CS146 - Week 10

Code

Hide

```
rm(list = ls())
set.seed(23)
```

provided:

Posterior for control group: $\text{Beta}(\theta_c \mid 58, 65)$.

Hide

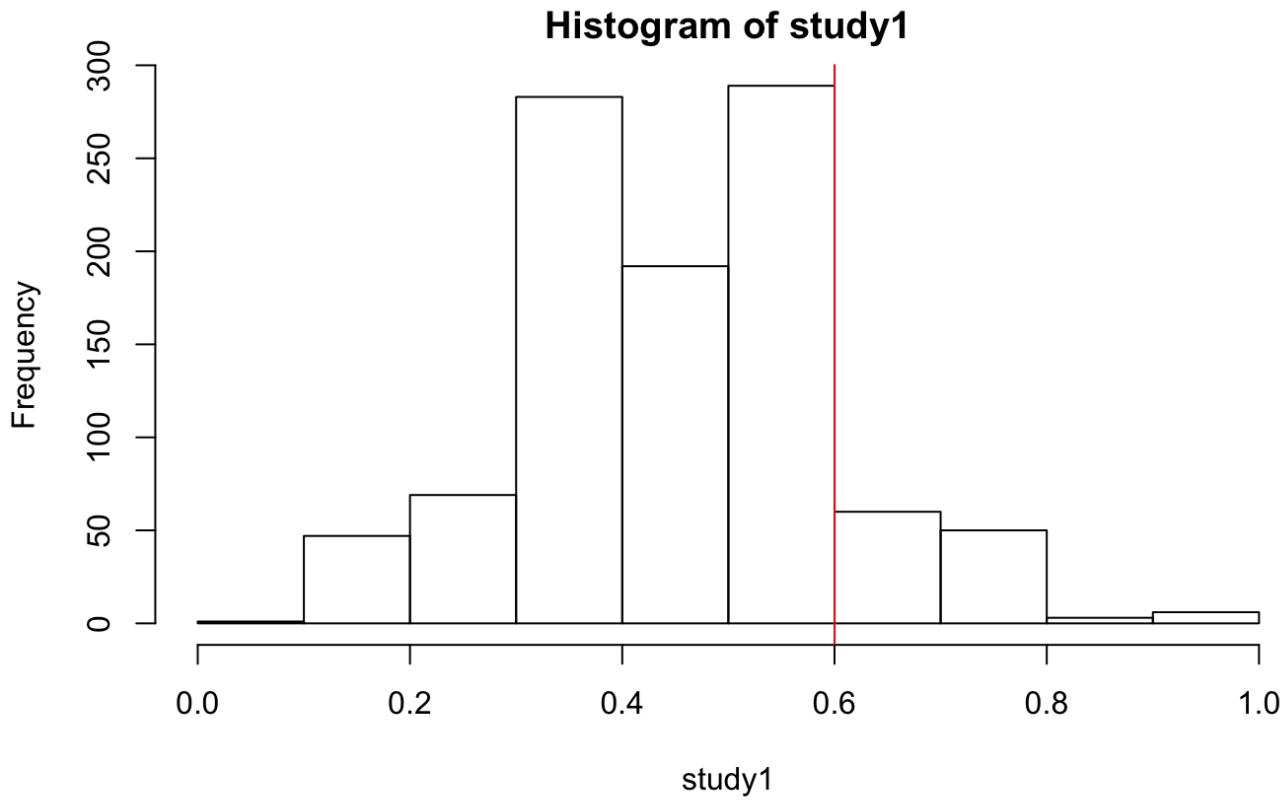
```
samples <- 1000 #number of simulations
produce_theta_from_posterior_control <- rbeta(samples, 58,65)
produce_theta_from_posterior_control
```

Studies (the data):

Study	Treatment group	Control group
Di Rienzo 2014	20 / 23	9 / 15
Galli 1994	10 / 16	11 / 18
Kaufman 1974	13 / 16	4 / 10
Qin 2014	35 / 45	21 / 39
Sanchez 2012	22 / 31	12 / 29
Silny 2006	7 / 10	0 / 10
Totals	107 / 141	57 / 121

Hide

```
study <- function(observations,total,success_is_theta){
  rbinom(observations,total,success_is_theta)
  }
#Generating the fake data per experiment:
#normalizing would be the test statistic. (T)
#the test statistic used is S/N (success over totals)
study1 <- study(samples,15,produce_theta_from_posterior_control)/15
study2 <- study(samples,18,produce_theta_from_posterior_control)/18
study3 <- study(samples,10,produce_theta_from_posterior_control)/10
study4 <- study(samples,39,produce_theta_from_posterior_control)/39
study5 <- study(samples,29,produce_theta_from_posterior_control)/29
study6 <- study(samples,10,produce_theta_from_posterior_control)/10
#Study 1:
hist(study1)
abline(v=9/15,col="red")
```



