

太好了,我们已经改进了...还是没有?

吴伟博士

许多公司每年在客户调查上要用掉大量的钱。客户调查结果的结果被用来修正战略,设计新产品和服务,聚焦于改进活动...以及庆祝成功。由于客户服务结果的影响非常重要,数据驱动的结论应该是可信的。问题是:我们总能够信赖我们看到的吗?

一家寿险公司-让我们把它称为 MyInsurance -有着全球的市场,在 2006 年庆祝他们改进客户满意的成功。他们骄傲的展示结果:“在泰国,和 2005 年只有 54% 的客户满意相比较,我们获得了 58% 的客户满意。”这听起来很好。对吧?在市场上数以百万的客户市场,4% 的满意度增长意味着,愿意购买 MyInsurance 的回头客的数量已经增长了几万。

这样的结论会很快得到。为什么?显而易见的原因,MyInsurance 没有真正的向数百万的客户询问他们的意见和建议。他们紧紧管理从 280 个客户收集意见和建议。而且,这被叫做样本。这样的方法每天被多次应用在每种行业的每个公司。

样本

样本是基于客户比较小的数量,并且者被用来制定关于“全体”的结论。在这个例子中,全体是指我们感兴趣的所有的客户库。样本有一个大优点:它既省钱又省时间,而且尤其是当测试的流程毁坏测试对象的时候,比如手机的跌落试验。这个优点和缺陷并存:那就是“误差幅度”或者“置信区间”。

误差幅度-置信区间

置信区间是一个范围,是我们期望可能的总体值的集合。由于我们研究样本,我们只能猜测“真值”是什么。在样本中,我们从不知道。这个置信区间不能被避免,即使是在“理想情况”下的极好的样本。然而,这个区间可以随着样本量的增加和在总体中减小变异而被减小。因此,唯一的选择是对期望的置信区间决定最小的样本数量。

在 MyInsurance 实例中这意味着什么呢?我们可以基于我们的样本量简单的统计计算出置信区间:

2005 年,“真实”的客户满意度等级是 48%到 60%之间。

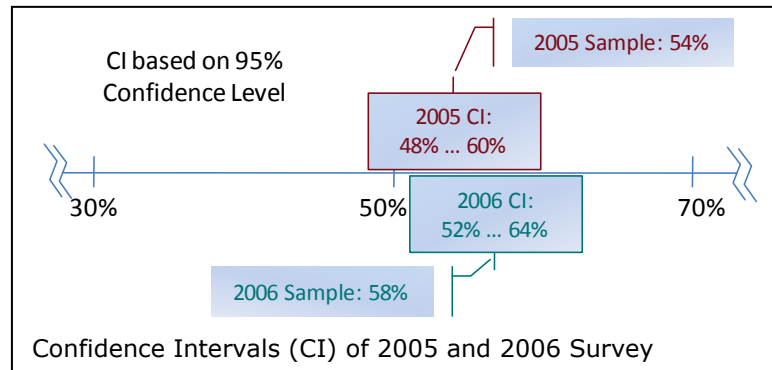
2006 年,这个值基于 52%到 64%。那么,我们是不是可以下结论我们改进了呢?对这个假设的风险是 35%。我们要不要承担这个风险呢?

如果 MyInsurance 希望区分 54%和 58%的客户满意度等级,那么他们需要 54%和 58%的置信区间没有重叠。如果他们有重叠,那么我们不能区分它们。因此,我们需要+/-2%的置信区间或者更低。

对这个需求的样本量的估算告诉我们,我们需要在客户满意度调查中引入 2500 个客户。基于 280 个客户的样本,我们能够轻易的得出结论,在客户满意调查中没有变化或者更坏。直到我们得到更多的数据,我们才能得到更好的结果。

You are interested in reading more articles from COE? Try here: www.COE-Partners.com/Articles.htm
For more information about our services, please visit our website www.COE-Partners.com or email us

不幸的是，在我们例子中，MyInsurance 没有理由根据客户满的增加来庆祝成功。如果他们在 2005 年取得其他一些样本，可能就会得到异于 54% 的其他一些结论，或高或低。没有更多的数据给我们更多的信息，我们没办法摆脱这个怀疑。



结论

很多情况，重要的结论是基于很少的样本数据而得出的。我们很多时候不关心这些数据的置信度，我们也不关心在我们生活中的很多变异。大多数情况我们关心平均值。这个平均值很好计算，并且每个人都是他表示什么。每个平均值都来自于样本，并且只是对于样本是正确的，对于我们试图得出结论的总体来说是“错的”。

管理层可能基于看到的数据方式的不同而改变结论：不要相信你看到的良品率，要了解置信区间是多少。Don't trust means, they are lies. 不要基于小样本数据而下结论投资，要了解你要投资的项目最小的改进是什么。不要相信平均值，他们不可信。

关于作者

吴伟博是悟为(北京)管理咨询有限公司的常务董事。悟为业务聚焦在亚洲市场。我们通过三方面帮助您的组织提高你的效率：识别客户的真正需求，设计和改进流程交付客户需求物，在组织中开发领导力和发展人才。想了解更多有关悟为，请查 www.coe-partners.com。

吴伟博在许多行业具有丰富的流程执行和组织改进的经验，尤其在六西格玛质量改进和战略执行方面。他在通用电气金融部门获得六西格玛黑带大师资格。你可以通过电子邮件 Uwe.Kaufmann@COE-Partners.com 与他联系。

样本 M&M

一个简单实验帮助你理解样本意味着什么：

每一包 200 克的巧克力 M&M. 打开包装，数一数里面有多少个糖豆。这个数字-总体-我数的是 233。现在，请计算黄色的数量。在我的实验中，这个数字是 43。这意味着，我在总体中有 18.5% 的黄色糖豆。

样本就是在总体中拿小数量的 M&M 的代表。我用一个碗来盛我们 M&M。在一些摇动和晃动后，我计算碗里的 20 个 M&M。第一次，我这里没有黄色的糖豆。我把 20 个糖豆放回，重新拿 20 个。第二次，我拿到了 4 个黄的糖豆。又做了八次，分别得到 2, 3, 3, 6, 3, 5, 4 和 3 的结果。

计算结果，我的样本给我的结果是 0%, 20%, 10%, 15%, 15%, 30%, 15%, 25% 和 15%。哪一个是正确的？没有正确的。所有的样本都只给了黄色糖豆在总体中的一个指示。

样本结果的改变甚至没有把所有的糖豆都计算一下。这样下结论将付出沉重的代价。