CHAPTER 4 - HYPOTHESIS TESTING - SIMPLE NULL HYPOTHESES

§4.1 Review of hypothesis testing and the Neyman-Pearson Lemma

Suppose $X_1, \ldots, X_n$ have joint distribution $f_\theta(\mathbf{x})$ for $\theta \in \Theta$ and we are interested in hypotheses of the form

$$H_0 : \theta \in \Theta_0 \quad \text{vs} \quad H_1 : \theta \in \Theta_1$$

where $\Theta_1 = \Theta \backslash \Theta_0$.
If the distribution $f_\theta(\mathbf{x})$ is completely determined under either the null or the alternative hypotheses (for example, say $\Theta_0 = \{\theta_0\}$), then that hypothesis is said to be simple. Otherwise it is known as compound. The most commonly encountered scenarios seen in earlier units are simple versus simple tests and simple versus compound tests.

Definition 4.1

Let C be a critical region for the hypothesis test

$$H_0 : \theta \in \Theta_0 \quad \text{vs} \quad H_1 : \theta \in \Theta_1$$

Size The size of the associated test is defined to be $\max_{\theta \in \Theta_0} \mathsf{P}_\theta(\mathbf{X} \in C)$ (i.e. the maximum probability of rejecting H_0 when the value of θ falls under H_0).

Power The power of the associated test is defined to be $\mathsf{P}_\theta(\mathbf{X} \in C)$ (i.e. the probability of rejecting H_0 at this value of θ).

Uniformly Most Powerful A critical region C^* and associated test are said to be Uniformly Most Powerful (UMP) at size α if

$$\begin{aligned} \max_{\theta \in \Theta_0} \mathsf{P}_\theta(\mathbf{X} \in C^*) &= \alpha \\ \text{and } \mathsf{P}_\theta(\mathbf{X} \in C^*) &\geq \mathsf{P}_\theta(\mathbf{X} \in C) \; \forall \theta \in \Theta_1 \end{aligned}$$

and for all critical regions C of size α.

In simple versus simple hypothesis tests, MA20226 introduced the Neyman-Pearson lemma as a way to generate uniformly most powerful tests at size α. Using the notation that $\Theta_0 = \{\theta_0\}$ and $\Theta_1 = \{\theta_1\}$, the Neyman-Pearson lemma states if $X_1, \ldots, X_n$ have joint distribution $f_\theta(\mathbf{x})$ then the UMP test at size α of

$$H_0 : \theta = \theta_0 \quad \text{vs} \quad H_1 : \theta = \theta_1$$

is the test using critical region

$$C^*_\alpha = \left\{ \mathbf{x} : \frac{f_{\theta_0}(\mathbf{x})}{f_{\theta_1}(\mathbf{x})} < k_\alpha \right\}$$

where $\mathsf{P}_{\theta_0}(\mathbf{X} \in C^*) = \alpha$.

Example

Suppose $X_1, \ldots, X_n$ are iid Bern(p) random variables and we wish to test

$$H_0 : p = \frac{1}{4} \quad \text{vs} \quad H_1 : p = p_1$$

where $p_1 > \frac{1}{4}$. The Neyman-Pearson lemma tells us that the uniformly most powerful test at size α is given by

$$\begin{aligned}
C^*_\alpha &= \left\{ \mathbf{x} : \frac{f_{\frac{1}{4}}(\mathbf{x})}{f_{p_1}(\mathbf{x})} < k \right\} \\
&= \left\{ \mathbf{x} : \frac{\left(\frac{1}{4}\right)^{\sum_{i=1}^n x_i} \left(\frac{3}{4}\right)^{n-\sum_{i=1}^n x_i}}{p_1^{\sum_{i=1}^n x_i} (1-p_1)^{n-\sum_{i=1}^n x_i}} < k \right\} \\
&= \left\{ \mathbf{x} : \left(\frac{1}{4p_1}\right)^{\sum_{i=1}^n x_i} \left(\frac{3}{4(1-p_1)}\right)^{n-\sum_{i=1}^n x_i} < k \right\} \\
&= \left\{ \mathbf{x} : \left(\frac{1-p_1}{3p_1}\right)^{\sum_{i=1}^n x_i} \left(\frac{3}{4}\right)^{n} < k \right\} \\
&= \left\{ \mathbf{x} : \left(\frac{1-p_1}{3p_1}\right)^{\sum_{i=1}^n x_i} < k_2 \right\} \\
&= \left\{ \mathbf{x} : \sum_{i=1}^n x_i \ln\left(\frac{1-p_1}{3p_1}\right) < k_3 \right\} \\
&= \left\{ \mathbf{x} : \sum_{i=1}^n x_i > k_\alpha \right\} \quad \text{since } 1 - p_1 < 3p_1 \text{ when } p_1 > \frac{1}{4}.
\end{aligned}$$

