MULTIDIMENSIONAL TOLERANCE REGIONS BASED ON A LARGE SAMPLE

NOBUYUKI NAKAJIMA AND KEIITI ISII

(Received Apr. 18, 1967)

1. Introduction

Let $X=(X^1, \cdots, X^p)$ be a random vector with a probability density function $f(x;\theta)=f(x^1, \cdots, x^p; \theta_1, \cdots, \theta_k)$, where $\theta=(\theta_1, \cdots, \theta_k)$ is an unknown parameter. Let $X_1=(X_1^1, \cdots, X_1^p), \cdots, X_n=(X_n^1, \cdots, X_n^p)$ be a sample of size n.

Wald [1] considered a tolerance region of a rectangular type $[\varphi^1, \psi^1] \times \cdots \times [\varphi^p, \psi^p]$ for X (at asymptotic confidence level β when $n \to \infty$, with content γ), where φ^i and ψ^i $(i=1, \cdots, p)$ are functions of the sample. Though rectangular regions are fairly natural when $p=1$, they are not so in general. For instance an ellipsoidal one will be more natural when the population is normal.

The purpose of the present paper is to give tolerance regions of a general type at asymptotic confidence level β as $n \to \infty$, with content γ. Main theorems are stated in section 3, and the proof is given in section 4. In advance of deriving the results, we shall state the theorem of Stokes briefly in section 2 which we use as a mathematical tool. In the final section we shall give two examples; the first is the case of a rectangular type (the result, of course, coincides with Wald [1]), and the second is that of an ellipsoidal type for a normal population.

2. The theorem of Stokes

By $S^p=(a_0, a_1, \cdots, a_p)$ we denote a p-dimensional oriented simplex in R^p whose vertices are $a_0, a_1, \cdots, a_p$. That a simplex is oriented means that its vertices are ordered. The orientation of $(a_{i_0}, a_{i_1}, \cdots, a_{i_p})$ is the same as that of $S^p=(a_0, a_1, \cdots, a_p)$ (in this case we regard the simplices as identical) when the permutation τ: $(0, 1, \cdots, p) \to (i_0, i_1, \cdots, i_p)$ is even, while it is opposite to that of S^p when τ is odd. In the latter case we denote $-S^p=(a_{i_0}, a_{i_1}, \cdots, a_{i_p})$. For convenience we say that the orientation of S^p is positive or negative according as the determinant

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ a_0^1 & a_1^1 & \cdots & a_p^1 \\ \vdots & \vdots & & \vdots \\ a_0^p & a_1^p & \cdots & a_p^p \end{vmatrix}$$

is positive or negative, where a_i^j is the jth coordinate of a_i.

Let π be a differentiable mapping from some open set in R^p containing S^p (as a set) into R^p, and let $\sigma^p=(S^p, \pi)$ be the image of S^p by π, which is called an oriented singular simplex. The formal sum of a finite number of singular simplices with integer coefficients $c=\sum k_j\sigma_j^p$ is called a p-singular chain, where $\sigma_j^p=(S_j^p, \pi_j)$ and $(-1)\sigma_j^p=(-S_j^p, \pi_j)$. An oriented singular simplex itself is also a p-singular chain.

The $(p-1)$-dimensional oriented simplex $S_i^{p-1}=(-1)^i(a_0, a_1, \overset{i}{\vee}\cdots, a_p)$ (we mean by $\overset{i}{\vee}$ that a_i is omitted, and we shall use the similar symbol throughout the paper) is called a $(p-1)$-dimensional oriented face of $S^p=(a_0, a_1, \cdots, a_p)$ whose orientation is determined in accordance with that of S^p. The boundary operator ∂ is defined by, for a p-singular chain $\sigma^p=(S^p, \pi)$, $\partial\sigma^p=\sum_{i=0}^{p}(S_i^{p-1}, \pi)$, and for $c=\sum k_j\sigma_j^p$, $\partial c=\sum k_j(\partial\sigma_j^p)$. We call ∂c the boundary of c.

In R^p we consider a signed volume element $dx^{i_1}\wedge\cdots\wedge dx^{i_q}$ of degree q. This product of $dx^{i_1}, \cdots, dx^{i_q}$ with respect to $\wedge$ is assumed to be q-linear and antisymmetric. For $a_{i_1,\dots,i_q}(x)$ which are functions of $x=(x^1, \cdots, x^p)$ we define a q-differential form ω by

$$\sum_{1\leq i_1<\cdots<i_q\leq p} a_{i_1,\dots,i_q}(x)\, dx^{i_1}\wedge\cdots\wedge dx^{i_q}\ ,$$

(a precise definition of a differential form is given, for example, in [3], [4]). We define the exterior derivative $d\omega$ of ω by

$$\sum_{1\leq i_1<\cdots<i_q\leq p} da_{i_1,\dots,i_q}(x)\wedge dx^{i_1}\wedge\cdots\wedge dx^{i_q}\ ,$$

where $da(x)=\sum(\partial a(x)/\partial x^j)\,dx^j$. Hence $d\omega$ is a $(q+1)$-differential form. In particular, if $q=p-1$, and if $\beta=\sum a_i(x)\,dx^1\wedge\overset{i}{\vee}\cdots\wedge dx^p$, then $d\beta=\{\sum(-1)^{i+1}(\partial a_i(x)/\partial x^i)\}\,dx^1\wedge\cdots\wedge dx^p$.

