PSYCHOMETRIKA—VOL. 70, NO. 4, 759–772
DECEMBER 2005
DOI: 10.1007/s11336-001-0916-1

VARIABILITY OF THE MAX AND MIN STATISTIC: A THEORY OF THE QUANTILE SPREAD AS A FUNCTION OF SAMPLE SIZE

JAMES T. TOWNSEND

INDIANA UNIVERSITY

HANS COLONIUS

UNIVERSITAET OLDENBURG

The maximum and minimum of a sample from a probability distribution are extremely important random variables in many areas of psychological theory, methodology, and statistics. For instance, the behavior of the mean of the maximum or minimum processing time, as a function of the number of component random processing times (n), has been studied extensively in an effort to identify the underlying processing architecture (e.g., Townsend & Ashby, 1983; Colonius & Vorberg, 1994). Little is known concerning how measures of variability of the maximum or minimum change with n. Here, a new measure of random variability, the quantile spread, is introduced, which possesses sufficient strength to define distributional orderings and derive a number of results concerning variability of the maximum and the minimum statistics. The quantile spread ordering may be useful in many venues. Several interesting open problems are pointed out.

Key words: Extremal distributions; Extremal variability; Quantile spread; Parallel processing.

1. Introduction

Consider a psychological response time experiment where two processes are in series and both must be completed before the response (or the next stage) can be made. The total random time is given as the sum of two random variables. Whether or not the component random variables are independent, the mean and variance of the sum can easily be derived from the well-known laws of probability and statistics (e.g., Luce, 1986; Townsend & Ashby, 1983). Of course, laws of sums of independent random variables are among the most studied of any in probability and statistics.

On the other hand, the situation when the two processes function in parallel, and the response waits for the longer of the two, is considerably more nebulous. When the component processes are statistically independent and under relatively weak conditions, it can be shown that the mean total (maximum) completion time increases in a negatively accelerated fashion as the number of components, n, increases (e.g., Townsend & Ashby, 1983, pp. 92–93). The fastest increasing of all such functions is itself also negatively accelerated (see Hartley & David, 1951; Sternberg, 1966). For a general approach to distribution effects on the maximum random variable, see Colonius and Vorberg (1994) and with regard to the minimum, see Colonius (1990) and Townsend and Nozawa (1995, 1997).[1] There are many other situations in the social and biological sciences where information about extrema can be of considerable importance.

This work was supported by an NIH Grant R01 MH57717 to the first author. Some of the collaboration took place during the year 2000 while J.T. Townsend was a Fellow at the Hanse Institute for Advanced Study (HWK), sponsored by H. Colonius at Oldenburg University.

Requests for reprints should be sent to James T. Townsend, Department of Psychology, Indiana University, 1101 E. 10th St. Bloomington, IN 47405-7007. E-mail: jtownsen@indiana.edu.

[1]For a recent investigation using variance to identify architecture see Donnelley, Found, & Müller (1999) and a related commentary on this topic by Townsend (2001).

© 2005 The Psychometric Society

However, there is a paucity of laws concerning the behavior of the variance, or any measure of variability, of comparable generality to the serial case (although Townsend (1990b) has shown that in the case of symmetric random variables, the variance for the maximum always decreases from $n = 1$ to 2). Indeed, in his well-known work on extreme value distributions, Gumbel (1958) remarks that the variance can increase, decrease, or act nonmonotonically as n increases depending on the exact parent distributions (i.e., the distribution of a single component). Similarly, the minimum time statistic has also been of considerable interest in recent years, also in, for instance, response time methodology (e.g., Bernstein, 1970; Colonius, 1990; Egeth, 1966; Miller, 1982; Mordkoff & Yantis, 1991; Townsend & Nozawa, 1995). The minimum time random variable, viewed as representing the winner of a "horse race" of parallel channels or multi-item processing can also be compared with the standard serial case where the minimum time distribution is simply that which describes processing of any single (and arbitrary, since they are all presumed to be the same) item or channel.

In addition to Gumbel (1958), another pioneering effort in related directions is Gnedenko (1943), and Harter (1978) offered a thorough bibliography of the literature on extreme values up to that time. Moriguti (1951) explored the situation in symmetric distributions and located the upper and lower bounds for the expectation, coefficient of variation, and the variance of the maximum statistic. The domain of application of his approach was broadened with a refined version of the Schwarz inequality provided in a subsequent paper (Moriguti, 1953). Other notable contributions of late to the general literature in this domain have been Munoz-Perez (1990), Belzunce (1999), and Kleiber (2002). Since the present study is not a review paper, we are not able to outline these other studies in detail.

Powerful methods for investigating variability and other distributional properties of order statistics have evolved in recent years (e.g., Arnold, 1987; Shaked & Shantikumar, 1994), and the general literature on extrema is now enormous. Nevertheless, there is still relatively little known concerning the conditions under which measures of variability show an increase or decrease as the maximum (or minimum) is taken over increasingly larger n.

As a measure of variability, the second central moment or variance statistic is ubiquitously employed in experimentation, methodology, and theory in the behavioral sciences, perhaps even more than the mean as the first moment is used as a measure of central tendency. However, there are compelling reasons to consider other measures of variability. Measures that are more tightly related to properties of the parent distribution can be of value both in theoretical ventures as well as in experimental comparisons (Dzhafarov (1992), e.g., theories of nonparametric statistics; and Townsend (1990a,b), notwithstanding the close linkage of variance to the theory of parametric statistics. In this paper, we introduce a measure of variability that has apparently not yet been studied in the literature. Because of its close relationship to distributional properties, it permits us to derive a variability ordering between random variables and to study some of its properties. It is also a generalization of the well-known range statistic as will be pointed out later.

