Data transformation and Parameter Transformations in NONMEM

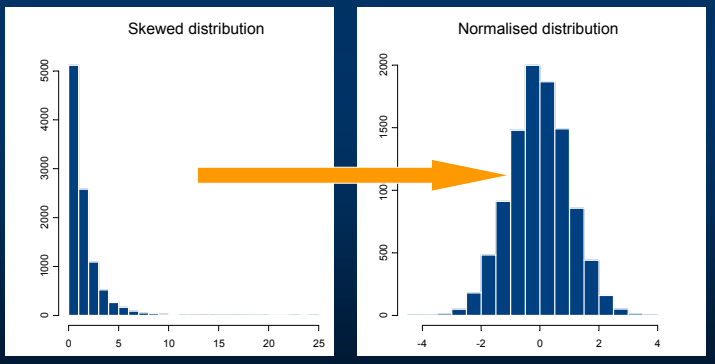
Pascal Girard
Pharsight Corp.
Pgirard@pharsight.com

Acknowledgements

Stuart Beal, UCSF
Lewis Sheiner, UCSF
Janet Wade, Exprimo LLP
NM-User group discussions

Purposes of Parameter & Data Transformation

- **Get unconstrained parameters or data; e.g.:**
  - From constrained 0 → 1 to unconstrained –INF→+INF
- **Normalize distributions**



Examples of Data and Parameter Transformations

Parameters

- Log transformation
- Power transformation
- Logit transformation
- Jupp transformation
- …

Data

- Box and Cox transformation
- Transformation Both Sides: Log-Log (special case of Box-Cox)
- Power transformation
- Logit transformation
- …

Logit transformation:

- Let X so that a < X < b
- Then $$Logit = \mathrm{Log}\left(\frac{X}{1-X}\right) \in [-\infty;+\infty]$$
- And inverse logistic is: $$X = \frac{(b-a)\exp(Logit)}{1+\exp(Logit)}$$

Example: NONMEM implementation for Emax [0-100%]:

```
LOGIT=THETA(1)+THETA(2)*SEX+AGE*THETA(3)+ETA(1)
EMAX=EXP(LOGIT) / (1 + EXP(LOGIT))
```


The Jupp transformation, initially developed to estimate knot locations of a spline, is useful to estimate series of increasing parameters

(Jupp (1978), SIAM *Journal of Numerical Analysis*; 15: 328–343.)

Constraints:

- $a < \theta_1 < \theta_2 < \ldots < \theta_p < b$

Unconstrained parameters γ

- γ_i can vary from –Inf to +Inf and are defined as
- $\gamma_i = \log(h_{i+1}/h_i)$ with $h_i=(\theta_i - \theta_{i-1})/(b-a)$ for i=1,p

Applications in PK-PD:

- rate constants $ka>k_{12}>k_{13}>0$;
- means of mixture distribution: $0 < \mu_1 < \mu_2 < \mu_3 < b$
- …

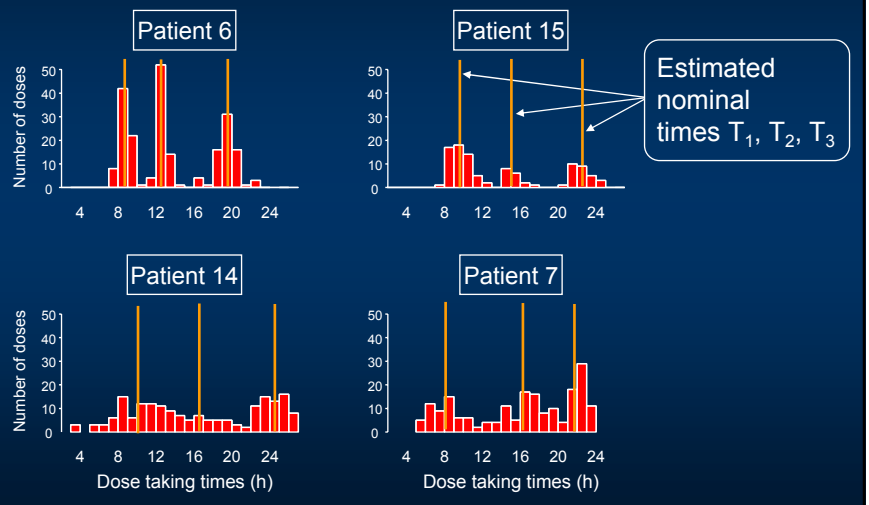
Example: Estimate of drug taking times T1, T2, T3 for a TID dose regimen, from compliance data.

