Objectives

Random vs. Fixed Definitions

Deciding whether to treat a factor as fixed vs. random in practice.

Implications of fixed vs. random on the model

Identify when to treat a factor as random vs. fixed in an experiment

Give examples of a 1 way random factor (typing speed, ammunition(?),

Write the model of a 1 way random factor anova, completely randomized vs. the fixed effects version of the model

Go both ways: workers constantly coming and going vs. stable staff

Random Factors

In the designs we have dealt with so far, we have been primarily focused on the effect of a treatment. These treatment factors have been fixed effects. The term “fixed” is meant to reinforce the fact that the levels of these factors is what we are interested in studying. The levels of these factors are chosen deliberately. They represent the population of factor levels we would like to study.

In contrast, levels of a random factor represent a subset of all possible levels the factor can take on. We would like to observe all levels of the factor, but that is impossible to do because there are too many levels. So a sample of the factor levels is taken. Actually, at least one random factor has been present in every design we have worked with: the residual factor is a random factor. The residual factor represents the effect of each observational unit. The units of a study generally represent a (random) sample from the entire population.

When a random factor is included in a study, the researcher is typically interested in measuring the variability in the random factor, rather than estimating the effect of each level in the study. The residual factor is a good example. The mean effect of each individual is not of interest, rather estimating the mean squared error is the focus.

\*\*Call out\*\*

Definition of fixed and random factors

\*\*

Deciding whether a factor should be treated as fixed or random depends on the context and purposes of the research. Consider the following example, which illustrates how the same factor could be treated as fixed or random, depending on the situation.

A researcher was interested in estimating the mean typing speed of college freshman at her university. A random, representative sample of 300 freshman was obtained. The researcher plans to ask each freshman to type a passage that 200 characters long. Her colleague points out the mean typing speed could be highly dependent on the difficulty of the passage freshman are asked to type. The researcher uses artificial intelligence (AI) to create 10 different passages. The 10 passages represent a sample of an infinite number of passages that could have been created. Each freshman is randomly assigned one of the 10 passages to type.

The researcher is not interested in the effect of any one passage in particular, just like they are not interested in any one student’s typing speed. Rather, the researcher can estimate how much variability in a person’s typing speed is due to variability in the passage being typed. The null hypothesis in this case is that the variance in typing speed between the passages is zero.

In the example above, the passages were created without consideration of content. However, if each of the 10 passages was created to represent a specific type of passage, then passage should be treated as a fixed variable. For example, one passage may incorporate many numbers. Another could have been created to include lots of punctuation. Another could include lots of unusual letters such as ‘q’ and ‘z’. Yet another passage might include a lot of capitalization, and so on. When each passage is deliberately meant to represent a particular category of interest, it may be more appropriate to treat the passage factor as fixed.

Consider another example that illustrates the difference between fixed and random factors.

A manufacturer of widgets is interested in improving the quality of its products. The operator of the machine that makes the product is a major source of potential variation. The manufacturing company has designed an experiment that will gather widget quality data made by 5 different machine operators.

*Machine operator as a fixed factor*: There are only a handful of people at the company that operate the machine. They have been with the company for quite some time. If this is the case, treating the machine operator factor as fixed makes a lot of sense. The 5 levels of machine operator may represent the entire population (or nearly so) of operators. Knowing which operator is producing sub-par quality would be important so that training could be provided to that individual.

Each machine operator will have their own quality distribution for the widgets they made. The picture below shows the quality distribution for the 4 operators in the study. In this case, we would gather a few observations from each of the distributions. We could use the data to estimate the true mean and effect of each operator.

PICTURE

*Machine operator as a random factor*: There is quite a lot of churn at the machine operator position. No one person stays for very long. In this situation, 5 operators would represent a small sample from the population of machine operators that have previously worked, currently work, or may work the machine in the future. If the factor for machine operator turns out to be significant, the company may improve overall processes or hiring practices, rather than identifying individual machine operators to train.

The picture below illustrates that there are many operators in the population, each with their own distribution of widget quality. However, not all of them can be observed. We will use observations from 4 randomly selected distributions (shown in color) to estimate the variance in machine operator means.

PICTURE

Implications of Including a Random Effect

In the picture above, where machine operator is treated as a random factor, it can be seen that there is more variability in the overall distribution of the response than can be observed from looking at just 4 operators. There are essentially 2 levels of sampling: first the operator is sampled, second the widgets produced by that operator will be sampled. The extra random sampling step of choosing operators adds uncertainty, i.e. increased standard error, in parameter estimates. In the typing example there are also two samples contributing variability: sampling of the freshman, and a sample of passages to be typed.

