Biometrical Journal **60** (2018) zzz, zzz–zzz / DOI: 10.1002/bimj.200100000

Simultaneous confidence sets for several effective doses

Daniel M. Tompsett [1], **Stefanie Biedermann**[*2], and **Wei Liu** [2]

[1] MRC Biostatistics Unit, Cambridge Institute of Public Health, Forvie Site, Robinson Way, Cambridge Biomedical Campus, Cambridge CB2 0SR, UK

[2] Mathematical Sciences, University of Southampton, Highfield Campus, University Road, Southampton SO17 1BJ, UK

Received September 2017, revised January 2018, accepted zzz

Construction of simultaneous confidence sets for several effective doses currently relies on inverting the Scheffé type simultaneous confidence band, which is known to be conservative. We develop novel methodology to make the simultaneous coverage closer to its nominal level, for both two-sided and one-sided simultaneous confidence sets. Our approach is shown to be considerably less conservative than the current method, and is illustrated with an example on modeling the effect of smoking status and serum triglyceride level on the probability of the recurrence of a myocardial infarction.

Key words: Logistic Regression; Effective Dose; Simultaneous Confidence Band; Simultaneous Inference; Inverse Dose Response Curve;

1 Introduction

Logistic regression has wide applications in dichotomous response studies, relating various stimuli to the probability of observing a response in a subject; see its application to medicine and biology statistics in Finney (1971), to quantitative risk and hazard assessment in Piegorsch et al. (2005) and Peng et al. (2015), and to drug dose response curves in Carter et al. (1986) and Bretz et al. (2008). A key goal in any such study is to identify the stimulus, usually the dose of a substance, needed to elicit a particular probability of response in a subject based on the logistic model, known as the Effective Dose (ED) or Lethal Dose.

Construction of confidence sets for a single ED is well studied. The two most notable methods are due to Fieller (1954) and Cox (1990). Construction methods for confidence sets for multidimensional single EDs are provided, for example, by Li, Zhang, Nordheim & Lehner (2008) in a Bayesian framework, and by Li et al. (2010) and Li & Wong (2011) using asymptotic theory and bootstrap methods. However, it is often of interest to identify multiple EDs at once. For example when looking to identify a minimal effective dose and maximum safe dose, or when information on safe doses for weaning patients off of a drug is warranted. The focus of this paper is therefore on *simultaneous* inference, establishing confidence sets for several EDs such that their simultaneous confidence level is at least some nominal $1 - \alpha$.

Currently such confidence sets are established by inverting the Scheffé type simultaneous confidence band (cf. Carter et al. (1986)), but are well noted (cf. Al-Saidy et al. (2003), Nitcheva et al. (2005), Piegorsch et al. (2005), Li, Nordheim, Zhang & Lehner (2008)) to be unduly conservative in terms of simultaneous confidence level. We therefore develop a method by constructing confidence sets tailored to offer simultaneous coverage for a particular number of EDs at once, which will be less conservative. Section 3 studies two sided confidence sets, and in Section 4 we consider one sided confidence sets. We illustrate our methodology, and demonstrate the improvement, with a real data example in Section 5, where we find simultaneous confidence sets for effective doses of serum triglyceride level for the recurrence of myocardial infarction. While we develop our approach for the logistic model, it is immediately applicable

[*]Corresponding author: e-mail: S.Biedermann@soton.ac.uk, Phone: +44-02380-593672, Fax: +44-02380-595147

to any generalised linear model with an asymptotically normal MLE and arbitrary link function for the linear predictor. This, and several other possible extensions, are further discussed in Section 6.

2 Statistical setting

2.1 The Logistic model and the effective dose

Assume a Bernoulli distributed response variable Y, where the probability of observing a success $(Y = 1)$, p , depends on a vector of q independent explanatory variables $x = (x_1, \ldots, x_q)^\intercal$ via the logistic regression model

$$p(x) = \frac{\exp(\boldsymbol{x}^\intercal \boldsymbol{\beta})}{1 + \exp(\boldsymbol{x}^\intercal \boldsymbol{\beta})} \tag{1}$$

or, equivalently,

$$\pi(p(x)) = \log\left(\frac{p(x)}{1 - p(x)}\right) = \boldsymbol{x}^\intercal \boldsymbol{\beta} \tag{2}$$

where $\boldsymbol{x}^\intercal = (1, x_1, \ldots, x_q) = (1, x^\intercal)$ and $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_q)^\intercal$ is the unknown parameter vector of interest. Note that to distinguish between the regression vector, $\boldsymbol{x}$, in the linear predictor of (1) and a combination of values of the explanatory variables $x = (x_1, \ldots, x_q)^\intercal$, the former is expressed in bold face. In practice, N observations of the response and its corresponding covariate values, denoted as Y_i and $x_i = (x_{i1}, \ldots, x_{iq})^\intercal$, are used to estimate $\boldsymbol{\beta}$ by its maximum likelihood estimator (MLE) $\hat{\boldsymbol{\beta}}$. We assume a sufficiently large N so that we have the large sample asymptotic normality property (cf. Walter (1983), Faraway (2016)),

$$\sqrt{N}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \xrightarrow[N\to\infty]{d} \mathcal{N}_{q+1}(\mathbf{0}, \boldsymbol{\Sigma}),$$

where $\mathcal{N}_{q+1}$ denotes the multivariate normal distribution of dimension $q + 1$. The asymptotic covariance matrix $\boldsymbol{\Sigma}$ can be consistently estimated by $N\boldsymbol{J}^{-1}$ where $\boldsymbol{J}^{-1}$ is the observed covariance matrix of $\hat{\boldsymbol{\beta}}$, which is routinely provided in any statistical software. Therefore, for sufficiently large N, the unknown $\boldsymbol{\Sigma}$ can be replaced by the observable $N\boldsymbol{J}^{-1}$, and we have the approximate distributional result

$$(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \overset{d}{\approx} \mathcal{N}_{q+1}\left(\mathbf{0}, \boldsymbol{J}^{-1}\right). \tag{3}$$

Define the effective dose (ED), x_p, as the value(s) of the covariates required to elicit a **specific** probability of success p,

$$x_p = \left\{x : \boldsymbol{x}^\intercal \boldsymbol{\beta} = \pi(p) = \log\left(\frac{p}{1-p}\right)\right\} = \left\{x : \frac{\exp(\boldsymbol{x}^\intercal \boldsymbol{\beta})}{1 + \exp(\boldsymbol{x}^\intercal \boldsymbol{\beta})} = p\right\}. \tag{4}$$

For a multivariate model ($q > 1$), the ED is a set of x which satisfies the equation in (4). Li, Nordheim, Zhang & Lehner (2008) consider the ED for a single covariate x_i, conditioned on values for the remaining $q-1$ covariates, known as the conditioning effective dose (CED). For example, for given $(x_2^*, \ldots, x_q^*)$, the CED is defined as

$$x_p^{CED} = \left\{x^* = (x_1, x_2^*, \ldots, x_q^*) : \boldsymbol{x}^{*\intercal} \boldsymbol{\beta} = \pi(p) = \log\left(\frac{p}{1-p}\right)\right\}.$$

Point estimation of the ED or CED can be achieved by plugging in the MLE. It is of greater interest, however, to establish a confidence region around the true EDs. For a single effective dose, confidence sets are given in Fieller (1954) and Cox (1990), but they are unsuitable for simultaneous inference.

2.2 The Scheffé Band method

For simultaneous inference, Brand et al. (1973) consider regression models that have only one covariate. This is extended in Walter (1983) to the case of several covariates. Under the asymptotic distribution (3), one can construct (see, e.g., Carter et al. (1986)) a $(1-\alpha)$-level Scheffé type simultaneous confidence band for the logistic regression line of the form

$$\pi(p(x)) = \boldsymbol{x}^\intercal\boldsymbol{\beta} \in \boldsymbol{x}^\intercal\hat{\boldsymbol{\beta}} \pm \sqrt{\chi^\alpha_{q+1}}\sqrt{\boldsymbol{x}^\intercal\boldsymbol{J}^{-1}\boldsymbol{x}}\ \forall\ x \in \mathbb{R}^q \tag{5}$$

where the critical constant $c = \sqrt{\chi^\alpha_{q+1}}$ is set to guarantee that

$$\mathrm{P}\left\{\boldsymbol{x}^\intercal\boldsymbol{\beta} \in \boldsymbol{x}^\intercal\hat{\boldsymbol{\beta}} \pm \sqrt{\chi^\alpha_{q+1}}\sqrt{\boldsymbol{x}^\intercal\boldsymbol{J}^{-1}\boldsymbol{x}}\ \forall\ x \in \mathbb{R}^q\right\} = 1-\alpha.$$

An equivalent $(1-\alpha)$-confidence band for $p(x)$ is then given by

$$p(x) \in \frac{\exp\left(\boldsymbol{x}^\intercal\hat{\boldsymbol{\beta}} \pm \sqrt{\chi^\alpha_{q+1}}\sqrt{\boldsymbol{x}^\intercal\boldsymbol{J}^{-1}\boldsymbol{x}}\right)}{1+\exp\left(\boldsymbol{x}^\intercal\hat{\boldsymbol{\beta}} \pm \sqrt{\chi^\alpha_{q+1}}\sqrt{\boldsymbol{x}^\intercal\boldsymbol{J}^{-1}\boldsymbol{x}}\right)}\ \forall\ x \in \mathbb{R}^q. \tag{6}$$

Originally considered by Scheffé (1953), for a normal-error linear model, any candidate for the true logistic model $\boldsymbol{x}^\intercal\boldsymbol{\beta}$ may be considered plausible only if it lies fully within the band over all $x \in \mathbb{R}^q$.