For the q-differential form ω defined on R^p whose exterior derivative is zero, there exists a $(q-1)$-differential form χ on R^p such that $d\chi=\omega$ (Poincaré's lemma).

Let $\omega=\sum a_{i_1,\dots,i_q}(x)\,dx^{i_1}\wedge\cdots\wedge dx^{i_q}$ be a q-differential form, and π be a differentiable mapping from some open set $U\subset R^q$ into R^p. For $u=(u^1, \cdots, u^q)\in U$, and $x(u)=\pi(u)=(x^1(u), \cdots, x^p(u))$, we define $\pi^*\omega$ by

$$\sum a_{i_1,\dots,i_q}(x(u))\left\{\sum_{j_1=1}^{q}\frac{\partial x^{i_1}}{\partial u^{j_1}}du^{j_1}\right\}\wedge\cdots\wedge\left\{\sum_{j_q=1}^{q}\frac{\partial x^{i_q}}{\partial u^{j_q}}du^{j_q}\right\}. \tag{2.1}$$

Then $\pi^*\omega$ is a q-differential form on U, and therefore can be written as $b(u)\,du^1\wedge\cdots\wedge du^q$ with some function $b(u)$ of $u=(u^1,\cdots,u^q)$. Now let U be an open set containing a positively oriented q-simplex S^q (as a set) in R^q, and $\sigma^q=(S^q,\pi)$. We define an integral of ω on σ^q by

$$\int_{\sigma^q}\omega=\int_{S^q}\pi^*\omega=\int_{S^q}b(u)\,du^1\cdots du^q\,, \tag{2.2}$$

where the integral of the last term is the usual one. Moreover, for $c=\sum k_i\sigma_i^q$, we define

$$\int_c\omega=\sum k_i\int_{\sigma_i^q}\omega\,. \tag{2.3}$$

THEOREM OF STOKES. *Let c be a p-singular chain and let β be a $(p-1)$-differential form. Then*

$$\int_c d\beta=\int_{\partial c}\beta\,.$$

The proof of this theorem is found in, for example, [5].

3. Theorems

Let $f(x;\theta)=f(x^1,\cdots,x^p;\theta_1,\cdots,\theta_k)$ be the probability density function of $X=(X^1,\cdots,X^p)$, where the domain Θ of the unknown parameter θ is an open set in R^k. Let $\hat\theta=(\hat\theta_1,\cdots,\hat\theta_k)$ be an estimator, with range in Θ, of $\theta=(\theta_1,\cdots,\theta_k)$ based on a sample $(X_1,\cdots,X_n)$.

Suppose that a region $c(\theta,\xi)$ is defined for every (θ,ξ), $\theta\in\Theta$, $0<\xi<1$, and satisfies

$$\int_{c(\theta,\xi)}f(x;\theta)\,dx=\xi\,, \tag{3.1}$$

where dx stands for $dx^1\cdots dx^p$. Let

$$I(\hat\theta,\theta,\xi)=\int_{c(\hat\theta,\xi)}f(x;\theta)\,dx\,. \tag{3.2}$$

Then it is clear that $I(\hat\theta,\theta,\xi)=\xi$.

We set the following regularity assumptions.

ASSUMPTION 1. The limiting joint distribution of $\sqrt{n}(\hat\theta_1-\theta_1),\cdots,\sqrt{n}(\hat\theta_k-\theta_k)$, when $n\to\infty$, is the k-variate normal distribution with zero means and non-singular covariance matrix $\|\sigma_{ij}(\theta)\|$ $(i,j=1,\cdots,k)$ for any $\theta=(\theta_1, \cdots, \theta_k)$ in some neighborhood of a point $\theta^0=(\theta_1^0, \cdots, \theta_k^0)$ in Θ, where $\sigma_{ij}(\theta)$ are continuous functions of $\theta_1, \cdots, \theta_k$ in a neighborhood of θ^0.

ASSUMPTION 2. $[\partial I(\hat{\theta}, \theta, \xi)/\partial\hat{\theta}_i]_{\hat{\theta}=\theta}$ $(i=1, \cdots, k)$ are continuous functions of (θ, ξ) in some neighborhood of (θ^0, γ).

ASSUMPTION 3. For at least one i, $[\partial I(\hat{\theta}, \theta^0, \gamma)/\partial\hat{\theta}_i]_{\hat{\theta}=\theta^0}$ is not zero.

Define $\sigma^2(\theta, \xi)$ by

$$\sigma^2(\theta, \xi)=\sum_{i=1}^{k}\sum_{j=1}^{k} \left.\frac{\partial I(\hat{\theta}, \theta, \xi)}{\partial\hat{\theta}_i}\right|_{\hat{\theta}=\theta} \left.\frac{\partial I(\hat{\theta}, \theta, \xi)}{\partial\hat{\theta}_j}\right|_{\hat{\theta}=\theta} \sigma_{ij}(\theta)\ , \tag{3.3}$$

and λ_β by

$$\int_{\lambda_\beta}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt=\beta\ . \tag{3.4}$$

Furthermore, set

$$\bar{\xi}=\bar{\xi}(\beta, \gamma, \hat{\theta})=\gamma-\lambda_\beta\frac{\sigma(\hat{\theta}, \gamma)}{\sqrt{n}}\ . \tag{3.5}$$

Consider the region $c(\hat{\theta}, \bar{\xi})$ by substituting $\hat{\theta}$ and $\bar{\xi}$ into $c(\theta, \xi)$, but let $c(\hat{\theta}, \bar{\xi})$ represent the empty set when $\bar{\xi}\leqq 0$, and the whole space R^p when $\bar{\xi}\geqq 1$. Then we obtain the following lemma which is essentially based on a result of Wald [1].