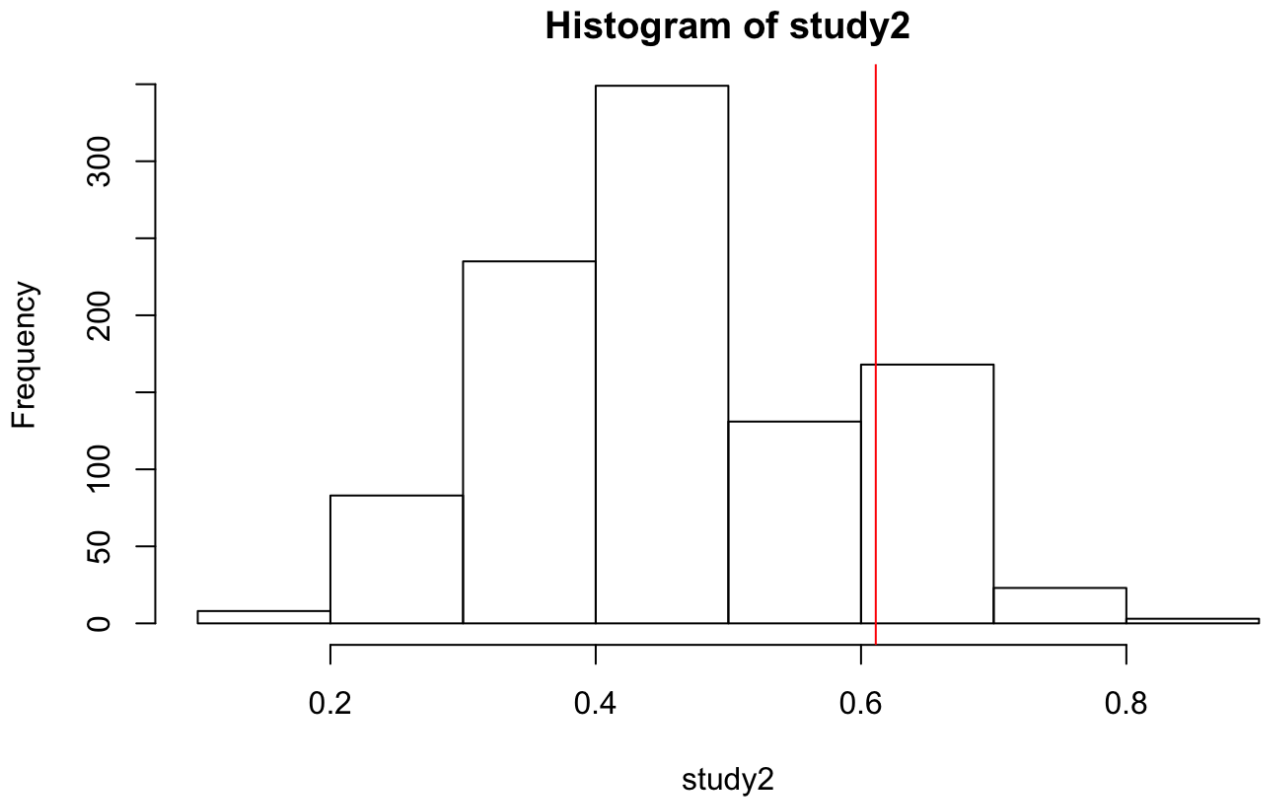
Hide

```
#P value:
mean(study1<.6)
```

```
[1] 0.763
```