Since $\sum_{i=1}^n X_i \sim Bin(n, 1/4)$ under H_0, k_α should be chosen as the top α quantile of a $Bin(n, 1/4)$. Notice that as this is a discrete distribution, it may not be possible to get the exact size α desired.

§4.2 Composite alternative hypotheses - one-sided

How do we find UMP tests when the null hypothesis is simple but the alternative hypothesis is compound and one-sided, for example

$$H_0 : \theta = \theta_0 \quad \text{vs} \quad H_1 : \theta > \theta_0?$$

Consider again the Bernoulli example of the previous section, where $C^*_\alpha = \{\mathbf{x} : \sum_{i=1}^n x_i > k_\alpha\}$ with k_α the upper α quantile of a $Bin(n, 1/4)$. The Neyman-Pearson lemma guarantees that this C^*_α is the most powerful size α test for the particular alternative $p_1 \in (1/4, 1)$. However this test is unchanged as p_1 varies across the range $(1/4, 1)$ and so it is also the Uniformly Most Powerful (UMP) test of the pair of hypotheses

$$H_0 : p = \frac{1}{4} \quad \text{vs} \quad H_1 : p > \frac{1}{4}.$$

(It satisfies the two conditions we require; firstly that the worst case scenario for rejecting H_0 when H_0 is true is α, and secondly that for any parameter value under H_1, the probably of rejecting H_0 is higher using this than using any other test of the same size.)

We use this approach more generally; if for the simple versus simple hypotheses, the Neyman-Pearson lemma generates a critical region which is independent of the alternative parameter, then that test will give us the UMP test of the corresponding one-sided compound alternative. This does not guarantee that a UMP test can always be found, a counterexample is given on Problem Sheet 7.

§4.3 Composite alternative hypotheses - two-sided

How do we develop tests when the alternative hypothesis is both composite and two sided

$$H_0 : \theta = \theta_0 \quad \text{vs} \quad H_1 : \theta \neq \theta_0?$$

The Neyman-Pearson lemma will not give the same C^* when we apply it to the alternative $H_1 : \ \theta = \theta_1$ if $\theta_1 > \theta_0$ as it does if $\theta_1 < \theta_0$. This means there is no UMP test for the composite two-sided alternative. Instead we will opt for a class of test which at least has the property that the probability of rejecting H_0 when H_0 is false is at least as big as the probability of rejecting H_0 when it is true (that is, the power function is minimised under H_0).

Definition 4.2

Unbiased test A test is said to be unbiased if its power under H_1 is no smaller than its size.

Uniformly most powerful unbiased test A test is said to be Uniformly Most Powerful Unbiased (UMPU) if it is uniformly most powerful within the class of unbiased tests.

Theorem 4.1

A test generated by the Neyman-Pearson lemma will be unbiased
Proof Let $f_0(\mathbf{x})$ be the joint density under H_0 and $f_1(\mathbf{x})$ be the joint density under H_1. Using the Neyman-Pearson Lemma, the critical region C^* is defined by

$$\begin{aligned} C^* &= \left\{\mathbf{x} : \frac{f_0(\mathbf{x})}{f_1(\mathbf{x})} \leq k_\alpha\right\} \\ &= \left\{\mathbf{x} : f_1(\mathbf{x}) \geq \frac{f_0(\mathbf{x})}{k_\alpha}\right\} \\ \text{and so } \bar{C^*} &= \left\{\mathbf{x} : \frac{f_0(\mathbf{x})}{f_1(\mathbf{x})} > k_\alpha\right\} \\ &= \{\mathbf{x} : f_0(\mathbf{x}) > f_1(\mathbf{x}) k_\alpha\} \end{aligned}$$

