***T*-Tests and Related Statistics: SAS[[1]](#footnote-1)©**

 Run the **T.sas** program, which is available on my [SAS programs page](http://core.ecu.edu/psyc/wuenschk/SAS/SAS-Programs.htm). You will need first to download the required [data files](http://core.ecu.edu/psyc/wuenschk/StatData/StatData.htm), which are: **Howell.dat** and **Tunnel2.dat.** Before running the program, be sure to edit the INFILE statements to point to the place where you deposited the downloaded data files.

## One-Sample *T*-Tests

 Look at the program and the output for the first analysis, that using [the Howell data](http://core.ecu.edu/psyc/wuenschk/SPSS/Howell-Variables.htm). We wish to test the null hypothesis that mean IQ among students in Vermont is the same as it is nationally, 100. If that hypothesis is true, and we were to subtract 100 from every IQ score in Vermont, the mean would be zero. **PROC MEANS** can be made to test the null hypothesis that a population has a mean of zero, so we just transform our sample scores, subtracting 100 from each, and then, for the resulting transformed variable, IQ\_diff, test the null hypothesis that the population mean is zero. Look at the output: As you can see, the mean IQ of students in Vermont is not significantly different from 100. Code from [Conf-Interval-d1.sas](http://core.ecu.edu/psyc/wuenschk/SAS/Conf-Interval-d1.sas) is then used to compute an estimate of Cohen’s *d* and construct a 95% confidence interval for the standardardized difference between the true mean and the hypothesized mean. The CLM option in PROC MEANS has already given you the unstandardized confidence interval.

 Another way to get the one-sample *t*-test done is to use the **H0=m** option in **PROC TTEST**, where "m" is the value of the mean in the null hypothesis. Consult SAS' help facility for more details if you wish to use this option.

# Correlated *T*-Tests

 Look at the program and the output for the second analysis, that using the Mus data. A little background will help you understand the research which generated these data. In the first experiment of my doctoral dissertation, wild‑strain house mice were, at birth, cross‑fostered onto house‑mouse (*Mus*), deer mouse (*Peromyscus*) or rat (*Rattus*) nursing mothers. Ten days after weaning, each subject was tested in an apparatus that allowed it to enter tunnels scented with clean pine shavings or with shavings bearing the scent of *Mus*, *Peromyscus*, or *Rattus*. One of the variables measured was how long each subject spent in each of the four tunnels during a twenty minute test. Also measured were the number of visits to each tunnel and the latency to first visit of each tunnel. Analysis of the data showed that the avoidance response of house mice to the scent of rats was reduced by pre-weaning experience with rats (see the related article, “Wuensch, K. L., Fostering house mice onto rats and deer mice: Effects on response to species odors. *Animal Learning and Behavior*, 1992, *20*, 253‑258.”).

 There are several mechanisms which could be involved in this change in response to rat‑scents.

* It may be that mere exposure to rat‑scents across the nursing period leads to the habituation of a fear response.
* The association of rat‑scent with positive reinforcers (milk, warmth, contact comfort, tactile stimulation, etc.) may counter-condition fear responses.
* The effects of being reared by rats may not depend at all upon associative or nonassociative exposure to rat‑scents. It may be that the parental care offered by a rat differs from that offered by house mice and deer mice in ways that produce lower emotionality and neophobia, or other changes that are more general, less stimulus‑bound than just a change in response to the scent of rats.

 Experiment 2 was designed to determine whether exposing *Mus*‑nursed pups to rat‑scents during their first 25 days of life would alter their responsiveness to rat‑scents.

 House mice and rats were bred, housed, and fostered as in the first experiment, with 16 litters being fostered onto *Mus* and 8 onto *Rattus*. One half of the *Mus*‑fostered litters were maintained in cages atop the maternity rack in the *Rattus*‑colony room (Group MR), with the other half in the *Mus*‑colony room (Group MM). Each day, approximately 50 ml of bedding was removed from each maternity cage and replaced with bedding (including freshly excreted feces) collected from the nesting area of another mother and litter in the same stage of nursing (same number of days since birth of the pups) as the receiving mother and litter.

 For the *Rattus*‑nursed litters (Group RR) and for Group MR the transferred bedding was collected from *Rattus* mothers and pups. For Group MM the bedding was collected from *Mus* mothers and pups. The transferred bedding was placed directly in the nesting area of each maternity cage. The transfer was done on Groups MM and RR to control the degree of daily disruption among the groups.

 Pups were weaned into a reversed light cycle isolation room, as was done in Experiment 1, and tested in the same apparatus as in Experiment 1, with procedural details remaining the same, except as here noted. The test was a two‑scent test, with *Rattus*‑scented bedding being placed in two tunnels 180 degrees from one another, and *Mus*‑scented bedding in the remaining two tunnels. Tests were not for a fixed 20 minute period as in Experiment 1, but rather continued for 15 minutes after each subject’s first entry into a tunnel. Sixteen subjects from each group were tested.

