Equivalence of regression curves

Frank Bretz (Novartis)

Based on: Dette H, Möllenhoff K, Volgushev S, Bretz F (2017)
Equivalence of regression curves. *Journal of the American Statistical Association* (in press)

Workshop on Design of Healthcare Studies
Singapore – 2017, July 05

FP7 HEALTH 2013 - 602552

Motivation

- An application: Populations of different geographic regions may bear differences in efficacy (or safety) dose response
  $\longrightarrow$ Objective: Ability to extrapolate study results
  $\longrightarrow$ Demonstrating similarity of curves becomes an issue
- Another application: Comparison of dose response relationships for two regimens
  $\longrightarrow$ For example, demonstrate that once-daily (o.d.) and twice-daily (b.i.d.) applications of a drug are similar over a given dose range
- Yet another application: Comparison of different drugs containing the same active substance in order to claim bioequivalence.
  $\longrightarrow$ Traditional approaches based on AUC or Cmax may be misleading
  $\longrightarrow$ Objective: Develop a test which takes the whole curve into account
- IDEAL project: Focus on small population clinical trials (e.g. rare diseases)
  $\longrightarrow$ Methodology should work for small sample sizes

Motivation

- An application: Populations of different geographic regions may bear differences in efficacy (or safety) dose response
  $\longrightarrow$ Objective: Ability to extrapolate study results
  $\longrightarrow$ Demonstrating similarity of curves becomes an issue
- Another application: Comparison of dose response relationships for two regimens
  $\longrightarrow$ For example, demonstrate that once-daily (o.d.) and twice-daily (b.i.d.) applications of a drug are similar over a given dose range
- Yet another application: Comparison of different drugs containing the same active substance in order to claim bioequivalence.
  $\longrightarrow$ Traditional approaches based on AUC or Cmax may be misleading
  $\longrightarrow$ Objective: Develop a test which takes the whole curve into account
- IDEAL project: Focus on small population clinical trials (e.g. rare diseases)
  $\longrightarrow$ Methodology should work for small sample sizes

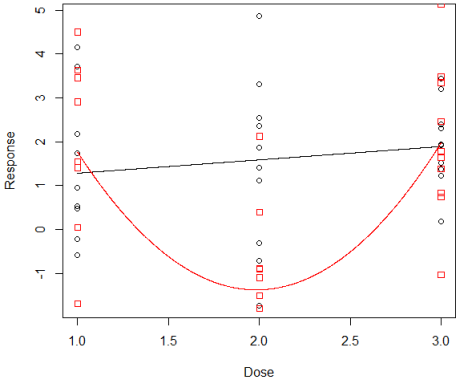
Motivation

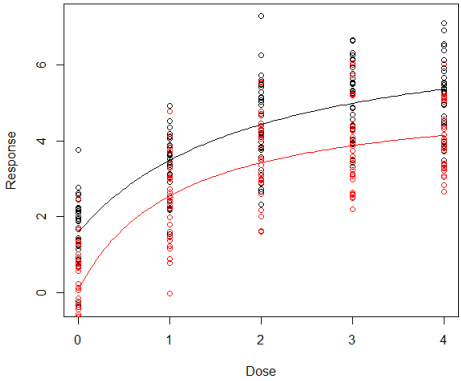
- An application: Populations of different geographic regions may bear differences in efficacy (or safety) dose response
  $\longrightarrow$ Objective: Ability to extrapolate study results
  $\longrightarrow$ Demonstrating similarity of curves becomes an issue
- Another application: Comparison of dose response relationships for two regimens
  $\longrightarrow$ For example, demonstrate that once-daily (o.d.) and twice-daily (b.i.d.) applications of a drug are similar over a given dose range
- Yet another application: Comparison of different drugs containing the same active substance in order to claim bioequivalence.
  $\longrightarrow$ Traditional approaches based on AUC or Cmax may be misleading
  $\longrightarrow$ Objective: Develop a test which takes the whole curve into account
- IDEAL project: Focus on small population clinical trials (e.g. rare diseases)
  $\longrightarrow$ Methodology should work for small sample sizes

