Multivariate Symmetry: Distribution-free Testing via Optimal Transport

Bodhisattva Sen[1]
Department of Statistics
Columbia University, New York

Kantorovich Initiative Seminar

13 April, 2023

[1]Supported by NSF grant DMS-2015376

Collaborator

Zhen Huang
PhD student (Columbia University, Statistics)

Testing for symmetry

- **Data**: $\{X_i\}_{i=1}^n$ iid $X \sim P$ (abs. cont.) on $\mathbb{R}$
- **Test** the hypothesis of symmetry, i.e.,

$$\mathrm{H}_0 : X \overset{d}{=} -X \qquad \text{versus} \qquad \mathrm{H}_1 : \text{not } \mathrm{H}_0$$

Testing for symmetry

- **Data**: $\{X_i\}_{i=1}^n$ iid $X \sim P$ (abs. cont.) on $\mathbb{R}$
- **Test** the hypothesis of symmetry, i.e.,

$$\mathrm{H}_0 : X \stackrel{d}{=} -X \qquad \text{versus} \qquad \mathrm{H}_1 : \text{not } \mathrm{H}_0$$

Distribution-free testing for symmetry

- Sign test [Arbuthnot (1710)]: "...the first use of significance tests..." (first nonparametric test)
- Wilcoxon signed-rank (WSR) test [Wilcoxon (1945)]: Created the field of (classical) nonparametrics

Testing for symmetry

- **Data**: $\{X_i\}_{i=1}^n$ iid $X \sim P$ (abs. cont.) on $\mathbb{R}$
- **Test** the hypothesis of symmetry, i.e.,

$$\mathrm{H}_0 : X \overset{d}{=} -X \qquad \text{versus} \qquad \mathrm{H}_1 : \text{not } \mathrm{H}_0$$

Distribution-free testing for symmetry

- Sign test [Arbuthnot (1710)]: "...the first use of significance tests..." (first nonparametric test)
- Wilcoxon signed-rank (WSR) test [Wilcoxon (1945)]: Created the field of (classical) nonparametrics
- Arises with paired (matched) data; when normality can be violated

Testing for symmetry

- **Data**: $\{X_i\}_{i=1}^n$ iid $X \sim P$ (abs. cont.) on $\mathbb{R}$
- **Test** the hypothesis of symmetry, i.e.,

$$\mathrm{H}_0 : X \overset{d}{=} -X \qquad \text{versus} \qquad \mathrm{H}_1 : \text{not } \mathrm{H}_0$$

Distribution-free testing for symmetry

- Sign test [Arbuthnot (1710)]: "...the first use of significance tests..." (first nonparametric test)
- Wilcoxon signed-rank (WSR) test [Wilcoxon (1945)]: Created the field of (classical) nonparametrics
- Arises with paired (matched) data; when normality can be violated

Long history: Arbuthnot (1710), Wilcoxon (1945), Hodges & Lehmann (1956), Chernoff & Savage (1958), McWilliams (1990) ...

Goal: Develop distribution-free testing for multivariate symmetry

Multivariate symmetry

There are many notions of symmetry in $\mathbb{R}^p$, for $p \geq 2$

Multivariate symmetry

There are many notions of symmetry in $\mathbb{R}^p$, for $p \geq 2$

- **Central**: Test $\mathrm{H}_0 : \mathbf{X} \stackrel{d}{=} -\mathbf{X}$

Multivariate symmetry

There are many notions of symmetry in $\mathbb{R}^p$, for $p \geq 2$

- **Central**: Test $\mathrm{H}_0 : \mathbf{X} \overset{d}{=} -\mathbf{X}$
- **Sign**: Test $\mathrm{H}_0 : \mathbf{X} \overset{d}{=} D\mathbf{X}$, $D = \mathrm{diag}(\pm 1, \ldots, \pm 1) \in \mathbb{R}^{p \times p}$

Multivariate symmetry

There are many notions of symmetry in $\mathbb{R}^p$, for $p \geq 2$

- **Central**: Test $\mathrm{H}_0 : \mathbf{X} \stackrel{d}{=} -\mathbf{X}$
- **Sign**: Test $\mathrm{H}_0 : \mathbf{X} \stackrel{d}{=} D\mathbf{X}$, $D = \mathrm{diag}(\pm 1, \ldots, \pm 1) \in \mathbb{R}^{p \times p}$
- **Spherical**: Test $\mathrm{H}_0 : \mathbf{X} \stackrel{d}{=} Q\mathbf{X}$, $Q \in \mathbb{R}^{p \times p}$ is any orthogonal matrix

Multivariate symmetry

There are many notions of symmetry in $\mathbb{R}^p$, for $p \geq 2$

- **Central**: Test $\mathrm{H}_0 : \mathbf{X} \stackrel{d}{=} -\mathbf{X}$
- **Sign**: Test $\mathrm{H}_0 : \mathbf{X} \stackrel{d}{=} D\mathbf{X}$, $D = \mathrm{diag}(\pm 1, \ldots, \pm 1) \in \mathbb{R}^{p \times p}$
- **Spherical**: Test $\mathrm{H}_0 : \mathbf{X} \stackrel{d}{=} Q\mathbf{X}$, $Q \in \mathbb{R}^{p \times p}$ is any orthogonal matrix

- $\mathrm{O}(p)$: group of all orthogonal matrices on $\mathbb{R}^{p \times p}$
- $\mathcal{G}$: compact subgroup of $\mathrm{O}(p)$
- **Goal**: Develop distribution-free testing for $\mathcal{G}$-symmetry, i.e.,

$$\mathrm{H}_0 : \mathbf{X} \stackrel{d}{=} Q\mathbf{X} \quad \forall\, Q \in \mathcal{G}, \qquad \text{versus} \qquad \mathrm{H}_1 : \text{not } \mathrm{H}_0$$

Multivariate symmetry

There are many notions of symmetry in $\mathbb{R}^p$, for $p \geq 2$

- **Central**: Test $\mathrm{H}_0 : \mathbf{X} \stackrel{d}{=} -\mathbf{X}$
- **Sign**: Test $\mathrm{H}_0 : \mathbf{X} \stackrel{d}{=} D\mathbf{X}$, $D = \mathrm{diag}(\pm 1, \dots, \pm 1) \in \mathbb{R}^{p \times p}$
- **Spherical**: Test $\mathrm{H}_0 : \mathbf{X} \stackrel{d}{=} Q\mathbf{X}$, $Q \in \mathbb{R}^{p \times p}$ is any orthogonal matrix

- $\mathrm{O}(p)$: group of all orthogonal matrices on $\mathbb{R}^{p \times p}$
- $\mathcal{G}$: compact subgroup of $\mathrm{O}(p)$
- **Goal**: Develop distribution-free testing for $\mathcal{G}$-symmetry, i.e.,

$$\mathrm{H}_0 : \mathbf{X} \stackrel{d}{=} Q\mathbf{X} \quad \forall\, Q \in \mathcal{G}, \qquad \text{versus} \qquad \mathrm{H}_1 : \text{not } \mathrm{H}_0$$

Long history: Weyl (1952), Hodges (1955), Watson (1961), Bickel (1965), Randles (1989), Baringhaus (1991), Chaudhuri & Sengupta (1993), Beran & Millar (1997), Marden (1999), Zuo & Serfling (2000), Hallin & Paindaveine (2002), Oja (2010), Serfling (2014), ...

Data: $X_1, \ldots, X_n$ iid $X \sim P$ (X abs. cont.) on $\mathbb{R}$ (i.e., $p = 1$)

Goal: Distribution-free testing of $\mathrm{H}_0 : X \stackrel{d}{=} -X$

Sign test [Arbuthnot (1710)]

- **Sign**: $S_i := \begin{cases} +1 & \text{if } X_i \geq 0 \\ -1 & \text{if } X_i < 0 \end{cases}$ Under H_0, $S_i \stackrel{iid}{\sim} \pm 1$ w.p. $\frac{1}{2}$
- Rejects H_0 when $\sum_{i=1}^n S_i$ is significantly different from 0

Data: $X_1, \ldots, X_n$ iid $X \sim P$ (X abs. cont.) on $\mathbb{R}$ (i.e., $p = 1$)

Goal: Distribution-free testing of $\mathrm{H}_0 : X \overset{d}{=} -X$

Sign test [Arbuthnot (1710)]

- **Sign**: $S_i := \begin{cases} +1 & \text{if } X_i \geq 0 \\ -1 & \text{if } X_i < 0 \end{cases}$ Under H_0, $S_i \overset{iid}{\sim} \pm 1$ w.p. $\frac{1}{2}$
- Rejects H_0 when $\sum_{i=1}^n S_i$ is significantly different from 0
- Under H_0: $\frac{1}{2} \sum_{i=1}^n (S_i + 1) \sim \mathrm{Bin}(n, \frac{1}{2})$
- **Distribution-freeness**: The null distribution of $\sum_{i=1}^n S_i$ is universal — does not depend on the underlying distribution of the data
- Leads to an exact and distribution-free test valid for all sample sizes

Data: $X_1, \ldots, X_n$ iid $X \sim P$ (X abs. cont.) on $\mathbb{R}$ (i.e., $p = 1$)

Goal: Distribution-free testing of $\mathrm{H}_0 : X \stackrel{d}{=} -X$

Sign test [Arbuthnot (1710)]

- **Sign**: $S_i := \begin{cases} +1 & \text{if } X_i \geq 0 \\ -1 & \text{if } X_i < 0 \end{cases}$ Under H_0, $S_i \stackrel{iid}{\sim} \pm 1$ w.p. $\frac{1}{2}$
- Rejects H_0 when $\sum_{i=1}^n S_i$ is significantly different from 0
- Under H_0: $\frac{1}{2} \sum_{i=1}^n (S_i + 1) \sim \mathrm{Bin}(n, \frac{1}{2})$
- **Distribution-freeness**: The null distribution of $\sum_{i=1}^n S_i$ is universal — does not depend on the underlying distribution of the data
- Leads to an exact and distribution-free test valid for all sample sizes
- **Issue**: Actually testing for $\mathrm{H}_0 : \mathbb{P}(X \geq 0) = \frac{1}{2}$; does not take into account the magnitude of the X_i's

Wilcoxon signed-rank test [Wilcoxon (1945)]

- Let R_i^+ be the absolute rank of X_i, i.e., the rank of $|X_i|$ in the sample of absolute values $|X_1|, \ldots, |X_n|$

- Rejects H_0 when $\sum_{i=1}^n S_i R_i^+$ is significantly different from 0

Wilcoxon signed-rank test [Wilcoxon (1945)]

- Let R_i^+ be the absolute rank of X_i, i.e., the rank of $|X_i|$ in the sample of absolute values $|X_1|, \ldots, |X_n|$
- Rejects H_0 when $\sum_{i=1}^n S_i R_i^+$ is significantly different from 0
- Under H_0, the distribution of $\sum_{i=1}^n S_i R_i^+$ is completely known

Distribution-freeness

- $\left(R_1^+, \ldots, R_n^+\right)$ are uniform over all $n!$ permutations of $\left\{\frac{1}{n}, \ldots, \frac{n}{n}\right\}$
- $(S_1, \ldots, S_n)$ independent of $\left(R_1^+, \ldots, R_n^+\right)$ under $\mathrm{H}_0 : X \stackrel{d}{=} -X$

Wilcoxon signed-rank test [Wilcoxon (1945)]

- Let R_i^+ be the absolute rank of X_i, i.e., the rank of $|X_i|$ in the sample of absolute values $|X_1|, \ldots, |X_n|$
- Rejects H_0 when $\sum_{i=1}^n S_i R_i^+$ is significantly different from 0
- Under H_0, the distribution of $\sum_{i=1}^n S_i R_i^+$ is completely known

Distribution-freeness

- $\left(R_1^+, \ldots, R_n^+\right)$ are uniform over all $n!$ permutations of $\left\{\frac{1}{n}, \ldots, \frac{n}{n}\right\}$
- $(S_1, \ldots, S_n)$ independent of $\left(R_1^+, \ldots, R_n^+\right)$ under $\mathrm{H}_0 : X \stackrel{d}{=} -X$

- Leads to an exact and distribution-free test valid for all sample sizes

Wilcoxon signed-rank test [Wilcoxon (1945)]

- Let R_i^+ be the absolute rank of X_i, i.e., the rank of $|X_i|$ in the sample of absolute values $|X_1|, \ldots, |X_n|$
- Rejects H_0 when $\sum_{i=1}^n S_i R_i^+$ is significantly different from 0
- Under H_0, the distribution of $\sum_{i=1}^n S_i R_i^+$ is completely known

Distribution-freeness

- $\left(R_1^+, \ldots, R_n^+\right)$ are uniform over all $n!$ permutations of $\left\{\frac{1}{n}, \ldots, \frac{n}{n}\right\}$
- $(S_1, \ldots, S_n)$ independent of $\left(R_1^+, \ldots, R_n^+\right)$ under $\mathrm{H}_0 : X \overset{d}{=} -X$

- Leads to an exact and distribution-free test valid for all sample sizes
- Consistent against location shift alternatives: $X_1, \ldots, X_n$ iid $f(\cdot - \theta)$; here f (unknown) is symmetric ($\mathrm{H}_0 : X \overset{d}{=} -X \quad \Leftrightarrow \quad \mathrm{H}_0 : \theta = 0$)

Wilcoxon signed-rank test [Wilcoxon (1945)]

- Let R_i^+ be the absolute rank of X_i, i.e., the rank of $|X_i|$ in the sample of absolute values $|X_1|, \ldots, |X_n|$
- Rejects H_0 when $\sum_{i=1}^n S_i R_i^+$ is significantly different from 0
- Under H_0, the distribution of $\sum_{i=1}^n S_i R_i^+$ is completely known

Distribution-freeness

- $\left(R_1^+, \ldots, R_n^+\right)$ are uniform over all $n!$ permutations of $\left\{\frac{1}{n}, \ldots, \frac{n}{n}\right\}$
- $(S_1, \ldots, S_n)$ independent of $\left(R_1^+, \ldots, R_n^+\right)$ under $\mathrm{H}_0 : X \stackrel{d}{=} -X$

- Leads to an exact and distribution-free test valid for all sample sizes
- Consistent against location shift alternatives: $X_1, \ldots, X_n$ iid $f(\cdot - \theta)$; here f (unknown) is symmetric $(\mathrm{H}_0 : X \stackrel{d}{=} -X \quad \Leftrightarrow \quad \mathrm{H}_0 : \theta = 0)$
- Powerful for heavy-tailed data, robust to outliers & contamination

