13th International Conference on Probabilistic Safety Assessment and Management (PSAM 13)
2~7 October, 2016 • Sheraton Grande Walkerhill • Seoul, Korea • www.psam13.org

NEW METHODS FOR SENSITIVITY TESTS OF EXPLOSIVE DEVICES

Amit Teller[1], David M. Steinberg[2], Lina Teper[1], Rotem Rozenblum[2], Liran Mendel[2], and Mordechai Jaeger[2]

[1] RAFAEL, POB 2250, Haifa, 3102102, Israel
[2] Tel-Aviv University, Tel-Aviv, 69978, Israel

The objective of this research is to examine two methods of experimental design for sensitivity testing, in which the goal is to study the probability of a response (go/no go) as a function of a stimulus. The comparison was carried out between the widely used Bruceton (up-down) method and the Dror-Steinberg method which is based on the Bayesian approach and is one of the modern approaches for such tests.

The utilization of new approaches for carrying out sensitivity tests brings opportunities to make the tests more efficient in terms of the number of trials required to achieve the desired level of confidence. It also makes it possible to provide a better level of confidence for reliability estimation. The main focus of this study is on a special experiment that was aimed at comparing simultaneously the two methods

The results show that the Bruceton procedure provides good estimation for the mean threshold value for response; however, as expected, improvements in the estimation of the standard deviation of the threshold distribution can be potentially achieved by using the new Bayesian approach. Under cautious design the new approach has also a potential for reducing the number of trials required to achieve the desired level of confidence for the mean, the standard deviation ans various quantiles..

I. INTRODUCTION

I.A. Review of Design Approaches

The focus of this study is to advance research on experimental design for sensitivity testing, in which the goal is to study the probability of a response (go/no go) as a function of a stimulus. Typical examples for such required investigations are the probability of explosion as a function of voltage in a trip wire or the probability of squib ignition as a function of the current. Other examples for such systems are smoke detectors, car safety belts and more.

The testing plan involves deciding which level to use as stimuli. The most widely used method is the so-called Bruceton (or up-and-down) test, developed by Dixon and Mood (Ref. 2). The Bruceton protocol has been used for many years in various industrial applications including the manufacture of explosive devices and the development of protective systems. Examples for the utilization of Bruceton tests are MIL-STD-1751A (Ref. 5) for qualification of explosives and ASTM standard No. D2463 (Ref. 6) for drop testing of plastic containers. The Bruceton method calls for running tests on a grid of stimulus values, which are expected to span a range from almost certain non-response to almost certain response. The protocol balances testing over this range by "stepping up" to the next highest stimulus value after a non-response and "stepping down" to the next smallest stimulus value after a response. In some experiments, the step size is changed during the course of the experiment if it proves to be too large or too small. The Bruceton method tends to concentrate most of the experimental stimuli near the median of the sensitivity distribution (i.e. the stimulus that has a 50% response probability). In several studies it has been shown to give good estimates of the median, but poor estimates of the spread (e.g. SD) of the distribution (Ref. 7).

It is common to assume that the sensitivity distribution has a parametric form, and usually that it is a normal distribution with mean μ and standard deviation σ, so that the probability of response for a stimulus x is $\Psi[(x-\mu)/\sigma]$. Then the study design can be directed toward estimating the parameters. Neyer (1994) followed this approach (Ref. 7), using the D-optimality criterion from the statistical design of experiments to guide the choice of stimuli. The criterion looks at the covariance matrix of the joint estimators of μ and σ and attempts to minimize the determinant of that matrix. The choice of stimuli (in fact the covariance matrix itself) depends on the values of μ and σ, so some preliminary knowledge is needed about their values. Neyer (Ref. 7) proposed an initial phase of the experiment that would be sufficient to estimate these parameters from the data and then chose subsequent stimuli by optimizing the design criterion, assuming that the current parameter estimates are the actual values. He showed that his method led to much better estimators than the Bruceton protocol. The Neyer (Ref. 7) method is referred in MIL-STD-331C (Ref. 6) as one of the optional sensitivity test protocols.

