Probability theory

Justifying inference in regression settings

Possibilities:

Assumptions about distribution of error terms

- Normal distribution
- t distribution

Asymptotic results - when number of observations go to infinity

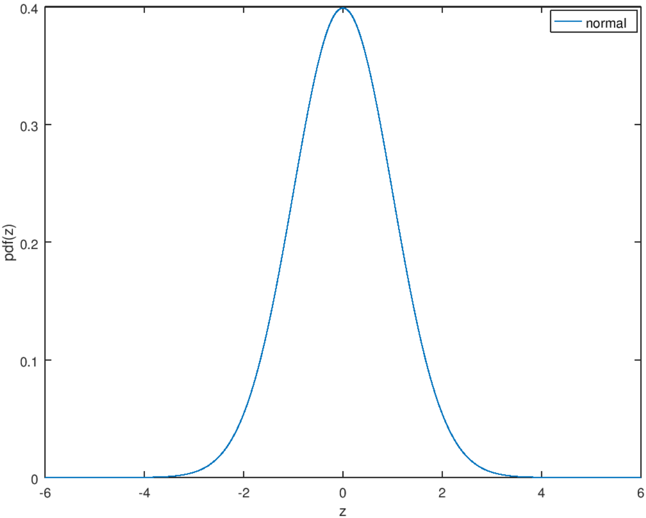
- Asymptotic normality

Probability distributions important for statistics

- Normal
- t-distribution
- The χ^2 distribution
- The F distribution

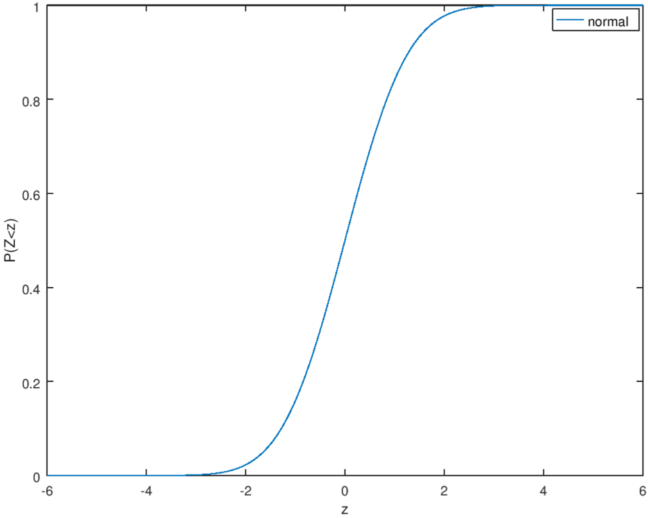
The unit normal distribution

Is defined over $(-\infty, \infty)$, but most of the probability mass sentered at zero



The unit normal distribution

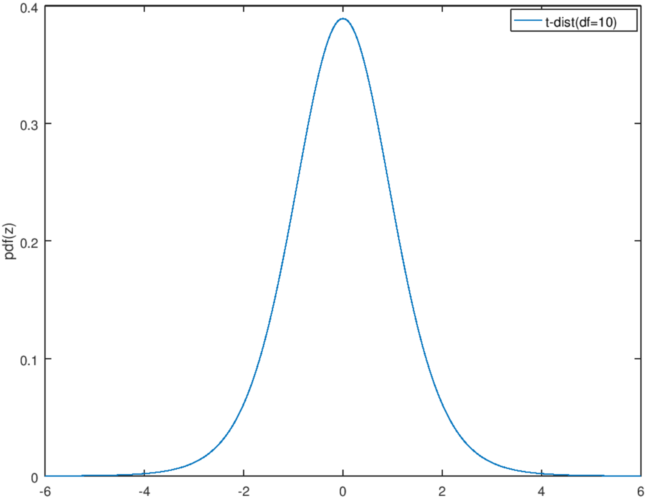
The cumulative probability function of a normal distribution



