Analysis of High Dimensional Repeated Measures Designs: The Two-Sample Statistic

M. Rauf Ahmad

Department of Medical Statistics
Center of Statistics
University of Göttingen

IBS/ROeS Seminar, Bern

September 09-13, 2007

Formulation of the problem: one sample case

- Let $\mathbf{X}_k = (X_{k1}, \ldots, X_{kd})'$ be a vector of d repeated measurements; $k = 1, \ldots, n$ independent subjects
- Assume $\mathbf{X}_k \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \boldsymbol{\Sigma} > 0$
- Aim: Test $H_0 : \mathbf{H}\boldsymbol{\mu} = \mathbf{0}$ when $d > n$;
- **H**: any appropriate hypothesis
- When $d < n$, classical multivariate procedures applicable without imposing any structure on $\boldsymbol{\Sigma}$
- Examples: Wald-type statistic, Hotelling's T^2 statistic, ...

Formulation of the problem: one sample case

- But they do not work for $d > n$, the case of high dimensional data
- For ex., T^2 totally collapses while Wald-type statistic is extremely liberal even for $n > d$
- **Idea:** The ANOVA-type statistic, based on the modification to the Box's (1954) approximation

Introduction

- Box approximation: Let $\mathbf{X} \sim \mathcal{N}_d(\mathbf{0}, \mathbf{\Sigma})$ and let $\mathbf{T}$ be any symmetric matrix. Then $Q = \mathbf{X}'\mathbf{T}\mathbf{X} \sim \sum_{i=1}^{d} \lambda_i C_i$ where λ_i are the eigenvalues of $\mathbf{T}\mathbf{\Sigma}$ and the $C_i \sim \chi_1^2$ are independent.
- The idea is to approximate the distribution of Q with that of a scaled, $g\chi_f^2$ distribution, where g and f are chosen such that the first two moments of Q and $g\chi_f^2$ coincide. (Known as Box approximation)
- Following the approximation, we solve

$$gf = \sum_i \lambda_i = \text{tr}(\mathbf{T}\mathbf{\Sigma}) \quad \text{and} \quad 2g^2 f = 2\sum_i \lambda_i^2 = 2\text{tr}(\mathbf{T}\mathbf{\Sigma})^2$$

and get $f = [\text{tr}(\mathbf{T}\mathbf{\Sigma})]^2/\text{tr}(\mathbf{T}\mathbf{\Sigma})^2$ and $g = \text{tr}(\mathbf{T}\mathbf{\Sigma})^2/\text{tr}(\mathbf{T}\mathbf{\Sigma})$

$$\text{so that} \quad \mathbf{X}'\mathbf{T}\mathbf{X}/\text{tr}(\mathbf{T}\mathbf{\Sigma}) \approx \chi_f^2/f$$

(Box, 1954; Brunner, 2001)

Introduction

- The ANOVA-type Statistic(ATS): $\mathbf{X}'\mathbf{T}\mathbf{X}/tr(\mathbf{T}\boldsymbol{\Sigma}) \sim \chi^2_f/f$
- For known $\boldsymbol{\Sigma}$, the quality of approximation is very good for $d > n$

$(n = 10)$	**Covariance Structure/**$1-\alpha$					
	CS			AR(0.6)		
d	0.90	0.95	0.99	0.90	0.95	0.99
5	0.8952	0.9498	0.9888	0.8852	0.9404	0.9848
10	0.9006	0.9480	0.9890	0.8966	0.9432	0.9856
50	0.9016	0.9526	0.9918	0.9016	0.9448	0.9848
100	0.8994	0.9454	0.9880	0.8978	0.9458	0.9874
200	0.8918	0.9432	0.9884	0.8968	0.9428	0.9860

- $\Longrightarrow$ Need appropriate estimators of f and g
- $\Longrightarrow$ Need appropriate estimators of $\text{tr}(\mathbf{T}\boldsymbol{\Sigma})$, $[\text{tr}(\mathbf{T}\boldsymbol{\Sigma})]^2$ and $\text{tr}(\mathbf{T}\boldsymbol{\Sigma})^2$

