

Problem Set 2

Kaylee, Carrie, Kritphong, Filipe and Klint
EDUC 252L

February 2, 2018

1 Breaking the Classical Test Theory Model

1.1 Coin Flips

Coin flips should not be reliable data - they're random! To look at this a little more analytically:

$$\alpha = \frac{K}{K-1} \left(1 - \frac{\sum_{i=1}^K p_i(1-p_i)}{\sigma_X^2} \right)$$

The interesting thing to note here is that the probability of flipping heads is:

$$p_i = 0.5$$

And the variance on the sum of K coin flips will be:

$$\sigma_X = 0.25K$$

Substituting in the formula for Cronbach's Alpha:

$$\alpha = \frac{K}{K-1} \left(1 - \frac{\sum_{i=1}^K (0.5)(1-0.5)}{0.25K} \right)$$

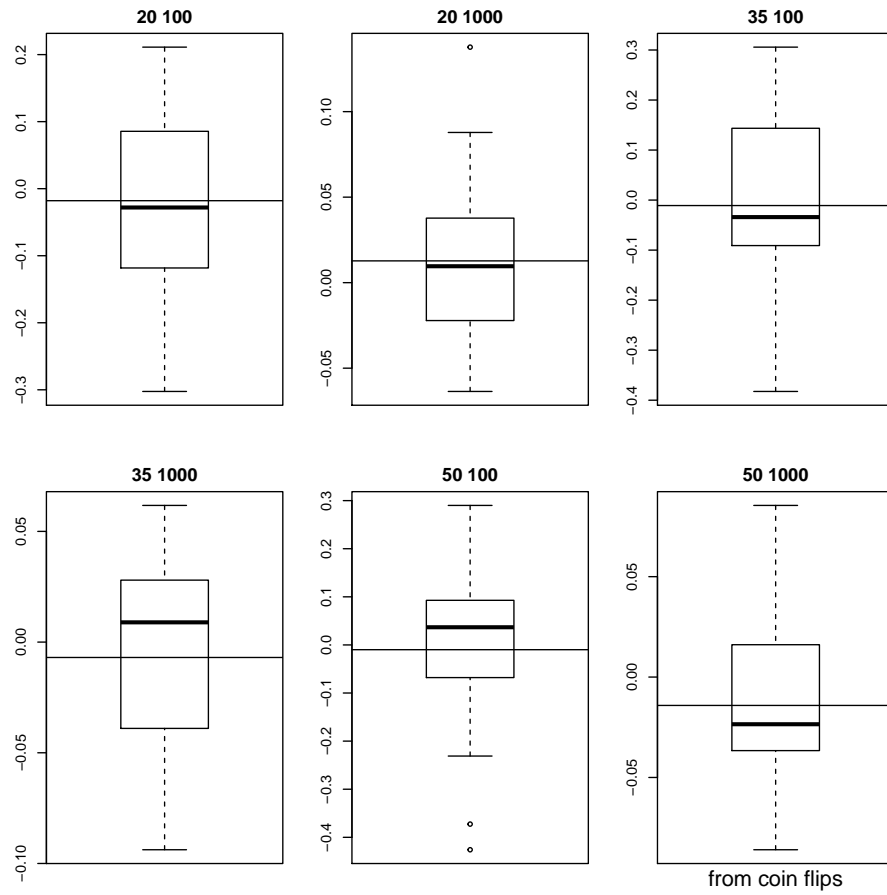
Cleaning up:

