Stat 209B-- Lectures, Course Files, and Readings

Week 0

Course introduction (slides and audio posted on main page)

Background readings (not required, but of interest if you haven't seen these before)

1. Correlation and Causation: A Comment, Stephen Stigler *Perspectives in Biology and Medicine*, volume 48, number 1 supplement (winter 2005)
2. Secret to Winning a Nobel Prize? Eat More Chocolate (Time)
   Publication: Chocolate Consumption, Cognitive Function, and Nobel Laureates Franz H. Messerli, M.D. N Engl J Med 2012; 367:1562-1564 October 18, 2012
3. *David Freedman chapters*.
   From Association to Causation: Some Remarks on the History of Statistics;
   Statistical Models for Causation: A critical review
   Statistical Models and Shoe Leather, *Sociological Methodology*, Vol. 21. (1991), pp. 291-313. JStor link

Week 1

Lecture slides, week 1 (pdf)
Audio companion, week 1
parta partb partc

1. Encouragement Designs: example of potential outcomes formulation.

Lecture Topics

Illustration using encouragement design representation in Holland (1988). copies of selected overheads.
Encouragement Designs. Potential outcomes formulation and IV parameter estimation in Holland (1988). Estimation handout
Do regression methods (path analysis) identify causal effects? Demonstrations of failure for Holland's encouragement design. class handout Encouragement design slides

Primary Readings

Paul Holland, Causal Effects and Encouragement Designs. Causal Inference, Path Analysis, and Recursive Structural Equations Models
Paul W. Holland Sociological Methodology, Vol. 18. (1988), pp. 449-484. (Encouragement design results; sections 3-5)
Holland Appendix (esp pp. 475-480) presents the potential outcomes formulation.

Abstract Rubin's model for causal inference in experiments and observational studies is enlarged to analyze the problem of "causes causing causes" and is compared to path analysis and recursive structural equations models.
A special quasi-experimental design, the encouragement design, is used to give concreteness to the discussion by focusing on the simplest problem that involves both direct and indirect causation.
It is shown that Rubin's model extends easily to this situation and specifies conditions under which the parameters of path analysis and recursive structural equations models have causal interpretations.

Encouragement Design research examples:

Sesamee Street evaluation
Gelman-Hill text sec 10.5; Data Analysis Using Regression and Multilevel/Hierarchical Models
Salt and Blood Pressure clinical trial
Publication: Feasibility and efficacy of sodium reduction in the Trials of Hypertension Prevention, phase I Trials of Hypertension Prevention Collaborative Research Group. S K Kumanyika, P R Hebert, J A Cutler, V I Lasser, C P Sugars, L Steffen-Batey, A A Brewer, MI. Hypertension doi: 10.1161/01.HYP.22.4.5021993;22:502-512

2. Mediating (process) variables

Lecture Topics

Historical (Barron-Kenny) methods David Kenny web page
R-implementations: mediating variables data analysis example data file
Barron-Kenny method via Sobel function in the multilevel package.
More extensive implementation (incl BCa bootstrapping) function `mediation` in package `MBESS` Ken Kelley;
power and sample size calculations in package `powerMediation`
`mediation` package. takes the topic up a large level of complexity/capabilities

Primary Readings

Vignette for `mediation` package Causal Mediation Analysis Using R
Mediation Analysis David P. MacKinnon, Amanda J. Fairchild, and Matthew S. Fritz Department of Psychology, Arizona State University, Tempe, Arizona 85287-1104; Annu. Rev. Psychol. 2007. 58:593-614

Mediation research examples:

Framing experiment
Brader T, Valentino NA, Suhat E (2008). What Triggers Public Opposition to Immigration? Anxiety, Group Cues, and Immigration." American Journal of Political Science, 52(4), 959-978. jstor link
Data in `mediation` package; data description and analyses in `mediation` package vignette (linked below)
Bench Science vs Path Analysis: Exercise and Alzheimers
The irisin bench-science mediation example is discussed at the beginning of Week 2 lecture for recap and because I couldn't find it at the time.
NYTimes:How Exercise May Help Keep Our Memory Sharp .
Publication: Exercise-linked FNDC5/irisin rescues synaptic plasticity and memory defects in Alzheimer's models *Nature Medicine* volume 25, pages165-175 (2019)
Mediated moderation?
Stanford Medicine Common opioids less effective for patients on SSRI antidepressants Publication: Predicting inadequate postoperative pain management in depressed patients: A machine learning approach Arjun Parthipan,Imon Banerjee,Keith Humphreys,Steven M. Asch,Catherine Curtin,Ian Carroll ,Tina Hernandez-Boussard Published: February 6, 2019https://doi.org/10.1371/journal.pone.0210575
New Yorker. December 23, 2013. The Power of the Hoodie-Wearing C.E.O. Publication: The Red Sneakers Effect: Inferring Status and Competence from Signals of Nonconformity Author(s): Silvia Bellezza, Francesca Gino, and Anat Keinan Source: *Journal of Consumer Research*

Additional Resources

Mediators and Moderators of Treatment Effects in Randomized Clinical Trials Helena Chmura Kraemer; G. Terence Wilson; Christopher G. Fairburn; W. Stewart Agras Arch Gen Psychiatry. 2002;59:877-883
additional technical papers. Causal Mediation Analysis Using R K. Imai, L. Keele, D. Tingley, and T. Yamamoto American Political Science Review Vol. 105, No. 4 November 2011
Unpacking the Black Box of Causality: Learning about Causal Mechanisms from Experimental and Observational Studies
MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M.,West, S. G., Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. Psychological Methods, 7, 83-104.
Useful expositions Using R
Chapter 14: Mediation and Moderation Alyssa Blair
Mediation and Moderation Analyses with R - OSF presentation slides

Week 1 Review Questions

Question 1. Mediating Variable Computations: Class example continued
The data set shown in class example ss423 is linked above and in the legacy directory http://web.stanford.edu/~rag/stat209/ss423
for predictor (IV) 'belong' outcome 'depress' and (potential) mediating variable 'master' The class example showed you the Baron-Kenny analysis using functions from the multilevel and MBESS packages.
Here just use 'lm' basic regression and the recipees from the class handout to recreate point estimates and asymptotic standard errors, significance tests for the mediating variable effect.
Compare your result with the class example posting.
Extra: also try out the more 'sophisticated' functions in the mediation package.

Solution for question 1

Question 2. Potential Outcomes, Encouragement Design Estimation and (Causal) Mediation
Task 1. Create a potential Outcomes dataset following the first ALICE specification in the posted slides (week 3) ## ALICE example beta = 3 rho = 3 tau = 1, delta = 3 (I did n=400; larger would be better so I redid with n = 6400)

Task 2. Use the artificial data to show the results for the mediation (indirect) effect by hand doing the 3 regressions using multilevel package (sobel) using MBESS package using the causal mediation estimation ACME from the mediation package and compare with rho*beta

Task 3 estimate beta by the Wald estimator (assuming tau = 0) and estimate mediation effect

Solution for question 2

Question 3. Sesame Street: Encouragement Design research example
Sesame Street research setting and data description given pdf p.30 of Lecture 1 (also Gelman text).
For this exercise use `postnumb` : posttest on numbers (0-54), along with the measures encour and regular from the class example in Lecture 1.
Use the encouragement design formulation to estimate the effect on child cognitive development (postnumb here) of watching more Sesame Street.
What assumption is necessary for the IV estimation in this design?
Obtain a point and interval estimate for the effect of viewing (use `ivreg` as in class example).
From simple descriptives reproduce this instrumental variables estimate (Wald estimator).
The second approach (path analysis) analyzed by Holland requires what assumption?
Obtain the path analyses (regression) estimate for the effect on child cognitive development (postnumb here) of watching more Sesame Street.
Compare with the IV estimate (which employs different assumptions).

Solution for question 3

Week 2

Moderating Variables in experimental studies (heterogeneous treatment effects)

To be posted
Lecture slides, week 2 (pdf)
Audio companion, week 2
parta partb partc

Lecture topics
0. Moderation, mediation recap slide
1. Review: formulation and purposes of analysis of covariance
basic (old) ancova exposition slides ancova and extensions, math notes
High School and Beyond (observational study) school means data example HSB ancova handout (ascii version) data for HSB ancova HSB ancova, scanned pdf
2. Moderating variables, Heterogeneous Treatment Effects (CATE).
Analyzing treatment effects as a function of covariate(s)
CNRL, including Johnson-Neyman technique cnrl data cnrl analysis (extended)

Primary Readings
Ancova and extensions
Rogosa, D. R. (1980). Comparing nonparallel regression lines. *Psychological Bulletin*, 88, 307-321. [a better quality scan from the APA site]
R resources (below).

Moderation research examples:
Gender differences in effectiveness of aspirin.
Aspirin may be less effective heart treatment for women than men
Publication: Aspirin Resistance in Patients with Stable Coronary Artery Disease, in the *Annals of Pharmacotherapy* April 2007
Moderating variables can be your friend (statistics is the only friend you need) music: I've got friends in low places
Wash Post: Why smart people are better off with fewer friends .
Publication: Country roads, take me home... to my friends: How intelligence, population density, and friendship affect modern happiness. *British Journal of Psychology* 2016
ATI research
Snow R.E. (1978) Aptitude-Treatment Interactions in Educational Research. In: Pervin L.A., Lewis M. (eds) Perspectives in Interactional Psychology. Springer, Boston, MA. https://doi.org/10.1007/978-1-4613-3997-7_10
Family SES as a moderating variable in nature/nuture:
Why Rich Parents Don't Matter UTexas press release: Being Poor Can Suppress Children's Genetic Potentials Publication: Emergence of a Gene x Socioeconomic Status Interaction on Infant Mental Ability Between 10 Months and 2 years DOI: 10.1177/0956797610392926 Psychological Science published online 17 December 2010 Elliot M. Tucker-Drob, Mijke Rhemtulla, K. Paige Harden, Eric Turkheimer and David Fask

R implementations and Resources
package `probemod` manual
package `interactions` intro vignette: Exploring interactions with continuous predictors in regression models manual

Additional Resources, Ancova and extensions
Improving Present Practices in the Visual Display of Interactions Advances in Methods and Practices in Psychological Science
analysis of covariance: Background/historical papers:
Covariance Adjustment in Randomized Experiments and Observational Studies Paul R. Rosenbaum *Statistical Science*, Vol. 17, No. 3. (Aug., 2002), pp. 286-304. Jstor
Some Aspects of Analysis of Covariance, A Biometrics Invited Paper with Discussion. D. R. Cox; P. McCullagh *Biometrics*, Vol. 38, No. 3, (Sep., 1982), pp. 541-561. Jstor
Analysis of Covariance: Its Nature and Uses William G. Cochran *Biometrics*, Vol. 13, No. 3, Special Issue on the Analysis of Covariance. (Sep., 1957), pp. 261-281. Jstor
The Use of Covariance in Observational Studies W. G. Cochran *Applied Statistics*, Vol. 18, No. 3. (1969), pp. 270-275. Jstor
Estimation of the Slope and Analysis of Covariance when the Concomitant Variable is Measured with Error James S. Degracie; Wayne A. Fuller *Journal of the American Statistical Association*, Vol. 67, No. 340. (Dec., 1972), pp. 930-937. Jstor
Deep background Neter-Wasserman text (Applied linear statistical models. Neter, Kutner, Nachtsheim and Wasserman 1996. Fifth edition. Homewood IL: Irwin, Inc.) chapters 22 and 8.
Johnson-Neyman technique and aptitude-treatment interaction (ATI)
Cronbach, L. J., & Snow, R. E. (1977). Aptitudes and instructional methods: A handbook for research on interactions. Irvington
Regions of Significant Criterion Differences in Aptitude-Treatment-Interaction Research Leonard S. Cahen; Robert L. Linn *American Educational Research Journal*, Vol. 8, No. 3. (May, 1971),

pp. 521-530. Jstor
Identifying Regions of Significance in Aptitude-by-Treatment-Interaction Research Ronald C. Serlin; Joel R. Levin *American Educational Research Journa*l, Vol. 17, No. 3. (Autumn, 1980), pp. 389-399. Jstor
Defining Johnson-Neyman Regions of Significance in the Three-Covariate ANCOVA Using Mathematica Steve Hunka; Jacqueline Leighton *Journal of Educational and Behavioral Statistics*, Vol. 22, No. 4. (Winter, 1997), pp. 361-387. Jstor
discussion of substantive issues: Trait-Treatment Interaction and Learning David C. Berliner; Leonard S. Cahen *Review of Research in Education*, Vol. 1. (1973), pp. 58-94. Jstor

Week 2 Review Questions

Question 1. Background: standard analysis of covariance.(no moderating variable)

A researcher is studying the effect of an incentive on the retention of subject matter and is also interested in the role of time devoted to study.
Subjects are randomly assigned to two groups, one receiving (C3 = 1) and the other not receiving (C3 = 0) an incentive. Within these groups, subjects are randomly assigned to 5, 10, 15, or 20 minutes of study (C2) of a passage specifically prepared for the experiment. At the end of the study period, a test of retention is administered.
Treat the study time as a covariate for investigating the differential effects of the incentive. Does using the covariate improve precision in estimating the effect of incentive?
Does the ancova assumption of a constant treatment effect at levels of StudyMin appear reasonable?
full data are in file retention.dat formerly located at http://statweb.stanford.edu/~rag/stat209/retention.dat
***note*:** As of January 2022 Statistics Dept. servers eliminated--files linked at statweb.stanford.edu/~rag/stat209/[file] or www-stat.stanford.edu/~rag/stat209/[file] now reside at rag.su.domains/stat209/[file].
Linked materials resolve to rag.su.domains seemlessly but to read in data files to R requires using the new file location.
Solution for question 1

Question 2. Revisit High School and Beyond ancova from Week 2 lecture

In the class example we used school level (mean, gradient) outcomes and used school mean ses as a covariate. Investigate the usefulness of that covariate by comparing the ancova in class example with just a simple t-test (sector) on these school level outcomes. What is the difference in precision between using the covariate or not? As this is not an RCT (revisit in Unit 2), also look at differences in the estimate of the sector effect (bias?).
Solution for question 2

Question 3. Comparing Regressions (demonstration data, not an RCT)

Let's give recognition to the guys who made S (and R) and take some data from Venables, W. N. and Ripley, B. D. (1999) Modern Applied Statistics with S-PLUS. Third Edition. Springer (now up to 4th edition). Chap 6 section 1 considers analysis of the data set whiteside (available as part of MASS subset of VR package) to access
> library(MASS) # do need to load library, MASS is part of base R > data(whiteside) > ?whiteside
Description
Mr Derek Whiteside of the UK Building Research Station recorded the weekly gas consumption and average external temperature at his own house in south-east England for two heating seasons, one of 26 weeks before, and one of 30 weeks after cavity-wall insulation was installed. The object of the exercise was to assess the effect of the insulation on gas consumption.
Format The whiteside data frame has 56 rows and 3 columns.:
Insul A factor, before or after insulation.
Temp Purportedly the average outside temperature in degrees Celsius. (These values is far too low for any 56-week period in the 1960s in South-East England. It might be the weekly average of daily minima.)
Gas The weekly gas consumption in 1000s of cubic feet.
Source. A data set collected in the 1960s by Mr Derek Whiteside of the UK Building Research Station. Reported by Hand, D. J., Daly, F., McConway, K., Lunn, D. and Ostrowski, E. eds (1993) A Handbook of Small Data Sets. Chapman & Hall, p. 69.

carry out a comparing regressions analysis with Insul as the group variable, Gas as outcome, and Temp as within-group predictor.
construct a 95% confidence interval for the effect of insul on on gas with temp = 4 (pick-a-point procedure)
for what values of temp does there appear to be an effect of Insul on Gas (simultaneous region of significance)
Solution for question 3

Question 4. R packages `interactions` and `probemod`
In lecture there was short mention of these two R-packages that whose main functions are to carry out the pick-a-point and Johnson-Neyman claculations, which are developed in Rogosa(1980).
Try out these functions using the cnrl dataset (also from Rogosa,1980) which we worked out in the lecture materials.
Solutions spoiler alert: no joy from these packages.
Solution for question 4

Week 3

---

mediation

The New York Times

PHYS ED

How Exercise May Help Keep Our Memory Sharp

Irisin, a hormone that is released during exercise, may improve brain health and lessen the damage that occurs during Alzheimer's disease.

By Gretchen Reynolds

Jan. 16, 2019

A hormone that is released during exercise may improve brain health and lessen the damage and memory loss that occur during dementia, a new study finds. The study, which was published this month in Nature Medicine, involved mice, but its findings could help to explain how, at a molecular level, exercise protects our brains and possibly preserves memory and thinking skills, even in people whose pasts are fading.

Considerable scientific evidence already demonstrates that exercise remodels brains and affects thinking. Researchers have shown in rats and mice that running ramps up the creation of new brain cells in the hippocampus, a portion of the brain devoted to memory formation and storage. Exercise also can improve the health and function of the synapses between neurons there, allowing brain cells to better communicate.

In people, epidemiological research indicates that being physically active reduces the risk for Alzheimer's disease and other dementias and may also slow disease progression.

But many questions remain about just how exercise alters the inner workings of the brain and whether the effects are a result of changes elsewhere in the body that also happen to be good for the brain or whether the changes actually occur within the brain itself.

Those issues attracted the attention of an international consortium of scientists — some of them neuroscientists, others cell biologists — all of whom were focused on preventing, treating and understanding Alzheimer's disease.

Those concerns had brought a hormone called irisin into their sphere of interest. Irisin, first identified in 2012 and named for Iris, the gods' messenger in Greek mythology, is produced by muscles during exercise. The hormone jump-starts multiple biochemical reactions throughout the body, most of them related to energy metabolism.

[Read** more about irisin**. | Sign up for** the Well newsletter**.]

Because Alzheimer's disease is believed to involve, in part, changes in how brain cells use energy, the scientists reasoned that exercise might be helping to protect brains by increasing levels of irisin there.

But if so, they realized, irisin would have to exist in human brains. To see if it did, they gathered tissues from brain banks and, using sophisticated testing, found irisin there. Gene expression patterns in those tissues also suggested that much of this irisin had been created in the brain itself. Levels of the hormone were especially high in the brains of people who were free of dementia when they died, but were barely detectable in the brains of people who had died with Alzheimer's.

Those tests, however, though interesting, could not tell scientists what role irisin might be playing in brains. So the researchers now turned to mice, some healthy and others bred to develop a rodent form of Alzheimer's.

They infused the brains of the animals bred to have dementia with a concentrated dose of irisin. Those mice soon began to perform better on memory tests and show signs of improved synaptic health.

At the same time, they soaked the brains of the healthy animals with a substance that inhibits production of irisin and then pumped in a form of beta amyloid, a protein that clumps together to form plaques in the brains of those with Alzheimer's. In effect, they gave the mice dementia. And, without any irisin in their brains, the once-healthy mice soon showed signs of worsening memory and poor function in the synapses between neurons in their hippocampus.

The scientists also looked inside individual neurons from healthy mice and found that, when they added irisin to the cells, gene expression changed in ways that would be expected to lessen damage from beta amyloid.

Finally and perhaps most important, the scientists had healthy mice work out, swimming for an hour almost every day for five weeks. Beforehand, some of the animals also were treated with the substance that blocks irisin production.

In the untreated animals, irisin levels in the brain blossomed during the exercise training and later, after the animals' brains were exposed to beta amyloid, they seemed to fight off its effects, performing significantly better on memory tests than sedentary control mice that likewise had been exposed.

But the animals that had been unable to create irisin did not benefit much from exercise. After exposure to beta amyloid, they performed about as poorly on memory tests as sedentary animals with beta amyloid in their brains.

Taken as a whole, these experiments suggest that exercise may protect against dementia in part by triggering an increase in the amount of irisin in the brain, says Ottavio Arancio, a professor of pathology and cell biology at Columbia University, who conducted the research along with two dozen colleagues from the Federal University of Rio de Janeiro in Brazil, Queen's University in Canada and other institutions.

But the experiments, although elaborate and multipronged, used mice, and so cannot tell us if exercise and irisin will work similarly in people, or how much and what types of exercise might be best for brain health. The results also do not show whether exercise and irisin can prevent Alzheimer's, but only that they seem to allay some of the effects of the disease in mice once it begins.

The scientists involved in the study hope soon to test a pharmaceutical form of irisin as a treatment for dementia in animals and eventually people, especially those who have lost the ability to exercise, Dr. Arancio says.

But for now, he says, the overarching lesson of the study would seem to be that "if you can, go for a walk."

A version of this article appears in print on , on Page D4 of the New York edition with the headline: Be Kind to Your Brain. Work Out.

nature medicine

ARTICLES

https://doi.org/10.1038/s41591-018-0275-4

Exercise-linked FNDC5/irisin rescues synaptic plasticity and memory defects in Alzheimer's models

Mychael V. Lourenco[1,2,3], Rudimar L. Frozza[1,4,19], Guilherme B. de Freitas[1,5,19], Hong Zhang[3], Grasielle C. Kincheski[1,2], Felipe C. Ribeiro[1,2], Rafaella A. Gonçalves[5], Julia R. Clarke[1,6], Danielle Beckman[1], Agnieszka Staniszewski[3], Hanna Berman[3], Lorena A. Guerra[1,2], Letícia Forny-Germano[1], Shelby Meier[7], Donna M. Wilcock[7], Jorge M. de Souza[8,9], Soniza Alves-Leon[8,9], Vania F. Prado[10,11,12], Marco A. M. Prado[10,11,12], Jose F. Abisambra[7], Fernanda Tovar-Moll[13,14], Paulo Mattos[13,15], Ottavio Arancio[3,16,17]*, Sergio T. Ferreira[1,2]* and Fernanda G. De Felice[1,5,18]*

Defective brain hormonal signaling has been associated with Alzheimer's disease (AD), a disorder characterized by synapse and memory failure. Irisin is an exercise-induced myokine released on cleavage of the membrane-bound precursor protein fibronectin type III domain-containing protein 5 (FNDC5), also expressed in the hippocampus. Here we show that FNDC5/irisin levels are reduced in AD hippocampi and cerebrospinal fluid, and in experimental AD models. Knockdown of brain FNDC5/irisin impairs long-term potentiation and novel object recognition memory in mice. Conversely, boosting brain levels of FNDC5/irisin rescues synaptic plasticity and memory in AD mouse models. Peripheral overexpression of FNDC5/irisin rescues memory impairment, whereas blockade of either peripheral or brain FNDC5/irisin attenuates the neuroprotective actions of physical exercise on synaptic plasticity and memory in AD mice. By showing that FNDC5/irisin is an important mediator of the beneficial effects of exercise in AD models, our findings place FNDC5/irisin as a novel agent capable of opposing synapse failure and memory impairment in AD.

