



社会科学研究三个基本原理

我认为只有对社会科学研究的方法有深刻的理解，我们才能具体地运用统计的方法。如果没有这种基础知识支持的话，统计方法是没有太大用途的。很多人认为社会科学方法就是怎么样搞统计。懂统计固然重要，因为我们如果不懂统计、不会使用电脑的话就很难有成果，但是搞电脑、搞统计必须基于你对社会科学方法有很好的理解。今天的讲座中很重要的一个内容就是社会科学研究三个基本原理。

我认为在社会学一百多年的历史中，最重要、最伟大的人物就是 Otis Dudley Duncan。Duncan 以前是我们学校的一位社会学家。我和他的想法很接近，所以说在很大程度上我的“三个基本原理”是在他的基础上发展出来的。他说：“社会学不像物理学。世上惟有物理学像物理学，因为对世界上的任何物理学的理解都已成了物理学的一部分。”把社会学当作物理学的一个延伸是完全错误的。社会学不像物理学，也不能像物理学。如果社会学是像物理学的东西，那么，既然已经有物理学了，还要社会学做什么？喜欢物理学的人可能不赞同这个观点，但是假如你理解社会学和物理学的差异，就会知道其中的道理。

我先介绍一下术语的定义。我讲的“社会科学研究”是指定量的社会科学研究；“基本原理”是指可以普遍适用于实际研究的方法，而不是由实际研究（即使是经典的研究）提供的结果。

我以前是研究科学史的。这一领域被自然科学史所主宰，大概 80% ~ 90% 的科学史都是关于自然科学史的研究。你要是读自然科学史，就能体会到

柏拉图对西方科学和哲学有极大的影响。Alfred Whitehead, 一位哲学家、数学家曾这样说：“欧洲哲学传统最可靠的特征是，它是由关于柏拉图的一系列注释所组成的。”所有欧洲传统的东西甚至包括牛顿的东西，都是对柏拉图的注解，他就把柏拉图上升到一个这么重要的位置。

为什么柏拉图在西方科学史上有如此重要的地位呢？柏拉图的一个很重要的贡献就是区分了本质的世界（world of being）和形成的世界（world of becoming）。这两个世界对柏拉图来说是很重要的。柏拉图为什么要做这样的区分呢？柏拉图说，你想做一个好的哲学家（也就是科学家），就要有很好的理解能力。你必须超越你能够观察到的世界，即当前面对的这个形成的世界，这样你才能对本质的世界有很好的理解。你不能看表面而要看实质，实质就是你看不到的东西，也是永恒的东西。对柏拉图来讲，真正的知识存在于普遍而且永恒的法则之中。对柏拉图来说，这就是真理。真理是普遍的、永恒的。真理不是存在于一个具体的事物之中。他把本质的世界分出来，就是说看世界不能看具体的东西，研究具体的东西你就不能看到实质。我们中学时候都学过几何，研究过圆。圆是什么？在欧氏几何里，圆是一个平面上离一个点距离相同的点的集合。谁见过圆？比如说我们看见那个钟是圆的。但是对于柏拉图来说那不是圆，因为它不可能是理想的圆，它不是由到定点等距离的点组成的。圆是抽象的，对柏拉图来讲，要研究圆的特性，你就不能研究那个圆（钟），即不能研究任何一个能够看得到的东西。圆的特征是永恒的、不变的、稳定的。这个特征只有在哲学家的脑袋里才有，具有这个特征的只能是抽象出来的东西。所以说，柏拉图让我们做的是透过现象看本质，要把形成的世界忽略，而把本质的世界看透。

为什么这样好？因为你一旦看透了本质的东西，就可以把它应用到生活中的任何方面。比如说，圆的特性你已经知道了，那么尽管有无穷无尽的具体的圆，但它们都遵循着同样的规律，就是你已经知道的圆的特性。这就是科学的一个很大的用途。为什么科学那么神圣？为什么大家都在讲科学这个、科学那个？历史上，科学在中国文化里面并不是占据很高的地位。在西方的哲学里，从柏拉图开始，科学就从哲学中分离出来，这是因为科学是永恒的，它的永恒性和普遍性是放之四海而皆准的东西。柏拉图认为法则是存在的，它有一个创造者。他这里的创造者不是宗教意义上的创造者。因为有了这些法则的存在，我们才可以发现。大家可以理解“科学发现”这个词实际上有需要推敲的地方。什么叫发现？发现意味着本来就有东西存在于那里。假如没有东西在那

里，
西藏
有规
界的
理发
：
在实
的，
差异
说我
说是
就是
异，
你按
这就
最终
样。
不
质是
各个
图的
最重
上，
的潜
出的一
什么
漂亮
不可
验真
的东
你
们不

里，就应该叫做发明。发现是指以前就有一个人很聪明、很伟大，很早就把东西藏在那里，然后等我们去发现、去整理。那些被藏起来的東西是很好的、很有规律的东西，也就是普遍的、永恒的真理，是关于圆、关于三角形、关于世界的所有的科学的东西。牛顿也好，爱因斯坦也好，把原有的美满、永恒的真理发现出来，这就是科学的发现。

