

Choosing the More Likely Hypothesis

Richard Startz
Department of Economics
University of California, Santa Barbara, USA
startz@ucsb.edu

now

the essence of knowledge

Boston — Delft

Foundations and Trends[®] in Econometrics

Published, sold and distributed by:

now Publishers Inc.
PO Box 1024
Hanover, MA 02339
United States
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:

now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is

R. Startz. *Choosing the More Likely Hypothesis*. Foundations and Trends[®] in Econometrics, vol. 7, no. 2, pp. 119–189, 2014.

This Foundations and Trends[®] issue was typeset in L^AT_EX using a class file designed by Neal Parikh. Printed on acid-free paper.

ISBN: 978-60198-899-7
© 2014 R. Startz

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, mechanical, photocopying, recording or otherwise, without prior written permission of the publishers.

Photocopying. In the USA: This journal is registered at the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923. Authorization to photocopy items for internal or personal use, or the internal or personal use of specific clients, is granted by now Publishers Inc for users registered with the Copyright Clearance Center (CCC). The 'services' for users can be found on the internet at: www.copyright.com

For those organizations that have been granted a photocopy license, a separate system of payment has been arranged. Authorization does not extend to other kinds of copying, such as that for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale. In the rest of the world: Permission to photocopy must be obtained from the copyright owner. Please apply to now Publishers Inc., PO Box 1024, Hanover, MA 02339, USA; Tel. +1 781 871 0245; www.nowpublishers.com; sales@nowpublishers.com

now Publishers Inc. has an exclusive license to publish this material worldwide. Permission to use this content must be obtained from the copyright license holder. Please apply to now Publishers, PO Box 179, 2600 AD Delft, The Netherlands, www.nowpublishers.com; e-mail: sales@nowpublishers.com

Foundations and Trends[®] in Econometrics
Volume 7, Issue 2, 2014
Editorial Board

Editor-in-Chief

William H. Greene
New York University
United States

Editors

Manuel Arellano
CEMFI, Spain

Wiji Arulampalam
University of Warwick

Orley Ashenfelter
Princeton University

Jushan Bai
Columbia University

Badi Baltagi
Syracuse University

Anil Bera
University of Illinois

Tim Bollerslev
Duke University

David Brownstone
UC Irvine

Xiaohong Chen
Yale University

Steven Durlauf
University of Wisconsin

Amos Golan
American University

Bill Griffiths
University of Melbourne

James Heckman
University of Chicago

Jan Kiviet
University of Amsterdam

Gary Koop
University of Strathclyde

Michael Lechner
University of St. Gallen

Lung-Fei Lee
Ohio State University

Larry Marsh
University of Notre Dame

James MacKinnon
Queens University

Bruce McCullough
Drexel University

Jeff Simonoff
New York University

Joseph Terza
Purdue University

Ken Train
UC Berkeley

Pravin Trivedi
Indiana University

Adonis Yatchew
University of Toronto

Editorial Scope

Topics

Foundations and Trends[®] in Econometrics publishes survey and tutorial articles in the following topics:

- Econometric models
- Simultaneous equation models
- Estimation frameworks
- Biased estimation
- Computational problems
- Microeconometrics
- Treatment modeling
- Discrete choice modeling
- Models for count data
- Duration models
- Limited dependent variables
- Panel data
- Time series analysis
- Latent variable models
- Qualitative response models
- Hypothesis testing
- Econometric theory
- Financial econometrics
- Measurement error in survey data
- Productivity measurement and analysis
- Semiparametric and nonparametric estimation
- Bootstrap methods
- Nonstationary time series
- Robust estimation

Information for Librarians

Foundations and Trends[®] in Econometrics, 2014, Volume 7, 4 issues. ISSN paper version 1551-3076. ISSN online version 1551-3084. Also available as a combined paper and online subscription.

Foundations and Trends® in Econometrics
Vol. 7, No. 2 (2014) 119–189
© 2014 R. Startz
DOI: 10.1561/08000000028



Choosing the More Likely Hypothesis¹

Richard Startz
Department of Economics,
University of California, Santa Barbara, USA
startz@ucsb.edu

¹Portions of this paper appeared in an earlier working paper titled, “How Should An Economist Do Statistics?” Advice from Jerry Hausman, Shelly Lundberg, Gene Savin, Meredith Startz, Doug Steigerwald, and members of the UCSB Econometrics Working Group is much appreciated.

