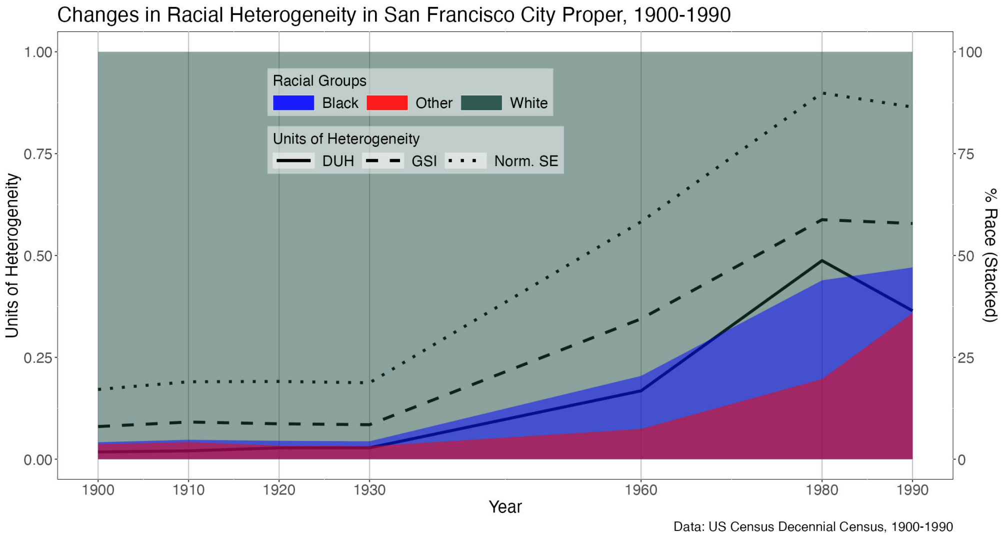
Descriptive Units of Heterogeneity: An Axiomatic Approach to Measuring Heterogeneity

Willy Chen

MSU Graduate Student Seminar

Empirical Motivation



What is Heterogeneity in a System?

Let $\Theta = \{\theta_1, \ldots, \theta_G\}$ be a **universe** of G distinct groups/categories. A system S is a mapping from Θ to $\mathbb{Z}_+^G$ such that $S = (n_1, \ldots, n_g \ldots, n_G)$ where n_g is the positive integer that represents the number of elements in θ_g in the system S.

For example, let S be the population of Michigan State University, S maps the groups faculty, staff, and students to the number of faculty, staff, and students.

What is Heterogeneity in a System?

Let $\Theta = \{\theta_1, \ldots, \theta_G\}$ be a **universe** of G distinct groups/categories. A system S is a mapping from Θ to $\mathbb{Z}_+^G$ such that $S = (n_1, \ldots, n_g \ldots, n_G)$ where n_g is the positive integer that represents the number of elements in θ_g in the system S.

For example, let S be the population of Michigan State University, S maps the groups faculty, staff, and students to the number of faculty, staff, and students.

Goal: For any two systems S_1 and S_2 with G groups, $\Phi : \mathbb{Z}_+^G \to \mathbb{R}$ measures the heterogeneity of S_1 and S_2 such that

$$\Phi(S_1) \geq \Phi(S_2) \iff S_1 \text{ is weakly more heterogeneous than } S_2.$$

Maximum and Minimum Heterogeneity

A system S_{max} with G groups is said to achieve **maximum heterogeneity** if it can be represented as a scalar multiple of the identity vector of size $G \in \mathbb{N}$:

$$S_{max} = (n, n, \ldots, n) = n \cdot (1, 1, \ldots, 1).$$

A system S with G groups is said to achieve **minimal heterogeneity** it can be presented as a $1 \times G$ vector where all but one entries are 0

$$S_{min} = (0, 0, \ldots, 0, n, 0, \ldots, 0) = n \cdot (0, 0, \ldots, 0, 1, 0, \ldots, 0).$$

Typically, $\Phi(S_{max}) = 1$ and $\Phi(S_{min}) = 0$.