在柏拉图的哲学思想中，真理是永恒的而不是变异（variation）的。但是在实际生活中，就是柏拉图自己也知道变异是存在的。所有的圆都是不一样的，所有的东西都是不一样的。任何事情要是做比较的话，你都会发现变异和差异。柏拉图对变异的理解是什么呢？变异是对本质的世界的拙劣复制。比如说我们有个蓝图，要做一个零件，你不可能每个都做得一模一样。虽然蓝图上说是5公分，但是做出来的零件不可能刚好是5公分。根据柏拉图的观点，这就是形成的世界和本质的世界的关系。本质的世界是一个真理，它排除了差异，但是你复制的时候总会出现变异。比如说你做这个钟，圆的直径是多少，你按照这个把它做出来。一旦你先做出一个模子，这个模子必然不是标准的，这就发生了变异。模子造出来之后再去做具体的钟的时候，还是会发生变异，最终导致现实的差异，也就是每个钟的直径都不一样，每个钟的快慢也不一样。

在柏拉图看来，这些变异都不好，也都是不重要的。在他看来，世界的本质是不连续的、抽象的、单一形式的，而具体的东西总是连续的，也就是说，各个具体的东西之间总有各种差异相衔接。物理学中的伟大成就就是遵循柏拉图的类型逻辑思维的结果。类型逻辑思维（typological thinking）就是说类型是最重要的，真理存在于物与物之间类型的差别上，而不是存在于量化的差别上，也不是存在于具体的物的差异上。柏拉图的哲学也同样解决了科学和宗教的潜在矛盾。举个例子来简单地说明一下。有人说哥白尼的天文学是对宗教提出的一个很大的挑战。实际上，哥白尼的天文学是受到宗教影响的。哥白尼为什么要提出日心说？因为日心说比地心说更加简化、更加理性化、更加漂亮。漂亮的世界一定是上帝创造的。而地心说的数学模型非常复杂，哥白尼认为这不可能是上帝创造出来的。他对科学的解释不像我们以前所说的“实践是检验真理的唯一标准”那么简单。科学的可信性在于简化。科学的东西、美妙的东西一定是简化的。伽利略、牛顿对科学的解释与哥白尼是很相似的。

偏差（deviation）是什么？按照类型逻辑思维，偏差是不好的东西，是我们不希望得到的结果。通过抽象思维把这些偏差忽略掉之后才能得到真正的知

识。比如说你研究具体的圆，因为它不是标准的圆，所以得不到真理，但是如果把偏差忽略了，就可以得到真正的知识。你想想，中学的时候，老师画的每一个圆其实都是不圆的，但是我们把那些不圆的表象都忽略了。

在统计学方面，有两个很伟大的统计学家：一个是 Jacob Bernoulli，他最大的贡献是大数定律（Law of Large Numbers）；另一个是 Pierre-Simon Laplace，他的最大贡献是中心极限定理（Central Limit Theorem）。他们提供了测量不确定性的数学方法。比如说刚才那个做零件的例子，现在假设有一个零件实际上是 5 公分，让 20 个学生用很精确的仪器来测量。每个人测出来的一定不是 5 公分，有的多一点而有的少一点。这就奇怪了：为什么对同一个东西会测量出不同的常数呢？这里就存在一个误差（error），这个误差是人为的。怎么办呢？把所有的误差加起来平均一下，平均数的基数越大，平均数就越可靠，与真实值也就越接近。这就是大数定律的内容。现在假设每次都随机地找 n 个人来测量，从而每次都能得到一个基数为 n 的平均数。如果 n 比较大，但是相对于总数来说又很小（这个例子中总数是无穷大，因为你可以让无穷多的人来测量），那么做了无数次的平均数之后，就会形成正态分布，它们的平均值就是真实值。这就是中心极限定理的意义。实际上这个例子中的误差是真实的误差。为什么这样说呢？真实的常数只有一个，因为很多人在不同时候、不同温度下测量，所以结果不一样，这种误差可以通过统计补偿（statistical compensation）来抵消。

Quetelet 说，我们知道社会数据存在很大的不确定性，但是我们可以通过测量理论来解决。任何一个现象的量数都有差异。比如人的高度，每个人的高矮是不一样的。但我们把所有人的高度放在一起，就会形成一个分布，而且是一个正态分布。可能在今天和明天做的抽样中，你和他的身高不一样，但是所有人身高的平均值始终是一样的。于是他提出一个普通人（average man）的理论。柏拉图说真理是永恒的、不变的，那么社会科学里什么是永恒的、不变的东西呢？Quetelet 的办法就是取平均值（mean/average），他认为平均值就是社会科学追求的那个不变的、永恒的真理。为什么可以这么讲？这是因为测量理论可以应用到社会现象。任何一个社会现象的量数都是有差异的，但是这些量数的平均数是永恒的，这些量数都遵守正态分布。Quetelet 说：“偶然原因法则是可以应用到个体和群体的一般法则，它支配着我们的道德、智力素质，正如它支配我们的身体素质一样。”他关注的内容包括不同国家、不同地区、不同组织、不同年龄间平均数值的差异。平均数是不变的，那是因为存在一个

