The Annals of Statistics
1999, Vol. 27, No. 5, 1666–1683

SECOND-ORDER CORRECTNESS OF THE POISSON BOOTSTRAP

By G. Jogesh Babu,[1] P. K. Pathak and C. R. Rao[2]

Pennsylvania State University, Michigan State University and Pennsylvania State University

Rao, Pathak and Koltchinskii have recently studied a sequential approach to resampling in which resampling is carried out sequentially one-by-one (with replacement each time) until the bootstrap sample contains $m \approx (1 - e^{-1})n \approx 0.632n$ distinct observations from the original sample. In our previous work, we have established that the main empirical characteristics of the sequential bootstrap go through, in the sense of being within a distance $O(n^{-3/4})$ from those of the usual bootstrap. However, the theoretical justification of the second-order correctness of the sequential bootstrap is somewhat difficult. It is the main topic of this investigation. Among other things, we accomplish it by approximating our sequential scheme by a resampling scheme based on the Poisson distribution with mean $\mu = 1$ and censored at $X = 0$.

1. Introduction. Efron (1979) introduced the bootstrap method of resampling as a ubiquitous sampling technique of estimating the variance of an estimator and sampling distribution of a given statistic. Singh (1981) showed, using Edgeworth expansions in the case of univariate sample mean, that the bootstrap is more accurate than the central limit theorem when higher-order population moments exist. In a fundamental paper, Bhattacharya and Ghosh (1978) have demonstrated that Edgeworth expansion for a wide class of statistics can be derived from Edgeworth expansions for multivariate sample means. These ideas are further exploited by Babu and Singh (1983, 1984) to show the superiority of the bootstrap method and by Babu and Singh (1985) to obtain Edgeworth expansions for the ratio statistic and similar statistics based on samples from finite populations. The method is also used by Babu and Singh (1989) to obtain global Edgeworth expansions for functions of means of random vectors, when one of the coordinates has a lattice distribution and the remaining part of the vector has a strongly nonlattice distribution. Later, Giné and Zinn (1990) showed that in a certain weak sense, the bootstrap method is valid (consistent) if and only if the central limit theorem holds. In fact, the central limit theorem furnishes accuracy of approximation $o(1)$, while if the third population moment exists, one can expect, in many

Received March 1998; revised April 1999.

[1] Supported in part by NSA Grant MDA904-97-1-0023 and NSF Grant DMS-96-26189.

[2] Supported in part by Army Research Office Grant DAAH04-96-1-0082.

AMS 1991 *subject classifications*. Primary 62G09; secondary 60F05

Key words and phrases. Edgeworth expansions, bootstrap, breakdown point, expansions for conditional distributions, lattice distribution.

commonly encountered populations, the accuracy of the bootstrap method to be $o(n^{-1/2})$, where n denotes the sample size. Thus while the bootstrap method has the potential of being second-order accurate, the central limit approximation is not so. This is one of the several reasons for the current interest and preference in the literature for those methods of resampling that are second-order accurate, that is, accurate $o(n^{-1/2})$.

Stemming from Efron's observation that the information content of a bootstrap sample is based on approximately $(1 - e^{-1})100\% \approx 63\%$ of the original sample, Rao, Pathak and Koltchinskii (1997) have introduced a sequential resampling method in which sampling is carried out one-by-one (with replacement) until $(m + 1)$ distinct original observations appear, where m denotes the largest integer not exceeding $(1 - e^{-1})n$. It has been shown that the empirical characteristics of this sequential bootstrap are within a distance $O(n^{-3/4})$ from the usual bootstrap. The authors provide a heuristic argument in favor of their sampling scheme and establish the consistency of the sequential bootstrap; however the question of second-order correctness is not addressed.

The main object of this paper is to examine the second-order correctness of the sequential bootstrap. The theoretical justification of this is somewhat more difficult because of the dependence among the bootstrap sample units. At this time, a rigorous Edgeworth expansion under this kind of dependence is unavailable in the literature. A cumbersome approach based on computation of cumulants, under the (unsubstantiated) assumption that a formal Edgeworth expansion is valid, may be given along the lines of Hall and Mammen (1994). This does not lead to a complete solution, as the Edgeworth expansions are not known. Instead, we first approximate the sequential bootstrap by another sequential resampling scheme based on the Poisson distribution. Under the new scheme the "independence" of sample units under resampling is preserved. A rigorous justification of the Edgeworth expansion can now be given more easily. In this paper we provide details for the sample mean. Edgeworth expansions for statistics that can be represented as smooth functions of multivariate sample means are considered in Section 4.

2. Sequential resampling scheme. Let $S = (X_1, X_2, \dots, X_n)$ be a random sample from a distribution F and $\theta(F)$ a parameter of interest. Let F_n denote the empirical distribution function based on S and suppose that $\theta(F_n)$ is to be used as an estimator of $\theta(F)$. The Efron bootstrap method approximates the sampling distribution of a standardized version of $\sqrt{n}\,(\theta(F_n) - \theta(F))$ by the resampling distribution of a corresponding statistic $\sqrt{n}\,(\theta(\hat{F}_n) - \theta(F_n))$ based on a bootstrap sample $\hat{S}_n$ in which the original F has been replaced by the empirical distribution based on the original sample S and F_n of the former statistic has been replaced by the empirical distribution based on a bootstrap sample $\hat{F}_n$. In Efron's bootstrap resampling scheme, $\hat{S}_n = (\hat{X}_1, \hat{X}_2, \dots, \hat{X}_n)$ is a random sample of size n drawn from S by simple random sampling with replacement (SRSWR). In the Rao, Pathak and

Koltchinskii (1997) sequential scheme, observations are drawn from S sequentially by SRSWR until there are $(m+1) = [n(1-e^{-1})] + 2$ distinct original observations in the bootstrap sample; the last observation is discarded to ensure technical simplicity. Thus an observed bootstrap sample under the Rao–Pathak–Koltchinskii scheme admits the form

$$\hat{S}_{N_1} = \left(\hat{X}_1, \hat{X}_2, \dots, \hat{X}_{N_1}\right) \tag{2.1}$$

in which $\hat{X}_1, \hat{X}_2, \dots, \hat{X}_{N_1}$ have $m \approx n(1-e^{-1})$ distinct observations from S. The random sample size N_1 admits the following decomposition in terms of the independent random variables:

$$N_1 = I_1 + I_2 + \cdots + I_m$$

in which $m = [n(1-e^{-1})] + 1$; $I_1 = 1$ and for each k, $2 \le k \le m$,

$$P(I_k = j) = \left(1 - \frac{k-1}{n}\right)\left(\frac{k-1}{n}\right)^{j-1}.$$

Although we have established the consistency of this sampling scheme, a rigorous proof of its second-order correctness requires an Edgeworth expansion for dependent random variables; such an expansion is unavailable in the literature at the present time. An alternative approach that can be used is to slightly modify the preceding resampling scheme so that existing techniques on Edgeworth expansion, such as those of Babu and Bai (1996), Bai and Rao (1991, 1992), Babu and Singh (1989) and others, can be employed. A modification of our previous resampling scheme that allows the second-order correctness to go through is as follows.

