Order Statistics and Pitman Closeness

Katherine F. Davies

Department of Statistics
University of Manitoba

October 12, 2010
1. Introduction

2. Pitman Closeness of Order Statistics to Population Quantiles

3. SCP of Order Statistics to Population Quantiles

4. SCP Plotting Points

5. Conclusions
We recognize that much interest centers on comparisons of optimal estimators by different criteria.

For example, one may want to compare an unbiased estimator to a minimum mean squared error estimator.

If one uses unbiasedness or mean squared error, the outcome is obvious.

In this regard, Rao recommended comparison by an alternative criterion, that of Pitman’s measure of closeness [15].
Pitman introduced his measure of closeness in 1937 as an alternative criterion in parameter estimation, for instance as an alternative to MSE.

The measure is based on the probabilities of the relative closeness of competing estimators to an unknown parameter.

There are two important notions that can be defined using Pitman's measure of closeness.

The first notion is that of the Pitman-closer estimator; the second is Pitman-closest estimator.
Two Important Definitions

- **Definition 1**: When comparing two estimators $\hat{\theta}_1$ and $\hat{\theta}_2$ of $\theta$, then $\hat{\theta}_1$ is a Pitman-closer estimator than $\hat{\theta}_2$ if

  \[
  \Pr(|\hat{\theta}_1 - \theta| < |\hat{\theta}_2 - \theta|) \geq 1/2
  \]

  $\forall \theta \in \Omega$ with strict inequality for at least one $\theta$. [2]

- This formulation gives rise to the Pitman-closest estimator.
For the Pitman-closest estimator, we have the following definition:

**Definition 2:** Let \( A \) be a nonempty class of estimators of a common parameter \( \theta \). Then, \( \hat{\theta}_* \) is **Pitman-closest** among estimators in \( A \) provided for every \( \hat{\theta} \in A \), such that \( \hat{\theta} \neq \hat{\theta}_* \), i.e.:

\[
P(\theta; \hat{\theta}_*, \hat{\theta}) = \Pr(|\hat{\theta}_* - \theta| < |\hat{\theta} - \theta|) \geq 1/2
\]

\( \forall \theta \in \Omega \) with strict inequality for at least one \( \theta \), where \( P(\theta; \hat{\theta}_*, \hat{\theta}) \) is Pitman closeness or nearness of \( \hat{\theta}_* \) to \( \hat{\theta} \) given \( \theta \). [14]

For further details, we refer those interested to the monograph by Keating, Mason and Sen [14] for an elaborate discussion on this topic.
Recently, many problems involving ordered data and Pitman closeness have been investigated.

Balakrishnan et al. [6] looked at comparing best linear unbiased and invariant estimators for the exponential mean parameter using Pitman closeness criterion, and similarly, a comparison of best linear unbiased and invariant predictors have been compared in Balakrishnan et al. [7].

Pitman closeness of records to population quantiles was explored in Ahmadi and Balakrishnan [1].
Motivating Work

- In 2008, Balakrishnan et al. [5] performed a similar investigation to that of Ahmadi and Balakrishnan [1]; they investigated Pitman closeness of sample median to population median.
- Further interesting results regarding the sample median and Pitman closeness were provided in a followup paper by Iliopoulous and Balakrishnan [11].
- The population median is a particular quantile and a natural followup problem is to look at population quantiles in general and that is our primary focus here.
Throughout this talk we shall assume the following common conditions:

1. Let $Y_1, \ldots, Y_n$ be a random sample of size $n$ from a continuous population with a cumulative distribution function (cdf) $G(y)$ and density function (pdf) $g(y)$ and $Y_{1:n} \leq \cdots \leq Y_{n:n}$ denote the corresponding order statistics.

2. Let $\xi_p$ be the $p^{th}$ quantile of $G(y)$, i.e.,

$$\Pr \left( Y \leq \xi_p \right) = p \quad \text{for } p \in (0, 1).$$

Furthermore, we shall make an assumption about the distribution $G(\cdot)$ which will restrict our attention.
An Assumption

- Suppose $G(\cdot)$ belongs to the location-scale family of distributions, viz.,

\[ G(y) = F\left(\frac{y - \mu}{\sigma}\right) \quad \text{and} \quad g(y) = \frac{1}{\sigma} f\left(\frac{y - \mu}{\sigma}\right) \quad \forall \ y \in \mathbb{R}, \]

where $\mu \in \mathbb{R}$ is the location parameter and $\sigma > 0$ is the scale parameter, then we have the following additional conditions:

1. We have $\xi_p = (\xi_p^* - \mu)/\sigma$ is the $p^{th}$ quantile of $F(x)$.
2. We shall let $X_1, \cdots, X_n$ be a random sample of size $n$ from the population with cdf $F(x)$ and pdf $f(x)$, and $X_{1:n} \leq \cdots \leq X_{n:n}$ denote the corresponding order statistics.
Our first primary objective is to extend the work on population median to the case of population quantiles.

Our goal is to consider the probability with which one order statistic will be \textit{Pitman-closer} to a quantile (say, the \( p^{th} \) quantile) than another order statistic from the same sample (i.e. pairwise comparisons).

These closeness probabilities will lead to identification of that order statistic which is \textit{Pitman-closest}.
A Definition

In the context of order statistics, we have an analogue to Definition 2 for comparing the closeness of order statistics to population quantiles.

**Definition 3:** The $\ell^{th}$ order statistic (for some $\ell \in \{1, \cdots, n\}$) is the Pitman-closest order statistic to the population quantile $\xi_p$ if:

$$\Pr (|X_{\ell:n} - \xi_p| < |X_{i:n} - \xi_p|) \geq \frac{1}{2}$$

$\forall \ i \in \{1, \cdots, n\}\setminus\{\ell\}$.

Knowing that order statistics have the ordering preserving property under location-scale transforms, we have the following result.
Result 1: Suppose $X_{\ell:n}$ (for some $\ell \in \{1, \cdots, n\}$) is the Pitman-closest order statistic to $\xi_p$, i.e.,

$$\Pr(|X_{\ell:n} - \xi_p| < |X_{i:n} - \xi_p|) \geq \frac{1}{2}$$

$\forall \ i \in \{1, \cdots, n\}\backslash\{\ell\}$. Then, $Y_{\ell:n}$ is the Pitman-closest order statistic to $\xi_p^*$.

Remark: Because of Result 1, it suffices to look for the Pitman-closest order statistic to a quantile $\xi_p$ for the standard distribution $F(\cdot)$.
Another Important Result

In the case when the population distribution is symmetric about origin, we can easily establish the following symmetry property.