 The bedding transfer did disturb the mothers, but they did not remove the alien feces etc. from the nesting area, so the procedure was effective in producing long term exposure to *Rattus*‑scents in Group MR.

 Cumulative time data were normalized by a square root of (X + .5) transformation and latency data by a log (X + 1) transformation. Transformation of the number of visits data was not required.

 In the data step you can see the transformations done on the raw data (L1, L2, t1 t2). The multiplication by 1.575 was required to convert the event recorder data (lengths of lines drawn on a polygraph paper) from units of length to units of time (seconds). The square root and log transformations were needed to reduce positive skewness.

 You know that a correlated *t*-test is nothing more than a one-sample *t*-test of the null hypothesis that difference scores have a mean of zero in the population. Note how I computed those difference scores (t\_diff, L\_diff, and v\_diff) in the data step. Because I wanted to test the animals' response to the two types of tunnels separately for each nursing group, I sorted by nursing group and then did the **t**-tests **BY NURS**.

 To reduce the amount of statistical output for this lesson, I restricted the analysis to the visits data. I also placed an asterisk at the begging of each command involving the time or the latency data to suppress execution of those commands.

 It is also possible to obtain correlated t-tests by using the **PAIRED** command with the **TTEST** procedure. However, that does not give one any information on the shape of the distribution of the difference scores, so I did not go that route.

 I ask that you trust me that the transformations were successful in reducing the skewness of the time and latency data, but feel welcome to edit the program to restore the commented-out transformations and get skewness and kurtosis statistics on both the untransformed and the transformed time and latency data.

 Look at the output. The within-group distributions of the difference scores (which we assume to be normally distributed) have skewness which is within acceptable limits for what we would expect in a sample drawn from a normally distributed population. The distributions are light in their tails, but that does not concern me. If they were heavy in their tails I would be looking for outliers.

 If you look at the means for the visits, time, and latency variables, you will see that for the MM and the MR mice, the mouse-scented tunnels were more attractive than the rat-scented tunnels (they visited them more often – were we to analyze the time and latency data we would also see that they spent more time in the mouse-scented tunnels and entered them with sooner than entering the rat-scented tunnels). Are these differences statistically significant? The "t value" and "PR > |t|" columns give us the values of the correlated *t* and its *p*. Of course, we only look at that for the difference score variables. As you can see, for these two groups, the differences are significant (they are also for the time and latency variables). When you look at the results for the RR group, you find that the difference is not significant (and was not for time and latency either). It appears that it is being reared by a rat mother that causes the loss of the fear of rat-scented spaces which is observed in normal mice (keep in mind that in the wild, rats eat mice), and that mere exposure to the scent of rats early in life is not sufficient to remove this fear.

 We should estimate *d* for each difference that we have tested. It is generally not a good idea to estimate *d* from the correlated samples *t*. \*It is usually better to compute the estimate as if we had independent samples. I shall illustrate using the visits data from the MM group. , a quite large difference.

 SAS code for constructing confidence intervals for the standardized difference between two related means can be found at my [SAS Programs Page](http://core.ecu.edu/psyc/wuenschk/SAS/SAS-Programs.htm), under “Confidence Interval for **d**, Two or More Related Samples.

 I have included in T.sas only the code for the confidence interval for the group of mice reared by mice, using the Conf-Interval-Correlated-d-Algina.sas code. We are 95% confident that size of the effect is somewhere between large and very large.

## Independent Samples *T*-Tests

 Now, look at the next part of the program. Here I create a new data set, Mus2, by telling SAS to take the Mus data set and create a new variable, "Mom," which has a value of "Mouse" for those animals who were reared by a mouse (the two groups that were NE (not equal to) "RR", that is, the MM and MR groups), and a value of "Rat" for those animals who were reared by a rat. Independent samples *t*-tests are conducted with **PROC TTEST**. The **CLASS** statement identifies the grouping (independent) variable, Mom. Since the criterion (dependent) scores are difference scores, I am testing to see if the differential attractiveness of Mouse- versus Rat-scented tunnels is the same in mice reared by mice as it is in mice reared by a rat.

 Look at the output. We are given means and standard deviations with unstandardized confidence intervals. We are also given *t*, *df*, and *p* the difference between groups, both with a pooled t-test, and with a separate variances (Satterthwaite) t-test. We see that the preference for mouse over rat scent is significantly greater in the mouse-reared animals than in the rat-reared animals.