Properties of sign and WSR tests when $p = 1$ [van der Vaart (1998)]:

1. **Distribution-freeness**:
   - S_i's are iid uniform over $\{-1, 1\}$, under $\mathrm{H}_0 : X \stackrel{d}{=} -X$
   - $\left(R_1^+, \ldots, R_n^+\right)$ are uniform over all $n!$ permutations of $\left\{\frac{1}{n}, \ldots, \frac{n}{n}\right\}$

2. **Independence**: $(S_1, \ldots, S_n)$ independent of $\left(R_1^+, \ldots, R_n^+\right)$ under H_0

Properties of sign and WSR tests when $p = 1$ [van der Vaart (1998)]:

1. **Distribution-freeness**:
   - S_i's are iid uniform over $\{-1, 1\}$, under $\mathrm{H}_0 : X \stackrel{d}{=} -X$
   - $(R_1^+, \ldots, R_n^+)$ are uniform over all $n!$ permutations of $\{\frac{1}{n}, \ldots, \frac{n}{n}\}$

2. **Independence**: $(S_1, \ldots, S_n)$ independent of $(R_1^+, \ldots, R_n^+)$ under H_0

3. **Asymptotic normality**: Both $\sum_{i=1}^n S_i$ and $\sum_{i=1}^n S_i R_i^+$ are asymptotically normal under H_0

Properties of sign and WSR tests when $p = 1$ [van der Vaart (1998)]:

1. **Distribution-freeness**:
   - S_i's are iid uniform over $\{-1, 1\}$, under $\mathrm{H}_0 : X \stackrel{d}{=} -X$
   - $\left(R_1^+, \ldots, R_n^+\right)$ are uniform over all $n!$ permutations of $\left\{\frac{1}{n}, \ldots, \frac{n}{n}\right\}$

2. **Independence**: $(S_1, \ldots, S_n)$ independent of $\left(R_1^+, \ldots, R_n^+\right)$ under H_0

3. **Asymptotic normality**: Both $\sum_{i=1}^n S_i$ and $\sum_{i=1}^n S_i R_i^+$ are asymptotically normal under H_0

4. **Asymptotic relative efficiency** (ARE) for location shift alternatives
   - Hodges-Lehmann (1956): ARE of WSR test w.r.t. t-test ≥ 0.864
   - Chernoff-Savage (1958): ARE of a Gaussian score transformed WSR test against the t-test is lower bounded by 1

Properties of sign and WSR tests when $p = 1$ [van der Vaart (1998)]:

1. **Distribution-freeness**:
   - S_i's are iid uniform over $\{-1, 1\}$, under $\mathrm{H}_0 : X \overset{d}{=} -X$
   - $\left(R_1^+, \ldots, R_n^+\right)$ are uniform over all $n!$ permutations of $\left\{\frac{1}{n}, \ldots, \frac{n}{n}\right\}$

2. **Independence**: $(S_1, \ldots, S_n)$ independent of $\left(R_1^+, \ldots, R_n^+\right)$ under H_0

3. **Asymptotic normality**: Both $\sum_{i=1}^n S_i$ and $\sum_{i=1}^n S_i R_i^+$ are asymptotically normal under H_0

4. **Asymptotic relative efficiency** (ARE) for location shift alternatives
   - Hodges-Lehmann (1956): ARE of WSR test w.r.t. t-test ≥ 0.864
   - Chernoff-Savage (1958): ARE of a Gaussian score transformed WSR test against the t-test is lower bounded by 1

5. Obtain distribution-free confidence sets for the "center" of X

Properties of sign and WSR tests when $p = 1$ [van der Vaart (1998)]:

1. **Distribution-freeness**:
   - S_i's are iid uniform over $\{-1, 1\}$, under $\mathrm{H}_0 : X \stackrel{d}{=} -X$
   - $\left(R_1^+, \ldots, R_n^+\right)$ are uniform over all $n!$ permutations of $\left\{\frac{1}{n}, \ldots, \frac{n}{n}\right\}$

2. **Independence**: $(S_1, \ldots, S_n)$ independent of $\left(R_1^+, \ldots, R_n^+\right)$ under H_0

3. **Asymptotic normality**: Both $\sum_{i=1}^n S_i$ and $\sum_{i=1}^n S_i R_i^+$ are asymptotically normal under H_0

4. **Asymptotic relative efficiency** (ARE) for location shift alternatives
   - Hodges-Lehmann (1956): ARE of WSR test w.r.t. t-test ≥ 0.864
   - Chernoff-Savage (1958): ARE of a Gaussian score transformed WSR test against the t-test is lower bounded by 1

5. Obtain distribution-free confidence sets for the "center" of X

Question: Can we derive tests with analogous properties when $p > 1$?

The distribution-free nature of signs and absolute ranks (under H_0) were crucial to developing distribution-free inference for symmetry when $p = 1$

Question: Can we define distribution-free (generalized) signs and ranks and develop distribution-free multivariate tests for $\mathcal{G}$-symmetry?

The distribution-free nature of signs and absolute ranks (under H_0) were crucial to developing distribution-free inference for symmetry when $p = 1$

Question: Can we define distribution-free (generalized) signs and ranks and develop distribution-free multivariate tests for $\mathcal{G}$-symmetry?

(Multivariate) ranks defined via optimal transport (OT) [Hallin (2017)] lead to distribution-free testing

Chernozhukov et al. (2017), De Valk & Segers (2018), Hallin, del Barrio, Cuesta-Albertos, Matrán (2018), Shi, Drton & Han (2019), Deb & S. (2019), Ghosal & S. (2019), Hallin, La Vecchia & Liu (2019), Hallin, Hlubinka, & Hudecová (2020), Deb, Ghosal & S. (2020), Shi, Hallin, Drton & Han (2020), Deb, Bhattacharya & S. (2021) ...

Outline

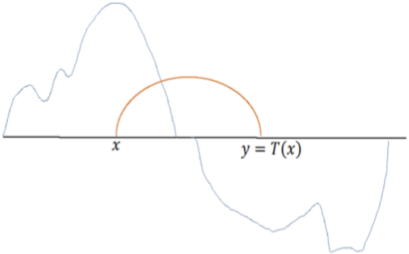
1. Generalized Signs and Ranks
   - Connection to Optimal Transport
   - Generalized Signs, Ranks and Signed-ranks
   - Population Analogues
2. Multivariate Distribution-free tests for Symmetry
   - Generalized Sign test and Wilcoxon Signed-rank test
   - Lower bounds on Asymptotic (Pitman) Relative Efficiency

Outline

1. Generalized Signs and Ranks
   - Connection to Optimal Transport
   - Generalized Signs, Ranks and Signed-ranks
   - Population Analogues
2. Multivariate Distribution-free tests for Symmetry
   - Generalized Sign test and Wilcoxon Signed-rank test
   - Lower bounds on Asymptotic (Pitman) Relative Efficiency

Optimal Transport: Monge's problem

Gaspard Monge (1781): What is the cheapest way to transport a pile of sand to cover a sinkhole?



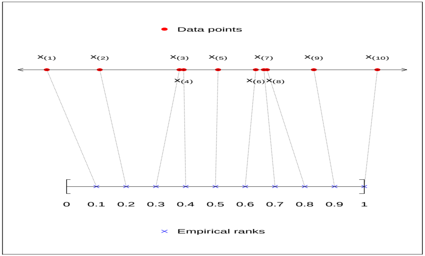
Goal: $\inf_{\mathbf{T}:\mathbf{T}(\mathbf{X})\sim\nu} \mathbb{E}_P[c(\mathbf{X}, \mathbf{T}(\mathbf{X}))]$ $\quad \mathbf{X} \sim P$

- P ("data" dist.) and ν ("reference" dist.)
- $c(\mathbf{x}, \mathbf{y}) \geq 0$: cost of transporting $\mathbf{x}$ to $\mathbf{y}$ (e.g., $c(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|^2$)
- $\mathbf{T}$ transports P to ν: $\mathbf{T}_{\#}P = \nu$ (i.e., $\mathbf{T}(\mathbf{X}) \sim \nu$ where $\mathbf{X} \sim P$)

Sample Ranks as Optimal Transport (OT) maps

- **Data**: $X_1, \ldots, X_n$ iid P (cont. dist.) on $\mathbb{R}$

- Let $P_n := \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i}$ and $\nu_n := \frac{1}{n}\sum_{i=1}^{n} \delta_{\frac{i}{n}}$



- **Sample rank map**: $\hat{R} : \{X_1, X_2, \ldots, X_n\} \longrightarrow \{\frac{1}{n}, \frac{2}{n}, \ldots, \frac{n}{n}\}$ solves

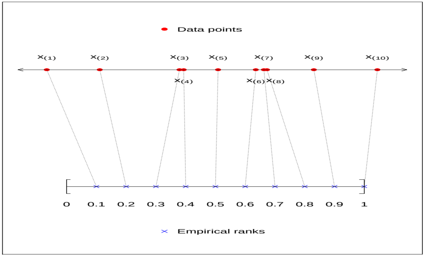
i.e., $$\hat{R} := \operatorname*{arg\,min}_{T : T_{\#}P_n = \nu_n} \frac{1}{n}\sum_{i=1}^{n} |X_i - T(X_i)|^2$$

Sample Ranks as Optimal Transport (OT) maps

- **Data**: $X_1, \ldots, X_n$ iid P (cont. dist.) on $\mathbb{R}$

- Let $P_n := \frac{1}{n} \sum_{i=1}^{n} \delta_{X_i}$ and

  $\nu_n := \frac{1}{n} \sum_{i=1}^{n} \delta_{\frac{i}{n}}$



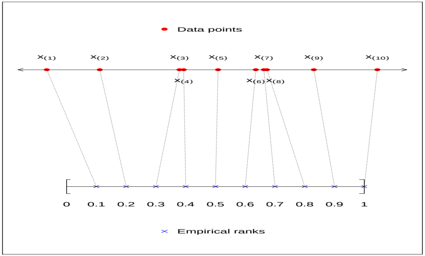
- **Sample rank map**: $\hat{R} : \{X_1, X_2, \ldots, X_n\} \longrightarrow \{\frac{1}{n}, \frac{2}{n}, \ldots, \frac{n}{n}\}$ solves

  i.e., $$\hat{R} := \operatorname*{arg\,min}_{T : T_{\#}P_n = \nu_n} \frac{1}{n} \sum_{i=1}^{n} |X_i - T(X_i)|^2 = \operatorname*{arg\,max}_{T : T_{\#}P_n = \nu_n} \frac{1}{n} \sum_{i=1}^{n} X_{(i)} T(X_{(i)})$$

Sample Ranks as Optimal Transport (OT) maps

- **Data**: $X_1, \ldots, X_n$ iid P (cont. dist.) on $\mathbb{R}$

- Let $P_n := \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i}$ and $\nu_n := \frac{1}{n}\sum_{i=1}^{n} \delta_{\frac{i}{n}}$



- **Sample rank map**: $\hat{R} : \{X_1, X_2, \ldots, X_n\} \longrightarrow \{\frac{1}{n}, \frac{2}{n}, \ldots, \frac{n}{n}\}$ solves

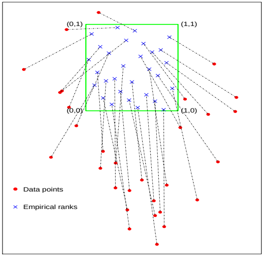
$$\text{i.e., } \hat{R} := \arg\min_{T : T_{\#}P_n = \nu_n} \frac{1}{n}\sum_{i=1}^{n} |X_i - T(X_i)|^2 = \arg\max_{T : T_{\#}P_n = \nu_n} \frac{1}{n}\sum_{i=1}^{n} X_{(i)} T(X_{(i)})$$

- $\hat{\sigma} := \arg\min_{\sigma \in \mathcal{S}_n} \frac{1}{n}\sum_{i=1}^{n} \left|X_{\sigma(i)} - \frac{i}{n}\right|^2$ where $\mathcal{S}_n$ is the set of all permutations of $\{1, \ldots, n\}$

- **Sample rank map**: $\hat{R}(X_i) = \frac{\hat{\sigma}^{-1}(i)}{n}$

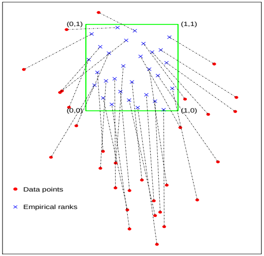
Multivariate Ranks as OT maps in $\mathbb{R}^p$ $(p \geq 1)$

- **Data**: $\mathbf{X}_1, \ldots, \mathbf{X}_n$ iid P (abs. cont.); $\nu \sim$ Unif$([0,1]^p)$ or $N(0, I_p)$
- Empirical rank map $\hat{R}$: $\{\mathbf{X}_1, \ldots, \mathbf{X}_n\} \to \{\mathbf{h}_1, \ldots, \mathbf{h}_n\} \subset [0,1]^d$ — sequence of "uniform-like" points (or quasi-Monte Carlo sequence)



Multivariate Ranks as OT maps in $\mathbb{R}^p$ ($p \geq 1$)

- **Data**: $\mathbf{X}_1, \ldots, \mathbf{X}_n$ iid P (abs. cont.); $\nu \sim$ Unif$([0,1]^p)$ or $N(0, I_p)$
- Empirical rank map $\hat{R}$: $\{\mathbf{X}_1, \ldots, \mathbf{X}_n\} \to \{\mathbf{h}_1, \ldots, \mathbf{h}_n\} \subset [0,1]^d$ — sequence of "uniform-like" points (or quasi-Monte Carlo sequence)



- Sample multivariate rank map [Hallin (2017), Deb & S. (2019)] is defined as the OT map s.t.

$$\hat{\sigma} := \arg\min_{\sigma \in \mathcal{S}_n} \frac{1}{n} \sum_{i=1}^{n} \|\mathbf{X}_{\sigma(i)} - \mathbf{h}_i\|^2; \qquad \hat{R}(\mathbf{X}_i) := \mathbf{h}_{\hat{\sigma}^{-1}(i)}$$

- Assignment problem (can be reduced to a linear program — $O(n^3)$)

Signs and absolute ranks via OT when $p = 1$

- **Data**: $X_1, \ldots, X_n$ iid P (cont. dist.) on $\mathbb{R}$
- $\mathrm{H}_0 : X \overset{d}{=} QX \quad \forall\, Q \in \mathcal{G} = \{+1, -1\}$

- Sign test: $\sum_{i=1}^n S_i$ [recall: $S_i := \mathrm{sign}(X_i)$]
- WSR test: $\sum_{i=1}^n S_i R_i^+$

Question: Can the signs and absolute ranks be obtained via OT?