Hide

```
#Study 2:
hist(study2)
abline(v=11/18,col="red")
```



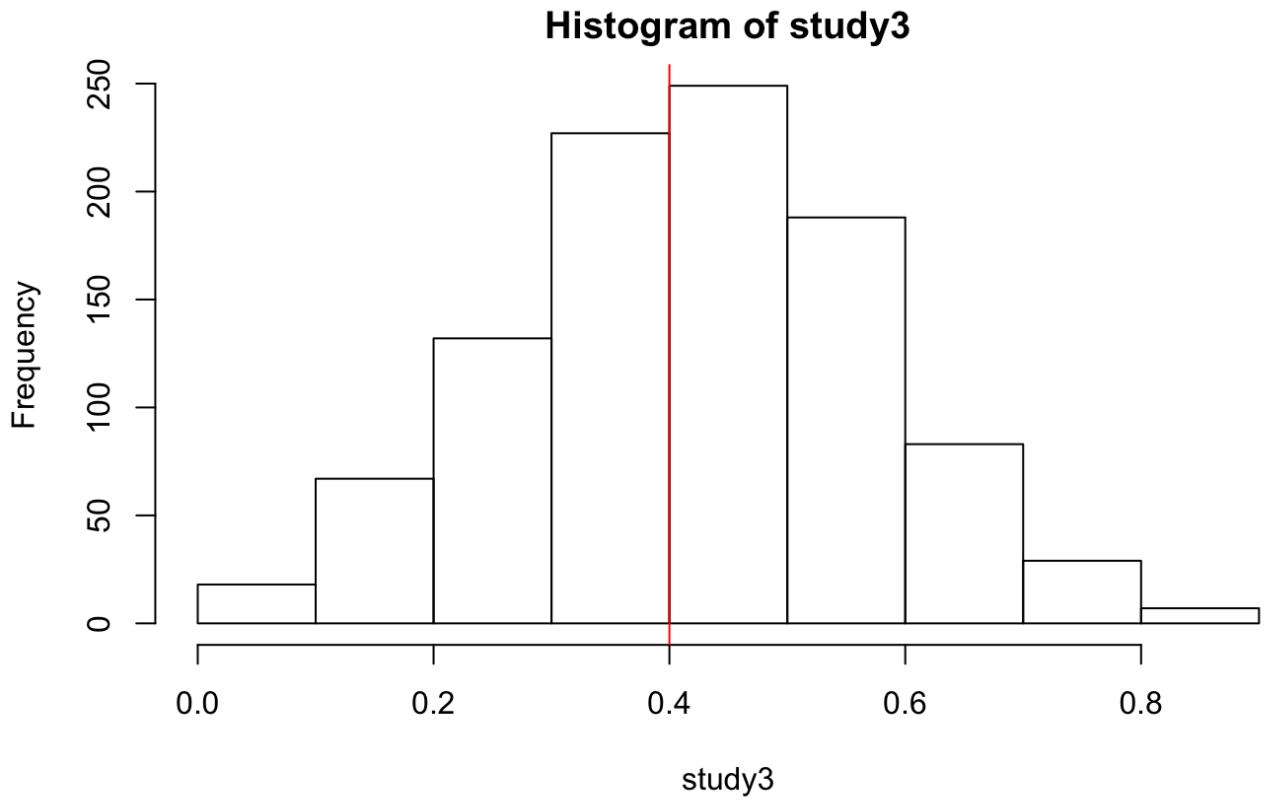
Hide

```
#P value:
mean(study1<11/18)
```

```
[1] 0.881
```

Hide

```
#Study 3:
hist(study3)
abline(v=4/10, col="red")
```



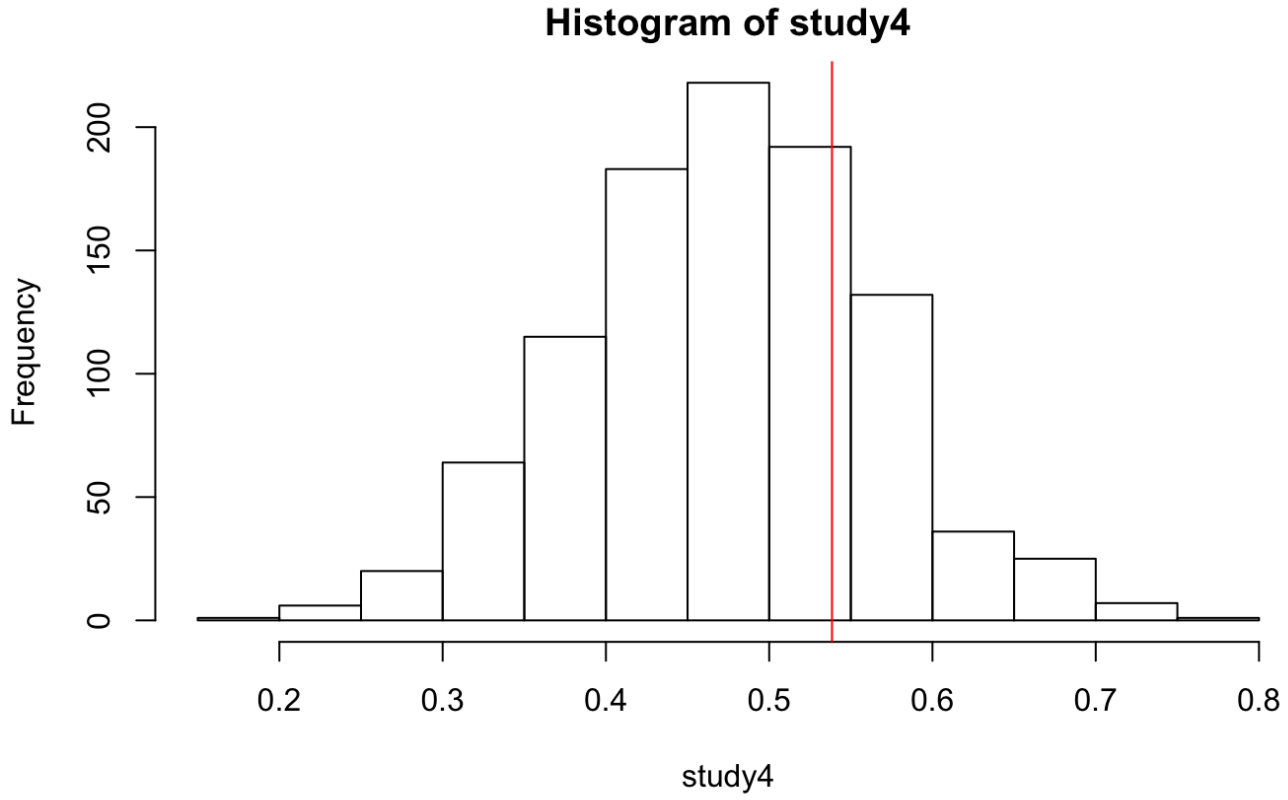
Hide

```
#P value:
mean(study1<4/10)
```

```
[1] 0.234
```

