Statistical Tools

Dr. K. A. Korb
University of Jos

Descriptive Statistics

- Three main types of descriptive statistics are:
  - Frequency (and Percentage)
    - Count the number of times that each category occurs
  - Measures of central tendency (Average: Mean, Median, Mode)
    - One number that best represents the entire set of scores
  - Measures of dispersion (Range, Standard Deviation)
    - Degree to which scores vary (deviate) around the average

Frequency

- **Frequency:** Tally the number of people in each category
- **Percentage**
  - Step 1: Frequency in a category divided by the total number
  - Step 2: Multiply by 100%

Calculating Percentage

Educational Qualification	Freq.	Divided by total	Times 100%	2 decimal points
NCE	37	0.486842	48.68421%	48.68%
First Degree	22	0.289474	28.94737%	28.95%
PGDE	15	0.197368	19.73684%	19.74%
MEd/PhD	2	0.026316	2.631579%	2.63%
Total	**76**	**1.00000**	**100.000%**	**100.00%**

Central Tendency

Note: N = Number in the Sample

- **Mode:** Most frequent response
- **Mean**
  - Step 1: Sum all of the scores
  - Step 2: Divide by N

Years Experience as a Teacher

Data

10	**7**	**24**	**1**	**2**	**8**	**2**	**3**	**5**	**8**

Step 1: Order from Lowest to Highest

1	**2**	**2**	**3**	**5**	**7**	**8**	**8**	**10**	**24**

- **Mode =** 2 and 8
- **Mean**: 1+2+2+3+5+7+8+8+10+24 = 70; 70 divided by 10 (N) = 7
  - **Mean = 7**

Measures of Dispersion

- **Range:** Highest minus lowest score
  - Generally also report the highest and lowest score
- **Standard Deviation:** Average (mean) "deviation" between each score and the mean based on the sample

Types of Variables

- **Discrete Variables:** Variables with distinct categories
- **Continuous Variables:** Variables that can be expressed numerically along a continuum

Identify Discrete and Continuous Variables

1. Gender
2. Age
3. Teaching subject
4. Income
5. Geopolitical zone
6. Intelligence
7. Marital Status
8. Level of Education
9. How much time a student studies.
10. Achievement in CRK
11. Has a teacher attended training on CA?
12. Level of intrinsic motivation

Appropriate Descriptive Statistics

- **Discrete Variables:** Frequency and percentage
  - Percentage is easier to interpret because it means "If there were 100 cases EXACTLY, how many would fit in the category."
  - Mode (most frequent response) can also be reported as a measure of central tendency
  - When developing a Table, it is often easiest to interpret by sorting the categories from highest to lowest

Appropriate Descriptive Statistics

- **Continuous Variables:** Mean is the best representation of the data
  - Also report the possible range of data
  - Report the standard deviation

Making Comparisons:

Scores on WAEC

Note: Scores were generated by a random number table.

	Gov't	**Private**
English	24.13	28.18
Maths	58.90	27.87
Integrated Science	61.19	47.89
Chemistry	34.38	34.25
Social Studies	99.10	86.97
CRK/IRK	70.90	70.53

How do we know when the differences between government and private schools are large enough to be meaningful, or significant?

Foundation of Inferential Statistics

- It is rarely possible to conduct a study with the entire population
  - Researchers choose a sample of the population
  - Researchers want to draw conclusions *(inferences)* about the entire population based on the results from the sample
- The purpose of *inferential* statistics is to determine whether the findings from the sample can generalize to the entire population
  - *Are the results large – or significant – enough to conclude that the results from the sample also apply to the population?*

Overview of Inferential Statistics

- There will always be differences between groups in a research study
- Inferential statistics determines whether the difference between the two groups in the sample is large enough to be able to say that the findings are *significant.*
- If the findings are significant, the conclusions from the sample can be applied (generalized) to the entire population
- If the difference between the groups is very small, then the findings are *not significant* and therefore were simply the result of chance

