Additivity Tests

George Karabatsos

Volume 1, pp. 25–29

in

Encyclopedia of Statistics in Behavioral Science

ISBN-13: 978-0-470-86080-9
ISBN-10: 0-470-86080-4

Editors

Brian S. Everitt & David C. Howell

Additivity Tests

When a test of additivity is performed on a set of data, the null hypothesis that a dependent variable is an additive, noninteractive function of two (or more) independent variables and the alternative hypothesis of nonadditivity are characterized by one or more interactions between the independent variables. If the dependent variable is on a quantitative (interval or ratio) scale (*see* **Measurement: Overview**), it is possible to perform a test of additivity in the context of **analysis of variance** (ANOVA). A more general test of additivity is achieved in the context of additive conjoint measurement theory. According to this theory, in order for additivity to hold on some monotonic transformation of the dependent variable, such that combinations of the independent variables are measurable on a common interval scale, it is necessary for data to be consistent with a hierarchy of (qualitative) cancellation axioms. The following two sections describe tests of additivity that are based on ANOVA and additive conjoint measurement, respectively.

Testing for Additivity, Assuming Quantitative Measurement

Suppose that IJ exchangeable sequences $\{Y_{1_{ij}}, \ldots, Y_{n_{ij}}, \ldots, Y_{N_{ij}}; i = 1, \ldots, I, j = 1, \ldots, J\}$ of data are observed from a two-factor experimental design, where $Y_{n_{ij}}$ refers to the nth observation of a quantitative dependent variable in cell ij, corresponding to a level $i \in \{1, \ldots, I\}$ of one independent variable and level $j \in \{1, \ldots, J\}$ of a second independent variable. It is natural to model such data by a two-way ANOVA, given by

$$Y_{n_{ij}} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{n_{ij}} \qquad (1)$$

for all levels $i = 1, \ldots, I$ and $j = 1, \ldots, J$, where μ is the grand mean of the dependent variable, the population parameter α_i represents the main effect of level i, the parameter β_j is the main effect of level j, the parameter γ_{ij} is the interaction effect of levels i and j, and $\varepsilon_{n_{ij}}$ is error assumed to be a random sample from a $N(0,\sigma^2_{\text{Error}})$ normal distribution (*see* **Analysis of Variance**; **Multiple Linear Regression**; **Repeated Measures Analysis of Variance**). In an ANOVA, the well-known F-statistic

$$F_{\text{Int}} = \frac{\sigma^2_{\text{Int}}}{\sigma^2_{\text{Error}}} = \frac{SS_{\text{Int}}/df_{\text{Int}}}{SS_{\text{Error}}/df_{\text{Error}}} \qquad (2)$$

provides a test of the null hypothesis of additivity $H_0: \{\gamma_{ij} = 0, \forall i, j\}$ versus the alternative hypothesis of nonadditivity $H_1: \{\gamma_{ij} \neq 0, \text{ for some } i, j\}$. Under H_0, statistic (2) follows an F distribution with $\{df_{\text{Int}}, df_{\text{Error}}\}$ degrees of freedom, where σ^2_{Int} and SS_{Int} are the variance and sums-of-squares due to interaction, respectively, and SS_{Error} is the error sums-of-squares (e.g., [34]) (*see* **Catalogue of Probability Density Functions**). Under a chosen Type I error rate, the additive null hypothesis is rejected when the value of F_{Int} is unusually large. The F test (2) can be extended to test for interactions between three or more independent variables, and/or to test for interactions in two or more dependent variables (*see* **Multivariate Analysis of Variance**). Also, there are alternative tests of additivity, such as those [2, 3] based on the rank of the interaction matrix $\gamma = (\gamma_{ij})$, as well as distribution-free tests.

When there is exactly one observation per cell ij, the ANOVA model is saturated, with zero degrees of freedom left ($df_{\text{Error}} = 0$) to perform the F test of additivity. To circumvent a saturated model, several researchers have proposed testing the additivity hypothesis $H_0: \{\gamma_{ij} = 0, \forall i, j\}$, by restricting each of the interaction parameters (*see* **Interaction Effects**) by some specific function, under the nonadditive alternative hypothesis H_1 [7, 10, 11, 18, 22–25, 27, 28, 33, 35, 36]. For example, Tukey [36, 33] proposed testing $H_1: \{\gamma_{ij} = \lambda\alpha_i\beta_j \neq 0, \text{ some } i, j\}$, while Johnson and Graybill [11] proposed $H_1: \{\gamma_{ij} = \lambda\delta_i\xi_j \neq 0, \text{ some } i, j\}$, where the so-called 'free' parameters λ, δ_i, and ξ_j represent sources of interaction that are not due to main effects. Alternatively, Tusell [37] and Boik [4,5] proposed tests of additivity that do not require the data analyst to assume any particular functional form of interaction, and, in fact, they are sensitive in detecting many forms of nonadditivity [6].

