

## ***Screening Design Selection***

### **Summary**

The STATGRAPHICS experimental design section can create a wide variety of designs intended to screen out the most important factors affecting a process. Included are two-level factorial designs, fractional factorials, mixed level fractions, irregular fractions, and Plackett-Burman designs. Each design may be replicated or augmented by the addition of centerpoints, depending on the magnitude of the effects that the experimenter wishes to detect.

This procedure is designed to help select an appropriate design from that collection, including the number of replicates and centerpoints. Once the desired effect and maximum block size are specified, a list of the best designs of each type is generated. The power curves for each suggested design are also plotted. The desired design may then be created by selecting the *Create Design* procedure on the experimental design menu.

**Sample StatFolio:** *screendesign.sgp*

### **Sample Data:**

None.

## Data Input

The first data input dialog box for this procedure is shown below.

- **Number of Factors:** the number of factors  $k$  to be varied during the experiment, between 2 and 16. The primary goal of most screening experiments is to select the most important factors from a larger list of possible candidates.
- **Designs to Consider:** the type of screening designs to consider in the selection process. Designs on the list include:

*Full 2-level factorials* – designs consisting of all combination of 2 levels for each of the  $k$  factors in the experiment. Full factorials can estimate the main effects of each factor and all interactions amongst the factors.

*Fractional factorials* – designs containing a subset of the runs in the full 2-level factorial. These designs are classified by their *resolution*, which refers to the type of effects that can be estimated:

**Resolution V:** the design can estimate all main effects and 2-factor interactions. These designs are preferable if they can be done in an acceptable number of runs.

**Resolution IV:** the design can estimate all main effects, but the 2-factor interactions are confounded amongst themselves. These designs can sense the

presence of interactions amongst the factors but cannot resolve which interactions are present.

**Resolution III:** main effects are clear of each other but are confounded with the 2-factor interactions. These designs should be used with care since any interactions can distort the interpretation of the main effects of the factors.

*Irregular fractions* – designs involving a non-orthogonal subset of the runs in the full 2-level factorials. Although there is some confounding between the various effects, the designs included have confounding patterns that make them “nearly” the resolution indicated.

*Mixed level fractions* – designs involving 3 levels of one factor and 2 levels of the other factors. Such designs can estimate curvature with respect to the factor that is run at 3 levels.

*Plackett-Burman designs* – subsets of the full 2-level factorial where, unlike the fractional factorials, the number of runs is not a power of 2. These designs are useful for screening large number of factors, although their confounding pattern is not as clean as the fractional factorials.

- **Maximum runs per block:** if the design is to be run in blocks, the maximum number of runs in a single block. Blocking is performed to eliminate the effect of an external factor that could impact the results, such as batch of raw material when all runs cannot be done with the same batch. This field may be left blank if the design will not be blocked.
- **Minimum centerpoints per block:** the smallest number of centerpoints per block. Centerpoints are runs done at a middle level of each of the factors, often replicated in order to estimate the magnitude of the experimental error. This setting insures that a minimum number of such runs will be generated.
- **Experimental error sigma:** the standard deviation  $\sigma$  of the experimental error. The experimental error is a measure of the extent to which the results of a single run might vary if the run was repeated at the same conditions of the experimental factors. The larger the experimental error, the more runs required to estimate an effect of a given magnitude. Normally, you would enter an estimate based on the results of previous experiments. If no estimate is available, you may leave this setting at 1.0. The “effects” displayed will then be expressed as a multiple of  $\sigma$ .

The dialog box shown above requests the generation of experimental designs for studying  $k = 5$  factors. Only designs that are of at least Resolution V (or nearly so) are to be considered. No blocking is to be done, nor is there any restriction on the minimum number of centerpoints. The experimental error  $\sigma$  is set to 1, so that the effects will be expressed in units of  $\sigma$ .

The second data input dialog box elicits information about the goal of the experiment:

- **Control:** specifies the goal from among the following choices:
  1. **Absolute error:** a confidence interval for each effect is to be constructed. That interval should not deviate from the point estimate in either direction by more than the absolute distance  $W$  indicated. The “effect” of a factor is defined as the magnitude of the change in the response over the range in which that factor is changed:
 
$$\text{effect} = (\text{response at high level of a factor}) - (\text{response at low level of a factor})$$
 If  $\sigma = 1$ , then  $W = 1$  implies a desired confidence interval for each effect of  $\pm$  one sigma.
  2. **Relative error:** a confidence interval for each effect is to be constructed. That interval should not deviate from the point estimate in either direction by more than the relative percentage  $P$  indicated. This is identical to the absolute error case with  $W$  set equal to  $P$  times the specified effect.
  3. **Power:** a hypothesis test is to be performed to determine whether or not each effect is significantly different than zero. The power of the test  $100(1-\beta)\%$  should equal the percentage specified when the true value of the effect equals  $\Delta = \text{Effect to Detect}$ . *Power* is defined as the probability of rejecting the null hypothesis when it is false, i.e., the probability that a factor will be judged to be statistically significant if the true effect equals  $\Delta$ .
  4. **Total Number of Runs:** the maximum number of runs that can be performed. If this selection is chosen, then the procedure will determine the best design of each type in which the number of runs does not exceed this limit.
- **Confidence Level:** the level of confidence  $100(1-\alpha)\%$  used when constructing confidence intervals. The value  $\alpha$  is also used as the level of Type I error when testing hypotheses. A

Type I error occurs when the null hypothesis is falsely rejected. In this case, it refers to the act of identifying an effect that does not actually exist.

