Introduction to Hypothesis Testing

Practice in defining hypotheses

For each of the following, determine whether it represents a null hypothesis claim or an alternative hypothesis claim:

1. King cheetahs on average run the same speed as standard spotted cheetahs.

2. For a particular student, the probability of correctly answer a 5-option multiple choice test is larger than 0.2 (i.e. better than guessing)

3. The probability of getting in a car accident is the same if using a cell phone then if not using a cell phone.

4. The number of hours that grade-school children spend doing homework predicts their future success on standardized tests.

Write out the null and alternative hypotheses in words and also in statistical notation for each of the following situations:

1. New York is known as "the city that never sleeps' '. A random sample of 25 New Yorkers were asked how much they sleep they get per night. Does these data providing convincing evidence that New Yorkers on average sleep less than 8 hours per night?

2. A study suggests that 25% of 25 year-olds have gotten married. You believe that this is incorrect and decide to conduct your own analysis.

Do a minority of Middlebury students drink coffee regularly?

Hypotheses, significance level, and data (steps 1 and 2)

H_0 :

H_A :

$\alpha =$

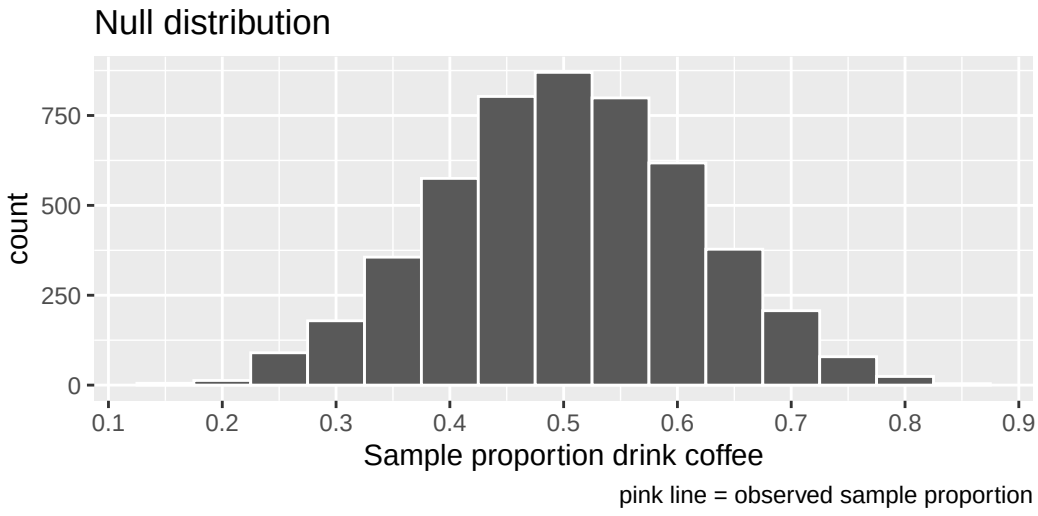
The following is our data:

```
x <- c(rep(1, 7), rep(0, 13))
p_hat <- mean(x)
```

Simulating under the null (step 3)

```
set.seed(2)
B <- 5000
n <- length(x)
null_props <- rep(NA, B)
for(b in 1:B){
  null_samp <- sample(x = c("pink", "purple"),
                      size = n,
                      replace = T,
                      prob = c(0.5, 0.5))
  null_props[b] <- sum(null_samp == "pink")/n
}
```

Visualize the null distribution of $\hat{p}$:



What does “as or more extreme” mean in this problem?

p-value:

Conclusion in context (step 4)

Decision and rationale:

Conclusion in context:

Possible scenarios

		State of world	
		H_0 true	H_0 false
Decision	Fail to reject H_0		
	Reject H_0		