1 Distribution Theory for Normal Samples

We have stated before the following proposition, which we prove again here to remind ourselves.

Definition 1. An orthonormal matrix O satisfies the identity $O^t O = I_n$, where I_n is the $n \times n$ identity matrix. [Here M^t denotes the *transpose* of the matrix M. The entry in the ith row and jth column of M^t is $M^t_{i,j} = M_{j,i}$.]

The rows of an orthnormal matrix form an orthnormal vector basis of $\mathbb{R}^n$. Also, an orthnormal matrix preserves length:

$$\sum_{i=1}^{n} (O\boldsymbol{y})_i^2 = \|O\boldsymbol{y}\|^2 = \|\boldsymbol{y}\|^2 = \sum_{i=1}^{n} y_i^2 \,. \tag{1}$$

Proposition 1. *Let $\boldsymbol{X} = (X_1, \ldots, X_n)$ be a vector of i.i.d. standard normal random variables, and let O be an orthonormal transformation. The random vector $\boldsymbol{Y} = O\boldsymbol{X}$ is also a vector of independent standard normal random variables*

Proof. First, if $T : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation, represented by multiplication by the matrix M, then T is invertible if and only if $\det(M) \neq 0$, and then the T^{-1} is obtained by matrix multiplication by M^{-1}. Moreover, the Jacobian of T^{-1} is the determinant of M^{-1}.

By Definition 1,

$$1 = \det(I_n) = \det(O^t O) = \det(O)^2 \,,$$

and $\det(O) = 1$. We conclude that the Jacobian of O^{-1} is 1. Also, by the definition of orthonormality, we see that $O^{-1} = O^t$, which is also an orthonormal matrix.

We thus have by the transformation formula,

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = (2\pi)^{-n/2} \exp\left(-\frac{1}{2} \sum_{i=1}^{n} (O^t \boldsymbol{y})_i^2 \right) = (2\pi)^{-n/2} \exp\left(-\frac{1}{2} \sum_{i=1}^{n} y_i^2 \right) \,.$$

The second inequality follows from (1). This proves that the density of $\boldsymbol{Y}$ is a product of standard normal densities, hence proving the Proposition. $\square$

Theorem 2. *Let $X_1, \ldots, X_n$ be a vector of i.i.d. standard normal random variables, and let*

$$\bar{X} \stackrel{\text{def}}{=} \frac{1}{n}\sum_{i=1}^{n} X_i\,, \quad \text{and} \quad S^2 \stackrel{\text{def}}{=} \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2\,.$$

Then $\bar{X}$ and S^2 are independent.

Proof. It is always possible, given a vector of unit length in $\mathbb{R}^n$, to find $n-1$ other vectors so that the collection of n vectors forms an orthonormal basis in $\mathbb{R}^n$. In particular, given that the first row of a matrix is equal to the vector $(n^{-1/2}, n^{-1/2}, \ldots, n^{-1/2})$, we can complete this matrix so that it is an orthonormal matrix O. [A matrix is orthonormal if and only if its rows form an orthonormal basis in $\mathbb{R}^n$.]

Let $\boldsymbol{Y} = O\boldsymbol{X}$. By the rules for matrix multiplication, we have that $Y_1 = n^{1/2}\bar{X}$. Since an orthonormal transformation preserves the length of a vector, we have

$$\begin{aligned}
\sum_{i=1}^{n} Y_i^2 &= \sum_{i=1}^{n} X_i^2 \\
Y_1^2 + \sum_{i=2}^{n} Y_i^2 &= \sum_{i=1}^{n} X_i^2 \\
\sum_{i=1}^{n} X_i^2 - n\bar{X}^2 &= \sum_{i=2}^{n} Y_i^2 \\
S^2 &= \frac{1}{n-1}\sum_{i=2}^{n} Y_i^2\,.
\end{aligned}$$

To summarize, we have shown that

$$\begin{aligned}
\bar{X} &= n^{-1/2} Y_1\,, \\
S^2 &= \frac{1}{n-1}\sum_{i=2}^{n} Y_i^2\,.
\end{aligned}$$

