The t- and F-distributions

Case when σ^2 is not known

An outline proof of this results is as follows:

~~1.~~ We first prove that S^2 is an unbiased estimator for σ^2. (Theorem 4)

$E[S^2] = \sigma^2$

~~2.~~ We then derive the distribution of S^2 under the assumption that $Y_i \sim N(\mu, \sigma^2)$. (Lemma 4, Definition 2.5, Theorem 5)

$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$

$S^2 \perp \bar{Y}$

3. We find the distribution of $\boxed{\frac{\bar{Y}-\mu}{S/\sqrt{n}}}$. $\sim t_{n-1}$

(Definition 2.6, Theorem 6)

Definition 2.6

Suppose $Z \sim N(0,1)$ and $X \sim \chi_n^2$ independently, and let

$$T = \frac{Z}{\sqrt{\frac{X}{n}}},$$

then T is said to have a t-distribution with n degrees of freedom and we write

$$X \sim t_n.$$

$$f(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{x^2}{n}\right)^{-\left(\frac{n+1}{2}\right)} \quad \text{for } -\infty < x < \infty$$

$E(X) = 0$ if $n > 1$

$var(X) = \frac{n}{n-2}$ if $n > 2$

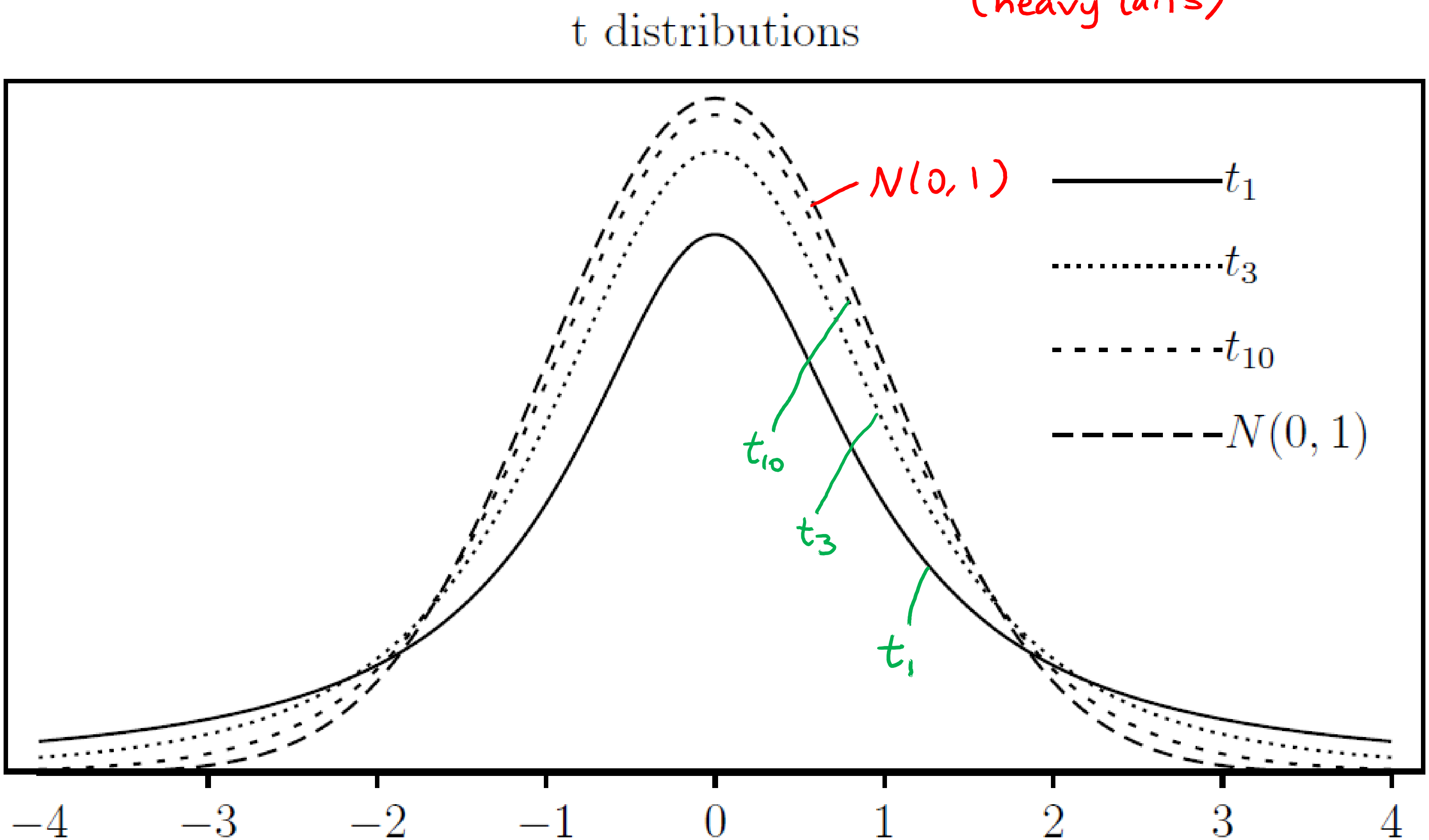
MGF does not exist

The t-distribution

Symmetric

flatter tails than normal

(heavy tails)



As $n \to \infty$, $t_n \to N$

Theorem 6

Suppose $Y_1, Y_2, \dots, Y_n$ are i.i.d. $N(\mu, \sigma^2)$ random variables.

