ARTICLE IN PRESS

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Vision Research xxx (2007) xxx–xxx

Vision Research

www.elsevier.com/locate/visres

Commentary

Singularities explained: Response to Klein

Mikhail Katkov [a,b], Misha Tsodyks [a], Dov Sagi [a,*]

[a] *Department of Neurobiology/Brain Research, The Weizmann Institute of Science, Rehovot 76100, Israel*
[b] *Department of Psychology, New York University, New York, NY 10003, USA*

Received 17 October 2006; received in revised form 30 October 2006

Abstract

Klein [Klein, A. S. (2006). Separating transducer nonlinearities and multiplicative noise in contrast discrimination. *Vision Research, 46*, 4279–4293] questions the existence of intrinsic singularities in two-alternative force-choice (2AFC) Signal Detection Theory (SDT) models, suggesting that the singularities found in Katkov et al. [Katkov, M., Tsodyks, M., & Sagi, D. (2006a). Singularities in the inverse modeling of 2AFC contrast discrimination data. *Vision Research, 46*, 259–266; Katkov, M., Tsodyks, M., & Sagi, D. (2006b). Analysis of two-alternative force-choice Signal Detection Theory model. *Journal of Mathematical Psychology, 50*, 411–420] are due to discarding higher order terms in the Taylor expansion of d' and/or limited to steep psychometric functions. Here we provide some simple intuitive examples that illustrate the results described in Katkov et al. (2006a, 2006b). We show, for the constant noise model, that singularities exist when exact values of d' are computed and that the singularities are not limited to steep psychometric functions. In these cases the disambiguation of the different models requires millions of trials.

Keywords: SDT; 2AFC; Contrast response function; Signal to noise; d'

1. 2AFC SDT models

We consider a standard 2AFC experiment where on each trial an observer is presented with two stimuli that differ on a single dimension (e.g. contrast, intensity) and reports which one of the two has a higher value. The analysis provided in Katkov, Tsodyks, and Sagi (2006a, 2006b) is aimed at finding the range of models that can describe the same experimental results given a finite number of 2AFC trials. Here we provide a brief description of the problem and some illustrative examples to further clarify the implications for psychophysics. Following Signal Detection Theory (SDT), we assume that each stimulus evokes a scalar internal response that varies across trials so that the observer's performance depends on the distributions of these internal responses. In addition, we assume that (1) the distribution of internal responses is Gaussian, and (2) the decision is made by comparing internal responses corresponding to the two stimuli. The percentage of correct discrimination responses under these assumptions is given by Green and Swets (1966)

$$P_{s_1,s_2} = \Phi\left(\frac{R_{s_2} - R_{s_1}}{\sqrt{\sigma_{s_1}^2 + \sigma_{s_2}^2}}\right), \quad (1)$$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt,$$

where $\Phi(x)$ is a standard normal cumulative distribution function, R_{s_1}, R_{s_2} are mean internal responses, σ_{s_1}, σ_{s_2} are trial-by-trial standard deviations of the internal responses, and P_{s_1,s_2} is the probability of reporting that the stimulus s_2 has higher contrast than stimulus s_1. The values of P_{s_1,s_2} are estimated from the experiment, whereas those of R and σ are the parameters of the model. Eq. (1) has four unknown variables for one pair of stimulus contrasts, leading to an ambiguous solution for the values of R and σ. Two unknowns can be fixed, such as $R_{s_1} = 0$, $\sigma_{s_1} = 1$, since P_{s_1,s_2}

* Corresponding author. Fax: +972 8 934 4131.
E-mail address: Dov.Sagi@Weizmann.ac.il (D. Sagi).

0042-6989/$ - see front matter
doi:10.1016/j.visres.2006.10.030

Please cite this article in press as: Katkov, M. et al., Singularities explained: Response to Klein, *Vision Research* (2007), doi:10.1016/j.visres.2006.10.030

does not depend on a particular choice of the origin and on the scale of the internal responses. However, there are still two unknowns left to be solved with one equation and an ambiguous solution is expected. This situation reflects a well-known problem in psychophysics, according to which, performance is a function of the signal-to-noise ratio and neither the signal nor the noise can be separately estimated. Consequently, there is a trade-off between the choice of contrast response function and noise amplitudes, which Klein (2006) calls singularity. For any choice of noise amplitudes, one can always find a contrast response function that will explain the same TvC (threshold vs. contrast) curve *exactly*—i.e. many models can explain the same data. The singularities found in Katkov et al. (2006a, 2006b) are of a different nature—many models can explain the data within an experimental error in conditions where the number of measurements (and equations) exceeds the number of unknown parameters (R and σ).