One may construct simultaneous confidence sets for the EDs or CEDs from either (5) or (6) by inverting the confidence band at a particular value of $\pi(p)$ (or equivalently p). Specifically we define the confidence set, for a particular value of p, as

$$\mathbf{C}_p = \left\{x : \pi(p) \in \boldsymbol{x}^\intercal\hat{\boldsymbol{\beta}} \pm \sqrt{\chi^\alpha_{q+1}}\sqrt{\boldsymbol{x}^\intercal\boldsymbol{J}^{-1}\boldsymbol{x}}\right\} = \left\{x : \frac{|\boldsymbol{x}^\intercal\hat{\boldsymbol{\beta}} - \pi(p)|}{\sqrt{\boldsymbol{x}^\intercal\boldsymbol{J}^{-1}\boldsymbol{x}}} < \sqrt{\chi^\alpha_{q+1}}\right\}. \tag{7}$$

By definition $\mathbf{C}_p$ contains all such x that the Scheffé type confidence band at x includes $\pi(p)$. We refer to this as the Scheffé Band method. It can be shown that the simultaneous coverage probability of **any** k effective doses x_{p_i} lying in the corresponding $\mathbf{C}_{p_i}$ for $i = 1, \ldots, k$ is at least $1-\alpha$ in the following way:

$$\begin{aligned}
&\mathrm{P}\left\{x_{p_i} \in \mathbf{C}_{p_i} \text{ for } i = 1, \ldots, k\right\}\\
=&\mathrm{P}\left\{\frac{|(\boldsymbol{x}_{\boldsymbol{p_i}})^\intercal\hat{\boldsymbol{\beta}} - \pi(p_i)|}{\sqrt{(\boldsymbol{x}_{\boldsymbol{p_i}})^\intercal\boldsymbol{J}^{-1}\boldsymbol{x}_{\boldsymbol{p_i}}}} < \sqrt{\chi^\alpha_{p+1}} \text{ for } i = 1, \ldots, k\right\}\\
\geq&\mathrm{P}\left\{\frac{|(\boldsymbol{x}_{\boldsymbol{p}})^\intercal(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})|}{\sqrt{(\boldsymbol{x}_{\boldsymbol{p}})^\intercal\boldsymbol{J}^{-1}\boldsymbol{x}_{\boldsymbol{p}}}} < \sqrt{\chi^\alpha_{p+1}}\ \forall p \in (0,1)\right\} \qquad (8)\\
=&\mathrm{P}\left\{\frac{|\boldsymbol{x}^\intercal(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})|}{\sqrt{\boldsymbol{x}^\intercal\boldsymbol{J}^{-1}\boldsymbol{x}}} < \sqrt{\chi^\alpha_{q+1}}\ \forall x \in \mathbb{R}^q\right\} \qquad (9)\\
=&\mathrm{P}\left\{\boldsymbol{x}^\intercal\boldsymbol{\beta} \in \boldsymbol{x}^\intercal\hat{\boldsymbol{\beta}} \pm \sqrt{\chi^\alpha_{p+1}}\sqrt{\boldsymbol{x}^\intercal\boldsymbol{J}^{-1}\boldsymbol{x}}\ \forall\ x \in \mathbb{R}^q\right\} = 1-\alpha.
\end{aligned}$$

Here the inequality in (8) occurs since $(\boldsymbol{x}_{\boldsymbol{p_i}})^\intercal\boldsymbol{\beta} = \pi(p_i)$, and the equality in (9) is formed because the set of all effective doses is the same as the set of all $x \in \mathbb{R}^q$. Consequently we must have $\mathrm{P}\left\{x_{p_i} \in \mathbf{C}_{p_i} \text{ for } i = 1, \ldots, k\right\} \geq 1-\alpha$. Clearly these confidence sets rely on the asymptotic distribution in (3); coverage accuracy will depend on the closeness of the distribution of $\hat{\boldsymbol{\beta}}$ to the normal. It is worth emphasizing that all our analytical results become exact if the distributional assumption in (3) is exact.

Note that we are guaranteed at least the nominal coverage $1-\alpha$ **regardless** of the choice of k. However, from the proof above, exact simultaneous coverage $1-\alpha$ is only possible when sets are sought for **every** effective dose at once. Thus for a small k, coverage tends to be significantly larger than $1-\alpha$. A simulation study conducted in Li, Nordheim, Zhang & Lehner (2008) showed that, for small k, coverage sometimes can be much larger than $1-\alpha$. This is also indicated by the studies in Al-Saidy et al. (2003), Piegorsch et al. (2005) and Nitcheva et al. (2005).

In this paper we therefore propose the following adaptation. We construct confidence sets of the form in (7) but where the critical constant c is set so that the minimal simultaneous coverage is guaranteed for a **pre-specified** number, k, of effective doses. By construction, these confidence sets will have closer to nominal confidence level than currently possible.

3 Two-sided simultaneous confidence sets for k effective doses

In this section we establish simultaneous confidence sets of the same form as in (7)

$$\mathbf{C}_p = \left\{ x : \frac{|\boldsymbol{x}^\intercal \hat{\boldsymbol{\beta}} - \pi(p)|}{\sqrt{\boldsymbol{x}^\intercal \boldsymbol{J}^{-1} \boldsymbol{x}}} < c \right\}, \tag{10}$$

where the critical constant c is suitably chosen so that the simultaneous coverage of the $\mathbf{C}_{p_i}$'s $(i = 1, \cdots, k)$ for a given value of k is at least $1-\alpha$. For the case $k = 2$ our result is for a multivariate model (i.e. $q \geq 1$). For $k \geq 3$, our result is for a univariate model with $q = 1$.

3.1 Simultaneous confidence sets for $k = 2$ effective doses

We begin with the simplest situation to construct a confidence set of the form in (10) for $k = 2$. In this case, c is chosen such that for any two effective doses x_{p_1} and x_{p_2} we have

$$\mathrm{P}\left\{x_{p_i} \in \mathbf{C}_{p_i} \text{ for } i = 1, 2\right\} \geq 1 - \alpha. \tag{11}$$

To find the smallest value of c satisfying (11), it is sufficient to find the two effective doses x_{p_1} and x_{p_2} that minimise the probability in (11), for any value of $\boldsymbol{\beta} \in \mathbb{R}^{q+1}$ or $c > 0$, that is

$$\min_{x_{p_1},\, x_{p_2}} \mathrm{P}\left\{x_{p_i} \in \mathbf{C}_{p_i} \text{ for } i = 1, 2\right\} \ \forall \boldsymbol{\beta} \in \mathbb{R}^{q+1}.$$

The value of c is then set to satisfy

$$\min_{x_{p_1},\, x_{p_2}} \mathrm{P}\left\{x_{p_i} \in \mathbf{C}_{p_i} \text{ for } i = 1, 2\right\} = 1 - \alpha.$$

Note that

$$\mathrm{P}\left\{x_{p_i} \in \mathbf{C}_{p_i} \text{ for } i = 1, 2\right\} = \mathrm{P}\left\{ \frac{|(\boldsymbol{x}_{\boldsymbol{p_i}})^\intercal (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})|}{\sqrt{(\boldsymbol{x}_{\boldsymbol{p_i}})^\intercal \boldsymbol{J}^{-1} \boldsymbol{x}_{\boldsymbol{p_i}}}} < c;\ i = 1, 2 \right\} = \mathrm{P}\left\{ |\boldsymbol{Z}_i| < c;\ i = 1, 2 \right\},$$

where $\boldsymbol{Z}_i = \frac{(\boldsymbol{x}_{\boldsymbol{p_i}})^\intercal (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})}{\sqrt{(\boldsymbol{x}_{\boldsymbol{p_i}})^\intercal \boldsymbol{J}^{-1} \boldsymbol{x}_{\boldsymbol{p_i}}}}$ is a standard normal variable due to (3). It follows immediately from Sidak's Inequality (cf. Hsu (1996)) that

$$\min_{x_{p_1},\, x_{p_2}} \mathrm{P}\left\{x_{p_i} \in \mathbf{C}_{p_i} \text{ for } i = 1, 2\right\} = \left(\mathrm{P}\left\{|\boldsymbol{Z}_i| < c\right\}\right)^2,$$

provided there exists x_{p_1} and x_{p_2} for which

$$Cov(\boldsymbol{Z}_1, \boldsymbol{Z}_2) = \frac{(\boldsymbol{x_{p_1}})^\intercal \boldsymbol{J}^{-1} \boldsymbol{x_{p_2}}}{\sqrt{(\boldsymbol{x_{p_1}})^\intercal \boldsymbol{J}^{-1} \boldsymbol{x_{p_1}}} \sqrt{(\boldsymbol{x_{p_2}})^\intercal \boldsymbol{J}^{-1} \boldsymbol{x_{p_2}}}} = 0,$$

i.e. $(\boldsymbol{x_{p_1}})^\intercal \boldsymbol{J}^{-1} \boldsymbol{x_{p_2}} = 0$. Since we allow the effective doses x_{p_1} and x_{p_2} to lie over the whole real space, there must exist x_{p_1} and x_{p_2} such that $(\boldsymbol{x_{p_1}})^\intercal \boldsymbol{J}^{-1} \boldsymbol{x_{p_2}} = 0$. Therefore c is set to satisfy $\{\mathrm{P}\{|\boldsymbol{Z}_i| < c\}\}^2 = 1 - \alpha$ and so $c = z_{\left(1 - \frac{1-\sqrt{1-\alpha}}{2}\right)}$, where z_γ is the γ^{th} quantile of the standard normal distribution.

