Sharp Nonasymptotic Bounds and Three Term Asymptotic Expansions for the Mean and Median of a Gaussian Sample Maximum

Anirban DasGupta and Jordan M. Stoyanov

June 15, 2012

Abstract For an iid standard normal sample of size n, the article first provides bounds on the expectation and the median of the sample maximum $X_{(n)}$ that are valid for all fixed $n \geq n_0$ for a small n_0, for example $n_0 = 7$. These fixed n bounds on the expectation and the median are derived by first deriving bounds on the standard normal quantile function $\Phi^{-1}(1-p)$ itself. The bounds resemble terms in the Edgeworth expansion for the CDF of a sample mean, in an appropriate sense. They are the currently best available explicit nonasymptotic bounds, and are of the correct asymptotic order up to the number of terms present in the given bounds.

Next, the article gives exact three term asymptotic expansions for the expectation and median of $X_{(n)}$. This is done by a technique of reducing the extreme value problem to a problem about sample means. This technique is general, and should apply to other distributions.

A potential application of a subset of our results is made to the Donoho-Johnstone estimate for the problem of detecting sparse weak signals.

1 Introduction

Let $X_{(n)}$ denote the maximum of an iid sample $X_1, \cdots, X_n$ from a univariate normal distribution, say, the standard normal. Distributional properties of $X_{(n)}$, such as the expectation, have become extraordinarily important in several frontier areas in theory as well as applications. A few instances are the widespread use of properties of $X_{(n)}$ for variable selection in sparse high dimensional regression, in studying the hard and soft thresholding estimators of Donoho and Johnstone (Donoho and Johnstone (1994)) for sparse signal detection and false discovery, and in analyzing or planning for extreme events of a diverse nature in practical enterprises, such as climate studies, finance, and hydrology.

We do know quite a bit about distributional properties of $X_{(n)}$ already. For example, we know the asymptotic distribution on suitable centering and norming (e.g., Galambos (1978)). We know that to the first order, the expectation, median, and the mode of $X_{(n)}$ are all asymptotic to $\sqrt{2 \log n}$, and that convergence to the asymptotic distribution is very slow (Hall (1979)). There is also very original and useful nonasymptotic work in Lai and Robbins (1976) on the expectation of $X_{(n)}$. The Lai-Robbins bounds were generalized to other order statistics, including the maximum, in Gascuel and Caraux (1992) and Rychlik (1998), who considers general L-estimates. In this article, we first provide the currently best available nonasymptotic bounds on the mean and the median of $X_{(n)}$. The bounds would remind one of Edgeworth expansions in the world of sample means; but they are valid for fixed n. As an example, Theorem 4.1 in this article shows that for $n \geq 7$,

$$E(X_{(n)}) \leq \sqrt{2 \log n} - \frac{\log 4\pi + \log \log n - 2}{2\sqrt{2 \log n}} + \frac{k(\log 4\pi + \log \log n)}{(2 \log n)^{3/2}},$$

where the constant k is explicit and can be taken to be 1.5 (or anything bigger). If simpler nonasymptotic bounds, although numerically less accurate, were desired, they could be easily extracted out of the above bound. For example, it follows from the above bound that for all $n \geq 10$,

$$E(X_{(n)}) \leq \sqrt{2 \log n} - \frac{\log 4\pi + \log \log n - 6}{6\sqrt{2 \log n}}.$$

While valid for fixed n, the successive terms in the above bound go down at increasing powers of $\frac{1}{\sqrt{2 \log n}}$, just as successive terms in an Edgeworth expansion for the

CDF of a sample mean go down at increasing powers of $\frac{1}{\sqrt{n}}$. However, the greatest utility of these bounds is that they are nonasymptotic, the sharpest ones available to date, and cover both the mean and the median of $X_{(n)}$. Bounds on the median are stated in Section 3, and the bounds on the mean in Section 4.

On the technical side, the principal ingredients for the bounds on the mean and the median are the Mitrinovic-Szarek-Werner type inequalities of the form $1 - \Phi(x) \leq \frac{a\phi(x)}{bx+\sqrt{c^2x^2+d}}, x > 0$, for suitable constants a, b, c, d, the Lai-Robbins inequality $E(X_{(n)}) \leq F^{-1}(1 - \frac{1}{n}) + n\int_{F^{-1}(1-\frac{1}{n})}^{\infty}[1 - F(x)]dx$, and various other analytic inequalities, e.g., Jensen. A principal strategy is to first analytically bound the standard normal quantile function $z_p = \Phi^{-1}(1 - p)$ itself, and then transform those to bounds on the mean and the median of $X_{(n)}$. For example, it is proved in this article (Corollary 2.1) that for $p \leq 0.1$,

$$z_p \leq \sqrt{2\log t} - \frac{\log 4\pi + \log\log t}{2\sqrt{2\log t}}(1 - \frac{k}{\log t}),$$

where $t = \frac{1}{p}$ and k can be taken to be 0.6403. Once again, if one wishes, simpler bounds can be extracted out of the above bound, at the cost of some loss in accuracy. In fact, the bounds on z_p are bi-directional, and may quite possibly be of some independent utility. The bounds on z_p are stated in Section 2.

In Section 5, we move on to deriving higher order asymptotic expansions for the mean and the median of $X_{(n)}$. This is achieved by first accomplishing a three term asymptotic expansion for z_p as $p \to 0$ (equation (42)). We then use the celebrated Réyni representation to reduce the problem about $X_{(n)}$ to a problem about sample means of an iid standard exponential sequence. Integral representation (39) is the key, and it is applicable in general, not the normal case alone. Very careful collection of terms is then needed to produce the ultimate three term asymptotic expansions for the mean and the median of $X_{(n)}$ up to an error of order $O(\frac{(\log\log n)^2}{(\log n)^{5/2}})$ (Theorem 5.1, 5.2). For example, Theorem 5.2 gives the result that the median of $X_{(n)}$ admits the three term asymptotic expansion

$$\text{med}(X_{(n)}) = \sqrt{2\log n} - \frac{\log\log n + \log 4\pi + 2\log\log 2}{2\sqrt{2\log n}}$$

$$+\frac{1}{8(2\log n)^{3/2}} \times \Big[(\log\log n - 2(1+\log\log 2))^2 + (\log 4\pi - 2)^2 - 4\log\log 2(\log 4\pi + 2\log\log 2)\Big]$$

$$+O(\frac{(\log\log n)^2}{(\log n)^{5/2}}).$$

Apart from the difficulty of the intermediate calculation, the asymptotic expansions can be practically useful when n is truly fantastically large and therefore the mean or the median cannot be reliably calculated.

The article ends with hinting at a potential application. The $\sqrt{2\log n}$ threshold used in the Donoho-Johnstone hard thresholding estimate is replaced by some substitutes guided by the theorems of this article. Our investigations on this potential application are intentionally limited in this article. The conclusion is that although theoretically $E(X_{(n)}) \sim \sqrt{2\log n}$, the $\sqrt{2\log n}$ proxy for $E(X_{(n)})$ is indeed a poor one practically, being too much of an overestimate (e.g., Table 3). We assert that this mythical approximation should be discontinued. Better proxies are suggested and justified to a limited extent in our example of Section 5. This example gives an initial but convincing evidence that suitably adjusting the threshold in the Donoho-Johnstone estimate makes it more effective in a practical sense.

All the proofs are deferred to the appendix in Section 6 for easy reading of the nontechnical parts.

2 Bi-directional Analytic Bounds on the Normal Quantile

Upper and lower bounds on the standard normal quantile function $z_p = \Phi^{-1}(1-p)$ are derived in this section as a precursor to the bounds on the mean and the median of $X_{(n)}$ in the next sections. These analytic bounds on z_p could be of some independent theoretical interest.

We assume that $p \le p_0$(specified), and use the following notation:

$$t_0 = \frac{1}{p_0}; \beta_0 = -\frac{\log \frac{z_{p_0}}{\sqrt{2\log t_0}}}{1-\frac{z_{p_0}}{\sqrt{2\log t_0}}}. \tag{1}$$

Also, for any nonnegative constants a, b, c, d satisfying

$$1-\Phi(z) \le \frac{a\phi(z)}{bz+\sqrt{c^2z^2+d}}, z>0,$$

(such sets of constants a, b, c, d do exist)

denote

$$\alpha = b^2 + c^2, \delta = \frac{d}{2\alpha z_{t_0}^2}. \tag{2}$$

Then, we have the following upper bound on the standard normal quantile z_p.