It is clear that if one wants to obtain an unambiguous solution for Eq. (1), one has to design an experiment where the number of measurements is at least greater than the number of parameters. The maximum number of independent measurements in discrimination experiments with a fixed set of stimuli can be achieved by performing all pairwise comparisons. For example, for five stimuli there are eight independent parameters, and ten independent measurements. In this case the presence of singularity is not obvious. There should be a good reason why several different models with eight parameters can equally describe the same ten measurements. In Katkov et al. (2006b) we provide such an analysis and describe under what conditions an over-completed (the number of measurements is greater than the number of parameters) system of equations such as Eq. (1) has many solutions, taking into account *experimental errors* due to finite sampling. In other words, the main question here is: what is the range of models that can fit the same experimental data.

For example, in Table 1 we present three models—one in which the noise amplitude does not depend on contrast and two models where the noise is an increasing function of contrast. We do not specify the values of $c_1, c_2, \ldots, c_n$ explicitly, since the following discussion does not depend on the particular choice of these contrast levels. Therefore, by adequately placing c_i, it is possible to describe any shape of contrast response function. Consequently, we discuss all possible models with arbitrary contrast response functions and constant noise. A more detailed analysis in Katkov et al. (2006b) shows that only the relationship between the contrast response function and the noise function is relevant. The specific values of the model parameters were computed using Matlab programs, which are attached as Supplementary material. Technically, we were looking for a model with parameters that best fit the performances calculated for the constant noise model shown in Table 1, with an additional constraint defining the ratio between the first and the last noise amplitudes. The ratios used were 10^4 and 2 for the first and the second alternative models, respectively, thus forcing them to have non-constant noise. The power model in Table 1 can be seen as a limit of an infinite ratio with additional constraint on the form of noise amplitudes (power function of contrast response).

Table 2 presents a comparison between exact expected probabilities [as computed by Eq. (1)] of correct responses for stimulus pairs that are represented by the rows (s_2) and the columns (s_1), for the three models presented in Table 1. Assume that we test a constant noise null hypothesis with these probabilities estimated using 100,000 trials per pair (totaling 1 million trials), the Pearson X^2 statistics can be computed as:

$$X^2 = \sum_{(i,k)} \frac{(P_{i,k} - \hat{P}_{i,k})^2}{P_{i,k}(1 - P_{i,k})} N, \quad (2)$$

where $P_{i,k}$ is the expected probability of the tested model—the top number in each cell of Table 2, $P_{i,k} - \hat{P}_{i,k}$ is the difference between this expected probability and one expected from another model (one of the two lower numbers in the corresponding cell of Table 2), and $N = 10^5$ is the number of trials used to estimate each probability. The resulting X^2 values are shown in Table 1. Even with 2 degrees of freedom the data produced by the alternative (multiplicative noise) models with one million trials are compatible with

Table 1
Example of a constant noise singular model

	c_1	c_2	c_3	c_4	c_5	X^2 ($N = 10^5$)
Constant noise model						
R	0	1	2	3	4	
σ	1	1	1	1	1	0
Alternative (multiplicative noise) model						
R	0	3978.86	10618.95	18847.43	28306.86	
σ	1	5626.30	7510.71	8884.14	10000.00	0.045
Power fit of noise amplitudes						
496.503 $R^{0.293}$	0	5630.35	7506.45	8880.42	10004.22	0.083
Another alternative model						
R	0	1.194	2.6883	4.4114	6.3252	
σ	1	1.3606	1.6167	1.8234	2	0.0005

Please cite this article in press as: Katkov, M. et al., Singularities explained: Response to Klein, *Vision Research* (2007), doi:10.1016/j.visres.2006.10.030

