

# Economics 103 – Statistics for Economists

Francis J. DiTraglia

University of Pennsylvania

Lecture # 13

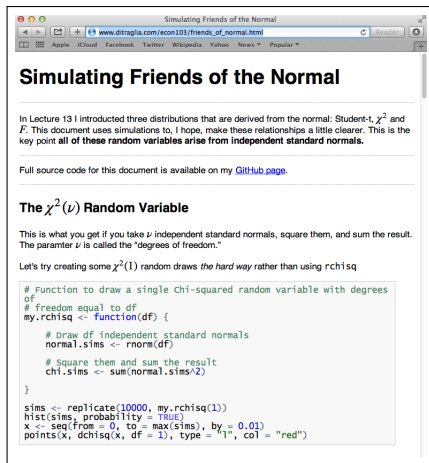
# Continuous RVs – Part III



Figure: The Normal Distribution and Friends.

[http://ditraglia.com/Econ103Public/Rtutorials/friends\\_of\\_normal.html](http://ditraglia.com/Econ103Public/Rtutorials/friends_of_normal.html)

Source Code on my [Github Page](#)



Simulating Friends of the Normal

In Lecture 13 I introduced three distributions that are derived from the normal: Student-t,  $\chi^2$  and  $F$ . This document uses simulations to, I hope, make these relationships a little clearer. This is the key point **all of these random variables arise from independent standard normals**.

Full source code for this document is available on my [Github page](#).

### The $\chi^2(\nu)$ Random Variable

This is what you get if you take  $\nu$  independent standard normals, square them, and sum the result. The parameter  $\nu$  is called the "degrees of freedom."

Let's try creating some  $\chi^2(1)$  random draws *the hard way* rather than using `rchisq`

```
# Function to draw a single Chi-squared random variable with degrees
# of
# freedom equal to df
my.rchisq <- function(df) {
  # Draw df independent standard normals
  normal.sims <- rnorm(df)

  # Square them and sum the result
  chi.sims <- sum(normal.sims^2)
}

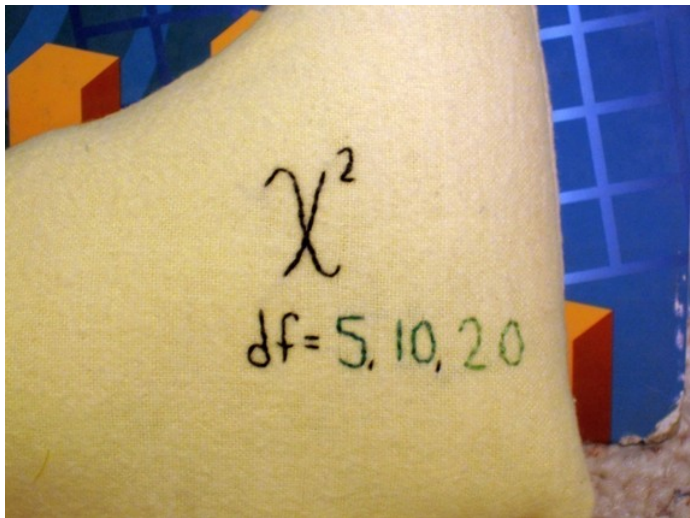
sims <- replicate(10000, my.rchisq(1))
hist(sims, probability = TRUE)
x <- seq(from = 0, to = max(sims), by = 0.01)
points(x, dchisq(x, df = 1), type = "l", col = "red")
```

## Functions of Independent RVs are Independent

If  $X$  and  $Y$  are independent random variables and  $g$  and  $h$  are functions, then the random variables  $g(X)$  and  $h(Y)$  are also independent.



