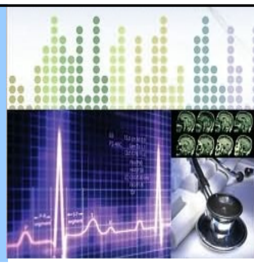



Statistical analysis of biological assays

Yossi Levy

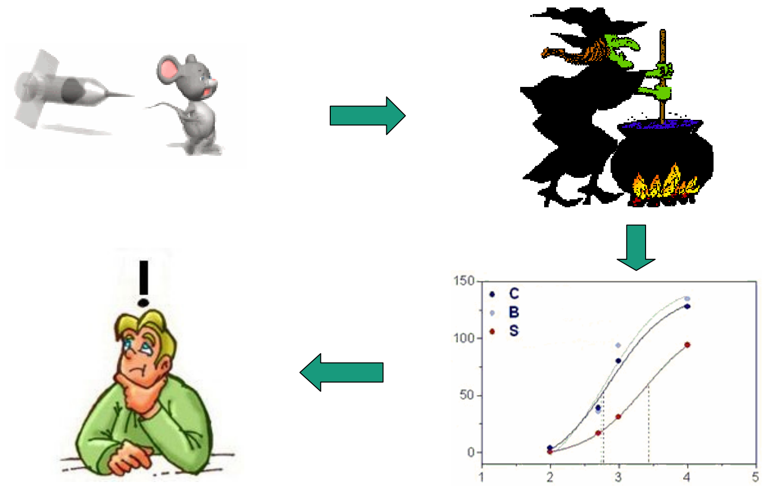
What is a bioassay?


Bioassays are for estimating the potency of a drug by utilizing the reaction caused by its application to live experimental subjects.

Bioassay always compares a test substance to a standard substance

- Assumptions
  - Comparable organisms
  - Same active compound
  - Only concentration can vary

Bioassay



Role of statistics in bioassay

- Advise on the general statistical principles underlying the assay method
- Devise a good experimental design that gives the most useful and reliable results
- Analyze the data making use of all the evidence on potency

Types of bioassay

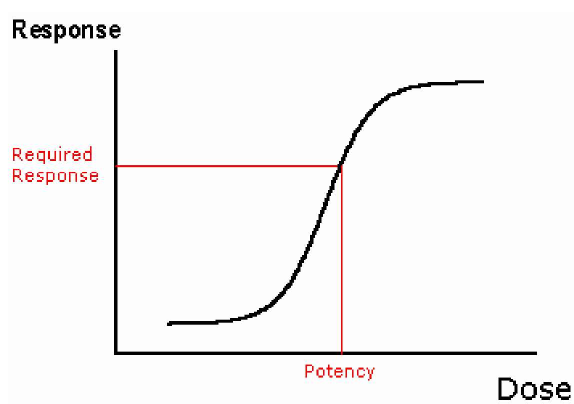
- Direct assay: response is directly measured
- Indirect assay
  - Quantitative
  - Binary

Examples

- Direct assay: Measure insulin level in blood
- Indirect assay
  - Quantitative: change in weight of a certain organ
  - Binary: dead or alive

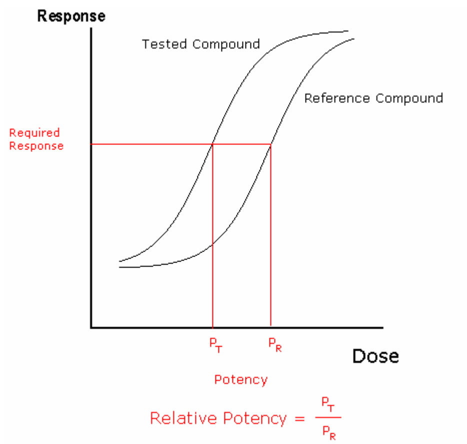
Potency

- Potency is the dose of a compound required to cause a particular response



Relative potency

- Relative potency is the ratio of the tested compound potency to the reference standard potency



Statistical models for quantitative assays

- Parallel line model
- Logistic model (4 or 5 parameters)
- Slope Ratio model

Basic requirements for applying a quantitative bioassay model

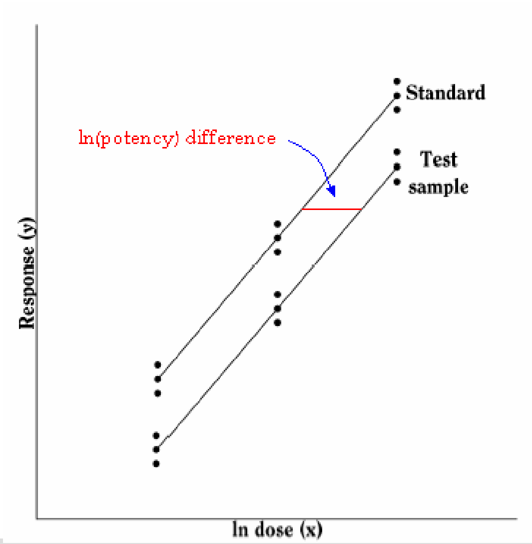
- Randomization
- Responses are Normally distributed
- Homogenous variances

A logarithmic transformation of the response measure is recommended to improve compliance with second and third requirements when necessary.

