



Overview of Non-Parametric Statistics

PRESENTER: ELAINE EISENBEISZ
OWNER AND PRINCIPAL, OMEGA STATISTICS

About Omega Statistics



Private practice consultancy based in Southern California,



Medical and Clinical Research: Epidemiology data, clinical trials phases I through IV, design and analysis of experiments, journal article assistance



Business Statistics: Econometrics, predictive modeling



Dissertation Assistance: Methods and Results Chapters



Survey Design



Projects of any size. Individual researchers, small business and start-ups, large corporations, government contracts.



Your one stop shop for all things measurable

Parametric Tests for Normal Data

Parametric tests such as t-tests, ANOVA, linear regressions...require assumptions to be met in order to make accurate estimates and inferences.

- ❖ Independence
- ❖ Normal Distribution
- ❖ Homogenous Variances Between Groups
- ❖ Continuous (ratio/interval) Dependent Variable

Non-Normal Data is NOT Abnormal Data

Nonparametric techniques do not have the stringent assumptions of parametric techniques, such as a need for a normal distribution.

This is why nonparametric techniques are also referred to as “distribution-free” tests.

Nonparametric techniques are useful when data are measured on a nominal (categorical) or ordinal (ranked) scale. They are also useful when sample sizes are small.

Distribution-Free Tests - General Assumptions

❖ Random Samples

❖ Independence

Note.

- Many studies and fields of research do not make use of random sampling and this assumption is often relaxed.
- Independence is not assumed for repeated measures tests.

Alternative for Independent Samples t-test: The Mann-Whitney U Test

Also called the Mann-Whitney Test, the Wilcoxon Rank Sum Test, or Wilcoxon Test.

Tests if two independent populations have the same center. Uses medians of ranked data instead of means and standard deviations.

When to use:

- ❖ Responses can only be ranked (data is at least ordinal)
- ❖ Variances between two groups are not equal
- ❖ Normality assumption is violated

Example: Mann-Whitney U Test

A study was conducted to test diastolic blood pressure readings between patients with sedentary lifestyles (Group A) and those who exercised 60 minutes at least 2 times per week (Group B).

It was not clear whether the assumptions for a t-test were valid, so the researcher employed nonparametric methods

Group A	85	89	86	91	77	93	100	82	92	86
Group B	83	73	65	65	90	77	78	97	85	75

data_t-test_mwu_paired.sav [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

Reports
Descriptive Statistics
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
Survival
Multiple Response
Quality Control
ROC Curve...

One Sample...
Independent Samples...
Related Samples...
Legacy Dialogs
Chi-square...
Binomial...
Runs...
1-Sample K-S...
2 Independent Samples...
K Independent Samples...
2 Related Samples...
K Related Samples...

Visible: 3 of 3 Variables

	id	diastolic																
1	1.00	85.0																
2	2.00	89.0																
3	3.00	86.0																
4	4.00	91.0																
5	5.00	77.0																
6	6.00	93.0																
7	7.00	100.0																
8	8.00	82.0																
9	9.00	92.0																
10	10.00	86.0																
11	11.00	83.0																
12	12.00	73.0																
13	13.00	65.0																
14	14.00	65.0																
15	15.00	90.0	2.00															
16	16.00	77.00	2.00															
17	17.00	78.00	2.00															
18	18.00	97.00	2.00															
19	19.00	85.00	2.00															
20	20.00	75.00	2.00															
21																		
22																		
23																		
24																		
25																		
26																		
27																		
28																		
29																		
30																		
31																		

Data View Variable View

2 Independent Samples...

IBM SPSS Statistics Processor is ready

9:07 AM
10/1/2012

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
diastolic	20	83.4500	9.58878	65.00	100.00
exercise	20	1.5000	.51299	1.00	2.00

Ranks

	exercise	N	Mean Rank	Sum of Ranks
diastolic	1.00	10	13.40	134.00
	2.00	10	7.60	76.00
	Total	20		

