Hypothesis Testing - Relationships

Session 03

Lecture Outline

- Correlational Research
  - Scattergrams.
  - The Correlation Coefficient.
  - An example.
  - Considerations.
- Hypothesis Testing
  - One and Two-tailed Tests.
  - Errors.
  - Power.
- Hypothesis Testing for Relationships

Correlational Research

Correlational research is concerned with the relationships between variables.

Whether high scores on one variable go with high (or low) scores on another, or whether there is no observable pattern.

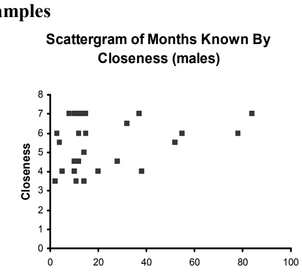
Scattergrams

Scattergram: method for gaining an impression of the nature of a relationship between two variables.

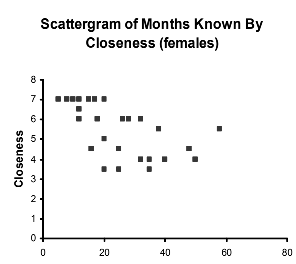
Construction:
Independent variable goes on the horizontal axis (X) and **dependent variable** on the vertical axis (Y).
Determine the range of raw scores and mark them on the axes from the lowest to highest (from the origin).
Plot each cases score on the Y axis with their corresponding score on the X axis.

Scattergrams

Some examples

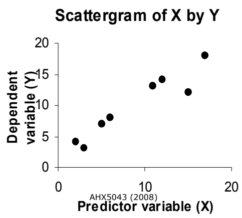


Scattergrams



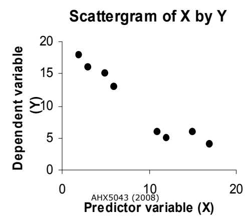
Scattergrams

Positive linear correlation *(high scores go with high, low scores go with low).*



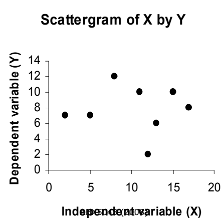
Scattergrams

Negative linear correlation *(high scores go with low, low scores go with high).*



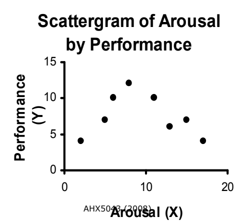
Scattergrams

Weak or no correlation *(no clear visible relationship).*



Scattergrams

Curvilinear correlation (*eg. the Inverted -U curve*)



Scattergrams

Scattergrams do no provide a precise statistic of the degree of correlation between the variables.

Often you will construct a scattergram and see no correlation when in fact there may be a weak correlation.

The Correlation Coefficient

Pearson's *r* : a measure of the degree of correlation between two linear variables.

Calculation of Pearson's *r* using Z scores:

Change all scores to Z scores.

Multiply out each pair of Z scores (for each individual) resulting in cross products.

Sum the cross products.

Divide by number of cases.

The Correlation Coefficient

Z score formula for Pearson's *r* :

$$r = \frac{\sum Z_X Z_Y}{N}$$

The sum of cross products divided by the number of pairs of scores.

The Correlation Coefficient

Underlying Logic:

Positive correlation - positive values on one variable correspond with positive values on the other, and negative values on one variable correspond with negative values on the other resulting in a positive sum of cross products.

Negative correlation - positive values on one variable correspond with negative values on the other, and visa versa resulting in a negative sum of cross products.

Weak or no correlation - mixture of positive and negative values canceling each other out resulting in a near zero sum of cross products.

The Correlation Coefficient

The logic of Pearson's *r*:

Using Z scores converts the two variables onto the same scale.

Dividing by the number of cases gives the average of the cross products.

Results can range from -1 to +1.

-1 = perfect negative linear correlation.

0 = no correlation.

+1 = perfect positive linear correlation.

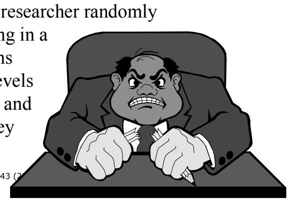
The Correlation Coefficient

Steps in calculating Pearson's *r*:

1. Construct a scatter diagram.
2. Determine if curvilinear - if so **do not** use Pearson's *r*.
3. Estimate degree and direction of correlation.
4. Compute correlation coefficient.
5. Check with estimation.

An example . . .

