Assignment 2-3: t-tests and nonparametric equivalents

David Kent Norman

University of Central Florida

1. We randomly assigned nine children to each of two groups. One group was given training in creative problem solving and the other was not. All children were then given a series of problems and asked to generate possible solutions. The number of solutions generated by each subject was:

Training	12	16	19	8	10	13	9	15	14
No Training	15	5	11	8	9	5	6	11	10

(A) Assume the data is normally distributed.

(1.) State the null

There is no difference in the mean number of problem solutions between the trained group of children and the untrained group.

(2.) What statistic would you use?

Independent t-test

Why?

The data is presumed to be normally distributed interval/ratio and there are two independent groups

(3.) Run the Test

Group Statistics

	group	N	Mean	Std. Deviation	Std. Error Mean
var	Training	9	12.89	3.551	1.184
	No Training	9	8.89	3.296	1.099

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
var	Equal variances assumed	.061	.808	2.477	16	.025	4.000	1.615	.576	7.424
	Equal variances not assumed			2.477	15.912	.025	4.000	1.615	.575	7.425

(4.) Interpret the Results

There is a statistically significant difference ($t=2.477$, $df=16$, $p<0.03$) between the trained ($\underline{M} = 12.89$, $s=3.551$) and untrained ($\underline{M} = 8.89$, $s=3.296$) groups. The 95% confidence interval ranges is 0.58-7.42, with a true mean difference of 4.00.

(5.) What is the role of random assignment in this activity?

Random assignment controls for factors not systematically varied or controlled in the experiment. The groups differ only by chance on all the possible variables of difference between the groups.

(B) Assume the data is not normally distributed.

(1.) State the null

There is no difference in the rank number of problem solutions between the trained group of children and the untrained group.

(2.) What statistic would you use?

Mann-Whitney

Why?

The data is presumed to be not normally distributed and there are two independent, ranked groups.

(3.) Run the Test

Ranks

	group	N	Mean Rank	Sum of Ranks
var	Training	9	12.11	109.00
	No Training	9	6.89	62.00
	Total	18		

Test Statistics[b]

	var
Mann-Whitney U	17.000
Wilcoxon W	62.000
Z	-2.082
Asymp. Sig. (2-tailed)	.037
Exact Sig. [2*(1-tailed Sig.)]	.040[a]

a. Not corrected for ties.

b. Grouping Variable: group

(4.) Interpret the Results

There is a statistically significant difference in number of creative solutions ranking ($z=-2.082$, $p<0.05$) between the trained ($\underline{M}_{rank} = 12.11$) and untrained ($\underline{M}_{rank} = 6.89$) groups.

2. It has been argued that first-born children tend to be more independent than later-born children. Suppose we develop a 25-point scale of independence and rate each of 20 first-born children and their second-born siblings using our scale. We do this when both siblings are adults, thus eliminating obvious age effects. The data on independence are as follows (a higher score means that the person is more independent):

Pair	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**

First-Born:	12	18	13	17	8	15	16	5	8	12
Second-Born:	10	12	15	13	9	12	13	8	10	8

Pair (Continued)	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
First-Born:	13	5	14	20	19	17	2	5	15	18
Second-Born	8	9	8	10	14	11	7	7	13	12

(A) Assume the data is **not** normally distributed

(1.) State the null

There is no difference in independence rating between first and second born children on the independence scale.

(2.) What statistic would you use?

Wilcoxon

Why?

The data is presumed to be not normally distributed (ranked) and the data is paired or related groups.

(3.) Run the Test

Ranks

		N	Mean Rank	Sum of Ranks
Second Born - First Born	Negative Ranks	13[a]	12.62	164.00
	Positive Ranks	7[b]	6.57	46.00
	Ties	0[c]		
	Total	20		

a. Second Born < First Born

b. Second Born > First Born

c. Second Born = First Born

Test Statistics[b]

	Second Born - First Born
Z	-2.211[a]
Asymp. Sig. (2-tailed)	.027

a. Based on positive ranks.

b. Wilcoxon Signed Ranks Test

(4.) Interpret the Results

There is a statistically significant difference ($z=-2.211$, $p<0.03$) in ranking between the first born and second born groups. First born adults had higher independence scores than second 13 of 20 times. There were only 7 second born with independence scores higher than first born. There were no ties in ranking.

(B) Assume the data is normally distributed

(1.) State the null

There is no statistically significant mean difference in independence score between first and second born children on the independence scale.

