Assessing the Similarity of Bioanalytical Methods

Jason Liao, Ph.D.
Merck Research Laboratories

Joint work with Yu Tian

2010 MBSW

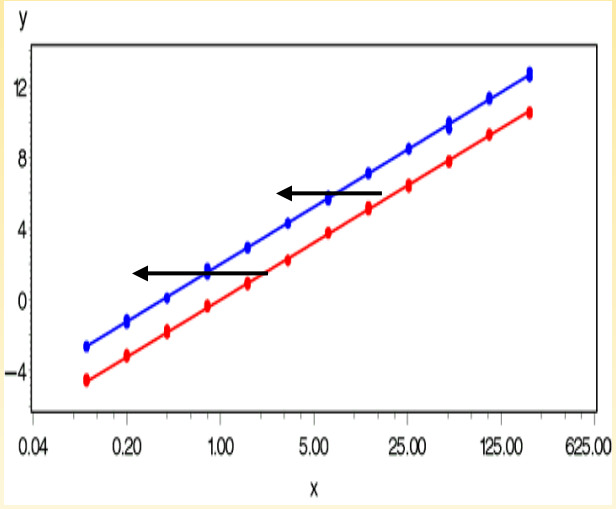
Outline

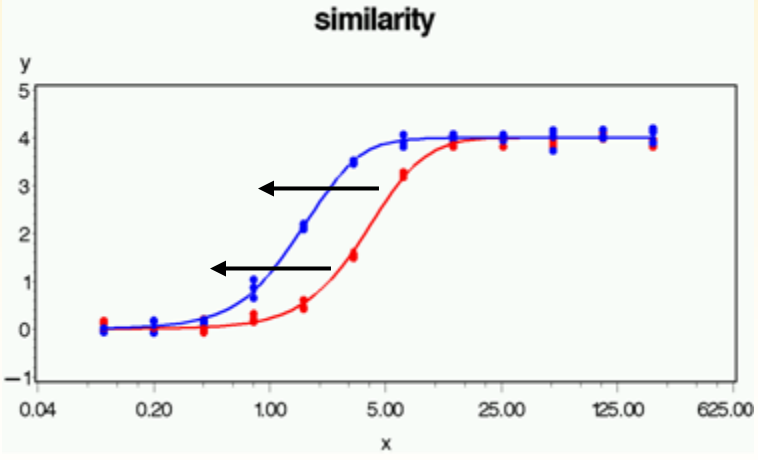
- Introduction
- Existing methods & potential problems
- New method
  - Linear
  - Nonlinear
- An Example
- Summary

Introduction

- Why Similarity
  - A key assumption
  - RP

- Definition of similarity
  - Mathematically : **f (x)=g (px)**
  - Parallelism
  - P: relative potency.





Introduction (cont.)

- Target: Sufficiently Similar → Single RP

- Assessing the Similarity
  - Mathematically.

  - The assessment of the degree of similarity is very tricky between two sparse, noisy sets of non-linear dose response data sets.

  - No universal strategy

  - Two kinds of existing method:
    (1) Significance Test
    (2) Equivalence Test.

Significance Test

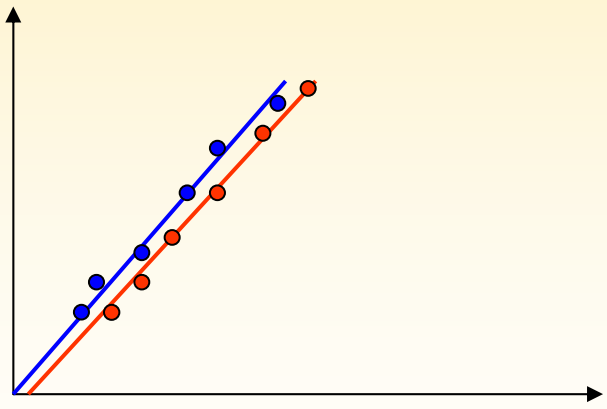
(linear case: Similarity⇔ Same Slopes)

- Significance Test

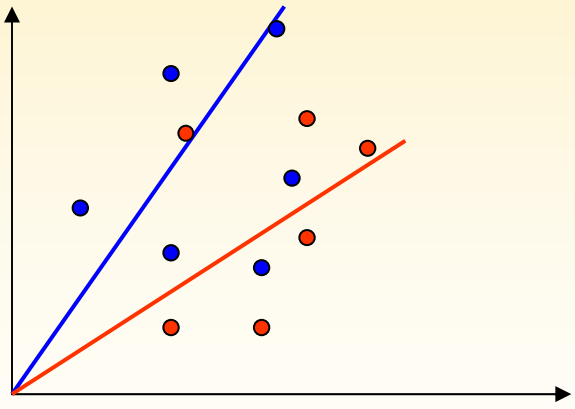
  Ho: Two slopes are exactly same.

  Ha: Two slopes are different;

When the precision increases….



Lab A: Good Precision
→ Not similar



Lab B: Poor Precision
→Similar

Equivalence Test

(linear case: Similarity= Same Slopes)

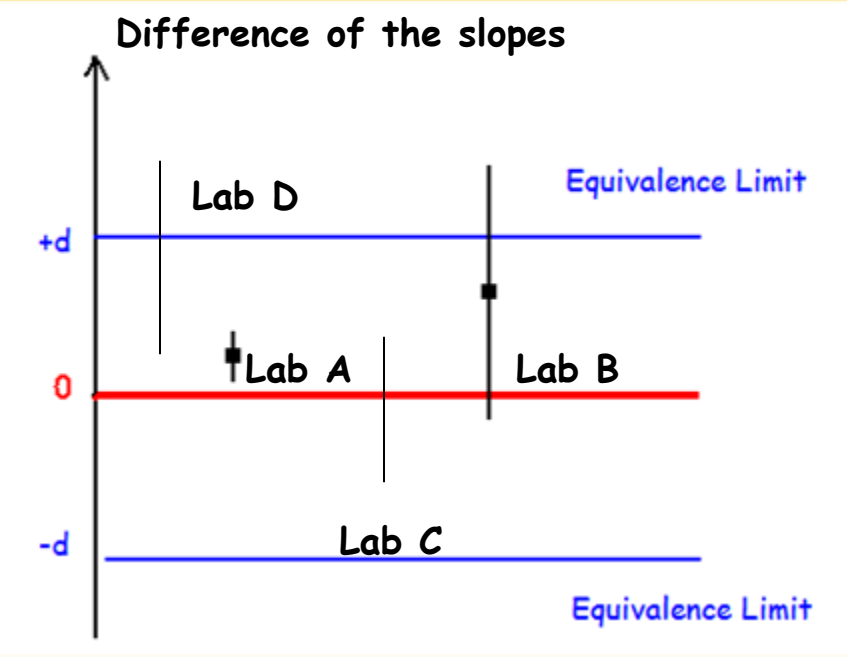
- Equivalence test

Ho: |difference| >=D Ha: |difference| <D

- The equivalence limits define differences between test and standard preparations that are considered unimportant.
  - Step1: CI for difference
  - Step2: CI vs. Equivalence limit

Equivalence Test

(linear case: Similarity= Same Slopes)



	Lab A	Lab B	Lab C	Lab D
Significance test	Not similar	Similar	Similar	Not Similar
Equivalence Test	Similar	Not Similar	Similar	Not Similar

Equivalence Test

(linear case: Similarity= Same Slopes)

- Fixed equivalence limit (Not vary with application)
  - **[0.8,1.25]** for the ratio of the slopes

- Capability based equivalence limit (Tolerance limit)
  - Manage the rate at which we falsely detect non-similarity
  - Can be assessed by evaluating reference material relative to itself.
    1. For complete reference data
    2. Pair them in all possible combination
    3. set up CI….
    4. use the most extreme boundary.

