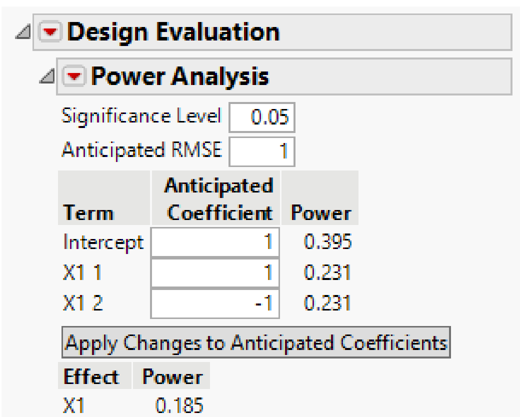
<u>**Power Analysis: Why use alternating coefficients for categorical factors with more than 2-levels?**</u>

<u>Background and Short Answer</u>

In Design Evaluation > Power Analysis, anticipated coefficients for categorical factors with more than two levels have default values of alternating +/- 1.

```
d = DOE(
	Custom Design,
	{Add Response( Maximize, "Y", ., ., . ),
	Add Factor( Categorical, {"L1", "L2", "L3"}, "X1", 0 ),
	Set Random Seed( 100 ), Number of Starts( 100 ),
	Add Term( {1, 0} ), Add Term( {1, 1} ), Set Sample Size( 6 ),
	Simulate Responses( 1 ), Save X Matrix( 0 ), Make Design}
);
```


Design Evaluation
Power Analysis
Significance Level 0.05
Anticipated RMSE 1

Term	Anticipated Coefficient	Power
Intercept	1	0.395
X1 1	1	0.231
X1 2	-1	0.231

Apply Changes to Anticipated Coefficients

Effect	Power
X1	0.185

These coefficients are generated by providing **Delta** (default=2), which is the desired difference you want to detect in the response due to changing from the low to high level (continuous factors) or between categories (categorical factor). See JMP 18 Help for more information. **JMP sets the coefficients for categorical factors as alternating +/- Delta/2 because this ensures, as we change between categories of the categorical factor, the difference in the mean response is Delta.**

<u>Statistical Details</u>

We define a linear model

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{e}$$

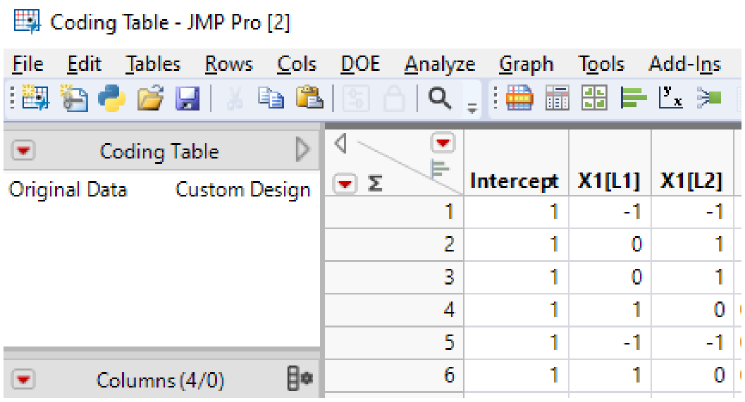
Where $\boldsymbol{y}$ is an nx1 vector of responses, **X** is an nxp design matrix, $\boldsymbol{\beta}$ is a px1 vector of parameters, and $\boldsymbol{e}$ is an nx1 vector of iid random error.

For a design with categorical factors, you can see the X matrix used in the power analysis calculations by saving the coding table from Fit Model.

```
dt = d << Make Table;
obj = (dt << Run Script("Model")) << Run;
dtCoding = obj << Save Coding Table;
```

In the above example, since X1 is a 3-level categorical factor, we need 2 columns in the coding table to define X1.

- In the Custom Design, whenever X1=L1, the coding table sets X1[L1] = 1 and X1[L2]= 0.
- In the Custom Design, whenever X1=L2, the coding table sets X1[L1] = 0 and X1[L2]= 1.
- In the Custom Design, whenever X1=L3, the coding table sets X1[L1] = -1 and X1[L2]= -1.



	Intercept	X1[L1]	X1[L2]
1	1	-1	-1
2	1	0	1
3	1	0	1
4	1	1	0
5	1	-1	-1
6	1	1	0

Using the coding table and the Anticipated Coefficients, we can calculate the mean response for each level, based on the model, as follows.

$$E(Y|X1 = L1) = [1 \quad 1 \quad 0] \times \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} = 2$$

$$E(Y|X1 = L2) = [1 \quad 0 \quad 1] \times \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} = 0$$

$$E(Y|X1 = L3) = [1 \quad -1 \quad -1] \times \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} = 1$$

Therefore, the maximum change in mean response across the levels of X1 is equal to 2, which is the desired difference that we want to detect (Delta).

If we set $\beta_A = [1, 1, +1]$ instead, these expectations would be 2, 2, and -1 respectively. So the change in mean response across levels of the factor would be 2 – (-1) = 3, which is *not equal* to the desired difference we want to detect.