Sequential LND Sensitivity Test for Binary Response Data

Lei Wang*, Yukun Liu*, Wei Wu* and Xiaolong Pu**

*aSchool of Finance and Statistics, East China Normal University, 500 Dongchuan Road, Shanghai, 200241, China

(email: leiwang.stat@gmail.com, ykliu@sfs.ecnu.edu.cn, wuwei0881@gmail.com, xlpu@stat.ecnu.edu.cn)

(Received 00 Month 200x; in final form 00 Month 200x)

Sensitivity tests are used to make inferences about a sensitivity, a characteristic property of some products that can not be observed directly. For binary response sensitivity data (dead or alive, explode or unexplode), the Langlie and Neyer are two well-known sensitivity tests. The priorities of the Langlie and Neyer tests are investigated in this paper. It is showed that the Langlie test has an advantage in getting an overlap, while the Neyer test has better estimation precision. Aiming at improving both the speed of getting an overlap and the estimation precision, we propose a new sensitivity test which replaces the first part of the Neyer test with the Langlie test. Our simulation studies indicate that the proposed test outperforms the Langlie, Neyer and Dror and Steinberg tests, from the viewpoints of estimation precision and probability of obtaining an overlap.

Keywords: sensitivity test; D-optimality criterion; MLE; Fisher information matrix; Langlie test; Neyer test.

1. Introduction

Binary response data or dichotomous data are important (see [3][8]) and commonly used in biology and initiating explosive device study (e.g., dead or alive, explode or unexplode). Sensitivity refers to the critical value of a latent continuous variable in the binary response data. The problem of interest is to make inferences about the sensitivity distribution. To this end, sensitivity tests are conducted to gather information about the sensitivity.

Sensitivity tests can be applied in various research fields such as biological, pharmaceutical, psychological and engineering research and so on. However, the scarcity of sensitivity information poses a challenge in constructing efficient sensitivity tests in various settings. The Probit test and the Bruceton test studied by Bliss [2] and Dixon and Mood [5], respectively, were two primary attempts in this direction. They are simple but inefficient. Robbins and Monro [15] proposed the Robbins-Monro test to make inferences about sensitivity quantiles. A more efficient sensitivity test is the Langlie test proposed by Langlie [11]. There are also some recent works about the sensitivity tests, see Wu [17] for efficient sequential designs with binary data, Neyer [14] for the D-optimality based sensitivity test (the Neyer test), Dror and Steinberg [6] for generalized linear models (the D-S test for short). Other discussions can be found in Davis [4], McLeish and Tosh [12], Joseph [9] and Karvanen et al. [10].
Sequential LND sensitivity test

Among others, the Langlie and Neyer tests are two well-known sensitivity testing methods. The Langlie test has many nice properties. For example, it does not put any distribution assumption on data; it is simple and easy to implement, etc. Moreover, the Langlie test was developed to find design points corresponding to 50% responses, so it could provide a good estimation of the population mean but could not provide an ideal estimation of the population variance of a symmetric distribution. Furthermore, it owns blindness in finding the follow-up design points. Unlike the Langlie test, the Neyer test assumes that the sensitivity follows a parametric model and has a clear goal in finding the follow-up design points. Roughly speaking, it consists of two parts. The first part is designed to make an overlap come up. By an overlap, we mean that there is an overlap between stimuli that produce responses and those that produce non-responses, or more specifically the smallest response is smaller than the largest non-response. An overlap guarantees the existence of a maximum likelihood estimate (MLE) of the unknown parameters (Silvapulle [16]), which will be used in the second part. In the second part, by maximizing the determinant of the Fisher information matrix of the parameters, the Neyer test renders the resulting MLEs to have approximately the smallest variances (Abdelbasit and Plackett [1]).

Intuitively, the Neyer test is more efficient than the Langlie test. And the efficiency of the Neyer test lies in the second part, which can extract as much information as possible from a population. However, the cost is that the Neyer test needs to specify one more quantity $\sigma_{\text{guess}}$ — a guess value of the standard deviation — at the beginning. It can be expected that the choice of such a guess value will affect the performance of the Neyer test, which will be seen from our simulation results: a bad guess value $\sigma_{\text{guess}}$ often makes an overlap come up late, resulting in a waste of samples. Unfortunately, it is not easy to give an accurate guess value of the standard deviation in practice, which plays down the efficiency of the Neyer test. On the contrary, the Langlie test is shown to be a stable method which can guarantee the speed of obtaining an overlap.

As pointed out earlier, an efficient sensitivity test should require less initial knowledge of the parameters and use all current data to determine the next design points. According to the above discussion and motivated by the Langlie and Neyer tests, we propose a new efficient sensitivity test in this paper, which replaces the first part of the Neyer test with the Langlie test, denoted as the LND test. The proposed test requires only a lower bound $x_L$ and an upper bound $x_U$ of the mean beforehand. Moreover, it inherits the advantages of both the Langlie test and the Neyer test in getting an overlap and having nice estimation precision, respectively.

The rest of this paper is organized as follows. In Section 2, we review the Langlie and Neyer tests briefly, then compare their properties by simulations. Section 3 presents the proposed test. Simulations are conducted in Section 4 to compare the proposed test and three competing tests, i.e., the Langlie, Neyer and D-S tests. A brief summary about the proposed test is given in Section 5.

2. The Langlie and Neyer tests

When one makes inferences about sensitivities, parametric models are usually assumed. In this article, according to Silvapulle [16], Neyer [13, 14] and Dror and Steinberg [6], we assume the cumulative distribution function (CDF) of sensitivity takes a location-scale form $f((x - \mu)/\sigma)$, where $f(\cdot)$ is a known function, $\mu$ and $\sigma > 0$ are unknown location and scale parameters, respectively. Special cases of $f(\cdot)$ include normal, log-normal, logistic and log-logistic distributions.