Existing Measures of Heterogeneity

Two main types of units used to measure heterogeneity in a population:

- Dispersion units such as the Gini coefficient measure how well-dispersed a population is, compared to a benchmark distribution.
- Concentration units such as HHI and Entropy measure the proportion of a population concentrated in certain groups.

Issues with Current Units:

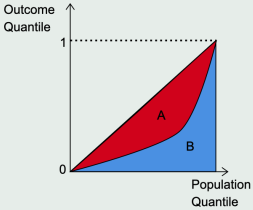
- Too much or too little restriction on the environment.
- Lack of comparability across systems with different number of groups.
- Concentration units do not reflect well changes in small groups.

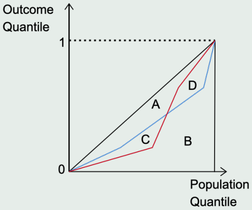
Solving these problems will make the unit “descriptive”.

Dispersion Units

Gini Coefficient

$$GI = \frac{A}{A + B}$$

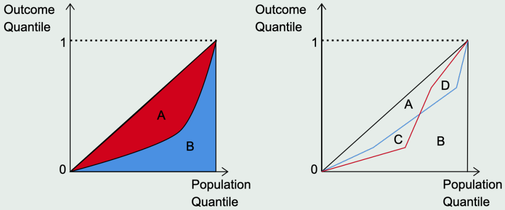




Dispersion Units

Gini Coefficient

$$GI = \frac{A}{A+B}$$



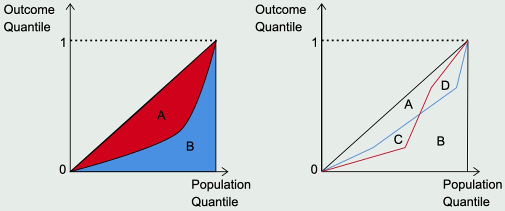
- Atkinson (1970) shows that when the Lorenz criterion is violated, functions can be chosen to rank the two distributions arbitrarily.

Dispersion Units

Gini Coefficient

$$GI = \frac{A}{A+B}$$

$$\frac{A+C}{A+B+C+D} \ vs. \ \frac{A+D}{A+B+C+D}$$



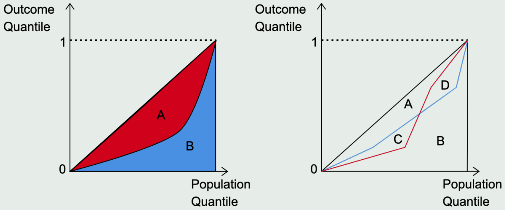
- Atkinson (1970) shows that when the Lorenz criterion is violated, functions can be chosen to rank the two distributions arbitrarily.
- Newbery (1970) shows that no linearly additive measure induces the same order as Gini.

Dispersion Units

Gini Coefficient

$$GI = \frac{A}{A+B}$$

$$\frac{A+C}{A+B+C+D} \; vs. \; \frac{A+D}{A+B+C+D}$$



- Atkinson (1970) shows that when the Lorenz criterion is violated, functions can be chosen to rank the two distributions arbitrarily.
- Newbery (1970) shows that no linearly additive measure induces the same order as Gini.
- Schwartz and Winship (1980) show that the Lorenz criterion is often neglected, leading to paradoxical inferences.

Concentration Units

Let $n_S = \sum_{g=1}^{G} n_g$,

HHI/GSI

$$HHI = \sum_{g=1}^{G} \left(\frac{n_g}{n_S}\right)^2, \ GSI = 1 - HHI$$

Concentration Units

Let $n_S = \sum_{g=1}^{G} n_g$,

HHI/GSI

$$HHI = \sum_{g=1}^{G} \left(\frac{n_g}{n_S}\right)^2, \; GSI = 1 - HHI$$

Shannon Entropy

$$SE = -\sum_{g=1}^{G} \left[\frac{n_g}{n_S} \cdot ln\left(\frac{n_g}{n_S}\right)\right]$$

