Order and Percentiles

Andrew Nobel

October, 2024

Minima and Maxima

Multiplication, Addition, and Signs

Recall: For any numbers $a, b \in \mathbb{R}$

(1) If $a, b \geq 0$ or $a, b \leq 0$ then $ab \geq 0$

(2) If $a \geq 0$ and $b \leq 0$ or vice-versa then $ab \leq 0$

(3) If $a, b \geq 0$ then $a + b \geq 0$

(4) If $a, b \leq 0$ then $a + b \leq 0$.

Note: (1)-(4) continue to hold if we replace $\leq$ and $\geq$ by $<$ and $>$, respectively

The Usual Order Relation

Definition: For $a, b \in \mathbb{R}$ write $a \leq b$ if $(b - a) \geq 0$ and $a < b$ if $(b - a) > 0$

Basic Properties

1. If $a \leq b$ and $b \leq a$ then $a = b$
2. If $a \leq b$ then $-b \leq -a$
3. If $a \leq b$ and $c \leq d$ then $a + c \leq b + d$
4. If $0 \leq a \leq b$ and $0 \leq c \leq d$ then $ac \leq bd$

Note: (2)-(4) continue to hold if we replace $\leq$ by $<$

Maxima and Minima of Finite Sequences

Definition: Let $A = \{a_1, \ldots, a_n\} \subseteq \mathbb{R}$ be a finite set

- $\max\{a_1, \ldots, a_n\}$ is any element a_j such that $a_i \leq a_j$ for $i = 1, \ldots, n$
- $\min\{a_1, \ldots, a_n\}$ is any element a_j such that $a_i \geq a_j$ for $i = 1, \ldots, n$

Other Notation

- $\max_{1 \leq i \leq n} a_i$ or simply $\max_i a_i$
- $\min_{1 \leq i \leq n} a_i$ or simply $\min_i a_i$

Maxima and Minima, cont.

Basic Properties: Let $a_1, \ldots, a_n \in \mathbb{R}$ and $b_1, \ldots, b_n \in \mathbb{R}$ be finite sequences

1. $\min_i a_i \;\leq\; a_j \leq \max_i a_i$ for $j = 1, \ldots, n$
2. $-\min_i a_i \;=\; \max_i(-a_i)$ and $-\max_i a_i \;=\; \min_i(-a_i)$
3. If $a_i \leq b_i$ for each i, then $\max_i a_i \;\leq\; \max_i b_i$ and $\min_i a_i \;\leq\; \min_i b_i$
4. If $c \geq 0$ and b are constants then $c \max_i a_i + b \;=\; \max_i(c\, a_i + b)$
5. $\max_i(a_i + b_i) \;\leq\; \max_i a_i + \max_i b_i$
6. $\min_i(a_i + b_i) \;\geq\; \min_i a_i + \min_i b_i$
7. $|\max_i a_i - \max_i b_i| \;\leq\; \max_i |a_i - b_i|$
8. $|\min_i a_i - \min_i b_i| \;\leq\; \max_i |a_i - b_i|$

Maxima and Minima of Nested Sets

Notation: Let $A \subseteq \mathbb{R}$ be finite. Define

1. $\min(A) := \min\{x : x \in A\}$ (smallest number in A)
2. $\max(A) := \max\{x : x \in A\}$ (largest number in A)

Fact: Let $A, B \subseteq \mathbb{R}$ be finite sets. If $A \subseteq B$ then

1. $\min(A) \geq \min(B)$
2. $\max(A) \leq \max(B)$

Order Statistics

Order Statistics

Def'n: The order statistics of a finite data set $x_1, \ldots, x_n \in \mathbb{R}$ are the ordered values

$$x_{(1)} \leq \cdots \leq x_{(n)}$$

where $x_{(j)}$ is jth smallest data point. Note that $x_{(j)}$ depends on the *entire* data set

Stochastic Setting: The order statistics of iid observations $X_1, \ldots, X_n$ are written as

$$X_{(1)} \leq \cdots \leq X_{(n)}$$

where $X_{(j)}$ is the jth smallest observation. Note that

- $X_{(1)}, \ldots, X_{(n)}$ are a random reordering of $X_1, \ldots, X_n$
- $X_{(1)} = \min(X_1, \ldots, X_n)$ and $X_{(n)} = \max(X_1, \ldots, X_n)$

Order Statistics, cont.