3.2 Simultaneous confidence sets for $k(\geq 3)$ effective doses

Note that $q = 1$ in this case. We construct confidence sets of the same form as in (10) for a univariate logistic model ($q = 1$), where now c is chosen so that for a specific $k \geq 3$

$$\min_{-\infty < x_{p_1}, \ldots, x_{p_k} < \infty} \mathrm{P}\{x_{p_i} \in \mathbf{C}_{p_i} \text{ for } i = 1, \ldots, k\} = 1 - \alpha, \tag{12}$$

where $-\infty < x_{p_i} < \infty \; \forall i = 1, \ldots, k$. We note that the method for $k = 2$ relies on the independence of the two $\boldsymbol{Z}_k$, which cannot be achieved, however, for $k \geq 3$. Let $\boldsymbol{P}^2 = \boldsymbol{J}^{-1}$ and let $\boldsymbol{N} = \boldsymbol{P^{-1}}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \sim N(\mathbf{0}, \boldsymbol{I}_2)$ due to (3). Note that we may write

$$\begin{aligned}
&\mathrm{P}\{x_{p_i} \in \mathbf{C}_{p_i} \text{ for } i = 1, \ldots, k\} \\
=&\mathrm{P}\left\{ \frac{|(\boldsymbol{x_{p_i}})^\intercal (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})|}{\sqrt{(\boldsymbol{x_{p_i}})^\intercal \boldsymbol{J}^{-1} \boldsymbol{x_{p_i}}}} < c \text{ for } i = 1, \ldots, k \right\} \\
=&\mathrm{P}\left\{ \frac{\left| \left\{ \boldsymbol{P} \begin{pmatrix} 1 \\ x_{p_i} \end{pmatrix} \right\}^\intercal \boldsymbol{N} \right|}{\left\| \boldsymbol{P} \begin{pmatrix} 1 \\ x_{p_i} \end{pmatrix} \right\|} < c \text{ for } i = 1, \ldots, k \right\} \\
=&\mathrm{P}\{\boldsymbol{N} \in \mathbb{V}(x_{p_i}) \text{ for } i = 1, \ldots, k\} = \mathrm{P}\{\boldsymbol{N} \in \mathbb{V}_k\}
\end{aligned}$$

where

$$\mathbb{V}(x_p) = \left\{ \boldsymbol{N} : \frac{\left| \left\{ \boldsymbol{P} \begin{pmatrix} 1 \\ x_p \end{pmatrix} \right\}^\intercal \boldsymbol{N} \right|}{\left\| \boldsymbol{P} \begin{pmatrix} 1 \\ x_p \end{pmatrix} \right\|} < c \right\}$$

is the region given in the $\boldsymbol{N} = (n_1, n_2)^\intercal$-plane by the stripe bounded by the two parallel lines that are perpendicular to the directional vector $\boldsymbol{P} \begin{pmatrix} 1 \\ x_p \end{pmatrix} = \boldsymbol{P}\boldsymbol{x}_p$ and c distance from the origin, and $\mathbb{V}_k = \cap_{i=1}^k \mathbb{V}(x_{p_i})$ is a $2k$-sided polygonal region, depicted in Figure 1 for $k = 4$. Figure 1 (and also Figures 2, 3 and 5) have been created using the Ipe software http://ipe.otfried.org/.

Denote the angle between any two directional vectors $\boldsymbol{P}\boldsymbol{x}_{\boldsymbol{p_i}}$ and $\boldsymbol{P}\boldsymbol{x}_{\boldsymbol{p_j}}$ by θ_{ij} as shown in Figure 1. It is clear from Figure 1 that $\mathrm{P}\{\boldsymbol{N} \in \mathbb{V}_k\}$ is equal to the probability of $\boldsymbol{N}$ in the parallelogram region ABCD less twice the probability of $\boldsymbol{N}$ in the upper right shaded region. Manipulation similar to Liu (2010) (pages

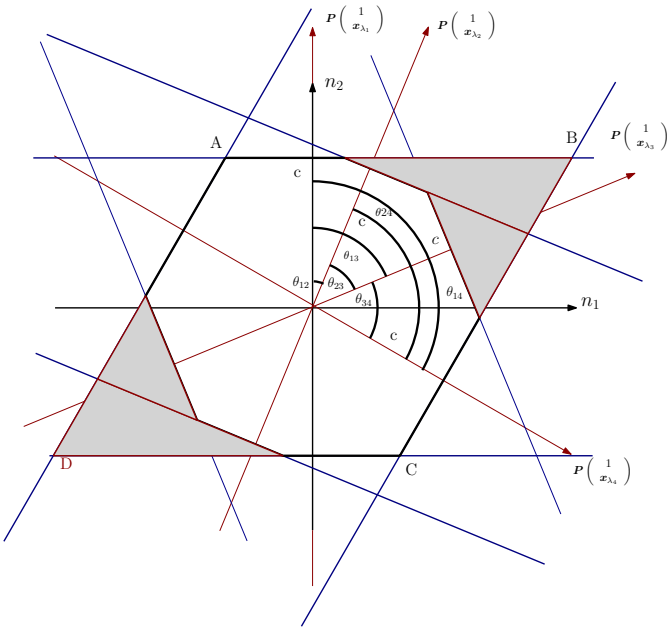

Figure 1 The region $\mathbb{V}_4$, as the parallelogram ABCD less the four grey shaded regions.

36-39) gives that $\mathrm{P}\{\boldsymbol{N} \in \mathbb{V}_k\}$ can be expressed as

$$\frac{1}{\pi}\left\{\int_{\frac{\theta_{1k}-\pi}{2}}^{\frac{\theta_{1k}}{2}}\left(1-\exp\left\{-\frac{c^2}{2cos(\theta)^2}\right\}\right)\mathrm{d}\theta+\int_{\frac{-\theta_{1k}}{2}}^{\frac{\pi-\theta_{1k}}{2}}\left(1-\exp\left\{-\frac{c^2}{2cos(\theta)^2}\right\}\right)\mathrm{d}\theta\right\}$$
$$-2\sum_{i=2}^{k-1}\int_{l_3(i)}^{c}\phi(n_2)\left[\Phi\left(-n_2cot(\theta_{(i-1)k})+\frac{c}{sin(\theta_{(i-1)k})}\right)\right.$$
$$\left.-\Phi\left(-n_2cot(\theta_{(i-1)i})+\frac{c}{sin(\theta_{(i-1)i})}\right)\right]\mathrm{d}n_2 \qquad (13)$$

with $l_3(i) = \frac{c(sin(\theta_{(i-1)k})-\sin(\theta_{(i-1)i}))}{sin(\theta_{(i-1)k}-\theta_{(i-1)i})}$. This is possible since the probability of $\boldsymbol{N} \in \mathbb{V}_k$ is rotation invariant at the origin and therefore for arbitrary x_{p_i} we may assume without loss of generality that the vectors $\boldsymbol{P}\boldsymbol{x}_{\boldsymbol{p_i}}$ are always arranged as in Figure 1.

The expression in (13) clearly depends on the k effective doses through the angles θ_{ij}. For $-\infty < x_{p_i} < \infty$ we have that $0 \leq \theta_{ij} \leq \pi$. Therefore it is sufficient to minimise the expression in (13) with respect to the θ_{ij} under the condition that $0 \leq \theta_{ij} \leq \pi$. This is given by the following theorem.

Theorem 3.1 *The probability* $P\{\boldsymbol{N} \in \mathbb{V}_k\}$ *is minimised with respect to the* θ_{ij} *when* $\mathbb{V}_k$ *takes a regular* $2k$*-sided polygonal shape, that is when* $\theta_{ij} = \frac{(j-i)\pi}{k}$ *for all* $1 \leq i \leq j \leq k$ *.*

The proof of Theorem 3.1 hinges on the following result.

Lemma 1 Suppose that $k-1$ directional vectors $\boldsymbol{P}\boldsymbol{x}_{\boldsymbol{p_i}}$ are fixed, and only one $\boldsymbol{P}\boldsymbol{x}_{\boldsymbol{p_j}}$ is allowed to vary between adjacent $\boldsymbol{P}\boldsymbol{x}_{\boldsymbol{p_{j-1}}}$ and $\boldsymbol{P}\boldsymbol{x}_{\boldsymbol{p_{j+1}}}$. Then $\mathrm{P}\{\boldsymbol{N} \in \mathbb{V}_k\}$ is minimised when $\boldsymbol{P}\boldsymbol{x}_{\boldsymbol{p_j}}$ lies halfway between $\boldsymbol{P}\boldsymbol{x}_{\boldsymbol{p_{j-1}}}$ and $\boldsymbol{P}\boldsymbol{x}_{\boldsymbol{p_{j+1}}}$.

The proof of Lemma 1 is given in the appendix, and the proof of Theorem 3.1 is then immediate by the following argument. Note that the essence of Lemma 1 is we may choose any three adjacent directions

available and always improve $\mathrm{P}\{\boldsymbol{N} \in \mathbb{V}_k\}$ by setting the middle direction halfway in between the other two. Therefore one cannot reduce the probability further only when all $2k$ directions lie equally spaced from each other, which occurs when $\theta_{ij} = \frac{(j-i)\pi}{k}$ as given by Theorem 3.1.

Therefore the value of c which satisfies (12) can be solved numerically from $\mathrm{P}\{\boldsymbol{N} \in \mathbb{V}_k\} = 1-\alpha$ using the expression in (13) with $\theta_{ij} = \frac{(j-i)\pi}{k}$.

4 One-sided simultaneous confidence sets for k effective doses

Often the focus in effective dose problems is to find a worst or best case scenario, that is, the smallest or largest plausible candidate for the ED. In this case upper or lower **one**-sided confidence sets are more informative. These are easily constructed (cf. Deutsch & Piegorsch (2012)) by inverting the upper or lower one sided simultaneous confidence bands. In this section we adapt the methods of Section 3 to construct one sided confidence sets having $1-\alpha$ simultaneous coverage for a specific k. We focus again on the case $q = 1$.

