Hypothesis Testing

Composite Hypothesis Tests

What if the distribution of Y under H_0 and/or H_1 depends on an unknown parameter?

1. Terminology/Formulae/Concepts/Caveats:

   (a) composite hypothesis:

$$\begin{aligned} H_0 &: Y \sim p(y|\theta, H_0) \qquad \theta \in \Lambda_0 \\ H_1 &: Y \sim p(y|\theta, H_1) \qquad \theta \in \Lambda_1 \end{aligned} \qquad \theta \text{ unknown!}$$

   (b) Bayesian approach:
   Say θ is random with density $w(\theta|H_j)$ under H_j. Then since

$$p_j(y) = \int_{\Lambda_j} p(y|\theta, H_j) w(\theta|H_j) d\theta$$

   we can apply the Bayesian decision rule developed earlier (assuming known costs C_{ij} and priors π_j):

$$\frac{p_1(y)}{p_0(y)} \gtrless \tau = \frac{\pi_0}{\pi_1} \frac{C_{10} - C_{00}}{C_{01} - C_{11}}$$

   (c) uniformly most powerful (UMP) test:
   When θ is treated as an unknown *deterministic* variable, we might try a NP-like criterion. Say that

$$\max_{\delta} P_D(\delta, \theta) \ \text{ s.t. } \ P_F(\delta, \theta) \leq \alpha$$

   is solved assuming fixed known θ, and the resulting rule can be expressed in a way that is *not* dependent on θ. Since such a rule is NP-optimal *for all* θ, it is called a "uniformly most powerful" (UMP) test. In many problems, however, a UMP test does not exist. (**Caution:** Derive the NP rule explicitly in terms of α and reduce it to the simplest form before claiming a UMP test doesn't exist!)

   (d) locally most powerful (LMP) test:
   If the UMP test does not exist but the hypotheses take the form

$$\begin{aligned} H_0 &: Y \sim p(y|\theta) \qquad \theta = \theta_0 \\ H_1 &: Y \sim p(y|\theta) \qquad \theta \in (\theta_0, \infty) \end{aligned} \qquad \theta \text{ deterministic and unknown} \tag{1}$$

   it may be worthwhile to adopt an NP-optimal strategy assuming that $\theta \approx \theta_0$, i.e., the "low-SNR" region. This is called the "locally most powerful" (LMP) test. Using a Taylor-series expansion of $P_D(\delta, \theta)$ around the point $\theta = \theta_0$, we found that the LMP test has the form

$$\left. \frac{\partial}{\partial \theta} \frac{p(y|\theta)}{p(y|\theta_0)} \right|_{\theta=\theta_0} \gtrless \eta \qquad \text{for } \eta \text{ chosen to yield } P_F = \alpha$$

   If the UMP test exists, since it is optimal for all θ, it must be optimal for $\theta \approx \theta_0$ and thus must be equivalent to the LMP test.

   (e) generalized likelihood ratio test (GLRT)
   Sometimes the UMP doesn't exist and the LMP is not practical. For example, the hypotheses may not be of the form (**??**), or the LMP may behave very poorly when the hypotheses satisfy (**??**) and $\theta \gg \theta_0$. An alternate strategy is to first estimate θ based on the observation y, then form a LRT based on these estimates. This strategy, when using "maximum likelihood" estimates, is known as the GLRT:

$$\frac{p(y|\hat{\theta}_1(y))}{p(y|\hat{\theta}_0(y))} \gtrless \eta \quad \text{for} \quad \begin{aligned} \hat{\theta}_1(y) &= \arg\max_{\theta \in \Lambda_1} p(y|\theta) \\ \hat{\theta}_0(y) &= \arg\max_{\theta \in \Lambda_0} p(y|\theta) \end{aligned}$$

   While not optimal in a Bayesian or NP sense, the GLRT typically yields good P_D-vs.-P_F performance.