06/30/20

Summer, 2020 • Math 5710
Definitions, Related Shorthand Notations, Axioms, and Theorems

Language and Logic

000. Note that in the world of mathematics, it is critical to distinguish between the name of an entity (e.g., the number 9 or the person Emmy Noether) and names used to refer to that entity (e.g., the numeral "9" or the numeral "8 + 1" or "Emmy Noether" or "Ms. Noether," or "the algebraist who specialized in mathematical physics," or "Amalie Emma Noether") because the name of an entity is not that same entity as the entity itself; to conduct the business of mathematics we need language precision. Thus, in mathematics when referring to an entity itself (e.g., 9 or Emmy Noether) we do not put the word in quotation marks but when we refer to the word itself (e.g., "3" or "Emmy") we use the quotation marks. So "3+ 3+ 3" is not the same name as "6 + 3" which means that "3+ 3+ 3" ≠ "6 + 3." But 3+ 3+ 3 = 6 + 3 because 3 + 3 + 3 is the same number as 6 + 3.

001. Definitions for *proposition* and two types of reasoning w/r determining truth values of propositions:

A. A *proposition* is a statement with a definitive truth value (i.e., it is either true or false).

B. Note the following definitions for *inductive reasoning* and *deductive reasoning*:

i. *Inductive reasoning* is the cognitive process by which people formulate generalizations (e.g., propose a conjecture) based on patterns gleaned from a finite string of specific examples or experiences.

ii. *Deductive reasoning* is the cognitive process by which people determine whether or not a specific example is subsumed by an accepted generalization. Aristotlean syllogism is the paradigm for logical deductions: (major premise $\wedge$ minor premise) $\Rightarrow$ conclusion

C. Note: In the world of mathematics, *inductive reasoning* is often used to formulate conjectures regarding propositions but to prove that a proposition is true (i.e, for the proposition to be elevated from the status of a *conjecture* to the status of a *theorem*), the proof must be based on *deductive reasoning*.

002. Notes w/r propositions:

A. Given p is a proposition, "$\overline{p}$" is read "the negation of p."

B. Given p and q are propositions, "$p \wedge q$" is read "p and q."

C. Given p and q are propositions, "$p \vee q$" is read "p or q."

D. Given p and q are propositions, "$p \veebar q$" is read "either p or q."

E. Given p and q are statements, "$p \Rightarrow q$" is read "p implies q" or "If p is true, then q is true."

Note: Propositions of the form $p \Rightarrow q$ where p and q are statements are referred to as "*conditional propositions*."

F. Given p and q are statements, "$p \Leftrightarrow q$" is read "p implies and is implied q" or "p is true if and only if q is true" (i.e., p iff q).

Note: Propositions of the form $p \Leftrightarrow q$ where p and q are statements are referred to as "*bi-conditional propositions*."

G. Given p and q are statements, the *converse of* $p \Rightarrow q$ is $q \Rightarrow p$ and that the *contrapositive of* $p \Rightarrow q$ is $\overline{q} \Rightarrow \overline{p}$.

H. The truth table for statements p and q:

p	q	$\overline{p}$	$p \vee q$	$p \wedge q$	$p \veebar q$	$p \Rightarrow q$	$q \Rightarrow p$	$\overline{q} \Rightarrow \overline{p}$	$p \Leftrightarrow q$
T	T	F	T	T	F	T	T	T	T
T	F	F	T	F	T	F	T	F	F
F	T	T	T	F	T	T	F	T	F
F	F	T	F	F	F	T	T	T	T

003. Three shorthand notations:

A. "$\exists$" is read "there exist."

B. "$\exists!$" is read "there uniquely exist."

C. "$\forall$" is read "for every."

Naive Set Theory

004. Two undefined words in our set-theoretic world:

“*Set*” and “*Element*”

005. Some fundamental set notations, ideas, words, and phrases:

A. “$x \in A$” is read “x is an element of set A.”

“$x, y, z \in A$” is read “x, y, z are elements of set A.”

Given n is a natural number, “$x_1, x_2, x_3, \ldots, x_n \in A$” is read “$x_1$, x_2, x_3, and so on through x_n are elements of set A.”

B. “$\{x : f(x) = y\}$” is read “the set of all x such that f of x is equal to y.”

006. Note: “$\varnothing$” is a symbol for the empty set.

Definition for the *empty set*:

$A = \varnothing \Leftrightarrow \nexists\, x \ni x \in A.$

007. Note: “V” is a symbol for the *universe* or the *universal set*:

Definition for the *universe* (i.e, *the universal set*):

$A = \mathrm{V} \Leftrightarrow \nexists\, x \ni x \notin \mathrm{V}$

008. Some shorthand conventions for naming subsets of numbers:

A. “$\mathbb{N}$” is read “the set of all natural numbers” or “the set of all positive integers.”

Thus, $\mathbb{N} = \{1, 2, 3, \ldots\}$

B. “ω” is read “the set of all whole numbers” or “the set of all non-negative integers.”

Thus, $\omega = \{0, 1, 2, 3, \ldots\}$

C. “$\mathbb{Z}$” is read “the set of all integers.”

Thus, $\mathbb{Z} = \{\ldots, {}^{-}3, {}^{-}2, {}^{-}1, 0, 1, 2, 3, \ldots\}$

D. "$\mathbb{Q}$" is read "the set of all rational numbers."

Here is a definition for *rational number*:

$$\mathbb{Q} = \{ q : \exists\, n, d \in \mathbb{Z} \ni \frac{n}{d} = q \}$$

E. "$\mathbb{R}$" is read "the set of all real numbers."

F. "$\mathbb{I}$" is read "the set of all irrational numbers."

Here is a definition for *irrational number*:

$$m \in \mathbb{I} \Leftrightarrow (m \in \mathbb{R} \wedge m \notin \mathbb{Q})$$

G. Given $a, b \in \mathbb{R} \ni a \leq b$, "$[a, b]$," depending on the context, is read "the closed interval from a to b."

Here is the definition for a *closed real interval*:

Given $a, b \in \mathbb{R} \ni a \leq b$, $[a, b] = \{ x : x \in \mathbb{R} \wedge a \leq x \leq b \}$

Given $a, b \in \mathbb{R} \ni a \leq b$, "$(a, b)$," depending on the context, is read "the open interval from a to b."

