

试验设计 (DOE)

概述

“协助”DOE 包括核心 Minitab 中所提供的 DOE 功能的子集，并使用一个连续的试验过程来简化设计的创建和分析过程。该过程首先筛选设计以确定最重要的因子。然后，我们提供分辨率更高的设计以查找曲率，并确定可用于识别对响应进行优化的因子设置的最终模型。

在本文中，我们概述了试验过程的各个步骤。我们提供有关如何确定“协助”中提供哪些设计的信息，包括功效的作用。我们还讨论了检测过程和数据中的拟合曲率。本文还介绍了用于分析数据并确定最佳模型的方法。

本文还提供了有关在“协助报告卡”中进行以下数据检查的更多信息：

- 区组
- 异常数据
- 检测能力

方法

连续试验过程

“协助”中的 DOE 功能将引导用户，通过一个连续的过程来设计和分析一个或多个试验，从而确定最重要的因子，并找到可优化响应的因子设置。这种连续的试验方法使用一组较小的试验数据，其中每个阶段的试验结果均将引导用户通往下一阶段的试验。此连续方法的一个优点是，在每一个阶段，只会执行少量实验性试验，从而不太可能在非生产性试验阶段浪费资源。

“协助”以结构化的形式提供核心 Minitab 中所包含的 DOE 功能的子集，可简化设计的创建和分析过程。此过程中的步骤如下：

1. 为 6 至 15 个因子创建筛选设计。
2. 拟合包括主效应的筛选模型并分析结果，以找出最重要的因子。
3. 根据步骤 2 的结果（其中包括 2-5 个最重要的因子）创建建模设计。
4. 拟合包括主效应和双向交互作用的线性模型并分析结果，然后查找因子和响应关系中的曲率证据。
5. 如果在步骤 4 中未检测到曲率，则使用该模型确定可优化响应的因子设置。
6. 如果在步骤 4 中检测到曲率，“协助”将建议您在设计中添加曲率点。
7. 拟合包括平方项的二次模型，以创建曲率模型并分析结果。
8. 使用最终模型，确定可优化响应的因子设置。

以下各节提供有关“协助”DOE 在下述方面的更详细的信息：

- 筛选设计
- 建模设计
- 拟合模型

筛选设计

通常情况下，您将根据大量的潜在因子开始连续的试验过程，然后淘汰那些对响应效应不大的因子。筛选设计是一种实验性设计，旨在从一大组因子中找出几个最重要的因子。在“协助”中，将为 6 至 15 个因子提供筛选设计。

设计类型

“协助”中的筛选设计是 Plackett-Burman 设计（一种特殊类型的 Resolution III 2 级设计）。Plackett-Burman（另类筛选因子）设计有两个主要优点：

- 它们允许使用非常少的实验性试验（少至 12 个）来估计因子的主效应。因为实验性试验的执行成本可能比较高昂，所以这使得这些设计更具成本效益。
- 主效应和二因子交互作用之间只有部分或少量混杂设计。不能彼此独立进行估计的效应称为混杂。在 Plackett-Burman 设计中，只有一部分混杂设计，因为每个效应的贡献只是全部交互作用的一小部分。

我们确定，为了进行筛选，合理的方法是使用仅估计主效应而不估计交互项的 Plackett-Burman 设计。筛选设计旨在包括大量因子。模型中的每一项至少需要运行一次试验，而且交互项的数量增长速度比主效应快，因此在大多数情况下，将模型与交互项拟合不切实际或不具有成本效益。此外，在大多数情况下，只有少数因子可以解释对响应的大部分效应。筛选设计的目标是要找出这些因子，而 Plackett-Burman 设计能够让用户识别出这些重要的主效应。此外，如前所述，因为 Plackett-Burman 设计中只有部分混杂设计，显著的主效应不太可能就是显著的二因子交互作用。

功效和折叠

当我们创建设计目录时，我们的目标只是让现有的设计具有足够的功效。我们计算了所有设计的功效并淘汰了某些功效较低的设计，包括适用于 10 或 11 个因子的 12 次 Plackett-Burman 实验设计。对于具有 10 或 11 个因子的设计，只有 20 次 Plackett-Burman 实验设计可用。我们还淘汰了功耗低和试验次数较多的 16、17 和 18 个因子的设计。有关设计的特定功效的更多信息，请参见“检测能力”部分。

对于具有 6 至 9 个因子的设计，我们允许折叠，这将增加实验性试验的次数，从而提高设计的精度和功效。在某些情况下，最好增加设计的试验次数，以提高检测到重要效应的可能性。通过折叠，新的试验将被添加到设计中，其中部分或全部因子水平可通过切换因子的高低水平来反转。折叠还可消除主效应和二因子交互作用之间的部分混杂，从而减小受混杂影响的主效应估计的偏差。“创建筛选设计总结报告”中的“检测能力”部分提供相关信息，以帮助用户确定设计是否有足够的功效检测到足够大的效应。

