

# Proving the $t$ -Distribution Arises from Sampling Normally-Distributed Populations

Benjamin Posnick, bmp53@cornell.edu

May 7, 2020

## Introduction

We want to prove the following claim:

**Claim 1.** *The  $t$ -distribution arises from sampling normal populations when the variance is unknown and is taken from the sample itself.*

The probability density function  $f_T(\cdot)$  of the  $t$ -distribution is given as follows:

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

where  $\Gamma(\cdot)$  is the gamma function. Our goal will be to show that a random variable  $T$  constructed from sampling  $n$  points from normally-distributed populations of unknown variance has exactly this probability density function, with  $\nu = n - 1$  degrees of freedom.

## The Random Variable $T$

Let  $X_1, \dots, X_n$  be independent and identically distributed (i.i.d.) normal random variables, with unknown mean  $\mu$  and unknown variance  $\sigma^2$ . We define the random variable  $T$  as follows:

$$T = \frac{\bar{X} - \mu}{\sqrt{S^2/n}} \quad (2)$$

where  $S$  is the sample standard deviation of  $X_1, \dots, X_n$ :

$$S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2} \quad (3)$$

We can rewrite  $T$  as follows, where  $Z$  is the standard normal random variable, i.e.  $Z \sim \mathcal{N}(0, 1)$ :

$$\begin{aligned}
T &= \frac{\bar{X} - \mu}{S/\sqrt{n}} \\
&= \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}} \cdot \frac{\sigma}{\sigma}} \\
&= \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}} \cdot \frac{S}{\sigma}} \\
&= Z \frac{1}{\frac{S}{\sigma}}
\end{aligned}$$

where the last line comes from the fact that  $\mathbb{E}[\bar{X}] = \mu$  and  $Std(\bar{X}) = \sqrt{Var(\bar{X})} = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}}$ . Using equation (3), we obtain:

$$\begin{aligned}
T &= Z \frac{1}{\frac{1}{\sigma} \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}} \\
&= Z \frac{1}{\sqrt{\frac{1}{n-1} \sum_{i=1}^n \frac{(X_i - \bar{X})^2}{\sigma^2}}} \\
&= Z \frac{1}{\sqrt{\frac{V}{n-1}}}
\end{aligned}$$

where  $V = \sum_{i=1}^n \frac{(X_i - \bar{X})^2}{\sigma^2}$ . Since  $V = \sum_{i=1}^n \frac{(X_i - \bar{X})^2}{\sigma^2} = \frac{(n-1)S^2}{\sigma^2}$ , by Cochran's theorem, it follows that  $V \sim \chi_{n-1}^2$ , i.e.  $V$  has chi-square distribution with  $n-1$  degrees of freedom. This comes from the fact that the square of the sum of two normals is exponentially-distributed, and the sum of exponentials is gamma-distributed, where the chi-square distribution is the special case of this exponential.

## The Probability Density Function of $T$

We now want to find the probability density function of  $T$ . If we define  $X = \sqrt{\frac{V}{n-1}} = \sqrt{\frac{(n-1)S^2}{n-1}}$ , then our expression for  $T$  becomes:

$$T = \frac{Z}{X}$$

The probability density function of  $X$  is found as follows:

$$\begin{aligned}
f_X(x) &= \frac{d}{dx} \left[ \mathbb{P} \left( \sqrt{\frac{V}{n-1}} \leq x \right) \right] \\
&= \frac{d}{dx} \left[ \mathbb{P} \left( V \leq (n-1)x^2 \right) \right] \\
&= \frac{d}{dx} \left[ F_V((n-1)x^2) \right] \\
&= 2(n-1)x \cdot f_V((n-1)x^2) \\
&= \frac{2(n-1)x}{2^{(n-1)/2} \Gamma(\frac{n-1}{2})} [(n-1)x^2]^{\frac{n-1}{2}-1} \exp \left( -\frac{(n-1)x^2}{2} \right)
\end{aligned}$$

Setting  $\nu = n - 1$  and simplifying, we obtain:

$$f_X(x) = \frac{2^{1-\nu/2}}{\Gamma(\frac{\nu}{2})} \nu^{\frac{\nu}{2}} x^{\nu-1} \exp\left(-\frac{\nu x^2}{2}\right)$$

