

Allotaxonomy

Last updated: 2026/08/19, 12:31:02 EDT

DOI:

Principles of Complex Systems,
Vols. 1, 2, 3D, 4 Fourever, V for Vendetta

Deliverator: Calla Beauregard
Showrunner: Prof. Peter Sheridan Dodds

Computational Story Lab | Vermont Complex Systems Institute
University of Vermont | Santa Fe Institute



Licensed under the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/)

The PoCVerse
Allotaxonomy
1 of 73

A plenitude of
distances

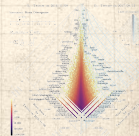
Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

Nutshell

References



These slides are brought to you by:

Sealie & Lambie Productions



The PoCVerse
Allotaxonomy
2 of 73

A plenitude of
distances

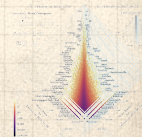
Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

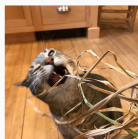
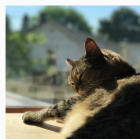
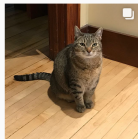
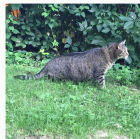
Nutshell

References



These slides are also brought to you by:

Special Guest Executive Producer



 On Instagram at [pratchett_the_cat](https://www.instagram.com/pratchett_the_cat) 

The PoCverse
Allotaxonomy
3 of 73

A plenitude of
distances

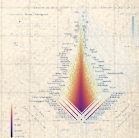
Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

Nutshell

References



Outline

A plenitude of distances

Rank-turbulence divergence

Probability-turbulence divergence

Explorations

Nutshell

References

The PoCSverse
Allotaxonomy
4 of 73

A plenitude of
distances

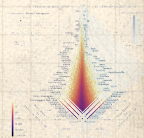
Rank-turbulence
divergence

Probability-
turbulence divergence

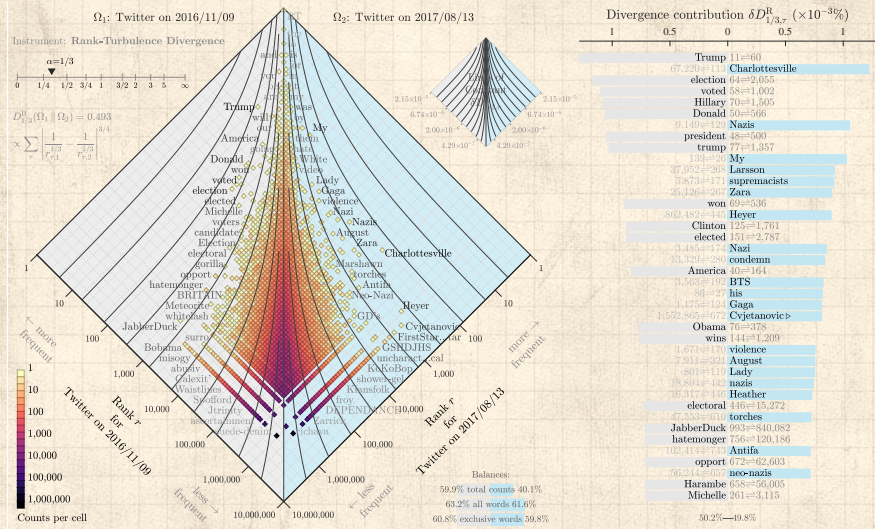
Explorations

Nutshell

References



Goal—Understand this:



A plenitude of
distances

Rank-turbulence
divergence

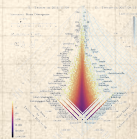
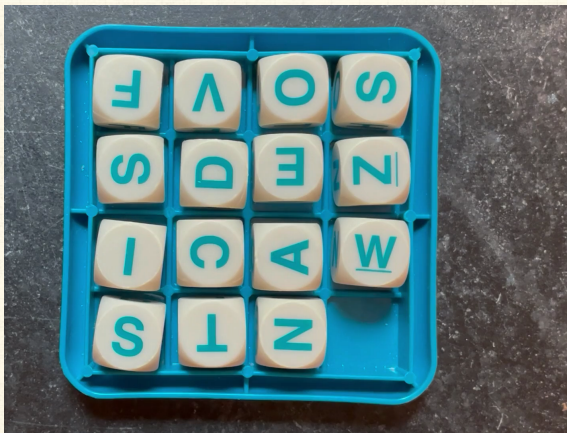
Probability-
turbulence divergence

Explorations

Nutshell

References

The Boggoracle Speaks:



Site (papers, examples, code):

<http://compstorylab.org/allotaxonomy/> 

Foundational papers:



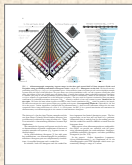
“Allotaxonomy and rank-turbulence divergence:
A universal instrument for comparing complex
systems” 

Dodds et al.,

EPJ Data Science, **12**, 1–42, 2023. ^[5]

EPJ Data Science version 

arXiv version 



“Probability-turbulence divergence: A tunable
allotaxonomic instrument for comparing
heavy-tailed type-size distributions” 

Dodds et al.,

, 2020. ^[6]

Basic science = Describe + Explain:



Dashboards of single scale instruments helps us understand, monitor, and control systems.

The PoCserve
Allotaxonomy
8 of 73

A plenitude of
distances

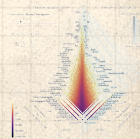
Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

Nutshell

References



Basic science = Describe + Explain:



Dashboards of single scale instruments helps us understand, monitor, and control systems.



Archetype: Cockpit dashboard for flying a plane

The PoCVerse
Allotaxonomy
8 of 73

A plenitude of
distances

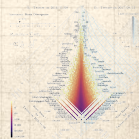
Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

Nutshell

References



Basic science = Describe + Explain:



Dashboards of single scale instruments helps us understand, monitor, and control systems.



Archetype: Cockpit dashboard for flying a plane



Okay if comprehensible.

The PoCVerse
Allotaxonomy
8 of 73

A plenitude of
distances

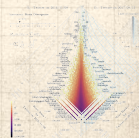
Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

Nutshell

References



Basic science = Describe + Explain:

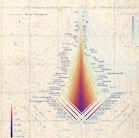
 Dashboards of single scale instruments helps us understand, monitor, and control systems.

 Archetype: Cockpit dashboard for flying a plane

 Okay if comprehensible.

 Complex systems present two problems for dashboards:

1. Scale with internal diversity of components: We need meters for every species, every company, every word.
2. Tracking change: We need to re-arrange meters on the fly.



Basic science = Describe + Explain:

 Dashboards of single scale instruments helps us understand, monitor, and control systems.

 Archetype: Cockpit dashboard for flying a plane

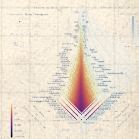
 Okay if comprehensible.

 Complex systems present two problems for dashboards:

1. Scale with internal diversity of components: We need meters for every species, every company, every word.
2. Tracking change: We need to re-arrange meters on the fly.

 Goal—Create comprehensible, dynamically-adjusting, differential dashboards showing two pieces:¹

1. ‘Big picture’ map-like overview,
2. A tunable ranking of components.



Basic science = Describe + Explain:

 Dashboards of single scale instruments helps us understand, monitor, and control systems.

 Archetype: Cockpit dashboard for flying a plane

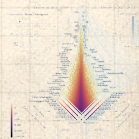
 Okay if comprehensible.

 Complex systems present two problems for dashboards:

1. Scale with internal diversity of components: We need meters for every species, every company, every word.
2. Tracking change: We need to re-arrange meters on the fly.

 Goal—Create comprehensible, dynamically-adjusting, differential dashboards showing two pieces:¹

1. ‘Big picture’ map-like overview,
2. A tunable ranking of components.



¹See the [lexicocalorimeter](#) 

Baby names, much studied: ^[12]

just a decade or so. If you were born in the United States around this year, these are names that are more likely to seem common and generic to you, but are distinctive generational markers.

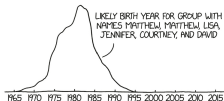
1880 Will, Maude, Minnie, May, Caro, Ida, Lela, Hattie, Jessie, Ada
 1885 Grecco, Maude, Will, Minnie, Lela, Effie, May, Caro, Lela, Nettie
 1890 Maude, May, Minnie, Effie, Maude, Jessie, Nettie, Hattie, Lela, Caro
 1895 Maude, Maude, Minnie, Jessie, Minnie, Hattie, Pearl, Ethel, Bertha
 1900 Maude, Myrtle, Jessie, Mamie, Pearl, Blanche, Gertrude, Ethel, Minnie, Gladys
 1905 Gladys, Viola, Maude, Myrtle, Gertrude, Pearl, Jessie, Blanche, Mamie, Ethel
 1910 Thelma, Gladys, Viola, Mildred, Beatrice, Lucille, Gertrude, Agnes, Hazel, Ethel
 1915 Mildred, Lucille, Thelma, Helen, Bernice, Pauline, Eleanor, Beatrice, Ruth, Dorothy
 1920 Marjorie, Dorothy, Mildred, Lucille, Warren, Thelma, Bernice, Virginia, Helen, June
 1925 Doris, June, Betty, Marjorie, Dorothy, Lorraine, Lela, Norma, Virginia, Joanne
 1930 Dolores, Betty, Jean, Billie, Doris, Norma, Lela, Billy, Jean, Marilyn
 1935 Shirley, Marlene, Joan, Dolores, Marilyn, Bobby, Betty, Billy, Joyce, Beverly
 1940 Corde, Judith, Judy, Carol, Joyce, Barbara, Joan, Carolyn, Shirley, Jerry
 1945 Judy, Judith, Linda, Carol, Sharon, Sandra, Carolyn, Larry, Janice, Dennis
 1950 Linda, Deborah, Gail, Judy, Gary, Larry, Diane, Dennis, Brenda, Annie
 1955 Debra, Deborah, Cathy, Kathy, Pamela, Randy, Kim, Cynthia, Diane, Cheryl
 1960 Debbie, Kim, Terri, Cindy, Kathy, Cathy, Laurie, Lori, Debra, Ricky
 1965 Lisa, Tammy, Lori, Todd, Kim, Mikaela, Tracy, Tina, Susan, Michele, Lisa
 1970 Tammy, Tracy, Tracy, Todd, Dawn, Tina, Sherry, Stacy, Michele, Lisa
 1975 Chad, Jason, Tanya, Heather, Jennifer, Amy, Shari, Shannon, Sherry, Tara
 1980 Brendy, Crystal, April, Jason, Jeremy, Erin, Tiffany, Jamie, Melissa, Jennifer
 1985 Crystal, Lindsay, Ashley, Lindsey, Doreen, Austin, Amanda, Tiffany, Crystal, Amber
 1990 Brittany, Chelsea, Kelsey, Cindy, Ashley, Courtney, Kayla, Kyle, Megan, Jessica
 1995 Taylor, Kelsey, Dakota, Austin, Haley, Cindy, Tyler, Shelby, Brittany, Kayla
 2000 Destiny, Madison, Hailey, Sydney, Kaitlyn, Hailey, Brianna, Hannah, Alyssa
 2005 Aidan, Diego, Gavin, Hailey, Ethan, Madison, Alex, Isabella, Ayden, Aidan
 2010 Jayden, Aidan, Neveah, Addison, Brayden, London, Peyton, Isabella, Ava, Liam
 2015 Ariu, Harper, Scarlett, Jason, Gracyn, Lincoln, Hudson, Liam, Zoey, Layla

If kids in your class were named Jeff, Lisa, Michael, Karen, and David, then you were probably born in the mid-1960s. If they were named Jayden, Isabella, Sophia, Ava, and Ethan, then you were probably born somewhere around 2010.

But names can reveal things about age in other ways.

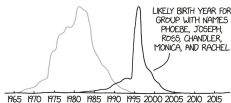
The mid-1990s TV show *Friends* featured six roommates, played by actors, named Matthew, Jennifer, Courtney, Lisa, David, and another Matthew. Each of those names has its own popularity curve; if we combine them all, we can guess what years the group of actors was likely born:

HOW TO: ABSURD SCIENTIFIC ADVICE FOR COMMON REAL-WORLD PROBLEMS



The actors were actually born in the late 1960s, on the very early edge of the popularity of their names. In other words, the actors all have names that were a little before their time. Courtney Cox and Jennifer Aniston had names that didn't really become popular until a decade later. (Maybe people with trendy parents are more likely to wind up in acting.) But the names are generally consistent with their era, if a little ahead of the curve.

We get something very different if we look at the names of their characters—Phoebe, Joseph, Ross, Chandler, Rachel, and Monica:



The show debuted in 1994. There's a clear spike in popularity of the names in 1995 and 1996, which can probably be attributed to the show putting the names in the minds of new parents. But it's not just the show—that name combination was clearly on the rise in the years before *Friends* premiered. It's possible that parents looking for good names for their children are influenced by some of the same cultural trends as TV writers looking for good names for their characters.

The PoCVerse
 Allotaxonomy
 9 of 73

A plenitude of
 distances

Rank-turbulence
 divergence

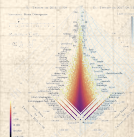
Probability-
 turbulence divergence

Explorations

Nutshell

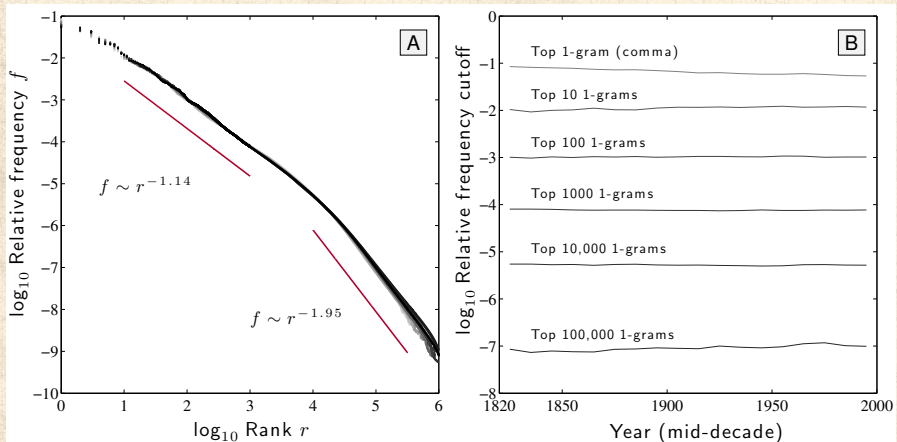
References

How to build a dynamical dashboard that helps sort through a massive number of interconnected time series?





