ON THE ASYMPTOTIC DISTRIBUTION

OF THE

LIKELIHOOD RATIO IN SOME PROBLEMS

ON

MIXED VARIATE POPULATIONS

by

J. Ogawa, M. D. Moustafa and S. N. Roy

This research was sponsored partly by the Office of Naval Research under Contract No. Nonr-855(06) for research in probability and statistics at Chapel Hill and partly by the United States Air Force through the Air Force Office of Scientific Research of the Air Research and Development Command, under Contract No. 18(600)-83.

Institute of Statistics
Mimeograph Series No. 180
August, 1957

ON THE ASYMPTOTIC DISTRIBUTION
OF THE
LIKELIHOOD RATIO IN SOME PROBLEMS
ON
MIXED VARIATE POPULATIONS*

by

J. Ogawa, M. D. Moustafa and S. N. Roy
Institute of Statistics, University of North Carolina
Chapel Hill, N. C.

Summary and Introduction

Let the likelihood function of the population under consideration be $P(X \mid H_0)$ and $P(X \mid H)$ under the hull-hypothesis H_0 and the alternative H respectively, and put

$$\lambda = \max P(X \mid H_0) \; / \; \max P(X \mid H) \quad .$$

Then it is well-known that, under certain conditions, the random variable $-2 \log_e \lambda$ has the X^2-distribution with suitable degrees of freedom in the limit as the sample size tends to infinity, provided the null hypothesis H_0 is true.

S. S. Wilks [6] stated this result early in 1938, and gave a sketch of a proof based on J. L. Doob's work [9]. Later, in 1943,

* This research was sponsored partly by the Office of Naval Research under Contract No. Nonr-855(06) for research in probability and statistics at Chapel Hill and partly by the United States Air Force through the Air Force Office of Scientific Research of the Air Research and Development Command, under Contract No. 18(600)-83. Reproduction in whole or in part for any purpose of the United States Government is permitted.

A. Wald [5] obtained the same result starting from somewhat stronger assumptions. However, as far as the present authors are concerned, they have not so far seen any complete proof of this proposition along Wilks' line, or, in other words, a complete proof based upon Doob's paper. Thus it seems worth while to give a completely rigorous proof of this proposition along the lines stated. In this note, the authors are mainly concerned with the asymptotic distribution of the statistic $-2 \log_e \lambda$ defined, as before, for testing null-hypotheses of certain types on a mixed variates population which was discussed by one of the authors in another paper [3]. Toward that end, in section 1, the authors present a rigorous proof of the fact that $-2 \log \lambda$ on Doob's assumptions has an asymptotic X^2-distribution. In section 2 they establish a theorem which guarantees the consistency and the uniqueness of the solution of the maximum likelihood equation and some discussion on the consistency of the maximum likelihood estimate. In section 3 the authors explain the problem of testing hypotheses in the 3-variates (X,Y,Z) population, where X,Y are continuous and Z is categorical. Then in section 4, the validity of Doob's conditions and the assumption of Theorem 2.1 will be verified in the case of the population which was explained in section 3.

1. Doob's Theorem and Wilks' Theorem. We shall start with the definition of the n^{th} approximation of the maximum likelihood estimate of a parameter or simply the maximum likelihood estimate of a parameter calculated from a random sample of size n. For the sake of simplicity of explanation, we shall begin with the single parameter case. For each value of θ in a non-degenerate point set

Ω, which is called the parameter space, let $f(x,\theta)$ be a probability density on the infinite interval $-\infty < x < \infty$. Let x be a chance variable whose distribution is determined by the probability density $f(x,\theta_o)$, which is called the true density. Without any loss of generality, we can assume that $\theta_o = 0$, and put

$$f(x) \equiv f(x,0) .$$

Suppose that for each set of numbers $x_1, x_2, \ldots, x_n$, $n = 1,2,\ldots,$ it is possible to find a value of θ in Ω, i.e., a function of $x_1, x_2, \ldots, x_n$, such that

$$(1.1) \qquad \prod_{j=1}^{n} f(x_j,\theta) \geq \prod_{j=1}^{n} f(x_j) .$$

We call θ the n^{th} approximation of the maximum likelihood estimate of θ or simply the maximum likelihood estimate of θ calculated from the first n observations.

In the following, we shall assume that the likelihood function $\prod_{j=1}^{n} f(x_j,\theta)$ has a relative maximum at $\theta = \hat{\theta}$ for fixed $x_1,x_2,\ldots,x_n$.

Proposition 1.1 (Doob) For each value of θ in some neighborhood $|\theta| \leq a_1, a_1 > 0$ of $\theta = 0$ (which is the true value of θ), let $f(x,\theta)$ be a probability density in the infinite interval $-\infty < x < \infty$. Let the true distribution of x be determined by the true probability density $f(x)$. Suppose (i) that $\log f(x,\theta)$ can be expressed in the form

$$(1.2) \qquad \log f(x,\theta) = \log f(x) + \theta\cdot\alpha(x) + \frac{\theta^2}{2}\beta(x) + \gamma(x,\theta) ,$$

where $\alpha(x)\ f(x)$, $\alpha^2(x)\ f(x)$, and $\beta(x)\ f(x)$ are Lebesque measurable and integrable over $-\infty < x < \infty$ and where

$$(1.3) \qquad \frac{\partial}{\partial\theta}\,\gamma(x,\theta) = \gamma_o(x,\theta)$$

exists for $\left|\theta\right| \leq a_2 \leq a_1, a_2 > 0$, and is continuous at $\theta = 0$.

(ii) that if

$$(1.4) \qquad \phi(x) = \underset{0<|\theta|\leq a_2}{\text{L. U. B.}} \left\{ \frac{\left|\gamma_o(x,\theta)\right|}{\theta^2} \right\} ,$$

then $\phi(x)\ f(x)$ is integrable over $-\infty < x < \infty$;

(iii) that if $\delta(x,\theta)$ is defined by

$$(1.5) \qquad f(x,\theta) = f(x)\left\{ 1 + \theta\alpha(x) + \frac{\theta^2}{2}\left[\beta(x) + \alpha(x)\right] + \delta(x,\theta) \right\} ,$$

then

$$(1.6) \qquad \lim_{\theta \to 0} \frac{1}{\theta^2} \int_{-\infty}^{\infty} \delta(x,\theta)\ f(x)\ dx = 0 .$$

Then

$$(1.7) \qquad \int_{-\infty}^{\infty} \alpha^2(x)\ f(x)\ dx + \int_{-\infty}^{\infty} \beta(x)\ f(x)\ dx = 0 .$$

Suppose that

$$(1.8) \qquad \sigma^2 = \int_{-\infty}^{\infty} \alpha^2(x)\ f(x)\ dx > 0 .$$

If the maximum likelihood estimate of θ is consistent, i.e.,

$$(1.9) \qquad \lim_{n\to\infty} \bar{P}_F(|\hat{\theta}| > \varepsilon) = 0^*$$

for every $\varepsilon > 0$, then it follows that

$$(1.10) \qquad \lim_{n\to\infty} \bar{P}_F(\sigma n^{\frac{1}{2}}\hat{\theta} < \lambda) = \lim_{n\to\infty} \underline{P}_F(\sigma n^{\frac{1}{2}}\hat{\theta} < x)$$
$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\lambda} e^{-\frac{t^2}{2}} dt \; {}^{**}$$

for every constant λ, uniformly in λ.

In the foregoing statements, it is to be observed that (1.7) and (1.10) are the conclusions, the rest being all premises or assumptions.

For the sake of convenience of later use, we shall state the same theorem in the multiparametric case and shall prove it.

Proposition 1.2. For each value of $\underline{\theta}' = (\theta_1, \ldots, \theta_s)$ in the neighborhood $|\theta| = \sqrt{\Sigma_{i=1}^{s} \theta_i^2} \leq a_1, a_1 > 0$ of $\underline{\theta} = \underline{0}$, let $f(x,\underline{\theta})$ be a probability density in the whole space R^m of the m-dimensional Euclidean space where

$$\underline{x}' = (x_1, x_2, \ldots, x_m)$$

denotes a point of R^m. Let the true distribution of $\underline{x}$ be determined

---

* For the definition of the probability measure P_F see Doob [2]. Since we have not assumed the measurability of the maximum likelihood estimate of θ, it is necessary to express the consistency in terms of the outer measure $\bar{P}_F$.

** $\underline{P}_F$ stands for the inner measure.

by the probability density $f(\underline{x}) \equiv f(\underline{x},\underline{0})$.

