# Supplementary Material for: Measuring Precision in Bioassays: Rethinking Assay Validation

by Michael P. Fay, Michael C. Sachs, and Kazutoyo Miura

## Summary

As a convenience for the reader, we collect some basic statistical results from normal models that are used in the main paper. The only non-standard result is that we have defined and derived the log-centered confidence intervals for the constant coefficient of variation model.

# S1 Normal Constant Standard Deviation Model

## S1.1 Tests and One-sided Confidence Limits on $\sigma$

Let $Y_1, \ldots, Y_n$ be independent with $Y_i \sim N(\mu, \sigma^2)$. Let $s^2$ be the usual unbiased variance estimate, and let $S^2$ be the associated random variable. Then

$$\frac{(n-1)S^2}{\sigma^2} \sim W_{n-1}, \tag{S1}$$

where $W_{n-1}$ is the cumulative distribution of a chi-squared distribution with $n-1$ degrees of freedom. Then a $100(1-\alpha)\%$ one-sided confidence interval for $\sigma^2$ is

$$\left[0, \frac{(n-1)s^2}{W_{n-1}^{-1}(\alpha)}\right], \tag{S2}$$

where $W_{n-1}^{-1}(a)$ be the $a$th quantile of that chi-squared distribution.

## S1.2 Confidence Intervals on $\mu$ When $\sigma$ is Known

If $\sigma$ is known then a $100(1-\alpha)\%$ confidence interval for $\mu$ is

$$C_\mu(y; 1-\alpha, \sigma) \;\; = \;\; \left[y - \Phi^{-1}(1-\alpha/2)\sigma, \quad y + \Phi^{-1}(1-\alpha/2)\sigma\right]. \tag{S3}$$

where $\Phi^{-1}(a)$ is the $a$th quantile of the standard normal distribution.

# S2 Normal Constant Coefficient of Variation Model

## S2.1 Testing and One-sided Limits for $\theta$

Consider the normal constant CV model, where $Y \sim N(\mu, \sigma^2)$, and the coefficient of variation is $\theta = \sigma/\mu$. We assume $\mu > 0$. Observe a sample of $n$ independent observations, $y_1, \ldots, y_n$.

Let $\hat{\theta} = \frac{\hat{\sigma}}{\bar{y}}$, where $\hat{\sigma}^2$ is the unbiased sample variance estimate and $\bar{y}$ is the sample mean. Then $T = \sqrt{n}/\hat{\theta}$ is distributed non-central t with $n-1$ degrees of freedom and non-centrality parameter $\sqrt{n}/\theta$. To test $H_0 : \theta \geq \theta_0$ versus $H_1 : \theta < \theta_0$, we reject at the $\alpha$ level when $T > t^{-1}_{n-1,\delta}(1-\alpha)$, where $\delta = \sqrt{n}/\theta_0$ and $t^{-1}_{df,ncp}(a)$ is the $a$th quantile of the non-central t with $df$ degrees of freedom and non-centrality parameter $ncp$. This is equivalent to rejecting when

$$\hat{\theta} \leq \frac{\sqrt{n}}{t^{-1}_{n-1,\delta}(1-\alpha)}.$$

A $100(1-\alpha_\ell)$% upper confidence limit for $\theta$ is the value $\bar{\theta}$ corresponding to the $\theta_0$ that just barely rejects, in other words the value $\bar{\theta}$ that solves,

$$t^{-1}_{n-1,\sqrt{n}/\bar{\theta}}(1-\alpha_\ell) = \frac{\sqrt{n}}{\hat{\theta}}. \tag{S4}$$

## S2.2 Log-centered CIs on $\mu$ When $\theta$ is Known

In this section, assume that $\theta$ is known. We derive a $100(1-\alpha)$% confidence interval for $\mu$ using the intersection of two one-sided confidence intervals,

$$C_\mu(y; 1-\alpha, \theta, f) = \left[\underline{\mu}(y; 1-\alpha_L, \theta), \bar{\mu}(y; 1-\alpha_U, \theta)\right].$$

where $\alpha_L + \alpha_U = \alpha$. First, a $(1-\alpha_L)$ one-sided lower confidence limit is

$$\underline{\mu}(y; 1-\alpha_L, \theta) = \frac{y}{\theta\Phi^{-1}(1-\alpha_L)+1} \tag{S5}$$

where $\Phi^{-1}(q)$ is the $q$th quantile of a standard normal. This can be shown using the fact that since $Y \sim N(\mu, \sigma^2)$ and we have assumed $\mu > 0$, then $Y/\mu \sim N(1, \theta^2)$. Thus,

$$\begin{aligned} 1-\alpha_L &= Pr\left[\frac{\frac{Y}{\mu}-1}{\theta} \leq \Phi^{-1}(1-\alpha_L)\right] \\ &= Pr\left[\frac{Y}{\mu} \leq \theta\Phi^{-1}(1-\alpha_L)+1\right] \\ &= Pr\left[\frac{Y}{\theta\Phi^{-1}(1-\alpha_L)+1} \leq \mu\right]. \end{aligned}$$

