6-2004

Computing Higher Central Moments for Interval Data

Vladik Kreinovich
The University of Texas at El Paso, vladik@utep.edu

Luc Longpre
The University of Texas at El Paso, longpre@utep.edu

Scott Ferson

Lev Ginzburg

Comments:

Technical Report: UTEP-CS-03-14b

Computing Higher Central Moments for Interval Data

Vladik Kreinovich[1], Luc Longpré[1],
Scott Ferson[2], and Lev Ginzburg[2]

[1]Computer Science Department, University of Texas at El Paso
El Paso, TX 79968, USA, {vladik,longpre}@cs.utep.edu
[2]Applied Biomathematics, 100 North Country Road
Setauket, NY 11733, USA, {scott,lev}@ramas.com

Abstract

Higher central moments are very useful in statistical analysis: the third moment M_3 characterizes asymmetry of the corresponding probability distribution, the fourth moment M_4 describes the size of the distribution's tails, etc. When we know the exact values $x_1, \ldots, x_n$, we can use the known formulas for computing the corresponding sample central moments. In many practical situations, however, we only know intervals $\mathbf{x}_1, \ldots, \mathbf{x}_n$ of possible values of x_i; in such situations, we want to know the range of possible values of M_m. In this paper, we propose algorithms that compute such ranges.

1 Formulation of the Problem

Higher moments are important. In engineering and science, when we have n measurement results $x_1, \ldots, x_n$, traditional statistical approach (see, e.g., [5, 21]) usually starts with computing their (sample) average

$$E = \bar{x} = \frac{x_1 + \ldots + x_n}{n}$$

and their (sample) variance

$$V = \frac{(x_1 - E)^2 + \ldots + (x_n - E)^2}{n}.$$

(Often, a standard deviation $\sigma \stackrel{\text{def}}{=} \sqrt{V}$ is used instead of V.) Variance is a particular case of a *central moment*

$$M_m = \frac{(x_1 - E)^m + \ldots + (x_n - E)^m}{n} \tag{1}$$

corresponding to $m = 2$. Higher moments – i.e., moments corresponding to $m = 3, 4, \ldots$ – are also used in engineering and science. For example, the third central moment M_3 is used to describe the asymmetry of the corresponding probability distribution, and the fourth central moment M_4 is used to describe the size of the distribution's tails. To be more precise, *skewness* M_3/σ^3 is used to characterize asymmetry, and *kurtosis* $M_4/\sigma^4 - 3$ is used to characterize the size of the tails (3 is subtracted so that kurtosis be equal to 0 for the practically frequent case of a normal distribution).

In addition to central moments, sometimes, non-central moments are also used:

$$M'_m = \frac{x_1^m + \ldots + x_n^m}{n}.$$

Case of interval data. When we know the exact values of x_i, then we can compute each moment by using the explicit formula (1). In some practical situations, however, we only have intervals $\mathbf{x}_i = [\underline{x}_i, \overline{x}_i]$ of possible values of x_i.

This happens, for example, if instead of observing the actual value x_i of the random variable, we observe the value $\widetilde{x}_i$ measured by an instrument with a known upper bound Δ_i on the measurement error (and no information about probabilities of different possible values of measurement error); then, the actual (unknown) value is within the interval $\mathbf{x}_i = [\widetilde{x}_i - \Delta_i, \widetilde{x}_i + \Delta_i]$. The practical importance of this question was emphasized, e.g., in [17, 18] on the example of processing geophysical data.

Another case is when we try to maintain privacy in statistical databases (see, e.g., [14]). If we collect a lot of data about people, we can extract important information that can be useful both to these people and to the economy. However, even if we strip off names and IDs from individual records, there is still enough information in, say, individual census records to enable a researcher to pinpoint a person and thus, to learn information about this person that would rather be private. Thus, while we should allow statistical queries, we cannot allow unrestricted statistical queries – they will violate privacy.

There exist many methods for maintaining privacy; most of these methods, however, do not provide an absolute guarantee that privacy will be maintained – only probabilistic guarantee. Besides, in these methods, each access to the statistical database restricts the potential future uses of this data.

One way to guarantee privacy is that, instead of keeping exact values of salary, age, weight, etc., we select a partition, and only keep the interval that contains the actual data. For example, we only keep the information that the age is between 40 and 50, or between 50 and 60, etc. Then, no matter how many queries we ask, we will never pinpoint the individual data more accurately.

In this case, if we are interested in a statistical characteristic $A(x_1,\ldots,x_n)$ and we only know the intervals $\mathbf{x}_i$ that contain the actual (unknown) values x_i, then it is natural to return the range of A, i.e., the interval

$$\{A(x_1,\ldots,x_n)\,|\,x_1\in\mathbf{x}_1,\ldots,x_n\in\mathbf{x}_n\}.$$

There are other practical cases when we need to apply statistical analysis to interval-valued data. For example, when we study the effect of a water pollutant on the fish life expectancy, we usually check on the fish, say, once a day. As a result, if, e.g., a fish was alive on Day 5 but dead on Day 6, all we know about the actual lifetime x_i of this fish i is that $x_i\in[5,6]$. The problem of statistical analysis of such interval data (called "censored data" in statistics) is exactly the same as for the privacy case.