Suppose

(i) that $\log f(\underline{x},\underline{\theta})$ can be expressed in the form

$$(1.11)\quad \log f(\underline{x},\underline{\theta}) = \log f(\underline{x}) + \underline{\theta}'\cdot\underline{\alpha}(\underline{x}) + \frac{1}{2}\underline{\theta}'\beta(\underline{x})\underline{\theta} + \gamma(\underline{x},\underline{\theta})$$

where

$$\underline{\alpha}(\underline{x}) = \begin{bmatrix} \alpha_1(\underline{x}) \\ \vdots \\ \alpha_s(\underline{x}) \end{bmatrix}, \quad \beta(\underline{x}) = (\beta_{ik}(\underline{x})) \quad , \quad \beta_{ik}(\underline{x}) = \beta_{ki}(\underline{x}) \ ,$$

and $\underline{\alpha}(\underline{x})\ f(\underline{x})$, $\underline{\alpha}(\underline{x})\ \underline{\alpha}'(\underline{x})\ f(\underline{x})$ and $\beta(\underline{x})\ f(\underline{x})$ are all Lebesque measurable and integrable over R^m, and where

$$(1.12)\qquad \frac{\partial}{\partial\theta_i}\gamma(\underline{x},\underline{\theta}) = \gamma_i(\underline{x},\underline{\theta}) \quad , \quad i = 1,2,\ldots,s \quad ,$$

exist for $|\underline{\theta}| \leq a_2 \leq a_1, a_2 > 0$ and are continuous functions of $\underline{\theta}$ at $\underline{\theta} = \underline{0}$;

(ii) that if

$$(1.13)\qquad \phi_i(x) = \underset{0<\theta\leq a_2}{\text{L. U. B.}} \left\{ \frac{|\gamma_i(\underline{x},\underline{\theta})|}{\underline{\theta}'\underline{\theta}} \right\}, \quad i = 1,2,\ldots,s \ ,$$

then $\underline{\phi}(\underline{x})\ f(\underline{x})$ are integrable over R^m, where

$$\underline{\phi}(\underline{x}) = \begin{bmatrix} \phi_1(\underline{x}) \\ \vdots \\ \phi_s(\underline{x}) \end{bmatrix} \quad ,$$

(iii) that if $\delta(\underline{x},\underline{\theta})$ is defined by

(1.14) $$f(\underline{x},\underline{\theta}) = f(\underline{x}) \left\{ 1 + \underline{\theta}'\underline{\alpha}(\underline{x}) + \frac{1}{2}\underline{\theta}'[\beta(\underline{x}) + \underline{\alpha}(\underline{x})\,\underline{\alpha}'(\underline{x})]\underline{\theta} + \delta(\underline{x},\underline{\theta}) \right\},$$

then

(1.15) $$\lim_{|\underline{\theta}| \to 0} \frac{1}{|\underline{\theta}|^2} \int_{R^m} \delta(\underline{x},\underline{\theta})\, f(\underline{x})\, d\underline{x} = 0 \quad .$$

Then we have

(1.16) $$\int_{R^m} [\beta(\underline{x}) + \underline{\alpha}(\underline{x})\,\underline{\alpha}'(\underline{x})]\, f(\underline{x})\, d\underline{x} = 0 \quad .$$

Suppose that the matrix

(1.17) $$V = - \int_{R^m} \beta(\underline{x})\, f(\underline{x})\, d\underline{x}$$

is positive difinite and symmetric. If the maximum likelihood estimate $\hat{\underline{\theta}}$ is consistent, i.e.,

$$\lim_{n \to \infty} P_F(|\hat{\underline{\theta}}| > \varepsilon) = 0$$

for every $\varepsilon > 0$, then

(1.18) $$\lim_{n \to \infty} P_F(\sqrt{n}\,\hat{\underline{\theta}} < \underline{\lambda}) = \lim_{n \to \infty} P_F(\sqrt{n}\,\hat{\theta} < \underline{\lambda})^*$$
$$= \frac{1}{(2\pi)^{\frac{s}{2}} |V|^{\frac{1}{2}}} \int_{-\infty}^{\lambda_1} \cdots \int_{-\infty}^{\lambda_s} \exp\left[-\frac{1}{2}\underline{t}' V^{-1} \underline{t}\right] d\underline{t} \ ,$$

for every constant vector $\underline{\lambda}$ uniformly in $\underline{\lambda}$.

Proof: Since $f(\underline{x},\underline{\theta})$ is a probability density, we get

$$\int_{R^m} f(\underline{x},\underline{\theta})\, d\underline{x} = 1$$

for all $\underline{\theta}$ in the neighborhood $|\underline{\theta}| \leq a_1$. Thus we have the relation

$$(1.19) \qquad \underline{\theta}'\int_{R^m} \underline{\alpha}(\underline{x})f(\underline{x})d\underline{x} + \frac{1}{2}\underline{\theta}'\int_{R^m}[\beta(\underline{x})+\underline{\alpha}(\underline{x})\underline{\alpha}'(\underline{x})]f(\underline{x})d\underline{x}\cdot\underline{\theta} + \int_{R^m}\delta(\underline{x},\underline{\theta})f(\underline{x})d\underline{x} = 0 \ .$$

If we choose $\underline{\theta}$ in the neighborhood in such a way that

$$\theta_i = \varepsilon\delta_{ik} \ , \quad \varepsilon = 1,2,\ldots,s \ ,$$

then the relation (1.19) turns out to be

$$(1.20) \qquad \varepsilon\int_{R^m} \alpha_i(\underline{x})f(\underline{x})d\underline{x} + \frac{1}{2}\varepsilon^2\int_{R^m}[\beta_{ii}(\underline{x})+\alpha_i^2(\underline{x})]f(\underline{x})d\underline{x} + \int_{R^m}\delta(\underline{x},\theta)f(\underline{x})d\underline{x} = 0 \ .$$

Dividing (1.20) by ε and letting $\varepsilon \to 0$, we obtain on account of the ocndition (1.15) that

$$\int_{R^m} \alpha_i(\underline{x})\, f(\underline{x})\, d\underline{x} = 0 \ , \quad i = 1,2,\ldots,s \ ,$$

---

* For two vectors $\underline{a}$ and $\underline{b}$, the notation $\underline{a} < \underline{b}$ stands for the simultaneous inequalities $a_1<b_1$, $a_2<b_2,\ldots,$ $a_s<b_s$.

or, in vector notation,

$$\int_{R^m} \underline{\alpha}(\underline{x})\, f(\underline{x})\, d\underline{x} = \underline{0} \quad . \tag{1.21}$$

Dividing (1.20) by ε^2 and letting $\varepsilon \to 0$, it follows on account of relations (1.15) and (1.21) that

$$\int_{R^m} [\beta_{ii}(\underline{x}) + \alpha_i^2(\underline{x})]\, f(\underline{x})\, d\underline{x} = 0 \ , \ i = 1,\ldots,s \ . \tag{1.22}$$

Next, choose $\underline{\theta}$ as follows:

$$\underline{\theta}' = (0 \ldots 0 \underset{i}{\varepsilon} 0 \ldots 0 \underset{k}{\varepsilon} 0 \ldots 0) \quad .$$

Then, in view of (1.21) and (1.22), the relation (1.19) can be expressed as

$$\varepsilon^2 \int_{R^m} [\beta_{ik}(\underline{x})+\alpha_i(\underline{x})\alpha_k(\underline{x})] f(\underline{x}) d\underline{x} + \int_{R^m} \delta(\underline{x},\underline{\theta}) f(\underline{x}) d\underline{x} = 0 \quad . \tag{1.23}$$

Dividing (1.23) by ε^2 and letting $\varepsilon \to 0$, we get

$$\int_{R^m} [\beta_{ik}(\underline{x})+\alpha_i(\underline{x})\alpha_k(\underline{x})]\, f(\underline{x})\, d\underline{x} = 0 \quad \text{for } i \neq k \ . \tag{1.24}$$

Now we can assume, without any loss of generality, that

$$-\int_{R^m} \beta_{ik}(\underline{x})\, f(\underline{x})\, d\underline{x} = \xi_i \delta_{ik} \quad , \quad \xi_i > 0 \quad , \quad i = 1,\ldots,s \ . \tag{1.25}$$

In fact, if (1.25) does not hold, we shall introduce a new set of parameters $\underline{\phi}$ defined by

$$\underline{\theta} = B\, \underline{\phi} \quad , \tag{1.26}$$

where B is an orthogonal matrix such that

$$(1.27) \qquad B' \, V \, B = D_\xi \quad ,$$

and where

$$D_\xi = \begin{bmatrix} \xi_1 & & 0 \\ & \ddots & \\ 0 & & \xi_s \end{bmatrix} , \quad \xi_i > 0 \quad .$$

Since we have assumed that the matrix V is positive definite, there certainly exists such a B.

For such a new parameter system $\underline{\phi}$, the probability density function becomes

$$f^*(\underline{x},\underline{\phi}) = f(\underline{x},\underline{0}) \left\{ 1 + \underline{\phi}'B'\underline{\alpha}(\underline{x}) + \frac{1}{2} \underline{\phi}'B'[\beta(\underline{x})+\underline{\alpha}(\underline{x})\underline{\alpha}'(\underline{x})]B\underline{\phi}\cdot\delta(\underline{x},B\underline{\phi}) \right\},$$

which can be written as

$$(1.28) \qquad f^*(\underline{x},\underline{\phi}) = f^*(\underline{x},\underline{0}) \left\{ 1+\underline{\phi}'\underline{\alpha}^*(\underline{x})+\frac{1}{2}\underline{\phi}'[\beta^*(\underline{x})+\underline{\alpha}^*(\underline{x})\underline{\alpha}^{*\prime}(\underline{x})]\underline{\phi}+\delta^*(\underline{x},\underline{\phi}) \right\} ,$$

and we shall write the logarithm of it as

$$(1.29) \qquad \log f^*(\underline{x},\underline{\phi}) = \log f^*(\underline{x},\underline{0})+\underline{\phi}'\underline{\alpha}^*(\underline{x})+\frac{1}{2}\underline{\phi}'\beta^*(\underline{x})\underline{\phi}+\gamma^*(\underline{x},\underline{\phi}) \quad .$$

It will easily be seen that

$$\gamma_i^*(x,\phi) \equiv \frac{\partial \gamma^*(\underline{x},\underline{\phi})}{\partial \phi_i} = \sum_{k=1}^{s} \frac{\partial \gamma(\underline{x},\underline{\phi})}{\partial \theta_k}\frac{\partial \theta_k}{\partial \phi_i} = \sum_{k=1}^{s} b_{ik} \frac{\partial \gamma(\underline{x},\underline{\theta})}{\partial \theta_k}$$

where $\quad B = (b_{ik}) \quad$, and

$$\phi_i^*(\underline{x}) = \underset{0<|\underline{\phi}|\le a_2}{\text{L. U. B.}}\left\{\frac{|\gamma_i^*(\underline{x},\underline{\phi})|}{|\underline{\phi}|^2}\right\} = \underset{0<|\theta|\le a_2}{\text{L. U. B.}}\left\{\frac{\left|\sum_{k=1}^{s} b_{ik}\gamma_k(\underline{x},\underline{\theta})\right|}{|\underline{\theta}|^2}\right\}$$

or

(1.30) $$\phi_i^*(\underline{x}) = \sum_{k=1}^{s} |b_{ik}|\, \phi_k(\underline{x}) \quad .$$

Thus $\underline{\phi}^*(\underline{x})\, f^*(\underline{x})$ is integrable over R^m. Furthermore,

(1.31) $$\lim_{|\underline{\phi}|\to 0} \frac{1}{\underline{\phi}'\underline{\phi}} \int_{R^m} \delta^*(\underline{x},\underline{\phi})\, f^*(\underline{x})\, d\underline{x} = \lim_{|\underline{\theta}|\to 0} \frac{1}{\underline{\theta}'\underline{\theta}} \int_{R^m} \delta(\underline{x},\underline{\theta})\, f(\underline{x})\, d\underline{x} = 0 \quad .$$

Since $\underline{\phi}$ is a linear combination of $\underline{\theta}$, if $\underline{\theta}$ has a consistent maximum likelihood estimate, then $\underline{\phi}$ has also a consistent maximum likelihood estimate. Thus all conditions in the theorem are satisfied with respect to $\underline{\phi}$.