Test Statistics^b

	diastolic
Mann-Whitney U	21.000
Wilcoxon W	76.000
Z	-2.195
Asymp. Sig. (2-tailed)	.028
Exact Sig. [2*(1-tailed Sig.)]	.029 ^a

a. Not corrected for ties.

b. Grouping Variable: exercise

The sedentary group had a significantly higher mean ranked diastolic blood pressure measurement ($p = .029$)

Alternative for One-Way Analysis of Variance (ANOVA)

The Kruskal-Wallis Test

The Kruskal-Wallis Test is set-up in the same manner as the Mann-Whitney U test (making use of ranks) but there are more than 2 independent groups to be tested.

Example:

The effect of supermarket shelf height was tested on a specific brand of ice cream. Three levels of shelf height were tested, (a) knee level, (b) waist level, and (c) eye level.

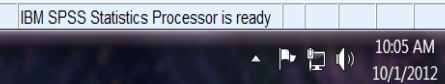
The shelf height of the ice cream was randomly changed three times per day for each of eight days.

Kruskal-Wallis Test (cont'd)

The total sales in hundreds of dollars was recorded for the 8-day period:

Knee	77	82	86	78	81	86	77	81
Waist	88	94	93	90	91	94	90	87
Eye	85	85	87	81	80	79	87	93

The same process is used in SPSS, except we now choose:



Ranks

	level	N	Mean Rank
sales	1.00	8	6.75
	2.00	8	19.69
	3.00	8	11.06
	Total	24	

Test Statistics^{a, b}

	sales
Chi-Square	13.971
df	2
Asymp. Sig.	.001

a. Kruskal Wallis
Test

b. Grouping
Variable: level

The results are statistically significant at the $\alpha = .05$ level, and mean rank of waist level is the highest, indicating more sales of the ice cream with the waist high placement.

Nonparametric Alternative for Paired t-test

The Wilcoxon Signed Rank Test

Subjects are measured at 2 different times, or under 2 different conditions, or are matched on a specific criteria. Also referred to as the Wilcoxon matched pairs signed ranks test.

Example:

A shoe manufacturer wants to compare the wear of a new sole to the shoe sole currently in use. A random sample of 10 cross country runners was chosen and the old and new soles were placed on each runner's shoes in random order (some had the new sole on the left foot, others the right foot)

Wilcoxon Signed Rank Test (cont'd)

Runner	Old Sole	New Sole
1	45.5	47.0
2	50.0	51.0
3	43.0	42.0
4	45.5	46.0
5	58.5	58.0
6	49.0	50.5
7	29.5	39.0
8	42.0	53.0
9	48.0	48.0
10	57.5	61.0

*data_wilcoxon.sav [DataSet0] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

12 :

	runner	old
1	1.00	45.
2	2.00	50.
3	3.00	43.
4	4.00	45.
5	5.00	58.
6	6.00	79.
7	7.00	29.
8	8.00	42.
9	9.00	48.
10	10.00	57.
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		
21		
22		
23		
24		
25		
26		
27		
28		
29		
30		
31		

Visible: 3 of 3 Variables

Reports
Descriptive Statistics
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
Survival
Multiple Response
Quality Control
ROC Curve...

One Sample...
Independent Samples...
Related Samples...
Legacy Dialogs
Chi-square...
Binomial...
Runs...
1-Sample K-S...
2 Independent Samples...
K Independent Samples...
2 Related Samples...
K Related Samples...

Data View Variable View

2 Related Samples...

IBM SPSS Statistics Processor is ready

10:56 AM
10/1/2012

Ranks

	N	Mean Rank	Sum of Ranks
new - old Negative Ranks	3 ^a	4.67	14.00
Positive Ranks	6 ^b	5.17	31.00
Ties	1 ^c		
Total	10		

a. new < old

b. new > old

c. new = old

Test Statistics^b

	new - old
Z	-1.009 ^a
Asymp. Sig. (2-tailed)	.313

a. Based on negative ranks.

b. Wilcoxon Signed Ranks
Test

Results are not statistically significant at the
 $\alpha = .05$ level.