2. The Quantile Spread

As suggested above, one of the possible difficulties in obtaining general characteristics for variances of extreme values may be the coarseness, relative to the fine-grain aspects of a probability distribution, of the functional map of the distribution into the variance statistic. An alternative approach is to focus on the concept of *quantile spread,* which we now define.

Definition 1. Suppose F is the absolutely continuous[2] distribution function (df) for some random variable X. Then the *quantile spread of F* is the function

$$QS_F(p) = F^{-1}(p) - F^{-1}(1-p) \tag{1}$$

defined for $0.5 < p < 1$. Equivalently, if dependence on the random variable is to be emphasized, we sometimes write $QS_X(p) = X(p) - X(1-p)$ where $X(p)$ denotes the pth quantile of F.

Of course, it is obvious that the ordinary interquantile range, expressed as the difference between the 75th and 25th percentiles, is a special case of quantile spread.

Sometimes it is more convenient to consider the survivor function (sf) S, $S(x) \equiv 1 - F(x)$ rather than the distribution function. Due to $F^{-1}(p) = S^{-1}(1-p)$, the quantile spread can also be written as

$$QS_S(p) = S^{-1}(1-p) - S^{-1}(p). \tag{2}$$

Example 1. (Weibull Distribution (e.g., Warburton, 1967)). Let $F(x) = 1 - \exp(-\beta x^\alpha)$, with $\alpha, \beta > 0$, be the cumulative distribution of some random variable X. For $F(x) = p$, it follows from simple algebra that

$$F^{-1}(p) = \left[\beta^{-1} \ln \frac{1}{1-p}\right]^{1/\alpha}. \tag{3}$$

This gives the quantile spread for the Weibull distribution

$$\begin{aligned} QS_X(p) &= F^{-1}(p) - F^{-1}(1-p) \\ &= \beta^{-1/\alpha}\left[\left(\ln \frac{1}{1-p}\right)^{1/\alpha} - \left(\ln \frac{1}{p}\right)^{1/\alpha}\right]. \end{aligned} \tag{4}$$

For $\alpha = 1$ the *exponential* distribution is obtained as a special case of the Weibull distribution with the corresponding exponential quantile spread

$$QS_X(p) = \beta^{-1} \ln \frac{p}{1-p}. \tag{5}$$

For two exponentially distributed random variables X and Y with parameters β_X and β_Y, respectively, $\beta_X < \beta_Y$ immediately implies $QS_X(p) > QS_Y(p)$ for all $p \in (0.5, 1)$. This ordering of the quantile spreads coincides with the fact that $\mathrm{Var}[X] = \beta_X^{-2} > \beta_Y^{-2} = \mathrm{Var}[Y]$. Note, however, that in general the quantile spread, as a statistic of variability, acts at a considerably finer grain level than the variance and hence is more sensitive to the form of the distribution. As a further example let us consider a symmetric (around 0) distribution, i.e., such that $F^{-1}(p) = -F^{-1}(1-p)$.

Example 2. (Logistic Distribution). The inverse of the cumulative distribution

$$F(x) = \frac{1}{1 + \exp[-x]}, \qquad -\infty < x < +\infty,$$

is

$$F^{-1}(p) = \ln \frac{p}{1-p},$$

[2]This assumption is made mainly for simplicity. Many results developed here could be stated more generally for right continuous distribution functions, using the right continuous inverse defined by $F^{-1}(p) \equiv \{\sup x : F(x) \leq p, p \in [0, 1]\}$.

yielding the logistic quantile spread

$$QS(p) = 2F^{-1}(p). \tag{6}$$

Comparing (6) with (5) reveals that, not surprisingly, the quantile spread function does not uniquely characterize the distribution function from which it is generated: the quantile spread for the logistic distribution is identical to the quantile spread of an exponential distribution with $\beta = 1/2$. In section 4 below we study the behavior of the quantile spread for extremal statistics as a function of sample size for a random sample from populations with various parent distributions F.

3. The Quantile Spread Order

The ordering of the quantile spreads for exponentially distributed random variables, with different parameters β_X and β_Y studied in the example above, suggests a more powerful definition of ordering random variables with respect to their quantile spreads.

Definition 2. Let X and Y be two random variables with absolutely continuous distribution functions E_X and F_Y, respectively. Assume that

$$QS_X(p) \leq QS_Y(p) \qquad \text{for all } p \in (0.5,\ 1). \tag{7}$$

Then X is called *smaller than Y in quantile spread order* (denoted as $X \leq_{QS} Y$).

It is conceptually clear that the order $\leq_{QS}$ indeed corresponds to a comparison of X and Y by variability because it requires the difference between the corresponding quantiles of X to be smaller than the corresponding quantiles of Y. The next proposition collects some elementary properties of the quantile spread order.

Proposition 1. *Let X and Y be two random variables with absolutely continuous distribution functions F_X and F_Y, respectively.*

(a) The order $\leq_{QS}$ is location-free, i.e.,

$$X \leq_{QS} Y \qquad \text{iff } X + c \leq_{QS} Y \quad \text{for any real } c.$$

(b) $X \leq_{QS} aX, \qquad a \geq 1.$

(c) $X \leq_{QS} Y, \qquad \text{iff} - X \leq_{QS} -Y.$

(d) Assume F_X and F_Y, are symmetric, then

$$X \leq_{QS} Y$$

if and only if

$$F_X^{-1}(p) \leq F_Y^{-1}(p) \qquad \text{for } p \in (0.5, 1).$$

Proof. Part (a) is obvious from the definition of quantile spread.

