The Annals of Probability
1981, Vol. 9, No. 5. 818–830

SOME CONDITIONAL LIMIT THEOREMS IN EXPONENTIAL FAMILIES

BY LARS HOLST[1]

Stanford University

Consider a sample from a p-dimensional exponential family and a random vector whose distribution is the same as that of the sample given a sufficient statistic. General limit theorems for functions of sum-type for such random vectors are obtained; simple sums, linear combinations, m-dependent sums, and U-statistics are considered. The results are illustrated by some examples.

1. Introduction. It is quite common in probability theory and statistics to come across a random variable (rv) whose distribution is the same as that of a sum of rv's conditioned on a sum of i.i.d. rv's. In this paper the asymptotic behavior of such rv's are studied in general situations for the exponential family. This situation covers many problems which have been studied separately. The results obtained below unify and extend those proved before by special methods for each case. Some general results for discrete situations are obtained in Holst (1979a). General questions of convergence of conditional distributions were studied by Steck (1957) and Chibisov (1972). Related problems are investigated in von Bahr and Svensson (1978), Barndorff-Nielsen and Cox (1979), and Michel (1979).

The organization of the paper is as follows. In Section 2 basic assumptions and notations are introduced, some facts for the exponential family are given, and a formula for a conditional characteristic function is obtained. Section 3 contains the symmetric situations with the sums of the simple type $\sum_{k=1}^{n} f_n(X_k)$. Linear combinations are considered in Section 4. A limit theorem for random sums of the form $\sum_{k=1}^{n-m} f(X_k, \cdots, X_{k+m})$ is given in Section 5. Conditional U-statistics both for the one and two sample cases are considered in Section 6. Examples and illustrations of the results obtained are given in Section 7.

2. General assumptions and the exponential family. Throughout this paper the assumptions and notations of this section will be used. First, recall some facts on exponential families. For proofs and a complete treatment see Barndorff-Nielsen (1978).

Let Ω be a sample space and denote by $X, X', X_1, X_1', \cdots$, i.i.d. rv's on Ω with distribution in a regular p-dimensional exponential family. That is, the probability measure on Ω is of the form

$$dP_\theta(x) = \exp(\theta' t(x) - K(\theta))\, d\mu(x),$$

where μ is a given measure on Ω and the natural parameter space

$$\Theta = \{\theta; |K(\theta)| < \infty\} \subset R^p$$

is an *open* set (all vectors are column vectors and $'$ denotes transponation). The cumulant generating function of $t(X)$ is

$$\ln E(\exp(u' t(X))) = K(u + \theta) - K(\theta).$$

Received July 1980.

[1] This research has been supported by a fellowship from the *American-Scandinavian Foundation* and grants from the *Swedish Natural Science Research Council.*

AMS 1970 *Subject Classifications.* Primary 60F05; secondary 62E20.

Key Words and Phrases. Conditional limit theorems, exponential families, U-statistics, m-dependence, urn models, gamma distribution, spacings.

For $\theta \in \Theta$ all moments of $t(X)$ exist, and

$$E_\theta t(X) = K'(\theta),$$

$$\mathrm{Var}_\theta\, t(X) = K''(\theta) > 0.$$

Let G be the minimal abelian subgroup in R^p of support of $t(X)$. It is assumed that either

$$G = R^p, \text{ the continuous case,}$$

or

$$G = Z^p, \text{ the lattice case.}$$

Set $H = R^p$ in the continuous case, and $H = (-\pi, \pi]^p$ in the lattice case. In the continuous case it is further assumed that

$$\int_{R^p} | E_\theta(\exp(i\eta' t(X))) |^{n_0} \, d\eta < \infty,$$

for all $\theta \in \Theta$ and for some n_0 (independent of θ). The assumption is equivalent to that the random vector

$$T_n = t(X_1) + \cdots + t(X_n)$$

has a bounded density with respect to Lebesgue measure in R^p for n sufficiently large; see, e.g., Bhattacharya and Rao (1976), Theorem 19.1. The density of T_n will in both the lattice and the continuous case be denoted by $f_{n,\theta}(t)$. It will be assumed that n is so large that it exists.

For the sample $\mathbf{X} = (X_1, \cdots, X_n)$ the likelihood equation can be written

$$E_\theta t(X) = (t(X_1) + \cdots + t(X_n))/n = \bar{T}_n.$$

It has a unique solution $\theta = \hat{\theta}$ if and only if $\bar{T}_n \in \Theta$. In typical applications $P_\theta(\bar{T}_n \in \Theta) = 1$ for all $\theta \in \Theta$, at least for n sufficiently large. For a full discussion on this matter see Barndorff-Nielsen (1978), Section 9.3. In the following, $t_n \in R^p$ denotes a possible value of T_n with $\bar{t}_n = t_n/n \in \Theta$. Thus there is a unique $\theta_n \in \Theta$ such that

$$E_{\theta_n} t(X) = \bar{t}_n = t_n/n.$$

For the sample $\mathbf{X} = (X_1, \cdots, X_n)$ a sufficient statistic is $T_n = \sum_{j=1}^n t(X_j)$. Thus $\mathscr{L}(\mathbf{X} \mid T_n = t_n)$ defines a unique distribution on Ω^n. Let $\mathbf{Y} = (Y_1, \cdots, Y_n)$ be a rv on Ω^n having this distribution, i.e.,

$$\mathscr{L}(\mathbf{Y}) = \mathscr{L}(\mathbf{X} \mid T_n = t_n).$$

For any measurable function u on Ω^n having values in R^q we have

$$\mathscr{L}(u(\mathbf{Y})) = \mathscr{L}(u(\mathbf{X}) \mid T_n = t_n).$$

For $n \ge n_0$, $\xi \in R^q$ and $\eta \in R^p$

$$\gamma_{\theta,n}(\xi, \eta) = E_\theta(\exp(i\xi' u(\mathbf{X}) + i\eta' T_n)) = \int_G \exp(i\eta' t) E(\exp(i\xi' u(\mathbf{X})) \mid T_n = t) f_{n,\theta}(t) \, d\nu(t),$$

where ν is the Lebesgue measure for $G = R^p$ and the counting measure for $G = Z^p$. Applying Fourier's inversion formula we obtain the ch.f. (characteristic function) of $u(\mathbf{Y})$. Thus we have:

PROPOSITION 2.1. *If*

$$\int_H | \gamma_{\theta,n}(\xi, \eta) | \, d\eta < \infty, \qquad \theta \in \Theta,$$

then

$$E(\exp(i\xi' u(\mathbf{Y}))) = (2\pi)^{-p} \int_H \exp(-i\eta' t_n)\gamma_{\theta,n}(\xi, \eta)\, d\eta / f_{n,\theta}(t_n).$$

Note that the distribution of $\mathbf{Y}$ does not involve θ. Thus, this parameter can be chosen arbitrarily in the above representation. It will be seen below that $\theta = \theta_n$ with $E_{\theta_n} t(X) = \bar{t}_n$ is a natural choice.

