Reprinted from IFPress (2000) 3(2)2,3

Population Thresholds

Daniel M. Ennis

Background: A surprising number of applied sensory problems can be classified as threshold problems. For example, the time after application when an air freshener can no longer be detected is a time threshold. Similarly, the moment at which an off-note can be detected in a food product is a threshold. This information can be used to establish the shelf-life of the product. Methods for measuring thresholds have been developed for individuals, but in many applications there is more interest in population thresholds than individual thresholds. Special problems arise when extending the ideas of individual thresholds to population thresholds. This report overviews a new approach to measuring population thresholds.

The Individual Threshold: Intuitively, a threshold is the transition point between no detection and detection. Intrinsic to the idea of the threshold is the assumption that the transition point is constant. In practice, however, responses are affected by psychological and physiological inputs so that shifts in the transition point may occur. This problem can be overcome by treating the threshold as a statistical quantity. Empirical data demonstrate that the detectability of a stimulus does not jump from 0% to 100% at some particular value. In fact, the probability of detection increases gradually as the intensity of the stimulus increases. Consequently, an individual threshold is often defined as the stimulus intensity that elicits a response from an individual 50% of the time. Table 1 provides a summary of many of the techniques used to measure individual thresholds.

Table 1. Threshold measurement methods.

METHOD	AUTHOR(S)	YEAR	REF #
Spearman-Karber	Spearman	1908	6
Spearman-Karber	Karber	1931	7
Probit	Bliss	1934	1
Probit	Bliss	1935	5
Probit	Berkson	1944	2
Moving Average	Thompson	1947	8
Up-and Down	Dixon	1948	13
Moving Average	Finney	1950	12
Moving Average	Bennett	1952	9
Moving Average	Weil	1952	11
Logistic	Berkson	1955	4
Moving Average	Bennett	1963	10
Probit	Finney	1971	3
Probit and Logit	Prentice	1976	14
Spearman-Karber	Hamilton	1977	15
Kernel	Copas	1983	16

The Population Threshold: As noted earlier, there is interest in the threshold value for a population. This population may consist of users of continuous action air fresheners, and the stimulus may be a brand of air freshener. In this case the population threshold is the time after first usage when the air freshener is no longer detectable. A common definition of the population threshold[17] is the median value of the individual thresholds for the members of a population. To use this definition, one must know the individual thresholds of members of the population. To estimate a population threshold in this case, many individual thresholds must be estimated[18]. This is time consuming and impractical. Moreover, since differences among individuals can be the main source of variation when estimating thresholds, increasing the number of tests on a fixed set of individuals will not lead to a more sensitive estimate of the population threshold.

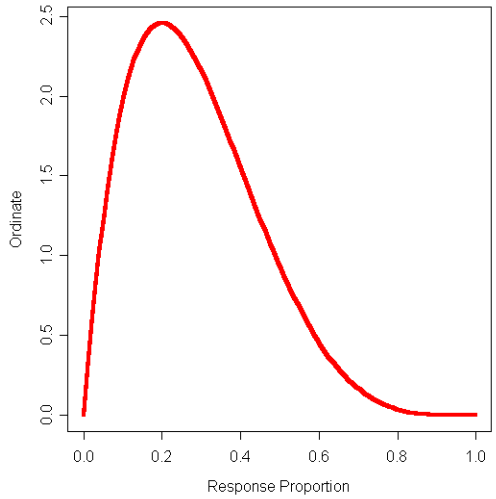
An alternate definition for the population threshold is the stimulus intensity level that produces a response in half of the population. According to this definition, individual thresholds need not be estimated to estimate the population threshold. The population threshold and its confidence intervals can be estimated directly from a sample drawn from the population.

Scenario: You want to know the moment when an off-note becomes detectable in a brand of snack food. From preliminary research you establish six time points for sampling. One hundred trained panelists evaluate five replicates at each of the six time points using a 'yes-no' task. From this data your goal is to estimate the population threshold time to off-note detection, allowing for differences in individual sensitivities.

A Method to Estimate the Population Threshold: The classical models for threshold estimation are based on the assumption that the response at each of the dosage steps follows a binomial distribution. This assumption is not always warranted. Population thresholds are measured over subjects and over trials so that pooled data can have both inter-subject and intra-subject variation. The variation in the pooled data is often larger than that in binomially distributed data. This phenomenon is known as *overdispersion*[19,20]. When overdispersion occurs, classical models, such as those involving probits or logits, are not adequate because the threshold error may be underestimated. One way of dealing with differences among subjects is to use the beta-binomial model. The beta-binomial model assumes that individuals differ in sensitivity and that individual response proportions follow a beta distribution. Figure 1 illustrates a beta distribution of these proportions at a particular stimulus dose. This model uses an overdispersion parameter to account for individual differences and allows the population threshold to be estimated. To use this method it is

