Estimation of Separable Representations in Psychophysical Experiments

Michele Bernasconi (mbernasconi@eco.uninsubria.it)
Christine Choirat (cchoirat@eco.uninsubria.it)
Raffaello Seri (rseri@eco.uninsubria.it)
Universit degli Studi dell'Insubria; Dipartimento di Economia; Via Ravasi 2
Varese, 21100 (VA) Italy

Abstract

Studying how individuals compare two given quantitative stimuli, say d_1 and d_2, is a fundamental problem. One very common way to address it is through *ratio magnitude estimation*, that is to ask individuals not to give values to d_1 and d_2 but rather to give their estimates of the ratio $p = d_1/d_2$. Several psychophysical theories (the most known being Stevens' power-law) claim that this ratio cannot be known directly and that there are cognitive distortions on the apprehension of the different quantities. These theories result in the so-called *separable representations* (which include Stevens' model). In this paper, we propose a general statistical framework that allows for testing in a rigorous way whether the separable representation theory is grounded or not. We conclude in favour of it, but strongly reject Stevens' model. As a byproduct, we provide estimates of the psychometric functions of interest.

Keywords: Psychophysical experiments; Steven's model; separable representation.

The Theoretical Framework

We propose to a subject two stimuli d_1 and d_2 out of a set of stimuli D and we ask him to state in what proportion p they are with respect to each other. According to Stevens' (1946) psychophysical framework, we can see this as a problem of *magnitude estimation*.[1] We say that a *separable representation* holds if a *psychophysical function* ψ and a *subjective weighting function* W exist such that the ratio p is in the following relation with d_1 and d_2:

$$\frac{\psi(d_1)}{\psi(d_2)} = W(p). \qquad (1)$$

Equation (1) corresponds to Narens' (1996) model apart from the fact that he does not necessarily suppose that p is a number (but a numeral). Usually the equality $W(1) = 1$ is supposed to hold, that is the individuals are able to correctly estimate ratios of equal stimuli; moreover, ψ is defined up to a multiplicative constant so that we can suppose that $\psi(1) = 1$. At last, it is easily seen that both ψ and W are defined up to a power transformation (i.e., if ψ and W are functions for which a separable representation holds, then so are also ψ^r and W^r for any real r); this will lead us to impose a restriction in our empirical investigation.

Several relations of this kind have been proposed in the literature. The original Stevens' model reduces simply to the case in which ψ and W are power functions (see e.g. Luce, 2001a, Section "Relations among the properties"). Remark that, as explained above, this case is undistinguishable from the case in which W is the identity function so that Stevens' model can be recovered as $\psi(d) = d^\kappa$ and $W(p) = p$ (or alternatively as $\psi(d) = d$ and $W(p) = p^{\frac{1}{\kappa}}$; this second formulation will be used for our empirical investigation because of its computational advantages). When criticizing what he calls *Stevens's Assumptions* (Narens, 1996, p. 109), Narens (1996) states (pp. 110-111) that the fact that $W(p) = p$ seems to be "anything more than a coincidence".

Starting from the representation (1), Luce (2001b) proposes a function of W based on hypotheses similar to Prelec's (1998) ones:

$$W(p) = \begin{cases} \exp\left[-\rho\left(-\ln p\right)^{\eta}\right] & p \in]0,1] \\ \exp\left[\rho'\left(\ln p\right)^{\eta'}\right] & p \in]1,\infty[\end{cases} \qquad (2)$$

Other examples of functions W are derived in Prelec (1998) and discussed in Luce (2001b). Luce (2002, pp. 526-528) proposes some other forms for ψ and W.

The objective of this paper is to test whether Stevens' power law model is appropriate or wether the separable representation holds in a simple ratio magnitude estimation experiment. Several papers have dealt with tests of behavioral properties (Ellermeier and Faulhammer, 2000, Zimmer, Luce and Ellermeier, 2001, Steingrimsson and Luce, 2003a, 2003b, based on the framework of Luce, 2002, 2004) or tests of particular functional forms (Hollands and Dyre, 2002). However, it seems to us that this is the first time in the literature that a formal and direct test of the formula (1) is conducted.

In the next Section we describe an experiment of ratio magnitude estimation involving 69 subjects. Then, we will show how the functions $W(\cdot)$ and $\psi(\cdot)$ can be estimated nonparametrically using polynomial regression and how the representation (1) can be tested. At last, the results are described.