Motivation

- An application: Populations of different geographic regions may bear differences in efficacy (or safety) dose response
  $\longrightarrow$ Objective: Ability to extrapolate study results
  $\longrightarrow$ Demonstrating similarity of curves becomes an issue
- Another application: Comparison of dose response relationships for two regimens
  $\longrightarrow$ For example, demonstrate that once-daily (o.d.) and twice-daily (b.i.d.) applications of a drug are similar over a given dose range
- Yet another application: Comparison of different drugs containing the same active substance in order to claim bioequivalence.
  $\longrightarrow$ Traditional approaches based on AUC or Cmax may be misleading
  $\longrightarrow$ Objective: Develop a test which takes the whole curve into account
- IDEAL project: Focus on small population clinical trials (e.g. rare diseases)
  $\longrightarrow$ Methodology should work for small sample sizes

Comparing curves - The setting





Comparing curves - The setting II

- Two dose response curves (from two samples)

$$Y_{\ell,i,j} = m_\ell(x_{\ell,i}, \vartheta_\ell) + \varepsilon_{\ell,i,j} \quad (j = 1, \ldots, n_{\ell,i}; i = 1, \ldots, k_\ell; \ell = 1, 2)$$

  - $\varepsilon_{\ell,i,j}$ independent $\sim \mathcal{N}(0, \sigma_\ell^2)$ $\quad (\ell = 1, 2)$
  - $x_{\ell,i} \in \mathcal{X}$ (dose levels)
  - $\vartheta_\ell \in \mathbb{R}^{d_\ell}$ (parameter in model m_ℓ)

Problem: Are the dose response curves m_1 and m_2 similar?

Comparing curves - The setting II

- Two dose response curves (from two samples)

$$Y_{\ell,i,j} = m_\ell(x_{\ell,i}, \vartheta_\ell) + \varepsilon_{\ell,i,j} \quad (j = 1, \ldots, n_{\ell,i}; i = 1, \ldots, k_\ell; \ell = 1, 2)$$

  - $\varepsilon_{\ell,i,j}$ independent $\sim \mathcal{N}(0, \sigma_\ell^2)$ $\quad (\ell = 1, 2)$
  - $x_{\ell,i} \in \mathcal{X}$ (dose levels)
  - $\vartheta_\ell \in \mathbb{R}^{d_\ell}$ (parameter in model m_ℓ)

Problem: Are the dose response curves m_1 and m_2 similar?

Comparing curves - The setting II

- Two dose response curves (from two samples)

$$Y_{\ell,i,j} = m_\ell(x_{\ell,i}, \vartheta_\ell) + \varepsilon_{\ell,i,j} \quad (j = 1, \ldots, n_{\ell,i}; i = 1, \ldots, k_\ell; \ell = 1, 2)$$

  - $\varepsilon_{\ell,i,j}$ independent $\sim \mathcal{N}(0, \sigma_\ell^2)$ $\quad (\ell = 1, 2)$
  - $x_{\ell,i} \in \mathcal{X}$ (dose levels)
  - $\vartheta_\ell \in \mathbb{R}^{d_\ell}$ (parameter in model m_ℓ)

Problem: Are the dose response curves m_1 and m_2 similar?

Measures of similarity

- d: a *metric* measuring the distance between m_1 and m_2.
- Hypothesis of similarity:

$$H_0 : d(m_1, m_2) \geq \epsilon \quad \text{versus} \quad H_1 : d(m_1, m_2) < \epsilon$$

(here ϵ is a pre-specified constant).

- Examples
  - maximum absolute difference

$$d_\infty(m_1, m_2) = \max_{x \in \mathcal{X}} |m_1(x, \vartheta_1) - m_2(x, \vartheta_2)|$$

  - squared L_2-distance

$$d_2(m_1, m_2) = \int_{\mathcal{X}} (m_1(x, \vartheta_1) - m_2(x, \vartheta_2))^2 dx$$

Measures of similarity

- d: a *metric* measuring the distance between m_1 and m_2.
- Hypothesis of similarity:

$$H_0 : d(m_1, m_2) \geq \epsilon \quad \text{versus} \quad H_1 : d(m_1, m_2) < \epsilon$$

(here ϵ is a pre-specified constant).

