STA261 Lecture 2 — 2017-07-10

Neil Montgomery

Last edited: 2017-07-10 18:42

t distributions

Theorem 1.6 Suppose $Z \sim N(0,1)$ and $U \sim \chi^2_\nu$. Define $T = Z/\sqrt{U/\nu}$. The density of T is:

$$f_T(t) = \frac{\Gamma[(\nu+1)/2]}{\sqrt{\nu\pi}\,\Gamma(\nu/2)}\left(1+\frac{t^2}{\nu}\right)^{-(\nu+1)/2}$$

t distributions

Theorem 1.6 Suppose $Z \sim N(0,1)$ and $U \sim \chi^2_\nu$. Define $T = Z/\sqrt{U/\nu}$. The density of T is:

$$f_T(t) = \frac{\Gamma[(\nu+1)/2]}{\sqrt{\nu\pi}\,\Gamma(\nu/2)} \left(1 + \frac{t^2}{\nu}\right)^{-(\nu+1)/2}$$

When T has this density we say "T has a t distribution with ν *degrees of freedom*" or $T \sim t_\nu$ for short. (The degrees of freedom parameter is "inherited" from the sum of squares in the denominator.)

t distributions

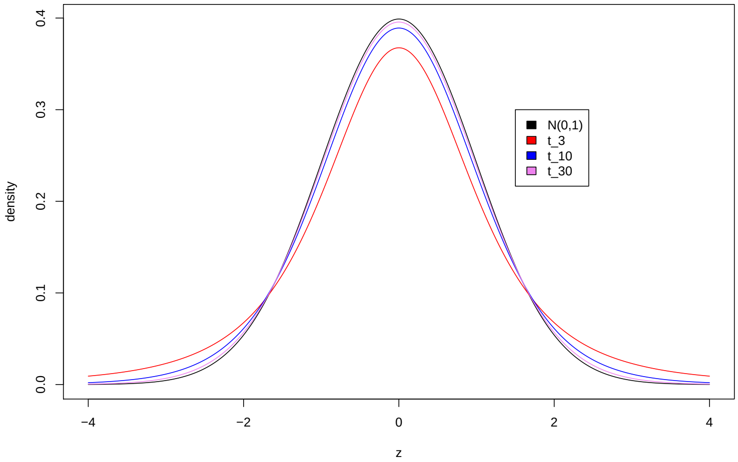
Theorem 1.6 Suppose $Z \sim N(0,1)$ and $U \sim \chi^2_\nu$. Define $T = Z/\sqrt{U/\nu}$. The density of T is:

$$f_T(t) = \frac{\Gamma[(\nu+1)/2]}{\sqrt{\nu\pi}\,\Gamma(\nu/2)}\left(1+\frac{t^2}{\nu}\right)^{-(\nu+1)/2}$$

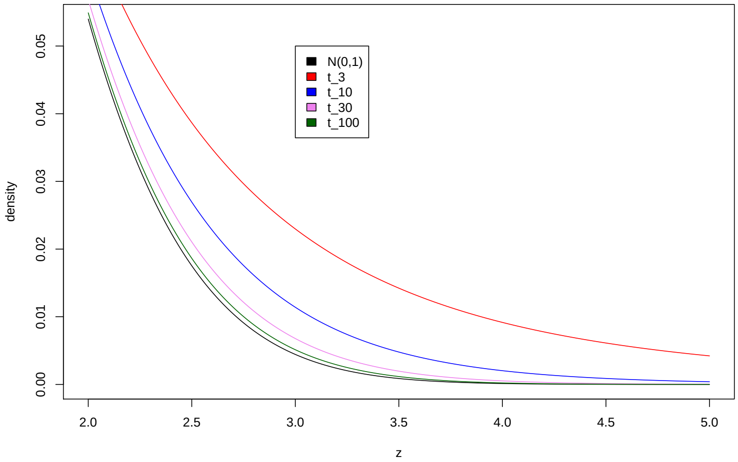
When T has this density we say "T has a t distribution with ν *degrees of freedom*" or $T \sim t_\nu$ for short. (The degrees of freedom parameter is "inherited" from the sum of squares in the denominator.)