4.1 Simultaneous one-sided sets for $k = 2$ effective doses

We construct a lower one-sided confidence set of the form

$$\mathbf{C}_p^- = \left\{ x : \frac{\boldsymbol{x}^\intercal \hat{\boldsymbol{\beta}} - \pi(p)}{\sqrt{\boldsymbol{x}^\intercal \boldsymbol{J}^{-1} \boldsymbol{x}}} < c \right\}, \tag{14}$$

and want to find the critical constant c so that

$$\min_{x_{p_1}, x_{p_2}} \mathrm{P}\left\{ x_{p_i} \in \mathbf{C}_{p_i}^- \text{ for } i = 1, 2 \right\} = 1 - \alpha.$$

Similar to the two-sided case we have

$$\mathrm{P}\left\{ x_{p_i} \in \mathbf{C}_{p_i}^- \text{ for } i = 1, 2 \right\} = \mathrm{P}\{\boldsymbol{N} \in \mathbb{V}_2^-\}$$

with$\mathbb{V}_2^- \equiv \mathbb{V}(x_{p_1})^- \cap \mathbb{V}(x_{p_2})^-$ where

$$\mathbb{V}(x_p)^- = \left\{ \boldsymbol{N} : \frac{\left\{ \boldsymbol{P} \begin{pmatrix} 1 \\ x_p \end{pmatrix} \right\}^\intercal \boldsymbol{N}}{\left\| \boldsymbol{P} \begin{pmatrix} 1 \\ x_p \end{pmatrix} \right\|} < c \right\}$$

is a region in the $\boldsymbol{N}$-plane, which includes the origin and is bounded by the line perpendicular to $\boldsymbol{P}\boldsymbol{x}_{\boldsymbol{p}}$ with distance c from the origin in the direction of $\boldsymbol{P}\boldsymbol{x}_{\boldsymbol{p}}$. Again, $\mathbb{V}_2^-$ is rotation invariant around the origin. Thus it may be represented as in Figure 2 with the angle between the two direction vectors denoted by θ_3.

Similar to the two-sided case, one can show by using manipulation similar to Liu (2010) that $\mathrm{P}\{\boldsymbol{N} \in \mathbb{V}_2^-\}$ can be expressed as

$$\frac{1}{2\pi} \left\{ \int_{-\frac{\theta_3}{2}}^{\frac{\pi}{2}} \left(1 - \exp\left\{ -\frac{c^2}{2cos(\theta)^2} \right\} \right) \mathrm{d}\theta + \int_{-\frac{\pi}{2}}^{\frac{\theta_3}{2}} \left(1 - \exp\left\{ -\frac{c^2}{2cos(\theta)^2} \right\} \right) \mathrm{d}\theta \right\} + \frac{\pi - \theta_3}{2\pi}. \tag{15}$$

Again, this expression depends on the effective doses through the angle θ_3. If we allow the effective doses to lie over the whole real line, it is sufficient to minimise this expression with respect to $0 < \theta_3 \leq \pi$, which is given by Lemma 2.

Lemma 2 Expression (15) is minimised with respect to $0 < \theta_3 \leq \pi$ at $\theta_3 = \pi$.

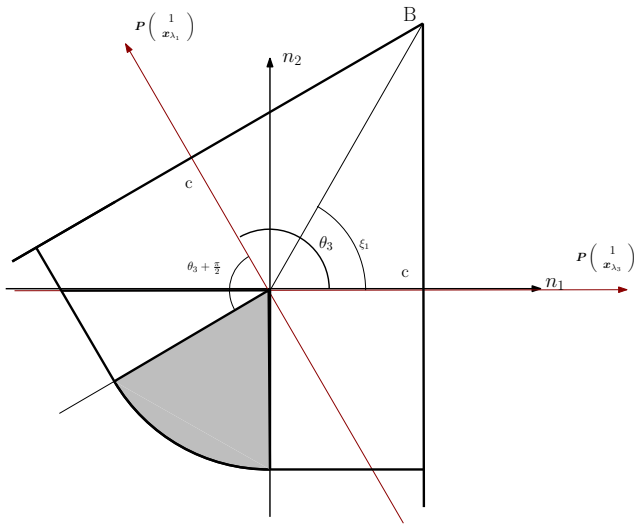

Figure 2 The region $\mathbb{V}_2^-$

P r o o f. We calculate the derivative of expression (15) with respect to θ_3 to get

$$\begin{aligned}
&\frac{1}{2\pi}\left\{\int_{-\frac{\theta_3}{2}}^{\frac{\pi}{2}} \frac{\mathrm{d}}{\mathrm{d}\theta_3}\left(1-\exp\left\{-\frac{c^2}{2cos(\theta)^2}\right\}\right)\mathrm{d}\theta-(-\frac{1}{2})\left(1-\exp\left\{-\frac{c^2}{2cos(-\frac{\theta_3}{2})^2}\right\}\right)\right\}\\
&+\frac{1}{2\pi}\left\{\int_{-\frac{\pi}{2}}^{\frac{\theta_3}{2}} \frac{\mathrm{d}}{\mathrm{d}\theta_3}\left(1-\exp\left\{-\frac{c^2}{2cos(\theta)^2}\right\}\right)\mathrm{d}\theta+\frac{1}{2}\left(1-\exp\left\{-\frac{c^2}{2cos(\frac{\theta_3}{2})^2}\right\}\right)\right\}-\frac{1}{2\pi}\\
=&\frac{1}{2\pi}\left\{1-\exp\left\{-\frac{c^2}{2cos(\frac{\theta_3}{2})^2}\right\}-1\right\}=-\frac{1}{2\pi}\exp\left\{-\frac{c^2}{2cos(\frac{\theta_3}{2})^2}\right\} \qquad (16)
\end{aligned}$$

as a result of cosine being an even function and the Leibniz integral rule. It is clear from (16) that the derivative is negative for all θ_3. Hence (15) is a monotonically decreasing function of $0 < \theta_3 \leq \pi$, and is therefore minimised in $\theta_3 = \pi$ as required.

□

Note that $\mathrm{P}\{\boldsymbol{N} \in \mathbb{V}_2^-\}$ at $\theta_3 = \pi$ is equal to the probability that $\boldsymbol{N}$ lies in the region bounded by the two lines parallel to the n_2 axis with distance c from the origin in the direction of the n_1 axis. Therefore we have

$$\min_{x_{p_1},\, x_{p_2}} \mathrm{P}\{\boldsymbol{N} \in \mathbb{V}_2^-\} = \mathrm{P}\{-c \leq n_1 \leq c, -\infty \leq n_2 \leq \infty\} = \mathrm{P}\{-c \leq n_1 \leq c\}.$$

Since n_1 has a univariate normal distribution it is clear that $c = z_{\left(1-\frac{\alpha}{2}\right)}$.

4.2 Simultaneous one-sided sets for $k(\geq 3)$ effective doses

For the general $k \geq 3$ case we want to find the value of c such that

$$\min_{x_{p_1},\, \ldots,\, x_{p_k}} \mathrm{P}\left\{x_{p_i} \in \mathbf{C}_{p_i}^- \text{ for } i = 1, \ldots, k\right\} = 1 - \alpha. \qquad (17)$$

It is clear that we may write

$$\mathrm{P}\left\{x_{p_i} \in \mathbf{C}_{p_i}^{-} \text{ for } i = 1, \ldots, k\right\} = \mathrm{P}\{\boldsymbol{N} \in \mathbb{V}_k^{-}\}$$

where $\mathbb{V}_k^{-} = \cap_{i=1}^{k} \mathbb{V}(x_{p_i})^{-}$. By rotational invariance, $\mathbb{V}_k^{-}$ may always take the form as shown in Figure 3, with the largest angle between any two directional vectors being θ_{1k}.

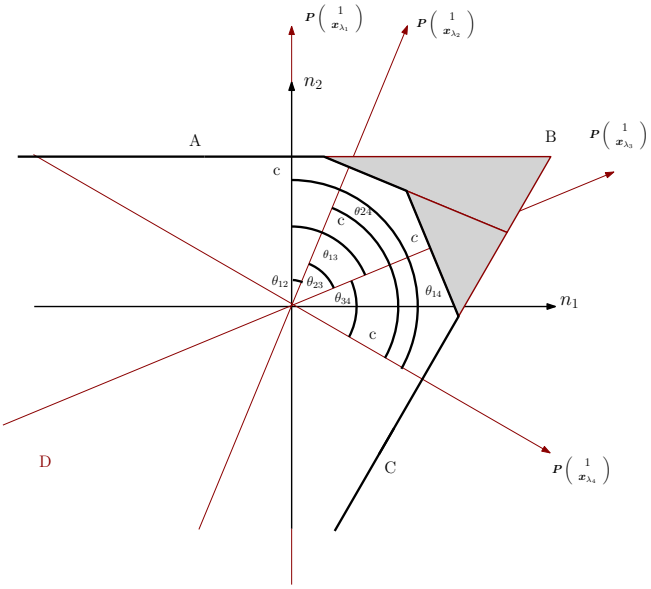

Figure 3 The region $\mathbb{V}_4^{-}$, expressed as the region bounded above by lines A and C, and the intercept B, less the grey shaded regions.

Following Liu (2010) (pages 41-44), we can show that $\mathrm{P}\{\boldsymbol{N} \in \mathbb{V}_k^{-}\}$ can be expressed as

$$\begin{aligned} & \frac{1}{2\pi}\left\{\int_{-\frac{\theta_{1k}}{2}}^{\frac{\pi}{2}}\left(1-\exp\left\{-\frac{c^2}{2cos(\theta)^2}\right\}\right)\mathrm{d}\theta+\int_{-\frac{\pi}{2}}^{\frac{\theta_{1k}}{2}}\left(1-\exp\left\{-\frac{c^2}{2cos(\theta)^2}\right\}\right)\mathrm{d}\theta\right\} \\ + \quad & \frac{\pi-\theta_{1k}}{2\pi}-\sum_{i=2}^{k-1}\int_{l_3}^{c}\phi(n_2)\left[\Phi\left(-n_2cot(\theta_{(i-1)k})+\frac{c}{sin(\theta_{(i-1)k})}\right)\right. \\ & \left.-\Phi\left(-n_2cot(\theta_{(i-1)i})+\frac{c}{sin(\theta_{(i-1)i})}\right)\right]\mathrm{d}n_2. \end{aligned} \tag{18}$$

As in the two-sided case, it is sufficient to minimise expression (18) with respect to $0 \leq \theta_{ij} \leq \pi$, which is given by the following theorem.