Dror and Steinberg (2008) proposed in Ref. 3 a design protocol that is rooted in Bayesian statistical analysis. Like Neyer, they emphasize precise estimation of the parameters in a statistical model and their goal is to eventually use optimal design theory to select each new stimulus. However, where Neyer collects initial data until it is possible to compute maximum likelihood estimates, Dror and Steinberg (D-S) take advantage of prior information from the experimenters to focus attention on plausible parameter values. Their method can outperform that of Neyer, especially when there is good prior information about the system under study.

There is also a family of methods that prefers to avoid making parametric modeling assumptions. The original article in this group was published by Robbins and Monro (Ref. 8) and developed a clever method for estimating a specified quantile x_p of the probability curve. The idea is to conduct a sequence of tests with a step-up and step-down scheme, as in the Bruceton protocol. However, whereas Bruceton uses a fixed step size, the Robbins-Monro method uses a decreasing step size, with the decrease at a particular rate. They proved that the sequence of stimuli in the experiment converges to the desired quantile x_p. Wu (Ref. 10) improved the method, exploiting a useful parallel to logistic regression models. Joseph (Ref. 4) made further improvements in the method and the latest contribution in this track was made by Wu and Tian (Ref. 11) who proposed a three-phase approach, called 3pod, to sensitivity testing for estimating a single quantile. Their protocol begins, like that of Neyer, with a small number of tests that permit fitting a parametric model. The second phase of their protocol, again like Neyer, selects stimuli for precise estimation of a parametric model. The final phase switches to Joseph's modified Robbins-Monro scheme (Ref. 4) for efficient estimation of x_p even when the parametric model used in the earlier phases was not valid.

I.B. Review of Analysis Approaches

Sensitivity experiments produce data of the form $\{x_i,y_i\}$, $i=1,\ldots,n$, where x_i is the value of the stimulus on the i'th test and y_i is 1 for a response and 0 for a non-response. The parametric models assume that the probability of a response is a function of the stimulus, $P(Y(x)=1)=g(x)$. For example, the probit model assumes that $P(Y(x)=1) = \Psi[(x-\mu)/\sigma]$. The data can be used to construct a likelihood function $L(\mu,\sigma)$ from which the parameters can be estimated by maximum likelihood. The method also provides approximate standard errors for the estimators. Statistical inference for any function of the parameters (e.g. response probabilities or quantiles) can be carried out via first-order Taylor series expansion.

Dror and Steinberg (Ref. 3) used a Bayesian analysis; in large samples this will give approximately the same answers as the maximum likelihood method, but in small samples it avoids the issue that maximum likelihood estimators may not exist and for extreme quantiles it may account better for lack of normality or symmetry. The Bayesian analysis begins with a prior distribution $\pi(\mu,\sigma)$ reflecting what is believed about the parameters before testing is begun. Once data are available, the prior distribution is multiplied by the likelihood and normalized to produce a posterior distribution. For sensitivity experiments, there is no simple analytical solution for the posterior distribution. Therefore a numerical scheme is needed to compute it. Dror and Steinberg used a discrete approximation in which a large number of μ,σ pairs is sampled from the prior and then are weighted proportionally to their likelihood. In Ref. 9 Steinberg et al. provided initial results indicating that the resulting summaries give reliable inference for both the parameters themselves and for derived quantities like response probabilities and quantiles.

The Robbins-Monro type procedures (Ref. 8), including 3pod (Ref. 11), differ from the above in that they avoid assuming a parametric model. They lead to good estimators for the specific quantile to which the design protocol is targeted. As the test stimuli tend to be strongly concentrated near the target quantile, they typically have little information about other quantiles.

The major benefit of the non-parametric approach is that it guarantees good results even when the sensitivity distribution does not match any standard family, or when we assume an incorrect parametric form. The drawback is that inference is far more limited than what is possible in the parametric framework. The size of qualification tests for explosives are often too small to successfully employ non-parametric methods. Thus in this research we focus here only on parametric methods.

II. CASE STUDY

II.A. Test Case Experiment for comparison between Bruceton and D-S methods

A test case was run to compare the standard Bruceton protocol and the design scheme proposed by Dror-Steinberg (D-S). The two methods were applied in parallel on an experiment that examined the sensitivity distribution of a production lot of detonators, with alternate trials from each of the two protocols.