(Girard et .al. (1998), *Stat in Med* ;17 : 2313-2334)

- Constrained dose taking times 0 < T1 < T2 < T3 < 24
- NONMEM code to get unconstrained parameters θ_1, θ_2, θ_3:

```
$PRED
     E1 = EXP(THETA(1))
     E2 = EXP(THETA(2))
     E3 = EXP(THETA(3))
     T1 = 24/(1+E1+E1*E2+E1*E2*E3)
     T2 = E1*T1+ T1
     T3 = E2*T2-T1 + T2
     etc …
$THETA
    0.1              ; THETA1
    0.11             ; THETA2
   -0.4              ; THETA3
```


Estimated nominal times T and histograms of (24h) actual dose times



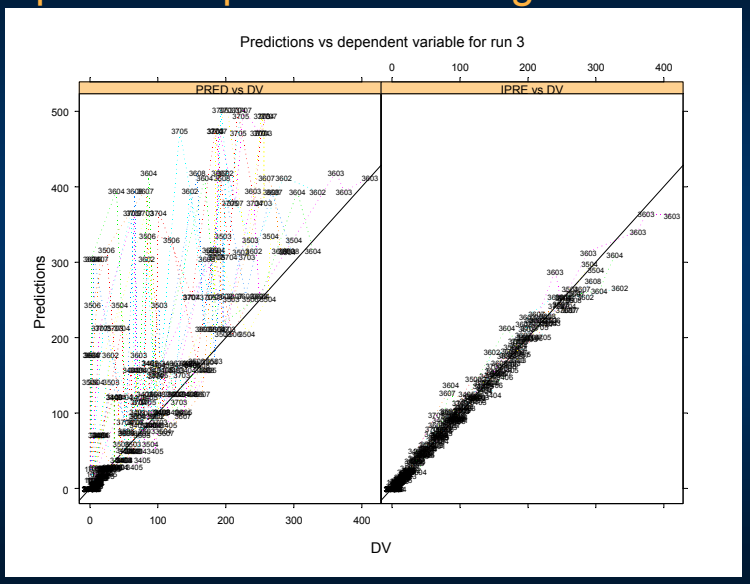
Data transformation

To log transform or to not log transform; that is the question. (Janet R. Wade)

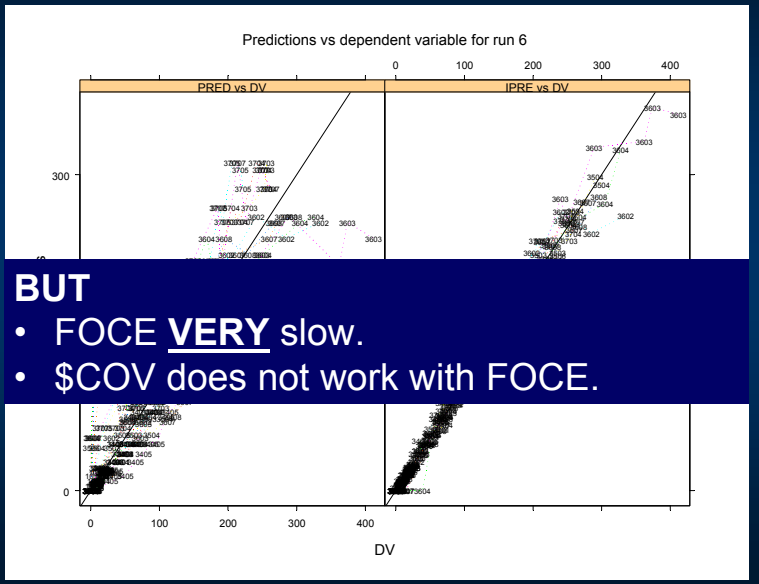
Background

- PK analysis
- Complex absorption
- Non-linear pharmacokinetics
  - Dose
  - Binding
- Rich data

Preliminary runs in NONMEM using FO and Slope intercept error model gave biased fit



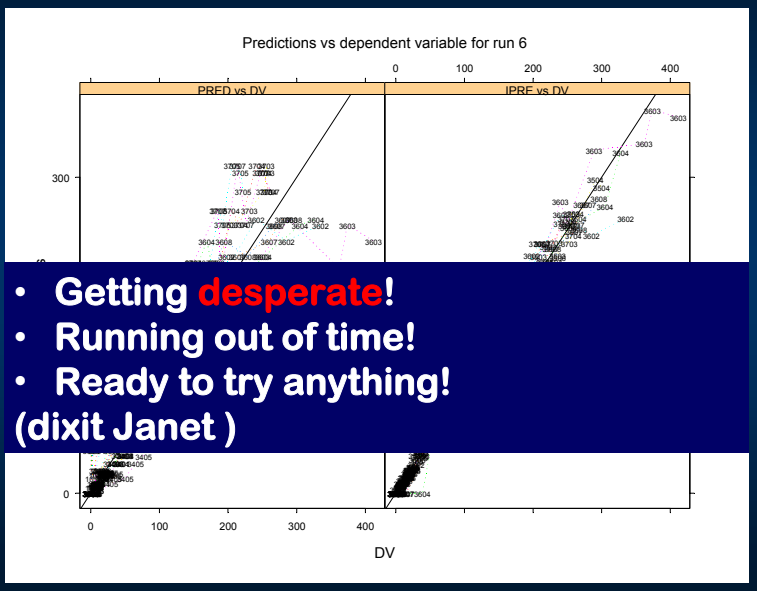
FOCE gave less bias in the PRED vs DV



BUT

- FOCE **VERY** slow.
- $COV does not work with FOCE.

