How to Design and Analyze Experiments with Pass/Fail Responses

Discovery Europe – March 2024

Donald McCormack – Global Technical Enablement

Design & Analysis of Pass/Fail Experiments

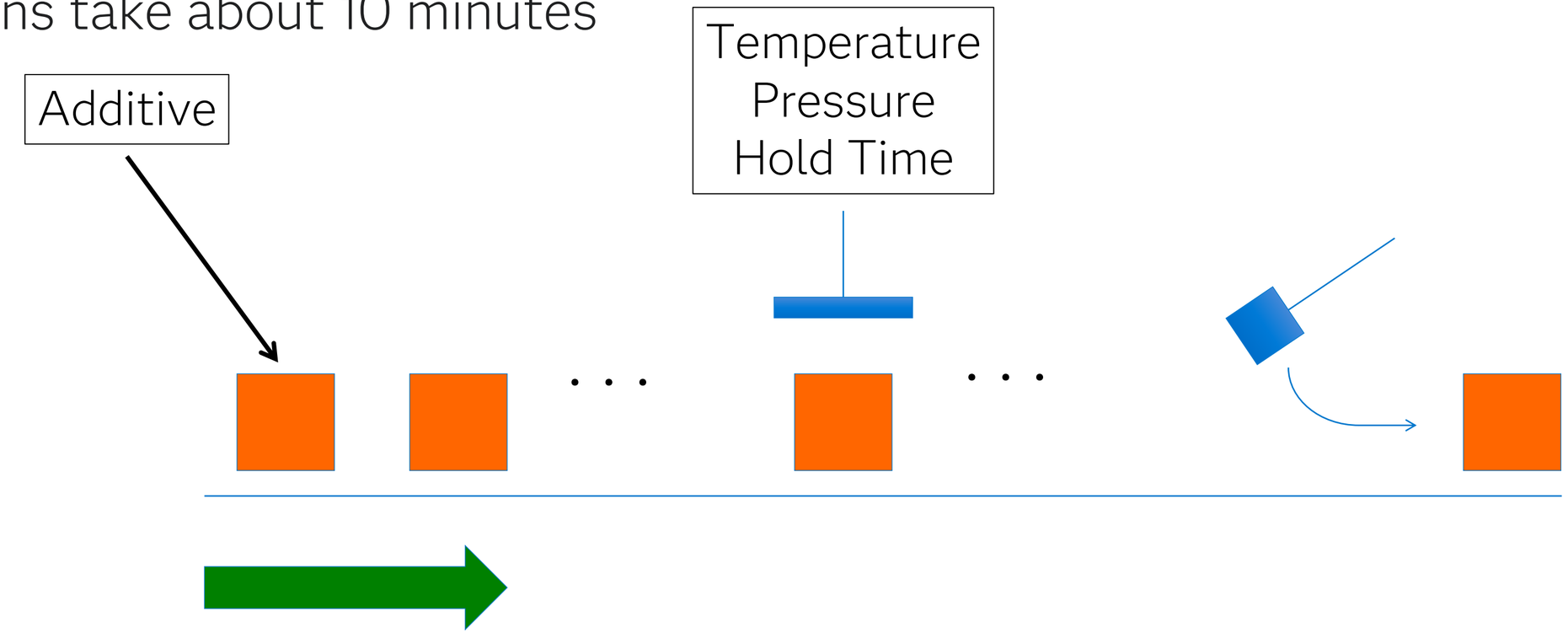
Introduction

- Widget Experiment Example
- It’s all about the model
- How to properly analyze experimental data
- How to right-size an experiment

Design & Analysis of Pass/Fail Experiments

Widget Experiment Example

- Does the widget have a defect? Current defect rate: 15%
- Goal: bring rate down to 5%
- Change of 3 – 5% or more is of practical importance
- Runs take about 10 minutes



