Permutation tests for the equality of covariance operators of functional data

with applications to evolutionary biology

Alessandra Cabassi

Joint work with
Dr Davide Pigoli, King's College London
Prof Piercesare Secchi, Politecnico di Milano

11 June 2018

MRC Biostatistics Unit

UNIVERSITY OF CAMBRIDGE

Outline

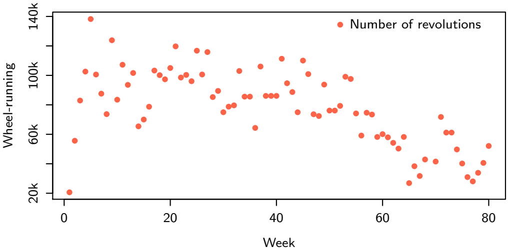
Evolutionary biology data set

Swallow et al. (1998)

Goal: Estimate the heritability of voluntary wheel-running behaviour in mice

Artificial selection of house mice for increased voluntary wheel running

- 10th generation
- 4 selected lines, 4 control lines, 20 mice per line

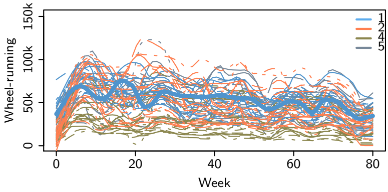


Mouse #90138.

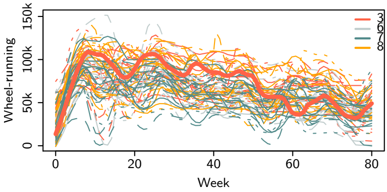
Functional data analysis

Functional variable	$X : \Omega \to L^2(I),\ I \subseteq \mathbb{R}$
Mean function	$\mu(t) = \mathbb{E}[X(t)],\ t \in I$
Covariance operator	$\Sigma(y) = \mathbb{E}[\langle X - \mathbb{E}(X), y\rangle (X - \mathbb{E}[X])],\ y \in L^2(I)$
Hilbert-Schmidt kernel of Σ	$\sigma(s,t) = \mathbb{E}[(X(t) - \mu(t)) \cdot (X(s) - \mu(s))],\ s, t \in I$

Evolutionary biology data set



(a) Control lines



(b) Selected lines

Number of revolutions run per week.
Smoothed and aligned data.

Objective

Consider q groups of samples of random curves, realisations of random processes with mean μ_i and covariance operator Σ_i

$$x_{i,1}, x_{i,2} \ldots, x_{i,n_i} \in L^2(I), \quad i = 1, \ldots, q.$$

We wish to test the hypothesis

$$H_0 : \{\Sigma_1 = \ldots = \Sigma_q\} \text{ against } H_1 : \{\text{at least one equality is not true}\}$$

and, if H_0 is rejected, perform pairwise comparisons between the groups.

Two-sample permutation test

Pigoli et al. (2014)

If $q = 2$, we can reformulate the test as follows

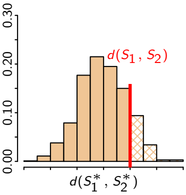
$$H_0 : d(\Sigma_1, \Sigma_2) = 0 \text{ against } H_1 : d(\Sigma_1, \Sigma_2) > 0$$

where $d(\cdot, \cdot)$ is some distance between two covariance operators.

Algorithm

1. If $\mu_1 \neq \mu_2$, let $\tilde{x}_{ij} = x_{ij} - m_i$, with m_i sample mean of group i
2. Compute $d(S_1, S_2)$, with S_i sample covariance operator of group i
3. Apply B random permutations to the labels of the sample curves
4. For each of them compute $d(S_1^*, S_2^*)$
5. The p-value of the test is

$$\lambda = \frac{\sum \mathbb{1}\left[d(S_1^*, S_2^*) \geq d(S_1, S_2)\right]}{B}$$



Distances between covariance operators

Pigoli et al. (2014)

- **Kernel distance**

$$d_L(\Sigma_1, \Sigma_2) = \|\sigma_1 - \sigma_2\|_{L^2(I\times I)} = \sqrt{\int_I \int_I (\sigma_1(s,t) - \sigma_2(s,t))^2 dsdt}$$

- **Square root matrix distance**

$$\begin{aligned} d_R(\Sigma_1, \Sigma_2) &= \|(\Sigma_1)^{1/2} - (\Sigma_2)^{1/2}\|_{HS} \\ &= \text{trace}[(\Sigma_1^{1/2} - \Sigma_2^{1/2})'(\Sigma_1^{1/2} - \Sigma_2^{1/2})] \end{aligned}$$

- **Procrustes size-and-shapes distance**

$$\begin{aligned} d_P(\Sigma_1, \Sigma_2)^2 &= \inf_{R\in\mathcal{O}(L^2)} \|L_1 - L_2 R\|_{HS}^2 \\ &= \|L_1\|_{HS}^2 + \|L_2\|_{HS}^2 - 2\sum_{n=1}^{+\infty} \rho_n \end{aligned}$$