Factors that Influence Significance

- **Sample size:** The larger the sample is, the more representative the sample is of the population
  - Larger sample sizes are more likely to obtain a significant result
- **The size of the difference between groups *(effect size).***
  - Larger effect sizes are more likely to obtain a significant result
    - Study variables that likely have a big effect size
    - In experimental studies, spend plenty of time making a very effective treatment to ensure a large effect size

Null Hypotheses

- There is **NO** significant difference between <groups> on <dependent variable>.
- There is **NO** significant effect of <independent variable> on <dependent variable>
- There is **NO** significant relationship between <variable 1> and <variable 2>
- Researchers want to assume that the results in a study are due to chance, unless there is compelling evidence otherwise
  - Prefer to err on the side of caution by saying there is **NO** significant result when there is, as opposed to saying there is a significant result when that is untrue

p Statistic

- **p:** Probability that the result is due to chance.
  - Ranges from 0.00000000000000000000 to 1.000.
  - The larger the p-value is (closer to 1.00), the more likely the result is due to chance
    - p=0.500 means the probability that the result is due to chance is 5 out of 10.
    - p=0.850 means the probability that the result is due to chance is 85 out of 100
    - p=0.050 means the probability that the result is due to chance is 5 out of 100

p Statistic

- Researchers want to be fairly certain of their conclusions, so they have decided the probability that the result is due to chance should be very small: less than 5 out of 100.
  - If p>.05, the probability is too great that the result is due to chance, so the null hypothesis is RETAINED
    - There is NO significant…
  - If p<.05, the probability is very low for the results to be due to chance, so the null hypothesis is REJECTED
    - There IS a significant…because results are most likely NOT due to chance
    - It can be concluded that a difference likely exists in the population, not just the sample

Inferential Statistics

- Comparisons between groups or between variables CANNOT be made without inferential statistics that determine whether the comparison is significant or not

Steps in Inferential Analyses

- In any inferential statistic, the first step is always to examine the p-value to determine if the result is significant
  - If not significant, analysis stops
  - If significant, then conduct post-hoc analyses if more than 2 groups are being compared
  - Examine descriptive statistics (mean, frequencies, correlation) to interpret the meaning of a significant result
    - This is oftentimes overlooked, but very important!

t-test

- ***t-test:*** Used when comparing two groups on a dependent variable
  - **Independent samples**: Two unique groups
  - **Correlated samples**: Groups are created by matched assignment or the same participants are compared on two variables
- Report the means and standard deviations for the two groups, the t, df, and two-tailed p

t-Test Example:

Comparison of Control (Group A) to Treatment (Group B) on Science Achievement

Data Summary

	A	B	Total
n	210	21	231
ΣX	58.1299999	7.68	65.8099999
ΣX^2	23.1461000	3.592	26.7381000
SS	7.0552	0.7833	7.9894
mean	0.2768	0.3657	0.2849

Results

$Mean_a - Mean_b$	t	df	p		
-0.0889	-2.1	229		one-tailed	0.0184115
				two-tailed	0.036823

t-Test Example

Comparison of Control (Group A) to Treatment (Group B) on Intrinsic Motivation

Data Summary

	A	B	Total
n	208	21	229
ΣX	137.049999	14.9299999	151.979999
ΣX^2	106.401699	12.2579000	118.659599
SS	16.1002	1.6434	17.7953
mean	0.6589	0.711	0.6637

Results

$Mean_a - Mean_b$	t	df	p		
-0.0521	-0.81	227		one-tailed	0.209394
				two-tailed	0.418788

Analysis of Variance (ANOVA)

- ANOVA: Compare three or more groups on one dependent variable
- One-way ANOVA: Compares multiple groups on the same DV
  - Significant p means that there is a significant difference between groups somewhere, NOT that there is a significant difference between all groups
  - A post-hoc test such as Tukey's HSD (Honestly Significant Difference) is needed to determine which differences are significant
  - Report means and standard deviations for all groups, ANOVA summary table, and the results of Tukey's HSD