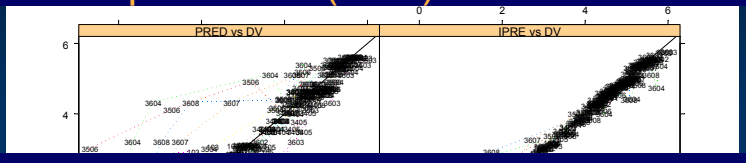
FOCE gave less bias in the PRED vs DV



- **Getting desperate!**
- **Running out of time!**
- **Ready to try anything!**

(dixit Janet)

Discuss with colleagues:
Suggestion made to log transform the data and the prediction (TBS).



Questions

- What indicators are there to suggest transforming your data?
- Can you get problems if the data are transformed when such a transformation isn't needed?

Log transform is a special case of Box and Cox transformation

(Box & Cox (1964) J. R. Stat. Soc. B; 26:211-46.)

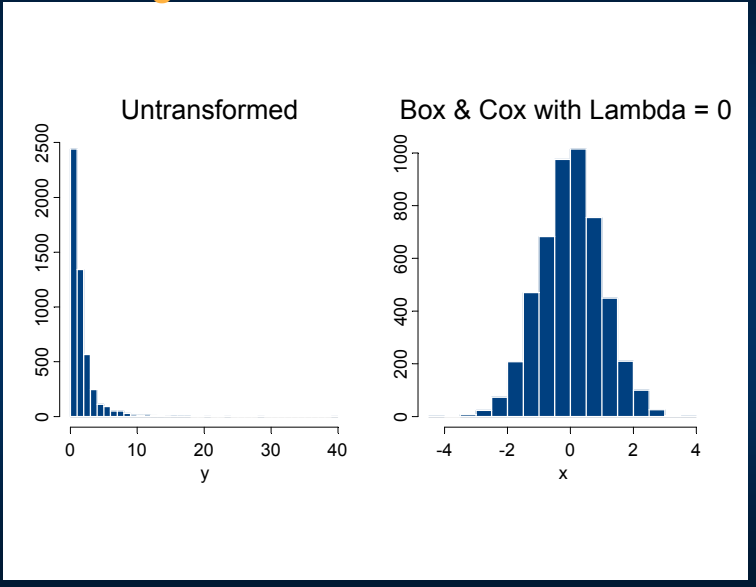
- Box & Cox transformation

$$h(Y,\lambda) = \left(Y^{\lambda} - 1\right)/\lambda \quad \text{if } \lambda \neq 0$$
$$= \log(Y) \quad \text{if } \lambda = 0$$

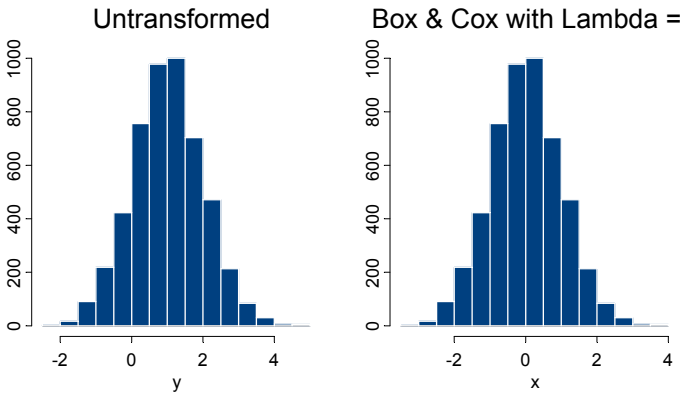
- Inverse Box & Cox transformation

$$h^{-1}(Z,\lambda) = \sqrt[\lambda]{(Z * \lambda + 1)} \quad \text{if } \lambda \neq 0$$
$$= \exp(Y) \quad \text{if } \lambda = 0$$

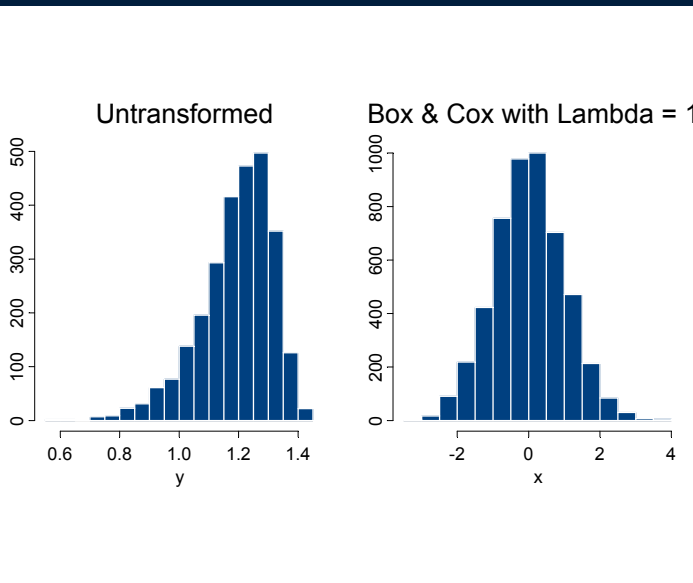
λ=0 $\rightarrow$ log-normal distribution
λ<1 $\rightarrow$ Right-skewed distributions