**Lemma 1.**  $Z$  and  $X$  are independent random variables.

*Proof.* We can prove this claim by showing that  $\bar{X}$  and  $S^2$  are independent random variables, since  $Z$  was derived from  $\bar{X}$ , since and  $X$  was derived from  $S^2$ . We know that:

$$\bar{X} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$$

We know that inside the summation over  $i \in \{1, \dots, n\}$  in the equation for the sample variance, we have the term  $X_i - \bar{X}$ :

$$\begin{aligned} X_i - \bar{X} &= X_i - \frac{1}{n}(X_1 + \dots + X_{i-1} + X_i + X_{i+1} + \dots + X_n) \\ &= \frac{(n-1)X_i}{n} - \frac{1}{n}(X_1 + \dots + X_{i-1} + X_{i+1} + \dots + X_n) \end{aligned}$$

Since the  $X_i$ 's are i.i.d. normally-distributed and the sample mean is also normally-distributed, it follows that their sum is also normally-distributed:

$$\begin{aligned} X_i - \bar{X} &\sim \mathcal{N}\left(\frac{n-1}{n}\mu - \frac{(n-1)}{n}\mu, \frac{(n-1)^2\sigma^2}{n^2} + \frac{(n-1)\sigma^2}{n^2}\right) \\ &\sim \mathcal{N}\left(0, \frac{\sigma^2}{n^2}((n-1)^2 + (n-1))\right) \\ &\sim \mathcal{N}\left(0, \frac{\sigma^2}{n^2}(n^2 - 2n + 1 + n - 1)\right) \\ &\sim \mathcal{N}\left(0, \frac{\sigma^2}{n^2}(n^2 - n)\right) \\ &\sim \mathcal{N}\left(0, \sigma^2\left(1 - \frac{1}{n}\right)\right) \\ &\sim \mathcal{N}\left(0, \frac{n-1}{n}\sigma^2\right) \end{aligned}$$

We also know  $\forall i \in \{1, \dots, n\}$ :

$$\begin{aligned} \text{Cov}(X_i - \bar{X}, \bar{X}) &= \text{Cov}(X_i, \bar{X}) - \text{Cov}(\bar{X}, \bar{X}) \\ &= \text{Cov}(X_i, \bar{X}) - \text{Var}(\bar{X}) \\ &= \mathbb{E}(X_i \bar{X}) - \mathbb{E}(X_i)\mathbb{E}(\bar{X}) - \text{Var}(\bar{X}) \\ &= \mathbb{E}\left(X_i \frac{1}{n}(X_1 + \dots + X_{i-1} + X_i + X_{i+1} + \dots + X_n)\right) - 0 \cdot \mu^2 - \frac{\sigma^2}{n} \\ &= \mathbb{E}\left(\frac{1}{n}(X_i X_1 + \dots + X_i X_{i-1} + X_i X_i + X_i X_{i+1} + \dots + X_i X_n)\right) - \frac{\sigma^2}{n} \\ &= \frac{1}{n}\left(\mathbb{E}(X_i)\mathbb{E}(X_1) + \dots + \mathbb{E}(X_i^2) + \dots + \mathbb{E}(X_i)\mathbb{E}(X_n)\right) - \frac{\sigma^2}{n} \\ &= \frac{1}{n}\left(0 + \dots + \text{Var}(X_i^2) + \dots + 0\right) - \frac{\sigma^2}{n} \\ &= \frac{\sigma^2}{n} - \frac{\sigma^2}{n} \\ &= 0 \end{aligned}$$

This implies that the covariance matrix  $\Sigma$  for the random vector  $V = (\bar{X}, X_1 - \bar{X}, \dots, X_n - \bar{X})^T$  is given by:

$$\Sigma = \begin{bmatrix} \text{Var}(\bar{X}) & \text{Cov}(\bar{X}, X_1 - \bar{X}) & \dots & \text{Cov}(\bar{X}, X_n - \bar{X}) \\ \text{Cov}(X_1 - \bar{X}, \bar{X}) & \text{Var}(X_1 - \bar{X}) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \text{Cov}(X_n - \bar{X}, \bar{X}) & \dots & \dots & \text{Var}(X_n - \bar{X}) \end{bmatrix}$$