Testing for Additivity, without Assuming Quantitative Measurement

Let an element of a (nonempty) product set $ax \in A_1 \times A_2$ denote the dependent variable that results after combining the effect of level $a \in A_1 = \{a, b, c, \ldots\}$ from one independent variable, and the effects of level $x \in A_2 = \{x, y, z, \ldots\}$ from another independent variable. According to the theory of additive conjoint measurement [20], the effects of two independent variables are additive if and only if

$$ax \succsim by \text{ implies } f_1(a) + f_2(x) \geq f_1(b) + f_2(y) \tag{3}$$

holds for all $ax, by \in A_1 \times A_2$, where $\succsim$ denotes a weak order, and the functions f_1 and f_2 map the observed effects of the independent variables onto interval scales. In order for the additive representation (3) to hold for some monotonic transformation of the dependent variable, a hierarchy of cancellation axioms must be satisfied [17, 20, 26]. For example, single cancellation (often called order-independence) is satisfied when

$$ax \succsim bx \text{ if and only if } ay \succsim by \tag{4a}$$

$$ax \succsim ay \text{ if and only if } bx \succsim by \tag{4b}$$

hold for all $a, b \in A_1$ and all $x, y \in A_2$. Double cancellation is satisfied when

$$ay \succsim bx \text{ and } bz \succsim cy \text{ implies } az \succsim cx \tag{5}$$

holds for all $a, b, c \in A_1$ and all $x, y, z \in A_2$. The additive representation (3) and cancellation axioms can be extended to any number of independent variables [17], and, of course, an additive representation is unnecessary when all independent variables have zero effects.

In evaluating the fit of data to the cancellation axioms, many researchers have either counted the number of axiom violations or employed multiple nonparametric test statistics (e.g., [8, 9, 19, 26, 29–31, 38]) (*see* **Binomial Confidence Interval**; **Binomial Distribution: Estimating and Testing Parameters**; **Median**; **Kendall's Coefficient of Concordance**; **Kendall's Tau – τ**). Unfortunately, such approaches to testing additivity are not fully satisfactory. They assume that different tests of cancellation are statistically independent, which they are not. Also, as is well-known, the Type I error rate quickly increases with the number of statistical tests performed.

These statistical issues are addressed with a model-based approach to testing cancellation axioms. Suppose that $\{Y_{1_k}, \ldots, Y_{n_k}, \ldots, Y_{N_k};\ k = 1, \ldots, m\}$ are m exchangeable sequences of N_k observations of a dependent variable, where Y is either a real-valued scalar or vector, and each sequence arises from some experimental condition $k \in \{1, \ldots, m\}$. For example, $m = IJ$ conditions may be considered in a two-factor experimental design. According to **de Finetti's** representation theorem (e.g., [1]), the following Bayesian model describes the joint probability of m exchangeable sequences:

$$\begin{aligned} &p(Y_{1_k}, \ldots, Y_{n_k}, \ldots, Y_{N_k}; k = 1, \ldots, m) \\ &= \int_{\Theta \subseteq \Omega} \prod_{k=1}^{m} \prod_{n_k=1}^{N_k} p(Y_{n_k} | \Theta_k) p(\Theta_1, \ldots, \Theta_k, \ldots, \Theta_m) \\ &\quad \times d\Theta_1, \ldots, \Theta_k, \ldots, \Theta_m, \end{aligned} \tag{6}$$

where $p(Y_{n_k} | \Theta_k)$ is the sampling likelihood at data point Y_{n_k} given the kth population parameter Θ_k, and $p(\Theta_1, \ldots, \Theta_k, \ldots, \Theta_m)$ is the prior distribution over the parameter vector $\Theta = (\Theta_1, \ldots, \Theta_k, \ldots, \Theta_m)$ (*see* **Bayesian Statistics**). The notation $\Theta \subseteq \Omega$ refers to the fact that any set of cancellation axioms implies order-restrictions on the dependent variable (as shown in (4) and (5)), such that the parameter vector Θ is constrained to lie within a proper subset Ω of its total parameter space. The form of the constraint Ω depends on the set of cancellation axioms under consideration. A test of a set of cancellation axioms is achieved by testing the fit of a set of data $\{Y_{1_k}, \ldots, Y_{n_k}, \ldots, Y_{N_k};\ k = 1, \ldots, m\}$ to the model in (6).