Then

$$T = \frac{\bar{Y} - \mu}{S/\sqrt{n}} \sim t_{n-1}.$$

We will show that T can be written as a ratio of a $N(0,1)$ random variable and a $\frac{\chi^2_{n-1}}{n-1}$ random variable.

Proof of Theorem 6

$$T = \frac{\bar{Y}-\mu}{\frac{S}{\sqrt{n}}} \cdot \frac{\frac{\sigma}{\sqrt{n}}}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$

$$= \frac{\bar{Y}-\mu}{\frac{\sigma}{\sqrt{n}}} \cdot \frac{\frac{\sigma}{\cancel{\sqrt{n}}}}{\frac{S}{\cancel{\sqrt{n}}}}$$

$$= \frac{\bar{Y}-\mu}{\frac{\sigma}{\sqrt{n}}} \cdot \frac{\sigma}{S}$$

$$= \frac{\bar{Y}-\mu}{\frac{\sigma}{\sqrt{n}}} \Big/ \sqrt{\frac{S^2}{\sigma^2}}$$

$$= \underbrace{\frac{\bar{Y}-\mu}{\frac{\sigma}{\sqrt{n}}}}_{N(0,1)} \Big/ \sqrt{\frac{\underbrace{\frac{S^2(n-1)}{\sigma^2}}_{\chi^2_{n-1}}}{(n-1)}}$$

$$= Z \Big/ \sqrt{\frac{X}{n-1}}$$

$$\sim t_{n-1}$$

Let $Z = \frac{\bar{Y}-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$

$X = \frac{S^2(n-1)}{\sigma^2} \sim \chi^2_{n-1}$

Definition 2.7

Let W_1 and W_2 be independent χ^2-distributed random variables with v_1 and v_2 degrees of freedom respectively, then

$$\boxed{F = \frac{W_1/\nu_1}{W_2/\nu_2}} \sim F_{\nu_1,\nu_2}$$

is said to have an F-distribution with ν_1 numerator degrees of freedom and ν_2 denominator degrees of freedom.

$$f(x) = \frac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)} \frac{\nu_1^{\frac{\nu_1}{2}} \nu_2^{\frac{\nu_2}{2}} x^{\frac{\nu_1}{2}-1}}{(\nu_1 x + \nu_2)^{\left(\frac{\nu_1+\nu_2}{2}\right)}}, \quad x > 0$$

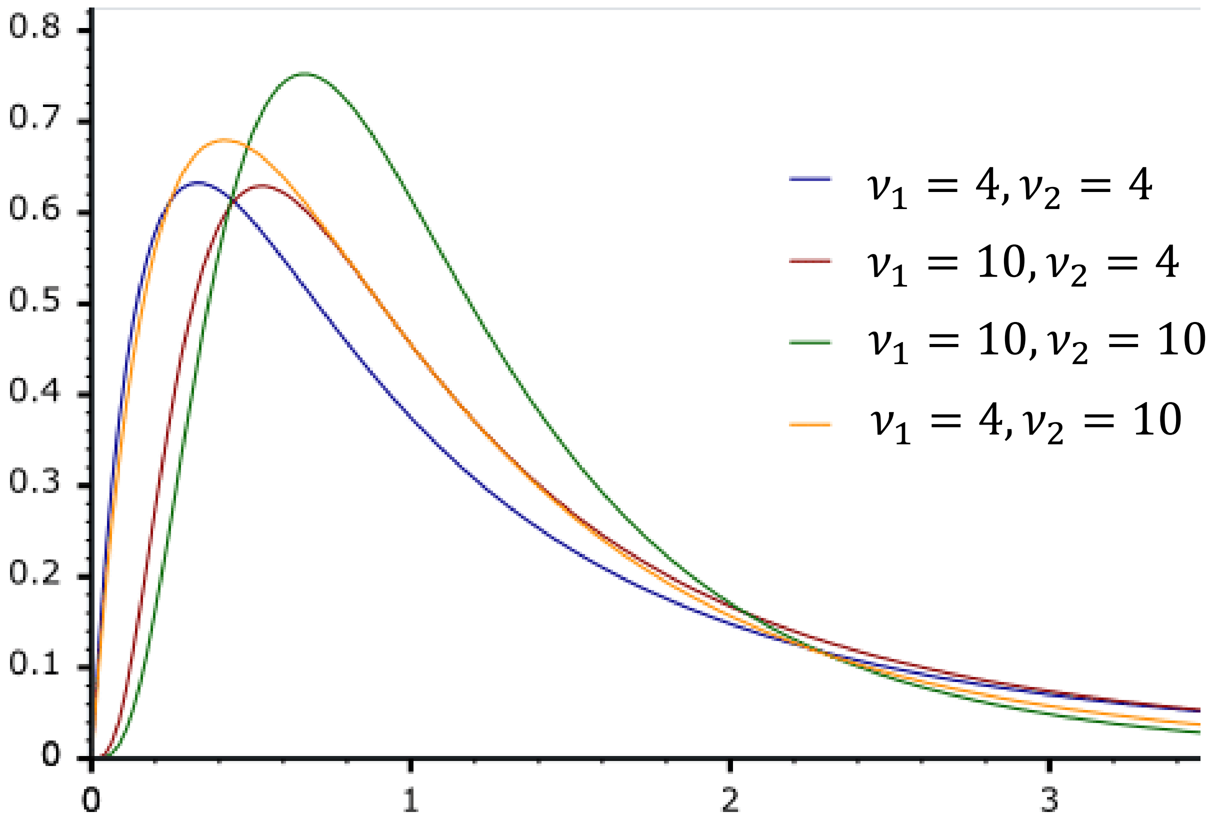
$$E(X) = \frac{\nu_2}{\nu_2-2} \quad \text{if } \nu_2 > 2$$