Signs and absolute ranks via OT when $p = 1$

- **Data**: $X_1, \ldots, X_n$ iid P (cont. dist.) on $\mathbb{R}$
- $\mathrm{H}_0 : X \stackrel{d}{=} QX \quad \forall\, Q \in \mathcal{G} = \{+1, -1\}$

- Sign test: $\sum_{i=1}^n S_i$ [recall: $S_i := \mathrm{sign}(X_i)$]
- WSR test: $\sum_{i=1}^n S_i R_i^+$

Question: Can the signs and absolute ranks be obtained via OT?

- Consider the optimization problem:

$$(\hat{Q}, \hat{\sigma}) := \arg\min \left\{ \sum_{i=1}^n \left| q_i X_{\sigma(i)} - \frac{i}{n} \right|^2 : \; Q = (q_i)_{i=1}^n \in \{\pm 1\}^n, \sigma \in \mathcal{S}_n \right\}$$

- The signs and absolute ranks are then given by:

$$S_i = \hat{Q}_{\hat{\sigma}^{-1}(i)}, \qquad R_i^+ = \frac{\hat{\sigma}^{-1}(i)}{n}$$

Signs and absolute ranks via OT when $p = 1$

- **Data**: $X_1, \ldots, X_n$ iid P (cont. dist.) on $\mathbb{R}$
- $\mathrm{H}_0 : X \overset{d}{=} QX \quad \forall\, Q \in \mathcal{G} = \{+1, -1\}$

- Sign test: $\sum_{i=1}^{n} S_i$ [recall: $S_i := \mathrm{sign}(X_i)$]
- WSR test: $\sum_{i=1}^{n} S_i R_i^+$

Question: Can the signs and absolute ranks be obtained via OT?

- Consider the optimization problem:

$$(\hat{Q}, \hat{\sigma}) := \arg\min \left\{ \sum_{i=1}^{n} \left| q_i X_{\sigma(i)} - \frac{i}{n} \right|^2 : \; Q = (q_i)_{i=1}^{n} \in \{\pm 1\}^n, \sigma \in \mathcal{S}_n \right\}$$

- The signs and absolute ranks are then given by:

$$S_i = \hat{Q}_{\hat{\sigma}^{-1}(i)}, \qquad R_i^+ = \frac{\hat{\sigma}^{-1}(i)}{n}$$

- The signed-rank for X_i is then defined as $S_i R_i^+$

- **Data**: $\mathbf{X}_1, \ldots, \mathbf{X}_n$ iid P (abs. cont.) on $\mathbb{R}^p$ ($p \geq 1$); $\mathcal{G} \subset \mathrm{O}(p)$
- Consider the following optimization problem:

$$(\hat{Q}, \hat{\sigma}) := \arg\min \left\{ \sum_{i=1}^{n} \| Q_i^\top \mathbf{X}_{\sigma(i)} - \mathbf{h}_i \|^2 : Q_i \in \mathcal{G}, \sigma \in \mathcal{S}_n \right\} \qquad (\star)$$

where $\{\mathbf{h}_1, \ldots, \mathbf{h}_n\}$ is discretization of the reference dist. ν

- **Data**: $\mathbf{X}_1, \ldots, \mathbf{X}_n$ iid P (abs. cont.) on $\mathbb{R}^p$ ($p \geq 1$); $\mathcal{G} \subset \mathrm{O}(p)$
- Consider the following optimization problem:

$$(\hat{Q}, \hat{\sigma}) := \arg\min \left\{ \sum_{i=1}^{n} \| Q_i^\top \mathbf{X}_{\sigma(i)} - \mathbf{h}_i \|^2 : Q_i \in \mathcal{G}, \sigma \in \mathcal{S}_n \right\} \qquad (\star)$$

where $\{\mathbf{h}_1, \ldots, \mathbf{h}_n\}$ is discretization of the reference dist. ν

Question: Can the above be seen as an OT problem?

- **Data**: $\mathbf{X}_1, \ldots, \mathbf{X}_n$ iid P (abs. cont.) on $\mathbb{R}^p$ ($p \geq 1$); $\mathcal{G} \subset \mathrm{O}(p)$
- Consider the following optimization problem:

$$(\hat{Q}, \hat{\sigma}) := \arg\min \left\{ \sum_{i=1}^{n} \|Q_i^\top \mathbf{X}_{\sigma(i)} - \mathbf{h}_i\|^2 : Q_i \in \mathcal{G}, \sigma \in \mathcal{S}_n \right\} \qquad (\star)$$

where $\{\mathbf{h}_1, \ldots, \mathbf{h}_n\}$ is discretization of the reference dist. ν

Question: Can the above be seen as an OT problem?

Define the cost function:

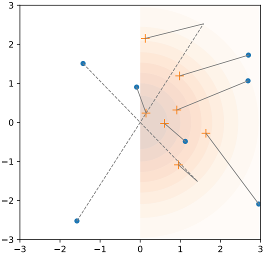
$$c(\mathbf{x}, \mathbf{h}) := \min_{Q \in \mathcal{G}} \|Q^\top \mathbf{x} - \mathbf{h}\|^2, \qquad \text{for } \mathbf{x}, \mathbf{h} \in \mathbb{R}^p.$$

Monge's problem (OT): $(\star) = \inf_{\mathbf{T}:\mathbf{T}_\# P_n = \nu_n} \frac{1}{n} \sum_{i=1}^{n} c(\mathbf{X}_i, \mathbf{T}(\mathbf{X}_i))$

where $\mathbf{T}$ transports $P_n := \frac{1}{n} \sum_{i=1}^{n} \delta_{\mathbf{X}_i}$ to $\nu_n := \frac{1}{n} \sum_{i=1}^{n} \delta_{\mathbf{h}_i}$

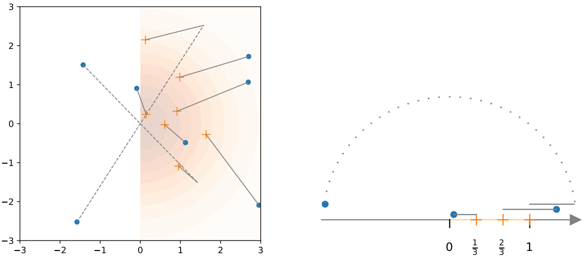
$$(\hat{Q}, \hat{\sigma}) := \arg\min \Big\{ \sum_{i=1}^{n} \| Q_i^\top \mathbf{X}_{\sigma(i)} - \mathbf{h}_i \|^2 : Q_i \in \mathcal{G}, \sigma \in \mathcal{S}_n \Big\} \qquad (\star)$$

Figure: Data points ("•") and their ranks ("+"). Here $\mathcal{G} = \{-I_p, I_p\}$.



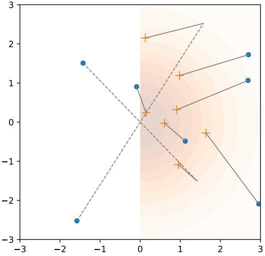
$$(\hat{Q}, \hat{\sigma}) := \arg\min \left\{ \sum_{i=1}^{n} \| Q_i^\top \mathbf{X}_{\sigma(i)} - \mathbf{h}_i \|^2 : Q_i \in \mathcal{G}, \sigma \in \mathcal{S}_n \right\} \qquad (\star)$$

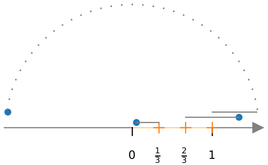
Figure: Data points ("•") and their ranks ("+"). Here $\mathcal{G} = \{-I_p, I_p\}$.



$$(\hat{Q}, \hat{\sigma}) := \arg\min \left\{ \sum_{i=1}^{n} \| Q_i^\top \mathbf{X}_{\sigma(i)} - \mathbf{h}_i \|^2 : Q_i \in \mathcal{G}, \sigma \in \mathcal{S}_n \right\} \qquad (\star)$$

Figure: Data points ("•") and their ranks ("+"). Here $\mathcal{G} = \{-I_p, I_p\}$.





$$(\hat{Q}, \hat{\sigma}) := \arg\min \left\{ \sum_{i=1}^{n} \| \mathbf{X}_{\sigma(i)} - Q_i \mathbf{h}_i \|^2 : Q_i \in \mathcal{G}, \sigma \in \mathcal{S}_n \right\}$$

- Define the generalized sign and generalized rank as:

$$S_n(\mathbf{X}_i) := \hat{Q}_{\hat{\sigma}^{-1}(i)}, \qquad R_n(\mathbf{X}_i) := \mathbf{h}_{\hat{\sigma}^{-1}(i)}$$

- The generalized signed-rank of $\mathbf{X}_i$ is $S_n(\mathbf{X}_i) R_n(\mathbf{X}_i)$ — it is the closest point to $\mathbf{X}_i$ in the orbit of $R_n(\mathbf{X}_i)$ (i.e., $\{Q R_n(\mathbf{X}_i) : Q \in \mathcal{G}\}$)

Uniqueness of generalized ranks & signed-ranks [Huang & S. (2023+)]

- The generalized rank — $R_n(\mathbf{X}_i)$ — is a.s. unique,[a] $\forall i \in [n]$

Uniqueness of generalized ranks & signed-ranks [Huang & S. (2023+)]

- The generalized rank — $R_n(\mathbf{X}_i)$ — is a.s. unique,[a] $\forall i \in [n]$

- The signed-rank — $S_n(\mathbf{X}_i)R_n(\mathbf{X}_i)$ — is a.s. unique, $\forall i \in [n]$
- **Recall**: the signed-rank is the point in the orbit of $R_n(\mathbf{X}_i)$ (i.e., $\{QR_n(\mathbf{X}_i) : Q \in \mathcal{G}\}$) that is closest to $\mathbf{X}_i$

[a] We assume that no two $\mathbf{h}_j$'s lie on a same orbit of $\mathcal{G}$.

Uniqueness of generalized ranks & signed-ranks [Huang & S. (2023+)]

- The generalized rank — $R_n(\mathbf{X}_i)$ — is a.s. unique,[a] $\forall i \in [n]$
- The signed-rank — $S_n(\mathbf{X}_i)R_n(\mathbf{X}_i)$ — is a.s. unique, $\forall i \in [n]$
- **Recall**: the signed-rank is the point in the orbit of $R_n(\mathbf{X}_i)$ (i.e., $\{QR_n(\mathbf{X}_i) : Q \in \mathcal{G}\}$) that is closest to $\mathbf{X}_i$

[a] We assume that no two $\mathbf{h}_j$'s lie on a same orbit of $\mathcal{G}$.

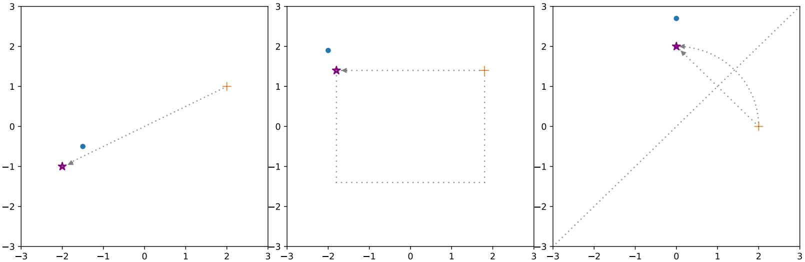


Figure: Data point ("•"), its rank ("+") and its signed-rank ("⋆").
Left: $\mathcal{G} = \{-I_p, I_p\}$ (central sym.). **Center**: $\mathcal{G}$ corresponds to sign symmetry.
Right: $\mathcal{G} = \mathrm{O}(2)$; the signed-rank ("⋆") is the (unique) point in $\{\mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x}\| = 2\}$ that is closest to the data point.

The sign $S_n(\mathbf{X}_i) = \arg\min_{Q \in \mathcal{G}} \|\mathbf{X}_i - QR_n(\mathbf{X}_i)\|^2$ may be not unique

Result If $\mathcal{G}$ is the group corresponding to central/sign symmetry, then the (generalized) sign $S_n(\mathbf{X}_i)$ is unique a.s.

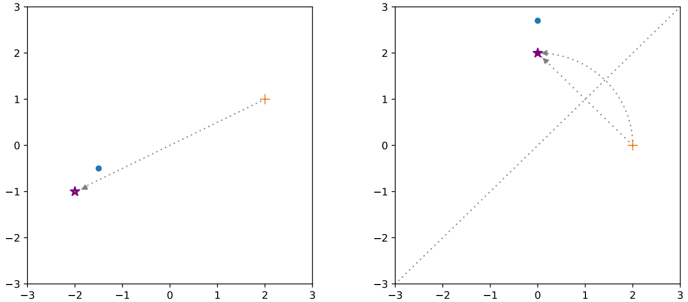


Figure: Data point ("•"), its rank ("+") and its signed-rank ("⋆"). **Left**: Here $\mathcal{G} = \{-I_p, I_p\}$ and sign is unique! **Right**: Here $\mathcal{G} = \mathrm{O}(2)$ and sign is not unique!

The sign $S_n(\mathbf{X}_i) = \arg\min_{Q \in \mathcal{G}} \|\mathbf{X}_i - QR_n(\mathbf{X}_i)\|^2$ may be not unique

Result If $\mathcal{G}$ is the group corresponding to central/sign symmetry, then the (generalized) sign $S_n(\mathbf{X}_i)$ is unique a.s.

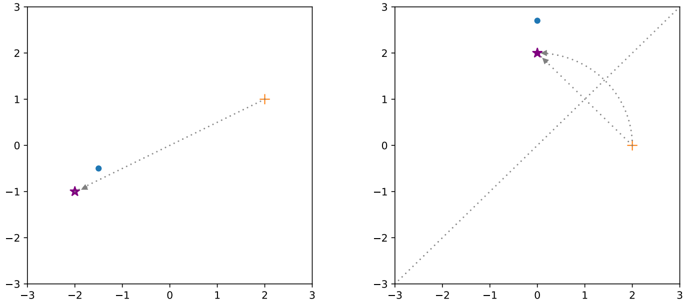


Figure: Data point ("•"), its rank ("+") and its signed-rank ("⋆"). **Left**: Here $\mathcal{G} = \{-I_p, I_p\}$ and sign is unique! **Right**: Here $\mathcal{G} = \mathrm{O}(2)$ and sign is not unique!

