M4056 Hypothesis Testing. **October 25, 2010**

In statistics, a *hypothesis* is a statement about a population parameter. Any hypothesis about a parameter θ, is equivalent to a statement of the form $\theta \in \Theta_0$, where Θ_0 is a subset of the set of all possible values of θ.

Hypothesis testing refers to the study rules for accepting or rejecting hypotheses, based on information from a sample. Obviously, rules can be made at whim unless we take into account the consequences of accepting or rejecting them. In due time, we will turn to these, but first we establish some vocabulary.

Examples of hypotheses.

1. This coin is fair, i.e., a given Bernoulli variable X has parameter $p = 1/2$.
2. More than 40% of the registered voters in Baton Rouge are Republicans, i.e., the Bernoulli variable X, defined to be 1 if a randomly chosen voter in Baton Rouge is Republican and 0 otherwise, has parameter $p \geq 0.4$.
3. The mean weight of adult cats is between 9 and 11 pounds, i.e., the approximately normal random variable X, defined to be the weight of a randomly chosen adult cat has parameter μ between 9 and 11.

In testing hypotheses, it is common to call one hypothesis the *null hypothesis*, (denoted H_0) and to refer to its negation as the *alternative hypothesis* (denoted H_1). For example, if θ is a parameter being tested, the null hypothesis might be $\theta \leq \theta_0$ and the alternative would be $\theta_0 < \theta$, where θ_0 is a fixed, known value. In general, the null hypothesis is of the form $\theta \in \Theta_0$ and the alternative hypothesis is $\theta \notin \Theta_0$.

Examples of hypotheses.

1. In the coin toss example, we might take the null hypothesis to be $p = 1/2$. If we tossed the coin 100 times and it never landed on heads, we would certainly reject the null hypothesis, as we probably would if it landed on heads only 25 times. On the other hand, if it landed on heads 45 times in 100 flips, we would be less inclined to reject the null, and if it landed on heads 49 times in 100 flips, there would be very little reason to reject the null.
2. This is different from example 1, since both the null and its alternative place θ in intervals. A reasonable procedure would be to choose an unbiased estimator $\hat{p}$ for p and accept the null

$$H_0 \;\Leftrightarrow\; p \leq 0.4,$$

   if and only if $\hat{p} \leq 0.4$. However, if there are benefits associated with choosing correctly and costs associated with making errors, then we might want to use a different criterion.
3. The null hypothesis might be $9 \leq \mu \leq 11$, and the alternative: $\mu < 9$ *or* $11 < \mu$. Or vice versa. The problem in this example has been posed with little context. The choice of which alternative to call the "null" is often a matter using context. For example, if we wanted to know if a particular diet had an effect on cats that tended to modify a population parameter (average weight), we might designate the parameter range for which we would be unwilling to claim an effect the *null range*. The null hypothesis might be that, in a population of cats on the diet, the parameter lies in the null range. To reject the null hypothesis in this case would be to accept the hypothesis of efficacy of the diet in raising weight.

A **testing procedure** is a rule that specifies the sample values for which the null hypothesis is to be accepted (the *acceptance region*) and for which it is to be rejected (the *rejection region*). Many tests are based on a statistic and specify acceptance/rejection regions in the range of the statistic, T (say). This, of course, is just a special case of the general procedure, since a statistic is a function on sample space. Thus, to call for acceptance if $T(\vec{x}) \in A$ is the same as to call for acceptance if $\vec{x} \in T^{-1}(A) := \{\, \vec{\omega} \in \Omega \mid T(\vec{\omega}) \in A \,\}$.

Likelihood ratio test. Suppose the values of a parameter θ lie in the set Θ. Suppose $\Theta_0 \subseteq \Theta$ and H_0 is the hypothesis $\theta \in \Theta_0$. The *likelihood ratio test statistic* is:

$$\lambda(\vec{x}) = \frac{\sup\{\, L(\theta|\vec{x}) \mid \theta \in \Theta_0 \,\}}{\sup\{\, L(\theta|\vec{x}) \mid \theta \in \Theta \,\}},$$

and the test itself is dependent on a number c, where we reject H_0 if $\lambda(\vec{x}) \leq c$.