The logarithm of the likelihood function obtained from the first n observations is given by

(1.32) $$Ln(\underline{\theta}) = \sum_{j=1}^{n} \log f(\underline{x}_j) + \underline{\theta}'\sum_{j=1}^{n} \underline{\alpha}(\underline{x}_j) + \frac{1}{2}\,\underline{\theta}'\sum_{j=1}^{n} \beta(\underline{x})\underline{\theta} + \sum_{j=1}^{n} \gamma(\underline{x}_j,\underline{\theta}) \quad ,$$

and since $Ln(\underline{\theta})$ was supposed to have a relative maximum at $\underline{\theta} = \hat{\underline{\theta}}$, we obtain the equations

(1.33) $$\sum_{j=1}^{n} \underline{\alpha}(\underline{x}_j) + \hat{\underline{\theta}}\sum_{j=1}^{n} \beta(\underline{x}_j) + \sum_{j=1}^{n} \underline{\gamma}(\underline{x}_j,\hat{\underline{\theta}}) = \underline{0} \quad ,$$

where $\underline{\gamma}'(\underline{x},\underline{\theta}) = (\gamma_1(\underline{x},\underline{\theta}), \ldots, \gamma_s(\underline{x},\underline{\theta}))$. This can be rewritten as follows:

$$\frac{1}{\sqrt{n}} \sum_{j=1}^{n} \alpha_i(x_j) + \sum_{k=1}^{s} \sqrt{n}\, \hat{\theta}_k \left(\frac{1}{n} \sum_{j=1}^{n} \beta_{ik}(x_j)\right) + \sqrt{n}\, |\hat{\underline{\theta}}|^2 \left(\frac{1}{n} \sum_{j=1}^{n} \frac{\gamma_i(\underline{x}_j,\hat{\underline{\theta}})}{|\hat{\underline{\theta}}|^2} \right) = 0 \ ,$$

$$i = 1,2,\ldots,s \ ,$$

or

$$(1.34) \qquad \sum_{k=1}^{s} \sqrt{n}\, \hat{\theta}_k \left\{ -\frac{1}{n} \sum_{j=1}^{n} \beta_{ik}(\underline{x}_j) - \hat{\theta}_k \left(\frac{1}{n} \sum_{j=1}^{n} \frac{\gamma_i(\underline{x}_j,\hat{\underline{\theta}})}{|\hat{\underline{\theta}}|^2} \right) \right\} = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \alpha_i(\underline{x}_j) \ , \quad i = 1,2,\ldots,s \ .$$

Since

$$-\frac{1}{n} \sum_{j=1}^{n} \beta_{ik}(\underline{x}_j) = \xi_i \delta_{ik} \quad \text{with probability 1} \ ,$$

as $n \to \infty$, and

$$\left| \frac{1}{n} \sum_{j=1}^{n} \frac{\gamma_i(\underline{x}_j,\hat{\underline{\theta}})}{|\hat{\underline{\theta}}|^2} \right| \leq \frac{1}{n} \sum_{j=1}^{n} \phi_i(\underline{x}_j)$$

and also

$$\frac{1}{n} \sum_{j=1}^{n} \phi_i(\underline{x}_j) \longrightarrow \int_{R^m} \phi_i(\underline{x})\, f(\underline{x})\, d\underline{x} = K_i < \infty \ ,$$

with probability 1, as $n \to \infty$, we can see that

$$(1.35)\qquad \sqrt{n}\,\xi_i\hat{\theta}_i = \frac{1}{\sqrt{n}}\Sigma_{j=1}^{n}\,\alpha_i(\underline{x}_j) + R_i \quad , \quad i = 1,\ldots,s$$

where $R_i \to 0$ in probability as $n \to \infty$.

Applying Doob's theorem [2] to (1.35) and the Central Limit Theorem [1] to $\frac{1}{\sqrt{n}}\,\Sigma_{j=1}^{n}\alpha_i(\underline{x}_j)$, we obtain

$$(1.36)\qquad \lim_{n\to\infty} P_F\left\{\sqrt{n}\,\hat{\theta}_1\xi_1 \le \mu_1', \ldots, \sqrt{n}\,\hat{\theta}_s\xi_s \le \mu_s'\right\}$$

$$= \lim_{n\to\infty} P_F\left\{\frac{1}{\sqrt{n}}\,\Sigma_{j=1}^{n}\alpha_1(\underline{x}_j) \le \mu_1', \ldots, \frac{1}{\sqrt{n}}\,\Sigma_{j=1}^{n}\alpha_s(\underline{x}_j) \le \mu_s'\right\}$$

$$= \frac{1}{(2\pi)^{\frac{s}{2}}\,(\xi_1\cdots\xi_s)^{\frac{1}{2}}}\int_{-\infty}^{\mu_1'}\cdots\int_{-\infty}^{\mu_s'} e^{-\frac{1}{2}\left(\frac{t_1^2}{\xi_1}+\cdots+\frac{t_s^2}{\xi_s}\right)}\,dt_1\cdots dt_s .$$

and consequently we get, by putting

$$\frac{\mu_i'}{\xi_i} = \lambda_i \quad , \quad i = 1,\ldots,s \quad ,$$

$$(1.37)\qquad \lim_{n\to\infty} P_F\left\{\sqrt{n}\,\hat{\underline{\theta}} < \underline{\lambda}\right\} = \frac{1}{(2\pi)^{\frac{k}{2}}\,|V^{-1}|^{\frac{1}{2}}}\int_{-\infty}^{\lambda_1}\cdots\int_{-\infty}^{\lambda_s} e^{-\frac{1}{2}\underline{t}'V\underline{t}}\,d\underline{t} .$$

In other words, $\sqrt{n}\,\hat{\underline{\theta}}$ has asymptotic normal distribution around the true value $\underline{\theta} = \underline{0}$ with variance-covariance matrix V^{-1}.

Q. E. D.

Using the above theorem, we shall now give a rigorous proof of the following result, stated and proved in sketch by Wilks [6], and also proved rigorously by Wald, on assumptions somewhat different from Doob's.

<u>Proposition 1.3 (Wilks)</u> To test a composite hypothesis

$$(1.38) \qquad H_o: \ \theta_1 = \theta_1^o , \ \ldots , \ \theta_r = \theta_r^o \qquad (r < s)$$

against the alternatives $H \neq H_o$ in the situation given in the previous Proposition, the statistic

$$(1.39) \qquad -2 \log \lambda = -2 \log \frac{\max \prod_{j=1}^{n} f(\underline{x}_j \mid H_o)}{\max \prod_{j=1}^{n} f(\underline{x}_j \mid H)}$$

has the X^2-distribution with r degrees of freedom, in the limit, as $n \to \infty$, no matter which simple hypothesis might be true under the composite null-hypothesis H_o.

<u>Proof:</u> For the time being, we assume that the true value of the parameter $\underline{\theta}$ is

$$\underline{\theta}^{o\prime} = (\theta_1^o \ldots \theta_r^o \ \theta_{r+1}^o \ldots \theta_s^o) \ .$$

Let the maximum likelihood estimate of $\underline{\theta}$ under H_o be

$$\underline{\theta}^{*\prime} = (\theta_1^o \ldots \theta_r^o \ \hat{\theta}_{r+1}^* \ldots \hat{\theta}_s^*) \quad ,$$

and let the maximum likelihood estimate of $\underline{\theta}$ under H be $\hat{\underline{\theta}}$. Then we have

$$(1.40) \qquad \log \max \prod_{j=1}^{n} f(\underline{x}_j \quad \underline{\theta}^o) = Ln(\hat{\underline{\theta}}^*)$$

$$= \Sigma_{j=1}^{n} \log f(\underline{x}_j \mid \underline{\theta}^o) + (\hat{\underline{\theta}}^* - \underline{\theta}^o)' \Sigma_{j=1}^{n} \underline{\alpha}(\underline{x}_j) + \tfrac{1}{2} (\hat{\underline{\theta}}^* - \underline{\theta}^o)' \Sigma_{j=1}^{n} \beta(\underline{x}_j) (\hat{\underline{\theta}}^* - \underline{\theta}^o)$$

$$+ \Sigma_{j=1}^{n} \gamma(\underline{x}_j, \hat{\underline{\theta}}^* - \underline{\theta}^o) \quad ,$$

and

$$(1.41) \qquad \log \max \prod_{j=1}^{n} f(\underline{x}_j \mid H) = Ln(\hat{\underline{\theta}})$$

$$= \sum_{j=1}^{n} \log f(\underline{x}_j \mid \underline{\theta}^o) + (\hat{\underline{\theta}}-\underline{\theta}^o)' \sum_{j=1}^{n} \underline{\alpha}(\underline{x}_j)$$

$$+ \frac{1}{2}(\underline{\theta}-\underline{\theta}^o)' \sum_{j=1}^{n} \beta(\underline{x}_j)(\hat{\underline{\theta}}-\underline{\theta}^o) + \sum_{j=1}^{n} \gamma(\underline{x}_j, \hat{\underline{\theta}}-\underline{\theta}^o) \quad ,$$

and we know by Prop. 1.2 that

$$(1.42) \qquad \sqrt{n}\,(\hat{\theta}_i^*-\theta_i^o)\cdot\xi_i = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \alpha_i(\underline{x}_j) + R_i^{*(n)} \quad , \quad i = r+1,\ldots,s \quad ,$$

and

$$(1.43) \qquad \sqrt{n}\,(\hat{\theta}_i-\theta_i^o)\cdot\xi_i = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \alpha_i(\underline{x}_j) + R_i^{(n)} \quad , \quad i = 1,2,\ldots,s \quad ,$$

where $R_i^{*(n)}$, $R_i^{(n)}$ converge to zero in probability as $n \to \infty$.