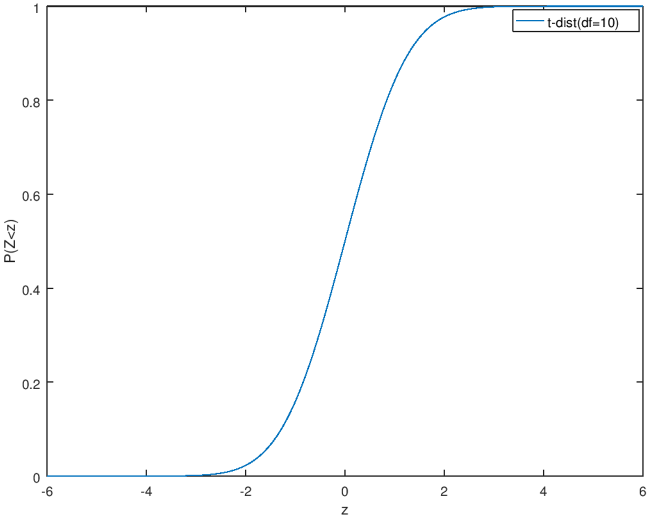
The T distribution

Is similar to the normal, defined over $(-\infty, \infty)$, but most of the probability mass sentered at zero

The probability density function of a t-distribution (df=10)

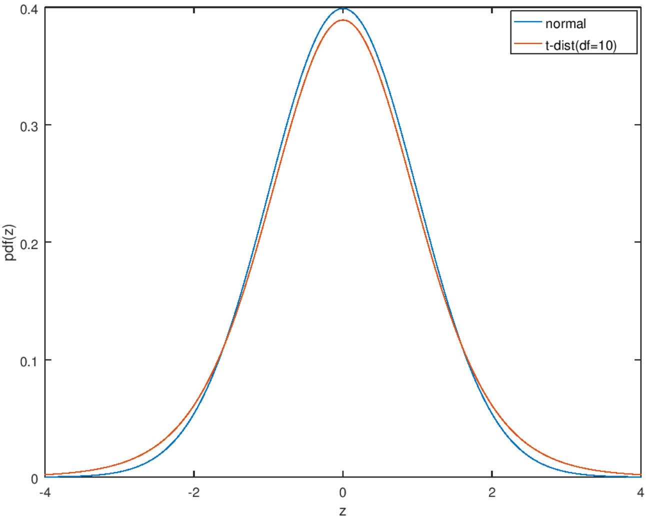


The cumulative probability function of a t-distribution (df=10)