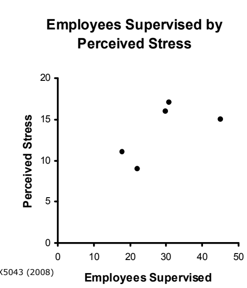
A researcher predicts a strong positive linear relationship between managers stress levels and the number of employees they supervise. In other words, the more employees managers supervise, the greater the managers perceived stress levels. The researcher randomly selects five managers working in a major corporation and obtains information on their stress levels (via a pencil and paper test), and the number of employees they supervise.



An example . . .

1. Construct a scattergram

Employees Supervised	Perceived Stress Level
45	15
30	16
18	11
22	9
31	17



An example . . .

2. Determine if curvilinear

- no looks roughly linear.

3. Estimate degree and direction of coefficient

- strong positive linear correlation (Pearson's *r* of approximately .7).

An example . . .

4. Compute correlation coefficient

X	X-M	$(X-M)^2$	Z_X	Y	Y-M	$(Y-M)^2$	Z_Y	Z_XZ_Y
45	15.8	249.6	1.70	15	1.4	1.96	0.46	0.78
30	0.8	0.64	0.09	16	2.4	5.76	0.78	0.07
18	-11.2	125.4	-1.21	11	-2.6	6.76	-0.85	1.02
22	-7.2	51.84	-0.78	9	-4.6	21.16	-1.50	1.16
31	1.8	3.24	0.19	17	3.4	11.56	1.11	0.21

$\Sigma Z_X Z_Y = 3.24$

M = 29.2
SS = 430.8
$SD^2 = 86.16$
SD = 9.28

M = 13.6
SS = 47.2
$SD^2 = 9.44$
SD = 3.07

r = .65

An example . . .

5. Check with estimation

Actual Pearson's *r* value of .65 is close to approximated *r* value of .7

These results suggests that there is a strong positive linear relationship between managers perceived stress and number of employees supervised as predicted.

Considerations

You cannot infer causality from a correlation.

If the range of possible values is restricted, you cannot apply the findings to the entire range the variable might have.

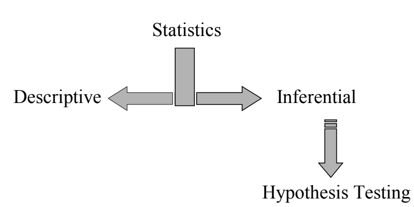
Size Conventions:

Large	approximately > .5
Moderate	approximately >.3 to .49
Low	approximately .1 to .3

Considerations

- The correlation coefficient is useless for non-linear relationships.
- It is fairly unreliable for smaller numbers of pairs of observations.
- Outliers have a major effect on Pearson's *r*.
- It does not enable the determination of cause and effect.

Hypothesis Testing



Hypothesis Testing

- The branch of statistics that helps you determine whether or not the prediction you made about something occurred by chance, or may actually represent a generalisable observation.
- A hypothesis test allows us to draw conclusions or make decisions regarding population data from sample data.

AHX5043 (2008) 25

Hypothesis Testing

A **hypothesis** is a claim or statement about a property of a population

A **hypothesis test** (or test of significance) is a standard procedure for testing a claim or statement about a property of a population.

- Hypothesis testing does not result in definitive conclusions. We are dealing in probabilities. We either conclude that the results we get are likely (or unlikely) to be due to chance.

AHX5043 (2008) 26

Hypothesis Testing

The **null hypothesis** (denoted H_0) is a statement that the value of a population parameter (such as proportion or mean) is equal to some claimed value.

The **alternative hypothesis** (denoted H_1) is a statement that the value of a population parameter somehow differs from the null hypothesis. The symbolic form must be a >, < or ≠ statement.

AHX5043 (2008) 27

Hypothesis Testing

Hypothesis testing is based on the preliminary assumption that the null hypothesis is true.

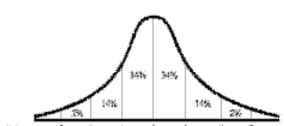
The null hypothesis generally represents a distribution based on no relationship or no difference between groups.

- *Similar to the notion of innocent until proven guilty.*

AHX5043 (2008) 28

Hypothesis Testing

Normal curve: distribution of scores that is characterised by a bell shaped curve in which the probability of a score drops off rapidly from the midpoint to the tails of the distribution.



AHX5043 (2008) 29

Hypothesis Testing

Normal curve tables enable us to determine precise areas under the normal curve.

AHX5043 (2008) 30

Hypothesis Testing

Probability: the expected relative frequency of a particular outcome, or how certain we are that a particular thing will happen.

