United States
Department of
Agriculture

Forest Service

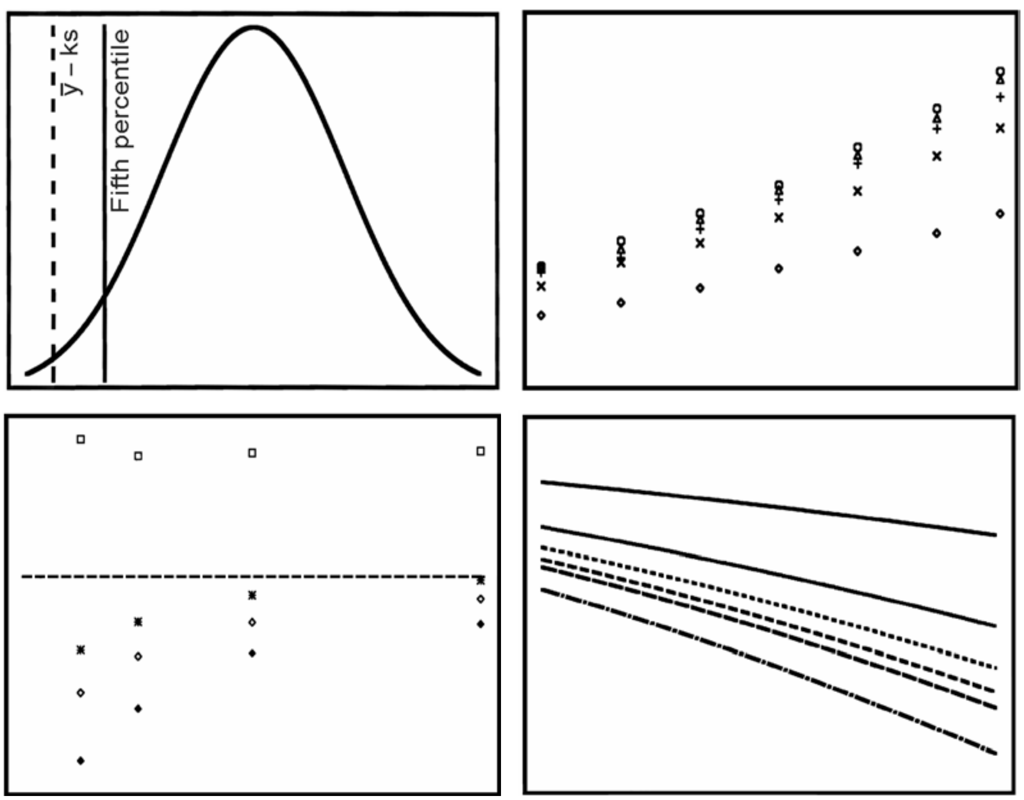
Forest
Products
Laboratory

Research
Paper
FPL–RP–623

Predictor Sort Sampling and Confidence Bounds on Quantiles I

Asymptotic Theory

Steve P. Verrill
Victoria L. Herian
David W. Green



Abstract

In this paper we examine the effect of predictor sort sampling on one-sided confidence bounds for normal quantiles. We have found that standard noncentral T theory that ignores the predictor sort nature of the sampling leads to $\overline{Y} - kS$ bounds that are overly conservative. On the other hand, maximum likelihood methods yield non-conservative bounds even for fairly large sample sizes. We provide an asymptotic result that yields the appropriate corrections for the standard noncentral T approach.

Keywords: concomitants of order statistics, tolerance bounds, asymptotics, allowable strength property

Contents

Page

August 2004

Verrill, Steve P.; Herian, Victoria L.; Green, David W. 2004. Predictor sort sampling and confidence bounds on quantiles I: Asymptotic theory. Res. Pap. FPL-RP-623. Madison, WI: U.S. Department of Agriculture, Forest Service, Forest Products Laboratory. 67 p.

A limited number of free copies of this publication are available to the public from the Forest Products Laboratory, One Gifford Pinchot Drive, Madison, WI 53726–2398. This publication is also available online at www.fpl.fs.fed.us. Laboratory publications are sent to hundreds of libraries in the United States and elsewhere.

The Forest Products Laboratory is maintained in cooperation with the University of Wisconsin.

Predictor Sort Sampling and Confidence Bounds on Quantiles I: Asymptotic Theory

Steve Verrill, Mathematical Statistician
Victoria L. Herian, Statistician
David W. Green, Supervisory Research General Engineer
Forest Products Laboratory, Madison, Wisconsin

1 Introduction

Designers working with lumber must try to ensure that the strengths of wood structural members exceed the loads to which the members will be subjected. One approach to this problem is to design so that expected loads do not exceed "allowable strength properties" associated with particular species and grades of lumber (ASTM D1990). An allowable property is commonly obtained experimentally by taking a sample from the lumber population in question, obtaining a lower one-sided confidence bound on the fifth percentile of the strength distribution of the population, and then dividing by a safety and duration-of-load factor.

If a normal strength distribution is assumed, engineers working with solid-sawn lumber (see, for example, ASTM D2915) can obtain a parametric one-sided lower confidence bound on the fifth percentile via the formula:

$$\bar{Y}_n - k_{n,\alpha,\beta} S_n \tag{1}$$

where we want to cover the α quantile with confidence $\beta \times 100\%$ and we have n replicates. Here $\bar{Y}_n$ denotes the average of n strength measurements, and S_n denotes the sample standard deviation of the measurements. Guttman (1970, Table 4.6) provides k values for $n = 2(1)100$, $(10)300$, $(25)500$, $(50)700$, $(100)1000$, $\alpha = 0.01$, 0.05, 0.10, and 0.25, and $\beta = 0.75$, 0.90, 0.95, and 0.99. He credits Owen (1963) for these tables.

Scientists in other areas (e.g., composite materials, groundwater monitoring, and soil remediation) also make use of formula 1 to obtain confidence bounds on quantiles. See, for example, MIL-HDBK-17-1 (2003), Gibbons (1994), and Michigan DEQ (1994).

For formula 1 to be valid, the sample of lumber (or composite material ...) must be a standard random sample. However, wood strength researchers commonly replace experimental unit allocation via random sampling with allocation via sorts based on non-destructive measurements of strength predictors such as modulus of elasticity and specific gravity. Warren and Madsen (1977) describe the procedure as follows:

> Specifically, then, all the boards in the experiment are ordered from weakest to strongest as nearly as can be judged from their moduli of elasticity, knot size, and slope of grain. To divide the material into J equivalent groups the first J boards, after ordering, are taken and randomly allocated one to each group. This is repeated with the second, third, fourth, etc., sets of J boards. The strength distributions of the resulting groups should then be essentially the same.

At the Forest Products Laboratory, this allocation procedure has come to be known as predictor sort sampling.

In an analysis of variance context, Cox (1957) compared seven procedures that one might use given the availability of a correlated predictor. Cox's calculations showed that the effective variance in these situations is $(1-\rho^2)\sigma_Y^2$, a fact also noted by Cochran (1957). Here σ_Y^2 is the variance of Y, and ρ is the correlation between the predictor X and the response Y.

As noted in Verrill (1993) (also see David and Gunnik 1997), the correlations among the order statistics of the predictor induce correlations among the responses so that the standard analysis of variance (ANOVA) assumptions are not satisfied for a predictor sort experiment. Verrill demonstrated that blocked ANOVAs are still essentially valid and that simply modified unblocked ANOVAs can also be performed on predictor sort data sets.

Verrill (1999) investigated the effects of predictor sort sampling on standard confidence intervals for the mean in an ANOVA context. He found that confidence intervals for the mean are overly conservative (confidence intervals are too wide and actual confidence levels are greater than nominal levels) if an unblocked analysis is performed and non-conservative (confidence intervals are too narrow and actual confidence levels fall below nominal levels) if a blocked analysis is performed. He obtained asymptotic results that yielded correct confidence interval coverage.

In the current paper we examine the effect of predictor sort sampling on one-sided confidence bounds for normal quantiles. We have found that standard noncentral T theory that ignores the predictor sort nature of the sampling leads to $\bar{Y}-kS$ bounds that are too low and thus statistically conservative (actual confidence levels are greater than nominal levels). On the other hand, maximum likelihood methods yield bounds that are too high and thus statistically non-conservative even for fairly large sample sizes. We provide an asymptotic result that yields the appropriate corrections for the standard noncentral T approach.

In a subsequent paper we will provide methods for calculating correct k values for small samples.

2 Poor Confidence Interval Coverage of the Standard Approach Given Predictor Sort Sampling

In Tables 1 to 24 we detail the coverages of four kinds of confidence interval for a variety of combinations of ρ, α, β, number of treatments (J), and number of replicates (n). The rows of the tables are based on separate 4,000 trial simulations. The four approaches that we consider are the incorrect standard approach, which ignores the dependencies induced by sorting on the predictor; two versions of the (correct) predictor sort $\bar{Y}-kS$ asymptotic approach; and a maximum likelihood approach. The maximum likelihood approach is developed in Appendix A. The two versions of the predictor sort approach differ in the estimate used for the correlation between the predictor and the response. Version 1 uses the consistent (see Section A.3) estimate

$$\hat{\rho} \equiv \sum_{j=1}^{J}\sum_{i=1}^{n}(X_{ij}-\bar{X}_{..})(Y_{ij}-\bar{Y}_{.j}) \Big/ \sqrt{\sum_{j=1}^{J}\sum_{i=1}^{n}(X_{ij}-\bar{X}_{..})^2 \sum_{j=1}^{J}\sum_{i=1}^{n}(Y_{ij}-\bar{Y}_{.j})^2}$$

Version 2 uses the maximum likelihood estimate of ρ. It is clear from the tables that the incorrect approach is overly conservative, that the problem becomes more severe as the correlation between the predictor and response variables increases, and that the problem does not vanish as sample sizes increase. It is also clear from the tables that version 2 of the predictor sort approach dominates the maximum likelihood approach in the sense that the actual coverage always approaches the nominal coverage more rapidly for the version 2 predictor sort approach than for the maximum likelihood approach. For smaller J, the version 1 predictor sort approach performs better than the maximum likelihood approach and the version 2 approach (see Figure 1). However, for large J and small n, the version 1 approach does not perform as well.

For smaller n the asymptotic approaches are non-conservative. k values that are appropriate for small sample sizes will appear in a future Forest Products Laboratory research report.

3 Sample Size Reductions Given Predictor Sort Sampling

In the course of the development of the asymptotic theory (see inequality (98)) we find that the correct k in the appropriate version of $\bar{Y} - kS$ is given by

$$k \approx -\Phi^{-1}(\alpha) + \Phi^{-1}(\beta) n^{-1/2} \sqrt{(\Phi^{-1}(\alpha))^2/(2J) + 1 - \rho^2 + \rho^2/J}$$

where Φ denotes the N(0,1) distribution function. Thus given higher ρ values, we can have smaller n values, and still have the same k. In fact if we set

$$n^{-1/2} \sqrt{(\Phi^{-1}(\alpha))^2/(2J) + 1 - \rho^2 + \rho^2/J}$$

equal to a constant we obtain

$$n \propto (\Phi^{-1}(\alpha))^2/(2J) + 1 - \rho^2 + \rho^2/J$$

Thus the approximate permissible sample size reduction factor obtained by using a predictor sort with a correlation of ρ between the predictor and the response is (here the denominator is the numerator with ρ set equal to 0)

$$\left((\Phi^{-1}(\alpha))^2/(2J) + 1 - \rho^2 + \rho^2/J\right) / \left((\Phi^{-1}(\alpha))^2/(2J) + 1\right)$$

Plots of this factor as function of ρ and J are provided in Figure 2. It is clear from the figure that practically significant sample size reductions (e.g., 30%) are attainable for reasonable correlations.

4 Incorrect "Allowable Properties" Given Predictor Sort Sampling and a Non-Predictor Sort Analysis

As noted in Section 1, in lumber strength applications "allowable properties" are calculated as b/f where b is a one-sided lower confidence bound on a fifth percentile and f is some "safety and duration of load factor." If b is too low then the allowable property will be too low. The ratio of the correct to incorrect allowable properties will be approximately equal to

$$r = (1 - k_{\mathrm{c}}\mathrm{CV})/(1 - k_{\mathrm{inc}}\mathrm{CV})$$

where k_{c} is the correct k value, k_{inc} is the incorrect k value, and $\mathrm{CV} \equiv \sigma/\mu$ for the normal distribution under consideration.

In Figures 3 through 5 we plot the ratio r versus the correlation ρ for $J = 2, 4, 6, 8, 10$ and $(n, \mathrm{CV}) = (10, 0.15), (10, 0.25), (20, 0.25)$. Since for small samples k_c must be determined by simulation, there is some irregularity in these curves. However it is clear that r increases as ρ, CV, or J increases, and r decreases as n increases.

In these plots, the ratio r is sometimes as high as 1.15 which is a figure that is large enough to attract the interest of lumber manufacturers. On the other hand, correlations for solid-sawn lumber (between MOR and MOE say) are probably not much greater than 0.70, so permissible increases in any (overly low) allowable properties that were calculated on the basis of predictor sort experiments are probably below 5%.

5 The Theorem that Yields the Asymptotically Correct k Values

Assume that the predictor variable and the variable of interest, Y, have a joint bivariate normal distribution with correlation ρ. Denote the variance of Y by σ_Y^2. Suppose that we have a two-way ANOVA design with n blocks and J treatments, and the allocation of samples is done via a predictor sort as described in Section 1.

Without loss of generality, we focus here on treatment 1. Let $\bar{Y}_{n,1}$, $S_{n,1}$ be defined as in (3) (see the next section), and let $\hat{\rho}$ be any consistent estimator of ρ. Then,

$$\text{Prob}\left(\bar{Y}_{n,1} - \hat{k}_n S_{n,1} \leq \mu_1 + \Phi^{-1}(\alpha)\sigma_Y\right) \to \beta$$

as $n \to \infty$, where μ_1 is the mean response for the first treatment, Φ^{-1} denotes the inverse of a standard normal cumulative distribution function,

$$\hat{k}_n \equiv \sqrt{(1-\hat{\rho}^2+\hat{\rho}^2/J)/n}\ \mathrm{F}^{-1}_{\mathrm{nct},\gamma_n(\hat{\rho}),nJ-1}(\beta),$$

$\mathrm{F}^{-1}_{\mathrm{nct},\gamma_n(\hat{\rho}),nJ-1}$ denotes the inverse of a noncentral T distribution with noncentrality parameter $\gamma_n(\hat{\rho})$ and $nJ-1$ degrees of freedom, and

$$\gamma_n(\hat{\rho}) \equiv -\Phi^{-1}(\alpha)\sqrt{n}(1-\hat{\rho}^2+\hat{\rho}^2/J)^{-1/2}$$

The proof is provided in Appendix B.

6 Heuristic Justification of the Theorem

In the *standard random sampling* case, the derivation of a confidence bound on the αth quantile proceeds as follows:

Let $\bar{Y}_n$ denote the mean of a sample of size n and let S_n denote the corresponding sample standard deviation. We want to find the k_n that satisfies

$$\text{Prob}(\bar{Y}_n - k_n S_n \leq \mu_Y + \Phi^{-1}(\alpha)\sigma_Y) = \beta$$

We have

$$\begin{aligned}
\text{Prob}(\bar{Y}_n - \mu_Y - \Phi^{-1}(\alpha)\sigma_Y \leq k_n S_n) &= \beta \\
\text{Prob}\left((\bar{Y}_n - \mu_Y - \Phi^{-1}(\alpha)\sigma_Y)/(\sigma_Y/\sqrt{n}) \leq k_n (S_n/\sigma_Y)\sqrt{n}\right) &= \beta \\
\text{Prob}(X_{\mathrm{nct},-\Phi^{-1}(\alpha)\sqrt{n},n-1} \leq k_n\sqrt{n}) &= \beta
\end{aligned}$$

where $X_{\mathrm{nct},-\Phi^{-1}(\alpha)\sqrt{n},n-1}$ denotes a random variable with distribution $\mathrm{F}_{\mathrm{nct},-\Phi^{-1}(\alpha)\sqrt{n},n-1}$. This holds if

$$\mathrm{F}^{-1}_{\mathrm{nct},-\Phi^{-1}(\alpha)\sqrt{n},n-1}(\beta) = k_n\sqrt{n}$$

or

$$\mathrm{F}^{-1}_{\mathrm{nct},-\Phi^{-1}(\alpha)\sqrt{n},n-1}(\beta)/\sqrt{n} = k_n \tag{2}$$

How must this derivation be altered in the predictor sort case?

Assume that we have n blocks and J treatments. We can think of a predictor sort specimen allocation in the following manner. A response value, Y, associated with a specimen is given by

$$Y = \mu_Y + \sigma_Y\left(\rho\,(P-\mu_P)/\sigma_P + \sqrt{1-\rho^2}Z\right)$$

where $(P - \mu_P)/\sigma_P$ and Z are independent N(0,1) random variables, and ρ is the correlation between P and Y. Prior to the experiment we have values for P. We rank the specimens on the basis of their associated P values and then randomly allocate the top J specimens to the first block, the next J to the second block, and so on.

Then, for $1 \le i \le n$, $1 \le j \le J$ we have

$$Y_{ij} = \mu_j + \sigma_Y(\rho X_{ij} + \sqrt{1-\rho^2} Z_{ij})$$

where the X_{ij}'s, $1 \le j \le J$, are a randomization of the ith group of order statistics from nJ iid $\mathrm{N}(0,1)$'s, the Z_{ij}'s are iid $\mathrm{N}(0,1)$, and the X's and Z's are independent.

Define

$$\begin{aligned}
W_{ij} &\equiv \rho X_{ij} + \sqrt{1-\rho^2} Z_{ij} \qquad (3)\\
\bar{Y}_{n,1} &\equiv \sum_{i=1}^{n} Y_{i1}/n = \mu_1 + \sigma_Y \sum_{i=1}^{n} W_{i1}/n \\
&= \mu_1 + \sigma_Y \left(\rho \bar{X}_{\cdot 1} + \sqrt{1-\rho^2}\bar{Z}_{\cdot 1}\right) \\
\bar{Y}_{n,2} &\equiv \mu_1 + \sigma_Y \left(\rho \bar{X}_{\cdot\cdot} + \sqrt{1-\rho^2}\bar{Z}_{\cdot 1}\right)
\end{aligned}$$

[Note that $\bar{Y}_{n,1}$ and $\bar{Y}_{n,2}$ are both estimators of μ_1. However, in $\bar{Y}_{n,2}$, an $\bar{X}_{\cdot\cdot}$ replaces the $\bar{X}_{\cdot 1}$ in $\bar{Y}_{n,1}$.]

$$\begin{aligned}
S^2_{n,1}/\sigma^2_Y &\equiv \sum_{j=1}^{J}\sum_{i=1}^{n}(W_{ij} - \bar{W}_{\cdot j})^2/(nJ-1) \\
S^2_{n,2}/\sigma^2_Y &\equiv \sum_{j=1}^{J}\sum_{i=1}^{n}\left(W_{ij} - [\rho\bar{X}_{\cdot\cdot} + \sqrt{1-\rho^2}\bar{Z}_{\cdot j}]\right)^2/(nJ-1) \\
S^2_{n,3}/\sigma^2_Y &\equiv \sum_{j=1}^{J}\sum_{i=1}^{n}\left(W_{ij} - [\rho\bar{X}_{\cdot\cdot} + \sqrt{1-\rho^2}\bar{Z}_{\cdot\cdot}]\right)^2/(nJ-1) \\
&= \sum_{j=1}^{J}\sum_{i=1}^{n}(W_{ij} - \bar{W}_{\cdot\cdot})^2/(nJ-1)
\end{aligned}$$

[Note that $S^2_{n,1}, S^2_{n,2}$, and $S^3_{n,3}$ are all estimators of σ^2_Y. However, in $S^2_{n,2}$, a $\bar{X}_{\cdot\cdot}$ replaces the $\bar{X}_{\cdot j}$ in $S^2_{n,1}$, and in $S^2_{n,3}$, a $\bar{Z}_{\cdot\cdot}$ replaces the $\bar{Z}_{\cdot j}$ in $S^2_{n,2}$.]

$$\begin{aligned}
V_{n,1} &\equiv (\bar{Y}_{n,1} - \mu_1)/\left((\sigma_Y/\sqrt{n})\sqrt{1-\rho^2+\rho^2/J}\right) \\
&= (\rho\bar{X}_{\cdot 1} + \sqrt{1-\rho^2}\bar{Z}_{\cdot 1})/\left(\sqrt{1-\rho^2+\rho^2/J}/\sqrt{n}\right) \\
V_{n,2} &\equiv (\bar{Y}_{n,2} - \mu_1)/\left((\sigma_Y/\sqrt{n})\sqrt{1-\rho^2+\rho^2/J}\right) \\
&= (\rho\bar{X}_{\cdot\cdot} + \sqrt{1-\rho^2}\bar{Z}_{\cdot 1})/\left(\sqrt{1-\rho^2+\rho^2/J}/\sqrt{n}\right) \\
U_{n,1} &\equiv \sqrt{nJ-1}(S_{n,1}/\sigma_Y - 1) \\
U_{n,2} &\equiv \sqrt{nJ-1}(S_{n,2}/\sigma_Y - 1)
\end{aligned}$$

and

$$U_{n,3} \equiv \sqrt{nJ-1}(S_{n,3}/\sigma_Y - 1)$$

Now we want to show that

$$\text{Prob}\left(\bar{Y}_{n,1} - \hat{k}_n S_{n,1} \le \mu_1 + \Phi^{-1}(\alpha)\sigma_Y\right) \to \beta$$

as $n \to \infty$ where

$$\hat{k}_n \equiv \sqrt{(1 - \hat{\rho}^2 + \hat{\rho}^2/J)/n}\ \mathrm{F}^{-1}_{\mathrm{nct},\gamma_n(\hat{\rho}),nJ-1}(\beta)$$

the noncentrality parameter, $\gamma_n(\hat{\rho})$, is given by

$$\gamma_n(\hat{\rho}) = -\Phi^{-1}(\alpha)\sqrt{n}/\sqrt{1 - \hat{\rho}^2 + \hat{\rho}^2/J}$$

and $\hat{\rho}$ is a reasonable estimator of the correlation between X and Y.

When we simply try to follow the steps that lead to equation 2, we quickly run into difficulties. Because of the predictor sort sampling and the resultant correlations among the X_{ij}, $\bar{Y}_{n,1}$ is no longer normally distributed, $(nJ-1)S^2_{n,1}$ is no longer distributed as a chi-squared, and $\bar{Y}_{n,1}$ and $S^2_{n,1}$ are no longer statistically independent. Further $\hat{k}_n$ is now a random variable.

We could try to establish that the ratio

$$(\bar{Y}_{n,1} - \mu_1 - \Phi^{-1}(\alpha)\sigma_Y)/(S_{n,1}/\sqrt{n})$$

is "close to" a noncentral T but this turns out to be difficult.

Instead we can (easily) show that

$$\text{Prob}(\bar{Y}_{n,1} - \hat{k}_n S_{n,1} \le \mu_1 + \Phi^{-1}(\alpha)\sigma_Y) = \text{Prob}(V_{n,1} \le \hat{c}_n + \hat{d}_n U_{n,1})$$

for a certain $\hat{c}_n$ and $\hat{d}_n$ where $\hat{c}_n \xrightarrow{p} c$ and $\hat{d}_n \xrightarrow{p} d$ for a certain c and d (the corollary to Lemma 12). Effectively this permits us to focus on

$$\text{Prob}(V_{n,1} \le c + dU_{n,1})$$

Now $V_{n,1}$ is "quite close" to $V_{n,2}$ (Lemma 3) and $V_{n,2}$ is normally distributed. Also $U_{n,1}$ is quite close to $U_{n,2}$ (Corollary 6.4), and $U_{n,2}$ is independent of $V_{n,2}$ (Lemma 14). $U_{n,2}$ in turn is quite close to $U_{n,3}$ (Corollary 6.3).

In the proof of the main theorem we essentially break the area under the curve $y = c + dx$ into squares (we have to be careful with squares that intersect the line and that adds some messiness to the proof). Because of the closeness of $U_{n,1}$ and $U_{n,2}$ and the closeness of $V_{n,1}$ and $V_{n,2}$, we have

$$\text{Prob}((U_{n,1}, V_{n,1}) \in \text{square}) \approx \text{Prob}((U_{n,2}, V_{n,2}) \in \text{square})$$

Since $U_{n,2}$ is independent of $V_{n,2}$,

$$\text{Prob}((U_{n,2}, V_{n,2}) \in \text{square}) = \text{Prob}(U_{n,2} \in \text{segment}_1) \times \text{Prob}(V_{n,2} \in \text{segment}_2)$$

Since $U_{n,2}$ is "close" to $U_{n,3}$,

$$\begin{aligned} & \text{Prob}(U_{n,2} \in \text{segment}_1) \times \text{Prob}(V_{n,2} \in \text{segment}_2) \\ \approx\ & \text{Prob}(U_{n,3} \in \text{segment}_1) \times \text{Prob}(V_{n,2} \in \text{segment}_2) \end{aligned}$$

Since $V_{n,2} \sim \mathrm{N}(0,1)$ and $U_{n,3} \sim \sqrt{nJ-1}\left(\sqrt{\chi^2_{nJ-1}/(nJ-1)} - 1\right)$, we can sum

$$\text{Prob}(U_{n,3} \in \text{segment}_1) \times \text{Prob}(V_{n,2} \in \text{segment}_2)$$

over all the squares under the $y = c + dx$ line and obtain

$$\mathrm{Prob}(\bar{Y}_{n,1} - \hat{k}_n S_{n,1} \le \mu_1 + \Phi^{-1}(\alpha)\sigma_Y) \approx \mathrm{Prob}(F_n \le \hat{c}_n + \hat{d}_n E_n)$$

where

$$\begin{aligned} E_n &\sim \sqrt{nJ-1}\left(\sqrt{\chi^2_{nJ-1}/(nJ-1)} - 1\right) \\ F_n &\sim \mathrm{N}(0,1) \end{aligned}$$

and E_n, F_n are independent. This comes close to completing the proof.

However, to show that we can use consistent estimates of ρ rather than the actual ρ in calculating $\hat{k}_n$ we also need to establish

$$\mathrm{Prob}\left(X_{\mathrm{nct},\gamma_n(\rho),nJ-1} \le \left(\sqrt{1-\hat{\rho}^2+\hat{\rho}^2/J}/\sqrt{1-\rho^2+\rho^2/J}\right)\mathrm{F}^{-1}_{\mathrm{nct},\gamma_n(\hat{\rho}),nJ-1}(\beta)\right) \to \beta$$

as $n \to \infty$ where $\hat{\rho}$ is a consistent estimator of ρ. This follows from Lemma 12. Lemma 12 in turn depends on Lemma 11. The proof of Lemma 11 is the most painful element of the overall proof. It is painful because we need to be careful in order to prove that a certain convergence is uniform in ρ.

