

Section 5: Randomization and the Basic Factorial Design

William Christensen



Randomization:

Whenever possible, use a chance device for any assigning or sampling. This applies not only to assigning treatments, but also to any other parts of the experimental protocol (detailed instructions or plans for executing an experiment).

EX Compare 10 rats eating beef-heavy diet with 10 rats eating tofu-heavy diet. What are things we will want to randomize (in addition to group assignment)?



Why randomize? (Review)

- 1 Protect against bias
- 2 Allows us to use probability and sampling distributions when analyzing the data.
 - ...because the errors will exhibit *chance-like* behavior

Balanced Designs

Balance refers to the presence of equal treatment group size:

EX BF[2]

EXERCISE

		low	moderate	intense
DIET	low-cal.	10 subj	10 subj	10 subj
	high-cal.	10 subj	10 subj	10 subj

- Often “unbalanced” is used to refer to the factor structure of an experiment.

EX Popcorn experiment with 5 treatments.

		“Brand of Oil”	
		No-name	Orville
“Amount of Oil”	0 tbsp	10 batches	
	1 tbsp	10 batches	10 batches
	2 tbsp	10 batches	10 batches

- Of course, we can always treat the above design as a balanced BF[1] with 5 levels of our factor called “Oil Contribution”
 - No oil
 - 1 tbsp no-name
 - 1 tbsp Orville
 - 2 tbsp no-name
 - 2 tbsp Orville

BF Designs

The experimental version of the BF is usually called a *completely randomized design* (CRD).

For the observational version of the BF, take a *simple random sample* (SRS) in each of the populations of interest.

- Def’n: A simple random sample (SRS) of size n is a randomly chosen subset of the population such that all possible samples of size n are equally likely.
 - Note that this condition implies that every member of the population has the same chance of being in the sample... but our condition above is a bit stronger.

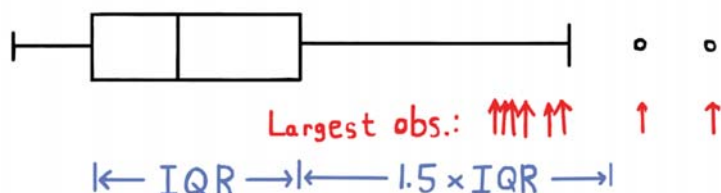
- Although we can use the BF structure and ANOVA to analyze observational data from non-random (convenience) samples, bias will taint the results.

Exploratory Data Analysis (EDA)

Usually advisable to begin our analysis with simple summaries and exploratory data analysis:

- Group means
- Plot data by group, looking for
 - group differences
 - outliers?
 - equal variances?
 - normality?

- I like to use *boxplots* to evaluate these issues



Note:

- “Whisker” reaches to observation that is most extreme among those within 1.5 IQR’s of the edge of the box, with dots (“o”) denoting outliers.

EX Make a boxplot for 100 random draws from a $N(0,1)$ distribution.

EX EDA for fish data

Definitions and Rules for Analysis

		Low Salt	High Salt
Canola {	Orville	36	63
	No-Name	47	67
Buttery {	Orville	28	81
	No-Name	42	85

Partition: A partition of the observations is a way of sorting them into groups.

[Note: we previously defined “factor” as a variable under examination in an experiment as a possible cause of variation. Below we give a definition for a different usage.]

Factor: a meaningful partition of the observations

Factors in BF [1] design:

Universal factor* $\rightarrow$ (1) Grand mean (benchmark)

Universal factor* $\rightarrow$ (2) Residual error

Structural factor $\rightarrow$ (3) **Treatment factor**

*These two universal factors occur in all designs

Example: Flavor scores for 3 methods of cooking fish (4 fish cooked per method)

	Fish			
	1	2	3	4
Method 1 (fry)	5.4	5.2	6.1	4.8
Method 2 (bake)	5.0	4.8	3.9	4.0
Method 3 (grill)	4.8	5.4	4.9	5.7

Notation for BF[1]

$$x_{ij} = \mu + \alpha_i + \varepsilon_{ij} \quad i = 1, \dots, I \quad j = 1, \dots, J$$

↑ ↑

I=3 for fish data J=4 for fish data

For fish data:

$$x_{23} = \underline{\hspace{2cm}} \quad x_{12} = \underline{\hspace{2cm}} \quad x_{41} = \underline{\hspace{2cm}}$$

Decomposition of Balanced BF[1] Design:

Estimated benchmark = Grand mean
= $\bar{x}_{..}$

“..” subscript indicates that we average over all values of i and j (total of IJ obs.)