Theorem 4.1 *Let θ_{ik} be the largest angle between all the $\boldsymbol{Px_{p_i}}$s which is spanned by $\boldsymbol{Px_{p_1}}$ and $\boldsymbol{Px_{p_k}}$, as depicted in Figure 3. Furthermore, allow these two vectors, and therefore θ_{ik}, to be fixed. Then the expression of $P\{\boldsymbol{N} \in \mathbb{V}_k^{-}\}$ in (18) is minimised , when $\boldsymbol{Px_{p_2}}, \ldots, \boldsymbol{Px_{p_{k-1}}}$ are allowed to vary freely between $\boldsymbol{Px_{p_1}}$ and $\boldsymbol{Px_{p_k}}$, at the configuration that the angle between two adjacent vectors $\boldsymbol{Px_{p_j}}$ and $\boldsymbol{Px_{p_{j+1}}}$ is equal to $\frac{\theta_{1k}}{k}$ for all $j = 1, \ldots, k-1$.*

Proof. We note that if θ_{1k} is fixed and say only $\boldsymbol{Px_{p_2}}$ is allowed to change between the two adjacent $\boldsymbol{Px_{p_1}}$ and $\boldsymbol{Px_{p_3}}$, then $\mathrm{P}\{\boldsymbol{N} \in \mathbb{V}_k^{-}\}$ is equal to the probability of $\boldsymbol{N}$ in the region R_{ABC} bounded by the

half lines BA and BC subtracting the probability of $\boldsymbol{N}$ in the grey shaded regions, as depicted in Figure 3. When $\boldsymbol{Px_{p_2}}$ changes between $\boldsymbol{Px_{p_1}}$ and $\boldsymbol{Px_{p_3}}$, the probability of $\boldsymbol{N}$ in R_{ABC} does not change, only the probability of lying in the grey shaded region, which is the same as in the two-sided case depicted in Figure 1. Therefore Lemma 1 applies and $\mathrm{P}\{\boldsymbol{N} \in \mathbb{V}_k^-\}$ is minimised when $\boldsymbol{Px_{p_2}}$ is halfway between $\boldsymbol{Px_{p_1}}$ and $\boldsymbol{Px_{p_3}}$. The theorem then follows from repeated applications of Lemma 1 on $\boldsymbol{Px_{p_2}} \ldots \boldsymbol{Px_{p_{k-1}}}$. □

To find the minimum probability in (17), we still need to minimise $\mathrm{P}\{\boldsymbol{N} \in \mathbb{V}_k^-\}$ in (18) with respect to $\theta_{1k} \in (0, \pi]$, with $\theta_{j(j+1)} = \frac{\theta_{1k}}{k}$ for all $j = 1, \ldots, k-1$. While we are unable to establish any analytical result, this minimisation can easily be done numerically since only a one-variable search is involved. Hence the minimum probability in (17) for a given c can be easily computed using Theorem 2 and a one-variable numeric search. By using the bisection method or other search algorithm on c, combined with a numeric search for the minimum at each step, the c which sets the minimum probability in (17) to $1-\alpha$ can be computed quickly and accurately.

4.3 A note on upper one-sided confidence sets

Suppose we look for an equivalent upper one-sided confidence set, formed from a lower one sided confidence band of the form

$$\mathbf{C}_p^+ = \left\{ x : \frac{-\left(\boldsymbol{x}^\intercal \hat{\boldsymbol{\beta}} - \pi(p)\right)}{\sqrt{\boldsymbol{x}^\intercal \boldsymbol{J}^{-1} \boldsymbol{x}}} < c \right\}. \tag{19}$$

Once again we may set

$$\mathrm{P}\left\{x_{p_i} \in \mathbf{C}_{p_i}^+ \text{ for } i = 1, \ldots, k\right\} = \mathrm{P}\{\boldsymbol{N} \in \mathbb{V}_k^+\},$$

where $\mathbb{V}_k^+ = \cap_{i=1}^k \mathbb{V}(x_{p_i})^+$ and

$$\mathbb{V}(x_p)^+ = \left\{ \boldsymbol{N} : \frac{\left\{-\boldsymbol{P}\begin{pmatrix} 1 \\ x_p \end{pmatrix}\right\}^\intercal \boldsymbol{N}}{\left\|\boldsymbol{P}\begin{pmatrix} 1 \\ x_p \end{pmatrix}\right\|} < c \right\}.$$

It is immediate that

$$\mathrm{P}\{\boldsymbol{N} \in \mathbb{V}_k^+\} = \mathrm{P}\{\boldsymbol{N} \in \mathbb{V}_k^-\}$$

by rotational invariance. Therefore the required c is the same as the c for the lower one-sided confidence sets.

5 Illustration

5.1 Values of c and comparison

In this section we compare the simultaneous confidence sets obtained by the Scheffé band method (S) to the two-sided confidence sets (AS2) of Section 3 and the upper one-sided confidence sets (AS1) of Section 4. We have computed the critical constants c for AS2, AS1 and S for $k = 2, 3$ and 4 in Table 1 below. Source code (files: "c values AS2.R" and "c values AS1.R") to reproduce the results is available as Supporting Information on the journal's web page (http://onlinelibrary.wiley.com/doi/xxx/suppinfo).

Table 1 Values of c for AS2, AS1 and S simultaneous confidence sets

		AS2	AS1	S for $q=1$	S for $q=2$
k	$1-\alpha$	c		$\sqrt{\chi^{\alpha}_2}$	$\sqrt{\chi^{\alpha}_3}$
2	0.99	2.806225	2.575829	3.034854	3.368214
	0.95	2.236477	1.96	2.447747	2.795483
	0.90	1.948822	1.644854	2.145966	2.500278
3	0.99	2.913494	2.712313	3.034854	3.368214
	0.95	2.343701	2.123498	2.447747	2.795483
	0.90	2.052293	1.823565	2.145966	2.500278
4	0.99	2.962385	2.787521	3.034854	3.368214
	0.95	2.38728	2.19572	2.447747	2.795483
	0.90	2.092173	1.89069	2.145966	2.500278

The relative size of each confidence set may be directly compared by the size of c. In each case a smaller c indicates a smaller confidence set or bound. Thus the relative improvement over S type confidence sets at level $1-\alpha$ is

$$\left|\frac{\sqrt{\chi^{\alpha}_{q+1}}-c}{\sqrt{\chi^{\alpha}_{q+1}}}\right| * 100.$$

Table 1 shows a clear reduction in the size of c by the new confidence sets proposed in this paper. For $\alpha = 0.05$ and $q = 1$, AS2 shows a relative improvement of 8.6%, 4.25% and 2.5% for $k = 2, 3$ and 4 respectively. For AS1 the corresponding improvements are 19.9%, 13.2% and 10.3%. AS2 for $k = 2$ can be used for multivariate models with $q \geq 2$. For $q = 2$ there is a significant jump in improvement, to approximately 23%.

5.2 Application to Myocardial Infarction data

To illustrate the benefits of our new methodology, we consider a real data example. The Ontario Exercise Heart Collaborative Study (Walter 1983) recorded data from 341 patients and measured the recurrence of a Myocardial Infarction (MI) over a four year period. Logistic analysis is performed on the recurrence of MI with respect to smoking status (x_1) and serum triglyceride level (x_2). The model gave results

$$\hat{\boldsymbol{\beta}} = (-2.2791, 0.7682, 0.001952)^{\intercal}, \quad \boldsymbol{J}^{-1} = \begin{pmatrix} 0.06511 & & \\ -0.04828 & 0.09839 & \\ -0.0001915 & -0.00003572 & 0.000002586 \end{pmatrix}.$$

We construct simultaneous confidence sets for the CEDs of serum triglyceride level for non smokers $(x_1 = 0)$, at $\alpha = 0.05$, for $k = 2$ when interest is in any two of the three CEDs for $p = 0.4, 0.5$ and 0.6, using the S and AS2 methods for $k = 2$. Hence the values of the critical constant c are $c = 2.236$ and $c = 2.448$ for the AS2 and the S method, respectively. To construct $\boldsymbol{C}_p$ for x_p (or x_p^{CED}) with critical constant c, the lower bound is the value of x (or x^* for the CED) that solves

$$p - \left(\frac{\exp\left(\boldsymbol{x}^{\intercal}\hat{\boldsymbol{\beta}} + c\sqrt{\boldsymbol{x}^{\intercal}\boldsymbol{J}^{-1}\boldsymbol{x}}\right)}{1+\exp\left(\boldsymbol{x}^{\intercal}\hat{\boldsymbol{\beta}} + c\sqrt{\boldsymbol{x}^{\intercal}\boldsymbol{J}^{-1}\boldsymbol{x}}\right)}\right) = 0,$$

and the upper bound solves

$$p - \left(\frac{\exp\left(\boldsymbol{x}^{\intercal}\hat{\boldsymbol{\beta}} - c\sqrt{\boldsymbol{x}^{\intercal}\boldsymbol{J}^{-1}\boldsymbol{x}}\right)}{1+\exp\left(\boldsymbol{x}^{\intercal}\hat{\boldsymbol{\beta}} - c\sqrt{\boldsymbol{x}^{\intercal}\boldsymbol{J}^{-1}\boldsymbol{x}}\right)}\right) = 0.$$

In this case we set $x^* = (0, x_2)$ to establish the appropriate CED and the confidence sets are shown in Table 2 below, and illustrated in Figure 4. Source code (file: "Case Study Heart Exercise.R") to reproduce the results is available as Supporting Information on the journal's web page (http://onlinelibrary.wiley.com/doi/xxx/suppinfo). The code will recreate Figure 4, and generate the lower bound values for AS2 and S type bands in Table 2.

Table 2 C_p for the S type, and AS2 type sets for $k = 2$

	AS2 for $k = 2$		S	
p	Lower	Upper	Lower	Upper
0.4	364.9	∞	315.9	∞
0.5	442	∞	384	∞
0.6	517.8	∞	450.4	∞

It is clear from Table 2 that the AS2 method demonstrates a noticeable improvement over the original S method.

