Effective Use of Advanced Statistical Methods in Research II

KEHINDE OLUWADIYA

Professor of Surgery (Orthopaedics)

Ekiti State University, Ado-Ekiti

CEO

POSK Educational Consult

www.oluwadiya.com

Preamble

This is the second of a 3-part lecture on the use of advanced statistical methods in medical research

Objectives

PART I

1. Why do you need to know statistics?
2. What you need for effective use of statistics
3. Data transformation

PART II

1. Limitations of P-value
2. Statistics for comparing 2 or more groups with continuous data
3. Regressions and Correlation

PART III

4. Factorial and Covariate designs
5. Risk Ratios and Odds Ratios
6. Survival Analysis
7. Sensitivity, Specificity and ROC Curves
8. Finding the right test for specific data

The problem with P

- P values provide less information than confidence intervals.
  - Statistical significance tells us that there is a difference
  - But it does not tell us the magnitude of the difference .
  - A P value provides only a probability that an estimate is due to chance
  - A P value could be statistically significant but of limited clinical significance.
    - A very large study might find that a difference of 0.5mmHg in BP between 2 rx groups is statistically significant but is this clinically relevant?

"A large study dooms you to statistical significance"

Anonymous Statistician

The problem with P

- Statistical tests in inferential statistics are designed to answer the question ‘‘how likely is the difference found in a sample due to chance?’’.
- This is the only purpose they serve: **the calculation of a probability value.**
- They do not indicate clinical significance!

Clinical significance

- The clinical significance of a research finding – the extent to which it may influence clinical practice – depends on many factors.
  - I. Many of these factors are related to study design.
  - II. are the adverse effects of a treatment also studied in addition to benefits?
  - III. Are the outcome measures clinically relevant (e.g., improvement in symptoms and functioning rather than laboratory measurements)?
  - IV. Are the effects lasting?
  - V. Is the cost of treatment worth the effect
  - VI. Can the study findings be generalized to patients across social and clinical settings?

Magnitude and clinical significance

- Cohen’s d
- Odds ratio
- Relative Risk
- Absolute Risk Reduction (ARR)
- Numbers needed to treat (NNT): this is the reciprocal of ARR

Magnitude and clinical significance

- Risk of response on a test drug = **0.6**
- Risk of response on placebo = **0.4**
- Absolute Risk Reduction = 0.2
- Number Needed to Treat (NNT) = 1/0.2 = **5**

	No response	**response**	**Total**
Placebo	A(60)	B(40)	A+B(100)
Drug	C(40)	D(60)	A+B(100)

STATISTICS FOR COMPARING TWO OR MORE GROUPS WITH CONTINUOUS DATA

Statistical tests and their appropriate data

Parametric tests

- Continuous data; normally distributed

Non-parametric tests

- Continuous data; not normally distributed (Categorical or Ordinal data)

Statistical tests and their appropriate data

Parametric tests

- Continuous data; normally distributed

Non-parametric tests

- Continuous data; not normally distributed (Categorical or Ordinal data)

Comparison of two sample mean

- Student's T test
  - Assumes normally distributed continuous dependent data.
  - Used when the independent variable has two categories e.g. Sex.
  - No need to do the math, commonly generated by most statistics software

But…

  - Understand the underlying theory and assumption

Student t test: Variable requirements

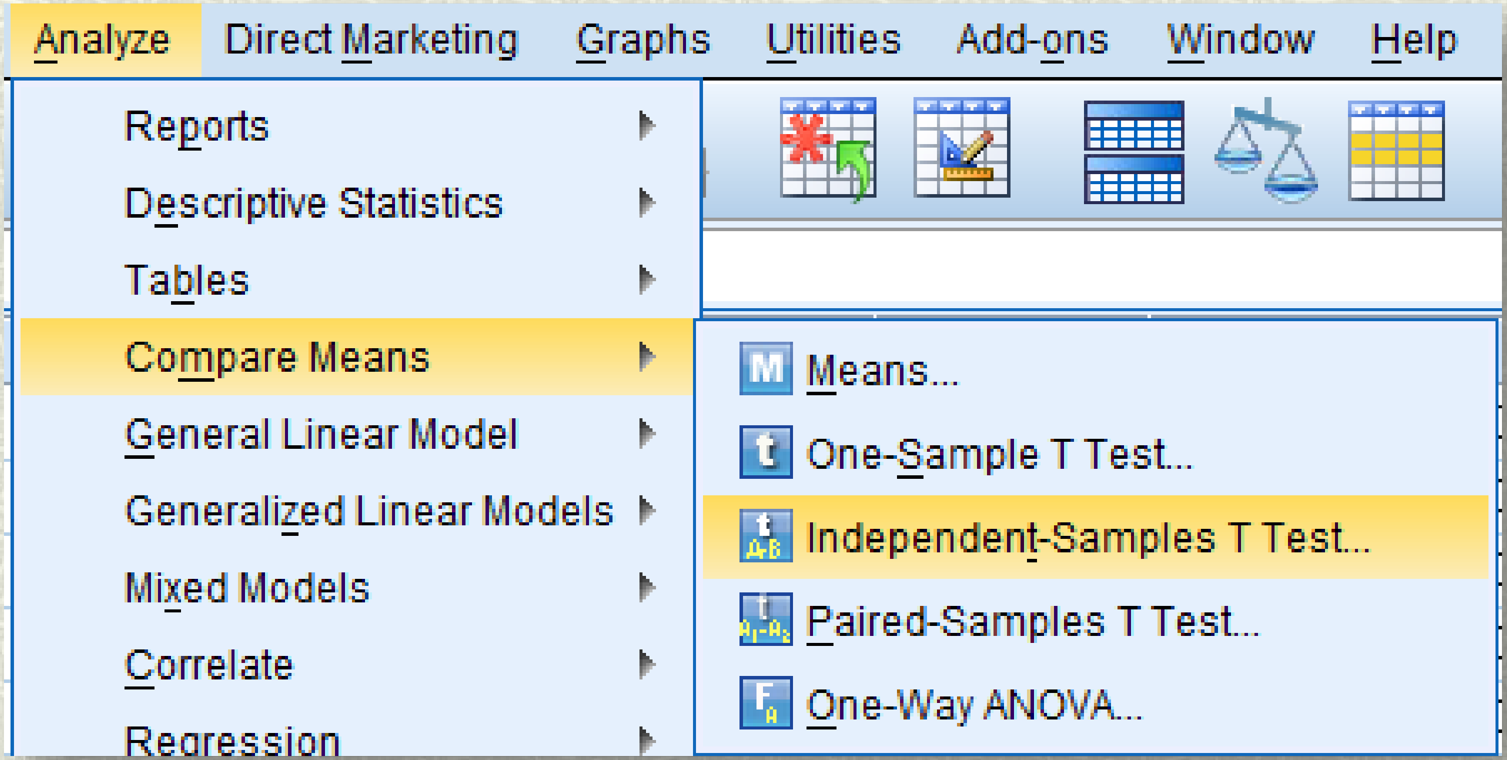
- Two variables are needed for the t-test procedure:

  A. One independent variable, which must be categorical e.g. *gender* (male/female).

  B. One continuous, dependent variable e.g. *pcv*.

Student t-test in SPSS

This is called Independent Samples T Test in SPSS



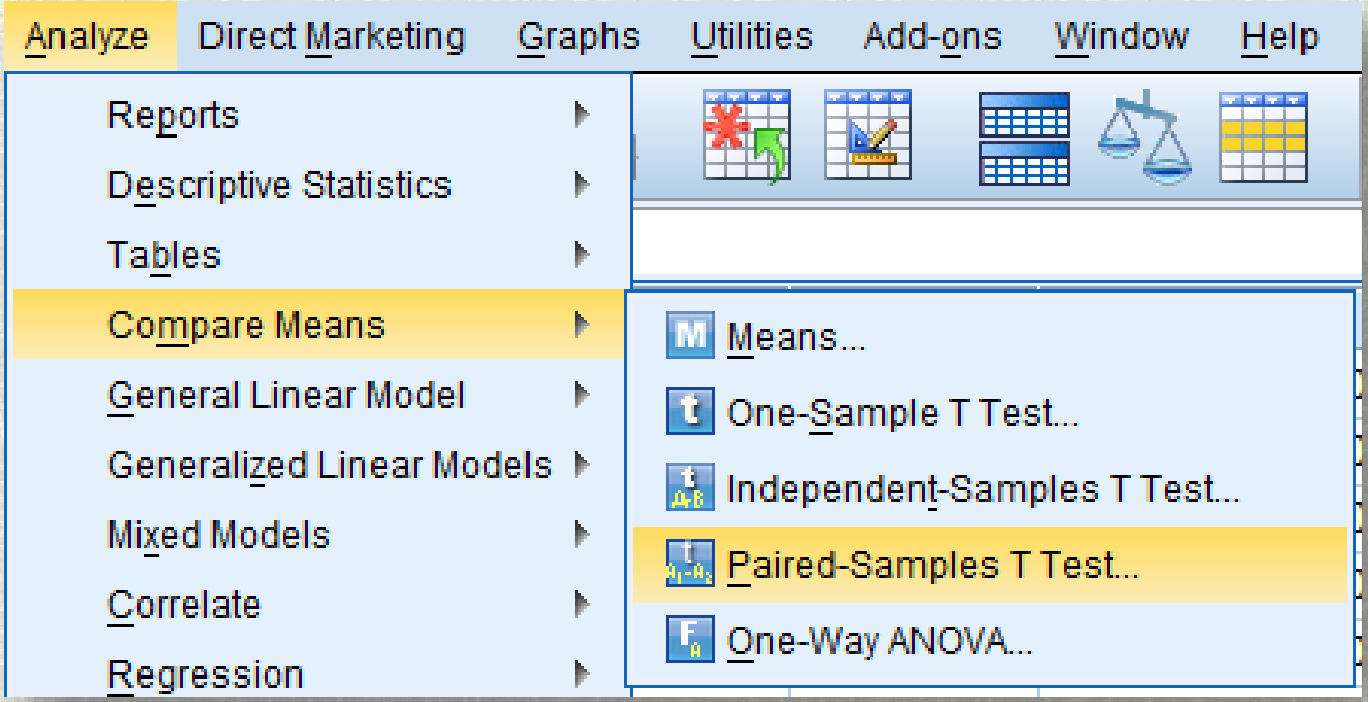
Comparison of two related samples

- **Paired T-Test**
  - Whereas T-test assumes there is independence of observations
  - Related samples i.e. Paired T-test is meant for “before” and “after” studies (crossover designs)