Equivalence Test

(linear case: Similarity= Same Slopes)

- DE (dilution effect)

$$DE=100\%(2^{1-bs/br}-1)<=20\%$$

  - A statistic called dilution effect was introduced in the industry to assess dilution similarity. (Schofield T. 2000).The dilution effect is a measure of the percent bias per 2-fold dilution in a test samples' value relative to that of the reference standard.
  - The absolute value of dilution effect less than 20% has been used in the industry to conclude dilution similarity (parallelism) between the test sample and the reference standard
  - Can lead to different conclusions if switching reference and sample positions.
  - Subset data easily lead to different conclusions: 4-fold dilution subset data vs. the 2-fold dilution whole data set

Equivalence Test

(Non-linear case)

- The data may not be linear…
- Equivalence test for the 5PL logistic function (Liao & Liu, 2009),

$$E(Y) = D + \frac{A - D}{\left[1 + \left(2^{1/g} - 1\right) \cdot \left(\frac{X}{C}\right)^{B}\right]^{g}}$$

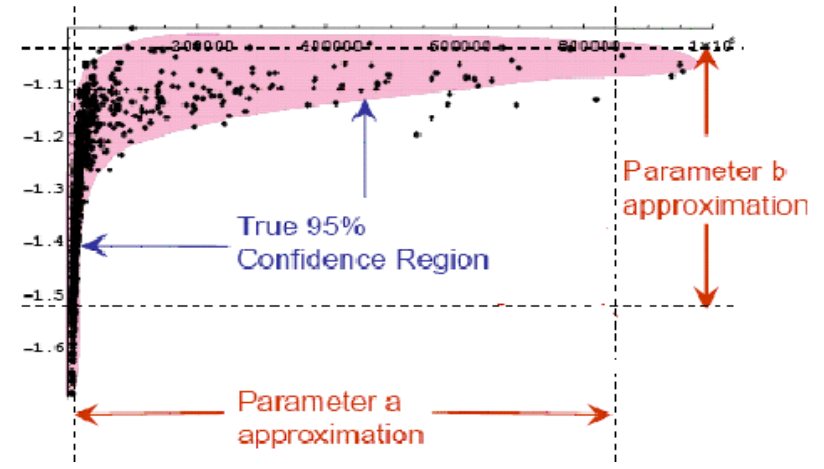
- It has been shown that 5PL curve1 can be parallel to 5PL curve2 if and only if A1= A2, B1= B2, D1= D2 , g1= g2 and C1= P•C2, where P is relative potency.

Equivalence Test (Non-linear case)

- Equivalence Test +4 parameters
- IUT: Intersection Union Test (Tian, et al. 2007, Jonkman & Sidik, 2009)
  - Ho: |A1-A2|>=a, |B1- B2|>=b, |D1- D2|>=d , |g1- g2|>=g .
  - Ha: |A1-A2|<a, |B1- B2|< b, |D1- D2|<d , |g1- g2|<g

"If the model is nonlinear, the contours are sometimes banana-shaped… sometimes stretch to infinity"
Drapper and Smith, Applied Regression Analysis p.543

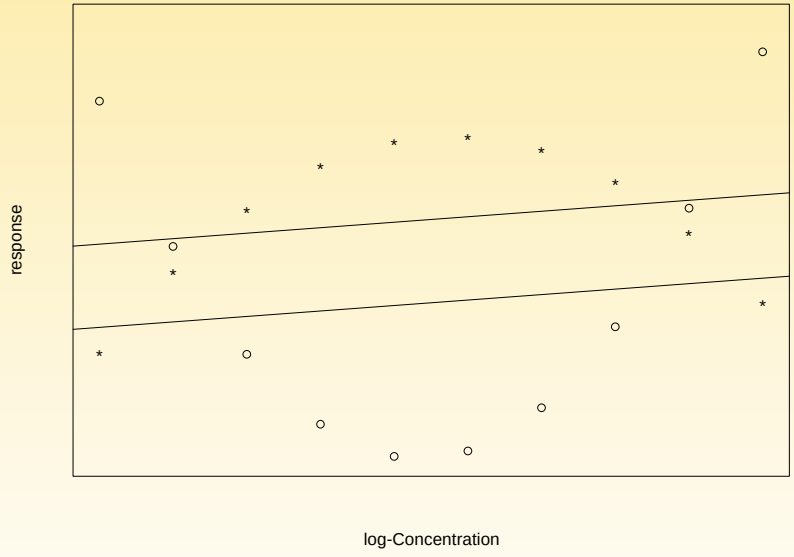
- Multidimensional confidence regions.
  - Intersection Union region misleading
  - Nearly impossible to do on 4 pars
  - Bootstrap may help



Joint Confidence Regions for 4PL,
Paul G.Gottschalk, Measuring Parallelism in Nonlinear Bioassays and Immunoassays

New Method

- Equivalence approach
- Overall difference of the response: shape of the curves



- Avoid the difficulty for obtaining the boundary/region for multi-parameters
- Boundary based on reference vs reference
  - Plus practical consideration

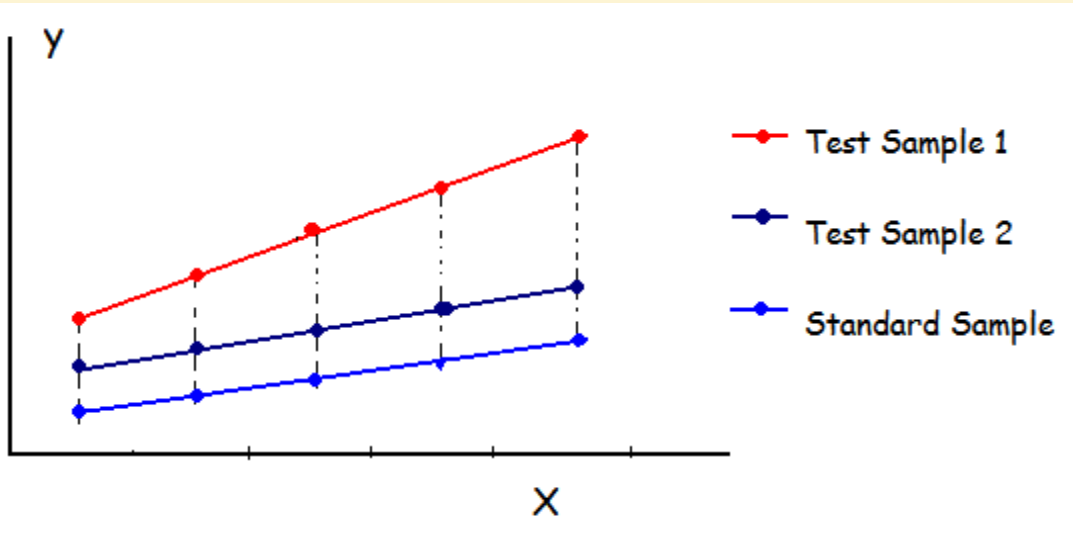
Linear Case

New Method: J-method

- J- method: $Y_{1i} - Y_{2i} \sim A_i + \omega_i \qquad \omega_i \sim N(0, \tau^2)$

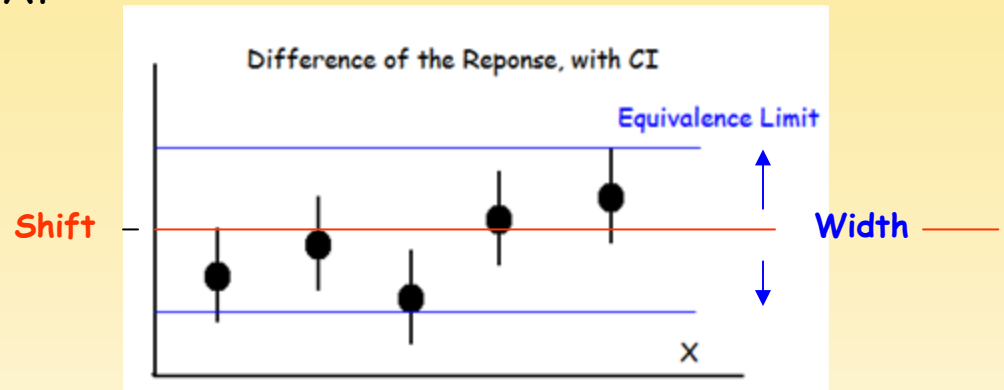
- Parallel ⇔ $A_1 = A_2 = ... = A_I = a_1 + b_1 x - (a_2 + b_2 x) = a_1 - a_2$

Considering variation….



New Method: J-method (cont.)

- Estimate the CI for Ai



- Construct Equivalence Limit

(1) Width-- Variation σ^2 that are considered to be acceptable Standard curve compare to itself (Tolerance Limit)

(2) Shift— $U = a_1 - a_2$

(3) The interval: $[U_L - 1.96 \cdot \sqrt{2\hat{\sigma}_u^2}, U_U + 1.96 \cdot \sqrt{2\hat{\sigma}_u^2}]$

New Method: J-method modification (cont.)