One-Way ANOVA Example

Comparison of Control (1), Treatment A (2), and Treatment B (3) on Science Achievement

Data Summary

	Samples					
	1	2	3	4	5	Total
N	75	53	80			208
ΣX	50.97	35.65	50.43			137.05
Mean	0.6796	0.672642	0.630375			0.658894
ΣX^2	39.6067	27.4241	39.3709			106.4017
Variance	0.067128	0.066239	0.095963			0.077779
Std.Dev.	0.259091	0.257369	0.309779			0.278889
Std.Err.	0.029917	0.035352	0.034634			0.019337

standard weighted-means analysis

ANOVA Summary Independent Samples k=3

Source	SS	df	MS	F	P
Treatment [between groups]	0.107239	2	0.053619	0.69	0.502737
Error	15.993007	205	0.078015		
Ss/Bl					
Total	16.100246	207			

Graph Maker

One-Way ANOVA Example

Comparison of Control (1), Treatment A (2), and Treatment B (3) on Intrinsic Motivation

Data Summary

	Samples					
	1	2	3	4	5	Total
N	75	60	82			217
ΣX	17.86	19.66	24.81			62.33
Mean	0.238133	0.327667	0.302561			0.287235
ΣX^2	6.4658	9.4942	9.7061			25.6661
Variance	0.029902	0.051733	0.027155			0.035939
Std.Dev.	0.172922	0.22745	0.164788			0.189575
Std.Err.	0.019967	0.029364	0.018198			0.012869

standard weighted-means analysis

ANOVA Summary Independent Samples k=3

Source	SS	df	MS	F	P
Treatment [between groups]	0.298167	2	0.149083	4.27	0.015192
Error	7.464574	214	0.034881		
Ss/Bl					
Total	7.762741	216			

Graph Maker

Tukey HSD Test

HSD[.05]=0.07; HSD[.01]=0.09
M1 vs M2 P<.05
M1 vs M3 nonsignificant
M2 vs M3 nonsignificant

M1 = mean of Sample 1
M2 = mean of Sample 2
and so forth.

HSD = the absolute [unsigned] difference between any two sample means required for significance at the designated level. HSD[.05] for the .05 level; HSD[.01] for the .01 level.

Analysis of Covariance (ANCOVA)

- **ANCOVA**: Compares post-test scores with pre-test scores factored out
  - *Concomitant variable* (CV) is the variable that should be controlled for (e.g., pre-test)
  - *Observed Means:* Actual means for the dependent variable (post-test).
  - *Adjusted Means:* Means that have been statistically manipulated based on the concomitant variable (pre-test)
- Report both the Observed and Adjusted Means as well as ANCOVA summary table
  - Present any figures with the Adjusted Means with a note so that readers are clear that these are Adjusted Means.

ANCOVA Example

Compare Treatment (A) to Control (B) on Science Achievement

Dependent Variable		
Sample		
A	B	Total
n		
240	240	480
Observed Means		
1.3304	1.8107	1.5706
Adjusted Means		
1.8663	1.2748	1.5706

Pre-test Scores
Group A: 0.28
Group B: 1.74

Aggregate Correlation within Samples: CV vs DV

r =	0.53	r^2 =	0.28

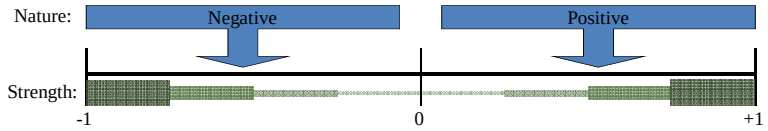
ANCOVA Summary

Source	SS	df	MS	F	P
adjusted means	17.12	1	17.12	33.58	<.0001
adjusted error	243.16	477	0.51		
adjusted total	260.28	478			

