# Analytic Permutation Testing via Kahane–Khintchine Inequalities

Adam B Kashlak, Sergii Myroshnychenko
University of Alberta, Edmonton, AB
Susanna Spektor
The Sheridan College Institute of Technology and Advanced Learning,
Mississauga, ON

5 September 2020

# What is Khintchine's inequality?

**Theorem (Khintchine's Inequality, 1923)**

*Let $\varepsilon_1, \ldots, \varepsilon_{2n} \stackrel{iid}{\sim}$ Rademacher.*[1] *Then, for any $p \in [1, \infty)$, there exist finite universal constants $A_p$ and $B_p$ such that for any fixed $x_1, \ldots, x_{2n} \in \mathbb{R}$ (or $\in \mathbb{C}$),*

$$A_p \left( \sum_{i=1}^{2n} x_i^2 \right)^{1/2} \leq \left\{ \mathrm{E} \left| \sum_{i=1}^{2n} \varepsilon_i x_i \right|^p \right\}^{1/p} \leq B_p \left( \sum_{i=1}^{2n} x_i^2 \right)^{1/2} .$$

[1] That is, $\mathrm{P}(\varepsilon_i = 1) = \mathrm{P}(\varepsilon_i = -1) = 1/2$.

# What is Khintchine's inequality?

**Theorem (Khintchine's Inequality, 1923)**

*Let $\varepsilon_1, \ldots, \varepsilon_{2n} \overset{iid}{\sim}$ Rademacher.*[1] *Then, for any $p \in [1, \infty)$, there exist finite universal constants $A_p$ and $B_p$ such that for any fixed $x_1, \ldots, x_{2n} \in \mathbb{R}$ (or $\in \mathbb{C}$),*

$$A_p \left( \sum_{i=1}^{2n} x_i^2 \right)^{1/2} \leq \left\{ \mathrm{E} \left| \sum_{i=1}^{2n} \varepsilon_i x_i \right|^p \right\}^{1/p} \leq B_p \left( \sum_{i=1}^{2n} x_i^2 \right)^{1/2}.$$

This has a lot of applications in the theory of $L^p$ spaces such as the existence of Hilbert subspaces of $L^p[0, 1]$. Optimal constants are known (Haagerup, 1982).

[1] That is, $\mathrm{P}(\varepsilon_i = 1) = \mathrm{P}(\varepsilon_i = -1) = 1/2$.

# What is the restricted Khintchine inequality?

**Theorem (Restricted Khintchine's Inequality (Spektor, 2018))**

*Let $\varepsilon_1, \ldots, \varepsilon_{2n}$ be dependent Rademacher random variables such that $\sum \varepsilon_i = 0$. Then, for any $p \in [2, \infty)$, there exists a finite universal constant $B_p$ such that for any fixed $x_1, \ldots, x_{2n} \in \mathbb{R}$,*

$$\left\{ \mathrm{E}_\varepsilon \left| \sum_{i=1}^{2n} x_i \varepsilon_i \right|^p \right\}^{1/p} \le B_p(\|x\|_2^2 - 2n\bar{x}^2)^{1/2}.$$

[2] Note: this coincides with the $2p$-th moments of a centred Gaussian.

# What is the restricted Khintchine inequality?

**Theorem (Restricted Khintchine's Inequality (Spektor, 2018))**

*Let $\varepsilon_1, \ldots, \varepsilon_{2n}$ be dependent Rademacher random variables such that $\sum \varepsilon_i = 0$. Then, for any $p \in [2, \infty)$, there exists a finite universal constant $B_p$ such that for any fixed $x_1, \ldots, x_{2n} \in \mathbb{R}$,*

$$\left\{ \mathrm{E}_\varepsilon \left| \sum_{i=1}^{2n} x_i \varepsilon_i \right|^p \right\}^{1/p} \leq B_p (\|x\|_2^2 - 2n\bar{x}^2)^{1/2}.$$

**... but isn't that a permuted t statistic?**

For $s^2$ the sample variance, divide by $n$ to get

$$\left\{ \mathrm{E}_\pi |\bar{x}_1^{(\pi)} - \bar{x}_2^{(\pi)}|^p \right\}^{1/p} \leq B_p s$$

where $B_{2p} = [(2p)!/p!]^{1/p}$.[2]

[2]Note: this coincides with the $2p$-th moments of a centred Gaussian.

# Statistical Hypothesis Testing

## Two sample test

A prototypical problem in statistical hypothesis testing is to determine whether or not two populations have the same mean value.