Here is the definition for an *open interval*:
Given $a, b \in \mathbb{R} \ni a \leq b$, $(a, b) = \{ x : x \in \mathbb{R} \wedge a < x < b\}$

Note that the shorthand notation for expressing real number intervals can include intervals open on one side and closed on the other:

Given $a, b \in \mathbb{R} \ni a \leq b$, $[a, b) = \{ x : x \in \mathbb{R} \wedge a \leq x < b\}$

and

Given $a, b \in \mathbb{R} \ni a \leq b$, $(a, b] = \{ x : x \in \mathbb{R} \wedge a < x \leq b\}$

Note the use of the symbol "∞" with the real number interval notation:

Given $a \in \mathbb{R}$, $[a, \infty) = \{ x : x \in \mathbb{R} \wedge a \leq x \}$

and

Given $a \in \mathbb{R}$, $(-\infty, a] = \{ x : x \in \mathbb{R} \;\; x \leq a \}$

and

Given $a \in \mathbb{R}$, $(a, \infty) = \{ x : x \in \mathbb{R} \wedge a < x \}$

and

Given $a \in \mathbb{R}$, $(-\infty, a) = \{ x : x \in \mathbb{R} \wedge x < a \}$

and

$(-\infty, \infty) = \mathbb{R}$

H. "$\mathbb{C}$" is read "the set of all complex numbers."

Here is a definition for $\mathbb{C}$:

Given $a, b \in \mathbb{R} \;\wedge\; i = \sqrt{-1}$, $(\mathbb{C} = \{ a + bi : a, b \in \mathbb{R} \})$

I. Definition for { imaginary numbers }:

Given $a, b \in \mathbb{R} \;\wedge\; i = \sqrt{-1}$, ({ imaginary numbers } $= \{ a + bi : a, b \in \mathbb{R} \wedge b \neq 0 \}$)

009. Note w/r { sets }:

{ sets } is a nebulous entity that sane mathematicians avoid passionately because of the disquieting paradoxes its existence generates. Further note that Jim is insane.

010. Relationships among elements of { sets }:

A. Given $A, B \in$ { sets }, note that "$A \subseteq B$" is read "A is a subset of B."

Definition for *subset*:

Given $A, B \in$ { sets }, $(A \subseteq B \Leftrightarrow (x \in A \Rightarrow x \in B))$

B. Given $A, B \in$ { sets }, note that "$A = B$" is read "A is equal to B."

Definition for *equal* sets:

Given $A, B \in$ { sets }, $(A = B \Leftrightarrow (A \subseteq B \wedge B \subseteq A))$

C. Given $A, B \in$ {sets}, note that "$A \subset B$" is read "A is a proper subset of B."

Definition for *proper subset*:

Given $A, B \in$ {sets}, $(A \subset B \Leftrightarrow (A \subseteq B \wedge A \neq B))$

011. Operations involving elements of { sets }:

A. Given $A, B \in$ { sets }, note that "$A \cup B$" is read "the union of A and B."

Definition for *union*:

Given $A, B \in$ { sets }, $(A \cup B = \{ x : x \in A \ \vee \ x \in B \})$

B. Given $A, B \in$ { sets }, note that "$A \cap B$" is read "the intersection of A and B."

Definition for *intersection*:

Given $A, B \in$ { sets }, $(A \cap B = \{ x : x \in A \ \wedge \ x \in B \})$

C. Given $a, n \in \omega \ni a \leq n \ \wedge S_i \in \{ \text{sets} \} \ \forall\, i \in \{ a, a+1, a+2, ..., n \})$,

$$\bigcup_{i=a}^{n} S_i = S_a \cup S_{a+1} \cup \ S_{a+2} \cup ... \cup S_n \ \ \wedge \ \ \bigcap_{i=a}^{n} S_i = S_a \cap S_{a+1} \cap \ S_{a+2} \cap ... \cap S_n$$

D. Given $A, B \in$ { sets }, note that "$A - B$" is read "A without B."

Definition for *without*:

Given $A, B \in$ { sets }, $(A - B = \{ x : x \in A \ \wedge \ x \notin B \})$

E. Given $A, B \in$ { sets }, note that "$A \times B$" is read "the cross product of A to B" or "the Cartesian product of A to B."

Definition for *cross product*:

Given $A, B \in$ { sets }, $(A \times B = \{ (x, y) : x \in A \wedge y \in B \})$

F. Given $A \in$ { sets }, note that "A^c " is read "the complement A" or "A complement."

Definition for *complement* of a set:

Given $A \in$ {sets} $\wedge$ V is the universe, $A^c = \mathrm{V} - A$

Algebra and Real Analysis

012. *Relation* from one set to another or from a set to itself:

A. Definition of a *relation from A to B*:

Given $A, B \in$ { sets }, (r is a *relation* from A to $B \Leftrightarrow r \subseteq A \times B$)

B. Definition of a *relation on a set*:

Given $A \in$ { sets }, (r is a *relation on A* $\Leftrightarrow r \subseteq A \times A$)

013. Relations that are *functions*:

A. Given $A, B \in$ {sets}, "$f: A \to B$" is read "f is a *function* from A to B"

Definition of *function*:

Given $A, B \in$ {sets}, ($f: A \to B \Leftrightarrow$ (r is a relation from A to B) $\wedge$
($\forall x \in A, \exists! y \in B \ni (x, y) \in f$))

B. Note on related terminology: "domain," "codomain," and "range":

Given $f : A \to B$, A *is* the *domain* of f $\wedge$ B is the *codomain* of f. Also, the *range* of f is defined as follows:

The *range* of f = { $y \in B : (x, y) \in f$ }

C. Note: "$f : A \xrightarrow{1:1} B$" is read "$f$ is a one-to-one function from A to B." It's also read, "f is an *injection* from A to B."

Definition for *one-to-one function* (i.e., *injection* or *injective function*):

Given $A, B \in$ {sets}, ($f : A \xrightarrow{1:1} B \Leftrightarrow f : A \to B \ni$($(x_1, y_1), (x_2, y_1) \in f \Rightarrow x_1 = x_2$))

D. Note: "$f : A \xrightarrow{onto} B$"is read "$f$ is an onto function from A to B." It's also read, "f is a *surjection* from A to B."

Definition for *onto function* (i.e., *surjection* or *surjective function*):

Given $A, B \in$ {sets}, ($f : A \xrightarrow{onto} B \Leftrightarrow f : A \to B \ni$ the *range* of f = B)

E. Note: "$f : A \xrightarrow[onto]{1:1} B$" or "$f : A \xrightarrow[1:1]{onto} B$" is read "*f* is a one-to-one and onto function from *A* to *B*." It's also read, "*f* is a *bijection* from *A* to *B*."