**Result 2:** Suppose $X_{\ell:n}$ is the Pitman-closest order statistic to the $p^{th}$ quantile $\xi_p$ of a distribution symmetric about origin, i.e.,

$$\Pr(|X_{\ell:n} - \xi_p| < |X_{i:n} - \xi_p|) \geq \frac{1}{2}$$

for all $i \in \{1, \cdots, n\} \setminus \{\ell\}$. Then, $X_{n-\ell+1:n}$ is the Pitman-closest order statistic to the $(1-p)^{th}$ quantile $\xi_{1-p}$, i.e.,

$$\Pr(|X_{n-\ell+1:n} - \xi_{1-p}| < |X_{i:n} - \xi_{1-p}|) \geq \frac{1}{2}$$

for all $i \in \{1, \cdots, n\} \setminus \{n - \ell + 1\}$. 
Some New Notation

- For some $p \in (0, 1)$, let us denote
  \[
  \pi(\ell)_i(p) = \Pr \left( |X_{\ell:n} - \xi_p| < |X_{i:n} - \xi_p| \right) \quad \text{for} \quad i = \{1, \ldots, n\} \setminus \{\ell\}
  \]
  for the probability of Pitman closeness to $\xi_p$ associated with any two order statistics.
- Using what we know about order statistics, we have the following general expressions for $\pi(\ell)_i(p)$. 
Pitman Closeness Probabilities

**Result 3:** For \( i = \ell + 1, \ell + 2, \ldots, n, \)

\[
\pi(\ell)_i(p) = 1 - I_p(\ell, n - \ell + 1) + k_{\ell,i,n} \sum_{j=0}^{i-\ell-1} (-1)^{i-\ell-1-j} \binom{i-\ell-1}{j} \frac{1}{n - \ell - j} \\
\times \int_{-\infty}^{\xi_p} \{F(x)\}^{\ell-1} \{1 - F(x)\}^j \{1 - F(-x + 2\xi_p)\}^{n-\ell-j} f(x)dx,
\]

and for \( i = 1, 2, \ldots, \ell - 1, \)

\[
\pi(\ell)_i(p) = I_p(\ell, n - \ell + 1) + k_{i,\ell,n} \sum_{j=0}^{\ell-i-1} (-1)^{\ell-i-1-j} \binom{\ell-i-1}{j} \frac{1}{\ell - j - 1} \\
\times \int_{\xi_p}^{\infty} \{F(y)\}^j \{1 - F(y)\}^{n-\ell} \{F(2\xi_p - y)\}^{\ell-j-1} f(y)dy,
\]
where

\[ I_x(\alpha, \beta) = \frac{1}{B(\alpha, \beta)} \int_0^x u^{\alpha-1}(1 - u)^{\beta-1} du, \quad 0 < x < 1, \]

is the incomplete beta ratio, \( B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)} \) is the complete beta function, and

\[ k_{\ell,i,n} = \frac{n!}{(\ell - 1)!(i - \ell - 1)!(n - i)!} \quad \text{for } 1 \leq \ell < i \leq n. \]
Three Cases of Interest

- As can be seen, the expressions for the probability of Pitman closeness are distribution dependent.
- We consider three location-scale distributions: the Uniform, exponential and power function.
- For each distribution we can determine the probability of Pitman closeness to $\xi_p$ associated with any two order statistics.
- With these probabilities, we can determine the Pitman-closest order statistic to a population quantile $\xi_p$. 
Let us consider the $U(-1, 1)$ distribution with pdf and cdf as

$$f(x) = \frac{1}{2} \quad \text{and} \quad F(x) = \frac{1 + x}{2} \quad \text{for} \quad -1 < x < 1,$$

and its $p^{th}$ quantile as $\xi_p = 2p - 1$ for $p \in (0, 1)$.

With these quantities, we can use Result 3 to yield the following expressions for the Pitman closeness probability associated with any two order statistics from a uniform sample, and then identify the closest order statistic in the Pitman sense.
Results

Result 4: For $i = \ell + 1, \ldots, n$,

$$\pi(\ell)_i(p) = 1 - l_p(\ell, n - \ell + 1) + k_{\ell, i, n} \sum_{j=0}^{i-\ell-1} (-1)^{i-\ell-1-j} \binom{i - \ell - 1}{j} \frac{1}{n - \ell - j}$$

$$\times \frac{1}{2^n} \sum_{a=0}^{j} \sum_{b=0}^{n-\ell-j} \binom{j}{a} \binom{n - \ell - j}{b} (1 - \xi_p)^{j-a} (-2\xi_p)^{n-\ell-j-b}$$

$$\times (1 + \xi_p)^{\ell+a+b} B(\ell + b, a + 1) \quad \text{for } 0 < p < \frac{1}{2}$$

$$= 1 - l_p(\ell, n - \ell + 1) + k_{\ell, i, n} \sum_{j=0}^{i-\ell-1} (-1)^{i-\ell-1-j} \binom{i - \ell - 1}{j} \frac{1}{n - \ell - j}$$

$$\times \frac{1}{2^n} \sum_{a=0}^{l-1} \sum_{b=0}^{j} \binom{l - 1}{a} \binom{j}{b} (2\xi_p)^{l-1-a} (1 - \xi_p)^{a+n-\ell+1}$$

$$\times B(a + n - \ell - j + 1, b + 1) \quad \text{for } \frac{1}{2} \leq p < 1.$$
Results cont’d

Similarly, for \( i = 1, \ldots, \ell - 1 \),

\[
\pi(\ell)_i(p) = I_p(\ell, n - \ell + 1) + k_{i,\ell,n} \sum_{j=0}^{\ell-i-1} (-1)^{\ell-i-1-j} \binom{\ell-i-1}{j} \frac{1}{\ell-j-1}
\]

\[
\times \frac{1}{2^n} \sum_{a=0}^{j} \sum_{b=0}^{n-\ell} \binom{j}{a} \binom{n-\ell}{b} (1 + \xi_p)^{\ell+b} (-2\xi_p)^{n-\ell-b}
\]

\[
\times B(a + 1, \ell - j + b) \quad \text{for} \quad 0 < p < \frac{1}{2}
\]

\[
= I_p(\ell, n - \ell + 1) + k_{i,\ell,n} \sum_{j=0}^{\ell-i-1} (-1)^{\ell-i-1-j} \binom{\ell-i-1}{j} \frac{1}{\ell-j-1}
\]

\[
\times \frac{1}{2^n} \sum_{a=0}^{j} \sum_{b=0}^{\ell-j-1} \binom{j}{a} \binom{\ell-j-1}{b} (1 + \xi_p)^{j-a} (2\xi_p)^{\ell-j-1-b}
\]

\[
\times (1 - \xi_p)^{a+n-\ell+b+1} B(a + 1, n - \ell + b + 1) \quad \text{for} \quad \frac{1}{2} \leq p < 1.
\]
<table>
<thead>
<tr>
<th>$\ell$</th>
<th>$p = 0.10$</th>
<th>$p = 0.25$</th>
<th>$p = 0.75$</th>
<th>$p = 0.90$</th>
<th>$p = 0.10$</th>
<th>$p = 0.25$</th>
<th>$p = 0.75$</th>
<th>$p = 0.90$</th>
<th>$p = 0.10$</th>
<th>$p = 0.25$</th>
<th>$p = 0.75$</th>
<th>$p = 0.90$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Closeness probabilities for Uniform for $n=10$
Figure 1: Closeness probabilities for uniform with $n = 10$, $p = 0.10, 0.25, 0.75, 0.90$
Table 2: Pitman-closest order statistic to the $p^{th}$ quantile of $\mathcal{U}(-1,1)$ from a sample of size $n$
Quantities of interest

- Let us consider the **standard exponential** distribution with pdf and cdf as

\[ f(x) = e^{-x} \quad \text{and} \quad F(x) = 1 - e^{-x} \quad \text{for} \ x > 0, \]

and its \( p^{th} \) quantile as \( \xi_p = -\ln(1 - p) \) for \( p \in (0, 1) \).

- Result 3 can again be used to derive the following expressions for the Pitman closeness probability associated with any two order statistics.

