Challenge the Statistics

Example for Today: Does Money Make People Happy?

Answer 1: A statistical test reveals no difference in happiness between people in the most common income categories.

Answer 2: A statistical test shows large difference in happiness by income

BOTH ARE CORRECT (SORT OF)

BOTH COME FROM THE SAME DATA

HOW?

Using the questions in this section, I'll discuss how we can resolve this apparent discrepancy

Objective Confirmation of Subjective Measures of Human Well-Being: Evidence from the U.S.A.

ANDREW J. OSWALD AND STEPHEN WU Authors Info & Affiliations

SCIENCE • 17 Dec 2009 • Vol 327, Issue 5965 • pp. 576-579 • DOI: 10.1126/science.1180606

The Best Things in Life Are Free?

Does money buy happiness? Answers to this question differ, depending, in part, on whether one asks an economist or a psychologist. The former would point to correlations between higher incomes and greater self-reported well-being, whereas the latter would argue that happiness shows little correlation with absolute material goods and is instead dictated largely by an individual's so-called set-point. Another strand of research invokes a hedonic treadmill, whereby income matters until subsistence requirements are met, at which point comparisons with one's neighbors are what influence one's sense of life satisfaction. **Oswald and Wu** (p. 576, published online 17 December; see the Perspective by **Layard**) establish that the subjective responses from 1 million adults, collected within health surveys

JOURNAL ARTICLE

Income and Happiness: Towards a Unified Theory

Richard A. Easterlin

The Economic Journal, Volume 111, Issue 473, July 2001, Pages 465–484,
https://doi.org/10.1111/1468-0297.00646
Published: 09 October 2008

Abstract

Material aspirations are initially fairly similar among income groups; consequently more income brings greater happiness. Over the life cycle, however, aspirations grow along with income, and undercut the favourable effect of income growth on happiness, although the cross-sectional happiness-income difference persists. People think they were less happy in the past and will be happier in the future, because they project current aspirations to be the same throughout the life cycle, while income grows. But since aspirations actually grow along with income, experienced happiness is systematically different from projected happiness. Consequently, choices turn out to be based on false expectations.

Challenging the Statistics

What is the context for these statistics?

What is the sample size?

What are you testing?

What is the null hypothesis?

What is the significance level?

How many tests are you doing?

Can I see the confidence intervals?

Is this practically significant?

Are you assuming causality?

What Is the Context for These Statistics?

How were the data collected?

How does this compare to the overall average?

What was the response rate?

Where else do we see effects in this range?

How much variability is there in the independent and dependent variables?

What is the sample size?

Sample size by itself is never enough

Population vs. Sample

Key question: was there randomization in sampling?

Who was dropped from this analysis? Why?

What are you testing?

Surprisingly difficult to answer

Don’t rely on broad concepts-- insist on specificity

Align tests with research questions

Specificity about exact test as well: why use this one?

What is the null hypothesis?

The null hypothesis is *posited*

While we take it for granted, it was never meant to be used this way (e.g. default assumption of a 0 difference)

Remind everyone to keep to the logic of inference

Reject the null

Fail to reject the null

NEVER accept the null

What is the significance level?

Significance is a measure of risk in repeated samples

Example: alpha/risk= .05, confidence=95%

Correct: in 5% of repeated samples, this procedure would falsely reject the null hypothesis

Incorrect: we can by 95% sure *in this sample* that the results are correct

It’s a statement of confidence in the overall method, not the result

Higher significance level= higher risk

Lower significance level= lower risk

Lower significance level does not mean more important results!

How many tests are you doing?

The canonical text: https://xkcd.com/882/

Since there's a possibility of a false positive with every test, repeating tests increases the likelihood of false positives.

Alpha=.05

1 test, 5% chance of false positive

10 tests, 50% chance

20 tests, 100% chance

This is a key question! Analysts will sometimes repeat tests on different parts of the data until they find what they started out looking for-- this is of course bad practice.

Can I See the Confidence Intervals?

Like significance, confidence intervals have a very specific interpretation

Correct: in 95% *of repeated samples*, a confidence interval calculated in this way will include the true value of the population parameter. In our case, the population mean.

Incorrect: we can be 95% certain that the true population interval is within the range we calculated for this sample.

We can never know this! The whole idea is that we only know what will happen in repeated samples.

Is this practically significant?

Maybe the most important question

How big is the difference?

How big is the effect?

What difference would this make in real people’s lives?

Not easy to answer-- analyst MUST know the context