Multivariate and multiple contrast testing in general covariate-adjusted factorial designs

A semiparametric resampling approach

Marléne Baumeister

No Science without Collaboration

- Prof. Dr. Markus Pauly, Konstantin Emil Thiel, Dr. Paavo Sattler (Department of Statistics, TU Dortmund University);
- Lynn Matits (Clinical & Biological Psychology, Institute of Psychology and Education, Ulm University);
- Dr. Georg Zimmermann (Team Biostatistics and Big Medical Data, IDA Lab Salzburg, Paracelsus Medical University).

Motivation

Motivation: A Statistical Model for Intervention Studies

Endpoints : highly correlated d-dimensional

Groups (k=2):   Control group n_1 Intervention group n_2
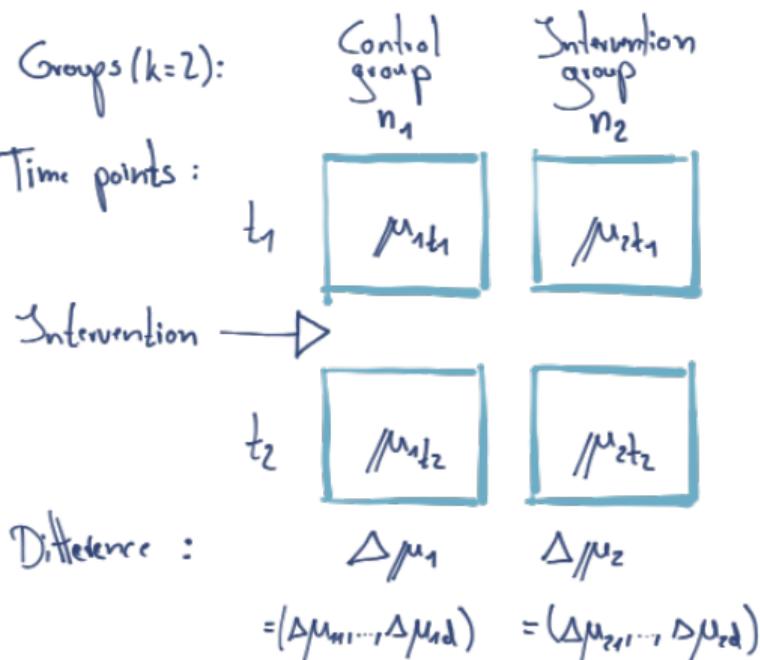
Covariables :
c-dimensional

attributes of the patients or environmental features

$\Rightarrow$ estimands are covariate-adjusted

Time points :

t_1 μ_{1t_1} μ_{2t_1}

Intervention $\longrightarrow$

t_2 μ_{1t_2} μ_{2t_2}

Hypotheses:
$H_0: \Delta\mu_1 = \Delta\mu_2$
$H_{0,l}: \Delta\mu_{1l} = \Delta\mu_{2l}$
$l \in \{1, \ldots, d\}$

Difference : $\Delta\mu_1$ $\Delta\mu_2$
$= (\Delta\mu_{11}, \ldots, \Delta\mu_{1d})$ $= (\Delta\mu_{21}, \ldots, \Delta\mu_{2d})$

Combining Two Trending Topics

Multiple Testing

- After rejecting a hypothesis in a complex factorial design we usually want to do post-hoc comparisons.
- $\rightarrow$ Simple solution: Multiple contrast test procedures (MCTPs, e.g. Bretz et al. (2001)).

Covariate Adjustment

- There is a desire or a requirement to adjust according to covariates.
- European Medicines Agency (EMA) recommends adjustments for covariates for example in randomized clinical trials.

Our example: less chronic stress through hypnosis?

- Stress is associated with changes in the heart rate variability
- Intervention: hypnosis.
- Data set:
  - Sample with control ($n_1 = 23$) and intervention ($n_2 = 22$) group,
  - multiple endpoints, that characterize the **heart rate variability**: SDNN[1], RMSSD[2], HF[3], LF[4], VLF[5].
  - Covariates: **chronic stress** and **suggestibility**: PSS, HGSH.

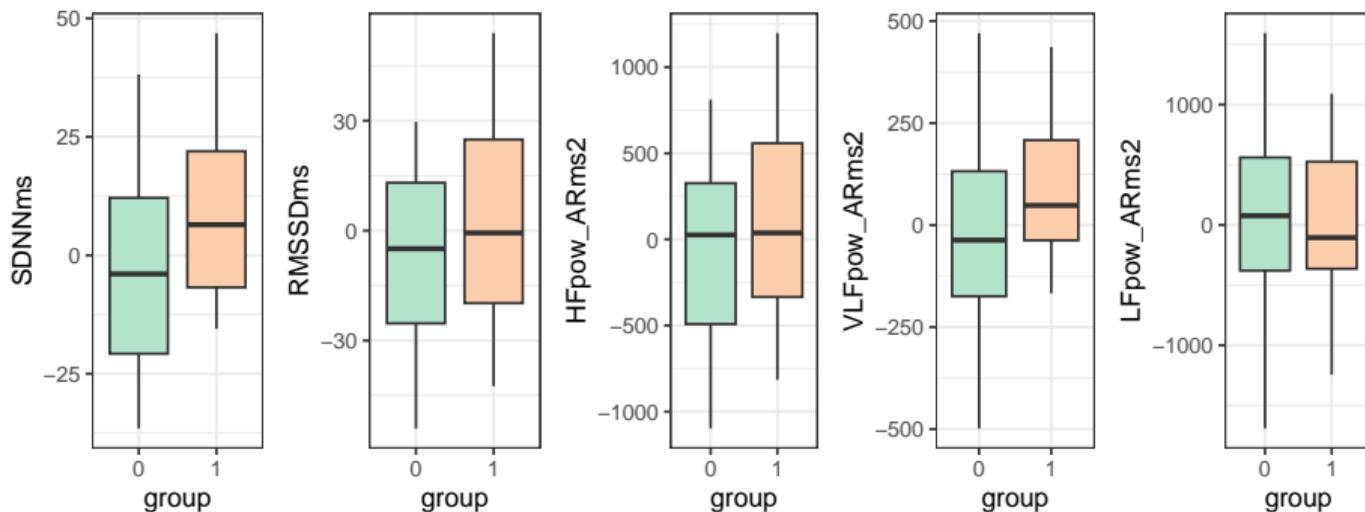[1] Standard Deviation of Normal-to-Normal Intervals, in ms.