“Is language evolution grinding to a halt? The scaling of lexical turbulence in English fiction suggests it is not” [↗](#)
Pechenick, Danforth, Dodds, Alshaabi, Adams, Reagan, Danforth, Frank, Reagan, and Danforth.
Journal of Computational Science, **21**, 24–37, 2017. ^[14]



For language, Zipf's law has two scaling regimes: ^[19]

$$f \sim \begin{cases} r^{-\alpha} & \text{for } r \ll r_b, \\ r^{-\alpha'} & \text{for } r \gg r_b, \end{cases}$$

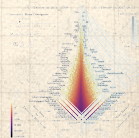
When comparing two texts, define Lexical turbulence as flux of words across a frequency threshold:

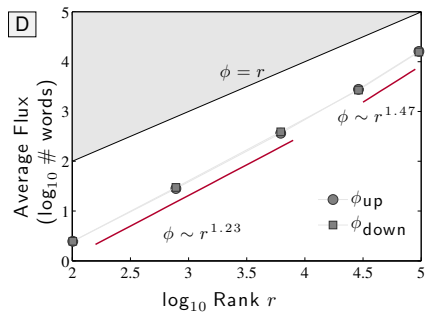
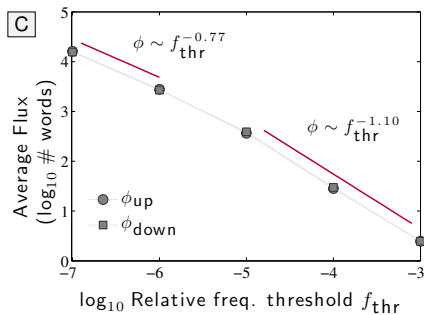
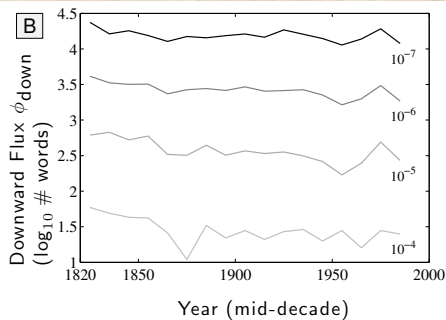
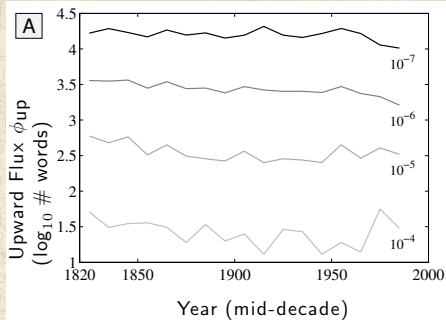
$$\phi \sim \begin{cases} f_{\text{thr}}^{-\mu} & \text{for } f_{\text{thr}} \ll f_b, \\ f_{\text{thr}}^{-\mu'} & \text{for } f_{\text{thr}} \gg f_b, \end{cases}$$

Estimates: $\mu \simeq 0.77$ and $\mu' \simeq 1.10$, and f_b is the scaling break point.

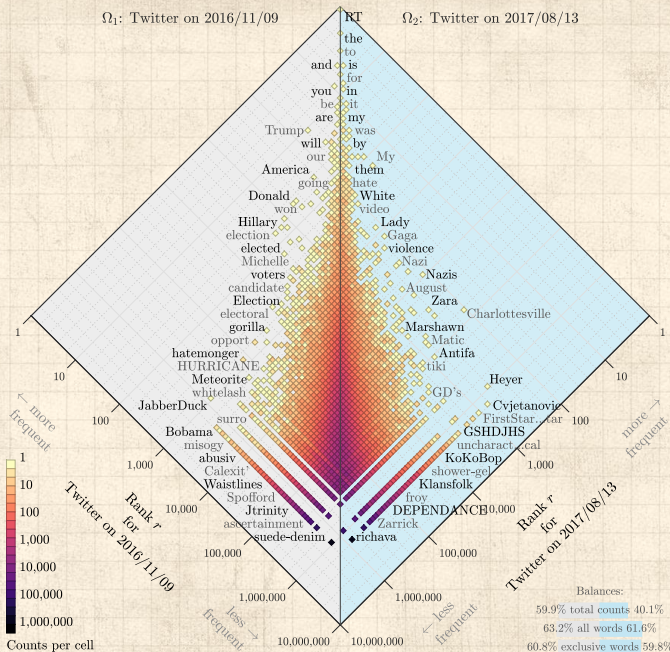
$$\phi \sim \begin{cases} r^\nu = r^{\alpha\mu'} & \text{for } r \ll r_b, \\ r^{\nu'} = r^{\alpha'\mu} & \text{for } r \gg r_b. \end{cases}$$

Estimates: Lower and upper exponents $\nu \simeq 1.23$ and $\nu' \simeq 1.47$.

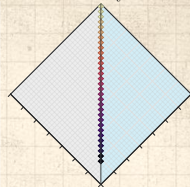




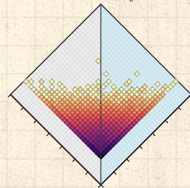
A. Rank-turbulence histogram:

 Ω_1 : Twitter on 2016/11/09 Ω_2 : Twitter on 2017/08/13

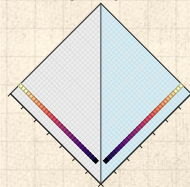
B. Identical systems:



C. Randomized systems:



D. Disjoint systems:



Balances:

59.9% total counts 40.1%

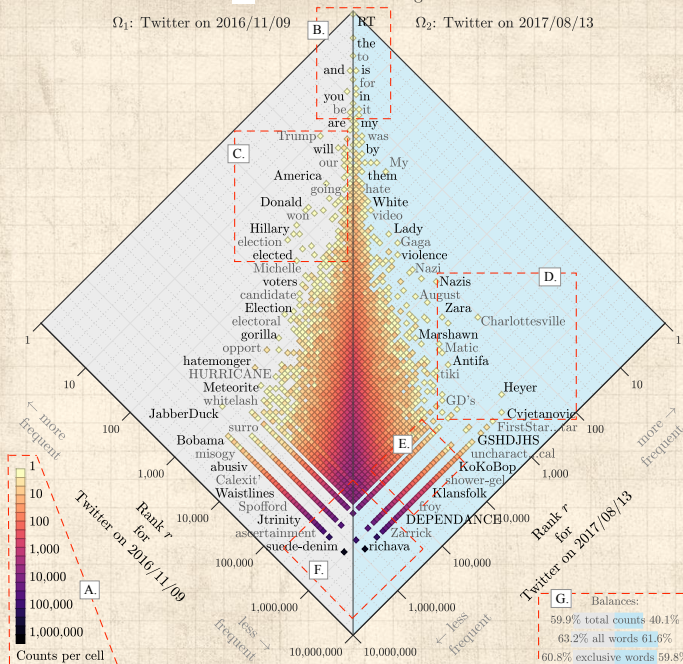
63.2% all words 61.6%

60.8% exclusive words 59.8%

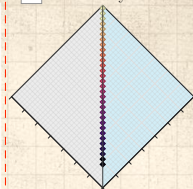
Rank-turbulence histogram:

Ω_1 : Twitter on 2016/11/09

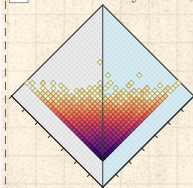
Ω_2 : Twitter on 2017/08/13



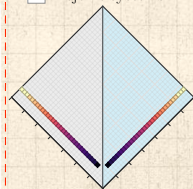
H. Identical systems:



I. Randomized systems:



J. Disjoint systems:



G. Balances:

59.9% total counts 40.1%

63.2% all words 61.6%

60.8% exclusive words 59.8%

Balances:

Top bar (optional)—Total size:



Relative balance of system sizes.



Examples: Total number of words in a book, total number of individuals in an ecology.

The PoCSverse
Allotaxonomy
15 of 73

A plenitude of
distances

Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

Nutshell

References

G

Balances:

59.9% total counts 40.1%

63.2% all words 61.6%

60.8% exclusive words 59.8%

Balances:

Top bar (optional)—Total size:

-  Relative balance of system sizes.
-  Examples: Total number of words in a book, total number of individuals in an ecology.

Middle bar—Types:

-  Fraction of types in each system as a percentage of the union of types from both systems.

The PoCverse
Allotaxonomy
15 of 73

A plenitude of
distances

Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

Nutshell

References

G	Balances:
59.9%	total counts 40.1%
63.2%	all words 61.6%
60.8%	exclusive words 59.8%

Balances:

Top bar (optional)—Total size:

-  Relative balance of system sizes.
-  Examples: Total number of words in a book, total number of individuals in an ecology.

Middle bar—Types:

-  Fraction of types in each system as a percentage of the union of types from both systems.

Bottom bar—Exclusive types:

-  Types that are present in one system only are ‘exclusive types’.
-  $\Omega^{(1)}$ -exclusive and $\Omega^{(2)}$ -exclusive indicate which system an exclusive type belongs to.
-  Percentage of exclusive types in a system relative to that system’s total number of types.

The PoCverse
Allotaxonomy
15 of 73

A plenitude of
distances

Rank-turbulence
divergence

Probability-
turbulence divergence

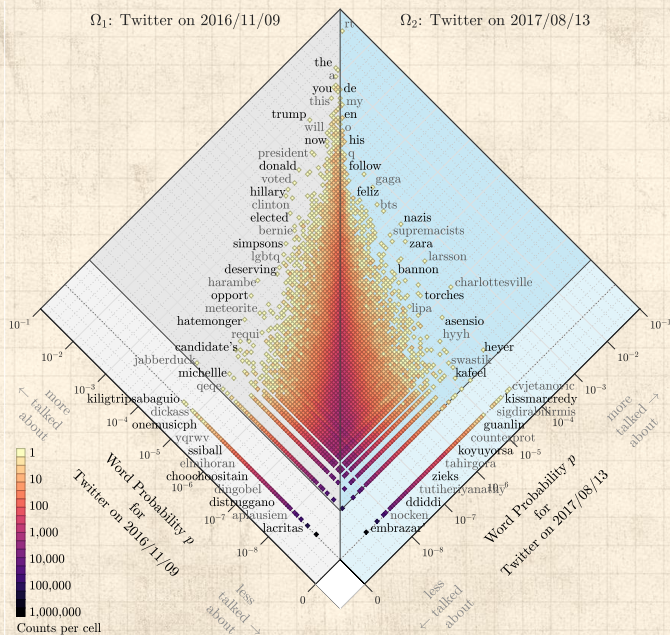
Explorations

Nutshell

References

G	Balances:
59.9%	total counts 40.1%
63.2%	all words 61.6%
60.8%	exclusive words 59.8%

Probability-turbulence histogram:



The PoCVerse
Allotaxonomy
16 of 73

A plenitude of
distances

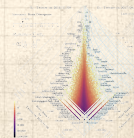
Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

Nutshell

References



So, so many ways to compare probability distributions:



Table 1. Summary of statistical tests for comparing two probability distributions

Test	Null Hypothesis	Alternative Hypothesis	Directionality	Power
Chi-square	Independence	Dependence	Two-sided	Low
Fisher's exact	Independence	Dependence	Two-sided	High
G-test	Independence	Dependence	Two-sided	High
Log-likelihood ratio	Independence	Dependence	Two-sided	High
Mantel-Haenszel	Independence	Dependence	Two-sided	High
McFadden's	Independence	Dependence	Two-sided	High
Nagelkerke's	Independence	Dependence	Two-sided	High
Pearson's	Independence	Dependence	Two-sided	Low
Spearman's	Independence	Dependence	Two-sided	High
Tau-b	Independence	Dependence	Two-sided	High
Tau-c	Independence	Dependence	Two-sided	High
Tau-d	Independence	Dependence	Two-sided	High
Tau-e	Independence	Dependence	Two-sided	High
Tau-f	Independence	Dependence	Two-sided	High
Tau-g	Independence	Dependence	Two-sided	High
Tau-h	Independence	Dependence	Two-sided	High
Tau-i	Independence	Dependence	Two-sided	High
Tau-j	Independence	Dependence	Two-sided	High
Tau-k	Independence	Dependence	Two-sided	High
Tau-l	Independence	Dependence	Two-sided	High
Tau-m	Independence	Dependence	Two-sided	High
Tau-n	Independence	Dependence	Two-sided	High
Tau-o	Independence	Dependence	Two-sided	High
Tau-p	Independence	Dependence	Two-sided	High
Tau-q	Independence	Dependence	Two-sided	High
Tau-r	Independence	Dependence	Two-sided	High
Tau-s	Independence	Dependence	Two-sided	High
Tau-t	Independence	Dependence	Two-sided	High
Tau-u	Independence	Dependence	Two-sided	High
Tau-v	Independence	Dependence	Two-sided	High
Tau-w	Independence	Dependence	Two-sided	High
Tau-x	Independence	Dependence	Two-sided	High
Tau-y	Independence	Dependence	Two-sided	High
Tau-z	Independence	Dependence	Two-sided	High

“Families of Alpha- Beta- and Gamma- Divergences:
Flexible and Robust Measures of Similarities” ↗

Cichocki and Amari,
Entropy, **12**, 1532-1568, 2010. ^[2]

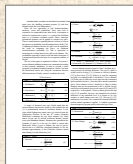


Table 1. Summary of statistical tests for comparing two probability distributions

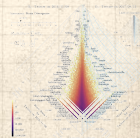
Test	Null Hypothesis	Alternative Hypothesis	Directionality	Power
Chi-square	Independence	Dependence	Two-sided	Low
Fisher's exact	Independence	Dependence	Two-sided	High
G-test	Independence	Dependence	Two-sided	High
Log-likelihood ratio	Independence	Dependence	Two-sided	High
Mantel-Haenszel	Independence	Dependence	Two-sided	High
McFadden's	Independence	Dependence	Two-sided	High
Nagelkerke's	Independence	Dependence	Two-sided	High
Pearson's	Independence	Dependence	Two-sided	Low
Spearman's	Independence	Dependence	Two-sided	High
Tau-b	Independence	Dependence	Two-sided	High
Tau-c	Independence	Dependence	Two-sided	High
Tau-d	Independence	Dependence	Two-sided	High
Tau-e	Independence	Dependence	Two-sided	High
Tau-f	Independence	Dependence	Two-sided	High
Tau-g	Independence	Dependence	Two-sided	High
Tau-h	Independence	Dependence	Two-sided	High
Tau-i	Independence	Dependence	Two-sided	High
Tau-j	Independence	Dependence	Two-sided	High
Tau-k	Independence	Dependence	Two-sided	High
Tau-l	Independence	Dependence	Two-sided	High
Tau-m	Independence	Dependence	Two-sided	High
Tau-n	Independence	Dependence	Two-sided	High
Tau-o	Independence	Dependence	Two-sided	High
Tau-p	Independence	Dependence	Two-sided	High
Tau-q	Independence	Dependence	Two-sided	High
Tau-r	Independence	Dependence	Two-sided	High
Tau-s	Independence	Dependence	Two-sided	High
Tau-t	Independence	Dependence	Two-sided	High
Tau-u	Independence	Dependence	Two-sided	High
Tau-v	Independence	Dependence	Two-sided	High
Tau-w	Independence	Dependence	Two-sided	High
Tau-x	Independence	Dependence	Two-sided	High
Tau-y	Independence	Dependence	Two-sided	High
Tau-z	Independence	Dependence	Two-sided	High

“Comprehensive survey on distance/similarity
measures between probability density functions” ↗

Sung-Hyuk Cha,
International Journal of Mathematical Models and
Methods in Applied Sciences, **1**, 300–307, 2007. ^[1]



Comparisons are distances, divergences ↗, similarities, inner products, fidelities ...



