Consulting Homework 10

Tim Vigers

27 November, 2018

1. Sample size calculations.

a.

i. Assuming a known SD

Table 1: Total absorbed zinc (TAZ) by dose

ID	2mg	5mg	10mg	15mg	20mg	30mg
1	1.81	3.55	8.28	10.33	6.7	13.28
2	1.77	5	8.33	11.28	9.77	12.14
3	2.09	NA	7.57	10.57	13.38	13.42
4	1.09	NA	7.99	12.4	10.95	9.83
5	1.66	3.81	7.94	10.46	16.32	11.54
6	1.91	4.33	5.4	7.18	4.74	9.66
7	1.31	1.28	6.71	6.25	17.13	7.52
8	1.28	3.13	7.04	7.54	8.68	12.36
SD	0.35002040756831	1.27288124609748	0.993777066118381	2.20778775830339	4.40949442356248	2.04896097362

Using the SD for the 2 mg group, because both the low and high zinc content groups are expected to absorb < 2 mg per day (90% power, alpha $= 0.05$):

$$n = \frac{\sigma^2(Z_{0.9} + Z_{0.975})^2}{(\text{detectable difference})^2} = \frac{0.3500204^2 * (1.28 + 1.96)^2}{(0.6)^2} = \frac{1.286106}{0.36} = 3.572516$$

So for 90% power to detect a difference in TAZ of 0.6 mg/day, assuming known variation, Jamie will need 4 participants in each group (since the planned analysis requires 4 difference measurements).

Check using an R function from methods:

```
findPowerZ <- function(diff = 5, sd = 1, n = 10, alpha = 0.05){
  z.alpha <- qnorm(1 - (alpha/2))
  power <- pnorm(diff/(sd/sqrt(n)) - z.alpha)
  return(power)
}
findPowerZ(diff = 0.6,sd = 0.3500204,n = 4)
```

```
## [1] 0.9290032
```

```
findPowerZ(diff = 0.6,sd = 0.3500204,n = 3)
```

```
## [1] 0.843535
```

ii. Assuming an unknown SD

Use R's built-in power calculation function:

```
power.t.test(delta = 0.6,sd = s,power = 0.9,type = "two.sample")
```

```
##
##      Two-sample t test power calculation
##
##               n = 8.242292
##           delta = 0.6
##              sd = 0.3500204
##       sig.level = 0.05
##           power = 0.9
##     alternative = two.sided
##
## NOTE: n is number in *each* group
```

With an unknown variance, R recommends a sample size of 9 participants per groups, which seems like a much better estimate to me (4 feels awfully low).