[2] Root Mean Square of Successive Differences, in ms.
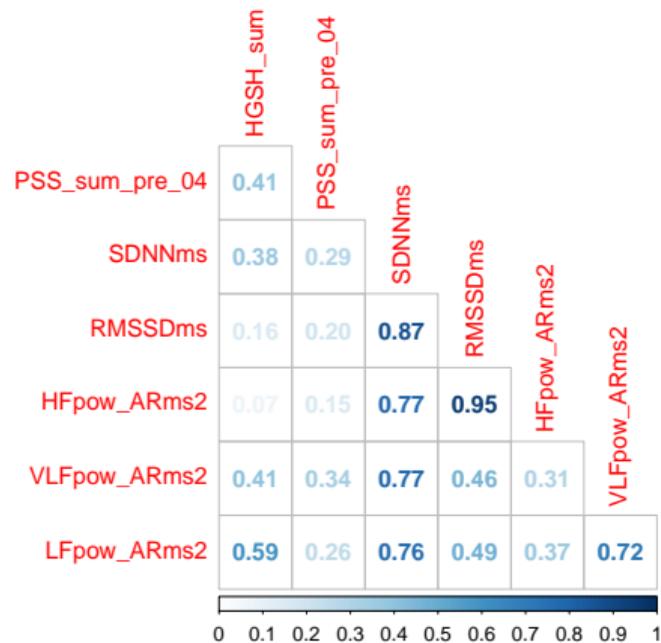
[3] High Frequency, in ms^2.

[4] Low Frequency, in ms^2.

[5] Very Low Frequency, in ms^2.

Heart Rate Parameters: Change from Baseline

Heart Rate Parameters: Correlation with Baseline Covariates

Statistical Methods

General Set-Up

■ MANCOVA model equation:

$$\mathbf{Y} = \tilde{\mathbf{M}}\boldsymbol{\mu} + \tilde{\mathbf{Z}}\boldsymbol{\nu} + \boldsymbol{\epsilon},$$

where $\tilde{\mathbf{M}} = \bigoplus_{i=1}^{k}(\mathbf{1}_{n_i} \otimes \mathbf{I}_d)$ and $\tilde{\mathbf{Z}} = \mathbf{Z} \otimes \mathbf{I}_d$.

- Groups $i \in \{1, \ldots, k\}$, subjects $j \in \{1, \ldots, n_i\}$, $n := \sum_{i=1}^{k} n_i$, dimensions $\ell \in \{1, \ldots, d\}$,
- $\mathbf{Y} = (\mathbf{Y}'_{11}, \ldots, \mathbf{Y}'_{kn_k})'$ vector of d-dimensional outcomes $\mathbf{Y}_{ij} = (Y_{ij1}, \ldots, Y_{ijd})'$,
- $\mathbf{Z} = (\mathbf{z}_{11}, \ldots, \mathbf{z}_{kn_k})'$ matrix of c covariates $\mathbf{z}_{ij} = (z_{ij1}, \ldots, z_{ijc})'$,
- $\boldsymbol{\epsilon} = (\boldsymbol{\epsilon}'_{11}, \ldots, \boldsymbol{\epsilon}'_{kn_k})'$, vector of k errors $\boldsymbol{\epsilon}_{ij} = (\epsilon_{ij1}, \ldots, \epsilon_{ijd})'$
- $\boldsymbol{\mu} = (\boldsymbol{\mu}'_1, \ldots, \boldsymbol{\mu}'_k)'$, vector of k adjusted means $\boldsymbol{\mu}_i = (\mu_{i1}, \ldots, \mu_{id})'$,
- $\boldsymbol{\nu} = (\boldsymbol{\nu}'_1, \ldots, \boldsymbol{\nu}'_c)'$, vector of c regression coefficients $\boldsymbol{\nu}_m = (\nu_{m1}, \ldots, \nu_{md})'$.

Assumptions

Semiparametric model:

- No distributional assumption.
- The errors ϵ_{ij} are mutually independent with $E(\epsilon) = \mathbf{0}$ and a group-specific covariance matrix $\mathbf{\Sigma}_i := \mathsf{Cov}(\epsilon_{ij})$, $E(\|\epsilon_{ij}\|^4) < \infty$.
- Positive variance: $\sigma_{i\ell}^2 := \mathsf{Var}(\epsilon_{ij\ell}) > 0$, thus heteroscedasticity and singularity are allowed.
- Other technical assumptions.