Theorem 2.1. For $p \leq p_0$ (specified),

$$z_p \leq \sqrt{2\log t} - \frac{\log \frac{4\alpha\pi}{a^2} + \log\log t}{2\sqrt{2\log t}}(1 - \frac{\beta_0}{2\log t}) - \frac{d(1-\delta)}{2\alpha(2\log t)^{3/2}}(1 - \frac{\beta_0}{2\log t}),$$

where $t = \frac{1}{p}$, and $a, b, c, d, \alpha, \delta, \beta_0$ are as defined in (1), (2).

A particularly important consequence of this general upper bound is the following corollary.

Corollary 2.1. For $p \leq 0.1$,

$$z_p \leq \sqrt{2\log t} - \frac{\log 4\pi + \log\log t}{2\sqrt{2\log t}}(1 - \frac{.6403}{\log t}).$$

The upper bound of Corollary 2.1 comes out in its best colors when p is small, i.e., when we want a fully analytic razor sharp upper bound on an extreme quantile. The table below testifies to it.

Table 1

p	z_p	Upper bound of Corollary 2.1
10^{-2}	2.326	2.459
10^{-3}	3.090	3.172
10^{-4}	3.719	3.777
10^{-5}	4.265	4.309
10^{-6}	4.753	4.789
10^{-7}	5.199	5.229
10^{-9}	5.998	6.019
10^{-11}	6.706	6.723
10^{-13}	7.349	7.362
10^{-15}	7.949	7.953

To attack the corresponding lower bound problem correctly, one has to restart with a lower bound on the standard normal probability $1 - \Phi(z)$. If we use the lower

bound $1-\Phi(z) \geq \phi(z)[\frac{1}{z}-\frac{1}{z^3}], z>0$, and carry out calculations similar to the ones we use in proving Theorem 2.1, then the following lower bound on z_p can be derived.

Theorem 2.2. For $p \leq p_0$(specified), such that $z_{p_0} > 1$,

$$z_p \geq \sqrt{2\log t} - \frac{\log 4\pi + \log\log t - 2\log k_0}{2\sqrt{2\log t}}(1 - \frac{1}{4\log t} + \frac{1}{16(\log t)^2}),$$

where $k_0 = \frac{z_{p_0}^2 - 1}{z_{p_0}^2}$.

3 Bounds on the Median of $X_{(n)}$

The bounds of section 2 on the standard normal quantile function are used in this section to obtain bounds on the median of $X_{(n)}$, the maximum of an iid standard normal sample. These bounds on the median do not follow immediately from the bounds of section 2. A fair amount of new work is still needed, although the bounds of section 2 form a critical ingredient. For brevity, we present only the upper bounds on the median, although lower bounds may also be derived. We let θ_n denote the median of $X_{(n)}$.

Theorem 3.1. Let p_0 be specified, and β_0 as defined in (1). Let

$$n_0 = \frac{\log 2}{-\log(1-p_0)}, x_0 = \log\frac{2^{1/n_0}}{2^{1/n_0}-1},$$

$$n_1 = \frac{\log 2}{\beta_0(3 + \frac{4}{\log x_0 + \log 4\pi - 2}) - \log[e^{\beta_0(3+\frac{4}{\log x_0+\log 4\pi-2})} - 1]}.$$

Then for $n \geq \max(n_0, n_1)$, one has

$$\theta_n \leq \sqrt{2\log n}(1 - \frac{2\log\log 2 - \frac{\log 4}{n}}{4\log n}) - \frac{\log 4\pi + \log\log n}{2\sqrt{2\log n}}(1 - \frac{\beta_0}{\log n})(1 - \frac{\frac{\log 2}{n} - \log\log 2}{2\log n}) \tag{3}$$

$$= \sqrt{2\log n} - \frac{\log 4\pi + \log\log n + 2\log\log 2 - \frac{\log 4}{n}}{2\sqrt{2\log n}}$$

$$+ \frac{(\log 4\pi + \log\log n)(\frac{\log 2}{n} - \log\log 2 + 2\beta_0)}{2(2\log n)^{3/2}} - \frac{\beta_0(\log 4\pi + \log\log n)(\frac{\log 2}{n} - \log\log 2)}{(2\log n)^{5/2}}. \tag{4}$$

A less accurate, but simpler and hence easier to use bound is the following corollary.

Corollary 3.1. For $n \geq 31$,

$$\theta_n \leq \sqrt{2\log n} - \frac{\log 4\pi + \log\log n + 2\log\log 2 - \frac{\log 4}{n}}{2\sqrt{2\log n}}(1 - \frac{1.29}{\log n})$$

$$\leq \sqrt{2\log n} - \frac{\log 4\pi + \log\log n - 1}{6\sqrt{2\log n}} \tag{5}$$

Once again, we first provide a small table on the efficacy of the simple bound (5).

Table 2

n	θ_n	Upper bound (5)
50	2.204	2.625
500	2.992	3.367
5000	3.636	3.979
25000	4.031	4.358
10^5	4.346	4.660
10^7	5.267	5.551

We see, as expected, that the upper bound in (5) is more accurate for large n.

4 Bounds on $E(X_{(n)})$

Intuitively, the mean of $X_{(n)}$ is approximately $\Phi^{-1}(1-\frac{1}{n})$, in the sense that $\frac{E(X_{(n)})}{\Phi^{-1}(1-\frac{1}{n})} \to 1$ as $n \to \infty$. Therefore, our bounds on the standard normal quantile function derived in section 2 should be useful in deriving nonasymptotic bounds on $E(X_{(n)})$. This is done in this section. In addition to the bounds in section 2, an inequality used in Lai and Robbins (1976) will be very useful, which is stated as a lemma.

Lemma 4.1. Let F be the CDF of a real valued random variable X with $E(|X|) < \infty$, and let $a_n = F^{-1}(1-\frac{1}{n})$. Then, for any $n \geq 2, E(X_{(n)}) \leq a_n + n\int_{a_n}^{\infty}[1-F(x)]dx$. A consequence of the lemma is the following bound on $E(X_{(n)})$ in the Gaussian case. We will refine this to a more explicit form later in this section.

Corollary 4.1. Let a, b, c, d be any nonnegative constants such that $1 - \Phi(z) \leq \frac{a}{bz+\sqrt{c^2z^2+d}}, z > 0$. Suppose $X_1, \cdots, X_n$ are iid standard normal. Then, for all $n \geq 7$,

$$E(X_{(n)}) \leq z_{\frac{1}{n}} + \frac{az_{\frac{1}{n}}}{(b+c)z_{\frac{1}{n}}^2 + \sqrt{c^2+d} - c}. \tag{6}$$

In particular, for all $n \geq 7$, one has the inequality

$$E(X_{(n)}) \leq z_{\frac{1}{n}} + \frac{1}{z_{\frac{1}{n}}}; \tag{7}$$

and the stronger inequality

$$E(X_{(n)}) \leq z_{\frac{1}{n}} + \frac{2z_{\frac{1}{n}}}{2z_{\frac{1}{n}}^2 + 1}. \tag{8}$$

Our next task is to convert Corollary 4.1 into the explicit bounds announced in the introduction. This is the content of the next theorem.

Theorem 4.1. Let $X_1, \cdots, X_n$ be iid standard normal. For all $n \geq 7$,

$$E(X_{(n)}) \leq \sqrt{2\log n} - \frac{\log 4\pi + \log\log n - 2}{2\sqrt{2\log n}} + C\frac{\log 4\pi + \log\log n}{(2\log n)^{3/2}}, \tag{9}$$

where the constant C may be taken to be 1.5 (or larger).

The bound in (9) has an interesting connection. It is known that in the maximally dependent case, $E(X_{(n)})$ satisfies the relationship

$$E(X_{(n)}) \leq \sqrt{2\log n} - \frac{\log 4\pi + \log\log n - 2}{2\sqrt{2\log n}} + O(\frac{(\log\log n)^2}{(\log n)^{3/2}});$$

(see equation (17) in Lai and Robbins (1976)). It is interesting that this *asymptotic representation* in the maximally dependent case appears exactly in the *nonasymptotic bound* on the mean of $X_{(n)}$ in Theorem 4.1.