2.1. *Poisson Bootstrap.* For the selection of a bootstrap sample with a given number m of distinct units, under the *Poisson bootstrap*, we provide a conceptual definition and a practical approach. Let us take a sample $\alpha_1, \dots, \alpha_n$ of n independent observations from $P(1)$, that is, Poisson distribution with mean 1. If there are exactly m nonzero values in the sample, we accept it and take

$$\hat{S} = \{(X_1, \alpha_1), (X_2, \alpha_2), \dots, (X_n, \alpha_n)\}, \tag{2.2}$$

that is, with the observation X_i repeated α_i times, as the bootstrap sample. If the number of nonzero values in $\alpha_1, \dots, \alpha_n$ is not exactly m, we reject the entire sample and draw another sample of size n. The bootstrap sample size N_2 of $\hat{S}$ as in (2.1) is a random variable

$$N_2 = \alpha_1 + \cdots + \alpha_n.$$

A practical way of implementing this resampling scheme is to first assign at random $(n-m)$ α's a value of zero and to the remaining $m\alpha$'s values independently chosen from the Poisson distribution with mean $\mu = 1$ and censored at $X = 0$. An outline of the equivalence of these two procedures is as follows.

THEOREM 2.1. *The moment generating function $M_{N_2}(t)$ of N_2, the sample size of the Poisson resampling scheme, is given by*

$$M_{N_2}(t) = \left((\exp((e^t - 1)) - e^{-1})/(1 - e^{-1})\right)^m. \tag{2.3}$$

PROOF. Let $Y_1, Y_2, \dots, Y_n$ be n Poisson variables with mean $\mu = 1$. Then it is easily seen that

$$P(N_2 = w) = \text{const}\, \Sigma_1 e^{-n}(\alpha_{1!}\alpha_{2!} \cdots \alpha_{m!})^{-1},$$

where the sum Σ_1 extends over all positive natural numbers $\alpha_1, \alpha_2, \dots, \alpha_m$ such that $\alpha_1 + \alpha_2 + \cdots + \alpha_m = w$. It then follows that [see Pathak (1962)]

$$\begin{aligned} P(N_2 = w) &= \text{const}\left(\frac{e^{-n}}{w!}\left(m^w - \binom{m}{1}(m-1)^w + \cdots \pm 1^w\right)\right) \\ &= \text{const}\frac{e^{-n}}{w!}(\Delta^m X^w|_{X=0}), \end{aligned} \tag{2.4}$$

where Δ is the difference operator with unit increment.

From (2.4) it follows that

$$P(N_2 = w) = \frac{1}{(e-1)^m}\Delta^m \frac{X^w}{w!}\bigg|_{X=0}.$$

Consequently, the moment generating function $M_{N_2}(t)$ of N_2 is given by

$$\begin{aligned} M_{N_2}(t) &= E(\exp(tN_2)) = \sum_{w \ge 0} \frac{e^{tw}}{(e-1)^m}\Delta^m \frac{X^w}{w!}\bigg|_{X=0} \\ &= \frac{1}{(e-1)^m}\sum_{w\ge 0}\Delta^m \frac{(e^t X)^w}{w!}\bigg|_{X=0} \\ &= \Delta^m (e-1)^{-m}\exp(Xe^t)|_{X=0} \\ &= (e-1)^{-m}\bigg\{\exp(me^t) - \binom{m}{1}\exp((m-1)e^t) \\ &\qquad + \binom{m}{2}\exp((m-2)e^t)\cdots\bigg\} \\ &= (e-1)^{-m}\exp(me^t)\left\{1 - \binom{m}{1}e^{-t} + \binom{m}{2}e^{-2t} - \cdots\right\} \\ &= (e-1)^{-m}\exp(me^t)(1 - \exp(-e^t))^m \\ &= ((\exp(e^t) - 1)/(e-1))^m = ((\exp((e^t - 1)) - e^{-1})/(1 - e^{-1}))^m. \end{aligned}$$

This completes the proof. □

The preceding theorem shows that the distribution of N_2 can be viewed as that of the sum of m i.i.d. random variables with a common distribution with the moment generating function given by the formula

$$m(t) = \big(\exp((e^t - 1)) - e^{-1}\big)(1 - e^{-1})^{-1}. \tag{2.5}$$

It is evident that $m(t)$ is the moment generating function of the Poisson distribution with mean 1 and censored at $X = 0$. Let T denote a random variable with moment generating function $m(t)$. Then $E(T) = 1/(1 - e^{-1})$ and $V(T) = e(e - 2)/(e - 1)^2$. Therefore

$$E(N_2) = mE(T) = n + O(1)$$

and

$$V(N_2) = mV(T) = n(e - 2)/(e - 1) + O(1).$$

2.2. *Advantages of Poisson bootstrap over classical bootstrap.* One of the main advantages of the sequential bootstrap over the classical fixed sample size bootstrap is to avoid situations where a bootstrap sample has several repeated observations which may give rise to a degenerate value of the statistic under consideration. Thus Poisson bootstrap avoids zero value for variance estimation.