Proposed modification:

define estimators of $\text{tr}(\mathbf{T\Sigma})$, $[\text{tr}(\mathbf{T\Sigma})]^2$ and $\text{tr}(\mathbf{T\Sigma})^2$ such that they are

- unbiased, consistent and of uniform rate of convergence w.r.t. d
- First choice, classical estimator $\hat{\mathbf{\Sigma}} = \frac{1}{n-1}\sum_i (\mathbf{X} - \overline{\mathbf{X}})(\mathbf{X} - \overline{\mathbf{X}})'$?
- With $\hat{\mathbf{\Sigma}}$, estimators of $[\text{tr}(\mathbf{T\Sigma})]^2$ and $\text{tr}(\mathbf{T\Sigma})^2$ are biased; and
- consistency depends on d and gets worse with increasing d

We use Chebychev inequality $P(|X| > c) \leq \text{E}(X^2)/c^2$, $c > 0$, such that, for an estimator $\hat{\theta}_{n,d}$ of functional θ_d

$$P\left(\left|\frac{\hat{\theta}_{n,d}}{\theta_d} - 1\right| > c\right) \;\leq\; \frac{1}{c^2}\left[\text{Var}\left(\frac{\hat{\theta}_{n,d}}{\theta_d}\right) + \left\{E\left(\frac{\hat{\theta}_{n,d}}{\theta_d}\right) - 1\right\}^2\right]$$

Proposed modification:

Based on Chebychev inequality, we define estimators such that

- $E(\hat{\theta}_{n,d}/\theta_d) = 1$, for all n, and for any fixed d
- $\lim_{n\to\infty} Var(\hat{\theta}_{n,d}/\theta_d) = 0$ for any fixed d

Further, for dimensional stability, ensure that

- $Var(\hat{\theta}_{n,d}/\theta_d)$ is uniformly bounded w.r.t. $d \Longrightarrow$
- If $n \to \infty$, then, for any fixed d, $d > n$ does not destroy the quality of approximation

Introduction

Estimators: one sample case

- Model: $\mathbf{X}_k \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$; Hypothesis: $H_0 : \mathbf{H}\boldsymbol{\mu} = \mathbf{0}$
- For any $\mathbf{H}$, $\mathbf{T} = \mathbf{H}'(\mathbf{H}\mathbf{H}')^{-}\mathbf{H}$ is unique, s.t. $\mathbf{H}\boldsymbol{\mu} = \mathbf{0} \Longleftrightarrow \mathbf{T}\boldsymbol{\mu} = \mathbf{0}$
- $\mathbf{T}$ can be any general linear hypothesis, including factorial hypotheses
- For $H_0 : \mathbf{T}\boldsymbol{\mu} = \mathbf{0}$, define $\mathbf{T}\mathbf{X}_k = \mathbf{Y}_k$ such that, under H_0, $E(\mathbf{Y}_k) = \mathbf{0}, \quad \text{Var}(\mathbf{Y}_k) = \mathbf{T}\boldsymbol{\Sigma}\mathbf{T} \ = \ \mathbf{S}$.
- Define: quadratic form: $A_k = \mathbf{Y}_k'\mathbf{Y}_k$ and bilinear form: $A_{kl} = \mathbf{Y}_k'\mathbf{Y}_l$
- Estimators of $\text{tr}(\mathbf{T}\boldsymbol{\Sigma})$, $[\text{tr}(\mathbf{T}\boldsymbol{\Sigma})]^2$ and $\text{tr}(\mathbf{T}\boldsymbol{\Sigma})^2$, respectively

$$B_0 = \frac{1}{n}\sum_{k=1}^{n} A_k; \quad B_1 = \frac{1}{n(n-1)} \underbrace{\sum_{k=1}^{n}\sum_{l=1}^{n}}_{k \neq l} A_k A_l; \quad B_2 = \frac{1}{n(n-1)} \underbrace{\sum_{k=1}^{n}\sum_{l=1}^{n}}_{k \neq l} A_{kl}^2$$

