Math 4501 - Probability and Statistics II

5.5 - Random functions associated with the normal distribution (continued!)

Student's t distribution

Theorem $Z \sim N(0,1)$, $U \sim \chi^2(r)$ independent

Let

IMPORTANT FACT $\rightsquigarrow$
$$\boxed{T = \frac{Z}{\sqrt{U/r}}},$$
r = # dof of U

where Z is a standard normal random variable, U is a $\chi^2(r)$ random variable, and Z and U are independent.

Then, the probability density function of T is given by

$$f(t) = \frac{\Gamma((r+1)/2)}{\sqrt{\pi r}\,\Gamma(r/2)} \frac{1}{(1 + t^2/r)^{(r+1)/2}}, \quad t \in \mathbb{R},$$

$\Rightarrow$ $f(t)$ is even!

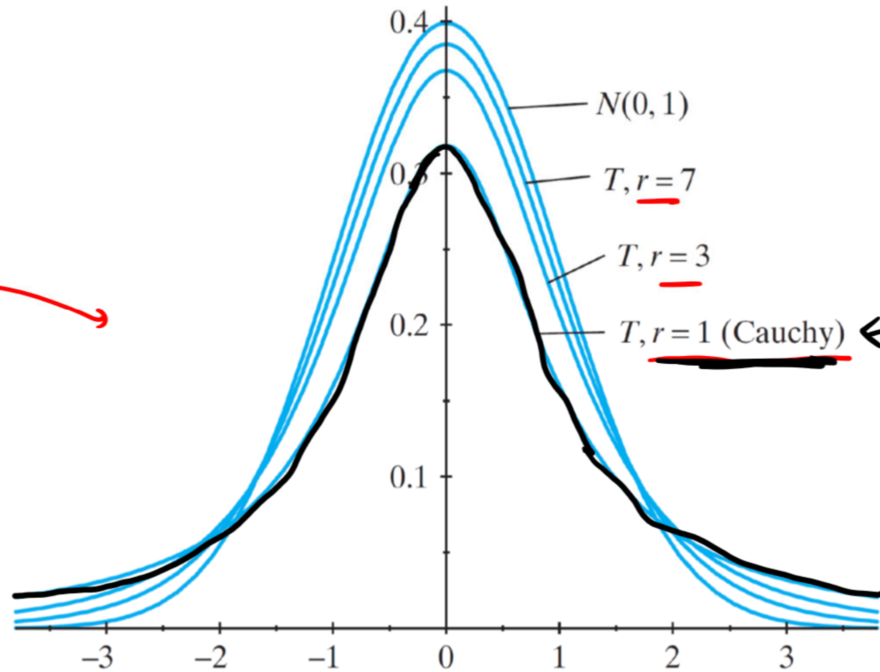
and T is said to have a Student's t distribution with r degrees of freedom, which we denote as $t(r)$.

Note: We use tables or appropriate software to evaluate probabilities associated with the t distribution.

Notes:

- The Cauchy distribution is a t distribution with 1 degree of freedom
- As $r \to \infty$, the t distribution with r degrees of freedom approaches the standard normal distribution
- The tails of the Cauchy distribution are always heavier than the standard normal distribution, and become lighter with increasing r

graph of t-student's pdfs with varying d.of.



← distribution with heavy tails!

pdf of $T \sim t(n)$ ← NOTATION FOR Student's t with n dofs

$$f(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\underbrace{\sqrt{\pi n}}_{n^{1/2}}\ \Gamma\left(\frac{n}{2}\right)} \cdot \frac{1}{\left(1+t^2/n\right)^{\frac{(n+1)}{2}}}$$

If $n = 1$

$\Gamma(1) = 0! = 1$

$$f(t) = \frac{\Gamma(1)}{\sqrt{\pi \cdot 1}\ \Gamma\left(\frac{1}{2}\right)} \cdot \frac{1}{\left(1+t^2/1\right)^{\frac{1+1}{2}}} = \frac{1}{\pi}\ \frac{1}{1+t^2}$$