 SAS code for constructing confidence intervals for the standardized difference between two independent means can be found at my [SAS Programs Page](http://core.ecu.edu/psyc/wuenschk/SAS/SAS-Programs.htm), under “Confidence Interval for *d*, Two or More Independent Samples.” I have included in T.sas the code for comparing the mice reared by mice to the mice reared by rats. Notice that both ends of the confidence interval indicate a very large effect.

 Please note that these confidence interval programs use the pooled *t* statistics, not the separate variance *t* statistics. Please see my document [Confidence Intervals, Pooled and Separate Variances T](http://core.ecu.edu/psyc/wuenschk/docs30/CI-T-Pooled-Separate.docx).

 Look back at the program. I created another data set, Mus3, with the subsetting if statement "if nurs NE 'RR' "-- this results in the data set containing only scores from the MM and the MR group. I then compared those two groups, to see if there was any effect of exposing the mice (in the MR group) to the scent of rats. This comparison, in combination with the comparison made earlier (MM and MR together versus RR) constitutes what statisticians call a complete, **orthogonal** set of contrasts. That means that the information in the first comparison is totally independent of the information in the second comparison. When you look at the output, you see that the MR group did not differ significantly from the MM group with respect to the strength of preference for the mouse-scented (rather than rat-scented) tunnels.

 Note that SAS also gives you tests of the hypothesis of equal variances. The "**Folded *F***" that SAS reports is nothing more than the "***Fma****x*" that we have discussed. As you know, for purposes of deciding whether to use a pooled *t* or a separate variances *t*, I prefer that you not use a *p* value from this *F*, but rather use this rule of thumb: If *Fmax* > 4 or 5 or if you have unequal sample sizes, use the separate variances *t*.

### Additional Exercises

## 1. The Weight Loss Data

 The last part of the program file analyzes the data that appeared in exercises 7.33 and 7.34 of the 3rd edition of Howell’s text. Twenty subjects enrolled in each of two weight loss programs, program A and program B. The loss variable is the amount of weight lost by those who completed six months of the program. Look at the output. The means differ significantly. I want you to look at the various statistics given by SAS for the two groups: Mean, standard deviation, sample size, and confidence intervals. I want you to identify just one of those statistics as being the most important statistic, important in terms of interpreting these results.

**2. The Howell Data**

 In the [HOWELL data file](http://core.ecu.edu/psyc/wuenschk/SPSS/Howell-Variables.htm) are four dichotomous variables, SEX, SOCPROB, REPEAT, and DROPOUT, and three continuous variables, ADDSC, IQ, and GPA. You are going to use these data to demonstrate to yourself that the pooled variances *t* test is just a special case of correlation/regression analysis, by conducting an analysis like that I show here:

options pageno=min nodate formdlim='-';

**Proc** **Format**; value drop **0**='Graduated' **1**='DroppedOut' ; **run**;

**data** howell; infile 'C:\D\StatData\howell.dat';

input addsc sex repeat iq engl engg gpa socprob dropout;

format dropout drop. ;

title1 'Compare dropouts with graduates, IQ'; **run**;

**proc** **corr**; var dropout; with IQ; **run**;

**proc** **ttest**; class dropout; var iq; **run**;

**data** CI3;

t = **2.93** ;

df = **86** ;

n1 = **78** ;

n2 = **10** ;

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*;

d = t/sqrt(n1\*n2/(n1+n2));

ncp\_lower = TNONCT(t,df,**.975**);

ncp\_upper = TNONCT(t,df,**.025**);

d\_lower = ncp\_lower\*sqrt((n1+n2)/(n1\*n2));

d\_upper = ncp\_upper\*sqrt((n1+n2)/(n1\*n2));

output; **run**; **proc** **print**; var d d\_lower d\_upper; **run**; **quit**;

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*;

**data** p;

 prob = 2\*PROBT(-**2.929**, **86**);

**proc** **print**; **run**;

------------------------------------------------------------------------------------------------

 Compare dropouts with graduates, IQ

 The CORR Procedure

 1 With Variables: iq

 1 Variables: dropout

 Simple Statistics

 Variable N Mean Std Dev Sum Minimum Maximum

 iq 88 100.26136 12.98496 8823 75.00000 137.00000

 dropout 88 0.11364 0.31919 10.00000 0 1.00000

 Pearson Correlation Coefficients, N = 88

 Prob > |r| under H0: Rho=0

 dropout

 iq -0.30122

 0.0043

, *p* = .004. We conclude that IQ is significantly related to dropout status.

