1 Hypothesis Testing - Uniformly Most Powerful Tests

We give the definition of a uniformly most powerful test in a general setting which includes one-sided and two-sided tests. We take the null hypothesis to be

\[ H_0 : \theta \in \Omega_0 \]

and the alternative to be

\[ H_1 : \theta \in \Omega_1 \]

We write the power function as \( \text{Pow}(\theta, d) \) to make its dependence on the decision function explicit.

**Definition:** A decision function \( d^* \) is a uniformly most powerful (UMP) decision function (or test) at significance level \( \alpha_0 \) if

1. \( \text{Pow}(\theta, d^*) \leq \alpha_0, \ \forall \theta \in \Omega_0 \)
2. For every decision function \( d \) which satisfies (1), we have \( \text{Pow}(\theta, d) \leq \text{Pow}(\theta, d^*), \ \forall \theta \in \Omega_1 \).

Do UMP tests ever exist? If the alternative hypothesis is one-sided then they do for certain distributions and statistics. We proceed by defining the needed property on the population distribution and the statistic.

**Definition:** Let \( T = t(X_1, X_2, \cdots, X_n) \) be a statistic. Let \( f(x_1, x_2, \cdots, x_n | \theta) \) be the joint density of the random sample. We say that \( f(x_1, x_2, \cdots, x_n | \theta) \) has a monotone likelihood ratio in the statistic \( T \) if for all \( \theta_1 < \theta_2 \) the ratio

\[
\frac{f(x_1, \cdots, x_n | \theta_2)}{f(x_1, \cdots, x_n | \theta_1)}
\]

depends on \( x_1, \cdots, x_n \) only through \( t(x_1, \cdots, x_n) \) and the ratio is an increasing function of \( t(x_1, \cdots, x_n) \).

**Example:** Consider a Bernoulli distribution for the population, i.e., we are looking at a population proportion. So each \( X_i = 0, 1 \) and \( p = P(X_i = 1) \). The joint density is

\[
f(x_1, \cdots, x_n | p) = p^{n\bar{x}}(1 - p)^{n-n\bar{x}}
\]
where

\[ \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \]

Let \( p_1 < p_2 \). We have

\[
\frac{f(x_1, \cdots, x_n|p_2)}{f(x_1, \cdots, x_n|p_1)} = \left[ \frac{p_2(1-p_1)}{p_1(1-p_2)} \right]^{n\bar{x}} \left[ \frac{1-p_2}{1-p_1} \right]^n
\]

So the ratio depends on the sample only through the sample mean \( \bar{x} \) and it is an increasing function of \( \bar{x} \). (It is an easy algebra exercise to check that if \( p_2 > p_1 \) then \( p_2(1-p_1)/(p_1(1-p_2)) > 1 \).)

**Example:** Now consider a normal population with unknown mean \( \mu \) and known variance \( \sigma^2 \). So the joint density is

\[
f(x_1, \cdots, x_n|\mu) = \frac{1}{(2\pi)^{n/2}\sigma^n} \exp\left( -\frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2 \right)
\]

Now let \( \mu_1 < \mu_2 \). A little algebra shows

\[
\frac{f(x_1, \cdots, x_n|\mu_2)}{f(x_1, \cdots, x_n|\mu_1)} = \exp\left( \frac{n}{\sigma^2} \bar{x}(\mu_2 - \mu_1) + \frac{(\mu_1^2 - \mu_2^2)n}{2\sigma^2} \right)
\]

So the ratio depends on \( x_1, x_2, \cdots, x_n \) only through \( \bar{x} \), and the ratio is an increasing function of \( \bar{x} \).

**Theorem 1.** Suppose \( f(x_1, \cdots, x_n|\theta) \) has a monotone likelihood ratio in the statistic \( T = t(X_1, \cdots, X_n) \). Consider hypothesis testing with alternative hypothesis \( H_a : \theta > \theta_1 \), and null hypothesis \( H_0 : \theta \leq \theta_0 \) or \( H_0 : \theta = \theta_0 \). Let \( \alpha_0, c \) be constants such that \( P(T \geq c) = \alpha_0 \). Then the test that rejects the null hypothesis if \( T \geq c \) is a UMP test at significance level \( \alpha_0 \).

**Example:** We continue the example of a normal population with known variance and unknown mean. We saw that the likelihood ratio is monotone in the sample mean. So if we reject the null hypothesis when \( \bar{X}_n \geq c \), this will be a UMP test with significance level \( \alpha = P(\bar{X}_n \geq c|\mu_0) \). Given a desired significance level \( \alpha \), we choose \( c \) so this equation holds. Then the theorem
tells us we have a UMP test. So for every $\mu > \mu_0$, our test makes $\text{Pow}(\mu)$ as large as possible.

**Example:** We continue the example of a Bernoulli distribution for the population (population proportion). To be concrete, suppose the null hypothesis is $p \leq 0.1$ and the alternative is $p > 0.1$. We have a random sample of size $n = 20$. Let $\hat{X}$ be the sample proportion. By what we’ve already done, the test that reject the null hypothesis when $\hat{X} \geq c$ will be a UMP test. We want to choose $c$ so that $P(\hat{X} \geq c) = \alpha_0$. However, $\hat{X}$ is a discrete RV (it can only be $0/20, 1/20, 2/20, \cdots, 19/20, 1$), so this is not possible. Suppose we want a significance level of 0.005 Using your favorite software (or a table of the binomial distribution) we find that $P(\bar{x} \geq 6/20|p = 0.1) = 0.0113$ and $P(\bar{x} \geq 7/20|p = 0.1) = 0.0024$. So we must take $c = 7/20$. Then the test that rejects the null if $\bar{x} \geq 7/20$ is a UMP test at significance level 0.005.

What about two-sided alternatives? It can be shown that there is no UMP test in this setting.