Paired sample t-test: Variable requirements

- Two variables, both continuous. But the measurements are taken in either of two ways:
  I. Measurements are taken twice in the same subject (Before and after an intervention)
  II. Matched-pairs or case-control studies, and the response for each are taken twice

Paired t-test in SPSS



Comparison of >2 sample means

- **AN**alysis **O**f **VA**riance (ANOVA)
  - Used to determine if two or more samples are from the same population i.e. no significant difference between their means

Requires that….

- Dependent variable is continuous data
- Independent variable is categorical data
- Independent variable = Grouping variable = Factor
  - This variable will consist of a number of categories or levels
  - There will be 2 or more of such categories or levels.
  - If there are only 2 categories, then the result will be identical to t-test

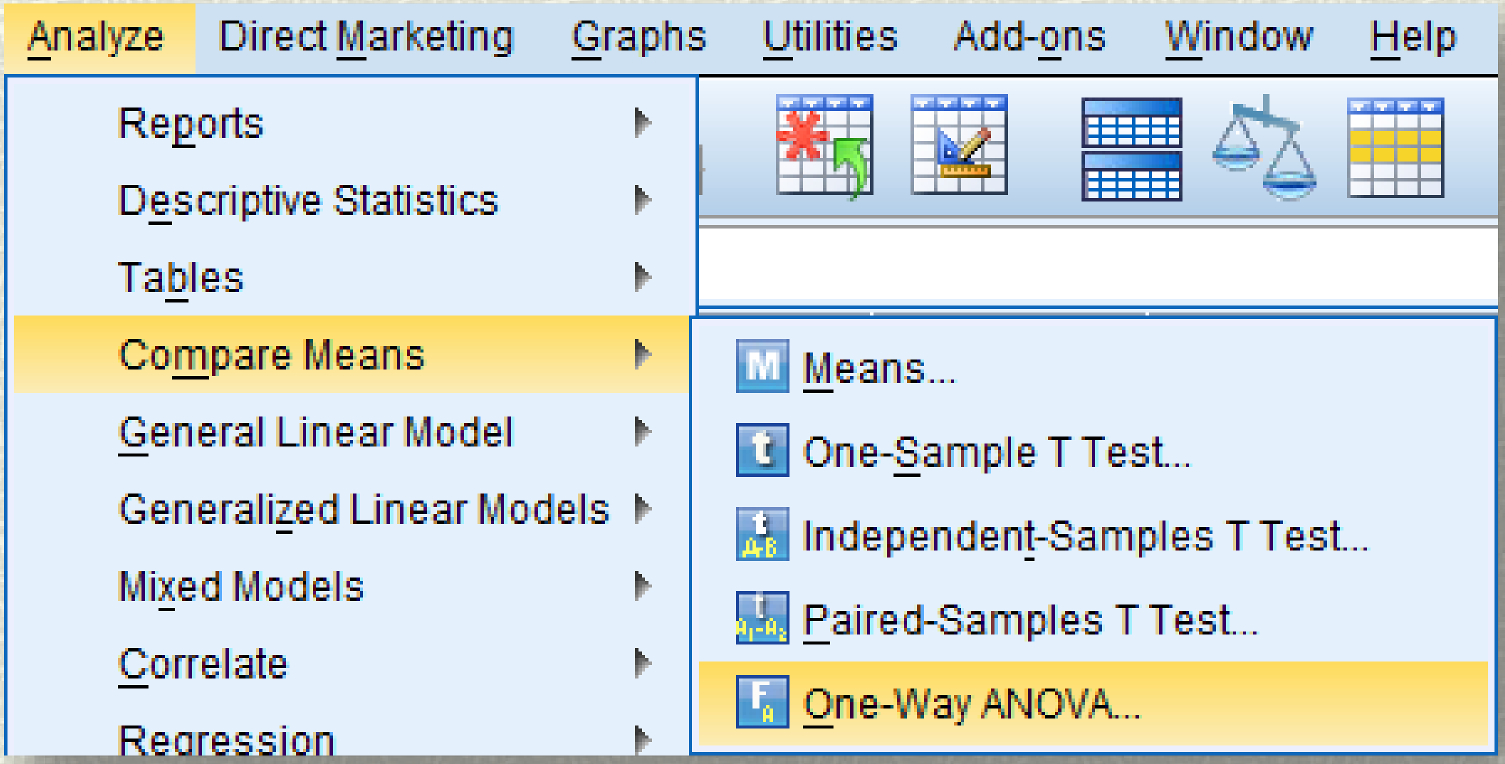
.

ANalysis **O**f **VA**riance

An example:

- You measure the PCV of 50 patients who have sustained fractures.
- Does the number of bones fractured affect the mean PCV of the patients ?
    - *number of bone fractured* is the independent variable or factor. It has 4 categories:
        - 0 bone fractured is one level of the factor
        - 1 bone fractured is one level of the factor
        - 2 bones fractured is one level of the factor
        - 3 bones fractured is one level of the factor
    - PCV is the dependent variable

ANOVA in SPSS



ANOVA in SPSS

Report

PCV

number of	Mean	N	Std. Deviation
0	39.83	6	7.387
1	32.17	12	8.277
2	25.20	5	7.362
3	24.20	5	5.975
Total	31.14	28	9.168

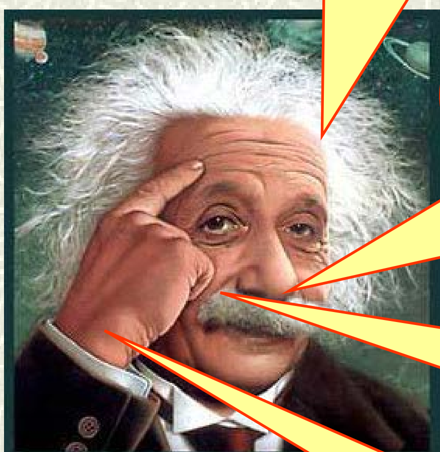
The objective is to determine if the mean PCV as shown is significantly different from each other, or they are due to chance

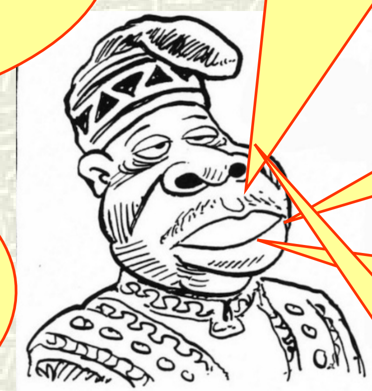
ANOVA

PCV

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	883.329	3	294.443	5.098	.007
Within Groups	1386.100	24	57.754		
Total	2269.429	27			

Significant result…now what?





Segun, there are 4 means!

Yes Albert! And there is a significant difference between the means, too. I'm so happy😂

Not so fast! Is the significance among all the means, or just some of them?

Don't know 😒

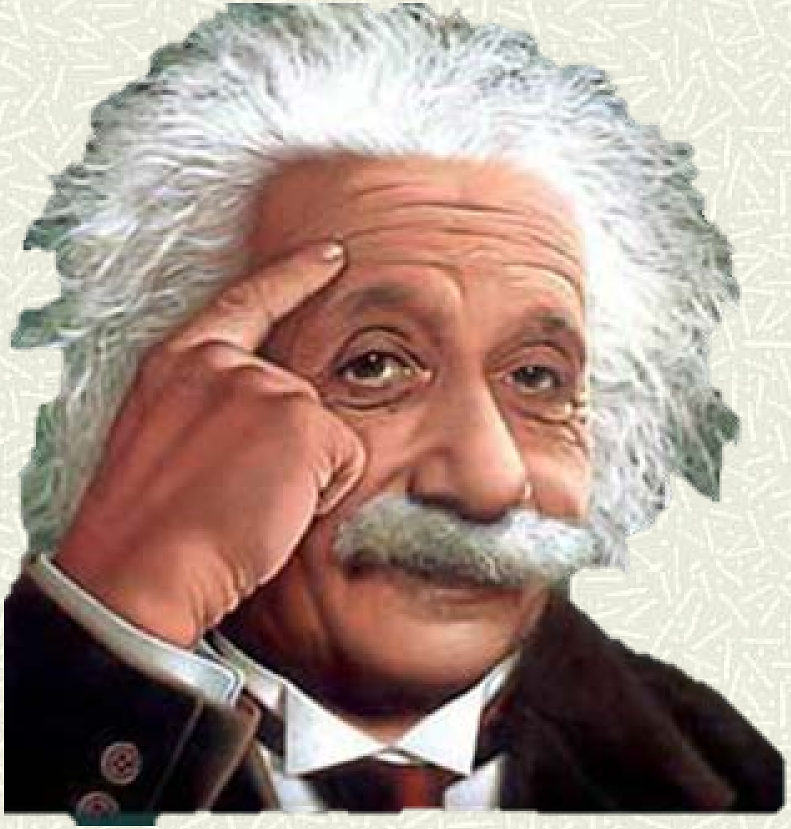
You should do a **post hoc** test my friend.