The General Multiple Testing Problem

- Consider r contrast vectors: $\mathbf{h}_s = (\mathbf{h}'_{s1}, \ldots, \mathbf{h}'_{sk})' \in \mathbb{R}^{kd}$, $\mathbf{h}_{si} = (h_{si1}, \ldots, h_{sid})'$, $s \in \{1, \ldots, r\}$.
    - $\mathbf{h}_s$ is a contrast vector, if $\sum_{i=1}^{k} \sum_{\ell=1}^{d} h_{si\ell} = 0$.
$\longrightarrow$ Combine them to a contrast matrix $\mathbf{H} = (\mathbf{h}_1, \ldots, \mathbf{h}_r)' \in \mathbb{R}^{r \times kd}$.
- Selected hypotheses: $\mathcal{H}_{0,s} : \mathbf{h}'_s \boldsymbol{\mu} = \mathbf{0}$ for every $s \in \{1, \ldots, r\}$.
- Global hypothesis $\mathcal{H}_0 : \mathbf{H}\boldsymbol{\mu} = \mathbf{0}$.

Estimators: adjusted mean

- Ordinary least squares (OLS) estimator $\hat{\boldsymbol{\mu}} = (\boldsymbol{\mu}_1', \dots, \boldsymbol{\mu}_k')'$ for $\boldsymbol{\mu}$:

$$\hat{\boldsymbol{\mu}}_i = \bar{\mathbf{Y}}_{i.} - \left(\bar{\mathbf{z}}_{i.} \otimes \mathbf{1}_d'\right) \hat{\boldsymbol{\nu}},$$

where $\hat{\boldsymbol{\nu}}$ is the OLS estimator of $\boldsymbol{\nu}$ (Eicker, 1963).

$\longrightarrow$ Consistent as a special form of the OLS estimator.

Estimators: sandwich covariance estimator

Adapt the sandwich covariance estimator of Eicker (1963) for multivariate data:

- HC4-adjusted (Cribari-Neto, 2004) squared residuals $\hat{\boldsymbol{\Sigma}}_{ij}$ to create a block diagonal matrix: $\hat{\mathbf{S}} = \bigoplus_{i=1}^{k} \bigoplus_{j=1}^{n_i} \hat{\boldsymbol{\Sigma}}_{ij}$.

$\longrightarrow$ Consistent estimator for $\boldsymbol{\Lambda}_{11}$:

$$\hat{\boldsymbol{\Lambda}}_{11} := n(\tilde{\mathbf{X}}'\tilde{\mathbf{X}})^{-1}\tilde{\mathbf{X}}'\hat{\mathbf{S}}\tilde{\mathbf{X}}(\tilde{\mathbf{X}}'\tilde{\mathbf{X}})^{-1},$$

with $\tilde{\mathbf{X}} = (\tilde{\mathbf{M}}, \tilde{\mathbf{Z}})$.

$\boldsymbol{\Lambda}$ is the limiting covariance matrix of $(\boldsymbol{\mu}, \boldsymbol{\nu})$ and $\boldsymbol{\Lambda}_{11}$ its upper left $dk \times dk$ dimensional block.

Estimators: zero-inflation

- In line with Friedrich and Pauly (2018) and Zimmermann et al. (2020): take only the diagonal elements of that covariance matrix estimator:

$$\hat{\mathbf{D}} := \mathrm{Diag}(\hat{\mathbf{\Lambda}}_{11}).$$

$\longrightarrow$ Consistent estimator for $\mathbf{D} := \mathrm{Diag}(\mathbf{\Lambda}_{11})$.

The Test's Construction: Test Statistics

- For every single null hypothesis $\mathcal{H}_{0,s}$, we consider one test statistic:

$$A_n(\mathbf{h}_s) = \sqrt{n} \frac{\mathbf{h}_s' \hat{\boldsymbol{\mu}}}{\sqrt{\mathbf{h}_s' \hat{\mathbf{D}} \mathbf{h}_s}} \xrightarrow{d} \mathbf{B} \sim \mathcal{N}\left(0, \frac{\mathbf{h}_s' \boldsymbol{\Lambda} \mathbf{h}_s}{\mathbf{h}_s' \mathbf{D} \mathbf{h}_s}\right). \tag{1}$$

- For the global test consider the maximum over all test statistics:

$$\max_{s \in \{1, \ldots, r\}} A_n(\mathbf{h}_s).$$

The Test's Construction: a Problem with the Classical Approach

- Classical MCTP approach: use the equicoordinate quantile of the joint distributions of $|A_n(\mathbf{h}_s)|$, $s \in \{1, \ldots, r\}$, for all single test comparisons.

$\longrightarrow$ As the limit distribution of **B** depends on the individual contrast, this is not possible in our set-up.

- Solution: adapt the approach of Munko et al. (2024).
    - Need for a bootstrap, e.g. the resampling methods proposed by Zimmermann et al. (2020) (parametric and wild bootstrap).

The Test's Construction: novel approach for the critical value

- Draw B bootstrap samples and consider for every sample $b \in \{1, \ldots, B\}$ the vector of test statistics $(A_n^{\circ,b}(\mathbf{h}_1), \ldots, A_n^{\circ,b}(\mathbf{h}_r))$.