Correlation

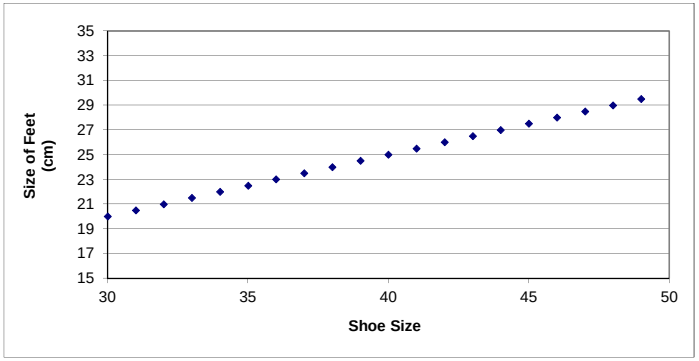
- **Correlation**: Examine the relationship between two variables within the same group of participants
  - Examine the p to determine if the correlation is significant
  - If this is significant, then the next step is to interpret the correlation itself, symbolized by r.
- Report the means and standard deviations for the two variables, the t, df, two-tailed p, and r

Interpreting Correlations

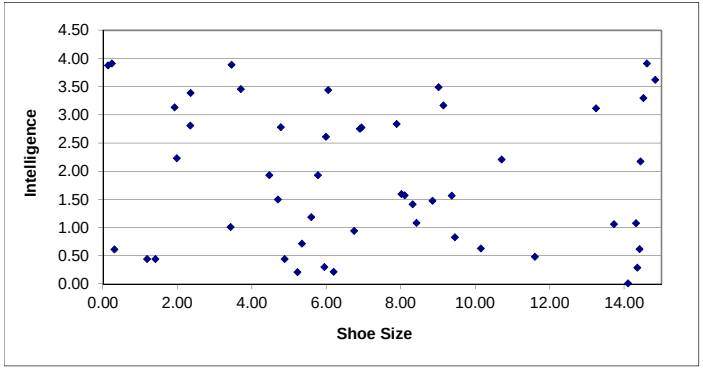


- Nature
  - **Positive**: Two variables increase together
  - **Negative**: As one variable increases, the other decreases
- Strength
  - Closer to -1 or +1 is stronger relationship
  - 0 is no relationship

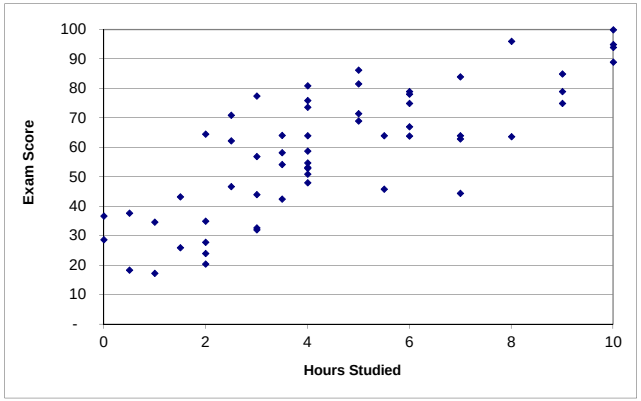
Correlation = 1.00



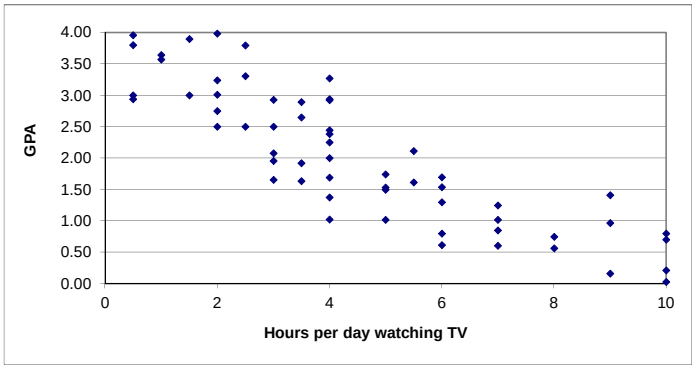
Correlation = .04



Correlation = .78



Correlation = -.86



Correlation Example

Relationship between Science Achievement and Intrinsic Motivation

Data Summary

ΣX =	68.63	ΣX^2 =	28.0123
ΣY =	435.3	ΣY^2 =	991.5978
ΣXY =	132.87		

	X	Y
N	242	
Mean	0.2836	1.7988
Variance	0.0355	0.8655
Std.Dev.	0.1883	0.9303
Std.Err.	0.0121	0.0598