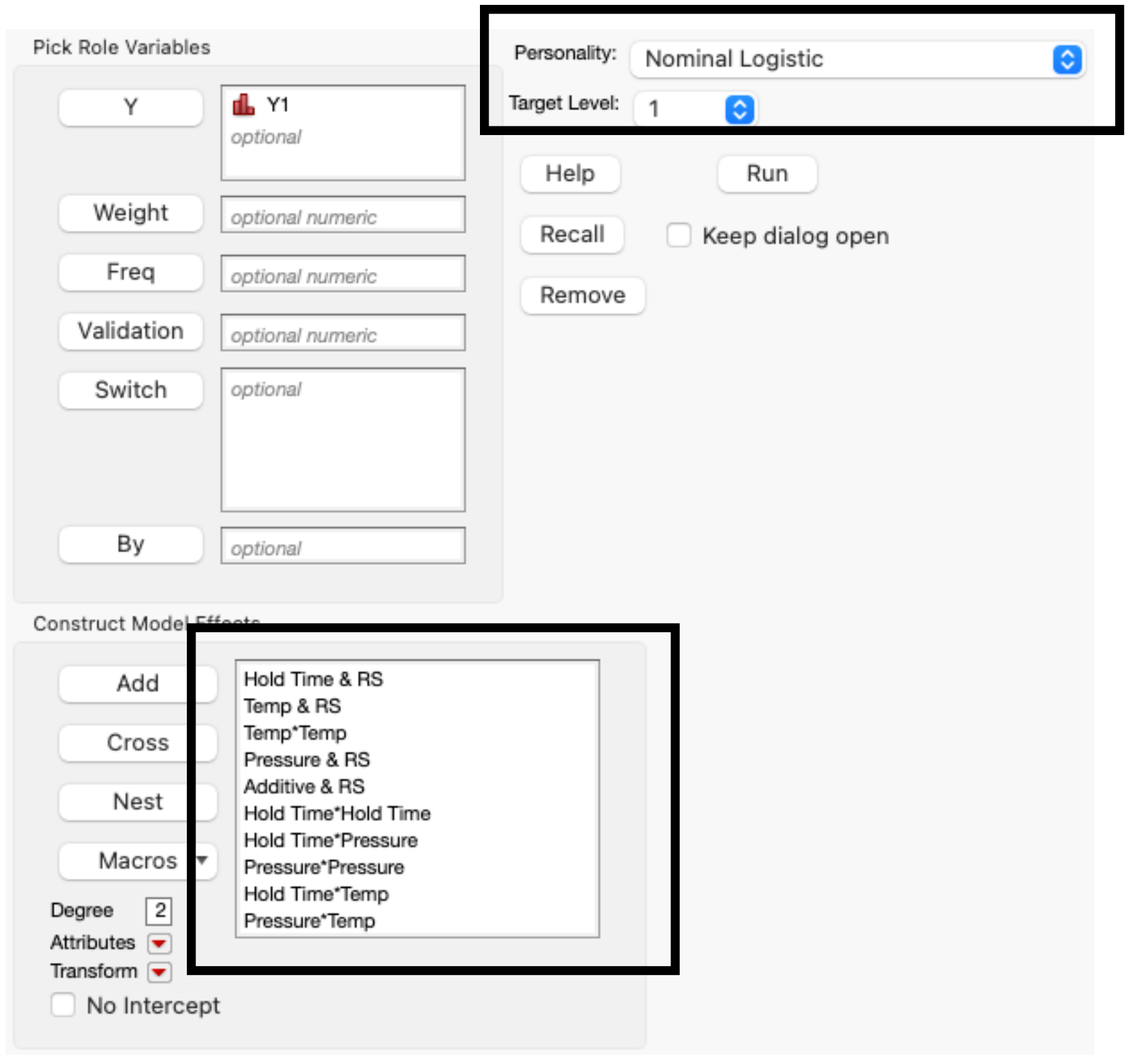
Design & Analysis of Pass/Fail Experiments

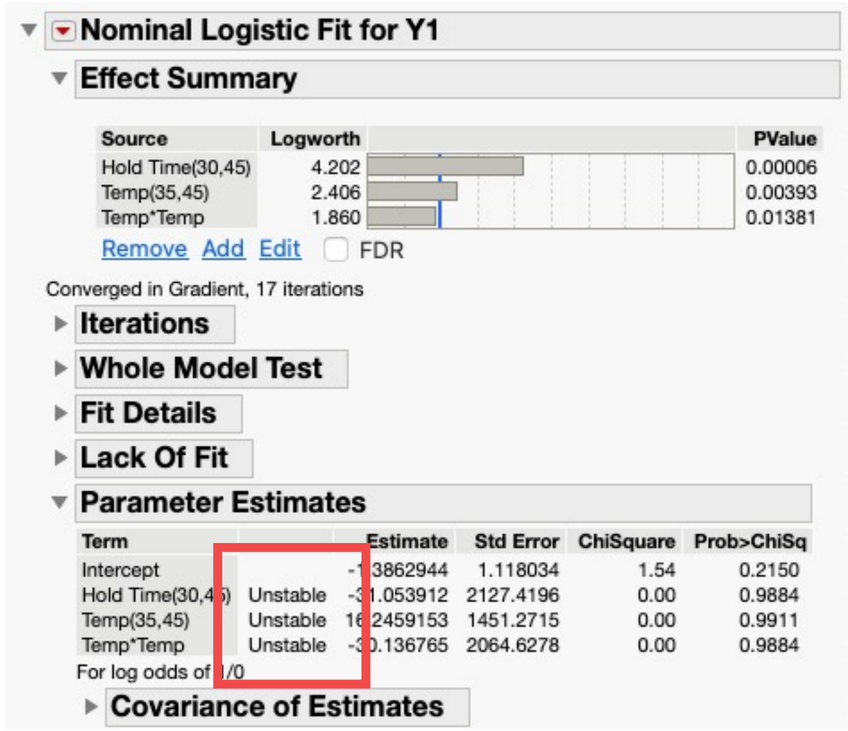
Widget Experiment Example

	Hold Time	Pressure	Temp	Additive	Y1
1	30	1.25	45	20	1
2	45	1	35	15	0
3	37.5	1.5	45	15	0
4	30	1	40	10	1
5	37.5	1.25	40	15	1
6	30	1	45	15	1
7	37.5	1.25	40	15	0
8	45	1	45	10	0
9	45	1.5	45	20	0
10	30	1.5	45	10	1
11	45	1.5	35	10	0
12	30	1.5	35	10	0
13	37.5	1	45	20	0
14	37.5	1.5	35	20	0
15	37.5	1	35	10	0
16	37.5	1.25	40	10	0
17	45	1.5	40	15	0
18	37.5	1.25	40	15	0
19	37.5	1.25	40	15	0
20	30	1.25	35	15	0
21	45	1	40	20	0
22	45	1.25	35	20	0
23	30	1.5	40	20	1
24	30	1	35	20	0

