ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Chernozhukov, Victor; Fernández-Val, Iván; Galichon, Alfred

Working Paper

Rearranging Edgeworth-Cornish-Fisher expansions

cemmap working paper, No. CWP19/07

Provided in Cooperation with:
The Institute for Fiscal Studies (IFS), London

Suggested Citation: Chernozhukov, Victor; Fernández-Val, Iván; Galichon, Alfred (2007) : Rearranging Edgeworth-Cornish-Fisher expansions, cemmap working paper, No. CWP19/07, Centre for Microdata Methods and Practice (cemmap), London, https://doi.org/10.1920/wp.cem.2007.1907

This Version is available at:
https://hdl.handle.net/10419/79274

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

cemmap
centre for microdata methods and practice

Rearranging Edgeworth-Cornish-Fisher Expansions

Victor Chernozhukov
Iván Fernández-Val
Alfred Galichon

The Institute for Fiscal Studies
Department of Economics, UCL

cemmap working paper CWP19/07

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL
An ESRC Research Centre

Institute for Fiscal Studies, 7 Ridgmount Street, London WC1E 7AE **Tel:** 44 (0)20 7291 4800 **Email:** info@cemmap.ac.uk **www.cemmap.ac.uk**

REARRANGING EDGEWORTH-CORNISH-FISHER EXPANSIONS

VICTOR CHERNOZHUKOV[†] IVÁN FERNÁNDEZ-VAL[§] ALFRED GALICHON[‡]

Abstract. This paper applies a regularization procedure called increasing rearrangement to monotonize Edgeworth and Cornish-Fisher expansions and any other related approximations of distribution and quantile functions of sample statistics. Besides satisfying the logical monotonicity, required of distribution and quantile functions, the procedure often delivers strikingly better approximations to the distribution and quantile functions of the sample mean than the original Edgeworth-Cornish-Fisher expansions.

Key Words: Edgeworth expansion, Cornish-Fisher expansion, rearrangement

---

Date: The results of this paper were first presented at the Statistics Seminar in Cornell University, October 2006. This version is of August 2007. We would like to thank Andrew Chesher, James Durbin, Ivar Ekeland, Xuming He, Joel Horowitz, Roger Koenker, Enno Mammen, Charles Manski, Ilya Molchanov, Francesca Molinari, and Peter Phillips for the helpful discussions and comments. We would also like to thank seminar participants at Cornell, University of Chicago, University of Illinois at Urbana-Champaign, Cowless Foundation 75th Anniversary Conference, Transportation Conference at the University of Columbia, and the CEMMAP "Measurement Matters" Conference for helpful comments.

[†] Massachusetts Institute of Technology, Department of Economics & Operations Research Center, and the University College of London, CEMMAP. E-mail: vchern@mit.edu. Research support from the Castle Krob Chair, National Science Foundation, the Sloan Foundation, and CEMMAP is gratefully acknowledged.

§ Boston University, Department of Economics. E-mail: ivanf@bu.edu.

‡ Harvard University, Department of Economics. E-mail: alfred.galichon@post.harvard.edu. Research support from the Conseil Général des Mines and the National Science Foundation is gratefully acknowledged.

1. Introduction

Higher order approximations to the distribution of sample statistics beyond the order $n^{-1/2}$ provided by the central limit theorem are of central interest in the theory of asymptotic statistics, see e.g. (Blinnikov and Moessner 1998, Cramér 1999, Bhattacharya and Ranga Rao 1976, Hall 1992, Rothenberg 1984, van der Vaart 1998). An important tool for performing these refinements is provided by the Edgeworth expansion (Edgeworth 1905, Edgeworth 1907), which approximates the distribution of the statistics of interest around the limit distribution (often the normal distribution) by a combination of Hermite polynomials, with coefficients defined in terms of moments. Inverting the expansion yields a related higher order approximation, the Cornish-Fisher expansion (Cornish and Fisher 1938, Fisher and Cornish 1960), to the quantiles of the statistic around the quantiles of the limiting distribution.

One of the important shortcomings of either Edgeworth or Cornish-Fisher expansions is that the resulting approximations to the distribution and quantile functions are not monotonic, which violates an obvious monotonicity requirement. This comes from the fact that the polynomials involved in the expansion are not monotone. Here we propose to use a procedure, called the rearrangement, to restore the monotonicity of the approximations and, perhaps more importantly, to improve the estimation properties of these approximations. The resulting improvement is due to the fact that the rearrangement necessarily brings the non-monotone approximations closer to to the true monotone target function.

The main findings of the paper can be illustrated through a single picture given as Figure 1. In that picture, we plot the true distribution function of a sample mean X based on a small sample, a third order Edgeworth approximation to that distribution, and a rearrangement of this third order approximation. We see that the Edgeworth approximation is sharply non-monotone and provides a rather poor approximation to the distribution function. The rearrangement merely sorts the value of the approximate

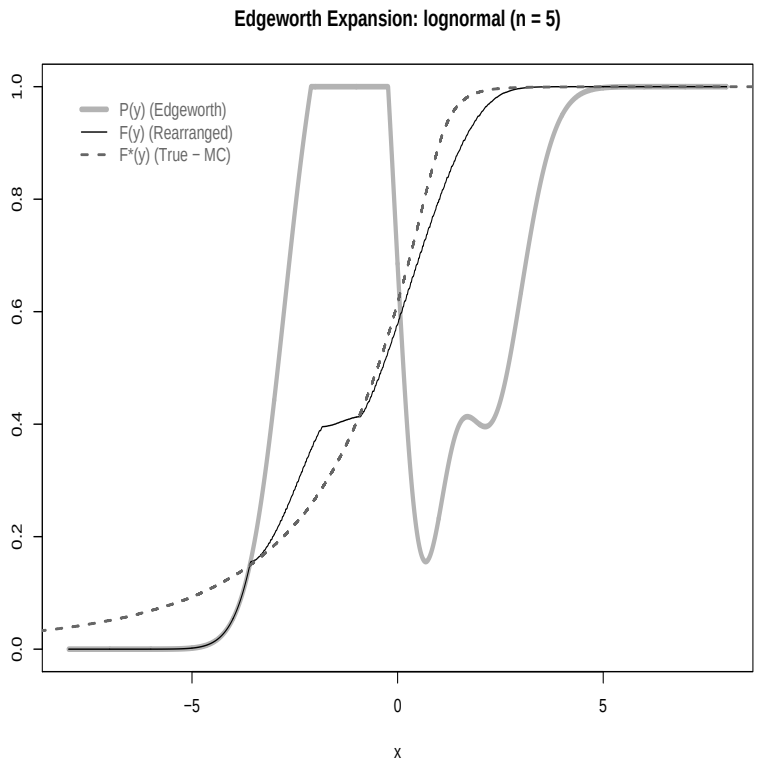

FIGURE 1. Distribution function for the standarized sample mean of log-normal random variables (sample size 5), the third order Edgeworth approximation, and the rearranged third order Edgeworth approximation.

distribution function in an increasing order. One can see that the rearranged approximation, in addition to being monotonic, is a much better approximation to the true function than the original approximation.