Definition for *bijection*:

Given $A, B \in$ {sets}, $(f : A \xrightarrow[onto]{1:1} B \Leftrightarrow f : A \xrightarrow{onto} B \wedge f : A \xrightarrow{1:1} B)$

014. Set equivalence and cardinality:

A. Definition for *one-to-one correspondence* between sets:

Given $A, B \in$ {sets}, there is a *one-to-one correspondence* between *A* and *B* $\Leftrightarrow$
$\exists f \ni f : A \xrightarrow[onto]{1:1} B$

B. Given $A \in$ {sets}, "$|A|$" is read "the *cardinality* of *A*."

Definition for *sets having the same cardinality*:

Given $A, B \in$ {sets}, $|A| = |B| \Leftrightarrow$
There is a *one-to-one correspondence* between *A* and *B*

Given $A, B \in$ { sets }, "$A \sim B$" is read "*A* is *equivalent* to *B*"

Definition for *equivalent* sets:

Given $A, B \in$ { sets }, $A \sim B \Leftrightarrow |A| = |B|$

015. Definitions for *constant* and *variable*:

A. A *constant* is a unique entity.

B. A *variable* is an unspecified element of a set that has more than one element.

016. Definition for *binary operation* on a set:

Given $A \in$ {sets}, (★ is a *binary operation* on $A \Leftrightarrow$ ★ $: A \times A \rightarrow A$)

017. Definition for *a partition of a set*:

Given $A, P \in$ {sets} $\ni (P = \{P_1, P_2, P_3, ..., P_k\} \wedge |P| = k)$,

$(P \in \{ \textit{partitions} \text{ of } A \} \Leftrightarrow (A = \bigcup_{i=1}^{k} P_i \wedge (P_i \cap P_j = \varnothing \; \forall \, i, j \in \{1, 2, ..., k\} \text{ where } i \neq j)))$

018. The eerie world of transfinite arithmetic:

A. Definition of *infinite set* that Georg Cantor (1845–1918) formulated:

$$S \in \{\text{infinite sets}\} \Leftrightarrow S \in \{\text{sets}\} \ni \exists\, T \in \{\text{sets}\} \ni (\, T \subset S \ \wedge\ T \sim S\,)$$

B. Definition of *finite set*:

$$S \in \{\text{finite sets}\} \Leftrightarrow (\, S \in \{\text{sets}\} \wedge\ S \notin \{\text{infinite sets}\})$$

C. Axiom w/r *cardinality of finite sets*:

$$S \in \{\text{finite sets}\} \Rightarrow |S| \in \ \omega$$

D. Note: "$\aleph_o$" is read "aleph-naught." And $\aleph_o$ is defined as follows:

Given $A \in \{\text{ sets }\}$, $(\ |A| = \ \aleph_o \Leftrightarrow A \sim \mathbb{N}\ )$

E. Definition of a *countably infinite set* (i.e., *countable set*):

$|A| = \ \aleph_o \Leftrightarrow A \in \{$ countably infinite sets $\}$)

F. Note: Some infinite sets (e.g., $\mathbb{I}$ and $\mathbb{R}$) are not countably infinite. If an infinite set is not countably infinite it is an element of { uncountable sets }. Definition of an *uncountably infinite set*:

Given $A \in \{$ infinite sets $\}$, (A is uncountably infinite $\Leftrightarrow$ A is not countably infinite).

Note: "A is uncountably infinite" is expressed "$|\ A\ | = \mathcal{C}$" and "$\mathcal{C}$" is referred to as "*the continuum*."

019. Definitions for *sequence* and *string*:

A. Definition for sequence s:

Given $A \in \{\text{sets}\}$, (s is a sequence $\Leftrightarrow s : \mathbb{N} \rightarrow A$)

Note: $\left(s(i)\right)_{i=1}^{\infty} = \left\{(1, s(1)), (2, s(2)), (3, s(3)), \ldots\right\}$

B. Definition for string t:

Given $A \in \{\text{sets}\}$ and $B = \{\ 1, 2, 3, \ldots, n\ \} \ni n \in \mathbb{N}$, (t is a string $\Leftrightarrow t : B \rightarrow A$)

020. Notations associated with sequences and strings:

A. "$\left(f(i)\right)_{i=1}^{\infty}$" is a notation for the sequence {(1, f(1)), (2, f(2)), (3, f(3)), ... } as is "(f(1)), f(2), f(3), ...)."

B. "$\left(f(i)\right)_{i=1}^{n}$" is a notation for the string {(1, f(1)), (2, f(2)), (3, f(3)), ..., (n, f(n)) } as is "(f(1)), f(2), f(3), ..., f(n))."

021. Definitions w/r monotonic sequences:

A. Definition for *nondecreasing sequence* :

Given $s : \mathbb{N} \rightarrow \mathbb{R}$, ($s \in$ { nondecreasing sequences } $\Leftrightarrow$ $s(i) \leq s(i+1)\ \forall\, i \in \mathbb{N}$)

B. Definition for *nonincreasing sequence* :

Given $s : \mathbb{N} \rightarrow \mathbb{R}$, ($s \in$ { nonincreasing sequences } $\Leftrightarrow$ $s(i+1) \leq s(i)\ \forall\, i \in \mathbb{N}$)

C. Definition for *monotonic sequence* (i.e., the sequence is *monotone*):
Given $s : \mathbb{N} \rightarrow \mathbb{R}$, ($s \in$ { monotonic sequences } $\Leftrightarrow$
$s \in$ { nonincreasing sequences } $\cup$ { nondecreasing sequences })

022. Definition and notations w/r convergent and divergent sequences:

A. Definition for a sequence *converging* to a real number:

Given $s : \mathbb{N} \rightarrow \mathbb{R}$, (s converges to the real number l $\Leftrightarrow$
$\forall\, \varepsilon \in (0, \infty), \exists\, m \in \mathbb{R} \ni (i > m \Rightarrow |\, s(i) - l \,| < \varepsilon)$)