7 Summary

Predictor sort experiments attempt to make use of the correlation between a predictor that can be measured prior to the start of an experiment and the response variable that we are investigating. Properly designed and analyzed, predictor sort experiments can reduce necessary sample sizes, increase statistical power, and reduce the lengths of confidence intervals. However, if the non-random nature of the predictor sort is not taken into account, problems can occur.

In particular, standard one-sided lower confidence bounds on quantiles of a normal distribution are overly conservative in a predictor sort situation. We have developed asymptotic theory that yields the correct k value in the $\bar{Y} - kS$ approach to obtaining a confidence bound. The resulting confidence bounds have coverages that approach the nominal values faster than bounds based on maximum likelihood estimation. In a subsequent paper we will provide k values that are appropriate for small sample sizes.

References

ASTM. (2003), Standard Practice for Establishing Allowable Properties for Visually-Graded Dimension Lumber from In-Grade Tests of Full-Size Specimens. ASTM D1990-00. *Annual Book of ASTM Standards*, Volume 04.10, American Society for Testing and Materials, West Conshohocken, Pennsylvania.

ASTM. (2003), Standard Practice for Evaluating Allowable Properties for Grades of Structural Lumber. ASTM D2915-03. *Annual Book of ASTM Standards*, Volume 04.10, American Society for Testing and Materials, West Conshohocken, Pennsylvania.

Chung, K.L. (1974), *A Course in Probability Theory*, Second Edition, Academic Press, New York.

Cochran, W.G. (1957), Analysis of Covariance: Its Nature and Uses, *Biometrics*, **13**, 261-281.

Cox, D.R. (1957), The Use of a Concomitant Variable in Selecting an Experimental Design, *Biometrika*, **44**, 150-158.

David, H.A. and Gunnik, J.L. (1997), The Paired t Test Under Artificial Pairing, *The American Statistician*, **51**, 9–12.

Gibbons, R.D. (1994), *Statistical Methods for Groundwater Monitoring*, John Wiley and Sons, New York.

Guttman, I. (1970), *Statistical Tolerance Regions: Classical and Bayesian*, Hafner Publishing Company, Darien, Connecticut.

Michigan DEQ (1994), *Guidance Document, Verification of Soil Remediation*, Hazardous Waste Program Section, Waste Management Division, Michigan Department of Environmental Quality, Lansing, Michigan.

MIL-HDBK-17-1 (2003), *Guidelines for Characterization of Structural Materials*, Department of Defense Single Stock Point, Philadelphia, Pennsylvania, `http://www.dodssp.daps.mil/`.

Owen, D.B. (1963), *Factors for one-sided tolerance limits and for variable sampling plans*, Sandia Corporation Monograph No. SCR-607 (19th edn).

Serfling, R.J. (1980), *Approximation Theorems of Mathematical Statistics*, John Wiley and Sons, New York, New York.

Verrill, S.P. (1993), Predictor Sort Sampling, Tight T's, and the Analysis of Covariance, *Journal of the American Statistical Association*, **88**, 119–124.

Verrill, S.P. (1999), When Good Confidence Intervals Go Bad: Predictor Sort Experiments and ANOVA, *The American Statistician*, **53**, 38–42.

Warren, W.G., and Madsen, B. (1977), Computer-Assisted Experimental Design in Forest Products Research: A Case Study Based on Testing the Duration-of-Load Effect, *Forest Products Journal*, **27**(3), 45-50.

Appendix A — Maximum Likelihood Estimation

The log likelihood in a one-way predictor sort is given by (here we neglect the constant term)

$$
\begin{aligned}
&-Jn(\ln(\sigma_X)+\ln(\sigma_Y)) - (Jn/2)\ln(1-\rho^2)\\
&-(1/(2(1-\rho^2)))\left(\sum_{j=1}^{J}\sum_{i=1}^{n}(X_{ij}-\mu_X)^2/\sigma_X^2 + \sum_{j=1}^{J}\sum_{i=1}^{n}(Y_{ij}-\mu_j)^2/\sigma_Y^2\right.\\
&\left.-2\rho\sum_{j=1}^{J}\sum_{i=1}^{n}(X_{ij}-\mu_X)(Y_{ij}-\mu_j)/(\sigma_X\sigma_Y)\right)
\end{aligned}
$$

The heuristic is

$$
\begin{aligned}
&\text{Prob(observed result)}\\
= \;&\text{Prob}(y\text{'s}|x\text{'s and random allocations of the } x\text{'s in the blocks})\\
&\times \text{Prob(random allocations of the } x\text{'s in the blocks)}\\
&\times \text{Prob}(x\text{'s})
\end{aligned}
$$

A.1 Partial Derivatives of the Log Likelihood Function

We have

$$\begin{aligned}
\partial/\partial\mu_X &= (1/(1-\rho^2))\left(\sum_{j=1}^{J}\sum_{i=1}^{n}(X_{ij}-\mu_X)/\sigma_X^2 - \rho\sum_{j=1}^{J}\sum_{i=1}^{n}(Y_{ij}-\mu_j)/(\sigma_X\sigma_Y)\right) && (4)\\
\partial^2/\partial\mu_X^2 &= -Jn/((1-\rho^2)\sigma_X^2)\\
E(\partial^2/\partial\mu_X^2) &= -Jn/((1-\rho^2)\sigma_X^2) && (5)
\end{aligned}$$

$$\begin{aligned}
\partial/\partial\mu_j &= (1/(1-\rho^2))\left(\sum_{i=1}^{n}(Y_{ij}-\mu_j)/\sigma_Y^2 - \rho\sum_{i=1}^{n}(X_{ij}-\mu_X)/(\sigma_X\sigma_Y)\right) && (6)\\
\partial^2/\partial\mu_j^2 &= -n/((1-\rho^2)\sigma_Y^2)\\
E(\partial^2/\partial\mu_j^2) &= -n/((1-\rho^2)\sigma_Y^2) && (7)
\end{aligned}$$

$$\begin{aligned}
\partial^2/\partial\mu_X\partial\mu_j &= \rho n/((1-\rho^2)\sigma_X\sigma_Y)\\
E(\partial^2/\partial\mu_X\partial\mu_j) &= \rho n/((1-\rho^2)\sigma_X\sigma_Y) && (8)
\end{aligned}$$

For $j_1 \neq j_2$,

$$\partial^2/\partial\mu_{j_1}\partial\mu_{j_2} = 0$$

so

$$E(\partial^2/\partial\mu_{j_1}\partial\mu_{j_2}) = 0 \qquad (9)$$

To calculate the expectations below we need to note that, because of the predictor sort sampling, we no longer have

$$E(X_{ij}-\mu_X) = 0$$

However, by a symmetry argument, we do have

$$E\left(\sum_{i=1}^{n}(X_{ij}-\mu_X)\right) = 0$$

Also,

$$Y_{ij}-\mu_j = \frac{\rho\sigma_Y}{\sigma_X}\times(X_{ij}-\mu_X) + \sigma_Y\sqrt{1-\rho^2}Z_{ij}$$

where the X's and Z's are independent and $E(Z_{ij}) = 0$.

$$\begin{aligned}
\partial/\partial\sigma_X &= -Jn/\sigma_X + (1/(1-\rho^2)) && (10)\\
&\quad\times\left(\sum_{j=1}^{J}\sum_{i=1}^{n}(X_{ij}-\mu_X)^2/\sigma_X^3 - \rho\sum_{j=1}^{J}\sum_{i=1}^{n}(X_{ij}-\mu_X)(Y_{ij}-\mu_j)/(\sigma_X^2\sigma_Y)\right)\\
\partial^2/\partial\sigma_X^2 &= Jn/\sigma_X^2 + (1/(1-\rho^2))\\
&\quad\times\left(-3\sum_{j=1}^{J}\sum_{i=1}^{n}(X_{ij}-\mu_X)^2/\sigma_X^4 + 2\rho\sum_{j=1}^{J}\sum_{i=1}^{n}(X_{ij}-\mu_X)(Y_{ij}-\mu_j)/(\sigma_X^3\sigma_Y)\right)\\
E(\partial^2/\partial\sigma_X^2) &= -(Jn/\sigma_X^2)(1+(1/(1-\rho^2))) && (11)
\end{aligned}$$

$$
\begin{aligned}
\partial/\partial\sigma_Y &= -Jn/\sigma_Y + (1/(1-\rho^2)) \\
&\quad \times \left(\sum_{j=1}^{J}\sum_{i=1}^{n}(Y_{ij}-\mu_j)^2/\sigma_Y^3 - \rho\sum_{j=1}^{J}\sum_{i=1}^{n}(X_{ij}-\mu_X)(Y_{ij}-\mu_j)/(\sigma_X\sigma_Y^2) \right) && (12) \\
\partial^2/\partial\sigma_Y^2 &= Jn/\sigma_Y^2 + (1/(1-\rho^2)) \\
&\quad \times \left(-3\sum_{j=1}^{J}\sum_{i=1}^{n}(Y_{ij}-\mu_j)^2/\sigma_Y^4 + 2\rho\sum_{j=1}^{J}\sum_{i=1}^{n}(X_{ij}-\mu_X)(Y_{ij}-\mu_j)/(\sigma_X\sigma_Y^3) \right) \\
E(\partial^2/\partial\sigma_Y^2) &= -(Jn/\sigma_Y^2)(1+(1/(1-\rho^2))) && (13)
\end{aligned}
$$

$$
\begin{aligned}
\partial^2/\partial\mu_X\partial\sigma_X &= (1/(1-\rho^2))\left(-2\sum_{j=1}^{J}\sum_{i=1}^{n}(X_{ij}-\mu_X)/\sigma_X^3 + \rho\sum_{j=1}^{J}\sum_{i=1}^{n}(Y_{ij}-\mu_j)/(\sigma_X^2\sigma_Y) \right) \\
E(\partial^2/\partial\mu_X\partial\sigma_X) &= 0 && (14)
\end{aligned}
$$

$$
\begin{aligned}
\partial^2/\partial\mu_j\partial\sigma_X &= (1/(1-\rho^2))\left(\rho\sum_{i=1}^{n}(X_{ij}-\mu_X)/(\sigma_X^2\sigma_Y) \right) \\
E(\partial^2/\partial\mu_j\partial\sigma_X) &= 0 && (15)
\end{aligned}
$$

$$
\begin{aligned}
\partial^2/\partial\mu_X\partial\sigma_Y &= (1/(1-\rho^2))\left(\rho\sum_{j=1}^{J}\sum_{i=1}^{n}(Y_{ij}-\mu_j)/(\sigma_X\sigma_Y^2) \right) \\
E(\partial^2/\partial\mu_X\partial\sigma_Y) &= 0 && (16)
\end{aligned}
$$

$$
\begin{aligned}
\partial^2/\partial\mu_j\partial\sigma_Y &= (1/(1-\rho^2))\left(-2\sum_{i=1}^{n}(Y_{ij}-\mu_j)/\sigma_Y^3 + \rho\sum_{i=1}^{n}(X_{ij}-\mu_X)/(\sigma_X\sigma_Y^2) \right) \\
E(\partial^2/\partial\mu_j\partial\sigma_Y) &= 0 && (17)
\end{aligned}
$$

$$
\begin{aligned}
\partial^2/\partial\sigma_X\partial\sigma_Y &= (\rho/(1-\rho^2))\left(\sum_{j=1}^{J}\sum_{i=1}^{n}(X_{ij}-\mu_X)(Y_{ij}-\mu_j)/(\sigma_X^2\sigma_Y^2) \right) \\
E(\partial^2/\partial\sigma_X\partial\sigma_Y) &= Jn\rho^2/((1-\rho^2)\sigma_X\sigma_Y) && (18)
\end{aligned}
$$

$$
\partial/\partial\rho = Jn\rho/(1-\rho^2) + (1/(1-\rho^2))\left(\sum_{j=1}^{J}\sum_{i=1}^{n}(X_{ij}-\mu_X)(Y_{ij}-\mu_j)/(\sigma_X\sigma_Y) \right) - (\rho/(1-\rho^2)^2)\times S \quad (19)
$$

where

$$S \equiv \left(\sum_{j=1}^{J} \sum_{i=1}^{n} (X_{ij} - \mu_X)^2/\sigma_X^2 + \sum_{j=1}^{J} \sum_{i=1}^{n} (Y_{ij} - \mu_j)^2/\sigma_Y^2 - 2\rho \sum_{j=1}^{J} \sum_{i=1}^{n} (X_{ij} - \mu_X)(Y_{ij} - \mu_j)/(\sigma_X\sigma_Y) \right)$$

$$\begin{aligned} \partial^2/\partial\rho^2 &= Jn(1+\rho^2)/(1-\rho^2)^2 + \left(4\rho/(1-\rho^2)^2\right) \sum_{j=1}^{J} \sum_{i=1}^{n} (X_{ij} - \mu_X)(Y_{ij} - \mu_j)/(\sigma_X\sigma_Y) \\ &\quad - \left(1/(1-\rho^2)^2 + 4\rho^2/(1-\rho^2)^3\right) \times S \\ E(\partial^2/\partial\rho^2) &= -Jn(1+\rho^2)/(1-\rho^2)^2 \end{aligned} \tag{20}$$

$$\begin{aligned} \partial^2/\partial\mu_X\partial\rho &= -1/(1-\rho^2) \sum_{j=1}^{J} \sum_{i=1}^{n} (Y_{ij} - \mu_j)/(\sigma_X\sigma_Y) \\ &\quad + 2\rho/(1-\rho^2)^2 \left(\sum_{j=1}^{J} \sum_{i=1}^{n} (X_{ij} - \mu_X)/\sigma_X^2 - \rho \sum_{j=1}^{J} \sum_{i=1}^{n} (Y_{ij} - \mu_j)/(\sigma_X\sigma_Y) \right) \\ E(\partial^2/\partial\mu_X\partial\rho) &= 0 \end{aligned} \tag{21}$$

$$\begin{aligned} \partial^2/\partial\mu_j\partial\rho &= -1/(1-\rho^2) \sum_{i=1}^{n} (X_{ij} - \mu_X)/(\sigma_X\sigma_Y) \\ &\quad + 2\rho/(1-\rho^2)^2 \left(\sum_{i=1}^{n} (Y_{ij} - \mu_j)/\sigma_Y^2 - \rho \sum_{i=1}^{n} (X_{ij} - \mu_X)/(\sigma_X\sigma_Y) \right) \\ E(\partial^2/\partial\mu_j\partial\rho) &= 0 \end{aligned} \tag{22}$$

$$\begin{aligned} \partial^2/\partial\sigma_X\partial\rho &= -1/(1-\rho^2) \sum_{j=1}^{J} \sum_{i=1}^{n} (X_{ij} - \mu_X)(Y_{ij} - \mu_j)/(\sigma_X^2\sigma_Y) \\ &\quad + 2\rho/(1-\rho^2)^2 \left(\sum_{j=1}^{J} \sum_{i=1}^{n} (X_{ij} - \mu_X)^2/\sigma_X^3 - \rho \sum_{j=1}^{J} \sum_{i=1}^{n} (X_{ij} - \mu_X)(Y_{ij} - \mu_j)/(\sigma_X^2\sigma_Y) \right) \\ E(\partial^2/\partial\sigma_X\partial\rho) &= \rho Jn/((1-\rho^2)\sigma_X) \end{aligned} \tag{23}$$

$$\begin{aligned} \partial^2/\partial\sigma_Y\partial\rho &= -1/(1-\rho^2) \sum_{j=1}^{J} \sum_{i=1}^{n} (X_{ij} - \mu_X)(Y_{ij} - \mu_j)/(\sigma_X\sigma_Y^2) \\ &\quad + 2\rho/(1-\rho^2)^2 \left(\sum_{j=1}^{J} \sum_{i=1}^{n} (Y_{ij} - \mu_j)^2/\sigma_Y^3 - \rho \sum_{j=1}^{J} \sum_{i=1}^{n} (X_{ij} - \mu_X)(Y_{ij} - \mu_j)/(\sigma_X\sigma_Y^2) \right) \\ E(\partial^2/\partial\sigma_Y\partial\rho) &= \rho Jn/((1-\rho^2)\sigma_Y) \end{aligned} \tag{24}$$

A.2 Maximum Likelihood Estimates

Setting $\partial/\partial\mu_X$ (equation (4)) to 0 we obtain

$$\hat{\mu}_X = \bar{X}_{..} - \hat{\rho}(\hat{\sigma}_X/\hat{\sigma}_Y)\sum_{j=1}^{J}(\bar{Y}_{.j} - \hat{\mu}_j)/J \tag{25}$$

For $j = 1, \dots, J$, setting $\partial/\partial\mu_j$ (equation (6)) to 0 we obtain

$$\hat{\mu}_j = \bar{Y}_{.j} - \hat{\rho}(\hat{\sigma}_Y/\hat{\sigma}_X)(\bar{X}_{.j} - \hat{\mu}_X) \tag{26}$$

or

$$\bar{Y}_{.j} - \hat{\mu}_j = \hat{\rho}(\hat{\sigma}_Y/\hat{\sigma}_X)(\bar{X}_{.j} - \hat{\mu}_X) \tag{27}$$

Combining (25) and (27) we obtain (for $\hat{\rho}^2 \neq 1$)

$$\hat{\mu}_X = \bar{X}_{..} \tag{28}$$

Now define

$$A \equiv \sum_{j=1}^{J}\sum_{i=1}^{n}(X_{ij} - \bar{X}_{..})^2/\hat{\sigma}_X^2 \tag{29}$$

$$B \equiv \sum_{j=1}^{J}\sum_{i=1}^{n}(Y_{ij} - \hat{\mu}_j)^2/\hat{\sigma}_Y^2 \tag{30}$$

and

$$C \equiv \sum_{j=1}^{J}\sum_{i=1}^{n}(X_{ij} - \bar{X}_{..})(Y_{ij} - \hat{\mu}_j)/(\hat{\sigma}_X\hat{\sigma}_Y) \tag{31}$$

Setting $\partial/\partial\sigma_X$ (equation (10) to 0 we obtain

$$Jn = (1/(1-\hat{\rho}^2))(A - \hat{\rho}C) \tag{32}$$

Setting $\partial/\partial\sigma_Y$ (equation (12) to 0 we obtain

$$Jn = (1/(1-\hat{\rho}^2))(B - \hat{\rho}C) \tag{33}$$

Setting $\partial/\partial\rho$ (19) to 0 we obtain

$$0 = \frac{Jn\hat{\rho}}{1-\hat{\rho}^2} + \frac{1}{1-\hat{\rho}^2} \times C - \frac{\hat{\rho}}{(1-\hat{\rho}^2)^2} \times (A + B - 2\hat{\rho}C) \tag{34}$$

Applying (32) and (33) to (34) we then obtain

$$0 = \frac{Jn\hat{\rho}}{1-\hat{\rho}^2} + \frac{1}{1-\hat{\rho}^2} \times C - \frac{\hat{\rho}}{(1-\hat{\rho}^2)^2} \times (2(1-\hat{\rho}^2)Jn)$$

or, after some algebra,

$$\hat{\rho} = \frac{C}{Jn} = \sum_{j=1}^{J}\sum_{i=1}^{n}(X_{ij} - \bar{X}_{..})(Y_{ij} - \hat{\mu}_j)/(Jn\hat{\sigma}_X\hat{\sigma}_Y) \tag{35}$$

From (32) and (35) we have

$$Jn = (1/(1-\hat{\rho}^2))(A - \hat{\rho}^2 Jn)$$

or

$$Jn - \hat{\rho}^2 Jn = A - \hat{\rho}^2 Jn$$

or

$$\hat{\sigma}_X^2 = \frac{\sum_{j=1}^{J}\sum_{i=1}^{n}(X_{ij} - \bar{X}_{..})^2}{Jn} \tag{36}$$

Similarly, from (33) and (35) we have

$$\hat{\sigma}_Y^2 = \frac{\sum_{j=1}^{J}\sum_{i=1}^{n}(Y_{ij} - \hat{\mu}_j)^2}{Jn} \tag{37}$$

Using (26) and (35) we obtain

$$\hat{\rho} = \sum_{j=1}^{J}\sum_{i=1}^{n}(X_{ij} - \bar{X}_{..})(Y_{ij} - \bar{Y}_{.j})/(Jn\hat{\sigma}_X\hat{\sigma}_Y(1-D)) \tag{38}$$

where

$$D \equiv \sum_{j=1}^{J}(\bar{X}_{.j} - \bar{X}_{..})^2/(J\hat{\sigma}_X^2) \tag{39}$$

From (26) and (37) we obtain

$$\hat{\sigma}_Y^2 = \sum_{j=1}^{J}\sum_{i=1}^{n}(Y_{ij} - \bar{Y}_{.j})^2/(Jn(1-\hat{\rho}^2 D)) \tag{40}$$

Solving for $\hat{\sigma}_Y^2$ in (38) and setting the result equal to (40), we obtain

$$\hat{\rho}^2 = E/((F_X F_Y (1-D)^2) + ED) \tag{41}$$

where

$$\begin{aligned}
E &\equiv \left(\sum_{j=1}^{J}\sum_{i=1}^{n}(X_{ij} - \bar{X}_{..})(Y_{ij} - \bar{Y}_{.j})\right)^2 \\
F_X &\equiv \sum_{j=1}^{J}\sum_{i=1}^{n}(X_{ij} - \bar{X}_{..})^2 \\
F_Y &\equiv \sum_{j=1}^{J}\sum_{i=1}^{n}(Y_{ij} - \bar{Y}_{.j})^2
\end{aligned}$$

and D is defined by (39).

$\hat{\mu}_X$, $\hat{\sigma}_X$, and $\hat{\rho}^2$ are given by (28), (36), and (41). Then $\hat{\sigma}_Y^2$ is given by (40), $\hat{\rho}$ by (38), and $\hat{\mu}_j$ by (26).

A.3 Proof that a Naive Estimator of ρ Is Consistent

Lemma

$$\Delta \equiv r_n^2 - \hat{\rho}^2_{\text{MLE}} \xrightarrow{\rho} 0$$

where

$$r_n^2 \equiv \left(\sum_{j=1}^{J} \sum_{i=1}^{n} (X_{ij} - \bar{X}_{\cdot\cdot})(Y_{ij} - \bar{Y}_{\cdot j}) \right)^2 \Big/ \left(\sum_{j=1}^{J} \sum_{i=1}^{n} (X_{ij} - \bar{X}_{\cdot\cdot})^2 \sum_{j=1}^{J} \sum_{i=1}^{n} (Y_{ij} - \bar{Y}_{\cdot j})^2 \right)$$

Proof:
In the notation of the preceding section

$$\begin{aligned}
\Delta &= E/(F_X F_Y) - E/\big(F_X F_Y (1-D)^2 + ED\big) \\
&= \big(E(F_X F_Y (1-D)^2 + ED) - E(F_X F_Y)\big) \big/ \big(F_X^2 F_Y^2 (1-D)^2 + F_X F_Y ED\big) \\
&= \big(E F_X F_Y (-2D + D^2) + E^2 D\big) \big/ \big(F_X^2 F_Y^2 (1-D)^2 + F_X F_Y ED\big) \qquad (42)
\end{aligned}$$

Dividing the numerator and the denominator in (42) by $F_X^2 F_Y^2$ we obtain

$$\Delta = \big((E/(F_X F_Y))(-2D + D^2) + (E/(F_X F_Y))^2 D\big) \big/ \big((1-D)^2 + (E/(F_X F_Y))D\big) \qquad (43)$$

Now Cauchy-Schwarz implies that $E/(F_X F_Y)$ is bounded by 1, and by Lemma 1 in Appendix B, $D \xrightarrow{p} 0$, so the lemma follows from (43).

A.4 One-Sided, Lower Maximum Likelihood Confidence Interval on a Quantile

A one-sided lower $\beta \times 100$ percent maximum likelihood confidence interval on an α quantile for the jth treatment population is

$$\hat{\mu}_j + \Phi^{-1}(\alpha)\hat{\sigma}_Y - z_\beta \sqrt{v_{11} + 2\Phi^{-1}(\alpha) v_{12} + (\Phi^{-1}(\alpha))^2 v_{22}}$$

where $\hat{\mu}_j$ is given by (26), $\hat{\sigma}_Y$ is given by (40), $z_\beta = \Phi^{-1}(\beta)$ and Φ is the $\mathrm{N}(0,1)$ cumulative distribution function, v_{11} is the asymptotic variance for $\hat{\mu}_j$, v_{22} is the asymptotic variance for $\hat{\sigma}_Y$, and v_{12} is the asymptotic covariance for $\hat{\mu}_j$, $\hat{\sigma}_Y$. Approximate values for v_{11}, v_{22}, and v_{12} can be obtained from the negative inverse of the matrix of expectations of second partials (whose elements are given by (5), (7), (8), (9), (11), (13), (14) – (18), (20) – (24)) evaluated at the maximum likelihood estimates.

Appendix B — Proofs

Assume that there are J levels of the factor and n "replicates" per level.
For $1 \le i \le n$, $1 \le j \le J$ we have

$$Y_{ij} = \mu_j + \sigma_Y (\rho X_{ij} + \sqrt{1-\rho^2} Z_{ij})$$

where the X_{ij}'s, $1 \le j \le J$ are a randomization of the ith group of order statistics from nJ iid $\mathrm{N}(0,1)$'s, the Z_{ij}'s are iid $\mathrm{N}(0,1)$, and the X's and Z's are independent.
Define W_{ij}, $\bar{Y}_{n,1}$, $\bar{Y}_{n,2}$, $S^2_{n,1}$, $S^2_{n,2}$, $S^2_{n,3}$, $V_{n,1}$, $V_{n,2}$, $U_{n,1}$, $U_{n,2}$, and $U_{n,3}$ as in (3).
Before proving the main result we first need to establish a series of lemmas.

Lemma 1

$$\sqrt{n}(\bar{X}_{\cdot j} - \bar{X}_{\cdot\cdot}) \xrightarrow{p} 0$$

as $n \to \infty$, j fixed.

Note that this lemma is at the heart of the matter. In the case of standard random sampling, we would have

$$\sqrt{n}(\bar{X}_{\cdot j} - \bar{X}_{\cdot\cdot}) \sim \mathrm{N}(0, (J-1)/J)$$

However in the predictor sort case, since each $\bar{X}_{\cdot j}$ is guaranteed to include one representative from each of the n blocks of adjacent order statistics, the $\bar{X}_{\cdot j}$'s tend to be more similar than in the case of standard random sampling, and Lemma 1 results.