λ=1 $\rightarrow$ no transformation



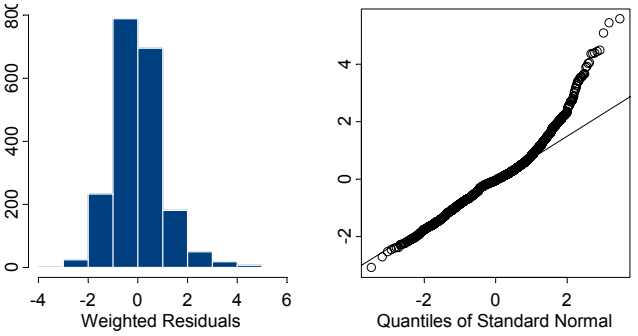
λ>1 $\rightarrow$ left-skewed distributions



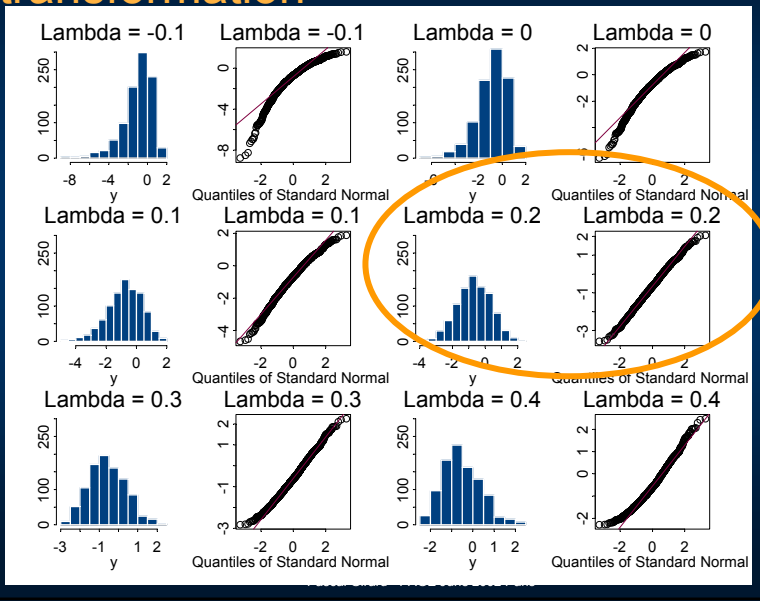
What indicators are there to suggest transforming your data?

Use of histogram and qqplot applied to Weighted Residuals to check normality

Weighted residuals distribution from a simulated model with exponential error



Look for an adequate Boxn & Cox transformation



Can you get problems if the data are transformed when such a transformation isn't needed?

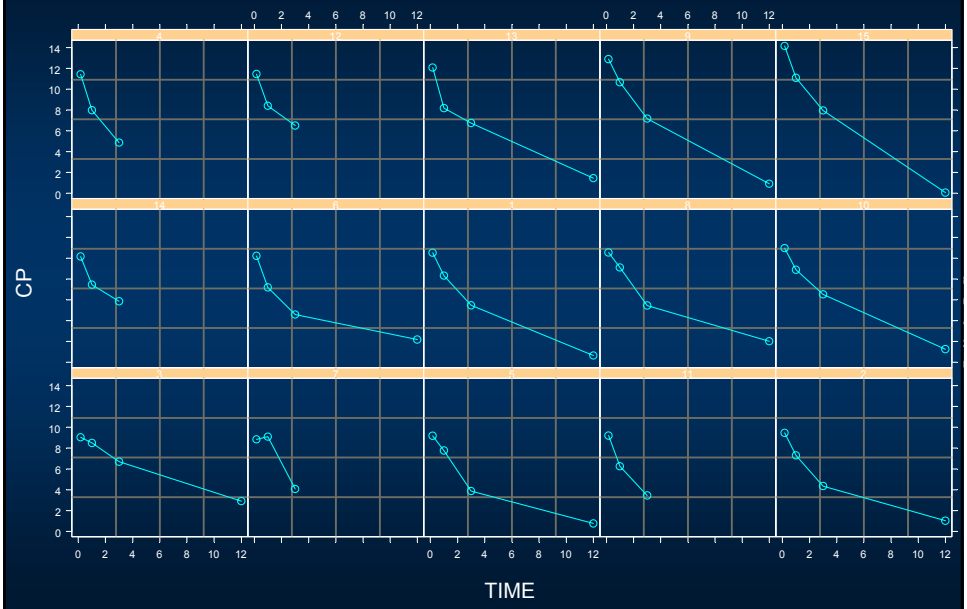
The answer is Yes!

Model with Transformation Both Sides using either Log or Box & Cox transformations

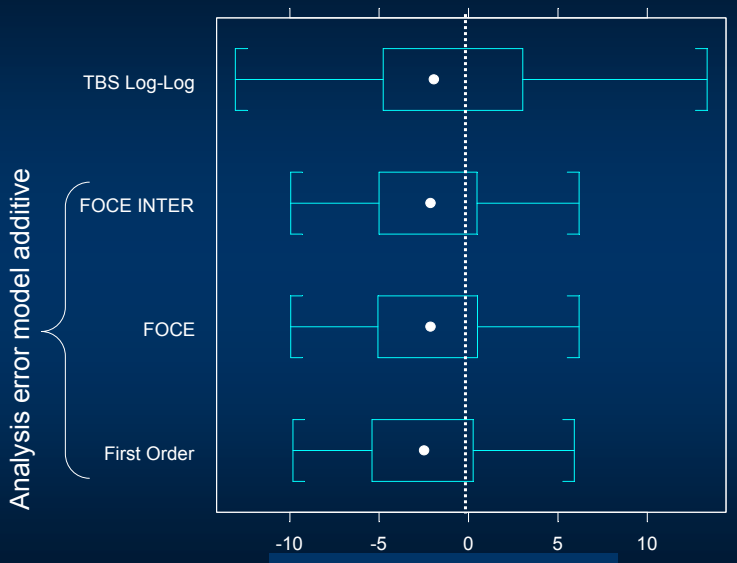
(Carrol & Ruppert (1988), Transformation and Weightinhg in regression. Chapman & Hall)

