Lecture 5: Determining Sample Size

Montgomery: Section 3.7 and 13.4
Choice of Sample Size: Fixed Effects

- Can determine the sample size for
  - Overall $F$ test
  - Contrasts of interest

- For simplicity, typically assume $n_i$’s constant, i.e.,
  $$n_1 = n_2 = \cdots = n_a = n$$

- Recall
  - Type I error rate: $\alpha = \Pr(\text{Reject } H_0 | H_0)$
  - Type II error rate: $\beta = \Pr(\text{Accept } H_0 | H_1)$
  - Power = $\Pr(\text{Reject } H_0 | H_1) = 1 - \beta$

- Need to know
  - Test Statistics
  - Distr. of test statistics under $H_0 \implies$ Reject Region (for given $\alpha$)
  - Distr. of test statistics under $H_1 \implies$ power = $\Pr(\text{Reject Region} | H_1)$
Determining Power for $F$ Test

- $\alpha = \Pr(F_0 > F_{\alpha,a-1,N-a}|H_0)$
- $\beta = \Pr(F_0 < F_{\alpha,a-1,N-a}|H_1)$

- Need to know distribution of $F_0$ when $H_1$ is true
  - Can show $F_0 = \frac{MS_{Trt}}{MS_E} \sim F_{a-1,N-a}(\delta)$
  - $\delta = n \sum \tau_i^2 / \sigma^2$ (non-centrality parameter)

- Recall $E(MS_{Trt}) = \sigma^2 + n \sum \tau_i^2 / (a - 1)$
  - $\delta = \{E(MS_{Trt}) - E(MS_E)\} \times df_{Trt} / E(MS_E)$

- Need to specify $\{\tau_i\}$ (Note the zero-sum constraint: $\sum_{i=1}^{a} \tau_i = 0$)

- Power will vary for different choices
Power Calculation for $F$ Test

- Given $\alpha$, $a$, and $n$, can determine $F_{\alpha,a-1,N-a}$

- Given some value of $\delta$, can use noncentral $F$ to compute power
  - In SAS, use function PROBF
  - Power$=1-\text{PROBF}(F_{\alpha,a-1,N-a},a-1,N-a,\delta)$

- Montgomery: OCC given in Chart V
  - Plots $\beta$ vs $\Phi$
  - $\Phi^2 = \delta/a = n \sum \tau_i^2 / (a\sigma^2)$
  - Can use charts to determine power or sample size
Methods to Determine $\delta$ or $\Phi^2$

1. Choose treatment means ($\mu + \tau_i$)
   - Solve for $\{\tau_i\}$ and compute $\Phi^2$ or $\delta$
   - Difficult to know what means to select

2. Take a minimum difference approach
   - Suppose there exists a pair of $(i, j)$ such that $|\tau_i - \tau_j| \geq D$
   - The minimum value: $\Phi^2 = nD^2/(2a\sigma^2)$ (e.g.,
     $\{\tau_i\} = \{-D/2, 0, \ldots, 0, D/2\}$)
   - Power of test is at least $1 - \beta$

3. Specify a standard deviation increase in percentage ($P$)
   - Under $H_1$, variance of a randomly chosen $y_i$ is $\sigma_y^2 = \sigma^2 + \sum \tau_i^2/a$
   - Randomly chosen $\tau_i$ has mean 0 and variance $\sum \tau_i^2/a$
   - $P = \left( \sqrt{\sigma^2 + \sum \tau_i^2/a} / \sigma - 1 \right) \times 100$
   - $\delta = an\{(1 + .01P)^2 - 1\}$
   - $\Phi^2 = n\{(1 + .01P)^2 - 1\}$
Power Calculation for Specific Contrast

- Often with an experiment, a researcher is primarily interested in just a few comparisons or contrasts. In these cases, it can be preferable to determine sample size for these rather than the overall $F$ test.