For a general discussion on conditional characteristic functions, see Zabell (1979).

3. Limit theorems for simple sums. In this section the asymptotic behavior of distributions of the form

$$\mathscr{L}(u_n(\mathbf{Y})) = \mathscr{L}(\textstyle\sum_{j=1}^n f_n(Y_j, t_n)) = \mathscr{L}(\sum_{j=1}^n f_n(X_j, t_n) \mid \sum_{j=1}^n t(X_j) = t_n)$$

is studied when $n \to \infty$. Here $\{f_n\}$ is a given sequence of real valued measurable functions. The main results are Theorem 3.1 and the Corollaries 3.5 and 3.6.

THEOREM 3.1. *Let* $\mathbf{X} = (X_1, \cdots, X_n)$, $\mathbf{Y} = (Y_1, \cdots, Y_n)$, $T_n = \sum_{j=1}^n t(X_j)$ *and* θ_n *be defined as in Section* 2, *and let the general assumptions of Section* 2 *be satisfied. Set*

$$C_n = \mathrm{Var}_{\theta_n} t(X) > 0.$$

Suppose that

(A1) *for* $n', n \to \infty$ *with* $n'/n \to \alpha$, $0 < \alpha < 1$,

$$|nC_n|^{1/2} \int_H |E_{\theta_n}(\exp(i\eta' t(X)))|^{n'}\, d\eta \to \int_{R^p} \exp(-\alpha\eta'\eta/2)\, d\eta,$$

(A2) *for* $n \to \infty$

$$\mathscr{L}_{\theta_n}\left(\sum_{j=1}^n \begin{pmatrix} f_n(X_j, t_n) \\ (nC_n)^{-1/2}(t(X_j) - \bar{t}_n) \end{pmatrix}\right) \to \mathscr{L}\begin{pmatrix} U \\ V \end{pmatrix}$$

where the ch.f. *of* (U, V) *is*

$$E(\exp(i\xi U + i\eta' V)) = h(\xi)\exp(-(A\xi^2 + 2\xi B'\eta + \eta'\eta)/2),$$

with $A, \xi \in R$, $B, \eta \in R^p$ *and where* $h(\xi)$ *is an infinitely divisible* ch.f. *without normal component. Then, when* $n \to \infty$,

$$\mathscr{L}(\textstyle\sum_{j=1}^n f_n(Y_j, t_n)) = \mathscr{L}(\sum_{j=1}^n f_n(X_j, t_n) \mid T_n = t_n) \to \mathscr{L}(W),$$

where the ch.f. *of* W *is*

$$E(\exp(i\xi W)) = h(\xi)\exp(-(A - B'B)\xi^2/2).$$

PROOF. A method due to LeCam (1958) will be used for proving the assertion. For

$$u_{n,n'}(\mathbf{x}) = \textstyle\sum_{j=n'+1}^n f_n(x_j, t_n),$$

Proposition 2.1 yields for $\xi \in R$ that

$$E(\exp(i\xi u_{n,n'}(\mathbf{Y}))) = (2\pi)^{-p} \int_H \exp(-i\eta' t_n) E_{\theta_n}(\exp(i\eta' \textstyle\sum_{j=1}^{n'} t(X_j)))$$

$$\cdot E_{\theta_n}(\exp(i\eta' \textstyle\sum_{j=n'+1}^n t(X_j) + i\xi u_{n,n'}(\mathbf{X})))\, d\eta / f_{n,\theta_n}(t_n)$$

$$= (2\pi)^{-p/2} \int_{(nC_n)^{1/2}H} E_{\theta_n}(\exp(i\eta'(nC_n)^{-1/2} \textstyle\sum_{j=1}^{n'} (t(X_j) - \bar{t}_n)))$$

$$\cdot \gamma_{\theta_n,n,n'}(\xi, \eta)\, d\eta / |2\pi n C_n|^{1/2} f_{n,\theta_n}(t_n),$$

where

$$\gamma_{\theta_n,n,n'}(\xi, \eta) = E_{\theta_n}(\exp(i\eta'(nC_n)^{-1/2} \textstyle\sum_{j=n'+1}^{n} (t(X_j) - \bar{t}_n) + i\xi u_{n,n'}(\mathbf{X}))).$$

From Assumption A1 it follows by Proposition 3.2 below that

$$|2\pi nC_n|^{1/2} f_{n,\theta_n}(t_n) \to 1, \qquad n \to \infty.$$

By Assumption A2 we have

$$\gamma_{\theta_n,n,n'}(\xi, \eta) \to (E(\exp(i\xi U + i\eta' V)))^{1-\alpha}.$$

By Assumption A1 and the extended form of Lebesgue's dominated convergence theorem, see Rao (1973), page 136, it follows that

$$E(\exp(i\xi u_{n,n'}(\mathbf{Y}))) \to g_\alpha(\xi) = (2\pi)^{-p/2} \int_{R^p} \exp(-\alpha\eta'\eta/2)(h(\xi)\exp(-(A\xi^2 + 2\xi B'\eta + \eta'\eta)/2))^{1-\alpha}\, d\eta.$$

Now

$$g_\alpha(\xi) \to 1, \qquad \alpha \uparrow 1,$$

$$g_\alpha(\xi) \to h(\xi)\exp(-(A - B'B)\eta^2/2), \qquad \alpha \downarrow 0.$$

Thus, by the argument of LeCam (1958), page 14, see also Billingsley (1968), page 25, it follows that in fact

$$E(\exp(i\xi \textstyle\sum_{j=1}^{n} f_n(Y_j))) \to h(\xi)\exp(-(A - B'B)\xi^2/2), \qquad n \to \infty,$$

which proves the assertion.

