Math 362: Mathmatical Statistics II

Le Chen
le.chen@emory.edu

Emory University
Atlanta, GA

Last updated on February 25, 2020

2020 Spring

Chapter 7. Inference Based on The Normal Distribution

§ 7.1 Introduction

§ 7.2 Comparing $\frac{\overline{Y}-\mu}{\sigma/\sqrt{n}}$ and $\frac{\overline{Y}-\mu}{S/\sqrt{n}}$

§ 7.3 Deriving the Distribution of $\frac{\overline{Y}-\mu}{S/\sqrt{n}}$

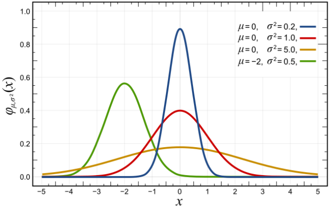
§ 7.1 Introduction

Carl Friedrich Gauss discovered the normal distribution in 1809 as a way to rationalize the method of least squares.

(1777-1855)

Marquis de Laplace proved the central limit theorem in 1810, consolidating the importance of the normal distribution in statistics.

(1749-1827)



Notation	$\mathcal{N}(\mu, \sigma^2)$
Parameters	$\mu \in \mathbb{R}$ = mean (location) $\sigma^2 > 0$ = variance (squared scale)
Support	$x \in \mathbb{R}$
PDF	$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$
CDF	$\frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right]$
Quantile	$\mu + \sigma\sqrt{2}\operatorname{erf}^{-1}(2p-1)$
Mean	μ
Median	μ
Mode	μ
Variance	σ^2
MAD	$\sigma\sqrt{2/\pi}$
Skewness	0
Ex. kurtosis	0
Entropy	$\frac{1}{2}\log(2\pi e\sigma^2)$
MGF	$\exp(\mu t + \sigma^2 t^2/2)$
CF	$\exp(i\mu t - \sigma^2 t^2/2)$
Fisher information	$\mathcal{I}(\mu,\sigma) = \begin{pmatrix} 1/\sigma^2 & 0 \\ 0 & 2/\sigma^2 \end{pmatrix}$ $\mathcal{I}(\mu,\sigma^2) = \begin{pmatrix} 1/\sigma^2 & 0 \\ 0 & 1/(2\sigma^4) \end{pmatrix}$
Kullback-Leibler divergence	$D_{\mathrm{KL}}(\mathcal{N}_0 \parallel \mathcal{N}_1) = \frac{1}{2}\left\{\left(\frac{\sigma_0}{\sigma_1}\right)^2 + \frac{(\mu_1-\mu_0)^2}{\sigma_1^2} - 1 + 2\ln\frac{\sigma_1}{\sigma_0}\right\}$

https://en.wikipedia.org/wiki/Normal_distribution

Test for normal parameters (one sample test)

Let $Y_1, \cdots, Y_n$ be a random sample from $N(\mu, \sigma^2)$.

Prob. 1 Find a test statistic Λ in order to test $H_0 : \mu = \mu_0$ v.s. $H_1 : \mu \neq \mu_0$.

When σ^2 is known: $$\Lambda = \frac{\overline{Y} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$$

When σ^2 is unknown: $\Lambda = ?$ $$\Lambda \overset{?}{=} \frac{\overline{Y} - \mu_0}{s/\sqrt{n}} \sim ?$$

Prob. 2 Find a test statistic Λ in order to test $H_0 : \sigma^2 = \sigma_0^2$ v.s. $H_1 : \sigma^2 \neq \sigma_0^2$.