Contents

1	Introduction	2
2	Choosing Between Hypotheses	6
3	Bayes Theorem	10
3.1	Bayes theorem applied to traditional estimators	10
3.2	Bayes theorem and power	12
3.3	Does the p -value approximate the probability of the null? .	14
4	A Simple Coin-Flipping Example	16
4.1	Point null and point alternative hypotheses	17
4.2	The probability of the null and the traditional p -value . . .	18
4.3	Choosing the null, the alternative, or remaining undecided	20
4.4	Implicit priors	21
4.5	The importance of the alternative	24
4.6	A continuous alternative for the coin-toss example	25
5	Regression Estimates	28
5.1	Uniform prior for the alternative hypothesis	29
5.2	Normal prior for the alternative hypothesis	34
5.3	One-sided hypotheses	38

6	Diffuse Alternatives and the Lindley “Paradox”	41
7	Is the Stock Market Efficient?	44
8	Non-sharp Hypotheses	49
9	Bayes Theorem and Consistent Estimation	52
10	More General Bayesian Inference	54
10.1	Use of the BIC	54
10.2	A light derivation of the BIC	55
10.3	Departures from the Bayesian approach	59
11	The General Decision-theoretic Approach	61
11.1	Wald’s method	61
11.2	Akaike’s method	63
12	A Practitioner’s Guide to Choosing Between Hypotheses	64
13	Summary	66
	References	68

Abstract

Much of economists' statistical work centers on testing hypotheses in which parameter values are partitioned between a null hypothesis and an alternative hypothesis in order to distinguish two views about the world. Our traditional procedures are based on the probabilities of a test statistic under the null but ignore what the statistics say about the probability of the test statistic under the alternative. Traditional procedures are not intended to provide evidence for the relative probabilities of the null versus alternative hypotheses, but are regularly treated as if they do. Unfortunately, when used to distinguish two views of the world, traditional procedures can lead to wildly misleading inference. In order to correctly distinguish between two views of the world, one needs to report the probabilities of the hypotheses given parameter estimates rather than the probability of the parameter estimates given the hypotheses. This monograph shows why failing to consider the alternative hypothesis often leads to incorrect conclusions. I show that for most standard econometric estimators, it is not difficult to compute the proper probabilities using Bayes theorem. Simple formulas that require only information already available in standard estimation reports are provided. I emphasize that frequentist approaches for deciding between the null and alternative hypothesis are not free of priors. Rather, the usual procedures involve an implicit, unstated prior that is likely to be far from scientifically neutral.

1

Introduction

Much of economists' statistical work centers on testing hypotheses in which parameter values are partitioned between a null hypothesis and an alternative hypothesis. In essence, we are trying to distinguish between two views about the world. We then ask where the estimated coefficient (or test statistic) lies in the distribution implied by the null hypothesis. If the estimated coefficient is so far out in the tail of the distribution that it is very unlikely we would have found such an estimate under the null, we reject the null and conclude there is significant evidence in favor of the alternative. But this is a terribly incomplete exercise, omitting any consideration of how unlikely it would be for us to see the estimated coefficient if the alternative were true. Pearson [1938, p. 242] put the argument this way,¹

[the] idea which has formed the basis of all the . . . researches of Neyman and myself . . . is the simple suggestion that the only valid reason for rejecting a statistical hypothesis is that some alternative hypothesis explains the events with a greater degree of probability.

¹As quoted by Weakliem [1999a,b, p. 363].

The principle that the probability of a realized coefficient under the alternative matters is at once well-understood and near-universally ignored by economists. What is less appreciated is the practical point: Our standard procedure of stating whether a coefficient is statistically significant (or equivalently whether the hypothesized value of a coefficient lies outside the confidence interval, or equivalently whether the p -value is small) can be a terribly misleading guide as to the odds favoring the null hypothesis relative to the alternative hypothesis. I give examples below to show just how misleading our usual procedures can be. Of course, for practice to change, there needs to be a better way to conduct inference. I present alternative procedures that can be easily implemented in our most common hypothesis testing situations.