Proof:

$$\begin{aligned}
|\sqrt{n}(\bar{X}_{\cdot j} - \bar{X}_{\cdot\cdot})| &= \sqrt{n}|X_{1j} - (\text{average of first block of X's}) + \ldots + \\
&\quad X_{nj} - (\text{average of } n\text{th block of X's})|/n \\
&\le (\text{largest in first block} - \text{smallest in first block} + \ldots \\
&\quad + \text{largest in } n\text{th block} - \text{smallest in } n\text{th block})/\sqrt{n} \\
&\le (X_{\max} - X_{\min})/\sqrt{n} \xrightarrow{p} 0
\end{aligned}$$

by the lemma in the appendix of Verrill (1993).

Lemma 2

$$V_{n,2} \sim \mathrm{N}(0,1)$$

and

$$U_{n,3} \sim \sqrt{nJ-1}\left(\sqrt{\chi^2_{nJ-1}/(nJ-1)} - 1\right)$$

Proof: Clear.

Lemma 3

$$\sqrt{n}(\bar{Y}_{n,1} - \bar{Y}_{n,2}) \xrightarrow{p} 0$$

Proof:

$$\sqrt{n}\left(\rho\bar{X}_{\cdot 1} + \sqrt{1-\rho^2}\bar{Z}_{\cdot 1} - (\rho\bar{X}_{\cdot\cdot} + \sqrt{1-\rho^2}\bar{Z}_{\cdot 1})\right) = \rho\sqrt{n}(\bar{X}_{\cdot 1} - \bar{X}_{\cdot\cdot})$$

which converges in probability to 0 by Lemma 1.

Lemma 4

$$\sqrt{nJ-1}(S^2_{n,1} - S^2_{n,3}) \xrightarrow{p} 0$$

Proof:

$$\begin{aligned}
\sqrt{nJ-1}\, S^2_{n,1}/\sigma^2_Y &= \sqrt{nJ-1}\sum_{j=1}^{J}\sum_{i=1}^{n}(W_{ij} - \bar{W}_{\cdot j})^2/(nJ-1) \\
&= \sqrt{nJ-1}\sum_{j=1}^{J}\sum_{i=1}^{n}(W_{ij} - \bar{W}_{\cdot\cdot} + \bar{W}_{\cdot\cdot} - \bar{W}_{\cdot j})^2/(nJ-1) \\
&= \sqrt{nJ-1}\left(\sum_{j=1}^{J}\sum_{i=1}^{n}(W_{ij} - \bar{W}_{\cdot\cdot})^2 - \sum_{j=1}^{J} n(\bar{W}_{\cdot j} - \bar{W}_{\cdot\cdot})^2\right)/(nJ-1)
\end{aligned}$$

Thus,

$$\sqrt{nJ-1}(S_{n,1}^2 - S_{n,3}^2)/\sigma_Y^2 = -\left(\sqrt{nJ-1}/(nJ-1)\right) \times \sum_{j=1}^{J} n(\bar{W}_{\cdot j} - \bar{W}_{\cdot\cdot})^2$$

which converges in probability to 0 by A.1 in Verrill (1993).

Lemma 5

$$\sqrt{nJ-1}(S_{n,2}^2 - S_{n,3}^2) \xrightarrow{p} 0$$

Proof:

$$\begin{aligned}
\sqrt{nJ-1}\, S_{n,2}^2/\sigma_Y^2 &= \sqrt{nJ-1}\left(\sum_{j=1}^{J}\sum_{i=1}^{n}\left(\rho X_{ij} + \sqrt{1-\rho^2}Z_{ij} - \left(\rho\bar{X}_{\cdot\cdot} + \sqrt{1-\rho^2}\bar{Z}_{\cdot j}\right)\right)^2\right)/(nJ-1) \\
&= \sqrt{nJ-1}\left(\sum_{j=1}^{J}\sum_{i=1}^{n}\left(\rho X_{ij} + \sqrt{1-\rho^2}Z_{ij} - \left(\rho\bar{X}_{\cdot\cdot} + \sqrt{1-\rho^2}\bar{Z}_{\cdot\cdot}\right)\right.\right. \\
&\quad \left.\left. + \left(\rho\bar{X}_{\cdot\cdot} + \sqrt{1-\rho^2}\bar{Z}_{\cdot\cdot}\right) - \left(\rho\bar{X}_{\cdot\cdot} + \sqrt{1-\rho^2}\bar{Z}_{\cdot j}\right)\right)^2\right)/(nJ-1) \\
&= \sqrt{nJ-1}\left(\sum_{j=1}^{J}\sum_{i=1}^{n}\left(W_{ij} - \bar{W}_{\cdot\cdot} + \sqrt{1-\rho^2}(\bar{Z}_{\cdot\cdot} - \bar{Z}_{\cdot j})\right)^2\right)/(nJ-1) \\
&= \sqrt{nJ-1}\sum_{j=1}^{J}\sum_{i=1}^{n}(W_{ij} - \bar{W}_{\cdot\cdot})^2/(nJ-1) \\
&\quad + 2\sqrt{1-\rho^2}\sqrt{nJ-1}\sum_{j=1}^{J}\sum_{i=1}^{n}(W_{ij} - \bar{W}_{\cdot\cdot})(\bar{Z}_{\cdot\cdot} - \bar{Z}_{\cdot j})/(nJ-1) \\
&\quad + (1-\rho^2)\sqrt{nJ-1}\sum_{j=1}^{J}\sum_{i=1}^{n}(\bar{Z}_{\cdot j} - \bar{Z}_{\cdot\cdot})^2/(nJ-1)
\end{aligned}$$

Thus,

$$\begin{aligned}
&\sqrt{nJ-1}(S_{n,2}^2 - S_{n,3}^2)/\sigma_Y^2 \qquad (44)\\
&= \sqrt{nJ-1}\left(2\sqrt{1-\rho^2}\sum_{j=1}^{J}\sum_{i=1}^{n}(W_{ij} - \bar{W}_{\cdot\cdot})(\bar{Z}_{\cdot\cdot} - \bar{Z}_{\cdot j}) + (1-\rho^2)\sum_{j=1}^{J}\sum_{i=1}^{n}(\bar{Z}_{\cdot j} - \bar{Z}_{\cdot\cdot})^2\right)/(nJ-1) \\
&= \sqrt{nJ-1}\left(2\sqrt{1-\rho^2}\sum_{j=1}^{J}\sum_{i=1}^{n}\left(\rho(X_{ij} - \bar{X}_{\cdot\cdot}) + \sqrt{1-\rho^2}(Z_{ij} - \bar{Z}_{\cdot\cdot})\right)(\bar{Z}_{\cdot\cdot} - \bar{Z}_{\cdot j})\right. \\
&\quad \left. + (1-\rho^2)\sum_{j=1}^{J} n(\bar{Z}_{\cdot j} - \bar{Z}_{\cdot\cdot})^2\right)/(nJ-1) \\
&= \sqrt{nJ-1}\left(2\sqrt{1-\rho^2}\sum_{j=1}^{J}\sum_{i=1}^{n}\rho(X_{ij} - \bar{X}_{\cdot\cdot})(\bar{Z}_{\cdot\cdot} - \bar{Z}_{\cdot j}) - (1-\rho^2)\sum_{j=1}^{J} n(\bar{Z}_{\cdot j} - \bar{Z}_{\cdot\cdot})^2\right)/(nJ-1)
\end{aligned}$$

Consider the first term in (44) (here we neglect constant multipliers):

$$\sum_{j=1}^{J} \left((\bar{Z}_{\cdot\cdot} - \bar{Z}_{\cdot j}) \times n(\bar{X}_{\cdot j} - \bar{X}_{\cdot\cdot}) \right) / \sqrt{nJ-1}$$

By Lemma 1 this converges in probability to 0. Since

$$\sum_{j=1}^{J} n(\bar{Z}_{\cdot j} - \bar{Z}_{\cdot\cdot})^2$$

is distributed as a chi-squared random variable with $J-1$ degrees of freedom, the second term in (44) also converges in probability to zero. This completes the proof.

Lemma 6

$$\sqrt{nJ-1}(S_{n,3}^2/\sigma_Y^2 - 1) \xrightarrow{D} \mathrm{N}(0,2)$$

Proof: Since $S_{n,3}^2/\sigma_Y^2 \sim \chi_{nJ-1}^2/(nJ-1)$, this is an immediate consequence of the central limit theorem.

Corollary 6.1

$$S_{n,3} \xrightarrow{p} \sigma_Y$$

Corollary 6.2

$$S_{n,1} \xrightarrow{p} \sigma_Y$$
$$S_{n,2} \xrightarrow{p} \sigma_Y$$

Proof: The corollary follows from Lemmas 4 and 5 and Corollary 6.1.

Corollary 6.3

$$\sqrt{nJ-1}(S_{n,1} - S_{n,3}) \xrightarrow{p} 0$$
$$\sqrt{nJ-1}(S_{n,2} - S_{n,3}) \xrightarrow{p} 0$$

Proof:

$$\sqrt{nJ-1}(S_{n,1} - S_{n,3}) = \sqrt{nJ-1}(S_{n,1}^2 - S_{n,3}^2)/(S_{n,1} + S_{n,3})$$

which converges in probability to 0 by Lemma 4 and Corollaries 6.1 and 6.2. The second part of the corollary follows in a similar fashion from Lemma 5 and Corollaries 6.1 and 6.2.

Corollary 6.4

$$\sqrt{nJ-1}(S_{n,1} - S_{n,2}) \xrightarrow{p} 0$$

Corollary 6.5

$$\sqrt{nJ-1}(S_{n,3}/\sigma_Y - 1) \xrightarrow{D} \mathrm{N}(0,1/2)$$

Proof:

$$\begin{aligned} \sqrt{nJ-1}(S_{n,3}/\sigma_Y - 1) &= \sqrt{nJ-1}\left((S_{n,3} - \sigma_Y)/\sigma_Y\right)(S_{n,3} + \sigma_Y)/(S_{n,3} + \sigma_Y) \\ &= \sqrt{nJ-1}(S_{n,3}^2/\sigma_Y^2 - 1)\sigma_Y/(S_{n,3} + \sigma_Y) \end{aligned}$$

which converges in distribution to a $\mathrm{N}(0, 1/2)$ by Lemma 6 and Corollary 6.1.

Lemma 7

Let R be an $r \times r$ square on the plane and let $\epsilon > 0$ be given. Then we can divide R into a $k \times k$ grid of subsquares such that given any line on the plane, the sum of the areas of the subsquares that intersect the line is less than ϵ.

Proof: $s \equiv r/k$ is the length of a side of a subsquare. Since the greatest distance between any two points in a subsquare is $\sqrt{2}s$, a subsquare can intersect a line only if all of its points lie within $\sqrt{2}s$ of the line. Thus all of the subsquares that intersect a line must lie within a rectangle of length $l + 2\sqrt{2}s$ and width $2\sqrt{2}s$ where l is the length of the intersection of the line with the square R. It is clear that $l \leq \sqrt{2}r$. Thus the area of the bounding rectangle is less than or equal to

$$(\sqrt{2}r + 2\sqrt{2}r/k)(2\sqrt{2}r/k)$$

and we can make this less than ϵ by making k large enough.

Lemma 8

Let $R = [q, q+r] \times [s, s+r]$ be an $r \times r$ square on the x, y plane with two of its sides parallel to the x axis. Let $\epsilon > 0$ be given. Let $y = c + dx$ be any line on the plane that intersects R. Suppose that $\hat{c}_n \xrightarrow{p} c$ and $\hat{d}_n \xrightarrow{p} d$. Then, given any $\delta > 0$ we can divide R into a $k \times k$ grid of subsquares, and find a fixed collection C_1 of the subsquares of area less than δ, and an N such that $n > N$ implies that with probability greater than $1 - \epsilon$, all the subsquares of R that intersect the line $y = \hat{c}_n + \hat{d}_n x$ lie in C_1.

Proof: Let $\epsilon > 0$ and $\delta > 0$ be given. Let k be chosen so that $4r^2/k < \delta/2$. Let $\hat{C}_{n,1}$ denote the collection of subsquares of R that intersect the line $y = \hat{c}_n + \hat{d}_n x$. Since $\hat{c}_n \xrightarrow{p} c$ and $\hat{d}_n \xrightarrow{p} d$ we can find an N such that $n > N$ implies

$$\text{Prob}\left(|\hat{c}_n + \hat{d}_n x - (c + dx)| < \delta/(2\sqrt{2}r)\ \forall x \in [q, q+r]\right) > 1 - \epsilon$$

It is clear that the perpendicular distance from a point in a subsquare that intersects $y = \hat{c}_n + \hat{d}_n x$ to the line $y = c + dx$ can be at most $\delta/(2\sqrt{2}r) + \sqrt{2}r/k$ when $|\hat{c}_n + \hat{d}_n x - (c + dx)| < \delta/(2\sqrt{2}r)$. Assume that $d \geq 0$. The proof is essentially the same in the other case. We can place lines, L_1 and L_2, that are perpendicular to $y = c + dx$ through the lower left and upper right corners of R. It is clear that lines L_1 and L_2 together with the two lines that are parallel to and a perpendicular distance of $\delta/(2\sqrt{2}r) + \sqrt{2}r/k$ from the line $y = c + dx$ form a rectangle R_0 that covers all the subsquares of R that intersect the line $y = \hat{c}_n + \hat{d}_n x$ when $|\hat{c}_n + \hat{d}_n x - (c + dx)| < \delta/(2\sqrt{2}r)$. Now let C_1 be the collection of all the subsquares of R that intersect R_0. Let R_1 denote the region covered by the elements of C_1. It is clear that L_1 and L_2 together with the two lines that are parallel to and a perpendicular distance of $\delta/(2\sqrt{2}r) + 2\sqrt{2}r/k$ from the line $y = c + dx$ form a rectangle R_2 that covers R_1. Further it is clear that the length of this rectangle is at most $\sqrt{2}r$.

Since the area of R_1 is less than or equal to the area of R_2 which is at most $\sqrt{2}r(\delta/(2\sqrt{2}r) + 2\sqrt{2}r/k) < \delta$, and for $n > N$

$$\text{Prob}(\hat{C}_{n,1} \subset C_1) \geq \text{Prob}(|\hat{c}_n + \hat{d}_n x - (c + dx)| < \delta/(2\sqrt{2}r)) > 1 - \epsilon$$

the proof is complete.

Corollary 8.1

Let E_0 denote the event

$$\{\text{all the subsquares of } R \text{ that lie strictly below } y = c + dx$$

$$\text{but not strictly below } y = \hat{c}_n + \hat{d}_n x \text{ lie in } C_1\}$$

Then for $n > N$, $\text{prob}(E_0) > 1 - \epsilon$.

Proof:

Let ϵ, δ be given and the subsquares constructed and N chosen as in the proof of Lemma 8. Let E_1 denote the event

$$\{|\hat{c}_n + \hat{d}_n x - (c + dx)| < \delta/(2\sqrt{2}r) \ \forall x \in [q, q + r]\}$$

Now let S be any subsquare that lies strictly below $y = c + dx$ but not strictly below $y = \hat{c}_n + \hat{d}_n x$. Since S does not lie strictly below $y = \hat{c}_n + \hat{d}_n x$, it must contain a point (x_0, y_0) such that $y_0 \geq \hat{c}_n + \hat{d}_n x_0$. Now if E_1 has occurred, (x_0, y_0) cannot lie more than a perpendicular distance of $\delta/(2\sqrt{2}r)$ below the line $y = c + dx$. Thus the points of S must lie below the line $y = c + dx$ but not more than $\delta/(2\sqrt{2}r) + \sqrt{2}r/k$ perpendicular units below the line. Thus S must intersect the rectangle R_0 defined in the proof of Lemma 8, so $S \in C_1$. As S was an arbitrary subsquare that lies strictly below $y = c + dx$ but not strictly below $y = \hat{c}_n + \hat{d}_n x$, we have $E_1 \subset E_0$, so

$$\text{Prob}(E_0) \geq \text{Prob}(E_1) > 1 - \epsilon$$

for $n > N$.

Corollary 8.2

Let E_0 denote the event

$$\{\text{all the subsquares of } R \text{ that lie strictly below } y = \hat{c}_n + \hat{d}_n x$$

$$\text{but not strictly below } y = c + dx \text{ lie in } C_1\}$$

Then for $n > N$, $\text{prob}(E_0) > 1 - \epsilon$.

Proof:

The proof is essentially the same as the proof of Corollary 8.1.

Lemma 9

Let X_n have a chi-squared distribution with n degrees of freedom. Then there exists a finite number M that bounds the probability density function (pdf) of $\sqrt{n}(\sqrt{X_n/n} - 1)$ for all n.

Proof: It is clear that the pdf of $\sqrt{n}(\sqrt{X_n/n} - 1)$ is bounded if the pdf of $\sqrt{X_n}$ is bounded. Let $x \geq 0$ and define

$$F_n(x) \equiv \text{Prob}(\sqrt{X_n} \leq x) = \text{Prob}(X_n \leq x^2) = G_n(x^2)$$

where G_n is the cumulative distribution function of a chi-squared random variable with n degrees of freedom. Then

$$\begin{aligned} F_n'(x) = G_n'(x^2)2x &= (1/2)^{n/2}(x^2)^{n/2-1}\exp(-x^2/2)2x/\Gamma(n/2) \\ &= (1/2)^{n/2-1}x^{n-1}\exp(-x^2/2)/\Gamma(n/2) \end{aligned}$$

where Γ denotes the gamma function.

Taking the derivative with respect to x and setting it equal to zero, we see that the pdf is maximized at $x = \sqrt{n-1}$. Thus the maximum value of the pdf is

$$\begin{aligned} & (1/2)^{n/2-1}\sqrt{n-1}^{\,n-1}\exp(-(n-1)/2)/((n/2-1)\Gamma(n/2-1)) \\ = \; & ((n-2)/2)^{(n-2)/2-1/2}(\exp(-(n-2)/2)/\Gamma((n-2)/2)) \times \sqrt{2}\sqrt{(n-1)/(n-2)} \\ & \times(1 + 1/(n-2))^{(n-2)/2} \times \exp(-1/2) \end{aligned}$$

which converges to $\sqrt{2}/\sqrt{2\pi}$ as $n \to \infty$ by Stirling's formula.

Lemma 10
If

$$X_n - X'_n \xrightarrow{p} 0$$

and

$$Y_n - Y'_n \xrightarrow{p} 0$$

then given any $\epsilon > 0$ and any $\delta > 0$ there exists an N such that $n > N$ implies

$$\begin{aligned}
& \mathrm{Prob}((X_n, Y_n) \in [a_1 + \delta, a_2 - \delta] \times [b_1 + \delta, b_2 - \delta]) - \epsilon \\
< \; & \mathrm{Prob}((X'_n, Y'_n) \in [a_1, a_2] \times [b_1, b_2]) \\
< \; & \mathrm{Prob}((X_n, Y_n) \in [a_1 - \delta, a_2 + \delta] \times [b_1 - \delta, b_2 + \delta]) + \epsilon
\end{aligned}$$

for arbitrary $a_1 < a_2$, $b_1 < b_2$.

Proof: Since $X_n - X'_n \xrightarrow{p} 0$ and $Y_n - Y'_n \xrightarrow{p} 0$ we can find an N such that $n > N$ implies

$$\mathrm{Prob}(|X_n - X'_n| \geq \delta) < \epsilon/2$$

and

$$\mathrm{Prob}(|Y_n - Y'_n| \geq \delta) < \epsilon/2$$

Now it is clear that

$$\begin{aligned}
& \{(X_n, Y_n) \in [a_1 + \delta, a_2 - \delta] \times [b_1 + \delta, b_2 - \delta]\} \cap \{|X_n - X'_n| < \delta\} \cap \{|Y_n - Y'_n| < \delta\} \\
\subset \; & \{(X'_n, Y'_n) \in [a_1, a_2] \times [b_1, b_2]\} \qquad (45)
\end{aligned}$$

Also

$$p = p_1 + p_2 \qquad (46)$$

where

$$\begin{aligned}
p_1 & \equiv \mathrm{Prob}(\{(X_n, Y_n) \in [a_1 + \delta, a_2 - \delta] \times [b_1 + \delta, b_2 - \delta]\} \cap \{|X_n - X'_n| < \delta\} \cap \{|Y_n - Y'_n| < \delta\}) \\
p_2 & \equiv \mathrm{Prob}(\{(X_n, Y_n) \in [a_1 + \delta, a_2 - \delta] \times [b_1 + \delta, b_2 - \delta]\} \cap [\{|X_n - X'_n| \geq \delta\} \cup \{|Y_n - Y'_n| \geq \delta\}])
\end{aligned}$$

and

$$p \equiv \mathrm{Prob}((X_n, Y_n) \in [a_1 + \delta, a_2 - \delta] \times [b_1 + \delta, b_2 - \delta])$$

Further, for $n > N$,

$$p_2 \leq \mathrm{Prob}(|X_n - X'_n| \geq \delta) + \mathrm{Prob}(|Y_n - Y'_n| \geq \delta) < \epsilon/2 + \epsilon/2 = \epsilon \qquad (47)$$

From (46) and (47) we have, for $n > N$,

$$p_1 = p - p_2 > p - \epsilon \qquad (48)$$

From (45) and (48) we have, for $n > N$,

$$\mathrm{Prob}((X'_n, Y'_n) \in [a_1, a_2] \times [b_1, b_2]) \geq p_1 > p - \epsilon \qquad (49)$$

Similarly, for $n > N$,

$$\mathrm{Prob}((X_n, Y_n) \in [a_1 - \delta, a_2 + \delta] \times [b_1 - \delta, b_2 + \delta]) > \mathrm{Prob}((X'_n, Y'_n) \in [a_1, a_2] \times [b_1, b_2]) - \epsilon \qquad (50)$$

Results (49) and (50) establish the lemma.

Corollary
If

$$X_n - X'_n \xrightarrow{p} 0$$

then given any $\epsilon > 0$ and any $\delta > 0$ there exists an N such that $n > N$ implies

$$\begin{aligned} & \mathrm{Prob}(X_n \in [a_1 + \delta, a_2 - \delta]) - \epsilon \\ \leq \; & \mathrm{Prob}(X'_n \in [a_1, a_2]) \\ \leq \; & \mathrm{Prob}(X_n \in [a_1 - \delta, a_2 + \delta]) + \epsilon \end{aligned}$$

for arbitrary $a_1 < a_2$.

Proof: The proof is similar to the proof of Lemma 10.

Lemma 11

$$\mathrm{Prob}\left(X_{\mathrm{nct},\gamma_n(\rho),nJ-1} \leq \gamma_n(\rho) + \Phi^{-1}(\beta)\sqrt{v(\rho)}\right) \to \beta$$

as $n \to \infty$ where $X_{\mathrm{nct},\gamma_n(\rho),nJ-1}$ is a noncentral T random variable with noncentrality parameter $\gamma_n(\rho)$ and $nJ - 1$ degrees of freedom,

$$\gamma_n(\rho) \equiv \nu(\rho)\sqrt{n}$$

and

$$v(\rho) \equiv \nu(\rho)^2/(2J) + 1 \tag{51}$$

where

$$\nu(\rho) \equiv -\Phi^{-1}(\alpha)/\sqrt{1 - \rho^2 + \rho^2/J}$$

Further the convergence is uniform in ρ in the sense that given any $\epsilon > 0$, there is an N such that $n > N$ implies that

$$\left|\mathrm{Prob}\left(X_{\mathrm{nct},\gamma_n(\rho),nJ-1} \leq \gamma_n(\rho) + \Phi^{-1}(\beta)\sqrt{v(\rho)}\right) - \beta\right| < \epsilon$$

regardless of the value of ρ.

This lemma is required to establish Lemma 12 which in turn is required to establish that, for the purposes of the main theorem, we can use any consistent estimator of ρ in our calculation of $\hat{k}_n$.

Proof:

Heuristic Argument

Because the rigorous details of this proof are painful to follow, it is worthwhile first to give a heuristic justification of the lemma.

We have

$$\begin{aligned} & \mathrm{Prob}\left(X_{\mathrm{nct},\gamma_n(\rho),nJ-1} \leq \gamma_n(\rho) + \Phi^{-1}(\beta)\sqrt{v(\rho)}\right) \\ = \; & \mathrm{Prob}\left(X_{\mathrm{nct},\gamma_n(\rho),nJ-1} - \gamma_n(\rho) \leq \Phi^{-1}(\beta)\sqrt{v(\rho)}\right) \\ = \; & \mathrm{Prob}\left(U_n(\rho) + V_n \leq \Phi^{-1}(\beta)\sqrt{v(\rho)}\right) \end{aligned} \tag{52}$$

where

$$\begin{aligned}
U_n(\rho) &\equiv \nu(\rho)\sqrt{n}(1/Y_n - 1) \\
V_n &\equiv F_n/Y_n \\
Y_n &\sim \sqrt{\chi^2_{nJ-1}/(nJ-1)} \\
F_n &\sim \mathrm{N}(0,1)
\end{aligned}$$

and F_n, Y_n are statistically independent.

Now since $Y_n \xrightarrow{p} 1$, V_n is asymptotically $\mathrm{N}(0,1)$. Also

$$U_n(\rho) = \nu(\rho)\left(\sqrt{n}/\sqrt{nJ-1}\right)\sqrt{nJ-1}(1-Y_n^2)/(Y_n(1+Y_n))$$

This converges in distribution to a $\mathrm{N}(0,\nu^2(\rho)/(2J))$. Thus since V_n and $U_n(\rho)$ are asymptotically independent (in this heuristic argument we do a a big hand wave here and rely on the fact that Y_n is essentially 1 for the purposes of V_n but has a distribution for the purposes of $U_n(\rho)$), $U_n(\rho)+V_n$ is asymptotically $\mathrm{N}(0,\nu^2(\rho)/(2J)+1)$ or $\mathrm{N}(0,v(\rho))$. Consequently,

$$\mathrm{Prob}\left(U_n(\rho)+V_n \le \Phi^{-1}(\beta)\sqrt{v(\rho)}\right) \approx \mathrm{Prob}(\mathrm{N}(0,1) \le \Phi^{-1}(\beta)) = \beta$$

To make this heuristic argument rigorous and to establish that it holds uniformly in ρ requires that we make use of an argument similar to that used to prove the main theorem. In particular we break the area under a line into squares and show that the probability that $(U_n(\rho), V_n)$ lies in a square is closely approximated by the probability that $(G_n(\rho), H_n)$ lies in the square where $G_n(\rho) \sim \mathrm{N}(0,\nu^2(\rho)/(2J))$, $H_n \sim \mathrm{N}(0,1)$, and $G_n(\rho)$ and H_n are independent.

