

# Stat Solver



The entry-level college statistics problem solver

## User's Manual and Release Information

Rough Draft Documentation For Stat Solver Version 1.0

Written August 1, 1989

Stat Solver application and documentation © 1989 by  
Jack A. Littleton

**Note: This is only a draft of the documentation for Stat Solver. The full documentation, including statistics problem examples, will be included with version 1.1, available sometime in September, 1989. More details will follow.**

**Introduction.** Why did I make Stat Solver? Surely it wasn't because of my love for classroom statistics (batting averages is a different question, however). There are three reasons why I wrote Stat Solver. The first is that I needed to write a useful application that used some routines that I developed in writing a Macintosh programming tutorial. The second is that I wanted to put what I learned in my statistics class to a practical use. The third, and most important, is that I got tired of checking my problems with the answers in the back of the textbook, only to find out that the answer in the back was incorrect.

**Disclaimer:** Even though this program went through several weeks of testing, and it correctly solved all the problems listed in my statistics workbook, I cannot be responsible for incorrect answers provided by Stat Solver. The program bugs that I am aware of are listed at the end of this paper, and other than those bugs, this program is (supposedly) error-free. If you do happen to find an error of any kind, write or E-mail to the addresses given at the end of the paper. In other words, take all answers from this program with a grain of salt!

**What this paper covers:** It is not the job of this paper to instruct you what the different statistics functions mean, or what they imply. That is your statistics instructor's job (it is not your instructors job, by the way, to inform you about the benefits of oat bran or why cholesterol is bad for you). What this paper does cover is how to enter the information needed by the statistics function this program calculates. This version (1.0, August, 1989) can solve the following functions:

- Binomial coefficients
- Binomial Distributions
- Binomial Histograms
- Normal Distributions
- Non-Standard Normal Distributions
- Confidence Intervals
- Tests of Hypotheses
- Sample Sizes
- Comparing Means

**How to get version 1.1.** At this time, a version 1.1 is in the works. The main difference between version 1.0 and 1.1 will be the enlarged documentation, including statistic problem examples, and the ability to print. There is a possibility of adding more functions, but this depends upon how long it takes to write the other portions of the program. To get information about getting version 1.1, register the program by writing to

Jack A. Littleton  
P.O Box 5301  
Chico, CA 59527

Please include a self-address stamped envelope so I can send you the particulars about version 1.1 (such as price and exactly what features have been added).

**The menus.** The use of the menus is pretty much straight forward. The file menu contains menu items that affect the way Stat Solver works. The help menu item provides a few choice words about the currently selected stat function.

The Options menu item brings up a dialog box that allows you change the number of decimal points and set the fonts used. The 'editText font' is the font used in the boxes in which you type the information given to Stat Solver. The 'Answer font' is the font the answers are drawn in.

To set the changes for the current session of Stat Solver, simply click on the 'OK' button. If you want to save the changes for posterity (or until you change your mind, whichever comes first) click on the 'Set Default Settings' button, then click on the 'OK' button. If you want to set the defaults I've set with my infinite wisdom, click on the 'Get Default Settings' button. If you don't want to change anything, click on the 'Cancel button'.

The next two menu items are reserved for version 1.1 of Stat Solver.

To quit from Stat Solver, select Quit from the menu item.

The Statistics menu reveals all the statistics functions available. To select a function, just select the menu item you want. Note that some of the items are hierarchical menus. These items have only slight

variations between them so that the grouping is somewhat logical.

The Tail menu (which is not present for all functions) tells Stat Solver about the area under a normal curve. Select the item which most closely represents the area you're interested in. The Case menu is used for functions that use different equations for different cases. How these two items are used will be explained in the paragraph for each function.

**The controls.** To find the answer to a problem (after you've filled in the boxes), click on the 'Calculate' button. To erase the data in the boxes and the answers, click on the 'Cancel' button.

**Binomial Coefficients.** This is sometimes referred to as an 'N choose R' problem, named such because of the variables used in the equation. This function finds how many subsets of size r can be constructed from a set of size n. To use this, put the number in the total set into the box labeled 'n' and the number in the subset into the box labeled 'r', like this:

Binomial Coefficients	
n =	<input type="text" value="15"/>
r =	<input type="text" value="3"/>
<div>Calculate</div> <div>Clear</div>	
$\binom{n}{r} = 455$	

In this example, a set of 15 items is broken up into subsets of size 3. As the answer indicates, there are 455 different ways a subset of 3 can be composed out of a set of 15.