Descriptive Units of Heterogeneity

Let $n_1 \geq n_2 > 0$, $P_1 = \frac{n_1}{n_S}$, and $\tilde{P}_g = \frac{n_g}{n_2 + \cdots + n_G}$. The Descriptive Units of Heterogeneity (DUH) is defined as:

$$DUH = \frac{ln(P_1) \cdot \left(\sqrt{\sum_{g=2}^{G} \left(\tilde{P}_g - \frac{1}{G-1} \right)^2} - 1 \right)}{ln(G)}$$

Fundamental Axioms

Axioms that induce partial ordering by pinning down

1. When two systems of G groups are equally heterogeneous
2. How to order the heterogeneity of two systems of G groups that are marginally different

[SYM] Group Symmetry

For any permutation $\pi(S)$ of S, $\Phi(S) = \Phi(\pi(S))$.

For example, take $n_a, n_b, n_c \in \mathbb{N}$,

G. \ S.	S_1		S_2		S_3
A	n_a		n_b		n_c
B	n_b	$\sim$	n_a	$\sim$	n_b
C	n_c		n_c		n_a

[INV] Scale Invariance

For any system S and a scalar $\lambda \in \mathbb{R}_{++}$, $\Phi(S) = \Phi(\lambda \cdot S)$.

For example, take $n_a, n_b, n_c, \in \mathbb{N}$, $\lambda \in \mathbb{R}_{++}$,

S. / G.	S_1		S_2		$S_1 + S_2$
A	n_a		λn_a		$(1+\lambda)n_a$
B	n_b	$\sim$	λn_b	$\sim$	$(1+\lambda)n_b$
C	n_c		λn_c		$(1+\lambda)n_c$

[PDT] Principle of Diminishing Transfers

Holding the order of groups constant, a transfer from a larger group to a smaller group increases heterogeneity. The increase increases in the difference between the two groups.

For example, take $n_a, n_b, n_c \in \mathbb{N}$ such that $n_a > n_b > n_c$ and $\varepsilon < \min\left\{n_b - n_c, \dfrac{n_a - n_b}{2}\right\}$.

G. \ S.	S_1		S_2		S_3
A	n_a		$n_a - \varepsilon$		$n_a - \varepsilon$
B	n_b	$\prec$	$n_b + \varepsilon$	$\prec$	n_b
C	n_c		n_c		$n_c + \varepsilon$

Building on the Fundamental Axioms

Existing units (satisfying *INV*) has the same arithmatic treatment for every group in the system, partially to satisfy *SYM*. However, said uniform treatment can be relaxed by treating the same type of groups the same.

Building on the Fundamental Axioms

Existing units (satisfying INV) has the same arithmatic treatment for every group in the system, partially to satisfy SYM. However, said uniform treatment can be relaxed by treating the same type of groups the same.

When comparing heterogeneity of groups, assuming $SYM+INV+PDT$ is equivalent to assuming Lorenz criterion, and Lorenz criterion only induces partial ordering between heterogeneity of systems.

Building on the Fundamental Axioms

Existing units (satisfying *INV*) has the same arithmatic treatment for every group in the system, partially to satisfy *SYM*. However, said uniform treatment can be relaxed by treating the same type of groups the same.

When comparing heterogeneity of groups, assuming $SYM + INV + PDT$ is equivalent to assuming Lorenz criterion, and Lorenz criterion only induces partial ordering between heterogeneity of systems.

$\Rightarrow$ To induce total ordering, *PDT* needs to be further refined.

What Matters in Measuring Heterogeneity?

The heterogeneity of system S is measured by $\Phi : S \to \mathbb{R}_+$ such that

$$\Phi(S_1) \geq \Phi(S_2) \iff S_1 \text{ is weakly more heterogeneous than } S_2.$$

What Matters in Measuring Heterogeneity?