法则，
该具有
是在金
长。对
不富也
。从
们是西
(popul
维。大
达尔文
人不惊
异》(《
(Varia
父母都
样的记
鱼，生
样就不
上一代
殖，它
要不然
础——
丰富的
繁殖复
这
要的內
讨真实
逻辑认
理。W
人的身
的数据

法则，它是不变的原因。他甚至把这个普通人上升到伦理学的角度，普通人应该具有善良、崇高、美好的品质，最好的人是普通人。我们一般都讲最好的人是在金字塔最高处的人，先是最高领导人，然后慢慢下来，是院长、所长、组长。对 Quetelet 来讲，最好的就是平均的，既不高也不矮，既不胖也不瘦，既不富也不穷。

从柏拉图的类型逻辑思维到 Quetelet 的普通人理论，在很长一段时间里它们是西方科学的主流，但是这个主流思想受到达尔文的总体逻辑思维 (population thinking) 的挑战和修正。我现在要讲的就是达尔文的总体逻辑思维。大家都知道达尔文提出的自然选择对生物学做出了很大的贡献。实际上，达尔文的贡献对社会科学也非常重要。对达尔文来讲，变异是真实的，不是令人不悦的失误部分。达尔文的《物种起源》第一章里就写《家养状况下的变异》(Variation under Domestication)，第二章就是《自然状况下的变异》(Variation under Nature)。达尔文的思想中最重要的一个观念就是变异。虽然父母都一样，环境也一样，但是每个个体的后代都是不一样的。假如他们都一样的话，就不可能有变异，更不可能有自然选择。为什么有自然选择？比如说鱼，生下来的时候有成千上万条，这些鱼都是不一样的，如果这些鱼都长得一样就不可能有选择，也不可能选择出与自然环境更匹配的。变异特征还可以从上一代传到下一代。比如说在寒冷的海域，抗寒的鱼活下来了，接着通过繁殖，它们下一代中有一部分就更抗寒。也就是说，下一代可能是很不一样的，要不然的话就没有自然选择了。这个很重要，因为它说明了自然选择的基础——每个个体之间的差异性。而个体的差异来自每一代的繁殖过程中产生的丰富的变异。每一代的变异很多，但是只有少部分的变异能够存活下来并通过繁殖复制自身。

这就是总体逻辑思维，它和统计有什么关系呢？在类型逻辑中平均数是主要的内容。在总体逻辑中重要的是差异，平均数只是总体的一个特征值，是探讨真实原因的手段，而不是原因本身。这是两者之间很重要的一个区别。总体逻辑认为平均数只是总体的一个特征值，而在类型逻辑看来平均数代表着真理。William Jevons 说其实 mean 和 average 是不一样的。比如说长度，同一个人的身高测量 50 次得到的数据是不一样的，50 个人的身高分别测量一次得到的数据也是不一样的。前者是 mean，后者是 average。^① 做这种区分是因为这

^① 现在一般也称后者为 mean。

两组数据出现差异的原因不同。前者实际上是一个测量的误差，它真实的东西只有一个；而后者则是 50 个人之间的真正的差别，这个差别并不是误差，而是实际上的差异。社会科学的重要性在于研究为什么个体和个体有差异，而不只是比较平均数。当然，平均数也是很重要的，比如说男性的平均工资比女性高，但是谁都知道有的女人的工资比男人高。我们也知道女性的平均寿命比较长，但也有女性早死。所以说平均的差异不能代表所有的个体和个体之间的差异，这是个很重要的概念。Francis Edgeworth 也做过类似的区分。

达尔文是生物学家，他没有做过具体的社会科学研究，那么社会科学研究是谁开拓的呢？是他的表弟 Galton，也是英国人。他把总体逻辑思维引入社会科学。对他来讲，平均数的价值是有限的，他认为个体差异是真正重要的东西，所以 Quetelet 的社会物理学^①看上去很漂亮，但是在社会科学中几乎没有什么用途。对社会科学不了解的人就会和 Quetelet 想的一样，重视平均数，他不知道平均数后面掩盖了多少重要的东西。平均数并不代表一切。Quetelet 在社会科学里的影响相对来讲是很小的。那么社会科学的侧重点在哪里呢？应该看差异在哪里，应该把重点集中在变异（variation）和共变（covariation）上。Galton 是现代统计学的始祖，我们先讲他的共变。假如不注意个体和个体之间的关系，你不可能讨论这些东西。正因为考虑到了个体和个体之间的差异，他才提出了对变异和共变的测量。有这样一个比较有名的例子，是关于父亲高矮与儿子高矮关系的。如果父亲高，儿子的身高也偏高。但是父亲高有可能是因为运气的原因，儿子可能没有父亲运气那么好，只是偏高，但是没有父亲高，这就是回归（regression）。假如父亲很矮，有矮的基因，那么儿子就会偏矮，但是不一定比父亲还矮。在每一个阶段都存在重新再分布。也就是说，你的身高取决于你父亲的身高，但你究竟有多高是不能决定的，只能是随机的。