The experiment began with a "step up" test of a single unit, which fired at current of 1.25 Amp. The Bruceton test began "stepping down" from 1.25 Amp in steps of 0.02 Amp. The 1.25 Amp current level was selected as being a reasonable first guess for the median, and based on previous experience the engineers running the experiment believed it very likely that the median was between 1.15 Amp and 1.35 Amp. For the purpose of this study it was quite reasonable to assume that the Standard Deviation (SD) would be about 5% of the median.

II.A.1. D-S Design

In order to apply the D-S method it is required to assume a prior distribution based on previous experience. For the purpose of comparing the methods the following prior was adopted:

$\mu \sim N(1.25, 0.04)$

The prior for the SD (σ) is derived using the proportionality to the median. The prior is found by defining $\sigma=\alpha\mu$ and then sampling α from a log-normal distribution with mean, on the log scale, of 0.05, and SD, on the log scale, of 2.5. The mean value reflects the belief that σ should be approximately 5% of μ; the spread gives a fairly broad prior distribution for σ.

Fig. 1 shows the results of 27 experiments where the D-S method was applied:

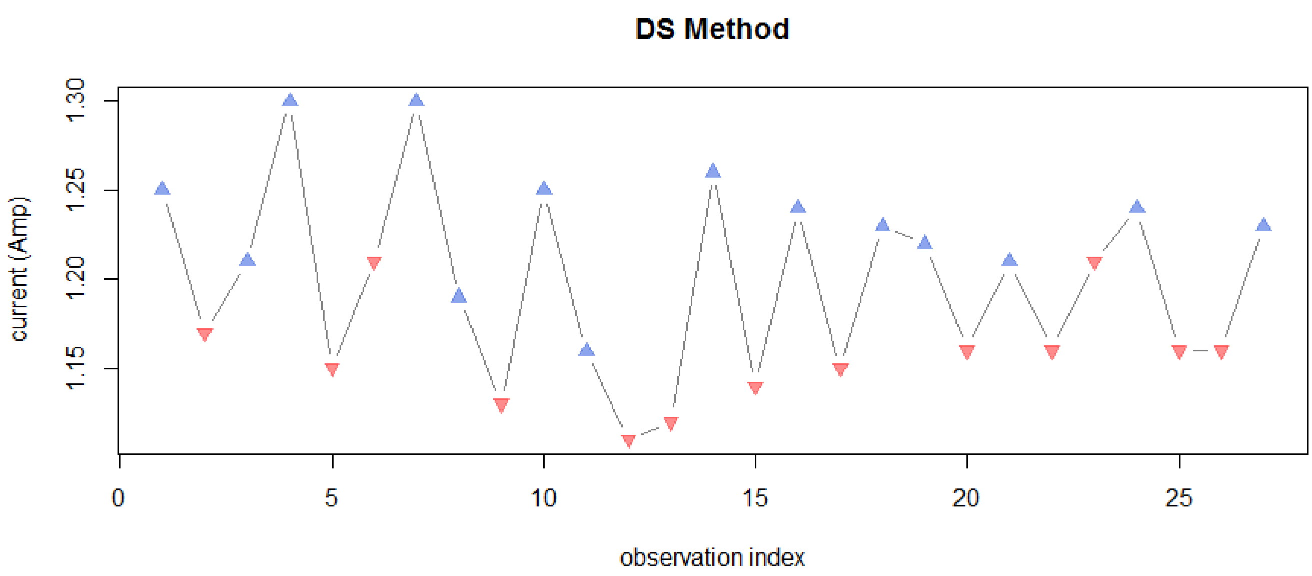

Fig. 1. Test results from applying the D-S method, Blue indicates "fire" and red "no fire".

The Bayesian analysis of D-S applied to all 27 trials gives the posterior distributions for the two parameters as shown in Figs. 2 and 3. The 95% credible interval for the mean μ is from 1.174 to 1.222 Amp. The posterior median and mean are both 1.196 Amp. As for σ, the median is 0.0328 Amp and the 95% credible interval is from 0.0181 to 0.0765 Amp.

