University of Waterloo W. H. Cherry

RELATIONSHIPS IN STATISTICS: The Protocol for Setting Levels and Interaction

The protocol for setting levels specifies the *values* to be taken by relevant explanatory variate(s) in a comparative Plan; the simplest case is *two* values of *one* focal variate but there is terminology to deal with the complications of more than two values of more than one focal variate. This terminology is used mainly in the context of *experimental* Plans.

- ∗ A **factor** is an explanatory variate; we distinguish an explanatory variate that is:
  - a *focal* variate; – a *non*-focal variate used as a **blocking factor**;
  - a *non*-focal variate whose value is managed for other reasons – see Note 6 overleaf on the lower half of page HL68.2.

  Our concern in this Statistical Highlight #68 is with factor(s) that are *focal* variate(s).
- ∗ Factor **levels** are the set of value(s) assigned to a factor – that is, (usually) the set of values assigned to the (or a) focal variate. Choosing the *values* for levels in the context of a particular investigation may require extra-statistical knowledge.
- ∗ A **treatment** is a *combination* of the levels of the factor(s) applied to a unit [in the sample (or the blocks)].
- ∗ A **run** is part of the Execution stage of an experimental Plan in which all the data are collected for *one* treatment.
- ∗ A **factorial** treatment structure involves *all* combinations of the levels of the (two or more) factors.
- ∗ The **(treatment) effect** of **X** on **Y** (usually) refers to the change in the *average* of **Y** for *unit* change in **X** and:
  - implies the **X**-**Y** relationship is (believed to be) *causal* – a change in **X** *causes* (brings about) a change in **Y**;
  - includes both the *magnitude* and *direction* of the relationship – for example, the *slope* and its *sign* for a *linear* relationship;
  - requires that all non-focal explanatory variates $\mathbf{Z}_i$ hold their (same) values when **X** changes;
  - is defined (the 'true' effect) over the elements of the *respondent population.*
- ∗ **Interaction** of two factors $\mathbf{X}_1$ and $\mathbf{X}_2$ is said to occur when the effect of one factor on a response variate **Y** depends on the level of the other factor. Interaction means the combined effect of two factors is *not* the sum of their individual effects.
  - Interaction is a key concept also in the discussions of Statistical Highlights #69 and #83.

Illustrations of this terminology are:

- ○ Levels of sex as a factor are *female* and *male*;
  the ranges used as levels of (human) age need careful consideration – ranges that are too *narrow* may consume unnecessary resources in attaining adequate replicating, while ranges that are too *broad* may obscure the effect(s) of age.
- ○ In a taste test of different brands of beer, the factor would be *brand of beer* and its levels would be the individual *brands.*
- ○ When there is only *one* focal variate, the treatments are its levels;
  when there are *two* focal variates, $\mathbf{X}_1$ (say) with *two* levels (denoted 1 and 2) and $\mathbf{X}_2$ with *three* levels (denoted A, B, C), there are $2 \times 3 = 6$ treatments (1A, 1B, 1C, 2A, 2B, 2C) in a factorial treatment structure;
  with four factors each at three levels, that are $4 \times 4 \times 4 = 4^3 = 64$ possible treatments.

NOTES: 1. Not all experimental Plans lead to an Execution stage in runs.

- ● Process improvement investigations often *do* – the Execution stage is then a set of runs, one for each treatment.
- ● A clinical trial of a drug usually does *not* involve runs – each participant takes the drug (or a placebo) [*i.e.*, the (two) treatments are applied to elements or units] for the *whole* period of the Execution stage.

When the Execution stage *does* involve runs, equiprobable assigning consists of equiprobable *ordering* of the *runs*, because unblocked, unknown and unmeasured non-focal explanatory variates are considered as being *time-dependent.*

2. Equiprobable assigning of treatments to units may not be feasible in an experimental Plan when one factor has hard-to-alter levels. For example, if pouring temperature (at two levels, say, of 1,450°F and 1,600°F) is a factor in an investigation to improve a process for making iron castings, the temperature of the furnace containing the molten iron cannot easily be altered; it may therefore be necessary to do *consecutively* all the runs at each temperature, instead of having the pouring temperature low or high under equiprobable assigning for each run. This *lack* of probability assigning *in*creases the limitation imposed on an Answer by comparison error.

- ● What is desirable statistically in data-based investigating may also be compromised in process improvement investigations by having to carry out the Execution stage under time pressure while the process continues normal operation; in addition to possible lack of equiprobable assigning, there may be limitations on Answers because:
  - there is not enough time to obtain adequate *replicating*;
  - the data reflect process operation only over a *limited* time period.

For a process with an *un*acceptably-high long-term scrap rate undergoing an investigation to try to reduce the rate, there have been instances of negative reaction from management to an investigation with an experimental Plan where some treatment(s) involve factor levels that would (temporarily) *in*crease the scrap rate.

3. In ordinary English, interaction customarily involves *two* entities; in statistics, *three* (or more) variates are involved – two (or more) focal variates and one response variate.

- ● *Confounding* also involves two explanatory variates and one response variate; it is compared and contrasted with interaction (and with other causal structures involving three variates) in Statistical Highlight #65.

 (*continued overleaf*)

NOTES: (cont.) 3. Interaction is not limited to *two* factors – k focal variates have $\binom{k}{i}$ possible i-factor interactions; for example, four focal variates have $\binom{4}{2}$ = 6 two-factor interactions, $\binom{4}{3}$ = 4 three-factor interactions, and $\binom{4}{4}$ = 1 four-factor interaction. When i = 1, the k '1-factor interactions' are the k **main effects**, the effects of the k factors *individually*.

