

Calculations notebook

Benjamin Banks

October 26, 2023,

Pre calculations

In the pre trail we have 20 total obs. Load the pre data:

```
library(readxl)
preData <- read_excel("C:/Users/Benja/dev/dtu/Image-segmentation/data/data.xlsx",
  sheet = "Initial", col_types = "numeric")
```

The means of them:

```
mean(preData$Accuracy)
```

```
## [1] 0.7604741
```

And the standard deviation

```
sd(preData$Accuracy)
```

```
## [1] 0.1508467
```

We will use the above information in a power analysis, to calculate how many observations our final experiment requires.

```
power.t.test(power=0.8, delta = 0.05, sd = 0.2159476, sig.level = 0.05)
```

```
##
##      Two-sample t test power calculation
##
##              n = 293.7792
##              delta = 0.05
##              sd = 0.2159476
##              sig.level = 0.05
##              power = 0.8
##      alternative = two.sided
##
## NOTE: n is number in *each* group
```

So the we will make 300 observations pr. model.

Final calculations

Load the final data:

```
library(readxl)
finalData <- read_excel("C:/Users/Benja/dev/dtu/Image-segmentation/data/data.xlsx",
  sheet = "Final", col_types = "numeric")
```

Find the mean:

```
colMeans(finalData)
```

```
## init_k_means_1 init_k_means_2
##      0.7716235      0.8648475
```

T-test with initial_k_means_iterations = 1 and initial_k_means_iterations = 2

```
t.test(finalData$init_k_means_1, finalData$init_k_means_2)
```

```
##
##  Welch Two Sample t-test
##
## data:  finalData$init_k_means_1 and finalData$init_k_means_2
## t = -7.6172, df = 498.88, p-value = 1.312e-13
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.11726972 -0.06917834
## sample estimates:
## mean of x mean of y
## 0.7716235 0.8648475
```