Table 2
Performances for constant noise and alternative models (see Table 1)

	c_1	c_2	c_3	c_4	c_5
c_1	0.5	0.23975 0.23972 0.23975	0.07865 0.07870 0.07866	0.01695 0.01694 0.01695	0.00234 0.00232 0.00233
c_2	0.76025 0.76028 0.76025	0.5	0.23975 0.23961 0.23974	0.07865 0.07869 0.07866	0.01695 0.01699 0.01695
c_3	0.92135 0.92130 0.92134	0.76025 0.76039 0.76026	0.5	0.23975 0.23969 0.23975	0.07865 0.07864 0.07865
c_4	0.98305 0.98306 0.98305	0.92135 0.92131 0.92134	0.76025 0.76031 0.76025	0.5	0.23975 0.23973 0.23976
c_5	0.99766 0.99768 0.99767	0.98305 0.98301 0.98305	0.92135 0.92136 0.92135	0.76025 0.76027 0.76024	0.5

The top number in each cell represents the performance for the constant noise model and the bottom numbers show performances for alternative models. It is obvious that it is impossible to distinguish these differences in an experiment of reasonable length. Note that $P_{c_i,c_k} + P_{c_k,c_i} = 1$.

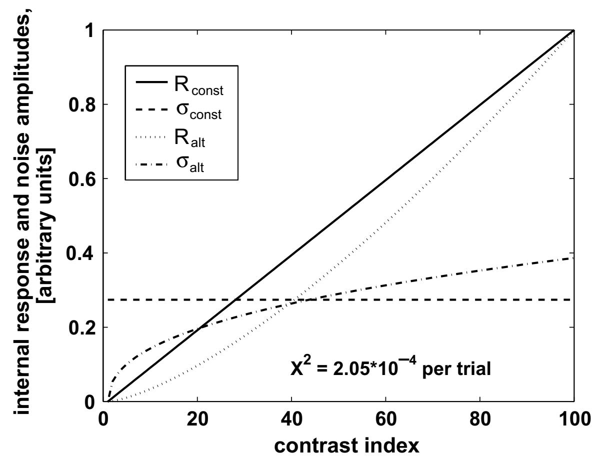

Fig. 1. An example of constant noise model and an alternative model using 100 stimuli and all pairwise comparisons. The noise amplitudes for the constant noise model was chosen to obtain maximum X^2 value for selected experimental design (SNR = 3.65, see Fig. 2 and the corresponding text). Here, for the presentation purpose, we rescaled the internal responses and noise amplitudes, setting $R_{\max} = 1$ in each model separately. Rescaling of all noise amplitudes (σ_i) and internal responses (R_i) by the same factor does not change the resulting performances [see Eq. (1)].

the constant noise model (the null hypothesis cannot be rejected). Here we used the X^2 metric, as was suggested in Klein (2006). We present here only two alternative models, but any model presented in the right column of Fig. 3 in Katkov et al. (2006b) has the same property, with power fit being the worst case, though X^2 is small even for noise amplitudes represented as a power fit of alternative models. Therefore, we show here an example of intrinsic singularity of the 2AFC SDT model: very different models produce almost indistinguishable results.

To ensure that this result is not due to the small number of contrast levels used, we carried out simulations using 100 contrast levels with results shown in Fig. 1 for one of the alternative models shown in Table 1 (max-to-min noise ratio of 10^4). X^2 in this case is 2.05×10^{-4}, requiring at least 10^4 trials per pair (totaling more than 10^7 trials) to disambiguate the models. Furthermore, the largest term in Eq. 2 is of order of 10^{-6}, therefore, even if it is possible to use only the more informative pairs, as was practically done by Kontsevich, Chen, and Tyler, 2002, million trials are required to disambiguate these models. Since 100 contrast levels are certainly sufficient to cover the contrast range used in experiments, modeling the data becomes impractical with any experimental design, unless some additional constraints are imposed on the tested model. Such constraints are frequently imposed on the models when assuming a specific functional form relating σ and R (Georgeson & Meese, 2006; Klein, 2006; Kontsevich et al., 2002).