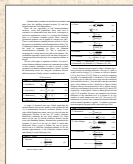
So, so many ways to compare probability distributions:



Test	Category	Properties
Chi-squared	Distance	Asymmetry, Invariance
Kullback-Leibler	Divergence	Asymmetry, Invariance
Jensen-Shannon	Divergence	Symmetry, Invariance
Wasserstein	Distance	Symmetry, Invariance
Earth Mover's	Distance	Symmetry, Invariance
Manhattan	Distance	Symmetry, Invariance
Euclidean	Distance	Symmetry, Invariance
Hamming	Distance	Symmetry, Invariance
Levenshtein	Distance	Symmetry, Invariance
Dynamic Time Warping	Distance	Symmetry, Invariance

“Families of Alpha- Beta- and Gamma- Divergences:
Flexible and Robust Measures of Similarities” 

Cichocki and Amari,
Entropy, **12**, 1532-1568, 2010. ^[2]



Test	Category	Properties
Chi-squared	Distance	Asymmetry, Invariance
Kullback-Leibler	Divergence	Asymmetry, Invariance
Jensen-Shannon	Divergence	Symmetry, Invariance
Wasserstein	Distance	Symmetry, Invariance
Earth Mover's	Distance	Symmetry, Invariance
Manhattan	Distance	Symmetry, Invariance
Euclidean	Distance	Symmetry, Invariance
Hamming	Distance	Symmetry, Invariance
Levenshtein	Distance	Symmetry, Invariance
Dynamic Time Warping	Distance	Symmetry, Invariance

“Comprehensive survey on distance/similarity
measures between probability density functions” 

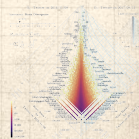
Sung-Hyuk Cha,
International Journal of Mathematical Models and
Methods in Applied Sciences, **1**, 300–307, 2007. ^[1]



Comparisons are distances, divergences , similarities, inner products, fidelities ...



60ish kinds of comparisons grouped into 10 families



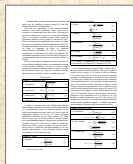
So, so many ways to compare probability distributions:



Test	Family	Robust	Flexible	Similarities
Chi-squared	Alpha	No	No	No
Fisher's	Beta	No	No	No
Gini	Gamma	No	No	No
...	...	...	...	...

“Families of Alpha- Beta- and Gamma- Divergences:
Flexible and Robust Measures of Similarities” 

Cichocki and Amari,
Entropy, **12**, 1532-1568, 2010. ^[2]



Test	Family	Robust	Flexible	Similarities
Chi-squared	Alpha	No	No	No
Fisher's	Beta	No	No	No
Gini	Gamma	No	No	No
...	...	...	...	...

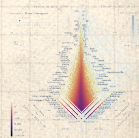
“Comprehensive survey on distance/similarity
measures between probability density functions” 

Sung-Hyuk Cha,
International Journal of Mathematical Models and
Methods in Applied Sciences, **1**, 300–307, 2007. ^[1]

 Comparisons are distances, divergences , similarities, inner
products, fidelities ...

 60ish kinds of comparisons grouped into 10 families

 A worry: Subsampled distributions with very heavy tails



Quite the festival:

Table 1. L_p Minkowski family		
1. Euclidean L_2	$d_{L_2} = \sqrt{\frac{1}{2} P - Q ^2}$	(1)
2. City block L_1	$d_{L_1} = \frac{1}{2} P - Q $	(2)
3. Minkowski L_p	$d_{L_p} = \sqrt[p]{\frac{1}{2} P - Q ^p}$	(3)
4. Chebyshev L_∞	$d_{L_\infty} = \min(P - Q)$	(4)

Table 2. L_1 family		
5. Sørensen	$d_{L_1} = \frac{\sum_i P_i - Q_i }{\sum_i (P_i + Q_i)}$	(5)

6. Gower	$d_{Gower} = \frac{1}{2} \sum_i \frac{ P_i - Q_i }{R_i}$	(6)
	$= \frac{1}{2} \sum_i \frac{ P_i - Q_i }{R_i}$	(7)

7. Soergel	$d_{Soergel} = \frac{\sum_i P_i - Q_i }{\sum_i \max(P_i, Q_i)}$	(8)
------------	--	-----

8. Kulczynski d	$d_{Kul} = \frac{\sum_i P_i - Q_i }{\sum_i \max(P_i, Q_i)}$	(9)
-------------------	--	-----

9. Canberra	$d_{Can} = \frac{\sum_i P_i - Q_i }{P_i + Q_i}$	(10)
10. Lorentzian	$d_{Lor} = \sum_i \ln(1 + P_i - Q_i)$	(11)

* L_1 family $\rightarrow$ Intersection (13), Wave Hodges (15), Czekanowski (16), Ruzicka (21), Tanimoto (23), etc.

Table 3. Intersection family		
11. Intersection	$x_{int} = \min(P_i, Q_i)$	(12)
	$d_{int} = 1 - x_{int} = \frac{1}{2} \sum_i P_i - Q_i $	(13)
12. Wave Hodges	$d_{WH} = \sum_i \frac{\min(P_i, Q_i)}{\max(P_i, Q_i)}$	(14)
	$= \sum_i \frac{ P_i - Q_i }{\max(P_i, Q_i)}$	(15)
13. Czekanowski	$x_{czk} = \frac{2 \sum_i \min(P_i, Q_i)}{\sum_i (P_i + Q_i)}$	(16)
	$d_{czk} = 1 - x_{czk} = \frac{\sum_i P_i - Q_i }{\sum_i (P_i + Q_i)}$	(17)

14. Motyka	$x_{mot} = \frac{\sum_i \min(P_i^2, Q_i)}{\sum_i (P_i + Q_i)}$	(18)
------------	--	------

	$d_{mot} = 1 - x_{mot} = \frac{\sum_i \max(P_i^2, Q_i)}{\sum_i (P_i + Q_i)}$	(19)
--	--	------

15. Kulczynski r	$x_{kulr} = \frac{1}{d_{kul}} \sum_i \min(P_i, Q_i)$	(20)
--------------------	--	------

16. Ruzicka	$x_{ruz} = \frac{\sum_i \min(P_i, Q_i)}{\sum_i \max(P_i, Q_i)}$	(21)
-------------	---	------

17. Tanimoto	$d_{tan} = \frac{\sum_i P_i + \sum_i Q_i - 2 \sum_i \min(P_i, Q_i)}{\sum_i P_i + \sum_i Q_i}$	(22)
--------------	---	------

	$= \frac{\sum_i \max(P_i, Q_i) - \min(P_i, Q_i)}{\sum_i \max(P_i, Q_i)}$	(23)
--	--	------

Table 4. Inner Product family		
18. Inner Product	$x_{ip} = P \cdot Q = \sum_i P_i Q_i$	(24)
19. Harmonic mean	$x_{hm} = \frac{2 \sum_i P_i Q_i}{\sum_i P_i + \sum_i Q_i}$	(25)
20. Cosine	$x_{cos} = \frac{\sum_i P_i Q_i}{\sqrt{\sum_i P_i^2} \sqrt{\sum_i Q_i^2}}$	(26)

21. Kumar-Hausbrook (PCE)	$x_{khu} = \frac{\sum_i P_i Q_i}{\sum_i P_i^2 + \sum_i Q_i^2 - \sum_i P_i Q_i}$	(27)
---------------------------	---	------

22. Jaccard	$x_{jac} = \frac{\sum_i P_i Q_i}{\sum_i P_i^2 + \sum_i Q_i^2 - \sum_i P_i Q_i}$	(28)
-------------	---	------

	$d_{jac} = 1 - x_{jac} = \frac{\sum_i (P_i - Q_i)^2}{\sum_i P_i^2 + \sum_i Q_i^2 - \sum_i P_i Q_i}$	(29)
--	---	------

23. Dice	$x_{dic} = \frac{2 \sum_i P_i Q_i}{\sum_i P_i^2 + \sum_i Q_i^2}$	(30)
----------	--	------

	$d_{dic} = 1 - x_{dic} = \frac{\sum_i (P_i - Q_i)^2}{\sum_i P_i^2 + \sum_i Q_i^2}$	(31)
--	--	------

	$d_{dic} = 1 - x_{dic} = \frac{\sum_i (P_i - Q_i)^2}{\sum_i P_i^2 + \sum_i Q_i^2}$	(32)
--	--	------

24. Fidelity	$x_{fid} = \sum_i \sqrt{P_i Q_i}$	(33)
--------------	-----------------------------------	------

25. Bhattacharyya	$d_b = -\ln \sum_i \sqrt{P_i Q_i}$	(34)
-------------------	------------------------------------	------

26. Hellinger	$d_h = \sqrt{\frac{1}{2} \sum_i P_i - Q_i }$	(35)
	$= \sqrt{\frac{1}{2} \sum_i P_i - Q_i }$	(36)

27. Mutanata	$d_{mut} = \sqrt{\sum_i P_i - Q_i }$	(37)
	$= \sqrt{2 \sum_i P_i - Q_i }$	(38)

28. Squared-chord	$d_{sqc} = \sqrt{\sum_i P_i - Q_i ^2}$	(39)
-------------------	---	------

	$x_{sqc} = 1 - d_{sqc}$	(40)
--	-------------------------	------

Table 6. Squared L_2 family or χ^2 family		
29. Squared Euclidean	$d_{sqe} = \sum_i (P_i - Q_i)^2$	(41)

30. Pearson χ^2	$d_{pc} = \sum_i \frac{(P_i - Q_i)^2}{Q_i}$	(42)
----------------------	---	------

31. Neyman χ^2	$d_{nc} = \sum_i \frac{(P_i - Q_i)^2}{P_i}$	(43)
---------------------	---	------

32. Squared χ^2	$d_{sq\chi^2} = \sum_i \frac{(P_i - Q_i)^2}{P_i + Q_i}$	(44)
----------------------	---	------

33. Probabilistic Symmetric χ^2	$d_{ps\chi^2} = \sum_i \frac{(P_i - Q_i)^2}{P_i + Q_i}$	(45)
--------------------------------------	---	------

34. Divergence	$d_{div} = 2 \sum_i \frac{(P_i - Q_i)^2}{(P_i + Q_i)^2}$	(46)
----------------	--	------

35. Clark	$d_{cl} = \sqrt{\sum_i \left(\frac{ P_i - Q_i }{P_i + Q_i} \right)^2}$	(47)
-----------	---	------

36. Additive Symmetric χ^2	$d_{as\chi^2} = \sum_i \frac{(P_i - Q_i)^2}{P_i + Q_i}$	(48)
---------------------------------	---	------

* Squared L_2 family $\rightarrow$ Jaccard (29), Dice (31)

Table 7. Shannon's entropy family		
37. Kullback-Leibler	$d_{kl} = \sum_i P_i \ln \frac{P_i}{Q_i}$	(49)

38. Jeffreys	$d_{j} = \sum_i (P_i - Q_i) \ln \frac{P_i}{Q_i}$	(50)
--------------	--	------

39. K. divergence	$d_{k} = \sum_i P_i \ln \frac{2P_i}{P_i + Q_i}$	(51)
-------------------	---	------

40. Topsoe	$d_{top} = \sum_i P_i \ln \left(\frac{2P_i}{P_i + Q_i} + Q_i \ln \left(\frac{2Q_i}{P_i + Q_i} \right) \right)$	(52)
------------	--	------

41. Jensen-Shannon	$d_{js} = \frac{1}{2} \sum_i P_i \ln \left(\frac{2P_i}{P_i + Q_i} \right) + \frac{1}{2} \sum_i Q_i \ln \left(\frac{2Q_i}{P_i + Q_i} \right)$	(53)
--------------------	--	------

42. Jensen difference	$d_{jd} = \frac{1}{2} \sum_i \left[P_i \ln \frac{P_i}{P_i + Q_i} + Q_i \ln \frac{Q_i}{P_i + Q_i} \right]$	(54)
-----------------------	--	------

Table 8. Combinations		
-----------------------	--	--

43. Taneja	$d_{tan} = \frac{\sum_i (P_i + Q_i)}{2} \ln \left(\frac{P_i + Q_i}{2 P_i Q_i} \right)$	(55)
------------	---	------

44. Kumar-Johnson	$d_{kj} = \sum_i \left(\frac{ P_i - Q_i }{2 P_i Q_i} \right)$	(56)
-------------------	--	------

45. Avg(L ₁ , L _∞)	$d_{av} = \frac{\sum_i P_i - Q_i + \max_i P_i - Q_i }{2}$	(57)
---	--	------

Table 10. Vicinostade		
Vicin-Wave Hodges	$d_{vwh} = \sum_i \frac{ P_i - Q_i }{\max(P_i, Q_i)}$	(58)

Vicin-Symmetric χ^2	$d_{vs\chi^2} = \sum_i \frac{(P_i - Q_i)^2}{\max(P_i, Q_i)}$	(59)
--------------------------	--	------

Vicin-Symmetric χ^2	$d_{vs\chi^2} = \sum_i \frac{(P_i - Q_i)^2}{\max(P_i, Q_i)}$	(60)
--------------------------	--	------

Vicin-Symmetric χ^2	$d_{vs\chi^2} = \sum_i \frac{(P_i - Q_i)^2}{\max(P_i, Q_i)}$	(61)
--------------------------	--	------

max-Symmetric	$d_{max} = \max \left(\sum_i \frac{(P_i - Q_i)^2}{P_i}, \sum_i \frac{(P_i - Q_i)^2}{Q_i} \right)$	(62)
---------------	--	------

min-symmetric	$d_{min} = \min \left(\sum_i \frac{(P_i - Q_i)^2}{P_i}, \sum_i \frac{(P_i - Q_i)^2}{Q_i} \right)$	(63)
---------------	--	------

The PoCVerse
Allotaxonomy
18 of 73

A plenitude of
distances

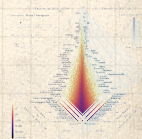
Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

Nutshell

References



Shannon tried to slow things down in 1956:



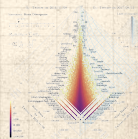
“The bandwagon” 

Claude E Shannon,

IRE Transactions on Information Theory, 2, 3,
1956. ^[16]



“Information theory has ... become something of a scientific bandwagon.”



Shannon tried to slow things down in 1956:



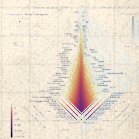
“The bandwagon” 

Claude E Shannon,

IRE Transactions on Information Theory, 2, 3,
1956. ^[16]

 “Information theory has ... become something of a scientific bandwagon.”

