Lecture 16

16.1 Fisher and Student distributions.

Consider \(X_1, \ldots, X_k\) and \(Y_1, \ldots, Y_m\) all independent standard normal r.v.

**Definition:** Distribution of the random variable

\[
Z = \frac{X_1^2 + \ldots + X_k^2}{Y_1^2 + \ldots + Y_m^2}
\]

is called Fisher distribution with degree of freedom \(k\) and \(m\), and it is denoted as \(F_{k,m}\).

Let us compute the p.d.f. of \(Z\). By definition, the random variables

\[
X = X_1^2 + \ldots + X_k^2 \sim \chi_k^2 \quad \text{and} \quad Y = Y_1^2 + \ldots + Y_m^2 \sim \chi_m^2
\]

have \(\chi^2\) distribution with \(k\) and \(m\) degrees of freedom correspondingly. Recall that \(\chi_k^2\) distribution is the same as gamma distribution \(\Gamma \left( \frac{k}{2}, \frac{1}{2} \right) \) which means that we know the p.d.f. of \(X\) and \(Y\):

- \(X\) has p.d.f. \(f(x) = \frac{(\frac{1}{2})^{\frac{k}{2}}}{\Gamma(\frac{k}{2})} x^{\frac{k}{2}-1} e^{-\frac{1}{2}x}\)
- \(Y\) has p.d.f. \(g(y) = \frac{(\frac{1}{2})^{\frac{m}{2}}}{\Gamma(\frac{m}{2})} y^{\frac{m}{2}-1} e^{-\frac{1}{2}y}\)

for \(x \geq 0\) and \(y \geq 0\). To find the p.d.f of the ratio \(\frac{X}{Y}\), let us first recall how to write its cumulative distribution function. Since \(X\) and \(Y\) are always positive, their ratio is also positive and, therefore, for \(t \geq 0\) we can write:

\[
P \left( \frac{X}{Y} \leq t \right) = P(X \leq tY) = E\{I(X \leq tY)\}
\]

\[
= \int_0^\infty \int_0^\infty I(x \leq ty)f(x)g(y)dx \, dy
\]

\[
= \int_0^\infty \left( \int_0^{ty} f(x)g(y) \, dx \right) dy
\]
LECTURE 16.

where \( f(x)g(y) \) is the joint density of \( X, Y \). Since we integrate over the set \( \{x \leq ty\} \) the limits of integration for \( x \) vary from 0 to \( ty \) (see also figure 16.1).

Since p.d.f. is the derivative of c.d.f., the p.d.f. of the ratio \( \frac{X}{Y} \) can be computed as follows:

\[
\frac{d}{dt} P\left( \frac{X}{Y} \leq t \right) = \frac{d}{dt} \int_0^\infty \int_0^{ty} f(x)g(y) \, dx \, dy = \int_0^\infty f(ty)g(y) \, dy
\]

\[
= \int_0^\infty \left( \frac{1}{2} \right)^{\frac{k}{2}} \Gamma\left(\frac{k}{2}\right) \left(\frac{1}{2}\right)^{\frac{m}{2}} \Gamma\left(\frac{m}{2}\right) e^{-\frac{1}{2}ty} \Gamma\left(\frac{m}{2}\right) \frac{1}{2} e^{-\frac{1}{2}y} \, dy
\]

\[
= \frac{\Gamma\left(\frac{m}{2}\right)}{\Gamma\left(\frac{k}{2}\right)} \frac{1}{2} \left(\frac{1}{2}\right)^{\frac{k}{2}} \Gamma\left(\frac{k}{2}\right) \Gamma\left(\frac{m}{2}\right) t^{-\frac{1}{2}} \int_0^\infty y^{\frac{k+m}{2}-1} e^{-\frac{1}{2}(t+1)y} \, dy
\]

