

Demo: Performing a One-Sample t Test Using PROC TTEST

Filename: **st101d02.sas**

In this demonstration we use the TTEST procedure to determine whether the population mean sale price of homes in Ames, Iowa, is \$135,000 given our sample. This procedure performs t tests and computes confidence limits. It can also use ODS Graphics to produce histograms, quantile-quantile plots, box plots, and confidence limit plots.



```
PROC TTEST DATA=SAS-data-set <options>;  
  VAR variables;  
RUN;
```

1. Open program st101d02.sas.



```
/*st101d02.sas*/  
ods graphics;  
proc ttest data=STAT1.ameshousing3  
          plots(shownull)=interval  
            H0=135000;  
  var SalePrice;  
  title "One-Sample t-test testing whether mean SalePrice=$135,000";  
run;  
title;
```

The PROC TTEST step analyzes the SalePrice variable. The H0= option specifies our null hypothesis value of 135,000. The INTERVAL option requests confidence interval plots for the means, and the SHOWNULL option displays a vertical reference line at the null value of 135,000.

2. Submit the code.
3. [Review the output.](#)

The Statistics table provides descriptive statistics of our sample, including sample size, mean, standard deviation, standard error, and minimum and maximum values of SalePrice.

The Confidence Limits table provides confidence limits for μ and σ . The default level is 95%, but you can change it with the ALPHA= option in the PROC TTEST statement. Set alpha equal to 1 minus the confidence level.

The T-Tests table provides the t test information, including the degrees of freedom, the t value and the p -value, 0.2460. Recall that if the t statistic is close to zero, and the p -value is greater than α , evidence suggests the hypothesized population parameter is statistically reasonable, and we can fail to reject the null hypothesis. Our t value is 1.16 and the p -value is greater than our α , 0.05, so we conclude that the mean sale price of homes in Ames, Iowa, is not statistically different from \$135,000.

The Mean of Sale Price Interval Plot shows the confidence interval around the mean estimate of sale price, and the vertical line references the null hypothesis value. Because the vertical reference line is within the bounds of the confidence interval, we conclude that the mean sale price of homes in Ames, Iowa, is not statistically different from \$135,000. Finally, we need to verify the validity of the test by checking that the distribution of the prices of houses is normal. We can use the histogram and Q-Q plot to verify this assumption.

In the Summary Panel, the Distribution of SalePrice histogram appears to be bell shaped, like a normal distribution. The normal and kernel density estimates are nearly overlapping, indicating that the estimated data distribution from our sample is nearly equivalent to a normal distribution.

If the data are normal, a Q-Q plot produces a relatively straight line with some deviations due to random sampling. In the Q-Q Plot of SalePrice, the sorted sale prices are plotted against quantiles from a standard normal distribution. The tail ends seem to be skewed due to possible outliers, but, overall, the plot fails to show departures from normality.

Based on the t test results, we can assume that the Student's t test is valid, and we can conclude that the mean sale price of homes in Ames, Iowa, is not statistically different from \$135,000.

Statistics 1: Introduction to ANOVA, Regression, and Logistic Regression

Copyright © 2019 SAS Institute Inc., Cary, NC, USA. All rights reserved.

Close