MM104/MM106/BM110
Statistics and Data Presentation

Lecture 6-5: Confidence Intervals
T intervals for the mean

Chris Robertson

University of Strathclyde Science

DEPARTMENT OF MATHEMATICS & STATISTICS

Confidence limits standard deviation unknown

- If the sample size is large – bigger than 30-50 – and the populations standard deviation is unknown you can use the sample standard deviation and the z value for the standard normal distribution to get a confidence interval for the mean
- What happens if the sample size is small – less than 30?
- Use the t- distribution

Confidence limits standard deviation unknown

- Why use the t- distribution?
- It is more variable than the standard normal and so takes into account the variation induced by estimating the standard deviation in small samples
- Both $\bar{x}$ and s will vary from sample to sample and the t distribution takes this into account whereas confidence intervals based upon the z values do not.

Confidence limits standard deviation unknown

- The usual situation that the population standard deviation is unknown when you are estimating the mean.
- In order to calculate the variance or standard deviation you need to know the mean.
- Situations where you know might the population standard deviation when estimating a mean
  - Variable measured repeatedly in a population, such as
  - IQ (15 units), Blood Pressure, heights, weights,
  - Manufacturing production lines

Confidence limits standard deviation unknown

- In virtually every other situation you will not know the standard deviation when estimating the mean and so you need to replace the z value in the equation for the confidence interval for the mean with a value from the t distribution

Replace with a value from the t distribution

Replace with the sample standard deviation, s

$$\bar{x} \pm z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

The degrees of freedom for the t distribution is given by $\nu = n - 1$, where n is the sample size

Confidence limits standard deviation unknown

- Different calculation of confidence intervals when the standard deviation is unknown and small sample size.
  - Estimate the standard deviation from the sample, s.
  - The $100(1-\alpha)\%$ confidence interval is defined by different statistics (t-distribution!):

$$\bar{x} \pm \boxed{t_{\alpha/2}^{\nu}} \left(\frac{s}{\sqrt{n}}\right)$$

Degrees of freedom

The t-distribution with constant $\boldsymbol{\nu = n\text{-}1}$ covers an area $(1\text{-}\alpha)$ between $-t$ and t.

Example

Imaginary airline Ryanjet covers a route overseas. In order to be fuel-efficient, the average speed should not go over 852 Km/h. The following numbers are cruising speeds in Km/h, taken on random different flights (same route) over a month. The speeds are thought to be Normally distributed. Find a 99% confidence interval for the mean cruising speed.