## Example (Heights of boys and girls at age 10)

Given a data set of 39 boys and 54 girls, the mean heights are

$$\hat{\mu}_{\text{boy}} = 142.2 \text{ cm, and } \hat{\mu}_{\text{girl}} = 140.7 \text{ cm}$$

with a difference of 1.5 cm and a two sample t-test gives a non-significant p-value of 0.26 where
p-value $= \mathrm{P}\left(\text{an observed difference} \geq 1.5 \,|\, \mu_{\text{boy}} = \mu_{girl}\right)$

### Two sample test

A prototypical problem in statistical hypothesis testing is to determine whether or not two populations have the same mean value.

### Example (Heights of boys and girls at age 10)

Given a data set of 39 boys and 54 girls, the mean heights are

$$\hat{\mu}_{\text{boy}} = 142.2 \text{ cm, and } \hat{\mu}_{\text{girl}} = 140.7 \text{ cm}$$

with a difference of 1.5 cm and a two sample t-test gives a non-significant p-value of 0.26 where
p-value $= \mathrm{P}\left(\text{an observed difference} \geq 1.5 \,|\, \mu_{\text{boy}} = \mu_{girl}\right)$

**Goal:** To use Kahane-Khintchine to construct bounds on such tail probabilities.

# Statistical Hypothesis Testing

Two statistical paradigms:

1. **Parametric Statistics**
   - Assume data has a specific distribution (e.g. normal)
   - Working with a finite number of parameters (e.g. mean and variance).
   - Easy to compute but strong assumptions.

# Statistical Hypothesis Testing

Two statistical paradigms:

1. **Parametric Statistics**
   - Assume data has a specific distribution (e.g. normal)
   - Working with a finite number of parameters (e.g. mean and variance).
   - Easy to compute but strong assumptions.
2. **Non-parametric Statistics**
   - Only assume general conditions on the data (e.g. exchangeable observations or permutationally invariant)
   - Working in an infinite dimensional space (e.g. log-concave distributions).
   - Hard to compute but weak assumptions.

# Statistical Hypothesis Testing

Two statistical paradigms:

1. **Parametric Statistics**
   - Assume data has a specific distribution (e.g. normal)
   - Working with a finite number of parameters (e.g. mean and variance).
   - Easy to compute but strong assumptions.
2. **Non-parametric Statistics**
   - Only assume general conditions on the data (e.g. exchangeable observations or permutationally invariant)
   - Working in an infinite dimensional space (e.g. log-concave distributions).
   - Hard to compute but weak assumptions.

A *permutation test* is a type of non-parametric statistical test.

# What is a permutation test?

## Two Sample: setup

Given $X_1, \ldots, X_n, X_{n+1}, \ldots, X_{2n}$ with $\mathrm{E}X_i = \mu_1$ for $i \leq n$ and $\mathrm{E}X_i = \mu_2$ for $i > n+1$,

$$H_0 : \mu_1 = \mu_2 \quad \text{vs} \quad H_1 : \mu_1 \neq \mu_2.$$

# What is a permutation test?

## Two Sample: setup

Given $X_1, \ldots, X_n, X_{n+1}, \ldots, X_{2n}$ with $\mathrm{E}X_i = \mu_1$ for $i \le n$ and $\mathrm{E}X_i = \mu_2$ for $i > n+1$,

$$H_0 : \mu_1 = \mu_2 \quad \text{vs} \quad H_1 : \mu_1 \neq \mu_2.$$

## Parametric Two Sample Test: t distribution

Compute $T_0 = \sqrt{n}(\bar{X}_1 - \bar{X}_2)/s$ for sample variance $s^2$. Then, for $T \sim t\,(2n-2)$,

$$\text{p-value} = \mathrm{P}\left(|T| > |T_0|\right).$$

# What is a permutation test?

## Two Sample: setup

Given $X_1, \ldots, X_n, X_{n+1}, \ldots, X_{2n}$ with $\mathrm{E}X_i = \mu_1$ for $i \leq n$ and $\mathrm{E}X_i = \mu_2$ for $i > n+1$,

$$H_0 : \mu_1 = \mu_2 \quad \text{vs} \quad H_1 : \mu_1 \neq \mu_2.$$

## Nonparametric Two Sample Test: permutation

For a uniformly random permutation $\pi$ on $2n$ elements, consider

$$T(\pi) = n^{-1}\left(\sum_{i=1}^{n} X_{\pi(i)} - \sum_{i=n+1}^{2n} X_{\pi(i)}\right)$$

and p-value $= \mathrm{P}\left(|T(\pi)| > |T_0| \,|\, X_1, \ldots, X_{2n}\right)$ for $T_0 = \bar{X}_1 - \bar{X}_2$.