- Examples
  - maximum absolute difference

$$d_\infty(m_1, m_2) = \max_{x \in \mathcal{X}} |m_1(x, \vartheta_1) - m_2(x, \vartheta_2)|$$

  - squared L_2-distance

$$d_2(m_1, m_2) = \int_{\mathcal{X}} (m_1(x, \vartheta_1) - m_2(x, \vartheta_2))^2 dx$$

Efficient tests for similarity

Tests based on distances

Basic idea: Estimate the distance between m_1 and m_2 directly and decide for similarity for small values of the resulting estimate

- (parametric) estimates of m_1 and m_2:

$$\hat{m}_1 = m_1(\cdot, \hat{\vartheta}_1), \;\; \hat{m}_2 = m_2(\cdot, \hat{\vartheta}_2)$$

- estimate of the distance between m_1 and m_2:

$$\hat{\mathbf{d}} = d(\hat{m}_1, \hat{m}_2)$$

Tests based on distances

$\hat{\mathbf{d}}$: estimate the distance between m_1 and m_2

- Decide for similarity, i.e. reject the hypothesis

$$H_0 : d(m_1, m_2) \geq \epsilon$$

whenever

$$\hat{\mathbf{d}} = d(\hat{m}_1, \hat{m}_2) < q$$

- **Problem:** how do we find the critical value q?

Estimating the distance between two curves – the squared L_2-distance

- Define an estimate by

$$\hat{\mathbf{d}}_{\mathbf{2}} = d_2(\hat{m}_1, \hat{m}_2) = \int_{\mathcal{X}} (m_1(x, \hat{\vartheta}_1) - m_2(x, \hat{\vartheta}_2))^2 dx$$

- Empirical process theory: $\left\{ m_1(x, \hat{\vartheta}_1) - m_2(x, \hat{\vartheta}_2) - (m_1(x, \vartheta_1) - m_2(x, \vartheta_2)) \right\}_{x \in \mathcal{X}}$ converges weakly to a centered Gaussian process (as $n_1, n_2 \to \infty$)

$$\longrightarrow d_2(\hat{m}_1, \hat{m}_2) - d_2(m_1, m_2) \overset{a}{\sim} \mathcal{N}\Big(0, \frac{\tau^2}{n_1 + n_2}\Big)$$

where

$$\tau^2 = \int_{\mathcal{X} \times \mathcal{X}} (m_1(x, \vartheta_1) - m_2(x, \vartheta_2)) \cdot (m_1(y, \vartheta_1) - m_2(y, \vartheta_2)) k(x, y) dx\, dy$$

Estimating the distance between two curves – the squared L_2-distance II

An asymptotic test for similarity of two dose response curves:

- If $\hat{\tau}^2$ is an estimate of τ^2, then

$$H_0 : d_2(m_1, m_2) \geq \epsilon_2$$

is rejected, whenever

$$d_2(\hat{m}_1, \hat{m}_2) < \epsilon_2 + \frac{\hat{\tau}}{\sqrt{n_1 + n_2}} u_\alpha$$

- **It is very difficult to estimate τ^2, especially for small sample sizes**

Estimating the distance between two curves – the maximum distance

- Define an estimate by

$$\hat{\mathbf{d}}_\infty = d_\infty(\hat{m}_1, \hat{m}_2) = \max_{x \in \mathcal{X}} |m_1(x, \hat{\vartheta}_1) - m_2(x, \hat{\vartheta}_2)|$$

- If the true absolute difference curve has only one extremal point, we have

$$d_\infty(\hat{m}_1, \hat{m}_2) - d_\infty(m_1, m_2) \overset{a}{\sim} \mathcal{N}\Big(0, \frac{\sigma^2_{d_\infty}}{n_1 + n_2}\Big),$$

  where $\sigma^2_{d_\infty}$ depends on the location of the extremal point
- Otherwise $d_\infty(\hat{m}_1, \hat{m}_2)$ is not asymptotically normal distributed
- **An asymptotic test is not always available!**

Bootstrap test for similarity

Generate data under the **null hypothesis**

- Estimate ϑ_1, ϑ_2 under the restriction of the null hypothesis, that is the estimates fulfill $d(m_1, m_2) = \epsilon$