**Binomial Distribution.** This function finds the probability of an experiment succeeding after a certain number of trials. Part of the equation includes a binomial coefficient as shown above. In this case the 'set' is the total number of trials to be run, and the 'subset' is the number of trials run before a experiment succeeds. You also need to know the probability of the experiment succeeding in a single trial, known as p. Put the number of trials into the box labeled 'n', and the number of trials run in the 'r' box. Put the probability of a trial succeeding into the 'p' box and the probability of a trial failing (which is 1-p) into the 'q' box:

n =	<input type="text" value="15"/>	<div>Calculate</div> <div>Clear</div>
r =	<input type="text" value="3"/>	
p =	<input type="text" value=".25"/>	
q =	<input type="text" value=".75"/>	
$\binom{n}{r} p^r q^{n-r} = 0.22520$		
$\mu = 3.75000$		
$\sigma^2 = 2.81250$		
$\sigma = 1.67705$		

Using the same n and r values as in the previous examples, and giving a probability of a trial succeeding as 25% (.25), we see that there is a 22.5% probability of the trial first succeeding on the 3rd trial.

**Binomial Histograms.** This function is much like the previous function, except that all values of r are shown (r is set equal from 1 to n). What this does is allow you to see the probability of a certain trial being the first to succeed:

n =

p =









**Calculate**

**Clear**

$(x)$	$P(x)$	$(x)$	$P(x)$
0	0.01336		
1	0.06682		
2	0.15591		
3	0.22520		
4	0.22520		
5	0.16515		
6	0.09175		
7	0.03932		
8	0.01311		
9	0.00340		
10	0.00068		
11	0.00010		
12	0.00001		
13	0.00000		
14	0.00000		
15	0.00000		

Note that  $x = 3$  shows the same value returned as in the previous example. If you want to know the probability of the 7th trial giving the first success, look under the  $x$  for 7, then look straight across under the  $p(x)$  column for the probability, which in this case is 3.932%.

**Normal Distributions.** This function returns the area (which correleates to a percentage) under the normal curve when delimited by two bounds (z-scores). To use this, you must give the left and right bounds, as well as select a normal curve template from the 'Tail' menu.

<b>Tail</b>	<b>Use When:</b>
 $0 \leq Z \leq Z_0$	<b>Left</b> $\leq 0$ <b>Right</b> $\neq \infty$
 $Z \leq Z_0$	<b>Left</b> $\leq 0$ <b>Right</b> $> 0$ or $\infty$
 $Z_0 \leq Z \leq 0$	<b>Left</b> $\neq \infty$ <b>Right</b> $\leq 0$
 $ Z  \leq Z_0$	<b> Left </b> = <b>Right</b>
 $ Z  \geq Z_0$	<b> Left </b> = <b>Right</b>
 $Z \geq Z_0$	<b>Left</b> $< 0$ <b>Right</b> $> 0$ or $\infty$
 $+Z \geq Z_0$	<b>Left</b> $> 0$ <b>Right</b> = $\infty$
 $-Z \leq Z_0$	<b>Left</b> = $\infty$ <b>Right</b> $< 0$

It is very important to select the correct tail before clicking on the 'Calculate' button. The best that can happen is you get an error, the worst is you get a wrong answer. Be careful!

Left Bounds

**Calculate**

Right Bounds

**Clear**

Probability = **0.1994**

In this example, the first tail menu was chosen, since  $\text{left} > 0$  and  $\text{right} \neq \infty$ . The area under the normal curve between z-scores .81 and 2.34 is 0.1994, or 19.94% of the total area under the normal curve.

For the tails that have the capability of infinity: leave the bound that goes to infinity blank. Do not type  $\infty$  (option-5), as it will confuse Stat Solver. Do not leave bounds of 0 blank, type in a 0.

For the absolute value tails: For  $|Z| < Z_0$  type the bound in the 'Right Bounds' box and leave the Left Bounds box empty. For  $|Z| > Z_0$  type the (positive) bound in the 'Left Bound' box and leave the Right Bounds box empty.

The last two tails, in which the left bound or right bound goes to infinity: Use these items only when one of the two tails goes to infinity, otherwise you will get wrong answers.

**Normal Distributions—Getting the Z values.** This function is the reverse of the previous function—Given a probability, a z-score is returned. Before starting the calculation, you need to pick out a tail template. But which template do you use? If the percentile value is greater than .5, you know that more than half the area of the normal curve is selected. The problems usually state to the left of or the right of the bounds desired as an answer.