5 Asymptotic Expansions and An Example

We remarked in the introduction that the common practice of approximating $E(X_{(n)})$ by $\sqrt{2\log n}$ is not practically sound. The upper bound of Theorem 4.1 offers an alternative approximation. A third approach is outlined below, and then the three approximations are compared in Table 3, which shows without any doubt how poor the $\sqrt{2\log n}$ approximation is. Here is the third approach.

The idea is to approximate $E(X_{(n)})$ by a suitable standard normal quantile $\Phi^{-1}(1-\delta_n)$. One may first consider using $\delta_n = \frac{1}{n}$ for at least two reasons. First, it follows from suitable strong laws for sample maximums that in the normal case $\frac{X_{(n)}}{\sqrt{2\log n}} \stackrel{a.s.}{\to} 1$ (see DasGupta (2008), Chapter 8 for a proof). A second reason is that we can write

$X_n = \Phi^{-1}(1 - W_n)$ with $W_n = 1 - \Phi(X_{(n)})$, and the quantile transformation shows that nW_n converges in distribution to a standard exponential random variable, say W. Thus, heuristically,

$$X_n = \Phi^{-1}(1 - W_n) \approx \Phi^{-1}(1 - \frac{W}{n}) \approx \Phi^{-1}(1 - \frac{1}{n}).$$

However, these heuristics are just not sufficiently effective. We can do better. In fact, the approximation $E(X_{(n)}) \approx \Phi^{-1}(1 - \frac{e^{-\gamma}}{n})$ where γ denotes Euler's constant is a *far better practical approximation to* $E(X_{(n)})$ *than either* $\sqrt{2\log n}$ *or* $\Phi^{-1}(1 - \frac{1}{n})$. Here is a precise result.

Proposition 5.1 Define δ_n by the equation

$$E(X_{(n)}) = \Phi^{-1}(1 - \delta_n) \Leftrightarrow \delta_n = 1 - \Phi(E(X_{(n)})) \tag{10}$$

Then $n\delta_n \to e^{-\gamma}$, where γ denotes Euler's constant.

This suggests that we may consider approximating $E(X_{(n)})$ by $\Phi^{-1}(1 - \frac{e^{-\gamma}}{n})$. We show a small table that demonstrates the very poor performance of the $\sqrt{2\log n}$ approximation, the much better performance of the analytic bound in Theorem 4.1, and the amazingly good approximation obtained by using $\Phi^{-1}(1-\frac{e^{-\gamma}}{n})$. The practice of approximating $E(X_{(n)})$ by $\sqrt{2\log n}$ should probably be discontinued.

Table 3

n	$E(X_{(n)})$	$\sqrt{2\log n}$	Bound of Theorem 4.1	$\Phi^{-1}(1 - \frac{e^{-\gamma}}{n})$
50	2.25	2.80	2.73	2.28
500	3.04	3.53	3.34	3.06
10^4	3.85	4.29	4.06	3.86
10^6	4.86	5.26	5.01	4.87
10^8	5.71	6.07	5.82	5.71
10^9	6.09	6.44	6.19	6.09

A higher order asymptotic expansion for the sequence δ_n defined above may be derived by very carefully putting together asymptotic expansions for $E(X_{(n)})$. This derivation, in turn, requires higher order asymptotic expansions with remainder terms for z_p as an intermediate step, and these, we believe deserve to be more well known than they are. Additionally, the derivation of the asymptotic expansion for $E(X_{(n)})$ to three terms is quite nontrivial. Also, we can imagine the higher order

expansion being useful in some way in extreme value theory, or for high dimensional Gaussian inference. These reasons motivate us to present the following first theorem of this section.

Theorem 5.1. Let $X_1, X_2, \cdots$ be an iid standard normal sequence and let γ denote Euler's constant. Then $E(X_{(n)})$ admits the following three term asymptotic expansion:

$$E(X_{(n)}) = \sqrt{2\log n} - \frac{\log\log n + \log 4\pi - 2\gamma}{2\sqrt{2\log n}}$$

$$+\frac{1}{8(2\log n)^{3/2}} \times \Big[(\log\log n - 2(1-\gamma))^2 + (\log 4\pi - 2)^2 + 4\gamma(\log 4\pi - 2\gamma) - \frac{2}{3}\pi^2\Big]$$

$$+O(\frac{(\log\log n)^2}{(\log n)^{5/2}}).$$

Our method of proof of Theorem 5.1 is such that it automatically delivers a three term asymptotic expansion also for the median of $X_{(n)}$. Here is the corresponding result; notice the remarkable similarity between the asymptotic expansion for the mean in Theorem 5.1 and that for the median in Theorem 5.2 below. The asymptotic expansions should also be compared to the bounds for fixed n in Theorem 4.1 and Corollary 3.1. We can see from these comparisons that the nonasymptotic bounds in Theorem 4.1 and Corollary 3.1 are rather tight, and that there is not much scope for improvement in those nonasymptotic bounds.

Theorem 5.2. Let $X_1, X_2, \cdots$ be an iid standard normal sequence. Then the median of $X_{(n)}$ admits the following three term asymptotic expansion:

$$\theta_n = \text{med}(X_{(n)}) = \sqrt{2\log n} - \frac{\log\log n + \log 4\pi + 2\log\log 2}{2\sqrt{2\log n}}$$

$$+\frac{1}{8(2\log n)^{3/2}} \times \Big[(\log\log n - 2(1+\log\log 2))^2 + (\log 4\pi - 2)^2 - 4\log\log 2(\log 4\pi + 2\log\log 2)\Big]$$

$$+O(\frac{(\log\log n)^2}{(\log n)^{5/2}}).$$

An important potential application below ends this article.

Example 5.1. Adjusting the Donoho-Johnstone Estimate The sparse signal detection problem is one of detecting a very small percentage of relatively weak signals hidden among a large number of pure noises. It is evidently a very difficult

problem, and usually we fail to recognize most of these weak signals. The theoretical development has largely concentrated on Gaussian signals. A popular model is $X_i \overset{indep.}{\sim} (1-\epsilon_n)N(0,1)+\epsilon_n N(\mu_{i,n},1)$. Here, ϵ_n is taken to be a sequence converging to zero at a suitable power rate. See Donoho and Jin (2004) for much more details on this model. Donoho and Johnstone, motivated by Gaussian extreme value theory, gave the *hard thresholding estimate*

$$\hat{\mu}_i = X_i I_{|X_i| > \sqrt{2\log n}}.$$

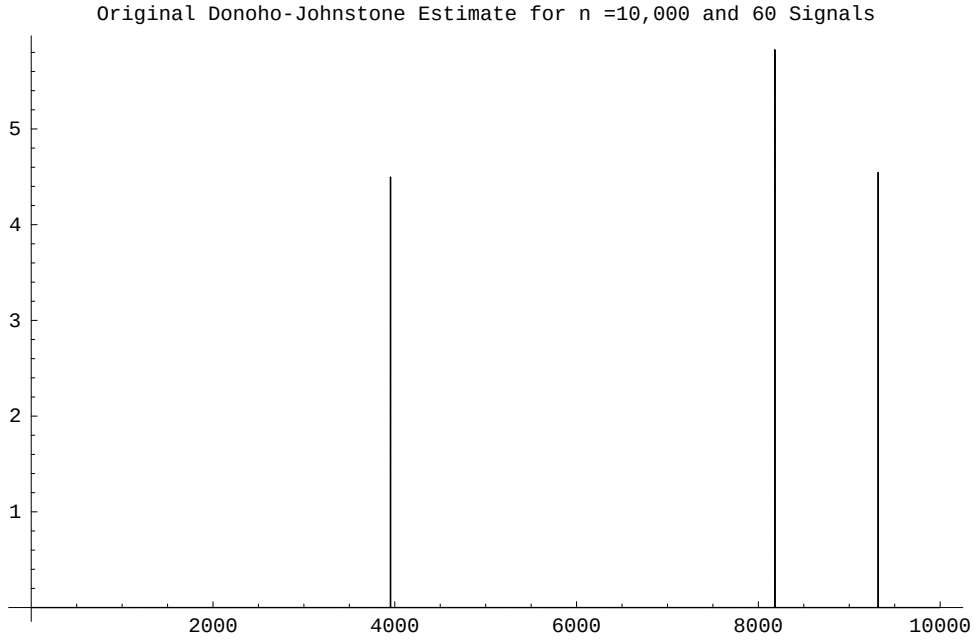
The rough motivation is that if the recorded signal exceeds $E(X_{(n)})$, which is the average value of the maximum among n pure noises, only then we treat it as a possible true signal.