Prob. 1 Find a test statistic for $H_0 : \mu = \mu_0$ v.s. $H_1 : \mu \neq \mu_0$, with σ^2 unknown

Sol. Composite-vs-composite test with:

$$\omega = \left\{(\mu, \sigma^2) : \mu = \mu_0,\ \sigma^2 > 0\right\}$$

$$\Omega = \left\{(\mu, \sigma^2) : \mu \in \mathbb{R},\ \sigma^2 > 0\right\}$$

The MLE under the two spaces are:

$$\omega_e = (\mu_e, \sigma_e^2) : \qquad \mu_e = \mu_0 \quad \text{and} \quad \sigma_e^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \mu_0)^2 \qquad \text{(Under } \omega\text{)}$$

$$\Omega_e = (\mu_e, \sigma_e^2) : \qquad \mu_e = \bar{y} \quad \text{and} \quad \sigma_e^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2 \qquad \text{(Under } \Omega\text{)}$$

$$L(\mu, \sigma^2) = (2\pi\sigma^2)^{-n} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{y_i - \mu}{\sigma}\right)^2\right)$$

$$L(\omega_e) = \cdots = \left[\frac{ne^{-1}}{2\pi\sum_{i=1}^{n}(y_i - \mu_0)^2}\right]^{n/2}$$

$$L(\Omega_e) = \cdots = \left[\frac{ne^{-1}}{2\pi\sum_{i=1}^{n}(y_i - \bar{y})^2}\right]^{n/2}$$

Hence

$$\lambda = \frac{L(\omega_e)}{L(\Omega_e)} = \left[\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \mu_0)^2}\right]^{n/2} = \cdots = \left[1 + \frac{n(\bar{y} - \mu_0)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}\right]^{-n/2}$$

$$= \left[1 + \frac{1}{n-1}\left(\frac{\bar{y} - \mu_0}{\sqrt{\frac{1}{n-1}\sum_{i=1}^n (y_i - \bar{y})^2} \Big/ \sqrt{n}}\right)^2\right]^{-n/2}$$

$$= \left[1 + \frac{1}{n-1}\left(\frac{\bar{y} - \mu_0}{s \,/\, \sqrt{n}}\right)^2\right]^{-n/2}$$

$$= \left[1 + \frac{t^2}{n-1}\right]^{-n/2}, \quad t = \frac{\bar{y} - \mu_0}{s \,/\, \sqrt{n}}$$

Finally, the test statistic is

$$\boxed{T = \frac{\overline{Y} - \mu_0}{S/\sqrt{n}}}$$

with $\quad \overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \quad$ and $\quad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)^2.$

Question: Find the exact distribution of T.

Prob. 2 Find a test statistic for $H_0 : \sigma^2 = \sigma_0^2$ v.s. $H_1 : \sigma^2 \neq \sigma_0^2$, with μ unknown

Sol. Composite-vs-composite test with:

$$\omega = \left\{(\mu, \sigma^2) : \mu \in \mathbb{R},\ \sigma^2 = \sigma_0^2\right\}$$
$$\Omega = \left\{(\mu, \sigma^2) : \mu \in \mathbb{R},\ \sigma^2 > 0\right\}$$

The MLE under the two spaces are:

$$\omega_e = (\mu_e, \sigma_e^2) : \qquad \mu_e = \bar{y} \quad \text{and} \quad \sigma_e^2 = \sigma_0^2 \qquad \text{(Under } \omega\text{)}$$

$$\Omega_e = (\mu_e, \sigma_e^2) : \qquad \mu_e = \bar{y} \quad \text{and} \quad \sigma_e^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2 \qquad \text{(Under } \Omega\text{)}$$

$$L(\mu, \sigma^2) = (2\pi\sigma^2)^{-n} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{y_i - \mu}{\sigma}\right)^2\right)$$

$$L(\omega_e) = (2\pi\sigma^2)^{-n} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{y_i - \bar{y}}{\sigma_0}\right)^2\right)$$

$$L(\Omega_e) = \cdots = \left[\frac{ne^{-1}}{2\pi\sum_{i=1}^{n}(y_i - \bar{y})^2}\right]^{n/2}$$