- This reduces problem back to the $t$ test situation

- Need to determine
  - Difference of importance
  - Standard error of comparison

- May want/need to adjust for multiple comparisons

- Montgomery describes confidence interval approach
  - Consider pairwise difference in treatment means
  - Specify length of $(1 - \alpha) \times 100\%$ confidence interval
  - Length/2 = $t_{\alpha/2, N-a} \sqrt{\frac{2\text{MS}_E}{n}}$
  - Based on the choice of $\text{MS}_E$, find $n$
Example 3.1 – Etch Rate (Page 75)

• Consider new experiment to investigate 5 RF power settings equally spaced between 180 and 200 W

• Wants to determine sample size to detect a mean difference of $D=30$ (Å/min) with 80% power

• Will use Example 3.1 estimates to determine new sample size

$$\hat{\sigma}^2 = 333.7, \quad D = 30, \quad \text{and} \quad \alpha = 0.05$$

• Using Table V: $V^2 = 900 \times n / (2 \times 5 \times 333.7) \approx 0.27 \times n$

<table>
<thead>
<tr>
<th>$n$</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Phi$</td>
<td>$\sqrt{2.43} \approx 1.56$</td>
<td>$\sqrt{2.70} \approx 1.64$</td>
<td>$\sqrt{3.0} \approx 1.72$</td>
</tr>
<tr>
<td>$df_E$</td>
<td>40</td>
<td>45</td>
<td>50</td>
</tr>
<tr>
<td>$\beta$</td>
<td>26%</td>
<td>20%</td>
<td>15%</td>
</tr>
<tr>
<td>power</td>
<td>74%</td>
<td>80%</td>
<td>85%</td>
</tr>
</tbody>
</table>
Using SAS : $\delta = a\Phi^2$

```sas
data new; a=5; alpha=.05; d=30; var=333.7;
    do n=5 to 15;
        df = a*(n-1);
        nc = n*d*d/(2*var);
        fcut = finv(1-alpha,a-1,df);
        beta = probf(fcut,a-1,df,nc);
        power = 1-beta; output;
    end;
proc print;
    var n df nc beta power; run;
```

<table>
<thead>
<tr>
<th>Obs</th>
<th>n</th>
<th>df</th>
<th>nc</th>
<th>beta</th>
<th>power</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>20</td>
<td>6.7426</td>
<td>0.57654</td>
<td>0.42346</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>25</td>
<td>8.0911</td>
<td>0.47884</td>
<td>0.52116</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>30</td>
<td>9.4396</td>
<td>0.39034</td>
<td>0.60966</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>35</td>
<td>10.7881</td>
<td>0.31289</td>
<td>0.68711</td>
</tr>
<tr>
<td>5</td>
<td>9</td>
<td>40</td>
<td>12.1366</td>
<td>0.24703</td>
<td>0.75297</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>45</td>
<td>13.4852</td>
<td>0.19234</td>
<td>0.80766 ***n=10 needed</td>
</tr>
<tr>
<td>7</td>
<td>11</td>
<td>50</td>
<td>14.8337</td>
<td>0.14788</td>
<td>0.85212</td>
</tr>
<tr>
<td>8</td>
<td>12</td>
<td>55</td>
<td>16.1822</td>
<td>0.11239</td>
<td>0.88761</td>
</tr>
<tr>
<td>9</td>
<td>13</td>
<td>60</td>
<td>17.5307</td>
<td>0.08451</td>
<td>0.91549</td>
</tr>
<tr>
<td>10</td>
<td>14</td>
<td>65</td>
<td>18.8792</td>
<td>0.06292</td>
<td>0.93708</td>
</tr>
<tr>
<td>11</td>
<td>15</td>
<td>70</td>
<td>20.2277</td>
<td>0.04641</td>
<td>0.95359</td>
</tr>
</tbody>
</table>
• Compare all pairs and detect any difference more than 30 (Å/min) with 80% power
• A pairwise comparison takes $t_0 = (\bar{Y}_i. - \bar{Y}_k.)/\sqrt{2}\text{MSE}/n$
• Consider Tukey’s adjustment: reject when $|t_0| > q_{.05}(5, df_E)/\sqrt{2}$
• $t_0 \overset{H_1}{\sim} t_{df_E}(\delta)$ with $\delta = (\mu_i - \mu_k)/\sqrt{2}\sigma^2/n$,
• Use PROBMC in SAS to get the quantile for multiple comparisons

```sas
data new1; a=5; alpha=.05; var=333.7; d=30;
do n=8 to 12;
   df = a*(n-1); nc = d/sqrt(var*2/n);
   /* crit = tinv(1-alpha/2,df); /*LSD approach*/
   crit = probmc("range",.,1-alpha,df,a)/sqrt(2); /*Tukey*/
   power=1-probt(crit,df,nc)+probt(-crit,df,nc); output;
end;
proc print; var n df power; run;
```