The incidence of AD, the most common form of dementia in older individuals, is increasing as the world population ages, with more than 35 million people now affected worldwide[1]. Currently, there is no effective treatment for AD[2]; notable efforts are aimed at developing strategies to counteract mechanisms leading to neuronal damage, synapse failure, and memory impairment in AD.

Consolidated evidence indicates that the central nervous system (CNS) is an important target for the actions of peripheral hormones, including insulin, leptin, glucagon-like peptide-1, glucocorticoids, and others[3–5]. Insulin, leptin, and glucagon-like peptide-1 stimulate neuronal survival and synaptic plasticity, and they contribute to higher brain functions, including cognition[6–9]. Failure of hormone-initiated signaling pathways has been associated with brain disorders, including AD[10]. For example, brain insulin signaling is decreased in AD[4,11–13], and strategies aimed at bolstering it are currently under clinical investigation[14,15].

Irisin was recently identified as a myokine released into the circulation on physical exercise that is capable of stimulating adipocyte browning and thermogenesis in mice and humans[16,17]. Irisin is cleaved from fibronectin type III domain-containing protein 5 (FNDC5), a transmembrane precursor protein expressed in muscle under the control of peroxisome proliferator-activated receptor-γ coactivator 1α (PGC-1α). FNDC5/irisin stimulates the expression of brain-derived neurotrophic factor (BDNF) in the hippocampus[18], a brain region centrally involved in learning and memory. This raises the possibility that FNDC5/irisin could play a neuroprotective role in brain disorders such as AD. In this study, we investigated FNDC5/irisin levels in the brain and cerebrospinal fluid (CSF) of patients with AD and in mouse models of AD, and we tested the hypothesis that FNDC5/irisin could be a key mediator of the beneficial effects of exercise on synaptic plasticity and memory in AD models, thus holding promise as a potential target for therapeutic intervention in AD.

[1]Institute of Medical Biochemistry Leopoldo de Meis, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil. [2]Institute of Biophysics Carlos Chagas Filho, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil. [3]Taub Institute for Research on Alzheimer's Disease and the Aging Brain, Columbia University, New York, NY, USA. [4]Oswaldo Cruz Institute, Oswaldo Cruz Foundation, FIOCRUZ, Rio de Janeiro, Brazil. [5]Centre for Neuroscience Studies, Queen's University, Kingston, Ontario, Canada. [6]School of Pharmacy, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil. [7]Sanders-Brown Center on Aging, University of Kentucky, Lexington, KY, USA. [8]Division of Neurosurgery, Clementino Fraga Filho University Hospital, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil. [9]Division of Neurology/Epilepsy Program, Clementino Fraga Filho University Hospital, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil. [10]Robarts Research Institute, University of Western Ontario, London, Ontario, Canada. [11]Department of Physiology and Pharmacology, University of Western Ontario, London, Ontario, Canada. [12]Department of Anatomy & Cell Biology, University of Western Ontario, London, Ontario, Canada. [13]D'Or Institute for Research and Education , Rio de Janeiro, Brazil. [14]Institute of Biomedical Sciences, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil. [15]Institute of Psychiatry, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil. [16]Department of Pathology & Cell Biology, Columbia University, New York, NY, USA. [17]Department of Medicine, Columbia University, New York, NY, USA. [18]Department of Psychiatry, Queen's University, Kingston, Ontario, Canada. [19]These authors contributed equally: Rudimar L. Frozza, Guilherme B. de Freitas. *e-mail: oa1@cumc.columbia.edu; ferreira@biof.ufrj.br; fernanda.defelice@queensu.ca

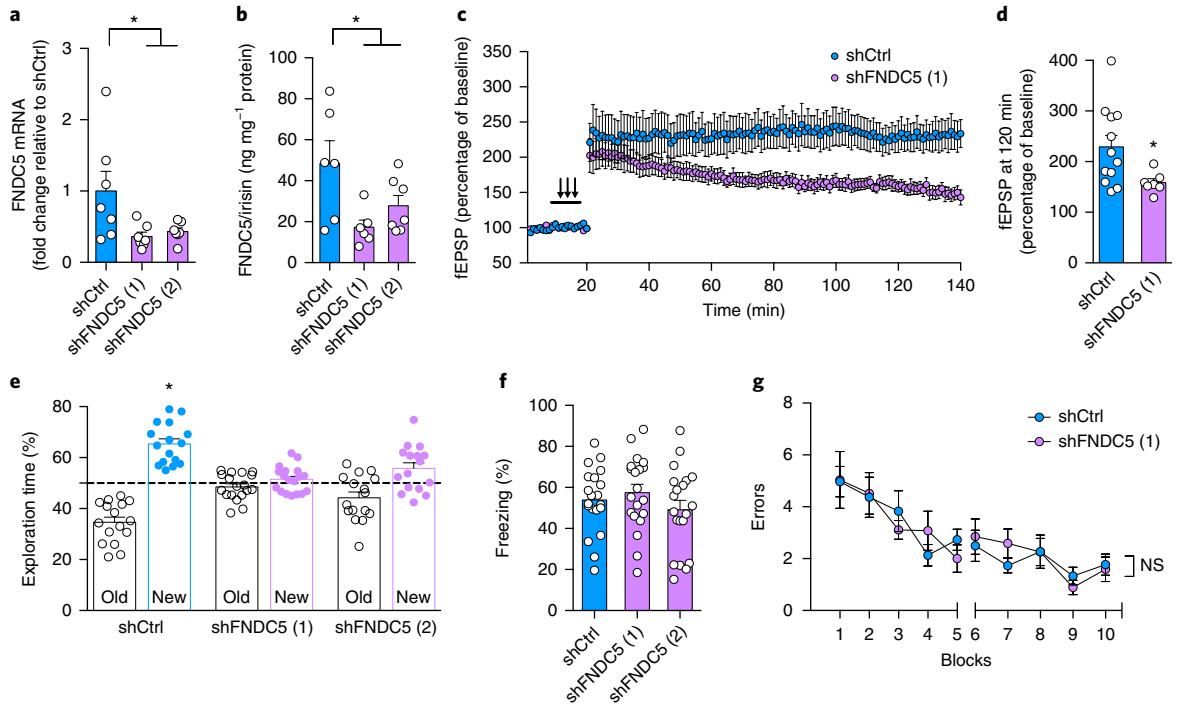


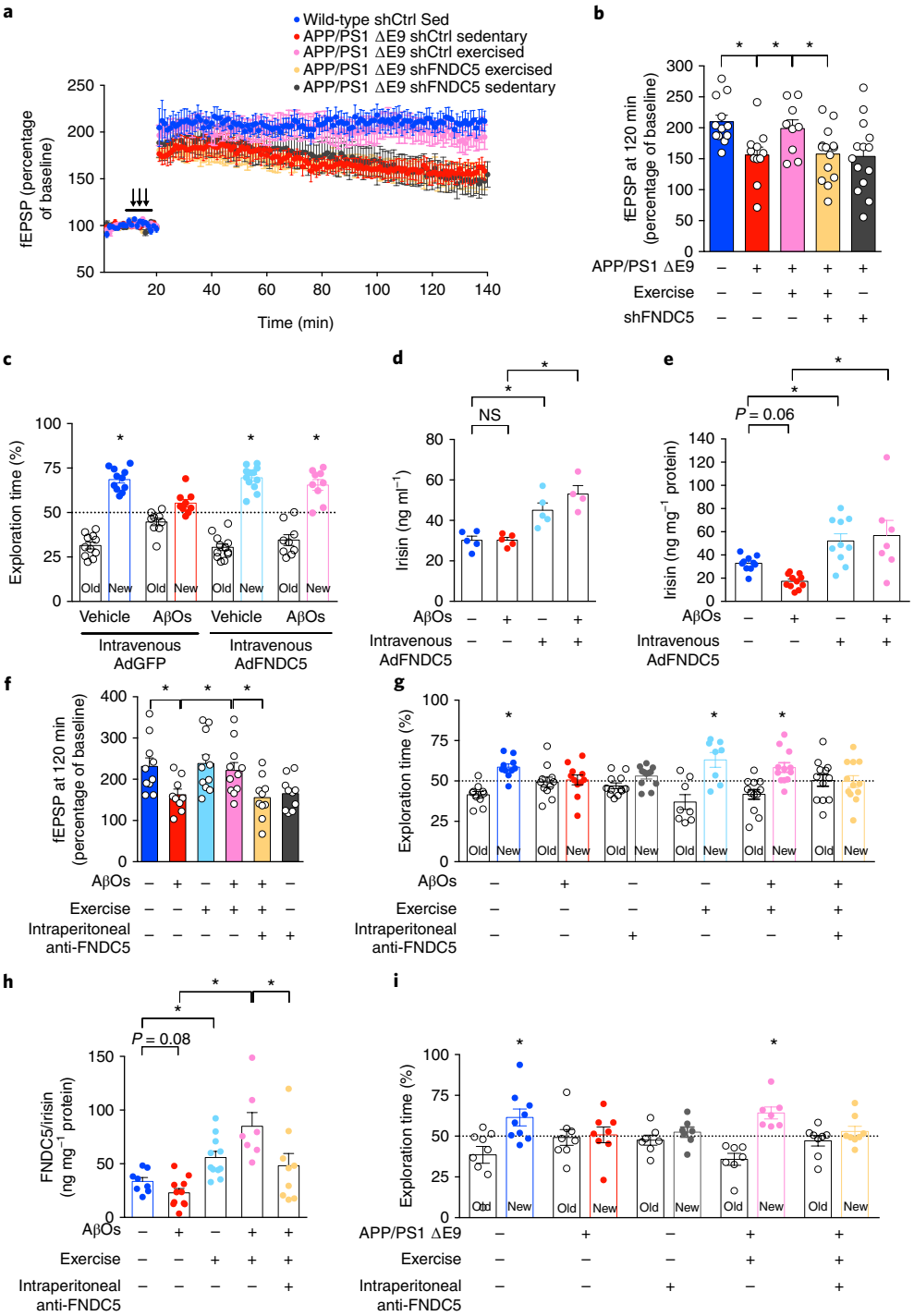
Fig. 3 | Downregulation of brain FNDC5/irisin impairs synaptic plasticity and object recognition memory in mice. Two distinct shRNAs targeting FNDC5 (shFNDC5 1 or 2) or shCtrl were injected intracerebroventricularly in C57BL/6 mice. **a**, Levels of FNDC5 mRNA compared to shFNDC5 (1)- or shFNDC5 (2)-injected mice ($N=7$ mice per group; $^*P<0.05$, one-way ANOVA with Holm–Šidák correction) in the frontal cortex. Bars express mean ± s.e.m. **b**, FNDC5/irisin protein in control (shCtrl) compared to shFNDC5 (1)- or shFNDC5 (2)-injected mice ($N=7$ mice per group; $^*P<0.05$, one-way ANOVA with Holm–Šidák correction) in the frontal cortex. Bars express mean ± s.e.m. **c**, Average traces for field excitatory postsynaptic potentials (fEPSPs) in hippocampal slices from each experimental group ($N=12$ slices for shCtrl, 7 slices for shFNDC5 (1) obtained from 3–4 mice per group). Traces represent mean ± s.e.m. per time. **d**, fEPSP at 120 min ($N=12$ slices for shCtrl, 7 slices for shFNDC5 obtained from 3–4 mice per group; $^*P<0.05$, two-sided repeated measures two-way ANOVA with Holm–Šidák correction). **e**, Summary quantification of novel object discrimination in the NOR task in shCtrl, shFNDC5 (1)- or shFNDC5 (2)-injected mice. $^*P<0.05$, statistically different from 50% (chance level) ($N=16$ mice for shCtrl, 18 for shFNDC5 (1), 16 for the shFNDC5 (2) group; one-sample Student's *t*-test). **f**, CFC in shCtrl or shFNDC5-infused C57BL/6 mice ($N=20$ mice per group; no significant difference was observed). Two-sided one-way ANOVA followed by Holm–Šidák correction. Bars express mean ± s.e.m. **g**, shCtrl or shFNDC5 (1)-infused C57BL/6 mice were assessed in a two-day RAWM task and presented similar error profiles across trials. Each block consisted of three consecutive trials. $N=9$ shCtrl, 11 shFNDC5 (1); repeated measures two-way ANOVA. Values are presented as mean ± s.e.m.

Neuroprotective actions of recombinant irisin in vitro. Because abnormal eukaryotic initiation factor 2α (eIF2α) phosphorylation and inhibition of protein synthesis have been recently described as key mechanisms driving synapse damage and memory failure in AD models[10,26,30–32], we examined the effects of recombinant irisin on phosphorylated eIF2α (eIF2α-P) and activating transcription factor 4 (ATF4) levels in cultured primary hippocampal neurons. Irisin prevented AβO-induced elevation in eIF2α-P and ATF4 (Extended Data Fig. 9a–c), as well as downregulation of de novo protein synthesis in hippocampal neurons (Extended Data Fig. 9d,e). Control measurements revealed that total eIF2α immunoreactivity remained unchanged (not shown).

We further found that recombinant irisin prevented dendritic spine loss in cultured hippocampal neurons exposed to AβOs (Extended Data Fig. 9f,g). Additional experiments determined that recombinant irisin reduced AβO binding to neurons (Extended Data Fig. 9h,i). Control binding studies revealed no direct interaction between AβOs and recombinant irisin (Extended Data Fig. 9j), ruling out the possibility that blockade of binding to neurons might be caused by sequestration of AβOs by recombinant irisin added to the medium. In addition, surface FNDC5/irisin and AβOs did not colocalize in dendrites of hippocampal neurons (Extended Data Fig. 9k). Results further showed that FNDC5/irisin overexpression reduced hippocampal soluble Aβ42 levels in APP/PS1 M146L mice (Extended Data Fig. 9l), but not insoluble Aβ42 in the hippocampus or cortex (Extended Data Fig. 9m–o).

We found that recombinant irisin stimulated the cyclic AMP (cAMP)–protein kinase A (PKA)–cAMP responsive element-binding protein (CREB) pathway in human cortical slices (Fig. 5a–c), a pathway that plays important roles in memory formation and has been found to be impaired in AD models[33–35]. Recombinant irisin further increased cAMP and phosphorylated CREB in mouse hippocampal slices (Fig. 5d,e). Irisin-induced CREB phosphorylation was abolished by PKA inhibition with myristoylated protein kinase inhibitor (PKI) 14–22, a selective PKA inhibitor (Fig. 5e). We also found that PKA activity mediated protection against nuclear translocation of ATF4 induced by AβOs (Fig. 5f,g). Irisin further induced transient phosphorylation of extracellular signal-regulated kinase in cultured neurons (data not shown). The effect of recombinant irisin was similar to forskolin, a direct activator of adenylyl cyclase (data not shown). Taken together, these results provide initial clues into the mechanisms by which recombinant irisin affords neuroprotection in experimental models of AD.

FNDC5/irisin mediates the protective actions of physical exercise on synaptic plasticity and memory in AD models. From a translational perspective, physical exercise could be a non-pharmacological strategy to increase hippocampal FNDC5/irisin in patients at



defects as disease progresses. Additional studies measuring irisin levels in the CSF during healthy aging and in patients with other neurological disorders are anticipated and will help to determine when irisin levels decrease during the course of AD.

Blockade of either brain or peripheral FNDC5/irisin in mice impaired LTP and NOR memory, implicating FNDC5/irisin in physiological memory processes. However, future studies are warranted to address the precise physiological roles of brain and peripheral FNDC5/irisin in the formation and consolidation of different types of memories.

It is noteworthy that peripheral administration of AdFNDC5 led to increases in hippocampal FNDC5/irisin and protected mice against memory impairment induced by AβOs. Most importantly, peripheral FNDC5/irisin is implicated in the preservation of hippocampal FNDC5/irisin levels, synaptic plasticity, and memory in AD mice. Collectively, these results provide mechanistic information on the beneficial actions of FNDC5/irisin in the brain and suggest that a cross talk between peripheral and central FNDC5/irisin influences synaptic plasticity and memory in mice.

Physical exercise has been previously shown to induce memory-related events in the brain[43–45] and has been proposed as an approach to reduce the risk of AD, potentially bringing about significant benefits to subjects with MCI and early AD[46–49]. Many efforts to identify the endogenous molecules responsible for the beneficial effects of exercise are underway. Brain PGC-1α and BDNF[28,50], as well as peripheral cathepsin B and β-hydroxybutyrate[51,52], have been

described as important molecules acting as mediators of exercise-induced neuroprotection. The results presented in this study add FNDC5/irisin to this list.

Our findings suggest that FNDC5/irisin could comprise an attractive novel therapy aimed to prevent dementia in patients at risk, as well as delay its progression in patients at the later stages, including those who can no longer exercise. Many patients with dementia are disabled due to other age-related conditions or comorbidities (for example, arthritis, heart disease, obesity, visual problems, depression) that preclude them from engaging in regular physical exercise. Therefore, the development of alternative approaches that build on the beneficial effects of exercise in the brain may benefit those patients.

In conclusion, our results demonstrate that FNDC5/irisin levels are reduced in human AD brains and CSF and in AD mouse models, and that boosting either brain or peripheral FNDC5/irisin levels attenuates synaptic and memory impairments in AD mouse models. We further show that FNDC5/irisin is a novel mediator of the beneficial effects of exercise on synapse function and memory in AD models. Bolstering brain FNDC5/irisin levels, either pharmacologically or through exercise, may thus constitute a novel therapeutic strategy to protect and/or repair synapse function and prevent cognitive decline in AD.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41591-018-0275-4.

Received: 13 October 2016; Accepted: 2 November 2018;
Published online: 7 January 2019