Another reason to prefer Poisson bootstrap is the robustness of variance estimation. Since the bootstrap utilizes all the data points, in general, the bootstrap estimator of variance of a statistic is not robust for robust statistics. The bootstrap estimator of variance of the sample median m_n based on the sample $X_1, \dots, X_n$ is given by

$$V_n^* = \tfrac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \big(X_{(i)} - X_{(j)}\big)^2 p_{i,n} p_{j,n},$$

where $X_{(1)} \leq \cdots \leq X_{(n)}$ is the ordering of the data,

$$p_{i,n} = n\binom{n-1}{r-1} \int_{(i-1)/n}^{i/n} u^r (1-u)^{n-r}\, du,$$

and $r = (n/2)$ if n is an even integer and $= (n+1)/2$ if it is an odd integer. Breakdown point is a widely used measure of robustness in modern statistical literature. The breakdown point of a statistic $\hat{\theta}_n = \hat{\theta}_n(X_1, \dots, X_n)$ is defined as (k/n), where k is the minimum number of data points needed to be replaced by worst possible outliers to move the statistic beyond any bound. In the case of bootstrap variance of sample median, the breakdown point is $(1/n)$.

However, in the case of sequential bootstrap or Poisson bootstrap, $X_{(i)}$ for $i < m - (n/2)$ do not enter the estimate of the variance of the sample median. Hence the breakdown point is $\approx \frac{1}{2} - e^{-1} = 0.132$.

We now proceed to establish the second-order correctness of the Poisson bootstrap, a sequential bootstrap based on the Poisson distribution.

3. Second order correctness.

Let $\{X_1, \dots, X_n\}$ be i.i.d. random variables with mean μ and variance σ^2. Suppose X_1 satisfies Cramér's condition

$$\limsup_{|t| \to \infty} \left| E(\exp(itX_1)) \right| < 1. \tag{3.1}$$

Let $Y_1, Y_2, \dots$ be a sequence of i.i.d. Poisson random variables with mean 1. Let

$$\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i, \qquad s_n^2 = \frac{1}{n} \sum_{i=1}^{n} \left(X_j - \overline{X}_n \right)^2, \qquad C_j = C_{j,n} = \left(X_j - \overline{X}_n \right)/s_n,$$

$$N = \sum_{i=1}^{n} Y_i, \qquad T_n = \sum_{i=1}^{n} \delta_i, \qquad q = e^{-1} = P(Y = 0) \text{ and } p = 1 - e^{-1},$$

where

$$\delta_i = \begin{cases} 1, & \text{if } Y_i > 0, \\ 0, & \text{otherwise.} \end{cases}$$

We shall obtain Edgeworth expansions for the distribution of $N^{-1/2} \sum_{j=1}^{n} C_j Y_j$ given $T_n = m$ and $X_1, \dots, X_n$. Then use this result to establish second-order correctness for the Poisson bootstrap in the simple case of $\sqrt{n}(\overline{X} - \mu)/\sigma$. This result can be extended to statistics which can be represented as a smooth function of a multivariate mean. Second-order correctness of Poisson bootstrap for such models is discussed in Section 4.

We now state the main theorem.

THEOREM 3.1. *Suppose* $E|X_1|^5 < \infty$ *and that the characteristic function of* X_1 *satisfies Cramér's condition* (3.1). *If* $m - np$ *is bounded, then*

$$\begin{aligned} &P\left(\frac{1}{\sqrt{N}} \sum_{i=1}^{n} (X_i - \overline{X}_n) Y_i \le x s_n | T_n = m; X_1, \dots, X_n \right) \\ &\qquad - P\left(\frac{1}{\sqrt{N}} \sum_{i=1}^{n} (X_i - E(x_1)) \le x\sigma \right) = O_p(n^{-1}), \end{aligned} \tag{3.2}$$

uniformly in x.

REMARK 1. With truncation and additional analysis, the moment condition on X_1 can be relaxed.

To prove Theorem 3.1, we first establish some notation and a preliminary proposition. Let ϕ denote the standard normal density, φ_0 denote the bivariate normal density with zero mean vector and dispersion matrix

$$\Sigma_0 = \begin{pmatrix} 1 & q \\ q & pq \end{pmatrix}.$$

Define

$$P_n(x) = \left(\frac{1}{6n}\sum_{j=1}^{n} C_j^3\right)(x^3 - x),$$

$$\begin{aligned} Q_n(x, y, \omega) &= P_n(x) + \tfrac{1}{6}E\big(Z'\Sigma_0^{-1}(y, \omega)'\big)^3 \\ &\quad - \tfrac{1}{2}\Big(E\big((Z'\Sigma_0^{-1}(y, \omega)')(Y-1)^2 pq + (q-p)\delta\big)/(q(p-q))\Big) \\ &\quad + \tfrac{1}{2}(x^2 - 1)E\big((Y-1)^2 Z'\Sigma_0^{-1}(y, \omega)'\big), \end{aligned}$$

where Y is a Poisson random variable with mean 1, $\delta = I_{\{Y>0\}}$ and $Z = (Y - 1, \delta - p)'$.

For brevity fix $X_1 \cdots X_n$ and define

$$y_r = (r - n)n^{-1/2}, \qquad \omega_m = (m - np)n^{-1/2}$$

and

$$F_n(x, r, m) = P\left(\sum_{i=1}^{n} C_i Y_i \le x\sqrt{n}\,, N = r, T_n = m\right).$$

Note that $\sum_{i=1}^{n} C_i = 0$ and $\sum_{i=1}^{n} C_i^2 = n$.

PROPOSITION 1. *Suppose for any $K > 0$, there exists a $0 < \gamma < 1$ such that*

$$\limsup_{n\to\infty} \sup_{K \le |t| \le n^3} \left|\frac{1}{n}\sum_{j=1}^{n} \exp(itC_j)\right| < \gamma.$$

Suppose $m - np$ is bounded and for some $M > 1$, $\sum_{i=1}^{n} |C_i|^5 < nM$. Then uniformly in x, r and m, we have

$$\begin{aligned} nP&\left(\sum_{i=1}^{n} C_i Y_i \le x\sqrt{n}\,, N = r, T_n = m\right) \\ &= \int_{-\infty}^{x} \Psi_n(v, y_r, \omega_n)\,dv + O(n^{-3/2}), \end{aligned} \tag{3.3}$$

where

$$\begin{aligned} \Psi_n(v, y, \omega) &= \phi(v)\varphi_0(y, \omega) \\ &\quad \times\big(1 + n^{-1/2}Q_n(v, y, \omega) + n^{-1}Q_{n1}(v, y, w)\big). \end{aligned} \tag{3.4}$$

Q_{n1} is a fourth degree polynomial in v, y, ω whose coefficients are bounded by a constant depending only on M.