PROPOSITION 3.2. *Suppose that Assumption* A1 *of Theorem* 3.1 *holds and that*

(A3) $$\mathscr{L}_{\theta_n}((nC_n)^{-1/2} \textstyle\sum_{j=1}^{n} (t(X_j) - \bar{t}_n)) \to N(\mathbf{0}, I).$$

Then,

$$|2\pi nC_n|^{1/2} f_{n,\theta_n}(t_n) \to 1, \qquad n \to \infty.$$

PROOF. Fourier's inversion formula

$$f_{n,\theta_n}(t_n) = (2\pi)^{-p} \int_H \exp(-i\eta' t_n)(E_{\theta_n}(\exp(i\eta' t(X))))^n\, d\eta$$

and changing coordinates in this integral gives

$$|2\pi nC_n|^{1/2} f_{n,\theta_n}(t_n) = (2\pi)^{-p/2} \int_{(nC_n)^{1/2}H} E_{\theta_n}(\exp(i\eta'(nC_n)^{-1/2} \textstyle\sum_{j=1}^{n} (t(X_j) - \bar{t}_n)))\, d\eta.$$

By Assumption A1 and Lebesgue's theorem the last integral converges to

$$(2\pi)^{-p/2} \int_{R^p} \exp(-\eta'\eta/2)\, d\eta = 1,$$

which proves the assertion.

In a similar way one obtains:

PROPOSITION 3.3. *Set*

$$T_{n'}^{s} = \textstyle\sum_{j=1}^{n'} (t(X_j) - t(X_j')).$$

If $T_{n'}^{s}$ *has a bounded density* $f_{n',\theta}^{s}(t)$ *and* A3 *holds, then the Assumption* A1 *is equivalent* *to*

$$\text{(A4)} \qquad |2\pi n C_n|^{1/2} f^s_{n',\theta_n}(\mathbf{0}) \to (\alpha/2)^{p/2},$$

when $n, n' \to \infty$ *such that* $n'/n \to \alpha$.

REMARK. In many cases it is easier to verify A4 than A1.

PROPOSITION 3.4. *Suppose that* $\{t_n\}$ *is such that* $\theta_n \to \theta_0 \in \Theta$, $n \to \infty$. *Then Assumption* (A1) *holds.*

PROOF. As Θ is open there is a closed ball $K_0 \subset \Theta$ and a number n_0 such that $\theta_n \in K_0$ for $n \geq n_0$. As $\text{Var}_\theta\, t(X) = K''(\theta) > 0$ is a continuous function, there exists a positive definite matrix D_0 such that $C_n = \text{Var}_{\theta_n}\, t(X) \geq D_0 > 0$. Using this uniformity the conventional proof of the local limit theorem with characteristic functions gives the assertion; see, e.g., Feller (1971), pages 516–517.

For most applications the following two corollaries are sufficient.

COROLLARY 3.5. *Suppose that Assumption* A1 *holds and that*

$$\text{(A5)} \qquad \mathscr{L}_{\theta_n}(\textstyle\sum_{j=1}^n f_n(X_j, t_n)) \to \mathscr{L}(U),$$

where U *has no normal component. Then,*

$$\mathscr{L}(\textstyle\sum_{j=1}^n f_n(Y_j, t_n)) \to \mathscr{L}(U).$$

PROOF. From the classical limit theorems for independent rv's it follows that the only possible limits of subsequences of the form

$$\mathscr{L}_{\theta_{n_k}}\left(\sum_{j=1}^{n_k} \begin{pmatrix} f_{n_k}(X_j, t_{n_k}) \\ (n_k C_{n_k})^{-1/2}(t(X_j) - \bar{t}_{n_k}) \end{pmatrix} \right)$$

have ch.f.'s of the type $h(\xi)\exp(-(A\xi^2 + 2\xi B'\eta + \eta'\eta)/2)$, where A and B can depend on the particular subsequence; see LeCam (1958), pages 9–10. But as U has no normal component, $A = 0$ and thus $B = 0$. Therefore, every subsequence has the same limit ch.f. $h(\xi)$. The assertion follows now from Theorem 3.1.

COROLLARY 3.6. *Suppose that* $\{t_n\}$ *is such that* $\theta_n \to \theta_0 \in \Theta$. *Let* f *be a given function with* $\text{Var}_\theta\, f(X) < \infty$ *for all* $\theta \in \Theta$. *Set*

$$A_0 = \text{Var}_{\theta_0} f(X),$$

$$B_0 = \text{Cov}_{\theta_0}(f(X), t(X)),$$

$$C_0 = \text{Var}_{\theta_0} t(X).$$

Then,

$$\mathscr{L}(n^{-1/2} \textstyle\sum_{j=1}^n (f(Y_j) - E_{\theta_n} f(X_j))) \to N(0, A_0 - B_0' C_0^{-1} B_0).$$

PROOF. By Proposition 3.4 the Assumption A1 holds. From the conventional form of the Central Limit Theorem it follows that

$$\mathscr{L}_{\theta_n}\left(n^{-1/2} \sum_{j=1}^n \begin{pmatrix} f(X_j) - E_{\theta_n} f(X_j) \\ t(X_j) - \bar{t}_n \end{pmatrix} \right) \to N\left(\mathbf{0}, \begin{pmatrix} A_0 & B_0' \\ B_0 & C_0 \end{pmatrix}\right),$$

recall $E_{\theta_n} t(X_j) = \bar{t}_n$. Thus, Assumption A2 is satisfied with $h(\xi) \equiv 1$. The assertion now follows from Theorem 3.1.

REMARK. It may be more natural to consider norming by $Ef(Y_j)$ instead of $E_{\theta_n}f(X_j)$. By Portnoy (1977) it follows that

$$Ef(Y_j) = E_{\theta_n}f(X_j) + O(1/n).$$

Thus, $E_{\theta_n}f(X_j)$ can be replaced by $Ef(Y_j) = E(f(X_j) \mid T_n = t_n)$ in the last corollary.