Post what?

You need to show me.

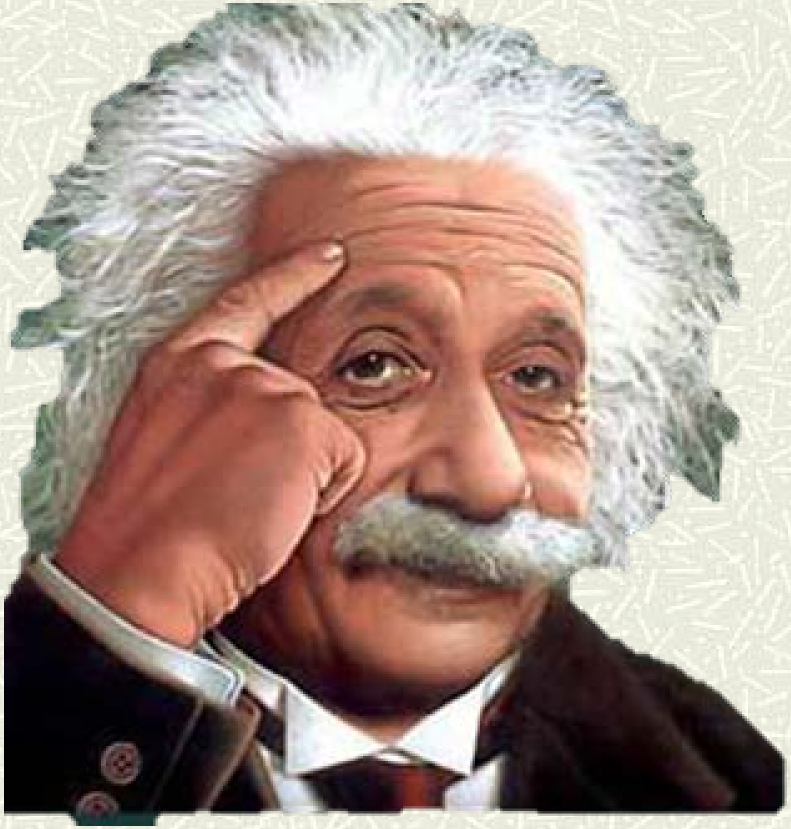
A Post Hoc test will show you which of the means are really different from each other

Post Hoc tests



- *After the Fact* comparisons of means used to identify which specific pairs of means are significantly different
- Designed to reduce **α** errors *regardless* of how many pairs of means are compared

Post Hoc tests



- Also called follow-up tests
  - Should be computed only after a significant ANOVA
  - They are like a collection of little t-tests
  - But they control overall type 1 error comparatively well
  - They do not have as much power as the omnibus test (the main ANOVA) – so you might get a significant ANOVA & no significant post hoc!
  - Purpose is to identify the means that are actually different from each other without increasing the probability of type 1 errors.

Post Hoc tests: Why it is preferred to multiple t-tests

- For a statistical test, e.g., t-test with a particular α value e.g. $\alpha = 0.05$, if the null hypothesis is true then the probability of not obtaining a significant result is $1 – 0.05 = 0.95$.
- Multiply 0.95 by the number of tests to calculate the probability of not obtaining one or more significant results across all tests.
- For two tests, the probability of not obtaining one or more significant results is $0.95 \times 0.95 = 0.9025$.
- Subtract that result from 1.00 to calculate the probability of making at least one type I error with two multiple tests:
- **1 – 0.9025 = 0.0975. = 9.75%**

Post Hoc tests: Why it is preferred to multiple t-tests

Formular is: $\mathbf{1-(1-\alpha)^n}$

α= alpha level, n= number of tests

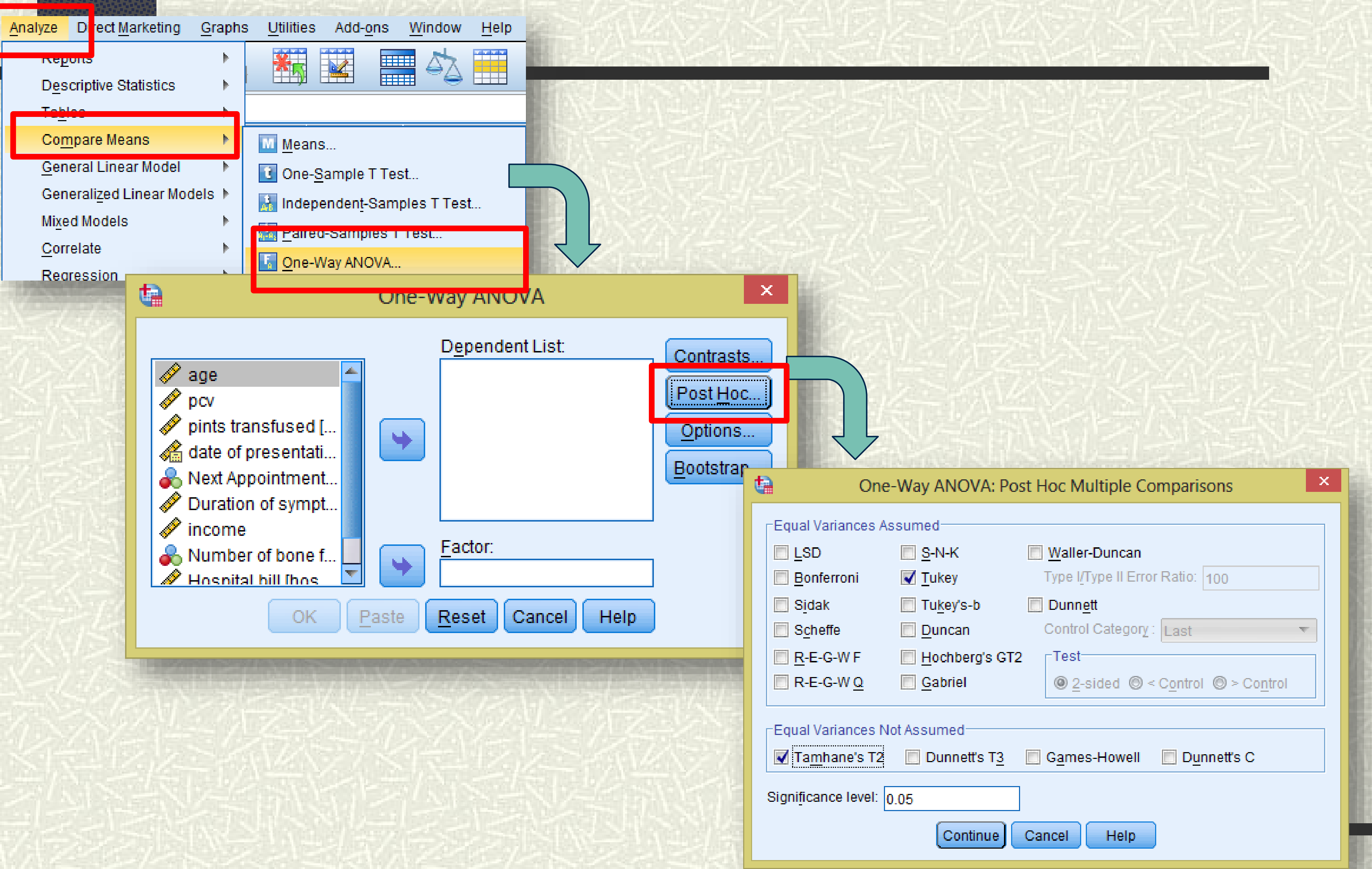
Example: You want to compare 4 groups (A, B, C, D).

You will have six pairs (α= 0.05 for each): A vs B, B vs C, C vs D, A vs C, A vs D, and B vs D.

Using the formula, the probability of not obtaining a significant result is 1 – (1 – 0.05)6 = **0.265**, which means your chances of incorrectly rejecting the null hypothesis (a type I error) is about 1 in 4 instead of 1 in 20 for a single t-test!!

Post hoc tests correct for these errors.

Post Hoc tests in SPSS



Post Hoc tests

Multiple Comparisons

Dependent Variable: PCV

	(I) number of bones fractured	(J) number of bones fractured	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	0	1	7.67	3.800	.210	-2.82	18.15
		2	14.63*	4.602	.020	1.94	27.33
		3	15.63*	4.602	.012	2.94	28.33
	1	0	-7.67	3.800	.210	-18.15	2.82
		2	6.97	4.045	.335	-4.19	18.13
		3	7.97	4.045	.227	-3.19	19.13
	2	0	-14.63*	4.602	.020	-27.33	-1.94
		1	-6.97	4.045	.335	-18.13	4.19
		3	1.00	4.806	.997	-12.26	14.26
	3	0	-15.63*	4.602	.012	-28.33	-2.94
		1	-7.97	4.045	.227	-19.13	3.19
		2	-1.00	4.806	.997	-14.26	12.26

Post Hoc tests: Homogeneous subset

PCV

	number of bones fractured	N	Subset for alpha = .05	
			1	2
Tukey HSD[a,b]	3	5	24.20	
	2	5	25.20	
	1	12	32.17	32.17
	0	6		39.83
	Sig.		.281	.312

Means for groups in homogeneous subsets are displayed.