} pdf of Cauchy

$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

special feature: Cauchy distribution has no moment $E[X^n]$, $n \in \mathbb{N}$

Case as $n \to \infty$.

$$f(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n}\;\Gamma\left(\frac{n}{2}\right)} \cdot \frac{1}{\left(1 + t^2/n\right)^{\frac{n+1}{2}}}$$

$\hookrightarrow \frac{1}{\sqrt{\pi}\cdot\sqrt{2}}$

$e^{-t^2/2}$

sum is $\frac{n+1}{2}$

Note :

(1) $$\left(1+\frac{t^2}{n}\right)^{\frac{n+1}{2}} = \left(1+\frac{t^2}{n}\right)^{\frac{n}{2}} \cdot \left(1+\frac{t^2}{n}\right)^{\frac{1}{2}} =$$

$$= \left[\left(1+\frac{t^2}{n}\right)^{n}\right]^{1/2} \cdot \left(1+\frac{t^2}{n}\right)^{\frac{1}{2}}$$

$$\downarrow n \to \infty \qquad\qquad \downarrow n \to \infty$$

$$\xrightarrow[\text{as } n \to \infty]{} \left(e^{t^2}\right)^{1/2} \cdot 1$$

Hint :

$$\left(1+\frac{1}{n}\right)^{n} \xrightarrow[n \to \infty]{} e$$

$$\left(1+\frac{x}{n}\right)^{n} \xrightarrow[n \to \infty]{} e^{x}$$

We've just proved that:

$$\frac{1}{\left(1+\frac{t^2}{n}\right)^{(n+1)/2}} \xrightarrow[n\to\infty]{} \frac{1}{e^{t^2/2}} = e^{-t^2/2}$$

remember the formula for pdf of $N(0,1)$

$$f(z) = \frac{1}{\sqrt{2\pi}}\, e^{-z^2/2}$$

Note 2: we can also prove that

$$\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n}\cdot\Gamma\left(\frac{n}{2}\right)} \to \frac{1}{\sqrt{2}}$$

why??

Property of gamma function : $\leftarrow$ generalizes factorial:

$n \in \mathbb{N}$: $\quad \Gamma(x+n) = (x+n-1)\cdot\Gamma(x+n-1)$

$$= (x+n-1)\cdot(x+n-2)\,\Gamma(x+n-2)$$

$$= (x+n-1)\cdot(x+n-2)\cdots x\,\Gamma(x)$$

$$\Rightarrow \frac{\Gamma(x+n)}{\Gamma(x)} = \overbrace{(x+n-1)(x+n-2)\cdots x}^{n \text{ factors}}$$

polynomials of degree n on numerator and denominator

$$\Rightarrow \boxed{\frac{\Gamma(x+n)}{x^n\cdot\Gamma(x)}} = \frac{(x+n-1)(x+n-2)\cdots x}{x^n} \xrightarrow[x\to\infty]{} 1$$

The previous statement is still true if we replace $n \in \mathbb{N}$ by a positive real number $\alpha > 0$

$$\Gamma(t) = \int_0^\infty y^{t-1} e^{-y}\, dy$$

$$\frac{\Gamma(x+\alpha)}{x^\alpha \cdot \Gamma(x)} \longrightarrow 1 \quad \text{as} \quad x \to \infty$$

$t = 1: \; \Gamma(1) = 1$

$t \in \mathbb{N}: \; t = n$

$\Gamma(n) = (n-1)\Gamma(n-1)$

$\Gamma(t) = (t-1)\Gamma(t-1)$

If we set $x = \frac{n}{2}$ and $\alpha = \frac{1}{2}$ we get

$$\to \frac{\Gamma\left(\frac{n+1}{2}\right)}{\left(\frac{n}{2}\right)^{1/2} \Gamma\left(\frac{n}{2}\right)} = \sqrt{2} \cdot \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n}\, \Gamma\left(\frac{n}{2}\right)} \Rightarrow \sqrt{2} \cdot \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n} \cdot \Gamma\left(\frac{n}{2}\right)} \to 1$$

$\downarrow$ as $n \to \infty$

1

$$\Rightarrow \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n}\, \Gamma\left(\frac{n}{2}\right)} \to \frac{1}{\sqrt{2}}$$

Important t distributed r.v.

