29.2: Comparing Two Means (Section 28.2)

To compare two means from independent samples, we can use the two-sample t-test. Let's say that we want to compare blood pressure of smokers and non-smokers; we don't have an expectation for the direction, so we will use a two-sided test. First let's perform a power analysis, again for a small effect:

```
0.20.8power_results_2sample <- pwr.t.test(d=, power=,
                                    type='two.sample'
                                    )
power_results_2sample
```

```
##
##      Two-sample t test power calculation
##
##               n = 393
##               d = 0.2
##       sig.level = 0.05
##           power = 0.8
##     alternative = two.sided
##
## NOTE: n is number in *each* group
```

This tells us that we need 394 subjects in each group, so let's sample 394 smokers and 394 nonsmokers from the NHANES dataset, and then put them into a single data frame with a variable denoting their smoking status.

```
nonsmoker_df <- NHANES_adult %>%
  dplyr::filter(SmokeNow=="Yes") %>%
  drop_na(BPSysAve) %>%
  dplyr::select(BPSysAve,SmokeNow) %>%
  sample_n(power_results_2sample$n)

smoker_df <- NHANES_adult %>%
  dplyr::filter(SmokeNow=="No") %>%
  drop_na(BPSysAve) %>%
  dplyr::select(BPSysAve,SmokeNow) %>%
  sample_n(power_results_2sample$n)

sample_df <- smoker_df %>%
  bind_rows(nonsmoker_df)
```

Let's test our hypothesis using a standard two-sample t-test. We can use the formula notation to specify the analysis, just like we would for `lm()` .

```
t.test(BPSysAve ~ SmokeNow, data=sample_df)
```

```
##
##  Welch Two Sample t-test
##
## data:  BPSysAve by SmokeNow
## t = 4, df = 775, p-value = 3e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  2.9 7.8
## sample estimates:
##  mean in group No mean in group Yes
##               125               120
```

This shows us that there is a significant difference, though the direction is surprising: Smokers have *lower* blood pressure!

Let's look at the Bayes factor to quantify the evidence:

```
sample_df <- sample_df %>%
  mutate(SmokeNowInt=as.integer(SmokeNow))
ttestBF(formula=BPSysAve ~ SmokeNowInt,
        data=sample_df)
```

```
## Bayes factor analysis
## --------------
## [1] Alt., r=0.707 : 440 ±0%
##
## Against denominator:
##   Null, mu1-mu2 = 0
## ---
## Bayes factor type: BFindepSample, JZS
```

This shows that there is very strong evidence against the null hypothesis of no difference.