r	r^2	Slope	Y Intercept	Std. Err. of Estimate
0.2231	0.0498	1.101987	1.486276	0.9088

t	df	P		
3.55	240		one-tailed	0.0002315
			two-tailed	0.000463

Correlation Example

Relationship between Science Achievement and English Achievement

Data Summary

ΣX =	156.11	ΣX^2 =	121.9003
ΣY =	5428	ΣY^2 =	128784
ΣXY =	3641.06		

	X	Y
N	233	
Mean	0.67	23.2961
Variance	0.0746	10.0542
Std.Dev.	0.2731	3.1708
Std.Err.	0.0179	0.2077

r	r^2	Slope	Y Intercept	Std. Err. of Estimate
0.0214	0.0005	0.24846	23.129632	3.177

t	df	P		
0.33	231		one-tailed	0.3708495
			two-tailed	0.741699

Correlation Example

Relationship between Science Achievement and Number of Days Absent from Class

Data Summary

$\Sigma X =$	421.25	$\Sigma X^2 =$	902.3475
$\Sigma Y =$	12532	$\Sigma Y^2 =$	668964
$\Sigma XY =$	20773.15		

	X	Y
N	242	
Mean	1.7407	51.7851
Variance	0.7016	82.9578
Std.Dev.	0.8376	9.1081
Std.Err.	0.0538	0.5855

r	r^2	Slope	Y Intercept	Std. Err. of Estimate
-0.5664	0.3208	-6.158946	62.505977	7.522

t	df	P		
-10.65	240		one-tailed	<.0001
			two-tailed	<.0001

Analysing Open-Ended Items

1. Identify **themes** in the responses
   - **Theme:** Concept that is frequently repeated in the data
2. Develop a detailed description of each theme
3. **Code** all responses to find themes
   - **Code:** Tally every statement that fits the theme
4. Calculate the frequency that each theme is mentioned in the items
5. Key interesting statements can be reported verbatim in the results section.

Analysing Open-Ended Items

RQ: What do teachers believe are the disadvantages of beating in the classroom?

Code	Theme	Description
S	Stubbornness	Hardens the student and/or student continues to misbehave
I	Injury	Beating could injure the student
F	Fear	Student begins to fear the teacher, school, and/or the class
NM	Negative Mood	The student is put into a negative mood and/or cannot focus
NL	Not Learn	Student not learn what they are being beaten for
OD	Other Disadvantage	A disadvantage that does not fit in another category

- Disadvantages, when a teacher goes to that extreme of beating a child without finding out what he did, and the child will not accept his fault, he may not know what he has done that is wrong, he will just feel that this person is beating me just like that.
- Good. Disadvantages. Yeah. I think they outweigh the advantage. One, it creates fear. Creates fear. And then where there is fear, there is no learning. The child's attention is on what harm the beating makes. They may even feel that whatever it does, the beating, so the fear is there constantly throughout the class that you are there talking. He is looking at you but he is not listening.
- The disadvantages is that it is not every offense that use cane and out of annoyance if you are not careful, you can hurt the child so it's why I don't like using cane.

Table 3. Teachers Beliefs of the Disadvantages of Beating.

Theme	Description	Percent
Stubbornness	Hardens the student and/or student continues to misbehave	50%
Injury	Beating could injure the student	40%
Fear	Student begins to fear the teacher, school, and/or the class	35%
Negative Mood	The student is put into a negative mood and/or cannot focus	25%
Other Disadvantage	A disadvantage that does not fit in another category	20%
Not Learn	Student not learn what they are being beaten for	10%