The Student’s t-distribution is a symmetric, unimodal distribution with declining, unbounded tails on both sides of the center. It originates from the estimation of the mean of a normally distributed population, where the population variance is unknown and the sample size is small. The Student’s t-distribution has one main feature that distinguishes it from the normal distribution: in general its tails are heavier, meaning they fall off like a power of $|x|$. The decay of the tails is governed by the only parameter $\nu$, the number of degrees of freedom. The similarity between the Student’s t- and standard normal distribution grows with $\nu$; if $\nu=\infty$, the Student’s t-distribution and the normal distribution are identical (@Walck).

\*\*Mathematical Definition:\*\*

\newline

Probability density function:

$$f\_X\left(x;\nu\right)=\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}=\frac{\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}}{\sqrt{\nu}\ \mathrm{B}\ \left(\frac{1}{2},\frac{\nu}{2}\right)}$$

Where $\Gamma(\cdot)$ and $\mathrm{B}(c,d) = \frac{\Gamma(c) \Gamma(d)}{\Gamma(c+d)}$ are respectively the gamma and the complete beta functions, needed to normalize the distribution.

Cumulative distribution function:

$$F\_X(x;\nu)= \begin{cases}\frac{1}{2}I\_\frac{\nu}{\nu+x^2}(\frac{\nu}{2},\frac{1}{2}) & \text{for } x \leq 0 \\1-\frac{1}{2}I\_\frac{\nu}{\nu+x^2}(\frac{\nu}{2},\frac{1}{2}) & \text{for } x > 0\end{cases}$$

Where $I\_y\left(c,d\right) = \frac{\mathrm{B}(y;c,d)}{\mathrm{B}(c,d)}$ is the regularized incomplete beta function with $\mathrm{B}(y;c,d)$ being the incomplete beta function.

Expected value:

$$\mathrm{E}(X)=\begin{cases}\text{undefined} & \text{for } \nu \leq 1 \\ 0 & \text{for } \nu > 1\end{cases}$$

Variance:

$$\mathrm{Var}(X)= \begin{cases}\text{undefined} & \text{for } \nu \leq 1 \\\infty & \text{for } 1 < \nu \leq 2 \\\frac{\nu}{\nu-2} & \text{for } \nu > 2\end{cases}$$

The above formula for the variance implies that when estimating $\nu$ for a model, it necessary to set a lower bound $b$, $2<b\leq \tilde{\nu}$ for the variance of the model to be finite.