TOOL: The random variable T, defined as follows, is the sampled data analogue of a standard normal (or gaussian) distribution. Because the value of T depends on S, however, T has a t-distribution that differs slightly from the standard normal (or gaussian) distribution.

$$T \equiv \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

where

$n \equiv$ number of data points, X_i, (which are independent and normally distributed with mean μ and variance σ)

$$\overline{X} \equiv \frac{1}{n}\sum_{i=1}^{n} X_i \equiv \text{sample mean}$$

$$S^2 \equiv \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 \equiv \text{sample variance}$$

The probability density function of T is a t-distribution with $\nu = n - 1$ degrees of freedom:

$$f_T(t) = \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{\pi\nu}}\left(1 + \frac{t^2}{\nu}\right)^{-(\nu+1)/2}$$

DERIV: The first step is to express T in terms of random variables with known distributions:

$$T = \frac{Z}{\chi/\sqrt{\nu}}$$

where

$$Z \equiv \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \qquad \text{and} \qquad \chi^2 \equiv \nu\frac{S^2}{\sigma^2}$$

As shown in other Conceptual Tools, Z has a standard normal (or gaussian) distribution, and χ^2 has a chi-squared distribution with $\nu = n - 1$ degrees of freedom:

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

$$f_{\chi^2}(x) = \begin{cases} \dfrac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{(\nu/2)-1} e^{-x/2} & x > 0 \\ 0 & x \le 0 \end{cases}$$

As shown in another tool, Z and χ^2 are also independent. This allows us to write the joint distribution of Z and χ^2 as the product of the respective probability density functions for Z and χ^2:

$$f(z,x) = \begin{cases} \dfrac{1}{\sqrt{2\pi}} e^{-z^2/2} \dfrac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{(\nu/2)-1} e^{-x/2} & x > 0 \\ 0 & x \le 0 \end{cases}$$

To find the probability density for T, we consider the cumulative distribution for T and take its derivative:

$$f_T(t) = \frac{d}{dt} F_T(t) = \frac{d}{dt} P(T \le t) = \frac{d}{dt} \int_{-\infty}^{\infty} \int_{\frac{z}{\sqrt{x}/\sqrt{\nu}} = t}^{\infty} f(z,x) dx dz$$

NOTE: We treat the case of $t > 0$. We have t in the lower limit in this case because larger values of x give smaller values of t. For the case of $t < 0$, we would integrate from $-\infty$ to t.

We move the derivative inside the outer integral and write the lower limit in terms of the x more clearly:

$$f_T(t) = \int_{-\infty}^{\infty} \left[\frac{d}{dt} \int_{x = \nu \frac{z^2}{t^2}}^{\infty} f(z,x) dx \right] dz$$

The chain rule yields an interpretation of the derivative of the integral:

$$\frac{d}{dt} \int_{x = \nu \frac{z^2}{t^2}}^{\infty} f(z,x) dx = \frac{d}{dx} \int_{x = \nu \frac{z^2}{t^2}}^{\infty} f(z,x) dx \cdot \frac{dx}{dt}$$

or

$$\frac{d}{dt} \int_{x = \nu \frac{z^2}{t^2}}^{\infty} f(z,x) dx = -f(z,x)\Big|_{x = \nu \frac{z^2}{t^2}} \cdot \nu \frac{-2z^2}{t^3}$$

Making this substitution, we have the following expression:

$$f_T(t) = \int_{-\infty}^{\infty} f(z,x)\Big|_{x=\nu\frac{z^2}{t^2}} \cdot \nu\frac{2z^2}{t^3}dz$$

Now we use the expression for $f(z, x)$, assuming $t > 0$:

$$f(z,x)\Big|_{x=\frac{\nu z^2}{t^2}} = \begin{cases} \frac{1}{\sqrt{2\pi}}e^{-z^2/2}\frac{1}{2^{\nu/2}\Gamma(\nu/2)}\left(\frac{\nu z^2}{t^2}\right)^{(\nu/2)-1} e^{-\left(\frac{\nu z^2}{t^2}\right)/2} & z > 0 \\ 0 & z \le 0 \end{cases}$$

NOTE: We have the condition that $z > 0$ because the definition of T requires $Z > 0$ to achieve $T > 0$. (X and $\sqrt{X}$ are always positive.)