 “While ... information theory is indeed a valuable tool ... [it] is certainly no panacea for the communication engineer or ... for anyone else.



Shannon tried to slow things down in 1956:



“The bandwagon” 

Claude E Shannon,

IRE Transactions on Information Theory, 2, 3,
1956. ^[16]



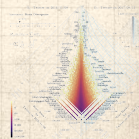
“Information theory has ... become something of a scientific bandwagon.”



“While ... information theory is indeed a valuable tool ... [it] is certainly no panacea for the communication engineer or ... for anyone else.



“A few first rate research papers are preferable to a large number that are poorly conceived or half-finished.”





We want two main things:

1. A measure of difference between systems
2. A way of sorting which types/species/words contribute to that difference

Table 1. L_p Minkowski family

1. Euclidean L_2	$d_{Euc} = \sqrt{\sum_{i=1}^d P_i - Q_i ^2}$	(1)
2. City block L_1	$d_{CB} = \sum_{i=1}^d P_i - Q_i $	(2)
3. Minkowski L_p	$d_{Mk} = \sqrt[p]{\sum_{i=1}^d P_i - Q_i ^p}$	(3)
4. Chebyshev L_∞	$d_{Cheb} = \max_i P_i - Q_i $	(4)

Table 2. L_1 family

5. Sørensen	$d_{sor} = \frac{\sum_{i=1}^d P_i - Q_i }{\sum_{i=1}^d (P_i + Q_i)}$	(5)
6. Gower	$d_{gow} = \frac{1}{d} \sum_{i=1}^d \frac{ P_i - Q_i }{R_i}$	(6)
	$= \frac{1}{d} \sum_{i=1}^d P_i - Q_i $	(7)
7. Soergel	$d_{sg} = \frac{\sum_{i=1}^d P_i - Q_i }{\sum_{i=1}^d \max(P_i, Q_i)}$	(8)
8. Kulczynski d	$d_{kul} = \frac{\sum_{i=1}^d P_i - Q_i }{\sum_{i=1}^d \min(P_i, Q_i)}$	(9)
9. Canberra	$d_{can} = \sum_{i=1}^d \frac{ P_i - Q_i }{P_i + Q_i}$	(10)
10. Lorentzian	$d_{lor} = \sum_{i=1}^d \ln(1 + P_i - Q_i)$	(11)
* L_1 family $\supset$ {Intersectoin (13), Wave Hedges (15), Czekanowski (16), Ruzicka (21), Tanimoto (23), etc}.		

The PoCSeve
Allotaxonomy
20 of 73

A plenitude of
distances

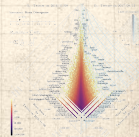
Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

Nutshell

References





We want two main things:

1. A measure of difference between systems
2. A way of sorting which types/species/words contribute to that difference



For sorting, many comparisons give the same ordering.

Table 1. L_p Minkowski family

1. Euclidean L_2	$d_{Euc} = \sqrt{\sum_{i=1}^d P_i - Q_i ^2}$	(1)
2. City block L_1	$d_{CB} = \sum_{i=1}^d P_i - Q_i $	(2)
3. Minkowski L_p	$d_{Mk} = \sqrt[p]{\sum_{i=1}^d P_i - Q_i ^p}$	(3)
4. Chebyshev L_∞	$d_{Cheb} = \max_i P_i - Q_i $	(4)

Table 2. L_1 family

5. Sørensen	$d_{sor} = \frac{\sum_{i=1}^d P_i - Q_i }{\sum_{i=1}^d (P_i + Q_i)}$	(5)
-------------	---	-----

6. Gower	$d_{gow} = \frac{1}{d} \sum_{i=1}^d \frac{ P_i - Q_i }{R_i}$	(6)
	$= \frac{1}{d} \sum_{i=1}^d P_i - Q_i $	(7)

7. Soergel	$d_{sg} = \frac{\sum_{i=1}^d P_i - Q_i }{\sum_{i=1}^d \max(P_i, Q_i)}$	(8)
------------	---	-----

8. Kulczynski d	$d_{kul} = \frac{\sum_{i=1}^d P_i - Q_i }{\sum_{i=1}^d \min(P_i, Q_i)}$	(9)
-------------------	--	-----

9. Canberra	$d_{can} = \sum_{i=1}^d \frac{ P_i - Q_i }{P_i + Q_i}$	(10)
-------------	--	------

10. Lorentzian	$d_{lor} = \sum_{i=1}^d \ln(1 + P_i - Q_i)$	(11)
----------------	---	------

* L_1 family $\supset$ {Intersectoin (13), Wave Hedges (15), Czekanowski (16), Ruzicka (21), Tanimoto (23), etc}.

The PoCverse
Allotaxonomy
20 of 73

A plenitude of
distances

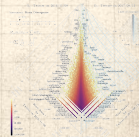
Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

Nutshell

References





We want two main things:

1. A measure of difference between systems
2. A way of sorting which types/species/words contribute to that difference



For sorting, many comparisons give the same ordering.



A few basic building blocks:

-  $|P_i - Q_i|$ (dominant)
-  $\max(P_i, Q_i)$
-  $\min(P_i, Q_i)$
-  $P_i Q_i$
-  $|P_i^{1/2} - Q_i^{1/2}|$ (Hellinger)

Table 1. L_p Minkowski family

1. Euclidean L_2	$d_{Euc} = \sqrt{\sum_{i=1}^d P_i - Q_i ^2}$	(1)
2. City block L_1	$d_{CB} = \sum_{i=1}^d P_i - Q_i $	(2)
3. Minkowski L_p	$d_{Mk} = \sqrt[p]{\sum_{i=1}^d P_i - Q_i ^p}$	(3)
4. Chebyshev L_∞	$d_{Cheb} = \max_i P_i - Q_i $	(4)

Table 2. L_1 family

5. Sørensen	$d_{sor} = \frac{\sum_{i=1}^d P_i - Q_i }{\sum_{i=1}^d (P_i + Q_i)}$	(5)
-------------	---	-----

6. Gower	$d_{gow} = \frac{1}{d} \sum_{i=1}^d \frac{ P_i - Q_i }{R_i}$	(6)
	$= \frac{1}{d} \sum_{i=1}^d P_i - Q_i $	(7)

7. Soergel	$d_{sg} = \frac{\sum_{i=1}^d P_i - Q_i }{\sum_{i=1}^d \max(P_i, Q_i)}$	(8)
------------	---	-----

8. Kulczynski d	$d_{kul} = \frac{\sum_{i=1}^d P_i - Q_i }{\sum_{i=1}^d \min(P_i, Q_i)}$	(9)
-------------------	--	-----

9. Canberra	$d_{can} = \sum_{i=1}^d \frac{ P_i - Q_i }{P_i + Q_i}$	(10)
-------------	--	------

10. Lorentzian	$d_{lor} = \sum_{i=1}^d \ln(1 + P_i - Q_i)$	(11)
----------------	---	------

* L_1 family $\supset$ {Intersectoin (13), Wave Hedges (15), Czekanowski (16), Ruzicka (21), Tanimoto (23), etc}.

The PoCServe
Allotaxonomy
20 of 73

A plenitude of
distances

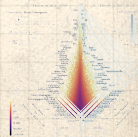
Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

Nutshell

References





Information theoretic
sortings are more opaque

Table 1. L_p Minkowski family		
1. Euclidean L_2	$d_{Euc} = \sqrt{\sum_{i=1}^d P_i - Q_i ^2}$	(1)
2. City block L_1	$d_{CB} = \sum_{i=1}^d P_i - Q_i $	(2)
3. Minkowski L_p	$d_{Mk} = \sqrt[p]{\sum_{i=1}^d P_i - Q_i ^p}$	(3)
4. Chebyshev L_∞	$d_{Cheb} = \max_i P_i - Q_i $	(4)

Table 2. L_1 family		
5. Sørensen	$d_{sor} = \frac{\sum_{i=1}^d P_i - Q_i }{\sum_{i=1}^d (P_i + Q_i)}$	(5)

6. Gower	$d_{gow} = \frac{1}{d} \sum_{i=1}^d \frac{ P_i - Q_i }{R_i}$	(6)
	$= \frac{1}{d} \sum_{i=1}^d P_i - Q_i $	(7)

7. Soergel	$d_{sg} = \frac{\sum_{i=1}^d P_i - Q_i }{\sum_{i=1}^d \max(P_i, Q_i)}$	(8)
------------	---	-----

8. Kulczynski d	$d_{kul} = \frac{\sum_{i=1}^d P_i - Q_i }{\sum_{i=1}^d \min(P_i, Q_i)}$	(9)
-------------------	--	-----

9. Canberra	$d_{can} = \sum_{i=1}^d \frac{ P_i - Q_i }{P_i + Q_i}$	(10)
-------------	--	------

10. Lorentzian	$d_{lor} = \sum_{i=1}^d \ln(1 + P_i - Q_i)$	(11)
----------------	---	------

* L_1 family $\supset$ {Intersectoin (13), Wave Hedges (15), Czekanowski (16), Ruzicka (21), Tanimoto (23), etc}.

The PoCSeve
Allotaxonomy
21 of 73

A plenitude of
distances

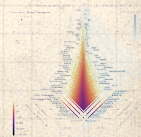
Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

Nutshell

References





Information theoretic
sortings are more opaque



No tunability

Table 1. L_p Minkowski family

1. Euclidean L_2	$d_{Euc} = \sqrt{\sum_{i=1}^d P_i - Q_i ^2}$	(1)
2. City block L_1	$d_{CB} = \sum_{i=1}^d P_i - Q_i $	(2)
3. Minkowski L_p	$d_{Mk} = \sqrt[p]{\sum_{i=1}^d P_i - Q_i ^p}$	(3)
4. Chebyshev L_∞	$d_{Cheb} = \max_i P_i - Q_i $	(4)

Table 2. L_1 family

5. Sørensen	$d_{sor} = \frac{\sum_{i=1}^d P_i - Q_i }{\sum_{i=1}^d (P_i + Q_i)}$	(5)
-------------	---	-----

6. Gower	$d_{gow} = \frac{1}{d} \sum_{i=1}^d \frac{ P_i - Q_i }{R_i}$	(6)
	$= \frac{1}{d} \sum_{i=1}^d P_i - Q_i $	(7)

7. Soergel	$d_{sg} = \frac{\sum_{i=1}^d P_i - Q_i }{\sum_{i=1}^d \max(P_i, Q_i)}$	(8)
------------	---	-----

8. Kulczynski d	$d_{kul} = \frac{\sum_{i=1}^d P_i - Q_i }{\sum_{i=1}^d \min(P_i, Q_i)}$	(9)
-------------------	--	-----

9. Canberra	$d_{can} = \sum_{i=1}^d \frac{ P_i - Q_i }{P_i + Q_i}$	(10)
-------------	--	------

10. Lorentzian	$d_{lor} = \sum_{i=1}^d \ln(1 + P_i - Q_i)$	(11)
----------------	---	------

* L_1 family $\supset$ {Intersectoin (13), Wave Hedges (15), Czekanowski (16), Ruzicka (21), Tanimoto (23), etc}.

The PoCSeve
Allotaxonomy
21 of 73

A plenitude of
distances

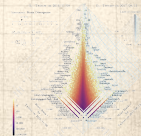
Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

Nutshell

References





Shannon's Entropy:

$$H(P) = \langle \log_2 \frac{1}{p_\tau} \rangle = \sum_{\tau \in R_{1,2;\alpha}} p_\tau \log_2 \frac{1}{p_\tau} \quad (1)$$

The PoCVerse
Allotaxonomy
22 of 73

A plenitude of
distances

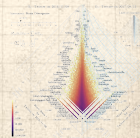
Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

Nutshell

References





Shannon's Entropy:

$$H(P) = \left\langle \log_2 \frac{1}{p_\tau} \right\rangle = \sum_{\tau \in R_{1,2;\alpha}} p_\tau \log_2 \frac{1}{p_\tau} \quad (1)$$



Kullback-Liebler (KL) divergence:

$$\begin{aligned} D^{\text{KL}}(P_2 \parallel P_1) &= \left\langle \log_2 \frac{1}{p_{2,\tau}} - \log_2 \frac{1}{p_{1,\tau}} \right\rangle_{P_2} \\ &= \sum_{\tau \in R_{1,2;\alpha}} p_{2,\tau} \left[\log_2 \frac{1}{p_{2,\tau}} - \log_2 \frac{1}{p_{1,\tau}} \right] \\ &= \sum_{\tau \in R_{1,2;\alpha}} p_{2,\tau} \log_2 \frac{p_{1,\tau}}{p_{2,\tau}}. \end{aligned} \quad (2)$$

The PoCverse
Allotaxonomy
22 of 73

A plenitude of
distances

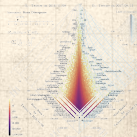
Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

Nutshell

References





Shannon's Entropy:

$$H(P) = \left\langle \log_2 \frac{1}{p_\tau} \right\rangle = \sum_{\tau \in R_{1,2;\alpha}} p_\tau \log_2 \frac{1}{p_\tau} \quad (1)$$

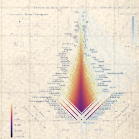


Kullback-Liebler (KL) divergence:

$$\begin{aligned} D^{\text{KL}}(P_2 \parallel P_1) &= \left\langle \log_2 \frac{1}{p_{2,\tau}} - \log_2 \frac{1}{p_{1,\tau}} \right\rangle_{P_2} \\ &= \sum_{\tau \in R_{1,2;\alpha}} p_{2,\tau} \left[\log_2 \frac{1}{p_{2,\tau}} - \log_2 \frac{1}{p_{1,\tau}} \right] \\ &= \sum_{\tau \in R_{1,2;\alpha}} p_{2,\tau} \log_2 \frac{p_{1,\tau}}{p_{2,\tau}}. \end{aligned} \quad (2)$$



Problem: If just one component type in system 2 is not present in system 1, KL divergence = ∞ .





Shannon's Entropy:

$$H(P) = \left\langle \log_2 \frac{1}{p_\tau} \right\rangle = \sum_{\tau \in R_{1,2;\alpha}} p_\tau \log_2 \frac{1}{p_\tau} \quad (1)$$



Kullback-Liebler (KL) divergence:

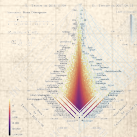
$$\begin{aligned} D^{\text{KL}}(P_2 \parallel P_1) &= \left\langle \log_2 \frac{1}{p_{2,\tau}} - \log_2 \frac{1}{p_{1,\tau}} \right\rangle_{P_2} \\ &= \sum_{\tau \in R_{1,2;\alpha}} p_{2,\tau} \left[\log_2 \frac{1}{p_{2,\tau}} - \log_2 \frac{1}{p_{1,\tau}} \right] \\ &= \sum_{\tau \in R_{1,2;\alpha}} p_{2,\tau} \log_2 \frac{p_{1,\tau}}{p_{2,\tau}}. \end{aligned} \quad (2)$$



Problem: If just one component type in system 2 is not present in system 1, KL divergence = ∞ .



