Chapter 5

Multiple Random Variables

$t$ Distribution
Outline

Studentization

The $t$ Distribution

Introduction to Asymptotic Analysis
Studentization

Given a independent sequence $X_1, \ldots, X_n$ with common distribution $F_x$, mean $\mu$, and standard deviation $\sigma$. Then, the standardized version of the sample mean $\bar{X}$

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}.$$ 

In most circumstances, the standard deviation is not known and thus we replace $Z$ with

$$T = \frac{\bar{X} - \mu}{S / \sqrt{n}}.$$ 

where $S$ is the sample standard deviation. The estimate $S / \sqrt{n}$ of the standard deviation of $\bar{X}$ is called the standard error. $T$ is called the studentized version of $\bar{X}$.

Our goal is to give the density of $T$ in the case that $X_1, \ldots, X_n$ are normally distributed.
The t Distribution

We begin by standardizing the observations $X_1, \ldots, X_n$.

$$Z_i = \frac{X_i - \mu}{\sigma}.$$  

Then,

$$T = \frac{\bar{Z}}{S/\sqrt{n}}.$$  

where

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (Z_i - \bar{Z})^2.$$
The \( t \)-Distribution

We next want to establish three facts

1. \( \bar{Z} \sim N(0, 1/n) \)

2. \( \sum_{i=1}^{n}(Z_i - \bar{Z})^2 \sim \chi^2_{n-1} \), the chi-square distribution with \( n - 1 \) degrees of freedom.

3. \( \bar{Z} \) and \( \sum_{i=1}^{n}(Z_i - \bar{Z})^2 \) are independent.

We have previously established 1. Moving on to 2 and 3, notice that for any \( i = 1, \ldots, n \),

\[
\text{Cov}(\bar{Z}, Z_i - \bar{Z}) = \text{Cov}(\bar{Z}, Z_i) - \text{Var}(\bar{Z})
\]

\[
= \frac{1}{n} \sum_{j=1}^{n} \text{Cov}(Z_j, Z_i) - \frac{1}{n} = \frac{1}{n} - \frac{1}{n} = 0.
\]

If we can show that, in this case, zero covariance implies independence, then we can conclude 3.
The \( t \) Distribution

As for 2, we return to the identity

\[
\sum_{i=1}^{n} (Z_i - \bar{Z})^2 = \sum_{i=1}^{n} Z_i^2 - n\bar{Z}^2.
\]

\[
\sum_{i=1}^{n} (Z_i - \bar{Z})^2 + n\bar{Z}^2 = \sum_{i=1}^{n} Z_i^2.
\]

The terms in the sum on the left are independent.

\[
\sum_{i=1}^{n} Z_i^2 \sim \chi_n^2, \quad \text{and} \quad n\bar{Z}^2 \sim \chi_1^2,
\]

Recall that the chi-square distribution with \( \nu \) degrees of freedom has moment generating function \( M_{\chi^2_\nu} t(t) = (1 - 2t)^{-\nu/2} \).
The \( t \) Distribution

Thus, we take the moment generating function of

\[
\sum_{i=1}^{n}(Z_i - \bar{Z})^2 + n(\bar{Z})^2 = \sum_{i=1}^{n} Z_i^2,
\]

to obtain

\[
M_{\sum_{i=1}^{n}(Z_i - \bar{Z})^2(t)} M_{n\bar{Z}^2(t)} = M_{\sum_{i=1}^{n} Z_i^2(t)},
\]

\[
M_{\sum_{i=1}^{n}(Z_i - \bar{Z})^2(t)}(1 - 2t)^{-1/2} = (1 - 2t)^{-n/2}
\]

\[
M_{\sum_{i=1}^{n}(Z_i - \bar{Z})^2(t)} = (1 - 2t)^{-(n-1)/2}.
\]

Because the moment generating function is one-to-one,

\[
\sum_{i=1}^{n}(Z_i - \bar{Z})^2 \sim \chi_{n-1}^2,
\]

establishing 3.
Linear Transformations of a Multivariate Normal

Our next step is to verify that a linear transformation of a multivariate normal is also normal. For $X$, a centered multivariate normal random variable, its density can be expressed by

$$f_X(x) = \frac{1}{\sqrt{\det \Sigma} (2\pi)^n} \exp \left( -\frac{x^T \Sigma^{-1} x}{2} \right)$$

where $\Sigma$ is a nonnegative definite, symmetric matrix.