L_i: $\Sigma_i = L_i L_i'$ for $i = 1, 2$
$\mathcal{O}(L^2)$: Space of unitary operators on $L^2(I)$
ρ_n: Eigenvalues of $L_2' L_1$

Non-parametric combination methodology

Global null hypothesis: **intersection** of partial null hypotheses

$$H_0 : \bigcap_{i \neq j} H_0^{ij}, \quad H_0^{ij} : \{\Sigma_i = \Sigma_j\}$$

Global alternative hypothesis: **union** of partial alternative hypotheses

$$H_1 : \bigcup_{i \neq j} H_1^{ij}, \quad H_1^{ij} : \{\Sigma_i \neq \Sigma_j\}$$

Idea: Combine all the k = q(q-1)/2 pairwise comparisons in a global test via the non-parametric combination methodology.

We define

$T_{ij} = T_{ij}(X)$ partial test statistics for testing H_0^{ij} against H_1^{ij}

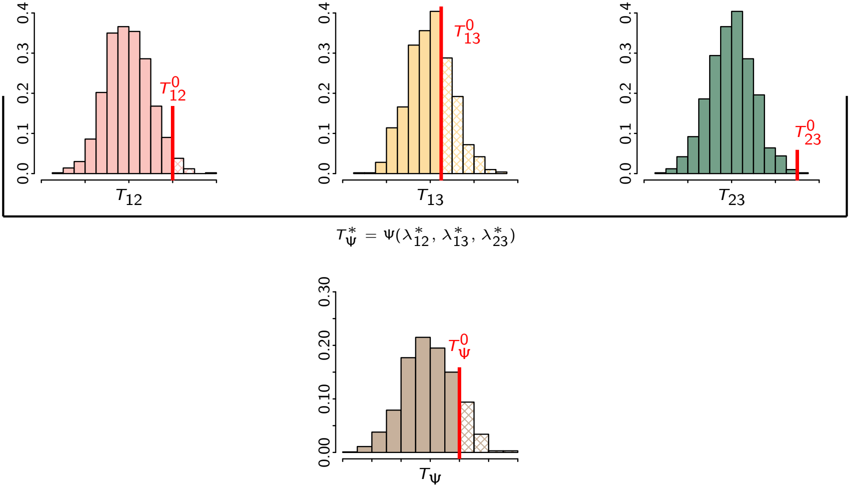
$\mathbf{T} = \mathbf{T}(X)$ k-dimensional vector of test statistics

$\Psi : \mathbb{R}^k \to \mathbb{R}^1$ combining function

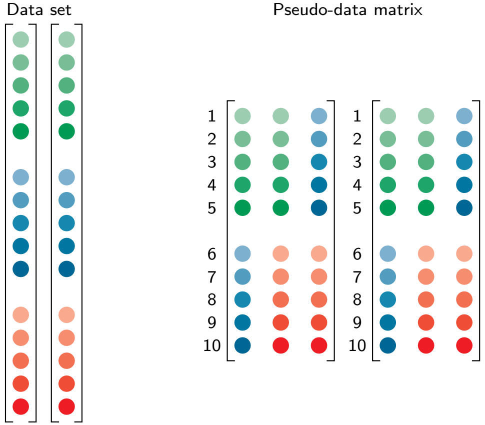
Non-parametric combination methodology

Pesarin and Salmaso (2010)

The NPC accounts for dependencies among partial statistics by obtaining their joint permutation null distribution.

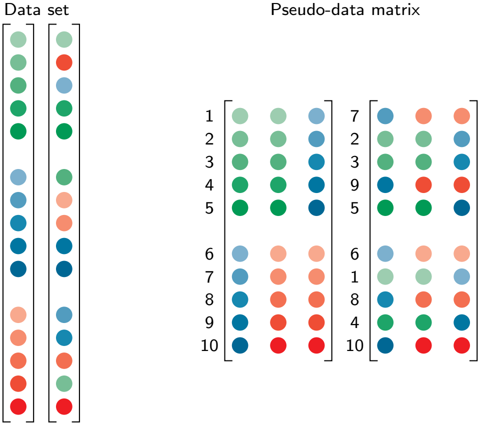


Solari et al. (2009)



Synchronized permutations allow both maintaining the dependencies among partial tests and involving the observations of each comparison at the same time.

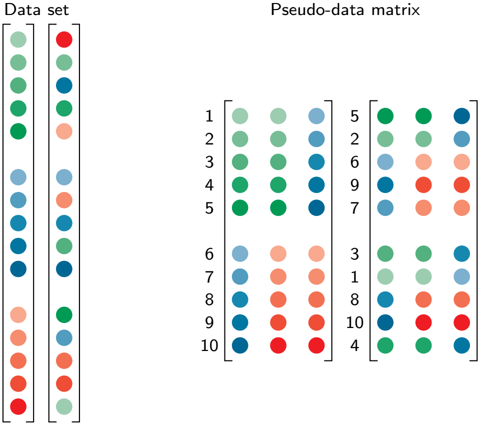
Solari et al. (2009)



Synchronized permutations allow both maintaining the dependencies among partial tests and involving the observations of each comparison at the same time.

