Tutorial: Confidence intervals with the t-distribution in Stata

Suppose t is a random variable that follows a t-distribution with n degrees of freedom.

`tden(n,t)`	returns the probability density function of Students t distribution
`ttail(n,t)`	returns the reverse cumulative (upper tail or survivor) Students t distribution
`invttail(n,p)`	returns the inverse reverse cumulative (upper tail or survivor) Students t distribution

Note that if `ttail(n,t)`$= p$, then `invttail(n,p)` $= t$.

Stata will calculate confidence intervals for you:
Calculator: `cii n mean sd, level(95)`
Function: `ci varlist, level(95)`

There is no Stata function for calculating confidence intervals for normally distributed data when the standard deviation is known, since this scenario doesn't really happen in practice.

1. Calculate the mean and standard deviation of BMI at baseline.

```
. summarize bmi1
```

2. **Take a sample of size 20 from the Framingham cohort.** Calculate the mean and standard deviation of BMI at baseline in the subsample (I use `set seed 2`, if you want to get the same sample as me). *We are interested in making inference about BMI at baseline in the total Framingham cohort using only the sample of size 20.*

```
. set seed 2
. drop if bmi1 == .
. sample 20, count
. sum bmi1
```

3. Assume that the sample standard deviation is known (and equal to the standard deviation in the Framingham cohort). Construct a 95% confidence interval for the mean BMI in your subsample. Note that if `normal(z)`$= p$, then `invnormal(p)` $= z$.

   95% CI: $\bar{x} \pm Z_{0.975}\sigma/\sqrt{n}$

```
. di 25.0 - invnormal(0.975)*4.1/sqrt(20)
. di 25.0 + invnormal(0.975)*4.1/sqrt(20)
```

4. Use `invttail` to construct a 95% confidence interval for the mean BMI in your subsample by hand, now assuming that the **sample standard deviation is unknown**.

```
. di 25.0 - invttail(19, 0.025)*3.2/sqrt(20)
. di 25.0 + invttail(19, 0.025)*3.2/sqrt(20)
```

5. Use `cii` to construct a 95% confidence interval for the mean BMI in your subsample.

```
. cii 20 25.0 3.2
```

6. Use `ci` to construct a 95% confidence interval for the mean BMI in your subsample.

```
. ci bmi1
```