LEMMA. *If*
(i) $\hat{\theta}=(\hat{\theta}_1, \cdots, \hat{\theta}_k)$ *satisfies assumption* 1, *and if*
(ii) $I(\hat{\theta}, \theta, \xi)$ *defined by* (3.2) *satisfies assumptions* 2 *and* 3, *then the region* $c(\hat{\theta}, \bar{\xi})$ *satisfies*

$$\Pr\left\{\int_{c(\hat{\theta}, \bar{\xi})} f(x;\, \theta^0)\, dx \geqq \gamma \,\middle|\, \theta^0\right\} \longrightarrow \beta \qquad (n\to\infty)\ ,$$

that is, $c(\hat{\theta}, \bar{\xi})$ *is a tolerance region at asymptotic confidence level* β *with content* γ.

Suppose that $f(x;\theta)=f(x^1, \cdots, x^p;\, \theta_1, \cdots, \theta_k)$ is continuously differentiable with respect to $x^1, \cdots, x^p$. We consider a p-differential form $\alpha(\theta)=f(x;\theta)\, dx^1\wedge\cdots\wedge dx^p$. Then it is trivial that $d\alpha(\theta)=0$. Hence, by Poincaré's lemma, there exist p functions $a_1(x;\theta), \cdots, a_p(x;\theta)$ such that the $(p-1)$-differential form $\beta(\theta)=\sum(-1)^{j+1} a_j(x;\theta)\, dx^1\wedge\overset{j}{\overset{\vee}{\cdots}}\wedge dx^p$ satisfies $d\beta(\theta)=\alpha(\theta)$. It is clear that $\sum\{\partial a_j(x;\theta)/\partial x^j\}=f(x;\theta)$.

We assume the following on a p-singular chain $c(\theta, \xi)=\sum \sigma_i^p(\theta, \xi)$ which satisfies

$$\int_{c(\theta,\xi)} \alpha(\theta)=\xi . \tag{3.6}$$

ASSUMPTION 4. $\sigma_i^p(\theta, \xi)$ is expressed as $\sigma_i^p(\theta, \xi)=(S_i^p, \pi_i(\theta, \xi))$ with an oriented p-simplex S_i^p and a differentiable one-to-one mapping $\pi_i(\theta, \xi)$ from some open set containing S_i^p (as a set) into R^p.

ASSUMPTION 5. By $|\sigma_i^p(\theta, \xi)|$ we represent $\sigma_i^p(\theta, \xi)$ as a set in R^p. Then $|c(\theta, \xi)|=\cup|\sigma_i^p(\theta, \xi)|$ is connected. If $|\sigma_i^p(\theta, \xi)|\cap|\sigma_{i'}^p(\theta, \xi)|$ $(i\neq i')$ is not empty, it is a singular $(p-1)$-simplex (considered as a set) which is a common face of $\sigma_i^p(\theta, \xi)$ and $\sigma_{i'}^p(\theta, \xi)$, with the orientations corresponding to them being opposite to each other.

By assumption 5, we can denote the boundary $\partial c(\theta, \xi)$ of $c(\theta, \xi)$ by a $(p-1)$-singular chain $\sum_{j=1}^{s} \sigma_j^{p-1}(\theta, \xi)$, and each of a $(p-1)$-singular simplices $\sigma_1^{p-1}(\theta, \xi), \cdots, \sigma_s^{p-1}(\theta, \xi)$ is a face of exactly one p-simplex. Moreover, we can clearly denote $\sigma_j^{p-1}(\theta, \xi)=(S_j^{p-1}, \rho_j(\theta, \xi))$ with an oriented $(p-1)$-simplex S_j^{p-1} in R^{p-1}, and a differentiable one-to-one mapping $\rho_j(\theta, \xi)$ from an open set U_j in R^{p-1} containing S_j^{p-1} into R^p. We further assume the following on $\rho_j(\theta, \xi)$ (therefore on $\pi_i(\theta, \xi)$).

ASSUMPTION 6. For $u=(u^1, \cdots, u^{p-1}) \in U_j$ let $x_j(u)=\rho_j(\theta, \xi)(u)=(x_j^1(u), \cdots, x_j^p(u))$. Then, for every i, $\partial^2 x_j^i/\partial u^s \partial u^t$ and $\partial^2 x_j^i/\partial\theta_l \partial u^s$ $(s\neq t$, $l=1, \cdots, k)$ are continuous functions of $(u^1, \cdots, u^{p-1}; \theta_1, \cdots, \theta_k)$ for $u \in U_j$ and θ in some neighborhood Δ of θ^0.