Rigorous Argument

Let $\epsilon > 0$ be given. Find an $r \times r$ square, R, on the x, y plane such that two of R's sides are parallel to the x axis and

$$\mathrm{Prob}((X,Y) \in R^c) < \epsilon \tag{53}$$

for (X,Y) having a

$$\mathrm{N}\left(\mathbf{0}, \begin{pmatrix} \nu(\rho)^2/(2J) & 0 \\ 0 & 1 \end{pmatrix}\right)$$

distribution. [Please note that we are in the process of using some fairly ugly notation. Earlier we were using $Y_{ij} = \mu_j + \sigma_Y(\rho X_{ij} + \sqrt{1-\rho^2}Z_{ij})$ in connection with the responses in a predictor sort. Here we use Y as the second variable in a bivariate normal, and below we use $Y_n \sim \sqrt{\chi^2_{nJ-1}/(nJ-1)}$. Our excuse is the limited size of the English alphabet. We hope and believe that context makes clear which Y is relevant.] It is clear that R can be chosen large enough so that (53) holds regardless of the value of ρ. Let $M \equiv \sqrt{2J}/(2\pi|\Phi^{-1}(\alpha)|)$ be the maximum value of the pdf of (X,Y) over all possible values of ρ.

Now by Lemma 7 we can divide the square R into a $k \times k$ grid with the following property: Given any line on the plane, the sum of the areas of all the subsquares of R that intersect the line is less than ϵ/M. Let $a_0 < a_1 < \ldots < a_k$ be the x coordinates of the vertices of the k^2 subsquares of R. Let $b_0 < b_1 < \ldots < b_k$ be the y coordinates of the vertices of the k^2 subsquares of R.

Next we express the central probability in the lemma in terms of $U_n(\rho)$ and V_n:

$$\begin{aligned}
& \operatorname{Prob}\left(X_{\mathrm{nct},\gamma_n(\rho),nJ-1} \le \gamma_n(\rho) + \Phi^{-1}(\beta)\sqrt{v(\rho)}\,\right) \\
= \;& \operatorname{Prob}\left(X_{\mathrm{nct},\gamma_n(\rho),nJ-1} - \gamma_n(\rho) \le \Phi^{-1}(\beta)\sqrt{v(\rho)}\,\right) \\
= \;& \operatorname{Prob}\left(V_n + U_n(\rho) \le \Phi^{-1}(\beta)\sqrt{v(\rho)}\,\right) \\
= \;& \operatorname{Prob}\left(V_n \le \Phi^{-1}(\beta)\sqrt{v(\rho)} - U_n(\rho)\right)
\end{aligned} \tag{54}$$

where

$$\begin{aligned}
U_n(\rho) &\equiv \nu(\rho)\sqrt{n}(1/Y_n - 1) \\
V_n &\equiv F_n/Y_n \\
Y_n &\sim \sqrt{\chi^2_{nJ-1}/(nJ-1)} \\
F_n &\sim \mathrm{N}(0,1)
\end{aligned}$$

and F_n, Y_n are statistically independent. (Thus F_n and $U_n(\rho)$ are statistically independent.)

It is clear that

$$F_n - V_n = F_n(1 - 1/Y_n) \xrightarrow{p} 0 \tag{55}$$

Next we demonstrate that

$$\operatorname{Prob}((U_n(\rho), V_n) \in [a_i, a_{i+1}] \times [b_j, b_{j+1}]) \approx \operatorname{Prob}(U_n(\rho) \in [a_i, a_{i+1}]) \times \operatorname{Prob}(\mathrm{N}(0,1) \in [b_j, b_{j+1}])$$

This argument concludes with result (64):

We claim that we can find a $\delta_1 > 0$ and an N_2 (N_1 is defined below) such that for $n > N_2$, all ρ, and $i \in \{0, 1, \ldots, k\}$

$$\operatorname{Prob}(U_n(\rho) \in [a_i, a_i + \delta_1]) < \epsilon/k^2 \tag{56}$$

and

$$\operatorname{Prob}(U_n(\rho) \in [a_i - \delta_1, a_i]) < \epsilon/k^2 \tag{57}$$

To prove result (56) consider

$$\begin{aligned}
& \operatorname{Prob}(U_n(\rho) \in [a_i, a_i + \delta_1]) \\
= \;& \operatorname{Prob}\left(\nu(\rho)\left(\sqrt{n}/\sqrt{nJ-1}\right)\sqrt{nJ-1}(1 - Y_n^2)/(Y_n(1+Y_n)) \in [a_i, a_i + \delta_1]\right) \\
= \;& \operatorname{Prob}(\nu(\rho)\left(\sqrt{n}/\sqrt{nJ-1}\right)\sqrt{nJ-1}(1 - Y_n^2)/(Y_n(1+Y_n)) \in [a_i, a_i + \delta_1] \\
& \cap\{Y_n \in [1 - \delta_1, 1 + \delta_1]\}) \\
& + \operatorname{Prob}(\nu(\rho)\left(\sqrt{n}/\sqrt{nJ-1}\right)\sqrt{nJ-1}(1 - Y_n^2)/(Y_n(1+Y_n)) \in [a_i, a_i + \delta_1] \\
& \cap(\{Y_n < 1 - \delta_1\} \cup \{Y_n > 1 + \delta_1\}))
\end{aligned} \tag{58}$$

The second probability in (58) is dominated by

$$\operatorname{Prob}(\{Y_n < 1 - \delta_1\} \cup \{Y_n > 1 + \delta_1\}) \tag{59}$$

Since $Y_n \xrightarrow{p} 1$ we can find a N_1 such that, regardless of the value of ρ, $n > N_1$ implies that the probability in (59) is less than $\epsilon/(2k^2)$.

Now assume that $a_i \geq 0$. (The proof is essentially the same in the other case.) Then the first probability in (58) is dominated by

$$\text{Prob}(\sqrt{nJ-1}(1-Y_n^2)$$
$$\in [a_i(1-\delta_1)(2-\delta_1)\sqrt{nJ-1}/(\sqrt{n}\nu(\rho)), (a_i+\delta_1)(1+\delta_1)(2+\delta_1)\sqrt{nJ-1}/(\sqrt{n}\nu(\rho))])$$

Thus by Pólya's Theorem (see, for example, Section 1.5.3 in Serfling (1980)) and the fact that $\nu(\rho)$ is bounded away from zero, we can find an $N_2 > N_1$ and a $\delta_2 > 0$ such that for $n > N_2$, all ρ, and $0 < \delta_1 < \delta_2$, this probability is less than $\epsilon/(2k^2)$. Thus for $n > N_2$, all ρ, and $0 < \delta_1 < \delta_2$, result (56) follows. Result (57) can be established in a similar fashion.

Next it is clear that we can find a $\delta_3 > 0$ such that for all n, for $j \in \{0, 1, \ldots, k\}$

$$\text{Prob}(F_n \in [b_j, b_j + \delta_3]) < \epsilon/k^2 \tag{60}$$

and

$$\text{Prob}(F_n \in [b_j - \delta_3, b_j]) < \epsilon/k^2 \tag{61}$$

Let $\delta \equiv \min(\delta_2, \delta_3)$.

By result (55), Lemma 10, and the statistical independence of F_n and $U_n(\rho)$, there exists an $N_3 > N_2$ such that for $n > N_3$ and arbitrary ρ, given any $i \in \{0, 1, \ldots, k-1\}$, $j \in \{0, 1, \ldots, k-1\}$,

$$\begin{aligned} & \text{Prob}(U_n(\rho) \in [a_i + \delta, a_{i+1} - \delta])\text{Prob}(F_n \in [b_j + \delta, b_{j+1} - \delta]) - \epsilon/k^2 \\ \leq\ & \text{Prob}((U_n(\rho), V_n) \in [a_i, a_{i+1}] \times [b_j, b_{j+1}]) \\ \leq\ & \text{Prob}(U_n(\rho) \in [a_i - \delta, a_{i+1} + \delta])\text{Prob}(F_n \in [b_j - \delta, b_{j+1} + \delta]) + \epsilon/k^2 \end{aligned} \tag{62}$$

Note that N_3 does not depend on ρ since, in the notation of Lemma 10, $X_n - X_n' = U_n(\rho) - U_n(\rho) = 0$, which does not depend on ρ, and $Y_n - Y_n' = F_n - V_n$, which does not depend on ρ.

Recalling that δ was chosen to satisfy (56), (57), (60), and (61), we can conclude from (62) that for $n > N_3$ (which does not depend on ρ) and $i \in \{0, 1, \ldots, k-1\}$, $j \in \{0, 1, \ldots, k-1\}$,

$$\begin{aligned} & [\text{Prob}(U_n(\rho) \in [a_i, a_{i+1}]) - \text{Prob}(U_n(\rho) \in [a_i, a_i + \delta]) - \text{Prob}(U_n(\rho) \in [a_{i+1} - \delta, a_{i+1}])] \\ & \times[\text{Prob}(F_n \in [b_j, b_{j+1}]) - \text{Prob}(F_n \in [b_j, b_j + \delta]) - \text{Prob}(F_n \in [b_{j+1} - \delta, b_{j+1}])] - \epsilon/k^2 \\ \leq\ & \text{Prob}((U_n(\rho), V_n) \in [a_i, a_{i+1}] \times [b_j, b_{j+1}]) \\ \leq\ & [\text{Prob}(U_n(\rho) \in [a_i, a_{i+1}]) + \text{Prob}(U_n(\rho) \in [a_i - \delta, a_i]) + \text{Prob}(U_n(\rho) \in [a_{i+1}, a_{i+1} + \delta])] \\ & \times[\text{Prob}(F_n \in [b_j, b_{j+1}]) + \text{Prob}(F_n \in [b_j - \delta, b_j]) + \text{Prob}(F_n \in [b_{j+1}, b_{j+1} + \delta])] + \epsilon/k^2 \end{aligned} \tag{63}$$

or (from (56), (57), (60), and (61)),

$$\begin{aligned} & \text{Prob}(U_n(\rho) \in [a_i, a_{i+1}])\text{Prob}(F_n \in [b_j, b_{j+1}]) - 9\epsilon/k^2 \\ \leq\ & \text{Prob}((U_n(\rho), V_n) \in [a_i, a_{i+1}] \times [b_j, b_{j+1}]) \\ \leq\ & \text{Prob}(U_n(\rho) \in [a_i, a_{i+1}])\text{Prob}(F_n \in [b_j, b_{j+1}]) + 9\epsilon/k^2 \end{aligned} \tag{64}$$

Next we show that

$$\text{Prob}(U_n(\rho) \in [a_i, a_{i+1}]) \approx \text{Prob}\left(\text{N}\left(0, \frac{\nu^2(\rho)}{2J}\right) \in [a_i, a_{i+1}]\right)$$

This argument concludes with result (73):

We have

$$\begin{aligned} &\text{Prob}(U_n(\rho) \in [a_i, a_{i+1}]) \\ &= \text{Prob}(\sqrt{nJ-1}(1-Y_n^2)/\sqrt{2} \in [x_L(\rho,n), x_U(\rho,n)]) \end{aligned} \tag{65}$$

where

$$x_L(\rho,n) \equiv a_i\sqrt{nJ-1}Y_n(1+Y_n)/(\nu(\rho)\sqrt{2n})$$

and

$$x_U(\rho,n) \equiv a_{i+1}\sqrt{nJ-1}Y_n(1+Y_n)/(\nu(\rho)\sqrt{2n})$$

Define

$$y_L(\rho) \equiv a_i\sqrt{2J}/\nu(\rho)$$

and

$$y_U(\rho) \equiv a_{i+1}\sqrt{2J}/\nu(\rho)$$

It is clear that we can find a $\delta_4 > 0$ such that regardless of the value of ρ,

$$\text{Prob}(\text{N}(0,1) \in [y_L(\rho)-\delta_4, y_U(\rho)+\delta_4]) - \text{Prob}(\text{N}(0,1) \in [y_L(\rho), y_U(\rho)]) < \epsilon/(3k^2) \tag{66}$$

and

$$\text{Prob}(\text{N}(0,1) \in [y_L(\rho), y_U(\rho)]) - \text{Prob}(\text{N}(0,1) \in [y_L(\rho)+\delta_4, y_U(\rho)-\delta_4]) < \epsilon/(3k^2) \tag{67}$$

Since $\nu(\rho)$ is bounded away from zero, and $Y_n \xrightarrow{p} 1$, it is clear that we can find an $N_4 > N_3$ such that for all ρ, $n > N_4$ implies that

$$\text{Prob}(\{|y_L(\rho) - x_L(\rho,n)| \geq \delta_4\} \cup \{|y_U(\rho) - x_U(\rho,n)| \geq \delta_4\}) < \epsilon/(3k^2) \tag{68}$$

Thus, for $n > N_4$, any ρ, we have

$$\begin{aligned} &\text{Prob}\left(\sqrt{nJ-1}(1-Y_n^2)/\sqrt{2} \in [y_L(\rho)+\delta_4, y_U(\rho)-\delta_4]\right) - \epsilon/(3k^2) \\ \leq\ & \text{Prob}(\sqrt{nJ-1}(1-Y_n^2)/\sqrt{2} \in [y_L(\rho)+\delta_4, y_U(\rho)-\delta_4] \\ & \cap\{|y_L(\rho) - x_L(\rho,n)| < \delta_4\} \cap \{|y_U(\rho) - x_U(\rho,n)| < \delta_4\}) \\ \leq\ & \text{Prob}\left(\sqrt{nJ-1}(1-Y_n^2)/\sqrt{2} \in [x_L(\rho,n), x_U(\rho,n)]\right) \\ \leq\ & \text{Prob}\left(\sqrt{nJ-1}(1-Y_n^2)/\sqrt{2} \in [y_L(\rho)-\delta_4, y_U(\rho)+\delta_4]\right) \\ & + \text{Prob}\left(\{|y_L(\rho) - x_L(\rho,n)| \geq \delta_4\} \cup \{|y_U(\rho) - x_U(\rho,n)| \geq \delta_4\}\right) \\ \leq\ & \text{Prob}\left(\sqrt{nJ-1}(1-Y_n^2)/\sqrt{2} \in [y_L(\rho)-\delta_4, y_U(\rho)+\delta_4]\right) + \epsilon/(3k^2) \end{aligned} \tag{69}$$

Now by Pólya's Theorem we can find an $N_5 > N_4$ such that $n > N_5$ implies that, for all ρ,

$$\begin{aligned} & \text{Prob}(\text{N}(0,1) \in [y_L(\rho)+\delta_4, y_U(\rho)-\delta_4]) - \epsilon/(3k^2) \\ \leq\ & \text{Prob}(\sqrt{nJ-1}(1-Y_n^2)/\sqrt{2} \in [y_L(\rho)+\delta_4, y_U(\rho)-\delta_4]) \\ \leq\ & \text{Prob}(\text{N}(0,1) \in [y_L(\rho)+\delta_4, y_U(\rho)-\delta_4]) + \epsilon/(3k^2) \end{aligned} \tag{70}$$

and

$$\begin{aligned} & \text{Prob}(\text{N}(0,1) \in [y_L(\rho)-\delta_4, y_U(\rho)+\delta_4]) - \epsilon/(3k^2) \\ \leq\ & \text{Prob}(\sqrt{nJ-1}(1-Y_n^2)/\sqrt{2} \in [y_L(\rho)-\delta_4, y_U(\rho)+\delta_4]) \\ \leq\ & \text{Prob}(\text{N}(0,1) \in [y_L(\rho)-\delta_4, y_U(\rho)+\delta_4]) + \epsilon/(3k^2) \end{aligned} \tag{71}$$

Results (65) and (69) – (71) imply that for $n > N_5$ and all ρ,

$$\begin{aligned}
& \text{Prob}(\text{N}(0,1) \in [y_L(\rho)+\delta_4, y_U(\rho)-\delta_4]) - 2\epsilon/(3k^2) && (72)\\
\leq\ & \text{Prob}(\sqrt{nJ-1}(1-Y_n^2)/\sqrt{2} \in [y_L(\rho)+\delta_4, y_U(\rho)-\delta_4]) - \epsilon/(3k^2) \\
\leq\ & \text{Prob}(U_n(\rho) \in [a_i, a_{i+1}]) \\
\leq\ & \text{Prob}(\sqrt{nJ-1}(1-Y_n^2)/\sqrt{2} \in [y_L(\rho)-\delta_4, y_U(\rho)+\delta_4]) + \epsilon/(3k^2) \\
\leq\ & \text{Prob}(\text{N}(0,1) \in [y_L(\rho)-\delta_4, y_U(\rho)+\delta_4]) + 2\epsilon/(3k^2)
\end{aligned}$$

Results (66), (67), and (72) then imply that for $n > N_5$ and all ρ,

$$\begin{aligned}
& \text{Prob}(\text{N}(0,1) \in [y_L(\rho), y_U(\rho)]) - \epsilon/k^2 \\
\leq\ & \text{Prob}(U_n(\rho) \in [a_i, a_{i+1}]) \\
\leq\ & \text{Prob}(\text{N}(0,1) \in [y_L(\rho), y_U(\rho)]) + \epsilon/k^2
\end{aligned}$$

or

$$\begin{aligned}
& \text{Prob}(E(\rho) \in [a_i, a_{i+1}]) - \epsilon/k^2 && (73)\\
\leq\ & \text{Prob}(U_n(\rho) \in [a_i, a_{i+1}]) \\
\leq\ & \text{Prob}(E(\rho) \in [a_i, a_{i+1}]) + \epsilon/k^2
\end{aligned}$$

where

$$E(\rho) \sim \text{N}(0, \nu(\rho)^2/(2J))$$

We now pull together results (64) and (73) to yield

$$\text{Prob}((U_n(\rho), V_n) \in \text{a particular square}) \approx \text{Prob}\left(\text{independent N}(0, \frac{\nu^2(\rho)}{2J}), \text{N}(0,1) \in \text{the square}\right)$$

This result is embodied in inequalities (74) and (75):

Results (64) and (73) imply that for $n > N_5$, all ρ, $i \in \{0,1,\ldots,k-1\}$, $j \in \{0,1,\ldots,k-1\}$

$$\begin{aligned}
& \text{Prob}(E(\rho) \in [a_i, a_{i+1}])\text{Prob}(F_n \in [b_j, b_{j+1}]) - 10\epsilon/k^2 && (74)\\
\leq\ & \text{Prob}((U_n(\rho), V_n) \in [a_i, a_{i+1}] \times [b_j, b_{j+1}]) \\
\leq\ & \text{Prob}(E(\rho) \in [a_i, a_{i+1}])\text{Prob}(F_n \in [b_j, b_{j+1}]) + 10\epsilon/k^2
\end{aligned}$$

Result (74) implies that for $n > N_5$, all ρ,

$$\begin{aligned}
& \text{Prob}\,(E(\rho) \in [a_0, a_k])\,\text{Prob}\,(F_n \in [b_0, b_k]) - 10\epsilon && (75)\\
\leq\ & \text{Prob}((U_n(\rho), V_n) \in R) \\
\leq\ & \text{Prob}\,(E(\rho) \in [a_0, a_k])\,\text{Prob}\,(F_n \in [b_0, b_k]) + 10\epsilon
\end{aligned}$$

Since $E(\rho) \sim \text{N}(0, \nu(\rho)^2/(2J))$ and $F_n \sim \text{N}(0,1)$, (53) and (75) imply that for $n > N_5$, all ρ,

$$\text{Prob}((U_n(\rho), V_n) \in R) > 1 - 11\epsilon \tag{76}$$

or

$$\text{Prob}((U_n(\rho), V_n) \in R^c) < 11\epsilon \tag{77}$$

Now we must be careful and show that the squares that intersect a particular line can be treated as negligible. This fact is expressed in result (78):

Define

$$C_{n,1} \equiv \text{ subsquares of } R \text{ that intersect the line } y = \Phi^{-1}(\beta)\sqrt{v(\rho)} - x$$

(see (54) for the source of this line),

$$\begin{aligned} R_{n,1} &\equiv \text{ the region covered by the elements of } C_{n,1} \\ C_{n,2} &\equiv \text{ subsquares of } R \text{ that lie entirely below the line } y = \Phi^{-1}(\beta)\sqrt{v(\rho)} - x \end{aligned}$$

and

$$R_{n,2} \equiv \text{the region covered by the elements of } C_{n,2}$$

Then by (74) and by the method of construction of the $k \times k$ grid, for $n > N_5$, all ρ,

$$\begin{aligned} & \text{Prob}((U_n(\rho), V_n) \in R_{n,1}) \\ \leq \;& \text{Prob}(\text{independent N}(0, \nu(\rho)^2/(2J)), \text{N}(0,1) \in R_{n,1}) + 10\epsilon \\ < \;& (\epsilon/M)M + 10\epsilon = 11\epsilon \end{aligned} \tag{78}$$

where M was introduced in the paragraph following (53).

Result (79) handles the squares that lie below the line:

Again by (74) for $n > N_5$, all ρ,

$$\begin{aligned} & \text{Prob}(\text{independent N}(0, \nu(\rho)^2/(2J)), \text{N}(0,1) \in R_{n,2}) - 10\epsilon \\ \leq \;& \text{Prob}((U_n(\rho), V_n) \in R_{n,2}) \\ \leq \;& \text{Prob}(\text{independent N}(0, \nu(\rho)^2/(2J)), \text{N}(0,1) \in R_{n,2}) + 10\epsilon \end{aligned} \tag{79}$$

We are now ready to complete the proof by introducing statistically independent $\text{N}(0, \nu^2(\rho)/(2J))$ **and** $\text{N}(0,1)$ **random variables:**

Let $G_n(\rho) \sim \text{N}(0, \nu(\rho)^2/(2J))$, $H_n \sim \text{N}(0,1)$, and let $G_n(\rho)$, H_n be independent.
By (53), for all ρ (see the remark following (53)),

$$\text{Prob}((G_n(\rho), H_n) \in R^c) < \epsilon \tag{80}$$

By the method of construction of the $k \times k$ grid, for all ρ,

$$\text{Prob}((G_n(\rho), H_n) \in R_{n,1}) < (\epsilon/M)M = \epsilon \tag{81}$$

Next note that

$$\begin{aligned} & \text{Prob}\left(H_n \leq \Phi^{-1}(\beta)\sqrt{v(\rho)} - G_n(\rho)\right) = \text{Prob}\left(G_n(\rho) + H_n \leq \Phi^{-1}(\beta)\sqrt{v(\rho)}\right) \\ = \;& \text{Prob}\left(\text{N}(0, v(\rho)) \leq \Phi^{-1}(\beta)\sqrt{v(\rho)}\right) = \beta \end{aligned} \tag{82}$$

Thus by (54) and (82),

$$\begin{aligned} & |\text{Prob}\left(X_{\text{nct},\gamma_n(\rho),nJ-1} \leq \gamma_n(\rho) + \Phi^{-1}(\beta)\sqrt{v(\rho)}\right) - \beta| \\ = \;& |\text{Prob}\left(V_n \leq \Phi^{-1}(\beta)\sqrt{v(\rho)} - U_n(\rho)\right) - \text{Prob}\left(H_n \leq \Phi^{-1}(\beta)\sqrt{v(\rho)} - G_n(\rho)\right)| \\ = \;& |\text{Prob}\left(V_n \leq \Phi^{-1}(\beta)\sqrt{v(\rho)} - U_n(\rho) \text{ and } (U_n(\rho), V_n) \in R^c\right) \end{aligned}$$

$$
\begin{aligned}
& + \operatorname{Prob}\left(V_n \le \Phi^{-1}(\beta)\sqrt{v(\rho)} - U_n(\rho) \text{ and } (U_n(\rho), V_n) \in R_{n,1}\right) \\
& + \operatorname{Prob}\left(V_n \le \Phi^{-1}(\beta)\sqrt{v(\rho)} - U_n(\rho) \text{ and } (U_n(\rho), V_n) \in R_{n,2}\right) \\
& - \operatorname{Prob}\left(H_n \le \Phi^{-1}(\beta)\sqrt{v(\rho)} - G_n(\rho) \text{ and } (G_n(\rho), H_n) \in R^c\right) \\
& - \operatorname{Prob}\left(H_n \le \Phi^{-1}(\beta)\sqrt{v(\rho)} - G_n(\rho) \text{ and } (G_n(\rho), H_n) \in R_{n,1}\right) \\
& - \operatorname{Prob}\left(H_n \le \Phi^{-1}(\beta)\sqrt{v(\rho)} - G_n(\rho) \text{ and } (G_n(\rho), H_n) \in R_{n,2}\right) | \\
\le \;& \operatorname{Prob}((U_n(\rho), V_n) \in R^c) + \operatorname{Prob}((G_n(\rho), H_n) \in R^c) \\
& + \operatorname{Prob}((U_n(\rho), V_n) \in R_{n,1}) + \operatorname{Prob}((G_n(\rho), H_n) \in R_{n,1}) \\
& + |\operatorname{Prob}((U_n(\rho), V_n) \in R_{n,2}) - \operatorname{Prob}((G_n(\rho), H_n) \in R_{n,2})|
\end{aligned}
$$

By results (77), (80), (78), (81), and (79), for $n > N_5$, all ρ, this last sum is less than

$$11\epsilon + \epsilon + 11\epsilon + \epsilon + 10\epsilon$$

As ϵ was arbitary, this completes the proof.

Corollary (An Aside)

$$\operatorname{Prob}\left(X_{\mathrm{nct},\gamma\sqrt{n},n} \le \gamma\sqrt{n} + \Phi^{-1}(\beta)\sqrt{\gamma^2/2+1}\ \right) \to \beta$$

as $n \to \infty$.

Proof: The proof is a simplified version of the proof of Lemma 10.

Note that this corollary implies that given any $\epsilon > 0$ we can find an $N > 0$ such that $n > N$ implies that

$$\gamma\sqrt{n} + \Phi^{-1}(\beta - \epsilon)\sqrt{\gamma^2/2+1} < \mathrm{F}^{-1}_{\mathrm{nct},\gamma\sqrt{n},n}(\beta) < \gamma\sqrt{n} + \Phi^{-1}(\beta + \epsilon)\sqrt{\gamma^2/2+1}$$

By contrast, the βth quantile for a $\mathrm{N}(\gamma\sqrt{n}, 1)$ random variable is $\gamma\sqrt{n} + \Phi^{-1}(\beta)$. Thus the asymptotic effect of dividing a $\mathrm{N}(\gamma\sqrt{n}, 1)$ random variable by an independent $\sqrt{\chi^2_n/n}$ is to expand the variance from 1 to $\gamma^2/2 + 1$.

Lemma 12

(Recall that this lemma is needed to establish that the main theorem holds for any consistent estimator $\hat{\rho}$ of ρ.)