Solution: If we can't compare a spork and a platypus directly, we create a fictional **spork-platypus hybrid**.





Shannon's Entropy:

$$H(P) = \left\langle \log_2 \frac{1}{p_\tau} \right\rangle = \sum_{\tau \in R_{1,2;\alpha}} p_\tau \log_2 \frac{1}{p_\tau} \quad (1)$$



Kullback-Liebler (KL) divergence:

$$\begin{aligned} D^{\text{KL}}(P_2 \parallel P_1) &= \left\langle \log_2 \frac{1}{p_{2,\tau}} - \log_2 \frac{1}{p_{1,\tau}} \right\rangle_{P_2} \\ &= \sum_{\tau \in R_{1,2;\alpha}} p_{2,\tau} \left[\log_2 \frac{1}{p_{2,\tau}} - \log_2 \frac{1}{p_{1,\tau}} \right] \\ &= \sum_{\tau \in R_{1,2;\alpha}} p_{2,\tau} \log_2 \frac{p_{1,\tau}}{p_{2,\tau}}. \end{aligned} \quad (2)$$



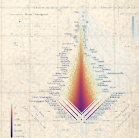
Problem: If just one component type in system 2 is not present in system 1, KL divergence = ∞ .



Solution: If we can't compare a spork and a platypus directly, we create a fictional **spork-platypus hybrid**.



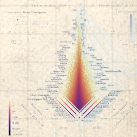
New problem: Re-read solution.



 Jensen-Shannon divergence (JSD): [9, 7, 13, 1]

$$\begin{aligned}
 D^{\text{JS}}(P_1 \parallel P_2) &= \frac{1}{2} D^{\text{KL}}\left(P_1 \parallel \frac{1}{2}[P_1 + P_2]\right) + \frac{1}{2} D^{\text{KL}}\left(P_2 \parallel \frac{1}{2}[P_1 + P_2]\right) \\
 &= \frac{1}{2} \sum_{\tau \in R_{1,2;\alpha}} \left(p_{1,\tau} \log_2 \frac{p_{1,\tau}}{\frac{1}{2}[p_{1,\tau} + p_{2,\tau}]} + p_{2,\tau} \log_2 \frac{p_{2,\tau}}{\frac{1}{2}[p_{1,\tau} + p_{2,\tau}]} \right). \quad (3)
 \end{aligned}$$

 Involving a third intermediate averaged system means JSD is now finite:
 $0 \leq D^{\text{JS}}(P_1 \parallel P_2) \leq 1.$



🔗 Jensen-Shannon divergence (JSD): [9, 7, 13, 1]

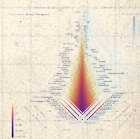
$$\begin{aligned}
 D^{\text{JS}}(P_1 \parallel P_2) &= \frac{1}{2} D^{\text{KL}}\left(P_1 \parallel \frac{1}{2}[P_1 + P_2]\right) + \frac{1}{2} D^{\text{KL}}\left(P_2 \parallel \frac{1}{2}[P_1 + P_2]\right) \\
 &= \frac{1}{2} \sum_{\tau \in R_{1,2;\alpha}} \left(p_{1,\tau} \log_2 \frac{p_{1,\tau}}{\frac{1}{2}[p_{1,\tau} + p_{2,\tau}]} + p_{2,\tau} \log_2 \frac{p_{2,\tau}}{\frac{1}{2}[p_{1,\tau} + p_{2,\tau}]} \right).
 \end{aligned} \tag{3}$$

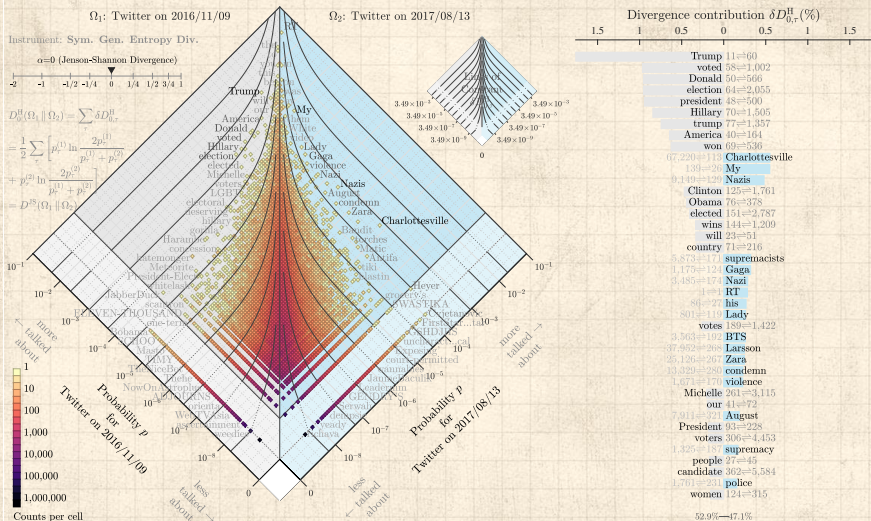
🔗 Involving a third intermediate averaged system means JSD is now finite:
 $0 \leq D^{\text{JS}}(P_1 \parallel P_2) \leq 1$.

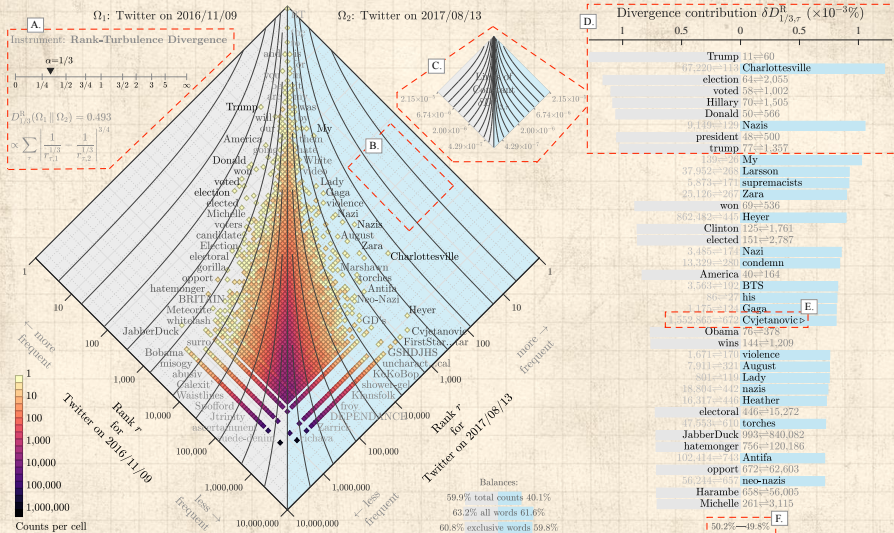
🔗 Generalized entropy divergence: [2]

$$\begin{aligned}
 D_{\alpha}^{\text{AS2}}(P_1 \parallel P_2) &= \\
 \frac{1}{\alpha(\alpha-1)} \sum_{\tau \in R_{1,2;\alpha}} &\left[\left(p_{\tau,1}^{1-\alpha} + p_{\tau,2}^{1-\alpha} \right) \left(\frac{p_{\tau,1} + p_{\tau,2}}{2} \right)^{\alpha} - (p_{\tau,1} + p_{\tau,2}) \right].
 \end{aligned} \tag{4}$$

Produces JSD when $\alpha \rightarrow 0$.







A plenitude of
distances

Rank-turbulence
divergence

Probability-
turbulence divergence

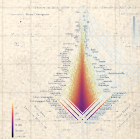
Explorations

Nutshell

References

Desirable rank-turbulence divergence features:

1. Rank-based.



A plenitude of
distances

Rank-turbulence
divergence

Probability-
turbulence divergence

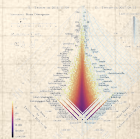
Explorations

Nutshell

References

Desirable rank-turbulence divergence features:

1. Rank-based.
2. Symmetric.



A plenitude of
distances

Rank-turbulence
divergence

Probability-
turbulence divergence

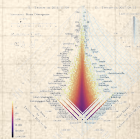
Explorations

Nutshell

References

Desirable rank-turbulence divergence features:

1. Rank-based.
2. Symmetric.
3. Semi-positive: $D_{\alpha}^R(\Omega_1 \parallel \Omega_2) \geq 0$.



A plenitude of
distances

Rank-turbulence
divergence

Probability-
turbulence divergence

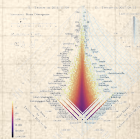
Explorations

Nutshell

References

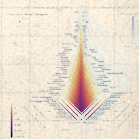
Desirable rank-turbulence divergence features:

1. Rank-based.
2. Symmetric.
3. Semi-positive: $D_{\alpha}^R(\Omega_1 \parallel \Omega_2) \geq 0$.
4. Linearly separable, for interpretability.



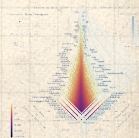
Desirable rank-turbulence divergence features:

1. Rank-based.
2. Symmetric.
3. Semi-positive: $D_{\alpha}^R(\Omega_1 \parallel \Omega_2) \geq 0$.
4. Linearly separable, for interpretability.
5. Subsystem applicable: Ranked lists of any principled subset may be equally well compared (e.g., hashtags on Twitter, stock prices of a certain sector, etc.).



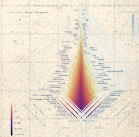
Desirable rank-turbulence divergence features:

1. Rank-based.
2. Symmetric.
3. Semi-positive: $D_{\alpha}^R(\Omega_1 \parallel \Omega_2) \geq 0$.
4. Linearly separable, for interpretability.
5. Subsystem applicable: Ranked lists of any principled subset may be equally well compared (e.g., hashtags on Twitter, stock prices of a certain sector, etc.).
6. Turbulence-handling: Suited for systems with rank-ordered component size distribution that are heavy-tailed.



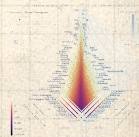
Desirable rank-turbulence divergence features:

1. Rank-based.
2. Symmetric.
3. Semi-positive: $D_{\alpha}^R(\Omega_1 \parallel \Omega_2) \geq 0$.
4. Linearly separable, for interpretability.
5. Subsystem applicable: Ranked lists of any principled subset may be equally well compared (e.g., hashtags on Twitter, stock prices of a certain sector, etc.).
6. Turbulence-handling: Suited for systems with rank-ordered component size distribution that are heavy-tailed.
7. Scalable: Allow for sensible comparisons across system sizes.



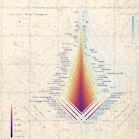
Desirable rank-turbulence divergence features:

1. Rank-based.
2. Symmetric.
3. Semi-positive: $D_{\alpha}^R(\Omega_1 \parallel \Omega_2) \geq 0$.
4. Linearly separable, for interpretability.
5. Subsystem applicable: Ranked lists of any principled subset may be equally well compared (e.g., hashtags on Twitter, stock prices of a certain sector, etc.).
6. Turbulence-handling: Suited for systems with rank-ordered component size distribution that are heavy-tailed.
7. Scalable: Allow for sensible comparisons across system sizes.
8. Tunable.



Desirable rank-turbulence divergence features:

1. Rank-based.
2. Symmetric.
3. Semi-positive: $D_{\alpha}^R(\Omega_1 \parallel \Omega_2) \geq 0$.
4. Linearly separable, for interpretability.
5. Subsystem applicable: Ranked lists of any principled subset may be equally well compared (e.g., hashtags on Twitter, stock prices of a certain sector, etc.).
6. Turbulence-handling: Suited for systems with rank-ordered component size distribution that are heavy-tailed.
7. Scalable: Allow for sensible comparisons across system sizes.
8. Tunable.
9. Story-finding: Features 1–8 combine to show which component types are most ‘important’



Some good things about ranks:

The PoCServe
Allotaxonomy
28 of 73

A plenitude of
distances

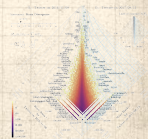
Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

Nutshell

References



Some good things about ranks:



Working with ranks is intuitive

The PoCSverse
Allotaxonomy
28 of 73

A plenitude of
distances

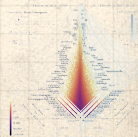
Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

Nutshell

References



Some good things about ranks:



Working with ranks is intuitive



Affords some powerful statistics (e.g., Spearman's rank correlation coefficient)

A plenitude of
distances

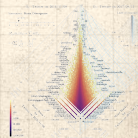
Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

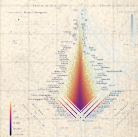
Nutshell

References



Some good things about ranks:

- Working with ranks is intuitive
- Affords some powerful statistics (e.g., Spearman's rank correlation coefficient)
- Can be used to generalize beyond systems with probabilities



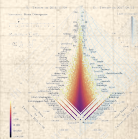
Some good things about ranks:

- Working with ranks is intuitive
- Affords some powerful statistics (e.g., Spearman's rank correlation coefficient)
- Can be used to generalize beyond systems with probabilities

A start:

$$\left| \frac{1}{r_{\tau,1}} - \frac{1}{r_{\tau,2}} \right|. \quad (5)$$

- Inverse of rank gives an increasing measure of 'importance'
- High rank means closer to rank 1
- We assign tied ranks for components of equal 'size'



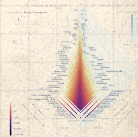
Some good things about ranks:

- Working with ranks is intuitive
- Affords some powerful statistics (e.g., Spearman's rank correlation coefficient)
- Can be used to generalize beyond systems with probabilities

A start:

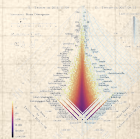
$$\left| \frac{1}{r_{\tau,1}} - \frac{1}{r_{\tau,2}} \right|. \quad (5)$$

- Inverse of rank gives an increasing measure of 'importance'
- High rank means closer to rank 1
- We assign tied ranks for components of equal 'size'
- Issue: Biases toward high rank components



We introduce a tuning parameter:

$$\left| \frac{1}{[r_{\tau,1}]^\alpha} - \frac{1}{[r_{\tau,2}]^\alpha} \right|^{1/\alpha} . \quad (6)$$

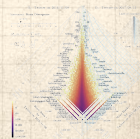


We introduce a tuning parameter:

$$\left| \frac{1}{[r_{\tau,1}]^{\alpha}} - \frac{1}{[r_{\tau,2}]^{\alpha}} \right|^{1/\alpha} . \quad (6)$$



As $\alpha \rightarrow 0$, high ranked components are increasingly dampened



We introduce a tuning parameter:

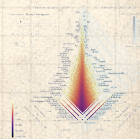
$$\left| \frac{1}{[r_{\tau,1}]^{\alpha}} - \frac{1}{[r_{\tau,2}]^{\alpha}} \right|^{1/\alpha} . \quad (6)$$



As $\alpha \rightarrow 0$, high ranked components are increasingly dampened



For words in texts, for example, the weight of common words and rare words move increasingly closer together.