The heterogeneity of system S is measured by $\Phi : S \to \mathbb{R}_+$ such that

$$\Phi(S_1) \geq \Phi(S_2) \iff S_1 \text{ is weakly more heterogeneous than } S_2.$$

- Relative size of the largest group: $P_1 = \frac{n_1}{n_1 + \cdots + n_G}$

What Matters in Measuring Heterogeneity?

The heterogeneity of system S is measured by $\Phi : S \to \mathbb{R}_+$ such that

$$\Phi(S_1) \geq \Phi(S_2) \iff S_1 \text{ is weakly more heterogeneous than } S_2.$$

- Relative size of the largest group: $P_1 = \frac{n_1}{n_1 + \cdots + n_G}$
- Relative size(s) of the minority group(s): $P_2, \ldots, P_G \left(= \frac{n_G}{n_1 + \cdots + n_G}\right)$

A unit of heterogeneity can be thought of as $\Phi = \Phi(\varphi, \psi)$ where
$\varphi(P_1)$ is the influence of the relative size of the largest group
$\psi(P_2, \ldots, P_G)$ the influence of the relative sizes of the minority group.

Characterization Axioms

Axioms that can be combined to uniquely characterize measures with total ordering and cardinal interpretation.

[IND] Independence

The influence of the relative size of the largest group on the unit of heterogeneity should be independent of the relative sizes of the minority groups, and vice versa.

$$\varphi(n_1, n_2, \ldots, n_G) = \varphi\left(\frac{n_1}{n_1 + \cdots + n_G}\right) = \varphi(P_1)$$

$$\psi(n_1, n_2, \ldots, n_G) = \psi(n_2, n_3, \ldots, n_G) = \psi\left(\frac{n_2}{n_2 + \cdots + n_G}, \ldots, \frac{n_G}{n_2 + \cdots + n_G}\right)$$

Using evenness for ψ

The omission of n_1 does not violate *INV*. Let $\tilde{P}_g = \frac{n_g}{n_2+\cdots+n_G}, \forall g \in \{2, \ldots, G\}$,

$$\tilde{P}_g = \frac{n_g}{n_2 + \cdots + n_G} = \frac{\frac{n_g}{n_1+\cdots+n_G}}{\frac{n_2}{n_1+\cdots+n_G} + \cdots + \frac{n_G}{n_1+\cdots+n_G}} = \frac{P_g}{P_2 + \cdots + P_G}$$

Lack of evenness is the distance[1] between the observed distribution of minority groups and the uniform distribution.

[1] Defined as the $p-$metric d_p on $\mathbb{R}^n$

Using Evenness for ψ

Definition: A function $\psi : \mathbb{R}_+^{G-1} \to \mathbb{R}_+$ is a measure of evenness in minority group distribution if it is of the form ($p \in \mathbb{N}$, $a, b \in \mathbb{R}_+$),

$$\psi_p(S) = \psi_p\left(\tilde{P}_2, \ldots, \tilde{P}_G\right) = 1 - \left(\sum_{g=2}^{G} \left|\tilde{P}_g - \frac{1}{G-1}\right|^p\right)^{\frac{1}{p}}$$

Using Evenness for ψ

Definition: A function $\psi : \mathbb{R}_+^{G-1} \to \mathbb{R}_+$ is a measure of evenness in minority group distribution if it is of the form ($p \in \mathbb{N}$, $a, b \in \mathbb{R}_+$),

$$\psi_p(S) = \psi_p\left(\tilde{P}_2, \ldots, \tilde{P}_G\right) = 1 - \left(\sum_{g=2}^{G} \left|\tilde{P}_g - \frac{1}{G-1}\right|^p\right)^{\frac{1}{p}}$$

Proposition: Let $G > 3$ and $\Phi_p(\varphi, \psi_p)$ satisfy SYM, INV, and IND. Holding P_1 constant, if Φ is strictly increasing in ψ_p, then Φ_p satisfies PDT if and if $p > 1$.