Requirements for PLA model

- The relationship between the logarithm of the dose and the response can be represented by a straight line.
- For any unknown (tested) substance the straight line is parallel to that of the standard.

1 picture = 1000 words



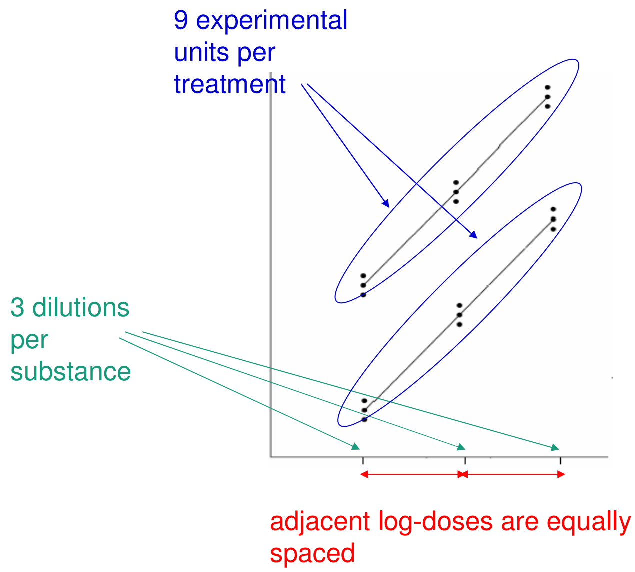
PLA in practice

- Design restrictions imposed by the ICH guidelines
- Experimental design
- Analysis of covariance
- Tests of validity
- Potency estimation and confidence limits
- Handling missing values
- Troubleshooting

Design restrictions

- Each substance must be tested with the same number of dilutions
- The ratio of adjacent doses must be constant for all treatments
- There must be an equal number of experimental units to each treatment

Design restrictions



Experimental design

Completely randomized design – if experimental units are reasonably homogeneous.

ICH guideline also discusses:

- Randomized block design
- Latin-square designs
- Cross-over designs

Tests of validity

The bioassay PLA model is valid if

- Assay must show response
- Response must be linear
- Response lines must be parallel

How to assess linearity?

- Just look at R^2 - naive
- Add a quadratic term to the model and verify that it is non-significant
- Model dose/dilution as a class variable, and compare the results to the "correct" model, using log-likelihood test
- Linear contrasts – compare slope between each two adjacent doses to the next slope

How to asses parallelism?

Let β_I be the regression coefficient of the interaction term, logdose*substance

Significance test approach: H_0: β_I=0

- If the corresponding p-value is less than α, we conclude non-parallelism
- If the corresponding p-value is greater than α, then what?

How to asses parallelism?

Equivalence test approach: H_0: $\beta_I \neq 0$

- Pre-determine acceptance limits for β_I: [-A, A]
- Calculate a 1- α confidence interval for β_I: (β_L, β_U)
- Reject H_0 if $-A < \beta_L$ and $\beta_U < A$

But, how one would determine A?

Which approach is better?

- It depends
- For more details see:

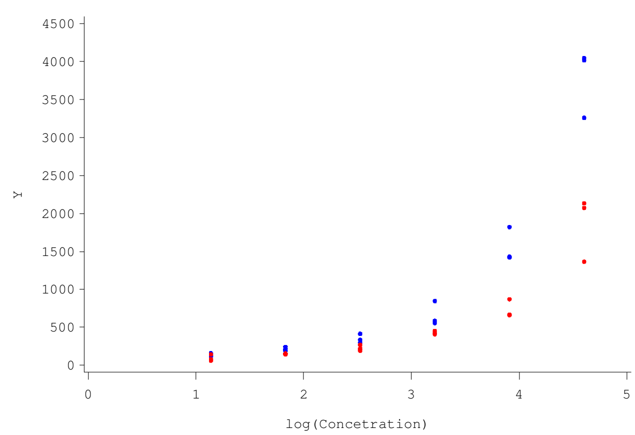
  Evaluations of Parallelism Test Methods Using ROC Analysis
  Harry Yang and Lanju Zhang, MedImmune
  2009 Non-clinical Biostatistics Conference, Boston, MA
  http://www.hsph.harvard.edu/ncb2009/files/ncb2009-c06-yang.pdf