Uniform Can choose $S_n(\mathbf{X}_i)$ 'uniformly' over all possible minimizing values

$$S_n(\mathbf{X}_i) = \arg\min_{Q \in \mathcal{G}} \| Q^\top \mathbf{X}_i - R_n(\mathbf{X}_i) \|^2 = \arg\min_{Q \in \mathcal{G}} \| \mathbf{X}_i - Q R_n(\mathbf{X}_i) \|^2$$

Question: When can we identify the (generalized) sign?

$$S_n(\mathbf{X}_i) = \arg\min_{Q \in \mathcal{G}} \|Q^\top \mathbf{X}_i - R_n(\mathbf{X}_i)\|^2 = \arg\min_{Q \in \mathcal{G}} \|\mathbf{X}_i - QR_n(\mathbf{X}_i)\|^2$$

Question: When can we identify the (generalized) sign?

- $\mathcal{G}$ acts freely if for $\mathbf{x} \in \mathbb{R}^p$ and $Q_1, Q_2 \in \mathcal{G}$,

$$Q_1\mathbf{x} = Q_2\mathbf{x} \quad \Rightarrow \quad Q_1 = Q_2$$

  (i.e., for any $\mathbf{x}$ in $\mathbb{R}^p$, we can identify the unique element in $\mathcal{G}$ that maps $\mathbf{x} \mapsto Q\mathbf{x}$)

- Free group action is available for central / sign symmetry

- For infinite groups $\mathcal{G}$ we may not have a free group action

Proposition [Huang & S. (2023+)]

Suppose that $\mathcal{G}$ acts freely and suppose no two $\mathbf{h}_j$'s lie on a same orbit of $\mathcal{G}$. Then $S_n(\cdot)$ is a.s. unique.

Computational complexity

- Cost function: $c_{i,j} \equiv c(\mathbf{X}_i, \mathbf{h}_j) := \min_{Q \in \mathcal{G}} \|Q^\top \mathbf{X}_i - \mathbf{h}_j\|^2, \quad \forall i, j \in [n]$
- OT problem: $\min \left\{ \sum_{i=1}^n c_{i,\sigma(i)} : \sigma \in \mathcal{S}_n \right\}$ — assignment problem

- If $\mathcal{G}$ is a finite group then $c_{i,j}$ can be computed in $O(1)$ time
- $\{R_n(\mathbf{X}_i)\}_{i=1}^n$ can be found by solving the assignment problem of $\{\mathbf{X}_1, \dots, \mathbf{X}_n\}$ to $\{\mathbf{h}_1, \dots, \mathbf{h}_n\}$ under cost $c(\cdot, \cdot)$ — complexity $O(n^3)$
- Sign: $S_n(\mathbf{X}_i) \equiv S_n(\mathbf{X}_i, R_n(\mathbf{X}_i)) := \arg\min_{Q \in \mathcal{G}} \|Q^\top \mathbf{X}_i - R_n(\mathbf{X}_i)\|^2$

Computational complexity

- Cost function: $c_{i,j} \equiv c(\mathbf{X}_i, \mathbf{h}_j) := \min_{Q \in \mathcal{G}} \|Q^\top \mathbf{X}_i - \mathbf{h}_j\|^2, \quad \forall i, j \in [n]$
- OT problem: $\min \left\{ \sum_{i=1}^n c_{i,\sigma(i)} : \sigma \in \mathcal{S}_n \right\}$ — assignment problem

- If $\mathcal{G}$ is a finite group then $c_{i,j}$ can be computed in $O(1)$ time
- $\{R_n(\mathbf{X}_i)\}_{i=1}^n$ can be found by solving the assignment problem of $\{\mathbf{X}_1, \ldots, \mathbf{X}_n\}$ to $\{\mathbf{h}_1, \ldots, \mathbf{h}_n\}$ under cost $c(\cdot, \cdot)$ — complexity $O(n^3)$
- Sign: $S_n(\mathbf{X}_i) \equiv S_n(\mathbf{X}_i, R_n(\mathbf{X}_i)) := \arg\min_{Q \in \mathcal{G}} \|Q^\top \mathbf{X}_i - R_n(\mathbf{X}_i)\|^2$

For some group $\mathcal{G}$, the computation can be much faster!

Spherical symmetry ($\mathcal{G} = \mathrm{O}(p)$)

- The computation time of the ranks (and signed-ranks): $O(n \log n)$
- $c(\mathbf{x}, \mathbf{h}) = \|\mathbf{x}\|^2 - 2 \max_{Q \in \mathcal{G}} \mathbf{x}^\top Q \mathbf{h} + \|\mathbf{h}\|^2 = \big(\|\mathbf{x}\| - \|\mathbf{h}\|\big)^2$

Spherical symmetry ($\mathcal{G} = \mathrm{O}(p)$)

- The computation time of the ranks (and signed-ranks): $O(n \log n)$
- $c(\mathbf{x}, \mathbf{h}) = \|\mathbf{x}\|^2 - 2 \max_{Q \in \mathcal{G}} \mathbf{x}^\top Q \mathbf{h} + \|\mathbf{h}\|^2 = \big(\|\mathbf{x}\| - \|\mathbf{h}\|\big)^2$
- If $\mathbf{X}_i$ has the j-th largest Euclidean norm among $\mathbf{X}_1, \ldots, \mathbf{X}_n$ and $\|\mathbf{h}_1\| < \ldots < \|\mathbf{h}_n\|$, then $\mathbf{X}_i$ will have $\mathbf{h}_j$ as its rank

Spherical symmetry ($\mathcal{G} = \mathrm{O}(p)$)

- The computation time of the ranks (and signed-ranks): $O(n \log n)$
- $c(\mathbf{x}, \mathbf{h}) = \|\mathbf{x}\|^2 - 2 \max_{Q \in \mathcal{G}} \mathbf{x}^\top Q \mathbf{h} + \|\mathbf{h}\|^2 = \big(\|\mathbf{x}\| - \|\mathbf{h}\|\big)^2$
- If $\mathbf{X}_i$ has the j-th largest Euclidean norm among $\mathbf{X}_1, \ldots, \mathbf{X}_n$ and $\|\mathbf{h}_1\| < \ldots < \|\mathbf{h}_n\|$, then $\mathbf{X}_i$ will have $\mathbf{h}_j$ as its rank
- The signed-rank of $\mathbf{X}_i$ is simply the vector in the direction of $\mathbf{X}_i$ with length $\|R_n(\mathbf{X}_i)\|$, i.e.,

$$S_n(\mathbf{X}_i) R_n(\mathbf{X}_i) = \|R_n(\mathbf{X}_i)\| \frac{\mathbf{X}_i}{\|\mathbf{X}_i\|}$$

- Given $\mathbf{X}_1, \ldots, \mathbf{X}_n$ iid P on $\mathbb{R}^p$ $(p \geq 1)$; $\quad \mathcal{G} \subset \mathrm{O}(p)$
- $\{\mathbf{h}_1, \ldots, \mathbf{h}_n\}$ is discretization of the reference dist. ν

- **OT**: $(\hat{Q}, \hat{\sigma}) := \arg\min \left\{ \sum_{i=1}^{n} \| Q_i^\top \mathbf{X}_{\sigma(i)} - \mathbf{h}_i \|^2 : Q_i \in \mathcal{G}, \sigma \in \mathcal{S}_n \right\}$
- Define the generalized sign and generalized rank as:

$$S_n(\mathbf{X}_i) := \hat{Q}_{\hat{\sigma}^{-1}(i)}, \qquad R_n(\mathbf{X}_i) := \mathbf{h}_{\hat{\sigma}^{-1}(i)}$$

- Given $\mathbf{X}_1, \ldots, \mathbf{X}_n$ iid P on $\mathbb{R}^p$ $(p \geq 1)$; $\mathcal{G} \subset \mathrm{O}(p)$
- $\{\mathbf{h}_1, \ldots, \mathbf{h}_n\}$ is discretization of the reference dist. ν

- **OT**: $(\hat{Q}, \hat{\sigma}) := \arg\min \left\{ \sum_{i=1}^{n} \| Q_i^\top \mathbf{X}_{\sigma(i)} - \mathbf{h}_i \|^2 : Q_i \in \mathcal{G}, \sigma \in \mathcal{S}_n \right\}$
- Define the generalized sign and generalized rank as:

$$S_n(\mathbf{X}_i) := \hat{Q}_{\hat{\sigma}^{-1}(i)}, \qquad R_n(\mathbf{X}_i) := \mathbf{h}_{\hat{\sigma}^{-1}(i)}$$

Theorem [Huang & S. (2023+)]

Result: $(R_n(\mathbf{X}_1), \ldots, R_n(\mathbf{X}_n))$ is uniformly distributed over the set of all $n!$ permutations of $\{\mathbf{h}_1, \ldots, \mathbf{h}_n\}$

Under $\mathrm{H}_0 : \mathbf{X} \stackrel{d}{=} Q\mathbf{X} \quad \forall\, Q \in \mathcal{G}$,

1. $S_n(\mathbf{X}_1), \ldots, S_n(\mathbf{X}_n)$ are iid Uniform$(\mathcal{G})$
2. $(R_n(\mathbf{X}_1), \ldots, R_n(\mathbf{X}_n))$ and $(S_n(\mathbf{X}_1), \ldots, S_n(\mathbf{X}_n))$ are independent

- Given $\mathbf{X}_1, \ldots, \mathbf{X}_n$ iid P on $\mathbb{R}^p$ $(p \geq 1)$; $\mathcal{G} \subset \mathrm{O}(p)$
- $\{\mathbf{h}_1, \ldots, \mathbf{h}_n\}$ is discretization of the reference dist. ν

- **OT**: $(\hat{Q}, \hat{\sigma}) := \arg\min \left\{ \sum_{i=1}^n \| Q_i^\top \mathbf{X}_{\sigma(i)} - \mathbf{h}_i \|^2 : Q_i \in \mathcal{G}, \sigma \in \mathcal{S}_n \right\}$
- Define the generalized sign and generalized rank as:

$$S_n(\mathbf{X}_i) := \hat{Q}_{\hat{\sigma}^{-1}(i)}, \qquad R_n(\mathbf{X}_i) := \mathbf{h}_{\hat{\sigma}^{-1}(i)}$$

Theorem [Huang & S. (2023+)]

Result: $(R_n(\mathbf{X}_1), \ldots, R_n(\mathbf{X}_n))$ is uniformly distributed over the set of all $n!$ permutations of $\{\mathbf{h}_1, \ldots, \mathbf{h}_n\}$

Under $\mathrm{H}_0 : \mathbf{X} \stackrel{d}{=} Q\mathbf{X} \quad \forall\, Q \in \mathcal{G}$,

1. $S_n(\mathbf{X}_1), \ldots, S_n(\mathbf{X}_n)$ are iid Uniform$(\mathcal{G})$
2. $(R_n(\mathbf{X}_1), \ldots, R_n(\mathbf{X}_n))$ and $(S_n(\mathbf{X}_1), \ldots, S_n(\mathbf{X}_n))$ are independent

Generalizes the distribution-freeness of signs and ranks beyond $p = 1$!

(Generalized) Wilcoxon signed-rank test: $W_n := \sum_{i=1}^n S_n(\mathbf{X}_i) R_n(\mathbf{X}_i)$

$\mathbf{X} \sim P$ (abs. cont.) on $\mathbb{R}^p$; ν: reference dist.; $\mathcal{G} \subset \mathrm{O}(p)$

Population OT problem [Kantorovich's relaxation]

$$\inf_{(\mathbf{X},\mathbf{H}):\mathbf{X}\sim P,\mathbf{H}\sim\nu} \mathbb{E}\left[c(\mathbf{X},\mathbf{H})\right], \qquad c(\mathbf{x},\mathbf{h}) := \min_{Q\in\mathcal{G}} \|Q^\top \mathbf{x} - \mathbf{h}\|^2$$

and $(\mathbf{X},\mathbf{H})$ runs over all joint dist. with marginals $\mathbf{X} \sim P$ and $\mathbf{H} \sim \nu$.

$\mathbf{X} \sim P$ (abs. cont.) on $\mathbb{R}^p$; ν: reference dist.; $\mathcal{G} \subset \mathrm{O}(p)$

Population OT problem [Kantorovich's relaxation]

$$\inf_{(\mathbf{X},\mathbf{H}):\mathbf{X}\sim P,\mathbf{H}\sim\nu} \mathbb{E}\left[c(\mathbf{X},\mathbf{H})\right], \qquad c(\mathbf{x},\mathbf{h}) := \min_{Q\in\mathcal{G}} \|Q^\top \mathbf{x} - \mathbf{h}\|^2$$

and $(\mathbf{X},\mathbf{H})$ runs over all joint dist. with marginals $\mathbf{X} \sim P$ and $\mathbf{H} \sim \nu$.

Assumption (A) (On ν and $\mathcal{G}$): $\exists\ B \subset \mathbb{R}^p$ with $\nu(B) = 1$ such that, for any $\mathbf{h} \in \mathbb{R}^p$, the orbit $\{Q\mathbf{h} : Q \in \mathcal{G}\}$ intersects B at one point at most.

- Central symmetry: $\mathbf{h}$ and $-\mathbf{h}$ cannot both be in B; we can take $B = (0,\infty) \times \mathbb{R}^{p-1}$;

$\mathbf{X} \sim P$ (abs. cont.) on $\mathbb{R}^p$; ν: reference dist.; $\mathcal{G} \subset \mathrm{O}(p)$

Population OT problem [Kantorovich's relaxation]

$$\inf_{(\mathbf{X},\mathbf{H}):\mathbf{X}\sim P,\mathbf{H}\sim\nu} \mathbb{E}\left[c(\mathbf{X},\mathbf{H})\right], \qquad c(\mathbf{x},\mathbf{h}) := \min_{Q\in\mathcal{G}} \|Q^\top \mathbf{x} - \mathbf{h}\|^2$$

and $(\mathbf{X},\mathbf{H})$ runs over all joint dist. with marginals $\mathbf{X} \sim P$ and $\mathbf{H} \sim \nu$.

Assumption (A) (On ν and $\mathcal{G}$): $\exists\ B \subset \mathbb{R}^p$ with $\nu(B) = 1$ such that, for any $\mathbf{h} \in \mathbb{R}^p$, the orbit $\{Q\mathbf{h} : Q \in \mathcal{G}\}$ intersects B at one point at most.

