STAT 581 - Problem Set 3.b

Alex Towell (atowell@siue.edu)

Problem 3

An experiment to compare a new drug to a standard is in the planning stages. The response variable of interest is the clotting time (in minutes) of blood drawn from a subject. The experimenters want to perform a two sample t test at level $\alpha = .05$, having power $\pi = .8$ at $\delta_A = 0.25$, for standard deviation $\sigma = 0.7$.

Part (a)

Determine the sample size for each drug in order to achieve the stated test specifications.

Selection of sample size is of fundamental importance in experimental design. We know that the CI for the difference in two means has a length given by

$$t_{\alpha/2,n_1+n_2-2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

and if $n_1 = n_2 = n$ then this simplifies to

$$t_{\alpha/2,2n-2} s_p \sqrt{\frac{2}{n}}.$$

We cannot control s_p but we can control n. While s_p will be relatively constant, $\sqrt{n^{-1}}$ will decrease as n increases, so the length of the CI is approximately $\mathcal{O}(\sqrt{n^{-1}})$.

A small CI is desirable, but in particular, in hypothesis testing, if H_A is true we may be concerned about type II errors, whose probability β is not only a function of $\delta_A = \mu_1 - \mu_2$ but also n. Generally, as n increases, β decreases.

The power of a test is given by $\pi = 1 - \beta$. We are interested in obtaining a power of $\pi = 0.8$ given all of the other specifications by increasing sample size n per population (so, the sample size is $2n$).

We may quickly compute an approximation of n with

$$n = 2(z_{\alpha/2} + z_\beta)\sigma^2/\delta_A^2 = 123.0704,$$

but of course we do not know σ. So, instead, we compute n with:

```
sd = .7
alpha = .05
h = .8  # power
v = .25 # alternative, a practical difference
power.t.test(n=NULL,delta=v,sd=sd,sig.level=alpha,power=h,type="two.sample")
```

```
##
##      Two-sample t test power calculation
##
##               n = 124.0381
##           delta = 0.25
##              sd = 0.7
```

```
##       sig.level = 0.05
##           power = 0.8
##     alternative = two.sided
##
## NOTE: n is number in *each* group
```

We see that $n' = 124.0381$. If the power h is a lower-bound, then let $n = \lceil n' \rceil = 125$. If the power specification is not a lower-bound, but an approximate specification, it seems appropriate to set $n = 124$.

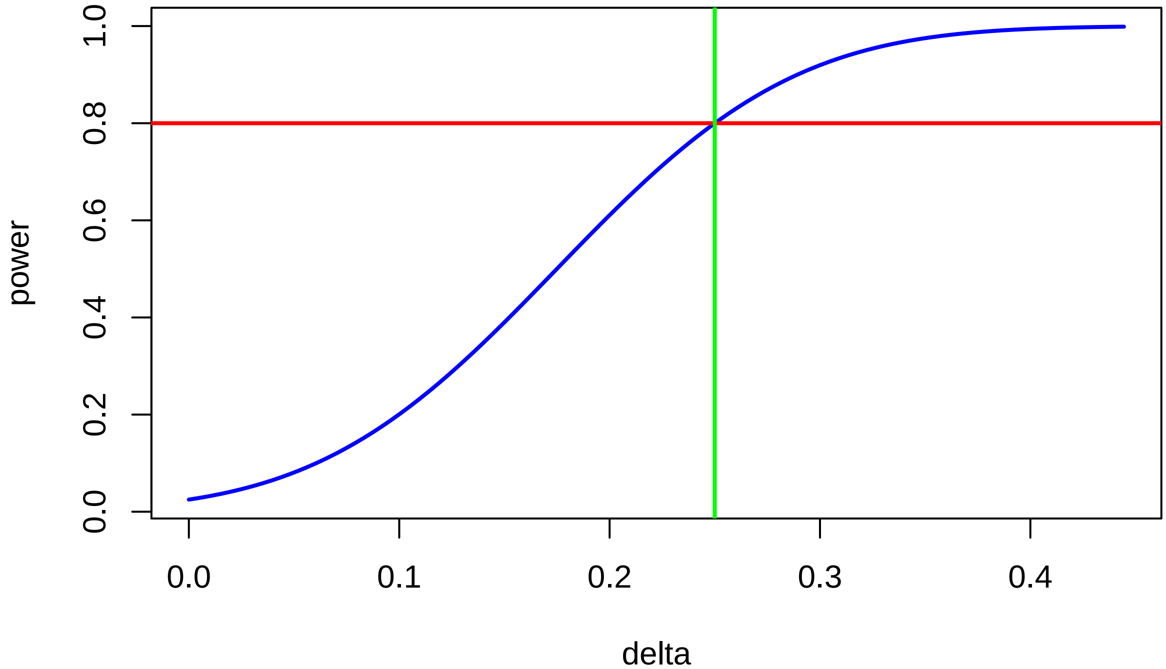
Part (b)