Estimated Treatment Effect = Treatment mean – Grand mean

$$= \bar{x}_{i.} - \bar{x}_{..}$$

“*i*.” subscript indicates that we choose an *i*
and average over *j* (total of *J* obs.)

$$\begin{aligned}\text{Residual Error} &= \text{Observed value} - \text{Treatment mean} \\ &= x_{ij} - \bar{x}_i.\end{aligned}$$

General decomposition rule

ex fish data

Observations	Benchmark (Grand Mean)	Method effect	Residuals
5.4 5.2 6.1 4.8	5 5 5 5	.375 .375 .375 .375	.025 -.175 .725 -.575
5.0 4.8 3.9 4.0	5 5 5 5	-.575 -.575 -.575 -.575	.575 .375 <u>-.525</u> -.425
4.8 5.4 4.9 5.7	5 5 5 5	.2 .2 .2 .2	-.4 .2 -.3 .5

Definition: One factor is inside another if each group of the first (inside) factor fits completely in the second (outside) factor.

- ? Is “method” inside “grand mean”?
- ? Is “residual error” inside “grand mean”?
- ? Is “residual error” inside “method”?

General rules for calculating estimated effect and df

- Estimated effect for a factor level
= (Average for the factor level)
– (sum of effects for all outside factors)
- df for a factor
= (number of levels for the factor)
– (sum of df for all outside factors)

Note: for BF[1]

- $df_{\text{Grand Mean}} = 1$
- $df_{\text{treatment}} = (\# \text{ of levels}) - 1$
- $df_{\text{residuals}} = \underset{\substack{\uparrow \\ \text{total \# of observations}}}{N} - (\# \text{ of treatment levels})$

ex fish data

Observations	Benchmark (Grand Mean)	Method effect	Residuals
5.4 5.2 6.1 4.8	5 5 5 5	.375 .375 .375 .375	.025 -.175 .725 -.575
5.0 4.8 3.9 4.0	5 5 5 5	-.575 -.575 -.575 -.575	.575 .375 <u>-525</u> -.425
4.8 5.4 4.9 5.7	5 5 5 5	.2 .2 .2 .2	-.4 .2 -.3 .5

Est'd Effect for method 2 (recall "method" is inside "benchmark")

$$= \underset{\substack{\uparrow \\ \text{avg for method 2}}}{4.425} - \underset{\substack{\uparrow \\ \text{benchmark effect}}}{5}$$

$$\text{df method} = \underset{\substack{\uparrow \\ \text{\# of levels for method}}}{3} - \underset{\substack{\uparrow \\ \text{df for benchmark}}}{1} = 2$$

Est'd Residual effect $\hat{\epsilon}_{23}$

$$= \underset{\substack{\uparrow \\ \text{3rd obs.} \\ \text{on 2nd method}}}{3.9} - \left(\underset{\substack{\uparrow \\ \text{benchmark} \\ \text{effect}}}{5} + \underset{\substack{\uparrow \\ \text{method effect} \\ \text{for method \#2}}}{(-.575)} \right)$$

$$\text{df residual} = \underset{\substack{\uparrow \\ N}}{12} - \left(\underset{\substack{\uparrow \\ \text{df benchmark}}}{1} + \underset{\substack{\uparrow \\ \text{df method}}}{2} \right) = 9$$

14/26

ex Popcorn Data

(A B means "B is inside A")