Hence

$$\lambda = \frac{L(\omega_e)}{L(\Omega_e)} = \left[\frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n\sigma_0^2}\right]^{n/2} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{y_i - \bar{y}}{\sigma_0}\right)^2 + \frac{n}{2}\right)$$

$$= \left[\frac{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2}{\frac{n}{n-1}\sigma_0^2}\right]^{n/2} \exp\left(-\frac{n-1}{2\sigma_0^2}\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2 + \frac{n}{2}\right)$$

$$= \left[\frac{s^2}{\frac{n}{n-1}\sigma_0^2}\right]^{n/2} \exp\left(-\frac{n-1}{2\sigma_0^2}s^2 + \frac{n}{2}\right)$$

Finally, the test statistic is

$$\boxed{S^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(Y_i - \overline{Y} \right)^2} \quad \text{with} \quad \overline{Y} = \frac{1}{n} \sum_{i=1}^{n} Y_i$$

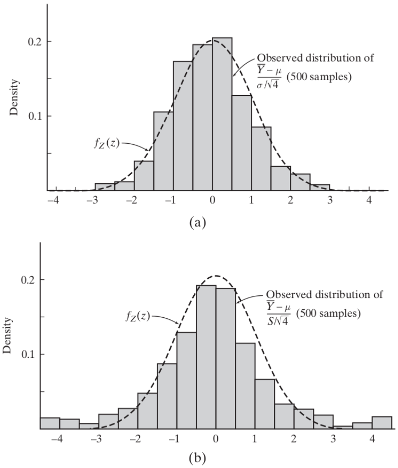
Question: Find the exact distribution of S^2.

§ 7.1 Introduction

§ 7.2 Comparing $\frac{\overline{Y}-\mu}{\sigma/\sqrt{n}}$ and $\frac{\overline{Y}-\mu}{S/\sqrt{n}}$

§ 7.3 Deriving the Distribution of $\frac{\overline{Y}-\mu}{S/\sqrt{n}}$

§ 7.2 Comparing $\frac{\overline{Y}-\mu}{\sigma/\sqrt{n}}$ and $\frac{\overline{Y}-\mu}{S/\sqrt{n}}$



Statistician William Sealy Gosset, known as "Student"

Ref. Student's t distribution comes from William Sealy Gosset's 1908 paper in Biometrika under the pseudonym "Student".

Gosset worked at the Guinness Brewery in Dublin, Ireland, and was interested in the problems of small samples – for example, the chemical properties of barley where sample sizes might be as few as 3.

V1 One version of the origin of the pseudonym is that Gosset's employer preferred staff to use pen names when publishing scientific papers instead of their real name, so he used the name "Student" to hide his identity.

V2 Another version is that Guinness did not want their competitors to know that they were using the t-test to determine the quality of raw material

§ 7.1 Introduction

§ 7.2 Comparing $\frac{\overline{Y}-\mu}{\sigma/\sqrt{n}}$ and $\frac{\overline{Y}-\mu}{S/\sqrt{n}}$

§ 7.3 Deriving the Distribution of $\frac{\overline{Y}-\mu}{S/\sqrt{n}}$

§ 7.3 Deriving the Distribution of $\frac{\overline{Y}-\mu}{S/\sqrt{n}}$

Def. **Sampling distributions**

Distributions of $\underbrace{\textit{functions of random sample}}_{\text{statistics / estimators}}$ of given size.

E.g. A random sample of size n from $N(\mu, \sigma^2)$ with σ^2 known.

Sample mean $\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \sim N(\mu, \sigma^2/n)$

Aim: Determine distributions for

Sample variance $S^2 := \frac{1}{n-1}\sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)^2$ *Chi square distr.*

$$T := \frac{\overline{Y}-\mu}{S/\sqrt{n}}$$ *Student t distr.*

$$\frac{S_1^2}{\sigma_1^2}\Big/\frac{S_1^2}{\sigma_1^2}$$ *F distr.*

Thm 1. Let $U = \sum_{i=1}^{m} Z_j^2$, where Z_j are independent $N(0,1)$ normal r.v.s. Then

$$U \sim \text{Gamma(shape=}m/2\text{,rate=}1/2\text{)}.$$

namely,

$$f_U(u) = \frac{1}{2^{m/2}\Gamma(m/2)} u^{\frac{m}{2}-1} e^{-u/2}, \qquad u \geq 0$$

Def 1. U in Thm 1 is called **chi square distribution** with m dgs of freedom.

