Problem 1 The Suspicious Card Game

Rakesh Vishwabrahmana

Question 1

Solution:

(a)

Let us define two wining population proportions p_{you} and $p_{opponent}$ for you and your opponent, respectively. We are interested in whether a game of chance is fair. Therefore, the expected winning proportion will be the same for both players, i.e., the expected earnings per play come to 0.

Therefore, the unknown population parameter of interest for this problem is the difference in population proportions $p_{opponent} - p_{you}$. Hence, the test statistic or the point estimate of interest is the difference in sample proportions $\hat{p}_{opponent} - \hat{p}_{you}$.

(b)

We can formulate our hypothesis test using the mathematical notation for our population parameter of interest, the difference in population proportions $p_{opponent} - p_{you}$.

$$
\begin{aligned}
H_0 \quad &: p_{opponent} - p_{you} = 0 \\
H_A \quad &: p_{opponent} - p_{you} = 0 > 0
\end{aligned}
$$

(c)

We have generated 1000 replicates of draws two cards for 5 rounds, assuming the null hypothesis is true. We can calculate test statistic $\hat{p}_{opponent} - \hat{p}_{you}$ for each of the replicate, which gives a vector of 1000 difference of sample proportion. The observed difference of proportion or observed test statistics $\hat{p}_{opponent} - \hat{p}_{you}$ is $4/5 - 0 = 0.8 = 80\%$. Now we calculate *p-values*, which state the test statistic is extreme or more extreme than the actual result from the scenario under the null hypothesis using a simulation experiment.

```r
# For computation convenience instead of alphabet(A, J, Q, K)
# we consider integer value (1, 11, 12, 13).
# Function for drawing 2 cards under null hypothesis
two_card <- function(){
  card <- sample(1:13, size = 2, replace = FALSE)
  return(max(card))
}
n_round <- 1000
# Function for Null estimate
prop_Ho <- function(n_round) {
you <- as.vector(replicate(n = n_round, two_card()))
```

```
opp <- as.vector(replicate(n = n_round, two_card()))
you_win <- vector(length = n_round)
for (i in 1:n_round) {
  if (you[i] > opp[i]) {
    you_win[i] <- "WIN"
  } else if (you[i] < opp[i]) {
    you_win[i] <- "LOSS"
  } else you_win[i] <- "DRAW"
}

res <- data.frame(you ,opp, you_win)
win_prob <- sum(res$you_win == "WIN")/n_round
loss_prob <- sum(res$you_win == "LOSS")/n_round
dif_prob <- win_prob - loss_prob
return(dif_prob)
}
test_stat <- replicate(n = 1000, prop_Ho(5))
p_value <- 1- pnorm(0.8,mean= mean(test_stat), sd = sd(test_stat))
hist(test_stat, probability = TRUE)
abline(v = 0.80, lty = 2, lwd = 2, col = 2)
```

The resulting *p-values* is 0.0309695.

(d)

Since this *p-value* is smaller than the significance level $\alpha = 0.05$, we conclude alternative hypothesis $p_{opponent} > p_{you}$. In other words, this *p-value* is sufficiently small to reject our hypothesis that your opponent may not be cheating.

Question 2

Solution:

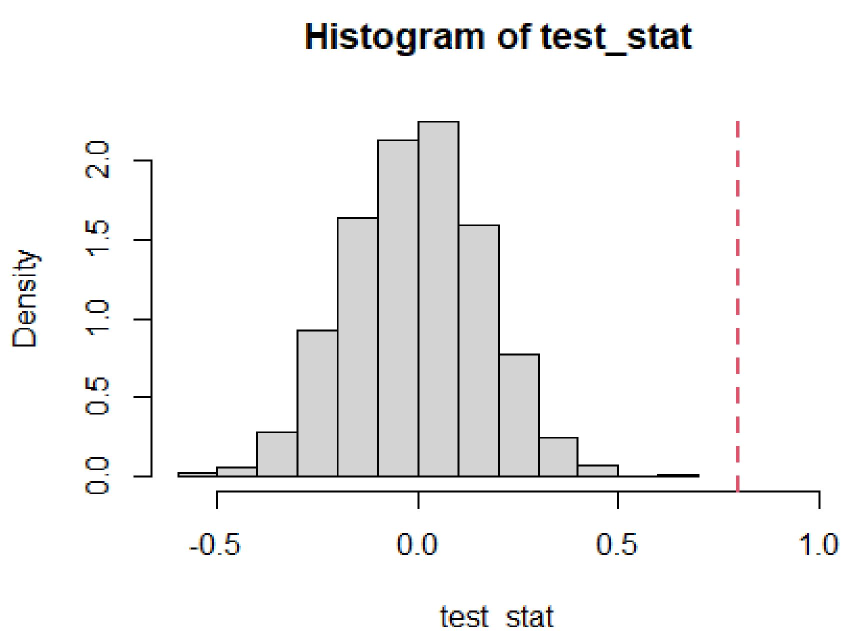
The procedure we followed fully satisfied the general statistical framework. Based on the observed data, we calculated observed test statistics and then found the test statistics distribution assuming the null hypothesis is true. Finally, we find the critical region of the test to make the decision.

Question 3

Solution:

Since the sample size is too small to test that your opponent may be cheating, we will consider several rounds to make our conclusion more secure. So let us consider 30 rounds instead of 5. But, first, let see how the *p-values* and the null distribution change.

```
test_stat <- replicate(n = 1000, prop_Ho(30))
p_value <- 1- pnorm(0.8,mean= mean(test_stat), sd = sd(test_stat))
hist(test_stat,xlim = c(min(test_stat),1), probability = TRUE)
abline(v = 0.80, lty = 2, lwd = 2, col = 2)
```



The *p-value* is reduced to 1.3302967^{-6} for 30 draws with stranger.