Probability (p) = the number of successful outcomes divided by the number of possible outcomes.

AHX5043 (2008) 31

Hypothesis Testing

Population: the entire set of things of interest.

Sample: the subset of the population about which you actually have information.

AHX5043 (2008) 32

Hypothesis Testing

In conducting inferential research we collect information from a **sample** to make **probabilistic** inferences about the parameters of a **normally distributed population**.

AHX5043 (2008) 33

Hypothesis Testing

A **test statistic** is a value computed from the sample data, used in making the decision whether or not to reject the null hypothesis.

- Examples are t, z, F, r etc.
- The test statistic indicates how far our sample deviates from the assumed population parameter.

AHX5043 (2008) 34

Hypothesis Testing

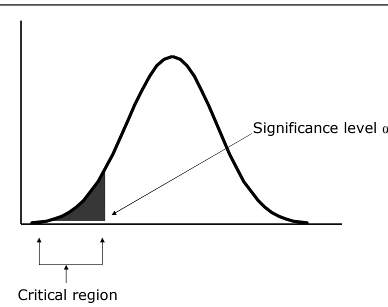
Critical region (or rejection region) is an area extremely different enough from the null hypothesis that would allow us to reject it.

Significance level (α) is the probability that the test statistic will fall in the critical region when the null hypothesis is actually true. Common values are 0.01, 0.05 (i.e. 1% & 5%).

A **critical value** is any value that separates the critical region from values of the test statistic that would not cause us to reject the null hypothesis.

AHX5043 (2008) 35

Hypothesis Testing



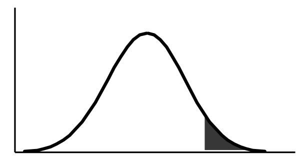
AHX5043 (2008) 36

One and Two-tailed Tests

One-tailed test: a situation where a researcher predicts a relationship or difference in a specific direction.

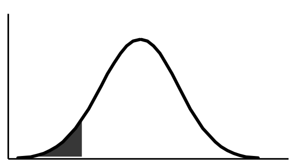
One and Two-tailed Tests

- E.g. A researcher predicts that there will be a significant positive relationship between height and weight.



One and Two-tailed Tests

- E.g. A researcher predicts there will be a significant negative relationship between body fat and fitness.

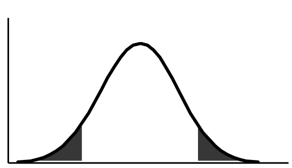


One and Two-tailed Tests

Two-tailed test: a situation where a researcher predicts a relationship or difference but does not specify in what direction.

One and Two-tailed Tests

- E.g. A researcher predicts that a particular drug will have a significant effect on body weight, but is unsure in which direction.



Errors in Hypothesis Testing

Type 1 error: The mistake of rejecting the null hypothesis when it is actually true. α is the probability of a type 1 error.

- Researcher concludes that treatment had an effect when it really did not.
- May be a serious problem. E.g. Stating that a drug is effective in treating cancer and it is really ineffective (can cost time, hope, & $).

Errors in Hypothesis Testing

Type 2 error: The mistake of failing to reject the null hypothesis when it actually false. The symbol β is used to represent the probability of a type 2 error

- Treatment or difference really exists, but hypothesis test fails to show this.
- E.g. Test for cancer (telling someone they do not have it when in fact they do).

Errors in Hypothesis Testing

Decision	Outcome if Ho is true	Outcome if Ho is false
Do not reject Ho	*Correct decision*	*Type II error*
Reject Ho	*Type I error*	*Correct decision*

Power

Power

- In relation to Type II error, power is define as 1 - beta.
- In other words, power is the probability of detecting a true significant difference.

Hypothesis Testing for Relationships

Assumptions for hypothesis testing using Pearson's *r*:

1. Both variables should be normally distributed.
2. Each variable should have approximately equal variance.

- Minor violations are OK.

Hypothesis Testing for Relationships

Five Stages in Hypothesis Testing:

1. Reframe the question into a research hypothesis and a null hypothesis about the populations.
2. Determine the characteristics of the comparison distribution.
3. Determine the cutoff sample score on the comparison distribution at which the null hypothesis should be rejected.
4. Determine the score of your sample on the comparison distribution.
5. Compare the scores obtained in steps 3 and 4 to decide whether or not to reject the null hypothesis.

