Order and metric geometry compatible stochastic processing

A traditional random variable $X$ is a function that maps from a stochastic process to the real line. Here, "real line" refers to the structure $(\mathbb{R}, \leq, |x-y|)$, where $\mathbb{R}$ is the set of real numbers, $\leq$ is the standard linear order relation on $\mathbb{R}$, and $d(x,y)=|x-y|$ is the usual metric on $\mathbb{R}$. The traditional expectation value $E(X)$ of $X$ is then often a poor choice of a statistic when the stochastic process that $X$ maps from is a structure other than the real line or some substructure of the real line. If the stochastic process is a structure that is not linearly ordered (including structures totally unordered) and/or has a metric space geometry very different from that induced by the usual metric, then statistics such as $E(X)$ are often of poor quality with regards to qualitative intuition and quantitative variance (expected error) measurements. For example, the traditional expected value of a fair die is $E(X)=(1/6)(1+2+...+6)=3.5$. But this result has no relationship with reality or with intuition because the result implies that we expect the value of $[ooo]$ (die face value "3") or $[oooo]$ (dice face value "4") more than we expect the outcome of say $[o]$ or $[oo]$. The fact is, that for a fair die, we would expect any pair of values equally. The reason for this is that the values of the face of a fair die are merely symbols with no order, and with no metric geometry other than the discrete metric geometry. On a fair die, $[oo]$ is not greater or less than $[o]$; rather $[oo]$ and $[o]$ are simply symbols without order. Moreover, $[o]$ is not "closer" to $[oo]$ than it is to $[ooo]$; rather, $[o]$, $[oo]$, and $[ooo]$ are simply symbols without any inherit order or metric geometry. This paper proposes an alternative statistical system, based somewhat on graph theory, that takes into account the order structure and metric geometry of the underlying stochastic process.
Order and metric geometry compatible stochastic processing

DANIEL J. GREENHOE

Abstract: A traditional random variable \( X \) is a function that maps from a stochastic process to the real line. Here, “real line” refers to the structure \((\mathbb{R}, |\cdot|, \leq)\), where \( \mathbb{R} \) is the set of real numbers, \( \leq \) is the standard linear order relation on \( \mathbb{R} \), and \( d(x, y) \triangleq |x - y| \) is the usual metric on \( \mathbb{R} \). The traditional expectation value \( E(X) \) of \( X \) is then often a poor choice of a statistic when the stochastic process that \( X \) maps from is a structure other than the real line or some substructure of the real line. If the stochastic process is a structure that is not linearly ordered (including structures totally unordered) and/or has a metric space geometry very different from that induced by the usual metric, then statistics such as \( E(X) \) are often of poor quality with regards to qualitative intuition and quantitative variance (expected error) measurements. For example, the traditional expected value of a fair die is \( E(X) = \frac{1}{6}(1 + 2 + \cdots + 6) = 3.5 \). But this result has no relationship with reality or with intuition because the result implies that we expect the value of \( \heartsuit \) or \( \diamondsuit \) more than we expect the outcome of say \( \clubsuit \) or \( \spadesuit \). The fact is, that for a fair die, we would expect any pair of values equally. The reason for this is that the values of the face of a fair die are merely symbols with no order, and with no metric geometry other than the discrete metric geometry. On a fair die, \( \heartsuit \) is not greater or less than \( \clubsuit \); rather \( \heartsuit \) and \( \clubsuit \) are simply symbols without order. Moreover, \( \spadesuit \) is not “closer” to \( \heartsuit \) than it is to \( \clubsuit \); rather, \( \spadesuit \), \( \heartsuit \), and \( \clubsuit \) are simply symbols without any inherit order or metric geometry. This paper proposes an alternative statistical system, based somewhat on graph theory, that takes into account the order structure and metric geometry of the underlying stochastic process.

2010 Mathematics Subject Classification 60B99, 60G05 (primary); 60E15, 60G20, 62C20, 62F03, 62P10, 92D20 (secondary)

Keywords: stochastic processes, stochastic signal processing, statistical signal processing, expectation, expectation operator

Contents

Table of Contents

1 Background: Relations

2 Background: Order

  2.1 Graph Theory .......................................................... 4
  2.2 Order relations .......................................................... 4

  2.2.1 Definitions ............................................................. 4
  2.2.2 Bounds on ordered sets .............................................. 5

2.3 Lattices ................................................................. 6

3 Background: power means ............................................ 7

Wednesday 18th February, 2015
# Background: Metric space

4 Background: Metric space

4.1 Definitions ................................................................. 8
4.2 Metric transforms ......................................................... 8
  4.2.1 Metric transforms on the domains of metrics .......... 8
  4.2.2 Metric preserving functions ................................. 9
  4.2.3 Product metrics .................................................. 11
4.3 Examples ................................................................. 13

5 Background: random processes

6 Ordered metric spaces ....................................................... 16
  6.1 Definitions ............................................................. 16
  6.2 Examples ............................................................... 17

7 Monotone functions on ordered sets ..................................... 17
  7.1 Definitions ............................................................. 17
  7.2 Properties ............................................................... 18

8 Outcome subspaces .......................................................... 22
  8.1 Definitions ............................................................. 22
  8.2 Examples ............................................................... 23

9 Random variables on outcome subspaces .............................. 42
  9.1 Definitions ............................................................. 42
  9.2 Properties ............................................................... 42
  9.3 Problem statement ................................................... 45
  9.4 Examples ............................................................... 45
    9.4.1 Fair die examples .............................................. 45
    9.4.2 Real die examples ............................................. 49
    9.4.3 Spinner examples ............................................. 52
    9.4.4 Pseudo-random number generator (PRNG) examples .. 55
    9.4.5 Genomic signal processing (GSP) examples .......... 60

10 Operations on outcome subspaces ..................................... 64
    10.1 Summation ........................................................... 64
    10.2 Multiplication ...................................................... 71
    10.3 Metric transformation ........................................... 72

Bibliography ............................................................... 74

## 1 Background: Relations

Definition 1.1 Let \( \mathbb{R} \) be the set of real numbers. Let \( \mathbb{R}^+ \triangleq \{ x \in \mathbb{R} | x \geq 0 \} \) be the set of non-negative real numbers. Let \( \mathbb{R}^+ \triangleq \{ x \in \mathbb{R} | x > 0 \} \) be the set of positive real numbers. Let \( \mathbb{R}^* \triangleq \mathbb{R} \cup \{ -\infty, \infty \} \) be the set of extended real numbers.\(^1\) Let \( \mathbb{Z} \) be the set of integers. Let \( \mathbb{N} \triangleq \{ n \in \mathbb{Z} | n \geq 1 \} \) be the set of integers.

\(^1\) \([94]\), pages 385–388, \(\langle\text{Appendix A}\rangle\)
natural numbers. Let \( \mathbb{Z}^* \triangleq \mathbb{Z} \cup \{-\infty, \infty\} \) be the extended set of integers.

**Definition 1.2** Let \( X \) be a set. The quantity \( 2^X \) is the power set of \( X \) such that
\[
2^X \triangleq \{ A \subseteq X \} \quad \text{(the set of all subsets of } X)\.
\]

**Definition 1.3** Let \( X \) and \( Y \) be sets. The Cartesian product \( X \times Y \) of \( X \) and \( Y \) is the set \( X \times Y \triangleq \{(x, y) \mid x \in X \text{ and } y \in Y\} \). An ordered pair \((x, y)\) on \( X \) and \( Y \) is any element in \( X \times Y \). A relation \( @ \) on \( X \) and \( Y \) is any subset of \( X \times Y \) such that \( @ \subseteq X \times Y \). The set \( 2^{X\times Y} \) is the set of all relations in \( X \times Y \). A relation \( f \in 2^{X\times Y} \) is a function if \((x, y_1) \in f \) and \((x, y_2) \in f \implies y_1 = y_2 \). The set \( Y^X \) is the set of all functions in \( 2^{X\times Y} \).

**Definition 1.4** Let \( @ \in 2^{X\times Y} \) be a relation (Definition 1.3 page 3).

- The domain of \( @ \) is \( D(\@) \triangleq \{ x \in X \mid \exists y \text{ such that } (x, y) \in @ \} \).
- The image set of \( @ \) is \( I(\@) \triangleq \{ y \in Y \mid \exists x \text{ such that } (x, y) \in @ \} \).
- The null space of \( @ \) is \( N(\@) \triangleq \{ x \in X \mid (x, 0) \in @ \} \).
- The range of \( @ \) is any set \( R(\@) \) such that \( I(\@) \subseteq R(\@) \).

**Definition 1.5** Let \( f \) be a function in \( Y^X \) and \( g \) a function in \( Z^Y \) (Definition 1.3 page 3). The composite function \( \star \) of \( g \) and \( f \) is \( g \circ f \triangleq \{(x, z) \in X \times Z \mid \exists x \in X \text{ such that } g[f(x)] = z \} \).

**Definition 1.6** The set function \( | A | \triangleq \begin{cases} \infty & \text{if } A \text{ is infinite} \\ |A| & \text{otherwise} \end{cases} \) \( \forall A \in 2^X \) is the cardinality of \( A \) such that
\[
| A | \triangleq \begin{cases} \infty & \text{if } A \text{ is infinite} \\ \text{number of elements in } A & \text{otherwise} \end{cases} \forall A \in 2^X
\]

**Definition 1.7** Let \( | X | \) be the cardinality (Definition 1.6 page 3) of a set \( X \).

The structure \( @ \) is the empty set, and is a set such that \( |@| = 0 \).

**Definition 1.8** Let \( f \) be a function in \( Y^X \) (Definition 1.3 page 3).

- \( f \) is surjective (also called onto) if \( f(X) = Y \).
- \( f \) is injective (also called one-to-one) if \( f(x) = f(y) \implies x = y \).
- \( f \) is bijective (also called one-to-one and onto) if \( f \) is both surjective and injective.

---

2 [74], page 4, [53], pages 26–30, [105], page 86, [65], page 10, [15], [14], page 7, [23], page 4, (\( Y^X \)); The notation \( Y^X \) and \( 2^{X\times Y} \) is motivated by the fact that for finite \( X \) and \( Y \), \( |Y^X| = |Y|^{|X|} \) and \( |2^{X\times Y}| = 2^{|X|\cdot|Y|} \).

3 [84], page 16, [65], page 7

4 [78], page 16, (1.2.15)

5 [87], page 8, (Definition 2.3: extended real-valued set function), [52], page 30, (§7. MEASURE ON RINGS)

6 [78], pages 14–15, [46], page 2, [23], page 5, [36], pages 16–17
2 Background: Order

2.1 Graph Theory

Definition 2.1 Let \(2^X\) be the set of all relations (Definition 1.3 page 3) on a set \(X\). The pair \((X, E)\) is a graph if \(E \subseteq 2^X\). A graph \((X, E)\) is **undirected** if \((x, y) \in E \implies (y, x) \in E\). A graph \((X, E)\) is **directed** if it is not undirected. A graph that is directed is a **directed graph**. A graph that is undirected is an **undirected graph**. The elements of \(X\) are the **vertices** and the ordered pairs of \(E\) are the **arcs** of a graph \((X, E)\). The element \(x\) is the **tail** and \(y\) is the **head** of an arc \((x, y)\).

Definition 2.2 The tuple \((X, E, d, w)\) is a **weighted graph** if \((X, E)\) is a graph (Definition 2.1 page 4), \(d(x, y)\) is a function in \((\mathbb{R}^+)^{X \times X}\) (Definition 1.3 page 3), and \(w(x)\) is a function in \((\mathbb{R}^+)^X\).

The function \(d\) is called the **edge weight**, and the function \(w\) is called the **vertex weight**.

Definition 2.3 Let \(G \triangleq (X, E, d, w)\) be a weighted graph (Definition 2.2 page 4). The **center** \(\mathcal{C}(G)\) of \(G\) is \(\mathcal{C}(G) \triangleq \arg\min_{x \in X} \max_{y \in X} d(x, y) w(y)\).

2.2 Order relations

2.2.1 Definitions

Definition 2.4 Let \(X\) be a set. A relation \(\leq\) is an **order relation** in \(2^X\) (Definition 1.3 page 3) if

1. \(x \leq x\)
2. \(x \leq y\) and \(y \leq z\) \(\implies\) \(x \leq z\) (reflexive) and
3. \(x \leq y\) and \(y \leq x\) \(\implies\) \(x = y\) (transitive) and
4. \(x \leq y\) and \(y \leq x\) \(\implies\) \(x = y\) (anti-symmetric).

An ordered set is the pair \((X, \leq)\). If \(\leq = \emptyset\) (Definition 1.7 page 3), then \((X, \leq)\) is an **unordered set**. The set \(X\) is called the **base set** of \((X, \leq)\). If \(x \leq y\) or \(y \leq x\), then elements \(x\) and \(y\) are said to be **comparable**; otherwise they are **incomparable**.

Definition 2.5 Let \((X, \leq)\) be an ordered set (Definition 2.4 page 4).

The **dual order relation** \(\geq\) of \(\leq\) is \(\geq \triangleq \{(x, y) \in X \times X \mid y \leq x\}\).

The **quasi-order relation** \(<\) of \(\leq\) is \(< \triangleq \{(x, y) \in X \times X \mid x \neq y\}\).

The **dual quasi-order relation** \(>\) of \(<\) is \(> \triangleq \{(x, y) \in X \times X \mid y < x\}\).
Definition 2.6  An order relation \( \leq \) is a linear order relation on \( X \) if \( x \leq y \) or \( y \leq x \) \( \forall x,y \in X \) (comparable). A linearly ordered set is the pair \( (X, \leq) \).

Definition 2.7  In an ordered set \( (X, \leq) \),
the set \( [x, y) \) is a closed interval and
the set \( (x, y) \) is a half-open interval and
the set \( (x, y] \) is an open interval.

Example 2.8  (Coordinatewise order relation) Let \( (X, \leq) \) be an ordered set. Let \( x \triangleq \{x_1, x_2, \ldots, x_N\} \) and \( y \triangleq \{y_1, y_2, \ldots, y_N\} \) be \( N \)-tuples. The coordinatewise order relation \( \preceq \) is defined as
\[
\forall x,y \in X^N \quad \begin{cases} 
 x \preceq y \iff \{x_1 \leq y_1 \text{ and } x_2 \leq y_2 \text{ and } \ldots \text{ and } x_N \leq y_N\} 
\end{cases}
\]
This relation is also called the dictionary order relation or alphabetic order relation.

2.2.2 Bounds on ordered sets

Definition 2.10  Let \( (X, \leq) \) be an ordered set and \( 2^X \) the power set of \( X \). For any set \( A \subseteq 2^X \), \( c \) is an upper bound of \( A \) in \( (X, \leq) \) if \( x \leq c \ \forall x \in A \). An element \( b \) is the least upper bound of \( A \) in \( (X, \leq) \) if \( b \) and \( c \) are upper bounds of \( A \implies b \leq c \). The least upper bound of the set \( A \) is denoted \( \bigvee A \). The join \( x \lor y \) of \( x \) and \( y \) is defined as \( x \lor y \triangleq \bigvee \{x,y\} \).

Definition 2.11  Let \( (X, \preceq) \) be an ordered set and \( 2^X \) the power set of \( X \). For any set \( A \subseteq 2^X \), \( p \) is a lower bound of \( A \) in \( (X, \preceq) \) if \( p \preceq x \ \forall x \in A \). An element \( a \) is the greatest lower bound of \( A \) in \( (X, \preceq) \) if \( a \) and \( p \) are lower bounds of \( A \implies p \preceq a \). The greatest lower bound of the set \( A \) is denoted \( \bigwedge A \). The meet \( x \land y \) of \( x \) and \( y \) is defined as \( x \land y \triangleq \bigwedge \{x,y\} \).

Lemma 2.12  Let \( (X, \preceq) \) be an ordered set. Let \( \bigvee A \) be the least upper bound (Definition 2.10 page 5) of a set \( A \subseteq 2^X \) (Definition 1.2 page 3). Let \( \bigwedge A \) be the greatest lower bound (Definition 2.11 page 5) of a set \( A \subseteq 2^X \).
\[
\{A = X\} \implies \begin{cases} 
1. \bigvee A = \{a \in X \mid x \leq a, \forall x, a \in X\} \\
2. \bigwedge A = \{a \in X \mid a \leq x, \forall x, a \in X\} 
\end{cases}
\]
2.3 Lattices

The structure available in an ordered set (Definition 2.4 page 4) tends to be insufficient to ensure “well-behaved” mathematical systems. This situation is greatly remedied if every pair of elements in the ordered set has both a least upper bound and a greatest lower bound (Definition 2.11 page 5) in the set; in this case, that ordered set is a lattice (next definition).

**Definition 2.13** An algebraic structure $L \triangleq (X, \lor, \land; \leq)$ is a lattice if
1. $(X, \leq)$ is an ordered set (Definition 2.4 page 4) and
2. $x, y \in X \implies x \lor y \in X$ (Definition 2.10 page 5) and
3. $x, y \in X \implies x \land y \in X$ (Definition 2.11 page 5).

The lattice $L$ is linear if $(X, \leq)$ is a linearly ordered set (Definition 2.6 page 5).

**Theorem 2.14** $(X, \lor, \land; \leq)$ is a lattice (Definition 2.13 page 6)

\[
\begin{align*}
    x \lor x &= x & x \land x &= x & \forall x \in X \quad \text{(Idempotent) and} \\
    x \lor y &= y \lor x & x \land y &= y \land x & \forall x, y \in X \quad \text{(Commutative) and} \\
    (x \lor y) \lor z &= x \lor (y \lor z) & (x \land y) \land z &= x \land (y \land z) & \forall x, y, z \in X \quad \text{(Associative) and} \\
    x \lor (x \land y) &= x & x \land (x \lor y) &= x & \forall x, y \in X \quad \text{(Absorptive).}
\end{align*}
\]

**Theorem 2.15** (minimax inequality) Let $(X, \lor, \land; \leq)$ be a lattice (Definition 2.13 page 6).

\[
\bigvee_{i=1}^{m} \bigwedge_{j=1}^{n} x_{ij} \leq \bigwedge_{j=1}^{n} \bigvee_{i=1}^{m} x_{ij} \quad \forall x_{ij} \in X
\]

Special cases of the minimax inequality include three distributive inequalities (next theorem). If for some lattice any one of these inequalities is an equality, then all three are equalities; and in this case, the lattice is called a distributive lattice.

**Theorem 2.16** (distributive inequalities) $(X, \lor, \land; \leq)$ is a lattice (Definition 2.13 page 6)

\[
\begin{align*}
    x \land (y \lor z) &\geq (x \land y) \lor (x \land z) & \forall x, y, z \in X \quad \text{(Join Super-Distributive) and} \\
    x \lor (y \land z) &\leq (x \lor y) \land (x \lor z) & \forall x, y, z \in X \quad \text{(Meet Sub-Distributive) and} \\
    (x \lor y) \lor (x \land z) \lor (y \lor z) &\leq (x \lor y) \land (x \lor z) \land (y \lor z) & \forall x, y, z \in X \quad \text{(Median Inequality).}
\end{align*}
\]

Besides the distributive property, another consequence of the minimax inequality is the modularity inequality (next theorem). A lattice in which this inequality becomes equality is said to be modular.

**Theorem 2.17** (Modular inequality) Let $(X, \lor, \land; \leq)$ be a lattice (Definition 2.13 page 6).

\[
x \leq y \implies x \lor (y \land z) \leq y \land (x \lor z)
\]
3 Background: power means

One of the most well known inequalities in mathematics is Minkowski’s Inequality. In 1946, H.P. Mul-
holland submitted a result that generalizes Minkowski’s Inequality to an equal weighted \( \phi \)-mean.\(^{20}\)
And Milovanović and Milovanovć (1979) generalized this even further to a weighted \( \phi \)-mean (next).

**Theorem 3.1** \(^{21}\) Let \( N \) be a non-negative tuple of weighting values such that \( \sum_{n=1}^{N} \lambda_n = 1 \). A
\[ \phi^{-1}\left(\sum_{n=1}^{N} \lambda_n \phi(x_n + y_n)\right) \leq \phi^{-1}\left(\sum_{n=1}^{N} \lambda_n \phi(x_n)\right) + \phi^{-1}\left(\sum_{n=1}^{N} \lambda_n \phi(y_n)\right) \]

**Definition 3.2** \(^{22}\) Let \( M_{\phi(x,r)}(\langle x_n \rangle) \) be the \( \langle x_n \rangle \) weighted \( \phi \)-mean of a non-negative tuple \( \langle x_n \rangle \). A
mean \( M_{\phi(x,r)}(\langle x_n \rangle) \) is a power mean with parameter \( r \) if \( \phi(x) \equiv x^r \). That is,
\[ M_{\phi(x,r)}(\langle x_n \rangle) = \left(\sum_{n=1}^{N} \lambda_n (x_n)^r\right)^{\frac{1}{r}} \]

**Theorem 3.3** \(^{23}\) Let \( M_{\phi(x,r)}(\langle x_n \rangle) \) be the power mean with parameter \( r \) of an \( N \)-tuple \( \langle x_n \rangle \).
\[ M_{\phi(x,r)}(\langle x_n \rangle) = \begin{cases} \max & \text{for } r = +\infty \\ \prod_{n=1}^{N} x_n^\lambda_n & \text{for } r = 0 \\ \min & \text{for } r = -\infty \end{cases} \]

**Corollary 3.4** \(^{24}\) Let \( \langle x_n \rangle \) be a tuple. Let \( \langle \lambda_n \rangle \) be a tuple of weighting values such that \( \sum_{n=1}^{N} \lambda_n = 1 \).
\[ \min \langle x_n \rangle \leq \left(\sum_{n=1}^{N} \lambda_n \frac{1}{x_n}\right)^{-1} \leq \prod_{n=1}^{N} x_n^{\lambda_n} \leq \sum_{n=1}^{N} \lambda_n x_n \leq \max \langle x_n \rangle \]

\(^{20}\) \cite{81}, page 115, \cite{82}, \cite{55}, (Theorem 24), \cite{108}, page 7, \cite{76}, page 258, \cite{20}, page 44, \cite{16}, page 179

\(^{21}\) \cite{80}, \cite{16}, page 306, (Theorem 9)

\(^{22}\) \cite{16}, page 175, \cite{13}, page 6

\(^{23}\) \cite{16}, pages 175–177, (see also page 203), \cite{13}, pages 6–8, \cite{7}, \cite{8}, page 68, \cite{50}

\(^{24}\) \cite{16}, page 71, \cite{13}, page 5, \cite{21}, pages 457–459, (Note II, theorem 17), \cite{60}, page 183, \cite{50}
4 Background: Metric space

4.1 Definitions

Definition 4.1 A function \( d \in \mathbb{R}^{X \times X} \) (Definition 1.3 page 3) is a quasi-metric on a set \( X \) if

1. \( d(x, y) \geq 0 \quad \forall x, y \in X \) (non-negative) and
2. \( d(x, y) = 0 \iff x = y \quad \forall x, y \in X \) (non-degenerate) and
3. \( d(x, y) \leq d(x, z) + d(z, y) \quad \forall x, y, z \in X \) (subadditive/triangle inequality).

The pair \((X, d)\) is a quasi-metric space if \( d \) is a quasi-metric on \( X \).

Definition 4.2 A quasi-metric (Definition 4.1 page 8) \( d \in \mathbb{R}^{X \times X} \) is a metric on a set \( X \) if

4. \( d(x, y) = d(y, x) \quad \forall x, y \in X \) (symmetric)

A quasi-metric space \((X, d)\) is a metric space if \( d \) is a metric.

Theorem 4.3 (metric characterization) Let \( d \) be a function in \((\mathbb{R}^{+})^{X \times X}\).

\[
\begin{align*}
\{ \text{d(x, y) is a metric} \} & \iff \left\{ \begin{array}{l}
1. \quad d(x, y) = 0 \iff x = y \quad \forall x, y \in X \\
2. \quad d(x, y) \leq d(z, x) + d(z, y) \quad \forall x, y, z \in X
\end{array} \right.
\end{align*}
\]

4.2 Metric transforms

Metrics induce other metrics, as demonstrated by the following:

\(\text{Theorem 4.5} \quad \text{(page 8)}: \quad \text{generate a metric using an isometry.}\)
\(\text{Theorem 4.6} \quad \text{(page 8)}: \quad \text{generate a metric using a monotone function.}\)
\(\text{Theorem 4.9} \quad \text{(page 9)}: \quad \text{generate a metric using a metric preserving function.}\)
\(\text{Theorem 4.19} \quad \text{(page 11)}: \quad \text{generate an \(N\)-dimensional metric from weighted 1-dimensional metrics.}\)

4.2.1 Metric transforms on the domains of metrics

Definition 4.4 Let \((X, d)\) and \((Y, p)\) be metric spaces (Definition 4.2 page 8).

The function \( f \in Y^X \) is an isometry on \((Y, p)(X, d)\) if \( d(x, y) = p(f(x), f(y)) \quad \forall x, y \in X \).

The spaces \((X, d)\) and \((Y, p)\) are isometric if there exists an isometry on \((Y, p)(X, d)\).

Theorem 4.5 Let \((X, d)\) and \((Y, p)\) be metric spaces. Let \( f \) be a bijective function (Definition 1.8 page 3) in \( Y^X \) and \( f^{-1} \) its inverse in \( X^Y \) (Definition 1.3 page 3).

\[
\{ f \text{ is an isometry on } (Y, p)(X, d) \} \iff \{ f^{-1} \text{ is an isometry on } (X, d)(Y, p) \}
\]

Theorem 4.6 (Pullback metric \( g \)-transform metric) Let \( X \) and \( Y \) be sets and \( g \) a function in \( Y^X \).

\[
\{ \begin{array}{l}
1. \quad p \text{ is a metric in } \mathbb{R}^Y \\
2. \quad g \text{ is injective}
\end{array} \quad \text{and} \quad \left\{ \begin{array}{l}
d(x, y) = p(g(x), g(y)) \quad \forall x, y \in X \\
d(x, y) \leq d(z, x) + d(z, y) \quad \forall x, y, z \in X
\end{array} \right. \}
\]

\(\text{References:} [114], \text{page 675}, \langle 1. \rangle, [95], [66], \text{page 71}, \langle \text{Introduction} \rangle, [89], [104], \text{page 65}, [48], \text{page 3}, \langle \text{Introduction} \rangle, [37], \langle \text{triangle inequality—Book I Proposition 20} \rangle\)

\(\text{References:} [34], \text{page 28}, [24], \text{page 21}, [57], \text{page 109}, [43], [42], \text{page 30}\)

\(\text{References:} [78], \text{page 264}, [47], \text{page 18}, [50]\)

\(\text{References:} [106], \text{page 153}, \langle \text{definition 19.4} \rangle, [47], \text{page 124}, \langle \text{Definition 6.22} \rangle, [67], \text{page 15}, \langle \text{Definition 2.4} \rangle, [71], \text{page 110}\)

\(\text{References:} [106], \text{page 153}, \langle \text{theorem 19.5} \rangle\)

\(\text{References:} [33], \text{page 81}\)
4.2.2 Metric preserving functions

**Definition 4.7** 31 Let \( \mathcal{M} \) be the set of all metric spaces (Definition 4.2 page 8) on a set \( X \). \( \phi \in \mathbb{R}^+ \) is a **metric preserving function** if \( (X, \phi) \) is a metric on \( X \) for all \( (X, \phi) \in \mathcal{M} \)

Theorem 4.8 (next theorem) presents some necessary conditions for a function \( \phi \) to be metric preserving. Theorem 4.9 (page 9) presents some sufficient conditions. But first some conditions that are not necessary:

1. It is *not* necessary for \( \phi \) to be continuous (see Example 4.14 page 10).
2. It is *not* necessary for \( \phi \) to be increasing (see Example 4.16 page 10).
3. It is *not* necessary for \( \phi \) to be monotone (see Example 4.17 page 11).

**Theorem 4.8** (necessary conditions) 32 Let \( \mathcal{R} \) be the **range** (Definition 1.4 page 3) of a function \( \phi \).

\[
\phi \text{ is a metric preserving function (Definition 4.7 page 9)} \implies \begin{cases} 
\text{1. } & \phi^{-1}(0) = \{0\} \\
\text{2. } & \mathcal{R} \subseteq \mathbb{R}^+ \\
\text{3. } & \phi(x + y) \leq \phi(x) + \phi(y) \quad (\phi \text{ is subadditive})
\end{cases}
\]

**Theorem 4.9** (sufficient conditions) 33 Let \( \phi \) be a function in \( \mathbb{R}^+ \).

\[
\begin{cases} 
\text{1. } & x \geq y \implies \phi(x) \geq \phi(y) \quad (\text{isotone}) \\
\text{2. } & \phi(0) = 0 \\
\text{3. } & \phi(x + y) \leq \phi(x) + \phi(y) \quad (\text{subadditive})
\end{cases} \implies \phi \text{ is a metric preserving function (Definition 4.7 page 9)}.
\]

![Figure 3: metric preserving functions](image)

**Example 4.10** \( (\alpha \text{-scaled metric/dilated metric}) \) 34 Let \( (X, d) \) be a metric space (Definition 4.2 page 8). \( \phi(x) \triangleq \alpha x, \alpha \in \mathbb{R}^+ \) is a metric preserving function (Figure 3 page 9 (A))

\[\text{Proof:} \quad \text{The proof for Example 4.10–Example 4.15 (page 10) follow from Theorem 4.9 (page 9).} \]

31 [110], page 849, (Definition 1.1), [25], page 309, [33], page 80
32 [25], page 310, (Proposition 2.1), [33], page 80
33 [25], (Proposition 2.3), [33], page 80, [65], page 131, (Problem C)
34 [32], page 44
Example 4.11  (power transform metric/snowflake transform metric)  \(^{35}\) Let \((X, d)\) be a metric space (Definition 4.2 page 8). \(\phi(x) \triangleq x^a, \ a \in (0, 1]\) is a metric preserving function (see Figure 3 page 9 (B)).

Example 4.12  \((\alpha\text{-truncated metric/radar screen metric})\)  \(^{36}\) Let \((X, d)\) be a metric space (Definition 4.2 page 8). \(\phi(x) \triangleq \min \{a, x\}, \ a \in \mathbb{R}^+\) is a metric preserving function (see Figure 3 page 9 (C)).

Example 4.13  (bounded metric)  \(^{37}\) Let \((X, d)\) be a metric space (Definition 4.2 page 8). \(\phi(x) \triangleq \frac{x}{1 + x}\) is a metric preserving function (see Figure 3 page 9 (D)).

