

# 正确决策

政府工作人员基于证据的决策指南



保罗·马克西姆，任·加里斯，戴雷·普利卡斯，梦娜·戴维斯  
译：张婷婷

如果需要本书或作者的其他信息，请与菲沙河谷大学(University of the Fraser Valley)的任·加里斯先生联系：

电子邮件：[len.garis@ufv.ca](mailto:len.garis@ufv.ca)

电话：604-543-6701

更多有关公众安全的研究成果请参看大学网站：<http://cjr.ufv.ca>



# 正确决策

政府工作人员基于证据的决策指南

保罗·马克西姆，任·加里斯，戴雷·普利卡斯，梦娜·戴维斯

译：张婷婷

© 2015 保罗·马克西姆



## 作者简介

### 保罗·马克西姆 (Paul Maxim) ， 劳里埃大学教授

保罗·马克西姆于渥太华大学获犯罪学硕士学位，之后于宾夕法尼亚大学获社会学博士学位，专攻犯罪学和研究方法。目前，他在加拿大安大略省滑铁卢的劳里埃大学 (Wilfrid Laurier University) 巴尔西利国际关系学院经济系担任教授，主要研究领域为人口和劳动经济学。

### 任·加里斯 (Len Garis) ， 菲沙河谷大学客座教授

任·加里斯是卑诗省素里市消防局局长，被菲沙河谷大学 (University of the Fraser Valley) 犯罪学及刑事司法学院聘任为客座教授，同时也是纽约江杰刑事司法学院 (John Jay College of Criminal Justice) 附属研究院成员，以及西蒙菲沙大学 (Simon Fraser University) 加拿大城市研究所成员。他着重于研究如何通过进行基于证据的决策与创新来应对公共安全领域面临的挑战。

### 戴雷·普利卡斯 (Darryl Plecas) ， 菲沙河谷大学荣誉教授

戴雷·普利卡斯现为菲沙河谷大学犯罪学及刑事司法学院的荣誉教授，退休前他曾担任加拿大皇家骑警大学资深研究员。他曾获大学杰出教授荣誉，并于2003年荣获第十四届大学教学国际会议颁发的教学及科研杰出创新人才奖。他发表了大量论述加拿大刑事司法领域的相关问题的文章，同时在应用政策、项目评估以及有效决策上的研究也很深入。

### 梦娜·戴维斯 (Mona Davies) ， 法律分析师

梦娜·戴维斯获伦敦大学法学学士学位，随后于多伦多大学获政治科学硕士学位，并拥有由特许秘书及行政人员研究所 (Institute of Chartered Secretaries and Administrators) 颁发的专业证书。她在国际法律领域具有超过10年的工作经验，工作过的单位包括私营和公立机构。她在投资者—国家仲裁方面也颇具专长。

# 目录

2	前言	41	统计
3	简介	51	设计实验
11	定义问题	61	评估项目
21	批判性思考	73	成本分析
31	收集证据	85	决策



此项目由卑诗省素里市政府委托制作，它属于素里市的新兴领导计划，该计划旨在通过教育、工作经验与师友计划来培养出表现杰出的工作人员。其中一个独立的部分为在商业项目实例中获得启发而得到新的技能，并且将其应用在工作中。

正确决策：基于证据的决策指南的这个版本正是该计划的一项重要的成果，本书有益于市政府、该计划的参与人员以及从中有所收获的全体市政府工作人员。其他的合作伙伴同时包括卑诗省菲莎河谷大学。

此手册通过加拿大安全保卫部门的支持、加拿大安全科学防卫研究与发展中心的领导以及加拿大公众安全部门的合作，根据正确决策：消防专业人员基于证据的决策指南而改编。

我们感谢在编写本指南过程中提供帮助的以下人员：

- Yalda Asadian, 素里市娱乐与文化部特别项目经理
- Tammy Britton, 素里市财务与技术部技术领导
- Ron Gill, 素里市规划与发展部市区北部规划经理
- Trent Hatfield, 素里市工程部环境技术专家
- 素里市新兴领导人计划
- Alex Tyakoff, 素里市战略规划分析师
- Vivian Zhang, 维多利亚大学土木工程系本科在读学生
- Dr. Charles Jennings, 纽约江杰刑事司法学院应急响应研究组主任

# 前言

就职于市级，省级或者联邦政府的工作人员，都在为所处区域的居民提供高效率的服务。无论是帮助公众的前台人员，或是帮助执政官员制定有效的政策和项目的幕后人员都是如此。政府工作人员—公务员—是健全的政府的一个关键组成部分。这些工作人员以促使政府以及其政策有效运行为己任，帮助完善民用项目，并确保其按照计划实行。

身为被政府聘用的职员的同时他们也是国家公民，因此公务员们对确保政府项目达到预期的效果有着很大的兴趣。一般来说，公务员们既要要对大众负责，也要对执政官员负责。特别是管理阶层更负有确保各种政策的重点、执行的项目有效率的达标的责任。虽然这里提到的职责都是很明显的，但是在现实中却是一件不容易贯彻的任务。

通过检验项目和法规的成果来衡量我们的项目和服务的效率是唯一可靠的方法。换句话说，有哪些证据可以证明我们使用了正确的方法做着正确的事情呢？

基于证据的决策方法便是其中一个衡量效率的框架。这个方法通过将一系列的基于真凭实据的策略（以数据的形式）综合起来，从而来衡量我们的成效。基于证据的决策方法是一个帮助我们在发展，培养和维持政府项目中做出更有效决策的透明的工具。

此手册为政府部门的工作人员提供了基于证据的决策方法的综述和介绍。配套的练习册中提供的素里市的实例可以帮助读者全面地理解手册中的相关理论。通过熟悉手册里的方法与技巧，我相信你会发现你的决策过程会变的更加有效。由数据主导，或者基于证据的方法，同时也让我们更加有效的验证我们的工作效率和成果。通过考察客观的指标，使公务员、政治家以及公众得到了一个衡量我们政策和项目的有效性的坚实基础，并确保了公众享有由我们付出的努力与来自于公众的税收一起创造出的最高价值的服务。



文森·兰隆德 (Vincent Lalonde),  
理学硕士，工程师  
素里市市执行长

# 简介

## 在不断变化的世界中的有效决策方法

尽管政府的主要功能一直都是“为公民提供服务”，但是各部门提供服务的方式一直在不断的改进，政府项目也日渐变的复杂。此外，公众多方面的需求也使得各个部门之间需要进行功能性的合作以便提供更多综合性服务。这些服务项目往往需要更丰富的资源、合理的处理程序以及更优或者更专的人才。

因此，领导和经理不断地面对这个问题：我们怎样才可以在资源和资金有限的情况下为公众提供他们所需求的高质量服务？我们需要在各种方法中做出选择和权衡，同时也需要考虑各种后果。当政府官员，利益相关方，最重要的是，当公众细察这些决策的效率的时候，决策者的压力加大了。如今的政府官员和纳税人不会因为表面意义就同意动用更多的资源。政府官员、项目管理阶层和各个阶层的行政机构渐渐地经常被迫在有限的资源内做出决策。

政府官员比以前更加需要以透明、合理的方式做出决策。我们认为，优秀的决策方法需要由尽可能多的证据、调查研究以及合理的信息来支持。我们将这种方法称为基于证据的决策方法。我们通过引用独立支持和可被证实的事实根据来证明基于证据的决策方法的合理关键性。这种方法可以帮助并确保我们做出的决策是健全并且合理的。使用有效的基于证据的决策方法可以帮助你得到你所寻求的结果。

---

如今的政府官员和纳税人不会因为表面意义就同意动用更多的资源。

---

基于证据的调查研究为什么如此重要呢？这种解决问题的方法为什么与政府工作人员相关呢？

下面列举了一些主要原因：  
缺乏被合理的证据引导的政策和项目往往既昂贵又浪费资源，或者导致不理想的未知效果；  
外部的财务决策者在审批项目申请的时候会因为报告中缺乏实证而认为该项目不合理；和  
基于证据的政策和策略往往能产生更理想的效果，这样你的可信度将会提高，你所在的部门以及全体人员的工作将会获得更多的支持。

此手册将会帮助你理解怎样找寻并且使用做出基于证据的决策所需要的信息和研究方法。它同时也将帮助你把你的决策置于一个令人信服的框架中以使其他人看到它们的价值。

当然，不是所有的决定都是或者都能够基于证据的。在职场和我们的私人生活中，我们往往更喜欢做出基于道德伦理、个人喜好和政治立场的决定。单纯的相信或者否定都是低估了社会生活的复杂程度。尽管如此，就算在这样的情况下，基于证据的决策方法仍然可以帮助你价值观、原则性以及意识形态相互关联起来，因此以帮助你的部门进行独立的取证和支持性的调查研究。

我们使用的证据来源多样。有些是政府和其他正规部门例行收集的的管理行政数据，有些是从正规的政策和项目评估过程中产生的，有些则来自于私人分析师和学者的工作成果，其他信息来源也包括你所属的机构或者部门。

## 学会在证据的世界中遨游

此手册将会帮助你大胆地遨游在证据的世界中。正如我们即将了解到的那样，不是所有的证据或者数据都具有同等的价值。就算是有用的信息，我们也需要将其置放于合适它的背景中来衡量它的准确性和意义。换句话说，此手册将提供生成数据时你所需要了解的相关知识。

除了学习如何评估证据，我们也将讨论如何使用证据形成一个具有说服力的逻辑论证，毕竟数据本身并不能够表达与支持你的决策。为了做出逻辑合理和连贯的政策或者项目决策，我们需要形成一个公共证明框架。从战略上来说，不合理的项目申请或者商业计划往往并不会成功。我们也会发现有許多被提出的论证或者辩护完全不合理。因此，我们将仔细学习一些常常出现但必须极力避免的逻辑错误。

### 此手册将会帮助你：

1. 找寻并且使用做出基于证据的决策所需要的信息和研究方法。
2. 将你的决策置于一个令人信服的框架中以使其他人看到它们的价值。

此手册也将解释如何进行环境扫描和SWOT分析（一个对机构中存在的优势，弱势，外在机会和威胁而进行的分析方法）。你将会了解到实行新的政策或项目，或者确定一个战略重点之前，这些分析往往是信息收集过程的重要组成部分。你也将会了解到在资源有限的时候，战略规划的关键部分就是成本收益分析和成本研究。

通过使用加拿大政府服务中的例子，此手册将会为你展示如何定义一个问题的方法。它将帮助你批判性地、创造性地思考，以及找寻到可支持你的决策的有效证据。此外，它将简单说明各种研究方法，让你学会如何以及何时去使用它们来支持你的观点。

在我们开始之前，仔细探讨一些更深层的问题会对接下来的学习很有帮助。基于证据的决策方法是怎样以及为什么变的如此的重要？你或者别人为什么要关注这个方法？

## 医疗和健康服务已经做出了示范<sup>1</sup>

在公共范围内，基于证据的决策方法的起源可以被追踪到1980年代。英国政府当时面临着严重的财政危机，他们开始意识到需要使用可信的证据和科学研究来支持决策和实践。决策者曾经浪费了大量的资源在只有很少数据支持的决策上，哪怕他们都相信这个决策。他们做出的决策常常取决于个人喜好、传统的解决方法或者当时流行的方式。

任何在当时的英国政府工作过一段时间的人都知道，世界上从来都不缺兜售最新管理万灵药的骗子。英国政府知道国内需要投资，但必须是有效的投资，而非仅仅是一些没有经过实践的主意。<sup>2</sup>

这种方法影响了许多其他的领域，但是受影响最大的是医学领域，因为研究人员可以将糟糕的实践方法和对病人的伤害程度直接联系起来。基于证据的医学是基于将学术研究和临床实践更紧密的联系起来的演变而来的。在理想的情况下，这能保证病人得到最好的疗效和最合适的照顾。研究人员和医护人员通过仔细审查他们的政策和程序从而尝试找出更有效的医疗机构的运作方法。<sup>3</sup>

医疗机构的运作方法很明显需要改进。举个例子，一个大型调查结果显示将研究结果与建议的政策相结合这个过程花费了大约15年。

作为一个生动的示例，让我们考虑到某种癌症的疗法的研究理论依据可能早已存在。但是从发现理论到临床实践哪怕只有一部分都要滞后15年。而且在浪费了那么长的时间之后，只有40%的医护人员能真正地应用了那些有价值的信息。<sup>4</sup>

与此同时，因为医护人员没有及时的充分利用那些有用的研究发现，从而使本可以受益的病人却持续的遭受病痛的折磨或者因此去世。更糟糕的是，其他团体可能因为看到“管理”疾病比真正地治愈它可以获得更大的利益而故意不去治愈它。

医疗领域中，基于证据的决策方法被持续地推进。在其他的国家，这种方法也成为了公共医疗政策发展中的奠基石。在全世界的政策主张群体，研究人员和其他的利益相关团体的共同努力下，使用基于证据的决策方法的需求渐渐地得到了认可。<sup>5</sup>

基于证据的决策方法是尝试通过使用出自研究、实验、观察以及事实产生的最可用的信息来帮助制定最好的决策和政策。有时候就如同上文的例子一样，这样做会与其他势力的价值观和利益直接产生冲突。



## 实例

正如我们已经指出的，医疗领域中已经存在大量的并且数量仍在持续增加的应用基于证据的决策方法的文献记载。刑事司法的领域中同样从实质上接受了并且应用着基于证据的决策方法。<sup>6</sup>

而在政府机构中，仅仅有着少量的公开的正式应用案例。可能内部工作人员听说过一些先例，但是这些先例的细节往往没有被公开。

早在2000年的时候，安大略省级政府曾经尝试着解决护士界的一些问题，就是一个大型并且相对记录完整的先例。<sup>7</sup> 在1998年，省政府建立了一个护士工作组以解决几个问题，包括“帮助渥太华留住已有的并且吸引新的护士，改善护士的工作环境，并且确保护士在日渐复杂的环境中具有足够的专业技能来提供关怀服务。”

护士工作组的欧布莱恩·帕乐斯和博曼在一份综述中总结并归纳了其中的基于证据的元素是如何联系且聚集在一起的。<sup>8</sup> 当时该工作组非常幸运地利用了几个大型数据库，包括渥太华政府的一个行政数据库，渥太华护士专科学校中有关注册学生资料的数据库，以及加拿大统计局和加拿大医疗信息机构的数据。

这些数据将一系列的采访、各方利益相关者递交的应征意见书以及描述了在职护士所面临的挑战的文献综述相结合，工作组可以由此来研究与护士相关的供求关系以及工作环境中存在的问题。正如的欧布莱恩·帕乐斯和博曼的总结所述：

此过程协调了研究人员提供的‘事实’和参与访谈的人员的价值观和信念。最终，由决策者（资深政府人士，经理，专业服务人员），知识传播者（工会，协会，公众代表，护士），以及研究人员（国家研究中心的研究人员和他们的同事）类似的组合而发展成的可靠体系承担了对各种建议的实施进展以及结果进行不断监控的责任。<sup>9</sup>

最终，人员短缺，员工士气以及患者服务方面的相关问题得到了显著地改善。

## 做出有效的决策：优秀领导和经理的任务

做出决策是领导和经理们的任务之一。他们的决策指引着所在部门的发展方向，并且影响着员工的士气和心理健康状态。糟糕的决策会增加矛盾、削弱士气。优秀的决策会带来切实的、正面的结果，从而使得部门更加成功，并且能够鼓舞员工的士气。

尽管领导和经理认可基于证据的决策方法的价值，在实施过程中依然存在着一些阻碍因素。一些行政人员由于时间紧迫不得不在缺乏完整的信息的情况下迅速做出决策，另一些则可能使用过时的信息。因为政府部门是一个快节奏的环境，常常需要工作人员短时间内迅速完成工作任务。

然而，通过清醒思考以及充分参考最新的信息数据从而做出合理的决策，以及进行更有效的实践才是管理工作的正确方法。此外，很多人会根据个人经验、观察、或者直觉做出决定。作为受过培训的政府工作人员，我们的个人经验和判断往往是合理的，但是它们只是决策所依据的一部分。认知科学指出，我们常常只注意到我们期待看到的东西，当需要面对具有不确定性的复杂问题带来的挑战时，我们的大脑并不是一个有效的处理器。

使用基于证据的研究调查可以帮助将我们的经验和观点融入一个最终更有说服力并且更广阔的信息环境中。此外，实践方法在不断地进化。政府专业人员在20世纪初期使用的方法或许在21世纪的今天已经不再适用。

---

糟糕的决策会增加矛盾、削弱士气。优秀的决策会带来切实的、正面的结果，从而使得部门更加成功，并且能够鼓舞员工士气。

---

当制定一项新策略或新政策的时候，最好去评估一下你所了解的信息、你身边其他人所知道的信息以及调查研究的结果。在你开始执行一项新的政策或者实施一个计划之后，谨慎的做法包括持续的评估其效果，这样可以生成可展示其成效的证据，使整个领域进步。通过你部门的研究可以告知其他部门哪些方法可行，哪些不可行，以及其中的原因。我们常常不愿意去评估某个项目或者实践是因为我们担心发现它并没有成效。但这并没有关系，因为无论作为个人还是作为一个社会集体，我们从失败中学习到的比在成功中学习到的更多。

## 实证研究的本质

当我们谈到证据的时候具体指的是什么？在这里的语境中，证据有它特定的意思，并不是指像警方在犯罪现场调查得到的证据。这里所指的证据是对由正规评估、实验或者其他通过研究模型等方法输出的系统性的数据集合进行实证研究从而得出的结果。它是一种通过处理具有可复制性、可观察性、可信性、可确认性并且可支持性的信息或者事实来回答一个研究性问题的系统方法。

当评估可用的研究方法的时候，得到的结果可能是：

- 定量数据，生成数字和统计的结果；
- 定性数据，生成有助于判断喜好，价值观，或者答题者的出发点的有效的主观信息。

这两种方法都可以生成有效的数据。关键在于你要知道在何时以及何处使用何种证据，以及判断出这些证据能否适当的为你的当前目标服务。

尽管证据的来源有许多，但是学术研究的优势在于它们不断的被与项目无利益关系的外界仔细地审查着，这表明其他独立的学者和研究人员在审查着研究是否可信以及其设计是否良好。但是，这并不代表研究工作要么完美要么失败。

不过，这样做可以提高研究结果的可信度。研究结果必须通过同行评审之后才能发表在大多数的学术期刊上。其中一些学术期刊具有极高的专业性，令领域之外的人对它们生畏。不过不必担心，多数重大的学术成果都会做一个总结或者将用通俗的语言将结果表述出来。

## 常见的研究方法

在医学领域中，研究的黄金标准是随机控制实验法。研究人员给个人随机分配各种预防，治疗或者诊断干涉，然后跟踪衡量干涉的效果。然而，其中任何一种干涉都有可能是完全无效的。这样允许研究人员将对照组（接受完全无效干涉的）与接受各种干涉的测试小组进行比较。药物测试往往就是用这样的方法进行的。我们在之后的章节中将审查不同取证的构架以及讨论研究人员之所以将随机的控制实验方法置于如此高的地位的原因。

社会学中也可以实现包含一个对照组的随机实验。比如，我们可以随机的为一些房屋配置安全警报系统，将它们作为测试小组，与其他随机的没有警报系统的房屋（对照组）做比较。这是回答并研究“配置了安全警报系统的房屋是否比没有警报系统的房屋更少的被小偷破门而入？”这个问题的方法之一。

道路和交通部门也可以通过做实验来评估安装于交通路口或者高承载率车辆(HOV)通道的摄影机的效果。

研究人员通过设计这样的实验来“控制”大量可能扭曲结果的外界因素。这样提高了实验的正确率，因此你对衡量标准和结果会更加有信心。研究人员对实验结果的可靠性也十分地关心——意思是：如果重复的持续做这项实验，我们会得到同样的结果吗？如果在不同的社区中做这项实验，我们会得到同样的结果吗？还是，只有在某个特定的社区才会得到这样的结果，以及如果是这样的原因是什么？研究必须正确且可靠，这样你才能相信它产生的结果是合理可信的而不是因为侥幸或者巧合。

## 做出更棒的决策

到目前为止，或许你已经了解到了通过合理可靠的研究而做出决策的益处了。很简单，如果你做了功课，那么你很可能会做出一个更加有理可依的决策。

由于决策的过程更加的透明，且是建立在你的直觉、猜测或个人见解之外的依据之上，你在为决策辩护的时候也会更加简单容易。

不过，我们应该意识到基于证据的决策方法最适合回答客观的问题。在这个章节的开始我们提到了其他方面的决策主要是基于我们的喜好、价值观或者信念，而这些和研究没有太大的关系。

然而在我们希望寻找一种符合我们的价值观的最有效的解决方法的时候，这两者往往会融合在一起。比如，社会的价值观给予了我们采取不同的方法做事情的动力，其一便是通过提供高质量的警务系统和减少犯罪起数来提高我们的生活质量。基于证据的研究方法帮助我们了解到需要做些什么和如何去做。

## 注释及参考文献

1. Tranfield, D., Denyer, D. and Smart, P. (2003) "Towards a methodology for developing evidence-informed management knowledge by means of systematic review," *British Journal of Management*, 14: 207-222.
2. Tranfield, D., Denyer, D. and Smart, P. (2003).
3. Vishwanath V. Baba, Farimah HakemZadeh, (2012) "Toward a theory of evidence based decision making," *Management Decision*, 50: 832-867.
4. Antman, E.M., Lau, J., Kupelnick, B., Mosteller, F., Chalmers, T.C. (1992) "A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. Treatments for myocardial infarction." *Journal of the American Medical Association*, 268: 240-248.
5. This includes Canada. See, for example, Kiefer, L., et al., (2005) "Fostering Evidence-based Decision-making in Canada. Examining the Need for a Canadian Population and Public Health Evidence Centre and Research Network." *Canadian Journal of Public Health*, 96: I-1 to I-40.
6. The US National Criminal Justice Reference Service has an extensive data base of relevant material. See <https://www.ncjrs.gov/App/Publications/AlphaList.aspx#>
7. See Ontario Ministry of Health and Long-term Care (1999) Good Nursing, Good Health: an Investment for the 21st Century. Report of the Nursing Task Force. [http://www.health.gov.on.ca/en/common/ministry/publications/reports/nurserep99/nurse\\_rep.aspx#appC](http://www.health.gov.on.ca/en/common/ministry/publications/reports/nurserep99/nurse_rep.aspx#appC) ; JPNC Implementation Monitoring Subcommittee (2003) Good Nursing, Good Health: The Return on Our Investment. [http://www.health.gov.on.ca/en/common/ministry/publications/reports/nursing\\_roi\\_04/jpnc\\_roi\\_2004.pdf](http://www.health.gov.on.ca/en/common/ministry/publications/reports/nursing_roi_04/jpnc_roi_2004.pdf) .
8. O'Brien-Pallas, L., and A. Baumann (2000) "Toward evidence-based policy decisions: a case study of nursing health human resources in Ontario, Canada." *Nursing Inquiry*, 7: 248-57.
9. O'Brien-Pallas, L., and A. Baumann (2000) p. 254.