[1] Sometimes *magnitude estimation* denotes the case in which a stimulus d_1 and a ratio p are known and the subject has to choose another stimulus d_2 such that the ratio of the two stimuli is given by p, even if this should be more correctly called *magnitude production*.

The Experiment

The experiment was conducted on 69 Italian students (from 1st to 4th university year in Economics). The language of the experiment was Italian. The experiment was made on personal computers with 3 rounds of 23 subjects each. Each round lasted 1h30min and subjects had the possibility of using pocket calculators and sheets of papers if they wanted to. The first 30 minutes were dedicated to reading together the written instructions that each subject received when entering the computer room and answering subjects' questions.

Subjects were asked to rate on a 1-9 integer scale the values of ratios of known probabilities, ratios of distances of pairs of Italian cities from a reference point and ratios of rainfalls in pairs of European cities; the ratios of the real stimuli were choosen to lie within the stated range 1-9. We chose not to randomize the order in which the three sub-experiments were to be performed. We namely decided to assign deterministically each of the 6 possible orders so that we knew from the start that none of the subjects' direct neighbors would work on the same sub-experiment at the same time.

It was decided to use a monetary reward as an incentive for subjects to perform the experiment as well as possible. Subjects were explained at the beginning of the experiment how the payment would be calculated. Namely, payment was proportional to good performance in the experiment.

Since in the following we will consider only the sub-experiment concerning distances, we will discuss this one in greater detail. We presented to the subjects 10 pairs of Italian cities and we asked them to estimate the ratio of their distances with respect to Milan: the 10 pairs are given by all the possible combinations out of the five cities Turin, Venice, Rome, Naples and Palermo. The range of the stimuli goes from 124 to 885 km and the range of the real distance ratios from 2 to 7.137.

Estimation and Inference

Log-log Transformation

The main problem with representation (1) is that it is not directly amenable to statistical estimation. Therefore, in order to get a simpler formulation, we write it as follows:

$$\ln W\left[\exp\left(\ln p\right)\right] = \ln\psi\left[\exp\left(\ln d_1\right)\right] - \ln\psi\left[\exp\left(\ln d_2\right)\right].$$

This is equivalent to a log-log transformation (see Luce, 2002, p. 526). We define:

$$\begin{aligned} \pi &= \ln p \\ \delta_i &= \ln d_i \\ \ln W\left[\exp\left(\cdot\right)\right] &= w\left(\cdot\right) \\ \ln\psi\left[\exp\left(\cdot\right)\right] &= \Psi\left(\cdot\right); \end{aligned}$$

the constraints $W(1) = 1$ and $\psi(1) = 1$ become respectively $w(0) = 0$ and $\Psi(0) = 0$. This means that we can write the separable representation (1) as:

$$\begin{aligned} w\left(\pi\right) &= \Psi\left(\delta_1\right) - \Psi\left(\delta_2\right) \\ \pi &= w^{-1}\left[\Psi\left(\delta_1\right) - \Psi\left(\delta_2\right)\right]. \end{aligned}$$

It is sensible to suppose that, at least because of discretization errors, the relation holds approximately, that is we can write:

$$\pi = w^{-1}\left[\Psi\left(\delta_1\right) - \Psi\left(\delta_2\right)\right] + \varepsilon, \qquad (3)$$

where ε is an error term. Remark that this is a special case of a much more general specification, in which the dependence of π on δ_1 and δ_2 is left unrestricted:

$$\pi = f\left(\delta_1, \delta_2\right) + \varepsilon. \qquad (4)$$

We will call SEP (for *separable*) the model of equation (3) and UNR (for *unrestricted*) the model of equation (4).

Even though the model of Equation (3) is restricted, it is general enough to recover the main theoretical models that can be found in the literature. Indeed, Stevens' model is obtained simply by putting $\Psi(\delta) = \kappa\delta$ and $w(\pi) = \pi$ (or alternatively $\Psi(\delta) = \delta$ and $w(\pi) = \frac{\pi}{\kappa}$). Model of Equation (2) proposed in Luce (2001b) and Prelec (1998) becomes, in our logarithmic formulation:

$$w^{-1}(x) = \begin{cases} -\left(-\frac{x}{\rho}\right)^{\frac{1}{\eta}} & p \in \left]0,1\right] \\ \left(\frac{x}{\rho'}\right)^{\frac{1}{\eta'}} & p \in \left]1,\infty\right[\end{cases}$$

When ρ or ρ' is equal to 1, this reduces to a power law. The statistical test of these functional forms is left to further work.