Example 4.14  (discrete metric preserving function)  \(^{38}\) Let \(\phi\) be a function in \(\mathbb{R}^\mathbb{R}\). \(\phi(x) \triangleq \begin{cases} 0 & \text{for } x \leq 0, \\ 1 & \text{otherwise} \end{cases}\) is a metric preserving function (see Figure 3 page 9 (E)).

Example 4.15  Let \(\phi\) be a function in \(\mathbb{R}^\mathbb{R}\). \(\phi(x) \triangleq \begin{cases} x & \text{for } 0 \leq x < 1, \\ 1 & \text{for } 1 \leq x \leq 2, \\ x - 1 & \text{for } 2 < x < 3, \\ 2 & \text{for } x \geq 3 \end{cases}\) is a metric preserving function (see Figure 3 page 9 (F)).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{example4}
\caption{non-monotone metric preserving functions}
\end{figure}

Example 4.16  Let \(\phi\) be a function in \(\mathbb{R}^\mathbb{R}\). \(\phi(x) \triangleq \begin{cases} 0 & \text{for } x = 0, \\ 1 + \frac{1}{x+1} & \text{for } x > 0 \end{cases}\) is a metric preserving function (see Figure 4 page 10 (A)).

\begin{itemize}
\item \(\text{Proof:}\)
\item (1) Note that \(\phi \ast d(x, x) = 0 \iff x = 0.\)
\item (2) lemma: \(\frac{1}{a + b} \leq \frac{1}{a} \leq \frac{1}{a} + \frac{1}{b}\) \(\forall a, b \in \mathbb{R}^+\)
\end{itemize}

\(^{35}\) [33], page 81, [32], page 45
\(^{36}\) [47], page 33, [32], pages 242–243
\(^{37}\) [110], page 849, [3], page 39
\(^{38}\) [25], page 311
(3) Proof that $\phi \circ \delta$ is subadditive:

$$
\phi \circ \delta (x, y) = 1 + \frac{1}{1 + \delta (x, z) + \delta (z, y)}
$$

by definition of $\phi$

$$
\leq 1 + \frac{1}{1 + \delta (x, z) + 1 + \delta (z, y)}
$$

by subadditive property of $\delta$ (Definition 4.2 page 8)

$$
\leq 1 + \frac{1}{1 + \delta (x, z) + 1 + \frac{1}{1 + \delta (z, y)}}
$$

because $\frac{1}{x} \leq \frac{1}{x + 1} \forall x \in \mathbb{R}$

$$
\leq 1 + \frac{1}{1 + \delta (x, z)} + \frac{1}{1 + \frac{1}{1 + \delta (z, y)}}
$$

because $x \leq x + 1 \forall x \in \mathbb{R}$

$$
= \phi \circ \delta (x, z) + \phi \circ \delta (z, y)
$$

by definition of $\phi$

(4) Therefore, by Theorem 4.3 (page 8), $\phi \circ \delta (x, y)$ is a metric and $\phi$ is a metric preserving function.

Example 4.17 ³³ Let $\phi$ be a function in $\mathbb{R}^\mathbb{R}$.

$$
\phi(x) \triangleq \begin{cases}
    x & \text{for } x \leq 2 \\
    1 + \frac{1}{x - 1} & \text{for } x > 2
\end{cases}
$$

is a metric preserving function (see Figure 4 page 10 (B)).

Example 4.18 Let $\phi$ be a function in $\mathbb{R}^\mathbb{R}$.

$$
\phi(x) \triangleq \begin{cases}
    x & \text{for } 0 \leq x \leq 2 \\
    -x + 4 & \text{for } 2 < x < 3 \\
    1 & \text{for } x \geq 3
\end{cases}
$$

is a metric preserving function (see Figure 4 page 10 (C)).

4.2.3 Product metrics

Theorem 4.19 (Power mean metrics) Let $X$ be a set. Let $\langle x_n \rangle \in X_1^N$ and $\langle y_n \rangle \in X_1^N$ be $N$-tuples on $X$.

\[
\begin{align*}
1. \quad p & \text{ is a metric on } X \quad \text{and} \\
2. \quad \sum_{n=1}^{N} \lambda_n = 1
\end{align*}
\]

\[
\Rightarrow \quad d (\langle x_n \rangle, \langle y_n \rangle) \triangleq \left( \sum_{n=1}^{N} \lambda_n p^r (x_n, y_n) \right)^{\frac{1}{r}}, \quad r \in [1, \infty]
\]

is a metric on $X$.

Moreover, if $r = \infty$, then

$$
d (\langle x_n \rangle, \langle y_n \rangle) = \max_{n=1, \ldots, N} \delta (x_n, y_n).
$$

Proof:

³³ [25], page 309, [35], page 25, (Example 1), [63]
4.2 METRIC TRANSFORMS

(1) Proof that \( \|x_n\| = \|y_n\| \implies d(\|x_n\|, \|y_n\|) = 0 \) for \( r \in [1, \infty) \):

\[
d(\|x_n\|, \|y_n\|) \triangleq \left( \sum_{n=1}^{N} \lambda_n p'(x_n, y_n) \right)^\frac{1}{r}
\]

by definition of \( d \)

\[
eq \left( \sum_{n=1}^{N} \lambda_n p'(x_n, x_n) \right)^\frac{1}{r}
\]

by \( \|x_n\| = \|y_n\| \) hypothesis

\[
eq \left( \sum_{n=1}^{N} 0 \right)^\frac{1}{r} = 0
\]

because \( p \) is nondegenerate

(2) Proof that \( \|x_n\| = \|y_n\| \iff d(\|x_n\|, \|y_n\|) = 0 \) for \( r \in [1, \infty) \):

\[
0 = d(\|x_n\|, \|y_n\|)
\]

by \( d(\|x_n\|, \|y_n\|) = 0 \) hypothesis

\[
\triangleq \left( \sum_{n=1}^{N} \lambda_n p'(x_n, y_n) \right)^\frac{1}{r}
\]

by definition of \( d \)

\[
\implies \left( p(x_n, y_n) \right)^\frac{1}{r} = 0 \text{ for } n = 1, 2, \ldots, N
\]

because \( p \) is non-negative

\[
\implies \|x_n\| = \|y_n\|
\]

because \( p \) is nondegenerate

(3) Proof that \( d \) satisfies the triangle inequality property for \( r = 1 \):

\[
d(\|x_n\|, \|y_n\|) \triangleq \left( \sum_{n=1}^{N} \lambda_n p'(x_n, y_n) \right)^\frac{1}{r}
\]

by definition of \( d \)

\[
= \sum_{n=1}^{N} \lambda_n p(x_n, y_n)
\]

by \( r = 1 \) hypothesis

\[
\leq \sum_{n=1}^{N} \lambda_n[ p(z_n, x_n) + p(z_n, y_n) ]
\]

by triangle inequality

\[
= \sum_{n=1}^{N} \lambda_n p(z_n, x_n) + \sum_{n=1}^{N} \lambda_n p(z_n, y_n)
\]

\[
= \left( \sum_{n=1}^{N} \lambda_n p'(z_n, x_n) \right)^\frac{1}{r} + \left( \sum_{n=1}^{N} \lambda_n p'(z_n, y_n) \right)^\frac{1}{r}
\]

by \( r = 1 \) hypothesis

\[
\triangleq d(\|z_n\|, \|x_n\|) + d(\|z_n\|, \|y_n\|)
\]

by definition of \( d \)

(4) Proof that \( d \) satisfies the triangle inequality property for \( r \in (1, \infty) \):

\[
d(\|x_n\|, \|y_n\|)
\]

\[
\triangleq \left( \sum_{n=1}^{N} \lambda_n p'(x_n, y_n) \right)^\frac{1}{r}
\]

by definition of \( d \)

\[
\leq \left( \sum_{n=1}^{N} \lambda_n [ p(z_n, x_n) + p(z_n, y_n) ]^r \right)^\frac{1}{r}
\]

by subadditive property (Definition 4.2 page 8)
4 BACKGROUND: METRIC SPACE

4.3 Examples

Example 4.20 (usual metric) Let \(|\cdot| \in \mathbb{R}^{\geq 0}\) be an absolute value function on a ring \(R\). The function \(d(x, y) \triangleq |x - y|\) is a metric on \(\mathbb{R}\), called the usual metric.

\[d(x, y) \leq d(x, z) + d(z, y)\] by subadditive property (Definition 4.2 page 8)

\[d(x, y) \leq 0\] by Minkowski's inequality

\[d(x, y) = 0\] because \(p\) is nondegenerate

4.3 Examples

Example 4.20 (usual metric) Let \(|\cdot| \in \mathbb{R}^{\geq 0}\) be an absolute value function on a ring \(R\). The function \(d(x, y) \triangleq |x - y|\) is a metric on \(\mathbb{R}\), called the usual metric.

\[d(x, y) \leq d(x, z) + d(z, y)\] by subadditive property (Definition 4.2 page 8)

\[d(x, y) \leq 0\] by Minkowski's inequality

\[d(x, y) = 0\] because \(p\) is nondegenerate

4.3 Examples

Example 4.20 (usual metric) Let \(|\cdot| \in \mathbb{R}^{\geq 0}\) be an absolute value function on a ring \(R\). The function \(d(x, y) \triangleq |x - y|\) is a metric on \(\mathbb{R}\), called the usual metric.

\[d(x, y) \leq d(x, z) + d(z, y)\] by subadditive property (Definition 4.2 page 8)

\[d(x, y) \leq 0\] by Minkowski's inequality

\[d(x, y) = 0\] because \(p\) is nondegenerate
Example 4.21 (discrete metric) Let $X$ be a set and $d \in \mathbb{R}^{X \times X}$.

$$d(x, y) \triangleq \begin{cases} 0 & \text{for } x = y \\ 1 & \text{otherwise} \end{cases}$$

is a metric on $X$, called the discrete metric.

5 Background: random processes

Definition 5.1 Let $X$ be a set. A function $P \in \mathbb{R}^{+X}$ is a probability function if

1. $P(1) = 1$ (normalized) and
2. $P(x) \geq 0 \forall x \in X$ (nonnegative) and
3. $x \land y = 0 \implies P(x \lor y) = P(x) + P(y) \forall x, y \in X$ (additive).

Remark 5.2 The previous definition of a probability function $P$ (Definition 5.1 page 14) only requires additivity rather than $\sigma$-additivity. This differs from the measure-theoretic probability function, due to A. N. Kolmogorov, which is defined as

1. $P(1) = 1$ (normalized) and
2. $P(x) \geq 0 \forall x \in X$ (nonnegative) and
3. $\bigwedge_{n=1}^{\infty} x_n = 0 \implies P\left(\bigvee_{n=1}^{\infty} x_n\right) = \sum_{n=1}^{\infty} P(x_n) \forall x_n \in X$ ($\sigma$-additive).

The advantage of this definition is that $P$ is a measure, and hence all the power of measure theory is subsequently at one’s disposal in using $P$. However, it has often been argued that the requirement of $\sigma$-additivity is unnecessary for a probability function. Even as early as 1930, de Finetti argued against it, in what became a kind of polite running debate with Fréchet. In fact, Kolmogorov himself provided some argument against $\sigma$-additivity when referring to the closely related Axiom of Continuity saying, “Since the new axiom is essential for infinite fields of probability only, it is almost impossible to elu-cidate its empirical meaning...For, in describing any observable random process we can obtain only finite fields of probability...” But in its support he added, “This limitation has been found expedient in researches of the most diverse sort.”

Definition 5.3 The triple $(\Omega, \mathcal{E}, P)$ is a probability space if $\Omega$ is a set, $\mathcal{E}$ is a $\mathcal{B}$-algebra on $\Omega$, and $P$ is a probability function (Definition 5.1 page 14) in $[0, 1]^\mathcal{E}$. In this case, $\Omega$ is called the set of outcomes.

Before defining what a random variable is (Definition 5.4 page 14), here are two things that it is not:

1. A random variable is not random.
2. A random variable is not a variable.

Definition 5.4 Let $(\Omega, \mathcal{E}, P)$ be a probability space (Definition 5.3 page 14).

A traditional random variable $X$ on $(\Omega, \mathcal{E}, P)$ is any function in the set $\mathbb{R}^{\Omega}$ (Definition 1.3 page 3).
Definition 5.5 Let \((\Omega, E, P)\) be a probability space (Definition 5.3 page 14) and \(X \in \mathbb{R}^\Omega\) a random variable (Definition 5.4 page 14). A probability density function \(p\) is a function in \([0, 1]^E\) such that
\[
P(A) = \int_A p(x) \, dx \quad \forall A \in E
\]

Definition 5.6 Let \((\Omega, E, P)\) be a probability space (Definition 5.3 page 14) and \(X \in \mathbb{R}^\Omega\) a random variable (Definition 5.4 page 14).

The traditional expected value \(E(X)\) of \(X\) is
\[
E(X) \triangleq \int_{\mathbb{R}} x \, p(x) \, dx.
\]
The traditional variance \(\text{Var}(X)\) of \(X\) is
\[
\text{Var}(X) \triangleq \int_{\mathbb{R}} [x - E(X)]^2 p(x) \, dx.
\]

Proposition 5.7 Let \((\Omega, E, P)\) be a probability space (Definition 5.3 page 14), \(X \in \mathbb{R}^\Omega\) a traditional random variable (Definition 5.4 page 14), \(E(X)\) the traditional expected value of \(X\), and \(\text{Var}(X)\) the traditional variance of \(X\) (Definition 5.6 page 15).

\[
\begin{align*}
\{ P(x) = 0 \forall x \notin \mathbb{Z} \} & \implies \left\{ \begin{array}{l}
1. E(X) = \sum_{x \in \mathbb{Z}} x P(x) \\
2. \text{Var}(X) = \sum_{x \in \mathbb{Z}} [x - E(X)]^2 P(x)
\end{array} \right\
\end{align*}
\]

Proposition 5.8 Let \((\Omega, E, P), X,\) and \(E\) be defined as in Proposition 5.7 (page 15).

\[
P(y-x) = P(y+x) \quad \forall x \in \mathbb{R} \implies \{ E(X) = y \}
\]

\((P\text{ is symmetric about a point } y)\)

Proof:
\[
E(X) \triangleq \int_{-\infty}^{\infty} x p(x) \, dx \quad \text{by definition of } E \quad (\text{Definition 5.6 page 15})
\]
\[
= \int_{-\infty}^{x+y} x p(x) \, dx + \int_{x+y}^{\infty} x p(x) \, dx \quad \text{by additive property of Lebesgue integration op.}^{47}
\]
\[
= \int_{u+y=-\infty}^{u+y=\infty} (u + x) p(u + x) \, du + \int_{u+y=\infty}^{u+y=-\infty} (u + x) p(u + x) \, du \quad \text{where } u \triangleq x - y \implies x = u + y
\]
\[
= -\int_{-\infty}^{\infty} (u + x) p(u + x) \, du + \int_{0}^{\infty} (u + x) p(u + x) \, du
\]
\[
= \int_{0}^{\infty} (-v + x) p(-v + x) \, dv + \int_{0}^{\infty} (u + x) p(u + x) \, du \quad \text{where } v \triangleq -u
\]
\[
= -\int_{0}^{\infty} v p(-v + x) \, dv + \int_{0}^{\infty} u p(u + x) \, du + \gamma \int_{0}^{\infty} p(-v + x) \, dv + \gamma \int_{0}^{\infty} p(u + x) \, du
\]
\[
= -\gamma \int_{0}^{\infty} v p(-v + x) \, dv + \int_{0}^{\infty} u p(u + x) \, du + \gamma \left( \int_{0}^{\infty} p(-v + x) \, dv + \gamma \int_{0}^{\infty} p(u + x) \, du \right)
\]
\[
\text{by symmetry hypothesis; cancels to 0}
\]
\[
\gamma \left( \int_{0}^{\infty} p(u + x) \, du + \gamma \int_{0}^{\infty} \int_{0}^{\infty} p(u + x) \, du \right)
\]
\[
\gamma \int_{0}^{\infty} p(u + x) \, du = \gamma \int_{-\infty}^{\infty} p(u) \, du = \gamma
\]

\(^{47}[17]\), page 35
6 Ordered metric spaces

6.1 Definitions

**Definition 6.1** A triple $\mathcal{G} = (X, d, \leq)$ is an ordered quasi-metric space if $(X, d)$ is a quasi-metric space (Definition 4.1 page 8) and $(X, \leq)$ is an ordered set (Definition 2.4 page 4).

$\mathcal{G}$ is an ordered metric space if $d$ is a metric (Definition 4.2 page 8).

$\mathcal{G}$ is an unordered quasi-metric space if $\leq = \emptyset$.

$\mathcal{G}$ is an unordered metric space if $d$ is a metric and $\leq = \emptyset$.

**Remark 6.2** Note that the four structures defined in Definition 6.1 are not mutually exclusive. For example, by Definition 6.1,

\[
\{\text{unordered metric space}\} \subset \{\text{unordered quasi-metric space}\} \subset \{\text{ordered quasi-metric space}\}
\]

\[
\{\text{unordered metric space}\} \subset \{\text{ordered metric space}\} \subset \{\text{ordered quasi-metric space}\}.
\]

**Remark 6.3** The use of the quasi-metric rather than exclusive use of the more restrictive metric in Definition 6.1 is motivated by state machines, where metrics measuring distances between states are in some cases by nature non-symmetric. One such example is the linear congruential pseudo-random number generator (Example 9.19 page 59).

**Remark 6.4** This paper makes extensive reference to the real line (next definition). There are several ways to define the real line. In particular, there are many possible ordering relations on $\mathbb{R}$ and several possible topologies on $\mathbb{R}$.48 In fact, order and topology are closely related in that an order relation $\leq$ (Definition 2.4 page 4) on a set always induces a topology (called the order topology); and in the case of the real line, a topology induces an order structure up to the order relation’s dual (Definition 2.5 page 4).49 This paper uses a fairly standard structure, as defined next.

---

48 [2], page 31, (“six topologies on the real line”), [101], pages 64–70, (Weird topologies on the real line), [85], page 53, (“often used topologies on the real line”), [61], pages 85–91, (§4.2 Examples of Topological Spaces).

49 [58], page 52, (2–5 The interval and the circle), [101], pages 69–70, (5.75 Note: Ordering and topology on $\mathbb{R}$)

---
7  MONOTONE FUNCTIONS ON ORDERED SETS

Definition 6.5  The triple \((\mathbb{R}, |\cdot|, \leq)\) is the real line if \(\mathbb{R}\) is the set of real numbers (Definition 1.1 page 2), \(d(x, y) \triangleq |x - y|\) is the usual metric on \(\mathbb{R}\) (Example 4.20 page 13), and \(\leq\) is the standard linear order relation (Definition 2.6 page 5) on \(\mathbb{R}\).

Definition 6.6  The triple \((\mathbb{Z}, |\cdot|, \leq)\) is the integer line if \(\mathbb{Z}\) is the set of integers (Definition 1.1 page 2), \(d(m, n) \triangleq |m - n|\) is the usual metric (Example 4.20 page 13) on \(\mathbb{R}\) restricted to \(\mathbb{Z}\), and \(\leq\) is the standard linear order relation on \(\mathbb{Z}\) as induced by Peano’s Axioms.\(^{50}\)

6.2  Examples

Example 6.7  The integer line (Definition 6.6 page 17) is an ordered metric space (Definition 6.1 page 16), and is illustrated in Figure 5 page 16 (A).

Example 6.8  The real line (Definition 6.5 page 17) is an ordered metric space (Definition 6.1 page 16), and is illustrated in Figure 5 page 16 (B).

Example 6.9  The complex plane \((\mathbb{C}, |\cdot|, \leq)\) is an ordered metric space (Definition 6.1 page 16) where \(\mathbb{C} \triangleq \mathbb{R}^2\) is the set of complex numbers, \(d(x, y) \triangleq |x - y| \triangleq \sqrt{\Re x - \Re y)^2 + (\Im x - \Im y)^2}\), \(\Re x \triangleq \Re (a, b) \triangleq a \forall (a, b) \in \mathbb{C}\) (\(\Re x\) is the real part of \(x\)), \(\Im x \triangleq \Im (a, b) \triangleq b \forall (a, b) \in \mathbb{C}\) (\(\Im x\) is the imaginary part of \(x\)), and \(\leq\) is any order relation defined on \(\mathbb{C}\). Possible order relations include the coordinatewise order relation (Example 2.8 page 5), the lexicographical order relation (Example 2.9 page 5), and \(\leq = \emptyset\) (in which case the complex plane is unordered). The complex plane is illustrated in Figure 5 page 16 (C).

Example 6.10  A 6 element ring \((\{0, 1, 2, 3, 4, 5\}, d, \emptyset)\) is an unordered metric space (Definition 6.1 page 16) where the metric \(d\) is defined on a ring as illustrated in Figure 5 page 16 (D), with each line segment representing a distance of 1.

Example 6.11  A 6 element discrete metric \((\{0, 1, 2, 3, 4, 5\}, d, \emptyset)\) is an unordered metric space (Definition 6.1 page 16) where the metric \(d\) is the discrete metric (Example 4.21 page 14). This structure is illustrated in Figure 5 page 16 (E).

Example 6.12  Figure 5 page 16 (F) illustrates a linear congruential pseudo-random number generator induced by the equation \(y_{n+1} = (y_n + 2) \mod 5\) with \(y_0 = 1\). The structure is an unordered quasi-metric space. See Example 9.17 (page 55)–Example 9.19 (page 59) for further demonstration.

7  Monotone functions on ordered sets

7.1  Definitions

Definition 7.1  \(^{51}\) Let \((X, \leq)\) and \((Y, \sqsubseteq)\) be ordered sets (Definition 2.4 page 4, Definition 2.5 page 4). Let \(\phi\) be a function in \(Y^X\) (Definition 1.3 page 3).
\( \phi \) is isotone \hspace{1cm} \text{in } (Y, \sqsubseteq) (X, \leq) \text{ if } x \leq y \implies \phi(x) \sqsubseteq \phi(y) \hspace{1cm} \forall x, y \in X.

\( \phi \) is strictly isotone \hspace{1cm} \text{in } (Y, \sqsubseteq) (X, \leq) \text{ if } x < y \implies \phi(x) \sqsubset \phi(y) \hspace{1cm} \forall x, y \in X.

\( \psi \) is antitone \hspace{1cm} \text{in } (Y, \sqsubseteq) (X, \leq) \text{ if } x \leq y \implies \psi(y) \sqsubseteq \psi(x) \hspace{1cm} \forall x, y \in X.

\( \psi \) is strictly antitone \hspace{1cm} \text{in } (Y, \sqsubseteq) (X, \leq) \text{ if } x < y \implies \psi(y) \sqsubset \psi(x) \hspace{1cm} \forall x, y \in X.

A function is monotone if it is isotone or antitone and strictly monotone if it is strictly isotone or strictly antitone. An isotone function in \((Y, \sqsubseteq) (X, \leq)\) is also said to be order preserving in \((Y, \sqsubseteq) (X, \leq)\).

### 7.2 Properties

**Lemma 7.2** Let \((X, \lor, \land; \leq)\) and \((Y, \sqcup, \sqcap; \sqsubseteq)\) be lattices (Definition 2.13 page 6).

Let \(f\) be a function in \(X^X\). Let \(\phi\) be a function in \(Y^X\) (Definition 1.3 page 3).

\[
\left\{ \begin{array}{l}
\text{\(\phi\) is isotone (Definition 7.1 page 17)} \\
\text{\(\phi\) is antitone (Definition 7.1 page 17)}
\end{array} \right. \quad \Rightarrow \quad \left\{ \begin{array}{l}
1. \ \arg \bigvee_{x \in X} f(x) \subseteq \arg \bigwedge_{x \in X} \phi[f(x)] \quad \text{and} \\
2. \ \arg \bigwedge_{x \in X} f(x) \subseteq \arg \bigvee_{x \in X} \phi[f(x)]
\end{array} \right.
\]

\(\) \hspace{2cm} Proof:

\[
\begin{align*}
\arg \bigvee_{x \in X} f(x) &= \arg \bigvee_{x \in X} \left\{ f(x) \mid f(y) \leq f(x) \hspace{1cm} \forall x, y \in X \right\} \\
&\subseteq \arg \bigvee_{x \in X} \left\{ f(x) \mid \phi[f(y)] \leq \phi(f(x)) \hspace{1cm} \forall x, y \in X \right\} \\
&= \arg \bigvee_{x \in X} \left\{ \phi[f(x)] \mid \phi[f(y)] \leq \phi(f(x)) \hspace{1cm} \forall x, y \in X \right\} \\
&= \arg \bigvee_{x \in X} \phi[f(x)] \quad \text{because } f \in X^X \text{ and by Lemma 2.12 page 5}
\end{align*}
\]

\[
\begin{align*}
\arg \bigwedge_{x \in X} f(x) &= \arg \bigwedge_{x \in X} \left\{ f(x) \mid f(y) \leq f(x) \hspace{1cm} \forall x, y \in X \right\} \\
&\subseteq \arg \bigwedge_{x \in X} \left\{ f(x) \mid \phi[f(y)] \leq \phi(f(x)) \hspace{1cm} \forall x, y \in X \right\} \\
&= \arg \bigwedge_{x \in X} \left\{ \phi[f(x)] \mid \phi[f(y)] \leq \phi(f(x)) \hspace{1cm} \forall x, y \in X \right\} \\
&= \arg \bigwedge_{x \in X} \phi[f(x)] \quad \text{because } f \in X^X \text{ and by Lemma 2.12 page 5}
\end{align*}
\]

**Remark 7.3** Let \((X, \leq)\) and \((Y, \sqsubseteq)\) be ordered sets (Definition 2.4 page 4). Let \(\sqsubseteq\) be the quasi-order relation of \(\sqsubseteq\) (Definition 2.5 page 4). Let \(\phi\) be a function in \(Y^X\) (Definition 1.3 page 3). Note that even if \(\phi\) is bijective (Definition 1.8 page 3) and strictly isotone (Definition 7.1 page 17),

\[
x < y \iff \phi(x) \sqsubset \phi(y) \hspace{1cm} \forall x, y \in X.
\]

An example is illustrated to the right where \(\phi(l) \sqsubset \phi(r)\), but \(l \not\sqsubseteq r\).

**Lemma 7.4** Let \(X \triangleq (X, \leq)\) and \(Y \triangleq (Y, \sqsubseteq)\) be ordered sets. Let \(\phi\) be a function in \(Y^X\).

\[
\left\{ \begin{array}{l}
A. \ \phi \text{ is strictly isotone} \quad \text{and} \\
B. \ X \text{ and } Y \text{ are linearly ordered}
\end{array} \right. \quad \Rightarrow \quad \left\{ \begin{array}{l}
1. \ x \leq y \iff \phi(x) \sqsubseteq \phi(y) \hspace{1cm} \forall x, y \in X \quad \text{and} \\
2. \ x < y \iff \phi(x) \sqsubset \phi(y) \hspace{1cm} \forall x, y \in X
\end{array} \right.
\]

\(\) \hspace{2cm} [19], page 10
**Lemma 7.5** Let $X \triangleq (X, \lor, \land; \leq)$ and $Y \triangleq (Y, \lor, \land; \leq)$ be lattices (Definition 2.13 page 6).

Let $f$ be a function in $X^X$. Let $\phi$ be a function in $Y^X$ (Definition 1.3 page 3).

A. $\phi$ is strictly isotone (Definition 7.1 page 17) and

B. $X$ is linearly ordered (Definition 2.6 page 5) and

C. $Y$ is linearly ordered (Definition 2.6 page 5) \[ \implies \quad \begin{aligned} 1. \quad \bigvee_{x \in X} \phi[f(x)] &= \phi \left[ \bigvee_{x \in X} f(x) \right] \\
&= \{ \phi[f(a)] \mid f(x) \leq f(a) \quad \forall x, a \in X \} \\
&= \{ \phi[f(a)] \mid \phi[f(x)] \subseteq \phi[f(a)] \quad \forall x \in X \} \\
&\triangleq \bigvee_{x \in X} \phi[f(x)] \\
2. \quad \bigwedge_{x \in X} \phi[f(x)] &= \phi \left[ \bigwedge_{x \in X} f(x) \right] \\
&= \{ \phi[f(a)] \mid f(a) \leq f(x) \quad \forall x \in X \} \\
&= \{ \phi[f(a)] \mid \phi[f(a)] \subseteq \phi[f(x)] \quad \forall x \in X \} \\
&\triangleq \bigwedge_{x \in X} \phi[f(x)] \end{aligned} \]

Proof:

\[ \phi \left[ \bigvee_{x \in X} f(x) \right] = \phi[ \{ f(a) \mid f(x) \leq f(a) \quad \forall x, a \in X \} ] \]

because $f \in X^X$ and by Lemma 2.12 page 5

\[ = \{ \phi[f(a)] \mid f(x) \leq f(a) \quad \forall x \in X \} \]

by Lemma 7.4 (page 18)

\[ = \{ \phi[f(a)] \mid \phi[f(x)] \subseteq \phi[f(a)] \quad \forall x \in X \} \]

by definition of $(Y, \lor, \land; \leq)$

\[ \phi \left[ \bigwedge_{x \in X} f(x) \right] = \phi[ \{ f(a) \mid f(a) \leq f(x) \quad \forall x \in X \} ] \]

because $f \in X^X$ and by Lemma 2.12 page 5

\[ = \{ \phi[f(a)] \mid f(a) \leq f(x) \quad \forall x \in X \} \]

by Lemma 7.4 (page 18)

\[ = \{ \phi[f(a)] \mid \phi[f(a)] \subseteq \phi[f(x)] \quad \forall x \in X \} \]

by definition of $(Y, \lor, \land; \leq)$

**Lemma 7.6** Let $X \triangleq (X, \lor, \land; \leq)$ and $Y \triangleq (Y, \lor, \land; \leq)$ be lattices. Let $f$ be a function in $X^X$.

Let $\psi$ be a function in $Y^X$.