The proposition is proved in the Appendix.

PROOF OF THEOREM 3.1. Using Proposition 1, we shall first show that uniformly in x,

$$
\begin{aligned}
&P\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)Y_i \le xs_n \mid T_n=m, X_1,\dots,X_n\right)\\
&\qquad = \Phi(x)+\frac{1}{6\sqrt{n}}\gamma_3(1-x^2)\phi(x)+O_p\left(\frac{1}{n}\right),
\end{aligned}
\tag{3.5}
$$

where Φ is the standard normal distribution function and $\gamma_3=\sigma^{-3}E(X_1-\mu)^3$. Let $H_n=\{|n^{-1/2}\sum_{i=1}^{n}(Y_i+1)|\le \log n\}$. Then by a moderate deviation result,

$$
1-P(H_n)=O(n^{-10}). \tag{3.6}
$$

Suppose $\sum_{j=1}^{n}|X_i-\overline{X}_n|^5 \le nMs_n^5$ for some $M>1$; then by Proposition 1,

$$
\begin{aligned}
&\sqrt{n}\,P\left(\frac{1}{\sqrt{N}}\sum_{j=1}^{n}\left(X_j-\overline{X}_n\right)Y_j \le xs_n, T_n=m\right)\\
&\quad=\sum_{|r-n|\le\sqrt{n}\log n}\sqrt{n}\,P\left(\frac{1}{\sqrt{r}}\sum_{j=1}^{n}C_jY_j\le x, N=r, T_n=m\right)+O(n^{-9})\\
&\quad=\frac{1}{\sqrt{n}}\sum_{|r-n|\le\sqrt{n}\log n}\int_{-\infty}^{x\sqrt{r/n}}\Psi_n(v,y_r,\omega_m)\,dv+O(n^{-1})\\
&\quad=\frac{1}{\sqrt{n}}\sum_{|r-n|\le\sqrt{n}\log n}\int_{-\infty}^{x}\Psi_n\left(u\left(1+\frac{1}{\sqrt{n}}y_r\right)^{1/2},y_r,\omega_m\right)\\
&\qquad\qquad\times\left(1+\frac{1}{\sqrt{n}}y_r\right)^{1/2}du\\
&\quad=\int_{-\infty}^{x}\left(\int_{|y|<\log n}\Psi_n\left(u\left(1+\frac{1}{\sqrt{n}}y\right)^{1/2},y,\omega_m\right)\right.\\
&\qquad\qquad\left.\times\left(1+\frac{1}{\sqrt{n}}y\right)^{1/2}dy\right)du\\
&\qquad+O(n^{-1}).
\end{aligned}
\tag{3.7}
$$

By Theorem 13 on local Edgeworth expansion on pages 205 and 206 of Petrov (1975), we have

$$
\begin{aligned}
\sqrt{npq}\,P(T_n=m)&=\phi(\omega_m)+\sum_{j=1}^{2}n^{-j/2}q_j(\omega_m)\phi(\omega_m)+O(n^{-1})\\
&=\phi(\omega_m)+n^{-1/2}q_1(\omega_m)\phi(\omega_m)+O(n^{-1}),
\end{aligned}
\tag{3.8}
$$

where q_1 and q_2 are polynomials and

$$q_1(y) = \tfrac{1}{6}E(\delta_1 - p)^3(pq)^{-3/2}(y^3 - 3y).$$

The estimate (3.5) follows from (3.6), (3.7) and (3.8) after some simple algebra. The theorem now follows from the standard Edgeworth expansion

$$P\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(X_i - \mu) \le \sigma x\right) = \Phi(x) + \frac{1}{\sqrt{n}}\gamma_3(1 - x^2)\phi(x) + O(n^{-1})$$

and

$$P\left(\sum_{j=1}^{n}|X_i - \bar{X}_n|^5 > nMs_n^5\right) = O\left(\frac{1}{n}\right) \quad \text{for some } M > 1.$$

This completes the proof. □

4. Smooth functional model. Proposition 1 and Theorem 3.1 can be extended to multivariate cases and to statistics which can be expressed as smooth functions of multivariate means. Let $X_1 \cdots X_n$ be a sequence of i.i.d. random vectors with mean μ and dispersion Σ. Let Σ_n denote the sample dispersion matrix of $X_1, \dots, X_n$. With some additional effort and using the ideas and lemmas from Babu and Bai (1996), we can establish the following results.

THEOREM 4.1. *Suppose the characteristic function of X_1 satisfies Cramér's condition and $E\|X_1\|^4$. Let H be a three times continuously differentiable function in a neighborhood of μ. Let $l(y)$ denote the vector of first-order partial derivatives at y and $l(\mu) \ne 0$. If $m - n(1 - e^{-1})$ is bounded, then for almost all sample sequences $\{X_j\}$, we have*

$$\sup_x \sqrt{n}\left|P\left(\frac{\sqrt{N}\left(H(N^{-1}\Sigma_{i=1}^{n}X_iY_i) - H(\bar{X}_n)\right)}{\sqrt{l'(\bar{X}_n)\Sigma_n l(\bar{X}_n)}} \le x|T_n = m;\, X_1, \dots, X_n\right)\right.$$
$$\left. -P\left(\sqrt{n}\left(H(\bar{X}_n) - H(\mu)\right) \le x\sqrt{l'(\mu)\Sigma l(\mu)}\right)\right| \to 0,$$

as $n \to \infty$.

The next two results are more suitable for applications to Studentized statistics.

THEOREM 4.2. *Let $\{X_n\}$ be as in Theorem* 4.1. *Suppose the function H is three times continuously differentiable in a neighborhood of the origin and $H(0) = 0$. If $m - n(1 - e^{-1})$ is bounded, then for almost all sample sequences $\{X_j\}$, we have*

$$\sup_x \sqrt{n}\left| P\left(\sqrt{N} H\left(N^{-1} \sum_{i=1}^{n} \left(X_i - \bar{X}_n\right) Y_i \right) \le x\sqrt{l'(0)\Sigma_n^2 l(0)} \,|T_n = m; \; X_1, \dots, X_n \right) - P\left(\sqrt{n}\, H\left(\bar{X}_n - \mu\right) \le x\sqrt{l'(0)\Sigma l(0)}\right) \right| \to 0,$$

as $n \to \infty$.