# 定义问题

## 决策是各式各样的

每天我们都在不断的做出各种决定，从早餐吃什么，在哪一个房间里举行会议，到是否购买一辆新的轿车。许多决定都非正式且无关紧要，没有什么不好的后果，然而其他的一些决定却有一定的风险或者不确定因素。我们每天都在冒险，合理的风险并不会妨碍大部分人的日常生活。使用正规的步骤来解决这些问题通常需要花费太多的时间和资源，从而导致我们的生活停滞不前。

与此同时，我们会面对私人生活中的重大决定，以及在职场中我们会需要通过和别人合作而做出一些会产生不小的后果的决定。这样的情况值得我们花费时间和精力去仔细地审查问题的细节，并且使用正规的步骤加以分析。一般来说，正规的分析步骤包括一个清晰的问题定义、可选方案的概述、以及选择任何方案的相关的成本和效益衡量结果。

基于证据的决策方法的其中一个优势是我们可以通过使用已知的结论来预测出一个可量化的结果。所幸的是所有的事物都是可量化的。无论量化有多么的模糊，只要它能提高你对某事物的了解，它就是一个量化。<sup>1</sup>

在事件发生之前，没有人可以预知一个决策会导致的实际影响。不过，我们通过借鉴经验和可用的证据，可得出一个合理可辨的预计结果。

---

基于证据的决策方法可以帮助我们在需要的时候做出重要的经济、社会或者政治方面的决定。

---

每个人都时不时的做出可能导致不良结果的决定，这就是现实生活。事实上，我们做出一个错误的选择与做出一个糟糕的决定是不同的。没有做出正确的选择与做出糟糕的决定是有区别的。正如我们之前说过，糟糕的决定是可以避免的。一个好的决定而导致的坏的结果却是我们无法控制的。那么如何区别一个好的决定和一个糟糕的决定呢？很简单，好的决定出自于清晰表达出来的问题。同时也从当我们引入尽可能多的合理且可用的证据的地方产生。一个好的决定是当你带着清晰的意识回顾时，在与当初同样的情况下，且在同样的证据支持下，你可以坚信你依然会做出同样的选择。

虽然因为一个好的决定导致了一个不那么理想的结果是不幸的，不过因为有一个好的决定的优势是我们可以从这样的失败中学习。当决策的过程是透明的時候，就有可能了解到产生负面结果的原因是什么。是我们的假设不正确吗？是我们遗漏了些重要的信息吗？还是我们的逻辑有瑕疵呢？

在此章节中，我们将思考以下几个问题：

- 问题是什么，我们应该怎样定义它？
- 我们怎样确定可选及替代方案？
- 我们怎样创造性的思考并生成新的方案？
- 我们怎样生成替代方案？

## 问题是什么？

通常，理性且系统的决策者都会从列出替代方案着手。列一个表是十分有效的方法。不过，正如约翰·洛克菲尔曾经在另一种情景中所说，“一张列表不足以称之为计划。”

在我们开始生成选项之前，我们需要提问：这个决策的目的是什么？我们预期的目标是什么？这些问题包含在对问题的分析中。分析的框架往往是通过制作一个战略规划或者商业计划。不经过计划就做出决策是很常见的。但是，正如格言所说，“如果我们不能计划，那么我们就计划着失败。”如果没有一个明确的计划，我们一般将无法得知一个不够理想的结果是因为一个糟糕的方法产生的，还是因为我们遇到了新的或者不同的情况。通过一个公开正规的分析过程，我们可以积累知识以避免在将来再犯同样的错误。

通常，没有经过计划的决策都无法得到好的结果。通过计划我们可以合乎逻辑的且系统的做出决策。这样，我们可以向评论家说明在当时的情况下我们做出的是理智合理的决策。

当我们问“问题是什么？”的时候，本质上我们在问，“我们的决策如何适合且促进我们机构的授权？”

## 在做任何事情之前，问一句“为什么？”

很明显，这名员工在尝试着迫使单位做出一个错误的选择。在这样的情况下，我们把它称为假二歧式因为这个提问假设了只有A或者B两种可能的选项。事实上，存在着有许多的选择：员工可能被转到其他部门，解雇或者什么也得不到。当然，我们的假设是这是一个在合理的好员工的业绩评估中出现的常规问题。

在考虑各种可能性之前，评估这名员工对工作单位的贡献是一个好的开始。理想的情况下，工作单位应该已有一个合理的业绩评估的政策。如果没有的话，你可以问五个问题。基于你的业绩，你为什么应该被奖赏？在提高你部门的工作效率上，你贡献了什么？我们在哪儿可以看到你做的贡献？你的部门中，你今年帮助或者支持了谁的工作？我们什么时候可以看到你业绩的回报？

或许在不同的情况下，有不同的合适的问题，但是基本的意思就是这些。主要的概念是将员工的要求与部门或者单位的大目标联系起来，保证我们考虑的选择与大目标的方向一致。一般来说，我们尝试着保证构成我们选择的基本原则是与大目标相关的。仅仅因为员工很友好，按时上班，或者着装整洁而给予奖励就很难说的过去。

适当的计划使得决策更合乎情理，哪怕最后的结果和预期的不同，在政府工作人员职责日益增加的环境下，这点特别的重要。

当这些分析思路都行不通的时候，你可以问问自己，“我能向单位中的其他人，我的老板，或者公众为我的决定辩护吗？”一位昔日的同事曾经说过，“我会假设我所做出的每一个重大决定都会是第二天报纸的头条，如果我觉得可以接受，那我就很可能在合理的情况下做出了一个合理的决定。”

授权是我们的战略规划的正常组成部分。有时候嵌于我们的操作计划或者标准操作程序中(SOPs)。

比如，一个利益团体希望为长者与儿童争取更多的娱乐服务。表面上，这是一个有价值的要求，但是这仅仅是有可能提出来的有价值的要求之一。市级的战略规划的目标中可能明确了保证交通与基础设施方面的建设需求为优先考虑的事情。因此，城市的紧急需求是升级公共汽车队或者由于洪涝而导致需要建一个排水系统。因为参考了计划，我们可以了解到为长者与儿童提供更多的娱乐服务并没有那么的重要。而且，与升级基础设施相比，在某一个地区增量投资给整个城市的生活质量带来的提高实在太少了。

这个问题的本质上取决于所请求的投资是否与整个社区所定义的需求相符合。并不是说，我们没有成功的增加整个社区的娱乐服务，实际上问题其实是如何在有限的资源和有竞争的需求下最好的满足社区的需求。

不可否认，经理会带着许多请求去见委员会，接着最受大家欢迎的项目会通过。重点在于被考虑的项目不仅仅是娱乐服务那一项。这里的关键是提出的项目要符合机构的操作重点，或者在最理想的情况下要符合机构的战略规划。

重申，将决策嵌于一个已有的计划，或者操作框架之中，那么以战略性评估的相关根据，该决策就是可辩护的。这样的话，暂缓建立一个新的提供娱乐设施的项目是合理的。

将决策嵌于一个已有的计划，或者操作框架之中，那么以战略性评估的相关根据，该决策就是可辩护的。

## 生成方案

可选选项常常是显而易见的。我们在设备还是人工上花费更多？我们的数据处理设备是否到了预期的使用寿命？在其他的情况下，可选选项并没有那么的显而易见，因为并不总是选A或者不选A这样的问题。在之后的章节中，我们将学习如何进行环境扫描以及SWOT（优势，劣势，机会和威胁）分析。这些都属于相对正规的分析方法，用来系统的检查其他人在类似的情况下做了以及没有做了什么。

在求助于这些方法之前，还有几种适合生成可选项的方法。你可以考虑以下的几种途径。

### 与你日常圈子之外的人聊天

我们常常将自己的社交与专业圈子局限在我们已经认识的或者一同工作的人之中。这样做常常会产生在有限的选择中强化信念的团体思考的思维方式。此外，相对于为你提供不一样的建议，同事和下属往往更加希望赞同你的话语或者告诉你他们认为你希望听到的话。而圈子外的人或许也会面对相似的情况但是会用完全不同的方式着手处理问题。

## 参加小组头脑风暴会议

除了有可能出现团体思维，有时候最好的想法来自于你身边的人。他们了解所在的机构以及问题所在。而且，咨询他们比聘请咨询师更实惠因为已经付了他们工资。可以单独的咨询他们的意见。也可以一个小组讨论，在会议上希望大家提出一些“疯狂”的想法会有意想不到的效果。

一些表面上的竞争有时候可以刺激人们想出新的方案。记住，今天的发明曾经是昨天的不可能。头脑风暴会议可以是随意的也可以是有组织的，其目的都是生成尽可能多的方案并且找到困扰已久的问题的解决方案。

## 阅读更多的书籍与期刊；上网

你的阅读量越大，特别是你领域之外的内容，你就会有更多新奇的方案。商业书籍明显是一个好选择但是有时候更好的方案来自于小说。大部人都希望不断的自我提升，阅读专业期刊是了解目前的发展趋势的一个好办法。网络通常混乱无序以及如格言所说，“你付出多少，便得到多少”。尽管这样，如今的搜索引擎已经十分优秀，可以帮助你发现珍宝。如 Stephen Covey 所说，“磨刀不误砍柴功”<sup>2</sup>

## 专注于你服务的人群—外部与内部

从你服务的外部与内部人群的视角去看这个世界。他们看到的你的机构或许与你和你的同事看到的截然不同。除了你服务的人群，你的专业协会，社区组织，教育机构和政府部门的其他机构都可以为你提供一个好方案。这些关系往往都有很大的价值。了解外部人的看法可以给你带来很大的好处。

## 聘请一个有声誉的咨询师

往往因为在当地的领域中你就是专家，因此你有了今天的职位，但是有时候你要做的决策在你的领域之外。大部分商业机构都会与外部的设计所，营销代理，网络设计员或者管理咨询师合作。关键在于明确你所需要的专业领域。一旦明确，就可以请你的同事们推荐一个咨询公司或者个人。通常，小型公司更加有创意并且收费较低，不过创意是与其他服务一样的生意可以被购买。

当然，你需要愿意接受新的视角。别让你的成见成为障碍。因为你不看好的某个人并不代表他们团体的方案都是糟糕的。同时，也别因为其他人的主意比你的好就觉得受到威胁。特别是当那个人是你的下属时，有这样一个有创意的员工在你的团队里是你的功劳。

最后，有时候你要接受最好的方案是最显而易见的那个。一个咨询师交给你一份结论你已经知道的报告，并不是因为他的懒惰或者没创意，或许是因为你认为最显而易见的方案就是最好的方案。就当作为证实了你的猜疑吧。

## 制定一个计划

制定的计划不论大小与复杂程度，每个机构都可以从中受益。无论我们将它称为战略性、机构性或者商业计划，目的都是相同的：一个机构需要了解他们需要做什么，为什么而做，方向是什么，以及希望采取怎样的方式达到目的？

如果没有计划，人们会随意的做决定。最好的结局是这些决定不一致，最糟糕的结局是它们互相矛盾。一个计划不能保证机构将成功地或者有效率地运行。但是如果缺少计划，机构的运行必然会变的平庸或者最终失败。

书店以及网络上有着许多资料概括了如何建立一个机构的整体规划。主题包含了从项目管理实践与原则到最新的管理案例分析。花费时间阅读类似的资料是一项非常值得的投资。

基本上，一个计划包含以下四个要素：

- 一份机构价值的概述；
- 一份宗旨与目标的概述；
- 一份关于机构希望如何达到目标的大纲；
- 一份评估绩效的指标。

每个计划的复杂程度不同但是尽可能的保持简单会带来有一些优势。复杂的计划往往难以记住并且有很多的约束。

计划的四个组成部分：



如同战场上的将军了解的那样，一旦开始行动，往往很少时候能和预期的一样。将军往往最希望的是战士们知道他们在打哪一场仗，他们能够铭记整体的目标，部队长接受过足够的训练以应对突发的战略性挑战及挫折。所以，这足能表明保持简单的重要性。

---

构思良好的陈述可以鼓动人心，且可以制作成鼓动人心的广告版置放在办公室的墙上以及年度报告会场中。粗糙的陈述只能给人轻浮的感觉。

---

### 陈述机构价值

现今的管理专家们常常会花大量的时间来确定机构的基本价值。通常，我们将机构的基本价值概括归为以下一种或多种：使命陈述，愿景陈述，价值陈述。

构思良好的陈述可以鼓动人心，且可以制作成鼓动人心的广告版置放在办公室的墙上以及年度报告会场中。粗糙的陈述只能给人轻浮的感觉。最实用的建议总是将事物保持简单、直接。简单的说，清楚的陈述容易被记住且容易遵循。

价值陈述在本质上概述了机构存在的原因。这也叫做使命陈述。很多机构，例如交通部的使命是很显而易见的。它存在的理由是为社区居民提供便捷、优惠的交通服务。使命陈述回答了这个问题，“你的目标是什么？”

在价值陈述中，还应该提供出机构预计在接下来的三至五年中的期望。换句话说便是机构中期的愿景是什么？例如，也许你希望你的业绩成为当地的模范。

最后，价值陈述中表明了你的核心信念。这是应该是基础而具有启发性的。谷歌的信念是“不伤害别人。”而你的可能是“为社区服务。”尽管听上去有些老套，不过当决策开始更多的专注于扩大机构的最大利益而非为客户服务的时候，价值陈述能够很快的提醒人们机构的核心信念。这个情况下，你所做的不应该是为了机构的利益，而是为你所在的社区服务。

### 陈述宗旨和目标

一个机构的宗旨和目标陈述包括它自设的指标。机构的宗旨是大方向上的指标；目标是为了达到宗旨而自设的中期步骤。大方向的目标包括为社区提供可用且价格实惠的住房，减少贫困以及创建一个安全和谐的社区。

为了实现减少贫困的愿景，我们常常首先需要列出达到中期目标的各个步骤。例如目标之一是建立相关的倡导策略且支持主要的服务行业。<sup>3</sup>



## 拿破仑的愿景与目标

人们常常混淆愿景与目标的概念。有时候还将它们交换着使用。尽管它们相互关联，但是它们各自有着特定的概念。我们看看1799年拿破仑·波拿巴的意图。

### 愿景

统治整个  
欧洲

### 目标

成为法国首领  
征服意大利  
征服西班牙  
打败普鲁士军队  
打败奥匈军队  
将波兰并入法国  
征服俄罗斯

然而讽刺的是，拿破仑实现了除了征服俄罗斯之外的所有目标。尽管得到了让人印象如此深刻的成就，他最终还是没能实现他的总体愿景。他低估了俄罗斯恶劣无情的冬天给军队带来的负面影响。

## 执行程序

我们之前提到过，一个列表并不足以构成一个规划。概述机构的愿景与目标是完整规划过程的必要非充分部分。一个真正的规划包括我们将怎样实现目标的探讨。哪个机制或者怎样的过程已经落实到位以达到期望的结果？比如目标是减少工伤发生的次数。我们可以将此与整体的愿景“提高员工的安全”联系起来。

需要思考的重点是：我们该怎么做才能达到目的呢？很明显，我们选择的机制取决于环境。也许我们可以为加强员工训练而提供更多的资源，但也有可能员工已经接受了足够的训练却没有机会练习那些步骤。另一个机制是加强安全部门的人员与一线员工之间的信息共享以及工作关系。

这个方法对规划中所有确认的愿景和目标都适用。无论是“软性”目标，比如提高员工的士气，还是“硬性”目标，比如减少工伤发生次数或者公共财产，建筑或者机构的破坏。

实施过程是指我们规划中的可操作项目。战略性或者商业规划往往列出机构期望实现的目标，但是缺少提供达到目标的方法。换句话说，如果比喻愿景与目标是一句话中的名词，实施过程则是可变动的部分或者动词。

## 评估结果

评估结果本质上是在填写一张计分卡。在进行评估之前，在你的规划中首先必须有你将使用的明确的特定的业绩指标。这些指标应该紧密的与规划中的特定的目标相联系，一般来说，指标应与整体愿景相联系。正如杰出的管理大师，彼得·达克曾经说过：“可以被评估的事，就可以做好。”

很显然，清晰的数量测定使用起来最简单，如同呼叫请求服务的数量变化，申请福利的数量变化，交通事故次数的变化，或者犯罪率的变化。不过，我们也不应忽略定性分析。比如社区满意指数或对犯罪的恐惧程度就很难被量化，但是这些对所有的服务提供部门来说都是非常重要的业绩元素。

一般来说，评估结果包括的内容非常的广，从内部业绩打分，提供服务的水平，到财务问责。很多战略规划建议建立一张表格，将操作目标置入一列，同时将相对的评估标准置入旁边一列。其中的相关性属于个人的判断，但是复杂的目标一般需要匹配比简单单一的评估更多样化的指标。

战略性或者商业规划往往列出机构期望实现的目标，但是缺少提供达到目标的方法。

由于愿景的性质相对于目标的期限更长，程度更高，确认特定的评估的难度就越高。此外，愿景往往相对于中期目标更加需要一个定性的衡量标准。要记住的是，虽然将评估结果与愿景、目标相联系非常重要，但是无需完美的契合。

有时某人的目标都达到了也却没有实现愿景也是可能的。同理，一两个目标没有达到也不代表机构没有实现它的整体愿景。实践的战略或者商业规划有时候会包含其他的事项或者在某一方面提供更多的细节。

我们也可以将机构为你所在社区带来“增值”的组成部分的有关组成部分的细节描述到位，或者该机构与其他类似机构或者服务提供部门有哪些不同。这些内容是否有用取决于机构所在的特定的环境和所处的情况。无论如何，这些成分是我们需要做出重要决策的时候而参考的要素列单的一部分。

我们常常模糊的呈现问题或决策。一个优秀的决策者会将问题的本质以及问题与机构的规划之间的关联说明的清清楚楚。在这之后，他会问一些附属问题，比如：要解决这个问题是否需要采取行动？如果需要，我们什么时候打算采取行动？这个问题是否紧急、重要？

## 基于证据的决策方法

基于证据的优秀决策与该机构的规划紧密的相互关联着。这不意味着有的时候我们就不会做出超过规划范围的决策。尽管我们需要专注在核心业务功能上，但是环境在变化且新的问题可能会出现。这个世界并不是静态的。

不过，有效率的经理都具备足够的灵活性且能够处理这些状况。尽管如此，通过练习规划往往可以获得广泛的视角或足够的指引，我们规划的框架中也可以包括这些突发状况。

一个优秀的规划为决策者在衡量他们选定的选项的时机以及原因提供了辩护依据。

他人可以轻易地批评决策者选定的选项。事实上，很多的决策有着几乎无数的选项可选。如果我们指出有些建议尽管有价值，但是它们在战略规划的范围之外的话，那么我们就可以大幅度的缩减可选项的数量。

通过参考一个优秀的规划，我们可以明确的知道相关问题是哪些，值得考虑的建议是哪些，因此，我们需要考虑利用哪些证据来衡量这些选择。

---

## 注释及参考文献

1. Douglas W. Hubbard (2014) *How to Measure Anything: Finding the Intangibles in Business*. New Jersey, NJ: John Wiley & Sons.
2. Stephen R. Covey (1989). *7 Habits of Highly Effective People*, New Jersey: Simon & Schuster.
3. City of Cornwall Strategic Plan, 2013.  
<http://www.cornwall.ca/en/cao/resources/CityCornwallStrategicPlan20132015.pdf>

# 批判性思考

## 清晰的思路

要做出优秀有用的决策，单凭证据与数据是不够的。我们提出一个论点或解释的过程与我们提供的信息的质量同等重要。当我们考虑使用基于证据的决策方法的时候，我们需将以下两个方面牢牢记住。首先，在任何方案的创建过程中，根本的论点必须基于有效的逻辑。论点常常导向至很多通常无效的结论。其次，创建一个方案的最有效的方式通常不是基于大部分人对证据支撑论点的理解。

这两点描述了人性。有些人喜爱争论，哪怕有人证实了他的论点是错误的或不和逻辑的，他也不会改变自己的看法。人类是一种固执的动物，有着为自己贪爱的虚假而辩护的趋势，而不愿接受最合理的推论与证据。

很多证据显示了人们最喜欢的社交活动之一就是相互辩论。在星期六晚上的任何一家体育酒吧中，你都会见到无穷的证据。

接下来还有网络。自涂鸦墙被发明之后，崛起的网络沦为了考虑不周的念头，阴谋论以及它们的追随者们绝佳场地。基本上，证据与合理的逻辑很难影响人们的看法。上一次在你指出了某人方案的一些不足之后，他回复以下的话是什么时候？比如“谢谢你指出了我不合逻辑的地方。我了解到我在这个问题上错了，并且从此以后在这点上我会改变我的看法。”

---

人类是一种固执的动物，有着为自己贪爱的虚假而辩护的趋势，而不愿接受最合理的推论与证据。

---

通常比较好的结果是他们转移话题；不好的结果是他们转身就一直嘀咕你以及与你一类的人都是些笨蛋。

事实上，有些讨论是没有结论的，无论是逻辑性的或以经验为依据的。比如神是否存在；谁是最美貌的演员；或者海伦姑妈烤的蛋糕是不是全世界最棒的，这些争论永远都得不到解决。

一般来说，价值相关的问题取决于个人感性的偏好。另一方面，我们也有被证据与理性影响的情形（或者说，至少我们当中的某些人在某些时候）。比如说经济问题，通常指挥着我们的理性情绪，健康问题，以及生与死——比如给你的孩子们注射麻疹疫苗，——倾向于引起理性的回应。虽然，这些以及所有的领域中都滋生着大量的江湖骗子。

本章节着重关注的情形是那些无论作为个人还是团体，我们都愿意考虑在决策过程中使用理性与基于证据的方法。这些情形在世事中很少出现，因此我们应该不要错过使用合理逻辑做出正确决策的机会。

## 逻辑谬论

逻辑性的陈述一般为这样的句式：如果由A可以得到结论B，且由B可以得到结论C，那么由A可以得到结论C。逻辑谬论是指论证过程中的那些未被承认或者未被质疑的存在着天生的缺口，互相矛盾或者简单的缺乏联系的部分。一些逻辑学家与哲学家以列出几乎无穷尽的逻辑谬论种类为职业(可参考网络)。不过，在大多数情况下，逻辑谬论都属于同一小类。学会辨认这些谬论，你将不太容易被刻意或无心的误导而引入歧途。

### 诉诸权威

没有人可以从零开始地创造全人类的知识。小的时候，我们学会了听从家长，老师和其他“专家”的正确教导。人类与其他动物的不同之一在于我们将已懂得的知识代代地积累且传递下去。这使得我们发明了抗生素，修建了摩天大厦以及往被我们成为手机地小盒子里发送垃圾信息。如果不从权威机构学习知识，世界上将没有文明存在。

不管怎样，尽管我们愿意接受大师专家们传授的智慧，但是也不应该盲目，因为大师专家们也不总是正确的。要求一些权威机构为他们的声明提供进一步的证据是很合理的。

我们没有那么多的时间质疑所有的权威机构，但是有些申述应该会引起你的怀疑。

你应该怀疑这样的句子，“但是，一直都是这样做事的，”或者“专家告诉我们…”

以下是常见的应该引起你的怀疑的句子：

- “但是，一直都是这么做事的。”
- “每个人都知道应该像这样做。”
- “你（我们）懂什么？他是专家。”
- “科学告诉我们…”
- “专家同意…”

在这些的情况下，我们可以很合乎情理的这样回答，“如果是这样的话，那么显然存在一些真凭实据可以支持这个观点。或许我们应该仔细核查相关的细节。”或者，“天哪，这真有趣，因为有些（科学家，专家等等）持完全相反的观点。我们应该怎么解决这个问题呢？”

一般来说，诉诸权威的人都持有这样的心态，“我懒得去核查，”或者，“我不想惹麻烦。”

## 人身攻击

人身攻击在拉丁语中的意思是“针对个人。”本质上，人身攻击的辩论指的是某些人攻击对方的人格而非对方持有的观点。通常为质疑对方的心智是否健全，道德操守是否端正或者怀疑对方的血统。人身攻击的辩论试图通过贬低对方的可信度来达到打击对方的目的。一些相对文雅的人身攻击的语句如：

- “你期待几个法西斯分子（或社会主义者，自由主义者，学者等等）告诉你什么？”
- “这很显然是跟不上时代的人持有的典型观点。”
- “这是典型的男权主义者（或女权）的反应。”
- “天哪，按他 / 她这样做事的方式，你居然认为他 / 她是一个专家。”
- “你在这个领域里工作了几年？”

这里强调的重点是将发言人与他持有的观点或者主张分别看待。仅仅因为某个人对另一个人的评价很低，不代表他的观点一定是错误的或者无关的。有时候很难做到，但是我们要尝试着做到尽管认为对方不值得尊重，也要尊重对方的观点。