Key conclusion:
An optimal cut off value, in terms of test statistic, p- value or equivalence bound can be chosen to make best trade-off between sensitivity and specificity

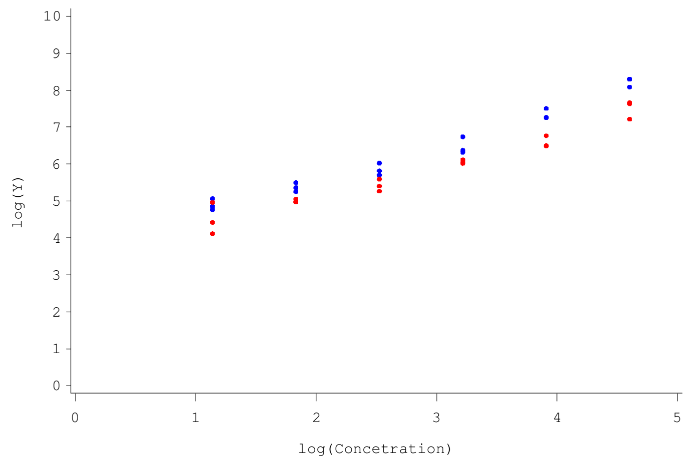
Example

```
data in;
  input substance $ conc y1 y2 y3;
   lconc=log(conc);
cards;
RS  100   4050.538 4019.029 3260.831
RS  50    1432.281 1823.191 1422.876
RS  25    558.284  587.956  848.65
RS  12.5  302.114  336.969  414.975
RS  6.25  191.442  244.982  213.579
RS  3.125 158.749  128.868  118.364
TB  100   1366.585 2134.742 2075.934
TB  50    660.938  669.61   872.149
TB  25    453.385  412.586  424.543
TB  12.5  269.963  193.644  222.505
TB  6.25  145.862  145.862  156.593
TB  3.125 143.725  83.434   61.609
;
run;
```

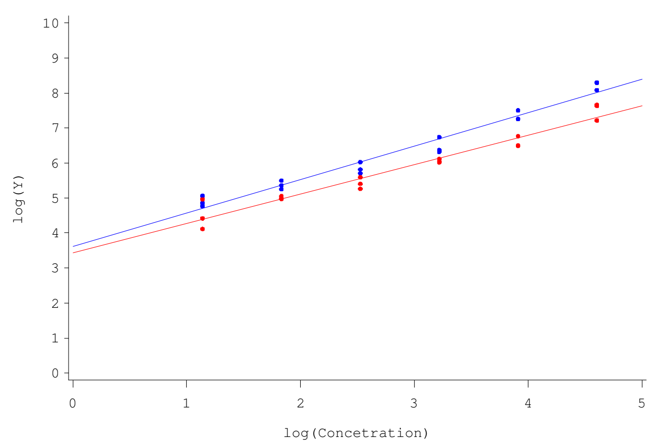

Take a look



Apply log transformation to Y



Add regression lines



Testing linearity

- Run a quadratic model for each substance:
  - $E[\log(Y)]=\beta_0+\beta_1 \cdot \log(\text{concentration}) + \beta_2 \cdot \log(\text{concentration})^2$
- Reject linearity if β_2 is significantly different from zero
- In order to assess linearity, non-linearity must be rejected for each of the substances
- What is the problem in this approach?

Test linearity - RS

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	23.53750642	11.76875321	553.95	<.0001
Error	15	0.31867988	0.02124533		
Corrected Total	17	23.85618630			

R-Square	Coeff Var	Root MSE	ly Mean
0.986642	2.290293	0.145758	6.364154

Source	DF	Type I SS	Mean Square	F Value	Pr > F
lconc	1	23.07441121	23.07441121	1086.09	<.0001
lconc2	1	0.46309521	0.46309521	21.80	0.0003

Test linearity - TB

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	18.08296613	9.04148307	203.15	<.0001
Error	15	0.66759614	0.04450641		
Corrected Total	17	18.75056228			

R-Square	Coeff Var	Root MSE	ly Mean
0.964396	3.606697	0.210965	5.849269

Source	DF	Type I SS	Mean Square	F Value	Pr > F
lconc	1	17.85632170	17.85632170	401.21	<.0001
lconc2	1	0.22664444	0.22664444	5.09	0.0394

Select sub-range

- Linearity of the whole concentration range was not assessed
- It is possible that the response is linear in a sub-range of at least 4 concentrations:
  - 3.125-50 (5 concentrations)
  - 6.25-100 (5 concentrations)
  - 3.125-25 (4 concentrations)
  - Etc..
- The guideline allows selecting the “best range”

How to select the "best range"?

