Factorial design in LimeSurvey: approaches of realization

Mykola Sydorov, Oleksiy Sereda

Taras Shevchenko National University of Kyiv, Faculty of Sociology

BaNoCOSS 2019 (Örebro, Sweden)

Factorial Design

Example

- Dimensions 7
- Levels 2,2,2,5,5,2,5
- Vignette

You were offered a job in your specialty with a high degree of professional development and personal experience with a low career potential. Official employment without a social package and informal salary. The schedule is full time in the office. The office is far from home. The salary is more than 20% of the average in industry.

Vignette Sample

- Vignette space 2*2*2*5*5*2*5=2000
- Deck size 6-10 (4+4=8 – 2 blocks)
- Mails DB = 347 (bachelors, masters and graduates)
- Total vignettes number 347*4=1388 for each of 2 blocks
- Method D-efficiency

+

- Open source survey, multiadmins, offline work
- Multilingual, 28 different Q types
- WYSIWYG, logic (conditions), expression manager
- respondent DB, invites+reminds, security
- Export to SPSS, SAS, R, csv etc.
- Save survey structure and archive…

www.limesurvey.org

-

- Needs extra knowledge

- library(AlgDesign)
- vign<-gen.factorial(c(2,2,2,5,5,2,5), factors="all", varNames=c("d1","d2","d3", "d4","d5", "d6","d7"))
- desD<-optFederov(frml=~.,data=vign, approximate=TRUE, criterion="D")
- desD2<-optFederov(frml=~., data=vign, nTrials=1388,nRep=100, criterion="D")

Approach #1

+

- Tokens for every mail (mail DB)
- Different deck for every respondent
- 3 categories: bachelors, masters, graduates

-

- Forbiddance of Mail.ru (15May2017)
- Low RR of masters and very low RR of graduates
- As a result: 544 vignettes (4 in deck), 136 of 347 respondents

Dim	Levels	Vignette sample (abs)	Survey data (abs)	Dim	Levels	Vignette sample (abs)	Survey data (abs)
1	1	694	89	5	1	278	188
	2	694	455		2	278	224
2	1	694	86		3	278	34
	2	694	458		4	276	52
3	1	694	452		5	278	46
	2	694	92	6	1	694	431
4	1	278	136		2	694	113
	2	278	140	7	1	276	0
	3	278	192		2	278	0
	4	278	39		3	278	105
	5	276	37		4	278	249
					5	278	190

Approach #1. Randomization

- Tokens for every mail (mail DB)
- Set of decks less than number of respondent (deck multiple use)
- Vignettes randomisation before deck generation
- Less number of vignettes/decks

Dim	Levels	Vignette sample (abs)	Survey data (abs)	Dim	Levels	Vignette sample (abs)	Survey data (abs)
1	1	15	1042	7	1	16	1118
	2	15	615		2	14	539
2	1	16	1092	8	1	16	995
	2	14	565		2	14	662
3	1	15	1055	9	1	16	1114
	2	15	602		2	14	543
4	1	15	1046	10	1	8	812
	2	15	611		2	7	280
5	1	15	1073		3	7	251
	2	15	584		4	8	314
6	1	15	1073				
	2	15	584				

Approach #1. Randomization

+

- No lost levels in dimensions
- Invitation/reminder (mail DB)
- Bias control, RR

-

- Bias

Approach #2.

- Vignette sample random mixing online
- 6 blocks of 10 different vignettes in each
- Less vignettes than estimated number of respondents
- Social Network distribution (no DB) -> more respondents

Dim	Levels	Vignette sample (abs)	Survey data (abs)	Dim	Levels	Vignette sample (abs)	Survey data (abs)
1	1	30	861	6	1	30	828
	2	30	1078		2	30	1111
2	1	30	1048	7	1	30	791
	2	30	891		2	30	1148
3	1	30	875	8	1	30	881
	2	30	1064		2	30	1058
4	1	30	899	9	1	30	1051
	2	30	1040		2	30	888
5	1	30	816	10	1	20	764
	2	30	1123		2	20	578
					3	20	597

Approach #2.

+

- Vignette sample random mixing online
- Social Network distribution (no DB) -> more respondents
- Item nonresponse
- Similar distribution of generated and received vignettes

-

- Impossible RR and Bias estimation
- No invitation/reminder
- Very few different vignettes

Approach #3

- Using R in survey structure modification
- 480 vignettes total in vignette space, 268 selected
- 6 blocks of 44-45 vignettes in each (the number may be larger)
- Random vignette selection in each block (1 in block)
- Random selection of blocks' queue

Dim	Levels	Vignette sample (abs)	Survey data (abs)	Dim	Levels	Vignette sample (abs)	Survey data (abs)
1	1	124	139	5	1	135	148
	2	144	167		2	133	158
2	1	95	111	6	1	148	160
	2	93	89		2	120	146
	3	80	106	7	1	134	155
3	1	89	100		2	134	151
	2	97	122	8	1	120	133
	3	82	84		2	148	173
4	1	124	139				
	2	144	167				

Approach #3

<row> <qid><![CDATA[11048]]></qid>
<parent_qid><![CDATA[0]]></parent_qid>
<sid><![CDATA[888992]]></sid>
<gid><![CDATA[515]]></gid>
<type><![CDATA[X]]></type>
<title><![CDATA[v1t15]]></title>
<question><![CDATA[VIGNETTE TEXT]]></question> <preg/>
<help/> <other><![CDATA[N]]></other> <mandatory/>
<question_order><![CDATA[15]]></question_order>
<language><![CDATA[uk]]></language>
<scale_id><![CDATA[0]]></scale_id>
<same_default><![CDATA[0]]></same_default>
<relevance><![CDATA[(((!is_empty(888992X515X11023.NAOK) && (888992X515X11023.NAOK == 15)))))]]></relevance>
</row>

Approach #3

+

- Vignette sample random selection online
- Random block queue selection
- Item nonresponse, Unit nonresponse
- Similar distribution of generated and received vignettes
- Bias estimation
- Invitation/reminder (mail DB)
- Different vignette number ~ RR*db_size

-

- Complicate survey structure

Now online

- Vignette sample random mixing online
- Mail DB
- Item nonresponse, Unit nonresponse, RR
- Invitation/reminder
- 480 vignettes in 6 sets by 80 + core questionnaire

Thank you for your patience