Proof First take $m = 1$. For any $u \geq 0$,

$$F_{Z^2}(u) = P(Z^2 \leq u) = P\left(-\sqrt{u} \leq Z \leq \sqrt{u}\right) = 2P\left(0 \leq Z \leq \sqrt{u}\right)$$

$$= \frac{2}{\sqrt{2\pi}} \int_0^{\sqrt{u}} e^{-z^2/2}\, dz$$

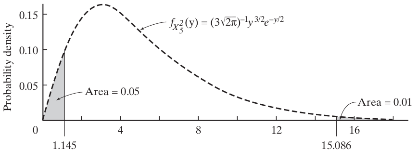
Differentiating both sides of the equation for $F_{Z^2}(u)$ gives $f_{Z^2}(u)$:

$$f_{Z^2}(u) = \frac{d}{du} F_{Z^2}(u) = \frac{2}{\sqrt{2\pi}} \frac{1}{2\sqrt{u}} e^{-u/2} = \frac{1}{2^{1/2}\Gamma\left(\frac{1}{2}\right)} u^{(1/2)-1} e^{-u/2}$$

Notice that $f_U(u) = f_{Z^2}(u)$ has the form of a gamma pdf with $r = \frac{1}{2}$ and $\lambda = \frac{1}{2}$. By Theorem 4.6.4, then, the sum of m such squares has the stated gamma distribution with $r = m\left(\frac{1}{2}\right) = \frac{m}{2}$ and $\lambda = \frac{1}{2}$. $\square$

Chi Square Table

	p							
df	.01	.025	.05	.10	.90	.95	.975	.99
1	0.000157	0.000982	0.00393	0.0158	2.706	3.841	5.024	6.635
2	0.0201	0.0506	0.103	0.211	4.605	5.991	7.378	9.210
3	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345
4	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277
5	0.554	0.831	1.145	1.610	9.236	11.070	12.832	15.086
6	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812
7	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475
8	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090
9	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666
10	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209
11	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725
12	3.571	4.404	5.226	6.304	18.549	21.026	23.336	26.217



$$\mathbb{P}(\chi_5^2 \leq 1.145) = 0.05 \iff \chi^2_{0.05,5} = 1.145$$

$$\mathbb{P}(\chi_5^2 \leq 15.086) = 0.99 \iff \chi^2_{0.99,5} = 15.086$$

```
> pchisq(1.145, df = 5)
[1] 0.04995622
> pchisq(15.086, df = 5)
[1] 0.9899989
```

```
> qchisq(0.05, df = 5)
[1] 1.145476
> qchisq(0.99, df = 5)
[1] 15.08627
```

Thm 2. Let $Y_1, \cdots, Y_n$ be a random sample from $N(\mu, \sigma^2)$. Then

(a) S^2 and $\overline{Y}$ are independent.

(b) $\dfrac{(n-1)S^2}{\sigma^2} = \dfrac{1}{\sigma^2}\sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)^2 \quad \sim \quad$ Chi Square$(n-1)$.