For part (b) note that $p = F_{aX}(x') \equiv P(aX \leq x') = P(X \leq x'/a) \equiv F_X(x'/a)$. Thus, $F_{aX}^{-1}(p) = aF_X^{-1}(p)$. This immediately implies the hypothesis.

For part (c), it is easy to show that $F_X^{-1}(p) = S_{-X}^{-1}(p) = F_{-X}^{-1}(1-p)$, from which the hypothesis follows from the definition of quantile order.

In the case of (d), note that the symmetry around 0 implies $F^{-1}(p) = -F^{-1}(1-p)$, from which the hypothesis follows immediately. The case of distribution functions, symmetric around a point different from 0, can be dealt with by using the location-free property of (a). □

There exist several different order concepts in the literature that characterize the variability of random variables (for an overview, see Shaked and Shanthikumar (1994); Kochar, 1997). Closely related to the quantile spread order is the following recent concept:

Definition 3. Let X and Y be two random variables with absolutely continuous[3] distribution functions F_X and F_Y, respectively. Assume that

$$F_X^{-1}(\alpha) - F_X^{-1}(\beta) \leq F_Y^{-1}(\alpha) - F_Y^{-1}(\beta) \tag{8}$$

whenever $0 < \beta \leq \alpha < 1$. Then X is called *smaller than Y in dispersive order* (denoted as $X \leq_{\text{disp}} Y$).

Obviously, $X \leq_{\text{disp}} Y$ **implies** $X \leq_{QS} Y$, but not vice versa. If X and Y have finite means, it can be shown (see Shaked & Shantikumar, 1994, p. 74) that $X \leq_{\text{disp}} Y$ implies $\text{Var}(X) \leq \text{Var}(Y)$. In view of the discussion above the dispersive order concept is too strong for our purposes. It stands to reason that our weaker order, where their α equals our p and their β equals our $1 - p$, will accommodate more distributions that obey the ordering principle, although the extent to which this holds has yet to be explored. Furthermore, the dispersive order has not been employed toward the goals on which we labor here. Nevertheless, for symmetric distributions, our quantile spread order is sufficient to order the variances.

Proposition 2. *Let X and Y be two random variables with absolutely continuous symmetric distribution functions F_X and F_Y, respectively. Then*

$$X \leq_{QS} Y \quad \textit{implies} \quad \text{Var}\,(X) \leq \text{Var}\,(Y).$$

Proof. Without loss of generality, let us assume symmetry around 0. By definition,

$$\text{Var}\,(X) = E(X^2) = \int_0^1 x^2 dF(x) = \int_0^1 [x(p)]^2 dp = 2\int_{.5}^1 [x(p)]^2 dp,$$

where the last equality results from symmetry, $x(p) = -x(1-p)$. Now assume $X \leq_{QS} Y$. By Proposition 1(d), this implies $x(p) \leq y(p)$ for $p \in (0.5, 1)$. Inserting in the above integral immediately gives the hypothesis. □

4. Quantile Spread for Extreme Value Statistics

4.1. Primary Notions and Examples

As discussed earlier we are interested in the behavior of a measure of variability for the minimum (or maximum) as the number of components, i.e., sample size, increases. As it turns out, for the quantile spread it is relatively straightforward in many cases to determine whether the quantile spread increases or decreases[4] with sample size. We consider several examples after which some more general results will be presented.

Let X_i, $i = 1, \ldots, n$, be independent and identically distributed (iid) random variables with distribution function F. From here on, our prime focus is on the quantile spread for the minimum and the maximum statistics, respectively,

$$X_{1:n} = \min\{X_1, \ldots, X_n\}$$

[3] Again, absolute continuity is usually not required.

[4] In this paper, "increasing" and "decreasing" mean "nondecreasing" and "nonincreasing," respectively.

and

$$X_{n:n} = \max\{X_1, \ldots, X_n\}.$$

The distribution function for the minimum is

$$F_{\text{MIN}}(x) = 1 - (1 - F(x))^n$$

or, with $S(x) = 1 - F(x)$,

$$S_{\text{MIN}}(x) = (S(x))^n.$$

With $S_{\text{MIN}}(x) = p$ it follows that

$$S_{\text{MIN}}^{-1}(p) = S^{-1}(p^{1/n}).$$

Hence,

$$QS_{\text{MIN}}(p;\, n) = S^{-1}[(1 - p)^{1/n}] - S^{-1}(p^{1/n})$$

with $p \in (0.5, 1)$.

Next, we consider the quantile spread for the maximum. The distribution function for the maximum is $F_{\text{MAX}}(x) = (F(x))^n$. With $F_{\text{MAX}}(x) = p$ it follows that $F_{\text{MAX}}^{-1}(p) = F^{-1}(p^{1/n})$. Hence,

$$QS_{\text{MAX}}(p; n) = F^{-1}(p^{1/n}) - F^{-1}[(1 - p)^{1/n}]$$

with $p \in (0.5, 1)$.