- Define $q_{s,\beta}^{\circ}$ the $(1 - \beta)$-quantile of $|A_n^{\circ,b}(\mathbf{h}_s)|$ and optimize the local level β in a way that the level α is controlled globally.

$\longrightarrow$ Get optimized critical value $\beta_n(\alpha)$.
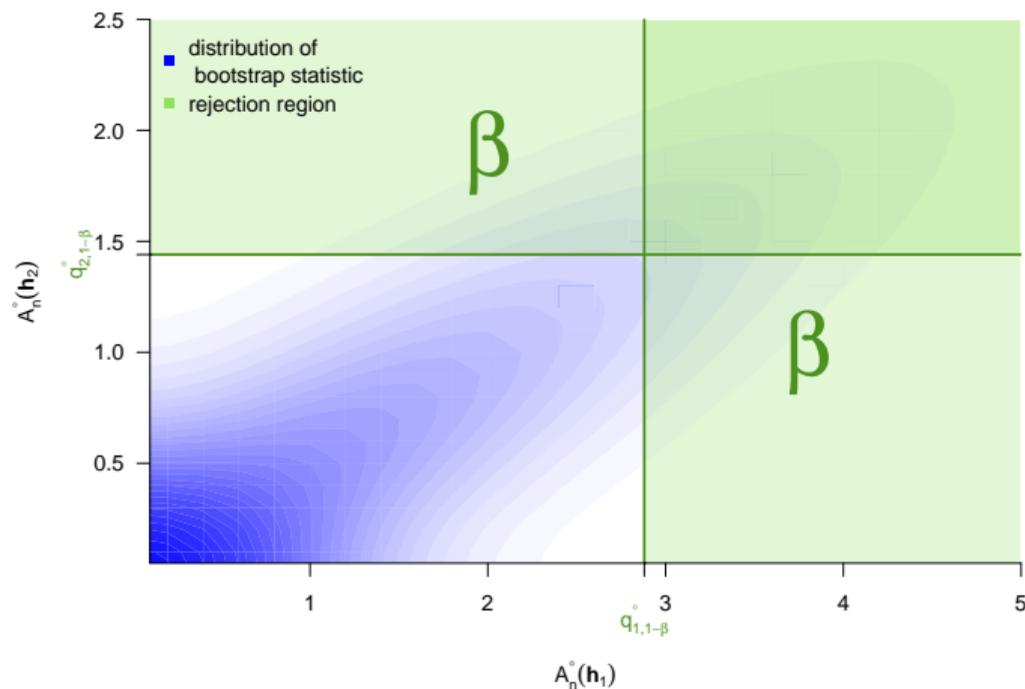
The Test's Construction: Testing Procedure

- One single hypothesis $\mathcal{H}_{0,s}$
    - is rejected if and only if $|A_n(\mathbf{h}_s)| > q^{\circ}_{s,\beta_n(\alpha)}$ or equivalently $|A_n(\mathbf{h}_s)|/q^{\circ}_{s,\beta_n(\alpha)} > 1$.
    - Tests $(0/0 := 0)$:

$$\varphi^{\circ}_{n,s} = \mathbf{1}\{|A_n(\mathbf{h}_s)| > q^{\circ}_{s,\beta_n(\alpha)}\}.$$

- The global hypothesis $\mathcal{H}_0$
    - is rejected, if and only if at least one $\mathcal{H}_{0,s}$ is rejected.
    - Global test:

$$\varphi^{\circ}_n = \mathbf{1}\left\{ \max_{s \in \{1,\dots,r\}} \frac{|A_n(\mathbf{h}_s)|}{q^{\circ}_{s,\beta_n(\alpha)}} > 1 \right\}.$$

Simulations

Simulation Set-Up

- Simulate FWER for the testing problem with random generated data for $k = 3$ group and $d = 3$ dimensions:

$$\mathbf{Y}_i = \mathbf{\Sigma}_i^{\frac{1}{2}} \mathbf{E}_i + \mathbf{Z}_i \cdot \boldsymbol{\nu}.$$

- Choose 5 distributions for $\mathbf{E}_i$, $i \in \{1, 2\}$,: standard normal, standard log normal, t_3, χ_3^2, double exponential,
- $c = 2$ covariates: $\mathbf{Z}_i \in \mathbb{R}^{n_i \times 2}$,
- a fixed $\boldsymbol{\nu}$, which ensures the correlation between $\mathbf{Z}_i$ and $\mathbf{Y}_i$,
- 3 different $\mathbf{\Sigma}_i$, $i \in \{1, 2\}$: homoscedastic (1), heteroscedastic (2) and singular (3),
- samples for $\mathbf{E}_i$: $K \cdot \{(10, 10, 10), (20, 10, 10), (10, 10, 20)\}$, $K \in \{1, 2, 3, 4\}$.