In the theorem, there was nothing about ν being an integer. It only has to be bigger than 0.

pictures of t densities (overall)



detail of the tails



properties of t_ν

The density is:

$$f_T(t) = \frac{\Gamma[(\nu+1)/2]}{\sqrt{\nu\pi}\,\Gamma(\nu/2)}\left(1+\frac{t^2}{\nu}\right)^{-(\nu+1)/2}$$

Note 2.0: T distributions are symmetric around 0.

Proposition 2.1: If $T \sim t_\nu$ and $\nu > 1$, then $E(T) = 0$, and if $0 < \nu \leqslant 1$ $E(T)$ does not exist.

Proposition 2.2: $\mathrm{Var}(T) = \frac{\nu}{\nu-2}$ when $\nu > 2$ and does not exist when $0 < \nu \leqslant 2$.

recall “convergence in distribution”

A sequence $X_1, X_2, X_3, \ldots$ is said to *converge in distribution* to X (notation: $X_n \Rightarrow X$)

$$\lim_{n \to \infty} F_{X_n}(x) = F_X(x)$$

for all x where F_X is continuous.

recall "convergence in distribution"

A sequence $X_1, X_2, X_3, \ldots$ is said to *converge in distribution* to X (notation: $X_n \Rightarrow X$)

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$$

for all x where F_X is continuous.

This is the most useful kind of convergence in statistical applications.

It is the mode of convergence in the Central Limit Theorem (proved in STA257 using moment generating functions.)

recall "convergence in distribution"

A sequence $X_1, X_2, X_3, \dots$ is said to *converge in distribution* to X (notation: $X_n \Rightarrow X$)

$$\lim_{n \to \infty} F_{X_n}(x) = F_X(x)$$

for all x where F_X is continuous.

This is the most useful kind of convergence in statistical applications.

It is the mode of convergence in the Central Limit Theorem (proved in STA257 using moment generating functions.)

Fact 2.3: If continuous random variables $X_1, X_2, X_3, \dots$ have densities $f_1(x), f_2(x), \dots$ which converge to a density $f(x)$ for X, at all x, then $X_n \Rightarrow X$

properties of t_ν

Proposition 2.4: If $T_n \sim t_n$, then $T_n \Rightarrow Z$ where $Z \sim N(0,1)$.

properties of t_ν

Proposition 2.4: If $T_n \sim t_n$, then $T_n \Rightarrow Z$ where $Z \sim N(0,1)$.

I'll only sketch the demonstration starting with:

$$f_T(t) = \frac{\Gamma[(\nu+1)/2]}{\sqrt{\nu\pi}\,\Gamma(\nu/2)}\left(1+\frac{t^2}{\nu}\right)^{-(\nu+1)/2}$$

properties of t_ν

Proposition 2.4: If $T_n \sim t_n$, then $T_n \Rightarrow Z$ where $Z \sim N(0,1)$.

I'll only sketch the demonstration starting with:

$$f_T(t) = \frac{\Gamma[(\nu+1)/2]}{\sqrt{\nu\pi}\,\Gamma(\nu/2)} \left(1 + \frac{t^2}{\nu}\right)^{-(\nu+1)/2}$$

The interested student can play with "Stirling's formula" for $\Gamma(x)$. . . .

properties of t_ν

Proposition 2.4: If $T_n \sim t_n$, then $T_n \Rightarrow Z$ where $Z \sim N(0,1)$.

I'll only sketch the demonstration starting with:

$$f_T(t) = \frac{\Gamma[(\nu+1)/2]}{\sqrt{\nu\pi}\,\Gamma(\nu/2)}\left(1+\frac{t^2}{\nu}\right)^{-(\nu+1)/2}$$

The interested student can play with "Stirling's formula" for $\Gamma(x)$. . . .