- Deal with large variation

$$\log(y_1) - \log(y_2) = \log(y_1/y_2)$$

- Control the width of equivalence limit
  - 2-fold difference boundary

$$y_1/y_2 <= 2 \rightarrow \log(y_1) - \log(y_2) <= \log(2) \approx 0.7$$

$$y_1/y_2 >= 0.5 \rightarrow \log(y_1) - \log(y_2) >= \log(0.5) \approx -0.7$$

  - 3-fold difference boundary

$$y_1/y_2 <= 3 \rightarrow \log(y_1) - \log(y_2) <= \log(3) \approx 1.1$$

$$y_1/y_2 >= 1/3 \rightarrow \log(y_1) - \log(y_2) >= \log(1/3) \approx -1.1$$

Advantages of the new method

- Capture the non-linear trend
- Capture the non-homogenous variance
  - Easy to modify for non-homogenous variance
- Control the rate of falsely declaiming non-similarity
  - With practical subject knowledge consideration

Simulation Study

- Linear Case
  - F
  - DE
  - J, J2, J3

Simulation Settings

Test Sample $Y_i = a_1 + b_1 X_i + \varepsilon_i$

Standard Sample $Y_i = a_2 + b_2 X_i + \delta_i$

$\varepsilon_i \sim N(0, \sigma^2) \quad \delta_i \sim N(0, \sigma^2)$

$X_i = \log(1, 2, 4, 8, 16)$ 2 fold dilution (16 to 1)