[PPT] Principle of Proportional Transfers

Holding the order of groups constant, a transfer from the largest group proportionally to the minority groups that reduces P_1 to $(P_1)^\alpha$ increases heterogeneity by a factor of α.

For example, $\Phi(S_1) < \alpha\Phi(S_1) = \Phi(S_2)$ if

G. \ S.	S_1		S_2
θ_1	P_1		P_1^α
θ_2	P_2	$\prec$	$P_2 + \frac{P_2}{P_2+P_3}(P_1 - P_1^\alpha)$
θ_3	P_3		$P_3 + \frac{P_3}{P_2+P_3}(P_1 - P_1^\alpha)$

[CON] Contractibility

Φ satisfies contractibility if adding one zero-group to a system of G groups decreases the heterogeneity of the system.

$$\Phi(n_2, \ldots, n_G, 0) < \Phi(n_2, \ldots, n_G).$$

Descriptive Units of Heterogeneity

Let $n_1 \geq n_2 > 0$, $P_1 = \frac{n_1}{n_S}$, and $\tilde{P}_g = \frac{n_g}{n_2+\cdots+n_G}$. The Descriptive Units of Heterogeneity (DUH) is defined as:

$$DUH = ln(P_1) \cdot \left[\left(\sum_{g=2}^{G} \left| \tilde{P}_g - \frac{1}{G-1} \right|^p \right)^{\frac{1}{p}} - 1 \right]$$

Descriptive Units of Heterogeneity

Let $n_1 \geq n_2 > 0$, $P_1 = \frac{n_1}{n_S}$, and $\tilde{P}_g = \frac{n_g}{n_2 + \cdots + n_G}$. The Descriptive Units of Heterogeneity (DUH) is defined as:

$$DUH = ln(P_1) \cdot \left[\left(\sum_{g=2}^{G} \left| \tilde{P}_g - \frac{1}{G-1} \right|^p \right)^{\frac{1}{p}} - 1 \right]$$

Theorem

DUH is the unique class of indices, up to positive scalar multiplication, that satisfies the *fundamental Axioms*, *IND*, *PPT*, and *CON* and uses ψ_p for evenness.

Existing Measures Satisfy Some Axioms

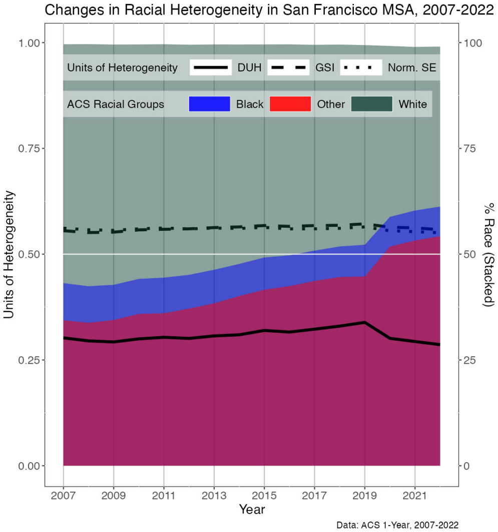
Type	Axiom	Gini	DUH	HHI	SE
Fundamental	Type Symmetry	✓	✓	✓	✓
	Scale Invariance	✓	✓	✓	✓
	Principle of Diminishing Transfers	✓	✓	✓	✓
Characterization	Independence	×	✓	✓	✓
	Principle of Proportional Transfers	×	✓	×	×
	Contractibility	✓	✓	×	×
	Expandability	×	×	✓	✓
	Replication Principle	×	×	✓	×
	Shannon's Additivity	×	×	×	✓

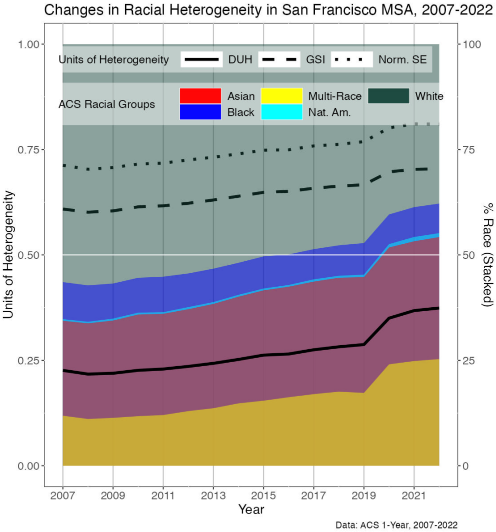
Reasonable Universe of Groups

Determining the elements of Θ is a framing problem and also a judgment call by the researcher. The use of any units of heterogeneity requires justifying the reasonable groupings.