Let $X_1, X_2, \ldots, X_n$ be a random sample from a normal distribution $N(\mu, \sigma^2)$. $\Rightarrow$ $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

i i d

We know that

$$\rightarrow \quad Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

← LAST TIME
(ALSO 3501)

and

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

← LAST TIME

are independent random variables.

Looks almost like

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

with σ replaced by S

Hence, we obtain that

We shall use a lot in Secs 7+8 →

$$T = \frac{Z}{\sqrt{U/(n-1)}} = \frac{\frac{\bar{X}-\mu}{\cancel{\sigma}/\sqrt{n}}}{\sqrt{\frac{\cancel{(n-1)}S^2}{\cancel{\sigma^2}}/\cancel{(n-1)}}} = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t(n-1)$$

$\downarrow$ as $n \to \infty$

$N(0,1)$

has a t distribution with $n-1$ degrees of freedom.

MOTIVATION:

$$X_1, X_2, \dots, X_n \sim N(\mu, \sigma^2)$$

$$\Rightarrow \quad \bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \quad \leftarrow \text{we want } \bar{X} \text{ to estimate } \mu$$

In general, μ and σ^2 are not known!

$\bar{X}$ is an estimate for μ, $\frac{\sigma^2}{n}$ gives a sense of the accuracy of the estimate

If we don't know σ^2, we might want to replace it by S^2 (estimate for σ^2

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N(0,1) \xrightarrow[\text{replaced by}]{} T = \frac{\bar{X} - \mu}{S / \sqrt{n}} \sim t(n-1)$$

$$T = \frac{Z}{\sqrt{V/n-1}}$$

Percentiles for the t distribution

Suppose T is $t(r)$, and let $\alpha \in (0,1)$ (usually $\alpha < 0.5$).

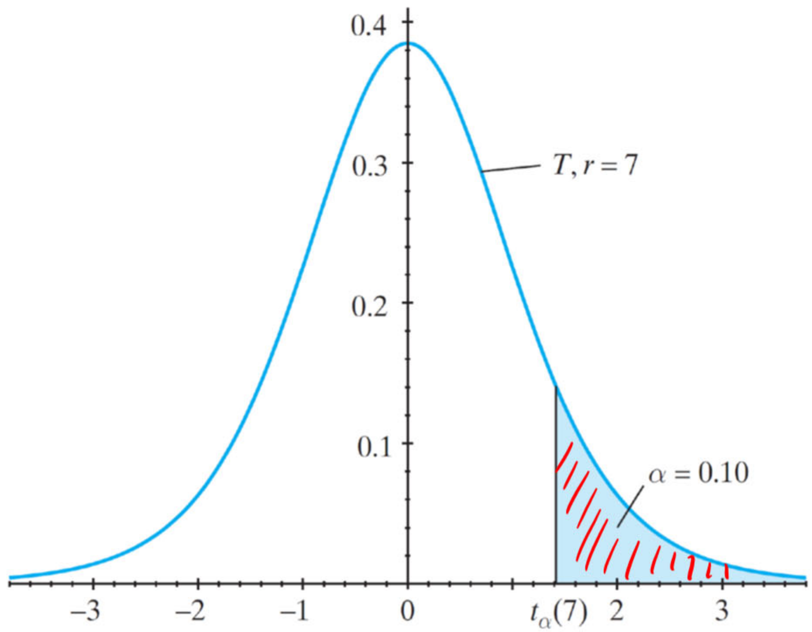
The $100(1-\alpha)$th percentile of its distribution is the (unique) number $t_\alpha(r)$ such that

$$P\,[T \geq t_\alpha(r)] = \alpha \ ,$$

that is, the probability to the right of $t_\alpha(r)$ is α.

z_α for $N(0,1)$

$\chi^2_\alpha(r)$ for $\chi^2(r)$



Example

Suppose T is $t(11)$.

