**6.1**

*There is no applicable example from my scientific research that requires multiple testing, the example below comes from Chapter 8 of Modern Statistics for Modern Biology by Susan Holmes & Wolfgang Huber.*

The goal of this multiple testing example to identify genes that are differentially abundant between the treated & untreated cells. Specifically, the data came from an experiment on *Drosophila melanogaster* cell cultures that investigated the effect of RNAi knock-down of the splicing factor *pasilla* [1].

The null hypothesis is that there is not any gene that is differentially abundant between the treated & untreated cells under *t*-test. As shown in Figure 1, the false discovery rate is roughly 11%.

A screenshot of a social media post

Description automatically generated

Figure 1. *p* value histogram

R code:

*#pasilla experiment path*

*fn = system.file("extdata", "pasilla\_gene\_counts.tsv",*

*package = "pasilla", mustWork = TRUE)*

*#expression counts for each gene and treatment combination*

*counts = as.matrix(read.csv(fn, sep = "\t", row.names = "gene\_id"))*

*#meta info file path*

*annotationFile = system.file("extdata",*

*"pasilla\_sample\_annotation.csv",*

*package = "pasilla", mustWork = TRUE)*

*#meta info for the count data above*

*pasillaSampleAnno = readr::read\_csv(annotationFile)*

*#clean up*

*library("dplyr")*

*pasillaSampleAnno = mutate(pasillaSampleAnno,*

*condition = factor(condition, levels = c("untreated", "treated")),*

*type = factor(sub("-.\*", "", type), levels = c("single", "paired")))*

*mt = match(colnames(counts), sub("fb$", "", pasillaSampleAnno$file))*

*stopifnot(!any(is.na(mt)))*

*library("DESeq2")*

*pasilla = DESeqDataSetFromMatrix(countData = counts,colData = pasillaSampleAnno[mt, ], design = ~condition)*

*#run mutiple -t- tests*

*pasilla = DESeq(pasilla)*

*res = results(pasilla)*

*#p value histogram*

*library(ggplot2)*

*ggplot(as(res, "data.frame"), aes(x = pvalue)) +*

*geom\_histogram(binwidth = 0.01, fill = "Royalblue", boundary = 0)+geom\_hline(yintercept=110, linetype="dashed", color = "red")*

Reference

[1] Holmes S, Wolfgang H: Modern Statistics for Modern Biology. Cambridge University Press. 2018.