**Chapter 12: Experimental Design and Optimization**

Table of Contents

[Factorial Design 3](#_Toc87178408)

[Randomness of Effects 6](#_Toc87178409)

In this chapter, we will be dealing with **input parameters**, more specifically, how each input parameter **affects the output**. For a queueing system for example, we might be interested in how changing just the number of servers to see how it affects the output, keeping other factors constant, or perhaps for all combinations of other factors.

Here, we will be using the term **factors** to refer to input parameters, system configurations or structural assumptions that affect the output results, or performance measure, termed **response**. The factors could be **controllable** or **uncontrollable**.

We will be performing **factor analysis**, also called **sensitivity screening**, which just means that we will identify the factor that has the most impact on the response.

## Factorial Design

Consider that we have just **one factor** for which we are setting different values, called **levels**. To compare performance, we need **at least two levels**. If we run the simulation for  **replications** for each level, then we have a total of replications. For  **factors**, where , assuming each factor has  **levels**, we will have  **configurations**, meaning we will have replications in total.

If we have factors, we want to know how the th factor affects the response. One way we could do this is by keeping the other factors **fixed** while **changing** the th factor. Since we have just **two levels**, we will consider one as a **negative level**, which has a lower value (not necessarily a negative value) and the other as a **positive level**, which has a higher value.

The question now arises about where to fix the other factors. The other factors will have different combinations. Which combination should we choose?

Say we have **two factors**, and , each with **two levels**, and and and respectively. If we want to examine the **effect of** , we must **fix** . Suppose we fix at . Thus, we will have **two responses**, and . The difference between these two values will be the **change in response**. Similarly, we could have fixed at , which would get us the responses and . However, neither of these would be accurate, since each case, we are essentially not considering **all possible configurations** for the other factor.

To deal with this problem, we will take the **average** of both configurations, i.e.

Note that we **must** subtract the **negative** value from the **positive** one.

To make sure we are taking all configurations into account, we could use a **design matrix**, such as the one below, which is for three factors.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Design Point | Factor 1 | Factor 2 | Factor 3 | Response |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

For example, if we are measuring the effect of Factor 1, , we have 4 sets of 2 configurations each. Thus, the equation we will have is

The easiest way to figure this out is to just write it down serial. The **sign** will depend on whether the **factor** we are considering is positive or negative.

We can also calculate the effect of **two factors**. Suppose we want to calculate the effect of Factor 1 and Factor 3, . Again, we will write down the responses serially, and the **sign** will depend on the **product** of the two factors, i.e. two negatives make a positive, a positive and a negative make a negative and so on.

Notice that the **denominator** does not change.

We can extend upon this for the effect of even more factors.

Example

In an inventory system, we have two factors, , the threshold at which orders are placed, and , the difference between the maximum inventory capacity and .

|  |  |  |
| --- | --- | --- |
| Factor |  |  |
|  |  |  |
|  |  |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Design Point |  |  |  | Response |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

The responses given here are the average cost per month for each of the stated factors in the system.

Based on this,

Thus, we can conclude that increase increases cost, increasing decreases cost and increasing both increases cost.

### Randomness of Effects

Since the **responses** are **random variables**, the **observed effects** are also **random**. This means we can measure the **average effect**.

Say we repeat the simulation times and measure the same effect. Thus,

This produces a **confidence interval** of

|  |  |  |
| --- | --- | --- |
| Design Point | Sample Mean | Sample Variance |
| , |  |  |
| , |  |  |
| , |  |  |
| , |  |  |

|  |  |
| --- | --- |
| Expected Effect | percent confidence interval |
|  |  |
|  |  |
|  |  |

One interesting thing is that if the **range** of the interval for an effect includes the value **zero**, then there is a chance that the effect might be **meaningless**.