Section 3.6. t- and F-Distributions

Note. In this section we give two more distributions, the Student's t-distribution and the F-distribution. These play a role in statistical inference.

Note 3.6.A. Let W and V be independent random variables where W has a standard normal distribution, $N(0,1)$, and V has a $\chi^2(r)$ distribution. The joint probability density function of W and V by Definition 2.4.1, so that

$$h(w,v) = \begin{cases} \frac{1}{\sqrt{2\pi}} e^{-w^2/2} \frac{1}{\Gamma(r/2)2^{r/2}} v^{r/2-1} e^{-v/2} & \text{for } -\infty < w < \infty, 0 < v < \infty \\ 0 & \text{elsewhere.} \end{cases}$$

Define random variable $T = W/\sqrt{V/r}$. As in Note 3.3.E and Example 3.3.6, we use the "transformation technique" and the Jacobian to find the probability density function $g_1(t)$ of T. Introduce u_1 and u_2 as

$$t = u_1(w,v) = \frac{w}{\sqrt{v/r}} \text{ and } u = u_2(w,v) = v.$$

We have

$$w = t\sqrt{v/r} = t\sqrt{u/r} = v_1(t,u) \text{ and } v = u = v_2(t,u),$$

and the Jacobian

$$J(w,v) = J(v_1,v_2) = \begin{vmatrix} \frac{\partial v_1}{\partial t} & \frac{\partial v_1}{\partial u} \\ \frac{\partial v_2}{\partial t} & \frac{\partial v_2}{\partial t} \end{vmatrix} = \begin{vmatrix} \frac{\partial w}{\partial t} & \frac{\partial w}{\partial u} \\ \frac{\partial v}{\partial t} & \frac{\partial v}{\partial t} \end{vmatrix}$$

$$= \begin{vmatrix} \sqrt{u/r} & \frac{t}{2\sqrt{u/r}} \\ 0 & 1 \end{vmatrix} = \sqrt{u/r} \neq 0 \text{ for } 0 < u < \infty.$$

So the transformation $(w, v) \mapsto (t, u)$ is one to one (injective; this is due to the nonzero Jacobian) and maps set $\mathcal{S} = \{(w, v) \mid -\infty < w < \infty, 0 < v < \infty\}$ one to one and onto $\mathcal{T} = \{(t, u) \mid -\infty < t < \infty, 0 < u < \infty\}$. As in Note 3.3.E, the joint probability density function of T and $U = V$ can be obtained from the joint probability density function of W and V by replacing w with $t\sqrt{u/r}$ and replacing v with u, and introducing a factor of $|J(v, w)| = \sqrt{u/r}$. This gives the joint probability density function of W and V on its support $\mathcal{T}$ as

$$\begin{aligned} g(t,u) &= \frac{1}{\sqrt{2\pi}} e^{-(t\sqrt{u/r})^2/2} \frac{1}{\Gamma(r/2)2^{r/2}} u^{r/2-1} e^{-u/2} \sqrt{u/r} \\ &= \frac{1}{\sqrt{2\pi r}\Gamma(r/2)2^{r/2}} u^{(r+1)/2-1} e^{-\frac{u}{2}-\frac{t^2 u}{2r}}, \end{aligned}$$

and $g(t, u) = 0$ elsewhere. The marginal probability function of T is then

$$\begin{aligned} g_1(t) &= \int_{-\infty}^{\infty} g(t,u) = \int_0^{\infty} \frac{1}{\sqrt{2\pi r}\Gamma(r/2)2^{r/2}} u^{(r+1)/2-1} e^{-\frac{u}{2}(1+t^2/r)}\,du \\ & \qquad [\text{Let } z = u(1+t^2/r)/2 \text{ and } dz = (1+t^2/r)/2\,du] \\ &= \int_0^{\infty} \frac{1}{\sqrt{2\pi r}\Gamma(r/2)2^{r/2}} \left(\frac{2z}{1+t^2/r}\right)^{(r+1)/2-1} e^{-z} \frac{2}{1+t^2/r}\,dz \\ &= \frac{\sqrt{2}}{\sqrt{2\pi r}\Gamma(r/2)(1+t^2/r)^{(r+1)/2}} \int_0^{\infty} z^{(r+1)/2-1} e^{-z}\,dz \\ &= \frac{\Gamma((r+1)/2)}{\sqrt{\pi r}\Gamma(r/2)(1+t^2/r)^{(r+1)/2}} \text{ for } -\infty < t < \infty. \end{aligned}$$

That is, if W is $N(0, 1)$, V is $\chi^2(r)$, and W and V are independent, then $T = W/\sqrt{V/r}$ has the probability function $g_1(t)$. This distribution is the t-distribution.