Simulation: Benchmark methods

- Asymptotic MCTPs: maximum test statistics compared with an asymptotic equicoordinate quantile from a normal and a t-distribution with $n - k$ degrees of freedom.
$\longrightarrow$ Not applicable in settings with singular covariance (Cov: 3)
- Bonferroni-corrected MANCATS: semiparametric MANCOVA by Zimmermann et al. (2020).
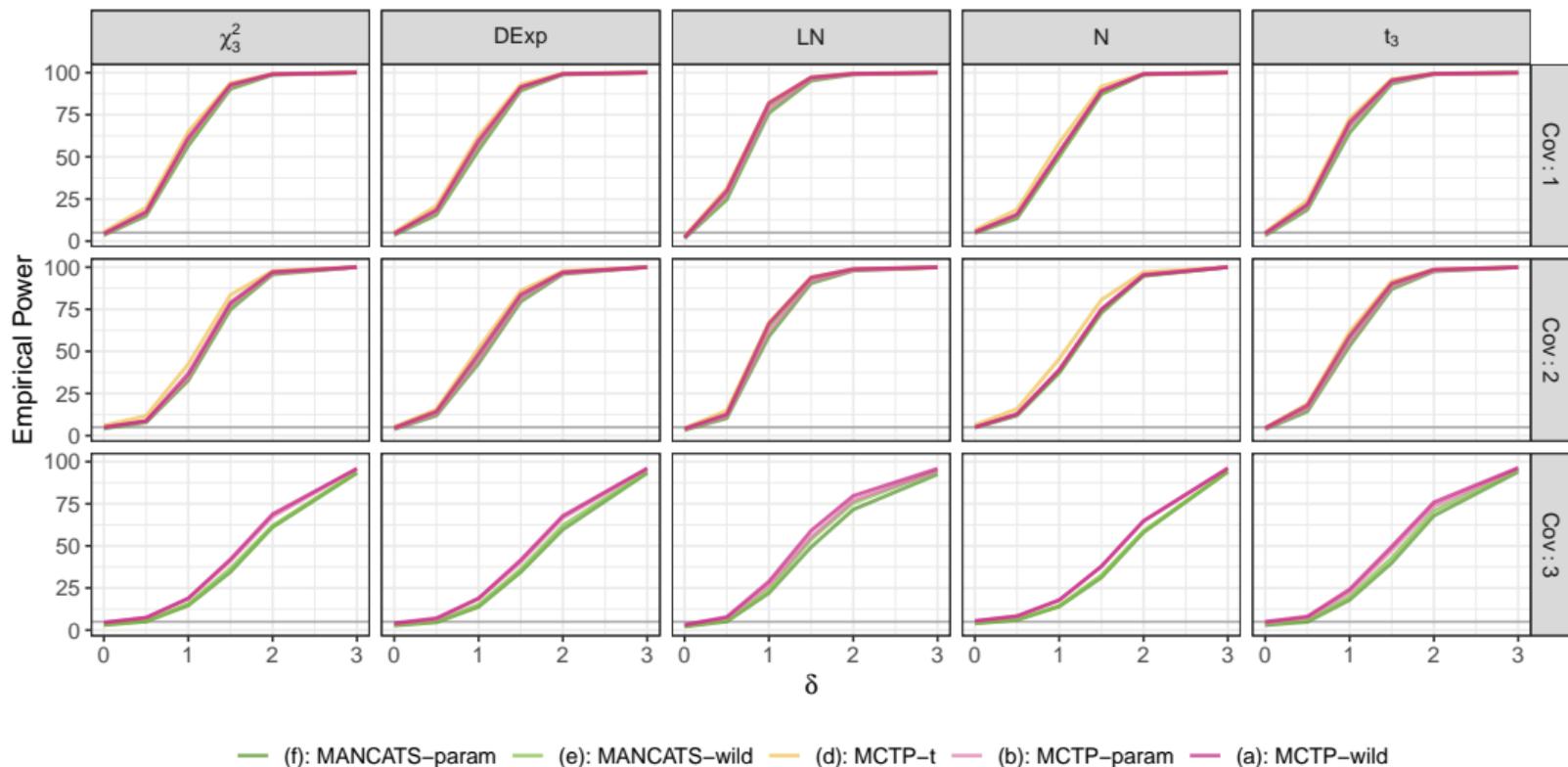$\longrightarrow$ Known to be too conservative regarding the FWER.