现在讲社会科学中变异性的独特之处。我认为社会科学的变异性比生物学的变异性更复杂，也更重要。这可能是一个新的观点。许多变异，特别是达尔文说的变异实际上是指当个体生下来的时候就已经完成了的变异。就拿刚才说的鱼来讲，一条鱼是否通过变异产生了抗寒的性质是在出生之前通过遗传和环境的相互作用就决定了的。而我们所说的变异，不仅是指个体间的变异，也包括同一个体在成长过程中的变异。我们说的这个变异是指你不但和你的父母不

^① 与 Auguste Comte 的社会物理学没有关系，是两个概念。

一样，人没有变异性。而达尔文说，变异性的影响。你可以调节的。政治、理性的行为。如男性基础，它们所做的事对将来，也可能通过不了。变的是一影响。社会科学的多数的真正可能就是说男女个体有差

一样，而且你今天和明天也会不一样。比如说你今天听了我的课，虽然明天你人没有变，你的基因没有变，环境也没有变，但是你的行为变化了。我们说的变异性除了指个体与生俱来的变异性外，还有个体在发展中产生的变异性。后面这种变异是受不确定因素影响的。人类行为的变异性不一定是可以遗传的，而达尔文讲的变异性是指可以从上一代遗传下来的变异性。实际上我们说的变异性不一定是可以遗传的，也不一定是物质性载体的。比如说你到这里来听我的课，你学到的东西是遗传不下去的，除非你教给你的子女，但这并不能说你可以通过基因的形式遗传下去。还有一个很重要的，就是马克思主义理论强调的一点，即人类能够改变影响他们自身的环境，这就是我们讲的社会运动和政治、经济制度的改造等等。下面一条，经济学家很重视，就是说人的行为是理性的，人能以可以预期的结果作为行动的基础。比如说上学，你认为上学对你有利，你才来上学。1992年诺贝尔经济学奖得主 Gary Becker，就是以理性行为的理论获奖的。他说婚姻就是理性的，结婚会有好处，如果不好，比如男性没有钱、没有经济依靠等就不会结婚。结婚不结婚就以理性因素做基础，上不上大学，来不来听课，你和领导的关系，请客吃饭等等，都有理性基础。也就是说，所有的行为不是随机的而是经过理性计算的。还有一个就是我们所说的历史性路径依赖（historical path dependence），也就是说以前发生的事对现在还有影响。我们不可能把以前的东西取消重来。以前发生的事情影响将来，这是社会科学中变异性的一个独特之处，即使是偶然发生的事情对将来也可能会有很重要的作用。结过婚的人可能知道，你当时认识你的爱人可能是通过非常偶然的的机会，认识之后就发生了很多事情，以后就是你想改变也改变不了。发生了就是发生了，以后的事都是以发生了的事为基础，这是不可能改变的事实。你听了这门课，或者做错了什么事，也是不可能去掉重来的。这就是一个路径依赖的问题。路径依赖是很重要的，等一下我要讲它会造成怎样的影响。

以上讲的是社会科学研究中与哲学、与历史有关的部分，下面我开始讲社会科学研究的三个基本原理。

第一个原理叫变异性原理（Variability Principle）。变异性是社会科学研究真正本质。我们不是研究类型的，而是研究变异和差异的。当然，差异也有可能是组与组之间的差异。比如说我刚才讲了，男女之间的收入不一样，假如说男女收入是一样的话，那么我们就不会认为性别在研究收入这个现象中是一个有趣的、重要的东西。我们之所以要研究性别收入差，就是因为有差别的存

在。虽然这是一个组间差别，但是我们的重点还是在差异上，比如说种族之间的差异、家庭背景之间的差异、教育的差异、家庭婚姻的差异等等。你仔细想一想，我们研究的东西是差异，而不是共性。类型之间的差异也是差异，只不过是一个特例而已。下面我就会讲到，类型之间的、组与组之间的差异是随着研究对象的变化而变化的。

下面讲第二个原理：社会分组原理（Social Grouping Principle）。为什么要分组？因为社会分组可减少组内差异。个体之间是很不一样的，但是个体可以分成组，比如分成年龄组、性别组、家庭背景境况组等等。分组显示了组与组之间的差异，这意味着每个组里面有相对组外来说更高的共同性。如果一个组当中没有共同性的话，那么组和组之间就没有差异性。如果女人没有共同性的话，那么就不可能有男和女的差异。同样的道理，每一个组当中有相对来说比较相似的特点。我们分组是基于组员之间有一定的共同性，分组以后组与组之间就有差异性。