Obviously, we can write $X_{1:n} \leq_{QS} X_{1:n-1}$ and $X_{n:n} \leq_{QS} X_{n-1:n-1}$ if $QS_{\text{MIN}}(p; n) \leq QS_{\text{MIN}}(p; n - 1)$ and $QS_{\text{MAX}}(p; n) \leq QS_{\text{MAX}}(p; n - 1)$, respectively, for all $p \in (0.5, 1)$.

Expanding on earlier examples we have

Example 1. (con'td): (Weibull Minimum). Using equation (3) we find that

$$QS_{\text{MIN}}(p; n) = (n\beta)^{-1/\alpha} \left[\left(\ln \frac{1}{1 - p} \right)^{1/\alpha} - \left(\ln \frac{1}{p} \right)^{1/\alpha} \right]. \tag{9}$$

Given that the difference of the logarithm terms is positive for any $p \in (0.5, 1)$, the quantile spread for the Weibull minimum is easily seen to decrease as a function of sample size n, $X_{1:n} \leq_{QS} X_{1:n-1}$.

Example 1. (cont'd) (Weibull Maximum). Using equation (3) it follows from simple algebra that

$$QS_{\text{MAX}}(p; n) = \beta^{-1/\alpha} \left[\left(\ln \frac{1}{1 - p^{1/n}} \right)^{1/\alpha} - \left(\ln \frac{1}{1 - (1 - p)^{1/n}} \right)^{1/\alpha} \right]. \tag{10}$$

In this case, the behavior of the quantile spread can be shown to increase or to decrease depending on the specific parameter values of $X_{n:n}$ if is not *QS* ordered.

We return to a prototypical symmetric distribution.

Example 2. (con'td): (Logistic Maximum). Using equation (6) it follows that the quantile spread for the maximum is

$$QS_{\text{MAX}}(p) = \ln \left[\frac{p^{1/n}[1 - (1 - p)^{1/n}]}{(1 - p^{1/n})(1 - p)^{1/n}} \right]. \tag{11}$$

It is easy to show that $X_{n:n} \leq X_{n-1:n-1}$.

Rather than considering the minimum for the logistic separately, this can be deduced from the following general result on symmetric distributions.

Proposition 3. *For a symmetric parent distribution the quantile spread for the maximum is identical to the quantile spread for the minimum.*

Proof. By definition,

$$QS_{\text{MAX}}(p;n) = F^{-1}(p^{1/n}) - F^{-1}[(1-p)^{1/n}].$$

By symmetry, $F^{-1}(p) = -F^{-1}(1-p)$ for any p. Inserting this result in the above equation, and using the relation $F^{-1}(p) = S^{-1}(1-p)$, yields

$$\begin{aligned} QS_{\text{MAX}}(p;n) &= F^{-1}[1-(1-p)^{1/n}] - F^{-1}(1-p^{1/n}) \\ &= S^{-1}[(1-p)^{1/n}] - S^{-1}(p^{1/n}) \\ &= QS_{\text{MIN}}(p;n) \end{aligned}$$

by definition. □

Therefore, the quantile spread on the minimum statistic also decreases for the logistic distribution.

A nonsymmetric distribution, the double exponential, turns out to be particularly interesting.

Example 3. (Double Exponential (or Extreme Value) Distribution *(Warburton, 1967)*). The inverse of the cumulative distribution

$$F(x) = \exp[-\exp(-x)], \qquad -\infty < x < +\infty,$$

is

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

yielding the double exponential quantile spread

$$\ln\left[\frac{\ln(1-p)}{\ln p}\right]. \tag{12}$$

Again, simple algebra yields the quantile spread for the minimum and the maximum statistics,

$$QS_{\text{MIN}}(p;n) = \ln\left[\frac{\ln(1-p^{1/n})}{\ln[1-(1-p)^{1/n}]}\right] \tag{13}$$

and

$$QS_{\text{MAX}}(p;n) = \ln\left[\frac{\ln(1-p)}{\ln p}\right]. \tag{14}$$

Note that, intriguingly, the quantile spread for the maximum does not vary ($X_{n:n} =_{QS} X_{n-1,n-1}$) with sample size n, while the minimum quantile spread can be shown to be a decreasing function of n, $X_{1:n} <_{QS} X_{1:n-1}$.

4.2. The Fundamental Lemmas

In the following, we derive some further more general results on the behavior of the quantile spread for the minimum and the maximum statistics. Note that for convenient notation we write $q = 1 - p$, so that always $q \in (0,\ 0.5)$ and $p \in (0.5,\ 1)$. We shall let $X(r)$ be the r-quantile figured as the inverse of survivor functions, and $X'(r)$ be the r-quantile figured as the inverse of distribution functions. $X(r)$ is then a decreasing function of r in contrast to $X'(r)$ which is an increasing function of r.

Lemma 1.