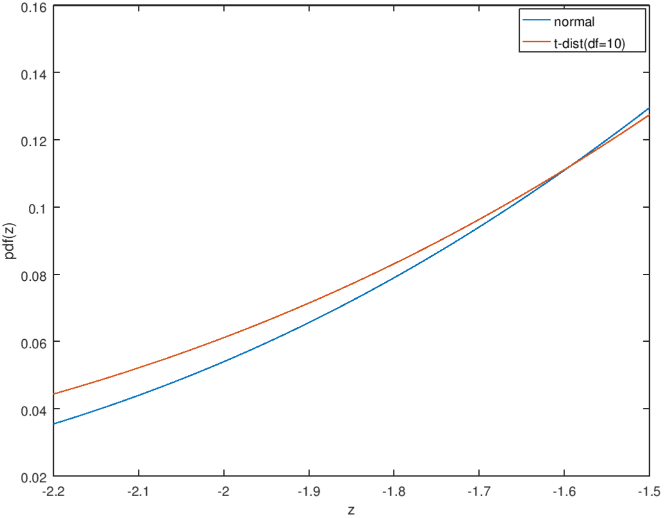


Comparing normal distribution and t distribution

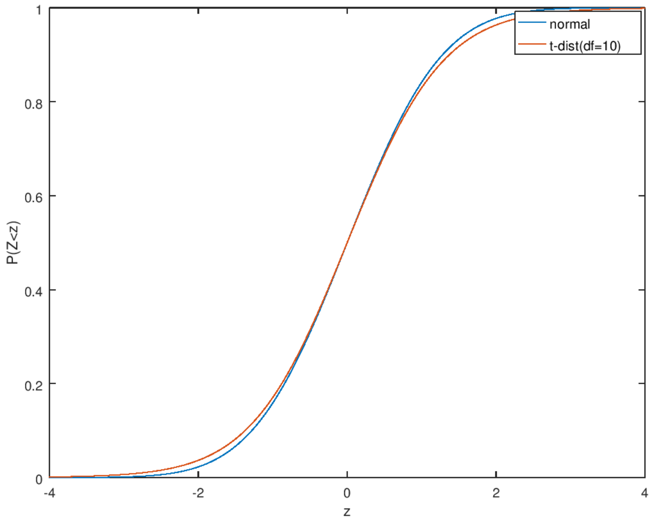
Comparing probability density functions of a normal distribution and a t-distribution (df=10)



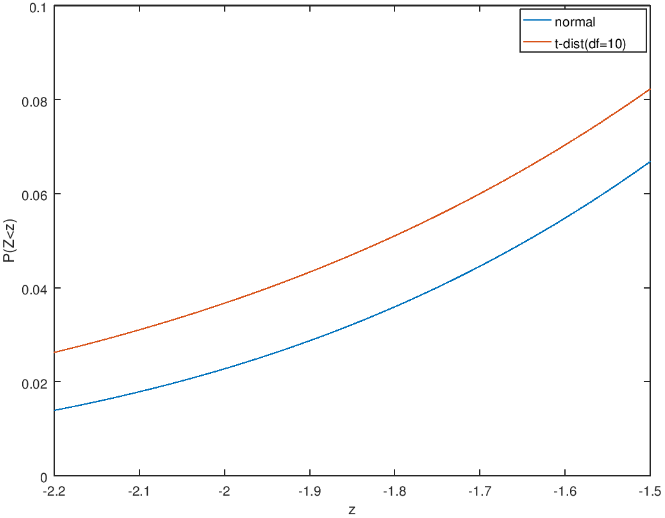
Comparing probability density functions of a normal distribution and a t-distribution (df=10) - Detail



Comparing cumulative probability functions of a normal distribution and a t-distribution (df=10)

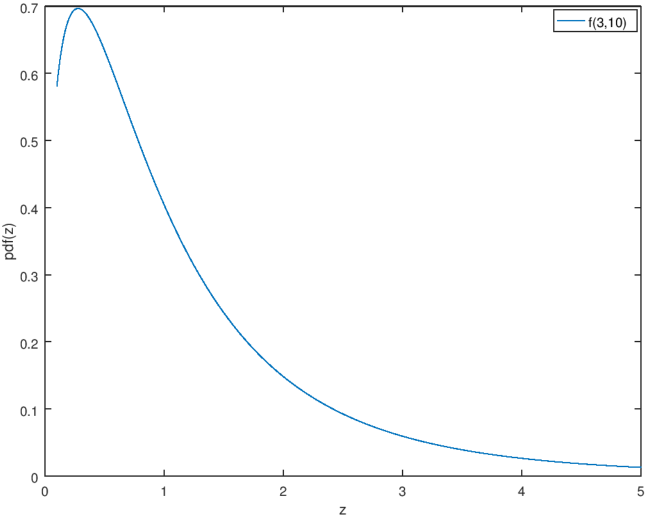


Comparing cumulative probability functions of a normal distribution and a t-distribution (df=10) - Detail



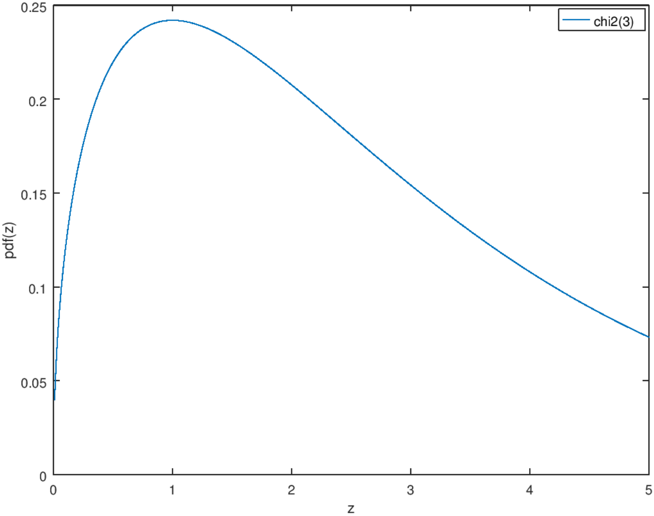
F dist

The probability density function of a F-distribution (df=3,10)