Suppose experiments have been conducted sequentially at $x_1, x_2, \ldots, x_N$. Let
δ
i
be the experimental result at
x
i
, where
δ
i
= 1 denotes a response and
δ
i
= 0 a non-response. The likelihood function based on
x
i
and
δ
i
(i = 1, · · · , N) is denoted as

\[ L(\mu, \sigma) = \prod_{i=1}^{N} f(z_i)^{\delta_i} (1 - f(z_i))^{1-\delta_i}, \quad z_i = (x_i - \mu)/\sigma. \] (1)

The parameters
\( \mu \) and
\( \sigma \) are usually estimated by their MLEs (Silvapulle [16]), say
\( \hat{\mu} \) and
\( \hat{\sigma} \), respectively. Any function of
\( \mu \) and
\( \sigma \) say
\( h(\mu, \sigma) \) can also be estimated by its MLE
\( \hat{h}(\hat{\mu}, \hat{\sigma}) \). Take a quantile of the sensitivity distribution for example. Let
\( L_q \) be the qth percentile of the sensitivity distribution and
\( z_q \) be the solution to
\( f(z_q) = q \). Then
\( L_h \) can be estimated by
\( \hat{L}_q = \hat{\mu} + z_q \hat{\sigma} \).

In the remaining of this section, we will briefly describe the Langlie and Neyer tests and investigate their properties through simulations.

2.1 The Langlie test

The Langlie test is a mechanical procedure which does not depend on the sensitivity distribution. When applied in practice, it requires the experimenter to specify a lower bound
\( x_L \) and an upper bound
\( x_U \) of the mean beforehand, at which the experimental results are most probably non-response and response, i.e., a 5% and a 95% chance of responding, respectively. The Langlie test consists of the following steps.

Step 1. Conduct the first experiment at
\( x_1 = (x_L + x_U)/2 \) and set
\( n = 1 \). Let
\( \delta_1 \) denote the experimental result, where
\( \delta_1 = 1 \) denotes a response and
\( \delta_1 = 0 \) a non-response. Go to Step 2.

Step 2. \( x_{n+1} \) is determined according to the follow rule. Let
\( k_0 \) be the smallest \( k \) such that
\( k = 2s_k \) where
\( s_k = \sum_{i=1}^{k} \delta_{n+1-i} \) (\( k = 1, 2, \cdots, n \)). If such a
\( k_0 \) exists, then \( x_{n+1} = (x_n + x_{k_0})/2 \). Otherwise, \( x_{n+1} = (x_n + x_L)/2 \) if \( \delta_n = 1 \); \( x_{n+1} = (x_n + x_U)/2 \) if \( \delta_n = 0 \). Set
\( n = n+1 \), and repeat Step 2 until the number of experiments \( n \) is equal to the pre-specified sample size \( N \).

2.2 The Neyer test

The Neyer test roughly consists of two parts. The first part is an initial test designed to get an overlap. The second part is a D-optimization procedure designed to find the follow-up design points by maximizing the Fisher information of the parameters. Below we give a brief introduction to the Fisher information and the D-optimization procedure, as they are vital components of the algorithm of the Neyer test.

**Fisher information matrix** The Fisher information is a way of measuring the amount of information that the data carries about an unknown parameter. Let \((x_i, \delta_i)\) \( (i = 1, 2, \cdots, n) \) be the first \( n \) experimental results, then the Fisher information matrix of the parameter \( \theta = (\mu, \sigma) \) under the parametric model \( f((x-\mu)/\sigma) \) is

\[ I_n(\theta) = \begin{pmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{pmatrix} = \sum_{i=1}^{n} \frac{(f'(z_i))^2}{f(z_i)(1-f(z_i))\sigma^2} \begin{pmatrix} 1 & z_i \\ z_i & z_i^2 \end{pmatrix}, \] (2)

where \( f'(\cdot) \) denotes the derivative of \( f(\cdot) \) and \( z_i = (x_i - \mu)/\sigma, \) \( i = 1, \cdots, n \).
Sequential LND sensitivity test

\[ x_{n+1} = \arg \max_x |I_{n+1}(x|\hat{\theta}_n)|, \]  

where  
\[ I_{n+1}(x|\hat{\theta}_n) = I_n(\hat{\theta}_n) + \frac{(f(z))^2}{f(z)(1-f(z))\sigma^2} \left( \frac{1}{z^2} \right) \]  

\[ z = (x - \hat{\mu}_n)/\hat{\sigma}_n \] and \( |I_{n+1}(x|\hat{\theta}_n)| \) is the determinant of \( I_{n+1}(x|\hat{\theta}_n) \).

Specifically, the procedure of the Neyer test is as follows. Keep in mind that the symbol \( \sigma_{\text{guess}} \) denotes a guess value of the standard deviation of the sensitivity.

**Step 1.** Conduct the first experiment at \( x_1 = (x_L + x_U)/2 \). Let \( \delta_1 \) denote the experimental result, where \( \delta_1 = 1 \) denotes a response and \( \delta_1 = 0 \) a non-response. If \( \delta_1 = 1 \), \( x_2 = (x_L + x_1)/2 \); otherwise \( x_2 = (x_U + x_1)/2 \). Conduct an experiment at \( x_2 \) and record the experimental result \( \delta_2 \). Set \( n = 2 \) and \( r = 1 \).

**Step 2.** The next design point \( x_{n+1} = 2x_n - x_1 \) if \( \delta_1 = \delta_2 = \cdots = \delta_n \); otherwise a binary search is performed between the minimum response and the maximum non-response until the difference between them is less than \( \sigma_{\text{guess}} \). If no overlap comes up, go to Step 3; otherwise go to Step 4.

**Step 3.** Let \( \hat{\mu}_n \) be the mean of the minimum response point and the maximum non-response point based on the first \( n \) data and \( \hat{\sigma}_n = 0.8\sigma_{\text{guess}} \). The next design point is \( x_{n+1} = \arg \max_x |I_{n+1}(x|\hat{\mu}_n, \hat{\sigma}_n)| \). Conduct an experiment at \( x_{n+1} \) and record the experimental result. If no overlap comes up, set \( n = n + 1 \), \( r = r + 1 \) and repeat Step 3; otherwise go to Step 4.

**Step 4.** Denote \( \hat{\theta}_n \) as the MLE of \( \theta \) based on the first \( n \) experimental data. The next design point \( x_{n+1} = \arg \max_x |I_{n+1}(x|\hat{\theta}_n)| \). Set \( n = n + 1 \), and repeat Step 4 until the number of experiments \( n \) is equal to a pre-specified sample size \( N \).