Computing moments for interval data is a particular case of a general problem of combining interval and probabilistic uncertainty. In all the above cases, we know the values $x_1,\ldots,x_n$ with interval uncertainty and therefore, the sets of possible values of E, V, and M_k are also intervals. How can we compute these intervals?

This is a specific problem related to a combination of interval and probabilistic uncertainty. Such problems – and their potential applications – have been described, in a general context, in the monographs [12, 22]; for further developments, see, e.g., [1, 2, 3, 4, 6, 9, 13, 16, 19, 20, 23] and references therein.

The specific problem of estimating the range of a moment under interval uncertainty was formulated by C. Manski; his research is summarized in [14] (see also [10]).

Case of non-central moments. For the *first* moment E, the solution is easy: since E is an increasing function of each of the variables x_i, it is easy to compute the interval $\mathbf{E}=[\underline{E},\overline{E}]$ of possible values of E:

$$\underline{E}=\frac{\underline{x}_1+\ldots+\underline{x}_n}{n};\quad \overline{E}=\frac{\overline{x}_1+\ldots+\overline{x}_n}{n}.$$

For example, for three intervals $\mathbf{x}_1=\mathbf{x}_2=\mathbf{x}_3=[-1,1]$, we have $\mathbf{E}=[-1,1]$. This result makes sense: e.g., since all we know about the values x_i is that $x_i\in[-1,1]$, it is quite possible that all three values x_i are equal to 1, in which case $E=1$. Similarly, it is possible that $x_1=x_2=x_3=-1$, so the resulting value $E=-1$ is also possible. On the other hand, since $|x_i|\le 1$, the average E of the three values x_i must always lie within the interval $[-1,1]$.

Similarly, we can easily compute the exact bounds for *non-central* moments M'_m for odd m:

$$\underline{M'_m}=\frac{(\underline{x}_1)^m+\ldots+(\underline{x}_n)^m}{n};\quad \overline{M'_m}=\frac{(\overline{x}_1)^m+\ldots+(\overline{x}_n)^m}{n}.$$

For example, for $m = 3$ and $\mathbf{x}_1 = \mathbf{x}_2 = \mathbf{x}_3 = [-1, 1]$, we get $\mathbf{M}'_3 = [-1, 1]$.

For even m, it is known (see, e.g., [11, 15]) that the range of x^m when $x \in [\underline{x}, \overline{x}]$ is equal to $[(\min(|\underline{x}|, |\overline{x}|))^m, (\max(|\underline{x}|, |\overline{x}|))^m]$ when $0 \notin [\underline{x}, \overline{x}]$ and to $[0, (\max(|\underline{x}|, |\overline{x}|))^m]$ otherwise. Thus,

$$\underline{M}'_m = \frac{F^m(\underline{x}_1, \overline{x}_1) + \ldots + F^m(\underline{x}_n, \overline{x}_n)}{n},$$

where $F(a, b) = 0$ if $a \le 0 \le b$ and $\min(|a|, |b|)$ otherwise, and

$$\overline{M}'_m = \frac{(\max(|\underline{x}_1|, |\overline{x}_1|))^m + \ldots + (\max(|\underline{x}_n|, |\overline{x}_n|))^m}{n}.$$

For example, for $[-1, 1]$ and $m = 4$, we have $[-1, 1]^4 = [0, 1]$ hence $\mathbf{M}'_4 = [0, 1]$ – which again makes perfect sense: this moment can be 0 if all the values x_i are equal to 0; it can be equal to 1 if all the values are equal to 1; and it cannot be larger than 1 because it is the average of three values x_i^4, each of which is ≤ 1.

Case of central moments: what is known? What is the interval $[\underline{M}_m, \overline{M}_m]$ of possible values for the *central* moment M_m? In [7, 8], we have analyzed this problem for the case of variance ($m = 2$). In this case, we have shown that the general problem of computing the interval $[\underline{V}, \overline{V}]$ exactly is NP-hard. We also showed:

- that there exists a quadratic-time algorithm for computing $\underline{V}$; and
- that there exists a quadratic-time algorithm for computing $\overline{V}$ for the case when the intervals do not all group together, i.e., crudely speaking, when for some C, every C different intervals $\mathbf{x}_{i_1}, \ldots, \mathbf{x}_{i_C}$ have an empty intersection.

In particular, for the above three intervals $\mathbf{x}_i = [-1, 1]$, we get the range $\mathbf{M}_2 = [0, 8/9]$. The smallest value 0 is attained when all x_i are equal, the largest value 8/9 is attained, e.g., when $x_1 = x_2 = 1$ and $x_3 = -1$ – and one can check that for all other values $x_i \in [-1, 1]$, the central moment M_2 is always within the interval $[0, 8/9]$.