Nonparametric Alternative to the One-Way Repeated Measures ANOVA The Freidman Test

Use the Freidman test when you measure the same subjects at 3 or more time points, or under 3 conditions.

Example:

A farmer tested the effects of 3 different fertilizers on his avocado yield. She chose 5 locations in her field and divided each of these five locations into 3 plots. She then randomly assigned one of the three fertilizers to one of the 3 plots in each location such that all 3 fertilizers were represented one time at each location. After one year she measured the yield of avocados for each plot.

Friedman Test (cont'd)

Here is the data:

Location	Fertilizer A	Fertilizer B	Fertilizer C
1	11	14	10
2	13	17	15
3	16	20	13
4	10	12	10
5	12	16	12



*friedman.sav [DataSet0] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

7 : fertilizerC

	location	fertilizerA
1	1.00	11.0
2	2.00	13.0
3	3.00	16.0
4	4.00	10.0
5	5.00	12.0
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		
21		
22		
23		
24		
25		
26		
27		
28		
29		
30		
31		

Visible: 4 of 4 Variables

var var var var var var var var var var var var var var var var

Reports
Descriptive Statistics
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
Survival
Multiple Response
Quality Control
ROC Curve...

One Sample...
Independent Samples...
Related Samples...
Legacy Dialogs
Chi-square...
Binomial...
Runs...
1-Sample K-S...
2 Independent Samples...
K Independent Samples...
2 Related Samples...
K Related Samples...

Data View Variable View

K Related Samples...

IBM SPSS Statistics Processor is ready

11:17 AM
10/1/2012

Friedman Test (cont'd)

Ranks

	Mean Rank
fertilizerA	1.60
fertilizerB	3.00
fertilizerC	1.40

Test Statistics^a

N	5
Chi-Square	8.444
df	2
Asymp. Sig.	.015

a. Friedman Test

We have statistical significance at the $\alpha = .05$ level, and it appears that fertilizer B is the best.

Nonparametric Alternative to Pearson's Product Moment Correlation: Spearman's Rank Order Correlation

Pearson's correlation can be used with continuous or dichotomous data. Spearman's Rank Order correlation can be used when the data is at least **ordinal**.

As with the other techniques presented, Spearman's correlation coefficient is determined by ranking data.