If the parameter of interest is \( \theta = \mu \) or \( \theta = \sigma \), we only need to modify the above procedure by replacing \( I_{n+1}(x|\theta) \) with \((I_{11}I_{22} - I_{12}^2)/I_{22}\) and \((I_{11}I_{22} - I_{12}^2)/I_{11}\), respectively.

2.3 Comparison of the Langlie and Neyer tests

As a preparation for further discussion, we investigate the performances of the Langlie and Neyer tests, from the viewpoints of the probability of getting an overlap and estimation precision. The true sensitivity model is assumed to be \( N(0, 1) \).

To proceed, we present a definition of the term “overlap”. Given a series of sensitivity data \((X_i, \delta_i) (i = 1, 2, \cdots, n)\), let \( X_{0U}^{(n)} \) and \( X_{1L}^{(n)} \) denote the largest stimulus that produces a non-response and the smallest stimulus that produces a response, respectively. We say that an overlap comes up, if and only if \( X_{0U}^{(n)} \geq X_{1L}^{(n)} \). Thus the probability of overlap after \( n \) experiments is defined as \( P(X_{0U}^{(n)} \geq X_{1L}^{(n)}) \).
If $X_{0U}^{(n)}$ or $X_{1L}^{(n)}$ has no definition, i.e., all experiments respond or none respond, then the probability of overlap will be zero.

### 2.3.1 Probability of obtaining an overlap

According to Silvapulle [16], unique MLE exist if and only if sensitivity data have an overlap. Thus, the probability of obtaining an overlap can be regarded as a rough criteria of evaluating different tests.

Since both the Langlie and Neyer tests need a lower bound $x_L$ and an upper bound $x_U$ of the mean, four arbitrary choices of $[x_L, x_U]$ are considered in our simulations, i.e., $[-3, 3], [-4, 5], [-5, 5]$ and $[-6, 6]$. Besides, noting that the Neyer test depends on $\sigma_{\text{guess}}$, we consider two choices of $\sigma_{\text{guess}}$, i.e., 1 and 3 (ND1, ND3 for short). For sample size $N$ is equal to 1, 2, · · · , 50, through 10,000 replications, we calculate the proportion of the samples in which there exists an overlap.

The simulation results are shown in Figure 1. It can be seen: (1) the ND1 test has uniformly larger probabilities than the ND3 test, which shows the probability of obtaining an overlap of the Neyer test is affected by the $\sigma_{\text{guess}}$; (2) the Langlie test has a comparable performance with the ND1 test and they have no much differences when sample size is greater than 10; and (3) the Langlie test has uniformly better performance than the ND3 test.

Note that the sample size is usually more than 10 in practice and it is not easy to get an accurate guess of $\sigma$, which plays down the efficiency of the Neyer test. Therefore, from the viewpoint of probabilities of getting an overlap, the Langlie test is promising and better than the Neyer test in most cases. In addition, it is also seen that the performance of the Neyer test is affected by $\sigma_{\text{guess}}$.

### 2.3.2 Precision of parameter estimates

In this subsection, we compare the two sensitivity tests from estimation precision of MLEs. Let $\hat{\theta}$ be the MLE of the parameter $\theta$, which may be $\theta = (\mu, \sigma)$, $\theta = \mu$ or $\theta = \sigma$. And define

$$SSB_{\theta} \equiv \frac{||\theta - \hat{\theta}||^2}{\sigma^2},$$

where $\sigma$ is the scale parameter. We take the average of $SSB_{\theta}$, denoted $\text{MSE}_{\theta}$, as the criteria of evaluating estimation precision.

In our simulations, four choices of sample size $N$, i.e., 20, 25, 30 and 40, and ten pairs of lower and upper bounds $[x_L, x_U]$, i.e., $[-3, 3], [-3, 5], [-4, 4], [-4, 6], [-4, 8], [-5, 5], [-5, 7], [-5, 10], [-6, 6]$ and $[-8, 10]$ are considered. We generate 2,000 data-sets from $N(0, 1)$ and estimate $\theta = (\mu, \sigma)$ by the Langlie and Neyer tests, respectively. The simulation results are presented in Table 1.

From Table 1, it is apparent that: (1) the precision of Neyer test is affected by the choice of $\sigma_{\text{guess}}$ since the ND1 test has uniformly better estimation precision than the ND3 test; (2) the Langlie and ND1 tests have very close performances when sample size is small such as 20, while the Langlie test is significantly inferior to the ND1 test for large sample sizes such as 25, 30 and 40; and (3) the Langlie test has an advantage than the ND3 test for small sample sizes, but the ND3 test gets better than the Langlie test as the sample size increases.

Overall, the Langlie test has better estimation precision for small samples, while the Neyer test has better estimation precision for large samples. This is probably because the Langlie test can get an overlap more quickly and therefore estimates the parameters by their MLEs earlier than the Neyer test, and the Neyer test is
more efficient in extracting information about the latent distribution after getting an overlap. However, the Langlie test may suffer from the inefficient choices of design points, and the performance of the Neyer test is affected by bad choices of $\sigma_{\text{guess}}$. We propose a new test in the next section, which inherits the advantages of the Langlie and Neyer tests, and overcomes their weaknesses.

3. The LND test

In Section 2, we know the efficiency of the Neyer test lies in the optimization procedure part, but it suffers from the bad choices of $\sigma_{\text{guess}}$ and the inefficiency of the first part, whose purpose is to get an overlap. On the other hand, the Langlie test can get an overlap stably and quickly and it does not need a pre-specified value of $\sigma_{\text{guess}}$, but it is inefficient to find the follow-up design points. Motivated by the Neyer and Langlie tests, we propose a new efficient sensitivity test, which replaces the first part of the Neyer test with the Langlie test, denoted as the LND test.

The new test requires only a lower bound $x_L$ and an upper bound $x_U$ for the mean. In our proposed test, the Langlie test is employed first to get an overlap as soon as possible. Once an overlap is obtained, the D-optimization procedure of the Neyer test is followed to find the optimal follow-up design points. The detailed procedures of the LND test are given as follows.