(2.) What statistic would you use?

Dependent t-test

Why?

The data is presumed to be normally distributed and the data is in two related pairs or groups

(3.) Run the Test

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	First Born	12.60	20	5.365	1.200
	Second Born	10.45	20	2.438	.545

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	First Born & Second Born	20	.678	.001

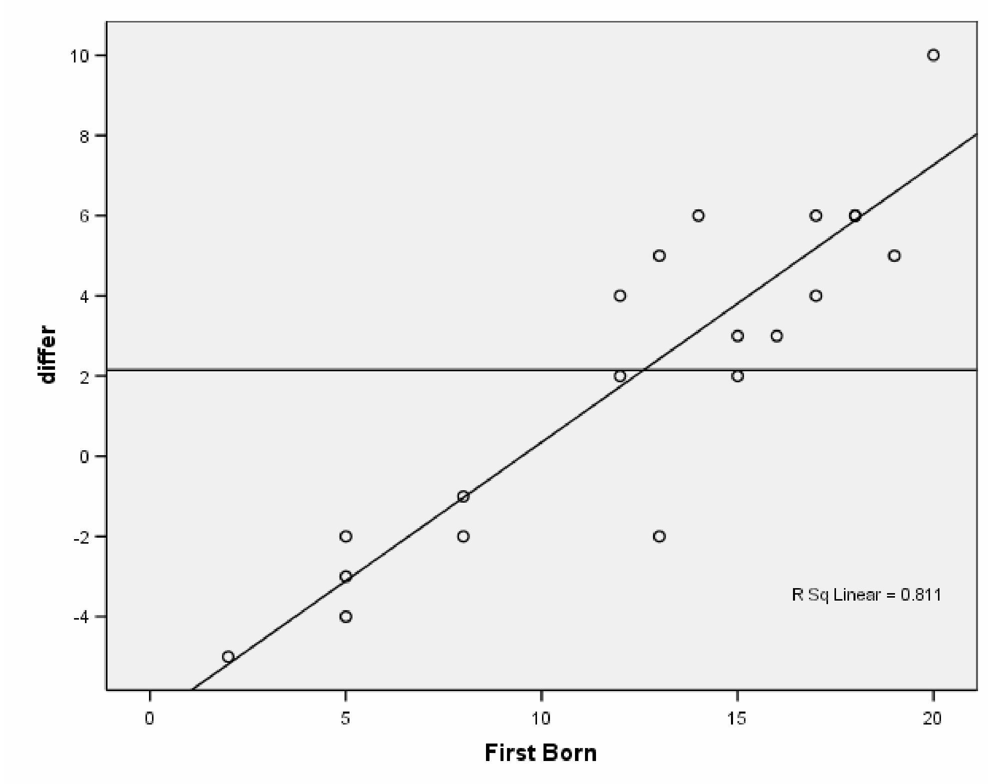
Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper			
Pair 1	First Born - Second Born	2.150	4.120	.921	.222	4.078	2.334	19	.031

(4.) Interpret the Results

There is a statistically significant mean difference (t=2.334, df=19, p<0.04) between the first born (M = 12.6, s=5.365) and second born (M = 10.45, s=2.438) groups. The 95% confidence interval suggests the true mean difference is 0.22-4.08.

(C) The results in this Exercise are not quite as clear-cut as we might like. Plot the differences as a function of the first-born's score. What does this figure suggest?



There is a linear relationship to suggest as the independence of one in the pair increases, the other decreases. Eighty-one percent of the variance in the difference was explained by the first-born's score, therefore there is strong evidence to suggest if the first-born shows higher independence on the independence scale, the second-born will have lower independence on the independence scale.

Use a random sample of 500 from the NELS88 data for questions 3-10.

For each analysis:

State the null

What procedure should be used Why

Run the procedure

Interpret the results

3. Is there a difference in reading standardized test scores between males and females?

Null: There is no difference in reading standardized test scores between males and females.