Since  $\text{Cov}(X_i - \bar{X}, \bar{X}) = \text{Cov}(\bar{X}, X_i - \bar{X}) = 0$ ,  $\forall i \in \{1, \dots, n\}$ ,  $\Sigma$  is a diagonal matrix, and because it is also square, it follows that it is symmetric too, i.e.  $\Sigma = \Sigma^T$ . Since  $\Sigma$  is upper-triangular, it follows that the diagonal entries are its eigenvalues. Since all of the eigenvalues are variances, all of the eigenvalues are non-negative. Because of this and the fact that  $\Sigma$  is symmetric,  $\Sigma$  is also positive semi-definite. Therefore, it follows that  $V$  is multivariate normal with mean vector  $M = (\mu, \mu, \dots, \mu)^T$ .

Since  $V$  is multivariate normal, any pair of its components which are uncorrelated are also independent. Since we showed  $\text{Cov}(X_i - \bar{X}, \bar{X}) = \text{Cov}(\bar{X}, X_i - \bar{X}) = 0$ ,  $\forall i \in \{1, \dots, n\}$ , it follows that all pairs of random variables in  $V$  are independent. This also implies then that  $\bar{X}$  and  $U = (X_1 - \bar{X}, \dots, X_n - \bar{X})^T$  are independent. Since we know that  $U^T U = \sum_{i=1}^n (X_i - \bar{X})^2 = (n-1)S^2$ , where  $S^2$  is the sample variance, we have thus shown that  $\bar{X}$  and  $S^2$  are independent.  $\square$

The random variable  $T$  is the quotient of  $Z$  and  $X$ , where  $X$  and  $Z$  are independent by lemma 1, so its probability density function is given by the following, where  $f_Z(\cdot)$  is the probability density function of the standard normal random variable:

$$\begin{aligned} f_T(t) &= \int_{-\infty}^{\infty} |s| f_Z(ts) f_X(s) ds \\ &= \int_0^{\infty} s f_Z(ts) f_X(s) ds \\ &= \int_0^{\infty} s \left[ \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(ts)^2}{2}\right) \right] \left[ \frac{2^{1-\nu/2}}{\Gamma(\frac{\nu}{2})} \nu^{\frac{\nu}{2}} s^{\nu-1} \exp\left(-\frac{\nu s^2}{2}\right) \right] ds \\ &= \frac{1}{\sqrt{2\pi}} \frac{2^{1-\nu/2}}{\Gamma(\frac{\nu}{2})} \nu^{\frac{\nu}{2}} \int_0^{\infty} s \left[ \exp\left(-\frac{(ts)^2}{2}\right) \right] \left[ s^{\nu-1} \exp\left(-\frac{\nu s^2}{2}\right) \right] ds \\ &= \frac{1}{\sqrt{2\pi}} \frac{2^{1-\nu/2}}{\Gamma(\frac{\nu}{2})} \nu^{\frac{\nu}{2}} \int_0^{\infty} s^{\nu} \exp\left(\frac{-(ts)^2 - \nu s^2}{2}\right) ds \\ &= \frac{1}{\sqrt{2\pi}} \frac{2^{1-\nu/2}}{\Gamma(\frac{\nu}{2})} \nu^{\frac{\nu}{2}} \int_0^{\infty} s^{\nu} \exp\left(-\frac{(t^2 + \nu)s^2}{2}\right) ds \end{aligned}$$

We can use u-substitution, with  $u \triangleq s^2 \implies du = 2s ds \implies ds = \frac{du}{2s} \implies s = \sqrt{u}$ :

$$f_T(t) = \frac{1}{\sqrt{2\pi}} \frac{2^{1-\nu/2}}{\Gamma(\frac{\nu}{2})} \nu^{\frac{\nu}{2}} \frac{1}{2} \int_0^{\infty} u^{(\nu-1)/2} \exp\left(-\frac{(t^2 + \nu)u}{2}\right) du \quad (4)$$

The integrand of this integral has a similar form to the gamma probability density function  $f_G(\cdot)$ , for some arbitrary random variable  $G \sim \text{Gamma}(k, \theta)$ , where  $k$  is the shape parameter and  $\theta$  is the scale parameter:

$$f_G(u) = \frac{1}{\Gamma(k)\theta^k} u^{k-1} \exp\left(-\frac{u}{\theta}\right) \quad (5)$$