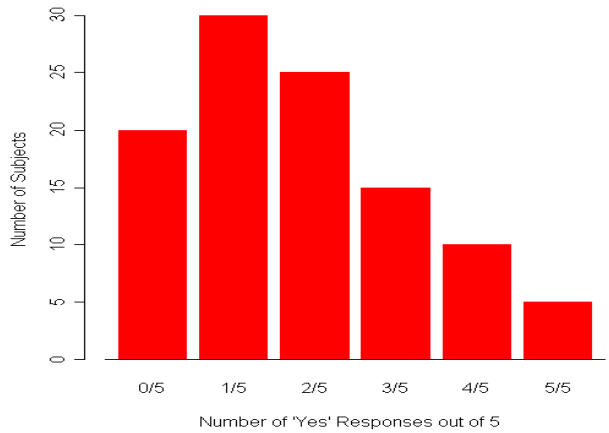
not necessary to estimate individual thresholds and individual information need not be extensive.

Figure 1. Beta distribution of response proportions across individuals at a particular dosage.



The Population Off-Note Time Threshold: At each time point in the aged snack food samples, you determine the number of 'yes' responses for each subject. For example, Figure 2 illustrates the number of 'yes' responses out of five for the subjects tested at the two day point. Using the data from the six time points and an equation from Bi and Ennis[21], you find the time threshold for off-note detection to be 6.5 days. You also estimate the variance to be 0.25. From these estimates, you can say with 95% confidence that the time to detection of the off-note is between 5.52 and 7.48 days. Note that it was not necessary to estimate individual thresholds; the estimate of the population threshold took individual differences into account. For a more detailed account of threshold measurement, refer to the paper by Bi and Ennis[21].

Figure 2. Off-note response proportions in the 'yes-no' task at the 2-day time point.



References:

1. Bliss, C.I. (1934). The method of probits. *Sci*., **79**, 38-39.
2. Berkson, J. (1944). Application of the logistic function to bioassay. *J. Am. Stati. Assoc.,* **39**, 357-365.
3. Finney, D. J. (1971). *Probit Analysis* 3rd ed., Cambridge University Press**.**
4. Berkson, J. (1955). Maximum likelihood and minimum c^2 estimates of the logistic function. *J. Am. Stati. Assoc.,* **50**, 206-209.
5. Bliss, C. I. (1935). The calculation of the dosage-mortality curve. *Ann. Appl. Biol*. **22**, 134-167.
6. Spearman, C. (1908). The method of right and wrong cases (constant stimuli) without Gauss's formulae. *Brit. J. Psychol*., **2**, 227-242.
7. Karber, G. (1931). Beitrag zur kollektiven behandlung pharmakologischer reihenversuche. *Arch. Exp. Path. Pharmak*, **162**, 480-487.
8. Thompson, W. R. (1947). Use of moving averages and interpolation to estimate median effective dose. I: Fundamental formulas, estimation of error, and relation to other methods, Bacteriol. *Rev*., **11**, 116-145.
9. Bennett, B. M. (1952). Estimation of the LD50 by moving averages. *J. Hygiene*, **50**, 157-164.
10. Bennett, B. M. (1963). Optimum moving averages for the estimation of median effective dose in bioassay. *J. Hygiene*, **61**, 401-406.
11. Weil, C. S. (1952). Tables for convenient calculation of median-effective dose (LD50 or ED50) and instructions in their use. *Biometrics,* **8**, 249-263.
12. Finney, D. J. (1950). The estimation of the mean of a normal tolerance distribution. *Sankhya* **10**, 341-360.
13. Dixon, W.J. and Mood, A.M. (1948). A method for obtaining and analyzing sensitivity data. *J. Am. Stati. Assoc*., **43**, 109-126.
14. Prentice, R. L. (1976). Generalization of the probit and logit methods for dose response curves. *Biometrics*, **32**, 761-768.
15. Hamilton, M. A., Russo, R. C. and Thurston, R. V. (1977). Trimmed Spearman-Karber method for estimating median lethal concentrations in toxicity bioassay. *Envir. Sci. & Technol*., **11**, 714-719.
16. Copas, J. B. (1983). Plotting *p* against *x*. *Applied Statistics*, **32**, 25-31.
17. ASTM Committee E-18. (1990). Standard definitions of terms relating to sensory evaluation of materials and products. In *Annual Book of ASTM Standards*. E-253-90. American Society for Testing and Materials, Philadelphia, PA.
18. ASTM Committee E-18. (1991). Standard practice for defining and calculating individual and group sensory thresholds from forced-choice data sets of intermediate size, in *Annual Book of ASTM Standards*. E-1432-91. American Society for Testing and Materials. Philadelphia, PA.
19. Cox, D. R. (1983). Some remarks on overdispersion. *Biometrika,* **70**(1), 296-374.
20. Ennis, D. M. and Bi, J. (1998). The beta-binomial model: Accounting for inter-intral variation in replicated difference and preference tests. *Journal of Sensory Studies,* **13**, 389-412.
21. Bi, J., and Ennis, D. M. (1998). Sensory thresholds: Concepts and methods. *Journal of Sensory Studies,* **13**, 133-148.