- Simulate 50 patients with1 compartment PK (Emax PD) models and error
  - Additive
  - Exponential
  - Exponential + additive
- Analyze with true model, using either FO, FOCE, FOCE INTER
- Analyze with TBS log-log
- Perform 100 replications

True simulation model for 15 individuals: PK & additive residuals



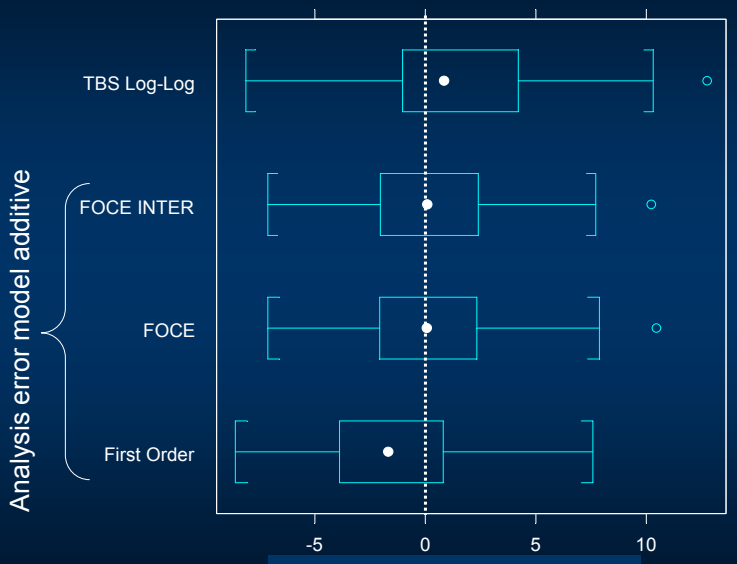
True simulation model: PK & additive residuals



% Bias on rate K

25

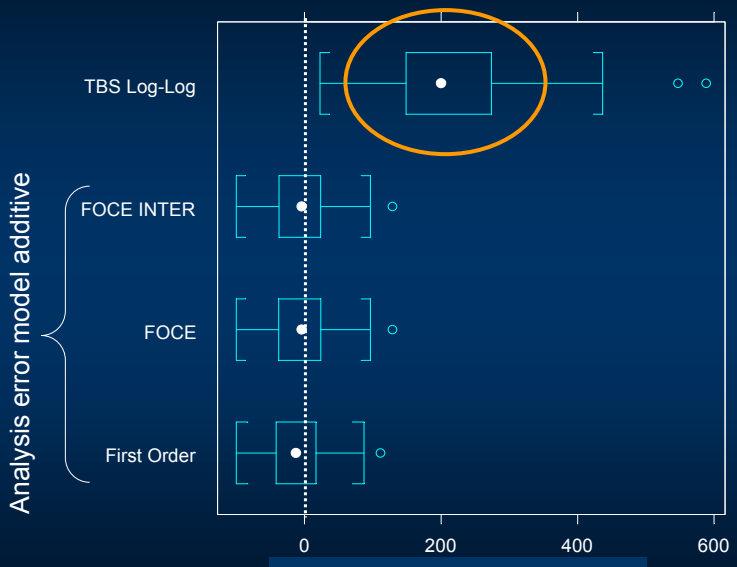
True simulation model: PK & additive residuals



% Bias on volume

26

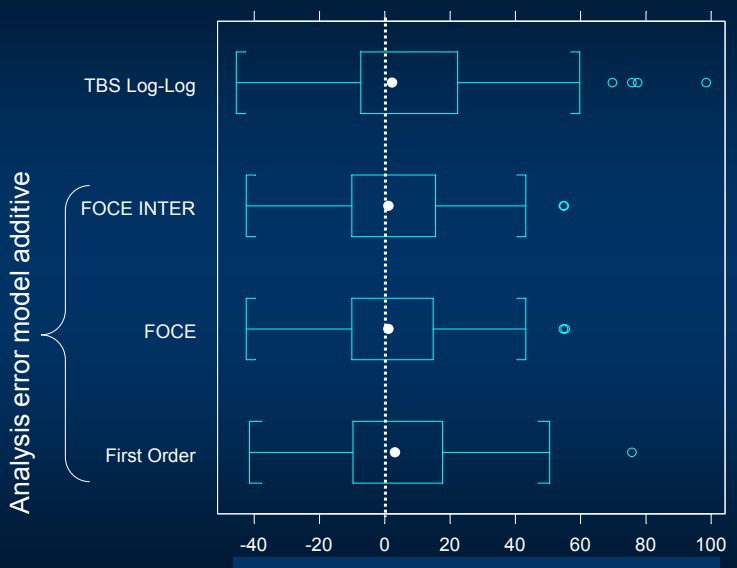
True simulation model: PK & additive residuals