- Estimators unbiased with variances of order $O\left(\frac{1}{n}\right)$

Modified ANOVA-type Statistic (ATS)

- With $Q_n = n\overline{\mathbf{X}}'\mathbf{T}\overline{\mathbf{X}}$, the ATS is defined as $\tilde{F} = Q_n/B_0$
- Using Taylor series approximation, the first two moments of $\tilde{F}$ are

$$\mathrm{E}(\tilde{F}) = 1;\ \mathrm{Var}(\tilde{F}) = \frac{2}{\tilde{f}}\left(1 - \frac{1}{n}\right)$$

- Compare with first two moments of $\chi^2_{\tilde{f}}/\tilde{f}$: 1 and $\frac{2}{\tilde{f}}$
- $\implies \tilde{F} \sim \chi^2_{\tilde{f}}/\tilde{f}$ or $\frac{Q_n}{B_0}\frac{B_1}{B_2} \sim \chi^2_{\tilde{f}}$ with $\tilde{f} = B_1/B_2$
- The approximation very accurate for n as small as 10 and any d
- For any cov. structure; special cases: CS, AR(0.2) and AR(0.6)
- Also accurate for non-normal distribution; typically evaluated for Exponential, Uniform and Bernoulli distributions
- Very high power

The Model

- Let $\mathbf{X}_{1k} = (X_{1k1}, \ldots, X_{1kd})' \sim \mathcal{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$, and
- $\mathbf{X}_{2l} = (X_{2l1}, \ldots, X_{2ld})' \sim \mathcal{N}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$
- $\mathbf{X}_{1k}$, $\mathbf{X}_{2l}$: independent vectors of d repeated measures each, so that
- $(\mathbf{X}'_{1k}, \mathbf{X}'_{2l})' \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with $\boldsymbol{\mu} = (\boldsymbol{\mu}'_1, \boldsymbol{\mu}'_2)'$, $\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_1 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_2 \end{pmatrix}$
- Corresponding sample statistics: $\overline{\mathbf{X}}_{.} = (\overline{\mathbf{X}}'_{1.}, \overline{\mathbf{X}}'_{2.})'$, $\hat{\boldsymbol{\Sigma}} = \begin{pmatrix} \hat{\boldsymbol{\Sigma}}_1 & \mathbf{0} \\ \mathbf{0} & \hat{\boldsymbol{\Sigma}}_2 \end{pmatrix}$
- No special structure assumed for $\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2$

The Hypotheses

$$
\begin{array}{llcll}
\mathbf{H}_0^{AB} & :(\mathbf{P}_2 \otimes \mathbf{P}_d)\boldsymbol{\mu} & = & \mathbf{0} & \text{(Interaction Hypothesis)} \\
\mathbf{H}_0^{A} & :(\mathbf{P}_2 \otimes \frac{1}{d}\mathbf{1}_d')\boldsymbol{\mu} & = & \mathbf{0} & \text{(Group Hypothesis)} \\
\mathbf{H}_0^{B} & :(\frac{1}{2}\mathbf{1}_2' \otimes \mathbf{P}_d)\boldsymbol{\mu} & = & \mathbf{0} & \text{(Time Hypothesis)}
\end{array}
$$

$\mathbf{P}_d = \mathbf{I}_d - \frac{1}{d}\mathbf{J}_d$: centering matrix; $\mathbf{1}$: vector of 1s

- Hypotheses of interest: $\mathbf{H}_0^{AB}$ and $\mathbf{H}_0^{B}$ $\rightarrow$ affected by d
- $\mathbf{H}_0^{A}$ is just a two-sample hypothesis (t-test); unaffected by d
- Formulation as general hypothesis: $\mathbf{T} = \mathbf{H}'\,(\mathbf{H}\mathbf{H}')^{-}\,\mathbf{H}$

$$
\mathbf{T}^{AB} \;=\; \mathbf{P}_2 \otimes \mathbf{P}_d; \quad \mathbf{T}^{B} \;=\; \tfrac{1}{2}\mathbf{J}_2' \otimes \mathbf{P}_d; \quad (\mathbf{J}_2 = \mathbf{1}_2\mathbf{1}_2')
$$