建模设计

确定 2 至 5 个重要因子后，Minitab 便建议创建一个建模设计，以获得一个模型，用于识别可优化响应的因子设置。

设计类型

2 至 5 个因子的建模设计都是全因子或分辨率 V 设计。这些设计可用于拟合所有主效应和二因子交互项，而不会在两项之间产生任何混杂。部分或全部高阶项（例如，三因子交互作用）可以与模型中的项混杂。然而，与主效应和二因子交互项相比，高阶项通常可以假定为忽略不计。

当我们创建设计目录时，我们的目标只是让现有的设计具有足够的功效。因此，我们淘汰了 4 次二因子实验设计，改用一个重复的 4 次二因子实验设计。

中心点和建模曲率

“协助”中的建模设计还包括中心点，检查数据中存在的曲率。所有的连续因子都在这些点的高低设置值之间设置。如果没有曲率，则中心点的平均响应等于各因子高低设置值的平均响应的平均值（设计空间的各个角）。当中心点的平均响应的平均值显著大于或小于各因子高低设置值的平均响应的平均值时，系统会检测到曲率。

虽然中心点可以检测到曲率，但它们不提供足够的信息来创建曲率模型。要创建曲率模型，需要平方项，从而需要在设计中添加更多的点。这些额外的点会将设计转换为以面为中心的复合设计。这是一种响应表面设计，从而可以拟合具有线性主效应的二次模型、所有二因子相互作用以及所有连续因子的平方项。

使用后向选择的拟合模型

我们探讨了拟合模型的几种方法，并确定使用 α 值为 0.10 的后向选择是最好的办法。当您拟合模型时，Minitab 首先会包括所有可能的项。然后，Minitab 会逐个移除最不显著的项，同时维护模型的层次结构。层次结构意味着，如果交互项显著，则形成交互的两个因子的线性项也必须包括在模型中。这是一种后向选择，是为了自动执行通常通过手工完成的模型选择过程。在“协助”DOE 的所有设计中，所有项都各自独立（比如平方项）或近似独立。因此，因子之间彼此关联的多重共线性不太可能发生。多重共线性将导致逐步搜寻程序错过最佳模型。用 $\alpha = 0.10$ ，而不是常用的 $\alpha = 0.05$ ，有助于提高检验的功效，从而提高重要项保留在模型中的可能性。

数据检查

区组

实验性设计中将使用区组，以尽量减少由试验的外部影响导致的偏差和误差，如噪声变量、缺少变量，或者每个区组中的试验执行方式的差异。通过将一起执行的实验性试验放入区组，您可以确定区组之间是否存在差异，并在模型中考虑这些差异。

目标


在“协助”DOE 中，您可以在创建建模设计时包括设计的仿行，也可以为建模设计添加轴向点，以拟合模型曲率。通常，仿行和轴向点的执行时间或条件不同于基本设计中的试验。当在不同的时间或条件下执行试验时，最好先考虑不同的条件可能产生的效应。

方法

为了解释实验条件下，仿行或轴向点和基本设计可能存在的差异，Minitab 会将仿行和轴向点放入不同的区组。具体而言，在建模设计中，Minitab 会将基本设计的仿行放入模型中的各个区组中。在二次设计时，Minitab 会将用于检测曲率的轴向点设计放在单独的区组中。

结果

为了与模型中其他项的处理方式保持一致，将使用后向淘汰法评估区组。报告卡指示区组项是否统计显著，从而表明整个区组中都存在差异。如果区组之间存在差异，则考虑调查原因，以确定实验条件或程序是否存在任何不一致。

状态	条件
	<p>区组在最终模型中</p> <p>区组显著。因为每个区组中的试验通常在不同的时间段执行，区组中的显著差异表明条件可能已随着时间的推移发生了改变。这种差异可能由试验的外部影响导致，如噪声变量、缺少试验中应包括的变量，或者每个区组中的试验执行方式的差异。考虑调查区组之间存在差异的原因。</p> <p>区组不在最终模型中</p> <p>区组不显著。因为每个区组中的试验通常在不同的时间段执行，此结果表明，没有证据表明，随着时间的推移，实验条件会存在差异。</p>

异常数据

在“协助”DOE 程序中，我们将异常数据定义为大型标准化残差的观测值，这是一种用于识别模型拟合过程中的异常数据的通用方法 (Neter et al., 1996)。由于异常数据会对结果产生巨大的影响，您可能需要更正数据，使分析有效。