Design & Analysis of Pass/Fail Experiments

Widget Experiment Example

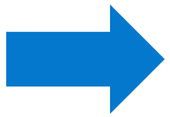


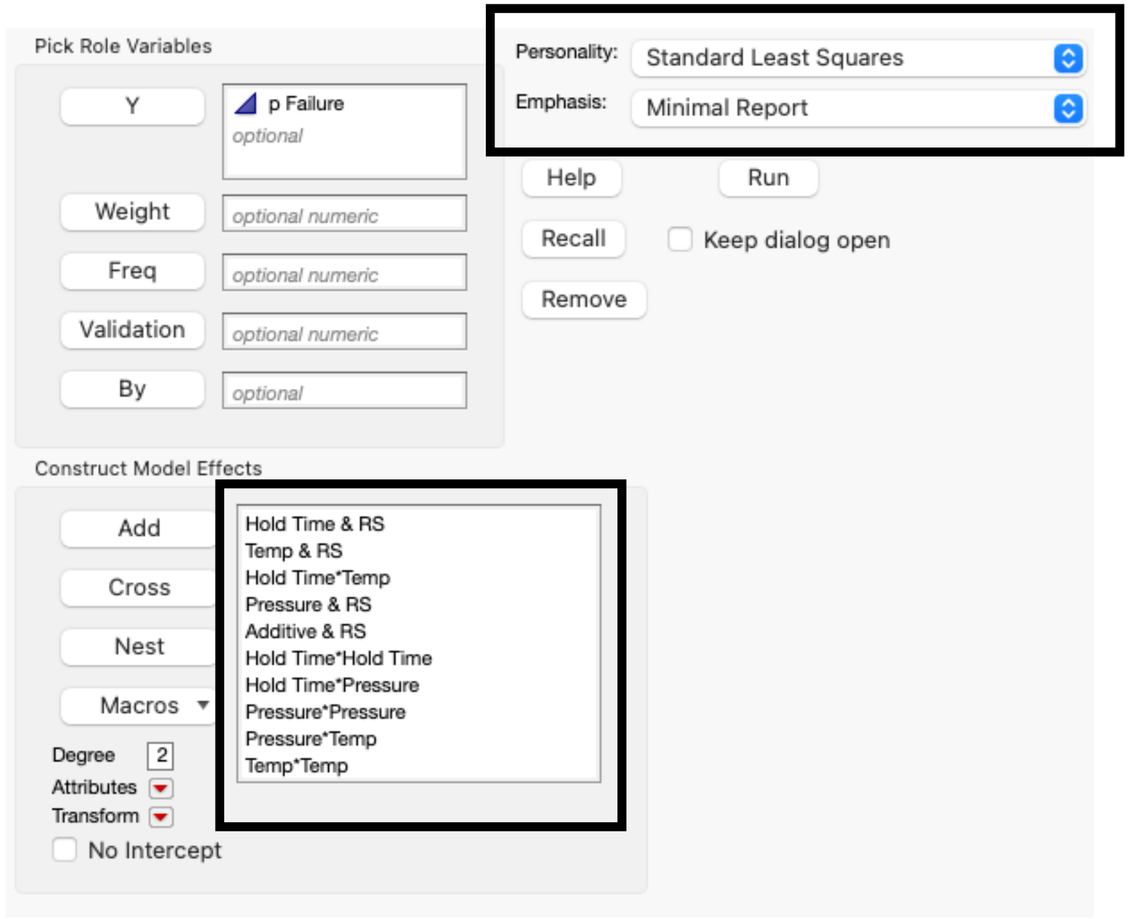


Design & Analysis of Pass/Fail Experiments

Widget Experiment Example

	Hold Time	Pressure	Temp	Additive	Y1	Y5	N	p Failure
1	30	1.25	45	20	1	3	5	0.6
2	45	1	35	15	0	0	5	0
3	37.5	1.5	45	15	0	2	5	0.4
4	30	1	40	10	1	1	5	0.2
5	37.5	1.25	40	15	1	0	5	0
6	30	1	45	15	1	5	5	1
7	37.5	1.25	40	15	0	4	5	0.8
8	45	1	45	10	0	0	5	0
9	45	1.5	45	20	0	0	5	0
10	30	1.5	45	10	1	4	5	0.8
11	45	1.5	35	10	0	1	5	0.2
12	30	1.5	35	10	0	0	5	0
13	37.5	1	45	20	0	3	5	0.6
14	37.5	1.5	35	20	0	0	5	0
15	37.5	1	35	10	0	2	5	0.4
16	37.5	1.25	40	10	0	0	5	0
17	45	1.5	40	15	0	0	5	0
18	37.5	1.25	40	15	0	0	5	0
19	37.5	1.25	40	15	0	0	5	0
20	30	1.25	35	15	0	0	5	0
21	45	1	40	20	0	0	5	0
22	45	1.25	35	20	0	2	5	0.4
23	30	1.5	40	20	1	2	5	0.4
24	30	1	35	20	0	1	5	0.2