$$\alpha = \frac{K}{K-1} \left(1 - \frac{0.25K}{0.25K} \right)$$

$$\alpha = \frac{K}{K-1} (1-1)$$

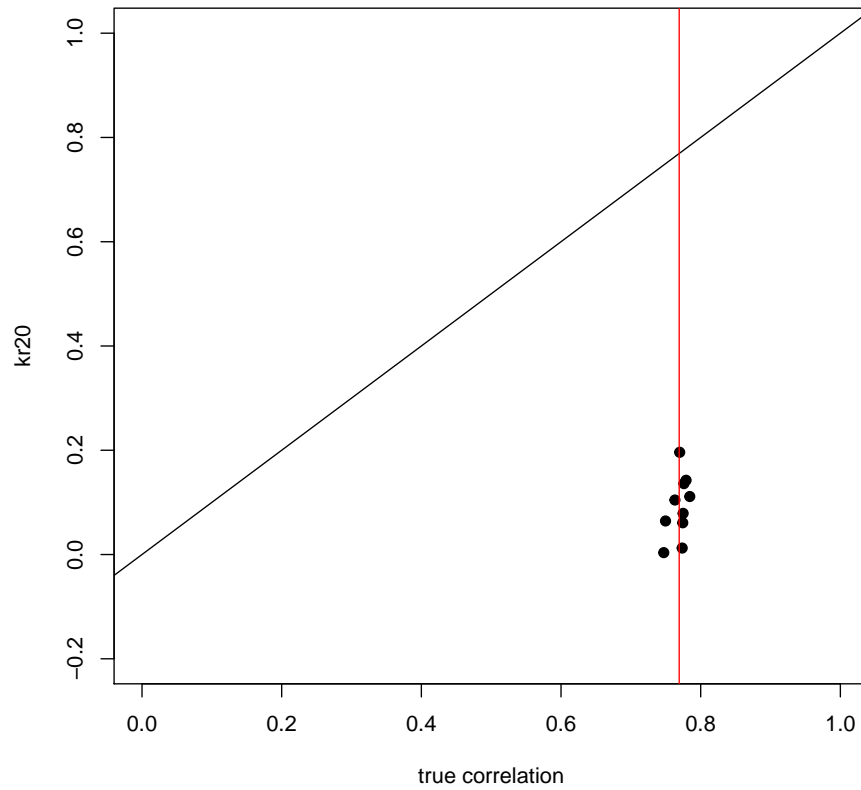
$$\alpha = 0$$

The expectation, then, is that α should be zero for each situation.

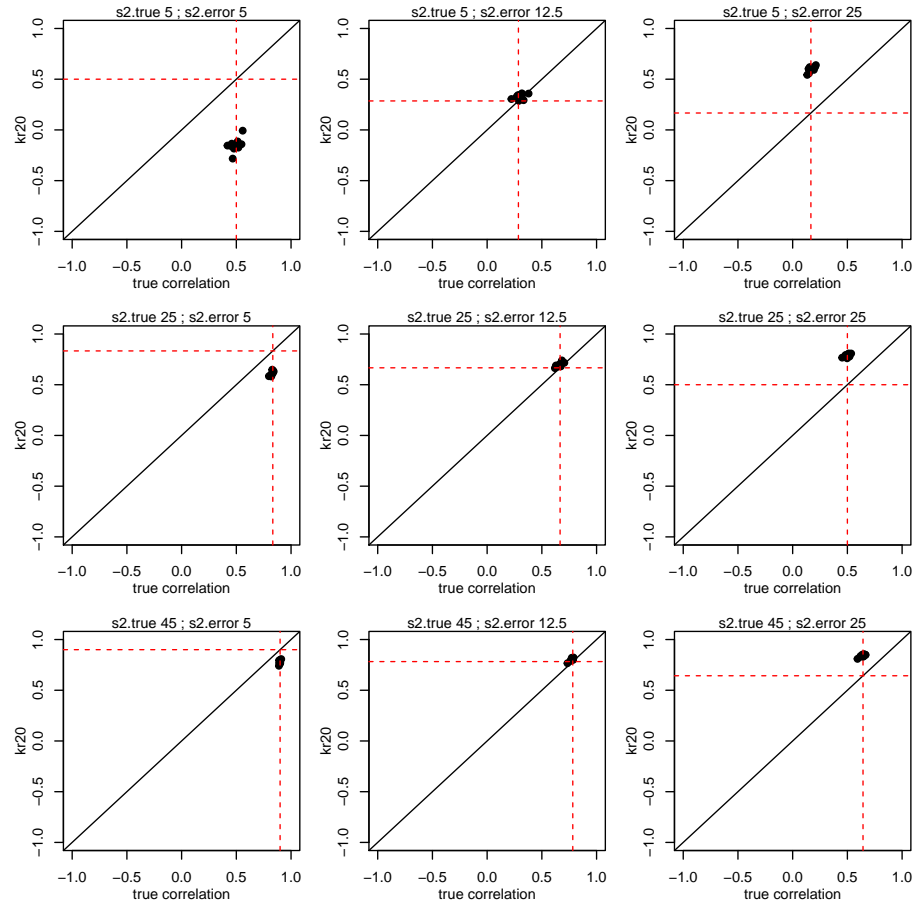


These α plots make sense - they are centered around zero, as predicted, and as the number of items increases, α is more tightly clustered around zero.

