Selection of Sample Size[1]

STA305 Winter 2014

[1] See last slide for copyright information.

Background Reading

Optional

- *Data analysis with SAS*, Chapter 8.
- For technical details about the non-central distributions, Rencher and Schaalje's *Linear models in statistics*, Section 5.4

How many subjects?

- In planning an experiment, one of the biggest decisions is how many experimental units you need.
- Need for what?

We never know the truth

- We wish we knew the exact values of all the parameters, but we never do.
- All we can hope for is to make good decisions (guesses) with high probability.
- We might guess the exact value of a parameter: Estimation.
- We might guess whether some statement about the parameter is right: Testing.

Estimate the value of a linear combination

$$\begin{aligned} \ell &= a_1\mu_1 + a_2\mu_2 + \cdots + a_p\mu_p \\ \hat{\ell} &= a_1\overline{Y}_1 + a_2\overline{Y}_2 + \cdots + a_p\overline{Y}_p \end{aligned}$$

- The estimate will be wrong, with probability one.
- What is the probability that the estimate will be wrong by less than some margin of error m?
- Choose sample size to make this probability acceptably high.

Expected value and variance

$$\begin{aligned} \ell &= a_1\mu_1 + a_2\mu_2 + \cdots + a_p\mu_p \\ \hat{\ell} &= a_1\overline{Y}_1 + a_2\overline{Y}_2 + \cdots + a_p\overline{Y}_p \end{aligned}$$

$$\begin{aligned} E(\hat{\ell}) &= \ell \\ Var(\hat{\ell}) &= \frac{\sigma^2}{n}\sum_{j=1}^{p}\frac{a_j^2}{f_j} \end{aligned}$$

- $f_j = \frac{n_j}{n}$ are relative sample sizes that add to one.
- f_j can be chosen to minimize the variance for any σ^2 and n.
- Assume they have been chosen in a smart way, and treat them as fixed.

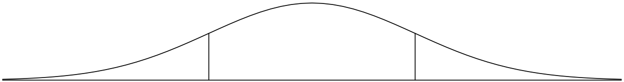
Probability that $\hat{\ell}$ is close to ℓ

First, what is the distribution of $\hat{\ell}$?

$$
\begin{aligned}
Pr\{|\hat{\ell}-\ell|<m\} &= Pr\left\{\left|\frac{\hat{\ell}-\ell}{\sqrt{\frac{\sigma^2}{n}\sum_{j=1}^{p}\frac{a_j^2}{f_j}}}\right| \leq \frac{m}{\sqrt{\frac{\sigma^2}{n}\sum_{j=1}^{p}\frac{a_j^2}{f_j}}}\right\} \\
&= Pr\left\{|Z| < \frac{\sqrt{n}m}{\sigma\sqrt{\sum_{j=1}^{p}\frac{a_j^2}{f_j}}}\right\}
\end{aligned}
$$

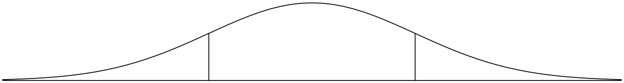
Make the probability as large as you like

$$Pr\left\{|Z| < \frac{\sqrt{n}m}{\sigma\sqrt{\sum_{j=1}^{p}\frac{a_j^2}{f_j}}}\right\}$$



The probability increases to one as $n \to \infty$.

The probability $Pr\left\{|Z| < \frac{\sqrt{n}m}{\sigma\sqrt{\sum_{j=1}^{p}\frac{a_j^2}{f_j}}}\right\}$



- The probability increases to one as $n \to \infty$.
- Set right hand side to 1.96 to and solve for n get a probability of 0.95, etc.
- The a_j and f_j are known.
- The value of m depends on what you mean by "close."
- But σ is tougher.

Do you really have to know σ?

- Maybe you have a good idea of σ from other, similar studies. This is most likely if you are planning a follow-up study. Use $\sqrt{MSE}$.
- Or maybe you can give a high guess and a low guess of σ, and get a *range* of sample sizes you might need.
- This is a lot of guessing.
- There is another possibility.

Another way out

Try to express the desired margin of error in *units of the common standard deviation* σ.