# What is a permutation test?

The Permutation Test:

- The distribution of $T(\pi)$ is discrete supported on the symmetric group $\mathbb{S}_{2n}$.
- Generally, the p-value is estimated via Monte Carlo simulation as $(2n)!$ is large.
- It is an *exact test*—i.e. achieves correct significance level for any $n$.
- Main assumption is *exchangeability*—i.e. distribution of $T$ is invariant under permutations under $H_0$.
- Often applied when the asymptotic behaviour of the test statistic is not known.

### Research Question

1. What is the distribution of $T : \mathbb{S}_{2n} \to \mathbb{R}$ for $\pi$ uniformly distributed on $\mathbb{S}_{2n}$?
2. What is a good upper bound for $\mathrm{P}\left(T(\pi) > T_0\right)$?

### Research Question

1. What is the distribution of $T : \mathbb{S}_{2n} \to \mathbb{R}$ for $\pi$ uniformly distributed on $\mathbb{S}_{2n}$?
2. What is a good upper bound for $\mathrm{P}\left(T(\pi) > T_0\right)$?

### The Permutationless Permutation Test

I want to run a permutation test without ever generating a single permutation!

### Theorem (K, Spektor, Myroshnychenko)

*For $X_1, \ldots, X_n$ with $X_i$, $i \le m_1$, from sample 1 and $X_i$, $i > m_1$, from sample 2 with $s$ the sample standard deviation of $X_1, \ldots, X_n$ and test statistic*

$$T(\pi) = \frac{1}{s}\left[\frac{1}{m_1}\sum_{i=1}^{m_1} X_{\pi(i)} - \frac{1}{m_2}\sum_{i=m_1+1}^{n} X_{\pi(i)}\right]$$

*with $m_1 = \kappa m_2$ for some $\kappa \ge 1$, then*

$$\mathrm{P}\left(T(\pi) \ge t\right) \le \exp\left(-nt^2/2\lceil\kappa+1\rceil^3\right).$$

# Theorem 2.1, Univariate Data

**Proof Sketch:**

1. First prove a weighted restricted variant of Khintchine's inequality.

## Theorem (K, Spektor, Myroshnychenko)

*For $m_1 > m_2 > 0$, let $n = m_1 + m_2$ and $M = m_1 - m_2$ and let $\kappa m_2 = m_1$ for some rational $\kappa > 1$. Let $\delta_1, \ldots, \delta_n$ be weighted dependent Rademacher random variables such that marginally $\mathrm{P}(\delta_i = 1/m_1) = \mathrm{P}(\delta_i = -1/m_2) = 1/2$ and such that $\sum \delta_i = 0$—i.e. precisely $m_1$ of the $\delta_i$ equal $1/m_1$ and $m_2$ equal $-1/m_2$. For any $p \in [2, \infty)$, there exists a positive finite constant $B_p$ such that for any sequence $x_1, \ldots, x_n \in \mathbb{R}$,*[3]

$$\mathrm{E}|\textstyle\sum_{i=1}^n \delta_i x_i|^p \le B_p \left( \lceil \kappa + 1 \rceil^2 s_n^2 / 2m_2 \right)^{p/2}$$

$s_n^2 = (n-1)^{-1} \sum_{i=1}^n (x_i - \bar{x})^2$ *is the sample variance of* $x$.

[3] This theorem is also valid for $x_i \in \mathbb{C}$ after standard alterations are made in the proof.

**Proof Sketch:**

2. Use Markov's / Chernoff's inequality to translate moment bounds into tail probability bounds.

## Lemma

*For a centred symmetric univariate random variable* $Z \in \mathbb{R}$ *such that* $\mathrm{E}|Z|^{2p} \le (2p)!C^p/p!$ *for some constant* $C > 0$*. Then,*

$$\mathrm{P}\left(Z > t\right) \le \mathrm{e}^{-t^2/4C}.$$

## Proof.

The moment generating function is bounded as follows:

$$\mathrm{E}\mathrm{e}^{\lambda Z} = \sum_{p=0}^{\infty} \frac{\lambda^p \mathrm{E}Z^p}{p!} = \sum_{p=0}^{\infty} \frac{\lambda^{2p} \mathrm{E}Z^{2p}}{(2p)!} \le \sum_{p=0}^{\infty} \frac{\lambda^{2p} C^p}{p!} \le \mathrm{e}^{\lambda^2 C}.$$

The result follows as $\mathrm{P}\left(Z > t\right) \le \mathrm{e}^{-\lambda t}\mathrm{E}\mathrm{e}^{\lambda Z}$. □

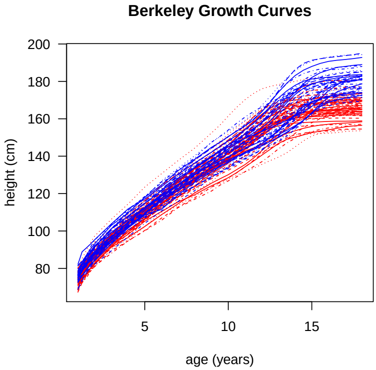
# A More General Setting

**Problem:** Univariate data (i.e. $X \in \mathbb{R}$) is boring! Statisticians care about more complex data (e.g. $X \in L^p$).

