**M4056 Hypothesis Testing: Examples.** **November 10, 2010**

**Problem.** Tony's Seafood is having a special on 20-pound packages of frozen shrimp. The shrimp come in two sizes: jumbo (12 to the pound) and large (16 to the pound). You figure the jumbo are a good deal, but not the large. Unfortunately, the packages are not marked. Fortunately, the sales clerk will weigh four shrimp from the package you select, at your request. How can you use this information to decide whether or not to buy? (Assume both kinds of shrimp are normally distributed by weight, with a standard deviation of 1/4 ounce.)

*Solution.* Presumably, purchasing a package of large shrimp entails a loss, which you wish to avoid. To make this vivid, imagine the following scenario: Jumbo shrimp are selling at Frank's Seafood for \$3.20 per pound and large for \$2.80. Tony's is selling 20-pound packs for \$60.00 each, but you have to purchase without being sure of the size of the shrimp in the pack, so if you purchase from Tony, you risk losing \$4.00, the difference in price between 20 pounds of large shrimp purchased from Frank and 20 pounds from Tony. You decide this risk is tolerable if you can decrease the probability to 1/20. You pick a package, and get the weight of 4 shrimp from it. Depending on this weight, should you buy or not?

Note that on any purchase, you have a chance of a \$4.00 discount on jumbo shrimp. You might wish to factor that into your decision-making, perhaps allowing gains to compensate for losses. However, the hypothesis-testing paradigm we are examining is based on the supposition that you are willing follow a decision procedure that will limit your losses to a level of probability $\alpha$ that you specify, and among such procedures you demand the one that will let the fewest opportunities for bargains escape, i.e., the test with greatest power.

So, we're at Tony's and the shrimp are on the sales counter. We need to choose between the two hypotheses:

$H_0$: the package in front of us contains **large** shrimp;

$H_1$: the package in front of us contains **jumbo** shrimp.

We need to use the weight of 4 shrimp from the package to choose between these hypotheses. We will demand that the probability of making a purchase of large shrimp (as a consequence of falsely concluding that they are jumbo) must be at most $\alpha = 1/20$. In other words, we want to bound the probability of rejecting the null when it is true—a Type I error—by 1/20. This is the *level* of the test we want, often denoted $\alpha$.

The Neyman-Pearson Lemma tells us that the "best" (i.e., most powerful) test among those with level $\alpha$ is the likelihood ratio test of size $\alpha$. Now the likelihood ratio test has a rejection region

$$R_k = \{\, \vec{x} \mid f(\vec{x}|\theta_1) > k f(\vec{x}|\theta_0) \,\}. \qquad (T_k),$$

where $f(\vec{x}|\theta_i)$ is the distribution of the data under hypothesis $H_i$, $i = 1, 2$. Under $H_0$ the shrimp weight is normal with $\mu = 1$ ounce and $\sigma = 1/4$ ounce. Under $H_1$, the shrimp weight is normal with $\mu = 4/3$ ounce and $\sigma = 1/4$ ounce. Recalling that the mean is a sufficient statistic for samples from a normal population of known variance, we will base our decision on $M$, the mean weight of the four shrimp. Under $H_0$, $M$ is normal$(\mu, \sigma^2/4) =$ normal$(1, 1/64)$. Under $H_1$, $M$ is normal$(4/3, 1/64)$. Thus, we have

$$f(m \mid \theta_0) = \frac{8}{\sqrt{2\pi}} e^{-32(m-1)^2},$$

$$f(m \mid \theta_1) = \frac{8}{\sqrt{2\pi}} e^{-32(m-4/3)^2},$$

and

$$\lambda(m) = \frac{e^{-32(m-4/3)^2}}{e^{-32(m-1)^2}} = e^{32\left((m-1)^2-(m-4/3)^2\right)} = e^{32(6m-7)/9}.$$

The rejection region is defined by $\lambda(m) > k$. Notice that $\lambda(m)$ is an increasing function of $m$, so the rejection region has the form $(b, \infty)$, and $\alpha = \int_b^\infty f(m \mid \theta_0)\, dm$. If $b = 1.206$, then $\alpha$ is about 0.0497. (0.206 is about 1.65 standard deviations ($\sigma = 1/8$) above the mean the null distribution.)

**Conclusion.** By the shrimp if the sample weighs at least 4.824 ounces. There is less than a 1/20 chance of getting large shrimp instead of jumbo.

**Comment.** The power of this test is the probability of rejecting the null when it's false. This is $\int_{1.206}^\infty f(m \mid \theta_1)\, dm \cong 0.845$. This means that using this test will result in a purchase about 84.5% of the time when jumbo shrimp are presented.

**Comment.** If we knew the proportion of packages of shrimp of each kind, large or jumbo, we could calculate the long-term payoff of using a test of any given $\alpha$. Each possible outcome—both erroneous and true—would have a payoff (positive or negative), and we could adjust $\alpha$ to produce the maximum average payoff over a large number of purchases. But if we do not know the proportion of packages of each kind, then we must specify $\alpha$ according to other criteria. This is quite typical of real life. In many real life situations, $\alpha$ is chosen based simply on tradition or subjective preference.