- "Say something like "I want my estimate to be within one-tenth of a standard deviation of the correct answer, with probability 0.90.
- To get an idea of what this means, the mean heights of Canadian men and women are about two standard deviations apart.
- A poll estimates that 40% intend to vote for a candidate, and says "These results are expected to be accurate within four percentage points, 19 times out of 20." They are estimating a margin of error of around 0.08 of a standard deviation, with probability 0.95.

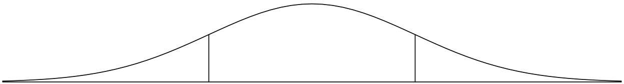
Express the desired margin of error in units of σ

Replace m with $m\sigma$

$$Pr\left\{|Z| < \frac{\sqrt{n}m\cancel{\sigma}}{\cancel{\sigma}\sqrt{\sum_{j=1}^{p}\frac{a_j^2}{f_j}}}\right\} = Pr\left\{|Z| < \frac{\sqrt{n}m}{\sqrt{\sum_{j=1}^{p}\frac{a_j^2}{f_j}}}\right\}$$

To get probability $1-\alpha$, set

$$z_{\alpha/2} = \frac{\sqrt{n}m}{\sqrt{\sum_{j=1}^{p}\frac{a_j^2}{f_j}}}$$



Solve for n

$$z_{\alpha/2} = \frac{\sqrt{n}m}{\sqrt{\sum_{j=1}^{p} \frac{a_j^2}{f_j}}} \quad \Leftrightarrow \quad n = \frac{z_{\alpha/2}^2 \sum_{j=1}^{p} \frac{a_j^2}{f_j}}{m^2}$$

And take the next higher integer. If you "know" σ, let

$$n = \frac{\sigma^2 \, z_{\alpha/2}^2 \sum_{j=1}^{p} \frac{a_j^2}{f_j}}{m^2}$$

Estimate $\mu_1 - \mu_2$

We want the estimate to be accurate to within $\frac{\sigma}{10}$, with probability 0.95.

- $m = \frac{1}{10}$ standard deviations.
- $a_1 = 1$, $a_2 = -1$
- Equal sample sizes, so $f_1 = f_2 = \frac{1}{2}$
- $z_{\alpha/2} = 1.96$

$$\begin{aligned} n &= \frac{z^2_{\alpha/2} \sum_{j=1}^{p} \frac{a_j^2}{f_j}}{m^2} \\ &= \frac{1.96^2(\frac{1^2}{1/2} + \frac{(-1)^2}{1/2})}{(1/10)^2} \\ &= 1536.64 \end{aligned}$$

So need $n = 1537$, or $n_1 = n_2 = 1538/2 = 769$ per group.

Testing (null) hypotheses

- Goal is to make correct decisions with high probability.
- When H_0 is true, probability of a correct decision (don't reject) is $1 - \alpha$. That's guaranteed if the model is correct.
- When H_0 if false, we want to reject it with high probability.
- The probability of rejecting the null hypothesis when the null hypothesis is false is called the *power* of the test.
- Power is one minus the probability of a Type II error.
- It is a function of the true parameter values.
- And also the design, including total sample size.

Power is an increasing function of sample size

- Usually, when H_0 is false, larger sample size yields larger power.
- If power goes to one as a limit when H_0 is false (regardless of the exact parameter values) the test is called *consistent*.
- Most commonly used tests are consistent, including the general linear F-test.
- This means that if H_0 is false, you can make the power as high as you wish by making the sample size bigger.

Strategy

- Pick an effect you'd like to be able to detect. An "effect" means a way that H_0 is wrong. It should be just over the boundary of interesting and meaningful.
- Pick a desired power – a probability with which you'd like to be able to detect the effect by rejecting the null hypothesis.
- Start with a fairly small n and calculate the power. Increase the sample size until the desired power is reached.

Distribution theory

- We need to study the distribution of the test statistic when the null hypothesis is *false*.
- All the distributions you've seen (Z, t, χ^2, F) were derived under the assumption that H_0 is *true*.
- Here we go.

Intermediate Goal

- For the regression model $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$
- Testing $H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{t}$
- We need to know the distribution of the F statistic when H_0 is false.
- It will be called the "non-central" F distribution.