(a) The quantile spread $QS_{\mathrm{MIN}}(1-q\,;n)$ decreases in n for any $q \in (0, 0.5)$ iff

$$\left.\frac{dX}{dr}\right|_{r=q^{1/n}} \div \left.\frac{dX}{dr}\right|_{r=(1-q)^{1/n}} \geq \frac{(1-q)^{1/n}\ln(1-q)}{q^{1/n}\ln q}. \tag{15}$$

(b) The quantile spread $QS_{\mathrm{MAX}}(p;n)$ decreases in n for any $p \in (0.5, 1)$ iff

$$\left.\frac{dX'}{dr}\right|_{r=p^{1/n}} \div \left.\frac{dX'}{dr}\right|_{r=(1-p)^{1/n}} \geq \frac{(1-p)^{1/n}\ln(1-p)}{p^{1/n}\ln p}. \tag{16}$$

Proof. (a) We can legitimatize this inequality by treating n as a continuous variable, since the arguments are differentiable functions of n. Then,

$$\frac{d}{dn}[X(q^{1/n}) - X((1-q)^{1/n})] =$$

$$-\frac{1}{n^2}\left(\left.\frac{dX}{dr}\right|_{r=q^{1/n}} q^{1/n}\ln q - \left.\frac{dX}{dr}\right|_{r=(1-q)^{1/n}} (1-q)^{1/n}\ln(1-q)\right). \tag{17}$$

Notice that $dX/dr < 0$. Now, comparing this expression to 0 establishes the result.
Part (b) is proven in the same fashion. □

Lemma 2. *For $q \in (0, 0.5)$, $p \in (.5, 1)$:*

(a)

$$|q^{1/n}\ln q| > |(1-q)^{1/n}\ln(1-q)|.$$

(b)

$$|p^{1/n}\ln p| < |(1-p)^{1/n}\ln(1-q)|. \tag{18}$$

Proof. (a) Expanding the log functions the above inequality turns into

$$q^{1/n}\sum_{i=1}^{\infty}\frac{(1-q)^i}{i} > (1-q)^{1/n}\sum_{i=1}^{\infty}\frac{q^i}{i}\ .$$

Comparing these two series term by term, with $k \geq 1$, we find

$$q^{1/n}(1-q)^k > (1-q)^{1/n}q^k$$

or

$$(1-q)^{nk-1} > q^{nk-1},$$

which holds because $0 < q < 0.5$. Hence, the inequality holds for the entire series and the lemma is true.

Part (b) is then also implied. □

Hence, from Lemma 2 the terms in (a) act to decrease QR_{MIN} in (17) since the right-hand side of (15) is less than 1 and $dX/dr < 0$ for the minimum.

Lemma 3 gives sufficient conditions for the minimum or maximum to decrease as n grows.

Lemma 3.

(a) For the minimum, because $dX/dr < 0$, *in the case of the minimum (because the survivor function is decreasing), if*

$$\left.\frac{dX}{dr}\right|_{q^{1/n}} < \left.\frac{dX}{dr}\right|_{(1-q)^{1/n}}$$

these terms act in the same direction as the power-log terms in Lemma 2 and QS_{MIN} *decreases. Otherwise, they act in contrast and more information must be garnered.*

(b) Analogously, because $dX'/dr > 0$, *in the case of the maximum if* $dX'/dr|_{p^{1/n}} < dX'/dr|_{(1-p)^{1/n}}$ *then these terms act in the same direction as the power-log terms and* QS_{MAX} *decreases. Otherwise, they act in contrast and more information must be garnered.*

Proof. Evident. □

We next turn to more subtle aspects of density shape and how it influences quantile spread.

4.3. New Measures of Skew and their Relationship to Quantile Spread

The next definition establishes an aspect of the shape of the parent density, relating it to the changing independent variate (through the survivor or distribution inverse as n varies) and an inequality that will be of immediate use.

Definition 4.

(a) A parent density f is *locally positively min-skewed* if, for a given q, n,

$$f(S^{-1\cdot}((1-q)^{1/n})) \geq f(S^{-1}(q^{1/n})).$$

It is *globally min-skewed* if the above holds for all $n \geq 1$, $q \in (0, 0.5)$, and if the inequality is strict for $n \geq 2$.

(b) A parent density f is locally negatively max-skewed if for a given p, n, $f(F^{-1}(p^{1/n})) \geq f(F^{-1}(1-p)^{1/n})$. It is globally max-skewed if the above holds for all $n \geq 1$, $p \in (0.5, 1)$, and if the inequality is strict for $n \geq 2$.

Observe that Definition 4(a) indicates that the left tail of the density tends to be greater than the right tail in a certain sense; and Definition 4(b) indicates the opposite. Thus, these conditions intuitively represent a form of positive and negative skew, respectively. The exact relations to traditional more coarse definitions of skew such as Pearson's β_1 or δ_1, or modern notions (e.g., Arnold & Groeneveld, 1995), are unknown at present, but some subsequent results will again support some type of relationship.

The next proposition offers sufficient conditions for $X_{1:n} <_{QS} X_{1:n-1}$ **or** $X_{n:n} <_{QS} X_{n-1:n-1}$ respectively.

Proposition 4.

(a) If the parent density is globally positive min-skewed, then $X_{1:n} <_{QS} X_{1:n-1}$, *that is,* QS_{MIN} *is always decreasing for* $n \geq 2$.

(b) If the parent density is globally negatively max-skewed, then $X_{\mathrm{MIN}} < X_{n-1;n-1}$, *that is* QS_{MAX} *is always decreasing for* $n \geq 2$.

Proof. (a) Let$S^{-1}(r) = x$. By elementary calculus,

$$\frac{dX}{dr} = \frac{dS^{-1}(r)}{dr} = \left(\frac{dS}{dx}\right)^{-1} = -\frac{1}{f(x)} = -\frac{1}{f(S^{-1}(r))}\ .$$

From Lemma 3,

$$\left.\frac{dX}{dr}\right|_{q^{1/n}} = -\frac{1}{f(S^{-1}(q^{1/n}))} < -\frac{1}{f(S^{-1}((1-q)^{1/n}))} = \left.\frac{dX}{dr}\right|_{(1-q)^{1/n}}$$

for $n \geq 2$, where the inequality holds by virtue of Definition 2.

