

Model II (or random effects) one-way ANOVA:

As noted earlier, if we have a random effects model, the treatments are chosen from a larger population of treatments;

we wish to generalize to this larger population.

Thus, experimental replications do not need to have the same treatments administered;

different samples from the population of treatment could be used.

Model II (random effects) analysis-of-variance model:

$$Y_{ij} = \mu + a_i + \epsilon_{ij} ,$$

where Y_{ij} is the j th observation in the i th group, μ is the overall mean, a_i is a random variable indicating an effect shared by all values in group i , and ϵ_{ij} is a random variable representing error.

The two random variables, a_i and ϵ_{ij} , are assumed uncorrelated within and between themselves with expected values of zero, and variances of σ_a^2 and σ_ϵ^2 , respectively.

Also, we will only deal with equal sample sizes; each group is assumed to have n observations:

thus, $1 \leq i \leq r$ and $1 \leq j \leq n$ and the total sample size, $n_T = nr$

Computationally, the same Sum-of-Squares can be obtained as in the fixed-effects model, i.e., Sum-of-Squares Treatment (SSTR); Sum-of-Squares Error (SSE); Sum-of-Squares Total (SSTO)

The null hypothesis of interest is $H_0 : \sigma_a^2 = 0$, i.e., in the population of possible effects, the variance is zero

We can show the following:

$E(MSE) = \sigma_\epsilon^2$; so, MSE is an unbiased estimate of σ_ϵ (always)

$E(MSTR) = n\sigma_a^2 + \sigma_\epsilon^2$; so, MSTR is an unbiased estimate of σ_ϵ when $\sigma_a^2 = 0$

So, the null hypothesis, $H_0 : \sigma_a^2 = 0$, can be tested as in Model I ANOVA:

$$\frac{MSTR}{MSE} \sim F_{r-1, r(n-1)}$$

To estimate the variance components, we note that MSE is always an unbiased estimate of σ_ϵ^2 ;

also, for an unbiased estimate of σ_a^2 , we use

$$\frac{MSTR - MSE}{n}$$

Noting that the total variance $Var(Y_{ij})$ is

$$\sigma_a^2 + \sigma_\epsilon^2,$$

we can form the ratio of treatment variance to total variance:

$$\frac{\sigma_a^2}{\sigma_a^2 + \sigma_\epsilon^2} \equiv \rho_I ,$$

which is called the intraclass correlation coefficient.

In the context of Model II, it is a measure of treatment effect strength

To estimate ρ_I , use a ratio:

(Mean Square Treatment – Mean Square Error) divided by (Mean Square Treatment + $(n - 1)$ Mean Square Error).

Later on we will show how to put a confidence interval on ρ_I

The Intraclass correlation (ICC) more generally:

This different type of correlational measure, an intraclass correlation coefficient (ICC), can be used when quantitative measurements are made on units organized into groups, typically of the same size.

It measures how strongly units from the same group resemble each other.

Here, we will emphasize only the case where group sizes are all 2, possibly representing data on a set of N twins, or two raters assessing the same N objects.

The basic idea generalizes, however, to an arbitrary number of units within each group.

Early work on the ICC from R. A. Fisher and his contemporaries conceptualized the problem as follows:

let (x_i, x'_i) , $1 \leq i \leq N$, denote the N pairs of observations (thus, we have N groups with two measurements in each).

The usual correlation coefficient cannot be computed, however, because the order of the measurements within a pair is unknown (and arbitrary).

As an alternative, we first double the number of pairs to $2N$ by including both (x_i, x'_i) and (x'_i, x_i) .

The Pearson correlation is then computed using the $2N$ pairs to obtain an ICC.

Heritability and the ICC:

In studying heritability, we need two central terms:

Phenotype: the manifest characteristics of an organism that result from both the environment and heredity; these characteristics can be anatomical or psychological, and are generally the result of an interaction between the environment and heredity.

Genotype: the fundamental hereditary (genetic) makeup of an organism; as distinguished from (phenotypic) physical appearance.

Based on the random effects model (Model II) for describing a particular phenotype, symbolically we have: $\text{Phenotype}(P) = \text{Genotype}(G) + \text{Environment}(E)$,

or in terms of variances, $\text{Var}(P) = \text{Var}(G) + \text{Var}(E)$, assuming that the covariance between G and E is zero.

The ICC in this case is the heritability coefficient,

$$H^2 = \frac{\text{Var}(G)}{\text{Var}(P)} .$$

As an aside we might note that in the context of Classical Test Theory where an “observed score” is composed of “true score” plus “error” and the latter are uncorrelated, the heritability coefficient is the reliability:

the ratio of “true score variance” to “observed score variance” (Cronbach’s α is a lower-bound, as you will learn in any test theory class)

Heritability estimates are often misinterpreted, even by those who should know better.

In particular, heritability refers to the proportion of variation between individuals in a population influenced by genetic factors.

Thus, because heritability describes the population and not the specific individuals within it, it can lead to an aggregation fallacy when one tries to make an individual-level inference from a heritability estimate.

For example, it is incorrect to say that because the heritability of a personality trait is, say, .6, that therefore 60% of a specific person's personality is inherited from parents and 40% comes from the environment.

The term “variation” in the phrase “phenotypic variation” is important to note.

If a trait has a heritability of .6, it means that of the observed phenotypic variation, 60% is due to genetic variation.

It does not imply that the trait is 60% caused by genetics in a given individual.

Nor does a heritability coefficient imply that any observed differences between groups (for example, a supposed 15 point I.Q. test score difference between blacks and whites) is genetically determined.

As noted explicitly in Stephen Jay Gould's *The Mismeasure of Man* (1996), it is a fallacy to assume that a (high) heritability coefficient allows the inference that differences observed between groups must be genetically caused.

As Gould succinctly states:

“[V]ariation among individuals within a group and differences in mean values between groups are entirely separate phenomena. One item provides no license for speculation about the other” .

For an in-depth and cogent discussion of the distinction between heritability and genetic determination, the reader is referred to Ned Block,

“How Heritability Misleads About Race” (*Cognition*, 1995, 56, 99–128).

Confidence Intervals on ρ_I :

We do this indirectly: first put a confidence interval on $\theta = \frac{\sigma_a^2}{\sigma_\epsilon^2}$

and then use the relation that $\rho_I = \frac{\theta}{1+\theta}$ to put the confidence interval on the intraclass correlation

The following is true all the time (given the Model II assumptions):

$$\frac{MSTR/(\sigma_\epsilon^2 + n\sigma_a^2)}{MSE/\sigma_\epsilon^2} \sim F_{r-1, r(n-1)}$$

This can be re-expressed:

$$\frac{MSTR}{MSE} \left(\frac{1}{1 + n\theta} \right) \sim F_{r-1, r(n-1)}$$

Find a and b so:

$$P\left(a \leq \frac{MSTR}{MSE} \left(\frac{1}{1 + n\theta} \right) \leq b\right) = .95$$

Doing some algebra:

$$\text{Let } x = \frac{1}{n} \left(\frac{MSTR}{bMSE} - 1 \right)$$

and

$$y = \frac{1}{n} \left(\frac{MSTR}{aMSE} - 1 \right)$$

Then, $P(x \leq \theta \leq y)$ and inverting

$$P\left(\frac{x}{1+x} \leq \rho_I \leq \frac{y}{y+1}\right) = 1 - \alpha$$