Design & Analysis of Pass/Fail Experiments

Widget Experiment Example

Response p Failure

Effect Summary

Source	Logworth	PValue
Hold Time*Temp	2.713	0.00194
Hold Time(30,45)	2.059	0.00874 ^
Temp(35,45)	1.589	0.02575 ^

Remove Add Edit FDR ('^' denotes effects with containing effects above them)

Lack Of Fit

Summary of Fit

RSquare	0.583531
RSquare Adj	0.521061
Root Mean Square Error	0.216935
Mean of Response	0.25
Observations (or Sum Wgts)	24

Analysis of Variance

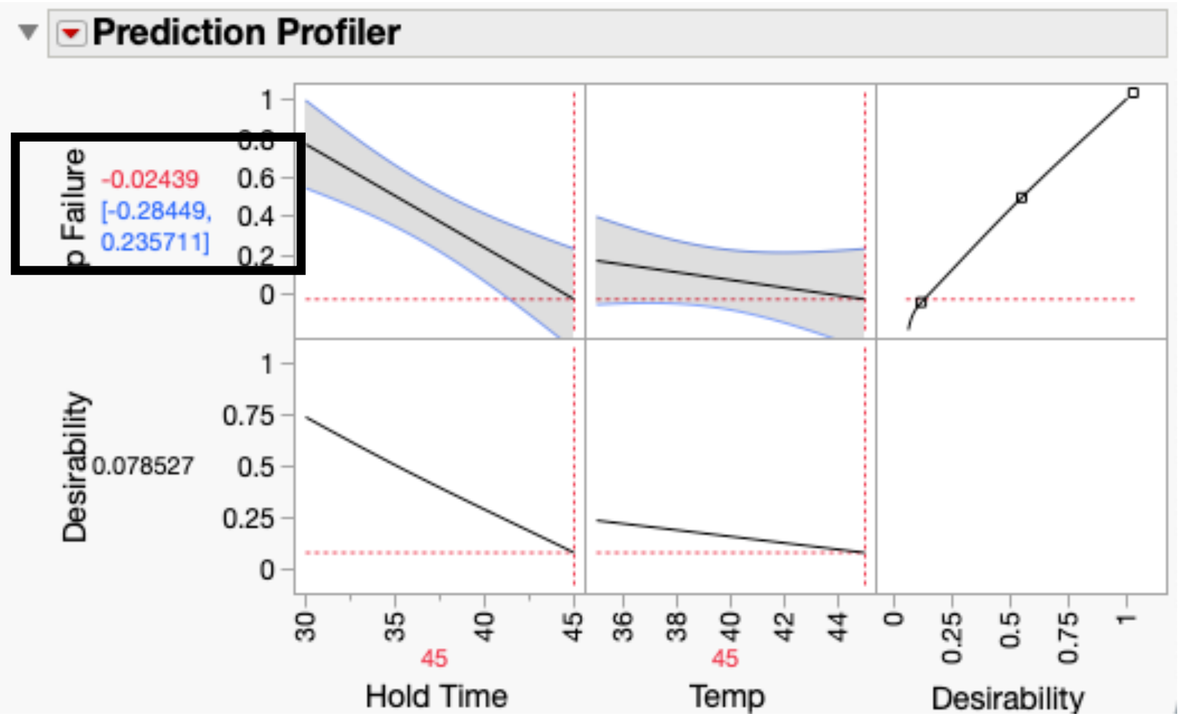
Source	DF	Sum of Squares	Mean Square	F Ratio
Model	3	1.3187805	0.439593	9.3409
Error	20	0.9412195	0.047061	Prob > F
C. Total	23	2.2600000		0.0005*

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	0.2390244	0.044527	5.37	<.0001*
Hold Time(30,45)	-0.164024	0.056444	-2.91	0.0087*
Temp(35,45)	0.1359756	0.056444	2.41	0.0258*
Hold Time*Temp	-0.235366	0.066004	-3.57	0.0019*

