

Chapter 7 实验法 (二): 析因设计与离散选择实验

授课教师: 吴翔

邮箱: wuhsiang@hust.edu.cn

- 1 析因设计
- 2 分式析因设计
- 3 离散选择实验

Section 1

析因设计

析因设计概述

很多实验要研究两个或多个因子的效应。此时，析因设计是最有效的。

析因设计 (factorial design) 是指，实验的每一次完全试验或每一次重复中，这些因子水平的所有可能的组合都被研究到。例如，当因子 A 有 a 个水平和因子 B 有 b 个水平时，则每次重复都包含全体 ab 个处理组合。析因设计可以估计**主效应和所有交互效应**。

析因设计的优点

析因设计具备一些独特的优点：

- 它比一次一因子实验效率高。
- 存在交互作用时，为了避免产生令人误解的结论，必须使用析因设计。
- 析因设计容许因子相对于其它各个因子的几个水平来估计其效应，所得结论在实验条件的范围内是有效的。

2^k 析因设计

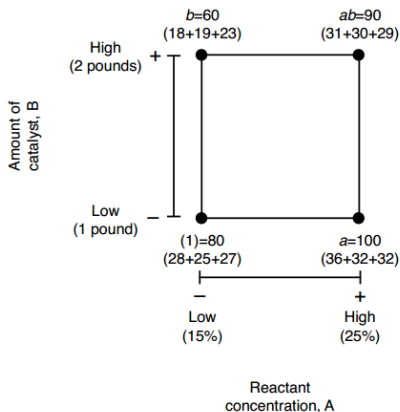
析因设计有一些特殊情况，它们广泛应用于研究工作，也是其它一些有重要实践价值的设计的基础。例如：

- 有 k 个因子，每个因子仅有两个水平。这些水平可以是定量的值，也可以是诸如 {高、低}、{有、无} 这样定性的水平。其因子组合共有 2^k 个，因而称之为 2^k 析因设计。

在实验工作早期阶段，可能有很多因子需要研究时， 2^k 设计就特别有用。它只需要最少的实验次数就可以研究完全析因设计的 k 个因子。因而，这些设计广泛地运用于因子筛选实验 (factor screening experiment)。

2² 析因设计

在最简单的 $2 \times 2 = 2^2$ 析因设计中，我们通常使用小写字母来表示不同的因子水平组合，这一记法通用于 2^k 序列。在二水平析因设计中，记号 $\{(1), a, b, ab\}$ 代表那一处理组合取 n 次重复的总和。



2² 析因设计：因子效应估计

因子的平均效应定义为，该因子的水平变化产生的响应变化在另一因子的水平上取平均值。例如，A 在 B 的低水平上的效应为 $[a - (1)]/n$ ，A 在 B 的高水平上的效应为 $[ab - b]/n$ 。因此，A 的主效应为：

$$A = \frac{1}{2n} \{[ab - b] + [a - (1)]\} = \frac{1}{2n} [ab + a - b - (1)].$$

B 的主效应类似。AB 的交互效应为：

$$AB = \frac{1}{2n} \{[ab - b] - [a - (1)]\} = \frac{1}{2n} [ab + (1) - a - b].$$

2² 析因设计：对照

由效应估计方程，可以定义 A 的对照 (contrast)：

$$\text{Contrast}_A = ab + a - b - (1).$$

因而效应 A 可以表示为：

$$A = \frac{\text{Contrast}_A}{2n}.$$

按照 $\{(1), a, b, ab\}$ 的顺序写出处理组合， $\{-1, 1, -1, 1\}$ 就称为对照 A 的对照系数。

其余效应的对照以及对照系数可以类似得出。

2² 析因设计：代数符号

通常会按照顺序记录这些组合，同时使用如下代数符号：

按顺序 (1), a , b , ab 写出处理组合常常很方便. 称这一顺序为标准顺序(或由 Frank Yates 提出的 Yates 顺序). 用标准顺序, 得出用来估计效应的对照系数是

