Basic and Advanced Statistical Tests

Writing Results Sections and Creating Tables and Figures

Amanda Ross and Victor L. Willson

SensePublishers

Basic and Advanced Statistical Tests

Basic and Advanced Statistical Tests

Writing Results Sections and Creating Tables and Figures

Amanda Ross
A. A. Ross Consulting and Research, USA

and

Victor L. Willson
Texas A&M University, USA

SENSE PUBLISHERS
ROTTERDAM/BOSTON/TAIPEI

A C.I.P. record for this book is available from the Library of Congress.

ISBN: 978-94-6351-084-4 (paperback)
ISBN: 978-94-6351-085-1 (hardback)
ISBN: 978-94-6351-086-8 (e-book)

Published by: Sense Publishers,
P.O. Box 21858,
3001 AW Rotterdam,
The Netherlands
https://www.sensepublishers.com/

Printed on acid-free paper

TABLE OF CONTENTS

INTRODUCTION

Many researchers have difficulty knowing how to properly write a results section for a scholarly work. They may comb through countless journals, trying to find tables and/or figures that match the statistical test used in their design. Statistics courses do a fine job of teaching graduate students how to run various basic and advanced tests. However, students are often left to their own devices to determine how to write the results section and create the appropriate tables and/or figures in APA format.

This book presents a concise look at basic and advanced statistical tests, providing a brief description of each and examples of research problems that yield themselves to each particular test. Revealing the core strength of each chapter, a sample scenario and results section write-up are provided. Depending upon the test and need, a sample table and/or figure may be provided.

This book serves as a reference manual for all professionals in the psychology, sociology and education fields. Actually, this book can assist professional researchers in any field, including medical research. By providing specific examples of situations aligned with each type of test, the reader can become adept at understanding what different variables (nominal, ordinal, interval, and ratio) look like in real-world problem situations.

While other books provide an in-depth examination of each test, how to run each test, and ways to interpret results, this book provides the missing piece in the literature out there. Graduate students, teachers, faculty, social workers, psychologists, and other professional researchers now have access to a guide that includes the name of the test, a brief description (without all of the technical jargon), examples, a sample scenario, a sample results section write-up, and accompanying tables and/or figures! When having difficulty knowing which information to extract from the output, simply turn to each results section. If you aren't sure which information to place in the write-up and which to place in a table, simply look at the examples!

We do hope this book/reference manual is helpful to you. As a graduate student researcher, this sort of book would have helped me immensely throughout my research endeavors. It provides us great pleasure to share this reference with you. Good luck with all your current and future research studies!

For many of our analyses, we used the BASC data set, *Results of Dataset analysis from the Behavior Assessment System for Children, Second Edition (BASC-2).*

PART I

BASIC STATISTICAL TESTS

CHAPTER 1

DESCRIPTIVE STATISTICS

BRIEF DESCRIPTION

The field of statistics can be divided into the categories of descriptive statistics and inferential statistics. Descriptive statistics are used to characterize a sample and include computations that do not require any inference about a population, such as frequencies, percentages, median, mean, mode, standard deviation, *z*-scores, variance, range, and interquartile range. Such statistics prove very useful when sample sizes are not large enough to warrant other tests.

Frequencies and percentages are low-level quantitative approaches that simply report either a count or a ratio. They can be easily represented, using bar graphs or histograms. Scores are often grouped, such as into fourths, termed quartiles, or tenths, termed deciles.

Measures of the center of a distribution of data include the median, mean, and mode. The median represents the middle value (or average of the two central values) of the data, when data are arranged in ascending order. The median is not affected by extreme scores and should be used as the measure of center when data includes extreme outliers or is not approximately normally distributed. The mean is equal to the ratio of the sum of all scores to the total number of scores. The mean is affected by extreme scores. The mean is appropriate when data is normally or symmetrically distributed. The mode is simply the data point that occurs most often in a data set. The mode is often considered the least useful measure of center, due to the possibility that the most frequent score occurs well away from the center and is not very representative of most scores.

Measures of deviation or spread of scores include standard deviation, variance, range, and interquartile range. The standard deviation of a set of scores is related to the typical deviation of a score from the mean, somewhat analogous to an average distance from the mean. The variance of a set of scores is simply the square of the standard deviation. The range is equal to the difference between the maximum and minimum score. The interquartile range is the difference between the score representing the third quartile, and the score representing the first quartile. The first quartile score represents the median of the lower half of the scores, while the third quartile score represents the median of the upper half of the scores.