Observed Values	Grand Mean	Salt	Oil	Popcorn Brand	Error
36 63	56.125 56.125	-17.875 17.875	-2.875 -2.875	-4.125 -4.125	4.75 -4
47 67	56.125 56.125	-17.875 17.875	-2.875 -2.875	4.125 4.125	7.5 -8.25
28 81	56.125 56.125	-17.875 17.875	2.875 2.875	-4.125 -4.125	-9 8.25
42 85	56.125 56.125	-17.875 17.875	2.875 2.875	4.125 4.125	-3.25 4

model:

$$Y_{ijk} = \underset{\substack{\uparrow \\ \text{Grand} \\ \text{Mean}}}{\mu} + \underset{\substack{\uparrow \\ \text{Salt} \\ \text{effect} \\ i = 1, 2}}{\alpha_i} + \underset{\substack{\uparrow \\ \text{Oil} \\ \text{effect} \\ j = 1, 2}}{\beta_j} + \underset{\substack{\uparrow \\ \text{Brand} \\ \text{effect} \\ k = 1, 2}}{\gamma_k} + \epsilon_{ijk}$$

15/26

Est'd effect for buttery oil

$$= \underset{\substack{\uparrow \\ \text{avg for buttery}}}{59} - \underset{\substack{\uparrow \\ \text{benchmark effect}}}{56.125} = 2.875$$

$$df_{\text{oil}} = \underset{\substack{\uparrow \\ \text{levels for oil}}}{2} - \underset{\substack{\uparrow \\ \text{df for benchmark}}}{1}$$

	Means	Effect Size
overall:	56.125	—
low salt:	38.25	} +/- 17.875
high salt:	74	
canola:	53.25	} +/- 2.875
buttery:	59	
orville:	52	} +/- 4.125
no-name:	60.25	

Est'd Residual effect for ε_{212}

high salt, canola oil, no-name brand

$$= \underset{\substack{\uparrow \\ y_{212} \text{ value}}}{67} - (\underset{\substack{\uparrow \\ \text{benchmark}}}{56.125} + \underset{\substack{\uparrow \\ \text{salt}}}{17.875} + \underset{\substack{\uparrow \\ \text{oil}}}{(-2.875)} + \underset{\substack{\uparrow \\ \text{brand}}}{4.125}) = -8.25$$

$$df_{\text{residual}} = 8 - (1 + 1 + 1 + 1) = 4$$

EX Extended example of a BF[1]:

Effect of male fruitfly reproductive behavior (Nature, 1981)

/*****

Columns	Description
1- 2	Serial No. (1-25) within each group of (the order in which data points were abstracted)
4	Number of companions (0, 1, or 2)
6	Type of companion 0: newly pregnant female 1: virgin female 9: not applicable (when PARTNERS=0)
8- 9	Lifespan, in days
11-14	Length of thorax, in mm (x.xx)
16-17	Percentage of each day spent sleeping

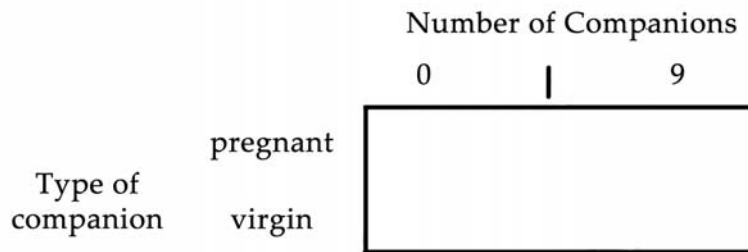
*****/

n=25 per treatment group

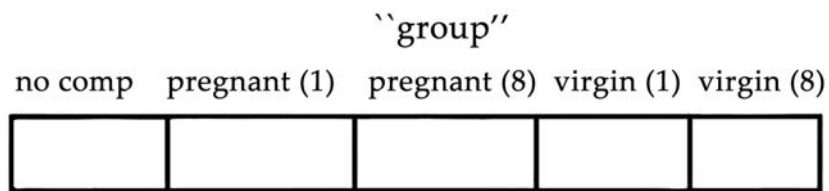
Two factors

? Is this a balanced two-factor design?

