Which did you find most compelling about the “price of life” readings

a) unintended consequences of HIV testing the entire US population
b) unintended consequences of raising plane tickets to improve air traffic safety versus car accident statistics
c) costs per life saved of asbestos removal versus pap smears
d) poverty and lack of education can lead to reduced options/poorer decisions regarding personal health (and correlation to an earlier death)
e) personal risk—“weight, exercise, sex, drugs, smoking, and investments”

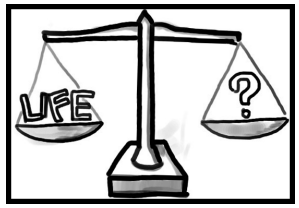


Image Credit: Linda Cai http://cdn1.theodysseyonline.com/files/2015/07/20/635730278800703110204526443_price-of-life-by-linda-cai.png

Deciding Public Policy

The problem with testing the entire US population for HIV is that

a) a positive result becomes relatively meaningless on its own because of all the false positives

b) other

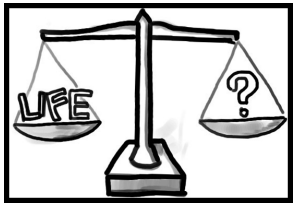


Image Credit: Linda Cai http://cdn1.theodysseyonline.com/files/2015/07/20/

635730278800703110204526443_price-of-life-by-linda-cai.png

Deciding Public Policy

The problem with testing the entire US population for HIV is that

a) a positive result becomes relatively meaningless on its own because of all the false positives

b) other

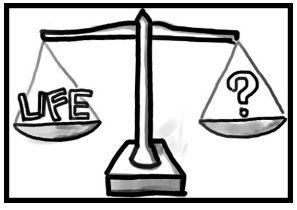


Image Credit: Linda Cai `http://cdn1.theodysseyonline.com/files/2015/07/20/635730278800703110204526443_price-of-life-by-linda-cai.png`

What it does at least somewhat reveal is to change the probability that a person is HIV-positive from roughly 3 in 1000 (general population) to roughly 1 in 4 true positives in the test (but 3 in 4 would be false positives). Testing other populations would require a different analysis.

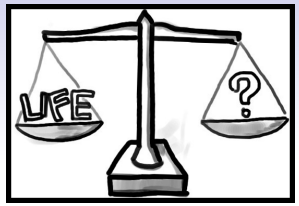


Image Credit: Linda Cai

- What are the pros and cons of HIV testing all of the US? all of Eswatini? of other populations?
  —If a test is 95% accurate for people who have a disease then it correctly tests positive 95% of the time, but incorrectly tests negative for them 100%-95%= 5% of the time (false negative). *Sensitivity* is .95.
  –If a test is 99% accurate for people who don't have a disease then it correctly tests negative 99% of the time, but incorrectly tests positive for them 1% of the time (false positive). *Specificity* is .99.

	Test+	Test-
Person is HIV +	HIV+ people $\times$ probability they test +	
Person is HIV -	# of false positives	
Total		

A retail version of OraQuick costs $30 and gives results in about 20 to 40 minutes. About half of US states test every inmate for HIV on admission or during incarceration. Voluntary testing programs are often ineffective because prisoners do not want to admit to high-risk behaviors. Given this and your statistical analyses from the case studies worksheet, consider whether we should support mandatory HIV testing of newly admitted prison inmates, as you **respond to all of these**:

a) Summarize what you believe is the strongest argument from the “yes” side

b) Summarize what you believe is the strongest argument from the “no” side

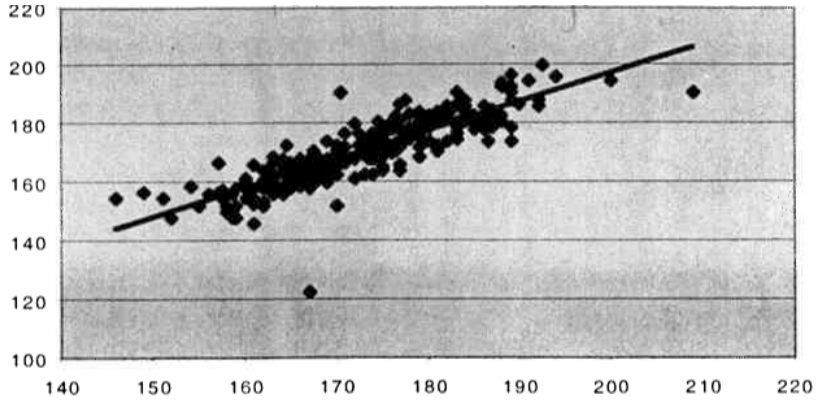
c) What do you think—yes or no?

Connections to Geometry Segment

Russell Knightley, http://www.rkm.com.au/VIRUS/HIV

Does Armspan Predict Height?

y is height

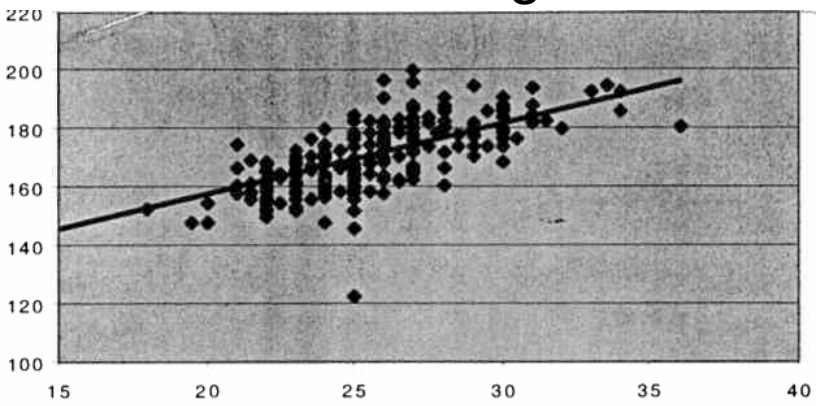


$$y = .9847x + .2357$$
$$R^2 = .7354$$

What else could predict height?