A. $\psi$ is strictly antitone (Definition 7.1 page 17) and

B. $X$ is linearly ordered (Definition 2.6 page 5) and

C. $Y$ is linearly ordered (Definition 2.6 page 5) \[ \implies \quad \begin{aligned} 1. \quad \bigwedge_{x \in X} \psi[f(x)] &= \psi \left[ \bigwedge_{x \in X} f(x) \right] \\
&= \{ \psi[f(a)] \mid f(a) \leq f(x) \quad \forall x \in X \} \\
&= \{ \psi[f(a)] \mid \psi[f(a)] \subseteq \psi[f(x)] \quad \forall x \in X \} \\
&\triangleq \bigwedge_{x \in X} \psi[f(x)] \\
2. \quad \bigvee_{x \in X} \psi[f(x)] &= \psi \left[ \bigvee_{x \in X} f(x) \right] \\
&= \{ \psi[f(a)] \mid f(x) \leq f(a) \quad \forall x, a \in X \} \\
&= \{ \psi[f(a)] \mid \phi[f(x)] \subseteq \phi[f(a)] \quad \forall x \in X \} \\
&\triangleq \bigvee_{x \in X} \psi[f(x)] \end{aligned} \]
\[ \psi \left[ \bigvee_{x \in X} f(x) \right] = \psi \left[ \{ f(a) | f(x) \leq f(a) \quad \forall x \in X \} \right] \quad \text{by definition of } (X, \vee, \wedge ; \leq) \]
\[ = \{ \psi[f(a)] | f(x) \leq f(a) \quad \forall x \in X \} \]
\[ = \{ \psi[f(a)] | \psi[f(x)] \leq \psi[f(a)] \quad \forall x \in X \} \quad \text{by definition of } \text{strictly antitone} \text{ (Definition 7.1 page 17)} \]
\[ \triangleq \bigvee_{x \in X} \psi[f(x)] \quad \text{by definition of } (Y, \cup, \cap ; \leq) \]

\[ \psi \left[ \bigwedge_{x \in X} f(x) \right] = \psi \left[ \{ f(a) | f(x) \leq f(x) \quad \forall x \in X \} \right] \quad \text{by definition of } (X, \vee, \wedge ; \leq) \]
\[ = \{ \psi[f(a)] | f(a) \leq f(x) \quad \forall x \in X \} \]
\[ = \{ \psi[f(a)] | \psi[f(x)] \leq \psi[f(a)] \quad \forall x \in X \} \quad \text{by definition of } \text{strictly antitone} \text{ (Definition 7.1 page 17)} \]
\[ \triangleq \bigwedge_{x \in X} \psi[f(x)] \quad \text{by definition of } (Y, \cup, \cap ; \leq) \]

Lemma 7.7 Let \( X \triangleq (X, \vee, \wedge ; \leq) \) and \( Y \triangleq (Y, \cup, \cap ; \leq) \) be lattices (Definition 2.13 page 6). Let \( f \) be a function in \( X^X \). Let \( \phi \) and \( \psi \) be functions in \( Y^X \).

\[ \begin{align*}
\{ & \text{A. } \phi \text{ is strictly isotone } \text{ (Definition 7.1 page 17)} \text{ and } \\
& \text{B. } X \text{ is linearly ordered } \text{ (Definition 2.6 page 5)} \}
\end{align*} \]
\[ \implies \begin{cases} 
1. \ \text{arg} \bigvee_{x \in X} f(x) = \text{arg} \bigwedge_{x \in X} \phi[f(x)] \quad \text{and} \\
2. \ \text{arg} \bigwedge_{x \in X} f(x) = \text{arg} \bigvee_{x \in X} \phi[f(x)] 
\end{cases} \]

\[ \begin{align*}
\text{PROOF:} & \quad \text{arg} \bigvee_{x \in X} f(x) \triangleq \text{arg} \bigvee \{ f(x) | x \in X \} \\
& = \text{arg} \{ f(a) | f(x) \leq f(a) \quad \forall x, a \in X \} \quad \text{because } f \in X^X \text{ and by Lemma 2.12 page 5} \\
& = \text{arg} \{ f(a) | \phi[f(x)] \leq \phi[f(a)] \quad \forall x, a \in X \} \quad \text{by hypothesis (A) and Lemma 7.6 page 19} \\
& = \text{arg} \{ \phi[f(a)] | f(x) \leq f(a) \quad \forall x, a \in X \} \quad \text{because } \text{arg}_x \{ f(x) | P(x) \} = \text{arg}_x \{ g[f(x)] | P(x) \} \\
& \triangleq \bigvee_{x \in X} \phi[f(x)] \quad \text{because } f \in X^X \text{ and by Lemma 2.12 page 5} \\
\end{align*} \]

\[ \begin{align*}
\text{PROOF:} & \quad \text{arg} \bigwedge_{x \in X} f(x) \triangleq \text{arg} \bigwedge \{ f(x) | x \in X \} \\
& = \text{arg} \{ f(a) | f(a) \leq f(x) \quad \forall x, a \in X \} \quad \text{because } f \in X^X \text{ and by Lemma 2.12 page 5} \\
& = \text{arg} \{ f(a) | \phi[f(a)] \leq \phi[f(x)] \quad \forall x, a \in X \} \quad \text{by hypothesis (A) and Lemma 7.6 page 19} \\
& = \text{arg} \{ \phi[f(a)] | f(a) \leq f(x) \quad \forall x, a \in X \} \quad \text{because } \text{arg}_x \{ f(x) | P(x) \} = \text{arg}_x \{ g[f(x)] | P(x) \} \\
& \triangleq \bigwedge_{x \in X} \phi[f(x)] \quad \text{because } f \in X^X \text{ and by Lemma 2.12 page 5} \\
\end{align*} \]

Remark 7.8 Using the definitions of Lemma 7.7 (page 20), and letting \( g \) be a function in \( X^X \), and despite the results of Lemma 7.7, note that
\[ \begin{align*}
\text{A. } \phi \text{ is strictly isotone } \text{ and } \\
\text{B. } X \text{ is linearly ordered } \end{align*} \]
\[ \implies \begin{cases} 
\text{arg} \bigwedge_{x \in X} [f(x)g(x)] = \text{arg} \bigwedge_{x \in X} \phi[f(x)]g(x) 
\end{cases} \]

For example, let \( f(x) \triangleq x, g(x) \triangleq -x + 2, \text{ and } \phi(x) \triangleq x^2 \). Then
arg \left\{ \begin{array}{l}
\bigwedge_{x \in X} \left[ f(x)g(x) \right] = \arg \bigwedge_{x \in X} \left[ -x^2 + 2x \right] = 1 \neq \frac{4}{3} = \arg \bigcap_{x \in X} \left[ -x^3 + 2x^2 \right] = \arg \bigcap_{x \in X} \phi(f(x))g(x)
\end{array} \right.

\textbf{Lemma 7.9} Let \( X \triangleq (X, \lor, \land; \leq) \) and \( Y \triangleq (Y, \sqcup, \sqcap; \sqsubseteq) \) be LATTICES (Definition 2.13 page 6). Let \( f \) be a function in \((X \times X)^X\). Let \( \phi \) and \( \psi \) be functions in \( Y^X \).

\[ \begin{array}{ll}
A. & \phi \text{ is STRICTLY ISOTONE and } \\
B. & X \text{ is LINEARLY ORDERED } \implies \\
1. & \arg \bigwedge_{x \in X} \bigvee_{y \in X} f(x, y) = \arg \bigcap_{x \in X} \bigvee_{y \in X} \phi[f(x, y)] \quad \forall x, y \in X \quad \text{and} \\
2. & \arg \bigwedge_{x \in X} \bigvee_{y \in X} f(x, y) = \arg \bigcap_{x \in X} \bigvee_{y \in X} \phi[f(x, y)] \quad \forall x, y \in X \\
3. & \arg \bigvee_{x \in X} \bigwedge_{y \in X} f(x, y) = \arg \bigvee_{x \in X} \bigwedge_{y \in X} \psi[f(x, y)] \quad \forall x, y \in X \\
\end{array} \]

\textit{Proof:}

\[ \begin{array}{l}
\arg \bigcap_{x \in X} \bigcup_{y \in X} \phi[f(x, y)] = \arg \bigcap_{x \in X} \phi\left[ \bigvee_{y \in X} f(x, y) \right] \quad \text{by Lemma 7.5 (page 19)} \\
= \arg \bigwedge_{x \in X} \bigvee_{y \in X} f(x, y) \quad \text{by Lemma 7.2 (page 18)} \\
\arg \bigcap_{x \in X} \bigcup_{y \in X} \phi[f(x, y)] = \arg \bigcap_{x \in X} \phi\left[ \bigwedge_{y \in X} f(x, y) \right] \quad \text{by Lemma 7.5 (page 19)} \\
= \arg \bigwedge_{x \in X} \bigwedge_{y \in X} f(x, y) \quad \text{by Lemma 7.2 (page 18)} \\
\arg \bigcup_{x \in X} \bigcap_{y \in X} \phi[f(x, y)] = \arg \bigcup_{x \in X} \phi\left[ \bigwedge_{y \in X} f(x, y) \right] \quad \text{by Lemma 7.5 (page 19)} \\
= \arg \bigwedge_{x \in X} \bigwedge_{y \in X} f(x, y) \quad \text{by Lemma 7.2 (page 18)} \\
\end{array} \]

\textbf{Remark 7.10} Let \((X, \leq)\) be an ordered set (Definition 2.4 page 4). Let \( \phi \) be a function in \( X^X \). If \( \phi \) is strictly isotone then

\[ \begin{array}{ll}
\sum_{x \in X} f(x) < \sum_{x \in X} g(x) & \implies \phi\left[ \sum_{x \in X} f(x) \right] < \phi\left[ \sum_{x \in X} g(x) \right] \quad \text{but} \\
\sum_{x \in X} \phi[f(x)] < \sum_{x \in X} \phi[g(x)] & \implies \phi\left[ \sum_{x \in X} f(x) \right] < \phi\left[ \sum_{x \in X} g(x) \right].
\end{array} \]

\textit{Proof:}

(1) Proof for (1): this follows directly from the definition of isotone (Definition 7.1 page 17).

(2) Proof for (2): Let \( f(x) = \left( \frac{1}{2}, \frac{1}{2}, \frac{1}{2} \right) \), \( g(x) = (0, 0, 1) \), and \( \phi(x) = x^2 \).

Then \[ \sum_{x \in X} \phi[f(x)] = \left( \frac{1}{2} \right)^2 + \left( \frac{1}{2} \right)^2 + \left( \frac{1}{2} \right)^2 = \frac{3}{4} < 1 = 0^2 + 0^2 + 1^2 = \sum_{x \in X} \phi[g(x)] \]

but \[ \phi\left[ \sum_{x \in X} f(x) \right] = \left( \frac{1}{2} + \frac{1}{2} + \frac{1}{2} \right)^2 = \frac{9}{4} > 1 = (0 + 0 + 1)^2 = \phi\left[ \sum_{x \in X} g(x) \right]. \]
8 Outcome subspaces

8.1 Definitions

Traditional probability theory is performed in a probability space \( (\Omega, \mathcal{E}, \mathbb{P}) \). This section extends the probability space structure to include what herein is called an outcome subspace (next definition).

**Definition 8.1** An extended probability space is the tuple \( (\Omega, \mathcal{X}, \preceq, \mathcal{E}, \mathbb{P}) \) where \( (\Omega, \mathcal{E}, \mathbb{P}) \) is a probability space (Definition 5.3 page 14) and \( (\Omega, \mathcal{X}, \preceq) \) is an ordered quasi-metric space (Definition 6.1 page 16). The 4-tuple \( (\Omega, \mathcal{X}, \preceq, \mathcal{E}, \mathbb{P}) \) is an outcome subspace of the extended probability space \( (\Omega, \mathcal{X}, \preceq, \mathcal{E}, \mathbb{P}) \).

**Definition 8.2** Let \( G \triangleq (\Omega, \mathcal{X}, \preceq, \mathcal{E}, \mathbb{P}) \) be an outcome subspace (Definition 8.1 page 22). The \( n \)-th-moment \( m_n(x, y) \) from \( x \) to \( y \) in \( G \) is

\[
\mathbb{E}(x, y) := [d(x, y)]^n \mathbb{P}(y) \quad \forall x, y \in \Omega, n \in \mathbb{N}
\]

The moment \( m(x, y) \) from \( x \) to \( y \) in \( G \) is

\[
m(x, y) := m_1(x, y) \quad \forall x, y \in \Omega
\]

This paper introduces a quantity called the outcome center of an outcome subspace (next definition) which is in essence the same as the center of a graph (Definition 2.3 page 4).

**Definition 8.3** Let \( G \triangleq (\Omega, \mathcal{X}, \preceq, \mathcal{E}, \mathbb{P}) \) be an outcome subspace (Definition 8.1 page 22).

\[
\hat{c}(G) \triangleq \arg\min_{x \in \Omega} \max_{y \in \Omega} d(x, y) \mathbb{P}(y) \quad \text{is the outcome center of } G.
\]

The following additional definitions are of interest due in part to Corollary 3.4 (page 7) and the minimax inequality (Theorem 2.15 page 6). They are illustrated in several examples in this section. However, most of them are not used outside this section.

**Definition 8.4** Let \( G \triangleq (\Omega, \mathcal{X}, \preceq, \mathcal{E}, \mathbb{P}) \) be an outcome subspace (Definition 8.1 page 22).

\[
\hat{c}_a(G) \triangleq \arg\min_{x \in \Omega} \sum_{y \in \Omega} d(x, y) \mathbb{P}(y) \quad \text{is the arithmetic center of } G.
\]

\[
\hat{c}_g(G) \triangleq \arg\min_{x \in \Omega} \prod_{y \in \Omega \setminus \{x\}} [d(x, y)^\mathbb{P}(y)] \quad \text{is the geometric center of } G.
\]

\[
\hat{c}_h(G) \triangleq \arg\min_{x \in \Omega} \left( \sum_{y \in \Omega \setminus \{x\}} \frac{1}{d(x, y)} \mathbb{P}(y) \right)^{-1} \quad \text{is the harmonic center of } G.
\]

\[
\hat{c}_m(G) \triangleq \arg\min_{x \in \Omega} \min_{y \in \Omega \setminus \{x\}} d(x, y) \mathbb{P}(y) \quad \text{is the minimal center of } G.
\]

\[
\hat{c}_m(G) \triangleq \arg\max_{x \in \Omega} \min_{y \in \Omega \setminus \{x\}} d(x, y) \mathbb{P}(y) \quad \text{is the maxmin center of } G.
\]

In a manner similar to the traditional variance function (Definition 5.6 page 15), the outcome variance (next) is a kind of measure of the quality of the outcome center as a representative estimate of all the values of the outcome subspace. Said another way, it is in essence the expected error of the center measure.
Definition 8.5 Let $G \triangleq (\Omega, d, \leq, P)$ be an outcome subspace (Definition 8.1 page 22). The outcome variance $\text{Var}(G; \hat{\mathbf{c}}_x)$ of $G$ with respect to $\hat{\mathbf{c}}_x$ is \[ \text{Var}(G; \hat{\mathbf{c}}_x) \triangleq \sum_{x \in \Omega} d^2(\hat{\mathbf{c}}_x(G), x) P(x). \]

where $\hat{\mathbf{c}}_x$ is any of the operators defined in Definition 8.3 or Definition 8.4 (page 22). Moreover, $\text{Var}(G) \triangleq \text{Var}(G; \hat{\mathbf{c}})$, where $\hat{\mathbf{c}}$ is the outcome center (Definition 8.3 page 22).

Remark 8.6 The quantity $\mathbf{c}(x, \Omega) \triangleq \sum_{y \in \Omega} d(x, y) P(y)$ in the arithmetic center $\hat{\mathbf{c}}(G)$ (Definition 8.4 page 22) is itself a metric (Definition 4.2 page 8). Thus, $\hat{\mathbf{c}}(G)$ is the $x$ that produces the minimum of all the metrics with center $x$.

Proof: This follows directly from power mean metrics theorem with $r = 1$ (Theorem 4.19 page 11).

8.2 Examples

Example 8.7 (fair die outcome subspace / discrete outcome subspace)

Let $G \triangleq ({\Box}, {\bigcirc}, {\blacklozenge}, {\blacklozenge}, {\blacklozenge}, {\blacklozenge}, {\blacklozenge}, {\blacklozenge}, {\blacklozenge}, {\blacklozenge}, {\blacklozenge}, {\blacklozenge})$, $d, \leq, P$ be the outcome subspace (Definition 8.1 page 22) generated by a fair die where $d$ is the discrete metric (Example 4.21 page 14), $\leq \triangleq \emptyset$ (completely unordered set), and $P(\Box) = P(\bigcirc) = \cdots = P(\blacklozenge) = \frac{1}{6}$. This is illustrated by the weighted graph (Definition 2.2 page 4) to the right, where each line segment represents a distance of 1, and elements in the outcome center $\hat{\mathbf{c}}(G)$ are shaded. The structure has the following geometric values:

$\hat{\mathbf{c}}(G) = \hat{\mathbf{c}}_\bigcirc(G) = \hat{\mathbf{c}}_{\Box}(G) = \hat{\mathbf{c}}_{\blacklozenge}(G) = \hat{\mathbf{c}}_{\bigcirc}(G) = \hat{\mathbf{c}}_{\Box}(G) = \emptyset$ (definition of $\hat{\mathbf{c}}$ (Definition 8.3 page 22))

$\text{Var}(G) = \text{Var}(G; \hat{\mathbf{c}}_{\bigcirc}) = \text{Var}(G; \hat{\mathbf{c}}_{\Box}) = \text{Var}(G; \hat{\mathbf{c}}_{\blacklozenge}) = \text{Var}(G; \hat{\mathbf{c}}_{\bigcirc}) = \text{Var}(G; \hat{\mathbf{c}}_{\Box}) = \text{Var}(G; \hat{\mathbf{c}}_{\blacklozenge}) = \text{Var}(G; \hat{\mathbf{c}}_{\bigcirc}) = 0$

Proof:

$\hat{\mathbf{c}}(G) \triangleq \arg\min_{x \in G} \max_{y \in G} d(x, y) \frac{1}{6}$

by definition of $\hat{\mathbf{c}}$ (Definition 8.3 page 22)

$= \arg\min_{x \in G} \max_{y \in G} d(x, y) \frac{1}{6}$

by definition of $G$

$= \arg\min_{x \in G} \frac{1}{6} \{1, 1, 1, 1, 1, 1\}$

by definition of discrete metric (Example 4.21 page 14)

$= \{\Box, \bigcirc, \blacklozenge, \bigcirc, \blacklozenge, \bigcirc, \blacklozenge\}$

by definition of $G$

$\hat{\mathbf{c}}(G) \triangleq \arg\min_{x \in G} \sum_{y \in G} d(x, y) P(y)$

by definition of $\hat{\mathbf{c}}_x$ (Definition 8.4 page 22)

$= \arg\min_{x \in G} \sum_{y \in G} d(x, y) \frac{1}{6}$

by definition of $G$

$= \arg\min_{x \in G} \frac{1}{6} \{5, 5, 5, 5, 5, 5\}$

by definition of discrete metric (Example 4.21 page 14)

$= \{\Box, \bigcirc, \blacklozenge, \bigcirc, \blacklozenge, \bigcirc, \blacklozenge\}$

by definition of $G$

$\hat{\mathbf{c}}_x(G) \triangleq \arg\min_{x \in G} \prod_{y \in G \setminus \{x\}} d(x, y) P(y)$

by definition of $\hat{\mathbf{c}}_x$ (Definition 8.4 page 22)

$= \arg\min_{x \in G} \prod_{y \in G \setminus \{x\}} 1 \frac{1}{6}$

by definition of $G$ and discrete metric (Example 4.21 page 14)
8.2 EXAMPLES

Example 8.8 (real die outcome subspace)
Let $G \triangleq (\{\Box, \Diamond, \lozenge, \clubsuit, \heartsuit, \spadesuit\}, \emptyset, \leq, P)$ be the outcome subspace (Definition 8.1 page 22) generated by a real die. On a real die (as opposed to a fair die Example 8.7 page 23), some symbols (die faces) are physically closer (in real 3 dimensional space) than others. Instead of using the discrete metric, we define the distance $d(x, y)$ from face $x$ to face $y$ to be the number of physical die edges that must be crossed to go from $x$ to $y$. In this case, it is still true that $d(\Box, \Box) = 0$, and

\[
d(\Box, \Diamond) = d(\Box, \lozenge) = d(\Box, \clubsuit) = d(\Box, \heartsuit) = d(\Box, \spadesuit) = 1, \quad d(\Diamond, \lozenge) = d(\Diamond, \clubsuit) = d(\Diamond, \heartsuit) = d(\Diamond, \spadesuit) = 2, \quad d(\lozenge, \clubsuit) = d(\lozenge, \heartsuit) = d(\lozenge, \spadesuit) = 0
\]

This structure is illustrated by the weighted graph (Definition 2.2 page 4) to the right, where each line segment represents a distance of 1. The structure has the following geometric values:

\[
\bar{e}(G) = \hat{c}(G) = \hat{e}(G) = \hat{c}(G) = \hat{m}(G) = \hat{m}(G) = \{\Box, \Diamond, \lozenge, \clubsuit, \heartsuit, \spadesuit\}
\]

\[
\operatorname{Var}(G) = \operatorname{Var}(G; \hat{c}) = \operatorname{Var}(G; \hat{c}) = \operatorname{Var}(G; \hat{c}) = \operatorname{Var}(G; \hat{m}) = \operatorname{Var}(G; \hat{m}) = \operatorname{Var}(G) = \operatorname{Var}(G; \hat{m}) = \operatorname{Var}(G; \hat{m}) = 0
\]
PROOF:

\[ \hat{\mathcal{C}}(G) \triangleq \arg \min_{x \in G} \max_{y \in G} d(x, y) \mathcal{P}(y) \]  
by definition of \( \hat{\mathcal{C}} \) (Definition 8.3 page 22)

\[ = \arg \min_{x \in G} \max_{y \in G} d(x, y) \frac{1}{6} \]  
by definition of \( G \)

\[ = \arg \min_{x \in G} \left\{ \frac{1}{6} \{2, 2, 2, 2, 2, 2\} \right\} \]  
because for each \( x \), there is a \( y \) such that \( d(x, y) = 2 \)

\[ = \{ \Box, \Box, \Box, \Box, \Box, \Box \} \]  
by definition of \( G \)

\[ \hat{C}_6(G) \triangleq \arg \min_{x \in G} \sum_{y \in G} d(x, y) \mathcal{P}(y) \]  
by definition of \( \hat{\mathcal{C}} \) (Definition 8.4 page 22)

\[ \triangleq \arg \min_{x \in G} \frac{1}{6} \sum_{y \in G} d(x, y) \frac{1}{6} \]  
by definition of \( G \)

\[ = \arg \min_{x \in G} \left\{ \frac{1}{6} \left\{ \begin{array}{c} 0 + 1 + 1 + 1 + 1 + 2 \\
1 + 0 + 1 + 1 + 2 + 1 \\
1 + 1 + 1 + 2 + 0 + 1 + 1 \\
1 + 2 + 1 + 1 + 1 + 1 + 0 
\end{array} \right\} \right\} \]  
by definition of \( G \)

\[ \{ \Box, \Box, \Box, \Box, \Box, \Box \} \]  
by definition of \( \mathcal{P} \)

\[ \hat{\mathcal{C}}(\mathcal{X}) \triangleq \arg \min_{x \in \mathcal{G}(x)} \left\{ \sum_{y \in \mathcal{G}(x)} \frac{1}{d(x, y)} \mathcal{P}(y) \right\}^{-1} \]  
by definition of \( \hat{\mathcal{C}} \) (Definition 8.4 page 22)

\[ = \arg \min_{x \in \mathcal{G}(x)} \left\{ \sum_{y \in \mathcal{G}(x)} \frac{1}{d(x, y)} \frac{1}{6} \right\}^{-1} \]  
by definition of \( G \)

\[ = \arg \min_{x \in \mathcal{G}(x)} \left\{ \sum_{y \in \mathcal{G}(x)} \frac{1}{d(x, y)} \right\}^{-1} \]  

\[ = \arg \max_{x \in \mathcal{G}(x)} \frac{1}{6} \sum_{y \in \mathcal{G}(x)} \frac{2}{d(x, y)} \]  

\[ = \arg \min_{x \in \mathcal{G}(x)} \left\{ \begin{array}{c} 0 + 2 + 2 + 2 + 2 + 1 \\
2 + 2 + 0 + 2 + 2 + 1 + 1 \\
2 + 2 + 1 + 0 + 2 + 2 + 1 \\
1 + 2 + 2 + 2 + 0 + 2 + 1 \\
1 + 2 + 1 + 2 + 2 + 0 + 2 \\
1 + 2 + 2 + 2 + 0 + 2 + 1 
\end{array} \right\} \]  

\[ = \arg \min_{x \in \mathcal{G}(x)} \frac{1}{6} \left\{ \begin{array}{c} 9 \\
9 \\
9 \\
9 \\
9 \\
9 
\end{array} \right\} \]  

\[ = \{ \Box, \Box, \Box, \Box, \Box, \Box \} \]  
by definition of \( \mathcal{P} \)
\[ \hat{c}_m(G) \triangleq \arg \min_{x \in G} \min_{y \in G \setminus \{x\}} \delta(x, y) \] by definition of \( \hat{c}_m \) (Definition 8.4 page 22)

\[ = \arg \min_{x \in G} \min_{y \in G \setminus \{x\}} 1 \times \frac{1}{6} \] by definition of \( \hat{c}_m \) (Definition 8.3 page 22)

\[ = \frac{1}{6} \{2, 2, 2, 2, 2, 2\} \] by definition of \( G \)

\[ \hat{c}_M(G) \triangleq \arg \max_{x \in G} \min_{y \in G \setminus \{x\}} \psi(x, y) \] by definition of \( \hat{c}_M \) (Definition 8.4 page 22)

\[ = \{0, 0, 0, 0, 0, 0\} \] by \( \hat{c}_M(G) \) result

\[ \mathcal{V}(G; \hat{c}_M) = \mathcal{V}(G; \hat{c}_M) = \mathcal{V}(G; \hat{c}_M) = \mathcal{V}(G; \hat{c}_M) = \mathcal{V}(G) \]

\[ \triangleq \sum_{x \in G} \[d(\hat{c}_M(G), x)]^2 \] by definition of \( \mathcal{V} \) (Definition 8.5 page 23)

\[ = \sum_{x \in G} (0) \frac{1}{6} \] because \( \hat{c}_M(G) = G \)

\[ = 0 \] by field property of additive identity element 0

**Remark 8.9** Let \( G \triangleq (\Omega, d, \leq, P) \) be the *fair die outcome subspace* (Example 8.7 page 23). Let \( H \triangleq (\Omega, p, \leq, P) \) be the *real die outcome subspace* (Example 8.8 page 24). These two subspaces are identical except for their metrics \( d \) and \( p \). So we can say that \( G \) and \( H \) are distinguished by their metrics. However, note that they are *indistinguishable* by the topologies induced by their metrics, because they both induce the same topology—the *discrete topology* \( 2^\Omega \) (Definition 1.2 page 3). That is, the geometric distinction provided in metric spaces is in general lost in topological spaces. Thus, topological spaces are arguably too general for the type of stochastic processing presented in this paper; rather, the stochastic processing discussed in this paper calls for metric space structure. And in this paper, this type of metric space structure is referred to as *metric geometry*.