From (1.40), (1.41) we obtain, on account of (1.42), (1.43)

$$(1.44) \qquad Ln(\hat{\underline{\theta}}^*) = \sum_{j=1}^{n} \log f(\underline{x}_j, \underline{\theta}^o) + \frac{1}{2} \sum_{i=r+1}^{s} \frac{1}{\xi_i}\left(\frac{1}{\sqrt{n}} \sum_{j=1}^{n} \alpha_i(\underline{x}_j)\right) + Rn^*$$

$$(1.45) \qquad Ln(\hat{\underline{\theta}}) = \sum_{j=1}^{n} \log f(\underline{x}_j, \underline{\theta}^o) + \frac{1}{2} \sum_{i=1}^{s} \frac{1}{\xi_i}\left(\frac{1}{\sqrt{n}} \sum_{j=1}^{n} \alpha_i(\underline{x}_j)\right) + Rn \quad ,$$

where Rn^* and Rn tend to zero in probability as $n \to \infty$. Thus we obtain

$$(1.46) \qquad -2 \log \lambda = -2\left[Ln(\hat{\underline{\theta}}^*) - Ln(\hat{\underline{\theta}})\right] = \sum_{i=1}^{r} \frac{1}{\xi_i}\left(\frac{1}{\sqrt{n}} \sum_{j=1}^{n} \alpha_i(\underline{x}_j)\right)^2 + Rn \ ,$$

where Rn converges in probability to zero as $n \to \infty$.

Since, by virtue of the Central Limit Theorem, the random vector

$$\begin{bmatrix} \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \alpha_1(\underline{x}_j) \\ \vdots \\ \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \alpha_s(\underline{x}_j) \end{bmatrix}$$

has asymptotically the normal distribution $N(\underline{O};D_\xi)$, in the limit as $n \to \infty$, it follows from (1.46) that $-2 \log \lambda$ has the limiting X^2-distribution with r degrees of freedom, as $n \to \infty$. It is easily seen that this limiting distribution is independent of the specific value of $\underline{\theta}^o$.

Q. E. D.

2. A Theorem on Maximum Likelihood Estimate. We have seen in the preceding section that if the maximum likelihood estimate which was defined by (1.1) exists and is consistent, then under the conditions posed by Prop. 1.7 or 1.2, the maximum likelihood estimate has the asymptotic normality and further that $-2 \log \lambda$ has limiting X^2-distribution.

If we define the maximum likelihood estimate by

$$(2.1) \qquad \prod_{j=1}^{n} f(\underline{x}_j,\hat{\underline{\theta}}) = \max_{\underline{\theta} \in \Omega} \prod_{j=1}^{n} f(\underline{x}_j,\underline{\theta})$$

and if it exists at all, this must satisfy the likelihood equation, i.e.,

$$(2.2) \qquad \left[\frac{\partial \operatorname{Ln}(\underline{\theta})}{\partial \theta_i}\right]_{\underline{\theta}=\hat{\underline{\theta}}} = 0 \quad , \quad i = 1,2,\ldots,s .$$

We can state the following

<u>Theorem 2.1</u> Under the same situation as Prop. 1.2, if $\gamma(\underline{x},\underline{\theta})$ has mixed second order partial derivatives with respect to $\underline{\theta}$ and they are continuous functions of $\underline{\theta}$ with finite expectations such that

$$(2.3) \qquad K = \sup_{i,k} \sup_{|\underline{\theta}| \le a_2} \left| \int_{R^m} \frac{\partial^2 \gamma(\underline{x},\underline{\theta})}{\partial \theta_i \partial \theta_k} f(\underline{x})\, d\underline{x} \right| < \min_i \xi_i = \xi ,$$

then it follows that the maximum likelihood equation has one and only one consistent solution for sufficiently large values of n. Thus, if the maximum likelihood estimate $\hat{\underline{\theta}}$ exists at all, then this should be consistent in this case.

<u>Proof:</u> In this situation since the likelihood function $\prod_{j=1}^{n} f(\underline{x}_j,\underline{\theta})$ is a continuous function of $\underline{\theta}$ in the closed set $|\underline{\theta}| \le a_2$, it attains its maximum at $\underline{\theta} = \hat{\underline{\theta}}$, and $\hat{\underline{\theta}}$ must satisfy the maximum likelihood equation, i.e.,

$$(2.4) \qquad \sum_{j=1}^{n} \underline{\alpha}(\underline{x}_j) + \hat{\underline{\theta}} \sum_{j=1}^{n} \beta(\underline{x}_j) + \sum_{j=1}^{n} \underline{\gamma}(\underline{x}_j,\underline{\theta}) = 0 ,$$

where

$$\underline{\gamma}(\underline{x},\underline{\theta}) = \begin{bmatrix} \gamma_1(\underline{x},\underline{\theta}) \\ \vdots \\ \gamma_s(\underline{x},\underline{\theta}) \end{bmatrix}, \quad \text{and} \quad \gamma_i(\underline{x},\underline{\theta}) = \frac{\partial}{\partial \theta_i} \gamma(\underline{x},\underline{\theta}) .$$

Now we shall show by an iterative method that the equation (2.4) has a consistent solution.

We shall define a sequence of successive approximations $\{\hat{\underline{\theta}}^{(\nu)}\}$

by the following equations,

$$(2.5) \qquad \sum_{j=1}^{n} \underline{\alpha}(\underline{x}_j) + \hat{\underline{\theta}}^{(o)} \sum_{j=1}^{n} \beta(\underline{x}_j) = 0 \quad ,$$

$$(2.6) \qquad \sum_{j=1}^{n} \underline{\alpha}(\underline{x}_j) + \underline{\theta}^{(\nu)} \sum_{j=1}^{n} \underline{\gamma}(\underline{x}_j, \underline{\theta}^{(\nu-1)}) = 0 \ , \ \nu = 1,2,\dots \ .$$

Here again we shall assume, without any loss of generality, that

$$(2.7) \qquad - \int_{R^m} \beta(\underline{x}) \ f(\underline{x}) \ d\underline{x} = D_\xi \ .$$

Since $- \lim_{n \to \infty} \frac{1}{n} \sum_{j=1}^{n} \beta(x_j) = D_\xi$ with probability one as $n \to \infty$, we can choose n sufficiently large so that

$$(2.8) \qquad - \frac{1}{n} \sum_{j=1}^{n} \beta_{ik}(\underline{x}_j) = \xi_i \ \delta_{ik} + \varepsilon_{ik} \quad ,$$

where $\varepsilon_{ik} \to 0$ with probability one as $n \to \infty$, thus from (2.5) we get

$$(2.9) \qquad D_\xi \cdot \sqrt{n} \ \hat{\underline{\theta}}^{(o)} = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \underline{\alpha}(\underline{x}_j) + \underline{\varepsilon} \quad ,$$

where $\underline{\varepsilon}' = (\varepsilon_1, \dots, \varepsilon_s)$ and $\varepsilon_i \to 0$ with probability one as $n \to \infty$.
Hence by the Central Limit Theorem we obtain

$$(2.10) \qquad \lim_{n \to \infty} P_F \left\{ \sqrt{n} \ \hat{\theta}_1^{o} \le \lambda_1, \ \dots \ , \sqrt{n} \ \hat{\theta}_s^{(o)} \le \lambda_s \right.$$

$$= \frac{(\xi_1 \dots \xi_s)^{\frac{1}{2}}}{(2\pi)^{\frac{k}{2}}} \int_{-\infty}^{\lambda_1} \dots \int_{-\infty}^{\lambda_s} e^{-\frac{1}{2}(\xi_1 t_1^2 + \dots + \xi_s t_s^2)} \ dt_1 \dots dt_s \ .$$

This expresses the fact that, for each i, $\hat{\theta}_i^{(o)}$ is distributed asymptotically as an $N(0,\xi_i/n)$; and therefore $\hat{\underline{\theta}}^{(o)}$ also is a consistent estimate of $\underline{\theta}$.

Next, we shall show that $\hat{\underline{\theta}}^{(\nu)}$ converges, as $\nu \to \infty$, to a solution $\hat{\underline{\theta}}$ of the equation (2.4), and $\hat{\underline{\theta}}$ has the same limiting distribution as that of $\hat{\underline{\theta}}^{(o)}$.