Thus, we can match parameters between our integrand and the gamma PDF to simplify our expression as follows:

$$\begin{aligned}\frac{1}{\theta} &= \frac{t^2 + \nu}{2} \implies \theta = \frac{2}{t^2 + \nu} \\ k - 1 &= (\nu - 1)/2 \implies k = \frac{\nu + 1}{2}\end{aligned}$$

Rearranging (5) and plugging in the above parameters, this implies that:

$$\begin{aligned}\Gamma(k)\theta^k f_G(u) &= u^{k-1} \exp\left(-\frac{u}{\theta}\right) \\ \Gamma\left(\frac{\nu + 1}{2}\right)\left(\frac{2}{t^2 + \nu}\right)^{\frac{\nu+1}{2}} f_G(u) &= u^{(\nu-1)/2} \exp\left(-\frac{(t^2 + \nu)u}{2}\right)\end{aligned}$$

Plugging back into (4) and simplifying, we obtain:

$$\begin{aligned}f_T(t) &= \frac{1}{\sqrt{2\pi}} \frac{2^{1-\nu/2}}{\Gamma(\frac{\nu}{2})} \nu^{\frac{\nu}{2}} \frac{1}{2} \int_0^\infty \Gamma\left(\frac{\nu + 1}{2}\right) \left(\frac{2}{t^2 + \nu}\right)^{\frac{\nu+1}{2}} f_G(u) du \\ &= \frac{1}{\sqrt{2\pi}} \frac{2^{1-\nu/2}}{\Gamma(\frac{\nu}{2})} \nu^{\frac{\nu}{2}} \frac{1}{2} \Gamma\left(\frac{\nu + 1}{2}\right) \left(\frac{2}{t^2 + \nu}\right)^{\frac{\nu+1}{2}} \underbrace{\int_0^\infty f_G(u) du}_{=1, \text{ gamma defined on } (0, \infty)} \\ &= \frac{1}{\sqrt{2\pi}} \frac{2^{1-\nu/2}}{\Gamma(\frac{\nu}{2})} \nu^{\frac{\nu}{2}} \frac{1}{2} \Gamma\left(\frac{\nu + 1}{2}\right) \left(\frac{2}{t^2 + \nu}\right)^{\frac{\nu+1}{2}} \\ &= \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})} \frac{1}{\sqrt{2\pi}} \frac{2^{1-\nu/2}}{2} \nu^{\frac{\nu}{2}} \frac{2^{\frac{\nu+1}{2}}}{(t^2 + \nu)^{\frac{\nu+1}{2}}} \\ &= \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})} \frac{1}{\sqrt{2\pi}} 2^{\nu/2} \nu^{\frac{\nu}{2}} \frac{(t^2 + \nu)^{-\frac{\nu+1}{2}}}{2^{-\frac{\nu+1}{2}}} \\ &= \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})} \frac{1}{\sqrt{2\pi}} \frac{2^{\nu/2}}{2^{-\frac{\nu+1}{2}}} \nu^{\frac{\nu}{2}} \left(\frac{t^2}{\nu} + \frac{\nu}{\nu}\right)^{-\frac{\nu+1}{2}} \\ &= \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})} \frac{\sqrt{2}}{\sqrt{2\pi}} \frac{\nu^{\frac{\nu}{2}}}{\nu^{-\frac{\nu+1}{2}}} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}} \\ &= \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})} \frac{1}{\sqrt{\pi}} \frac{1}{\sqrt{\nu}} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}} \\ &= \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi} \cdot \Gamma(\frac{\nu}{2})} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}\end{aligned}$$

We see that this final result is exactly equal to (1), thus proving the claim.  $\square$

## References

1. <https://statisticaloddsandends.wordpress.com/2019/08/25/proof-that-sample-mean-is-independent-of-sample-variance-under-normality/>
2. <https://math.stackexchange.com/questions/474733/derivation-of-the-density-function-of-student-t-distribution-from-this-big-integ>

3. [https://math.tntech.edu/ISR/Introduction\\_to\\_Probability/Distributions\\_of\\_Functions/thispage/node16.html](https://math.tntech.edu/ISR/Introduction_to_Probability/Distributions_of_Functions/thispage/node16.html)
4. <https://online.stat.psu.edu/stat414/node/174/>