Find $t_{0.05}(11)$.

$n = 11$

$\alpha = 0.05$

Check on the table the entry corresponding to

$n = 11$ and $\alpha = 0.05$

$$t_{0.05}(11) = 1.796$$

Example

Suppose T is $t(11)$.

Find $P(-1.363 < T \leq 2.201)$.

Let $T \sim t(11)$ and denote by F the cdf of $t(11)$.

Then $P(-1.363 < T \leq 2.201) = F(2.201) - F(\underline{-1.363})$

symmetry!

$= F(2.201) - (1 - F(1.363))$

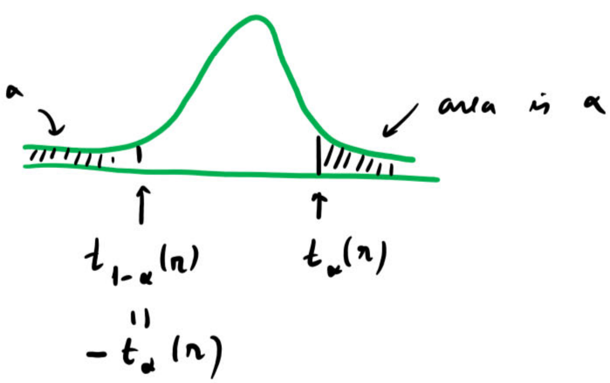
$= F(2.201) + F(1.363) - 1$

$= 0.975 + 0.9 - 1$

table!

$= 0.875$

pdf of $t(n)$ distribution is an even function!



Math 4501 - Probability and Statistics II

5.2 - The F distribution

The F distribution

Theorem

$U \sim \chi^2(r_1)$, $V \sim \chi^2(r_2)$ independent !

Let U and V be independent chi-square distributed random variables with r_1 and r_2 degrees of freedom, respectively.
Then

\# of dofs from numerator

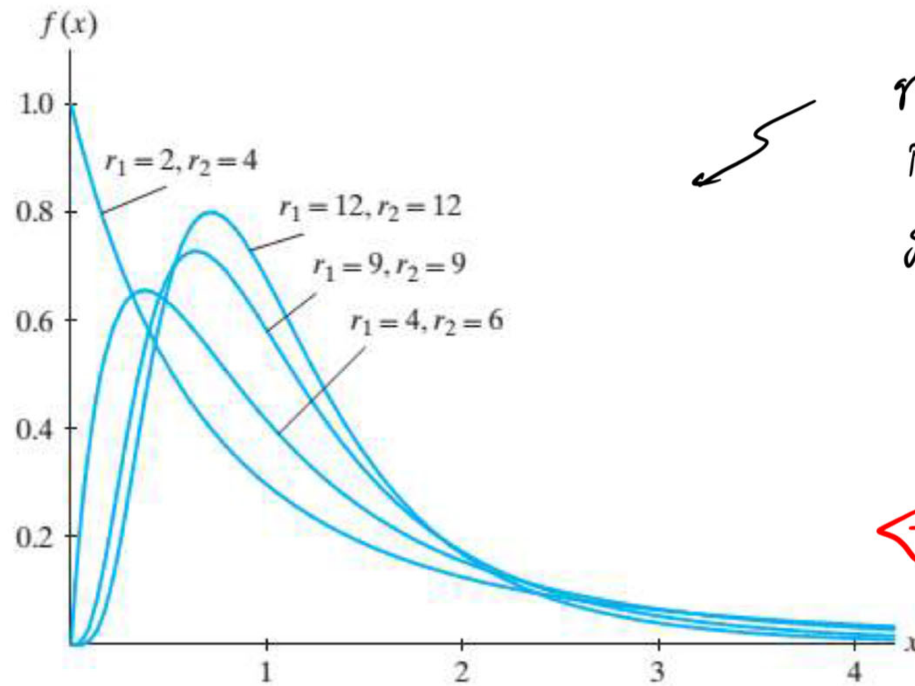
IMPORTANT FACT $\rightarrow$ $$\boxed{W = \frac{U/r_1}{V/r_2} \sim F(r_1, r_2)}$$ \# dofs from denominator