Fact: Let $X_1, \ldots, X_n$ be iid with expectation $\mathbb{E}X$

1. The order statistics $X_{(1)}, \ldots, X_{(n)}$ are *not* independent
2. For each fixed i and j we have $\mathbb{P}(X_{(i)} = X_j) = 1/n$
3. If $i \leq j$ then $\mathbb{E}X_{(i)} \leq \mathbb{E}X_{(j)}$
4. $\mathbb{E}X_{(1)} \leq \mathbb{E}X$ and $\mathbb{E}X_{(n)} \geq \mathbb{E}X$
5. $\sum_{i=1}^{n} X_{(i)} = \sum_{j=1}^{n} X_j$
6. $\mathbb{E}(\sum_{i=1}^{n} X_{(i)}) = n\mathbb{E}X$

Fact: If X_i are continuous, then $\mathbb{P}(X_i = X_j) = 0$ for all i, j. Thus there are no ties, and the order statistics $X_{(1)} < \cdots < X_{(n)}$ are strictly increasing

Densities of Order Statistics

Fact: Let $X_1, \ldots, X_n$ be iid with CDF F and density f

1. The rth order statistic $X_{(r)}$ has density

$$f_{(r)}(u) = r\binom{n}{r} F(u)^{r-1}(1 - F(u))^{n-r} f(u)$$

2. The maximum $X_{(n)} = \max\{X_i\}$ has density

$$f_{(n)}(u) = nF^{n-1}(u)f(u)$$

3. The order statistics $X_{(1)}, \cdots, X_{(n)}$ have *joint* density

$$g(x_1, \ldots, x_n) = n! f(x_1) \cdots f(x_n) \mathbb{I}(x_1 < \ldots < x_n)$$

Percentile Function

Percentiles

Overview: Let X have CDF $F(x) = \mathbb{P}(X \leq x)$. For $\alpha \in (0, 1)$ a number x is said to be an α-percentile of X if

$$\mathbb{P}(X \leq x) \; = \; \alpha$$

In other words, x is the solution to $F(x) = \alpha$. Immediate issues

- If F has jumps, no solution x may exist
- If F has has flat bits, there may be many solutions x

Inverse of the CDF: If F is invertible at α then

$$\mathbb{P}(X \leq F^{-1}(\alpha)) = F(F^{-1}(\alpha)) = \alpha$$

In this case $F^{-1}(\alpha)$ is the α-percentile of X. Goal: a generalization of F^{-1}

Percentile Function

Definition: Let X have CDF F. The percentile function $\varphi : (0,1) \to \mathbb{R}$ of X is given by

$$\varphi(u) = \min\{x : F(x) \geq u\}$$

Key fact: Right continuity of F ensures that $\{x : F(x) \geq u\} = [\varphi(u), \infty)$

Basic Properties

1. If $u \leq v$ then $\varphi(u) \leq \varphi(v)$ (φ is non-decreasing)
2. $\varphi(u) \leq x$ iff $u \leq F(x)$
3. $F(\varphi(u)) \geq u$, with equality if F is continuous

Random Number Generation Using Percentile Function

Fact: Let X have percentile function φ. If $U \sim$ Unif$(0,1)$ then $\varphi(U) \stackrel{d}{=} X$

Application: To simulate a random variable $X \sim F$ proceed as follows

- Calculate or approximate the percentile function φ of X
- Generate a uniform random variable U (easy using random coin flips)
- Output $X = \varphi(U)$ (to some precision)

Partial converse: If X has a continuous CDF F then $F(X) \sim$ Unif$(0,1)$