## 扯开话题——“红鲱鱼”

红鲱鱼是指在讨论中某些人提出与主题无关的话题。比如，在委员会议中，有人认为出差，管理员的设备购置或者加班的费用过高。接着，他又指出是几年前我们没有购置适当的科技设备导致了这些过高的开销。

这里的问题其实是没有经过审核且不合适的花费是因为缺乏财务监督。在罗马时代和电脑的发明之前，有效的财务监督就存在了。购置适当的科技设备也许会在监督过程中起到帮助作用，但是这不能保证监督本身的过程有效。仅仅使问题更麻烦而不能真正解决问题的所谓解决方案都属于此例。

解决红鲱鱼的关键在于问一问鲱鱼与待解决的问题之间有什么联系？如何通过使用科技加强监督？有可用的合适软件吗？审计师是否进行过合适的设备培训以提高他们的工作绩效？毕竟，电脑只是听从了我们的命令。

## 粉色鲱鱼

真正的红鲱鱼是那些明显与待解决的问题无关的事项。有时候，某些人会提出一个可以解决问题的充分但是非必要事项。我们把这些称之为粉色鲱鱼。也许管理员会提议的粉色鲱鱼是机构中存在的最大问题是缺乏资金。

资金可以购买资源。但是很多时候，拥有更多的资金并不能改善什么。资金本身并不是解决问题的方法。适当的监督财政开销，更合理的使用已有的物质与人力资源，或者一个更具新意的方法都比简单的投入资金更加有效。将已有或将获得的资源直接投入在开发或加强与问题相关的机制上才是必要的措施。

如同解决红鲱鱼一样，我们在回应要求更多资金的要求的时候应该问清楚资金的使用计划。得到的回答很可能是打算用来配置更多的设备或者聘请更多的人员。那么附带问题就变成了：那么设备或者人员将以怎样的方式增强一个此刻出现漏洞或无效的运作程序？

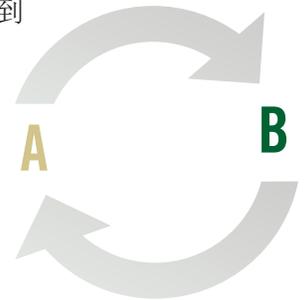
## 循环论证

循环论证的句式是：如果A导致了B，因为B是A的结果。与政治相关的辩论中常常出现循环论证。老师们最常遇到的循环论证的例子是，一些学生在考试之后来找老师说，他们不可能考试只得到C的成绩因为他们是优等生。（那么，请你解释如果你是一名优等生，你是怎么只考了C的？）

还有一个有代表性的例子有时候会在薪资协商的过程中出现。谈判单位常常坚持他们需要更多的加薪因为他们在类似机构中一向是最高薪的人群。如果你不给他们加薪，他们怎么会是最高薪的？在劳资协商中出现的影响最大的棘轮效应就是基于循环论证。A组的合同中注明他们要求薪资比管辖区内其他群体高百分之十，因为他们区域的生活费最高。而B组提出无论生产率高低或其他任何因素，他们要求薪资在A组的上下浮动百分之十之内以保持竞争力。因此任何一组的薪资变化都会导致另一组的变化。

有时候我们使用回避问题实质来描述一个循环论证。论证的形式本质上都是一样的：“你知道的，某个行为之所是不合法的是因为它违背了法律。”“违背了法律”与“不合法”是相同的意思，所以此人在辩论某个行为是不合法的因为它不合法。

同样的，在市议会常常听到的意见是某个特定团体不支持增税因为他们将此作为纲领的一部分。当我们问为什么这是他们纲领的一部分，得到的回答是团体组员不支持增税。



要打破这样的循环，我们需要知道一个群体最初为什么通过了法律：它的初衷是什么？同样的，我们需要知道不增税对选民有哪些好处？不增税的话很有可能导致否定纳税人非常需要的某一项服务项目。在这样的情况下，我们需要问清楚，到底应该使用哪一种经济机制？

## 其他的谬论

当人们无法运用理性与证据的时候，便求助于各种其他类型的逻辑谬论。从青少年终年的诉诸大众：“但是学校里每个人都有，”到诉诸自然：“那就不是自然的。”并行逻辑，当然在专业领域中数不胜数。每一个城市或部门都有一个诺沃·维吉特，因此我们也需要一个。如今一个备受官僚喜爱的例子是在解释为什么我们保留现有的实践方法或者为什么事物保持原样的时候，挂在嘴边的是：“就是这样，”代替了之前被滥用了的，“机构的本质就是那样。”这些都是对我们毫无用处的逻辑谬论。

仅仅因为两件事物看上去有关联，不代表其中一件是另一件的起因——或者，事实上它们之间没有任何因果关系。

## 因果关系

我们还没有提到的一种谬论的正式的名称是因果颠倒，后此谬论，它的意思是“在它之后，也正是由于它。”学习过统计学的人会认出这是谬论“相关不代表有因果关系”的变体。

仅仅因为两件事物看上去有关联，不代表其中一件是另一件的起因——或者，事实上它们之间没有任何因果关系。其中，消防车谬论中就出现了这样荒谬的假设。故事出自于有一位火星星人来到了地球，他注意到有火灾发生的地方就一定有一辆消防车在场。所以，这位火星星人错误的认为是消防车引起了火灾。当然，这种谬论也应用在警车与犯罪现场，以及救护车与伤亡事故的关系中。

显然，相关或者相联的两件事物与它们之间的因果关系有一定的关系。问题是我们怎样辨别或者辨认因果关系。

这是一件重要的事因为在我们的每天看到的世界中思维形式上和想象中的因果关系比比皆是。每当我们遇见一些不完全了解的事物时，我们的第一个倾向是问一问，这事是怎么发生的？换句话说，原因是什么？

从历史的角度来看，因果思维方法是一个较新的理念。大部分学者使用大卫·胡姆的文字作为解释原因以及如何辨认因果关系的起点。

胡姆是1700年代早期至中期的一位苏格兰哲学家。他没有对问题作不必要的说明，他辨认出了形成因果关系的三个必要条件。第一个条件是因果必须是巧合的或如他所说的，“联合起来”。这里指出了两件通常同时出现的事物的相关性。



大卫·胡姆（1711年4月26日—1776年8月25日）是一名苏格兰哲学家，历史学家，经济学家，和以哲学经验主义和怀疑态度为代表的散文作家。他是苏格兰启蒙和西方哲学领域的重要代表人物。他被誉为“公认的用英文写作的最伟大的人物。”胡姆与约翰·洛克，乔治·贝克莱，以及其他一起被归类为英国经验主义的领导。

第二个条件是起因必须在结果之前发生。因为，如果上文提到的火星人在现场多逗留一会儿的话，那么他就会注意到先出现了火灾，消防车后到现场。所以，是因为火灾引起了消防员过来救火；而非消防车的存在导致了火灾。

第三个条件是最难分辨的，我们将它成为无虚假关系前提。无虚假关系是指起因不仅是结果的充分条件，且是导致结果的必要条件。反过来解释会更容易理解。无虚假关系的意思是没有第三方因素导致起因与结果同时出现。比如说，犯罪起数，警察人数，与管辖区内的人口三者之间的密切相关性。警察人数与区内犯罪起数之间并没有因果关系，但是两者都可能受潜在的人口密度变化影响而变化。

虚假关系是指两个或以上的事件是巧合的关系。而真正的起因却是潜在的第三个事件。这里的问题是尽管我们排除了那个明显的起因，依然是同样的结果。因此，在犯罪现场与警车的例子中，如果警察因为接到恶作剧的电话而出警，那么无论是否有犯罪现场，警车都会出现。从评估员或科学家的角度来看，无虚假关系是最难控制的因素。观察两件事物是否巧合，以及一件事在另一事之前发生都并不困难。

困难的是辨别是否存在其他潜在的机制导致了这两件事的发生。本质上，我们发明了两种处理虚假关系的方法。第一种是尝试建立解释性理论来解释某事如何或为什么导致了另一件事。术语就是，我们需要找出一个因果机制。逻辑语言里就是，X为什么能够导致Y？也如我们在生意场上问的，“是个什么故事？”

## 胡姆提出的满足因果关系的三个条件：

1. 因和果必须是巧合。
2. 起因必须在结果之前发生。
3. 没有第三方因素导致起因与结果同时出现。

处理非虚假关系的第二种方法就是通过人为的操控物质条件。就是说，我们能够通过使用物质复制出结果吗？我们将这种操控称之为实验。

随着时间的积累，为了能够将我们相信的起因与结果因素从其他可能的或虚假的影响中隔离开来，我们发明了一系列的实验设计或者操控状况的方法。在之后的章节中，我们会重点描述这些方法。

总之，我们充分的了解到只有满足了这三个条件，我们才能合理的相信两件事之间确实存在着因果关系。这些条件是发生巧合或相联的事件；起因在结果之间发生的时间顺序；以及，无其他潜在机制导致了显性的起因与结果的非虚假关系的前提。

不幸的是，我们进行过的很多研究都没有完全满足这三个条件。因此，我们常常会听说一些宣称某些因素（比如，石榴）可以降低患癌症的风险的医疗调查问卷。通常，这些属于相关性研究，比如有人通过调查发现了在吃石榴的人群中的患癌病人较少。我们可能可以确认吃石榴发生在患癌症之前的时间顺序。

这些研究通常不排除虚假的或混合的因素。比如，吃石榴的人群也许不抽烟，更注重锻炼，饮食均衡以及通常比不吃石榴的人群有更加健康的生活方式。这些更可能是影响致癌几率的真正因素。而与吃不吃石榴并没有关系。

当然，一旦我们开始相信吃石榴与致癌几率相关，我们可以生成许多的可能性因果解释。比如，我们可以辩论石榴中的大量维生素C或抗氧化剂可以抵制癌症。

## 将证据与解释联系起来

一种常见的误区是人们常常认为只要收集到充分的证据，他就可以“证明”那个假设或理论是正确的。事实上，一个合理的解释与组成证据的成分之间的关系十分复杂。

要证明一段关系，我们通常在需要使用数据或证据。首先，我们想寻求的解释必须符合与我们手头的大多数证据或事实。如果得到的解释不适用于我们大部分已知的事，那么它不太可能是我们需要的那个解释。一般我们有两种使用数据或证据的方式来证明一段关系。

一旦我们将貌似合理的解释范围缩小到了那些逻辑合理并且与现有证据相符的解释，那么我们就需要进行二次实验来证实这些解释在关键的情况下还站得住脚。显然，我们选择了一个与现有证据相符的解释，因此在同样的情况下收集到的数据不会为我们提供更确凿的证据。

### 要证明一个理论：

1. 我们像寻求的解释必须符合我们手头的大多数证据或事实。
2. 我们需要进行二次实验来证实这些解释在关键的情况下还站得住脚。

比如，一个事实是涂鸦乱画的市中心社区的犯罪率高于其他社区，但是这不能证明是涂鸦乱画的现象引起了犯罪行为。回到火星人的例子，经常看到火灾与消防车出现在一个地点不能证明它们间的因果关系。另一方面，只需要几次消防车没有出现灾火灾现场就可以反驳这个假设。

这是胡姆关于因果关系的讨论中最重要的一点。证明一件事情是真的远比证明它不是真的要困难许多。

胡姆使用的另一个例子是，仅仅因为太阳自从远古时期开始就从东方升起，从西方落下，这也不能证明明天也必然将会这样。另一方面，我们只需要找到一次太阳没有从东方升起就可以反驳这个规律的有效性。尽管这是一个特别的例子，但是它足以说明证明一段关系，与反驳一段关系所需证据的数量远远不等。

## 理论假设与虚无假设

在科学研究中，常常使用测试虚无假设真伪的方法来建立新的知识。这种方法首先需要提出一个能够回答问题的可信的真（或者假）的答案陈述。我们可以由假设的测试结果得到某些结论。<sup>1</sup> 比如，一个理论假设为在市中心交通要道繁忙时段征收交通拥堵费可以减少私人用车以及缓减交通拥堵。这听上去很有道理同时符合基本经济理论。另一方面，征收费用也可能没有效果。认为这个方法无效是因为许多人为了方便的驾驶私人汽车而愿意支付一些费用。此外，运货卡车司机这类的司机并没有选择的余地，无论是否会造成交通堵塞，他们都必须运送货物。

首先我们将看看与假设相反的理论是不是真的，以此作为理论假设是否为真的证据。就是说，征收费用对交通状况无影响。

这为我们带来了虚无假设：征收交通堵塞费与交通状况之间没有显示统计意义上的显著差异。如果我们无法证明虚无假设是虚假的，那么我们必须逻辑上拒绝理论假设，即征收交通堵塞费有效是真实的结论。

这是科学家们用来测试假设与理论的方法。我们无法直接的证明理论假设的真伪。但是如果证明了虚无假设（即，征收交通堵塞费对结果无影响）是真的，那么我们就采用理论假设。或者，至少我们需要严肃的重新思考理论假设到底是什么。或许纯粹因为征收的费用太少了。如果我们无法证明虚无假设，我们有强烈的理由相信理论假设是真的。当剩余的可能性越来越少的时候，理论假设的可信度越高。

## 案例分析：注射器交换项目

当地卫生局面临的重大挑战之一是静脉注射毒品使用者之间的艾滋病毒，肝炎以及其他传染病的传播。许多专业人士认为，注射毒品的次生影响比毒品本生对使用者造成的有害影响要严重的多。

为了减少或减轻注射器造成的有害影响，许多管辖区都成立了注射器交换与所谓的“安全”注射点。尽管有关注射器交换项目的很多的文章都提出这些措施很大程度的减少了艾滋病与其他血液传播的疾病，但是依然有许多的毒品使用者没有参加这些项目。在卑诗省的温哥华市，温哥华地区毒品使用者网建立了一个员工运行的延展服务网以补充更传统的交换项目模式。凯纳·哈雅士与她的同事们对这个项目是否达到了预期的效果进行了评估。<sup>2</sup>

### 理论假设：

在这个特定的实例中，理论假设是同伴实施的项目将减少通常不去传统交换点的静脉注射毒品使用者中重复使用注射器的次数。

### 虚无假设：

虚无假设是指对等网对重复使用注射器这一行为没有影响。

我们可以围绕着理论与虚无假设的而展开逻辑上以及行为解释的讨论。逻辑以及行为解释都可以作为提出的理论与虚无假设的依据。同伴干预的有效性的依据很大程度建立在这样的概念上：同伴是机动的，了解难以接触的人群的行为，以及可以在需要的地点为他们提供注射器。虚无假设的依据则基于这些难以接触的人群流动性大（通常没有稳定的住所），以及使用毒品使他们极度不愿意去理解注射器卫生的重要性。然而，最终，哈雅士与其他人断定“降低重复使用注射器与提供这项服务的享受机会有关。”



# 收集证据

## 环境扫描

我们制定的规划帮助我们设置优先事项，影响基于证据的决策方法，以及影响机构履行职责的能力。当出现问题，做出决策的时候，我们需要依照证据来辨别出我们的决定可能会带来的影响或者效应。政府部门可以使用这个方法来提高绩效，且保持领先于公众的期待。

收集这类信息的一种常用的方法是进行环境扫描。简单的说，环境扫描可以为我们提供机构内现状的有依据的，全面的情形。我们因此而了解内外部的真实情况，重要问题，以及对机构造成影响的趋势。此类信息可以证实或反驳我们的感知，同时可以用它们指导我们规划未来，政策上优先安排以及规划预算。在决定未来策略与发展合适、充分的响应方面，环境扫描也是非常有用的。

进行环境扫描对机构有那些益处？我们为什么要花费时间与精力去做扫描？以下列出了进行环境扫描可以带来的一些显然优势：

- 在看待机构的愿景与授权的范围内的问题上，提供一个新的、客观的角度，同时将它们最有效的进行排序；
- 提供一个访问信息、研究结果、统计数值以及其他第三方收集的数据的机会；
- 提供一个社区利益相关方、机构、个人参与并提供与他们相关事项的建议、看法以及意见的合作机会；

环境扫描可以为我们提供机构内现状的有依据的，全面的情形。我们因此而了解内外部的真实情况，重要问题，以及对机构造成影响的趋势。

- 在较大的团体中，提供一个发掘可解决问题的优势与有利条件的机会；
- 可以更好的了解自己机构内的优势与有利条件的比较框架或某一点；
- 提供一个了解你的机构的项目与实践对其他组织、机构、个人或团体造成的影响，以及这些项目与实践在多大的程度上有效的实现了授权的机会。

进行环境扫描是一个序列性过程，它包括从辅助源中收集信息，包括已发表的研究报告，统计数值，或其他。也可以使用你亲自与个人或者团体联系而得到的第一手或主源信息进行补充。此类信息的分析结果可作为你所在机构在更广泛的社会生态中定位的基础。

环境扫描与其他许多的管理过程不同，它并没有正规的格式。我们将为你提供这个过程的概述以及一些适用工具的建议。

## 环境扫描的类型

本质上有两种环境扫描。第一种方法不那么正式，你可以基于你自己的知识以及你或助手所能收集到的资料而自己进行。第一步是将你了解到的在相似情况下其他人采取的处理方式写在纸上。换句话说，你在观察与你身处同一个社会环境中的其他人的做事方式。

一般情况下，与同事关系紧密，交换阅读文献，以及定期参加会议与讲习班的人们更容易进行此步骤。

第二步包括了更正式些的评审。根据待解决的问题，你可以去阅读相关的期刊或新闻文章。网络是这些内容的好来源。像谷歌，冰以及网络爬虫这样的搜索引擎可以很快的检索到大量的信息。使用搜索引擎的挑战之一在于电脑很难从搜索结果中辨认出你真正需要的内容，因此，搜索引擎产生的糠多于麦子。

## 使用网络搜索引擎

使用搜索引擎有一些技巧。如果你足够幸运可以访问一名市区图书管理员或一所当地大学图书馆，那么那里通常有拥有丰富经验的人可以为你提供帮助。在下一页中，我们提供了一些使用谷歌搜索的窍门。

无论在网上或亲自访问一个图书馆，都可以找到专业文献。业内杂志与期刊通常覆盖了普遍性的问题。

一个好的开端可以是去与当地大学的一名熟悉你所在机构的使命的图书管理员，或机构内部图书馆的知识渊博的工作人员交谈。

想获取更详细的信息来源，你将需要进入正式的研究或学术文献的世界。因为许多种撰写学术文献的方法令人在阅读它们的时候让人感到畏惧。有些文章通俗易懂，而另一些则需要读者懂得有关主题的大量知识。关键是不要因此感到气馁。

有时候值得花些时间进行更广泛的搜索。比如加拿大政府联盟的官方网站上就有着丰富的信息。许多与当地政府合作的大学院系也是很好的信息来源，并且有一些潜在的联合项目。

## 有效的适用谷歌搜索

1. 务必具体。  
网站:[网页地址](site:[website URL])与你搜索的词组来找寻网站内的页面,用作者:[姓名](author:[name])搜索作者,以及标题内:[词](intitle:[word])搜索包含标题内的词的页面
2. 格式。  
使用文件格式:[档案或其他格式](filetype:[pdf or other extension])搜索图片以及所有格式的文件(比如文档与压缩过的图像文件)。
3. 扩大搜索范围。  
使用星号(\*)做为填补空白的通配符号。
4. 通过去除多余条件而缩小搜索范围。  
在你不希望包括的事项前加上一个减号。比如说,防盗一铃的搜索结果中将不包括“铃”。  
使用(两个点)做为范围指示来限制搜索结果,比如说,“二手的铲雪机2010..2014”将只会显示这4年内的结果。
5. 使用特定的搜索引擎。  
秀工具。网络爬虫则包含了一系列的搜索引擎。同时amazine([www.amazines.com](http://www.amazines.com))的网站提供免费文章的数据库。

## 信息源

专业与商业杂志是极好的信息源。虽然政府服务机构的杂志并不常见，但是政府内的各个部门，比如说消防局、人力资源部等都有自己的出版物。尽管美国的相关内容较多，谷路普（Govloop）网站是搜索针对政府部门内的相关问题的一系列文章的一个好起点。<sup>1</sup>

如果你需要与社区相关的特质信息与统计数值，加拿大统计局的网站上提供了大量的此类信息。加拿大统计局收集了多种主题的大量信息。<sup>2</sup> 同时，一些省政府也与统计局合作或独立的收集信息。本页底部列出了一些主要的省级政府的网址。

如果你希望进行数据比较而扩大搜索范围，那么，可以参考英国社区与当地市政府机构收集并发表与贫困，火灾与救

援服务、住房与无家可归、当地政府财政、以及发展规划与土地使用相关的正式统计结果。<sup>3</sup>

为了进行更有效的扫描，你务必需要通过多种方式收集信息。这样做你可以检查在多种信息源中是否凸显了同一个问题的存在。

有时候，与其他部门合作进行一项正式的扫描程序会给你带来很大的收获。这样的情况下，你可以考虑请一位外部的引导师来进行正式的扫描。在SWOT分析的相关章节的第二部分中描述了正式扫描的相关步骤。

环境扫描与SWOT分析的主要区别在于环境扫描中关注的问题覆函范围更加广阔。SWOT分析通常局限于机构面临的挑战与机会的相关问题。

### 省级政府数据源网址

卑诗省 <http://www.bcstats.gov.bc.ca/Home.aspx>  
 阿尔伯塔省 <http://finance.alberta.ca/aboutalberta/osi/>  
 萨斯喀彻温省 <http://www.stats.gov.sk.ca/>  
 曼尼托巴省 <http://www.gov.mb.ca/mbs/>  
 安大略省 <http://www.ontario.ca/government/ontario-open-data>  
 魁北克省 [http://www.stat.gouv.qc.ca/statistiques/index\\_an.html](http://www.stat.gouv.qc.ca/statistiques/index_an.html)  
 新不伦瑞克省 <http://www.snb.ca/e/0001e.asp>  
 新斯科舍省 <http://novascotia.ca/sns/access/vitalstats.asp>  
 纽芬兰与拉布拉多省 <http://www.stats.gov.nl.ca/>  
 西北地区 <http://www.statsnwt.ca/>  
 育空地区 <http://www.eco.gov.yk.ca/stats/>  
 努勒维特地区 <http://www.stats.gov.nu.ca/en/home.aspx>

## 制定你的环境扫描框架

你需要收集的信息取决于你想解决的问题。制定一个环境扫描的框架，首先我们从提出几个主要的问题开始，比如：

- 关键的问题是什么？
- 我们需要了解这个问题哪些方面的信息？
- 有哪些规律与驱动影响着这些因素？

一旦你围绕这些问题建立起一个框架，并且从主要信息源及辅助信息源获取到了研究结果，接下来就可以开始进行分析。

首先，我们需要考虑辅助信息源的研究中显现的主题、概念、问题或者隐忧。换句话说，其他的团体，组织，社区或其他地区的政府受到了这个问题的哪些影响？他们是怎样将这些问题按轻重程度排序的？

将调查结果与从你选定的焦点群体中生成的定性数据相比较。参考人们在一对一的面谈中表达出的想法。出现了哪些普遍的主题？调查结果呈现出了怎样的一致性与重复性？

试着想出这些人是怎样给在间接调查中出现的问题排序的？他们的想法与结果吻合吗？还是他们提出了在间接调查结果之外的不同的想法、方案或问题？

一旦你与你的团队在问题轻重排序的意见上达成一致，从最严重与紧急的问题着手，接下来你就可考虑能够帮助你解决问题的策略，项目活动以及实践方法。同时，你也需要考虑与实现战略重点相关的预算问题。

正如我们所说，SWOT分析常常伴随着环境扫描，它确定出影响机构达成其使命能力的内部与外部的优势、劣势、机会与威胁。本章节的第二部分将对SWOT分析方法做出更全面的解释。

下面介绍了一个正在进行的环境扫描例子。

## 例：加拿大招聘与社会发展：2015年春季安大略省劳工市场

加拿大统计局对安大略劳工市场进行一年一度的环境扫描。扫描结果包含了人口分布、经济状况与劳工市场环境的概况，且为识别加拿大服务部所提供的项目与服务中存在的潜在压力提供了信息。