- We can then identify for different sample sizes and quantiles that order statistic which is Pitman-closest.
Result 5: For \( i = \ell + 1, \cdots, n, \)

\[
\pi(\ell)_i(p) = 1 - l_p(\ell, n - \ell + 1) + k_{\ell,i,n} \sum_{j=0}^{i-\ell-1} (-1)^{i-\ell-1-j} \binom{i-\ell-1}{j} \frac{1}{n-\ell-j} \\
\times (1 - p)^{2(n-\ell-j)} \sum_{a=0}^{\ell-1} (-1)^a \binom{\ell-1}{a} H_{a,j,\ell,n}(p),
\]

where

\[
H_{a,j,\ell,n}(p) = \begin{cases} 
\frac{1}{a+2j+\ell-n+1} \left[ 1 - (1 - p)^{a+2j+\ell-n+1} \right] & \text{if } a + 2j + \ell - n + 1 \neq 0 \\
\xi_p & \text{if } a + 2j + \ell - n + 1 = 0.
\end{cases}
\]
Similarly, for \( i = 1, \cdots, \ell - 1 \),

\[
\pi_{(\ell)i}(p) = l_p(\ell, n - \ell + 1) + k_{i,\ell,n} \sum_{j=0}^{\ell-i-1} (-1)^{\ell-i-1-j} \binom{\ell - i - 1}{j} \frac{1}{\ell - j - 1} \\
\times \sum_{a=0}^{\ell-j-1} \sum_{b=0}^{j} (-1)^{a+b} \binom{\ell - j - 1}{a} \binom{j}{b} (1 - p)^{2a} J_{a,b,\ell,n}(p),
\]

where

\[
J_{a,b,\ell,n}(p) = \begin{cases} \\
\frac{1}{n-\ell-a+b+1} \left[ (1 - p)^{n-\ell-a+b+1} - (1 - p)^{2(n-\ell-a+b+1)} \right] \\
\xi_p \quad \text{if } n - \ell - a + b + 1 \neq 0 \\
\xi_p \quad \text{if } n - \ell - a + b + 1 = 0.
\end{cases}
\]
A Remark

Suppose we have two independent standard exponential random variables, and look at the corresponding order statistics $X_{1:2}$ and $X_{2:2}$.

Then, intuition suggests that $X_{1:2}$ will be closer to $\xi_p$ for $p$ up to some value after which $X_{2:2}$ will be closer to the quantile $\xi_p$.

With $n = 2$, $\ell = 1$ and $i = 2$, we find from our expression above that

$$\pi_{(1)2}(p) = \Pr(|X_{1:2} - \xi_p| < |X_{2:2} - \xi_p|)$$

$$= 1 - I_p(1, 2) + 2(1 - p)^2\{-\ln(1 - p)\}$$

$$= 1 - \{1 - (1 - p)^2\} - 2(1 - p)^2 \ln(1 - p)$$

$$= (1 - p)^2\{1 - 2\ln(1 - p)\}.$$
Remark Cont’d

- Since
  \[
  \frac{\partial \pi_{(1)^2}(p)}{\partial p} = 4(1 - p)\ln(1 - p) < 0,
  \]
  we see that \(\pi_{(1)^2}(p)\) is a monotonic decreasing function in \(p\).

- Moreover, since \(\pi_{(1)^2}(0) = 1\), we can find a value of \(p\), say \(p_0\), such that \(\pi_{(1)^2}(p) \geq \frac{1}{2}\) for \(0 < p \leq p_0\) and \(\pi_{(1)^2}(p) < \frac{1}{2}\) for \(p_0 < p < 1\).

- In fact, we determine that point to be \(p_0 = 0.5675\), i.e. we have \(X_{1:2}\) to be Pitman-closer (than \(X_{2:2}\)) to \(\xi_p\) whenever \(p > 0.5675\).
<table>
<thead>
<tr>
<th>$\ell$</th>
<th>$p = 0.10$</th>
<th>$p = 0.25$</th>
<th>$p = 0.75$</th>
<th>$p = 0.90$</th>
<th>$p = 0.10$</th>
<th>$p = 0.25$</th>
<th>$p = 0.75$</th>
<th>$p = 0.90$</th>
<th>$p = 0.10$</th>
<th>$p = 0.25$</th>
<th>$p = 0.75$</th>
<th>$p = 0.90$</th>
<th>$p = 0.10$</th>
<th>$p = 0.25$</th>
<th>$p = 0.75$</th>
<th>$p = 0.90$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.05969</td>
<td>0.1197</td>
<td>0.0000</td>
<td>0.1000</td>
<td>0.05969</td>
<td>0.1197</td>
<td>0.0000</td>
<td>0.1000</td>
<td>0.05969</td>
<td>0.1197</td>
<td>0.0000</td>
<td>0.1000</td>
<td>0.05969</td>
<td>0.1197</td>
<td>0.0000</td>
<td>0.1000</td>
</tr>
<tr>
<td>2</td>
<td>0.4031</td>
<td>0.8803</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.4031</td>
<td>0.8803</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.4031</td>
<td>0.8803</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.4031</td>
<td>0.8803</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>3</td>
<td>0.1865</td>
<td>0.7561</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.1865</td>
<td>0.7561</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.1865</td>
<td>0.7561</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.1865</td>
<td>0.7561</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>4</td>
<td>0.0635</td>
<td>0.5660</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.0635</td>
<td>0.5660</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.0635</td>
<td>0.5660</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.0635</td>
<td>0.5660</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>5</td>
<td>0.0158</td>
<td>0.3494</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.0158</td>
<td>0.3494</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.0158</td>
<td>0.3494</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.0158</td>
<td>0.3494</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>6</td>
<td>0.0028</td>
<td>0.1688</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.0028</td>
<td>0.1688</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.0028</td>
<td>0.1688</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.0028</td>
<td>0.1688</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>7</td>
<td>0.0004</td>
<td>0.0605</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.0004</td>
<td>0.0605</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.0004</td>
<td>0.0605</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.0004</td>
<td>0.0605</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>8</td>
<td>0.0000</td>
<td>0.0150</td>
<td>0.9687</td>
<td>0.9998</td>
<td>0.0000</td>
<td>0.0150</td>
<td>0.9687</td>
<td>0.9998</td>
<td>0.0000</td>
<td>0.0150</td>
<td>0.9687</td>
<td>0.9998</td>
<td>0.0000</td>
<td>0.0150</td>
<td>0.9687</td>
<td>0.9998</td>
</tr>
<tr>
<td>9</td>
<td>0.0000</td>
<td>0.0023</td>
<td>0.7584</td>
<td>0.9944</td>
<td>0.0000</td>
<td>0.0023</td>
<td>0.7584</td>
<td>0.9944</td>
<td>0.0000</td>
<td>0.0023</td>
<td>0.7584</td>
<td>0.9944</td>
<td>0.0000</td>
<td>0.0023</td>
<td>0.7584</td>
<td>0.9944</td>
</tr>
<tr>
<td>10</td>
<td>0.0000</td>
<td>0.0022</td>
<td>0.9879</td>
<td>0.8933</td>
<td>0.0000</td>
<td>0.0022</td>
<td>0.9879</td>
<td>0.8933</td>
<td>0.0000</td>
<td>0.0022</td>
<td>0.9879</td>
<td>0.8933</td>
<td>0.0000</td>
<td>0.0022</td>
<td>0.9879</td>
<td>0.8933</td>
</tr>
</tbody>
</table>

Table 3: Closeness probabilities for the exponential for $n=10$.
A Useful Table

<table>
<thead>
<tr>
<th>$n$</th>
<th>$p$</th>
<th>0.10</th>
<th>0.25</th>
<th>0.75</th>
<th>0.90</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>3</td>
<td>8</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>2</td>
<td>4</td>
<td>12</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>2</td>
<td>5</td>
<td>16</td>
<td>19</td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Pitman-closest order statistic to the $p^{th}$ quantile of the standard exponential from a sample of size $n$
Quantities of interest

- Let us consider the power function distribution with pdf and cdf as
  \[ f(x) = \alpha x^{\alpha - 1} \quad \text{and} \quad F(x) = x^\alpha \quad \text{for} \quad x \in (0, 1), \alpha > 0, \]
  and its \( p^{th} \) quantile as \( \xi_p = p^{1/\alpha} \) for \( p \in (0, 1) \).