794
880
797
901
818
834
879
833
862
816
804
812
861
838
818
826
898
835

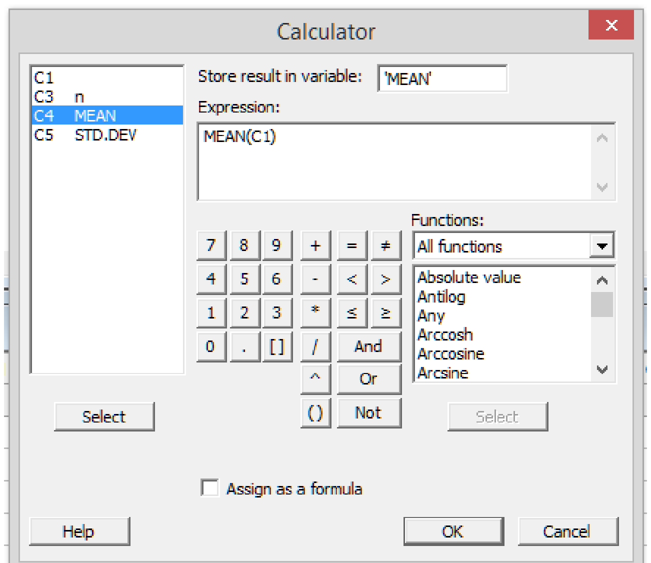
99% confidence level so

$$\alpha/_2 = 0.005$$

Example

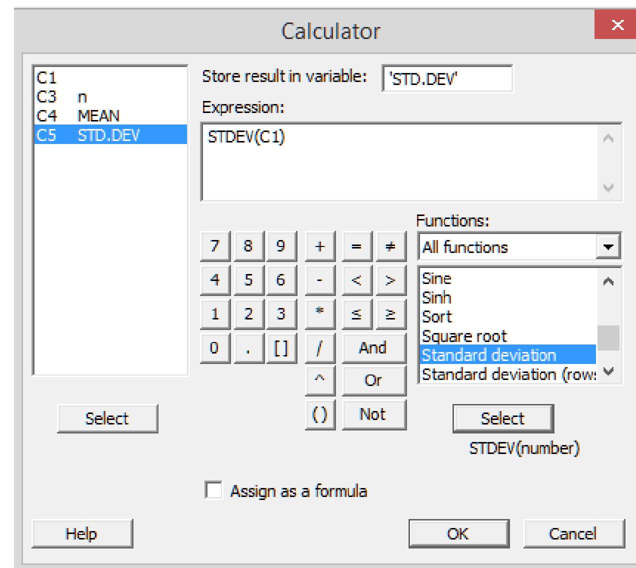
Paste the data into Minitab

Use the calculator to get the sample mean and standard deviation



n	MEAN	STD.DEV
18	839.222	33.4480

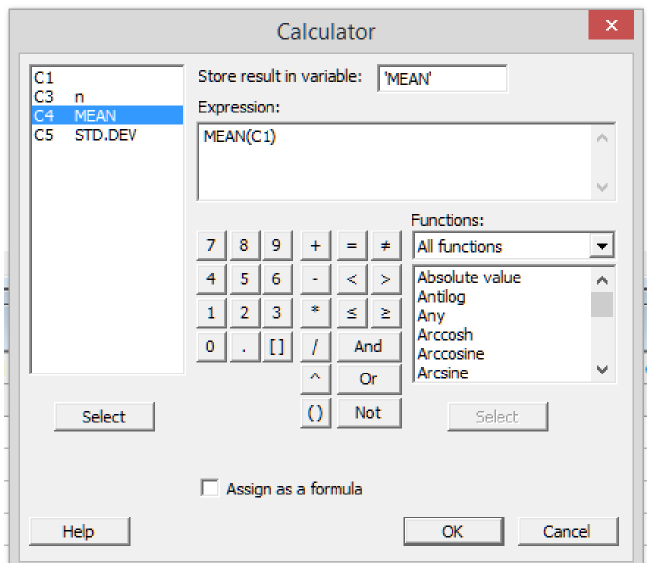
$$\bar{x} \pm t^{\nu}_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)$$



