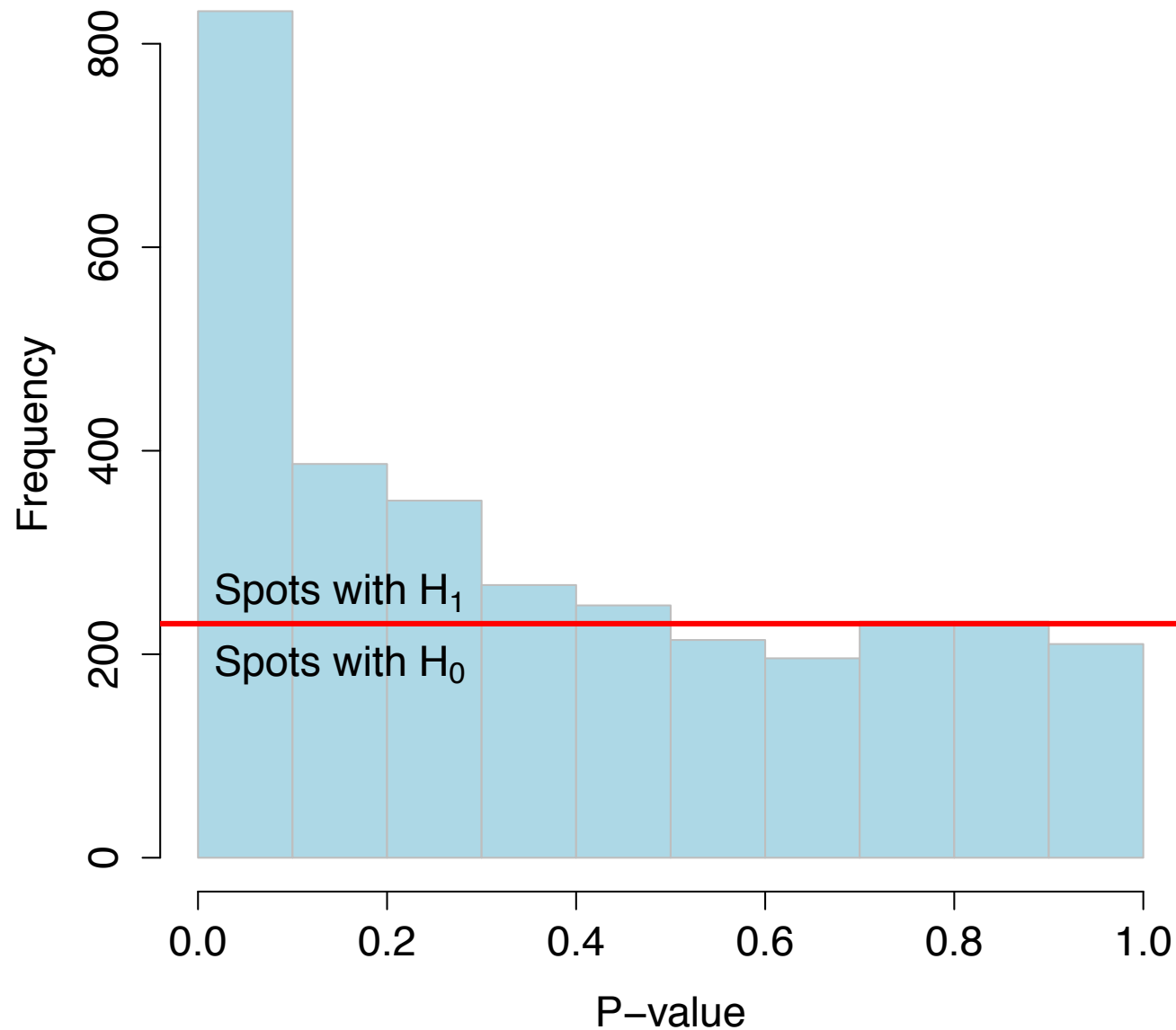


# Microarrays – False Discovery Rate

Prof. Tesler

Math 186  
Winter 2019

# $P$ -value histogram for Hedenfalk data



The distribution is approximately uniform on  $[\cdot 3, 1]$  but not on  $[0, \cdot 3]$ .

# $P$ -value distribution

- **One definition of  $P$ -value:** under  $H_0$ , what is the probability of seeing data whose test statistic is “at least this extreme”?
- **Apply this definition to the  $P$ -value itself:**  
 $P = .08$  means only 8% of the cases will be at least as extreme as the observed data. So,  $\text{Prob}(P \leq .08) = .08$ .  
In general,  $\text{Prob}(P \leq \alpha) = \alpha$ , so  **$P$  is uniform on  $[0, 1]$** .
- This assumes the data really comes from the distribution for which the “Accept  $H_0$ ” decision rule was designed.
- If the null is “true” (e.g.,  $\mu_X = \mu_Y$ ) but the distribution is not what the decision rule was designed for (e.g., not normal distribution, or incorrect  $\sigma$ ), the  $P$ -value distribution will not be uniform because the  $P$ -values were computed incorrectly or only approximately.
- Some spots follow the null while some follow the alternative. Tests should be designed so that data actually generated by the alternative has small  $P$ -values.

# Error rate for multiple hypothesis tests on an array

- At significance level  $\alpha = 0.05$ , we expect  $\approx 5\%$  of spots with no biological difference in expression levels between BRCA1 & 2 tumors will nonetheless appear to exhibit such a difference in the experiment.
- In this experiment, the arrays have  $\approx 6500$  spots, but useable data was only available for  $\approx 3200$  spots (due to image defects, etc.).
- We don't know how many of these 3200 are *truly*  $H_0$  or *truly*  $H_1$ . Most of them should be  $H_0$ , so the estimated number of false positives is  $.05(3200) = 160$ .
- There were 565  $p$ -values under 0.05. Additional mathematical and/or (labor-intensive) biological tests are required to determine which of these 565 spots are false positives.

# Multiple hypothesis tests on an array

- We simultaneously do a separate hypothesis test for every spot:  
 $H_0^{(i)}$  vs.  $H_1^{(i)}$  at sig. level  $\alpha_i$  for  $i = 1, \dots, r$ ,  
Each spot has its own Type I and Type II error.
- The *False Discovery Rate* (FDR) is the fraction of positives that are false positives.
- For  $\alpha = 0.05$ , our estimated FDR is  $\frac{160}{565} = 0.28$ .
- Vary  $\alpha$ . Estimate FDR in the same way as  $\alpha$  varies.
- For each  $\alpha$ , we can determine how many positives and estimate what fraction (FDR) are false positives, then pick how many we have the resources to do the additional tests for.

# Estimated FDR as $\alpha$ varies