Since $\pi_i(\theta, \xi)$ is a one-to-one mapping for any θ, $\partial x_j^i/\partial\theta_l$ can be considered as a function of $x=(x^1, \cdots, x^p)$ for any θ, instead of a function of $u=(u^1, \cdots, u^{p-1})$.

ASSUMPTION 7. $\partial x_j^i/\partial\theta_l$ is a continuous function of $(x^1, \cdots, x^p; \theta_1, \cdots, \theta_k)$ on $\sigma_j^{p-1}(\theta, \xi)$, and its value is independent of j at points x common to some j's.

Then we obtain the following theorem.

THEOREM 1. *If the p-singular chain* $c(\theta, \xi)$ *satisfies* (3.6) *and assumptions* 4–7, *then* $I(\hat\theta, \theta, \xi)$ *defined by*

$$I(\hat\theta, \theta, \xi)=\int_{c(\hat\theta,\xi)} \alpha(\theta) \tag{3.7}$$

is continuously differentiable with respect to $\hat\theta_1, \cdots, \hat\theta_k$ *for* $\hat\theta$ *in* Δ, *and*

$$(3.8)\quad \frac{\partial I(\hat{\theta}, \theta, \xi)}{\partial \hat{\theta}_l} = \sum_{j=1}^{s} \int_{S_j^{p-1}} f(x_j(u; \hat{\theta}); \theta) \frac{\partial(x_j^1, \cdots, x_j^p)}{\partial(\hat{\theta}_l, u^1, \cdots, u^{p-1})} du^1 \wedge \cdots \wedge du^{p-1} .$$

The proof will be given in the next section.

COROLLARY. *Under the same conditions as in theorem* 1, *we have, for* θ *in* Δ,

$$(3.9)\quad \left.\frac{\partial I(\hat{\theta}, \theta, \xi)}{\partial \hat{\theta}_l}\right|_{\hat{\theta}=\theta} = \int_{\partial c(\theta, \xi)} f(x; \theta) \sum_{i=1}^{p} (-1)^{i+1} \frac{\partial x^i}{\partial \theta_l} dx^1 \wedge \overset{i}{\overset{\vee}{\cdots}} \wedge dx^p .$$

The lemma and theorem 1 imply the following.

THEOREM 2. *Suppose that the estimator* $\hat{\theta} = (\hat{\theta}_1, \cdots, \hat{\theta}_k)$ *satisfies assumption* 1, $c(\theta, \xi)$ *satisfies* (3.6) *and assumptions* 4–7, *and that* $\partial I(\hat{\theta}, \theta, \xi)/\partial \hat{\theta}_l$ *given by* (3.8) *satisfies assumptions* 2 *and* 3. *Then the* p-*singular chain* $c(\hat{\theta}, \bar{\xi})$, *which is constructed from the sample by using* $\hat{\theta}$ *and* $\bar{\xi}$ *in* (3.5), *gives a* γ-*content tolerance region at asymptotic level* β *as* $n \to \infty$, *when* θ^0 *is the true value of the parameter.*

4. Proof of theorem 1

Since $\alpha(\theta) = d\beta(\theta)$, we can apply the theorem of Stokes to (3.7) and obtain

$$(4.1)\quad I(\hat{\theta}, \theta, \xi) = \int_{c(\hat{\theta}, \xi)} d\beta(\theta) = \int_{\partial c(\hat{\theta}, \xi)} \beta(\theta) = \sum_{j=1}^{s} \int_{\sigma_j^{p-1}(\hat{\theta}, \xi)} \beta(\theta) .$$

Put

$$(4.2)\quad A_j(u; \hat{\theta}, \theta, \xi) = \begin{vmatrix} a_1(x_j(u; \hat{\theta}); \theta) \dfrac{\partial x_j^1(u; \hat{\theta})}{\partial u^1} & \cdots & \dfrac{\partial x_j^1(u; \hat{\theta})}{\partial u^{p-1}} \\ \vdots & \vdots & \vdots \\ a_p(x_j(u; \hat{\theta}); \theta) \dfrac{\partial x_j^p(u; \hat{\theta})}{\partial u^1} & \cdots & \dfrac{\partial x_j^p(u; \hat{\theta})}{\partial u^{p-1}} \end{vmatrix} .$$

Then, by means of (2.1) and (2.2), (4.1) is rewritten as

$$(4.3)\quad \sum_{j=1}^{s} \int_{S_j^{p-1}} (\rho_j(\hat{\theta}, \xi))^* \beta(\theta)$$

$$= \sum_{j=1}^{s} \int_{S_j^{p-1}} \sum_{i=1}^{p} a_i(x_j(u; \hat{\theta}); \theta) \left\{ \sum_{j_1=1}^{p-1} \frac{\partial x_j^1}{\partial u^{j_1}} du^{j_1} \right\} \wedge \overset{i}{\overset{\vee}{\cdots}} \wedge \left\{ \sum_{j_p=1}^{p-1} \frac{\partial x_j^p}{\partial u^{j_p}} du^{j_p} \right\}$$

$$= \sum_{j=1}^{s} \int_{S_j^{p-1}} A_j(u; \hat{\theta}, \theta, \xi) du^1 \wedge \cdots \wedge du^{p-1} .$$

