

Using SPSS for One Sample Tests

SPSS isn't as good as Stata for one sample tests. As far as I know, it can't handle Case I at all. It does not have anything like Stata's calculator functions, so you have to have raw data. More information is sometimes available in Stata's output. Nonetheless, SPSS is probably adequate for most needs.

A. Single Sample Tests Case I: Sampling distribution of \bar{X} , Normal parent population (i.e. X is normally distributed), σ is known.

I don't know of any way to do Case I in SPSS.

B. Case II: Sampling distribution for the binomial parameter p .

Problem. The mayor contends that 25% of the city's employees are black. Various left-wing and right-wing critics have claimed that the mayor is either exaggerating or understating the number of black employees. A random sample of 120 employees contains 18 blacks. Test the mayor's claim at the .01 level of significance.

SPSS Solution. In SPSS, you use the NPAR TESTS command with the BINOMIAL option. (On the SPSS menus, it is Analyze/Nonparametric Tests/Binomial.) This gives you results that are identical (as far as I can tell) to Stata's `bitest` command. For the dichotomy, the default is for the lower value to correspond to success while the higher value stands for failure (which is the opposite of the 0-1 failure/success coding used by `bitest`).

* Enter the data. I am including the data in the syntax, but it is probably
* easier just to use the SPSS Data editor.

```
Data List Free / X Wgt.
```

```
Begin Data.
```

```
1 18
```

```
2 102
```

```
End Data.
```

```
Weight by Wgt.
```

```
Value Labels X 1 "Success" 2 "Failure".
```

```
NPAR TEST
```

```
  /BINOMIAL (.25)= X
```

```
  /MISSING ANALYSIS.
```

NPar Tests

Binomial Test

	Category	N	Observed Prop.	Test Prop.	Asymp. Sig. (1-tailed)	
X	Group 1	Success	18	.15	.25	.005645 ^{a,b}
	Group 2	Failure	102	.85		
	Total		120	1.00		

a. Alternative hypothesis states that the proportion of cases in the first group < .25.

b. Based on Z Approximation.

Note that, unless $p = .5$, SPSS assumes that you want to do a 1-tailed test. Since we want to do a two-tailed test, you should double the reported significance level (although unfortunately, I don't think this gives results that are exactly right; use Stata if you want the exact 2-tailed probability.)

This is identical (albeit less detailed) to what Stata says:

```
. bitesti 120 18 .25
```

```

      N   Observed k   Expected k   Assumed p   Observed p
-----
      120           18           30       0.25000       0.15000

Pr(k >= 18)           = 0.997208   (one-sided test)
Pr(k <= 18)           = 0.005645   (one-sided test)
Pr(k <= 18 or k >= 43) = 0.011020   (two-sided test)

```

C. Case III: Sampling distribution of \bar{X} , normal parent population, σ unknown.

Problem. The Deans contend that the average graduate student makes \$8,000 a year. Zealous administration budget cutters contend that the students are being paid more than that, while the Graduate Student Union contends that the figure is less. A random sample of 6 students has an average income (measured in thousands of dollars) of 6.5 and a sample variance of 2. Using both confidence intervals and significance tests, test the Deans' claim at the .10 and .02 levels of significance.

SPSS Solution. In SPSS, you need to analyze raw data. Once you have such data, you can use the T-TEST command (using the menus, it is Analyze/ Compare Means/ One-Sample T-Test). For the .10 level of significance,

```
T-TEST
  /TESTVAL = 8
  /MISSING = ANALYSIS
  /VARIABLES = pay
  /CRITERIA = CI(.90) .
```

T-Test

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
pay	6	6.5000	1.41421	.57735

One-Sample Test

	Test Value = 8					
	t	df	Sig. (2-tailed)	Mean Difference	90% Confidence Interval of the Difference	
					Lower	Upper
pay	-2.598	5	.048	-1.50000	-2.6634	-.3366

Compare this to Stata, where we got

```
. ttest pay = 8, level(90)
```

One-sample t test

```
-----
Variable |      Obs      Mean   Std. Err.   Std. Dev.   [90% Conf. Interval]
-----+-----
      pay |         6         6.5   .5773503   1.414214   5.336611   7.663389
-----
```

Degrees of freedom: 5

Ho: mean(pay) = 8

Ha: mean < 8
t = -2.5981
P < t = 0.0242

Ha: mean != 8
t = -2.5981
P > |t| = 0.0484

Ha: mean > 8
t = -2.5981
P > t = 0.9758

The results are the same, except that SPSS reports the CI for the difference, i.e. it subtracts the hypothesized value from the upper and lower bounds. Hence, in SPSS, if 0 does not fall within the CI of the difference, you reject the null. For the 98% CI in SPSS,

T-TEST

```
/TESTVAL = 8
/MISSING = ANALYSIS
/VARIABLES = pay
/CRITERIA = CI(.98) .
```

T-Test

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
pay	6	6.5000	1.41421	.57735

One-Sample Test

	Test Value = 8					
	t	df	Sig. (2-tailed)	Mean Difference	98% Confidence Interval of the Difference	
					Lower	Upper
pay	-2.598	5	.048	-1.50000	-3.4427	.4427

Again, this matches up with Stata's results:

```
. ttest pay = 8, level(98)
```

One-sample t test

```
-----+-----
Variable |      Obs      Mean   Std. Err.   Std. Dev.   [98% Conf. Interval]
-----+-----
      pay |         6       6.5   .5773503   1.414214   4.557257   8.442743
-----+-----
```

Degrees of freedom: 5

Ho: mean(pay) = 8

Ha: mean < 8
t = -2.5981
P < t = 0.0242

Ha: mean != 8
t = -2.5981
P > |t| = 0.0484

Ha: mean > 8
t = -2.5981
P > t = 0.9758