our present situation, we have

Theorem 7.1. *The necessary and sufficient condition that (7.02) holds true is that $h(\cdot)$ satisfies the relation*

$$(7.08) \quad b(h(\tau)(\tau_1 - \tau_2) + \tau) = h(\tau)(b(\tau_2) - b(\tau_1)) + b(\tau)$$

for all τ .

Proof: The left hand of (7.07) can be written as

$$(7.09) \quad \int_R \exp\{-h(\tau)(\tau_1 - \tau_2) + \tau v + h(\tau)(b(\tau_2) - b(\tau_1)) + b(\tau) + a_1(v)\} d\mu(v).$$

Let us put $\tau^* = h(\tau)(\tau_1 - \tau_2) + \tau$. In order to satisfy the condition that (7.04) is equal to 1 for all τ , we should have

$$(7.10) \quad h(\tau)(b(\tau_2) - b(\tau_1)) + b(\tau) = b(\tau^*),$$

which gives the relation to be proved.

Example 7.1. For the normal distribution $N(\tau, 1)$, we have $b(\tau) = -\tau^2/2$. Then (7.08) gives us $h(\tau) = (\tau_2 + \tau_1 - 2\tau)/(\tau_1 - \tau_2)$.

Example 7.2. For the case when $b(\tau) = \log \tau$, we have $h(\tau)(\tau_1 - \tau_2) + \tau = \tau(\tau_2 \tau_1^{-1})^{h(\tau)}$.

KYUSHU UNIVERSITY

References

- [1] R. A. FISHER: Statistical methods and scientific inference, Oliver and Boyd, London and Edinburgh (1956).
- [2] T. KITAGAWA: *Successive process of statistical inferences*, (1), Mem. Fac. Sci. Kyushu Univ. Ser. A, **5** (1950), 139-180.
- [3] T. KITAGAWA: *Successive process of statistical inferences*, (2), Mem. Fac. Sci. Kyushu Univ. Ser. A, **6** (1951), 55-95.
- [4] T. KITAGAWA: *Successive process of statistical inferences*, (3), Mem. Fac. Sci. Kyushu Univ. Ser. A, **6** (1952), 131-155.
- [5] T. KITAGAWA: *Successive process of statistical inferences*, (4), Bull. Math. Stat., **5** (1952), 35-50.
- [6] T. KITAGAWA: *Successive process of statistical inferences*, (5), Mem. Fac. Sci. Kyushu Univ. Ser. A, **7** (1953), 81-106.
- [7] T. KITAGAWA: *Successive process of statistical inferences*, (6), Mem. Fac. Sci. Kyushu Univ. Ser. A, **8** (1953), 1-29.
- [8] T. KITAGAWA: *Empirical functions and interpenetrating sampling procedures*, Mem. Fac. Sci. Kyushu Univ. Ser. A, **8** (1954), 109-152.
- [9] T. KITAGAWA: *Some contributions to the design of sample surveys*, Sankhya, **17** (1956), 1-36.
- [10] B. O. KOOPMAN: *On distributions admitting a sufficient statistics*, Trans. Amer. Math. Soc., **39** (1936), 399-409.
- [11] E. L. LEHMANN and H. SCHEFFÉ: *Completeness, similar regions and unbiased estimation*. Part I, Sankhya, **10** (1950), 305-340, Part II, Sankhya, **15** (1955), 219-236.
- [12] E. J. G. PITMAN: *Sufficient statistics and intrinsic accuracy*. Proc. Camb. Phil. Soc., **32** (1936), 567-569.
- [13] C. R. RAO: *Sufficient statistics and minimum variance estimates*. Proc. Camb. Phil. Soc., **45** (1949), 213-218.
- [14] A. WALD: *Sequential analysis*. John Wiley and Sons, New York (1947).
- [15] Y. WASHIO, H. MORIMOTO and N. IKEDA: *Unbiased estimation based on sufficient statistics*, Bull. Math. Stat., **6** (1956), 69-93.