Proof. We will prove the case $n = 2$.

$$\overline{Y} = \frac{Y_1 + Y_2}{2}, \qquad Y_1 - \overline{Y} = \frac{Y_1 - Y_2}{2}, \qquad Y_2 - \overline{Y} = \frac{Y_2 - Y_1}{2}$$

$$S^2 = ... = \frac{1}{2}\left(Y_1 - Y_2\right)^2$$

(a) It is equivalanet to show $Y_1 + Y_2 \perp Y_1 - Y_2$. Since they are normal, it suffices to show that

$$\mathbb{E}[(Y_1 + Y_2)(Y_1 - Y_2)] = \mathbb{E}[Y_1 + Y_2]\mathbb{E}[Y_1 - Y_2]$$

(b) $\frac{(n-1)S^2}{\sigma^2} = \left(\frac{Y_1 - Y_2}{\sqrt{2}\sigma}\right)^2$ and $\frac{Y_1 - Y_2}{\sqrt{2}\sigma} \sim N(0, 1)$ ... $\square$

Def 2. If $U \sim$ Chi Square(n) and $V \sim$ Chi Square(m), and $U \perp V$, then

$$F := \frac{V/m}{U/n}$$

follows the **(Snedecor's) F distribution** with m and n degrees of freedom.

Thm 3. Let $F_{m,n} = \frac{V/m}{U/n}$ be an F r.v. with m and n degrees of freedom. Then

$$f_{F_{m,n}}(w) = \frac{\Gamma\left(\frac{m+n}{2}\right) m^{m/2} n^{n/2}}{\Gamma(m/2)\Gamma(n/2)} \times \frac{w^{m/2-1}}{(n+mw)^{(m+n)/2}}, \quad w \geq 0$$

Equivalently,

$$f_{F_{m,n}}(w) = B(m/2, n/2)^{-1} \left(\frac{m}{n}\right)^{\frac{m}{2}} w^{\frac{m}{2}-1} \left(1 + \frac{m}{n} w\right)^{-\frac{m+n}{2}}$$

where $B(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$

Recall

Theorem 3.8.4 *Let X and Y be independent continuous random variables, with pdfs $f_X(x)$ and $f_Y(y)$, respectively. Assume that X is zero for at most a set of isolated points. Let $W = Y/X$. Then*

$$f_W(w) = \int_{-\infty}^{\infty} |x|\, f_X(x)\, f_Y(wx)\, dx$$

Proof of Thm 3.

$\square$

Proof We begin by finding the pdf for V/U. From Theorem 7.3.1 we know that $f_V(v) = \frac{1}{2^{m/2}\Gamma(m/2)} v^{(m/2)-1} e^{-v/2}$ and $f_U(u) = \frac{1}{2^{n/2}\Gamma(n/2)} u^{(n/2)-1} e^{-u/2}$.

From Theorem 3.8.4, we have that the pdf of $W = V/U$ is

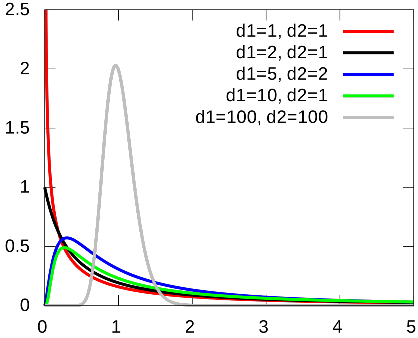
$$\begin{aligned}
f_{V/U}(w) &= \int_0^\infty |u| f_U(u) f_V(uw)\, du \\
&= \int_0^\infty u \frac{1}{2^{n/2}\Gamma(n/2)} u^{(n/2)-1} e^{-u/2} \frac{1}{2^{m/2}\Gamma(m/2)} (uw)^{(m/2)-1} e^{-uw/2}\, du \\
&= \frac{1}{2^{(n+m)/2}\Gamma(n/2)\Gamma(m/2)} w^{(m/2)-1} \int_0^\infty u^{\frac{n+m}{2}-1} e^{-[(1+w)/2]u}\, du
\end{aligned}$$