Structure of the Data

Now we describe the structure of the data provided by the experiment. For any individual $h = 1, \ldots, H$ with $H = 69$, we observe a vector of log-ratios $\boldsymbol{\pi}^h = \left(\pi_1^h, \ldots, \pi_C^h\right)^T$ where $C = 10$ is the number of possible pairwise comparisons. For any stated log-ratio π_i^h (which corresponds to the $i-$th comparison for the $h-$th individual), we know also the values of the lograrithms of stimuli, say $\delta_{i,1} = \log(d_{i,1})$ and $\delta_{i,2} = \log(d_{i,2})$; remark that the stimuli do not depend on the individual h. We suppose the existence of a relation of the form

$$\pi_i^h = f\left(\delta_{i,1}, \delta_{i,2}\right) + \varepsilon_i^h.$$

The case of interest is the one in which the function f takes the form (3), but we will also consider below a more general framework in which f is left unrestricted: this will allow us to test the restriction through statistical techniques. The residuals for the individual h are stacked in a vector $\boldsymbol{\varepsilon}^h = \left(\varepsilon_1^h, \ldots, \varepsilon_C^h\right)$: we suppose that the mean and the covariance matrix are respectively $\mathbb{E}\left(\boldsymbol{\varepsilon}^h\right) = \mathbf{0}$ and $\mathbb{V}\left(\boldsymbol{\varepsilon}^h\right) = \boldsymbol{\Sigma}$. The matrix $\boldsymbol{\Sigma}$ can take several alternative forms, but we will not pursue the topic here. Details are available from the authors upon request.

Statistical Theory

Our aim is to estimate functions $w(\cdot)$ and $\Psi(\cdot)$ using statistical methods. Different theories assume different forms for these functions. In a first time, we would like to compare the unrestricted model of Equation (4) and the restricted model of Equation (3) to statistically test whether there is enough empirical evidence that supports the restricted model. Should it be the case, we would like in a second time to compare the restricted model to its further particulars, namely the models of Stevens and the one of Luce and Prelec.

The key idea to perform a rigorous statistical analysis of the problem is to use nonparametric methods. We assume namely that the functions $f(\cdot)$, $w^{-1}(\cdot)$ and $\Psi(\cdot)$ are smooth enough to be approximated by a polynomial expansion. (Remark that this requirement is not very constraining since all the proposed theorical models so far suppose infinitely smooth functions.)

Let M, L and N be the respective orders of the polynomials used to approximate $f(\cdot)$, $w^{-1}(\cdot)$ and $\Psi(\cdot)$ respectively. Remark that $f(\cdot)$ is a function of two arguments δ_1 and δ_2; the polynomials that approximate it must therefore contains all the powers of δ_1 and δ_2 up to order M. Remark also that the function of interest is $w(\cdot)$, so that the approximation of $w^{-1}(\cdot)$ must be inverted to get the the approximation of $w(\cdot)$. It is easy to see that the assumption $W(1) = 1$ (which is always assumed) implies that these polynomials have no constant term.

Summing up, for the UNR model the parameters to be estimated are $\mathbf{\Sigma}$ and the $\frac{(M+1)(M+2)}{2}$ polynomial parameters, while for the SEP model the parameters are $\mathbf{\Sigma}$, and the $L+N-1$ polynomial parameters. Stevens' model is a restriction of SEP in which $N = L = 1$.

Supposing that the errors $\boldsymbol{\varepsilon}$ are distributed according to a multivariate normal distribution, we can write the loglikelihood and maximize it numerically in order to get the maximum likelihood estimates. The UNR model is quite simple to estimate, while the SEP model is very complex and requires a particular algorithm derived by the authors at this aim. More details are available upon request.