It is easily seen that the second-order correctness of the Poisson bootstrap of a pivot such as

$$\pi_N^* = \sqrt{N}\left(\sum_{i=1}^{n} \left(X_j - \bar{X}_n\right) Y_j \right) \Big/ s_n$$

follows from Theorem 4.2. The one-term correction captures the skewness of the underlying distribution.

The most commonly used statistics, especially the Studentized versions, are of the type

$$t_n = \sqrt{n}\left(H(\bar{X}) - H(\mu)\right)/\xi\left(\frac{1}{n} \sum_{i=1}^{n} \lambda(X_i) \right), \tag{4.1}$$

where λ is a function on $\mathbb{R}^k \to \mathbb{R}^r$ and ξ is a smooth real-valued function on $\mathbb{R}^r$. The classical Student's t is an example of this type of statistic. If X_i are univariate, then

$$t_{n1} = \sqrt{n}\left(\bar{X}_n - E(X_1)\right)/s_n,$$

satisfies (4.1) with $H(x) = x$, $\lambda(x) = (x^2, x)$, $\xi(x, y) = \max(0, (x - y^2))^{1/2}$ and $s_n^2 = (1/n)\Sigma_{i=1}^n (X_i - \bar{X}_n)^2$. The version corresponding to (4.1) under the Poisson scheme is generally of the type

$$t_n(Y) = \sqrt{N}\left(H\left(\frac{1}{n} \sum_{i=1}^{n} X_i Y_i \right) - H\left(\bar{X}_n\right) \right) \Big/ \xi\left(\frac{1}{N} \sum_{i=1}^{n} \lambda(X_i) Y_i \right).$$

As in Theorem 4 of Babu and Singh (1984), we can derive the following.

THEOREM 4.3. *Let*

$$\xi(E(\lambda(X_1))) = \sqrt{l'(\eta)\Sigma l(\eta)}, \qquad \xi\left(\bar{X}_n\right) = \sqrt{l'\left(\bar{X}_n\right)\Sigma_n l\left(\bar{X}_n\right)}$$

and let $L(X_i)$ be a linearly independent subcollection of $(X_i, \lambda(X_i))$ with the property that all the coordinates of $(X_i, \lambda(X_i))$ can be expressed as linear combinations of those of $L(X_i)$. If the characteristic function of $L(X_1)$ satisfies

Cramér's condition $E\|L(X_1)\|^4 < \infty$, *and if* $m - np$ *is bounded then*

$$\sup_x \sqrt{n}\,|P(t_n(Y) \le x | T_n = m, X_1, \dots, X_n) - P(t_n \le x)| \to 0,$$

as $n \to \infty$, *for almost all sample sequences* $\{X_j\}$.

APPENDIX

To establish Proposition 1 of Section 3, we require some preliminary results. First note that by Theorem 10.1 of Bhattacharya and Ranga Rao (1986), there exists a random variable V with distribution J such that $E(V^{10}) < \infty$, $P(|V| \ge 1) < \frac{1}{4}$ and the characteristic function $\hat{J}$ of J vanishes outside the interval $[-c, c]$, for some $c > 0$, that is,

$$\hat{J}(t) = 0 \quad \text{for } |t| > c. \tag{A.1}$$

For any $\varepsilon > 0$, let J_ε denote the distribution of $V\varepsilon$. The next lemma is a trivial consequence of Lemma 24.1 of Bhattacharya and Ranga Rao (1986). The inequality is similar to (4.1) of Babu and Singh (1984). See also (4.2) of Babu and Singh (1984).

LEMMA 1. *Suppose* h *is a Borel measurable function on the real line* $\mathbb{R}$, *bounded by* 1. *Let* μ *be a finite measure and* ν *a finite signed measure. Then for any* $0 < \varepsilon < 1$,

$$\begin{aligned}\left|\int h(y)(\mu - \nu)(dy)\right| \le 15\Bigg(&|(\mu - \nu) * J_\varepsilon|(\mathbb{R}) + 3^{-\varepsilon^{1/4}}(\mu(\mathbb{R}) + \nu^+(\mathbb{R}))\\ &+(\mu(\mathbb{R}) + \nu^+(\mathbb{R}))P(|V| > \varepsilon^{-1/2})\\ &+ \sup_{|z| \le \varepsilon^{1/4}} \int \omega(h, 2\varepsilon, z + y)\nu^+(dy)\Bigg),\end{aligned}$$

where $\nu = \nu^+ - \nu^-$ *is the Jordan decomposition*, $|\nu| = \nu^+ + \nu^-$ *is the total variation measure and*

$$\omega(h, \varepsilon, z) = \sup(|h(z + y') - h(z + y'')| : |y'| < \varepsilon, |y''| < \varepsilon).$$

Further, *we have for any* $0 < v < 1$, $0 < \varepsilon < 1$,

$$\begin{aligned}&\int \omega(h, \varepsilon, v - y)\phi(y)\,dy\\ &\qquad \le 3\int \omega(h, \varepsilon, y)\phi(y)\,dy + O\left(v^{-3}\exp\left(-\tfrac{1}{8}v^{-2}\right)\right).\end{aligned} \tag{A.2}$$

The inequality (A.2) is the relation (35) of Sweeting (1977).

The next lemma helps in estimating the total variation norm of $(\mu - \nu) * J_\varepsilon$. Let $D^k h(t, s, \omega)$ denote the kth partial derivative of h with respect to t.