(b) Same kind of demonstration.

A corollary gives a special case of some interest. □

Corollary to Proposition 4.

(a) *If f is decreasing, then* $X_{1:n} <_{QS} X_{1:n-1}$ for all n.
(b) *If $f(x)$ increases, then $X_{n:n} <_{QS} X_{n-1:n-1}$ for all n.*

Proof.

(a) By f decreasing,

$$f(S^{-1}(q^{1/n})) < f(S^{-1}((1-q)^{1/n}))$$

for all q, n. This is the definition of global positive skewness, so the conclusion ensues.
Part (b) follows likewise. □

Again we see that the characteristics of f that guarantee $X_{1:n} <_{QS} X_{1:n-1}$ are very different than those which guarantee $X_{n:n} <_{QS} X_{n-1::n-1}$. This result further generalizes the earlier conclusion that the exponential distribution has a decreasing minimum quantile spread (see Example 1). The following example shows that the dX/dr term can override the power-log terms and cause an increase in QS_{MIN}.

Example 4. (Negative Exponential Distribution). Let

$$F(x) = \begin{cases} \exp(x) & \text{for}\, x \leq 0, \\ 1 & \text{for}\, x > 0. \end{cases} \tag{19}$$

This yields the quantile spread $QS(p) = \ln(1-q)/q$ so that

$$QS_{\text{MIN}}(1-q; n) = \ln\left[\frac{(1-q)^{1/n}}{1-(1-q)^{1/n}}\right]. \tag{20}$$

This can readily be shown to be increasing in n for all $q \in (0, 0.5)$.

It is equally easy to show that this distribution yields a decreasing QS_{MAX}, in line with the corollary to Proposition 4. Similarly, it was earlier pointed out, as part of the Weibull Maximum example (also see (1)), that the positive exponential distribution produces an increasing QS_{MAX} and a decreasing QS_{MIN}. Hence, it would appear that what guarantees a decreasing QS_{MAX} tends to facilitate, but does not always guarantee, an increasing QS_{MIN} and what ensures a decreasing QS_{MIN} tends to act in the direction of facilitating an increasing QS_{MAX}. Thus, it is intriguing to learn that a type of distribution exists that guarantees both positive min-skew and negative max-skew and, therefore, simultaneously delivers global decreases in both the MAX and MIN quantile spreads.

Proposition 5.

(a) *If $f(x)$ is symmetric, decreases to the median, and increases thereafter, then* $X_{1:n} \leq_{QS} X_{1:n-1}$ for $n \geq 2$.
(b) $X_{n:n} \leq_{QS} X_{n-1:n-1}$ for $n \geq 2$.

Proof. (a) By Proposition 4, if

$$f(S^{-1}(q^{1/n})) \leq f(S^{-1}((1-q)^{1/n}))$$

for $n > 1$ (see Definition 2), then Q_{MIN} is always decreasing. Now, $f(S^{-1}(q)) = f(S^{-1}(1-q))$, i.e., the above holds with equality at $n = 1$ due to symmetry.

However, because $1-(1-q)^{1/n} < q^{1/n} < (1-q)^{1/n}$ for $1 < n$ and $q < 1/2$, always,

$$f(S^{-1}(q^{1/n})) < f(S^{-1}((1-q)^{1/n})) \qquad \text{for } n \geq 2.$$

(b) The proof is similar but values of $f(F^{-1}(p^{1/n}))$ and $f(F^{-1}((1-p)^{1/n}))$ are ordered and moving to the left rather than toward the right. □

It may be observed in the proof of Proposition 5 that the combination of the convexity of f, with its leftward (MIN) vs. rightward (MAX) movement as a function of n, permits satisfaction as both the MAX and MIN quantile spreads decrease with n.

The new notions of skew have proven to be of use in exploring the native properties of distribution with regard to the behavior of extremal quantile spread behavior. However, outside the strictly increasing, strictly decreasing, or U-shaped types of distribution, do any of the previous examples obey the sufficient skew conditions?

Example 1. *(cont'd):* (Weibull Minimum). It may be recalled that QS_{MIN} always decreases as a function of n. The main parameter driving behavior is α, so we first set $\beta = 1$. Then numerical computation indicates that the positive min-skew condition holds for $\alpha < 1.0$ and $n = 1, 10$. It already begins to fail at $\alpha = 1.1$ for larger values of n, $q \in (0.5)$. In general, it seems to be correlated, in terms of α, to the traditional measure of skew γ_1 (e.g., Cox, 1962), in that as α increases $\gamma_1$becomes negative and positive min-skew (Definition 2) begins to fail. The reverse is true as α decreases toward 0. Furthermore, for fixed α, higher values of q, n reduces positive min-skew (e.g., measured by the difference of the two quantities in Definition 2).

Example 2. *(cont'd):* (Logistic). The logistic distribution is symmetric, so the behavior of MAX, and MIN must be the same. Straightforward analytic investigation shows that the sufficient condition never holds. That is, $f(S^{-1}((1-q)^{1/n})) = (1-q)^{1/n}(1-(1-q)^{1/n}) \leq q^{1/n}(1-q)^{1/n} = f(S^{-1}(q^{1/n}))$, hence always violating positive min-skew. Yet, we know, and voluminous computations illustrate, that $X_{1:n} \leq_{QS} X_{1:n-1}$. Also, the necessary and sufficient condition from Lemma 1 is supported by extensive numerical computation, which finds the satisfaction of $(1-(1-q)^{1/n})/(1-q^{1/n}) > \ln(1-q)/\ln(q)$. It appears that min- and max-skews fail too weakly to override the strong tendency of the power log terms to produce extremal decreases in n.