这项扫描识别了以下四个关键点<sup>4</sup>

### 概况 / 经济环境

- 2014年的国际经济增长不均匀且不确定，导致经济衰退前的增长在现在看来无法持续。
- 在美国与英国经济呈现出明显优势的对比下，亚洲各国的经济状况不一，市场发展缓慢，以及欧元区因面临不断加大的经济与政治风险而停滞。
- 安大略最大的对外贸易伙伴—美国—在2014年预期经济增长为2.4%，同时预测在2015年的经济增长进一步加强3.6%，成为世界上经济最发达的国家。
- 2014年加拿大经济预计增长了2.4%，主要基于外贸优势，但是由于受到油价下跌的影响，而抵消了部分外贸带来的经济增长，在2015年预计增长仅仅为2.3%。

- 油价下跌可能导致安大略等由制造业为主导的省份赶上西部省份2015年的GDP增长。
- 由于对外贸易，特别是制造业相关的贸易将猛烈增长，安大略省的经济预计将在2014年增长2.3%，2015年增长2.7%。

### 省级与当地劳工市场环境

- 2014年安大略省失业率持续下降至后经济衰退期的7.3%，但是主要是由于找工作人数的减少。
- 2015年，预测的就业率将上升大约1.0%，预测的平均失业率为6.8%。
- 2014年安大略省15岁至29岁的人群就业市场有所改进。
- 服务业的就业增长率普遍高于比产品制造业，特别是在贸易（+28000），与专业、科研以及技术服务行业（+17200）。
- 慕斯科卡至卡瓦萨斯经济地区的就业率增长最快（+10.6%），与此同时多伦多地区持平（0.0%）。

基于这些与其他信息，你可能决定需要重新调整你部门或服务的项目。很明显，有几种选择。取决于你所在的地理位置（一个主要城市地区或者一个小型社区），你可能希望扩展或重组所提供服务的范围。

另一个选择是辨认出其他部门参与的四五个主要地区，并且将其作为你提供核心服务的重点。此外，你从环境扫描中得到的取决于你在一开始试图解决的问题。

## SWOT 分析

SWOT 使用起来非常简单。它与环境扫描可以很好的互补。只需稍加努力，SWOT 就可以生成至关重要的信息，同时它将那些信息构造成了一个可作为进一步分析的基础的框架。它是一个优秀的辅助决策过程的工具，并且帮助我们做出重要的决策——特别是正确的决策。

正如我们在本章节开始讨论的那样，环境扫描为你提供可以辨别出与研究主题相关的问题与关注的事物的主要与次级信息。结合这些信息与SWOT分析的结果，你将有更好的条件来辨别你策略优先顺序与未来方向。

通过部门内多人参与讨论存在的优势与劣势，SWOT分析提供了环境扫描之外的结果。通过观察部门的外部环境，你可以发掘出可被加以利用以扩大部门利益的机会。

SWOT分析同时探索存在的威胁：会对部门中正在进行的工作造成困难的外部因素、现实情况或者趋势。

有时候，我们与一名引导者一起进行小组SWOT分析。请每个组员提前完成问卷，可以使他们在小组讨论之前进行独立的评估。更简单的做法可以是给每个组员一张空白的SWOT模版，让他们提前填写，且带到会议中讨论。

### 进行SWOT 分析

尽管你独自就可以完成一项SWOT分析，但是如果由机构中多人参与其中则可以充分发挥该练习的优势。

管理者面对的矛盾之一是，员工与其他人期待领导做出决策，同时又希望可以参与到决策过程中来。与其他的活动一样，咨询是需要付出代价的。当你向员工咨询的时候，他们就没有在做本职的工作。

而且，团队动态可以产生意外的结果。

团体政治会显现出来而且红鲤鱼现象将浪费许多的时间。因此，由外部的引导者来带领这个练习常常很有效。咨询可带来的益处则数不胜数。

首先，团队通常能够产生出至关重要的方案，而一名管理者或者甚至一个管理团队往往却会忽视。其次，机构内不同部门的人们与不同的受众、供应商、社区团体、客户或消费者、监管者、以及其他服务供应者打着交道。

这为机构提供了不同的视角，特别是外在影响的方面。第三，有些参与者的意见尽管没有被采纳，他们也会觉得自己在决策过程中有一定的发言权。通常这样会提高士气以及在需要做出选择与执行不同的选项的时候，制造出更多的“大宗买进”。

在以团体的形式进行分析的时候，首先被提出的问题之一是：将会参与讨论的有哪些人？多种多样背景、职位的参与者可确保得到最全面的评估结果。

尽管不能完全保证，但是这样可以增加关键方面不被忽略的可能性。一般来说，SWOT分析应该至少由中层管理人员，最好是更高级别的管理者领导进行。

此外，这项分析应该包含机构中各部门的代表员工与关键位置上的管理人。不过，虽然并非总是如此，但是机构中的领导往往对外部与内部的问题更有洞悉。这来自他们的经验以及与机构内外各种人的关系。

在开始分析与填写矩阵之前，将环境扫描的结果提供给团队成员阅读常常是值得的。

确保你将环境扫描与SWOT分析背后的研究指导问题包含在内，它将组成讨论的内容框架。为讨论会定下些有帮助的基本规则。

## 例：洛维奇镇的SWOT分析

以下是洛维奇镇为了调查当地企业老板与主要社区利益相关者的意见与观点而进行的一项SWOT分析。<sup>5</sup>

这份报告的目标是通过社区利益相关者的视角，提供对洛维奇镇经济发展情形与潜力更深入的了解。

### 劳动力与教育

优势	机会
<ul style="list-style-type: none"> <li>在制造业与商业内，镇区内有技术性的劳动力—许多当地企业老板能够聘请到能胜任的职业。</li> <li>伍德斯道科与英格索的汽车业为当地的青少年提供了技术性贸易机会。</li> <li>许多当地的企业都有实习或学徒机会。</li> <li>镇区以当地学校为主要特征。</li> <li>洛维奇区高中提供创造性的课程。</li> <li>洛维奇区高中具备优秀的学生实习项目—许多当地公司都支持该项目。</li> <li>镇区附近有好几个专校和大学。</li> </ul>	<ul style="list-style-type: none"> <li>该区内汽车业的增长将为当地居民提供额外的工作岗位以及吸引更多的劳动力—推广小镇生活方式。</li> <li>加强社区保留洛维奇高中的意愿，通过认识到需要一个新的模型。以下包括了识别的可选项：               <ul style="list-style-type: none"> <li>- 使用7-12年级的模型扩张年级；</li> <li>- 在高中内建立一个农业培训项目做为独特的课程；</li> <li>- 考虑向当地居民增收附加税做为保留学校的费用。</li> </ul> </li> <li>加大推广高中实习项目给当地公司。</li> <li>将事业日当成一个重要的帮助青少年在当地就业的重要途径。</li> <li>为当地公司提供劳动力规划和吸引 / 涉及 / 培训员工的相关培训机会。</li> </ul>
劣势	威胁
<ul style="list-style-type: none"> <li>吸引 / 聘请 / 保留一些技术性的专业人士比较困难（比如，经理，工程师，焊工）。</li> <li>洛维奇高中和当地小学的未来有不确定性。</li> <li>损失学校实习项目会导致降低青少年接触当地就业机会以及当地缺乏技术性劳动力。</li> <li>当地社区中仅有少量青少年就业机会。</li> <li>对高中实习机会感兴趣的人数超过市场上可提供的工作机会。</li> </ul>	<ul style="list-style-type: none"> <li>农业劳动力减少。</li> <li>技术性劳动力因为认为大公司可以提供更好的工作机会而搬离镇区。</li> <li>为学生工作项目提供资金对许多公司都十分重要—失去资金会降低这些机会的数量。</li> <li>失去洛维奇高中 / 学校重组：               <ul style="list-style-type: none"> <li>- 当地教育类工作减少。</li> <li>- 依赖学生的当地公司损失付得起的劳动力。</li> <li>- 失去实习项目以及青少年接触当地工作的机会。</li> <li>- 对留在 / 搬来该区的家庭和公司的抑制更大。</li> </ul> </li> </ul>

## SWOT 讨论会的基本规则

- 每次重点讨论一个观点。
- 聆听、理解，并感谢他人的意见，避免打断或批评他人的发言。
- 设置合理的时间限制来保证讨论的持续。互相尊重—接受不同的观点与视角的出现。
- 对如何管理手机与助理人员带来的干扰达成共鸣。我们建议将手机关机且行政人员只对紧急情况做出回应。
- 机密性：哪些内容可以对外谈论？最终这些信息将去到哪里？将怎样保护匿名？
- 所有的成员都应该参与讨论。

在团队对环境扫描的结果中显现的问题与隐忧进行考虑的时候，请他们仔细考虑SWOT分析中的每个方面以评估该怎样更全面的解决这些问题与隐忧。

在你进行 SWOT 分析的时候，切记以下几个因素：

- SWOT分析是一个主观的过程，并不是某项科学。不过，由环境扫描中得到的定量与定性的数据将帮助参与人员相信所得到结论是有理可依的。
- 通过重点考虑少数几个问题来保持其简单性。如果显现出其他的问题，你可以在之后的过程中去解决它们。如果不设置这些限制，这个过程将会牵涉到过多数据与信息。
- 列出机构的实际优势与劣势。建立安全与透明的会议环境，这样参与人员会诚实的发言。

总之，综合使用SWOT分析与环境扫描以建立真实的、经过调查的、以及得到内部员工与外界利益相关者支持的战略规划。基于证据的决策方法得益于使用类似的工具，使得计划与决策基于可靠的事实与研究结果，同时由多种视角与意见指引。

## 注释及参考文献

1. <https://data.govloop.com/Government/Government-Trade-Magazines/ztqm-a4sy>
2. <http://www.statcan.gc.ca>
3. <https://www.gov.uk/government/organisations/department-for-communities-and-local-government/about/statistics>
4. See <http://www.esdc.gc.ca/eng/jobs/lmi/publications/e-scan/on/on-escan-spring2015.pdf> for the full document in pdf format.
5. Township of Norwich SWOT Analysis Report June 2008. Prepared for: The Corporation of the Township of Norwich by Harry Cummings & Associates Inc.  
<http://hcaconsulting.ca/pdfs/2008%20Norwich%20Township%20SWOT%20Analysis%20Report.pdf>

# 统计学

## 一种决策的工具

统计学可能是最常被误解的学科。大多数的大学生都害怕学这门课，同时大多数教授也不太愿意教这门课。此外，这个学科常常被辱骂为骗子的手段。马克吐温曾经说过，“世界上充满了谎言，该死的谎言和统计数据。”但是，如果用的恰当，统计学可以成为决策者的工具箱中最有用与强大的工具之一。

我们怀疑统计学的坏名声来自两个源头。其一，许多人认为它是数学的一种——大多数人在高中就学不好的科目。第二，许多教统计学的老师本身并不是统计学家，因此他们可以熟练的使用技术细节，但是却很少能够掌握根本的逻辑。统计学中需要应用到一些数学知识，但是涉及的难度不会高于平衡一个支票账户所需要的数学。理解统计学的关键在于将它看成通过一种整理的方法，以及它本质是试图将充满了不确定性的世界体现出意义。事实上，统计学的其中一个定义就是在不确定的条件下做出决策的科学。

---

理解统计学的关键在于将它看成通过一种整理的方法，以及它本质是试图将充满了不确定性的世界体现出意义。

---

对大多数的决策者来说，重点不在于纠缠于统计分析的细节，而是理解某个活动背后的本质原理或逻辑。本质的原理一般只有几个，而且通常很简单。一旦理解了，哪怕缺乏所涉及的数学或技术方面的知识，也可以充分的应用统计学的原理进行分析。

统计学包括两种基本的活动。第一种是出于想描述某件事的目的而收集数据。第二种是在使用数据的辅助下做出决定或推断。第一种活动我们称之为描述性统计；第二种我们称之为推论统计。

## 关于测量的讨论

我们常常将观察与纪录数据的过程成为测量。统计学家与普通人对测量的看法差异在于，统计学家假设所有的测量都存在误差。换句话说，在统计学中，出现100%正确的测量结果的原因是运气，而不是高水平的管理。从统计学的角度来看，测量中存在的误差有两个基本来源：固有误差或不稳定性，以及运算 / 操作误差。

固有误差或不稳定性是指被测量的对象的属性。比如，你请别人告诉你他们对工作的满意度，用百分制来描述（假设0代表最不满意，100代表完全满意），他们可能回复71。如果你在不同的情况下反复问他们同一个问题，他们给你的答复很可能在接近71的范围之内。

实际上很多人都知道自己对工作有多满意，但是很难给出一个准确的数字。此外，哪怕他们对工作完全满意，但是在不同因素的影响下，他们确切的满意度会变化。比如一天中的不同时段、他们是否刚刚与上司发生了争执、甚至天气。因此尽管在某个范围内是稳定的，但是大部分人对工作的实际满意度是天生不稳定的。

体内酒精测定器分析也是一样。呼气分析法是酒后驾车中测量血液中酒精含量(BAC)的最常用的方法。假设警察两次分别测量了同一个司机。如果司机刚刚打了个嗝或者呕吐了；或者手机或警方的无线电对测定器造成了电子干扰；或者周围环境中有人抽烟，有很多的灰尘，或者湿度很高；都会造成前后两次测量结果的差异。因此，从统计学的角度来看，血液中的酒精含量也是天生不稳定的。

固有误差是被测量的对象的属性—比如体内酒精测定器，可能会被你口中的酒精含量影响。

操作误差与我们如何进行测量有关—比如，测量器的问题或者我们读取的方法。

在固有偏差之外，还有运算 / 操作误差的存在。比如警察忘记了手动校准仪器；电池中的电不足；仪器被使用的不当；测定器的类型不完全正确；纪录结果的时候将0.08误写为0.8；或者在测量与纪录的间隔中，警察忘记了实际数字是多少。

重点是，我们尽可能的去尝试得到完全准确的测量结果，但是往往是十分的困难，如果不是完全不可能的话。相信我们可以做到完全准确是自欺欺人。此外，在大多情况下，“接近某个数值”已经足够说明问题了。酒精测量结果是0.08或者0.085有什么影响？统计学之所以强大的原因之一就是它假设了我们的测量中存在一些误差。

统计学的另一个优势在于，当适当的使用它的时候，我们可以预计到在测量过程中存在多少误差。

统计学家认为，相信能够得到完全准确的测量结果的人就如同埋头在沙子中的鸵鸟一样。接受测量中到处都可能存在误差这个事实是更明智的选择，那么为什么不承认并且尝试着预测出误差的大小呢？我们可以怎样做出预测？答案是需要同时进行多次相同的测量，或者测量多个类似的事件。



统计学家认为，相信能够得到完全准确的测量结果的人就如同埋头在沙子中的鸵鸟一样。

## 描述性统计

记得一件事物的特性相对简单，无论这个事物是一个人，是一个类似月食这样的事件，或是某个人的汽车颜色。同样，大部分人都可以轻易的回忆起几件事物的特性。事件的数量越多，我们越难回忆起群体中某一个单独事件的特性。比如，我们记得办公室中同事的年龄，但是几乎不可能记得一个地区或城市中所有志愿者的年龄。如果我们希望说些关于志愿者年龄的什么，那么我们需要合计或概况这些数据。这时候就需要使用描述性统计了。

描述性统计概况一个群体的特性以使大量的数据有一定的意义。就算我们能记得，但是列出600个人的年龄并没有太大的用处。描述性统计使我们可以辨别出列出数据的一些有用的特性。通常，在一系列的观察结果中我们希望了解的两个方面是：典型的特性是什么，以及有多大的可变性？

算数平均数或均值被最常用来衡量典型的特性。比如，我们算出志愿者们平均年龄62岁。其他的衡量典型的特性的方法包括中位数与模式。中位数是年龄在其之上与在其之下的人数相同的数值。这个例子中的中位数可能是55岁。换句话说，志愿者中有一半的人数的年龄大于55岁，另一半小于55岁。模式是指最多人数的年龄。我们也可以将这些测量看作是年龄分布列表中的中心定位点。

衡量可变性可以告诉我们数据的范围分布有多广或者变化有多大。了解到志愿者的平均年龄为62岁是一方面；但是了解到他们之中的大部分人在55岁与70岁之间，而不是在60岁到65岁之间则是另外一方面的信息。最常见的可变性的衡量方法叫做统计范围与统计方差。

统计范围指测量两点之间的距离。比如，志愿者中最年轻的是24岁，最年长的是71岁。范围则是72减去24，或48年。测量范围就是找出分布中最小值与最大值的差别。

最小—最大值范围是有用的，但是有时候会给人一种错觉。比如，志愿者中最年长的是78岁，但是其他上年纪的志愿者都不到60岁。这样的情况下，我们将78岁称之为数据中的异常值。

当处理数据分布中存在的古怪极端的情形时，我们有时候使用一种叫做四分范围的统计方法。首先我们需要找出在25%的人数之上的年龄，以及在75%的人数之下的年龄。四分范围就是这两个年龄数值之差。最小—最大值范围，四分范围使我们更多的了解了年龄分布的状况。

在范围之外，我们常常通过统计方差来了解数据相对集中趋势是怎样分布的。最常用的两种统计方差是方差与标准偏差。猛一看，这些统计方法有些难度，但是其实非常的简单。重点是要理解基本概念而非数学知识。

描述性统计概况一个群体的特性以使大量的数据有一定的意义。

算数平均数或均值被最常用用来衡量典型的特性。

衡量可变性可以告诉我们数据的范围分布有多广或者变化有多大。

## 衡量典型性的例子

在决策的过程中，最简单的描述性统计都十分有帮助。我们首先来审查一周内在两个临近的镇中开出的停车罚单（7天）。汇报的每天开出的罚单列在了下面的表格中。

		A镇	B镇
	时期	停车罚单的数量	
	周一	10	9
	周二	11	16
	周三	8	7
中值点	周四	9	7
	周五	9	8
	周六	9	7
	周日	5	7
总数		61	61
平均数		9	9
中位数		9	7
众数		9	7

我们可以计算出平均值或算数平均数，中位数与众数来衡量典型性。平均值或算数平均数就是罚单数的总值除以总天数。中位数处在一个有50%的数字各在其之上与之下的点上。众数则是出现次数最多的罚单数。

表格中的数据代表了开出的停车罚单的实际数字。从这有限的信息中，有好几个兴趣点。首先，两个镇一周内开出罚单的总数都是61张。平均每天9张。我们注意到B镇有一天就开出了16张。在统计学中，我们将这样的特殊数值称为异常值。算数平均数受异常值的影响特别大。如果我们将16换做30的话，就很显而易见。其他的数值都保持不变，但是平均数将增加到10.7。

受异常值影响较小的是中位值（有时候也叫中位点）。正如我们已经解释过的，中位数是将数据分布分为50%之上与之下的数值。在表格中，中位数或中位点是9，刚好是一个星期四，一周中间的那天。A镇的中位数是9，B镇的是7。

B镇的中位数较小是因为除了一个16的异常值，总体上它比A镇开出的罚单数要少。因为我们的数据较少，因此很显而易见。在较大的数据组中就没有这么明显。尽管原理都是一样的。

## 衡量可变性的例子

在这个例子中，我们将继续使用上面的停车罚单的数据。我们得知两个镇开出的典型的或平均的罚单数量是相同的。可是，通过观察原始数据，发现B镇中的可变性可能要高于A镇。B镇中的平均数与中位数略有不同从数据上支持了这个观点。

	A镇		
	罚单数量	平均偏差	偏差平方
周一	10	1	1
周二	11	2	4
周三	8	-1	1
周四	9	0	0
周五	9	0	0
周六	9	0	0
周日	5	-4	16
平均数	9	0	2.9
	B镇		
周一	9	0	0
周二	16	7	49
周三	7	-2	4
周四	7	-2	4
周五	8	-1	1
周六	7	-2	4
周日	7	-2	4
平均数	9	0	9.4

范围是衡量可变性的指标之一。A镇的罚单数量是5至11，它的范围是6。B镇的罚单数最少是7，最多是16，范围是9。

另两个常用的衡量指标是方差与标准偏差。尽管猛一看有些复杂，其实它们的理论非常简单。我们将每个独立的罚单数减去平均值得到上表格中第二竖行的数字结果。比如，在A镇中，第一个平均偏差是 $10-9=1$ 。以此类推。

接着将平均偏差平方就得到了第三竖行的数字结果。将每一个原始数据都这样处理之后，我们发现了两个现象。第一，平均数的平均方差为零。这将总是如此，因为平均数是数据分布的中心，正的平均偏差与负的会互相抵消。因此，我们计算出了第三竖行的数字结果：偏差平方。

偏差平方的平均值叫做方差。A镇的方差是2.9，B镇的是9.4。这说明B镇每日开出的罚单数量的可变性较大。统计学中的方差非常的重要。在稍微高级的统计学中，我们的目标是尝试找出出现了方差或者一组数据的方差大于另一组的原因。或者B镇在周末有大型的户外农贸市场，因此吸引了许多外地人，这也可以解释周末开出的罚单数较多。这些都属于我们可能想测验的概念或假设。

因为偏差平方通常是大的数字，因此我们通常比较偏差平方的根。这样数字结果与原始数据差不多大。偏差平方的根叫做均方差。A镇的均方差是1.7，B镇是3.1。这说明了B镇每日开出的罚单数量的可变性略大于A镇。

## 题外话

我们也可以通过使用统计方差来预测测量中存在着多少误差。比如，两个人平均需要花费25分钟完成一件活动。其中一个人的方差是8分钟，另一个人是3分钟。这两个人的表现相当，但是方差说明了第二个人的表现相对稳定，也就是说更好。从管理层的角度来看，有趣的问题是为什么一个人的表现比另一个人更稳定。

再深一步的调查结果也许会表明第一个人需要在不同的环境下执行任务，而第二个人较少面对环境的挑战。也可能是因为第一个人开始的时候并不那么认真，而之后努力赶上了进度。

不管怎样，方差的差异可以为我们提供平均数或中心趋势的差异之外的信息。

## 推论统计

统计学的第二回合被我们称为推论统计。推论统计帮助我们得出结论与做出决策。与大部分的描述性统计不同的是，推论统计背后的数学可能会变得足够复杂。因此，我们重点关注在推论统计背后的逻辑与考察如何使用它帮助我们做出决策。通常从书本中自学描述性统计并不简单。对于缺乏这个区域的背景知识的读者来说，去当地的大学修一门一学期的课程是很值得的。否则，理解它的概念就足够了；把细节留给专家们吧。

推论统计的用途非常的多。然而，最主要的两个是从样本中预测或推论出群体的特性，并且预测两个或更多的群体或样本之间是否存在显著性差异。

## 人口估计

我们首先来讨论从人口中的样本进行推论。如果想知道一个城市中一氧化碳(CO)探测器的人口比例，我们可以与每个家庭联系，并提出这个问题。

向一个管辖范围中的每一个人收集信息叫做人口普查。在一个有30万户家庭的都市中，这是一个昂贵并且耗时的建议。这也是极少并且在限定条件下才进行人口普查的原因。幸运的是，在二十世纪初期，统计学家想出了如何从子群或样本推测出整体（群体）的特性的方法。

从人口中收集样本是这个方法的关键所在。本质上，使用“任何旧样本”都抄不了近路。样本必须通过特定的方法从人口中收集。主题有时会有些不同，但是在这里我们保持简单并且考虑基本的情况。统计学家有时候把这种情况称为简单随机样本。简单随机样本指的是人口中的每一户家庭被选中的机会同等，并且它被选中的机会独立于其他家庭的被选中机会。我们将这个定义的组成部分分解为：随机选择，同等机会，以及独立性。