Let $\hat{\rho}$ be a consistent estimator of ρ. Then

$$\operatorname{Prob}\left(X_{\mathrm{nct},\gamma_n(\rho),nJ-1} \le \left(\sqrt{1-\hat{\rho}^2+\hat{\rho}^2/J}/\sqrt{1-\rho^2+\rho^2/J}\right)\mathrm{F}^{-1}_{\mathrm{nct},\gamma_n(\hat{\rho}),nJ-1}(\beta)\right) \to \beta$$

as $n \to \infty$ where

$$
\begin{aligned}
\gamma_n(\rho) &\equiv \nu(\rho)\sqrt{n} \\
\nu(\rho) &\equiv -\Phi^{-1}(\alpha)/\sqrt{1-\rho^2+\rho^2/J}
\end{aligned}
$$

$X_{\mathrm{nct},\gamma_n(\rho),nJ-1}$ denotes a noncentral T random variable with noncentrality parameter $\gamma_n(\rho)$ and $nJ-1$ degrees of freedom, and $\mathrm{F}_{\mathrm{nct},\gamma_n(\hat{\rho}),nJ-1}$ denotes the distribution function of a noncentral T random variable with noncentrality parameter $\gamma_n(\hat{\rho})$ and $nJ-1$ degrees of freedom.

Proof: Let $\epsilon > 0$ be given. By Lemma 11 we can find an N_1 such that for $n > N_1$, all ρ,

$$\text{Prob}\left(X_{\text{nct},\gamma_n(\rho),nJ-1} \leq \gamma_n(\rho) + \Phi^{-1}(\beta - \epsilon)\sqrt{v(\rho)}\right) < \beta - \epsilon/2$$

and

$$\text{Prob}\left(X_{\text{nct},\gamma_n(\rho),nJ-1} \leq \gamma_n(\rho) + \Phi^{-1}(\beta + \epsilon)\sqrt{v(\rho)}\right) > \beta + \epsilon/2$$

where

$$v(\rho) \equiv \nu(\rho)^2/(2J) + 1$$

Thus, for $n > N_1$, all ρ,

$$\gamma_n(\rho) + \Phi^{-1}(\beta - \epsilon)\sqrt{v(\rho)} < \text{F}^{-1}_{\text{nct},\gamma_n(\rho),nJ-1}(\beta) < \gamma_n(\rho) + \Phi^{-1}(\beta + \epsilon)\sqrt{v(\rho)} \tag{83}$$

So for $n > N_1$,

$$\begin{aligned}
& \text{Prob}\left(X_{\text{nct},\gamma_n(\rho),nJ-1} \leq \left(\sqrt{1-\hat\rho^2+\hat\rho^2/J}/\sqrt{1-\rho^2+\rho^2/J}\right)\left(\gamma_n(\hat\rho) + \Phi^{-1}(\beta - \epsilon)\sqrt{v(\hat\rho)}\right)\right) \\
\leq\ & \text{Prob}\left(X_{\text{nct},\gamma_n(\rho),nJ-1} \leq \left(\sqrt{1-\hat\rho^2+\hat\rho^2/J}/\sqrt{1-\rho^2+\rho^2/J}\right)\text{F}^{-1}_{\text{nct},\gamma_n(\hat\rho),nJ-1}(\beta)\right) \qquad (84) \\
\leq\ & \text{Prob}\left(X_{\text{nct},\gamma_n(\rho),nJ-1} \leq \left(\sqrt{1-\hat\rho^2+\hat\rho^2/J}/\sqrt{1-\rho^2+\rho^2/J}\right)\left(\gamma_n(\hat\rho) + \Phi^{-1}(\beta + \epsilon)\sqrt{v(\hat\rho)}\right)\right)
\end{aligned}$$

Now, since $\hat\rho \xrightarrow{p} \rho$, we can find a $N_2 > N_1$ such that $n > N_2$ implies

$$\text{Prob}\left(\Phi^{-1}(\beta - 2\epsilon)\sqrt{v(\rho)} \leq \Phi^{-1}(\beta - \epsilon)\sqrt{v(\hat\rho)}\sqrt{1-\hat\rho^2+\hat\rho^2/J}/\sqrt{1-\rho^2+\rho^2/J}\right) > 1 - \epsilon$$

and

$$\text{Prob}\left(\Phi^{-1}(\beta + 2\epsilon)\sqrt{v(\rho)} \geq \Phi^{-1}(\beta + \epsilon)\sqrt{v(\hat\rho)}\sqrt{1-\hat\rho^2+\hat\rho^2/J}/\sqrt{1-\rho^2+\rho^2/J}\right) > 1 - \epsilon$$

Thus for $n > N_2$,

$$\begin{aligned}
& \text{Prob}\left(X_{\text{nct},\gamma_n(\rho),nJ-1} \leq \gamma_n(\rho) + \Phi^{-1}(\beta - 2\epsilon)\sqrt{v(\rho)}\right) - \epsilon \qquad (85) \\
\leq\ & \text{Prob}\left(X_{\text{nct},\gamma_n(\rho),nJ-1} \leq \gamma_n(\rho) + \Phi^{-1}(\beta - \epsilon)\sqrt{v(\hat\rho)}\left(\sqrt{1-\hat\rho^2+\hat\rho^2/J}/\sqrt{1-\rho^2+\rho^2/J}\right)\right)
\end{aligned}$$

and

$$\begin{aligned}
& \text{Prob}\left(X_{\text{nct},\gamma_n(\rho),nJ-1} \leq \gamma_n(\rho) + \Phi^{-1}(\beta + \epsilon)\sqrt{v(\hat\rho)}\left(\sqrt{1-\hat\rho^2+\hat\rho^2/J}/\sqrt{1-\rho^2+\rho^2/J}\right)\right) \\
\leq\ & \text{Prob}\left(X_{\text{nct},\gamma_n(\rho),nJ-1} \leq \gamma_n(\rho) + \Phi^{-1}(\beta + 2\epsilon)\sqrt{v(\rho)}\right) + \epsilon \qquad (86)
\end{aligned}$$

Results (84) through (86) and the fact that $\gamma_n(\rho) = \gamma_n(\hat\rho)\sqrt{1-\hat\rho^2+\hat\rho^2/J}/\sqrt{1-\rho^2+\rho^2/J}$ imply that for $n > N_2$,

$$\begin{aligned}
& \text{Prob}\left(X_{\text{nct},\gamma_n(\rho),nJ-1} \leq \gamma_n(\rho) + \Phi^{-1}(\beta - 2\epsilon)\sqrt{v(\rho)}\right) - \epsilon \qquad (87) \\
\leq\ & \text{Prob}\left(X_{\text{nct},\gamma_n(\rho),nJ-1} \leq \left(\sqrt{1-\hat\rho^2+\hat\rho^2/J}/\sqrt{1-\rho^2+\rho^2/J}\right)\text{F}^{-1}_{\text{nct},\gamma_n(\hat\rho),nJ-1}(\beta)\right) \\
\leq\ & \text{Prob}\left(X_{\text{nct},\gamma_n(\rho),nJ-1} \leq \gamma_n(\rho) + \Phi^{-1}(\beta + 2\epsilon)\sqrt{v(\rho)}\right) + \epsilon
\end{aligned}$$

Finally, by Lemma 11, we can find a $N_3 > N_2$ such that $n > N_3$ implies

$$\text{Prob}\left(X_{\text{nct},\gamma_n(\rho),nJ-1} \leq \gamma_n(\rho) + \Phi^{-1}(\beta - 2\epsilon)\sqrt{v(\rho)}\right) > \beta - 3\epsilon \tag{88}$$

and

$$\text{Prob}\left(X_{\text{nct},\gamma_n(\rho),nJ-1} \leq \gamma_n(\rho) + \Phi^{-1}(\beta + 2\epsilon)\sqrt{v(\rho)}\right) < \beta + 3\epsilon \tag{89}$$

Results (87) to (89) imply that for $n > N_3$,

$$\beta - 4\epsilon \leq \text{Prob}\left(X_{\text{nct},\gamma_n(\rho),nJ-1} \leq \left(\sqrt{1-\hat{\rho}^2+\hat{\rho}^2/J}/\sqrt{1-\rho^2+\rho^2/J}\right)\text{F}^{-1}_{\text{nct},\gamma_n(\hat{\rho}),nJ-1}(\beta)\right) \leq \beta + 4\epsilon$$

As ϵ was arbitrary, the lemma is established.

Corollary

Let $\hat{\rho}$ be a consistent estimator of ρ. Define

$$\hat{k}_n \equiv \left(\sqrt{1-\hat{\rho}^2+\hat{\rho}^2/J}/\sqrt{n}\right)\text{F}^{-1}_{\text{nct},\gamma_n(\hat{\rho}),nJ-1}(\beta) \tag{90}$$

where

$$\gamma_n(\hat{\rho}) \equiv \nu(\hat{\rho})\sqrt{n} \equiv \left(-\Phi^{-1}(\alpha)/\sqrt{1-\hat{\rho}^2+\hat{\rho}^2/J}\right)\sqrt{n} \tag{91}$$

Define

$$\hat{c}_n \equiv (\hat{k}_n + \Phi^{-1}(\alpha))\sqrt{n}/\sqrt{1-\rho^2+\rho^2/J} \tag{92}$$

and

$$\hat{d}_n \equiv \hat{k}_n\sqrt{n}/\left(\sqrt{1-\rho^2+\rho^2/J}\sqrt{nJ-1}\right) \tag{93}$$

Then

$$\hat{k}_n \to -\Phi^{-1}(\alpha) \tag{94}$$

as $n \to \infty$,

$$\hat{c}_n \xrightarrow{p} c \equiv \Phi^{-1}(\beta)\sqrt{v(\rho)} \tag{95}$$

and

$$\hat{d}_n \xrightarrow{p} d \equiv -\Phi^{-1}(\alpha)/\left(\sqrt{1-\rho^2+\rho^2/J}\sqrt{J}\right) = \nu(\rho)/\sqrt{J} \tag{96}$$

as $n \to \infty$ where

$$v(\rho) \equiv \nu^2(\rho)/(2J) + 1 \tag{97}$$

and $\nu(\rho)$ is defined in (91).

Proof:

Let $\epsilon > 0$ be given. Then by (83), we can find an N such that $n > N$ implies

$$\gamma_n(\hat{\rho}) + \Phi^{-1}(\beta - \epsilon)\sqrt{v(\hat{\rho})} < \text{F}^{-1}_{\text{nct},\gamma_n(\hat{\rho}),nJ-1}(\beta) < \gamma_n(\hat{\rho}) + \Phi^{-1}(\beta + \epsilon)\sqrt{v(\hat{\rho})}$$

Thus, for $n > N$,

$$\begin{aligned} & \left(\sqrt{1-\hat{\rho}^2+\hat{\rho}^2/J}/\sqrt{n}\right)\left(\gamma_n(\hat{\rho}) + \Phi^{-1}(\beta-\epsilon)\sqrt{v(\hat{\rho})}\right) \\ < \ & \hat{k}_n \\ < \ & \left(\sqrt{1-\hat{\rho}^2+\hat{\rho}^2/J}/\sqrt{n}\right)\left(\gamma_n(\hat{\rho}) + \Phi^{-1}(\beta+\epsilon)\sqrt{v(\hat{\rho})}\right) \end{aligned}$$

or

$$\begin{aligned} & -\Phi^{-1}(\alpha)+\Phi^{-1}(\beta-\epsilon)\sqrt{1-\hat\rho^2+\hat\rho^2/J}\sqrt{v(\hat\rho)}/\sqrt{n} \\ < \; & \hat{k}_n \\ < \; & -\Phi^{-1}(\alpha)+\Phi^{-1}(\beta+\epsilon)\sqrt{1-\hat\rho^2+\hat\rho^2/J}\sqrt{v(\hat\rho)}/\sqrt{n} \end{aligned} \tag{98}$$

This establishes (94) which in turn establishes (96).

From (98) and definition (92) we have, for $n > N$,

$$\begin{aligned} & \left(\Phi^{-1}(\beta-\epsilon)\sqrt{1-\hat\rho^2+\hat\rho^2/J}\sqrt{v(\hat\rho)}/\sqrt{n}\right)\sqrt{n}/\sqrt{1-\rho^2+\rho^2/J} \\ < \; & \hat{c}_n \\ < \; & \left(\Phi^{-1}(\beta+\epsilon)\sqrt{1-\hat\rho^2+\hat\rho^2/J}\sqrt{v(\hat\rho)}/\sqrt{n}\right)\sqrt{n}/\sqrt{1-\rho^2+\rho^2/J} \end{aligned}$$

or

$$\begin{aligned} & \Phi^{-1}(\beta-\epsilon)\sqrt{v(\hat\rho)}\left(\sqrt{1-\hat\rho^2+\hat\rho^2/J}/\sqrt{1-\rho^2+\rho^2/J}\right) \\ < \; & \hat{c}_n \\ < \; & \Phi^{-1}(\beta+\epsilon)\sqrt{v(\hat\rho)}\left(\sqrt{1-\hat\rho^2+\hat\rho^2/J}/\sqrt{1-\rho^2+\rho^2/J}\right) \end{aligned}$$

Since $\hat\rho \xrightarrow{p} \rho$ and ϵ was arbitrary, this establishes (95) and completes the proof of the corollary.

The Independence of $V_{n,2}$ and $U_{n,2}$

To be rigorous we need to be fairly careful as we establish the independence of $V_{n,2}$ and $U_{n,2}$. First, recall the following lemma.

Lemma 13

Let $Z_1, \ldots, Z_n$ be independent N(0,1) random variables. Then there exist independent N(0,1) random variables, $T_1, \ldots, T_n$, such that $\bar{Z}$ is a continuous function of T_1 and $Z_1 - \bar{Z}, \ldots, Z_n - \bar{Z}$ are continuous functions of $T_2, \ldots, T_n$.

Proof:

Let $\mathbf{u}_1 = (1\ldots1)^T/\sqrt{n}$ and let $\mathbf{u}_2, \ldots, \mathbf{u}_n$ be an orthonormal extension of $\mathbf{u}_1$ to a basis of Euclidean n-space. Then

$$(\mathbf{u}_1 \ldots \mathbf{u}_n)^T\mathbf{Z} \sim \mathrm{N}(\mathbf{0}, \mathbf{I}_{nxn})$$

so $\mathbf{u}_1^T\mathbf{Z}, \ldots, \mathbf{u}_n^T\mathbf{Z}$ are stochastically independent. Now

$$\bar{Z} = \mathbf{u}_1^T\mathbf{Z}/\sqrt{n}$$

is a continuous function of $\mathbf{u}_1^T\mathbf{Z}$. Also

$$Z_1 - \bar{Z} = \mathbf{v}_1^T\mathbf{Z}$$

where $\mathbf{v}_1^T = (1\,0\ldots0) - (1\ldots1)/n$. Since $\mathbf{v}_1^T\mathbf{1} = 0$, $\mathbf{v}_1$ lies in the linear span of $\mathbf{u}_2, \ldots, \mathbf{u}_n$. Thus $Z_1 - \bar{Z} = \mathbf{v}_1^T\mathbf{Z}$ is a linear combination of $\mathbf{u}_2^T\mathbf{Z}, \ldots, \mathbf{u}_n^T\mathbf{Z}$. A similar argument holds for $Z_2 - \bar{Z}, \ldots, Z_n - \bar{Z}$. This completes the proof of the lemma.

Now to complete the justification of the stochastic independence of $V_{n,2}$ and $U_{n,2}$ we would simply like to invoke Lemma 13 and Theorem 3.3.2 of Chung (1974). However to do so we need to build the predictor sort randomization into our probability model. We can do this by adding an extra dimension to the $2 \times nJ$ (n here differs from the n in Lemma 13) dimensional product space upon which we are (implicitly) generating the nJ X_{ij}'s and nJ Z_{ij}'s of the main theorem.

Lemma 14

$V_{n,2}$ and $U_{n,2}$ are stochastically independent.

Proof:

Let R_{ij}'s and Z_{ij}'s, $i = 1, \ldots, n$, $j = 1, \ldots, J$ be independent N(0,1)'s on a $2 \times nJ$ product space. Extend the product space by "multiplying" by the interval [0,1], where the associated measure is the standard uniform measure on [0,1]. Let $\mathbf{w}$ be an element of this product space (so its $2 \times nJ + 1$th component lies in [0,1]). Let W_k be the random variable defined by $W_k(\mathbf{w}) = w_k$. We have $R_{ij}(\mathbf{w}) = w_{n(j-1)+i} = W_{n(j-1)+i}(\mathbf{w})$ and $Z_{ij}(\mathbf{w}) = w_{nJ+n(j-1)+i} = W_{nJ+n(j-1)+i}(\mathbf{w})$, and the probability density at $\mathbf{w}$ is

$$\left(\prod_{k=1}^{2nJ} \frac{1}{\sqrt{2\pi}} \exp(-w_k^2/2) \right) \times 1$$

Now the R's and Z's are all statistically independent. By Lemma 13 we can find statistically independent $S_1, \ldots, S_{nJ}$ on this product space such that the S's are functions of $W_1, \ldots, W_{nJ}$, $\bar{R}_{\cdot\cdot}(\mathbf{w}) = \sum_{k=1}^{nJ} W_k(\mathbf{w})$ is a function of S_1, and the $R_{ij}(\mathbf{w}) - \bar{R}_{\cdot\cdot}(\mathbf{w})$'s are functions of $S_2, \ldots, S_{nJ}$. Also by Lemma 13 we can find statistically independent $T_{1j}, \ldots, T_{nj}$ for $j = 1, \ldots, J$ such that the T_{ij}'s are functions of $W_{nJ+n(j-1)+1}, \ldots, W_{nJ+nj}$, $\bar{Z}_{\cdot j}(\mathbf{w})$ is a function of T_{1j}, and the $Z_{ij}(\mathbf{w}) - \bar{Z}_{\cdot j}(\mathbf{w})$'s are functions of $T_{2j}, \ldots, T_{nj}$. Clearly the S's and the T's are independent of each other and of W_{2nJ+1}.

Consider indicator variables $I_{(d_{i-1}, d_i)}(W_{2nJ+1}(\mathbf{w}))$ where $i = 1, \ldots, J!^n$ and $d_i = i/(J!^n)$. $I_{(d_{i-1}, d_i)}(W_{2nJ+1}(\mathbf{w})) = 1$ if $W_{2nJ+1}(\mathbf{w})$ lies in (d_{i-1}, d_i), 0 otherwise.

Then, in the notation of (3),

$$S_{n,2}^2(\mathbf{w})/\sigma_Y^2 = \sum_{k=1}^{J!^n} I_{(d_{k-1}, d_k)}(W_{2nJ+1}(\mathbf{w})) \times$$

$$\sum_{j=1}^{J} \sum_{i=1}^{n} \left(\rho \left(X_{kij}(\mathbf{w}) - \bar{X}_{\cdot\cdot}(\mathbf{w}) \right) + \sqrt{1-\rho^2} \left(Z_{ij}(\mathbf{w}) - \bar{Z}_{\cdot j}(\mathbf{w}) \right) \right)^2$$

where $\mathbf{w}$ is a point in the product space, and the $X_{kij}(\mathbf{w}) - \bar{X}_{\cdot\cdot}(\mathbf{w})$'s are the randomization of the $R_{ij}(\mathbf{w}) - \bar{R}_{\cdot\cdot}(\mathbf{w})$'s that corresponds to the kth of the possible $J!^n$ predictor sort randomizations. It is clear that $S_{n,2}^2(\mathbf{w})/\sigma_Y^2$ is a Borel measurable function of $\{S_2, \ldots, S_{nJ}\}$, $\{T_{2j}, \ldots, T_{nj}\}$ for $j = 1, \ldots, J$, and W_{2nJ+1}. Also it is clear that $\bar{Y}_{n,2}$ is a Borel measurable function of S_1 and T_{11}.

Thus in the product probability space, $\bar{Y}_{n,2}$ and $S_{n,2}^2$ are Borel measurable functions of disjoint sets of stochastically independent random variables. So by Chung's Theorem 3.3.2, $\bar{Y}_{n,2}$ and $S_{n,2}^2$ are stochastically independent.

Proof of the Main Result

Theorem.

Let $\hat{\rho}$ be a consistent estimator of ρ. Then

$$\text{Prob}\left(\bar{Y}_{n,1} - \hat{k}_n S_{n,1} \le \mu_1 + \Phi^{-1}(\alpha)\sigma_Y \right) \to \beta$$

as $n \to \infty$ where $\bar{Y}_{n,1}$ and $S_{n,1}$ are defined in (3),

$$\hat{k}_n \equiv \sqrt{(1-\hat{\rho}^2+\hat{\rho}^2/J)/n}\ \mathrm{F}^{-1}_{\mathrm{nct},\gamma_n(\hat{\rho}),nJ-1}(\beta),$$

the noncentrality parameter, $\gamma_n(\hat{\rho})$, is given by

$$\gamma_n(\hat{\rho}) \equiv \nu(\hat{\rho})\sqrt{n}$$

and

$$\nu(\rho) \equiv -\Phi^{-1}(\alpha)/\sqrt{1-\rho^2+\rho^2/J}$$

Proof: Let $\epsilon > 0$ be given. Find an $r \times r$ square, R, on the plane such that

$$\mathrm{Prob}((X,Y) \in R^c) < \epsilon \tag{99}$$

for (X,Y) having a

$$\mathrm{N}\left(\mathbf{0}, \begin{pmatrix} 1/2 & 0 \\ 0 & 1 \end{pmatrix}\right)$$

distribution.

Let X_m have a chi-squared distribution with m degrees of freedom. By Lemma 9 we can find a finite M that bounds the probability density function of $\sqrt{m}(\sqrt{X_m/m}-1)$ for all m.

Define

$$\hat{c}_n \equiv \left(\hat{k}_n + \Phi^{-1}(\alpha)\right)\sqrt{n}/\sqrt{1-\rho^2+\rho^2/J}$$

and

$$\hat{d}_n \equiv \hat{k}_n\sqrt{n}/\left(\sqrt{1-\rho^2+\rho^2/J}\sqrt{nJ-1}\right)$$

By the corollary to Lemma 12,

$$\hat{c}_n \xrightarrow{p} c = \Phi^{-1}(\beta)\sqrt{v(\rho)}$$

and

$$\hat{d}_n \xrightarrow{p} d = -\Phi^{-1}(\alpha)/\left(\sqrt{1-\rho^2+\rho^2/J}\sqrt{J}\right)$$

where

$$v(\rho) \equiv \frac{\nu^2(\rho)}{2J}+1$$

Thus by Lemma 8 we can divide R into a $k \times k$ grid of subsquares, and find a fixed collection C_1 of subsquares of R with total area less than $\epsilon/\left(M/\sqrt{2\pi}\right)$, and an N_0 such that $n > N_0$ implies that with probability greater than $1-\epsilon$, all of the subsquares of R that intersect the line $y = \hat{c}_n + \hat{d}_n x$ lie in C_1. Let R_1 denote the region covered by the elements of C_1. Let $a_0 < a_1 < \ldots < a_k$ be the x coordinates of the vertices of the k^2 subsquares of R. Let $b_0 < b_1 < \ldots < b_k$ be the y coordinates of the vertices of the k^2 subsquares of R.

Let $V_{n,1}$, $V_{n,2}$, $U_{n,1}$, $U_{n,2}$, and $U_{n,3}$ be defined as in (3).

Now we express the central probability in the theorem in terms of $U_{n,1}$ and $V_{n,1}$:

We have

$$\begin{aligned}
& \mathrm{Prob}(\bar{Y}_{n,1} - \hat{k}_n S_{n,1} \le \mu_1 + \Phi^{-1}(\alpha)\sigma_Y) \qquad (100)\\
= \; & \mathrm{Prob}(\bar{Y}_{n,1} - \mu_1 \le \Phi^{-1}(\alpha)\sigma_Y + \hat{k}_n S_{n,1}) \\
= \; & \mathrm{Prob}\left((\bar{Y}_{n,1}-\mu_1)/\left((\sigma_Y/\sqrt{n})\sqrt{1-\rho^2+\rho^2/J}\right) \le \hat{c}_n + \hat{d}_n\sqrt{nJ-1}(S_{n,1}/\sigma_Y - 1)\right) \\
= \; & \mathrm{Prob}(V_{n,1} \le \hat{c}_n + \hat{d}_n U_{n,1})
\end{aligned}$$

Next we show that

$$\begin{aligned} & \text{Prob}((U_{n,1}, V_{n,1}) \in [a_i, a_{i+1}] \times [b_j, b_{j+1}]) \\ \approx \; & \text{Prob}(U_{n,2} \in [a_i, a_{i+1}]) \times \text{Prob}(V_{n,2} \in [b_j, b_{j+1}]) \end{aligned}$$

This argument concludes with result (103):

By Lemmas 2 and 9 we can find a $\delta_1 > 0$ such that for all n

$$\text{Prob}(U_{n,3} \in \text{arbitrary interval of length smaller than } \delta_1) < \epsilon/k^2 \tag{101}$$

and since (Lemma 2) $V_{n,2} \sim \text{N}(0,1)$, we can find a $\delta_2 > 0$ such that for all n

$$\text{Prob}(V_{n,2} \in \text{arbitrary interval of length smaller than } \delta_2) < \epsilon/k^2 \tag{102}$$

Let $\delta \equiv \min(\delta_1, \delta_2)/2$.