# A More General Setting

**Problem:** Univariate data (i.e. $X \in \mathbb{R}$) is boring! Statisticians care about more complex data (e.g. $X \in L^p$).

Instead of comparing the heights of girls and boys at age 10, compare the entire growth curves or compare the covariance operators.



# What is the Kahane-Khintchine Inequality?

**Theorem (Kahane-Khintchine Inequality, 1964)**

*Let $(\mathcal{X}, \|\cdot\|))$ be a Banach space. For any $p, p' \in [1, \infty)$, there exists a universal finite constant $C_{p,p'} > 0$ such that for any sequence of $X_1, \ldots, X_n \in \mathcal{X}$*

$$\left\{ \mathrm{E} \left\| \sum_{i=1}^{n} \varepsilon_i X_i \right\|^{p} \right\}^{1/p} \le C_{p,p'} \left\{ \mathrm{E} \left\| \sum_{i=1}^{n} \varepsilon_i X_i \right\|^{p'} \right\}^{1/p'}$$

*where $\varepsilon_i$ are iid Rademacher random variables.*

# What is the Kahane-Khintchine Inequality?

**Theorem (Kahane-Khintchine Inequality, 1964)**

*Let $(\mathcal{X}, \|\cdot\|)$ be a Banach space. For any $p, p' \in [1, \infty)$, there exists a universal finite constant $C_{p,p'} > 0$ such that for any sequence of $X_1, \ldots, X_n \in \mathcal{X}$*

$$\left\{ \mathrm{E} \left\| \sum_{i=1}^{n} \varepsilon_i X_i \right\|^{p} \right\}^{1/p} \leq C_{p,p'} \left\{ \mathrm{E} \left\| \sum_{i=1}^{n} \varepsilon_i X_i \right\|^{p'} \right\}^{1/p'}$$

*where $\varepsilon_i$ are iid Rademacher random variables.*

Optimal constants are not currently known for this inequality in our setting of interest. Though, optimal constants are known in other cases (Latała & Oleszkiewicz, 1994).

# Theorem 2.2, Commutative $L^q$ Spaces

### Theorem (K, Spektor, Myroshnychenko)

*Let $m = n/2$, $\|\cdot\|_{S^q}$ be the q-Schatten norm for matrices or operators, and $\varepsilon_1, \ldots, \varepsilon_n$ be Rademacher random variables such that $\sum_{i=1}^{n} \varepsilon_i = 0$. Let $X_1(t), \ldots, X_n(t)$ be continuous functions on a compact interval with empirical covariance operator $\hat{\Sigma}(s,t) = (n-1)^{-1} \sum_{i=1}^{n} X_i(s) X_i(t)$. Let $q \in [1, \infty)$ with norm $\|\cdot\|_{L^q}$. For $T(\pi) = \|\sum_{i=1}^{n} \varepsilon_i X_i\|_{L^q}$, there exists a universal constant $c > 0$ such that*

$$\mathrm{P}\left(T(\pi) \geq t\right) \leq \exp\left(-t^2 / c \|\hat{\Sigma}^{1/2}\|_{S^q}^2\right).$$

# Theorem 2.3, Non-commutative $L^q$ Spaces

## Theorem (K, Spektor, Myroshnychenko)

*Let* $\|\cdot\|_{S^q}$ *be the q-Schatten norm for a matrix or operator and* $\varepsilon_1, \ldots, \varepsilon_n$ *be Rademacher random variables such that* $\sum_{i=1}^n \varepsilon_i = 0$. *For* $d, d' > 1$, *let* $X_1, \ldots, X_n \in \mathbb{R}^{d \times d'}$ *be a collection of n fixed matrices (or let* $X_1, \ldots, X_n$ *be a collection of bounded linear operators). For* $T(\pi) = \|\sum_{i=1}^n \varepsilon_i X_i\|_{S^q}$, *there exists a universal constant* $c > 0$ *such that*

$$\mathrm{P}\left(T(\pi) > t\right) \leq \exp\left(-t^2/c\mathcal{S}^2\right)$$

*where* $\mathcal{S} = \max\left\{\|(\sum_{i=1}^n X_i X_i^*)^{1/2}\|_{S^q}, \|(\sum_{i=1}^n X_i^* X_i)^{1/2}\|_{S^q}\right\}$ *with* $X_i^*$ *the adjoint operator.*

**Proof Sketch:**

1. Use similar tricks for inserting the dependency condition into Kahane's inequality in the commutative and non-commutative settings.
2. Use Markov / Chernoff and the bound on the moment generating function to bound the tail probability.