This increased variability can be seen when comparing the mathematical expression of the models. If machine operator is treated as a *fixed* factor, the fixed effects model can be written as

See page 36 of orange

The observed responses will have a mean of mu and a variance of sigmasquared.

If machine operator is treated as a *random* effect, the model is nearly identical:

See page 598

The key difference is that the machine operator effects are now treated as an observations from a random variable, which has its own mean (zero) and standard deviation (sigma\_T). MARGIN?: Note that unlike the fixed effect model, where the sum of the factor level effects (alpha\_i) summed to zero, the sum of the observed random effects, A\_i, do not necessarily sum to zero.

Since both random effects, A and sigma, have a mean of zero, the mean of y is still mu. However, the variance of y needs to incorporate the variance contributed by both random effects, sigma\_y = sigma\_A + sigma\_epsilon. MARGIN: This assumes that epsilon and alpha are mutually independent, in other words, their covariance is zero.

The null hypothesis for a random factor is that the variance of the factor level means is zero and can be written as

Ho: sigma\_A = 0.

**How to Decide**

In reality, it can be hard to determine whether a factor should be treated as fixed or random. Here are some rules of thumb that may help you make the distinction of whether to treat a variable as random or fixed. (Cobb, p. 561)

1. Observational units (i.e. the residual factor) should always be random
2. Blocks, and other nuisance factors, are usually random
3. Factors nested inside of another factor is usually random
4. Experimental factors (factors that are purposefully manipulated in an experiment) are usually fixed

Still having trouble deciding? Here are a few additional questions you can ask yourself to decide. There is always an exception, but these are good rules of thumb to follow: <https://dynamicecology.wordpress.com/2015/11/04/is-it-a-fixed-or-random-effect/>, and Cobb (p. 559).

* Do the levels represent a tiny fraction of all possible levels (random), or a nearly exhaustive list of all possible levels (fixed)?
* Is the underlying nature of the variable continuous? If so, it should be treated as fixed. The reason being is that you most likely chose levels to span a range of reasonable values on the continuum, the levels were (hopefully) not chosen at random.
  + If the variable is put into the model as a continuous variable (not factor levels), then it is also a fixed variable.
* Are there 4 or fewer levels? If so, it often makes sense to treat the variable as fixed, even if it is truly random. This is because estimating a variance from a sample of just 3 or 4 is not very reliable and it greatly complicates the model.
* Are you interested in reporting the mean of each level (fixed) or the variance between levels (random)?
* If you repeated the experiment, would/could the levels of the factor stay the same (fixed), or change (random)?

**What about interactions?**

When a model has fixed and random factors it is called a “mixed model”. An example of this is the SP/RM model. If two random factors are crossed, their interaction is treated as a random factor. If two fixed factors are crossed, their interaction is treated as a fixed factor. If a random factor and a fixed factor are crossed, the result can be called a mixed factor. Dealing with mixed factors is outside the scope of this book. Crossing a fixed and a random factor is outside the scope of this book.

Revisit the F Test

On this page the F test is explained as a ratio of variances. Specifically, the ratio variance between factor level means to the variance of individual observations within factor level means. We can break this ratio down a bit further. In a simple one-factor ANOVA, there are actually two components contributing to the variance between factor level means: the variance due to different treatments and the variance due to individuals (which is also referred to as residual error). Recognizing that “Mean Squares” is synonymous with “variance”, we can write the following to express this concept:

MStreatment = treatment effects + residual error

The denominator in our F statistic has traditionally been the Mean Squared error (or residual error). Thus, the F statistic amounts to (treatment effects + residual error) / residual error. When the treatment effects are small or insignificant, we get **expect** the F statistic close to be (0 + residual error) / residual error = residual error / residual error = 1.

In this calculation, the only difference between the numerator and the denominator is the variance due to treatment. Thus, the F statistic ratio is how we isolate the error that is coming from the factor of interest. Our approach is to construct the F statistic such that the only difference in the numerator and denominator is the addition to the numerator of the factor of interest.

Of course, due to random sampling, the F statistic will not always be exactly 1 even if the effect of a factor truly was zero. However, if the true effect was zero and we were able to observe all possible samples (in some fantasy, hypothetical world), the mean of the F statistics would be 1. This notion of resampling and getting different statistics is called a sampling distribution. Each variance estimate has its own sampling distribution (just like a sample mean has a sampling distribution…see Math221 chapter 6). The term Expected Mean Square is used to refer to the mean of a sampling distribution of the variance. Rather than calculate a value for the expected mean square, our goal is simply to identify factors of a design that contribute to a factor’s EMS. This will help our understanding of the appropriate F statistic calculation.