Example 3. *(Cont'd):* (Double Exponential). From our earlier result with the double exponential QS_{MAX}, we have $X_{n:n} =_{QS} X_{n-1:n-1}$. This equivalence across n implies that negative max-skew cannot occur and investigation substantiates that inference. With regard toQS_{MIN}, computations for $q = 0.01$ to 0.50 and $n = 1, 10$, suggest global decreasing order. This hypothesis is backed up by calculations supporting the necessary and sufficient condition in Lemma 1. Nevertheless, numerical investigation indicates that the sufficient condition captured by positive min-skew is locally obeyed only at $n = 1$ for some values of q, and fails for $n = 2, 10$ when $q \in (0, 0.5)$.

5. General Discussion

The quantile spread appears to offer a viable and fruitful approach to the study of the variability of extreme values. The properties of the entire distribution functions permit considerably more power than is true for functionals such as the variance. The quantile spread permits investigation of local and global orderings relative, for example, to families of distributions. For instance, the quantile spread for the minimum or the maximum of n random variables often acts in an orderly fashion and can be assessed locally for certain p, n and sometimes imposes the global ordering. Symmetric distributions act the same way in this quantile spread ordering for the maximum and minimum.

It was discovered that for both the maximum and the minimum random variables, there is a strong tendency for the QS to decrease, as shown by the power-log terms in Lemma 2. However, the shape of the distribution can either facilitate this tendency or counteract it. Facilitation is associated with the pair of sufficient conditions we designated as positive min-skew or negative max-skew (Lemma 3, Definition 2, and Proposition 4) in the case of the minimum and maximum, respectively. If a counteraction (given by failure of the skew measure) is not sufficiently great to overcome the tendency toward decrease, then the necessary and sufficient condition given by Lemma 1 is obeyed. Our examples indicated that in some common distributions, the ordering held despite (weak) failures of the sufficient conditions. Any symmetric U-shaped density satisfies both sufficient conditions; otherwise, increasing densities guarantee decreasing quantile spread on the maximum, and tend to work against decreases in the quantile spread on the minimum. Conversely, decreasing densities guarantee decreasing quantile spread on the minimum, and tend to act against decreasing quantile spread on the maximum.

All of these properties should assist experimentalists, for instance, those studying mental architecture as well as psychometricians and others interested (concerned with performance extrema, for example), in test behavior. However, we should now consider issues that evoke deliberation of both empirically substantive, as well as mathematical, features. A major question concerns the possibility of immediate application to data. At present, this is not feasible. Perhaps the most obvious obstacle, but one that is remediable, is that no one has collected data that are suitable for application. First of all, a paucity of studies have acquired even variances along with the means, and few, if any, have presented quantile data in such a way as to permit our treatment.

Furthermore, in any application to data, a fair test should first ascertain that certain conditions obtain. The present distributional properties assume that the component distributions are independent and identically distributed and invariant across values of n.

What about the "independence" aspect of independent and identically distributed? Currently, it is difficult to isolate and test the independence assumption in response time data mostly because of the presence of extraneous (e.g., sensory, motor, etc.) processing time components that are present along with the psychological mechanisms under study (but see Townsend & Wenger, 2004). On the other hand, accuracy/confusion approaches can test independence in a straightforward manner (Ashby & Townsend, 1986).

Next, consider the aspect of independent and identically distributed that the marginal distributions are identical. We expect that this condition will not be critical in the general behavior of extrema. In typical experiments, for instance, in memory or visual search experiments, that stipulation would imply that there are no position effects with regard to the placement of a search-target.

The fact that the parent distribution is constant, not only inside samples $1, 2, \ldots, n$, but also as n itself is varied, implies the information processing assumption of unlimited capacity. That is, each individual processing channel in a task performs just as efficiently, with $n = 5$ objects altogether, as when only one is operating. Many human perceptual, cognitive, and motor tasks are limited capacity rather than unlimited capacity, but the latter is often assumed even in relatively pristine areas like psychophysics and, at least sometimes, appears to be correct (e.g., see Hughes & Townsend, 1998).

The above paragraphs outline some of the obstacles in applying the present results, at least to data involving internal psychological processes. We do not view them as insurmountable although some are certainly nontrivial. However, several decades ago virtually nothing was known about the mathematics of parallel and serial processes and how these descriptions could lead to testable differences, and the present status of these archtectures is heartening (e.g., Townsend & Ashby, 1983; Townsend, 1990b). So, we regard the present work as embryonic and propaedeutic to an expanding knowledge base relative to engaging actual data.

Finally, we would like to see the present work, as well as other modern approaches that possess a finer-grained articulation of distributional properties into the statistical measure of interest, play a stronger role in current methodology and modeling. Using analogous concepts with regard to a dominance order of distributional properties (e.g., an order of cdf's implies an order of means an medians), has proven of benefit in psychological theory and methodology (e.g., Townsend, 1990c; Van Zandt, 2002; Schweickert, Giorgini, & Dzhafarov, 2000).

References

Arnold, B.C. (1987). *Majorization and the Lorenz order: A brief introduction*. New York: Springer-Verlag.