Non-central chi-square with $df = 1$

Let $Z \sim N(\mu, 1)$. Then $W = (Z - \mu)^2 \sim \chi^2(1)$, and $M_W(t) = (1 - 2t)^{-\frac{1}{2}}$. If Z is not "centered" by subtracting off μ,

$$
\begin{aligned}
M_{Z^2}(t) &= E(e^{Z^2 t}) \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{z^2 t}\, e^{-\frac{1}{2}(z-\mu)^2}\, dz \\
&= (1 - 2t)^{-\frac{1}{2}}\, e^{\mu^2 t/(1-2t)}
\end{aligned}
$$

for $t < 1/2$.

Definition of the non-central chi-squared distribution

Generalizing $M_W(t) = (1-2t)^{-\frac{1}{2}} e^{\mu^2 t/(1-2t)}$

The positive random variable W is said to have a non-central chi-squared distribution with degrees of freedom $\nu > 0$ and non-centrality parameter $\lambda \geq 0$ if

$$M_W(t) = (1-2t)^{-\frac{\nu}{2}} e^{\frac{\lambda t}{1-2t}}$$

for $t < 1/2$.

- If $\lambda = 0$, this reduces to the ordinary central chi-squared.
- We have seen that if $Z \sim N(\mu, 1)$, then $W = Z^2 \sim \chi^2(\nu = 1, \lambda = \mu^2)$.

Sum of independent chi-squares

Use $M_W(t) = (1-2t)^{-\frac{\nu}{2}} e^{\frac{\lambda t}{1-2t}}$

Let $Z_1, \ldots, Z_p \stackrel{ind}{\sim} N(\mu_j, 1)$, and $W = \sum_{j=1}^p Z_j^2$. Then

$$\begin{aligned} M_W(t) &= \prod_{j=1}^p M_{Z_j^2}(t) \\ &= \prod_{j=1}^p (1-2t)^{-\frac{1}{2}} e^{\mu_j^2 t/(1-2t)} \\ &= (1-2t)^{-\frac{p}{2}} e^{\frac{\left(\sum_{j=1}^p \mu_j^2\right)t}{1-2t}} \end{aligned}$$

So $W = \mathbf{Z}'\mathbf{Z} \sim \chi^2\left(p, \lambda = \sum_{j=1}^p \mu_j^2\right)$.

Non-central F

Let $W_1 \sim \chi^2(\nu_1, \lambda)$ and $W_2 \sim \chi^2(\nu_2)$ be independent. Then

$$F^* = \frac{Y_1/\nu_1}{Y_2/\nu_2} \sim F(\nu_1, \nu_2, \lambda)$$

is said to have a *non-central* F distribution with degrees of freedom ν_1 and ν_2, and non-centrality parameter λ. Write $F^* \sim F(\nu_1, \nu_2, \lambda)$.

- Reduces to the ordinary central F when $\lambda = 0$.
- Good numerical numerical methods are available for calculating the probabilities.

Theorem

Proof omitted

If $F^* \sim F(\nu_1, \nu_2, \lambda)$, then

- F^* is stochastically increasing in λ, meaning that for every $x > 0$, $Pr\{F^* > x|\lambda\}$ is an increasing function of λ.
- That is, the bigger the non-centrality parameter, the greater the probability of getting F^* above any point (such as a critical value).
- $lim_{\lambda \to \infty} Pr\{F^* > x|\lambda\} = 1$.

Heading for the general linear test of $H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{t}$

With $F = \frac{(\mathbf{C}\widehat{\boldsymbol{\beta}}-\mathbf{t})'(\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}')^{-1}(\mathbf{C}\widehat{\boldsymbol{\beta}}-\mathbf{t})}{q\, MSE}$

- Let $\mathbf{Y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.
- Recall $(\mathbf{Y} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{Y} - \boldsymbol{\mu}) \sim \chi^2(p)$
- When $\mathbf{Y}$ is not centered,

$$\mathbf{Y}'\boldsymbol{\Sigma}^{-1}\mathbf{Y} \sim \chi^2(p, \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}).$$

Let $\mathbf{Y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Then $\mathbf{Y}'\boldsymbol{\Sigma}^{-1}\mathbf{Y} \sim \chi^2(p, \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu})$.

