PSYC*6060: Differences between two groups (t-test)

1 Traditional Power: Assuming you know the population effect size

Consider an example, you are interested in the effect of a drug and conducted a study with a drug group and a control group. For a variety of reasons, you are convinced that the true population effect size is $d = .90$. Typically, we want an 85% chance of finding such an effect if it exists. Use R to determine the appropriate same size for this study.

Note that is you are not sure of the effect size then use .20 (a small effect) to be conservative (or one of the various more contextualized small effect alternatives we have discussed). In this case though, we believe the population effect size is .90.

```
library(pwr)
pwr.t.test(d=0.90,power=0.85,sig.level=0.05,type="two.sample",alternative="greater")
```

```
##
##      Two-sample t test power calculation
##
##               n = 18.4679
##               d = 0.9
##       sig.level = 0.05
##           power = 0.85
##     alternative = greater
##
## NOTE: n is number in *each* group
```

2 Safeguard power: Basing your effect size on a study result

You want to conduct an independent-samples t-test. Past research indicates that the typical effect size in your research area is $d = 0.90$ using $N = 15$ per group (i.e., the standardized mean difference is .90).

If a d-value is not reported in previous research you can easily calculate one using the command below. To calculate a d-value you need for each group the mean (Mean.1, Mean.2), the standard deviations (s.1, s.2), and the group sizes (n.1, n.2). Note that the command below *smd* stands for *standardized mean difference* which is just another way of saying d-value. More specifically, a d-value is an index of the standardized mean difference.

```
library(MBESS)
smd(Mean.1=3.4, Mean.2=2.499, s.1=.95,s.2=1.05, n.1=15,n.2=15)
```

```
## [1] 0.8998759
```

Note that d-values can also be calculated from raw data as we will do below.

Previous research indicated a $d = .90$ with $N = 15$ per cell. The first step is to calculate a confidence interval:

```
library(MBESS)
ci.smd(smd=0.90, n.1=15, n.2=15)
```

```
## $Lower.Conf.Limit.smd
## [1] 0.1395921
##
## $smd
## [1] 0.9
##
## $Upper.Conf.Limit.smd
## [1] 1.645916
```

Based on this analysis we see the sample d =.90 could plausibly been caused by a population d-value as low as 0.1396 or as high as 1.646. Power analyses are very sensitive to decimals so we use the full range of decimals from this confidence interval in our power analysis.

We use the lower bound of the confidence interval (0.1396) in our power analysis.

```
library(pwr)
pwr.t.test(d=0.1396,power=0.85,sig.level=0.05,type="two.sample",alternative="greater")
```

```
##
##      Two-sample t test power calculation
##
##               n = 738.4894
##               d = 0.1396
##       sig.level = 0.05
##           power = 0.85
##     alternative = greater
##
## NOTE: n is number in *each* group
```

Thus, a safeguard analysis indicates that we need $N = 739$ per cell. Note if the original study had used more than $N = 15$ per cell, the confidence interval would have been much narrower and consequently the safeguard N much smaller.

3 Confidence Interval Width: Assuming you know the population effect size

Often it's better to think in terms of how wide a confidence interval you would be comfortable with if your predicted effect size is correct. In this example you believe the population d-value is .50. Therefore you want to ensure the width of the confidence interval isn't wider than .50.

.

```
library(MBESS)
ss.aipe.smd.full(delta=.90, conf.level=.95,width=.90)
```

```
## [1] 42
```

Thus, to ensure the width of the confidence interval does not exceed the effect size, you need $N = 42$ participants PER GROUP (assuming the population effect size is $d = .90$)

4 Safeguard Confidence Interval Width: Basing your effect size on a study result

Previous research indicated a $d = .90$ with $N = 15$ per cell. The first step is to calculate a confidence interval:

```
library(MBESS)
ci.smd(smd=0.90, n.1=15, n.2=15)
```

```
## $Lower.Conf.Limit.smd
## [1] 0.1395921
##
## $smd
## [1] 0.9
##
## $Upper.Conf.Limit.smd
## [1] 1.645916
```

Based on this analysis we see the sample $d = .90$ could plausibly been caused by a population d-value as low as 0.1396 or as high as 1.646. But if you are basing the .90 on a sample estimate, with 95% CI [0.1396, 1.646], then you would do the following.

```
library(MBESS)
ss.aipe.smd.full(delta=0.1396, conf.level=.95,width=0.1396)
```

```
## [1] 1581
```

This more realistic estimate suggests you need a sample size of 1581 PER GROUP. You can see sample size requirement increase substantially when the population effect size is small.

4.1 What if the first study had a better design?

Imagine a scenario where the original study used a larger sample size but observed the same effect ($d = 0.90$). Specifically, 100 per cell was used in the original study rather than 15 per cell. The resulting confidence interval would be smaller:

```
ci.smd(smd=0.90, n.1=100, n.2=100)
```

```
## $Lower.Conf.Limit.smd
## [1] 0.6079717
```

```
##
## $smd
## [1] 0.9
##
## $Upper.Conf.Limit.smd
## [1] 1.189957
```

If the original study had used the large $N = 100$ per cell, the confidence interval would be .60 to 1.19. Thus, the safeguard power analysis would be:

```
library(MBESS)
ss.aipe.smd.full(delta=0.60, conf.level=.95,width=0.60)
```

```
## [1] 90
```

This sample size of 90 PER cell, in comparison to 1581, seems quite reasonable. However, it only occurs if the original study used a large N per cell. Thus, if you conduct a sample N study not only do you create problems for interpreting your own data, you create problems for subsequent researchers that do a safeguard sample size analysis.