Arnold, B.C., & Groeneveld, R.A. (1995). Measuring skewness with respect to the mode. *The American Statistician, 49*, 34–38.

Ashby, F.G., & Townsend, J.T. (1986). Varieties of perceptual independence. *Psychological Review, 93*, 154–179.

Belzunce, F. (1999). On a characterization of right spread order by the increasing convex order. *Statistics & Probability Letters, 45*, 103–110.

Bernstein, I.H. (1970). Can we see and hear at the same time? Some recent studies of intersensory facilitation of reaction time. *Acta Psychologica, 33*, 21–35.

Colonius, H. (1990). Possibly dependent probability summation of reaction time. *Journal of Mathematical Psychology, 34*, 253–275.

Colonius, H., & Vorberg, D. (1994). Distribution inequalities for parallel models with unlimited capacity. *Journal of Mathematical Psychology, 38*, 35–58.

Cox, D.R. (1962). *Renewal theory*. London: Methuen.

Donnelley, N., Found, A., & Müller, H.J. (1999). Searching for impossible objects: Processing form and attributes in early vision. *Perception & Psychophysics, 61*, 675–690.

Dzhafarov, E.N. (1992). The structure of simple reaction time to step-function signals. *Journal of Mathematical Psychology, 36*, 235–268.

Egeth, H.E. (1966). Parallel vs. serial processes in multidimensional stimulus discrimination. *Perception & Psychophysics, 1*, 245–252.

Gnedenko, B.V. (1943). Sur la distribution limited u terme maximum d'une serie aleatorie. *Annals of Mathematics, 44*, 423–453.

Gumbel, E.J. (1958). *Statistics of extremes*. New York: Columbia University Press.

Harter, H.L. (1978). A bibliography of extreme-value theory. *International Statistical Review, 46*, 279–306.

Hartley, H.O., & David, H.A. (1951). Universal bounds for mean spread and extreme observation. *Annals of Mathematical Statistics, 25*, 85–99.

Hughes, H.C., & Townsend, J.T. (1998). Varieties of binocular interaction in human vision. *Psychological Science, 9*, 53–60.

Kleiber, C. (2002). Variability ordering of heavy-tailed distributions with applications to order statistics. *Statistics and Probability Letters, 58*, 381–388.

Kochar, S.C. (1997). Dispersive ordering of order statistics. Unpublished manuscript.

Luce, R.D. (1986). *Response times: Their role in inferring elementary mental organization*. New York: Oxford University Press.

Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology, 14*, 247–279.

Mordkoff, J.T., & Yantis, S. (1991). An interactive race model of divided attention. *Journal of Experimental Psychology, 17*, 520–538.

Moriguti, S. (1951). Extremal properties of extreme value distributions. *Annals of Mathematical Statistics, 22*, 523–536.
Moriguti, S. (1953). A modification of Schwarz's inequality with applications to distributions. *Annals of Mathematical Statistics, 24,* 107–113.
Munoz-Perez, J. (1990). Dispersive ordering by the spread function. *Statistics & Probability Letters, 10*, 407–410.
Schweickert, R., Giorgini, M., & Dzhafarov, E. (2000). Selective influence and response time cumulative distributed functions in serial-parallel task networks. *Journal of Mathematical Psychology, 44*, 504–535.
Shaked, M., & Shanthikumar, J.G. (1994). *Stochastic orders and their applications*. San Diego, CA: Academic Press.
Sternberg, S. (1966). High-speed scanning in human memory. *Science, 153*, 652–654.
Townsend, J.T. (1990a). *A potpourri of ingredients for Horse (Race) Soup*. Technical Report #32. Cognitive Science Program. Indiana University, Bloomington, IN.
Townsend, J.T. (1990b). Serial vs. parallel processing: Sometimes they look like Tweedledum and Tweedledee but they can (and should) be distinguished. *Psychological Science, 1*, 46–54.
Townsend, J.T. (1990c). The truth and consequences of ordinal differences in statistical distributions: Toward a theory of hierarchical inference. *Psychological Bulletin, 108*, 551–567.
Townsend, J.T. (2001). A clarification of self-terminating vs. exhaustive variances in serial and parallel models. *Perception & Psychophysics, 63*, 1101–1106.
Townsend, J.T., & Ashby, F.G. (1983). *The stochastic modeling of elementary psychological processes*. Cambridge, MA: Cambridge University Press.
Townsend, J.T., & Nozawa, G. (1995). Spatio-temporal properties of elementary perception: An investigation of parallel, serial, and coactive theories. *Journal of Mathematical Psychology, 39*, 321–359.
Townsend, J.T., & Nozawa, G. (1997). Serial exhaustive models can violate the race model inequality: Implications for architecture and capacity. *Psychological Review, 104*, 595–602.
Townsend, J.T., & Wenger, M.J. (2004). A theory of interactive processing: New capacity measures and predictions for a response time inequality series. *Psychological Review*. *111*, 1003-1035.
Van Zandt, T. (2002). Analysis of response time distributions. In J.T. Wixted (Ed.) *Stevens handbook of experimental psychology* (3rd ed.). Methodology in Experimental Psychology (Vol. 4, pp. 461–516) (H. Pashler, ed.). New York: Wiley.
Warburton, C.E., Jr. (1967). On the application of the double exponential and Weibull distributions to failure data. *Texfile Research Journal 37*, 703–704.

Manuscript received 31 JUL 2001
Final version received 6 MAY 2004