Technical Report on "The assessment of non-inferiority in a gold standard design with censored, exponentially distributed endpoints"

M. Mielke

Department of Mathematical Stochastics, University Göttingen
e-mail: mmielke@math.uni-goettingen.de

A. Munk

Department of Mathematical Stochastics, University Göttingen
e-mail: munk@math.uni-goettingen.de

A. Schacht

Lilly Deutschland GmbH, Bad Homburg
e-mail: schacht_alexander@lilly.com

1. Introduction

In this paper, we present some technical details required by Mielke et al. [1]. They derived the Wald-type test for the assessment of non-inferiority in a three-armed model with exponentially distributed, right censored endpoints. The presented test procedures are based on the asymptotic normality of the ML-estimators. In *Theorem A.1* the asymptotic normality is derived by means of the classical theory of exponential families. *Theorem A.2* provides the optimal allocation to the three groups in terms of minimizing the resulting asymptotic variance of the ML-estimator for a given total sample size. We start in section 2.1 with an introduction of the model considered by [1].

2. Model and Proofs of Theorems

2.1. Model and Hypothesis

The following introduction of the model is parallel to Section 2 of [1]. We are concerned with a three-armed clinical trial where one observes T_{ki}, $i = 1, \dots, n_k$, independent and exponentially distributed survival times with parameters λ_k, $k = R, T, P$, where R, T, and P abbreviates reference, treatment and placebo group, respectively. To fix the relation of the parameters λ_k and the distribution of the survival times we assume that $\mathrm{E}[T_{ki}] = \lambda_k$. Further, let the corresponding censoring times U_{ki} be independent distributed according to G_k, where U_{ki} is independent of T_{ki} for $i = 1, \dots, n_k$ and $k = R, T, P$. The observations consist of pairs (X_{ki}, δ_{ki}), where $X_{ki} = \min\{T_{ki}, U_{ki}\}$ are the observed survival times and $\delta_{ki} = \mathbf{1}_{\{T_{ki} \le U_{ki}\}}$, $i = 1, \dots, n_k$, $k = R, T, P$, are the corresponding censoring indicators. Hence, $\delta_{ki} = 1$ stands for an uncensored observation. Moreover, none of the groups should asymptotically vanish, i.e. for $k = R, T, P$ and $n = n_R + n_T + n_p$

$$\frac{n_k}{n} \longrightarrow w_k \tag{1}$$

holds for $n_R, n_T, n_p \to \infty$ and some $w_k \in (0, 1)$. Further, we assume that the probabilities for an uncensored observation should be positive, i.e.

$$p_k := P(\delta_{ki} = 1) > 0,$$

imsart-generic ver. 2007/04/13 file: mielke2008_TechnicalReport.tex date: April 29, 2008

for $k = R, T, P$.

Mielke et al. [1] consider the assessment of non-inferiority of a new treatment to a reference one in terms of a retention of effect hypothesis on the log scale, i.e.

$$\begin{array}{ll} & H_0^N : \log \lambda_T - \log \lambda_P \geq \Delta \left(\log \lambda_R - \log \lambda_P \right) \\ \text{vs.} & K_0^N : \log \lambda_T - \log \lambda_P < \Delta \left(\log \lambda_R - \log \lambda_P \right) \end{array} \tag{2}$$

with $\Delta \in [0, \infty)$. The alternative K_0^N means that the test treatment T achieves more than $\Delta \times 100\%$ of the active control effect, where both are compared to placebo and the effect is measured via the log relative risk (cf. Rothmann et al. [2]). The hypothesis (2) is equivalent to

$$H_0^N : \eta := \log \lambda_T - \Delta \log \lambda_R + (\Delta - 1) \log \lambda_P \geq 0\,. \tag{3}$$

The ML-estimator for η is given by

$$\hat{\eta} = \log \hat{\lambda}_T - \Delta \log \hat{\lambda}_R + (\Delta - 1) \log \hat{\lambda}_P \tag{4}$$

by plugging in the ML-estimators

$$\hat{\lambda}_k = \frac{\sum_{i=1}^{n_k} X_{ki}}{\sum_{i=1}^{n_k} \delta_{ki}}$$

for $k = R, T, P$.