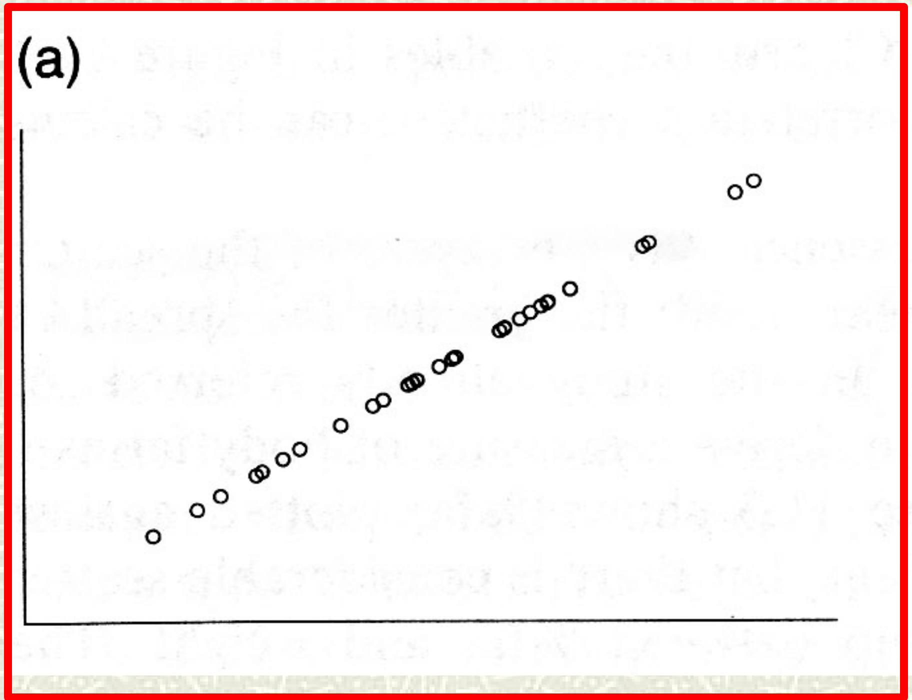
a. Uses Harmonic Mean Sample Size = 6.154.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

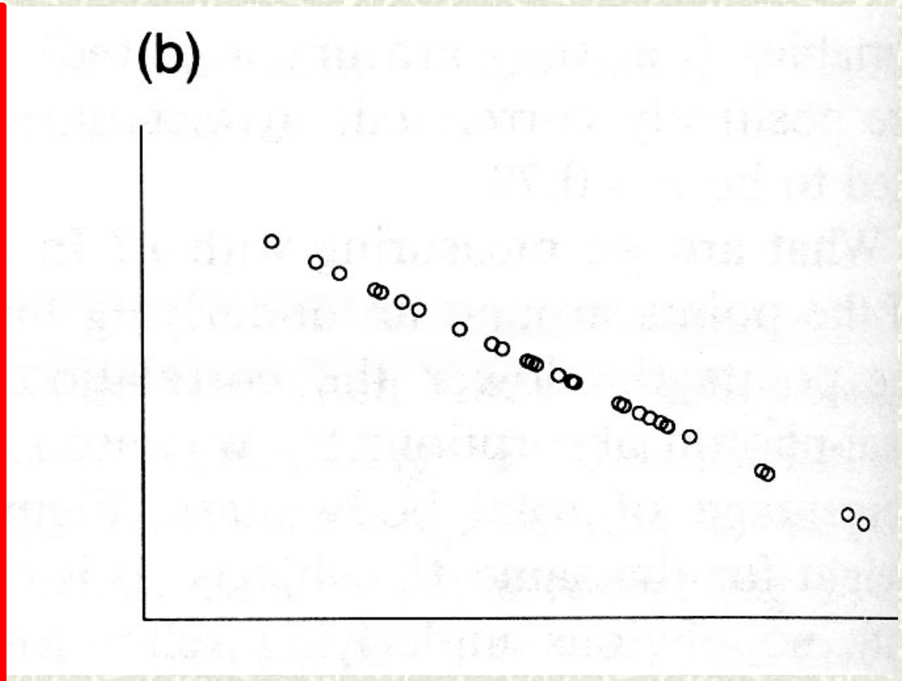
Correlation

- Assesses the linear relationship between two continuous variables
  - Example: height and weight
- Strength of the association is described by a correlation coefficient- *r*
  - *r* = 0 - .2 low, probably meaningless
  - *r* = .2 - .4 low, possible importance
  - *r* = .4 - .6 moderate correlation
  - *r* = .6 - .8 high correlation
  - *r* = .8 - 1 very high correlation
- Can be positive or negative
- Pearson’s or Spearman’s correlation coefficient
- Tells nothing about causation