Figure: PDF for  $\chi^2$ -Distribution



## $\chi^2$ Random Variable

Let  $X_1, \dots, X_\nu \sim \text{iid } N(0, 1)$ . Then,

$$(X_1^2 + \dots + X_\nu^2) \sim \chi^2(\nu)$$

where the parameter  $\nu$  is the *degrees of freedom*



## $\chi^2$ Random Variable

Let  $X_1, \dots, X_\nu \sim \text{iid } N(0, 1)$ . Then,

$$(X_1^2 + \dots + X_\nu^2) \sim \chi^2(\nu)$$

where the parameter  $\nu$  is the *degrees of freedom*

Support =  $(0, \infty)$

# $\chi^2$ PDFs

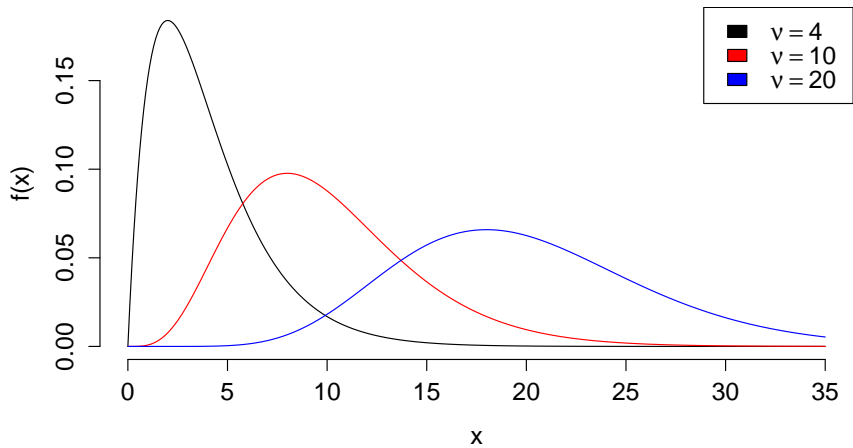
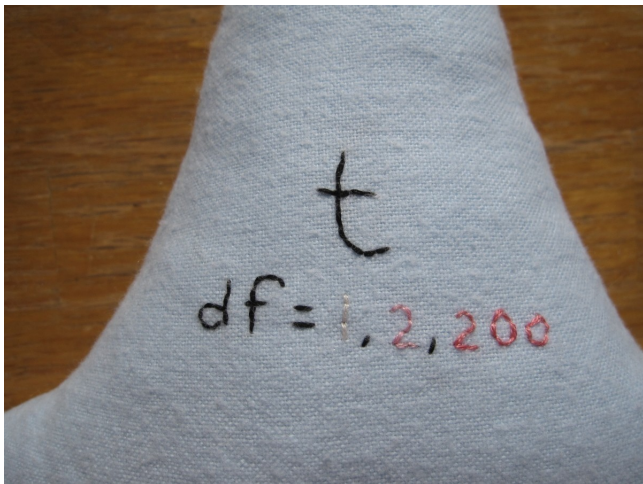




Figure: PDF for Student-t Distribution



## Student-t Random Variable

Let  $X \sim N(0, 1)$  independent of  $Y \sim \chi^2(\nu)$ . Then,

$$\frac{X}{\sqrt{Y/\nu}} \sim t(\nu)$$

where the parameter  $\nu$  is the degrees of freedom.

# Student-t Random Variable

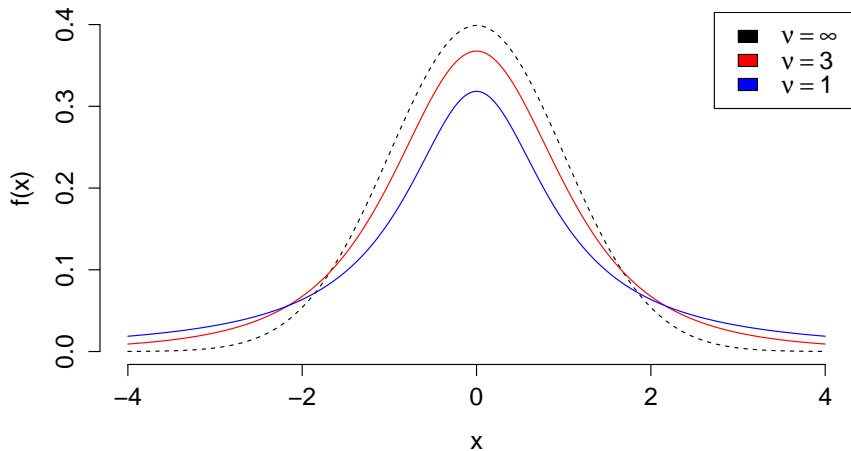
Let  $X \sim N(0, 1)$  independent of  $Y \sim \chi^2(\nu)$ . Then,

$$\frac{X}{\sqrt{Y/\nu}} \sim t(\nu)$$

where the parameter  $\nu$  is the degrees of freedom.

- ▶ Support =  $(-\infty, \infty)$
- ▶ As  $\nu \rightarrow \infty$ ,  $t \rightarrow$  Standard Normal.
- ▶ Symmetric around zero, but mean and variance may not exist!
- ▶ Degrees of freedom  $\nu$  control “thickness of tails”