% Bias on ω^2 for K

27

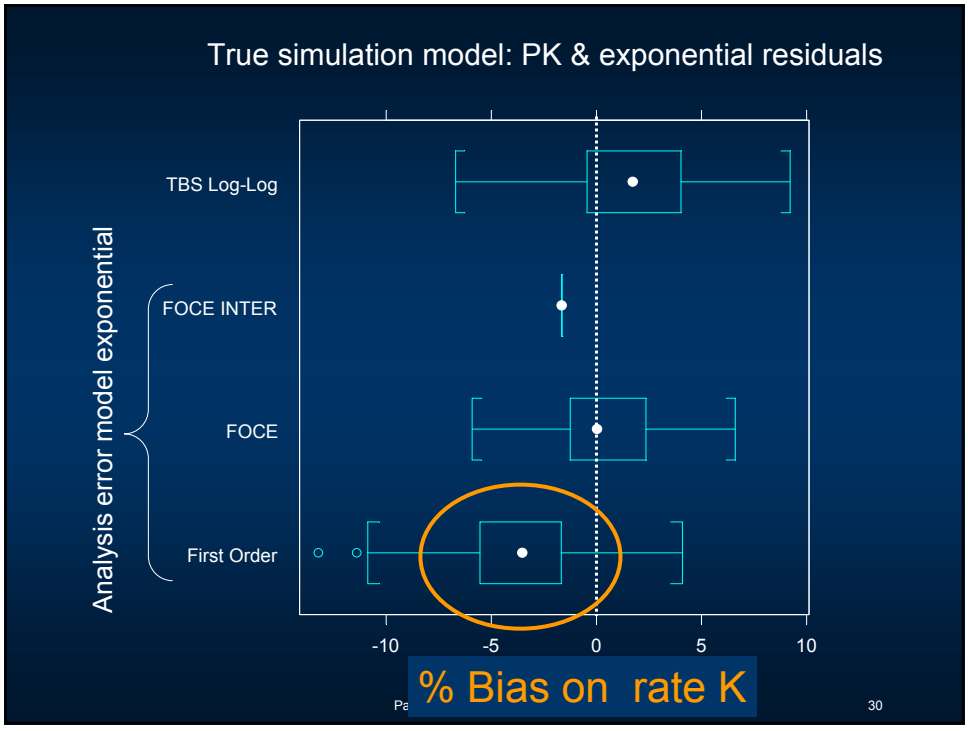
True simulation model: PK & additive residuals



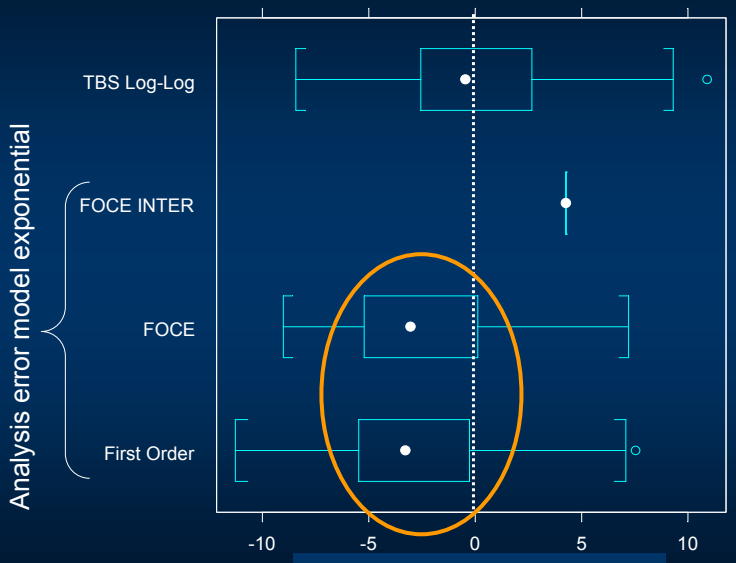
% Bias on ω^2 for volume

28

Simulation with exponential error



True simulation model: PK & exponential residuals



Simulation summary of the bias for PK

Analysis / Simul	FO	FOCE	FOCE INTER	TBS Log-Log
Additive	OK	OK	OK	OM² for K >>>
Exponential	K, V <	V <	?	OK
Exponential + additive	K, V <	OK	OK	OM² for K >>>

Simulation summary of the bias for PD

Simul \ Analysis	FO	TBS Log-Log
Additive	OK	EC50 <<< OM^2 >>>
Exponential	Imprecise OM^2	OK