4. A limit theorem for linear combinations.

THEOREM 4.1. *Let $\{a_{jn}; j = 1, \cdots, n, n = 1, 2, \cdots\}$ be a double array of real numbers such that*

$$\sum_{j=1}^n a_{jn} = 0,$$

$$\sum_{j=1}^n a_{jn}^2/n \to 1, \qquad n \to \infty,$$

$$\max_{1\le j\le n} a_{jn}^2/n \to 0, \qquad n \to \infty.$$

Let the sequence $\{t_n\}$ be such that $\theta_n \to \theta_0 \in \Theta$, $n \to \infty$. Suppose that there exists n_0 such that for all $\xi \in R^p$ and real numbers $\delta_1, \cdots \delta_{n_0}$

$$\int_H \left| \exp\left(\sum_{j=1}^{n_0} K(\theta + i(\eta + \delta_j\xi))\right) \right| d\eta < \infty.$$

Then

$$\mathscr{L}\left(n^{-1/2} \sum_{j=1}^n a_{jn}t(Y_j)\right) = \mathscr{L}\left(n^{-1/2} \sum_{j=1}^n a_{jn}t(X_j) \mid \sum_{j=1}^n t(X_j) = t_n\right) \to N(\mathbf{0}, \operatorname{Var}_{\theta_0} t(X)), \ n \to \infty.$$

PROOF. Consider for $\xi, \eta \in R^p$ the characteristic function

$$\begin{aligned}\exp(-i\eta' t_n/n^{1/2})\gamma_{\theta_n,n}(\xi, \eta/n^{1/2}) &= E_{\theta_n}\Big(\exp\Big(i\xi' \sum_{j=1}^n a_{jn}t(X_j)n^{1/2} \\ &\quad + i\eta' \sum_{j=1}^n (t(X_j) - \bar{t}_n)/n^{1/2}\Big)\Big) \\ &= \exp\Big(\sum_{j=1}^n (K(\theta_n + i(a_{jn}\xi + \eta)/n^{1/2}) - K(\theta_n) \\ &\quad - i(a_{jn}\xi + \eta)'K'(\theta_n)/n^{1/2})\Big)\end{aligned}$$

which by the assumptions is integrable in η for $n \ge n_0$. As $\theta_n \to \theta_0 \in \Theta$ it follows for fixed $\eta \in R^p$ that, when $n \to \infty$,

$$\exp(-i\eta' t_n/n^{1/2})\gamma_{\theta_n,n}(\xi, \eta/n^{1/2}) \to \exp(-(\xi' K''(\theta_0)\xi + \eta' K''(\theta_0)\eta)/2).$$

By Proposition 2.1 we have

$$\begin{aligned}&E\Big(\exp\Big(i\xi' \sum_{j=1}^n a_{jn}t(Y_j)/n^{1/2}\Big)\Big) \\ &\qquad = (2\pi)^{-p/2} \int_{n^{1/2}H} \exp(-i\eta' t_n/n^{1/2})\gamma_{\theta_n,n}(\xi, \eta/n^{1/2})\, d\eta/(2\pi n)^{p/2} f_{n,\theta_n}(t_n).\end{aligned}$$

Propositions 3.2 and 3.4 yield

$$(2\pi n)^{p/2} f_{n,\theta_n}(t_n) \to | K''(\theta_0) |^{-1/2}.$$

Hence the assertion is proved if

$$\begin{aligned}&(2\pi)^{-p/2} | K''(\theta_0) |^{1/2} \int_{n^{1/2}H} \exp(-i\eta' t_n/n^{1/2})\gamma_{\theta_n,n}(\xi, \eta/n^{1/2})\, d\eta \\ &\qquad \to (2\pi)^{-p/2} | K''(\theta_0) |^{1/2} \int_{R^p} \exp(-(\xi' K''(\theta_0)\xi + \eta' K''(\theta_0)\eta)/2)\, d\eta \\ &\qquad = \exp(-\xi' K''(\theta_0)\xi/2), \quad n \to \infty.\end{aligned}$$

As the integrand of the left hand side converges pointwise to that of the right hand side, it only remains to validate the interchange of integration and taking limits. This can be done in the standard way by splitting the domain of integration $n^{1/2}H$ into different parts and considering each part separately; for example, see Feller (1971), page 516 and Theorem 3 of Holst (1979a).

5. A limit theorem for m-dependence.

THEOREM 5.1. *Let f be real valued and measurable with*

$$\mathrm{Var}_\theta f(X_1, \cdots, X_m) < \infty, \qquad \textit{all } \theta \in \Theta.$$

Suppose that $\{t_n\}$ is such that $\theta_n \to \theta_0 \in \Theta$, $n \to \infty$.
Then

$$\mathscr{L}(n^{-1/2} \textstyle\sum_{j=1}^{n-m+1} (f(Y_j, \cdots, Y_{j+m-1}) - \mu_n)) \to N(0, \sigma_0^2),$$

where

$$\mu_n = E_{\theta_n} f(X_1, \cdots, X_m)$$

$$\sigma_0^2 = A_0 - B_0' C_0^{-1} B_0$$

with

$$A_0 = \mathrm{Var}_{\theta_0} f(X_1, \cdots, X_m) + 2 \textstyle\sum_{j=2}^{m} \mathrm{Cov}_{\theta_0} (f(X_1, \cdots, X_m), f(X_j, \cdots, X_{j+m-1}))$$

$$B_0 = \mathrm{Cov}_{\theta_0} (f(X_1, \cdots, X_m), t(X_1) + \cdots + t(X_m))$$

$$C_0 = \mathrm{Var}_{\theta_0} t(X_1).$$

PROOF. By the central limit theorem for m-dependent sequences it follows for $n', n \to \infty$ with $n'/n \to \alpha > 0$

$$\mathscr{L}_{\theta_n}\left(n^{-1/2} \textstyle\sum_{j=1}^{n'-m} \begin{pmatrix} f(X_j, \cdots, X_{j+m-1}) - \mu_n \\ t(X_j) - \bar{t}_n \end{pmatrix}\right) \to N\left(\mathbf{0}, \alpha \cdot \begin{pmatrix} A_0 & B_0' \\ B_0 & C_0 \end{pmatrix}\right),$$

where A_0, B_0 and C_0 are given in the assertion. The rest of the proof proceeds in the same way as in Corollary 3.6.

