Example: Is Yawning Contagious? I

	Yawn Seed	No Seed
Yawned	10	4
No Yawn	24	12

	Yawn Seed	No Seed	Sum
Yawned	10	4	14
No Yawn	24	12	36
Sum	34	16	50

Joint distribution

(proportions for each combination of the two variables)

	Yawn Seed	No Seed
Yawned	0.20	0.08
No Yawn	0.48	0.24

Example: Is Yawning Contagious? II

	Yawn Seed	No Seed	Sum
Yawned	10	4	14
No Yawn	24	12	36
Sum	34	16	50

Conditional proportions

By rows? By columns?

Which should we use as evidence yawning is contagious?

	Yawn Seed	No Seed
Yawned	0.714	0.286
No Yawn	0.667	0.333

	Yawn Seed	No Seed
Yawned	0.294	0.25
No Yawn	0.706	0.75

Example: Is Yawning Contagious? III

	Yawn Seed	No Seed	Sum
Yawned	10	4	14
No Yawn	24	12	36
Sum	34	16	50

	Yawn Seed	No Seed	Sum
Yawned	0.714	0.286	1
No Yawn	0.667	0.333	1

	Yawn Seed	No Seed
Yawned	0.294	0.25
No Yawn	0.706	0.75
Sum	1.000	1.00

Testing the Hypothesis I

Researchers hypothesis: Yawning is contagious.

"Null Hypothesis" (objective starting point, no effect of yawning)

If no association between yawn seed and yawning,
then 14 yawners were just tired
and would have yawned anyway,
no matter which group they were assigned to.

In that case (null hypothesis, no effect):
the only source of variability in the number who yawn (or not)
in this study comes from the experimenters' intervention:
randomizing people to one group (yawn seed) or the other (no seed).

Testing the Hypothesis II

Set up "cards" to represent the 14 yawners (red) and 36 non-yawners (black)

```
cards <- c(rep("red", 14), rep("black", 36))
cards
```

```
 [1] "red"   "red"   "red"   "red"   "red"   "red"   "red"
 [8] "red"   "red"   "red"   "red"   "red"   "red"   "red"
[15] "black" "black" "black" "black" "black" "black" "black"
[22] "black" "black" "black" "black" "black" "black" "black"
[29] "black" "black" "black" "black" "black" "black" "black"
[36] "black" "black" "black" "black" "black" "black" "black"
[43] "black" "black" "black" "black" "black" "black" "black"
[50] "black"
```

Shuffle (sample) 34 cards to represent randomization of people into the yawn seed treatment group

Testing the Hypothesis III

```
shuffle.ttt <- sample(cards, 34)
shuffle.ttt
```

```
 [1] "black" "red"   "black" "black" "red"   "red"   "red"
 [8] "black" "black" "black" "red"   "black" "black" "black"
[15] "black" "black" "black" "black" "black" "black" "black"
[22] "black" "black" "black" "red"   "red"   "red"   "black"
[29] "black" "black" "black" "black" "black" "black"
```

Testing the Hypothesis IV

Count how many red cards (yawners) end up in this group by chance alone.

```
shuffle.ttt == "red"
```

```
 [1] FALSE  TRUE FALSE FALSE  TRUE  TRUE  TRUE FALSE FALSE
[10] FALSE  TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[19] FALSE FALSE FALSE FALSE FALSE FALSE  TRUE  TRUE  TRUE
[28] FALSE FALSE FALSE FALSE FALSE FALSE FALSE
```

```
sum(shuffle.ttt == "red")
```

```
[1] 8
```

```
ttt.count <- sum(shuffle.ttt == "red")
ttt.count
```

```
[1] 8
```

Testing the Hypothesis V

Repeat this simulation (shuffle) of the cases 20 times

```
ttt.count.20 <- replicate(20, sum(sample(cards, 34)=="red"))
ttt.count.20
```

```
 [1] 11  8  8  8  9  9 10  8 11 11  8 10 10 10  6  9 11 10
[19] 11 10
```

How often did we get the same result as in the Mythbusters' data?

```
sum(ttt.count.20 == 10)
```