References

1. Prince, M. et al. The global prevalence of dementia: a systematic review and metaanalysis. *Alzheimers Dement.* **9**, 63–75.e62 (2013).
2. Ferreira, S. T., Lourenco, M. V., Oliveira, M. M. & De Felice, F. G. Soluble amyloid-β oligomers as synaptotoxins leading to cognitive impairment in Alzheimer's disease. *Front. Cell. Neurosci.* **9**, 191 (2015).
3. Fernandez, A. M. & Torres-Alemán, I. The many faces of insulin-like peptide signalling in the brain. *Nat. Rev. Neurosci.* **13**, 225–239 (2012).
4. Biessels, G. J. & Reagan, L. P. Hippocampal insulin resistance and cognitive dysfunction. *Nat. Rev. Neurosci.* **16**, 660–671 (2015).
5. McEwen, B. S. Preserving neuroplasticity: role of glucocorticoids and neurotrophins via phosphorylation. *Proc. Natl Acad. Sci. USA* **112**, 15544–15545 (2015).
6. During, M. J. et al. Glucagon-like peptide-1 receptor is involved in learning and neuroprotection. *Nat. Med.* **9**, 1173–1179 (2003).
7. Chiu, S.-L., Chen, C.-M. & Cline, H. T. Insulin receptor signaling regulates synapse number, dendritic plasticity, and circuit function in vivo. *Neuron* **58**, 708–719 (2008).
8. Grillo, C. A. et al. Hippocampal insulin resistance impairs spatial learning and synaptic plasticity. *Diabetes* **64**, 3927–3936 (2015).
9. Irving, A. J. & Harvey, J. Leptin regulation of hippocampal synaptic function in health and disease. *Philos. Trans. R. Soc. Lond. B* **369**, 20130155 (2013).
10. Lourenco, M. V., Ferreira, S. T. & De Felice, F. G. Neuronal stress signaling and eIF2α phosphorylation as molecular links between Alzheimer's disease and diabetes. *Prog. Neurobiol.* **129**, 37–57 (2015).
11. Bomfim, T. R. et al. An anti-diabetes agent protects the mouse brain from defective insulin signaling caused by Alzheimer's disease-associated Aβ oligomers. *J. Clin. Invest.* **122**, 1339–1353 (2012).
12. Talbot, K. et al. Demonstrated brain insulin resistance in Alzheimer's disease patients is associated with IGF-1 resistance, IRS-1 dysregulation, and cognitive decline. *J. Clin. Invest.* **122**, 1316–1338 (2012).
13. De Felice, F. G., Lourenco, M. V. & Ferreira, S. T. How does brain insulin resistance develop in Alzheimer's disease? *Alzheimers Dement.* **10**, S26–S32 (2014).
14. Wadman, M. US government sets out Alzheimer's plan. *Nature* **485**, 426–427 (2012).
15. De Felice, F. G. Alzheimer's disease and insulin resistance: translating basic science into clinical applications. *J. Clin. Invest.* **123**, 531–539 (2013).
16. Boström, P. et al. A PGC1-α-dependent myokine that drives brown-fat-like development of white fat and thermogenesis. *Nature* **481**, 463–468 (2012).
17. Jedrychowski, M. P. et al. Detection and quantitation of circulating human irisin by tandem mass spectrometry. *Cell Metab.* **22**, 734–740 (2015).
18. Wrann, C.D. et al. Exercise induces hippocampal BDNF through a PGC-1α/FNDC5 pathway. *Cell Metab.* **18**, 649–659 (2013).
19. Chen, K. et al. Irisin protects mitochondria function during pulmonary ischemia/reperfusion injury. *Sci. Transl. Med.* **9**, eaao6298 (2017).
20. Lee, P. et al. Irisin and FGF21 are cold-induced endocrine activators of brown fat function in humans. *Cell Metab.* **19**, 302–309 (2014).
21. Schumacher, M. A., Chinnam, N., Ohashi, T., Shah, R. S. & Erickson, H. P. The structure of irisin reveals a novel intersubunit β-sheet fibronectin type III (FNIII) dimer: implications for receptor activation. *J. Biol. Chem.* **288**, 33738–33744 (2013).
22. Mucke, L. & Selkoe, D. J. Neurotoxicity of amyloid β-protein: synaptic and network dysfunction. *Cold Spring Harb. Perspect. Med.* **2**, a0063381 (2012).
23. Sebollela, A. et al. Amyloid-β oligomers induce differential gene expression in adult human brain slices. *J. Biol. Chem.* **287**, 7436–7445 (2012).
24. Colaianni, G. et al. The myokine irisin increases cortical bone mass. *Proc. Natl Acad. Sci. USA* **112**, 12157–12162 (2015).
25. Figueiredo, C. P. et al. Memantine rescues transient cognitive impairment caused by high-molecular-weight Aβ oligomers but not the persistent impairment induced by low-molecular-weight oligomers. *J. Neurosci.* **33**, 9626–9634 (2013).
26. Lourenco, M. V. et al. TNF-α mediates PKR-dependent memory impairment and brain IRS-1 inhibition induced by Alzheimer's β-amyloid oligomers in mice and monkeys. *Cell Metab.* **18**, 831–843 (2013).
27. Jankowsky, J. L. et al. Co-expression of multiple transgenes in mouse CNS: a comparison of strategies. *Biomol. Eng.* **17**, 157–165 (2001).
28. Cheng, A. et al. Involvement of PGC-1α in the formation and maintenance of neuronal dendritic spines. *Nat. Commun.* **3**, 1250 (2012).
29. Holcomb, L. et al. Accelerated Alzheimer-type phenotype in transgenic mice carrying both mutant amyloid precursor protein and presenilin 1 transgenes. *Nat. Med.* **4**, 97–100 (1998).
30. Ma, T. et al. Suppression of eIF2α kinases alleviates Alzheimer's disease-related plasticity and memory deficits. *Nat. Neurosci.* **16**, 1299–1305 (2013).
31. Yang, W. et al. Repression of the eIF2α kinase PERK alleviates mGluR-LTD impairments in a mouse model of Alzheimer's disease. *Neurobiol. Aging* **41**, 19–24 (2016).
32. Trinh, M. A. & Klann, E. Translational control by eIF2α kinases in long-lasting synaptic plasticity and long-term memory. *Neurobiol. Learn. Mem.* **105**, 93–99 (2013).
33. Gong, B. et al. Persistent improvement in synaptic and cognitive functions in an Alzheimer mouse model after rolipram treatment. *J. Clin. Invest.* **114**, 1624–1634 (2004).
34. Vitolo, O. V. et al. Amyloid β-peptide inhibition of the PKA/CREB pathway and long-term potentiation: reversibility by drugs that enhance cAMP signaling. *Proc. Natl Acad. Sci. USA* **99**, 13217–13221 (2002).
35. Schaefer, N. et al. The malleable brain: plasticity of neural circuits and behavior: a review from students to students. *J. Neurochem.* **142**, 790–811 (2017).
36. Katsnelson, A., De Strooper, B. & Zoghbi, H. Y. Neurodegeneration: from cellular concepts to clinical applications. *Sci. Transl. Med.* **8**, 364ps318 (2016).
37. Selkoe, D. J. Alzheimer's disease is a synaptic failure. *Science* **298**, 789–791 (2002).
38. Lepeta, K. et al. Synaptopathies: synaptic dysfunction in neurological disorders: a review from students to students. *J. Neurochem.* **138**, 785–805 (2016).
39. Zhang, Y. et al. Irisin stimulates browning of white adipocytes through mitogen-activated protein kinase p38 MAP kinase and ERK MAP kinase signaling. *Diabetes* **63**, 514–525 (2014).
40. Timmons, J. A., Baar, K., Davidsen, P. K. & Atherton, P. J. Is irisin a human exercise gene? *Nature* **488**, E9–E10 (2012).
41. Albrecht, E. et al. Irisin: a myth rather than an exercise-inducible myokine. *Sci. Rep.* **5**, 8889 (2015).
42. Martin, K. C. & Kandel, E. R. Cell adhesion molecules, CREB, and the formation of new synaptic connections. *Neuron* **17**, 567–570 (1996).
43. Suwabe, K. et al. Rapid stimulation of human dentate gyrus function with acute mild exercise. *Proc. Natl Acad. Sci. USA* **115**, 10487–10492 (2018).
44. van Praag, H., Fleshner, M., Schwartz, M. W. & Mattson, M. P. Exercise, energy intake, glucose homeostasis, and the brain. *J. Neurosci.* **34**, 15139–15149 (2014).
45. Neufer, P. D. et al. Understanding the cellular and molecular mechanisms of physical activity-induced health benefits. *Cell Metab.* **22**, 4–11 (2015).
46. Baker, L. D. et al. Effects of aerobic exercise on mild cognitive impairment: a controlled trial. *Arch. Neurol.* **67**, 71–79 (2010).
47. Buchman, A. S. et al. Total daily physical activity and the risk of AD and cognitive decline in older adults. *Neurology* **78**, 1323–1329 (2012).

Question 1. Mediating Variable Computations: Class example continued
The data set shown in class example ss423 is linked above and in the legacy directory http://web.stanford.edu/~rag/stat209/ss423
for predictor (IV) 'belong' outcome 'depress' and (potential) mediating variable 'master' The class example showed you the Baron-Kenny analysis using functions from the multilevel and MBESS packages.
Here just use 'lm' basic regression and the recipees from the class handout to recreate point estimates and asymptotic standard errors, significance tests for the mediating variable effect.
Compare your result with the class example posting.
Extra: also try out the more 'sophisticated' functions in the mediation package.

Solution for question 1

Question 2. Potential Outcomes, Encouragement Design Estimation and (Causal) Mediation
Task 1. Create a potential Outcomes dataset following the first ALICE specification in the posted slides (week 3) ## ALICE example beta = 3 rho = 3 tau = 1, delta = 3 (I did n=400; larger would be better so I redid with n = 6400)

Task 2. Use the artificial data to show the results for the mediation (indirect) effect by hand doing the 3 regressions using multilevel package (sobel) using MBESS package using the causal mediation estimation ACME from the mediation package and compare with rho*beta

Task 3 estimate beta by the Wald estimator (assuming tau = 0) and estimate mediation effect

Solution for question 2

Question 3. Sesame Street: Encouragement Design research example
Sesame Street research setting and data description given pdf p.30 of Lecture 1 (also Gelman text).
For this exercise use `postnumb` : posttest on numbers (0-54), along with the measures encour and regular from the class example in Lecture 1.
Use the encouragement design formulation to estimate the effect on child cognitive development (postnumb here) of watching more Sesame Street.
What assumption is necessary for the IV estimation in this design?
Obtain a point and interval estimate for the effect of viewing (use `ivreg` as in class example).
From simple descriptives reproduce this instrumental variables estimate (Wald estimator).
The second approach (path analysis) analyzed by Holland requires what assumption?
Obtain the path analyses (regression) estimate for the effect on child cognitive development (postnumb here) of watching more Sesame Street.
Compare with the IV estimate (which employs different assumptions).

Solution for question 3

Week 2

Moderating Variables in experimental studies (heterogeneous treatment effects)

~~To be posted~~
Lecture slides, week 2 (pdf)
Audio companion, week 2
parta partb partc

Lecture topics
0. Moderation, mediation recap slide
1. Review: formulation and purposes of analysis of covariance
basic (old) ancova exposition slides ancova and extensions, math notes
High School and Beyond (observational study) school means data example HSB ancova handout (ascii version) data for HSB ancova HSB ancova, scanned pdf
2. Moderating variables, Heterogeneous Treatment Effects (CATE).
Analyzing treatment effects as a function of covariate(s)
CNRL, including Johnson-Neyman technique cnrl data cnrl analysis (extended)

Primary Readings
Ancova and extensions
Rogosa, D. R. (1980). Comparing nonparallel regression lines. *Psychological Bulletin*, 88, 307-321. [a better quality scan from the APA site]
R resources (below).

Moderation research examples:
Gender differences in effectiveness of aspirin.
Aspirin may be less effective heart treatment for women than men
Publication: Aspirin Resistance in Patients with Stable Coronary Artery Disease, in the *Annals of Pharmacotherapy* April 2007
Moderating variables can be your friend (statistics is the only friend you need) music: I've got friends in low places
Wash Post: Why smart people are better off with fewer friends .
Publication: Country roads, take me home... to my friends: How intelligence, population density, and friendship affect modern happiness. *British Journal of Psychology* 2016
ATI research
Snow R.E. (1978) Aptitude-Treatment Interactions in Educational Research. In: Pervin L.A., Lewis M. (eds) Perspectives in Interactional Psychology. Springer, Boston, MA. https://doi.org/10.1007/978-1-4613-3997-7_10
Family SES as a moderating variable in nature/nuture:
Why Rich Parents Don't Matter UTexas press release: Being Poor Can Suppress Children's Genetic Potentials Publication: Emergence of a Gene x Socioeconomic Status Interaction on Infant Mental Ability Between 10 Months and 2 years DOI: 10.1177/0956797610392926 Psychological Science published online 17 December 2010 Elliot M. Tucker-Drob, Mijke Rhemtulla, K. Paige Harden, Eric Turkheimer and David Fask

R implementations and Resources
package `probemod` manual
package `interactions` intro vignette: Exploring interactions with continuous predictors in regression models manual

Additional Resources, Ancova and extensions
Improving Present Practices in the Visual Display of Interactions Advances in Methods and Practices in Psychological Science
analysis of covariance: Background/historical papers:
Covariance Adjustment in Randomized Experiments and Observational Studies Paul R. Rosenbaum *Statistical Science*, Vol. 17, No. 3. (Aug., 2002), pp. 286-304. Jstor
Some Aspects of Analysis of Covariance, A Biometrics Invited Paper with Discussion. D. R. Cox; P. McCullagh *Biometrics*, Vol. 38, No. 3, (Sep., 1982), pp. 541-561. Jstor
Analysis of Covariance: Its Nature and Uses William G. Cochran *Biometrics*, Vol. 13, No. 3, Special Issue on the Analysis of Covariance. (Sep., 1957), pp. 261-281. Jstor
The Use of Covariance in Observational Studies W. G. Cochran *Applied Statistics*, Vol. 18, No. 3. (1969), pp. 270-275. Jstor
Estimation of the Slope and Analysis of Covariance when the Concomitant Variable is Measured with Error James S. Degracie; Wayne A. Fuller *Journal of the American Statistical Association*, Vol. 67, No. 340. (Dec., 1972), pp. 930-937. Jstor
Deep background Neter-Wasserman text (Applied linear statistical models. Neter, Kutner, Nachtsheim and Wasserman 1996. Fifth edition. Homewood IL: Irwin, Inc.) chapters 22 and 8.
Johnson-Neyman technique and aptitude-treatment interaction (ATI)
Cronbach, L. J., & Snow, R. E. (1977). Aptitudes and instructional methods: A handbook for research on interactions. Irvington
Regions of Significant Criterion Differences in Aptitude-Treatment-Interaction Research Leonard S. Cahen; Robert L. Linn *American Educational Research Journal*, Vol. 8, No. 3. (May, 1971),

pp. 521-530. Jstor
Identifying Regions of Significance in Aptitude-by-Treatment-Interaction Research Ronald C. Serlin; Joel R. Levin *American Educational Research Journa*l, Vol. 17, No. 3. (Autumn, 1980), pp. 389-399. Jstor
Defining Johnson-Neyman Regions of Significance in the Three-Covariate ANCOVA Using Mathematica Steve Hunka; Jacqueline Leighton *Journal of Educational and Behavioral Statistics*, Vol. 22, No. 4. (Winter, 1997), pp. 361-387. Jstor
discussion of substantive issues: Trait-Treatment Interaction and Learning David C. Berliner; Leonard S. Cahen *Review of Research in Education*, Vol. 1. (1973), pp. 58-94. Jstor

Week 2 Review Questions

Question 1. Background: standard analysis of covariance (no moderating variable)

A researcher is studying the effect of an incentive on the retention of subject matter and is also interested in the role of time devoted to study.
Subjects are randomly assigned to two groups, one receiving (C3 = 1) and the other not receiving (C3 = 0) an incentive. Within these groups, subjects are randomly assigned to 5, 10, 15, or 20 minutes of study (C2) of a passage specifically prepared for the experiment. At the end of the study period, a test of retention is administered.
Treat the study time as a covariate for investigating the differential effects of the incentive. Does using the covariate improve precision in estimating the effect of incentive?
Does the ancova assumption of a constant treatment effect at levels of StudyMin appear reasonable?
full data are in file retention.dat formerly located at http://statweb.stanford.edu/~rag/stat209/retention.dat
***note*:** As of January 2022 Statistics Dept. servers eliminated--files linked at statweb.stanford.edu/~rag/stat209/[file] or www-stat.stanford.edu/~rag/stat209/[file] now reside at rag.su.domains/stat209/[file].
Linked materials resolve to rag.su.domains seemlessly but to read in data files to R requires using the new file location.

Solution for question 1

Question 2. Revisit High School and Beyond ancova from Week 2 lecture

In the class example we used school level (mean, gradient) outcomes and used school mean ses as a covariate. Investigate the usefulness of that covariate by comparing the ancova in class example with just a simple t-test (sector) on these school level outcomes. What is the difference in precision between using the covariate or not? As this is not an RCT (revisit in Unit 2), also look at differences in the estimate of the sector effect (bias?).

Solution for question 2

Question 3. Comparing Regressions (demonstration data, not an RCT)

Let's give recognition to the guys who made S (and R) and take some data from Venables, W. N. and Ripley, B. D. (1999) Modern Applied Statistics with S-PLUS. Third Edition. Springer (now up to 4th edition). Chap 6 section 1 considers analysis of the data set whiteside (available as part of MASS subset of VR package) to access
> library(MASS) # do need to load library, MASS is part of base R > data(whiteside) > ?whiteside
Description
Mr Derek Whiteside of the UK Building Research Station recorded the weekly gas consumption and average external temperature at his own house in south-east England for two heating seasons, one of 26 weeks before, and one of 30 weeks after cavity-wall insulation was installed. The object of the exercise was to assess the effect of the insulation on gas consumption.
Format The whiteside data frame has 56 rows and 3 columns.:
Insul A factor, before or after insulation.
Temp Purportedly the average outside temperature in degrees Celsius. (These values is far too low for any 56-week period in the 1960s in South-East England. It might be the weekly average of daily minima.)
Gas The weekly gas consumption in 1000s of cubic feet.
Source. A data set collected in the 1960s by Mr Derek Whiteside of the UK Building Research Station. Reported by Hand, D. J., Daly, F., McConway, K., Lunn, D. and Ostrowski, E. eds (1993) A Handbook of Small Data Sets. Chapman & Hall, p. 69.

carry out a comparing regressions analysis with Insul as the group variable, Gas as outcome, and Temp as within-group predictor.
construct a 95% confidence interval for the effect of insul on on gas with temp = 4 (pick-a-point procedure)
for what values of temp does there appear to be an effect of Insul on Gas (simultaneous region of significance)

Solution for question 3

Question 4. R packages `interactions` and `probemod`
In lecture there was short mention of these two R-packages that whose main functions are to carry out the pick-a-point and Johnson-Neyman claculations, which are developed in Rogosa(1980).
Try out these functions using the cnrl dataset (also from Rogosa,1980) which we worked out in the lecture materials.
Solutions spoiler alert: no joy from these packages.

Solution for question 4

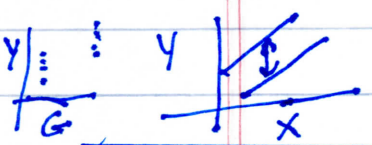
Week 3

THIS PAGE LEFT BLANK

FOR PAGINATION

Recap, RCT, G = T/C (0,1) Y outcome

Before ... ATE via t-test Y ~ G est $\mu_1 - \mu_0$
[stat 60] (ACE) or using X (concomitant var)
ancova (or blocking) Y ~ G + X
for precision: coef G est $\mu_1 - \mu_0$



[next page ATE]

But there's more ...

mediation WHY ATE?
via path analysis ?? Baron Kenny mess
week ~~2,3~~ even in RCT

moderation CATE (cnvrl)
individual diffs in response to intervention, heterogeneous treat
via pick-a-point (subgroups)
~~week 5~~ J-N region of significance
[ATE says all X in region, or no X]

---

even more ...

> > My name is Patrick Forscher, and I am the lab instructor for a data analysis course in the UW-Madison Psychology Department. The instructor of record, Markus Brauer, and I are teaching the course in R. We have used your the mediate() function in the mediation package to demonstrate how to test for simple mediation (a la Baron & Kenny, 1986). However, we'd also like to teach the students to test for moderated mediation (when a mediation effect varies across levels of a third variable; Preacher, Rucker, & Hayes, 2007) and mediated moderation (when a variable provides the causal mechanism through which an interaction exerts its effect on a dependent variable; Muller, Judd, & Yzerbyt, 2005). Is the mediate() function able to test for moderated mediation or mediated moderation? If not, would you be able to recommend a package that can test for these cases?
> >
> > Thanks for your time and advice!

indiv diffs in mediation [M/F in framing anxiety]

why indiv diffs in effect? [aspiring M/F]

• Two parameters are of primary interest. The **average treatment effect** (**ATE**) is

$$\tau_{ate} = E[Y(1) - Y(0)]. \tag{2}$$

The expected gain for a randomly selected unit from the population. This is sometimes called the *average causal effect*.

• The **average treatment effect on the treated** (**ATT**) is the average gain from treatment for those who actually were treated:

$$\tau_{att} = E[Y(1) - Y(0)|W = 1] \tag{3}$$

Chapter 14: Mediation and Moderation

Code

Alyssa Blair

1 What are Mediation and Moderation?

Mediation analysis tests a hypothetical causal chain where one variable X affects a second variable M and, in turn, that variable affects a third variable Y. Mediators describe the how or why of a (typically well-established) relationship between two other variables and are sometimes called intermediary variables since they often

describe the process through which an effect occurs. This is also sometimes called an indirect effect. For instance, people with higher incomes tend to live longer but this effect is explained by the mediating influence of having access to better health care.

In R, this kind of analysis may be conducted in two ways: Baron & Kenny's (1986) 4-step indirect effect method and the more recent *mediation* package (Tingley, Yamamoto, Hirose, Keele, & Imai, 2014). The Baron & Kelly method is among the original methods for testing for mediation but tends to have low statistical power. It is covered in this chapter because it provides a very clear approach to establishing relationships between variables and is still occassionally requested by reviewers. However, the *mediation* package method is highly recommended as a more flexible and statistically powerful approach.

Moderation analysis also allows you to test for the influence of a third variable, Z, on the relationship between variables X and Y. Rather than testing a causal link between these other variables, moderation tests for when or under what conditions an effect occurs. Moderators can stength, weaken, or reverse the nature of a relationship. For example, academic self-efficacy (confidence in own's ability to do well in school) moderates the relationship between task importance and the amount of test anxiety a student feels (Nie, Lau, & Liau, 2011). Specifically, students with high self-efficacy experience less anxiety on important tests than students with low self-efficacy while all students feel relatively low anxiety for less important tests. Self-efficacy is considered a moderator in this case because it interacts with task importance, creating a different effect on test anxiety at different levels of task importance.

In general (and thus in R), moderation can be tested by interacting variables of interest (moderator with IV) and plotting the simple slopes of the interaction, if present. A variety of packages also include functions for testing moderation but as the underlying statistical approaches are the same, only the "by hand" approach is covered in detail in here.

Finally, this chapter will cover these basic mediation and moderation techniques only. For more complicated techniques, such as multiple mediation, moderated mediation, or mediated moderation please see the *mediation* package's full documentation.

1.1 Getting Started

If necessary, review the Chapter on regression. Regression test assumptions may be tested with *gvlma*. You may load all the libraries below or load them as you go along. Review the help section of any packages you may be unfamiliar with ? (packagename).

PHYSORG.COM
SCIENCE : PHYSICS : TECH : NANO : NEWS

Published: 7 hours ago, 16:41 EST, **April 27, 2007**

Aspirin may be less effective heart treatment for women than men

Interaction--it depends

A new study shows that aspirin therapy for coronary artery disease is four times more likely to be ineffective in women compared to men with the same medical history.

Historically, studies have shown that aspirin therapy is less effective in women than in men, but it has remained unclear how much less effective and whether this affects patient outcomes, said Michael Dorsch, clinical pharmacist and adjunct clinical instructor at the University of Michigan College of Pharmacy.

Dorsch is the lead author of the paper, "Aspirin Resistance in Patients with Stable Coronary Artery Disease," which appears online today in the Annals of Pharmacotherapy.

Originally, Dorsch and his team set out to determine if patients with a history of heart attacks were more apt to be aspirin resistar than those with coronary artery disease but no history of heart attack. They found that gender and not medical history was a predictor for aspirin resistance, Dorsch said. The results surprised him.

"I was surprised by how big of a difference it was for females," said Dorsch, who has appointments at the U-M Health System an the U-M College of Pharmacy, and started the study as a resident at the University of North Carolina. "This is another piece of information that affirms we need more studies in women."

Aspirin therapy is a cornerstone in managing heart disease because it inhibits blood clotting. Aspirin therapy can reduce the risk of a nonfatal heart attack or stroke by about 23 percent, and an estimated 20 million men and women take a low dose of aspirin (81-325 mg daily) to control heart disease. But despite its effectiveness, there is evidence that aspirin is less effective in some patients, and researchers don't really know why. This can be frightening because most doctors do not check for aspirin resistanc before prescribing aspirin therapy and therefore presume it's working in the patient when it may not be, he said.

There isn't enough evidence to show if people who are aspirin resistant can simply take larger doses, but Dorsch warns that people taking aspirin on the advice of a doctor shouldn't stop therapy on account of these results.