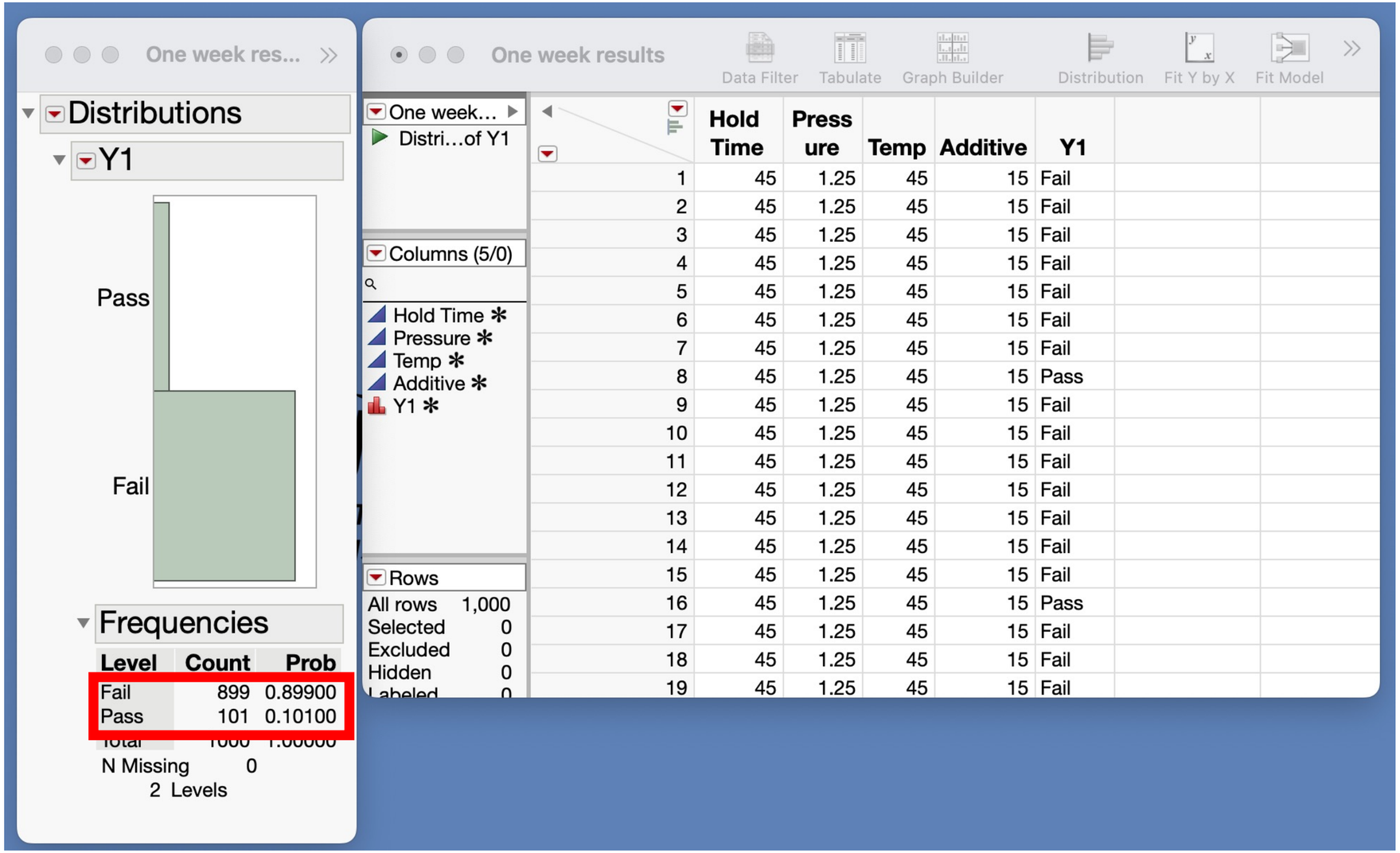
Effect Tests

Effect Details



Design & Analysis of Pass/Fail Experiments

Widget Experiment Example



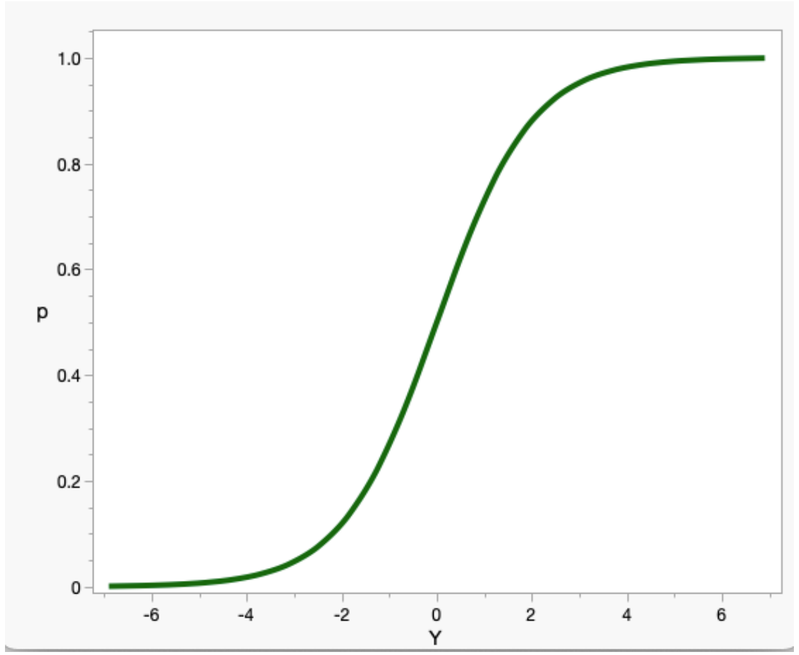
Design & Analysis of Pass/Fail Experiments

It's all about the model

- The typical model

$$P[fail] = f(X) = \beta_0 + \beta_1 \text{Hold Time} + \beta_2 \text{Pressure} + \ \ldots$$

- Assumptions of typical model:
  - Linear relationship
  - Unbounded response range
  - Normal residuals
- Can we fix these issues with a transformation on the failure probability or by using a different distribution?