## 随机选择

随机选择指的是使用一种随机机制来选出样本中的家庭—比如投掷硬币或者电脑随机数生成器。换句话说，不能依据可用性或者大门的颜色来选择。比如说一个城市中登记目录上列有1至30万户的家庭，我们的样本需要1200户家庭，随机选择就是指我们使用随机数字生成器从1到30万中选择出1200个数字。一旦我们得到了这些数字，我们就可以从列表中选择出所出位置与之相对的家庭。

## 同等机会

同等机会指的是每一户家庭被选中的机会是同等的。在有重复的地址或者漏掉了某种家庭（比如说，所有的公寓或者某个地区中的所有家庭）的列表中，会导致某些家庭得到更多被选中的机会，或者完全没有机会。

## 独立性

这是指是否选择某一户家庭不能决定或影响其他的家庭。比如，选择样本的人可能会注意到列表上出现两户家庭都居住在同一条街区，或相邻。研究员可能因为两户太相似而只留下一户。这是不能被接受的，凡事随机选中的都必须被包括在列表之中。

如果我们不遵守这些规则，那么就有可能遗漏掉某些从样本中推测到的群体的特性。其他一些需要被考虑的如，样本的大小，但是这些属于最好与专家讨论的细节。如果我们尊重基本的规则，我们可以预测出使用一氧化碳探测器，以家庭为单位的群体比例的大概范围。

换句话说，通过样本可以预测出群体中确切存在的近似特性，但是可能无法达到精确。统计学样本与其他的方法不同的是它可能可以预测出群体特性所处的范围。因此，我们可以从发出的20份问卷，回复的19份中推断出家庭中装置了一氧化碳探测器的比例是百分之X，误差正负Y个百分点。

不知情的人常常因为统计预测不能提供确切的数值而低估它。但是，正如我们之前所说的，统计学的基本假设就是所有的衡量都存在着误差，因此，我们能做的最好的是得到一个估值点以及其精确程度的合理解。这是其他的方法做不到的。通过统计预测，你可以了解到某个预测是否足够精确而对你有用或变化太大或对某种实际用途来说不够精确。

还有许多可用的预测方法，但是它们无法告诉你预测结果是接近群体中的确切数值还是更接近火星。

## 显著差异

推论统计的另一个用途是可以预测出两份样本是否相似或不同。比如，一位警察局长或消防局长可能想了解一年中不同的局站中的响应时间是否存在差异。通常，响应时间的数据由自动派遣系统收集。在一段时间结束之后，可以计算出平均数或平均响应时间。正如前面所讨论的，平均值是根据容易出错的数据预测出来的，而且有一些数值分布在预测平均值的附近。因此，问题是，如果一个部门的响应时间是8分钟，另一个是9分钟，它们之间相差的那一分钟是真实的差异，还是可能存在的测量误差？

某些巨大的，有重大意义的差异不需要使用统计方法来帮助我们做出决策。比如，如果响应时间的差异是10分钟，那么我们就知道这是一个真实存在的且重要的差异。但是，如果是一分钟的话，我们很难分辨这是真实的还是属于正常变异的范围。

统计学可以告诉我们某个差异是处在正常变异的范围之内还是之外。如果在范围之外，那么我们将这个差异称为统计显著性差异。

我们要注意的，仅仅因为某件事存在着统计显著性差异，不代表这个差异有巨大的意义。比如，甲市居民平均每年比乙市居民多赚100元。

这个差异可能有统计显著性差异，但是如果两个城市的居民平均年薪在7万元的话，基本上没有人会认为这个差异有什么重要性。另一方面，如果某件事不存在统计显著性差异的话，那么我们应该将它当作正常变异范围内的差异来看待，因此，也不具有巨大的意义。

当我们面对多种对比的时候，推论统计更加有效。大城市中通常有上百个十字路口。这些路口的交通事故发生率是否显著性的不同？更高级的技术可以帮助我们辨别出哪些因素与这些差异相关。这把我们带到了本章节的最后一个话题，就是统计模型的角色。

---

## 统计模型

对于大多数的决策者来说，统计学的真正的能力在于它可以模拟社会，自然与机械过程。统计模型为我们调查受到多种因素影响的复杂的问题的时候提供了帮助。比如，统计模型一直以来都用于研究附近公路上的交通噪音。交通噪音是一个地区的整体的噪音污染的一部分，这是公认的事实。高速公路上的交通噪音给附近地区造成了干扰，特别是高流量与高速度的区域。有些人认为这些噪音是对环境健康的严重威胁。<sup>1</sup>

在统计模型中，我们集中了大部分的注意力在尝试解释存在的变异性上。

因此，我们回到基本的统计理论之一上一变异性。所以，比如说，我们可能提问这样的问题：在不同的地段或者不同的时间，有哪些因素或许影响了交通噪音？是交通流量因素，道路因素，交通工具因素还是人为因素？<sup>2</sup>根据分析结果以及使用其他的模型，很有可能辨别出哪种干涉方式可以有效的降低噪音，以及哪种方式无效。

决策者的工作还包括改善资源分配与使用效率。不论模型或者背后的程序有多么的复杂，统计模型都能帮助我们获取可预测准确率的许多有用结果。

以下列出了一些主要的可通过统计模型解决的问题：

- 整体模型是否准确的反应了我们尝试描述或模仿的过程？换句话说，它是否具有统计显著性？
- 模型解释了多少输出因素的变异性？
- 模型中的哪些元素具有统计显著性？哪些不具有？
- 模型中的各种组成部分对输出因素造成的影响相对大小或排序是什么？
- 所造成的这些影响是否具有重大的或原则上的意义？
- 模型中不同的替代组成部分如何互动以影响结果？

正如我们指出的，统计学并不是所有决策的灵丹妙药。但是，适当的使用统计技巧可以对我们所研究的问题提出深入的见解。

决策是一个复杂的过程，且最好的过程是我们在多种分析工具的帮助下得出某个结论。我们常常需要权衡利弊。某些因素可能呈现统计显著性但是没有重大的意义。同样的，仅仅某个选择相对更加有效，但是不代表它在社会学或经济学方面合乎情理。不管怎样，知道某件事是否会造成实际的影响是个很好的开始。

---

## 注释及参考文献

1. Subramani, T., M. Kavitha and K.P. Sivaraj (2012) “Modelling of traffic noise pollution.” International Journal of Engineering Research and Applications 2: 3175-3182.
2. Subramani, T., M. Kavitha and K.P. Sivaraj (2012)

# 设计实验

## 我们怎样知道它是什么意思？

此书潜在的一个概念是基于证据的决策方法具有特有的优势。这里所指的证据是可观察与可测量的“事实”或数据。尽管我们主张一般情况下有用事实根据是件好事，但是单独的一个事实或一小部分数据或信息碎片基本上没有什么意义。因为除非与其他事物的比较，否则没有什么具备意义的。

假设，在你去印度的旅行中看到出售的一双鞋子的价格为2859卢比。如果你不熟悉印度的鞋子出售价格，你可能不清楚这是否物有所值。除非你有其他可以对比的事物，否则这双鞋子的价格为2859卢比这个事实对你毫无用处。可以对比的事物包括其他商品的价格或印度平均的时薪或在其他货币中同等购买力。当前，2900卢比大约等同于60加币。只有通过比较，鞋子的相对价值才有意义。

同样，你的当地警察局总体犯罪破获率为百分之40，其中百分之70是暴力犯罪。一次市议会中，有人提问这样的业绩是否在可接受的范围之内。市民的期望平均是消灭或破获百分之90以上的可被指控的犯罪活动。加拿大统计局的全国报告数据显示，加拿大全国范围内犯罪破获率大约为百分之40，其中暴力犯罪占百分之60。通过比较，很明显你的当地部门的总体犯罪破获率与国内其他地区一致，且在消灭暴力犯罪方面做的更好一些。

---

我们需要适当的比较点来理解一件事实的意思。

---

这里强调的重点是我们需要适当的比较点来理解一件事实的意思。

在基于证据的决策方法的框架中，一个我们需要问自己的关键的问题是：最适当的比较点是什么？一个辅助问题可能是：做出比较的最好方式是什么？答案是通过使用一个被程序评价员与应用统计学家称作实验设计的标准框架。简单来说，实验设计是帮助我们做出适当比较的另一种方法。

本章节的余下部分将集中讲解一些基本的实验设计，我们通常用来评估信息或数据对解决与决策相关的问题的价值。在应用研究中，实验的设计可能会变得非常复杂。但是，无论设计复杂与否，本质上都有一些有价值基本原则。

## 反事实

当我们做出或观察一些事物的时候，我们会问：如果这件事没有发生的话会怎样？在第二次世界大战中，如果核心国赢了会怎样？如果执政党输了上一次的选举会怎样？如果不是市政府而是保险公司负责提供消防服务会怎样？

将这些事件与理论模型做出比较。虽然不能为我们直接提供证明，但作为一项脑力练习，它迫使我们辨别出某项政策或项目中的重要因素。有哪些关系重大或有可能造成一定影响的相关或有效因素？这些与已有的结果或事实相反的概念叫做“反事实”。

阿尔伯特·爱因斯坦将对于反事实的思考称为思想实验。思想实验包括通过我们头脑中进行的分析，来思考一件特定的事件或结果存在的潜在影响与后果。爱因斯坦的思考实验与简单幻想或推理的差别在于前者专注于我们采取怎样的方法来检验思考实验。

思想实验的一个例子是，我们可以思考因为工作性质而引起的压力大小如何影响了员工的工作表现；比如说，感知风险，长时间工作，工作时段以及责任级别。我们认出这些因素都是压力源，但是实际上它们对员工的工作表现会造成影响吗？在我们的思考实验中我们也要考虑到其他的因素，例如机构中的压力源。

根据经验我们可以得知对工作造成影响的也有其他的因素。给员工带来的最大压力可能并不是工作本身的性质，而是机构与机构中的人事引起的压力。或许是缺乏对出色的表现给予奖励或赞誉影响了工作表现。

我们也应该考虑工作环境中其他可能带来办公室压力的因素，比如机构结构与机构活动中的各个方面（比如，同事间的关系，培训，资源，领导以及监督）。通过这个思考的过程我们得出的结论是：工作的性质并不是导致工作压力的唯一源头，而例如工作环境这样的其他压力源也是主要元凶。

在仔细思考之后，我们得到了一个合理的结论。尽管，本身逻辑合理的结论并不总能解决看得见的世界中的问题。我们需要的是基于反复观察到的坚实证据—这些证据不光存在于我们的头脑中，而是可以被其他人看的到，分享以及评估。

## 有力证据的构成是什么？

当我们从事基于证据的决策的时候，最本质的问题是：适当证据的构成是什么？如果我们将科学看成是寻找某个问题的“真正的”解释的一种机制，那么可以在诉讼案件的背景下思考这个问题。在法庭中，如同在科学中一样，有着数量各异的证据。

即使某件事情本质上是真实的，我们将一些证据看的比另一些更合理，更可靠，以及更相关。科学也是一样。好的证据经得起严格的推敲。但是，好的证据是由什么组成的？

好证据的其中一个特性是经得起反复严格的测试验证。在科学的框架中，测试一个方案的基本机制是实验设计。实验设计是逻辑的物理应用，因此让我们分析实验设计背后的逻辑。

假设我们希望评估防盗报警器对住宅闯入犯罪率的影响。其中一个方法是为一个社区中的每一户住宅安装防盗报警器。我们就可以观察到闯入犯罪率在安装报警器的前后是否存在差异。

不幸的是，任何被观察到的差异都有可能是由其他因素引起的（回想我们之前对虚假论点的讨论）。比如说，因为巧合，住宅闯入犯罪率可能由于社区中年轻人口的减少，或者一个更积极的工作市场，警察巡逻的改变，或者一个更活跃的邻里守望计划而降低。我们知道这些因素与住宅闯入事件密切相关。

理想的情况下，我们希望能够同时观察到同一个社区在安装与未安装报警器的情况。换句话说，我们希望能够基于同一个社区参与与未参与安装报警器计划的结果差异，来评估这项计划的有效性。不过，我们了解这是不可能的。同一事物不可能同时存在于两种状态中。在任何一个时间点，社区要么参与了这项计划，要么没有参与。

这种无法观察到同一实体同时在不同的状态下被称为“反事实问题”。就是说，如果存在另一种情况的话，我们如何去进行估量？

如果相反或者反事实情况发生的话，但是我们却无法评估该状况下的结果，那么我们如何确认安装报警器改善了住宅闯入犯罪率，还是由于其他因素？科学家们与项目评估人员采取的方法是找到一个与学习小组尽量相似的对照小组。对照小组与学习小组的相似性对获取的结论的可信度有着至关重要的影响。

获取或者设计对照小组有许多的渠道，其中有一些比另一些更有优势。

本身逻辑合理的结论并不总能解决看得见的世界中的问题。我们需要的是基于反复观察到的坚实证据。

评估人员花费了许多的时间与精力考虑一个对照小组是否合适。

比如说，我们可能找到一个邻近目标社区的“姐妹”社区来做对照。另一方面，我们可以为每相邻两户中的一户提供报警器，或者仅仅为社区的南边住户提供。我们甚至可以考虑将目标社区与省份或地区中的所有其他的社区进行比较。

所有的方法都可以为我们提供一个评判目标社区中报警器的潜在影响的参考点。

然而，问题是，所有的选择都有潜在的限制。一些条件或者情况会导致目标与对照小组产生本质上的差别。有时候我们可以注意到这些差别。比如说，在选择“姐妹”社区的时候，那个社区的居民年纪普遍较大，而且整体上的安全设计较差。这对一个业余观察者来说都是显而易见的差别。然后，存在的差别却并没有这么的明显。

本章节的剩余篇章将重点讲述我们可以辨别有效对照小组的不同方法以复制或模仿反事实。无论我们评估的是哪一种项目，辨别出对照小组都是关键。简单的说，在不具备反事实的有效估计的情况下，我们无法建立起可信的项目的影响。

一步测试法无法从其他的角度解释结果。

## 与目标做对比（一步测试法）

其中一个最简单的设计是我们需要将整体与一个特定的目标或者标准相比较。一般而言，政策方针建立在立法标准或者由研究最优实践而设定的目标之上的。目标因背景而异。比如说，某社区的目标可能是五年内减少百分之二十的交通事故。某个制造商实施了六标准差的体制，具体是指所生产的每一百万产品中不超过3.4个次品。人力资源政策也可能指示机构聘请的员工中有一定的百分比人数是少数群体。

接下来的关键在于将我们的感兴趣的总体与一个理论上可行或可达到的目标相比较。

一旦我们开始实施某项行动，那么关心的问题就变成了我们是否达到了目标。如果我们达到了目标，我们就有理由相信这个行动（通常指实施的某项政策或者项目）成功了。当然，我们将使用一个统计过程来帮助我们判定是否离目标足够近从而称作达到目标。

方法论的文献中有时将这种方法称为一步测试法。就是说，将开展的某个行动、政策或项目与某个标准相比较，如果它达到了标准，那么一般情况下我们就假定这项行动是成功的。这样的证据看上去合理可信。不过遗憾的是，一步测试法有它的局限性。在接下来的例子中我们可以看到其中一个主要的局限性。

### 例：一步测试法

假设某个社区中因火警事故的死亡率为每百万人口中九个人，他们希望在接下来的三年中可以将其降低至每百万人口中五个人。消防局长可能认为要达到这个目标而需要采取的性能价格比率最高的措施是派放免费的烟雾报警器。他实施了这个项目且死亡率在三年之后降低到了每百万人口中5.1个人，按照该社区的规模，该结果在数据上可以等同达到了目标。我们可以由此推断死亡率降低的原因是派放了烟雾报警器吗？证据看上去很有力。

事实上，也可能存在着其他的解释。免费派放烟雾报警器活动在当地媒体中得到了大量的宣传。消防员与志愿者们挨门挨户的派放烟雾报警器。若没人在家，门上会留有一张通知户主去各类零售店领取的通知单。该活动使当地居民意识到了有关住宅消防安全问题。因为宣传，当地居民意识到了要注意消防安全以及在他们的家中做出了一些改变。有些人清理了火炉边的杂物，更少的人在睡觉后还使用取暖器，以及更多的人计划了万一发生火宅后逃离路线。

换句话说，通过加强当地火灾的意识，无论家中是否安装了烟雾报警器，社区成员都做出了降低因为火灾而导致人员死亡的行动。

这里的重点并非争辩烟雾报警器对降低死亡没有贡献。而是可能存在其他解释或者巧合因素从而达到了目标。其他解释的可信度取决于其他的因素。首先，其他的解释逻辑上是否合理？如果之前的公共宣传活动没有带来显著的影响，那么我们可能坚持烟雾报警器是有效的机制。

另一方面，如果公共宣传活动使其他社区的死亡率明显降低，那么我们就更加相信其他的解释。进一步的解释可能是因为建筑法规中的长期改变，全民防火意识加强，吸烟人数降低，或者其他等等原因而导致火宅引起的死亡率整体下降。因此不管怎样，死亡率都会降低。

## 前后比较设计

前后比较设计是一步测试法或者目标设计的一个变异。再次用之前的例子来说，我们手头有一组期望制造影响的兴趣小组或社区。我们预先考量局面，运用某些干涉措施，然后再看结果如何。这里的假设是前后结果的差异完全归因于所采取的干涉措施的影响。前后比较设计与一步测试法不同的是，后者的比较对象是一个政策上的目标，而前者的潜在比较对象是采取措施之前的基线。前后比较设计与一步测试法有着同样的优势与弱势。具体来说，我们永远无法确定结果的差异是采取的措施带来的影响还是由于简单的巧合。比如说，司法部希望降低年轻司机中的汽车事故率。该部门决定将25岁以下的司机最大法定允许的血液中的酒精浓度从0.08降低到0.05。通过审查实施该措施后的三年内的相关数据，一位审查者注意到了年轻司机中的事故发生率确实降低了。

再次申明，我们可能会认为司法改变是至关因素。另一方面，年轻司机中的饮酒次数与血液中酒精浓度也可能并没有什么变化。

差异可能仅仅是因为警察监控更加严格，他们在尝试加强执行新司法的过程中将年轻司机做为了主要关注目标。

就算没有新的司法，加强警察监控也很可能产生同样的结果。也就是说，重要的因素并非立了新法，而是单单加强警察对年轻司机的监控便可以起到威慑的作用。

### 忽略限制

这些设计中存在的局限性不代表所收集的证据是无关的。我们有充足的理由相信如果我们在大量社区中实施这项措施，在不同的情况下仍将会产生同样的结果。同样，先实施某项措施，接着再将之撤销也可以为我们提供许多信息。如果该措施产生了期望的结果并且撤销该措施之后回到了实施前的基线，那么我们就得到了一个说明该措施是根本原因的更加强大的论据。我们需要记住的是证据极少能够具备绝对性。就如同在法庭中一样，一些形式的证据比其他的可信度更高。

考虑到这些设计的固有弱点，我们可以探索可以使用哪些方法来解决问题。到此为止，评估员与科学家们之中的黄金标准是我们称之为的经典实验设计。

## 经典设计

一个经验法则是世事都是不完美的且确定性是抓不住的目标。另一方面，人们极少是由于对自己的死亡风险不确定而参与俄罗斯转盘。同样，某人在股市中失去所有的财产所存在的风险是百分之一，与百分之八十以上的风险对投资行为的影响是完全不同的。因此，如果我们无法达到完美，那么当下实验设计的典范或黄金标准是什么？

至今为止，评估员与科学家们依赖两组，前后设计以提供最有效以及最可靠的证据。首先我们从上文提到的前后设计着手。接着我们用一个做为反事实的对比或对照组进行补充。换句话说，我们有一个小组接受了某种治疗且另一组没有。如果接受了某种待遇的小组呈现出了某种显著的改变，且这种改变不存在于对照组中，那么我们有非常强烈的理由相信这种干涉起到了一定的作用。

这种设计的主要优势在于它可以保证对照组与实验组或者治疗组的初始状态相同。这可以追溯到我们之前对于反事实的讨论中谈到的，我们希望可以同时观察到相同的病人接受治疗以及没有接受治疗的情况。这个情况从物理上来说是无法实现的。不过，我们可以尽可能的确保治疗与比较小组的初始状态相同。这是如何实现的呢？

这种设计的主要优势在于它可以保证对照组与实验组或者治疗组的初始状态相同。

其中的一个方法是选取相似的人群（或者社区以及其他），并且将他们分为两组。但是，除非他们是一摸一样的克隆人，我们永远无法确保他们的所有的性质都相同。幸运的是，虽然我们在工作上基本上不使用克隆或完全相同的配对，但是我们将研究对象分为两个统计上相等的小组。正如我们之前提到的，统计上的相等并不同于绝对的相同，但是也足以说明，一般来说，两个小组不存在统计上的显著差别。换句话说，从实用角度来说，他们足够相似而可以被当成完全相同。

确保统计上的相等的方法是首先选取初始组，接着将组内成员随机的分配至治疗小组以及对照组。

随机分配的意思是，比如通过抛硬币（一枚公平的硬币）的方法或者由一个随机数字生成器来分配。在初始组足够大的情况下，产生了两个子组将在统计上相等。意思是，组内的个体中存在的任何差异都会互相抵消。在一定程度上，初始组越大，两个子组就显得越相似。

任何可能影响结果的系统性误差（不可控的）都将分布在两组中。因此，这两个子组的所有相关的性质都相同，除了一组将被干预或者治疗，而另一组保持原样这点之外。

## 避免样本选择偏见

设计一个强大经典的设计的关键在于研究者随机分配实验与对照组的成员。像我们没有随机分配治疗与对照组的成员的情况存在着潜在的样本选择偏见。意思就是治疗与对照组的某个相关性质可能并不相同。比如说，我们可能希望进行一项住宅是否安装了防盗报警器的相关研究。

如果安装了防盗报警器的住宅中的犯罪率更低，但是这样的差别不一定是主要或者完全因为报警器的存在。很可能是因为安装了报警器的人们比没有安装的居民更小心谨慎。

换句话说，安装了报警器的人们同时也是安装高质量的锁或者窗门的群体，以及邻里守望项目中更积极的志愿者。

通常，在任何有人或者研究主题志愿或者被选入治疗小组的时候都应该被怀疑。主要研究对象常常志愿参与一个项目因为他们的动机更加强烈或者认为治疗对他们更加有利。只有评估员或者研究者对治疗与控制状况进行随机分配，才能避免样本选择偏见的发生。说到这里，研究者们进行真正的随机分配是十分必要的。众所周知，有些研究者们将他们认为最有合作可能或者最可能成功的对象选进治疗小组，而不是对照组。

## 不够理想的变异

有时候我们不能够随机的分配政策或者项目干预小组与控制组的成员。比如说由政府决定立法的情况下。就性质而言，社会政策是通过管辖区贯彻的，而并非随机的分配给某一个地区的。比如说，如果卑诗省希望采用一套新的关于就医等待时间的绩效标准会发生什么？很显然，我们可以应用前后比较模型，但是我们清楚它的局限性。有没有办法通过使用经典设计的框架来克服这些局限性呢？

### 匹配对比设计

答案是，有些方法不如经典模型理想，但是可能比简单的应用前后对比方法更可行。由于我们无法随机的将管辖区分配到不同的响应标准中，其中一个方法是寻找潜在的克隆。就是说，我们所了解的（或者假设的）管辖区虽然有着不同的标准，但是它们所有的或者大部分的相关方面都相似。就卑诗省来说，我们可能会考虑选择华盛顿州和俄勒冈州，以及阿尔伯塔省做为其比较对象。这里做出的假设是这些管辖区虽然有着不同的响应标准但是它们的所处地理位置以及社会人口方面的特质与卑诗省相似。

如果不能产生不同的结果的话，那么不断重复地做同一件事情毫无意义。

我们将这个方法称为匹配对比的过程。我们尝试寻找到与实验对象尽可能相似的匹配管辖区做出相关对照组。当然，样本选择偏见的问题在这里也可能会出现，因为决定实施某项政策的管辖区一定在某些地方与那些没有实施的区域不同。就如同简单的前后对比方法一样，我们需要比使用经典设计的黄金标准更加谨慎的对待得出的结果。

不管怎样，由匹配对比设计得出了可证明某些实践确实有效的可信证据。或许最好的例子之一是通过白天开启汽车车灯来减少交通事故的早期研究。<sup>1</sup>从另一方面来说，匹配对比研究的同时也提议了某些政策在实施之后并没有能够达到期望的效果。这里的一个例子是对于实施死刑与降低谋杀率之间的关系的研究。

跨司法管辖区而收集到的证据的优势在于尽管死刑可能平息人们的报复心理，但实际发生的谋杀案起数并没有因此而减少。

在我们需要做出决策的时候，证据越有力，做出的决定也越可能是正确的。我们也可能犯错，但是错误的决定也可以帮助我们了解到问题出在了哪里。如果结论不合理的话，那么反复做同一件事情也是毫无意义的。当我们当下的做法很明显无法达到期望的效果的时候，逻辑上来说我们应该尝试其他的方法。终究我们会找到可行的方法。当我们面对相反的证据的时候，更重要的是我们必须愿意改变我们的观点。

我们经常循规蹈矩地进行某种同样的行为模式，哪怕证据已经表明了这样做并不能产生我们期望达到的效果。几个世纪以来，医生们在反面证据已经存在的情况下依然采用放血疗法，仅仅是因为这样做看上去符合常规。却不考虑有许多的病人因此而更加衰弱以及导致死亡的事实。

---

## 本质特质

这里讨论的最重要的一点是如何收集证据—用来生成数据的框架或者设计—是帮助我们决定证据的有效性的一个重要元素。另一个重要因素则是我们之前提到的事物只有通过比较才有意义的概念。

换句话说，凡事都需要一个比较对象来体现出它的意义。一种介入或者行动只有与另一种行动或者无行动（什么都不做）相比较的时候才有意义。那个比较对象叫做反事实。

因为同一件事物无法同时存在在两种不同的状态中，因此我们必须寻找到最合适的比较方法。正如我们已经看到的，克隆十分难得，因此至今我们设计出的最好方法是随机实验，即将被实验者或实验对象随机的分配至一个治疗小组以及一个合适的对比或对照组中。随机分配的过程帮助我们排除了系统性的样本选择偏见的可能性。

有的情况下，无法实现随机分配，因此我们会尝试着创造出尽可能理想的条件。

我们应该一直对通过这些方法得到的证据保持怀疑的态度，但是如果这些方法看上去合理并且产生的结果并没有其他符合逻辑的解释，那么一般来说我们就可以给予证据合理的重视，直到我们找到更有说服力的证据。

哪怕是由设计最优的实验得出的结果，它也有不可信的时候。我们需要考虑的其中一方面是实验设计的合理性；另一方面则是其所产生的影响或者效果的大小。很明显，相对产生细微以及边缘的影响的干涉方法，我们更有理由使用来自能够产生重大影响的干涉方法的证据来做为决定的依据。但是这使我们进入下一个章节的主题—例如政策评估以及成本效益或者成本效率分析等其他方面的考虑。

## 注释及参考文献

1. See, for example, Elvik, R. (1993) “The effects on accidents of compulsory use of daytime running lights for cars in Norway” *Accident Analysis and Prevention*, 25: 383-398.

