Homework 2

Jiaming Chen

12/19/2017

1

calculate n

n = 13.09777

```
power.t.test(delta = 1, sd = 1, power = 0.8, type = "two.sample", alternative = "one.sided")
```

```
##
##      Two-sample t test power calculation
##
##               n = 13.09777
##           delta = 1
##              sd = 1
##       sig.level = 0.05
##           power = 0.8
##     alternative = one.sided
##
## NOTE: n is number in *each* group
```

How n change along with Δ?

n decrease when Δ increase.

```
power.t.test(delta = 2, sd = 1, power = 0.8, type = "two.sample", alternative = "one.sided")
```

```
##
##      Two-sample t test power calculation
##
##               n = 3.987012
##           delta = 2
##              sd = 1
##       sig.level = 0.05
##           power = 0.8
##     alternative = one.sided
##
## NOTE: n is number in *each* group
```

```
power.t.test(delta = 3, sd = 1, power = 0.8, type = "two.sample", alternative = "one.sided")
```

```
##
##      Two-sample t test power calculation
##
##               n = 2.462449
##           delta = 3
##              sd = 1
##       sig.level = 0.05
##           power = 0.8
##     alternative = one.sided
##
```

```
## NOTE: n is number in *each* group
```

```
power.t.test(delta = 4, sd = 1, power = 0.8, type = "two.sample", alternative = "one.sided")
```

```
##
##      Two-sample t test power calculation
##
##               n = 1.993898
##           delta = 4
##              sd = 1
##       sig.level = 0.05
##           power = 0.8
##     alternative = one.sided
##
## NOTE: n is number in *each* group
```

How n change along with σ_A^2?

n increase when σ_A^2 increase.

```
power.t.test(delta = 1, sd = 0.7, power = 0.8, type = "two.sample", alternative = "one.sided")
```

```
##
##      Two-sample t test power calculation
##
##               n = 6.85034
##           delta = 1
##              sd = 0.7
##       sig.level = 0.05
##           power = 0.8
##     alternative = one.sided
##
## NOTE: n is number in *each* group
```

```
power.t.test(delta = 1, sd = 1, power = 0.8, type = "two.sample", alternative = "one.sided")
```

```
##
##      Two-sample t test power calculation
##
##               n = 13.09777
##           delta = 1
##              sd = 1
##       sig.level = 0.05
##           power = 0.8
##     alternative = one.sided
##
## NOTE: n is number in *each* group
```

```
power.t.test(delta = 1, sd = 2, power = 0.8, type = "two.sample", alternative = "one.sided")
```

```
##
##      Two-sample t test power calculation
##
##               n = 50.1508
##           delta = 1
##              sd = 2
##       sig.level = 0.05
##           power = 0.8
```

```
##     alternative = one.sided
##
## NOTE: n is number in *each* group
```

2

We should use Chi-Square test, because we're checking if a random number follows a certain distribution.

N=101

	0 six	1 six	2 six	3 six
expected	58.58	35.35	7.07	0.4646
observed	48	35	15	3

$\chi^2 = 1.91 + 0.0035 + 8.89 + 13.84 = 24.6435$

$df = (2-1) \times (4-1) = 3$

p-value is 1.833077e-05

```
1-pchisq(24.6435,3)
```

```
## [1] 1.833077e-05
```

3

Feature engineering:

The new features can be generated by combining existing features or splitting existing features. For example, for the categorical features with too many levels, we can combine some levels to form a new categorical feature with less levels. We can also combine correlated features to avoid colinearity.