To derive the $100(1-\alpha_U)$% upper limit, first consider the case where $\theta\Phi^{-1}(1-\alpha_U) < 1$, so that

$$\begin{aligned} 1-\alpha_U &= Pr\left[-\Phi^{-1}(1-\alpha_U) \leq \frac{\frac{Y}{\mu}-1}{\theta}\right] \\ &= Pr\left[-\theta\Phi^{-1}(1-\alpha_U)+1 \leq \frac{Y}{\mu}\right] \\ &= Pr\left[\mu \leq \frac{Y}{1-\theta\Phi^{-1}(1-\alpha_U)}\right]. \end{aligned}$$

so that (when $\theta\Phi^{-1}(1-\alpha_U) < 1$)

$$\bar{\mu}(y; 1-\alpha_U, \theta) = \frac{y}{1-\theta\Phi^{-1}(1-\alpha_U)} \tag{S6}$$

Let $\alpha_L = f\alpha$ and $\alpha_U = (1-f)\alpha$, where $f \in (0,1)$. Then

$$C_\mu(y; 1-\alpha, \theta, f) = \left[\frac{y}{1+\theta\Phi^{-1}(1-\alpha_L)}, \quad \frac{y}{1-\theta\Phi^{-1}(1-\alpha_U)}\right] \tag{S7}$$

if $\theta\Phi^{-1}(1-\alpha_U) < 1$, otherwise we define the upper limit as $\infty$ and set $f = 1$. If $f = 0.5$ so that $\alpha_L = \alpha_U = \alpha/2$, we call this interval a central one, meaning that both one-sided error rates are equal and bounded at $\alpha/2$.

We define a log-centered interval, as a confidence interval for $\log(\mu)$ of the form $\log(y) \pm r$. Taking the log of both confidence limits in equation S7, we get

$$\left[\log(y) - \log\left\{1+\theta\Phi^{-1}(1-\alpha_L)\right\}, \log(y) - \log\left\{1-\theta\Phi^{-1}(1-\alpha_U)\right\}\right]$$

or equivalently,

$$\left[\log(y) - \log\left\{1+\theta\Phi^{-1}(1-\alpha f)\right\}, \log(y) + \log\left\{\frac{1}{1-\theta\Phi^{-1}(1-\alpha[1-f])}\right\}\right].$$

In other words, if there is an $f$ that gives

$$1+\theta\Phi^{-1}(1-\alpha f) = \frac{1}{1-\theta\Phi^{-1}(1-\alpha[1-f])},$$

then we define that $f$ as $f_{logc}$ and get a log-centered interval with

$$r(\theta, 1-\alpha) = \log\left\{1+\theta\Phi^{-1}(1-\alpha f_{logc})\right\} = \log\left\{\frac{1}{1-\theta\Phi^{-1}(1-\alpha[1-f_{logc}])}\right\}. \tag{S8}$$

If a log-centered interval exists, it will not be a central interval.

As mentioned previously, if $\theta\Phi^{-1}(1-\alpha) > 1$ then we cannot get a log-centered interval and we set $\bar{\mu} \equiv \infty$. Additionally, if $f$ gets very close to 0, then the lower limit approaches 0. So we may have $\underline{\mu} = 0$ due to computer rounding, in which case we cannot write the confidence interval as a log-centered one.