# 项目评估

## 序

各级政府部门的年预算大多都花费在提供服务上——公共安全服务，消防服务，环境保护服务，社会公益服务，交通运输服务，公共医疗卫生服务，公园管理服务，维护服务等。通过调查政府部分的工作，你将发现大部分都与提供服务相关。因此，政府在不断的确定是否满足了他们所服务的人群的需求。因此，政府部门会定期重新审查各级服务以保证它们具备适当的针对性，同时也会对服务的结构与操作过程的效果与效率进行评估。所做的这一切都是为了确保税款运用的恰当。

评估服务交付的效果与效率并不简单。事情可能在短期内变得复杂。政府极少拥有能够满足服务需求的足够资源是主要问题之一。而且，当政府希望在服务交付方面做出改变时，他们往往面临着各种限制，例如基础设施的短缺，劳动力合约，管辖区涉及的事项，法律方面的需求，以及许多潜在政治压力。

这解释了为什么削减，改动，或者增加服务一直都是一项难题。结论是在政府想做的与能够做的之间有着一道鸿沟。因此，评估政府的服务水平是一个敏感的问题，也不奇怪政府对待评估事项十分的谨慎，特别是没有考虑具体的环境而进行的评估。

当考虑到政府各级提供的许多服务是由外部机构交付的时候，评估服务则变得更加敏感。

---

评估服务交付的效果与效率并不简单。

---

这些服务由商业机构，独立承包商，以及各类非营利的服务机构提供。实际上，这些机构的大多数都不具备资源或者内部的专业知识以适当的评估他们所提供的服务。此外，进行自我评价存在着一个固有问题，因为大多数机构会为了既得利益而尽可能的将自己展现的积极正面。

另一方面，政府没有足够的资金来对他们投资的每个机构提供的服务进行彻底的评估。一般的情况下，拨款安排通常是在很短的时期内完成并且金额有限。简单的说，这样的情况下，政府虽然可能希望进行评估，但是由于资源有限而无法进行。因此，政府常常依赖于某个人或者机构的声誉，并且停留在承包人提供的服务的表面意义上。幸运的是，政府的大部分外界承包都是有限的，同时承包商们明白他们必须保持一定的标准以再次获得资金。

说到此，政府部门在评估许多内部与外部提供的服务项目的效果与价值方面还是有空间的。评估不需要很复杂，昂贵或费力。可以根据政府评估其所提供的服务而需要考虑的敏感度来进行。

考虑到这些限制，首要的事情是认识到所有的政府服务都可以被当成某种项目。它们可能叫做倡议，社会企业，课程，或者简单的服务，但是我们可以把它们都当成自成体系的可被评估的实体看待。全部都应该以完成某个任务的方式交付出某个产品或者服务。此外，那些结果应该以最有效的方式实现。在理想的世界里，我们可以将感兴趣的项目与备选方案比较，并且决定哪一种更出众。从这个角度来说，我们说的是同一种被称为项目评估的技术。了解项目评估的基本方法可帮助你评估某项服务是否有效以及更好的服务纳税人。

本章节将复习在项目评估中需要提问的问题。很明显，有些项目规模庞大并且涉及多个管辖区（比如说，一些美国的计划），因此我们这里无法提供对其评估所需的背景。同样的，有些项目的目的是多样性的，以及评价其结果所采用的方法十分的复杂。不管怎样，我们这里讲到的方法对评估大多数政府部门中的典型项目都适用。

在本章节的末尾我们讨论项目的逻辑模型将为你在评估的过程中提供指引。首先，我们来了解一些在评估之前所需要考虑的基本问题。

## 第一个问题：具体是什么项目？

这里的关键词是“具体”，因为如果你不清楚某个项目的细节，实际上你无法衡量它完整的效果，当然你也无法决定它是否是以有效的方式在运行。此外，你无法确定你是否选择了合适的对比对象，因为你可能在不知不觉地比较苹果和橘子。

话虽如此，很多时候我们发现第一个问题并没有得到解答——人们常常认为一个项目的总体描述信息已经足够多，实际上远远不够。你需要了解该项目的各个组成的足够细节，这样才能避免沟通中的误会。比如说，某个机构可能声明他们为社区提供一项恢复性司法项目。这样的声明并没有问题，但是类似的项目多种多样且区别非常的大，因此当你确定它们实际上做什么之后，你会发现它们之间很难找到相似的两种。

一份项目描述必须详细清晰的说明服务对象期待的结果。通常，当项目各组成部分的说明是清晰的，且具有可测量的

一份项目描述必须详细清晰的展现服务对象期待的结果。

定义的时候，你就可以因此而鉴定这是一份优秀的描述。否则，就无法很好的回答在项目评估中出现的问题。不管怎样，在我们逐渐全面的考虑评估过程的时候，具备一个清晰的项目描述的重要性将更加的明显。

或许拿减肥计划打比方可以很好的说明这点。你需要以使外人能够衡量参与者们预期的付出与得到。众所周知，减肥计划的时间长度各异并且包含各种形式的组成部分（比如，饮食，运动，训练员，以及补充剂），而且一眼看上去，许多部分容易混淆。因此，一份泛泛的计划描述是不够的。

## 第二个问题：该项目达到了预期的效果的吗？

同样的，这是个很少被提及的问题。但是，如果某个项目在建立的初期就将怎样交付的计划纪录在文档中的话，通常就不难回答这个问题。而该项目的接收对象可能会被问到该项目是否真的在各方面都达到了所声称的标准，也可能被问到接收到了该项目中各个组成部分的多少，甚至还可能被问到他们是否参与到了该项目之中。

为了更好的赏析提出的这个重点，你只需要回忆在高中或者大学时期上过的某一门课程。你将想起不是所有的课程都符合课程大纲中的描述，并且有一位老师在教室中并不代表课程内容是以学生们实际上学到了一些知识的方式来授课的。此外，就算按照计划传授了课程内容，也不代表每个登记的学生都参与了。有些学生在课堂上睡觉，有些在做白日梦，以及有些干脆不来上课。通常，学生对同所大学中教同一课程的教授的评价存在着巨大差异。

有些学生表明某个课程提供的比预期的要少；比如，一本从来没有被提及的必备教科书，与课程大纲毫无关系的讲课，或者难以理解授课教授在说些什么。

在一个相对固定的环境中，多年来为特质一致的学生群体，教授了一门大学课程这样的一个简单的项目都存在这样的差异，那么你可以想象在接收对象更加广泛，由多种服务提供部门，在各种环境中提供的同一种项目的交付中会存在多大的差异。此外，没人可以假设项目是如同预期的那样交付的或者接收的。若想知道到底发生了什么，你需要审核包括交付衡量标准的交付索赔。

重申，这里的重点是仅仅因为某个人在某个项目中并不代表他们按照规定参与了，或者他们如预期般的利用了每个组成部分。第二个问题需要你想办法确认参与者接收到了多少部分以及是否如预期般的完成了整个项目。



## 第三个问题：该项目的参与者是否真正的从中收益了？

当你知道某个项目的确切性质以及它实际交付的程度后，你就应该准备好了面对最根本的问题：结果。这里的关键在于决定项目的接收对象的范围（比如，城市，机构，社区，目标团体，以及个体）感受到的某种改变（比如，条件，满意度，态度，技术，能力）

的前后衡量标准。这个改变应该与该项目具体打算实现的目标相关。

在这里，改变前的衡量标准特别重要。它们表明了项目接收对象的起始状态，因此为该项目可能产生的某种影响提供了比较的基础。

---

改变前后的衡量标准对于测量改变的程度非常重要。

---

同时也反应了接收对象的起始的程度不同的事实。

通常，我们综合现有的数据与和之后的衡量标准直接相关的指标来探讨以确定之前的衡量标准。没有这些之前与之后的衡量标准的话，你基本上无法得知该项目是否达到了预期的效果。也就是说，如果你考虑周全的选择前后衡量标准的话，你很可能能够确定在哪些情况下，在什么时候和哪儿，对哪个接收对象所采取的方法的哪个方面最有效。

为了帮助你更全面的看待前后衡量标准一事，让我们来考虑一个社区守卫犯罪预防项目，它只会在邻里之间互相熟悉，平时留意邻居的房屋，及时报告可疑行为，以及采取一系列保护自家房屋的防患措施的前提下，犯罪率才会降低。但是，项目评估的第一部分的主要关注不应该在于犯罪率是否降低。我们首先需要确认我们可以回答第一个关于社区守卫以及它所有的组成部分的问题。也就是说，实施的计划中是否包括了社区会议，印刷资料以及一位组织群众活动与为群众及时提供消息的社区守卫队长？

接下来，正如第二个问题中所指示的，我们需要知道各有多少个社区住户参与了该项目的每个组成部分。也就是说，他们参加会议了吗？他们是否努力的去熟悉自己的邻居？以及，他们是否遵守了住宅安全建议并且按照建议中所说的锁好门窗？

一旦我们确认了居民们曾经被邀请来参与整个社区守卫项目中，而且事实上他们也参与了，我们接下来就需要解决第三个问题来判定社区守卫是否引发了居民新的作为，如果答案是肯定的，他们做到了怎样的程度。

之前衡量标准，在项目的初始时期，可能包括询问目标居民周边的居民数量以及他们与多少邻居交谈过。也可能包括询问居民他们采取了哪些针对住宅安全防患的措施。如果这些问题看上去有些过分的话，那么实际上，在一个关于社区守卫有效性的研究中，说明了该项目中几乎所有的试图去做的事情在附近的未参与社区守卫项目的社区中已经开展开来了。

上文提到的研究并没有包括之前衡量标准，只包括了之后衡量标准。如果这份研究分析包含了完整的前后衡量标准，那么我们就有信心说这份分析显示了该项目并没有改变参与者的任何相关行为。与此同时，参与该项目的市政府一直为一名进行项目协调的全职员工支付薪水—很显然的是在浪费税金。

我们用减肥计划的例子来再一次强调以上的要点，很明显我们希望了解参与者初始与结束时的体重各是多少，这样可以得知他们的减肥效果。最好我们能够同时了解到参与者其他方面的一些背景资料，这样我们可以得知该计划对那一类型的人群最有效。我们还希望能够肯定该项目是为明显肥胖的人群服务的，而且他们的减肥效果不是因为其他的因素。

## 第四个问题：是否实现了终极的利益？

人们通常将“终极利益”这个问题与参与者讨论他们的状况是否因为参与某个项目而有所改变的第三个问题混淆。我们再次回顾社区守卫项目来感受第三个与第四个问题之间的区别。通过回答第三个问题，我们尝试确定参与者们是否因为参与了该计划而在相关行为上有所改变。这可以表明该项目是否达到了预期的效果。接下来的问题与项目的主旨相关，在社区守卫的例子中，即是防止或者减少犯罪案件。

重要的是，在回答这里的“终极利益”的问题之前，我们必须首先回答第三个问题，因为许多因素都可能对终极目标造成影响。就是说，你可能永远都无从得知是该项目影响了终极目标还是其他的因素。比如说，我们可能认定某个项目如预期般有效，但是，最后它并没有对终极目标造成太大的影响。拿社区守卫项目来说，在这项研究中也发现了参与了社区守卫的社区中的犯罪率其实与附近没有参与的类似社区相同。如上面提到的，我们已经通过回答第三个问题而得知这里开展的社区守卫并没有达到预期的效果，因此我们也不指望它能够降低犯罪率。

在评估终极利益这件事上，通过使用一个比较小组或者情况来确定该项目所造成的实际影响这一步骤至关重要。有时候有些项目看上去有效的实现了终极利益，而事实上却是与它无关的其他因素导致的。比如许多声称有效的防止犯罪项目就在本质上忽略了西方社会的犯罪率本身就在不断降低的事实。

在任何情况下，通过回答第三个问题来确认某个项目如同预期的有效实行是一回事，而确认该项目对达到某些终极目标有贡献是另一回事。这需要进行两种分析，包括实施前与实施后的两套衡量标准以及，理想的情况下，两套对照组一各与第三个与第四个问题对应。

我们可以通过思考减肥计划的问题来重申这里的重点。第三个问题需要我们确定哪些参与者成功减了多少重，而第四个问题需要我们确定减重对一些其他例如改善健康状况这样高于一切的规划性目标有多大的实际贡献。后面这项考量可以以好几种方法落实，比如总体上少生病，看医生的次数减少，请的病假日减少，或者分配的药物减少。换句话说，减肥这件事只有对其他事情做出了实际上的贡献才有意义。

## 第五个问题：那又怎样？

每当完成一份评估的时候，我们都应该思考是否存在比该项目更加有效的另一项目存在？尽管我们确定了该项目达到了预期的效果，但是我们依然应该找寻是否存在更好的方法。但是，还有一个重要的部分：你同时也需要思考是否存在能够更有效的达到终极利益的另一项目。因此，你需要将你的结果与其他的项目进行比较。

有很多进行比较的方法，其中，一个好的开始是进行与该项目相关的主题范围内的文献综述。文献是指项目评估的完整报告，同时，你可以轻易得到指明哪些有效与哪些无效的信息。运气好的话，你甚至可能找到一份包含了一系列的相关项目的综合对比分析。

需要注意的是，进行比较的对象要适宜。理想的情况下，在进行评估前你需要首先检索文献资料，以及在这个过程中发现其他人是如何进行类似的评估的。

基于这一点，减肥计划是一个非常好的例子。或许该计划对减肥有帮助，但是有许多其他的计划更加有效的达到了同样的目标。要记住在改善健康状况是终极目标上，或许其他的例如调整简单的饮食或者改变一些生活方式或许会更加有效。不管怎样，文献中有着许多的与减肥和其他类似计划的例子。这里的目标纯粹是确保可以确认被评估的项目是达到终极目标的最好方法之一。

---

## 第六个问题：该项目是否在尽可能有效的运行？

直到现在，我们一直在强调在学术上被描绘为“结果评估”的意义。就是说，我们强调确定某个项目是否如预期般的运行（中期结果），以及是否如预期般的对更广的目标做出了贡献（终极结果）。项目评估中的一个同等重要的部分是“过程评估”。过程评估是一项对运行过程以及与项目相关的系统中的每一步骤进行的评估，它的主要目的在于衡量该项目是否在尽可能有效的运行。因此，它包括了深入全面的观察所使用的资源，对使用的数量、质量以及应用进行评估，并且确定它们是否最好的满足了该项目的需要。有时候，有些项目缺乏足够的人员与资金。

另一些时候，它们在某一方面的资源过剩。还有些时候，它们需要重新分配资源。资源管理出现误差也很常见。

也可能需要修正项目的内容。领导阶层，接收过程，推荐系统，数据系统，科技，问责制度，沟通问题，劳动力问题，以及相关利益方的参加程度，都可能需要被仔细考察。这里的总体目标是确保该项目的需求被满足从而实现项目目标。

进行过程评估至关重要。我们都见过一些在运行上表里不一的项目。人们很容易被某个项目的潜力误导，因为在实施过程中存在着隐藏的弱点。

每一个优秀的项目在过程评估之后都有被改善的潜力。我们需要记住的是当某个项目运行的没有效率的话，它的效果往往也不会太好。

过程评估往往会令项目运行部门感到受威胁。但是，没有必要这样。并非要将项目的每一个方面都放在显微镜下观察。评估可以从最近的任务与系统或者有问题的方面开始着手。

如果时间与资源允许的话，你也可以进一步研究问题出在了哪里。再来，评估可以以“成型评估”的方式进行，这里的目标是为增加效率与效益而提出建议。

过程评估是一项对运行过程以及与项目相关的系统中的每一步骤进行的评估，它的主要目的在于衡量该项目是否在尽可能有效的运行。

再次回到减肥计划的例子，我们可以看到过程评估并非注重在结果上，而是考察该项目的管理与执行是如何进行的。这可以包括考察记录参与者总的参与程度的方法，以及记录参与者享用饮食，运动与其他项目中的因素的方法。这里的目标是为了使该项目更加友好与有效而提出建议。

## 第七个问题：该项目值得花钱吗？

我们提到了我们应该探索是否存在能够使项目达到终极目标的更好的机制。我们也应该探索是否存在同等有效但是成本更低的备选项目。在粗略的级别上，这是一个简单的练习：首先确立该项目的成本，然后看看竞争或者备用选项的成本。在略详细的级别上，这个练习通常需要可观的经验与技巧，特别是当你试着将间接成本，现物出资，乘数效应，以及其他等计算在成本内的时候。无论如何，这些都可以归纳到本书稍后会介绍的成本收益或成本效率分析中。

成本分析不仅仅是比较两个项目的成本多少。它也包括确认该项目是否如预期的节约资源。

项目常常代表着人们期待的较便宜的做事方式。也就是说，它们首先打算代表着一种成本节约。

比如在减肥计划的例子中来进行成本分析，我们可能希望直到，比如说，该项目的成本是否比类似的项目更低。我们也想知道通过任何减重而得到的整体的健康益处是否代表着该项目中的投资成本节约。我们甚至也可以考虑是否存在着可以达到该项目中达到的健康收益的成本效率更高的方法。此外，这里的目标确保该项目值得投资。

## 组织你的评估的一种方法：使用一个逻辑模型

到现在为止，我们已经在一些基础问题的层面讨论了评估方法。同时，不管怎样，这些问题可以做为一个“逻辑模型”的基础——一个在评估过程中帮助引导评估员的框架。逻辑模型的复杂程度不一。你可以通过考虑下面的几点来了解模型中所包含的因素。

- 1. 项目活动:**这里如同第一个问题一样，首先需要明确为了产生项目的每个预期结果而设计的具体活动。因此，你需要考虑所应用的技巧、产品和科技，以及该项目为了产生每个预期结果而使用的运行策略。比如说，如果被评估的项目是一个生活技能培训项目，那么你需要知道所使用的课程有哪些事项：交付的方法；授课时间有多少个小时；引导员或者指导员的任职资格是什么；交付的形式；交付时间表；以及所使用的教材。通常，如果有人没有参加该项目而通过资料就可以了解该项目的话，那么你可以确定手头上的描述简介非常优秀。知情的局外人也应该了解活动是怎样和为什么与预期结果相联系的。
- 2. 输出:**与第二个问题一致，这里的重点是确认项目按计划中的数量与质量交付。在生活技能的例子中，比如说，你会希望确认该项目中使用的格式，使用的材料，所授课程的种类，以及参与者从课堂中吸取到了知识与技能的多少。看待这件事情的另一种方式是，项目活动是审查某个项目的组成部分，输出是指衡量与审查项目是否如同期望般的交付。
- 3. 直接结果:**这里我们关注的是第三个问题，并且希望确认参与者们从该项目中有所得益。基本上，这项练习就是衡量参与者们因参与了该项目而产生的任何变化。在生活技能项目的例子中，这就包括衡量参与者们是否学到了某种新的生活技能、知识、行为与态度。更加精细的衡量方法则包括在一段时间之后，他们保留了多少所学。进一步说，如果收集了参与者的背景信息，我们还可以将那些信息与他们的学习情况联系起来。
- 4. 终极结果:**如同第四个问题中说明的，关键的评估目标是确认该项目产生了预期中的终极利益。重申，直接结果并非实施项目的初始原因——一般来说，初始原因的目标更加广泛。在这里我们需要将参与者和未参与者们进行对比。比如说在生活技能项目中，终极目标可能是提高犯罪者的再就业率，加强家庭稳定，减少药物滥用，或者降低再犯罪率。那么我们需要衡量的是经过一段时间之后，相对没有参与该项目的犯罪者群体，参与了该项目的犯罪者们的就业率是否更高，他们的家庭是否更加稳定，药物滥用现象是否减少了，以及再犯罪率是否降低了。

5. **比较结果**: 这里的任务则如同第五个问题中提出的, 旨在考察是否存在某种更好的选项。在这一点上, 可以拿其他地区实施的类似项目做为比较对象, 或者发表的文献中也记录了一些其他选项的实施结果。不管怎样, 我们应该记录通过与其他选项的结果比较来衡量被评估的项目是最好的方法。如果存在能够更有效的达到目标的其他的选项的话, 那么我们就无法确认被评估的项目达到了终极的利益。拿生活技能的例子来说, 这里需要做的是通过查询文献中记录的类似项目的结果, 用做比较的对象。
6. **活动效率**: 第六个问题提出了该项目是否按照计划有效率的运行? 因此, 我们需要检查所使用的资源与运作流程是否最有效率的达到了目标? 生活技能项目的例子中, 我们会检查该项目的的时间是否可以缩短, 材料与课堂时间是否被充分的利用了, 以及在不影响效果的前提下, 课程人数是否可以增加。
7. **成本收益比较**: 第七个问题指明了通过使用成本收益分析来解决两个问题。首先, 是否存在着某个性价比更高的项目或者途径? 第二, 实施某个项目的成本是多少? 拿生活技能项目的例子来说, 这将包括确定它的成本, 接下来与其他类似项目的成本相比较, 以及与完全不实施该项目的成本相比较。换句话说, 实施该项目的成本是大于或者小于未参与该项目的犯罪者们再犯罪而造成的经济损失?

