**Causal Factor**: x, **Criterion**: y

Pre-Experimental (observational) study: subjects assign themselves to different groups on their own

True-experimental designs (randomized controlled) study: scientist assigns subjects to treatment and control groups at random

Quasi-experimental designs: scientist unable to achieve complete control over scheduling of treatments or cannot randomly assign respondents to experimental treatment conditions, but has stringer controls than with basic observational studies

**--------------------------------------------------------------------**

**Notation to Represents Experimental Designs**

**X**: exposure of individual, group, or other entity to the experimental treatment

**O**: observation or measurement of the test unit

**R**: randomized assignment

Movement through time represented by horizontal arrangement of X’s and O’s from left to right. Simultaneous exposure or measurement by a vertical arrangement

Factors other than the experimental variable that affect the dependent variable are called nuisance, confounding, or extraneous factors/vars. When nuisance factors are not properly controlled, they are said to cofound the effects of the experimental variable

**Internal validity:** The ability of the experiment to unambiguously show a cause-and-effect relationship, i.e., to what extent can we attribute the effect that was observed to the experimental variable and not the other (cofounding) factors?

**External validity:** The extent to which the results of the experiment can be generalized to other people, settings (e.g., geography), and time (seasonality)

**--------------------------------------------------------------------**

|  |  |
| --- | --- |
| *X* | *O* |

**After-only design** (one-shot case study)

**Threats to IV**: history, maturation, selection

|  |  |  |
| --- | --- | --- |
| *O1* | *X* | *O2* |

**One-group Pretest-Posttest Design** (Before-after)

**Result of interest:** D\_hat = *O2* – *O1*

**Analysis:** Paired-sample t test

**Threats to IV:** history, maturation, pre-measurement, placebo effect

**Threats to Internal Validity (One-group PP Design)**

**Interaction** (or interactive testing) **effect:** when a pre-measure changes the respondent’s sensitivity or responsiveness to the independent variable(s). This is only a threat to external validity

**Placebo effect:** respondents acts differently because they know that they are being exposed to the treatment

|  |  |  |
| --- | --- | --- |
| *TG:* | *X* | *O1* |
| *CG:* |  | *O2* |

**Static group comparison:**

**Result of interest:** D\_hat = *O2* – *O1*

**Threats to IV:** selection, maturation / mortality (if treatment is unpleasant)

**Threats to Internal Validity (Static group compare)**

**Experimental mortality:** Differential loss of respondents from different groups

**Selection bias:** When the groups formed for the purposes of the experiment are initially unequal with respect to the dependent variable or in the propensity to respond to the independent variable

Remedies:

**- Randomization:** assign subjects to treatment and control group using a random procedure.

- **Matching:** match treatment and control groups with respect to variables which you suspect influence response (used with small sample sizes). \* Form blocks of units that are similar \* Randomly assign units within a block to treatment and control groups \* Blocking is closely related to stratification. Blocking is used in experiments and stratification in surveys

- ***Control*** for other causal factors (forks) with regression

- ***Propensity score*** models for observational studies. Find matched “twin(s)” for each treated case that is as similar as possible prior to self-selection into treatment

**Threats to Internal Validity (General)**

**Statistical regression:** When individuals are assigned to groups because of their scores on some measurement, such as initial attitude towards a brand. (Also called the *regression effect*)

**Threats to External Validity (General)**

All previous threats to internal validity are also threats to external validity. In addition, there are the following:

**Surrogate situation:** When the environment, the population samples, and/or the treatments are different copy those that will be encountered in the actual situation, e.g., copy testing … forced attention

**Measurement timing:** When pre-measurements are made at an inappropriate time to indicate the effect of the experimental treatment, e.g., effect of temporary price cut on forward buying

**True Design experiments**

**After-only vs Before-after with control group**

|  |  |  |  |
| --- | --- | --- | --- |
| *TG (R):* |  | *X* | *O1* |
| *CG (R):* |  |  | *O2* |

**After-only:**

|  |  |  |  |
| --- | --- | --- | --- |
| *TG (R):* | *O1* | *X* | *O2* |
| *CG (R):* | *O3* |  | *O4* |

**Before-after:**

**Result of interest:** D\_hat = (*O2* – *O1*) – (*O4* – *O3*)

**Analysis:** Compute differences between post and pre measures and compare with independent sample *t* test

**Threats to IV:** Interactive testing effect

**Quasi-Experimental Designs**

|  |  |  |  |
| --- | --- | --- | --- |
| *TG:* | *O1* | *X* | *O2* |
| *CG:* | *O3* |  | *O4* |

**Before-after with Control Quasi Design**

**Result of interest:** D\_hat = (*O2* – *O1*) – (*O4* – *O3*)

**Analysis:** Independent-sample *t* test on differences

- Sometimes, treatment and control *matched*: units assigned to treatment and control based on key factors

**Threats to Validity:** selection, interactive testing effects

**--------------------------------------------------------------------**

**Regression Terms and Symbols:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Term** | **ACT** | **JWHT** | **Other** |
| Sum of squared errors | **SSE** | **RSS** |  |
| Total sum of squares | **SST** | **TSS** |  |
| Mean squared error | **MSE** |  |  |
| Residual Standard Error |  | **RSE** |  |

**SSE =**

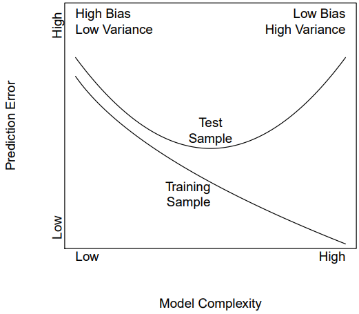
**MSE** =

**Penalized Estimates**

or AIC = deviance + 2p

**--------------------------------------------------------------------**

**Model Complexity**



**Training data:** Used to estimate model parameters

**Validation data:** Used to select model hyperparameters (or use K-fold cross validation if you are data poor)

**Test data:** Used for final, inter-model comparisons

**Model fit (SSE and deviance) adding predictors will:**

- Always improve SSE on *estimation* sample

- Not necessarily improve SSE on *validation* data

**Iterative Model Selection**

**Forward selection:** Begin with no vars. Add var the yields greatest significant improvement in SSE. Repeat until no significant improvement in SSE.

**Backward elimination:** Begin with all candidate vars. Drop variable that causes smallest non-significant increase in SSE.

**Stepwise selection:** (Usually) begin with no vars. Drop var that causes smallest *non-significant* increase in SSE. Add variable that yields greatest significant improvement in SSE. Repeat both previous steps until no improvement in SSE.

**Shrinkage Estimation**

**Min = MSE + lambda \* Penalty**

**Lambda -> infinity, big punishment for non 0 betas**

**High bias, low variance**

**Lambda -> 0, max strength model**

**Low bias, high variance**