has an F distribution with r_1 and r_2 degrees of freedom, for which the pdf is given by

$$\boxed{f(w) = \frac{\Gamma\left[(r_1 + r_2)/2\right](r_1/r_2)^{r_1/2} w^{r_1/2-1}}{\Gamma(r_1/2)\Gamma(r_2/2)\left[1 + (r_1/r_2)w\right]^{(r_1+r_2)/2}}, \qquad w > 0 .}$$

Notation: if W has an F distribution with r_1 and r_2 degrees of freedom, we say that the distribution of W is $F(r_1, r_2)$ and denote it as $W \sim F(r_1, r_2)$.

Note: We use tables or appropriate software to evaluate probabilities associated with the F distribution.



Several examples of F distribution pdfs

graphs look like pdfs of gamma or χ^2 (special kind of gamma)

MOTIVATION: Why will we need the F distribution ??

random sample $X_1, X_2, \ldots, X_{n_x} \sim N(\mu_x, \sigma_x^2) \Rightarrow U = \frac{(n_x-1)S_x^2}{\sigma_x^2} \sim \chi^2(n_x-1)$

LAST TIME

+

another independent

random sample $Y_1, Y_2, \ldots, Y_{n_y} \sim N(\mu_y, \sigma_y^2) \Rightarrow V = \frac{(n_y-1)S_y^2}{\sigma_y^2} \sim \chi^2(n_y-1)$

Define: r.v

$$W = \frac{U/(n_x-1)}{V/(n_y-1)} = \frac{\frac{(n_x-1)S_x^2}{\sigma_x^2}\Big/(n_x-1)}{\frac{(n_y-1)S_y^2}{\sigma_y^2}\Big/(n_y-1)} = \frac{S_x^2/\sigma_x^2}{S_y^2/\sigma_y^2} = \frac{S_x^2}{S_y^2}\cdot\frac{\sigma_y^2}{\sigma_x^2}$$

$$W \sim F(\underbrace{n_x-1}_{r_1}, \underbrace{n_y-1}_{r_2})$$

$\frac{S_x^2}{S_y^2} \xrightarrow[\text{information about}]{\text{infer}} \frac{\sigma_x^2}{\sigma_y^2}$

↳ Secs 7 and 8

Percentiles for F distribution

$W = \frac{U/r_1}{V/r_2}$ $U \sim \chi^2(r_1)$ $V \sim \chi^2(r_2)$

Suppose W is $F(r_1, r_2)$, and let $\alpha \in (0, 1)$ (usually $\alpha < 0.5$).

The $100(1 - \alpha)$th percentile is the number $F_\alpha(r_1, r_2)$ such that

$$\boxed{P\left[W \geq F_\alpha(r_1, r_2)\right] = \alpha} \ ,$$

that is, the right-tail probability of $F_\alpha(r_1, r_2)$ is α.

The 100α percentile is the number $F_{1-\alpha}(r_1, r_2)$ such that

$$P\left[W \leq F_{1-\alpha}(r_1, r_2)\right] = \alpha \ ,$$

that is, the right-tail probability of $F_{1-\alpha}(r_1, r_2)$ is $1 - \alpha$.