- Similar to the previous two examples, we can use Result 3 to derive the following expressions for the Pitman closeness probability associated with any two order statistics.

- Varying \( \alpha \) and \( n \), we can then determine the Pitman-closest order statistic to the quantiles of interest.
Results

**Result 6:** For $i = \ell + 1, \ldots, n$,

$$
\pi_{(\ell)}(p) = 1 - l_p(\ell, n - \ell + 1) + k_{\ell,i,n} \sum_{j=0}^{i-\ell-1} (-1)^{i-\ell-1-j} \binom{i - \ell - 1}{j} \frac{1}{n - \ell - j} 
$$

$$
\times \alpha \sum_{a=0}^{j} \sum_{b=0}^{n-\ell-j} (-1)^{a+b} \binom{j}{a} \binom{n - \ell - j}{b} (2\xi_p)^{\alpha(\ell+a+b)} B(\alpha b + 1, \alpha(a + \ell)) 
$$

$$
\times \left\{ 1 - l_{1\frac{1}{2}}(\alpha b + 1, \alpha(a + \ell)) \right\} \text{ for } 0 < p < \frac{1}{2\alpha},
$$

$$
= 1 - l_p(\ell, n - \ell + 1) + k_{\ell,i,n} \sum_{j=0}^{i-\ell-1} (-1)^{i-\ell-1-j} \binom{i - \ell - 1}{j} \frac{1}{n - \ell - j} 
$$

$$
\times \alpha \sum_{a=0}^{j} \sum_{b=0}^{n-\ell-j} (-1)^{a+b} \binom{j}{a} \binom{n - \ell - j}{b} (2\xi_p)^{\alpha(a+\ell+b)} B(\alpha(a + \ell), \alpha b + 1) 
$$

$$
\times \left\{ l_{1\frac{1}{2}}(\alpha(a + \ell), \alpha b + 1) - l_{1\frac{1}{2}}(\alpha(a + \ell), \alpha b + 1) \right\} \text{ for } \frac{1}{2\alpha} \leq p < 1.
$$
Similarly, for \( i = 1, \ldots, \ell - 1 \),

\[
\pi_{(\ell)i}(p) = l_p(\ell, n - \ell + 1) + k_{i,\ell,n} \sum_{j=0}^{\ell-i-1} (-1)^{\ell-i-1-j} \binom{\ell-i-1}{j} \frac{1}{\ell-j-1}
\]

\[
\times \alpha \sum_{a=0}^{n-\ell} (-1)^a \binom{n-\ell}{a} (2\xi_p)^{\alpha(a+\ell)} B(\alpha(j + a + 1), \alpha(\ell - j - 1) + 1)
\]

\[
\times \left\{ 1 - l_{\frac{1}{2\xi_p}}(\alpha(j + a + 1), \alpha(\ell - j - 1) + 1) \right\} \quad \text{for } 0 < p < \frac{1}{2\alpha},
\]

\[
= l_p(\ell, n - \ell + 1) + k_{i,\ell,n} \sum_{j=0}^{\ell-i-1} (-1)^{\ell-i-1-j} \binom{\ell-i-1}{j} \frac{1}{\ell-j-1}
\]

\[
\times \alpha \sum_{a=0}^{n-\ell} (-1)^a \binom{n-\ell}{a} (2\xi_p)^{\alpha(a+\ell)} B(\alpha(j + a + 1), \alpha(\ell - j - 1) + 1)
\]

\[
\times \left\{ l_{\frac{1}{2\xi_p}}(\alpha(j + a + 1), \alpha(\ell - j - 1) + 1) \right\} \quad \text{for } \frac{1}{2\alpha} \leq p < 1.
\]
### A Useful Table

<table>
<thead>
<tr>
<th>$n$</th>
<th>$p$</th>
<th>0.10</th>
<th>0.25</th>
<th>0.75</th>
<th>0.90</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>$\alpha = 0.01$</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>0.03</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>0.07</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>0.10</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>10</td>
<td>$\alpha = 0.01$</td>
<td>1</td>
<td>2</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>0.03</td>
<td>1</td>
<td>2</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>1</td>
<td>2</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>0.07</td>
<td>1</td>
<td>3</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>0.10</td>
<td>1</td>
<td>3</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>1</td>
<td>3</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>15</td>
<td>$\alpha = 0.01$</td>
<td>1</td>
<td>4</td>
<td>11</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>0.03</td>
<td>1</td>
<td>4</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>1</td>
<td>4</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>0.07</td>
<td>1</td>
<td>4</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>0.10</td>
<td>1</td>
<td>4</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>2</td>
<td>4</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>20</td>
<td>$\alpha = 0.01$</td>
<td>2</td>
<td>5</td>
<td>15</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>0.03</td>
<td>2</td>
<td>5</td>
<td>15</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>2</td>
<td>5</td>
<td>15</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>0.07</td>
<td>2</td>
<td>5</td>
<td>15</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>0.10</td>
<td>2</td>
<td>5</td>
<td>15</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>2</td>
<td>5</td>
<td>16</td>
<td>19</td>
</tr>
</tbody>
</table>

**Table 5:** Pitman-closest order statistic to the $p^{th}$ quantile of the power function distribution from a sample of size $n$
Observations

- The extension from population median to population quantiles was straightforward.
- Among all distributions, we observed the following:
  1. We found the same order statistic to be closest to each quantile for the uniform distribution and exponential distribution; this reveals the natural robustness aspect of the Pitman closeness concept in this regard.
  2. In the case of the power function distribution, we started to see changes, and yet they are quite consistent with the corresponding results for the uniform and exponential distributions.
  3. The slight difference in the Pitman-closest order statistics in the case of the power function distribution are seen to occur for extreme values of $\alpha$, viz., $0 < \alpha < 0.25$. 
Is there another way?

- Notice the comparisons in our Pitman closeness probabilities involve two estimators at a time, and consequently one may be required to conduct as many as \( m(m - 1)/2 \) paired comparisons if \( m \) estimators are compared as a class.

- With so many paired comparisons, Blyth [9] advises us to reduce the number of comparisons by considering the joint distributions of their respective loss functions.

- In simultaneous comparisons, consider the following two simultaneous-closeness criteria given by Blyth [9].
Blyth’s Criteria

**Criterion 1**

Choose \( \hat{\theta}_i \) from among estimators in the class \( C \) for which

\[
\max_{i \in \mathcal{K}} \Pr[L_i = \min_{j \in \mathcal{K}} (L_j)],
\]

where \( \mathcal{K} \) is an index set for the class \( C \), and \( L_i = |\hat{\theta}_i - \theta| \). This criterion chooses the estimator within \( C \) which is most-frequently closest to the value of an unknown parameter \( \theta \).