- Central symmetry: $\mathbf{h}$ and $-\mathbf{h}$ cannot both be in B; we can take $B = (0,\infty) \times \mathbb{R}^{p-1}$; when $p = 1$, $B = (0,1)$ and $\nu =$ Unif(0,1)

$\mathbf{X} \sim P$ (abs. cont.) on $\mathbb{R}^p$; ν: reference dist.; $\mathcal{G} \subset \mathrm{O}(p)$

Population OT problem [Kantorovich's relaxation]

$$\inf_{(\mathbf{X},\mathbf{H}):\mathbf{X}\sim P,\mathbf{H}\sim\nu} \mathbb{E}\left[c(\mathbf{X},\mathbf{H})\right], \qquad c(\mathbf{x},\mathbf{h}) := \min_{Q\in\mathcal{G}} \|Q^\top \mathbf{x} - \mathbf{h}\|^2$$

and $(\mathbf{X},\mathbf{H})$ runs over all joint dist. with marginals $\mathbf{X} \sim P$ and $\mathbf{H} \sim \nu$.

Assumption (A) (On ν and $\mathcal{G}$): $\exists\ B \subset \mathbb{R}^p$ with $\nu(B) = 1$ such that, for any $\mathbf{h} \in \mathbb{R}^p$, the orbit $\{Q\mathbf{h} : Q \in \mathcal{G}\}$ intersects B at one point at most.

- Central symmetry: $\mathbf{h}$ and $-\mathbf{h}$ cannot both be in B; we can take $B = (0,\infty) \times \mathbb{R}^{p-1}$; when $p = 1$, $B = (0,1)$ and $\nu = \mathrm{Unif(0,1)}$
- Sign symmetry: We can take $B = (0,\infty)^p$
- Spherical symmetry ($\mathcal{G} = \mathrm{O}(p)$): We can take $B = (0,\infty) \times \{0\}^{p-1}$;

$\mathbf{X} \sim P$ (abs. cont.) on $\mathbb{R}^p$; ν: reference dist.; $\mathcal{G} \subset \mathrm{O}(p)$

Population OT problem [Kantorovich's relaxation]

$$\inf_{(\mathbf{X},\mathbf{H}):\mathbf{X}\sim P,\mathbf{H}\sim\nu} \mathbb{E}\left[c(\mathbf{X},\mathbf{H})\right], \qquad c(\mathbf{x},\mathbf{h}) := \min_{Q\in\mathcal{G}} \|Q^\top \mathbf{x} - \mathbf{h}\|^2$$

and $(\mathbf{X},\mathbf{H})$ runs over all joint dist. with marginals $\mathbf{X} \sim P$ and $\mathbf{H} \sim \nu$.

Assumption (A) (On ν and $\mathcal{G}$): $\exists\ B \subset \mathbb{R}^p$ with $\nu(B) = 1$ such that, for any $\mathbf{h} \in \mathbb{R}^p$, the orbit $\{Q\mathbf{h} : Q \in \mathcal{G}\}$ intersects B at one point at most.

- Central symmetry: $\mathbf{h}$ and $-\mathbf{h}$ cannot both be in B; we can take $B = (0,\infty) \times \mathbb{R}^{p-1}$; when $p = 1$, $B = (0,1)$ and $\nu =$ Unif(0,1)
- Sign symmetry: We can take $B = (0,\infty)^p$
- Spherical symmetry ($\mathcal{G} = \mathrm{O}(p)$): We can take $B = (0,\infty) \times \{0\}^{p-1}$; thus ν is not abs. cont. here

Quotient map for cost $c(\mathbf{x}, \mathbf{h}) := \min_{Q \in \mathcal{G}} \|\mathbf{x} - Q\mathbf{h}\|^2$

- Orbit of $\mathbf{h}$ is $\{Q\mathbf{h} : Q \in \mathcal{G}\}$; every point in an orbit has the same cost

Quotient map for cost $c(\mathbf{x}, \mathbf{h}) := \min_{Q \in \mathcal{G}} \|\mathbf{x} - Q\mathbf{h}\|^2$

- Orbit of $\mathbf{h}$ is $\{Q\mathbf{h} : Q \in \mathcal{G}\}$; every point in an orbit has the same cost
- Image of group action of $\mathcal{G}$ on B: $\mathcal{G}B = \{Q\mathbf{h} : Q \in \mathcal{G}, \mathbf{h} \in B\} \subset \mathbb{R}^p$

For any point in $\mathcal{G}B$, quotient map picks the representative point in B:

$$q : \mathcal{G}B \to B \qquad \text{where} \quad q(Q\mathbf{h}) = \mathbf{h} \;\; \text{for} \;\; \mathbf{h} \in B, Q \in \mathcal{G}.$$

If Assumption (A) holds, then $q(\cdot)$ is well-defined

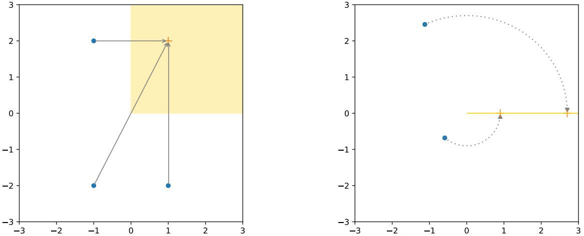


Figure: Shows the action of the quotient map q on: (i) (**Left**) 3 points when $\mathcal{G}$ corresponds to the group for sign symmetry, and (ii) (**Right**) on 2 points for $\mathcal{G}$ corresponding to the group for spherical symmetry (here $q(\mathbf{x}) = (\|\mathbf{x}\|, 0)$)

Population generalized rank map [Huang & S. (2023+)]

Let $\mathbf{X} \sim P$ (abs. cont.), $\mathbf{H} \sim \nu$ and suppose Assumption (A) holds.

Then, $\exists$ (P-a.e.) unique map $R : \mathbb{R}^p \to \mathbb{R}^p$ that solves the OT problem of transporting P to ν ($R_\# P = \nu$), i.e., Monge's problem = Kantorovich's relaxation:

$$\inf_{(\mathbf{X},\mathbf{H})\sim\pi\in\Pi(P,\nu)} \mathbb{E}_\pi \left[c(\mathbf{X},\mathbf{H})\right] = \mathbb{E}_P \left[c(\mathbf{X}, R(\mathbf{X}))\right], \quad c(\mathbf{x},\mathbf{h}) := \min_{Q\in\mathcal{G}} \|\mathbf{x} - Q\mathbf{h}\|^2$$

Population generalized rank map [Huang & S. (2023+)]

Let $\mathbf{X} \sim P$ (abs. cont.), $\mathbf{H} \sim \nu$ and suppose Assumption (A) holds.

Then, $\exists$ (P-a.e.) unique map $R : \mathbb{R}^p \to \mathbb{R}^p$ that solves the OT problem of transporting P to ν ($R_\# P = \nu$), i.e., Monge's problem = Kantorovich's relaxation:

$$\inf_{(\mathbf{X},\mathbf{H})\sim\pi\in\Pi(P,\nu)} \mathbb{E}_\pi\left[c(\mathbf{X},\mathbf{H})\right] = \mathbb{E}_P\left[c(\mathbf{X},R(\mathbf{X}))\right], \quad c(\mathbf{x},\mathbf{h}) := \min_{Q\in\mathcal{G}} \|\mathbf{x} - Q\mathbf{h}\|^2$$

Even if P and ν do not have second order moments, the following hold:

(i) $\exists$ a P-a.e. unique map $R : \mathbb{R}^p \to \mathbb{R}^p$ s.t. $(\mathbf{X}, R(\mathbf{X}))$ has the unique distribution in $\Pi(P,\nu)$ with a c-cyclically monotone support.

(ii) $\exists$ a l.s.c. convex function ψ such that $R(\mathbf{x}) = q(\nabla\psi(\mathbf{x}))$ (P-a.e.)

$\mathbf{X} \sim P$ (abs. cont.), $\mathbf{H} \sim \nu$ and suppose Assumption (A) holds.

Cost function: $c(\mathbf{x}, \mathbf{h}) := \min_{Q \in \mathcal{G}} \|\mathbf{x} - Q\mathbf{h}\|^2$

Population rank and signed-rank maps [Huang & S. (2023+)]

(i) $\exists$ a P-a.e. unique map $R : \mathbb{R}^p \to \mathbb{R}^p$ s.t. $(\mathbf{X}, R(\mathbf{X}))$ has the unique distribution in $\Pi(P, \nu)$ with a c-cyclically monotone support.

(ii) $\exists$ a l.s.c. convex function ψ such that $R(\mathbf{x}) = q(\nabla\psi(\mathbf{x}))$ (P-a.e.)

(iii) Here, $\nabla\psi(\cdot)$ is the P-a.e. unique gradient of a convex function s.t. $\nabla\psi(G\mathbf{X}) \sim G\mathbf{H}$, where $G \sim \text{Uniform}(\mathcal{G})$ is indep. of $\mathbf{X}$ & $\mathbf{H}$

$\mathbf{X} \sim P$ (abs. cont.), $\mathbf{H} \sim \nu$ and suppose Assumption (A) holds.

Cost function: $c(\mathbf{x}, \mathbf{h}) := \min_{Q \in \mathcal{G}} \|\mathbf{x} - Q\mathbf{h}\|^2$

Population rank and signed-rank maps [Huang & S. (2023+)]

(i) $\exists$ a P-a.e. unique map $R : \mathbb{R}^p \to \mathbb{R}^p$ s.t. $(\mathbf{X}, R(\mathbf{X}))$ has the unique distribution in $\Pi(P, \nu)$ with a c-cyclically monotone support.

(ii) $\exists$ a l.s.c. convex function ψ such that $R(\mathbf{x}) = q(\nabla\psi(\mathbf{x}))$ (P-a.e.)

(iii) Here, $\nabla\psi(\cdot)$ is the P-a.e. unique gradient of a convex function s.t. $\nabla\psi(G\mathbf{X}) \sim G\mathbf{H}$, where $G \sim \text{Uniform}(\mathcal{G})$ is indep. of $\mathbf{X}$ & $\mathbf{H}$

(iv) $\nabla\psi(\mathbf{X}) \overset{a.s.}{=} S(\mathbf{X}, R(\mathbf{X}))R(\mathbf{X})$ — the (generalized) signed-rank; here

$$S(\mathbf{x}, \mathbf{h}) := \arg\min_{Q \in \mathcal{G}} \|\mathbf{x} - Q\mathbf{h}\|^2$$

(v) $\nabla\psi(\cdot)$ is equivariant under the group action of $\mathcal{G}$, i.e.,

$$\nabla\psi(Q\mathbf{x}) = Q\nabla\psi(\mathbf{x}) \qquad \text{for all} \quad Q \in \mathcal{G}, \quad \text{and} \quad \mathbf{x} \text{ (a.e.)}$$

Convergence of generalized signs, ranks and signed-ranks

Fix some $k > 0$. **Assume**: (i) $\nu_n := \frac{1}{n}\sum_{i=1}^n \delta_{\mathbf{h}_i} \xrightarrow{d} \nu$ as $n \to \infty$; (ii) for $\mathbf{H}_n \sim \nu_n$, $\mathbb{E}[\|\mathbf{H}_n\|^k] \to \mathbb{E}[\|\mathbf{H}\|^k]$, as $n \to \infty$.

1. (Convergence of signed-ranks)

$$\frac{1}{n}\sum_{i=1}^n \|S_n(\mathbf{X}_i)R_n(\mathbf{X}_i) - \nabla\psi(\mathbf{X}_i)\|^k \xrightarrow{a.s.} 0.$$

Convergence of generalized signs, ranks and signed-ranks

Fix some $k > 0$. **Assume**: (i) $\nu_n := \frac{1}{n}\sum_{i=1}^n \delta_{\mathbf{h}_i} \xrightarrow{d} \nu$ as $n \to \infty$;
(ii) for $\mathbf{H}_n \sim \nu_n$, $\mathbb{E}[\|\mathbf{H}_n\|^k] \to \mathbb{E}[\|\mathbf{H}\|^k]$, as $n \to \infty$.

1. (Convergence of signed-ranks)

$$\frac{1}{n}\sum_{i=1}^n \|S_n(\mathbf{X}_i)R_n(\mathbf{X}_i) - \nabla\psi(\mathbf{X}_i)\|^k \xrightarrow{a.s.} 0.$$

2. (Convergence of ranks) If $q(\cdot)$ is continuous, then

$$\frac{1}{n}\sum_{i=1}^n \|R_n(\mathbf{X}_i) - R(\mathbf{X}_i)\|^k \xrightarrow{a.s.} 0.$$

Convergence of generalized signs, ranks and signed-ranks

Fix some $k > 0$. **Assume**: (i) $\nu_n := \frac{1}{n}\sum_{i=1}^n \delta_{\mathbf{h}_i} \xrightarrow{d} \nu$ as $n \to \infty$;
(ii) for $\mathbf{H}_n \sim \nu_n$, $\mathbb{E}[\|\mathbf{H}_n\|^k] \to \mathbb{E}[\|\mathbf{H}\|^k]$, as $n \to \infty$.

1. (Convergence of signed-ranks)

$$\frac{1}{n}\sum_{i=1}^{n} \|S_n(\mathbf{X}_i)R_n(\mathbf{X}_i) - \nabla\psi(\mathbf{X}_i)\|^k \xrightarrow{a.s.} 0.$$

2. (Convergence of ranks) If $q(\cdot)$ is continuous, then

$$\frac{1}{n}\sum_{i=1}^{n} \|R_n(\mathbf{X}_i) - R(\mathbf{X}_i)\|^k \xrightarrow{a.s.} 0.$$

3. (Convergence of signs) If $\mathcal{G}$ acts freely[a] on $\mathcal{G}B$, then

$$\frac{1}{n}\sum_{i=1}^{n} \|S_n(\mathbf{X}_i) - S(\mathbf{X}_i, R(\mathbf{X}_i))\|_F^k \xrightarrow{a.s.} 0,$$

where $S(\mathbf{x}, \mathbf{h}) := \arg\min_{Q \in \mathcal{G}} \|Q^\top \mathbf{x} - \mathbf{h}\|^2$; $\|\cdot\|_F$ is the Frobenius norm.