效应	(1)	a	b	ab
A :	-1	+1	-1	+1
B :	-1	-1	+1	+1
AB :	+1	-1	-1	+1

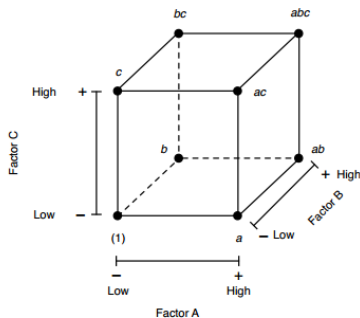
表 6.2 计算 2² 设计的效应的代数符号

处理组合	析因效应			
	I	A	B	AB
(1)	+	-	-	+
a	+	+	-	-
b	+	-	+	-
ab	+	+	+	+

图 2: 2² factorial design: notation

2³ 析因设计

在 $2 \times 2 \times 2 = 2^3$ 析因设计中, 我们类似地使用小写字母来表示不同的因子水平组合。记号 $\{(1), a, b, ab, c, ac, bc, abc\}$ 代表那一处理组合取 n 次重复的总和, 由此列出的 8 次试验方案称之为**设计矩阵** (design matrix)。



(a) Geometric view

Run	Factor		
	A	B	C
1	-	-	-
2	+	-	-
3	-	+	-
4	+	+	-
5	-	-	+
6	+	-	+
7	-	+	+
8	+	+	+

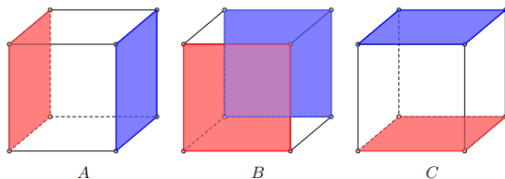
(b) Design matrix

图 3: 2^3 factorial design

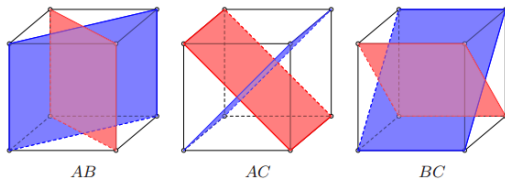
2³ 析因设计：主效应及交互效应

$$\begin{aligned}
 A &= \bar{y}_{A+} - \bar{y}_{A-} \\
 &= \frac{a + ab + ac + abc}{4n} - \frac{(1) + b + c + bc}{4n}.
 \end{aligned}$$

The effect for B and C are computed in a similar manner.



(a) Main effects



(b) Two-factor interaction

2³ 析因设计：代数符号

类似地，通常按照顺序记录这些组合，同时使用如下代数符号：

表 6.3 计算 2³ 设计的效应的代数符号

处理组合	析因效应							
	<i>I</i>	<i>A</i>	<i>B</i>	<i>AB</i>	<i>C</i>	<i>AC</i>	<i>BC</i>	<i>ABC</i>
(1)	+	-	-	+	-	+	+	-
<i>a</i>	+	+	-	-	-	-	+	+
<i>b</i>	+	-	+	-	-	+	-	+
<i>ab</i>	+	+	+	+	-	-	-	-
<i>c</i>	+	-	-	+	+	-	-	+
<i>ac</i>	+	+	-	-	+	+	-	-
<i>bc</i>	+	-	+	-	+	-	+	-
<i>abc</i>	+	+	+	+	+	+	+	+

图 5: 2³ factorial design: notation

2³ 设计的 8 个处理组合间有 7 个自由度，3 个自由度与 A、B、C 的主效应有关，4 个自由度与交互效应有关。交互效应包括 3 个二阶交互 AB，AC 和 BC，以及 1 个三阶交互 ABC。

2³ 析因设计：效应性质

以上表格有几个有趣的性质：

- 除了列 I 以外，每列加号的个数和减号的个数相等。
- 任意两列符号乘积之和为 0，亦即任意两列是正交的。
- 列 I 与任一系列相乘，该列符号不变，即列 I 是单位元素 (identity element)。
- 任意一系列与自身相乘，得到单位元素列 I，即 $A \times A = I$ 。
- 任意两列相乘，得到表中某一系列。例如 $A \times B = AB$,
 $AB \times B = AB^2 = A$ 。

2³ 析因设计：对照

由效应估计方程，可以定义 A 的对照 (contrast)：

$$\text{Contrast}_A = a + ab + ac + abc - (1) - b - c - bc.$$

因而效应 A 可以表示为：

$$A = \frac{\text{Contrast}_A}{4n}.$$

以及 A 的效应的平方和可以表示为：

$$SS_A = \frac{(\text{Contrast}_A)^2}{8n}.$$

按照 $\{(1), a, b, ab, c, ac, bc, abc\}$ 的顺序写出处理组合，相应的系数就称为对照 A 的对照系数。其余效应的对照以及对照系数可以类似得出。