- It must demonstrate linearity response and parallelism
- If there is more that one such sub-range, the best one should be chosen
- Most commercial software select the range with highest R^2
- Better approach: select range with highest signal to noise ratio:

$$S/N = \frac{Y_{\max} - Y_{\min}}{MSE}$$

*The example will continue with the upper range: 12.5-100

Linearity testing – Upper range

substance=RS

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	9.64902455	4.82451228	186.22	<.0001
Error	9	0.23317353	0.02590817		
Corrected Total	11	9.88219809			

R-Square	Coeff Var	Root MSE	ly Mean
0.976405	2.306705	0.160960	6.977927

Source	DF	Type I SS	Mean Square	F Value	Pr > F
lconc	1	9.59920698	9.59920698	370.51	<.0001
lconc2	1	0.04981757	0.04981757	1.92	0.1989

substance=TB

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	6.95727004	3.47863502	114.99	<.0001
Error	9	0.27226780	0.03025198		
Corrected Total	11	7.22953784			

R-Square	Coeff Var	Root MSE	ly Mean
0.962340	2.719261	0.173931	6.396259

Source	DF	Type I SS	Mean Square	F Value	Pr > F
lconc	1	6.89957386	6.89957386	228.07	<.0001
lconc2	1	0.05769618	0.05769618	1.91	0.2006

Testing response

- Linearity has been assessed. The quadratic term can be removed from the model
- Run a linear model for each substance:
  - $E[\log(Y)]=\beta_0+\beta_1\cdot\log(\text{concentration})$
- Reject null hypothesis of no response if β_1 is significantly different from zero
- In order to assess response, null hypothesis must be rejected for each of the substances

Results of response test

substance=RS

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	9.59920698	9.59920698	339.21	<.0001
Error	10	0.28299110	0.02829911		
Corrected Total	11	9.88219809			

R-Square	Coeff Var	Root MSE	ly Mean
0.971364	2.410793	0.168223	6.977927

Source	DF	Type I SS	Mean Square	F Value	Pr > F
lconc	1	9.59920698	9.59920698	339.21	<.0001

substance=TB

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	6.89957386	6.89957386	209.10	<.0001
Error	10	0.32996398	0.03299640		
Corrected Total	11	7.22953784			

R-Square	Coeff Var	Root MSE	ly Mean
0.954359	2.839927	0.181649	6.396259

Source	DF	Type I SS	Mean Square	F Value	Pr > F
lconc	1	6.89957386	6.89957386	209.10	<.0001

Parallelism test

- Response has been assessed. Add substance and its interaction with concentration to model
- Run a linear model for whole data over the chosen range:
  - $E[\log(Y)]=\beta_0 + \beta_1 \cdot \log(\text{concentration}) + \beta_2 \cdot \text{Substance} + \beta_3 \cdot \log(\text{concentration}) * \text{Substance}$
  - Substance is modeled as a 0-1 variable
- Either test whether β_3 is significantly different from zero,
- Or better: construct a confidence interval for β_3

Parallelism test results

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	18.52880442	6.17626814	201.52	<.0001
Error	20	0.61295509	0.03064775		
Corrected Total	23	19.14175951			

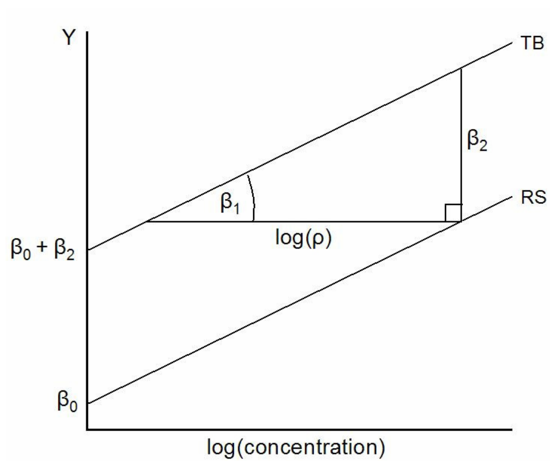
R-Square	Coeff Var	Root MSE	ly Mean
0.967978	2.617954	0.175065	6.687093

Source	DF	Type I SS	Mean Square	F Value	Pr > F
lconc	1	16.38759892	16.38759892	534.71	<.0001
substance	1	2.03002358	2.03002358	66.24	<.0001
lconc*substance	1	0.11118192	0.11118192	3.63	0.0713