目标

我们想要确定标准化残差必须多大才表明数据点异常。



方法

我们根据 Minitab 中的 DOE 标准程序（统计 > DOE > 因子 > 分析因子设计和统计 > DOE > 响应面 > 分析响应面设计），制定了用于识别异常观测值的准则。

结果

标准化残差等于残差值 e_i 除以其标准差的估计值。一般来说，如果标准化残差的绝对值大于 2，观测值将被认为异常，不过这项准则有些保守。您可能希望所有大数据集观测值中正好约有 5% 的几率满足这一标准（如果错误数据呈正态分布）。然而，对于小型实验数据集，几乎不可能标记任何观测值，最好调查出现异常值的原因。

当检查异常数据时，“协助报告卡”会显示以下状态指标：

状态	条件
	无标准化残差 ≥ 2 没有异常数据点。异常数据会对结果产生巨大的影响。
	一个标准化残差 ≥ 2 一个数据点有一个大残差，无法通过模型实现良好的拟合。该点在诊断报告及工作表的第 X 行中被标记为红色。因为异常数据点会对结果产生巨大的影响，请尽量找出异常原因。纠正任何数据输入或测量错误。考虑再次执行与特殊原因有关的试验并重新进行分析。 多个标准化残差 ≥ 2 X 数据点具有较大的残差，无法通过模型实现良好的拟合。这些点在诊断报告上被标记为红色。您可以将鼠标悬停在某一点上或使用 Minitab 的笔刷功能识别工作表中的行。因为异常数据点会对结果产生巨大的影响，请尽量找出异常原因。纠正任何数据输入或测量错误。考虑再次执行与特殊原因有关的试验并重新进行分析。

检测能力

进行设计性实验时，有必要在收集数据之前，了解设计可能检测到的效应值。如果设计功能不足以检测到所需的效应值，可能需要在设计中包括更多试验。不过，由于在设计中包括多个试验的成本比较高昂，确定是否有必要增加功效很重要。

目标

我们希望向用户提供有关设计可以 60% 和 80% 的功效水平检测到的效应值的信息。我们还希望向用户提供有关包括当前可用的其他试验的设计的效应值信息。对于 6 至 9 个因子的筛选设计，用户可以选择在其设计中包括 12 或 24 个试验。对于建模设计，用户可以包括基本设计的仿行，来增加设计中的试验总数。

方法

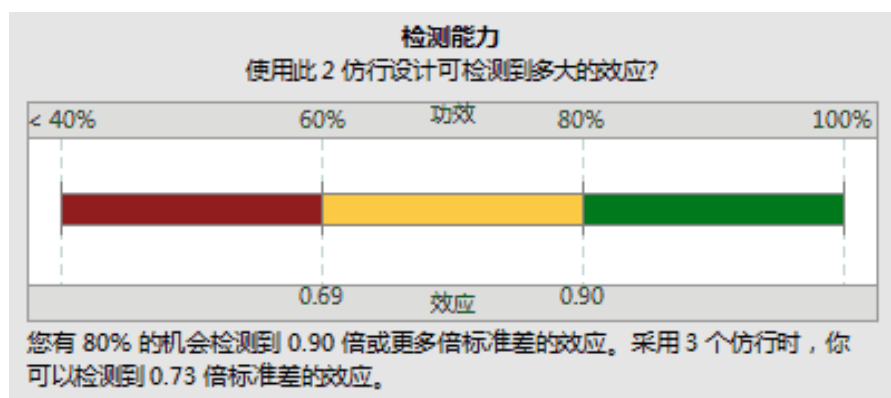
我们计算出可以在“协助”中检测到的每个设计的功效和效应值。功效是发现统计显著的因子效应的概率。效应值采用标准差单位。

结果

“总结”报告显示可以通过功效为 60% 和 80% 的设计检测到的效应值。对于有一个大型（折叠）设计可用的筛选设计，该报告还将指出可以通过大型设计中 80% 的功效检测到的效应值。对于提供更多仿行的建模设计，该报告指出可以用附加复制 80% 的功效检测到的效应值。然后，用户可以判断所选设计是否合适，并权衡使用具有更多可用试验的设计的利弊。

请参见附录 A，了解有关每个设计可以 60% 和 80% 的功效检测到的效应值的具体信息。

下图是“总结”报告中提供的功效信息的示例。



参考书

Neter, J., Kutner, M.H., Nachtsheim, C.J., & Wasserman, W. (1996). *Applied linear statistical models*. Chicago: Irwin.