By Corollary 6.4 and Lemmas 3, 10, and 14, there exists an $N_1 > N_0$ such that for $n > N_1$, given any $i \in \{0, 1, \ldots, k-1\}$, $j \in \{0, 1, \ldots, k-1\}$,

$$\begin{aligned} & \text{Prob}(U_{n,2} \in [a_i + \delta, a_{i+1} - \delta])\text{Prob}(V_{n,2} \in [b_j + \delta, b_{j+1} - \delta]) - \epsilon/k^2 \\ \leq \; & \text{Prob}((U_{n,1}, V_{n,1}) \in [a_i, a_{i+1}] \times [b_j, b_{j+1}]) \\ \leq \; & \text{Prob}(U_{n,2} \in [a_i - \delta, a_{i+1} + \delta])\text{Prob}(V_{n,2} \in [b_j - \delta, b_{j+1} + \delta]) + \epsilon/k^2 \end{aligned} \tag{103}$$

We now show that

$$\begin{aligned} & \text{Prob}((U_{n,1}, V_{n,1}) \in [a_i, a_{i+1}] \times [b_j, b_{j+1}]) \\ \approx \; & \text{Prob}(U_{n,3} \in [a_i, a_{i+1}]) \times \text{Prob}(V_{n,2} \in [b_j, b_{j+1}]) \end{aligned}$$

This argument concludes with result (108):

By Corollary 6.3 and the corollary to Lemma 10, we can find an $N_2 > N_1$ such that for $n > N_2$, given any $i \in \{0, 1, \ldots, k-1\}$,

$$\text{Prob}(U_{n,3} \in [a_i + 2\delta, a_{i+1} - 2\delta]) - \epsilon/k^2 \leq \text{Prob}(U_{n,2} \in [a_i + \delta, a_{i+1} - \delta]) \tag{104}$$

and

$$\text{Prob}(U_{n,2} \in [a_i - \delta, a_{i+1} + \delta]) \leq \text{Prob}(U_{n,3} \in [a_i - 2\delta, a_{i+1} + 2\delta]) + \epsilon/k^2 \tag{105}$$

Thus from (103), (104), and (105), for $n > N_2$ and $i \in \{0, 1, \ldots, k-1\}$, $j \in \{0, 1, \ldots, k-1\}$,

$$\begin{aligned} & \text{Prob}(U_{n,3} \in [a_i + 2\delta, a_{i+1} - 2\delta])\text{Prob}(V_{n,2} \in [b_j + \delta, b_{j+1} - \delta]) - 2\epsilon/k^2 \\ \leq \; & \text{Prob}((U_{n,1}, V_{n,1}) \in [a_i, a_{i+1}] \times [b_j, b_{j+1}]) \\ \leq \; & \text{Prob}(U_{n,3} \in [a_i - 2\delta, a_{i+1} + 2\delta])\text{Prob}(V_{n,2} \in [b_j + \delta, b_{j+1} - \delta]) + 2\epsilon/k^2 \end{aligned} \tag{106}$$

We can conclude from (106) that for $n > N_2$ and $i \in \{0, 1, \ldots, k-1\}$, $j \in \{0, 1, \ldots, k-1\}$,

$$\begin{aligned} & [\text{Prob}(U_{n,3} \in [a_i, a_{i+1}]) - \text{Prob}(U_{n,3} \in [a_i, a_i + 2\delta]) - \text{Prob}(U_{n,3} \in [a_{i+1} - 2\delta, a_{i+1}])] \\ & \times[\text{Prob}(V_{n,2} \in [b_j, b_{j+1}]) - \text{Prob}(V_{n,2} \in [b_j, b_j + \delta]) - \text{Prob}(V_{n,2} \in [b_{j+1} - \delta, b_{j+1}])] - 2\epsilon/k^2 \\ \leq \; & \text{Prob}((U_{n,1}, V_{n,1}) \in [a_i, a_{i+1}] \times [b_j, b_{j+1}]) \\ \leq \; & [\text{Prob}(U_{n,3} \in [a_i, a_{i+1}]) + \text{Prob}(U_{n,3} \in [a_i - 2\delta, a_i]) + \text{Prob}(U_{n,3} \in [a_{i+1}, a_{i+1} + 2\delta])] \\ & \times[\text{Prob}(V_{n,2} \in [b_j, b_{j+1}]) + \text{Prob}(V_{n,2} \in [b_j - \delta, b_j]) + \text{Prob}(V_{n,2} \in [b_{j+1}, b_{j+1} + \delta])] + 2\epsilon/k^2 \end{aligned} \tag{107}$$

or (recall that 2δ was chosen to satisfy (101) and (102))

$$\begin{aligned} & \operatorname{Prob}(U_{n,3} \in [a_i, a_{i+1}])\operatorname{Prob}(V_{n,2} \in [b_j, b_{j+1}]) - 10\epsilon/k^2 \\ \leq\ & \operatorname{Prob}((U_{n,1}, V_{n,1}) \in [a_i, a_{i+1}] \times [b_j, b_{j+1}]) \\ \leq\ & \operatorname{Prob}(U_{n,3} \in [a_i, a_{i+1}])\operatorname{Prob}(V_{n,2} \in [b_j, b_{j+1}]) + 10\epsilon/k^2 \end{aligned} \tag{108}$$

Next we show that with high probability, $U_{n,1}$, $V_{n,1}$ lies inside the square R introduced in the second line of the proof. This argument ends with inequality (113):

From (108) we have, for $n > N_2$,

$$\begin{aligned} & \operatorname{Prob}(U_{n,3} \in [a_0, a_k])\operatorname{Prob}(V_{n,2} \in [b_0, b_k]) - 10\epsilon \\ \leq\ & \operatorname{Prob}((U_{n,1}, V_{n,1}) \in R) \\ \leq\ & \operatorname{Prob}(U_{n,3} \in [a_0, a_k])\operatorname{Prob}(V_{n,2} \in [b_0, b_k]) + 10\epsilon \end{aligned} \tag{109}$$

By Corollary 6.5, $U_{n,3} \xrightarrow{D} \mathrm{N}(0, 1/2)$. Thus we can find an $N_3 > N_2$ such that for $n > N_3$,

$$\begin{aligned} & \operatorname{Prob}(\mathrm{N}(0,1/2) \in [a_0, a_k]) - \epsilon \\ \leq\ & \operatorname{Prob}(U_{n,3} \in [a_0, a_k]) \\ \leq\ & \operatorname{Prob}(\mathrm{N}(0,1/2) \in [a_0, a_k]) + \epsilon \end{aligned} \tag{110}$$

So for $n > N_3$,

$$\begin{aligned} & \operatorname{Prob}(\mathrm{N}(0,1/2) \in [a_0, a_k])\operatorname{Prob}(V_{n,2} \in [b_0, b_k]) - 11\epsilon \\ \leq\ & \operatorname{Prob}((U_{n,1}, V_{n,1}) \in R) \\ \leq\ & \operatorname{Prob}(\mathrm{N}(0,1/2) \in [a_0, a_k])\operatorname{Prob}(V_{n,2} \in [b_0, b_k]) + 11\epsilon \end{aligned} \tag{111}$$

Since (Lemma 2) $V_{n,2} \sim \mathrm{N}(0,1)$, (99) and (111) imply that for $n > N_3$,

$$\operatorname{Prob}((U_{n,1}, V_{n,1}) \in R) > 1 - 12\epsilon \tag{112}$$

or

$$\operatorname{Prob}((U_{n,1}, V_{n,1}) \in R^c) < 12\epsilon \tag{113}$$

Now define

$$\begin{aligned} \hat{C}_{n,1} &\equiv \text{ subsquares of } R \text{ that intersect the line } y = \hat{c}_n + \hat{d}_n x \\ \hat{R}_{n,1} &\equiv \text{ the region covered by the elements of } \hat{C}_{n,1} \\ \hat{C}_{n,2} &\equiv \text{ subsquares of } R \text{ that lie entirely below the line } y = \hat{c}_n + \hat{d}_n x \\ \hat{R}_{n,2} &\equiv \text{ the region covered by the elements of } \hat{C}_{n,2} \\ C_{n,2} &\equiv \text{ subsquares of } R \text{ that lie entirely below the line } y = c + dx \end{aligned}$$

and

$$R_{n,2} \equiv \text{the region covered by the elements of } C_{n,2}.$$

Next we demonstrate that the probability that $(U_{n,1}, V_{n,1})$ lies in $\hat{R}_{n,1}$ is small. This argument concludes at result (116):

Now (C_1 and R_1 are defined just prior to equation (100)), for $n > N_3(> N_0)$, by the method of construction of the subsquares,

$$\begin{aligned}\mathrm{Prob}((U_{n,1},V_{n,1})\in\hat{R}_{n,1}) &= \mathrm{Prob}(\{(U_{n,1},V_{n,1})\in\hat{R}_{n,1}\}\cap\{\hat{C}_{n,1}\subset C_1\}^c)\\ &\quad+\mathrm{Prob}(\{(U_{n,1},V_{n,1})\in\hat{R}_{n,1}\}\cap\{\hat{C}_{n,1}\subset C_1\})\\ &< \epsilon+\mathrm{Prob}(\{(U_{n,1},V_{n,1})\in\hat{R}_{n,1}\}\cap\{\hat{C}_{n,1}\subset C_1\})\end{aligned}\tag{114}$$

By Lemma 2, inequality (108), and the method of construction of the subsquares, for $n > N_3$,

$$\begin{aligned}&\mathrm{Prob}(\{(U_{n,1},V_{n,1})\in\hat{R}_{n,1}\}\cap\{\hat{C}_{n,1}\subset C_1\})\\ \le\ &\mathrm{Prob}\left(\text{independent}\sqrt{nJ-1}\left(\sqrt{\chi^2_{nJ-1}/(nJ-1)}-1\right),\mathrm{N}(0,1)\in R_1\right)+10\epsilon\\ <\ &[\epsilon/(M/\sqrt{2\pi})]M/\sqrt{2\pi}+10\epsilon=11\epsilon\end{aligned}\tag{115}$$

Thus by results (114) and (115), for $n > N_3$

$$\mathrm{Prob}((U_{n,1},V_{n,1})\in\hat{R}_{n,1})<12\epsilon\tag{116}$$

Next we establish that the probability that $(U_{n,1},\ V_{n,1})$ lies below the line $y=\hat{c}_n+\hat{d}_nx$ is very close to the probability that $(U_{n,1},\ V_{n,1})$ lies below the line $y=c+dx$. This argument concludes at result (118):

It is clear that

$$\begin{aligned}&|\mathrm{Prob}((U_{n,1},V_{n,1})\in\hat{R}_{n,2})-\mathrm{Prob}((U_{n,1},V_{n,1})\in R_{n,2})|\\ \le\ &\mathrm{Prob}(\{(U_{n,1},V_{n,1})\in\hat{R}_{n,2}\}\cap\{(U_{n,1},V_{n,1})\in R^c_{n,2}\})\\ &+\mathrm{Prob}(\{(U_{n,1},V_{n,1})\in\hat{R}^c_{n,2}\}\cap\{(U_{n,1},V_{n,1})\in R_{n,2}\})\end{aligned}$$

Let E_0 denote the event

$$\{\text{all of the subsquares of } R \text{ that lie strictly below } y=\hat{c}_n+\hat{d}_nx$$
$$\text{but not strictly below } y=c+dx \text{ lie in } C_1\}.$$

We have

$$\begin{aligned}&\mathrm{Prob}(\{(U_{n,1},V_{n,1})\in\hat{R}_{n,2}\}\cap\{(U_{n,1},V_{n,1})\in R^c_{n,2}\})\\ =\ &\mathrm{Prob}(\{(U_{n,1},V_{n,1})\in\hat{R}_{n,2}\}\cap\{(U_{n,1},V_{n,1})\in R^c_{n,2}\}\cap E_0)\\ &+\mathrm{Prob}(\{(U_{n,1},V_{n,1})\in\hat{R}_{n,2}\}\cap\{(U_{n,1},V_{n,1})\in R^c_{n,2}\}\cap E^c_0)\end{aligned}\tag{117}$$

By Corollary 8.2 and the construction of the subsquares of R, the second term on the right hand side in (117) is less than ϵ. Clearly the first term in (117) is bounded by $\mathrm{Prob}(\{(U_{n,1},V_{n,1})\in R_1\})$. Similarly, by Corollary 8.1,

$$\mathrm{Prob}(\{(U_{n,1},V_{n,1})\in\hat{R}^c_{n,2}\}\cap\{(U_{n,1},V_{n,1})\in R_{n,2}\})\le\mathrm{Prob}(\{(U_{n,1},V_{n,1})\in R_1\})+\epsilon$$

Thus for $n > N_3$, result (108) implies that

$$\begin{aligned}&|\mathrm{Prob}((U_{n,1},V_{n,1})\in\hat{R}_{n,2})-\mathrm{Prob}((U_{n,1},V_{n,1})\in R_{n,2})|\\ \le\ &\mathrm{Prob}(\{(U_{n,1},V_{n,1})\in\hat{R}_{n,2}\}\cap\{(U_{n,1},V_{n,1})\in R^c_{n,2}\})\\ &+\mathrm{Prob}(\{(U_{n,1},V_{n,1})\in\hat{R}^c_{n,2}\}\cap\{(U_{n,1},V_{n,1})\in R_{n,2}\})\\ \le\ &2\times\mathrm{Prob}((U_{n,1},V_{n,1})\in R_1)+2\epsilon\\ \le\ &2\times\mathrm{Prob}\left(\text{independent }\sqrt{nJ-1}\left(\sqrt{\chi^2_{nJ-1}/(nJ-1)}-1\right),\mathrm{N}(0,1)\in R_1\right)+20\epsilon+2\epsilon\\ <\ &2[\epsilon/(M/\sqrt{2\pi})]M/\sqrt{2\pi}+22\epsilon=24\epsilon\end{aligned}\tag{118}$$

Next we will establish that for the "easy" variables E_n, F_n defined below, we obtain results (119), (120), and (121) analogous to results (113), (116), and (118) for $U_{n,1}, V_{n,1}$. We also obtain result (122) which relates the $U_{n,1}, V_{n,1}$ probability content of the subsquares below the line to the E_n, F_n content. Result (123) establishes that E_n and F_n combine with quantities $\hat{c}_n$ and $\hat{d}_n$ to yield the correct asymptotic coverage:

Let

$$E_n \sim \sqrt{nJ-1}\left(\sqrt{\chi^2_{nJ-1}/(nJ-1)} - 1\right)$$

and

$$F_n \sim \mathrm{N}(0,1),$$

and let E_n, F_n be independent. Then by Corollary 6.5 and the construction of R (R is defined on the second line of the proof),

$$\begin{aligned}
& \mathrm{Prob}((E_n, F_n) \in R) \\
= \;& \mathrm{Prob}\left(\sqrt{nJ-1}\left(\sqrt{\chi^2_{nJ-1}/(nJ-1)} - 1\right) \in [a_0, a_k]\right) \mathrm{Prob}(\mathrm{N}(0,1) \in [b_0, b_k]) \\
\to \;& \mathrm{Prob}(\mathrm{N}(0,1/2) \in [a_0, a_k])\mathrm{Prob}(\mathrm{N}(0,1) \in [b_0, b_k]) > 1-\epsilon
\end{aligned}$$

So we can find an $N_4 > N_3$ such that $n > N_4$ implies

$$\mathrm{Prob}((E_n, F_n) \in R^c) < \epsilon \tag{119}$$

Next note that by reasoning similar to that used to establish (116), for $n > N_4$,

$$\mathrm{Prob}((E_n, F_n) \in \hat{R}_{n,1}) < 2\epsilon \tag{120}$$

By reasoning similar to that used to establish (118), for $n > N_4$,

$$|\mathrm{Prob}((E_n, F_n) \in \hat{R}_{n,2}) - \mathrm{Prob}((E_n, F_n) \in R_{n,2})| < 4\epsilon \tag{121}$$

By Lemma 2 and result (108), for $n > N_4$,

$$|\mathrm{Prob}((U_{n,1}, V_{n,1}) \in R_{n,2}) - \mathrm{Prob}((E_n, F_n) \in R_{n,2})| < 10\epsilon \tag{122}$$

Now, by the definitions of $\hat{k}_n$, $\hat{c}_n$, and $\hat{d}_n$ (definitions (90), (92), and(93)), and by Lemma 12,

$$\begin{aligned}
& \mathrm{Prob}(F_n \le \hat{c}_n + \hat{d}_n E_n) \qquad (123) \\
= \;& \mathrm{Prob}\Big(\mathrm{N}(0,1) \le \left(\hat{k}_n + \Phi^{-1}(\alpha)\right)\sqrt{n}/\sqrt{1-\rho^2+\rho^2/J} \\
& + \left(\hat{k}_n/\sqrt{1-\rho^2+\rho^2/J}\right)\sqrt{n/(nJ-1)}\sqrt{nJ-1}\left(\sqrt{\chi^2_{nJ-1}/(nJ-1)} - 1\right)\Big) \\
= \;& \mathrm{Prob}\left(\mathrm{N}(0,1) \le \Phi^{-1}(\alpha)\sqrt{n}/\sqrt{1-\rho^2+\rho^2/J} + (\hat{k}_n\sqrt{n}/\sqrt{1-\rho^2+\rho^2/J}\,)\sqrt{\chi^2_{nJ-1}/(nJ-1)}\right) \\
= \;& \mathrm{Prob}\Big(\mathrm{N}(0,1) \le \Phi^{-1}(\alpha)\sqrt{n}/\sqrt{1-\rho^2+\rho^2/J} \\
& + \left(\sqrt{1-\hat{\rho}^2+\hat{\rho}^2/J}/\sqrt{1-\rho^2+\rho^2/J}\right)\mathrm{F}^{-1}_{\mathrm{nct},\gamma_n(\hat{\rho}),nJ-1}(\beta)\sqrt{\chi^2_{nJ-1}/(nJ-1)}\Big) \\
= \;& \mathrm{Prob}\left(X_{\mathrm{nct},\gamma_n(\rho),nJ-1} \le \left(\sqrt{1-\hat{\rho}^2+\hat{\rho}^2/J}/\sqrt{1-\rho^2+\rho^2/J}\right)\mathrm{F}^{-1}_{\mathrm{nct},\gamma_n(\hat{\rho}),nJ-1}(\beta)\right) \to \beta
\end{aligned}$$

as $n \to \infty$.

Now we will assemble all of the preceding pieces to yield the desired result:

We have

$$
\begin{aligned}
& |\text{Prob}\left(\bar{Y}_{n,1} - \hat{k}_n S_{n,1} \leq \mu_1 + \Phi^{-1}(\alpha)\sigma_Y\right) - \beta| \qquad (124)\\
\leq\ & |\text{Prob}\left(\bar{Y}_{n,1} - \hat{k}_n S_{n,1} \leq \mu_1 + \Phi^{-1}(\alpha)\sigma_Y\right) - \text{Prob}\left(F_n \leq \hat{c}_n + \hat{d}_n E_n\right)| + |\text{Prob}\left(F_n \leq \hat{c}_n + \hat{d}_n E_n\right) - \beta|
\end{aligned}
$$

Now by (100), the first term on the right hand side of (124) equals

$$
\begin{aligned}
& |\text{Prob}\left(V_{n,1} \leq \hat{c}_n + \hat{d}_n U_{n,1}\right) - \text{Prob}\left(F_n \leq \hat{c}_n + \hat{d}_n E_n\right)| \\
=\ & |\text{Prob}\left(V_{n,1} \leq \hat{c}_n + \hat{d}_n U_{n,1} \text{ and } (U_{n,1}, V_{n,1}) \in R^c\right) \\
& + \text{Prob}\left(V_{n,1} \leq \hat{c}_n + \hat{d}_n U_{n,1} \text{ and } (U_{n,1}, V_{n,1}) \in \hat{R}_{n,1}\right) \\
& + \text{Prob}\left(V_{n,1} \leq \hat{c}_n + \hat{d}_n U_{n,1} \text{ and } (U_{n,1}, V_{n,1}) \in \hat{R}_{n,2}\right) \\
& - \text{Prob}\left(F_n \leq \hat{c}_n + \hat{d}_n E_n \text{ and } (E_n, F_n) \in R^c\right) \\
& - \text{Prob}\left(F_n \leq \hat{c}_n + \hat{d}_n E_n \text{ and } (E_n, F_n) \in \hat{R}_{n,1}\right) \\
& - \text{Prob}\left(F_n \leq \hat{c}_n + \hat{d}_n E_n \text{ and } (E_n, F_n) \in \hat{R}_{n,2}\right)| \\
\leq\ & \text{Prob}\left((U_{n,1}, V_{n,1}) \in R^c\right) + \text{Prob}\left((E_n, F_n) \in R^c\right) \\
& + \text{Prob}\left((U_{n,1}, V_{n,1}) \in \hat{R}_{n,1}\right) + \text{Prob}\left((E_n, F_n) \in \hat{R}_{n,1}\right) \\
& + |\text{Prob}\left((U_{n,1}, V_{n,1}) \in \hat{R}_{n,2}\right) - \text{Prob}\left((U_{n,1}, V_{n,1}) \in R_{n,2}\right)| \\
& + |\text{Prob}\left((U_{n,1}, V_{n,1}) \in R_{n,2}\right) - \text{Prob}\left((E_n, F_n) \in R_{n,2}\right)| \\
& + |\text{Prob}\left((E_n, F_n) \in R_{n,2}\right) - \text{Prob}\left((E_n, F_n) \in \hat{R}_{n,2}\right)|
\end{aligned}
$$

By results (113), (119), (116), (120), (118), (122), and (121), for $n > N_4$, this last sum is less than

$$12\epsilon + \epsilon + 12\epsilon + 2\epsilon + 24\epsilon + 10\epsilon + 4\epsilon \qquad (125)$$

As ϵ was arbitrary, results (124), (125), and (123) complete the proof.

Table 1: Observed confidence interval coverage of the 0.01 quantile when the nominal coverage equals 0.75, $J = 2$

		coverage[1]			
ρ	n	incorrect	version 1	version 2	MLE
0.70	5	0.7762	0.6903	0.6410	0.5663
	10	0.7695	0.7218	0.6823	0.6258
	20	0.7728	0.7258	0.6965	0.6585
	40	0.7798	0.7400	0.7215	0.6970
0.75	5	0.8003	0.7083	0.6540	0.5663
	10	0.7752	0.7175	0.6760	0.6245
	20	0.7853	0.7325	0.7027	0.6690
	40	0.7915	0.7500	0.7275	0.7067
0.80	5	0.7940	0.7105	0.6570	0.5733
	10	0.7917	0.7302	0.6937	0.6358
	20	0.7973	0.7505	0.7202	0.6867
	40	0.7970	0.7392	0.7190	0.6950
0.85	5	0.8065	0.7222	0.6670	0.5850
	10	0.8163	0.7422	0.7007	0.6452
	20	0.7915	0.7372	0.7103	0.6760
	40	0.7967	0.7372	0.7188	0.6975
0.90	5	0.8215	0.7235	0.6743	0.5887
	10	0.8247	0.7432	0.7055	0.6512
	20	0.8157	0.7412	0.7130	0.6770
	40	0.8253	0.7508	0.7250	0.7045
0.95	5	0.8320	0.7278	0.6803	0.5980
	10	0.8417	0.7450	0.7003	0.6462
	20	0.8337	0.7425	0.7127	0.6775
	40	0.8365	0.7502	0.7288	0.7043
0.99	5	0.8423	0.7285	0.6810	0.5915
	10	0.8582	0.7585	0.7175	0.6600
	20	0.8525	0.7528	0.7255	0.6883
	40	0.8515	0.7560	0.7348	0.7120

[1]"incorrect" denotes the incorrect standard approach; "version 1" denotes a predictor sort approach using the section A.3 consistent estimator of ρ; "version 2" denotes a predictor sort approach using the maximum likelihood estimator of ρ; and "MLE" denotes the full maximum likelihood approach presented in Appendix A.

Table 2: Observed confidence interval coverage of the 0.01 quantile when the nominal coverage equals 0.75, $J = 4$

		coverage			
ρ	n	incorrect	version 1	version 2	MLE
0.70	5	0.8025	0.6657	0.6292	0.5903
	10	0.7945	0.6940	0.6735	0.6452
	20	0.7883	0.7107	0.6943	0.6747
	40	0.7840	0.7208	0.7090	0.6947
0.75	5	0.8090	0.6760	0.6405	0.5940
	10	0.8047	0.7030	0.6775	0.6465
	20	0.8073	0.7143	0.6943	0.6717
	40	0.7925	0.7248	0.7127	0.6970
0.80	5	0.8243	0.6855	0.6550	0.6112
	10	0.8140	0.7087	0.6885	0.6520
	20	0.8230	0.7250	0.6980	0.6755
	40	0.8153	0.7345	0.7235	0.7095
0.85	5	0.8518	0.6963	0.6647	0.6148
	10	0.8393	0.7170	0.6913	0.6530
	20	0.8415	0.7362	0.7165	0.6937
	40	0.8430	0.7475	0.7335	0.7153
0.90	5	0.8548	0.6957	0.6735	0.6195
	10	0.8618	0.7248	0.7000	0.6610
	20	0.8652	0.7338	0.7153	0.6963
	40	0.8565	0.7378	0.7265	0.7100
0.95	5	0.8895	0.7100	0.6933	0.6450
	10	0.8962	0.7390	0.7123	0.6770
	20	0.8878	0.7470	0.7288	0.7067
	40	0.8882	0.7502	0.7368	0.7192
0.99	5	0.9042	0.7160	0.6993	0.6442
	10	0.9130	0.7388	0.7155	0.6785
	20	0.9245	0.7502	0.7302	0.7035
	40	0.9257	0.7498	0.7410	0.7208

Table 3: Observed confidence interval coverage of the 0.01 quantile when the nominal coverage equals 0.75, $J = 8$

		coverage			
ρ	n	incorrect	version 1	version 2	MLE
0.70	5	0.8025	0.6298	0.6108	0.5787
	10	0.8067	0.6735	0.6617	0.6432
	20	0.7985	0.7050	0.7005	0.6853
	40	0.7965	0.7348	0.7260	0.7173
0.75	5	0.8285	0.6442	0.6292	0.5998
	10	0.8170	0.6717	0.6633	0.6432
	20	0.8190	0.7060	0.6950	0.6833
	40	0.8055	0.7290	0.7225	0.7147
0.80	5	0.8243	0.6550	0.6472	0.6155
	10	0.8413	0.6983	0.6820	0.6635
	20	0.8275	0.7103	0.7015	0.6863
	40	0.8317	0.7400	0.7302	0.7220
0.85	5	0.8550	0.6623	0.6645	0.6348
	10	0.8582	0.6985	0.6883	0.6677
	20	0.8650	0.7415	0.7328	0.7183
	40	0.8482	0.7235	0.7150	0.7053
0.90	5	0.8935	0.6773	0.6785	0.6430
	10	0.8900	0.7137	0.7073	0.6877
	20	0.8805	0.7212	0.7070	0.6935
	40	0.8790	0.7382	0.7292	0.7165
0.95	5	0.9160	0.6885	0.6850	0.6577
	10	0.9187	0.7087	0.7037	0.6785
	20	0.9203	0.7368	0.7322	0.7165
	40	0.9267	0.7495	0.7380	0.7268
0.99	5	0.9515	0.7115	0.7183	0.6833
	10	0.9583	0.7365	0.7282	0.7055
	20	0.9603	0.7345	0.7258	0.7083
	40	0.9617	0.7440	0.7340	0.7220

Table 4: Observed confidence interval coverage of the 0.01 quantile when the nominal coverage equals 0.75, $J = 30$

		coverage			
ρ	n	incorrect	version 1	version 2	MLE
0.70	5	0.8300	0.5883	0.5982	0.5913
	10	0.8143	0.6510	0.6530	0.6472
	20	0.8125	0.6930	0.6953	0.6930
	40	0.8035	0.6915	0.6903	0.6883
0.75	5	0.8235	0.6148	0.6185	0.6078
	10	0.8280	0.6633	0.6667	0.6610
	20	0.8310	0.6875	0.6883	0.6853
	40	0.8160	0.7205	0.7170	0.7145
0.80	5	0.8445	0.6110	0.6295	0.6180
	10	0.8478	0.6615	0.6667	0.6597
	20	0.8370	0.7017	0.6997	0.6947
	40	0.8343	0.7320	0.7330	0.7300
0.85	5	0.8715	0.6160	0.6345	0.6242
	10	0.8708	0.6880	0.6885	0.6827
	20	0.8682	0.7020	0.7005	0.6957
	40	0.8758	0.7332	0.7282	0.7262
0.90	5	0.8985	0.6375	0.6607	0.6500
	10	0.9022	0.6883	0.6915	0.6850
	20	0.9075	0.7005	0.7047	0.7000
	40	0.9045	0.7208	0.7250	0.7202
0.95	5	0.9267	0.6450	0.6880	0.6710
	10	0.9455	0.6990	0.7133	0.7037
	20	0.9517	0.7230	0.7288	0.7228
	40	0.9470	0.7228	0.7210	0.7163
0.99	5	0.9715	0.6530	0.7095	0.6933
	10	0.9865	0.7007	0.7305	0.7190
	20	0.9918	0.7320	0.7392	0.7310
	40	0.9940	0.7395	0.7392	0.7350