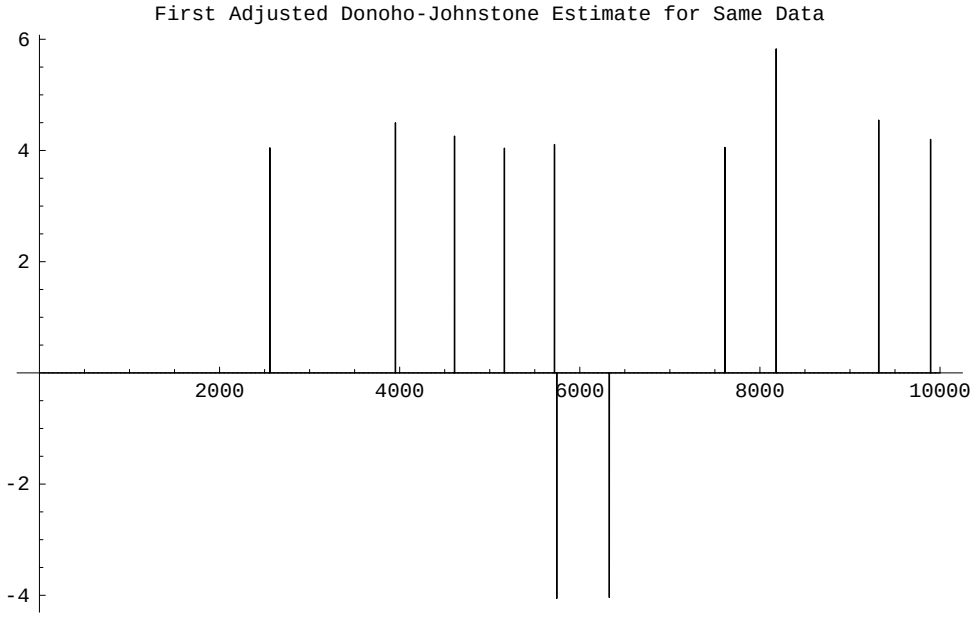
We present a small simulation to suggest that the $\sqrt{2\log n}$ proxy for $E(X_{(n)})$ in the Donoho-Johnstone estimate is practically speaking too conservative, in the sense that a great many of the true signals will not be picked up by using $\sqrt{2\log n}$ as the yardstick. In contrast, if we use, for example, $\Phi^{-1}(1-\frac{e^{-\gamma}}{n})$ as the threshold, then every signal picked up by the original Donoho-Johnstone estimate is picked up, and about four times as many true signals get picked up by readjusting the threshold. Formally, we look at the following three adjustments of the usual Donoho-Johnstone estimate:

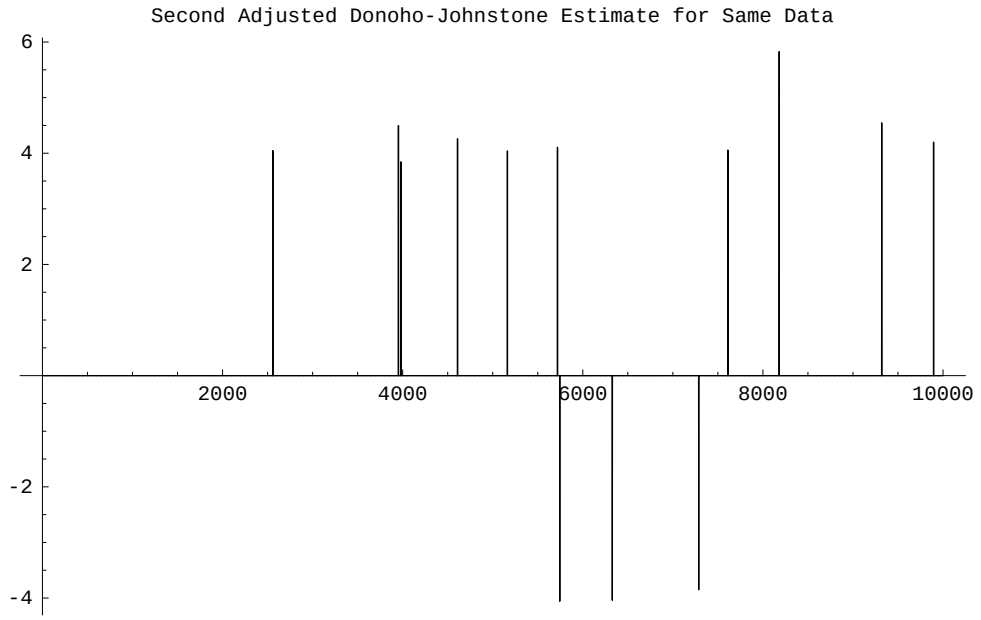
$$\hat{\mu}_i^{(1)} = X_i I_{|X_i| > \Phi^{-1}(1-\frac{e^{-\gamma}}{n})};$$
$$\hat{\mu}_i^{(2)} = X_i I_{|X_i| > \Phi^{-1}(1-\frac{\log 2}{n})};$$
$$\hat{\mu}_i^{(3)} = X_i I_{|X_i| > \sqrt{2\log n} - \frac{\log\log n}{2\sqrt{2\log n}}}.$$

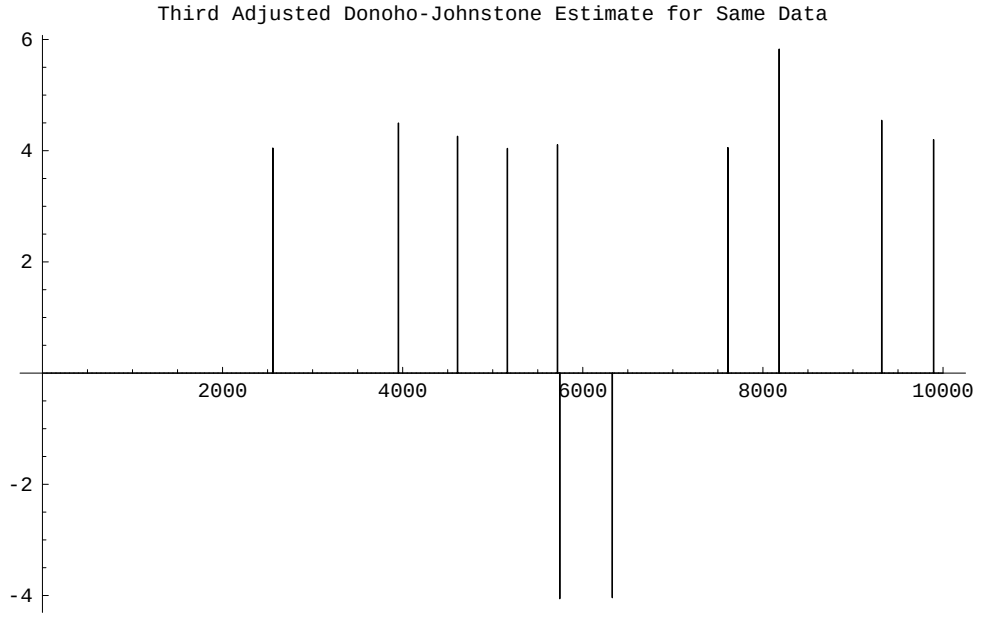
The improved performance of the three adjusted estimates is demonstrated in the figures below. We see that the original Donoho-Johnstone estimate picks up 3 of a total of 60 true signals, and the adjusted Donoho-Johnstone estimates pick up all of those, and additional true signals. The second adjusted estimate with an adjusted threshold of $\Phi^{-1}(1-\frac{\log 2}{n})$, the choice being motivated by Theorem 5.2, picks up the most, namely 13 signals. The common value of the signals was taken to be $\mu_i \equiv \sqrt{\log n} = 3.03$, with $n = 10,000$. There is scope for demonstrating this advantage theoretically, but it will not be pursued in this article because of space

considerations, and also because it will take us too far from the primary theme of this article.









6 *Appendix* Proofs

Proof of Corollary 2.1 Consider the inequality $1 - \Phi(z) \leq \frac{\phi(z)}{z}, z > 0$. In the notation of the theorem, this makes $a = 1, b = 1, c = d = 0, \alpha = 1, \delta = 0$, and since $t_0 = \frac{1}{p_0} = 10$, we have $\beta_0 = -\frac{\log \frac{z_{.1}}{\sqrt{2 \log 10}}}{1 - \frac{z_{.1}}{\sqrt{2 \log 10}}} = 1.28052$, and hence, $\frac{\beta_0}{2} = .64026 < .6403$.