- This simultaneous criterion can be thought of as a *max-min* criterion in that we are maximizing the probability of \( \hat{\theta}_i \) having the smallest loss among all estimators in \( C \).
Blyth’s Criteria

Criterion 2

Choose $\hat{\theta}_i$ from among estimators in the class $\mathcal{C}$ for which

$$\min_{i \in \mathcal{K}} \Pr[\mathcal{L}_i = \max_{j \in \mathcal{K}} (\mathcal{L}_j)].$$

(2)

This criterion directs us to choose the estimator within a class $\mathcal{C}$ which is least-frequently farthest from $\theta$

- This can therefore be thought of as a *min-max* criterion in that we are minimizing the probability that $\hat{\theta}_i$ has the maximum loss among the estimators in $\mathcal{C}$.
- Whenever the size of the index set, $\mathcal{K}$, is two, these criteria are equivalent to the definition of a Pitman-closer estimator.
Other criteria

- Another closely related criterion has been suggested by Banks [8]:
  - For a given $\epsilon > 0$, he suggests that we prefer an estimator, $\delta_1(x)$, where $x$ is a vector of data obtained from a sample, over a competing estimator, $\delta_2(x)$, if

  $\Pr[\| \delta_1(x) - \theta \| < \epsilon] > \Pr[\| \delta_2(x) - \theta \| < \epsilon]$ \hspace{1cm} (3)

  $\forall \theta \in \Theta$, where $\| x \|$ is a loss function.

- Banks’ criterion can be generalized to the simultaneous comparison of multiple estimators of $\theta$ as well.

- The list goes on but it is Blythe’s criteria which we will find most useful here.
For the computations required in Blyth’s criteria, we turn to the work of Fountain et al. [10].

While the geometry of their arguments is complex, we consider a special case in which estimators are ordered, as in the case of order statistics.

Combining the concepts of Pitman closeness and simultaneous comparisons, we have a new goal and with it, introduce a new concept and some subsequent some new results.
Aim and Method

Objective

We would like to determine the probability with which each order statistic, $X_{i:n}$ for each $i \in \{1, 2, \cdots, n\}$, is *simultaneously Pitman-closest* to $\theta$ when compared with the remaining order statistics.

- Our procedure for calculating our probabilities of interest is based on partitioning the essential range of a random vector of observations into regions in which each order statistic is the “best”.
In accordance with Blyth’s first criterion, we have the following definition for the simultaneous-closeness probability.

**Definition**

The *simultaneous-closeness probability (SCP)* of $X_{i:n}$, $i \in \{1, \cdots, n\}$, among the order statistics, $X_{1:n}, \cdots, X_{n:n}$, in the estimation of a population parameter $\theta$ is

$$
\pi_{i:n}(\theta) = \Pr \left\{ |X_{i:n} - \theta| < \min_{j, j \neq i} |X_{j:n} - \theta| \right\}. \quad (4)
$$
A Useful Result

- Using the geometric arguments of by Fountain et al. [10], this probability can be found by carrying out just two comparisons:

\[
\pi_{i:n}(\theta) = P(X_{i:n}, X_{i-1:n}|\theta, n) - P(X_{i+1:n}, X_{i:n}|\theta, n),
\]

(5)

where \( P(X_{i:n}, X_{i-1:n}|\theta, n) = Pr(|X_{i:n} - \theta| < |X_{i-1:n} - \theta|) \)

- We can define the probability of interest in the case of both the unbounded and bounded support.

- Just as we looked at applications of Pitman closeness probabilities in the previous section, we can do the same here.
Settings

- The results to follow are so determined based on the support of the random variable $X$ (i.e., whether the support is the entire real line or a bounded open interval).
- In particular, we are interested in parameters with a parameter space that coincides with the support of $X$.
- We first present general results for random variables with complete support (i.e., $\mathbb{R}$) and then present similar ones for random variables with bounded support.
Unbounded Case

**Theorem**

Assume that the common conditions hold for $X$, with support $\mathbb{R}$. Then for $i = 2, \cdots, n - 1$, the simultaneous-closeness probability, $\pi_{i:n}(\theta)$, of $X_{i:n}$ to $\theta$ is given by

$$
\pi_{i:n}(\theta) = \binom{n}{i-1} [F(\theta)]^{i-1} [1 - F(\theta)]^{n-i+1} \\
+ \int_0^{F(\theta)} \frac{n!}{(i-1)!(n-i)!} \left[ \bar{F}(2\theta - F^{-1}(u)) \right]^{n-i} u^{i-1} du \\
- \int_0^{F(\theta)} \frac{n!}{(i-2)!(n-i+1)!} \left[ \bar{F}(2\theta - F^{-1}(u)) \right]^{n-i+1} u^{i-2} du.
$$

(6)
Bounded Case

Theorem

Assume the common conditions for $X$, where the support of $X$ is bounded on the interval $(a, b)$. Then, the simultaneous-closeness probability $\pi_{i:n}(\theta)$ is given by

$$
\pi_{i:n}(\theta) = n \binom{n - 1}{i - 2} \int_a^{b^*} f(x) [F(x)]^{i-2} \left\{ [F(h_2(x))]^{n-i+1} - [F(h_1(x))]^{n-i+1} \right\} dx,
$$

where $b^* = \min(b, 2\theta - a)$, $h_1(x) = \max(a, x)$ and $h_2(x) = \min(b, 2\theta - x)$. 


Some Results

Population Quantiles

- By taking $\theta = \xi_p$, we can derive explicit expression for the SCP probabilities as seen in Eqs. (6) and (7).

- First, for the unbounded case, assuming that the common conditions hold for $X$, with support $\mathbb{R}$. Then for $i = 2, \cdots, n - 1$,

$$
\pi_{i:n}(p) = \binom{n}{i - 1} p^{i-1} (1 - p)^{n-i+1} \\
+ \frac{n!}{(i - 1)! (n - i)!} \int_0^p \left\{ 1 - F \left[ 2F^{-1}(p) - F^{-1}(u) \right] \right\}^{n-i} u^{i-1} du \\
- \frac{n!}{(i - 2)! (n - i + 1)!} \int_0^p \left\{ 1 - F \left[ 2F^{-1}(p) - F^{-1}(u) \right] \right\}^{n-i+1} u^{i-2} du.
$$
Some Notes

- Noting the previous expression is free of $\mu$ and $\sigma$, this suggests that the probabilities can be computed and cataloged for various choices of $p$ and $n$ for different families such as Normal, extreme-value, Cauchy, Laplace, logistic, and so on.

- These simultaneous-closeness probabilities may then be used in developing efficient goodness-of-fit methods and also as “good” plotting points in graphical model validity methods.
A Useful Observation

We have the following useful corollary which simplifies the computation in the case when the standard distribution is symmetric about the origin.

**Corollary**

If the standard pdf is symmetric about the origin (i.e., \( f(z) = f(-z) \), \( F(-z) = 1 - F(z) \) and \( F^{-1}(p) = -F^{-1}(q) \), where \( q = 1 - p \)), then we have

\[
\pi_{i:n}(p) = \pi_{n-i+1:n}(q) \tag{9}
\]

for \( i = 1, \cdots, n \).
Bounded Case

- Similarly, we have developed a general expression for $\pi_{i:n}(\theta)$ when the support of $X$ is bounded on the interval $(a, b)$.

- If $X$ is bounded on the interval $(a, b)$, then $Z = (X - \mu)/\sigma$ is bounded on the interval $(a', b')$, where $b' = (b - \mu)/\sigma$ and $a' = (a - \mu)/\sigma$. Let $z_p = (\xi_p - \mu)/\sigma$.