From assumption 6, it is clear that $\partial A(u;\hat{\theta},\theta,\xi)/\partial\hat{\theta}_l$ is continuous. Hence $I(\hat{\theta},\theta,\xi)$ is continuously differentiable with respect to $\hat{\theta}_1,\cdots,\hat{\theta}_k$, and

$$\frac{\partial I(\hat{\theta},\theta,\xi)}{\partial\hat{\theta}_l}=\sum_{j=1}^{s}\int_{S_j^{p-1}}\frac{\partial A_j(u;\hat{\theta},\theta,\xi)}{\partial\hat{\theta}_l}\,du^1\wedge\cdots\wedge du^{p-1}\,. \tag{4.4}$$

Suppose in general, that $x^1,\cdots,x^p$ are differentiable functions of $t=(t^1,\cdots,t^p)$, and that $a_1(x),\cdots,a_p(x)$ are differentiable functions of $x=(x^1,\cdots,x^p)$. Let $J=\partial(x^1,\cdots,x^p)/\partial(t^1,\cdots,t^p)$ and let $D_{ik}(t)$ be the (i,k)-cofactor of J. Then

$$\sum_{k=1}^{p}\frac{\partial}{\partial t^k}\left\{\sum_{i=1}^{p}a_i(x(t))D_{ik}(t)\right\}=\sum_{k=1}^{p}\sum_{i=1}^{p}\sum_{l=1}^{p}\frac{\partial a_i}{\partial x^l}\frac{\partial x^l}{\partial t^k}D_{ik}+\sum_{i=1}^{p}a_i\sum_{k=1}^{p}\frac{\partial D_{ik}}{\partial t^k}\,. \tag{4.5}$$

The first term of the right-hand side becomes

$$\sum_{l=1}^{p}\sum_{i=1}^{p}\frac{\partial a_i}{\partial x^l}\sum_{k=1}^{p}\frac{\partial x^l}{\partial t^k}D_{ik}=\sum_{i=1}^{p}\frac{\partial a_i}{\partial x^i}\cdot J\,,$$

because $\sum_{k=1}^{p}(\partial x^l/\partial t^k)D_{ik}=\delta_{il}\cdot J$, where δ_{il} is the Kronecker symbol. It is easily seen that the second term vanishes if $\partial^2x^i/\partial t^k\partial t^l=\partial^2x^i/\partial t^l\partial t^k$ $(i,k,l=1,\cdots,p)$. Therefore (4.5) becomes

$$\sum_{k=1}^{p}\frac{\partial}{\partial t^k}\left\{\sum_{i=1}^{p}a_i(x(t))D_{ik}(t)\right\}=\sum_{i=1}^{p}\frac{\partial a_i}{\partial x^i}(t)\cdot J\,. \tag{4.6}$$

Apply (4.6) to our problem with $\hat{\theta}_l=t^1$, $u^1=t^2,\cdots,u^{p-1}=t^p$. Since $\sum a_i(x(u;\hat{\theta});\theta)D_{i1}=A(u;\hat{\theta},\theta,\xi)$, we obtain

$$\frac{\partial A(u;\hat{\theta},\theta,\xi)}{\partial\hat{\theta}_l}+\sum_{k=1}^{p-1}\frac{\partial}{\partial u^k}\left\{\sum_{i=1}^{p}a_iD_{i,k+1}\right\}=f(x(u;\hat{\theta});\theta)\cdot J\,. \tag{4.7}$$

Hence, (4.4) becomes

$$\begin{aligned}\frac{\partial I(\hat{\theta},\theta,\xi)}{\partial\hat{\theta}_l}=&\sum_{j=1}^{s}\int_{S_j^{p-1}}f(x_j(u;\hat{\theta});\theta)\frac{\partial(x_j^1,\cdots,x_j^p)}{\partial(\hat{\theta}_l,u^1,\cdots,u^{p-1})}du^1\wedge\cdots\wedge du^{p-1}\\&-\sum_{j=1}^{s}\int_{S_j^{p-1}}\sum_{k=1}^{p-1}\frac{\partial}{\partial u^k}\left\{\sum_{i=1}^{p}a_iD_{i,k+1}\right\}du^1\wedge\cdots\wedge du^{p-1}\,.\end{aligned} \tag{4.8}$$

The second term of the right-hand side, by applying the theorem of Stokes again, proves to be

$$\sum_{j=1}^{s}\sum_{t=0}^{p-1}\int_{S_{j,t}^{p-2}}\sum_{k=1}^{p-1}\left\{\sum_{i=1}^{p}a_iD_{i,k+1}\right\}du^1\wedge\overset{\overset{k}{\vee}}{\cdots}\wedge du^{p-1}\,, \tag{4.9}$$

where $S_{j,0}^{p-2}, \cdots, S_{j,p-1}^{p-2}$ are faces of S_j^{p-1} (the orientation of $S_{j,t}^{p-2}$ corresponds to that of S_j^{p-1}). By assumptions 5 and 7 we can find a continuous $(p-2)$-differential form $\omega(\hat{\theta}, \theta)$ which satisfies

$$(4.10) \qquad (\rho_j(\hat{\theta}, \xi))^* \omega(\hat{\theta}, \theta) = \sum_{k=1}^{p-1} \left\{ \sum_{i=1}^{p} a_i D_{i,k+1} \right\} du^1 \wedge \overset{k}{\overset{\vee}{\cdots}} \wedge du^{p-1}$$