Step 1. Conduct the first experiment at $x_1 = (x_L + x_U)/2$ and set $n = 1$. Let $\delta_1$ denote the experimental result, where $\delta_i = 1$ denotes a response and $\delta_i = 0$ a non-response.

Step 2. $x_{n+1}$ is determined according to the follow rule. Let $k_0$ be the smallest $k$ such that $k = 2s_k$ where $s_k = \sum_{i=1}^k \delta_{n+1-i}$ ($k = 1, 2, \cdots, n$). If such a $k_0$ exists, then $x_{n+1} = (x_n + x_{k_0})/2$. Otherwise, $x_{n+1} = (x_n + x_L)/2$ if $\delta_n = 1$; $x_{n+1} = (x_n + x_U)/2$ if $\delta_n = 0$. Conduct an experiment at $x_{n+1}$ and record the experimental result. If an overlap comes up, go to Step 3; otherwise set $n = n + 1$ and repeat Step 2.

Step 3. Denote $\hat{\theta}_n$ as the MLE of $\theta$ based on the first $n$ experimental data. The next design point $x_{n+1} = \arg \max_x I_{n+1}(x|\hat{\theta}_n)$. Set $n = n + 1$, and repeat Step 3 until the number of experiments $n$ is equal to the sample size $N$.

4. Simulation Studies

In this section, we compare the performances of the proposed LND test, the Langlie, the Neyer and the D-S tests by simulations. The simulation data are generated from two common sensitivity models–Probit and Logit models. And evaluation criteria include probability of obtaining an overlap and estimation precision in terms of $\text{MSE}_\theta$.

Besides the $\text{MSE}_\theta$, in order to assess the overall performance between different settings, we adopt the relative mean index (RMI) which was used by Han and Tsung [7], Zou and Qiu [18]. Define

$$RMI = \frac{1}{K} \sum_{i=1}^K \frac{\text{MSE}^{(i)}_{\theta} - \text{MMSE}^{(i)}_{\theta}}{\text{MMSE}^{(i)}_{\theta}},$$

where $\text{MMSE}^{(i)}_{\theta}$ is the smallest $\text{MSE}^{(i)}_{\theta}$ of the $i$-th setting ($i = 1, \cdots, K$) among
all the test plans (the Langlie, the Neyer, the D-S and the proposed tests). Then, \( \frac{MSE_\theta^{(i)} - MMSE_\theta^{(i)}}{MMSE_\theta^{(i)}} \) can be considered as a relative efficiency measure of the \( MSE_\theta^{(i)} \), compared to the best one, and RMI is the average of all the relative efficiency values. A test plan with a smaller RMI value is considered better in its overall performance.

4.1 Probit model

Suppose the true model is \( N(0, 1) \), and the simulation setups are the same as those in subsection 2.3. For D-S test, we consider four prior distributions for \( \mu \) and \( \sigma \) (source code http://www.math.tau.ac.il/~dms/GLMDesign). The probabilities of overlap and the estimation precision of the D-S test (see Appendix A) are the best ones of the four priors.

From the viewpoint of getting an overlap, the LND test has the same capability as the Langlie test, is slightly inferior to the ND1 and D-S tests, but prior to the ND3 test. See Figure 1.

When evaluate a test method from estimation precision, we consider three types of parameters, i.e., \( \theta = (\mu, \sigma) \), \( \theta = \mu \) and \( \theta = \sigma \) in our simulations. The corresponding results are presented in Table 1, Figure 2 and Table 2, respectively.

After carefully study the simulation results for \( \theta = (\mu, \sigma) \), we have the following findings. Firstly, according to the RMI, the LND test has almost the same estimation precision as the ND1 test, the ideal Neyer test with \( \sigma_{\text{guess}} \) being the true value of \( \sigma \). The main reason is that the LND and ND1 share the same strategy after an overlap comes up. As the experiment number increases, these two tests have more and more in common, therefore their performances are getting closer and closer to each other. This implies that the Neyer test with a general \( \sigma_{\text{guess}} \) can not have better estimation precision than the LND test. Secondly, compared with Langlie and ND3 tests, the LND test always has better RMI. Since the strategy of the Langlie test produces an overlap very quickly and then estimates the parameters by MLEs, the LND test performs better than the ND3 test when the sample size is small. The LND test then adopts the D-optimization procedure to find follow-up design points after getting an overlap. As the D-optimization procedure is the most efficient way to extract information about the unknown parameters, it is not surprised to see that the LND test is better than the Langlie test in estimation precision. Those findings coincide with the design purpose of the LND test, i.e. having as good property of getting an overlap as the Langlie test for small sample sizes and as good estimation precision as the Neyer test for relatively large sample sizes. Finally, the propose test is also better than the D-S test for large sample size. As a by-product, it is found that the estimation precision of the D-S test are affected by the priors in Appendix A.

Since the D-S test relies on the priors and the computations are complicated, in the case of \( \theta = \mu \) and \( \theta = \sigma \), we only compare the Langlie, ND1, ND3 and LND tests based on 2,000 replications, where the \( I_{n+1}(x|\theta) \) should be replaced with \( (I_{11}I_{22} - I_{12}^2)/I_{22} \) and \( (I_{11}I_{22} - I_{12}^2)/I_{11} \), respectively.

In the case of \( \theta = \mu \), the four tests have similar and nice performances, which are illustrated in Figure 2. In the case of \( \theta = \sigma \), the LND test has uniformly better estimation precision than the Langlie test, by using the D-optimization procedure to determine the follow-up design points. Although the LND test is inferior to the ND1 test especially for small sample sizes, it is much better than the ND3 test. The latter is even worse than the Langlie test, which may be due to the bad guess value \( \sigma_{\text{guess}} = 3 \). For reasonable comparison, we add the simulation results of the
Sequential LND sensitivity test

Neyer test with $\sigma_{\text{guess}} = 1.5$ (ND1.5) in Table 2. It can be seen that the LND test also has better estimation precision than the ND1.5. And the ND1.5 is severely inferior to the ND1 although $\sigma_{\text{guess}} = 1.5$ is very close to the true value.