Table 5: Observed confidence interval coverage of the 0.05 quantile when the nominal coverage equals 0.75, $J = 2$

ρ	n	coverage			
		incorrect	version 1	version 2	MLE
0.70	5	0.7795	0.7075	0.6595	0.5915
	10	0.7847	0.7228	0.6865	0.6420
	20	0.7730	0.7235	0.6947	0.6623
	40	0.7802	0.7460	0.7262	0.7047
0.75	5	0.7935	0.7013	0.6583	0.5897
	10	0.7917	0.7318	0.6990	0.6530
	20	0.7812	0.7340	0.7053	0.6733
	40	0.7788	0.7272	0.7145	0.6955
0.80	5	0.7950	0.7153	0.6690	0.5978
	10	0.7990	0.7395	0.6993	0.6530
	20	0.7913	0.7352	0.7113	0.6833
	40	0.7903	0.7410	0.7218	0.7007
0.85	5	0.8103	0.7225	0.6770	0.6028
	10	0.8010	0.7395	0.7010	0.6562
	20	0.8095	0.7355	0.7117	0.6787
	40	0.8047	0.7405	0.7240	0.7053
0.90	5	0.8147	0.7165	0.6670	0.5942
	10	0.8123	0.7292	0.6943	0.6468
	20	0.8190	0.7460	0.7208	0.6913
	40	0.8133	0.7402	0.7242	0.7045
0.95	5	0.8350	0.7332	0.6895	0.6095
	10	0.8357	0.7472	0.7117	0.6603
	20	0.8307	0.7422	0.7143	0.6833
	40	0.8237	0.7418	0.7245	0.7045
0.99	5	0.8525	0.7485	0.7053	0.6332
	10	0.8430	0.7470	0.7133	0.6597
	20	0.8452	0.7505	0.7235	0.6950
	40	0.8360	0.7388	0.7232	0.7013

Table 6: Observed confidence interval coverage of the 0.05 quantile when the nominal coverage equals 0.75, $J = 4$

		coverage			
ρ	n	incorrect	version 1	version 2	MLE
0.70	5	0.7903	0.6730	0.6448	0.6068
	10	0.8117	0.7150	0.6957	0.6733
	20	0.7895	0.7230	0.7123	0.6940
	40	0.7943	0.7340	0.7235	0.7085
0.75	5	0.7987	0.6827	0.6545	0.6185
	10	0.8103	0.7037	0.6840	0.6610
	20	0.8057	0.7208	0.7060	0.6877
	40	0.7975	0.7328	0.7200	0.7087
0.80	5	0.8345	0.7050	0.6727	0.6360
	10	0.8220	0.7255	0.7025	0.6773
	20	0.8257	0.7315	0.7180	0.6970
	40	0.8270	0.7470	0.7358	0.7222
0.85	5	0.8365	0.7060	0.6863	0.6415
	10	0.8407	0.7140	0.6953	0.6645
	20	0.8330	0.7305	0.7127	0.6923
	40	0.8355	0.7442	0.7338	0.7202
0.90	5	0.8552	0.7027	0.6743	0.6385
	10	0.8570	0.7295	0.7165	0.6840
	20	0.8550	0.7342	0.7150	0.6980
	40	0.8528	0.7450	0.7308	0.7188
0.95	5	0.8882	0.7120	0.7003	0.6515
	10	0.8902	0.7212	0.7060	0.6735
	20	0.8832	0.7438	0.7290	0.7080
	40	0.8818	0.7460	0.7345	0.7215
0.99	5	0.9155	0.7358	0.7190	0.6680
	10	0.9140	0.7355	0.7240	0.6867
	20	0.9187	0.7388	0.7220	0.7047
	40	0.9247	0.7598	0.7485	0.7350

Table 7: Observed confidence interval coverage of the 0.05 quantile when the nominal coverage equals 0.75, $J = 8$

		coverage			
ρ	n	incorrect	version 1	version 2	MLE
0.70	5	0.8077	0.6432	0.6348	0.6165
	10	0.8095	0.6895	0.6830	0.6700
	20	0.8137	0.7093	0.7030	0.6937
	40	0.7997	0.7302	0.7262	0.7180
0.75	5	0.8205	0.6645	0.6565	0.6335
	10	0.8337	0.6963	0.6857	0.6683
	20	0.8243	0.7218	0.7170	0.7073
	40	0.8125	0.7250	0.7212	0.7137
0.80	5	0.8363	0.6673	0.6625	0.6408
	10	0.8393	0.7040	0.6977	0.6815
	20	0.8375	0.7252	0.7170	0.7065
	40	0.8355	0.7280	0.7225	0.7160
0.85	5	0.8565	0.6765	0.6780	0.6515
	10	0.8610	0.7047	0.6955	0.6760
	20	0.8580	0.7278	0.7198	0.7110
	40	0.8530	0.7365	0.7260	0.7202
0.90	5	0.8888	0.6840	0.6877	0.6575
	10	0.8880	0.7137	0.7117	0.6927
	20	0.8860	0.7305	0.7208	0.7087
	40	0.8810	0.7372	0.7328	0.7252
0.95	5	0.9185	0.7003	0.7037	0.6713
	10	0.9143	0.7160	0.7105	0.6880
	20	0.9215	0.7298	0.7215	0.7080
	40	0.9253	0.7328	0.7188	0.7095
0.99	5	0.9463	0.6950	0.7027	0.6690
	10	0.9650	0.7308	0.7290	0.7097
	20	0.9645	0.7518	0.7375	0.7218
	40	0.9633	0.7405	0.7368	0.7258

Table 8: Observed confidence interval coverage of the 0.05 quantile when the nominal coverage equals 0.75, $J = 30$

		coverage			
ρ	n	incorrect	version 1	version 2	MLE
0.70	5	0.8155	0.6342	0.6360	0.6298
	10	0.8087	0.6743	0.6727	0.6687
	20	0.8090	0.7085	0.7025	0.7003
	40	0.8073	0.7200	0.7177	0.7163
0.75	5	0.8315	0.6268	0.6320	0.6248
	10	0.8225	0.6607	0.6665	0.6610
	20	0.8317	0.7060	0.7060	0.7037
	40	0.8263	0.7270	0.7272	0.7255
0.80	5	0.8538	0.6365	0.6555	0.6485
	10	0.8440	0.6807	0.6920	0.6875
	20	0.8492	0.7192	0.7192	0.7143
	40	0.8470	0.7418	0.7390	0.7360
0.85	5	0.8750	0.6542	0.6755	0.6657
	10	0.8765	0.6863	0.6975	0.6923
	20	0.8715	0.7095	0.7113	0.7073
	40	0.8770	0.7372	0.7312	0.7300
0.90	5	0.9045	0.6502	0.6823	0.6725
	10	0.9005	0.6967	0.7045	0.6983
	20	0.9030	0.7133	0.7173	0.7137
	40	0.9083	0.7242	0.7248	0.7230
0.95	5	0.9347	0.6535	0.6910	0.6777
	10	0.9537	0.6920	0.7157	0.7110
	20	0.9577	0.7308	0.7400	0.7360
	40	0.9537	0.7285	0.7285	0.7252
0.99	5	0.9790	0.6573	0.7315	0.7177
	10	0.9875	0.6977	0.7348	0.7260
	20	0.9918	0.7173	0.7310	0.7248
	40	0.9945	0.7335	0.7450	0.7395

Table 9: Observed confidence interval coverage of the 0.01 quantile when the nominal coverage equals 0.90, $J = 2$

ρ	n	coverage			
		incorrect	version 1	version 2	MLE
0.70	5	0.9135	0.8702	0.8375	0.7218
	10	0.9217	0.8865	0.8600	0.7887
	20	0.9193	0.8925	0.8740	0.8315
	40	0.9257	0.8932	0.8835	0.8558
0.75	5	0.9165	0.8782	0.8445	0.7330
	10	0.9233	0.8935	0.8725	0.8075
	20	0.9340	0.8882	0.8735	0.8320
	40	0.9283	0.8958	0.8835	0.8555
0.80	5	0.9283	0.8785	0.8488	0.7318
	10	0.9363	0.8942	0.8755	0.8093
	20	0.9293	0.8915	0.8780	0.8280
	40	0.9305	0.9020	0.8900	0.8598
0.85	5	0.9323	0.8878	0.8528	0.7428
	10	0.9417	0.8910	0.8680	0.7965
	20	0.9460	0.9000	0.8802	0.8413
	40	0.9350	0.8920	0.8810	0.8550
0.90	5	0.9420	0.8875	0.8605	0.7528
	10	0.9377	0.8925	0.8728	0.7980
	20	0.9527	0.8988	0.8828	0.8367
	40	0.9443	0.8892	0.8768	0.8472
0.95	5	0.9497	0.8742	0.8450	0.7382
	10	0.9537	0.9000	0.8718	0.7985
	20	0.9597	0.8908	0.8772	0.8333
	40	0.9577	0.8990	0.8885	0.8595
0.99	5	0.9620	0.8882	0.8560	0.7498
	10	0.9637	0.8938	0.8702	0.7980
	20	0.9673	0.8935	0.8785	0.8293
	40	0.9630	0.9028	0.8930	0.8620

Table 10: Observed confidence interval coverage of the 0.01 quantile when the nominal coverage equals 0.90, $J = 4$

ρ	n	coverage			
		incorrect	version 1	version 2	MLE
0.70	5	0.9285	0.8380	0.8170	0.7470
	10	0.9350	0.8712	0.8542	0.8113
	20	0.9263	0.8832	0.8708	0.8413
	40	0.9270	0.8780	0.8700	0.8560
0.75	5	0.9395	0.8595	0.8403	0.7678
	10	0.9437	0.8708	0.8532	0.8095
	20	0.9407	0.8832	0.8705	0.8430
	40	0.9413	0.8920	0.8840	0.8658
0.80	5	0.9533	0.8622	0.8420	0.7788
	10	0.9407	0.8758	0.8638	0.8195
	20	0.9485	0.8818	0.8702	0.8435
	40	0.9520	0.8952	0.8888	0.8708
0.85	5	0.9550	0.8695	0.8492	0.7840
	10	0.9580	0.8802	0.8695	0.8237
	20	0.9567	0.8902	0.8772	0.8508
	40	0.9593	0.8998	0.8900	0.8702
0.90	5	0.9593	0.8770	0.8630	0.7887
	10	0.9698	0.8852	0.8728	0.8320
	20	0.9710	0.8922	0.8795	0.8505
	40	0.9722	0.8940	0.8872	0.8672
0.95	5	0.9735	0.8775	0.8682	0.8083
	10	0.9832	0.8918	0.8810	0.8367
	20	0.9870	0.9002	0.8928	0.8670
	40	0.9858	0.9020	0.8932	0.8778
0.99	5	0.9888	0.8922	0.8860	0.8170
	10	0.9908	0.8968	0.8892	0.8478
	20	0.9918	0.8965	0.8878	0.8568
	40	0.9932	0.8952	0.8855	0.8640

Table 11: Observed confidence interval coverage of the 0.01 quantile when the nominal coverage equals 0.90, $J = 8$

ρ	n	coverage			
		incorrect	version 1	version 2	MLE
0.70	5	0.9340	0.8220	0.8130	0.7738
	10	0.9355	0.8620	0.8530	0.8283
	20	0.9375	0.8798	0.8745	0.8560
	40	0.9345	0.8820	0.8768	0.8665
0.75	5	0.9395	0.8263	0.8215	0.7778
	10	0.9517	0.8665	0.8555	0.8337
	20	0.9465	0.8798	0.8720	0.8552
	40	0.9447	0.8898	0.8870	0.8735
0.80	5	0.9500	0.8410	0.8355	0.7920
	10	0.9535	0.8608	0.8538	0.8320
	20	0.9545	0.8862	0.8805	0.8615
	40	0.9560	0.8918	0.8888	0.8778
0.85	5	0.9680	0.8480	0.8485	0.8045
	10	0.9725	0.8695	0.8640	0.8353
	20	0.9680	0.8812	0.8742	0.8538
	40	0.9800	0.8995	0.8920	0.8815
0.90	5	0.9708	0.8662	0.8618	0.8237
	10	0.9798	0.8702	0.8668	0.8430
	20	0.9822	0.8822	0.8798	0.8620
	40	0.9788	0.8975	0.8890	0.8778
0.95	5	0.9860	0.8672	0.8702	0.8275
	10	0.9922	0.8855	0.8830	0.8578
	20	0.9918	0.8962	0.8935	0.8725
	40	0.9940	0.9000	0.8955	0.8832
0.99	5	0.9962	0.8730	0.8790	0.8370
	10	0.9982	0.8845	0.8842	0.8540
	20	0.9992	0.8952	0.8922	0.8702
	40	0.9995	0.9065	0.8988	0.8865

Table 12: Observed confidence interval coverage of the 0.01 quantile when the nominal coverage equals 0.90, $J = 30$

		coverage			
ρ	n	incorrect	version 1	version 2	MLE
0.70	5	0.9373	0.7830	0.7895	0.7745
	10	0.9400	0.8343	0.8370	0.8290
	20	0.9417	0.8530	0.8515	0.8450
	40	0.9457	0.8762	0.8728	0.8690
0.75	5	0.9445	0.7870	0.8040	0.7885
	10	0.9475	0.8343	0.8385	0.8297
	20	0.9495	0.8712	0.8705	0.8632
	40	0.9493	0.8828	0.8798	0.8762
0.80	5	0.9523	0.7907	0.8043	0.7890
	10	0.9587	0.8430	0.8502	0.8390
	20	0.9643	0.8642	0.8630	0.8565
	40	0.9623	0.8702	0.8680	0.8655
0.85	5	0.9623	0.8033	0.8273	0.8113
	10	0.9695	0.8423	0.8528	0.8413
	20	0.9752	0.8675	0.8660	0.8590
	40	0.9782	0.8870	0.8835	0.8800
0.90	5	0.9775	0.8145	0.8430	0.8247
	10	0.9835	0.8535	0.8655	0.8552
	20	0.9842	0.8688	0.8745	0.8672
	40	0.9868	0.8872	0.8898	0.8852
0.95	5	0.9920	0.8173	0.8570	0.8410
	10	0.9958	0.8562	0.8745	0.8628
	20	0.9980	0.8790	0.8875	0.8782
	40	0.9980	0.8928	0.8930	0.8895
0.99	5	0.9995	0.8215	0.8835	0.8612
	10	1.0000	0.8625	0.8915	0.8792
	20	1.0000	0.8828	0.8950	0.8870
	40	1.0000	0.8932	0.8992	0.8930

Table 13: Observed confidence interval coverage of the 0.05 quantile when the nominal coverage equals 0.90, $J = 2$

		coverage			
ρ	n	incorrect	version 1	version 2	MLE
0.70	5	0.9225	0.8810	0.8550	0.7505
	10	0.9193	0.8855	0.8628	0.8047
	20	0.9225	0.8905	0.8760	0.8383
	40	0.9140	0.8872	0.8768	0.8490
0.75	5	0.9233	0.8822	0.8502	0.7578
	10	0.9295	0.8890	0.8702	0.8125
	20	0.9263	0.8965	0.8805	0.8415
	40	0.9220	0.8920	0.8828	0.8605
0.80	5	0.9237	0.8760	0.8485	0.7548
	10	0.9317	0.8898	0.8655	0.8060
	20	0.9323	0.8935	0.8805	0.8417
	40	0.9313	0.8955	0.8868	0.8582
0.85	5	0.9395	0.8842	0.8525	0.7550
	10	0.9405	0.8915	0.8658	0.8080
	20	0.9387	0.8888	0.8770	0.8380
	40	0.9420	0.9087	0.9000	0.8760
0.90	5	0.9415	0.8832	0.8600	0.7558
	10	0.9427	0.8875	0.8695	0.8155
	20	0.9567	0.9028	0.8870	0.8458
	40	0.9495	0.8998	0.8912	0.8655
0.95	5	0.9573	0.8878	0.8625	0.7668
	10	0.9605	0.8978	0.8835	0.8225
	20	0.9550	0.8945	0.8792	0.8387
	40	0.9553	0.8948	0.8845	0.8605
0.99	5	0.9620	0.8990	0.8715	0.7695
	10	0.9640	0.8948	0.8788	0.8217
	20	0.9683	0.9022	0.8888	0.8485
	40	0.9677	0.8970	0.8892	0.8605

Table 14: Observed confidence interval coverage of the 0.05 quantile when the nominal coverage equals 0.90, $J = 4$

ρ	n	coverage			
		incorrect	version 1	version 2	MLE
0.70	5	0.9270	0.8578	0.8415	0.7817
	10	0.9257	0.8752	0.8635	0.8275
	20	0.9350	0.8802	0.8712	0.8490
	40	0.9267	0.8905	0.8850	0.8710
0.75	5	0.9317	0.8538	0.8357	0.7827
	10	0.9425	0.8792	0.8648	0.8337
	20	0.9463	0.8882	0.8808	0.8608
	40	0.9467	0.8972	0.8922	0.8765
0.80	5	0.9387	0.8505	0.8390	0.7860
	10	0.9457	0.8732	0.8612	0.8260
	20	0.9527	0.8958	0.8840	0.8638
	40	0.9530	0.8980	0.8905	0.8785
0.85	5	0.9543	0.8625	0.8500	0.7963
	10	0.9620	0.8868	0.8750	0.8425
	20	0.9565	0.8938	0.8868	0.8652
	40	0.9690	0.9018	0.8985	0.8788
0.90	5	0.9615	0.8792	0.8690	0.8077
	10	0.9708	0.8942	0.8820	0.8460
	20	0.9750	0.8955	0.8850	0.8635
	40	0.9720	0.8960	0.8890	0.8732
0.95	5	0.9748	0.8708	0.8648	0.8047
	10	0.9800	0.8938	0.8848	0.8458
	20	0.9852	0.8968	0.8888	0.8598
	40	0.9860	0.8952	0.8910	0.8732
0.99	5	0.9865	0.8818	0.8752	0.8183
	10	0.9922	0.8918	0.8860	0.8427
	20	0.9918	0.9020	0.8925	0.8695
	40	0.9918	0.8968	0.8908	0.8702

Table 15: Observed confidence interval coverage of the 0.05 quantile when the nominal coverage equals 0.90, $J = 8$

		coverage			
ρ	n	incorrect	version 1	version 2	MLE
0.70	5	0.9357	0.8275	0.8210	0.7887
	10	0.9377	0.8678	0.8612	0.8445
	20	0.9403	0.8760	0.8698	0.8562
	40	0.9445	0.8972	0.8935	0.8855
0.75	5	0.9423	0.8363	0.8363	0.8007
	10	0.9463	0.8662	0.8638	0.8385
	20	0.9487	0.8832	0.8795	0.8672
	40	0.9427	0.8800	0.8770	0.8675
0.80	5	0.9497	0.8485	0.8468	0.8105
	10	0.9507	0.8725	0.8665	0.8475
	20	0.9537	0.8780	0.8718	0.8552
	40	0.9580	0.8852	0.8795	0.8712
0.85	5	0.9600	0.8518	0.8570	0.8260
	10	0.9683	0.8690	0.8700	0.8458
	20	0.9680	0.8850	0.8808	0.8690
	40	0.9690	0.8900	0.8852	0.8740
0.90	5	0.9748	0.8450	0.8518	0.8160
	10	0.9808	0.8742	0.8732	0.8465
	20	0.9815	0.8872	0.8878	0.8758
	40	0.9858	0.8955	0.8908	0.8802
0.95	5	0.9882	0.8615	0.8745	0.8317
	10	0.9920	0.8898	0.8882	0.8640
	20	0.9940	0.8960	0.8955	0.8778
	40	0.9952	0.9000	0.8930	0.8850
0.99	5	0.9982	0.8692	0.8880	0.8470
	10	0.9982	0.8725	0.8868	0.8630
	20	0.9990	0.8938	0.8880	0.8692
	40	0.9992	0.8930	0.8918	0.8800

Table 16: Observed confidence interval coverage of the 0.05 quantile when the nominal coverage equals 0.90, $J = 30$

		coverage			
ρ	n	incorrect	version 1	version 2	MLE
0.70	5	0.9403	0.7987	0.8055	0.7963
	10	0.9445	0.8508	0.8540	0.8480
	20	0.9440	0.8725	0.8732	0.8680
	40	0.9470	0.8800	0.8800	0.8772
0.75	5	0.9517	0.8073	0.8217	0.8115
	10	0.9510	0.8478	0.8498	0.8420
	20	0.9527	0.8750	0.8778	0.8735
	40	0.9595	0.8805	0.8808	0.8768
0.80	5	0.9570	0.8023	0.8215	0.8100
	10	0.9567	0.8492	0.8505	0.8455
	20	0.9675	0.8730	0.8742	0.8695
	40	0.9688	0.8828	0.8858	0.8840
0.85	5	0.9690	0.8200	0.8345	0.8197
	10	0.9738	0.8530	0.8625	0.8528
	20	0.9758	0.8750	0.8748	0.8685
	40	0.9778	0.8818	0.8812	0.8790
0.90	5	0.9800	0.8235	0.8492	0.8383
	10	0.9875	0.8540	0.8730	0.8622
	20	0.9910	0.8782	0.8808	0.8752
	40	0.9905	0.8885	0.8905	0.8872
0.95	5	0.9930	0.8247	0.8692	0.8562
	10	0.9958	0.8525	0.8802	0.8692
	20	0.9972	0.8755	0.8822	0.8755
	40	0.9995	0.8978	0.8960	0.8932
0.99	5	0.9995	0.8085	0.8890	0.8705
	10	1.0000	0.8612	0.8948	0.8848
	20	1.0000	0.8720	0.8868	0.8762
	40	1.0000	0.8925	0.8970	0.8908

Table 17: Observed confidence interval coverage of the 0.01 quantile when the nominal coverage equals 0.95, $J = 2$

		coverage			
ρ	n	incorrect	version 1	version 2	MLE
0.70	5	0.9617	0.9367	0.9143	0.7985
	10	0.9587	0.9423	0.9285	0.8562
	20	0.9583	0.9443	0.9327	0.8892
	40	0.9640	0.9523	0.9455	0.9185
0.75	5	0.9660	0.9345	0.9145	0.8033
	10	0.9665	0.9473	0.9307	0.8670
	20	0.9635	0.9483	0.9400	0.8980
	40	0.9698	0.9530	0.9475	0.9205
0.80	5	0.9613	0.9405	0.9215	0.8087
	10	0.9657	0.9395	0.9290	0.8688
	20	0.9673	0.9505	0.9393	0.8915
	40	0.9718	0.9473	0.9365	0.9120
0.85	5	0.9698	0.9387	0.9183	0.8083
	10	0.9735	0.9495	0.9365	0.8728
	20	0.9738	0.9447	0.9347	0.8895
	40	0.9732	0.9453	0.9400	0.9175
0.90	5	0.9695	0.9337	0.9157	0.8067
	10	0.9775	0.9495	0.9370	0.8702
	20	0.9810	0.9495	0.9395	0.9008
	40	0.9792	0.9483	0.9407	0.9147
0.95	5	0.9842	0.9410	0.9250	0.8123
	10	0.9825	0.9527	0.9400	0.8775
	20	0.9862	0.9510	0.9425	0.9070
	40	0.9860	0.9487	0.9410	0.9157
0.99	5	0.9875	0.9460	0.9320	0.8225
	10	0.9870	0.9485	0.9383	0.8752
	20	0.9900	0.9445	0.9337	0.8910
	40	0.9900	0.9515	0.9455	0.9185

Table 18: Observed confidence interval coverage of the 0.01 quantile when the nominal coverage equals 0.95, $J = 4$

		coverage			
ρ	n	incorrect	version 1	version 2	MLE
0.70	5	0.9667	0.9100	0.9000	0.8257
	10	0.9725	0.9313	0.9205	0.8785
	20	0.9712	0.9417	0.9373	0.9117
	40	0.9663	0.9395	0.9365	0.9183
0.75	5	0.9685	0.9183	0.9058	0.8377
	10	0.9752	0.9347	0.9243	0.8902
	20	0.9712	0.9380	0.9330	0.9073
	40	0.9800	0.9435	0.9377	0.9207
0.80	5	0.9780	0.9273	0.9137	0.8455
	10	0.9778	0.9340	0.9250	0.8812
	20	0.9770	0.9433	0.9360	0.9120
	40	0.9822	0.9487	0.9433	0.9257
0.85	5	0.9782	0.9330	0.9227	0.8495
	10	0.9865	0.9375	0.9295	0.8920
	20	0.9830	0.9420	0.9370	0.9080
	40	0.9860	0.9490	0.9460	0.9303
0.90	5	0.9845	0.9405	0.9335	0.8650
	10	0.9900	0.9335	0.9250	0.8862
	20	0.9912	0.9447	0.9383	0.9097
	40	0.9902	0.9450	0.9395	0.9210
0.95	5	0.9910	0.9377	0.9277	0.8612
	10	0.9968	0.9427	0.9363	0.8930
	20	0.9962	0.9445	0.9377	0.9127
	40	0.9965	0.9515	0.9483	0.9300
0.99	5	0.9978	0.9430	0.9407	0.8798
	10	0.9992	0.9503	0.9427	0.9030
	20	0.9982	0.9515	0.9467	0.9195
	40	0.9995	0.9513	0.9457	0.9295

Table 19: Observed confidence interval coverage of the 0.01 quantile when the nominal coverage equals 0.95, $J = 8$

		coverage			
ρ	n	incorrect	version 1	version 2	MLE
0.70	5	0.9712	0.8968	0.8928	0.8485
	10	0.9728	0.9177	0.9193	0.8920
	20	0.9750	0.9290	0.9240	0.9077
	40	0.9715	0.9375	0.9335	0.9243
0.75	5	0.9690	0.8972	0.8952	0.8512
	10	0.9772	0.9263	0.9230	0.8970
	20	0.9815	0.9313	0.9280	0.9113
	40	0.9750	0.9347	0.9315	0.9220
0.80	5	0.9755	0.9080	0.9080	0.8600
	10	0.9785	0.9345	0.9325	0.9075
	20	0.9838	0.9380	0.9350	0.9215
	40	0.9852	0.9487	0.9447	0.9355
0.85	5	0.9808	0.9090	0.9115	0.8688
	10	0.9875	0.9287	0.9297	0.9035
	20	0.9878	0.9317	0.9293	0.9130
	40	0.9912	0.9427	0.9413	0.9280
0.90	5	0.9922	0.9270	0.9303	0.8892
	10	0.9950	0.9407	0.9347	0.9115
	20	0.9952	0.9437	0.9397	0.9215
	40	0.9965	0.9440	0.9413	0.9305
0.95	5	0.9945	0.9247	0.9270	0.8838
	10	0.9980	0.9350	0.9350	0.9058
	20	0.9990	0.9510	0.9490	0.9297
	40	0.9990	0.9480	0.9467	0.9385
0.99	5	1.0000	0.9290	0.9373	0.8958
	10	1.0000	0.9490	0.9483	0.9187
	20	1.0000	0.9490	0.9463	0.9277
	40	1.0000	0.9467	0.9450	0.9305