For example, the dialog box above indicates that the following tests will be performed after the experiment is run:

Null hypothesis  $H_0$ : effect = 0  
 Alt. hypothesis  $H_A$ : effect  $\neq$  0

The probability of a Type I error (rejecting a true null hypothesis) is set to  $\alpha = 5\%$ , while the probability of a Type II error (not rejecting a false null hypothesis) is set to  $\beta = 10\%$  when the true absolute effect equals 2 (sigma). In practice, this will be done by looking at the ANOVA table and identifying as statistically significant any effect that has a *P-Value* less than 0.05.

The following table may be helpful in remembering how to set the error probabilities.

	Conclude that effect is not significant	Conclude that effect is significant
$H_0$ is true (effect = 0)	Correct decision	Type I error (false positive) risk = $\alpha$
$H_0$ is false (effect $\neq$ 0)	Type II error (false negative) risk = $\beta$	Correct decision power = $1 - \beta$

### Analysis Summary

The *Analysis Summary* shows the input parameters and the smallest available designs of each type:

<u>Screening Design Selection</u>								
Input								
Number of Factors	Min. Centerpoints per Block	Max. Runs per Block	Exp. Error Sigma	Desired Power	Target Effect	Confidence Level		
5	0		1.0	90.0%	2.0	95.0%		
Selected Designs								
Design	Runs	Resol.	Corner Points	Center Points	Error D.F.	Reps.	Blocks	Power (%)
Factorial $2^5$	32	V+	32	0	16	1	1	99.863
Half fraction $2^{5-1}$	22	V	16	6	6	1	1	91.4633
Mixed level fraction $3 \times 2^{4-1}$	24	~V	16	0	7	1	1	92.7197

The output includes:

- **Design:** the type of design
- **Runs:** the total number of runs to be performed.
- **Resol.:** the resolution of the design. V+ indicates a design of resolution greater than V. ~V means a design of nearly resolution V.
- **Corner Points:** the number of runs *N* placed at the vertices of the experimental region, i.e., runs in which all factors are set either at their highest level or their lowest level.

- **Center Points:** the number of runs  $C$  placed at the center of the experimental region, i.e., runs in which all factors are set midway between their low and high level.
- **Error D.F.:** the number of degrees of freedom  $\nu$  available for estimating the experimental error  $\sigma$ . The higher the degrees of freedom, the better the estimate. In calculating this value, it is assumed that all estimable main effects and two-factor interactions will be estimated, including combinations of 2-factor interactions in the case of the fractional factorials. For the mixed level factorials, a single quadratic effect is also accounted for.
- **Reps.:** the number of replicates of the runs in the base design.
- **Blocks:** the number of blocks into which the runs have been divided.
- **Power:** the probability of identifying as statistically significant any effect of size equal to the *Effect to detect* specified on the data input dialog box.

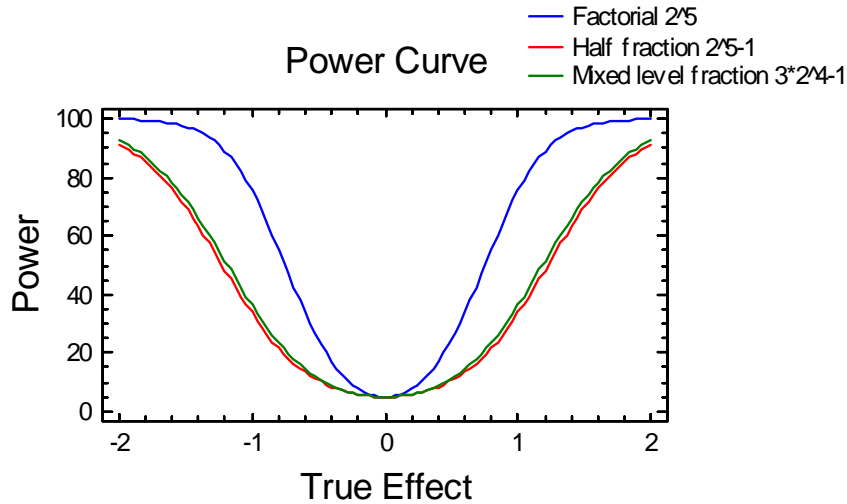
In the current example, three designs have been identified:

1. *Factorial 2<sup>5</sup>*: this is a full factorial design containing all  $N = 32$  combinations of the high and low level of the factors. Since the power is well in excess of 90%, no centerpoints were added.
2. *Factorial 2<sup>5</sup>-1*: a fractional factorial design containing  $N = 16$  of the 32 combinations of the runs in the full factorial. 6 centerpoints were added in order to achieve the desired power.
3. *Mixed level fraction 3\*2<sup>4</sup>-1*: a fractional factorial design involving 3 levels of one factor and 2 levels of the other four factors. There are  $N = 16$  corner points, plus 8 additional points at a middle level of the factor that is run at 3 levels. This design is not perfectly orthogonal, due to the presence of the third level of the one factor, but it is close enough to be easily interpreted.

The latter 2 designs are considerably smaller than the full factorial. If the desired power is 90%, one of these two designs should be selected. The choice between the two depends on whether or not the experimenter wishes to perform 2 extra runs in order to estimate possible curvature with respect to one selected factor.

## Power Curve

The Power Curve shows the probability that each of the selected designs will identify as statistically significant effects of a given magnitude.



The half fraction and mixed level fraction have very similar power.

### Calculations

- **Selection Algorithm**

The program begins with one replicate of the base design. It then adds points one at a time until the desired precision or power is achieved. Initially, centerpoints are added until enough additional points have been added such that a full replicate of the base design can be created instead. For blocked designs, the only designs considered are those with an equal number of centerpoints in all blocks.

- **Standard Error of an Effect**

$$\sigma_{effect} = \sqrt{\frac{4\sigma^2}{N}} \tag{1}$$

- **Confidence Interval for an Effect**

$$estimate \pm t_{\alpha/2, v} \hat{\sigma}_{effect} \tag{2}$$