Karabatsos [12, 15, 16] implemented this approach for testing several cancellation axioms, in the case where, for $k = 1, \ldots, m$, the dependent variable is dichotomous $Y_{n_k} \in \{0, 1\}$ and Θ_k is a binomial parameter. For example, Karabatsos [12] tested single cancellation (4) by evaluating the fit of dichotomous data $\{Y_{1_k}, \ldots, Y_{n_k}, \ldots, Y_{N_k};\ k = 1, \ldots, m\}$ to the model in (6), where all $m = IJ$ binomial parameters were subject to the constraint Ω that $\Theta_{ij} \leq \Theta_{i+1,j}$ for all ordered levels $i = 1, \ldots, I - 1$ of the first independent variable and $\Theta_{ij} \leq \Theta_{i,j+1}$ for all ordered levels $j = 1, \ldots, J - 1$ of the second independent variable.

Karabatsos [14] later generalized this binomial approach, by considering a vector of multinomial parameters $\Theta_k = (\theta_{1k}, \ldots, \theta_{rk}, \ldots, \theta_{Rk})$ for each experimental condition $k = 1, \ldots, m$, where θ_{rk} refers to the probability of the rth response pattern. Each response pattern is characterized by a particular weak order defined over all elements of $A_1 \times A_2$. In this context, Ω refers to the sum-constraint $\sum_{r_k \in \sim V_k} \theta_{rk} \geq C$ for each experimental condition k, and some chosen threshold $C \in [1/2, 1]$, where $\sim V_k$ is the set of response patterns that do not violate a given cancellation axiom.

Karabatsos [13] proposed a slightly different multinomial model, as a basis for a Bayesian bootstrap [32] approach to isotonic (inequality-constrained) regression (*see* **Bootstrap Inference**). This procedure can be used to estimate the nonparametric posterior distribution of a discrete- or continuous-valued dependent variable Y, subject to the order-constraints of the set of all possible linear orders (for example, $Y_1 \leq Y_2 \leq \cdots \leq Y_k \leq \cdots \leq Y_m$) that satisfy the *entire hierarchy* of cancellation axioms. Here, a test of additivity is achieved by evaluating the fit of the observed data $\{Y_{1_k}, \ldots, Y_{n_k}, \ldots, Y_{N_k};\ k = 1, \ldots, m\}$ to the corresponding order-constrained posterior distribution of Y.

Earlier, as a non-Bayesian approach to additivity testing, Macdonald [21] proposed isotonic regression to determine the least-squares **maximum-likelihood estimate** (MLE) of the dependent variable $\{\hat{Y}_k;\ k = 1, \ldots, m\}$, subject to a linear order-constraint (e.g., $Y_1 \leq Y_2 \leq \cdots \leq Y_k \leq \cdots \leq Y_m$) that satisfies a given cancellation axiom (*see* **Least Squares Estimation**). He advocated testing each cancellation axiom separately, by evaluating the fit of the observed data $\{Y_{1_k}, \ldots, Y_{n_k}, \ldots, Y_{N_k};\ k = 1, \ldots, m\}$ to the MLE $\{\hat{Y}_k;\ k = 1, \ldots, m\}$ under the corresponding axiom.

Acknowledgments

Karabatsos's research is supported by National Science Foundation grant SES-0242030, program of Methodology, Measurement, and Statistics. Also, this work is supported in part by Spencer Foundation grant SG2001000020.

References

[1] Bernardo, J.M. (2002). *Bayesian Theory* (second reprint), Wiley, New York.

[2] Boik, R.J. (1986). Testing the rank of a matrix with applications to the analysis of interactions in ANOVA, *Journal of the American Statistical Association* **81**, 243–248.

[3] Boik, R.J. (1989). Reduced-rank models for interaction in unequally-replicated two-way classifications, *Journal of Multivariate Analysis* **28**, 69–87.

[4] Boik, R.J. (1990). Inference on covariance matrices under rank restrictions, *Journal of Multivariate Analysis* **33**, 230–246.

[5] Boik, R.J. (1993a). Testing additivity in two-way classifications with no replications: the locally best invariant test, *Journal of Applied Statistics* **20**, 41–55.

[6] Boik, R.J. (1993b). A comparison of three invariant tests of additivity in two-way classifications with no replications, *Computational Statistics and Data Analysis* **15**, 411–424.

[7] Boik, R.J. & Marasinghe, M.G. (1989). Analysis of non-additive multiway classifications, *Journal of the American Statistical Association* **84**, 1059–1064.

[8] Coombs, C.H. & Huang, L.C. (1970). Polynomial psychophysics of risk, *Journal of Mathematical Psychology* **7**, 317–338.

[9] Falmagne, J.-C. (1976). Random conjoint measurement and loudness summation, *Psychological Review* **83**, 65–79.

[10] Harter, H.L. & Lum, M.D. (1962). An interpretation and extension of Tukey's one-degree of freedom for non-additivity, *Aeronautical Research Laboratory Technical Report, ARL*, pp. 62–313.