Definition. A random variable T with probability density function

$$g(t) = \frac{\Gamma((r+1)/2)}{\sqrt{\pi r}\Gamma(r/2)(1+t^2/r)^{(r+1)/2}},\ -\infty < t < \infty,$$

(where $r > 0$) has a *t-distribution* with r degrees of freedom.

Note. The only appearance of t in the probability density function of a t-distribution only involves the term t^2. So $g(t)$ is symmetric about 0 and hence the median is 0. Differentiation shows that g has a single maximum at $t = 0$. We will show in Section 5.2. Convergence in Distribution (using the Lebesgue Dominated Convergence Theorem; see Example 5.2.3) that the limit as the degrees of freedom $r \to \infty$ of $g(t)$ is the standard normal distribution $N(0, 1)$.

Note. William S. Gosset (June 13, 1876–October 16, 1937) attended New College Oxford and received a mathematics degree in 1897 and a chemistry degree in 1899. He started working as a chemist at Arthur Guinness Son and Company in 1899; this is the Guinness brewery in Dublin, Ireland. Gosset was involved in statistical analysis. He studied with Karl Pearson at University College in London in 1906–07. The Guinness Board of Directors allowed its scientists to publish, as long as they did not mention beer, Guinness, or their own surname. Gosset had record data in a notebook with "The Student's Science Notebook" printed on the front. As a result, Gosset submitted his papers under the name "Student" (this story appears in Stephen Ziliak's "Guinnessometrics: The Economic Foundations of 'Student's' t" *Journal of Econometric Perspectives*, **22**(4), 199–216 (2008), a copy of which is available on the American Economic Association webpage [Accessed 7/17/2021]; see page 203). He presented the t-distribution (or "Student's t-distribution," but surprisingly not "Gossett's t-distribution") in "The Probable Error of a Mean," *Biometrika* **6**(1), 1-25 (March 1908). The original paper is available online from JSTOR. The t-distribution is useful in estimating the mean of a normally distributed population when sample size is small and when the population standard deviation is unknown. Gosset published 21 papers and collaborated with Fisher, Neyman, and Pearson.

Gossett image from MacTutor History of Mathematics Archive Gossett biography webpage and Guinness image from E-Bay.

This biographical information is from MacTutor History of Mathematics Archive Gossett biography webpage and the Wikipedia page on Gosset.

Example 3.6.1. Let random variable T have a t-distribution with r degrees of freedom. As shown in Note 3.6.A, $T = W(V/r)^{-1/2}$ where W has a standard normal, $N(0,1)$, distribution, and V has a $\chi^2(r)$ distribution and W and V are independent. If $k < r$ (and so $(-k/2) > -r/2$) then

$$\begin{aligned} E(T^k) &= E\left[W^k\left(\frac{V}{r}\right)^{-k/2}\right] \\ &= E(W^k)E\left[\left(\frac{V}{r}\right)^{-k/2}\right] \text{ by Theorem 2.4.4, since } W \text{ and } V \text{ are independent} \\ &= E(W^k)\frac{1}{r^{-k/2}}E(V^{-k/2}) \text{ by the linearity of expectation, Theorem 1.8.2} \end{aligned}$$

$$\begin{aligned} &= E(W^k)\frac{1}{r^{-k/2}}\frac{2^{-k/2}\Gamma(r-k/2)}{\Gamma(r/2)} \text{ by Theorem 3.3.2} \\ &\quad \text{(with } k \text{ of Theorem 3.3.2 replaced with } -k/2 \text{ here)} \\ &= E(W^k)\frac{2^{-k/2}\Gamma(r/2-k/2)}{\Gamma(r/2)r^{-k/2}} \text{ if } k < r. \end{aligned}$$