- The simultaneous-closeness probability of $X_{i:n}$ to $\xi_p$ in this case is given by

$$\pi_{i:n}(p) = n \binom{n - 1}{i - 2} \int_{a'}^{b^{**}} f(z) [F(z)]^{i-2}$$

$$\times \left\{ [F(h_2(z))]^{n-i+1} - [F(h_1(z))]^{n-i+1} \right\} dz,$$

where $b^{**} = \min(b', 2z_p - a')$, $h_1(z) = \max(a', z)$ and $h_2(z) = \min(b', 2z_p - z)$. 
Two Cases

- Being the forementioned results are predicated on the support of the random variable and this is further separated into two categories (bounded and unbounded), we can look at applications in these two cases.

- Here we consider applying the unbounded results to the normal distribution and the bounded results to the exponential distribution.
SCP for Normal

- In the Normal distribution, the standard forms are

\[ f(z) = \phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right), \quad F(z) = \Phi(z) \quad \text{and} \quad F^{-1}(p) = \Phi^{-1}(p). \]

- When we substitute these expressions into Eq. (8) and use the fact that \( \Phi(-z) = 1 - \Phi(z) \), we obtain

\[
\pi_{i:n}(p) = \binom{n}{i-1} p^{i-1} (1 - p)^{n-i+1} \\
+ \frac{n!}{(i-1)!(n-i)!} \int_0^p \{ \Phi \left[ \Phi^{-1}(u) - 2\Phi^{-1}(p) \right] \}^{n-i} u^{i-1} du \\
- \frac{n!}{(i-2)!(n-i+1)!} \int_0^p \{ \Phi \left[ \Phi^{-1}(u) - 2\Phi^{-1}(p) \right] \}^{n-i+1} u^{i-2} du.
\]
We have tabulated these probabilities for values of $i = 1, \ldots, 10$, $p = 0.05(0.05)0.50$ and $n = 10$ in Table 6.

Then, the values for $p = 0.55(0.05)0.95$ can be found readily by using the symmetry property in the corollary just mentioned.
SCP of Order Statistics to Population Quantiles

Applications

SCP for Normal

Table 6: SCP of order statistics for Normal distribution when $n = 10$ for various values of $p$ and $i$

<table>
<thead>
<tr>
<th>$p \setminus i$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>0.78921</td>
<td>0.18365</td>
<td>0.02478</td>
<td>0.00222</td>
<td>0.00014</td>
<td>0.00001</td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.00000</td>
</tr>
<tr>
<td>0.10</td>
<td>0.52231</td>
<td>0.34097</td>
<td>0.11082</td>
<td>0.02252</td>
<td>0.00308</td>
<td>0.00029</td>
<td>0.00026</td>
<td>0.00026</td>
<td>0.00026</td>
<td>0.00026</td>
</tr>
<tr>
<td>0.15</td>
<td>0.31927</td>
<td>0.37855</td>
<td>0.21151</td>
<td>0.07179</td>
<td>0.01613</td>
<td>0.00247</td>
<td>0.00153</td>
<td>0.00153</td>
<td>0.00153</td>
<td>0.00153</td>
</tr>
<tr>
<td>0.20</td>
<td>0.18403</td>
<td>0.33583</td>
<td>0.28126</td>
<td>0.14078</td>
<td>0.04616</td>
<td>0.01024</td>
<td>0.00029</td>
<td>0.00029</td>
<td>0.00029</td>
<td>0.00029</td>
</tr>
<tr>
<td>0.25</td>
<td>0.10058</td>
<td>0.25937</td>
<td>0.30255</td>
<td>0.20883</td>
<td>0.09375</td>
<td>0.02833</td>
<td>0.00765</td>
<td>0.00765</td>
<td>0.00765</td>
<td>0.00765</td>
</tr>
<tr>
<td>0.30</td>
<td>0.05210</td>
<td>0.18049</td>
<td>0.28047</td>
<td>0.25626</td>
<td>0.15155</td>
<td>0.06011</td>
<td>0.01598</td>
<td>0.01598</td>
<td>0.01598</td>
<td>0.01598</td>
</tr>
<tr>
<td>0.35</td>
<td>0.02548</td>
<td>0.11483</td>
<td>0.23100</td>
<td>0.27212</td>
<td>0.20682</td>
<td>0.10513</td>
<td>0.03573</td>
<td>0.03573</td>
<td>0.03573</td>
<td>0.03573</td>
</tr>
<tr>
<td>0.40</td>
<td>0.01169</td>
<td>0.06712</td>
<td>0.17152</td>
<td>0.25607</td>
<td>0.24613</td>
<td>0.15794</td>
<td>0.06765</td>
<td>0.06765</td>
<td>0.06765</td>
<td>0.06765</td>
</tr>
<tr>
<td>0.45</td>
<td>0.00499</td>
<td>0.03599</td>
<td>0.11543</td>
<td>0.21606</td>
<td>0.26008</td>
<td>0.20878</td>
<td>0.11177</td>
<td>0.03848</td>
<td>0.03848</td>
<td>0.03848</td>
</tr>
<tr>
<td>0.50</td>
<td>0.00195</td>
<td>0.01758</td>
<td>0.07031</td>
<td>0.16406</td>
<td>0.24609</td>
<td>0.24609</td>
<td>0.16406</td>
<td>0.07031</td>
<td>0.01758</td>
<td>0.00195</td>
</tr>
</tbody>
</table>

Table 6 presents a clear picture of the most-preferred order statistic for the $p^{th}$ quantile and we do observe from this table that the most-preferred order statistic, with index $i^*$, increases from 1 to 10 as $p$ increases from 0.05 to 0.95.
In the case of the standard exponential distribution, we have $a' = 0$ and $b' = \infty$, and the standard forms, for $z > 0$, in this case are

$$f(z) = e^{-z}, \quad F(z) = 1 - e^{-z} \quad \text{and} \quad F^{-1}(z) = -\ln(1 - z).$$

We can use expressions from Balakrishnan et al. [2] to express the simultaneous probabilities for $\xi_p$ as

$$\pi_{i:n}(p) = P(X_{i:n}, X_{i-1:n}\mid p, n) + P(X_{i:n}, X_{i+1:n}\mid p, n) - 1 = \pi_{(i)\mid i-1}(p) + \pi_{(i)\mid i+1}(p) - 1,$$

where $\pi_{(\ell),i}(p)$ is given in Result 7 (in Eq. (23) for $i = \ell + 1, \ldots, n$ and Eq. (25) for $i = 1, \ldots, \ell - 1$) of Balakrishnan et al. [2].
The combined expression for the simultaneous-closeness probabilities in the exponential distribution can be simplified using the following integral:

\[ I_{\ell,m}(p) = \int_0^p \frac{u^{\ell-1}}{(1-u)^m} \, du \quad \text{for} \quad \ell, m = 1, 2, \ldots . \tag{12} \]

With the integral in Eq. (12), \( \pi_{i:n}(p) \) can be expressed in this case as

\[
\pi_{i:n}(p) = \binom{n}{i-1} p^{i-1}(1-p)^{n-i+1} + a_{i:n} q^{2(n-i)} I_{i,n-i}(p) - b_{i:n} q^{2(n-i+1)} I_{i-1,n-i+1}(p),
\]

where \( a_{i:n} = \frac{n!}{(i-1)!(n-i)!} \) and \( b_{i:n} = \frac{n!}{(i-2)!(n-i+1)!} \).