Note that the convergence is *really fast.*

F distributions

Proposition 2.5: Suppose U and V are independent with $U \sim \chi_n^2$ and $V \sim \chi_m^2$. Then the density of $F = \frac{U/n}{V/m}$ is:

$$f_F(x) = \frac{\Gamma\left(\frac{n+m}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{m}{2}\right)}\left(\frac{m}{n}\right)^{n/2}\frac{x^{n/2-1}}{\left(1+(n/m)\,x\right)^{(n+m)/2}}, \quad x > 0$$

F distributions

Proposition 2.5: Suppose U and V are independent with $U \sim \chi^2_n$ and $V \sim \chi^2_m$. Then the density of $F = \frac{U/n}{V/m}$ is:

$$f_F(x) = \frac{\Gamma\left(\frac{n+m}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{m}{2}\right)} \left(\frac{m}{n}\right)^{n/2} \frac{x^{n/2-1}}{\left(1 + (n/m)\, x\right)^{(n+m)/2}}, \quad x > 0$$

This uses **Factoid 2.6:** if X has density $f_X(x)$, then the density of X/a for $a > 0$ is $af_X(ax)$.

F distributions

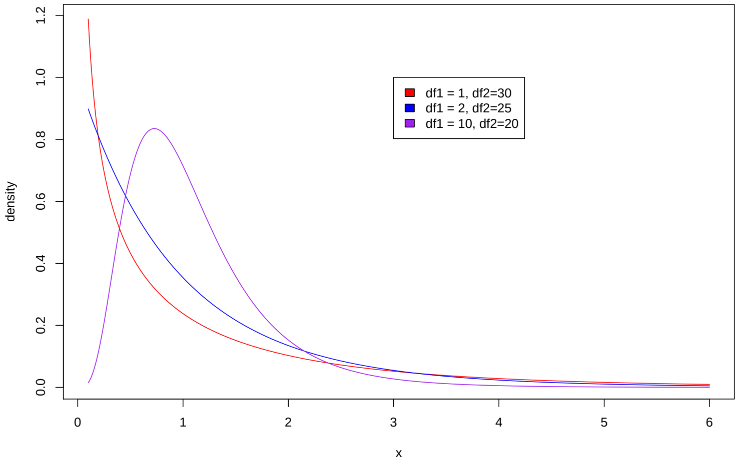
Proposition 2.5: Suppose U and V are independent with $U \sim \chi^2_n$ and $V \sim \chi^2_m$. Then the density of $F = \frac{U/n}{V/m}$ is:

$$f_F(x) = \frac{\Gamma\left(\frac{n+m}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{m}{2}\right)} \left(\frac{m}{n}\right)^{n/2} \frac{x^{n/2-1}}{(1+(n/m)\,x)^{(n+m)/2}}, \quad x > 0$$

This uses **Factoid 2.6:** if X has density $f_X(x)$, then the density of X/a for $a > 0$ is $af_X(ax)$.

In words: "...has an F distribution with n and m degrees of freedom" or $F_{n,m}$.

pictures of F distributions



the square of a t_ν distribution

Proposition 2.7: If $T \sim t_\nu$, then $T^2 \sim F_{1,\nu}$

sample mean and sample variance from a $N(\mu, \sigma^2)$ sample

$\overline{X}$ and S^2

Suppose $X_1, \ldots, X_n$ is a *sample* from a population $X \sim N(\mu, \sigma^2)$. Define the *sample mean*:

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n},$$

and the *sample variance*:

$$S^2 = \sum_{i=1}^{n} \frac{\left(X_i - \overline{X}\right)^2}{n-1}.$$

distribution of $\overline{X}$

From last class $\overline{X} \sim N(\mu, \sigma^2/n)$.