y is height in both cases

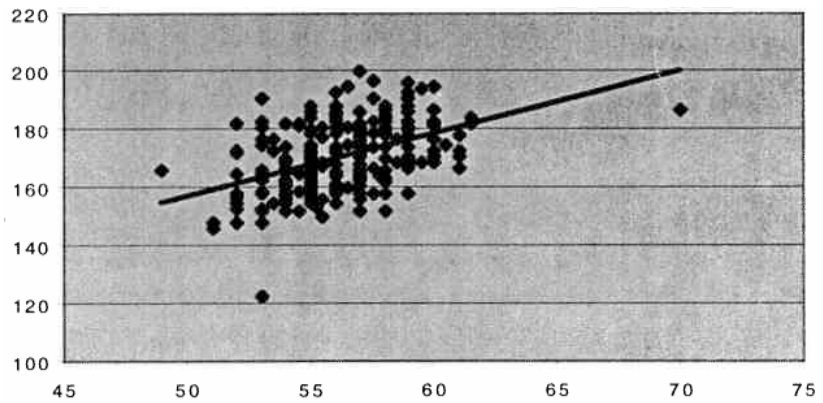
Shoe Length



$y = 2.4042x + 109.23$

$R^2 = .4131$

Head Circumference



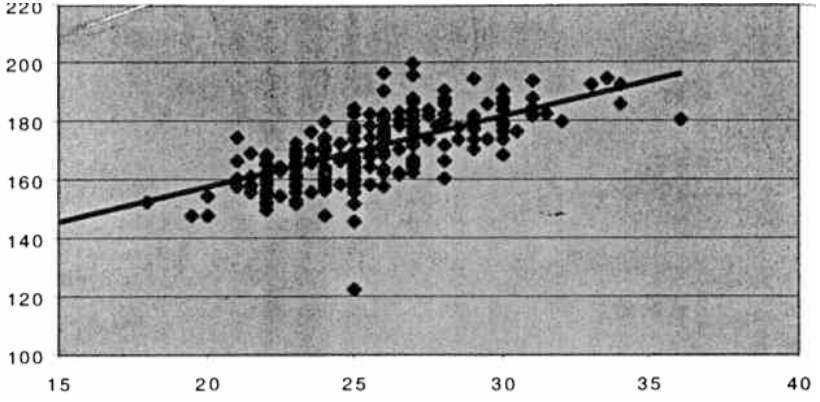
$y = 2.1505x + 49.472$

$R^2 = .2129$

What else could predict height?

y is height in both cases

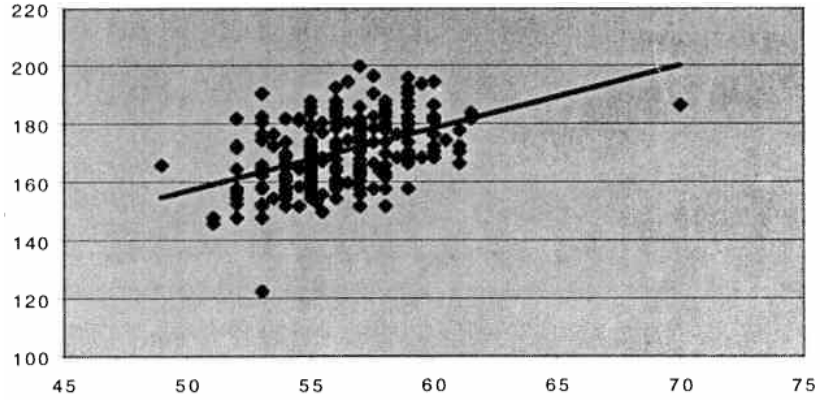
Shoe Length



$y = 2.4042x + 109.23$

$R^2 = .4131$

Head Circumference



$y = 2.1505x + 49.472$

$R^2 = .2129$

Say a break-in occurred. The perpetrator left a shoe print that was approximately 26 cm in length, and a head stocking was found near the scene that was stretched to a diameter of 18 cm (circumference = $\pi \times$ diameter). Use both pieces of information to come up with a prediction of the person's height.

Stereotype Threat

- White men performed worse on a test of mathematical abilities when reminded of Asian-Americans' superior performance in mathematics [Aronson, 1999].
- Asian women performed better on a mathematics test when 'cued' as Asians, but they performed worse when their gender identity was 'cued' [Shih, 1999].

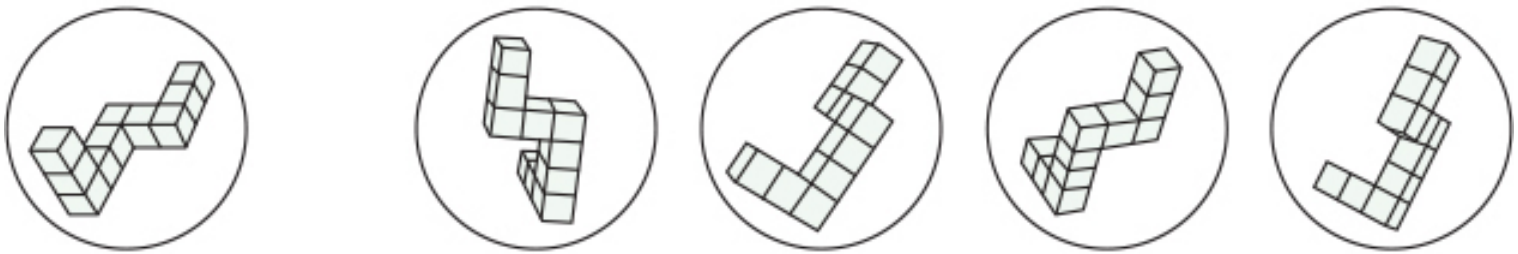
Think about whether you or someone you know have ever experienced something similar to stereotype threat as part of some kind of group (for example, gender, race, sibling, hair color, athlete, southern accent...) where external expectations from someone else (teacher, society, parents, friends...) affected your performance.

