KAAP 602

Statistics

Lecture #2 Notes

1. Review SPSS homework on descriptive statistics
2. Explain Standard Error of the Mean
	1. SEM = s/sqrt(n)
3. Introduce correlation
	1. Measures relationship between 2 variables
	2. Used for paired scores (scores must come from same “subject”)
	3. Values range between -1 to +1
	4. Used to measure:
		1. Reliability
		2. Validity
		3. Relationship between measures
		4. Make predictions
	5. Scattergrams describe relationship. Draw:
		1. No relation
		2. Weak positive
		3. Weak negative
		4. Strong positive
		5. Strong negative
	6. Caution about use of word “cause”
	7. Caution about bivariate distributions
	8. Caution about non-linear distributions
	9. Correlation for ordinal data vs interval/rational data
		1. Ordinal data, most common is Rho, or Spearman’s Rho

 $ρ=1-\frac{6Σd^{2}}{N\left(N^{2}-1\right)}$

* + 1. Interval/rational data: Pearson’s product-moment

Z-score equation: $r=\frac{\left(Z\_{x}\*Z\_{y}\right)}{\left(n-1\right)}$

Raw score equation: $r=\frac{Σ\left(X-\overbar{X}\right)\left(Y-\overbar{Y}\right)}{\sqrt{Σ\left(X-\overbar{X}\right)^{2}\left(Y-\overbar{Y}\right)^{2}}}$

* 1. Evaluating a correlation coefficient
		1. Statistical significance (significant if p < 0.05)
		2. Issues with test of statistical significance
			1. As N increases, the magnitude of the correlation coefficient necessary to achieve statistical significance decreases.
		3. General guidelines for evaluating the strength of a relationship\*
			1. 0.0-0.1: No Relation
			2. 0.1-0.39: Weak
			3. 0.4-0.69: Moderate
			4. 0.7-1.0: Strong

\*Dancey CP, Reidy J. *Statistics without Maths for Psychology*. Fifth. Prentice Hall, London; 2011.

1. Introduce linear regression
	1. General concept: correlation provides knowledge of one variable from another variable.
	2. Correlation coefficient is used to “temper” the estimate toward the mean
	3. Standardized simple regression equation: Z’y = r \* Zx
	4. Expand to equation for raw data

$$Y^{'}= r\left(\frac{S\_{y}}{S\_{x}}\right)\left(X-\overbar{X}\right)$$

* 1. Explain concept of residuals
	2. Explain standard error of estimate
		1. $S\_{est}= \sqrt{\frac{Σ(y-y^{'})^{2}}{N-1}}$
	3. Use residuals to determine if relationship is linear (Plot pred vs resid).
	4. Running in SPSS
		1. Correlation
		2. Scatterplots
		3. Simple linear regression
		4. Transforming variables
		5. Polynomial regression