From (2.6) it follows that

$$(\hat{\underline{\theta}}^{(\nu)} - \hat{\underline{\theta}}^{(\nu-1)}) \sum_{j=1}^{n} \beta(\underline{x}_j) + \sum_{j=1}^{n} [\underline{\gamma}(\underline{x}_j, \hat{\underline{\theta}}^{(\nu-1)}) - \underline{\gamma}(\underline{x}_j, \hat{\underline{\theta}}^{(\nu-2)})] = 0$$

or

(2.11)

$$(\hat{\underline{\theta}}^{(\nu)} - \hat{\underline{\theta}}^{(\nu-1)}) \frac{1}{n} \sum_{j=1}^{n} \beta(\underline{x}_j) + \frac{1}{n} \sum_{j=1}^{n} [\underline{\gamma}(\underline{x}_j, \hat{\underline{\theta}}^{(\nu-1)}) - \underline{\gamma}(\underline{x}_j, \hat{\underline{\theta}}^{(\nu-2)})] = 0 .$$

Now since

$$\gamma_i(\underline{x}, \hat{\underline{\theta}}^{(\nu-1)}) - \gamma_i(\underline{x}, \hat{\underline{\theta}}^{(\nu-2)}) = \sum_{k=1}^{s} (\hat{\theta}_k^{(\nu-1)} - \hat{\theta}_k^{(\nu-2)}) \gamma_{ik}(\underline{x}, \hat{\underline{\theta}}^{*(\nu-2)}) ,$$

where $\underline{\theta}^{*(\nu-1)}$ stands for a certain point on the segment connecting $\underline{\theta}^{(\nu-1)}$ and $\underline{\theta}^{(\nu-2)}$ in Ω, and by our assumption (2.3)

$$(2.12) \qquad \lim_{n \to \infty} \frac{1}{n} \sum_{j=1}^{n} \gamma_{ik}(\underline{x}_j, \hat{\underline{\theta}}^{*(\nu-1)}) = \int_{R^m} \gamma_{ik}(\underline{x}, \hat{\underline{\theta}}^{*(\nu-1)}) f(\underline{x}) \, d\underline{x}$$

$$= K_{ik}(\hat{\underline{\theta}}^{*(\nu-1)}) \leq K ,$$

with probability one as $n \to \infty$, therefore, for sufficiently large values of n, we obtain from (2.11) that

$$(2.13)\qquad \left|\hat{\theta}_i^{(\nu)} - \hat{\theta}_i^{(\nu-1)}\right| \leq \frac{K}{\xi_\mu}\left|\underline{\hat{\theta}}^{(\nu-1)} - \underline{\hat{\theta}}^{(\nu-2)}\right| \quad , \; i = 1,2,\ldots,s \quad ,$$

and consequently

$$(2.14)\qquad \left|\underline{\hat{\theta}}^{(\nu)} - \underline{\hat{\theta}}^{(\nu-1)}\right| \leq \frac{K}{\xi}\left|\underline{\hat{\theta}}^{(\nu-1)} - \underline{\hat{\theta}}^{(\nu-2)}\right| \quad , \quad \nu = 2,3,\ldots .$$

and

$$(2.15)\qquad \left|\underline{\hat{\theta}}^{(1)} - \underline{\hat{\theta}}^{(o)}\right| \leq \frac{K}{\xi}\left|\underline{\hat{\theta}}^{(o)}\right| \quad .$$

Putting

$$(2.16)\qquad H(n) = \left|\underline{\hat{\theta}}^{(o)}\right| \quad ,$$

it will easily be seen that $H(n) \to 0$ in probability as $n \to \infty$. Combining (2.14), (2.15) and (2.16) we get

$$(2.17)\qquad \left|\underline{\hat{\theta}}^{(\nu)} - \underline{\hat{\theta}}^{(\nu-1)}\right| \leq \left(\frac{K}{\xi}\right)^{\nu-1} H(n) \quad .$$

If we define $\underline{\theta}$ by

$$(2.18)\qquad \underline{\hat{\theta}} = \underline{\hat{\theta}}^{(o)} + (\underline{\hat{\theta}}^{(1)} - \underline{\hat{\theta}}^{(o)}) + (\underline{\hat{\theta}}^{(2)} - \underline{\hat{\theta}}^{(1)}) + \ldots \quad ,$$

the right hand side of (2.18) converges absolutely for sufficiently large values of n, and $\underline{\hat{\theta}}$ turns out to be a solution of (2.4). Furthermore, as is easily seen, $\left|\underline{\hat{\theta}} - \underline{\hat{\theta}}^{(o)}\right|$ is dominated by the quantity

$$(2.19)\qquad \frac{K}{\xi} H(n) \,/\, \left(1 - \frac{K}{\xi}\right) \quad ,$$

which converges in probability to zero os $n \to \infty$. Thus $\underline{\hat{\theta}}$ has the same limiting distribution as that of $\underline{\hat{\theta}}^{(o)}$, i.e., $\underline{\hat{\theta}}$ is a consistent estimate of $\underline{\theta}$.

Finally we shall show the uniqueness of the solution of (2.4) for sufficiently large values of n.

Suppose that (2.4) has two solutions $\hat{\underline{\theta}}$ and $\hat{\underline{\theta}}^*$; then we have

$$(\hat{\underline{\theta}} - \hat{\underline{\theta}}^*) \cdot \frac{-1}{n} \sum_{j=1}^{n} \beta(\underline{x}_j) = \frac{1}{n} \sum_{j=1}^{n} (\underline{\gamma}(\underline{x}_j, \hat{\underline{\theta}}) - \underline{\gamma}(\underline{x}_j, \hat{\underline{\theta}}^*))$$

or

$$(2.20) \qquad \left| \frac{-1}{n} \sum_{j=1}^{n} \beta(\underline{x}_j) \right| \left| \hat{\underline{\theta}} - \hat{\underline{\theta}}^* \right| \leqq K \left| \hat{\underline{\theta}} - \hat{\underline{\theta}}^* \right| .$$

Letting $n \to \infty$, we get

$$(2.21) \qquad 1 \leqq \frac{K}{\xi} ,$$

unless $\left| \hat{\underline{\theta}} - \hat{\underline{\theta}}^* \right| \to 0$ as $n \to \infty$, and (2.21) is a contradiction to our assumption (2.3). Q. E. D.

3. Testing Hypotheses on Certain Multivariate Populations, Some of Whose Variates are Continuous and the Rest are Categorical [3].

We shall consider the case of a 3-variate-(X,Y,Z) population, where X,Y are continuous variates and Z is categorical. Suppose Z can/over range r categories and suppose that the conditional distribution of X,Y, given that Z belongs to the i-th category, is a bivariate normal distribution with the mean

$$\underline{\mu}(i) = \begin{bmatrix} \mu_1(i) \\ \mu_2(i) \end{bmatrix},$$

and with the variance-covariance matrix

$$V(i) = \begin{bmatrix} v_{11}(i) & v_{12}(i) \\ v_{12}(i) & v_{22}(i) \end{bmatrix}.$$

Suppose we have a sample of n observations, where n is fixed from sample to sample, such that n_i individuals belong to the i-th category, $i = 1,2,\ldots,r$, and hence $\Sigma_{i=1}^{r} n_i = n$. Every individual belonging to the i-th category has two measurements (X_{ij}, Y_{ij}), $j = 1,2,\ldots,n_i$.

We shall consider the following two cases separately:

I. Z is a random variable. In this case n_i are random variables subject to the restriction $\Sigma_{i=1}^{r} n_i = n$ (fixed).

II. Z is a way of classification. In this case n_i $(i = 1,2,\ldots,r)$ are fixed numbers subject to the restriction $\Sigma_{i=1}^{r} n_i = n$.

Case I.

Let p_i $(i = 1,2,\ldots,r)$ be the probability that an observed individual belongs to the i-th category and hence $\Sigma_{i=1}^{r} p_i = 1$. In this case, the likelihood function is given by

$$(3.1) \qquad P\left\{ U, Z=\underline{n} \right\} = P\left\{ U \,\middle|\, Z=\underline{n} \right\} P(Z=\underline{n})$$

$$= \prod_{i=1}^{r} \left(\frac{1}{2\pi \left| V(i) \right|^{\frac{1}{2}}} \right)^{n_i} \exp\left\{ -\frac{1}{2} \sum_{i=1}^{r} \sum_{j=1}^{n_i} (\underline{U}_{ij} - \underline{\mu}(i))' V(i)^{-1} (\underline{U}_{ij} - \underline{\mu}(i)) \right\}$$

$$\times \frac{n!}{\prod_{i=1}^{r} n_i!} \prod_{i=1}^{r} p_i^{n_i} ,$$

where $\underline{U}_{ij}' = (X_{ij}, Y_{ij})$ and U stands for all $\underline{U}_{ij}$.

We shall be concerned with testing of null-hypotheses of the following types.

1a. Conditional independence between X and Y, given Z.

This amounts to testing the null-hypothesis

$$(3.2) \qquad H_o: \; V_{12}(i) = 0 \text{ for all } i \text{ against } H \neq H_o .$$

The statistic $-2 \log \lambda$ in this case turns out to be

$$(3.3) \qquad -\sum_{i=1}^{r} n_i \log (1 - \hat{r}^2(i)) \quad ,$$

where $\hat{r}(i)$ is the estimate of the ordinary correlation coefficient calculated from the sample belonging to the i-th category. We can show that the quantity (3.4) is equal in probability to

$$(3.5) \qquad \sum_{i=1}^{r} n_i \hat{r}^2(i) \quad .$$

It will be shown in the next section that the population probability distribution in this case will satisfy all conditions of Prop. 1.2 and Theorem 2.1. Hence the statistic given by (3.5) has a limiting distribution which is the X^2-distribution with degrees of freedom r.

1b. Independence between (X,Y) and Z.

This amounts to testing the null-hypothesis

$$(3.6) \qquad H_o: \underline{\mu}(i) = \underline{\mu} \equiv \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \quad V(i) = V \equiv \begin{bmatrix} v_{11} & v_{12} \\ v_{12} & v_{22} \end{bmatrix}$$

for all $i = 1,2,\ldots,r$, against $H \neq H_o$, where $\underline{\mu}$ and V consist of five arbitrary nuisance parameters.