In this paper, we extend these algorithms to higher central moments.

2 First Result: Computing $\underline{M}_m$ for Even m

Theorem 1. *For every even m, there exists an algorithm that computes $\underline{M}_m$ in quadratic time.*

This algorithm is as follows:

- First, we sort all $2n$ values $\underline{x}_i$, $\overline{x}_i$ into a sequence $x_{(1)} \le x_{(2)} \le \ldots \le x_{(2n)}$. This sequence divides the real line into $2n+1$ segments $[x_{(k)}, x_{(k+1)}]$, where $k = 0, \ldots, 2n$, $x_{(0)} \stackrel{\text{def}}{=} -\infty$, and $x_{(2n+1)} \stackrel{\text{def}}{=} +\infty$.
- For each of these segments $[x_{(k)}, x_{(k+1)}]$, we do the following:
  - First, we define $x_1, \ldots, x_n$ as the following linear functions of a new auxiliary variable α:
    * if $\underline{x}_i \ge x_{(k+1)}$, we take $x_i = \underline{x}_i$ (independent of α);
    * if $\overline{x}_i \le x_{(k)}$, we take $x_i = \overline{x}_i$ (independent of α);
    * in all other cases, we take $x_i = \alpha$.
  - Based on these expressions for x_i, we find the expression for

    $$E = \frac{x_1 + \ldots + x_n}{n}$$

    as a linear function of α.
  - Then, we substitute these expressions for x_i and E into the equation

    $$\frac{1}{n} \cdot \sum_{i=1}^{n} (x_i - E)^{m-1} = (\alpha - E)^{m-1}.$$

    This equation is is a polynomial equation of order $\le m-2$ in terms of the unknown α (terms proportional to α^{m-1} cancel out), so it has $\le m-2$ solutions. We compute these solutions.
  - For each of the solutions that is inside the segment $[x_{(k)}, x_{(k+1)}]$, we substitute the corresponding value α into the formulas for x_i and E, thus, we compute x_i and E; based on these values, we compute

    $$M_m = \frac{1}{n} \cdot \sum_{i=1}^{n} (x_i - E)^m.$$

- The smallest of thus computed values M_m is the desired lower bound $\underline{M}_m$.

In particular, when $\mathbf{x}_1 = \mathbf{x}_2 = \mathbf{x}_3 = [-1, 1]$ and $m = 4$, this algorithm produces the correct bound $\underline{M}_4 = 0$ – that is attainable if, e.g., $x_1 = x_2 = x_3 = 0$.

For reader's convenience, the proof that this algorithm always computes $\underline{M}_m$ correctly, and that it requires quadratic time is placed (as well as all other proofs) in the special Proofs section at the end of this paper.

3 Second Result: Computing $\overline{M}_m$ for Even m in Time 2^n

For small number of measurements, we can use the following algorithm to compute $\overline{M}_m$:

Theorem 2. *For every even m, there exists an algorithm that computes $\overline{M}_m$ in time $O(2^n)$.*

The algorithm is as follows: for each i, we select either $x_i = \underline{x}_i$ or $x_i = \overline{x}_i$. For each i from 1 to n, there are two options, so totally, we have 2^n combinations to try. For each of these combinations, we compute M_m; the largest of the resulting 2^n values is the desired upper bound $\underline{M}_m$.

In particular, for $\mathbf{x}_1 = \mathbf{x}_2 = \mathbf{x}_3 = [-1, 1]$ and $m = 4$, this algorithm produces the bound $\overline{M}_4 = 32/27 \approx 1.19$ – that is attainable if, e.g., $x_1 = x_2 = 1$ and $x_3 = -1$. Once can check that for all other combinations $x_i \in [-1,]$, we get smaller (or equal) value of the 4th central moment M_4.

4 Third Result: Computing $\overline{M}_m$ for Even m in Quadratic Time (Case When Intervals Do Not Group Together)

Sets $S_1, \ldots, S_n$ are called *pairwise disjoint* if every pair has an empty intersection, i.e., if $S_i \cap S_j = \emptyset$ for all $i \neq j$. We can generalize this definition from pairs to tuples of arbitrary size C:

Definition 1. *Let $C \geq 2$ be an integer. We say that a sequence of sets $S_1, \ldots, S_n$ is C-*wise disjoint *if for every C different indices $i_1, \ldots, i_C$, we have $S_{i_1} \cap \ldots \cap S_{i_C} = \emptyset$.*

Theorem 3. *For every even m and for every $C \geq 2$, there exists an algorithm that computes $\overline{M}_m$ in quadratic time when the input intervals $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are C-wise disjoint.*

This algorithm is as follows:

- First, we sort all $2n$ values $\underline{x}_i, \overline{x}_i$ into a sequence $x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(2n)}$. This sequence divides the real line into $2n+1$ segments $[x_{(k)}, x_{(k+1)}]$, where $k = 0, \ldots, 2n$, $x_{(0)} \stackrel{\text{def}}{=} -\infty$, and $x_{(2n+1)} \stackrel{\text{def}}{=} +\infty$.
- For each of these segments $[x_{(k)}, x_{(k+1)}]$, we do the following:
  - First, we describe several combinations $(x_1, \ldots, x_n)$ as follows:
    - if $\overline{x}_i < x_{(k)}$, we take $x_i = \underline{x}_i$;

    * if $\underline{x}_i > x_{(k+1)}$, we take $x_i = \overline{x}_i$;
    * for all other indices i (there are $\le C$ of them), we consider all possible combinations of $x_i = \underline{x}_i$ and $x_i = \overline{x}_i$.