<table>
<thead>
<tr>
<th>Obs</th>
<th>n</th>
<th>df</th>
<th>power</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8</td>
<td>35</td>
<td>0.65814</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>40</td>
<td>0.73085</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>45</td>
<td>0.79139</td>
</tr>
<tr>
<td>4</td>
<td>11</td>
<td>50</td>
<td>0.84057</td>
</tr>
<tr>
<td>5</td>
<td>12</td>
<td>55</td>
<td>0.87971</td>
</tr>
</tbody>
</table>
• Interested in comparing each setting to 200 W, and detect a difference more than 30 (Å/min) with 80% power

• A pairwise comparison takes \( t_0 = (\bar{Y}_i - \bar{Y}_c) / \sqrt{2 \text{MSE}/n} \)

• Consider using Dunnett’s adjustment: reject when \( |t_0| > d_{.05}(4, df_E) \)

• \( t_0 \overset{H_1}{\sim} t_{df_E}(\delta) \) with \( \delta = (\mu_i - \mu_k) / \sqrt{2\sigma^2/n}, \)

\[
\begin{align*}
data \text{ new1; a=5; alpha=.05; var=333.7; d=30;}
do \ n=7 \to 12;
\quad \text{df = a*(n-1); nc = d/sqrt(var*2/n);}
\quad \text{crit = probmc("dunnett2",.,1-alpha,df,a-1); /*Two-Sided Dunnett*/}
\quad \text{power=1-probt(crit,df,nc)+probt(-crit,df,nc); output;}
end;
\end{align*}
\]

\[
\begin{align*}
\text{proc print;}
\quad \text{var n df power; run;}
\end{align*}
\]

\[
\begin{align*}
\text{Obs} & \quad n & \quad df & \quad \text{power} \\
1 & 7 & 30 & 0.68794 \\
2 & 8 & 35 & 0.76201 \\
3 & 9 & 40 & 0.82136 \quad ***\text{Only 9 replicates needed here} \\
4 & 10 & 45 & 0.86780 \\
5 & 11 & 50 & 0.90341 \\
6 & 12 & 55 & 0.93024
\end{align*}
\]
Use of **PROC POWER** in SAS

- **PROC POWER** and **PROC GLMPOWER** both calculate $\sum_{i=1}^{a} \tau_i^2$ using pre-specified treatment means $\mu_i$

```sas
proc power;
   onewayanova test=overall power=.80 npergroup=. stddev=18.27
    groupmeans = -15|0|0|0|15;
run; quit;
```

---

**Overall F Test for One-Way ANOVA**

<table>
<thead>
<tr>
<th>Fixed Scenario Elements</th>
<th>Exact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method</td>
<td>Exact</td>
</tr>
<tr>
<td>Group Means</td>
<td>-15</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>18.27</td>
</tr>
<tr>
<td>Nominal Power</td>
<td>0.8</td>
</tr>
<tr>
<td>Alpha</td>
<td>0.05</td>
</tr>
</tbody>
</table>

| Computed N Per Group                     |
|------------------------------------------|-------|
| Actual Power                             | N Per Group |
| 0.808                                    | 10     |
Neither PROC POWER nor PROC GLMPOWER can easily do multiple comparison adjustment.

```plaintext
proc power;
    onewayanova test=contrast power=.80 npergroup=. stddev=18.27
    groupmeans = -15|0|0|0|15
    contrast = (1 0 0 0 -1);
run; quit;
```

---

**Single DF Contrast in One-Way ANOVA**

<table>
<thead>
<tr>
<th>Fixed Scenario Elements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method</td>
</tr>
<tr>
<td>Contrast Coefficients</td>
</tr>
<tr>
<td>Group Means</td>
</tr>
<tr>
<td>Standard Deviation</td>
</tr>
<tr>
<td>Nominal Power</td>
</tr>
<tr>
<td>Number of Sides</td>
</tr>
<tr>
<td>Null Contrast Value</td>
</tr>
<tr>
<td>Alpha</td>
</tr>
</tbody>
</table>

Computed N Per Group

<table>
<thead>
<tr>
<th>Actual Power</th>
<th>N Per Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.844</td>
<td>7</td>
</tr>
</tbody>
</table>
**Use of PROC GLMPOWER in SAS**

data trtmeans; input trt resp @@; datalines;
1 -15 2 0 3 0 4 0 5 15
;

proc glmpower data=trtmeans;
  class trt;
  model resp = trt;
  contrast 'trt1 vs. trt5' trt 1 0 0 0 -1;
  power stddev=18.27 alpha=0.05 ntotal=; power=0.8;
run; quit;