Hypothesis Testing for Relationships

1. Reframe the question into a research hypothesis and a null hypothesis about the populations.

- Population 1: Managers like those in this study.
- Population 2: Managers for whom there is no correlation between numbers of employees supervised and stress.

- Ho: The two populations have the same correlation.
- H1: Population 1 has a significantly higher correlation than Population 2.

Hypothesis Testing for Relationships

2. Determine the characteristics of the comparison distribution.

- The comparison distribution is an *r* distribution with a mean of 0.

AHX5043 (2008) 49

Hypothesis Testing for Relationships

3. Determine the cutoff sample score on the comparison distribution at which the null hypothesis should be rejected.

Read the appropriate critical *r* value from the table. You need to know:

- The number of pairs of scores (n = 5).
- The level of significance (α = .05).
- Type of hypothesis (directional).

- In this instance the critical *r* score is .805

AHX5043 (2008) 50

Hypothesis Testing for Relationships

4. Determine the score of your sample on the comparison distribution.

- Calculate *r*.
- In the example *r* was calculated to be .65

AHX5043 (2008) 51

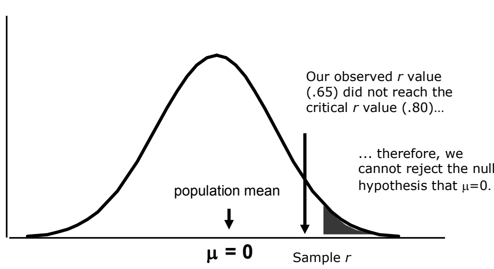
Hypothesis Testing for Relationships

5. Compare the scores obtained in steps 3 and 4 to decide whether or not to reject the null hypothesis.

- The *r* value calculated (r = .65) was not more extreme than the critical *r* value found in the table (r = .805) therefore we cannot reject the null hypothesis. In terms of statistical significance, the results of this study remain inconclusive.

AHX5043 (2008) 52

Hypothesis Testing for Relationships



AHX5043 (2008) 53

Hypothesis Testing for Relationships

Correlation Matrix: A table in which each variable is listed across the top and down the left side, and the calculated r values between each pair of variables are shown inside the table.

Table 1

Correlations between the CTAI-2 (cognitive), CTAI-2 (somatic), CSAI-2 (cognitive) CSAI-2 (somatic) scores.

Measure	1		2		3		4	
	r	n	r	n	r	n	r	n
1. CTAI-2 (cognitive)	—							
2. CTAI-2 (somatic)	.40***	93	—					
3. SAS (worry)	.60***	93	.30**	93	—			
4. SAS (concentration d.)	.36***	93	.07	93	-.07	94	—	
5. SAS (somatic)	.28**	93	.58***	93	.22*	94	.29**	94

* p < .05. ** p < .005. *** p < .0005.

AHX5043 (2008) 54

Research Examples

Cromwell, R. L., & Newton, R. A. (2004). Relationship Between Balance and Gait Stability in Healthy Older Adults. Journal of Aging and Physical Activity, 2004, 11, 90-100.

Research Examples

- Gait patterns of young adults are characterised by phases of instability that allow for efficient forward progression and lateral shifting of the body's center of mass with each step
- With aging, adaptations in older adults' walking pattern increase stability and decrease the capacity for moving the body forward.
- Previous research has revealed that age-related changes in gait create a more stable walking pattern, and measures of balance are related to walking performance.

Research Examples

- The purpose of this study was to determine the relationship between walking stability and measures of balance in healthy older adults.

Research Examples

Method

- 17 older and 20 young adults performed the Berg Balance Test (BBT) and walked 10 m.
- Walking velocity (WV) and cadence were measured, and a gait-stability ratio (GSR) was calculated. Higher GSR indicated that a greater portion of the gait cycle was spent in double-limb support.

Research Examples

Table 3 Correlation Matrix for Walking and Balance Measures

	Cadence	Walking velocity	GSR
Berg Balance Test total score	−.01	.20	−.36
Item 12	.37	.58*	−.74**
Item 13	.04	.24	−.39
Item 14	.04	.16	−.27

*Significantly different from the correlation for walking velocity and Berg Balance Test Item 12 ($p < .05$).
*Significant correlation ($p < .05$).

Research Examples

Conclusions

- A strong inverse relationship was observed between walking stability and dynamic balance in healthy older adults.
- Significant relationships were determined for walking velocity and GSR with BBT Item 12 (alternate stepping on a stool) for older adult participants. Results suggest that GSR provides a better indication than walking velocity for assessing balance deficits during walking in older adults.