- Main effects and interaction effects are instances of **treatment effects**, and are represented by (response) *model parameters*. Any *linear combination* of such parameters where the coefficients sum to *zero* is called a **contrast**.
- For four focal variates, there are 4 + 6 + 4 + 1 = 15 treatment effects potentially of interest; these effects can *all* be estimated with a 16-run experimental Plan involving a factorial treatment structure.
- A *two*-factor interaction is the effect of one factor on the effect of another factor on a response variate; a *three*-factor interaction is the effect of one factor on the effect of another factor on the effect of a third factor on a response variate, and so on.

4. When there are two or more focal variates, 'lurking variates' criterion (1) [see the Appendix below] entails all *non*-focal variates be kept the same but, to allow interaction effect(s) to be estimated, the *focal* variates must be changed *together* according to the balanced scheme of a factorial treatment structure. However, confounding *may* then arise as outlined in Note 5 below.

- A ***mis***understanding of criterion (1) is to extend the *ensuring everything stays the same* precept to the *focal* variates and to only change them ***one*** *at a time*. For example, for *two* factors each with *two* levels (denoted *Lo* and *Hi*), have one run with both X_1 and X_2 set 'Lo', another run with X_2 set 'Hi' and another with X_2 back at 'Lo' and X_1 set 'Hi'; the resulting data, shown as three response variate averages in Table HL68.1 at the right, do *not* allow the X_1-X_2 interaction effect to be estimated, because there is no run with both factors set 'Hi'.

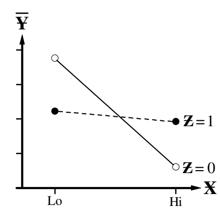
Table HL68.1	X_2 Lo	X_2 Hi
X_1 Lo	$\bar{Y}_{Lo,Lo}$	$\bar{Y}_{Lo,Hi}$
X_1 Hi	$\bar{Y}_{Hi,Lo}$	No data

Such a Plan, if it required four replicates for each treatment, would involve 12 runs. With a *factorial* treatment structure, only *4* runs provide the *same* level of replicating *and* an estimate of the interaction effect.

5. The idea in Note 3 above of estimating 15 treatment effects from a 16-run experimental Plan can be adapted to *fewer* estimates (7, say) from *fewer* (say 8 of the 16) runs – this is called a **fractional factorial** treatment structure (here, a **half** fraction). Under such a Plan, it is only possible to estimate *combinations* of treatment effects, like the main effect of one factor *and* one three-factor interaction. Because we cannot separate such combinations into their individual effects without data for *all 16* runs, there is *confounding* within the combinations.

- Inability to separate *treatment* effects under a Plan involving a *fractional* factorial treatment structure would be better called *perfect* confounding, to distinguish it from *partial* confounding (see Section 2 in Statistical Highlight #3), where the association of X and Z typically has a correlation with magnitude *less* than 1. As discussed in Statistical Highlight #3, both cases are usually (unwisely) simply called 'confounding' without distinction.

6. An idea, associated with the name of Taguchi, for *exploiting* interaction is illustrated by improvement of a process for manufacturing ceramic tiles; the diagram at the right for an X-Y relationship displays an interaction effect, because the *slope* of the (linear) relationship between X and the *average* of Y is *different* (here, smaller negative magnitude) when (non-focal) explanatory variate Z = 1 ('Hi') than when Z = 0 ('Lo'). In the tile-manufacturing process, if:



- Y is tile size *after* firing in an oven,
- X is oven temperature, whose variation from 'Lo' to 'Hi' over position within the oven causes tiles of the *same* initial size, but fired in different oven positions, to have different *final* sizes,
- Z is amount of clay in the ingredient mix used for the tiles,

by managing the amount of clay in the ingredient mix (*i.e.*, setting Z = 1), the manufacturing process is improved by making variation in tile final size *less* sensitive to variation in firing temperature due to tile position within the oven. This *indirect* approach exploiting interaction avoids the (more expensive) *direct* approach of making the temperature more uniform within the oven; of course, the properties of the tiles must remain acceptable when Z = 1 and clay must not be too expensive an ingredient.

Appendix: Criteria for Causation in Statistics (see Statistical Highlight #62)

To define ***formally*** in statistics what it means to say (a change in) X *causes* (a change in) Y in a ***target*** population, we state three criteria (useful in practice when establishing causation or quantifying the effect of X on Y):

(1) **LURKING VARIATES:** Ensure *all other* explanatory variates Z_1, Z_2, ..., Z_k hold their (same) values for *every* population element when X = 0 and X = 1 (sometimes phrased as: *Hold all the* Z_i ***fixed*** *for.....*).

(2) **FOCAL VARIATE:** Observe the population Y-values, and calculate an appropriate attribute value, under *two* conditions:
- ⊙ with *every* element having X = 0;
- ⊙ with *every* element having X = 1.

(3) **ATTRIBUTE:** Attribute(Y, perhaps some of Z_1, Z_2, ..., Z_k | X = 0) ≠ Attribute(Y, perhaps some of Z_1, Z_2, ..., Z_k | X = 1); those of Z_1, Z_2, ..., Z_k *included* in the attribute will have the *same* values when X = 0 and X = 1 under (1).