MIXED MODEL PROCEDURES TO ASSESS POWER, PRECISION, AND SAMPLE SIZE IN THE DESIGN OF EXPERIMENTS
Introduction

Two Main Questions Statistical Consultants are Asked:

- “How should I analyze my data?”
- “Here is my design – how many replicates do I need?”

To answer: need the objectives of the study

- Treatment Design
- Experiment Design

- GOAL: efficient design, cost effectiveness
Stroup’s Method

- Uses PROC MIXED to:
  - Clarify the treatment design and comparisons among treatments needed
  - Choose among competing designs
  - Decide on a sample size

- Pre-Stroup: Could only calculate n for CRD, RCB
  - What if design is more complex?
Background Useful Facts:

- Power is the probability of detecting an effect given that an effect exists.
- Can calculate F statistics for hypotheses which are estimable functions.
- Non centrality parameter is used to find power.
  - Power = 1 - prob(type 2 error)
- Power can be used to determine minimum sample size to detect effect.
4 step process in SAS

- Create data set with desired structure
  - Instead of observed data, use means which reflect the desired effect

- In a data step:
  - # of treatments
  - #replicates
  - subsampling/stripping/splitting
Example

- Design: Split Plot
  - 2 Drugs (Standar, Experimental)
  - 3 Doses (Low, Med, High)
    - Looking to see if linear response is the same in both drugs (linear dose x treatment contrast)

- Data effect_sizes sets the treatments
- Data trials generates the data sets with the desired structure
Step 2

- Run PROC MIXED with variance-covariance components set to the values you want

- Need to use noiter and noprofile commands to keep it from running REML iterations and use your variances

- Use variances in the range you think reasonably fits your data
Example continued

proc mixed data=trials noprofile;
title2 'Estimate the non-centrality parameter ';
by set n_whole;
class &wp_fact &sp_fact &wpeu;
model mu = &wp_fact | &sp_fact;
random &wpeu(&wp_fact);
parms (&wp_var) (&sp_var) / noiter;
  /* this is where the estimated variance components are specified */
contrast 'lin x trt' &wp_fact.*&sp_fact 1 0 -1 -1 0 1;
  /* save the results to the ods datasets */
ods output tests3=power_effects; ods output contrasts =power_contrasts;
run;
Step 3

- Use model and contrast statements to compute F statistics of the effects you are interested in.

- These are the non-centrality parameters (we will use these for our power analysis).
Example continued

```plaintext
proc mixed data=trials noprofile;
title2 'Estimate the non-centrality parameter';
by set n_whole;
class &wp_fact &sp_fact &wpeu;
model mu = &wp_fact | &sp_fact;
random &wpeu(&wp_fact);
parms (&wp_var) (&sp_var) / noiter;
contrast 'lin x trt' &wp_fact.*&sp_fact 1 0 -1 -1 0 1;
ods output tests3=power_effects;
ods output contrasts =power_contrasts;
run;
```
Step 4

- Use function statements for the F distribution to compute the power

- Approximates the Non Centrality Parameter from the F statistics calculated in Step 3, combines these estimates to give the estimated power
/* now to compute approximations to the power */
data power; set power;
cnc = numdf*Fvalue;
/* approximate non-centrality parameter */
ficrit = finv(1-&alpha, numDF, denDF, 0);
/* estimate critical value */
power = 1 - probf(fcrit, numdf, dendf, nc);
/* estimated power */
attrib power label='Power' format=7.2;
attrib nc label='Non-centrality' format=7.1;
attrib ficrit label='F-critical value' format=7.2;
drop probF;
run;
### Table: ANOVA Summary

