Session 6:

Correlation

Correlation Analysis and Covariance

Aims

Measuring Relationships

- **Scatterplots**
- **Covariance**
- **Pearson's Correlation Coefficient**

Nonparametric measures

- **Spearman's Rho**
- **Kendall's Tau**

What is a Correlation?

- It is a way of measuring the extent to which two variables are related.
- It measures the pattern of responses across variables.

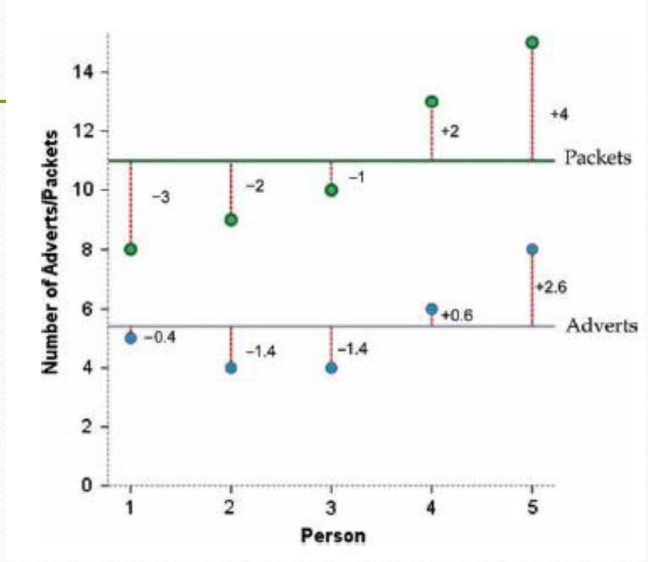
Measuring Relationships

- We need to see whether as one variable increases, the other increases, decreases or stays the same.

- This can be done by calculating the Covariance.

Covariance

- Calculate the error between the mean and each subject's score for the first variable (x).
- Calculate the error between the mean and their score for the second variable (y).
- Multiply these error values.
- Add these values and you get the cross product deviations
- The covariance is the average cross-product deviations:



$$Cov(x,y) = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{N-1}$$

Problems with Covariance

It depends upon the units of measurement.

- E.g. The Covariance of two variables measured in Miles might be 4.25, but if the same scores are converted to Km, the Covariance is 11.

One solution: **standardize it!**

- Divide by the standard deviations of both variables.

The standardized version of Covariance is known as the Correlation coefficient.