- In general, $\mathbf{T}$ can be any general linear hypothesis

The Estimators

- Define the differences $\mathbf{X}_{1k} - \mathbf{X}_{2l}$, $\forall\ k = 1, \ldots, n_1$ and $l = 1, \ldots, n_2$ so that
$$\mathbf{X}_{1k} - \mathbf{X}_{2l} \sim \mathcal{N}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2, \boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)$$
- Then for $\mathbf{H}_0^{AB} : \mathbf{T}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) = \mathbf{0}$, define $\mathbf{Y}_{1k} - \mathbf{Y}_{2l} = \mathbf{T}(\mathbf{X}_{1k} - \mathbf{X}_{2l})$, such that
$$\mathbf{Y}_{1k} - \mathbf{Y}_{2l} \sim \mathcal{N}(\mathbf{0}, \mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T})$$
- Estimator of covariance matrix, $\boldsymbol{\Sigma}_d = \mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}$, under $\mathbf{H}_0^{AB}$, is
$$\hat{\boldsymbol{\Sigma}}_d = \frac{1}{n_1 n_2} \sum_{k=1}^{n_1} \sum_{l=1}^{n_2} (\mathbf{Y}_{1k} - \mathbf{Y}_{2l})(\mathbf{Y}_{1k} - \mathbf{Y}_{2l})'$$
- $\mathrm{E}\left(\hat{\boldsymbol{\Sigma}}_d\right) = \mathbf{T}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\mathbf{T}$

The Estimators

$$\begin{aligned} A_{kl} &:= (\mathbf{Y}_{1k} - \mathbf{Y}_{2l})'(\mathbf{Y}_{1k} - \mathbf{Y}_{2l}) \qquad \text{(Quadratic Form)} \\ A_{klrs} &:= (\mathbf{Y}_{1k} - \mathbf{Y}_{2l})'(\mathbf{Y}_{1r} - \mathbf{Y}_{2s}) \qquad \text{(Bilinear Form)} \end{aligned}$$

Estimators of $\text{tr}[\mathbf{T}(\mathbf{\Sigma}_1 + \mathbf{\Sigma}_2)\mathbf{T}]$, $\left[\text{tr}[\mathbf{T}(\mathbf{\Sigma}_1 + \mathbf{\Sigma}_2)\mathbf{T}]\right]^2$ and $\text{tr}[\mathbf{T}(\mathbf{\Sigma}_1 + \mathbf{\Sigma}_2)\mathbf{T}]^2$

$$\begin{aligned} B_0 &= \frac{1}{n_1 n_2} \sum_{k=1}^{n_1} \sum_{l=1}^{n_2} A_{kl} \\ B_1 &= \frac{1}{n_1 n_2 (n_1 - 1)(n_2 - 1)} \underbrace{\sum_{k=1}^{n_1} \sum_{l=1}^{n_2} \sum_{r=1}^{n_1} \sum_{s=1}^{n_2}}_{k \neq r,\ l \neq s} A_{kl} A_{rs} \\ B_2 &= \frac{1}{n_1 n_2 (n_1 - 1)(n_2 - 1)} \underbrace{\sum_{k=1}^{n_1} \sum_{l=1}^{n_2} \sum_{r=1}^{n_1} \sum_{s=1}^{n_2}}_{k \neq r,\ l \neq s} A_{klrs}^2 \end{aligned}$$

respectively

Approximating Distribution

- Estimators B_0, B_1, B_2 unbiased and consistent
- Variances of order $O\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$
- Variances uniformly bounded with respect to d

- $Q_n = (\overline{\mathbf{X}}_{1.} - \overline{\mathbf{X}}_{2.})'\mathbf{T}(\overline{\mathbf{X}}_{1.} - \overline{\mathbf{X}}_{2.})$ be the quadratic form of the test statistic; Then, for $\mathbf{H}_0^{AB}$, the modified ANOVA-type statistic is

$$\widetilde{F} \ = \ \frac{Q_n}{B_0}$$

- First two moments of $\widetilde{F}$, based on Taylor series approximation
- Since $\text{Cov}(\overline{\mathbf{X}}_{1.} - \overline{\mathbf{X}}_{2.}) = \frac{1}{n_1}\mathbf{\Sigma}_1 + \frac{1}{n_2}\mathbf{\Sigma}_2, \quad \Longrightarrow$
- no general closed form result tractable
- Assume: either $n_1 = n_2$ or $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2$