LEMMA 2. *Let* g *be a real-valued function on* $\mathbb{R} \times \mathbb{Z}^2$ *satisfying*

$$\sum_{r=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \int_{-\infty}^{\infty} (1 + x^2)|g(x, r, j)|\,dx < \infty.$$

Then for all integers r and j,

$$\int_{-\infty}^{\infty} |g(x,r,j)|\,dx \le \max_{k=0,2} \int_G \int_{-\infty}^{\infty} \left|D^k \hat{g}(t,s,\omega)\right| dt\,ds\,d\omega,$$

where $G = [-\pi,\pi] \times [-\pi,\pi]$ and $\hat{g}$ denotes the Fourier transform of g, that is,

$$\hat{g}(t,s,\omega) = \sum_{r=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \exp(isr + i\omega j) \int_{-\infty}^{\infty} e^{itx} g(x,r,j)\,dx.$$

PROOF. Following the proof of Lemma 11.6 of Bhattacharya and Ranga Rao (1986), let $A = \{x: g(x,r,j) \ge 0\}$. If for each r, j,

$$\hat{g}(t;r,j) = \int_{-\infty}^{\infty} e^{itx} g(x,r,j)\,dx,$$

then the Fourier transform $\hat{h}_{r,j}$ of the function

$$h_{r,j}(x) = (1+x^2) g(x,r,j)$$

is given by

$$\hat{h}_{r,j}(t) = \hat{g}(t;r,j) - D^2\hat{g}(t;r,j).$$

Thus by Fourier inversion and Fubini's theorem we have

$$\begin{aligned}(1+x^2)g(x,r,j) = (2\pi)^{-3} \int_{-\infty}^{\infty} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \big(\hat{g}(t,s,\omega) - D^2\hat{g}(t,s,\omega)\big)& \\ \times \exp(-i(sr + \omega j + tx))\,ds\,d\omega\,dt.&\end{aligned}$$

Hence we have

$$\begin{aligned}
&\int_{-\infty}^{\infty} |g(x,r,j)|\,dx \\
&\quad = \int_A g(x,r,j)\,dx - \int_{\mathbb{R}-A} g(x,r,j)\,dx \\
&\quad = \left(\int_A - \int_{\mathbb{R}-A}\right)(1+x^2)g(x,r,j)\,dx \\
&\quad = (2\pi)^{-3}\left(\int_A - \int_{\mathbb{R}-A}\right)(1+x^2)^{-1}\left(\int_{-\infty}^{\infty}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}(\hat{g}(t,s,\omega)\right. \\
&\qquad\qquad \left. -D^2\hat{g}(t,s,\omega))\exp(-i(sr+\omega j+tx)\,ds\,d\omega\,dt)\right) dx \\
&\quad \le (2\pi)^{-3}\left(\int_{-\infty}^{\infty}(1+x^2)^{-1}\,dx\right)\int_{-\infty}^{\infty}\int_G \big(|\hat{g}(t,s,\omega)| \\
&\qquad\qquad + \left|D^2\hat{g}(t,s,\omega)\right|\big)\,ds\,d\omega\,dt \\
&\quad \le \max_{k=0,2}\int_{-\infty}^{\infty}\int_G \left|D^k\hat{g}(t,s,\omega)\right| ds\,d\omega\,dt.
\end{aligned}$$

This completes the proof of the lemma. □

The next lemma is Lemma 2 of Babu and Singh (1984) and is stated here for ready reference.

LEMMA 3. *Suppose* X_1 *satisfies Cramér's condition* (3.1). *For any* $K > 0$, *there exists a* $0 < \gamma = \gamma(K) < 1$ *and* $\eta > 0$ *such that for almost all sample sequences* $\{X_i\}$,

$$\limsup_{n\to\infty} \sup_{K<|t|<e^{\eta n}} \left| \frac{1}{n} \sum_{j=1}^{n} \exp\Bigl(iu\bigl(X_j - \overline{X}_n\bigr)\Bigr) \right| < \gamma.$$

The next three lemmas are similar to Lemmas 2 and 3 of Babu and Bai (1996). To state these, let $d_j = d_{j,n} = c_j n^{-1/2}$ satisfy for some $M' > 1$,

$$\sum_{1}^{n} d_j = 0, \qquad \sum_{1}^{n} d_j^2 = 1 \quad \text{and} \quad \sum_{j=1}^{d} |d_j|^3 \le \bigl(M'/\sqrt{n}\,\bigr). \tag{A.3}$$

In proving Proposition 1, we apply the lemmas with $d_j = C_j n^{-1/2}$. For each fixed r and m, the Fourier transform of

$$F_n(x, r, m) = P\left(\sum_{j=1}^{n} d_j Y_j \le x,\ \sum_{j=1}^{n} Y_j = r,\ \sum_{j=1}^{n} \delta_j = m \right) \tag{A.4}$$

is given by

$$\begin{aligned} h(t; r, m) &= E\left(\exp\left(it \sum_{j=1}^{n} d_j \right) I_{(\sum_{j=1}^{n} Y_j = r,\, \sum_{j=1}^{n} \delta_j = m)} \right) \\ &= \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} f_n(t, s, \omega) \exp(-isr - i\omega m)\, ds\, d\omega, \end{aligned} \tag{A.5}$$

where f_n is the characteristic function of $\sum_{j=1}^{n}(d_j Y_j, Y_j, \delta_j)$. Note that

$$f_n(t, s, \omega) = \prod_{j=1}^{n} f\bigl(td_j + s, \omega\bigr),$$

where

$$f(u, \omega) = E\bigl(\exp(iuY_1 + i\omega\delta_1)\bigr) = e^{-1}\bigl(1 - e^{i\omega} + \exp(i\omega - e^{iu})\bigr).$$

Let for any subset R of $\{1, \dots, n\}$,

$$f_{n,R}(t, s, \omega) = \prod_{j\in R} f\bigl(td_j + s, \omega\bigr)$$

and for any $0 \le k \le 2$, let

$$f_{n,k}(t, s, \omega) = \max\bigl\{ | f_{n,R}(t, s, \omega) | \bigr\},$$

where the maximum is taken over all subsets R of $\{1, \dots, n\}$ of size $n - k$. Finally, let

$$d_n(t) = \left| \frac{1}{n} \sum_{j=1}^{n} \exp\bigl(itd_j\bigr) \right|. \tag{A.6}$$

LEMMA 4. *For any* $0 \le k \le 2$,

$$f_{n,k}(t,s,\omega) \le \exp\bigl(1 - 0.2n(1 - d_n(t))\bigr). \tag{A.7}$$

PROOF. Observe that

$$\begin{aligned}
\frac{1}{n}\sum_{j=1}^{n} \bigl|f(td_j + s, \omega)\bigr|^2 &= \frac{1}{n}\sum_{j=1}^{n} E\Bigl(\exp\bigl(i(td_j + s)(Y_1 - Y_2) + i\omega(\delta_1 - \delta_2)\bigr)\Bigr) \\
&= E\left(\left(\frac{1}{n}\sum_{j=1}^{n} \exp\bigl(it(Y_1 - Y_2)d_j\bigr)\right)\right. \\
&\qquad \left. \times \exp\bigl(is(Y_1 - Y_2) + i\omega(\delta_1 - \delta_2)\bigr)\right) \\
&\le E\bigl(d_n(t(Y_1 - Y_2))\bigr) \\
&\le P(|Y_1 - Y_2| \ne 1) + d_n(t)P(|Y_1 - Y_2| = 1) \\
&\le 1 - (1 - d_n(t))P(|Y_1 - Y_2| = 1) \\
&\le 1 - (1 - d_n(t))0.4.
\end{aligned}$$