Correlation

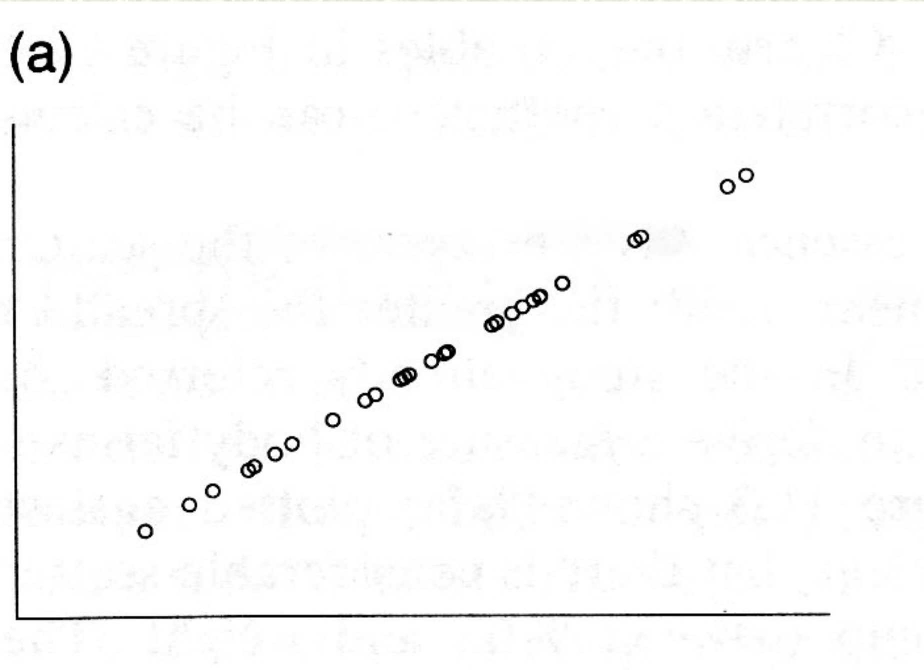


Positive Correlation

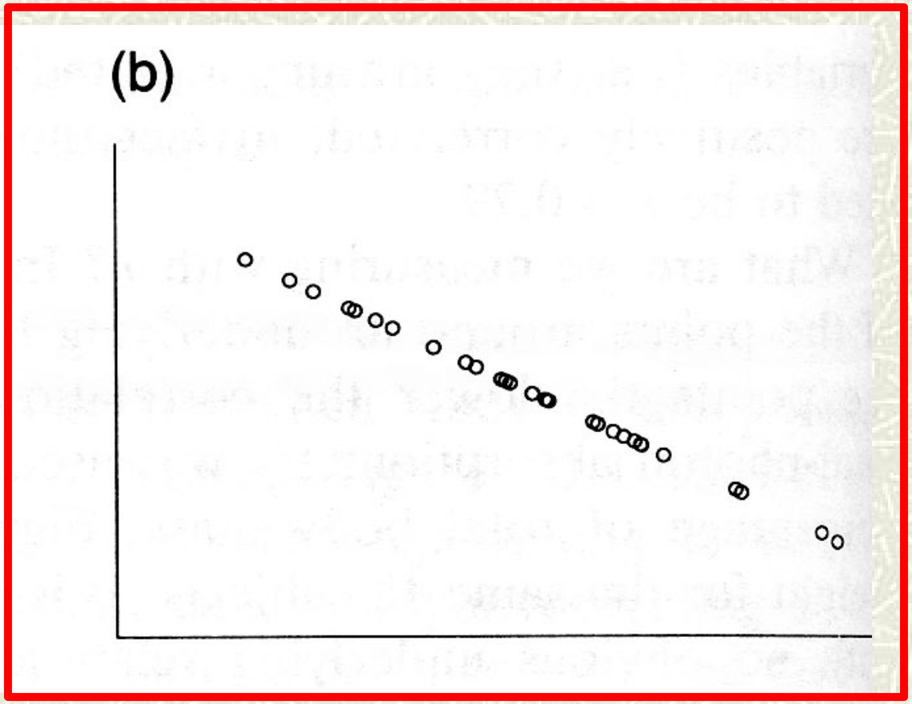


Negative Correlation

Correlation

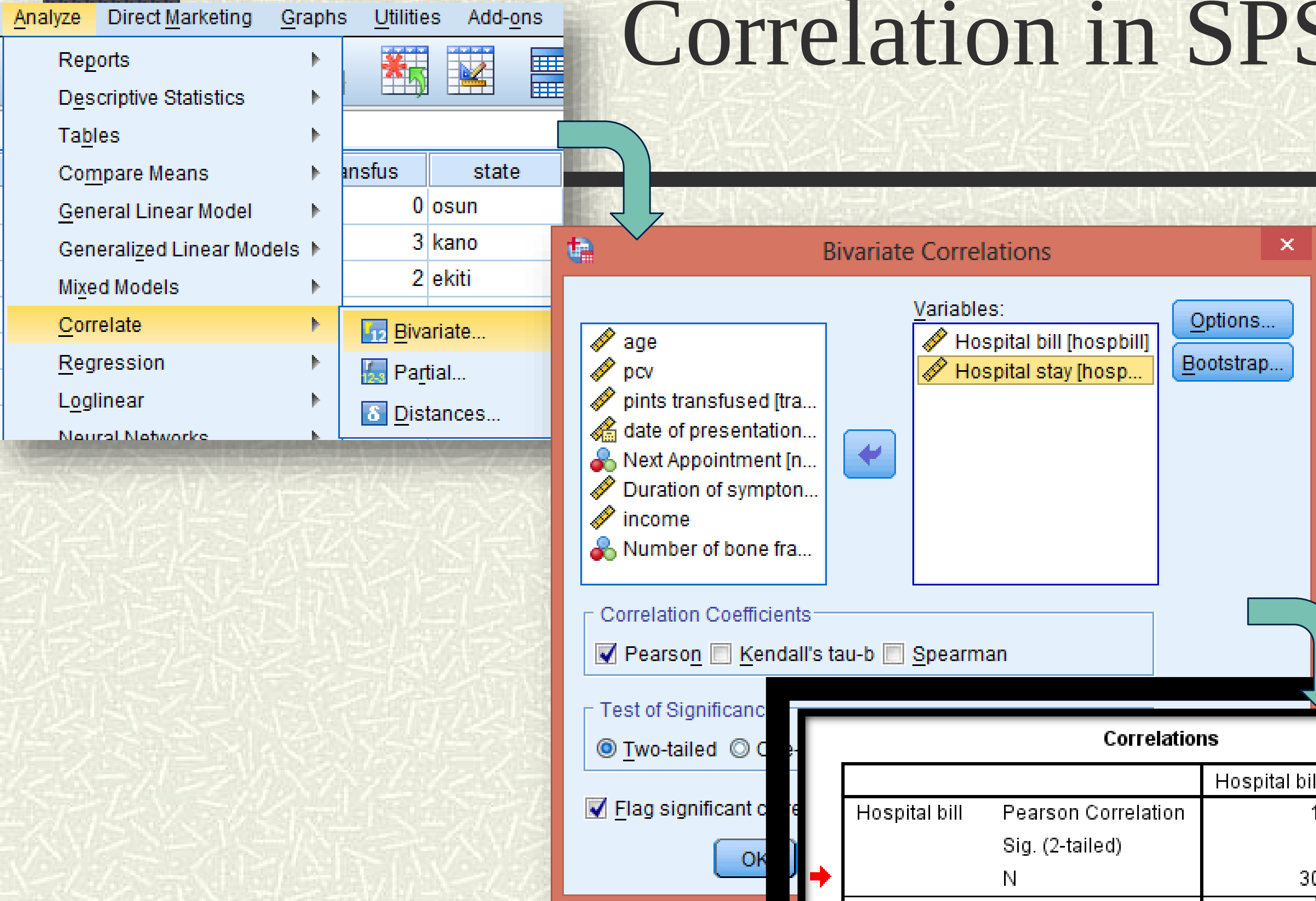


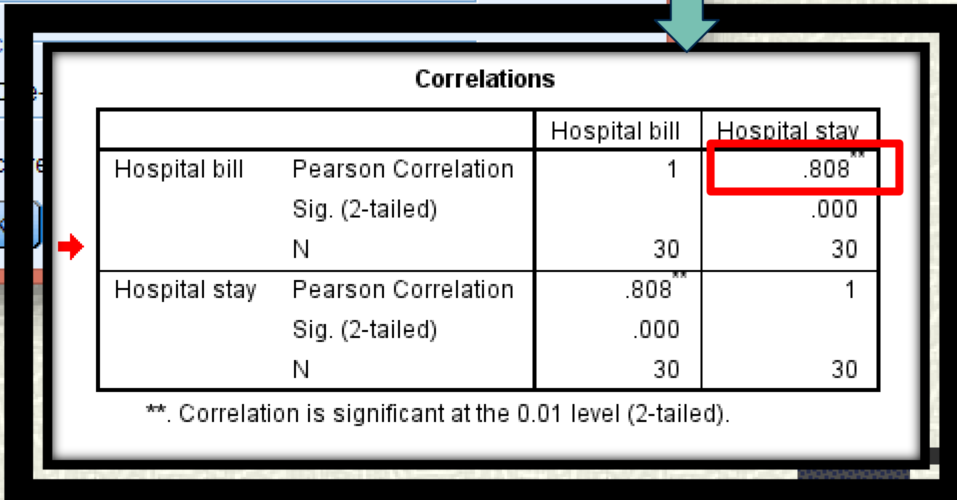
Positive Correlation



Negative Correlation

Correlation in SPSS





Correlations

		Hospital bill	Hospital stay
Hospital bill	Pearson Correlation	1	.808**
	Sig. (2-tailed)		.000
	N	30	30
Hospital stay	Pearson Correlation	.808**	1
	Sig. (2-tailed)	.000	
	N	30	30

**. Correlation is significant at the 0.01 level (2-tailed).

Regression

- Based on fitting a line to data
  - Provides a regression coefficient, which is the slope of the line
    - For example; $y = \beta_0 + \beta_1 x$ (simple linear regression)
  - Use to predict a dependent variable's value based on the value of an independent variable.
    - E.g., in analysis of height and weight, for a known height, one can predict weight.
- Much more useful than correlation
  - Allows prediction of values of y rather than just knowing the relationship between the two variables.

Regression

- Types of regression
  - Linear and Multiple - uses continuous data to predict continuous data outcome
  - Logistic- uses continuous and categorical data to predict probability of a dichotomous outcome
  - Poisson regression- predict a dependent variable that consists of "count data" given one or more independent variables.
  - Cox proportional hazards regression- survival analysis.

Simple Linear Regression Equation

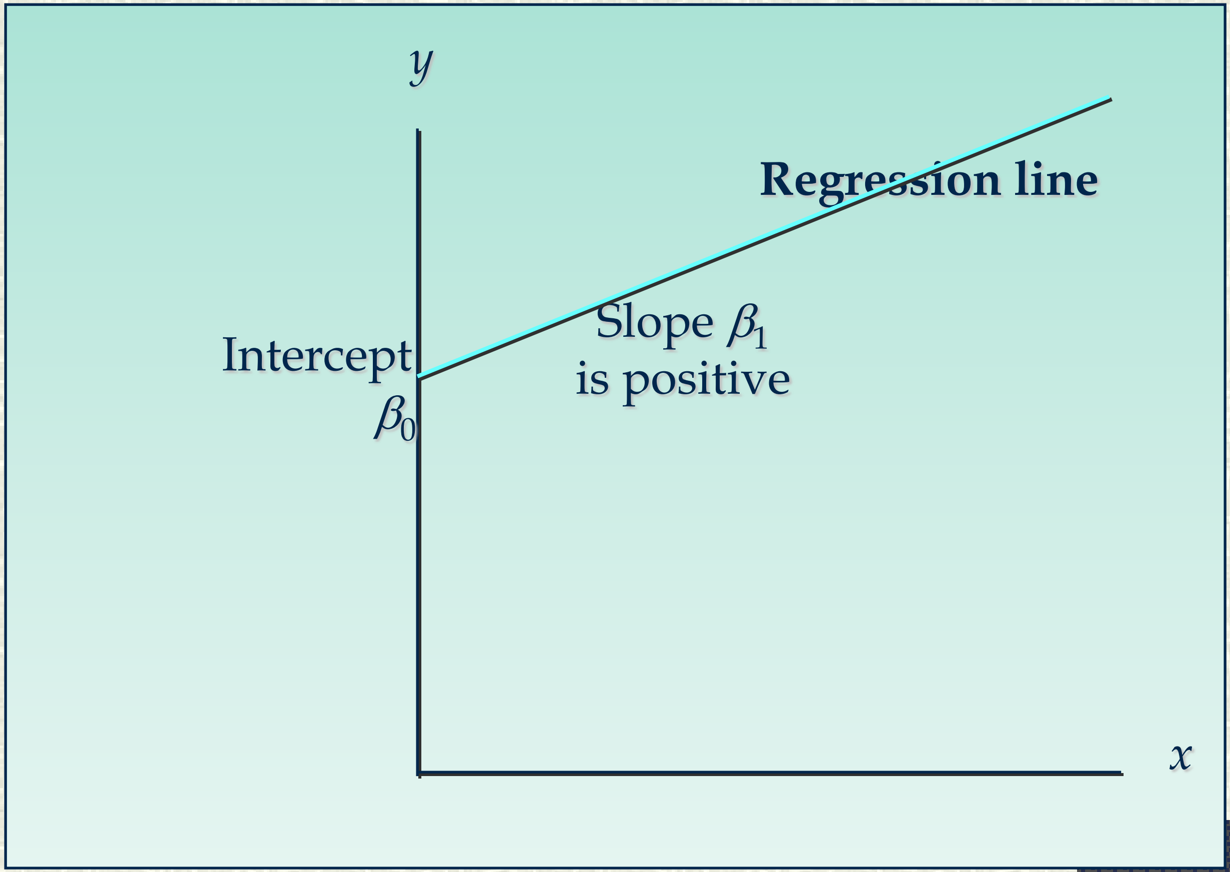
- The simple linear regression equation is:

$$y = \beta_0 + \beta_1 x$$

  - Graph of the regression equation is a straight line.
  - β_0 is the intercept of the regression line on the y axis.
  - β_1 is the slope of the regression line.
  - y is the expected value of y for a given x value.