REMARK. Portnoy (1977) proves

$$E_{\theta_n} f(X_1, \cdots, X_m) = E(f(X_1, \cdots, X_m) \mid \textstyle\sum_{j=1}^{n} t(X_j) = t_n) + O(n^{-1}).$$

Thus μ_n can be replaced by $E\, f(Y_1, \cdots, Y_m)$ in Theorem 5.

6. Limit Theorems for U-Statistics.

Let f be a given function with

$$\mathrm{Var}_\theta f(X_1, X_2) < \infty.$$

In this section asymptotic normality of U-statistics with kernel f, conditional on the sufficient statistic $T_n = \sum_{j=1}^{n} t(X_j) = t_n$, is obtained both for a two-sample and a one-sample situation.

Define ψ_{10} and ψ_{01} through

$$\psi_{10}(x) = E_\theta(f(X_1, X_2) \mid X_1 = x) - E_\theta f(X_1, X_2),$$

and

$$\psi_{01}(x) = E_\theta(f(X_1, X_2) \mid X_2 = x) - E_\theta f(X_1, X_2).$$

THEOREM 6.1. *Suppose that* $\{t_n\}$ *satisfies* $\theta_n \to \theta_0 \in \Theta$, $n \to \infty$. *Set*

$$\sigma_{10}^2 = \mathrm{Var}_{\theta_0}\psi_{10}(X),$$

$$\sigma_{01}^2 = \mathrm{Var}_{\theta_0}\psi_{01}(X),$$

$$B_{01} = \mathrm{Cov}_{\theta_0}(\psi_{01}(X), t(X)),$$

$$B_{10} = \mathrm{Cov}_{\theta_0}(\psi_{10}(X), t(X)),$$

$$C_0 = \mathrm{Var}_{\theta_0} t(X).$$

Let $n = n_1 + n_2$ *and* $n_1, n_2 \to \infty$ *such that*

$$n_1/n_2 \to \lambda, \qquad 0 < \lambda < \infty.$$

Then, when $n \to \infty$,

$$\mathscr{L}(n_1^{1/2} \textstyle\sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} (f(X_i, X_j) - E_{\theta_n} f(X_i, X_j))/n_1 n_2 \,|\, \sum_{k=1}^{n} t(X_k) = t_n)$$
$$\to N(0, \sigma_{10}^2 + \lambda\, \sigma_{01}^2 - (\lambda/(1+\lambda))(B_{10} + B_{01})' C_0^{-1}(B_{10} + B_{01})).$$

PROOF. To facilitate notations set

$$V_{ij} = f(X_i, X_j) - E_{\theta_n} f(X_i, X_j).$$

Let $\nu \to \infty$ so $\nu/n_1 \to \alpha$, $0 < \alpha \le 1$. By Proposition 2.1 one obtains for $\xi \in R$

$$E(\exp(i\xi\, n_1^{1/2} \textstyle\sum_{i=\nu+1}^{n_1} \sum_{j=n_1+1}^{n} V_{ij}/n_1 n_2 \,|\, T_n = t_n) = (2\pi)^{-p}(f_{n,\theta_n}(t_n))^{-1} n^{-p/2}$$
$$\cdot \int_{n^{1/2}H} E_{\theta_n}(\exp(i\eta' \textstyle\sum_{k=1}^{\nu} (t(X_k) - \bar{t}_n)/n^{1/2})) H_\nu(\xi, \eta)\, d\eta,$$

where

$$H_\nu(\xi, \eta) = E_{\theta_n}(\exp(i\xi\, n_1^{1/2} \textstyle\sum_{i=\nu+1}^{n_1} \sum_{j=n_1+1}^{n} V_{ij}/n_1 n_2 + i\eta' \sum_{k=\nu+1}^{n} (t(X_k) - \bar{t}_n)/n^{1/2}))\,.$$

As in Lehmann (1975), pages 363–365 one proves that, when $n \to \infty$,

$$\mathrm{Var}_{\theta_n}(\textstyle\sum_{i=\nu+1}^{n_1} \sum_{j=n_1+1}^{n} (f(X_i, X_j) - \psi_{10}(X_i) - \psi_{01}(X_j))/n_1^{1/2}\, n_2) \to 0.$$

By Propositions 3.2 and 3.4

$$(2\pi)^{p/2} |nC_0|^{1/2} f_{n,\theta_n}(t_n) \to 1.$$

Thus in order to obtain convergence of the conditional characteristic function it is sufficient to get convergence of

$$A_\nu(\xi) = (2\pi)^{-p/2} \int_{n^{1/2} C_0^{1/2} H} E_{\theta_n}(\exp(i\eta'(nC_0)^{-1/2} \textstyle\sum_{k=1}^{\nu} (t(X_k) - \bar{t}_n))) \cdot G_\nu(\xi, \eta)\, d\eta,$$

where

$$G_\nu(\xi, \eta) = E_{\theta_n}(\exp(i n_1^{1/2}\, \xi(\textstyle\sum_{i=\nu+1}^{n_1} \psi_{10}(X_i)/n_1 + \sum_{j=n_1+1}^{n} \psi_{01}(X_j)(n_1 - \nu)/n_1 n_2)$$
$$+ i n^{-1/2}\, \eta' \textstyle\sum_{k=\nu+1}^{n} C_0^{-1/2} (t(X_k) - \bar{t}_n))).$$

By the Central Limit Theorem it follows that

$$G_\nu(\xi, \eta) \to G_\alpha(\xi, \eta) = \exp(-((1-\alpha)(\xi^2 \sigma_{10}^2 + 2(\lambda/(1+\lambda))^{1/2} \eta' C_0^{-1/2} B_{10}\xi + (\lambda/(1+\lambda))\eta'\eta)$$
$$+ (1-\alpha)^2 \lambda \xi^2 \sigma_{01}^2 + 2(1-\alpha)(\lambda/(1+\lambda))^{1/2} \eta' C_0^{-1/2} B_{01}\xi + (1+\lambda)^{-1}\eta'\eta)/2).$$

The extended form of Lebesgue's dominated convergence theorem yields

$$A_m(\xi) \to A_\alpha(\xi) = (2\pi)^{-p/2} \cdot \int_{R^p} \exp(-(\lambda\alpha/(1+\lambda))\eta'\eta/2)G_\alpha(\xi, \eta)\, d\eta.$$

Here

$$A_\alpha(\xi) \to A_1(\xi) = 1 \qquad \alpha\uparrow 1,$$

and

$$A_\alpha(\xi) \to A_0(\xi) = (2\pi)^{-p/2} \int_{R^p} G_0(\xi, \eta)\, d\eta$$

$$= \exp(-\xi^2(\sigma_{10}^2 + \lambda\sigma_{01}^2 - (\lambda/(1+\lambda))(B_{10} + B_{01})'\, C_0^{-1}\, (B_{10} + B_{01}))/2)$$

when $\alpha \downarrow 0$. By the argument of LeCam (1958), c.f. the proof of Theorem 3.1; the assertion follows.

