ADVANCED BAYESIAN MODELING

ROBUST INFERENCE:
THE *t* DISTRIBUTION

In classical statistics, the t distribution is often used for normal-sample inference.

A generalized version also has properties that make it a robust alternative to the normal distribution ...

Location-Scale t Distribution

The t distribution with degrees of freedom $\nu > 0$, location μ, and scale $\sigma > 0$, denoted

$$t_\nu(\mu, \sigma^2)$$

has continuous density

$$p(x) \quad \propto \quad \left(1 + \frac{(x-\mu)^2}{\nu\sigma^2}\right)^{-(\nu+1)/2}$$

for all x.

When $\nu > 1$, the mean is μ.

For $\nu \leq 1$, the mean is undefined, but μ is still the median.

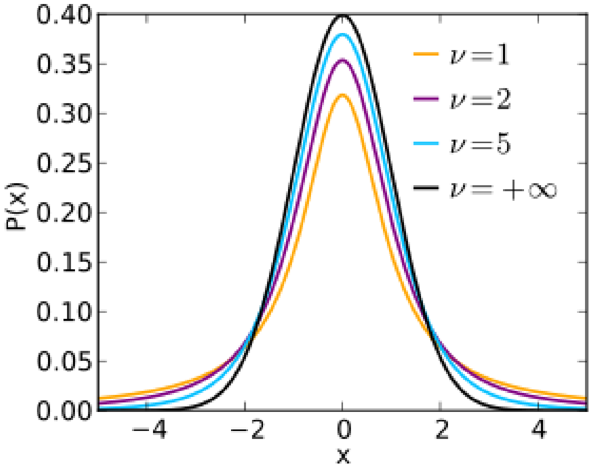
The variance is actually

$$\frac{\nu}{\nu - 2}\,\sigma^2$$

if $\nu > 2$, and otherwise undefined.

$t_1(\mu, \sigma^2)$ is the (location-scale) Cauchy distribution that was used earlier in logistic regression.

$t_\nu(0, 1)$ is the standard t distribution with ν df, used in classical statistics ...



From: Student's t-distribution. (2017, November 30). In *Wikipedia, The Free Encyclopedia*. Retrieved November 30, 2017, from
`https://en.wikipedia.org/w/index.php?title=Student%27s_t-distribution&oldid=812931585`

Degrees of Freedom and Tails

For small ν, the t distribution has **long (heavy) tails**, which are thicker than those of a normal.

For larger ν, the tails become shorter, and the distribution becomes more like a normal. In fact,

$$t_\nu(\mu, \sigma^2) \quad \underset{\nu \to \infty}{\longrightarrow} \quad \mathrm{N}(\mu, \sigma^2)$$

So a t distribution could have long or short tails, as controlled by ν.

Note: ν need not be an integer – it can be any positive value.

To illustrate the long tails when $\nu = 2$:

k	Prob. a $t_2(\mu, \sigma^2)$ variable is $> k\sigma$ from μ
1	0.4226497
2	0.1835034
3	0.0954660
4	0.0571910
5	0.0377496

(Compare with table for the short-tailed $\mathrm{N}(0, 1)$ earlier.)

Long Tails and Outliers

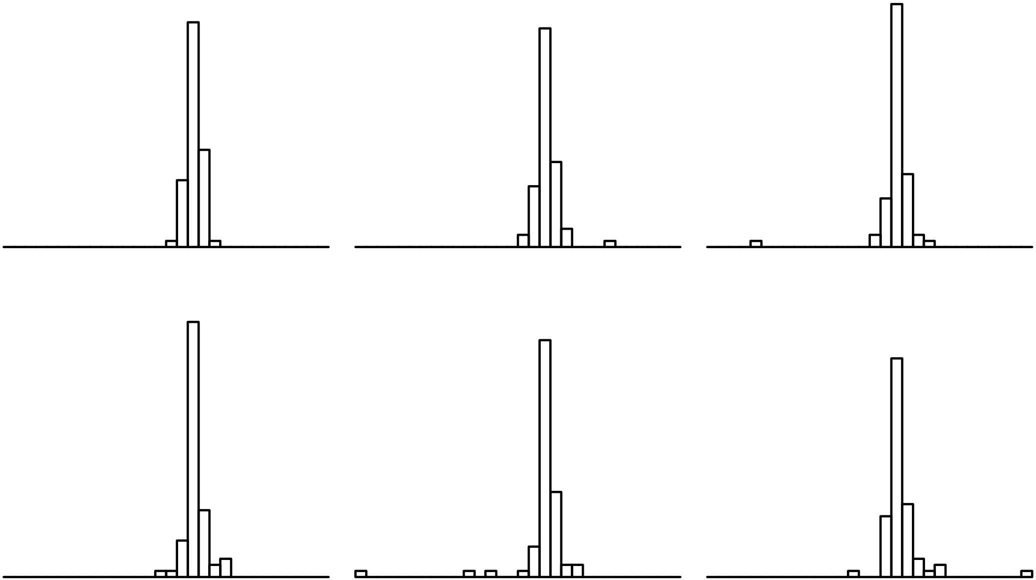
A continuous distribution with long tails tends to produce samples that have outliers, relative to a normal distribution.

For example, let's simulate 6 samples of size 66 from $t_2(0, 1)$:

```
> n <- 66

> tsamp <- matrix(rt(6*n, 2), n, 6)  # 6n indep standard t's with df=2
```

Now make a histogram of each column ...



Remark: R functions for the t distribution do not admit a location or a scale parameter – they assume $\mu = 0$ and $\sigma^2 = 1$.

You can use

$$X \;\sim\; t_\nu(0,1) \qquad \Rightarrow \qquad \mu + \sigma X \;\sim\; t_\nu(\mu, \sigma^2)$$

For example, to simulate n independent draws from $t_\nu(\mu, \sigma^2)$ in R, use

```
mu + sigma * rt(n, nu)
```