# Student-t PDFs



## F Random Variable

Suppose  $X \sim \chi^2(\nu)$  independent of  $Y \sim \chi^2(\omega)$ . Then,

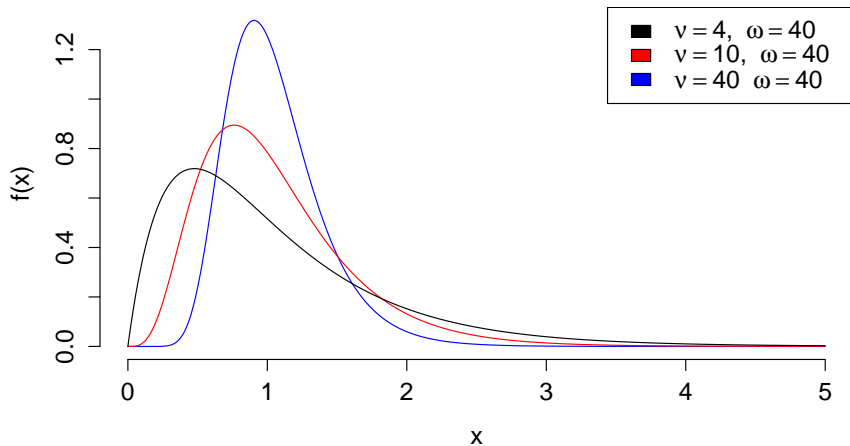
$$\frac{X/\nu}{Y/\omega} \sim F(\nu, \omega)$$

where  $\nu$  is the numerator degrees of freedom and  $\omega$  is the denominator degrees of freedom.

Support =  $(0, \infty)$



## F PDFs



## R Commands – CDFs and Quantile Functions

$F(x) = P(X \leq x)$  is the CDF,  $Q(p) = F^{-1}(p)$  the Quantile Function

	$F(x)$	$Q(p)$
$N(\mu, \sigma^2)$	<code>pnorm(x, mean = <math>\mu</math>, sd = <math>\sigma</math>)</code>	<code>qnorm(p, mean = <math>\mu</math>, sd = <math>\sigma</math>)</code>
$\chi^2(\nu)$	<code>pchisq(x, df = <math>\nu</math>)</code>	<code>qchisq(p, df = <math>\nu</math>)</code>
$t(\nu)$	<code>pt(x, df = <math>\nu</math>)</code>	<code>qt(p, df = <math>\nu</math>)</code>
$F(\nu, \omega)$	<code>pf(x, df1 = <math>\nu</math>, df2 = <math>\omega</math>)</code>	<code>qf(p, df1 = <math>\nu</math>, df2 = <math>\omega</math>)</code>

Mnemonic: “p” is for Probability, “q” is for Quantile.

## R Commands – PDFs and Random Draws

	$f(x)$	Make n iid Random Draws
$N(\mu, \sigma^2)$	<code>dnorm(x, mean = <math>\mu</math>, sd = <math>\sigma</math>)</code>	<code>rnorm(n, mean = <math>\mu</math>, sd = <math>\sigma</math>)</code>
$\chi^2(\nu)$	<code>dchisq(x, df = <math>\nu</math>)</code>	<code>rchisq(n, df = <math>\nu</math>)</code>
$t(\nu)$	<code>dt(x, df = <math>\nu</math>)</code>	<code>rt(n, df = <math>\nu</math>)</code>
$F(\nu, \omega)$	<code>df(x, df1 = <math>\nu</math>, df2 = <math>\omega</math>)</code>	<code>rf(n, df1 = <math>\nu</math>, df2 = <math>\omega</math>)</code>

Mnemonic: “d” is for Density, “r” is for Random.

Example:  $X_1, X_2, X_3 \sim \text{iid } N(0, 1)$

What is the distribution of  $Y_1 = X_1^2 + X_2^2$ ?

Example:  $X_1, X_2, X_3 \sim \text{iid } N(0, 1)$

What is the distribution of  $Y_1 = X_1^2 + X_2^2$ ?

Sum of squares of two indep. std. normals  $\Rightarrow Y_1 \sim \chi^2(2)$

Example:  $X_1, X_2, X_3 \sim \text{iid } N(0, 1)$

What is the distribution of  $Y_1 = X_1^2 + X_2^2$ ?

Sum of squares of two indep. std. normals  $\Rightarrow Y_1 \sim \chi^2(2)$

What is the distribution of  $Y_2 = (Y_1/2)/(X_3^2)$ ?

Example:  $X_1, X_2, X_3 \sim \text{iid } N(0, 1)$

What is the distribution of  $Y_1 = X_1^2 + X_2^2$ ?

Sum of squares of two indep. std. normals  $\Rightarrow Y_1 \sim \chi^2(2)$

What is the distribution of  $Y_2 = (Y_1/2)/(X_3^2)$ ?

$Y_1 \sim \chi^2(2)$  and  $X_3^2 \sim \chi^2(1)$

Example:  $X_1, X_2, X_3 \sim \text{iid } N(0, 1)$

What is the distribution of  $Y_1 = X_1^2 + X_2^2$ ?

Sum of squares of two indep. std. normals  $\Rightarrow Y_1 \sim \chi^2(2)$

What is the distribution of  $Y_2 = (Y_1/2)/(X_3^2)$ ?

