Normal Distributions

November 18, 2008

1 Chi-square Distributions

For any continuous random variable X, let $Y = X^2$.

$$F_Y(y) = P\{Y \le y\} = P\{-\sqrt{y} \le X \le \sqrt{y}\} = F_X(\sqrt{y}) - F_X(-\sqrt{y}).$$

Thus, the density

$$f_Y(y) = \frac{1}{2\sqrt{y}}(f_X(\sqrt{y}) - f_X(-\sqrt{y})).$$

If the density f_X is symmetric, then

$$f_Y(y) = \frac{1}{\sqrt{y}} f_X(\sqrt{y}).$$

If we take $Y = Z^2$ with Z the standard normal, then

$$f_Y(y) = \frac{1}{\sqrt{2\pi y}} \exp -\frac{y}{2}.$$

We recognize this as the density of a gamma random variable with parameters $\alpha = 1/2$ and $\beta = 2$. Consequently, for $Z_1, Z_2, \ldots, Z_n$ independent standard normals, the sum $Y = Z_1^2 + Z_2^2 + \cdots + Z_n^2$ is a a gamma random variable with parameters $\alpha = n/2$ and $\beta = 2$.

This random variable, called **a chi-square random variable with n degress of freedom** χ^2_n, has density

$$f_Y(y) = \frac{1}{2^{n/2}\Gamma(n/2)} y^{(n/2)-1} \exp -\frac{y}{2}.$$

2 The t Distribution

The t-test is based on understanding the t statistic.

$$T(x) = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

We shall accomplish this in four steps.

Step 1. $\sqrt{n}(\bar{X} - \mu_0)/\sigma$ is a standard normal random variable.

For this, notice that $\bar{X}$ is a normal random variable with mean μ_0 and standard deviation $\sigma/\sqrt{n}$

Step 2. For each i, $X_i - \bar{X}$ and $\bar{X}$ are independent.

For normal random variables, uncorrelated random variables are independent.
Thus, it suffices to show that the covariance is 0. To that end, note that

$$\mathrm{Cov}(X_i - \bar{X}, \bar{X}) = \mathrm{Cov}(X_i, \bar{X}) - \mathrm{Cov}(\bar{X}, \bar{X}).$$

For the first term, use the fact that $\mathrm{Cov}(X_1, X_j) = 0$ if $i \neq j$ and $\mathrm{Cov}(X_1, X_i) = \mathrm{Var}(X_i) = \sigma^2$. Then,

$$\mathrm{Cov}(X_i, \bar{X}) = \frac{1}{n} \sum_{j=1}^{n} \mathrm{Cov}(X_i, X_j) = \frac{1}{n}\sigma^2.$$

From Step 1, we know that

$$\mathrm{Cov}(\bar{X}, \bar{X}) = \mathrm{Var}(\bar{X}) = \frac{1}{n}\sigma^2.$$

Now combine to see that $\mathrm{Cov}(X_i - \bar{X}, \bar{X}) = 0$.

Step 3. $\sum_{i=1}^{n}(X_i - \bar{X})^2/\sigma^2$ is a χ-square random variable with $n-1$ degrees of freedom.

Let $Z_i = (X_i - \mu)/\sigma$ and $\bar{Z}$ be the average of the Z_i. Then Z_i are independent standard normal random variables.

$$\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \sum_{i=1}^{n}(Z_i - \bar{Z})^2 = \sum_{i=1}^{n} Z_i^2 - n\bar{Z}^2$$

or

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

Let's write this

$$Y = U + V.$$

By step 2, the sum is of independent random variables. So, if we use the properties of moment generating functions

$$M_Y(r) = M_U(r) \cdot M_V(r).$$

Now Y is a χ^2_n random variable. So, $M_Y(r) = (1-2r)^{-n/2}$. V is a χ^2_1 random variable. So, $M_Y(r) = (1-2r)^{1/2}$. Consequently,

$$M_U(r) = \frac{M_Y(r)}{M_V(r)} = \frac{(1-2r)^{-n/2}}{(1-2r)^{-1/2}} = (1-2r)^{-(n-1)/2}$$

and U is a χ^2_{n-1} random variable.

In summary, we can write

$$T = \frac{Z}{\sqrt{U/(n-1)}}$$

where Z is a standard random variable, U is a χ^2_{n-1} random variable, and Z and U are independent. Consequently, their densities are

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \quad \text{and} \quad f_U(u) = \frac{u^{n/2-3/2}}{2^{(n-1)/2}\Gamma((n-1)/2)} e^{-u/2}.$$

Step 4. Finding the density of T, $f_T(t)$.

Z and U have joint density

$$f_{Z,U}(z,u) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \frac{u^{n/2-3/2}}{2^{(n-1)/2}\Gamma((n-1)/2)} e^{-u/2}.$$

Define the one to one transformation

$$t = \frac{z}{\sqrt{u/(n-1)}} \quad \text{and} \quad v = u.$$

Then, the inverse transformation

$$z = \frac{t\sqrt{v}}{\sqrt{n-1}} \quad \text{and} \quad u = v.$$

the joint density

$$f_{T,V}(t,v) = f_{Z,U}(\frac{t\sqrt{v}}{\sqrt{n-1}}, v)|J(t,v)|.$$

where $|J(z,u)|$ is the absolute value of the Jacobian of the inverse transformation.

In this case,

$$J(t,v) = \det \begin{bmatrix} \partial z/\partial t & \partial z/\partial v \\ \partial u/\partial t & \partial u/\partial v \end{bmatrix} = \det \begin{bmatrix} \sqrt{v}/\sqrt{n-1} & t/(2\sqrt{v(n-1)} \\ 0 & 1 \end{bmatrix} = \frac{\sqrt{v}}{\sqrt{n-1}}.$$

Then,

$$f_{T,V}(t,v) = \frac{1}{\sqrt{2\pi}2^{(n-1)/2}\Gamma((n-1)/2)} v^{n/2-3/2} \exp\left(-\frac{v}{2}\left(1+\frac{t^2}{n-1}\right)\right) \frac{\sqrt{v}}{\sqrt{n-1}}$$

Finally, to find the marginal density for T, we integrate with respect to v to obtain

$$f_T(t) = \frac{1}{\sqrt{2\pi}2^{(n-1)/2}\Gamma((n-1)/2)\sqrt{n-1}} \int_0^\infty v^{n/2-1} \exp\left(-\frac{v}{2}\left(1+\frac{t^2}{n-1}\right)\right) \, dv.$$

Change variables by setting $w = v(1 + t^2/(n-1))/2$.

$$\begin{aligned} f_T(t) &= \frac{1}{\sqrt{2\pi}2^{(n-1)/2}\Gamma((n-1)/2)\sqrt{n-1}} \int_0^\infty \left(\frac{2w}{1+t^2/(n-1)}\right)^{n/2-1} e^{-w} \left(\frac{2}{1+t^2/(n-1)}\right) \, dw \\ &= \frac{1}{\sqrt{\pi(n-1)}\Gamma((n-1)/2)} \left(1+\frac{t^2}{n-1}\right)^{-n/2} \int_0^\infty w^{n/2-1} e^{-w} \, dw \\ &= \frac{\Gamma(n/2)}{\sqrt{\pi(n-1)}\Gamma((n-1)/2)} \left(1+\frac{t^2}{n-1}\right)^{-n/2}. \end{aligned}$$

Note that

$$\left(1+\frac{t^2}{n-1}\right)^{-n/2} \to \exp -\frac{t^2}{2}$$

and $n \to \infty$. Thus, for large n, the t density is very close to the density of a standard normal.