As $x \le e^{x-1}$ for all $0 \le x \le 1$, for any subset $R \subset \{1, \dots, n\}$ with at least $n - 2$ integers, we have

$$\begin{aligned}
|f_{n,R}(t,s,\omega)|^2 &= \prod_{j \in R} \bigl|f(td_j + s, \omega)\bigr|^2 \\
&\le \exp\left(2 - n + \sum_{j \in R} \bigl|f(td_j + s, \omega)\bigr|^2\right) \\
&\le \exp\left(2 - n + \sum_{j=1}^{n} \bigl|f(td_j + s, \omega)\bigr|^2\right) \\
&\le \exp\bigl(2 - 0.4n(1 - dn(t))\bigr).
\end{aligned}$$

This completes the proof. □

LEMMA 5. *For* $|t| \le \sqrt{n}/M'$, *we have*

$$d_n(t) = \left|\frac{1}{n}\sum_{j=1}^{n} \exp(itd_j)\right| \le 1 - \frac{1}{3n}t^2 \tag{A.8}$$

and for $0 \le k \le 2$, $|s| \le \pi$, $|\omega| \le \pi$, $|t| \le \sqrt{n}/M'$,

$$f_{n,2}(t,s,\omega) \le \exp\left(1 - \frac{1}{15}t^2\right). \tag{A.9}$$

PROOF. For $|t| \le \sqrt{n}/M'$,

$$\begin{aligned} d_n(t) &= \left| 1 - \frac{t^2}{2n} + \frac{1}{n} \sum_{j=1}^{n} \left(\exp(itd_j) - 1 - itd_j + \frac{1}{2} t^2 d_j^2 \right) \right| \\ &\le 1 - \frac{1}{2n} t^2 + \frac{M'}{6} |t|^3 n^{-3/2} \\ &\le 1 - \frac{1}{3n} t^2. \end{aligned}$$

The second part follows from (A.7). □

LEMMA 6. *Suppose $|t| \le \log n$, $|s| \le \pi$, $|\omega| \le \pi$ and $s^2 + \omega^2 \ge (\log n)^2 n^{-1}$. Then there exist constants k_1 and k_2 depending only on M' such that for $0 \le k \le 2$,*

$$f_{n,k}(t, s, \omega) \le k_1 \exp\left(-k_2 (\log n)^2\right). \tag{A.10}$$

PROOF. As $E(Y_1^3) = 5$,

$$\begin{aligned} &\left| \exp(-iu - ivp) E(\exp(iuY_1 - iu\delta_1)) - 1 + \tfrac{1}{2}(u,v)\Sigma_0(u,v)' \right| \\ &\quad \le (u^2 + v^2)^{3/2}. \end{aligned}$$

Since

$$(u,v)\Sigma_0(u,v)' \ge q\left(\tfrac{1}{2} - q\right)(u^2 + v^2),$$

there exists a positive $\eta < (q/16)(\frac{1}{2} - q)$ such that for all $|u| < 4\eta$, $|v| < 4\eta$, we have

$$\begin{aligned} \frac{1}{2} &\le \left| E(\exp(iuY_1 + iv\delta_1)) \right| \\ &\le 1 - 2^{-1}(u,v)\Sigma_0(u,v)' + (u^2 + v^2)^{3/2} \\ &\le 1 - \frac{q}{2}\left(\frac{1}{2} - q \right)(u^2 + v^2) + (u^2 + v^2)^{3/2} \\ &\le 1 - \left(\frac{q}{2}\left(\frac{1}{2} - q \right) - 4\eta \right)(u^2 + v^2) \\ &\le 1 - \eta(u^2 + v^2) \\ &\le \exp(-\eta(u^2 + v^2)). \end{aligned} \tag{A.11}$$

If $|t| \le (\log n)$ then $|td_j| \le (M')^{1/3} n^{-1/6} \log n \le \eta$ for all $n > n(M', \eta)$. In this case if $|s| < 2\eta$, $|\omega| < 2\eta$, then $|s + td_j| \le 3\eta$ for all $n \ge n(M', \eta)$. So by (A.11),

$$\begin{aligned} f_{n,k}(t, s, \omega) &\le 4 |f_n(t, s, \omega)| \\ &\le 4 \exp(-\eta n(s^2 + \omega^2) - \eta t^2) \\ &\le 4 \exp(-\eta n(s^2 + \omega^2)) \\ &\le 4 \exp\left(-\eta (\log n)^2\right), \end{aligned}$$

provided $s^2 + \omega^2 \ge (\log n)^2 n^{-1}$ and $|s| < 2\eta$, $|\omega| < 2\eta$.

To establish inequality (A.10) for the remaining s and ω, note that b given by

$$b(u) = E(\exp(iuY_1)|Y_1 > 0)$$

is the characteristic function of a lattice distribution of span 1. So for $0 < \eta < \pi/10$,

$$\begin{aligned} |E(\exp(iuY_1 + iv\delta_1))| &= |q + pe^{iv} r(u)| \\ &\le q + p|r(u)| < 1 - \gamma \end{aligned}$$

for some $0 < \gamma < 1$, whenever $\eta < |u| < \pi + \eta$. Hence for $|t| \le \log n$, $2\eta \le |s| < \pi$, all ω and for all large n,

$$|f(td_j + s, \omega)| \le 1 - \gamma.$$

Now consider the case $|s| < 2\eta$ and $|v| \ge 2\eta$. Let $0 < \theta < \eta$ be such that $|1 - r(u)| < \eta^2 q$ whenever $|u| < \theta$. Hence as $0 < \eta < 1$,

$$\begin{aligned} |E(\exp(iuY_1 + iv\delta_1))| &= |q + pe^{iv} r(u)| \\ &\le |q + pe^{iv}| + p|1 - r(u)| \\ &\le |q + pe^{iv}| + \eta^2 pq \\ &= (1 - 2pq(1 - \cos u))^{1/2} + \eta^2 pq \\ &= (1 - 2pq(1 - \cos \eta))^{1/2} + \eta^2 pq \\ &= 1 + pq(\cos 2\eta - 1 + \eta^2) \\ &\le 1 - \eta^2 pq \le 1 - \theta^2 pq. \end{aligned}$$