Next the one-sample case is considered. Let f be symmetric. Set

$$\psi_0(x) = \psi_{10}(x) = \psi_{01}(x),$$

with $E\, \psi_0(X) = 0$.

THEOREM 6.2. *Suppose that* $\{t_n\}$ *satisfies* $\theta_n \to \theta_0 \in \Theta$, $n \to \infty$. *Set*

$$\sigma_0^2 = \mathrm{Var}_{\theta_0}\psi_0(X),$$

$$B_0 = \mathrm{Cov}_{\theta_0}(\psi_0(X), t(X)),$$

$$C_0 = \mathrm{Var}_{\theta_0}t(X).$$

Then, when $n \to \infty$,

$$\mathscr{L}\left(n^{1/2} \textstyle\sum_{1\le i<j\le n}(f(X_i, X_j) - E_{\theta_n} f(X_i, X_j))\Big/\binom{n}{2} \,\Big|\, \sum_{k=1}^n t(X_k) = t_n\right) \to N(0, 4(\sigma_0^2 - B_0'\, C_0^{-1}\, B_0)).$$

PROOF. Set

$$V_{ij} = f(X_i, X_j) - E_{\theta_n} f(X_i, X_j)$$

and consider for $1 < n_1 < n$ the sums

$$Z_{1n} = n^{1/2} \textstyle\sum_{1\le i<j\le n_1} V_{ij}\Big/\binom{n}{2}$$

$$Z_{2n} = n^{1/2} \textstyle\sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} V_{ij}\Big/\binom{n}{2}$$

$$Z_{3n} = n^{1/2} \textstyle\sum_{n_1<i<j\le n} V_{ij}\Big/\binom{n}{2}.$$

Suppose that $n_1, n \to \infty$ so $n_1/n \to \gamma$, $0 < \gamma < 1$. In the same way as in the previous proof one shows that

$$\mathscr{L}(Z_{3n} \mid T_n = t_n) \to \mathscr{L}(Z_{3\gamma}) = N(0, 4(1-\gamma)^3\, \sigma_0^2 - 4(1-\gamma)^4\, B_0'\, C_0^{-1}\, B_0).$$

and by symmetry we also have

$$\mathscr{L}(Z_{1n} \mid T_n = t_n) \to \mathscr{L}(Z_{1\gamma}) = N(0, 4\gamma^3\, \sigma_0^2 - 4\,\gamma^4\, B_0'\, C_0^{-1}\, B_0).$$

From Theorem 6.1 we obtain

$$\mathscr{L}(Z_{2n} \mid T_n = t_n) \to \mathscr{L}(Z_{2\gamma}) = N(0, 4\gamma(1-\gamma)(\sigma_0^2 - 4\gamma(1-\gamma)B_0'C_0^{-1}B_0)).$$

When $\gamma \downarrow 0$,

$$Z_{1\gamma} \to_D 0,$$

$$Z_{2\gamma} \to_D 0,$$

$$Z_{3\gamma} \to_D N(0, 4(\sigma_0^2 - B_0' C_0^{-1} B_0)).$$

Thus by LeCam (1958)

$$\mathscr{L}(Z_{1n} + Z_{2n} + Z_{3n} \mid T_n = t_n) \to N(0, 4(\sigma_0^2 - B_0' C_0^{-1} B_0)),$$

proving the assertion.

REMARK. $E_{\theta_n} f(X_1, X_2)$ can be replaced by $E(f(X_1, X_2) \mid T_n = t_n)$ in Theorems 6.1 and 6.2.

7. Examples. In this section the results of Sections 3–6 are illustrated by some examples. More detailed analysis of specific applications will be and are given elsewhere.

EXAMPLE 7.1. Let $X_1, X_2, \cdots, X_n$ be i.i.d. with a Poisson distribution with expectation m (corresponding to $\theta = \ln m$). The conditional distribution

$$\mathscr{L}((X_1, \cdots, X_n) \mid \sum_{k=1}^n X_k = r) = \mathscr{L}(Y_1, \cdots, Y_n)$$

is the multinomial $(r, 1/n, \cdots, 1/n)$-distribution. Consider the classical occupancy problem, i.e., the number of empty cells, or

$$\mathscr{L}(\sum_{k=1}^n I(Y_k = 0)) = \mathscr{L}(\sum_{k=1}^n I(X_k = 0) \mid \sum_{k=1}^n X_k = r),$$

Choose $\theta_n = \ln(r/n)$. Now suppose that $r, n \to \infty$ so

$$n\, e^{-r/n} \to a,$$

which implies

$$\mathscr{L}_{\theta_n}(\sum_{k=1}^n I(X_k = 0)) \to \text{Poisson } (a).$$

As the limit has no normal component, Corollary 3.5 yields,

$$\mathscr{L}(\sum_{k=1}^n I(Y_k = 0)) \to \text{Poisson } (a).$$

This result goes back to von Mises and can be proved by direct calculations; see Feller (1968), page 105. In Holst (1979a), (1979b), and (1980b) problems related to discrete situations, in particular urn models, are studied.