下面的表格概述了这个过程。

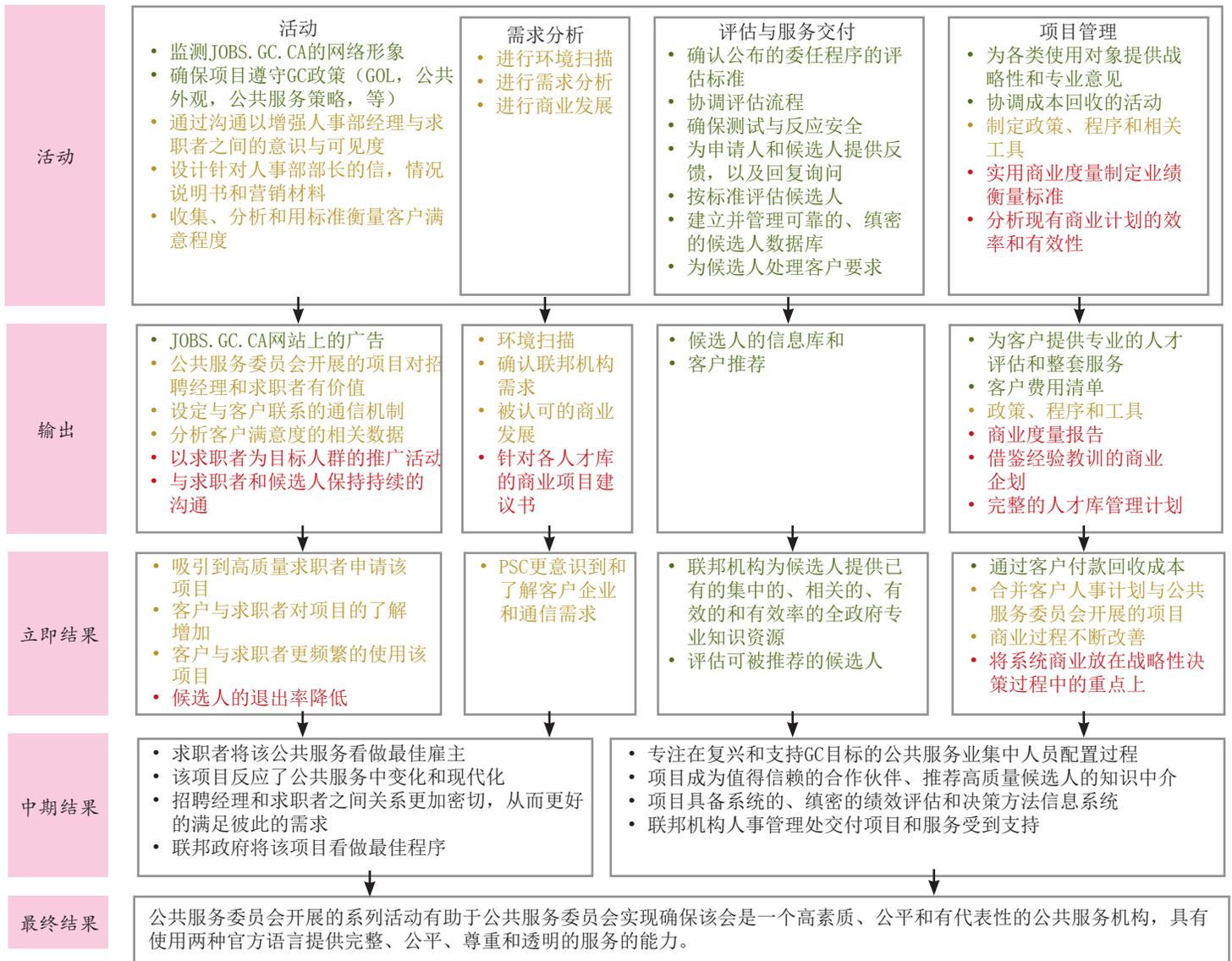


## 例：公共服务委员会逻辑模型

以下的逻辑模型<sup>1</sup>从图表上展现了一个计划的输入，活动，输出与结果。这是加拿大公共服务委员会基于该会开展的项目目标来分析与鉴定其的优势与劣势而特别制作的。该会开展的人才匹配项目

是指通过新的和创新的方案来实现该会的职责以及辅助其他的公共服务项目，比如人员配置，评估与其他提前测试候选人等系列事项，比如专上公职招聘。<sup>2</sup>

这些列于活动信息中。因为公共服务委员会开展的系列活动较新，同时不断改善，操作小组在学习的同时还可以尝试着缩短计划与实际操作之间的差距。



公共服务委员会，企业管理分支，评估部门

在逻辑模型中，不同的颜色表示了初步差距分析：目前为止已完成的部分，部分完成的部分以及未完成的部分。

从上一页的练习中，我们识别出了以下的差距：<sup>3</sup>

## 沟通与拓展

其中发现的最大差距存在于求职者与公共服务委员会开展的活动的关系当中。问卷调查的结果表明求职者对活动的过程了解甚少。求职者普遍认为委员会未能做到及时告知他们在职位匹配中的状态（百分之66）。

## 需求分析

环境扫描，需求分析与商业开发机会的识别必须从某些地区开始，并且在其他地区规范化。每个区域的商业建议已经系统的发展起来了。此时这些活动正在以其独特的方式进行中。

## 评估与服务交付

因为服务交付是公共服务委员会开展的人才匹配项目的主旨，实施队伍的主要精力与资源都集中在这个方面。但是，在开展活动方面依然存在着一些不足，主要在于为求职者与候选人提供反馈信息。而这些活动应该是该项目的最强的领域。

## 管理计划

商业制度和其他数据源，例如信息管理过程是评估与衡量业绩的主要工具。到2009年11月为止，业绩衡量数据的范围从有限到不足，并且缺少国家级标准程序。这种状况使业绩评估更加复杂。

我们已经列出的框架并非是唯一的。通过文献搜索，你可以找到许多其他的种类。关键在于使用系统性的框架来检查某个项目的设计。同时，也存在着一份项目评估不包括此处列出的每一个组成部分的其他的原因。我们这里展示出的只作为指引。

## 总结

如本章节的开始所说，评估政府投资的服务的过程十分复杂。但是若将它们当作项目来看的话会有所帮助。事实上，我们可以说大多数政府服务可以被当作项目来看，以及被当作项目来评估。此外，我们看到评估员是带着一系列的基本问题来开始评估的过程，但是并非需要去解决这里列出的每一个问题。但是，如果目标是评估一个服务是否如同预期般交付，也就是说，该服务是否有效，以及它的性价比较高的话，那么就 need 考虑到每一个问题。

回答我们在这里列出的问题仅仅是开始的步骤。实际工作在于建立可以使你获取基于证据的可信的研究设计。

因此，这包括考虑我们在本书中概括的其他的其他问题。正如有经验的研究者会告诉你的那样，基本上很少人能够如同他们希望的那么全面的做出评估。通常会被缺乏数据，无法得到项目的具体信息，受限于时间与预算，以及其他不可控因素影响。不过，我们的目标是在条件允许的情况下尽可能的严谨，并且在描述与展示结果的时候仔细的考虑并提及所存在的限制与注意事项。

---

## 注释及参考文献

1. Public Service Commission of Canada  
<http://www.psc-cfp.gc.ca/abt-aps/inev-evin/2010/pools-bassins/img/figure4-eng.jpg>
2. <http://www.psc-cfp.gc.ca/abt-aps/inev-evin/2010/pools-bassins/index-eng.htm#ex-sum>
3. <http://www.psc-cfp.gc.ca/abt-aps/inev-evin/2010/pools-bassins/index-eng.htm#appC>

# 成本分析

## 基本概念

成本分析有两种变异。第一种考虑的是做某件事情的费用。比如说，决定购买一辆汽车需要考虑的不仅仅是购买汽车的花费，还包括修理之类的维护费用，加油之类的消耗品购买，以及保险之类的支持成本。在有些情况下，还会出现其他的成本，比如盖建停车库或者租用停车位的花费。如果只看与所有权相关的成本，我们还需要考虑到汽车的贬值（希望在几年内出售汽车的时候我们还能够收回些成本）以及购买汽车的花费上的利息。

另一种分析成本的方法是成本收益或者成本效率分析。在这个方法中，我们将做出某个决定的成本与预期利益的价值联系起来。比如说，某个部门可能打算投资在深入培训上。这里出现的问题就是：该投资的收益是多少？如果培训的项目是在一个完全不使用互联网的社区内阻止网络犯罪的话，那么收益将是零。事实上，这是一个只有成本的情况。

另一方面，如果培训的内容与员工的健康和安全相关，那么收益可能显现这些方面上，如在更低的事与受伤率，减少病假，降低保险费，增加员工工作效率以及提升员工士气等。我们可以将这些收益的相对价值与提供培训的花费相比较来估计出投资回报率。

---

成本研究可以帮助我们辨别出与某个决定相关的总成本，以及因此得到的回报或者利益。

---

经济学中的一个基本思想是机会成本的概念。假设你预算有限，决定做某件事情的话必然排除了做另一件事情的可能。比如说，已知某个部门的资本预算，他们的决定是购买一辆轻型货车。做了这个决定之后，其他的选择一越野车，轿车，摩托车等等一都被排除了。也就是说，选择其他车型的机会都没有了。不仅仅是失去了选择其他车型的实际机会，我们同时还放弃了与那些选择相关的利益。

成本研究可以帮助我们辨别出与某个决定相关的总成本，以及因此得到的回报或者利益。此外，我们还可以研究第二或者第三个选择的相关成本与收益，并且将它们与首选比较。有时候通过这项练习，我们会发现之前认为不够好的选项其实比首选更好。

金钱形式的成本并不是，也不应该是，我们做出决定的时候唯一需要考虑的因素。政治与其他社会考虑因素也影响着我们的决定。不过，金钱形式的成本非常重要以及通常容易计算。大部分的商品与服务都有着一定的金钱成本。但是同时政治与社会成本常常与经济上的决定有着紧密关联。正如可以正式的衡量金钱成本一样，也可以通过通用的成本框架来衡量某个决定引起的非金钱形式的影响。它们的区别在于在后者的情况下，成本与收益往往更难量化。不管怎样，决策者可以并且使用定性的数据来衡量这类决策的影响。

无论我们进行的是不是一个直接的成本分析，成本效率或成本收益分析，都需要考虑以下五个基本步骤。

## 需要考虑的步骤

1. 从部门的操作或者战略性计划中辨别出与问题或者分析有关的部分；
2. 设定我们期望达到的目标；
3. 辨别出存在的选项或者选择；
4. 进行所选选择或者考虑的选择的财政（成本收益或成本效率）分析；
5. 准备一份总结的会计报表。

这些步骤看上去是在重述我们之前的描述。事实上也是如此。不过，我们需要将有效的基于证据的决策看成一个全面框架的一部分，这个框架的起始是思考我们在做什么以及为什么而做，可选的选择有哪些，以及我们使用了哪些证据来帮助我们做出了这个决定。如果不知道我们在做什么以及为什么而做，那么我们不可能辨别出适当的信息。这样的话，我们可能会收集许多数据但是却没有足够的证据。

## 成本分析

直接的成本研究包括估计某个设备或者服务的总生命周期成本。生命周期的意思是我们使用这个商品或者服务的时间段。比如说，一辆汽车的实际的平均预期生命大约是12年，之后可能就会被废弃。某个人或者机构可能会决定购买汽车，使用五年后将它出售。在这个情况下，该汽车的生命周期是五年。

这里的关键是进行准确的成本分析以保证我们包括了所有的适当的成本。一般来说，设备或者资本货物的成本基本上都可以包括在以下的几个种类内：

- 贬值，
- 资本利息，
- 维护费用（购买与维修），
- 许可或管理费用，以及
- 操作费用。

虽然分析师时常在分析中不考虑操作成本，但是其实应该考虑，哪怕最终的决策不受它们的影响。如果该设备是用做替换某个部分的，那么相关的操作成本是从被替代的设备那里延续下来的。不管怎样，假设某个市级政府决定购买一支运盐卡车和一辆平路机，增加的车辆可能需要额外的操作与维护人员，我们需要将这些费用考虑在分析之内。

可能会有人对我们的列表中包括了贬值而非起始的资金成本感到不解。这里的假设是那个设备在生命周期到期后将被出售。因此，这里的资金成本部分是购买价格减去出售价格，也就是我们所说的贬值。

不同的设备的贬值速度不同，但是一般来说是每年百分之20-30。我们以剩余价值的方式计算贬值，比如一个价值\$100,000的设备以每年百分之20的速度贬值，一年后它的价值是\$80,000。第二年的贬值是\$80,000x.2，或者\$16,000。因此，两年后的贬值则是\$20,000+\$16,000，或者\$36,000，以及该设备的剩余价值则是\$100,000-\$36,000，或者\$64,000。

在成本分析中常常被忽略的是购买资金的利息。文献中常常将利率称为贴现率。当通过借贷来购买的时候，需要考虑利率的理由很明显，因为银行或者金融机构会将该部分的费用包括进去。不管怎样，无论该设备是否全额付清，我们都需要考虑“租用”资本的费用。因为如果没有购买该设备的话，我们可以将这笔资金进行投资或者做为它用。这是另一种形式的机会成本。

## 获取可靠的成本预估

无论是聘请某人还是购买某个设备的总成本，进行优秀的成本研究的关键在于确保我们包含了所有的项目，并且获取这些项目的最准确和可靠的成本预估。各机构所处的环境不同，因此通常最好的信息来自于经验。回顾你所在机构的财物记录可以给你带来启发。因为它们反应了真实的经验，通过它们可以很容易的看到意外支出最常出现在哪些方面。

这里的关键是进行准确的成本分析以保证我们包括了所有的适当的成本。

不要将意外支出看成独特的或者偶然一次的事件而忽略它们，应将它们放在你分析表中的单独项目之中。

当无法借鉴机构的过去的经验的时候，我们还可以从其他的地方得到信息。通常，供应商可以提供与竞争对手的产品的成本比较。不过，要注意的是这些分析常常是片面的或者已经排除了某些“不便”的单独项目。要确保你手头的对比信息是全面的。当你缺少某方面的信息的时候，可以要求对方补充信息。

许多独立机构也进行各种项目的成本分析，特别是专业或者贸易协会。同样的，非政府机构与其他公共机构也常常公开他们的预算和成本分析研究。大部分都可以在网络上或者当地图书馆中查询的到。有时候只需简单地打一通电话就可以得到宝贵的数据。

下一页的表格展示了一个直接成本分析的例子。这里，我们研究的是一辆标准的，全尺寸轻型卡车的购买与五年内操作成本。但这个例子中没有考虑驾驶员的工资。

直接成本分析是通过预估生命周期成本以决定购买某个商品的负担能力。在比较不同商品的时候非常有用。比如说，某品牌的轻型卡车的资金成本比其他品牌的更高，但是它的维护成本却更低。那么问题就成为了：哪个选择更好？

## 购买崭新轻型私用卡车的成本

三年内购买与操作一辆轻型卡车的成本：

项目	成本(\$)
购买价格	23,500
售出价格	9,500
贬值	18,577
融资	3,387
燃油	10,079
保险	3,471
税与许可费用	3,650
维护	2,069
修理	821
<b>总成本</b>	<b>42,054</b>
<b>每千米成本</b>	<b>0.47</b>

### 假设：

- 每年驾驶20,000千米
- 首付\$2,750.30 融资利率百分之2.7
- 油费每升\$1.25
- 每100千米耗油10.46升

类似的分析可以用于决定是购买一辆崭新的汽车还是二手更加划算，或者是租用还是全额买断。很明显，在这些不同的情况下，我们必须做出有关预期生命周期，操作费用与贬值的假设。在这里我们重申，一般来说，成本分析中使用的金额都是预计的。重点是，通过一些调查和经验可以将误差减少到最少。不管怎样，许多预期的开支项目，比如出售价格和实际操作成本，都是基于不可控的假设。

我们已经讨论了财物资本分析，同时我们也可以对人员成本的做出类似分析。基本的概念都是一样的。通常，我们在决定是否聘请某人的时候注意力在他的薪酬上，但是从属成本也可能很大。当退休金，税，保险，福利与其他薪酬有关的事项也考虑进来的时候，一般相当于总薪酬的额外百分之15-30。这是培训与维持那个人之外的成本。维持成本包括那个人的工作区域和工作所需的设备和物资。在之前的例子中，我们指出了通常需要为设备配一名操作员。因此，同样的，人员也需要使用设备来工作。

## 成本收益的有关注释

在之前的分析中，我们的注意力集中在了在一辆汽车的生命周期之内持有与操作它的总成本。在决策过程中，了解某件事物的总成本十分重要。可是，通常只了解总成本还不够。大多数设备或者其他项目都可以产生某些形式的输出或者产品。拿汽车来说，输出是运输。在这里，了解行驶每千米的费用比总成本更加有用。

在上一页中提到的例子中，预计的汽车行驶每千米成本大约是\$0.47。我们将每单位输出的价格或者成本称为它的成本效率。尽管成本效率与总成本很明显的有联系，在决策的过程中我们依然应该将它看作一个独立的因素。通常，总成本的差别可能不会造成什么影响。重要的是每单位成本。单位成本与总成本不同的其中一个原因是总成本包含两个组成部分：固定或沉没成本，以及可变成本。固定成本与一次性购买费用之类的相关。可变成本通常与操作和维护费用相关。一个设备的固定成本可能较高，但是如果它比一个价格低些的更加有效率的话，那么通常它的单位成本会更低。

成本效率分析中的一个重点因素是能够辨别出合适的衡量输出的标准并且能够以正确的方法去衡量它们。

人员成本也是一样的。如果高薪酬的员工更加有效率，更少请假并且工作质量更高的话，那么与将工作外包给低廉的地区而省下些费用相比较的话，高薪酬的员工其实更加有价值。最重要的是生产了多少件产品，服务了多少个人，以及输出或者服务的质量。

成本效率分析中的一个重点因素是能够辨别出合适的衡量输出的标准并且能够以正确的方法去衡量它们。再重申一次，机构的运作或者战略规划在这里将起到重要的作用。这些文件中应该概述了机构的目标与运作目的。一个机构的效率与它要交付的产品或者服务有着直接或者间接的关系。

## 成本收益分析

成本收益分析通常是简单的成本效率分析的延伸。一个主要的区别在于成本收益分析考虑的投资回报的范围更加广泛。大部分成本收益分析包括不那么容易被量化的效果（收益）或者有着更广泛社会影响的输出。

成本收益分析是以福利经济为依据的。它和经济学中大部分的分支不同，因为它不仅仅重视顾客和商行的决定，而且关注公共决策给整个社区带来的经济利益方面的影响。因此，成本收益分析常常专注于例如生活质量或者环境质量这样的问题。进行成本收益分析中基本挑战之一是怎样衡量利益从而使它们在整个问题中都具有可比性。在商品中，苹果不是电压器。但是，两者可以同时出现在同一个市场中，并且可以给两者都标价。货币是一个常见的交换单位，它使种出苹果的人可以购买电压器，哪怕制造电压器的人并不想用电压器换一个苹果。

许多公共商品和服务中出现的难题是现实中不存在为这些事项建立起金钱价值的开放市场。此外，因为意识形态方面的原因，许多人反对为公共商品赋予金钱价值。一个经常出现的说法是“环境是无价的”或者“人的生命是无价的。”事实上，两者都可以被赋予金钱价值。问题是不存在为这两者定价的独立的或者漠不关心的市场存在。不管怎样，相对于直接成本分析，这是成本收益分析的一个本质弱点。

## 显示性与陈述性偏好

是否可以用金钱衡量人的生命这样的哲学问题尽管没有答案，但福利经济学家仍有两个广泛的处理工具。他们将其中一种称为显示性偏好方法。显示性偏好与人们在面对某种定性的现象时的真实反应有关，例如将某处地产与类似的但是有个消防栓在前园的地产相比。孩子教育在父母眼中的价值可能影响他们愿意在孩子学费上的花费是总收入的百分之多少。

经济学家的第二个工具叫做陈述性偏好。陈述性偏好就是：某人愿意告诉你的他们愿意为某个事物花费多少钱？

我们可能会评价环境在人们心中的价值，比如说，通过增加多少税收使他们愿意支持洁净空气或者保护自然的倡议。通常，陈述性偏好是通过问卷等类似形式确定的。

尽管显示性和陈述性偏好各有价值，但是两者都有局限性。通过使用个人的人生保险金额来衡量他有多么爱惜自己的生命看上去是一个不错的显示性偏好。但是，他们购买的保额受到他们能支付的起的保险金的限制。而且，一个很爱惜生命的人可能不希望看到他的亲戚从他的死亡中收益从而选择不购买保险。在上一个世纪中，社会学家大量的研究了各个方面的陈述性偏好。他们最终的结论上人们所做的和所说的差别相当的大。

然而，成本收益分析是我们可以用来衡量各种政策和项目的广泛影响的几个方法之一。它可以帮助我们明确问题，辨别出组成的部分，以及与问题相关的证据。这个方法被大多数的公共部门接受并且在许多政府部门是强制要求使用的。比如说，加拿大财政部委员会的使命是任何省政府实施的监管框架必须建立在成本收益分析之上。

它的目的是“部门和机构估算监管或者非监管的选项以使整个社会的利益最大化。因此，所有的监管部门和机构必须展示出他们所推荐的选项可使加拿大公民、商业与政府的经济、环境以及社会福利比任何其他类型的监管或者非监管项目收获最大的利益。”<sup>1</sup>

总的来说，我们可以在多方面使用成本收益分析。比如说，用作：

- 决定是否应该实行某个提议的项目或者计划；
- 决定是否应该继续某个已有的项目或者计划；或者，
- 在备选项目或者计划中做出选择。

我们可以在多方面使用成本收益分析。比如说，用作：

- 决定是否应该实行某个提议的项目或者计划；
- 决定是否应该继续某个已有的项目或者计划；或者，
- 在备选项目或者计划中做出选择。

## 成本利益分析的组成部分

在建立和执行一个成本利益分析的时候，需要遵从下列几个步骤。

这些步骤包括：

### 1. 辨别问题

这是与你的操作或者战略计划有关的问题陈述。

### 2. 辨别限制因素

这里讨论的是行政方面的规定以及你可能面对的其他挑战。这些包括财政限制，管理或者人员方面的挑战，环境以及其他规定，以及任何你需要解决的其他障碍。

### 3. 列出备用选项

每一个提议都有备用选项，包括什么都不做或者维持现状。比如说，如果问题是是否关闭某个办公厅，那么备用选项就包括与另一个部门合并，与其他服务部共享空间，或者扩展业务。

### 4. 列出益处

列出的备用选项的投资回报是什么？是否有金钱回报还是提升生产效率？问题也可能并非希望增加总收入，而是减少或者避免花费。是否有健康，安全或者环境收益？这里的问题可能与整体的生活质量相关。在设备，时间或者人员方面是否节约了开支？

### 5. 如何量化成本和开支？

很明显，市场或者金钱价值是衡量商品和服务的最简单的方法。我们已经列出了提供市场价格的难处。但是，仍可能找到某个成本或者收益的代理价格。社会科学家已经建立了预测个人生命价值的方法。<sup>2</sup>与社区噪音水平或者高交通流量有关成本，可以通过比较喧闹的与安静的社区的房价来预计，或者比较交通流量高与低的房价。

通常，我们可以通过检阅适当的文献来找到方法来衡量难以用金钱计量的问题的价值。我们已经讨论了进行更集中的网络搜索的方法。借用当地专院和大学的经济学家和其他社会科学家的专业知识也可以帮助到你。