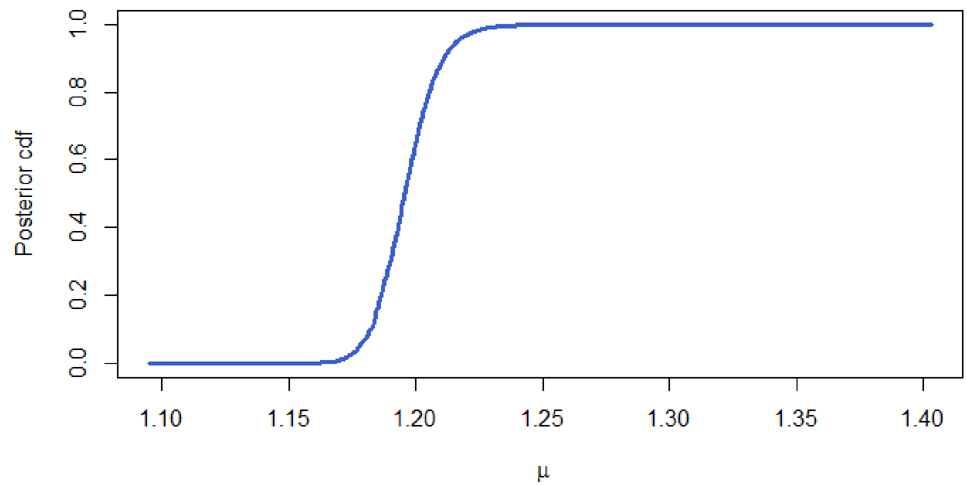

Fig. 2. Posterior Cumulative Function (CDF) for μ.

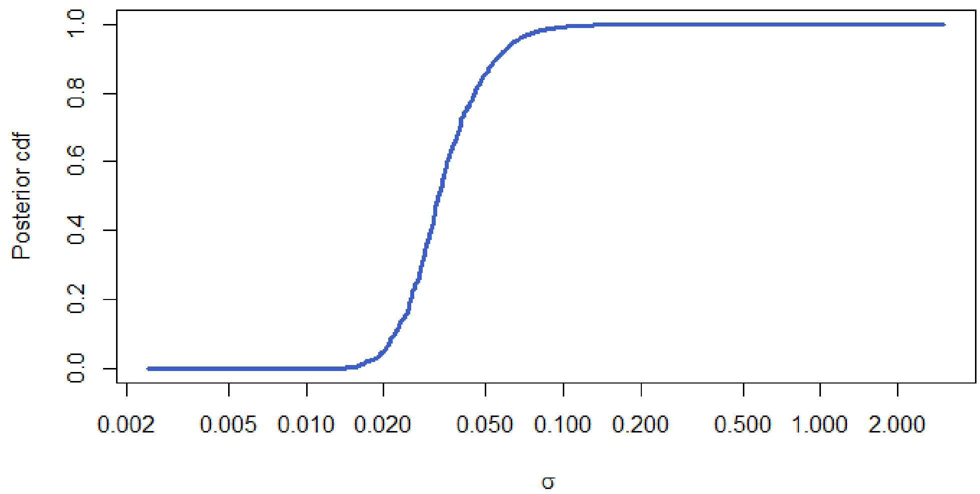

Fig. 3. Posterior Cumulative Distribution Function (CDF) for σ (plotted on a log scale).

The D-S data were also analyzed using a standard probit regression model. The resulting estimates based on the probit analysis for μ and σ were 1.192 and 0.0268 Amp, respectively, with direct standard errors (from Taylor expansions) of 0.0098 and 0.0083 Amp. Inference for the parameters is shown in Table I, alongside the inference from the DS analysis. For σ, Table I shows inference based on a normal approximation to the distribution of $\hat{\sigma}$ and also on a normal approximation to the distribution of log($\hat{\sigma}$). In Table I upper and lower bounds are 95% credible limits (for the D-S analysis) and approximate 95% confidence intervals (for the probit analysis).