Relative potency estimation

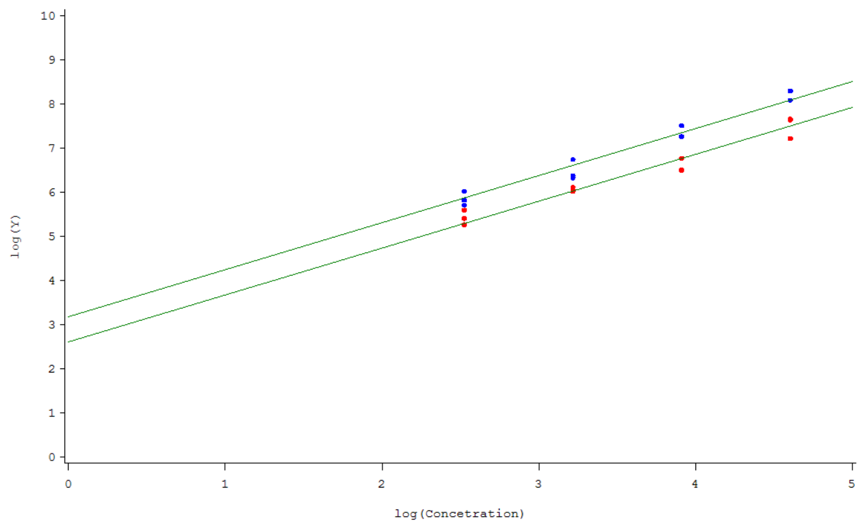
- The final bioassay model over the chosen range is:
  - $E[\log(Y)] = \beta_0 + \beta_1 \cdot \log(\text{concentration}) + \beta_2 \cdot \text{Substance}$
  - Substance is modeled as a 0-1 variable
- This model implies that for the RS (substance=0)
  - $E[\log(Y)] = \beta_0 + \beta_1 \cdot \log(\text{concentration})$
- And for the TB (substance=1)
  - $E[\log(Y)] = (\beta_0 + \beta_2) + \beta_1 \cdot \log(\text{concentration})$

Relative Potency estimation



$$\log \hat{\rho} = \frac{\hat{\beta}_2}{\hat{\beta}_1}$$

The final bioassay model



Relative Potency estimation

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	18.41762250	9.20881125	267.06	<.0001
Error	21	0.72413701	0.03448271		
Corrected Total	23	19.14175951			

R-Square	Coeff Var	Root MSE	ly Mean
0.962170	2.776920	0.185695	6.687093

Parameter	Estimate		Standard Error	t Value	Pr > |t|
Intercept	2.594488957	B	0.18244577	14.22	<.0001
lconc	1.066280765		0.04891191	21.80	<.0001
substance RS	0.581667656	B	0.07580976	7.67	<.0001
substance TB	0.000000000	B	.	.	.

The X'X matrix has been found to be singular, and a generalized inverse was used to solve the normal equations. Terms whose estimates are followed by the letter 'B' are not uniquely estimable.

$$\log\hat{\rho} = \frac{-0.5817}{1.0662} = -0.5455$$

$$\hat{\rho} = \exp(-0.5455) = 0.5795 \approx 58\%$$

How to calculate a confidence interval?

- You can't.
- Estimators of β are linear combinations of the Ys, that are normally distributed
- Therefore, estimators of β are normally distributed
- So the distribution of $\log \hat{\rho} = \frac{\hat{\beta}_2}{\hat{\beta}_1}$ is Cauchy
- Instead of a confidence interval, we calculate fiducial limits

Fieller's theorem

Let μ and v be two unknown parameters, and let ρ= μ/v.

Let a and b be unbiased estimators for μ and v, respectively, that are linear in observations that are normally distributed.

Let the variances and covariance estimates of a and b be $v_{11}s^2$, $v_{22}s^2$ and $V_{12}s^2$, respectively, where s^2 is an error mean square having m degrees of freedom.

Let t be the $\alpha/2$ critical value from a t distribution with m-1 degrees of freedom, and let $g=t^2s^2v_{22}/b^2$.

Let R=a/b and estimate for ρ. Then upper and lower confidence limits for ρ are:

Fieller's theorem

$$R_L, R_U = \frac{\left\{ R - \frac{g v_{12}}{v_{22}} \pm \frac{ts}{b} \cdot \left[v_{11} - 2R v_{22} + R^2 v_{22} - g\left(v_{11} - \frac{v_{12}^2}{v_{22}} \right) \right]^{1/2} \right\}}{1 - g}$$

Proof of Fieller's theorem

Let U=a-ρb. Then EU=0 and its estimated variance is

$s^2(v_{11}-2\rho v_{12}+ \rho^2 v_{12})$ with m degrees of freedom.