Simulation: Divided by Sample Size Factor
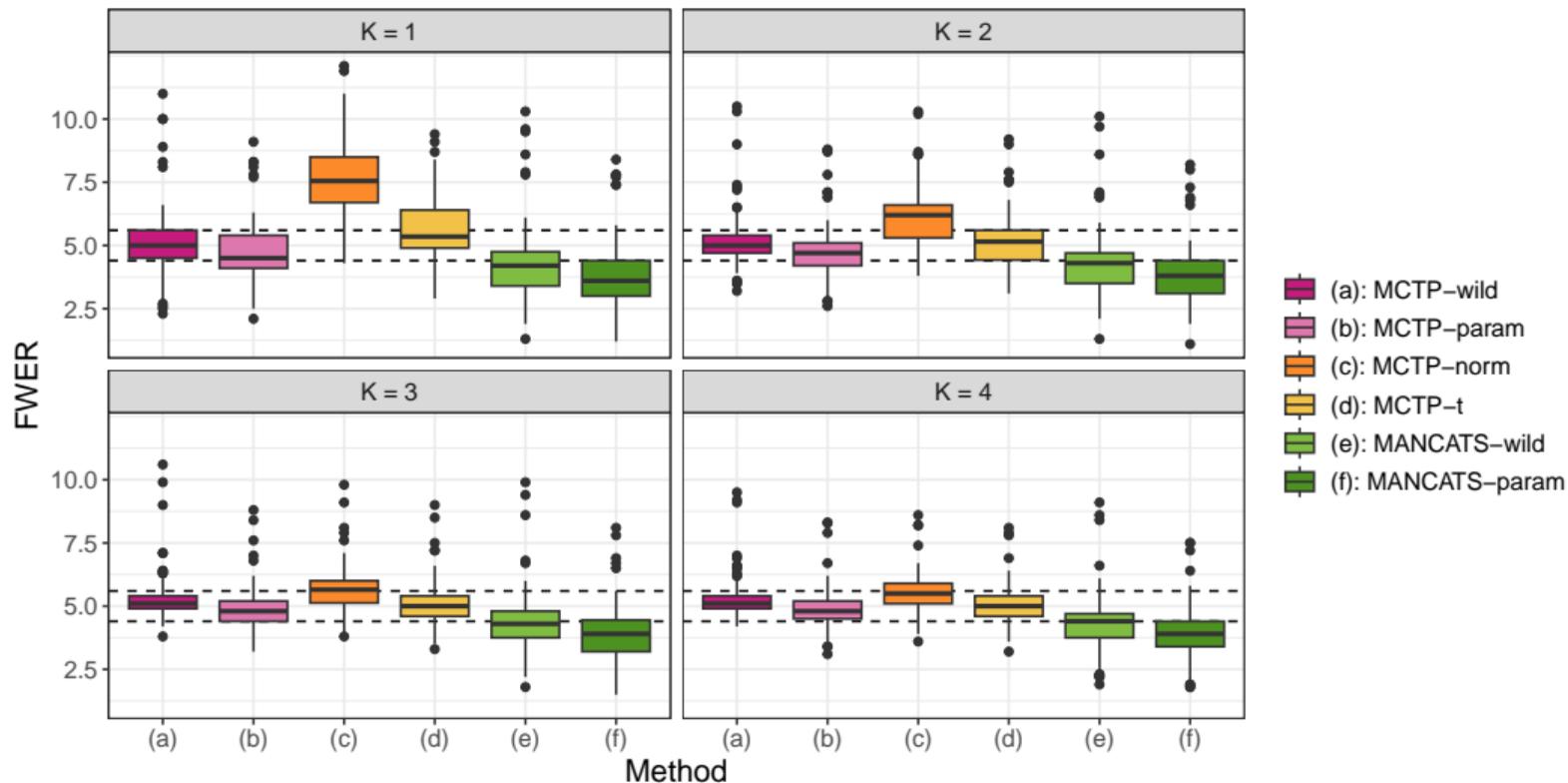
Simulation: Divided by Covariance Structure
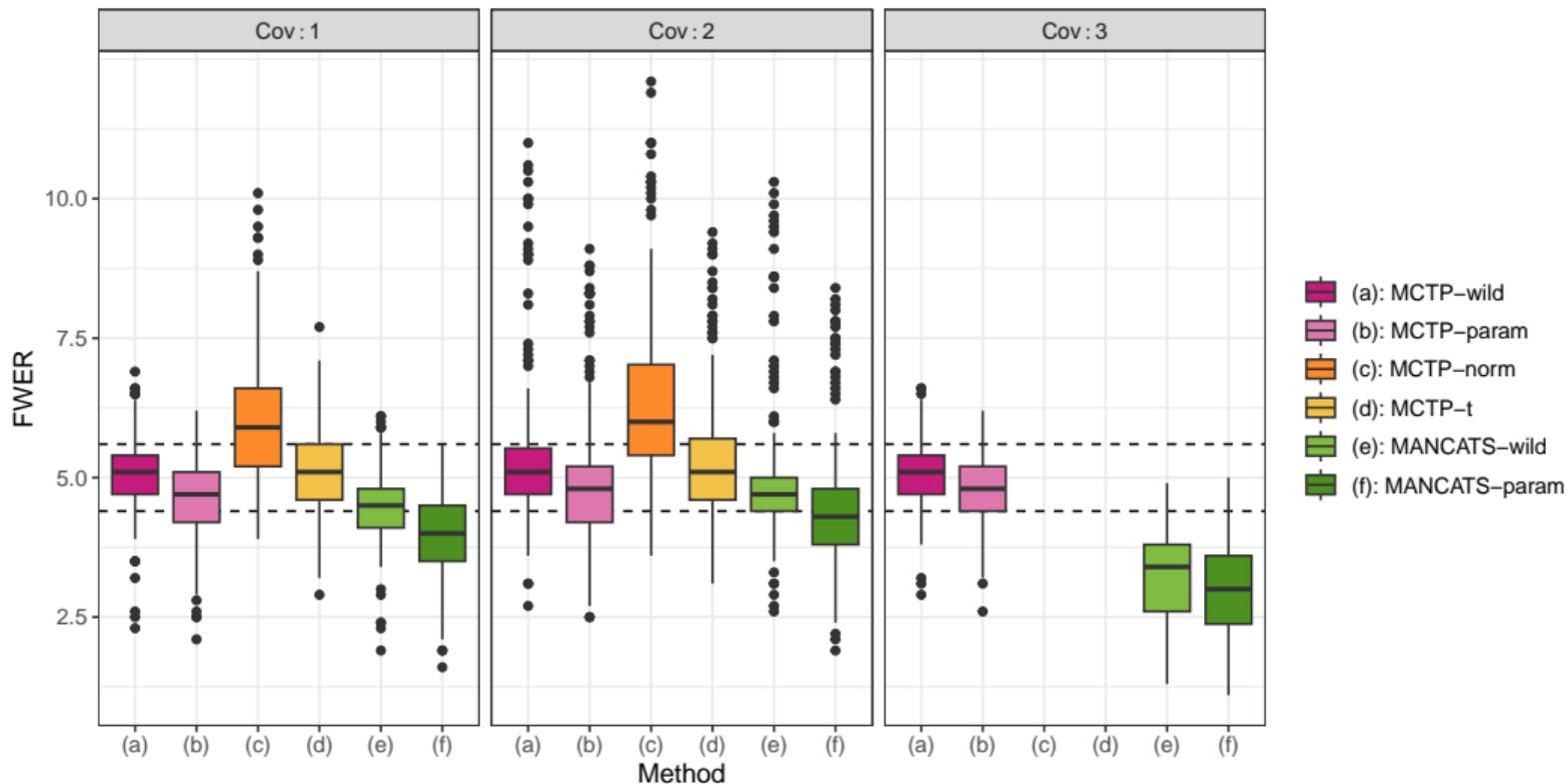
Simulation: Power