If we plug these values into the bound of Theorem 2.1, the corollary follows.

Proof of Theorem 2.1 Since the constants a, b, c, d are such that $1-\Phi(z) \leq \frac{a\phi(z)}{bz+\sqrt{c^2z^2+d}}$ for all $z > 0$, we have, using $z = z_p$,

$$1 - \Phi(z) = p \leq \frac{a}{\sqrt{2\pi}e^{z^2/2}(bz + \sqrt{c^2z^2 + d})} \tag{11}$$

$$\Rightarrow e^{z^2/2}(bz + \sqrt{c^2z^2 + d}) \leq \frac{a}{\sqrt{2\pi}p} = \frac{at}{\sqrt{2\pi}} \tag{12}$$

$$\Rightarrow e^{z^2/2}(\sqrt{b^2z^2 + c^2z^2 + d}) \leq \frac{at}{\sqrt{2\pi}} \tag{13}$$

$$\Rightarrow 2\log t + \log\frac{a^2}{2\pi} \geq z^2 + \log(\alpha z^2 + d) = z^2 + \log\alpha + \log z^2 + \log(1 + \frac{d}{\alpha z^2}) \tag{14}$$

(recall the notation $\alpha = b^2 + c^2$).

Using the inequality $\log(1 + y) \geq (1 - \delta)y$ for $y \leq 2\delta$ with $y = \frac{d}{\alpha z^2}$, we now have, from (14),

$$2\log t + \log\frac{a^2}{2\pi\alpha} \geq z^2 + 2\log z + (1 - \delta)\frac{d}{\alpha z^2}. \tag{15}$$

We now use a critical technical trick of writing $z = z_p$ in the form

$$z = \sqrt{2\log t} - \frac{\rho}{\sqrt{2\log t}},$$

where $\rho = \rho(t) > 0$. This enables us to transform the bounds derived on z into bounds on ρ, which will turn out to be a useful trick.

From this representation of z, we get

$$z^2 = 2\log t + \frac{\rho^2}{2\log t} - 2\rho,$$

$$\log z = \frac{1}{2}\log 2 + \frac{1}{2}\log\log t + \log(1 - \frac{\rho}{2\log t}),$$

and the inequalities

$$\frac{1}{z^2} \geq \frac{1}{2\log t}, \log z \geq \frac{1}{2}\log 2 + \frac{1}{2}\log\log t - \beta_0\frac{\rho}{2\log t}.$$

(recall the notation $\beta_0 = -\frac{\log\frac{z_{p_0}}{\sqrt{2\log t_0}}}{1-\frac{z_{p_0}}{\sqrt{2\log t_0}}}$ and that it has been assumed that $p \leq p_0$ and thus, $t = \frac{1}{p} \geq t_0 = \frac{1}{p_0}$)

Therefore, from (15),

$$2\log t + \frac{\rho^2}{2\log t} - 2\rho + \log 2 + \log\log t - \beta_0\frac{\rho}{\log t} + \frac{d(1-\delta)}{2\alpha\log t} \leq 2\log t + \log\frac{a^2}{2\pi\alpha} \tag{16}$$

$$\iff \rho^2 - [2\beta_0 + 4\log t]\rho + 2\log t[\log\frac{4\alpha\pi}{a^2} + \log\log t + \frac{d(1-\delta)}{2\alpha\log t}] \leq 0. \tag{17}$$

This is a convex quadratic in the argument ρ with two real roots, given by

$$(2\log t + \beta_0) \pm \sqrt{(2\log t + \beta_0)^2 - 2\log t\{\log\frac{4\alpha\pi}{a^2} + \log\log t + \frac{d(1-\delta)}{2\alpha\log t}\}}. \tag{18}$$

Therefore, by virtue of the convex nature of the quadratic, ρ must be greater than or equal to the lower root, i.e.,

$$\rho \geq (2\log t + \beta_0) - \sqrt{(2\log t + \beta_0)^2 - 2\log t\{\log\frac{4\alpha\pi}{a^2} + \log\log t + \frac{d(1-\delta)}{2\alpha\log t}\}} \tag{19}$$

$$= \frac{2\log t\{\log\frac{4\alpha\pi}{a^2} + \log\log t + \frac{d(1-\delta)}{2\alpha\log t}\}}{(2\log t + \beta_0) + \sqrt{(2\log t + \beta_0)^2 - 2\log t\{\log\frac{4\alpha\pi}{a^2} + \log\log t + \frac{d(1-\delta)}{2\alpha\log t}\}}} \tag{20}$$

(by writing $x - \sqrt{y}$ as $\frac{x^2-y}{x+\sqrt{y}}$)

$$\geq \frac{2\log t\{\log\frac{4\alpha\pi}{a^2} + \log\log t + \frac{d(1-\delta)}{2\alpha\log t}\}}{2(2\log t + \beta_0)}$$

$$\geq \frac{\log\frac{4\alpha\pi}{a^2} + \log\log t + \frac{d(1-\delta)}{2\alpha\log t}}{2}(1 - \frac{\beta_0}{2\log t}). \tag{21}$$

Plug this lower bound (21) on ρ to derive

$$\Rightarrow z = z_p = \sqrt{2\log t} - \frac{\rho}{\sqrt{2\log t}}$$

$$\leq \sqrt{2\log t} - \frac{\log\frac{4\alpha\pi}{a^2} + \log\log t + \frac{d(1-\delta)}{2\alpha\log t}}{2\sqrt{2\log t}}(1 - \frac{\beta_0}{2\log t}), \tag{22}$$

and this is the same as the upper bound in the statement of Theorem 2.1.

Proof of Corollary 3.1 First note that if we choose $p_0 = 0.1$, then we get $\beta_0 = 0.6403$. and the conditions $n \geq n_0$ and $n \geq n_1$ work out respectively to $n \geq 6.58$ and $n \geq 30.55$. Thus, we need $n \geq 31$ for the corollary.

Since the term corresponding to $(2\log n)^{-5/2}$ in Theorem 3.1 is negative, we can ignore it and get the upper bound

$$\theta_n \le \sqrt{2\log n} - \frac{\log 4\pi + \log\log n + 2\log\log 2 - \frac{\log 4}{n}}{2\sqrt{2\log n}}$$

$$+\frac{(\log 4\pi + \log\log n)(\frac{\log 2}{n} - \log\log 2 + 2\beta_0)}{2(2\log n)^{3/2}} \tag{23}$$

$$= \sqrt{2\log n} - \frac{\log 4\pi + \log\log n + 2\log\log 2 - \frac{\log 4}{n}}{2\sqrt{2\log n}}\Big[1 - \frac{\frac{(\log 4\pi+\log\log n)(\frac{\log 2}{n}-\log\log 2+2\beta_0)}{\log 4\pi+\log\log n+2\log\log 2-\frac{\log 4}{n}}}{2\log n}\Big] \tag{24}$$

Now use the fact that the ratio $\frac{\frac{(\log 4\pi+\log\log n)(\frac{\log 2}{n}-\log\log 2+2\beta_0)}{\log 4\pi+\log\log n+2\log\log 2-\frac{\log 4}{n}}}{2}$ is increasing for $n \ge 7$ and is lower bounded by 1.29, and this gives the first statement of the corollary. For the second statement, use the bounds $2\log\log 2 - \frac{\log 4}{n} \ge 2\log\log 2 - \frac{\log 4}{7} > -1$, and $1 - \frac{1.29}{\log n} \ge 1 - \frac{1.29}{\log 7} > \frac{1}{3}$. Plug these two bounds into the first statement of the corollary, and the second statement follows.

Proof of Theorem 3.1 We start with noting that the median θ_n satisfies $\Phi^n(\theta_n) = \frac{1}{2}$, and hence, $\theta_n = z_p = \Phi^{-1}(1-p)$, where $p = p_n = 1 - 2^{-1/n}$. Thus, in the notation of Theorem 2.1, we have $t = \frac{1}{p} = \frac{2^{1/n}}{2^{1/n}-1}$. Theorem 3.1 is going to be obtained by using the bound of Theorem 2.1 on z_p as the principal tool, although additional specific algebraic inequalities will also be needed. Referring to Theorem 2.1, with $t = \frac{2^{1/n}}{2^{1/n}-1}$, we will upper bound $\sqrt{2\log t}$ and $-\frac{\log 4\pi+\log\log t}{2\sqrt{2\log t}}(1 - \frac{\beta_0}{\log t})$ separately. Combining these two separate upper bounds will produce the upper bound on θ_n of Theorem 3.1.