  $\rightsquigarrow$ **constrained** estimates $\hat{\hat{\vartheta}}_1, \hat{\hat{\vartheta}}_2$

  e.g. for $d = d_2$, $\hat{\hat{\vartheta}}_1, \hat{\hat{\vartheta}}_2$ satisfy $\int_{\mathcal{X}} (m_1(x, \hat{\hat{\vartheta}}_1) - m_2(x, \hat{\hat{\vartheta}}_2))^2 dx = \epsilon_2$
- Generate bootstrap data (parametric bootstrap)

$$Y^*_{\ell,i,j} = m_\ell(x_{\ell,i}, \hat{\hat{\vartheta}}_\ell) + \hat{\sigma}_\ell \varepsilon^*_{\ell,i,j}$$

where $\varepsilon^*_{\ell,i,j}$ are i.i.d. $\sim \mathcal{N}(0,1)$

Bootstrap test for similarity

Generate data under the **null hypothesis**

- Estimate ϑ_1, ϑ_2 under the restriction of the null hypothesis, that is the estimates fulfill $d(m_1, m_2) = \epsilon$

  $\rightsquigarrow$ **constrained** estimates $\hat{\hat{\vartheta}}_1, \hat{\hat{\vartheta}}_2$

  e.g. for $d = d_2$, $\hat{\hat{\vartheta}}_1, \hat{\hat{\vartheta}}_2$ satisfy $\int_{\mathcal{X}} (m_1(x, \hat{\hat{\vartheta}}_1) - m_2(x, \hat{\hat{\vartheta}}_2))^2 dx = \epsilon_2$
- Generate bootstrap data (parametric bootstrap)

$$Y^*_{\ell,i,j} = m_\ell(x_{\ell,i}, \hat{\hat{\vartheta}}_\ell) + \hat{\sigma}_\ell \varepsilon^*_{\ell,i,j}$$

where $\varepsilon^*_{\ell,i,j}$ are i.i.d. $\sim \mathcal{N}(0, 1)$

Bootstrap test for similarity II

Bootstrap test:

- Calculate $\hat{\vartheta}_1^*, \hat{\vartheta}_2^*$ from the bootstrap data $Y_{\ell,i,j}^*$
- Calculate $\hat{d}^* = d(\hat{m}_1^*, \hat{m}_2^*)$, for $d = d_2$ that is

$$\hat{\mathbf{d}}_2^* = \int_{\mathcal{X}} (m_1(x, \hat{\vartheta}_1^*) - m_2(x, \hat{\vartheta}_2^*))^2 dx$$

- Repeat this procedure B times
- If $\hat{\mathbf{d}}^{*(\mathbf{1})}, \ldots, \hat{\mathbf{d}}^{*(\mathbf{B})}$ denote the ordered bootstrap replicates the hypothesis

$$H_0 : d(m_1, m_2) \geq \epsilon$$

is rejected, whenever

$$\hat{\mathbf{d}} < \hat{\mathbf{d}}^{*(\lfloor \mathbf{B}\alpha \rfloor)}$$

Bootstrap test for similarity III

- Theoretical properties:
  - the bootstrap test has asymptotic level α
  - the bootstrap test is consistent
- More precisely: for increasing sample sizes $n_1, n_2 \to \infty$

$$\begin{aligned} d(m_1, m_2) = \epsilon &\rightsquigarrow \quad \mathbb{P}(\text{“rejection”}) \to \alpha \\ d(m_1, m_2) > \epsilon &\rightsquigarrow \quad \mathbb{P}(\text{“rejection”}) \to 0 \\ d(m_1, m_2) < \epsilon &\rightsquigarrow \quad \mathbb{P}(\text{“rejection”}) \to 1 \end{aligned}$$

- In case of $d = d_\infty$ and the true difference curve having more than one extremal point the bootstrap test is still valid but more conservative

Finite sample properties

- $m_1(x, \vartheta_1) = \delta + \frac{5x}{x+1}, \qquad m_2(x, \vartheta_2) = \frac{5x}{1+x}, \qquad \mathcal{X} = [0, 4]$

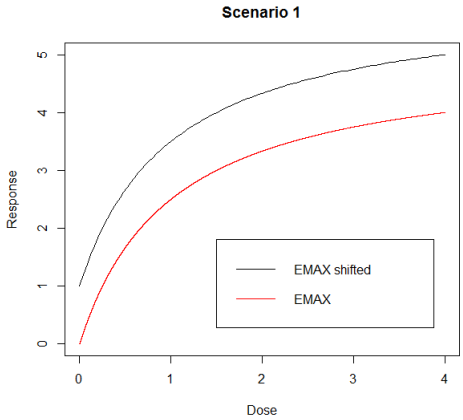


Figure: Two shifted EMAX models.