Since $E(W) = 0$ then $E(T) = 0$ (from the inequality with $k = 1$), as long as $r > 1$. With $k = 2$ in this inequality, since $E(W^2) = 1$ by Example 3.3.4 (with $j = 2$), we have

$$\begin{aligned} \text{Var}(T) &= E((T-E(T))^2) = E(T^2) = (1)\frac{2^{-1}\Gamma(r/2-1)}{\Gamma(r/2)r^{-1}} \\ &= \frac{r\Gamma(r/2-1)}{2(r/2-1)\Gamma(r/2-1)} \text{ since } \Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1) \text{ for } \alpha > 0 \\ &= \frac{r}{2(r/2-1)} = \frac{r}{r-2}, \end{aligned}$$

where $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$ for $\alpha > 0$ is given in *Thomas' Calculus, Early Transcendentals*, 14th Edition, "Chapter 8, Techniques of Integration," "Additional and Advanced Exercises" Number 43. That is, a t-distribution with $r > 2$ degree of freedom has a mean of 0 and a variance of $r/(r-2)$.

Note 3.6.B. To introduce the F-distribution, consider two independent chi-square variables U and V with r_1 and r_2 degrees of freedom, respectively. Since U and V are independent, then by Definition 2.4.1, the joint probability density function of U and V is a product of two chi-square probability density functions:

$$h(u,v) = \begin{cases} \frac{1}{\Gamma(r_1/2)\Gamma(r_2/2)2^{(r_1+r_2)/2}} u^{r_1/2-1} v^{r_2/2-1} e^{-(u+v)/2} & \text{where } 0 < u, v, < \infty \\ 0 & \text{elsewhere.} \end{cases}$$

Define random variable $W = \dfrac{U/r_1}{V/r_2}$. We now find the marginal probability density function $g_1(w)$ of W. As above, for the t-distribution (and as in Note 3.3.E and

Example 3.3.6) we use the "transformation technique." Introduce u_1 and u_2 as

$$w = u_1(u,v) = \frac{u/r_1}{v/r_2} \text{ and } z = u_2(u,v) = v.$$

We have

$$u = (r_1/r_2)vw = (r_1/r_2)zw = v_1(w,z) \text{ and } v = z = v_2(w,z),$$

and the Jacobian

$$J(u,v) = J(v_1,v_2) = \begin{vmatrix} \frac{\partial v_1}{\partial w} & \frac{\partial v_1}{\partial z} \\ \frac{\partial v_2}{\partial w} & \frac{\partial v_2}{\partial z} \end{vmatrix} = \begin{vmatrix} \frac{\partial u}{\partial w} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial w} & \frac{\partial v}{\partial z} \end{vmatrix} = \begin{vmatrix} (r_1/r_2)z & (r_1/r_2)w \\ 0 & 1 \end{vmatrix}$$
$$= (r_1/r_2)z \neq 0 \text{ for } 0 < z < \infty.$$

So the transformation $(u,v) \mapsto (w,z)$ is one to one (injective; this is due to the nonzero Jacobian) and maps set $\mathcal{S} = \{(u,v) \mid 0 < u < \infty, 0 < v < \infty\}$ onto the set $\mathcal{T} = \{(w,z) \mid 0 < w < \infty, 0 < z < \infty\}$. As in Note 3.3.E, the joint probability density function of W and $Z = V$ can be obtained from the joint probability density function of U and V by replacing u with $(r_1/r_2)zw$, replacing v with z, and introducing a factor of $|J(v,w)| = (r_1/r_2)z$. This gives the joint probability density function of W and Z on its support $\mathcal{T}$ as

$$g(w,z) = \frac{1}{\Gamma(r_1/2)\Gamma(r_2/2)2^{(r_1+r_2)/2}} \left(\frac{r_1 zw}{r_2}\right)^{(r_1-2)/2} z^{(r_2-2)/2} e^{-\left(\frac{r_1 zw}{r_2}+z\right)/2} \frac{r_1 z}{r_2}$$

and $g(wmz) = 0$ elsewhere. The marginal probability density function of W is then

$$\begin{aligned} g_1(w) &= \int_{-\infty}^{\infty} g(w,z)\,dz \\ &= \int_0^{\infty} \frac{(r_1/r_2)^{r_1/2} w^{r_1/2-1}}{\Gamma(r_1/2)\Gamma(r_2/2)2^{(r_1+r_2)/2}} z^{(r_1+r_2)/2-1} e^{-\frac{z}{2}\left(\frac{r_1 w}{r_2}+1\right)}\,dz \\ &\quad \left[\text{let } t = \frac{z}{2}\left(\frac{r_1 w}{r_2}+1\right) \text{ and } dy = \frac{1}{2}\left(\frac{r_1 w}{r_2}+1\right)dz\right] \end{aligned}$$

