Variability as constant coefficient of variation: Can we right two decades in error?

Jeroen Elassaiss-Schaap and Siem H. Heisterkamp

Objective

Derive mathematically correct equations that express variability components of PK-PD models as constant coefficients of variation

Introduction

- Variability estimates by NONMEM are 'difficult'.
- Therefore transformed into CV values
  - *CV = coefficient of variation*
- Variability mostly estimated as log-normal
  - *e.g. CL = THETA(x) * EXP(ETA(y))*
- CV mostly approximated
  - *by taking the square root of OMEGA values*
- This approximation is valid when variability is small.

Historical background

- First applications of NONMEM in the 1980s: PK analysis
- Application of NONMEM gradually evolved
- From sparsely sampled PK
  – *Typically Phase III datasets*
- And rich PK-PD
  – *e.g. anesthetics research*
- To general population PK-PD of biomarkers and endpoints
- Variability in PK often in 10-40% range
- Variability in PK-PD often in 50-200% range

Practical issue

- NONMEM application evolved from low to high variability situations.
- In early days approximation of CV% justified
  – *Because variability estimates were low*
- Nowadays variability estimates are often high.
- So how about the validity of the square-root approximation?

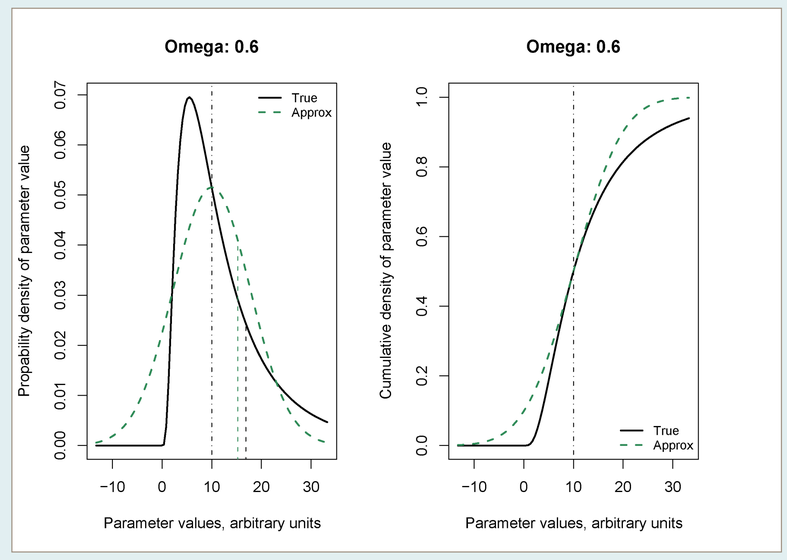

Figure 1. Difference between the true and approximated distribution. The probability density and the cumulative probability density of the true distribution and the distribution described by a constant coefficient of variation approximated by the square root of OMEGA. Vertical lines denote the mean and the 75th percentile of the two distributions.

Algebraic derivation of expressions

- Most commonly used NONMEM expression for a parameter, e.g.:
  CL = THETA(1) * EXP(ETA(1))
- Where
  – *CL = clearance*
  – *THETA(1) = fixed effect parameter*
  – *ETA(1) = random variable drawn from a normal distribution with estimated variance OMEGA(1)*
- In this derivation OMEGA is notated as ω2 to stress it is a variance
- So formally:

$$Par = \theta \cdot e^{\eta} \qquad \eta \propto Normal\left(0, \omega^2\right)$$

- The definitions for mean and variance are:

$$\begin{cases} E\left[f\left(x\right)\right] \underset{D}{=} \int f\left(s\right) \cdot prob.dens\left(s\right) ds \\ Var\left[f\left(x\right)\right] \underset{D}{=} \int f\left(s\right)^2 \cdot prob.dens\left(s\right) ds - \left(E\left[f\left(x\right)\right]\right)^2 \end{cases}$$

