

### Algorithm for Computing the False Discovery Rate (FDR) procedure:

Consider testing  $m$  number of hypotheses. Let  $P_k$  denote the ordered p-values for the tests from the smallest to the largest ( $1 \leq k \leq m$ ). Let  $\alpha$  denote the overall (or family-wise) rate for falsely declaring a statistical significant test such as  $\alpha = 0.05$ . The procedure determines the statistical significance of a test based on the following steps:

1. Compare

$$P_i \leq \frac{i}{m} \alpha$$

2. Find the  $k$  that is the largest  $i$  for the above to hold true, i.e.,

$$P_k \leq \frac{k}{m} \alpha \text{ and } P_{k+1} > \frac{k+1}{m} \alpha$$

3. Then, reject the hypotheses corresponding to the first  $k$  smallest p-values.

The FDR procedure will ensure an FRD rate not exceeding the prescribed  $\alpha$  level.

**Example.** In the “FDR.sas” program, we considered testing a set of 21 hypotheses for a real study data. Beta estimates, standard errors and t statistics for the 21 tests are saved in the variables, a, b, c and are used to find the p-values for the tests. These p-values are then ordered and saved in the variable pt.

Since  $P_9 = 0.04964 > 0.021429 = \frac{9}{21}(0.05)$ ,  $k = 8$  and we reject the nulls corresponding to the first 8 tests ordered by the p-values.

The SAS IML is then used to compare the Bonferroni type I error level with FDR. For this example, Bonferroni type I = 0.0024 and FDR = 0.019.