Klein (2006) suggests that steep psychometric functions may produce singularities. In our analysis, the shape of the contrast response function is irrelevant since we do not specify the contrast levels used (c_i). However, problems may arise if the contrasts are not properly chosen so that their spacing is too small or too large relative to the slope of the corresponding psychometric function. In the constant noise model it is possible to define "steepness" in terms of the ratio between the span of contrast response function and the noise amplitude—maximum signal to noise ratio, $\mathrm{SNR} = \frac{R_{\max}-R_{\min}}{\sigma}$, with higher SNRs corresponding to steeper psychometric functions. Fig. 2 shows the dependency of X^2 on SNR. In this computation, we used 15 stimuli (contrast levels) and all pairwise comparisons. We assumed a constant noise model and looked for the best fitting alternative model with a max-to-min noise ratio of 10^4. We found that X^2 is maximal with SNR = 3.4

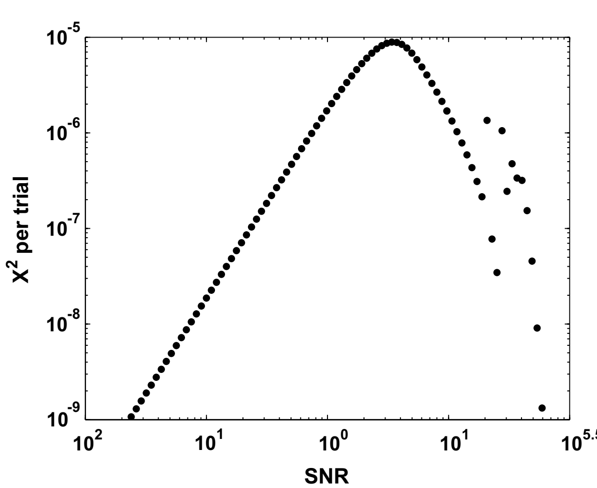

Fig. 2. The effect of SNR on X^2 of an alternative model. 15 stimuli and all pairwise comparisons were used in these computations. The alternative model has a max-to-min noise ratio of 10^4. The curve is maximal at SNR = 3.40. The fluctuations seen on the right can be attributed to numerical errors, possibly due to the large number of local minima with large SNR.

Please cite this article in press as: Katkov, M. et al., Singularities explained: Response to Klein, *Vision Research* (2007), doi:10.1016/j.visres.2006.10.030

having a value of 10^{-5} per trial. Thus, even in the worst case, rejection of the alternative model (with "multiplicative" noise) requires 10^5 trials per measurement and thus, having 105 pairwise comparisons, impractical.

2. Equations analyzed

The results presented here in Tables 1 and 2 and in our previous reports (Fig. 3 in Katkov et al., 2006b, Figs. 4 and 5 in Katkov et al., 2006a) are from simulations that compute and compare exact values of performances. Thus, contrary to Klein's claim, our results do not rely on the Taylor expansion of d', and the d' values in the equations are solved exactly, without using linear approximation. We do use the Taylor expansion when analyzing the effects on d' when there are small changes in the parameters around the solution, i.e. linearizing differences in d' in a small neighborhood of the "true" model, while computing d' values exactly. In the analysis of d' differences three singularities (in addition to the continuous symmetry) are revealed in the first-order term of the Taylor expansion. For example, suppose we carry out a 2AFC experiment having two stimuli—s_1 and s_2. Furthermore, suppose the true parameters of the internal response distributions are R_{s_1} and σ_{s_1}, and R_{s_2} and σ_{s_2}, respectively. In a long experiment, we expect to measure a performance P corresponding to

$$z = \frac{R_{s_2} - R_{s_1}}{\sqrt{\sigma_{s_1}^2 + \sigma_{s_2}^2}}.$$

Let us now examine the alternative model, which differs from the true one by the singular eigenvector corresponding to the constant noise model (Katkov et al., 2006b, Table 1, second line). Namely, we set $R'_{s_1} = R_{s_1} + \gamma R_{s_1}^2$ and $\sigma'_{s_1} = \sigma_{s_1}(1 + \gamma R_{s_1}\sigma_{s_1})$ for the stimulus s_1 and $R'_{s_2} = R_{s_2} + \gamma R_{s_2}^2$ and $\sigma'_{s_2} = \sigma_{s_2}(1 + \gamma R_{s_2}\sigma_{s_2})$ for the stimulus s_2. Now we can compute a z_{alt} score for the alternative model:

$$z_{\text{alt}} = \frac{R'_{s_2} - R'_{s_1}}{\sqrt{(\sigma'_{s_1})^2 + (\sigma'_{s_2})^2}},$$

$$z_{\text{alt}} = \frac{R_{s_2} + \gamma R_{s_2}^2 - R_{s_1} - \gamma R_{s_1}^2}{\sqrt{(\sigma_{s_1}(1 + \gamma R_{s_1}\sigma_{s_1}))^2 + (\sigma_{s_2}(1 + \gamma R_{s_2}\sigma_{s_2}))^2}},$$

with a corresponding performance P'. Applying Taylor expansion with respect to γ leads to

$$z_{\text{alt}} = z + \frac{(R_{s_2} - R_{s_1})^2\left(\sigma_{s_1}^2 - \sigma_{s_2}^2\right)}{\left(\sigma_{s_1}^2 + \sigma_{s_2}^2\right)^{3/2}}\gamma + O(\gamma^2).$$

