Derivation of Student's T Distribution II

Part I: The F and T Distributions

Bruce M. Boghosian

Department of Mathematics

Tufts University

Outline

In our last episode. . .

Bruce M. Boghosian

- We reviewed the gamma and beta functions.
  - The gamma function $\Gamma(r)$ has one argument, r.
  - The beta function $B(r, s)$ has two arguments, r and s.
- We reviewed the gamma and beta distributions.
  - The gamma distribution has two parameters, r and λ.
  - The beta distribution has two parameters, r and s.
- Sum of squares of n iid $N(0, 1)$ r.v.s is gamma distributed
  - Special case of gamma distribution with $r = n/2$ and $\lambda = 1/2$
  - Called "chi squared distribution with n degrees of freedom"

In our last episode. . . I

Bruce M. Boghosian

- Suppose that
  - $Y_1, \ldots, Y_n$ are iid $N(\mu, \sigma^2)$ r.v.s.
  - $\overline{Y} = \frac{1}{n}\sum_{j=1}^n Y_j$
  - $S = \frac{1}{n-1}\frac{1}{n}\sum_{j=1}^n \left(Y_j - \overline{Y}\right)^2$
  - $X_j := \frac{Y_j-\mu}{\sigma}$ is standardized version of Y_j
  - $X_1, \ldots, X_n$ are iid $N(0,1)$ r.v.s.
- We then used some linear algebra to show
  - We can transform from $\langle X_1, \ldots, X_n\rangle$ to $\langle Z_1, \ldots, Z_n\rangle$
  - where $Z_1, \ldots, Z_n$ are also iid $N(0,1)$ r.v.s.
  - and where $Z_n = \sqrt{n}\,\overline{X} = \frac{\overline{Y}-\mu}{\sigma/\sqrt{n}}$

In our last episode... II

- It then follows that

$$\sum_{j=1}^{n} Z_j^2 = \sum_{j=1}^{n-1} Z_j^2 + n\overline{X}^2 = \sum_{j=1}^{n} \left(X_j - \overline{X}\right)^2 + n\overline{X}^2.$$

- Hence

$$\sum_{j=1}^{n-1} Z_j^2 = \sum_{j=1}^{n} \left(X_j - \overline{X}\right)^2 = (n-1)S^2$$

- So S^2 depends only on $Z_1\, ldots, Z_{n-1}$, and is independent of

$$Z_n = \sqrt{n}\,\overline{X} = \frac{\overline{Y} - \mu}{\sigma/\sqrt{n}}.$$

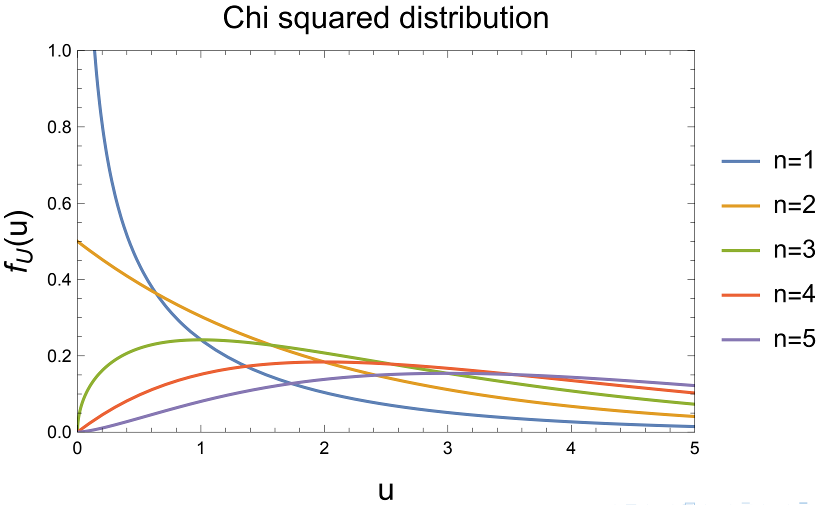
- Hence S^2 is independent of $\overline{X}$.
- Hence S^2 is independent of $\overline{Y}$.

In our last episode. . . III

Bruce M. Boghosian

- We have
$$\sum_{j=1}^{n-1} Z_j^2 = \sum_{j=1}^{n} \left(X_j - \overline{X}\right)^2 = (n-1)S^2$$
- Hence $\frac{(n-1)S^2}{\sigma^2}$ is chi squared distributed with $n-1$ degrees of freedom.
- In summary: If $Y_1, \dots, Y_n$ is a random sample from $N(\mu, \sigma^2)$, then
  - S^2 and $\overline{Y}$ are independent
  - $\frac{(n-1)S^2}{\sigma^2}$ is chi squared distributed with $n-1$ degrees of freedom.