**Proof:**

1. Every *metric space* \((\Omega, d)\) (Definition 4.2 page 8) induces a *topological space* \((\Omega, T)\).
2. In particular, a *metric* \( d \) induces an *open ball* \( B(x, r) \in (2^\Omega)^{2 \times \mathbb{R}^+} \) centered at \( x \) with radius \( r \) such that \( B(x, r) \triangleq \{ y \in \Omega \mid d(x, y) < r \} \).
3. At each outcome \( x \) in \( G \), only two open balls are possible: \( B(x, r) = \left\{ \begin{array}{ll} \{x\} & \text{for } 0 < r \leq 1 \\ \Omega & \text{for } r > 1 \end{array} \right\} \).
4. Let \( x' \) represent the die face which, when its numeric value is summed “in the usual way” with the numeric value of the die face \( x \), equals 7. Then at each point \( x \) in \( H \), three open balls are possible:
   \( B(x, r) = \left\{ \begin{array}{ll} \{x\} & \text{for } 0 < r \leq 1 \\ \{x, x'\} & \text{for } 1 < r \leq 2 \\ \Omega & \text{for } r > 2 \end{array} \right\} \).
5. The open balls of \((\Omega, d)\) or \((\Omega, p)\) in turn induce a *base* for a topology \( T \), such that \( T = \{ U \in 2^\Omega \mid U \text{ is a union of open balls} \} \). The topology induced by \( G \) is the *discrete topology* \( 2^\Omega \) (Definition 1.2 page 3). The topology induced by \( H \) is also the *discrete topology* \( 2^\Omega \).
6. So the metrics of \( G \) and \( H \) are different. And the balls induced by \( G \) and those induced by \( H \) are different. However, the topologies induced by \( G \) and \( H \) are the same.
Example 8.10 (weighted die outcome subspace) A weighted die generates an outcome subspace \( G \) (Definition 8.1 page 22). The weighted die illustrated to the right has the following geometric values:

\[
\hat{\mathcal{C}}(G) = \mathcal{C}_a(G) = \{\Box\} \\
\mathcal{C}_b(G) = \{\square\} \\
\mathcal{C}_m(G) = \{\bigcirc\} \\
\mathcal{C}_m(G) = \{\bigcirc, \bigcirc, \bigcirc\} \\
\mathcal{C}_m(G) = \{\bigcirc, \bigcirc, \bigcirc\}
\]

\[\begin{align*}
\mathbb{V}(G) &= 0.33 \\
\mathbb{V}(G; \mathcal{C}_b) &= 0.847 \\
\mathbb{V}(G; \mathcal{C}_m) &= 0.22 \\
\mathbb{V}(G; \hat{\mathcal{C}}_b) &= 0.767
\end{align*}\]

Note that the outcome center \( \hat{\mathcal{C}}(G) \) and arithmetic center \( \mathcal{C}_b(G) \) again yield identical results. Also note that of the four center measures of cardinality \( |\mathcal{C}(G)| = |\mathcal{C}_b(G)| = |\mathcal{C}_m(G)| = 1 \) (Definition 1.6 page 3), \( \hat{\mathcal{C}} \) and \( \mathcal{C}_b \) yield by far the lowest variance measures.

\[\mathbb{P} \text{PROOF:}\]

\[
\hat{\mathcal{C}}(G) \triangleq \arg \min_{x \in G} \max_{y \in G} d(x, y) \mathbb{P}(y)
\]

by definition of \( \hat{\mathcal{C}} \) (Definition 8.3 page 22)

\[
\triangleq \arg \min_{x \in G} \max_{y \in G} \frac{1}{300} d(x, y) \mathbb{P}(y) 300
\]

by Lemma 7.9 (page 21)

\[
= \arg \min_{x \in G} \max_{y \in G} \begin{cases}
0 \times 30 & 1 \times 15 & 1 \times 10 & 1 \times 180 & 1 \times 6 & 2 \times 10 \\
1 \times 30 & 0 \times 15 & 1 \times 10 & 1 \times 180 & 2 \times 6 & 1 \times 10 \\
1 \times 30 & 1 \times 15 & 0 \times 10 & 2 \times 180 & 1 \times 6 & 1 \times 10 \\
1 \times 30 & 1 \times 15 & 2 \times 10 & 0 \times 180 & 1 \times 6 & 1 \times 10 \\
2 \times 30 & 1 \times 15 & 1 \times 10 & 1 \times 180 & 1 \times 6 & 0 \times 10
\end{cases}
\]

by definition of \( \hat{\mathcal{C}}_b \) (Definition 8.4 page 22)

\[
\hat{\mathcal{C}}_b(G) \triangleq \arg \min_{x \in G} \sum_{y \in G} d(x, y) \mathbb{P}(y)
\]

by definition of \( \hat{\mathcal{C}}_b \) (Definition 8.4 page 22)

\[
\frac{1}{300} \sum_{y \in G} d(x, y) \mathbb{P}(y) 300
\]

by definition of \( G \)

\[
= \arg \min_{x \in G} \sum_{y \in G} d(x, y) \mathbb{P}(y) 300
\]

by Lemma 7.2 (page 18)

\[
= \arg \min_{x \in G} \begin{cases}
0 + 15 + 10 + 180 + 6 + 20 \\
30 + 0 + 10 + 180 + 12 + 10 \\
30 + 15 + 20 + 0 + 6 + 10 \\
30 + 30 + 10 + 180 + 0 + 10 \\
60 + 15 + 10 + 180 + 6 + 0
\end{cases}
\]

by definition of \( \hat{\mathcal{C}}_b \) (Definition 8.4 page 22)

\[
\hat{\mathcal{C}}_b(G) \triangleq \arg \min_{x \in G} \prod_{y \in \mathcal{Q}_x} \left[ d(x, y)^{\mathbb{P}(y)} \right]
\]

by definition of \( \hat{\mathcal{C}}_b \) (Definition 8.4 page 22)

\[
= \arg \min_{x \in G} \prod_{y \in \mathcal{Q}_x} \left[ d(x, y)^{300 \mathbb{P}(y)} \right]^{\frac{1}{30}}
\]

(by definition of \( \hat{\mathcal{C}}_b \) (Definition 8.4 page 22)
\[
\hat{b}_n(G) \triangleq \arg \min_{x \in G} \min_{y \in G(G)} \prod_{y \in G(G)} \left[ d(x, y)^{300P(y)} \right]
\]
by Lemma 7.2 (page 18)

\[
\hat{c}_n(G) \triangleq \arg \min_{x \in G} \left\{ \sum_{y \in G(G)} \frac{1}{d(x, y)} P(y) \right\}^{-1}
\]
by definition of \(\hat{c}_n\) (Definition 8.4 page 22)

\[
\hat{c}_m(G) \triangleq \arg \min_{x \in G} \left\{ \sum_{y \in G(G)} \frac{300P(y)}{d(x, y)} \right\}
\]
by Lemma 7.2 (page 18)

\[
\hat{c}_m(G) \triangleq \arg \min_{x \in G} \min_{y \in G(G)} \left\{ \sum_{y \in G(G)} \frac{1}{d(x, y)} P(y) \right\}
\]
by definition of \(\hat{c}_m\) (Definition 8.4 page 22)

\[
\hat{c}_m(G) \triangleq \arg \min_{x \in G} \left\{ \sum_{y \in G(G)} \frac{1}{d(x, y)} P(y) \right\}
\]
by \(\hat{c}_m(G)\) result

\[
\hat{c}_b(G) \triangleq \arg \min_{x \in G} \left\{ \sum_{y \in G(G)} \frac{1}{d(x, y)} P(y) \right\}
\]
by \(\hat{c}_b(G)\) result

\[
\hat{c}_b(G) \triangleq \arg \min_{x \in G} \left\{ \sum_{y \in G(G)} \frac{1}{d(x, y)} P(y) \right\}
\]
by definition of \(\hat{c}_b\) (Definition 8.5 page 23)

\[
\hat{c}_b(G) \triangleq \arg \min_{x \in G} \left\{ \sum_{y \in G(G)} \frac{1}{d(x, y)} P(y) \right\}
\]
by \(\hat{c}_b(G)\) result
\[
\begin{align*}
&= 1^2 \times \frac{1}{10} + 0^2 \times \frac{1}{20} + 1^2 \times \frac{1}{30} + 1^2 \times \frac{3}{5} + 2^2 \times \frac{1}{50} + 1^2 \times \frac{1}{30} \\
&= \frac{1}{300} (30 + 0 + 10 + 180 + 24 + 10) = \frac{254}{300} = 0.847
\end{align*}
\]

\[
\text{Var}(\mathcal{G}; \hat{\epsilon}_h) \triangleq \sum_{x \in \mathcal{G}} d^2(\hat{\epsilon}_h(\mathcal{G}), x) \mathcal{P}(x)
\]

\[
= \sum_{x \in \mathcal{G}} d^2(\square, \square, \square, \square, \square, \square, x) \mathcal{P}(x)
\]

\[
= 1^2 \times \frac{1}{10} + 2^2 \times \frac{1}{20} + 1^2 \times \frac{1}{30} + 1^2 \times \frac{3}{5} + 0^2 \times \frac{1}{50} + 1^2 \times \frac{1}{30}
\]

\[
= \frac{1}{300} (30 + 60 + 10 + 180 + 0 + 10) = \frac{290}{300} = 0.967
\]

\[
\text{Var}(\mathcal{G}; \hat{\epsilon}_m) \triangleq \sum_{x \in \mathcal{G}} d^2(\hat{\epsilon}_m(\mathcal{G}), x) \mathcal{P}(x)
\]

\[
= \sum_{x \in \mathcal{G}} d^2(\square, \square, \square, \square, \square, \square, x) \mathcal{P}(x)
\]

\[
= 1^2 \times \frac{1}{10} + 0^2 \times \frac{1}{20} + 1^2 \times \frac{1}{30} + 1^2 \times \frac{3}{5} + 0^2 \times \frac{1}{50} + 1^2 \times \frac{1}{30}
\]

\[
= \frac{1}{300} (30 + 0 + 10 + 180 + 0 + 10) = \frac{230}{300} = 0.767
\]

\[
\text{Example 8.11} \quad \text{(board game spinner outcome subspace)}
\]

The six value \textit{spinner outcome subspace} (Definition 8.1 page 22) illustrated to the right has the following geometric values:

\[
\hat{\text{C}}(\mathcal{G}) = \hat{\text{C}}_a(\mathcal{G}) = \hat{\text{C}}_b(\mathcal{G}) = \hat{\text{C}}_h(\mathcal{G}) = \hat{\text{C}}_m(\mathcal{G}) = \hat{\text{C}}_M(\mathcal{G}) = \{1, 2, 3, 4, 5, 6\}
\]

\[
\text{Var}(\mathcal{G}) = \text{Var}(\mathcal{G}; \hat{\text{C}}_a) = \text{Var}(\mathcal{G}; \hat{\text{C}}_b) = \text{Var}(\mathcal{G}; \hat{\text{C}}_h) = \text{Var}(\mathcal{G}; \hat{\text{C}}_m) = \text{Var}(\mathcal{G}; \hat{\text{C}}_M) = 0
\]

\[
\text{Proof:}
\]

\[
\hat{\text{C}}(\mathcal{G}) \triangleq \arg \min_{x \in \mathcal{G}} \max_{y \in \mathcal{G}} d(x, y) \mathcal{P}(y)
\]

\[
= \arg \min_{x \in \mathcal{G}} \max_{y \in \mathcal{G}} d(x, y) \frac{1}{6}
\]

\[
= \arg \min_{x \in \mathcal{G}} \max_{y \in \mathcal{G}} d(x, y)
\]

because \(f(x) = \frac{1}{6}x\) is \textit{strictly isoton}e and by Lemma 7.9 (page 21)

\[
= \arg \min_{x \in \mathcal{G}} \max_{y \in \mathcal{G}} \begin{pmatrix}
\text{d(1,1)} & \ldots & \text{d(1,6)} \\
\text{d(2,1)} & \ldots & \text{d(2,6)} \\
\text{d(3,1)} & \ldots & \text{d(3,6)} \\
\text{d(4,1)} & \ldots & \text{d(4,6)} \\
\text{d(5,1)} & \ldots & \text{d(5,6)} \\
\text{d(6,1)} & \ldots & \text{d(6,6)}
\end{pmatrix}
\]

\[
= \text{arg min}_{x \in \mathcal{G}} \max_{y \in \mathcal{G}} \begin{pmatrix}
0 & 1 & 2 & 3 & 2 & 1 \\
1 & 0 & 1 & 2 & 3 & 2 \\
2 & 1 & 0 & 1 & 2 & 3 \\
3 & 2 & 1 & 0 & 1 & 2 \\
2 & 3 & 2 & 1 & 0 & 1 \\
1 & 2 & 3 & 2 & 1 & 0
\end{pmatrix}
\]

\[
= \text{arg min}_{x \in \mathcal{G}} \begin{pmatrix}
3 \\
3 \\
3 \\
3 \\
3 \\
3
\end{pmatrix}
\]

\[
\hat{\text{C}}_a(\mathcal{G}) \triangleq \arg \min_{x \in \mathcal{G}} \sum_{y \in \mathcal{G}} d(x, y) \mathcal{P}(y)
\]

by definition of \(\hat{\text{C}}_a\) (Definition 8.4 page 22)
\[ = \arg \min_{x \in G} \sum_{y \in G} d(x, y) \frac{1}{6} \quad \text{by definition of } G \]

\[ = \arg \min_{x \in G} \sum_{y \in G} d(x, y) \quad \text{because } f(x) = \frac{1}{6} x \text{ is strictly isotone and by Lemma 7.2 (page 18)} \]

\[ = \arg \min_{x \in G} \left\{ \begin{array}{c}
0 + 1 + 2 + 3 + 2 + 1 \\
1 + 0 + 1 + 2 + 3 + 2 \\
3 + 2 + 1 + 0 + 1 + 2 \\
2 + 3 + 2 + 1 + 0 + 1 \\
1 + 2 + 3 + 2 + 1 + 0 
\end{array} \right\} = \arg \min_{x \in G} \left\{ \begin{array}{c}
9 \\
9 \\
9 \\
9 \\
9 
\end{array} \right\} \]

\[ = \sum_{x \in G} \left( \begin{array}{c}
\prod_{y \in G} d(x, y) \end{array} \right) \]

\[ = \arg \min_{x \in G} \left\{ \begin{array}{c}
1 \\
1 \\
2 \\
3 \\
2 \\
1 
\end{array} \right\} = \arg \min_{x \in G} \left\{ \begin{array}{c}
12 \\
12 \\
12 \\
12 \\
12 \\
12 
\end{array} \right\} \]

\[ = \arg \max_{x \in G} \sum_{y \in G} \frac{1}{d(x, y)} \quad \text{by Lemma 7.2 (page 18)} \]

\[ = \arg \max_{x \in G} \left\{ \begin{array}{c}
+ 1 + 1 + 1 + 1 + 1 + 1 \\
+ 1 + 1 + 1 + 1 + 1 + 1 \\
+ 1 + 1 + 1 + 1 + 1 + 1 \\
+ 1 + 1 + 1 + 1 + 1 + 1 \\
+ 1 + 1 + 1 + 1 + 1 + 1 \\
+ 1 + 1 + 1 + 1 + 1 + 1 
\end{array} \right\} = \arg \max_{x \in G} \left\{ \begin{array}{c}
20 \\
20 \\
20 \\
20 \\
20 \\
20 
\end{array} \right\} \]

\[ \hat{\mu}_n(G) \triangleq \arg \min_{y \in G} d(x, y) P(y) \quad \text{by definition of } \hat{\mu}_n \text{ (Definition 8.4 page 22)} \]

\[ = \arg \min_{x \in G} \left\{ \begin{array}{c}
1, 1, 1, 1, 1 
\end{array} \right\} \]

\[ = \arg \min_{x \in G} \left\{ \begin{array}{c}
1, 1, 1, 1, 1 
\end{array} \right\} \]

\[ = \arg \min_{x \in G} \left\{ \begin{array}{c}
1, 1, 1, 1, 1 
\end{array} \right\} \]

\[ = \arg \min_{x \in G} \left\{ \begin{array}{c}
1, 1, 1, 1, 1 
\end{array} \right\} \]

\[ = \arg \min_{x \in G} \left\{ \begin{array}{c}
1, 1, 1, 1, 1 
\end{array} \right\} \]

\[ = \arg \min_{x \in G} \left\{ \begin{array}{c}
1, 1, 1, 1, 1 
\end{array} \right\} \]

\[ = \arg \min_{x \in G} \left\{ \begin{array}{c}
1, 1, 1, 1, 1 
\end{array} \right\} \]

\[ = \arg \min_{x \in G} \left\{ \begin{array}{c}
1, 1, 1, 1, 1 
\end{array} \right\} \]
\[\{\Box, \Diamond, \lozenge, \circ, \square, \bigcirc\}\] by definition of \(G\)

\[\hat{\mu}_m(G) \triangleq \arg\max_{x \in G} \min_{y \in \mathcal{O}(x)} d(x, y)P(y)\] by definition of \(\hat{\mu}_m\) (Definition 8.4 page 22)

\[= \arg\max_{x \in G} \{1, 1, 1, 1, 1, 1\}\] by \(\hat{\mu}_m(G)\) result

\[= \{\Box, \Diamond, \lozenge, \circ, \square, \bigcirc\}\] by definition of \(G\)

\[\var(G; \hat{\mu}_G) = \var(G; \hat{\mu}_k) = \var(G; \hat{\mu}_m) = \var(G; \hat{\mu}_h) = \var(G)\]

\[= \sum_{x \in G} \left[d\left(\hat{\mu}(G), x\right)\right]^2P(x)\] by definition of \(\var\) (Definition 8.5 page 23)

\[= \sum_{x \in G} (0^2) \frac{1}{6}\] because \(\hat{\mu}(G) = G\)

\[= 0\] by field property of additive identity element 0

\[\text{Example 8.12 (weighted spinner outcome subspace)}\]

The six value \textit{weighted spinner outcome subspace} \(G\) (Definition 8.1 page 22) illustrated to the right has the following geometric values:

\[
\begin{align*}
\hat{\mu}(G) &= \hat{\mu}_k(G) = \hat{\mu}_m(G) = \{1, 6\} \\
\var(G) &= \hat{\mu}_h(G) = \{2, 5\} \\
\var(G; \hat{\mu}_h) &= \{1, 2, 3, 4, 5, 6\} \\
\hat{\mu}(G) &= \hat{\mu}_m(G) = \{1, 6\}
\end{align*}
\]

The \textit{outcome center} result is used later in Example 9.16 (page 54). Note that, unlike the \textit{weighted die outcome subspace} (Example 8.10 page 27), of the center measures of cardinality 2 or less, the \textit{harmonic center} \(\hat{\mu}_h(G)\) yields the lowest \textit{outcome variance} (Definition 8.5 page 23). This is surprising since it suggests that \(\hat{\mu}_h(G)\) is superior to all the other \textit{center measures} (Definition 8.3 page 22, Definition 8.4 page 22), but yet unlike the other center measures, it yields center values that are not maximally likely.

\[\text{\textit{PROOF}}:\]

\[
\begin{align*}
\hat{\mu}(G) &\triangleq \arg\min_{x \in G} \max_{y \in \mathcal{O}(x)} d(x, y)P(y) \\
= \arg\min_{x \in G} \sum_{y \in G} \left[d(x, y)P(y)\right] \\
\hat{\mu}(G) &\triangleq \arg\min_{x \in G} \sum_{y \in \mathcal{O}(x)} d(x, y)P(y) \\
= \arg\min_{x \in G} \prod_{y \in \mathcal{O}(x)} d(x, y)^{P(y)}
\end{align*}
\] by definition of \(\hat{\mu}\) (Definition 8.3 page 22), \(\hat{\mu}_h\) (Definition 8.4 page 22)
= \arg \min_{x \in G} \prod_{y \in \mathcal{Q}(x)} \left[ d(x, y)^{6p(y)} \right]^{\frac{1}{6}}

= \arg \min_{x \in G} \left( \prod_{y \in \mathcal{Q}(x)} [d(x, y)^{6p(y)}] \right)^{\frac{1}{6}}

= \arg \min_{x \in G} \prod_{y \in \mathcal{Q}(x)} [d(x, y)^{6p(y)}] \quad \text{by Lemma 7.2 (page 18)}

\hat{\xi}(G) \triangleq \arg \min_{x \in G} \left( \sum_{y \in \mathcal{Q}(x)} \frac{1}{d(x, y)} P(y) \right)^{-1}

\hat{\zeta}_n(G) \triangleq \arg \min_{x \in G} \left( \sum_{y \in \mathcal{Q}(x)} \frac{1}{d(x, y)} P(y) \right)^{-1}

= \arg \max_{x \in G} \sum_{y \in \mathcal{Q}(x)} \frac{1}{d(x, y)} P(y) \quad \text{because } \phi(x) \triangleq x^{-1} \text{ is strictly antitone and by Lemma 7.6 page 19}

= \arg \max_{x \in G} \sum_{y \in \mathcal{Q}(x)} \frac{6P(y)}{d(x, y)} \quad \text{by Lemma 7.2 (page 18)}

\hat{\zeta}_m(G) \triangleq \arg \min_{x \in G} d(x, y) P(y) \quad \text{by definition of } \hat{\zeta}_m \text{ (Definition 8.4 page 22)}

= \arg \min_{x \in G} \min_{y \in \mathcal{Q}(x)} \left[ \begin{array}{ccc} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 6 & 1 & 2 \end{array} \right] \quad = \arg \min_{x \in G} \left[ \begin{array}{c} 1 \\ 1 \\ 6 \end{array} \right] \quad = \left[ \begin{array}{c} 2 \\ 3 \\ 5 \end{array} \right]

\hat{\zeta}_M(G) \triangleq \arg \max_{x \in G} d(x, y) P(y) \quad \text{by definition of } \hat{\zeta}_M \text{ (Definition 8.4 page 22)}

= \arg \max_{x \in G} \left\{ 1, 1, 1, 1, 1 \right\} \quad \text{by } \hat{\zeta}_m(G) \text{ result}

= \left\{ 1, 2, 3, 4, 5, 6 \right\}

\hat{\var}(G; \hat{\zeta}_n) = \hat{\var}(G; \hat{\zeta}_m) = \hat{\var}(G) \quad \text{by definition of } \hat{\var} \text{ (Definition 8.5 page 23)}

= \sum_{x \in G} d^2(\hat{\zeta}(G), x) P(x) \quad \text{by definition of } \hat{\var}(G) \text{ result}

= \sum_{x \in G} d^2(\left\{ 1, 6 \right\}, x) P(x)

= \left( 0 \right)^2 \frac{3}{6} + (1)^2 \frac{1}{6} + (2)^2 \frac{1}{6} + (2)^2 \frac{1}{6} + (1)^2 \frac{1}{6} + (0)^2 \frac{3}{6}
\[ \alpha = \frac{10}{6} = \frac{5}{3} = \frac{2}{3} \approx 1.667 \]

\[ \mathbf{Var}(G; \hat{\mu}_h) \triangleq \sum_{x \in G} d^2(\hat{\mu}_h(G), x) \mathbb{P}(x) \quad \text{by definition of } \mathbf{Var} \text{ (Definition 8.5 page 23)} \]

\[ = \sum_{x \in G} d^2([2, 5], x) \mathbb{P}(x) \quad \text{by } \hat{\mu}_h(G) \text{ result} \]

\[ = (1)^2 \frac{3}{6} + (0)^2 \frac{1}{6} + (1)^2 \frac{1}{6} + (1)^2 \frac{1}{6} + (0)^2 \frac{1}{6} + (1)^2 \frac{3}{6} \]

\[ = \frac{8}{6} = \frac{4}{3} \approx 1.333 \]

\[ \mathbf{Var}(G; \hat{\mu}_m) = \mathbf{Var}(G; \hat{\mu}_n) \]

\[ \triangleq \sum_{x \in G} d^2(\hat{\mu}_m(G), x) \mathbb{P}(x) \quad \text{by definition of } \mathbf{Var} \text{ (Definition 8.5 page 23)} \]

\[ = \sum_{x \in G} d^2([1, 2, 3, 4, 5, 6], x) \mathbb{P}(x) \quad \text{by } \hat{\mu}_n(G) \text{ result} \]

\[ = \sum_{x \in G} 0^2 \mathbb{P}(x) = 0 \]

\[ \mathbf{Example 8.13} \text{ (weighted ring)} \text{ The weighted five element ring illustrated to the right has the geometric values below. The outcome center result is used later in Example 9.17 (page 55).} \]

\[ \hat{\xi}(G) = \{4\} \quad \mathbf{Var}(G) = \frac{11}{9} \approx 1.222 \]

\[ \hat{\xi}_g(G) = \{3, 4\} \quad \mathbf{Var}(G; \hat{\xi}_g) = \frac{7}{9} \approx 0.778 \]

\[ \hat{\xi}_b(G) = \{3\} \quad \mathbf{Var}(G; \hat{\xi}_b) = \frac{16}{9} \approx 1.778 \]

\[ \hat{\xi}_m(G) = \{1, 2, 3\} \quad \mathbf{Var}(G; \hat{\xi}_m) = \frac{5}{9} \approx 0.556 \]

\[ \hat{\xi}_m(G) = \{1, 2\} \quad \mathbf{Var}(G; \hat{\xi}_m) = \frac{2}{9} \approx 0.222 \]

\[ \hat{\xi}_p(G) = \{1, 2\} \quad \mathbf{Var}(G; \hat{\xi}_p) = \frac{16}{9} \approx 1.778 \]

\[ \text{⚠️ \text{PROOF:}} \]

\[ \hat{\xi}(G) \triangleq \arg \min_{x \in G} \max_{y \in G} d(x, y) \mathbb{P}(y) \quad \text{by definition of } \hat{\xi} \text{ (Definition 8.3 page 22)} \]

\[ = \arg \min_{x \in G} \max_{y \in G} \frac{1}{9} d(x, y) \mathbb{P}(y) \]

\[ = \arg \min_{x \in G} \max_{y \in G} d(x, y) \mathbb{P}(y) \]

\[ = \arg \min_{x \in G} \max_{y \in G} \begin{bmatrix} d(0, 0) \mathbb{P}(0) & d(0, 1) \mathbb{P}(1) & d(0, 2) \mathbb{P}(2) & d(0, 3) \mathbb{P}(3) & d(0, 4) \mathbb{P}(4) \\ d(1, 0) \mathbb{P}(0) & d(1, 1) \mathbb{P}(1) & d(1, 2) \mathbb{P}(2) & d(1, 3) \mathbb{P}(3) & d(1, 4) \mathbb{P}(4) \\ d(2, 0) \mathbb{P}(0) & d(2, 1) \mathbb{P}(1) & d(2, 2) \mathbb{P}(2) & d(2, 3) \mathbb{P}(3) & d(2, 4) \mathbb{P}(4) \\ d(3, 0) \mathbb{P}(0) & d(3, 1) \mathbb{P}(1) & d(3, 2) \mathbb{P}(2) & d(3, 3) \mathbb{P}(3) & d(3, 4) \mathbb{P}(4) \\ d(4, 0) \mathbb{P}(0) & d(4, 1) \mathbb{P}(1) & d(4, 2) \mathbb{P}(2) & d(4, 3) \mathbb{P}(3) & d(4, 4) \mathbb{P}(4) \end{bmatrix} \]

\[ = \arg \min_{x \in G} \begin{bmatrix} 4 \\ 6 \\ 4 \\ 6 \\ 4 \end{bmatrix} \quad = \begin{bmatrix} 4 \\ 2 \\ 4 \end{bmatrix} \]

\[ \hat{\xi}_m(G) \triangleq \arg \min_{x \in G} \sum_{y \in G} d(x, y) \mathbb{P}(y) \quad \text{by definition of } \hat{\xi}_m \text{ (Definition 8.4 page 22)} \]

\[ = \arg \min_{x \in G} \sum_{y \in G} \frac{1}{9} d(x, y) \mathbb{P}(y) \]
\[
\hat{\gamma}(G) \triangleq \arg \min_{x \in \mathcal{G}} \prod_{y \in \mathcal{G}(x)} d(x, y)^{P(y)}
\]
by definition of \( \hat{\gamma} \) (Definition 8.4 page 22)

\[
\hat{\gamma}(G) \triangleq \arg \min_{x \in \mathcal{G}} \left[ \prod_{y \in \mathcal{G}(x)} d(x, y)^{\phi(y)} \right]^\frac{1}{\phi}
\]

because \( f(x) \triangleq x^\frac{1}{\phi} \) is strictly isotone and by Lemma 7.2 (page 18)

\[
\hat{\gamma}(G) \triangleq \arg \max_{x \in \mathcal{G}} \left( \sum_{y \in \mathcal{G}} \frac{1}{d(x, y)} P(y) \right)^{-1}
\]
by definition of \( \hat{\gamma} \) (Definition 8.4 page 22)

8.2 EXAMPLES

because \( f(x) = \frac{1}{9} x \) is strictly isotone and by Lemma 7.2 (page 18)

\[
\hat{\gamma}(G) \triangleq \arg \max_{x \in \mathcal{G}} \sum_{y \in \mathcal{G}} \frac{\phi(y)}{d(x, y)}
\]

because \( \phi(x) \triangleq x^{-1} \) is strictly antitone and by Lemma 7.6 page 19

\[
\hat{\gamma}(G) \triangleq \min_{x \in \mathcal{G}(x)} d(x, y) P(y)
\]
by definition of \( \hat{\gamma} \) (Definition 8.4 page 22)

because \( f(x) = \frac{1}{9} x \) is strictly isotone and by Lemma 7.2 (page 18)
\[
\text{Example 8.14} \quad \text{The weighted five element structure illustrated to the right has the following geometric values:}
\]

\[
\begin{align*}
\hat{\mathcal{U}}(G) &= \{3\} & \hat{\mathcal{U}}(G) &= \{1, 2, 3, 4\} & \text{Var}(G; \hat{\mathcal{U}}) &= \frac{10}{9} \\ 
\hat{\mathcal{U}}_{\mathcal{S}}(G) &= \{3, 4\} & \hat{\mathcal{U}}(G) &= \{1, 2, 3\} & \text{Var}(G; \hat{\mathcal{U}}_{\mathcal{S}}) &= \frac{1}{9} \\ 
\hat{\mathcal{U}}_{\mathcal{K}}(G) &= \{1, 2, 3\} & \hat{\mathcal{U}}(G) &= \{0, 3, 4\} & \text{Var}(G; \hat{\mathcal{U}}_{\mathcal{K}}) &= \frac{1}{9} \\ 
\hat{\mathcal{U}}_{\mathcal{M}}(G) &= \{1, 2\} & \hat{\mathcal{U}}(G) &= \{1, 2\} & \text{Var}(G; \hat{\mathcal{U}}_{\mathcal{M}}) &= \frac{2}{9}
\end{align*}
\]
The *outcome center* result is used later in Example 9.18 (page 58). Note that only the operators \( \hat{\mathcal{C}} \) and \( \hat{\mathcal{C}}_g \) were able to successfully isolate a single center point (\( | \hat{\mathcal{C}}(G) | = | \hat{\mathcal{C}}_g(G) | = | \{3\} | = 1 \)).