Hide

```
#Study 4:
hist(study4)
abline(v=21/39, col="red")
```



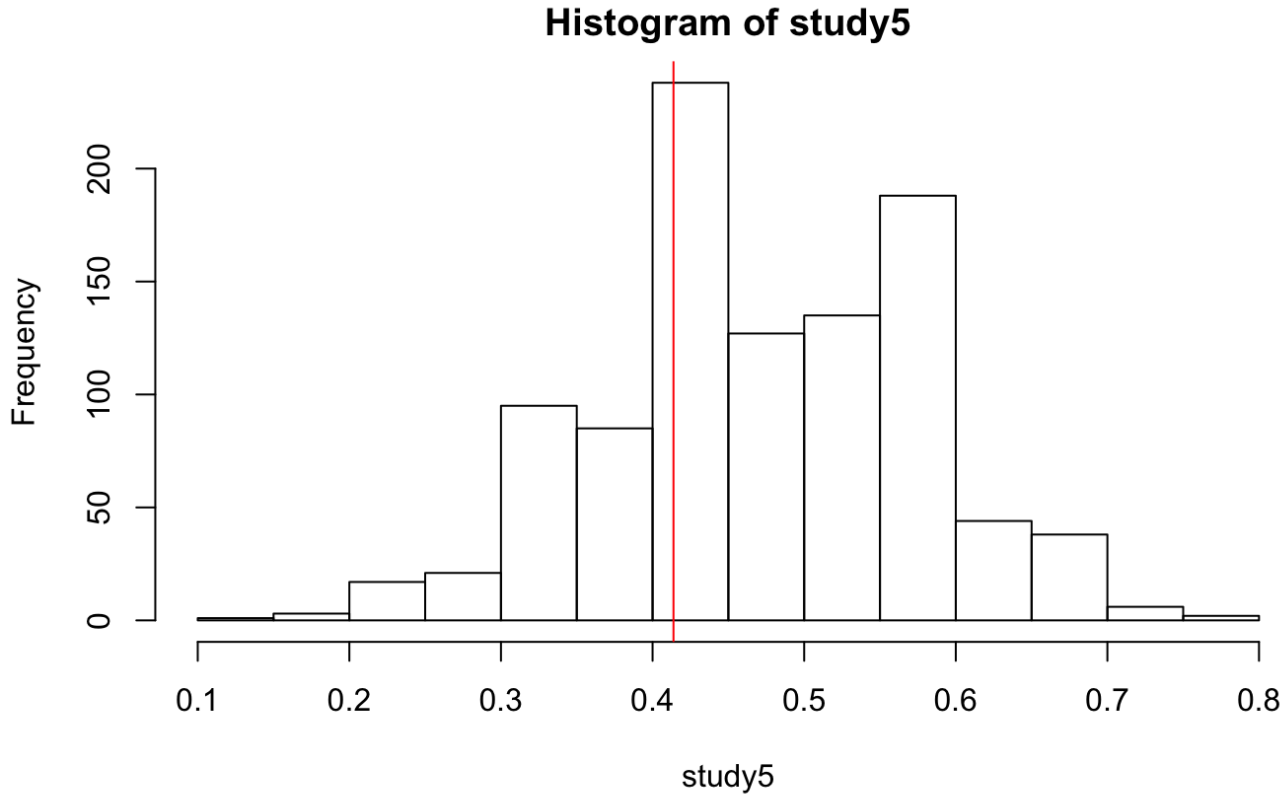
Hide

```
#P value:
mean(study1<21/39)
```

```
[1] 0.763
```

