STA305/1004-Class 16

March 13, 2017

Today's Class

- My office hours: 1-2 instead of 12-1.
- Deadline for HW#3 extended to March 20, 22:00

- Sample size for ANOVA
- Factorial designs at two levels
- Cube plots
- Calculation of factorial effects

Sample size for ANOVA - Designing a study to compare more than two treatments

- Consider the hypothesis that k means are equal vs. the alternative that at least two differ.
- What is the probability that the test rejects if at least two means differ?
- Power $= 1 - P(\text{Type II error})$ is this probability.

Sample size for ANOVA - Designing a study to compare more than two treatments

The null and alternative hypotheses are:

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k \text{ vs. } H_1 : \mu_i \neq \mu_j.$$

The test rejects at level α if

$$MS_{Treat}/MS_E \geq F_{k-1,N-K,\alpha}.$$

The power of the test is

$$1 - \beta = P\left(MS_{Treat}/MS_E \geq F_{k-1,N-K,\alpha}\right),$$

when H_0 is false.

Sample size for ANOVA - Designing a study to compare more than two treatments

When H_0 is false it can be shown that:

- MS_{Treat}/σ^2 has a non-central Chi-square distribution with $k-1$ degrees of freedom and non-centrality parameter δ.
- MS_{Treat}/MS_E has a non-central F distribution with the numerator and denominator degrees of freedom $k-1$ and $N-k$ respectively, and non-centrality parameter

$$\delta = \frac{\sum_{i=1}^{k} n_i \left(\mu_i - \bar{\mu}\right)^2}{\sigma^2},$$

where n_i is the number of observations in group i, $\bar{\mu} = \sum_{i=1}^{k} \mu_i / k$, and σ^2 is the within group error variance .

This is dentoted by $F_{k-1,N-k}(\delta)$.

Non-Centrality parameter

numerator df

denominator df

Direct calculation of Power

- The power of the test is

$$P\left(F_{k-1,N-k}(\delta) > F_{k-1,N-K,\alpha}\right).$$

- The power is an increasing function δ
- The power depends on the true values of the treatment means μ_i, the error variance σ^2, and sample size n_i.
- If the experimentor has some prior idea about the treament means and error variance the sample size (number of replications) that will guaruntee a pre-assigned power of the test.

Within group variance.

Blood coagulation example - sample size

Suppose that an investigator would like to replcate the blood coagulation study with only 3 animals per diet. In this case $k = 4$, $n_i = 3$. The treatment means from the initial study are:

Diet	A	B	C	D
Average	61	66	68	61

```r
anova(lm.diets)
```

this information can be used to calculate the power for a new study.

```
## Analysis of Variance Table
##
## Response: y
##           Df Sum Sq Mean Sq F value    Pr(>F)
## diets      3    228    76.0  13.571 4.658e-05 ***
## Residuals 20    112     5.6
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Blood coagulation example - sample size

- $\mu_1 = 61$, $\mu_2 = 66$, $\mu_3 = 68$, $\mu_4 = 61$.
- The numerator df = 4-1=3, and the denominator df = 12-4=8.
- The error variance σ^2 was estimated as $MS_E = 5.6$.
- Assuming that the estimated values are the true values of the parameters, the non-centrality parameter of the F distribution is

$$\delta = 3 \times \left((61 - 64)^2 + (66 - 64)^2 + (68 - 64)^2 + (61 - 64)^2 \right) / 5.6 = 20.35714$$

Blood coagulation example - sample size

If we choose $\alpha = 0.05$ as the significance level then $F_{3,20,0.05} = 4.0661806$. The power of the test is then

$$P\left(F_{3,8}(20.35714) > 4.066181\right) = 0.8499.$$

This was calculated using the CDF for the F distribution in R `pf()`.