The statistic $-2 \log \lambda$ in this case turns out to be

$$(3.7) \qquad \sum_{i=1}^{r} n_i \log \frac{\hat{v}_{11}\hat{v}_{22}(1-\hat{r}^2)}{\hat{v}_{11}(i)\hat{v}_{22}(i)(1-\hat{r}^2(i))} ,$$

and this is equal in probability to

$$(3.8) \qquad \sum_{i=1}^{r} n_i(\underline{\mu}(i)-\underline{\mu})' \hat{U}_1(\underline{\mu}(i)-\underline{\mu}) + \sum_{i=1}^{r} n_i(\hat{\underline{v}}^*(i)-\hat{\underline{v}}^*)' \hat{U}_2(\hat{\underline{v}}^*(i)-\hat{\underline{v}}^*) ,$$

where

$$(3.9) \qquad \hat{\underline{v}}^{*\prime} = (\hat{v}_{11}, \hat{v}_{22}, \hat{v}_{12}) ,$$

$$(3.10) \qquad \hat{U}_1 = \begin{bmatrix} \hat{v}_{11} & \hat{v}_{12} \\ \hat{v}_{12} & \hat{v}_{22} \end{bmatrix}^{-1} \equiv \begin{bmatrix} \hat{v}^{11} & \hat{v}^{12} \\ \hat{v}^{12} & \hat{v}^{22} \end{bmatrix} ,$$

and

$$(3.11) \qquad \hat{U}_2 = \begin{bmatrix} \frac{1}{2}(\hat{v}^{11})^2 & \frac{1}{2}(\hat{v}^{12})^2 & \hat{v}^{11}\hat{v}^{12} \\ \frac{1}{2}(\hat{v}^{12})^2 & \frac{1}{2}(\hat{v}^{22})^2 & \hat{v}^{12}\hat{v}^{22} \\ \hat{v}^{11}\hat{v}^{12} & \hat{v}^{22}\hat{v}^{12} & \hat{v}^{11}\hat{v}^{22}+(\hat{v}^{12})^2 \end{bmatrix} ,$$

and $\hat{r}$ is the estimated correlation coefficient calculated from the data pooled from all the categories. This statistic (3.8) has an asymptotic X^2-distribution with $5(r-1)$ degrees of freedom.

__1c.__ Independence between (X,Z) and Y.

It can be shown that this amounts to testing the null-hypothesis

$$(3.12) \qquad H_o: \; \mu_2(i) = \mu_2 \; , \; v_{22}(i) = v_{22} \; , \; v_{12}(i) = 0 \quad \text{for all } i = 1,2,\ldots,r$$

against $H \neq H_o$.

The statistic $-2 \log \lambda$ in this case turns out to be

(3.13) $$\sum_{i=1}^{r} n_i \log \frac{\hat{v}_{22}}{\hat{v}_{22}(i)(1-\hat{r}^2(i))} ,$$

and this is equal in probability to

(3.14) $$\sum_{i=1}^{r} \left[\frac{n_i(\hat{\mu}_2(i) - \hat{\mu}_2)^2}{\hat{v}_{22}} + \frac{n_i(\hat{v}_{22}(i) - \hat{v}_{22})^2}{2\hat{v}_{22}^2} + n_i\hat{r}^2(i) \right] .$$

Thus this statistic has an asymptotic X^2-distribution with (3r-2) degrees of freedom.

<u>Id.</u> Total independence.

What we are interested in is to test the null-hypothesis

(3.15) $H_o: \underline{\mu}(i) = \underline{\mu}$, $V(i) = V$ where $v_{12} = 0$ for all $i = 1,2,\ldots,r$,

against $H \neq H_o$. The statistic $-2 \log \lambda$ is given by

(3.16) $$\sum_{i=1}^{r} n_i \log \frac{\hat{v}_{11}\hat{v}_{22}}{\hat{v}_{11}(i)\hat{v}_{22}(i)(1-\hat{r}^2(i))} ,$$

and this is equal in probability to

(3.17) $$\sum_{k=1}^{2}\left[\sum_{i=1}^{r} n_i \left\{ \frac{(\hat{\mu}_k(i)-\hat{\mu}_k)^2}{\hat{v}_{kk}} + \frac{(\hat{v}_{kk}(i)-\hat{v}_{kk})^2}{2\hat{v}_{kk}^2} \right\}\right] + \sum_{i=1}^{r} \frac{n_i(\hat{v}_{12}(i))^2}{\hat{v}_{11}\hat{v}_{22}} .$$

This statistic has an asymptotic X^2-distribution with (5r-4) degrees of freedom.

Case II

In this case we are dealing with r independent bivariate normal populations. The hypothesis corresponding to Ia will be that X and Y are independent in all the r different bivariate normal populations, that corresponding to Ib will be that (X,Y) will have the same distribution in all different r populations, that corresponding to Id will be that X and Y are independent and have the same distribution in all the r populations. Ic has no analogue in case II.

For each of these problems, the statistic and the asymptotic distribution (on the null hypothesis) are the same as the corresponding ones in case I. However, the asymptotic power of the test for any problem in case II would differ from that of the corresponding test in case I.

4. Verification of the Validity of Doob's Conditions for Some Mixed Variates Population

We shall show in this section that all conditions of Prop. 1.2 and Theorem 2.1 are satisfied in the special case which was treated in section 3.

For the sake of simplicity of notation, we shall put

$$\theta_1(i) = \mu_1(i)\ ,\ \theta_2(i) = \mu_2(i)\ ,\ \theta_3(i) = v_{11}(i)\ , \tag{4.1}$$

$$\theta_4(i) = v_{22}(i)\ ,\ \theta_5(i) = v_{12}(i)\ ,$$

$$i = 1, 2, \ldots, r\ .$$

Assuming that $\underline{\theta}^o$ is the true value of the parameter, and dropping the categorical symbol i for the time being, we can express the logarithm of the density function as follows:

$$(4.2) \qquad \log f(X,Y,\underline{\theta}) = \log f(X,Y,\underline{\theta}^o) + (\underline{\theta}-\underline{\theta}^o)'\underline{\alpha}(X,Y) + \tfrac{1}{2}(\underline{\theta}-\underline{\theta}^o)'\beta(X,Y)(\underline{\theta}-\underline{\theta}^o) + \gamma(X,Y,\underline{\theta}) \quad ,$$

where

$$(4.3) \qquad \alpha_i(X,Y) = \left(\frac{\partial \log f(X,Y,\underline{\theta})}{\partial \theta_i}\right)_{\underline{\theta} = \underline{\theta}^o} \quad , \; i = 1,\dots,5 \quad ,$$

$$\beta_{ik}(X,Y) = \left(\frac{\partial^2 \log f(X,Y,\underline{\theta})}{\partial \theta_i \partial \theta_k}\right)_{\underline{\theta} = \underline{\theta}^o} \quad , \; i,k=1,\dots,5 \quad ,$$

$$(4.4) \qquad \gamma(X,Y,\underline{\theta}) = \log \frac{f(X,Y,\underline{\theta})}{f(X,Y,\underline{\theta}^o)} - (\underline{\theta}-\underline{\theta}^o)'\underline{\alpha}(X,Y) - \tfrac{1}{2}(\underline{\theta}-\underline{\theta}^o)'\beta(X,Y)(\underline{\theta}-\underline{\theta}^o) \quad .$$

Since

$$(4.5) \qquad \log f(X,Y,\underline{\theta}) = C - \tfrac{1}{2}\log(\theta_3\theta_4-\theta_5^2) - \frac{1}{2(\theta_3\theta_4-\theta_5^2)} \times \left\{\theta_4(X-\theta_1)^2-2\theta_5(X-\theta_1)(Y-\theta_2)+\theta_3(Y-\theta_2)^2\right\} \quad ,$$

it follows that

$$\alpha_1(X,Y) = \left(\frac{\partial \log f(X,Y,\underline{\theta})}{\partial \theta_1}\right)_{\underline{\theta}^o} = \frac{\theta_4^o(X-\theta_1^o)-\theta_5^o(Y-\theta_2^o)}{\theta_3^o\theta_4^o-\theta_5^{o2}} \quad ,$$

$$\alpha_2(X,Y) = \left(\frac{\partial \log f(X,Y,\underline{\theta})}{\partial \theta_2}\right)_{\underline{\theta}^o} = \frac{-\theta_5^o(X-\theta_1^o)+\theta_3^o(Y-\theta_2^o)}{\theta_3^o\theta_4^o-\theta_5^{o2}} \quad ,$$

$$(4.6) \qquad \alpha_3(X,Y) = \left(\frac{\partial \log f(X,Y,\underline{\theta})}{\partial \theta_3}\right)_{\underline{\theta}^o} = -\frac{\theta_4^o}{2(\theta_3^o\theta_4^o-\theta_5^{o2})} + \frac{\theta_4^o}{2(\theta_3^o\theta_4^o-\theta_5^{o2})^2} \times \left\{\theta_4^o(X-\theta_1^o)^2-s\theta_5^o(X-\theta_1^o)(Y-\theta_2^o)+\theta_3^o(Y-\theta_2^o)^2\right\} - \frac{(Y-\theta_2^o)^2}{2(\theta_3^o\theta_4^o-\theta_5^{o2})} \quad ,$$

$$\alpha_4(X,Y) = \left(\frac{\partial \log f(X,Y,\underline{\theta})}{\partial \theta_4}\right)_{\underline{\theta}^o} = -\frac{\theta_3^o}{2(\theta_3^o\theta_4^o-\theta_5^{o2})} + \frac{\theta_3^o}{2(\theta_3^o\theta_4^o-\theta_5^{o2})}$$

$$\times \left\{\theta_4^o(X-\theta_1^o)^2-2\theta_5^o(X-\theta_1^o)(Y-\theta_2^o)+\theta_3^o(Y-\theta_2^o)^2\right\} - \frac{(X-\theta_1^o)^2}{2(\theta_3^o\theta_4^o-\theta_5^{o2})^2} ,$$

$$\alpha_5(X,Y) = \left(\frac{\partial \log f(X,Y,\underline{\theta})}{\partial \theta_5}\right)_{\underline{\theta}^o} = \frac{\theta_5^o}{\theta_3^o\theta_4^o-\theta_5^{o2}} - \frac{\theta_5^o}{(\theta_3^o\theta_4^o-\theta_5^{o2})^2}$$

$$\times \left\{\theta_4^o(X-\theta_1^o)^2-2\theta_5^o(X-\theta_1^o)(Y-\theta_2^o)+\theta_3^o(Y-\theta_2^o)^2\right\} + \frac{(X-\theta_1^o)(Y-\theta_2^o)}{\theta_3^o\theta_4^o - \theta_5^{o2}} .$$