$a_1 = a_2 = 1$

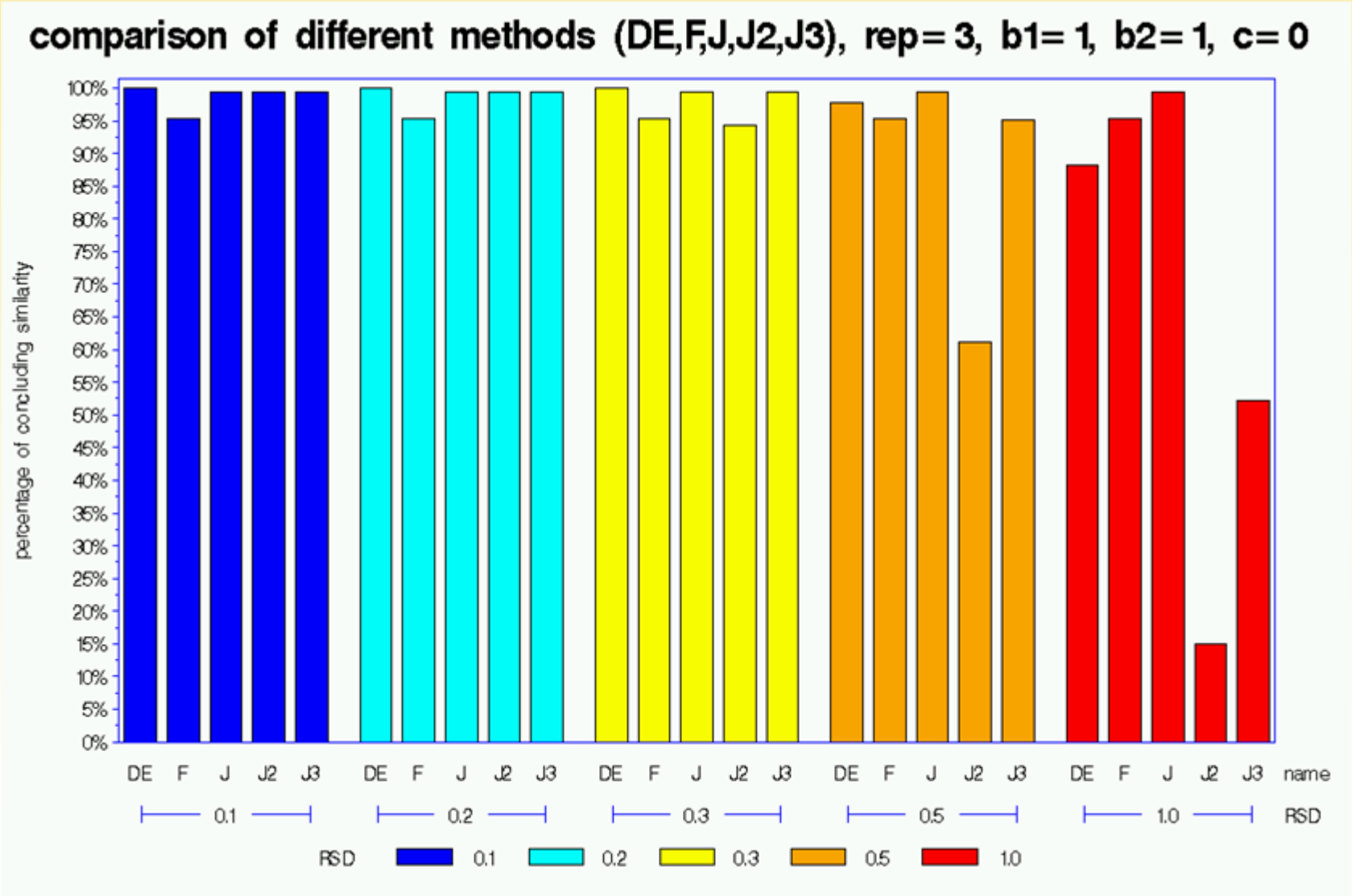
$b_2 = 1$

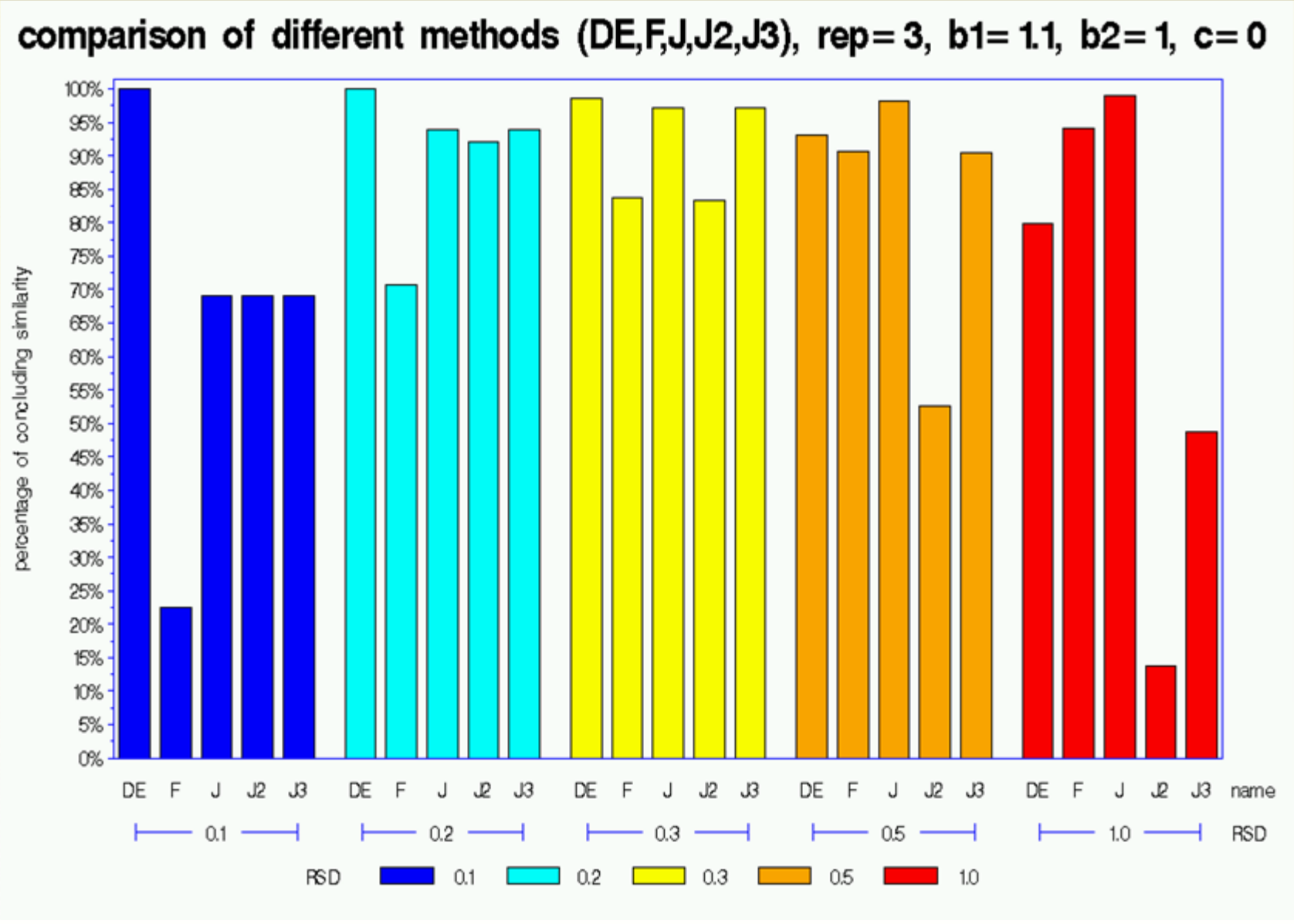
$b_1 = 1,\ 1.01,\ 1.1,\ 1.3,\ 1.5,\ 2$

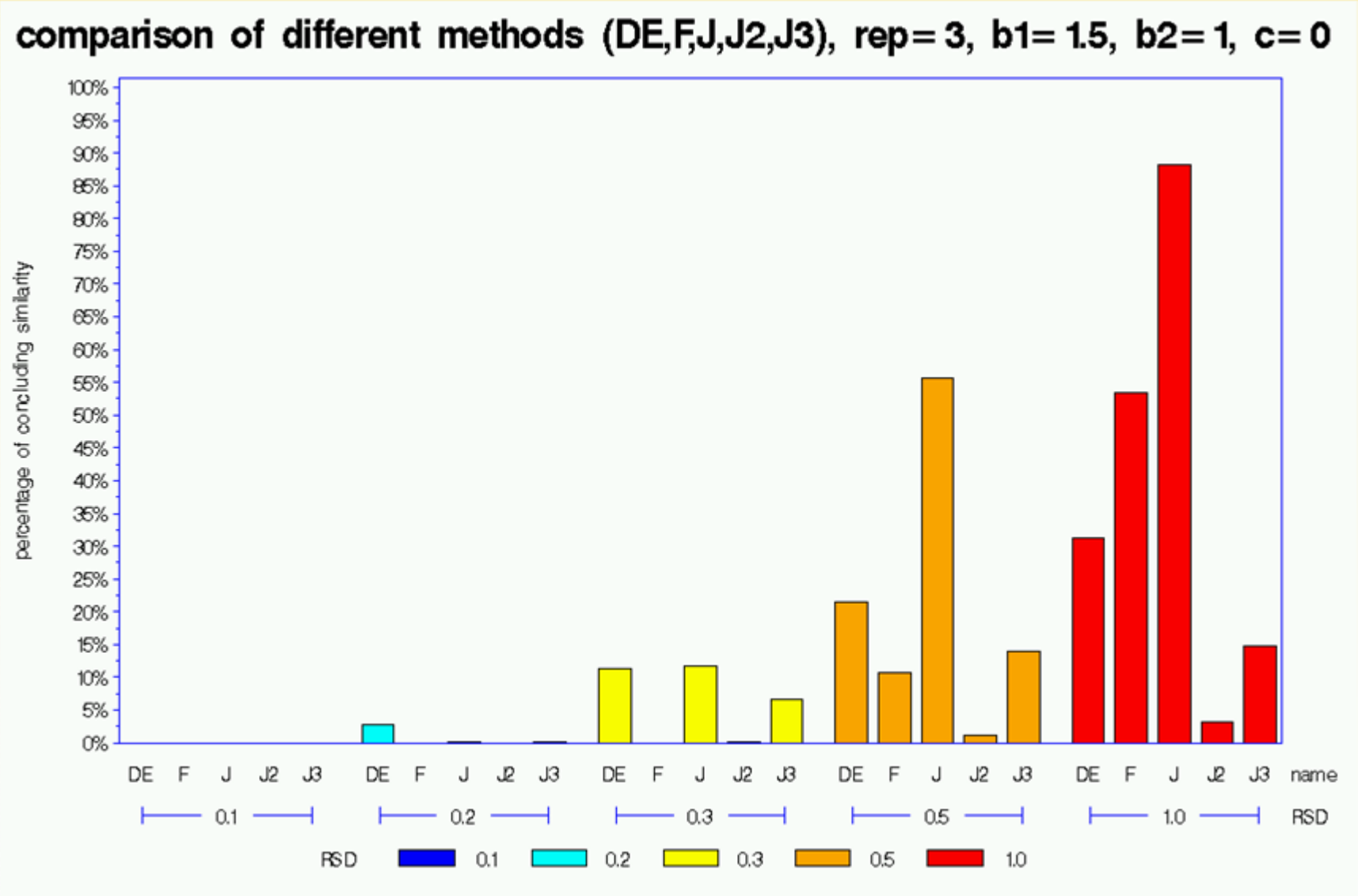
$\text{RSD} = 5\%, 10\%, 20\%, 30\%, 50\%, 100\%$

3 Replicates

numbers of simulation $= 1000$ for each parameter setting


comparison of different methods (DE,F,J,J2,J3), rep= 3, b1= 1, b2= 1, c= 0
percentage of concluding similarity
100%
95%
90%
85%
80%
75%
70%
65%
60%
55%
50%
45%
40%
35%
30%
25%
20%
15%
10%
5%
0%
DE F J J2 J3 DE F J J2 J3 DE F J J2 J3 DE F J J2 J3 DE F J J2 J3
name
0.1
0.2
0.3
0.5
1.0
RSD
RSD 0.1 0.2 0.3 0.5 1.0


comparison of different methods (DE,F,J,J2,J3), rep= 3, b1= 1.1, b2= 1, c= 0
percentage of concluding similarity
100%
95%
90%
85%
80%
75%
70%
65%
60%
55%
50%
45%
40%
35%
30%
25%
20%
15%
10%
5%
0%
DE F J J2 J3 DE F J J2 J3 DE F J J2 J3 DE F J J2 J3 DE F J J2 J3
name
0.1
0.2
0.3
0.5
1.0
RSD
RSD 0.1 0.2 0.3 0.5 1.0


comparison of different methods (DE,F,J,J2,J3), rep=3, b1=1.5, b2=1, c=0
percentage of concluding similarity
100%
95%
90%
85%
80%
75%
70%
65%
60%
55%
50%
45%
40%
35%
30%
25%
20%
15%
10%
5%
0%
DE F J J2 J3 DE F J J2 J3 DE F J J2 J3 DE F J J2 J3 DE F J J2 J3
name
0.1 0.2 0.3 0.5 1.0
RSD
RSD 0.1 0.2 0.3 0.5 1.0

Summary

- F: variance ↗, tends to conclude parallel

- DE: variance ↗, may not be able to get the right conclusion

- J, J2, J3: J3 works well, even when variance ↗;
  - Catch the non-linear trend
  - Catch the non-homogenous variance
  - Control the rate of false non-similarity

- New method is based on reference vs reference capability with possible practically significant boundary

Non-linear Case

New Method

- Similar idea with linear case: work on the difference of response directly instead of the curve parameters using equivalence approach
  - Avoid high dimension issues/complicated parameter region
  - Using scientific judgement about the boundary
  - Can add practical boundary into account

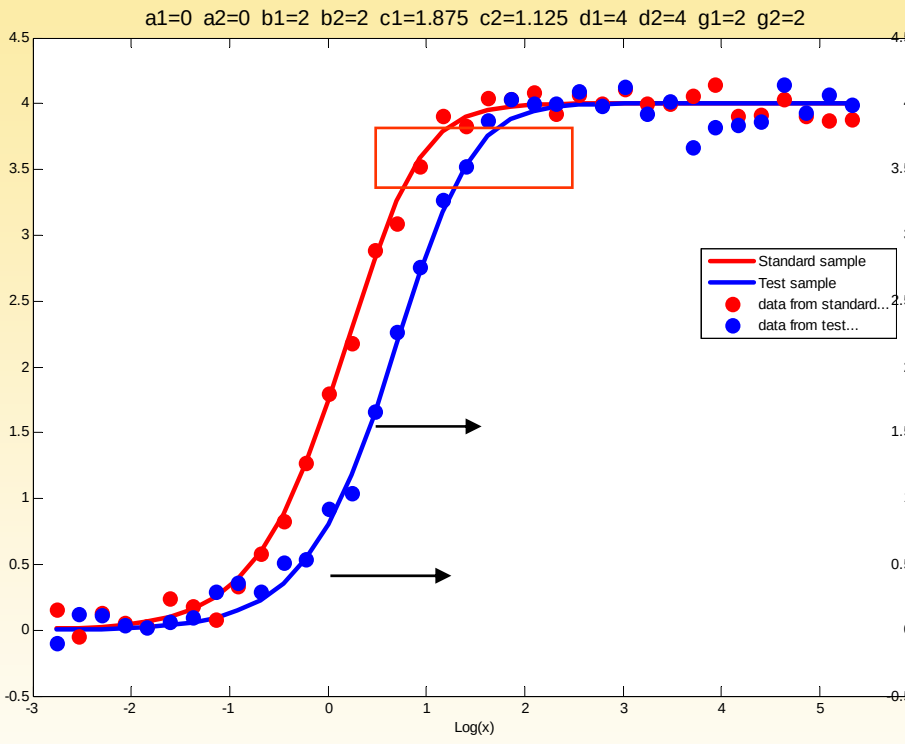
- However, the method cannot be applied directly on the original curve. Need to wok on the transformed curve
  - Similarity:

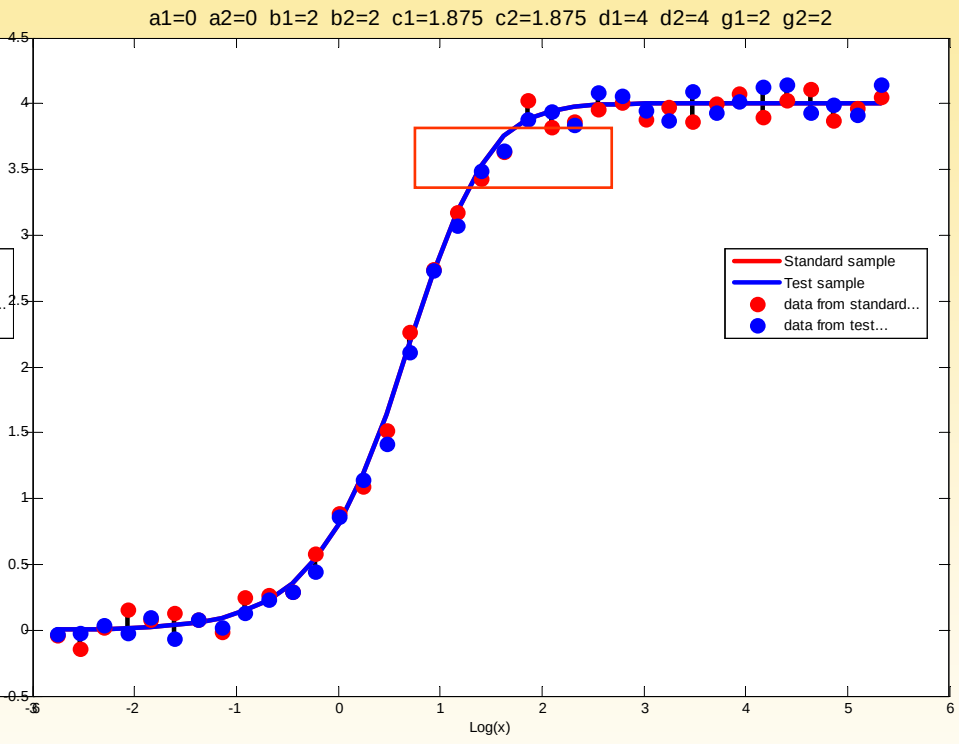
$$\mathbf{f_1(x)=f_2(px)} \longleftrightarrow \mathbf{f_1(x)-f_2(px)=0} \text{ for any x.}$$

- Evaluate the distance of two samples after transformation:

  Metric: $\mathbf{||f_1(x)-f_2(px)||}$

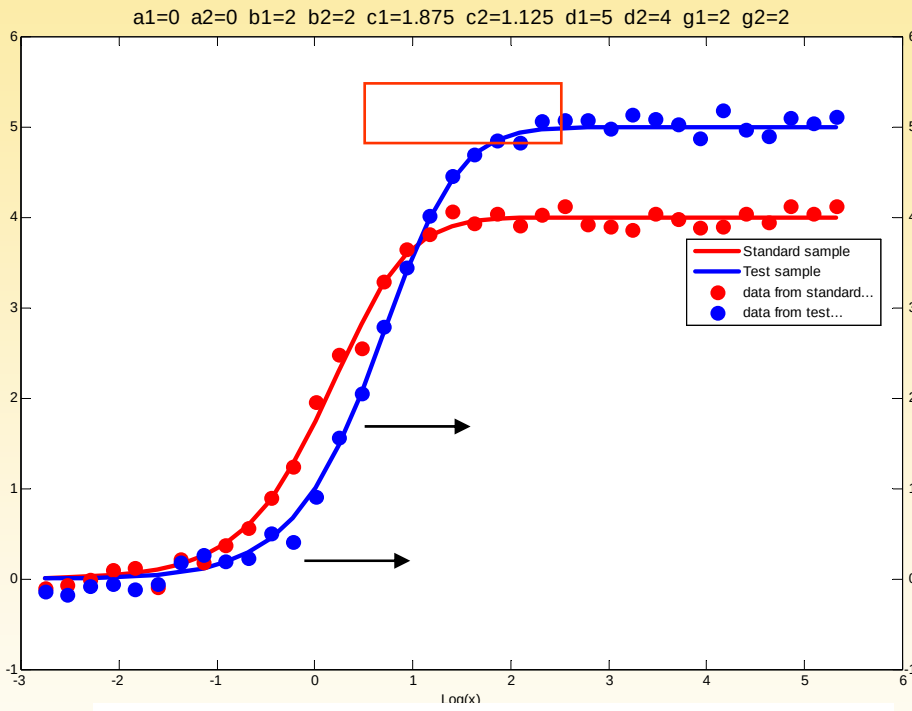
  - Not the one $\mathbf{||f_1(x)-f_2(x)||}$ (which is a constant in linear case, but not in a non-linear case)



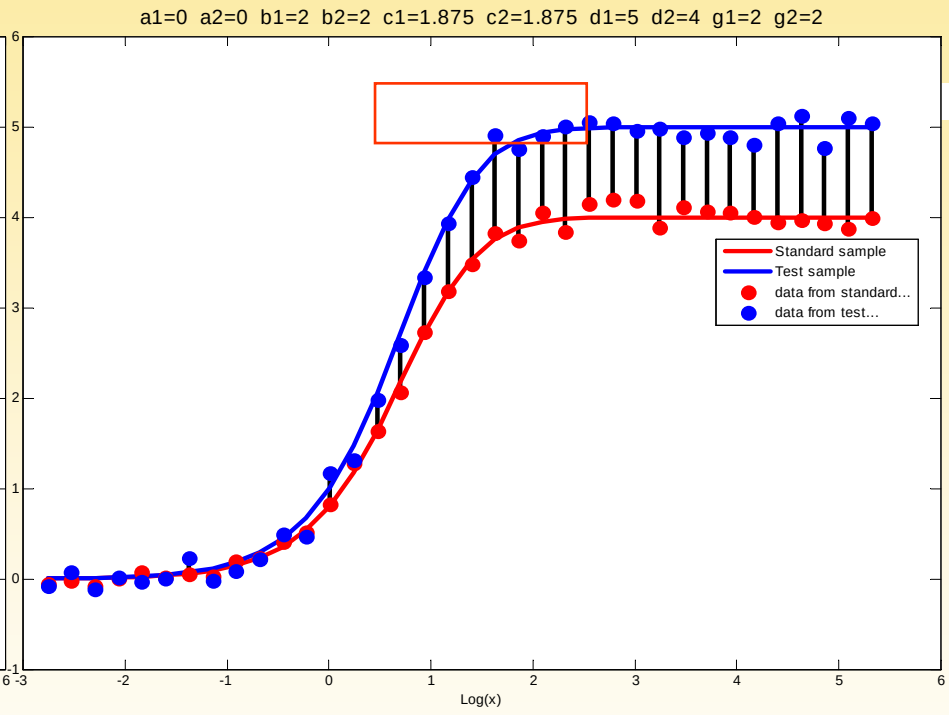


Before Transformation

After Transformation



Before Transformation



After Transformation

Math for transformed curve

Test Sample: $\hat{y}_{1i} = y_{1i}$

Estimation of the transformed standard sample:

$$\hat{y}_{2i} = f_2(\hat{p}x) + \hat{\delta}_i \qquad \text{where } \hat{P} = \hat{C}_2/\hat{C}_1$$

$$= \hat{D}_2 + \frac{\hat{A}_2 - \hat{D}_2}{\left[1 + \left(2^{1/\hat{g}_2} - 1\right)\cdot\left(\hat{p}\,\frac{X}{\hat{C}_2}\right)^{\hat{B}_2}\right]^{\hat{g}_2}} + \hat{\delta}_i$$

$$= \hat{D}_2 + \frac{\hat{A}_2 - \hat{D}_2}{\left[1 + \left(2^{1/\hat{g}_2} - 1\right)\cdot\left(\frac{X}{\hat{C}_1}\right)^{\hat{B}_2}\right]^{\hat{g}_2}} + \hat{\delta}_i$$

New Methods

- A fair measure of $||f_1(x)-f_2(px)||$
- Method I: J-method
- Method II: Yu-method $\left\|\hat{y}_1(x) - \hat{y}_2(x)\right\|_2 = \sum_{i=1}^{n} (\hat{y}_{1i} - \hat{y}_{2i})^2$
  - Tolerance Limit
  - → $\left\|\hat{y}_1(x) - \hat{y}_2(x)\right\|_2 = \sum_{i=1}^{n} (\hat{y}_{1i} - \hat{y}_{2i})^2 \quad < \quad 2n\hat{\sigma}^2$

Simulation Study

- Nonlinear Case
  - AIC, AIC-c
  - F
  - DE
  - J, J2, J3
  - Yu

Nonlinear: 5PL model

- 5PL (Liao and Liu, 2009):

$$Y = D + \frac{A - D}{\left[1 + \left(2^{1/g} - 1\right) \cdot \left(\frac{X}{C}\right)^{B}\right]^{g}}$$

- Var(y_i)=constant , RSD=0.1,0.5
- a_1 b_1 c_1 g_1 d_1 for the test sample
  a_2 b_2 c_2 g_2 $d_2$for the standard sample
- "standard" and "test" are simulated with RP=0.8
- 2-fold dilutions from 204.8 to 0.1, total 12 dilution levels
- Triplicate observations
- Number of simulation=500