Procedure: Use independent sample t-test because the data is assumed to be normally distributed with two independent groups and an interval/ratio dependent variable

Run

Group Statistics

	COMPOSITE SEX	N	Mean	Std. Deviation	Std. Error Mean
READING STANDARDIZED SCORE	MALE	237	48.7271	10.18662	.66169
	FEMALE	212	50.5628	9.52039	.65386

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
READING STANDARDIZED SCORE	Equal variances assumed	2.724	.100	-1.966	447	.050	-1.83565	.93377	-3.67077	-.00053
	Equal variances not assumed			-1.973	446.133	.049	-1.83565	.93025	-3.66388	-.00743

There is a statistically significant mean difference (equal variances t=-1.966, df=447, p=0.05) between the male (M = 48.7271, s=10.18662) and female (M = 212, s=9.52039) reading standardized test scores. The 95% confidence interval of the difference between means is between -3.67 and 0.00.

4. Can you successfully predict reading standardized test score by gender?

Null: There is no relationship between math score and gender.

Procedure: bivariate regression because we are trying to predict; gender is nominal

Run

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	COMPOSITE SEX[a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: READING STANDARDIZED SCORE

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.093[a]	.009	.006	9.87774

a. Predictors: (Constant), COMPOSITE SEX

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	377.067	1	377.067	3.865	.050[a]
	Residual	43613.644	447	97.570		
	Total	43990.711	448			

a. Predictors: (Constant), COMPOSITE SEX

b. Dependent Variable: READING STANDARDIZED SCORE

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	46.891	1.452		32.305	.000
	COMPOSITE SEX	1.836	.934	.093	1.966	.050

a. Dependent Variable: READING STANDARDIZED SCORE

There is a relationship between reading test scores (R=.093, $F_{1,447}$=3.865, p<0.01) and gender. Sex accounts for less than 1% of the variance in reading score. The equation for predicting reading score from sex is:

Reading' = 46.891 + 1.836 (comp sex)

Compare your answers to questions 3 and 4.

The t associated with the regression for problem 4 is 1.97 and the t for problem 3 is the same. Squaring the equal variances t value in question 3 (-1.966) is equal to the F value in question 4 (3.865).

5. Do the respondents average history scores differ from the Bozo state mean of 61?

Null: The respondents' average history scores do not differ from the state mean of 61.

Procedure: one sample t-test because there is one group tested against a hypothesized mean

Run:

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
HISTORY/CIT/GEOG STANDARDIZED SCORE	446	49.5910	10.10117	.47830

One-Sample Test

	Test Value = 61					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
HISTORY/CIT/GEOG STANDARDIZED SCORE	-23.853	445	.000	-11.40904	-12.3491	-10.4690

There is a statistically significant difference ($t=-23.853$, $df=445$, $p<.001$) in means between the average history scores ($\underline{M} = 49.5910$, $s=10.10117$) and the hypothesized Bozo state mean of 61.

6. How much time do respondents spend watching tv on the weekend?

Null: there no null with descriptives

Procedure: descriptives because the question asks for a simple overview of the ordinal sample