什么是社会分组（social grouping）？我并不想在分组是唯名的还是唯实的之间有所选择。组有可能是实在的，也有可能是唯名的。这和马克思主义的阶级观念不一样，马克思主义认为社会关系建立在所有权上。我认为到底用名称还是真正的物质基础来划分不是很重要，社会分组只有根据社会结果（social outcomes）来分才会是有意义的。我不是看你分组的原则是什么，是真的还是假的，是符合这个理论还是那个理论。从统计的观点来讲，分组的意义在于它有利于研究社会结果的差异。我刚才讲的收入，假如男女之间没有收入差异的话，那么对于收入而言，根据性别分组就不重要。我是研究人口的，比如我们要研究人的死亡率，这就要考虑性别，因为不同性别之间的死亡率差很多。反过来说，假如性别之间没有死亡率的差异，那我们就不要考虑以性别来分组。分组的意义在于它是否能解释差异，所以社会分组应用于不同的社会结果或许会有不同的意义。比如说种族、民族、社会阶层、家庭背景、党派、政治面貌等很多分组，这些分组在不同的社会结果下的意义是不一样的。比如说你的政治态度和你的死亡没有关系，有关系的可能是政治态度与社会经济地位的关系，政治态度对你行为的影响。比如说，政治运动来了，你到底是参与还是不参与？这时你的政治倾向就和分组很有关系。而政治倾向和死亡率根本没有关系，至少没有直接的关系，而只是伪相关（spurious correlation）。假如说我要研究你的政治行为，那么你的政治态度和分组就很重要。

社会分组能减少社会结果的差异性，减少得越多，社会分组就越有意义。

下面
性。
的也
到。
的信
亿个
消化
简单
息，
信息
一样
教育
数越
据越
种妥
告诉
既想
作妥
体。
经告
就越
组之
代表
好，
会分
体的
很多
盾。
下面
的。
释的

下面我会讲为什么会这样。我刚才说了，社会科学所要理解和解释的就是差异性。实际上社会科学最想做到的和柏拉图提出的并没有什么两样，我们想提供的也是很简单的、永恒不变的东西。我们能做到的话，很好，但是我们却做不到。那做不到怎么办呢？我们只能做一个妥协，也就是说，我们不可能把所有的信息都告诉人们，那么就告诉人们大量信息中很少的一部分。比如说有 10 亿个人，有 10 亿个信息，这么多信息我根本不能告诉你，连我自己都不能够消化。我只能把很大的总体中的一部分信息告诉你，用的参数越少，信息就越简单，就越容易理解和讨论。比如说平均数，它也是大的总体里面的一个信息，这个信息可以理解，可以在相互之间传播。当然，平均数是一个比较粗的信息，也可以不用平均数。我只用两三个数值就可以告诉你男女之间有什么不一样，如受教育程度不一样，死亡率不一样，收入也不一样。或者我就告诉你教育的增加和收入的增加有什么关系，这只是一个参数。我想概括的总体的参数越少越好，与柏拉图讲的科学是一样的。也就是说，告诉你的信息里参数数据越少越好，越精越好，这和自然科学是一样的。但是你应该知道，这只是一种妥协而已。我们没有办法把大批量的、整体的信息告诉他人，就只好作妥协告诉他一部分。当然，告诉他的这一部分是有意义的一部分。在这种条件下，既想刻画整个总体，又要用简化的方法告诉对方很少的信息，在这两个要求下作妥协，也就是说，我虽然只告诉你一部分信息，但是这一部分概括了总体。比如说平均值和方差，这两个参数不能告诉你整个群体的情况，但是已经告诉你很多的情况。比如说性别社会分组能够解释的东西越多，它的分组就越好。如果世界上只有 10 种人，所有的人之间的差异都归于 10 组中组与组之间的差异，而每组内部没有差异，那么我告诉你每个组的平均值，就能代表所有的人。可惜这是不可能的。社会分组能够解释的社会变异越多越好，但是组内差异永远是存在的，你永远不能把组内差异解释穷尽，这是社会分组所不能解释的差异性。

社会科学的复杂性就在这里。我们想刻画总体，但是我们不能把所有的总体的东西都刻画出来，只能告诉你一部分参数。告诉你一部分参数就能告诉你很多关于总体的东西，但是还是有很多东西没有解释，这就是我们遇到的矛盾。分组以后，你知道组之间的差异，但组当中还有差别，怎么处理？这就是下面我要讲到的，由于没有控制其他的变量，这样做可能是粗糙的甚至是错误的。比如说，死亡率和政治倾向在表面上好像有关系，但是结果你会发现你解释的也许并不是政治倾向，而是经济地位，因为经济地位与死亡率有关系，与

政治倾向也有关系。这里，简化（parsimony）是一个很重要的词。中世纪有一个神学家，他认为上帝是理性的，他创造的世界一定是简化的，假如搞不清的话，你就用一把剃刀把多余的部分从知识中剔除出去。^① 我们社会分组的原则也是简化。我刚才说了，一个总体有千千万万的信息，而我要告诉你的是一些简化的东西。简化会有代价，在社会科学当中就体现为误差。这种误差并不是自然科学中测量的误差，不可以忽略不计。社会科学中的这个误差是一个真实性的、理解性的误差，是知识上的一个缺陷，是真实的而不是可以忽略不计的。这就是和自然科学或者说物理学不一样的地方。比如说，物理学的误差经过测量许多次取平均数就可以忽略掉，而在社会科学中误差就是知识上的缺陷。