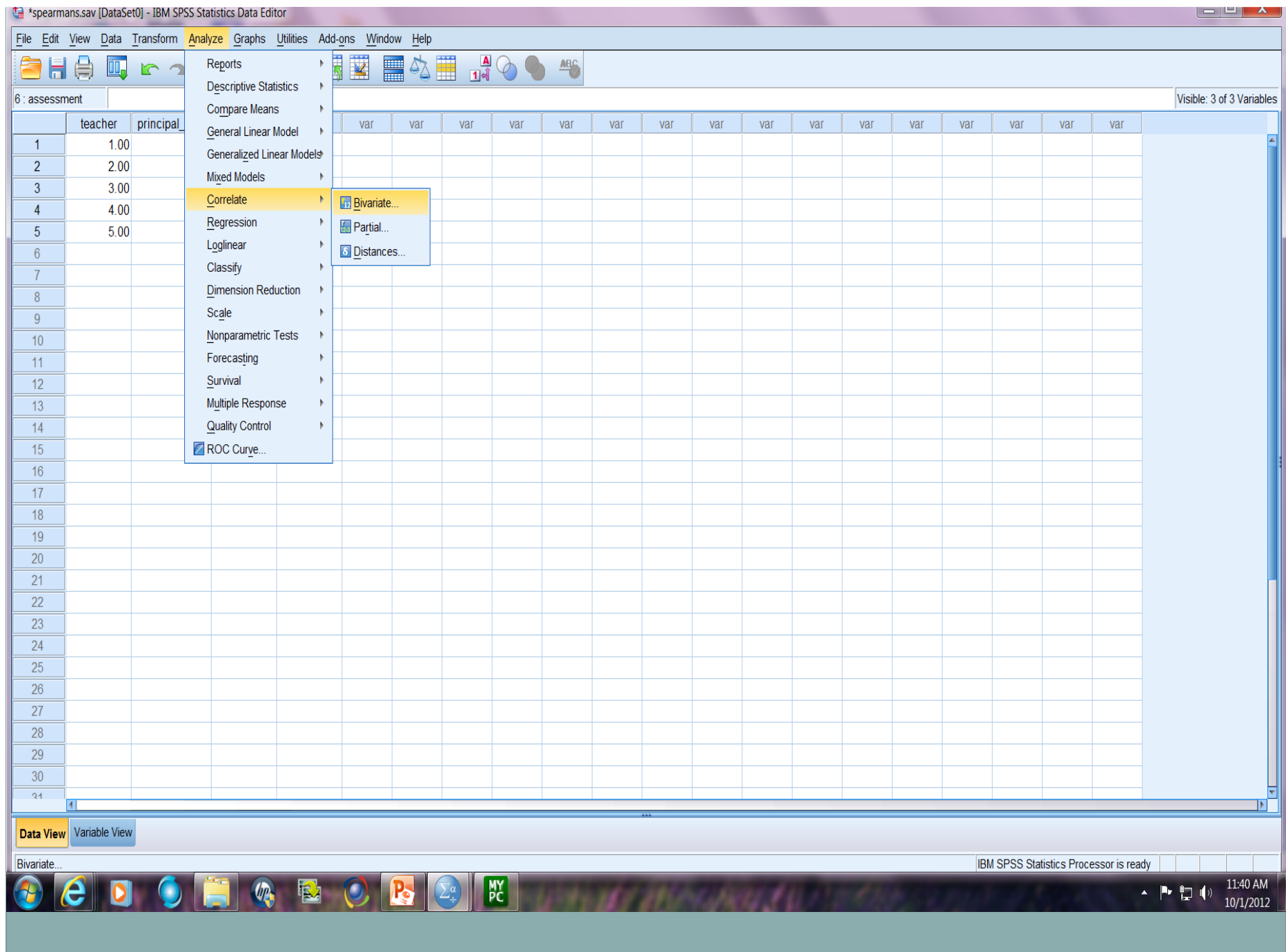
Spearman's Rank Order Correlation (Cont'd)

Example:

Five high school mathematics teachers were ranked by their principal according to their teaching ability. The teachers also took an exam that assessed their knowledge of the required mathematics subject matter. Is there agreement between the principal's ranking and the assessment scores?

Spearman's Rank Order Correlation (cont'd)

Teacher	Principal's Rank	Assessment Score
1	4	72
2	2	69
3	3	82
4	1	93
5	5	80



Spearman's Rank Order Correlation (cont'd)

Correlations

			principal_rank	assessment
Spearman's rho	principal_rank	Correlation Coefficient	1.000	-.300
		Sig. (2-tailed)	.	.624
		N	5	5
	assessment	Correlation Coefficient	-.300	1.000
		Sig. (2-tailed)	.624	.
		N	5	5

Not significant at the $\alpha = .05$ level ($p = .624$)

So what do I do when my data is Nominal?

The Chi-Square Test of Independence

Use the chi-square test of independence when you want to explore the relationship between 2 categorical variables.

Assumptions for the chi-square test of independence are:

- ❖ Independence

- ❖ Each in the contingency table must have an *expected value* of 5 or more. Some statistician relax this assumption a bit and require that at least 80% of the cells have an expected value of 5 or more.

Chi-Square Test of Independence (cont'd)

Example: From SPSS Survival Manual, (Pallant ,2007)

Is there a relationship between gender and smoking behavior?