Statistics

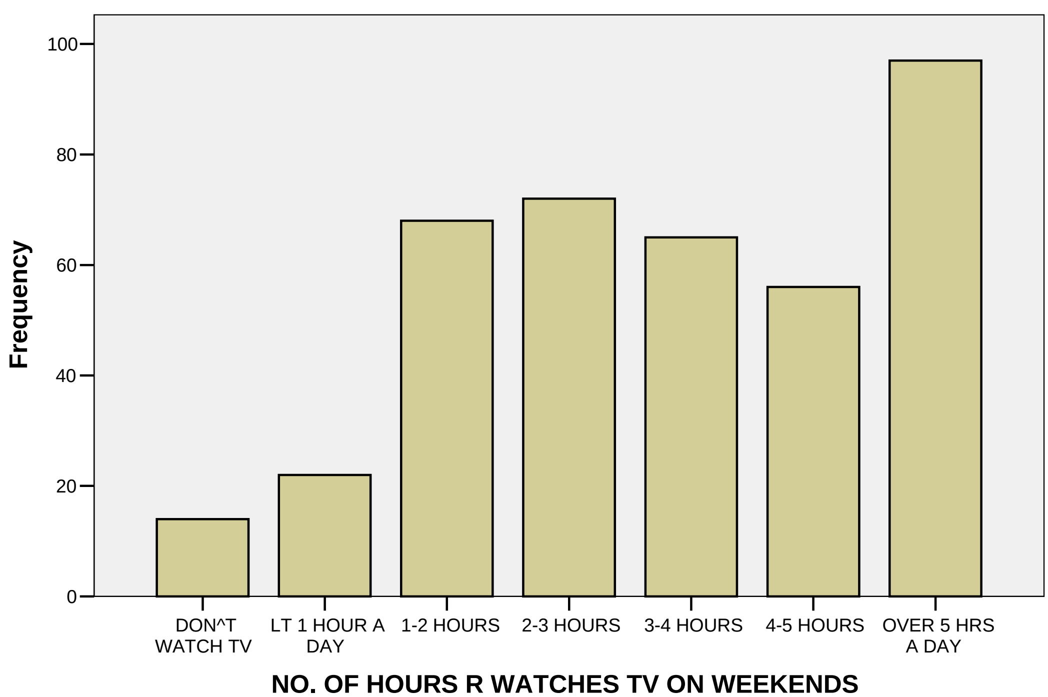
NO. OF HOURS R WATCHES TV ON WEEKENDS

N	Valid	394
	Missing	106
Median		4.00
Minimum		0
Maximum		6

NO. OF HOURS R WATCHES TV ON WEEKENDS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	DON^T WATCH TV	14	2.8	3.6	3.6
	LT 1 HOUR A DAY	22	4.4	5.6	9.1
	1-2 HOURS	68	13.6	17.3	26.4
	2-3 HOURS	72	14.4	18.3	44.7
	3-4 HOURS	65	13.0	16.5	61.2
	4-5 HOURS	56	11.2	14.2	75.4
	OVER 5 HRS A DAY	97	19.4	24.6	100.0
	Total	394	78.8	100.0	
Missing	{MULTIPLE RESPNSE}	47	9.4		
	{MISSING}	27	5.4		
	System	32	6.4		
	Total	106	21.2		
Total		500	100.0		

NO. OF HOURS R WATCHES TV ON WEEKENDS



The median number of hours the respondents watch TV on weekends is 3-4 hours. The most frequent response was "over 5 hours a day" (n = 97).

7. Do respondents who report mostly As in Math have better scores on the math standardized test than those who report mostly Cs?

Null: Respondents who report mostly As in Math do not differ in scores on the math standardized test compared to those who report mostly Cs.

Procedure: The independent t-test is best because the dependent variable is interval/ratio and the respondents have either an A or C.

Run:

Group Statistics

	MATH GRADES FROM GRADE 6 UNTIL NOW	N	Mean	Std. Deviation	Std. Error Mean
MATHEMATICS STANDARDIZED SCORE	MOSTLY AS	166	54.8166	10.77132	.83602
	MOSTLY CS	81	45.3904	7.22691	.80299

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
MATHEMATICS STANDARDIZED SCORE	Equal variances assumed	23.835	.000	7.128	245	.000	9.42626	1.32236	6.82161	12.03090
	Equal variances not assumed			8.132	221.338	.000	9.42626	1.15919	7.14180	11.71071

There is a statistically significant difference (unequal variances t=8.132, df=221.338, p<.001) between students with mostly As (M = 54.8166, s=10.77132) and mostly Cs (M = 45.3904, s=7.22691). The 95% confidence interval is 7.14-11.71, with a mean difference of 9.426.

8. Is there a relationship between weekday tv time and Math grades?

Null: There is no statistically significant relationship between weekday tv time and Math grades.

Procedure: Spearman's because the dependent variable is ranked data. Relationship is the keyword.

Run:

Correlations

			NO. OF HOURS R WATCHES TV ON WEEKDAYS	MATH GRADES FROM GRADE 6 UNTIL NOW
Spearman's rho	NO. OF HOURS R WATCHES TV ON WEEKDAYS	Correlation Coefficient	1.000	.064
		Sig. (2-tailed)	.	.202
		N	414	404
	MATH GRADES FROM GRADE 6 UNTIL NOW	Correlation Coefficient	.064	1.000
		Sig. (2-tailed)	.202	.
		N	404	456

Result: There is no statistically significant relationship ($r = 0.064$, $p>.05$) between weekday TV time and Math grades.

9. Is there a difference between English and Math grades for the respondents?

Null: There is no difference between English and Math grades for the respondents.

Procedure: Wilcoxon because the dependent data is ranked and it is related groups or pairs.