Z-scores are standardized scores, with a mean of 0 and standard deviation of 1. Z-scores may be useful when comparing an individual score to a population mean, or when comparing a sample mean to a population mean, when given the population standard deviation. They can be used with any set of scores, often to interpret or compare two scores computed using different metrics. For example, a student might have a mean of 88 on one test and 83 on another. However, the mean for the first test was 74, standard deviation 12, while the mean of the second test was 66, standard deviation 15. The respective z-scores are 1.17 and 1.13, virtually identical in relation to their means, even though it was not obvious from the original data.

Note. For a normal distribution in a population, a z-distribution based on the normal curve, may be used whenever the population standard deviation is known. A z-score may be interpreted as the number of standard deviations a value falls above or below a mean. For example, suppose a student's score on a statistics exam shows a z-score of −1.5. This z-score reveals that the student's score is 1.5 standard deviations below the class mean for the exam.

The formula, $z = \frac{X - \mu}{\sigma}$, may be used to compare an individual score to a population mean, where X represents the individual score, μ represents the population mean, and σ represents the population standard deviation. Note. If only a sample mean and population standard deviation are available, the sample mean may be substituted for the population mean, in the formula shown above. However, if the population standard deviation is not known, and only the sample mean and sample standard deviation are available, a t-distribution should be used, instead. This distribution is a bit wider and flatter than the normal or "bell" curve since it accounts for the uncertainty in the standard deviation.

When comparing a sample mean to a population mean, the formula, $z = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$, may be used, where $\overline{X}$ represents the sample mean and n represents the sample size.

Let's look at a couple of examples, where descriptive statistics may be used.

Example 1

You want to compare the number of articles that are mixed methods to those that are not mixed methods. You decide to compute frequencies and then calculate percentages from those frequencies.

articles mixed = 12
articles not mixed = 34
% mixed = 12/(12 + 34) ≈ 0.26

Example 2

You want to determine the deviation of math test scores from the class mean. You decide to calculate the mean, standard deviation, variance, and *z*-score.

Scores: 44, 47, 58, 63, 74, 77, 79, 83, 84, 88, 91, 93, 96, 97
Average = (44 + 47 + 58 + 63 + 74 + 77 + 79 + 83 + 84 + 88 + 91 + 93 + 96 + 97)/14

$= 1074/14 \approx 76.71$
Variance $= [\ (44 - 76.71)^2 + (47 - 76.71)^2 + \ldots + (97 - 76.71)^2\ ] / (14 - 1)$
$= (3976.86) / 13 \approx 305.91$
Standard deviation = square root(Variance) ≈ 17.5
z-scores for Scores:
–1.87, –1.70, –1.07, –0.78, –0.15, 0.02, 0.13, 0.36, 0.42, 0.65, 0.82, 0.93, 1.10, 1.16

So, the score of 44 is 1.87 standard deviations below the mean.

Note. Since we calculated the mean that we used for the variance calculation, the averaging used to get the variance is reduced by 1 unit, corresponding to the information we already accounted for by calculating the mean. This is called degrees of freedom, corresponding to the amount of information we started with, in this case 14 scores.

WRITING A RESULTS SECTION

> *Scenario Used*
>
> We examined descriptive statistics for the variable of attitude to school.

Let's now look at a sample results section write-up for reporting descriptive statistics, using the scenario described above.

The steps we used in SPSS were:

Analyze
Descriptive Statistics
Descriptives
Options
Check Descriptives options you need (includes options of skewness and kurtosis)

Manually calculate the z-scores, which show the number of standard deviations a score is above or below the mean. (We could have also calculated the interquartile range, standard deviation, and variance, as we did for Example 2.)

We calculated the z-scores for the median, minimum score, and maximum score. These were calculated as:

$$z = \frac{X - \mu}{\sigma} = \frac{49 - 49.72}{9.819} \approx -0.07$$

$$z = \frac{X - \mu}{\sigma} = \frac{38 - 49.72}{9.819} \approx -1.19$$

$$z = \frac{X - \mu}{\sigma} = \frac{74 - 49.72}{9.819} \approx 2.47$$

The SPSS frequency output is shown below:

ATTITUDE TO SCHOOL

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	38	64	16.3	16.3	16.3
	41	55	14.0	14.0	30.4
	45	64	16.3	16.3	46.7
	49	54	13.8	13.8	60.5
	52	40	10.2	10.2	70.7
	56	33	8.4	8.4	79.1
	60	27	6.9	6.9	86.0
	63	20	5.1	5.1	91.1
	67	11	2.8	2.8	93.9
	71	15	3.8	3.8	97.7
	74	9	2.3	2.3	100.0
	Total	392	100.0	100.0	

RESULTS SECTION WRITE-UP

The mean score for attitude to school was 49.72, with a standard deviation of approximately 9.82 points. The minimum and maximum attitude to school scores were 38 and 74, respectively. The mode is both 38 and 45. Since the mean is greater than the median, and the mean is pulled towards the tail of a distribution, this distribution is positively skewed. Refer to Figure 1.1 to see the histogram of the data.