Now let $Y = AX$ for $A$, an $n \times p$ matrix. Then,

$$\Sigma_Y = \text{Cov}(Y) = A^T \text{Cov}(X) A = A^T \Sigma A$$

$\Sigma_Y$ is a covariance matrix and so must be a nonnegative definite, symmetric matrix. As long as $\Sigma_Y$ is invertible, we can write

$$f_Y(y) = \frac{1}{\sqrt{\det \Sigma_Y} (2\pi)^p} \exp \left( -\frac{y^T \Sigma_Y^{-1} y}{2} \right)$$

and is a centered multivariate normal density,
Linear Transformations of a Multivariate Normal

Now consider the linear transformation \( Y = AZ \)

\[
(Z_1, \ldots, Z_n) \mapsto (Z_i, Z_i - \bar{Z}) = (Y_1, Y_2)
\]

Then,

\[
\Sigma_Y = \text{Cov}(Y) = \begin{pmatrix} \sigma_1^2 & \rho \sigma_1 \sigma_2 \\ \rho \sigma_1 \sigma_2 & \sigma_2^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 + 1/n \end{pmatrix}
\]

\[
\Sigma_Y^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & n/(n+1) \end{pmatrix}, \quad \det(\Sigma_Y) = \frac{n+1}{n}
\]

and

\[
f_Y(y) = \sqrt{\frac{n}{(n+1)(2\pi)^2}} \exp -\frac{1}{2} \left( y_1^2 + (n+1)^2 y_2^2 / n^2 \right)
\]

and the density factors into a function of \( y_1 \) and \( y_2 \). Indeed, this shows that any uncorrelated random variables in a multivariate normal are independent.
The $t$ Distribution

Taking stock, we can now write

$$T = \frac{\bar{X} - \mu}{S_X / \sqrt{n}} = \frac{\bar{Z}}{S_Z / \sqrt{n}} = \frac{Y_1}{\sqrt{Y_2 / \nu}}.$$ 

where

- $Y_1 \sim N(0, 1)$, $Y_2 \sim \chi^2_{\nu}$,
- $\nu = n - 1$, $Y_1$ and $Y_2$ are independent.

The joint density

$$f_{Y_1, Y_2}(y_1, y_2) = f_{Y_1}(y_1) \cdot f_{Y_2}(y_2) = \frac{1}{\sqrt{2\pi}} e^{-y_1^2/2} \cdot \frac{1}{\Gamma(\nu/2)2^{\nu/2}} y_2^{\nu/2-1} e^{-y_2/2}.$$

To find the density for $T$, we use the transformation

$$(t, u) = g(y_1, y_2) = \left( \frac{y_1}{\sqrt{y_2 / \nu}}, y_2 \right) \quad \text{and} \quad (y_1, y_2) = g^{-1}(t, u) = (t \sqrt{u / \nu}, u).$$
The \( t \) Distribution

With \((y_1, y_2) = g^{-1}(t, u) = (t \sqrt{u/\nu}, u)\), we have the Jacobian matrix

\[
J(t, u) = \begin{pmatrix}
\frac{\partial y_1}{\partial t} & \frac{\partial y_1}{\partial u} \\
\frac{\partial y_2}{\partial t} & \frac{\partial y_2}{\partial u}
\end{pmatrix} = \begin{pmatrix}
\sqrt{u/\nu} & t/(2\sqrt{uv}) \\
0 & 1
\end{pmatrix}, \quad \det J(u, \nu) = \sqrt{\frac{u}{\nu}}
\]

Then,

\[
f_{T,U}(t, u) = f_{Y_1}(t \sqrt{u/\nu}) \cdot f_{Y_2}(u) \cdot \det J(t, u) = \\
\frac{1}{\sqrt{2\pi}} e^{-t^2 u/2\nu} \cdot \frac{1}{\Gamma(\nu/2)2^{\nu/2}} u^{\nu/2-1} e^{-u/2} \sqrt{\frac{u}{\nu}} \cdot 1 \\
= \frac{1}{\sqrt{2\pi}} \frac{1}{\Gamma(\nu/2)2^{\nu/2} u^{\nu/2}} e^{-(1+t^2/\nu)u/2} u^{(\nu+1)/2-1}
\]
The t Distribution