This fact needs to flow in your blood.

the independence of $\overline{X}$ and S^2 when $n = 2$

An example from last class showed that when X_1 and X_2 are i.i.d. $N(0, 1)$, then $X_1 + X_2$ and $X_1 - X_2$ are independent.

the independence of $\overline{X}$ and S^2 when $n = 2$

An example from last class showed that when X_1 and X_2 are i.i.d. $N(0, 1)$, then $X_1 + X_2$ and $X_1 - X_2$ are independent.

There was nothing special about $N(0, 1)$. It could have been any $N(\mu, \sigma^2)$ (just more tedious to show.)

the independence of $\overline{X}$ and S^2 when $n = 2$

An example from last class showed that when X_1 and X_2 are i.i.d. $N(0, 1)$, then $X_1 + X_2$ and $X_1 - X_2$ are independent.

There was nothing special about $N(0, 1)$. It could have been any $N(\mu, \sigma^2)$ (just more tedious to show.)

Theorem 2.8: When X_1 and X_2 are i.i.d. $N(\mu, \sigma^2)$, $\overline{X}$ and S^2 are indepedent.

the independence of $\overline{X}$ and S^2 in general

Proposition 2.9: S^2 can be expressed in terms of $\left(X_2 - \overline{X}, \ldots, X_n - \overline{X}\right)$, using the fact that $\sum_{i=1}^{n}(X_i - \overline{X}) = 0$.

Proposition 2.10: When $X_1, \ldots, X_n$ are i.i.d. $N(\mu, \sigma^2)$, $\overline{X} \perp S^2$

the distribution of S^2

Proposition 2.11: If $X_1, \ldots, X_n$ are i.i.d. $N(\mu, \sigma)$, then $\frac{n-1}{\sigma^2} S^2 \sim \chi^2_{n-1}$.

Proposition 2.12: $E(S^2) = \sigma^2$ and $\text{Var}(S^2) = \frac{2(\sigma^2)^2}{n-1}$

what's the point of all this?

Revisit my desire to know the average height of students in the class.

The distribution of heights might be $N(\mu, \sigma^2)$. I'm interested in μ.I am going to sample n people from the list, to obtain a sample $X_1, \ldots, X_n$.

what's the point of all this?

Revisit my desire to know the average height of students in the class.

The distribution of heights might be $N(\mu, \sigma^2)$. I'm interested in μ.I am going to sample n people from the list, to obtain a sample $X_1, \ldots, X_n$.

It seems reasonable to use $\overline{X}$ as a guess for the average (an *estimator for the mean* μ.) How close might I be to the true value?

what's the point of all this?

Revisit my desire to know the average height of students in the class.

The distribution of heights might be $N(\mu, \sigma^2)$. I'm interested in μ.I am going to sample n people from the list, to obtain a sample $X_1, \ldots, X_n$.

It seems reasonable to use $\overline{X}$ as a guess for the average (an *estimator for the mean* μ.) How close might I be to the true value?

We could make probability calculations involving $\overline{X} \sim N(\mu, \sigma^2)$, or equivalently $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0, 1)$.

what's the point of all this?

Revisit my desire to know the average height of students in the class.

The distribution of heights might be $N(\mu, \sigma^2)$. I'm interested in μ.I am going to sample n people from the list, to obtain a sample $X_1, \dots, X_n$.

It seems reasonable to use $\overline{X}$ as a guess for the average (an *estimator for the mean* μ.) How close might I be to the true value?

We could make probability calculations involving $\overline{X} \sim N(\mu, \sigma^2)$, or equivalently $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$.

But we don't know σ. What do we do?

the distribution of $\frac{\overline{X}-\mu}{s/\sqrt{n}}$

Proposition 2.13: When $X_1, \ldots, X_n$ are i.i.d. $N(\mu, \sigma^2)$,

$$\frac{\overline{X} - \mu}{s/\sqrt{n}} \sim t_{n-1}$$