**Proof:**

\[
\hat{\mathcal{C}}(G) \triangleq \arg \min_{x \in G} \max_{y \in G} d(x, y) \mathcal{P}(y)
\]

by definition of \( \hat{\mathcal{C}} \) (Definition 8.3 page 22)

\[
= \arg \min_{x \in G} \frac{1}{9} \max_{y \in G} d(x, y) \mathcal{P}(y)
\]

because \( f(x) = \frac{1}{9} x \) is *strictly isotone* and by Lemma 7.9 (page 21)

\[
= \arg \min_{x \in G} \frac{1}{9} \max_{y \in G} \begin{bmatrix}
 d(0, 0) \mathcal{P}(0) \\
 d(0, 1) \mathcal{P}(1) \\
 d(1, 0) \mathcal{P}(0) \\
 d(1, 1) \mathcal{P}(1)
\end{bmatrix}
\]

\[
= \arg \min_{x \in G} \frac{1}{9} \max_{y \in G} \begin{bmatrix}
 0 \times 2 + 2 \times 1 + 1 \times 1 + 1 \times 2 + 2 \times 3 \\
 2 \times 2 + 0 \times 1 + 2 \times 1 + 1 \times 2 + 1 \times 3 \\
 1 \times 2 + 2 \times 1 + 0 \times 1 + 2 \times 2 + 1 \times 3 \\
 1 \times 2 + 1 \times 1 + 2 \times 1 + 0 \times 2 + 1 \times 3 \\
 2 \times 2 + 1 \times 1 + 1 \times 1 + 1 \times 2 + 0 \times 3
\end{bmatrix}
\]

because \( \mathcal{P}(x) = \frac{1}{9} x \) is *strictly isotone* and by Lemma 7.2 (page 18)

\[
= \arg \min_{x \in G} \begin{bmatrix}
 6 \\
 4 \\
 3 \\
 4 \\
 3
\end{bmatrix}
\]

\[
= \arg \min_{x \in G} \begin{bmatrix}
 3 \\
 4 \\
 3 \\
 4
\end{bmatrix}
\]

\[
\hat{\mathcal{C}}_g(G) \triangleq \arg \min_{x \in G} \prod_{y \in G \setminus \{x\}} d(x, y) \mathcal{P}(y)
\]

by definition of \( \hat{\mathcal{C}}_g \) (Definition 8.4 page 22)

\[
= \arg \min_{x \in G} \prod_{y \in G \setminus \{x\}} d(x, y) \mathcal{P}(y)
\]

because \( f(x) \triangleq x^{\frac{1}{1}} \) is *strictly isotone* and by Lemma 7.2 (page 18)

\[
= \arg \min_{x \in G} \prod_{y \in G \setminus \{x\}} d(x, y)
\]

\[
= \arg \min_{x \in G} \begin{bmatrix}
 2^2 \times 1^1 \times 1^1 \times 2^1 \times 1^2 \\
 2^2 \times 1^1 \times 2^1 \times 1^1 \\
 2^2 \times 1^1 \times 1^2 \times 1^1
\end{bmatrix}
\]

\[
= \arg \min_{x \in G} \begin{bmatrix}
 2^4 \\
 2^3 \\
 2^1
\end{bmatrix}
\]

\[
= \arg \min_{x \in G} \begin{bmatrix}
 3 \\
 4 \\
 3
\end{bmatrix}
\]

\[
\hat{\mathcal{C}}_r(G) \triangleq \arg \min_{x \in \Omega} \left( \sum_{y \in \Omega} \frac{1}{d(x, y)} \mathcal{P}(y) \right)^{-1}
\]

by definition of \( \hat{\mathcal{C}}_r \) (Definition 8.4 page 22)

\[
= \arg \max_{x \in \Omega} \left( \sum_{y \in \Omega} \frac{1}{d(x, y)} \mathcal{P}(y) \right)
\]

because \( \phi(x) \triangleq x^{-1} \) is *strictly antitone* and by Lemma 7.6 page 19

\[
= \arg \max_{x \in \Omega} \left( \frac{1}{9} \sum_{y \in \Omega} \frac{1}{d(x, y)} \mathcal{P}(y) \right)
\]

\[
= \arg \max_{x \in \Omega} \frac{\mathcal{P}(y)}{d(x, y)}
\]

because \( f(x) = \frac{1}{9} x \) is *strictly isotone* and by Lemma 7.2 (page 18)
\[ \hat{\nu}_m(G) \triangleq \arg \max_{x \in G} \min_{y \in G(x)} d(x, y) P(y) \]

by definition of \( \hat{\nu}_m \) (Definition 8.4 page 22)

\[ = \arg \max_{x \in G} \min_{y \in G(x)} \frac{1}{9} d(x, y) P(y) \]

because \( f(x) = \frac{1}{9} x \) is strictly isotone and by Lemma 7.9 (page 21)

\[ = \arg \max_{x \in G} \min_{y \in G(x)} \left\{ \begin{array}{ll} 2 \times 2 & 1 \times 1 \\ 2 \times 1 & 1 \times 2 \\ 1 \times 2 & 2 \times 1 \\ 2 \times 2 & 1 \times 1 \end{array} \right\} = \arg \max_{x \in G} \left\{ \begin{array}{l} 2 \\ 2 \\ 1 \end{array} \right\} = \left\{ \begin{array}{l} 3 \\ 4 \end{array} \right\} \]

\[ \hat{\nu}_m(G) \triangleq \arg \max_{x \in G} \min_{y \in G(x)} d(x, y) P(y) \]

by definition of \( \hat{\nu}_m \) (Definition 8.4 page 22)

\[ = \arg \max_{x \in G} \{1, 2, 2, 1, 1\} \]

by \( \hat{\nu}_m(G) \) result

\[ = \{1, 2\} \]

\[ \hat{\nu}(G) = \hat{\nu}_m(G) \]

by \( \hat{\nu}(G) \) and \( \hat{\nu}_m(G) \) results

\[ = \arg \max_{x \in G} \{1, 2, 2, 1, 1\} \]

by definition of \( \hat{\nu}(G) \) (Definition 8.5 page 23)

\[ \hat{\nu}(G) = \hat{\nu}_m(G) = \{1, 2\} \]

by definition of \( \hat{\nu}(G) \) (Definition 8.5 page 23)

\[ \hat{\nu}(G) = \hat{\nu}_m(G) = \{1, 2\} \]

by definition of \( \hat{\nu}(G) \) (Definition 8.5 page 23)

\[ \hat{\nu}(G) = \hat{\nu}_m(G) = \{1, 2\} \]

by definition of \( \hat{\nu}(G) \) (Definition 8.5 page 23)

\[ \hat{\nu}(G) = \hat{\nu}_m(G) = \{1, 2\} \]

by definition of \( \hat{\nu}(G) \) (Definition 8.5 page 23)

\[ \hat{\nu}(G) = \hat{\nu}_m(G) = \{1, 2\} \]

by definition of \( \hat{\nu}(G) \) (Definition 8.5 page 23)

\[ \hat{\nu}(G) = \hat{\nu}_m(G) = \{1, 2\} \]

by definition of \( \hat{\nu}(G) \) (Definition 8.5 page 23)

\[ \hat{\nu}(G) = \hat{\nu}_m(G) = \{1, 2\} \]

by definition of \( \hat{\nu}(G) \) (Definition 8.5 page 23)

\[ \hat{\nu}(G) = \hat{\nu}_m(G) = \{1, 2\} \]

by definition of \( \hat{\nu}(G) \) (Definition 8.5 page 23)
\[
\sum_{x \in \mathcal{G}} d^2((1, 2), x) P(x) = (1)^2 \frac{2}{9} + (0)^2 \frac{1}{9} + (0)^2 \frac{1}{9} + (1)^2 \frac{2}{9} + (1)^2 \frac{3}{9} = \frac{7}{9} \approx 0.778
\]

**Example 8.15** The outcome subspace (Definition 8.1 page 22) illustrated to the right, with a quasi-metric (Definition 4.1 page 8) has the following geometric values:

\[
\mathring{\mathcal{C}}(\mathcal{G}) = 1, \quad \mathring{\mathcal{C}}(\mathcal{G}) = \{3\}, \quad \mathcal{V}(\mathcal{G}) = \frac{12}{3} \approx 1.333
\]

\[
\mathring{\mathcal{M}}(\mathcal{G}) = \{0, 4\}, \quad \mathcal{V}(\mathcal{G}; \mathring{\mathcal{M}}) = \frac{10}{9} \approx 1.111
\]

\[
\mathcal{I}(\mathcal{G}) = \{0, 4\}, \quad 1 \times 10 \approx 16.667
\]

\[
\mathcal{M}(\mathcal{G}) = \{1, 2, 3\}, \quad \mathcal{V}(\mathcal{G}; \mathcal{M}) = \frac{5}{9} \approx 0.555
\]

This is the first example in this section to use a directed graph (rather than an undirected graph (Definition 2.1 page 4)) and to require the use of a quasi-metric (Definition 4.1 page 8) that is not a metric. Unlike Example 8.14 (page 35), which had neither of these restrictions, twice as many center operators (4 rather than 2) were able to successfully isolate a single center point. The outcome center result is used later in Example 9.19 (page 59).

**PROOF:**

\[
\mathring{\mathcal{U}}(\mathcal{G}) \overset{\text{def}}{=} \arg \min_{x \in \mathcal{G}} \max_{y \in \mathcal{G}} d(x, y) P(y) \quad \text{by definition of } \mathring{\mathcal{U}} \text{ (Definition 8.3 page 22)}
\]

\[
\mathring{\mathcal{M}}(\mathcal{G}) \overset{\text{def}}{=} \arg \min_{x \in \mathcal{G}} \max_{y \in \mathcal{G}} \frac{1}{9} d(x, y) P(y) \quad \text{because } f(x) = \frac{1}{9} x \text{ is \emph{strictly isotone} and by Lemma 7.9 (page 21)}
\]

\[
\mathcal{I}(\mathcal{G}) \overset{\text{def}}{=} \arg \min_{x \in \mathcal{G}} \prod_{y \in \mathcal{G} \setminus \{x\}} d(x, y) P(y) \quad \text{by definition of } \mathcal{I} \text{ (Definition 8.4 page 22)}
\]
\[
\begin{align*}
\arg \min_{x \in G} \left[ \prod_{y \in G \setminus \{x\}} d(x, y)^{q_p(y)} \right]^\frac{1}{q_p(y)} \\
= \arg \min_{x \in G} \prod_{y \in G \setminus \{x\}} d(x, y)^{q_p(y)} \quad \text{because } f(x) \triangleq x^\frac{1}{q} \text{ is strictly isotone} \text{ and by Lemma 7.2 (page 18)}
\end{align*}
\]
The physical geometry inducing a stochastic process and the *metric geometry* (Remark 8.9 page 26) of the stochastic process itself may be very different, as illustrated in the next two examples.

**Example 8.16** (archery) Consider the archery target illustrated to the left. It consists of several concentric circles, each with a different point value. However, its' *outcome subspace* structure, has a very different geometry, as illustrated to the right. Assuming uniform distribution, the graph center is shaded in the illustration to right.

**Example 8.17** (darts) Consider the simplified dart board illustrated to the left and *outcome subspace* illustrated to the right. Unlike the archery *outcome subspace* (Example 8.16 page 40), the outcome subspace is non-linear. Assuming uniform distribution, the graph center is shaded in the illustration.

**Example 8.18** (DNA) *Genomic Signal Processing* (GSP) analyzes biological sequences called *genomes*. These sequences are constructed over a set of 4 symbols that are commonly referred to as A, T, C, and G, each of which corresponds to a nucleobase (adenine, thymine, cytosine, and guanine, respectively). A typical genome sequence contains a large number of symbols (about 3 billion for humans, 29751 for the SARS virus).
Let $\mathcal{G} \triangleq (\{A, T, C, G\}, d, \leq, \mathcal{P})$ be the outcome subspace (Definition 8.1 page 22) generated by a genome where $d$ is the discrete metric (Example 4.21 page 14), $\leq \triangleq \emptyset$ (completely unordered set), and $\mathcal{P}(A) = \mathcal{P}(T) = \mathcal{P}(C) = \mathcal{P}(G) = \frac{1}{4}$. This space is illustrated by the graph (Definition 2.1 page 4) to the right with shaded center (Definition 8.3 page 22).

The graph has the following geometric values:

- $\mathcal{C}(\mathcal{G}) = \{A, T, C, G\}$ (shaded in illustration)
- $\mathcal{C}_0(\mathcal{G}) = \{A, T, C, G\}$ (shaded in illustration)
- $\mathsf{Var}(\mathcal{G}) = 0$ (Definition 8.5 page 23)

Proof:

- $\mathcal{C}(\mathcal{G}) = \arg \min_{x \in \mathcal{G}} \max_{y \in \mathcal{G}} d(x, y)$

  - by definition of $\mathcal{C}$ (Definition 8.3 page 22)

- $\mathcal{C}(\mathcal{G}) = \arg \min_{x \in \mathcal{G}} \max_{y \in \mathcal{G}} d(x, y) \frac{1}{4}$

  - by definition of $\mathcal{G}$

- $\mathcal{C}(\mathcal{G}) = \arg \min_{x \in \mathcal{G}} \max_{y \in \mathcal{G}} d(x, y)$

  - because $f(x) = \frac{1}{4} x$ is strictly isotone and by Lemma 7.9 (page 21)

- $\mathcal{C}(\mathcal{G}) = \{A, T, C, G\}$

  - by definition of $\mathcal{G}$

- $\mathcal{C}_0(\mathcal{G}) = \arg \min_{x \in \mathcal{G}} \sum_{y \in \mathcal{G}} d(x, y)$

  - by definition of $\mathcal{C}_0$ (Definition 8.4 page 22)

- $\mathcal{C}_0(\mathcal{G}) = \arg \min_{x \in \mathcal{G}} \sum_{y \in \mathcal{G}} d(x, y) \frac{1}{6}$

  - by definition of $\mathcal{G}$

- $\mathcal{C}_0(\mathcal{G}) = \arg \min_{x \in \mathcal{G}} \sum_{y \in \mathcal{G}} d(x, y)$

  - because $f(x) = \frac{1}{4} x$ is strictly isotone and by Lemma 7.9 (page 21)

- $\mathcal{C}_0(\mathcal{G}) = \arg \min_{x \in \mathcal{G}} \left\{ \begin{array}{c} 0 + 1 + 1 + 1 \\ 1 + 0 + 1 + 1 \\ 1 + 1 + 0 + 1 \\ 1 + 1 + 1 + 0 \end{array} \right\}$

  - $\{A, T, C, G\}$

- $\mathsf{Var}(\mathcal{G}) = \sum_{x \in \mathcal{G}} [d(\mathcal{C}(\mathcal{G}), x)]^2 \mathcal{P}(x)$

- by definition of $\mathsf{Var}$ (Definition 8.5 page 23)

- $\mathsf{Var}(\mathcal{G}) = \sum_{x \in \mathcal{G}} [d(\mathcal{C}(\mathcal{G}), x)]^2 \mathcal{P}(x)$

- because $\mathcal{C}(\mathcal{G}) = \mathcal{G}$

- $\mathsf{Var}(\mathcal{G}) = \sum_{x \in \mathcal{G}} \frac{1}{6}$

- by field property of additive identity element $0$

Example 8.19 (DQPSK) \[55\] In digital communications, there are several modulation techniques available. Most of these manipulate (“modulate”) the parameters a sinusoidal signal (called the carrier) at the transmitter to “carry” information (such as a person’s voice) to a receiver where under certain reasonable conditions the information can be recovered with an acceptably low error rate (e.g. $\leq 0.0001\%$). Parameters of the sinusoid that may be manipulated include the sinusoid’s amplitude (Amplitude Shift Keying or ASK), frequency (Frequency Shift Keying or FSK), or phase (Phase Shift Keying or PSK). The information to be carried is first encoded into a sequence of “symbols”, with each symbol carrying $N$ bits

\[55\] [113]
(typically \(N = 1, 2, \) or 3). All of modulation techniques generate a code space with \(2^N\) code points. In each modulation technique, the receiver must somehow have a reference by which it can recover the information from the carrier. A receiver may easily generate an amplitude reference for ASK modulation by using a simple low pass filter to find the 0 Hertz component of the received signal. A receiver may generate a frequency reference for FSK modulation by using an oscillator circuit that oscillates at the same frequency as the unmodulated carrier. For PSK, generating a reference cannot be done without assistance from the transmitter. The transmitter may provide this assistance by also transmitting a reference sinusoid (one that does not change its phase), or by encoding the reference signal into the information sequence itself.

One way to do the latter using \(N = 2\) bit encoding is a modulation technique called \textit{Differential Quadrature Phase Shift Keying} or DQPSK. In DQPSK, consecutive code points at the transmitter cannot change by more than 1 bit. In the illustration, this means that they cannot “jump” across the square to the opposite corner. That is, each symbol is partly a function of the previous symbol. By doing so, the phase information can be encoded into the sequence.

\[
\begin{align*}
\Psi(00,10) &= 1 \\
\Psi(01,10) &= 2 \\
\Psi(00,01) &= 1 \\
\Psi(01,11) &= 1 \\
\Psi(10,11) &= 1 \\
\Psi(00,11) &= 2
\end{align*}
\]

9 Random variables on outcome subspaces

9.1 Definitions

The traditional \textit{random variable} (Definition 5.4 page 14) is a mapping from a \textit{probability space} (Definition 5.3 page 14) to the \textit{real line} (Definition 6.5 page 17). This paper extends this definition to include functions with additional structure in the domain and expanded structure in the range (next definition).

\textbf{Definition 9.1} A function \(X \in H^G\) (Definition 1.3 page 3) is an \textbf{outcome random variable} if \(G\) is an \textit{outcome subspace} (Definition 8.1 page 22) and \(H\) is an \textit{ordered quasi-metric space} (Definition 6.1 page 16).

The definitions of \textit{outcome expected value} and \textit{outcome variance} (next definition) of an \textit{outcome random variable} are, in essence, identical to the \textit{outcome center} (Definition 9.2 page 42) and \textit{outcome variance} (Definition 9.2 page 42) of \textit{outcome subspaces} (Definition 8.1 page 22) that \textit{outcome random variables} map from and by induction, to.

\textbf{Definition 9.2} Let \(G\) be an \textit{outcome subspace} (Definition 8.1 page 22), \(H\) an \textit{ordered quasi-metric space} (Definition 6.1 page 16), and \(X\) be an \textit{outcome random variable} (Definition 9.1 page 42) in \(\in H^G\). Let \(H \triangleq (\Omega, d, \leq, P)\) be the \textit{outcome subspace} induced by \(H, G,\) and \(X\). Let \(\hat{E}_X\) be a function from \(\Omega\) to the power set \(\mathcal{P}\).

The \textbf{outcome expected value} \(\hat{E}(X)\) of \(X\) is \(\hat{E}(X) \triangleq \arg \min_{x \in \Omega} \max_{y \in \Omega} d(x, y) P(y)\).

The \textbf{outcome variance} \(\hat{\var}(X; E_x)\) of \(X\) is \(\hat{\var}(X) \triangleq \sum_{x \in \Omega} d^2(E_x(X), x) P(x)\).

Moreover, \(\hat{\var}(X) \triangleq \hat{\var}(X; \hat{E})\), where \(\hat{E}\) is the \textit{outcome expected value} function.

9.2 Properties

\textbf{Theorem 9.3} Let \(X \in H^G\) be a \textit{random variable} (Definition 5.4 page 14) on an \textit{ordered quasi-metric space} (Definition 6.1 page 16) \(H\). Let \(H \triangleq (\Omega, d, \leq, P)\) be the \textit{outcome subspace} (Definition 8.1 page 22) induced by \(H, G,\) and...
Let \( \text{Var}(X) \) be the traditional variance \( (\text{Definition 5.6 page 15}) \) of \( X \). Let \( \hat{\text{Var}}(X) \) be the outcome subspace variance of \( X \) \( (\text{Definition 9.2 page 42}) \).

\[
\{ \; H \triangleq (\Omega, \cdot, \leq) \; \text{is the real line} \;
\implies
\{ \; \text{Var}(X; E) = \text{Var}(X) \; \}
\]

**Proof:**

\[
\hat{\text{Var}}(X; E) = \sum_{x \in H} d^2(E(X), x) P(x)
\]

by definition of \( \text{Var} \) \( (\text{Definition 9.2 page 42}) \)

\[
= \sum_{x \in \mathbb{R}} |E(X) - x|^2 P(x)
\]

by definition of real line \( H \) \( (\text{Definition 6.5 page 17}) \)

\[
= \int_{\mathbb{R}} (x - E(X))^2 P(x) \, dx
\]

by definition of Lebesgue integration on \( \mathbb{R} \)

\[
= \text{Var}(X)
\]

by definition of \( \text{Var} \) \( (\text{Definition 5.6 page 15}) \)

**Remark 9.4** Despite the correspondence of traditional variance and outcome variance on the real line as demonstrated in Theorem 9.3 \( (\text{page 42}) \), the situation is different for expected values. Even when both are calculated on the same real line, the traditional expected value \( E(X) \) \( (\text{Definition 5.6 page 15}) \) and the outcome expected value \( \hat{E}(X) \) \( (\text{Definition 9.2 page 42}) \) don’t always yield the same value. Demonstrations of this include Example 9.14 \( (\text{page 51}) \) and Example 9.17 \( (\text{page 55}) \). However, there is one common situation in which the two statistics do correspond (next theorem).

**Theorem 9.5** Let \( X, H, G \) be defined as in Theorem 9.3 \( (\text{page 42}) \). Let \( E(X) \) be the traditional expected value \( (\text{Definition 5.6 page 15}) \) and \( \hat{E}(X) \) the outcome expected value of \( X \) \( (\text{Definition 9.2 page 42}) \).

\[
\{ \;
1. \; H \triangleq (\Omega, d, \leq) \; \text{(real line \text{Definition 6.5 page 17}) and} \\
2. \; P(a - x) = P(a + x) \; \forall x \in \mathbb{R} \; \text{(symmetric about a)} 
\}
\implies
\{ \; \hat{E}(X) = a = E(X) \; \}

**Proof:**

\[
\hat{E}(X) \triangleq \arg \min_{y \in \mathbb{R}} \max_{x \in \mathbb{R}} |x - y| P(y)
\]

by definition of \( \hat{E} \) \( (\text{Definition 9.2 page 42}) \)

\[
= \arg \min_{x \in \mathbb{R}} \max_{y \in \mathbb{R}} |x - y| P(y)
\]

by definition of real line \( (\text{Definition 6.5 page 17}) \)

\[
= a
\]

because \( h(x) \triangleq \max_{y \in \mathbb{R}} |x - y| P(y) \) is minimized when \( x = a \)

by Proposition 5.8 \( (\text{page 15}) \)

**Theorem 9.6**

Let \( G \triangleq (\Omega_G, d_G, \leq_G, P_G) \) be an outcome subspace \( (\text{Definition 8.1 page 22}) \).

Let \( H \triangleq (\Omega_H, d_H, \leq_H) \) be an ordered quasi-metric space \( (\text{Definition 6.1 page 16}) \).

Let \( K \triangleq (\Omega_K, d_K, \leq_K) \) be an ordered quasi-metric space \( (\text{Definition 6.1 page 16}) \).

Let \( X \in H^G \) be a random variable from \( G \) onto \( H \) \( (\text{Definition 9.1 page 42}) \).

Let \( f \in \Omega_K \) be a function from \( \Omega_H \) onto \( \Omega_K \) \( (\text{pullback}) \) \( (\text{Theorem 4.6 page 8}) \).

Let \( \phi \in \mathbb{R}^K \) be a function from \( \mathbb{R} \) into \( \mathbb{R} \) \( (\text{pushforward}) \) \( (\text{Definition 4.7 page 9}) \).

Let \( H \triangleq (\Omega_H, d_H, \leq_H, P_H) \) be an outcome subspace induced by \( G, H, \) and \( X \).

Let \( K \triangleq (\Omega_K, d_K, \leq_K, P_K) \) be an outcome subspace induced by \( K, H, \) and \( f \).
9.2 PROPERTIES

\[
\begin{aligned}
1. & \quad \mathbb{P} \text{ is INJECTIVE} \\
2. & \quad \phi \text{ is STRICTLY ISOTONE} \\
3. & \quad d_H(f(x), f(y)) \mathcal{P}(y) = \phi[d_h(x, y) \mathcal{P}(y)]
\end{aligned}
\]

\[\implies \{ \hat{E}[f(X)] = f(\hat{E}(X)) \}\]

\textbf{Proof:}

\[
\hat{E}[f(X)] = \arg\min_{x \in \mathcal{H}} \max_{y \in \mathcal{H}} d_k(x, y) \mathcal{P}(y)
\]

by definition of \(\hat{E}\) (Definition 9.2 page 42) and \(K\)

\[
= f \left[ \arg\min_{x \in \mathcal{H}} \max_{y \in \mathcal{H}} d_k(f(x), f(y)) \mathcal{P}(f(y)) \right]
\]

by \(f\) bijection hypothesis

\[
= f \left[ \arg\min_{x \in \mathcal{H}} \max_{y \in \mathcal{H}} d_h(f(x), f(y)) \mathcal{P}(y) \right]
\]

by \(f\) bijection hypothesis

\[
= f \left[ \arg\min_{x \in \mathcal{H}} \max_{y \in \mathcal{H}} \phi[d_h(x, y) \mathcal{P}(y)] \right]
\]

by \(d_h\) hypothesis

\[
= f \left[ \arg\min_{x \in \mathcal{H}} \max_{y \in \mathcal{H}} \phi[d_h(x, y) \mathcal{P}(y)] \right]
\]

by \(\phi\) is strictly isotone hypothesis and Lemma 7.9 page 21

\[
= f \hat{E}(X)
\]

by definition of \(\hat{E}\) (Definition 9.2 page 42) and \(X\)

\textbf{Corollary 9.7} Let \(H\) be an ORDERED METRIC SPACE (Definition 6.1 page 16) and \(X \in H^G\) a RANDOM VARIABLE (Definition 9.1 page 42) onto \(H\). Let \((\mathbb{R}, |\cdot|, \leq)\) be the REAL LINE ORDERED METRIC SPACE (Definition 6.5 page 17).

\[
H = (\mathbb{R}, |\cdot|, \leq) \implies \{ \hat{E}(aX) = a\hat{E}(X) \quad \forall a \in \mathbb{R}^+ \}
\]

\textbf{Proof:}

(1) Proof for \(a = 0\) case:

\[
\hat{E}(0 \cdot X) = \arg\min_{x \in \mathcal{H}} d(x, y) \mathcal{P}(y)
\]

by definition of \(\hat{E}\) (Definition 9.2 page 42)

\[
= \arg\min_{x \in [0]} d(x, y) \mathcal{P}(y)
\]

\[
= \arg\min_{x \in [0]} d(0, 0) \mathcal{P}(y)
\]

\[
= \arg\min_{x \in [0]} 0 \mathcal{P}(y)
\]

by nondegenerate property of \(d\) (Definition 4.2 page 8)

\[
= 0
\]

\[
= 0 \cdot \hat{E}(X)
\]

(2) Proof for \(a > 0\) case: \(d(f(x), f(y)) \mathcal{P}(y) \triangleq |ax - ay| \mathcal{P}(y) = |a||x - y| \mathcal{P}(y) \triangleq |a||d(x, y) \mathcal{P}(y)|\)

\[
\hat{E}(aX) = a\hat{E}(X) \quad \text{because } f(x) = ax \text{ is strictly isotone on the real line and by Theorem 9.6 (page 43)}
\]
9.3 Problem statement

The traditional random variable \( X \) (Definition 5.4 page 14) is a function that maps from a stochastic process to the real line (Definition 6.5 page 17). The traditional expectation value \( \mathbb{E}(X) \) of \( X \) is then often a poor choice of a statistic when the stochastic process that \( X \) maps from is a structure other than the real line or some substructure of the real line. There are two fundamental problems:

1. A traditional random variable \( X \) maps to the linearly ordered real line. However, \( X \) often maps from a random process that is non-linearly ordered (or even unordered Definition 2.4 page 4, Definition 2.6 page 5).

2. A traditional random variable \( X \) maps to the real line with a metric geometry (Remark 8.9 page 26) induced by the usual metric (Example 4.20 page 13). But many random processes have a fundamentally different metric geometry, a common one being that induced by the discrete metric (Example 4.21 page 14).

Thus, the order structure of the domain and range of \( X \) are often fundamentally dissimilar, leading to statistics, such as \( \mathbb{E}(X) \), that are of poor quality with regards to qualitative intuition and quantitative variance (expected error) measurements, and of dubious suitability for tasks such as decision making, prediction, and hypothesis testing.

Remark 9.8 Unlike in traditional statistical processing, it in general not true that \( \mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y) \). See Example 10.11 (page 74) for a counter example.

Remark 9.9 A possible solution to the traditional random variable order and metric geometry problem is to allow the random variable to map into the complex plane (Example 6.9 page 17) with the usual metric, rather than into the real line only. However, this is a poor solution, as demonstrated in Example 9.21 (page 61).

9.4 Examples

9.4.1 Fair die examples

Example 9.10 (fair die mappings to real line and integer line) Let \( G \) be the fair die outcome subspace (Example 8.7 page 23). Let \( X \in (\mathbb{R}, |\cdot|, \leq)^G \) be a random variable (Definition 5.4 page 14, Definition 9.1 page 42) mapping from \( G \) to the real line (Definition 6.5 page 17), and \( Y \in (\mathbb{Z}, |\cdot|, \leq)^G \) be a random variable (Definition 5.4 page 14, Definition 9.1 page 42) mapping from \( G \) to the integer line (Definition 6.6 page 17), as illustrated in Figure 6 (page 45). Let \( \mathbb{E} \) be the traditional expected value function (Definition 5.6 page 15), \( \text{Var} \) the traditional variance function.
(Definition 5.6 page 15), \( \hat{E} \) the outcome expected value function (Definition 9.2 page 42), and \( \text{Var} \) the outcome variance function (Definition 9.2 page 42). This yields the following statistics:

- **Geometry of** \( \mathcal{G} \):
  \[ \hat{\mathcal{C}}(\mathcal{G}) = \{ \square, \square, \lozenge, \bigcirc, \triangle, \triangleleft \} \]
- **Traditional statistics on real line**:
  \[
  \begin{align*}
  E(X) &= 3.5 \\
  \text{Var}(X; E) &= \frac{35}{12} \approx 2.917
  \end{align*}
  \]
- **Outcome subspace statistics on real line**:
  \[
  \begin{align*}
  \hat{E}(X) &= \{ 3.5 \} \\
  \text{Var}(X; \hat{E}) &= \frac{35}{12} \approx 2.917
  \end{align*}
  \]
- **Outcome subspace statistics on integer line**:
  \[
  \begin{align*}
  \hat{E}(Y) &= \{ 3, 4 \} \\
  \text{Var}(Y; \hat{E}) &= \frac{35}{12} \approx 1.667
  \end{align*}
  \]

**Proof:**

\[
\begin{align*}
\hat{\mathcal{C}}(\mathcal{G}) &= \{ \square, \square, \lozenge, \bigcirc, \triangle, \triangleleft \} \\
E(X) &\triangleq \sum_{x \in \mathbb{R}} x \cdot P(x) \\
&= \sum_{x \in \mathbb{R}} \frac{1}{6} x = \frac{1 + 2 + 3 + 4 + 5 + 6}{6} = \frac{21}{6} = \frac{7}{2} = 3.5
\end{align*}
\]
\[
\begin{align*}
\text{Var}(X; E) &= \text{Var}(X) \\
&= \sum_{x \in \mathbb{R}} (x - E(X))^2 \cdot P(x) \\
&= \sum_{x \in \mathbb{R}} \left( x - \frac{7}{2} \right)^2 \cdot \frac{1}{6} \\
&= \left[ \left( 1 - \frac{7}{2} \right)^2 + \left( 2 - \frac{7}{2} \right)^2 + \left( 3 - \frac{7}{2} \right)^2 + \left( 4 - \frac{7}{2} \right)^2 + \left( 5 - \frac{7}{2} \right)^2 + \left( 6 - \frac{7}{2} \right)^2 \right] \frac{1}{6} \\
&= \frac{25 + 9 + 1 + 9 + 25 + 25}{24} = \frac{70}{24} = \frac{35}{12} \approx 2.917
\end{align*}
\]
\[
\begin{align*}
\hat{E}(X) &= E(X) \\
&= \{ \frac{7}{2} \} = \{ 3.5 \}
\end{align*}
\]
\[
\begin{align*}
\text{Var}(X; \hat{E}) &= \sum_{x \in \mathbb{R}} d^2(\hat{E}(X), x) \cdot P(x) \\
&= \sum_{x \in \mathbb{R}} d^2(E(X), x) \cdot P(x) \\
&= \text{Var}(X) \\
&= \frac{35}{12} \approx 2.917
\end{align*}
\]
\[
\begin{align*}
\hat{E}(Y) &= \text{arg min} \max_{x \in \mathbb{Z}, y \in \mathbb{H}} d(x, y) \cdot P(y) \\
&= \text{arg min} \max_{x \in \mathbb{Z}, y \in \mathbb{H}} |x - y| \cdot \frac{1}{6} \\
&= \text{arg min} \max_{x \in \mathbb{Z}, y \in \mathbb{H}} |x - y| \\
&= \text{arg min} \max_{x \in \mathbb{Z}, y \in \mathbb{H}} \begin{bmatrix} 0 & 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 5 & 6 & 7 & 8 & 9 \\ 5 & 6 & 7 & 8 & 9 & 10 \end{bmatrix} = \text{arg min} \begin{bmatrix} 5 \\ 4 \\ 3 \\ 2 \\ 1 \\ 0 \end{bmatrix}
\end{align*}
\]
\[
\begin{align*}
\text{Var}(Y; \hat{E}) &= \sum_{x \in \mathbb{Z}} d^2(\hat{E}(Y), x) \cdot P(x) \\
&= \text{Var}(Y)
\end{align*}
\]
The random variable mappings in Example 9.11 (page 45) have two fundamental problems:

1. The order structure of the fair die and the order structure of real line are inherently dissimilar in that while the bijective (Definition 1.8 page 3) mapping X is trivially order preserving (Definition 7.1 page 17), its inverse is not order preserving. And this is a problem. In the linearly ordered (Definition 2.6 page 5) range of X, it is true that X(⚀) = 1 < 2 = X(⚁). But in the unordered domain of X {⚀, ⚁, ⚂, ⚃, ⚄, ⚅}, it is not true that ⚀ < ⚁; rather ⚀ and ⚁ are simply symbols without order. This causes problems when we attempt to use the random variable to make statistical inferences involving moments (Definition 8.2 page 22). The traditional expected value (Definition 5.6 page 15) of a fair die (Example 8.7 page 23) is E(X) = \frac{1}{6}(1 + 2 + \ldots + 6) = 3.5. This implies that we expect the outcome of ⚀ or ⚁ more than we expect the outcome of say ⚁ or ⚂. But these results have no relationship with reality or with intuition because the values of a fair die are merely symbols. For a fair die, we would expect any pair of values equally. We would not expect the outcome [⚀ or ⚁] more than we would expect the outcome [⚁ or ⚂], or more than we would expect any other outcome pair.

