Estimations of Distance to Moon Anchored by Radius of Earth and Distance to Sun

Abstract

Anchoring is demonstrated in a Bunker Hill class of 24 students. Each student was asked to estimate the distance to the moon; group 1 was anchored on the distance to the sun (9e7 miles) while group 2 anchored with radius of Earth (3e3 miles). Due to heavy skew in group 1, the normality-based pvalue was 0.1 while permutation-based methods allowed us to conclude that anchoring caused an effect (p-value = 0.03).

Background

Say some things about anchoring.

Statistical tests often assume sampling distributions are normal. This condition is easily met with moderate sample sizes from nearly any population (even with a population as unnormal as bimodal or exponential distributions). However, with small sample sizes, the normal approximation to sampling distributions easily fails.

Luckily, permutation tests are more rubust. With a modern version of R, permutation tests are intuitive to code. The 24 students were tasked with determining the results of a permutation test on their own original data; 10 were able to implement the permutation test compared to 15 who were able to implement a two-proportion t-test.

```
prop.test(x=c(10,15),n=c(24,24))
```

```
##
##  2-sample test for equality of proportions with continuity
##  correction
##
## data:  c(10, 15) out of c(24, 24)
## X-squared = 1.3357, df = 1, p-value = 0.2478
## alternative hypothesis: two.sided
## 95 percent confidence interval:
##  -0.5264379  0.1097712
## sample estimates:
##    prop 1    prop 2
## 0.4166667 0.6250000
```

It seems (from the p-value) more data is necessary to determine a preference. It should be mentioned that the class these students were in focused on normal-based techniques (like standard introductory classes usually do). However, the concept of permutation (randomization) tests was introduced in the first chapter with a card-shuffling description/embodiment of the procedure. (make reference to OpenIntro)

Experimental Methods

Every participant was asked to estimate the distance to the moon. By having the instructions on cards, and asking the responses to be written on the cards, all participants could participate simultaneously. The participants were given approximately 30 seconds to write an answer. All participants responded.

The two types of cards gave different hints. The groups (Earth and Sun) were randomly assigned by shuffling the cards. By using cards, it was easy to guarantee equal sample sizes. Half of the participants were told the radius of Earth (3,000 miles) while the other half were told the distance to the Sun (90,000,000 miles). The cards were made with the printout that can be generated (or modified) with the R script below.

```
#This code will create a page that can be printed (saved in the current working directory).
#The resulting pdf/printout is in the supplemental materials.
pdf("Moon_cards.pdf",width=8.5,height=11,paper="letter",title = "Moon cards printout")
par(mar = c(0,0,0,0))
plot.new()
fig.align="center"
for(xpos in seq(0,1,1/3)){
  abline(v = xpos)
}
for(ypos in seq(0,1,1/4)){
  abline(h = ypos)
}
for(x in seq(0,2/3,1/3)){
  for(y in seq(0,3/4,1/4)){
    text(x+0.16,y+0.20,"Please estimate the distance\nto the moon \nin miles.", cex=0.8)
  lines(c(x+0.05,x+0.3333-0.05),c(y+0.07,y+0.07))
  if (y %in% c(0,0.5)){
    text(x+0.16,y+0.03,"Hint:\nThe sun is 90,000,000 miles away.", cex=0.8)
  } else {
    text(x+0.16,y+0.03,"Hint:\nThe radius of Earth is 3,000 miles.", cex=0.8)
  }
  }
}
dev.off()
```

