# Multiple Testing Project: p-value

## 2.2 P-value

### 1. Distribution of the p-value under the null

**1.1 Show that for any $\alpha$, $c_\alpha = F_H^{-1}(1-\alpha)$**

We know that $\alpha = 1 - F_H(c_\alpha)$ from the definition of $c_\alpha$.

$\alpha = 1 - F_H(c_\alpha)$

$\alpha - 1 = -F_H(c_\alpha)$

$-(\alpha - 1) = F_H(c_\alpha)$

$1 - \alpha = F_H(c_\alpha)$

$F_H^{-1}(1-\alpha) = c_\alpha$

$c_\alpha = F_H^{-1}(1-\alpha)$

Q.E.D.

**1.2 Show that the p-value of the test, as a function of the data $\mathbf{X}$ used, is given by $p(\mathbf{X}) = 1 - F_H(T(\mathbf{X}))$.**

The p-value is defined as $p - value = inf\{\alpha : T(X) \in R_\alpha\}$

Which is to say that the p-value is the *smallest* $\alpha$ for which $T(X)$ is in the region $R_\alpha$ of the probability distribution $P_H$

So the p-value is an instance of $\alpha$, which is defined as $\alpha = 1 - F_H(c_\alpha)$ where $c_\alpha$ is chosen so that the equation is true. Therefore, if we replace $c_\alpha$ with our test statistic $T(\mathbf{X})$, we get a $p(\mathbf{X}) = 1 - F_H(\mathbf{X})$.

This matters because it highlights that the choice of $\alpha$ sets the minimum p-value for $\mathbf{X}$ as $p(\mathbf{X})$ such that one rejects the hypothesis that X is from the same distribution as Y when $T(\mathbf{X}) > c_\alpha$.

From the previous question we know, $F_H(c_\alpha) = 1 - \alpha$

$F_H(c_\alpha) = 1 - P_H[T(\mathbf{X}) > c_\alpha]$

$P_H[T(\mathbf{X}) > c_\alpha] = 1 - F_H(c_\alpha)$

$P_H[T(\mathbf{X}) > \mathbf{x}] = 1 - F_H(\mathbf{X})$

$p(\mathbf{X}) = 1 - F_H(\mathbf{X})$

**1.3 Show that for any univariate random variable $y$ with continuous distribution function $F$, the random variables $F(y)$ and $1 - F(y)$ follow the uniform distribution.**

We know that the CDF of a uniform distribution is as follows,

$$F(y)\begin{cases} 0 & y \leq a \\ \frac{y-a}{b-a} & a \leq y \leq b \\ 1 & y \leq b \end{cases}$$

$$1 - F(y)\begin{cases} 1 & y \leq a \\ 1 - \frac{y-a}{b-a} & a \leq y \leq b \\ 0 & y \leq b \end{cases}$$

$$1 - F(y) \begin{cases} 1 & y \leq a \\ \frac{b-y}{b-a} & a \leq y \leq b \\ 0 & y \leq b \end{cases}$$

$$1 - F(y) \begin{cases} 1 & y \leq a \\ \frac{y-b}{a-b} & b \leq y \leq a \\ 0 & y \leq b \end{cases}$$

which is the CDF for another unifrom random distribution.

**1.4 Using the above results, show that the p-value follows the uniform distribution under H.**

$\alpha = 1 - F_H(c_\alpha)$

If $F_H(c_\alpha)$ follows a uniform distribution, $1 - F_H(c_\alpha)$ does too, as proven in the previous exercise.