The integrand is the variable part of a gamma density with $r = (n+m)/2$ and $\lambda = (1+w)/2$. Thus, the integral equals the inverse of the density's constant. This gives

$$f_{V/U} = \frac{1}{2^{(n+m)/2}\Gamma(n/2)\Gamma(m/2)} w^{(m/2)-1} \frac{\Gamma\left(\frac{n+m}{2}\right)}{[(1+w)/2]^{\frac{n+m}{2}}} = \frac{\Gamma\left(\frac{n+m}{2}\right)}{\Gamma(n/2)\Gamma(m/2)} \frac{w^{(m/2)-1}}{(1+w)^{\frac{n+m}{2}}}$$

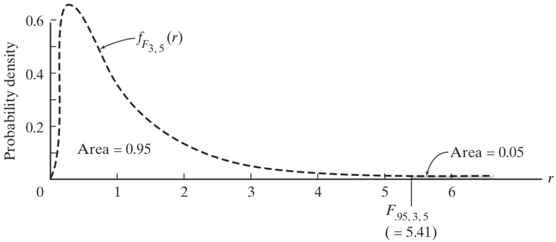
The statement of the theorem, then, follows from Theorem 3.8.2:

$$f_{\frac{V/m}{U/n}}(w) = f_{\frac{n}{m}V/U}(w) = \frac{1}{n/m} f_{V/U}\left(\frac{w}{n/m}\right) = \frac{m}{n} f_{V/U}\left(\frac{m}{n}w\right) \qquad \square$$




```
# Draw F density
x=seq(0,5,0.01)
pdf= cbind(df(x, df1 = 1, df2 = 1),
df(x, df1 = 2, df2 = 1),
df(x, df1 = 5, df2 = 2),
df(x, df1 = 10, df2 = 1),
df(x, df1 = 100, df2 = 100))
matplot(x,pdf, type = "l")
 title ("F with various dgrs of freedom")
```

F Table



$$\mathbb{P}(F_{3,5} \leq 5.41) = 0.95 \quad \Longleftrightarrow \quad F_{0.95,3,5} = 5.41$$

```
> pf(5.41,df1 = 3, df2 = 5)
[1] 0.9500093
```

```
> qf(0.95,df1 = 3, df2 = 5)
[1] 5.409451
```

Def 3. Suppose $Z \sim N(0,1)$, $U \sim$ Chi Square(n), and $Z \perp U$. Then

$$T_n = \frac{Z}{\sqrt{U/n}}$$

follows the **Student's t-distribution** of n degrees of freedom.

Remark $T_n^2 \sim F$-distribution with 1 and n degrees of freedom.

Thm 4. The pdf of the Student t of degree n is

$$f_{T_n}(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)} \times \left(1 + \frac{t^2}{n}\right)^{-\frac{n+2}{2}}, \quad t \in \mathbb{R}.$$

Proof Note that $T_n^2 = \frac{Z^2}{U/n}$ has an F distribution with 1 and n df. Therefore,

$$f_{T_n^2}(t) = \frac{n^{n/2}\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{n}{2}\right)} t^{-1/2} \frac{1}{(n+t)^{(n+1)/2}}, \quad t > 0$$

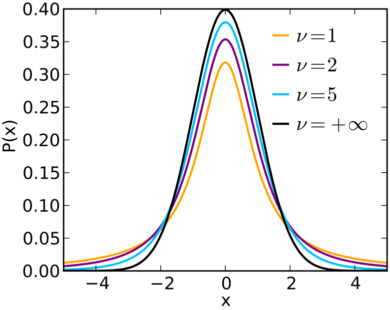
Suppose that $t > 0$. By the symmetry of $f_{T_n}(t)$,

$$\begin{aligned} F_{T_n}(t) = P(T_n \le t) &= \frac{1}{2} + P(0 \le T_n \le t) \\ &= \frac{1}{2} + \frac{1}{2} P(-t \le T_n \le t) \\ &= \frac{1}{2} + \frac{1}{2} P\left(0 \le T_n^2 \le t^2\right) \\ &= \frac{1}{2} + \frac{1}{2} F_{T_n^2}(t^2) \end{aligned}$$