2³ 析因设计：效应估计方法

可以使用方差分析方法估计 2³ 析因设计的各个效应及显著性。方差分析表如下且 $a = b = c = 2$ 。

同时，也可以使用线性回归分析方法，同时纳入主效应和交互项。

表 5.12 三因子固定效应模型的方差分析

方差来源	平方和	自由度	均 方	期望均方	F_0
A	SS_A	$a - 1$	MS_A	$\sigma^2 + \frac{bcn \sum \tau_i^2}{a-1}$	$F_0 = \frac{MS_A}{MS_E}$
B	SS_B	$b - 1$	MS_B	$\sigma^2 + \frac{acn \sum \beta_j^2}{b-1}$	$F_0 = \frac{MS_B}{MS_E}$
C	SS_C	$c - 1$	MS_C	$\sigma^2 + \frac{abn \sum \gamma_k^2}{c-1}$	$F_0 = \frac{MS_C}{MS_E}$
AB	SS_{AB}	$(a-1)(b-1)$	MS_{AB}	$\sigma^2 + \frac{cn \sum \sum (\tau\beta)_{ij}^2}{(a-1)(b-1)}$	$F_0 = \frac{MS_{AB}}{MS_E}$
AC	SS_{AC}	$(a-1)(c-1)$	MS_{AC}	$\sigma^2 + \frac{bn \sum \sum (\tau\gamma)_{jk}^2}{(a-1)(c-1)}$	$F_0 = \frac{MS_{AC}}{MS_E}$
BC	SS_{BC}	$(b-1)(c-1)$	MS_{BC}	$\sigma^2 + \frac{an \sum \sum (\beta\gamma)_{jk}^2}{(b-1)(c-1)}$	$F_0 = \frac{MS_{BC}}{MS_E}$
ABC	SS_{ABC}	$(a-1)(b-1)(c-1)$	MS_{ABC}	$\sigma^2 + \frac{n \sum \sum \sum (\tau\beta\gamma)_{ijk}^2}{(a-1)(b-1)(c-1)}$	$F_0 = \frac{MS_{ABC}}{MS_E}$
误差	SS_E	$abc(n-1)$	MS_E	σ^2	
总和	SS_T	$abcn - 1$			

2^k 析因设计：概述

推广到 2^k 析因设计的情形，其统计模型包含 k 个主效应， $\binom{k}{2}$ 个二因子交互作用， $\binom{k}{3}$ 个三因子交互作用，...，以及一个 k 因子交互作用，全模型共有 $2^k - 1$ 个效应。

例如， $k = 6$ 时，这一 2^6 析因设计的单次完全重复需要做 64 次试验。在此设计中，共有 63 个自由度，其中 6 个与主效应对应，15 个与二因子交互作用对应，而其余 42 个自由度则与三阶及以上的交互作用对应。

2^k 析因设计：对照

尽管这些效应的计算大多数可以利用计算机进行，但有时手工计算效应估计或效应的平方和也是必要的。对于较大的 k 值，可以用展开下式右边的方法：

$$\text{Contrast}_{AB\dots K} = (a \pm 1)(b \pm 1)\dots(k \pm 1).$$

在展开上式时，按初等代数方法计算，而在最后的表达式中，用 (1) 代替 1 。小括号中符号的选取方法是：当因子包含在效应中时，取负号；不在时，取正号。

例如， 2^3 析因设计中 AB 的对照是：

$$\text{Contrast}_{AB} = (a - 1)(b - 1)(c + 1) = abc + ab + c + (1) - ac - bc - a - b.$$

2^k 析因设计：对照（续）

进一步，在 2^5 设计中，ABCD 的对照是：

$$\text{Contrast}_{ABCD} = (a - 1)(b - 1)(c - 1)(d - 1)(e + 1).$$

一旦计算出效应的对照，就可以根据下式估计效应并计算平方和：

$$AB...K = \frac{1}{2^{(k-1)} \times n} (\text{Contrast}_{AB...K})$$

以及

$$SS_{AB...K} = \frac{1}{2^k \times n} (\text{Contrast}_{AB...K})^2.$$

Section 2

分式析因设计

分式析因设计：概述

当 2^k 析因设计中的因子个数 k 增加时，设计的一次完全重复所需的试验次数 2^k 迅速增大，以至于超出大多数实验者拥有的资源。例如， $2^{10} = 1024$ 。然而，这 $2^k - 1$ 个自由度之中，仅有 k 个与主效应对应，仅有 $\binom{k}{2}$ 个与二阶交互作用对应，而其余均是三阶及以上的交互作用。