The skewness is 0.731, with standard error of skewness of 0.123. The kurtosis is −0.307, with standard error of kurtosis of 0.246. Skewness and kurtosis values between ±2 indicate that the distribution may be considered to be approximately normal. In other words, it isn't too skewed in either direction and has an approximately normal shape. Since the skewness and kurtosis are in the acceptable range of values, it makes sense that the standard errors would be small. Figure 1.2 shows the overlay of the normal curve on this distribution.

The median of 49 is approximately 0.07 standard deviations below the mean. The minimum score of 38 is approximately 1.19 standard deviations below the mean. The maximum score of 74 is approximately 2.47 standard deviations above the mean. Therefore, the median is located within 68% of the scores, the minimum score is located within 95% of the scores, and the maximum score is located within 99.7% of the scores. In other words, the maximum score is in the interval that contains only 5% of the scores. Stated another way, this value is significantly different from the mean of 49.72, if using 0.05 as the alpha level.

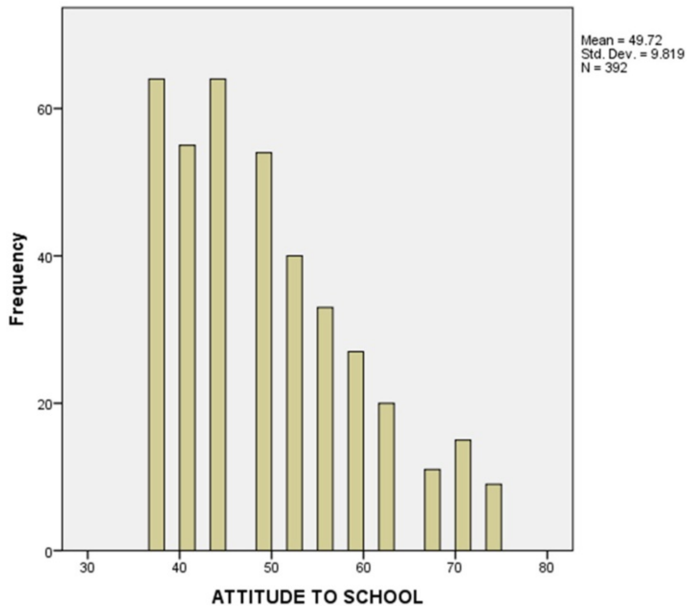

Figure 1.1. Attitude to school histogram

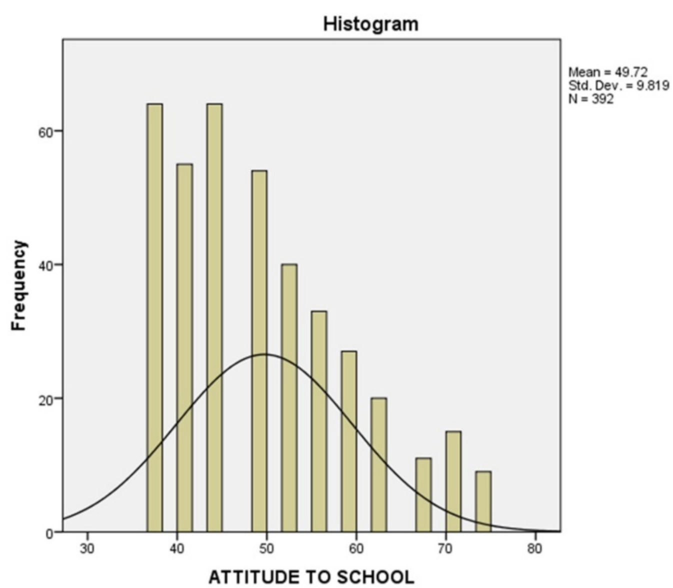

Figure 1.2. Attitude to school histogram with normal curve overlay

CHAPTER 2

ONE-SAMPLE T-TEST

BRIEF DESCRIPTION

A one-sample t-test compares the mean of a sample to an a priori score (or population mean). The test uses either a known population standard deviation or a sample standard deviation. The test analyzes interval scores. Normal distribution is not required but if data are badly skewed, a nonparametric test is preferred, such as a *binomial test* with each case scored as above (1) or below (0) the a priori mean. Note. When only a sample standard deviation is known, a *t*-test should be used, or if a population standard deviation can be specified, a *z*-test should be used.

