An Introduction to Split-Plot Experiments with Application to Bone Tissue Engineering

Byran J. Smucker
Miami University, Department of Statistics

CERM 2019, Midland, MI

Outline

Bioplotter

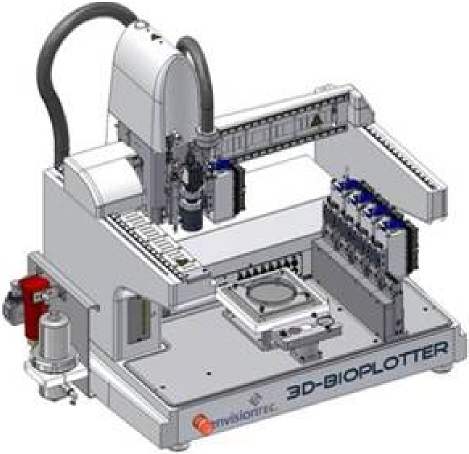

Figure: Via Amy Yousefi

The Bioplotter Makes Scaffolds for Bone Tissue

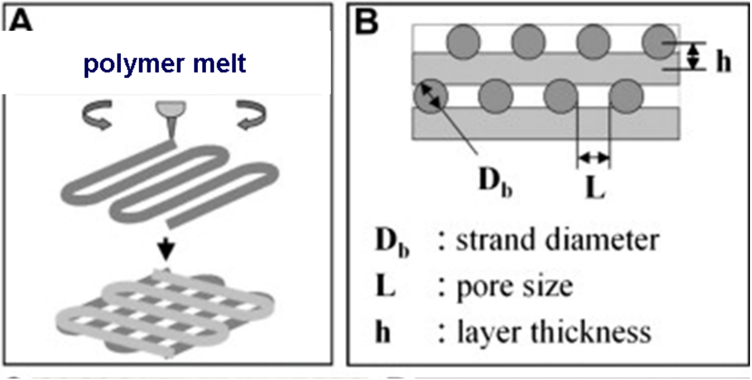

Figure: Via Amy Yousefi

Completely Randomized Experiment

Run	Needle Diameter (mm)	Temperature (°C)	Pressure (Bar)	Dispensing Speed (mm/s)
1	0.6	100	4.5	3
2	0.4	130	3.5	3
3	0.2	115	3.5	2
4	0.6	130	4.5	1
5	0.2	100	3.5	3
6	0.2	130	3.5	1
7	0.2	130	4.5	3
8	0.2	130	2.5	3
9	0.4	100	2.5	2
10	0.6	130	2.5	1
11	0.4	115	4.5	2
12	0.4	115	3.5	1
13	0.4	115	3.5	2
14	0.6	130	3.5	2
15	0.2	100	2.5	1
16	0.2	100	4.5	1
17	0.6	100	3.5	1
18	0.6	115	2.5	3
19	0.6	130	4.5	3
20	0.6	130	2.5	2
21	0.6	130	3.5	3
22	0.6	100	3.5	2

A More Convenient Ordering of the Experiment

Run	Needle Diameter (mm)	Temperature (°C)	Pressure (Bar)	Dispensing Speed (mm/s)
1	0.2	100	3.5	3
2	0.2	100	2.5	1
3	0.2	100	4.5	1
4	0.2	115	3.5	2
5	0.2	130	3.5	1
6	0.2	130	4.5	3
7	0.2	130	2.5	3
8	0.4	100	2.5	2
9	0.4	115	4.5	2
10	0.4	115	3.5	1
11	0.4	115	3.5	2
12	0.4	130	3.5	3
13	0.6	100	4.5	3
14	0.6	100	3.5	1
15	0.6	100	3.5	2
16	0.6	115	2.5	3
17	0.6	130	4.5	1
18	0.6	130	2.5	1
19	0.6	130	3.5	2
20	0.6	130	4.5	3
21	0.6	130	2.5	2
22	0.6	130	3.5	3

But There is a Problem with the Convenient Reordering

Undoes the effect of randomization,

But There is a Problem with the Convenient Reordering

Undoes the effect of randomization, which leads to

- Runs no longer statistically independent (so standard statistical analysis (linear regression) will yield incorrect variance estimates)
- Vulnerability to bias caused by lurking variables

Outline

We All Want Valid Experiments

Here are your options, if you want to do a valid experiment:

We All Want Valid Experiments

Here are your options, if you want to do a valid experiment:

(1) isolate the factor you want to study by controlling *all possible* alternative explanations.

We All Want Valid Experiments

Here are your options, if you want to do a valid experiment:

(1) isolate the factor you want to study by controlling *all possible* alternative explanations.

Or ...

We All Want Valid Experiments

Here are your options, if you want to do a valid experiment:

(1) isolate the factor you want to study by controlling *all possible* alternative explanations.

Or ...

(2) *randomize* the order of the runs in the experiment (or more generally, randomize the experimental units to the treatments), which allows the effects of alternative explanations to be canceled out

Randomization

"Nature ... will best respond to a logical and carefully thought out questionnaire; indeed, if we ask her a single question, she will often refuse to answer until some other topic has been discussed." R.A. Fisher

Randomization

"Nature ... will best respond to a logical and carefully thought out questionnaire; indeed, if we ask her a single question, she will often refuse to answer until some other topic has been discussed." R.A. Fisher

So many possible "other topics" to discuss, so Fisher's clever solution was to randomize so the other topics become irrelevant.

Randomization

"Nature ... will best respond to a logical and carefully thought out questionnaire; indeed, if we ask her a single question, she will often refuse to answer until some other topic has been discussed." R.A. Fisher

So many possible "other topics" to discuss, so Fisher's clever solution was to randomize so the other topics become irrelevant.