We organize the rest of the paper as follows. In Section 2, we describe the rearrangement and qualify the approximation property it provides for monotonic functions. In section 3, we introduce the rearranged Edgeworth-Cornish-Fisher expansions and explain how these produce better estimates of distributions and quantiles of statistics. In Section 4, we illustrate the procedure with several additional examples.

2. Improving Approximations of Monotone Functions by Rearrangement

In what follows, let $\mathcal{X}$ be a compact interval. We first consider an interval of the form $\mathcal{X} = [0, 1]$. Let $f(x)$ be a measurable function mapping $\mathcal{X}$ to K, a bounded subset of $\mathbb{R}$. Let $F_f(y) = \int_{\mathcal{X}} 1\{f(u) \leq y\}du$ denote the distribution function of $f(X)$ when X follows the uniform distribution on $[0, 1]$. Let

$$f^*(x) := Q_f(x) := \inf \{y \in \mathbb{R} : F_f(y) \geq x\}$$

be the quantile function of $F_f(y)$. Thus,

$$f^*(x) := \inf \left\{ y \in \mathbb{R} : \left[\int_{\mathcal{X}} 1\{f(u) \leq y\}du \right] \geq x \right\}.$$

This function f^* is called the increasing rearrangement of the function f. The rearrangement is a tool that is extensively used in functional analysis and optimal transportation (see e.g. Hardy, Littlewood, and Pólya (1952) and Villani (2003).) It originates in the work of Chebyshev, who used it to prove a set of inequalities (Bronshtein, Semendyayev, Musiol, Muehlig, and Mühlig 2004). Here we use this tool to improve approximations of monotone functions, such as the Edgeworth-Cornish-Fisher approximations to the distribution and quantile functions of the sample mean.

The rearrangement operator simply transforms a function f to its quantile function f^*. That is, $x \mapsto f^*(x)$ is the quantile function of the random variable $f(X)$ when $X \sim U(0, 1)$. Another convenient way to think of the rearrangement is as a sorting operation: Given values of the function $f(x)$ evaluated at x in a fine enough mesh of equidistant points, we simply sort the values in an increasing order. The function created in this way is the rearrangement of f.

Finally, if $\mathcal{X}$ is of the form $[a, b]$, let $\bar{x}(x) = (x-a)/(b-a) \in [0, 1]$, $x(\bar{x}) = a+(b-a)\bar{x} \in [a, b]$, and $\bar{f}^*$ be the rearrangement of the function $\bar{f}(\bar{x}) = f(x(\bar{x}))$ defined on $\bar{\mathcal{X}} = [0, 1]$. Then, the rearrangement of f is defined as

$$f^*(x) := \bar{f}^*(\bar{x}(x)).$$

The following result is detailed in Chernozhukov, Fernandez-Val, and Galichon (2006).

Proposition 1 (Improving Approximation of Monotone Functions)**.** *Let $f_0 : \mathcal{X} = [a, b] \to K$ be a weakly increasing measurable function in x, where K is a bounded subset of $\mathbb{R}$. This is the target function that we want to approximate. Let $\widehat{f} : \mathcal{X} \to K$ be another measurable function, an initial approximation or an estimate of the target function f_0.*

1. For any $p \in [1, \infty]$, the rearrangement of $\widehat{f}$, denoted $\widehat{f}^*$, weakly reduces the estimation error:

$$\left[\int_{\mathcal{X}} \left|\widehat{f}^*(x) - f_0(x)\right|^p dx\right]^{1/p} \leq \left[\int_{\mathcal{X}} \left|\widehat{f}(x) - f_0(x)\right|^p dx\right]^{1/p}. \tag{2.1}$$

2. Suppose that there exist regions $\mathcal{X}_0$ and $\mathcal{X}_0'$, each of measure greater than $\delta > 0$, such that for all $x \in \mathcal{X}_0$ and $x' \in \mathcal{X}_0'$ we have that (i) $x' > x$, (ii) $\widehat{f}(x) > \widehat{f}(x') + \epsilon$, and (iii) $f_0(x') > f_0(x) + \epsilon$, for some $\epsilon > 0$. Then the gain in the quality of approximation is strict for $p \in (1, \infty)$. Namely, for any $p \in [1, \infty]$,

$$\left[\int_{\mathcal{X}} \left|\widehat{f}^*(x) - f_0(x)\right|^p dx\right]^{1/p} \leq \left[\int_{\mathcal{X}} \left|\widehat{f}(x) - f_0(x)\right|^p dx - \delta_{\mathcal{X}}\eta_p\right]^{1/p}, \tag{2.2}$$

where $\eta_p = \inf\{|v - t'|^p + |v' - t|^p - |v - t|^p - |v' - t'|^p\}$ and $\eta_p > 0$ for $p \in (1, \infty)$, with the infimum taken over all v, v', t, t' in the set K such that $v' \geq v + \epsilon$ and $t' \geq t + \epsilon$; and $\delta_{\mathcal{X}} = \delta/(b - a)$.

Corollary 1 (Strict Improvement)**.** *If the target function f_0 is increasing over $\mathcal{X}$ and $\widehat{f}$ is decreasing over a subset of $\mathcal{X}$ that has positive measure, then the improvement in L_p norm, for $p \in (1, \infty)$, is necessarily strict.*

The first part of the proposition states the weak inequality (2.1), and the second part states the strict inequality (2.2). As an implication, the Corollary states that the inequality is strict for $p \in (1, \infty)$ if the original estimate $\widehat{f}(x)$ is decreasing on a subset of $\mathcal{X}$ having positive measure, while the target function $f_0(x)$ is increasing on $\mathcal{X}$ (by increasing, we mean strictly increasing throughout). Of course, if $f_0(x)$ is constant, then

the inequality (2.1) becomes an equality, as the distribution of the rearranged function $\widehat{f}^*$ is the same as the distribution of the original function $\widehat{f}$, that is $F_{\widehat{f}^*} = F_{\widehat{f}}$.

