Test for the efficacy for hormone therapy in prostate cancer.

Shane Crinion

27 September 2018

The male androgen hormones testosterone and dihydrotestosterone contribute to the progression in prostate cancer, and the use of androgen deprivation therapy (ADT) is considered a standard treatment protocol for metastatic disease. Treatment using LHRH agonists typically reduces the blood serum concentration of testosterone from an average of 540 ng/dl to what is considered castration levels of ~ 50 ng/dl. Working with colleagues you wish to design an experiment that assesses the efficacy of a novel LHRH agonist, with two groups of otherwise healthy male patients - one group will get the LHRH agonist, the other a placebo. Prior to the study, the mean/standard deviation of blood serum for all participants is 540/180. Assuming the drug works, the treated group would be expected to have a testosterone level of at most 50 ng/dl.

We can use the R library 'pwr' to implement a range of power calculations - for example, to determine a power calculation using a t-distribution we would use the command:

```
pwr.t.test(n = NULL, d = NULL, sig.level = 0.05, power = NULL,
    type = c("two.sample", "one.sample", "paired"), alternative = c("two.sided",
        "less", "greater"))
```

- **n** is the number of observations that are either specified or tested to find the number required to meet the statistical significance defined by other parameters.
- **d** his refers to the effect size. The effect size is a measurement of the trend between two variables
- **sig.level** is the p value. It is the significance threshold and accounts for the level of results that can result purely due to chance.
- **power** is the level ability of the statistical test to find the desired results. Higher statistical power decreases type II errors and is particularly important in genome analysis due to the large data sizes.
- **type** indicates whether the test is one tailed or two tailed. The one tailed test will test the possibility of significant relationship in one direction and will complete disregard the other direction. The two tailed test will test for a relationship in both directions ie. significantly higher or lower than x.
- **alternative** is the alternative hypothesis to the null which is used to null the hypothesis. The "two sided" indicates that the value be either side of the mean. The "less" indicates less than mean and vice versa for "greater"

Load the 'pwr' library:

```
library(pwr)
```

(i) Perform a power calculation to determine the optimum number of volunteers to include in the study, assuming the study is powered to 80% and efficacy is assessed at the 5% level of significance, and that the same standard deviation is associated with the castration-effective response.

```
# BLOOD SERUM CONC LEVELS
avg_testos_ng <- 540
```

```r
castration_ng <- 50
sd_testos_ng <- 180

# CALCULATE EFFECT SIZE
d = (avg_testos_ng - castration_ng)/sd_testos_ng

# PERFORM T-TEST TO DETERMINE OPTIMUM NUMBER OF VOLUNTEERS
test1 <- pwr.t.test(n = NULL, sig.level = 0.05, d = d, power = 0.8)

test1
```

```
##
##      Two-sample t test power calculation
##
##               n = 3.403697
##               d = 2.722222
##       sig.level = 0.05
##           power = 0.8
##     alternative = two.sided
##
## NOTE: n is number in *each* group
```

The two sample t-test is used to compare the average difference between the two groups. The number of samples required for accurate results is dependant on the level of differences between the two samples. The results indicate that N = 3.4 samples are required to obtain accurate results in the study.

(ii) Actual studies indicate the standard deviation of the measured testosterone blood concentration for castration-like conditions is typically 10% of the measured value - so for a 50 ng/dl cutoff, this would be 5 ng/dl. Given that the differences in standard deviations require a modified calculation of the effect size, re-calculate the optimum number of volunteers for the same power and significance level.

```r
# NEW CALCULATION FOR EFFECT SIZE, d norm conditions
sd_testos_ng_norm_sqrd <- sd_testos_ng * sd_testos_ng

# castration conditions
sd_testos_ng_cast <- sd_testos_ng * 0.05
sd_testos_ng_cast_sqrd <- sd_testos_ng_cast * sd_testos_ng_cast


# NEW CALCULATION FOR SD
sd_testos_pooled_sqrt <- sqrt((sd_testos_ng_norm_sqrd + sd_testos_ng_cast_sqrd)/2)


d_corrected <- (avg_testos_ng - castration_ng)/sd_testos_pooled_sqrt

# NEW POWER TEST FOR CORRECTED n
test2 <- pwr.t.test(n = , sig.level = 0.05, d = d_corrected,
    power = 0.8)
test2
```

```
##
##      Two-sample t test power calculation
##
##               n = 2.481832
```

```
##               d = 3.845
##       sig.level = 0.05
##           power = 0.8
##     alternative = two.sided
##
## NOTE: n is number in *each* group
```

The re-calculated number of samples (n = 2.48) is reduced by the larger difference between the healthy and non-healthy samples, making accurate variation measurable with smaller values.