Bias_PD.ppt

Conclusion and important recommendations

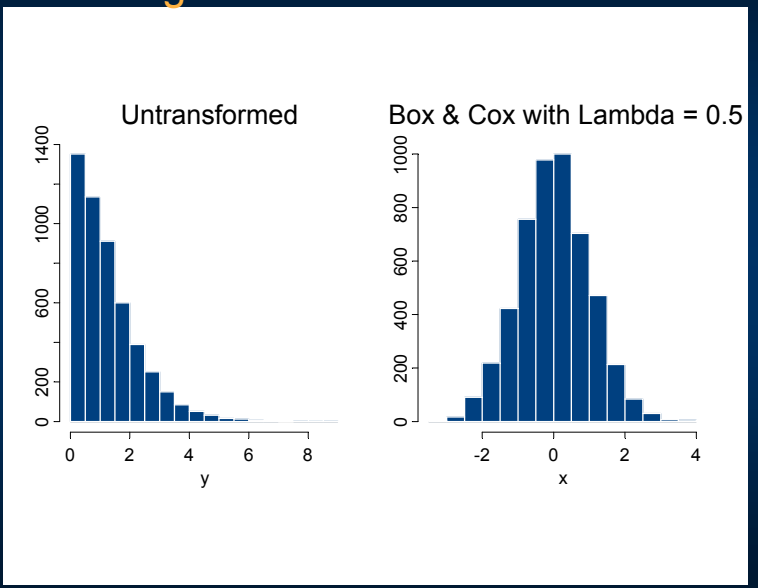
- Analyse carrefully weighted residuals using graphics
- Look for data transformations and use preferably TBS rather than FOCE to re-estimate **all** parameters

But

- Search for best residual model has not to be substituted for search of best structural model.
- Otherwise one will try to compensate model miss-specifications by residual error model sophistication.
- So this search for residual model has to happen after the best possible structural model has been identified.

Backup slides

Example 2: λ=0.5 $\rightarrow$ square-root
λ<1 $\rightarrow$ Right-skewed distributions



Why transform?

Regression model makes 3 basic assumptions:

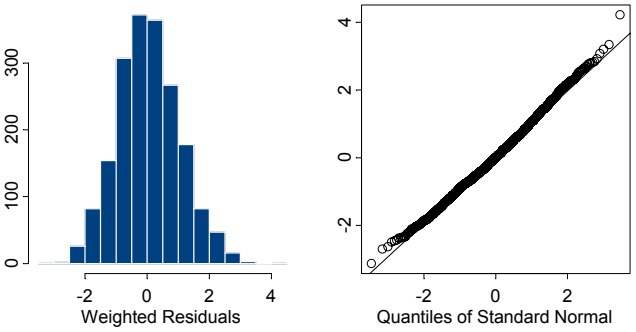
- $E(Y)=f(X, \theta, \eta)$
- $Y - f(X, \theta, \eta) = \varepsilon$, $Variance(Y) = Variance(\varepsilon)=\sigma^2$
- Given X, the errors $\varepsilon=Y - f(X, \theta, \eta)$ are independently distributed

- Furthermore, if one wants certain property for the parameter estimates and to be able to test hypothesis, he will need a 5th assumptions:

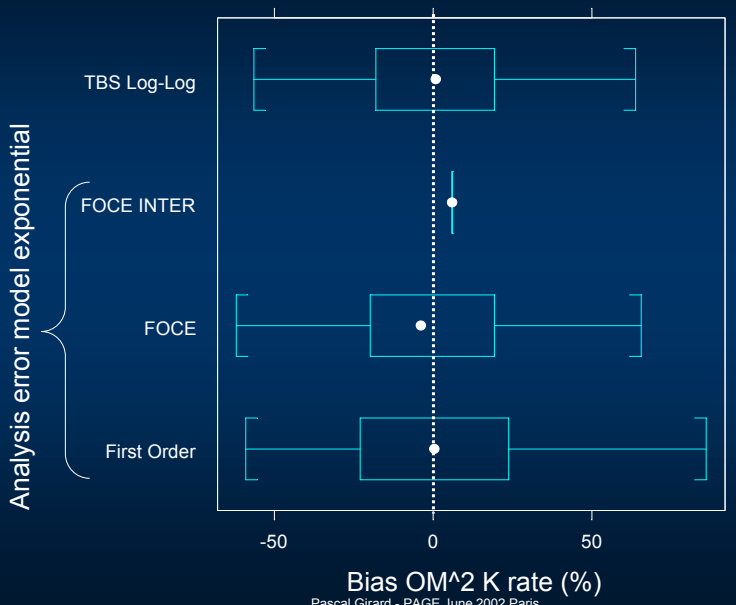
  The errors are normally distributed

Use of histogram and qqplots applied to Weighted Residuals to check normality

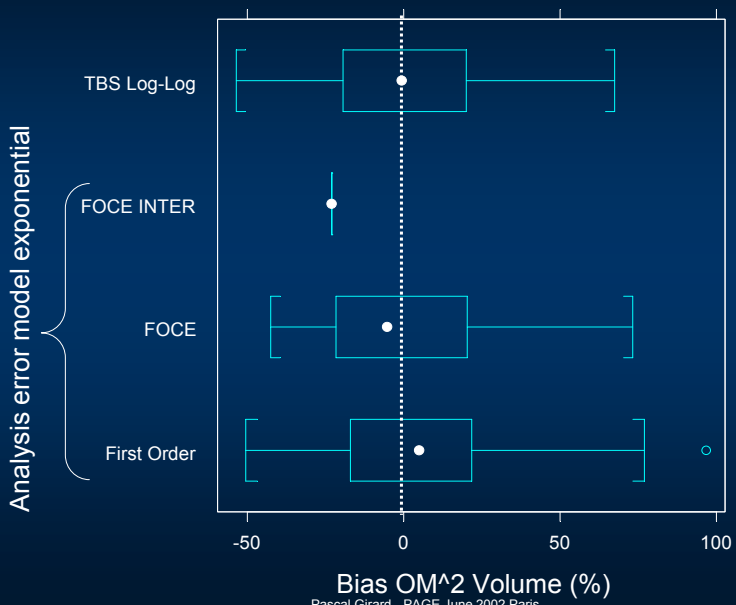
Weighted residuals distribution from a simulated model with additive errors