Therefore:

$$P\left[U^2 \le t^2 s^2 \left(v_{11} - 2\rho v_{12} + \rho^2 v22\right)\right] = 1 - \alpha$$

The results follows for solving the quadratic equation in ρ.

Fiducial limits calculation

R-Square	Coeff Var	Root MSE	ly Mean
0.962170	2.776920	0.185695	6.687093

s

X'X Generalized Inverse (g2)					
	Intercept	lconc	substance RS	substance TB	ly
Intercept	0.96530855	-0.247367194	-0.083333333	0	2.5944889575
lconc	-0.247367194	0.069378966	8.21612E-17	0	1.0662807646
substance RS	-0.083333333	8.21612E-17	0.1666666667	0	0.5816676564
substance TB	0	0	0	0	0
ly	2.5944889575	1.0662807646	0.5816676564	0	0.7241370111

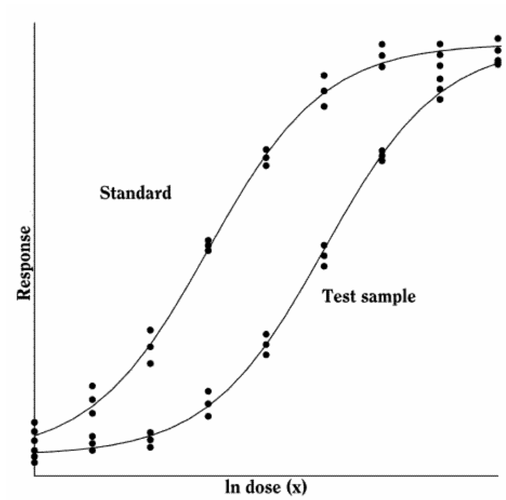
v_{22} v_{11}

Troubleshooting

- Problem: Exceptionally high residual error (MSE)
- Solution: this is an indication of technical problem – check the bioassay process

- Problem: Exceptionally low residual error may cause F values to exceed critical values.
- Solution: replace residual error by estimate from historical data.

Non linear response



4-parameter model

$$EY = \delta + \frac{\alpha - \delta}{1 + \exp\{-\beta(x - \gamma)\}}$$

- α - upper asymptote
- δ - lower asymptote
- β – slope factor
- γ - horizontal location

- Validity of model: α, δ and β – are same for RS and TB
- Log(relative potency) = γ_{RS} - γ_{TB}

Quantal Bioassay

- Response is discrete
  - Often a binary response: e.g. Dead/Alive
- Dose response function is sometimes called "Tolerance Distribution"
- A logistic distribution is a natural model for such data

Example: Bacterial tolerance

Bacterial Dose	Dead	Alive
$1.2 \cdot 10^3$	0	5
$1.2 \cdot 10^4$	0	5
$1.2 \cdot 10^5$	2	3
$1.2 \cdot 10^6$	4	2
$1.2 \cdot 10^7$	5	1
$1.2 \cdot 10^8$	5	0

Modeling

Probability of death at level x_i of drug (or bacterial concentration) is

$$P(Y_i \le x_i) = p_i = \frac{\exp\{\alpha + \beta x_i\}}{1 + \exp\{\alpha + \beta x_i\}}$$

Where Y_i is the tolerance for subject i.

Then

$$\log\left(\frac{p_i}{1 - p_i}\right) = \alpha + \beta x_i$$

LD50/ED50

- The dose at which 50% of subjects produce a response is called LD50 or ED50 (LD=lethal dose, ED=effective dose)
- Let x_{50}=log(LD50) and p_{50}=0.5 (probability of response at the median of the tolerance distribution). Then

$$\log\left(\frac{p_{50}}{1 - p_{50}}\right) = 0 = \hat{\alpha} + \hat{\beta} x_{50}$$

