§18. t-Random Variables

$t-$ distribution is a member of a one parameter family of random variables. The member of the $t-$family have heavier tails than the normal random variables.

Remark
The Gamma function is defined as

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$$

for $x > 0$
Integration by parts shows that on the natural numbers $\Gamma(n) = (n-1)!$

Definition 1

A random variable X is **t-distributed** with $\nu-$ degrees of freedom, $X \sim \text{Stud}(\nu)$ if it has pdf

$$p(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{x^2}{2}\right)^{-\frac{\nu+1}{2}} \qquad , \qquad x \in \mathbb{R}$$

Example 1

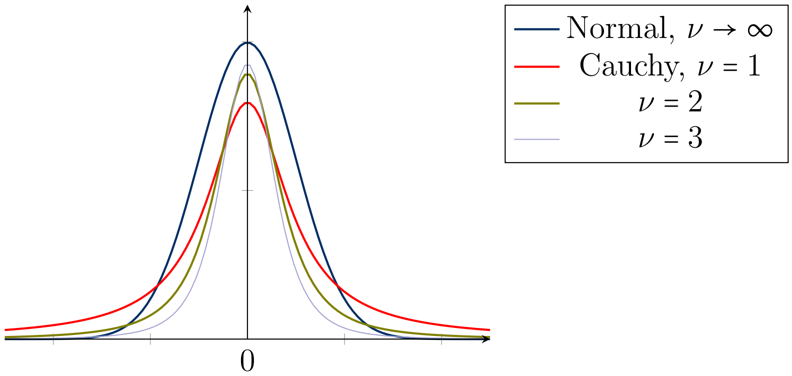
$$\nu = 1 \qquad p(x) = \frac{1}{\pi\left(1+x^2\right)} \quad ; \quad \text{so this is Cauchy;}$$

$$\nu = 2 \qquad p(x) = \frac{1}{2\sqrt{2}\left(1+\frac{x^2}{2}\right)^{3/2}}$$

$$\nu = 3 \qquad p(x) = \frac{2}{\pi\sqrt{3}\left(1+\frac{x^2}{3}\right)^2}$$

$$\vdots$$

$$\nu \to \infty \qquad p(x) \to \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \quad ; \quad \text{so this is the standard normal.}$$



t-distributed random variables are usually deployed to study the sample mean of a sample drawn from a normal distribution when the variance of the Gaussian is unknown.

Let $X_1, X_2, \ldots, X_n \sim N(\mu, \sigma^2)$ and let $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ be the sample mean (it is a random variable as well). Then

$$Z = \frac{X - \mu}{\sigma/\sqrt{n}} \sim N(0,1) \qquad \text{and} \qquad T = \frac{\overline{X} - \mu}{s/\sqrt{n}} \sim \text{Stud}(\nu) \quad , \quad \nu = n - 1$$

where $s^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$ is the sample variance.

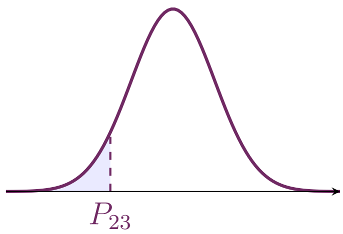
Example 2

A US company claims that their whole house water filters last on average 300 days. Assume that the life span of these filters is normally distributed. Yvan has (unwittingly) collected a sample of 15 such filters. The sample average is 290 days with sample standard deviation of 50 days. Compute the percentile value of this sample.

Solution

$t = \dfrac{\bar{x} - \mu}{s/\sqrt{n}} = \dfrac{290 - 300}{50/\sqrt{15}} = -0.7746$ with $\nu = n - 1 = 14\,df$

$P(T < -0.7746) = 0.2257 \sim 23$ percentile.



Example 3

The historical daily high for Montreal on April 1^{st} is $7°C$. Assume that these April 1^{st} daily highs are normally distributed. A sample of daily highs in April 1^st from the last 6 years shows:

6.8	7.2	8.9	11.2	5.9	10.4

Compute the percentile position of this sample (do this in excel).

Solution

$\bar{x} = 8.4 \quad ; \quad s = 2.1138$

$t = \dfrac{84 - 7}{2.1138/\sqrt{6}} = 1.622 \quad$ with $\quad \nu = 5\,df$

$P(T < 1.622) = 0.9171 \sim P_{92}$