This proposition establishes that the rearranged estimate $\widehat{f}^*$ has a smaller estimation error in the L_p norm than the original estimate whenever the latter is not monotone. This is a very useful and generally applicable property that is independent of the sample size and of the way the original approximation $\widehat{f}$ to f_0 is obtained.

Remark 1. *An indirect proof of the weak inequality (2.1) is a simple but important consequence of the following classical inequality due to Lorentz (1953): Let q and g be two functions mapping $\mathcal{X}$ to K, a bounded subset of $\mathbb{R}$. Let q^* and g^* denote their corresponding increasing rearrangements. Then,*

$$\int_{\mathcal{X}} L(q^*(x), g^*(x), x)dx \leq \int_{\mathcal{X}} L(q(x), g(x), x)dx,$$

for any submodular discrepancy function $L : \mathbb{R}^3 \mapsto \mathbb{R}$. Set $q(x) = \widehat{f}(x)$, $q^*(x) = \widehat{f}^*(x)$, $g(x) = f_0(x)$, and $g^*(x) = f_0^*(x)$. Now, note that in our case $f_0^*(x) = f_0(x)$ almost everywhere, that is, the target function is its own rearrangement. Moreover, $L(v, w, x) = |w - v|^p$ is submodular for $p \in [1, \infty)$. This proves the first part of the proposition above. For $p = \infty$, the first part follows by taking the limit as $p \to \infty$. In the Appendix, for completeness, we restate the proof of Chernozhukov, Fernandez-Val, and Galichon (2006) of the strong inequality (2.2) as well as the direct proof of the weak inequality (2.1).

Remark 2. *The following immediate implication of the above finite-sample result is also worth emphasizing: The rearranged estimate $\widehat{f}^*$ inherits the L_p rates of convergence from the original estimates $\widehat{f}$. For $p \in [1, \infty]$, if $\lambda_n = [\int_{\mathcal{X}} |f_0(x) - \widehat{f}(x)|^p du]^{1/p} = O_P(a_n)$ for some sequence of constants a_n, then $[\int_{\mathcal{X}} |f_0(x) - \widehat{f}^*(x)|^p du]^{1/p} \leq \lambda_n = O_P(a_n)$. However, while the rate is the same, the error itself is smaller.*

Remark 3. *Finally, one can also consider weighted rearrangements that can accentuate the quality of approximation in various areas. Indeed, consider an absolutely continuous*

distribution function Λ *on* $\mathcal{X} = [a, b]$*, then we have that for* $(\widehat{f} \circ \Lambda^{-1})^*(u)$ *denoting the rearrangement of* $u \mapsto \widehat{f}(\Lambda^{-1}(u))$

$$\left[\int_{[0,1]} \left|(\widehat{f} \circ \Lambda^{-1})^*(u) - f_0(\Lambda^{-1}(u))\right|^p du\right]^{1/p} \leq \left[\int_{[0,1]} \left|\widehat{f}(\Lambda^{-1}(u)) - f_0(\Lambda^{-1}(u))\right|^p du\right]^{1/p}, \tag{2.3}$$

or, equivalently, by a change of variable, setting

$$\widehat{f}^*_\Lambda(x) = (\widehat{f} \circ \Lambda^{-1})^*(\Lambda(x))$$

we have that

$$\left[\int_{\mathcal{X}} \left|\widehat{f}^*_\Lambda(x) - f_0(x)\right|^p d\Lambda(x)\right]^{1/p} \leq \left[\int_{\mathcal{X}} \left|\widehat{f}(x) - f_0(x)\right|^p d\Lambda(x)\right]^{1/p}. \tag{2.4}$$

Thus, the function $x \mapsto \widehat{f}^*_\Lambda(x)$ *is the weighted rearrangement that provides improvements in the quality of approximation in the norm that is weighted according to the distribution function* Λ*.*

In the next section, we apply rearrangements to improve the Edgeworth-Cornish-Fisher and related approximations to distribution and quantile functions.

3. Improving Edgeworth-Cornish-Fisher and Related expansions

3.1. **Improving Quantile Approximations by Rearrangement.** We first consider the quantile case. Let Q_n be the quantile function of a statistic X_n, i.e.

$$Q_n(u) = \inf\{x \in \mathbb{R} : Pr[X_n \leq x] \geq u\},$$

which we assume to be strictly increasing. Let $\widehat{Q}_n$ be an approximation to the quantile function Q_n satisfying the following relation on an interval $\mathcal{U}_n = [\varepsilon_n, 1 - \varepsilon_n] \subseteq [0, 1]$:

$$Q_n(u) = \widehat{Q}_n(u) + \epsilon_n(u), \quad |\epsilon_n(u)| \leq a_n, \quad \text{for all } u \in \mathcal{U}_n. \tag{3.1}$$

The leading example of such an approximation is the inverse Edgeworth, or Cornish-Fisher, approximation to the quantile function of a sample mean. If X_n is is a standardized sample mean, $X_n = \sum_{i=1}^n (Y_i - E[Y_i])/\sqrt{Var(Y_i)}$ based on an i.i.d. sample

$(Y_1, ..., Y_n)$, then we have the following J-th order approximation

$$\begin{aligned} &Q_n(u) = \widehat{Q}_n(u) + \epsilon_n(u) \\ &\widehat{Q}_n(u) = R_1(\Phi^{-1}(u)) + R_2(\Phi^{-1}(u))/n^{1/2} + ... + R_J(\Phi^{-1}(u))/n^{(J-1)/2}, \\ &|\epsilon_n(u)| \leq Cn^{-J/2}, \quad \text{for all } u \in \mathcal{U}_n = [\varepsilon_n, 1-\varepsilon_n], \\ &\text{for some } \varepsilon_n \searrow 0 \text{ and } C > 0, \end{aligned} \tag{3.2}$$

provided that a set of regularity conditions, specified in, e.g., Zolotarev (1991), hold. Here Φ and Φ^{-1} denote the distribution function and quantile function of a standard normal random variable. The first three terms of the approximation are given by the polynomials,

$$\begin{aligned} R_1(z) &= z, \\ R_2(z) &= \lambda(z^2 - 1)/6, \\ R_3(z) &= (3\kappa(z^3 - 3z) - 2\lambda^2(2z^3 - 5z))/72, \end{aligned} \tag{3.3}$$