Proof

- Because $\boldsymbol{\Sigma}$ is symmetric and positive definite, the symmetric matrix $\boldsymbol{\Sigma}^{-1/2}$ exists.
- Let $\mathbf{Z} = \boldsymbol{\Sigma}^{-1/2}\mathbf{Y}$. Note $\mathbf{Z}'\mathbf{Z} = \mathbf{Y}'\boldsymbol{\Sigma}^{-1}\mathbf{Y}$.
- $\mathbf{Z}$ is multivariate normal with $E(\mathbf{Z}) = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\mu}$ and

$$
\begin{aligned}
cov(\mathbf{Z}) &= \boldsymbol{\Sigma}^{-1/2} cov(\mathbf{Y}) (\boldsymbol{\Sigma}^{-1/2})' \\
&= \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-1/2} \\
&= \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Sigma}^{1/2}\boldsymbol{\Sigma}^{1/2}\boldsymbol{\Sigma}^{-1/2} \\
&= \mathbf{I} \cdot \mathbf{I} = \mathbf{I}
\end{aligned}
$$

- Thus the elements of $\mathbf{Z}$ are independent normal with variance one, and $\sum_{j=1}^p Z_j^2 = \mathbf{Z}'\mathbf{Z}$ is non-central chi-squared with $df = p$ and non-centrality parameter $\lambda = \sum_{j=1}^p E(Z_j)^2 = E(\mathbf{Z})'E(\mathbf{Z}) = \left(\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\mu}\right)' \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\mu} = \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$. ■

General linear test

- Assume the linear model $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$
- $H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{t}$ may be false.

$$F^* = \frac{(\mathbf{C}\widehat{\boldsymbol{\beta}} - \mathbf{t})'(\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}')^{-1}(\mathbf{C}\widehat{\boldsymbol{\beta}} - \mathbf{t})/q}{MSE} \sim F\left(q, n-p, \lambda\right)$$

where

$$\lambda = \frac{(\mathbf{C}\boldsymbol{\beta} - \mathbf{t})'(\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}')^{-1}(\mathbf{C}\boldsymbol{\beta} - \mathbf{t})}{\sigma^2}$$

To show it

Recall that if $\mathbf{Y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then $\mathbf{Y}'\boldsymbol{\Sigma}^{-1}\mathbf{Y} \sim \chi^2(p, \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu})$.

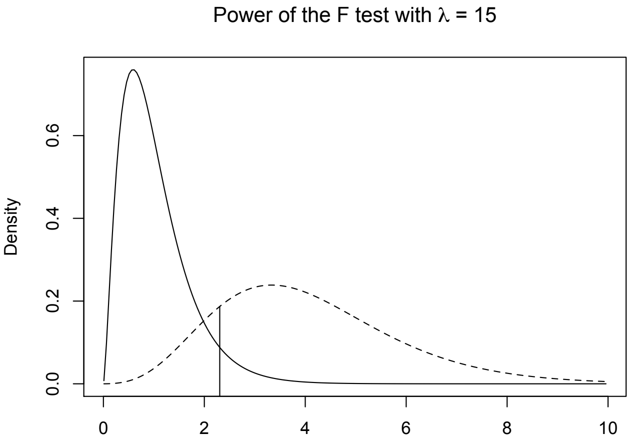
- $\widehat{\boldsymbol{\beta}} \sim N_p\left(\boldsymbol{\beta}, \sigma^2(\mathbf{X}'\mathbf{X})^{-1}\right)$
- $\mathbf{Y} = \mathbf{C}\widehat{\boldsymbol{\beta}} - \mathbf{t} \sim N_q(\mathbf{C}\boldsymbol{\beta} - \mathbf{t}, \sigma^2\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}')$
- So with $\boldsymbol{\mu} = \mathbf{C}\boldsymbol{\beta} - \mathbf{t}$ and $\boldsymbol{\Sigma} = \sigma^2\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}'$, the numerator is non-central chi-squared divided by df.

$$F^* = \frac{(\mathbf{C}\widehat{\boldsymbol{\beta}} - \mathbf{t})'(\mathbf{C}(\sigma^2\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}')^{-1}(\mathbf{C}\widehat{\boldsymbol{\beta}} - \mathbf{t})/q}{MSE/\sigma^2}$$

- $SSE/\sigma^2 \sim \chi^2(n-p)$ regardless of H_0.
- And numerator and denominator are independent because $\widehat{\boldsymbol{\beta}}$ and SSE are independent.