The pieces of a model that contribute to a factor’s expected mean square (EMS) are the treatment factor itself and all **random, inside** factors. This means there will be cases when the denominator of the F statistic is not simply the Mean Squared Error (MSE).

Let’s consider 3 different examples:

* A completely randomized two-way ANOVA, where both experimental factors are fixed
* A completely randomized two-way ANOVA, where both experimental factors are random
* A SP/RM[1,1] design, where blocks are random.

Example 1: BF[2] where both factors are fixed

This example is identical to the analysis presented on the BF[2] page. We can build a table to verify that we are calculating the F statistic correctly. The first column lists each of the factors in the design. The second column shows the EMS, or in other words, what factors are contributing to the observed variability in that factor.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | EMS | F stat | Denominator of F stat |  |  |  |
| Grand Mean | Grand Mean + E | Grand Mean + E / E | Mean Squared error |  |  |  |
| Factor A | A + E | A + E / E | Mean Squared error |  |  |  |
| Factor B | B + E | B + E / E | Mean Squared error |  |  |  |
| A x B | AB + E | AB + E | Mean Squared error |  |  |  |
| Error term (E) (random) | E | - |  |  |  |  |

Notice that in every F statistic the denominator is just the error term. Because all the factors are fixed, the only other source of variation contributing to that factor’s variance estimate is the error term. When both factors are random, this is not the case.

Example 2: Both Factors are Random

Remember, the expected mean square (EMS) is composed of the treatment itself and all **random, inside** factors

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | EMS | F stat | Denominator of F Stat |  |  |  |
| Grand Mean | Grand Mean + A + B + AB + E | NA | NA |  |  |  |
| Factor A (random) | A + AB + E | (A + AB +E) / (AB + E) | Mean Squared for AB |  |  |  |
| Factor B (random) | B + AB +E | B + AB + E / (AB + E) | Mean Squared for AB |  |  |  |
| A x B (random) | AB + E | AB + E / E | Mean Squared for E |  |  |  |
| Error term (E) (random) | E | - |  |  |  |  |

First, notice that there is no appropriate test for the grand mean. To correctly isolate the effect of the grand mean in the F test numerator, we would need a factor with an expected mean square of A + B + AB + E to put in the denominator. None of the factors have that for their EMS, and therefore the F test for grand mean cannot be done. This is not too concerning though because we generally already assume the value is not zero, or simply don’t care.

The F test for Factor A and Factor B both use the Mean Squares of AB interaction in the denominator of the F test. This is needed so that the only difference between the numerator and denominator of the F statistic is the contribution of the factor being tested (in the numerator). The interaction term, AB, uses the Mean Squared Error for its denominator for the same reason.

Example 3: Split Plot/Repeated Measures

Let’s look at a bit more complex (and realistic) scenario.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | EMS | F stat | Denominator of F Stat |  |  |  |
| Grand Mean | Grand Mean + S + E | (Grand Mean + S + E) / (S + E) | MS-S |  |  |  |
| Factor A | A + S + E | (A + S +E) / (S+ E) | MSS |  |  |  |
| Block Factor, S (random) | S + E | S + E / E | MSE |  |  |  |
| Factor B | B + E | B + E / (E) | MSE |  |  |  |
| A x B | AB + E | AB + E / E | MSE |  |  |  |
| Error term (E) | E | - |  |  |  |  |

In this hierarchical (two levels of experimental units), mixed model, the mean squares for block (S) is used as the denominator for the parts of the model that are at the “higher” level of the hierarchy. In other words, the blocks are the experimental units to which Factor A is applied, and so their Mean squares (i.e. variance) is the denominator for the F test. The block factor and everything else uses the Mean Squared error in the denominator of their F test. Unless you are using a software/package designed for mixed models, the software may output the incorrect F statistic by default.

These theoretical definitions are helpful

Visually

Blocks are also random factors usually. This is because often the blocks are subjects, or the blocks represent a subset of all possible blocks that could have been created

I like the picture of the fixed effects, showing 3 distributions – where observatios are drawn from the 3 distributions, and then fixed effects shows that there are many distributions and then I randomly select which distributions I will be drawing observations from.

I could talk about my experience with the DBL experiment. I have multiple sections and each student received a treatment. I could quantify different levels to describe the sections (fixed effects), or I could simply put section in as a random factor because each section is one possible way to group students that could have happened (but didn’t).

We have actually been dealing with random factors in all the previous designs, we just didn’t refer to them that way.