5 Conducting the t-test

Load and view the data:

```
library(tidyverse)
my.data <- read_csv("drugData.csv")
head(my.data)
```

```
## # A tibble: 6 × 3
##      ID Arousal Group
##   <int>   <dbl> <int>
## 1     1    39.1     0
## 2     2    38.0     0
## 3     3    14.9     0
## 4     4    20.7     0
## 5     5    19.5     0
## 6     6    32.2     0
```

Ensure you convert categorical variables to factors (this is a critical step).

```
my.data$Group <- as.factor(my.data$Group)
glimpse(my.data)
```

```
## Observations: 64
## Variables: 3
## $ ID      <int> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, ...
## $ Arousal <dbl> 39.1, 38.0, 14.9, 20.7, 19.5, 32.2, 11.0, 20.7, 26.4, 35.7, 26.4, 28.8, 33.4, 1...
## $ Group   <fctr> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ...
```

Now assign factor levels:

```
levels(my.data$Group) <- list("Drug"=0, "Control"=1)
```

The first step in actually conducting the t-test is to determine if the variances are equal for the two groups. There are different t-test formulas that are used if the variances are equal or not.

5.1 Descriptive Statistics

```
psych::describeBy(my.data$Arousal,group=my.data$Group)
```

```
## group: Drug
##    vars  n mean   sd median trimmed   mad min  max range skew kurtosis   se
## 1     1 35   24 12.2   20.7   23.38 13.79 5.3 54.1  48.8 0.47    -0.69 2.06
## ------------------------------------------------------------------------
## group: Control
##    vars  n mean   sd median trimmed   mad   min  max range  skew kurtosis   se
## 1     1 29 16.5 11.8     18   17.05 10.23 -15.5 35.8  51.3 -0.64      0.1 2.19
```

You will need the mean (M) and standard deviation (SD) from this table for each group when reporting results.

5.2 Homogeneity of Variance Assumption Test

The test requires the *car* package (Companion to Applied Regression) - be sure to install it. Because the *car* package conflicts with the *tidyverse* package we don't use the library command when using it. Rather we use the notation below.

```
car::leveneTest(my.data$Arousal, group=my.data$Group,center="median")
```

```
## Levene's Test for Homogeneity of Variance (center = "median")
##       Df F value Pr(>F)
## group  1  0.2561 0.6146
##       62
```

In this case, the homogeneity of variance assumption is true (i.e., Levene's test was non-significant). So we proceed accordingly.

Regardless of the approach used though we need to select select the score for the drug group and control group separately. We do this we the commands below to select subgroups based on the column *Group* in my.data.

```
library(tidyverse)

#Arousal scores for drug group
exp.group.rows <- my.data %>% filter(Group=="Drug")

#Arousal scores for control group
control.group.rows <- my.data %>% filter(Group=="Control")
```

5.3 Equal variances assumed t-test (i.e., Levene's non-significant)

Since Levene's was non-significant we use the command below.

```
t.test(x=exp.group.rows$Arousal,
       y=control.group.rows$Arousal,var.equal=TRUE)
```

```
##
##  Two Sample t-test
##
## data:  exp.group.rows$Arousal and control.group.rows$Arousal
## t = 2.4837, df = 62, p-value = 0.01572
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   1.46298 13.53012
## sample estimates:
## mean of x mean of y
##  24.00000  16.50345
```

Then calculate the effect size with confidence interval (when variance are equal):

```
library(MBESS)
d.value <- smd(Group.1=exp.group.rows$Arousal,
               Group.2=control.group.rows$Arousal, Unbiased=TRUE)
print(d.value)
```

```
## [1] 0.616084
```

```
ci.smd(smd=d.value,
       n.1=length(exp.group.rows$Arousal),
       n.2=length(control.group.rows$Arousal))
```

```
## $Lower.Conf.Limit.smd
## [1] 0.1097988
##
## $smd
## [1] 0.616084
##
## $Upper.Conf.Limit.smd
## [1] 1.11763
```

5.4 Equal variances NOT assumed t-test (i.e., Levene's significant)

If Levene's test had been significant we would have used the command below. Note the change in the last argument for var.equal.

```
t.test(x=exp.group.rows$Arousal,
       y=control.group.rows$Arousal,var.equal=FALSE)
```

Then we would calculate the effect size with confidence interval (when variance are **not** equal). When the homogeneity of variance assumption is violated we typically use one group as reference group (i.e., control group / placebo group). In doing so, we use the standard deviation of that group as the denominator for the d-value (the control group in this case). Note that we use smd.c now not smd. Likewise, we use ci.smd.c not ci.smd (both in MBESS library).

```
d.value <- smd.c(Group.T=exp.group.rows$Arousal,
                 Group.C=control.group.rows$Arousal)
print(d.value)


ci.smd.c(smd.c=d.value,n.E=length(exp.group.rows$Arousal),
         n.C=length(control.group.rows$Arousal))
```

5.5 Writing it up

Participants exposed to the drug ($M = 24.00$, $SD = 12.20$) had more Activation/Arousal than participants not exposed to the drug ($M = 16.50$, $SD = 11.8$), $d = 0.62$, 95% CI [0.11, 1.12], $t(62) = 2.48$, $p = .02$. The confidence interval width indicates a plausible population d-value range of 0.11 (a very small effect) to 1.12 (a very large effect) suggesting the results provide strong evidence the direction but not the strength of the effect.