The greater the non-centrality parameter λ, the greater the power

If H_0 is false



What makes λ big?

$$\lambda = \frac{(\mathbf{C}\boldsymbol{\beta} - \mathbf{t})'(\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}')^{-1}(\mathbf{C}\boldsymbol{\beta} - \mathbf{t})}{\sigma^2}$$

- Small σ^2.
- Null hypothesis very wrong.
- Relative sample sizes.
- Total sample size big.
- But sample size is hidden in the $\mathbf{X}'\mathbf{X}$ matrix.

With cell means coding

- Assume there are p treatment combinations.
- The $\mathbf{X}$ matrix has exactly one 1 in each row, and all the rest zeros.
- There are n_j ones in each column.

$$\mathbf{X}'\mathbf{X} = \left[\begin{array}{cccc} n_1 & 0 & \cdots & 0 \\ 0 & n_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & n_p \end{array} \right]$$

Multiplying and dividing by n

$\lambda = \frac{(\mathbf{C}\boldsymbol{\beta}-\mathbf{t})'(\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}')^{-1}(\mathbf{C}\boldsymbol{\beta}-\mathbf{t})}{\sigma^2}$

$$\lambda = n \times (\frac{\mathbf{C}\boldsymbol{\beta}-\mathbf{t}}{\sigma})'(\mathbf{C}\left[\begin{array}{cccc} 1/f_1 & 0 & \cdots & 0 \\ 0 & 1/f_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/f_p \end{array}\right]\mathbf{C}')^{-1}(\frac{\mathbf{C}\boldsymbol{\beta}-\mathbf{t}}{\sigma})$$

- $f_1, \ldots f_p$ are relative sample sizes: $f_j = n_j/n$
- $\mathbf{C}\boldsymbol{\beta}-\mathbf{t}$ is an *effect*, a particular way in which the null hypothesis is wrong. It is naturally expressed in units of the common within-treatment standard deviation sigma, and in general there is no reasonable way to avoid it.
- Almost always, $\mathbf{t} = \mathbf{0}$.
- The non-centrality parameter is sample size times a quantity that is sometimes called “effect size.”
- The idea is that effect size represents how wrong H_0 is.

Example: Comparing two means

Suppose we have a random sample of size n_1 from a normal distribution with mean μ_1 and variance σ^2, and independently, a second random sample from a normal distribution with mean μ_2 and variance σ^2. We wish to test $H_0 : \mu_1 = \mu_2$ versus the alternative $H_a : \mu_1 \neq \mu_2$. If the true means are a half a standard deviation apart, we want to be able to detect it with probability 0.80.

We'll do it with cell means coding, letting $x_{i,1} = 1$ if observation i is from treatment one (and zero otherwise), and $x_{i,2} = 1$ if observation i is from treatment two (and zero otherwise).

- The model is $Y_i = \beta_1 x_{i,1} + \beta_2 x_{i,2} + \epsilon_i$
- $\beta_1 = \mu_1$, $\beta_2 = \mu_2$
- $H_0 : \beta_1 - \beta_2 = 0$

$H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{t}$

$$
\begin{aligned}
\mathbf{C} &= (1, -1) \\
\boldsymbol{\beta} &= \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \\
t &= (0) \\
\lambda &= n \times \left(\frac{\mathbf{C}\boldsymbol{\beta} - \mathbf{t}}{\sigma}\right)' \left(\mathbf{C} \begin{pmatrix} 1/f_1 & 0 \\ 0 & 1/(1-f_1) \end{pmatrix} \mathbf{C}'\right)^{-1} \left(\frac{\mathbf{C}\boldsymbol{\beta} - \mathbf{t}}{\sigma}\right)
\end{aligned}
$$