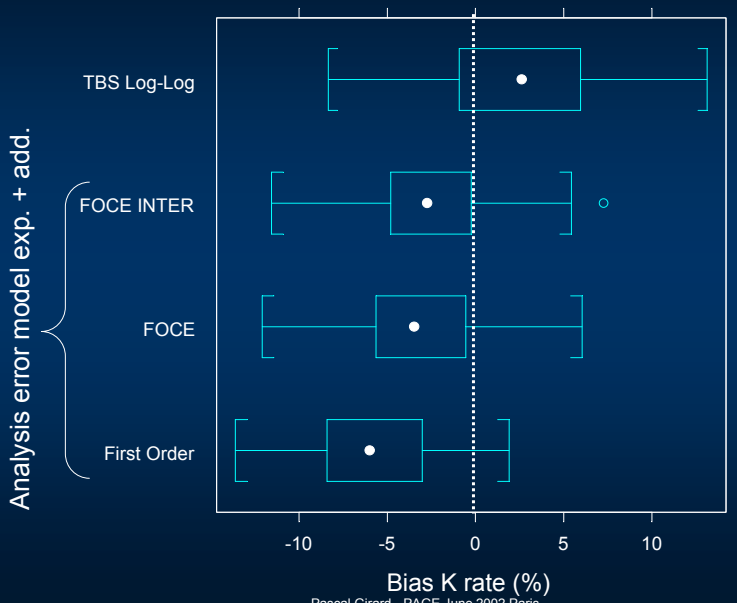
True simulation model: PK & exponential residuals



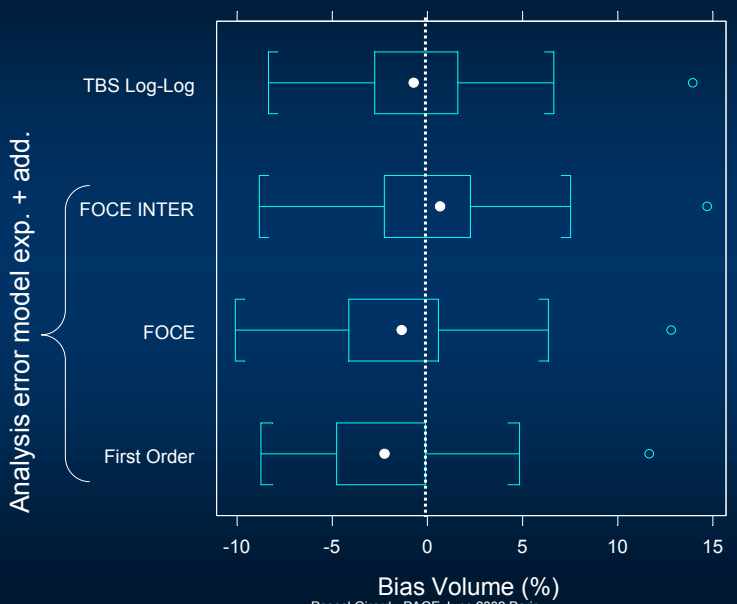
True simulation model: PK & exponential residuals



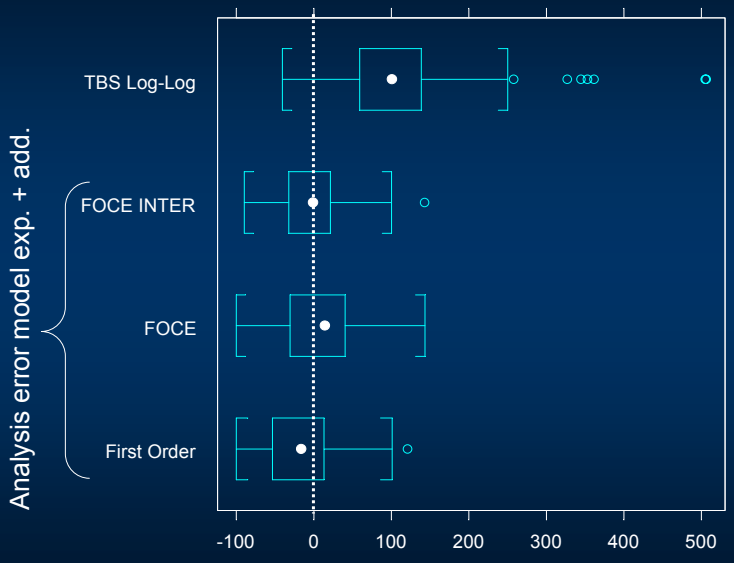
True simulation model: PK & exponential + additive residuals



True simulation model: PK & exponential + additive residuals



True simulation model: PK & exponential + additive residuals



True simulation model: PK & exponential + additive residuals