STA 674

Regression Analysis And Design Of Experiments

Advanced Concepts – Lecture 4

Power Analysis and Sample Size

STA 674, RADOE:
Power Analysis

- Last time, talked about avoiding a false replication—pseudoreplication—when we take repeated measurements from a single experimental unit.
- This time, we'll consider one of the most important themes in any experiment—money—in the form of "how many replications do I need to achieve a desired level of significance?" or, the converse, "with ___ replicates, how likely am I to detect a significant difference?"

STA 674, RADOE:
Power Analysis

Sample size

- There is one very important question we haven't asked so far:

How many replicates do I need?

- What does this mean?

How many replicates do I need to attain a specified power?

STA 674, RADOE:
Power Analysis

Power

- The power of a test is the probability that the null hypothesis is rejected ***given specific values of all the unknown quantities.***
- Example: One-Sample T-test

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

- Here, **power** is the probability of rejecting H_0 for set values of μ, μ_0, σ, n, and α.

....correctly rejecting...

and you can only do that by adjusting n

STA 674, RADOE:
Power Analysis

Example – blood pressure

- A single person's blood pressure is not constant but fluctuates with a standard deviation of about 5 mm of hg. The mean SBP for a normal individual is about 120 mm of hg.
- Suppose that an individual with mild cardiovascular diseased (CVD) has a mean systolic blood pressure (SBP) of $\mu = 130$ mm of hg. A doctor will take 2 measurements from this individual and test the hypothesis that the patient has elevated blood pressure:

$$H_0: \mu \leq 120 \text{ versus } H_0: \ \mu > 120$$

- What is the probability of rejecting H_0 at the $\alpha = 0.05$ level of significance?

STA 674, RADOE:
Power Analysis

Example – blood pressure

- What is the probability that $t > t_{\alpha,n-1}$ if:

$$t = \frac{\bar{x}-\mu_0}{s/\sqrt{n}} \text{ and } \bar{X} \sim Normal\left(\mu, \sigma^2/n\right)$$

- with $\mu = 130$, $\mu_0 = 120$, $\sigma = 5$, $n = 2$, and $\alpha = 0.05$?

SAS Code

```
PROC POWER;
ONESAMPLEMEANS sides=1 nullmean=120 mean=130
stddev=5 ntotal=2 alpha=.05 power=.;
RUN;
```

STA 674, RADOE:
Power Analysis

Example – blood pressure

Interpretation?

If the doctor were to repeat this procedure many, many times taking 2 samples from the same patient, then the doctor would reject the hypothesis that the patient has normal blood pressure 34% of the time.

The POWER Procedure
One-Sample t Test for Mean

Fixed Scenario Elements	
Distribution	Normal
Method	Exact
Number of Sides	1
Null Mean	120
Alpha	0.05
Mean	130
Standard Deviation	5
Total Sample Size	2

Computed Power
Power
0.342

STA 674, RADOE:
Power Analysis

Example – blood pressure

- What if we have a desired level of power?

SAS Code and Output

```
PROC POWER;
ONESAMPLEMEANS sides=1 nullmean=120 mean=130
stddev=5 ntotal=. alpha=.05 power=.80;
RUN;
```

The POWER Procedure
One-Sample t Test for Mean

Fixed Scenario Elements	
Distribution	Normal
Method	Exact
Number of Sides	1
Null Mean	120
Alpha	0.05
Mean	130
Standard Deviation	5
Nominal Power	0.8

Computed N Total	
Actual Power	N Total
0.908	4

STA 674, RADOE:
Power Analysis

Power Analysis for ANOVA

- For the analysis of variance we will consider the power of the overall F-test.

Hypotheses:

- H_0: all treatment means are equal
- H_a: at least one pair of means differs

Power:

- The power is the probability that we reject H_0 given:
- values for the treatment means (e.g., μ_1, μ_2, μ_3),
- the standard deviation of the errors (σ), and
- the number of replicates.

STA 674, RADOE:
Power Analysis

How do we get the values we need?