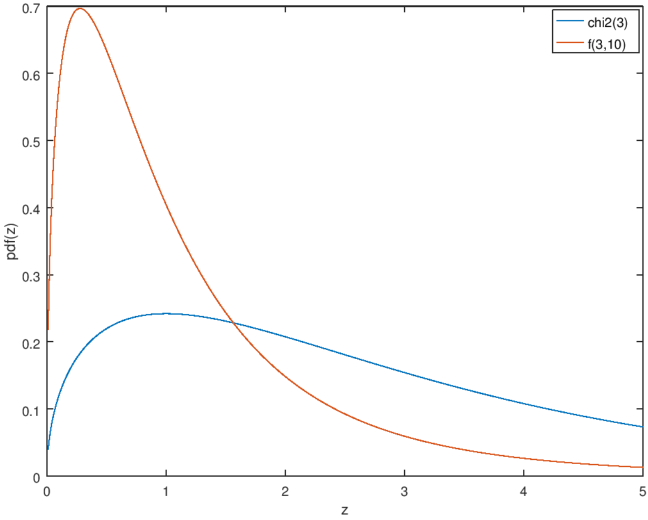
χ^2 dist

The probability density function of a χ^2-distribution (df=3)

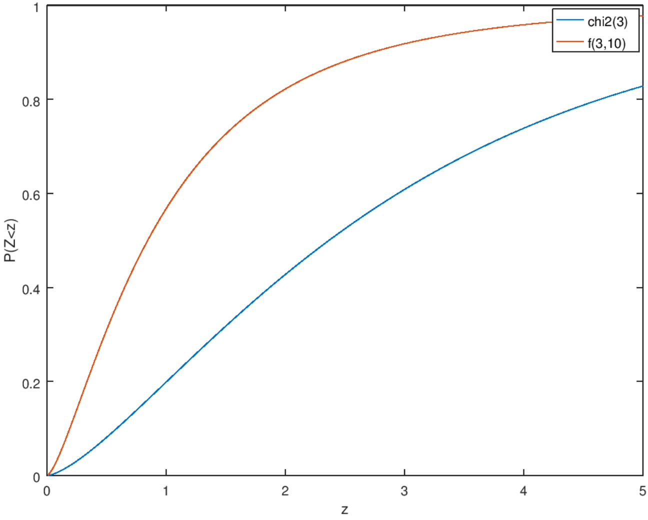


Comparing F and χ^2 distributions

The probability density functions of a F and χ^2-distribution (df=3)



The cumulative probability functions of a F and χ^2-distribution (df=3)



Motivating the concept of convergence in probability.

$$\tilde{x}_i \sim N(0,1)$$

Consider drawing T samples of x_i, and take the average

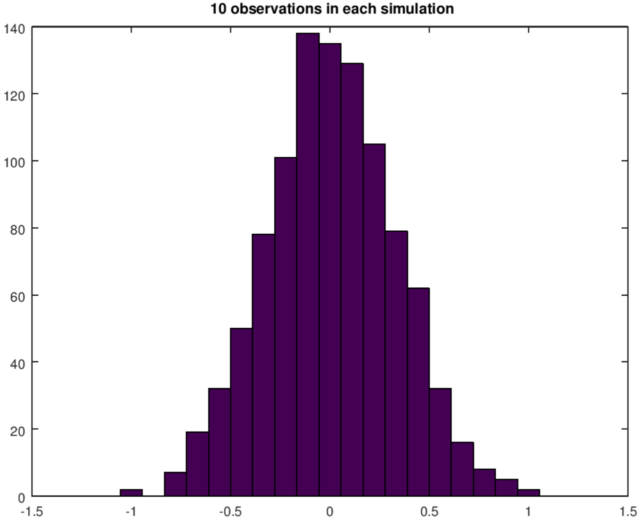
$$X_T = \sum_{t=1}^{T} x_t$$

Do this as a simulation, where in each simulation we draw T $\tilde{x}_i$ observations.
Make a histogram of the resulting averages for values of T of 10, 100 and 1000

