

Methods Homework 6

Tim Vigers

October 10, 2018

A) Load `gvhd.txt` into R, then subset the data to focus on only transplant recipients with an HLA-matched sibling donor.

```
# Read in and subset the data.
gvhd <- read.table("~/School/UC Denver/Biostatistics/Biostatistical Methods 1/Homework 6/gvhd.txt", sep=";", as.is=T)
# Subset to just HLA matched siblings.
hla.matched <- gvhd[gvhd$hla.matched.sibling == 1, c(1,3,4)]
```

B) Calculate the proportion of recipients that got GvHD in the Treatment A group. Repeat for Treatment B.

```
# Subset by treatment.
treat.a <- hla.matched[hla.matched$treatment == "A",]
treat.b <- hla.matched[hla.matched$treatment == "B",]
# Calculate proportions.
prop.treat.a <- sum(treat.a$outcome) / length(treat.a$outcome)
prop.treat.b <- sum(treat.b$outcome) / length(treat.b$outcome)
prop.treat.a
```

```
## [1] 0.2905983
```

```
prop.treat.b
```

```
## [1] 0.220339
```

C) Among transplant recipients with HLA-matched donors, is there a significant association between treatment and GvHD at the 5% level of significance? Carry this test out using both a permutation test, and either an exact or asymptotic method, as appropriate. Summarize your results and comment on differences, if any, between the two methods you applied.

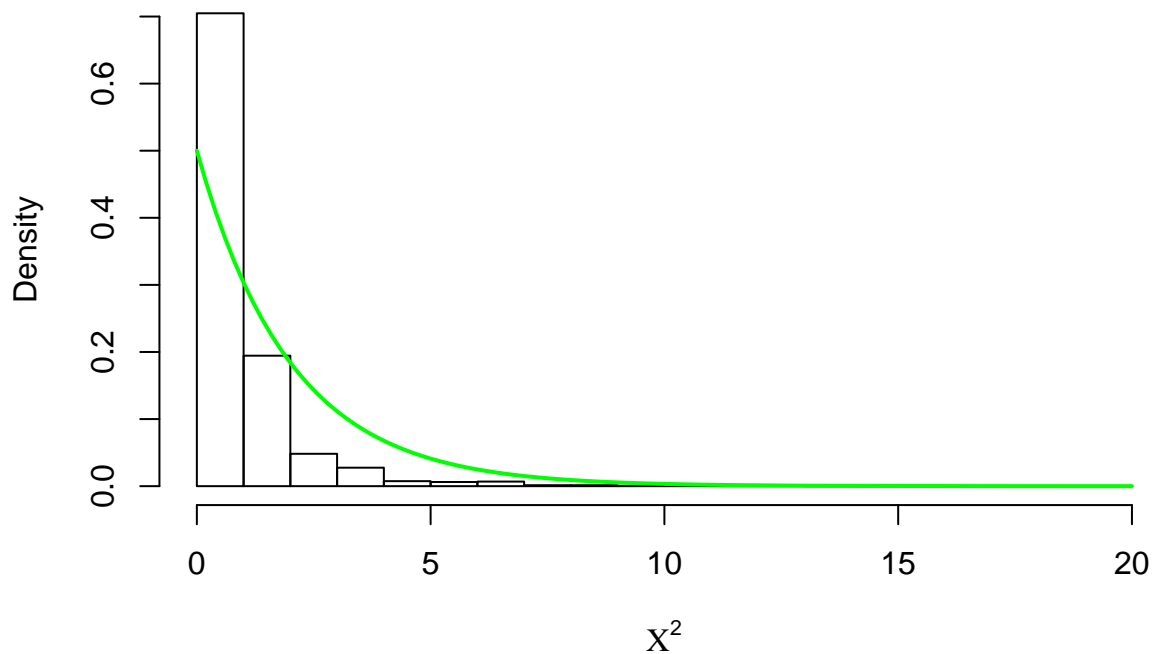
```
# Make a vector to store permutation results.
B <- 10^5 - 1
result <- numeric(B)
# Repeat the chi-square test on B permutations.
for (i in 1:B) {
  permuted <- sample(hla.matched$treatment)
  table <- table(permuted, hla.matched$outcome)
```

```

test <- chisq.test(table)
result[i] <- test$statistic
}
# Plot.
hist(result, freq=FALSE, xlab = expression(Chi^2), main="Permutation distribution for chi-square statis
curve(dchisq(x, 2), add=TRUE, col="green", lwd=2)

```

Permutation distribution for chi-square statistic



```

# Find the test statistic for the original table.
observed <- chisq.test(table(hla.matched[,2],hla.matched[,3]))$statistic
# Compare the p value from the permutation distribution to the p-value from the
# chi-square distribution.
perm.p <- (sum(result >= observed)+1)/(B + 1)
dist.p <- 1-pchisq(observed, df=1)

```