Recap of chi-square

Souvik Roy

October 5, 2022

1 Introduction

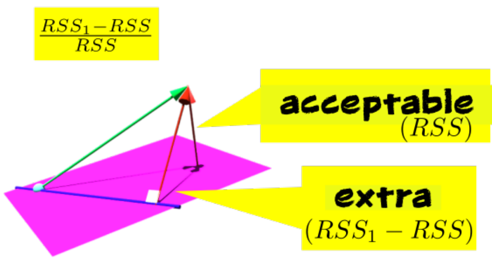
Suppose we are given that our data already fits the model,

$$\vec{y} = X\vec{\beta} + \vec{\varepsilon}$$

with $\vec{\varepsilon} \sim N(\vec{0}, \sigma^2 I)$, $\vec{\beta} \in \mathbb{R}^p$ and $\sigma^2 > 0$. Now we want to test the already familiar hypothesis,

$$\begin{aligned} H_0 \quad : \quad & \vec{y} = X_0\vec{\beta} + \vec{\varepsilon}, \\ & \vec{\varepsilon} \sim N(\vec{0}, \sigma^2 I), \\ & \vec{\beta} \in \mathbb{R}^q, \sigma^2 > 0, \\ & \text{ where } \mathcal{C}(X_0) \subseteq \mathcal{C}(X). \end{aligned}$$

against the alternative hypothesis that H_0 does not hold. In this case, it is implied that the initial model, namely $\vec{y} = X\vec{\beta} + \vec{\varepsilon}$ with the appropriate assumptions, is **still holds**. We must remember that the initial model, whatever X might be, is **given** to us, and is taken to be true throughout. In the hypothesis test, we test if a **subset** (or subspace to be more accurate) of the given model is true or not. We have already seen the above setup, and we are also familiar with the geometric intuition. That is, if RSS and RSS_1 denote the residual sum of squares under the original model and the null model respectively, then we shall look at the relative error deviation $\dfrac{RSS_1 - RSS}{RSS}$ as a measure to see if our null model is good enough. The RSS is the acceptable level of error since that is obtained from the original model that is taken to be true.



2 Recap of chi-square

The χ^2 distribution is defined to be the distribution of the sum of squares of iid standard normal random variables. Formally, if $X_1, X_2, \ldots, X_n \sim N(0,1)$, and $Z = \sum_{i=1}^{n} X_i^2$, then, $Z \sim \chi^2_{(n)}$, where n is known as the number of degrees of freedom.
This can be rewritten as follows. If $\vec{X} \sim N_n(\vec{0}, I)$, then $||\vec{X}||^2 \sim \chi^2_n$, since each component of $\vec{X}$ is independent and follows standard normal. Again, if $\vec{X} \sim N_n(\vec{0}, \sigma^2 I)$, with $\sigma^2 > 0$, then, clearly, each X_i follows $N(0, \sigma^2)$ and is independent of the other components. Therefore, $\frac{1}{\sigma}\vec{X} \sim N(\vec{0}, I)$, and thus, $\dfrac{1}{\sigma^2}||\vec{X}||^2 \sim \chi^2_{(n)}$.
We introduce a new notation for the distribution of $||\vec{X}||$ in this case. We write,

$$||\vec{X}||^2 \sim \sigma^2 \chi^2_{(n)}.$$

Again, if the X_i's are independent, but each of them follow a normal with mean μ_i, but the same variance, σ^2, then, upon stacking the X_i's up as a vector $\vec{X}$, we have $\vec{X} \sim N_n(\vec{\mu}, \sigma^2 I)$. Then, we again look at the distribution of $||\vec{X}||^2$, and whatever that is, we write it as $\sigma^2 \chi^2_{(n)}(||\vec{\mu}||^2)$. Therefore, in this case,

$$||\vec{X}||^2 \sim \sigma^2 \chi^2_{(n)}(||\vec{\mu}||^2).$$

We call the above distribution to be a **non-central chi-squared** and the parameter $||\vec{\mu}||^2$ to be our **non centrality parameter**. This however is not an entirely standard notation, as some authors might prefer including the σ^2 into the non-centrality parameter, however we shall use our notation here.