Problem set

Xinyi Lin

9/25/2019

## Problem 1

### Question a

To calculate a sample size, we need additional information including:

1. The dependent variable is approximately normally distributed within each group.
2. The data is collected from a representative, randomly selected portion of the total population.

### Question b

If we assume sample size of two groups are the same. As In order to have 5% significance,

## Problem 1

### Question a

14\*pbinom(0, 14, 0.05) + 20\*(1-pbinom(0, 14, 0.05))

## [1] 17.07395

The expected value of the sample size is around 17.07.

### Question b

Let be the number of responses in the first or second stages, where .

The probability of a “go” decision is

.

pbinom(3, 20, 0.05, lower.tail = FALSE)\*dbinom(0, 14, 0.05) + pbinom(2, 20, 0.05, lower.tail = FALSE)\*dbinom(1, 14, 0.05) + pbinom(1, 20, 0.05, lower.tail = FALSE)\*dbinom(2, 14, 0.05) + pbinom(0, 20, 0.05, lower.tail = FALSE)\*dbinom(3, 14, 0.05) + pbinom(4, 14, 0.05, lower.tail = FALSE)

## [1] 0.08438283

So the probability of a “go” decision is around 0.084.

### Question c

pbinom(3, 20, 0.2, lower.tail = FALSE)\*dbinom(0, 14, 0.2) + pbinom(2, 20, 0.2, lower.tail = FALSE)\*dbinom(1, 14, 0.2) + pbinom(1, 20, 0.2, lower.tail = FALSE)\*dbinom(2, 14, 0.2) + pbinom(0, 20, 0.2, lower.tail = FALSE)\*dbinom(3, 14, 0.2) + pbinom(4, 14, 0.2, lower.tail = FALSE)

## [1] 0.7580235

When teh true response rate is 20%, the probability of a “go” decision is around 0.758.