$Y_1 \sim \chi^2(2)$  and  $X_3^2 \sim \chi^2(1)$

Hence  $Y_2 = \text{ratio of two indep. } \chi^2 \text{ RVs, each divided by its degrees of freedom} \Rightarrow Y_2 \sim F(2, 1)$



Example:  $X_1, X_2, X_3 \sim \text{iid } N(0, 1)$

What is the distribution of  $Y_1 = X_1^2 + X_2^2$ ?

Sum of squares of two indep. std. normals  $\Rightarrow Y_1 \sim \chi^2(2)$

What is the distribution of  $Y_2 = (Y_1/2)/(X_3^2)$ ?

$Y_1 \sim \chi^2(2)$  and  $X_3^2 \sim \chi^2(1)$

Hence  $Y_2 = \text{ratio of two indep. } \chi^2 \text{ RVs, each divided by its degrees of freedom} \Rightarrow Y_2 \sim F(2, 1)$

What is the distribution of  $Z = X_3/\sqrt{Y_1/2}$ ?

Example:  $X_1, X_2, X_3 \sim \text{iid } N(0, 1)$

What is the distribution of  $Y_1 = X_1^2 + X_2^2$ ?

Sum of squares of two indep. std. normals  $\Rightarrow Y_1 \sim \chi^2(2)$

What is the distribution of  $Y_2 = (Y_1/2)/(X_3^2)$ ?

$Y_1 \sim \chi^2(2)$  and  $X_3^2 \sim \chi^2(1)$

Hence  $Y_2 =$  ratio of two indep.  $\chi^2$  RVs, each divided by its degrees of freedom  $\Rightarrow Y_2 \sim F(2, 1)$

What is the distribution of  $Z = X_3/\sqrt{Y_1/2}$ ?

Ratio of standard normal and square root of independent  $\chi^2$  RV divided by its degrees of freedom  $\Rightarrow Z \sim t(2)$

Suppose  $X_1, X_2, \sim \text{iid } N(\mu, \sigma^2)$



Let  $Y = (X_1 - \mu)^2 + (X_2 - \mu)^2$ . What is the distribution of  $Y/\sigma^2$ ?

- (a)  $F(2, 1)$
- (b)  $\chi^2(2)$
- (c)  $t(2)$
- (d)  $N(\mu, \sigma)$
- (e) None of the above

$$Y_1 \sim \chi^2(2), \quad Y_2 \sim F(2, 1), \quad Z \sim t(2)$$

What is the median of  $Y_1$ ?

$$Y_1 \sim \chi^2(2), \quad Y_2 \sim F(2, 1), \quad Z \sim t(2)$$

What is the median of  $Y_1$ ?

`qchisq(0.5, df = 2)`  $\approx 1.4$

$$Y_1 \sim \chi^2(2), \quad Y_2 \sim F(2, 1), \quad Z \sim t(2)$$

What is the median of  $Y_1$ ?

`qchisq(0.5, df = 2)`  $\approx 1.4$

What is  $P(Y_2 \leq 5)$ ?

$$Y_1 \sim \chi^2(2), \quad Y_2 \sim F(2, 1), \quad Z \sim t(2)$$

What is the median of  $Y_1$ ?

`qchisq(0.5, df = 2)`  $\approx 1.4$

What is  $P(Y_2 \leq 5)$ ?

`pf(5, df1 = 2, df2 = 1)`  $\approx 0.7$

$$Y_1 \sim \chi^2(2), \quad Y_2 \sim F(2, 1), \quad Z \sim t(2)$$

What is the median of  $Y_1$ ?

`qchisq(0.5, df = 2)`  $\approx 1.4$

What is  $P(Y_2 \leq 5)$ ?

`pf(5, df1 = 2, df2 = 1)`  $\approx 0.7$

What value of  $c$  gives  $P(-c \leq Z \leq c) = 0.5$ ?

Use Symmetry (like normal)

`c = qt(0.75, df = 2)`  $\approx 0.8$



$$Y_1 \sim \chi^2(2), \quad Y_2 \sim F(2, 1), \quad Z \sim t(2)$$

What is the median of  $Y_1$ ?

$$\text{qchisq}(0.5, \text{df} = 2) \approx 1.4$$

What is  $P(Y_2 \leq 5)$ ?

$$\text{pf}(5, \text{df1} = 2, \text{df2} = 1) \approx 0.7$$

What value of  $c$  gives  $P(-c \leq Z \leq c) = 0.5$ ?

Use Symmetry (like normal)

$$c = \text{qt}(0.75, \text{df} = 2) \approx 0.8$$

$$\text{or equivalently } -c = \text{qt}(0.25, \text{df} = 2) \approx -0.8$$