Parameter settings

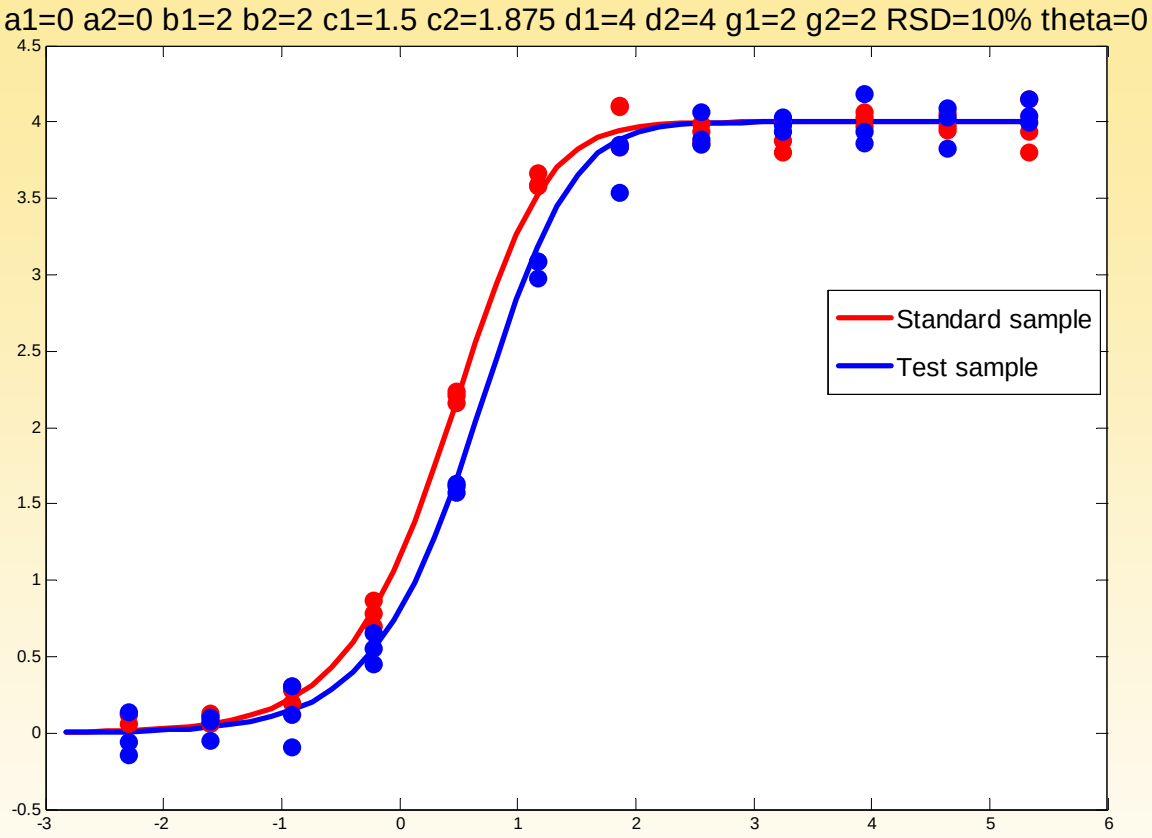
	A	B	C	D	G	
Standard	0	2	1.5	4	1	
Test	0	2	1.875	4	1	
Standard	0	2	1.5	4	2	
Test	0	2	1.875	4	2	
Standard	0	2	1.5	4	1	A
Test	0.2	2	1.875	4	1	
Standard	0	2	1.5	4	2	
Test	0.2	2	1.875	4	2	
Standard	0	2	1.5	4	1	
Test	1	2	1.875	4	1	
Standard	0	2	1.5	4	2	
Test	1	2	1.875	4	2	

	A	B	C	D	G	
Standard	0	2	1.5	4	1	D
Test	0	2	1.875	4.2	1	
Standard	0	2	1.5	4	2	
Test	0	2	1.875	4.2	2	
Standard	0	2	1.5	4	1	
Test	0	2	1.875	4.5	1	
Standard	0	2	1.5	4	2	
Test	0	2	1.875	4.5	2	

	A	B	C	D	G	
Standard	0	2	1.5	4	1	B
Test	0	2.2	1.875	4	1	
Standard	0	2	1.5	4	2	
Test	0	2.2	1.875	4	2	
Standard	0	2	1.5	4	1	
Test	0	3	1.875	4	1	
Standard	0	2	1.5	4	2	
Test	0	3	1.875	4	2	
Standard	0	2	1.5	4	1	G
Test	0	2	1.875	4	2	
Standard	0	2	1.5	4	1.2	
Test	0	2	1.875	4	2	

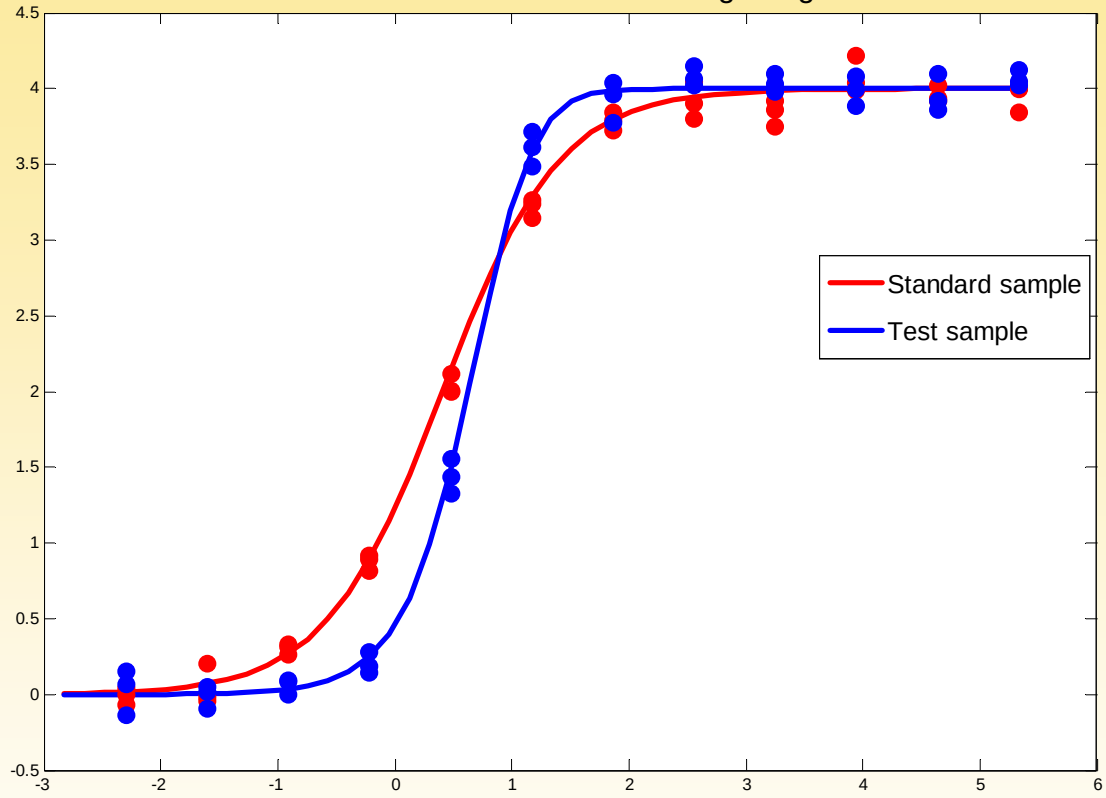
	A	B	C	D	G	
Standard	0	2	1.5	4	2	B.G
Test	0	3	1.875	4	1	
Standard	0	2	1.5	4	1	
Test	0	3	1.875	4	2	

RSD=10%
RSD=50%



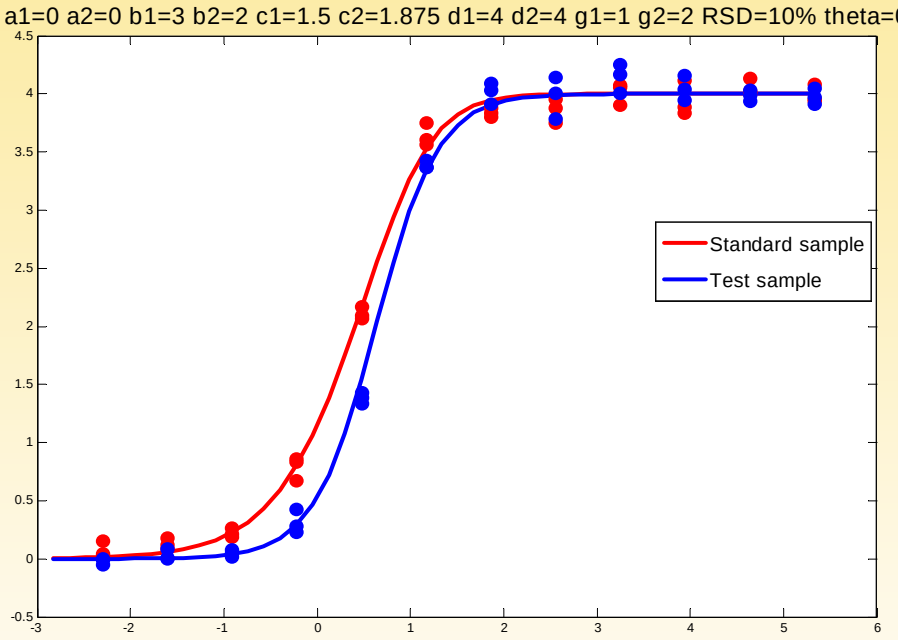
Method	J	J2	J3	F	DE	AIC	AIC-c	Yu
Percent	95.4%	95.4%	95.4%	94.6%	99.4%	82.8%	94.4%	100%