  As a result, we get $\le 2^C$ different combinations.
  - For each resulting combination $(x_1, \ldots, x_n)$, we compute E as the average of all the values x_i, then we compute

$$M_{m-1} = \frac{1}{n} \cdot \sum_{i=1}^{n} (x_i - E)^{m-1},$$

  and $\alpha = E + M_{m-1}^{1/(m-1)}$.
  - For each combination for which the resulting value α is within the segment $[x_{(k)}, x_{(k+1)}]$, we compute

$$M_m = \frac{1}{n} \cdot \sum_{i=1}^{n} (x_i - E)^m.$$

- The largest of thus computed values M_m is the desired upper bound $\overline{M}_m$.

In particular, for $\mathbf{x}_1 = \mathbf{x}_2 = \mathbf{x}_3 = [-1, 1]$, $m = 4$, and $C = 4$, we get the correct bound $\overline{M}_4 = 32/27 \approx 1.19$.

5 Fourth Result: Computing $\underline{M}_m$ and $\overline{M}_m$ for Odd m in Cubic Time (Case When Intervals Do Not Group Together)

Theorem 4. *For every odd m and for every $C \ge 2$, there exists an algorithm that computes $\underline{M}_m$ in cubic time when the input intervals $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are C-wise disjoint.*

The algorithm for computing $\underline{M}_m$ is as follows:

- First, we sort all $2n$ values $\underline{x}_i$, $\overline{x}_i$ into a sequence $x_{(1)} \le x_{(2)} \le \ldots \le x_{(2n)}$. This sequence divides the real line into $2n+1$ segments $[x_{(k)}, x_{(k+1)}]$, where $k = 0, \ldots, 2n$, $x_{(0)} \stackrel{\text{def}}{=} -\infty$, and $x_{(2n+1)} \stackrel{\text{def}}{=} +\infty$.
- For each pair of segments $[x_{(k)}, x_{(k+1)}]$ and $[x_{(l)}, x_{(l+1)}]$, $k \le l$, we do the following:
  - First, we describe several combinations $(x_1, \ldots, x_n)$ as linear functions of α^- and α^+ as follows:
    * if $\overline{x}_i < x_{(k)}$, we take $x_i = \underline{x}_i$;

* if $\underline{x}_i > x_{(k+1)}$, we take $x_i = \overline{x}_i$;
* for all other indices i (there are $\le 2C$ of them), we consider all possible combinations of $x_i = \underline{x}_i$, $x_i = \overline{x}_i$, $x_i = \alpha^-$ (if $[x_{(k)}, x_{(k+1)}] \subseteq \mathbf{x}_i$) and $x_i = \alpha^+$ (if $[x_{(l)}, x_{(l+1)}] \subseteq \mathbf{x}_i$).

As a result, we get $\le 4^{2C}$ different combinations.

– For each resulting combination $(x_1, \ldots, x_n)$, we find the expression for E as the average of all the values x_i. Then, we substitute the expressions for x_i and E into the system of equations

$$\frac{1}{n} \cdot \sum_{i=1}^{n} (x_i - E)^{m-1} = (E - \alpha^-)^{m-1};$$

$$\frac{1}{n} \cdot \sum_{i=1}^{n} (x_i - E)^{m-1} = (\alpha^+ - E)^{m-1}.$$

We compute all solutions of this system of polynomial equations with unknowns α^- and α^+.

– For each of the solutions for which $\alpha^- \in [x_{(k)}, x_{(k+1)}]$ and $\alpha^+ \in [x_{(l)}, x_{(l+1)}]$, we substitute the corresponding values α^- and α^+ into the formulas for x_i and E, thus, we compute x_i and E; based on these values, we compute

$$M_m = \frac{1}{n} \cdot \sum_{i=1}^{n} (x_i - E)^m.$$

- The smallest of thus computed values M_m is the desired lower bound $\underline{M}_m$.

Theorem 5. *For every odd m and for every $C \ge 2$, there exists an algorithm that computes $\overline{M}_m$ in cubic time when the input intervals $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are C-wise disjoint.*

Since m is odd, the m-th central moment of the values $x_1, \ldots, x_n$ is equal to minus the m-th moment of the values $-x_1, \ldots, -x_n$. Turning to $-x_i$ changes largest and smallest values and vice versa. Thus, to compute $\overline{M}_m$ for the intervals $\mathbf{x}_i = [\underline{x}_i, \overline{x}_i]$, it is sufficient to compute the lower bound $\underline{M}_m$ for the intervals $-\mathbf{x}_i = [-\overline{x}_i, -\underline{x}_i]$, and then change the sign of the resulting bound. Since we can use the above cubic-time algorithm to compute $\underline{M}_m$, we thus get a cubic-time algorithm for computing $\overline{M}_m$.