 Compare dropouts with graduates, IQ

 The TTEST Procedure

 Variable: iq

 dropout N Mean Std Dev Std Err Minimum Maximum

 Graduated 78 101.7 12.9583 1.4672 75.0000 137.0

 DroppedOut 10 89.4000 6.7363 2.1302 79.0000 98.0000

 Diff (1-2) 12.2538 12.4537 4.1830

 dropout Method Mean 95% CL Mean Std Dev 95% CL Std Dev

 Graduated 101.7 98.7322 104.6 12.9583 11.1955 15.3852

 DroppedOut 89.4000 84.5811 94.2189 6.7363 4.6335 12.2979

 Diff (1-2) Pooled 12.2538 3.9383 20.5694 12.4537 10.8384 14.6392

 Diff (1-2) Satterthwaite 12.2538 6.8412 17.6665

 Method Variances DF t Value Pr > |t|

 Pooled Equal 86 2.93 0.0043

 Satterthwaite Unequal 19.064 4.74 0.0001

 Equality of Variances

 Method Num DF Den DF F Value Pr > F

 Folded F 77 d 3.70 0.0383

------------------------------------------------------------------------------------------------

 Compare dropouts with graduates, IQ

 Obs d d\_lower d\_upper

 1 0.98415 0.30691 1.65593

 Notice that the value of *t* and *p* are identical for 1) the test of the null that the ρ between IQ and dropout status is zero, and 2) the pooled variances *t* test of the null that μDroppedOut = μGraduated. We have demonstrated that the usual *t* test is just a special case of the significance test used with correlation analysis. Remember that the next time some fool tells you that you cannot make causal attributions on the basis of a correlation analysis (the fools usually say something like “correlation does not imply causation”), but you can from the results of a *t* test or ANOVA. It is how you gathered the data (experimentally or not) that determines whether or not it is appropriate to make a causal attribution. In this case, since we manipulated nothing, it is not appropriate to make a causal attribution regardless of whether we use a correlation analysis or a *t* test.

 Also note that there is a significant difference in the groups’ variances, with the one being almost four times the other. Regardless of that, we should be using a separate variances test here, given the great difference in sample sizes.

 APA-style summary statement: High-school dropouts (*n* = 10) differed significantly from graduates (*n* = 78) on both variance in IQ and mean IQ. IQ was significantly more variable among graduates (*s* = 12.96) than among dropouts (*s* = 6.74), *F*(77, 9) = 3.70, *p* = .038. A separate variances *t* test showed that mean IQ was significantly greater among graduates (*M* = 101.7) than among dropouts (*M* = 89.4), *t*(19.1) = 4.74, *p* < .001, *d* = .98, 95% CI [.31, 1.66].

 You chose a pair of variables and conduct the same analysis – but do not chose the same pair I chose. Use SAS to conduct a *t*-test testing the null hypothesis that there is no relationship between the dichotomous variable and the continuous variable. Present your results in an APA-style summary statement. In addition to using the TTEST procedure, I want you to analyze these data a second way, with PROC CORR. Here is the statement: **PROC CORR; VAR X Y;** -- where "X" and "Y" are the two variables you chose. The CORR procedure gives you the Pearson *r* correlation coefficient. When *r* is computed between a dichotomous variable and a continuous variable, as we have done here, it is called a **point-biserial correlation coefficient**. One can use *t* to test the null hypothesis that a sample *r* was computed on data randomly drawn from a bivariate population in which *ρ* is zero. The *t* is computed this way:

  *n* = number of pairs of scores

 Compute that *t* and compare it to the pooled variances independent samples *t* SAS computed for you. You will demonstrate that the pooled variances independent samples *t*-test is a special case of a correlation/regression analysis!

## 3. Your Personal Data

 Choose one of the dichotomous variables and one of the continuous variables from your Personal Data Set. Conduct an independent samples *t*-test to see if the two variables are related. Report sample statistics, *t*, *df*, *d*, a confidence interval for *d*, and your conclusion as if you were writing the results section of an APA-style manuscript.

**4. The Geography of IQ, Income, and Voting Behavior**

Complete the exercise in the document [IQ, Income, and the 2004 US Presidential Election](http://core.ecu.edu/psyc/wuenschk/StatData/Bush-Kerry2004.doc).

**Data Transformations**

 When one obtains a sample of scores that are distinctly not normal, e may try one or more transformations to “normalize” the data. Please run my program Transform.sas, which is available on my [SAS programs page](http://core.ecu.edu/psyc/wuenschk/SAS/SAS-Programs.htm), and which illustrates the use of transformations to reduce positive skewness in data. The results should allow you to answer the questions posed at the end of the program file.

\*References

Dunlap, W. P., Cortina, J. M., Vaslow, J. B., and Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods*, *1*, 170–177. doi: 10.1037/1082-989X.1.2.170

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Frontiers in Psychology. <https://doi.org/10.3389/fpsyg.2013.00863>

Copyright 2021, Karl L. Wuensch - All rights reserved.

1. © Copyright 2021, Karl L. Wuensch - All rights reserved. [↑](#footnote-ref-1)