Note: Given $s : \mathbb{N} \rightarrow \mathbb{R}$, note that the following statement is equivalent to stating that s converges to the real number l :

$$\lim_{i\to\infty} s(i) = l$$

B. Definition for *convergent* sequence:

Given $s : \mathbb{N} \rightarrow \mathbb{R}$, ($s \in$ { convergent sequences } $\Leftrightarrow \exists\, l \in \mathbb{R} \ni \lim_{i\to\infty} s(i) = l$

C. Definition for *divergent* sequence:

Given $s : \mathbb{N} \rightarrow \mathbb{R}$, ($s \in$ { divergent sequences } $\Leftrightarrow s \notin$ { convergent sequences }
(i.e., Given $s : \mathbb{N} \rightarrow \mathbb{R}$, ($s \in$ { divergent sequences } $\Leftrightarrow \nexists\, l \in \mathbb{R} \ni \lim_{i\to\infty} s(i) = l$)

023. Definition and notations w/r convergent and divergent sequences:

A. Definition for a sequence *converging* to a real number:

Given $s : \mathbb{N} \to \mathbb{R}$, (s converges to the real number $l \Leftrightarrow$
$\forall\, \varepsilon \in (0, \infty), \exists\, m \in \mathbb{R} \ni (i > m \Rightarrow | s(i) - l | < \varepsilon))$

Note: Given $s : \mathbb{N} \to \mathbb{R}$, note that the following statement is equivalent to stating that s converges to the real number l :

$$\lim_{i\to\infty} s(i) = l$$

B. Definition for *convergent* sequence:

Given $s : \mathbb{N} \to \mathbb{R}$, ($s \in$ { convergent sequences } $\Leftrightarrow \exists\, l \in \mathbb{R} \ni \lim_{i\to\infty} s(i) = l$

C. Definition for *divergent* sequence:

Given $s : \mathbb{N} \to \mathbb{R}$, ($s \in$ { divergent sequences } $\Leftrightarrow s \notin$ { convergent sequences }
(i.e., Given $s : \mathbb{N} \to \mathbb{R}$, ($s \in$ { divergent sequences } $\Leftrightarrow \nexists\, l \in \mathbb{R} \ni \lim_{i\to\infty} s(i) = l$)

024. Definitions, notes, and theorems w/r sequences that diverge to $+\infty$ or $-\infty$:

A. Definitions:

i. Given $s : \mathbb{N} \to \mathbb{R}$, (s diverges to $+\infty \Leftrightarrow$
$\forall\, u \in (0, \infty), \exists\, m \in \mathbb{R} \ni (i > m \Rightarrow s(i) > u))$

ii. Given $s : \mathbb{N} \to \mathbb{R}$, (s diverges to $-\infty \Leftrightarrow$
$\forall\, u \in (-\infty, 0), \exists\, m \in \mathbb{R} \ni (i > m \Rightarrow s(i) < u))$

B. Note:

{{ convergent sequences }, { sequences that diverge to $+\infty$ }, { sequences that diverge to $-\infty$}, { divergent sequences that do not diverge to $+\infty$ or to $-\infty$ }} $\in$
{ partitions of { sequences }}

C. Note: Given $s : \mathbb{N} \to \mathbb{R}$, ("$s$ diverges to $+\infty$" may be written "$\lim_{i\to\infty} s(i) = +\infty$" and

"s diverges to $-\infty$" may be written "$\lim_{i\to\infty} s(i) = -\infty$")

025. *Cauchy sequences*:

A. Definition for *Cauchy sequence*: Given $s : \mathbb{N} \to \mathbb{R}$, ($s \in$ { Cauchy sequences } $\Leftrightarrow$ $\forall\, \varepsilon \in (0, \infty), \exists\, m \in \mathbb{R} \ni ((\, i > m \wedge j > m\,) \Rightarrow |\, s(i) - s(j)\,| < \varepsilon\,))$

B. Theorem 0: { convergent sequences } = { Cauchy sequences }

026. Definitions and a note related to *limit of a function as domain values approach a real number*:

A. For two-sided limits:

Given $f: D \to \mathbb{R} \wedge D \subseteq \mathbb{R} \wedge l, x_o \in \mathbb{R}$,
$(\lim\limits_{x \to x_0} f(x) = l \Leftrightarrow \forall\, \varepsilon \in (0, \infty), \exists\, \delta \in (0, \infty) \ni |\, x - x_o\,| < \delta \Rightarrow |\, f(x) - l\,| < \varepsilon\,)$

B. For left-sided limits:

Given $f: D \to \mathbb{R} \wedge D \subseteq \mathbb{R} \wedge l, x_o \in \mathbb{R}$,
$(\lim\limits_{x \to x_0^-} f(x) = l \Leftrightarrow \forall\, \varepsilon \in (0, \infty), \exists\, \delta \in (0, \infty) \ni x_o - x < \delta \Rightarrow |\, f(x) - l\,| < \varepsilon\,)$

C. For right-sided limits:

Given $f: D \to \mathbb{R} \wedge D \subseteq \mathbb{R} \wedge l, x_o \in \mathbb{R}$,
$(\lim\limits_{x \to x_0+} f(x) = l \Leftrightarrow \forall\, \varepsilon \in (0, \infty), \exists\, \delta \in (0, \infty) \ni x - x_o < \delta \Rightarrow |\, f(x) - l\,| < \varepsilon\,)$

027. Definition for *continuous function* at a point:

Given $x_0, x_1 \in \mathbb{R} \ni (\, x_0 < x_1 \wedge f : [\, x_0, x_1\,] \to \mathbb{R} \wedge c \in (\, x_0, x_1\,)\,)$,
(f is continuous at $c \Leftrightarrow \lim\limits_{x \to c} f(x) = f(c)\,)$

028. Definitions for limits of functions as domain values either get large without limit or small without limit:

A. Given $f: D \to \mathbb{R} \wedge D \subseteq \mathbb{R} \wedge l \in \mathbb{R}$,
$(\lim\limits_{x \to +\infty} f(x) = l \Leftrightarrow \forall\, \varepsilon \in (0, \infty), \exists\, m \in \mathbb{R} \ni x > m \Rightarrow |\, f(x) - l\,| < \varepsilon\,)$

B. Given $f: D \to \mathbb{R} \wedge D \subseteq \mathbb{R} \wedge l \in \mathbb{R}$,
$(\lim\limits_{x \to -\infty} f(x) = l \Leftrightarrow \forall\, \varepsilon \in (0, \infty), \exists\, m \in \mathbb{R} \ni x < m \Rightarrow |\, f(x) - l\,| < \varepsilon\,)$

Probability w/r Discrete Random Variables

029. Experiments, simulations, outcomes, random outcomes, sample spaces, events, and complements of events:

A. Experiment (undefined term)

B. A less-than rigorous definition for *simulation*:

A simulation is a procedure that facilitates addressing questions about real problems by running experiments that are algebraically equivalent to the real problem.