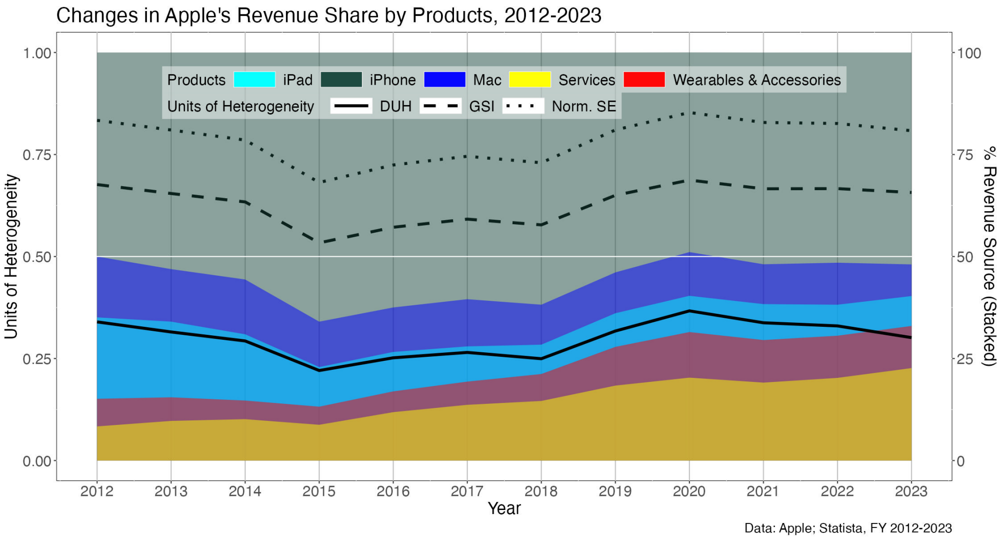
Keeping this in mind, let us consider examples regarding when DUH can be used and why they should be used. For simplicity, I will use my preferred version of DUH, where $p = 2$ so that ψ uses the Euclidean distance.

Reasonable Universe of Groups





Empirical Example: Revenue Stream Heterogeneity



Thank You!

Uniqueness of φ

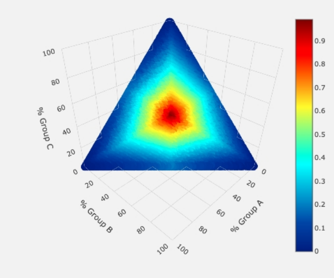
Lemma: If an index $\Phi(\varphi, \psi)$ of heterogeneity satisfies *Scale Invariance* and *Principle of Transfers*, it is monotonically decreasing in P_1, and therefore a positive monotonic transformation of $\frac{1}{P_1}$.

Lemma: If an index $\Phi(\varphi, \psi)$ of heterogeneity satisfies *Independence*, *Scale Invariance*, *Principle of Transfers*, and *Principle of Proportional Transfers*, then φ and ψ must be multiplicatively separable.