如果实验者可以合理地假定某些高阶交互作用可被忽略，则关于主效应和低阶交互效应的信息就可以只做完全析因实验的一部分而求得。这一设计方法称为分式析因设计 (fractional factorial designs)。

分式析因设计：概述（续）

分式析因设计主要用于**因子筛选实验** (factor screening experiment)。因子筛选实验的目的是，在众多因子中识别出效应较大的那些因子。它通常在项目的早期阶段实施，那时，一开始考虑的很多因子有可能对响应只有小的效应或没有效应。因此，需要识别出重要因子，在随后的实验中更深入地探讨。

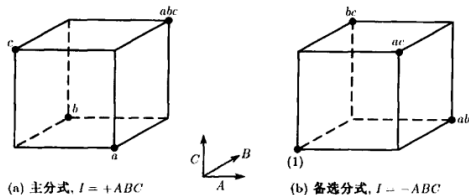
分式析因设计的成功应用基于**3 个关键思想**：

- **效应稀疏原理**：当有很多变量时，系统或过程很可能被少量几个主效应和低阶交互效应所主导。
- **投影性质**：分式析因设计可以投影到由显著性因子的子集组成的设计中去。
- **序贯实验**：可以将两个或多个分式析因设计序贯地组合为一个设计，用来估计所感兴趣的因子效应和交互效应。

2^{3-1} 分式析因设计

表 8.1 2^3 析因设计的加减符号表

处理组合	因子效应							
	I	A	B	C	AB	AC	BC	ABC
a	+	+	-	-	-	-	+	+
b	+	-	+	-	-	+	-	+
c	+	-	-	+	+	-	-	+
abc	+	+	+	+	+	+	+	+
ab	+	+	+	-	+	-	-	-
ac	+	+	-	+	-	+	-	-
bc	+	-	+	+	-	-	+	-
(1)	+	-	-	-	+	+	+	-

图 8.1 2^3 析因设计的两个 $1/2$ 分式设计图 7: 2^{3-1} fractional factorial design

2^{3-1} 分式析因设计：定义关系

2^{3-1} 设计由选取 ABC 列中带加号的处理组合组成。换言之，这一分式析因设计的定义关系 (defining relation) 是

$$I = ABC.$$

因为，单位列 I 全是加号，而我们选取了 ABC 列中带加号的处理组合。

如此，ABC 称为这一特定分式析因设计的生成元 (generator)。 2^{3-1} 分式析因设计有 3 个自由度，可以用来估计 3 个主效应。例如，

$$[A] = \frac{1}{2}(a - b - c + abc).$$

2^{3-1} 分式析因设计：别名

同时，二因子交互效应 BC 的估计为：

$$[BC] = \frac{1}{2}(a - b - c + abc).$$

因此， $[A] = [BC]$ 。类似地， $[B] = [AC]$ ， $[C] = [AB]$ 。因此，不可能区分这些效应。当我们估计 A, B, C 时，实际是在估计 $A+BC$, $B+AC$, $C+AB$ 。具有这一性质的两个或多个效应称为**别名** (alias)。

别名可由定义关系 $I = ABC$ 推导而来，即

$$A = AI = A^2BC = BC.$$

本例中，A 和 BC 互为别名。因而，符号上记为：

$$[A] \rightarrow A + BC; [B] \rightarrow B + AC; [C] \rightarrow C + AB.$$

2^{3-1} 分式析因设计：构造方法

2^{3-1} 分式析因设计的构造方法是，先写出由完全 2^2 析因设计的试验组成的基本设计 (basic design)。然后使用定义关系，获得因子 C 的设计水平。

试验	完全的 2^2 析因设计 (基本设计)		$2^{3-1}_{III}, I = ABC$		
	A	B	A	B	$C = AB$
1	-	-	-	-	+
2	+	-	+	-	-
3	-	+	-	+	-
4	+	+	+	+	+