Continuing the calculation

$$
\begin{aligned}
\lambda &= n \left(\frac{\mathbf{C}\boldsymbol{\beta}-\mathbf{t}}{\sigma}\right)' \left(\mathbf{C}\begin{pmatrix} 1/f_1 & 0 \\ 0 & 1/(1-f_1) \end{pmatrix}\mathbf{C}'\right)^{-1} \left(\frac{\mathbf{C}\boldsymbol{\beta}-\mathbf{t}}{\sigma}\right) \\
&= n \left(\frac{\mu_1-\mu_2}{\sigma}\right) \left(\frac{1}{f_1}+\frac{1}{1-f_1}\right)^{-1} \left(\frac{\mu_1-\mu_2}{\sigma}\right) \\
&= n\ f_1(1-f_1) \left(\frac{\mu_1-\mu_2}{\sigma}\right)^2 \\
&= n\ f_1(1-f_1)\ d^2
\end{aligned}
$$

$\lambda = nf(1-f)d^2$, where $f = \frac{n_1}{n}$ and $d = \frac{|\mu_1 - \mu_2|}{\sigma}$

- For two-sample problems, d is usually called effect size. The effect size specifies how wrong the null hypothesis is, by expressing the absolute difference between means in units of the common within-cell standard deviation.
- The non-centrality parameter (and hence, power) depends on the three parameters μ_1, μ_2 and σ^2 only through the effect size d.
- Power depends on sample size, effect size and an aspect of design allocation of relative sample size to treatments. Equal sample sizes yield the highest power in the 2-sample case.

Back to the problem

$\lambda = nf(1-f)d^2$

We wish to test $H_0 : \mu_1 = \mu_2$ versus the alternative $H_a : \mu_1 \neq \mu_2$. If the true means are a half a standard deviation apart, we want to be able to detect it with probability 0.80.

$$\begin{aligned}
\lambda &= nf(1-f)\left(\frac{|\mu_1 - \mu_2|}{\sigma}\right)^2 \\
&= n\frac{1}{2}\left(1 - \frac{1}{2}\right)\left(\frac{1}{2}\right)^2 \\
&= \frac{n}{16}
\end{aligned}$$

SAS proc iml

```
/*****************  fpow1.sas ********************/
options linesize=79 noovp formdlim='_' nodate;
title 'Two-sample power analysis';

proc iml; /* Replace alpha, q, p, d and wantpow below    */
     alpha = 0.05;  /* Signif. level for testing H0: C Beta = t  */
     q = 1;         /* Numerator df = # rows in C matrix         */
     p = 2;         /* There are p beta parameters               */
     d = 1/2;       /* d = |mu1-mu2|/sigma */
     wantpow = .80; /* Find n to yield this power                */
     power = 0; n = p; oneminus = 1-alpha; /* Initializing ... */
     do until (power >= wantpow);
        n=n+1 ;
        ncp = n * 1/4 * d**2;
        df2 = n-p;
        power = 1-probf(finv(oneminus,q,df2),q,df2,ncp);
     end;
     print alpha p q d wantpow;
     print "Required sample size is " n;
     print "For a power of " power;
/*************************************************************
```

Output

```
                Two-sample power analysis

        alpha          p          q          d    wantpow

         0.05          2          1        0.5        0.8


                                               n

           Required sample size is           128


                                   power

                For a power of  0.8014596
```

To do a power analysis for any factorial design

$H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{t}$

All you need is a vector of relative sample sizes and a vector of numbers representing the differences between $\mathbf{C}\boldsymbol{\beta}$ and $\mathbf{t}$ in units of sigma.

$$\lambda = n \times \left(\frac{\mathbf{C}\boldsymbol{\beta} - \mathbf{t}}{\sigma} \right)' \left(\mathbf{C} \begin{bmatrix} 1/f_1 & 0 & \cdots & 0 \\ 0 & 1/f_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/f_p \end{bmatrix} \mathbf{C}' \right)^{-1} \left(\frac{\mathbf{C}\boldsymbol{\beta} - \mathbf{t}}{\sigma} \right)$$

