Neyman-Pearson Lemma

Lemma 1.1 *Let X be a random variable with unknown distribution, either $f_0(x)$ or $f_1(x)$, and consider the test situation*

$$H_0 : X \sim f_0(x) \text{ versus } H_1 : X \sim f_1(x) \tag{1}$$

(simple versus simple). To test H_0 versus H_1, the most powerful size α test is

$$\varphi(x) = \begin{cases} 1 & x \ni \ f_0(x)/f_1(x) < k_\alpha \\ \gamma_\alpha & \qquad\qquad\qquad = k_\alpha \\ 0 & \qquad\qquad\qquad > k_\alpha \end{cases} \tag{2}$$

where k_α and γ_α are chosen such that $\mathrm{E}_{f_0}(\varphi(x)) = \alpha$.

This lemma is the basis for future testing where we move on to composite hypotheses.

Proof 1.1 *Consider two tests, φ and φ^*. The difference of the powers of these tests is*

$$\begin{aligned}
(1-\beta_\varphi) - (1-\beta_{\varphi^*}) &= \mathrm{E}_{f_1}(\varphi) - \mathrm{E}_{f_1}(\varphi^*) && (3)\\
&= \mathrm{E}_{f_1}(\varphi - \varphi^*) && (4)\\
&= \int_{\mathfrak{X}} (\varphi(x) - \varphi^*(x)) f_1(x)dx && (5)\\
&= \int_{x \ni f_0/f_1 < k} (\varphi(x) - \varphi^*(x)) f_1(x)dx + \int_{x \ni f_0/f_1 = k} (\varphi(x) - \varphi^*(x)) f_1(x)dx \\
&\quad + \int_{x \ni f_0/f_1 > k} (\varphi(x) - \varphi^*(x)) f_1(x)dx && (6)\\
&\geq \frac{1}{k}\int_{x \ni f_0/f_1 < k} (\varphi(x) - \varphi^*(x)) f_0(x)dx + \frac{1}{k}\int_{x \ni f_0/f_1 = k} (\varphi(x) - \varphi^*(x)) f_0(x)dx \\
&\quad + \frac{1}{k}\int_{x \ni f_0/f_1 > k} (\varphi(x) - \varphi^*(x)) f_0(x)dx && (7)\\
&\geq \frac{1}{k_\alpha}\mathrm{E}_{f_0}(\varphi(X) - \varphi^*(X)) && (8)\\
&\geq 0 && (9)
\end{aligned}$$

Thus, the power of φ is no less than the power of φ^*.

The transition between alternative and null hypotheses in equations 6 and 7 is due to the fact that in the first integral we are in the rejection region so $\varphi - \varphi^* = 1 - \varphi^* \geq 0$ and $f_0/f_1 < k$ means that $f_1 \geq f_0/k$. Similarly, in the middle integral we are in the randomization region so $f_0/f_1 = k$ means that $f_1 = f_0/k$. Finally, for the last integral we are in the "acceptance" region so $\varphi - \varphi^* = 0 - \varphi^* \leq 0$ and $f_0/f_1 > k$ means that $f_1 \geq f_0/k$.

Equation 9 follows from 8 follows since $k_\alpha \geq 0$ and under H_0 we have that $\mathrm{E}_0(\varphi(X)) \geq \mathrm{E}_0(\varphi^*(X))$ or $\alpha \geq \alpha^*$ by size.

Example 1.1 *Let* $X_i \overset{iid}{\sim} B(1,p)$. *To test*

$$H_0 : p = 0.5 \text{ versus } H_1 : p = 0.75$$

we look at

$$\frac{f_0(x)}{f_1(x)} = \frac{1^{\sum x_i} 1^{5-\sum x_i}/2^5}{3^{\sum x_i} 1^{5-\sum x_i}/4^5} \tag{10}$$

$$= \frac{2^5}{3^{\sum x_i}} \tag{11}$$

To make $\alpha = 0.5$ *we note that*

$$\frac{f_0(x)}{f_1(x)} = \frac{2^5}{3^{\sum x_i}} \tag{12}$$

is decreasing in $\sum x_i$. *So we will reject* H_0 *if* $T = \sum X_i$ *is too large. Now note that* $T = \sum X_i \sim B(5,p)$ *so that under the null hypothesis*

$$f_0(t) = \binom{5}{t} \left(\frac{1}{2}\right)^t \left(\frac{1}{2}\right)^{5-t} \tag{13}$$

Thus

$t = \sum x_i$	0	1	2	3	4	5
$f_0(t)$	0.0313	0.1563	0.3125	0.3125	0.1563	0.0313

If we reject for $\sum x_i = 5$ $(f_0/f_1 = 0.1317)$ *then we reject with probability 0.0313 under* H_0 *(not enough). Using* $\sum x_i \geq 4$ $(f_0/f_1 = .3951)$ *we reject with probability* $0.1563 + 0.0313 = 0.1876$ *(too much). So, we should reject if* $\sum x_i = 5$ *and randomize for* $\sum x_i = 4$. *To get exact size* α *we need* γ_α.

Note that

$$\mathrm{E}_{1/2}(\varphi(X)) = \mathrm{P}_{1/2}[\text{reject } H_0] \tag{14}$$

$$= \gamma \mathrm{P}_{1/2}\left(\sum X_i = 4\right) + \mathrm{P}_{1/2}\left(\sum X_i = 5\right) \tag{15}$$

$$= \gamma 0.1563 + 0.0313 \tag{16}$$

$$= 0.05 \tag{17}$$

Solving for γ *we obtain* $\gamma = 0.1196$ *and our level* $\alpha = 0.05$ *test is*

$$\varphi(\mathbf{X}) = \begin{cases} 1 & \sum X_i = 5 \\ 0.1196 & \sum X_i = 4 \\ 0 & \text{else} \end{cases} \tag{18}$$

Finally, note that if $p = \frac{3}{4}$*, the power of this test is*

$$1 - \beta = \left(\frac{3}{4}\right)^5 + \gamma \binom{5}{4} \left(\frac{3}{4}\right)^4 \left(\frac{1}{4}\right) \tag{19}$$

$$= 0.2373 + \gamma 0.3955 \tag{20}$$

$$= 0.2846 \tag{21}$$

which is pretty low. We might want to use a few more observations.