EXAMPLE 7.2. Let $X, X_1, \cdots, X_n$ be i.i.d. from a gamma (α, β) distribution, i.e., the density with respect to Lebesgue-measure is

$$f(x) = (\beta^\alpha \Gamma(\alpha))^{-1} x^{\alpha-1} e^{-x/\beta}, \qquad x, \alpha, \beta > 0.$$

Various inference problems concerning the parameters α, β have been studied; see, e.g., Engelhardt and Bain (1977), Gross and Clark (1975), and the references therein. Corollary 3.6 is useful for obtaining approximations of the distributions involved as the following may indicate. The gamma distribution with both parameters unknown defines a regular exponential family with

$$\theta = (\theta_1, \theta_2)' = (\alpha, -1/\beta)', \qquad \theta_1 > 0, \theta_2 < 0,$$

$$t(X) = (\ln X, X)',$$

$$E_\theta t(X) = (\psi(\alpha) + \ln \beta, \alpha\beta)',$$

$$\text{Var}_\theta t(X) = \begin{pmatrix} \psi'(\alpha) & \beta \\ \beta & \alpha\beta^2 \end{pmatrix},$$

where $\psi(\alpha) = \Gamma'(\alpha)/\Gamma(\alpha)$. Furthermore,

$$E_\theta X^2 = \alpha(1+\alpha)\beta^2,$$

$$\text{Var}_\theta X^2 = \alpha(1+\alpha)(6+4\alpha)\beta^4,$$

$$\text{Cov}_\theta(X^2, t(X)) = (2\alpha(1+\alpha)\beta^3, (1+2\alpha)\beta^2)'.$$

Suppose that $t_n/n \to (a_0, b_0)'$, with $a_0 \in R$ and $b_0 > 0$. It is easy to see that this is equivalent to $\theta_n = (\alpha_n, -1/\beta_n) \to \theta_0 = (\alpha_0, -1/\beta_0)$, where $\alpha_0, \beta_0 > 0$. Thus, Corollary 3.6 yields for the dispersion statistic

$$\mathscr{L}(n^{-1/2} \textstyle\sum_{k=1}^n (X_k^2 - \alpha_n(1+\alpha_n)\beta_n^2) \mid T_n = t_n)$$
$$\to N(0, \beta_0^4(2\alpha_0(1+\alpha_0) - \alpha_0/(\alpha_0\,\psi'(\alpha_0) - 1))), \quad n \to \infty.$$

This can also be stated as: the ratio of the sample variance and the predicted variance $\alpha_n\beta_n^2$, given the sufficient statistic, obeys

$$(n-1)^{-1} \textstyle\sum_{k=1}^n (X_k - \bar{X}_n)^2/\alpha_n\beta_n^2 \sim \text{As}N(1, \sigma_n^2/n)$$

where $\bar{X}_n$ is the sample mean and

$$\sigma_n^2 = (2(1+\alpha_n) - (\alpha_n\,\psi'(\alpha_n) - 1)^{-1})/\alpha_n.$$

This can be used to test the fit of the gamma distribution. Barndorff-Nielsen and Cox (1979) gave the corresponding asymptotic result for the *density* function.

EXAMPLE 7.3. An experiment has two outcomes S and F and is repeated until n S's have occurred. Suppose that the probability for S is, $p = 1 - q$ (unknown) until S occurs for the first time, $1 - Kq$ $(0 < K \le 1)$ between the first and second occurrence of S, $1 - K^2q$ between the second and third, etc. Otherwise, trials are independent. Let $X_1, \cdots, X_n$ be the number of F's between the successive S's. The log-likelihood-ratio statistic for testing $H_0{:}K = 1$ against $H_1{:}K < 1$ is essentially the linear combination $\sum_{k=1}^n kX_k$. Under H_0 the X's are independent with a geometric distribution with mean $\mu = q/p$, and $T_n = \sum_{k=1}^n X_k$ is a sufficient statistic. By the strong law of large numbers $\sum_{k=1}^n X_k/n \to \mu$ with probability one. Therefore, let us suppose that $t_n/n \to \mu > 0$. Using Theorem 4.1 it follows that

$$\mathscr{L}((\textstyle\sum_{k=1}^n kX_k - t_n \cdot (n+1)/2)/(nt_n(n+t_n)/3)^{1/2} \mid T_n = t_n) \to N(0, 1), \qquad n \to \infty.$$

The conditional distribution is the same as that of a linear transformation of the Wilcoxon-statistic in the non-parametric two-sample problem (under the null hypothesis); c.f., Holst (1979a), Example 2.

In general, conditional distributions of linear combinations occur in connection with testing a hypothesis against specific local alternatives; c.f., Michel (1979). Thus, Theorem 4.1 is useful to obtain approximations of such test statistics under the null hypothesis. One may also remark that the asymptotic distribution for simple random sampling without replacement from a finite population follows from Theorem 4.1; see Holst (1979a), Example 3.

EXAMPLE 7.4. Let $X_1, \cdots, X_n$ be i.i.d. from an exponential distribution with mean $-1/\theta > 0$. Consider the sum $\sum_{k=1}^{n-1} X_kX_{k+1}$, conditional on $T_n = \sum_{k=1}^n X_k = n$. Simple calculations give

$$E_{-1}(X_1X_2) = 1,$$

$$\text{Var}_{-1}(X_1X_2) = 3,$$

$$\text{Cov}_{-1}(X_1X_2, X_2X_3) = 1,$$

$$\text{Cov}_{-1}(X_1X_2, X_1 + X_2) = 2,$$

$$\text{Var}_{-1}X_1 = 1.$$

Theorem 5.1 yields

$$\mathscr{L}(n^{-1/2} \sum_{k=1}^{n-1}(X_k X_{k+1} - 1) \mid T_n = n) \to N(0, 1).$$

Let $S_1, \cdots, S_n$ be the spacings of $n - 1$ points from a uniform distribution on (0, 1), that is the successive distances between the points including the endpoints 0 and 1. It is well known that

$$\mathscr{L}(X_1, \cdots, X_n \mid \sum_{k=1}^{n} X_k = n) = \mathscr{L}(nS_1, \cdots, nS_n).$$

Hence, the limit above can also be stated as

$$\mathscr{L}(n^{-1/2} \sum_{k=1}^{n-1} (n^2 S_k S_{k+1} - 1)) \to N(0, 1).$$

Further aspects on spacings and their use can be found in Holst (1979c), (1980a) and the references therein.