where λ is the skewness and κ is the kurtosis of the random variable Y. The Cornish-Fisher approximation is one of the central approximations of the asymptotic statistics. Unfortunately, inspection of the expression for polynomials (3.3) reveals that this approximation does not generally deliver a monotone estimate of the quantile function. This shortcoming has been pointed and discussed in detail e.g. by Hall (1992). The nature of the polynomials constructed is such that there always exists a large enough range $\mathcal{U}_n$ over which the Cornish-Fisher approximation is not monotone, cf. Hall (1992). As an example, in the case of the second order approximation ($J = 2$) we have that for $\lambda < 0$

$$\widehat{Q}_n(u) \searrow -\infty, \text{ as } u \nearrow 1, \tag{3.4}$$

that is, the Cornish-Fisher "quantile" function $\widehat{Q}_n$ is decreasing far enough in the tails. This example merely suggests a potential problem that may apply to practically relevant ranges of probability indices u. Indeed, specific numerical examples given below show that in small samples the non-monotonicity can occur in practically relevant ranges. Of

course, in sufficiently large samples, the regions of non-monotonicity are squeezed quite far into the tails.

Let Q_n^* be the rearrangement of $\widehat{Q}_n$. Then we have that for any $p \in [1, \infty]$, the rearranged quantile function reduces the approximation error of the original approximation:

$$\left[\int_{\mathcal{U}_n} \left|\widehat{Q}_n^*(u) - Q_n(u)\right|^p du\right]^{1/p} \leq \left[\int_{\mathcal{U}_n} \left|\widehat{Q}_n(u) - Q_n(u)\right|^p du\right]^{1/p} \leq (1 - 2\varepsilon_n)a_n, \quad (3.5)$$

with the first inequality holding strictly for $p \in (1, \infty)$ whenever $\widehat{Q}_n$ is decreasing on a region of $\mathcal{U}_n$ of positive measure. We can give the following probabilistic interpretation to this result. Under condition (3.1), there exists a variable $U = F_n(X_n)$ such that both the stochastic expansion

$$X_n = \widehat{Q}_n(U) + O_p(a_n), \quad (3.6)$$

and the expansion

$$X_n = \widehat{Q}_n^*(U) + O_p(a_n), \quad (3.7)$$

hold,[1] but the variable $\widehat{Q}_n^*(U)$ in (3.7) is a better coupling to the statistic X_n than $\widehat{Q}_n(U)$ in (3.6), in the following sense: For each $p \in [1, \infty]$,

$$[E1_n[X_n - \widehat{Q}_n^*(U)]^p]^{1/p} \leq [E1_n[X_n - \widehat{Q}_n(U)]^p]^{1/p}, \quad (3.8)$$

where $1_n = 1\{U \in \mathcal{U}_n\}$. Indeed, property (3.8) immediately follows from (3.5).

The above improvements apply in the context of the sample mean X_n. In this case, the probabilistic interpretation above is directly connected to the higher order central limit theorem of Zolotarev (1991), which states that under (3.2), we have the following higher-order probabilistic central limit theorem,

$$X_n = \widehat{Q}_n(U) + O_p(n^{-J/2}). \quad (3.9)$$

[1] $\widehat{Q}_n^*(U)$ is defined only on $\mathcal{U}_n$, so we can set $\widehat{Q}_n^*(U) = Q_n(U)$ outside $\mathcal{U}_n$, if needed. Of course, $U \notin \mathcal{U}_n$ with probability going to zero.

The term $\widehat{Q}_n(U)$ is Zolotarev's high-order refinement over the first order normal terms $\Phi^{-1}(U)$. Sun, Loader, and McCormick (2000) employ an analogous higher-order probabilistic central limit theorem to improve the construction of confidence intervals.

The application of the rearrangement to the Zolotarev's term in fact delivers a clear improvement in the sense that it also leads to a probabilistic higher order central limit theorem

$$X_n = \widehat{Q}_n^*(U) + O_p(n^{-J/2}), \tag{3.10}$$

but the leading term $\widehat{Q}_n^*(U)$ is closer to X_n than the Zolotarev's term $Q_n(U)$, in the sense of (3.8).

We summarize the above discussion into a formal proposition.

Proposition 2. *If expansion (3.1) holds, then the improvement (3.5) necessarily holds. The improvement is necessarily strict if $\widehat{Q}_n$ is decreasing over a region of $\mathcal{U}_n$ that has a positive measure. In particular, this improvement property applies to the inverse Edgeworth approximation defined in (3.2) to the quantile function of the sample mean.*

3.2. **Improving Distributional Approximations by Rearrangement.** We next consider distribution functions. Let $F_n(x)$ be the distribution function of a statistic X_n, and $\widehat{F}_n(x)$ be the approximation to this distribution such that the following relation holds:

$$F_n(x) = \widehat{F}_n(x) + \epsilon_n(x), \quad |\epsilon_n(x)| \leq a_n, \text{ for all } x \in \mathcal{X}_n, \tag{3.11}$$

where $\mathcal{X}_n = [-b_n, b_n]$ is an interval in $\mathbb{R}$ for some sequence of positive numbers b_n possibly growing to infinity.

The leading example of such an approximation is the Edgeworth expansion of the distribution function of a sample mean. If X_n is a standardized sample mean, $X_n = \sum_{i=1}^n (Y_i - E[Y_i])/\sqrt{Var(Y_i)}$ based on an i.i.d. sample $(Y_1, ..., Y_n)$, then we have the

following J-th order approximation

$$\begin{aligned} F_n(z) &= \widehat{F}_n(z) + \epsilon_n(z) \\ \widehat{F}_n(x) &= P_1(x) + P_2(x)/n^{1/2} + ... + P_J(x)/n^{(J-1)/2}, \\ |\epsilon_n(x)| &\leq Cn^{-J/2}, \quad \text{for all } x \in \mathcal{X}_n, \end{aligned} \tag{3.12}$$

for some $C > 0$, provided that a set of regularity conditions, specified in e.g. Hall (1992), hold. The first three terms of the approximation are given by

$$\begin{aligned} P_1(x) &= \Phi(x), \\ P_2(x) &= -\lambda(x^2 - 1)\phi(x)/6, \\ P_3(x) &= -(3\kappa(x^3 - 3x) + \lambda^2(x^5 - 10x^3 + 15x))\phi(x)/72, \end{aligned} \tag{3.13}$$

where Φ and ϕ denote the distribution function and density function of a standard normal random variable, and λ and κ are the skewness and kurtosis of the random variable Y. The Edgeworth approximation is one of the central approximations of the asymptotic statistics. Unfortunately, like the Cornish-Fisher expansion it generally does not provide a monotone estimate of the distribution function. This shortcoming has been pointed and discussed in detail by (Barton and Dennis 1952, Draper and Tierney 1972, Sargan 1976, Balitskaya and Zolotuhina 1988), among others.