Example

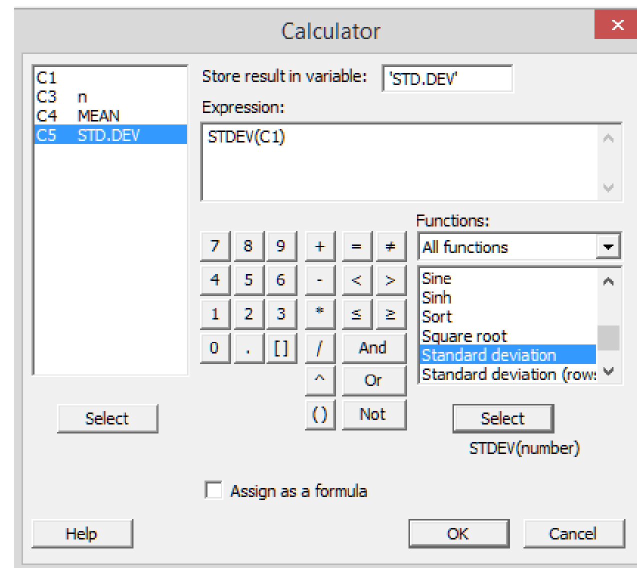
Paste the data into Minitab

Use the calculator to get the sample mean and standard deviation



n	MEAN	STD.DEV
18	839.222	33.4480

$$\bar{x} \pm t^{\nu}_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)$$

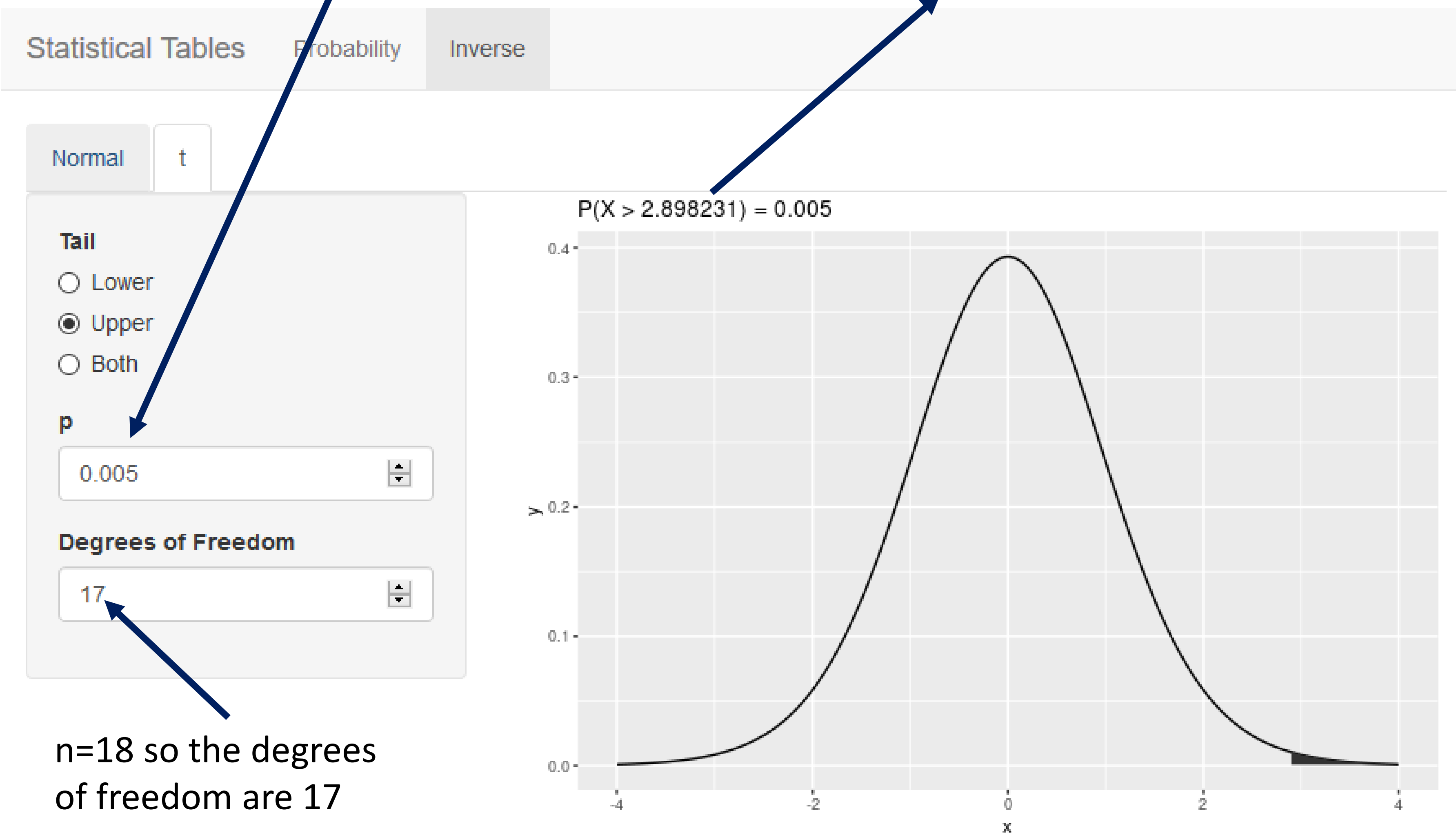


n=18 so the degrees of freedom are 17

Example

99% confidence level so
$^{\alpha}/_{2} = 0.005$

Value to use in the t interval



n=18 so the degrees
of freedom are 17

Example

In order to be fuel-efficient, the average speed should not go over 852 Km/h. Find a 99% confidence interval for the mean cruising speed.

$$\bar{x} \pm t^{\nu}_{\alpha/2}\left(\frac{S}{\sqrt{n}}\right)$$

$\bar{x} = 839.222$

$\frac{s}{\sqrt{n}} = \frac{33.4480}{\sqrt{18}} = 7.884$

$\nu = n - 1 = 17$

$(1 - \alpha) = 0.99;$

$t^{\nu}_{\alpha/2} = 2.898231$

UPPER C.L.

$$\bar{x} + t^{\nu}_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right) = \boxed{862.07}$$

LOWER C.L.

$$\bar{x} - t^{\nu}_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right) \quad \boxed{816.37}$$

The average speed is 839 km/h with a 99% confidence interval of (816, 862)km/h

Key Points

- Confidence limits for a mean in small samples where the variable is normally distributed are given by

$$\bar{x} \pm t_{\alpha/2}^{v}\left(\frac{s}{\sqrt{n}}\right)$$

The degrees of freedom for the t distribution is given by $\nu = n - 1$, where *n* is the sample size

In large samples, n > 30 the values from the t distribution are similar to the z vales from the standard normal distribution and so t intervals are really only used in small samples.

The z intervals can be used in large samples even if the original variable is not normal

To use the t distribution in small samples the variable must follow a normal distribtuion