In particular, for $\mathbf{x}_1 = \mathbf{x}_2 = \mathbf{x}_3 = [-1, 1]$, $m = 3$, and $C = 4$, we get the bounds $\overline{M}_3 = 80/81 \approx 0.988$ and $\overline{M}_3 = -80/81$. The largest value $\overline{M}_3$ is attained when $x_1 = x_2 = 1$ and $x_3 = -1$; the smallest value $\underline{M}_3$ is attained, e.g., when $x_1 = x_2 = -1$ and $x_3 = 1$.

6 Proofs

Proof of Theorem 1. The central moment M_m is a continuous function of n variables; thus, its smallest possible value on a compact box $\mathbf{x}_1 \times \ldots \times \mathbf{x}_n$ is attained at some point $(x_1^{(0)}, \ldots, x_n^{(0)})$. Since the function M_m is also smooth, for each variable i for which the interval $\mathbf{x}_i$ is non-degenerate (i.e., of finite width), the minimum is attained either when x_i is inside the corresponding interval $(\underline{x}_i, \overline{x}_i)$ and the derivative $\partial M_m/\partial x_i$ is equal to 0, or when $x_i^{(0)}$ coincides with one of the endpoints of this interval. To be more precise, we must have one of the following three situations:

- either $x_i^{(0)} \in (\underline{x}_i, \overline{x}_i)$ and $\partial M_m/\partial x_i = 0$;
- or $x_i^{(0)} = \underline{x}_i$ and $\partial M_m/\partial x_i \geq 0$;
- or $x_i^{(0)} = \overline{x}_i$ and $\partial M_m/\partial x_i \leq 0$.

Differentiating M_m w.r.t. x_i, and taking into consideration that $\partial E/\partial x_i = 1/n$, we conclude that

$$\frac{\partial M_m}{\partial x_i} = \frac{1}{n} \cdot m \cdot (x_i - E)^{m-1} + \frac{1}{n} \cdot \sum_{j=1}^{n} m \cdot (x_j - E)^{m-1} \cdot \left(-\frac{1}{n}\right).$$

The sum in this formula is proportional to the $(m-1)$-st central moment M_{m-1}, so the above formula can be simplified into:

$$\frac{\partial M_m}{\partial x_i} = \frac{m}{n} \cdot ((x_i - E)^{m-1} - M_{m-1}). \tag{2}$$

Due to this formula:

- if $\partial M_m/\partial x_i = 0$, then $(x_i - E)^{m-1} = M_{m-1}$, hence $x_i = \alpha \stackrel{\text{def}}{=} E + M_{m-1}^{1/(m-1)}$;
- if $\partial M_m/\partial x_i \geq 0$, then $(x_i - E)^{m-1} \geq M_{m-1}$, hence (since the function $z \to z^{1/(m-1)}$ is increasing for even m) $x_i \geq \alpha$;
- if $\partial M_m/\partial x_i \leq 0$, then $(x_i - E)^{m-1} \leq M_{m-1}$, hence $x_i \leq \alpha$.

Therefore, the above conditions on $x_i^{(0)}$ can be reformulated as follows:

- either $x_i^{(0)} \in (\underline{x}_i, \overline{x}_i)$ and $x_i^{(0)} = \alpha$; in this case, $\underline{x}_i < \alpha < \overline{x}_i$;
- or $x_i^{(0)} = \underline{x}_i$ and $x_i^{(0)} = \underline{x}_i \geq \alpha$;
- or $x_i^{(0)} = \overline{x}_i$ and $x_i^{(0)} = \overline{x}_i \leq \alpha$.

Hence, once we know α, we can determine all n values $x_i^{(0)}$ as follows:

- if $\overline{x}_i \le \alpha$, then we cannot have the first case (when $\alpha < \overline{x}_i$) or the second case (when $\alpha \le \underline{x}_i$ hence $\alpha < \overline{x}_i$); therefore, we can only have the third case, when $x_i^{(0)} = \overline{x}_i$;
- similarly, if $\alpha \le \underline{x}_i$, then we must have $x_i^{(0)} = \underline{x}_i$;
- finally, when $\underline{x}_i < \alpha < \overline{x}_i$, then we must have $x_i^{(0)} = \alpha$.

The only thing that remains is to find α. Once we know to which of $2n+1$ segments $[x_{(k)}, x_{(k+1)}]$ the value α belongs, we can uniquely describe all the values x_i as linear functions of α, and then define α from the condition that $\alpha = E + M_{m-1}^{1/(m-1)}$, i.e., equivalently, that $M_{m-1} = (\alpha - E)^{m-1}$. This is exactly what our algorithm does.