$$\begin{aligned}
&= \frac{(r_1/r_2)^{r_1/2} w^{r_1/2-1}}{\Gamma(r_1/2)\Gamma(r_2/2)2^{(r_1+r_2)/2}} \int_0^\infty \left(\frac{2y}{r_1 w/r_2+1}\right)^{(r_1+r_2)/2-1} e^{y} \left(\frac{2}{r_1 w/r_2+1}\right) dy \\
&= \frac{(r_1/r_2)^{r_1/2} w^{r_1/2-1}}{\Gamma(r_1/2)\Gamma(r_2/2)2^{(r_1+r_2)/2}} \left(\frac{2}{r_1 w/r_2+1}\right)^{(r_1+2_2)/2} \int_0^\infty y^{(r_1+r_2)/2-1} e^{-y}\, dy \\
&= \begin{cases} \frac{\Gamma((r_1+r_2)/2)(r_1/r_2)^{r_1/2}}{\Gamma(r-1/2)\Gamma(r_2/2)} \frac{w^{r_1/2-1}}{(1+r_1 w/r_2)^{(r_1+r_2)/2}} & \text{for } 0 < w < \infty \\ 0 & \text{elsewhere.} \end{cases}
\end{aligned}$$

That is, if U and V are independent chi-square variables with r_1 and r_2 degrees of freedom, respectively, then $W = (U/r_1)/(V/r_2)$ has the probability density function $g_1(w)$. This distribution is an F-distribution.

Definition. A random variable W with probability density function

$$g(w) = \begin{cases} \frac{\Gamma((r_1+r_2)/2)(r_1/r_2)^{r_1/2}}{\Gamma(r-1/2)\Gamma(r_2/2)} \frac{w^{r_1/2-1}}{(1+r_1 w/r_2)^{(r_1+r_2)/2}} & \text{for } 0 < w < \infty \\ 0 & \text{elsewhere} \end{cases}$$

(where $r_1, r_2 > 0$) has an F*-distribution.*

Example 3.6.2. Let F have an F-distribution with r_1 and r_2 degrees of freedom. As shown in Note 3.6.B, we have $F = (r_2/r_1)(U/V)$ where U and V are independent χ^2 random variables with r_1 and r_2 degrees of freedom, respectively. Since U and V are independent, the by Theorem 2.4.4

$$E(F^k) = E\left(\left(\frac{r_2}{r_1}\frac{U}{V}\right)^k\right) = \left(\frac{r_2}{r_1}\right)^k E(U^k)E(V^{-k}),$$

provided $E(U^k)$ and $E(V^{-k})$ exists. Since $k > -(r_1/2)$ (because $k \in \mathbb{N}$ and $r_1 > 0$), then by Theorem 3.3.2, $E(U^k)$ exists. Also by Theorem 3.3.2, $E(V^{-k})$ exists provided $-k > -r_2/2$ (or $2k < r_2$). If this holds, then by Theorem 3.3.2 with

$k = 1$, the mean of F is

$$\begin{aligned}
E(F) &= \frac{r_2}{r_1} E(U) E(V^{-1}) \\
&= \frac{r_2}{r_1} r_1 \frac{2^{-1}\Gamma(r_2/2 - 1)}{\Gamma(r_2/2)} \text{ since the mean of } \chi^2(r) \text{ is } r \\
&= \frac{r_2}{2} \frac{\Gamma(r_2/2 - 1)}{(r_2/2 - 1)\Gamma(r_2/2 - 1)} \text{ since } \Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1) \text{ for all } \alpha > 0 \\
&= \frac{r_2}{r_2 - 2}.
\end{aligned}$$

Notice that for r_2 large then $E(F)$ is near 1.

Theorem 3.6.1. Student's Theorem.

Let $X_1, X_2, \ldots, X_n$ be identical in distribution ("iid") random variables each having a normal distribution with mean μ and variance σ^2. Define the random variables

$$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \text{ and } S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2.$$

Then

(a) $\overline{X}$ has a $N(\mu, \sigma^2/n)$ distribution.

(b) $\overline{X}$ and S^2 are independent.

(c) $(n-1)S^2/\sigma^2$ has a $\chi^2(n-1)$ distribution.

(d) The random variable $T = \dfrac{\overline{X} - \mu}{S/\sqrt{n}}$ has a Student t-distribution with $n - 1$ degrees of freedom.