Put Table 2 about here

4.2 Logit model

Suppose the data are from logistic distribution $\text{Logistic}(\mu, \sigma)$, where $\mu$ and $\sigma$ are the location and scale parameters of logistic distribution, respectively. We consider four choices of sample size $N = 20, 25, 30, 40$ in simulation studies. Six pairs of lower and upper bounds $[x_L, x_U] = [2 - \pi/\sqrt{3}C, 2 + c\pi/\sqrt{3}C]$ are taken, where $C = 3, 3.5, 4, 4.5, 5, 5.5$.

In each setting, we generate 2,000 data-sets from $\text{Logistic}(2, 1)$ by the Langlie test, the two Neyer tests (ND1, ND3), the D-S test and the LND test, respectively. For the D-S test, according to Dror and Steinberg [6, pp. 296], our prior for the location parameter $\mu$ is uniform on $[x_L, x_U]$, and the prior for the scale parameter $\sigma$ is lognormal with a mean of $\log(1)$ and a standard deviation of 1. For each data-set, we estimate $\theta = (\mu, \sigma)$ by its MLE $\hat{\theta} = (\hat{\mu}, \hat{\sigma})$ and get the MSE $\text{MSE}(\mu, \sigma)$ of the five tests in Table 3. Furthermore, for sample size $N$ is equal to $1, 2, \cdots, 50$, through 10,000 replications, we get the probabilities of overlap of the Langlie (or the LND), the ND1, the ND3 and the D-S tests when $C = 3, 3.5, 4, 4.5$ in Figure 3, respectively.

From Figure 3, from the viewpoint of getting an overlap, we find that the L-ND test is slightly inferior to the ND1 test but prior to the ND3 test, which are consistent with the normal distribution. And it performs better than the D-S test. From Table 3, it is clear that the LND test has comparable estimation precision with the ND1, but is better than ND3. It can also be seen that the ND3 is severely inferior to the ND1, which shows that the performance of the Neyer test is affected by $\sigma_{\text{guess}}$ again.

4.3 Other models

Compared with the Neyer test, the proposed test does not involve a scale parameter estimation, which makes it more robust. Our simulation results under the above two location-scale models confirm this point. In this subsection, we show that this property makes the proposed test applicable even in non-location-scale models.

We consider two sensitivity distributions, neither of which is of the location-scale form $f((x - \mu)/\sigma)$. The first is a normal mixture distribution $0.5N(\mu_1, 0.5^2) + 0.5N(\mu_2, 0.5^2)$, where $\theta = (\mu_1, \mu_2)'$ is the parameter of interest. The second is a Weibull distribution $(\alpha, \beta)$, with $\alpha$ and $\beta$ being the scale and shape parameters, respectively. Here $\theta = (\alpha, \beta)'$ is of interest. As both the Neyer test and the D-S test were designed for location-scale distributed sensitivities, neither of them can be applied directly in the two models. Thus in this simulation, only the proposed LND test and the Langlie test are compared.

We generate 2,000 data-sets from both models with sample size $N = 20, 25, 30, 40$. In application of the LND test and the Langlie test, we choose lower and upper bounds to be $[x_L, x_U] = [-\sqrt{5c/2}, \sqrt{5c/2}]$ for the mixture model, and $[x_L, x_U] = [0, 0.5 + 0.5c]$ for the Weibull model. Six choices of the constant $c$ are considered, i.e., $c = 3, 3.5, 4, 4.5, 5, 5.5$. We compute the MSEs of the MLEs of the parameter $\theta$ under both models with data from the Langlie and LND tests, respectively. The simulation results are presented in Tables 4 and 5. We find that although the performances of the resulting estimators are influences by different
choices of the constant $c$, the proposed LND test always produces more efficient estimators than the Langlie test. This once again provides evidence for the priority of the LND test.

5. Conclusion

In this paper, based on the Langlie and Neyer test, we propose a new sensitivity test, the LND test. It inherits the nice properties of the Langlie and the Neyer tests in getting an overlap and having nice estimation precision, respectively. It is easy to implement and requires few initial parameters. Our simulation results provide the evidence that the proposed test has some superiority over the Langlie, Neyer and D-S tests.

According to the numerical examples, the proposed method is better than the Langlie and D-S tests, and has close performance with the ideal Neyer test for moderate and large sample sizes from the viewpoints of estimation precision and probability of obtaining an overlap. Therefore, our new test can be seen as a good all-around sensitivity test towards the Langlie and Neyer test. Meanwhile, in contrast to the Neyer and D-S tests, the LND is more robust since it does not need a guess value of the standard dispersion or the priors for the parameters.

Note that the LND test is general since it can be applied to the arbitrary sensitivity distributions. Although we assume the true sensitivity distributions are $N(0,1)$ and Logistic$(2,1)$ in our simulations, like some other tests, the LND test can generally be applied in making inferences for normal, log-normal, logistic and log-logistic and other parametric sensitivity populations. The only modification is to change the determinant of the Fisher information matrix in the LND test procedure.

Acknowledgements

The authors thank the Associate Editor and two anonymous referees for their valuable and constructive comments which significantly improves the manuscript. This work was supported by National Science Fund of China (11271135, 11001083, 11101156) and China Scholarship Council.

References


Appendix A: Simulation results of Dror and Steinberg test

Dror and Steinberg [6] considered the problem of experimental design when the response is modeled by a generalized linear model. It also can be applied to multifactor experiments. Based on a D-optimality criterion and on a Bayesian analysis that exploits a discretization of the parameter space to efficiently represent the posterior distribution, D-S test shows some superiorities in efficiency to Neyer test and Bruceton test (Dixon and Mood [5]), especially for the small sample size. However, it typically includes more than just two inputs. The priors are obviously important and will affect the performance of the D-S test, which can be found in our simulation results.