Not only did the study quantify how much more effective aspirin therapy is for men than for women, it is also the first study that Dorsch knows of to measure aspirin resistance in men and women with stable coronary artery disease. Previous studies have looked at the impact of aspirin therapy on people who have had a heart attack.

For the study, researchers randomly selected 100 patients who were visiting their cardiologist for a regularly scheduled appointment. All had coronary artery disease but only half had a history of heart attack. Researchers used a device called VerifyNow Aspirin Assay to test the percentage of platelet reactivity after blood samples were exposed to a chemical that causes clotting.

Aspirin works by causing platelet inhibition in the blood, which means that platelets cannot stick together and this slows the formation of blood clots that cause a heart attack or stroke.

"This does happen in women, but it doesn't happen in as many women and it's not as effective," Dorsch said. The testing device uses an optical sensor to "see" what percentage of the platelets in the blood sample clump together. Anything less than 40 percent platelet inhibition is considered aspirin resistant.

"We really don't know the mechanism," Dorsch said. "It could be that women have a more active platelet system in the body so it less likely that platelet action would be inhibited."

In the future, researchers hope to look at aspirin therapy outcomes in women only and see if those outcomes can be changed. The majority of testing for aspirin therapy has been on men, so not much is known about how women respond.

"Heart disease is the number one killer of women in the United States. Future research should be aimed at finding out the cause of this increase in aspirin resistance and the effect on outcomes in women with heart disease." Dorsch said.

Cardiology

Aspirin Resistance in Patients with Stable Coronary Artery Disease with and without a History of Myocardial Infarction

Michael P Dorsch, Jin Sun Lee, Donald R Lynch, Steven P Dunn, Jo E Rodgers, Todd Schwartz, Emily Colby, Debbie Montague, and Susan S Smyth

BACKGROUND: Aspirin therapy is a cornerstone in the prevention of atherothrombotic events, but recurrent vascular events are estimated to occur in 8–18% of patients taking aspirin for secondary prevention after 2 years. Estimates of biologic aspirin resistance vary from 5% to 60%, depending on the assay used. However, the relationship between biologic measurements of aspirin resistance and adverse clinical events remains unclear.

OBJECTIVE: To determine whether patients with documented myocardial infarction (MI) while on aspirin therapy (cases) were more likely to be aspirin resistant than were patients with coronary artery disease (CAD) who had no history of MI (controls) and to assess clinical predictors of aspirin resistance in patients with stable CAD.

METHODS: This case–control study examined aspirin responses using the VerifyNow Aspirin Assay system in 50 cases and 50 controls who had taken a dose of aspirin within 48 hours of presentation to the clinic visit. Odds ratios were estimated to determine the association between aspirin resistance and MI. Independent predictors of aspirin resistance were determined using univariate and multivariate analyses.

RESULTS: An increase in the prevalence of aspirin resistance among cases (16% vs 12% in controls) was not observed (OR 1.40; 95% CI 0.45 to 4.37; p = 0.566). In the overall CAD population, female sex was independently associated with aspirin resistance (OR 4.01; 95% CI 1.15 to 13.92; p = 0.029).

CONCLUSIONS: Additional large studies are required to understand whether biologically defined aspirin resistance is associated with increased risk for cardiovascular events, with special attention paid to sex differences.

KEY WORDS: aspirin resistance, coronary artery disease, myocardial infarction.

Ann Pharmacother 2007;41:xxxx.

Published Online, 24 Apr 2007, *www.theannals.com*, DOI 10.1345/aph.1H621

Aspirin therapy is a cornerstone in the management of acute coronary syndromes and in the prevention of atherothrombotic events due to its ability to block arachidonic acid metabolism by irreversibly inhibiting cyclooxygenase-1 and reducing the production of thromboxane A_2, a platelet activator and vasoconstrictor. Aspirin reduces the risk of nonfatal myocardial infarction (MI), nonfatal stroke, and vascular death in a broad range of high-risk patients by approximately 23%.[1]

Nonetheless, some individuals experience thrombotic events while taking aspirin, and recurrent vascular events are estimated to occur in 8–18% of patients taking aspirin for secondary prevention after 2 years.[2] Measurements of platelet function in individuals taking aspirin revealed heterogeneous effects of the drug; in some individuals, aspirin fails to inhibit arachidonic acid–mediated platelet aggregation. These observations have led to the concept of aspirin resistance, which has been defined as the occurrence of atherothrombotic events despite adequate aspirin therapy (clinical resistance) or the failure of aspirin to inhibit platelet function as measured by a variety of assays (biologic/pharmacodynamic resistance). Estimates of clinical aspirin resistance, as extrapolated from event rates among those on aspirin therapy, are approximately 13%, while estimates of biologic aspirin resistance range from 5% to 60% depending on the population studied and the assay used.[3]

The relationship between biologic measurements of aspirin resistance and adverse clinical events is not clear, in part because assays of platelet function tend to be technically demanding or unreliable. In this report, we used the VerifyNow Aspirin Assay to determine whether patients with documented prior MI while on aspirin therapy were more likely to be aspirin resistant than a cohort of patients with coronary artery disease (CAD) who had no history of prior MI.

Author information provided at the end of the text.

vention that included women, a 42% reduction in risk of MI was observed in men taking 75 mg of aspirin; however, no significant reduction occurred in women.[10]

Together, these findings suggest that women do not receive clinical benefits from aspirin therapy in terms of primary protection from MI that are similar to those of men. Differences in inflammatory state, vessel size, platelet reactivity, or other factors may contribute to a dissimilar response to aspirin therapy between the sexes. In a study of nearly 1300 healthy volunteers, low-dose aspirin (81 mg daily) reduced platelet function in both men and women, with a slightly greater decrease in arachidonic acid–induced platelet aggregation occurring in females.[11] This finding is consistent with the observation that aspirin concentrations are higher in women due to slower drug clearance.[12] However, because women had higher baseline platelet reactivity, they tended to retain slightly higher platelet aggregation while on aspirin therapy compared with men. It is tempting to speculate that in patients on chronic aspirin therapy for CAD, where platelet reactivity may be heightened, such differences could translate into higher levels of biologically defined aspirin resistance in women. If biologically defined aspirin resistance is related to clinical events, then a higher incidence of resistance among women could translate into less protection from MI.

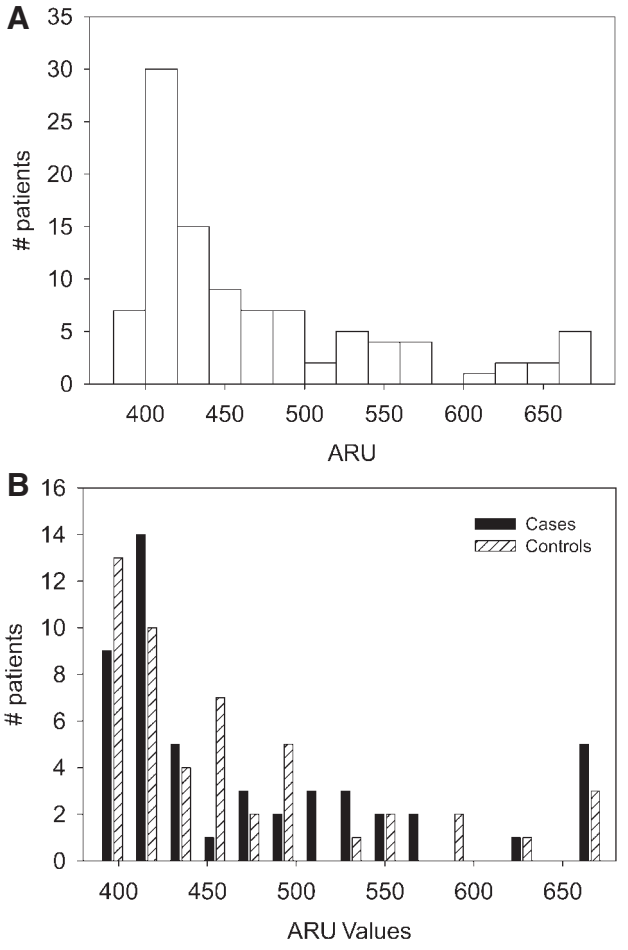


Figure 1. ARU in patients with CAD. **A** = all patients; **B** = cases (history of MI on aspirin) and controls.
ARU = aspirin-resistant units; CAD = coronary artery disease; MI = myocardial infarction.

Currently it is not clear what, if anything, should be done for patients who are identified in a biology assay as aspirin resistant. In several studies, low-dose aspirin was associated with aspirin resistance.[7] The results of our small (n = 6) follow-up substudy suggest that patients on low-dose aspirin who are found to be aspirin resistant in the VerifyNow Aspirin Assay may become aspirin responsive if higher doses are used. However, it is unknown whether those higher doses will increase either protection from cardiovascular events or occurrence of adverse effects in this population.

Limitations

Several factors limit drawing firm conclusions from this study. First, the incidence of biologic aspirin resistance was much lower than expected. Previous studies have indicated that the incidence of aspirin resistance as measured by the VerifyNow Aspirin Assay in patients on chronic aspirin therapy is 23% and is higher in those with known CAD. It is not clear why the incidence of biologic resistance that we observed (14%) was lower than that previously reported. Nonetheless, our study was underpowered to detect a 4% absolute difference in biologic aspirin resistance between cases and controls. A much larger trial would be required to determine whether this difference in aspirin resistance is real. Nonetheless, the clinical significance of such a small difference must be questionable.

A second concern relates to patient adherence to medical therapy both at the time of the initial MI and at the time of platelet function testing. To address this, the patients were screened in an attempt to verify their adherence to aspirin therapy within 48 hours of testing. Third, samples were collected throughout the day, and it is possible that the time of the blood collection could have affected platelet function. Lastly, it has been documented that aspirin responsiveness may change over time.[7,8] It is possible that patients who suffer an MI while taking aspirin therapy may be more likely to be aspirin resistant during that period of time. Likewise, longitudinal studies would be required to determine whether aspirin resistance in our control population is associated with increased ischemic vascular events.

Conclusions

Using a case–control design, an increase in the incidence of aspirin resistance as assessed by the VerifyNow Aspirin Assay among patients on aspirin and with a history of MI was not observed. Female sex was associated with

The Washington Post

warning: Observational Study, associations

Ph.D. , yes
posse, no

V nmjaknf

Why smart people are better off with fewer friends

Ax!Bgqhrsnogdq!Hmfq`g`l L`dpg 07

Hell might actually be other people -- at least if you're really smart.

That's the implication of fascinating new research published last month in the British Journal of Psychology. Evolutionary psychologists Satoshi Kanazawa of the London School of Economics and Norman Li of Singapore Management University dig in to the question of what makes a life well-lived. While traditionally the domain of priests, philosophers and novelists, in recent years survey researchers, economists, biologists and scientists have been tackling that question.

Kanazawa and Li theorize that the hunter-gatherer lifestyles of our ancient ancestors form the foundation for what make us happy now. "Situations and circumstances that would have increased our ancestors' life satisfaction in the ancestral environment may still increase our life satisfaction today," they write.

They use what they call "the savanna theory of happiness" to explain two main findings from an analysis of a large national survey (15,000 respondents) of adults aged 18 to 28.

First, they find that people who live in more densely populated areas tend to report less satisfaction with their life overall. "The higher the population density of the immediate environment, the less happy" the survey respondents said they were. Second, they find that the more social interactions with close friends a person has, the greater their self-reported happiness.

But there was one big exception. For more intelligent people, these correlations were diminished or even reversed.

"The effect of population density on life satisfaction was therefore more than twice as large for low-IQ individuals than for high-IQ individuals," they found. And "more intelligent individuals were actually less satisfied with life if they socialized with their friends more frequently."

Let me repeat that last one: When smart people spend more time with their friends, it makes them less happy.

Now, the broad contours of both findings are largely uncontroversial. A large body of previous research, for instance, has outlined what some have called an "urban-rural happiness gradient." Kanazawa and Li explain: "Residents of rural areas and small towns are happier than those in suburbs, who in turn are happier than those in small central cities, who in turn are happier than those in large central cities."

Why would high population density cause a person to be less happy? There's a whole body of sociological research addressing this question. But for the most visceral demonstration of the effect, simply take a 45-minute ride on a crowded

rush-hour Red Line train and tell me how you feel afterward.

Kanazawa and Li's second finding is a little more interesting. It's no surprise that friend and family connections are generally seen as a foundational component of happiness and well-being. But why would this relationship get turned on its head for really smart people?

I posed this question to Carol Graham, a Brookings Institution researcher who studies the economics of happiness. "The findings in here suggest (and it is no surprise) that those with more intelligence and the capacity to use it ... are less likely to spend so much time socializing because they are focused on some other longer term objective," she said.

Think of the really smart people you know. They may include a doctor trying to cure cancer or a writer working on the great American novel or a human rights lawyer working to protect the most vulnerable people in society. To the extent that frequent social interaction detracts from the pursuit of these goals, it may negatively affect their overall satisfaction with life.

But Kanazawa and Li's savanna theory of happiness offers a different explanation. The idea starts with the premise that the human brain evolved to meet the demands of our ancestral environment on the African savanna, where the population density was akin to what you'd find today in, say, rural Alaska (less than one person per square kilometer). Take a brain evolved for that environment, plop it into today's Manhattan (population density: 27,685 people per square kilometer), and you can see how you'd get some evolutionary friction.

Similarly with friendship: "Our ancestors lived as hunter–gatherers in small bands of about 150 individuals," Kanazawa and Li explain. "In such settings, having frequent contact with lifelong friends and allies was likely necessary for survival and reproduction for both sexes." We remain social creatures today, a reflection of that early reliance on tight-knit social groups.

The typical human life has changed rapidly since then -- back on the savanna we didn't have cars or iPhones or processed food or "Celebrity Apprentice" -- and it's quite possible that our biology hasn't been able to evolve fast enough to keep up. As such, there may be a "mismatch" between what our brains and bodies are designed for, and the world most of us live in now.

To sum it all up: You've heard of the paleo-diet. But are you ready for paleo-happiness?

There's a twist, though, at least as Kanazawa and Li see it. Smarter people may be better equipped to deal with the new (at least from an evolutionary perspective) challenges present-day life throws at us. "More intelligent individuals, who possess higher levels of general intelligence and thus greater ability to solve evolutionarily novel problems, may face less difficulty in comprehending and dealing with evolutionarily novel entities and situations," they write.

If you're smarter and more able to adapt to things, you may have an easier time reconciling your evolutionary predispositions with the modern world. So living in a high-population area may have a smaller effect on your overall well-being -- that's what Kanazawa and Li found in their survey analysis. Similarly, smarter people may be better-equipped to jettison that whole hunter-gatherer social network -- especially if they're pursuing some loftier ambition.

British Journal of Psychology (2016)

The British Psychological Society

www.wileyonlinelibrary.com

Country roads, take me home. . . to my friends: How intelligence, population density, and friendship affect modern happiness

Norman P. Li[1] and Satoshi Kanazawa[2]*

[1]School of Social Sciences, Singapore Management University, Singapore
[2]Managerial Economics and Strategy Group, Department of Management, London School of Economics and Political Science, UK

We propose the savanna theory of happiness, which suggests that it is not only the current consequences of a given situation but also its ancestral consequences that affect individuals' life satisfaction and explains why such influences of ancestral consequences might interact with intelligence. We choose two varied factors that characterize basic differences between ancestral and modern life – population density and frequency of socialization with friends – as empirical test cases. As predicted by the theory, population density is negatively, and frequency of socialization with friends is positively, associated with life satisfaction. More importantly, the main associations of life satisfaction with population density and socialization with friends significantly interact with intelligence, and, in the latter case, the main association is reversed among the extremely intelligent. More intelligent individuals experience lower life satisfaction with more frequent socialization with friends. This study highlights the utility of incorporating evolutionary perspectives in the study of subjective well-being.

Positive psychology and evolutionary psychology are two subfields of psychology that have made significant advances in the last few decades (Cosmides & Tooby, 2013; Diener, 2012). While several evolutionary psychologists have written on happiness (Buss, 2000; Hill & Major, 2013; Lewis, Al-Shawaf, Russell, & Buss, 2015; Nesse, 2004), with only a couple of exceptions (Diener, Kanazawa, Suh, & Oishi, 2015; Heintzelman & King, 2014), positive psychologists have not drawn on insights from evolutionary psychology. At the same time, while positive psychologists have accumulated an impressive amount of empirical knowledge in the last few decades about who is happier than whom, when, and how, there are few systematic general theories of happiness – evolutionary or otherwise – that can explain *why* some individuals are happier than others. In this study, we propose an evolutionary psychological theory of subjective well-being that we call *the savanna theory of happiness,* and provide empirical support for the theory.

*Correspondence should be addressed to Satoshi Kanazawa, Managerial Economics and Strategy Group, Department of Management, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, UK (email: S.Kanazawa@lse.ac.uk).

DOI:10.1111/bjop.12181

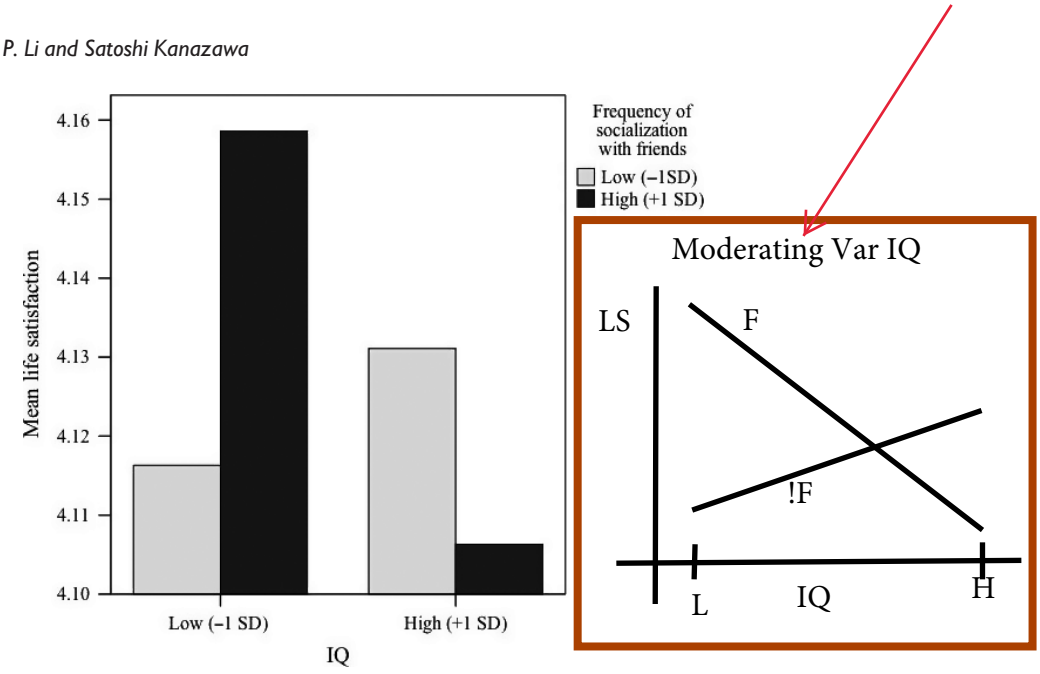


Figure 2. Interaction effect between frequency of socialization with friends and intelligence on life satisfaction.

(M = 4.1063) than those who socialized with friends less frequently (M = 4.1311). The statistical interaction was such that more intelligent individuals were actually *less* satisfied with life if they socialized with their friends more frequently. Among low-IQ individuals, the mean difference in life satisfaction between those who socialize with friends more and less frequently (0.423) translated to d = .05, and that among high-IQ individuals (−0.0248) to d = −.03.

Given that our data are correlational and frequency of socialization with friends and life satisfaction were measured at the same time, we cannot rule out an opposite causal order to what we hypothesize, where happier people choose to socialize with their friends more frequently. This may potentially be a problem because our measure of frequency of socialization with friends referred to recent past ('In the past 7 days'), while the measure of life satisfaction was global ('as a whole'). We are sure there are some mutual influences between life satisfaction and frequency of socialization with friends, but there are a few considerations, suggesting that the results largely reflect our hypothesized causality. For instance, Baker, Cahalin, Gerst, and Burr (2005) showed that the positive effect of seeing family and friends on subjective well-being remained even after controlling for the earlier level of life satisfaction in a previous wave of a longitudinal survey. Similarly, in our data, frequency of socialization with friends was still significantly associated with life satisfaction even after happiness at Waves I and II (measured by the question 'How often was each of the following things true during the past seven days? You were happy'. 0 = *never or rarely*, 1 = *sometimes*, 2 = *a lot of the time*, 3 = *most of the time or all of the time*), in addition to current marital status, was controlled (b = .018, p = .016).

Discussion

Consistent with our prediction derived from the savanna theory of happiness, and with past empirical studies, frequency of socialization with friends had a significantly positive

Table 2. Frequency of socialization with friends and life satisfaction

	(1)	(2)	(3)	(4)
Frequency of socialization with friends	.008	.031***	.017*	.103**
	(.006)	(.007)	(.007)	(.035)
Currently married		.663***	.773***	.765***
		(.043)	(.045)	(.045)
Age			−.060***	−.059***
			(.009)	(.009)
Sex			.126***	.125***
			(.031)	(.032)
Ethnicity				
African American			−.194***	−.203***
			(.038)	(.040)
Asian American			−.301***	−.301***
			(.057)	(.058)
Native American			−.215**	−.210**
			(.068)	(.069)
Education			.145***	.147***
			(.008)	(.009)
Intelligence				.003
				(.002)
Intelligence*frequency of socialization with friends				−.001*
				(.000)
Threshold				
Y = 1	−5.060	−4.876	−4.389	−4.028
	(.108)	(.109)	(.236)	(.282)
Y = 2	−3.113	−2.928	−2.434	−2.084
	(.049)	(.051)	(.215)	(.264)
Y = 3	−1.573	−1.382	−.866	−.518
	(.035)	(.037)	(.213)	(.262)
Y = 4	.590	.806	1.364	1.710
	(.033)	(.036)	(.213)	(.262)
−2Log Likelihood	255.142	408.003***	15615.357***	29472.587***
Cox & Snell pseudo-R^2	.000	.016	.041	.041
Number of cases	15,111	15,111	15,047	14,513

Note. Main entries are unstandardized regression coefficients.
Numbers in parentheses are standard errors.
*p < .05; **p < .01; ***p < .001.

stronger among less intelligent individuals than among more intelligent individuals. The interaction term between intelligence and frequency of socialization with friends was significantly negative ($b = -.001$, $p = .014$).