# S3 Lognormal Constant Coefficient of Variation Model

## S3.1 Confidence Interval on $\theta$ from Log Normal Model with Constant Coefficient of Variation

Assume $Y_1, \ldots, Y_n$ are independent and all distributed lognormal, so that $Z_i = \log(Y_i) \sim N(\xi, \nu)$. In terms of $mu$ and $\sigma$ we have

$$E(Y_i) = \mu = \exp(\xi + \frac{\nu}{2}) \qquad \text{and} \qquad Var(Y_i) = \sigma^2 = \mu^2\left\{\exp(\nu) - 1\right\}$$

so that

$$\theta \;=\; \frac{\sigma}{\mu} = \sqrt{\exp(\nu) - 1}. \tag{S9}$$

Then since $Z_1, \ldots, Z_n$ are normal with variance $\nu$, and since $\theta$ is a monotonic function of $\nu$, we can use the normal methods (see Section S1.1) to test the null that $\theta \geq \theta_0$. In other words, the null, $H_0 : \theta \geq \theta_0$, can be equivalently be written as the null, $H_0 : \sqrt{\exp(\nu) - 1} \geq \sqrt{\exp(\nu_0) - 1}$ where $\theta_0 = \sqrt{\exp(\nu_0) - 1}$. Or equivalently, as $H_0 : \nu \geq \nu_0$.

Let $s_z^2$ be the sample variance of the $Z_i$. Then following Section S1.1 we reject the null (in any of its parameterizations) at the $\alpha_\ell$ level whenever

$$s_z^2 \;\leq\; \frac{\nu_0 W_{n-1}^{-1}(\alpha_\ell)}{n-1}$$

or equivalently

$$\tilde{\theta} = \sqrt{\exp(s_z^2) - 1} \;\leq\; \sqrt{(\theta_0^2 + 1)^{\frac{W_{n-1}^{-1}(\alpha_\ell)}{n-1}} - 1}. \tag{S10}$$

When $\frac{W_{n-1}^{-1}(\alpha)}{n-1} = 1$ or equivalently, when $\alpha = \alpha_\ell = W_{n-1}(n-1)$ then this is the same as rejecting when $\tilde{\theta} \leq \theta_0$. Thus, a $(1-\alpha_\ell)\%$ one-sided confidence limit on $\theta$ is the value of $\theta_0$ that just barely rejects. In other words,

$$\bar{\theta}(1-\alpha_\ell) \;=\; \sqrt{-1 + exp\left(\frac{(n-1)s_z^2}{W_{n-1}^{-1}(\alpha_\ell)}\right)}. \tag{S11}$$

## S3.2 Confidence Interval on $\mu$ from Log Normal Model with Known Constant Coefficient of Variation

Now consider the confidence interval for $\mu$ when $\nu$ is known. First, let $\underline{\xi}(z) = z - \sqrt{\nu}\Phi^{-1}(1-\alpha_L)$ and $\bar{\xi}(z) = z + \sqrt{\nu}\Phi^{-1}(1-\alpha_U)$ be the lower and upper limits of a $100(1-\alpha)\%$ confidence interval for $\xi$, where $\alpha_L + \alpha_U = \alpha$. Then

$$\begin{aligned} 1 - \alpha \;&=\; Pr\left[\underline{\xi}\{Z\} \leq \xi \leq \bar{\xi}\{Z\}\right] \\ &=\; Pr\left[\exp\left(\underline{\xi}\{Z\} + \frac{\nu}{2}\right) \leq \mu \leq \exp\left(\bar{\xi}\{Z\} + \frac{\nu}{2}\right)\right] \end{aligned}$$

since $\mu = \exp(\xi + \nu/2)$. We use $\nu = \log(\theta^2 + 1)$ to get the confidence interval. In other words, the $100(1-\alpha)\%$ confidence interval for $\mu$ is the union of the two one-sided confidence intervals, giving $(\underline{\mu}(y; 1-\alpha_L), \bar{\mu}(y; 1-\alpha_U))$. Using $\alpha_L = f\alpha$ and $\alpha_U = (1-f)\alpha$, we rewriting the intervals as

$$\underline{\mu}(y; 1-\alpha f) \;=\; \exp\{\log(y) - r_L\},$$

where $r_L = \sqrt{\nu}\Phi^{-1}(1-\alpha f) - \nu/2$, and

$$\bar{\mu}(y; 1-\alpha(1-f)) \;=\; \exp\left\{\log(y) - r_U\right\},$$

where $r_U = \sqrt{\nu}\Phi^{-1}(1-\alpha(1-f)) + \nu/2$. We can get a log-centered interval by solving for the $f \in (0, 0.5)$ that solves $r_L = r_U$, or equivalently the $f_{logc}$ that solves

$$\Phi^{-1}(1-\alpha f_{logc}) - \Phi^{-1}(1-\alpha(1-f_{logc})) \;=\; \sqrt{\nu}.$$

Thus, the effective standard deviation of $\log(Y)$ is $r(\theta, 0.6827)$, where

$$r(\theta, 1-\alpha) \;=\; \sqrt{\nu}\Phi^{-1}(1-\alpha(1-f)) + \nu/2, \qquad \text{(S12)}$$

and from equation S9, $\nu = \log(\theta^2 + 1)$.