图 8: 2^{3-1} design: construction

2^{3-1} 分式析因设计：投影性质

2^{3-1} 分式析因设计，可以投影为 3 个 2^2 完全析因设计。

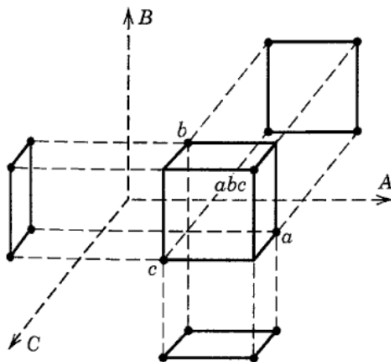


图 9: 2^{3-1} design: projection

2^{5-1} 分式析因设计：案例描述

在一集成电路生产线上用 2^{5-1} 设计来研究 5 个因子，以改进生产量。5 个因子及水平是：

变量	名称	-1 水平	+1 水平
<i>A</i>	孔径设定	小	大
<i>B</i>	曝光时间	-20%	+20%
<i>C</i>	冲洗时间	30 s	40 s
<i>D</i>	掩膜尺寸	小	大
<i>E</i>	蚀刻时间	14.5 min	15.5 min

图 10: factors and levels

拟试验 16 次，应如何做 2^{5-1} 分式析因设计？别名关系是怎样的？

2^{5-1} 分式析因设计：实验设计

类似地，先写出 2^4 析因设计的试验组成的基本设计。然后使用定义关系，获得因子 E 的设计水平。

试验	基本设计				$E = ABCD$	处理组合	产量
	A	B	C	D			
1	-	-	-	-	+	e	8
2	+	-	-	-	-	a	9
3	-	+	-	-	-	b	34
4	+	+	-	-	+	abe	52
5	-	-	+	-	-	c	16
6	+	-	+	-	+	ace	22
7	-	+	+	-	+	bce	45
8	+	+	+	-	-	abc	60
9	-	-	-	+	-	d	6
10	+	-	-	+	+	ade	10
11	-	+	-	+	+	bde	30
12	+	+	-	+	-	abd	50
13	-	-	+	+	+	cde	15
14	+	-	+	+	-	acd	21
15	-	+	+	+	-	bcd	44
16	+	+	+	+	+	abcde	63

图 11: design matrix

2^{5-1} 分式析因设计：效应估计

使用对照，可以方便地估计各个效应及平方和：

变量	回归系数	效应估计	平方和
总平均值	30.312 5		
<i>A</i>	5.562 5	11.125 0	495.062
<i>B</i>	16.937 5	33.875 0	4 590.062
<i>C</i>	5.437 5	10.875 0	473.062
<i>D</i>	-0.437 5	-0.875 0	3.063
<i>E</i>	0.312 5	-0.625 0	1.563
<i>AB</i>	3.437 5	6.875 0	189.063
<i>AC</i>	0.187 5	0.375 0	0.563
<i>AD</i>	0.562 5	1.125 0	5.063
<i>AE</i>	0.562 5	1.125 0	5.063
<i>BC</i>	0.312 5	0.625 0	1.563
<i>BD</i>	-0.062 5	-0.125 0	0.063
<i>BE</i>	-0.062 5	-0.125 0	0.063
<i>CD</i>	0.437 5	0.875 0	3.063
<i>CE</i>	0.187 5	0.375 0	0.563
<i>DE</i>	-0.687 5	-1.375 0	7.563

图 12: effect estimates

2^{5-1} 分式析因设计：方差分析

根据各个效应和平方和的情况，辅以合理推断，可以忽略一些效应较小的效应，尤其是高阶效应。从而，可以得到如下方差分析表。因而，可以在后续实验中，在 A、B、C 三个因子上设置更多的水平，来探讨其效应。

方差来源	平方和	自由度	均 方	F_0	P 值
A(孔径设定)	495.062 5	1	495.062 5	193.20	< 0.000 1
B(曝光时间)	4 590.062 5	1	4 590.062 5	1 791.24	< 0.000 1
C(冲洗时间)	473.062 5	1	473.062 5	184.61	< 0.000 1
AB	189.062 5	1	189.062 5	73.78	< 0.000 1
误差	28.187 5	11	2.562 5		
总和	5 775.437 5	15			

图 13: ANOVA: 2^{5-1} design

Section 3

离散选择实验

经典的现实问题

请思考面临的一些现实决策问题：

- 出差时，选择什么交通工具到天河机场？准备预订什么酒店？
- 应当购买什么商业健康保险？或者还是不要购买？
- 患者去哪一家医院就医？
- 应当接种哪一种 HPV 疫苗？或者还是暂时不接种 HPV 疫苗？
- ...