Figure 2 presents the statistical interaction graphically. Among less intelligent individuals (with a mean IQ of 81.39), frequency of socialization with friends had a significantly positive effect on life satisfaction. Those who socialized with friends more frequently (6.71, nearly every day) had a significantly higher life satisfaction ($M = 4.1586$) than those who socialized with friends less frequently (1.95, less than twice a week) ($M = 4.1163$). In contrast, among more intelligent individuals (with a mean IQ of 115.57), those who socialized with friends more frequently were actually *less satisfied with life*

Dependent variable: Global life satisfaction

Add Health asked its respondents 'How satisfied are you with your life as a whole?': 1 = *very dissatisfied*, 2 = *dissatisfied*, 3 = *neither satisfied nor dissatisfied*, 4 = *satisfied*, and 5 = *very satisfied* (reverse coded). We used this measure of life satisfaction as the dependent variable in our ordinal regression analysis.[1]

Independent variable: Population density

Add Health measured the population density at the level of census block group (a subdivision of a census tract and the smallest geographic unit for which the Census Bureau tabulates aggregate data), census tract, county, and state. It was measured as the number of persons in thousands/km^2.

The distributions of population density at all levels were extremely positively skewed (skewness: block group = 6.780; census tract = 7.449; county = 8.702; state = 17.460). We therefore took the natural log of the measures of population density, which nearly eliminated the skewness (skewness after log normalization: block group = −.809; census tract = −.684; county = .023; state = −.121). We used the natural logs of measures of population density in our regression analyses below.

Independent variable: Intelligence

Add Health measured respondents' intelligence by an abbreviated version of the Peabody Picture Vocabulary Test. Their raw scores were transformed into the standard IQ metric, with a mean of 100 and a standard deviation of 15. The Peabody Picture Vocabulary Test is properly a measure of verbal intelligence. However, verbal intelligence is known to be highly correlated with and thus heavily load on general intelligence (Huang & Hauser, 1998; Miner, 1957; Wolfle, 1980).

Control variables

In addition, we controlled for the following characteristics of the respondent: sex (0 = *female*, 1 = *male*); age (in years); ethnicity (with three dummies for African American, Asian American, and Native American, with White American as the reference category); education (in years of formal schooling); and current marital status (1 = *currently married*, 0 = *otherwise*). All of these variables are known correlates of happiness (Dolan *et al.*, 2008). Preliminary analysis showed that the respondent's earnings had no association with life satisfaction among Add Health respondents (r = .013, p = .116, n = 14,414), perhaps because of their relative youth and little variance in earnings (M = 11,744, median = 8,000 SD = 17,289, IQR = 16,500, n = 14,425). This was consistent with earlier studies, which showed that variance in earnings generally increased with age (Beach, Finnie, & Gray, 2010; Caswell & Kluge, 2015; Lam & Levison, 1992).

[1] *More sophisticated statistical procedures like structural equation modelling (SEM) or multilevel modelling (MLM) are not feasible with our data. SEM is not feasible because we have only one indicator each for all of the variables in our analysis, and MLM is not feasible because, while we know the population density of the county or the state of residence, for example, we do not know in which county or state the Add Health respondents reside. (Add Health is extremely concerned about privacy issues and does not make much individually identifiable information available in the data.) So we cannot perform MLM by nesting individual respondents in the county or state of their residence.*

Table 1. (*Continued*)

	(5) Block group	(6) Census tract	(7) County	(8) State
$Y = 2$	−2.258	−2.254	−2.105	−2.173
	(.215)	(.215)	(.218)	(.222)
$Y = 3$	−.690	−.686	−.536	−.606
	(.212)	(.212)	(.216)	(.220)
$Y = 4$	1.539	1.543	1.694	1.621
	(.212)	(.212)	(.216)	(.220)
−2Log Likelihood	32503.061***	31976.678***	27210.467***	21051.828***
Cox & Snell pseudo-R^2	.042	.042	.043	.041
Number of cases	14,811	14,811	14,811	14,811

	(9) Block group	(10) Census tract	(11) County	(12) State
Population density	−.248***	−.222***	−.284***	−.435**
	(.054)	(.054)	(.059)	(.141)
Age	−.050***	−.051***	−.049***	−.054***
	(.010)	(.010)	(.010)	(.009)
Sex	.118***	.118***	.119***	.120***
	(.032)	(.032)	(.032)	(.032)
Ethnicity				
African American	−.192***	−.191***	−.188***	−.197***
	(.041)	(.041)	(.041)	(.040)
Asian American	−.256***	−.263***	−.255***	−.288***
	(.059)	(.059)	(.059)	(.059)
Native American	−.194**	−.198**	−.203**	−.220**
	(.069)	(.069)	(.069)	(.069)
Education	.153***	.153***	.154***	.149***
	(.009)	(.009)	(.009)	(.009)
Currently married	.710***	.710***	.708***	.716***
	(.045)	(.045)	(.045)	(.045)
Intelligence	−.001	.000	.003*	.010*
	(.001)	(.001)	(.001)	(.004)
Intelligence*population density	.002***	.002***	.002***	.003*
	(.001)	(.001)	(.001)	(.001)
Threshold				
$Y = 1$	−4.227	−4.167	−3.765	−3.098
	(.253)	(.253)	(.266)	(.469)
$Y = 2$	−2.258	−2.197	−1.795	−1.129
	(.233)	(.233)	(.247)	(.458)
$Y = 3$	−.691	−.630	−.227	.436
	(.230)	(.230)	(.244)	(.457)
$Y = 4$	1.537	1.596	2.001	2.661
	(.231)	(.231)	(.245)	(.458)
−2Log Likelihood	31731.139***	31710.645***	31321.322***	30674.162***
Cox & Snell pseudo-R^2	.043	.042	.043	.041
Number of cases	14,278	14,278	14,278	14,278

Note. Main entries are unstandardized regression coefficients.
Numbers in parentheses are standard errors.
*$p < .05$; **$p < .01$; ***$p < .001$.

that for A. Individuals with high aptitude generally do better if assigned to Treatment B. Treatment A tends to be superior for persons below X^* (the "crossover point" where the two lines intersect). Despite the difference at the mean, Treatment A is not superior in general.

Investigators should attend to differences in mean outcome and also to differences in relations of aptitudes to outcomes. The interaction study combines the correlational and experimental modes of inquiry and goes beyond them. Where policy makers have a commitment to universal education and a set of common educational goals, a principal aim of research on instruction becomes one of finding interactions and capitalizing on them.

Cronbach, beyond the two disciplines.. ATI research

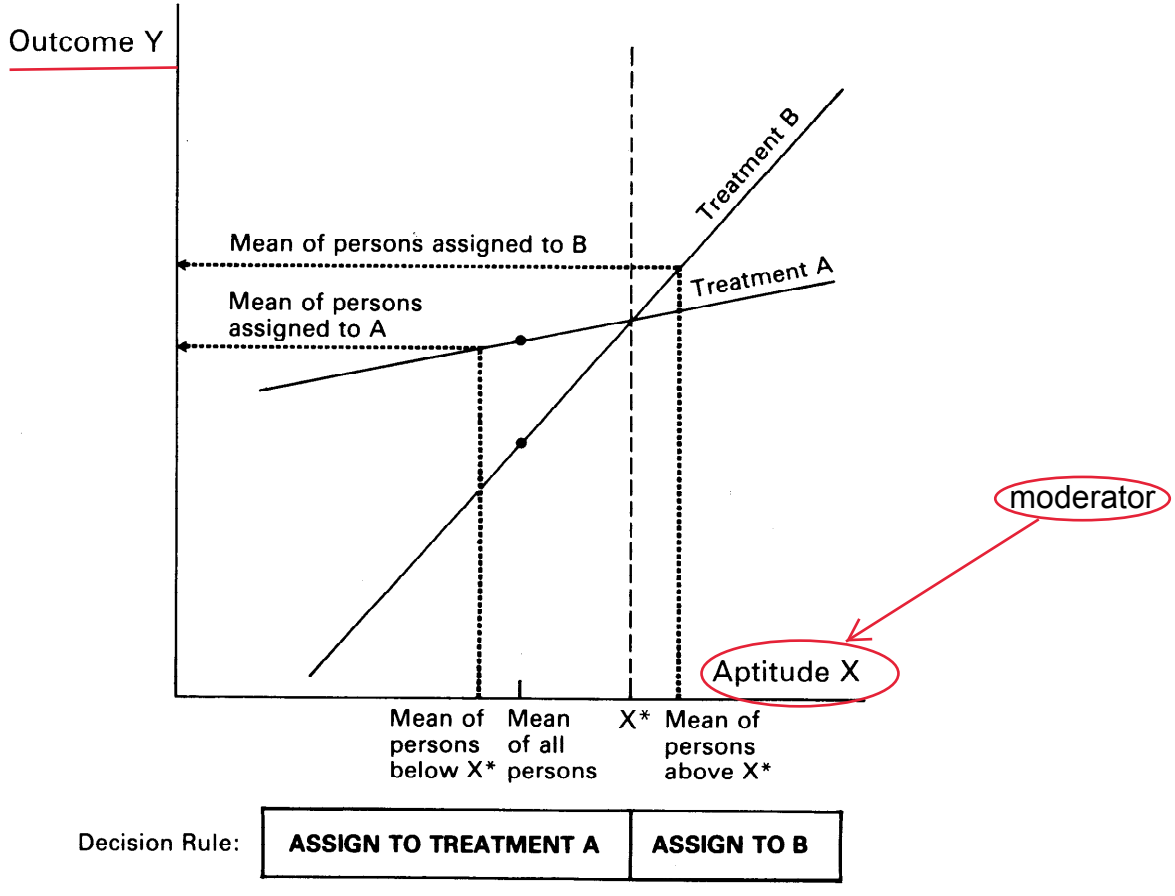


Figure 2.3.
The scheme for examining aptitude-treatment interaction.

ATI Cronbach-Snow disordinal interaction differential assignment uncertainty around X* J-N region of sig complement of CI for X*

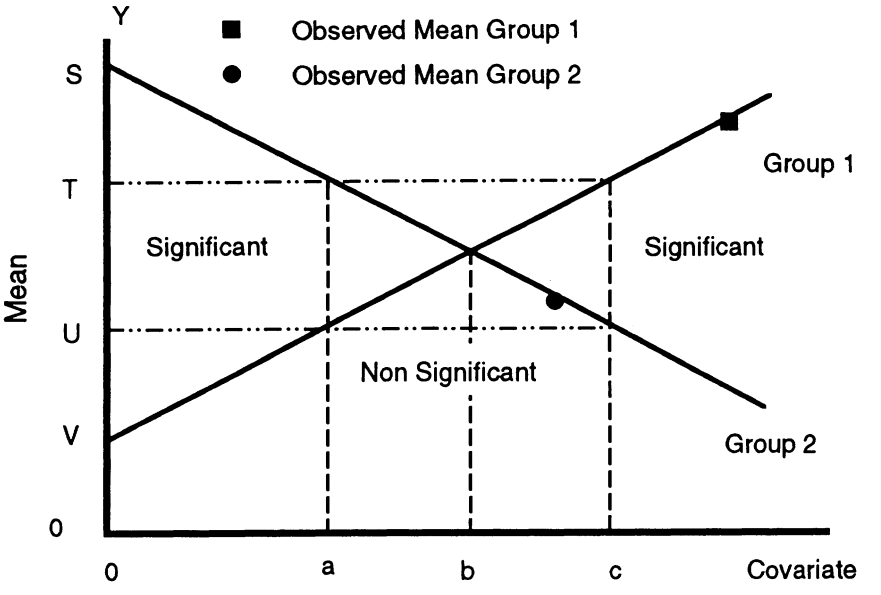


FIGURE 1. *Nonhomogeneous regressions for two groups*

covariate value of *b*, no differences exist between the adjusted group means. Experimental results that suggest the sort of group differences represented in Figure 1 can vary from nonsignificant to significant and can favor different groups, depending upon the value of the covariate to which the groups are adjusted. If the regression lines are parallel—that is, if the assumption of homogeneity of the regression slopes is tenable—then it makes no difference to what value of the covariate the means are adjusted, because the adjustment is identical for all values of the covariate. Many computer program packages which carry out an ANCOVA implicitly adjust the group means to a covariate value of 0 or do not tell the researcher to what value the adjustment was made.

In the case of nonhomogeneous regression slopes, the problem to be solved is to find the values of *a* and *c* in Figure 1 such that they separate the boundaries of statistical significance and nonsignificance for a nominal value of alpha in testing the differences between the adjusted group means of the dependent variable. The values of *a* and *c* are said to delimit *regions of significance.* For a single covariate the values of *a* and *c* can be found by solving a quadratic in one unknown. For two or more covariates the problem becomes more difficult, because the regions of significance are defined by point sets of covariate values. A modification to the usual ANCOVA to accommodate nonhomogeneity of regression slopes was first detailed by Johnson and Neyman (1936) and is usually referred to as the Johnson-Neyman ANCOVA or the Johnson-Neyman procedure.

Recent critical discussion of the ANCOVA including the Johnson-Neyman procedure (Maxwell & Delaney, 1990; Maxwell, O'Callaghan, & Delaney, 1993; Rutherford, 1992) suggests that this method of analysis continues to be of

event than those in C. The directionality of mediation and moderation is important to note. Moderators always precede what they moderate, which in turn precedes

A posttreatment measure (not a moderator) uncorrelated with treatment (not a mediator) that has a main effect but no interaction might also be called a nonspecific predictor of

Moderator as covariate, extreme non-parallel ex

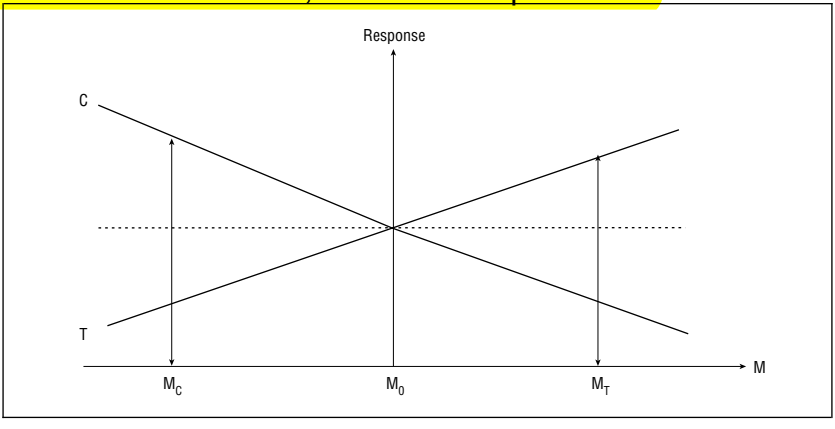


Figure 2. A special case in which there is no main effect of treatment, no main effect of moderator or mediator (M), and no overall effect of treatment, but in which treatment may change not only the level but also the action of M on the outcome, a mediating effect. T indicates treatment group; C, control or comparison group; M_C, the mean of M in C; M_T, the mean of M in T; and M_0, the midpoint of these two.

Summary of Population Definitions Relating Target Measure to Treatment and Outcome

Target Measure	Correlation With Treatment	Relationship to Outcome in Linear Model
Pretreatment	No (by definition)	Interaction with or without main effect
Pretreatment	No (by definition)	Main effect only
Posttreatment	Yes	Main effect or interaction
Posttreatment	Yes	Neither main effect nor interaction
Posttreatment	No	Interaction with or without main effect
Posttreatment	No	Main effect only
Pretreatment or posttreatment	No	Neither main effect nor interaction

week2, Kraemer

BIOMETRICS 56, 882–885
September 2000

Unit–Treatment Interaction and Its Practical Consequences

Gary L. Gadbury

Department of Mathematical Sciences, University of North Carolina–Greensboro, Greensboro, North Carolina 27402, U.S.A.
email: glgadbur@uncg.edu

and

Hari K. Iyer

Department of Statistics, Colorado State University, Fort Collins, Colorado 80523, U.S.A.

SUMMARY. Most statistical characterizations of a treatment effect focus on the average effect of the treatment over an entire population. However, average effects may provide inadequate information, sometimes misleading information, when a substantial unit–treatment interaction is present in the population. It is even possible that a nonnegligible proportion of the individuals in the population experience an unfavorable treatment effect even though the treatment might appear to be beneficial when considering population averages. This paper examines the extent to which information about unit–treatment interaction can be extracted using observed data from a two-treatment completely randomized experiment. A method for utilizing the information from an available covariate is proposed. Although unit–treatment interaction is a nonidentifiable quantity, we show that mathematical bounds for it can be estimated from observed data. These bounds lead to estimated bounds for the probability of an unfavorable treatment effect. Maximum likelihood estimators of the bounds and their corresponding large-sample distributions are given. The use of the estimated bounds is illustrated in a clinical trials data example.

KEY WORDS: Additivity; Clinical trial; Counterfactual; Heterogeneity; Potential response; Subject–treatment interaction.

1. Introduction

We begin by considering a clinical trial where a new treatment, T_1, is being compared with a standard treatment, T_2. An example of such a trial is discussed in Section 4. Let X denote what the response would be if a randomly chosen individual from the population is subjected to treatment T_1 and Y denote what the response of this individual would be if subjected to treatment T_2. X and Y are called potential responses, and the quantity $D = X - Y$ may be defined as the effect of treatment T_1 relative to T_2 for the chosen individual. What is generally estimated in clinical trials is an average effect, $\mathrm{E}(X - Y)$, where the expectation $\mathrm{E}(\cdot)$ is with respect to the population of interest.

It is not always the case that every subject in the population will experience a beneficial effect due to treatment T_1. We will interpret the phrase "T_1 has a beneficial effect" to mean $X - Y \geq \tau$, where τ is some specified constant. Without loss of generality, we take τ to be zero. It is important to note that the effect of T_1 relative to T_2 could appear beneficial when considering the average effect even though a nonnegligible proportion of individuals could be experiencing an unfavorable effect. In fact, Longford (1999) suggested that the validity of current clinical trial design and analysis is greatly eroded when treatment effects are heterogeneous. A key parameter, the value of which should be considered when making decisions concerning the use of a new treatment T_1, is the proportion of the individuals in the population for whom the value of $D = X - Y$ is negative. We denote this proportion as $P_- = P(D < 0)$. If P_- is nonnegligible, it becomes important to identify the subset of individuals in the population who actually benefit from T_1. For instance, in our discussion of a clinical trials data set in Section 4 involving epilepsy patients, we show that the maximum likelihood estimates for lower and upper bounds for P_- are 5 and 36%, respectively. We also show how to obtain approximate confidence intervals for these bounds, which then allows us to test the statistical significance of these estimates. We then obtain similar bounds and confidence intervals for P_- conditioned on the value of a suitable covariate. This latter procedure is particularly useful in identifying the range of values of the covariate for which the treatment may be safely recommended.

We will assume that the random vector $\{(X, Y)\}$ of potential responses follows a bivariate normal distribution (possibly after a suitable monotonic transformation). The value of P_- is then nonzero when $\sigma_D^2 = \mathrm{var}(X - Y)$ is nonzero, i.e., when unit–treatment interaction is present. The problem of estimating P_- is tied to the problem of estimating σ_D^2. If the treatments T_1 and T_2 are assigned to a random sam-

which subjects benefit?

Original Articles

Estimating Conditional Average Treatment Effects

Jason Abrevaya, Yu-Chin Hsu & Robert P. Lieli

Pages 485-505

Received 01 Jul 2012, Accepted author version posted online: 30 Oct 2014,Published online: 27 Oct 2015

Download citation https://doi.org/10.1080/07350015.2014.975555 Check for updates

Abstract

We consider a functional parameter called the conditional average

observational study example

unconfoundedness assumption applies. In contrast to quantile regressions, the subpopulations of interest are defined in terms of the possible values of a set of continuous covariates rather than the quantiles of the potential outcome distributions. We show that the CATE parameter is nonparametrically identified under unconfoundedness and propose inverse probability weighted estimators for it. Under regularity conditions, some of which are standard and some are new in the literature, we show (pointwise) consistency and asymptotic normality of a fully nonparametric and a semiparametric estimator. We apply our methods to estimate the average effect of a first-time mother's smoking during pregnancy on the baby's birth weight as a function of the mother's age. A robust qualitative finding is that the expected effect becomes stronger (more negative) for older mothers.

KEY WORDS: Birth weight, Inverse probability weighted estimation, Nonparametric method, Treatment effect heterogeneity

Psychological Bulletin
1980, Vol. 88, No. 2, 307-321

Comparing Nonparallel Regression Lines

David Rogosa
Department of Education, University of Chicago

Statistical comparisons of experimental groups frequently are based on the within-group regressions of an outcome variable on a concomitant variable. I present a comprehensive strategy for the statistical comparison of within-group regressions that is suitable for both parallel and nonparallel regression lines. New results are obtained and new interpretations are formulated for standard statistical procedures such as analysis of covariance. New procedures for comparing regressions are proposed and illustrated.

The setting for this investigation is an experiment in which two groups are given alternative treatments. The groups are formed by random assignment of individual cases to the groups. One of these groups may be considered a control group, as in a study to evaluate the effectiveness of a curricular innovation relative to the effectiveness of a standard curriculum. The purpose of the experiment is to assess the treatment effect—the differential effectiveness of the two treatments.

A number of initial characteristics (usually called covariates or predictor variables) are recorded for each case before the initiation of the treatment; outcome variables are measured at the end of the treatment period. I limit consideration to one outcome variable (Y) and one predictor variable (X). The raw data from this experiment are measures of X and Y that are obtained from the members of the two experimental groups.