C. Outcome (undefined term)

D. Note and definition for *sample space*:

Note: “Ω” is often used as a symbol to denote a sample space.

Definition for sample space Ω:

$\Omega = \{ \text{ outcomes of an experiment } \}$

E. Definition for *event*:

$A \in \{ \text{ events of } \Omega \} \Leftrightarrow A \subseteq \Omega$

F. Definition and note for *complement of an event*:

Note: “A^c ” is read “the complement of event A”

Definition:

$A^c = \{ X : X \subseteq \Omega \wedge X \cap A = \varnothing \}$

030. Probability Measures (i.e., probability distributions or probability functions)

Definition for *probability measure*:

Given Ω is a sample space $\wedge E = \{ \text{ events of } \Omega \}$, ($p \in \{ \text{ probability measures on } \Omega \}$
$\Leftrightarrow$
$p : E \to [\, 0, 1 \,] \ni (p(\Omega) = 1 \wedge (A_1 \subseteq E \wedge A_2 \subseteq E \wedge$
$A_1 \cap A_2 = \varnothing) \Rightarrow p(A_1 \cup A_2)) = p(A_1) + p(A_2)))$

031. Four theorems w/r $p \ni (p \in \{$ probability measures on Ω where $| \Omega | \in \omega \cup \{ \aleph_o \}$:

A. Theorem 01: $p(A^c) = 1 - p(A)$

B. Theorem 02: $p(\varnothing) = 0$

C. Theorem 03: $A \subseteq B \Rightarrow p(A) \leq p(B)$

D. Theorem 04: $p(A \cup B) = p(A) + p(B) - p(A \cap B)$

032. Random outcome:

A. Definition for *random outcome*: : The outcomes of Ω are *random* $\Leftrightarrow$ $(p \in \{$ probability measures on $\Omega \} \wedge (p(\{ x \}) = p(\{ y \}) \; \forall \; x, y \in \Omega))$

B. Note: Unless specified otherwise for a particular entry in this Glossary, subsequent references to Ω are meant to designate a sample space containing only random outcomes.

033. A taste of psychometrics:

A. Measurements

Definition for *measurement*: A measurement is a process by which data or information are collected via empirical observations.

Note: A measurement is an objective process focusing on a quantitative variable.

B. Assessments

S

Definition for *assessment*: An assessment is a judgment or evaluation drawn from inferences that extend beyond what has been empirically observed.

Note: An assessment is a subjective process focusing on a qualitative variable.

C. The relations between a measurement and an assessment:

Note: Assessments are influenced by data (i.e., measurement results) but accurate assessments also depend on the sagaciousness and ethics of the people who conduct the assessments.

D. Discrete data strings and data sequences:

Definition for *discrete data*:

$m \in \{$ discrete data $\} \Leftrightarrow ((m \in \{$ strings $\} \veebar m \in \{$ sequences $\}) \wedge$ the codomain of m is a subset of $\mathbb{R}$).

E. Measurement usefulness:

i. Note: “$D_0 = D_t + D_E$” is read “data from a measurement equals the sum of what the data would be if the measurement were perfectly valid and the influence of measurement error.”

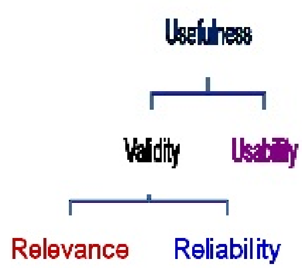
ii. Definition for *measurement relevance*: A measurement is *relevant* to the same degree that the data it generates are pertinent to the assessment that is influenced by those data.

iii. Definition for *measurement reliability*: A measurement is *reliable* to the same degree that it generates data that are internally consistent (i.e., the data reflects a non-contradictory pattern).

iv. Definition for *measurement validity*: A measurement is *valid* to the same degree that it both relevant and reliable.

v. Definition for *measurement usability*: A measurement is *usable* to the degree that it is practical to administer.

v. Definition for *measurement usefulness*: A measurement is *useful* to the same degree that it both valid and useable.



034. A taste of Counting:

A. Multiplication principle theorem:

Theorem 05: Independent experiments #1 and #2 with respective finite sample spaces Ω_1 and Ω_2 are conducted $\Rightarrow |\ \Omega_1 \cap\ \Omega_2\ | = |\ \Omega_1\ | \cdot |\ \Omega_2\ |$

B. Definition of a *permutation* of a finite set:

Given $A \in \{$ finite sets $\}$, $(f \in \{$ permutations of $A\ \} \Leftrightarrow f : A \underset{onto}{\overset{1:1}{\longrightarrow}} A\ )$

C. Note: Given $n, r \in \omega \ni r \le n$, "${}_nP_r$" is read "the number of all possible permutations of n elements taken r at a time.

D. Theorem 06: ${}_nP_r = \dfrac{n!}{(n-r)!}$

E. Definition of a *combination* on a finite set:

Given $A \in \{\text{ finite sets }\} \wedge\ n, r \in \omega\ \wedge r \le n \wedge |A| = n$,
((A combination of r on A) $= B \Leftrightarrow\ B \subseteq A \ni |B| = r$)

F. Note: Given $n, r \in \omega \ni r \le n$, "${}_nC_r$" or "$\binom{n}{r}$" is read "the number of all possible combinations of n elements taken r at a time.

G. Theorem 07: ${}_nC_r = \dfrac{n!}{r!(n-r)!}$

035. Theorem 08 (Binomial Theorem):

Given $c, d \in \mathbb{R}\ \wedge t \in \omega$, ($(c+d)^t = \sum_{i=0}^{t}\left[\binom{t}{i}c^{t-i}d^i\right]$)

036. Conditional Probability:

A. Note: Given $A \subseteq \Omega \wedge B \subseteq \Omega \ni p(B) \neq 0$, "$p(A|B)$" is read "the conditional probability of A given that B."