$$\hat{x}_{50} = -\hat{\alpha} / \hat{\beta}$$

$$LD50 = \exp\{-\hat{\alpha} / \hat{\beta}\}$$

Confidence interval for LD50

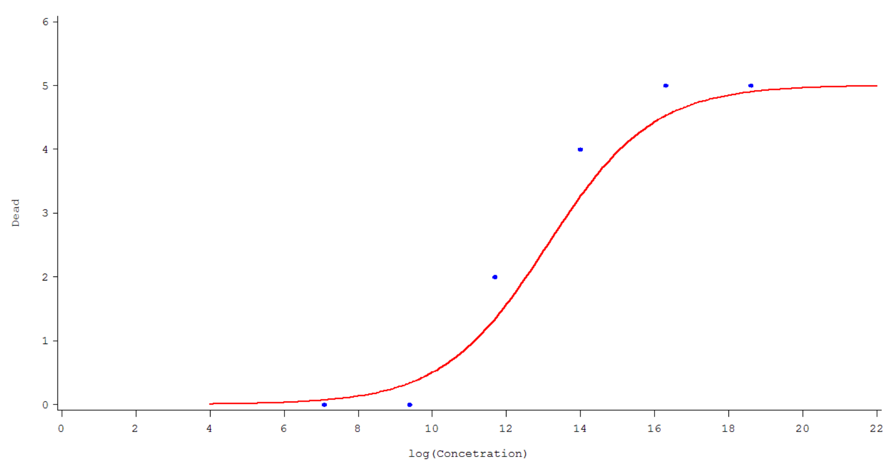
Using Taylor series expansion (the delta method):

$$V(\hat{x}_{50}) = \hat{x}_{50}^2 \cdot \left(\frac{V(\hat{\alpha})}{\hat{\alpha}^2} - \frac{2\operatorname{cov}(\hat{\alpha},\hat{\beta})}{\hat{\alpha}\hat{\beta}} + \frac{V(\hat{\beta})}{\hat{\beta}^2} \right)$$

Then a 100(1-α)% CI for log(LD50) is

$$\hat{x}_{50} \pm z_{1-\alpha/2}\sqrt{V(\hat{x}_{50})}$$

Data and model



proc logistic output

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	22.8356	1	<.0001
Score	17.8025	1	<.0001
Wald	9.0223	1	0.0027

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-9.2680	3.1630	8.5857	0.0034
logdose	1	0.7071	0.2354	9.0223	0.0027

Estimated Covariance Matrix		
Parameter	Intercept	logdose
Intercept	10.00458	-0.73338
logdose	-0.73338	0.055418

$$LD50 = \exp\{9.268/0.7071\} = 488942 \approx 4.8 \cdot 10^5$$

CI for $\log(LD50)$ is

$$13.1 \pm 1.96\sqrt{0.6005} = (11.6, 14.6)$$

CI for $LD50$ is

$$\exp(11.6, 14.6) \approx (10.9 \cdot 10^4, 2.1 \cdot 10^6)$$

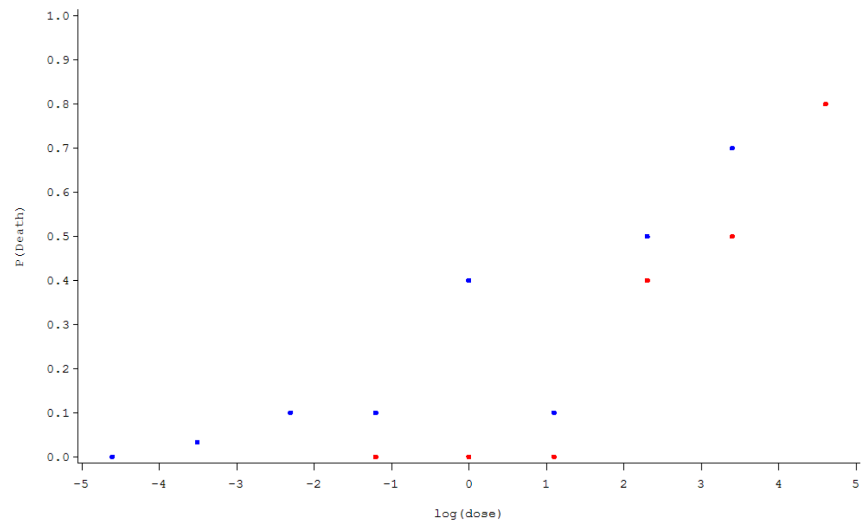
Comparing two drugs

	Drug	Dose	Dead	Alive	Total
1	N	0.01	0	30	30
2	N	0.03	1	29	30
3	N	0.10	1	9	10
4	N	0.30	1	9	10
5	S	0.30	0	10	10
6	N	1.00	4	6	10
7	S	1.00	0	10	10
8	N	3.00	1	9	10
9	S	3.00	0	10	10
10	N	10.00	5	5	10
11	S	10.00	4	6	10
12	S	30.00	5	5	10
13	N	30.00	7	3	10
14	S	100.00	8	2	10