**Proof Sketch:**

1. Use similar tricks for inserting the dependency condition into Kahane's inequality in the commutative and non-commutative settings.
2. Use Markov / Chernoff and the bound on the moment generating function to bound the tail probability.

And we're done! We have a general method for two sample testing within general Banach spaces except...

# Theorems 2.2 & 2.3

**Proof Sketch:**

1. Use similar tricks for inserting the dependency condition into Kahane's inequality in the commutative and non-commutative settings.
2. Use Markov / Chernoff and the bound on the moment generating function to bound the tail probability.

And we're done! We have a general method for two sample testing within general Banach spaces except...
The stats journals were unhappy that these bounds are based on unknown constants.

# The Beta Transformation in $\mathbb{R}$

In statistics, we want a p-value to be a $\text{Uniform}[0, 1]$ random variable under the null hypothesis, but our bound on the p-value follows a more general beta distribution.

# The Beta Transformation in $\mathbb{R}$

In statistics, we want a p-value to be a $\mathrm{Uniform}\,[0,1]$ random variable under the null hypothesis, but our bound on the p-value follows a more general beta distribution.

## Proposition (K, Spektor, Myroshnychenko)

*Under the setting of Theorem 2.1 (univariate data) with n sufficiently large,*

$$\mathrm{P}\left(\exp\left\{-nT(\pi)^2/2\lceil\kappa+1\rceil^3\right\}<u\right)\le C_0 I\left(u;\frac{\lceil\kappa+1\rceil^3}{(2+\kappa+\kappa^{-1})},\frac{1}{2}\right)$$

*where $I(u;\alpha,\beta)$ is the regularized incomplete beta function and*

$$C_0=\frac{\left(\frac{\lceil\kappa+1\rceil^3}{2+\kappa+\kappa^{-1}}\right)^{1/2}\Gamma\left(\frac{\lceil\kappa+1\rceil^3}{2+\kappa+\kappa^{-1}}\right)}{\Gamma\left(\frac{1}{2}+\frac{\lceil\kappa+1\rceil^3}{2+\kappa+\kappa^{-1}}\right)}.$$

# Beta Transformation, Proof Sketch

1. Note that $nT_0^2/(2+\kappa+\kappa^{-1})$ is approximately $\chi^2(1)$ via the CLT.
2. For $Z \sim \chi^2(1)$, bound

$$\begin{aligned}
\mathrm{P}\left(\mathrm{e}^{-Z/c} \leq u\right) &= \mathrm{P}\left(Z \geq -c \log u\right) \\
&= (2\pi)^{-1/2} \int_{-c\log u}^{\infty} x^{-1/2}\mathrm{e}^{-x/2} dx \\
&= \left(\frac{c}{2\pi}\right)^{1/2} \int_0^u (-\log y)^{-1/2} y^{c/2-1} dy \\
&\leq \left(\frac{c}{2\pi}\right)^{1/2} \int_0^u (1-y)^{1/2-1} y^{c/2-1} dy \\
&= \frac{(c/2)^{1/2}\Gamma(c/2)}{\Gamma((c+1)/2)} I(u; c/2, 1/2)
\end{aligned}$$

**Theorem (K, Spektor, Myroshnychenko)**

*Let $(\mathcal{X}, \|\cdot\|)$ be a Banach space with separable dual space $\mathcal{X}^*$, and let $h : \mathbb{R} \to \mathbb{R}$ be monotonically increasing. For any random variable $X$ taking values in $\mathcal{X}$ such that $\mathrm{E}h(\|X\|)^2 < \infty$ and $\sup_{X\in\mathcal{X}} h(\|X\|) < U < \infty$ and for $u \in (0,1)$ and some constants $C, c, \alpha, \beta > 0$,*

$$\mathrm{P}\left(\mathrm{e}^{-h(\|X\|)/c} < u\right) \le CI(u; \alpha, \beta)$$

*where $I(u; \alpha, \beta)$ is the incomplete beta function for $c$ sufficiently large.*