Summary

Data Example Revisited

- global $\alpha = 0.05$
- local p-values compared to adjusted local levels β:

	β	i	ii	iii	iv	v
t-test	0.01	0.051	0.193	0.364	0.084	0.66
wild bootstrap MANCATS	0.01	**0.004**	0.118	0.354	**0.008**	0.610
wild bootstrap MCTP	0.0145	**0.005**	0.127	0.348	**0.011**	0.591

Summary

- Asymptotically valid MCTP method for MANCOVA.
- Semiparametric model that allow for potential singularity.
- Covers various multivariate testing procedures.
- → Possibility of multivariate modelling in intervention studies.
- Holds the level of significance better than Bonferroni-adjustment with a comparable power.
- Method will be available in `MANOVA.RM` soon.

Outlook

What I am currently working on I

WEAVE-Project OUTCLASS- powerful tools for functional data in complex factorial designs with missing values

$\longrightarrow$ First goal: permutation tests in a functional matched pairs design with missing values.

- Consider two-dimensional functional data:

$$\mathbf{x}_1, \ldots, \mathbf{x}_n : \mathcal{T} \to \mathbb{R}^2, \; \mathbf{x}_j(t) = (x_{1j}(t), x_{2j}(t))', \; j \in \{1, \ldots, n\},$$

with missing values:

$$\underbrace{\begin{pmatrix} x_{11}^{(c)}(t) \\ x_{21}^{(c)}(t) \end{pmatrix}, \ldots, \begin{pmatrix} x_{1n_1}^{(c)}(t) \\ x_{2n_1}^{(c)}(t) \end{pmatrix}}_{\mathbf{x}^{(c)}}, \underbrace{\begin{pmatrix} x_{11}^{(m)}(t) \\ --- \end{pmatrix}, \ldots, \begin{pmatrix} x_{1n_2}^{(m)}(t) \\ --- \end{pmatrix}, \begin{pmatrix} --- \\ x_{21}^{(m)}(t) \end{pmatrix}, \ldots, \begin{pmatrix} --- \\ x_{2n_3}^{(m)}(t) \end{pmatrix}}_{\mathbf{x}^{(m)}}.$$

What I am currently working on II

- Consider hypotheses regarding the mean function $\boldsymbol{\eta} = (\eta_1, \eta_2) : \mathcal{T} \to \mathbb{R}^2$:

$$\mathcal{H}_0 : \forall \, t \in \mathcal{T} : \eta_1(t) = \eta_2(t), \quad \mathcal{H}_1 : \exists \, t \in \mathcal{T} : \eta_1(t) \neq \eta_2(t).$$

- To deal with the missing values, adapt the idea of Amro and Pauly (2016) for functional data to consider two different type of test statistics:
    - a pointwise paired t-type test statistic for the complete part $\mathbf{x}^{(c)}$,
    - and a pointwise Welch-t-type test statistic for the incomplete part $\mathbf{x}^{(m)}$.
- $\longrightarrow$ Combine them to one test statistic $K_n(\mathbf{x}, t)$ through a weighted sum of squares and consider as overall test statistics:

$$I_n(\mathbf{x}) = \int_{\mathcal{T}} K_n(\mathbf{x}, t) dt, \quad S_n(\mathbf{x}) = \sup_{t \in \mathcal{T}} K_n(\mathbf{x}, t).$$

What I am currently working on III

- Consider two different versions of permutation (one for the paired and one for the unpaired test statistic) to derive critical values.
- I am currently working on the proof of the permutation techniques.

Thank you for your attention!

References I

📄 Amro, Lubna and Markus Pauly (2016). "Permuting Incomplete Paired Data: A Novel Exact and Asymptotic Correct Randomization Test". In: *Journal of Statistical Computation and Simulation* 87.6, pp. 1148–1159.

📄 Bretz, Frank, Alan Genz, and Ludwig A. Hothorn (2001). "On the Numerical Availability of Multiple Comparison Procedures". In: *Biometrical Journal* 43.5, pp. 645–656.