B. Definition for *conditional probability* : Given $A \subseteq \Omega \wedge B \subseteq \Omega \ni p(B) \neq 0$,
($p(A|B) = \dfrac{p(A \cap B)}{p(B)}$)

C. Theorem 09 (the little version of Bayes' Theorem):

Given $A \subseteq \Omega \wedge B \subseteq \Omega \ni p(B) \neq 0$, ($p(B|A) = \dfrac{p(A|B)p(B)}{p(\mathrm{A})}$)

D. Theorem 09 (the expanded version of Bayes' Theorem):

($A, B_1,\ B_1, B_1, \ldots B_n, \in \{\text{ events of }\ \Omega\} \wedge B_i \cap B_j = \varnothing\ \forall\, i, j \in \{1, 2, 3, \ldots, n\} \ni i \neq j\ \wedge$

$\bigcup_{i=1}^{n} B_i = \Omega\ \wedge p(B_i) > 0\ \ \forall\, i \in \{1, 2, 3, \ldots, n\}) \Rightarrow P(B_j|A) = \dfrac{P(A|B_j)P(B_j)}{\sum_i^n P(A|B_i)P(B_i)}$ }

037. Independent, dependent, and mutually-exclusive events:

A. Definition for *independent events*:

Given $A \subseteq \Omega \wedge B \subseteq \Omega$, (A and B are independent of one another $\Leftrightarrow$ $p(A \cap B) = p(A)\cdot p(B)$)

B. Definition for *dependent events*:

Given $A \subseteq \Omega \wedge B \subseteq \Omega$, (A and B are dependent of one another $\Leftrightarrow$ $p(A \cap B) \neq p(A)\cdot p(B)$)

C. Theorem 10: Given $A \subseteq \Omega \wedge B \subseteq \Omega$, ($p(A \cap B) \neq p(A)\cdot p(B) \Rightarrow$ $p(A \cap B) = p(A)\cdot p(B|A)$)

D. Definition of *mutually exclusive events*:

Given $A \subseteq \Omega \wedge B \subseteq \Omega$, (A and B are mutually-exclusive relative to one another $\Leftrightarrow$ $p(A \cap B) = 0$)

038. Discrete Random Variables:

Definition for *discrete random variable* : $X \in$ { discrete random variables of Ω } $\Leftrightarrow$ ($| \Omega |, | X | \in \{ \aleph_o, 0, 1, 2, 3, ... \} \wedge E =$ { events of Ω } $\wedge X : E \to \mathbb{R}$)

039. Discrete probability distribution and discrete uniform probability uniform distribution:

A. Definition for *discrete probability distribution* :

Given $X \in$ { discrete random variables of Ω },
($m \in$ { discrete probability distribution for X } $\Leftrightarrow$ $m : X \to [0, 1] \wedge \sum_{x \in E} m(x) = 1$)

B. Definition for discrete *uniform probability distributions* :

Given $m \in$ { discrete distribution functions for X on Ω } $\wedge | \Omega | = n \ni n \in \mathbb{N}$,
($m \in$ { discrete uniform probability distribution functions for X on Ω } $\Leftrightarrow$
$m(x) = \dfrac{1}{n} \ \forall x \in X$

040. The relationship between probabilities and odds:

Definition for the odds of an event occurring:

Given E is an event of Ω , (the odds in favor of E occurring =
$r : s$ where $r, s \in [0, \infty) \Leftrightarrow \dfrac{r}{s} = \dfrac{p(E)}{1 - p(E)}$)

041. Expected values, variances, and standard deviation of discrete random variables:

A. Definition for *expected value* for a discrete random variable X:

Given $X \in$ { discrete random variables of Ω } $\wedge$ ($m \in$ { discrete probability distribution for X }, ($E(X)$ is the *expected value* of $X \Leftrightarrow E(X) = \sum_{x\in\Omega} xm(x)$ provided that the series converges absolutely).

Note: If the series does not converge absolutely, then X does not have an expected value.

B. Definition for *variance* for a discrete random variable X:

Given $X \in$ { discrete random variables of Ω } $\wedge$ the expected value of X is $E(X)$, ($V(X)$ = *variance of* $X \Leftrightarrow V(X) = E((X - E(X))^2)$)

C. Definition for *standard deviation* for a discrete random variable X:

Given $X \in$ { discrete random variables of Ω } $\wedge$ the variance of X is $V(X)$, ($D(X)$ = *standard deviation of* $X \Leftrightarrow D(X) = \sqrt{V(X)}$

042. Bernoulli random variable :

Definition for a *Bernoulli random variable*:

$X \in$ { Bernoulli random variables of Ω } $\Leftrightarrow$
$X \in$ { discrete random variables of Ω } $\wedge$ the range of X = { 0, 1 }

043. Binomial experiments and binomial probability distributions:

A. Definition for a binomial random variable: :

Given $n \in \mathbb{N} \wedge \Omega = \{ 1, 2, 3, ..., n \} \wedge X \in$ { Bernoulli random variables of { 0, 1}} $\wedge$ (A string of n experiments are conducted with $X \ni (X(i) = 0 \veebar (X(i) = 1$ depending on the results of the i^{th} experiment $\wedge |\{ (i, X(i)) : X(i) = 1 \}| = k$), ($Y \in$ { binomial random variables of Ω } $\Leftrightarrow$

$Y \in$ { discrete random variables of Ω } $\ni Y(i) = \sum_{i}^{n} X(i) = k$

B. Theorem 11: ($X \in$ { Bernoulli random variables of { 0, 1 }} $\wedge\, m : X \to [0, 1] \wedge$ ($\Omega = \{ 0, 1, 2, ..., n \} \wedge Y \in$ { binomial random variables of Ω }) $\wedge\, p : Y \to [0, 1]) \Rightarrow$
$p(k) = \binom{n}{k} m(1)^k (1 - m(1))^{n-k}$

044. Notes w/r geometric, negative binomial, hypergeometric, and Poisson experiments and probability distributions:

A. Note on these Notes #044B–E: Geometric, negative binomial, hypergeometric, and Poisson random variables and their respective probability distributions, like a myriad of other random variables and probability functions are worthy of our attention. These particular four (random variable, probability distributions) ordered pairs are applicable to addressing important real-life problems and for our purposes of Math 5710 they are vital in our quest toward the *Central Limit Theorem*. However, We won't attend to them as we'd like because of course time constraints but we will become aware of their existence for future reference and to help us develop a better feel why Glossary Entry #49 is what it is.