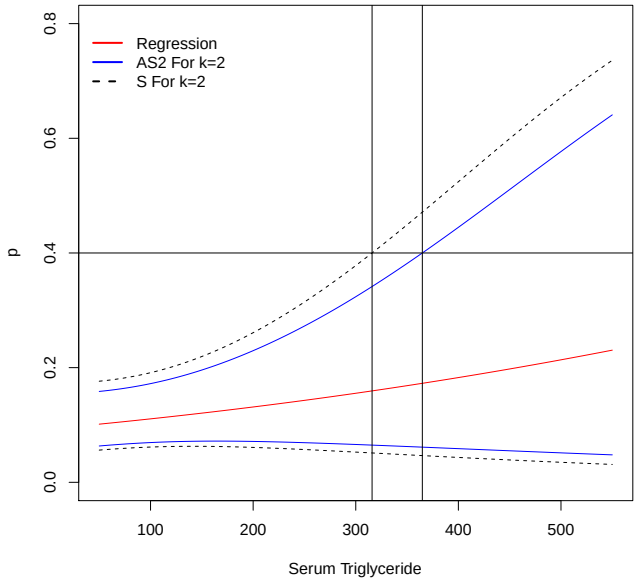

Figure 4 The confidence bands for the S and AS2 methods for $k = 2$ at $x_1 = 0$, for the Heart Exercise Data.

Figure 4 shows the Scheffé and the AS2 confidence bands for p as a function of the serum triglyceride level, and illustrates the construction of the confidence set for the effective level $x_{0.4}$, i.e. the level with expected response rate $p = 0.4$. From the intersection of the horizontal line at $p = 0.4$ with the respective upper confidence band, we move vertically to the x-axis. The level of serum triglyceride corresponding to this point is the lower bound of the confidence set for $x_{0.4}$.

We see that the confidence band that generates C_p, for AS2 sets, clearly has a significantly smaller average width over the whole range than the simultaneous confidence band of the S method. This results in smaller confidence sets, as demonstrated by Table 2, in which the sets constructed from the AS2 method have an improved lower bound of around 50-70 units. There is no upper bound as a result of the shape of

the logistic curve. In fact, it is well known (cf. Fieller (1954)) that $\boldsymbol{C}_p$ is not guaranteed to be a closed or even a single interval.

6 Discussion

We have proposed a new method to construct improved simultaneous confidence sets on effective doses, and have demonstrated that our approach shows notable reductions in the size of simultaneous confidence sets. AS2 sets for $k \geq 3$ may exhibit only a small improvement for $q = 1$. However, for $k = 2$, AS2 confidence sets may be obtained for multivariate models and show significant improvement, particularly for large values of q. As such, AS2 sets offer excellent utility in minimum and maximum effective dose problems. AS1 sets show a further and more significant reduction in the size of the appropriate bound than the equivalent AS2 set as expected including for $k = 2$. Establishing AS2 and AS1 sets at $k > 2$ for a multivariate model represents a significant challenge, however since many challenging practical models involve multivariate regression, there exists motivation for further study on this subject.

Though the focus of this paper is on logistic regression, the methods described here require only that there exists a regression line of the form $\boldsymbol{x}^\intercal\hat{\boldsymbol{\beta}} = f(p(x))$ with $f()$ being some link function and an approximately normal $\hat{\boldsymbol{\beta}}$ with some covariance matrix $\boldsymbol{J}^{-1}$. The simultaneous confidence sets considered in this paper immediately extend to the probit model, log logistic or even the Weibull models such as those described in Buckley & Piegorsch (2008) and Deutsch & Piegorsch (2012). Indeed these methods also extend to the normal linear regression model in which the unknown variance σ^2 needs to be estimated too. We look to apply the AS methods to a practical study of different models as further work.

A further interesting topic for future research will be to investigate simultaneous confidence sets for several EDs or CEDS when the covariates are constrained to some bounded region, e.g. the dose of a drug may have an upper limit for safety reasons.

Acknowledgements: The authors would like to thank the Editor, the Associate Editor and the two reviewers, whose insightful suggestions have led to a considerable improvement of an earlier version of this paper.

The authors have declared no conflict of interest.

7 Appendix

Proof of Lemma 1

To prove Lemma 1, we will first focus on the case $k = 3$. Denote the angle between $\boldsymbol{Px_{p_1}}$ and $\boldsymbol{Px_{p_2}}$ as θ_1 and the angle between $\boldsymbol{Px_{p_1}}$ and $\boldsymbol{Px_{p_3}}$ as θ_3 as shown in Figure 5. It is then immediate from (13) that $\mathrm{P}\{\boldsymbol{N} \in \mathbb{V}_3\}$ is equal to

$$\frac{1}{\pi}\left\{\int_{(\theta_3-\pi)/2}^{\theta_3/2}\left(1-\exp\left\{-\frac{c^2}{2cos(\theta)^2}\right\}\right)\mathrm{d}\theta+\int_{-\theta_3/2}^{(\pi-\theta_3)/2}\left(1-\exp\left\{-\frac{c^2}{2cos(\theta)^2}\right\}\right)\mathrm{d}\theta\right\}$$
$$-2\int_l^c \phi(n_2)\left[\Phi\left(-n_2cot(\theta_3)+\frac{c}{sin(\theta_3)}\right)-\Phi\left(-n_2cot(\theta_1)+\frac{c}{sin(\theta_1)}\right)\right]\mathrm{d}n_2 \quad (20)$$

where $l = \frac{c(sin(\theta_3)-\sin(\theta_1))}{sin(\theta_3-\theta_1)}$. As a next step, we will show that for fixed value of θ_3, the minimum in (20) is attained at $\theta_1 = \theta_3/2$; see Lemma 3 below.

Lemma 3 For given $0 < \theta_3^* \leq \pi$ the probability in Equation (20) is minimised with respect to $\theta_1 \in (0, \theta_3^*)$, when $\theta_1 = \frac{\theta_3^*}{2}$.

Proof of Lemma 3

Note that the probability of $\boldsymbol{N}$ lying in the parallelogram ABCD depends on θ_3 but not θ_1. As a result, with a fixed value of θ_3, Equation (20) varies only with the probability of lying in the grey shaded region (see Figure 5).

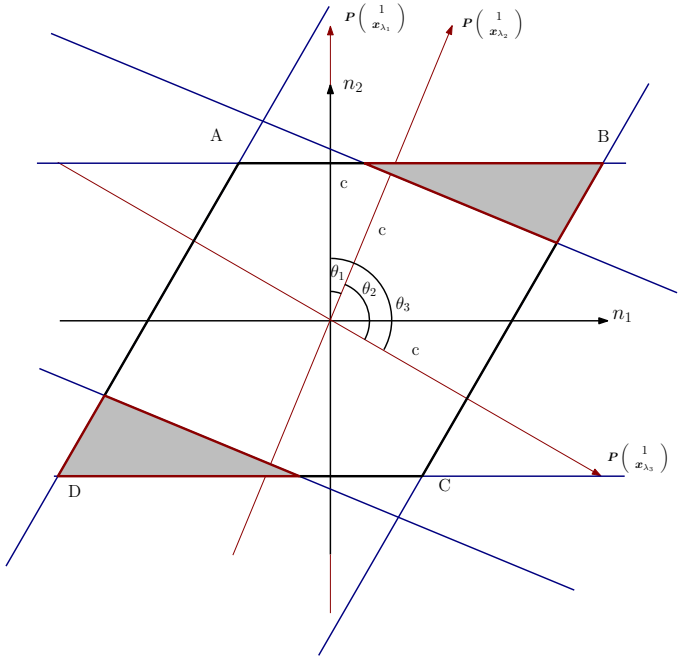

Figure 5 An expression of the region $\mathbb{V}_3$

It is therefore clear that to prove Lemma 3 is is sufficient to find $\theta_1 \in (0, \theta_3^*)$ which maximises the expression

$$\int_{\frac{c(sin(\theta_3^*)-\sin(\theta_1))}{sin(\theta_3^*-\theta_1)}}^{c} \phi(n_2) \left[\Phi\left(-n_2 cot(\theta_3^*) + \frac{c}{sin(\theta_3^*)}\right) - \Phi\left(-n_2 cot(\theta_1) + \frac{c}{sin(\theta_1)}\right)\right] \mathrm{d}n_2. \quad (21)$$

This is done in two steps, Result 1 and Result 2, which are proven at the end of the Appendix:

Result 1 The differential of the expression (21) with respect to θ_1 is given by

$$\frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{c^2}{2}\right\} \left[\phi\left(\frac{c - ccos(\theta_1)}{sin(\theta_1)}\right) - \phi\left(\frac{\frac{c(sin(\theta_3^*)-\sin(\theta_1))}{sin(\theta_3^*-\theta_1)} - ccos(\theta_1)}{sin(\theta_1)}\right)\right]. \quad (22)$$

Result 2 As a function of $\theta_1 \in (0, \theta_3^*)$ with $\theta_3^* < \pi$, the expression (22), denoted $D(\theta_1)$, has only one zero point at $\theta_1 = \frac{\theta_3^*}{2}$. Furthermore, $D(\theta_1) > 0$ for $\theta_1 \in (0, \frac{\theta_3^*}{2})$ and $D(\theta_1) < 0$ for $\theta_1 \in (\frac{\theta_3^*}{2}, \theta_3)$.