# S4 Proof that Coverage of Log-Centered Confidence Intervals Do Not Depend on $\mu$ Values

We want to show that, given $\theta$, the following does not depend on $\mu(x_1), \ldots, \mu(m)$ or $\mu^* = \mu(x^*)$:

$$\begin{aligned}\Pi(1-\gamma) &= Pr\left[\log(Y^*) - r\left(\bar{\theta}(\mathbf{Y}, q), 1-\gamma\right) \le \log(\mu^*) \le \log(Y^*) + r\left(\bar{\theta}(\mathbf{Y}, q), 1-\gamma\right)\right] \\ &= Pr\left[\log(Y^*/\mu) - r\left(\bar{\theta}(\mathbf{Y}, q), 1-\gamma\right) \le 0 \le \log(Y^*/\mu) + r\left(\bar{\theta}(\mathbf{Y}, q), 1-\gamma\right)\right].\end{aligned}$$

The proof is completed if we show three things: (i) the distribution of $Y^*/\mu^*$ only depends on $\theta$, (ii) the distribution of $\bar{\theta}(\mathbf{Y}, q)$ only depends on $\theta$, and $r(\theta, 1-\gamma)$ only depends on $\theta$. We take these one-at-a-time.

**(i)** Show the distribution of $Y^*/\mu$ only depends on $\theta$. We do this separately for the two models.

**Normal constant CV model:** We have $Y^* \sim N(\mu^*, \sigma^2(x^*))$. Dividing by $\mu^*$ we get $Y^*/\mu^* \sim N(1, \sigma^2(x^*)/\mu^{*2}) = N(1, \theta^2)$.

**Lognormal constant CV model:** We have $\log(Y^*) \sim N(\xi(x^*), \nu) = N(\log(\mu^*) - \nu(x^*)/2, \nu(x^*))$, so that $\log(Y^*) - \log(\mu^*) = \log(Y^*/\mu^*) \sim N(-\nu(x^*)/2, \nu(x^*))$. Finally, $\nu(x^*) = \log(\theta^2 + 1)$.

**(ii)** Show the distribution of $\bar{\theta}(\mathbf{Y}, q)$ only depends on $\theta$. Since $\bar{\theta}(\mathbf{Y}, q) = \max_{j=1,\ldots,m} \bar{\theta}_j(\mathbf{Y}_j, 1-\alpha_\ell)$ where $\alpha_\ell = (1-q)^m$, we just need to show that $\bar{\theta}_j(\mathbf{Y}_j, 1-\alpha_\ell)$ for each $j$ depends only on $\theta$. We do this separately for the two models.

**Normal constant CV model:** Equation S4 gives $\bar{\theta}_j(\mathbf{Y}_j, 1-\alpha_\ell)$ for any level $j$ (the $j$s are suppressed in that section). The only random variable in equation S4 is $\hat{\theta}$, and we know that the distribution of $T = \sqrt{n}/\hat{\theta}$ is non-central t with $n-1$ degrees of freedom and non-centrality parameter $\sqrt{n}/\theta$. So the distribution of $\hat{\theta}$ only depends on $n$ and $\theta$.

**Lognormal constant CV model:** From equation S11 (the $j$s are suppressed in that section), we see that the only random variable that $\bar{\theta}_j(\mathbf{Y}_j, 1-\alpha_\ell)$ depends on is $S_z^2$ which is the sample variance of $Z_1, \ldots, Z_n$ where for each $i = 1, \ldots, n$, $Z_i \sim N(\xi(x_j), \nu(x_j))$. By standard results showing the independence of the sample variance from the mean, $S_z^2$ depends only on $\nu(x_j) = \log(\theta^2 + 1)$.

**(iii)** Show $r(\theta, 1-\gamma)$ only depends on $\theta$.

**Normal constant CV model:** From equation S8 we see that $r(\theta, 1-\gamma)$ only depends on $\theta$, since the calculation of $f_{logc}$ only uses $\theta$.

**Lognormal constant CV model:** From equation S12 we see that $r(\theta, 1-\gamma)$ only depends on $\theta$, since $\nu = \log(\theta^2 + 1)$ and the calculation of $f$ only uses $\nu$.