6.4: Distributions Based on a Normal Random Sample

Taylor

University of Virginia

Definitions

This is the density of a chi-squared (χ^2_ν) distribution with ν degrees of freedom.

$$p(x) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{(\nu/2)-1} \exp(-x/2), \qquad x > 0$$

Sometimes we'll write this for shorthand $X \sim \chi^2_\nu$.

chi-squared distribution

If $X \sim \chi^2_\nu$, then

- $E(X) = \nu$,
- $V(X) = 2\nu$ and
- $M_X(t) = (1 - 2t)^{-\nu/2}$

chi-squared distribution

They often say that chi-squared rvs with $\nu = 1$ are formed by squaring standard normal random variables. We can prove that. Let Z be a standard normal rv. Let $X = Z^2$. Then

$$\begin{aligned} P(X \le x) &= P(Z^2 \le x) \\ &= P(-\sqrt{x} \le Z \le \sqrt{x}) \\ &= 2P(0 \le Z \le \sqrt{x}) \\ &= 2\Phi(\sqrt{x}) - 2\Phi(0) \end{aligned}$$

Then take the derivative...

$$\begin{aligned} p(x) &= 2\phi(\sqrt{x}).5x^{-.5} \\ &= 2\frac{1}{\sqrt{2\pi}}e^{-.5x}(.5x^{-.5}) \\ &= \frac{1}{2^{1/2}\Gamma(1/2)}x^{1/2-1}e^{-x/2} \end{aligned}$$

That just takes care of χ_1^2. What about χ_ν^2 with $\nu > 1$? Let $X_1, \ldots X_\nu \overset{iid}{\sim} \chi_1^2$

$$M_{\sum_{i=1}^{\nu} X_i}(t) = \prod_{i=1}^{\nu} M_{X_i}(t) = \prod_{i=1}^{\nu} [(1-2t)^{-1/2}] = (1-2t)^{-\nu/2}$$

So we can add ν independent χ_1^2s to get this.

chi-squared distribution

Actually it's more general. As long as some chi-squared rvs are independent, we can add them together nicely. It doesn't matter what their degrees of freedom are.

This is the proposition on page 316. The proof uses the same technique as above.

Note

So if we have $X_1, \ldots X_n \overset{iid}{\sim} \text{Normal}(\mu, \sigma^2)$...

$$\sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma} \right)^2 \sim \chi_n^2$$

But we don't know μ usually so this doesn't always help...

A trick

check this:

$$\sum_i (X_i - \mu)^2 = \sum_i (X_i - \bar{X} + \bar{X} - \mu)^2 = \sum_i (X_i - \bar{X})^2 + n(\bar{X} - \mu)^2$$

which implies this:

$$\begin{aligned}
\sum_i \left(\frac{X_i - \mu}{\sigma} \right)^2 &= \frac{1}{\sigma^2} \sum_i (X_i - \mu)^2 \\
&= \frac{1}{\sigma^2} \sum_i (X_i - \bar{X})^2 + \frac{n}{\sigma^2} (\bar{X} - \mu)^2 \\
&= \sum_i \left(\frac{X_i - \bar{X}}{\sigma} \right)^2 + \left(\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \right)^2
\end{aligned}$$

We know the LHS is a χ^2_n and the right side of the RHS is a χ^2_1. Does this mean $\sum_i \left(\frac{X_i - \bar{X}}{\sigma}\right)^2$ is a χ^2_{n-1}?

$$\sum_i \left(\frac{X_i - \mu}{\sigma}\right)^2 = \sum_i \left(\frac{X_i - \bar{X}}{\sigma}\right)^2 + \left(\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}\right)^2$$

We need to show two things:

- $\sum_i \left(\frac{X_i - \bar{X}}{\sigma}\right)^2$ is indep. of $\left(\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}\right)^2$
- $\sum_i \left(\frac{X_i - \bar{X}}{\sigma}\right)^2$ is distributed as a χ^2_{n-1}

First Bullet Point

It is possible to show that $\bar{X}$ and $X_i - \bar{X}$ $i = 2, \ldots, n$ are independent by using the transformation theorem.

It is also possible to use the transformation theorem again to show that $\sum_i \left(\frac{X_i - \bar{X}}{\sigma}\right)^2$ and $\left(\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}\right)^2$ are independent (because they're functions of independent rvs the Jacobian factors).

Second Bullet Point

Let's rewrite

$$\sum_i \left(\frac{X_i - \mu}{\sigma}\right)^2 = \sum_i \left(\frac{X_i - \bar{X}}{\sigma}\right)^2 + \left(\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}\right)^2$$

as

$$X = Y + Z$$

then

$$M_X(t) = M_Y(t) M_Z(t)$$

or more specifically

$$(1 - 2t)^{-n/2} = M_Y(t)(1 - 2t)^{-1/2}$$

After rearranging we see $Y = \sum_i \left(\frac{X_i - \bar{X}}{\sigma}\right)^2 \sim \chi^2_{n-1}$

Recap

We just proved $\sum_i \left(\frac{X_i - \bar{X}}{\sigma}\right)^2 \sim \chi^2_{n-1}$. Sometimes people write this as

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$

We're going to use this all the time.