Lemma 3 follows now immediately from Result 2, which tells us that the point $\theta_1 = \frac{\theta_3^*}{2}$ is a maximum point of (21), and subsequently, (20) is also minimized at the stationary point for a fixed θ_3. □

We may now prove Lemma 1. Form $\mathbb{V}_k$ as the intersection of $\mathbb{V}(x_{p_{j-1}})$ and $\mathbb{V}(x_{p_{j+1}})$ less the remaining $k-2$ regions of the form $\mathrm{P}\left\{\boldsymbol{N} \in \mathbb{V}(x_{p_{j-1}}) \cap \mathbb{V}(x_{p_{j+1}}) \cap \mathbb{V}^c(x_{p_i})\right\}; i \neq j-1, j+1$. If only $\boldsymbol{P}\boldsymbol{x}_{\boldsymbol{p_j}}$ can vary then the probability of $\boldsymbol{N}$ in $\mathbb{V}_k$ depends only on the grey shaded region $\mathrm{P}\left\{\boldsymbol{N} \in \mathbb{V}(x_{p_{j-1}}) \cap \mathbb{V}(x_{p_{j+1}}) \cap \mathbb{V}^c(x_{p_j})\right\}$. By rotational invariance this is equivalent to the region in (21) with $\theta_{j-1,j+1} = \theta_3$, and the proof is then immediate from the proof of Lemma 3. □

Proof of Result 1

Let

$$\phi(n_2)\left[\Phi\left(-n_2cot(\theta_3^*)+\frac{c}{sin(\theta_3^*)}\right)-\Phi\left(-n_2cot(\theta_1)+\frac{c}{sin(\theta_1)}\right)\right]=g(n_2,\theta_1)$$

and $l_3(\theta_1)=\frac{c(sin(\theta_3)-\sin(\theta_1))}{sin(\theta_3-\theta_1)}$. Applying the Leibniz integral rule to the differential of (21) yields

$$\frac{\mathrm{d}}{\mathrm{d}\theta_1}\int_{l_3(\theta_1)}^{c} g(n_2,\theta_1)\mathrm{d}n_2=\int_{l_3(\theta_1)}^{c}\frac{\mathrm{d}g(n_2,\theta_1)}{\mathrm{d}\theta_1}\mathrm{d}n_2-g(l_3,\theta_1)\frac{\mathrm{d}l_3(\theta_1)}{\mathrm{d}\theta_1}.$$

Further noting that

$$g(l_3,\theta_1)=\phi(l_3)\left[\Phi\left(-l_3cot(\theta_3^*)+\frac{c}{sin(\theta_3^*)}\right)-\Phi\left(-l_3cot(\theta_1)+\frac{c}{sin(\theta_1)}\right)\right]=0,$$

the differential of (21) becomes

$$\begin{aligned}
&\int_{l_3}^{c}\frac{\mathrm{d}}{\mathrm{d}\theta_1}\phi(n_2)\left[\Phi\left(-n_2cot(\theta_3^*)+\frac{c}{sin(\theta_3^*)}\right)-\Phi\left(-n_2cot(\theta_1)+\frac{c}{sin(\theta_1)}\right)\right]\mathrm{d}n_2\\
=&\int_{l_3}^{c}\frac{\mathrm{d}}{\mathrm{d}\theta_1}\phi(n_2)\Phi\left(-n_2cot(\theta_3^*)+\frac{c}{sin(\theta_3^*)}\right)\mathrm{d}n_2\\
&-\int_{l_3}^{c}\frac{\mathrm{d}}{\mathrm{d}\theta_1}\phi(n_2)\Phi\left(-n_2cot(\theta_1)+\frac{c}{sin(\theta_1)}\right)\mathrm{d}n_2\\
=&-\int_{l_3}^{c}\frac{\mathrm{d}}{\mathrm{d}\theta_1}\phi(n_2)\Phi\left(-n_2cot(\theta_1)+\frac{c}{sin(\theta_1)}\right)\mathrm{d}n_2\\
=&\int_{l_3}^{c}\phi(n_2)\phi\left(-n_2cot(\theta_1)+\frac{c}{sin(\theta_1)}\right)\left[\frac{c(cos(\theta_1))-n_2}{sin(\theta_1)^2}\right]\mathrm{d}n_2. \qquad (23)
\end{aligned}$$

Through manipulation involving the normal pdf, we have

$$\begin{aligned}
\phi(n_2)\phi\left(-n_2cot(\theta_1)+\frac{c}{sin(\theta_1)}\right)=&\frac{1}{2\pi}\exp\left\{-\frac{n_2^2}{2}-\frac{(-n_2cot(\theta_1)+\frac{c}{sin(\theta_1)})^2}{2}\right\}\\
=&\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{c^2}{2}\right\}\phi(z)
\end{aligned}$$

where $z=\frac{n_2-ccos(\theta_1)}{sin(\theta_1)}$ and $\phi(z)$ is the standard normal pdf for z. Hence (23) becomes

$$\begin{aligned}
&\int_{l_3}^{c}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{c^2}{2}\right\}\left[\frac{c(cos(\theta_1))-n_2}{sin(\theta_1)^2}\right]\phi(z)\mathrm{d}n_2\\
=&\int_{l_3}^{c}\frac{ccos(\theta_1)}{sin(\theta_1)^2\sqrt{2\pi}}\exp\left\{-\frac{c^2}{2}\right\}\phi(z)\mathrm{d}n_2-\int_{l_3}^{c}\frac{n_2}{sin(\theta_1)^2\sqrt{2\pi}}\exp\left\{-\frac{c^2}{2}\right\}\phi(z)\mathrm{d}n_2. \qquad (24)
\end{aligned}$$

(24) consists of two terms, which we evaluate individually. For the first term we change the variable of integration to z, noting that $\mathrm{d}z/\mathrm{d}n_2=1/sin(\theta_1)$, which gives

$$\begin{aligned}
&\int_{l_3}^{c}\frac{ccos(\theta_1)}{sin(\theta_1)^2\sqrt{2\pi}}\exp\left\{-\frac{c^2}{2}\right\}\phi(z)\mathrm{d}n_2=\int_{\frac{l_3-ccos(\theta_1)}{sin(\theta_1)}}^{\frac{c-ccos(\theta_1)}{sin(\theta_1)}}\frac{ccos(\theta_1)}{sin(\theta_1)\sqrt{2\pi}}\exp\left\{-\frac{c^2}{2}\right\}\phi(z)\mathrm{d}z\\
=&\left[\frac{ccos(\theta_1)}{sin(\theta_1)\sqrt{2\pi}}\exp\left\{-\frac{c^2}{2}\right\}\Phi(z)\right]_{t_2}^{t_1}=\frac{ccos(\theta_1)}{sin(\theta_1)\sqrt{2\pi}}\exp\left\{-\frac{c^2}{2}\right\}[\Phi(t_1)-\Phi(t_2)] \qquad (25)
\end{aligned}$$

where $t_1 = \frac{c - ccos(\theta_1)}{sin(\theta_1)}$ and $t_2 = \frac{l_3 - ccos(\theta_1)}{sin(\theta_1)}$. For the second term, we apply integration by parts to give

$$\begin{aligned}
&\int_{l_3}^{c} \frac{n_2}{sin(\theta_1)^2\sqrt{2\pi}} \exp\left\{-\frac{c^2}{2}\right\} \phi(z) \mathrm{d}n_2 \\
=&\left[\frac{n_2}{sin(\theta_1)\sqrt{2\pi}} \exp\left\{-\frac{c^2}{2}\right\} \Phi(z)\right]_{l_3}^{c} - \int_{l_3}^{c} \frac{1}{sin(\theta_1)\sqrt{2\pi}} \exp\left\{-\frac{c^2}{2}\right\} \Phi(z) \mathrm{d}n_2 \\
=&\frac{1}{sin(\theta_1)\sqrt{2\pi}} \exp\left\{-\frac{c^2}{2}\right\} [n_2\Phi(z)]_{l_3}^{c} - \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{c^2}{2}\right\} \int_{t_2}^{t_1} \Phi(z) \mathrm{d}z \\
=&\frac{1}{sin(\theta_1)\sqrt{2\pi}} \exp\left\{-\frac{c^2}{2}\right\} [n_2\Phi(z)]_{l_3}^{c} - \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{c^2}{2}\right\} [z\Phi(z) + \phi(z)]_{t_2}^{t_1}
\end{aligned} \tag{26}$$

where the result $\int_{t_2}^{t_1} \Phi(z) \mathrm{d}z = [z\Phi(z) + \phi(z)]_{t_2}^{t_1}$ is an immediate consequence of integration by parts used in Ng & Murray (1969) (Section 4.1, Result 1). By using the two expressions (25) and (26) above, (24) can be expressed as

$$\begin{aligned}
&\frac{ccos(\theta_1)}{sin(\theta_1)\sqrt{2\pi}} \exp\left\{-\frac{c^2}{2}\right\} [\Phi(z)]_{t_2}^{t_1} - \frac{1}{sin(\theta_1)\sqrt{2\pi}} \exp\left\{-\frac{c^2}{2}\right\} [n_2\Phi(z)]_{l_3}^{c} \\
&+\frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{c^2}{2}\right\} [z\Phi(z) + \phi(z)]_{t_2}^{t_1} \\
=&\left[\left(\frac{ccos(\theta_1)}{sin(\theta_1)\sqrt{2\pi}} \exp\left\{-\frac{c^2}{2}\right\} \Phi(t_1)\right) - \left(\frac{ccos(\theta_1)}{sin(\theta_1)\sqrt{2\pi}} \exp\left\{-\frac{c^2}{2}\right\} \Phi(t_2)\right)\right] \\
&-\left[\left(\frac{c}{sin(\theta_1)\sqrt{2\pi}} \exp\left\{-\frac{c^2}{2}\right\} \Phi(t_1)\right) - \left(\frac{l_3}{sin(\theta_1)\sqrt{2\pi}} \exp\left\{-\frac{c^2}{2}\right\} \Phi(t_2)\right)\right] \\
&+\left[\left(\frac{t_1}{\sqrt{2\pi}} \exp\left\{-\frac{c^2}{2}\right\} \Phi(t_1)\right) - \left(\frac{t_2}{\sqrt{2\pi}} \exp\left\{-\frac{c^2}{2}\right\} \Phi(t_2)\right)\right] \\
&+\left[\left(\frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{c^2}{2}\right\} \phi(t_1)\right) - \left(\frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{c^2}{2}\right\} \phi(t_2)\right)\right] \\
=&\frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{c^2}{2}\right\} [\phi(t_1) - \phi(t_2)] \\
=&\frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{c^2}{2}\right\} \left[\phi\left(\frac{c - ccos(\theta_1)}{sin(\theta_1)}\right) - \phi\left(\frac{\frac{c(sin(\theta_3^*) - \sin(\theta_1))}{sin(\theta_3^* - \theta_1)} - ccos(\theta_1)}{sin(\theta_1)}\right)\right]
\end{aligned}$$

as required. □

Proof of Result 2

The differential is the difference of two standard normal pdf values, multiplied by some constant term. Consequently, the behaviour of the differential with respect to θ_1 hinges on the relationship between the absolute values of the two arguments of the pdfs. Specifically, we are comparing

$$\left|\frac{c - ccos(\theta_1)}{sin(\theta_1)}\right| \text{ to } \left|\frac{\frac{c(sin(\theta_3^*) - \sin(\theta_1))}{sin(\theta_3^* - \theta_1)} - ccos(\theta_1)}{sin(\theta_1)}\right|.$$