We introduce a tuning parameter:

$$\left| \frac{1}{[r_{\tau,1}]^{\alpha}} - \frac{1}{[r_{\tau,2}]^{\alpha}} \right|^{1/\alpha} . \quad (6)$$



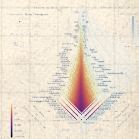
As $\alpha \rightarrow 0$, high ranked components are increasingly dampened



For words in texts, for example, the weight of common words and rare words move increasingly closer together.



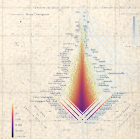
As $\alpha \rightarrow \infty$, high rank components will dominate.



We introduce a tuning parameter:

$$\left| \frac{1}{[r_{\tau,1}]^{\alpha}} - \frac{1}{[r_{\tau,2}]^{\alpha}} \right|^{1/\alpha} . \quad (6)$$

- As $\alpha \rightarrow 0$, high ranked components are increasingly dampened
- For words in texts, for example, the weight of common words and rare words move increasingly closer together.
- As $\alpha \rightarrow \infty$, high rank components will dominate.
- For texts, the contributions of rare words will vanish.



Trouble:



The limit of $\alpha \rightarrow 0$ does not behave well for

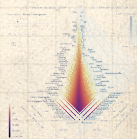
$$\left| \frac{1}{[r_{\tau,1}]^{\alpha}} - \frac{1}{[r_{\tau,2}]^{\alpha}} \right|^{1/\alpha}.$$



The leading order term is:

$$\left(1 - \delta_{r_{\tau,1} r_{\tau,2}}\right) \alpha^{1/\alpha} \left| \ln \frac{r_{\tau,1}}{r_{\tau,2}} \right|^{1/\alpha}, \quad (7)$$

which heads toward ∞ as $\alpha \rightarrow 0$.



Trouble:



The limit of $\alpha \rightarrow 0$ does not behave well for

$$\left| \frac{1}{[r_{\tau,1}]^\alpha} - \frac{1}{[r_{\tau,2}]^\alpha} \right|^{1/\alpha}.$$



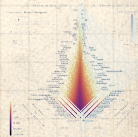
The leading order term is:

$$\left(1 - \delta_{r_{\tau,1} r_{\tau,2}}\right) \alpha^{1/\alpha} \left| \ln \frac{r_{\tau,1}}{r_{\tau,2}} \right|^{1/\alpha}, \quad (7)$$

which heads toward ∞ as $\alpha \rightarrow 0$.



Oops.



Trouble:



The limit of $\alpha \rightarrow 0$ does not behave well for

$$\left| \frac{1}{[r_{\tau,1}]^\alpha} - \frac{1}{[r_{\tau,2}]^\alpha} \right|^{1/\alpha}.$$



The leading order term is:

$$\left(1 - \delta_{r_{\tau,1} r_{\tau,2}}\right) \alpha^{1/\alpha} \left| \ln \frac{r_{\tau,1}}{r_{\tau,2}} \right|^{1/\alpha}, \quad (7)$$

which heads toward ∞ as $\alpha \rightarrow 0$.



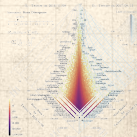
Oops.



But the insides look nutritious:

$$\left| \ln \frac{r_{\tau,1}}{r_{\tau,2}} \right|$$

is a nicely interpretable log-ratio of ranks.



Some reworking:

$$\delta D_{\alpha,\tau}^R(R_1 \parallel R_2) \propto \frac{\alpha+1}{\alpha} \left| \frac{1}{[r_{\tau,1}]^\alpha} - \frac{1}{[r_{\tau,2}]^\alpha} \right|^{1/(\alpha+1)}. \quad (8)$$

The PoCVerse
Allotaxonomy
31 of 73

A plenitude of
distances

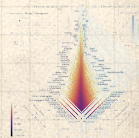
Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

Nutshell

References



Some reworking:

$$\delta D_{\alpha,\tau}^R(R_1 || R_2) \propto \frac{\alpha + 1}{\alpha} \left| \frac{1}{[r_{\tau,1}]^\alpha} - \frac{1}{[r_{\tau,2}]^\alpha} \right|^{1/(\alpha+1)}. \quad (8)$$



Keeps the core structure.

The PoCVerse
Allotaxonomy
31 of 73

A plenitude of
distances

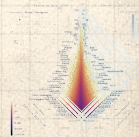
Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

Nutshell

References



Some reworking:

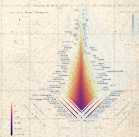
$$\delta D_{\alpha,\tau}^R(R_1 || R_2) \propto \frac{\alpha + 1}{\alpha} \left| \frac{1}{[r_{\tau,1}]^\alpha} - \frac{1}{[r_{\tau,2}]^\alpha} \right|^{1/(\alpha+1)}. \quad (8)$$



Keeps the core structure.



Large α limit remains the same.



Some reworking:

$$\delta D_{\alpha, \tau}^R(R_1 \parallel R_2) \propto \frac{\alpha + 1}{\alpha} \left| \frac{1}{[r_{\tau, 1}]^\alpha} - \frac{1}{[r_{\tau, 2}]^\alpha} \right|^{1/(\alpha+1)}. \quad (8)$$



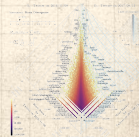
Keeps the core structure.



Large α limit remains the same.



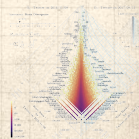
$\alpha \rightarrow 0$ limit now returns log-ratio of ranks.



Some reworking:

$$\delta D_{\alpha, \tau}^R(R_1 \parallel R_2) \propto \frac{\alpha + 1}{\alpha} \left| \frac{1}{[r_{\tau, 1}]^\alpha} - \frac{1}{[r_{\tau, 2}]^\alpha} \right|^{1/(\alpha+1)}. \quad (8)$$

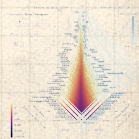
- Keeps the core structure.
- Large α limit remains the same.
- $\alpha \rightarrow 0$ limit now returns log-ratio of ranks.
- Next: Sum over τ to get divergence.



Some reworking:

$$\delta D_{\alpha,\tau}^R(R_1 \parallel R_2) \propto \frac{\alpha+1}{\alpha} \left| \frac{1}{[r_{\tau,1}]^\alpha} - \frac{1}{[r_{\tau,2}]^\alpha} \right|^{1/(\alpha+1)}. \quad (8)$$

-  Keeps the core structure.
-  Large α limit remains the same.
-  $\alpha \rightarrow 0$ limit now returns log-ratio of ranks.
-  Next: Sum over τ to get divergence.
-  Still have an option for normalization.



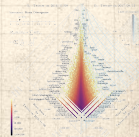
Some reworking:

$$\delta D_{\alpha,\tau}^R(R_1 \parallel R_2) \propto \frac{\alpha+1}{\alpha} \left| \frac{1}{[r_{\tau,1}]^\alpha} - \frac{1}{[r_{\tau,2}]^\alpha} \right|^{1/(\alpha+1)}. \quad (8)$$

-  Keeps the core structure.
-  Large α limit remains the same.
-  $\alpha \rightarrow 0$ limit now returns log-ratio of ranks.
-  Next: Sum over τ to get divergence.
-  Still have an option for normalization.

Rank-turbulence divergence:

$$D_{\alpha}^R(R_1 \parallel R_2) = \frac{1}{\mathcal{N}_{1,2;\alpha}} \sum_{\tau \in R_{1,2;\alpha}} \delta D_{\alpha,\tau}^R(R_1 \parallel R_2) \quad (9)$$



A plenitude of
distances

Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

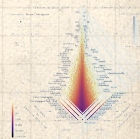
Nutshell

References

Normalization:

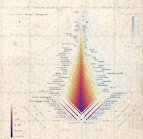


Take a data-driven rather than analytic approach to determining $\mathcal{N}_{1,2;\alpha}$.



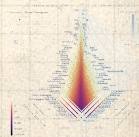
Normalization:

- Take a data-driven rather than analytic approach to determining $\mathcal{N}_{1,2;\alpha}$.
- Compute $\mathcal{N}_{1,2;\alpha}$ by taking the two systems to be disjoint while maintaining their underlying Zipf distributions.



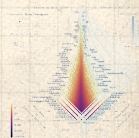
Normalization:

- Take a data-driven rather than analytic approach to determining $\mathcal{N}_{1,2;\alpha}$.
- Compute $\mathcal{N}_{1,2;\alpha}$ by taking the two systems to be disjoint while maintaining their underlying Zipf distributions.
- Ensures: $0 \leq D_{\alpha}^R(R_1 \parallel R_2) \leq 1$



Normalization:

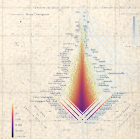
- Take a data-driven rather than analytic approach to determining $\mathcal{N}_{1,2;\alpha}$.
- Compute $\mathcal{N}_{1,2;\alpha}$ by taking the two systems to be disjoint while maintaining their underlying Zipf distributions.
- Ensures: $0 \leq D_{\alpha}^R(R_1 \parallel R_2) \leq 1$
- Limits of 0 and 1 correspond to the two systems having identical and disjoint Zipf distributions.



Rank-turbulence divergence:

Summing over all types, dividing by a normalization prefactor $\mathcal{N}_{1,2;\alpha}$ we have our prototype:

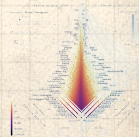
$$D_{\alpha}^R(R_1 \parallel R_2) = \frac{1}{\mathcal{N}_{1,2;\alpha}} \frac{\alpha + 1}{\alpha} \sum_{\tau \in R_{1,2;\alpha}} \left| \frac{1}{[r_{\tau,1}]^{\alpha}} - \frac{1}{[r_{\tau,2}]^{\alpha}} \right|^{1/(\alpha+1)} \quad (10)$$



General normalization:

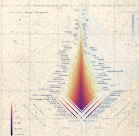


If the Zipf distributions are disjoint, then in $\Omega^{(1)}$'s merged ranking, the rank of all $\Omega^{(2)}$ types will be $r = N_1 + \frac{1}{2}N_2$, where N_1 and N_2 are the number of distinct types in each system.



General normalization:

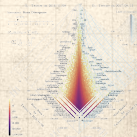
- ❏ If the Zipf distributions are disjoint, then in $\Omega^{(1)}$'s merged ranking, the rank of all $\Omega^{(2)}$ types will be $r = N_1 + \frac{1}{2}N_2$, where N_1 and N_2 are the number of distinct types in each system.
- ❏ Similarly, $\Omega^{(2)}$'s merged ranking will have all of $\Omega^{(1)}$'s types in last place with rank $r = N_2 + \frac{1}{2}N_1$.



General normalization:

- ✖ If the Zipf distributions are disjoint, then in $\Omega^{(1)}$'s merged ranking, the rank of all $\Omega^{(2)}$ types will be $r = N_1 + \frac{1}{2}N_2$, where N_1 and N_2 are the number of distinct types in each system.
- ✖ Similarly, $\Omega^{(2)}$'s merged ranking will have all of $\Omega^{(1)}$'s types in last place with rank $r = N_2 + \frac{1}{2}N_1$.
- ✖ The normalization is then:

$$\begin{aligned} \mathcal{N}_{1,2;\alpha} = & \frac{\alpha+1}{\alpha} \sum_{\tau \in R_1} \left| \frac{1}{[r_{\tau,1}]^\alpha} - \frac{1}{[N_1 + \frac{1}{2}N_2]^\alpha} \right|^{1/(\alpha+1)} \\ & + \frac{\alpha+1}{\alpha} \sum_{\tau \in R_2} \left| \frac{1}{[N_2 + \frac{1}{2}N_1]^\alpha} - \frac{1}{[r_{\tau,2}]^\alpha} \right|^{1/(\alpha+1)}. \end{aligned} \quad (11)$$



Limit of $\alpha \rightarrow 0$:

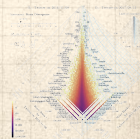
$$D_0^R(R_1 \parallel R_2) = \sum_{\tau \in R_{1,2;\alpha}} \delta D_{0,\tau}^R = \frac{1}{\mathcal{N}_{1,2;0}} \sum_{\tau \in R_{1,2;\alpha}} \left| \ln \frac{r_{\tau,1}}{r_{\tau,2}} \right|, \quad (12)$$

where

$$\mathcal{N}_{1,2;0} = \sum_{\tau \in R_1} \left| \ln \frac{r_{\tau,1}}{N_1 + \frac{1}{2}N_2} \right| + \sum_{\tau \in R_2} \left| \ln \frac{r_{\tau,2}}{\frac{1}{2}N_1 + N_2} \right|. \quad (13)$$



Largest rank ratios dominate.



Limit of $\alpha \rightarrow \infty$:

$$D_{\infty}^R(R_1 \parallel R_2) = \sum_{\tau \in R_{1,2;\alpha}} \delta D_{\infty,\tau}^R$$

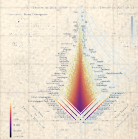
$$= \frac{1}{\mathcal{N}_{1,2;\infty}} \sum_{\tau \in R_{1,2;\infty}} \left(1 - \delta_{r_{\tau,1} r_{\tau,2}}\right) \max_{\tau} \left\{ \frac{1}{r_{\tau,1}}, \frac{1}{r_{\tau,2}} \right\}. \quad (14)$$

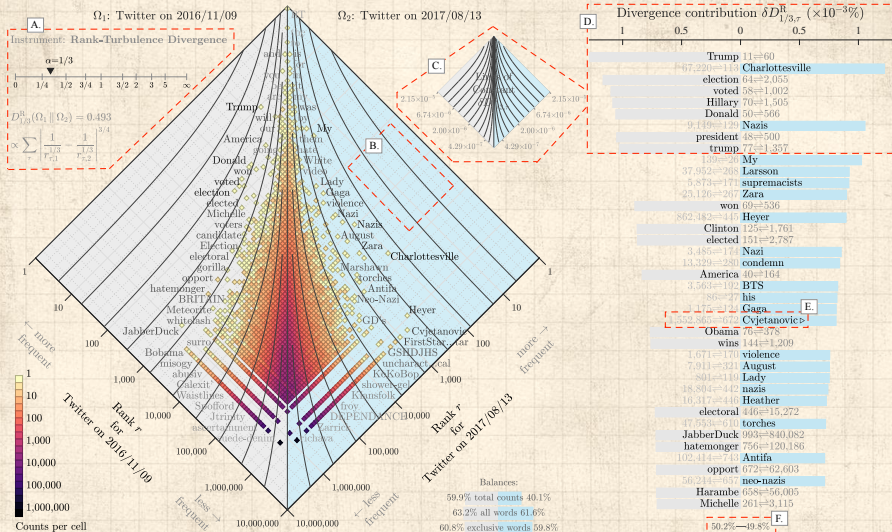
where

$$\mathcal{N}_{1,2;\infty} = \sum_{\tau \in R_1} \frac{1}{r_{\tau,1}} + \sum_{\tau \in R_2} \frac{1}{r_{\tau,2}}. \quad (15)$$



Highest ranks dominate.