TABLE I. Inference for μ and σ for the D-S method

	Estimate	**Lower Bound**	**Upper Bound**
D-S results for μ	**1.196**	**1.174**	**1.222**
Probit results for μ	**1.192**	**1.173**	**1.211**
D-S results for σ	**0.0328**	**0.0181**	**0.0765**
Probit results for σ			
Direct results	**0.0268**	**0.0105**	**0.0431**
Log scale results	**0.0268**	**0.0146**	**0.0490**

II.A.2. Bruceton probit analysis

Fig. 4 shows the results of 25 experiments where the D-S method was applied:

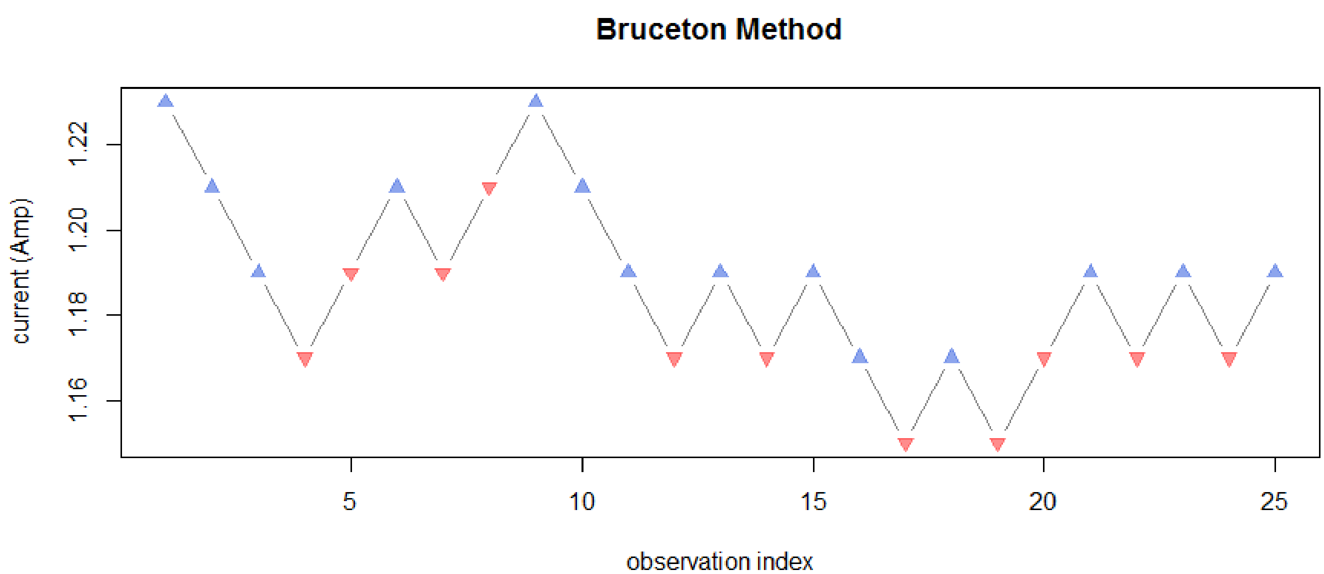

Fig. 4. Test results from applying the Bruceton method, Blue indicates “fire” and red “no fire”.

Based on the classic Bruceton method, the resulting estimates for μ and σ are 1.182 Amp and 0.0215 Amp, respectively, with direct standard errors (from Taylor expansions) of 0.0062 Amp and 0.0084 Amp. Inference for the parameters is shown in Table II. For σ, Table II shows inference based on a normal approximation to the distribution of $\hat{\sigma}$ and also on a normal approximation to the distribution of log($\hat{\sigma}$).

TABLE II - Inference for μ and σ for the Bruceton method

	Estimate	Lower Bound	Upper Bound
Probit results for μ	**1.182**	**1.170**	**1.194**
Probit results for σ			
Direct results	**0.0215**	**0.0049**	**0.0380**
Log scale results	**0.0215**	**0.0099**	**0.0464**

II.A. Comparison of the D-S and the Bruceton designs

The probit results from the two designs are similar. The D-S method (with two additional trials compared to Bruceton) has a higher standard error for estimating μ and almost the same standard error for estimating σ as compared to Bruceton. We expected Bruceton to give better results for the center of the distribution, so the first result is expected. The second result does not correspond to our expectations – the D-S method is claimed to give better standard errors for σ. This disagreement is explained as follows: The approximate standard errors for the two methods are proportional to the estimated value of σ and, in the Bruceton trials, that estimate was lower by 20% than in the D-S trials. If the same estimate of σ (say the estimate from analyzing all 52 trials together, not shown here) were used to compute the standard error for both sets of data, the D-S results would have a standard error that was smaller by 20%.