Plotting the chi squared distribution



Finding the pdf of a quotient of random variables I

- **Lem.:** Let X and Y have pdfs $f_X(x)$ and $f_Y(y)$, respectively. Assume that $X = 0$ for at most a set of isolated points. Let $W = Y/X$. Then

$$f_W(w) = \int_{-\infty}^{+\infty} dx\, |x|\, f_X(x) f_Y(wx).$$

- **Pf.:** First compute the cdf:

$$\begin{aligned}
F_W(w) =& P(Y/X \leq w) \\
=& P(Y/X \leq w,\, X \geq 0) + P(Y/X \leq w,\, X < 0) \\
=& P(Y \leq wX,\, X \geq 0) + P(Y \geq wX,\, X < 0) \\
=& P(Y \leq wX,\, X \geq 0) + 1 - P(Y \leq wX,\, X < 0) \\
=& \int_0^{\infty} dx \int_{-\infty}^{wx} dy\ f_X(x) f_Y(y) + 1 - \int_{-\infty}^{0} dx \int_{-\infty}^{wx} dy\ f_X(x) f_Y(y)
\end{aligned}$$

Finding the pdf of a quotient of random variables II

- **Pf. (continued):** We have

$$F_W(w) = \int_0^\infty dx \int_{-\infty}^{wx} dy\ f_X(x) f_Y(y) + 1 - \int_{-\infty}^{0} dx \int_{-\infty}^{wx} dy\ f_X(x) f_Y(y)$$

- Differentiate with respect to w

$$f_W(w) = \int_0^\infty dx\ x\ f_X(x) f_Y(wx) + \int_{-\infty}^{0} dx\ (-x)\ f_X(x) f_Y(wx)$$

- So we finally have that which was to be shown

$$\boxed{f_W(w) = \int_{-\infty}^{+\infty} dx\ |x|\ f_X(x) f_Y(wx)} \quad \square$$

Definition of the F distribution

- **Def.:** Suppose that U and V are independent chi squared r.v.s with n and m degrees of freedom, respectively. A random variable of the form $\frac{V/m}{U/n}$ is said to have an *F distribution with m and n degrees of freedom*.
- **Thm.:** Suppose $F_{m,n} = \frac{V/m}{U/n}$ denotes an F r.v. with m and n degrees of freedom. The pdf of $F_{m,n}$ has the form

$$\boxed{f_{F_{m,n}}(w) = \frac{\Gamma\left(\frac{m+n}{2}\right) m^{m/2} n^{n/2} w^{(m/2)-1}}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)(n+mw)^{(m+n)/2}} = \frac{m^{m/2} n^{n/2} w^{(m/2)-1}}{B\left(\frac{m}{2},\frac{n}{2}\right)(n+mw)^{(m+n)/2}}}$$

for $w \geq 0$.

Proof of the form of the F distribution I

- **Thm.:** The pdf of r.v. $F_{m,n} = \frac{V/m}{U/n}$ is $f_{F_{m,n}}(w)$, given on last slide.
- **Pf.:** We know that U and V are chi squared distributed with n and m df, resp.

$$f_V(v) = \frac{1}{2^{m/2}\Gamma(m/2)} v^{(m/2)-1} e^{-v/2} \quad \text{and} \quad f_U(u) = \frac{1}{2^{n/2}\Gamma(n/2)} u^{(n/2)-1} e^{-u/2}$$

- Begin by finding the pdf of $\frac{V}{U}$, using theorem for pdf of a quotient

$$\begin{aligned} f_{V/U}(w) &= \int_0^\infty du\ |u|\ f_U(u) f_V(wu) \\ &= \int_0^\infty du\ u\ \frac{1}{2^{n/2}\Gamma(n/2)} u^{(n/2)-1} e^{-u/2} \frac{1}{2^{m/2}\Gamma(m/2)} (wu)^{(m/2)-1} e^{-wu/2} \end{aligned}$$