# Beta Transform, Proof Sketch

1. For $Z = h(\|X\|) \in \mathbb{R}^+$, rewrite as a countable supremum $Z = \sup_{\phi \in B^*} h(\phi(X))$ and apply Talagrand's concentration inequality: $\mathrm{P}\left(Z \geq \mathrm{E}Z + t\right) \leq \exp\left(\frac{-t^2}{a+bt}\right)$ for positive constants $a$ and $b$ depending on $\mathrm{E}h(\|X\|)^2$ and $\sup_{X \in \mathcal{X}} h(\|X\|)$.
2. Use a bunch of inequalities and calculus to get that

$$\exp\left(\frac{-t^2}{a+bt}\right) \leq \mathrm{e}^{-t/b}\left[\left(1+\frac{bt}{a}\right)^a\right]^{1/b^2}$$
$$\leq \mathrm{e}^{-t/b}\sum_{k=0}^{a}\frac{1}{k!}\left(\frac{t}{b}\right)^k = \int_t^\infty \frac{x^{a-1}\mathrm{e}^{-x/b}}{b^a\Gamma(a)}dx$$

3. Then proceed similarly to the univariate case to get

$$\mathrm{P}\left(\mathrm{e}^{-(Z-\mathrm{E}Z)/c} \leq u\right) \leq \frac{c^a}{b^a\Gamma(a)}\int_0^u (1-y)^{a-1}y^{c/b-a-1}dy$$
$$= CI(u; a, c/b-a)$$

# The Empirical Beta Transform

In practice, we do not know the universal constants emerging from Kahane's inequality and Talagrand's inequality, so instead we perform a small simulation to estimate the parameters of the beta distribution $\alpha$ and $\beta$.

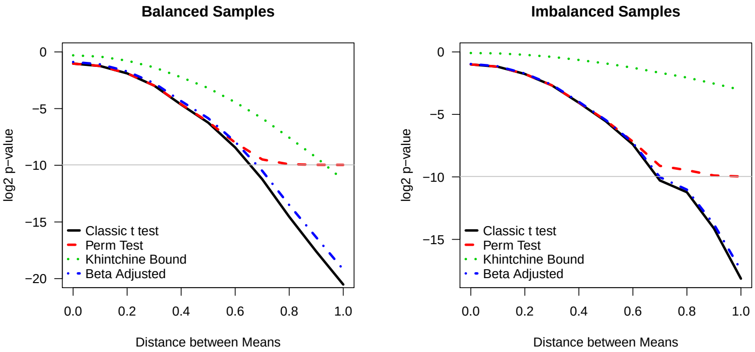
In practice, we do not know the universal constants emerging from Kahane's inequality and Talagrand's inequality, so instead we perform a small simulation to estimate the parameters of the beta distribution $\alpha$ and $\beta$.

**Empirical Beta Transform**

Compute p-value $p_0$ from test statistic $T_0$.
Choose $r > 1$, the number of permutations to simulate—e.g. $r = 10$.
Draw $\pi_1, \ldots, \pi_r$ from $\mathbb{S}_n$ uniformly at random.
Compute p-values $p_1, \ldots, p_r$ from test statistics $T_{\pi_1}, \ldots, T_{\pi_r}$.
Find the method of moments estimator for $\alpha$ and $\beta$.
  Estimate first and second central moments of the $p_i$ by $\bar{p}$ and $s^2$.
  Estimate $\hat{\alpha} = \bar{p}^2(1 - \bar{p})/s^2 - \bar{p}$.
  Estimate $\hat{\beta} = [\bar{p}(1 - \bar{p})/s^2 - 1][1 - \bar{p}]$.
Return the adjusted p-value $I(p_0; \hat{\alpha}, \hat{\beta})$.

# But does it work on simulated data?

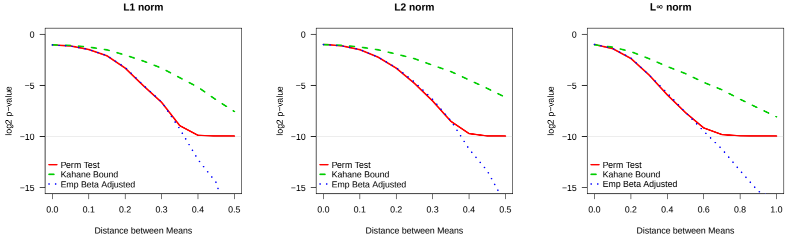
## Simulated Univariate Gaussian Data




On the left, $m_1 = m_2 = 100$. On the right, $m_1 = 140, m_2 = 60$.

# But does it work on simulated data?

## Simulated Multivariate Gaussian Data




For data in $\mathbb{R}^{12}$ with sample sizes $m_1 = m_2 = 50$ for $\ell^1$, $\ell^2$, and $\ell^\infty$ norms.

