Problem 1

According to the CDC, the average height of US women increased from 62 to 63 inches from 1964 to 2016. Let’s assume that the population of heights are normally distributed with a standard deviation of 3 inches, so we’ll be working with z-distributions.

This first problem is about the power associated with detecting this 1 inch increase in mean height. We’ll be filling in the probabilities in the following table:

In order to determine if women are significantly taller than they were in 1964 we will test the null hypothesis that the mean in 2016 is the same as the mean in 1964, which was 62. We will then compare this number to an observed mean drawn from the 2016 distribution. Use an alpha value of $\alpha = 0.05$.

<table>
<thead>
<tr>
<th>Probabilities ($\alpha = 0.05$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_0$ True</td>
</tr>
<tr>
<td>Fail to reject $H_0$</td>
</tr>
<tr>
<td>Reject $H_0$</td>
</tr>
</tbody>
</table>

**a)** You can already fill in half of the table above. If $H_0$ is true, then the probability of rejecting $H_0$ when it is true is defined to be $\alpha$. Since we will either reject or fail to reject $H_0$, the probability of failing to reject $H_0$ when it’s actually true is $1 - \alpha$. Fill in those numbers in the table above.

**b)** Suppose you were to draw 16 samples from the 1964 population. What is the standard error of the mean?

Next you’ll be calculating power by drawing normal distributions and filling in areas. Note that the tutorial this is done with respect to $z$-distributions. This time you’ll be drawing distributions with respect to mean heights. Below is a bell curve that represents the distribution of means from the null hypothesis (1964) population. It has a mean of 62 and the standard deviation you calculated in problem b:
c) Find the range of heights for the upper tail of this distribution which has an area of $\alpha = 0.05$.

d) Shade this region in the curve above.

e) Suppose you were to draw a mean from the 1964 population that fell in this shaded region. If you were testing the null hypothesis that the population mean was equal to 62, what would our decision be? Would this be a correct decision? If not, what type of error would it be?
f) Let’s assume that $H_0$ is false and that the ‘true’ distribution of heights is that measured in 2016. Now draw a normal distribution on the graph above that represents this ‘true’ distribution of mean heights drawn from the 2016 population. This should be a normal distribution with mean 63 but with the same standard deviation as for $H_0$.

g) Take the rejection region that you found from problem c and lightly shade the area under this ‘true’ distribution (on top of the previously shaded region).

h) If you were to draw a mean of 16 samples from the true (2016) population that landed in the this shaded region, what would you decision be? Would this be a correct decision? If not, what type of error would we be making?

i) What is the area of this newly shaded region?

j) The area you just calculated in part i is the power of our test. It’s the probability of rejecting $H_0$ when $H_0$ is false. That is, it is the probability of obtaining a mean from the 2016 distribution that is significantly greater than the null hypothesis mean of 62 inches.

Now fill in the other two numbers in the table from part a above.
k) Now, redo the math and calculate the power for the same mean but using an α value of 0.01. Did power go up or down? Explain why. Here’s a new table and graph for you to fill out.

<table>
<thead>
<tr>
<th></th>
<th>$H_0$ True</th>
<th>$H_0$ False</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fail to reject $H_0$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reject $H_0$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Probabilities (alpha = 0.01)

![Graph showing mean heights distribution]

Mean heights (in)
**Problem 2** Suppose you think there might be a difference in self confidence between genders. In the survey, I asked what you thought your score would be on Exam 1. Splitting the class by gender, the estimated Exam 1 score for the 122 female students had a mean of 84.72 and a standard deviation of 9.4004. Exam 1 score predictions for the 30 male students had a mean of 86.67 and a standard deviation of 8.1847. Let’s see if these two means are significantly different from each other. We’ll use an $\alpha$ value of 0.05.

a) Calculate the pooled standard deviation:

$$sp = \sqrt{\frac{(n_x-1)s_x^2 + (n_y-1)s_y^2}{n_x-1 + n_y-1}}$$

b) Calculate the pooled standard error of the mean:

$$s_{\bar{x} - \bar{y}} = sp \sqrt{\frac{1}{n_x} + \frac{1}{n_y}}$$

c) Calculate the t-statistic:

$$t = \frac{\bar{x} - \bar{y}}{s_{\bar{x} - \bar{y}}}$$

d) Find the critical value of $t$
e) State your decision as a sentence in APA format.