This proves that our algorithm is correct. To complete the proof, we must also show that this algorithm requires quadratic time.

Indeed, sorting requires $O(n \cdot \log(n))$ steps, and the rest of the algorithm requires linear time ($O(n)$) for each of $2n$ segments, i.e., the total quadratic time. The theorem is proven.

Proof of Theorem 2. Similarly to the proof of Theorem 1, we an conclude that for every i, the maximum of M_m over the interval $\mathbf{x}_i$ is attained either inside the interval (when the partial derivative is 0) or at one of the endpoint of this interval. Thus, to prove that our algorithm is correct, we must show that the maximum of M_m cannot be attained for $x_i \in (\underline{x}_i, \overline{x}_i)$, when $\partial M_m/\partial x_i = 0$. Indeed, in the maximum point, the second derivative $\partial^2 M_m/\partial x_i^2$ must be non-positive. In the proof of Theorem 1, we have already derived an explicit formula (2) for $\partial M_m/\partial x_i$. The formula (2) describes this derivative in terms of M_{m-1}, so when we differentiate both sides of the formula (2), we can use the same expression for the derivative of M_{m-1}. As a result, we get the following:

$$\frac{\partial^2 M_m}{\partial x_i^2} = \frac{m}{n} \cdot T,$$

where

$$T = (m-1)\cdot(x_i - E)^{m-2} - (m-1)\cdot(x_i-E)^{m-2}\cdot\frac{1}{n} -$$
$$\frac{m-1}{n}\cdot(x_i-E)^{m-2} + \frac{m-1}{n}\cdot M_{m-2} =$$
$$\frac{m-1}{n}\cdot\left((n-2)\cdot(x_i-E)^{m-2} + M_{m-2}\right).$$

In the trivial case of $n = 1$, all central moments are 0. When $n \ge 2$, both terms are non-negative, so the second derivative is non-negative. We know that the second derivative must be non-positive, so it must be equal to 0. Since the sum of two non-negative numbers is equal to 0, both numbers are equal to 0, in particular,

$$M_{m-2} = \frac{1}{n} \cdot \sum_{i=1}^{n} (x_i - E)^{m-2} = 0.$$

Therefore, all the values x_i are identically equal to E. In this case, $M_m = 0$, so this cannot be where the largest possible value of M_m is attained. This contradiction shows that the maximum cannot be attained inside the interval $\mathbf{x}_i$, hence it attained at the endpoints. The theorem is proven.

Proof of Theorem 3. We have already proven, in Theorem 2, that maximum can only be attained at one of the endpoints of the interval $[\underline{x}_i, \overline{x}_i]$, i.e., when $x_i^{(0)} = \underline{x}_i$ or $x_i^{(0)} = \overline{x}_i$. Hence, for each i, we have one of the following two situations:

- either $x_i^{(0)} = \underline{x}_i$ and $\partial M_m/\partial x_i \le 0$;
- or $x_i^{(0)} = \overline{x}_i$ and $\partial M_m/\partial x_i \ge 0$.

We already know, from the proof of Theorem 1, that the condition $\partial M_m/\partial x_i \le 0$ is equivalent to $x_i \le \alpha$, and the condition $\partial M_m/\partial x_i \ge 0$ is equivalent to $x_i \ge \alpha$. Thus, the above two situations can be reformulated a follows:

- either $x_i^{(0)} = \underline{x}_i$ and $x_i^{(0)} = \underline{x}_i \le \alpha$;
- or $x_i^{(0)} = \overline{x}_i$ and $x_i^{(0)} = \overline{x}_i \ge \alpha$.

Hence:

- if $\overline{x}_i < \alpha$, then we cannot have the second case (when $\overline{x}_i \ge \alpha$) and therefore, we can only have the first case, when $x_i^{(0)} = \underline{x}_i$;
- similarly, if $\alpha < \underline{x}_i$, then we must have $x_i^{(0)} = \overline{x}_i$.

The only case when the knowledge of α does not help us determine x_i is the case when $\underline{x}_i \le \alpha \le \overline{x}_i$, i.e., when $\alpha \in \mathbf{x}_i$.

Since intervals are C-wise disjoint, for each α, there can be no more than C such intervals, so we can try all 2^C possible assignments for each segment. In other words, the time increases by a constant ($\le 2^C$) over the running time of the algorithm described in Theorem 1. This justifies the algorithm and proves that it runs in quadratic time.

Proof of Theorem 4. Similarly to the proof of Theorem 1, we conclude that for the point where the function M_m attains its minimum, we have:

- either $x_i^{(0)} \in (\underline{x}_i, \overline{x}_i)$ and $\partial M_m/\partial x_i = 0$;

- or $x_i^{(0)} = \underline{x}_i$ and $\partial M_m/\partial x_i \ge 0$;
- or $x_i^{(0)} = \overline{x}_i$ and $\partial M_m/\partial x_i \le 0$.