附录 A：检测能力

我们计算出可以在“协助”中检测到的每个设计的功效和效应值。功效是发现统计显著的因子效应的概率。效应值采用标准差单位。

与模型项有关的效应值是真正的模型方程中项系数的两倍。在筛选模型中，效应值有一个简单的解释：当因子从低水平变成高水平时，平均值响应中发生的变化。

表 1 下表显示“协助”中提供的筛选设计的效应值。

因子	试验次数	功效为 60% 的效应	功效为 80% 的效应
6	12	1.27325	1.67693
6	24	0.80721	1.05805
7	12	1.3282	1.75498
7	24	0.80936	1.06092
8	12	1.43101	1.90493
8	24	0.8118	1.06420
9	12	1.68682	2.29728
9	24	0.81462	1.06797
10	20	0.919135	1.20607
11	20	0.928949	1.21945
12	20	0.941923	1.23725
13	20	0.959863	1.26206
14	20	0.986258	1.29895
15	20	1.02882	1.3594

表 2 下表显示“协助”中提供的建模设计的效应值。

因子总数	类别因子	仿行	功效为 60% 的效应	功效为 80% 的效应
2	0	2	1.517676	1.995488
2	0	4	0.991255	1.299356
2	0	6	0.795584	1.042572
2	1	2	1.489312	1.956272

因子总数	类别因子	仿行	功效为 60% 的效应	功效为 80% 的效应
2	1	4	0.986308	1.292747
2	1	6	0.79336	1.039623
2	2	2	1.626711	2.149402
2	2	4	1.012132	1.32742
2	2	6	0.805317	1.055524
3	0	1	1.752624	2.333048
3	0	2	1.001723	1.313394
3	0	3	0.79841	1.046325
3	0	4	0.685061	0.89768
3	0	5	0.609738	0.798946
3	1	1	1.626711	2.149402
3	1	2	0.994252	1.303368
3	1	3	0.795584	1.042572
3	1	4	0.683497	0.895612
3	1	5	0.608716	0.797597
3	2	1	1.468798	1.928128
3	2	2	0.977848	1.281481
3	2	3	0.788844	1.033647
3	2	4	0.679641	0.890522
3	2	5	0.606149	0.794214
3	3	1	3.804252	5.7928
3	3	2	1.038597	1.363392
3	3	3	0.811803	1.064195
3	3	4	0.692413	0.907434
3	3	5	0.614534	0.805288
4	0	1	1.053102	1.383293
4	0	2	0.689744	0.903887

因子总数	类别因子	仿行	功效为 60% 的效应	功效为 80% 的效应
4	0	3	0.556612	0.729334
4	0	4	0.47976	0.628615
4	0	5	0.428010	0.560802
4	1	1	1.038597	1.363392
4	1	2	0.688304	0.901977
4	1	3	0.556027	0.728562
4	1	4	0.479427	0.628176
4	1	5	0.427789	0.560511
4	2	1	1.006462	1.319772
4	2	2	0.684233	0.896585
4	2	3	0.554302	0.726288
4	2	4	0.478427	0.626861
4	2	5	0.427119	0.559631
4	3	1	0.982394	1.287529
4	3	2	0.679988	0.890980
4	3	3	0.552383	0.723762
4	3	4	0.477284	0.625358
4	3	5	0.426341	0.558609
4	4	1	1.102670	1.452267
4	4	2	0.694658	0.910421
4	4	3	0.558674	0.732059
4	4	4	0.480955	0.630190
4	4	5	0.428812	0.561858
5	0	1	1.460831	1.989497
5	0	2	0.694658	0.910421
5	0	3	0.557797	0.730899
5	0	4	0.480244	0.629252

因子总数	类别因子	仿行	功效为 60% 的效应	功效为 80% 的效应
5	0	5	0.428261	0.561133
5	1	1	1.239292	1.649714
5	1	2	0.692413	0.907434
5	1	3	0.557051	0.729913
5	1	4	0.47985	0.628733
5	1	5	0.428010	0.560802
5	2	1	1.053102	1.383293
5	2	2	0.686516	0.899606
5	2	3	0.554925	0.727108
5	2	4	0.478694	0.627212
5	2	5	0.427261	0.559817
5	3	1	0.994252	1.303368
5	3	2	0.680992	0.892303
5	3	3	0.552683	0.724156
5	3	4	0.477418	0.625533
5	3	5	0.426414	0.558704
5	4	1	0.970149	1.271267
5	4	2	0.676819	0.886805
5	4	3	0.550801	0.721681
5	4	4	0.476297	0.624062
5	4	5	0.425652	0.557704
5	5	2	0.703042	0.92162
5	5	3	0.560538	0.734525
5	5	4	0.481695	0.631166
5	5	5	0.429191	0.562356

© 2015, 2017 Minitab Inc. All rights reserved.
Minitab®, Quality. Analysis. Results.® and the Minitab® logo are all registered trademarks of Minitab, Inc., in the United States and other countries. See minitab.com/legal/trademarks for more information.