and so the left tail of $F_{1-\alpha}(r_1, r_2)$ is α

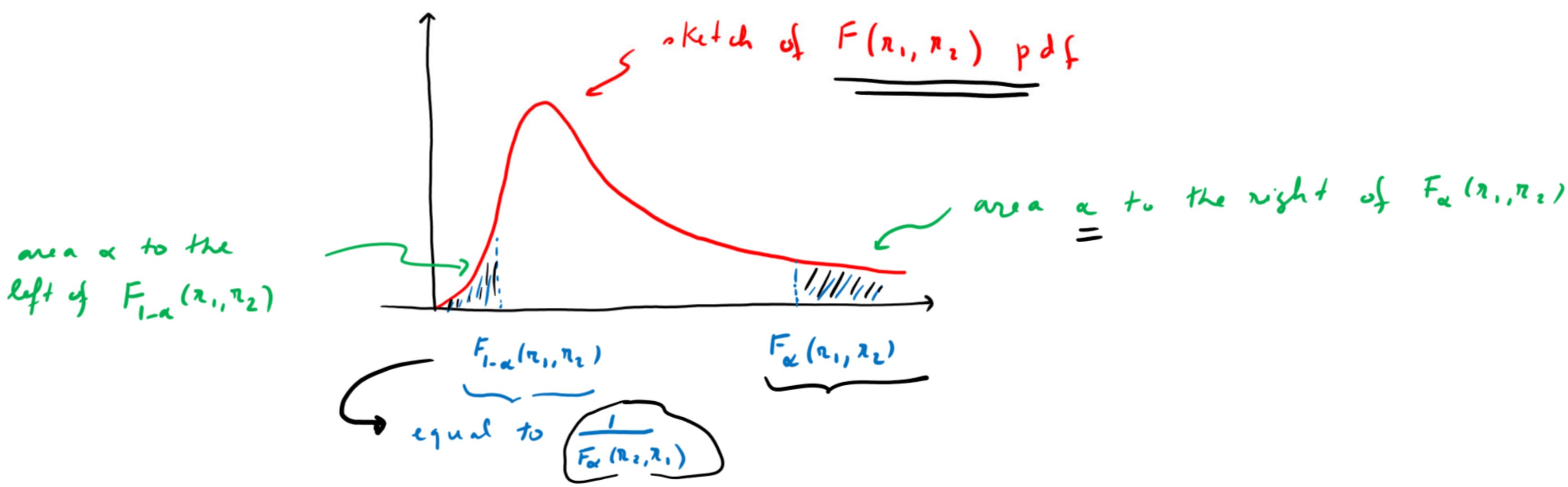
Property: If W is $F(r_1, r_2)$, then $1/W$ is $F(r_2, r_1)$.

Consequence: Since $\boxed{W = \dfrac{U/r_1}{V/r_2}, \quad U \sim \chi^2(r_1),\ V \sim \chi^2(r_2) \text{ ind.}} \Rightarrow \dfrac{1}{W} = \dfrac{V/r_2}{U/r_1} \sim F\underbrace{(r_2, r_1)}_{\text{dof in opposite order to those of } W}$

$$\alpha = P\left[W \le \underbrace{F_{1-\alpha}(r_1, r_2)}\right] = P\left[1/W \ge \underbrace{1/F_{1-\alpha}(r_1, r_2)}_{F_\alpha(r_2, r_1)}\right]$$

it follows that

$$\frac{1}{F_{1-\alpha}(r_1, r_2)} = F_\alpha(r_2, r_1) \quad \text{or} \quad \boxed{F_{1-\alpha}(r_1, r_2) = \frac{1}{F_\alpha(r_2, r_1)}}.$$



Example

Let $W \sim F(4,6)$.

Evaluate $P(1/15.21 \leq W \leq 9.15)$. $=$???

Let $G_{4,6}$ denote the cdf of $F(4,6)$

then $P\left(\frac{1}{15.21} \leq W \leq 9.15\right) = G_{4,6}(9.15) - G_{4,6}\left(\frac{1}{15.21}\right)$

$= \; ? \; - \; ?$

$= 0.99 - 0.01 = 0.98$

use table:

$G_{4,6}(9.15) = P(W \leq 9.15) = 0.99$

$G_{4,6}\left(\frac{1}{15.21}\right) = P\left(W \leq \frac{1}{15.21}\right) = P\left(\frac{1}{W} > 15.21\right) = 0.01$

$\frac{1}{W} \sim F(6,4)$