Run:

Ranks

		N	Mean Rank	Sum of Ranks
MATH GRADES FROM GRADE 6 UNTIL NOW - ENGLISH GRADES FROM GRADE 6 UNTIL NOW	Negative Ranks	141[a]	125.54	17701.00
	Positive Ranks	115[b]	132.13	15195.00
	Ties	192[c]		
	Total	448		

a. MATH GRADES FROM GRADE 6 UNTIL NOW < ENGLISH GRADES FROM GRADE 6 UNTIL NOW

b. MATH GRADES FROM GRADE 6 UNTIL NOW > ENGLISH GRADES FROM GRADE 6 UNTIL NOW

c. MATH GRADES FROM GRADE 6 UNTIL NOW = ENGLISH GRADES FROM GRADE 6 UNTIL NOW

Test Statistics[b]

	MATH GRADES FROM GRADE 6 UNTIL NOW - ENGLISH GRADES FROM GRADE 6 UNTIL NOW
Z	-1.124[a]
Asymp. Sig. (2-tailed)	.261

a. Based on positive ranks.

b. Wilcoxon Signed Ranks Test

There is no statistically significant difference in ranking ($z=-1.124$, $p>0.05$) between English and Math grades for the respondents. The ranks of the data showed 141 respondents ranked higher in English than math, 115 ranked higher in Math than English, and 192 tied in rank.

10. Does a respondent's self concept influence their math standardized test score?

Null: Self-concept does not influence math scores.

Procedure: regression because we want to find the relationship between 2 or more interval/ratio variables. Influence is a strong word to define it from using Pearson's instead.

Run:

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.103[a]	.011	.008	10.46710

a. Predictors: (Constant), SELF CONCEPT 1

b. Dependent Variable: MATHEMATICS STANDARDIZED SCORE

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	526.245	1	526.245	4.803	.029[a]
	Residual	48644.679	444	109.560		
	Total	49170.925	445			

a. Predictors: (Constant), SELF CONCEPT 1

b. Dependent Variable: MATHEMATICS STANDARDIZED SCORE

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	50.324	.496		101.529	.000	49.350	51.298
	SELF CONCEPT 1	1.374	.627	.103	2.192	.029	.142	2.607

a. Dependent Variable: MATHEMATICS STANDARDIZED SCORE

There is a statistically significant difference ($R=0.103$, $F_{1,444}=4.803$, $p<.03$) between self-concept and math standardized score. Only 1.1% of the variance in math can be explained by self-concept. As self-concept increases by 1, math score increases by 1.374:

Math' = 50.324 + 1.374(self-concept)

11. When do you use the following? What do the results tell you?

One sample t-test – when you want to test against a hypothesized mean value to see if there is a statistically significant difference from the hypothesized mean in some interval/ratio data. There is only one group to compare to the hypothesized mean. The results show how many standard error units, represented by the t value, the sample mean is from the hypothesized mean. The p value shows how likely the difference is likely to occur by chance.

Independent t-test – when you want look for a difference between two independent groups of data to say if they are statistically different or not in some normally distributed interval/ratio data. The results show how many standard error units, represented by t, the mean of one group is from the other group. The p value shows how likely the difference in means is to occur by chance.

Dependent t-test – when you want to look for a difference between two related groups to say if they are statistically different or not in normally distributed interval/ratio data. Paired or related data could be defined as a pair of pre-test and post-test scores. The results show how many standard error units the first mean (pretest) is from the second mean (posttest). The test accounts for the correlation between the two groups and calculates a p value to say how likely the difference between the two groups is to occur by chance.

Pearson's correlation – discover the relationship between two variables, where the variables are normally distributed interval/ratio data. The results show the statistical significance of the relationship of the variables through the values r and r^2. The value of r is a measure of magnitude of the relationship in ranked form.

Regression - discover the relationship between two or more variables, where the dependent variable is normally distributed interval/ratio data. The results provide information to create a least squares regression like that shows the relationship between the dependent variable and the independent variable. The regression line formula allows the researcher to predict values for the independent variable as, "for each increase in the dependent variable, the independent variable should increase/decrease by X." A regression test also provides a means of statistical significance between the variables.

Mann Whitney U – determine differences in ranking in data that is not normally distributed and that is in independent groups to find the difference between two groups of ordinal data. The results show a z statistic with a significance value to predict the likelihood that the change in the ranking represented by z occurred by chance. It is the nonparametric equivalent to the independent t-test.

Wilcoxon – use with data that is not normally distributed and that is in dependent groups or pairs to find the difference between two groups of ordinal data. It is the nonparametric equivalent to the dependent t-test. Like the Mann Whitney, the results produce a z statistic and a probability level, along with a "scoreboard" of how many times each variable was ranked above, below, or tied with the other.