Proof of the form of the F distribution II

- Continuing

$$
\begin{aligned}
f_{V/U}(w) &= \frac{w^{(m/2)-1}}{2^{(n+m)/2}\Gamma(n/2)\Gamma(m/2)} \int_0^\infty du\ u^{\frac{n+m}{2}-1} e^{-[(w+1)/2]u} \\
&= \frac{w^{(m/2)-1}\left(\frac{w+1}{2}\right)^{-(n+m)/2}}{2^{(n+m)/2}\Gamma(n/2)\Gamma(m/2)} \int_0^\infty dz\ z^{\frac{n+m}{2}-1} e^{-z} \\
&= \frac{w^{(m/2)-1}\left(w+1\right)^{-(n+m)/2}}{\Gamma(n/2)\Gamma(m/2)} \int_0^\infty dz\ z^{\frac{n+m}{2}-1} e^{-z} \\
&= \frac{\Gamma\left(\frac{n+m}{2}\right)}{\Gamma(n/2)\Gamma(m/2)} \frac{w^{(m/2)-1}}{(w+1)^{(n+m)/2}}
\end{aligned}
$$

Proof of the form of the F distribution II—

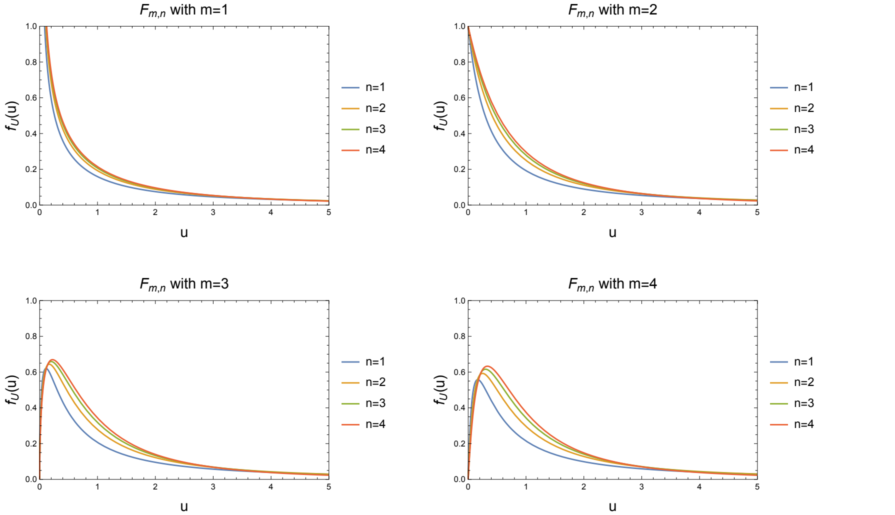
- We have

$$f_{V/U}(w) = \frac{\Gamma\left(\frac{n+m}{2}\right)}{\Gamma(n/2)\Gamma(m/2)} \frac{w^{(m/2)-1}}{(w+1)^{(n+m)/2}}$$

- Finally, if we include the scaling, we arrive at what was to be shown

$$\begin{aligned} f_{\frac{V/m}{U/n}}(w) =& f_{\frac{n}{m}\frac{V}{U}}(w) = \frac{1}{n/m} f_{V/U}\left(\frac{w}{n/m}\right) = \frac{m}{n} f_{V/U}\left(\frac{m}{n}w\right) \\ =& \frac{m}{n} \frac{\Gamma\left(\frac{n+m}{2}\right)}{\Gamma(n/2)\Gamma(m/2)} \frac{\left(\frac{mw}{n}\right)^{(m/2)-1}}{\left(\frac{mw}{n}+1\right)^{(n+m)/2}} \\ =& \frac{\Gamma\left(\frac{m+n}{2}\right) m^{m/2} n^{n/2} w^{(m/2)-1}}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)(n+mw)^{(m+n)/2}} \qquad \square \end{aligned}$$

Plotting the F distribution



Definition and symmetry of the T distribution

- **Def.:** Let Z be an r.v. sampled from $N(0,1)$, and let U be a chi squared r.v., independent of Z with n degrees of freedom. The Student T ratio with n df is denoted T_n where $T_n = \frac{Z}{\sqrt{U/n}}$.
- **Lem.:** The pdf $f_{T_n}(t)$ is a symmetric function. That is $\forall t: \; f_{T_n}(-t) = f_{T_n}(t)$.
- **Pf.:** We know f_Z is symmetric. Let $V = \sqrt{U/n}$ so

$$f_{T_n}(t) = \int_0^\infty dv \; f_V(v) f_Z(tv) = \int_0^\infty dv \; f_V(v) f_Z(-tv) = f_{T_n}(-t) \qquad \square$$