To recap: two implications of randomization:

1. Reduce/eliminate bias and thus allow simpler identification of causal effects
2. Make statistical independence a plausible assumption

Outline

Simple Example

TABLE 1. Split-Plot Design and Data for Studying the Corrosion Resistance of Steel Bars (Box et al. (2005))

Whole-plot	Temperature (°C)	Coating (randomized order)			
1.	360	C_2 73	C_3 83	C_1 67	C_4 89
2.	370	C_1 65	C_3 87	C_4 86	C_2 91
3.	380	C_3 147	C_1 155	C_2 127	C_4 212
4.	380	C_4 153	C_3 90	C_2 100	C_1 108
5.	370	C_4 150	C_1 140	C_3 121	C_2 142
6.	360	C_1 33	C_4 54	C_2 8	C_3 46

From Jones and Nachtsheim (2009), "Split-Plot Designs: What, Why, and How", *Journal of Quality Technology*

Two Types of Factors so Two Randomizations

hard-to-change (whole plot): you don't want to change the levels of these factors for every run;

Two Types of Factors so Two Randomizations

hard-to-change (whole plot): you don't want to change the levels of these factors for every run; (needle diameter and temperature)

Two Types of Factors so Two Randomizations

hard-to-change (whole plot): you don't want to change the levels of these factors for every run; (needle diameter and temperature)

easy-to-change (split plot): no problem changing the levels of these factors for every run;

Two Types of Factors so Two Randomizations

hard-to-change (whole plot): you don't want to change the levels of these factors for every run; (needle diameter and temperature)

easy-to-change (split plot): no problem changing the levels of these factors for every run; (pressure and dispensing speed)

Two Types of Factors so Two Randomizations

hard-to-change (whole plot): you don't want to change the levels of these factors for every run; (needle diameter and temperature)

easy-to-change (split plot): no problem changing the levels of these factors for every run; (pressure and dispensing speed)

Randomize the hard-to-change treatment combinations to the whole plots and randomize the easy-to-change treatment combinations to the subplots.

Constructing a Split-Plot Experiment

Simple case: construct a completely randomized design for both the whole plot and split-plot factors.

Constructing a Split-Plot Experiment

Simple case: construct a completely randomized design for both the whole plot and split-plot factors.

More realistic case: use software to construct the design (google "split-plot design [software]", where "[software]" might be "Minitab", "Design Expert", "JMP", or others)

Constructing a Split-Plot Experiment

Simple case: construct a completely randomized design for both the whole plot and split-plot factors.

More realistic case: use software to construct the design (google “split-plot design [software]”, where “[software]” might be “Minitab”, “Design Expert”, “JMP”, or others)

Ideal case: collaborate with a statistician

The Bioplotter Design

Whole Plots	Diameter	Temperature	Speed	Pressure
1	0.6	100	1	3.5
1	0.6	100	3	2.5
1	0.6	100	3	4.5
2	0.4	115	2	3.5
2	0.4	115	3	4.5
2	0.4	115	2	3.5
3	0.6	130	1	2.5
3	0.6	130	3	3.5
3	0.6	130	1	4.5
4	0.2	115	2	4.5
4	0.2	115	1	3.5
4	0.2	115	3	2.5
5	0.4	115	1	2.5
5	0.4	115	2	3.5
5	0.4	115	2	3.5
6	0.2	100	1	4.5
6	0.2	100	3	3.5
6	0.2	100	1	2.5
7	0.2	130	3	4.5
7	0.2	130	2	2.5
7	0.2	130	1	3.5
8	0.4	130	1	3.5
8	0.4	130	3	2.5
8	0.4	130	2	4.5

Analyzing a Split-Plot Experiment

Since two randomizations, we need two error terms!

Analyzing a Split-Plot Experiment

Since two randomizations, we need two error terms!

Regression model (non-split-plot) for i^{th} experimental run:

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \epsilon_i$$

Analyzing a Split-Plot Experiment

Since two randomizations, we need two error terms!

Regression model (non-split-plot) for i^{th} experimental run:

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \epsilon_i$$

Regression model (split-plot) for i^{th} experimental run, from the j^{th} whole plot:

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \gamma_j + \epsilon_i$$

Analyzing a Split-Plot Experiment

Since two randomizations, we need two error terms!

Regression model (non-split-plot) for i^{th} experimental run:

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \epsilon_i$$

Regression model (split-plot) for i^{th} experimental run, from the j^{th} whole plot:

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \gamma_j + \epsilon_i$$

So to analyze a split-plot experiment, you use basic regression but add a "random effect" associated with the whole plot.

Analyzing a Split-Plot Experiment

Since two randomizations, we need two error terms!

Regression model (non-split-plot) for i^{th} experimental run:

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \epsilon_i$$

Regression model (split-plot) for i^{th} experimental run, from the j^{th} whole plot:

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \gamma_j + \epsilon_i$$

So to analyze a split-plot experiment, you use basic regression but add a "random effect" associated with the whole plot.

Use online resources from software companies or a consultation with a statistician to ensure correct analysis.

Final Thoughts

Take-aways:

1. Randomize your experiment, unless you are willing to believe (and defend) that you have eliminated *every* possible alternative causal explanation
2. If you can't completely randomize because some factors are hard-to-change, design a split-plot experiment
3. Use software resources or statistics experts to construct and analyze split-plot experiments

Contact Information

Byran Smucker
Affiliation: Miami University, Department of Statistics
E-mail: smuckerb@miamioh.edu
Twitter: @ByranSmucker