Let F_n^* be the rearrangement of $\widehat{F}_n$. Then, we have that for any $p \in [1, \infty]$, the rearranged Edgeworth approximation reduces the approximation error of the original Edgeworth approximation:

$$\left[\int_{\mathcal{X}_n} \left|\widehat{F}_n^*(x) - F_n(x)\right|^p dx\right]^{1/p} \leq \left[\int_{\mathcal{X}_n} \left|\widehat{F}_n(x) - F_n(x)\right|^p dx\right]^{1/p} \leq 2b_n a_n, \tag{3.14}$$

with the first inequality holding strictly for $p \in (1, \infty)$ whenever $\widehat{F}_n$ is decreasing on a region of $\mathcal{X}_n$ of positive measure.

Proposition 3. *If expansion (3.11) holds, then the improvement (3.14) necessarily holds. The improvement is necessarily strict if $\widehat{F}_n$ is decreasing over a region of $\mathcal{X}_n$*

that has a positive measure. In particular, this improvement property applies to Edgeworth approximation defined in (3.12) to the distribution function of the sample mean.

3.3. **Weighted Rearrangement of Cornish-Fisher and Edgeworth Expansions.** In some cases, it could be worthwhile to weigh different areas of support differently than the Lebesgue (flat) weighting prescribes. For example, it might be desirable to rearrange $\widehat{F}_n$ using F_n as a weight measure. Indeed, using F_n as a weight, we obtain a better coupling to the P-value: $P = F_n(X_n)$ (in this quantity, X_n is drawn according to the true F_n). Using such weight will provide a probabilistic interpretation for the rearranged Edgeworth expansion, in analogy to the probabilistic interpretation for the rearranged Cornish-Fisher expansion. Since the weight is not available we may use the standard normal measure Φ as the weight measure instead. We may also construct an initial rearrangement with the Lebesgue weight, and use it as weight itself in a further weighted rearrangement (and even iterate in this fashion). Using non-Lebesgue weights may also be desirable when we want the improved approximations to weight the tails more heavily. Whatever the reason might be for further non-Lebesgue weighting, we have the following properties, which follow immediately in view of Remark 3.

Let Λ be a distribution function that admits a positive density with respect to the Lebesgue measure on the region $\mathcal{U}_n = [\varepsilon_n, 1-\varepsilon_n]$ for the quantile case and region $\mathcal{X}_n = [-b_n, b_n]$ for the distribution case. Then if (3.1) holds, the Λ-weighted rearrangement $\widehat{Q}^*_{n,\Lambda}$ of the function $\widehat{Q}$ satisfies

$$\left[\int_{\mathcal{U}_n} \left|\widehat{Q}^*_{n,\Lambda}(u) - Q_n(u)\right|^p d\Lambda(u)\right]^{1/p} \leq \left[\int_{\mathcal{U}_n} \left|\widehat{Q}_n(u) - Q_n(u)\right|^p d\Lambda(u)\right]^{1/p} \tag{3.15}$$

$$\leq (\Lambda[1-\varepsilon_n] - \Lambda[\varepsilon_n])a_n, \tag{3.16}$$

where the first equality holds strictly when $\widehat{Q}$ is decreasing on a subset of positive Λ-measure. Furthermore, if (3.11) holds, then the Λ-weighted rearrangement $\widehat{F}^*_{n,\Lambda}$ of the

function $\widehat{F}$ satisfies

$$\left[\int_{\mathcal{X}_n}\left|\widehat{F}^*_{n,\Lambda}(x)-F_n(x)\right|^p d\Lambda(x)\right]^{1/p} \leq \left[\int_{\mathcal{X}_n}\left|\widehat{F}_n(x)-F_n(x)\right|^p d\Lambda(x)\right]^{1/p} \quad (3.17)$$

$$\leq (\Lambda[b_n]-\Lambda[-b_n])a_n. \quad (3.18)$$

4. Numerical Examples

In addition to the lognormal example given in the introduction, we use the gamma distribution to illustrate the improvements that the rearrangement provides. Let $(Y_1, ..., Y_n)$ be an i.i.d. sequence of Gamma(1/16,16) random variables. The statistic of interest is the standardized sample mean $X_n = \sum_{i=1}^n (Y_i - E[Y_i])/\sqrt{Var(Y_i)}$. We consider samples of sizes $n = 4, 8, 16$, and 32. In this example, the distribution function F_n and quantile function Q_n of the sample mean X_n are available in a closed form, making it easy to compare them to the Edgeworth approximation $\widehat{F}_n$ and the Cornish-Fisher approximation $\widehat{Q}_n$, as well as to the rearranged Edgeworth approximation $\widehat{F}^*_n$ and the the rearranged Cornish-Fisher approximation $\widehat{Q}^*_n$. For the Edgeworth and Cornish-Fisher approximations, as defined in the previous section, we consider the third order expansions, that is we set $J = 3$.

Figure 2 compares the true distribution function F_n, the Edgeworth approximation $\widehat{F}_n$, and the rearranged Edgeworth approximation $\widehat{F}^*_n$. We see that the rearranged Edgeworth approximation not only fixes the monotonicity problem, but also consistently does a better job at approximating the true distribution than the Edgeworth approximation. Table 1 further supports this point by presenting the numerical results for the L_p approximation errors, calculated according to the formulas given in the previous section. We see that the rearrangement reduces the approximation error quite substantially in most cases.

Figure 3 compares the true quantile function Q_n, the Cornish-Fisher approximation $\widehat{Q}_n$, and the rearranged Cornish-Fisher approximation $\widehat{Q}^*_n$. Here too we see that the

rearrangement not only fixes the non-monotonicity problem, but also brings the approximation closer to the truth. Table 2 further supports this point numerically, showing that the rearrangement reduces the L_p approximation error quite substantially in most cases.

5. Conclusion

In this paper, we have applied the rearrangement procedure to monotonize Edgeworth and Cornish-Fisher expansions and other related expansions of distribution and quantile functions. The benefits of doing so are twofold. First, we have obtained estimates of the distribution and quantile curves of the statistics of interest which satisfy the logical monotonicity restriction, unlike those directly given by the truncation of the series expansions. Second, we have shown that doing so resulted in better approximation properties.