2.2. Asymptotic Normality of the MLE and Optimal Allocation

Asymptotic Normality of the ML-estimator, Theorem A.1: The ML-estimator $\hat{\eta}$ given in (4) is asymptotically normally distributed, i.e. $\sqrt{n}\,(\hat{\eta} - \eta) \xrightarrow{\mathfrak{D}} \mathcal{N}(0, \sigma^2)$ with variance

$$\sigma^2 = \frac{1}{w_T\, p_T} + \frac{\Delta^2}{w_R\, p_R} + \frac{(\Delta - 1)^2}{w_P\, p_P}. \tag{5}$$

Proof. For $k = R, T, P$ the density for an observation (X_{ki}, δ_{ki}) can be written as

$$h_k(\lambda_k, x, \delta) = \lambda_k^{-\delta}\, \mathrm{e}^{-\frac{x}{\lambda_k}}\, \tilde{h}_k(x, \delta)$$

with $\tilde{h}_k(x, \delta) = g_k(x)^{1-\delta}(1 - G_k)^{\delta} \mathbf{1}_{\{x \geq 0\}}$, g_k and G_k the density and distribution function, respectively, of the censoring times U_{ki}. Hence, the densities can be written as an exponential family

$$h_k(\lambda_k, x, \delta) = \mathrm{e}^{-\frac{x}{\lambda_k} - \delta \log \lambda_k}\, \tilde{h}_k(x, \delta).$$

Therefore the required regularity conditions to obtain asymptotic normality of $\hat{\lambda}_k$ are satisfied and we have (confer for example [3, Theorem 4.6])

$$\sqrt{n_k}\,(\hat{\lambda}_k - \lambda_k) \xrightarrow{\mathfrak{D}} \mathcal{N}(0, I_k^{-1})$$

with Fisher information matrices I_k, which can be computed by

$$I_k = -E_{\lambda_k}\left[\frac{\partial^2}{\partial^2 \lambda_k} \log h_k(\lambda_k, X, \delta) \right] = \frac{p_k}{\lambda_k^2}$$

with $p_k = P(\delta_{ki} = 1) > 0$ by assumption. Consequently, the transformed ML-estimator $\sqrt{n_k}(\log \hat{\lambda}_k - \log \lambda_k)$ has asymptotic variance p_k^{-1}, which yields together with the independence of the groups and (1), that $\sqrt{n}(\hat{\eta} - \eta)$ is asymptotically normal with variance σ^2 (5).

□

Thus, based on Theorem A.1 we reject H_0^N for a given significance level α if

$$\sqrt{n}\,\frac{\hat{\eta}}{\hat{\sigma}} \leq z_\alpha \,,$$

where z_α denotes the α-quantile of the standard normal distribution and $\hat{\sigma}^2$ is a consistent estimator for σ^2. Based on Theorem A.1 Mielke et al. [1] point out that H_0^N is rejected with a probability of at least $1-\beta$, if for the total sample size n

$$n \;\geq\; \frac{\sigma^2}{\eta^2}\;(z_\alpha - z_{1-\beta})^2 \tag{6}$$

holds, whereas the significance level α, λ_k, p_k, Δ and hence also η given in (3) are pre-specified in planning a clinical trial. Thus, each term on the right hand side other than σ^2 is fixed. The asymptotic variance σ^2 (5) depends on the allocation of the samples. Theorem A.2 presents the optimal allocation in terms of minimizing σ^2 and therewith the total required sample size n, confer (6).