Since Dror and Steinberg [6] advised making the spread of the prior large, for normal sensitivity distribution in Section 4, a natural choice of D-S test is to take a prior with $\mu_1 = 0.5(x_L + x_U)$ as the prior mean for $\mu$, and $\mu_2 = 0.1(x_U - x_L)$ as the prior mean for $\sigma$, so that we can make the interval corresponding to a $10\sigma$ spread. Therefore, the priors in the D-S test can be written as follows:

$$\mu \sim \text{normal}(\mu_1, \sigma_1^2) \quad \text{and} \quad \sigma \sim \text{lognormal}(\mu_2, \sigma_2^2).$$

In addition, the prior variances are selected to ensure that $f(x_L)$ and $f(x_U)$ are with high prior probabilities close to 0 and 1, respectively. Then, four choices of $(\sigma_1, \sigma_2)$ are considered: (I) $(x_U - x_L)/6, \mu_2/2$; (II) $(x_U - x_L)/10, \mu_2/2$; (III) $(x_U - x_L)/8, \mu_2/4)$ and (IV) $(x_U - x_L)/10, \mu_2/4$.

For sample size equal to 1, 2, ⋯, 50, through 10,000 replications, we get the probabilities of overlap of these four priors. The results are shown in Figure 1, where the probabilities is based on the prior IV. The corresponding simulation results on estimation precision based on 2,000 replications are presented in Table 7.
Table 1. MSE(\(\sigma, \mu\))(\(\times 10^{-2}\)) and RMI comparisons for the parameter of interest \(\theta = (\mu, \sigma)\) when the population distribution is \(N(0,1)\).

<table>
<thead>
<tr>
<th>Tests</th>
<th>Langlie</th>
<th>ND1</th>
<th>ND3</th>
<th>D-S</th>
<th>LND</th>
<th>Langlie</th>
<th>ND1</th>
<th>ND3</th>
<th>D-S</th>
<th>LND</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bounds</td>
<td>N=20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>N=25</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[-3, 3]</td>
<td>32.82</td>
<td>32.02</td>
<td>37.57</td>
<td>25.60</td>
<td>31.20</td>
<td>26.54</td>
<td>24.08</td>
<td>28.37</td>
<td>24.18</td>
<td>25.08</td>
</tr>
<tr>
<td>[-3, 5]</td>
<td>34.01</td>
<td>33.93</td>
<td>38.21</td>
<td>25.87</td>
<td>33.88</td>
<td>26.34</td>
<td>25.92</td>
<td>28.74</td>
<td>22.86</td>
<td>25.40</td>
</tr>
<tr>
<td>[-4, 4]</td>
<td>34.60</td>
<td>32.77</td>
<td>44.84</td>
<td>17.96</td>
<td>32.77</td>
<td>27.58</td>
<td>25.70</td>
<td>34.24</td>
<td>16.31</td>
<td>25.30</td>
</tr>
<tr>
<td>[-4, 6]</td>
<td>34.20</td>
<td>30.37</td>
<td>39.23</td>
<td>25.62</td>
<td>33.86</td>
<td>25.82</td>
<td>23.34</td>
<td>29.30</td>
<td>21.78</td>
<td>25.74</td>
</tr>
<tr>
<td>[-4, 8]</td>
<td>34.90</td>
<td>33.71</td>
<td>40.12</td>
<td>46.18</td>
<td>35.61</td>
<td>27.18</td>
<td>24.59</td>
<td>30.13</td>
<td>34.80</td>
<td>26.15</td>
</tr>
<tr>
<td>[-5, 5]</td>
<td>34.41</td>
<td>33.26</td>
<td>41.95</td>
<td>18.67</td>
<td>34.21</td>
<td>28.04</td>
<td>25.43</td>
<td>30.13</td>
<td>16.99</td>
<td>27.31</td>
</tr>
<tr>
<td>[-5, 7]</td>
<td>35.63</td>
<td>33.15</td>
<td>40.23</td>
<td>31.49</td>
<td>34.56</td>
<td>27.00</td>
<td>23.35</td>
<td>30.45</td>
<td>27.06</td>
<td>25.87</td>
</tr>
<tr>
<td>[-5, 10]</td>
<td>35.21</td>
<td>35.35</td>
<td>41.61</td>
<td>65.03</td>
<td>36.36</td>
<td>28.73</td>
<td>26.87</td>
<td>31.85</td>
<td>50.80</td>
<td>25.62</td>
</tr>
<tr>
<td>[-6, 6]</td>
<td>33.57</td>
<td>31.55</td>
<td>41.95</td>
<td>25.11</td>
<td>33.97</td>
<td>26.67</td>
<td>24.57</td>
<td>30.13</td>
<td>22.40</td>
<td>27.92</td>
</tr>
<tr>
<td>[-8, 10]</td>
<td>40.07</td>
<td>35.73</td>
<td>40.23</td>
<td>56.04</td>
<td>36.91</td>
<td>29.45</td>
<td>27.22</td>
<td>30.45</td>
<td>46.83</td>
<td>30.03</td>
</tr>
</tbody>
</table>

RMI 0.333 0.267 0.567 0.179 0.308 0.235 0.142 0.379 0.219 0.193

Table 2. MSE(\(\sigma\))(\(\times 10^{-2}\)) and RMI comparisons for the parameter of interest \(\theta = \sigma\) when the population distribution is \(N(0,1)\).

<table>
<thead>
<tr>
<th>Tests</th>
<th>Langlie</th>
<th>ND1</th>
<th>ND1.5</th>
<th>ND3</th>
<th>D-S</th>
<th>LND</th>
<th>Langlie</th>
<th>ND1</th>
<th>ND1.5</th>
<th>ND3</th>
<th>D-S</th>
<th>LND</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bounds</td>
<td>N=30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>N=40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[-5, 10]</td>
<td>23.90</td>
<td>21.65</td>
<td>24.58</td>
<td>42.93</td>
<td>20.65</td>
<td>17.56</td>
<td>15.25</td>
<td>16.32</td>
<td>30.95</td>
<td>14.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[-8, 10]</td>
<td>23.68</td>
<td>22.50</td>
<td>23.29</td>
<td>39.79</td>
<td>23.38</td>
<td>17.33</td>
<td>15.56</td>
<td>15.83</td>
<td>30.81</td>
<td>15.80</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

RMI 0.203 0.084 0.261 0.278 0.123 0.239 0.039 0.146 0.411 0.061
Table 3. MSE \((\mu, \sigma) \times 10^{-2}\) and RMI comparisons for the parameter of interest \(\theta = (\mu, \sigma)\) when the population distribution is Logistic(2.1).