Design & Analysis of Pass/Fail Experiments

It's all about the model

- Three options

$$Log\left(\frac{P[fail]}{P[pass]}\right) = Log\left(\frac{P[fail]}{1-P[fail]}\right) \qquad \text{Logit}$$

$$\Phi^{-1}\left[P[fail]\right] = Normal\ Quantile(P[fail]) \qquad \text{Probit}$$

$$Log\{-Log(1-P[fail])\} \qquad \text{Complementary Log-Log}$$

  - All three assume Binomial (i.e., not Normal) errors

- The first two are built into JMP platforms. The last one can be modeled using the Nonlinear platform.

Design & Analysis of Pass/Fail Experiments

How to properly analyze experimental data – Overview

- Fit Model – Nominal Logistic Regression (JMP)
- Fit Model – Generalized Linear Model (JMP)
- Fit Model – Generalized Regression (JMP Pro)
- Fit Y by X (one Y and one X only)/Nonlinear

Design & Analysis of Pass/Fail Experiments

How to properly analyze experimental data – Data Organization

- How should the data be organized? Example: 2 factor, four run experiment with five trials per unique treatment condition

	Time	Temperature
1	30	37
2	30	20
3	45	20
4	45	37

Design & Analysis of Pass/Fail Experiments

How to properly analyze experimental data – Data Organization

	Treatment Condition	Time	Temperature	Y
1	1	30	20	Green
2	1	30	20	Red
3	1	30	20	Green
4	2	30	37	Green
5	4	45	37	Green
6	1	30	20	Red
7	3	45	20	Red
8	3	45	20	Green
9	1	30	20	Red
10	2	30	37	Green
11	3	45	20	Green
12	2	30	37	Red
13	4	45	37	Green
14	3	45	20	Red
15	2	30	37	Green
16	2	30	37	Red
17	4	45	37	Red
18	3	45	20	Green
19	4	45	37	Green
20	4	45	37	Green

Raw

	Treatment Condition	Time	Temperature	Y	N
1	1	30	20	Green	2
2	2	30	37	Green	3
3	3	45	20	Green	3
4	4	45	37	Green	4
5	1	30	20	Red	3
6	2	30	37	Red	2
7	3	45	20	Red	2
8	4	45	37	Red	1

Summarized Stacked

	Treatment Condition	Time	Temperature	N Green	N Total
1	1	30	20	2	5
2	2	30	37	3	5
3	3	45	20	3	5
4	4	45	37	4	5

Summarized Split

Live demo here!

Fitting models to pass/fail data

Design & Analysis of Pass/Fail Experiments

How to right-size an experiment

- Use the Custom Design platform to create the intended design.
- Before generating the final design table turn on **Simulate Responses** under the hotspot in the top outline.
- Create the JMP Data Table containing the design. A dialog box will appear letting you change the default coefficients and error distribution.
  - Change the coefficients to the desired magnitude. Details are on the next slide.
  - Change the error type to Binomial and set the sample size to the desired value.
  - Click **Apply**
- The equation to use for the simulation is saved as a formula in the **Y Simulated** column.

Design & Analysis of Pass/Fail Experiments

How to right-size an experiment – Calculate Coefficients

- The coefficient values will depend on the underlying model, the value at the baseline probability and the probability value at which an observation is considered important. The absolute difference between these last two values corresponds to the effect size to use in the dialog
- Logit baseline

$$Log\left(\frac{P[target]}{1-P[target]}\right) = Log\left(\frac{0.15}{1-0.15}\right) = -1.73$$

- If we are interested in detecting a change to at least a 10% failure rate

$Log\left(\frac{0.1}{1-0.1}\right) = -2.2$ Making the coefficient $Abs[-1.73-(-2.2)] \approx 0.5$

Continue live demo!

Sizing an experiment through simulation

Design & Analysis of Pass/Fail Experiments

Summary

- Use the Fit Model platform for analysis. There are three options
  - Logistic Regression (+Stepwise Regression)
  - Generalized Linear Model
  - Generalized Regression (JMP Pro)
- Size the experiment properly by simulating experimental runs from the proposed design using JMP's built in bootstrap simulation feature (JMP Pro).

