Introduction

Any research study requires the researcher to conduct one or more experiments to make confident claims based on their results. A thorough plan is essential on how an experiment should be conducted is under a particular important aspect of the statistical theory called \emph{experimental design}.

Among the advantages that DOE can provide is the increased amount of information per experiment compared to an ad hoc approach. The second benefit occurs in providing an organized approach toward analysis and interpretation of results, thus facilitating communication. Another advantage is the ability to identify interactions among factors, leading to more reliable prediction of response in areas not directly covered by experimentation. The fourth benefit is in the assessment of information reliability in light of experimental and analytical variation.

Statistical design of experiments (DOE), or simply experimental design, is a proven technique used extensively today in many industrial-manufacturing processes. Considering that this method was originally conceived to identify genetic variation in crops, it has not, until recently, been widely taken up by life scientists.

As more research disciplines are using automation and microfluidics to obtain faster results, an increasing number of scientists are now recognizing the assistance that experimental design can provide. Consequently, this technique is finding increasing acceptance in many areas beyond its origins in genetics.

(Talks about High-throughput biotechnology)

The type of experiment that this thesis is focusing on is high-throughput biotechnologies experiment.

The key aspect of high-throughput biotechnologies experiment is parallelization

High-throughput biotechnologies have improved rapidly within a last decade; on the other hand, the statistical methods, for analysing the data generated from these technologies, are falling further behind (need some references here).

\emph{Experimental design}, introduced by \citep{Fisher1935}, is a set of procedures that outline how an experiment should be conducted

maximised the resultant information

as efficient as possible

which allows us to define the model on how the analysis should be performed.

variation arising as much

the complexities of the measurment rocedure

One of the most important statistical theories during the process of collecting data is \emph{experimental design} which was initially introduced by \citep{Fisher1935}.

Experimental design is

(Talks about High-throughput biotechnology)

High-throughput biotechnologies have improved rapidly within a last decade; on the other hand, the statistical methods, for analysing the data generated from these technologies, are falling further behind (need some references here).

There are many different families of designs exits in the published literature

(Links High-throughput biotechnology with two-phase experiment)

two-phase experiments

(Using an example of proteomics experiment to discuss two-phase experiment)

Using an example of proteomics experiment

The main focuses of this thesis is proteomics experiments, which is the study of proteins.

The Phase 1 experiment involves the organisms that are to be perturbed by the experimental conditions of interest. Since the abundance of proteins cannot be measured directly from the organisms, the Phase 2 experiment uses multiplexing techniques such as iTRAQ peptide labelling, coupled with liquid chromatography-mass spectrometry (LC-MS), to measure the abundance of proteins in samples extracted from the organisms in the Phase 1 experiment. These two steps of experimentation are also known as two-phase experiments.

Most of the experiments involves these high-throughput biotechnologies have a two-phase structure, when the responses of experimental units to treatments cannot be measured directly in a single experiment. Subsequent processing (Phase 2) of the initial (Phase 1) experiment is necessary in order for the measurements to be made.