Option 1: Prospective Power Analysis

In a prospective power analysis estimates of the treatment means and σ are obtained from the analysis of previously collected data (pilot studies) or expert opinion.

Option 2: Retrospective Power Analysis

In a retrospective power analysis estimates of the treatment means and σ are obtained from the current data.

STA 674, RADOE:
Power Analysis

Example – lettuce

A horticulturalist is interested in studying the effects of nitrogen on the production of lettuce. She plans to plant lettuce in 20 plots and randomly treats the 4 plots with each of 5 different levels of ammonium nitrate: 0, 50, 100, 150, and 200 lb/acre.

She projects that the mean number of heads of lettuce produced at each level of ammonium nitrate (AN) will be:

Level of AN	0	50	100	150	200
Heads of lettuce	100	115	130	145	160

and believes that the standard deviation of the number of heads from plots with the same treatment will be about 25. What is the expected power of this experiment?

STA 674, RADOE:

Power Analysis

Interpretation

If the true treatment means are as given in the table above and the true standard deviation of the residuals was 25, the probability that the null hypothesis is rejected would be .743.

If she repeated the experiment many times under these conditions then the hypothesis that all treatments produce the same number of lettuces on average would be rejected 74.3% of the time.

```
DATA MEANS1;
INPUT NITROGEN LETTUCE;
CARDS;
0 100
50 115
100 130
150 145
200 160
RUN;
PROC GLMPOWER DATA=MEANS1;
CLASS nitrogen;
MODEL lettuce=nitrogen;
POWER stddev=25 ntotal=20 power=.;
RUN;
```

The GLMPOWER Procedure

Fixed Scenario Elements	
Dependent Variable	LETTUCE
Source	NITROGEN
Error Standard Deviation	25
Total Sample Size	20
Alpha	0.05
Test Degrees of Freedom	4
Error Degrees of Freedom	15

Computed Power
Power
0.743

STA 674, RADOE:
Power Analysis

Example – lettuce

Suppose in the lettuce experiment that the horticulturalist wanted a power of .9. How many lettuce would she need to grow in each plot?

STA 674, RADOE:

Power Analysis

Interpretation

If the true treatment means are as given in the table above and the true standard deviation of the residuals was 25, then the horticulturalist would need to use 30 plots to attain a power of .80 or greater. She should use 6 plots per treatment.

If she repeated the experiment many times under these conditions then the hypothesis that all treatments produce the same number of lettuces on average would be rejected > 90% of the time if she used 6 plots per treatment. In fact, the hypothesis would be rejected 94% of the time.

```
DATA MEANS1;
INPUT NITROGEN LETTUCE;
CARDS;
0 100
50 115
100 130
150 145
200 160
RUN;
PROC GLMPOWER DATA=MEANS1;
CLASS nitrogen;
MODEL lettuce=nitrogen;
POWER stddev=25 power=.9 ntotal=.
RUN;
```

The GLMPOWER Procedure

Fixed Scenario Elements	
Dependent Variable	LETTUCE
Source	NITROGEN
Error Standard Deviation	25
Nominal Power	0.9
Alpha	0.05
Test Degrees of Freedom	4

Computed N Total		
Error DF	Actual Power	N Total
25	0.941	30

STA 674, RADOE:

Power Analysis

```
DATA MEANS2;
INPUT BLOCK NITROGEN LETTUCE;
CARDS;
1 0 110
1 50 125
1 100 140
1 150 155
1 200 170
2 0 90
2 50 105
2 100 120
2 150 135
2 200 150
RUN;
/* Computing power */;
PROC GLMPOWER DATA=MEANS2;
CLASS nitrogen block;
MODEL lettuce=nitrogen block;
POWER stddev=10 power=. ntotal=20;
RUN;
```

The GLMPOWER Procedure

Fixed Scenario Elements	
Dependent Variable	LETTUCE
Error Standard Deviation	10
Total Sample Size	20
Alpha	0.05
Error Degrees of Freedom	14

Computed Power			
Index	Source	Test DF	Power
1	NITROGEN	4	>.999
2	BLOCK	1	0.986