2. The metric geometry (Remark 8.9 page 26) of the fair die outcome subspace is very dissimilar to the metric geometry of the real line (Definition 6.5 page 17) that it is mapped to by the random variable X. And this is a problem. In the metric geometry of the fair die induced by the discrete metric (Example 4.21 page 14), ⚀ is no closer to ⚁ than it is to ⚂ (d(⚀, ⚁) = 1 = d(⚀, ⚂)). However in the metric geometry of the real line induced by the usual metric d(x, y) \triangleq |x − y| (Example 4.20 page 13), X(⚀) = 1 is closer to X(⚁) = 2 than it is to X(⚂) = 3 (|1 − 2| = 1 ≠ 2 = |1 − 3|).

Example 9.11 (fair die mapping to isomorphic structure) Let G \triangleq ( {⚀, ⚁, ⚂, ⚃, ⚄, ⚅}, d, \emptyset, P) be a fair die outcome subspace (Example 8.7 page 23), and H \triangleq ( {1, 2, 3, 4, 5, 6}, d, \emptyset) be an unordered metric space (Definition 6.1 page 16). Example 9.10 (page 45) presented mappings from G to structures with structures dissimilar to G. Figure 7 page 47 (A) illustrates a mapping to the isomorphic structure H \triangleq ( {1, 2, 3, 4, 5, 6}, d, \emptyset, X(P)), yielding the following statistics:

\[ \bar{E}(X) = \{1, 2, 3, 4, 5, 6\} \quad \bar{Var}(X) = 0 \]

Here, \bar{E}(X) equals the entire base set of H, indicating a statistic carrying no information about an expected outcome. That is, there is no best guess concerning outcome. This is much different than the traditional probability of 3.5 (Example 9.10 page 45) which deceptively suggests a likely outcome of ⚁ or ⚂. And one could easily argue that no information is much better than misleading information.
9.4 EXAMPLES

PROOF:

\[ \hat{E}(X) \triangleq \arg \min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} d(x, y) \mathcal{P}(y) \]

by definition of \( \hat{E} \) (Definition 9.2 page 42)

\[ = \arg \min_{x \in \mathcal{G}} \max_{y \in \mathcal{Y}} d(x, y) \mathcal{P}(y) \]

because \( \mathcal{G} \) and \( \mathcal{H} \) are isomorphic

\[ = X[\hat{C}(\mathcal{G})] \]

by definition of \( \hat{C} \) (Definition 8.3 page 22)

\[ = X[\{ \Box, \lozenge, \lozenge, \Diamond, \lozenge, \lozenge \}] \]

by Example 9.10 (page 45)

\[ = \{1, 2, 3, 4, 5, 6\} \]

by definition of \( X \)

\[ \text{Var}(X) \triangleq \sum_{x \in \mathcal{X}} d^2(\hat{E}(X), x) \mathcal{P}(x) \]

by definition of \( \text{Var} \) (Definition 9.2 page 42)

\[ = \sum_{x \in \mathcal{G}} d^2(X(\hat{C}(\mathcal{G}), x) \mathcal{P}(x) \]

because \( \mathcal{G} \) and \( \mathcal{H} \) are isomorphic

\[ \triangleq \text{Var}(\mathcal{G}) \]

by definition of \( \text{Var} \) (Definition 8.5 page 23)

\[ = 0 \]

by Example 9.10 (page 45)

Although all the coefficients of the polynomial equation \( x^2 - 2x + 2 = 0 \) are in the set of real numbers \( \mathbb{R} \), the solutions of the equation (\( x = 1 + i \) and \( x = 1 - i \)) are not. Rather, the two solutions are in the complex plane \( \mathbb{R}^2 \) (Example 6.9 page 17), of which \( \mathbb{R} \) is a substructure. This is an example of extending a structure (from \( \mathbb{R} \) to \( \mathbb{R}^2 \)) to achieve more useful results. The same idea can be applied to a random variable \( X \in \mathcal{H}^G \). The definition of an outcome random variable (Definition 9.1 page 42) does not require a bijection between \( \mathcal{G} \) and \( \mathcal{H} \); rather, it only requires that the mapping be “into” the base set of \( \mathcal{H} \) (Definition 1.8 page 3). In Example 9.11 (page 47) in which \( \mathcal{G} \) is isomorphic to \( \mathcal{H} \), the expected value of \( X \) is a set with six values. However, we could extend \( \mathcal{H} \), while still preserving the order and metric geometry of \( \mathcal{G} \), to produce a random variable with a simpler expected value (next example).

Example 9.12 (fair die mapping with extended range) Let \( \mathcal{G} \triangleq (\{ \Box, \lozenge, \lozenge, \Diamond, \lozenge, \lozenge \}, d, \emptyset, \mathcal{P}) \) be a fair die outcome subspace (Example 8.7 page 23), and \( \mathcal{H} \triangleq (\{1, 2, 3, 4, 5, 6, 0\}, p, \emptyset) \) be an unordered metric space (Definition 6.1 page 16). Figure 7 page 47 (B) illustrates a random variable mapping \( \mathcal{X} \) from \( \mathcal{G} \) to the extended structure \( \mathcal{H} \), yielding the following statistics:

\[ \hat{E}(X) = \{0\} \]

\[ \text{Var}(X) = \frac{1}{4} \]

As in Example 9.11, order and metric geometry are still preserved. Here, an expected value of \( \{0\} \) simply means that no real physical value is expected more or less than any other real physical value. Note also that the variance (expected error) is more than 11 times smaller than that of the corresponding statistical estimates on the real line (\( \chi^2 \) versus 3\( \chi^2 \) Example 9.10 page 45).

PROOF:

\[ \hat{E}(X) \triangleq \arg \min_{x \in \mathcal{H}} \max_{y \in \mathcal{Y}} d(x, y) \mathcal{P}(y) \]

by definition of \( \hat{E} \) (Definition 9.2 page 42)

\[ = \arg \min_{x \in \mathcal{H}} \max_{y \in \mathcal{N}[0]} d(x, y) \mathcal{P}(y) \]

because \( \mathcal{P}(0) = 0 \)

\[ = \arg \min_{x \in \mathcal{H}} \max_{y \in \mathcal{N}[0]} d(x, y) \frac{1}{6} \]

\[ = \arg \min_{x \in \mathcal{H}} \max_{y \in \mathcal{N}[0]} d(x, y) \]

because \( f(x) = \frac{1}{6}x \) is strictly isotone and by Lemma 7.9 (page 21)
\[
\begin{align*}
\bar{\var}(X) & \triangleq \sum_{x \in \mathcal{H}} d^2(\bar{E}(X), x) P(x) \quad \text{by definition of } \bar{\var} \ (\text{Definition 9.2 page 42}) \\
& = \sum_{x \in \mathcal{H}} d^2(0, x) P(x) \quad \text{by } \bar{E}(X) \text{ result} \\
& = \sum_{x \in \mathcal{H}/\{0\}} d^2(0, x) \frac{1}{6} \quad \text{by definition of } \mathcal{G} \\
& = 6 \left(\frac{1}{2}\right)^2 \frac{1}{6} = \frac{1}{4} \quad \text{by definition of } \mathcal{H}
\end{align*}
\]

9.4.2 Real die examples

![Diagram of random variable mappings]

Figure 8: random variable mappings from the real die outcome subspace to several ordered metric spaces (Example 9.13 page 49)

**Example 9.13** (real die mappings) Let \( G \) be the real die outcome subspace (Example 8.8 page 24). Let \( W, X, Y \) and \( Z \) be random variable (Definition 9.1 page 42) mappings as illustrated in Figure 8 (page 49). Let \( E, \bar{\var}, \bar{E}, \) and \( \bar{\var} \) be defined as in Example 9.10 (page 45). This yields the following statistics:
geometry of $\mathcal{G}$: $\hat{\mathcal{G}} = \{\Box, \triangle, \square, \lozenge, \diamondsuit, \bigcirc\}$

traditional statistics on real line: $E(W) = \frac{7}{2} = 3.5$ \quad $\text{Var}(W; E) = \frac{35}{12} \approx 2.917$

outcome subspace statistics on real line: $E(W) = \{3.5\}$ \quad $\text{Var}(W; E) = \frac{35}{12} \approx 2.917$

outcome subspace statistics on integer line: $E(W) = \{3, 4\}$ \quad $\text{Var}(W; E) = \frac{35}{12} \approx 1.667$

outcome subspace statistics on isomorphic structure: $E(Y) = \{1, 2, 3, 4, 5, 6\}$ \quad $\text{Var}(Y; E) = 0$

outcome subspace statistics on extended structure: $E(Z) = \{0\}$ \quad $\text{Var}(Z; E) = 1$

Similar to Example 9.11 (page 47), the statistic $E(Z) = \{0\}$ indicates a statistic carrying no information about an expected outcome. Again, one could easily argue that no information is much better than misleading information.

Proof:

\[
\hat{\mathcal{G}} = \{\Box, \triangle, \square, \lozenge, \diamondsuit, \bigcirc\} \quad \text{by Example 8.8 (page 24)}
\]
\[
E(W) = \frac{7}{2} = 3.5 \quad \text{by } E(X) \text{ result of Example 9.10 page 45}
\]
\[
\text{Var}(W; E) = \frac{35}{12} \approx 2.917 \quad \text{by } \text{Var}(X) \text{ result of Example 9.10 page 45}
\]
\[
E(W) = \{3.5\} \quad \text{by } E(X) \text{ result of Example 9.10 page 45}
\]
\[
\text{Var}(W; E) = \frac{35}{12} \approx 2.917 \quad \text{by } \text{Var}(X; E) \text{ result of Example 9.10 page 45}
\]
\[
E(X) = \{3, 4\} \quad \text{by } E(Y) \text{ result of Example 9.10 page 45}
\]
\[
\text{Var}(X; E) = \frac{5}{3} \approx 1.667 \quad \text{by } \text{Var}(Y; E) \text{ result of Example 9.10 page 45}
\]
\[
E(Y) = Y(\{\Box, \triangle, \square, \lozenge, \diamondsuit, \bigcirc\}) = \{1, 2, 3, 4, 5, 6\} \quad \text{by } \hat{\mathcal{G}} \text{ result of Example 8.8 page 24}
\]
\[
\text{Var}(Y; E) \triangleq \sum_{x \in \mathcal{H}} d^2(E(Y), x) P(x) \quad \text{by definition of } \text{Var} \text{ (Definition 9.2 page 42)}
\]
\[
= \sum_{x \in \mathcal{H}} d^2(H, x) P(x) \quad \text{by } E(Y) \text{ result}
\]
\[
= \sum_{x \in \mathcal{H}} 0^2 x P(x) \quad \text{by nondegenerate property of quasi-metrics (Definition 4.1 page 8)}
\]
\[
= 0
\]
\[
E(Z) \triangleq \arg \min_{x \in \mathcal{K}} \max_{y \in \mathcal{K}} d(x, y) P(y) \quad \text{by definition of } E \text{ (Definition 9.2 page 42)}
\]
\[
= \arg \min_{x \in \mathcal{K}} \max_{y \in \mathcal{K} \setminus \{0\}} d(x, y) P(y) \quad \text{because } P(0) = 0
\]
\[
= \arg \min_{x \in \mathcal{K}} \max_{y \in \mathcal{K} \setminus \{0\}} d(x, y) \frac{1}{6} \quad \text{by definition of } \mathcal{G} \text{ and } Z
\]
\[
= \arg \min_{x \in \mathcal{K}} \max_{y \in \mathcal{K} \setminus \{0\}} d(x, y) \quad \text{because } f(x) = \frac{1}{6} x \text{ is strictly isotone and by Lemma 7.9 (page 21)}
\]
\[
= \arg \min_{x \in \mathcal{K}} \max_{y \in \mathcal{K}} \begin{bmatrix}
    d(1, 1) & d(1, 2) & \ldots & d(1, 6) \\
    d(2, 1) & d(2, 2) & \ldots & d(2, 6) \\
    d(3, 1) & d(3, 2) & \ldots & d(3, 6) \\
    d(4, 1) & d(4, 2) & \ldots & d(4, 6) \\
    d(5, 1) & d(5, 2) & \ldots & d(5, 6) \\
    d(6, 1) & d(6, 2) & \ldots & d(6, 6)
\end{bmatrix} = \arg \min_{x \in \mathcal{K}} \begin{bmatrix}
    0 & 1 & 1 & 1 & 1 & 2 \\
    1 & 0 & 1 & 1 & 2 & 1 \\
    1 & 1 & 0 & 1 & 1 & 2 \\
    2 & 1 & 1 & 1 & 0 & 1 \\
    1 & 1 & 1 & 1 & 1 & 1
\end{bmatrix} = \arg \min_{x \in \mathcal{K}} \begin{bmatrix}
    2 \\
    2 \\
    2 \\
    2 \\
    1 \\
    0
\end{bmatrix} = \{0\}
\]
\[
\text{Var}(Z) \triangleq \sum_{x \in \mathcal{K}} d^2(E(Z), x) P(x) \quad \text{by definition of } \text{Var} \text{ (Definition 9.2 page 42)}
\]
Example 9.14 (weighted die mappings) Let $G$ be weighted die outcome subspace (Example 8.10 page 27), and $X$ and $Y$ be random variables, as illustrated in Figure 9 (page 51). Let $E$, $\mathcal{V}$, $\mathcal{E}$, and $\mathcal{V}_E$ be defined as in Example 9.10 (page 45). This yields the following statistics:

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geometry of $G$:</td>
<td>$\hat{\mathcal{C}}(G) = { \heartsuit }$</td>
</tr>
<tr>
<td>Traditional statistics on real line:</td>
<td>$E(X) = 3$</td>
</tr>
<tr>
<td>Outcome subspace statistics on real line:</td>
<td>$\hat{E}(X) = { \frac{25}{7} } \approx {3.57}$</td>
</tr>
<tr>
<td>Outcome subspace statistics on isomorphic structure</td>
<td>$\hat{E}(Y) = {4}$</td>
</tr>
</tbody>
</table>

The statistic $E(X) = 3$ evaluated on the real line is arguably very poor because it suggests that we “expect” the event $\heartsuit$ rather than $\diamondsuit$, even though $P(\heartsuit)$ is very large, $P(\diamondsuit)$ is very small, and the physical distance $d(\heartsuit, \diamondsuit) = 2$ on the die from $\heartsuit$ to $\diamondsuit$ is twice as much as it is to any of the other four die faces. If we retain use of the real line but replace the traditional expected value $E(X)$ with the outcome expected value $\hat{E}(X)$, a small but significant improvement is made ($\mathcal{V}_E(X; \hat{E}) \approx 1.143 < 1.43 = \mathcal{V}_E(X; E)$). Arguably a better choice still is to abandon the real line altogether in favor of the isomorphic structure $K$ and the statistic $\hat{E}(Y) = \{4\}$ evaluated on $K$, yielding not only an intuitively better result but also a variance $\mathcal{V}_E(Y; \hat{E})$ that is more than 4 times smaller than that of $E(X)$ ($\mathcal{V}_E(Y; \hat{E}) \approx 0.337 < 1.43 = \mathcal{V}_E(X; E)$).

\[ \hat{E}(X) = \sum_{x \in \mathbb{Z}} xP(x) \]
\[ = 1 \times \frac{1}{10} + 2 \times \frac{1}{20} + 3 \times \frac{1}{30} + 4 \times \frac{3}{5} + 5 \times \frac{1}{50} + 6 \times \frac{1}{30} \]
\[ = \frac{1}{300} (1 \times 30 + 2 \times 15 + 3 \times 10 + 4 \times 180 + 5 \times 6 + 6 \times 10) = \frac{900}{300} = 3 \]

$\mathcal{V}_E(X; \hat{E}) \approx \mathcal{V}_E(Y; \hat{E})$
\[
= \sum_{x \in \mathbb{Z}} (x - E(X))^2 \mathbb{P}(x)
= \frac{1}{10} (1 - 3)^2 + \frac{1}{20} (2 - 3)^2 + \frac{1}{30} (3 - 3)^2 + \frac{3}{50} (4 - 3)^2 + \frac{1}{30} (5 - 3)^2 + \frac{1}{30} (6 - 3)^2
= \frac{1}{300} (120 + 15 + 0 + 180 + 24 + 90) = \frac{429}{300} = 1.43
\]

\[\hat{E}(X) \triangleq \arg \min_{x \in \mathbb{R}} \max_{y \in \mathbb{R}} |x - y| \mathbb{P}(y)\]
by definition of \(\hat{E}\) (Definition 9.2 page 42)

\[\hat{E}(X) = \left\{ \frac{25}{7} \right\} \approx \{3.5714\}\]

\[\var(X; \hat{E}) \triangleq \sum_{x \in \mathbb{R}} d^2(\hat{E}(X), x) \mathbb{P}(x)\]
by \(\hat{E}(X)\) result

\[\var(X; \hat{E}) = \left( \frac{25}{7} - 1 \right)^2 \frac{1}{10} + \left( \frac{25}{7} - 2 \right)^2 \frac{1}{20} + \left( \frac{25}{7} - 3 \right)^2 \frac{1}{30} + \left( \frac{25}{7} - 4 \right)^2 \frac{3}{5} + \left( \frac{25}{7} - 5 \right)^2 \frac{1}{50} + \left( \frac{25}{7} - 6 \right)^2 \frac{1}{30}\]

\[= \frac{16805}{49 \times 300} = \frac{3361}{2940} \approx 1.143\]

\[\mathbb{E}(Y) = \mathbb{Y}(\hat{\mathbb{U}}(\mathbb{G}))\]
\[\mathbb{E}(Y) = \mathbb{Y}([\mathbb{E}])\]
\[\mathbb{E}(Y) = \{4\}\]

\[\var(Y; \hat{E}) = \var(\mathbb{G}) = \frac{101}{300} \approx 0.337\]

\[\because \quad \because \quad \because \]

9.4.3 Spinner examples

Example 9.15 (spinner mappings) A six value board game spinner has a cyclic structure as illustrated in Figure 10 (page 53). Again, the order and metric geometry of the real line mapped to by the random variable \(\mathbb{X}\) is very dissimilar to that of the outcome subspace that it is supposed to represent. Therefore, statistical inferences based on \(\mathbb{X}\) will likely result in values that are arguably unacceptable. Both random variables \(\mathbb{Y}\) and \(\mathbb{Z}\) map to structures in which order and metric geometry are preserved. The mappings yield the following statistics:

geometry of \(\mathbb{G}\):
\[\hat{\mathbb{U}}(\mathbb{G}) = \{1, 2, 3, 4, 5, 6\}\]

traditional statistics on real line:
\[\mathbb{E}(\mathbb{W}) = 3.5 \quad \var(\mathbb{W}; \hat{E}) = \frac{35}{14} \approx 2.917\]

outcome subspace statistics on real line:
\[\hat{\mathbb{E}}(\mathbb{W}) = \{3.5\} \quad \var(\mathbb{W}; \hat{\mathbb{E}}) = \frac{35}{14} \approx 2.917\]

outcome subspace statistics on integer line:
\[\hat{\mathbb{E}}(\mathbb{X}) = \{3, 4\} \quad \var(\mathbb{X}; \hat{\mathbb{E}}) = \frac{35}{14} \approx 2.25\]

outcome subspace statistics on isomorphic structure:
\[\hat{\mathbb{E}}(\mathbb{Y}) = \{1, 2, 3, 4, 5, 6\} \quad \var(\mathbb{Y}; \hat{\mathbb{E}}) = 0\]

outcome subspace statistics on extended structure:
\[\hat{\mathbb{E}}(\mathbb{Z}) = \{0\} \quad \var(\mathbb{Z}; \hat{\mathbb{E}}) = \frac{9}{4} = 2.25\]
Proof:

\[ \hat{U}(G) = \{ \text{①, ②, ③, ④, ⑤, ⑥} \} \]  

by Example 8.11 (page 29)

\[ E(W) = \sum_{x \in \mathbb{R}} xP(x) \]  

by definition of \( E \) (Definition 5.6 page 15)

\[ = \frac{7}{2} = 3.5 \]  

by fair die example Example 9.10 (page 45)

\[ \tilde{\text{Var}}(W; E) = \text{Var}(X) \]  

by Theorem 9.3 page 42

\[ \Delta \sum_{x \in \mathbb{R}} [x - E(X)]^2P(x) \, dx \]  

by definition of \( \text{Var} \) (Definition 5.6 page 15)

\[ = \frac{35}{12} \approx 2.917 \]  

by fair die example Example 9.10 (page 45)

\[ E(W) = E(W) \]  

because on real line, \( P \) is symmetric, and by Theorem 9.5 page 43

\[ = \{3.5\} \]  

by \( E(W) \) result

\[ \tilde{\text{Var}}(W; E) = \tilde{\text{Var}}(W; E) \]  

because \( \tilde{E}(W) = E(W) \)

\[ = \frac{35}{12} \approx 2.917 \]  

by \( \tilde{\text{Var}}(W; E) \) result

\[ E(X) = \arg \min_{x \in \mathbb{Z}} \max_{y \in \mathbb{Z}} d(x, y) P(y) \]  

by definition of \( E \) (Definition 9.2 page 42)

\[ \Delta \arg \min_{x \in \mathbb{Z}} \max_{y \in \mathbb{Z}} |x - y|^\frac{1}{6} \]  

by definition of integer line (Definition 6.6 page 17) and \( G \)

\[ = \{3, 4\} \]  

by fair die example Example 9.10 (page 45)

\[ \tilde{\text{Var}}(X; E) = \frac{5}{3} \approx 1.667 \]  

by fair die example Example 9.10 (page 45)

\[ \tilde{E}(Y) = Y[\hat{U}(G)] \]  

because \( G \) and \( H \) are isomorphic under mapping \( Y \)

\[ = Y[\{\text{①, ②, ③, ④, ⑤, ⑥}\}] \]  

by \( \hat{U}(G) \) result

\[ = \{1, 2, 3, 4, 5, 6\} \]  

by definition of \( Y \)

\[ \tilde{\text{Var}}(Y; \tilde{E}) = \text{Var}(G) = 0 \]  

by spinner outcome subspace example (Example 8.11 page 29)
\[ \bar{E}(Z) \triangleq \arg \min_{x \in \mathcal{H}} \max_{y \in \mathcal{H}} d(x, y) \] 

by definition of \( \bar{E} \) (Definition 9.2 page 42)

= \arg \min_{x \in \mathcal{H}} \max_{y \in \mathcal{H}[0]} d(x, y) \] 

because \( P(0) = 0 \)

= \arg \min_{x \in \mathcal{H}} \max_{y \in \mathcal{H}[0]} \frac{1}{6} \] 

by definition of \( G \)

= \arg \min_{x \in \mathcal{H}} \max_{y \in \mathcal{H}[0]} d(x, y) \] 

because \( f(x) = \frac{1}{6} x \) is strictly isotone and by Lemma 7.9 (page 21)

\[ \bar{\var}(Z) \triangleq \sum_{x \in \mathcal{H}} d^2(\hat{\mathcal{U}}(G), x) P(x) \] 

by definition of \( \bar{\var} \) (Definition 9.2 page 42)

= \sum_{x \in \mathcal{H}} d^2(\emptyset, x) P(x) \] 

by \( \hat{E}(X) \) result

= \sum_{x \in \mathcal{H}[0]} \left( \frac{3}{2} \right)^2 \frac{1}{6} = |\mathcal{H}\setminus\{0\}| \left( \frac{3}{2} \right)^2 \frac{1}{6} = 6 \left( \frac{3}{2} \right)^2 \frac{1}{6} = \frac{9}{4} \]

\[ \text{Figure 11: weighted spinner mappings (Example 9.16 page 54)} \]

**Example 9.16** (weighted spinner mappings) Let \( G \) be weighted spinner outcome subspace (Example 8.12 page 31) with random variable mappings as illustrated in Figure 11 (page 54). This yields the following statistics:

- geometry of \( G \): \( \hat{\mathcal{U}}(G) = \{1, 6\} \)
- traditional statistics on real line \( (\mathbb{R}, |\cdot|, \leq) \): \( \hat{E}(X) = 3.5 \) \( \hat{\var}(W; \hat{E}) = \frac{17}{4} \approx 4.25 \)
- outcome subspace statistics on real line \( (\mathbb{R}, |\cdot|, \leq) \): \( \hat{E}(X) = \{3.5\} \) \( \hat{\var}(W; \hat{E}) = \frac{17}{4} \approx 4.25 \)
- outcome subspace statistics on isomorphic structure \( \mathcal{H} \): \( \hat{E}(Y) = \{1, 6\} \) \( \hat{\var}(Y; \hat{E}) = \frac{3}{2} \approx 1.67 \)
- outcome subspace statistics on continuous structure \( \mathcal{K} \): \( \hat{E}(Z) = \{0.5\} \) \( \hat{\var}(Z; \hat{E}) = \frac{37}{20} = 1.85 \)

Note that based on the variance values, the statistic \( \hat{E}(Z) \) on the continuous ring \( \mathcal{K} \) is arguably a much better statistic than \( \hat{E}(X) \) on the (continuous) real line \( (\mathbb{R}, |\cdot|, \leq) \).

\[ \hat{\mathcal{U}}(G) = \{1, 6\} \] 

by weighted spinner outcome subspace example (Example 8.12 page 31)
\[ E(X) = \sum_{x \in \mathbb{Z}} x \cdot P(x) = 1 \times \frac{3}{10} + 2 \times \frac{1}{10} + 3 \times \frac{1}{10} + 4 \times \frac{1}{10} + 5 \times \frac{1}{10} + 6 \times \frac{1}{10} = \frac{35}{10} = \frac{7}{2} = 3.5 \]

\[ \text{Var}(X; E) = \text{Var}(X) \quad \text{by Theorem } 9.3 \text{ page 42} \]

\[ = \sum_{x \in \mathbb{Z}} (x - E(X))^2 P(x) \quad \text{by definition of Var (Definition 5.6 page 15)} \]

\[ = \left(1 - \frac{7}{2}\right)^2 \frac{3}{10} + \left(2 - \frac{7}{2}\right)^2 \frac{1}{10} + \left(3 - \frac{7}{2}\right)^2 \frac{1}{10} + \left(4 - \frac{7}{2}\right)^2 \frac{1}{10} + \left(5 - \frac{7}{2}\right)^2 \frac{1}{10} + \left(6 - \frac{7}{2}\right)^2 \frac{3}{10} \]

\[ = \frac{1}{10} \left[ \left(\frac{-5}{2}\right)^2 \times 3 + \left(\frac{-3}{2}\right)^2 + \left(-\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + \left(\frac{3}{2}\right)^2 \times 3 \right] \]

\[ = \frac{1}{10} \left[ 75 + 9 + 1 + 9 + 75 \right] = \frac{170}{40} = \frac{17}{4} = 4.25 \]

\[ E(Y) = Y[\hat{U}(\mathcal{G})] \quad \text{because } \mathcal{G} \text{ and } H \text{ are isomorphic under } Y \]

\[ = Y[\{1, 6\}] \quad \text{by } \hat{U}(\mathcal{G}) \text{ result} \]

\[ \text{Var}(Y; \hat{E}) = \text{Var}(\mathcal{G}) \quad \text{because } \mathcal{G} \text{ and } H \text{ are isomorphic under } Y \]

\[ = \frac{5}{3} \approx 1.667 \quad \text{by weighted spinner outcome subspace example (Example 8.12 page 31)} \]

\[ \hat{E}(Z) = \arg \min_{x \in \mathcal{K}} \max_{y \in \mathcal{K}} d(x, y) P(y) \quad \text{by definition of } \hat{E} \text{ (Definition 9.2 page 42)} \]

\[ = 0.5 \]

\[ \text{Var}(Z; \hat{E}) = \sum_{x \in \mathcal{K}} d^2(\hat{E}(Z), x) P(x) \quad \text{by definition of } \text{Var} \text{ (Definition 9.2 page 42)} \]

\[ = \sum_{x \in \mathcal{K}} d^2\left(\frac{1}{2}, x\right) P(x) \quad \text{by } \hat{E}(Z) \text{ result} \]

\[ = \left(\frac{1}{2}\right)^2 \frac{3}{10} + \left(\frac{3}{2}\right)^2 \frac{1}{10} + \left(\frac{5}{2}\right)^2 \frac{1}{10} + \left(\frac{5}{2}\right)^2 \frac{1}{10} + \left(\frac{3}{2}\right)^2 \frac{1}{10} + \left(\frac{1}{2}\right)^2 \frac{3}{10} \]

\[ = \frac{1}{40}(3 + 9 + 25 + 25 + 9 + 3) = \frac{74}{40} = \frac{37}{20} = 1.85 \]

9.4.4 Pseudo-random number generator (PRNG) examples

Example 9.17 (LCG mappings, standard ordering)
The equation \( x_{n+1} = (7x_n + 5) \mod 9 \) with \( x_0 = 1 \) is a linear congruential (LCG) pseudo-random number
generator (PRNG) that has full period\(^{56}\) of 9 values. These 9 values can be mapped, using a surjective (Definition 1.8 page 3) function \(s \in G^0\) to the 5 element set \(\{0, 1, 2, 3, 4\}\) to “shape” the distribution from a uniform distribution to non-uniform:\(^{57}\)

<table>
<thead>
<tr>
<th>(n)</th>
<th>0</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>
<th>11</th>
<th>…</th>
</tr>
</thead>
<tbody>
<tr>
<td>(y_n \triangleq s(x_n))</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>…</td>
</tr>
</tbody>
</table>

Let \(G\) be the outcome subspace and \(X, Y, \) and \(Z\) be the outcome random variables illustrated in Figure 12 (page 56). This yields the following statistics:

- **Geometry of \(G_9\):**
  \[\hat{\mathcal{C}}(G) = \{0, 1, 2, 3, 4, 5, 6, 7, 8\}\]

- **Geometry of \(G\):**
  \[\hat{\mathcal{C}}(G) = \{4\}\]

- **Traditional statistics on real line:**
  \[E(X) = \frac{7}{3} \approx 2.333 \quad \text{Var}(X; E) = \frac{22}{3} \approx 2.444\]

- **Outcome subspace statistics on real line:**
  \[\hat{E}(X) = \{\frac{12}{5} = 2.4\} \quad \text{Var}(Y; \hat{E}) = \frac{5551}{1000} \approx 2.449\]

- **Outcome subspace statistics on integer line:**
  \[\hat{E}(Y) = \{2, 3\} \quad \text{Var}(Y; \hat{E}) = \frac{16}{9} \approx 1.778\]

- **Outcome subspace statistics on isomorphic structure:**
  \[\hat{E}(Z) = \{4\} \quad \text{Var}(Z; \hat{E}) = \frac{4}{4} = 1\]

Note that unlike the statistics \(E(X)\) and \(\hat{E}(X)\) on the real line, the statistic \(\hat{E}(Z)\) on the isomorphic structure \(K\) yields the maximally likely result, and a much smaller variance as well.