Probability-turbulence divergence:

$$D_{\alpha}^p(P_1 \parallel P_2) = \frac{1}{\mathcal{N}_{1,2;\alpha}^p} \frac{\alpha + 1}{\alpha} \sum_{\tau \in R_{1,2;\alpha}} \left| [p_{\tau,1}]^{\alpha} - [p_{\tau,2}]^{\alpha} \right|^{1/(\alpha+1)}. \quad (16)$$

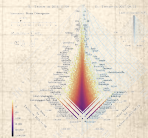
 For the unnormalized version ($\mathcal{N}_{1,2;\alpha}^p=1$), some troubles return with 0 probabilities and $\alpha \rightarrow 0$.

 Weep not: $\mathcal{N}_{1,2;\alpha}^p$ will save the day.

Normalization:

With no matching types, the probability of a type present in one system is zero in the other, and the sum can be split between the two systems' types:

$$\mathcal{N}_{1,2;\alpha}^P = \frac{\alpha + 1}{\alpha} \sum_{\tau \in R_1} [p_{\tau,1}]^{\alpha/(\alpha+1)} + \frac{\alpha + 1}{\alpha} \sum_{\tau \in R_2} [p_{\tau,2}]^{\alpha/(\alpha+1)} \quad (17)$$

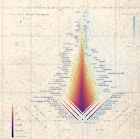


Limit of $\alpha = 0$ for probability-turbulence divergence

if both $p_{\tau,1} > 0$ and $p_{\tau,2} > 0$ then

$$\lim_{\alpha \rightarrow 0} \frac{\alpha + 1}{\alpha} \left| [p_{\tau,1}]^{\alpha} - [p_{\tau,2}]^{\alpha} \right|^{1/(\alpha+1)} = \left| \ln \frac{p_{\tau,2}}{p_{\tau,1}} \right|. \quad (18)$$

But if $p_{\tau,1} = 0$ or $p_{\tau,2} = 0$, limit diverges as $1/\alpha$.

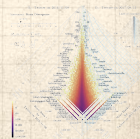


Limit of $\alpha=0$ for probability-turbulence divergence

 Normalization:

$$\mathcal{N}_{1,2;\alpha}^p \rightarrow \frac{1}{\alpha} (N_1 + N_2) . \quad (19)$$

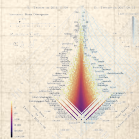
 Because the normalization also diverges as $1/\alpha$, the divergence will be zero when there are no exclusive types and non-zero when there are exclusive types.



Combine these cases into a single expression:

$$D_0^p(P_1 \parallel P_2) = \frac{1}{(N_1 + N_2)} \sum_{\tau \in R_{1,2;0}} (\delta_{p_{\tau,1},0} + \delta_{0,p_{\tau,2}}). \quad (20)$$

-  The term $(\delta_{p_{\tau,1},0} + \delta_{0,p_{\tau,2}})$ returns 1 if either $p_{\tau,1} = 0$ or $p_{\tau,2} = 0$, and 0 otherwise when both $p_{\tau,1} > 0$ and $p_{\tau,2} > 0$.
-  Ratio of types that are exclusive to one system relative to the total possible such types,

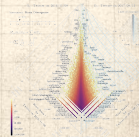


Type contribution ordering for the limit of $\alpha=0$

✉ In terms of contribution to the divergence score, all exclusive types supply a weight of $1/(N_1 + N_2)$. We can order them by preserving their ordering as $\alpha \rightarrow 0$, which amounts to ordering by descending probability in the system in which they appear.

✉ And while types that appear in both systems make no contribution to $D_0^P(P_1 \parallel P_2)$, we can still order them according to the log ratio of their probabilities.

✉ The overall ordering of types by divergence contribution for $\alpha=0$ is then: (1) exclusive types by descending probability and then (2) types appearing in both systems by descending log ratio.

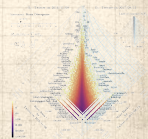


Limit of $\alpha=\infty$ for probability-turbulence divergence

$$D_{\infty}^p(P_1 \parallel P_2) = \frac{1}{2} \sum_{\tau \in R_{1,2;\infty}} (1 - \delta_{p_{\tau,1}, p_{\tau,2}}) \max(p_{\tau,1}, p_{\tau,2}) \quad (21)$$

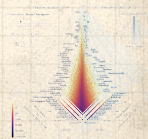
where

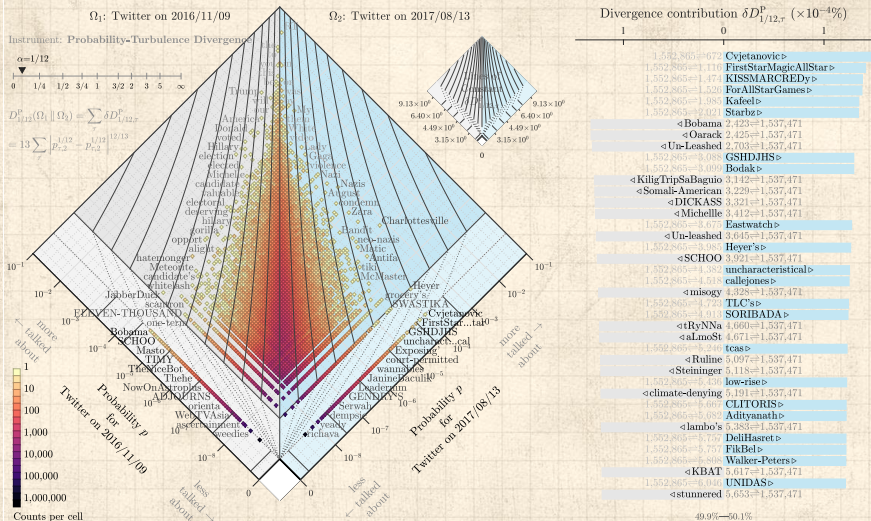
$$\mathcal{N}_{1,2;\infty}^p = \sum_{\tau \in R_{1,2;\infty}} (p_{\tau,1} + p_{\tau,2}) = 1 + 1 = 2. \quad (22)$$

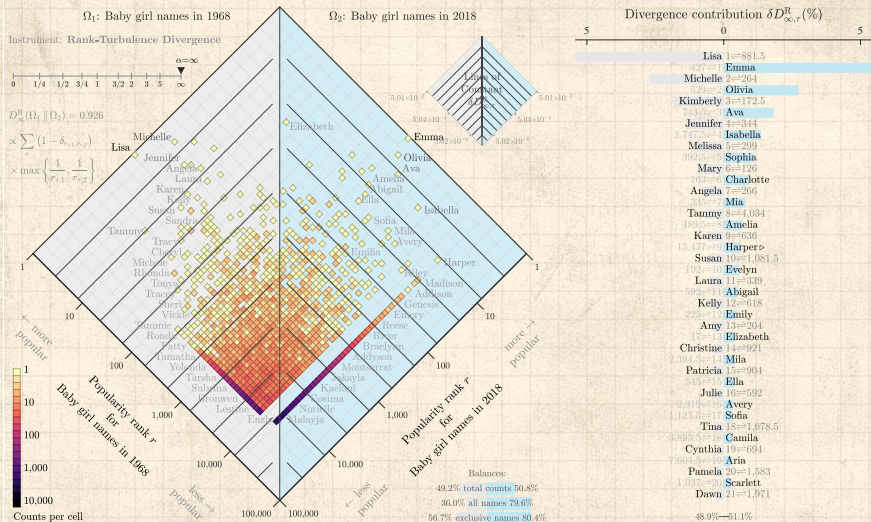


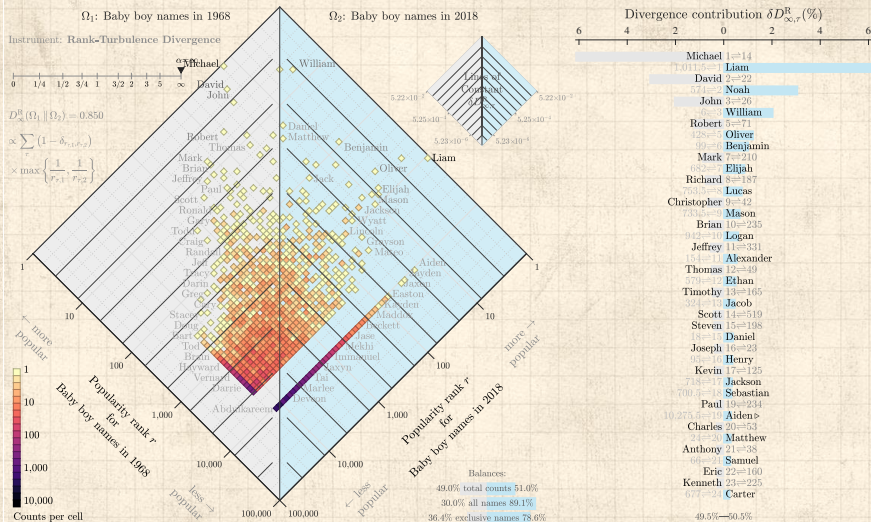
Connections for PTD:

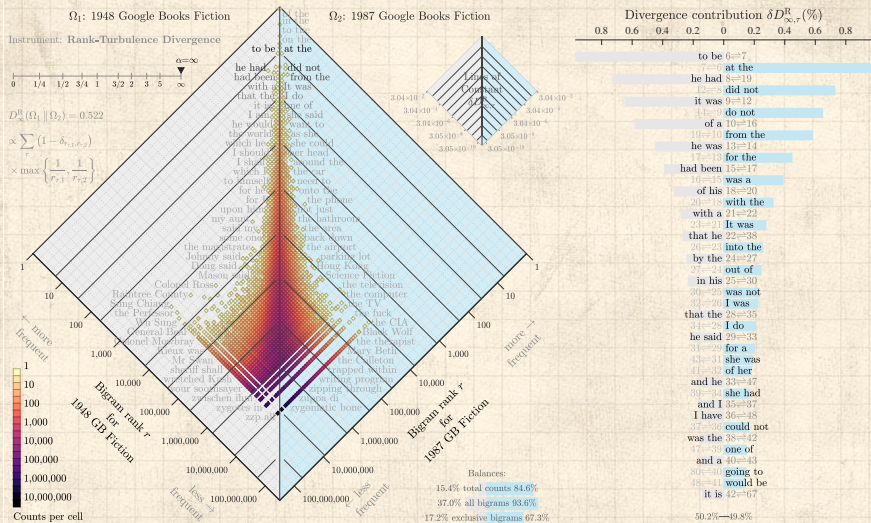
-  $\alpha = 0$: Similarity measure Sørensen-Dice coefficient ^[4, 17, 10], F_1 score of a test's accuracy ^[18, 15].
-  $\alpha = 1/2$: Hellinger distance ^[8] and Mautusita distance ^[11].
-  $\alpha = 1$: Many including all $L^{(p)}$ -norm type constructions.
-  $\alpha = \infty$: Motyka distance ^[3].



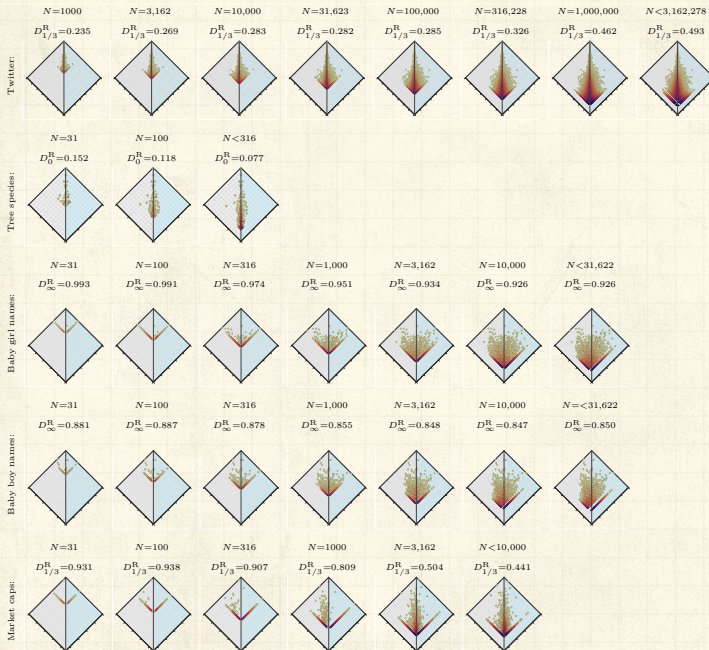








Effect of subsampling:



The PoCserve
Allotaxonomy
54 of 73

A plenitude of
distances

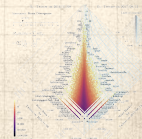
Rank-turbulence
divergence

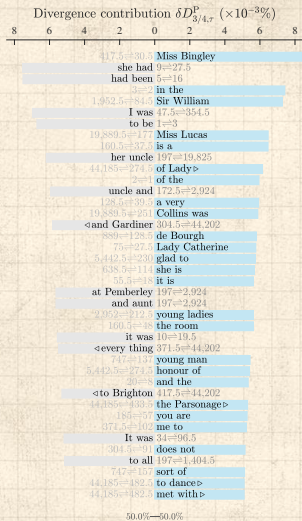
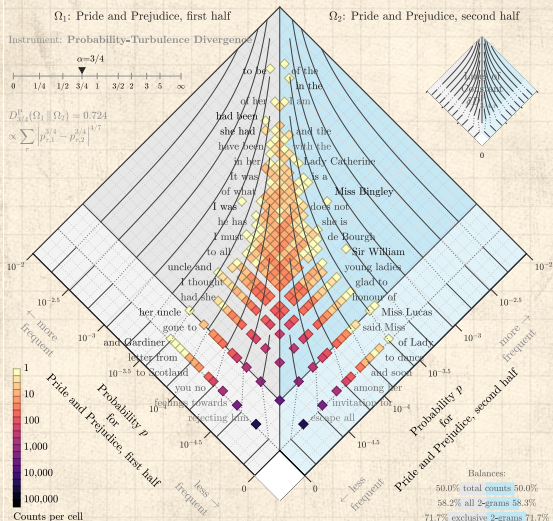
Probability-
turbulence divergence