The d_2 bootstrap test – simulation of Type I error rates

- $m_1(x, \vartheta_1) = \delta + \frac{5x}{x+1}, \qquad m_2(x, \vartheta_2) = \frac{5x}{1+x}, \qquad \mathcal{X} = [0, 4]$
- $H_0 : d_2(m_1, m_2) \geq 1$ versus $H_1 : d_2(m_1, m_2) < 1$

			$\alpha = 0.05$			$\alpha = 0.1$		
			(σ_1^2, σ_2^2)			(σ_1^2, σ_2^2)		
(n_1, n_2)	δ	d_2	$(0.25, 0.25)$	$(0.5, 0.5)$	$(0.25, 0.5)$	$(0.25, 0.25)$	$(0.5, 0.5)$	$(0.25, 0.5)$
$(10, 10)$	1	4	0.000	0.000	0.000	0.000	0.000	0.000
$(10, 10)$	0.75	2.25	0.004	0.002	0.001	0.000	0.002	0.000
$(10, 10)$	0.5	1	0.051	0.064	0.052	0.101	0.120	0.118
$(20, 20)$	1	4	0.000	0.000	0.000	0.000	0.000	0.000
$(20, 20)$	0.75	2.25	0.001	0.002	0.000	0.004	0.005	0.001
$(20, 20)$	0.5	1	0.057	0.058	0.050	0.125	0.107	0.097
$(50, 50)$	1	4	0.000	0.000	0.000	0.000	0.000	0.000
$(50, 50)$	0.75	2.25	0.001	0.000	0.000	0.002	0.000	0.000
$(50, 50)$	0.5	1	0.057	0.048	0.054	0.097	0.114	0.093

Table: Simulated Type I error rates of the d_2-bootstrap test (1000 simulations, $B = 300$).

The d_2 bootstrap test – simulation of Type I error rates

- $m_1(x, \vartheta_1) = \delta + \frac{5x}{x+1}, \qquad m_2(x, \vartheta_2) = \frac{5x}{1+x}, \qquad \mathcal{X} = [0, 4]$
- $H_0 : d_2(m_1, m_2) \geq 1$ versus $H_1 : d_2(m_1, m_2) < 1$

			$\alpha = 0.05$			$\alpha = 0.1$		
			(σ_1^2, σ_2^2)			(σ_1^2, σ_2^2)		
(n_1, n_2)	δ	d_2	$(0.25, 0.25)$	$(0.5, 0.5)$	$(0.25, 0.5)$	$(0.25, 0.25)$	$(0.5, 0.5)$	$(0.25, 0.5)$
$(10, 10)$	1	4	0.000	0.000	0.000	0.000	0.000	0.000
$(10, 10)$	0.75	2.25	0.004	0.002	0.001	0.000	0.002	0.000
$(10, 10)$	0.5	1	0.051	0.064	0.052	0.101	0.120	0.118
$(20, 20)$	1	4	0.000	0.000	0.000	0.000	0.000	0.000
$(20, 20)$	0.75	2.25	0.001	0.002	0.000	0.004	0.005	0.001
$(20, 20)$	0.5	1	0.057	0.058	0.050	0.125	0.107	0.097
$(50, 50)$	1	4	0.000	0.000	0.000	0.000	0.000	0.000
$(50, 50)$	0.75	2.25	0.001	0.000	0.000	0.002	0.000	0.000
$(50, 50)$	0.5	1	0.057	0.048	0.054	0.097	0.114	0.093

Table: Simulated Type I error rates of the d_2-bootstrap test (1000 simulations, $B = 300$).