Synchronised permutations

Solari et al. (2009)



Synchronized permutations allow both maintaining the dependencies among partial tests and involving the observations of each comparison at the same time.

Some admissible combining functions

Pesarin and Salmaso (2010)

- Fisher's

$$T_F = -2 \sum_i \log(\lambda_i)$$

- Liptack's

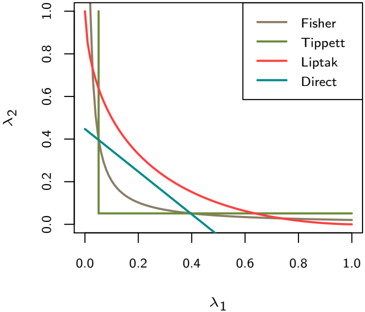
$$T_L = \sum_i \phi^{-1}(1 - \lambda_i)$$

- Tippett's

$$T_T = \min_{1 \le i \le k} \lambda_i$$

- Direct combination

$$T_D = \sum_i \lambda_i$$



Post-hoc comparisons

Closed testing procedure, Marcus et al. (1976)

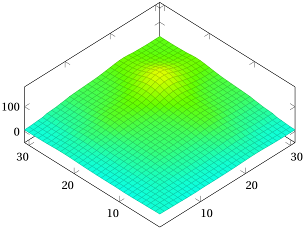
- Can be used with any combining function
- High number of steps
- Very conservative

Step-down procedure, Westfall and Young (1993)

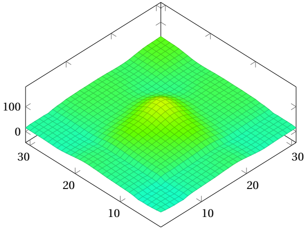
- Only for Tippett combining function
- Iterative procedure, only k steps
- Less conservative

Ramsay and Silverman (2005), Berkeley growth study data set

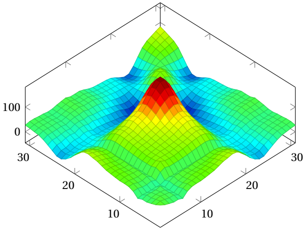
$N = 20, p = 31$



$\gamma = 0$

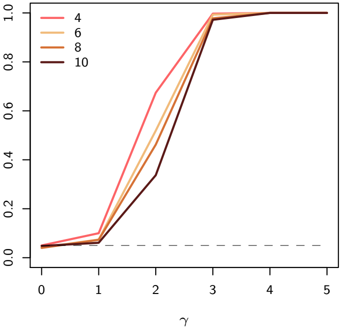


$\gamma = 2$

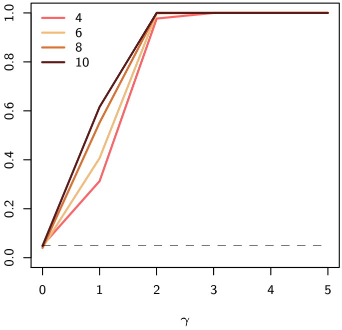


$\gamma = 5$

Empirical power of the test



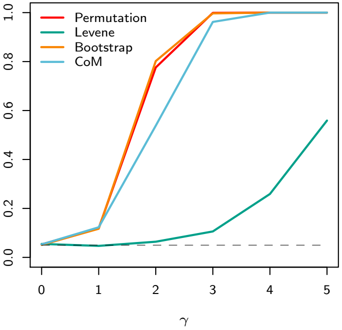
(a) One of the groups has covariance operator equal to Σ_1, the others $\Sigma(\gamma)$.



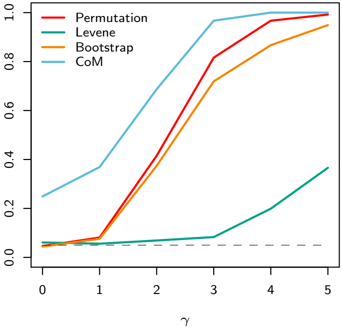
(b) Half of the groups have covariance equal to Σ_1, the others $\Sigma(\gamma)$.

Synchronised permutation global tests applied to the first case study using Tippett combining function, with an increasing number of data samples.

Comparison with similar tests



(a) Gaussian data.



(b) Non-Gaussian data.

Synthetic data are generated from a **Gaussian** process and a multivariate **t-Student**. The combining function is Tippett.

Missing datum

A mouse died of unknown causes at the beginning of the experiment.

Missing datum

A mouse died of unknown causes at the beginning of the experiment.