(iii) Finally, we might imagine that we would want to be as stringent as possible in avoiding false negatives - if we enrolled 10 volunteers for each group, and tested the hypothesis that the drug had no effect at the 5% significance level, how powered would our experiment be? What about at the 0.1% level?

```
# PERFORM T-TEST WITH CORRECTED VARIABLES FOR n AND sig.level
test3 <- pwr.t.test(n = 10, sig.level = c(0.001, 0.05), d = d_corrected,
    power = )
test3
```

```
##
##      Two-sample t test power calculation
##
##               n = 10
##               d = 3.845
##       sig.level = 0.001, 0.050
##           power = 0.9999527, 1.0000000
##     alternative = two.sided
##
## NOTE: n is number in *each* group
```

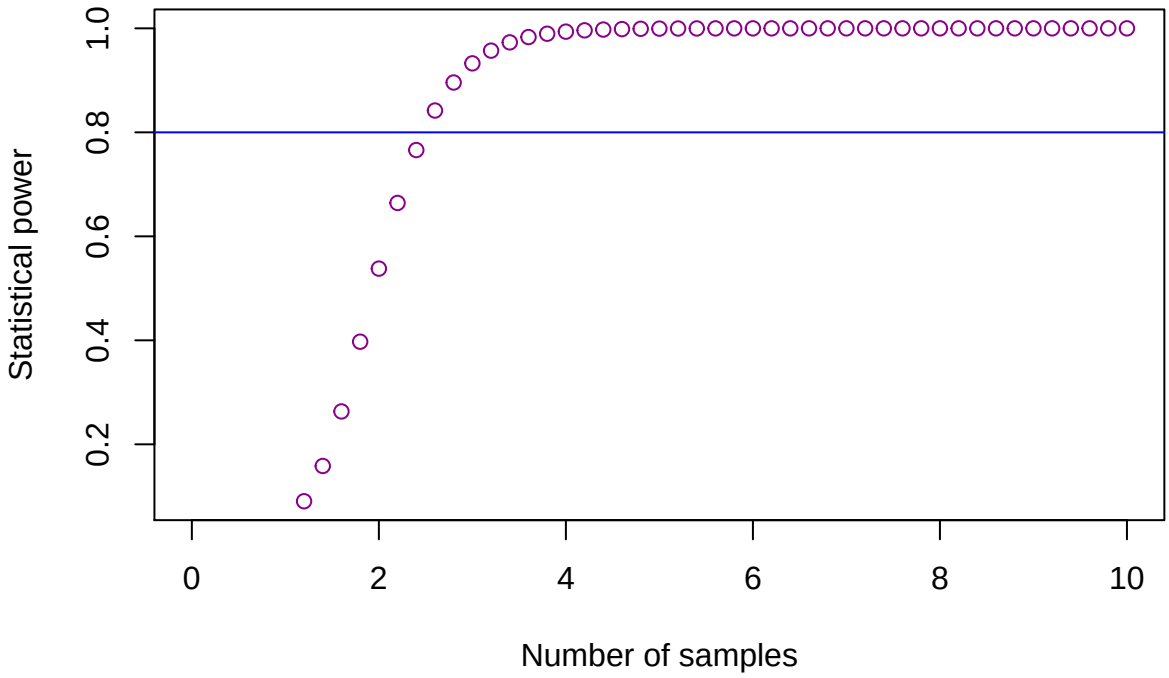
Illustrate of plots to demonstrate how power varies as a function of numbers of patients enrolled.

```
# NEW T-TEST FOR POWER USING A SEQUENCE OF NUMBERS
test4 <- pwr.t.test(n = seq(0, 10, 0.2), sig.level = 0.05, d = d_corrected,
    power = )
```

```
## Warning in qt(sig.level/tside, nu, lower = FALSE): NaNs produced
```

```
# PLOT OF SEQUENCE OF VALUES
plot(x = test4$n, y = test4$power, xlab = "Number of samples",
    ylab = "Statistical power", col = "darkmagenta")

abline(h = 0.8, col = "blue")
```



With 10 samples for 0.1% and 5% significance level, the test would have power of 0.9999527 and 1 respectively. 10 samples would be highly accurate with little or no expected false negatives.

Most clinical trials for LHRH agonists would use 50-60 patients patients per sample group which is a considerably high number which would provide almost complete accuracy given the same statistical power. Using less patients but testing other molecular mechanisms or drugs may be more valuable. Alternatively, using a diverse group will provide valuable population response statistics.