The dilution assumption:

$$Z_S = \rho Z_N$$

Comparing two drugs



Comparing two drugs

- Dilution assumption: $z_S=\rho\cdot z_N$ for doses of S and N with the same probability of response
- If x represents log of dose, then $x_S=\log\rho+x_N$
- Logistic model for drug S is:

$$p_S(x_{Si})=\frac{1}{1+\exp\{-(\alpha_S+\beta x_{Si})\}}$$

Comparing two drugs

Therefore, for drug N, remembering that $x_S = \log\rho + x_N$, a logistic model for drug N is:

$$p_N(x_{Ni}) = p_S(\log\rho + x_{Ni}) =$$
$$= \frac{1}{1+\exp\{-(\alpha_S + \beta(\log\rho + x_{Ni}))\}} =$$
$$= \frac{1}{1+\exp\{-((\alpha_S + \beta\log\rho) + \beta x_{Ni})\}} =$$
$$= \frac{1}{1+\exp\{-(\alpha_N + \beta x_{Ni})\}} =$$

Comparing two drugs

Therefore, the dilution assumption implies that:

$$\log\left\{\frac{p_S(x_{Si})}{1-p_S(x_{Si})}\right\} = \alpha_S + \beta x_{Si}$$
$$\log\left\{\frac{p_N(x_{Ni})}{1-p_N(x_{Ni})}\right\} = \alpha_N + \beta x_{Ni} = \alpha_S + \beta\log\rho + \beta x_{Ni}$$

The assumption can be tested by fitting a model with separate intercepts and slopes and then testing for common slope

Testing for common slope – SAS output

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-2.6943	0.5549	23.5731	<.0001
drug	N	1	1.3234	0.5549	5.6872	0.0171
logdose		1	0.9130	0.1764	26.7940	<.0001
logdose*drug	N	1	-0.3167	0.1764	3.2227	0.0726

Interaction term is not statistically significant

=> common slope

=> the dilution assumption holds

Run model without interaction

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-2.0429	0.3288	38.6153	<.0001
drug	N	1	0.5504	0.2539	4.6974	0.0302
logdose		1	0.7363	0.1254	34.4644	<.0001

Estimated Covariance Matrix			
Parameter	Intercept	drugN	logdose
Intercept	0.108082	-0.02826	-0.0294
drugN	-0.02826	0.064481	0.01336
logdose	-0.0294	0.01336	0.015732

The model:

$$\log\{p/(1-p)\} = -2.0429 + 0.5504 \cdot I(\text{Drug} = N) + 0.7363 \cdot \log dose$$

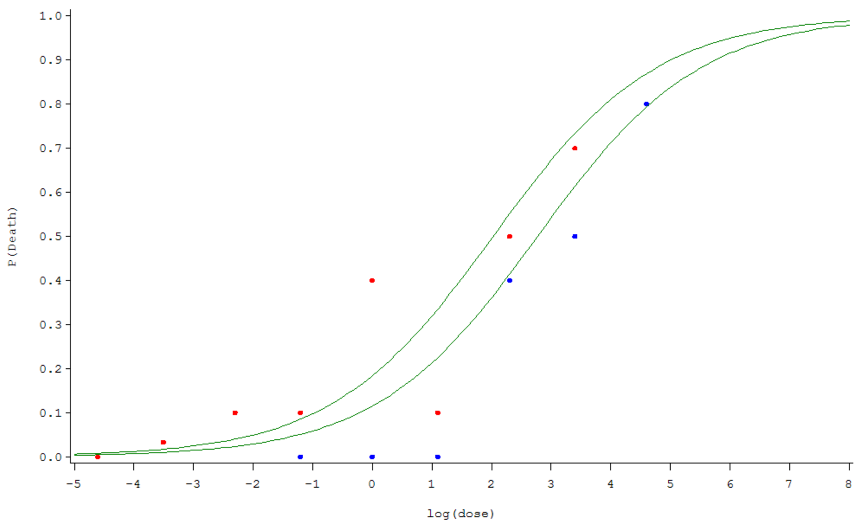
For drug S:

$$\log\{p/(1-p)\} = -2.0429 + 0.7363 \cdot \log dose$$

For drug N:

$$\log\{p/(1-p)\} = -1.4925 + 0.7363 \cdot \log dose$$

The model



Parameter estimates

$$\log LD50_N = -\frac{-1.4925}{0.7363} = 2.027 \qquad LD50_N = 7.60$$

$$\log LD50_S = -\frac{-2.0429}{0.7363} = 2.775 \qquad LD50_S = 16.04$$

$$\log \rho = \frac{\hat{\alpha}_N - \hat{\alpha}_S}{\hat{\beta}} = \frac{0.5504}{0.7363} = 0.7475$$

$$\hat{\rho} = 2.11$$