决策与（离散）选择









Hebert Simon 认为:

Decision making is the heart of administration, and the vocabulary of administrative theory should be derived from the logic and psychology of human choice. – By Hebert Simon

或者决策的通行定义:

A decision is choice out of several alternatives (options) made by the decision maker to achieve some objectives in a given situation.

手机消费例子

	Levels			
Attribute	1	2	3	4
<i>Style</i>	Candy bar 	Slide phone 	Flip phone 	Touch screen 
<i>Brand</i>	Blackberry 	Nokia 	LG 	Samsung 
<i>Talk-time</i>	3 h	5 h	7 h	9 h
<i>Weight</i>	100 g	115 g	130 g	145 g
<i>Camera quality</i>	2 Mega pixels	3 Mega pixels	6 Mega pixels	8 Mega pixels

Below is a description of the fixed specifications of your Smart Phone:

Telecom: **Data services** WAP, E-Mail, WWW, SMS, MSN Messenger, MMS; **Modem** Integrated Wireless cellular modem; and **Cellular enhancement protocol** EDGE, GPRS, GSM

Operating System/Software: **OS provided** Microsoft Windows Mobile 6.0 Professional; **Software included** GPS Catcher, Alarm, ASUS Zip, Contacts, Location Courier, Windows Live, Business Card Recognition, Backup, ActiveSync, Microsoft Word Mobile, Microsoft PowerPoint Mobile, Notes, Windows Media Player, Travelog, Auto Cleaner, Meeting Time Planner Newstation, Calendar, Microsoft Excel Mobile, Calculator, Microsoft Internet Explorer Mobile, ASUS Remote Presenter, MSN Messenger, File Explorer

Memory: **Installed RAM** 64 MB; **RAM technology** SDRAM; **Installed ROM** 128 MB Flash Phone: **Call features** Call hold, Call history, Call forwarding

Processor: **Processor** Texas Instruments 200 MHz OMAP1.850

Input Device: **Input device type** Jog Dial, Keypad, Stylus, 5-way navigation button

图 14: smartphone consumption decision

疫苗接种决策例子

	Plan A	Plan B
Effectiveness	50	90
Protective duration (Month)	18	6
Adverse event	Fever 1-2 days	None
Frequency of injection	3	2
Price (CNY)	200	0
Place of origin	Imported	Domestic
Which one is more attractive	<input type="checkbox"/>	<input type="checkbox"/>

图 15: vaccination decision

- 候选项 (alternatives)
- 属性 (attributes)

有限理性假设

管理即是决策，而决策即是评估少数几个候选项 (alternatives)，并从中作出选择 (choice)。其心理学基础是：

- 决策资源是有限的，决策者是有限理性的
- 决策资源：时间成本，经济成本，认知资源，信息
- 信息时代的挑战：信息过载导致了注意力（亦即认知资源）稀缺

Daniel McFadden 与离散选择模型

美国计量经济学家 Daniel McFadden 与 James Heckman 共同获得 2000 年诺贝尔经济学奖。Daniel McFadden 的获奖理由是“因为他发展了分析离散选择的理论和方法” (“for his development of theory and methods for analyzing discrete choice”)。

离散选择模型 (discrete choice models, DCM) 始于交通领域，后来逐渐扩展到其他个人和家庭层面的决策行为领域。而当纳入价格变量之后，又随后扩展到**支付意愿** (willingness to pay, WTP) 研究领域。

如何决策：效用最大化原理

第 i 个决策者（例如消费者、患者等）面临 $J \geq 2$ 个候选项，仔细评估了这些候选项之后，选择了第 j ($1 \leq j \leq J$) 个候选项。我们应该如何理解这一决策？

经济学解释：假设第 j ($1 \leq j \leq J$) 个候选项对于第 i 个决策者而言，效用为 U_{ij} 。那么，决策者所选择的候选项，必定是对他而言效用最大的，即**效用最大化原理**：

$$U_{ij} \geq U_{ik}, 1 \leq k \leq J \text{ and } k \neq j.$$

决策者视角 vs 研究者视角

我们需要区分两个视角：

- 决策者视角：总效用 U_{ij} 是已知的。
- 研究者视角：仅能观测到部分效用，而无法观测到总效用。

因而，就研究者视角而言，需要将总效用分解为两部分：

$$U_{ij} = V_{ij} + \epsilon_{ij}.$$

其中 V_{ij} 是可以观测到的效用 (observed utility)，而误差项 ϵ_{ij} 是不可观测的效用 (unobserved utility)。