The main problem is to select the number of terms in the polynomial regressions (N, L and M). In order to do so, we use the BIC (Bayesian Information Criterion; see Schwarz, 1978), a method to penalize the likelihood taking into account the number of parameters. In the UNR model we just have to choose M, while in the SEP model we have to choose both N and L. Our strategy for the present case is to estimate several polynomial regression models for different values of M and (N, L), to choose the best UNR and SEP models according to the BIC and to test the restrictions imposed by the separable representation of equation (1), either through the BIC or through likelihood ratio tests. Once the model has been choosen, we can get the following estimates of the functions some algebra gives the nonparametric estimates of ψ and W.

The Results

Comparing the estimated models according to the BIC, it turns out that the best model of class UNR arises for $M = 2$, while the best model of class SEP for $(N, L) = (2, 3)$. The comparison of the best model of the UNR type and the best model of the SEP type shows that the SEP model appears to have a slight advantage over the UNR one. This suggests that a separable representation is supported by our experiment. As we remarked above, the Stevens' model arises when $(N, L) = (1, 1)$; since the values of N and L that appear as best are both different from 1, we cannot provide support for the Stevens' power-law model. This finding is confirmed by the BIC value (that is -0.6566 for the model in (3) and -0.7926 for Stevens' model) and by a likelihood ratio test (that takes the value 207.357 with 3 degrees of freedom, that is a $p-$value of $1.1 \cdot 10^{-44}$).

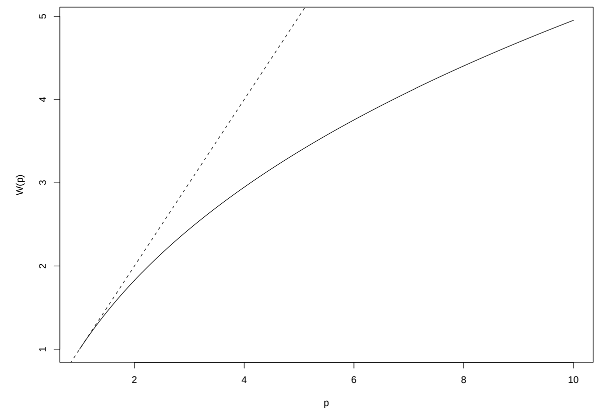

Figure 1: The subjective weighting function W.

Now we come to a deeper analysis of SEP, UNR and Stevens' models. Using the estimated parameters of the model with $(N, L) = (2, 3)$, it is possible to get an estimate of the functions ψ and W. Figures 1 and 2 show respectively the *subjective weighting function* and the *psychophysical function* for the SEP model (solid line) and Stevens' model (dashed line). The functions have been rescaled in order to have the same origin and the same slope at the origin, but the different behavior of the functions is evident. Figure 3 displays a three-dimensional representation of the expected value of p as stated by an individual (i.e. $W^{-1}\left(\frac{\psi(d_1)}{\psi(d_2)}\right)$ for d_1 and d_2 varying in the range of the data) and Figure 4 shows the same graph from above: the contour lines represent equal integer values of p. The deviation from the true value of the ratio (*i.e.* $\frac{d_1}{d_2}$) is difficult to see. Therefore, in Figure 5 we have displayed a three-dimensional representation of the deviation of the expected value of p from the true one (i.e. $W^{-1}\left(\frac{\psi(d_1)}{\psi(d_2)}\right) - \frac{d_1}{d_2}$); Figure 6 shows the same graph from above, with contour lines separated by 0.5 units.

Also Stevens' model has been estimated. The parameter value is $\phi_1 = 0.835$ and this corresponds to $\psi(d) = d^{0.835}$ and $p = \left(\frac{d_1}{d_2}\right)^{0.835}$ (this has to be compared with the estimated exponents in Stevens, 1961). Figures 7, 8, 9 and 10 show the expected value of p and its deviation from the true value: the graphs are quite far from the corresponding ones for the SEP model.

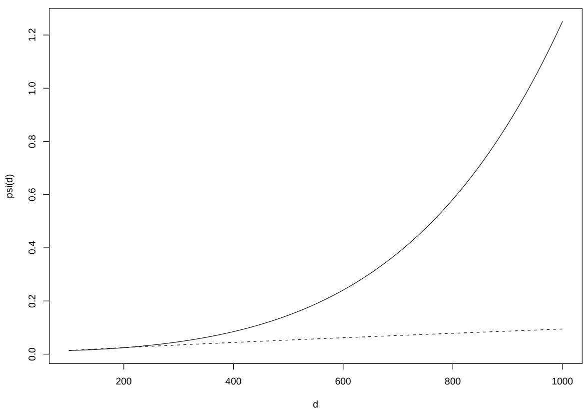

Figure 2: The psychophysical function ψ.