第三个原理是社会情境原理（Social Context Principle）。群体变异性的模式会随着社会情境（social context）的变化而变化，这种社会情境常常是由时间和空间来界定的。也就是说，社会情境不一样，变异性就不一样。比如说，教育对收入的影响是每个社会都有的，教育高，收入高，但是并不是所有的人教育高收入一定就高。大街上，有些小学、中学没有毕业的人也有可能比你还有钱，这就是差异，但教育对收入的影响可能随着社会的变化而不一样。比如说社会制度，改革之前和改革之后就不一样，中国和英国不一样，美国和日本也不一样，中世纪的英国和现在的英国也不一样。比如说有的经济学家认为，经济发展快的话，教育的回报率高，为什么？因为机会多。教育高，知识多，容易接受新的事物。而社会经济停滞不前，就会有饱和的现象，回报率就会降低。我想讲的是，随着社会情境的变化，变异性的规律和模式也会变化，我下面讲的一些统计模型会用到这些。

社会情境不同于社会分组。因为社会情境是有边界的，是一个独立的社会系统，比如说社会主义制度、资本主义制度、美国和英国，以前的中世纪和现在，而社会分组是没有边界的。男女生活在同一个家庭，在一起工作，是没有界限的，只是因为社会研究的统计和调查的需要把他们分出来。那么，社会情境的边界是怎么来的呢？边界有两个：一个是时间，一个是空间。社会科学经常要用到时空的概念。时空很重要，达尔文讲进化论的时候也用到了时空。他不可能找到物种以前的状态，也就是说，历史已经发生了，他可能找到一些诸

^① 这位神学家是 William Ockham，新唯名论的创始人。“Ockham 剃刀”是一种经济思维原则，就是要把既不能在逻辑上自明，又没有经验证据的东西从知识中剔除出去。

如
是
再
经
意
和
关
有
怎
些
地
会

的
关
会
用
的
政
在
比
作
一
能
的
不
这
可
之
们

学

如化石的东西，但是他不能证明。他不能找到时间的差异，那还有一种办法就是找到空间的差异。南美可能有些东西还没有进化，与欧洲、非洲的不一样。再比如，想看看以前经济不发达的中国是怎么样的，可是在中国的发达地区已经找不到这样的地方了，那就到落后的地方去找。这就是说时和空有很重要的意义。因为你要知道社会情境对社会关系的影响可能是不一样的，就是因为时和空是不一样的。我会讲一些具体的、实际的方法，但是今天只是给你们一个关于时和空的杠杆，使你们得到一些你们本来不能得到的东西。比如，你们只有中国计划经济时期的总体数据，不知道成熟的市场经济制度对收入的影响，怎么办？那就去找一些不同的点、不同的地区。如果你认为深圳更市场化一些，有些地方保留更多的计划经济特征，你不可能接触到将来，但是假设其他地区以后会像深圳这样，那你就做一下地区之间的差异的比较。这个差异可能会帮助你预测将来其他地区会怎么样，当然，这样做你还必须做很多假设。

个体变异的模式可以被个体之间的关系所支配，这也是社会分组无法解释的。社会分组只是讲把一部分人摆在一起，把另外一部分人摆在一起。而社会关系可能随环境而变，有些关系不是个人的，而是一个系统的属性，比如说社会制度，这是社会这个系统的属性而不是个人的属性。人们居住在一起可能采用不同的制度、不同的法律，拥有不同的文化，这都是环境的属性而不是个人的属性。所以，个体变异的模式也可以被宏观的条件所支配，比如社会结构、政治结构和文化结构。这些条件也许是间断的，即在某段时期内是确定的，但在时期之间是跳跃的。比如通过一项关于婚姻的或是关于社会制度的法律，再比如中华人民共和国1949年成立，都是一下子发生的条件变化。而这些条件作用的结果是连续的、不断变异的。说结果是连续的有两层意思：第一，任何一个确定的条件下的结果都是一个连续体，也就是说结果是各种各样的，不可能是单一的；第二，在条件突然发生变化时，结果的变化并不都是一下子显现的，而是可能和条件变化在时间上相分离的。有一些社会条件是突然变化的，不是确定会发生的。人的行为可以导致宏观条件的变化和人们关系间的变化，这就是社会变迁的主要来源。这是生物学没有的。我们是理性的，我们在一起可能会做一些或者理性或者看起来是不理性的事情。我们生活在一个社会系统之中，有统治阶层和被统治阶层，人们之间总是会有矛盾的，在这个时候，人们的行为会引发法律、政治和文化的变化，这些变化会造成社会变迁。

