Estimated True Scores and the Confidence Interval

We can never know the true score associated with a psychological attribute. Further, if someone takes a sound psychological test, the best “point” estimate of the psychological attribute measured by the test will be the person’s score on the test. For example, if you take a standard intelligence test and get a score of say 115 on the test, the 115 is the best estimate of your level of intelligence that can be had from that test.

However, as we will see, if you are going to use a person’s score on a test to predict their performance on a subsequent test of the same kind, there is another kind of point estimate, the estimated true score \( X_{te} \) that might be a better estimate than the observed test score for making the prediction.

\[
X_{te} = R_{XX} \left( X - \bar{X} \right) + \bar{X}
\]

Using data from Table 1:

http://www1.appstate.edu/~bacharachvr/table_5.pdf

\[ R_{XX} = .48 \]
\[ \bar{X} = 105 \]
Using Denise’s observed score = 85

\[ X_{te} = .48(85 - 105) + 105 = 95.4 \]

Using Bob’s observed score = 145
(I suggest you do the calculations)

\[ X_{te} = 123.2 \]

Notice that in each case the estimated true score is closer than the observed score to the mean of the distribution of test scores. This effect is called regression to the mean.

If \( R_{XX} = 1 \) then each observed score would equal its corresponding estimated true score. If \( R_{XX} = 0 \) then each observed score would equal the mean of the distribution of observed scores. (Use Denise’s observed score and calculate estimated true scores for her assuming the \( R_{XX} = 1 \) and then assuming that \( R_{XX} = 0 \)).
Confidence Intervals

Because an observed score on any test is an estimate of that person’s true score, many test manuals report score intervals around each possible observed score. Each interval represents a range of score values that is likely to contain the true score. For example, suppose that someone has an observed score of 50 on a self-esteem test, and that the $se_m$ for the test scores is 3. You can calculate an interval or range of scores around the observed score, a confidence interval, that has a specified probability of containing that person’s true score:

$$X_O \pm (z_{68\%})(se_m)$$

Where $X_O$ is the observed score; $z_{68\%}$ is the $z$-score from a normal distribution table corresponding to a $z$-score below which 68 percent of the area of the normal distribution falls… in this case, $z_{68\%} = 1.0$. In our example, the $se_m = 3$:

$$50 \pm (1)(3) = 47.00 - 53.00$$

In this example, we could conclude that there is a 68% chance that the true score associated with the observed score of 50 falls in a score interval that ranges from 47 to 53.

If we wanted to establish a 95 percent confidence interval around our observed score of 50:

$$50 \pm (1.96)(3) = 44.12 - 55.88$$

because 95% of the area of the normal distribution falls below a $z$-score of 1.96.

Now let’s compute a 90% confidence interval around Bob’s observed score from Table 2:

$$X_{Bob} = 130$$

$$se_m = 20.21 \sqrt{1 - .71} = 10.88$$

$$z_{90\%} = 1.28$$

$$130 \pm 13.93 = 116.07 - 143.93$$
We can conclude that there is a 90% chance that Bob’s true score falls somewhere between the scores of 116.07 and 143.93. Try computing the 90% confidence interval around Bob’s observed score from Table 1. The interval will be much larger than the interval we just calculated … why?

Notice that as test score reliability increases, that is as measurement error decreases, confidence intervals shrink in size. If $R_{\chi\chi} = 1$, then the observed score would equal the true score. The smaller the confidence interval for any given $z\%$, the more accurate the observed score.

A confidence interval can also be constructed around the true score in the same way that a confidence interval can be constructed around an observed score.