My goal here is to offer a perspective on how economists should choose between hypotheses. While some of the points are original, many are not. After all, much of the paper comes down to saying “remember Bayes theorem,” which has likely been around since Bayes [1763]; or according to the delightful account by McGrayne [2011], at least since Laplace [1774]. While it is entirely clear that economists do choose between hypotheses using statistical tests as if Bayes theorem does not exist, it is not because we have not been reminded of the danger of such practice. It seems the advice didn’t take. Leamer [1983a,b] laid out much of the argument in the very first volume of the *Handbook of Econometrics*. McCloskey [1992] reports a discussion in which Ken Arrow said, “Statistical significance in its usual form is indefensible.” In an influential article in the medical literature, Ioannidis [2005] reminds medical researchers “. . . the probability that a research finding is indeed true depends on the prior probability of it being true. . . , the statistical power of the study, and the level of statistical significance.” Kass and Raftery [1995] offer some of the theory behind what’s said below. The discussion in this monograph is at least foreshadowed in Pearson [1938] and Arrow [1960] and parts are pretty explicit in Leamer [1978, 1983a] and Raftery [1986a, 1995]. The hope is that by (a) giving blunt examples of the consequences of ignoring Bayes theorem and (b) offering very easy ways to adjust frequentist statistics to properly account for Bayes theorem, econometric practice may change more than it has in the past.

This monograph is aimed primarily at the classical, frequentist, econometrician who needs to choose between hypotheses. Most results are illustrated in the context of the most simple econometric situation, one where we have a normally distributed estimator, $\hat{\theta} \sim N(\theta, \sigma_{\hat{\theta}}^2)$, and a null hypothesis $\theta = \theta_0$ versus an alternative $\theta \neq \theta_0$. The canonical example is a test of a regression coefficient. There are five major points.

1. The traditional use of classical hypothesis testing to choose between hypotheses leads to misleading results. As a practical matter, standard practice can be very, *very* misleading. It is entirely possible to strongly reject the null in cases where the null is more likely than the alternative, and vice versa.
2. Choosing between hypotheses requires invoking Bayes theorem. For the most common empirical applications at least, those where the estimated coefficients are approximately normal, applying Bayes theorem is very easy.
3. Once one acknowledges that one wants to compare a null hypothesis to an alternative, something has to be said about the likelihood of particular values of the parameter of interest under the alternative. Use of Bayes theorem does require specifying some prior beliefs. Sometimes this can be done in a way in which the specified priors take a neutral stance between null and alternative; sometimes a completely neutral stance is more difficult.
4. The notion that frequentist procedures specify a null and then take a neutral stance with regard to parameter values under the alternative is wrong. Frequentist decision rules are equivalent to adopting an implicit prior. The implicit prior is often decidedly non-neutral.
5. Economic hypotheses are usually best distinguished by some parameter being small or large, rather than some parameter being exactly zero versus non-zero. Application of Bayes theorem permits the former, preferred, kind of hypothesis comparison by considering non-sharp nulls. The calculations required for choosing between non-sharp hypotheses are straightforward.

All of this is, obviously, related to frequentist versus Bayesian approaches to econometrics. This paper is addressed to the frequentist econometrician. Nothing addresses any of the philosophical differences between frequentists and Bayesians. Some Bayesian tools are used, although these are really just statements of probability theory and should be uncontroversial. A succinct statement of the goal of the monograph is this:

After running a regression, the empirical economist should be able to draw an inference about the probability of a null versus an alternative hypothesis that is both correct and easy to make.

References

- K. J. Arrow. Decision theory and the choice of a level of significance for the t -test. In C. Olkin, M. Hoeffding, and Mann, editors, *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*. Stanford University Press, Stanford, CA, 1960.
- M. S. Bartlett. A comment on D.V. Lindley's Statistical Paradox. *Biometrika*, 44:533–534, 1957.
- T. Bayes. An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions (Royal Society of London)*, 53:269–271, 1763.
- K. P. Burnham and D. R. Anderson. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer, New York, 2nd edition, 2002.
- G. Casella and R. L. Berger. Reconciling bayesian and frequentist evidence in the one-sided testing problem. *Journal of the American Statistical Association*, 82(387):106–111, 1987.
- M. H. Degroot. Doing what comes naturally: Interpreting a tail area as a posterior probability or as a likelihood ratio. *Journal of the American Statistical Association*, 68(344):966–969, 1973.
- J. M. Dickey. The weighted likelihood ratio, linear hypotheses on normal location parameters. *The Annals of Mathematical Statistics*, 42(1):204–223, 1971.
- J. M. Dickey. Is the tail area useful as an approximate bayes factor? *Journal of the American Statistical Association*, 72(357):138–142, 1977.