The raw data are summarized through estimation of the within-group regression lines, the average outcome for a given initial characteristic value within each group. The summary of the raw data consists of (a) the separate sample within-group regression lines and (b) estimates of the sampling variances and covariances associated with these sample regressions.

This article presents and evaluates statistical procedures for comparing the within-group regression lines. I view the treatment effect as a function of X and define the treatment effect as the difference between the population regression lines. The usual dichotomy between parallel and nonparallel regressions is shunned. The difference between the sample regression lines is used to estimate this treatment effect, and this estimate will depend on X (to some degree) whenever the sample within-group regressions are not parallel. Although this approach seems natural for problems of comparing regressions, the development and exposition of traditional statistical methods has proceeded along other lines.

Two types of assessments of the treatment effect are sought. First, an overall treatment effect, in which dependence on X of the treatment effect is ignored, can be estimated by evaluating the difference between the sample regression lines at a prespecified value of X. These procedures are called pick-a-point procedures. Second, an assessment of the difference between the regressions over the entire range of X can be used to evaluate the treatment effect as a function of X. The Johnson-Neyman technique as extended by Potthoff (1964) is one procedure in which the regressions are compared over the range of X. Nonsimultaneous inference procedures are associated with the first type of assessment, and simultaneous inference procedures are associated with the second type of assessment.

Requests for reprints should be sent to David Rogosa, who is now at the School of Education, Stanford University, Stanford, California 94305.

 0033-2909/80/8802-0307$00.75

Package 'probemod'

April 22, 2015

Title Statistical Tools for Probing Moderation Effects

Version 0.2.1

Description Contains functions that are useful for probing moderation effects (or interactions) including techniques such as pick-a-point (also known as spotlight analysis) and Johnson-Neyman (also known as floodlight analysis). Plot function is also provided to facilitate visualization of results from each of these techniques.

Depends R (>= 3.1.2)

License GPL-3

LazyData no

Author Jiat Chow Tan [aut, cre]

Maintainer Jiat Chow Tan <w110013@ntu.edu.sg>

NeedsCompilation no

Repository CRAN

Date/Publication 2015-04-22 22:01:37

R topics documented:

`pickapoint` *Pick-A-Point Technique*

Description

Probe moderation effect using the Pick-A-Point technique

Usage

```
pickapoint(model, dv, iv, mod, points, method = "meansd", alpha = 0.05,
  yas = "none")
```

Arguments

`model`	Regression model (lm, glm, list).
`dv`	Dependent variable (character).
`iv`	Independent variable (character).
`mod`	Moderator variable(s) (character or character vector).
`points`	List of points to test for each moderator variable (list).
`method`	Method to use. Possible values are: `"meansd"`, `"percentiles"`, `method="meansd"` by default.
`alpha`	Alpha level to use (numeric).
`yas`	Show y (or conditional effect) as: `"none"`, `"ratio"`, `"probability"`, `"percentage"`, `yas="none"` by default.

Value

A list with the elements

References

Spiller, S. A., Fitzsimons, G. J., Lynch, J. G., Jr, & McClelland, G. H. (2013). Spotlights, floodlights, and the magic number zero: Simple effects tests in moderated regression. Journal of Marketing Research, 50(2), 277-288.

Aiken, L. S., & West, S. G. (1991). Multiple regression: Testing and interpreting interactions. Thousand Oaks, CA: Sage Publications.

Examples

```
## Not run:
myModel <- lm('dv ~ iv + mod', data=someData)
pickapoint(myModel, dv='DV', iv='IV', mod='MOD')
pickapoint(myModel, dv='DV', iv='IV', mod='MOD', alpha=.01)
pickapoint(myModel, dv='DV', iv='IV', mod='MOD', method='percentiles')
pickapoint(myModel, dv='DV', iv='IV', mod='MOD', points=c(1,2,3))

## End(Not run)
```

jn *Johnson-Neyman Technique*

Description

Probe moderation effect using the Johnson-Neyman technique

Usage

```
jn(model, dv, iv, mod, mrange, alpha = 0.05, yas = "none")
```

Arguments

model	Regression model (lm, glm, list).
dv	Dependent variable (character).
iv	Independent variable (character).
mod	Moderator variable(s) (character or character vector).
mrange	Range of values that jn should examine for moderator variable. Uses the current range of moderator values by default (numeric vector).
alpha	Alpha level to use (numeric).
yas	Show y (or conditional effect) as: "none", "ratio","probability","percentage", yas="none" by default.

Value

A list with the elements

References

Spiller, S. A., Fitzsimons, G. J., Lynch, J. G., Jr, & McClelland, G. H. (2013). Spotlights, floodlights, and the magic number zero: Simple effects tests in moderated regression. Journal of Marketing Research, 50(2), 277-288.

Bauer, D. J., & Curran, P. J. (2005). Probing interactions in fixed and multilevel regression: Inferential and graphical techniques. Multivariate Behavioral Research, 40(3), 373-400.

Examples

```
## Not run:
myModel <- lm('DV ~ IV + MOD', data=someData)
jnresults <- jn(myModel, dv='DV', iv='IV', mod='MOD')
jnresults <- jn(myModel, dv='DV', iv='IV', mod='MOD', alpha=.01)
plot(jnresults)

## End(Not run)
```

Stat 209 [redacted]

ANCOVA, CNRL equations

~~Week 6~~ math notes

precursor: t-test $Y = \beta_0 + \beta_1 G + e$ $\hat{\beta}_1 / se(\hat{\beta}_1)$ pooled t-test (not Welch)

$G = 0, 1$ group membership

$$Y = \gamma_0 + \gamma_1 G + \gamma_2 X + e$$

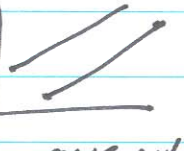
link w/ data (A)

ancova

$$\hat{\gamma}_1 = \bar{Y}_1 - \bar{Y}_0 - \hat{\gamma}_p (\bar{X}_1 - \bar{X}_0)$$

constant treatment effect

$\hat{\gamma}_2 = \hat{\gamma}_p$ ave w/in group slopes



cf. overheads derive on board HSB school-level ex

---

more general model (CNRL)

(B) $Y = \beta_1 + \beta_2 G + \beta_3 X + \beta_4 XG + e$

"interaction" term

see Berk, Barnow ATI intros

Group Membership (Dummy Variables)

indicator var's
Nwk p. 351 eq 10.3

<Ident(sec 10.1 >

Define $G_i = 1$ if unit i is in group 1 (e.g. treatment)

$G_i = 0$ otherwise

(i.e. unit i in group 0, control)

Y_i outcome for unit i.

$E(Y_i | G = 1) = \mu_1$; $E(Y_i | G = 0) = \mu_0$

$E(Y|G) = \beta_0 + \beta_1 G$

regression formulation for pooled t-test

$\beta_0 = \mu_0$ $\beta_1 = \mu_1 - \mu_0$

Analysis of Covariance

(NKNW Ch. 10 Indicator vars; Chap 25 ancova ?
one-way 25.1-23.4?)

Using concomitant (pretest) information to improve precision of group comparisons.

Comparing conditional expectations $E(Y|x)$ (NKNW 25.1)

more precise than comparing (unconditional) group means ($\bar{Y}$).

Single classification: Regression Approach (Kutner p. 869)
outcome Y, covariate X

$I-1$ Group membership vars
$G_1, \ldots G_{I-1}$. (NKNW 25.3 crackers)

Stat 209 [redacted]

ANCOVA, CNRL equations

[redacted] math notes

precursor: t-test $Y = \beta_0 + \beta_1 G + e$ $\hat{\beta}_1 / se(\hat{\beta}_1)$

$G = 0, 1$ group membership

$$Y = \gamma_0 + \gamma_1 G + \gamma_2 X + e$$

pooled t-test (not Welch)

link w/ data (A)

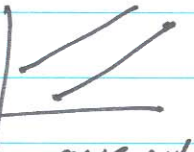
ancova

$$\hat{\gamma}_1 = \bar{Y}_1 - \bar{Y}_0 - \hat{\delta}_p (\bar{X}_1 - \bar{X}_0)$$

constant treatment effect

$$\hat{\gamma}_2 = \hat{\delta}_p$$

ave w/in group slopes



cf. overheads derive on board HSB school-level ex

more general model (CNRL)

(B) $$Y = \beta_1 + \underline{\beta_2 G} + \beta_3 X + \beta_4 XG + e$$

"interaction" term

see Berk, Data ATI intros

Comparing Conditional Expectations: Ancova

2 groups (1 and 0)
outcome Y, covariate X

Straight-line regression function within each group

$$E(Y|X) = \alpha_1 + \gamma_1 X \quad \text{(Group 1)}$$

$$E(Y|X) = \alpha_0 + \gamma_0 X \quad \text{(Group 0)}$$

[ancova restrictions: $\gamma_1 = \gamma_0 = \gamma$
$\sigma_1^2 = \sigma_0^2 = \sigma^2$ (equal residual variances)]

Fit a straight-line within each group

$$\hat{Y} = \hat{\alpha}_1 + \hat{\gamma}_1 X = \bar{Y}_1 + \hat{\gamma}_1 (X - \bar{X}_1) \quad G1$$

$$\hat{Y} = \hat{\alpha}_0 + \hat{\gamma}_0 X = \bar{Y}_0 + \hat{\gamma}_0 (X - \bar{X}_0) \quad G0$$

Analysis of covariance presumes equal w/in group slopes (11 lines) so $\hat{\gamma}_1, \hat{\gamma}_0$ estimate common γ

Replace w/in group slopes by ⟨ called gw NWK 23.35 ⟩

$$\hat{\gamma}_p = \frac{\hat{\gamma}_1 SSX_1 + \hat{\gamma}_0 SSX_0}{SSX_1 + SSX_0}$$

to obtain two parallel w/in group regressions

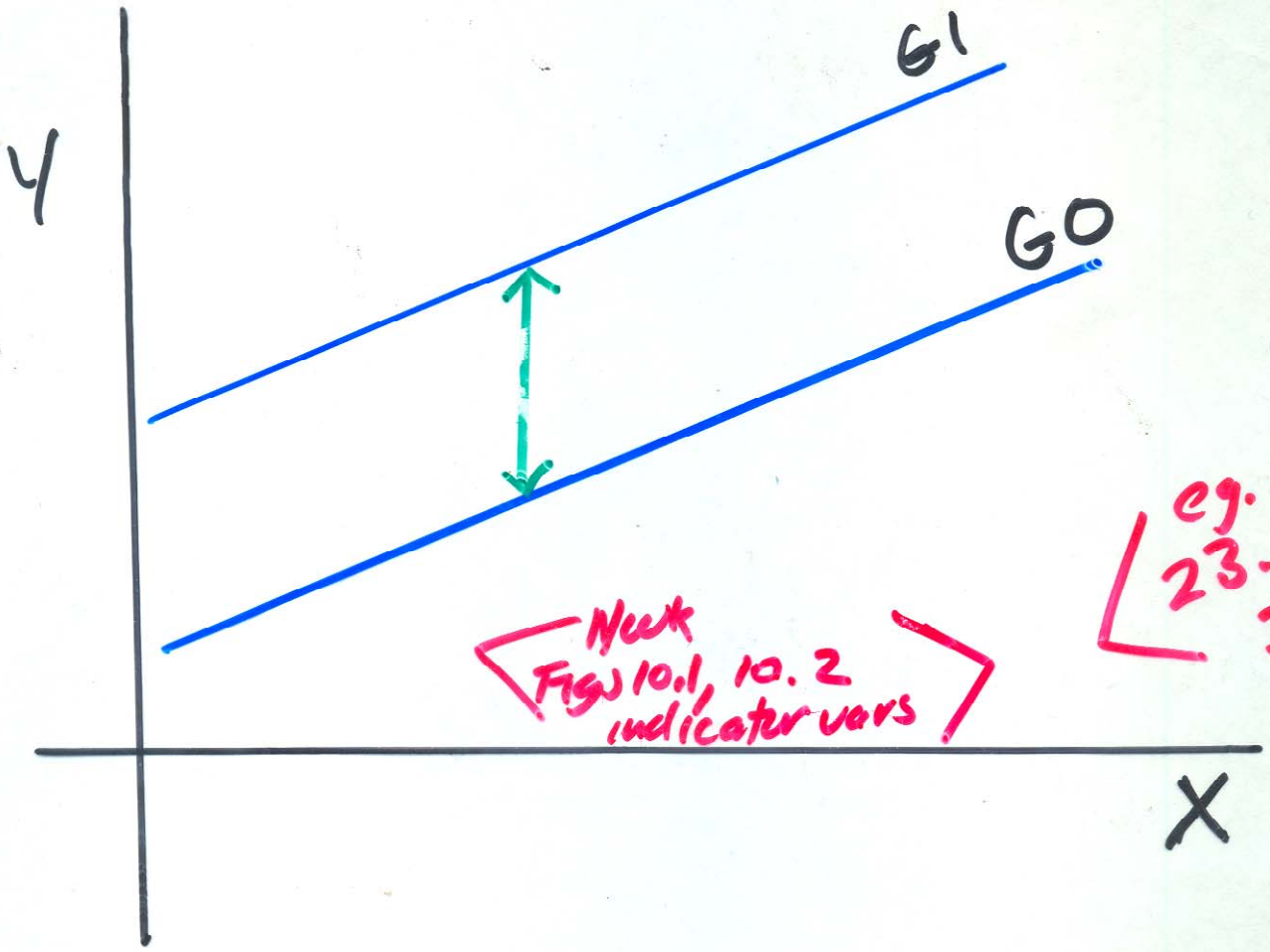
$$\hat{Y} = \bar{Y}_1 + \hat{\gamma}_p (X - \bar{X}_1)$$

$$\hat{Y} = \bar{Y}_0 + \hat{\gamma}_p (X - \bar{X}_0)$$

Adjusted mean difference is vertical distance between 11 lines

$$\bar{Y}_1 + \hat{\gamma}_p (X - \bar{X}_1) - [\bar{Y}_0 + \hat{\gamma}_p (X - \bar{X}_0)]$$

$$= \bar{Y}_1 - \bar{Y}_0 - \hat{\gamma}_p (\bar{X}_1 - \bar{X}_0)$$ ← ancova adjustment.



eg. Nwlk 23.3 Fig I=3

Nwlk Figs 10.1, 10.2 indicator vars

↕ adjusted mean difference
ancova estimate of treatment effect (G1 vs G0 membership)

<diff of adj treatment means Nwlk 23.40, Fig 23.9>

precision of group comparison (s.c.)
improved in ancova to the extent that Y|X more precise than $\bar{Y}$
(i.e. magnitude of $\sqrt{1-r^2_{xy}}$

Ancova via Regression w/ Group membership variables.

<KNNL 23.2>

For each individual we have Y, X, group membership ($G_1, \ldots G_{I-1}$)

Two groups $G=1, G=0$

General model:

product term.

$$E(Y|X,G) = \beta_0 + \beta_1 G + \beta_2 X + \beta_3 XG$$

<KNNL sec 10.2> <KNNL eq 10.7>

w/in groups

$$E(Y|X, G=1) = (\beta_0 + \beta_1) + (\beta_2 + \beta_3)X$$

$$E(Y|X, G=0) = \beta_0 + \beta_2 X$$

Thus difference between w/in group slopes $\beta_3 = (\gamma_1 - \gamma_0)$

Ancova assumes $\beta_3 = 0$

Thus ancova model is

$$E(Y|X,G) = \delta_0 + \delta_1 G + \delta_2 X$$

< Nwls eqs 10.4, 10.5 >

$\hat{\delta}_2 = \hat{\gamma}_p$ (common w/in group slope)

$\hat{\delta}_1$ = adjusted mean difference (ancova treatment effect)

Ancova $H_0: \delta_1 = 0$ $H_a: \delta_1 \neq 0$

inferences for δ_1

test statistic $\hat{\delta}_1 / s.e.(\hat{\delta}_1)$

interval estimate for δ_1.

famous observational data

HLM Example : 2-level linear model

Also Lab2

Introduction

In this example we analyze data from the *High School and Beyond* Survey. The data contain information on students' mathematical achievement. Additional information available includes the school sector (Catholic versus public) and school mean social class, denoted by the variables SECTOR and MEANSES respectively.

The 7185 level-1 units are the students, and at this level we use the social class (denoted by the variable SES) to model the math achievement. At level-2, each of the 160 schools' intercept and slope are predicted by school sector and school mean social class.

The information above implies that the level-1 model will have two coefficients for each student: the intercept (π_0) and the SES slope (π_1), as shown below:

$$\text{MATHACH} = \beta_0 + \beta_1 (\text{SES}) + r$$

At level-2, the intercept and SES slope are modeled as functions of sector and mean SES:

B-R reverse predictors

$$\beta_0 = \gamma_{00} + \gamma_{01}(\text{SECTOR}) + \gamma_{02}(\text{MEANSES}) + u_0$$
$$\beta_1 = \gamma_{10} + \gamma_{11}(\text{SECTOR}) + \gamma_{12}(\text{MEANSES}) + u_1$$

Note that both the intercept and the slope are modeled as having randomly varying residuals. The assumption is that the intercept and slope vary not only as a function of the two predictors SECTOR and MEANSES, but also as a function of a unique school effect.

Package 'MEMSS'

January 2, 2012

Version 0.9-0

Date 2011-02-15

Title Data sets from Mixed-effects Models in S

Author Douglas Bates <bates@stat.wisc.edu>, Martin Maechler <maechler@R-project.org> and Ben Bolker <bbolker@gmail.com>

Maintainer <lme4-authors@R-forge.wu-wien.ac.at>

Description Data sets and sample analyses from Pinheiro and Bates,''Mixed-effects Models in S and S-PLUS'' (Springer, 2000).

Depends R(>= 2.12.0), lme4 (>= 0.999375-36)

LazyData yes

License GPL (>= 2)

Repository CRAN

Repository/R-Forge/Project lme4

Repository/R-Forge/Revision 1265

Date/Publication 2011-02-17 13:06:45

R topics documented:

HSB data

From Lab2

```
> data(MathAchieve)
> MathAchieve[1:10,]
Grouped Data: MathAch ~ SES | School
   School Minority    Sex    SES MathAch MEANSES
1    1224       No Female -1.528   5.876  -0.428
2    1224       No Female -0.588  19.708  -0.428
3    1224       No   Male -0.528  20.349  -0.428
4    1224       No   Male -0.668   8.781  -0.428
5    1224       No   Male -0.158  17.898  -0.428
6    1224       No   Male  0.022   4.583  -0.428
7    1224       No Female -0.618  -2.832  -0.428
8    1224       No   Male -0.998   0.523  -0.428
9    1224       No Female -0.888   1.527  -0.428
10   1224       No   Male -0.458  21.521  -0.428
> dim(MathAchieve)
[1] 7185    6

  Public Catholic
    3642     3543
```

individ level file

```
> data(MathAchSchool)
> MathAchSchool[1:10,]
     School Size   Sector PRACAD DISCLIM HIMINTY MEANSES
1224   1224  842   Public   0.35   1.597       0  -0.428
1288   1288 1855   Public   0.27   0.174       0   0.128
1296   1296 1719   Public   0.32  -0.137       1  -0.420
1308   1308  716 Catholic   0.96  -0.622       0   0.534
1317   1317  455 Catholic   0.95  -1.694       1   0.351
1358   1358 1430   Public   0.25   1.535       0  -0.014
1374   1374 2400   Public   0.50   2.016       0  -0.007
1433   1433  899 Catholic   0.96  -0.321       0   0.718
1436   1436  185 Catholic   1.00  -1.141       0   0.569
1461   1461 1672   Public   0.78   2.096       0   0.683
> dim(MathAchSchool)
[1] 160   7
> table(MathAchSchool$Sector)
  Public Catholic
      90       70
```

school-level file

Lab2 creates a combined data set and then uses lme

finding your inner ancova

Linear Mixed Models

Appendix to *An R and S-PLUS Companion to Applied Regression*

John Fox

Stat 209 [redacted]

3.2 Fitting a Hierarchical Linear Model with lme

(HLM)

Following Bryk and Raudenbush (1992) and Singer (1998), I will fit a hierarchical linear model to the math-achievement data. This model consists of two equations: First, within schools, we have the regression of math achievement on the individual-level covariate SES; it aids interpretability of the regression coefficients to center SES at the school average; then the intercept for each school estimates the average level of math achievement in the school.

160 schools
7185 students

Using centered SES, the individual-level equation for individual j in school i is

Math ~ SES ← elitist?

$$\text{mathach}_{ij} = \alpha_{0i} + \alpha_{1i}\text{cses}_{ij} + \varepsilon_{ij} \quad (2)$$

M ~ SES egalitarian

ancova $\bar{Y}$

$\hat{\gamma}_{01} = 5.3$

$\hat{\gamma}_{02} = 1.22$

1) ancova on class means

2) ancova on slopes, Math/SES

ancova $\hat{\beta}_{slope}$

$\hat{\gamma}_{11} = 1.04$

$\hat{\gamma}_{12} = -1.64$

At the school level, also following Bryk and Raudenbush, I will entertain the possibility that the school intercepts and slopes depend upon sector and upon the average level of SES in the schools:

$\bar{Y}$ on S $\bar{SES}$

$$\begin{aligned}\alpha_{0i} &= \gamma_{00} + \gamma_{01}\text{meanses}_i + \gamma_{02}\text{sector}_i + u_{0i} \\ \alpha_{1i} &= \gamma_{10} + \gamma_{11}\text{meanses}_i + \gamma_{12}\text{sector}_i + u_{1i}\end{aligned} \quad (3)$$

q1 ↑ Public Catholic

This kind of formulation is sometimes called a *coefficients-as-outcomes* model.

