

## DESCRIBING RELATIONSHIP

- **SHAPE:** Linear (straight line) Non – Linear (curved line)
- **STRENGTH:** Pearson's r value
- Strong = 0.75 or greater
- Moderate = 0.45 – 0.75
- Weak = 0.25 – 0.45
- Extremely Weak = Less than 0.25
- **DIRECTION:** (based on whether Pearson's r is - / +)
- Positive relationship = 1 (perfect)
- Negative relationship = -1 (perfect)
- No relationship = 0

## SAMPLE SIZE

- Large sample size = Pearson's r is accurate
- Small sample size = Pearson's r unreliable & need to use P value

## LINEAR REGRESSION

- The variation in the DV due to differences in the IV
- Used primarily for prediction

## BIVARIATE REGRESSION: (1 DV)

### REGRESSION EQUATION (used primarily to predict)

- Determines the form of the relationship (slope / gradient of the line)
- Predicted  $Y = a + bX$  (for sample) use beta instead of b for population
- $Y = DV$
- a = intercept (where line crosses the y axis)
- b (slope / gradient value) (can be + or -)
- $X = IV$

**LINE OF BEST FIT:** The slope & vertical intercept which will minimise the prediction errors

### COEFFICIENT OF DETERMINATION ( $r^2$ ): (variation in DV due to linear relationship)

- Gives a better idea of how important the relationship is
- To determine how much of the variance is attributed to the linear relationship
- (Predicted Value – Mean Value)

### RESIDUAL (prediction error): (Variation in DV due to other factors)

- Difference between the DV & our line of best fit
- To determine how much of the variance is attributed to "other" factors
- (Actual (observed) Value – Variance due to Linear Relationship)

## INTERPRET THE SLOPE

- If positive = On average, for each additional unit increase in IV, our DV tends to be \_ units higher
- If negative = On average, for each additional unit increase in IV, our DV tends to be \_ units lower

## REGRESSION OUTPUT (SPSS)

- Predicted DV (name) = a (value for constant) b (+/- value for IV) x IV (name) (can substitute if given a value)
- Eg. Predicted life Optimism = 29.11 –0.50 x Neuroticism
- On average, for each additional unit increase in Neuroticism, Life Optimism would tend to decrease by 0.50 units

## INTERPRET VERTICAL INTERCEPT

- Vertical intercept is when  $x = 0$
- Eg. When children do not watch any television, their predicted aggression level would be.....

## RELATIONSHIP BETWEEN R (multiple R) & r

- **SPSS Anova output table**
- Regression = SS Regress: variation in DV due to IV (linear relationship)
- Residual = SS Resid: variation in DV due to other factors
- Total = SSy: total variation
- $R^2 = \frac{\text{SS Regress}}{\text{SS y}}$
- Value of multiple R given in regression SPSS output is always positive & has the same magnitude as r
- r value is (+ / -) based on Negative / Positive slope

## ASSUMPTIONS OF REGRESSION

**METRIC SCALE:** Assumed by ensuring both variables are measured on a metric scale

**INDEPENDENCE:** All of the observations are independent, assumed by random sampling

**LINEARITY:** The relationship between DV & IV must be linear, tested by producing Scatterplots to check for non-linearity

- **DV vs IV** – regular scatterplot (straight not curved)
- **Standardised Residual vs IV** – plot the residual on the vertical axis (rather than DV)
- Observed value above line = residual is positive