The function in the underbraced integral almost looks like a p.d.f. of gamma distribution \( \Gamma(\alpha, \beta) \) with parameters \( \alpha = (k + m)/2 \) and \( \beta = 1/2 \), only the constant in front is missing. If we multiply and divide by this constant, we will get that,

\[
\frac{d}{dt} P\left( \frac{X}{Y} \leq t \right) = \frac{\frac{1}{2}}{\Gamma\left(\frac{k}{2}\right)} \frac{1}{2} \left(\frac{1}{2}\right)^{\frac{k}{2}} \Gamma\left(\frac{k}{2}\right) \Gamma\left(\frac{m}{2}\right) t^{-\frac{1}{2}} \int_0^\infty y^{\frac{k+m}{2}-1} e^{-\frac{1}{2}(t+1)y} \, dy
\]

\[
= \frac{\Gamma\left(\frac{k+m}{2}\right)}{\Gamma\left(\frac{k}{2}\right) \Gamma\left(\frac{m}{2}\right)} t^{-\frac{1}{2}} (1 + t)^{-\frac{k+m}{2}}
\]

Since we integrate a p.d.f. and it integrates to 1.

To summarize, we proved that the p.d.f. of Fisher distribution with \( k \) and \( m \) degrees of freedom is given by

\[
f_{k,m}(t) = \frac{\Gamma\left(\frac{k+m}{2}\right)}{\Gamma\left(\frac{k}{2}\right) \Gamma\left(\frac{m}{2}\right)} t^{-\frac{1}{2}} (1 + t)^{-\frac{k+m}{2}}.
\]
Next we consider the following definition. The distribution of the random variable
\[ Z = \frac{X_1}{\sqrt{\frac{1}{m}(Y_1^2 + \cdots + Y_m^2)}} \]
is called the Student distribution or \( t \)-distribution with \( m \) degrees of freedom and it is denoted as \( t_m \).

Let us compute the p.d.f. of \( Z \). First, we can write,
\[ \mathbb{P}( -t \leq Z \leq t ) = \mathbb{P}( Z^2 \leq t^2 ) = \mathbb{P}\left( \frac{X_1^2}{Y_1^2 + \cdots + Y_m^2} \leq \frac{t^2}{m} \right). \]
If \( f_Z(x) \) denotes the p.d.f. of \( Z \) then the left hand side can be written as
\[ \mathbb{P}( -t \leq Z \leq t ) = \int_{-t}^{t} f_Z(x) \, dx. \]
On the other hand, by definition, \( \frac{X_1^2}{Y_1^2 + \cdots + Y_m^2} \) has Fisher distribution \( F_{1,m} \) with 1 and \( m \) degrees of freedom and, therefore, the right hand side can be written as
\[ \int_{0}^{\frac{t^2}{m}} f_{1,m}(x) \, dx. \]
We get that,
\[ \int_{-t}^{t} f_Z(x) \, dx = \int_{0}^{\frac{t^2}{m}} f_{1,m}(x) \, dx. \]
Taking derivative of both side with respect to \( t \) gives
\[ f_Z(t) + f_Z(-t) = f_{1,m}(\frac{t^2}{m}) \frac{2t}{m}. \]
But \( f_Z(t) = f_Z(-t) \) since the distribution of \( Z \) is obviously symmetric, because the numerator \( X \) has symmetric distribution \( N(0, 1) \). This, finally, proves that
\[ f_Z(t) = \frac{t}{m} f_{1,m}(\frac{t^2}{m}) = \frac{t}{m} \frac{\Gamma(\frac{m+1}{2})}{\Gamma(\frac{m}{2})} \left( \frac{t^2}{m} \right)^{-1/2} \left( 1 + \frac{t^2}{m} \right)^{-\frac{m+1}{2}} = \frac{\Gamma(\frac{m+1}{2})}{\Gamma(\frac{1}{2}) \Gamma(\frac{m}{2})} \frac{1}{\sqrt{m}} \left( 1 + \frac{t^2}{m} \right)^{-\frac{m+1}{2}}. \]