Substituting the school-level equation 3 into the individual-level equation 2 produces

$$\begin{aligned}\text{mathach}_{ij} &= \gamma_{00} + \gamma_{01}\text{meanses}_i + \gamma_{02}\text{sector}_i + u_{0i} \\ &\quad +(\gamma_{10} + \gamma_{11}\text{meanses}_i + \gamma_{12}\text{sector}_i + u_{1j})\text{cses}_{ij} + \varepsilon_{ij}\end{aligned}$$

Rearranging terms,

Cath schools higher achievement more egalatarian (compensatory)

$$\begin{aligned}\text{mathach}_{ij} &= \gamma_{00} + \gamma_{01}\text{meanses}_i + \gamma_{02}\text{sector}_i + \gamma_{10}\text{cses}_{ij} \\ &\quad +\gamma_{11}\text{meanses}_i\text{cses}_{ij} + \gamma_{12}\text{sector}_i\text{cses}_{ij} \\ &\quad +u_{0i} + u_{1i}\text{cses}_{ij} + \varepsilon_{ij}\end{aligned}$$

combined model eq. Bryk sec 10.3 (10.15)

Here, the γ's are fixed effects, while the u's (and the individual-level errors ε_{ij}) are random effects.

Finally, rewriting the model in the notation of the linear mixed model (equation 1),

$$\begin{aligned}\text{mathach}_{ij} &= \beta_1 + \beta_2\text{meanses}_i + \beta_3\text{sector}_i + \beta_4\text{cses}_{ij} \\ &\quad +\beta_5\text{meanses}_i\text{cses}_{ij} + \beta_6\text{sector}_i\text{cses}_{ij} \\ &\quad +b_{i1} + b_{i2}\text{cses}_{ij} + \varepsilon_{ij}\end{aligned} \quad (4)$$

even though there are interactions galore in combined model, none in the Level 2 ancova models

---

`lme` (linear mixed effects) function in the `nlme` library, however, employs the *Laird-Ware form* of the linear mixed model (after a seminal paper on the topic published by Laird and Ware, 1982):

mixed effects $y = X\beta + Zb + \varepsilon$

$$\begin{aligned}y_{ij} &= \beta_1 x_{1ij} + \cdots + \beta_p x_{pij} \\ &\quad +b_{i1} z_{1ij} + \cdots + b_{iq} z_{qij} + \varepsilon_{ij} \\ b_{ik} &\sim \text{N}(0, \psi_k^2), \text{Cov}(b_k, b_{k'}) = \psi_{kk'} \\ \varepsilon_{ij} &\sim \text{N}(0, \sigma^2\lambda_{ijj}), \text{Cov}(\varepsilon_{ij}, \varepsilon_{ij'}) = \sigma^2\lambda_{ijj'}\end{aligned} \quad (1)$$

where

- y_{ij} is the value of the response variable for the jth of n_i observations in the ith of M groups or clusters.
- $\beta_1, \ldots, \beta_p$ are the fixed-effect coefficients, which are identical for all groups.
- $x_{1ij}, \ldots, x_{pij}$ are the fixed-effect regressors for observation j in group i; the first regressor is usually for the constant, $x_{1ij} = 1$.
- $b_{i1}, \ldots, b_{iq}$ are the random-effect coefficients for group i, assumed to be multivariately normally distributed. The random effects, therefore, vary by group. The b_{ik} are thought of as random variables, not as parameters, and are similar in this respect to the errors ε_{ij}.

HSB by ancova

[redacted] Stat 209


```
> #let's do the hsb ancova
> hsbancdat = read.table(file="D:\\drr09\\stat209\\hsbancova", header = T)
> summary(hsbancdat)
     school       Intercept        csesslp          sector     meanses
> attach(hsbancdat)
> tapply(Intercept, sector, summary)
$C   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   7.336  13.200  14.470  14.200  15.900  19.720
$P  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   4.240   9.719  11.710  11.390  13.200  18.110
> tapply(csesslp, sector, summary)
$C    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  -2.0150  0.5698  1.5230  1.4680  2.4600  5.2580
$P   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  -1.014   1.695   2.922   2.772   3.824   6.266

> #initial differences on covariate?
> tapply(meanses, sector, summary)
$C   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
-0.7619 -0.1039  0.2388  0.1601  0.4346  0.8250
$P   Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
-1.19400 -0.40090 -0.09058 -0.13550  0.11370  0.68200
> #Cath higher on school-level SES
> # do the ancova on school-level outcomes (level, slope)
> hsbancdat$gr = 2 - as.numeric(sector) # code Cath = 1, Pub = 0 on sector
> intancova = lm(Intercept ~ gr + meanses)
> summary(intancova)
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  12.1195     0.2026  59.807  < 2e-16 ***
gr            1.2219     0.3169   3.855 0.000168 ***
meanses       5.3874     0.3810  14.140  < 2e-16 ***
---
Residual standard error: 1.859 on 157 degrees of freedom
Multiple R-squared: 0.6489,     Adjusted R-squared: 0.6444
F-statistic: 145.1 on 2 and 157 DF,  p-value: < 2.2e-16

> # compare with hlm/lme coeffs gr vs 1.226 (.306) df 157 t = 4.00
> # compare with hlm/lme coeffs meanses vs 5.33 (.369) df 157 t = 14.4

> slpancova = lm(csesslp ~ gr + meanses)
> summary(slpancova)
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   2.8886     0.1600  18.049  < 2e-16 ***
gr           -1.5580     0.2503  -6.224 4.22e-09 ***
meanses       0.8612     0.3009   2.862  0.00478 **
---
Residual standard error: 1.468 on 157 degrees of freedom
Multiple R-squared: 0.1999,     Adjusted R-squared: 0.1897
F-statistic: 19.61 on 2 and 157 DF,  p-value: 2.492e-08

> # compare with hlm/lme coeffs gr vs -1.64 (.239) t = -6.85
> # compare with hlm/lme coeffs meanses vs 1.03 (.299) df 157 t = 3.48

> # look at comparing regressions; neither outcome refutes parallel regressions
> intcnrl = lm(Intercept ~ gr + meanses + I(gr*meanses))
Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)      12.1825     0.2075  58.714  < 2e-16 ***
gr                1.2487     0.3168   3.942 0.000122 ***
meanses           5.8524     0.5139  11.388  < 2e-16 ***
I(gr * meanses)  -1.0261     0.7634  -1.344 0.180855
---
> slpcnrl = lm(csesslp ~ gr + meanses + I(gr*meanses))
Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)      2.9122     0.1646  17.692  < 2e-16 ***
gr              -1.5480     0.2513  -6.160 5.92e-09 ***
meanses          1.0351     0.4077   2.539   0.0121 *
I(gr * meanses) -0.3837     0.6056  -0.634   0.5272
---
```

Cath higher on $\bar{Y}$

Cath lower on $\hat{\alpha}_1$ school slope

so ancova will adjust a little bit

make 0,1 group var

school level

$\hat{\delta}_{01}$ close

alternative: individual level ancova? Y_{ij}, G, sess, csesa

$\hat{\delta}_{11}$ school slope

a little different, due to weighting.

dependence of scatter effect on SES

Comparing Nonparallel Regr Lines

Rogosa (1980) <see NWK p. ~~948~~ 1034 prob 23.1>

2 group comparison single background variable (X), outcome (Y)

G = 1, 0

Fit separate within group regr's via

$$E(Y|G,X) = \beta_0 + \beta_1 G + \beta_2 X + \beta_3 X*G$$

treatment effect as a function of X:

$$\Delta(X) = \beta_1 + \beta_3 X$$

draw inferences about $\Delta(X)$

ATI research (TTI)

(ancova assumes/requires $\beta_3 = 0$, why do it?)

point estimate $D(X) = \hat{\beta}_1 + \hat{\beta}_3 X$

<see cnrl.lis, .dat>
CNRL paper online

[redacted]

CNRL Math Notes

model $Y = \beta_1 + \beta_2 G + \beta_3 X + \beta_4 XG + \varepsilon$

Grp 1 $E(Y|X, G=1) = \beta_1 + \beta_2 + (\beta_3 + \beta_4)X$

Grp 0 $E(Y|X, G=0) = \beta_1 + \beta_3 X$

Treatment effect (diff of regressions) $\Delta(X) = \beta_2 + \beta_4 X$ — func of X iff $\beta_4 \neq 0$

abscissa of point of intersection (w/in group regressions) "cut-off" $X^0 = -\beta_2/\beta_4$

ATI research assignment on "aptitude" to differential instruction.

Inference

From sample obtain estimates $\hat{\beta}_i$ $i = 1, \ldots, 4$

$D(X) = \hat{\beta}_2 + \hat{\beta}_4 X$ $\quad s_{ij}$, elements of $\hat{Cov}(\hat{\beta}_i)$ 4×4

$\hat{X}^0 = -\hat{\beta}_2/\hat{\beta}_4$ ratio estimator (biased)

inference for $\Delta(X)$ pick-a-point

$D(C_a)$ ancova treatment effect; $D(\bar{X}_G)$ average treatment effect

$D(\bar{X}_{sub})$ average treatment effect for subgroup

sampling variance $s^2_{D(X)} = s_{22} + s_{24}C_a + s_{44}(X - C_a)^2$, where $C_a = -s_{24}/s_{44}$

$\Delta(X)$ inference: $D(X)/s_{D(X)}$ via t-distrib N-4 df, usual CI

Inference for $\Delta(X)$ J-N region of significance

X-values s.t. reject $\Delta(X) = 0$

R non-simultaneous $D(X) \pm \sqrt{F^{\alpha}_{1,N-4}\, s^2_{D(X)}}$ concatenate $D(X)/s_{D(X)}$ tests or CI

R' proper simultaneous Working-Hotelling band about $D(X)$

hyperbolas $D(X) \pm \sqrt{2F^{\alpha}_{2,N-4}\, s^2_{D(X)}}$ $\quad (R > R')$

R' {values on X-axis outside WH band} Potthoff 1966

Psychological Bulletin
1980, Vol. 88, No. 2, 307-321

Comparing Nonparallel Regression Lines

David Rogosa
Department of Education, University of Chicago

Statistical comparisons of experimental groups frequently are based on the within-group regressions of an outcome variable on a concomitant variable. I present a comprehensive strategy for the statistical comparison of within-group regressions that is suitable for both parallel and nonparallel regression lines. New results are obtained and new interpretations are formulated for standard statistical procedures such as analysis of covariance. New procedures for comparing regressions are proposed and illustrated.

The setting for this investigation is an experiment in which two groups are given alternative treatments. The groups are formed by random assignment of individual cases to the groups. One of these groups may be considered a control group, as in a study to evaluate the effectiveness of a curricular innovation relative to the effectiveness of a standard curriculum. The purpose of the experiment is to assess the treatment effect—the differential effectiveness of the two treatments.

A number of initial characteristics (usually called covariates or predictor variables) are recorded for each case before the initiation of the treatment; outcome variables are measured at the end of the treatment period. I limit consideration to one outcome variable (Y) and one predictor variable (X). The raw data from this experiment are measures of X and Y that are obtained from the members of the two experimental groups.

The raw data are summarized through estimation of the within-group regression lines, the average outcome for a given initial characteristic value within each group. The summary of the raw data consists of (a) the separate sample within-group regression lines and (b) estimates of the sampling variances and covariances associated with these sample regressions.

Requests for reprints should be sent to David Rogosa, who is now at the School of Education, Stanford University, Stanford, California 94305.

This article presents and evaluates statistical procedures for comparing the within-group regression lines. I view the treatment effect as a function of X and define the treatment effect as the difference between the population regression lines. The usual dichotomy between parallel and nonparallel regressions is shunned. The difference between the sample regression lines is used to estimate this treatment effect, and this estimate will depend on X (to some degree) whenever the sample within-group regressions are not parallel. Although this approach seems natural for problems of comparing regressions, the development and exposition of traditional statistical methods has proceeded along other lines.

Two types of assessments of the treatment effect are sought. First, an overall treatment effect, in which dependence on X of the treatment effect is ignored, can be estimated by evaluating the difference between the sample regression lines at a prespecified value of X. These procedures are called pick-a-point procedures. Second, an assessment of the difference between the regressions over the entire range of X can be used to evaluate the treatment effect as a function of X. The Johnson-Neyman technique as extended by Potthoff (1964) is one procedure in which the regressions are compared over the range of X. Nonsimultaneous inference procedures are associated with the first type of assessment, and simultaneous inference procedures are associated with the second type of assessment.

 0033-2909/80/8802-0307$00.75

[redacted]

CNRL Math Notes

model $Y = \beta_1 + \beta_2 G + \beta_3 X + \beta_4 XG + \varepsilon$

Grp 1 $E(Y|X, G=1) = \beta_1 + \beta_2 + (\beta_3 + \beta_4)X$

Grp 0 $E(Y|X, G=0) = \beta_1 + \beta_3 X$

Treatment effect (diff of regressions) $\Delta(X) = \beta_2 + \beta_4 X$ — func of X iff $\beta_4 \neq 0$

abscissa of point of intersection (w/in group regression) "cut-off" $X^0 = -\beta_2/\beta_4$

ATI research assignment on "aptitude" to differential instruction.

Inference

From sample obtain estimates $\hat{\beta}_i$ $i = 1, \ldots, 4$

$D(X) = \hat{\beta}_2 + \hat{\beta}_4 X$

s_{ij}, elements of $\widehat{Cov}(\hat{\beta}_i)$ 4×4

$\hat{X}^0 = -\hat{\beta}_2/\hat{\beta}_4$ ratio estimator (biased)

Inference for $\Delta(X)$ pick-a-point

$D(C_a)$ ancova treatment effect; $D(\bar{X}_G)$ average treatment effect

$D(\bar{X}_{sub})$ average treatment effect for subgroup

sampling variance $s^2_{D(X)} = s_{22} + s_{24}C_a + s_{44}(X - C_a)^2$, where $C_a = -s_{24}/s_{44}$

$\Delta(X)$ inference: $D(X)/s_{D(X)}$ via t-distrib N-4 df, usual CI

Inference for $\Delta(X)$ J-N region of significance

X-values s.t. reject $\Delta(X) = 0$

R non-simultaneous $D(X) \pm \sqrt{F^{\alpha}_{1,N-4}\, s^2_{D(X)}}$ — concatenate $D(X)/s_{D(X)}$ tests or CI

R' proper simultaneous Working-Hotelling band about $D(X)$

hyperbolas $D(X) \pm \sqrt{2F^{\alpha}_{2,N-4}\, s^2_{D(X)}}$ $(R > R')$

R' {values on X-axis outside WH band} Potthoff 1966

more general model (CNRL)

see Berk text ATI intros

(B) $Y = \beta_1 + \beta_2 G + \beta_3 X + \beta_4 XG + e$

"interaction" term

diff regression lines (cf Berk) w/in groups

$\hat{\beta}_2$, careful, diff of lines at $X=0$

$$D(X) = \beta_2 + \beta_4 X$$

treatment effect as function of X sample estimate

(C) $$\hat{D}(X) = \hat{\beta}_2 + \hat{\beta}_4 X$$

sampling variance $S^2_{\hat{D}(X)} = S^2_{\hat{D}(C_a)} + S_{44}(X - C_a)^2$ (G)

$\therefore \mathrm{Cov}(\hat{\beta}_i, \hat{\beta}_j)$

s_{ij}

(D) $C_a = -\frac{S_{24}}{S_{44}}$ center of accuracy

min

$\hat{D}(C_a) = \underset{\hat{\gamma}_1}{\text{ancova}}$ (E) ; $S^2_{\hat{D}(C_a)} = S_{22} + S_{24} C_a$

See reverse for more careful listing

$\hat{D}(X)/S_{\hat{D}(X)}$ pick-a-point

Regions of significance, J-N

(H) $D(X)$ "significantly" different from 0

simultaneous R' $2F_{2,n-4}$

cf CNRL paper p318 R' working-hotelling

Comparing Nonparallel Regression Lines

Stat209

```
> cnrldat = read.table(file="D:\\drr06\\stat209\\cnrl.dat", header = T)
> cnrldat
      Y    X G
01 2.23 0.28 1
02 4.99 0.97 1
03 3.37 1.25 1
04 8.54 2.46 1
05 8.40 2.51 1
06 3.70 1.17 1
07 7.93 1.78 1
08 2.43 1.21 1
09 5.40 1.63 1
10 8.44 1.98 1
11 3.25 2.36 0
12 5.30 2.11 0
13 1.39 0.45 0
14 4.69 1.76 0
15 6.56 2.09 0
16 3.00 1.50 0
17 5.85 1.25 0
18 1.90 0.72 0
19 3.85 0.42 0
20 2.95 1.53 0
```

$n_1 = 10$, $n_2 = 10$

Rogosa 1980

quick description

```
> tapply(Y,G, summary)
$"0"
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  1.390   2.963   3.550   3.874   5.148   6.560
$"1"
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  2.230   3.452   5.195   5.543   8.283   8.540
> tapply(X,G, summary)
$"0"
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.4200  0.8525  1.5150  1.4190  2.0070  2.3600
$"1"
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  0.280   1.180   1.440   1.524   1.930   2.510

> cor(X,G)   #point-biserial correlation
[1] 0.07939952
> cor.test(X,G)
        Pearson's product-moment correlation
data:  X and G
t = 0.3379, df = 18, p-value = 0.7393
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval: -0.3763445  0.5042046
sample estimates:  cor 0.07939952
```

no apparent or significant pre-existing diffs

standard ancova

link w/ math (A)

```
> cnrlancova = lm(Y ~ G + X) #ancova
> summary(cnrlancova)
Call:  lm(formula = Y ~ G + X)
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.6120     0.8870   0.690 0.499578
G             1.4276     0.6909   2.066 0.054391 .
X             2.2988     0.5225   4.400 0.000391 ***
---
Residual standard error: 1.54 on 17 degrees of freedom
Multiple R-Squared: 0.5975,     Adjusted R-squared: 0.5501
F-statistic: 12.62 on 2 and 17 DF,  p-value: 0.0004374
```

$\hat{\gamma}_1$ $\hat{\delta}_2$

```
> # just fails significance-- pooled slope 2.3
```

$\hat{\beta}_p = ?$

CNRL paper: plot, w/in group slopes

```
> #look at within-group slopes
> regrG1 = lm(Y[G==1] ~ X[G==1])
> coef(regrG1)
(Intercept)   X[G == 1]
  0.4971152   3.3109480
> regrG0 = lm(Y[G==0] ~ X[G==0])
> coef(regrG0)
(Intercept)   X[G == 0]
   2.010282    1.313402
> #within-group regressions rather different, significant (n=10)?
```

$\hat{\beta}_p$ 2.3 in between

```
> cnrlreg = lm(Y ~ G + X + I(X*G)) #allows two separate regression lines
> summary(cnrlreg)
Call: lm(formula = Y ~ G + X + I(X * G))
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   2.0103     1.0501   1.914   0.0736 .
G            -1.5132     1.5403  -0.982   0.3405
X             1.3134     0.6704   1.959   0.0677 .
I(X * G)      1.9975     0.9544   2.093   0.0527 .
---
Residual standard error: 1.407 on 16 degrees of freedom
Multiple R-Squared: 0.684,      Adjusted R-squared: 0.6247
F-statistic: 11.54 on 3 and 16 DF,  p-value: 0.0002816
```

(B)

not signif, but sample slopes differ

Using R's model formula notation

The model formula notation that R uses allows this to be done in a systematic manner. It is a bit confusing to learn, but this flexible notation is used by most of R's more advanced functions.

To illustrate, the above could be done by (if the data frame PlantGrowth is attached)

```
> boxplot(weight ~ group)
```

What does this do? It breaks the weight variable down by values of the group factor and hands this off to the boxplot command. One should read the line weight ~ group as "model weight *by* the variable group". That is, break weight down by the values of group.

When there are two variables involved things are pretty straightforward. The response variable is on the left hand side and the predictor on the right:

response ~ predictor (when two variables).

When there are more than two predictor variables things get a little confusing. In particular, the usual mathematical operators do not do what you may think. Here are a few different possibilities that will suffice for these notes.[9]

Suppose the variables are generically named Y, X1, X2

formula	meaning
Y ~ X1	Y is modeled by X1
Y ~ X1 + X2	Y is modeled by X1 and X2 as in multiple regression
Y ~ X1 * X2	Y is modeled by X1, X2 *and* X1*X2
(Y ~ (X1 + X2)^2)	Two-way interactions. Note usual powers
Y ~ X1+ I((X2^2)	Y is modeled by X1 and $X2^2$
Y ~ X1 | X2	Y is modeled by X1 conditioned on X2

The exact interpretation of "modeled by" varies depending upon the usage. For the boxplot command it is different than the lm command. Also notice that usual mathematical meanings are available, but need to be included inside the I function.

Ways to view multivariate data

Now that we can store and access multivariate data, it is time to see the large number of ways to visualize the datasets.

***n*-way contingency tables** Two-way contingency tables were formed with the table command and higher order ones are no exception. If w,x,y,z are 4 variables, then the command table(x,y) creates a two-way table, table(x,y,z) creates two-way tables x versus y for each value of z. Finally x,y,z,w will do the same for each combination of values of z and w. If the variables are stored in a data frame, say df then the command table(df) will behave as above with each variable corresponding to a column in the given order.

To illustrate let's look at some relationships in the dataset Cars93 found in the MASS library.