<table>
<thead>
<tr>
<th>Tests</th>
<th>Langlie</th>
<th>ND1</th>
<th>ND3</th>
<th>D-S</th>
<th>LND</th>
<th>Langlie</th>
<th>ND1</th>
<th>ND3</th>
<th>D-S</th>
<th>LND</th>
</tr>
</thead>
<tbody>
<tr>
<td>c N=20</td>
<td>N=25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.0</td>
<td>65.94</td>
<td>54.41</td>
<td>68.39</td>
<td>56.73</td>
<td><strong>58.24</strong></td>
<td>47.47</td>
<td>43.99</td>
<td>52.34</td>
<td>45.28</td>
<td><strong>46.80</strong></td>
</tr>
<tr>
<td>3.5</td>
<td>67.76</td>
<td>56.15</td>
<td>70.49</td>
<td>56.00</td>
<td><strong>57.30</strong></td>
<td>48.79</td>
<td>45.64</td>
<td>54.17</td>
<td>45.46</td>
<td><strong>45.93</strong></td>
</tr>
<tr>
<td>4.0</td>
<td>60.28</td>
<td>56.49</td>
<td>68.83</td>
<td>57.24</td>
<td><strong>60.98</strong></td>
<td>47.19</td>
<td>44.58</td>
<td>53.94</td>
<td>45.29</td>
<td><strong>48.30</strong></td>
</tr>
<tr>
<td>4.5</td>
<td>68.66</td>
<td>57.41</td>
<td>71.01</td>
<td>60.31</td>
<td><strong>59.53</strong></td>
<td>44.68</td>
<td>46.22</td>
<td>54.22</td>
<td>48.44</td>
<td><strong>46.61</strong></td>
</tr>
<tr>
<td>5.0</td>
<td>59.61</td>
<td>58.15</td>
<td>67.47</td>
<td>59.08</td>
<td><strong>59.79</strong></td>
<td>47.81</td>
<td>46.88</td>
<td>53.31</td>
<td>46.09</td>
<td><strong>47.56</strong></td>
</tr>
<tr>
<td>5.5</td>
<td>60.96</td>
<td>59.84</td>
<td>67.25</td>
<td>57.80</td>
<td><strong>59.41</strong></td>
<td>47.16</td>
<td>47.32</td>
<td>54.14</td>
<td>45.27</td>
<td><strong>47.15</strong></td>
</tr>
<tr>
<td>RMI</td>
<td>0.128</td>
<td>0.006</td>
<td>0.216</td>
<td>0.020</td>
<td><strong>0.044</strong></td>
<td>0.046</td>
<td>0.015</td>
<td>0.190</td>
<td>0.022</td>
<td><strong>0.043</strong></td>
</tr>
</tbody>
</table>

Table 4. MSE \((\mu_1, \mu_2) \times 10^{-2}\) for \(\theta = (\mu_1, \mu_2)\) when the population distribution is the normal mixture 0.5N\((\mu_1, 0.5^2)\) + 0.5N\((\mu_2, 0.5^2)\).

<table>
<thead>
<tr>
<th>Tests</th>
<th>Langlie</th>
<th>LND</th>
<th>Langlie</th>
<th>LND</th>
<th>Langlie</th>
<th>LND</th>
<th>Langlie</th>
<th>LND</th>
</tr>
</thead>
<tbody>
<tr>
<td>c N=20</td>
<td>N=25</td>
<td>N=30</td>
<td>N=40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.0</td>
<td>82.45</td>
<td>68.16</td>
<td>76.03</td>
<td>59.69</td>
<td>68.44</td>
<td>52.40</td>
<td>59.63</td>
<td>45.00</td>
</tr>
<tr>
<td>3.5</td>
<td>83.65</td>
<td>67.65</td>
<td>75.27</td>
<td>58.00</td>
<td>69.05</td>
<td>51.58</td>
<td>61.38</td>
<td>44.74</td>
</tr>
<tr>
<td>4.0</td>
<td>86.47</td>
<td>68.89</td>
<td>79.35</td>
<td>57.37</td>
<td>71.15</td>
<td>49.72</td>
<td>62.73</td>
<td>45.15</td>
</tr>
<tr>
<td>4.5</td>
<td>83.22</td>
<td>69.98</td>
<td>76.21</td>
<td>59.15</td>
<td>70.08</td>
<td>51.40</td>
<td>59.88</td>
<td>43.57</td>
</tr>
<tr>
<td>5.0</td>
<td>85.59</td>
<td>70.22</td>
<td>77.59</td>
<td>58.28</td>
<td>68.94</td>
<td>50.52</td>
<td>59.86</td>
<td>43.72</td>
</tr>
<tr>
<td>5.5</td>
<td>86.74</td>
<td>70.31</td>
<td>78.84</td>
<td>60.30</td>
<td>70.59</td>
<td>51.67</td>
<td>60.69</td>
<td>44.27</td>
</tr>
<tr>
<td>RMI</td>
<td>0.053</td>
<td>0.029</td>
<td>0.166</td>
<td>0.034</td>
<td><strong>0.010</strong></td>
<td>0.040</td>
<td>0.022</td>
<td>0.133</td>
</tr>
</tbody>
</table>

Table 5. MSE \((\alpha, \beta) \times 10^{-2}\) for \(\theta = (\alpha, \beta)\) when the population distribution is the Weibull model \((0.5, 1)\).