and express $\omega(\hat{\theta}, \theta)$ as

$$\sum_{m=1}^{p} \sum_{k \neq m} \varepsilon_{m,k} a_m(x; \theta) \frac{\partial x^k}{\partial \hat{\theta}_l} dx^1 \wedge \overset{m}{\vee} \cdots \overset{k}{\vee} \wedge dx^p$$

with $\varepsilon_{m,k}=1$ or -1. The integral in (4.9) is expressed as

$$\int_{\sigma_{j,t}^{p-2}(\hat{\theta}, \xi)} \omega(\hat{\theta}, \theta) ,$$

and hence (4.9) is equal to

$$(4.11) \qquad \sum_{j=1}^{s} \sum_{t=0}^{p-1} \int_{\sigma_{j,t}^{p-2}(\hat{\theta}, \xi)} \omega(\hat{\theta}, \theta) = \int_{\partial(\partial c(\hat{\theta}, \xi))} \omega(\hat{\theta}, \theta) .$$

Since $\partial c(\hat{\theta}, \xi)$ is a closed surface, it is clear that $\partial(\partial c(\hat{\theta}, \xi))=0$. Therefore (4.11), hence the second term of the right-hand side of (4.8) vanishes. Thus the proof of theorem 1 is completed.

5. Examples

I. The case of rectangular regions of Wald

In this case the region determined by $c(\hat{\theta}, \xi)$ is $[\varphi^1, \psi^1] \times \cdots \times [\varphi^p, \psi^p]$ where $\varphi^i = \varphi^i(\hat{\theta}, \xi)$ and $\psi^i = \psi^i(\hat{\theta}, \xi)$ $(i=1, \cdots, p)$. Clearly we can partition $c(\hat{\theta}, \xi)$ into a finite number of simplices, and express $c(\hat{\theta}, \xi) = \sum_{i=1}^{r} \sigma_i^p(\hat{\theta}, \xi)$, where each $\sigma_i^p(\hat{\theta}, \xi)$ is a positively oriented (not singular) p-simplex. We write $\partial c(\hat{\theta}, \xi) = \sum_{j=1}^{s} \sigma_j^{p-1}(\hat{\theta}, \xi)$. Then it is clear that $\sigma_j^{p-1}(\hat{\theta}, \xi)$ is positively oriented. On $\sigma_j^{p-1}(\hat{\theta}, \xi)$, $x^q = \varphi^q$ or $= \psi^q$ for some q $(1 \leqq q \leqq p)$. We denote by $\partial c(\hat{\theta}, \xi) \cap \{x^q = \varphi^q\}$ the sum $\sum \sigma_j^{p-1}(\hat{\theta}, \xi)$ of all $\sigma_j^{p-1}(\hat{\theta}, \xi)$'s on which $x^q = \varphi^q$. The symbol $\partial c(\hat{\theta}, \xi) \cap \{x^q = \psi^q\}$ is defined similarly.

It is easily seen that $dx^1 \wedge \overset{t}{\overset{\vee}{\cdots}} \wedge dx^p = \delta_{qt}(-1)^t dx^1 \overset{t}{\overset{\vee}{\cdots}} dx^p$ on $\partial c(\hat{\theta}, \xi) \cap \{x^q = \varphi^q\}$, and $dx^1 \wedge \overset{t}{\overset{\vee}{\cdots}} \wedge dx^p = \delta_{qt}(-1)^{t+1} dx^1 \overset{t}{\overset{\vee}{\cdots}} dx^p$ on $\partial c(\hat{\theta}, \xi) \cap \{x^q = \psi^q\}$, and in each case $\partial x^q / \partial \hat{\theta}_l$ is constant (equals $\partial \varphi^q / \partial \hat{\theta}_l$ or $\partial \psi^q / \partial \hat{\theta}_l$). Hence, using the Wald's notation

$$I_q(\hat{\theta}, \theta, \xi, x^q) = \int_{\varphi^p}^{\phi^p} \cdots \int_{\varphi^{q+1}}^{\phi^{q+1}} \int_{\varphi^{q-1}}^{\phi^{q-1}} \cdots \int_{\varphi^1}^{\phi^1} f(x; \theta)\, dx^1 \cdots dx^{q-1} dx^{q+1} \cdots dx^p ,$$

we have

$$\int_{\partial c(\hat{\theta}, \xi) \cap \{x^q = \varphi^q\}} f(x; \theta) \sum_{t=1}^{p} (-1)^{t+1} \frac{\partial x^t}{\partial \hat{\theta}_l}\, dx^1 \wedge \overset{t}{\overset{\vee}{\cdots}} \wedge dx^p$$

$$= \frac{\partial \varphi^q}{\partial \hat{\theta}_l} \int_{\partial c(\hat{\theta}, \xi) \cap \{x^q = \varphi^q\}} (-1) f(x; \theta)\, dx^1 \overset{q}{\overset{\vee}{\cdots}} dx^p = -\frac{\partial \varphi^q}{\partial \hat{\theta}_l} I_q(\hat{\theta}, \theta, \xi, \varphi^q) .$$