Using matlab to do simulation of convergence in probability

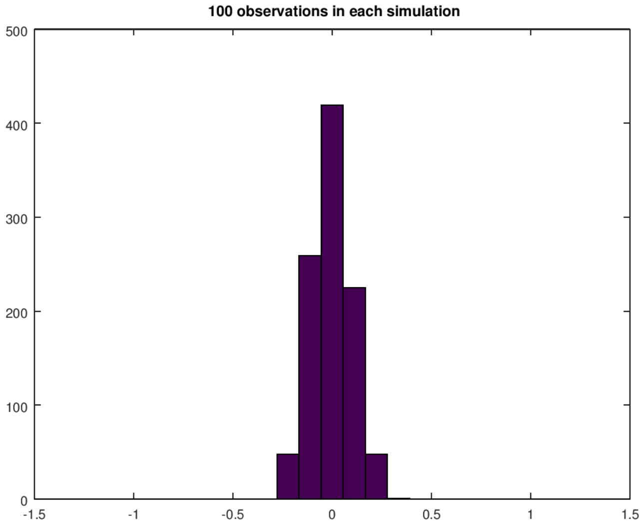
```
function con_prob(n, no_sims)
  X = randn(n,no_sims);
  bins = linspace(-1,1,19);
  m = mean(X);
  hist(m,bins);
endfunction;
```

$T = 10$

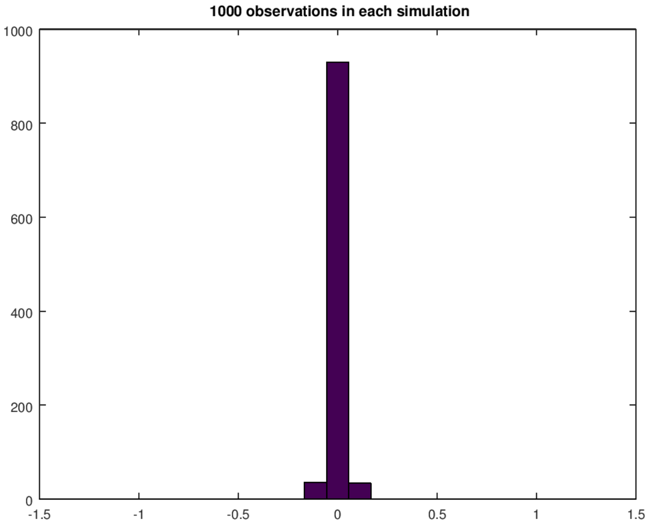


10 observations in each simulation
140
120
100
80
60
40
20
0
-1.5
-1
-0.5
0
0.5
1
1.5

$T = 100$


100 observations in each simulation
500
400
300
200
100
0
-1.5
-1
-0.5
0
0.5
1
1.5

$T = 1000$


1000 observations in each simulation
1000
800
600
400
200
0
-1.5
-1
-0.5
0
0.5
1
1.5

Convergence to a constant

We have just illustrated the concept of

- Convergence in probability

Used to justify estimation: In large sample, the sample mean will converge to its true value.