Fixed Scenario Elements

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>resp</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha</td>
<td>0.05</td>
</tr>
<tr>
<td>Error Standard Deviation</td>
<td>18.27</td>
</tr>
<tr>
<td>Nominal Power</td>
<td>0.8</td>
</tr>
</tbody>
</table>

Computed N Total

<table>
<thead>
<tr>
<th>Index</th>
<th>Type</th>
<th>Source</th>
<th>Test DF</th>
<th>Error DF</th>
<th>Actual Power</th>
<th>Power</th>
<th>N Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Effect</td>
<td>trt</td>
<td>4</td>
<td>45</td>
<td>0.808</td>
<td></td>
<td>50</td>
</tr>
<tr>
<td>2</td>
<td>Contrast</td>
<td>trt1 vs. trt5</td>
<td>1</td>
<td>30</td>
<td>0.844</td>
<td></td>
<td>35</td>
</tr>
</tbody>
</table>
Choice of Sample Size: Random Effects

- Can use central F distribution
  - \((N - a)MS_E/\sigma^2 \sim \chi^2_{N-a}\)
  - \((a - 1)MS_{Trt}/(\sigma^2 + n\sigma^2_\tau) \sim \chi^2_{a-1}\)
  - Thus \(F_0/\lambda^2 \sim F_{a-1,N-a}\), where \(\lambda^2 = \frac{E(MS_{Trt})}{E(MS_E)} = 1 + n\sigma^2_\tau/\sigma^2\)
  - Power: \(P(F_0 > F_{\alpha,a-1,N-a}\mid \sigma^2_\tau > 0) = P(F > F_{\alpha,a-1,N-a}/\lambda^2)\)
  - Can specify ratio of \(\sigma^2_\tau/\sigma^2\)
  - Can specify percentage increase \(P = (\sqrt{\sigma^2 + \sigma^2_\tau/\sigma} - 1) \times 100\)

- OCC given in Chart VI
  - Plots \(\beta\) vs \(\lambda\)

- Use SAS function \texttt{PROBF}
  - power = 1-\texttt{PROBF}(\(F_{\alpha,a-1,N-a}/\lambda^2,a-1,N-a\))
Example: Batch Example on Slide 8 of Lecture 4

- Consider new experiment with 5 batches: a random effects problem
  - The variance estimate is $\sigma^2 = 1.8$
  - Desire to detect situation when $\sigma^2 \geq 3.6 = 2\sigma^2$
  - Set power at 80% and $\alpha = .05$

- Using Table VI: $\lambda = \sqrt{1 + 2n}$

<table>
<thead>
<tr>
<th>$n$</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda$</td>
<td>$\sqrt{7} \approx 2.65$</td>
<td>$\sqrt{9} = 3$</td>
<td>$\sqrt{11} \approx 3.32$</td>
</tr>
<tr>
<td>$df_E$</td>
<td>10</td>
<td>15</td>
<td>20</td>
</tr>
<tr>
<td>$\beta$</td>
<td>28%</td>
<td>18%</td>
<td>15%</td>
</tr>
<tr>
<td>power</td>
<td>72%</td>
<td>82%</td>
<td>85%</td>
</tr>
</tbody>
</table>

- Appears $n = 4$ gives appropriate power
Using SAS

data new; a = 5; alpha=.05; ratiovar=2.0;
  do n=2 to 10;
    df = a*(n-1);
    lambdasq = 1+ratiovar*n;
    fcut = finv(1-alpha,a-1,df);
    beta=probf(fcut/lambdasq,a-1,df);
    power = 1-beta; output;
  end;

proc print;
  var n beta power; run;

<table>
<thead>
<tr>
<th>Obs</th>
<th>n</th>
<th>beta</th>
<th>power</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>0.52933</td>
<td>0.47067</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>0.26112</td>
<td>0.73888</td>
</tr>
</tbody>
</table>
| 3   | 4 | 0.15292 | 0.84708 *n=4 gives the power
| 4   | 5 | 0.10027 | 0.89973 |
| 5   | 6 | 0.07081 | 0.92919 |
| 6   | 7 | 0.05267 | 0.94733 |
| 7   | 8 | 0.04072 | 0.95928 |
| 8   | 9 | 0.03242 | 0.96758 |
| 9   | 10| 0.02643 | 0.97357 |