下面看一些具体的统计例子。比如说在社会科学中，至少在量化的社会科学研究中，最重要的方法就是回归分析（regression analysis）。通过对调查来的

数据进行回归分析，比如说最典型的教育对收入的影响即可以做一个回归分析。我现在就举这个例子，看一些具体的问题。在这个问题当中，我们假设一个由 N 个个体组成的总体，其中 N 可能趋向于无穷大。有一个被关注的结果或社会现象作为因变量 Y ，比如说是收入，并以实线来度量。我们感兴趣的自变量是 D ，它有两个数值：要么是 $D = 1$ ，表示干预 (treatment)；要么是 $D = 0$ ，表示控制 (control)，即不干预。比如说读大学， Y 就是收入， $D = 1$ 就是读大学， $D = 0$ 就是没有读大学。假定每个人都读完高中，一个总体中，任何个体要么是读大学的，要么是没有读大学的。我现在要看的的就是最简化的典型案例。在这种情况下，干预作为原因会引起怎样的因果关系 (causal relation)。对某一个个体 i 来说，我现在不讲抽样的问题，相对每一个人的观测数据，存在着一个反事实的结果 (counterfactual effect)。我想知道读大学对收入的影响，我就应该考虑对某一个人来说，有了大学教育和没有大学教育的收入差距。这一差距必须是大学教育的影响。对每一个人，我都想得到两个数据：一个是读大学之后的，另一个是不读大学的情况下的。但是实际上，我们只能看到其中一个：要么是某一个人读了大学之后的收入，要么是他没有读大学的收入。要想知道读大学对一个人的影响，必须同时知道这两个收入数据。但是实际上，数据只能告诉你一种情况。所以说，如果没有假设，社会科学的量化是做不成的。在个体层面上你根本不可能得到因果关系，因为你无法找到反事实情况下的同一个个体作为对照组。如果你已经读了大学，我就没有办法知道如果你不读大学的收入情况会是怎样的。

我给大家举一个极端的例子，一个自然科学的例子，自然科学怎么办？比如说我想知道温度对电灯泡的寿命的影响。如果说一个房间室温很高，而另一个房间室温很低，我们可以观察电灯泡的寿命在这两个房间是否一样。在这个例子中，我可以简单地假设这个房间的电灯泡和那个房间的电灯泡是一样的，事实上这个假设也经常是成立的。假如是一样的话，它们之间的差异在什么地方？就在于室内的温度不一样。如果说任何读了大学的人和任何没有读大学的人没有本质上的差异，他们的差异只是反映在有没有读大学上，并且任何读大学的人之间也没有本质上的差异，这就是个体同质性 (homogeneity) 假设。假如这个假设是成立的话，我们根本没有必要去抽样和做大型的调查，我们只要研究两个人就可以了，即一个读了大学的和另一个没有读大学的，因为所有读大学的人都是一样的，所有没有读大学的人也都是一样的。所有的人都一样，那么我们就没有必要做调查。对自然科学家来说，一个分子就是一个分

子，
房间的
以了。
会性的
样的，
门当户
门当户
现实的
门当户
结果可
-
笼统 |
(popu
搞农
植物，
如你
变异，
之间
析研
在于
是你
案可
不能
-
一部
是两
是 E
收入
 $E(Y^1)$
 $E(Y_0)$
在
可观

子，元素就是元素，而不管它们是今天的还是明天的，是这个房间的还是那个房间的。这就是同质性的问题。如果具有了同质性，我们只需要两个个体就可以了。所以我们说笑话，任何人都能做社会学，只要把两个人的差异看成是社会性的差异就行了。但是这样了解的只有两个个案。这里的逻辑是，所有的这样的人都是差不多的。比如说，一些人说门当户对的婚姻都不好，因为某某是门当户对的婚姻，结果不好；另外一些人说，门当户对的婚姻好，因为某某是门当户对的婚姻，结果很好。这些都隐含着依据非常不现实的假设做的非常不现实的评价。为什么这样说呢？因为所有门当户对的婚姻是不一样的，所有不门当户对的婚姻也是不一样的。假如有很大的异质性（heterogeneity）的话，结果可能就不一样了。

一般的人做社会科学看到的是局部的、片面的东西，所以他们的结论是笼统的，没有代表性。我讲的要义在于，正因为存在总体变异性（population variability），科学抽样（scientific sampling）才成为必要。这个是搞农学的人最早提出来的，因为每一棵植物是不一样的，假如说你只取一棵植物，你怎么知道另一棵植物有同样的结果？人也一样，人也有变异性。假如你接受第一个原理（变异性原理），你就能接受这个道理。正因为总体的变异，我们要做科学抽样，取得一个样本，而样本具有代表性。个体和个体之间是有差异的，没有差异就不需要代表性了。换句话说，我们强调定量分析研究方法是因为定性分析归根结底在描述变异的现象时是不可靠的，原因在于你不知道你的个案是不是有代表性。从个案中你当然可以看得很深，但是你做定性研究，始终不知道你的研究对象是不是具有代表性。你选择的个案可能没有代表性，可能是错的。因为总体大，一定有各种各样的人，个案不能说明总体。