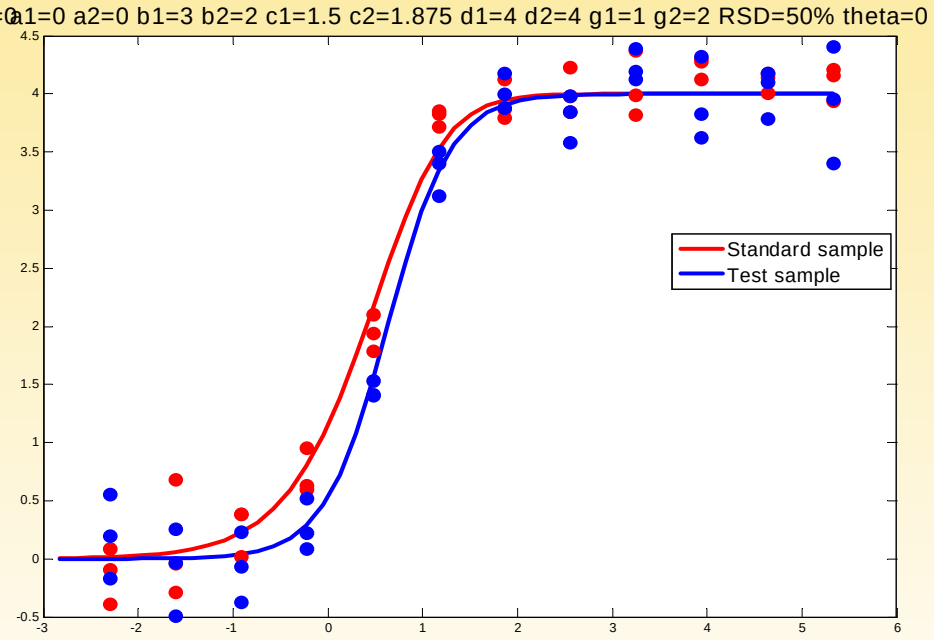
a1=0 a2=0 b1=3 b2=2 c1=1.5 c2=1.875 d1=4 d2=4 g1=2 g2=1 RSD=10% theta=0



Method	J	J2	J3	F	DE	AIC	AIC-c	Yu
Percent	0.8%	0.8%	0.8%	0%	7.4%	0%	0%	0%

Acceptable Assay: $(b_1,g_1)=(3,1)$ vs. $(b_2,g_2)=(2,2)$





Method	J	J2	J3	F	DE	AIC	AIC-c	Yu
50%RSD	95.2%	0%	13.9%	91.0%	44.8%	78.6%	90.4%	100%
10%RSD	85.2%	85.2%	85.2%	8.4%	37.6%	2.8%	7.4%	94.2%

Penalize the precise assay

Summary

- Nonlinear Case
  - AIC, AIC-c: closed to F in terms of determining similarity
  - F: variance ↘, tends to reject acceptable assay
  - DE: only when ADG are equal; variance ↗, same as linear
  - J, J2, J3: Better than F, boundary…
  - Yu: Better than F, boundary…

An example

Data from a stability study

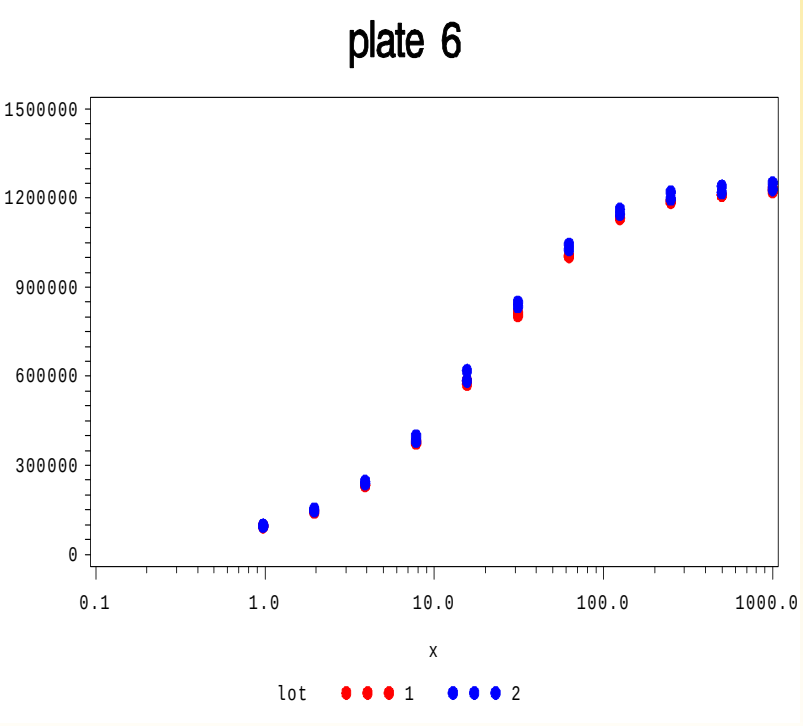


plate 6
1500000
1200000
900000
600000
300000
0
0.1
1.0
10.0
100.0
1000.0
x
lot
1
2

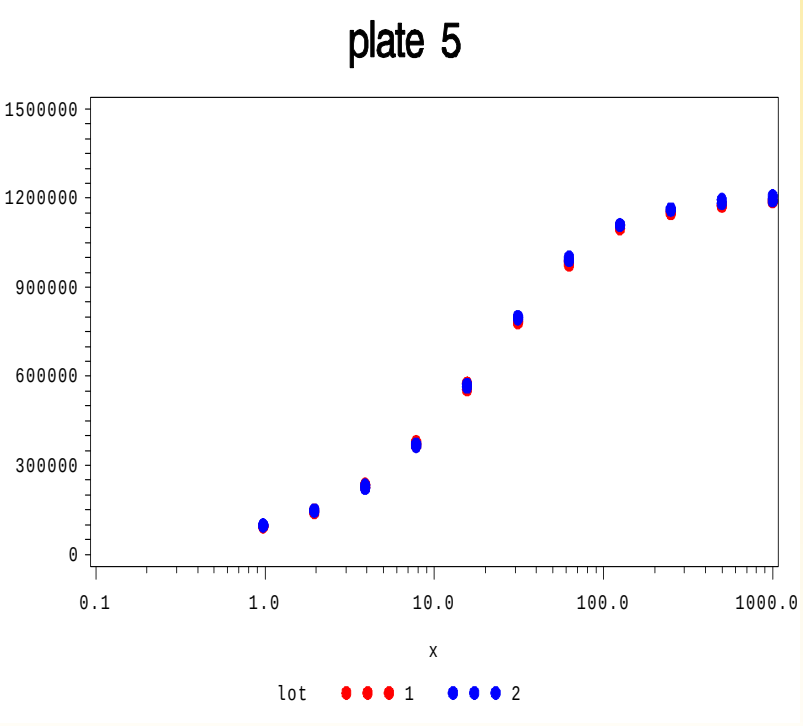


plate 5
1500000
1200000
900000
600000
300000
0
0.1
1.0
10.0
100.0
1000.0
x
lot
1
2

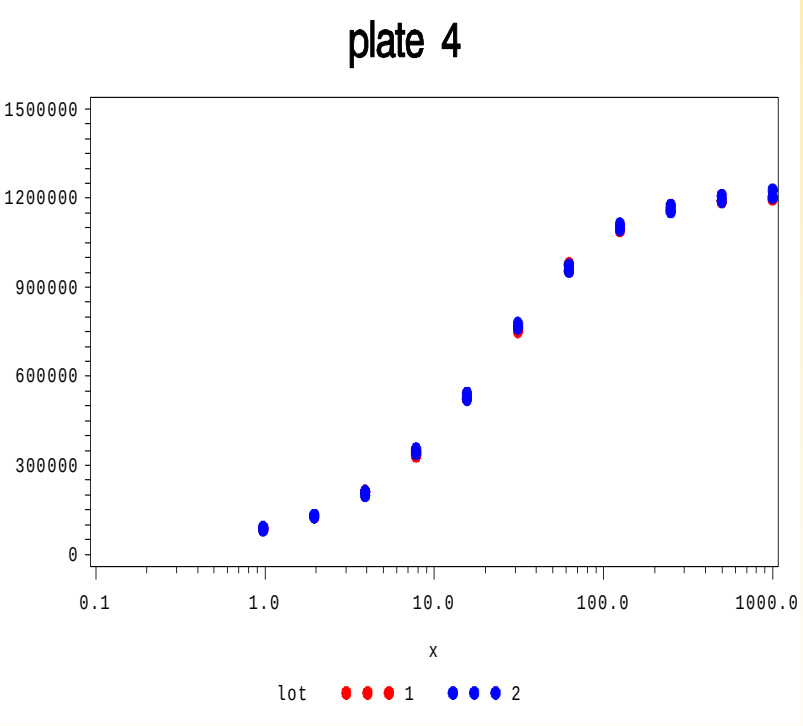


plate 4
1500000
1200000
900000
600000
300000
0
0.1
1.0
10.0
100.0
1000.0
x
lot
1
2