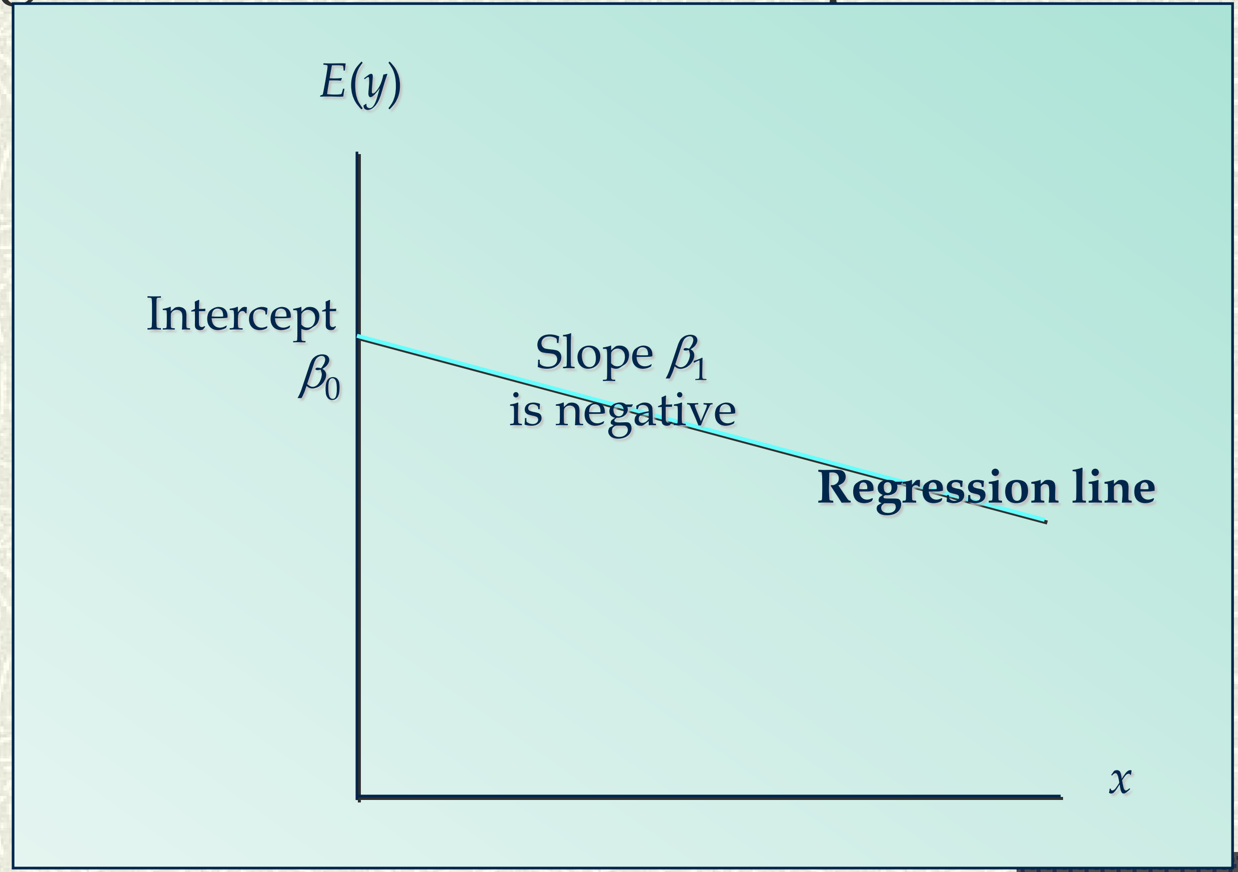
Simple Linear Regression Equation

- Positive Linear Relationship



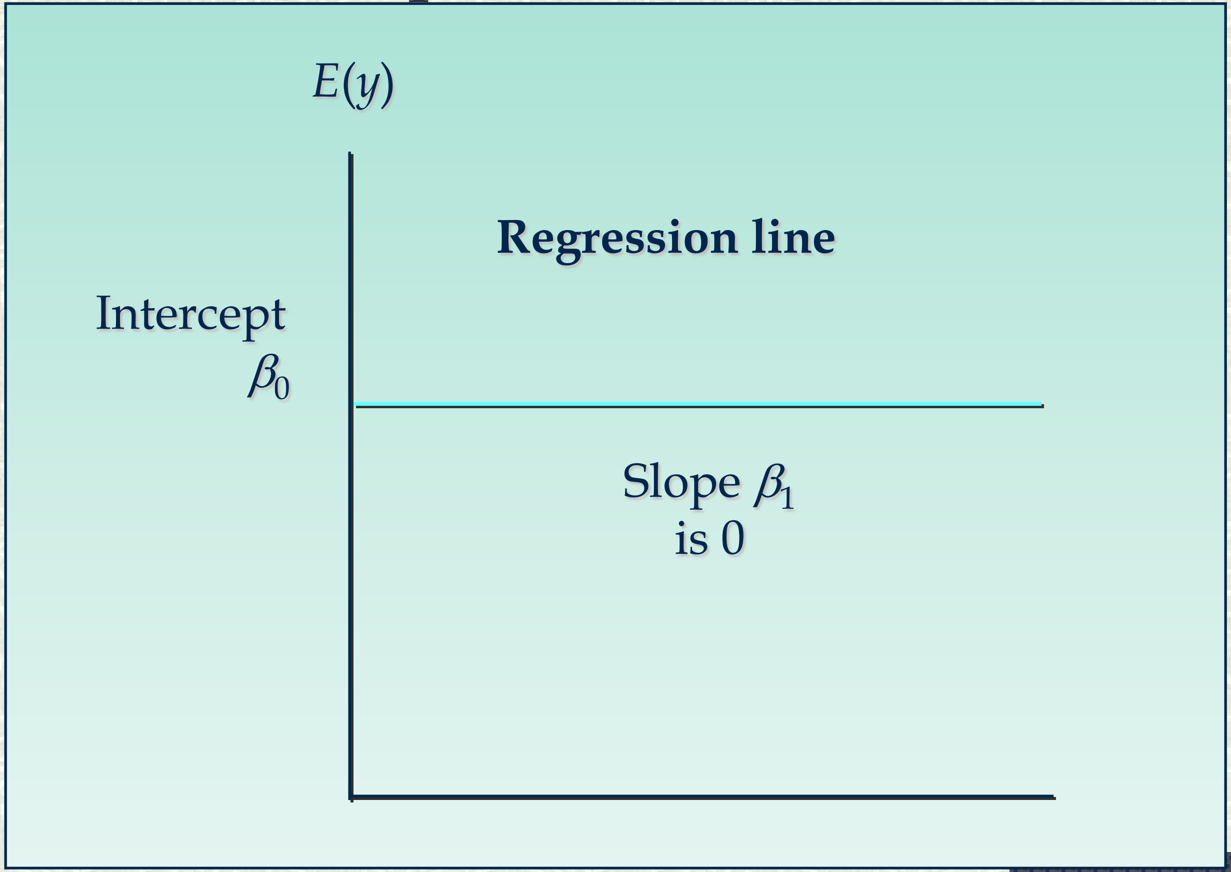
Simple Linear Regression Equation

- Negative Linear Relationship



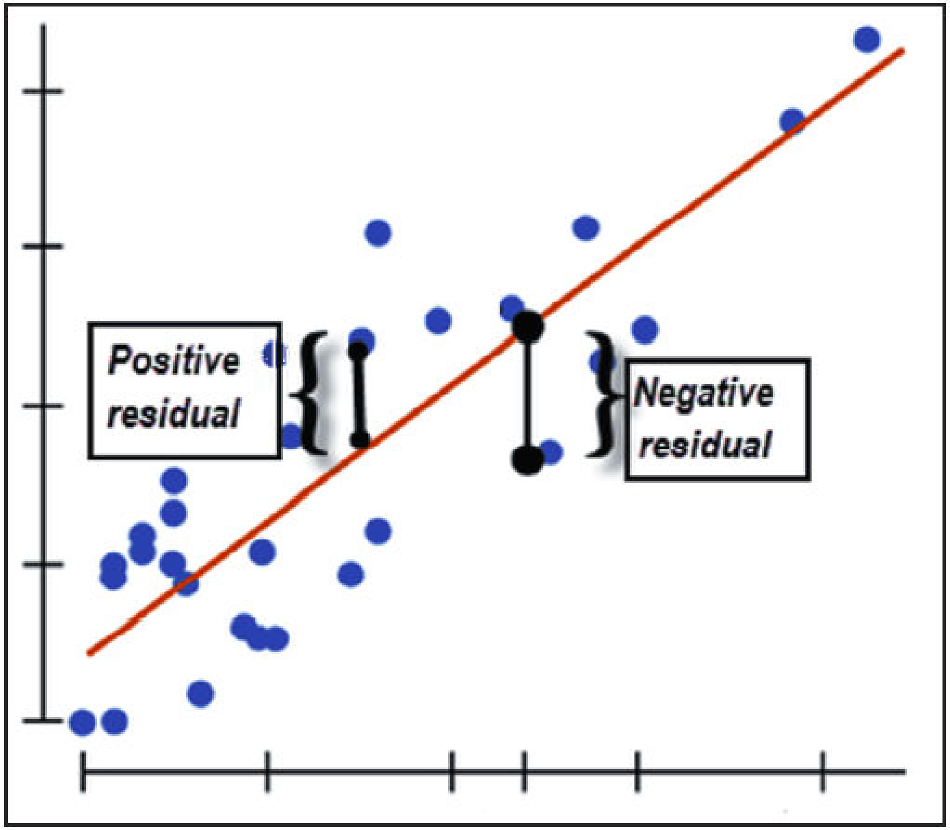
Simple Linear Regression Equation

- No Relationship



Important terms: Residuals

- What are the residuals?
- Residuals are the differences between the regression line and actual cases in the collected data.
- Large residuals mean that the model is not fitting the data very well
- Small residuals implies that the model is doing a better job at fitting the data



Important terms: Goodness of Fit

- The goodness of fit of a model is how well it fits or represents the actual data.
- Estimated by
  1. ANOVA
  2. The Coefficient of Determination (R^2)

Causes of poor model fit

- Two important factors may cause poor fit of a model
  1. Outliers
  2. Influential cases

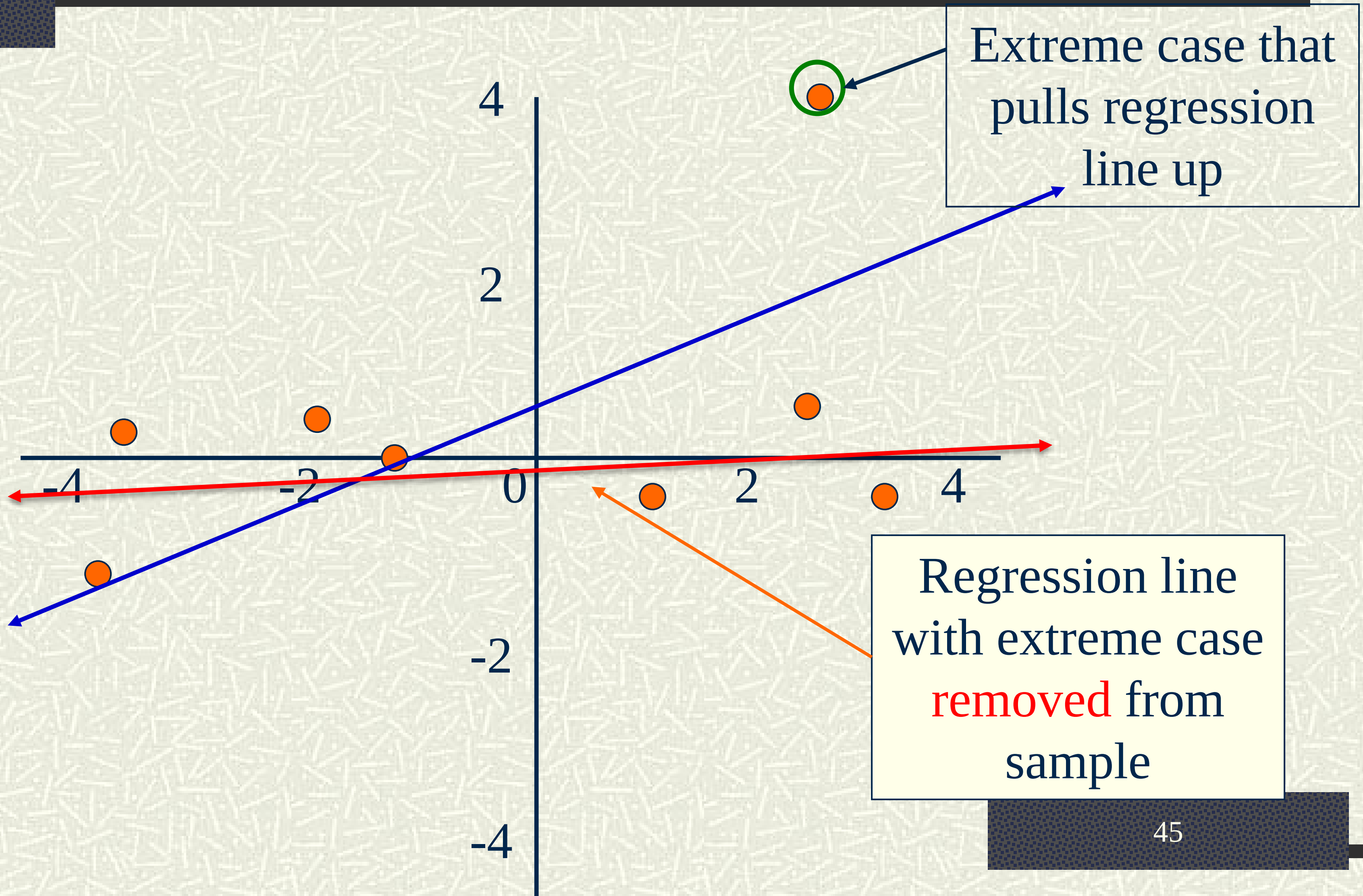
Outliers

- This are extreme cases
- They are cases with large residuals
- They pull the regression line towards themselves
- Some serious outlier scenarios include:
  1. If a case has a standardized residual greater than 3.
  2. If more than 1% of the cases have standardized residuals greater than 2.58.
  3. If more than 5% of cases have standardized residuals with an absolute value greater than 1.96.

Outliers