	Smoker?		
Gender	Yes	No	Total
Male	18	107	184
Female	26	74	252
Total	85	351	436

chisquare.sav [DataSet0] - IBM SPSS Statistics Data Editor

FileEditViewDataTransformAnalyzeGraphsUtilitiesAdd-onsWindowHelp

153 : smoker

	gender	smoker
1	1.00	1.00
2	1.00	1.00
3	1.00	1.00
4	1.00	1.00
5	1.00	1.00
6	1.00	1.00
7	1.00	1.00
8	1.00	1.00
9	1.00	1.00
10	1.00	1.00
11	1.00	1.00
12	1.00	1.00
13	1.00	1.00
14	1.00	1.00
15	1.00	1.00
16	1.00	1.00
17	1.00	1.00
18	1.00	1.00
19	1.00	.00
20	1.00	.00
21	1.00	.00
22	1.00	.00
23	1.00	.00
24	1.00	.00
25	1.00	.00
26	1.00	.00
27	1.00	.00
28	1.00	.00
29	1.00	.00
30	1.00	.00
31	1.00	.00

Visible: 2 of 2 Variables

Reports

Descriptive Statistics

Compare Means

General Linear Model

Generalized Linear Models

Mixed Models

Correlate

Regression

Loglinear

Classify

Dimension Reduction

Scale

Nonparametric Tests

Forecasting

Survival

Multiple Response

Quality Control

ROC Curve...

Frequencies...

Descriptives...

Explore...

Crosstabs...

Ratio...

P-P Plots...

Q-Q Plots...

Data ViewVariable View

Crosstabs...

IBM SPSS Statistics Processor is ready

12:05 PM10/1/2012

gender * smoker Crosstabulation

			smoker		Total
			.00	1.00	
gender	1.00	Count	107	18	125
		Expected Count	100.6	24.4	125.0
		% within gender	85.6%	14.4%	100.0%
		% within smoker	59.1%	40.9%	55.6%
	2.00	Count	74	26	100
		Expected Count	80.4	19.6	100.0
		% within gender	74.0%	26.0%	100.0%
		% within smoker	40.9%	59.1%	44.4%
Total		Count	181	44	225
		Expected Count	181.0	44.0	225.0
		% within gender	80.4%	19.6%	100.0%
		% within smoker	100.0%	100.0%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	4.752 ^a	1	.029	.042	.022
Continuity Correction ^b	4.043	1	.044		
Likelihood Ratio	4.730	1	.030		
Fisher's Exact Test					
Linear-by-Linear Association	4.731	1	.030		
N of Valid Cases	225				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 19.56.

b. Computed only for a 2x2 table

Some Final Notes

Q. Why don't we always use nonparametric techniques?

A. Because the parametric tests are more powerful, meaning you will have a better probability of seeing significance that is truly in your data

Parametric tests are very robust to deviations from the normality assumption and the presence of outliers if you have equal variance between your independent groups.

Final Notes (cont'd)

I've covered only a small amount of nonparametric tests which were based almost entirely on the use of ranking data. I suggest you research more into the subject and you will discover many other types of tests that can be used.

A Useful Reference:

UCLA's Academic Technology Department (ATS) has a great site with annotated output of tests for many statistical programs such as SPSS, R, and SAS.

<http://www.ats.ucla.edu/stat/>

Thank You for Attending!

Elaine Eisenbeisz, Owner and Principal
Omega Statistics

41690 Ivy Street, Suite D6
Murrieta, CA 92562

Phone: 951-461-7226

Fax: 951-461-7339

Toll Free: 877-461-7226

Email: elaine@OmegaStatistics.com

Website: www.OmegaStatistics.com

THE ANALYSIS
FACTOR

This presentation was part of **The Craft of Statistical Analysis Webinar Series**, a free program of The Analysis Factor.

You can request a free copy of the webinar recording at:
<http://www.theanalysisfactor.com/webinars/recordings>