<table>
<thead>
<tr>
<th>Tests</th>
<th>Langlie</th>
<th>LND</th>
<th>Langlie</th>
<th>LND</th>
<th>Langlie</th>
<th>LND</th>
<th>Langlie</th>
<th>LND</th>
</tr>
</thead>
<tbody>
<tr>
<td>c N=20</td>
<td>N=25</td>
<td>N=30</td>
<td>N=40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.0</td>
<td>55.50</td>
<td>43.82</td>
<td>48.21</td>
<td>35.01</td>
<td>43.54</td>
<td>30.08</td>
<td>35.43</td>
<td>22.80</td>
</tr>
<tr>
<td>3.5</td>
<td>57.01</td>
<td>42.82</td>
<td>48.84</td>
<td>36.20</td>
<td>43.19</td>
<td>29.50</td>
<td>36.32</td>
<td>20.93</td>
</tr>
<tr>
<td>4.0</td>
<td>54.33</td>
<td>39.52</td>
<td>47.08</td>
<td>31.73</td>
<td>41.88</td>
<td>27.77</td>
<td>35.98</td>
<td>19.86</td>
</tr>
<tr>
<td>4.5</td>
<td>56.93</td>
<td>39.82</td>
<td>50.20</td>
<td>33.08</td>
<td>45.82</td>
<td>27.20</td>
<td>37.63</td>
<td>21.42</td>
</tr>
<tr>
<td>5.0</td>
<td>56.21</td>
<td>42.71</td>
<td>49.14</td>
<td>35.57</td>
<td>44.09</td>
<td>29.04</td>
<td>35.86</td>
<td>21.19</td>
</tr>
<tr>
<td>5.5</td>
<td>55.72</td>
<td>41.75</td>
<td>49.75</td>
<td>33.40</td>
<td>43.30</td>
<td>28.08</td>
<td>36.22</td>
<td>21.19</td>
</tr>
</tbody>
</table>
Table 6. MSE($\mu, \sigma$)($\times 10^{-2}$) of the D-S test for the parameter of interest $\theta = (\mu, \sigma)$, N(0, 1) case.

<table>
<thead>
<tr>
<th>Priors</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IIIIV</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IIIIV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N=20</td>
<td></td>
<td></td>
<td></td>
<td>N=25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[--3, 5]</td>
<td>29.10</td>
<td>27.19</td>
<td>25.63</td>
<td>25.87</td>
<td>25.32</td>
<td>22.77</td>
<td>22.06</td>
<td>22.86</td>
</tr>
<tr>
<td>[--4, 4]</td>
<td>27.20</td>
<td>19.75</td>
<td>22.89</td>
<td>17.96</td>
<td>23.64</td>
<td>17.83</td>
<td>20.49</td>
<td>16.31</td>
</tr>
<tr>
<td>[--4, 6]</td>
<td>33.61</td>
<td>30.31</td>
<td>24.90</td>
<td>25.62</td>
<td>29.30</td>
<td>25.85</td>
<td>22.05</td>
<td>21.78</td>
</tr>
<tr>
<td>[--4, 8]</td>
<td>42.71</td>
<td>54.47</td>
<td>39.68</td>
<td>46.18</td>
<td>36.50</td>
<td>44.82</td>
<td>33.43</td>
<td>34.80</td>
</tr>
<tr>
<td>[--5, 10]</td>
<td>40.16</td>
<td>38.36</td>
<td>31.72</td>
<td>31.49</td>
<td>33.79</td>
<td>32.81</td>
<td>27.16</td>
<td>27.06</td>
</tr>
<tr>
<td>[--6, 10]</td>
<td>35.95</td>
<td>70.98</td>
<td>58.51</td>
<td>65.03</td>
<td>43.15</td>
<td>54.41</td>
<td>50.80</td>
<td>50.80</td>
</tr>
<tr>
<td>[--6, 10]</td>
<td>40.08</td>
<td>29.81</td>
<td>25.11</td>
<td>20.08</td>
<td>23.64</td>
<td>21.53</td>
<td>20.33</td>
<td>16.99</td>
</tr>
<tr>
<td>N=30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>N=40</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[--3, 5]</td>
<td>22.57</td>
<td>20.90</td>
<td>19.01</td>
<td>20.08</td>
<td>18.04</td>
<td>17.29</td>
<td>15.91</td>
<td>16.34</td>
</tr>
<tr>
<td>[--4, 6]</td>
<td>24.92</td>
<td>22.08</td>
<td>18.95</td>
<td>19.33</td>
<td>20.42</td>
<td>18.59</td>
<td>15.77</td>
<td>15.59</td>
</tr>
<tr>
<td>[--4, 8]</td>
<td>32.21</td>
<td>39.02</td>
<td>27.95</td>
<td>30.15</td>
<td>24.62</td>
<td>28.01</td>
<td>22.40</td>
<td>23.01</td>
</tr>
<tr>
<td>[--5, 10]</td>
<td>29.45</td>
<td>28.71</td>
<td>23.26</td>
<td>24.31</td>
<td>23.08</td>
<td>22.45</td>
<td>19.83</td>
<td>18.33</td>
</tr>
<tr>
<td>[--5, 10]</td>
<td>35.06</td>
<td>42.29</td>
<td>37.78</td>
<td>42.93</td>
<td>26.51</td>
<td>30.69</td>
<td>29.41</td>
<td>30.95</td>
</tr>
<tr>
<td>[--6, 6]</td>
<td>29.13</td>
<td>23.74</td>
<td>21.75</td>
<td>20.08</td>
<td>23.88</td>
<td>20.10</td>
<td>18.68</td>
<td>16.57</td>
</tr>
<tr>
<td>[--6, 10]</td>
<td>41.62</td>
<td>37.43</td>
<td>40.19</td>
<td>39.79</td>
<td>30.97</td>
<td>28.20</td>
<td>31.10</td>
<td>30.81</td>
</tr>
</tbody>
</table>

Figure 1. Comparison of probabilities of getting an overlap for different choices of $[x_L, x_U]$: (a) $[-3, 3]$, (b) $[-4, 4]$, (c) $[-5, 5]$ and (d) $[-6, 6]$ (N(0, 1) case).
Figure 2. Comparison of MSE$_\mu$ for different sample size $N$: (a) $N=20$, (b) $N=25$, (c) $N=30$ and (d) $N=40$ ($N(0, 1)$ case)
Figure 3. Comparison of probabilities of getting an overlap for different choices of $[x_L, x_U] = [2 - \frac{c\pi}{\sqrt{3}}, 2 + \frac{c\pi}{\sqrt{3}}]$: (a) $c=3$, (b) $c=3.5$, (c) $c=4$ and (d) $c=4.5$ (logistic(2, 1) case)