Graph the power curve for the chosen sample size. Explain how the power curve displays the desired properties of the test.

```r
# power.curve
#
# create the power curve for the chosen sample size.
#
# arguments:
#   n: sample size,
#   sd: standard deviation
#   alpha: significance level
#   h: power (shows as a horizontal line)
#   v: specific alternative (shows as a vertical line)
#
# output:
#   graph of power curve
power.curve = function(n, sd, alpha, h, v)
{
  df = 2*(n-1)
  delta = seq(from=0,to=5*sd/sqrt(n/2),length.out = 1000)
  power = 1 - pt(qt(1-alpha/2,df),df,ncp = sqrt(n/2)*(delta/sd))

  plot(delta,power,type = "l",lwd=2,col="blue")
  abline(h=h,col="red",lwd=2)
  abline(v=v,col="green",lwd=2)
}

power.curve(n=124,sd=sd,alpha=alpha,h=h,v=v)
```



First, we see that $\delta_A = 0.25$ obtains a power of $\pi(\delta_A) = 0.8$, as specified.

Second, and this is not explicitly shown on the graph, whether $H_0 : \delta = 0$ is true or $H_A : \delta = \delta_A$ is true, there is a low probability of committing an error since the 95% confidence interval under the null model, approximately

$$\pm 1.96\sigma/\sqrt{n} = [-0.123, 0.123],$$

does not intersect with an approximation of the 95% confidence interval under the alternative model,

$$\delta_A \pm 1.96\sigma/\sqrt{n} = [0.127, 0.373].$$

The small region between these confidence intervals may be classified as the "don't care" region.

Note that we do not know σ so these CIs are approximate (and slightly more narrow in length). However, since n is pretty large, it should be reasonably accurate, i.e., the limiting distribution of the t distribution is the standard normal as $n \to \infty$.

Part (c)

Provide a general explanation of how δ_A can be determined.

The specific alternative δ_A is chosen to represent an effect size that is expected (e.g., from past experience or related data), important (difference is non-negligible), and/or practical (cost considerations).

Problem 4

Refer back to the tensile strength example of problem 2. Use the data from this study to perform a power analysis for a main study. The experimenters desire a level $\alpha = .05$ test with power $\pi = .8$.

Part (a)

Determine the sample size for each group based on specifying the maximum difference in means.

```
library("readxl")
library("multcomp")
data = read_excel("./handout2data.xlsx")
```

```r
strength = na.omit(data$strength)
percent = na.omit(as.factor(data$percent))

pwr = .8
alpha = .05
a = 5
means = by(strength,percent,mean)
max.D = max(means) - min(means)
#summary(aov(strength~percent))
s2 = 8.06
power.anova.test(
  groups=a,
  between.var=max.D^2/2/(a-1),
  within.var=s2,
  power=pwr,
  sig.level=alpha,
  n=NULL)
```

```
##
##      Balanced one-way analysis of variance power calculation
##
##          groups = 5
##               n = 2.533845
##     between.var = 17.405
##      within.var = 8.06
##       sig.level = 0.05
##           power = 0.8
##
## NOTE: n is number in each group
```

We see that $n' = 2.534$. We propose using a sample size of $n = 3$ for each group.

Part (b)

Use a simulation to compute power at $n = 3$ using the pilot study to specify the model parameters.

```r
sim.size = 100000
decide.Ha = rep(NA,sim.size)
n = 3

for (k in 1:sim.size)
{
  y1 = rnorm(n,means[1],sqrt(s2))
  y2 = rnorm(n,means[2],sqrt(s2))
  y3 = rnorm(n,means[3],sqrt(s2))
  y4 = rnorm(n,means[4],sqrt(s2))
  y5 = rnorm(n,means[5],sqrt(s2))
  ybar1 = mean(y1)
  ybar2 = mean(y2)
  ybar3 = mean(y3)
  ybar4 = mean(y4)
  ybar5 = mean(y5)
  var1 = var(y1)
  var2 = var(y2)
```

```
  var3 = var(y3)
  var4 = var(y4)
  var5 = var(y5)
  F.stat = n*var(c(ybar1,ybar2,ybar3,ybar4,ybar5)) / mean(c(var1,var2,var3,var4,var5))
  decide.Ha[k] = (F.stat > qf(1-alpha,a-1,a*(n-1)))
}
power = mean(decide.Ha)
power
```

```
## [1] 0.97583
```

Part (c)

Comment on the use of pilot study data in a power analysis.

Specifying parameter values for a power analysis based on estimates from a pilot study, without accounting for estimation error, may lead to a hypothesis test that does not have adequate power.