Hence we can easily see that

$$(4.7) \qquad \int\!\!\int_{-\infty}^{\infty} \underline{\alpha}(x,y)\, f(x,y,\underline{\theta}^o)\, dx\, dy = \underline{0} .$$

Furthermore,

$$\beta_{11}(x,y,\underline{\theta}) = \frac{\partial^2 \log f(x,y,\underline{\theta})}{\partial \theta_1^2} = -\frac{\theta_4}{\theta_3\theta_4-\theta_5^2} ,$$

$$\beta_{12}(x,y,\underline{\theta}) = \beta_{21}(x,y,\underline{\theta}) = \frac{\partial^2 \log f(x,y,\underline{\theta})}{\partial \theta_1 \partial \theta_2} = -\frac{\theta_5}{\theta_3\theta_4-\theta_5^2} ,$$

$$\beta_{13}(x,y,\underline{\theta}) = \beta_{31}(x,y,\underline{\theta}) = \frac{\partial^2 \log f(x,y,\underline{\theta})}{\partial \theta_1 \partial \theta_3}$$

$$= -\frac{\theta_4}{(\theta_3\theta_4-\theta_5)^2}\left[\theta_4(x-\theta_1)-\theta_5(y-\theta_2)\right] ,$$

$$\beta_{14}(x,y,\underline{\theta}) = \beta_{41}(x,y,\underline{\theta}) = \frac{\partial^2 \log f(x,y,\underline{\theta})}{\partial \theta_1 \partial \theta_4}$$

$$= -\frac{\theta_3}{(\theta_3\theta_4-\theta_5^2)^2}\left[\theta_4(x-\theta_1)-\theta_5(y-\theta_2)\right] + \frac{x-\theta_1}{\theta_3\theta_4-\theta_5^2} \quad ,$$

$$\beta_{15}(x,y,\underline{\theta}) = \beta_{51}(x,y,\underline{\theta}) = \frac{\partial^2 \log f(x,y,\underline{\theta})}{\partial \theta_1 \partial \theta_5}$$

$$= \frac{2\theta_5}{(\theta_3\theta_4-\theta_5^2)^2}\left[\theta_4(x-\theta_1)-\theta_5(y-\theta_2)\right] - \frac{y-\theta_2}{\theta_3\theta_4-\theta_5^2} \quad ,$$

$$\beta_{22}(x,y,\underline{\theta}) = -\frac{\theta_3}{\theta_3\theta_4-\theta_5^2} \quad ,$$

$$\beta_{23}(x,y,\underline{\theta}) = \beta_{32}(x,y,\underline{\theta}) = \frac{\partial^2 \log f(x,y,\underline{\theta})}{\partial \theta_2 \partial \theta_3}$$

$$= -\frac{\theta_4}{(\theta_3\theta_4-\theta_5^2)^2}\left[-\theta_5(x-\theta_1)+\theta_3(y-\theta_2)\right] - \frac{y-\theta_2}{\theta_3\theta_4-\theta_5^2} \quad ,$$

$$\beta_{24}(x,y,\underline{\theta}) = \beta_{42}(x,y,\underline{\theta}) = \frac{\partial^2 \log f(x,y,\underline{\theta})}{\partial \theta_2 \partial \theta_4}$$

$$= -\frac{\theta_3}{(\theta_3\theta_4-\theta_5^2)^2}\left[-\theta_5(x-\theta_1)+\theta_3(y-\theta_2)\right] \quad ,$$

$$\beta_{25}(x,y,\underline{\theta}) = \beta_{52}(x,y,\underline{\theta}) = \frac{\partial^2 \log f(x,y,\underline{\theta})}{\partial \theta_2 \partial \theta_5}$$

$$= \frac{2\theta_5}{(\theta_3\theta_4-\theta_5^2)^2}\left[-\theta_5(x-\theta_1)+\theta_3(y-\theta_1)\right] - \frac{x-\theta_1}{\theta_3\theta_4-\theta_5^2} \quad ,$$

$$\beta_{33}(x,y,\underline{\theta}) = \frac{\partial^2 \log f(x,y,\underline{\theta})}{\partial \theta_3^2}$$

$$= \frac{\theta_4^2}{2(\theta_3\theta_4-\theta_5^2)^2} - \frac{\theta_4^2}{(\theta_3\theta_4-\theta_5^2)^2}\left[\theta_4(x-\theta_1)^2-2\theta_5(x-\theta_1)(y-\theta_2)+\theta_3(y-\theta_2)^2\right]$$

$$+ \frac{\theta_4}{(\theta_3\theta_4-\theta_5^2)^2}(y-\theta_2)^2 \quad ,$$

$$\beta_{34}(x,y,\underline{\theta}) = \beta_{43}(x,y,\underline{\theta}) = \frac{\partial^2 \log f(x,y,\underline{\theta})}{\partial \theta_3 \partial \theta_4}$$

$$= -\frac{1}{2(\theta_3\theta_4-\theta_5^2)} + \frac{\theta_3\theta_4}{2(\theta_3\theta_4-\theta_5^2)^2} + \left[\frac{1}{2(\theta_3\theta_4-\theta_5^2)^2} - \frac{\theta_3\theta_4}{(\theta_3\theta_4-\theta_5^2)^3}\right]$$

$$\times \left[\theta_4(x-\theta_1)^2-2\theta_5(x-\theta_1)(y-\theta_2)+\theta_3(y-\theta_2)^2\right]$$

$$+ \frac{\theta_4}{2(\theta_3\theta_4-\theta_5^2)^2}(x-\theta_1)^2 + \frac{\theta_3}{2(\theta_3\theta_4-\theta_5^2)^2}(y-\theta_2)^2 \quad ,$$

$$= \frac{\theta_5^2}{2(\theta_3\theta_4-\theta_5^2)^2} - \frac{\theta_3\theta_4+\theta_5^2}{2(\theta_3\theta_4-\theta_5^2)^3}\left[\theta_4(x-\theta_1)^2-2\theta_5(x-\theta_1)(y-\theta_2)+\theta_3(y-\theta_2)^2\right]$$

$$+ \frac{1}{2(\theta_3\theta_4-\theta_5^2)^2}\left[\theta_4(x-\theta_1)^2+\theta_3(y-\theta_2)^2\right] \quad ,$$

$$\beta_{35}(x,y,\underline{\theta}) = \beta_{53}(x,y,\underline{\theta}) = \frac{\partial^2 \log f(x,y,\underline{\theta})}{\partial \theta_3 \partial \theta_5}$$

$$= -\frac{1}{(\theta_3\theta_4-\theta_5^2)^2}\left[\theta_4\theta_5-2\theta_4\theta_5\,\frac{\theta_4(x-\theta_1)^2-2\theta_5(x-\theta_1)(y-\theta_2)+\theta_3(y-\theta_2)^2}{\theta_3\theta_4-\theta_5^2}\right.$$

$$\left. + \theta_4(x-\theta_1)(y-\theta_2) + \theta_5(y-\theta_2)^2\right] \quad ,$$

$$\beta_{44}(x,y,\underline{\theta}) = \frac{\partial^2 \log f(x,y,\underline{\theta})}{\partial \theta_4^2}$$

$$= \frac{\theta_3}{2(\theta_3\theta_4-\theta_5^2)^2} \left[\theta_3+2(x-\theta_1)^2-\theta_3\theta_4(x-\theta_1)^2+2\theta_3\theta_5(x-\theta_1)(y-\theta_2) \right.$$

$$\left. +\theta_5^2(y-\theta_2)^2 \right] ,$$

$$\beta_{45}(x,y,\underline{\theta}) = \beta_{54}(x,y,\underline{\theta}) = \frac{\partial^2 \log f(x,y,\underline{\theta})}{\partial \theta_4 \partial \theta_5}$$

$$= -\frac{1}{(\theta_3\theta_4-\theta_5^2)^2} \left[\theta_3\theta_5+\theta_3(x-\theta_1)(y-\theta_2)+\theta_5(x-\theta_1)^2 \right.$$

$$\left. -\frac{\theta_3\theta_5}{\theta_3\theta_4-\theta_5^2} \left\{ \theta_4(x-\theta_1)^2-2\theta_5(x-\theta_1)(y-\theta_2)+\theta_5(y-\theta_2)^2 \right\} \right] ,$$

$$\beta_{55}(x,y,\underline{\theta}) = \frac{\partial^2 \log f(x,y,\underline{\theta})}{\partial \theta_5^2}$$

$$= \frac{1}{(\theta_3\theta_4-\theta_5^2)^2} \left[\theta_3\theta_4+\theta_5^2 + 6\theta_5(x-\theta_1)(y-\theta_2)-\theta_4(x-\theta_1)^2-\theta_3(y-\theta_2)^2 \right.$$

$$\left. -\frac{4\theta_5^2}{\theta_3\theta_4-\theta_5^2} \left\{ \theta_4(x-\theta_1)^2-2\theta_5(x-\theta_1)(y-\theta_2)+\theta_3(y-\theta_2)^2 \right\} \right] .$$