Approximating Distribution

- Case I: $n_1 = n_2 = n$; $\mathbf{\Sigma}_1 \neq \mathbf{\Sigma}_2$
- Case II: $n_1 \neq n_2$; $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2 = \mathbf{\Sigma}$

Then

$$\mathrm{E}\left(\widetilde{F}^{AB}\right) \approx 1 \qquad \text{(Both Cases)}$$

$$\mathrm{Var}\left(\widetilde{F}^{AB}\right) \approx \frac{2}{\widetilde{\widetilde{f}}}\left(1 - \frac{1}{n}\right)^2 \qquad \text{(Case I)}$$

$$\approx \frac{2}{\widetilde{\widetilde{f}}}\left(1 - \frac{1}{n_1}\right)\left(1 - \frac{1}{n_2}\right) \qquad \text{(Case II)}$$

where $\widetilde{f} = \frac{B_1^{AB}}{B_2^{AB}}$ is the estimated degrees of freedom

- Variance of $\tilde{F}$ also uniformly bounded w. r. t. d

Approximating Distribution

For a $\frac{\chi^2_{\tilde{f}}}{\tilde{f}}$-distribution, $\mathrm{E}(\chi^2_{\tilde{f}}/\tilde{f}) = 1$, $\mathrm{Var}(\chi^2_{\tilde{f}}/\tilde{\tilde{f}}) = \frac{2}{\tilde{f}}$

Compare with the moments of the $\widetilde{F}$

$$
\begin{aligned}
\mathrm{E}\left(\widetilde{F}^{AB}\right) &\approx 1 && \text{(Both Cases)} \\
\mathrm{Var}\left(\widetilde{F}^{AB}\right) &\approx \frac{2}{\tilde{\tilde{f}}}\left(1 - \frac{1}{n}\right)^2 && \text{(Case I)} \\
&\approx \frac{2}{\tilde{\tilde{f}}}\left(1 - \frac{1}{n_1}\right)\left(1 - \frac{1}{n_2}\right) && \text{(Case II)}
\end{aligned}
$$

- $\Rightarrow$ $\widetilde{F} \sim \frac{\chi^2_f}{f}$, asymptotically
- or $F = \frac{Q_n}{B_0}\frac{B_1}{B_2} \sim \chi^2_{\tilde{f}}$ with $\tilde{f} = \frac{B_1}{B_2}$

The Time Effect

- For $\mathbf{H}_0^B : \mathbf{T}(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2) = \mathbf{0}$, define the sums $\mathbf{Y}_{1k} + \mathbf{Y}_{2l} = \mathbf{T}(\mathbf{X}_{1k} + \mathbf{X}_{2l})$
- Define B_0, B_1 and B_2, accordingly, using

$$A_{kl} = (\mathbf{Y}_{1k} + \mathbf{Y}_{2l})'(\mathbf{Y}_{1k} + \mathbf{Y}_{2l}); \quad A_{klrs} = (\mathbf{Y}_{1r} + \mathbf{Y}_{2s})'(\mathbf{Y}_{1r} + \mathbf{Y}_{2s})$$

- Quadratic form for the statistic

$$Q_n^B = (\overline{\mathbf{X}}_{1.} + \overline{\mathbf{X}}_{2.})'\mathbf{T}(\overline{\mathbf{X}}_{1.} + \overline{\mathbf{X}}_{2.})$$