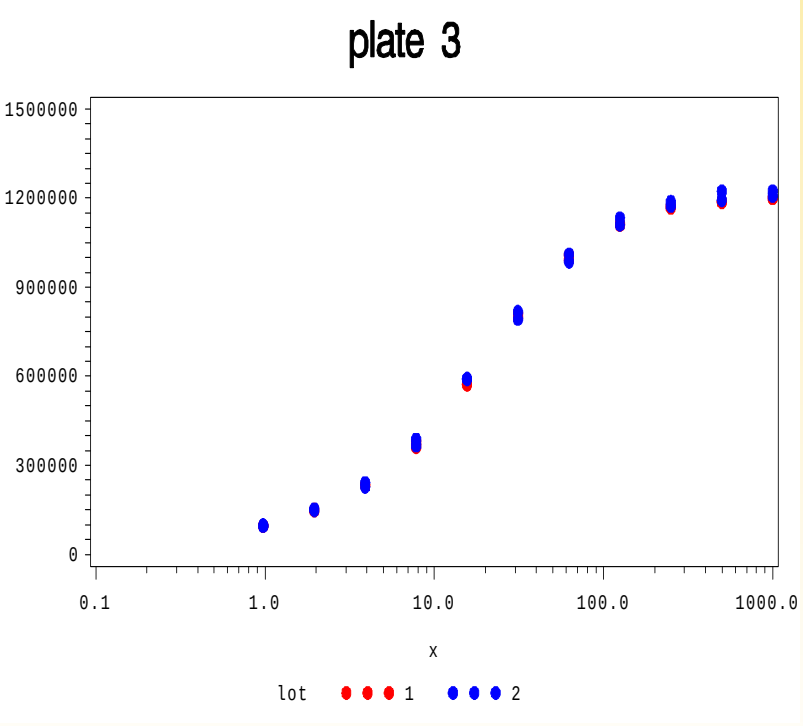


plate 3
1500000
1200000
900000
600000
300000
0
0.1
1.0
10.0
100.0
1000.0
x
lot
1
2

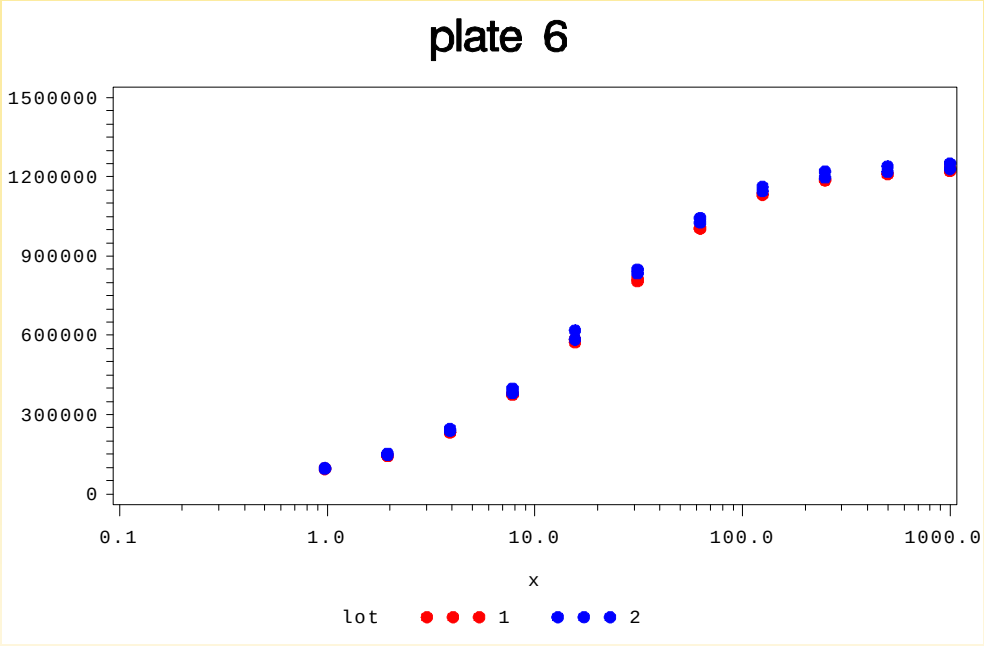


Plate	F	AIC	AIC-c	DE	J	Yu
3	Parallel	Not Parallel	Parallel	Parallel	Parallel	Parallel
4	Parallel	Parallel	Parallel	Parallel	Parallel	Parallel
5	Not Parallel	Not Parallel	Not Parallel	Parallel	Parallel	Parallel
6	Not Parallel	Not Parallel	Not Parallel	Parallel	Parallel	Parallel

Sensitive to trivial difference

Summary & Discussion

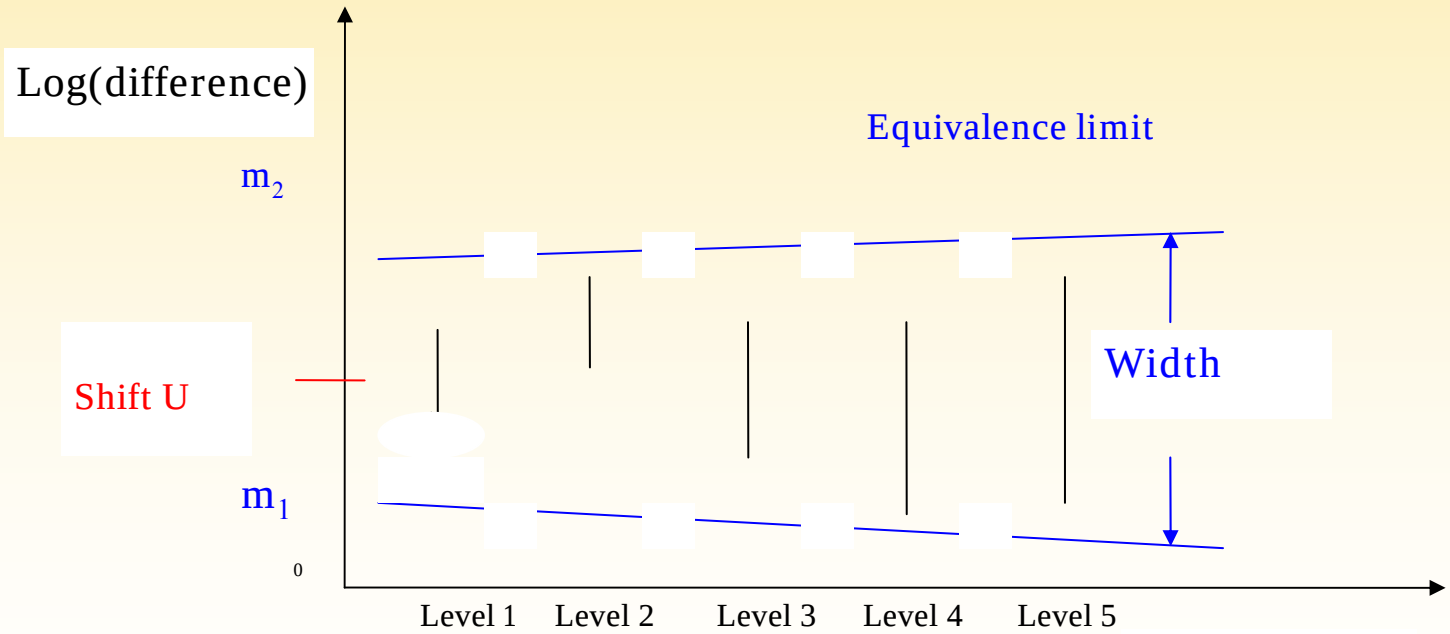
- New method using equivalence approach
  - On the difference of response directly
  - Not on the parameters

- Advantages of the new method
  - Have better performance than existing methods
  - Avoid high dimension issues/complicated parameter region
  - Use scientific judgement about the boundary
  - Can add practical boundary into account
  - Catch the non-linear trend
  - Catch the non-homogenous variance
  - Control the rate of false non-similarity

- Bootstrap method can be used to improve precision for small sample

Summary & Discussion (cont.)

- Heterogeneous Variance: $Var(y) = c \cdot (y)^{2r}$
  - New method can be easily applied




Thank You !