The following inequalities will be used in the proof of this theorem:

$$\log t = \log(\frac{2^{1/n}}{2^{1/n}-1}) = \frac{\log 2}{n} - \log(e^{\frac{\log 2}{n}} - 1)$$

$$\le \frac{\log 2}{n} - \log\log 2 + \log n \tag{25}$$

Hence,

$$\sqrt{2\log t} \le \sqrt{2\log n - 2\log\log 2 + \frac{\log 4}{n}}$$

$$= \sqrt{2\log n}\sqrt{1 - \frac{2\log\log 2 - \frac{\log 4}{n}}{2\log n}} \le \sqrt{2\log n}(1 - \frac{2\log\log 2 - \frac{\log 4}{n}}{4\log n}) \tag{26}$$

We now turn our attention to bounding $-\frac{\log 4\pi+\log\log t}{2\sqrt{2\log t}}(1-\frac{\beta_0}{\log t})$. For this, we will use the following calculus fact: The function $\frac{\log 4\pi+\log x}{\sqrt{x}}(1-\frac{\beta_0}{x})$ is monotone decreasing in x if x satisfies

$$x \geq \beta_0(3+\frac{4}{\log x+\log 4\pi-2}). \tag{27}$$

This is proved by a direct differentiation of $\frac{\log 4\pi+\log x}{\sqrt{x}}(1-\frac{\beta_0}{x})$.

In (27), we identify x with $\log t$. Since it has been already assumed that $x \geq x_0$ (recall the notation that $x_0 = \log\frac{2^{1/n_0}}{2^{1/n_0}-1}$), therefore, the condition $x \geq \beta_0(3+\frac{4}{\log x+\log 4\pi-2})$ holds if

$$x \geq \beta_0(3+\frac{4}{\log x_0+\log 4\pi-2}) \tag{28}$$

holds. On some algebra, this latest condition (28) is seen to be equivalent to

$$n \geq n_1 = \frac{\log 2}{\beta_0(3+\frac{4}{\log x_0+\log 4\pi-2})-\log[e^{\beta_0(3+\frac{4}{\log x_0+\log 4\pi-2})}-1]}. \tag{29}$$

(this explains why we have the condition $n \geq n_1$ in the statement of Theorem 3.1.) This monotone decreasingness result (under the condition $n \geq n_1$) together with the bound $\log t \leq \frac{\log 2}{n} - \log\log 2 + \log n$ (see (25) above) give us

$$\frac{\log 4\pi+\log\log t}{2\sqrt{2\log t}}(1-\frac{\beta_0}{\log t})$$

$$\geq \frac{\log 4\pi+\log(\log n+\frac{\log 2}{n}-\log\log 2)}{2\sqrt{2}\sqrt{\log n+\frac{\log 2}{n}-\log\log 2}}(1-\frac{\beta_0}{\log n+\frac{\log 2}{n}-\log\log 2}). \tag{30}$$

$$= \frac{\log 4\pi+\log(\log n+\frac{\log 2}{n}-\log\log 2)}{2\sqrt{2\log n}(1+\frac{\frac{\log 2}{n}-\log\log 2}{\log n})^{1/2}}(1-\frac{\beta_0}{\log n+\frac{\log 2}{n}-\log\log 2})$$

$$\geq \frac{\log 4\pi+\log(\log n+\frac{\log 2}{n}-\log\log 2)}{2\sqrt{2\log n}(1+\frac{\frac{\log 2}{n}-\log\log 2}{\log n})^{1/2}}(1-\frac{\beta_0}{\log n})$$

$$\geq \frac{\log 4\pi+\log(\log n+\frac{\log 2}{n}-\log\log 2)}{2\sqrt{2\log n}}(1-\frac{\frac{\log 2}{n}-\log\log 2}{2\log n})(1-\frac{\beta_0}{\log n})$$

$$\geq \frac{\log 4\pi+\log\log n}{2\sqrt{2\log n}}(1-\frac{\frac{\log 2}{n}-\log\log 2}{2\log n})(1-\frac{\beta_0}{\log n}), \tag{31}$$

and this is assertion (3) of Theorem 3.1. If we expand (3), we get assertion (4) of the theorem. This completes the proof of Theorem 3.1.

Proof of Corollary 4.1 We note that the inequality in (7) is the key Lai-Robins (1976)

bound, and that (8) is an improvement on (7). To prove Corollary 4.1, observe that by Lemma 4.1,

$$E(X_{(n)}) \leq z_{\frac{1}{n}} + n\int_{z_{\frac{1}{n}}}^{\infty}[1-\Phi(x)]dx = z_{\frac{1}{n}} + n\int_{z_{\frac{1}{n}}}^{\infty}\frac{1-\Phi(x)}{\phi(x)}\phi(x)dx$$

$$\leq z_{\frac{1}{n}} + n\int_{z_{\frac{1}{n}}}^{\infty}\frac{a}{bx+\sqrt{c^2x^2+d}}\phi(x)dx \leq z_{\frac{1}{n}} + n\frac{a}{bz_{\frac{1}{n}}+\sqrt{c^2z_{\frac{1}{n}}^2+d}}\int_{z_{\frac{1}{n}}}^{\infty}\phi(x)dx$$

$$= z_{\frac{1}{n}} + \frac{a}{bz_{\frac{1}{n}}+\sqrt{c^2z_{\frac{1}{n}}^2+d}} \tag{32}$$

Use now the easily provable inequality $\sqrt{c^2z_{\frac{1}{n}}^2+d} \geq cz_{\frac{1}{n}} + \frac{\sqrt{c^2+d}-c}{z_{\frac{1}{n}}}$ for all n such that $z_{\frac{1}{n}} \geq 1$, i.e., for $n \geq 7$. This is where the assumption $n \geq 7$ is needed. Plugging this bound on $\sqrt{c^2z_{\frac{1}{n}}^2+d}$ into (8), we get

$$E(X_{(n)}) \leq z_{\frac{1}{n}} + \frac{a}{bz_{\frac{1}{n}}+cz_{\frac{1}{n}}+\frac{\sqrt{c^2+d}-c}{z_{\frac{1}{n}}}} = z_{\frac{1}{n}} + \frac{a}{(b+c)z_{\frac{1}{n}}+\frac{\sqrt{c^2+d}-c}{z_{\frac{1}{n}}}}$$

$$= z_{\frac{1}{n}} + \frac{az_{\frac{1}{n}}}{(b+c)z_{\frac{1}{n}}^2+\sqrt{c^2+d}-c},$$

which is assertion (6). Assertion (7) follows from (6) by using $a=b=1, c=d=0$, by virtue of the inequality $1-\Phi(z) \leq \frac{\phi(z)}{z}, z>0$. (8) follows from (6) by using $a=4, b=3, c=1, d=8$, by virtue of the inequality $1-\Phi(z) \leq \frac{4\phi(z)}{3z+\sqrt{z^2+8}}, z>0$ (Szarek and Werner (1999)).

Proof of Theorem 4.1: We remark that we use inequality (7) to obtain the result in this theorem. Somewhat better results can be obtained by using the stronger inequality (8) instead.

The starting point for the proof is the result in Corollary 2.1 that

$$z_{\frac{1}{n}} \leq \bar{z}_{\frac{1}{n}} := \sqrt{2\log n} - \frac{\log 4\pi + \log\log n}{2\sqrt{2\log n}}(1-\frac{\beta_0}{\log n}), \tag{33}$$

where β_0 may be taken to be 0.6403. Note now that the function $x \longrightarrow x+\frac{1}{x}$ is increasing in x for $x \geq 1$. Since $z_{\frac{1}{n}} \geq 1$ for $n \geq 7$, this means

$$E(X_{(n)}) \leq z_{\frac{1}{n}} + \frac{1}{z_{\frac{1}{n}}} \leq \bar{z}_{\frac{1}{n}} + \frac{1}{\bar{z}_{\frac{1}{n}}}.$$

The result of this theorem follows by proving that $\bar{z}_{\frac{1}{n}} + \frac{1}{\bar{z}_{\frac{1}{n}}}$ is smaller than the expression in (9). This is done below.