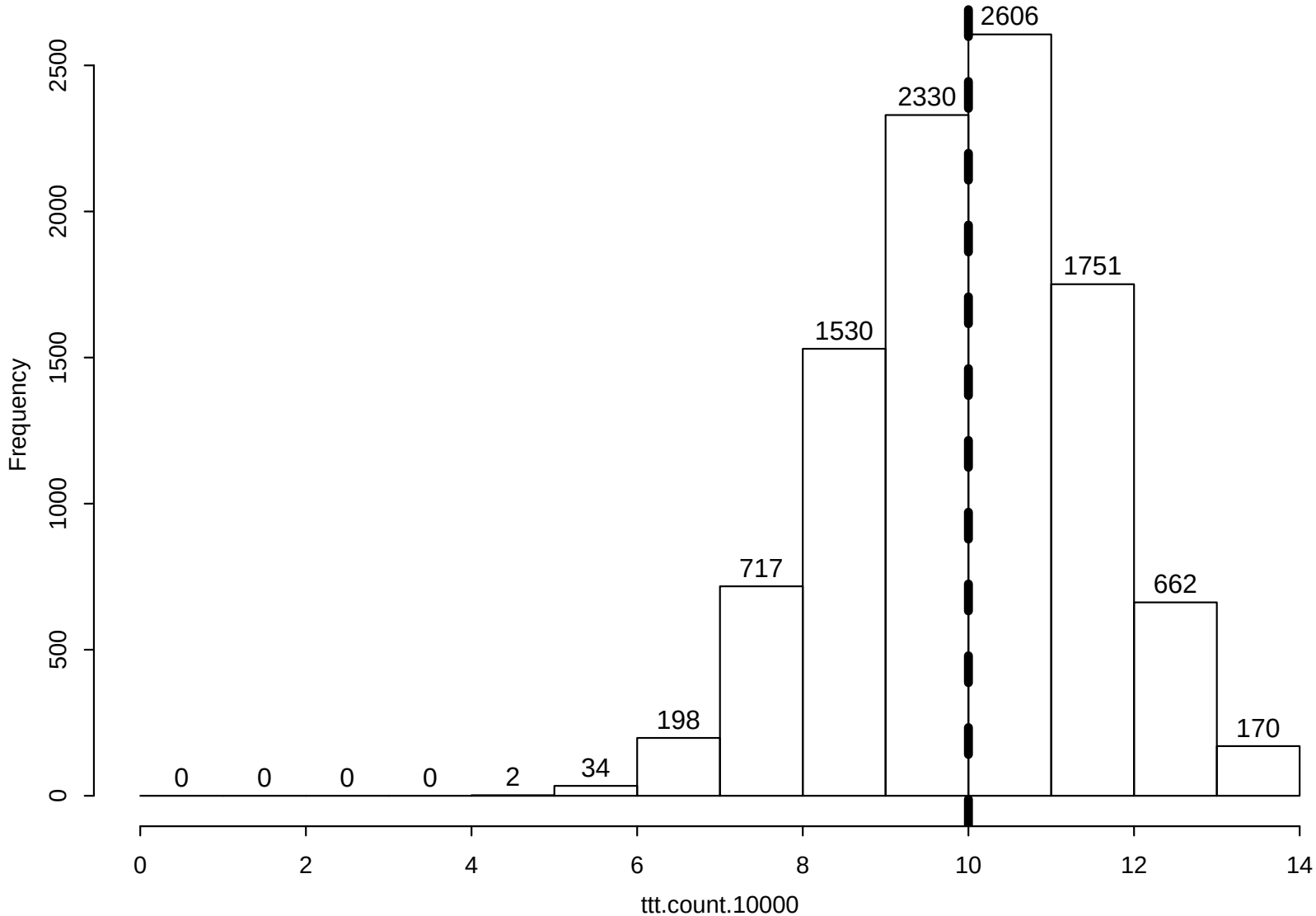
```
[1] 6
```

```
sum(ttt.count.20 == 10) / 20
```

```
[1] 0.3
```

Testing the Hypothesis VI

Now, repeat the simulation (shuffle) of the cases 10,000 times



Testing the Hypothesis VII

How statisticians measure "evidence" against the null hypothesis:

How often would we expect to observe data as (or more) "extreme" than the data we actually observed?
("as extreme" means "as far from expected")

```
sum(ttt.count.10000 >= 10)
```

```
[1] 5189
```

```
sum(ttt.count.10000 >= 10) / 10000
```

```
[1] 0.519
```

http://www.rossmanchance.com/applets/ChiSqShuffle.html?yawning=1