The d_2 bootstrap test – simulation of power

			$\alpha = 0.05$			$\alpha = 0.1$		
			(σ_1^2, σ_2^2)			(σ_1^2, σ_2^2)		
(n_1, n_2)	δ	d_2	$(0.25, 0.25)$	$(0.5, 0.5)$	$(0.25, 0.5)$	$(0.25, 0.25)$	$(0.5, 0.5)$	$(0.25, 0.5)$
$(10, 10)$	0.25	0.25	0.210	0.118	0.134	0.300	0.212	0.256
$(10, 10)$	0.1	0.04	0.294	0.132	0.186	0.427	0.250	0.312
$(10, 10)$	0	0	0.351	0.145	0.176	0.467	0.286	0.340
$(20, 20)$	0.25	0.25	0.392	0.171	0.225	0.534	0.302	0.382
$(20, 20)$	0.1	0.04	0.560	0.308	0.418	0.720	0.460	0.562
$(20, 20)$	0	0	0.610	0.314	0.390	0.757	0.462	0.555
$(50, 50)$	0.25	0.25	0.724	0.460	0.554	0.825	0.595	0.825
$(50, 50)$	0.1	0.04	0.961	0.691	0.821	0.982	0.824	0.973
$(50, 50)$	0	0	0.984	0.734	0.865	0.998	0.861	0.999

Table: Simulated power of the d_2-bootstrap test (1000 simulations, $B = 300$).

The d_∞ bootstrap test – simulation of Type I error rates

		$\alpha = 0.05$			$\alpha = 0.1$		
		(σ_1^2, σ_2^2)			(σ_1^2, σ_2^2)		
(n_1, n_2)	$d = d_\infty$	$(0.25, 0.25)$	$(0.5, 0.5)$	$(0.25, 0.5)$	$(0.25, 0.25)$	$(0.5, 0.5)$	$(0.25, 0.5)$
$(10, 10)$	1	0.000	0.004	0.001	0.007	0.019	0.010
$(10, 10)$	0.75	0.000	0.008	0.006	0.013	0.041	0.020
$(10, 10)$	0.5	0.015	0.040	0.016	0.050	0.104	0.054
$(20, 20)$	1	0.000	0.000	0.000	0.000	0.004	0.006
$(20, 20)$	0.75	0.000	0.002	0.000	0.003	0.010	0.002
$(20, 20)$	0.5	0.006	0.019	0.016	0.027	0.051	0.046
$(50, 50)$	1	0.000	0.000	0.000	0.000	0.000	0.001
$(50, 50)$	0.75	0.006	0.000	0.000	0.004	0.007	0.002
$(50, 50)$	0.5	0.003	0.005	0.004	0.018	0.027	0.034

Table: Simulated Type I error rates of the d_∞-bootstrap test for $\epsilon_\infty = 0.5$ (1000 simulations, $B = 300$).

The d_∞ bootstrap test – simulation of power

		$\alpha = 0.05$			$\alpha = 0.1$		
		(σ_1^2, σ_2^2)			(σ_1^2, σ_2^2)		
(n_1, n_2)	$d = d_\infty$	$(0.25, 0.25)$	$(0.5, 0.5)$	$(0.25, 0.5)$	$(0.25, 0.25)$	$(0.5, 0.5)$	$(0.25, 0.5)$
$(10, 10)$	0.25	0.062	0.050	0.053	0.147	0.118	0.118
$(10, 10)$	0.1	0.100	0.070	0.099	0.195	0.137	0.190
$(10, 10)$	0	0.109	0.090	0.092	0.216	0.143	0.176
$(20, 20)$	0.25	0.085	0.060	0.076	0.171	0.134	0.162
$(20, 20)$	0.1	0.158	0.090	0.112	0.309	0.184	0.220
$(20, 20)$	0	0.178	0.108	0.120	0.324	0.209	0.219
$(50, 50)$	0.25	0.162	0.086	0.098	0.283	0.178	0.218
$(50, 50)$	0.1	0.390	0.212	0.232	0.568	0.349	0.398
$(50, 50)$	0	0.457	0.211	0.266	0.630	0.363	0.438

Table: Simulated power of the d_∞-bootstrap test for $\epsilon_\infty = 0.5$ (1000 simulations, $B = 300$).

A case study – IBS data set

- Biesheuvel, E. and Hothorn, L. A. : Female and male patients with Irritable Bowel Syndrome (IBS), $n = 369$
- randomized to one of the five doses 0 (placebo), 1, 2, 3, and 4
- larger values corresponding to a better treatment effect
- fitted models
  - male: $m_1(x, \beta_1) = 0.398 + 0.043x$
  - female: $m_2(x, \beta_2) = 0.220 + 0.517 \frac{x}{1.396+x}$
- maximum distance: 0.1784