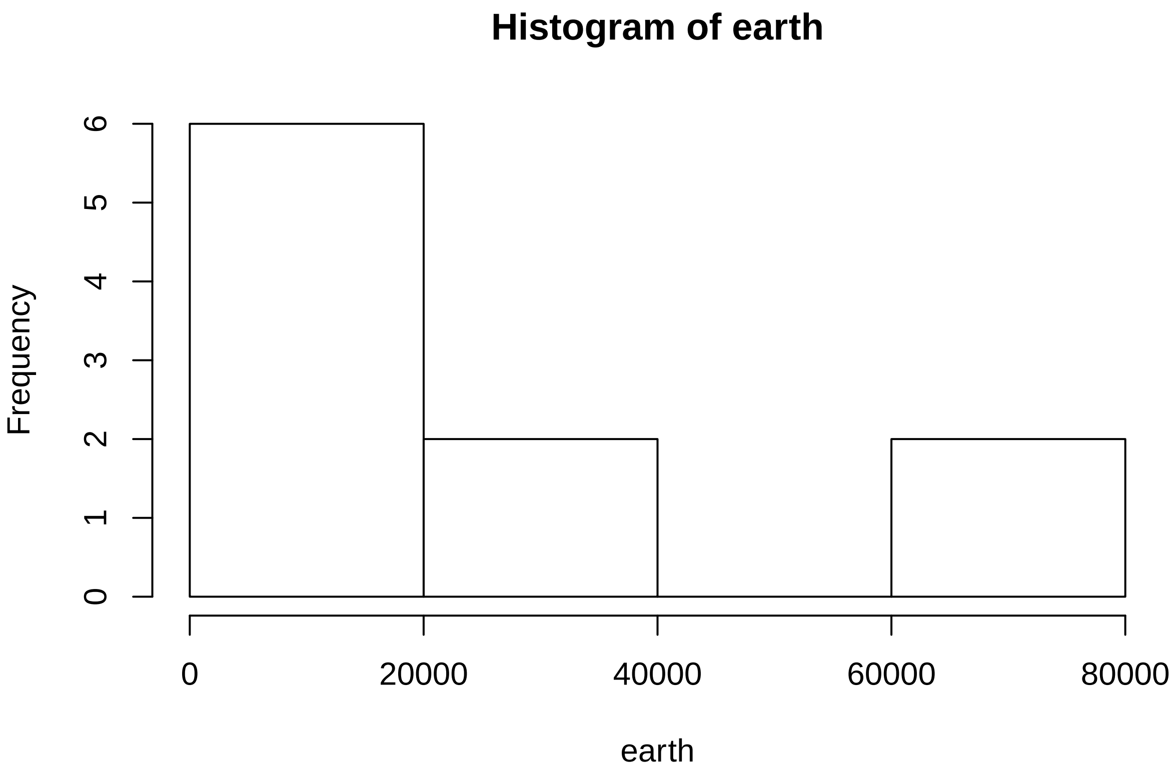
Please estimate the distance to the moon in miles. ____________________ Hint: The radius of Earth is 3,000 miles.	Please estimate the distance to the moon in miles. ____________________ Hint: The radius of Earth is 3,000 miles.	Please estimate the distance to the moon in miles. ____________________ Hint: The radius of Earth is 3,000 miles.
Please estimate the distance to the moon in miles. ____________________ Hint: The sun is 90,000,000 miles away.	Please estimate the distance to the moon in miles. ____________________ Hint: The sun is 90,000,000 miles away.	Please estimate the distance to the moon in miles. ____________________ Hint: The sun is 90,000,000 miles away.
Please estimate the distance to the moon in miles. ____________________ Hint: The radius of Earth is 3,000 miles.	Please estimate the distance to the moon in miles. ____________________ Hint: The radius of Earth is 3,000 miles.	Please estimate the distance to the moon in miles. ____________________ Hint: The radius of Earth is 3,000 miles.
Please estimate the distance to the moon in miles. ____________________ Hint: The sun is 90,000,000 miles away.	Please estimate the distance to the moon in miles. ____________________ Hint: The sun is 90,000,000 miles away.	Please estimate the distance to the moon in miles. ____________________ Hint: The sun is 90,000,000 miles away.

The statistical analysis took two forms: Welch's t test and an approximate permutation method.