Motivating the concept of convergence in distribution.

$$\tilde{x}_i \sim N(0,1)$$

Consider drawing T samples of x_i, and take the average

$$X_T = \sum_{t=1}^{T} x_t$$

Multiply this average with $\sqrt{T}$:

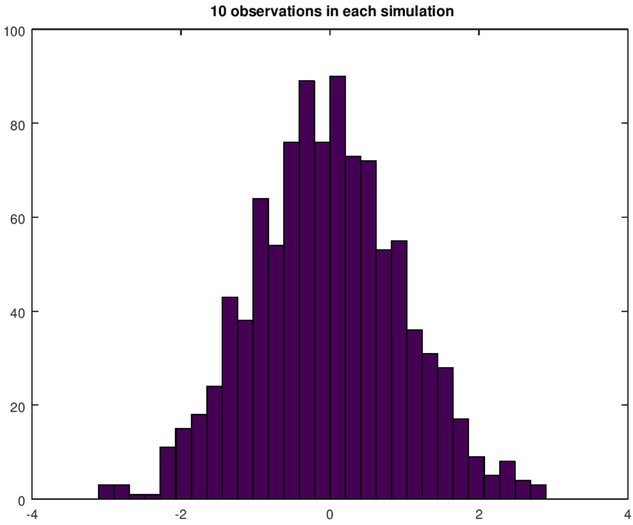
$$\sqrt{T} X_T$$

Do this as a simulation, where in each simulation we draw T $\tilde{x}_i$ observations.
Make a histogram of the resulting averages for values of T of 10, 100 and 1000

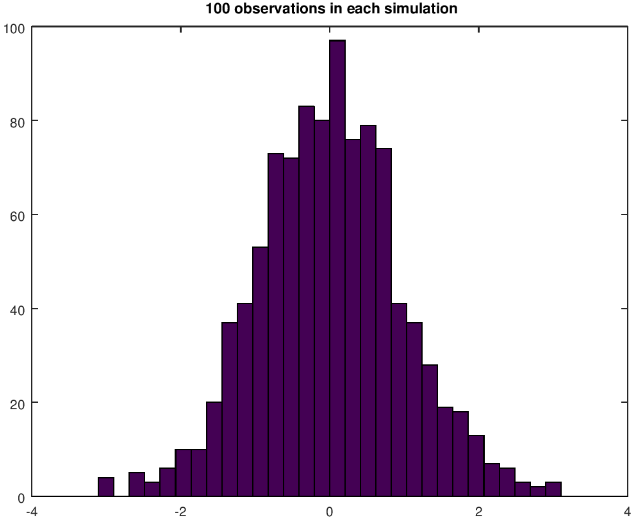
Using matlab to do simulation of convergence in distribution

```
function con_dist(n, no_sims)
  X = randn(n,no_sims);
  bins = linspace(-3,3,30);
  m = sqrt(n)*mean(X);
  hist(m,bins);
endfunction;
```

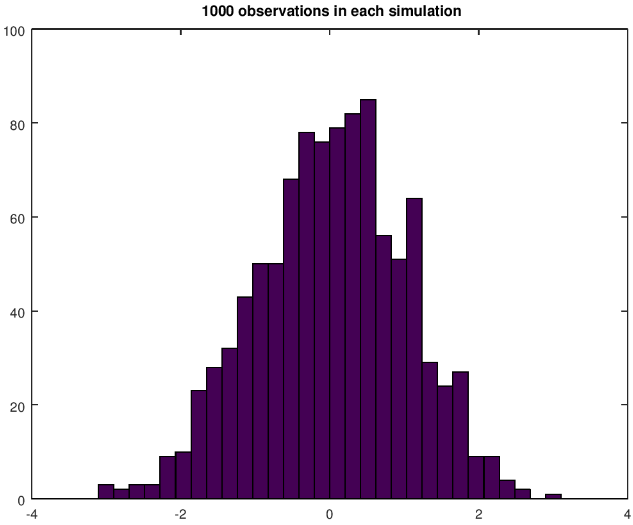
$T = 10$



$T = 100$



$T = 1000$



Convergence to a distribution

We have just illustrated the concept of

- Convergence in distribution

Used to justify probability statements: In large samples, a sample mean multiplied with the square root of the number of observations will not converge to a constant, it will have a distribution.
This is what one uses to evaluate the quality of an estimate: How sure are we of it – Quantified as a probability using this probability distribution.