Table 20: Observed confidence interval coverage of the 0.01 quantile when the nominal coverage equals 0.95, $J = 30$

		coverage			
ρ	n	incorrect	version 1	version 2	MLE
0.70	5	0.9705	0.8642	0.8712	0.8572
	10	0.9765	0.9083	0.9065	0.8990
	20	0.9772	0.9260	0.9257	0.9215
	40	0.9792	0.9335	0.9333	0.9300
0.75	5	0.9700	0.8808	0.8878	0.8730
	10	0.9810	0.9107	0.9115	0.9032
	20	0.9818	0.9277	0.9333	0.9267
	40	0.9818	0.9375	0.9365	0.9340
0.80	5	0.9805	0.8735	0.8880	0.8738
	10	0.9860	0.9217	0.9230	0.9145
	20	0.9865	0.9310	0.9335	0.9267
	40	0.9865	0.9395	0.9400	0.9367
0.85	5	0.9858	0.8795	0.9035	0.8878
	10	0.9888	0.9140	0.9187	0.9080
	20	0.9912	0.9387	0.9387	0.9337
	40	0.9932	0.9390	0.9345	0.9300
0.90	5	0.9945	0.8782	0.9073	0.8878
	10	0.9970	0.9163	0.9235	0.9137
	20	0.9975	0.9293	0.9315	0.9237
	40	0.9978	0.9367	0.9360	0.9313
0.95	5	0.9982	0.8955	0.9310	0.9133
	10	0.9998	0.9243	0.9320	0.9237
	20	0.9995	0.9435	0.9500	0.9425
	40	1.0000	0.9467	0.9487	0.9437
0.99	5	1.0000	0.8978	0.9395	0.9207
	10	1.0000	0.9250	0.9453	0.9335
	20	1.0000	0.9367	0.9417	0.9353
	40	1.0000	0.9405	0.9445	0.9387

Table 21: Observed confidence interval coverage of the 0.10 quantile when the nominal coverage equals 0.95, $J = 2$

		coverage			
ρ	n	incorrect	version 1	version 2	MLE
0.70	5	0.9605	0.9337	0.9157	0.8273
	10	0.9653	0.9465	0.9393	0.8918
	20	0.9635	0.9400	0.9303	0.8965
	40	0.9643	0.9465	0.9423	0.9215
0.75	5	0.9655	0.9437	0.9275	0.8375
	10	0.9650	0.9385	0.9267	0.8702
	20	0.9705	0.9530	0.9450	0.9127
	40	0.9705	0.9445	0.9425	0.9237
0.80	5	0.9670	0.9320	0.9170	0.8347
	10	0.9735	0.9420	0.9303	0.8860
	20	0.9715	0.9490	0.9415	0.9073
	40	0.9745	0.9467	0.9413	0.9237
0.85	5	0.9683	0.9395	0.9240	0.8353
	10	0.9738	0.9413	0.9345	0.8808
	20	0.9788	0.9443	0.9375	0.9062
	40	0.9800	0.9530	0.9483	0.9263
0.90	5	0.9740	0.9477	0.9345	0.8458
	10	0.9785	0.9483	0.9385	0.8845
	20	0.9775	0.9440	0.9360	0.9087
	40	0.9820	0.9480	0.9417	0.9227
0.95	5	0.9802	0.9485	0.9317	0.8455
	10	0.9838	0.9485	0.9373	0.8838
	20	0.9878	0.9520	0.9433	0.9135
	40	0.9842	0.9453	0.9390	0.9163
0.99	5	0.9862	0.9477	0.9370	0.8550
	10	0.9898	0.9495	0.9403	0.8905
	20	0.9892	0.9507	0.9405	0.9155
	40	0.9905	0.9515	0.9460	0.9210

Table 22: Observed confidence interval coverage of the 0.10 quantile when the nominal coverage equals 0.95, $J = 4$

		coverage			
ρ	n	incorrect	version 1	version 2	MLE
0.70	5	0.9683	0.9180	0.9093	0.8595
	10	0.9710	0.9337	0.9270	0.8950
	20	0.9702	0.9370	0.9333	0.9173
	40	0.9782	0.9497	0.9453	0.9357
0.75	5	0.9708	0.9260	0.9167	0.8698
	10	0.9782	0.9357	0.9285	0.9010
	20	0.9805	0.9493	0.9447	0.9295
	40	0.9790	0.9467	0.9435	0.9327
0.80	5	0.9718	0.9307	0.9267	0.8768
	10	0.9795	0.9403	0.9333	0.9038
	20	0.9860	0.9463	0.9400	0.9240
	40	0.9805	0.9423	0.9383	0.9270
0.85	5	0.9798	0.9277	0.9245	0.8728
	10	0.9868	0.9457	0.9380	0.9055
	20	0.9868	0.9403	0.9370	0.9200
	40	0.9895	0.9513	0.9467	0.9353
0.90	5	0.9865	0.9305	0.9275	0.8812
	10	0.9895	0.9383	0.9345	0.8998
	20	0.9920	0.9490	0.9447	0.9240
	40	0.9938	0.9495	0.9463	0.9330
0.95	5	0.9915	0.9345	0.9327	0.8785
	10	0.9945	0.9455	0.9385	0.9038
	20	0.9982	0.9490	0.9457	0.9247
	40	0.9960	0.9483	0.9473	0.9337
0.99	5	0.9978	0.9320	0.9330	0.8820
	10	0.9975	0.9385	0.9340	0.9022
	20	0.9995	0.9403	0.9377	0.9147
	40	0.9992	0.9455	0.9435	0.9305

Table 23: Observed confidence interval coverage of the 0.10 quantile when the nominal coverage equals 0.95, $J = 8$

		coverage			
ρ	n	incorrect	version 1	version 2	MLE
0.70	5	0.9695	0.9093	0.9075	0.8815
	10	0.9760	0.9333	0.9293	0.9133
	20	0.9762	0.9447	0.9427	0.9333
	40	0.9765	0.9485	0.9475	0.9417
0.75	5	0.9752	0.9113	0.9125	0.8818
	10	0.9762	0.9170	0.9165	0.8988
	20	0.9802	0.9400	0.9367	0.9265
	40	0.9812	0.9445	0.9430	0.9357
0.80	5	0.9810	0.9215	0.9160	0.8922
	10	0.9840	0.9360	0.9325	0.9163
	20	0.9848	0.9405	0.9400	0.9315
	40	0.9898	0.9427	0.9405	0.9325
0.85	5	0.9875	0.9153	0.9220	0.8935
	10	0.9915	0.9335	0.9310	0.9127
	20	0.9932	0.9415	0.9410	0.9293
	40	0.9920	0.9487	0.9445	0.9377
0.90	5	0.9910	0.9267	0.9305	0.9022
	10	0.9958	0.9413	0.9443	0.9237
	20	0.9955	0.9413	0.9445	0.9335
	40	0.9982	0.9537	0.9517	0.9440
0.95	5	0.9968	0.9215	0.9365	0.9030
	10	0.9992	0.9407	0.9447	0.9245
	20	0.9992	0.9473	0.9485	0.9350
	40	1.0000	0.9490	0.9493	0.9395
0.99	5	1.0000	0.9130	0.9377	0.9067
	10	0.9998	0.9360	0.9403	0.9235
	20	1.0000	0.9470	0.9455	0.9317
	40	1.0000	0.9435	0.9420	0.9335

Table 24: Observed confidence interval coverage of the 0.10 quantile when the nominal coverage equals 0.95, $J = 30$

		coverage			
ρ	n	incorrect	version 1	version 2	MLE
0.70	5	0.9758	0.8958	0.9012	0.8925
	10	0.9740	0.9203	0.9240	0.9180
	20	0.9768	0.9345	0.9337	0.9313
	40	0.9818	0.9425	0.9427	0.9413
0.75	5	0.9785	0.8938	0.9000	0.8925
	10	0.9830	0.9205	0.9240	0.9195
	20	0.9865	0.9337	0.9320	0.9293
	40	0.9830	0.9377	0.9387	0.9377
0.80	5	0.9832	0.9000	0.9103	0.9012
	10	0.9860	0.9157	0.9203	0.9133
	20	0.9905	0.9417	0.9450	0.9417
	40	0.9902	0.9417	0.9433	0.9405
0.85	5	0.9895	0.8930	0.9117	0.9038
	10	0.9908	0.9207	0.9315	0.9245
	20	0.9920	0.9317	0.9363	0.9310
	40	0.9965	0.9443	0.9445	0.9430
0.90	5	0.9968	0.8988	0.9217	0.9130
	10	0.9968	0.9273	0.9377	0.9337
	20	0.9985	0.9415	0.9440	0.9413
	40	0.9998	0.9503	0.9497	0.9477
0.95	5	0.9982	0.8932	0.9365	0.9245
	10	1.0000	0.9210	0.9430	0.9360
	20	1.0000	0.9323	0.9397	0.9355
	40	0.9995	0.9413	0.9397	0.9375
0.99	5	1.0000	0.8758	0.9427	0.9283
	10	1.0000	0.9085	0.9417	0.9343
	20	1.0000	0.9355	0.9493	0.9437
	40	1.0000	0.9400	0.9445	0.9417

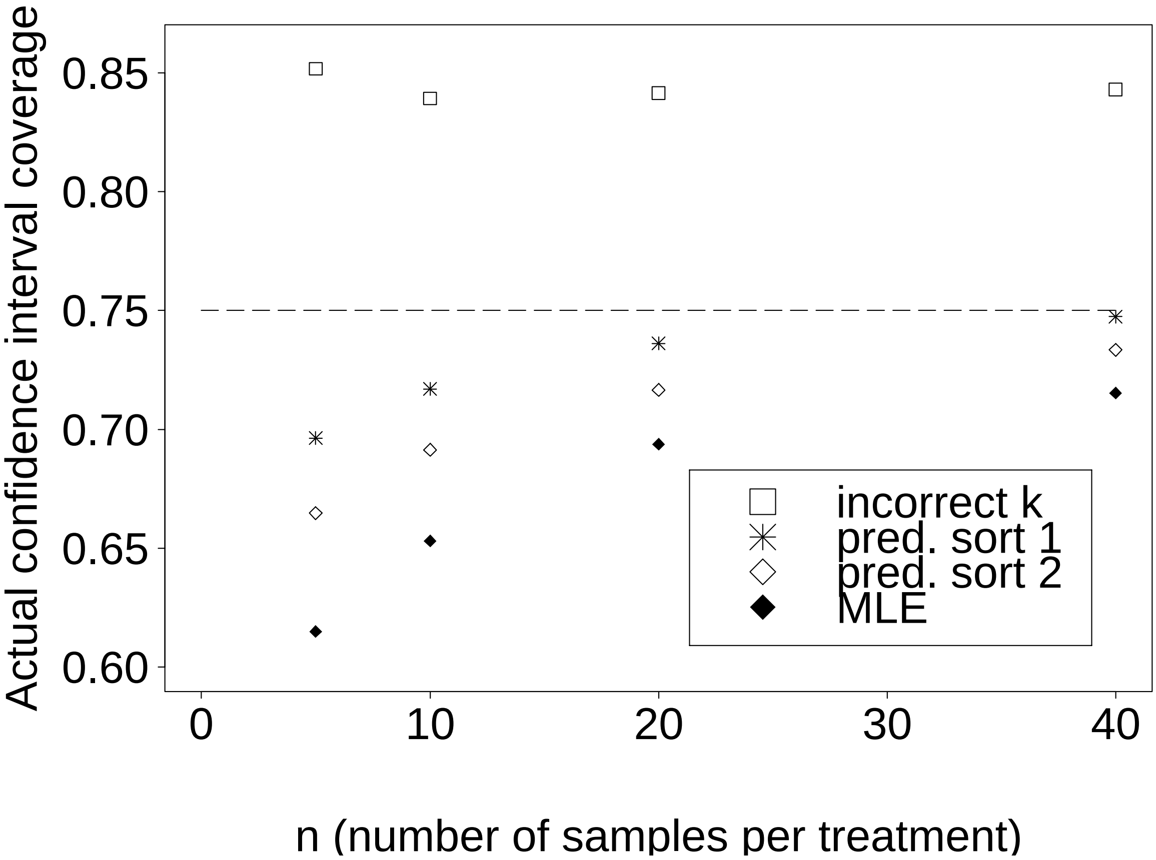


Figure 1: Approach of actual confidence interval coverage to nominal (.75) coverage as sample size increases. The correlation between the predictor and the response is 0.85. The number of treatments is four. The confidence interval is for the 0.01 quantile. (See Table 2.)

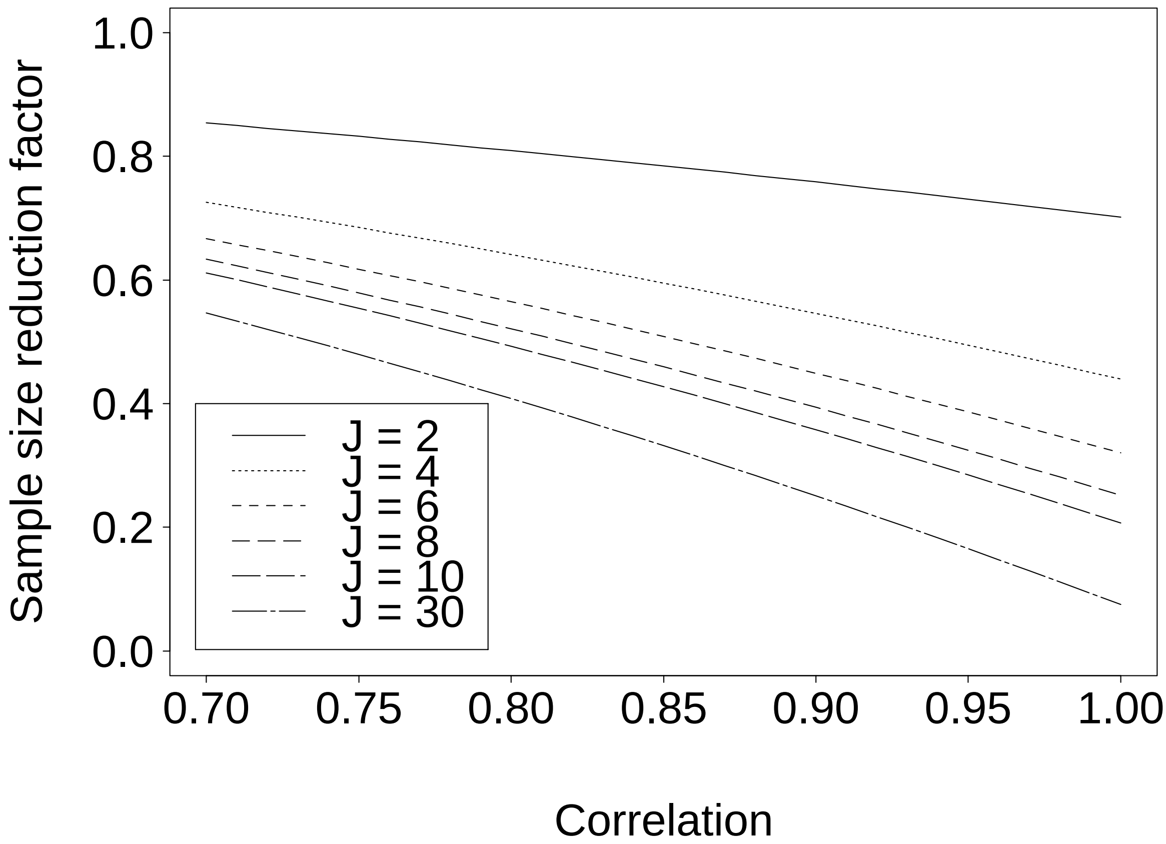


Figure 2: Sample size reduction factor as a function of the correlation between the predictor and the response, and the number of treatments J.

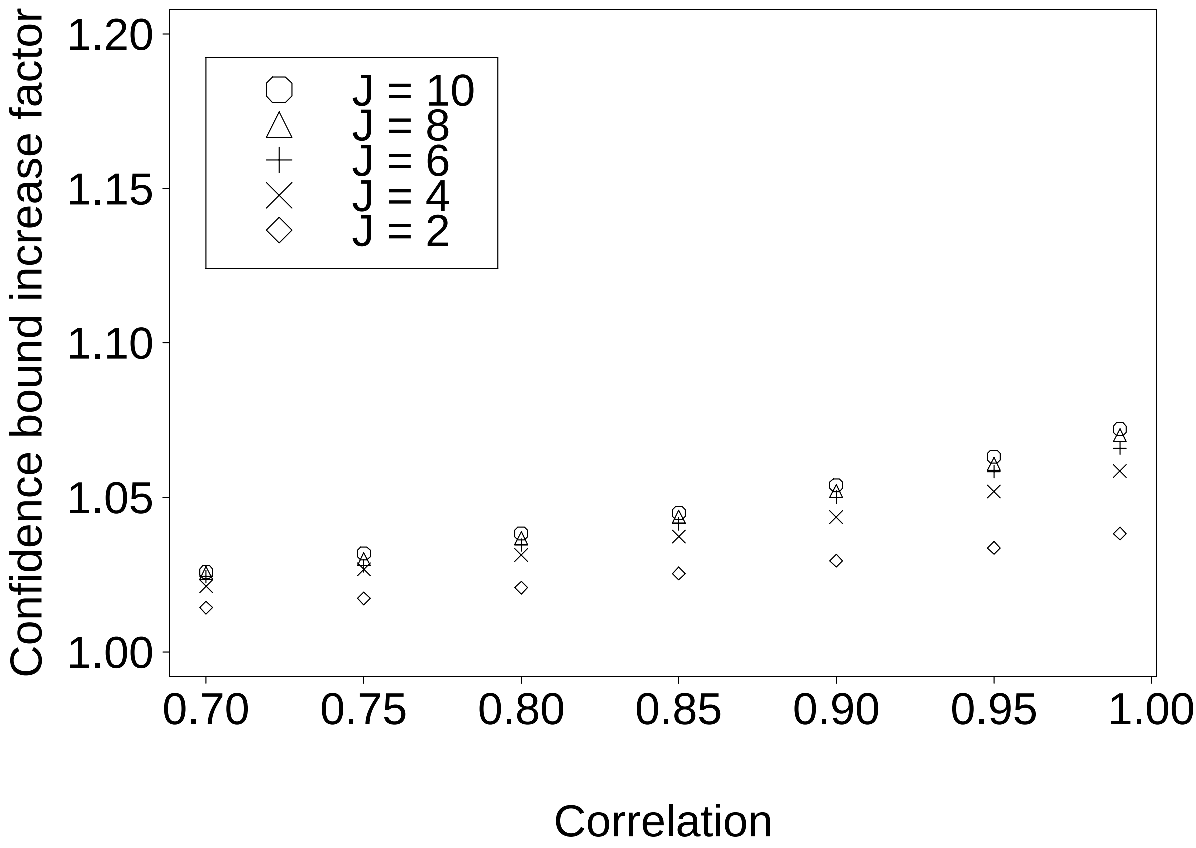


Figure 3: The ratio of the correct to the incorrect allowable property as a function of the correlation between the predictor and the response, and the number of treatments J. $n = 10$, $CV = 0.15$

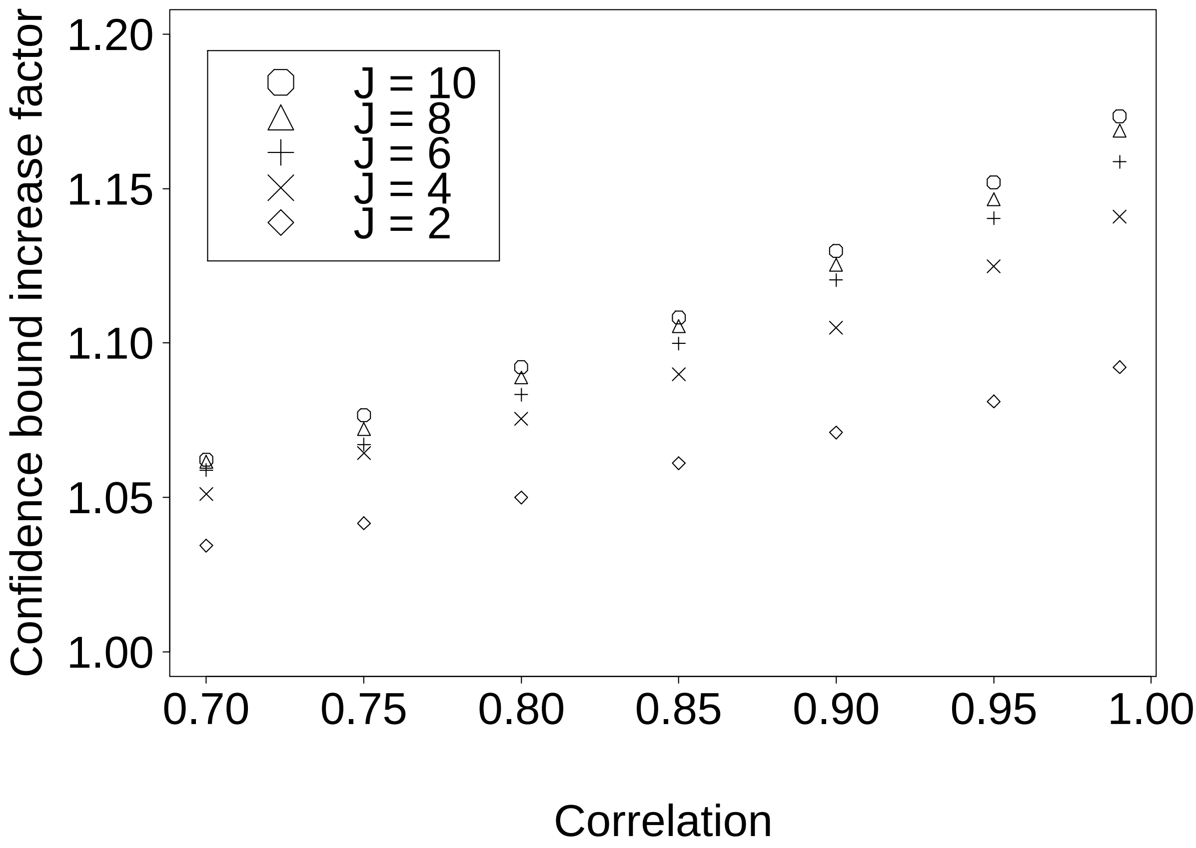


Figure 4: The ratio of the correct to the incorrect allowable property as a function of the correlation between the predictor and the response, and the number of treatments J. $n = 10$, $CV = 0.25$

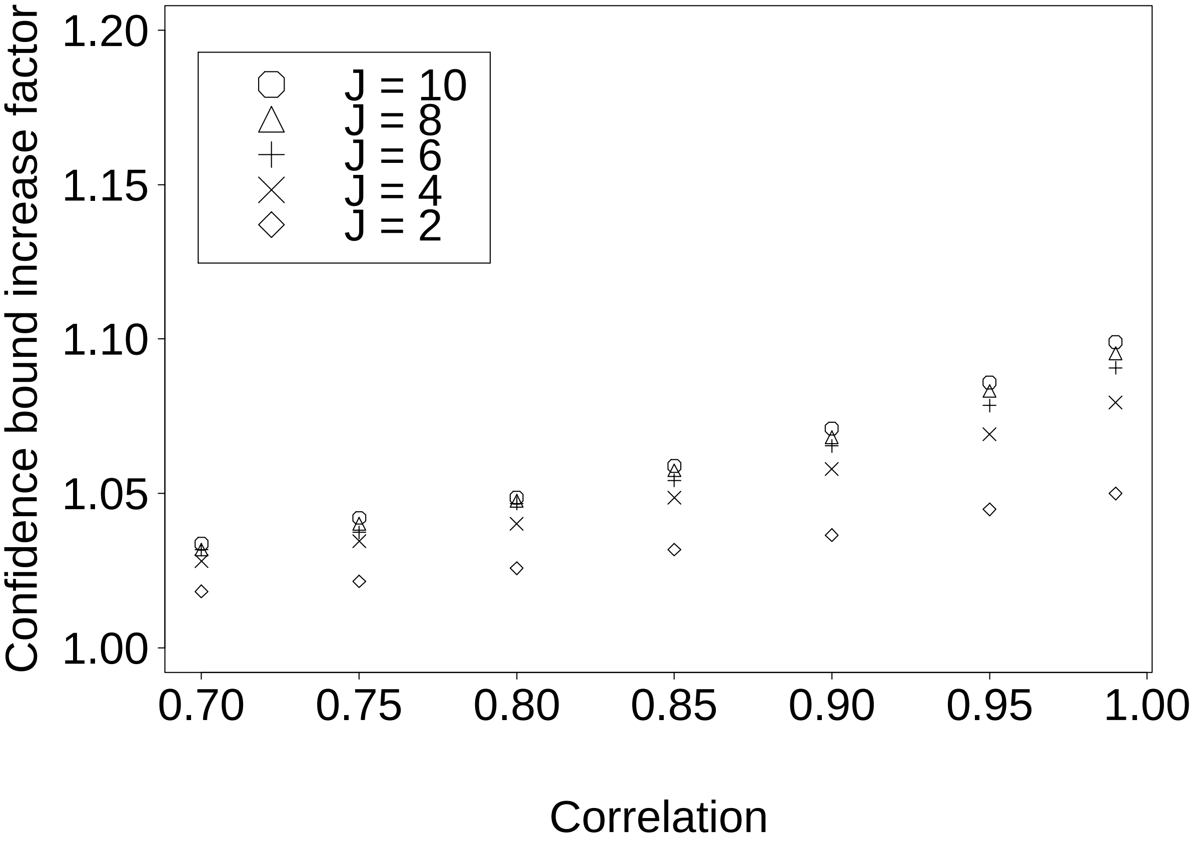


Figure 5: The ratio of the correct to the incorrect allowable property as a function of the correlation between the predictor and the response, and the number of treatments J. $n = 20$, $CV = 0.25$