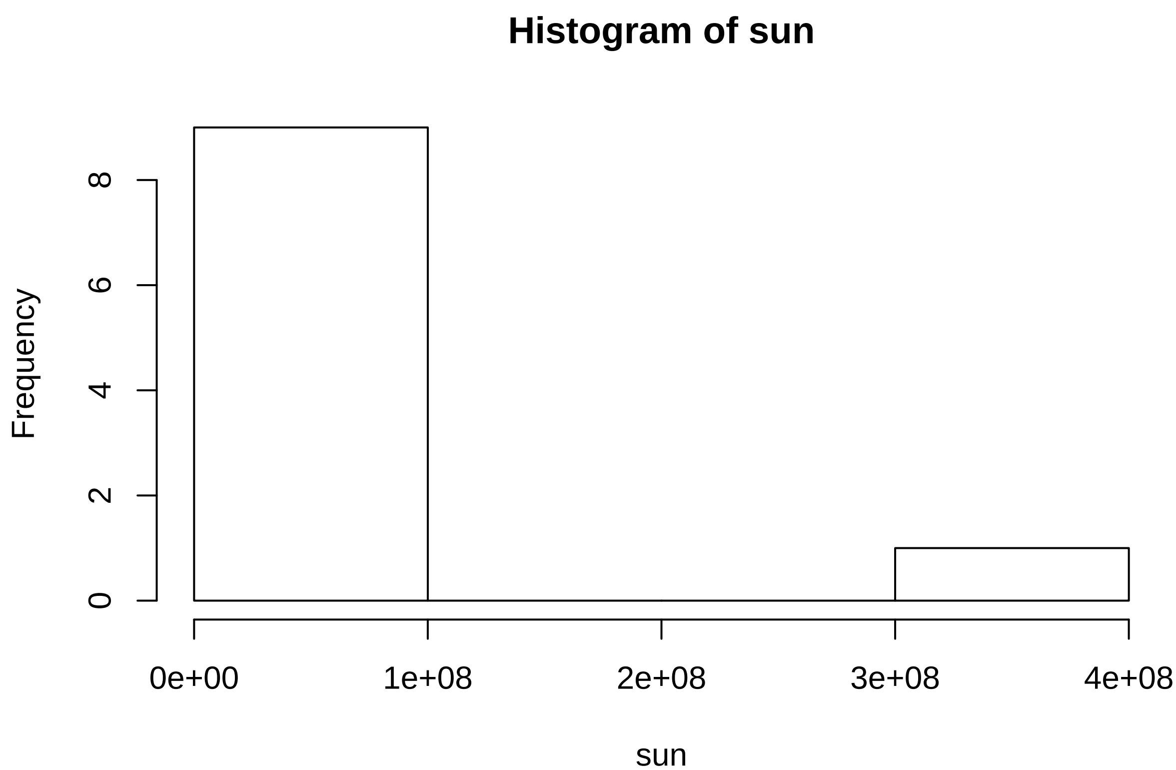
Results

The sample data are shown here.

```
earth = c(7.5e4,2e4,299,39590,3900,75000,30000,8197,15836,8100)
sun = c(1.23e7,1.85e5,2.2e5,45730003,1.2e5,6e5,4e8,2.05e7,48372213,3e5)
hist(earth)
```




```
hist(sun)
```



Notice both of these groups appear to have a heavy skew, so we should be wary of using an approach that assumes (repeated) small samples would be normally distributed. We continue with the Welch's t test as an exercise.

We find the means and sample standard deviations.

```
xbar1 = mean(earth)
s1 = sd(earth)
n1 = length(earth)
xbar2 = mean(sun)
s2 = sd(sun)
n2 = length(sun)
```

Group	Mean	Standard deviation	Size
Earth	2.75922×10^4	2.7694669×10^4	10
Sun	5.2832722×10^7	1.2341899×10^8	10

Discussion

```
t.test(earth,sun)
```

```
##
##  Welch Two Sample t-test
##
## data:  earth and sun
## t = -1.353, df = 9, p-value = 0.2091
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -141093758   35483499
## sample estimates:
```

```
##  mean of x  mean of y 
##    27592.2 52832721.6
```

```
nshuffles = 1000
od = xbar2-xbar1
rds = c()
for(i in 1:nshuffles)
{
    tot1 = sum(sample(c(earth,sun),n1))
    tot2 = sum(c(earth,sun))-tot1
    mean1 = tot1/n1
    mean2 = tot2/n2
    rds = c(rds,mean2-mean1)
}
print(sum(abs(rds)>abs(od))/nshuffles)
```

```
## [1] 0
```