Appendix A. Proof of Proposition 1

We consider the case where $\mathcal{X} = [0, 1]$ only, as the more general intervals can be dealt similarly. The first part establishes the weak inequality, following in part the strategy in Lorentz's (1953) proof. The proof focuses directly on obtaining the result stated in the proposition. The second part establishes the strong inequality.

Proof of Part 1. We assume at first that the functions $\widehat{f}(\cdot)$ and $f_0(\cdot)$ are simple functions, constant on intervals $((s-1)/r, s/r]$, $s = 1, ..., r$. For any simple $f(\cdot)$ with r steps, let f denote the r-vector with the s-th element, denoted f_s, equal to the value of $f(\cdot)$ on the s-th interval. Let us define the sorting operator $S(f)$ as follows: Let ℓ be an integer in $1, ..., r$ such that $f_\ell > f_m$ for some $m > l$. If ℓ does not exist, set $S(f) = f$. If ℓ exists, set $S(f)$ to be a r-vector with the ℓ-th element equal to f_m, the m-th element equal to f_ℓ, and all other elements equal to the corresponding elements of f. For any submodular function $L : \mathbb{R}^2 \to \mathbb{R}_+$, by $f_\ell \geq f_m$, $f_{0m} \geq f_{0\ell}$ and the definition of the

submodularity,

$$L(f_m, f_{0\ell}) + L(f_\ell, f_{0m}) \leq L(f_\ell, f_{0\ell}) + L(f_m, f_{0m}).$$

Therefore, we conclude that

$$\int_{\mathcal{X}} L(S(\widehat{f})(x), f_0(x))dx \leq \int_{\mathcal{X}} L(\widehat{f}(x), f_0(x))dx, \tag{A.1}$$

using that we integrate simple functions.

Applying the sorting operation a sufficient finite number of times to $\widehat{f}$, we obtain a completely sorted, that is, rearranged, vector $\widehat{f}^*$. Thus, we can express $\widehat{f}^*$ as a finite composition $\widehat{f}^* = S \circ ... \circ S(\widehat{f})$. By repeating the argument above, each composition weakly reduces the approximation error. Therefore,

$$\int_{\mathcal{X}} L(\widehat{f}^*(x), f_0(x))dx \leq \int_{\mathcal{X}} L(\underbrace{S \circ ... \circ S}_{\text{finite times}}(\widehat{f}), f_0(x))dx \leq \int_{\mathcal{X}} L(\widehat{f}(x), f_0(x))dx. \tag{A.2}$$

Furthermore, this inequality is extended to general measurable functions $\widehat{f}(\cdot)$ and $f_0(\cdot)$ mapping $\mathcal{X}$ to K by taking a sequence of bounded simple functions $\widehat{f}^{(r)}(\cdot)$ and $f_0^{(r)}(\cdot)$ converging to $\widehat{f}(\cdot)$ and $f_0(\cdot)$ almost everywhere as $r \to \infty$. The almost everywhere convergence of $\widehat{f}^{(r)}(\cdot)$ to $\widehat{f}(\cdot)$ implies the almost everywhere convergence of its quantile function $\widehat{f}^{*(r)}(\cdot)$ to the quantile function of the limit, $\widehat{f}^*(\cdot)$. Since inequality (A.2) holds along the sequence, the dominated convergence theorem implies that (A.2) also holds for the general case. □

Proof of Part 2. Let us first consider the case of simple functions, as defined in Part 1. We take the functions to satisfy the following hypotheses: there exist regions $\mathcal{X}_0$ and $\mathcal{X}_0'$, each of measure greater than $\delta > 0$, such that for all $x \in \mathcal{X}_0$ and $x' \in \mathcal{X}_0'$, we have that (i) $x' > x$, (ii) $\widehat{f}(x) > \widehat{f}(x') + \epsilon$, and (iii) $f_0(x') > f_0(x) + \epsilon$, for $\epsilon > 0$ specified in the proposition. For any strictly submodular function $L : \mathbb{R}^2 \to \mathbb{R}_+$ we have that

$$\eta = \inf\{L(v', t) + L(v, t') - L(v, t) - L(v', t')\} > 0,$$

where the infimum is taken over all v, v', t, t' in the set K such that $v' \geq v + \epsilon$ and $t' \geq t + \epsilon$.

We can begin sorting by exchanging an element $\widehat{f}(x)$, $x \in \mathcal{X}_0$, of r-vector $\widehat{f}$ with an element $\widehat{f}(x')$, $x' \in \mathcal{X}_0'$, of r-vector $\widehat{f}$. This induces a sorting gain of at least η times $1/r$. The total mass of points that can be sorted in this way is at least δ. We then proceed to sort all of these points in this way, and then continue with the sorting of other points. After the sorting is completed, the total gain from sorting is at least $\delta\eta$. That is,

$$\int_{\mathcal{X}} L(\widehat{f}^*(x), f_0(x))dx \leq \int_{\mathcal{X}} L(\widehat{f}(x), f_0(x))dx - \delta\eta.$$

We then extend this inequality to the general measurable functions exactly as in the proof of part one. □

References

Balitskaya, E. O., and L. A. Zolotuhina (1988): "On the representation of a density by an Edgeworth series," *Biometrika*, 75(1), 185–187.

Barton, D. E., and K. E. Dennis (1952): "The conditions under which Gram-Charlier and Edgeworth curves are positive definite and unimodal," *Biometrika*, 39, 425–427.

Bhattacharya, R. N., and R. Ranga Rao (1976): *Normal approximation and asymptotic expansions*. John Wiley & Sons, New York-London-Sydney, Wiley Series in Probability and Mathematical Statistics.

Blinnikov, S., and R. Moessner (1998): "Expansions for nearly Gaussian distributions," *Astron. Astrophys. Suppl. Ser.*, 130, 193–205.

Bronshtein, I. N., K. A. Semendyayev, G. Musiol, H. Muehlig, and H. Mühlig (2004): *Handbook of mathematics*. Springer-Verlag, Berlin, english edn.

Chernozhukov, V., I. Fernandez-Val, and A. Galichon (2006): "Improving Approximations of Monotone Functions by Rearrangement," MIT and UCL CEMMAP Working Paper, available at www.arxiv.org, cemmap.ifs.org.uk, www.ssrn.com.