EXAMPLE 7.5. Let $X_1, \cdots, X_n$ be one-dimensional random variables with the continuous distribution $F(x)$ belonging to an exponential family. For $1 < n_1 < n = n_1 + n_2$ consider the distribution

$$\mathscr{L}(Z_n) = \mathscr{L}(n_1^{1/2} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} (I(X_i < X_j) - ½)/n_1 n_2 \mid T_n = t_n),$$

of the Wilcoxon-Mann-Whitney statistic conditioned on the sufficient statistic. We have

$$\psi_{10}(x) = EI(x < X_1) - ½ = ½ - F(x),$$

$$\psi_{01}(x) = EI(X_1 < x) - ½ = F(x) - ½.$$

Thus, $\psi_{10}(X_1) = -\psi_{01}(X_1)$ is uniformly distributed on $(-½, ½)$ and $B_{10} = -B_{01}$. Theorem 6.1 yields

$$\mathscr{L}(Z_n) \to N(0, (1 + \lambda)/12),$$

when $n_1, n_2 \to \infty$ such that $n_1/n_2 \to \lambda$. Hence, the asymptotic distribution for the conditional statistic is the same as for the unconditional. We may remark that this can also be seen in the following more straightforward manner.

Introduce

$$\mathscr{L}(Y_1, \cdots, Y_n) = \mathscr{L}(X_1, \cdots, X_n \mid T_n = t_n).$$

As the Y's are continuous and exchangeable, the random variable $\sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} (I(Y_i < Y_j) - ½) + n_1(n + 1)/2$ has the same distribution as the sum of the ranks $\{Y_1, \cdots Y_{n_1}\}$ among $\{Y_1, \cdots, Y_n\}$, that is the ordinary (null) distribution for the Wilcoxon statistic.

EXAMPLE 7.6. Let $X_1, \cdots, X_n$ be i.i.d. from an exponential distribution with mean $-1/\theta$. Consider Dini's index dispersion $\sum_{1 \le i \le j \le n} |X_i - X_j|$ conditional on $T_n = \sum_{k=1}^{n} X_k = n$ (corresponding to $\theta_n = -1$). We have

$$E_{-1}|X_1 - X_2| = 1,$$

$$\psi_0(x) = E_{-1}(|X_1 - x| - 1) = 2e^{-x} + x - 2,$$

$$\sigma_0^2 = \mathrm{Var}_{-1}\, \psi_0(X_1) = ⅓,$$

$$B_0 = \mathrm{Cov}_{-1}(X_1, \psi_0(X_1)) = ½,$$

$$C_0 = \mathrm{Var}_{-1}\, X_1 = 1.$$

Theorem 6.2 yields

$$\mathscr{L}\left(n^{1/2} \sum_{1 \le i < j \le n} (|X_i - X_j| - 1)/\binom{n}{2} \,\middle|\, T_n = n\right) \to N(0, ⅓), \qquad n \to \infty.$$

The limit can also be stated as

$$\mathscr{L}\left(n^{1/2} \sum_{1 \le i < j \le n} (n|S_i - S_j| - 1)/\binom{n}{2}\right) \to N(0, \tfrac{1}{3}),$$

when $n \to \infty$, where $S_1, \cdots, S_n$ are spacings; c.f., Example 7.4 above.

REFERENCES

VON BAHR, B., AND SVENSSON, Å. (1978). A conditional limit theorem. Technical report, Inst. för Försäkringsmatematik och Matematisk Statistik, Stockholm University.

BARNDORFF-NIELSEN, O. (1978). *Information and Exponential Families in Statistical Theory.* Wiley, New York.

BARNDORFF-NIELSEN, O., AND COX, D. R. (1979). Edgeworth and saddle-point approximations with statistical applications. *J. Roy. Statist. Soc. Ser. B* **41** 279–312.

BHATTACHARYA, R. N., AND RANGA RAO, R. (1976). *Normal Approximation and Asymptotic Expansions.* Wiley, New York.

BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

CHIBISOV, D. M. (1972). On the normal approximation for a certain class of statistics. *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* **1** 153–174.

ENGELHARDT, M., AND BAIN, L. J. (1977). Uniformly most powerful unbiased tests on the scale parameter of a gamma distribution with nuisance shape parameter. *Technometrics* **19** 77–81.

FELLER, W. (1968). *An Introduction to Probability Theory and Its Applications.* **1.**, 3rd ed., Wiley, New York.

FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications.* **2.**, 2nd ed., Wiley, New York.

GROSS, A. J., AND CLARK, V. A. (1975). *Survival Distributions: Reliability Applications in the Biomedical Sciences.* Wiley, New York.

HOLST, L. (1979a). Two conditional limit theorems with applications. *Ann. Statist.* **7** 551–557.

HOLST, L. (1979b). A unified approach to limit theorems for urn models. *J. Appl. Probability* **16** 154–162.

HOLST, L. (1979c). Asymptotic normality of sum-functions of spacings. *Ann. Probability* **7** 1066–1072.

HOLST, L. (1980a). On multiple covering of a circle with random arcs. *J. Appl. Probability* **17** 284–290.

HOLST, L. (1980b). On matrix occupancy, committee, and capture-recapture problems. *Scand. J. Statist.*, **7** 139–146.

LECAM, L. (1958). Un théorème sur la division d'un intervalle par des points pris au hasard. *Publ. Inst. Statist. Univ. Paris* **7** 7–16.

LEHMANN, E. L. (1975). *Nonparametrics: Statistical Methods Based on Ranks.* Holden-Day, San Francisco.

MICHEL, R. (1979). On the asymptotic efficiency of conditional tests for exponential families. *Ann. Statist.* **7** 1256–1263.

PORTNOY, S. (1977). Asymptotic efficiency of minimum variance unbiased estimators. *Ann. Statist.* **5** 522–529.

RAO, C. R. (1973). *Linear Statistical Inference and Its Applications.* 2nd ed., Wiley, New York.

STECK, G. P. (1957). Limit theorems for conditional distributions. *Univ. Calif. Publ. Statist* **2** 237–284.

ZABELL, S. (1979). Continuous versions of regular conditional distributions. *Ann. Probability* **7** 159–165.

DEPARTMENT OF MATHEMATICS
UPPSALA UNIVERSITY
THUNBERGSV. 3
S-75238 UPPSALA
SWEDEN