We have tabulated these values for \( n = 10 \) and \( i = 1, \ldots , 10 \) for incremental values of \( p = 0.05(0.05)0.95 \) in Table 7.
Table 7: SCP of order statistics for exponential distribution when \( n=10 \) for various values of \( p \) and \( i \)
Observations

- Quite interestingly, we find the simultaneous-closest order statistic to the $p^{th}$ quantile in this case to be quite close to that of the Normal case and the corresponding simultaneous-closeness probabilities to be quite close as well; this reveals that the concept discussed here is naturally quite robust.

- In Table 7, we observe the lack of symmetry in the simultaneous-closeness probabilities since the standard exponential is skewed to the right.

- In addition, while the probabilities $\pi_{i:n}(p)$ and $\pi_{n-i+1:n}(q)$ are not equal, the range of quantiles over which an order statistic is preferred varies, although slightly as mentioned.

- For example, $X_{1:10}$ is most frequently closest to $\xi_p$, for $p = 0$ to $p = 0.15^+$, whereas $X_{10:10}$ is most frequently closest to $\xi_p$ after $p = 0.85$. 
Plotting points of order statistics are often used in the determination of goodness-of-fit of observed data to theoretical percentiles.

In the literature, there has been controversy on the choice of plotting points.

Plotting points are often determined using nonparametric methods which produce, for example, the mean- and median-ranks commonly used in practice.

To motivate the applicability of SCPs in the plotting points problem, consider the following diagram:
Figure 2: SCP for Normal order statistics when $n=10$ and $i=1,\cdots,10$. 
In Figure 2, for a given $i$, we can see the value of $p$ for which the SCP is maximized.

This leads to identification of optimal plotting points based on Pitman simultaneous-closeness probabilities.

That is, we use a distribution-based approach which selects plotting points (quantiles) based on the simultaneous-closeness of order statistics to population quantiles.
Other Plotting Points

- The concept of mean-rank, denoted by \( e_{i:n} \), is based on the fact that

\[
\mathbb{E} [F (X_{i:n})] = \frac{i}{n+1} = e_{i:n},
\]

since \( F(X_{i:n}) \overset{d}{=} U_{i:n} \), the \( i \)-th order statistic in a sample of size \( n \) from the uniform \( \mathcal{U}(0,1) \) distribution.

- In an alternative approach, letting \( M(X) \) denote the median of \( X \), one could use the median-rank, \( m_{i:n} \) of the \( i \)-th order statistic given by

\[
M [F (X_{i:n})] = b_{0.5;i,n-i+1} = m_{i:n},
\]

where \( b_{0.5;\alpha,\beta} \) is the median of a \( \mathcal{B}(\alpha, \beta) \) due to (1) and one can establish that

\[
e_{i:n} < m_{i:n} \quad \forall \quad i < \frac{n}{2} \quad \text{and} \quad e_{i:n} > m_{i:n} \quad \forall \quad i > \frac{n}{2}.
\]
A Reminder

- Recall the result for the simultaneous-closeness probability:

**Result 1** For $i = 2, \cdots, n - 1$,

\[
\pi_{i:n}(p) = \binom{n}{i-1} p^{i-1}(1-p)^{n-i+1} \\
+ \frac{n!}{(i-1)!(n-i)!} \int_0^p \{1 - F \left[2F^{-1}(p) - F^{-1}(u)\right]\}^{n-i} u^{i-1} du \\
- \frac{n!}{(i-2)!(n-i+1)!} \int_0^p \{1 - F \left[2F^{-1}(p) - F^{-1}(u)\right]\}^{n-i+1} u^{i-2} du
\]
A natural question that arises here is as follows.

For given values of $i$ and $n$ and the choice of the standard normal distribution, for example, for what value $p$ is the SCP $\pi_{i:n}(p)$ in (14) maximized.

Such a determination, as mentioned, would give an “optimal” plotting position corresponding to $X_{i:n}$ while assessing the adequacy of the fit of a normal distribution to the observed data.
Determining where the probability is maximized would require us to take the partial derivative of $\pi_{i:n}(p)$ in (14) with respect to $p$ and equate it to zero.

Now, setting $\frac{\partial \pi_{i:n}}{\partial p}$ equal to zero, the terms without integrals cancel and we obtain an objective equation.
The equation to be solved is:

\[
(n - i) \int_0^p \left\{ F \left[ F^{-1}(u) - 2F^{-1}(p) \right] \right\}^{n-I-1} f \left( F^{-1}(u) - 2F^{-1}(p) \right) u^{i-1} du
\]

\[
= (i - 1) \int_0^p \left\{ F \left[ F^{-1}(u) - 2F^{-1}(p) \right] \right\}^{n-I} f \left( F^{-1}(u) - 2F^{-1}(p) \right) u^{i-2} du,
\]

which is equivalent to

\[
\int_0^p \left\{ F \left[ F^{-1}(u) - 2F^{-1}(p) \right] \right\}^{n-I-1} f \left( F^{-1}(u) - 2F^{-1}(p) \right) u^{i-2}
\]

\[
\times \left\{ (n - i)u - (i - 1)F \left[ F^{-1}(u) - 2F^{-1}(p) \right] \right\} du = 0.
\]

We shall refer to the solution to this equation as SCP plotting point.
Many Cases

- For a specific $F(z)$ and $f(z)$, such as the standard normal distribution, we can solve Eq. (15) numerically to determine the value of $p$ that maximizes the SCP $\pi_{i:n}(p)$.
- Naturally, we would expect the plotting points derived by maximization of $\pi_{i:n}(p)$ with respect to $p$ not to differ markedly from plotting points based on mean- or median-ranks.
- We show you the results for the normal case and several others.
In the case of normal distribution, we need to solve the equation

\[
\int_0^p \left\{ \Phi \left[ \Phi^{-1}(u) - 2\Phi^{-1}(p) \right] \right\}^{n-i-1} \phi \left( \Phi^{-1}(u) - 2\Phi^{-1}(p) \right) u^{i-2} \\
\times \left\{ (n-i)u - (i-1)\Phi \left[ \Phi^{-1}(u) - 2\Phi^{-1}(p) \right] \right\} \, du = 0, \quad (16)
\]

where \( \phi(z) \) and \( \Phi(z) \) are the standard normal pdf and cdf, respectively.
Some Results

- We determine the SCP for values of \( p \) ranging over 0.001(0.001)0.999 for all values of \( i \) corresponding to different choices of \( n \).
- Denoting the solution of (16) by \( s_{i:n} \) for each \( i = 1, \ldots, n \), we first note that \( s_{1:n} \equiv 0 \) and \( s_{n:n} \equiv 1 \), as one would expect logically.
- In Table 8, we present, for \( n = 10 \) and \( i = 2(1)9 \), the values of \( s_{i:n} \) and the median-and mean-ranks.
Table 8: Plotting points for normal distribution when $n=10$

<table>
<thead>
<tr>
<th>$i$</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_{i:n}$</td>
<td>0.1433</td>
<td>0.2472</td>
<td>0.3487</td>
<td>0.4496</td>
<td>0.5504</td>
<td>0.6513</td>
<td>0.7528</td>
<td>0.8567</td>
</tr>
<tr>
<td>$m_{i:n}$</td>
<td>0.1623</td>
<td>0.2586</td>
<td>0.3551</td>
<td>0.4517</td>
<td>0.5483</td>
<td>0.6449</td>
<td>0.7414</td>
<td>0.8377</td>
</tr>
<tr>
<td>$e_{i:n}$</td>
<td>0.1818</td>
<td>0.2727</td>
<td>0.3636</td>
<td>0.4545</td>
<td>0.5455</td>
<td>0.6364</td>
<td>0.7273</td>
<td>0.8182</td>
</tr>
</tbody>
</table>
Similar results for the SCP plotting positions for the normal distribution when $n = 15$ are presented in Figure 3 and Table 9.