- A. C. Doyle. *Silver Blaze*. reprinted 1993 Oxford University Press, New York, 1893.
- D. Firth and J. Kuha. Comments on ‘a critique of the bayesian information criterion for model selection’. *Sociological Methods & Research*, 27(3):398–402, 1999.
- A. Gelman and D. B. Rubin. Evaluating and using methods in the social sciences: A discussion of ‘a critique of the bayesian information criterion for model selection’. *Sociological Methods & Research*, 27(3):403–410, 1999.
- J. Geweke. *Contemporary Bayesian Econometrics and Statistics*. John Wiley & Sons, Hoboken, NJ, 2005.
- E. Greenberg. *Introduction to Bayesian Econometrics*. Cambridge University Press, Cambridge, 2008.
- J. A. Hausman. Specification tests in econometrics. *Econometrica*, 46(6):1251–1271, 1978.
- J. P. A. Ioannidis. Why most published research findings are false. *PLoS Medicine*, 2(8):696–701, 2005.
- R. E. Kass and A. E. Raftery. Bayes factors. *Journal of the American Statistical Association*, 90(430):773–796, 1995.
- G. Koop. *Bayesian Econometrics*. John Wiley & Sons, 2003.
- P. Laplace. Mémoire sur la probabilité des causes par les événemens. *Savants étranges*, 6:621–656, 1774.
- E. E. Leamer. *Specification Searches: Ad Hoc Inference with Nonexperimental Data*. John Wiley and Sons, New York, 1978.
- E. E. Leamer. Model choice and specification analysis. In Z. Griliches and M. D. Intriligator, editors, *Handbook of Econometrics*, volume 1, pages 285–330. North Holland Publishing Co., 1983a.
- E. E. Leamer. Let’s take the con out of econometrics. *American Economic Review*, 73(1):31–43, 1983b.
- Y. Li, T. Zeng, and J. Yu. A new approach to bayesian hypothesis testing. *Journal of Econometrics*, 178(3):602–612, 2014.
- D. V. Lindley. A statistical paradox. *Biometrika*, 44:187–192, 1957.
- D. N. McCloskey. The loss function has been mislaid: The rhetoric of significance tests. *American Economic Review*, 75(2):201–205, 1985.
- D. N. McCloskey. Other things equal: The bankruptcy of statistical significance. *Eastern Economic Journal*, 18(3):359–361, 1992.

- D. N. McCloskey and Stephen T. Ziliak. The standard error of regressions. *Journal of Economic Literature*, XXXIV:97–114, 1996.
- S. B. McGrayne. *The Theory That Would Not Die: How Bayes' Rule Cracked the Enigma Code, Hunted Down Russian Submarines, and Emerged Triumphant from Two Centuries of Controversy*. Yale University Press, New Haven, CT, 2011.
- J. Neyman and E. S. Pearson. On the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 231:289–337, 1933.
- E. S. Pearson. 'Student' as a statistician. *Biometrika*, 30:210–250, 1938.
- D. Poirier. *Intermediate Statistics and Econometrics*. M.I.T. Press, Cambridge, 1995.
- A. E. Raftery. Choosing models for cross-classifications. *American Sociological Review*, 51(1):145–146, 1986a.
- A. E. Raftery. A note on bayes factors for log-linear contingency table models with vague prior information. *Journal of the Royal Statistical Society, Series B*, 48:249–250, 1986b.
- A. E. Raftery. Bayesian model selection in social research. *Sociological Methodology*, 25:111–163, 1995.
- A. E. Raftery. Bayes factors and BIC: Comment on 'a critique of the bayesian information criterion for model selection'. *Sociological Methods & Research*, 27(3):411–427, 1999.
- G. Schwarz. Estimating the dimension of a model. *The Annals of Statistics*, 6:461–464, 1978.
- H. Simon. Statistical tests as a basis for 'yes-no' choices. *Journal of the American Statistical Association*, 40(229):80–84, 1945.
- R. Startz. On the implicit uniform bic prior. *Economics Bulletin*, 34(2), April 2014.
- A. Wald. *Statistical Decision Functions*. Wiley, New York, 1950.
- D. L. Weakliem. A critique of the bayesian information criterion for model selection. *Sociological Methods & Research*, 27(3):359–397, 1999a.
- D. L. Weakliem. Reply to Firth and Kuha, Gelman and Rubin, Raftery, and Xie. *Sociological Methods & Research*, 27(3):436–443, 1999b.

References

71

- Y. Xie. The tension between generality and accuracy. *Sociological Methods & Research*, 27(3):428–435, 1999.
- S. T. Ziliak and D. N. McCloskey. *The Cult of Statistical Significance*. University of Michigan Press, Ann Arbor, MI, 2008.