Thus, $\bar{X}$ depends only on Y_1, and S^2 depends only on $(Y_2, \ldots, Y_n)$. Since Y_1 and $(Y_2, \ldots, Y_n)$ are independent, it follows that $\bar{X}$ and S^2 are independent. $\square$

Corollary 3. *Let $X_1, \ldots, X_n$ be a random sample from a $N(\mu, \sigma^2)$ distribution. Then $\bar{X}$ and S^2 are independent.*

Proof. Let $Z_i = (X_i - \mu)/\sigma$, for $i = 1, \dots, n$. Then $(Z_1, \dots, Z_n)$ is a vector of independent and identically distributed standard normal random variables. We can write

$$\begin{aligned}\bar{Z} &= \frac{\bar{X} - \mu}{\sigma} \\ \bar{X} &= \sigma\bar{Z} + \mu\,,\end{aligned}$$

and

$$\begin{aligned}\frac{1}{n-1}\left(\sum_{i=1}^{n}(Z_i - \bar{Z})^2\right) &= \frac{1}{\sigma^2}\frac{1}{n-1}\left(\sum_{i=1}^{n}(X_i - \bar{X})^2\right) \\ S^2 &= \frac{\sigma^2}{n-1}\left(\sum_{i=1}^{n}(Z_i - \bar{Z})^2\right)\,.\end{aligned}$$

Since $\bar{Z}$ and $\frac{1}{n-1}\sum_{i=1}^{n}(Z_i - \bar{Z})^2$ are independent by Theorem 2, it follows that $\bar{X}$ and S^2 are independent. □

We know determine the distribution of S^2:

Proposition 4. *If $X_1, \dots, X_n$ is a random sample from a $N(\mu, \sigma^2)$ distribution, then the distribution of $((n-1)/\sigma^2)S^2$ is chi-squared with $n-1$ degrees of freedom.*

Proof. We have

$$\begin{aligned}\underbrace{\sum_{i=1}^{n}(X_i - \mu)^2}_{V_3} &= \sum_{i=1}^{n}(X_i - \bar{X} + \bar{X} - \mu)^2 \\ &= \sum_{i=1}^{n}(X_i - \bar{X})^2 + 2(X_i - \bar{X})(\bar{X} - \mu) + (\bar{X} - \mu)^2 \\ &= \sum_{i=1}^{n}(X_i - \bar{X})^2 + 2(\bar{X} - \mu)\sum_{i=1}^{n}(X_i - \bar{X}) + n(\bar{X} - \mu)^2 \\ &= \underbrace{\sum_{i=1}^{n}(X_i - \bar{X})^2}_{V_1} + \underbrace{n(\bar{X} - \mu)^2}_{V_2}\,.\end{aligned}$$

Now V_3/σ^2 has a chi-squared distribution with n degrees of freedom, as it is the sum of n independent standard normals. V_2/σ^2 has a chi-squared distribution with 1 degree of freedom, as it is the square of a standard normal. Furthermore, V_1 and V_2 are independent by Corollary 3. If M_i denotes the moment generating function for V_i/σ^2, then we have that

$$M_3(t) = M_1(t)M_2(t)\,,$$

and so

$$M_1(t) = M_3(t)/M_2(t)\,.$$

The moment generating function for a chi-squared with ν degrees of freedom is $(1-2t)^{-\nu/2}$. Thus

$$M_1(t) = \frac{(1-2t)^{-n/2}}{(1-2t)^{-1/2}} = (1-2t)^{-(n+1)/2}\,.$$

It follows that V_2/σ^2 has a chi-squared distribution with $n-1$ degrees of freedom. □

We have seen, by moment generating functions, that the sum of independent normal random variables is again normal. Hence it follows that $\bar{X}$ is normal with mean μ and variance σ^2/n, when $X_1, \ldots, X_n$ is a sample from a $N(\mu, \sigma^2)$ distribution. We now have enough information to determine the joint distribution of $\bar{X}$ and S^2.

Corollary 5. *Let $X_1, \ldots, X_n$ be a random sample from a $N(\mu, \sigma^2)$ distribution. Then $\bar{X}$ and S^2 are independent, $\bar{X}$ is $N(\mu, \sigma^2/n)$, and $\frac{n-1}{\sigma^2}S^2$ is chi-squared with $n-1$ degrees of freedom.*