More information is available by analyzing the results at various possible early stopping points. Figs. 5 and 6 summarize the estimates and confidence intervals for the two parameters for 15, 20 and 25 trials. Fig. 5 shows the result for the mean (μ) and Fig. 6 uses the log scale confidence intervals for σ. As for the mean, the Bruceton experiment gave consistently narrower confidence intervals.

Similar to the comparison between the entire trials, the major difference between the two sets is that the Bruceton sequence led to lower estimates of σ, and correspondingly lower standard error for σ. When the confidence intervals are corrected for this feature, the D-S design gives more precise estimates of σ.

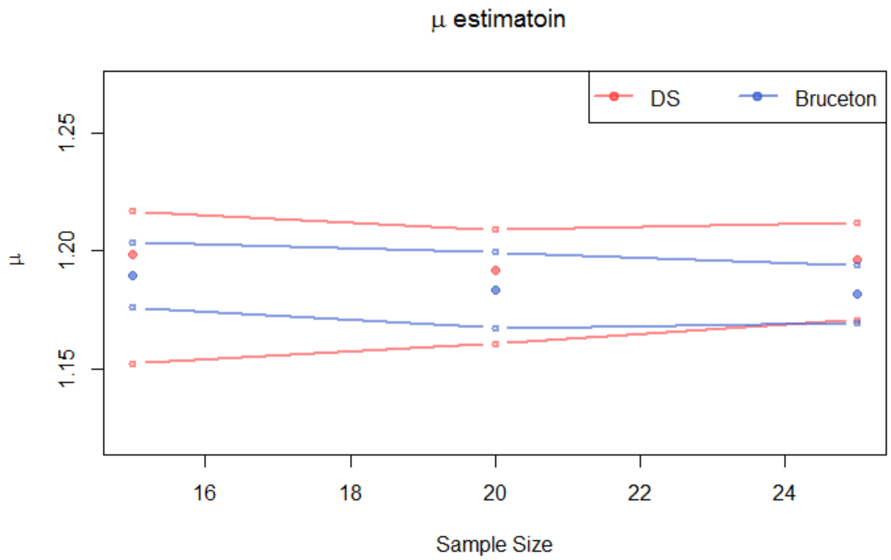

Fig. 5. The point estimates (dots) and confidence intervals (lines) for μ from the two experimental plans.

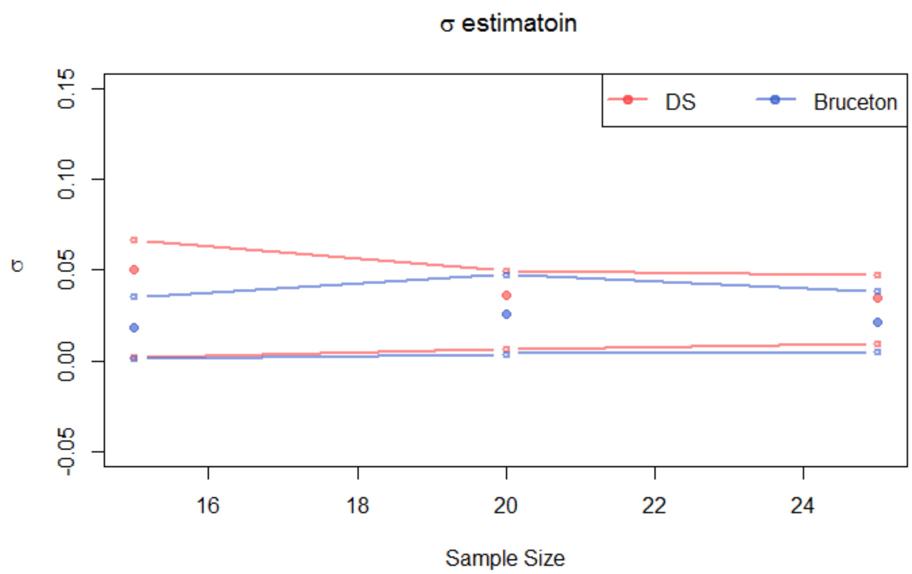

Fig. 6. The point estimates (dots) and confidence intervals (lines) for σ from the two experimental plans.