Pesarin and Salmaso (2010)

Notation

O: inclusion indicator associated to X

$F[\mathbf{t}|(\mathbf{X}, \mathbf{O})]$, $\mathbf{t} \in \mathbb{R}^k$: permutation distribution of the test statistic $\mathbf{T}$

$\boldsymbol{\kappa} = [\kappa_1, \ldots, \kappa_q]$: actual sample sizes of valid data in each group

Hypotheses

Data are missing completely at random.

Under H_0, the permutation distribution of T_i depends only on κ_i

$$H_0 : \{(\mathbf{X}_1, \mathbf{O}_1) \stackrel{d}{=} \ldots \stackrel{d}{=} (\mathbf{X}_q, \mathbf{O}_q)\} \quad \text{against} \quad H_1 : \{H_0 \text{ is not true}\}$$

Then, it suffices to prove that $F[\mathbf{t}|(\mathbf{X}, \boldsymbol{\kappa})] = F[\mathbf{t}|(\mathbf{X}, \boldsymbol{\kappa}^*)]$

Here: $T_i^* = d(S_i^*, S_j^*)$ symmetric, $k_i^* = 20, k_j^* = 19 \vee k_i^* = 19, k_j^* = 20$

Test on the covariance operators

	Line 1	Line 2	Line 3	Line 4	Line 5	Line 6	Line 7	Line 8
Line 1	**1**							
Line 2	**<0.001**	**1**						
Line 3	**0.76**	**0.4**	**1**					
Line 4	**0.72**	**0.98**	**0.71**	**1**				
Line 5	**0.76**	**0.91**	**0.99**	**0.99**	**1**			
Line 6	**0.13**	**0.56**	**0.41**	**0.99**	**0.78**	**1**		
Line 7	**0.34**	**0.76**	**0.99**	**0.99**	**0.99**	**0.99**	**1**	
Line 8	**0.01**	**0.99**	**0.71**	**0.85**	**0.99**	**0.76**	**0.99**	**1**

Partial *p*-values of the synchronised permutation test on the covariance operators.
Global *p*-value is less than 1/1000.

Summary and future work

Summary

- No assumptions on the data generating process
- Works with any distance
- Straightforward post-hoc comparisons if groups are balanced
- Computationally intensive

Future work

- Handling of missing data and unbalanced groups
- Implementation in C++

References

Alessandra Cabassi, Davide Pigoli, Piercesare Secchi and Patrick A Carter (2017).
Permutation tests for the equality of covariance operators of functional data with applications to evolutionary biology. *Electron. J. Statist. 11, no. 2, 3815–3840.*

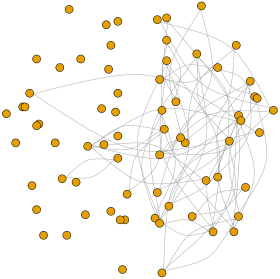
Alessandra Cabassi and Adam B Kashlak (2017).
fdcov: Analysis of Covariance Operators.
R package version 1.1.0.

Extension to network data

Joint work with

Matteo Fontana, Politecnico di Milano & Prof Alessio Farcomeni, La Sapienza Università di Roma.

Goal: Test for differences in brain functional connectivity of two or more groups of patients



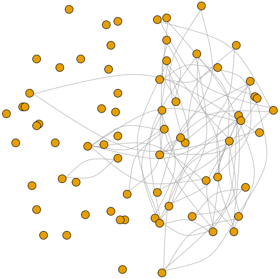
Functional network of a brain:
nodes = regions of interest,
edges = correlation of brain activity.

Extension to network data

Joint work with

Matteo Fontana, Politecnico di Milano & Prof Alessio Farcomeni, La Sapienza Università di Roma.

Goal: Test for differences in brain functional connectivity of two or more groups of patients



Functional network of a brain:
nodes = regions of interest,
edges = correlation of brain activity.

Consider q groups of functional brain networks, realisations of random processes with mean Γ_i

$$G_{i,1}, G_{i,2} \dots, G_{i,n_i}, \quad i = 1, \dots, q.$$

We wish to test the hypothesis

$$H_0 : \{\Gamma_1 = \dots = \Gamma_q\} \text{ against}$$

$$H_1 : \{\text{at least one equality is not true}\}$$

and, if H_0 is rejected, perform pairwise comparisons between the groups.

Extension to network data

Idea: Identify functional networks as positive semi-definite matrices.

So the test can be performed similarly to before:

- Test statistic:
$$d(\hat{\Gamma}_i, \hat{\Gamma}_j),$$
where $\hat{\Gamma}_i$ sample Fréchet mean of the functional networks $G_{i,m}$
$$\hat{\Gamma}_i = \arg\inf_{\Gamma} \sum_{m=1}^{n_i} d(G_{i,m}, \Gamma)^2.$$
- Distances: finite-dimensional versions of the kernel, square root, and Procrustes distances.

Extension to network data

Idea: Identify functional networks as positive semi-definite matrices.