```
1-pf(q = 4.066181,df1 = 3,df2 = 8,ncp = 20.35714)
```

```
## [1] 0.8499
```

Calculating power and sample size using the pwr library

- There are several libraries in R which can calculate power and sample size for statistical tests. The library `pwr()` has a function for ANOVA.
- `pwr.anova.test(k = NULL, n = NULL, f = NULL, sig.level = 0.05, power = NULL)`

for computing power and sample size.

- `k` Number of groups
- `n` Number of observations (per group)
- `f` Effect size

The effect size f

$$f = \sqrt{\frac{\sum_{i=1}^{k} \left(\mu_i - \bar{\mu}\right)^2 / k}{\sigma^2}},$$

is related to the non-centrality parameter δ via $\boxed{\delta = k \cdot n_i \cdot f^2.}$

- n_i is the number of observations in group i, $\bar{\mu} = \sum_{i=1}^{k} \mu_i / k$, and σ^2 is the within group error variance.

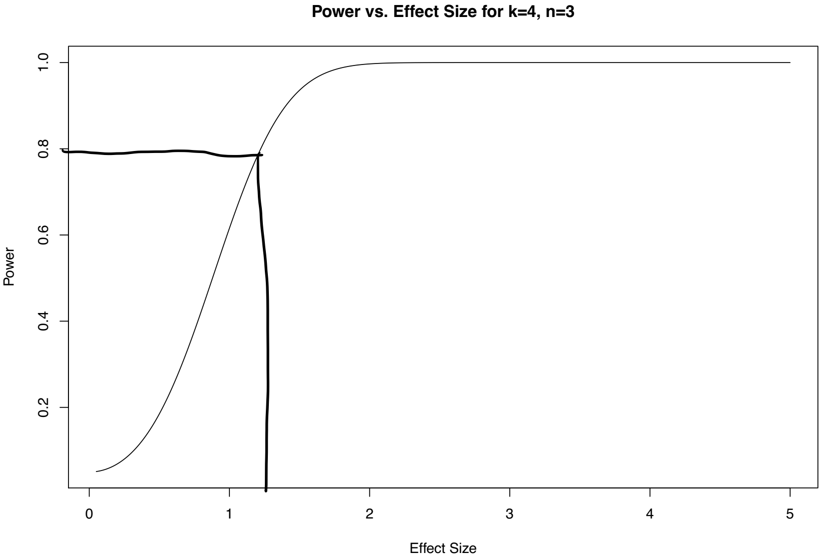
Calculating power and sample size using the pwr library

In the previous example $\delta = 20.35714$ so $f = \sqrt{\frac{\delta}{k \cdot n_i}} = \sqrt{20.35714/4 \cdot 3} = 1.3024701$.

```r
library(pwr)
pwr.anova.test(k = 4,n = 3,f = 1.30247)
```

```
##
##      Balanced one-way analysis of variance power calculation
##
##               k = 4
##               n = 3
##               f = 1.30247
##       sig.level = 0.05
##           power = 0.8499
##
## NOTE: n is number in each group
```

Calculating power and sample size using the pwr library



Calculating power using simulation

The general procedure for simulating power is:

1. Use the underlying model to generate random data with (a) specified sample sizes, (b) parameter values that one is trying to detect with the hypothesis test, and (c) nuisance parameters such as variances.
2. Run the estimation program (e.g., `t.test()`,`lm()`) on these randomly generated data.
3. Calculate the test statistic and p-value.
4. Do Steps 1–3 many times, say, N, and save the p-values. The estimated power for a level alpha test is the proportion of observations (out of N) for which the p-value is less than alpha.

Calculating power using simulation

One of the advantages of calculating power via simulation is that we can investigate what happens to power if, say, some of the assumptions behind one-way ANOVA are violated.

Calculating power using simulation - R program

```
#Simulate power of ANOVA for three groups

NSIM <- 1000 # number of simulations
res <- numeric(NSIM) # store p-values in res
```

μ_1 μ_2 μ_3

```
mu1 <- 2; mu2 <- 2.5;mu3 <- 2  # true mean values of treatment groups
```

$\sigma_\tau =$

```
sigma1 <- 1; sigma2 <- 1; sigma3 <- 1 #variances in each group
n1 <- 40; n2 <- 40; n3 <- 40 #sample size in each group

for (i in 1:NSIM) # do the calculations below N times
  {
# generate sample of size n1 from N(mu1,sigma1^2)
y1 <- rnorm(n = n1,mean = mu1,sd = sigma1)
# generate sample of size n2 from N(mu2,sigma2^2)
y2 <- rnorm(n = n2,mean = mu2,sd = sigma2)
# generate sample of size n3 from N(mu3,sigma3^2)
y3 <- rnorm(n = n3,mean = mu3,sd = sigma3)
y <- c(y1,y2,y3) # store all the values from the groups
# generate the treatment assignment for each group
trt <- as.factor(c(rep(1,n1),rep(2,n2),rep(3,n3)))
m <- lm(y~trt) # calculate the ANOVA
res[i] <- anova(m)[1,5] # p-value of F test
}
sum(res<=0.05)/NSIM # calculate p-value
```

```
## [1] 0.615
```

Calculating power using simulation - R program

- Assume non-equal variances.

μ_i

σ_i