1.2 Simulating Item Response Data



The feature of the data generation mechanism that makes the α values super low is that items are getting marked correctly (essentially) at random! Even though the item responses generated correct p-values and test-level correlations, the data generation disregarded any internal structure you would expect. More clearly stated, respondents of similar ability levels did not have similar item response profiles.

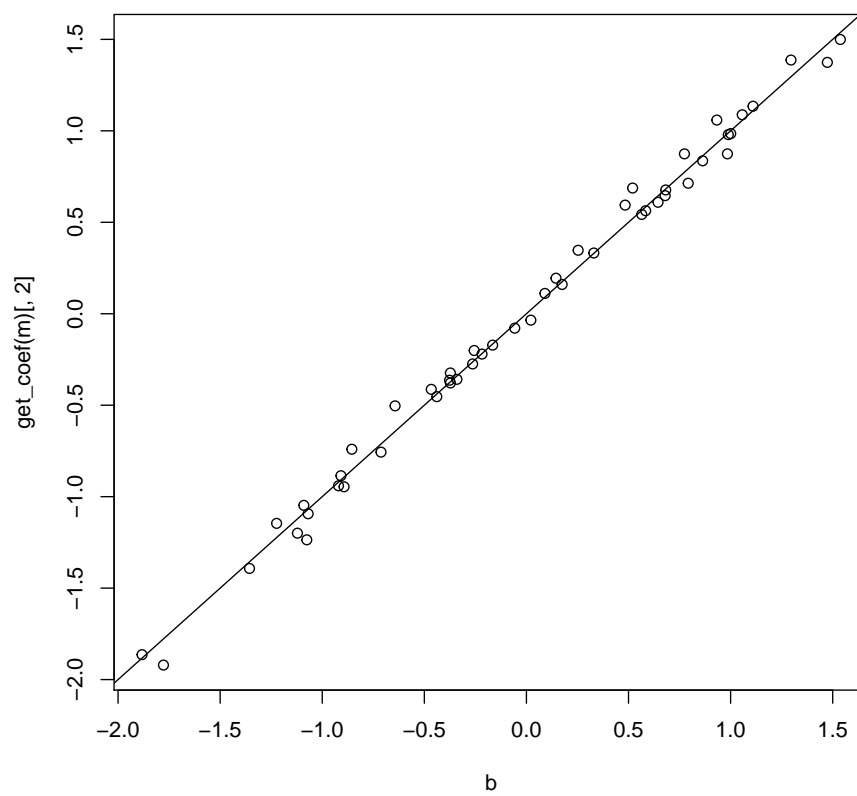


Looking at the resulting plots, it's clear that even with nonsensical item response data, the KR-20 estimate of reliability increases as a function of both true score variance and error variance. This *feels* very wrong. Increasing true score variance can be done by applying an instrument to a population it may not have been originally designed for. Increasing error variance can be done by adding more items or manipulating the quality of items. The challenge with feeling good about the CTT model is that KR-20 is both heavily valued and easily manipulated. The worst part is that some of the behaviors that would increase a KR-20 value could have negative impacts on the validity of the instrument.