📄 Cribari-Neto, Francisco (Mar. 2004). "Asymptotic Inference Under Heteroskedasticity of Unknown Form". In: *Computational Statistics & Data Analysis* 45.2, pp. 215–233.

📄 Eicker, F. (June 1963). "Asymptotic Normality and Consistency of the Least Squares Estimators for Families of Linear Regressions". In: *The Annals of Mathematical Statistics* 34.2, pp. 447–456.

📄 Friedrich, Sarah and Markus Pauly (May 2018). "MATS: Inference for Potentially Singular and Heteroscedastic MANOVA". In: *Journal of Multivariate Analysis* 165, pp. 166–179.

References II

Munko, Merle et al. (2024). "RMST-Based Multiple Contrast Tests in General Factorial Designs". In: *Statistics in Medicine* 43.10, pp. 1849–1866.

Smaga, Łukasz (Mar. 2020). "A Note on Repeated Measures Analysis for Functional Data". In: *AStA Advances in Statistical Analysis* 104.1, pp. 117–139.

Zimmermann, Georg, Markus Pauly, and Arne C Bathke (Dec. 2019). "Small-Sample Performance and Underlying Assumptions of a Bootstrap-Based Inference Method for a General Analysis of Covariance Model with Possibly Heteroskedastic and Nonnormal Errors". In: *Statistical Methods in Medical Research* 28.12, pp. 3808–3821.

— (May 2020). "Multivariate Analysis of Covariance with Potentially Singular Covariance Matrices and Non-Normal Responses". In: *Journal of Multivariate Analysis* 177, p. 104594.

Example: our problem of intervention research

- $k = 2$ groups and $d = 4$ dimensions.
- Estimand $\Delta\boldsymbol{\mu} = (\Delta\mu_{11}, \Delta\mu_{12}, \Delta\mu_{13}, \Delta\mu_{14}, \Delta\mu_{21}, \Delta\mu_{22}, \Delta\mu_{23}, \Delta\mu_{24})'$.
- To realize the testing problem $\mathcal{H}_{0,\ell} : \Delta\boldsymbol{\mu}_{1\ell} = \Delta\boldsymbol{\mu}_{2\ell}$, $\ell \in \{1, \ldots, 4\}$, use the matrix

$$
\mathbf{H} = \begin{pmatrix}
1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & -1
\end{pmatrix}.
$$

Example: Dunnett Testing Procedure

- $k = 3$ groups and $d = 2$ dimensions.
- Estimand $\boldsymbol{\mu} = (\mu_{11}, \mu_{12}, \mu_{21}, \mu_{22}, \mu_{31}, \mu_{32})'$.
- To realize the testing problem $\mathcal{H}_{0,s} : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_s$, $s \in \{2, 3\}$, use the matrix

$$
\mathbf{H} = \begin{pmatrix} 1 & 1 & -1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 & -1 & -1 \end{pmatrix}.
$$

Resampling Methods: Wild Bootstrap

- Generate n independent identically distributed Rademacher random variable T_{ij} ($P(T_{11} = -1) = P(T_{11} = 1) = 1/2$),
- Elements of the wild bootstrap sample:

$$\mathbf{Y}_{ij}^* := \hat{\epsilon}_{ij} \frac{T_{ij}}{\sqrt{1 - p_{ij,ij}}}.$$

- Calculate the wild bootstrap estimators: $\hat{\boldsymbol{\mu}}^*$, $\hat{\boldsymbol{\nu}}^*$ and $\hat{\mathbf{D}}^*$.
- Use them to calculate the wild bootstrap test statistic: $A_n^*(\mathbf{h}_s)$.
- $\hat{\mathbf{D}}^*$ is consistent for $\mathbf{D}$ and $A_n^*(\mathbf{h}_s)$ has asymptotically the same distribution as $A_n(\mathbf{h}_s)$ under $\mathcal{H}_{0,s}$ for all $s \in \{1, \ldots, r\}$ (Zimmermann et al., 2019, 2020).

$p_{ij,ij}$ is the ij-th diagonal element of the hat matrix $\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$, with $\mathbf{X} = (\mathbf{M}, \mathbf{Z})$.

Resampling Methods: Parametric Bootstrap

- Draw independent d-dimensional observation vectors from the normal distribution:

$$\mathbf{Y}_{ij}^{\star} \sim \mathcal{N}\left(\mathbf{0}, \hat{\mathbf{\Sigma}}_i\right),$$

  where

$$\hat{\mathbf{\Sigma}}_i = \frac{1}{n_i - c - 1} \sum_{j=1}^{n_i} \hat{\boldsymbol{\epsilon}}_{ij} \hat{\boldsymbol{\epsilon}}_{ij}'.$$

- Calculate the parametric bootstrap estimators: $\hat{\boldsymbol{\mu}}^{\star}$, $\hat{\boldsymbol{\nu}}^{\star}$ and $\hat{\mathbf{D}}^{\star}$.
- Use them to calculate the wild bootstrap test statistic: $A_n^{\star}(\mathbf{h}_s)$.
- $\hat{\mathbf{D}}^{\star}$ is consistent for $\mathbf{D}$ and $A_n^{\star}(\mathbf{h}_s)$ has asymptotically the same distribution as $A_n(\mathbf{h}_s)$ under $\mathcal{H}_{0,s}$ for all $s \in \{1, \ldots, r\}$ (Zimmermann et al., 2019, 2020).