Mental Rotations Test (MRT)–Is Time on your Side?

Which two of the pictures on the right are the same as the one on the right when rotated in 3-space? Write down two from a), b), c), and d)...

Mental Rotations Test (MRT)–Is Time on your Side?

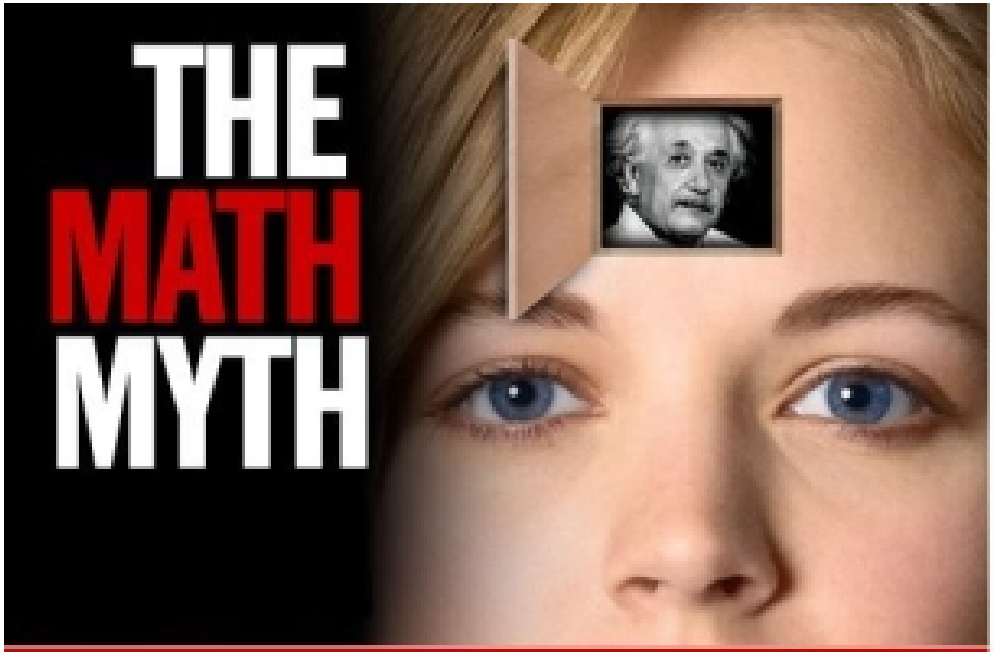
Which two of the pictures on the right are the same as the one on the right when rotated in 3-space? Write down two from a), b), c), and d)...



a) b) c) d)

Mental rotations, A Group Test of Three-Dimensional Spatial Visualization, Vandenberg and Kuse, 1978

Math Gene



(*Time* Magazine, 2005)

- *I was in Japan... Nobody could fathom the idea that if learning higher math didn't come easily, you weren't supposed to continue. You were supposed to work harder... it had nothing to do with some concept of a math gene.* [Lazarus, 2001]

Confounding Issues in Testing–What to Do?