选择概率

由于研究者无法观测到总效用，因而可观测效用最大的候选项，就**不一定是**决策者所选择的候选项。因而，需要进一步假定，来处理观测到的选择（observed choice） j 和可观测部分效用 V_{ij} 的关系。

当误差项 ϵ_{ij} 服从 Gumbel 分布（或称广义极值 I 型分布）时，第 i 个决策者选择第 j 个候选项的**概率**为：

$$\text{Prob}(ij) = \frac{\exp(V_{ij})}{\sum_{k=1}^J \exp(V_{ik})}.$$

这一概率用于后续极大似然估计法中的**似然函数构造**。

如何评估候选项：效用函数

决策者面临 $J \geq 2$ 个候选项时，是**如何评估**这些候选项的？换言之，如何构造效用函数？

效用函数 (utility function) 是**候选项属性**的函数，假定每个候选项有 P 个属性， $X_{jp} (1 \leq p \leq P)$ ，那么可观测部分效用 V_{ij} 通常建模为属性的线性函数：

$$V_{ij} = \sum_{p=1}^P \beta_{ip} \cdot X_{jp}.$$

效用函数计算示例

假定决策者考虑疫苗接种，并且评估两款疫苗的两个属性：保护效力和副作用。疫苗 A 的保护效力为 70%，副作用几乎没有；疫苗 B 的保护效力为 90%，有中等程度的副作用。决策者赋予保护效力和（中等程度）副作用的权系数分别为 10 和 -3，那么：

$$V_A = 10 \times 70\% + 0 = 7, V_B = 10 \times 90\% - 3 = 6.$$

因此，决策者接种疫苗 A 的概率是：

$$\text{Prob}(A) = \frac{\exp(7)}{\exp(7) + \exp(6)} = 0.73.$$

什么是重要的：属性偏好

决策者在评估这些候选项的时候，特定的属性发挥了什么作用？或者说，属性和效用是什么关系？

决策者赋予属性的权系数称为**偏好** (preferences)，即 β_{ip} 。偏好决定了属性如何构成效用。偏好是离散选择实验中需要估计的**参数**。

疫苗接种决策中，对潜在副作用更加在意的个体，其副作用的属性偏好也就越大。

偏好异质性：汝之蜜糖，彼之砒霜

不同的决策者（例如消费者、患者等），各自看重不同的属性。换言之，决策者对特定属性的偏好通常是**异质的** (heterogeneous)。

研究者可以权衡研究目标，对偏好做出进一步假定，通常包括三个方案：

- ① 偏好是同质的： $\beta_{ip} = \beta_p$ ，此时即是**多项 logit 模型**。
- ② 偏好在给定分组下是同质的：若第 i 个决策者属于第 s ($1 \leq s \leq S$) 组，则 $\beta_{ip} = \beta_{sp}$ ，此时即是**潜类别 logit 模型**。
- ③ 偏好是异质的： β_{ip} 取决于个体特征，此时是**随机参数 logit 模型**；可以进一步采用分层 logit 模型来解释偏好异质性的来源。

离散选择实验的微观经济学基础回顾

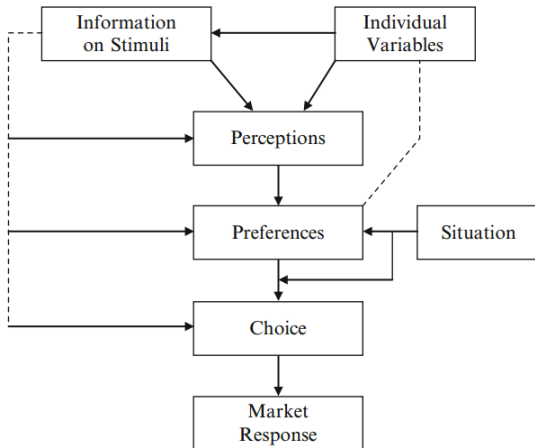


图 16: DCE decision-making framework

离散选择实验设计案例

考虑疫苗接种问题。假定我们纳入四个因子，每个有 2 个水平：

- 疫苗有效性：70% vs 90%
- 保护期限：6 个月 vs 12 个月
- 疫苗副作用：局部疼痛 vs 1~2 天内发热
- 价格：300 元 vs 500 元

请设计离散选择实验。