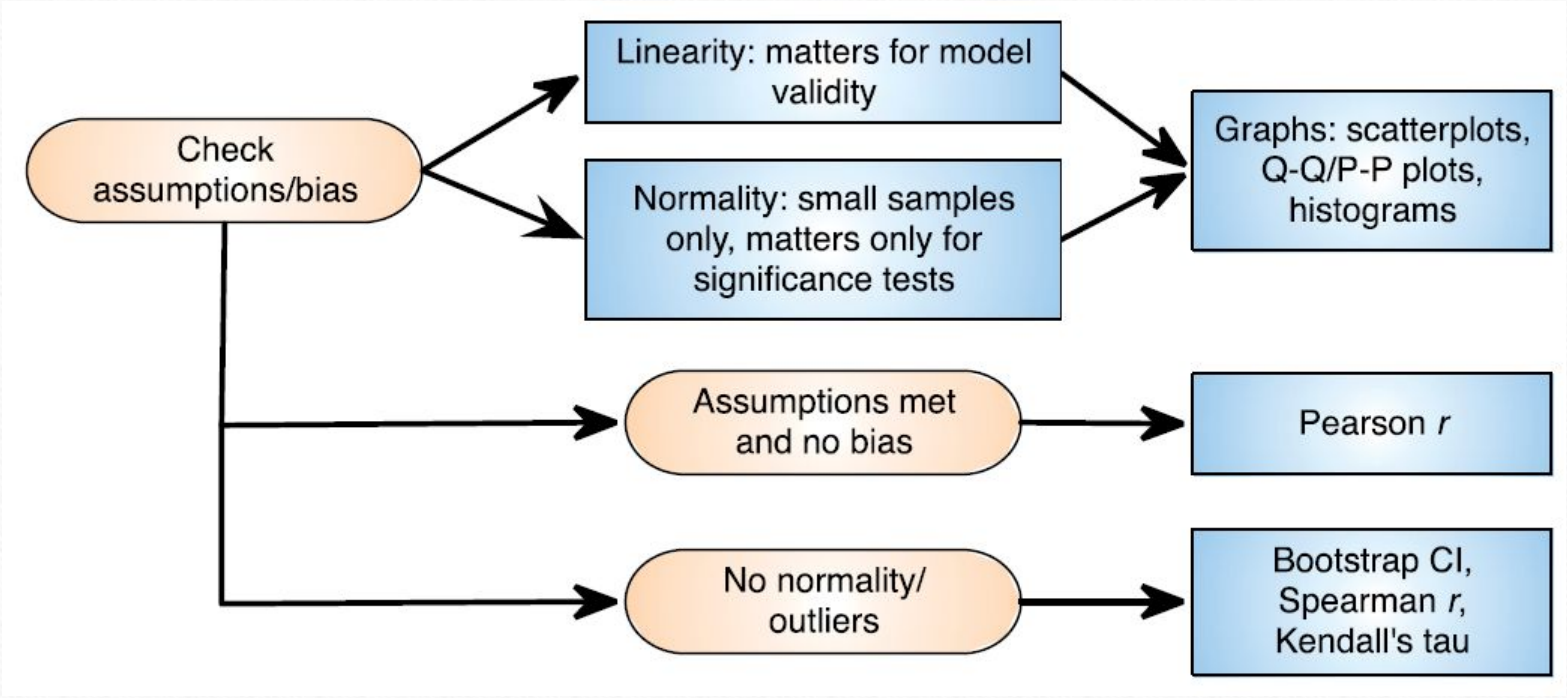
- It is relatively affected by units of measurement.

The Correlation Coefficient (Pearson)

$$r = \frac{Cov_{xy}}{s_x s_y}$$

$$= \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{(N-1)s_x s_y}$$

Conducting Correlation Analysis



Things to know about the Correlation

- It varies between -1 and +1
- 0 = no relationship

Coefficient of determination, r^2

- By squaring the value of r you get the proportion of variance in one variable shared by the other.

Interpretation of Correlation
(may vary by discipline)

Correlations

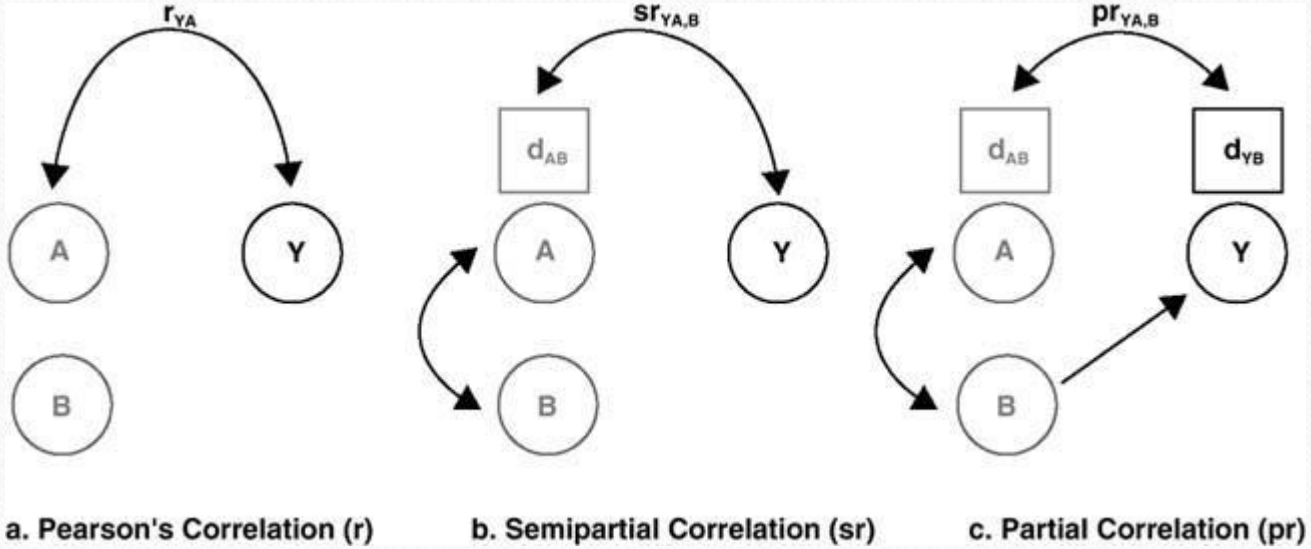
- From **0 to 0.25** (-0.25) = little (**weak**) or no relationship;
- From **0.25 to 0.50** (-0.25 to 0.50) = fair (**moderate**) degree of relationship;
- From **0.50 to 0.75** (-0.50 to -0.75) = moderate to good (**strong**) relationship;
- Greater than **0.75** (or -0.75) = very good to excellent (**very strong**) relationship.