We define ζ by the equation $1 - \frac{1}{\zeta} = \sup_{n \geq 7} \frac{\log 4\pi + \log\log n}{4 \log n}$. This gives $\zeta = 1.69693$. The constant ζ will be used in the proof below by using the inequality $\frac{1}{1-u} \leq 1 + \zeta u$ whenever $u \leq 1 - \frac{1}{\zeta}$, and by identifying u with $u = \frac{\log 4\pi + \log\log n}{4 \log n}$. The proof is completed by noting

$$\bar{z}_{\frac{1}{n}} + \frac{1}{\bar{z}_{\frac{1}{n}}} = \sqrt{2\log n} - \frac{\log 4\pi + \log\log n}{2\sqrt{2\log n}}(1 - \frac{\beta_0}{\log n}) + \frac{1}{\sqrt{2\log n} - \frac{\log 4\pi + \log\log n}{2\sqrt{2\log n}}(1 - \frac{\beta_0}{\log n})}$$

$$= \sqrt{2\log n} - \frac{\log 4\pi + \log\log n}{2\sqrt{2\log n}}(1 - \frac{\beta_0}{\log n}) + \frac{1}{\sqrt{2\log n}[1 - \frac{\log 4\pi + \log\log n}{4\log n}(1 - \frac{\beta_0}{\log n})]}$$

$$\leq \sqrt{2\log n} - \frac{\log 4\pi + \log\log n}{2\sqrt{2\log n}}(1 - \frac{\beta_0}{\log n}) + \frac{1}{\sqrt{2\log n}(1 - \frac{\log 4\pi + \log\log n}{4\log n})}$$

$$\leq \sqrt{2\log n} - \frac{\log 4\pi + \log\log n}{2\sqrt{2\log n}}(1 - \frac{\beta_0}{\log n}) + \frac{1}{\sqrt{2\log n}}(1 + \zeta\frac{\log 4\pi + \log\log n}{4\log n})$$

$$= \sqrt{2\log n} - \frac{\log 4\pi + \log\log n - 2}{2\sqrt{2\log n}} + \frac{\log 4\pi + \log\log n}{(2\log n)^{3/2}}(\beta_0 + \frac{\zeta}{2}),$$

which is bound (9), on writing $\beta_0 + \frac{\zeta}{2} = C$. Since β_0 can be taken to be 0.6403 and ζ can be takken to be 1.69693, it follows that we may take C to be $C = 0.6403 + \frac{1.69693}{2} = 1.48877 < 1.5$, as claimed.

Proof of Proposition 5.1 Let

$$a_n = \sqrt{2\log n} - \frac{\log\log n + \log 4\pi}{2\sqrt{2\log n}}, n \geq 2. \tag{34}$$

Let also Z have the standard Gumbel distribution with the CDF $G(z) = e^{-e^{-z}}, -\infty < z < \infty$. Then, it is well known that

$$\sqrt{2\log n}[X_{(n)} - a_n] \stackrel{\mathcal{L}}{\Rightarrow} Z,$$

and that $\{Z_n\}$ is uniformly integrable. Therefore, $E(Z_n) \to E(Z) = \gamma$, the Euler constant. By transposition, we have the implication

$$E(X_{(n)}) = \sqrt{2\log n} - \frac{\log\log n + \log 4\pi - 2\gamma}{2\sqrt{2\log n}} + o(\frac{1}{\sqrt{\log n}}) = b_n + o(\frac{1}{\sqrt{\log n}}), \tag{35}$$

where

$$b_n = \sqrt{2\log n} - \frac{\log\log n + \log 4\pi - 2\gamma}{2\sqrt{2\log n}}. \tag{36}$$

Therefore,

$$\delta_n = 1 - \Phi(b_n + o(\frac{1}{\sqrt{\log n}})). \tag{37}$$

Now, using the fact that $1 - \Phi(x) = \frac{\phi(x)}{x}(1 + O(x^{-2}))$ as $x \to \infty$, we get

$$\delta_n = \frac{1}{\sqrt{2\pi}} \frac{e^{-\frac{1}{2}[2\log n - (\log\log n + \log 4\pi - 2\gamma) + o(1)]}}{\sqrt{2\log n}(1 + o(1))}$$

$$= \frac{1}{\sqrt{2\pi}} \frac{1}{n} \sqrt{\log n}\, (2\sqrt{\pi}) e^{-\gamma}(1 + o(1)) = \frac{e^{-\gamma}(1 + o(1))}{n}, \tag{38}$$

which is the claim in the proposition.

Proof of Theorem 5.1 The representation $E(X_{(n)}) = b_n + o(\frac{1}{\sqrt{\log n}})$ in equation (35) is not sharp enough to produce an asymptotic expansion for $E(X_{(n)})$. We will derive such an higher order asymptotic expansion for $E(X_{(n)})$ by using the trick of reducing a problem in extreme value theory to a problem about sums. This will be achieved by using the celebrated *Réyni representation* for the vector of uniform order statistics. This trick may be much more generally useful in extreme value theory than the particular purpose for which it is used in this article.

Because the proof of this theorem is long and intricate, we break it into smaller main steps.

Step 1 Let $U_1, U_2, \cdots$ be an iid $U[0,1]$ sequence and let $U_{(n)} = \max\{U_1, U_2, \cdots, U_n\}$. Let $X_1, X_2, \cdots$ denote an iid standard normal sequence, and let $W_0, W_1, \cdots$ denote an iid standard exponential sequence. Finally, let $Z_n = \frac{W_1 + \cdots + W_n - n}{\sqrt{n}}$, and let $F_n(z)$ denote the CDF of Z_n.

By *Réyni's representation*, $U_{(n)} \stackrel{\mathcal{L}}{=} \frac{W_1 + \cdots + W_n}{W_0 + W_1 + \cdots + W_n}$, and therefore, by the quantile transformation

$$X_{(n)} \stackrel{\mathcal{L}}{=} \Phi^{-1}(U_{(n)}) \stackrel{\mathcal{L}}{=} \Phi^{-1}(1 - \frac{W_0}{W_0 + W_1 + \cdots + W_n}).$$

Hence,

$$\mu_n = E(X_{(n)}) = E[\Phi^{-1}(1 - \frac{W_0}{W_0 + W_1 + \cdots + W_n})]$$

$$= E[\Phi^{-1}(1 - \frac{W_0}{W_0 + n + \sqrt{n}Z_n})] = \int_{\mathcal{R}} \int_0^\infty \Phi^{-1}(1 - \frac{x}{x + n + \sqrt{n}z}) e^{-x} dx dF_n(z). \tag{39}$$

Step 2 For every given x, z, the argument $\frac{x}{x + n + \sqrt{n}z}$ in (39) is small for large n. Thus, the idea now is to obtain and use sufficiently high order asymptotic expansions for

$\Phi^{-1}(1-p)$ when $p \to 0$.

The asymptotic expansions for $\Phi^{-1}(1-p)$ are derived by inverting *Laplace's expansion* for the standard normal tail probability, which is

$$1 - \Phi(x) = \phi(x)[\frac{1}{x} - \frac{1}{x^3} + \frac{1 \times 3}{x^5} - \cdots + (-1)^k \frac{1 \times 3 \times (2k-1)}{x^{2k+1}} + R_k(x)], \quad (40)$$

where for any specific k, $R_k(x) = O(x^{-(2k+3)})$ as $x \to \infty$. By using Laplace's expansion (40) with $k = 1$, we get the following successive asymptotic expansions for $z_p = \Phi^{-1}(1-p)$ as $p \to 0$:

Writing $t = \frac{1}{p}$,

$$z_p^2 = 2\log t + O(\log\log t);$$

$$z_p^2 = 2\log t - \log\log t + O(1);$$

$$z_p^2 = 2\log t - \log\log t - \log 4\pi + O(\frac{\log\log t}{\log t});$$

$$z_p^2 = 2\log t - \log\log t - \log 4\pi + \frac{\log\log t}{2\log t} + O(\frac{1}{\log t});$$

$$z_p^2 = 2\log t - \log\log t - \log 4\pi + \frac{\log\log t + \log 4\pi - 2}{2\log t} + O(\frac{\log\log t}{(\log t)^2}). \quad (41)$$