- The probability density for a lognormal distribution is given by:

$$prob.dens\left(e^{\eta}\right) = \frac{1}{\eta\sqrt{2\pi\omega^2}} e^{-\frac{\eta^2}{2\omega^2}}$$

- The expression for the mean of the log-normal distribution, see [1]:

$$E\left[e^{\eta}\right] \underset{D}{=} \frac{1}{\sqrt{2\pi\omega^2}} \int_{-\infty}^{\infty} \frac{e^s}{s} e^{-\frac{s^2}{2\omega^2}} ds = e^{\frac{\omega^2}{2}}$$

- And for variance, see [1] for simplification of the integral:

$$Var\left(e^{\eta}\right) \underset{D}{=} \frac{1}{\sqrt{2\pi\omega^2}} \int_{-\infty}^{\infty} \frac{e^{2s}}{s} e^{-\frac{s^2}{2\omega^2}} ds - \left(E\left[e^{\eta}\right]\right)^2 = e^{\omega^2}\left(e^{\omega^2} - 1\right)$$

- From the latter the expression for CV follows:

$$CV\left(e^{\eta}\right) \underset{D}{=} \frac{1}{E\left[e^{\eta}\right]} \sqrt{Var\left(e^{\eta}\right)} = \frac{1}{e^{\frac{\omega^2}{2}}} \sqrt{e^{\omega^2}\left(e^{\omega^2} - 1\right)} = \sqrt{\left(e^{\omega^2} - 1\right)}$$

- The confidence interval of the CV can be calculated by applying the same transformation on the confidence interval of the variance
- This expression for CV approximated by a Taylor expansion:

$$CV\left(e^{\eta}\right) \approx \sqrt{\omega^2} \qquad (\text{if } \omega \leq 0.1)$$

- This derivation is also valid for models build on log-transformed data with additive error (transform-both-sides approach, TBS)
- When an error term is estimated indirectly with a THETA and a fixed random effect (e.g. $SIGMA 1 FIX), be careful to quadrate the standard deviation term before applying the formulas above

Discussions

- The default expression of CV% for lognormal variability
  `sqrt( OMEGA )`
  leads to unacceptable bias when variability is high, see Figures 1 and 2.
- An expression that correctly calculates CV% is available:
  `sqrt( exp(OMEGA) – 1 )`
- Pharmacometricians therefore should consider always applying the correct equation.
- Reports and papers on modeling should indicate how CV% values are calculated to avoid misinterpretation.

Interpretation of CV%

- Even the correct CV% is not directly interpretable on an arrhythmic scale, see Figure 3.
  – *It represents the lognormal distribution, i.e. it applies on a geometric scale.*
- An improved practice would be to tabulate relative values at informative fractions of the population, e.g. the 5th and 95th percentile.
  – *Such an approach can be applied for any parametric or semi-parametric transformation of ETA, e.g. a logit transformation*
- For example in R, for a lognormal OMEGA of 0.6:
  `exp( qnorm( c( .05, .95 ), sd=sqrt( .6 ) ))` or
  `qlnorm( c( .05, .95 ), sd=sqrt( .6 ) )`
  report 28 and 358% for the 5th and 95th percentile, respectively. This range includes 90% of the population.

Consequences for structural model

- A finding with consequences for model development: the population mean of e^{ETAi} in NONMEM is <u>not</u> equal to one but to $e^{OMEGAi/2}$.
  – *NB: TBS similarly implies a model for median, not mean, observations.*
- The difference is compensated in the corresponding THETA.
- This difference is already ~10% when OMEGA is 0.2.
- It is therefore advisable, under conditional estimation, to specify a THETA within the EXP() expression.
  – *This THETA can be interpreted as a median independent of OMEGA value.*
  – *It additionally has improved estimation and simulation properties.*
- NB: the FO method inherently approximates e^{ETAi} by (1 + ETA_i) and therefore is not affected, see [2].