Cornish, E. A., and S. R. A. Fisher (1938): "Moments and Cumulants in the Specification of Distributions," *Review of the International Statistical Institute*, 5, 307–320.

Cramér, H. (1999): *Mathematical methods of statistics*, Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, Reprint of the 1946 original.

Draper, N. R., and D. E. Tierney (1972): "Regions of positive and unimodal series expansion of the Edgeworth and Gram-Charlier approximations," *Biometrika*, 59, 463–465.

Edgeworth, F. Y. (1905): "The law of error," *Trans. Camb. Phil. Society*, 20, 36–65, 113–41.

——— (1907): "On the representatin of a statistical frequency by a series," *J. R. Stat. Soc.*, 70, 102–6.

Fisher, S. R. A., and E. A. Cornish (1960): "The percentile points of distributions having known cumulants," *Technometrics*, 2, 209–225.

Hall, P. (1992): *The bootstrap and Edgeworth expansion*, Springer Series in Statistics. Springer-Verlag, New York.

Hardy, G. H., J. E. Littlewood, and G. Pólya (1952): *Inequalities*. Cambridge University Press, 2d ed.

Lorentz, G. G. (1953): "An inequality for rearrangements," *Amer. Math. Monthly*, 60, 176–179.

Rothenberg, T. J. (1984): "Approximating the Distributions of Econometric Estimators and Test Statistics," in *Handbook of econometrics, Vol. II*, vol. 2. North-Holland, Amsterdam, edited by Zvi Griliches and Michael D. Intriligator.

Sargan, J. D. (1976): "Econometric estimators and the Edgeworth approximation," *Econometrica*, 44(3), 421–448.

Sun, J., C. Loader, and W. P. McCormick (2000): "Confidence bands in generalized linear models," *Ann. Statist.*, 28(2), 429–460.

van der Vaart, A. W. (1998): *Asymptotic statistics*. Cambridge University Press, Cambridge.

Villani, C. (2003): *Topics in optimal transportation*, vol. 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI.

Zolotarev, V. M. (1991): "Asymptotic expansions with probability 1 of sums of independent random variables," *Teor. Veroyatnost. i Primenen.*, 36(4), 770–772.

TABLE 1. Estimation errors for approximations to the distribution function of the standarized sample mean from a Gamma(1/16, 16) population.

	First Order	Third Order	TO - Rearranged	Ratio (TO/RTO)
			A. $n = 4$	
L_1	0.07	0.05	0.02	0.38
L_2	0.10	0.06	0.03	0.45
L_3	0.13	0.07	0.04	0.62
L_4	0.15	0.08	0.07	0.81
L_∞	0.30	0.48	0.48	1.00
			B. $n = 8$	
L_1	0.06	0.03	0.01	0.45
L_2	0.08	0.04	0.02	0.63
L_3	0.10	0.05	0.04	0.85
L_4	0.11	0.06	0.06	0.96
L_∞	0.23	0.28	0.28	1.00
			C. $n = 16$	
L_1	0.05	0.01	0.01	0.97
L_2	0.06	0.02	0.02	0.99
L_3	0.08	0.03	0.03	1.00
L_4	0.08	0.04	0.04	1.00
L_∞	0.15	0.11	0.11	1.00
			D. $n = 32$	
L_1	0.04	0.01	0.01	1.00
L_2	0.05	0.01	0.01	1.00
L_3	0.06	0.01	0.01	1.00
L_4	0.06	0.01	0.01	1.00
L_∞	0.09	0.03	0.03	1.00

TABLE 2. Estimation errors for approximations to the distribution function of the standarized sample mean from a Gamma(1/16, 16) population.

	First Order	Third Order	TO - Rearranged	Ratio (TO/RTO)
			A. $n = 4$	
L_1	0.50	0.24	0.09	0.39
L_2	0.59	0.32	0.11	0.35
L_3	0.69	0.42	0.13	0.31
L_4	0.78	0.52	0.15	0.29
L_∞	2.04	1.53	0.49	0.32
			B. $n = 8$	
L_1	0.39	0.08	0.03	0.37
L_2	0.47	0.11	0.04	0.35
L_3	0.56	0.16	0.05	0.32
L_4	0.65	0.21	0.06	0.30
L_∞	1.66	0.67	0.22	0.33
			C. $n = 16$	
L_1	0.28	0.02	0.02	0.97
L_2	0.35	0.04	0.03	0.84
L_3	0.43	0.05	0.04	0.71
L_4	0.51	0.07	0.04	0.63
L_∞	1.34	0.24	0.10	0.44
			D. $n = 32$	
L_1	0.20	0.01	0.01	1.00
L_2	0.26	0.01	0.01	1.00
L_3	0.31	0.02	0.02	1.00
L_4	0.37	0.02	0.02	1.00
L_∞	1.02	0.07	0.07	1.00