B. Note on *geometric* experiments and related discrete probability distributions:

John R. Rice on p. 40 of *Mathematical Statistics and Data Analysis* (3rd. Ed.), begins the following statement by comparing geometric experiments to binomial experiments: " The **geometric distribution** is also constructed from independent Bernoulli trials but from an infinite sequence. On each trial, a success occurs with probability p, and X is the total number of trials up to and including the first success. So that $X = k$, there must be $k - 1$ failures followed by a success. From the independence of the trials, this occurs with probability

$$P(k) = P(X = k) = (1 - p)^{k-1}p, \qquad k = 1, 2, 3, \ldots$$

Note that these probabilities sum to 1:

$$\sum_{k=1}^{\infty}(1-p)^{k-1}p = p\sum_{j=0}^{\infty}(1-p)^{j} = 1\text{"}$$

Further note that Rice's use of the notations "P" and "p", although a popular convention, differs from our usage. We will clarify during one of our class meetings.

C. Note on *negative binomial* experiments and related discrete probability distributions:

John R. Rice on p. 41 of *Mathematical Statistics and Data Analysis* (3rd. Ed.) states that, "The **negative binomial distribution** arises as a generalization of the geometric distribution. Suppose that a sequence of independent trials, each with probability of success p, is performed until there are r successes in all; let X denote the total number of trials. To find $P(X = k)$, we can argue in the following way: Any such sequence has probability $p^r(1-p)^{k-r}$, from the independence assumption. The last trial is a success, and the remaining $r - 1$ successes can be assigned to the remaining $k - 1$ trials in $\binom{k-1}{r-1}$ ways. Thus, $P(X = k) = \binom{k-1}{r-1}p^r(1-p)^{k-r}$."

D. Note on *hypergeometric* experiments and related discrete probability distributions:

A hypergeometric experiment is quite similar to a binomial experiment but with one crucial exception: In a binomial experiment, the selected events are independent from one another; whereas in a hypergeometric experiment the selected events are dependent on one another because events are selected one at a time without replacement. Consider the following example of hypergeometric experiment:

> Five cards are randomly selected from a standard 52-card poker deck and this is done *without replacement*. The goal of the experiment is to determine the probability that exactly two of the selected cards are red.

Now consider a similar experiment that is binomial rather than hypergeometric:

> Five cards are randomly selected one at a time from a standard 52-card poker deck; after the first card is selected, its color is recorded and returned to the deck, and the deck is reshuffled. The same algorithm is repeated until five cards have been drawn. Thus, this is done *with replacement*. The goal of the experiment is to determine the probability that exactly two of the selected cards are red.

For a hypergeometric random variable X, the formula for the discrete probability function p can be expressed as follows where N is the number of elements in the population (e.g., 52), k = the number of successful events in the population (e.g., the number of possible events in which exactly two cards are red), n = the number of element in each event (e.g., 5) , and x is the number of successes in the random sample (e.g., 2):

$$p(\, X = x\,) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}$$

E. Note on *Poisson* experiments and related discrete probability distributions:

Binomial random variables are applicable to a myriad of real-world problems. Often we need to calculate $p(\, k\,)$ with parameter specified by Glossary line 043A in the following formula:

$$p(\, k\,) = \binom{n}{k} m(1)^k\,(\,1 - m(1))^{n-k}$$

However, in many cases this calculation is inordinately time-consuming (even for a computer) due the magnitude of n! Poisson random variables and their discrete probability distributions provide approximations that can be within ε of the binomial probabilities (Yes, we are referring to ε from Glossary Lines 023A and 026A). The formula for a discrete probability distribution for a Poisson random variable X is as follows with λ a constant positive real number whose value is selected based on parameters of the experiment including the number of trials:

$$p(X-k)=\frac{\lambda^k}{k!}e^{-k} \text{ where } k \in \omega$$

Probability w/r Continuous Random Variables

045. Definition for *continuous random variable* :

Definition for *continuous random variable* : $X \in$ { continuous random variables of Ω } $\Leftrightarrow$ ($|\,\Omega\,|, |\,X\,| \in \{ \mathcal{C}^n : n \in \mathbb{N} \} \wedge E =$ { events of Ω } $\wedge X : E \to \mathbb{R}$)

046. Definitions for *density function of continuos random variables* and *cumulative density function of continuos random variables* and a theorem:

A. Given $X \in$ { continuous random variable },

($f \in$ { density functions of X } $\Leftrightarrow f : \mathbb{R} \to \mathbb{R} \ni p(\,a \le X \le b\,) = \int_a^b f(x)dx \;\forall\, a, b \in \mathbb{R}$)

B. Definition for cumulative distribution function of continuous random variables:
Given $X \in$ { continuous random variable },
($F_X \in$ { cumulative distribution function of X } $\Leftrightarrow F_X : \mathbb{R} \to \mathbb{R} \ni p(\,X \le x\,)$))

C. Theorem 12:

($X \in$ { continuous random variable } with density function f) $\Rightarrow$

(The cumulative distribution function of $X = F \ni (F(x) = \int_{-\infty}^{x} f(t)dt$)

$\wedge \dfrac{d}{dx}F(x) = f(x)$)

047. Expected values and variances of continuous random variables:

A. Definition for *expected value* for a continuous random variable X:

Given $X \in$ { continuous random variables of Ω } $\wedge$ ($f \in$ { density function for X },

($E(X)$ is the *expected value* of $X \Leftrightarrow E(X) = \int_{-\infty}^{\infty} xf(x)dx$ provided that the definite integral $\int_{-\infty}^{\infty} |x|\, f(x)dx$ }exists.