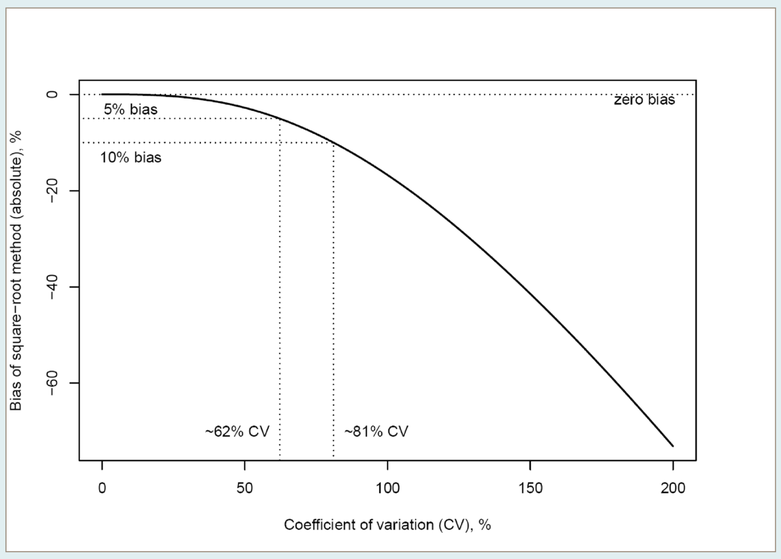

Figure 2. Bias of the square-root approximation plotted as a function of the true coefficient of variation. Bias is calculated absolutely, i.e. defined as the approximate CV minus the true CV value. Dotted lines indicate 5 and 10% bias values.

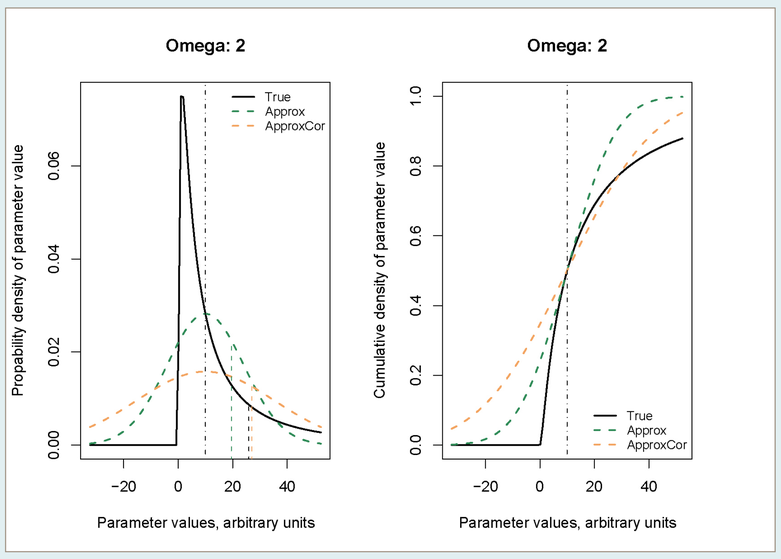

Figure 3. Effect of the approximating the lognormal distribution with large variability. The square-root method clearly underestimates the standard deviation. The approximation with the correct standard deviation, CV=sqrt(exp(OMEGA)-1), clearly improves the cumulative density above the mean. Moreover, the 75th percentile stays consistently close to the true distribution. Large inconsistencies are however visible below the mean. Plotting conventions as in Figure 1.

Conclusion

- **So the answer is yes, the community of pharmacometricians can and should explicitly use the correct expression for the coefficient of variation, $\sqrt{\left(e^{\omega^2} - 1\right)}$, instead of its Taylor approximation when a lognormal distribution has been modeled.**
- **An improved practice would however be to report relative values of the typical value estimate at informative fractions, e.g. at the 5th and 95th percentiles of the distribution.**
- **The lognormal distribution can be approximated by a proportional or normal distribution only when variance is low (CV smaller than ~50% or OMEGA smaller than ~0.2).**

Schering-Plough

References
1. Introduction to the theory of statistics AM Mood, FA Graybill DC Boes, 1974,McGraw-Hill
2. NONMEM Users guide, part VI, chapter III, section D

0922001