A *t*-value may be determined using the formula, $t = \dfrac{\overline{X} - \mu}{\dfrac{s}{\sqrt{n}}}$, where s represents the sample standard deviation. The resulting *t*-value may then be compared to the critical t-value, found in the *t*-distribution table, for the appropriate degrees of freedom (number of cases minus 1) and desired level of significance. If the calculated *t*-value is less than the critical *t*-value, the null hypothesis (hypothesis of no difference) should not be rejected, with no significant difference between the sample mean and population mean declared. However, if the calculated *t*-value is larger than the critical *t*-value, the null hypothesis should be rejected, with a significant difference in means declared.

For the population standard deviation-known situation, s in the formula above is replaced with σ, the known value, and t is replaced by z, the normal curve statistic. Then, the normal distribution is referenced to compare the observed significance of the z-statistic with the desired level.

A graphing calculator, excel spreadsheet, SPSS, or other software package, may be used to calculate the *t*-value and determine significance. These tools may also be used to calculate a p-value, which represents the significance of the difference. If the p-value is less than the level of significance, which is determined by the researcher, a significant difference may be declared (and the null hypothesis rejected). Often, a level of 0.05 is used. Recall that .05 is the probability that we might get an unusual sample that actually came from the distribution with the hypothesized mean μ but it was atypical and produced an extreme mean. This could happen in 5% of the samples we drew randomly from the population.

Let's look at a couple of examples.

Example 1

You want to know if algebra exam scores of a fall freshman class are different from a mean of 70. You will use a two-tailed test, for a 0.05 level of significance.

Scores: 42, 57, 58, 59, 59, 68, 69, 69, 72, 81, 87, 87, 93, 96, 98
Sample mean: 73
Sample standard deviation: 16.7

The *t*-value may be calculated by writing: $t = \dfrac{73-70}{\dfrac{16.7}{\sqrt{15}}}$, where $t \approx 0.7$.

The critical *t*-value is 2.145. (Since there are 15 scores, the number of degrees of freedom equals 14, i.e. ($n-1$).) Since the calculated *t*-value is *not* larger than the critical *t*-value, no significant difference between the sample exam scores and score of 70 should be declared.

A graphing calculator or software package may be used to verify the calculated *t*-value, as well as to determine the *p*-value. The *p*-value is approximately 0.5. The *p*-value being greater than 0.05 (the level of significance) is further evidence that the null hypothesis should *not* be rejected and *no* significant difference declared.

In Excel, one can type into a cell the function '=T.DIST(0.7,14,TRUE)', which will return the probability .752302. Thus, .25 is the proportion of random cases that will occur about the t-statistic. For the two-tailed test, we double this to about .5, which gives us the probability of a difference as large as 3 points either above or below 70. This is the cumulative probability that a t-statistic of 0.7 or smaller would be computed from a distribution with a mean of 70 given the calculated standard deviation and 14 degrees of freedom. For the two-tailed test at the .05 level, the probability would have needed either to be less than .025 or greater than .975 for significance.

Example 2

You want to compare the salaries of employees at a start-up instructional design company to the average employee salary of employees at an established company, which is recorded as $45,000. A 0.05 level of significance will be used.

Salaries of sample: $28,000; $30,300; $31,400; $33,600; $34,100; $40,200; $41,800; $43,600; $44,900; $45,200; $48,300; $52,000

Mean of sample: $39,450

Standard deviation of sample: $7,771.57

t-value: $t = \dfrac{39,450-45,000}{\dfrac{7,771.57}{\sqrt{12}}}$ or $t \approx -2.47$

critical *t*-value: 2.201

p-value: p ≈ 0.03

Since the calculated *t*-value has an absolute value that is greater than the critical *t*-value, we will reject the null hypothesis. Thus, a significant difference should be declared. Also, the *p*-value of 0.03, which is less than 0.05, also shows that a significant difference should be declared. In other words, there is a significant difference between the salaries of employees at the startup and the average annual salary, recorded for the established company.

WRITING A RESULTS SECTION

Scenario Used

We compared the mean of a sample of depression scores to a population mean.