Two-factor design:



One-factor design:

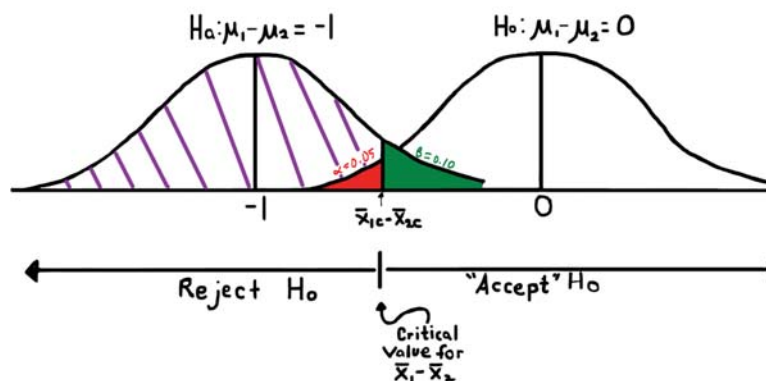


$$H_o : \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$$

H_a : at least one μ_i is different from others

Power and Sample Size for 1-Way ANOVA

Recall that in the 2-group case, as $\mu_1 - \mu_2$ moves away from 0, the **power** increases



Similar idea for ANOVA:

Under $H_0 : \mu_1 = \mu_2 = \dots = \mu_I \quad i = 1, \dots, I$
 (or $H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_I = 0$)

$$F^* = \frac{MS_{\text{group}}}{MS_{\text{error}}} \sim F_{I-1, N-I}$$

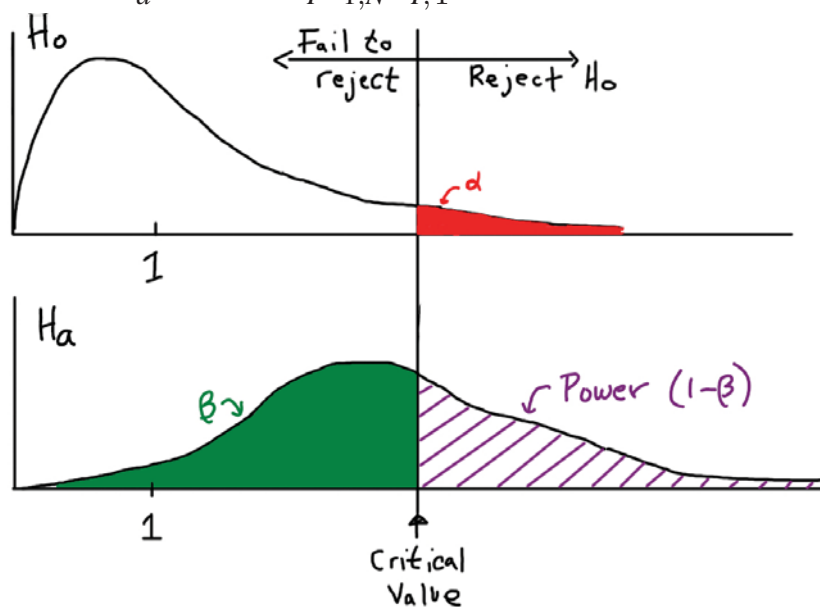
Note: When H_0 is true $\Phi^2 = \frac{n \sum_{i=1}^I \alpha_i^2}{I\sigma^2} = 0$

but as $\Phi^2 = \frac{n \sum_{i=1}^I \alpha_i^2}{I\sigma^2}$ gets larger, power increases

- Φ is the “non-centrality parameter” and governs the shape of the distribution of F^* under H_a

Under $H_0 : F^* \sim F_{I-1, N-I} \leftarrow \text{“F dist.” } (\Phi=0)$

Under $H_a : F^* \sim F_{I-1, N-I, \Phi} \leftarrow \text{“non-central F dist.”}$



$$\text{Power} = \Pr\{F^* > F_{\alpha, I-1, N-I} \mid \Phi\}$$

To find power, we need:

- 1 α (significance level)
- 2 n (observations per group)
- 3 I (number of groups)
- 4 values for $\alpha_1, \alpha_2, \dots, \alpha_I$
- 5 estimate of σ^2