Percentile

**Calculate**

Bounds = **0.850**

**Clear**

In this case,  $+Z \geq Z_0$  was selected, and it returns a z-score of 0.85, which means the area between 0.85 and infinity on the normal curve is .1994, or 19.94% of the total area.

**Non-Standard Normal Distributions.** A normal distribution has a mean and a standard deviation of 0. A non-standard normal distribution shifts the normal distribution according to a mean and standard deviation. Enter the un-adjusted bounds into the boxes labeled Left Bounds and Right Bounds, then enter the mean in the box labeled  $\mu$  and the standard deviation into the box labeled  $\sigma$ .

Left Bounds

**Calculate**

Right Bounds

**Clear**

$\mu$

$\sigma$

Probability = **0.4294**

In this case, a non standard normal curve is created with a mean of 3.1 and standard deviation of .8. The area between 3.2 and 4.7 is 0.4294, or 42.94% of the total area. You still need to pick a template from the tail menu. In this case, menu item #1 was picked from the tail item. How to you pick a tail template? Look at the tail menu chart above. Instead of comparing the left and right bounds with zero, compare it with the mean. Menu item # 1 was picked in this case because the left bound, 3.2, is greater than the mean of 3.1, and the right bound was not infinity.

**Non-Standard Normal Distribution—Getting Z Values.** This is the reverse of the previous function. Given a probability (area), and a mean and standard deviation, the function returns a bound:

Percentile

$\mu$

$\sigma$

Calculate

Clear

Bounds = **3.392**

In this case, the curve has a mean of 3.1 and standard deviation of .4. Menu item #7 was used, meaning that the area under the normal curve from 3.392 to infinity is equal to 0.2334.

**Confidence Intervals for  $\mu$ .** This function creates a confidence interval given a confidence level (actually alpha), a mean, standard deviation and number of trials to base the confidence interval on:

$\alpha$

$\bar{x}$

$\sigma$

n

Calculate

Clear

**Lower bound = 33.64938**

**Upper bound = 34.35062**

In this example the confidence level was .95 (alpha is 1 - CI, or .05), the sample mean (or  $\bar{x}$ ) was 34, the standard deviation was 1.2, and it took 45 samples to determine the mean and standard variation. This results in a confidence level between 33.64938 and 34.35062.

The Case menu is used in this function to use the right equation for the data given. Here's how the cases are used:

Case 1 & 2—Used for large samples (n is larger than 30) and the standard deviation is known.

Case 3—Used for small samples.

**Confidence Levels for p.** This function is much like the previous function except that it uses a binomial type variable. The information needed is basically the same, but slightly different; the confidence level, the number of successes and the number of trials:

$\alpha$

x

n

Calculate

Clear

**Lower limit = 0.37120**

**Upper Limit = 0.62880**

In this example, the confidence level is .99 (alpha = 1-CL), the number of successes (x) is 50 and number of trials (n) is 100. The confidence level ranges from 0.37120 to 0.62880.

**Test of Hypotheses for  $\mu$ .** This function advises to reject or not reject a hypothesis. You must enter the information for the hypothesis: a confidence level, mean, sample mean, standard deviation, and sample size. Along with this information you must choose from a modified tail menu and a case menu (as mentioned in the paragraph on confidence intervals for  $\mu$ ).

**Tail**

✓ 	Two-Tailed CR	⌘=	"Differs significantly from"
	Right Tailed CR	⌘R	"Significantly higher"
	Left Tailed CR	⌘L	"Significantly lower"

The phrases to the right of the menu are used to describe tests made on a mean. If the test is made to see if the mean is significantly different than the sample mean, select the two-tailed critical region. if the test is made to see if the mean is significantly higher than the mean, select the right tailed critical region. If the test is made to see if the mean is significantly lower than the mean, select the left tailed critical region.

$\alpha$

**Calculate**

$\mu$

**Clear**

$\bar{x}$

$\sigma$

$n$

**Critical Region = 1.8810**

**Computed Z = -2.6123**

**Reject Ho**

In this example, the significance level is .06, the mean is 9, the sample mean is 8.43, standard deviation is 1.38 and 40 trials were conducted to get the sample mean. The function returns the critical region, the computed z value, and whether or not to reject the hypothesis. In this case, the two-tailed critical region was selected, and case 1 was used.