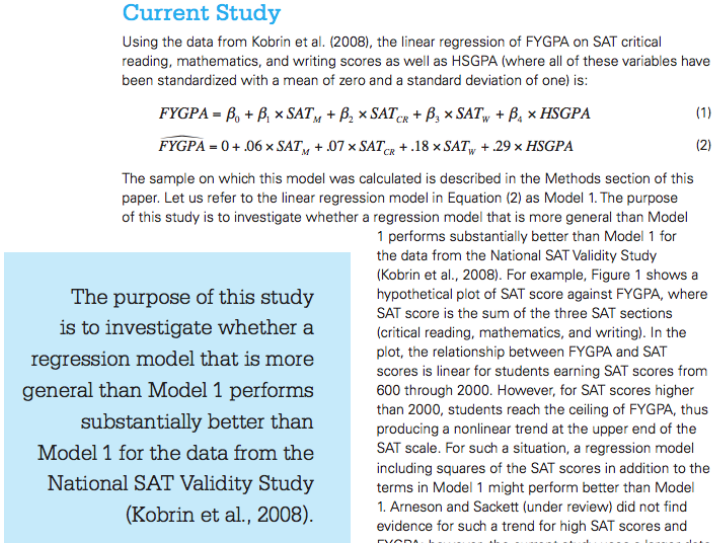
Here's Good News... SAT scores are declining at a slower rate

[from *The Simpsons*]

Confounding Issues in Testing–What to Do?

Here's Good News... SAT scores are declining at a slower rate [from *The Simpsons*]

- Predicting GPA for admission purposes?



Current Study

Using the data from Kobrin et al. (2008), the linear regression of FYGPA on SAT critical reading, mathematics, and writing scores as well as HSGPA (where all of these variables have been standardized with a mean of zero and a standard deviation of one) is:

$$FYGPA = \beta_0 + \beta_1 \times SAT_M + \beta_2 \times SAT_{CR} + \beta_3 \times SAT_W + \beta_4 \times HSGPA \quad (1)$$

$$\widehat{FYGPA} = 0 + .06 \times SAT_M + .07 \times SAT_{CR} + .18 \times SAT_W + .29 \times HSGPA \quad (2)$$

The sample on which this model was calculated is described in the Methods section of this paper. Let us refer to the linear regression model in Equation (2) as Model 1. The purpose of this study is to investigate whether a regression model that is more general than Model 1 performs substantially better than Model 1 for the data from the National SAT Validity Study (Kobrin et al., 2008). For example, Figure 1 shows a hypothetical plot of SAT score against FYGPA, where SAT score is the sum of the three SAT sections (critical reading, mathematics, and writing). In the plot, the relationship between FYGPA and SAT scores is linear for students earning SAT scores from 600 through 2000. However, for SAT scores higher than 2000, students reach the ceiling of FYGPA, thus producing a nonlinear trend at the upper end of the SAT scale. For such a situation, a regression model including squares of the SAT scores in addition to the terms in Model 1 might perform better than Model 1. Arneson and Sackett (under review) did not find evidence for such a trend for high SAT scores and

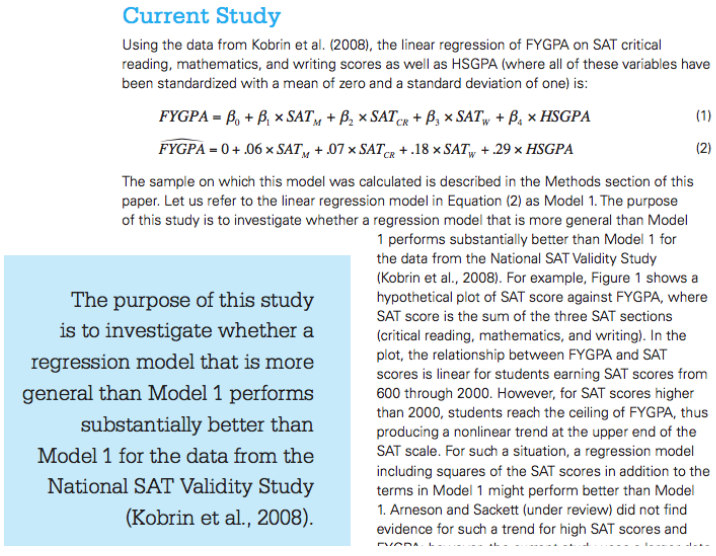
The purpose of this study is to investigate whether a regression model that is more general than Model 1 performs substantially better than Model 1 for the data from the National SAT Validity Study (Kobrin et al., 2008).

