

Sample size calculation for TMJ study

Xuelong Wang

2019-04-12

1 Background and Goal

1.1 TMJ

TMJ, temporomandibular joints, are the joints and jaw muscles that make it possible to open and close your mouth.

1.2 Goal

As we discussed in the meeting, not all dentist clinics accept the TMJ patients. So the primary goal, for my understanding, is to figure out for what reasons the TMJ accepting rate is very low. For the very beginning of the study, what we are interested in, again for my understanding, is to see if there is any difference of the accepting rate among k different geographic groups. Moreover, we also want also to calculate the sample size for that test. Please note, if we want to answer a more complicated question about the the TMJ accepting rate, we may need to conduct a designed survey.

2 Sample size calculation

2.1 Hypothesis test and sample size calculation for comparison 2 propotions

The question this Hypothesis is trying to answer is if the two proportions are same or different. In the Statistics, we say that the null hypothesis is they same, and alternative hypothesis is that they are different, that is

$$H_0 : p_1 = p_2 \quad H_A : p_1 \neq p_2.$$

Based on this test, we could derive the formula for sample size calculation, which is

$$n = \left(\frac{\Phi_{\alpha/2} \sqrt{2p(1-p)} + \Phi_{\beta} \sqrt{p_1(1-p_1) + p_2(1-p_2)}}{|\delta|} \right)^2,$$

Where Φ is the distribution function of standardized normal, α is the type-I error and β is the type-II error, $\delta = p_1 - p_2$ is difference between p_1 and p_2 . Note that we assume those two groups have the same sample size n . For all of those parameters mentioned above, p, p_1, p_2 are calculated from sample and α, β, δ need to be decided based on experience. For more details, see Agresti (2007), ch. 5.5; also Fleiss (2003), p69.

2.2 Hypothesis test for comparison K proportions

For K proportions comparison, we could just extend the comparison of 2 proportions. The basic idea is to perform $K(K-1)/2$ tests for each combination of p_i and p_j , and if there is one test rejects the null hypothesis, then we claim there is some difference among the K proportions. To simplify the test, we can just look the two groups which is the $p_{min} = \min(p_1, \dots, p_k)$ and $p_{max} = \max(p_1, \dots, p_k)$. Then the corresponding hypothesis is

$$H_0 : p_1 = \dots = p_k \iff p_{min} = p_{max} \quad H_A : p_{max} \neq p_{min}.$$

The sample size is

$$n = \left(\frac{\Phi_{\alpha/2\tau} \sqrt{2p(1-p)} + \Phi_{\beta} \sqrt{p_{max}(1-p_{max}) + p_{min}(1-p_{min})}}{|\delta|} \right)^2,$$

where $\tau = k(k-1)/2$, which is because of we have K different groups need to compare with. More details of multiple comparison could be found at (“Calculate Sample Size Needed to Compare K Proportions” 2013).

2.3 Sample size

As we mentioned above, to calculate the sample size we need to pre-specify the following parameters: $\alpha, \beta, k, p_{max}, p_{min}$. There is a built-in R function `power.prop.test` for calculate the sample size and power analysis. For instance, if we set $\alpha = 0.05, \beta = 0.2, k = 6, p_{max} = 0.6, p_{min} = 0.1$. Please note since we are doing a multiple test, we need to adjust for the type-I error, which is $\alpha_{adj} = 0.05/6(6-1)/2 = 0.0033$. Then we can plug those parameters in to the `power.prop.test` get following results:

```
k = 6
alpha = 0.05
beta = 0.2
p_max = 0.3
p_min = 0.1
power.prop.test(p1 = p_max,
                p2 = p_min,
                sig.level = alpha/(k*(k-1)/2),
                power = 1 - beta)
```

Two-sample comparison of proportions power calculation

```
n = 112.5057
p1 = 0.3
p2 = 0.1
sig.level = 0.003333333
power = 0.8
alternative = two.sided
```

NOTE: n is number in *each* group

Based on result, if the sample size is larger than 25, then we could test the effect size which is $p_{max} - p_{min} = 0.3 - 0.1 = 0.2$ with significant level 0.0033 and power 0.8. Note that if we have less number of groups, we will need less sample size for test a same effect size, as following result shows.

```
k = 3
alpha = 0.05
beta = 0.2
p_max = 0.3
p_min = 0.1
power.prop.test(p1 = p_max,
                p2 = p_min,
                sig.level = alpha/(k*(k-1)/2),
                power = 1 - beta)
```

Two-sample comparison of proportions power calculation

```
n = 82.37508
p1 = 0.3
p2 = 0.1
sig.level = 0.01666667
power = 0.8
alternative = two.sided
```

NOTE: n is number in *each* group

Note: previous experience, research results and more detail discussion may be needed in order to choose those parameters appropriately.

2.4 Power analysis table

I also compute a power table for a series of different conditions as following:

Table 1: Powers for different conditions

n/delta	0.1	0.15	0.2
80	0.068	0.196	0.409
100	0.091	0.266	0.532
120	0.117	0.34	0.642
150	0.159	0.45	0.772
200	0.237	0.616	0.904

- $\alpha = 0.05$ and $k = 5$,
- assume $p_{min} = 0.3$ and $\delta = \{0.1, 0.15, 0.2\}$, so the corresponding $p_{max} = \{0.4, 0.45, 0.5\}$
- $n = 80, 100, 120, 150, 200$.

3 Bias Issue

Based on the survey result, I find that there may be a bias problem for the TMJ study. According to the second question, which is asking if the person accepts the referral of patients with “TMJ” or not. The sample size is $n = 155$ and the accept rate is $\hat{p} = 0.79$. we can also construct an 95% confidence inter for it, which is $\hat{p} \pm 0.0641 = (0.73, 0.85)$. I don’t have much experience in this filed, but based on the sample, the accept rate seems very high to me. One possible reason is that the clinic which accepts TMJ patients is tending to reply to the survey. Therefore, research results based on this survey may be affected the bias issue.

Reference

- Agresti, Alan. 2007. *An Introduction to Categorical Data Analysis*. John Wiley & Sons.
- “Calculate Sample Size Needed to Compare K Proportions.” 2013. <http://powerandsamplesize.com/Calculators/Compare-k-Proportions/1-Way-ANOVA-Pairwise>.
- Fleiss, Joseph L. 2003. *Statistical Methods for Rates and Proportions, 3rd Edition*. John Wiley & Sons.