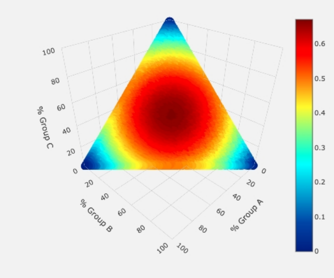
Theorem: If an index $\Phi(\varphi, \psi)$ of heterogeneity satisfies *Independence*, *Scale Invariance*, *Principle of Transfers*, and *Principle of Proportional Transfers*, then it must be $\Phi = \varphi(P_1) \cdot \psi(\tilde{P}_2, \ldots, \tilde{P}_G)$ where $\varphi(P_1) = -c ln(P_1)$, $c \in \mathbb{R}_{++}$

Table 1: Differences between DUH, GSI, and SE

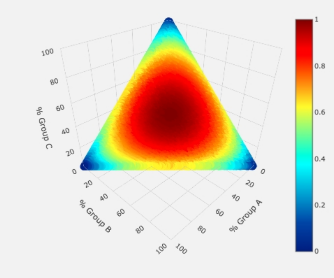
DUH ($p = 2$)



GSI



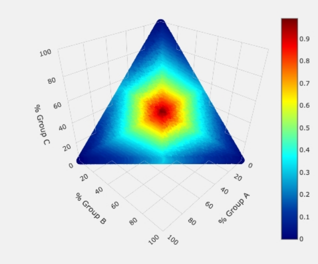
Norm. SE



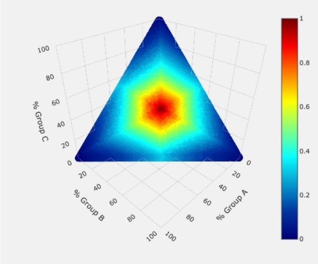
p in ψ_p Determines the "Weight" of Evenness of Minority

Table 2: DUH with Different p

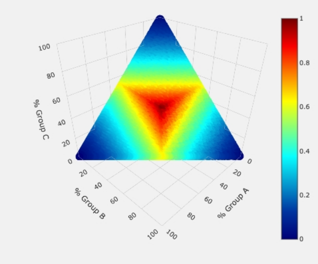
$p = 20$



$p = 200$



$p = 2000$



[EXP] Expandability

Adding any arbitrary number of zero-groups does not affect the measure Φ.

$\Phi(n_1, \ldots, n_G)$ satisfies *Expandability* if

$$\Phi(n_1, \ldots, n_G) = \Phi(n_1, \ldots, n_G, 0)$$

[REP] Replication Principle

$\Phi(n_1, \ldots, n_G)$ satisfies *Replication Principle* (for concentration) if $\forall k \in \mathbb{N}$

$$\frac{1}{k}\Phi(n_1, \ldots, n_G) = \Phi\left(\underbrace{\frac{n_1}{k}, \frac{n_1}{k}, \ldots, \frac{n_1}{k}}_{\text{Sum to } n_1}, \frac{n_2}{k}, \frac{n_2}{k}, \ldots, \underbrace{\frac{n_G}{k}, \ldots, \frac{n_G}{k}}_{\text{Sum to } n_G}\right)$$

[SADD] Shannon's Additivity

Define $n_{gj} \geq 0$ such that $n_g = \sum_{j=1}^{m_g} n_{gj}$, $\forall g \in \{1, \ldots, G\}$, $\forall j \in \{1, \ldots, m_g\}$

$\Phi(n_1, \ldots, n_G)$ satisfies *Shannon's Additivity* if

$$\Phi(n_{11}, \ldots, n_{Gm_G}) = \Phi(n_1, \ldots, n_G) + \sum_{g=1}^{G} \frac{n_g}{n_S} \cdot \Phi\left(\frac{n_{g1}}{n_g}, \ldots, \frac{n_{gm_g}}{n_g}\right)$$

which implies (by setting $m_g = 1$, $\forall g \in \{1, \ldots, G-1\}$ and $n_{G'} = n_G + n_{G+1}$),

$$\begin{aligned}\Phi(n_1, \ldots, n_G, n_{G+1}) =& \Phi(n_1, \ldots, n_{G'}) \\ &+ \frac{n_{G'}}{n_1 + \ldots + n_{G-1} + n_{G'}} \cdot \Phi\left(\frac{n_G}{n_{G'}}, \frac{n_{G+1}}{n_{G'}}\right)\end{aligned}$$