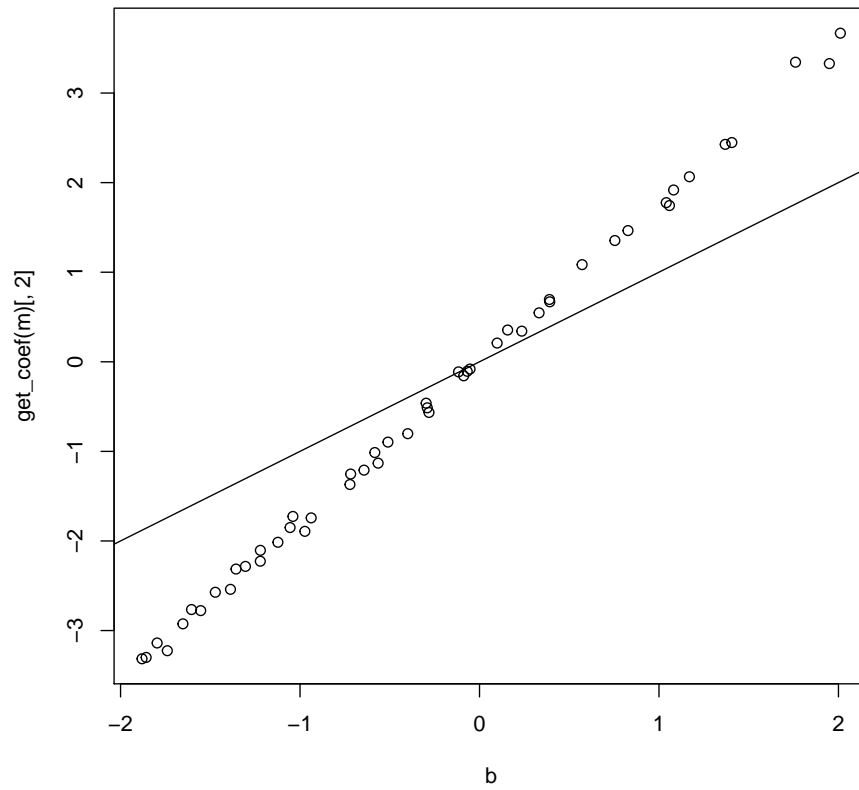
2 Different Link Functions

2.1 The Default

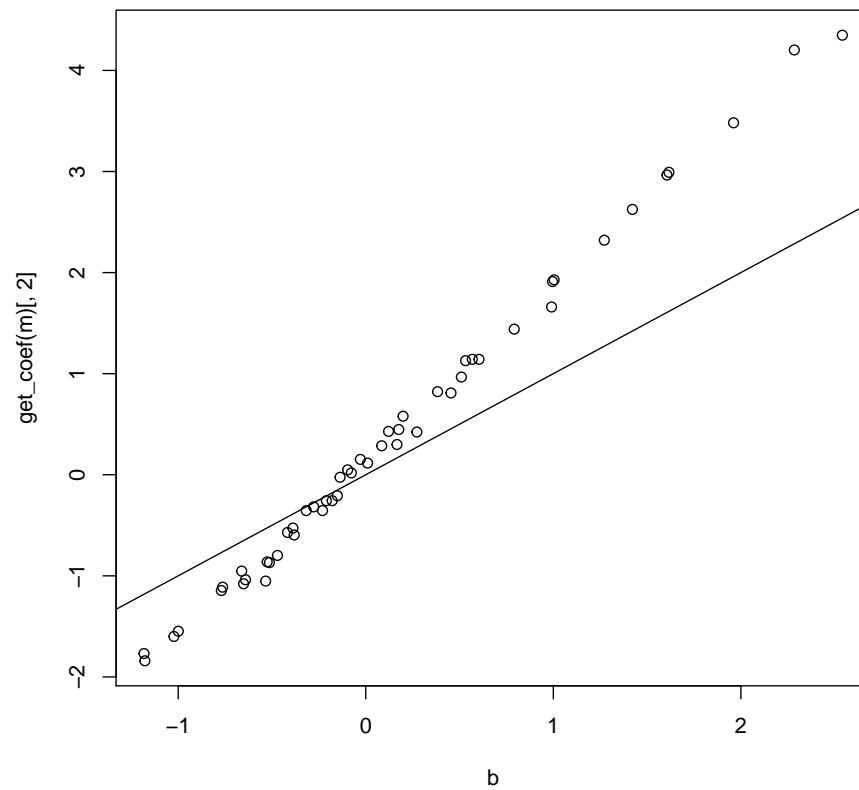
```
## Loading required package: stats4  
## Loading required package: lattice
```



2.2 The Normal



2.3 Heavy Tails



2.4 Skewed

```
##
## Attaching package: 'sn'
## The following object is masked from 'package:stats':
##
## sd
```