An Investigation of the Fit of Linear Regression Models to Data from an SAT Validity Study by Jennifer L. Kobrin, Sandip Sinharay, Shelby J. Haberman, and Michael Chajewski

Confounding Issues in Testing–What to Do?

Here's Good News... SAT scores are declining at a slower rate [from *The Simpsons*]

- Predicting GPA for admission purposes?



Current Study

Using the data from Kobrin et al. (2008), the linear regression of FYGPA on SAT critical reading, mathematics, and writing scores as well as HSGPA (where all of these variables have been standardized with a mean of zero and a standard deviation of one) is:

$$FYGPA = \beta_0 + \beta_1 \times SAT_M + \beta_2 \times SAT_{CR} + \beta_3 \times SAT_W + \beta_4 \times HSGPA \quad (1)$$

$$\widehat{FYGPA} = 0 + .06 \times SAT_M + .07 \times SAT_{CR} + .18 \times SAT_W + .29 \times HSGPA \quad (2)$$

The sample on which this model was calculated is described in the Methods section of this paper. Let us refer to the linear regression model in Equation (2) as Model 1. The purpose of this study is to investigate whether a regression model that is more general than Model 1 performs substantially better than Model 1 for the data from the National SAT Validity Study (Kobrin et al., 2008). For example, Figure 1 shows a hypothetical plot of SAT score against FYGPA, where SAT score is the sum of the three SAT sections (critical reading, mathematics, and writing). In the plot, the relationship between FYGPA and SAT scores is linear for students earning SAT scores from 600 through 2000. However, for SAT scores higher than 2000, students reach the ceiling of FYGPA, thus producing a nonlinear trend at the upper end of the SAT scale. For such a situation, a regression model including squares of the SAT scores in addition to the terms in Model 1 might perform better than Model 1. Arneson and Sackett (under review) did not find evidence for such a trend for high SAT scores and

The purpose of this study is to investigate whether a regression model that is more general than Model 1 performs substantially better than Model 1 for the data from the National SAT Validity Study (Kobrin et al., 2008).

An Investigation of the Fit of Linear Regression Models to Data from an SAT Validity Study by Jennifer L. Kobrin, Sandip Sinharay, Shelby J. Haberman, and Michael Chajewski

- When and if to mark gender, race, ethnicity?
  No fair, you changed the outcome by measuring it! [from *Futurama* about the Heisenberg uncertainty principle]

Gallup

What is your Gallup Poll which you analyzed in the lab (and has a Survey Methods section that includes a margin of sampling error %).

Name at least one item from your Gallup Poll that you found interesting or surprising, or that you had a question on.

Gallup

What is your Gallup Poll which you analyzed in the lab (and has a Survey Methods section that includes a margin of sampling error %).

Name at least one item from your Gallup Poll that you found interesting or surprising, or that you had a question on.

What are some recent examples of confidence intervals and statistically invalid Gallup statements (where they are possibly invalidated by taking into account the margin of error)?

General Electric Illumination Study


General Electric (GE) funded the National Research Council (NRC) of the National Academy of Sciences to study how worker productivity is tied to lighting during a 1924 illumination study. What do you think the NRC concluded?

a) Turn up the lighting slightly (but noticeably) and productivity goes up
b) Dim the lighting slightly (but noticeably) and productivity goes up
c) Productivity stays the same in both cases
d) Productivity goes up in both cases
e) Productivity goes down in both cases

General Electric Illumination Study


General Electric (GE) funded the National Research Council (NRC) of the National Academy of Sciences to study how worker productivity is tied to lighting during a 1924 illumination study. What do you think the NRC concluded?

a) Turn up the lighting slightly (but noticeably) and productivity goes up
b) Dim the lighting slightly (but noticeably) and productivity goes up
c) Productivity stays the same in both cases
d) Productivity goes up in both cases
e) Productivity goes down in both cases

No fair, you changed the outcome by measuring it! [from *Futurama* about the Heisenberg uncertainty principle]