Also remember a lot of this stuff isn't true in general – we had to assume Normality!

Refresher

Recall the gamma function:

$$\Gamma(z) = \int_0^\infty e^{-x} x^{z-1} dx$$

Refresher

These were the main properties:

1. if z is an integer, $\Gamma(z) = (z-1)!$
2. or more generally (not just if z is an integer), $\Gamma(z) = (z-1)\Gamma(z-1)$
3. we can go to someplace like wolframalpha.com and type in "Gamma(1/2)" and get $\sqrt{\pi}$

So a lot of times when we're doing integration, we want to be able to recognize a gamma function and use these results.

The t-Distribution

So now let's get back to the t-distribution. First, let's look at its representation.

Let Z be a standard normal, and let X be an independent χ^2_ν random variable. Then $T \sim \mathrm{t}_\nu$ if we can write it like this:

$$T = \frac{Z}{\sqrt{X/\nu}}$$

Example

In STAT 2120 we always talk about how $T = \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim \mathrm{t}_{n-1}$

$$\frac{\bar{X}-\mu}{S/\sqrt{n}} = \frac{\sigma/\sqrt{n}}{\sigma/\sqrt{n}} \left(\frac{\bar{X}-\mu}{S/\sqrt{n}} \right) = \left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \right) \div \sqrt{\frac{(n-1)S^2}{\sigma^2} / (n-1)}$$

Keep in mind we're using the fact that $\bar{X}$ and S^2 are independent for normal rvs. That's from slide 10.

Proof

This is straight off of page 321 in the text.

$$\begin{aligned}
F_T(t) &= P\left(\frac{Z}{\sqrt{X/\nu}} \le t\right) \\
&= P\left(Z \le t\sqrt{X/\nu}\right) \\
&= \int_0^\infty \int_{-\infty}^{t\sqrt{x/\nu}} f_X(x) f_Z(z) dz dx \\
&= \int_0^\infty f_X(x) \left[\int_{-\infty}^{t\sqrt{x/\nu}} f(z) dz\right] dx \\
&= \int_0^\infty f_X(x) \left[F_Z(t\sqrt{x/\nu})\right] dx
\end{aligned}$$

Proof (continued)

Let's take the derivative now. That last step is left as an exercise. Hint: recognize gamma functions.

$$
\begin{aligned}
f_T(t) &= \frac{d}{dt}\int_0^\infty f_X(x)\left[F_Z(t\sqrt{x/\nu})\right]dx \\
&= \int_0^\infty \frac{d}{dt} f_X(x)\left[F_Z(t\sqrt{x/\nu})\right]dx \\
&= \int_0^\infty f_X(x)\frac{d}{dt}\left[F_Z(t\sqrt{x/\nu})\right]dx \\
&= \int_0^\infty f_X(x)\left[f_Z(t\sqrt{x/\nu})\sqrt{x/\nu}\right]dx \\
&= \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\,\Gamma(\nu/2)}\left(1+t^2/\nu\right)^{-\left(\frac{\nu+1}{2}\right)}
\end{aligned}
$$

F distributions

Another commonly used distribution is an F distribution. Here is it's representation:

$$F_{\nu_1,\nu_2} = \frac{X_1/\nu_1}{X_2/\nu_2}$$

where $X_1 \sim \chi^2_{\nu_1}$, $X_2 \sim \chi^2_{\nu_2}$ and X_1 is independent of X_2.

Example

F distributions are sometimes used to test if two independent groups of randomly sampled data have the same variance. We know for the first group that $\frac{(n-1)S_1^2}{\sigma_1^2} \sim \chi^2_{n-1}$ and for the second group $\frac{(m-1)S_2^2}{\sigma_2^2} \sim \chi^2_{m-1}$. We can easily make an F random variable with this:

$$F_{m-1,n-1} = \frac{(m-1)S_2^2}{\sigma_2^2(m-1)} \div \frac{(n-1)S_1^2}{\sigma_1^2(n-1)} = \frac{S_2^2}{\sigma_2^2} \div \frac{S_1^2}{\sigma_1^2}$$

Notes

1. When we do hypothesis testing, sometimes $H_0 : \sigma_1^2 = \sigma_2^2$, which reduces our F statistic a bit further to S_1^2/S_2^2.
2. $1/F_{\nu_1,\nu_2} = F_{\nu_2,\nu_1}$ (check using it's representation)