Correlation and Causality

The third-variable problem: in any correlation, causality between two variables cannot be assumed because there may be other measured or unmeasured variables (i.e., covariates or control variables) affecting the results.

Direction of causality: Correlation coefficients say nothing about which variable causes the other to change

Partial vs Semi-Partial Correlations



Nonparametric Correlation

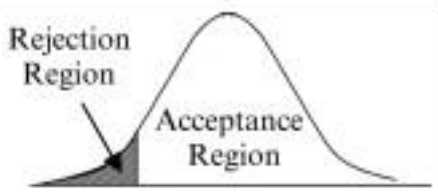
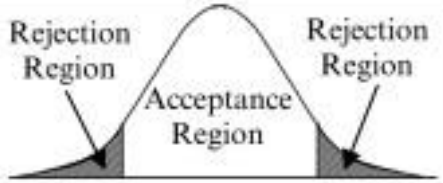
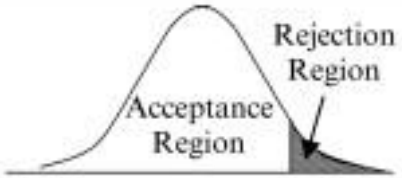
Spearman's Rho, r_s $(or\ \rho)$

Pearson's correlation on the ranked data

Kendall's Tau, τ

- Better than Spearman's for small samples

One-Tailed vs Two-Tailed Tests

One-Tailed Test (Left Tail)	Two-Tailed Test	One-Tailed Test (Right Tail)
$H_0 : \mu_X = \mu_0$ $H_1 : \mu_X < \mu_0$	$H_0 : \mu_X = \mu_0$ $H_1 : \mu_X \neq \mu_0$	$H_0 : \mu_X = \mu_0$ $H_1 : \mu_X > \mu_0$
		

2-Tailed Testing

Correlations

		Essay Mark (%)	Hours Spent on Essay
Essay Mark (%)	Pearson Correlation	1	.267
	Sig. (2-tailed)		.077
	N	45	45
Hours Spent on Essay	Pearson Correlation	.267	1
	Sig. (2-tailed)	.077	
	N	45	45

$$H_0: r = 0$$
$$H_1: r \neq 0$$

The correlation is $r(45) = 0.267$, $sig = 0.038 (< 0.05)$.
This test is significant.
Reject H_0.
Conclusion: There is a relationship.

1-Tailed Testing

Correlations

		Essay Mark (%)	Hours Spent on Essay
Essay Mark (%)	Pearson Correlation	1	.267*
	Sig. (1-tailed)		.038
	Sum of Squares and Cross-products	2009.060	216.101
	Covariance	45.660	4.911
	N	45	45
Hours Spent on Essay	Pearson Correlation	.267*	1
	Sig. (1-tailed)	.038	
	Sum of Squares and Cross-products	216.101	326.835
	Covariance	4.911	7.428
	N	45	45

*. Correlation is significant at the 0.05 level (1-tailed).

$$H_0: r = 0$$
$$H_1: r > 0$$

The correlation is $r(45) = 0.267$, $sig = 0.038 (< 0.05)$.
This test is significant.
Reject H_0.
Conclusion: There is a positive relationship.