[a] $\mathcal{G}$ acts freely on $\mathcal{G}B$, if for $\mathbf{h} \in B$ and $Q \in \mathcal{G}$, $Q\mathbf{h} = \mathbf{h} \quad \Rightarrow \quad Q = I_p$.

Outline

Data: $\{\mathbf{X}_i\}_{i=1}^n$ iid $\mathbf{X} \sim P$ (abs. cont.) on $\mathbb{R}^p$; test $\mathrm{H}_0 : \mathbf{X} \stackrel{d}{=} Q\mathbf{X} \ \ \forall Q \in \mathcal{G}$

Under H_0, the generalized signs $S_n(\mathbf{X}_1), \ldots, S_n(\mathbf{X}_n)$ are iid $\mathrm{Uniform}(\mathcal{G})$

Generalized sign test: When $\mathcal{G}$ is finite

Suppose $\mathcal{G} = \{g_1, \ldots, g_m\}$ is a finite group of size m which acts freely. Let

$$Y_j := \sum_{i=1}^n \mathbf{1}(S_n(\mathbf{X}_i) = g_j), \qquad j = 1, \ldots, m.$$

Under H_0,

$$(Y_1, \ldots, Y_m) \sim \mathrm{Multinomial}\left(n, \frac{1}{m}\mathbf{1}_m\right).$$

Distribution-free: Generalizes the usual sign test beyond $p = 1$!

Data: $\{\mathbf{X}_i\}_{i=1}^n$ iid $\mathbf{X} \sim P$ (abs. cont.) on $\mathbb{R}^p$; test $\mathrm{H}_0 : \mathbf{X} \stackrel{d}{=} Q\mathbf{X} \;\; \forall Q \in \mathcal{G}$

Under H_0, the generalized signs $S_n(\mathbf{X}_1), \ldots, S_n(\mathbf{X}_n)$ are iid $\mathrm{Uniform}(\mathcal{G})$

Generalized sign test: When $\mathcal{G}$ is finite

Suppose $\mathcal{G} = \{g_1, \ldots, g_m\}$ is a finite group of size m which acts freely. Let

$$Y_j := \sum_{i=1}^n \mathbf{1}(S_n(\mathbf{X}_i) = g_j), \qquad j = 1, \ldots, m.$$

Under H_0,

$$(Y_1, \ldots, Y_m) \sim \mathrm{Multinomial}\left(n, \frac{1}{m}\mathbf{1}_m\right).$$

Distribution-free: Generalizes the usual sign test beyond $p = 1$!

If m is large, take generalized sign test based on $V_n := \frac{1}{\sqrt{n}} \sum_{i=1}^n S_n(\mathbf{X}_i)$

Central symmetry: $\frac{1}{p}\|V_n\|_F^2 \xrightarrow{d} \chi_1^2$

Sign symmetry: $\|V_n\|_F^2 \xrightarrow{d} \chi_p^2$

Spherical symmetry: $p\|V_n\|_F^2 \xrightarrow{d} \chi_{p^2}^2$

Generalized Wilcoxon Signed-rank test

- The generalized Wilcoxon signed-rank statistic is

$$\mathbf{W}_n := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} S_n(\mathbf{X}_i) R_n(\mathbf{X}_i)$$

- $\mathbf{W}_n$ is distribution-free under $\mathrm{H}_0 : \mathbf{X} \stackrel{d}{=} Q\mathbf{X} \ \ \forall\, Q \in \mathcal{G}$

Generalized Wilcoxon Signed-rank test

- The generalized Wilcoxon signed-rank statistic is

$$\mathbf{W}_n := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} S_n(\mathbf{X}_i) R_n(\mathbf{X}_i)$$

- $\mathbf{W}_n$ is distribution-free under $\mathrm{H}_0 : \mathbf{X} \overset{d}{=} Q\mathbf{X} \;\; \forall\, Q \in \mathcal{G}$

Asymptotic normality of $\mathbf{W}_n$ [Huang & S. (2023+)]

Suppose: (i) $\nu_n := \frac{1}{n}\sum_{i=1}^{n} \delta_{\mathbf{h}_i} \xrightarrow{d} \nu$ & 2nd moment convergence;
(ii) $\mathbb{E}[G] = \mathbf{0}_{p \times p}$ where $G \sim \mathrm{Uniform}(\mathcal{G})$;
Then:

$$\mathbf{W}_n \xrightarrow{d} N\left(\mathbf{0}_p, \Sigma_{\mathrm{GH}}\right),$$

where Σ_{GH} be the covariance matrix of $G\mathbf{H}$, with $G \perp\!\!\!\perp \mathbf{H}$ (here $\mathbf{H} \sim \nu$).

Generalized Wilcoxon Signed-rank test

- The generalized Wilcoxon signed-rank statistic is

$$\mathbf{W}_n := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} S_n(\mathbf{X}_i) R_n(\mathbf{X}_i)$$

- $\mathbf{W}_n$ is distribution-free under $\mathrm{H}_0 : \mathbf{X} \stackrel{d}{=} Q\mathbf{X} \ \ \forall\, Q \in \mathcal{G}$

Asymptotic normality of $\mathbf{W}_n$ [Huang & S. (2023+)]

Suppose: (i) $\nu_n := \frac{1}{n} \sum_{i=1}^{n} \delta_{\mathbf{h}_i} \xrightarrow{d} \nu$ & 2nd moment convergence;
(ii) $\mathbb{E}[G] = \mathbf{0}_{p \times p}$ where $G \sim \mathrm{Uniform}(\mathcal{G})$;
Then:

$$\mathbf{W}_n \xrightarrow{d} N\left(\mathbf{0}_p, \Sigma_{\mathrm{GH}}\right),$$

where Σ_{GH} be the covariance matrix of $G\mathbf{H}$, with $G \perp\!\!\!\perp \mathbf{H}$ (here $\mathbf{H} \sim \nu$).

- The Wilcoxon signed-rank test rejects H_0 for

$$\mathbf{W}_n^{\top} \Sigma_{\mathrm{GH}}^{-1} \mathbf{W}_n \geq c_\alpha$$

- c_α is the universal cut-off; well-approximable by the χ^2_p-quantile

The generalized Wilcoxon signed-rank statistic is

$$\mathbf{W}_n := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} S_n(\mathbf{X}_i) R_n(\mathbf{X}_i)$$

Test for $\mathcal{G}$-symmetry: $\mathrm{H}_0 : \mathbf{X} \stackrel{d}{=} Q\mathbf{X} \ \forall\, Q \in \mathcal{G}, \quad$ vs. $\quad \mathrm{H}_1$: not H_0

Consistency of WSR for testing $\mathcal{G}$-symmetry [Huang & S. (2023+)]

Assume: (i) $\nu_n := \frac{1}{n}\sum_{i=1}^{n} \delta_{\mathbf{h}_i} \xrightarrow{d} \nu$; (ii) 1st moment convergence
Then, the Wilcoxon signed-rank test which rejects H_0 for

$$\mathbf{W}_n^\top \Sigma_{\mathrm{GH}}^{-1} \mathbf{W}_n \geq c_\alpha$$

is consistent against all alternatives for which

$$\mathbb{E}[\nabla\psi(\mathbf{X})] \neq \mathbf{0}.^{a}$$

[a] $\mathbb{E}[\nabla\psi(\mathbf{X})] \neq \mathbf{0}$ holds for location shift models if $\psi(\cdot)$ is strictly convex & $-I_p \in \mathcal{G}$.

Asymptotics under local alternatives [Huang & S. (2023+)]

Let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ be iid $f(\cdot - \theta)$ on $\mathbb{R}^p$; here f is $\mathcal{G}$-symmetric distribution. Consider testing:

$$\mathrm{H}_0 : \theta = \mathbf{0}_p \qquad \text{versus} \qquad \mathrm{H}_1 : \theta = \frac{\mu}{\sqrt{n}}; \quad \mu \neq 0_p \in \mathbb{R}^p$$

Under 'suitable' assumptions[a] and standard regularity conditions of the parametric family $\{f(\cdot - \theta)\}_{\theta \in \mathbb{R}^p}$ (e.g., QMD), we have, under H_1:

$$\mathbf{W}_n \xrightarrow{d} N(\gamma, \Sigma_{\mathrm{GH}}),$$

where $\gamma := \mathbb{E}_{\mathrm{H}_0}\left[\nabla\psi(\mathbf{X}) \frac{\mu^\top \nabla f(\mathbf{X})}{f(\mathbf{X})}\right] \in \mathbb{R}^p$.

[a](i) $\nu_n := \frac{1}{n}\sum_{i=1}^n \delta_{\mathbf{h}_i} \xrightarrow{d} \nu$ & 2nd moment convergence; (ii) $\mathbb{E}[G] = \mathbf{0}_{p\times p}$

- Generalized WSR test: $\mathbf{W}_n^\top \Sigma_{\mathrm{GH}}^{-1} \mathbf{W}_n \xrightarrow{d} \left\| \Sigma_{\mathrm{GH}}^{-1/2}\gamma + N(\mathbf{0}, I_p) \right\|^2$
- The non-centrality parameter of generalized WSR test is $\|\Sigma_{\mathrm{GH}}^{-1/2}\gamma\|^2$

Asymptotics under local alternatives [Huang & S. (2023+)]

Let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ be iid $f(\cdot - \theta)$ on $\mathbb{R}^p$; here f is $\mathcal{G}$-symmetric distribution. Consider testing:

$$\mathrm{H}_0 : \theta = \mathbf{0}_p \qquad \text{versus} \qquad \mathrm{H}_1 : \theta = \frac{\mu}{\sqrt{n}}; \quad \mu \neq 0_p \in \mathbb{R}^p$$

Under 'suitable' assumptions[a] and standard regularity conditions of the parametric family $\{f(\cdot - \theta)\}_{\theta \in \mathbb{R}^p}$ (e.g., QMD), we have, under H_1:

$$\mathbf{W}_n \xrightarrow{d} N(\gamma, \Sigma_{\mathrm{GH}}),$$

where $\gamma := \mathbb{E}_{\mathrm{H_0}}\left[\nabla\psi(\mathbf{X})\frac{\mu^\top \nabla f(\mathbf{X})}{f(\mathbf{X})}\right] \in \mathbb{R}^p$.

---

[a] (i) $\nu_n := \frac{1}{n}\sum_{i=1}^n \delta_{\mathbf{h}_i} \xrightarrow{d} \nu$ & 2nd moment convergence; (ii) $\mathbb{E}[G] = \mathbf{0}_{p\times p}$

- Generalized WSR test: $\mathbf{W}_n^\top \Sigma_{\mathrm{GH}}^{-1} \mathbf{W}_n \xrightarrow{d} \left\|\Sigma_{\mathrm{GH}}^{-1/2}\gamma + N(\mathbf{0}, I_p)\right\|^2$
- The non-centrality parameter of generalized WSR test is $\|\Sigma_{\mathrm{GH}}^{-1/2}\gamma\|^2$

Question: How does this compare with Hotelling's T^2 test?

- **Question**: How to compare two consistent tests S_n and T_n?
- **Asymptotic relative (Pitman) efficiency** (ARE) [Pitman (1948), Serfling (1980), Lehmann & Romano (2005), van der Vaart (1998)]

- **Question**: How to compare two consistent tests S_n and T_n?
- **Asymptotic relative (Pitman) efficiency** (ARE) [Pitman (1948), Serfling (1980), Lehmann & Romano (2005), van der Vaart (1998)]

- **Data**: $\mathbf{X}_1, \ldots, \mathbf{X}_n \overset{iid}{\sim} \mathrm{P}_\theta$
- $\{P_\theta\}_{\theta \in \Theta \subset \mathbb{R}^p}$: "smooth" (satisfies DQM) parametric family

- **Question**: How to compare two consistent tests S_n and T_n?
- **Asymptotic relative (Pitman) efficiency** (ARE) [Pitman (1948), Serfling (1980), Lehmann & Romano (2005), van der Vaart (1998)]

- **Data**: $\mathbf{X}_1, \ldots, \mathbf{X}_n \overset{iid}{\sim} \mathrm{P}_\theta$
- $\{P_\theta\}_{\theta \in \Theta \subset \mathbb{R}^p}$: "smooth" (satisfies DQM) parametric family

- **Test** $\mathrm{H}_0 : \theta = 0$ vs. $\mathrm{H}_1 : \theta = \Delta$; $\Delta \to 0$

- **Question**: How to compare two consistent tests S_n and T_n?
- **Asymptotic relative (Pitman) efficiency** (ARE) [Pitman (1948), Serfling (1980), Lehmann & Romano (2005), van der Vaart (1998)]

- **Data**: $\mathbf{X}_1, \ldots, \mathbf{X}_n \overset{iid}{\sim} \mathrm{P}_\theta$
- $\{P_\theta\}_{\theta \in \Theta \subset \mathbb{R}^p}$: "smooth" (satisfies DQM) parametric family

- **Test** $\mathrm{H}_0 : \theta = 0$ vs. $\mathrm{H}_1 : \theta = \Delta$; $\Delta \to 0$
- Fix $\alpha \in (0, 1)$ (level) and $\beta \in (\alpha, 1)$ (power)

- **Question**: How to compare two consistent tests S_n and T_n?
- **Asymptotic relative (Pitman) efficiency** (ARE) [Pitman (1948), Serfling (1980), Lehmann & Romano (2005), van der Vaart (1998)]

- **Data**: $\mathbf{X}_1, \ldots, \mathbf{X}_n \overset{iid}{\sim} \mathrm{P}_\theta$
- $\{P_\theta\}_{\theta \in \Theta \subset \mathbb{R}^p}$: "smooth" (satisfies DQM) parametric family

- **Test** $\mathrm{H}_0 : \theta = 0$ vs. $\mathrm{H}_1 : \theta = \Delta$; $\Delta \to 0$
- Fix $\alpha \in (0,1)$ (level) and $\beta \in (\alpha, 1)$ (power)
- Let $N_\Delta(T_\cdot) \equiv N_\Delta$ denote the minimum number of samples s.t.:

$$\mathbb{E}_{\mathrm{H}_0}[T_{N_\Delta}] = \alpha \qquad \text{and} \qquad \mathbb{E}_{\mathrm{H}_1}[T_{N_\Delta}] \geq \beta$$

- **Question**: How to compare two consistent tests S_n and T_n?
- **Asymptotic relative (Pitman) efficiency** (ARE) [Pitman (1948), Serfling (1980), Lehmann & Romano (2005), van der Vaart (1998)]