```r
#Simulate power of ANOVA for three groups
NSIM <- 1000 # number of simulations
res <- numeric(NSIM) # store p-values in res
mu1 <- 2; mu2 <- 2.5;mu3 <- 2  # true mean values of treatment groups
sigma1 <- 2; sigma2 <- 1; sigma3 <- 1 #variances in each group
n1 <- 40; n2 <- 40; n3 <- 40 #sample size in each group

for (i in 1:NSIM) # do the calculations below N times
  {
# generate sample of size n1 from N(mu1,sigma1^2)
y1 <- rnorm(n = n1,mean = mu1,sd = sigma1)
# generate sample of size n2 from N(mu2,sigma2^2)
y2 <- rnorm(n = n2,mean = mu2,sd = sigma2)
# generate sample of size n3 from N(mu3,sigma3^2)
y3 <- rnorm(n = n3,mean = mu3,sd = sigma3)
y <- c(y1,y2,y3) # store all the values from the groups
# generate the treatment assignment for each group
trt <- as.factor(c(rep(1,n1),rep(2,n2),rep(3,n3)))
m <- lm(y~trt) # calculate the ANOVA
res[i] <- anova(m)[1,5] # p-value of F test
}
sum(res<=0.05)/NSIM # calculate p-value
```

```
## [1] 0.338
```

Example of a factorial design

Suppose that an investigator is interested in examining three components of a weight loss intervention. The three components are:

1. Keeping a food diary (yes/no)
2. Increasing activity (yes/no)
3. Home visit (yes/no)

Factorial designs

3 factors – food diary
– physical activity
– Home visit.

this is an example of a 2^3 factorial design.

The investigator plans to investigate all $2x2x2 = 2^3 = 8$ combinations of experimental conditions.

The experimental conditions will be.

Expt condition	Keep food diary	Increase physical activity	Home visit	weight loss
1	No	No	No	y_1
2	No	No	Yes	y_2
3	No	Yes	No	y_3
4	No	Yes	Yes	y_4
5	Yes	No	No	y_5
6	Yes	No	Yes	y_6
7	Yes	Yes	No	y_7
8	Yes	Yes	Yes	y_8

Each factor has 2 levels.

Factorial designs at two levels

- To perform a factorial design, you select a fixed number of levels of each of a number of factors (variables) and then run experiments in all possible combinations.

Factorial designs at two levels

- The factors can be quantitative or qualitative.
- Two levels of a quantitative variable could be two different temperatures or two different concentrations.
- Qualitative factors might be two types of catalysts or the presence and absence of some entity.

Factorial design

The notation 2^3 identifies: - the number of factors (3) - the number of levels of each factor (2) - how many experimental conditions are in the design ($2^3 = 8$)

Factorial experiments can involve factors with different numbers of levels.

Factorial design

Consider a $4^2 x 3^2 x 2^1$ design.

1. How many factors? $5 = 2+2+1$
2. How many levels of each factor? $\longrightarrow$ 2 factors with 4 levels
   2 factors with 3 levels
   1 factor with 2 levels
3. How many experimental conditions (runs)?

$$\parallel$$

$$16 \times 9 \times 2 = 288$$

Difference between ANOVA and Factorial Designs

In ANOVA the objective is to compare the individual experimental conditions with each other. In a factorial experiment the objective is generally to compare combinations of experimental conditions.

Let's consider the food diary study above. What is the effect of keeping a food diary?

Expt condition	Keep food diary	Increase physical activity	Home visit	weight loss
1	No	No	No	y_1
2	No	No	Yes	y_2
3	No	Yes	No	y_3
4	No	Yes	Yes	y_4
5	Yes	No	No	y_5
6	Yes	No	Yes	y_6
7	Yes	Yes	No	y_7
8	Yes	Yes	Yes	y_8

$$\frac{y_1+y_2+y_3+y_4}{4}$$

$$\frac{y_5+y_6+y_7+y_8}{4}$$

We can estimate the effect of food diary by comparing the mean of all conditions where food diary is set to NO (conditions 1-4) and mean of all conditions where food diary set to YES (conditions 5-8). This is also called the **main effect** of food diary, the adjective *main* being a reminder that this average is taken over the levels of the other factors.

Difference between ANOVA and Factorial Designs

Expt condition	Keep food diary	Increase physical activity	Home visit	weight loss
1	No	No	No	y_1
2	No	No	Yes	y_2
3	No	Yes	No	y_3
4	No	Yes	Yes	y_4
5	Yes	No	No	y_5
6	Yes	No	Yes	y_6
7	Yes	Yes	No	y_7
8	Yes	Yes	Yes	y_8

The main effect of food diary is:

$$\frac{y_1 + y_2 + y_3 + y_4}{4} - \frac{y_5 + y_6 + y_7 + y_8}{4}.$$

The main effect of physical activity is:

Physical Activity=No

$$= \left\{ \frac{y_1 + y_2 + y_5 + y_6}{4} \right\} - \left\{ \frac{y_3 + y_4 + y_7 + y_8}{4}. \right\}$$

Physical Activity = Yes.

The main effect of home visit is:

$$\frac{y_1 + y_3 + y_5 + y_7}{4} - \frac{y_2 + y_4 + y_6 + y_8}{4}.$$