B. Definition for *variance* for a continuous random variable X:

Given $X \in$ { continuous random variables of Ω } $\wedge$ ($f \in$ { density function for X },
($V(X)$ = *variance of* $X \Leftrightarrow V(X) = E((X - E(X))^2)$)

C. Theorem 13:

Given $X \in$ { continuous random variables of Ω } $\wedge$ ($f \in$ { density function for X },
($E((X - E(X))^2)) = E(X^2) - (E(X))^2$)

048. Theorem 14 (the Principle of Large Numbers):

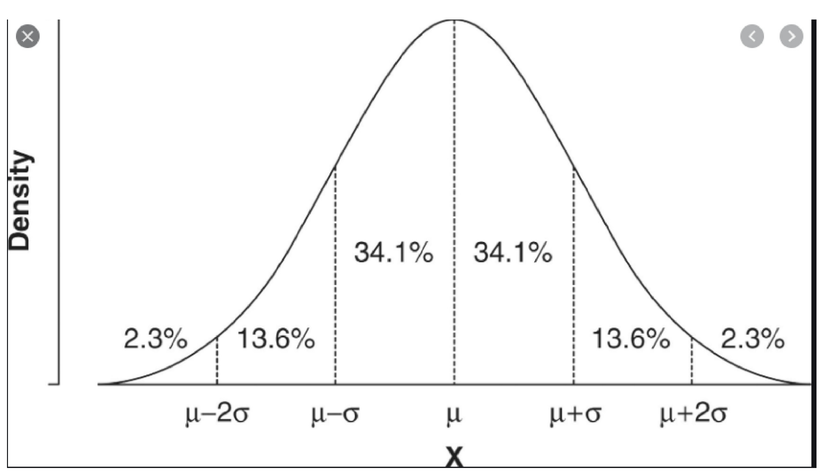
(($n \in \mathbb{N} \wedge X_1, X_2, X_3, \ldots, X_n \in$ { independent trials process with continuous function p } $\wedge$
($\mu \in \mathbb{R} \ni \mu$ is the expected value of $X_i \ \forall\, i \in \{1, 2, 3, \ldots, n\}$) $\wedge$
($\sigma \in [0, \infty) \ni \sigma^2$ is the expected variance of $X_i \ \forall\, i \in \{1, 2, 3, \ldots, n\}$)) $\Rightarrow$

$$\lim_{n\to\infty}\left(p\left(\left|\frac{\sum_{i=1}^{n} X_i}{n} - \mu\right|\right)\right) = 0$$

049. The family of *normal* (i.e, *Gaussian*) probability density functions:

A. Definition for normal probability density function: $f \in$ { normal probability density functions }
$\Leftrightarrow$ ($f \in$ { probability density functions } $\wedge$ ($f : \mathbb{R} \to \mathbb{R} \ni$ ($f(x)$ =
$\frac{1}{\sigma\sqrt{2\pi}} \cdot (e^{-(x-\mu^2)/2\sigma^2}) \ni$
(μ is the expected value of $f(x) \wedge \sigma^2$ is the expected variance of $f(x)$)).

B. Note that the following graph of a normal probability density function, depicts some of the features of a the function that is particularly useful in the world of inferential statistics:



050: Theorem 15 (the Central Limit Theorem):

(($n \in \mathbb{N} \wedge X_1,\ X_2,\ X_3, ...,\ X_n \in$ { independent trials process with continuous function p } $\wedge$ ($\mu \in \mathbb{R} \ni \mu$ is the expected value of $X_i\ \ \forall\, i \in$ { 1, 2, 3, ..., n }) $\wedge$ ($\sigma \in [\,0, \infty\,) \ni \sigma^2$ is the expected variance of $X_i\ \ \forall\, i \in$ { 1, 2, 3, ..., n })) $\Rightarrow$ ($\exists\ f\ \ni$ ($f \in$ { normal probability density functions } $\ni$ (μ is the expected value of $f\,(\,x\,)\ \wedge \sigma^2$ is the expected variance of $f\,(\,x\,)\ \wedge$ (as n increases without limit (i.e, $n \to \infty$), p varies so it is a closer and closer approximation of f (i.e, $p \to f$).

051. Population means, population standard deviations, sample means, sample standard deviations, and z-scores:

A. Given $t \in$ { data strings resulting from measurements of entire populations }, note the following:

i. "μ_t" is read "the population mean of t."

ii. Definition for *population mean*: $\mu_t = \dfrac{\sum_{i=1}^{N} t(i)}{N}$ where $N = |\ t\ |$

iii. "σ_t" is read "the population standard deviation of t."

iv. Definition for *population standard deviation*: $\sigma_t = \sqrt{\dfrac{\sum_{i=1}^{N} (t(i) - \mu_t)^2}{N}}$ where $N = |\ t\ |$

v. "$z_{t(i)}$" is read "the z-score associated with $t(\ i\ )$."

vi. Definition for $z_{t(i)}$: $z_{t(i)} = \dfrac{t(i) - \mu_t}{\sigma_t}$

B. Given $t \in$ { data strings resulting from measurements of samples drawn from sample drawn from populations }, note the following:

i. "$\overline{x}_t$" is read "the sample mean of t."

ii. Definition for *sample mean*: $\overline{x}_t = \dfrac{\sum_{i=1}^{n} t(i)}{n}$ where $n = |\ t\ |$

iii. "s_t" is read "the sample standard deviation of t."

iv. Definition for *sample standard deviation*: $s_t = \sqrt{\dfrac{\sum_{i=1}^{n} (t(i) - \overline{x}_t)^2}{n-1}}$ where $n = |\ t\ |$

052. Population statistics and inferential statistics:

A. Definition for *population data*: Data gathered via measurements employed on an entire population of interest are population data.

B. Definition for *population parameter*: Population parameters are statistics (e.g., μ, σ, ρ) that exists with respect to population data.

C. Definition for *sample data*: Data gathered via measurements employed on a sample randomly drawn from a population of interest are sample data.

D. Definition for *inferential statistics*: Inferential statistics are statistics (e.g., $\bar{x}$, s, r) computed from sample data for the purpose of assessing hypotheses with respect of population parameters.

053. Null hypotheses, type I error, type II error, and statistical power:

A. Definition of a *null hypothesis*: A null hypothesis (symbolized "H_o") is a supposition about the value of a population parameter.

B: Note: In the world of inferential statistics, the supposition that H_o is true establishes a probability density function (e.g., a normal distribution) to help assess whether or not H_o (e.g., $\mu_t = \mu_{\bar{x}}$) should be rejected. The underlying logic is somewhat similar to that of mathematical proofs by contradiction; however, instead of deducing an absolute logical contradiction, the probability about the truth value of H_o is computed.

C. Note: A *type I* error occurs whenever a true null hypothesis is rejected. A *type II* error occurs whenever a false null hypothesis is not rejected.

D. Note: In an experiment that employs inferential statistics, statistical test *A* of a null hypothesis has greater *statistical power* than statistical test *B* iff the likelihood of a Type II is lower for *A* than it is for *B* (i.e., the likelihood of a Type I error using statistical test A is greater than the likelihood of a Type I error using statistical test *B*).

054. *Pearson product-moment correlation coefficient*:

$$\rho_{(a,b)} = \frac{1}{N}\sum_{i=1}^{N} z_{a_i} z_{b_i}$$