一旦我们完成了这些步骤，我们可以准备一份总结这些要素的报告并且展示相关的成本与收益。

## 净现值

常言道，“一鸟在手胜于二手在林。”金钱也是如此。我们收取借贷利息的原因之一是给借方资金会导致贷方面临支付机会成本。金钱无法用在别的地方。为了弥补贷方的机会成本，借方必须支付利息。比如说，你购买了五年兑换的定期性存款单，五年后你得到的比你投资的多一些。一张利息为百分之三的\$1,000的单一的价值是 $\$1,000 \times 1.03 \times 1.03 \times 1.03 \times 1.03 \times 1.03 = \$1,000 \times 1.035 = \$1,159$ 。

我们也可以考虑到相反面。五年后你收到的\$2,000资助在今天的价值是多少？换句话说，你愿意为现在就拥有这份现金支付多少费用？

这是逆转抵押贷款背后的原理。如果你允许银行或者金融机构在几年后按市场价格出售你的房屋且保留所有收益的话，那么现在他们可以支付给你房屋价值的一小部分。这与上一个例子相反。在这样的情况下，我们将利率称为贴现率。在贴现率为百分之三的情况下，未来的\$2,000资助今天的价值是： $\$2,000 \times 1/1.035 = \$2,000 \times .863 = \$1,725$ 。

我们将一个未来的金额现金的价值称之为净现值或者NPV。净现值是未来金额的相反价值。因为项目和资本货物有预期生命周期，通常可以将成本标准化成今天的价值，也就是净现值。另一种看待净现值的方法是将它们等同于固定价值，恰恰与我们试图控制通货膨胀的时候市场价格相反。

在这些例子中，我们已经讨论了经济学家称作的私人时间偏好率，因为重点在于个人。在公共环境下，选择将公共资金投入某一项项目中常常就排除了其他可利于大众的项目的实施的可能性。在公共或者福利环境下，经济学家常常将这个退后的价值称之为社会机会成本。尽管术语不同，但是根本的原理是相似的。

## 收益—成本比例

对于持续了一段时间的项目，我们需要分期计算成本和收益。有时候，成本的持续时间不同于收益的预计生命周期。吉萨的大金字塔建立于公元前2550年左右而且大约在那时支付了成本。埃及旅游业从那时期就因此获得利益。所以，为了能够做出对比，分析师计算出了成本和福利的净现值。

我们将收益成本之间的比值称为收益成本比值或者BCR。假设某个项目的收益净现值是\$13.5百万，成本净现值是\$10百万，那么这里的收益成本比值就是：

$$\text{收益成本比值} = \frac{\text{收益净现值}}{\text{成本净现值}} = \frac{13.5}{10.0} = 1.35$$

理想的情况下，比值应该大于一。如果小于一的话则说明成本高于收益，以及其他因素不变，那么应该放弃该选择。通常我们选择有着最高收益成本比例的选项。如果某个项目的比例较小，却也被选择的话，那么我们需要将这样做的理由注明在账簿的收益那边。

## 例：加拿大卡尔加里格伦莫尔水库引水

于2013年六月，卡尔加里市邻近柏江的区域遭受了重大洪水。<sup>3</sup>这是无法预见的事件，因为柏江的水流超过了百年一次的极限。事实上，自1930年代后，就没有出现过类似的水流量。总之，因洪水引起的损失预计在\$445百万左右。另外，还拨了额外的\$55百万准备紧急反应物品。<sup>4</sup>

在这起事件之后，市政府聘请了一间咨询公司来估计在文化大街附近建造一个格伦莫尔水库引水道的费用。这是一个寻常土木工程项目，而且最初它看上去也没有什么挑战。通过这个例子，我们可以理解这类项目中的各组成部分。咨询公司提供了以下的预计成本。<sup>5</sup>

主要的成本（以百万元计）与施工引水道相关。

### 水流

资本成本组成部分	500立方米/秒	700立方米/秒
调动	32.7	32.3
入口	55.5	63.5
管道	132.6	146.6
出口	68.2	71.7
其他	0.9	0.9
小计	289.9	315
意外开支	72.5	78.8
合计	362.4	393.8

在这里，工程师们根据预计最大的水流量每秒500立方米和每秒700立方米两种情况提供了不同的场景。这里要做几点说明。首先，成本是预计的（根据2014年中的市场价格）而且一旦合同对外招标以及计算出实际人工和材料费用之后，实际成本很可能改变。另外，可能因为出现不可预见的挑战或者市府改变了设计，而产生在小计下面的“意外开支”。

与这类合同的常见的实践一致，意外开支大约在预计的资本总成本的百分之25左右。

成本预计中的另一项省略的是可能招致的商品和服务（增值）税。

预计资本成本不仅仅是工程本身的开支。绘测，工程，公用通道以及其他成本都加在总数之上。包括了这些成本后，咨询工程师提供了以下的预计的该项目资本总成本。

### 水流

资本成本组成部分	500立方米/秒	700立方米/秒
资本成本（施工）	362.4	393.8
环境保护	5.4	5.9
专业服务	90.6	98.4
公共通道	0.1	0.1
合计	458.5	498.2

在后面的表格中，我们看到施工成本之后的下一个最大成本是“专业服务”。这些包括施工管理费，设计费，许可证费和大型项目中的其他标准开支。咨询师使用了行业中的常见的实践，专业服务一般是资本成本的百分之25。公共通道进入权是在施工阶段必要的地役权。这里增值税和引水道在预计生命周期内所需的维护费用都没有被计算在内。

该引水道的净值大约在五亿元左右，与2013年洪水造成的估算损失差不多。

这里隐现的问题是是否值得进行这项投资？受到影响的居民和企业主很可能会同意。而城市内受影响较小的地区可能抱有不同的意见。很明显，这个问题将引起有关在不远的将来此类事件再次发生的机率是多少，以及可接受的政策选项有哪些？<sup>6</sup>

## 例：新西兰公共交通系统购买

另一个相关的例子是市府面临着做出为市交通系统购买哪一种公共汽车的决定。这里包括了价格和环境影响等多方面因素。

纳恩，瓦吉斯和阿德力考虑了作为新西兰未来公共交通系统基础的几种选项。<sup>7</sup>他们考虑的标准车型是柴油机系。柴油引擎排放量大而且噪音高。其他可选择的机系包括柴油 / 电力混合车型和电力车。纳恩和他的同事们注意到，表面上看起来，柴油公交车最

便宜。一辆柴油公交车的价格在\$300,000新币至\$450,000新币之间。

柴油 / 电力混合车型的价格大约在\$600,000新币左右，以及全电力车的价格大约在\$900,000新币至\$1,000,000新币之间。从资本成本的角度来看，传统的柴油车系是最好的选择。在研究了一些“整个生命周期”的预测成本之后，成本差别开始看上去很不一样了。以下的表格总结了小组的调查结果。

系统场景	购买公车 (包括固定成本)	燃油	维护和更新公车	司机	合计
更好的柴油车系	\$165.9百万	\$184.7百万	\$277.7百万	\$385.2百万	\$1013.5百万
混合型车系	\$220.1百万	\$159.9百万	\$276.3百万	\$385.2百万	\$1041.5百万
先柴油再电力	\$256.1百万	\$159.1百万	\$286.2百万	\$385.2百万	\$1085.5百万

在平均预期12年半的交通系统生命周期内，不同系统的整体成本之间的差别并不大。因此，决策者可能应该侧重一些类似排放量或者噪音，或者科技的可靠性等辅助因素上。

正如笔者所提及的，混合车系和全电力车系的公共汽车的排放量较小，并且普遍更安静。但是，这类科技还未经过长期的多次测试以及性能有时出问题。同时，现在可以支持此类新科技的基础设施并不多，这也是一个现实的问题。

## 总结

成本分析是基于证据的决策过程中生成数据的一种常用方法。基本上，成本分析为我们做三件事。首先，在正确完成成本分析之后，它们将我们希望评估的结果与操作和战略计划中的目标联系起来。本质上它们帮助我们专注于辨别该活动是否符合机构的使命。

其次，成本分析帮助我们专注于许多组成实际成本的单项开支。“粗略的”或者传统的成本分析常常忽略许多与活动相关的辅助成本。比如说，成本分析通常忽略需支付的利息或者增加人员的相关费用。

通过专注而详细的分析，我们更可能确保将这些事项包括在内。而且，通过将成本分析的结果展示给同事和其他人，也可为你提供宝贵的意见反馈。

第三，成本分析提供了一个透明并且公平的帮助我们决策的机械方法。

评估是相对可观且侧重主题的。因为假设的成本价值与单项事项对应，因此仔细的检查起来非常的方便。这个过程透明化为我们提供了一个更加合乎情理

的决定：可以被一个独立观察者替换的决定。而且，与单纯基于价值的决策不同，基于证据的决策迫使批评家们生成替换价值或者分析来做为批评所展示的分析的根据。

哪怕有人能够提供其他的证据，那份证据将为得到一份更加准确的状况评估做出贡献，因此，总体来说还是从中获益。最后，被提议的决策将是着有更好的根据。

## 注释及参考文献

1. Treasury Board of Canada (2007) Canadian Cost-Benefit Analysis Guide: Regulatory Proposals. Ottawa: Government of Canada. Catalogue No. BT58-5/2007.  
<http://www.tbs-sct.gc.ca/rtrap-parfa/analys/analys-eng.pdf>
2. Robinson, L.A., (2007) "How US government agencies value mortality risk reductions." Review of Environmental Economics and Policy, 1: 238-299.
3. For a report outlining the extent of the disaster, see Expert Management Panel on River Flood Mitigation (2014) Calgary's Flood Resilient Future. Available at: <http://www.calgary.ca/UEP/Water/Documents/Water-Documents/Flood-Panel-Documents/Expert-Management-Panel-Report-to-Council.PDF>
4. The report of the Expert Panel estimated that the total cost of flooding in Province of Alberta that year was in the range of \$5-6 billion. Only a fraction of that was covered by insurance.
5. See [http://www.calgary.ca/UEP/Water/Documents/Water-Documents/Flood-Panel-Documents/Appendix\\_G\\_Cost.pdf](http://www.calgary.ca/UEP/Water/Documents/Water-Documents/Flood-Panel-Documents/Appendix_G_Cost.pdf) ; more information on the project is available at the City of Calgary website at: <http://www.calgary.ca/UEP/Water/Pages/Flooding-and-sewer-back-ups/Flood-Mitigation-Panel/Flood-panel.aspx> [as at August 6, 2015].
6. For a nice summary of technical articles relating to disasters, see: Shreve, C.M. and I. Kelman (2014) "Does mitigation save? Reviewing cost-benefit analyses of disaster risk reduction." International Journal of Disaster Risk Reduction, 10: 213-135.
7. Nunns, P., J. Varghese and S. Adli (2015) "Better bus fleets for New Zealand: Evaluating costs and trade-offs." Presented at the IPENZ Transportation Group Conference, Christchurch, New Zealand, March 22-14. Available at: <http://conf.hardingconsultants.co.nz/workspace/uploads/paper-nunns-peter-better-54f39398eebe2.pdf>

# 做出决策

## 利用证据

我们在私人和专业生活中不断的做出决策。大部分时候，这些决定基于我们在培训过程中所学的知识，传统的智慧，或者按照传统实践。通常，质问传统实践只能重复以前的发现。但是，许多情况下，传统实践和常识并不管用。我们可能能够得到期望的结果，或者通过这样的方法得到的结果并没有效率。出于某些原因，人们不喜欢改变。我们是保守的物种。我们习惯于重复行事，尽管有时候结果并不理想。正如匿名戒酒手册上所说的，“重复行事却期望得到不同的结果是不理智的。”

历史上，我们可以原谅决策者们追求老朽的仪式。毕竟，正如雨师们所认为的那样，只要你常常跳舞，迟早会降雨。然而现代的天气预报已经足够的准确，雨师也已经不再是一个正当的职业。



---

基于证据的决策方法让决策过程变得透明公开—它不再是封闭而神秘的过程，而是成为了可以使观察者跟随着逻辑思路和证据的过程。

---

因为气象学已经积累了足够的系统性知识而能够相当准确的预测当地温度，降雨和其他气象。气象学家通过科学研究与其他形式的系统性学习从而达到了现有的程度。

依赖系统性的学习和收集数据，也就是科学方法，在其他的学科中造成了矛盾的破坏。这是不幸的，因为今天我们有大量的试验性证据可以帮助我们更好的决策。此外，当缺乏现有的分析的时候，进行一份可以帮助我们自己的决策的局部分析也不是难事。这并不是说人人都需要成为科学家—远非如此。我们需要做的是使用试验性结论来建立起可靠的证据。

基于证据的决策方法通常可以让你做出更好的决定。基于证据的决策的优势在于它的过程是透明公开的。局外人都可以熟悉决策的基础。它不再是封闭而神秘的过程，而是成为了可以使观察者跟随着逻辑思路和证据的过程。

基于证据的决策过程通过使用关于政府政策和服务的结果的最佳可用的研究与信息来实行方针以及评估机构、部门和人员。

我们并不是说你一直能够寻找到解决问题的最优方案，而是基于证据的决策方法帮助我们识别出不可行的选择和做法。在这些情况下，尝试新的方法对你来说并没有损失。大多数的情况下，查阅已有的证据或者你自己收集的数据可以为你指引一条更可能有收获的方向。

每个人都从证据中做出推论。推论是人类基本的技能。分析是一种技能，就如同绘画或者驾驶汽车一样。它可以被教授，可以被学习，并且可以通过练习而提高。不管怎样，和许多其他的技能一样，例如空手道，需要亲自练习和应用。这本手册，配对的练习册以及相关的案例研究将为你提供这样的机会。

总之，我们怎样将本书中的课程集合起来来规划一个优秀的基于证据的决策方法？基本上，可以归类于四个主要步骤。

---

离开正确的问题，再多的数据也无法为你提供答案。

---

## 辨别与构思问题

本书的前三个章节主要专注于辨别出恰当的问题。离开正确的问题，再多的数据也无法为你提供答案。我们重复强调了好的问题离不开合适的框架。理想的情况下，你需要从你所在组织的规划或者战略规划中提取这些问题。这可以帮助你专注在你的单位的主要目的和目标上。许多机构失败的主要原因在于失去了对他们的使命的洞见。他们尝试着对所有的人提供所有的服务。这是不可能达到的目标。

如果你缺乏一份组织或者战略规划的话，接下来能做的最好的事情就是将问题细化。问几个基本的问题：

- 我们提议这么做的理由是什么？
- 最可能的结果都有哪些？
- 这种作为和组织的使命有怎样的联系？
- 这种作为将为我的机构或者我们服务的人群带来什么利益？
- 是否存在成本效率或者成本收益更高的其他选择？
- 这种作为是否会造成长期或者短期的后果？
- 如果我们进行这种作为的话，我可能需要哪些其他的资源？

如果你提议的新的或者在你的机构使命的传统范围之外的话，你应该考虑准备一份集中的商业计划来支持或者证明这项提议是有正当理由的。

一旦你辨别出而且证明了合适的理由，接下来就概述选项。

通常应该有两三个可实行的替代选项。在其他的情况下，选择的范围和他们的相关价值则不是那么的显而易见。在这样的情况下，可以考虑进行环境扫描或者SWOT分析。如果问题很至关重要，那么咨询一个外部的引导师是值得的。

## 收集证据

通常最好的证据来自于你的机构内部。你记录了呼叫请求服务和金融账户。这些和其他的资源可以为你提供有价值的深刻见解。一般来说，内部数据可以提供一个好的基准或者维持现状的标准。

在你的机构之外，存在着其他的信息源。可以从专业机构和商业机构入手。供应商也可为你提供比较选项和服务生命周期内的预计成本。在网络上搜索。尽管网络上有许多的垃圾，但是也能够发现宝物。学会使用你喜欢的搜索引擎来尽可能的减少不相关的资料。不要害怕去管辖区之外的机构查找资料。

图书馆和当地的大学也是非常好的信息源。图书馆中可访问网络数据，可以搜索到学术文章和其他专业资料。如果我们不习惯使用这类的设施的话，可能会觉得有些胆怯。但是记住，图书馆员可以成为你的好朋友。你可以与公共图书馆员联系或者去当地的大学咨询专业的建议。

图书馆员同时也可以指引你在大量的统计数据中找到你需要的资料。大部分省政府和省级代理机构都收集并且公开当地的数据。尽管大部分的数据都是公开的，但是有一部分仅限于授权的代理机构访问。如果你在公共服务机构工作的话，很可能你的单位是其中被授权的机构之一。加拿大统计局网站同时也是一个非常有价值的信息来源。

记住，图书馆员可以  
成为你的好朋友。

一些专科学校和大学中有着研究与你单位使命相关的事项的实验室和科研小组或者机构。这些通过网络就能够搜索的到或者可以寻求当地图书馆员的帮助。

要记住的是，不管怎样，每个证据的价值不同。不要害怕去批评，或者持相反的意见，特别是当它所声称的与你的单位或者同事的经验不符。虽然不是每次都这样，但是如果某件事情太美好而显得难以置信的话，那么通常它不是真的。问问你自己，这个信息源是否可信。提供这些数据的机构运作是否公平或者独立，还是有着利己的意图？这个研究或者出版文章是否经过了外部评价过程？

## 整理证据

收集到足够的证据之后，你应该有序的整理好这些证据。成本研究可以由电子表格清晰的展现出来。其他的资料也可以用表格展示。确保记录信息的来源以及你是在哪里找到的。这样，万一有人质疑证据的真实性的话，你可以让他们参考原始出处。

展示数据的一个关键在于将它们放在合适的上下文中。记住，如果一件事物和其他的事物没有任何关系的话，那么它是毫无意义的。问问自己，“跟什么相比较？”

你的警察或者消防部门中的三分钟的平均救援呼叫的反应时间是否足够？我们是否可以辨别且专注于最重要的呼叫来提高效率？你的上级，当选的官员以及社区内的其他人一定会问你类似的问题。

展示数据的一个关键在于将它们放在合适的上下文中。

是否值得为一个网络注册系统花费一百万元？这是有必要的购买开支，万一没有适当的实施的话就成为了浪费资金？职员现有的培训水平是否足够？在不久的将来是否需要改变培训要求？如果答案是肯定的话，那么怎样改变？

这些问题只有在与基准对比的时候才可以回答。市场中的商品和服务的价格范围是什么？企业规范或者业绩标准是什么？是否存在着可以与你的部门对比的优良常规？

你越常使用基于证据的决策过程，它将变得越容易。

## 回顾决策过程

在你完成了分析过程之后，检查一遍整个决策过程是一个好习惯。你学到了些什么？可以如何精简这个过程或者使它更有效率？你越常使用基于证据的决策过程，它将变得越容易。知识是可以积累的。很快你就可以辨别出最好的信息来源。你将学会如何是这个过程更加有效率并且如何避免走题。

尽管基于证据的决策过程通常比其他方法花费更多的时间，但是它有自己的益处。基于证据的决定在面临详细审查的时候更加有弹性。为了我们自己和我们服务的社区，我们应该更多的将基于证据的决策方法用于思考和应用中。

在你向上级或者市议会的提议某个项目的过程中，使用有力的外部证据更有可能得到一个正面的回复。展示成功的或者具有可靠的投资回报预计的前例同样十分有效。最后，如果有人质疑的话，你可以说你提供了支持这个提议的证据。如果他们依然反对，那么你可以要求他们出示他们的数据。

# 他人对正确决策的评价

市政当局是我们经济的火车头以及大多数加拿大人的家。在地方政府中，决策与资源分配情况越来越复杂。培训我们的职员使用反应当地环境的基于证据的决策方法是完全必要的。这本可读性很强并且实用的手册是一个非常好的工具，适合所有级别的职员，并且是加强各级政府公共服务人员的绩效方面迈出的了不起的一步。

潘妮·贝恩 (Penny Ballem)，皇家医学院医学博士，温哥华市前任市执行长；卑诗大学医学院临床教授

虽然这本手册的重点在基于证据的决策方法，但是它同时也提醒了我们并不是所有的决策都是，或者都能够，绝对的基于事实上。需要被考虑的还有其他因素。有时候我们需要根据情况，内部和外部的因素，政治环境等做出做适合的决定，而不是绝对正确的决定。

正确决策：政府工作人员基于证据的决策指南概述的过程是建立一个高工作效率团队和创建一个持续进步的文化氛围的秘诀。

弗朗西斯·张 (Francis Cheung)，工程师，兰里市首席行政长

正确决策：政府工作人员基于证据的决策指南是系列手册之一，旨在确保政府提供确切需要的、创造卓越的价值、以及可分析框架内交付的服务。任何政府机构的职员都会从实用的循序渐进的项目发展方法和实际案例研究中收益。

乔治·邓肯 (George C. Duncan)，列治文市首席行政长

我有一些法律届的同事，因此我知道赢得一起案件的关键在于获得可支持我们论点的证据。作为一名地方政府的律师和警察董事会成员，我赞同政府中日益重视的基于证据的决策方法。正确的决策：政府工作人员基于证据的决策指南应该作为加拿大各级政府一地方，省级和联邦的现任工作人员和有抱负的政治家，以及他们的职员顾问的必读书目。

洛蕾娜·史戴普斯 (Lorena (Lori) Staples)，加拿大枢密院，洛蕾娜·史戴普斯律师行以及珊利奇区警察董事会成员

在一个快速变化，充满了相互竞争的利益冲突的复杂环境中，市议会及职员在全力解决问题的时候，基于证据的决策方法显得越来越重要。这本手册为追求在市政决策过程中使用更多证据的人士提供了优秀的资源。

大卫·斯图尔特 (David Stuart)，北温行政区首席行政长

这本手册是一部关于基于证据的决策方法的优秀的启蒙读物。该书措辞清晰，内容全面合理，并且提供了帮助理解的大量例子。许多新公务员的工作将从此书中收益。

洛里·沃纳梅克 (Lori Wanamaker)，反海外腐败法，金融行为监管局，卑诗省司法部副部长

如果更多的政府部门实践基于证据的决策方法，那么大部分的业务都不需要承包或者外包给私企，因为私企将毫无竞争力。

肯·维斯纳 (Ken Wiesner)，多地政府前首席行政长，加拿大市级行政行政长协会部长

## 做出正确的决策

作为一名政府工作人员，每天你都需要根据可用的资源来满足需求的重要决定。你应该采取怎样的决策方法，以及如何证明你的决策是有道理的呢？

在这本手册中，保罗·马克西姆教授，任·加里斯消防局长和教授，戴雷·普利卡斯荣誉教授，和梦娜·戴维斯法律分析师探索了基于证据而做出决策的本质、优势以及方法。

## 他人对正确决策的评价

完全版以及本书受到的其它支持，请参考封底内页。

政府机构面临着在有限的可利用资源内交付项目的日益严峻的挑战，这里真正的考验是在有效性和效率之间找到恰当的平衡。这本手册全面的考虑了如何通过基于证据的决策方法是两者兼顾，以及该方法是怎样确保公众从我们的努力和他们所交的税款中获取最大的利益的合适工具。

文森·兰隆德 (Vincent Lalonde)，理学硕士，工程师，素里市市执行长

这本可读性强的以及实用的手册是一个非常好的工具，适合所有级别的职员，并且是加强各级政府公共服务人员的绩效方面迈出了了不起的一步。

潘妮·贝恩 (Penny Ballem)，皇家医学院医学博士，温哥华市前任市执行长；卑诗大学医学院临床教授

正确决策：政府工作人员基于证据的决策指南概述的过程是建立一个高工作效率团队和创建一个持续进步的文化氛围的秘诀。

弗朗西斯·张 (Francis Cheung)，工程师，兰里市首席行政长

任何政府机构的职员都会从实用的循序渐进的项目发展方法和实际案例研究中收益。

乔治·邓肯 (George C. Duncan)，列治文市首席行政长

这本手册为追求在市政决策过程中使用更多证据的人士提供了优秀的资源。

大卫·斯图尔特 (David Stuart)，北温行政区首席行政长