Incorporating the constraint on z into the lower limit of the outer integral yields the following expression:

$$f_T(t) = \int_0^{\infty} \frac{1}{\sqrt{2\pi}}e^{-z^2/2}\frac{1}{2^{\nu/2}\Gamma(\nu/2)}\left(\frac{\nu z^2}{t^2}\right)^{(\nu/2)-1} e^{-\left(\frac{\nu z^2}{t^2}\right)/2} \nu\frac{2z^2}{t^3}dz$$

It is helpful to define a term for the constants:

$$k = 2\frac{1}{\sqrt{2\pi}}\frac{1}{2^{\nu/2}\Gamma(\nu/2)}\nu\cdot\nu^{(\nu/2)-1} = \frac{1}{\sqrt{\pi}}\frac{1}{2^{(\nu-1)/2}\Gamma(\nu/2)}\nu^{\nu/2}$$

Using this new term, we have the following expression:

$$f_T(t) = k\int_0^{\infty} e^{-z^2/2}\left(\frac{z^2}{t^2}\right)^{(\nu/2)-1} e^{-\left(\frac{\nu z^2}{t^2}\right)/2}\left(\frac{z^2}{t^2}\right)\frac{1}{t}dz$$

or

$$f_T(t) = k\int_0^{\infty} e^{-z^2\left(1+\frac{\nu}{t^2}\right)/2}\left(\frac{z^2}{t^2}\right)^{\nu/2}\frac{1}{t}dz$$

We define the following variable that allows us to use a convenient change of variables:

$$z_t \equiv z\sqrt{1+\frac{\nu}{t^2}}$$

We have the following relationships for the change of variables:

$$\frac{dz_t}{\sqrt{1+\frac{\nu}{t^2}}} = dz, \quad z = 0 \Rightarrow z_t = 0, \quad z \to \infty \Rightarrow z_t \to \infty$$

and

$$\frac{z^2}{t^2} = \frac{z_t^2}{1+\frac{\nu}{t^2}} \cdot \frac{1}{t^2} = z_t^2 \frac{1}{t^2+\nu} = z_t^2 \frac{1}{\nu} \frac{1}{1+\frac{t^2}{\nu}} = z_t^2 \nu^{-1}\left(1+\frac{t^2}{\nu}\right)^{-1}$$

In terms of z_t we have the following integral expression:

$$f_T(t) = k\int_0^\infty e^{-z_t^2/2}\left(z_t^2\right)^{\nu/2}(\nu)^{-\nu/2}\left(1+\frac{t^2}{\nu}\right)^{-\nu/2}\frac{1}{\sqrt{\nu}}\frac{dz_t}{\sqrt{1+\frac{t^2}{\nu}}}$$

or

$$f_T(t) = k\nu^{-(\nu+1)/2}\left(1+\frac{t^2}{\nu}\right)^{-(\nu+1)/2}\int_0^\infty e^{-z_t^2/2}\left(z_t^2\right)^{\nu/2} dz_t$$

We change variables again in order to make the integrand have the form of a chi-squared distribution:

$$w \equiv z_t^2$$

We have the following relationships for the change of variables:

$$\frac{dw}{2z_t} = \frac{dw}{2\sqrt{w}} = dz_t, \quad z_t = 0 \Rightarrow w = 0, \quad z_t \to \infty \Rightarrow w \to \infty$$

In terms of w we have the following integral expression:

$$f_T(t) = k\nu^{-(\nu+1)/2}\left(1+\frac{t^2}{\nu}\right)^{-(\nu+1)/2}\int_0^\infty w^{\nu/2}e^{-w/2}\frac{dw}{2\sqrt{w}}$$

By moving appropriate constants inside the integral, we obtain the integral of an entire chi-squared distribution with ν + 1 degrees of freedom. In other words the probability represented by the following integral must equal unity:

$$\int_0^\infty \frac{1}{2^{(\nu+1)/2}\Gamma((\nu+1)/2)} w^{((\nu+1)/2)-1} e^{-w/2} dw = 1$$

It follows that we have the following expression for the probability density function for T:

$$f_T(t) = k2^{(\nu-1)/2}\Gamma((\nu+1)/2)\nu^{-(\nu+1)/2}\left(1+\frac{t^2}{\nu}\right)^{-(\nu+1)/2}$$

Simplification of the constants yields the final result:

$$f_T(t) = \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{\pi\nu}}\left(1+\frac{t^2}{\nu}\right)^{-(\nu+1)/2}$$