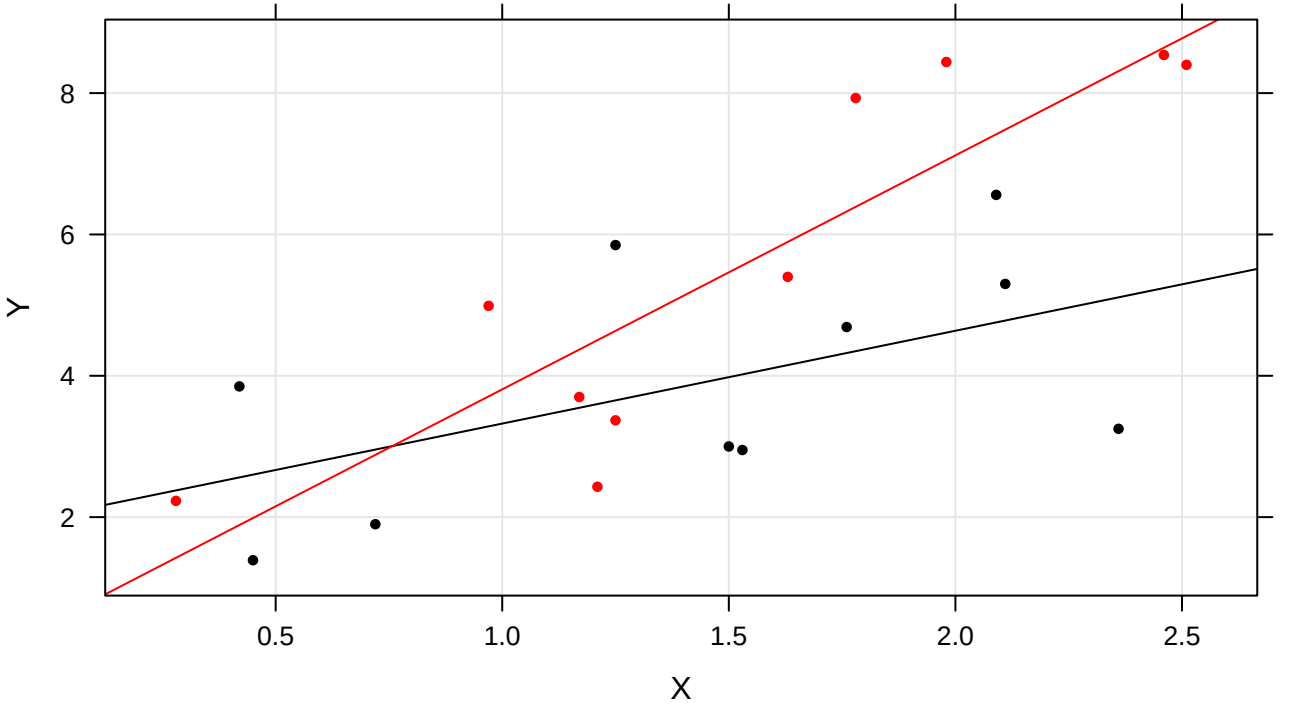
```
> library(MASS);data(Cars93);attach(Cars93)
## make some categorical variables using cut
> price = cut(Price,c(0,12,20,max(Price)))
> levels(price)=c("cheap","okay","expensive"))
> mpg = cut(MPG.highway,c(0,20,30,max(MPG.highway)))
> levels(mpg) = c("gas guzzler","okay","miser"))
## now look at the relationships
> table(Type)
Type
Compact   Large Midsize   Small  Sporty     Van
     16      11      22      21      14       9
> table(price,Type)
           Type
```

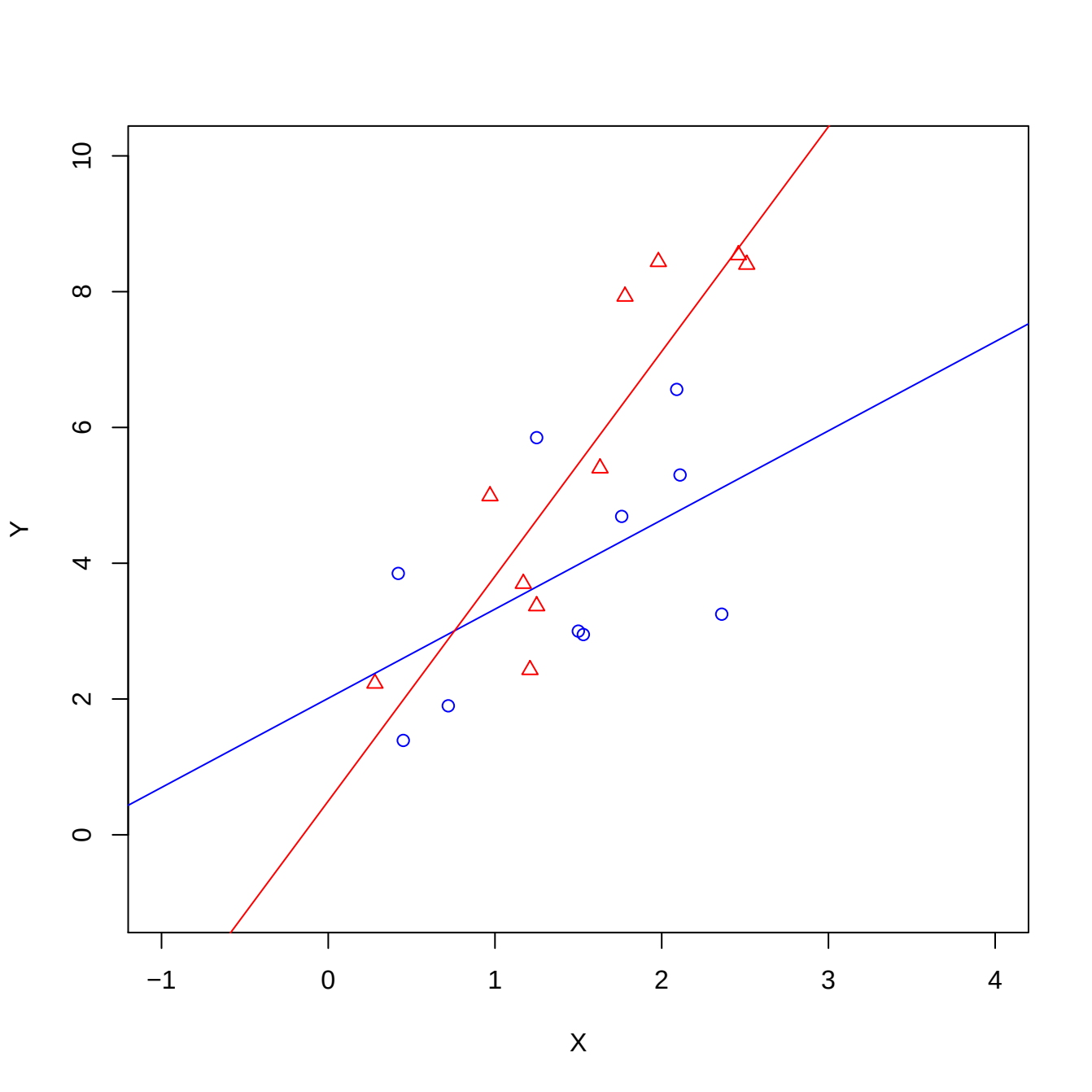
[9] A thorough explanation of the syntax and its usage is found in the manual "An Introduction to R" which accompanies the R software, and the contributed document "Using R for Data Analysis and Graphics" by Maindonald. See the appendix for more information on these.

```
R version 3.0.1 (2013-05-16) -- "Good Sport"
Copyright (C) 2013 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)



> dat = read.table("http://www-stat.stanford.edu/~rag/stat209/cnrl.dat",header=T)
Error in file(file, "rt") : cannot open the connection
> dat = read.table("http://statweb.stanford.edu/~rag/stat209/cnrl.dat",header=T)
> library(lattice)
> head(dat)
      Y    X G
01 2.23 0.28 1
02 4.99 0.97 1
03 3.37 1.25 1
04 8.54 2.46 1
05 8.40 2.51 1
06 3.70 1.17 1
> xyplot(Y ~ X, groups = G, dat, type = c("g", "p", "r"), aspect = .5, col = c("black", "red"))
> xyplot(Y ~ X, groups = G, dat, type = c("g", "p", "r"), aspect = .5, col = c("black", "red"), pch=20)
>
```


Y
8
6
4
2
0.5
1.0
1.5
2.0
2.5
X


Y
X
10
8
6
4
2
0
−1
0
1
2
3
4

CNRL (Rogosa 1980) p.2

Stat 209 [redacted]

```
> coef(cnrlreg)
(Intercept)           G           X      I(X * G)
   2.010282   -1.513167    1.313402    1.997546
> vcov(cnrlreg)
            (Intercept)          G          X   I(X * G)
(Intercept)   1.1027356 -1.1027356 -0.6376948  0.6376948
G            -1.1027356  2.3726001  0.6376948 -1.3411185
X            -0.6376948  0.6376948  0.4493973 -0.4493973
I(X * G)      0.6376948 -1.3411185 -0.4493973  0.9109614
```

$\hat{\beta}_i$ $i = 1, \ldots, 4$

s_{ij} $i, j = 1, \ldots, 4$

```
> # comparing regressions analysis
> DX = coef(cnrlreg)[2] + coef(cnrlreg)[4]*X
> Ca = -vcov(cnrlreg)[2,4]/vcov(cnrlreg)[4,4]
> Ca
[1] 1.472201
> mean(X)
[1] 1.4715
> DCa = coef(cnrlreg)[2] + coef(cnrlreg)[4]*Ca
> DCa
1.427622
> coef(cnrlancova)
(Intercept)           G           X
  0.6119525   1.4276223   2.2988354
```

(C) (D) (E)

center and difference at regressions at center
Average treatment effect

get variances of D(X) for inference

```
> VarDCa = vcov(cnrlreg)[2,2] + vcov(cnrlreg)[2,4]*Ca
> VarDCa
[1] 0.3982038
> DCa/sqrt(VarDCa)  #compare to t(16)
2.262354
> qt( .975, 16)
[1] 2.119905
> #pooling hurts even when not significant
> VarDX = VarDCa + vcov(cnrlreg)[4,4]*(X - Ca)^2  #variance for pick-a-point procedures
```

(F)

quasi-ancova beats, Fisher

(G)

for R' $qF(.95, 2, 16) = 3.6337$

mathematica

```
In[2]:= d[x_] := -1.513167 + 1.997546*x
```

D(X)

```
In[4]:= vard[x_] := 0.3982038 + 0.9109614*(x - 1.472201)^2

In[5]:= NSolve[(d[x])^2 - 2*3.633723*vard[x] == 0, x]

Out[5]= {{x → 1.63438}, {x → 3.4785}}
```

$\partial^2 D(x)$

end points of R'

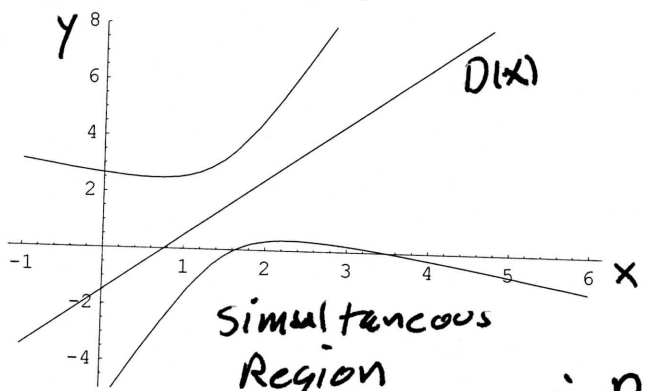
find J-N regions (simultaneous)

equiv to CI for X^0 intersection pt (Fieller's thm)

single part R' because

$\hat{\beta}_{4*}$ not big enough

(H)



Simultaneous Region

ascending powers of x coeff(x^0,x^1,x^2)

in R use

```
> polyroot(c(-14.953, 13.448, -2.6302))
[1] 1.634306+0i 3.478613-0i
```

or uniroot w/ functions

analysis continued in webposting

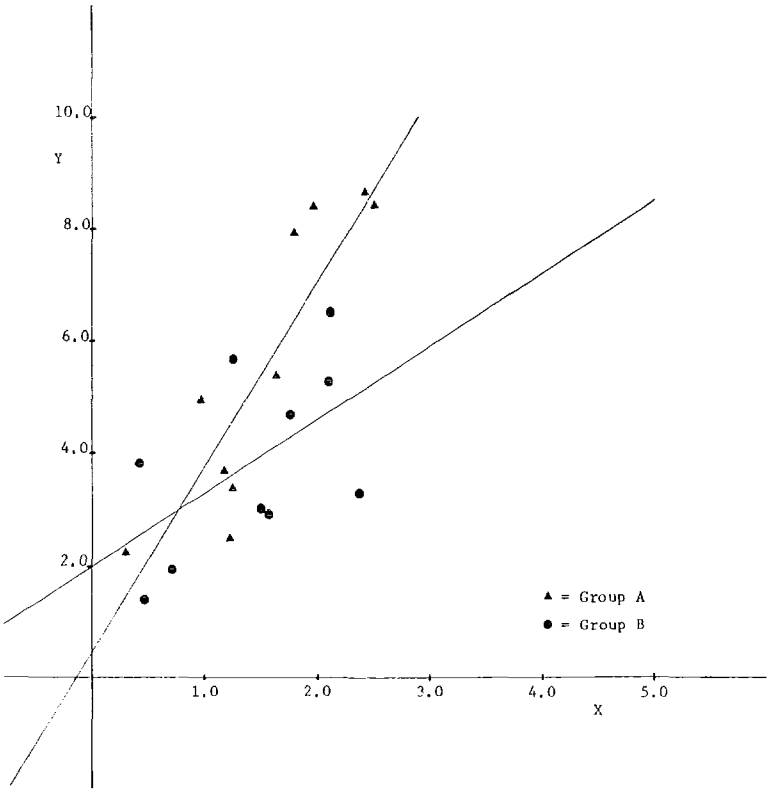


Figure 1. Plot of the data in Table 1 and the sample within-group regression lines.

culations. The end points of R' are the roots of the equation $[D(X)]^2 - 2F^{\alpha}(2, N-4)s^2_{D(X)} = 0$. This equation can be expanded as a quadratic equation in X:

$$[\hat{\beta}_4^2 - 2F^{\alpha}(2, N-4)s_{44}]X^2 + 2[\hat{\beta}_2\hat{\beta}_4 - 2F^{\alpha}(2, N-4)s_{24}]X + [\hat{\beta}_2^2 - 2F^{\alpha}(2, N-4)s_{22}] = 0. \quad (24)$$

Substitution of numerical values into Equation 24 yields an equation of the form $AX^2 + 2BX + C = 0$. If R' exists, the two distinct real roots of this equation are

$$\frac{-B \pm \sqrt{B^2 - AC}}{A}.$$

use polyroot

Denote these roots as X_+ for the larger root and X_- for the smaller root. When $\hat{\beta}_4^2/s_{44} > 2F^{\alpha}(2, N-4)$, R' is composed of the X values: $X > X_+$, $X < X_-$. When $\hat{\beta}_4^2/s_{44} < 2F^{\alpha}(2, N-4)$, R' is composed of the X values: $X_- < X < X_+$ (Rogosa, in press). Alternatively, a graphical procedure for R' is to plot the confidence functions in Expression 21 and find the X values outside the intersection of the X-axis and the simultaneous confidence band for $Y = \Delta(X)$.

The data in Table 1 illustrate these procedures. Summary statistics for each group are also presented in Table 1. The data and the within-group regression lines are plotted in Figure 1. The sample within-group regressions for Groups A and B, respectively, are

$$Y = .497 + 3.31X,$$
$$Y = 2.01 + 1.31X.$$

The sample regressions intersect at the point (.756, 3.00); $\widehat{X^0}(= -\hat{\beta}_2/\hat{\beta}_4)$ is larger than 4 of the 20 X values.

The difference of the sample regressions is

$$D(X) = -1.51 + 2.00X.$$

In these data $\bar{X}_G = C_a = 1.47$, and thus $D(\bar{X}_G) = D(C_a) = 1.43$. The estimated variance of $D(X)$ is $s^2_{D(X)} = .398 + .910(X$

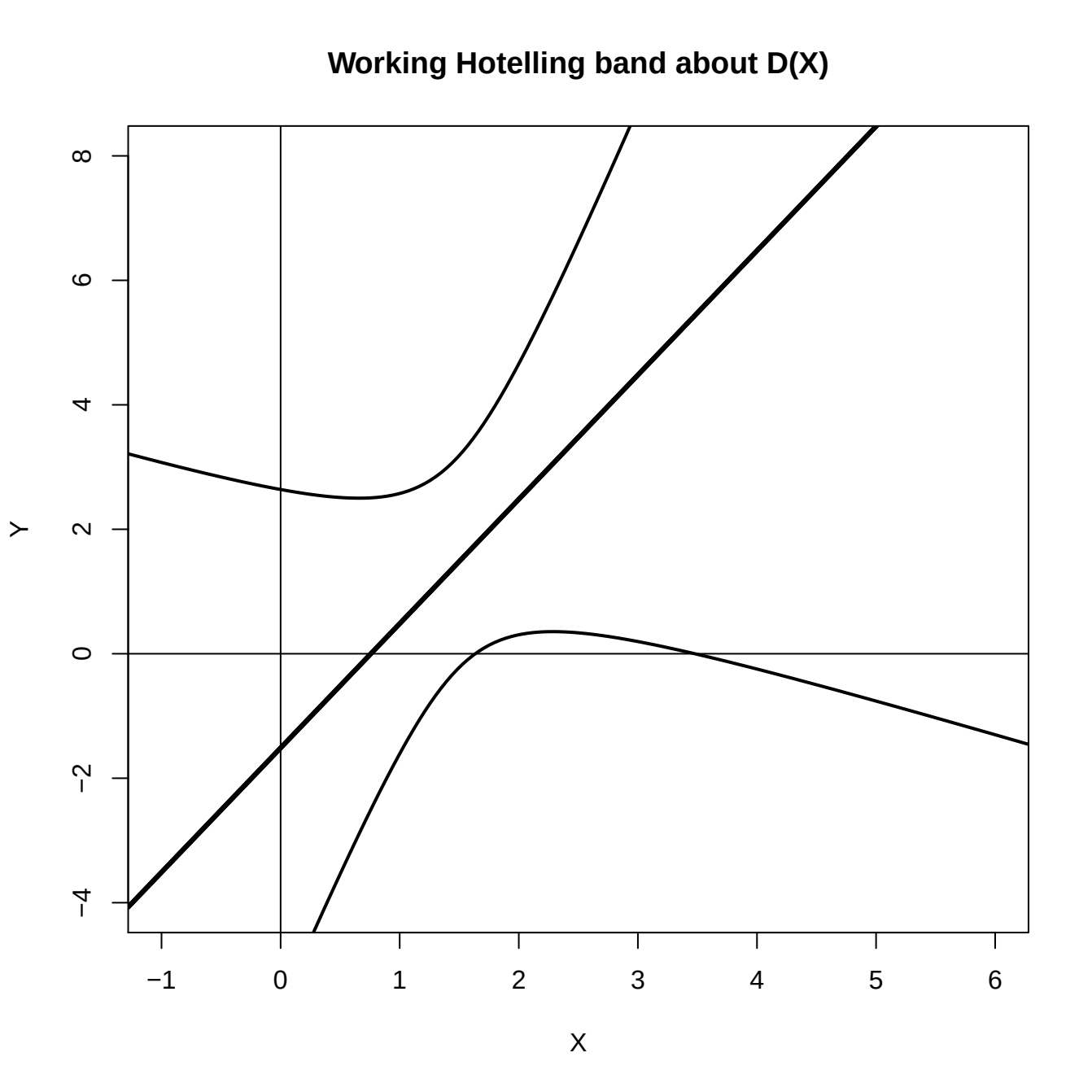
```
#redo cnrl analysis doing region of significance and plots in R

> dat = read.table("http://www-stat.stanford.edu/~rag/stat209/cnrl.dat",header=T)
> attach(dat)
> n = length(X)
> cnrl = lm(Y ~ G + X + I(X*G))
>
>
> vcov(cnrl)
            (Intercept)          G          X   I(X * G)
(Intercept)   1.1027356 -1.1027356 -0.6376948  0.6376948
G            -1.1027356  2.3726001  0.6376948 -1.3411185
X            -0.6376948  0.6376948  0.4493973 -0.4493973
I(X * G)      0.6376948 -1.3411185 -0.4493973  0.9109614
> Ca = -vcov(cnrl)[2,4]/vcov(cnrl)[4,4]
> Ca
[1] 1.472201
> DCa = coef(cnrl)[2] + coef(cnrl)[4]*Ca
> DCa
       G
1.427622
> #simultaneous region of significance
> #do this in R, solve quadratic equation p.318 CNRL paper
>
> qf(.95, 2, n-4)
[1] 3.633723
>
> cf = coef(cnrl)
> vc = vcov(cnrl)
> cF = 2*qf(.95, 2, n-4)
>
> #?polyroot
> #polyroot(z)
> #Arguments: z: the vector of polynomial coefficients in increasing order.
>
> cP = c(cf[2]^2 - cF*vc[2,2], 2*(cf[2]*cf[4] - cF*vc[2,4]), cf[4]^2 - cF*vc[4,4])
> cP
         G          G   I(X * G)
-14.953071  13.447774  -2.630175
> polyroot(cP) #endpoints in R'
[1] 1.634383+0i 3.478500-0i
>
> #Simultaneous region: values between 1.63 and 3.47
> #range of X measurements approx (0.25, 2.5)

> summary(X)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  0.280   1.120   1.515   1.472   2.007   2.510
> plot(NA,xlim=c(-1,6),ylim=c(-4,8),xlab="X",ylab="Y",main="Working Hotelling band abou
> abline(h=0)
> abline(v=0)

> VarDCa = vcov(cnrl)[2,2] + vcov(cnrl)[2,4]*Ca
> VarDCa
[1] 0.3982038
> lm0 = lm(Y ~ X, subset=(G==0))
> lm1 = lm(Y ~ X, subset=(G==1))
>
> abline(lm1$coef[1] - lm0$coef[1], lm1$coef[2] - lm0$coef[2],lwd=3)
```

```
> temp.seq = seq(-2,7,length=200)
>
>
> dx = (lm1$coef[1] - lm0$coef[1]) + temp.seq*(lm1$coef[2] - lm0$coef[2])
> band = sqrt(2*qf(0.95, 2, n-4) * (VarDCa + vcov(cnrl)[4,4]*(temp.seq - Ca)^2))
> lines(temp.seq, dx - band, lwd=2)
> lines(temp.seq, dx + band, lwd=2)
> # save your plot window as a pdf before closing with #dev.off()
```


Working Hotelling band about D(X)
Y
X
8
6
4
2
0
−2
−4
−1
0
1
2
3
4
5
6