There are two other selections in the same hierarchical menu that deal with hypothesis tests for  $\mu$ . The information that is entered is the same, but the output and the equations used to obtain the output are different. The P Value technique uses a binomial variable to reject (or not reject) the hypothesis. If the P value is less than the significance level, the hypothesis is rejected (using the same information from the examples above):

**P Value = 0.0088**

**Reject Ho**

The CI Method creates a confidence interval for the information given. If the mean does not fall in the confidence level, the hypothesis is rejected:

**Lower = 8.0196**

**Upper = 8.840**

**$\mu$  = 9.0000**

**Reject Ho**

**Test of hypotheses for p.** This function tests hypotheses for binomial variables. The standard four step method requires the significance level, the p value, the number of successes in the sample trial, and the number of trials in the sample. You also need to select a tail template from the Tail menu as described in the chapter on testing hypotheses for  $\mu$ :



$\alpha$	<input type="text" value=".05"/>
$p$	<input type="text" value=".72"/>
$\hat{p}$ $\swarrow$ $x$	<input type="text" value="269"/>
$\nwarrow$ $n$	<input type="text" value="347"/>

**Calculate**

**Clear**

**Critical Region = 1.960**

**Computed Z = 2.291**

**Reject  $H_0$**

Just as with the tests for  $\mu$ , this function returns the critical region and the computed z value, as well as whether or not to reject the hypothesis.

The P Value method and the CI method require the same information as the example above, and return the same information as the test for  $\mu$  examples.

**Comparing Means.** These functions determine if a mean is significantly different, greater than, or less than another mean. The information that you need to supply is the same as the hypothesis test functions, except the data for two means is supplied:

$\alpha$	<input type="text" value=".04"/>		
mean 1:		mean 2:	
$\bar{x}$	<input type="text" value="50.38"/>	$\bar{x}$	<input type="text" value="46.73"/>
$\sigma$	<input type="text" value="9.27"/>	$\sigma$	<input type="text" value="12.39"/>
$n$	<input type="text" value="35"/>	$n$	<input type="text" value="46"/>

**Calculate**

**Clear**

**Critical Region = 2.054**

**Computed Z = 1.517**

**Do not reject  $H_0$**

Enter the each mean's data in the appropriate column. In this example, mean 1 is 50.38 with a standard deviation of 9.27, with 35 trials needed to determine this information. Mean 2 is 46.73 with a standard deviation of 12.39, with 46 trials needed. The information returned is just like the other 4-step method functions. You need to select a tail template from the Tail menu, just as shown in the hypothesis test paragraph, as well as a case from the Case menu. Select case 1 if you are testing a large ( $n > 30$ ) sample, or case 3 if testing a small ( $n < 30$ ) sample.

The P-Value method requires the same data as the four-step method, and returns the same information that the previous P-Value functions returned:

**P Value = 0.129**

**Do Not Reject  $H_0$**

The CI method also requires the same data as the four-step method, and returns the same information that the previous CI method functions returned:

**Lower = -1.293**

**Upper = 8.593**

**$\mu = 0$**

**Do not reject  $H_0$**

Note that all three functions advise not to reject  $H_0$ , even though the methods used to determine this are different.

**Known bugs:**

Actually, there are no known bugs, and if there are they are very elusive. Listed below are bugs that I believe to have exterminated, plus some hints for using the program.

There was a very frustrating menu bug that I believe to have fixed. When switching from one function to another (without closing windows) the menu bar would occasionally become filled with random characters where a tail or case menu was supposed to be. If this ever happens, quit Stat Solver, launch it again, and pick the desired function. If you do not quit and click on the random characters, the program will crash.

In the normal and non-standard normal curve functions, an alert is drawn if either of the bounds is greater than 3.09, telling you that because of lookup-table constraints (which goes up to 3.09), you cannot use a value greater than 3.09. If you need to use a value greater than this, use 3.09 anyway. The answer will not be exact, but will be within a few decimal points.

While we're on the subject, there could be a rounding error problem in all the functions that use any of the lookup tables (all the functions on the statistics menu following the normal curve functions). The lookup tables are correct to three decimal places. The workbook that I used to confirm answers returned by Stat Solver used tables that were correct to only two places, which meant that the answers were very rarely exactly the same. You may have the same problem with your homework problems, and should be aware of the problem.

Please forgive me for any strange terminology that may have been used in either the documentation or the program itself. The function and variable names used in this program were the same as those used in the statistics class I took at California State University, Chico. The instructor apparently liked to create his own terminology. I did look up some of the functions in other statistics books and tried to achieve a happy medium.