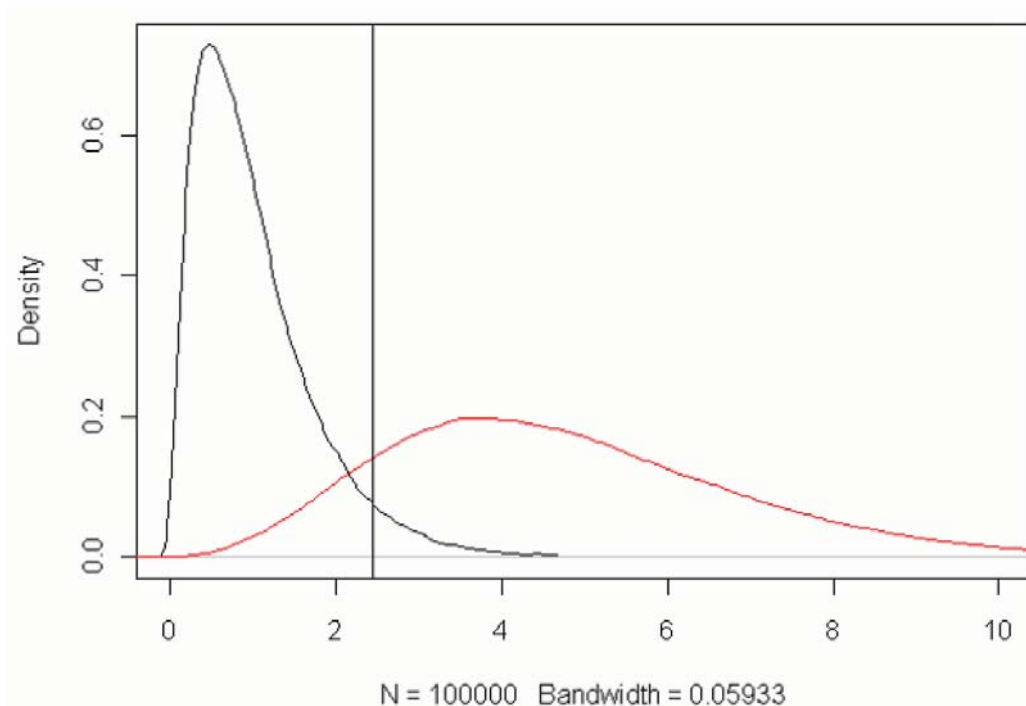
EX Suppose we are planning the fruitfly study.

- $\alpha = 0.05$
- planning on $n = 25$ flies per group
- $I = 5$ groups
- Would like to consider power for the case where four of the groups have the same $\alpha_i = c$, $i = (1, 2, 3, 4)$ and one group has $\alpha_5 = c + 10$.
 - Since $\sum \alpha_i = 0$
$$4c + (c + 10) = 0$$
$$\Rightarrow c = -2$$
$$\Rightarrow \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = -2 \text{ and } \alpha_5 = 8$$
 - Assume that nearly all ($\approx 99.7\%$) of flies live between 30 and 100 days.
$$\Rightarrow \sigma \approx 11.67 \text{ days} \quad \leftarrow \text{why } 11.67?$$

$$\Phi^2 = \frac{n \sum_{i=1}^I \alpha_i^2}{I\sigma^2} = \frac{25(4(-2)^2 + 8^2)}{5(11.67)^2} = 2.937$$

```
> n <- 25
> siglevel <- 0.05
> numgroups <- 5
> alphai <- c(-2,-2,-2,-2,8)
> sigma <- 11.67
> Phi2 <- n * sum(alphai^2) / (numgroups * sigma^2)
> Phi2
[1] 2.937097
> fcrit <- qf(1 - siglevel, numgroups-1, numgroups*n - numgroups)
> fcrit
[1] 2.447237
> power <- 1 - pf(fcrit, numgroups-1, numgroups*n - numgroups,
  numgroups*Phi2) # notice that the non-centrality parameter
  # required by R is numgroups*Phi2
> power
[1] 0.8694991
```

F(4, 120)[black] and F(4, 120, Phi)[red]



Choosing n:

We could plug in different values of n into the approach above, until we have the n needed for the desired power.

—or—

Use “power.anova.test” in R

```
> ## Give n and leave power out of list of arguments:
    gives you the power
> power.anova.test(groups=5,
  between.var=var(c(0,0,0,0,10)),
  within.var=11.67^2, sig.level=0.05, n=25)
      power = 0.8694991
> ## Give power and leave n out of list of arguments:
    gives you n
> power.anova.test(groups=5,
  between.var=var(c(0,0,0,0,10)),
  within.var=11.67^2, sig.level=0.05, power=.869)
```