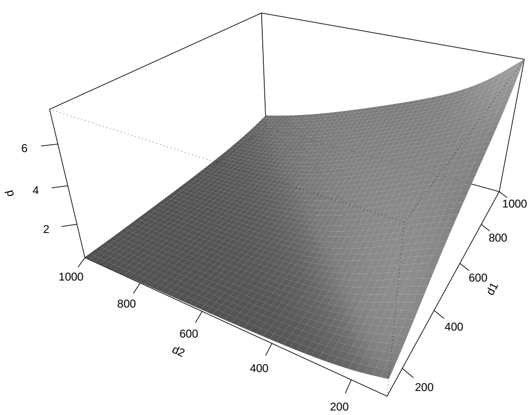

Figure 3: Expected value of p in SEP model.

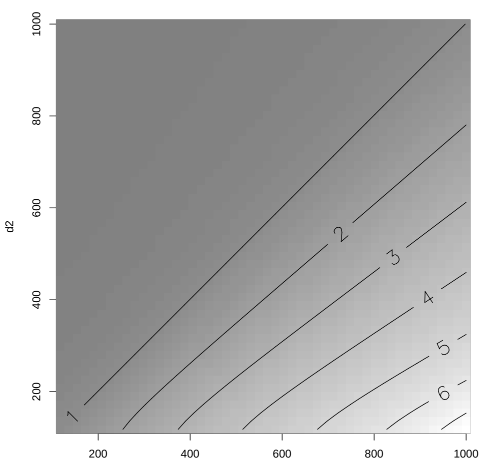

Figure 4: Expected value of p in SEP model with contour lines.

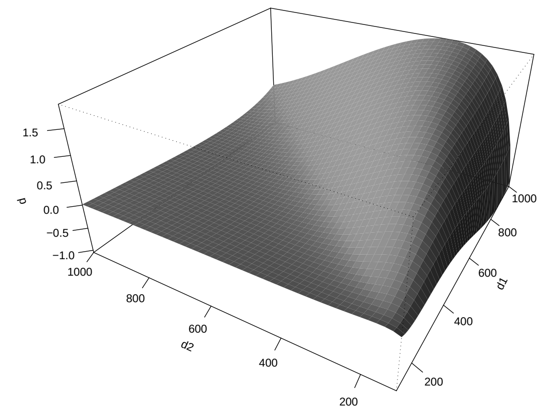

Figure 5: Expected minus true value of p in SEP model.

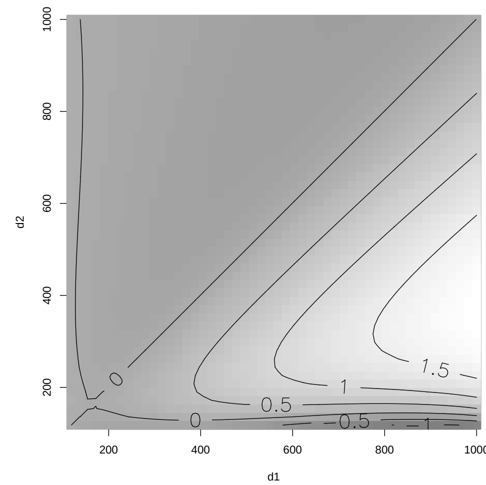

Figure 6: Expected minus true value of p in SEP model with contour lines.

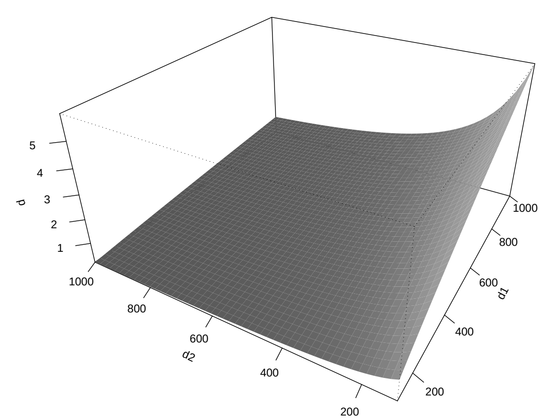

Figure 7: Expected value of p in Stevens' model.