So the test can be performed similarly to before:

- Test statistic:

$$d(\hat{\Gamma}_i, \hat{\Gamma}_j),$$

  where $\hat{\Gamma}_i$ sample Fréchet mean of the functional networks $G_{i,m}$

$$\hat{\Gamma}_i = \arg\inf_{\Gamma} \sum_{m=1}^{n_i} d(G_{i,m}, \Gamma)^2.$$

- Distances: finite-dimensional versions of the kernel, square root, and Procrustes distances.

Reference

Alessandra Cabassi, Alessandro Casa, Matteo Fontana, Massimiliano Russo and Alessio Farcomeni.
Three Testing Perspectives on Connectome Data.
Under review.

Thanks for listening!

alessandracabassi.wordpress.com

ac2051@cam.ac.uk

@sandy.cabassi

References

Alessandra Cabassi, Davide Pigoli, Piercesare Secchi and Patrick A Carter (2017).
Permutation tests for the equality of covariance operators of functional data with applications to evolutionary biology. *Electron. J. Statist. 11 (2017), no. 2, 3815–3840.*

Alessandra Cabassi and Adam B Kashlak (2016).
fdcov: Analysis of Covariance Operators.
R package version 1.0.0.

Ruth Marcus, Peritz Eric, and K Ruben Gabriel (1976).
On closed testing procedures with special reference to ordered analysis of variance.
Biometrika, 63(3):655–660.

Fortunato Pesarin and Luigi Salmaso (2010).
Permutation tests for complex data: theory, applications and software.
John Wiley and Sons.

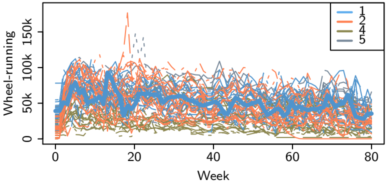
Davide Pigoli, John AD Aston, Ian L Dryden, and Piercesare Secchi (2014).
Distances and inference for covariance operators.
Biometrika, 101(2):409–422.

James O Ramsay and Bernard W Silverman (2005).
Functional Data Analysis.
Springer Series in Statistics.

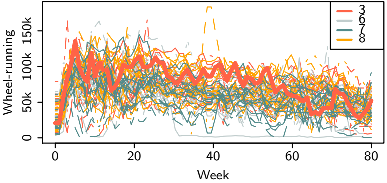
Aldo Solari, Luigi Salmaso, Fortunato Pesarin, and Dario Basso (2009).
Permutation Tests for Stochastic Ordering and ANOVA.
Springer New York.

Peter H Westfall and S Stanley Young (1993).
Resampling-based multiple testing: Examples and methods for p-value adjustment.
John Wiley and Sons, volume 279.

Raw data



(a) Control lines



(b) Selected lines

Number of revolutions run per week.
Raw data.

Smoothing

Ramsay and Silverman (2005), **Penalised Sum of Squared Errors**

$$\texttt{PENSSE}_\delta(x|\mathbf{y}) = [\mathbf{y} - x(\mathbf{t})]' W [\mathbf{y} - x(\mathbf{t})] + \delta \; \texttt{PEN}(x)$$

Sum of squared errors	$\texttt{SMSSE}(\mathbf{y}|\mathbf{c}) = (\mathbf{y} - \Pi\mathbf{c})' W (\mathbf{y} - \Pi\mathbf{c})$
Roughness penalty term	$\texttt{PEN}_2(x) = \|\ddot{x}\|^2 = \int_I [\ddot{x}(s)]^2 ds$

De Boor (2002)

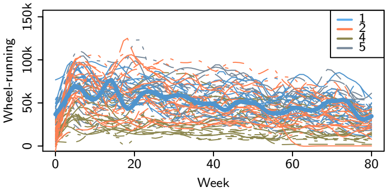
$\texttt{PENSSE}_\delta(x|\mathbf{y})$ is minimised by a cubic spline with breakpoints at the data points.

Craven and Wahba (1978)

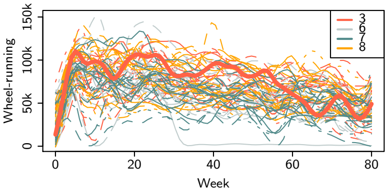
$$\hat{\delta} = \arg \min_{\delta \in (0,+\infty)} \texttt{GCV}(\delta) = \arg \min_{\delta \in (0,+\infty)} \left(\frac{p}{p - \mathrm{df}(\delta)} \right) \left(\frac{\texttt{SSE}}{p - \mathrm{df}(\delta)} \right)$$

$\mathrm{df}(\delta)$: number of degrees of freedom of the spline

Smoothing



(a) Control lines



(b) Selected lines

Number of revolutions run per week.
Smooth data.

Generalisation of Levene's test

Levene (1960), Anderson (2006)

Perform ANOVA analysis on the distances from individual points within each group to the group centroid, i.e. the sample functional mean:

$$z_{ij} = d(x_{ij}, \bar{x}_j), \quad \forall i, j.$$

In other words, compare the test statistic

$$T = \frac{n-q}{q-1} \frac{\sum_{i=1}^{q} (\bar{z}_i - \bar{z})^2}{\sum_{i=1}^{q} \sum_{j=1}^{n_i} (z_{ij} - \bar{z}_i)^2}$$

against $F(\alpha, q-1, n-q)$.