- Under $\mathbf{H}_0^B$,

$$\widetilde{F} = \frac{Q_n^B}{B_0^B}$$

- Final results remain exactly the same as for the interaction effect

The Time Effect

$$
\begin{aligned}
\mathrm{E}\left(\widetilde{F}^{AB}\right) &\approx 1 && \text{(Both Cases)} \\
\mathrm{Var}\left(\widetilde{F}^{AB}\right) &\approx \frac{2}{\widetilde{f}}\left(1-\frac{1}{n}\right)^2 && \text{(Case I)} \\
&\approx \frac{2}{\widetilde{f}}\left(1-\frac{1}{n_1}\right)\left(1-\frac{1}{n_2}\right) && \text{(Case II)}
\end{aligned}
$$

- $F = \frac{Q_n}{B_0}\frac{B_1}{B_2} \sim \chi^2_{\tilde{f}}$, asymptotically, with $\widetilde{f} = \frac{B_1^B}{B_2^B}$
- Estimators B_0, B_1, B_2 unbiased and consistent with variances, also of $\tilde{F}$, uniformly bounded w.r.t. d

Simulations: Test Size

$n_1 = n_2 = n$; 5,000 runs; Compound Symmetry

		Interaction Effect			Time Effect		
n	d	0.90	0.95	0.99	0.90	0.95	0.99
10	20	0.8970	0.9490	0.9942	0.8952	0.9492	0.9936
	50	0.8928	0.9506	0.9906	0.8912	0.9476	0.9930
	100	0.8920	0.9510	0.9926	0.8948	0.9484	0.9896
	200	0.9002	0.9522	0.9920	0.8962	0.9468	0.9912
	300	0.8986	0.9506	0.9834	0.8976	0.9474	0.9904
20	50	0.9034	0.9492	0.9884	0.9020	0.9518	0.9926
	100	0.8932	0.9448	0.9898	0.9042	0.9534	0.9900
	200	0.8970	0.9466	0.9898	0.9004	0.9498	0.9932
	300	0.9074	0.9516	0.9914	0.8990	0.9462	0.9896

Simulations: Test Size

$n_1 = n_2 = n$; Autoregressive(0.2); 5,000 runs

		Interaction Effect			Time Effect		
n	d	0.90	0.95	0.99	0.90	0.95	0.99
10	20	0.9020	0.9536	0.9940	0.8992	0.9526	0.9934
	50	0.8966	0.9486	0.9922	0.8862	0.9430	0.9882
	100	0.8908	0.9450	0.9942	0.8942	0.9496	0.9914
	200	0.8970	0.9548	0.9918	0.8994	0.9506	0.9912
	300	0.8968	0.9498	0.9912	0.8904	0.9418	0.9902
20	50	0.8992	0.9468	0.9882	0.8976	0.9486	0.9916
	100	0.8928	0.9466	0.9884	0.9084	0.9538	0.9898
	200	0.8974	0.9476	0.9908	0.9000	0.9494	0.9916
	300	0.8938	0.9482	0.9896	0.8962	0.9472	0.9904

Simulations: Test Size

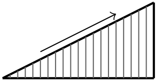
$n_1 = n_2 = n$; Autoregressive(0.6); 5,000 runs

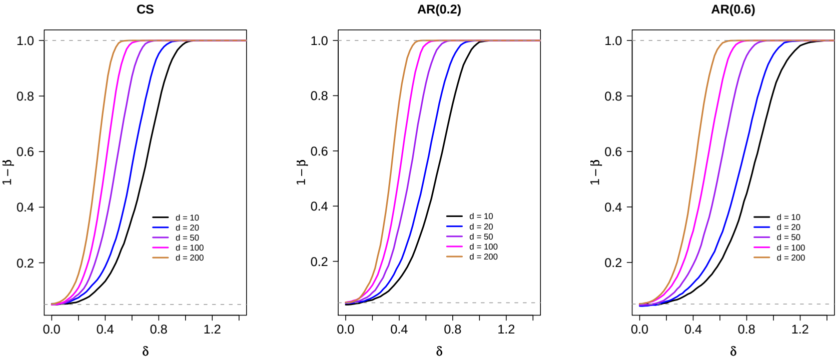
		Interaction Effect			Time Effect		
n	d	0.90	0.95	0.99	0.90	0.95	0.99
10	20	0.8950	0.9460	0.9894	0.9024	0.9522	0.9912
	50	0.8944	0.9454	0.9886	0.9008	0.9466	0.9900
	100	0.8844	0.9382	0.9896	0.8954	0.9414	0.9880
	200	0.8924	0.9472	0.9910	0.8938	0.9450	0.9902
	300	0.8900	0.9484	0.9908	0.8968	0.9544	0.9930
20	50	0.9042	0.9490	0.9906	0.8946	0.9436	0.9882
	100	0.9074	0.9524	0.9922	0.8978	0.9468	0.9882
	200	0.9026	0.9524	0.9888	0.8962	0.9430	0.9888
	300	0.8916	0.9462	0.9876	0.9020	0.9516	0.9888

