Hypothesis Testing

Neyman-Pearson Rule

What if we can't assign the priors (π_j) and costs (C_{ij})?
Idea: Tradeoff between P_F (prob of false alarm) and P_D (prob of detection).

1. Terminology/Formulae:
   (a) power of the test: P_D = probability of detection = $P_1(\Gamma_1)$
   (b) significance or level of the test: α = the maximum allowable probability of false alarm, $P_F = P_0(\Gamma_1)$ = actual probability of false alarm
   (c) receiver operating characteristic (ROC):
   The plot of P_D vs. P_F for a given rule.
   (d) power function:
   The plot of P_D vs. d for fixed α, where d is the "separation to noise ratio." For the location testing problem, recall $d = \frac{\mu_1 - \mu_0}{\sigma}$.
   (e) randomized decision rule:
   This is a generalization of the decision rule concept. The output of the function $\tilde{\delta}(\cdot)$ is the **probability** that the argument is assigned to H_1.
   $$\tilde{\delta}(y) = \begin{cases} 1 & \text{if } L(y) > \eta \\ \gamma & \text{if } L(y) = \eta \\ 0 & \text{if } L(y) < \eta \end{cases}$$
   With this formulation,
   $$P_F(\tilde{\delta}) = E_0\{\tilde{\delta}(Y)\} = \int_\Gamma \tilde{\delta}(y) p_0(y) \mu(dy) = P_0[L(Y) > \eta] + \gamma P_0[L(Y) = \eta]$$
   $$P_D(\tilde{\delta}) = E_1\{\tilde{\delta}(Y)\} = \int_\Gamma \tilde{\delta}(y) p_1(y) \mu(dy) = P_1[L(Y) > \eta] + \gamma P_1[L(Y) = \eta]$$
   (f) Neyman-Pearson rule:
   The NP rule is a likelihood ratio test where the threshold η_0 and the randomization γ_0 are chosen such that $P_F = \alpha$ exactly.

2. Method for determining the rule:
   (a) Determine the LRT $L(y) \gtrless \eta$.
   (b) If $L(Y)$ is a continuous random variable:
      i. convert $L(y) > \eta$ into a simple relationship between y and some threshold y_η
      ii. calculate $P_F = P_0(\Gamma_1)$ as a function of y_η
      iii. find y_η such that $P_F = \alpha$
      iv. calculate the resulting $P_D = P_1(\Gamma_1)$
   (c) If $L(Y)$ is not a continuous random variable:
      i. plot $P_0[L(Y) > \eta]$ versus η
      ii. determine the discontinuities
      iii. determine if randomization is necessary for the desired α; if not, treat as continuous-$L(Y)$ case above.
      iv. determine η_0, the smallest η such that $P_0[L(Y) > \eta] < \alpha$
      v. calculate the randomization:
      $$\gamma_0 = \frac{\alpha - P_0[L(Y) > \eta_0]}{P_0[L(Y) = \eta_0]}$$
      vi. calculate the resulting $P_D = P_1(\Gamma_1)$

3. Caveats:
   (a) Even when $L(Y)$ is a continuous random variable, Γ_1 may be the union of **disjoint** sets.
   (b) The relationship between the threshold (η_0) and the level of the test (α) may not be a single function for all values of α, *i.e.* for $\alpha \in [0, 0.3], \eta = f_1(\alpha)$ and for $\alpha \in (0.3, 1], \eta = f_2(\alpha)$.
   (c) Always doublecheck whether you need to randomize.
   (d) Note: a **complete solution** to a Neyman-Pearson detection problem includes the explicit statement of the decision rule and/or critical region **and** calculation of the probability of detection achieved by the rule.