Asymptotic approach

Dauxois et al. (1982)

If $\mathbb{E}\|X_{ij}\|^4 < \infty$, then $\sqrt{n_i}(S_i - \Sigma_i)$ converges in distribution to a zero-mean Gaussian random element of $\mathcal{F}$, the Hilbert space of Hilbert-Schmidt operators, with covariance operator Υ_i.

Boente et al. (2014)

Given a sample $x_{i,1}, \ldots, x_{i,n_i}$, let $\hat{\Upsilon}_i$ be consistent estimators of Υ_i $i = 1, 2$.

1. Define $\hat{\Upsilon} = \hat{\tau}_1^{-1}\hat{\Upsilon}_1 + \hat{\tau}_2^{-1}\hat{\Upsilon}_2$ with $\hat{\tau}_i = n_i/(n_1 + n_2)$.
2. For $1 \leq l \leq q_n$ denote by $\hat{\iota}_l$ the positive eigenvalues of $\hat{\Upsilon}$
3. Generate $Z_1^*, \ldots, Z_{q_n}^*$ i.i.d. such that $Z_i^* \sim \mathcal{N}(0, 1)$ and let

$$U_n^* = \sum_{j=1}^{q_n} \hat{\iota}_j Z_j^{*2}$$

4. Repeat the previous step B times, to get B values of U_n^*

Empirical bootstrap

Paparoditis and Sapatinas (2014)

The observed collection of random functions satisfy

$$x_{i,j}(t) = \mu_i(t) + \varepsilon_{i,j}(t), \quad \forall i,j \quad \forall t \in I.$$

Let T be a given test statistic of interest based on X.

1. Calculate $\bar{x}_{i,n_i} = n_i{}^{-1} \sum_{j=1}^{n_i} x_{i,j}, \quad \forall i$
2. Calculate $\varepsilon_{i,j} = x_{i,j} - \bar{x}_{i,n_i}, \quad \forall i,j.$
3. Generate bootstrap functional pseudo-observations

   $$x^*_{i,j} = \bar{x}_{i,n_i} + \varepsilon^*_{i,j}, \quad \forall i,j$$

   where $\varepsilon^*_{i,j} = \varepsilon_{I,J}$ and (I,J) is a pair of random variables.
4. Repeat B times and compute the test statistic T^*
5. Compute the approximated distribution of T^*
6. Reject H_0 iff $T > c^*_\alpha$, where $Pr(T^* > c^*_\alpha) = \alpha$.

Non-asymptotic test based on concentration inequalities

Kashlak et al. (2016)

p-Schatten norm

$$\|\Sigma\|_p^p = \begin{cases} \|\tilde{\phi}\|_{l^p}^p = \sum_{n=1}^{\infty} |\tilde{\phi}_n^p|, & \text{if } p \in [1, \infty), \\ \max_{n\in\mathbb{N}} |\tilde{\phi}_n|, & \text{if } p = \infty. \end{cases}$$

Confidence set

$$\left\{\Sigma : \|S - \Sigma\|_p \leq \|R_n\|_p + \varsigma \Big[-\frac{2}{n} \log(2\alpha) \Big]^{1/2} - \frac{\varsigma \log(2\alpha)}{3n} \right\},$$

where R_n is the Rademacher average

$$R_n = \frac{1}{n} \sum_{i=1}^{n} \chi_i \big[(x_i - m)^{\otimes 2} - S \big].$$

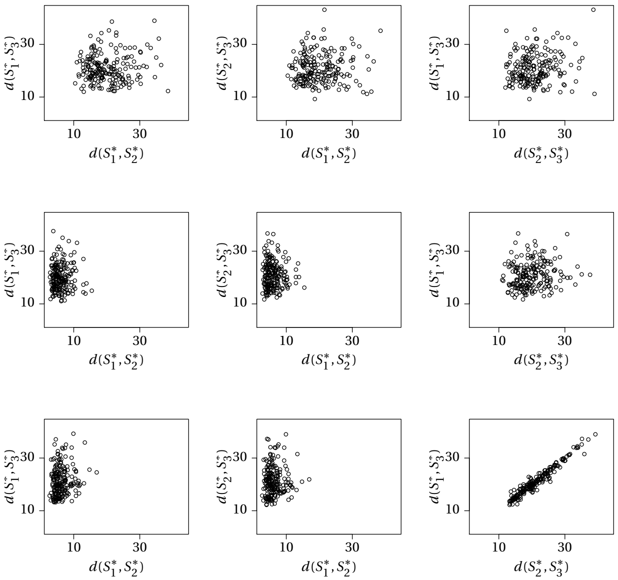
Solari et al. (2009)