# Extending to k sample testing

Under the one-way ANOVA model, $X_{i,j} = \mu + \tau_i + \varepsilon_{i,j}$ we can test

| | | |
|---|---|---|
| Pairwise | $H_0^{(ij)} : \tau_i = \tau_j$ | $H_1^{(ij)} : \tau_i \neq \tau_j$ |
| Global | $H_0 : \tau_1 = \ldots = \tau_k = 0$ | $H_1 : \exists \tau_i \neq 0.$ |

by applying the same permutation to each pairwise comparison.

# Extending to k sample testing

Under the one-way ANOVA model, $X_{i,j} = \mu + \tau_i + \varepsilon_{i,j}$ we can test

| | | |
|---|---|---|
| Pairwise | $H_0^{(ij)} : \tau_i = \tau_j$ | $H_1^{(ij)} : \tau_i \neq \tau_j$ |
| Global | $H_0 : \tau_1 = \ldots = \tau_k = 0$ | $H_1 : \exists \tau_i \neq 0.$ |

by applying the same permutation to each pairwise comparison.

## Extension to Banach Spaces

Note that the global test statistic can be considered as an $\ell^2$ direct sum of $\kappa = \binom{k}{2}$ Banach spaces. For a sequence $(\mathcal{X}_i, \|\cdot\|_i)$ and elements $X_i \in \mathcal{X}_i$,

$$(X_i)_{i=1}^n \in \left( \bigoplus_{i=1}^{\kappa} \mathcal{X}_i \right)_{\ell^2}$$

with norm $\|(X_i)_{i=1}^n\| = (\sum_{i=1}^n \|X_i\|_i^2)^{1/2}$.

# Synchronized Test: Analytic Variant

Let $X^{(i)}$ be the vector of observations from population $i$ and let

$$X = \begin{pmatrix} X^{(1)} & X^{(1)} & \dots & X^{(k-1)} \\ X^{(2)} & X^{(3)} & \dots & X^{(k)} \end{pmatrix}.$$

be the $(2m) \times \binom{k}{2}$ matrix of pairwise comparisons. Then, for $\varepsilon$ multivariate Rademacher such that $\sum \varepsilon_j = 0$, we have the test statistic $T = \|X^{\mathrm{T}}\varepsilon\|_{\ell^2}$

# Synchronized Test: Analytic Variant

Let $X^{(i)}$ be the vector of observations from population $i$ and let

$$X = \begin{pmatrix} X^{(1)} & X^{(1)} & \dots & X^{(k-1)} \\ X^{(2)} & X^{(3)} & \dots & X^{(k)} \end{pmatrix}.$$

be the $(2m) \times \binom{k}{2}$ matrix of pairwise comparisons. Then, for $\varepsilon$ multivariate Rademacher such that $\sum \varepsilon_j = 0$, we have the test statistic $T = \|X^{\mathrm{T}}\varepsilon\|_{\ell^2}$

The squared test statistic $T^2 = \sum_{i,j=1}^{2m} a_{i,j}\varepsilon_i\varepsilon_j$ is a degree 2 homogeneous Rademacher chaos with $a_{i,j}$ the $ij$th entry in $XX^{\mathrm{T}}$.

# Theorem 3.1, k-sample testing

**Theorem (K, Spektor, Myroshnychenko)**

*Let* $T = \|X^{\mathrm{T}}\varepsilon\|_{\ell^2}$ *for* $X$ *the above* $2m \times \binom{k}{2}$ *matrix and* $\varepsilon_i$ *such that* $\sum_{i=1}^{2m} \varepsilon_i = 0$*. Then, for some universal constant* $c$*,*

$$\mathrm{P}\left(T > t\right) \leq \exp\left[-t^2/c\mathcal{S}\right]$$

*where* $\mathcal{S} = \|XX^{\mathrm{T}}\|_{S^2}$*.*

**Proof Sketch**

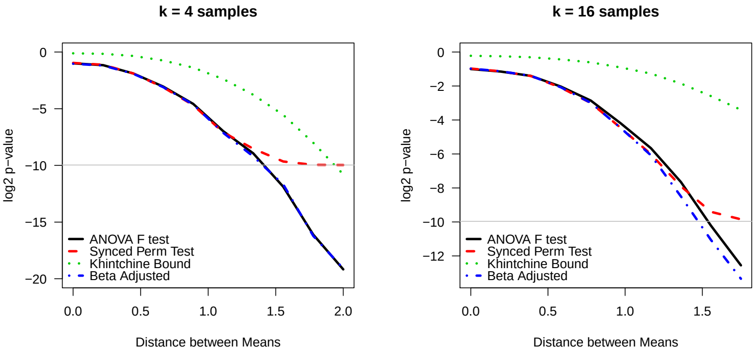
1. Extend the ideas in Spektor (2018) from Rademacher sums to homogeneous Rademacher chaoses of degree 2.
2. Apply a decoupling inequality from Kwapien (1987)