Here, the derivative $\partial M_m/\partial x_i$ is described by the same formula (2) as in the proof of Theorem 1. The difference is that m is now odd, so:

- if $\partial M_m/\partial x_i = 0$, then $(x_i - E)^{m-1} = M_{m-1}$, hence $|x_i - E| = M_{m-1}^{1/(m-1)}$, so either x_i is equal to $\alpha^- \stackrel{\text{def}}{=} E - M_{m-1}^{1/(m-1)}$, or x_i is equal to $\alpha^+ \stackrel{\text{def}}{=} E + M_{m-1}^{1/(m-1)}$;
- if $\partial M_m/\partial x_i \ge 0$, then $(x_i - E)^{m-1} \ge M_{m-1}$, hence $|x_i - E| \ge M_{m-1}^{1/(m-1)}$, so $x_i \le \alpha^-$ or $x_i \ge \alpha^+$;
- if $\partial M_m/\partial x_i \le 0$, then $(x_i - E)^{m-1} \le M_{m-1}$, hence $|x_i - E| \le M_{m-1}^{1/(m-1)}$, so $\alpha^- \le x_i \le \alpha^+$.

Therefore, the above conditions on $x_i^{(0)}$ can be reformulated as follows:

- in the first case, $x_i^{(0)} \in (\underline{x}_i, \overline{x}_i)$ and either $x_i^{(0)} = \alpha^-$ or $x_i^{(0)} = \alpha^+$; in this case, either $\underline{x}_i < \alpha^- < \overline{x}_i$ or $\underline{x}_i < \alpha^+ < \overline{x}_i$;
- in the second case, $x_i^{(0)} = \underline{x}_i$ and either $x_i^{(0)} = \underline{x}_i \le \alpha^-$ or $x_i^{(0)} = \underline{x}_i \ge \alpha^+$;
- in the third case, $x_i^{(0)} = \overline{x}_i$ and $\alpha^- \le x_i^{(0)} = \overline{x}_i \le \alpha^+$.

Hence, once we know α^- and α^+, we can determine (at least some) some values $x_i^{(0)}$ as follows:

- if $\alpha^- < \underline{x}_i < \overline{x}_i < \alpha^+$, then we cannot have the first case (when $\alpha^- > \underline{x}_i$ or $\alpha^+ < \overline{x}_i$), and we cannot have the second case (when $\underline{x}_i \le \alpha^-$ or $\underline{x}_i \ge \alpha^+$); therefore, we can only have the third case, when $x_i^{(0)} = \overline{x}_i$;
- if $\alpha^+ \le \underline{x}_i$, then we cannot have the first case (when either $\alpha^+ > \underline{x}_i$, or $\alpha^- > \underline{x}_i$ hence $\alpha^+ > \underline{x}_i$), and we cannot have the third case (when $\overline{x}_i \le \alpha^+$ hence $\underline{x}_i < \alpha^+$); therefore, we can only have the second case, when $x_i^{(0)} = \underline{x}_i$;
- similarly, if $\overline{x}_i < \alpha^-$, then we cannot have the first case (when either $\alpha^- < \underline{x}_i$, or $\alpha^+ < \overline{x}_i$ hence $\alpha^- < \overline{x}_i$), and we cannot have the third case (when $\alpha^- \le \overline{x}_i$); therefore, we can only have the second case, when $x_i^{(0)} = \underline{x}_i$.

We have described all the cases in which neither of the two auxiliary values α^- and α^+ is in the interval $\mathbf{x}_i$; in all these cases, we can uniquely determine the value $x_i^{(o)}$. The only cases when we cannot uniquely determine the value $x_i^{(0)}$ are the cases when either α^- or α^+ is within the interval $\mathbf{x}_i$.

Once we choose segments that contain α^- and α^+, we have no more than C intervals $\mathbf{x}_i$ that contain α^- and no more than C intervals that contain α^+. Thus, for the remaining $\geq n - 2C$ indices i, we can uniquely determine x_i. For the $2C$ indices, we try all possible combinations. This is exactly what we do in our algorithm. Thus, the algorithm is indeed correct.

The algorithm requires linear time $O(n)$ for each pair of segments; there are $O(n^2)$ pairs of segments, hence the algorithm requires cubic time. The theorem is proven.

Acknowledgments

This work was supported in part by NASA under cooperative agreement NCC5-209, by the Future Aerospace Science and Technology Program (FAST) Center for Structural Integrity of Aerospace Systems, effort sponsored by the Air Force Office of Scientific Research (AFOSR), Air Force Materiel Command, USAF, under grant F49620-00-1-0365, by NSF grants EAR-0112968 and EAR-0225670, by by Small Business Innovation Research grant 9R44CA81741 to Applied Biomathematics from the National Cancer Institute (NCI), a component of the National Institutes of Health (NIH), by a research grant from Sandia National Laboratories (SNL), as part of the Department of Energy Accelerated Strategic Computing Initiative (ASCI), by IEEE/ACM SC2001 and SC2002 Minority Serving Institutions Participation Grants, and by a grant from the Army Research Laboratories (ARL).