If the design is balanced, i.e. $n_1 = \cdots = n_C = n$, we permute the rows of the pseudo-data matrix

$$\begin{bmatrix} \mathbf{x_1} & \mathbf{x_1} & \dots & \mathbf{x_{C-1}} \\ \mathbf{x_2} & \mathbf{x_3} & \dots & \mathbf{x_C} \end{bmatrix} = \begin{bmatrix} x_1^1 & x_1^1 & \dots & x_1^{C-1} \\ x_2^1 & x_2^1 & \dots & x_2^{C-1} \\ \vdots & \vdots & & \vdots \\ x_n^1 & x_n^1 & \dots & x_n^{C-1} \\ \\ x_1^2 & x_1^3 & \dots & x_1^C \\ x_2^2 & x_2^3 & \dots & x_2^C \\ \vdots & \vdots & & \vdots \\ x_n^2 & x_n^3 & \dots & x_n^C \end{bmatrix}$$

Synchronized permutations allow both maintaining the dependencies among partial tests and involving the observations of each comparison at the same time.

Permutation strategies



Pesarin and Salmaso (2010)

1^{st} phase: estimating the k-variate distribution of $\mathbf{T}$

1. Calculate the vector of observed values of tests: $\mathbf{T}^0 = \mathbf{T}(X)$
2. Consider a random permutation $X^* \in X_{/X}$ of X and compute $\mathbf{T}^* = \mathbf{T}(X^*)$
3. Carry out B independent repetitions of the previous step: $\{\mathbf{T}^{(b)}\}_{b=1}^{B}$ is a random sampling from the permutation distribution of $\mathbf{T}$
4. A consistent estimate of the CDF $F(\mathbf{t}|X_{/X})$ is
$$\hat{F}(\mathbf{t}|X_{/X}) = \frac{\sum_b \mathbb{1}(\mathbf{T}^{(b)} \leq \mathbf{t})}{B}$$
5. A consistent estimate of $p_i = Pr\{T_i^* \geq t|X_{/X}\}$ is
$$\hat{p}_i(t|X_{/X}) = \frac{\sum_b \mathbb{1}(T_i^{(b)} \geq t)}{B}$$

Non-parametric combination methodology

Pesarin and Salmaso (2010)

2^{nd} phase: simulating a procedure for NPC

1. The k observed p-values are estimated by $\lambda_i = \hat{p}_i(T_i^0|X_{/X})$
2. The combined observed value of the test is $T_\Psi^0 = \Psi(\lambda_1, \dots, \lambda_k)$
3. The bth combined value is

$$T_\Psi^{(b)} = \Psi(\hat{p}_1^{(b)}, \dots, \hat{p}_k^{(b)}), \quad \hat{p}_i^{(b)} = \hat{p}_i(T_i^{(b)}|X_{/X}) \quad i = 1, \dots, k, \; b = 1, \dots, B$$

4. The p-value of the combined test T is estimated as

$$\lambda_\Psi = \frac{\sum_b \mathbb{1}(T_\Psi^{(b)} \geq T_\Psi^0)}{B}$$

5. If $\lambda_\Psi \leq \alpha$, H_0 is rejected

Marcus et al. (1976)

Closed testing procedure

Consider the closure of the set, which is the set of all possible intersection hypotheses.

1. Test all the hypotheses simultaneously by using permutation tests:
   - Calculate the statistics $T_{\bar{K}}$ for each non-empty $\bar{K} \subseteq \{1, \ldots, k\}$;
   - Perform B permutations
   - Compute the permutation statistics $T^*_{\bar{K}}$ for each non-empty $\bar{K}$
   - Calculate the raw p-values as

$$\lambda_{\bar{K}} = \frac{\sum_b \mathbb{1}\left[T^*_{\bar{K}}(b) \geq T_{\bar{K}}\right]}{B};$$

2. Reject any hypothesis H_{0i} when the test of H_{0i} itself is significant and the test of every intersection hypothesis that includes H_{0i} is significant.

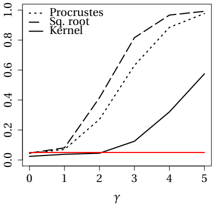
Westfall and Young (1993)

Step-down procedure for Tippett combining function

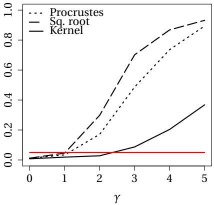
Let $\lambda_{(1)}, \dots, \lambda_{(k)}$ be the increasing ordered partial p-values.

1. $\tilde{\lambda}_{(1)} = \lambda_{(1),\dots,(k),\text{Tippett}}$
   - If $\tilde{\lambda}_{(1)} \leq \alpha$, reject $H_{0(1)}$ and continue;
   - Otherwise retain $H_{0(1)}, \dots, H_{0(k)}$ and stop.
2. $\tilde{\lambda}_{(i)} = \max\{\lambda_{(i),\dots,(k),\text{Tippett}}, \tilde{\lambda}_{(i-1)}\}$
   - If $\tilde{\lambda}_{(i)} \leq \alpha$, reject also $H_{0(i)}$ and continue;
   - Otherwise retain $H_{0(i)}, \dots, H_{0(k)}$ and stop.