- Outliers can result from:
  - Errors in coding or data entry (→**rectify**)
  - Highly unusual cases (→**exclude?**)
  - Or sometimes they reflect important “real” variation (→**include?)**

Outliers: Example



Influential Cases

- Some cases have unusually high effect on the regression model.
- A case is influential if omitting it from the analysis gives a very different model
- Influential cases can be evaluated by *Cook's Distance* which is a measure of the overall influence of a case on the model. It is significant when it is greater than 1.

Multicollinearity

- This is when some of the independent variables are highly correlated
- If two independent variables are closely related its difficult to estimate their regression coefficients because they may be predicting the same thing.
- Solution is to eliminate one of them

Multicollinearity

- Multicollinearity is evaluated by doing collinear diagnostics.
  - Do correlation which will show collinearity between individual variables
  - Check Tolerance statistics which will show collinearity between groups. It should be less than 1 and (usually) greater than 0.5

Homoscedasticity

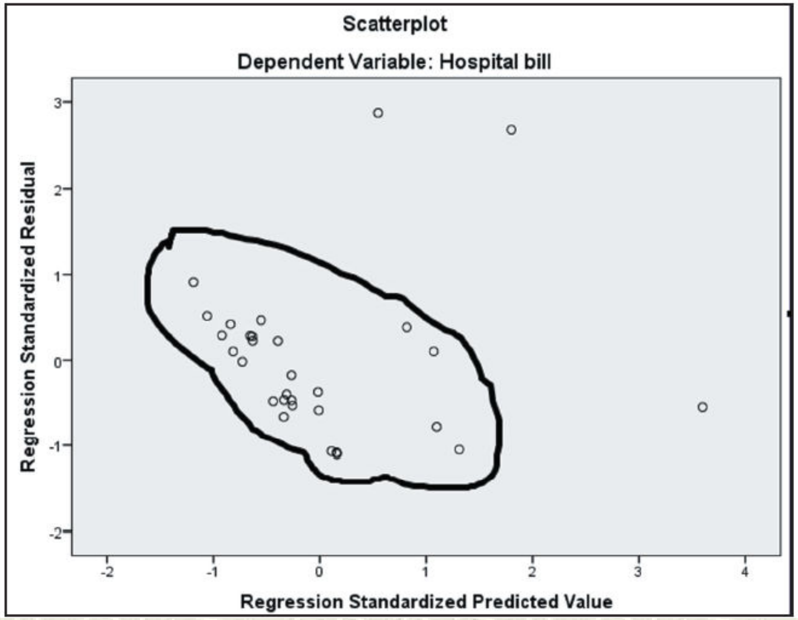
Also called equality of variance

This refers to the assumption that the dependent variable (γ) exhibits similar amounts of variance across the range of values for the independent variables (χ)