If $X \sim t_n$, then $X^2 \sim F_{1,n}$.

$$\text{var}(X) = \frac{2\nu_2^2(\nu_1+\nu_2-2)}{\nu_1(\nu_2-2)^2(\nu_2-4)} \quad \text{if } \nu_2 > 4$$

MGF does not exists

The F-distribution



skewed to the right

Example 2.4

Consider two independent random samples, with sample size n_1 and n_2, from normal distributions with variance σ_1^2 and σ_2^2 respectively. Let S_1^2 and S_2^2 be the sample variance of the corresponding samples. Show that

$$\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}$$

has an F-distribution.

$$\frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}} = \frac{\underbrace{\frac{S_1^2}{\sigma_1^2}(n_1-1)}_{\chi^2_{n_1-1}} \Big/ (n_1-1)}{\underbrace{\frac{S_2^2}{\sigma_2^2}(n_2-1)}_{\chi^2_{n_2-1}} \Big/ (n_2-1)} = \frac{\frac{W_1}{n_1-1}}{\frac{W_2}{n_2-1}} \sim F_{n_1-1,\, n_2-1}$$

where $W_1 = \frac{(n_1-1)S_1^2}{\sigma_1^2} \sim \chi^2_{n_1-1}$ and

$W_2 = \frac{(n_2-1)S_2^2}{\sigma_2^2} \sim \chi^2_{n_2-1}$ are independent

Example 2.5

Let S_1^2 denote the sample variance of a random sample of size 10 from Population I and let S_2^2 denote the sample variance for a random sample of size 8 from population II. The variance of Population I is three times the variance of Population II. Find the two numbers a and b such that $\boxed{P(a \le S_1^2/S_2^2 \le b) = 0.90}$, assuming S_1^2 is independent of S_s^2.

$n_1 = 10, \quad n_2 = 8, \quad \sigma_1^2 = 3\sigma_2^2$

$$F = \frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}} = \frac{\frac{S_1^2}{3\cancel{\sigma_2^2}}}{\frac{S_2^2}{\cancel{\sigma_2^2}}} = \frac{S_1^2}{3S_2^2} \sim F_{n_1-1,\, n_2-1} = F_{9,7}$$

$$P\left(\frac{S_1^2}{S_2^2} < a\right) = 0.05$$

$$P\left(\frac{S_1^2}{S_2^2} > b\right) = 0.05$$

Example 2.5

$$0.05 = P\left(\frac{S_1^2}{S_2^2} < a\right)$$

$$= P\left(\frac{S_1^2}{3S_2^2} < \frac{a}{3}\right)$$

$$= P\left(F < \frac{a}{3}\right) \text{ where } F \sim F_{9,7}$$

Use R qf(0.05, 9,7)

$$0.05 = P(F < 0.304)$$

$$\Rightarrow \frac{a}{3} = 0.304$$

$$a = 3(0.304)$$

$$= 0.912$$

$$0.05 = P\left(\frac{S_1^2}{S_2^2} > b\right)$$

$$= 1 - P\left(\frac{S_1^2}{S_2^2} \leq b\right)$$

$$= 1 - P\left(\frac{S_1^2}{3S_2^2} \leq \frac{b}{3}\right)$$

$$= 1 - P\left(F \leq \frac{b}{3}\right)$$

$$0.95 = P\left(F \leq \frac{b}{3}\right)$$

Using R qf(0.95, 9, 7)

$$0.95 = P(F \leq 3.677)$$

$$\Rightarrow \frac{b}{3} = 3.677$$

$$b = 11.03$$

$$\therefore \quad P\left(0.912 \leq \frac{S_1^2}{S_2^2} \leq 11.03\right) = 0.90$$