\(\triangleright\) **Proof:**

\[
\hat{C}(G_9) \triangleq \arg \min_{x \in G_9} \max_{y \in G_9} d(x, y) P(y)
\]

by definition of \(\hat{C}\) (Definition 8.3 page 22)

\[
= \arg \min_{x \in G_9} \max_{y \in G_9} d(x, y) \frac{1}{9}
\]

by definition of \(G_9\)

\[
= \arg \min_{x \in G_9} \max_{y \in G_9} d(x, y)
\]

because \(\phi(x) = \frac{1}{9} x\) is strictly isotone and by Lemma 7.9 page 21

\(^{56}\) [59], \(\triangleright\) [62], page 137, (Theorem 6.1), \(\triangleright\) [102], page 86, (Hull-Dobell Theorem)

\(^{57}\) The sequence \(\{1, 3, 0, 2, 4, 1, 3, 0, 2, 4, 1, 3, \ldots\}\) is generated by the equation \(y_{n+1} = (y_n + 2) \mod 5\) with \(y_0 = 1\)
= \arg \min_{x \in \mathcal{G}} \{4, 4, 4, 4, 4, 4, 4, 4, 4\} \quad \text{because the maximum distance in } \mathcal{G}_y \text{ from any } x \text{ is 4}

\hat{\mathcal{C}}(\mathcal{G}) = \{4\} \quad \text{by weighted ring outcome subspace example (Example 8.13 page 33)}

E(X) = \sum_{x \in \mathbb{R}} x \mathbb{P}(x)

= 0 \times \frac{2}{9} + 1 \times \frac{1}{9} + 2 \times \frac{1}{9} + 3 \times \frac{2}{9} + 4 \times \frac{3}{9} = \frac{21}{9} = \frac{7}{3} \approx 2.333

\mathbb{V} \mathbb{a}r(X; E) = \mathbb{V} \mathbb{a}r(X)

= \sum_{x \in \mathbb{R}} x^2 \mathbb{P}(x)

= \left(0 - \frac{7}{3}\right)^2 + \left(1 - \frac{7}{3}\right)^2 + \left(2 - \frac{7}{3}\right)^2 + \left(3 - \frac{7}{3}\right)^2 + \left(4 - \frac{7}{3}\right)^2 = \frac{198}{81} = \frac{22}{9} \approx 2.457

E(Y)

= \arg \min_{x \in \mathcal{H}} \max_{y \in \mathcal{H}} d(x, y) \mathbb{P}(y)

by definition usual metric on real line (Definition 6.5 page 17)

= \arg \min_{x \in \mathcal{H}} \max_{y \in \mathcal{H}} |x - y| \mathbb{P}(y)

by definition integer line (Definition 6.6 page 17)

= \frac{12}{5}

because expression is minimized at argument \( x = \frac{12}{5} \)

\mathbb{V} \mathbb{a}r(Y; \hat{\mathcal{C}}) = \sum_{x \in \mathbb{Z}} d^2(\hat{\mathcal{C}}(Y), x) \mathbb{P}(x)

by definition of \( \mathbb{V} \mathbb{a}r \) (Definition 9.2 page 42)

= \sum_{x \in \mathbb{Z}} d^2\left(\frac{12}{5}, x\right) \mathbb{P}(x)

by \( \mathcal{H} \) result

= \left(\frac{12}{5} - 0\right)^2 + \left(\frac{12}{5} - 1\right)^2 + \left(\frac{12}{5} - 2\right)^2 + \left(\frac{12}{5} - 3\right)^2 + \left(\frac{12}{5} - 4\right)^2

= \frac{551}{225} \approx 2.449

E(Z) = E[\mathbb{H} \setminus \hat{\mathcal{C}}(\mathcal{G})]

because \( \mathcal{G} \) and \( \mathcal{H} \) are isomorphic under \( \mathcal{Z} \)

by \( \hat{\mathcal{C}}(\mathcal{G}) \) result

= {0, 1, 2, \ldots, 8}

because the distances for values of \( x \) in \( \mathcal{G}_y \) are the same
\[ \bar{\text{Var}}(Z; \mathcal{E}) \triangleq \sum_{x \in \mathbb{N}} d^2(\mathcal{E}(Z), x) \mathbb{P}(x) \quad \text{by definition of } \bar{\text{Var}} \text{ (Definition 9.2 page 42)} \]

\[ = \sum_{x \in \mathbb{N}} d^2(\{4\}, x) \mathbb{P}(x) \quad \text{by } \mathcal{E}(Z) \text{ result} \]

\[ = 1^2 \times \frac{2}{9} + 2^2 \times \frac{1}{9} + 2^2 \times \frac{1}{9} + 0^2 \times \frac{3}{9} = \frac{12}{9} = \frac{4}{3} \approx 1.333 \]

---

**Example 9.18**  (LCG mappings, sequential ordering)

In Example 9.17 (page 55), the structures \( G_9, G \), and \( H \) were ordered as a standard ring of integers \( 0 < 1 < 2 < \cdots < 7 < 8 < 0 \) for \( G_9 \). In this current example, as illustrated in Figure 13 (page 58), these structures are ordered as they appear in the sequences generated by \( x_{n+1} = (7x_n + 5) \mod 9 \) and \( s \) (see Example 9.17 for sequence description). This yields the following statistics:

- **Geometry of \( G_9 \):** \( \hat{\mathbb{U}}(G_9) = \{0, 1, 2, 3, 4, 5, 6, 7, 8\} \)
- **Geometry of \( G \):** \( \hat{\mathbb{U}}(G) = \{3\} \)
- **Outcome subspace statistics on isomorphic structure:** \( \mathcal{E}(Z) = \{3\} \quad \bar{\text{Var}}(Z; \mathcal{E}) = \frac{10}{9} \approx 1.111 \)

Note that a change in ordering structure (from standard ring ordering to sequential ordering) yields a change in statistics (\( \mathcal{E}(Z) = \{3\} \) as opposed to \( \mathcal{E}(Z) = \{4\} \)). Intuitively, the sequential ordering of Example 9.18 should yield a better estimate than that of Example 9.17, because it more closely matches the way the PRNG produces a sequence. This intuition is also supported by the variance values \( \bar{\text{Var}}(Z) = \frac{12}{9} \) for standard ring ordering, \( \bar{\text{Var}}(Z) = \frac{10}{9} \) for sequential ordering). However, counterintuitively, the sequential ordering no longer yields the maximally likely result of \( \{4\} \).

---

**Proof:**

\[ \hat{\mathbb{U}}(G_9) = \{0, 1, 2, \cdots, 8\} \quad \text{by LCG mappings standard ordering example (Example 9.17 page 55)} \]

\[ \hat{\mathbb{U}}(G) = \{3\} \quad \text{by Example 8.14 (page 35)} \]

\[ E(Z) = Z[\hat{\mathbb{U}}(G)] \quad \text{because } G \text{ and } H \text{ are isomorphic under } Z \]

\[ = Z[\{3\}] \quad \text{by } \hat{\mathbb{U}}(G) \text{ result} \]

\[ = \{3\} \quad \text{by definition of } Z \]

\[ \bar{\text{Var}}(Z; \mathcal{E}) = \bar{\text{Var}}(G) \quad \text{because } G \text{ and } H \text{ are isomorphic under } Z \]
= \frac{10}{9} \approx 1.111 \quad \text{by Example 8.14 (page 35)}

Figure 14: LCG mappings to linear (X), non-linear discrete (Y) and non-linear continuous (Z) ordered metric spaces (Example 9.19 page 59)

Example 9.19  (LCG mappings, sequential directed graph)
Let \( G, H \) and \( Z \) be illustrated in Figure 14 (page 59). In Example 9.18 (page 58), the outcome values were ordered sequentially like a PRNG, but the metrics were commutative, which is unlike a PRNG. In this example, the outcomes are assigned quasi-metrics (Definition 4.1 page 8, Remark 6.3 page 16) that are non-commutative. For example in the shaped sequence \( s(x_n) = (\ldots, 3, 4, 4, 1, 3, \ldots) \), the "distance" from 3 to 4 is \( \rho(3, 4) = 1 \), but from 4 to 3 is \( \rho(4, 3) = 2 \). This yields the following statistics:

- geometry of \( G \):
  \[ \hat{\mathcal{C}}(G) = \{0, 1, 2, 3, 4, 5, 6, 7, 8\} \]

- geometry of \( G \):
  \[ \mathcal{C}(G) = \{3\} \]

- outcome subspace statistics on isomorphic structure:
  \[ \mathbb{E}(Z) = \{3\} \quad \text{Var}(Z; \mathbb{E}) = \frac{4}{3} \approx 1.333 \]

Note that this technique yields the same estimate \( \mathbb{E}(Z) = \{3\} \) as Example 9.18, but with a larger variance.

\textbf{Proof:}

\[ \hat{\mathcal{C}}(G_y) \triangleq \arg\min_{x \in G_y} \max_{y \in G_y} d(x, y) P(y) \]
by definition of \( \hat{\mathcal{C}} \) (Definition 8.3 page 22)

\[ = \arg\min_{x \in G_y} \max_{y \in G_y} d(x, y) \frac{1}{9} \]
by definition of \( G_y \)

\[ = \arg\min_{x \in G_y} \max_{y \in G_y} d(x, y) \]
because \( \phi(x) = \frac{1}{9} x \) is strictly isotone and by Lemma 7.9 page 21

\[ = \arg\min_{x \in G_y} \{8, 8, 8, 8, 8, 8, 8, 8\} \]
because the maximum distance in \( G_y \) from any \( x \) is 8

\[ = \{0, 1, 2, \ldots, 8\} \]
because the distances for values of \( x \) in \( G_y \) are the same

\[ \hat{\mathcal{C}}(G) = \{3\} \]
by Example 8.15 (page 38)

\[ \mathbb{E}(Z) = \mathbb{E}(\hat{\mathcal{C}}(G)) \]
because \( G \) and \( H \) are isomorphic under \( Z \)

\[ = \mathbb{E}(\{3\}) \]
by \( \hat{\mathcal{C}}(G) \) result

\[ = \{3\} \]
by definition of \( Z \)

\[ \text{Var}(Z; \mathbb{E}) = \text{Var}(\hat{\mathcal{C}}(G)) \]
because \( G \) and \( H \) are isomorphic under \( Z \)

\[ = \frac{4}{3} \approx 1.333 \]
by Example 8.15 (page 38)

\( \blacksquare \)
9.4.5 Genomic signal processing (GSP) examples

Example 9.20 (DNA to linear structures) Genomic Signal Processing (GSP) analyzes biological sequences called genomes. These sequences are constructed over a set of 4 symbols that are commonly referred to as A, T, C, and G, each of which corresponds to a nucleobase (adenine, thymine, cytosine, and guanine, respectively).58 A typical genome sequence contains a large number of symbols (about 3 billion for humans, 29751 for the SARS virus).59 Let $\mathcal{G} \triangleq (\{A, T, C, G\}, d, \varnothing, P)$ be the outcome subspace (Definition 8.1 page 22) generated by a genome where $d$ is the discrete metric (Example 4.21 page 14), $\leq = \varnothing$ indicates a completely unordered set (Definition 2.4 page 4), and $\mathbb{P}(A) = \mathbb{P}(T) = \mathbb{P}(C) = \mathbb{P}(G) = \frac{1}{4}$ (uniformly distributed).

Let $\mathcal{H} \triangleq (\mathbb{R}, |\cdot|, \leq)$ be the real line (Definition 6.5 page 17). This yields the following statistics:

- geometry of $\mathcal{G}$:
  $\mathcal{G} = \{A, T, C, G\}$
- traditional statistics on real line:
  $E(X) = 1.5$, $\text{Var}(X; E) = \frac{5}{4} = 1.25$
- outcome subspace statistics on real line:
  $\mathcal{E}(X) = 1.5$, $\mathcal{V}(X; \mathcal{E}) = \frac{5}{4} = 1.25$
- outcome subspace statistics on integer line:
  $\mathcal{E}(Y) = \{1, 2\}$, $\mathcal{V}(Y; \mathcal{E}) = \frac{1}{4} = 0.5$

The symbols A, T, C and G in general again have an order structure and a metric geometry (Remark 8.9 page 26) that is fundamentally dissimilar from that mapped to by the random variables $X$ and $Y$. Therefore, statistical inferences made using these random variables will likely lead to results that arguably have little relationship with intuition or reality.

**Proof:**

\[
\mathcal{E}(\mathcal{G}) = \{A, T, C, G\} \quad \text{by Example 8.18 page 40}
\]

\[
E(X) = \int_{\mathbb{R}} xP(x) \, dx \quad \text{by definition of } E \text{ (Definition 5.6 page 15)}
\]

\[
= \sum_{x \in \mathbb{Z}} xP(x) \quad \text{by definition of } P \text{ and Proposition 5.7 page 15}
\]

\[
= \frac{1}{4} \sum_{x \in \{0, 1, 2, 3\}} x \quad \text{by definitions of } \mathcal{G}, \mathcal{H} \text{ and } X
\]

\[
= \frac{1}{4} (0 + 1 + 2 + 3) = \frac{6}{4} = \frac{3}{2} = 1.5
\]

\[
\mathcal{V}(X; E) = \mathcal{V}(X) \quad \text{by Theorem 9.3 page 42}
\]

---

58 [77], (Mendel (1853): gene coding uses discrete symbols), [112], page 737, (Watson and Crick (1953): gene coding symbols are adenine, thymine, cytosine, and guanine), [111], page 965, [93], page 52

59 [1], (http://www.ncbi.nlm.nih.gov/genome/guide/human/), (Homo sapiens, NC_000001–NC_000022 (22 chromosome pairs), NC_000023 (X chromosome), NC_000024 (Y chromosome), NC_012920 (mitochondria)), [1], (http://www.ncbi.nlm.nih.gov/nuccore/30271926), (SARS coronavirus, NC_004718.3) [100], (homo sapien chromosome 1), [99], (SARS coronavirus)
\[
\int \mathcal{R} (x - \mathcal{E}(X))^2 \mathcal{P}(x) \\
= \sum_{x \in \mathcal{Z}} (x - \mathcal{E}(X))^2 \mathcal{P}(x) \\
= \frac{1}{4} \sum_{x \in \mathcal{H}} \left[ x - \frac{3}{2} \right]^2 \\
= \frac{1}{4} \left\{ (0 - \frac{3}{2})^2 + (1 - \frac{3}{2})^2 + (2 - \frac{3}{2})^2 + (3 - \frac{3}{2})^2 \right\} \\
= \frac{1}{4} \cdot \frac{20}{16} = \frac{5}{4} = 1.25 \\
\]

\[
\mathcal{E}(X) = \mathcal{E}(X) \\
= \frac{3}{2} = 1.5 \\
\mathcal{E}(X) \triangleq \arg \min_{x \in \mathcal{R}} \max_{y \in \mathcal{R}} d(x, y) \mathcal{P}(y) \\
= \arg \min_{x \in \mathcal{R}} \max_{y \in \mathcal{R}} |x - y| \mathcal{P}(y) \\
= \arg \min_{x \in \mathcal{R}} \max_{y \in \{0,1,2,3\}} |x - y| \frac{1}{4} \\
= \arg \min_{x \in \mathcal{R}} \left\{ \begin{array}{ll}
|x - 1| & \text{for } x \leq \frac{3}{2} \\
|x - 0| & \text{otherwise}
\end{array} \right\} \\
= \{ \frac{3}{2} \} = \frac{1}{2} = 1.5 \\
\mathcal{V} \mathcal{A} \mathcal{R}(X; \mathcal{E}) = \mathcal{V} \mathcal{A} \mathcal{R}(X; \mathcal{E}) \\
= \frac{5}{4} \\
\mathcal{E}(X) \triangleq \arg \min_{x \in \mathcal{Z}} \max_{y \in \mathcal{Z}} d(x, y) \mathcal{P}(y) \\
= \arg \min_{x \in \mathcal{Z}} \max_{y \in \mathcal{Z}} |x - y| \mathcal{P}(y) \\
= \arg \min_{x \in \mathcal{Z}} \max_{y \in \{0,1,2,3\}} |x - y| \frac{1}{4} \\
= \arg \min_{x \in \mathcal{Z}} \max_{y \in \{0,1,2,3\}} |x - y| \\
= \arg \min_{x \in \{0,1,2,3\}} \left\{ \begin{array}{ll}
0 & \text{for } x \leq \frac{3}{2} \\
1 & \text{otherwise}
\end{array} \right\} \\
= \left\{ \frac{3}{2} \right\} = \frac{3}{2} = 1.5 \\
\mathcal{V} \mathcal{A} \mathcal{R}(X; \mathcal{E}) = \mathcal{V} \mathcal{A} \mathcal{R}(X; \mathcal{E}) \\
= \frac{5}{4} \\
\mathcal{E}(X) \triangleq \arg \min_{x \in \mathcal{Z}} \max_{y \in \mathcal{Z}} d(x, y) \mathcal{P}(y) \\
= \arg \min_{x \in \mathcal{Z}} \max_{y \in \mathcal{Z}} |x - y| \mathcal{P}(y) \\
= \arg \min_{x \in \mathcal{Z}} \max_{y \in \{0,1,2,3\}} |x - y| \frac{1}{4} \\
= \arg \min_{x \in \mathcal{Z}} \max_{y \in \{0,1,2,3\}} |x - y| \\
= \arg \min_{x \in \{0,1,2,3\}} \left\{ \begin{array}{ll}
0 & \text{for } x \leq \frac{3}{2} \\
1 & \text{otherwise}
\end{array} \right\} \\
= \left\{ \frac{1}{2} \right\} = \frac{1}{2} = 0.5 \\
\]

**Example 9.21** (GSP to complex plane)
A possible solution for the GSP problem (Example 9.20 page 60) is to map \( \{A, T, C, G\} \) to the complex plane (Example 6.9 page 17) rather than the real line (Definition 6.5 page 17) such that (see also illustration to the right)
\[
\begin{align*}
A & \triangleq X(A) = -1 + i \\
T & \triangleq X(T) = 1 + i \\
C & \triangleq X(C) = -1 - i \\
G & \triangleq X(G) = 1 - i.
\end{align*}
\]
However, this solution also is arguably unsatisfactory for two reasons:

1. The order structures are dissimilar. Note that \( C < A \), but \( C \) and \( A \) are incomparable (Definition 2.4 page 4).

2. The metric geometries are dissimilar. Let \( \delta \) be the discrete metric and \( \rho \) the usual metric in \( \mathbb{C} \).

   Note that \( \delta (A, T) = \delta (A, C) = \delta (A, G) = 1 \), but \( \rho (a, t) = |a - t| = 2 \neq 2\sqrt{2} = |a - g| = \rho (a, g) \).

Example 9.22 (DNA mapping with extended range)
Example 9.20 (page 60) presented a mapping from a DNA structure to a linearly ordered lattices, but the order and metric geometry was not preserved. In this example, a different structure is used that does preserve both order and metric geometry (see illustration to the right). This yields the following statistics:
\[
\begin{align*}
\mathcal{E}(X) & = \{0\} \\
\mathcal{V}_\mathcal{A}(X) & = \frac{1}{4}
\end{align*}
\]

**Proof:**
\[
\begin{align*}
\mathcal{E}(H) & \triangleq \arg \min_{x \in \mathcal{H}} \max_{y \in \mathcal{H}} \delta (x, y) \rho (y) & \text{by definition of } \mathcal{E} \text{ (Definition 9.2 page 42)} \\
& = \arg \min_{x \in \mathcal{H}} \max_{y \in \mathcal{H}} \delta (x, y) \rho (y) & \text{because } \rho (0) = 0 \\
& = \arg \min_{x \in \mathcal{H}} \max_{y \in \mathcal{H}} \delta (x, y) \frac{1}{4} & \text{by definition of } \mathcal{G} \\
& = \arg \min_{x \in \mathcal{H}} \max_{y \in \mathcal{H}} \delta (x, y) & \text{because } \phi (x) = \frac{1}{4} \text{ is strictly isotone and by Lemma 7.9 page 21} \\
& = \{0\} & \text{because expression is minimized at } x = \{0\} \\
\mathcal{V}_\mathcal{A}(X) & = \sum_{x \in \mathcal{H}} \delta ^2 (\mathcal{E}(X), x) \rho (x) & \text{by definition of } \mathcal{V}_\mathcal{A} \text{ (Definition 9.2 page 42)} \\
& = \sum_{x \in \mathcal{H}} \delta ^2 (\{0\}, x) \rho (x) & \text{by } \mathcal{E}(X) \text{ result} \\
& = \sum_{x \in \mathcal{H}\setminus\{0\}} \left( \frac{1}{2} \right) ^2 \frac{1}{4} = |\mathcal{H}\setminus\{0\}| \left( \frac{1}{2} \right) ^2 \frac{1}{4} = 4 \left( \frac{1}{2} \right) ^2 \frac{1}{4} = \frac{1}{4}
\end{align*}
\]

Example 9.23 (GSP with Markov model)
Markov probability models have often been used in genomic signal processing (GSP). A change in the statistics in the sequence may in some cases mean
a change in function of the genomic sequence (DNA code). Finding such a change in statistics then is very useful in identifying functions of segments of genomic sequences. Let $G$ be an outcome subspace (Definition 8.1 page 22) representing a Markov model of depth 2 for a genomic sequence as illustrated in Figure 16 (page 63), with joint and conditional probabilities computed over a finite window. Let $H$ be an outcome subspace isomorphic to $G$, and $X$ be a random variable mapping $G$ to $H$. A change in the value of the statistic $\hat{E}(X)$ over the window then may indicate a change in function within the genomic sequence.
10 Operations on outcome subspaces

10.1 Summation

Example 10.1 (pair of dice outcome subspace) A pair of real dice has a structure as illustrated in Figure 17 (page 64). The values represent the standard sum of die faces and thus range from 2 to 12. The table in the figure provides the metric distances between summed values based on the number of edges that must be transversed to move from the first value to the second value. Alternatively, the distance is the number of times the dice must be rotated 90 degrees to move from the first value being face up to the second value being face up. This structure is also illustrated in the undirected graph on the right in Figure 17, along with each value's standard probability.

\[ \hat{\mathcal{C}}(G) \triangleq \arg \min_{x \in G} \max_{y \in G} d(x, y) \mathbb{P}(y) \]

by definition of \( \hat{\mathcal{C}} \) (Definition 8.3 page 22)

\[
= \arg \min_{x \in G} \max_{y \in G} \begin{cases} 
\frac{1}{36} \quad \text{for } y = 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 \\
\frac{1}{36} \quad \text{for } y = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 \\
\end{cases}
\]

Many many thanks to Katie L. Greenhoe and Jonathan J. Greenhoe for help computing these values.
10 OPERATIONS ON OUTCOME SUBSPACES

Daniel J. Greenhoe

The next two examples are examples of sums of outcome subspaces (Definition 8.1 page 22): Example 10.2 page 65 (sum of dice pair) and Example 10.4 page 68 (sum of spinner pair).

\[
\hat{c}(G) \triangleq \min_{x \in G} \sum_{y \in G} d(x, y)P(y)
\]

by definition of \(\hat{c}\) (Definition 8.4 page 22)

\[
\hat{c}(G) = \min_{x \in G} \frac{1}{36} \left\{ 0 + 2 + 3 + 4 + 5 + 12 + 10 + 8 + 6 + 6 + 4 \\
+ 1 + 0 + 3 + 4 + 5 + 6 + 5 + 8 + 6 + 4 + 2 \\
+ 1 + 2 + 3 + 4 + 0 + 6 + 5 + 4 + 3 + 4 + 2 \\
+ 2 + 2 + 3 + 4 + 5 + 0 + 5 + 4 + 3 + 2 + 2 \\
+ 2 + 4 + 3 + 4 + 5 + 6 + 0 + 4 + 3 + 2 + 1 \\
+ 2 + 4 + 6 + 4 + 5 + 6 + 5 + 0 + 3 + 2 + 1 \\
+ 2 + 4 + 6 + 8 + 5 + 6 + 5 + 4 + 0 + 2 + 1 \\
+ 3 + 4 + 6 + 8 + 10 + 6 + 5 + 4 + 3 + 0 + 1 \\
+ 4 + 6 + 6 + 8 + 10 + 12 + 5 + 4 + 3 + 2 + 0 \right\} 
\]

by definition of Var (Definition 9.2 page 42)

\[
\text{Var}(G) \triangleq \sum_{x \in G} [d(E(G), x)]^2P(x)
\]

by definition of \(\text{Var}\) (Definition 9.2 page 42)

\[
\text{Var}(G) = \sum_{x \in G} [d(7, x)]^2P(x)
\]

= \(\frac{1}{36} (2^2 \times 1 + 1^2 \times 2 + 1^2 \times 3 + 1^2 \times 4 + 1^2 \times 5 + 0^2 \times 6 + 1^2 \times 5 + 1^2 \times 4 + 1^2 \times 3 + 1^2 \times 2 + 2^2 \times 1) \)

= \(\frac{1}{36} (\overline{2} + 3 + 4 + 5 + 0 + 5 + 4 + 3 + 2 + 4) \)

= \(\frac{36}{36} \)

= 1

Figure 18: \textit{pair of dice mappings} (Example 9.14 page 51)
Example 10.2  (pair of dice and hypothesis testing)  Let \( G \) be the pair of dice outcome subspace (Example 10.1 page 64), \( X \in (\mathbb{R}, |\cdot|, \leq)^G \) an outcome random variable mapping from \( G \) to the real line (Definition 6.5 page 17), and \( X \in H^G \) a mapping to a structure \( H \) that is isomorphic to \( G \), as illustrated in Figure 18 (page 65). This yields the following statistics:

- **geometry of \( G \):** \( \hat{\mathcal{C}}(G) = \{7\} \)  
- **traditional statistics on real line \((\mathbb{R}, |\cdot|, \leq)\):** \( \mathcal{E}(X) = 7 \) \( \mathbb{V} \mathbb{A} \mathbb{R}(W; E) = \frac{35}{6} \approx 5.833 \)
- **outcome subspace statistics on real line \((\mathbb{R}, |\cdot|, \leq)\):** \( \hat{\mathbb{V}} \mathbb{A} \mathbb{R}(\mathcal{Y}) = \{7\} \) \( \hat{\mathbb{V}} \mathbb{A} \mathbb{R}(\mathcal{Y}; \hat{\mathbb{V}} \mathbb{A} \mathbb{R}) = 35\frac{6}{6} \approx 5.833 \)
- **outcome subspace statistics on isomorphic structure \( H \):** \( \hat{\mathbb{V}} \mathbb{A} \mathbb{R}(\mathcal{Y}) = \{7\} \) \( \hat{\mathbb{V}} \mathbb{A} \mathbb{R}(\mathcal{Y}; \hat{\mathbb{V}} \mathbb{A} \mathbb{R}) = 1 \)

Although the expected values of both outcome subspaces are essentially the same (7 and \( \{7\} \)), the isomorphic structure \( H \) yields a much smaller variance (a much smaller expected error). This is significant in statistical applications such as hypothesis testing. Suppose for example we have two pair of real dice (Example 8.8 page 24), one pair being made of two uniformly distributed die and one pair of weighted die. We want to know which pair is the uniform die. So we roll each pair one time. Suppose the outcome of the first pair is 11 and the outcome of the second pair is 6. Which pair is more likely to be the uniform pair? Using traditional statistical analysis, the answer is the second pair, because it is closer to the expected value \((0.414 \text{ standard deviations as opposed to } 1.656 \text{ standard deviations})\). However, this result is deceptive, because as can be seen in Figure 17 (page 64), the distance from the expected value to the values 11 and 6 are the same \( \hat{\mathbb{D}}(7, 11) = \hat{\mathbb{D}}(7, 6) = 1 \) standard deviation). So arguably the outcome of the single roll test would contribute nothing to a good decision algorithm.