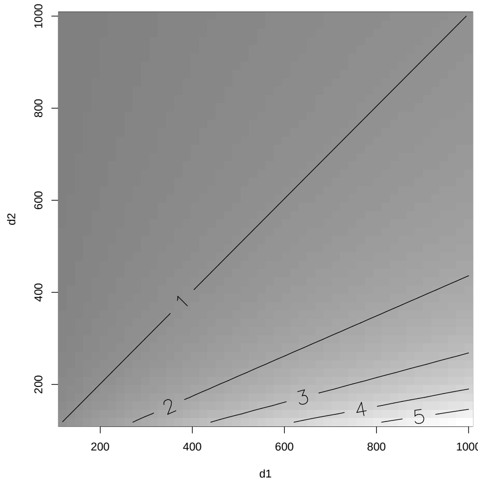

Figure 8: Expected value of p in Stevens' model with contour lines.

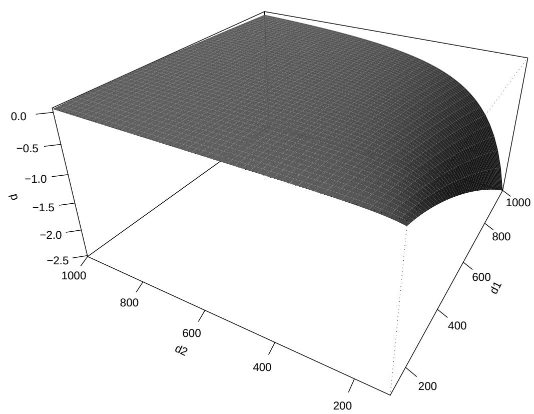

Figure 9: Expected minus true value of p in Stevens' model.

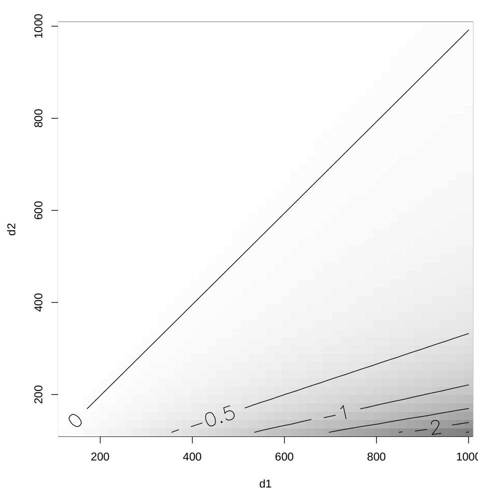

Figure 10: Expected minus true value of p in Stevens' model with contour lines.

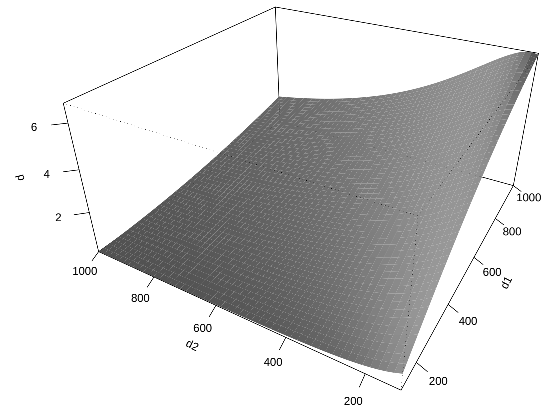

Figure 11: Expected value of p in UNR model.

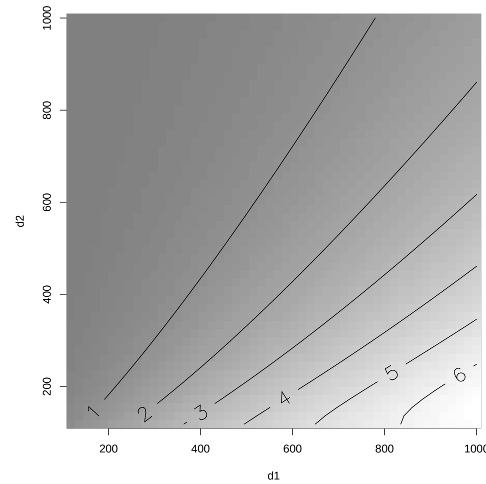

Figure 12: Expected value of p in UNR model with contour lines.