Differentiating $F_{T_n}(t)$ gives the stated result:

$$\begin{aligned} f_{T_n}(t) = F'_{T_n}(t) &= t \cdot f_{T_n^2}(t^2) \\ &= t \frac{n^{n/2}\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{n}{2}\right)} (t^2)^{-(1/2)} \frac{1}{(n+t^2)^{(n+1)/2}} \\ &= \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)} \cdot \frac{1}{\left[1+\left(\frac{t^2}{n}\right)\right]^{(n+1)/2}} \end{aligned}$$

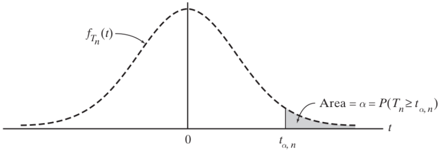
□




```
# Draw Student t−density
x=seq(−5,5,0.01)
pdf= cbind(dt(x, df = 1),
           dt(x, df = 2),
           dt(x, df = 5),
           dt(x, df = 100))
matplot(x,pdf, type = "l")
 title ("Student's t−distributions")
```

t Table

	α						
df	.20	.15	.10	.05	.025	.01	.005
1	1.376	1.963	3.078	6.3138	12.706	31.821	63.657
2	1.061	1.386	1.886	2.9200	4.3027	6.965	9.9248
3	0.978	1.250	1.638	2.3534	3.1825	4.541	5.8409
4	0.941	1.190	1.533	2.1318	2.7764	3.747	4.6041
5	0.920	1.156	1.476	2.0150	2.5706	3.365	4.0321
6	0.906	1.134	1.440	1.9432	2.4469	3.143	3.7074
⋮			⋮				
30	0.854	1.055	1.310	1.6973	2.0423	2.457	2.7500
∞	0.84	1.04	1.28	1.64	1.96	2.33	2.58



$$\mathbb{P}\left(T_3 > 4.541\right) = 0.01 \quad \Longleftrightarrow \quad t_{0.01,3} = 4.541$$

```
> 1-pt(4.541, df =3)
[1]  0.009998238
```

```
> alpha = 0.01
> qt(1-alpha, df = 3)
[1]  4.540703
```

Thm 5. Let $Y_1, \cdots, Y_n$ be a random sample from $N(\mu, \sigma^2)$. Then

$$T_{n-1} = \frac{\overline{Y} - \mu}{S/\sqrt{n}} \quad \sim \quad \text{Student's t of degree } n-1.$$

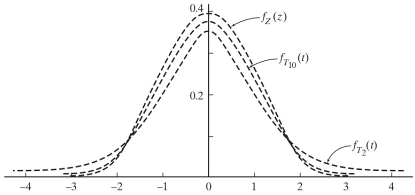
Proof.

$$\frac{\overline{Y} - \mu}{S/\sqrt{n}} = \frac{\dfrac{\overline{Y} - \mu}{\sigma/\sqrt{n}}}{\sqrt{\dfrac{(n-1)S^2}{\sigma^2(n-1)}}}$$

$$\frac{\overline{Y} - \mu}{\sigma/\sqrt{n}} \sim N(0,1) \qquad \perp \qquad \frac{(n-1)S^2}{\sigma^2} \sim \text{Chi Square}(n-1)$$

By Def. 2 ... $\square$

As $n \to \infty$, Students' t distribution will converge to $N(0,1)$:



Thm 6. $f_{T_n}(x) \to f_Z(x) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ as $n \to \infty$, where $Z \sim N(0,1)$.

Proof By Stirling's formula:

$$\Gamma(z) = \sqrt{\frac{2\pi}{z}} \left(\frac{z}{e}\right)^z (1 + O(1/z)) \qquad \text{as } z \to \infty$$

$$\Longrightarrow \quad \lim_{n\to\infty} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)} = \frac{1}{\sqrt{2\pi}}$$

...... $\square$