Similarly, we have

$$\int_{\partial c(\hat{\theta}, \xi) \cap \{x^q = \phi^q\}} f(x; \theta) \sum_{t=1}^{p} (-1)^{t+1} \frac{\partial x^t}{\partial \hat{\theta}_l}\, dx^1 \wedge \overset{t}{\overset{\vee}{\cdots}} \wedge dx^p = \frac{\partial \phi^q}{\partial \hat{\theta}_l} I_q(\hat{\theta}, \theta, \xi, \phi^q) .$$

Hence (3.9) becomes

$$\left. \frac{\partial I(\hat{\theta}, \theta, \xi)}{\partial \hat{\theta}_l} \right|_{\hat{\theta} = \theta} = \sum_{q=1}^{p} \frac{\partial \phi^q}{\partial \theta_l} I_q(\theta, \theta, \xi, \phi^q) - \sum_{q=1}^{p} \frac{\partial \varphi^q}{\partial \theta_l} I_q(\theta, \theta, \xi, \varphi^q) ,$$

which coincides with Wald's results.

II. The case of ellipsoidal regions for a normal population

Let $\boldsymbol{X} = (X^1, \cdots, X^p)$ have a nondegenerate normal density function with an unknown parameter $\theta = (\mu_1, \cdots, \mu_p; \sigma_{ij}, 1 \leqq i \leqq j \leqq p)$:

$$f(\boldsymbol{x}; \theta) = (2\pi)^{-p/2} |\Sigma|^{-1/2} \exp\left[-\frac{1}{2} (\boldsymbol{x} - \boldsymbol{\mu}) \Sigma^{-1} (\boldsymbol{x} - \boldsymbol{\mu})' \right]$$

where $\boldsymbol{\mu} = (\mu_1, \cdots, \mu_p)$ and $\Sigma = \|\sigma_{ij}\|_{i,j=1,\dots,p}$, $\sigma_{ji} = \sigma_{ij}$.

Siotani [7] and John [8] treated tolerance regions for a normal population independently. Each of their methods is based on a sampling distribution of sufficient statistics.

Now we consider the region $c(\theta, \xi)$ with the smallest volume under the condition

$$\int_{c(\theta, \xi)} f(\boldsymbol{x}; \theta)\, d\boldsymbol{x} = \xi . \tag{5.1}$$

Such $c(\theta, \xi)$ is clearly expressed as

$$c(\theta, \xi) = \{\boldsymbol{x} : (\boldsymbol{x} - \boldsymbol{\mu}) \Sigma^{-1} (\boldsymbol{x} - \boldsymbol{\mu})' \leqq k\} \tag{5.2}$$

where k is a constant determined by θ and ξ.

Since Σ^{-1} is symmetric and positive definite, there exists a non-singular matrix A such that $\Sigma^{-1}=AA'$. The condition $(\boldsymbol{X}-\boldsymbol{\mu})\Sigma^{-1}(\boldsymbol{X}-\boldsymbol{\mu})' \leqq k$ is equivalent to $\boldsymbol{ZZ}'\leqq k$ with $\boldsymbol{Z}=(\boldsymbol{X}-\boldsymbol{\mu})A$. Since the distribution $\boldsymbol{ZZ}'$ has the chi-square distribution with p degrees of freedom, the value of k in (5.2) is determined by $\Pr(\chi_p^2 \leqq k)=\xi$. We denote this k by $\chi_p^2(1-\xi)$.

Consider the transformation $\boldsymbol{x}=\boldsymbol{\mu}+\boldsymbol{z}A^{-1}$. When $\boldsymbol{z}$ moves in the sphere $S(\chi_p(1-\xi))$ with radius $\chi_p(1-\xi)$, $\boldsymbol{x}$ moves in $c(\theta, \xi)$, and $\boldsymbol{z}$ on $\partial S(\chi_p(1-\xi))$ corresponds to $\boldsymbol{x}$ on $\partial c(\theta, \xi)$. Now, take A so that A is continuously differentiable with respect to every σ_{ij} (for instance, take A as a triangular matrix with positive diagonal elements). Then, since $S(\chi_p(1-\xi))$ is obviously partitioned into a finite number of p-singular simplices, $c(\theta, \xi)$ satisfies assumptions 4–7 in section 3, with the above transformation.

Hence, since $f(\boldsymbol{x};\theta)=(2\pi)^{-p/2}|\Sigma|^{-1/2}\exp(-\chi_p^2(1-\xi)/2)$ on $\partial c(\theta, \xi)$, (3.9) becomes

$$\left.\frac{\partial I(\hat{\theta}, \theta, \xi)}{\partial \hat{\theta}_l}\right|_{\hat{\theta}=\theta} = (2\pi)^{-p/2} e^{-\chi_p^2(1-\xi)/2} \times \int_{\partial c(\theta,\xi)} \sum_{q=1}^{p} (-1)^{q+1} \frac{\partial x^q}{\partial \theta_l} |\Sigma|^{-1/2} dx^1 \wedge \overset{q}{\check{\cdots}} \wedge dx^p . \tag{5.3}$$