At last, we show the same graphs (see Figures 11, 12, 13 and 14) for the UNR model: in this case, it can be seen that the similarity with the corresponding SEP model is much stronger.

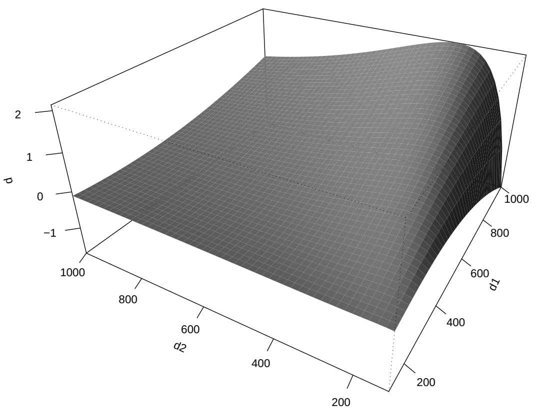

Figure 13: Expected minus true value of p in UNR model.

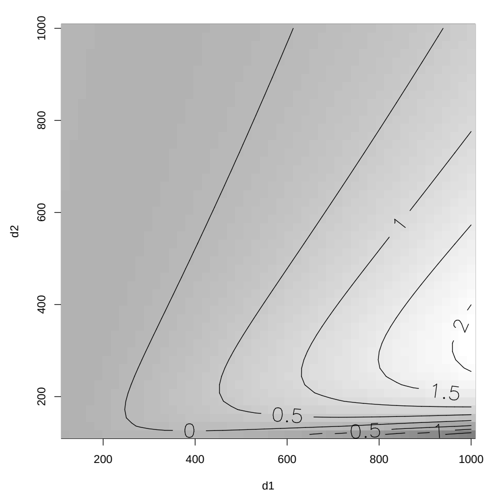

Figure 14: Expected minus true value of p in UNR model with contour lines.

References

Ellermeier, W. & Faulhammer, G. (2000). Empirical evaluation of axioms fundamental to Stevens' ratio-scaling approach. I. Loudness production. *Perception Psychophysics*, *62*, 1505–1511.

Hollands, J. G. & Dyre, B. P. (2000). Bias in proportion judgements: The cyclical power model. *Psychological Review*, *107*, 500–524.

Luce, R. D. (2001a). A way to blend Fechner and Stevens. In E. Sommerfeld & R. Kompass & T. Lachmann (Eds.), *Fechner Day 2001. Proceedings of the Seventeenth Annual Meeting of the International Society of Psychophysics*. Pabst Science Publishers.

Luce, R. D. (2001b). Reduction invariance and Prelec's weighting functions. *Journal of Mathematical Psychology*, *45*, 167–179.

Luce, R. D. (2002). A psychophysical theory of intensity proportions, joint presentations, and matches. *Psychological Review*, *109*, 520–532.

Luce, R. D. (2004). Symmetric and asymmetric matching of joint presentations. *Psychological Review*, *111*, 446–454.

Narens, L. (1996). A theory of ratio magnitude estimation. *Journal of Mathematical Psychology*, *40*, 109–129.

Prelec, D. (1998). The probability weighting function. *Econometrica*, *66*, 497–527.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*, 461–464.

Steingrimsson, R., & Luce, R. D. (2003a). *Evaluating a model of global psychophysical judgments: I. Behavioral properties of summations and productions*, *Working Paper*.

Steingrimsson, R., & Luce, R. D. (2003b). *Evaluating a model of global psychophysical judgments: II. Behavioral properties linking summations and productions*, *Working Paper*.

Stevens, S. S. (1946). On the theory of scales of measurement. *Science*, *103*, 677–680.

Stevens, S. S. (1961). To honor Fechner and repeal his law. *Science*, *133*, 80–86.

Zimmer, K., & Luce, R. D., & Ellermeier, W. (2001). Testing a new theory of psychophysical scaling: Temporal loudness integration. In E. Sommerfeld & R. Kompass & T. Lachmann (Eds.), *Fechner Day 2001. Proceedings of the Seventeenth Annual Meeting of the International Society of Psychophysics*. Pabst Science Publishers.