Supplementary Slides

Design & Analysis of Pass/Fail Experiments

How to properly analyze experimental data – Data Organization

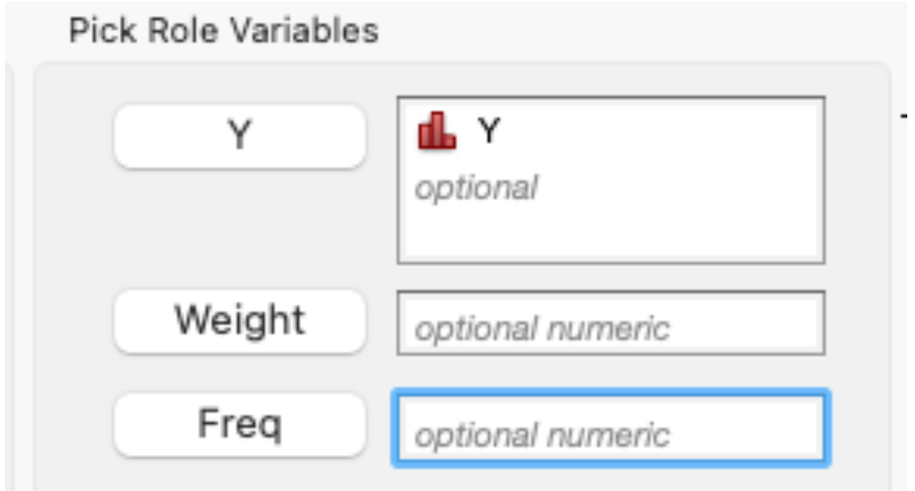
- Raw data: One column containing the nominal response level. One row per observations.
- Summarized Stacked: One column of counts aggregated over response level and unique treatment conditions, one column giving response level
- Summarized Split: One column of counts aggregated over one response level and unique treatment conditions, one column giving total counts for the unique treatment condition.

Design & Analysis of Pass/Fail Experiments

How to properly analyze experimental data – Data Organization

	Treatment Condition	Time	Temperature	Y
1	1	30	20	Green
2	1	30	20	Red
3	1	30	20	Green
4	2	30	37	Green
5	4	45	37	Green
6	1	30	20	Red
7	3	45	20	Red
8	3	45	20	Green
9	1	30	20	Red
10	2	30	37	Green
11	3	45	20	Green
12	2	30	37	Red
13	4	45	37	Green
14	3	45	20	Red
15	2	30	37	Green
16	2	30	37	Red
17	4	45	37	Red
18	3	45	20	Green
19	4	45	37	Green
20	4	45	37	Green

Raw

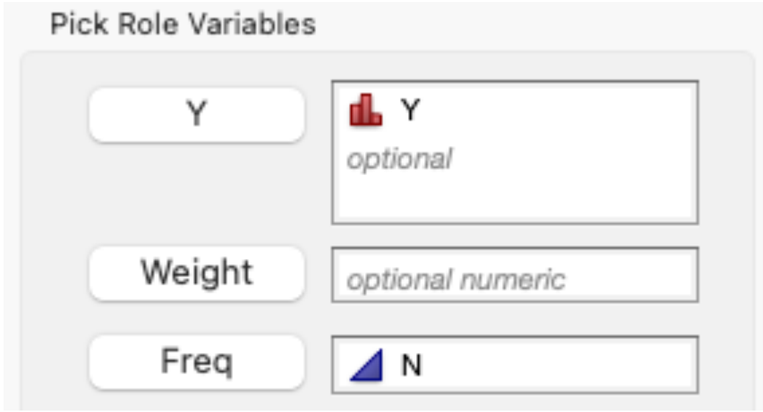


Design & Analysis of Pass/Fail Experiments

How to properly analyze experimental data – Data Organization

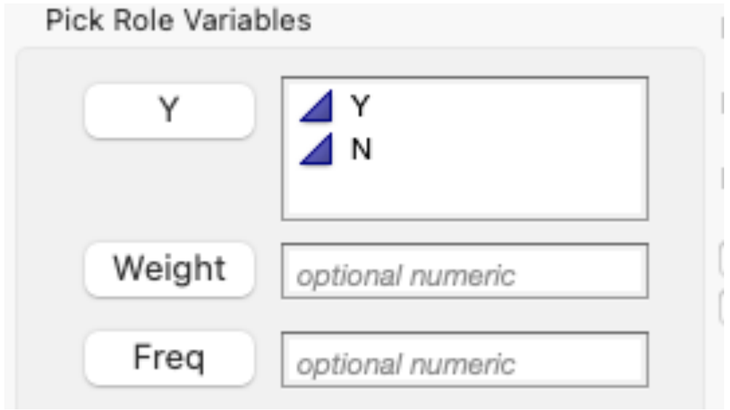
	Treatment Condition	Time	Temperature	Y	N
1	1	30	20	Green	2
2	2	30	37	Green	3
3	3	45	20	Green	3
4	4	45	37	Green	4
5	1	30	20	Red	3
6	2	30	37	Red	2
7	3	45	20	Red	2
8	4	45	37	Red	1

Summarized Stacked



	Treatment Condition	Time	Temperature	N Green	N Total
1	1	30	20	2	5
2	2	30	37	3	5
3	3	45	20	3	5
4	4	45	37	4	5