Let α be the size of the test, and β the power; we aim to show that $\beta - \alpha \geq 0$. By definition

$$\alpha = \int_{C^*} f_0(\mathbf{x}) d\mathbf{x} \tag{1}$$

$$\beta = \int_{C^*} f_1(\mathbf{x}) d\mathbf{x} \tag{2}$$

$$1 - \alpha = \int_{\bar{C^*}} f_0(\mathbf{x}) d\mathbf{x} \tag{3}$$

$$1 - \beta = \int_{\bar{C^*}} f_1(\mathbf{x}) d\mathbf{x} \tag{4}$$

Subtracting the first from the second of these equations, and recalling the definition of C^*, we see that

$$\begin{aligned}(2)-(1) \Rightarrow \quad \beta - \alpha &= \int_{C^*} (f_1(\mathbf{x}) - f_0(\mathbf{x}))d\mathbf{x} \\ &\geq \int_{C^*} (\frac{f_0(\mathbf{x})}{k_\alpha} - f_0(\mathbf{x}))d\mathbf{x} \\ &= (\frac{1}{k_\alpha} - 1)\int_{C^*} f_0(\mathbf{x})d\mathbf{x} \\ &= (\frac{1}{k_\alpha} - 1)\alpha \qquad (5)\end{aligned}$$

Subtracting the fourth from the third of the equations, and recalling the definition of $\bar{C}^*$, we see that

$$\begin{aligned}(3)-(4) \Rightarrow \quad \beta - \alpha &= \int_{\bar{C}^*} (f_0(\mathbf{x}) - f_1(\mathbf{x}))d\mathbf{x} \\ &> \int_{\bar{C}^*} (k_\alpha f_1(\mathbf{x}) - f_1(\mathbf{x}))d\mathbf{x} \\ &= (k_\alpha - 1)\int_{\bar{C}^*} f_1(\mathbf{x})d\mathbf{x} \\ &= (k_\alpha - 1)(1-\beta) \qquad (6)\end{aligned}$$

Now either $k_\alpha \leq 1$ or $k_\alpha \geq 1$. Since $\alpha \geq 0$, equation (5) gives us the desired lower bound if $k_\alpha \leq 1$. If $k_\alpha \geq 1$, since $1 - \beta \geq 0$, equation (6) gives us the desired lower bound.

The implication of Theorem 4.1 is that we can build UMPU tests using the Neyman-Pearson lemma.

Example

Returning to our previous Bernoulli example, $X_1, \ldots, X_n$ iid Bern(p) random variables, but now with a two-sided composite alternative:

$$H_0 : p = \frac{1}{4} \quad \text{vs} \quad H_1 : p \neq \frac{1}{4}.$$

From previous calculations, we can see that the UMP test for the test

$$\begin{aligned}H_0 : p = \frac{1}{4} \qquad &\text{vs} \qquad H_1 : p > \frac{1}{4} \\ \text{is } C^{\text{upper}} \quad &= \quad \left\{\mathbf{x} : \sum_{i=1}^n x_i > k_{\text{upper}}\right\}\end{aligned}$$

while for the hypotheses

$$\begin{aligned}H_0 : p = \frac{1}{4} \qquad &\text{vs} \qquad H_1 : p < \frac{1}{4} \\ \text{the UMP test is } C^{\text{lower}} \quad &= \quad \left\{\mathbf{x} : \sum_{i=1}^n x_i < k_{\text{lower}}\right\}\end{aligned}$$

In order to have a UMPU test of the two-sided composite alternative, we should use $C = C^{\text{upper}} \cup C^{\text{lower}}$, that is rejecting H_0 for either unusually large or unusually small values of $\sum_{i=1}^{n} x_i$. However, in order for such a test to have size α, k_{upper} and k_{lower} should be chosen in such a way that

$$\begin{aligned} \text{size } \mathsf{P}_{1/4}(\mathbf{X} \in C^*) &= \alpha \\ \text{i.e. } \mathsf{P}_{1/4}(\sum_{i=1}^{n} X_i < k_{\text{lower}} \cup \sum_{i=1}^{n} X_i > k_{\text{upper}}) &= \alpha. \end{aligned}$$

It is common to split the size equally between the two disjoint areas of the critical region, which in this example means setting k_{upper} and k_{lower} to be the upper and lower $\alpha/2$ quantiles of a $Bin(n, 1/4)$ respectively.