- **Data**: $\mathbf{X}_1, \ldots, \mathbf{X}_n \overset{iid}{\sim} \mathrm{P}_\theta$
- $\{P_\theta\}_{\theta \in \Theta \subset \mathbb{R}^p}$: "smooth" (satisfies DQM) parametric family

- **Test** $\mathrm{H}_0 : \theta = 0$ vs. $\mathrm{H}_1 : \theta = \Delta$; $\Delta \to 0$
- Fix $\alpha \in (0, 1)$ (level) and $\beta \in (\alpha, 1)$ (power)
- Let $N_\Delta(T_\cdot) \equiv N_\Delta$ denote the minimum number of samples s.t.:

$$\mathbb{E}_{\mathrm{H}_0}[T_{N_\Delta}] = \alpha \qquad \text{and} \qquad \mathbb{E}_{\mathrm{H}_1}[T_{N_\Delta}] \geq \beta$$

- The asymptotic (Pitman) efficiency of S_n w.r.t. T_n is given by

$$\mathrm{ARE}\,(S_n, T_n) := \lim_{\Delta \to 0} \frac{N_\Delta(T_\cdot)}{N_\Delta(S_\cdot)}$$

- **Question**: How to compare two consistent tests S_n and T_n?
- **Asymptotic relative (Pitman) efficiency** (ARE) [Pitman (1948), Serfling (1980), Lehmann & Romano (2005), van der Vaart (1998)]

- **Data**: $\mathbf{X}_1, \ldots, \mathbf{X}_n \overset{iid}{\sim} \mathrm{P}_\theta$
- $\{P_\theta\}_{\theta \in \Theta \subset \mathbb{R}^p}$: "smooth" (satisfies DQM) parametric family

- **Test** $\mathrm{H}_0 : \theta = 0$ vs. $\mathrm{H}_1 : \theta = \Delta$; $\Delta \to 0$
- Fix $\alpha \in (0, 1)$ (level) and $\beta \in (\alpha, 1)$ (power)
- Let $N_\Delta(T_\cdot) \equiv N_\Delta$ denote the minimum number of samples s.t.:

$$\mathbb{E}_{\mathrm{H}_0}[T_{N_\Delta}] = \alpha \qquad \text{and} \qquad \mathbb{E}_{\mathrm{H}_1}[T_{N_\Delta}] \geq \beta$$

- The asymptotic (Pitman) efficiency of S_n w.r.t. T_n is given by

$$\mathrm{ARE}\,(S_n, T_n) := \lim_{\Delta \to 0} \frac{N_\Delta(T_\cdot)}{N_\Delta(S_\cdot)}$$

$\mathrm{ARE}\,(S_n, T_n)$ can depend on α and β, but in some cases they don't!

Hotelling T^2: $n\bar{\mathbf{X}}^\top S_n^{-1}\bar{\mathbf{X}}$ where

$$S_n := \frac{1}{n-1}\sum_{i=1}^n(\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})^\top \xrightarrow{p} \Sigma_{\mathbf{X}} := \mathbb{E}\left(\mathbf{X} - \mathbb{E}\mathbf{X}\right)\left(\mathbf{X} - \mathbb{E}\mathbf{X}\right)^\top.$$

Generalized **WSR:** $\mathbf{W}_n^\top \Sigma_{\mathrm{GH}}^{-1}\mathbf{W}_n$

- $\mathbf{X}_1, \ldots, \mathbf{X}_n \overset{iid}{\sim} f(\cdot - \theta)$; f is $\mathcal{G}$-symmetric
- $\{f(\cdot - \theta)\}$: "smooth" (satisfies DQM) parametric family
- Consider $\mathrm{H}_0 : \theta = \mathbf{0}_p$ vs. $\mathrm{H}_1 : \theta = \frac{\mu}{\sqrt{n}}$; $\mu \neq 0 \in \mathbb{R}^p$

Result: $\mathrm{ARE}\left(\mathbf{W}_n, \bar{\mathbf{X}}_n\right) = \frac{\|\Sigma_{\mathrm{GH}}^{-1/2}\gamma\|^2}{\|\Sigma_{\mathbf{X}}^{-1/2}\mu\|^2}$; $\gamma = \mathbb{E}_{\mathrm{H}_0}\left[\nabla\psi(\mathbf{X})\frac{\mu^\top\nabla f(\mathbf{X})}{f(\mathbf{X})}\right]$.

Hotelling T^2: $n\bar{\mathbf{X}}^\top S_n^{-1}\bar{\mathbf{X}}$ where

$S_n := \frac{1}{n-1}\sum_{i=1}^n(\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})^\top \overset{p}{\to} \Sigma_{\mathbf{X}} := \mathbb{E}\left(\mathbf{X} - \mathbb{E}\mathbf{X}\right)\left(\mathbf{X} - \mathbb{E}\mathbf{X}\right)^\top$.

Generalized **WSR:** $\mathbf{W}_n^\top \Sigma_{\mathrm{GH}}^{-1}\mathbf{W}_n$

- $\mathbf{X}_1, \ldots, \mathbf{X}_n \overset{iid}{\sim} f(\cdot - \theta)$; f is $\mathcal{G}$-symmetric
- $\{f(\cdot - \theta)\}$: "smooth" (satisfies DQM) parametric family
- Consider $\mathrm{H}_0 : \theta = \mathbf{0}_p$ vs. $\mathrm{H}_1 : \theta = \frac{\mu}{\sqrt{n}}$; $\mu \neq 0 \in \mathbb{R}^p$

Result: $\mathrm{ARE}\left(\mathbf{W}_n, \bar{\mathbf{X}}_n\right) = \frac{\|\Sigma_{\mathrm{GH}}^{-1/2}\gamma\|^2}{\|\Sigma_{\mathbf{X}}^{-1/2}\mu\|^2}$; $\gamma = \mathbb{E}_{\mathrm{H}_0}\left[\nabla\psi(\mathbf{X})\frac{\mu^\top \nabla f(\mathbf{X})}{f(\mathbf{X})}\right]$.

Some observations

- Expression of $\mathrm{ARE}\left(\mathbf{W}_n, \bar{\mathbf{X}}_n\right)$ does not depend on α and β
- $\mathrm{ARE}\left(\mathbf{W}_n, \bar{\mathbf{X}}_n\right)$ can depend on ν [Deb, Bhattacharya & S. (2021)]

Hotelling T^2: $n\bar{\mathbf{X}}^\top S_n^{-1}\bar{\mathbf{X}}$ where

$S_n := \frac{1}{n-1}\sum_{i=1}^n(\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})^\top \xrightarrow{p} \Sigma_{\mathbf{X}} := \mathbb{E}\left(\mathbf{X} - \mathbb{E}\mathbf{X}\right)\left(\mathbf{X} - \mathbb{E}\mathbf{X}\right)^\top$.

Generalized **WSR:** $\mathbf{W}_n^\top \Sigma_{\mathrm{GH}}^{-1}\mathbf{W}_n$

- $\mathbf{X}_1, \ldots, \mathbf{X}_n \overset{iid}{\sim} f(\cdot - \theta)$; f is $\mathcal{G}$-symmetric
- $\{f(\cdot - \theta)\}$: "smooth" (satisfies DQM) parametric family
- Consider $\mathrm{H}_0 : \theta = \mathbf{0}_p$ vs. $\mathrm{H}_1 : \theta = \frac{\mu}{\sqrt{n}}$; $\mu \neq 0 \in \mathbb{R}^p$

Result: $\mathrm{ARE}\left(\mathbf{W}_n, \bar{\mathbf{X}}_n\right) = \frac{\|\Sigma_{\mathrm{GH}}^{-1/2}\gamma\|^2}{\|\Sigma_{\mathbf{X}}^{-1/2}\mu\|^2}$; $\gamma = \mathbb{E}_{\mathrm{H}_0}\left[\nabla\psi(\mathbf{X})\frac{\mu^\top\nabla f(\mathbf{X})}{f(\mathbf{X})}\right]$.

Some observations

- Expression of $\mathrm{ARE}\left(\mathbf{W}_n, \bar{\mathbf{X}}_n\right)$ does not depend on α and β
- $\mathrm{ARE}\left(\mathbf{W}_n, \bar{\mathbf{X}}_n\right)$ can depend on ν [Deb, Bhattacharya & S. (2021)]

Can we lower bound ARE for sub-classes of multivariate dists., i.e.,

$$\min_{\mathcal{F}} \mathrm{ARE}\left(\mathbf{W}_n, \bar{\mathbf{X}}_n\right) = ??$$

$\mathbf{X}_1, \ldots, \mathbf{X}_n$ be iid $f(\cdot - \theta)$ on $\mathbb{R}^p$; f is density of a $\mathcal{G}$-symmetric dist.

Gaussian case: f is density of $N(\mathbf{0}_p, \Sigma_X)$, where Σ_X is p.d. (unknown)

Theorem [Huang & S. (2023+)]

Suppose: (i) $\nu_n := \frac{1}{n}\sum_{i=1}^n \delta_{\mathbf{h}_i} \xrightarrow{d} \nu$ & 2nd moment convergence;
(ii) $\mathbb{E}[G] = \mathbf{0}_{p\times p}$ where $G \sim \mathrm{Uniform}(\mathcal{G})$.

If $G\mathbf{H} \sim N(\mathbf{0}_p, I_p)$, then

$$\mathrm{ARE}\,(\mathbf{W}_n, \bar{\mathbf{X}}_n) = 1.$$

$\mathbf{X}_1, \ldots, \mathbf{X}_n$ be iid $f(\cdot - \theta)$ on $\mathbb{R}^p$; f is density of a $\mathcal{G}$-symmetric dist.

Gaussian case: f is density of $N(\mathbf{0}_p, \Sigma_X)$, where Σ_X is p.d. (unknown)

Theorem [Huang & S. (2023+)]

Suppose: (i) $\nu_n := \frac{1}{n}\sum_{i=1}^n \delta_{\mathbf{h}_i} \xrightarrow{d} \nu$ & 2nd moment convergence;
(ii) $\mathbb{E}[G] = \mathbf{0}_{p\times p}$ where $G \sim \mathrm{Uniform}(\mathcal{G})$.

If $G\mathbf{H} \sim N(\mathbf{0}_p, I_p)$, then

$$\mathrm{ARE}\,(\mathbf{W}_n, \bar{\mathbf{X}}_n) = 1.$$

If $G\mathbf{H}$ has the spherical uniform distribution[a], then

$$\mathrm{ARE}\,(\mathbf{W}_n, \bar{\mathbf{X}}_n) = \kappa_p \geq \begin{cases} 0.95, & \text{for } p < 5 \\ 0.648, & \forall\ p \end{cases}$$

[a] $\kappa_1 = 3/\pi$ reduces to the classical ARE of the WSR test against the t-test.

$\mathbf{X}_1, \ldots, \mathbf{X}_n$ be iid $f(\cdot - \theta)$ on $\mathbb{R}^p$; f is density of a $\mathcal{G}$-symmetric dist.

Gaussian case: f is density of $N(\mathbf{0}_p, \Sigma_X)$, where Σ_X is p.d. (unknown)

Theorem [Huang & S. (2023+)]

Suppose: (i) $\nu_n := \frac{1}{n}\sum_{i=1}^n \delta_{\mathbf{h}_i} \xrightarrow{d} \nu$ & 2nd moment convergence;
(ii) $\mathbb{E}[G] = \mathbf{0}_{p\times p}$ where $G \sim \mathrm{Uniform}(\mathcal{G})$.

If $G\mathbf{H} \sim N(\mathbf{0}_p, I_p)$, then

$$\mathrm{ARE}\,(\mathbf{W}_n, \bar{\mathbf{X}}_n) = 1.$$

If $G\mathbf{H}$ has the spherical uniform distribution[a], then

$$\mathrm{ARE}\,(\mathbf{W}_n, \bar{\mathbf{X}}_n) = \kappa_p \geq \begin{cases} 0.95, & \text{for } p < 5 \\ 0.648, & \forall\ p \end{cases}$$

[a] $\kappa_1 = 3/\pi$ reduces to the classical ARE of the WSR test against the t-test.

- Generalizes Chernoff & Savage (1958)
- ARE can be arbitrarily large (can tend to $+\infty$) for heavy tailed dists.

$\mathbf{X}_1, \ldots, \mathbf{X}_n$ be iid $f(\cdot - \theta)$ on $\mathbb{R}^p$; here f is density of a $\mathcal{G}$-symmetric dist.

Independent components

$\mathcal{F}_{\mathrm{ind}} = \{f(\cdot - \theta)\}$ has density $f(z_1, \ldots, z_p) = \prod_{i=1}^p f_i(z_i)$

Theorem [Huang & S. (2023+)]

Suppose: (i) $\nu_n := \frac{1}{n}\sum_{i=1}^n \delta_{\mathbf{h}_i} \xrightarrow{d} \nu$ & 2nd moment convergence;
(ii) $\mathbb{E}[G] = \mathbf{0}_{p\times p}$ where $G \sim \mathrm{Uniform}(\mathcal{G})$.

If $G\mathbf{H} \sim \mathrm{Uniform}(-1, 1)^p$, then

$$\min_{\mathcal{F}_{\mathrm{ind}}} \mathrm{ARE}\,(\mathbf{W}_n, \bar{\mathbf{X}}_n) = 0.864.$$

$\mathbf{X}_1, \ldots, \mathbf{X}_n$ be iid $f(\cdot - \theta)$ on $\mathbb{R}^p$; here f is density of a $\mathcal{G}$-symmetric dist.

Independent components

$\mathcal{F}_{\mathrm{ind}} = \{f(\cdot - \theta)\}$ has density $f(z_1, \ldots, z_p) = \prod_{i=1}^p f_i(z_i)$

Theorem [Huang & S. (2023+)]

Suppose: (i) $\nu_n := \frac{1}{n} \sum_{i=1}^n \delta_{\mathbf{h}_i} \xrightarrow{d} \nu$ & 2nd moment convergence;
(ii) $\mathbb{E}[G] = \mathbf{0}_{p \times p}$ where $G \sim \mathrm{Uniform}(\mathcal{G})$.