The pdf of the T distribution I

- **Thm.:** The pdf for a Student t random variable with n degrees of freedom is given by

$$f_{T_n}(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)\left(1+\frac{t^2}{n}\right)^{(n+1)/2}} \qquad \text{for } t \in \mathbb{R}.$$

- **Pf.:** Note that $T_n^2 = \frac{Z^2}{U/n}$ has an F distribution with 1 and n df. Hence,

$$f_{T_n^2}(t) = \frac{n^{n/2}\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{n}{2}\right)} \frac{t^{-1/2}}{(n+t)^{(n+1)/2}}$$

The pdf of the T distribution II

- **Pf. (continued):** By the symmetry of T_n, we have

$$\begin{aligned} F_{T_n} = P(T_n \le t) &= \frac{1}{2} + P(0 \le T_n \le t) \\ &= \frac{1}{2} + \frac{1}{2} P(-t \le T_n \le t) \\ &= \frac{1}{2} + \frac{1}{2} P(0 \le T_n^2 \le t^2) \\ &= \frac{1}{2} + \frac{1}{2} F_{T_n^2}(t^2) \end{aligned}$$

- Differentiating yields

$$f_{T_n}(t) = \frac{1}{2} f_{T_n^2}(t^2)(2t) = t\, f_{T_n^2}(t^2).$$

The pdf of the T distribution III

- **Pf. (continued):** We may now complete the proof as follows

$$
\begin{aligned}
f_{T_n}(t) &= t\, f_{T_n^2}(t^2) \\
&= t\, \frac{n^{n/2}\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{n}{2}\right)} \frac{(t^2)^{-1/2}}{(n+t^2)^{(n+1)/2}} \\
&= \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)} \frac{1}{\left(1+\frac{t^2}{n}\right)^{(n+1)/2}}, \qquad \square
\end{aligned}
$$

as was to be shown.

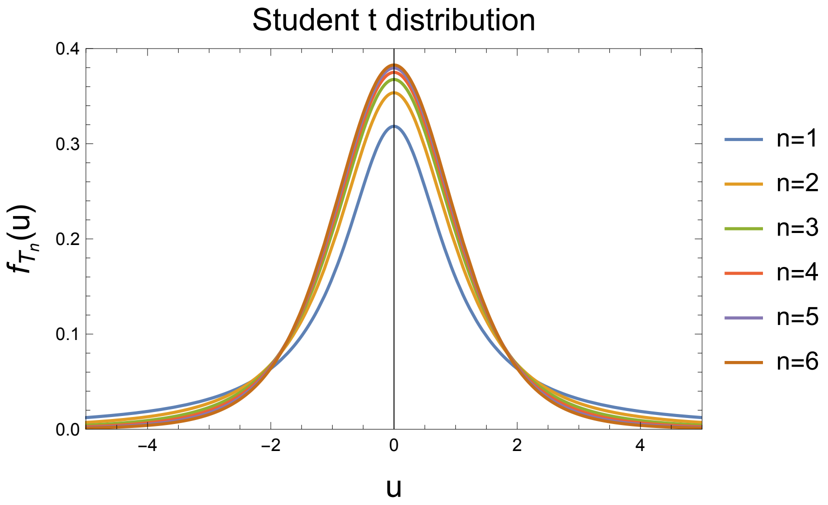
Demonstrating that the Y_j are Student t distributed

- **Thm.:** Let $Y_1, \ldots, Y_n$ be a random sample, each from $N(\mu, \sigma^2)$. Then, $T_{n-1} = \frac{\overline{Y}-\mu}{S/\sqrt{n}}$ has a Student t distribution with $n-1$ df.
- **Pf.:** Note that we can write

$$\frac{\overline{Y}-\mu}{S/\sqrt{n}} = \frac{\frac{\overline{Y}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{(n-1)S^2}{\sigma^2(n-1)}}}$$

- Numerator is in $N(0,1)$ and denominator has a chi squared distribution with $n-1$ df.
- Moreover, numerator and denominator are independent r.v.s.
- The statement of the theorem follows instantly. $\square$

Plotting the T distribution



Tables in the appendices

- You are now in a position to understand more of the tables in the back of the book.
- Table A.2 gives percentiles of the student T_n for various values of n and levels of confidence.
- Table A.3 does the same thing for chi squared distributions.
- Table A.4 does the same thing for F distributions.

Summary

- We have completed the derivation of the student T distribution.
- It can be used for sampling small-sample, normally distributed data.
- It is tabulated and it is possible to use it for interval estimation, hypothesis testing, etc.