Summarized Split

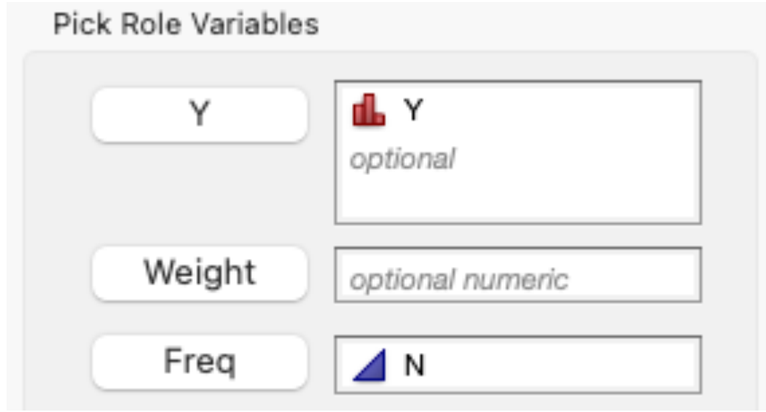


Design & Analysis of Pass/Fail Experiments

How to properly analyze experimental data – Data Organization

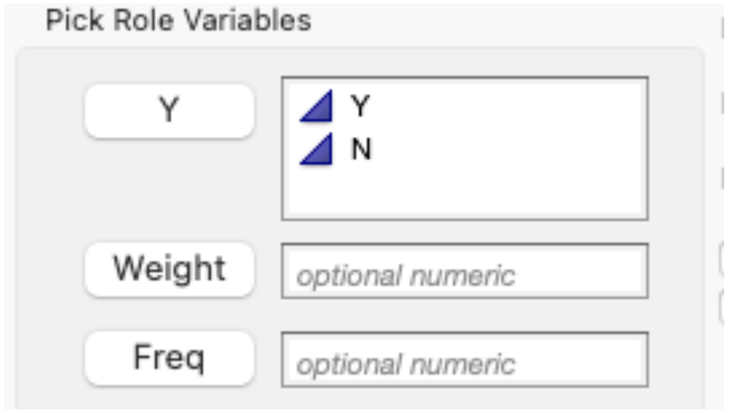
	Treatment Condition	Time	Temperature	Y	N
1	1	30	20	Green	2
2	2	30	37	Green	3
3	3	45	20	Green	3
4	4	45	37	Green	4
5	1	30	20	Red	3
6	2	30	37	Red	2
7	3	45	20	Red	2
8	4	45	37	Red	1

Summarized Stacked



	Treatment Condition	Time	Temperature	N Green	N Total
1	1	30	20	2	5
2	2	30	37	3	5
3	3	45	20	3	5
4	4	45	37	4	5

Summarized Split



Design & Analysis of Pass/Fail Experiments

Nominal Logistic Regression

- **Fit Model** (JMP)
- Data organized in raw or summarized stacked format
- **Nominal** response
- **Personality**: **Nominal Logistic** or **Stepwise**
- **Target Level**: associates the chosen level with the modeled probability . The other level is calculated as 1 minus the modeled probability.

Design & Analysis of Pass/Fail Experiments

Generalized Linear Model

- **Fit Model** (JMP)
- Data organized in any of the three formats
  - To use the Summarized Split format, both columns need to be **Continuous**. The first contains the counts for the target level, the other the total counts. The response needs to be **Nominal** for the other two formats.
- **Distribution: Binomial**
- **Link Function: Logit** or **Probit**

Design & Analysis of Pass/Fail Experiments

Generalized Regression

- **Fit Model** (JMP Pro)
- Any data organization. Similar set-up to **Generalized Linear Model**.
- **Target Level**: same as Logistic Regression
- **Distribution**: **Binomial**
- **Run** generates a report for a full logistic regression model. Reduced and alternative models are available under **Model Launch**.

Design & Analysis of Pass/Fail Experiments

How to right-size an experiment – Determine the Effect Directions

- The direction of the coefficients (i.e., whether they're positive or negative) will affect the results to a lesser extent and can also be added.
- Main effects: will the response increase (+) or decrease (–) as the factor increases?
- Two factor interactions: is the relationship between factors synergistic (+) or antagonistic (–)
- Quadratic effects: do you expect the response curve to produce a maxima (+) or minima (–).

Effect	Main Effect	Quadratic
Hold Time	+	+
Pressure	+	+
Temperature	–	+
Additive	+	–

Effect	Direction
Hold Time x Pressure	+
Hold Time x Temperature	–
Hold Time x Additive	+
Pressure x Temperature	+
Pressure x Additive	+
Temperature x Additive	+

Design & Analysis of Pass/Fail Experiments

How to right-size an experiment – Analyze Results/Run Simulation

- Using the Generalized Regression or Generalized Linear Model option, fit a model to the data. Logistic Regression can also be used, but only if the data is in raw format.
  - Data in Summarized Stacked format should be avoided because it requires complex changes be made to the simulation formula.
- Hover over a p-value column in a parameter estimates or effect tests table. Right click and select Simulate.
- In the resulting dialog, make sure **Column to Switch Out** and **Column to Switch In** are both selected to be the simulation column. **Set Number of Samples** to the desired value and click **OK**.
- **Simulate** will rerun the simulation column formula for each run and reanalyze the data using the method chosen above.

Design & Analysis of Pass/Fail Experiments

How to right-size an experiment – Analyze Simulation Results

- The resulting report window contains **Distribution** and **Power Analysis** scripts. The later is a subset of the former.
  - **Power Analysis** provides an estimate of effect p-values along with a confidence interval at different alpha values. It also counts the number of rejected effects at four different alpha values (0.01, 0.05, 0.10, 0.20) and provides credible intervals for these value.
  - **Distribution** adds the distribution graphs, quantiles and summary statistics
- Creating a summary counting the number of correctly and incorrectly identified effects broken down by type (main effect, two-factor interaction and quadratic) could also be beneficial.