Thus, on account of the relations

$$v_{ik} = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \beta_{ik}(x,y,\underline{\theta}^o)\, f(x,y,\underline{\theta}^o)\, dx\, dy ,$$

we obtain the following:

$$v_{11} = \frac{\theta_4^o}{\theta_3^o\theta_4^o-\theta_5^{o2}} , \quad v_{12} = \frac{-\theta_5^o}{\theta_3^o\theta_4^o-\theta_5^{o2}} , \quad v_{13} = 0 , \; v_{14} = 0 , \; v_{15} = 0 ,$$

(4.8)

$$v_{22} = \frac{\theta_3^o}{\theta_3^o\theta_4^o-\theta_5^{o2}} , \quad v_{23} = 0 , \; v_{24} = 0 , \; v_{25} = 0 ,$$

and

$$v_{33} = \frac{1}{2} \frac{\theta_4^{o2}}{(\theta_3^o\theta_4^o-\theta_5^{o2})^2} , \quad v_{34} = \frac{1}{2}\frac{\theta_5^{o2}}{(\theta_3^o\theta_4^o-\theta_5^{o2})^2} , \quad v_{35} = - \frac{\theta_4^o\theta_5^o}{(\theta_3^o\theta_4^o-\theta_5^{o2})^2} ,$$

(4.9)

$$v_{44} = \frac{1}{2}\frac{\theta_3^{o2}}{(\theta_3^o\theta_4^o-\theta_5^{o2})^2} , \quad v_{45} = - \frac{\theta_3^o\theta_5^o}{(\theta_3^o\theta_4^o-\theta_5^{o2})^2} ,$$

$$v_{55} = \frac{\theta_3^o\theta_4^o}{(\theta_3^o\theta_4^o-\theta_5^{o2})^2} + \frac{\theta_5^{o2}}{(\theta_3^o\theta_4^o-\theta_5^{o2})^2} .$$

It can be easily checked that

$$(4.10) \qquad \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [\beta(x,y) + \underline{\alpha}(x,y)\cdot\underline{\alpha}'(x,y)] \, f(x,y,\underline{\theta}^o) \, dx \, dy = 0$$

Next we have

$$(4.11) \qquad \gamma_i(x,y,\underline{\theta}) = \frac{\partial \gamma(x,y,\underline{\theta})}{\partial \theta_i} = \alpha_i(x,y,\underline{\theta}) - \alpha_i(x,y,\underline{\theta}^o) - \sum_{k=1}^{5} \beta_{ik}(x,y)(\theta_k-\theta_k^o)$$

$$= \sum_{k=1}^{5} [\beta_{ik}(x,y,\underline{\theta}^*)-\beta_{ik}(x,y,\underline{\theta}^o)](\theta_k-\theta_k^o) ,$$

where $\underline{\theta}^*$ stands for a certain point on the segment joining $\underline{\theta}$ and $\underline{\theta}^o$.

By means of Cauchy's inequality we get

$$\left|\gamma_i(x,y,\underline{\theta})\right| \le \left|\underline{\theta}-\underline{\theta}^o\right| \left[\sum_{k=1}^{5} [\beta_{ik}(x,y,\underline{\theta}^*)-\beta_{ik}(x,y,\underline{\theta}^o)]^2\right]^{1/2},$$

and hence

$$(4.12) \qquad \frac{\left|\gamma_i(x,y,\underline{\theta})\right|}{\left|\underline{\theta}-\underline{\theta}^o\right|^2} \le \left[\frac{\sum_{i=1}^{5}[\beta_{ik}(x,y,\underline{\theta}^*)-\beta_{ik}(x,y,\underline{\theta}^o)]^2}{\sum_{i=1}^{5}(\theta_k-\theta_k^o)^2}\right]^{1/2}.$$

Since, as was shown before, $\beta_{ik}(x,y,\underline{\theta})$ has continuous partial derivatives with respect to $\underline{\theta}$, again by using the mean value theorem and Cauchy's inequality, we obtain

$$(4.13) \qquad \frac{\left|\gamma_i(x,y,\underline{\theta})\right|}{\left|\underline{\theta}-\underline{\theta}^o\right|^2} \le \left[\sum_{k=1}^{5}\sum_{j=1}^{5}(\beta_{ikj}(x,y,\underline{\theta}^{**})^2\right]^{1/2}$$

where $\underline{\theta}^{**}$ stands for a certain point on the segment between $\underline{\theta}^o$ and $\underline{\theta}^*$.

We can find functions

$$(4.14) \qquad \phi_i(x,y) \ge \sum_{k=1}^{5}\sum_{j=1}^{5} \max_{\underline{\theta}} \left|\beta_{ikj}(x,y,\underline{\theta})\right|, \quad i = 1,2,\ldots,5 ,$$

such that $\phi_i(x,y)\, f(x,y,\underline{\theta}^o)$ are integrable, and

$$(4.15) \qquad \frac{\left|\gamma_i(x,y,\underline{\theta})\right|}{\left|\underline{\theta}-\underline{\theta}^o\right|^2} \le \phi_i(x,y) \quad .$$

Next we shall show that the condition (2.3) is also satisfied. For this purpose we must introduce the new parameter ϕ such that

$$\underline{\theta} = B\,\underline{\phi} \tag{4.16}$$
$$B'\,v\,B = D_{\xi}\ ,\quad B'\,B = I\ ;$$

then

$$\gamma^*(x,y,\underline{\phi}) \equiv \gamma(x,y,B\underline{\phi}) = \log \frac{f(x,y,B\underline{\phi})}{f(x,y,B\underline{\phi}^o)} - (\underline{\phi}-\underline{\phi}^o)'B'\underline{\alpha}(x,y) \tag{4.17}$$
$$- \tfrac{1}{2}(\underline{\phi}-\underline{\phi}^o)'B'\beta(x,y)B(\underline{\phi}-\underline{\phi}^o)\ ,$$

and

$$\gamma_i^*(x,y,\underline{\phi}) = \sum_{j=1}^{5} \frac{\partial\gamma(x,y,\underline{\theta})}{\partial\theta_j}\frac{\partial\theta_j}{\partial\phi_i} = \sum_{j=1}^{5}\frac{\partial\gamma(x,y,\underline{\theta})}{\partial\theta_j}\, b_{ji} \tag{4.18}$$
$$\gamma_{ik}^*(x,y,\underline{\phi}) = \sum_{j=1}^{5}\sum_{j'=1}^{5}\frac{\partial^2\gamma(x,y,\underline{\theta})}{\partial\theta_j\partial\theta_{j'}}\, b_{ji}b_{j'k}$$

and

$$\frac{\partial^2\gamma(x,y,\underline{\theta})}{\partial\theta_j\partial\theta_{j'}} = \gamma_{jj'}(x,y,\underline{\theta}) = \beta_{jj'}(x,y,\underline{\theta}) - \beta_{jj'}(x,y,\underline{\theta}^o) \tag{4.19}$$
$$\left|\gamma_{jj'}(x,y,\underline{\theta})\right| \le \left|\underline{\theta}-\underline{\theta}^o\right|\phi_{jj'}(x,y)\ ,$$

where $\phi_{jj'}(x,y)$ was defined in (4.14).

Thus we get

$$\int_{-\infty}^{\infty} \gamma_{ik}^*(x,y,\underline{\phi})\, f^*(x,y,\underline{\phi}^o)\,dx\,dy = \sum_{jj'=1}^{5} b_{ji}b_{j'i}$$
$$\times \int_{-\infty}^{\infty}\!\!\int \frac{\partial^2\gamma(x,y,\underline{\theta})}{\partial\theta_j\partial\theta_{j'}}\, f(x,y,\theta^o)\,dxdy\ ,$$

and hence

$$(4.20) \qquad K_{ik}(\underline{\phi}) \equiv \left| \int\int_{-\infty}^{\infty} \gamma^*_{ik}(x,y,\underline{\phi})\, f^*(x,y,\underline{\phi}^0)\, dx\, dy \right|$$

$$\leq \left| \underline{\theta} - \underline{\theta}^0 \right| \sum_{\jmath\jmath'=1}^{5} b_{\jmath i} b_{\jmath' k} \int\int_{-\infty}^{\infty} \phi_{\jmath\jmath'}(x,y) f(x,y,\underline{\theta}^0) dxdy \;.$$

We can choose $\left| \underline{\theta} - \underline{\theta}^0 \right| = \left| \underline{\phi} - \underline{\phi}^0 \right|$ sufficiently small so that

$$(4.21) \qquad K = \max_{i,k} \max_{\underline{\phi}} K_{ik}(\underline{\phi}) < \xi \quad ,$$

where ξ denotes the minimum characteristic root os V.

Thus we have shown that in the case which was mentioned in section 3, all conditions required by Doob and also the condition (2.3) are satisfied.

5. Concluding Remarks. The exact role of the consistency condition in this whole scheme, the precise relation between Doob's [2] and Wald's [5] papers and also between the present paper and Wald's paper [5] are worth a more careful consideration which is now under way. The authors hope to be able to discuss these in a later communication.

REFERENCES

1. Cramér, H., Mathematical Methods of Statistics, Princeton University Press, 1946.

   Cramér, H., Random Variables and Probability Distributions, Cambridge Tracts in Mathematics and Mathematical Physics No. 26, 1937.

2. Doob, J. L., "Probability and Statistics," Trans. Amer. Math. Soc., Vol. 36, pp. 759-775.

3. Moustafa, M. D., "Testing of Hypotheses on a Multivariate Population, Some of the Variates being Continuous and the Rest Categorical," U. N. C. Institute of Statistics Mimeo. Series No. 179, 1957.

4. Ogawa, J., "A limit theorem of Cramér and its generalizations," U. N. C. Institute of Statistics Mimeo. Series No. 163, 1957.

5. Wald, A., "Tests of statistical hypotheses concerning several parameters when the number of observations is large," Trans. Amer. Math. Soc., Vol. 54 (1943), pp. 426-482.

6. Wilks, S. S., "The large-sample distribution of the likelihood ratio for testing composite hypothesis," Ann. Math. Stat., Vol. 9 (1938), pp. 60-62.