Statistica Sinica **37** (2027), 1-17
doi:https://doi.org/10.5705/ss.202024.0224

GENERAL SLICED FACTORIAL DESIGNS FOR ONLINE EXPERIMENTS WITH MULTIPLE PLATFORMS

Zijian Han[1], Dongying Wang[2], Fasheng Sun*[1] and Peter Chien*[3]

[1]*Northeast Normal University,* [2]*Jilin University of Finance and Economics and* [3]*University of Wisconsin-Madison*

Abstract: Digital marketing is an integral part of digital transformation in industry. It is critical to use design of experiments to conduct online experiments for various forms of digital marketing including web design, email marketing, social network and others. Online experiments often involve multiple platforms, including desktop computers from different manufacturers, different types of mobiles and smart watches of different brands. A sliced factorial design is a suitable choice for designing online experiments with multiple platforms. We provide a general theory for sliced factorial designs and propose sliced generalized wordlength patterns to construct such designs for any number of platforms. The theory uses the characteristics of parallel flat design-based sliced factorial designs to construct optimal sliced factorial designs. For practical use, several constructed designs are provided in the Supplementary Material, enabling practitioners to efficiently design and analyze online experiments across multiple platforms.

Key words and phrases: Design of experiments, digital marketing, J-characteristics, sliced effect hierarchy principle.

1. Introduction

With the growth of the internet, online experiments are being conducted for various forms of digital marketing including website optimization, mobile apps, social networks, video recommendation and email campaigns (Haizler and Steinberg, 2021).

Typical A/B testing refers to experiments whose goal is to identify whether a proposed change is effective. The term "A/B testing" was invented for online experiments in which a sample population is randomly divided into two mutually exclusive sets (runs), A and B, allowing for comparative analysis. Some online experiments test features one-at-a-time, which is called "one-factor-at-a-time" (OFAT) in the design of experiments. Since the pioneering work of Fisher, it has been well known that factorial experiments outperform OFAT; see Section 4.7 of Wu and Hamada (2021). This form of testing is called a multivariate test in online experiments. In a simple example, consider the following three website factors, each at two levels, corresponding to two different versions: the

*Corresponding author. E-mail: (Sun) sunfs359@nenu.edu.cn; (Chien) peter.chien@wisc.edu