Example: Test for interaction

	Level of B		
Level of A	1	2	Average
1	μ_{11}	μ_{12}	$\mu_{1.}$
2	μ_{21}	μ_{22}	$\mu_{2.}$
3	μ_{31}	μ_{32}	$\mu_{3.}$
Average	$\mu_{.1}$	$\mu_{.2}$	$\mu_{..}$

$$H_0 : \mu_{11} - \mu_{12} = \mu_{21} - \mu_{22} = \mu_{31} - \mu_{32}$$

$H_0 : \mu_{11} - \mu_{12} = \mu_{21} - \mu_{22} = \mu_{31} - \mu_{32}$

$$\begin{array}{ccc} \mathbf{C} & \boldsymbol{\beta} & = & \mathbf{t} \\ \left(\begin{array}{rrrrrr} 1 & -1 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & -1 & 1 \end{array}\right) & \left(\begin{array}{c} \mu_{11} \\ \mu_{12} \\ \mu_{21} \\ \mu_{22} \\ \mu_{31} \\ \mu_{32} \end{array}\right) & = & \left(\begin{array}{c} 0 \\ 0 \end{array}\right) \end{array}$$

Suppose this null hypothesis is false in a particular way that we want to be able to detect.

Null hypothesis is wrong

Suppose that for $A = 1$ and $A = 2$, the population mean of Y is a quarter of a standard deviation higher when $B = 2$, but if $A = 3$, the population mean of Y is a quarter of a standard deviation higher for $B = 1$. Of course there are infinitely many sets of means satisfying these constraints, even if they are expressed in standard deviation units. But they will all have the same effect size. One such pattern is the following.

	Level of B	
Level of A	1	2
1	0.000	0.250
2	0.000	0.250
3	0.000	-0.250

Sample sizes are all equal, and we want to be able to detect an effect of this magnitude with probability at least 0.80.

All we need is true $\mathbf{C}\boldsymbol{\beta}$

$H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{0}$ is wrong.

	Level of B	
Level of A	1	2
1	0.000	0.250
2	0.000	0.250
3	0.000	-0.250

$$\mathbf{C}\boldsymbol{\beta} = \begin{pmatrix} 0 \\ -0.5 \end{pmatrix}$$

Matrix calculations with proc iml

$\lambda = \frac{(\mathbf{C}\boldsymbol{\beta}-\mathbf{t})'(\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}')^{-1}(\mathbf{C}\boldsymbol{\beta}-\mathbf{t})}{\sigma^2}$

```
/*****************  fpow2.sas ********************/
options linesize=79 noovp formdlim='_';
title 'Sample size calculation for the interaction example';

proc iml;
/********* Edit this input: Rows of matrices are separated by commas  ********/
     alpha = 0.05; wantpow = .80;
     f = {1,1,1,1,1,1};          /* Relative sample sizes       */
     C = { 1 -1 -1  1  0  0,       /* Contrast matrix              */
           0  0  1 -1 -1  1};
     eff = {0, 0.5};    /* In standard deviation units */
/*****************************************************************************/
```

fpow2.sas continued

$\lambda = \frac{(\mathbf{C}\boldsymbol{\beta}-\mathbf{t})'(\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}')^{-1}(\mathbf{C}\boldsymbol{\beta}-\mathbf{t})}{\sigma^2}$

```
    p = nrow(f) ; q = nrow(eff); f = f/sum(f);
    core = inv(C*inv(diag(f))*C‘);
    effsize = eff‘*core*eff;
    power = 0; n = p; oneminus = 1-alpha;  /* Initializing ...*/
    do until (power >= wantpow);
       n = n+1 ;
       ncp = n * effsize;
       df2 = n-p;
       power = 1-probf(finv(oneminus,q,df2),q,df2,ncp);
    end; /*  End Loop */
    print " ";
    print " Required sample size is " n " for a power of " power;
    print " " ;
/**********************************************************
```

Output

```
                Sample size calculation for the interaction example                 1


                                        n                       power

           Required sample size is        697   for a power of   0.8001726
```

$697/6 = 116.1667$ and $117 * 6 = 702$, so a total of $n = 702$ experimental units are needed for equal sample sizes.

Copyright Information