$$\mathrm{E}_\varepsilon|T^2|^p = \left[\mathrm{E}_\pi \mathrm{E}_\delta |\textstyle\sum_{i,j=1}^m \delta_i \delta_j b_{i,j,\pi}|^p\right]^{1/p}$$
$$\le B_p^2 C \left[\mathrm{E}_\pi \mathrm{E}_\delta |\textstyle\sum_{i,j=1}^m \delta_i \delta_j b_{i,j,\pi}|^2\right]^{1/2} \le 2B_p^2 C \|XX^{\mathrm{T}}\|_{S^2}$$

3. Then use Markov/Chernoff again with a little help from the Legendre duplication formula and Watson's formula to get the correct coefficients.

# And it also works on simulated data!

**Simulated k samples of univariate Gaussian data**



# Really Cool Data Application 1: Speech Sounds

Testing differences between

- 12 spoken vowel sounds (phonemes)
- Taken from recordings of 4 speakers while controlling for gender (male/female) and country of origin (Canada/China).
- Using experimental designs—i.e. the Latin square design and the randomized block design.

# Really Cool Data Application 1: Speech Sounds

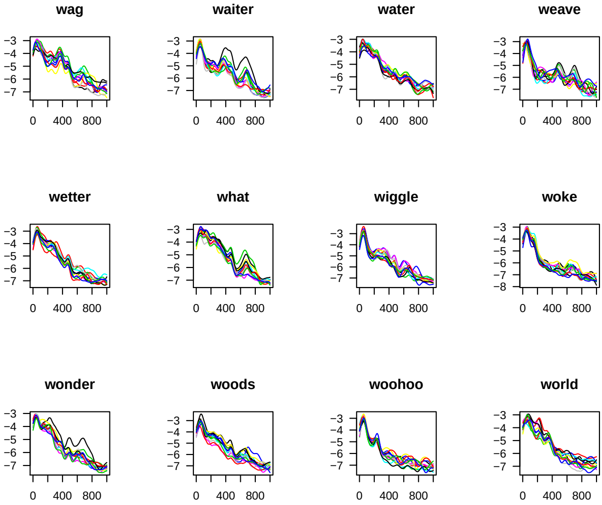
Testing differences between

- 12 spoken vowel sounds (phonemes)
- Taken from recordings of 4 speakers while controlling for gender (male/female) and country of origin (Canada/China).
- Using experimental designs—i.e. the Latin square design and the randomized block design.

The only other way to analyze this data would be to run a large scale simulation based permutation test requiring an estimated 35 million SVD calculations with estimated runtime of 74 days on my Intel Core i7-7567U CPU.

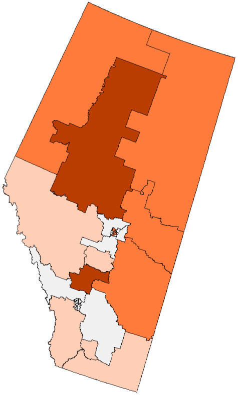
# Really Cool Data Application 1: Speech Sounds

## Log periodograms of one of the speakers



**Alberta 2019 Electoral Results:**

- Testing for local autocorrelation at each riding in Alberta—i.e. testing if election results correlated or uncorrelated with the neighbouring results.
- Test statistics take the form a measure of matrix association like $\gamma_{AB} := \sum_{i,j=1}^{n} a_{i,j} b_{i,j}$.
- Considering the map as a graph $\mathcal{G}$ with a measurement $y_i$ taken at each vertex, the permutation test permutes the location of the vertices.



# Summary of Results

- Adapted the restricted Khintchine inequality to bounding permutation p-values for the unpaired t-test.
- Extended this result for imbalanced samples
- Extended this via the Kahane-Khintchine inequality to
  - Commutative Banach Spaces: Multivariate, Functional Data
  - Non-Commutative Banach Spaces: Matrices and Operators
- Developed a k-sample test based on the synchronized permutation test.
- Extended this to more advanced experimental designs.
- Able to apply such tools to a diverse collection of data sets.

# Summary of Results

- Adapted the restricted Khintchine inequality to bounding permutation p-values for the unpaired t-test.
- Extended this result for imbalanced samples
- Extended this via the Kahane-Khintchine inequality to
  - Commutative Banach Spaces: Multivariate, Functional Data
  - Non-Commutative Banach Spaces: Matrices and Operators
- Developed a k-sample test based on the synchronized permutation test.
- Extended this to more advanced experimental designs.
- Able to apply such tools to a diverse collection of data sets.

Thank you for listening