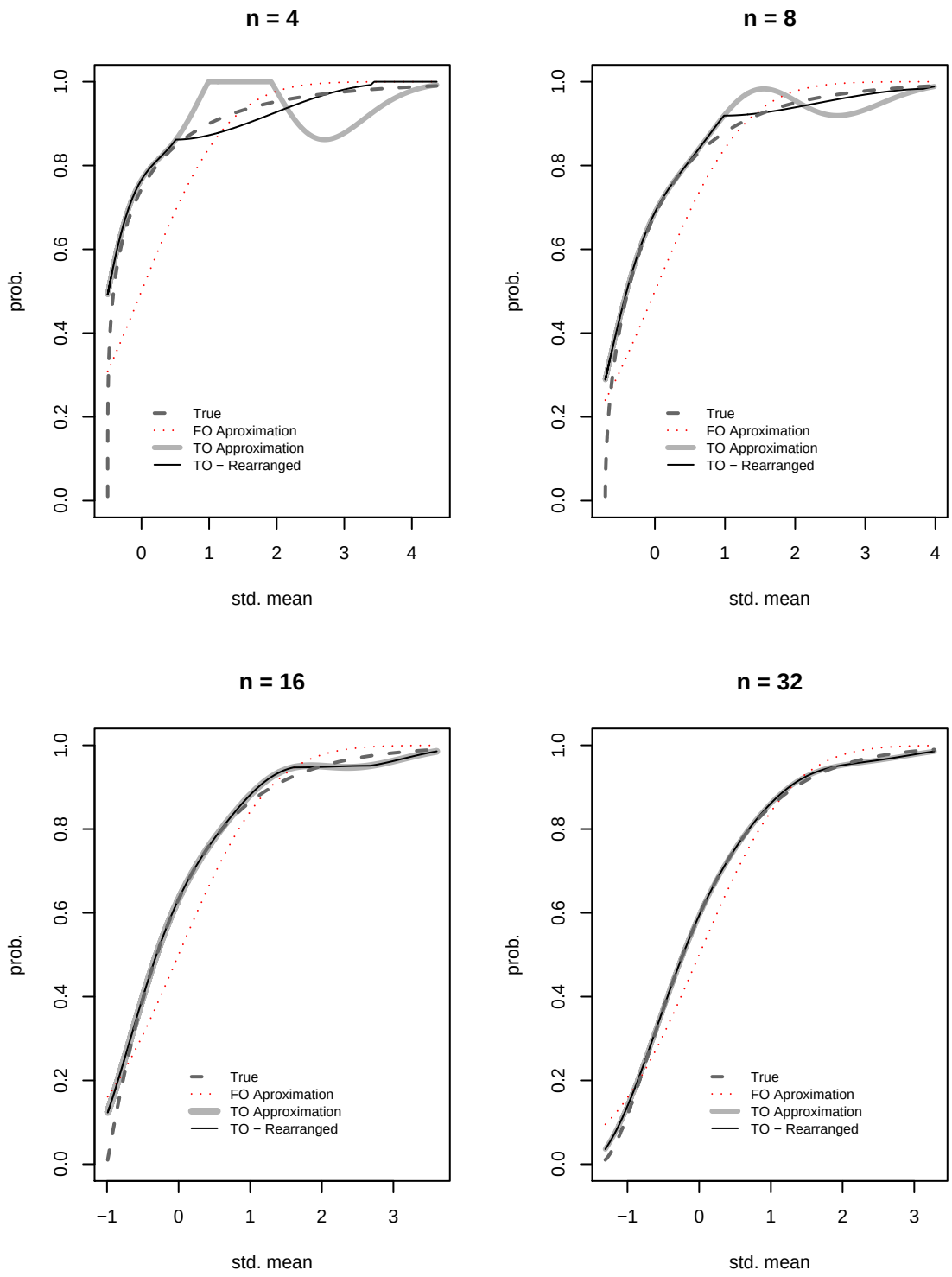

FIGURE 2. Distribution Functions, First Order Approximations, Third Order Approximations, and Rearrangements for the standarized sample mean from a Gamma(1/16, 16) population.

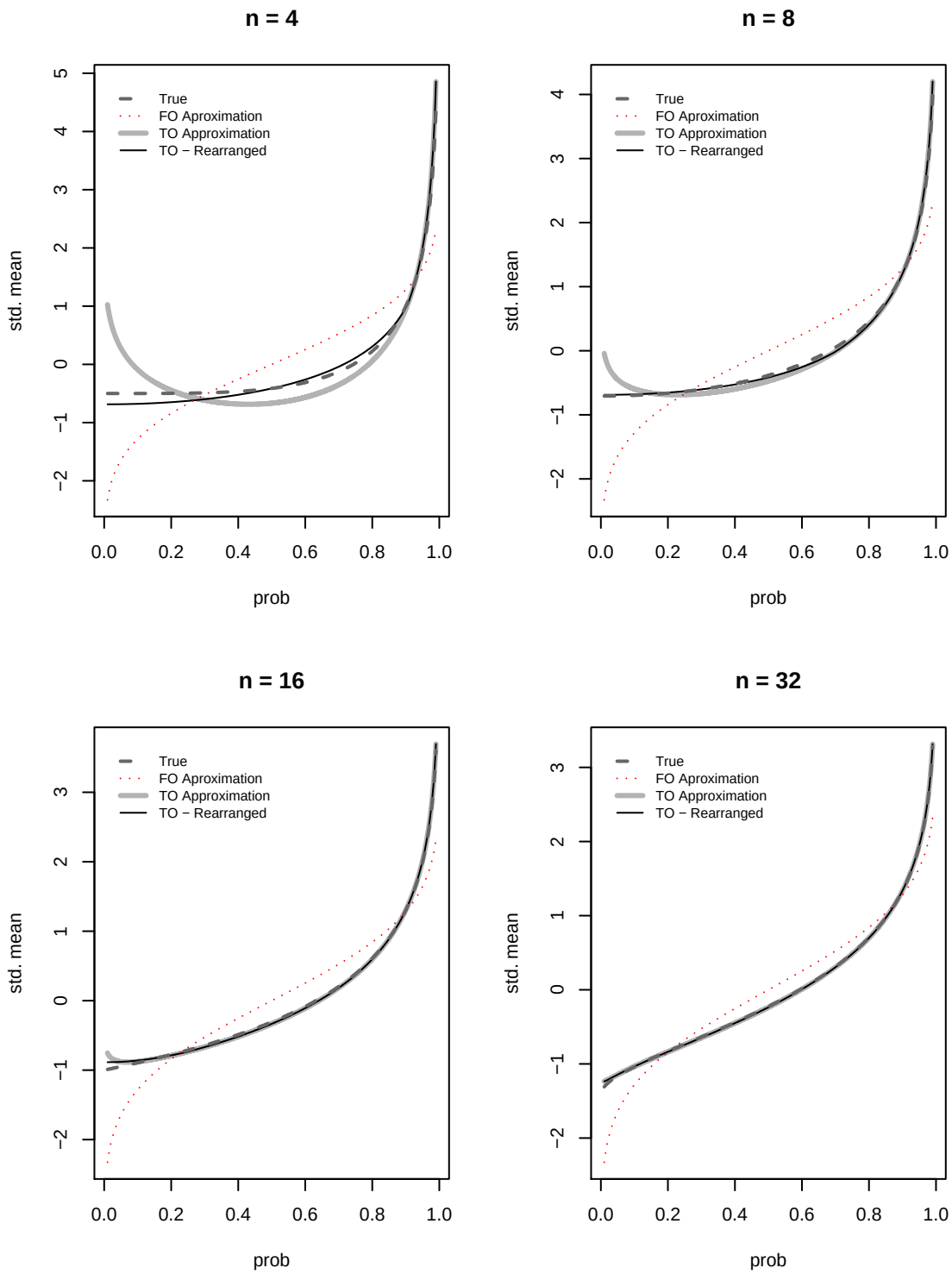

FIGURE 3. Quantile Functions, First Order Approximations, Third Order Approximations, and Rearrangements for the standarized sample mean from a Gamma(1/16, 16) population.