BASICS OF EXPERIMENTAL DESIGN

From a statistician's perspective, an experiment is performed to decide **(1)** whether the observed differences among the treatments (or sets of experimental conditions) included in the experiment are due only to change, and **(2)** whether the size of these differences is of practical importance. Statistical inference reaches these decisions by comparing the variation in response among those experimental units exposed to the same treatment (experimental error) with that variation among experimental units exposed to different treatments (treatment effect). Thus, the three principles of experimental design are:

• replication, to provide an estimate of experimental error;

• randomization, to ensure that this estimate is statistically valid; and

• local control, to reduce experimental error by making the experiment more efficient. The number of replications (sample size) is the number of experimental units that receive each treatment.

The sample size should be small enough that negligible treatment differences are not declared statistically significant and large enough that meaningful treatment differences are declared statistically significant. Repeated measurements on the same experimental unit may or may not constitute true replications; treating dependent observations as if they were independent is one of the most common statistical errors found in the scientific literature.

Randomization means the use of a random device to assign the treatments to the experimental units. Randomization prevents the introduction of systematic bias into the experiment and provides the link between the actual experiment and the statistical model that underlies the data analysis. Thus, randomization is essential to the valid use of statistical methods. Performing the experiment with more care is one way to exert local control. For example, the treatments should be applied uniformly and under standardized conditions. However, an experiment can also be made more efficient by the judicious choice of design structure and treatment structure. If the experimental units are homogeneous, the treatments can be assigned to units randomly. Although the completely randomized design (CRD) is very flexible and easy to analyze, it is not always possible to obtain enough homogeneous experimental units to make this an efficient design. However, it is often possible to sort the experimental units into homogenous groups (blocks). The arrangement of the experimental units into blocks is the design structure of the experiment. There are many types of block designs, including the randomized complete block design, balanced or partially balanced incomplete block designs, and the Latin square design. With all of these designs, the gain in efficiency (compared with the CRD) is expected to outweigh the loss in flexibility and the increased complexity of the statistical analysis. An experiment with n types of treatments (factors), each with two or more levels, is said to have an n-way treatment structure. For example, an experiment comparing diets with three levels of protein and four levels of fat would have a two-way treatment structure.

DESIGN OF EXPERIMENTS (DOE)

term **experiment** is defined as the systematic procedure carried out under controlled conditions in order to discover an unknown effect, to test or establish a hypothesis, or to illustrate a known effect. When analyzing a process, experiments are often used to evaluate which process inputs have a significant impact on the process output, and what the target level of those inputs should be to achieve a desired result (output). Experiments can be **designed** in many different ways to collect this information. **Design of Experiments (DOE)** is also referred to as **Designed Experiments** or **Experimental Design** - all of the terms have the same meaning.

Components of Experimental Design

Consider the following diagram of a cake-baking process. There are three aspects of the process that are analyzed by a designed experiment:

- **Factors**, or inputs to the process. Factors can be classified as either controllable or uncontrollable variables. In this case, the controllable factors are the ingredients for the cake and the oven that the cake is baked in. The controllable variables will be referred to throughout the material as factors. Note that the ingredients list was shortened for this example - there could be many other ingredients that have a significant bearing on the end result (oil, water, flavoring, etc). Likewise, there could be other types of factors, such as the mixing method or tools, the sequence of mixing, or even the people involved. People are generally considered a Noise Factor (see the glossary) - an uncontrollable factor that causes variability under normal operating conditions, but we can control it during the experiment using blocking and randomization. Potential factors can be categorized using the Fishbone Chart (Cause & Effect Diagram) availabl e from the Toolbox.

- **Levels**, or settings of each factor in the study. Examples include the oven temperature setting and the particular amounts of sugar, flour, and eggs chosen for evaluation.

- **Response**, or output of the experiment. In the case of cake baking, the taste, consistency, and appearance of the cake are measurable outcomes potentially influenced by the factors and their respective levels. Experimenters often desire to avoid optimizing the process for one response at the expense of another. For this reason, important outcomes are measured and analyzed to determine the factors and their settings that will provide the best overall outcome for the critical-to-quality characteristics - both measurable variables and assessable attributes.

Purpose of Experimentation Designed

experiments have many potential uses in improving processes and products, including:

- **Comparing Alternatives.** In

the case of our cake-baking example, we might want to compare the results from two different types of flour. If it turned out that the flour from different vendors was not significant, we could select the lowest-cost vendor. If flour were significant, then we would select the best flour. The experiment(s) should allow us to make an informed decision that evaluates both quality and cost.

- Identifying the **Significant Inputs** (Factors) Affecting an Output (Response) - **separating the vital few from the trivial many**. We might ask a question: "What are the significant factors beyond flour, eggs, sugar and baking?"
- Achieving an **Optimal Process Output**(Respons). "What are the necessary factors, and what are the levels of those factors, to achieve the exact taste and consistency of Mom's chocolate cake?
- **Reducing Variability**. "Can the recipe be changed so it is more likely to always come out the same?"
- **Minimizing, Maximizing, or Targeting an Output** (Response). "How can the cake be made as moist as possible without disintegrating?"
- Improving process or product "**Robustness**" - fitness for use under varying conditions. "Can the factors and their levels (recipe) be modified so the cake will come out nearly the same no matter what type of oven is used?"
- **Balancing Tradeoffs** when there are multiple Critical to Quality Characteristics (CTQC's) that require optimization. "How do you produce the best tasting cake with the simplest recipe (least number of ingredients) and shortest baking time?"

Experiment Design Guidelines

The Design of an experiment addresses the questions outlined above by stipulating the following:

- The **factors** to be tested.
- The **levels** of those factors.
- The **structure** and layout of experimental runs, or condition.