

Part 5 Lecture 2b Reliability





Standard Error of Measurement

□ ICC → relative measure of reliability □ No units

□ SEM → absolute index of reliability
 □ Same units as the measurement of interest
 □ Usually used to construct confidence intervals

□ The SEM is the standard error in estimating observed scores from true scores.





Calculating the SEM

2 basic ways to calculate SEM
 #1

$$SEM = SD\sqrt{1 - ICC}$$

$$SD = \sqrt{\frac{SS_{TOTAL}}{n-1}}$$





Calculating the SEM

2 basic ways to calculate SEM
 #2

 $SEM = \sqrt{MS_E}$





Standard Error of Measurement

We can report SEM values in addition to the ICC values and the results of the ANOVA

We can also construct 95% confidence intervals about a subject's estimated true score based on the SEM

□We can calculate the minimum difference (MD) that can be considered "real" between scores





Confidence Intervals

- First we must estimate the subjects true score (T):
- □ X = grand mean
- \Box S = observed score

$$T = \overline{X} + ICC(d)$$
$$d = S - \overline{X}$$





Confidence Intervals

□ Second, we must determine the standard error of estimate (SEM_{TS}):

 \Box SD = standard deviation

□ ICC = intraclass correlation coefficient

$$SEM_{TS} = SD\sqrt{1 - ICC} * \sqrt{ICC}$$





Confidence Intervals

□ We can calculate the 95% confidence intervals around the obtained score as:

$S \pm 1.96(SEM)$

We can calculate the 95% confidence intervals around the true score as:

$$T \pm 1.96(SEM_{TS})$$





Minimum Difference

The SEM can be used to determine the minimum difference (MD) to be considered "real" and can be calculated as follows:

$MD = SEM(1.96)(\sqrt{2})$





Examiner Reliability Statistics

Two measures used:

- Percentage Agreement
 - the percentage of judgements where the two examiners have agreed compared to the total number of judgements made

Kappa Statistic





Example – Percentage Agreement

Number of agreements = sum of diagonals

= 61

Total number of cases = overall total

= 100

$$PA = \frac{Number of agreements}{Total number of cases} \times 100$$

Percentage agreement = 61%





Kappa Statistic

The Kappa Statistic measures the agreement between the evaluations of two examiners when both are rating the same objects.

It describes agreement achieved beyond chance, as a proportion of that agreement which is possible beyond chance.





Kappa Statistic

Interpreting Kappa

□ The value of the Kappa Statistic ranges from 0 - 1.00, with larger values indicating better reliability.

□ A value of 1 indicates perfect agreement.

□ A value of 0 indicates that agreement is no better than chance.

□ Generally, a Kappa > 0.60 is considered satisfactory.





Kappa Statistic

The formula for calculating the Kappa Statistic is:

$$\text{Kappa} = \frac{P_O - P_E}{1 - P_E}$$

where:

 P_O = proportion of observed agreements P_F = proportion of agreements expected by chance

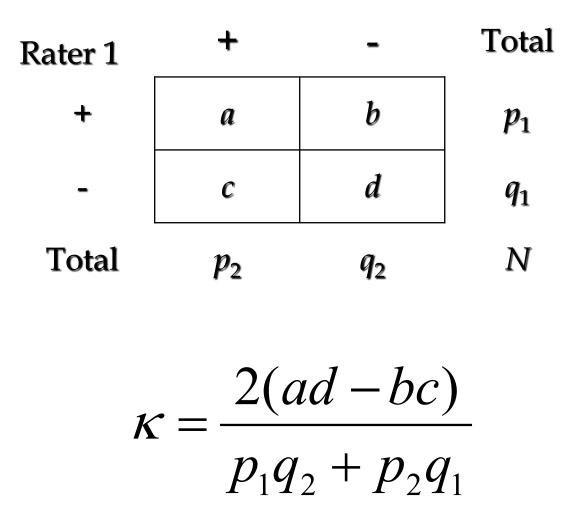
Observed agreement = (A+D) Expected agreement = (((A+B)*(A+C))+((C+D)*(B+D)))/(A+B+C+D) Kappa = ((Observed agreement) - (Expected agreement))/((A+B+C+D) - (Expected agreement))





Kappa Statistic (Formula)

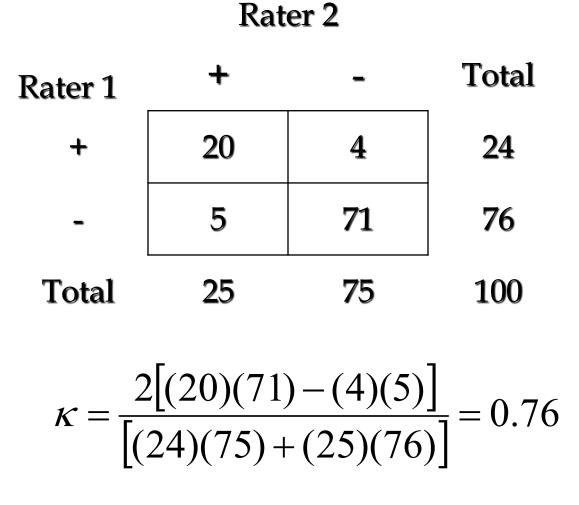
Rater 2







Kappa Statistic (Illustration)







Interpretation of Kappa

Percent agreement above chance
The closer to 1, better agreement

Range of Kappa	Interpretation
> .75	Excellent agreement
.40 to .75	Good agreement
< .40	Poor agreement







Next up in Part 6 Lecture 1: Survival Analysis