[11] Johnson, D.E. & Graybill, F.A. (1972). An analysis of a two-way model with interaction and no replication, *Journal of the American Statistical Association* **67**, 862–868.

[12] Karabatsos, G. (2001). The Rasch model, additive conjoint measurement, and new models of probabilistic measurement theory, *Journal of Applied Measurement* **2**, 389–423.

[13] Karabatsos, G. (2004a). *A Bayesian Bootstrap Approach To Testing The Axioms Of Additive Conjoint Measurement*. Manuscript under review.

[14] Karabatsos, G. (2004b). The exchangeable multinomial model as an approach to testing deterministic axioms of choice and measurement. To appear, *Journal of Mathematical Psychology*.

[15] Karabatsos, G. & Sheu, C.-F. (2004). Order-constrained Bayes inference for dichotomous models of non-parametric item-response theory, *Applied Psychological Measurement* **2**, 110–125.

[16] Karabatsos, G. & Ullrich, J.R. (2002). Enumerating and testing conjoint measurement models, *Mathematical Social Sciences* **43**, 485–504.

[17] Krantz, D.H., Luce, R.D., Suppes, P. & Tversky, A. (1971). *Foundations of Measurement: Additive and Polynomial Representations*, Academic Press, New York.

[18] Krishnaiah, P.R. & Yochmowitz, M.G. (1980). Inference on the structure of interaction in the two-way classification model, in *Handbook of Statistics*, Vol. 1, P.R. Krishnaiah, ed., North Holland, Amsterdam, pp. 973–984.

[19] Levelt, W.J.M., Riemersma, J.B. & Bunt, A.A. (1972). Binaural additivity of loudness, *British Journal of Mathematical and Statistical Psychology* **25**, 51–68.

[20] Luce, R.D. & Tukey, J.W. (1964). Additive conjoint measurement: a new type of fundamental measurement, *Journal of Mathematical Psychology* **1**, 1–27.

[21] Macdonald, R.R. (1984). Isotonic regression analysis and additivity, in *Trends in Mathematical Psychology*, E. Degreef & J. Buggenhaut, eds, Elsevier Science Publishers, North Holland, pp. 239–255.

[22] Mandel, J. (1961). Non-additivity in two-way analysis of variance, *Journal of the American Statistical Association* **56**, 878–888.

[23] Mandel, J. (1969). The partitioning of interaction in analysis of variance, *Journal of Research, National Bureau of Standards* **B 73B**, 309–328.

[24] Mandel, J. (1971). A new analysis of variance model for non-additive data, *Technometrics* **13**, 1–18.

[25] Marasinghe, M.G. & Boik, R.J. (1993). A three-degree of freedom test of additivity in three-way classifications, *Computational Statistics and Data Analysis* **16**, 47–61.

[26] Michell, J. (1990). *An Introduction to the Logic of Psychological Measurement*, Lawrence Earlbaum Associates, Hillsdale.

[27] Milliken, G.A. & Graybill, F.A. (1970). Extensions of the general linear hypothesis model, *Journal of the American Statistical Association* **65**, 797–807.

[28] Milliken, G.A. & Johnson, D.E. (1989). *Analysis of Messy Data*, Vol. 2, Van Nostrand Reinhold, New York.

[29] Nygren, T.E. (1985). An examination of conditional violations of axioms for additive conjoint measurement, *Applied Psychological Measurement* **9**, 249–264.

[30] Nygren, T.E. (1986). A two-stage algorithm for assessing violations of additivity via axiomatic and numerical conjoint analysis, *Psychometrika* **51**, 483–491.

[31] Perline, R., Wright, B.D. & Wainer, H. (1979). The Rasch model as additive conjoint measurement, *Applied Psychological Measurement* **3**, 237–255.

[32] Rubin, D.B. (1981). The Bayesian bootstrap, *Annals of Statistics* **9**, 130–134.

[33] Scheffé, H. (1959). *The Analysis of Variance* (Sixth Printing), Wiley, New York.

[34] Searle, S.R. (1971). *Linear Models*, John Wiley & Sons, New York.

[35] Tukey, J.W. (1949). One degree of freedom for non-additivity, *Biometrics* **5**, 232–242.

[36] Tukey, J.W. (1962). The future of data analysis, *Annals of Mathematical Statistics* **33**, 1–67.

[37] Tusell, F. (1990). Testing for interaction in two-way ANOVA tables with no replication, *Computational Statistics and Data Analysis* **10**, 29–45.

[38] Tversky, A. (1967). Additivity, utility, and subjective probability, *Journal of Mathematical Psychology* **4**, 175–201.

George Karabatsos