Hide

```
#Study 5:
hist(study5)
abline(v=12/29, col="red")
```



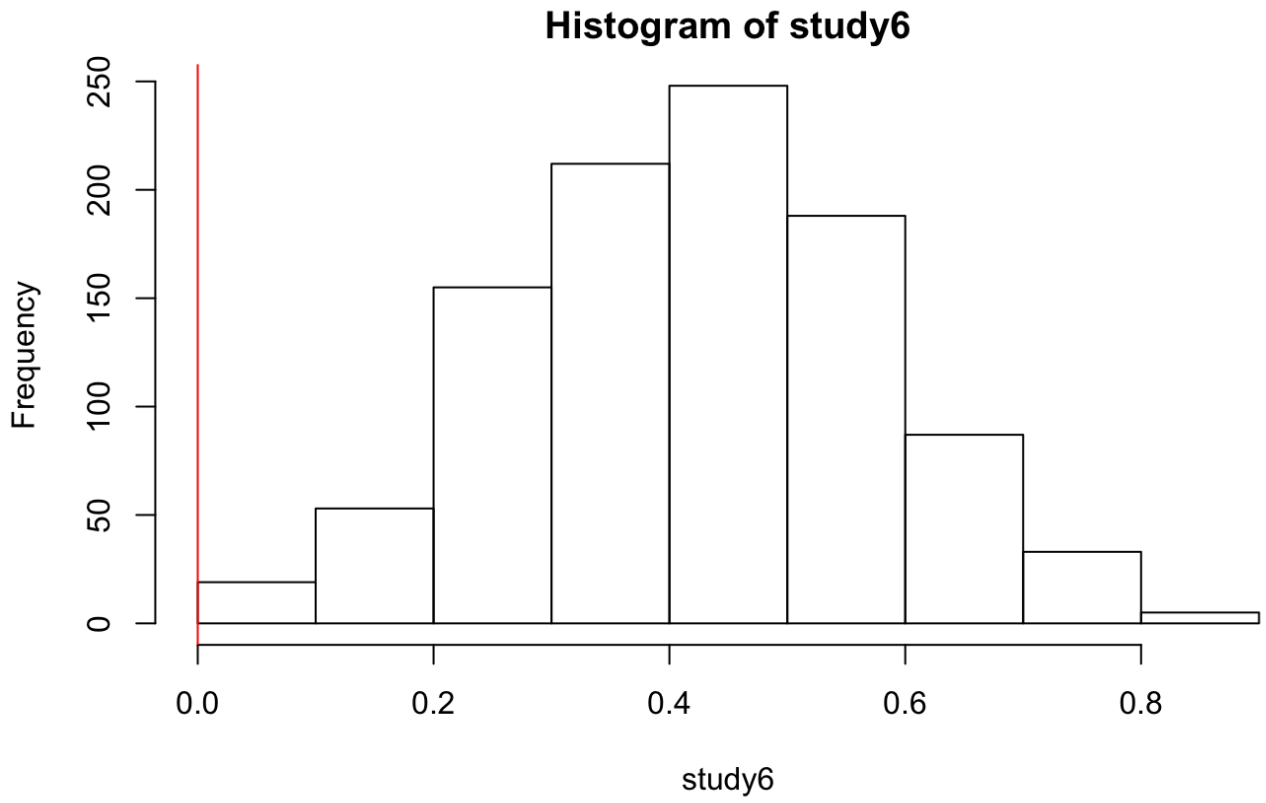
Hide

```
#P value:
mean(study1<12/29)
```

```
[1] 0.4
```

Hide

```
#Study 6:
hist(study6)
abline(v=0/10, col="red")
```



Hide

```
#P value:
mean(study1<0/10)
```

```
[1] 0
```

We see that the p-value, and the graph of the generated data and the real data of the 6th study point at the discrepancy: the expected value of the test statistic is is too far on the distribution of the generated values (p-value 0). This model cannot accuratly generate plausible data for this study.