f) What is the effect size, $d = \frac{\bar{x} - \bar{y}}{s_p}$

g) Suppose this were the true effect size. What would be the power of this test? (Use the power curves for $\alpha = 0.05$, two tailed, two means). Remember, each power curve corresponds to the average sample size for each mean (about 80 for this example).
h) Use the power curve below to determine, for this effect size, how large the sample size for each group would need to be to get a power of 0.8.
i) Draw a bar graph showing the two means with error bars representing ± one standard error of the mean. You’ll have to calculate the standard errors of the mean by dividing each standard deviation by the square root of each sample size.
**Problem 3 a)** Use R to run the same test that you did in problem 2, comparing expected score in Exam 1 across gender. Use $\alpha = 0.05$.

Hint: The expected score in Exam 1 is called 'Exam1'. Here's how to test for a significant difference in preferred temperature across gender:

```r
survey <- read.csv("http://www.courses.washington.edu/psy315/datasets/Psych315W18survey.csv")
x <- survey$temperature[survey$gender == "Male"]
y <- survey$temperature[survey$gender == "Female"]
x <- x[!is.na(x)]
y <- y[!is.na(y)]
out <- t.test(x,y,
alternative = "two.sided",
var.equal = TRUE)
out$p.value
[1] 0.3013393
```

**b)** Use R to calculate the effect size. Here's how (using the values from the temperature example), given that you already have your 'x' and 'y' variables.

```r
nx <- length(x)
yy <- length(y)
sp <- sqrt((nx-1)*sd(x)^2 + (ny-1)*sd(y)^2)/(nx-1+ny-1)
sp
[1] 9.138226
d <- abs(mean(x)-mean(y))/sp
d
[1] 0.2115354
```
c) Calculate the observed power of this test given the effect size that you calculated in part b, the fact that this is a two-tailed test, and the value of $\alpha = 0.05$. Here’s how (with values from the temperature example):

```r
# Find observed power from d, alpha and n
nx <- length(x)
y <- length(y)
out <- power.t.test(n = (nx+ny)/2,
d = d,
sig.level = 0.05,
power = NULL,
alternative = "two.sided",
type = "two.sample")
out$power

[1] 0.251863
```

d) Run these lines in R to generate a bar graph with error bars representing plus or minus one standard error of the mean. This example is for preferred temperature across gender, so be sure to change the labels to ‘Expected Exam 1 Score’. You don’t have to turn in a printout of the plot - I’ll trust you that you tried running the code yourself.

```
require(ggplot2)
Loading required package: ggplot2
# Generate a 'data frame' containing statistics for x and y:
summary <- data.frame(
  mean <- c(mean(x),mean(y)),
  n <- c(length(x),length(y)),
  sd <- c(sd(x),sd(y)))
summary$sem <- summary$sd/sqrt(summary$n)
colnames(summary) = c("mean","n","sd","sem")
row.names(summary) = c("Male","Female")
# This was a bit of work, but it creates a nice table:
summary
  mean n sd sem
Male 71.10000 30 11.50217 2.100000
Female 73.03306 121 8.468504 0.769864
# Once you have this summary table, the rest will give you a nice looking
# bar plot with error bars:
# Define y limits for the bar graph based on means and sem’s
ylim <- c(min(summary$mean-1.5*summary$sem),
  max(summary$mean+1.5*summary$sem))
# Plot bar graph with error bar as one standard error (standard error of the mean/SEM)
ggplot(summary, aes(x = row.names(summary), y = mean)) +
xlab("students") +
geom_bar(position = position_dodge(), stat="identity", fill="blue") +
geom_errorbar(aes(ymin=mean-sem, ymax=mean+sem),width = .5) +
theme_bw() +
theme(panel.grid.major = element_blank()) +
scale_y_continuous(name = "Preferred temperature") +
coord_cartesian(ylim=ylim)
```