Simulations: Power

Interaction effect:

$n_1 = n_2 = 10$; $\alpha = 0.05$; 5000 runs

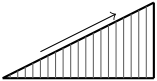


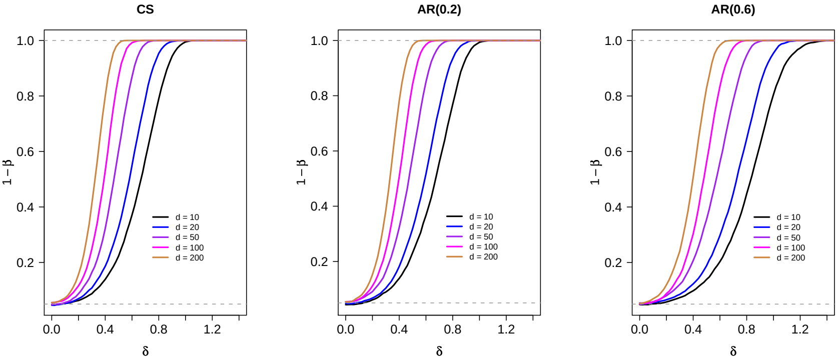


Simulations: Power

Time effect:

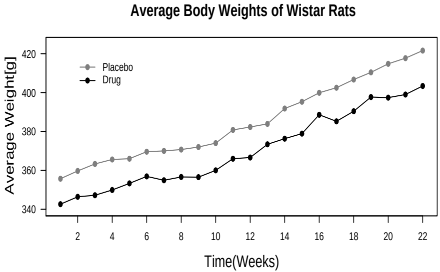
$n_1 = n_2 = 10$; $\alpha = 0.05$; 5000 runs





Example

Brunner, Domhof and Langer (2002)



- Effect of toxicity of a drug on weights of male Wistar rats
- Two independent groups, each of 10 rats
- Response Variable: Avergae body weight[g]
- Aim: Comparison of profiles of two groups over 22 weeks

Example

Analysis of weight data: $n_1 = n_2 = 10; d = 22$

- Hypothesis matrices
  - $\mathbf{T}^{AB} : \mathbf{P}_2 \otimes \mathbf{P}_{22}$
  - $\mathbf{T}^{B} : \frac{1}{2}\mathbf{J}_2 \otimes \mathbf{P}_{22}$

Effect	F	$\tilde{f}$	p-value
Interaction	0.51	3.68	0.7115
Time	9.56	1.28	0.0008

- For each effect, $F \sim \chi^2_{\tilde{f}}$

References

1. Box, G. E. P. (1954a). Some theorems on quadratic forms applied in the study of analysis of variance problems, I: Effect of inequality of variance in the one-way classification. *The Annals of Mathematical Statistics*, **25**(2), 290-302.
2. Brunner, E. (2001). Asymptotic and approximate analysis of repeated measuresd esigns under heteroscedasticity. In: *mathematical statistics with applications in biometry*, Eds.: Kunert, J. and G. Trenkler. Josef Eul Verlag, Lohmar.
3. Brunner, E., S. Domhof and F. Langer (2002). *Nonparametric analysis of longitudinal data in factorial experiments*. Wiley, N. Y.
4. Werner, C. and E. Brunner (2004). Dimensionsstabile Approximation für Verteilungen von zufälligen quadratischen Formen im Repeated Measures Design. Technical Report, Abt. Med. Stat., Universität Göttingen.