Thus if $|t| \le \log n$ and $|s| < \theta/2$, then $|rd_j + s| < \theta$ for all large n. So if $|\omega| > 2\eta$, $0 \le k \le 2$, then

$$f_{n,k}(t, s, \omega) \le k_3 \rho^n$$

for some $k_3 > 0$ and $0 < \rho < 1$. This completes the proof of the lemma. □

Proof of Proposition 1. Since for any $\varepsilon > 0$,

$$\int_{x-\varepsilon}^{x+\varepsilon} \phi(u)\, dv = O(\varepsilon)$$

uniformly in x, we have by Lemmas 1 and 2, with $\varepsilon = n^{-2}$,

$$\left| nF_n(x, r, m) - \int_{-\infty}^{x} \Psi_n(v, y_r, \omega_m)\, dv \right|$$

$$\leqslant n^{-2} + \max_{0 \le j \le 2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-cn^{5/2}}^{-cn^{5/2}} \Big| D^j\Big(n\hat{f}_n(t, s, \omega)\exp(-isr - i\omega m) - \hat{\Psi}_n(t, s, \omega)\Big)\Big|\, dt\, ds\, d\omega, \tag{A.12}$$

where $\hat{\Psi}_n$ and $\hat{f}_n$ denote the characteristic functions of Ψ_n and $\sum_{i=1}^{n}(n^{-1/2}C_jY_j, Y_j, \delta_j)$, and c is the constant in (A.1). Note that for each r and m, the Fourier transform of $F_n(\cdot, r, m)$ is given by

$$\hat{F}_n(t, r, m) = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \hat{f}_n(t, s, \omega)\exp(-isr - i\omega m)\, ds\, d\omega.$$

To estimate the last integral in (A.12), we divide the range of integration into four possibly overlapping regions:

(i) $|t| \le \log n$, $|s| \le (\log n)n^{-1/2}$, $|\omega| \le (\log n)n^{-1/2}$;
(ii) $|t| \le \log n$, $\omega^2 + s^2 \ge (\log n)^2 n^{-1}$, $|\omega| \le \pi$, $|s| \le \pi$;
(iii) $\log n < |t| \le \sqrt{n}/M$, $|s| \le \pi$ and $|\omega| \le \pi$;
(iv) $(\sqrt{n}/M) \le |t| \le cn^{5/2}$, $|s| \le \pi$ and $|\omega| \le \pi$.

We expand $nD^jf_n(r, s, \omega)$ in region (i) and estimate the integral in (A.12) as in the proof of Theorem 9.9 of Bhattacharya and Ranga Rao (1986). Lemma 6 is used for region (ii), Lemma 5 is used for region (iii) and Lemmas 3 and 4 are used for region (iv) to estimate the integral. These estimates lead to a bound of $0(n^{-3/2})$ for (A.12), completing the proof. □

REFERENCES

BABU, G. J. and BAI, Z. D. (1996). Mixtures of global and local Edgeworth expansions and their applications. *J. Multivariate Anal.* **59** 282–307.

BABU, G. J. and SINGH, K. (1983). Inference on means using the bootstrap. *Ann. Statist.* **11** 999–1003.

BABU, G. J. and SINGH, K. (1984). On the term Edgeworth correction by Efron's bootstrap. *Sankhyā Ser. A* **46** 219–232.

BABU, G. J. and SINGH, K. (1985). Edgeworth expansion for sampling without replacement from finite populations. *J. Multivariate Anal.* **17** 261–278.

BABU, G. J. and SINGH, K. (1989). On Edgeworth expansions in the mixture cases. *Ann. Statist.* **17** 443–447.

BAI, Z. D. and RAO, C. R. (1991). Edgeworth expansion of a function of sample means. *Ann. Statist.* **19** 1295–1315.

BAI, Z. D. and RAO, C. R. (1992). A note on the Edgeworth expansion for ratio of sample means. *Sankhyā Ser. A* **54** 309–322.

BHATTACHARYA, R. N. and GHOSH, J. K. (1978). On the validity of the formal Edgeworth expansion. *Ann. Statist.* **6** 434–451.

BHATTACHARYA, R. N. and RANGA RAO, R. (1986). *Normal Approximations and Asymptotic Expansions*. Krieger, Malabar, FL.

EFRON, B. (1979). Bootstrap methods: another look at the jackknife. *Ann. Statist.* **7** 1–26.

GINÉ, E. and ZINN, J. (1990). Bootstrapping general empirical measures. *Ann. Probab.* **18** 851–869.

HALL, P. and MAMMEN, E. (1994). On general resampling algorithms and their performance in distribution estimation. *Ann. Statist.* **22** 2011–2030.

PATHAK, P. K. (1962). On simple random sampling with replacement. *Sankhyā Ser. A* **24** 287–302.

PETROV, V. V. (1975). *Sums of Independent Random Variables*. Springer, New York.

RAO, C. R., PATHAK, P. K. and KOLTCHINSKII, V. I. (1997). Bootstrap by sequential resampling. *J. Statist. Plann. Inference* **64** 257–281.

SINGH, K. (1981). On the asymptotic accuracy of Efron's bootstrap. *Ann. Statist.* **9** 1187–1195.

SWEETING, T. J. (1977). Speeds of convergence for the multidimensional central limit theorem. *Ann. Probab.* **5** 28–41.

G. J. BABU
C. R. RAO
DEPARTMENT OF STATISTICS
326 THOMAS BUILDING
PENNSYLVANIA STATE UNIVERSITY
UNIVERSITY PARK, PENNSYLVANIA 16802
E-MAIL: babu@stat.psu.edu

P. K. PATHAK
DEPARTMENT OF STATISTICS
AND PROBABILITY
MICHIGAN STATE UNIVERSITY
EAST LANSING, MICHIGAN