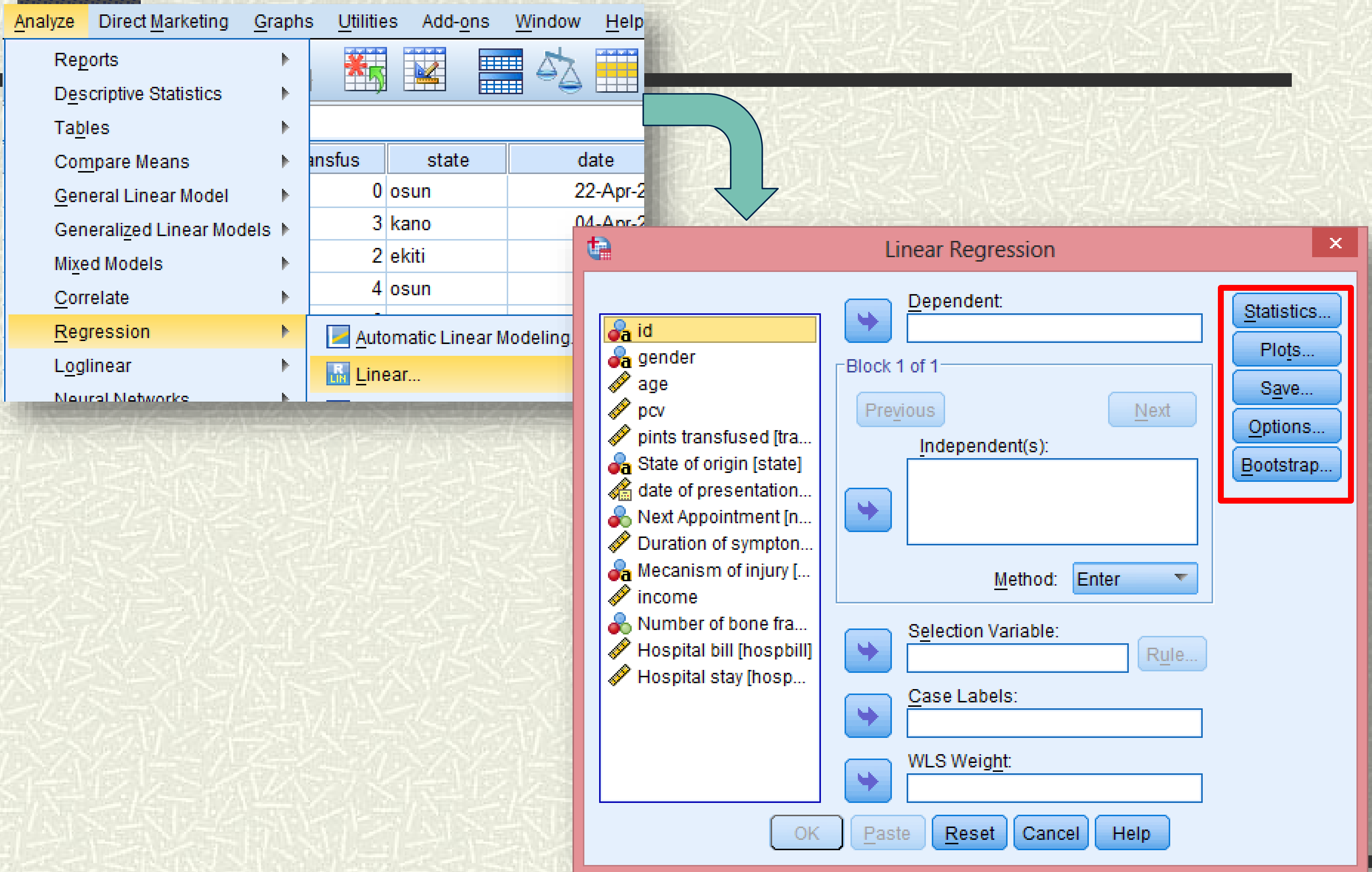
The opposite of homoscedasticity is referred to as heteroscedasticity.

Homoscedasticity

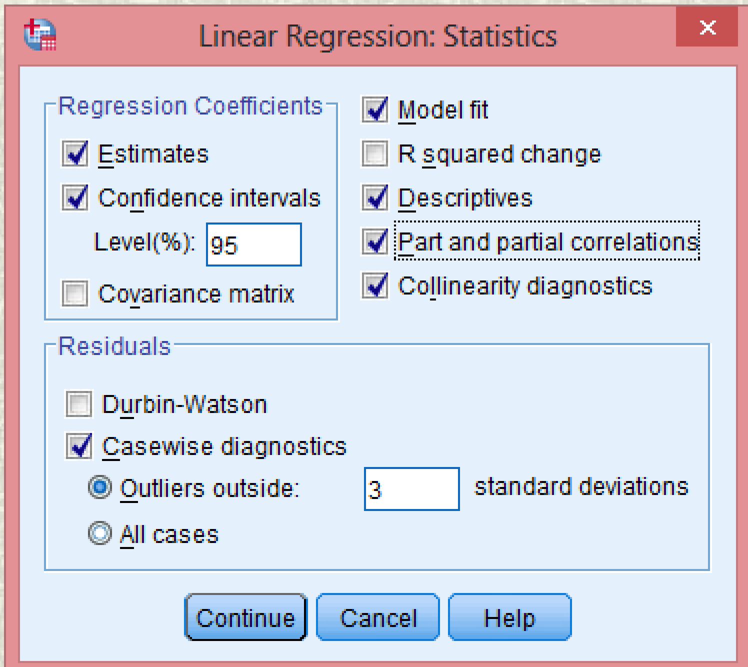
- It is checked by plotting standardized predicted values (ZPRED) against standardized residuals (ZRESID)
- There is homoscedasticity when the scatter plot shows no discernible pattern and is clustered around zero.

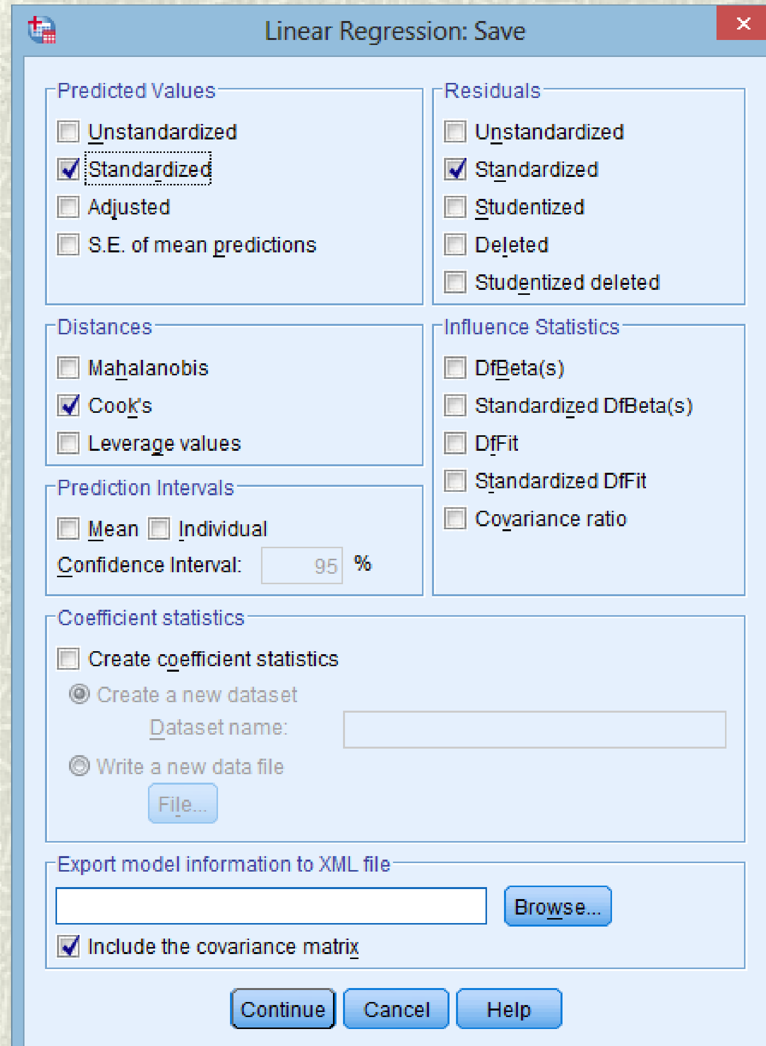


Multiple regression with SPSS



Multiple regression with SPSS





Multiple regression with SPSS

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.857[a]	.734	.703	5025.355

a. Predictors: (Constant), Duration of symptons in days, Hospital stay, Number of bone fractured

b. Dependent Variable: Hospital bill

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1808099847	3	602699948.9	23.865	.000[b]
	Residual	656609012.9	26	25254192.80		
	Total	2464708859	29			

Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	-442.701	1643.014		-.269	.790	-3819.965	2934.563
	Hospital stay	166.153	39.444	.568	4.212	.000	85.074	247.232
	Number of bone fractured	3540.993	1284.139	.379	2.757	.011	901.408	6180.578
	Duration of symptons in days	-67.621	62.542	-.112	-1.081	.290	-196.178	60.936

a. Dependent Variable: Hospital bill

Logistic Regression

- Non-parametric equivalent to Linear Regression
- Used when the dependent variable is dichotomous
- The independent variables may be categorical or continuous
- Odds ratio are calculated

What is the difference between crude and adjusted Odds Ratio

THIS BRINGS US TO THE END OF PART II

About Me

Oluwadiya Kehinde

- Professor of Surgery at the Ekiti State University, Ado-Ekiti
- Author of "**Getting to Know SPSS**", the best selling book on SPSS in Nigeria
- CEO of **POSK Educational Consult**, Consultancy Firm for Training in Statistical and Health Education

www.Oluwadiya.com

Thanks for your attention

To ask questions, please join the forum at www.oluwadiya.com