By the theorem of Stokes, the integral in the right-hand side is equal to

$$\int_{c(\theta,\xi)} \sum_{q=1}^{p} \frac{\partial}{\partial x^q} \frac{\partial x^q}{\partial \theta_l} |\Sigma|^{-1/2} dx^1 \wedge \cdots \wedge dx^p , \tag{5.4}$$

where $(\partial/\partial x^q)(\partial x^q/\partial\theta_l)$ should be interpreted as differentiating $\partial x^q(\boldsymbol{z};\theta)/\partial\theta_l$ (with respect to x^q), considering it as the function of $\boldsymbol{x}$ and θ by substituting $\boldsymbol{z}=(\boldsymbol{x}-\boldsymbol{\mu})A$. Since $(\partial/\partial x^q)(\partial x^q/\partial\mu_j)=0$ $(q, j=1, \cdots, p)$, it follows from (5.3) and (5.4) that

$$\left.\frac{\partial I(\hat{\theta}, \theta, \xi)}{\partial \hat{\mu}_j}\right|_{\hat{\theta}=\theta} = 0 \qquad (j=1, \cdots, p) .$$

On the other hand, since $\partial\boldsymbol{x}/\partial\sigma_{ij}=\boldsymbol{z}(\partial A^{-1}/\partial\sigma_{ij})=(\boldsymbol{x}-\boldsymbol{\mu})A(\partial A^{-1}/\partial\sigma_{ij})$,

$$\sum_{q=1}^{p} \frac{\partial}{\partial x^q} \frac{\partial x^q}{\partial \sigma_{ij}} = \operatorname{tr}\left(A\frac{\partial A^{-1}}{\partial \sigma_{ij}}\right) = \frac{1}{2}\operatorname{tr}\left(\Sigma^{-1}\frac{\partial \Sigma}{\partial \sigma_{ij}}\right)$$
$$= \begin{cases} \frac{1}{2}\sigma^{ii} & (i=j) \\ \sigma^{ij} & (i<j) , \end{cases}$$

where σ^{ij} is the (i, j)-element of Σ^{-1}. Hence (5.4) turns out to be

$$\left(1-\frac{1}{2}\delta_{ij}\right)\sigma^{ij}\int_{c(\theta,\xi)}|\Sigma|^{-1/2}\,dx^1\wedge\cdots\wedge dx^p\;.$$

Because of the fact

$$\int_{c(\theta,\xi)}|\Sigma|^{-1/2}\,dx^1\wedge\cdots\wedge dx^p=\int_{S(\chi_p(1-\xi))}dz^1\wedge\cdots\wedge dz^p=\frac{\pi^{p/2}\chi_p^p(1-\xi)}{\Gamma(p/2+1)}\;,$$

(5.3) proves to be

$$\text{(5.5)}\qquad \left.\frac{\partial I(\hat{\theta},\theta,\xi)}{\partial\hat{\sigma}_{ij}}\right|_{\hat{\theta}=\theta}=\left(1-\frac{1}{2}\delta_{ij}\right)2^{-p/2}e^{-\chi_p^2(1-\xi)/2}\frac{\chi_p^p(1-\xi)\sigma^{ij}}{\Gamma(p/2+1)}\;.$$

When we take the maximum likelihood estimate $(\hat{\mu},\hat{\Sigma})$ of (μ,Σ), $n\hat{\Sigma}$ is distributed as the Wishart distribution $W(\Sigma,n-1)$, and therefore the limiting distribution of $\sqrt{n}(\hat{\Sigma}-\Sigma)$ is the normal with zero means and the covariance matrix having $\sigma_{ik}\sigma_{jl}+\sigma_{il}\sigma_{jk}$ as the (ij,kl)-element (see, for instance, [6]). Hence assumptions 1-3 in section 3 are satisfied and (3.3) is explicitly expressed as

$$\begin{aligned}\sigma^2(\theta,\xi)&=\sum_{i\le j,k\le l}\left.\frac{\partial I}{\partial\hat{\sigma}_{ij}}\right|_{\hat{\theta}=\theta}\left.\frac{\partial I}{\partial\hat{\sigma}_{kl}}\right|_{\hat{\theta}=\theta}(\sigma_{ik}\sigma_{jl}+\sigma_{il}\sigma_{jk})\\&=\frac{e^{-\chi_p^2(1-\xi)}\{\chi_p^2(1-\xi)\}^p}{2^{p-1}\{\Gamma(p/2)\}^2}\;.\end{aligned}$$

When $p=1$ it coincides with Wald's result.

UNIVERSITY OF OSAKA PREFECTURE, OSAKA UNIVERSITY

REFERENCES

[1] A. Wald, "Setting of tolerance limits when the sample is large," *Ann. Math. Statist.*, 13 (1942), 389-399.
[2] A. Wald, "Tests of statistical hypotheses concerning several parameters when the number of observations is large," *Trans. Amer. Math. Soc.*, 54 (1943), 426-482.
[3] Y. Matsushima, *Tayotai Nyumon* (in Japanese), Shokabo, 1965.
[4] S. Chern, *Differentiable Manifolds*, Chicago Univ., 1959.
[5] G. de Rham, *Variétés Differentiables*, Nancago Univ., 1960.
[6] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, John Wiley, 1958.
[7] M. Siotani, "Tolerance regions for a multivariate normal populations," *Ann. Inst. Statist. Math.*, 16 (1964), 135-153.
[8] S. John, "A tolerance region multivariate for normal distributions," *Sankhyā*, 23 (1961), 363-368.