**Proof:**

\[
\hat{\mathcal{C}}(G) = \{7\} \quad \text{by Example 10.1 (page 64)}
\]
\[
\hat{\mathbb{V}} \mathbb{A} \mathbb{R}(G) = 1 \quad \text{by Example 10.1 (page 64)}
\]
\[
\mathcal{E}(X) = \int_{x \in \mathbb{R}} x \mathbb{P}(x) \, dx \quad \text{by definition of } \mathcal{E} \text{ (Definition 5.6 page 15)}
\]
\[
\Delta = \frac{1}{36} \sum_{x \in \mathbb{Z}} x \mathbb{P}(x) \quad \text{by definition of } \mathbb{P}
\]
\[
= \frac{1}{36} (2 \times 1 + 3 \times 2 + 4 \times 3 + 5 \times 4 + 6 \times 5 + 7 \times 6 + 8 \times 5 + 9 \times 4 + 10 \times 3 + 11 \times 2 + 12 \times 1)
\]
\[
= \frac{252}{36} = 7
\]
\[
\hat{\mathbb{V}} \mathbb{A} \mathbb{R}(X; \mathcal{E}) = \mathbb{V} \mathbb{A} \mathbb{R}(X) \quad \text{by Theorem 9.3 (page 42)}
\]
\[
= \int_{x \in \mathbb{R}} [x - \mathcal{E}(X)]^2 \mathbb{P}(x) \, dx \quad \text{by definition of } \mathbb{V} \mathbb{A} \mathbb{R} \text{ (Definition 5.6 page 15)}
\]
\[
= \sum_{x \in \mathbb{Z}} [x - \mathcal{E}(X)]^2 \mathbb{P}(x) \quad \text{by definition of } \mathbb{P}
\]
\[
= \sum_{x \in \mathbb{Z}} (x - 7)^2 \cdot \frac{1}{36} \quad \text{by } \mathcal{E}(X) \text{ result}
\]
\[
= \frac{2 \times 25 \times 1 + 16 \times 2 + 9 \times 3 + 4 \times 4 + 1 \times 5}{36} = \frac{35}{6} \approx 5.833
\]
\[
\mathcal{E}(X) = \mathcal{E}(X) \quad \text{by Theorem 9.5 (page 43)}
\]
\[
= \{7\} \quad \text{by } \mathcal{E}(X) \text{ result}
\]
\[
\hat{\mathbb{V}} \mathbb{A} \mathbb{R}(X; \mathcal{E}) = \mathbb{V} \mathbb{A} \mathbb{R}(X; \mathcal{E}) \quad \text{because } \mathcal{E}(X) = \mathcal{E}(X)
\]
\[
= \frac{35}{6} \approx 5.833 \quad \text{by } \mathbb{V} \mathbb{A} \mathbb{R}(X; \mathcal{E}) \text{ result}
\]
\[
\mathcal{E}(Y) = \gamma \{\hat{\mathcal{C}}(G)\} \quad \text{because } G \text{ and } H \text{ are isometric under } Y
\]
Figure 19: metrics based on a pair of spinners (Example 10.3 page 67)

**Example 10.3** (pair of spinners) A pair of spinners (Example 8.11 page 29) has a structure as illustrated in Figure 19 (page 67). The values represent the standard sum of spinner positions (1, 2, ..., 6) and thus range from 2 to 12. The table in the figure provides the metric distances between summed values based on how many positions one must traverse to get from one value to the next (in which ever direction is shortest). This structure is also illustrated in the undirected graph in the upper right of Figure 19, along with each value's standard probability.

**Proof:**

\[
\hat{c}(G) \triangleq \arg \min_{x \in G} \max_{y \in G} d(x, y)P(y) \quad \text{by definition of } \hat{c} \text{ (Definition 8.3 page 22)}
\]

\[
= \arg \min_{x \in G} \max_{y \in G} \begin{pmatrix}
    d(2,2)P(2) & d(2,3)P(3) & d(2,4)P(4) & \cdots & d(2,12)P(12) \\
    d(3,2)P(2) & d(3,3)P(3) & d(3,4)P(4) & \cdots & d(3,12)P(12) \\
    \vdots & \vdots & \vdots & \ddots & \vdots \\
    d(12,2)P(2) & d(12,3)P(3) & d(12,4)P(4) & \cdots & d(12,12)P(12)
\end{pmatrix}
\]

\[
= \arg \min_{x \in G} \max_{y \in G} \frac{1}{36} \begin{pmatrix}
    0 \times 1 & 1 \times 2 & 2 \times 3 & 3 \times 4 & 2 \times 5 & 1 \times 6 & 2 \times 5 & 3 \times 4 & 4 \times 3 & 3 \times 2 & 2 \times 1 \\
    1 \times 1 & 0 \times 2 & 1 \times 3 & 2 \times 4 & 3 \times 5 & 2 \times 6 & 1 \times 5 & 2 \times 4 & 3 \times 3 & 4 \times 2 & 3 \times 1 \\
    2 \times 1 & 1 \times 2 & 0 \times 3 & 1 \times 4 & 2 \times 5 & 3 \times 6 & 2 \times 5 & 1 \times 4 & 2 \times 3 & 3 \times 2 & 4 \times 1 \\
    3 \times 1 & 2 \times 2 & 1 \times 3 & 0 \times 4 & 1 \times 5 & 2 \times 6 & 3 \times 5 & 2 \times 4 & 3 \times 3 & 2 \times 2 & 3 \times 1 \\
    2 \times 1 & 3 \times 2 & 2 \times 3 & 1 \times 4 & 0 \times 5 & 1 \times 6 & 2 \times 5 & 3 \times 4 & 2 \times 3 & 1 \times 2 & 2 \times 1 \\
    1 \times 1 & 2 \times 2 & 3 \times 3 & 2 \times 4 & 1 \times 5 & 0 \times 6 & 1 \times 5 & 2 \times 4 & 3 \times 3 & 2 \times 2 & 1 \times 1 \\
    2 \times 1 & 1 \times 2 & 2 \times 3 & 3 \times 4 & 2 \times 5 & 1 \times 6 & 0 \times 5 & 1 \times 4 & 2 \times 3 & 3 \times 2 & 2 \times 1 \\
    3 \times 1 & 2 \times 2 & 1 \times 3 & 2 \times 4 & 3 \times 5 & 2 \times 6 & 1 \times 5 & 0 \times 4 & 1 \times 3 & 2 \times 2 & 3 \times 1 \\
    4 \times 1 & 3 \times 2 & 2 \times 3 & 1 \times 4 & 2 \times 5 & 3 \times 6 & 2 \times 5 & 1 \times 4 & 0 \times 3 & 1 \times 2 & 2 \times 1 \\
    3 \times 1 & 4 \times 2 & 3 \times 3 & 2 \times 4 & 1 \times 5 & 2 \times 6 & 3 \times 5 & 2 \times 4 & 1 \times 3 & 0 \times 2 & 1 \times 1 \\
    2 \times 1 & 3 \times 2 & 4 \times 3 & 3 \times 4 & 2 \times 5 & 1 \times 6 & 2 \times 5 & 3 \times 4 & 2 \times 3 & 1 \times 2 & 0 \times 1
\end{pmatrix}
\]
10.1 SUMMATION

\[ \begin{align*}
\hat{c}_q(G) &= \arg \min_{x \in G} \sum_{y \in G} d(x, y) P(y) \\
&= \arg \min_{x \in G} \frac{1}{36} \left\{ \begin{array}{c}
0 + 2 + 6 + 12 + 10 + 6 + 10 + 12 + 12 + 6 + 2 \\
1 + 0 + 3 + 8 + 15 + 12 + 5 + 8 + 9 + 8 + 3 \\
2 + 2 + 0 + 4 + 10 + 18 + 10 + 4 + 6 + 6 + 4 \\
3 + 4 + 3 + 0 + 5 + 12 + 15 + 8 + 3 + 4 + 3 \\
4 + 6 + 6 + 4 + 10 + 18 + 10 + 4 + 6 + 6 + 4 \\
3 + 8 + 9 + 8 + 5 + 12 + 15 + 8 + 3 + 4 + 3 \\
2 + 6 + 12 + 10 + 6 + 10 + 12 + 6 + 2 + 0 \\
\end{array} \right\} = \arg \min_{x \in G} \frac{1}{36} \left\{ \begin{array}{c}
78 \\
72 \\
66 \\
60 \\
54 \\
50 \\
46 \\
40 \\
34 \\
28 \\
22 \\
\end{array} \right\} = \{7\}
\end{align*} \]

\[ \text{Var}(G) \triangleq \sum_{x \in G} [d(\mathcal{E}(G), x)]^2 P(x) \]

\[ = \sum_{x \in G} [d(7, x)]^2 P(x) \]

\[ = \frac{1}{36} \left( 1^2 \times 1 + 2^2 \times 2 + 3^2 \times 3 + 2^2 \times 4 + 1^2 \times 5 + 0^2 \times 6 + 1^2 \times 5 + 2^2 \times 4 + 2^2 \times 3 + 2^2 \times 2 + 1^2 \times 1 \right) \]

\[ = \frac{1}{36} \left( 2 + 8 + 27 + 16 + 5 + 0 + 5 + 16 + 12 + 8 + 1 \right) \]

\[ = \frac{100}{36} = \frac{25}{9} = 2.778 \]

Figure 20: pair of spinner mappings (Example 10.4 page 68)

Example 10.4 (pair of spinner and hypothesis testing) Let \( G \) be a pair of spinners (Example 10.3 page 67), \( X \) a random variable mapping to the real line (Definition 6.5 page 17), and \( Y \) a random variable mapping to an ordered metric space (Definition 6.1 page 16) that is isomorphic to \( G \) under \( Y \), as illustrated in Figure 20 (page 68). This yields the following statistics:
geometry of $G$:

$$\hat{\mathcal{U}}(G) = \{7\} \quad \hat{\mathcal{V}}(G) = \{7\} \quad \text{Var}(G) = \frac{25}{9} \approx 2.778$$

traditional statistics on real line $(\mathbb{R}, |\cdot|, \leq)$:

$$E(X) = 7 \quad \text{Var}(W; E) = \frac{15}{36} = \frac{5}{12} \approx 5.833$$

outcome subspace statistics on real line $(\mathbb{R}, |\cdot|, \leq)$:

$$\hat{E}(X) = \{7\} \quad \hat{\text{Var}}(W; \hat{E}) = \frac{15}{36} \approx 5.833$$

outcome subspace statistics on isomorphic structure $H$:

$$\hat{\mathcal{U}}(H) = \{7\} \quad \hat{\mathcal{V}}(H) = \{7\} \quad \text{Var}(H; \hat{\mathcal{U}}) = 1$$

Although the expected value of both outcome subspaces are the same ($$E(X) = \hat{E}(Y) = 7$$), the isomorphic outcome subspace $H$ yields a much smaller variance (a much smaller expected error). This is significant in statistical applications such as hypothesis testing. Suppose for example we have two pair of spinners (Example 8.11 page 29), one pair being made of two uniformly distributed spinners, and one pair of weighted spinners. We want to estimate which is which. So we spin each pair one time. Suppose the outcome of the first pair is $12$ and the the outcome of the second pair is $10$. Which pair is more likely to be the uniform pair? Using traditional statistical analysis, the answer is the second pair, because it is closer to the expected value ($$d(7, 10) = |7 − 10| = 3 = 1.8 \text{ standard deviations as opposed to} d(7, 12) = |7 − 12| = 5 = 3 \text{ standard deviations}$$). However, this result is deceptive, because as can be seen in the table in Figure 19 (page 67), $12$ is actually closer to the expected value in $G$ than is $10$ ($$d(7, 12) = 1 < 3 = d(7, 10)$$). So arguably the better choice, based on this one trial, is the first pair.

Proof:

$$\hat{\mathcal{U}}(G) = \{7\} \quad \text{by Example 10.3 (page 67)}$$

$$\hat{\mathcal{V}}(G) = \{7\} \quad \text{by Example 10.3 (page 67)}$$

$$\text{Var}(G) = \frac{25}{9} \approx 2.778 \quad \text{by Example 10.3 (page 67)}$$

$$E(X) = \frac{252}{36} = 7 \quad \text{by Example 10.2 (page 65)}$$

$$\text{Var}(X) = \frac{35}{6} \approx 5.833 \quad \text{by Example 10.2 (page 65)}$$

$$\hat{E}(X) = E(X) = \{7\} \quad \text{because on real line, $P$ is symmetric, and by Theorem 9.5 page 43}$$

$$\text{Var}(X; \hat{E}) = \text{Var}(X; E) = \text{Var}(X) \quad \text{by Theorem 9.3 page 42}$$

$$\hat{E}(Y) = Y[\hat{\mathcal{U}}(G)] = Y[\{7\}] \quad \text{by Example 10.3 (page 67)}$$

$$= \{7\} \quad \text{by definition of $X$}$$

$$\text{Var}(Y; \hat{E}) = \text{Var}(G) \quad \text{because $G$ and $H$ are isomorphic under $Y$}$$

$$= 1 \quad \text{by Example 10.3 (page 67)}$$

Example 10.5 (linear addition) Let $X$ be a random variable (Definition 9.1 page 42) mapping to a real line ordered metric space (Definition 6.5 page 17) resulting in probability values of

$$P(0) = P(2) = \frac{1}{2}, \text{ and } P(x) = 0 \text{ otherwise.}$$

Let $Y$ be a random variable mapping to a real line ordered metric space resulting in probability values of

$$P(0) = P(1) = \frac{1}{4}, P(2) = \frac{1}{4}, \text{ and } P(x) = 0 \text{ otherwise.}$$

Let $Z \overset{d}{=} X + Y$ be the random variable mapping to the outcome subspace (Definition 8.1 page 22) induced by
adding $X$ and $Y$ resulting in probabilities
$$
\begin{array}{c|cccccc}
  z & 0 & 1 & 2 & 3 & 4 & 5 \\
  \mathbb{P}(z) & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\
\end{array},
$$
and $\mathbb{P}(z) = 0$ otherwise.

Note that although the traditional expectation $E$ (Definition 5.6 page 15) distributes over addition such that
$$
E(X + Y) = \frac{18}{8} = 1 + \frac{5}{4} = E(X) + E(Y),
$$
the alternative expectation $\hat{E}$ (Definition 9.2 page 42) does not:
$$
\hat{E}(X + Y) = \frac{8}{3} \neq \frac{7}{3} = 1 + \frac{4}{3} = \hat{E}(X) + \hat{E}(Y).
$$

\begin{proof}

$E(X) = \int_{x \in \mathbb{R}} x \mathbb{P}(x) \, dx$
by definition of $E$ (Definition 5.6 page 15)

$= \sum_{x \in \mathbb{Z}} x \mathbb{P}(x) \, dx$
by definition of $\mathbb{P}$ and Proposition 5.7 page 15

$= 0 \times \frac{1}{2} + 2 \times \frac{1}{2}$
by definition of $\mathbb{P}$

$= 1$

$E(Y) = \int_{y \in \mathbb{R}} y \mathbb{P}(y) \, dy$
by definition of $E$ (Definition 5.6 page 15)

$= \sum_{y \in \mathbb{Z}} y \mathbb{P}(y) \, dy$
by definition of $\mathbb{P}$ and Proposition 5.7 page 15

$= 0 \times \frac{1}{4} + 1 \times \frac{1}{4} + 2 \times \frac{1}{2}$
by definition of $\mathbb{P}$

$= \frac{5}{4}$

$E(Z) = \int_{z \in \mathbb{R}} z \mathbb{P}(z) \, dz$
by definition of $E$ (Definition 5.6 page 15)

$= \sum_{z \in \mathbb{Z}} z \mathbb{P}(z)$
by definition of $\mathbb{P}$ and Proposition 5.7 page 15

$= 0 \times \frac{1}{8} + 1 \times \frac{1}{8} + 2 \times \frac{3}{8} + 3 \times \frac{1}{8} + 4 \times \frac{2}{8}$
by definition of $\mathbb{P}$

$= \frac{9}{4}$

$\hat{E}(X) = \arg \min_{x \in \mathbb{R}} \max_{y \in \mathbb{R}} d(x, y) \mathbb{P}(y)$
by definition of $\hat{E}$ (Definition 9.2 page 42)

$= \arg \min_{x \in \mathbb{Z}} \max_{y \in \mathbb{Z}} d(x, y) \mathbb{P}(y)$
by definition of $\mathbb{P}(x)$

$= \arg \min_{x \in \mathbb{Z}} \left\{ \begin{array}{cl}
  |2 - x| & \text{for } x \leq 1 \\
  |x| & \text{otherwise}
\end{array} \right.$

$= \{1\}$
because $\max(x)$ is minimized at $x = 1$

\end{proof}
\[ \mathcal{E}(Y) = \arg\min_{x \in \mathbb{R}} \max_{y \in \mathbb{R}} d(x, y) \mathbb{P}(y) \]

by definition of \( \mathcal{E} \) (Definition 9.2 page 42)

\[ = \arg\min_{x \in \mathbb{Z}} \max_{y \in \mathbb{Z}} d(x, y) \mathbb{P}(y) \]

by definition of \( \mathbb{P}(x) \)

\[ = \left\{ \frac{4}{3} \right\} \]

because \( \max(y) \) is minimized at \( y = \frac{4}{3} \)

\[ \mathcal{E}(Z) = \arg\min_{x \in \mathbb{R}} \max_{y \in \mathbb{R}} d(x, y) \mathbb{P}(y) \]

by definition of \( \mathcal{E} \) (Definition 9.2 page 42)

\[ = \arg\min_{x \in \mathbb{Z}} \max_{y \in \mathbb{Z}} \left\{ \begin{array}{ll}
|2 - x| & \text{for } 4 \geq x \geq 4 \\
|x| & \text{otherwise}
\end{array} \right. \mathbb{P}(y) \]

\[ = \left\{ \frac{8}{3} \right\} \]

because \( \max(z) \) is minimized at \( z = \frac{8}{3} \) (see Figure 21 page 70)

\[ \mathcal{E}(2 \mathcal{E}(Z)) = \{4\} = \{2 \times 5 \mod 6\} = 2 \{5\} \mod 6. \]

\[ \overset{\text{Proof}}{=} \]

**10.2 Multiplication**

![Pair of spinner mappings](image)

**Figure 22: pair of spinner mappings** (Example 10.6 page 71)

**Example 10.6** (ring multiplication) Let \( X \in H^G \) be a random variable where \( G \) is the *weighted spinners* illustrated in Figure 22 (page 71). Note that, in agreement with Corollary 9.7 (page 44),

\[ \mathcal{E}(2X) = \{4\} = \{2 \times 5 \mod 6\} = 2 \{5\} \mod 6 = 2 \mathcal{E}(X) \mod 6. \]

\[ \overset{\text{Proof}}{=} \]

\[ \hat{\mathcal{C}}(G) \triangleq \arg\min_{x \in H} \max_{y \in H} d(x, y) \mathbb{P}(y) \]

by definition of \( \hat{\mathcal{C}} \) (Definition 8.4 page 22)

\[ = \arg\min_{x \in H} \max_{y \in H} \frac{1}{10} \left\{ \begin{array}{ccccccc}
0 \times 1 & 1 \times 1 & 2 \times 1 & 3 \times 1 & 2 \times 1 & 1 \times 5 \\
1 \times 1 & 0 \times 1 & 1 \times 1 & 2 \times 1 & 3 \times 1 & 2 \times 5 \\
2 \times 1 & 1 \times 1 & 0 \times 1 & 1 \times 1 & 2 \times 1 & 3 \times 5 \\
3 \times 1 & 2 \times 1 & 1 \times 1 & 0 \times 1 & 1 \times 1 & 2 \times 5 \\
2 \times 1 & 3 \times 1 & 2 \times 1 & 1 \times 1 & 0 \times 1 & 1 \times 5 \\
1 \times 1 & 2 \times 1 & 3 \times 1 & 2 \times 1 & 1 \times 1 & 0 \times 5
\end{array} \right\} = \arg\min_{x \in H} \frac{1}{10} \left\{ \begin{array}{c}
5 \\
10 \\
10 \\
5 \\
5 \\
3
\end{array} \right\} \]

\[ = \{5\} \]

\[ \tilde{\mathcal{C}}(G) \triangleq \arg\min_{x \in H} \sum_{y \in H} d(x, y) \mathbb{P}(y) \]

by definition of \( \tilde{\mathcal{C}} \) (Definition 8.4 page 22)
### 10.3 Metric transformation

It is possible to use a metric transform (Definition 4.7 page 9) to transform the structure of an outcome subspace (Definition 8.1 page 22) into a completely different outcome subspace. This is demonstrated in Example 10.8 (page 72)–Example 10.10 (page 74). Naturally, by doing so one can sometimes even change the geometric centers (Definition 8.3 page 22, Definition 8.4 page 22) of the outcome subspaces, and hence also the statistics of random variables that map to/from them. This is demonstrated in Example 10.10 (page 74)–Example 10.11 (page 74).

**Theorem 10.7**  Let \( \phi \) be a metric preserving function (Definition 4.7 page 9). Let \( \mathbf{G} \triangleq (\Omega, d, \leq, \mathbf{P}) \) and \( \mathbf{H} \) be outcome subspaces (Definition 8.1 page 22).

\[
\begin{align*}
\{ (1). \quad & \phi(\mathbf{H}) = \mathbf{G} \\ (2). \quad & \phi \text{ is strictly isotone} \\ (3). \quad & \mathbf{P} \text{ is uniform} \} & \implies \hat{\mathbf{H}}(\mathbf{H}) = \hat{\mathbf{G}}(\mathbf{G})
\end{align*}
\]

**Proof:**

\[
\begin{align*}
\hat{\mathbf{H}}(\mathbf{H}) &= \phi(\hat{\mathbf{H}}(\mathbf{H})) \\
&= \arg \min_{x \in \mathbf{H}} \max_{y \in \mathbf{H}} \phi[d(x, y)] \mathbf{P}(y) \quad \text{by definition of } \hat{\mathbf{H}} \text{ (Definition 8.3 page 22)} \\
&= \arg \min_{x \in \mathbf{G}} \max_{y \in \mathbf{G}} \phi[d(x, y)] \quad \text{by hypothesis (3) and Lemma 7.9 page 21} \\
&= \arg \min_{x \in \mathbf{G}} d(x, y) \quad \text{by hypothesis (2) and Lemma 7.9 page 21} \\
&= \hat{\mathbf{G}}(\mathbf{G}) \quad \text{by definition of } \hat{\mathbf{G}} \text{ (Definition 8.3 page 22)}
\end{align*}
\]

**Example 10.8** (discrete metric transform on outcome subspaces) Let \( \mathbf{g} \) be a function (a pullback function Theorem 4.6 page 8) such that \( \mathbf{g}(1) = \Box, \mathbf{g}(2) = \Box, \mathbf{g}(3) = \Box, \mathbf{g}(4) = \Box, \mathbf{g}(5) = \Box, \) and \( \mathbf{g}(\Box) = \Box \).
Then under the discrete metric preserving function $\phi$ (Example 4.14 page 10) the real die outcome subspace (Example 8.8 page 24) becomes the fair die outcome subspace (Example 8.7 page 23), and under $\phi \circ g$ the spinner outcome subspace (Example 8.11 page 29) also becomes the fair die outcome subspace, as illustrated in Figure 23 (page 73). This yields the following geometric statistics:

$$\hat{\mathbf{C}}(\mathbf{G}) = \hat{\mathbf{C}}(\mathbf{H}) = \{\heartsuit, \spadesuit, \clubsuit, \diamondsuit, \triangleleft\}.$$ 

Figure 24: Example 4.18 (page 11) metric preserving function $\phi_1$ and Example 4.15 (page 10) metric preserving function $\phi_2$ on spinner outcome subspaces (Example 10.9 page 73, Example 10.10 page 74)

Example 10.9 Let $\phi_1$ be the metric preserving function defined in Example 4.18 (page 11). Then under $\phi_1$, the spinner outcome subspace (Example 8.11 page 29) becomes what is here called the wagon wheel output subspace, as illustrated on the left in Figure 24 (page 73). Let $\mathbf{G}$ be the spinner outcome subspace and $\mathbf{H}$ the wagon wheel outcome subspace. This yields the following geometric statistics:

$$\hat{\mathbf{C}}(\mathbf{G}) = \{\heartsuit, \spadesuit\} \quad \hat{\mathbf{C}}(\mathbf{H}) = \{\spadesuit\}.$$ 

Note that the metric transform $\phi_1$ also moves the outcome center from one that is not maximally likely, to one that is.

\[\text{Proof:}\]

$$\hat{\mathbf{C}}(\mathbf{G}) \triangleq \arg \min_{x \in \mathbf{G}} \max_{y \in \mathbf{G}} d(x, y) P(y) \quad \text{by definition of } \hat{\mathbf{C}} (\text{Definition 8.3 page 22})$$

$$= \arg \min_{x \in \mathbf{G}} \max_{y \in \mathbf{G}} \frac{1}{10} \{0 \times 1 \quad 1 \times 2 \quad 2 \times 1 \quad 3 \times 1 \quad 2 \times 4 \quad 1 \times 1 \}$$

$$= \arg \min_{x \in \mathbf{G}} \frac{1}{10} \{8 \quad 12 \}$$

$$= \{\heartsuit\}$$

$$\hat{\mathbf{C}}(\mathbf{H}) \triangleq \arg \min_{x \in \mathbf{G}} \max_{y \in \mathbf{G}} d(x, y) P(y) \quad \text{by definition of } \hat{\mathbf{C}} (\text{Definition 8.3 page 22})$$

$$= \arg \min_{x \in \mathbf{G}} \frac{1}{10} \{8 \quad 4 \}$$

$$= \{\spadesuit\}$$
Example 10.10  Let \( \phi_2 \) be the metric preserving function defined in Example 4.15 (page 10). Let \( g \) be the function defined in Example 10.8 (page 72). Then under \( \phi_2 \circ g \), the spinner outcome subspace (Example 8.11 page 29) becomes the weighted die outcome subspace (Example 8.10 page 27), as illustrated on the right in Figure 24 (page 73). Let \( G \) be the spinner outcome subspace and \( H \) the weighted die outcome subspace. These structures have the following geometric statistics:

\[
\hat{\mathcal{C}}(G) = \{4, 6\} \quad \hat{\mathcal{C}}(H) = \{1, 3, 4, 5, 6\}.
\]

Note that in Example 10.9 (page 73), the metric transform \( \phi_1 \) results in a smaller (smaller cardinality \( |\hat{\mathcal{C}}(G)| = 2 > 1 = |\hat{\mathcal{C}}(H)| \)). But here, the metric transform \( \phi_2 \circ g \) results in a larger center. (\( |\hat{\mathcal{C}}(G)| = 2 < 5 = |\hat{\mathcal{C}}(H)| \)).

\[\boxed{\hat{\mathcal{C}}(G) = \{4, 6\}}\]

\[\hat{\mathcal{C}}(H) \triangleq \arg\min_{x \in G} \max_{y \in G} d(x, y) P(y)\]

by definition of \( \hat{\mathcal{C}} \) (Definition 8.3 page 22)

\[
\begin{align*}
\hat{\mathcal{C}}(H) &= \arg\min_{x \in H} \max_{y \in H} \frac{1}{10} \begin{bmatrix}
0 & 1 & 2 & 1 & 1 & 2 & 4 & 1 & 1 \\
1 & 0 & 2 & 1 & 1 & 2 & 4 & 1 & 1 \\
1 & 1 & 2 & 1 & 1 & 2 & 4 & 1 & 1 \\
2 & 1 & 1 & 2 & 1 & 1 & 4 & 1 & 1 \\
2 & 1 & 2 & 1 & 1 & 2 & 1 & 1 & 4 \\
1 & 1 & 2 & 2 & 1 & 1 & 4 & 1 & 1 \\
1 & 1 & 2 & 1 & 1 & 4 & 1 & 1 & 1 \\
1 & 2 & 1 & 1 & 4 & 1 & 1 & 1 & 1 \\
0 & 1 & 2 & 1 & 1 & 4 & 1 & 1 & 0
\end{bmatrix} \\
&= \arg\min_{x \in H} \frac{1}{10} \begin{bmatrix}
4 & 4 & 4 & 4 \\
4 & 4 & 4 & 4 \\
4 & 4 & 4 & 4 \\
4 & 4 & 4 & 4 \\
4 & 4 & 4 & 4 \\
4 & 4 & 4 & 4 \\
4 & 4 & 4 & 4 \\
4 & 4 & 4 & 4 \\
4 & 4 & 4 & 4 \\
\end{bmatrix} = \{5\}
\end{align*}
\]

\[\boxed{\hat{\mathcal{C}}(H) = \{5\}}\]

\[\hat{\mathcal{C}}(G) = \{4, 6\}\]

\[\hat{\mathcal{C}}(H) \triangleq \arg\min_{x \in G} \max_{y \in G} d(x, y) P(y)\]

by definition of \( \hat{\mathcal{C}} \) (Definition 8.3 page 22)

\[
\begin{align*}
\hat{\mathcal{C}}(H) &= \arg\min_{x \in H} \max_{y \in H} \frac{1}{10} \begin{bmatrix}
0 & 1 & 2 & 1 & 1 & 2 & 4 & 1 & 1 \\
1 & 0 & 2 & 1 & 1 & 2 & 4 & 1 & 1 \\
1 & 1 & 2 & 1 & 1 & 2 & 4 & 1 & 1 \\
2 & 1 & 1 & 2 & 1 & 1 & 4 & 1 & 1 \\
2 & 1 & 2 & 1 & 1 & 2 & 1 & 1 & 4 \\
1 & 1 & 2 & 2 & 1 & 1 & 4 & 1 & 1 \\
1 & 1 & 2 & 1 & 1 & 4 & 1 & 1 & 1 \\
1 & 2 & 1 & 1 & 4 & 1 & 1 & 1 & 1 \\
0 & 1 & 2 & 1 & 1 & 4 & 1 & 1 & 0
\end{bmatrix} \\
&= \arg\min_{x \in H} \frac{1}{10} \begin{bmatrix}
4 & 4 & 4 & 4 \\
4 & 4 & 4 & 4 \\
4 & 4 & 4 & 4 \\
4 & 4 & 4 & 4 \\
4 & 4 & 4 & 4 \\
4 & 4 & 4 & 4 \\
4 & 4 & 4 & 4 \\
4 & 4 & 4 & 4 \\
4 & 4 & 4 & 4 \\
\end{bmatrix} = \{5\}
\end{align*}
\]

\[\boxed{\hat{\mathcal{C}}(H) = \{5\}}\]
REFERENCES


[30] R Dedekind, *Was sind und was sollen die Zahlen?*, from: “Gesammelte mathematische Werke”, (R Fricke, E Noether, O Ore, editors), Druck and Verlag von Friedr. Vieweg and Sohn Akt.-Ges., Braunschweig (1888) 335–391What are and what should be numbers?


[37] Euclid, *Elements* (circa 300BC)


REFERENCES


[60] J L W V Jensen, *Sur les fonctions convexes et les inégalités entre les valeurs moyennes (On the convex functions and the inequalities between the average values)*, Acta Mathematica 30 (1906) 175–193


[90] G Peano, Arithmetices principa, nova methodo exposita (1889) The principles of arithmetic presented by a new method
[95] H Ribeiro, Sur les espaces à métrie faible, Portugalica mathematica 4 (1943) 21–40
[99] e a Runtao He, Analysis of multimerization of the SARS coronavirus nucleocapsid protein, Biochemical and Biophysical Research Communications 316 476–483
REFERENCES


**Telecommunications Engineering Department, National Chiao-Tung University, Hsinchu, Taiwan**

dgreenhoe@gmail.com

Received: aa bb 20YY Revised: cc dd 20ZZ