Let's now look at a sample results section write-up for analysis of the scenario described above.

First, let's look at the data analysis.

The steps we used in SPSS were:

Analyze

Compare Means

One-Sample T Test

Select Test Variable (depression)

Enter Test Value in Test Value box (population mean of 58)

OK

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
DEPRESSION	392	49.86	9.643	.487

The one-sample statistics table above shows descriptive statistics for the depression variable.

One-Sample Test

	Test Value = 58					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
DEPRESSION	-16.703	391	.000	-8.135	-9.09	-7.18

The One-Sample Test table above shows us the p-value (highlighted above) and gives us the calculated t-value and degrees of freedom, which are needed, if a t-table is used to determine significance, in lieu of the Sig. column.

RESULTS SECTION WRITE-UP

The sample mean for depression scores was 49.86, with a standard deviation of 9.643. There is a statistically significant difference ($p = 0.000$) between the sample mean and the population mean of 58. The critical t-value is 1.96 for a two-tailed test, alpha of 0.05, and degrees of freedom of 391. The calculated t-value is -16.703. The absolute value of the calculated t-value is greater than the critical t-value, so we will reject the null hypothesis of no difference and declare a statistically significant difference.

CHAPTER 3

INDEPENDENT SAMPLES T-TEST

BRIEF DESCRIPTION

An independent samples t-test compares the means of two groups. The data are interval for the groups. There is *not* an assumption of normal distribution (if the distribution of one or both groups is really unusual, the t-test will not give good results with unequal sample sizes), but there is an assumption that the two standard deviations are equal. If the sample sizes are equal or very similar in size, even that assumption is not critical.

The general formula for the t-value may be written as: $t = \dfrac{\overline{X_1} - \overline{X_2}}{s_{\overline{X_1} - \overline{X_2}}}$.

When the sample sizes are the same, the following formula may be used: $t = \dfrac{\overline{X_1} - \overline{X_2}}{\sqrt{\dfrac{s_1^{\ 2}}{n_1} + \dfrac{s_2^{\ 2}}{n_2}}}$.

However, when the sample sizes are *not* the same, a pooled variance estimate is used, giving the formula: $t = \dfrac{\overline{X_1} - \overline{X_2}}{\sqrt{s_p^{\ 2}\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$.

A reasonable rule of thumb for deciding whether or not the standard deviations are equal is if the ratio of the larger standard deviation to the smaller standard deviation is less than 2. If sample sizes are equal it won't matter, but if sample sizes are unequal, there is a problem in interpreting the t-test if the larger standard deviation occurs with the smaller sample size. In that case, the significance level can be quite different from what you expect, so the researcher alpha level should be made very small, say .001, to overcome the violation of the requirement.

Luckily, a graphing calculator or software package may be used to calculate this t-value.

Let's look at a couple of examples that include two independent samples and require an independent samples t-test.

Example 1

You want to compare the statistics test scores of males and females in a freshman-level statistics class.

Male scores: 56, 58, 59, 77, 81, 88, 89, 91, 91, 92, 93, 94, 95, 96, 96
Female scores: 62, 63, 64, 67, 67, 68, 77, 81, 87, 87, 93, 93

$$\overline{X_1} \approx 83.7$$

$$s_{x1} \approx 14.5$$

$$n1 = 15$$

$$\overline{X_2} = 75.75$$

$$s_{x2} \approx 12$$

$$n2 = 12$$

Note that the standard deviations are quite similar, less than the ratio of 2 discussed above, and the sample sizes are also similar, supporting use of the t-test.

Substituting these values into a graphing calculator, for a 2SampTTest, gives the following summary:

$$t \approx 1.5$$
$$p \approx 0.14$$
$$df = 25$$

Note. For this example, the alternate hypothesis used was $\mu_1 \neq \mu_2$. The variance was pooled.

Since the *p*-value is greater than 0.05, we may declare that there is *not* a significant difference in male and female statistics test scores. The *t*-value of approximately 1.5 is also less than the critical *t*-value of 2.06, revealing the same conclusion of failing to reject the null hypothesis. For the excel version discussed in Chapter 2, '=T.DIST(1.5,25,TRUE)', the probability returned was .926931, less than the .975 we would require for two-tailed significance at .05.

Example 2

You want to compare the number of articles submitted per year by assistant professors and associate professors at a particular university.

Number of articles submitted by assistant professors: 1, 1, 2, 2, 3, 4, 4, 4, 5, 6
Number of articles submitted by associate professors: 2, 2, 3, 4, 5, 7, 8, 8