The two approaches can be compared by looking at the results for the full set of D-S trials. The Bayesian credible intervals are clearly wider than the standard confidence intervals. This is a surprising result, as the Bayesian analysis adds "prior information" that was not present in the standard analysis. In principle, this should lead to tighter inferential statements. Further, if the asymptotic approximations that justify the standard analysis are valid, then we are guaranteed to get tighter statements with the Bayesian analysis. So the fact that the Bayesian results are more spread out is a strong suggestion that the asymptotics are not a good guide to inference with such small experiments. The standard analysis for σ assumes that its direct estimator has an approximately normal distribution. If that were true, we should have found that the posterior distribution for σ is approximately normal. In fact, we found that it is much closer to normal when shown on the log scale. This is a further indication that the standard inferential analysis for σ is not valid, but could be improved if it were conducted on the log scale.

II. SUMMARY AND CONCLUSIONS

In this study we compared two procedures for carrying out sensitivity tests for detonators of explosive devices: The Bruceton method and the Dror-Steinberg method. The test case is providing very useful raw material that help us both to study and to improve the methods. As expected from the literature, we found that the Bruceton protocol was slightly better for estimating the mean of the distribution, whereas D-S was better for the standard deviation. The advantage of D-S for the standard deviation was not as clear cut as we expected. One important reason for this is that the assessed precision of the SD is proportional to the estimated value. The Bruceton results led to a lower estimated SD and this translated directly into improved assessed precision. The lower estimate from the Bruceton trials was purely a matter of chance, as the detonators were assigned randomly to be used in one of the Bruceton trials or one of the DS trials. When precision is assessed using a common estimator of the SD (from combining all the trials), the DS protocol is found to have about 25% better precision than Bruceton (smaller SD), roughly equivalent to reducing the number of trials by about 40%.

In the next stage of the research we will run large-scale simulation, based on the results of the test case, to compare the design protocols. The simulation will also include the new 3pod method of design (Ref. 11). In addition to the design protocol, we will also study the quality of the statistical inferences obtained from them. In particular, we want to compare the properties of the interval estimates from the Bayesian analysis to those from the "classical" analysis, in which Taylor series expansions are used to compute error intervals for estimated features of the sensitivity distribution.

ACKNOWLEDGMENTS

The authors would like to thank Idan Ptashnik and his team in Rafael for their kind support in carrying out the sensitivity experiments.

This work was supported by an RSF (Rafael Science Fund) grant. This fund was established to enhance scientific research collaboration between industry and academic institutions.

REFERENCES

1. ASTM standard No. 2463, “Standard Test Method for Drop Impact Resistance of Blow-Molded Thermoplastic Containers” (2010).
2. W. J. DIXON and A. M. MOOD, “A method for obtaining and analyzing sensitivity data”, Journal of the American Statistical Association, **43**, 109–126 (1948).
3. H.A. DROR and D. M. STEINBERG, “Sequential experimental designs for generalized linear models”, Journal of the American Statistical Association, **103**, 288–298 (2008).
4. V. R. JOSEPH, “Efficient Robbins-Monro procedure for binary data”, Biometrika, **91**, 461–470 (2004).
5. MIL-STD-1751A, “Safety and Performance Tests for the Qualification of Explosives” (2001).
6. MIL-STD-331C, “Test Methods Standard for Fuze and Fuze Components” (2001).
7. B. T. NEYER, “D-Optimality-Based Sensitivity Test”, Technometrics, **36**, 61–70 (1994)
8. H. ROBBINS and S. MONRO, “A stochastic approximation method”, Annals of Mathematical Statistics, **29**, 373–405 (1951)
9. D. M. STEINBERG, H. DROR and L. VILLA, “Discussion of ``Three-phase optimal design of sensitivity experiments" by Wu and Tian”, Journal of Statistical Planning and Inference, **149**, 26–29 (2014).
10. C. F. J. WU, “Efficient sequential designs with binary data”, Journal of the American Statistical Association, **80**, 974–984 (1985).
11. C. F. J. WU and Y. TIAN, “Three-phase optimal design of sensitivity experiments”, Journal of Statistical Planning and Inference, **149**, 1–15 (2014).