<table>
<thead>
<tr>
<th>Obs</th>
<th>Plot</th>
<th>LEVEL</th>
<th>Effect</th>
<th>Num DF</th>
<th>Den DF</th>
<th>F Value</th>
<th>Non-centrality</th>
<th>F-critical value</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>type</td>
<td>1</td>
<td>6</td>
<td>5.33</td>
<td>5.3</td>
<td>5.99</td>
<td>0.49</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>dose</td>
<td>2</td>
<td>12</td>
<td>48.00</td>
<td>96.0</td>
<td>3.89</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>type*dose</td>
<td>2</td>
<td>12</td>
<td>5.33</td>
<td>10.7</td>
<td>3.69</td>
<td>0.73</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>lin x trt</td>
<td>1</td>
<td>12</td>
<td>10.67</td>
<td>10.7</td>
<td>4.75</td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>type</td>
<td>1</td>
<td>9</td>
<td>6.27</td>
<td>6.7</td>
<td>5.52</td>
<td>0.62</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>dose</td>
<td>2</td>
<td>16</td>
<td>60.00</td>
<td>130.0</td>
<td>3.63</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>5</td>
<td>type*dose</td>
<td>2</td>
<td>16</td>
<td>6.67</td>
<td>13.3</td>
<td>3.63</td>
<td>0.85</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>5</td>
<td>lin x trt</td>
<td>1</td>
<td>16</td>
<td>13.33</td>
<td>13.3</td>
<td>4.49</td>
<td>0.93</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>type</td>
<td>1</td>
<td>10</td>
<td>6.00</td>
<td>8.0</td>
<td>4.96</td>
<td>0.72</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>6</td>
<td>dose</td>
<td>2</td>
<td>20</td>
<td>72.00</td>
<td>144.0</td>
<td>3.49</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>6</td>
<td>type*dose</td>
<td>2</td>
<td>20</td>
<td>6.00</td>
<td>16.0</td>
<td>3.49</td>
<td>0.92</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>6</td>
<td>lin x trt</td>
<td>1</td>
<td>20</td>
<td>16.00</td>
<td>16.0</td>
<td>4.95</td>
<td>0.97</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>7</td>
<td>type</td>
<td>1</td>
<td>12</td>
<td>9.33</td>
<td>9.3</td>
<td>4.75</td>
<td>0.80</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>7</td>
<td>dose</td>
<td>2</td>
<td>24</td>
<td>94.00</td>
<td>168.0</td>
<td>3.40</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>7</td>
<td>type*dose</td>
<td>2</td>
<td>24</td>
<td>9.33</td>
<td>18.7</td>
<td>3.40</td>
<td>0.96</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>7</td>
<td>lin x trt</td>
<td>1</td>
<td>24</td>
<td>18.67</td>
<td>18.7</td>
<td>4.26</td>
<td>0.99</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>8</td>
<td>type</td>
<td>1</td>
<td>14</td>
<td>10.67</td>
<td>10.7</td>
<td>4.60</td>
<td>0.86</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>9</td>
<td>dose</td>
<td>2</td>
<td>20</td>
<td>96.00</td>
<td>192.0</td>
<td>3.84</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>9</td>
<td>type*dose</td>
<td>2</td>
<td>20</td>
<td>10.67</td>
<td>21.3</td>
<td>3.34</td>
<td>0.98</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>9</td>
<td>lin x trt</td>
<td>1</td>
<td>20</td>
<td>21.53</td>
<td>21.5</td>
<td>4.20</td>
<td>0.99</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>9</td>
<td>type</td>
<td>1</td>
<td>16</td>
<td>12.00</td>
<td>12.0</td>
<td>4.49</td>
<td>0.90</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>9</td>
<td>dose</td>
<td>2</td>
<td>32</td>
<td>100.00</td>
<td>216.0</td>
<td>3.29</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>9</td>
<td>type*dose</td>
<td>2</td>
<td>32</td>
<td>12.00</td>
<td>24.0</td>
<td>3.29</td>
<td>0.99</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>9</td>
<td>lin x trt</td>
<td>1</td>
<td>32</td>
<td>24.00</td>
<td>24.0</td>
<td>4.15</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>10</td>
<td>type</td>
<td>1</td>
<td>18</td>
<td>13.33</td>
<td>13.3</td>
<td>4.81</td>
<td>0.93</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>10</td>
<td>dose</td>
<td>2</td>
<td>16</td>
<td>120.00</td>
<td>240.0</td>
<td>3.26</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>10</td>
<td>type*dose</td>
<td>2</td>
<td>16</td>
<td>13.33</td>
<td>26.7</td>
<td>3.26</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>10</td>
<td>lin x trt</td>
<td>1</td>
<td>16</td>
<td>26.67</td>
<td>26.7</td>
<td>4.11</td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>
Why is this useful

- Competing designs can be compared for power efficiency
- Can be used for any design! Randomness is ok!
- Can compute variance of estimable functions – will tell you precision of your design
- If you can pick the most powerful design for your number of replicates, can get better results for the same price
Conclusions

1. Mixed model methods allow power/precision analysis for pretty much ANY design (yay!)

2. Client and consultant need to understand the steps that need to happen before sample size is determined
   - Precise definitions of objectives and how they will be addressed statistically
   - Specific estimable functions of interest
Needed to run Stoup’s Method:

- Agree on the smallest difference of scientific importance
  - What kind of difference is relevant?
  - What magnitude makes it important?
  - How do we recognize this?

- Variances
  - Consult literature, previous analyses
  - Work from understanding of what typically happens

- Will make research more cost effective – everyone wins!
Thanks!

(and Happy Birthday Harsha!)