Since $\phi(x) - \phi(y) \geq 0$ if and only if $|x| - |y| \leq 0$, we have $D(\theta_1) \leq 0 \Leftrightarrow d(\theta_1) \geq 0$ where

$$\begin{aligned} d(\theta_1) =& \left| \frac{c - ccos(\theta_1)}{sin(\theta_1)} \right| - \left| \frac{\frac{c(sin(\theta_3^*) - \sin(\theta_1))}{sin(\theta_3^* - \theta_1)} - ccos(\theta_1)}{sin(\theta_1)} \right| \\ =& \frac{c}{|sin(\theta_1)|} \left(|1 - cos(\theta_1)| - \left| \frac{sin(\theta_3^*) - \sin(\theta_1)}{sin(\theta_3^* - \theta_1)} - cos(\theta_1) \right| \right). \end{aligned}$$

Since $c > 0$ and $\theta_1 \in (0, \pi)$, it is sufficient to focus on the sign of

$$\begin{aligned} g(\theta_1) =& |1 - cos(\theta_1)| - \left| \frac{sin(\theta_3^*) - \sin(\theta_1)}{sin(\theta_3^* - \theta_1)} - cos(\theta_1) \right| \\ =& (1 - cos(\theta_1)) - \left| \frac{sin(\theta_3^*) - \sin(\theta_1)}{sin(\theta_3^* - \theta_1)} - cos(\theta_1) \right| \end{aligned}$$

as $(1 - cos(\theta_1))$ is always positive. Also, $\frac{sin(\theta_3^*) - \sin(\theta_1)}{sin(\theta_3^* - \theta_1)} - cos(\theta_1)$ is always negative, and thus it is sufficient to evaluate

$$\begin{aligned} g(\theta_1) =& 1 - cos(\theta_1) - \left(cos(\theta_1) + \frac{sin(\theta_1) - \sin(\theta_3^*)}{sin(\theta_3^* - \theta_1)} \right) \\ =& \frac{sin(\theta_3^* - \theta_1) - 2cos(\theta_1)sin(\theta_3^* - \theta_1) - sin(\theta_1) + \sin(\theta_3^*)}{sin(\theta_3^* - \theta_1)} \\ =& \frac{sin(\theta_3^* - \theta_1) + sin(2\theta_1 - \theta_3^*) - sin(\theta_1)}{sin(\theta_3^* - \theta_1)} \end{aligned}$$

using the product to sum formulae. It is clear that, at $\theta_1 = \frac{\theta_3 *}{2}$, $g(\theta_1) = 0$ and therefore $D(\theta_1) = 0$. Thus $\theta_1 = \frac{\theta_3 *}{2}$ is a stationary point of $D(\theta_1)$. For the sign of $g(\theta_1)$ over $0 < \theta_1 < \theta_3^* < \pi$, it suffices to focus on the numerator

$$h(\theta_1) = sin(\theta_3^* - \theta_1) + sin(2\theta_1 - \theta_3^*) - sin(\theta_1).$$

Let $\theta_1 = \frac{\theta_3^*}{2} - \epsilon$ for $\epsilon \in (0, \frac{\theta_3^*}{2})$, then

$$\begin{aligned} h(\frac{\theta_3^*}{2} - \epsilon) =& sin(\theta_3^* - \theta_1) + sin(2\theta_1 - \theta_3^*) - sin(\theta_1)|_{\theta_1 = \frac{\theta_3^*}{2} - \epsilon} \\ =& sin(\theta_3^* - (\frac{\theta_3^*}{2} - \epsilon)) + sin(2(\frac{\theta_3^*}{2} - \epsilon) - \theta_3^*) - sin(\frac{\theta_3^*}{2} - \epsilon) \\ =& sin(\frac{\theta_3^*}{2} + \epsilon) - sin(2\epsilon) - sin(\frac{\theta_3^*}{2} - \epsilon) \\ =& sin((\frac{\theta_3^*}{2} - \epsilon) + 2\epsilon) - sin(2\epsilon) - sin(\frac{\theta_3^*}{2} - \epsilon) \\ =& sin(\frac{\theta_3^*}{2} - \epsilon)cos(2\epsilon) + sin(2\epsilon)(cos(\frac{\theta_3^*}{2} - \epsilon) - 1) - sin(\frac{\theta_3^*}{2} - \epsilon) \\ <& sin(\frac{\theta_3^*}{2} - \epsilon)cos(2\epsilon) - sin(\frac{\theta_3^*}{2} - \epsilon) < 0 \end{aligned}$$

as $(cos(\frac{\theta_3^*}{2} - \epsilon) - 1)$ must always be negative. Furthermore, from

$$sin(\frac{\theta_3^*}{2} + \epsilon) - sin(2\epsilon) - sin(\frac{\theta_3^*}{2} - \epsilon) < 0$$

established above, it follows that

$$h(\frac{\theta_3^*}{2}+\epsilon) = -(sin(\frac{\theta_3^*}{2}+\epsilon) - sin(2\epsilon) - sin(\frac{\theta_3^*}{2}-\epsilon)) > 0.$$

This tells us that the function $d(\theta_1)$ is negative as the value of θ_1 approaches the stationary point, and hence $D(\theta_1)$ is positive, and that the reverse is true as we move away past the stationary point, as required. □

References

Al-Saidy, O. M., Piegorsch, W. W., West, R. W. & Nitcheva, D. K. (2003), 'Confidence bands for low-dose risk estimation with quantal response data', *Biometrics* **59**, 1056–1062.

Brand, R., Pinnock, D. & Jackson, K. (1973), 'Large sample confidence bands for the logistic response curve and its inverse', *The American Statistician* **27**, 157–160.

Bretz, B., Hsu, J., Pinheiro, J. & Liu, Y. (2008), 'Dose finding – a challenge in statistics', *Biometrical Journal* **50**, 480–504.

Buckley, B. & Piegorsch, W. (2008), 'Simultaneous confidence bands for abbott adjusted quantal response models', *Statistical Methodology* **5**, 209–219.

Carter, W., Chinchilli, V., Wilson, J., Campbell, E., Kessler, F. & Carchman, R. (1986), 'An asymptotic confidence region for the ed_{100p} from the logistic response surface for a combination of agents', *The American Statistician* **40**, 124–128.

Cox, C. (1990), 'Fieller's theorem, the likelihood and the delta method', *Biometrics* **46**, 709–718.

Deutsch, R. & Piegorsch, W. (2012), 'Benchmark dose profiles for joint action quantal data in quantitative risk assessment', *Biometrics* **68**, 1313–1322.

Faraway, J. (2016), *Extending the Linear Model with R, 2nd edition*, CRC Press.

Fieller, E. (1954), 'Some problems in interval estimation', *Journal of the Royal Statistical Society, Series B* **16**, 175–185.

Finney, D. (1971), *Probit Analysis, Second Edition*, Cambridge University Press.

Hsu, J. (1996), *Multiple Comparisons, Theory and Methods*, Chapman and Hall.

Li, J., Nordheim, E., Zhang, C. & Lehner, C. (2008), 'Estimation and confidence regions for multi-dimensional effective dose', *Biometrical Journal* **50**, 110–122.

Li, J. & Wong, W. (2011), 'Two-dimensional toxic dose and multivariate logistic regression, with application to decompression sickness', *Biostatistics* **12**, 143–155.

Li, J., Zhang, C., Doksum, K. & Nordheim, E. (2010), 'Simultaneous confidence intervals for semiparametric logistic regression and confidence regions for the multi-dimensional effective dose', *Statistica Sinica* **20**, 637–659.

Li, J., Zhang, C., Nordheim, E. & Lehner, C. (2008), 'On the multivariate predictive distribution of multi-dimensional effective dose: a bayesian approach', *Journal of Statistical Computation and Simulation* **78**, 429–442.

Liu, W. (2010), *Simultaneous inference in regression*, CRC Press.

Ng, E. & Murray, G. (1969), 'A table of integrals of the error functions', *Journal of Research of the National Bureau of Standards* **73B**.

Nitcheva, D. K., Piegorsch, W. W., West, R. W. & Kodell, R. L. (2005), 'Multiplicity-adjusted inference in risk assessment: Benchmark analysis with quantal response data', *Biometrics* **61**, 277–286.

Peng, J., Robichaud, M. & Alsubie, A. (2015), 'Simultaneous confidence bands for lower-dose risk estimation with quantal data', *Biometrical Journal* **57**, 27–38.

Piegorsch, W., West, R., Pan, W. & Kodell, R. (2005), 'Low dose risk estimation via simultaneous statistical inferences', *Journal of the Royal Statistical Society, Series C* **54**, 245–258.

Scheffé, H. (1953), 'A method for judging all contrasts in analysis of variance', *Biometrika* **40**, 87–104.

Walter, W. (1983), 'A note on confidence bands for the logistic response curve', *The American Statistician* **37**, 158–160.