A case study – IBS data set II

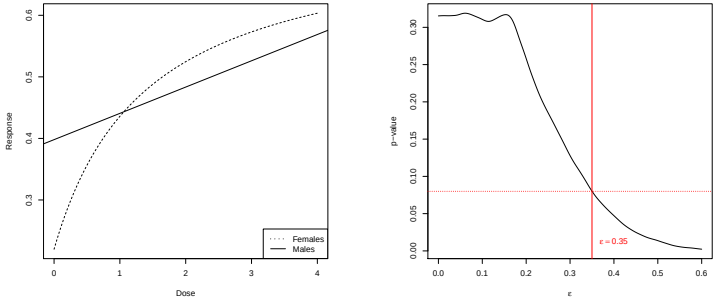


Figure: Left: Fitted dose response curves for male (linear model) and female (Emax model) patients. Right: p-values of the d_∞- bootstrap test for different values of the threshold ϵ_∞.

- the p-value corresponding to the choice $\epsilon_\infty = 0.35$ is given by 0.078

Further results and extensions

- the bootstrap d_∞-test was implemented in the `R` package `TestingSimilarity`
- we investigated the performance of all tests under the assumption of model misspecification and observed a robust performance
- the proposed tests were adapted to
  - the case of multiple curve comparison
  - dependent data
  - models with common parameters
- especially the adaption to dependent data offers a variety of new applications, e.g. the demonstration of bioequivalence

Further results and extensions II

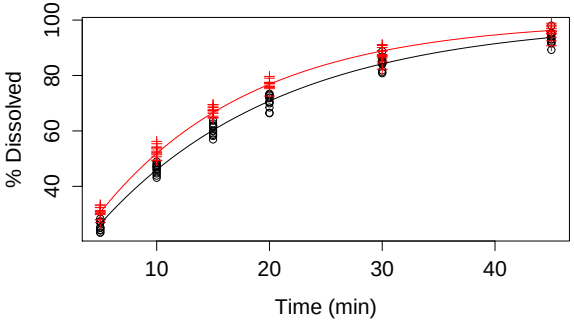


Figure: *Comparison of concentration profiles.*

Conclusions

- Two powerful tests have been proposed by
  - estimating the distance and the variance of the test statistic directly (asymptotic approach)
  - generating quantiles by a parametric bootstrap
- the bootstrap approach can be applied to any metric without deriving the asympotic distribution
- the tests are robust against misspecification of the functional form and can also be applied to dependent data

Conclusions

- Two powerful tests have been proposed by
  - estimating the distance and the variance of the test statistic directly (asymptotic approach)
  - generating quantiles by a parametric bootstrap
- the bootstrap approach can be applied to any metric without deriving the asympotic distribution
- the tests are robust against misspecification of the functional form and can also be applied to dependent data

Conclusions

- Two powerful tests have been proposed by
  - estimating the distance and the variance of the test statistic directly (asymptotic approach)
  - generating quantiles by a parametric bootstrap
- the bootstrap approach can be applied to any metric without deriving the asympotic distribution
- the tests are robust against misspecification of the functional form and can also be applied to dependent data

Acknowledgments

Thank you very much!

The research is embedded in the IDeAl EU-FP7 project, Grant-Agreement No. 602 552

References:

- Biesheuvel, E. and Hothorn, L. A. (2002). Many-to-one comparisons in stratified designs, Biometrical Journal, 44, 101–116
- Bjoern Bornkamp, Jose Pinheiro and Frank Bretz (2010) DoseFinding: Planning and Analyzing Dose Finding experiments. R package version 0.1-1. Available at http://cran.r-project.org/web/packages/DoseFinding/index.html
- Moellenhoff, K. (2015) TestingSimilarity: Bootstrap Test for Similarity of Dose Response Curves Concerning the Maximum Absolute Deviation. R package version 1.0, available at http://cran.r-project.org/web/packages/TestingSimilarity/index.html

Further References

References:

- Bretz, Moellenhoff, Dette, Liu, Trampisch (2016) Assessing the similarity of dose response and target doses in two non-overlapping subgroups. *Statistics in Medicine* (in press). Previous version available at arXiv:1607.05424
- Dette, Moellenhoff, Volgushev, Bretz (2016) Equivalence of regression curves. *Journal of the American Statistical Association* (in press). Previous version available at arXiv:1505.05266
- Moellenhoff, Dette, Bretz (2018) Equivalence of regression curves sharing common parameters. Manuscript in preparation.