If $G\mathbf{H} \sim \mathrm{Uniform}(-1, 1)^p$, then

$$\min_{\mathcal{F}_{\mathrm{ind}}} \mathrm{ARE}\left(\mathbf{W}_n, \bar{\mathbf{X}}_n\right) = 0.864.$$

If $G\mathbf{H} \sim N(\mathbf{0}_p, I_p)$, then

$$\min_{\mathcal{F}_{\mathrm{ind}}} \mathrm{ARE}\left(\mathbf{W}_n, \bar{\mathbf{X}}_n\right) = 1.$$

$\mathbf{X}_1, \ldots, \mathbf{X}_n$ be iid $f(\cdot - \theta)$ on $\mathbb{R}^p$; here f is density of a $\mathcal{G}$-symmetric dist.

Independent components

$\mathcal{F}_{\text{ind}} = \{f(\cdot - \theta)\}$ has density $f(z_1, \ldots, z_p) = \prod_{i=1}^p f_i(z_i)$

Theorem [Huang & S. (2023+)]

Suppose: (i) $\nu_n := \frac{1}{n}\sum_{i=1}^n \delta_{\mathbf{h}_i} \xrightarrow{d} \nu$ & 2nd moment convergence;
(ii) $\mathbb{E}[G] = \mathbf{0}_{p\times p}$ where $G \sim \text{Uniform}(\mathcal{G})$.

If $G\mathbf{H} \sim \text{Uniform}(-1, 1)^p$, then

$$\min_{\mathcal{F}_{\text{ind}}} \text{ARE}(\mathbf{W}_n, \bar{\mathbf{X}}_n) = 0.864.$$

If $G\mathbf{H} \sim N(\mathbf{0}_p, I_p)$, then

$$\min_{\mathcal{F}_{\text{ind}}} \text{ARE}(\mathbf{W}_n, \bar{\mathbf{X}}_n) = 1.$$

- Generalizes Hodges & Lehmann (1956), Chernoff & Savage (1958)
- ARE can be arbitrarily large (can tend to $+\infty$) for heavy tailed dists.

$\mathbf{X}_1, \ldots, \mathbf{X}_n$ be iid $f(\cdot - \theta)$ on $\mathbb{R}^p$; here f is density of a $\mathcal{G}$-symmetric dist.

Elliptically symmetric distributions

$\mathcal{F}_{\text{ell}} = \{f(\cdot - \theta)\}$ is class of elliptically symmetric distributions on $\mathbb{R}^p$, i.e.,

$$f(\mathbf{x}) \propto (\det(\Sigma_X))^{-\frac{1}{2}} \underline{f}\left(\mathbf{x}^\top \Sigma_X^{-1} \mathbf{x}\right), \quad \text{for all } \mathbf{x} \in \mathbb{R}^p$$

$\mathbf{X}_1, \ldots, \mathbf{X}_n$ be iid $f(\cdot - \theta)$ on $\mathbb{R}^p$; here f is density of a $\mathcal{G}$-symmetric dist.

Elliptically symmetric distributions

$\mathcal{F}_{\text{ell}} = \{f(\cdot - \theta)\}$ is class of elliptically symmetric distributions on $\mathbb{R}^p$, i.e.,

$$f(\mathbf{x}) \propto (\det(\Sigma_X))^{-\frac{1}{2}} \underline{f}\left(\mathbf{x}^\top \Sigma_X^{-1} \mathbf{x}\right), \quad \text{for all } \mathbf{x} \in \mathbb{R}^p$$

Theorem [Huang & S. (2023+)]

Suppose: (i) $\nu_n := \frac{1}{n}\sum_{i=1}^n \delta_{\mathbf{h}_i} \xrightarrow{d} \nu$ & 2nd moment convergence;
(ii) $\mathbb{E}[G] = \mathbf{0}_{p \times p}$ where $G \sim \text{Uniform}(\mathcal{G})$.

If $G\mathbf{H} \sim N(\mathbf{0}_p, I_p)$, then $\min_{\mathcal{F}_{\text{ell}}} \text{ARE}(\mathbf{W}_n, \bar{\mathbf{X}}_n) = 1$.

$\mathbf{X}_1, \ldots, \mathbf{X}_n$ be iid $f(\cdot - \theta)$ on $\mathbb{R}^p$; here f is density of a $\mathcal{G}$-symmetric dist.

Elliptically symmetric distributions

$\mathcal{F}_{\text{ell}} = \{f(\cdot - \theta)\}$ is class of elliptically symmetric distributions on $\mathbb{R}^p$, i.e.,

$$f(\mathbf{x}) \propto (\det(\Sigma_X))^{-\frac{1}{2}} \underline{f}\left(\mathbf{x}^\top \Sigma_X^{-1} \mathbf{x}\right), \quad \text{for all } \mathbf{x} \in \mathbb{R}^p$$

Theorem [Huang & S. (2023+)]

Suppose: (i) $\nu_n := \frac{1}{n} \sum_{i=1}^n \delta_{\mathbf{h}_i} \xrightarrow{d} \nu$ & 2nd moment convergence;
(ii) $\mathbb{E}[G] = \mathbf{0}_{p \times p}$ where $G \sim \text{Uniform}(\mathcal{G})$.

If $G\mathbf{H} \sim N(\mathbf{0}_p, I_p)$, then $\min_{\mathcal{F}_{\text{ell}}} \text{ARE}\,(\mathbf{W}_n, \bar{\mathbf{X}}_n) = 1$.

If $G\mathbf{H}$ has the spherical uniform distribution, then

$$\min_{\mathcal{F}_{\text{ell}}} \text{ARE}\,(\mathbf{W}_n, \bar{\mathbf{X}}_n) \geq 0.648.$$

$\mathbf{X}_1, \ldots, \mathbf{X}_n$ be iid $f(\cdot - \theta)$ on $\mathbb{R}^p$; here f is density of a $\mathcal{G}$-symmetric dist.

Elliptically symmetric distributions

$\mathcal{F}_{\mathrm{ell}} = \{f(\cdot - \theta)\}$ is class of elliptically symmetric distributions on $\mathbb{R}^p$, i.e.,

$$f(\mathbf{x}) \propto (\det(\Sigma_X))^{-\frac{1}{2}} \underline{f}\left(\mathbf{x}^\top \Sigma_X^{-1} \mathbf{x}\right), \quad \text{for all } \mathbf{x} \in \mathbb{R}^p$$

Theorem [Huang & S. (2023+)]

Suppose: (i) $\nu_n := \frac{1}{n} \sum_{i=1}^n \delta_{\mathbf{h}_i} \xrightarrow{d} \nu$ & 2nd moment convergence;
(ii) $\mathbb{E}[G] = \mathbf{0}_{p \times p}$ where $G \sim \mathrm{Uniform}(\mathcal{G})$.

If $G\mathbf{H} \sim N(\mathbf{0}_p, I_p)$, then $\min_{\mathcal{F}_{\mathrm{ell}}} \mathrm{ARE}(\mathbf{W}_n, \bar{\mathbf{X}}_n) = 1$.

If $G\mathbf{H}$ has the spherical uniform distribution, then

$$\min_{\mathcal{F}_{\mathrm{ell}}} \mathrm{ARE}(\mathbf{W}_n, \bar{\mathbf{X}}_n) \geq 0.648.$$

This generalizes the famous result of Chernoff and Savage (1958)

$\mathbf{X}_1, \ldots, \mathbf{X}_n$ be iid $f(\cdot - \theta)$ on $\mathbb{R}^p$; here f is density of a $\mathcal{G}$-symmetric dist.

Elliptically symmetric distributions

$\mathcal{F}_{\text{ell}} = \{f(\cdot - \theta)\}$ is class of elliptically symmetric distributions on $\mathbb{R}^p$, i.e.,

$$f(\mathbf{x}) \propto (\det(\Sigma_X))^{-\frac{1}{2}} \underline{f}\left(\mathbf{x}^\top \Sigma_X^{-1} \mathbf{x}\right), \quad \text{for all } \mathbf{x} \in \mathbb{R}^p$$

Theorem [Huang & S. (2023+)]

Suppose: (i) $\nu_n := \frac{1}{n} \sum_{i=1}^n \delta_{\mathbf{h}_i} \xrightarrow{d} \nu$ & 2nd moment convergence;
(ii) $\mathbb{E}[G] = \mathbf{0}_{p \times p}$ where $G \sim \text{Uniform}(\mathcal{G})$.

If $G\mathbf{H} \sim N(\mathbf{0}_p, I_p)$, then $\quad \min_{\mathcal{F}_{\text{ell}}} \text{ARE}\,(\mathbf{W}_n, \bar{\mathbf{X}}_n) = 1.$

If $G\mathbf{H}$ has the spherical uniform distribution, then

$$\min_{\mathcal{F}_{\text{ell}}} \text{ARE}\,(\mathbf{W}_n, \bar{\mathbf{X}}_n) \geq 0.648.$$

This generalizes the famous result of Chernoff and Savage (1958)

Similar lower bounds can also be obtained for other sub-classes of multivariate distributions (e.g., the model for ICA)

Distribution-free confidence set for the center of symmetry

- $\mathbf{X} \sim P$ on $\mathbb{R}^p$ has a $\mathcal{G}$-symmetric distribution with center of symmetry $\boldsymbol{\theta}^*$ (unknown) if

$$(\mathbf{X} - \boldsymbol{\theta}^*) \stackrel{d}{=} Q(\mathbf{X} - \boldsymbol{\theta}^*), \qquad \forall Q \in \mathcal{G}$$

- **Goal**: Given data $\mathbf{X}_1, \ldots, \mathbf{X}_n$ iid P, find a distribution-free confidence set for $\boldsymbol{\theta}^*$

Distribution-free confidence set for the center of symmetry

- $\mathbf{X} \sim P$ on $\mathbb{R}^p$ has a $\mathcal{G}$-symmetric distribution with center of symmetry $\boldsymbol{\theta}^*$ (unknown) if

$$(\mathbf{X} - \boldsymbol{\theta}^*) \stackrel{d}{=} Q(\mathbf{X} - \boldsymbol{\theta}^*), \qquad \forall Q \in \mathcal{G}$$

- **Goal**: Given data $\mathbf{X}_1, \ldots, \mathbf{X}_n$ iid P, find a distribution-free confidence set for $\boldsymbol{\theta}^*$

- **Idea**: Invert the collection of hypothesis tests

- Fix $\boldsymbol{\theta} \in \mathbb{R}^p$, and test

$$\mathrm{H}_{0,\boldsymbol{\theta}} : (\mathbf{X} - \boldsymbol{\theta}) \stackrel{d}{=} Q(\mathbf{X} - \boldsymbol{\theta}), \quad \forall Q \in \mathcal{G}$$

  using generalized Wilcoxon signed-rank test with $\{\mathbf{X}_i - \boldsymbol{\theta}\}_{i=1}^n$

- $\mathcal{C} := \{\boldsymbol{\theta} : \mathrm{H}_{0,\boldsymbol{\theta}} \text{ is accepted}\}$ — exact $(1-\alpha)$ confidence set for $\boldsymbol{\theta}^*$

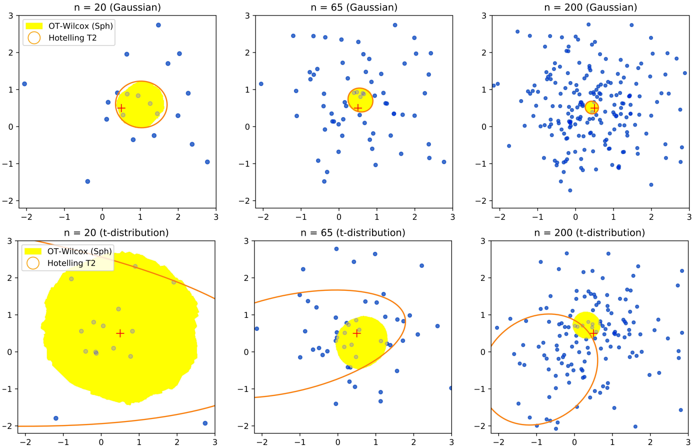


Figure: Confidence sets for $\boldsymbol{\theta}^*$ as the sample size n varies, obtained from (i) normal data (first row) and (ii) data from multivariate *t*-distribution with 1 degree of freedom (second row), for $\mathcal{G}$ corresponding to spherical symmetry.

Summary

- Framework for distribution-free testing for multivariate symmetry
- Developed notions of generalized signs, ranks and signed-ranks
- Proposed generalizations of sign and Wilcoxon signed-rank tests

Summary

- Framework for distribution-free testing for multivariate symmetry
- Developed notions of generalized signs, ranks and signed-ranks
- Proposed generalizations of sign and Wilcoxon signed-rank tests
- Proposed tests are: (i) distribution-free and have good efficiency, (ii) computationally feasible, (iii) more powerful for distributions with heavy tails, and (iv) robust to outliers and contamination

Summary

- Framework for distribution-free testing for multivariate symmetry
- Developed notions of generalized signs, ranks and signed-ranks
- Proposed generalizations of sign and Wilcoxon signed-rank tests
- Proposed tests are: (i) distribution-free and have good efficiency, (ii) computationally feasible, (iii) more powerful for distributions with heavy tails, and (iv) robust to outliers and contamination
- Can develop universally consistent, distribution-free tests for multivariate symmetry using kernel methods (ongoing work)

Thank you very much!

Questions?

Question: How to generate

$$S_n(\mathbf{X}_i) \equiv S(\mathbf{X}_i, R_n(\mathbf{X}_i)) := \underset{Q \in \mathcal{G}}{\arg\min} \|Q^\top \mathbf{X}_i - R_n(\mathbf{X}_i)\|^2$$

when it is not unique?

Spherical symmetry $\mathcal{G} = \mathrm{O}(p)$

Let

$$S(\mathbf{x}, \mathbf{h}) := \underset{Q \in \mathcal{G}}{\arg\min} \|Q^\top \mathbf{x} - \mathbf{h}\|^2.$$

If $\mathbf{h}, \mathbf{x} \neq \mathbf{0}$, let $\mathbf{w} = \frac{\mathbf{h}}{\|\mathbf{h}\|}$, and $\mathbf{v} = \frac{\mathbf{x}}{\|\mathbf{x}\|}$. Then, $S(\mathbf{x}, \mathbf{h})$ should be chosen uniformly from:

$$\{Q \in \mathrm{O}(p) : \ \mathbf{v} = Q\mathbf{w}\} = \{\mathbf{v}\mathbf{w}^\top + VUW^\top : \ U \in \mathrm{O}(p-1)\},$$

where V and W are $p \times (p-1)$ matrices such that $V^\top V = W^\top W = I_{p-1}$, $V^\top \mathbf{v} = W^\top \mathbf{w} = 0$.