The opinions expressed herein are those of the author and not necessarily those of the NASA, AFOSR, NSF, NCI, NIH, SNL, or ARL.

The authors are greatly thankful to the anonymous referees for valuable suggestions.

References

[1] D. Berleant, "Automatically verified arithmetic with both intervals and probability density functions", *Interval Computations*, No. 2, pp. 48–70.

[2] D. Berleant, "Automatically verified arithmetic on probability distributions and intervals", In: R. B. Kearfott and V. Kreinovich (eds.) *Applications of Interval Computations*, Kluwer, Dordrecht.

[3] D. Berleant and C. Goodman-Strauss, "Bounding the results of arithmetic operations on random variables of unknown dependency using intervals", *Reliable Computing*, 1998, Vol. 4, No. 2, pp. 147–165.

[4] D. Berleant, L. Xie, and J. Zhang, "Statool: A Tool for Distribution Envelope Determination (DEnv), an Interval-Based Algorithm for Arithmetic on Random Variables", *Reliable Computing*, 2003, Vol. 9, No. 2, pp. 91–108.

[5] G. Casella and R. Berger, *Statistical inference*, Duxbury Press, Belmont, California, 2001.

[6] S. Ferson, *RAMAS Risk Calc 4.0: Risk Assessment with Uncertain Numbers*, CRC Press, Boca Raton, Florida, 2002.

[7] S. Ferson, L. Ginzburg, V. Kreinovich, L. Longpré, and M. Aviles, "Computing Variance for Interval Data is NP-Hard", *ACM SIGACT News*, 2002, Vol. 33, No. 2, pp. 108–118.

[8] S. Ferson, L. Ginzburg, V. Kreinovich, L. Longpré, and M. Aviles, "Exact Bounds on Sample Variance of Interval Data", *Extended Abstracts of the 2002 SIAM Workshop on Validated Computing*, Toronto, Canada, May 23–25, 2002, pp. 67–69.

[9] S. Ferson, D. Myers, and D. Berleant, *Distribution-free risk analysis: I. Range, mean, and variance*, Applied Biomathematics, Technical Report, 2001.

[10] J. L. Horowitz, C. F. Manski, M. Ponomareva, and J. Stoye, "Computation of Bounds on Population Parameters When the Data Are Incomplete", *Reliable Computing*, 2003, Vol. 9, No. 6, pp. 419–440.

[11] L. Jaulin, M. Kieffer, O. Didrit, and E. Walter, *Applied Interval Analysis*, Springer-Verlag, Berlin, 2001.

[12] V. P. Kuznetsov, *Interval Statistical Models*, Radio i Svyaz, Moscow, 1991 (in Russian).

[13] W. A. Lodwick and K. D. Jamison, "Estimating and Validating the Cumulative Distribution of a Function of Random Variables: Toward the Development of Distribution Arithmetic", *Reliable Computing*, 2003, Vol. 9, No. 2, pp. 127–141.

[14] C. Manski, *Partial Identification of Probability Distributions*, Springer-Verlag, New York, 2003.

[15] R. E. Moore, *Interval Analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1966.

[16] R. E. Moore and W. A. Lodwick, "Interval Analysis and Fuzzy Set Theory", *Fuzzy Sets and Systems*, 2003, Vol. 135, No. 1, pp. 5–9.

[17] P. Nivlet, F. Fournier, and J. Royer, "A new methodology to account for uncertainties in 4-D seismic interpretation", *Proceedings of the 71st Annual International Meeting of the Society of Exploratory Geophysics SEG'2001*, San Antonio, Texas, September 9–14, 2001, pp. 1644–1647.

[18] P. Nivlet, F. Fournier, and J. Royer, "Propagating interval uncertainties in supervised pattern recognition for reservoir characterization", *Proceedings of the 2001 Society of Petroleum Engineers Annual Conference SPE'2001*, New Orleans, Louisiana, September 30–October 3, 2001, paper SPE-71327.

[19] H. Regan, S. Ferson, and D. Berleant, "Equivalence of five methods for bounding uncertainty", *Journal of Approximate Reasoning* (to appear).

[20] N. C. Rowe, "Absolute bounds on the mean and standard deviation of transformed data for constant-sign-derivative transformations", *SIAM Journal of Scientific Statistical Computing*, 1988, Vol. 9, pp. 1098–1113.

[21] H. M. Wadsworth, Jr. (eds.), *Handbook of statistical methods for engineers and scientists*, McGraw-Hill Publishing Co., New York, 1990.

[22] P. Walley, *Statistical Reasoning with Imprecise Probabilities*, Chapman & Hall, N.Y., 1991.

[23] R. Williamson and T. Downs, "Probabilistic arithmetic I: numerical methods for calculating convolutions and dependency bounds", *International Journal of Approximate Reasoning*, 1990, Vol. 4, pp. 89–158.