Optimal allocation, Theorem A.2: The asymptotic variance σ^2 in (5) is minimized in $W = \left\{(w_R, w_P) \in [0,1]^2 : w_R + w_P \leq 1\right\}$ for

$$w_R^* = \frac{\Delta\, p_P^{-1}}{p_T^{-1} + \Delta\, p_R^{-1} + |1-\Delta|\, p_P^{-1}} \quad \text{and} \quad w_P^* = \frac{|1-\Delta|\, p_P^{-1}}{p_T^{-1} + \Delta\, p_R^{-1} + \,|1-\Delta|\, p_P^{-1}}\,.$$

Proof. Equating partial derivatives of σ^2 with zero gives

$$\frac{\partial}{\partial w_R}\,\sigma^2 \;=\; \frac{1}{p_T^2(1-w_R-w_P)^2} - \frac{\Delta^2}{p_R^2 w_R^2} = 0 \tag{7}$$

$$\frac{\partial}{\partial w_P}\,\sigma^2 \;=\; \frac{1}{p_T^2(1-w_R-w_P)^2} - \frac{(1-\Delta)^2}{p_P^2 w_P^2} = 0\,. \tag{8}$$

This yields a polynomial of degree four and four possible roots,

$$\begin{aligned}
(w_R, w_P) \;=\; & ((1-\Delta)\, p_P^{-1} + \Delta\, p_R^{-1} - p_T^{-1})^{-1}\cdot(\Delta\, p_R^{-1}, (1-\Delta)\, p_P^{-1})\,,\\
& ((\Delta-1)\, p_P^{-1} + \Delta\, p_R^{-1} - p_T^{-1})^{-1}\cdot(\Delta\, p_R^{-1}, (\Delta-1)\, p_P^{-1})\,,\\
& ((1-\Delta)\, p_P^{-1} + \Delta\, p_R^{-1} + p_T^{-1})^{-1}\cdot(\Delta\, p_R^{-1}, (1-\Delta)\, p_P^{-1})\,,\\
& ((\Delta-1)\, p_P^{-1} + \Delta\, p_R^{-1} + p_T^{-1})^{-1}\cdot(\Delta\, p_R^{-1}, (\Delta-1)\, p_P^{-1})\,.
\end{aligned}$$

However,

$$(w_R^*, w_P^*) \;=\; \left(\frac{\Delta\, p_R^{-1}}{p_T^{-1} + \Delta\, p_R^{-1} + |1-\Delta|\, p_P^{-1}}\,,\, \frac{|1-\Delta|\, p_P^{-1}}{p_T^{-1} + \Delta\, p_R^{-1} + \,|1-\Delta|\, p_P^{-1}}\right)$$

is the only solution to (7) and (8) contained in W. Finally, the Hessian matrix of σ^2 with respect to (w_R, w_P) at (w_R^*, w_P^*) is positive definite since

$$\frac{\partial^2}{\partial^2 w_R}\,\sigma^2(w_R^*, w_P^*) = \frac{2\, p_T\, p_R (p_T^{-1} + \Delta p_R^{-1})\left(p_T^{-1} + \Delta\, p_R^{-1} + |1-\Delta|\, p_P^{-1}\right)^3}{\Delta} \;>\; 0$$

and the determinant of the Hessian matrix is equal to

$$\frac{4\, p_T\, p_R\, p_P\left(p_T^{-1} + \Delta\, p_R^{-1} + |1-\Delta|\, p_P^{-1}\right)^7}{|1-\Delta|\,\Delta} \;>\; 0\,.$$

Hence a local minimum is attained at (w_R^*, w_P^*), which is also the global minimum in W, because it is the only stationary point in W.

□

References

[1] M. Mielke, A. Munk, and A. Schacht. Planning and assessing non-inferiority in a gold standard design with censored, exponentially distributed endpoints. *Statistics in Medicine*, Accepted.

[2] M. Rothmann, N. Li, G. Chen, G. Chi, R. Temple, and H.-H. Tsou. Design and analysis of non-inferiority mortality trials in oncology. *Statistics in Medicine*, 22:239–264, 2003.

[3] A. W. van der Vaart. *Asymptotic statistics*, volume 3 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, 1998.