<table>
<thead>
<tr>
<th></th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_i:n$</td>
<td>0.0951</td>
<td>0.1644</td>
<td>0.2320</td>
<td>0.2992</td>
<td>0.3662</td>
<td>0.4331</td>
<td>0.5000</td>
</tr>
<tr>
<td>$m_i:n$</td>
<td>0.1094</td>
<td>0.1743</td>
<td>0.2394</td>
<td>0.3045</td>
<td>0.3697</td>
<td>0.4348</td>
<td>0.5000</td>
</tr>
<tr>
<td>$e_i:n$</td>
<td>0.1250</td>
<td>0.1875</td>
<td>0.2500</td>
<td>0.3125</td>
<td>0.3750</td>
<td>0.4375</td>
<td>0.5000</td>
</tr>
</tbody>
</table>

Table 9: Plotting points for normal distribution when $n=15$
Figure 3: SCP for normal order statistics when \( n=15 \) and \( i=1, \ldots, 15 \).
Other distributions

- As mentioned earlier, the SCP plotting points can also be determined for other distributions.
- For example, we can carry out the necessary computations for the logistic, Laplace and Cauchy distributions.
- The corresponding results for $n = 10$ and $n = 15$ are presented in Tables 10 and 11, respectively, for the aforementioned distributions, alongside the normal SCP plotting points and mean- and median-ranks.
- From these tables, we observe that the plotting points are nearly the same for the inner order statistics, while those for extreme order statistics vary.
Some More Results

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cauchy</td>
<td>0.1281</td>
<td>0.2375</td>
<td>0.3433</td>
<td>0.4479</td>
<td>0.5521</td>
<td>0.6567</td>
<td>0.7625</td>
<td>0.8819</td>
<td></td>
</tr>
<tr>
<td>Laplace</td>
<td>0.1381</td>
<td>0.2425</td>
<td>0.3443</td>
<td>0.4465</td>
<td>0.5535</td>
<td>0.6557</td>
<td>0.7575</td>
<td>0.8619</td>
<td></td>
</tr>
<tr>
<td>logistic</td>
<td>0.1405</td>
<td>0.2456</td>
<td>0.3479</td>
<td>0.4494</td>
<td>0.5506</td>
<td>0.6521</td>
<td>0.7544</td>
<td>0.8595</td>
<td></td>
</tr>
<tr>
<td>normal</td>
<td>0.1433</td>
<td>0.2472</td>
<td>0.3487</td>
<td>0.4496</td>
<td>0.5504</td>
<td>0.6513</td>
<td>0.7528</td>
<td>0.8567</td>
<td></td>
</tr>
<tr>
<td>uniform</td>
<td>0.1584</td>
<td>0.2542</td>
<td>0.3521</td>
<td>0.4506</td>
<td>0.5494</td>
<td>0.6479</td>
<td>0.7458</td>
<td>0.8416</td>
<td></td>
</tr>
<tr>
<td>m</td>
<td>0.1623</td>
<td>0.2586</td>
<td>0.3551</td>
<td>0.4517</td>
<td>0.5483</td>
<td>0.6449</td>
<td>0.7414</td>
<td>0.8377</td>
<td></td>
</tr>
<tr>
<td>e</td>
<td>0.1818</td>
<td>0.2727</td>
<td>0.3636</td>
<td>0.4545</td>
<td>0.5455</td>
<td>0.6364</td>
<td>0.7273</td>
<td>0.8182</td>
<td></td>
</tr>
</tbody>
</table>

Table 10: SCP plotting points for some symmetric families of distributions when $n=10$
Table 11: SCP plotting points for some symmetric families of distributions when \( n=15 \)

<table>
<thead>
<tr>
<th>( i )</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cauchy</td>
<td>0.0849</td>
<td>0.1572</td>
<td>0.2269</td>
<td>0.2957</td>
<td>0.3640</td>
<td>0.4320</td>
<td>0.5000</td>
</tr>
<tr>
<td>Laplace</td>
<td>0.0922</td>
<td>0.1620</td>
<td>0.2299</td>
<td>0.2971</td>
<td>0.3642</td>
<td>0.4313*</td>
<td>0.5000</td>
</tr>
<tr>
<td>logistic</td>
<td>0.0932</td>
<td>0.1632</td>
<td>0.2313</td>
<td>0.2987</td>
<td>0.3659</td>
<td>0.4330</td>
<td>0.5000</td>
</tr>
<tr>
<td>normal</td>
<td>0.0951</td>
<td>0.1644</td>
<td>0.2320</td>
<td>0.2992</td>
<td>0.3662</td>
<td>0.4331</td>
<td>0.5000</td>
</tr>
<tr>
<td>uniform</td>
<td>0.1058</td>
<td>0.1699</td>
<td>0.2353</td>
<td>0.3012</td>
<td>0.3674</td>
<td>0.4337</td>
<td>0.5000</td>
</tr>
<tr>
<td>( m )</td>
<td>0.1094</td>
<td>0.1743</td>
<td>0.2394</td>
<td>0.3045</td>
<td>0.3697</td>
<td>0.4348</td>
<td>0.5000</td>
</tr>
<tr>
<td>( e )</td>
<td>0.1250</td>
<td>0.1875</td>
<td>0.2500</td>
<td>0.3125</td>
<td>0.3750</td>
<td>0.4375</td>
<td>0.5000</td>
</tr>
</tbody>
</table>
Observations

- Looking at the results for the various symmetric distributions just considered, we observe that the SCP plotting points for the inner order statistics were close and consistent across a spectrum of distributions and that the greatest disparities occurred for the extreme order statistics at each end of the sample.

- The natural question then is how robust is this method to a sequence of distributions with tails that are progressively heavier?
What did we find?

- To answer this, one may consider the family of $t$-distributions with $f$ degrees of freedom.
- We observe that the values for interior order statistics are very nearly the same for all $f$, while those for the extreme order statistics vary a bit, thus displaying the inherent robustness of these plotting points.
- In addition, we also observe the behavior to be monotone with respect to $f$. 
The concept of Pitman Closeness has much history and continues to prove to be a useful tool in many contexts.

We have looked at its use in comparing order statistics as estimators and defined a new concept, that of simultaneous closeness.

The concept of simultaneous-closeness builds on previous ideas and yet finds a new use in goodness-of-fit testing in the form of plotting points.
What followed?

- Extending the work done on plotting points, a correlation-type test statistic was developed which uses SCP plotting points; work has been done for goodness-of-fit to Normal and Weibull.
- Current work is being carried out for Pitman closeness of order statistics to population quantiles based on progressively Type-II right censored samples.
I would just like to say thank you to...

**DR. BALAKRISHNAN FOR “INVITING” ME**

**AND,**

**WILLIAM VOLTERMAN FOR LISTENING TO MOST OF THIS MATERIAL FOR THE 3rd TIME**

**AND,**

**TO ALL OF YOU FOR YOUR ATTENTION!**