Step 3 From the final expression in (41),

$$z_p = \sqrt{2\log t}[1 - \frac{\log\log t + \log 4\pi}{2\log t} + \frac{\log\log t + \log 4\pi - 2}{(2\log t)^2} + O(\frac{\log\log t}{(\log t)^3})]^{1/2}$$

$$= \sqrt{2\log t}[1 - \frac{\log\log t + \log 4\pi}{4\log t} - \frac{(\log\log t + \log 4\pi)^2 - 4\log\log t - 4\log 4\pi + 8}{32(\log t)^2}$$

$$+ O(\frac{(\log\log t)^2}{(\log t)^3})]$$

$$= \sqrt{2\log t}[1 - \frac{\log\log t + \log 4\pi}{4\log t} - \frac{(\log\log t - 2)^2 + (\log 4\pi - 2)^2}{32(\log t)^2}$$

$$+ O(\frac{(\log\log t)^2}{(\log t)^3})] \quad (42)$$

Step 4 Equation (42) in conjunction with the integral representation in (39) will be key in producing the desired asymptotic expansion for μ_n. We begin with μ_n. Toward this, with (39) in mind, we use $p = \frac{x}{n+z\sqrt{n}+x}$ and $t = \frac{n+z\sqrt{n}+x}{x} = \frac{n}{x}[1 + \frac{z}{\sqrt{n}} + \frac{x}{n}]$. We will now use (42) and derive asymptotic expansions up to the needed order for each of the three terms in expression (42). Then, we will combine them to produce

the final asymptotic expansion for μ_n.

The first term in (42)

$$T_1 := \sqrt{2\log t} = \sqrt{2(\log n - \log x) + O_p(\frac{1}{\sqrt{n}})}$$

$$= \sqrt{2\log n}\sqrt{1 - \frac{\log x}{\log n} + O_p(\frac{1}{\sqrt{n}\log n})}$$

$$= \sqrt{2\log n} - \frac{\log x}{\sqrt{2\log n}} - \frac{(\log x)^2}{2(2\log n)^{3/2}} + O_p((\log n)^{-5/2}). \tag{43}$$

Next,

$$T_2 := \frac{\log\log t + \log 4\pi}{2\sqrt{2\log t}}$$

$$= \frac{\log(\log n - \log x + O_p(\frac{1}{\sqrt{n}})) + \log 4\pi}{2[\sqrt{2\log n} - \frac{\log x}{\sqrt{2\log n}} + O_p((\log n)^{-3/2})]}$$

$$= \frac{\log\log n - \frac{\log x}{\log n} + \log 4\pi + O_p((\log n)^{-2})}{2\sqrt{2\log n}[1 - \frac{\log x}{2\log n} + O_p((\log n)^{-2})]}$$

$$= \frac{\log\log n - \frac{\log x}{\log n} + \log 4\pi + \log x\frac{\log\log n}{2\log n} + \log x\frac{\log 4\pi}{2\log n} + O_p(\frac{\log\log n}{(\log n)^2})}{2\sqrt{2\log n}}. \tag{44}$$

It remains to handle the third term in (42), namely

$$T_3 := \frac{(\log\log t - 2)^2 + (\log 4\pi - 2)^2}{\frac{32(\log t)^2}{\sqrt{2\log t}}}.$$

For this, we use the following three observations:

$$(\log\log t)^2 = [\log\log n - \frac{\log x}{\log n} + O_p((\log n)^{-2})]^2$$

$$= (\log\log n)^2 - (2\log x)\frac{\log\log n}{\log n} + O_p(\frac{\log\log n}{(\log n)^2}); \tag{45}$$

$$-4\log\log t = -4\log\log n + 4\frac{\log x}{\log n} + O_p((\log n)^{-2}); \tag{46}$$

and,

$$\frac{1}{\frac{32(\log t)^2}{\sqrt{2\log t}}} = \frac{1}{16\sqrt{2}(\log t)^{3/2}} = \frac{1}{8(2\log t)^{3/2}}$$

$$= \frac{1 + \frac{3}{2}\frac{\log x}{\log n} + O_p((\log n)^{-2})}{8(2\log n)^{3/2}} \tag{47}$$

By combining (42), (43), (44), (45), (46), and (47), we get

$$\Phi^{-1}(1-\frac{x}{x+z\sqrt{n}+n}) = \sqrt{2\log n} - \frac{\log\log n + \log 4\pi + 2\log x}{2\sqrt{2\log n}}$$

$$+\frac{1}{8(2\log n)^{3/2}} \times \Big[(\log\log n)^2 - 4\log\log n - 4(\log x)\log\log n + (8-4\log 4\pi)(\log x)$$

$$-4(\log x)^2 + (\log 4\pi)^2 - 4\log 4\pi + 8\Big] + O_p(\frac{(\log\log n)^2}{(\log n)^{5/2}}). \tag{48}$$

Notice the interesting fact that no terms involving z appear in this three term expansion for $\Phi^{-1}(1-\frac{x}{x+z\sqrt{n}+n})$, although on introspection it is clear why the z terms do not appear.

Step 5 Using now the integration facts

$$\int_0^\infty (\log x)e^{-x}dx = -\gamma;\ \int_0^\infty (\log x)^2 e^{-x}dx = \gamma^2 + \frac{\pi^2}{6},$$

and (39), (48), and Bhattacharya and Rao (1986), we get the three term asymptotic expansion for $E(X_{(n)})$:

$$\mu_n = E(X_{(n)}) = \sqrt{2\log n} - \frac{\log\log n + \log 4\pi - 2\gamma}{2\sqrt{2\log n}}$$

$$+\frac{1}{8(2\log n)^{3/2}} \times \Big[(\log\log n)^2 - 4(1-\gamma)\log\log n + \gamma(4\log 4\pi - 8) - 4(\gamma^2 + \frac{\pi^2}{6})$$

$$+(\log 4\pi)^2 - 4\log 4\pi + 8\Big] + O(\frac{(\log\log n)^2}{(\log n)^{5/2}})$$

$$= \sqrt{2\log n} - \tfrac{\log\log n+\log 4\pi-2\gamma}{2\sqrt{2\log n}} + \tfrac{1}{8(2\log n)^{3/2}} \times \Big[(\log\log n - 2(1-\gamma))^2$$

$$+(\log 4\pi - 2)^2 + 4\gamma(\log 4\pi - 2\gamma) - \frac{2}{3}\pi^2\Big] + O(\frac{(\log\log n)^2}{(\log n)^{5/2}}), \tag{49}$$

and this establishes the theorem.

Proof of Theorem 5.2 Theorem 5.2 follows from our calculations in the proof of Theorem 5.1 after making a simple observation. Recall that $\theta_n = \Phi^{-1}(2^{-\frac{1}{n}})$. Therefore, $\theta_n = \Phi^{-1}(1-p)$, with

$$t = \frac{1}{p} = \frac{2^{\frac{1}{n}}}{2^{\frac{1}{n}}-1} = \frac{n}{\log 2} + O(1). \tag{50}$$

Therefore, the asymptotic expansion for θ_n up to an error of $O(\frac{(\log\log n)^2}{(\log n)^{5/2}})$ follows in a straightforward manner from formula (42). The details of the algebra will be omitted.

7 *Bibliography*

Bhattacharya, R. N. and Rao, R. R. (1986). Normal Approximation and Asymptotic Expansions, Wiley, New York.

Donoho, D. and Jin, J. (2004). Higher criticism for detecting sparse heterogenous mixtures, Ann. Stat., 32, 962-994.

Donoho, D. and Johnstone, I. M. (1994). Minimax risk over ℓ_p balls with ℓ_q error, Prob. Theory Related Fields, 2, 277-303.

Galambos, J. (1978). Asymptotic Theory of Extreme Order Statistics, Wiley, New York.

Gascuel, O. and Caraux, G. (1992). Bounds on expectations of order statistics via extremal dependences, Stat and Prob Letters, 15, 2, 143-148.

Hall, P. (1979). On the rate of convergence of normal extremes, Jour. Appl. Prob., 16, 2, 433-439.

Lai, T. L. and Robbins, H. (1976). Maximally dependent random variables, Proc. Nat. Acad. Sci., USA, 73, 286-288.

Rychlik, T. (1998). Bounds on expectations of L-estimates, in Handbook of Statistics, 16, Amsterdam, North-Holland, 105-145.

Szarek, S. and Werner, E. (1999), A nonsymmetric correlation inequality for Gaussian measures, Jour. Mult. Anal., 68, 193-211.