It can be seen that the first-order term remains unless $\sigma_{s_1} = \sigma_{s_2}$. In this case the difference $z_{\text{alt}} - z$ goes at least as γ^2. What does this mean in terms of hypothesis testing? Let us now examine the χ^2 goodness of fit. Suppose we performed N trials, but performances are obtained from an alternative model (or can be described exactly by an alternative model). Then, χ^2 can be written in the following way [Eq. (2) for a single pair]:

$$\chi^2 = \frac{(P - P')^2}{P(1 - P)}N.$$

Since [see Eq. (1)]

$$P = \Phi(z) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} e^{-\frac{t^2}{2}}\,dt$$

is an analytic function, then a Taylor expansion of χ^2 is

$$\chi^2 = \frac{\left(\frac{dP}{dz}\big|_z dz + o(dz)\right)^2}{P(1 - P)}N.$$

Ignoring insignificant high-order terms in γ:

$$\chi^2 \approx \left(\text{const}(\sigma_{s_1}^2 - \sigma_{s_2}^2)^2\gamma^2 + \text{const}\ \gamma^4\right)N.$$

According to statistics, when χ^2 reaches some threshold, which depends on the number of degrees of freedom, we would reject the null hypothesis (data described by an alternative model can be explained by an original model in our case). Thus, in fixing the values of χ^2 at this level, we can see the dependency of limiting γ on the number of trials in the constant noise true model case ($\sigma_{s_1} = \sigma_{s_2}$), and in the opposite case; $\gamma \propto 1/\sqrt{N}$ when $\sigma_{s_1} \neq \sigma_{s_2}$ otherwise $\gamma \propto 1/\sqrt[4]{N}$. For a large number of trials, this means that the critical value of γ is smaller than 1, and thus, it is much larger for the constant noise case. This means that if, for example, in a nonsingular case it is necessary to measure 100 trials to obtain a reliable estimation of the parameters, then in a singular case it would require 10 000 trials to obtain the same range of solutions.

3. Conclusion

Here we presented a discussion of singularities in the SDT model that describe 2AFC performance, clarifying some issues pointed out in Klein (2006). We described some simple examples that illustrate the concept of singularity used by us and the resulting implications. In Katkov et al. (2006b) we found another two classes of models that exhibit singular behavior in addition to the model associated with constant noise. Note that not all 2AFC SDT models are singular, and thus the corresponding data can be modeled with a small range of parameters (for example, as in the left column in Fig. 3 in Katkov et al. (2006b)). Therefore, the question about the usefulness of 2AFC procedures for estimating the contrast response function and noise amplitudes depends on how far the "true" unknown model is from one of the singular models. This is an empirical question but anyone using 2AFC methods should be aware of this. Our experimental results have properties that are typical to that of the constant noise singular model (Katkov et al., 2006a).

Acknowledgments

We thank Nitzan Censor for helpful comments on an earlier version of the manuscript. This research was supported by the Basic Research Foundation, administrated by the Israel Academy of Science.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.visres.2006.10.030.

References

Georgeson, M. A., & Meese, T. S. (2006). Fixed or variable noise in contrast discrimination? The jury's still out…. *Vision Research, 46*, 4294–4303.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

Katkov, M., Tsodyks, M., & Sagi, D. (2006a). Singularities in the inverse modeling of 2AFC contrast discrimination data. *Vision Research, 46*, 259–266.

Katkov, M., Tsodyks, M., & Sagi, D. (2006b). Analysis of two-alternative force-choice Signal Detection Theory model. *Journal of Mathematical Psychology, 50*, 411–420.

Klein, A. S. (2006). Separating transducer nonlinearities and multiplicative noise in contrast discrimination. *Vision Research, 46*, 4279–4293.

Kontsevich, L. L., Chen, C. C., & Tyler, C. W. (2002). Separating the effects of response nonlinearity and internal noise psychophysically. *Vision Research, 42*, 1771–1784.