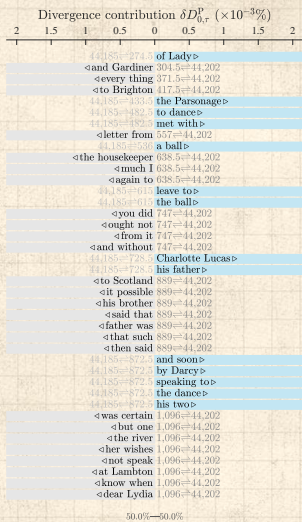
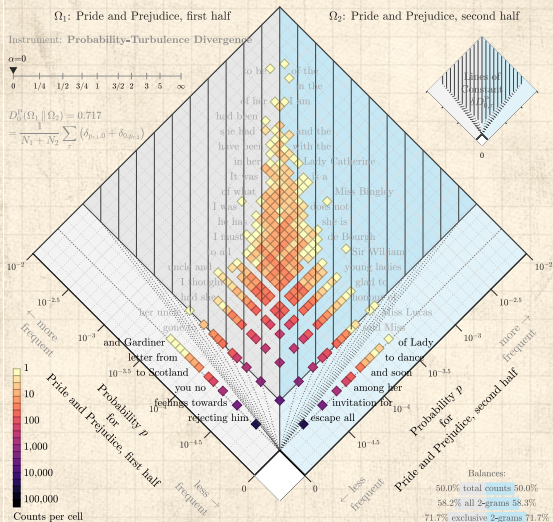
Explorations

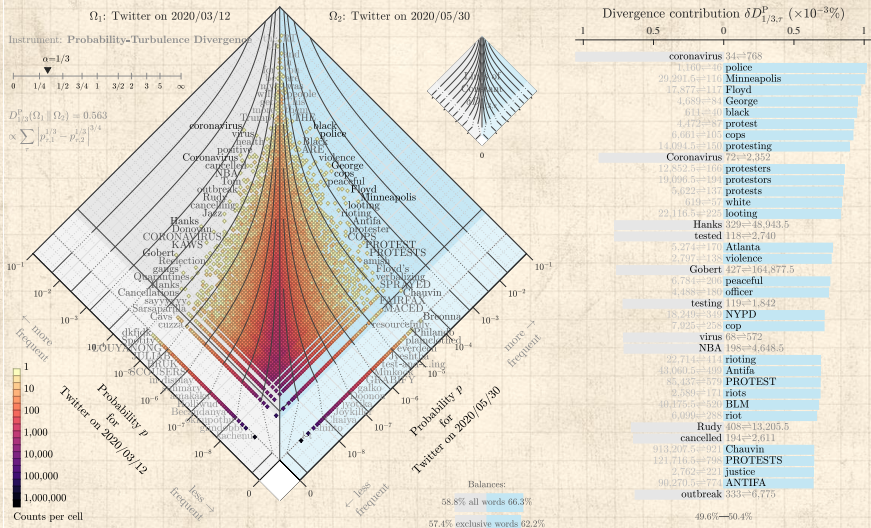
Nutshell

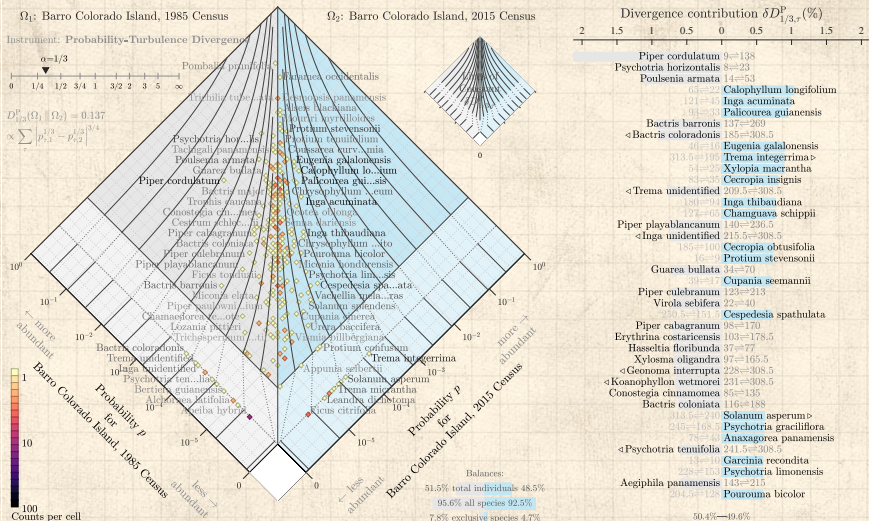
References











Flipbooks for RTD:



Twitter:

[allotaxonometer-flipbook-1-rank-div.pdf](#)  

[allotaxonometer-flipbook-2-probability-div.pdf](#)  

[allotaxonometer-flipbook-3-gen-entropy-div.pdf](#)  



Market caps:

[allotaxonometer-flipbook-4-marketcaps-6years-rank-div.pdf](#)  



Baby names:

[allotaxonometer-flipbook-5-babynames-girls-50years-rank-div.pdf](#)  

[allotaxonometer-flipbook-6-babynames-boys-50years-rank-div.pdf](#)  

[Baby girl names over time relative to 1950](#)  

[Baby boy names over time relative to 1950](#)  



Google books:

[allotaxonometer-flipbook-7-google-books-onegrams-rank-div.pdf](#)  

[allotaxonometer-flipbook-8-google-books-bigrams-rank-div.pdf](#)  

[allotaxonometer-flipbook-9-google-books-trigrams-rank-div.pdf](#)  

Flipbooks for PTD:



Jane Austen:

Pride and Prejudice, 1-grams  

Pride and Prejudice, 2-grams  

Pride and Prejudice, 3-grams  



Social media:

Twitter, 1-grams  

Twitter, 2-grams  

Twitter, 3-grams  



Ecology:

Barro Colorado Island  

A plenitude of
distances

Rank-turbulence
divergence

Probability-
turbulence divergence

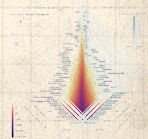
Explorations

Nutshell

References

Code:

<https://gitlab.com/compstorylab/allotaxonometer>



Claims, exaggerations, reminders:



Needed for comparing large-scale complex systems:
Comprehensible, dynamically-adjusting, differential
dashboards.

The PoCserve
Allotaxonomy
66 of 73

A plenitude of
distances

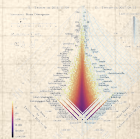
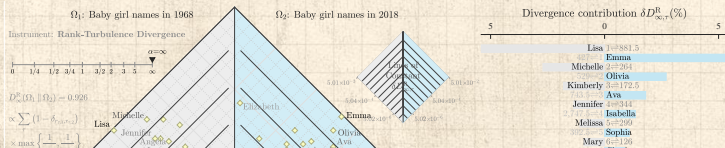
Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

Nutshell

References



Claims, exaggerations, reminders:



Needed for comparing large-scale complex systems:
Comprehensible, dynamically-adjusting, differential
dashboards.



Many measures seem poorly motivated and largely
unexamined (e.g., JSD).

The PoCverse
Allotaxonomy
66 of 73

A plenitude of
distances

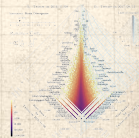
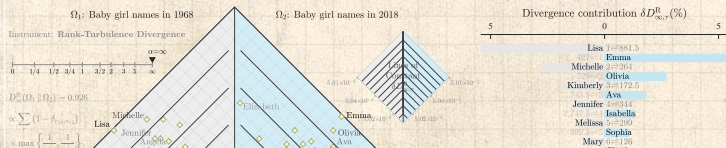
Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

Nutshell

References



Claims, exaggerations, reminders:



Needed for comparing large-scale complex systems:
Comprehensible, dynamically-adjusting, differential
dashboards.



Many measures seem poorly motivated and largely
unexamined (e.g., JSD).



Of value: Combining big-picture maps with ranked lists.

The PoCverse
Allotaxonomy
66 of 73

A plenitude of
distances

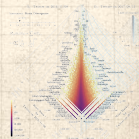
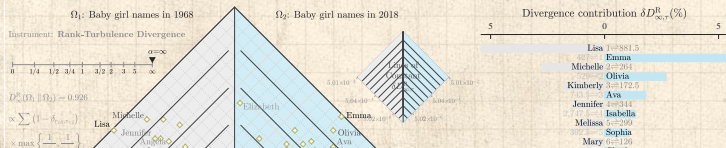
Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

Nutshell

References



Claims, exaggerations, reminders:

-  Needed for comparing large-scale complex systems: Comprehensible, dynamically-adjusting, differential dashboards.
-  Many measures seem poorly motivated and largely unexamined (e.g., JSD).
-  Of value: Combining big-picture maps with ranked lists.
-  Online tunable versions of rank-turbulence divergence now exist:

 App version: <https://allotax.vercel.app/> 

 Observable version:
<https://observablehq.com/@jstonge/allotaxonometer-4-all> 

 Github: <https://github.com/jstonge/allotaxp> 

The PoCverse
Allotaxonomy
66 of 73

A plenitude of
distances

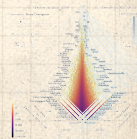
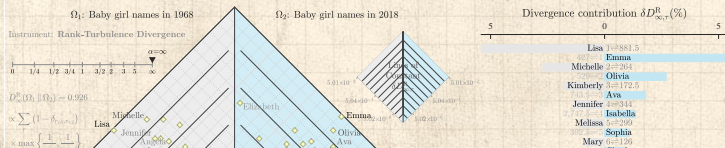
Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

Nutshell

References



Claims, exaggerations, reminders:

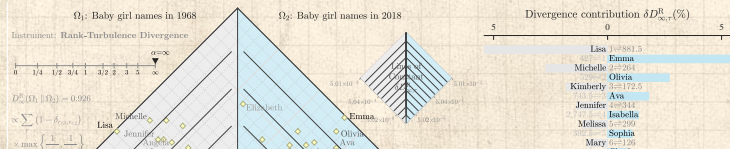
- Needed for comparing large-scale complex systems: Comprehensible, dynamically-adjusting, differential dashboards.
- Many measures seem poorly motivated and largely unexamined (e.g., JSD).
- Of value: Combining big-picture maps with ranked lists.
- Online tunable versions of rank-turbulence divergence now exist:

App version: <https://allotax.vercel.app/>

Observable version: <https://observablehq.com/@jstonge/allotaxonometer-4-all>

Github: <https://github.com/jstonge/allotaxp>

- Future: Could do all probability based instruments



The PoCverse
Allotaxonomy
66 of 73

A plenitude of
distances

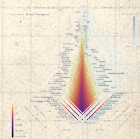
Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

Nutshell

References



References I

- [1] S.-H. Cha.
Comprehensive survey on distance/similarity measures
between probability density functions.
International Journal of Mathematical Models and Methods
in Applied Sciences, 1:300–307, 2007. pdf ↗
- [2] A. Cichocki and S.-i. Amari.
Families of Alpha- Beta- and Gamma- divergences: Flexible
and robust measures of similarities.
Entropy, 12:1532–1568, 2010. pdf ↗
- [3] M.-M. Deza and E. Deza.
Dictionary of Distances.
Elsevier, 2006.

The PoCverse
Allotaxonomy
67 of 73

A plenitude of
distances

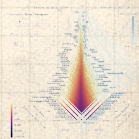
Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

Nutshell

References



References II

- [4] L. R. Dice.
Measures of the amount of ecologic association between species.
Ecology, 26:297–302, 1945.
- [5] P. S. Dodds, J. R. Minot, M. V. Arnold, T. Alshaabi, J. L. Adams, D. R. Dewhurst, T. J. Gray, M. R. Frank, A. J. Reagan, and C. M. Danforth.
Allotaxonomy and rank-turbulence divergence: A universal instrument for comparing complex systems.
EPJ Data Science, 12(37):1–42, 2023.
arXiv version: <https://arxiv.org/abs/2002.09770>. pdf ↗

The PoCverse
Allotaxonomy
68 of 73

A plenitude of
distances

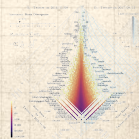
Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

Nutshell

References



References III

- [6] P. S. Dodds, J. R. Minot, M. V. Arnold, T. Alshaabi, J. L. Adams, A. J. Reagan, and C. M. Danforth.

Probability-turbulence divergence: A tunable allotaxonomic instrument for comparing heavy-tailed type-size distributions, 2020.

Available online at <https://arxiv.org/abs/2008.13078>. pdf ↗

- [7] D. M. Endres and J. E. Schindelin.

A new metric for probability distributions.

IEEE Transactions on Information Theory, 49(7):1858–1860, 2003. pdf ↗

- [8] E. Hellinger.

Neue begründung der theorie quadratischer formen von unendlichvielen veränderlichen.

Journal für die reine und angewandte Mathematik (Crelles Journal), 1909(136):210–271, 1909. pdf ↗

The PoCserve
Allotaxonomy
69 of 73

A plenitude of
distances

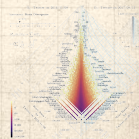
Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

Nutshell

References



References IV

[9] J. Lin.

Divergence measures based on the Shannon entropy.

IEEE Transactions on Information theory, 37(1):145–151,
1991. pdf ↗

[10] J. Looman and J. B. Campbell.

Adaptation of Sørensen's k (1948) for estimating unit
affinities in prairie vegetation.

Ecology, 41(3):409–416, 1960. pdf ↗

[11] K. Matusita et al.

Decision rules, based on the distance, for problems of fit, two
samples, and estimation.

The Annals of Mathematical Statistics, 26(4):631–640, 1955.
pdf ↗

The PoCverse
Allotaxonomy
70 of 73

A plenitude of
distances

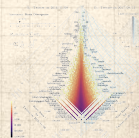
Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

Nutshell

References



References V

[12] R. Munroe.

How To: Absurd Scientific Advice for Common Real-World Problems.

Penguin, 2019.

[13] F. Osterreicher and I. Vajda.

A new class of metric divergences on probability spaces and its applicability in statistics.

Annals of the Institute of Statistical Mathematics,
55(3):639–653, 2003.

[14] E. A. Pechenick, C. M. Danforth, and P. S. Dodds.

Is language evolution grinding to a halt? The scaling of lexical turbulence in English fiction suggests it is not.

Journal of Computational Science, 21:24–37, 2017. pdf 

[15] Y. Sasaki.

The truth of the f -measure, 2007.

The PoCverse
Allotaxonomy
71 of 73

A plenitude of
distances

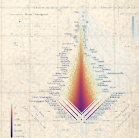
Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

Nutshell

References



References VI

- [16] C. E. Shannon.
The bandwagon.
IRE Transactions on Information Theory, 2(1):3, 1956.
[pdf](#) 
- [17] T. Sorensen.
A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons.
Videnski Selskab Biologiske Skrifter, 5:1–34, 1948.
- [18] C. J. Van Rijsbergen.
Information retrieval.
Butterworth-Heinemann, 2nd edition, 1979.

The PoCverse
Allotaxonomy
72 of 73

A plenitude of
distances

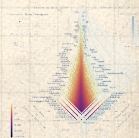
Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

Nutshell

References



References VII

The PoCVerse
Allotaxonomy
73 of 73

A plenitude of
distances

Rank-turbulence
divergence

Probability-
turbulence divergence

Explorations

Nutshell

References

[19] J. R. Williams, J. P. Bagrow, C. M. Danforth, and P. S. Dodds.

Text mixing shapes the anatomy of rank-frequency
distributions.

Physical Review E, 91:052811, 2015. pdf ↗