Integrate with respect to $u$ to obtain the marginal distribution for $t$

$$f_T(t) = \frac{1}{\sqrt{2\pi}} \frac{1}{\Gamma(\nu/2)2^{\nu/2}\nu^{1/2}} \int_0^\infty e^{-(1+t^2/\nu)u/2} u^{(\nu+1)/2-1} \, du$$

The integrand has the $u$ dependent part of the $\Gamma((\nu + 1)/2, (1 + t^2/\nu)/2)$ density. Therefore, the integral equals

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

and

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

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

The parameter $\nu$ is again called the degrees of freedom.
The $t$ Distribution

The density $f_T(t)$ is symmetric about the origin. Thus, $ET = 0$ provided the integral exists. The integrand in computing the mean $tf_T(t) = O(t^{-(\nu+1)/2})$ implying that the integral converges provided $\nu > 1$.

Exercise. $\text{Var}(T) = \nu/(\nu - 2), \nu > 2$.

The $t$-density (left) and $t$-distribution (right) with 4 (brown) and 8 (teal) degrees of freedom. The case of the normal random variable is shown in black.
Upper tail probabilities for \( t \)-distribution with 10\% (teal), 5\% (brown) and 1\% (black).

**Exercise.** The density of the \( t \) distribution with \( \nu \) degrees of freedom converges to the density of the standard normal.
Introduction to Asymptotic Analysis

Repeatedly roll a six-sided dice

- Count the number of pips and keep a running average.
- At the beginning we might see some larger fluctuations in our average.
- As they continue to roll dice, we expect to see this running average settle and converge to the true mean of the die.

This phenomena is informally known as the law of averages. In probability theory, we call this the law of large numbers.
The first three running averages seem to become closer to the $\mu = 3.5$, the mean on a fair dice. The next three running averages appear to be from a die with bias towards the higher values.
Introduction to Asymptotic Analysis

For the law of large numbers, the sample means from a sequence of independent random variables converge to their common distributional mean as the number $n$ of random variables increases.

$$\frac{1}{n} S_n = \bar{X}_n \to \mu \text{ as } n \to \infty.$$  

Moreover, the standard deviation of $\bar{X}_n$ is inversely proportional to $\sqrt{n}$. For example, for fair dice $\bar{X}_n$ converges to $\mu = \frac{7}{2}$.

Because the standard deviation $\sigma_{\bar{X}_n} \propto 1/\sqrt{n}$, we magnify the difference between the running average and the mean by a factor of $\sqrt{n}$ and investigate graphs of

$$\sqrt{n} \left( \frac{1}{n} S_n - \mu \right) \text{ versus } n.$$
Introduction to Asymptotic Analysis

Here are five simulations of \( \sqrt{n} \left( \frac{1}{n} S_n - \mu \right) \) for the fair dice.

So, we do not see a convergence through the simulations.
To see if the distributions of $\sqrt{n} \left( \frac{1}{n} S_n - \mu \right)$ has any structure, let’s look at 100 simulations.

```r
> sigma<-sqrt(35/12)
> zsim<-numeric(100)
> for (i in 1:100){x<-sample(6,200,replace=TRUE);
  zsim[i]<-sqrt(200)*(sum(x)/200-7/2)}
> plot(ecdf(zsim),xlim=c(-4,4),ylim=c(0,1),col="aquamarine3")
> par(new=TRUE)
> curve(pnorm(x/sigma),-4,4,ylim=c(0,1),xlab="",ylab="")
```
The empirical cumulative distribution function of $\sqrt{n \frac{1}{n} (S_n - \mu)}$ (teal).
The distribution function of $Z \sim N(0, \sigma^2)$ (black)
For a sequence of random variables, $Y_1, Y_2, \ldots$, and a distinguished random variable $Y$, we will consider a variety of notions of convergence of the sequence to $Y$.

One version focuses on the specific realizations $\omega$, the difference between $Y_n(\omega)$ and $Y(\omega)$ is small for sufficiently large $n$. We see this in the law of large numbers where the $Y_n$ is the average of the first $n$ observations and $Y$ is the constant, namely the mean $\mu$ of the random variables.

The second version is based on the differences between the distributions of the $Y_n$ and $Y$ is small for sufficiently large $n$. We see this in the second example with the distribution of $Y_n = \sqrt{n}(S_n/n - \mu)$ close to the distribution of a normal random variable.