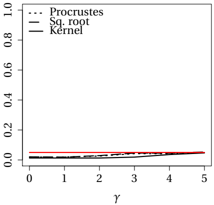
Partial tests



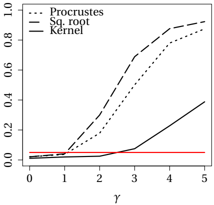
(a) Global test.



(b) Partial test, samples 1 and 2.



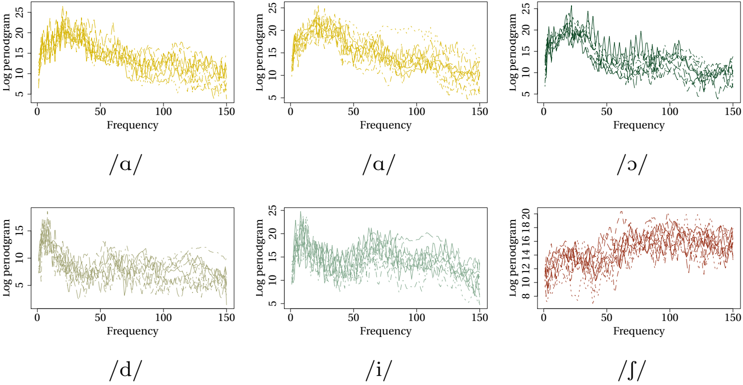
(c) Partial test, samples 2 and 3.



(d) Partial test, samples 1 and 3.

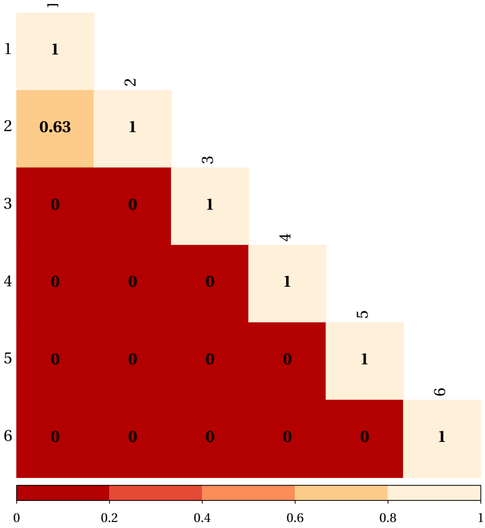
Hastie et al. (1995)

200 samples in each group, 150 frequencies.



Speech recognition data set, raw data.

Test on the covariance operators



Partial p-values of the synchronised permutation test on the covariance operators. Global p-value is less than $1/1000$.

References - II

Marti J Anderson (2006).
Distance-based tests for homogeneity of multivariate dispersions.
Biometrics, 62(1):245–253.

Graciela Boente, Daniela Rodriguez, and Mariela Sued (2014).
A test for the equality of covariance operators.
arXiv preprint arXiv:1404.7080.

Peter Craven and Grace Wahba (1978).
Smoothing noisy data with spline functions.
Numerische Mathematik, 31(4):377–403.

Jacques Dauxois, Alain Pousse, and Yves Romain (1982).
Asymptotic theory for the principal component analysis of a vector random function: some applications to statistical inference.
Journal of multivariate analysis, 12(1):136–154.

Carl De Boor (2002).
Spline basics.
Handbook of Computer Aided Geometric Design, Elsevier, 13:141–164.

Adam B Kashlak, John AD Aston, and Richard Nickl (2016).
Inference on covariance operators via concentration inequalities: k-sample tests, classification, and clustering via rademacher complexities.
arXiv preprint arXiv:1604.06310.

Howard Levene (1960).
Robust tests for equality of variances.
Contributions to probability and statistics: Essays in honor of Harold Hotelling, 2:278–292.

References - III

Ruth Marcus, Peritz Eric, and K Ruben Gabriel (1976).
On closed testing procedures with special reference to ordered analysis of variance.
Biometrika, 63(3):655–660.

Efstathios Paparoditis and Theofanis Sapatinas (2014).
Bootstrap-based testing for functional data.
arXiv preprint arXiv:1409.4317.

Fortunato Pesarin and Luigi Salmaso (2010).
Permutation tests for complex data: theory, applications and software.
John Wiley & Sons.

James O Ramsay and Bernard W Silverman (2005).
Functional Data Analysis.
Springer Series in Statistics.

Aldo Solari, Luigi Salmaso, Fortunato Pesarin, and Dario Basso (2009).
Permutation Tests for Stochastic Ordering and ANOVA.
Springer New York.

Peter H Westfall and S Stanley Young (1993).
Resampling-based multiple testing: Examples and methods for p-value adjustment.
John Wiley & Sons, volume 279.