让我们回到大学教育和收入的例子。把总体 P 划分为两个部分，比如说一部分 P_1 是被干预的，即 $D_i = 1$ ，另一部分 P_0 是未被干预的，即 $D_i = 0$ 。这是两个子群体。 q 为 P_0 在总体 P 中的比例。已经读了大学的人的收入平均值是 $E(Y_1^1) = E(Y^1 | D = 1)$ 。对已经读了大学的人，如果他们没读大学的平均收入是 $E(Y_1^0) = E(Y^0 | D = 1)$ 。没有读大学的人的收入平均值是 $E(Y_0^0) = E(Y^0 | D = 0)$ 。没有读大学的人的假如读了大学，他们的收入平均值是 $E(Y_0^1) = E(Y^1 | D = 0)$ 。

在这四个公式里面， $E(Y_1^1) = E(Y^1 | D = 1)$ 和 $E(Y_0^0) = E(Y^0 | D = 0)$ 是可观测到的，其他两个则是反事实的。为什么要这么想？ $E(Y_1^1) = E(Y^1 | D =$

1) 和 $E(Y_0^1) = E(Y^1 | D=0)$ 是不一样的, 因为它们代表了两个总体: $E(Y_1^1) = E(Y^1 | D=1)$ 是已经读了大学的人的平均收入, $E(Y_0^1) = E(Y^1 | D=0)$ 是没有读大学的人如果读了大学的平均收入。在同样读了大学的情况下, 这两组人的平均收入可能是不一样的。我们假设它们一样, 问题就容易了, 但这仅仅是假设而不是事实。

同样, $E(Y_1^0) = E(Y^0 | D=1)$ 和 $E(Y_0^0) = E(Y^0 | D=0)$ 也代表了两个总体。前者是已经读了大学的人如果不读大学, 他们的平均收入是多少, 后者是没有读大学的人的平均收入。这两组有可能是一样的, 但是也有可能是不一样的。社会科学研究当中, 我们常常已经做了这个假设而不说明。这个假设使得我们过高地估计了大学回报率。因为能力强的人、守纪律的人、听话的人、与领导关系相处好的人工资会比较高, 这些人上大学的可能性也比较大。因此, 上过大学的人即使没有上大学的话, 他们的收入也会比没有上过大学的人收入高。你相信吗? 假如这个世界把大学取消了, 两组人(原本读大学的和原本没有读大学的)中前者本来就应该比后者要好, 不是因为读大学他们的收入高, 而是因为他们本来就具有比较好的素质, 而这些素质和上大学是有相关性的。

如果你比较 $E(Y_1^1) = E(Y^1 | D=1)$ 和 $E(Y_0^0) = E(Y^0 | D=0)$, 实际上你假设了 $E(Y^1 | D=1) = E(Y^1 | D=0)$ 和 $E(Y^0 | D=1) = E(Y^0 | D=0)$, 就是没有上大学的人假如他们上了大学, 他们的收入水平应该是与那些已经上了大学的人的收入水平相似。现在上了大学的人, 假如他们不上大学, 他们的收入类似于没有上大学的人的收入。有了这两个假设你才可以做 $E(Y^1 | D=1)$ 和 $E(Y^0 | D=0)$ 的比较。这类比较是我们经常看到的, 只是我们并不一定意识到它背后的含义。

根据总期望值规则 (total expectation rule):

$$\begin{aligned} E(Y^1 - Y^0) &= E(Y_1^1 - Y_1^0)(1 - q) + E(Y_0^1 - Y_0^0)q \\ &= E(Y_1^1 - Y_0^0) - E(Y_1^0 - Y_0^0) - (\delta_1 - \delta_0)q \end{aligned}$$

其中: $\delta_1 = E(Y_1^1 - Y_1^0)$, $\delta_0 = E(Y_0^1 - Y_0^0)$ 。

假如说所有上大学的和不上大学的人的收入的平均差异是 $E(Y^1 - Y^0)$, 就是平均的回报因果关系。它可以分解为以下两部分: $E(Y_1^1 - Y_1^0)(1 - q)$ 和 $E(Y_0^1 - Y_0^0)q$ 。第一部分是已经上了大学的人和他们如果不上大学相比, 他们会有多少平均收入差异; 第二部分就是没有上大学的人和他们如果上了大学相

比,
 $E(Y_1^1)$
部分
两组
质性
即使
 $E(Y_1^1)$
 $Y_0^0)$
人上
好处

人的
设,

实际

这两:
(Y_1^1
要知:
 $E(Y_1^1$
 $\delta_0)q$:

学的
种偏
人的
人能
学等:
个假
的影
只有:

比，他们的平均收入差异是多少。它也可以再分解成三部分： $E(Y_1' - Y_0^c)$ 、 $E(Y_1^c - Y_0^c)$ 和 $(\delta_1 - \delta_0)q$ ，其中： $\delta_1 = E(Y_1' - Y_1^c)$ ， $\delta_0 = E(Y_0' - Y_0^c)$ 。第一部分就是上大学的和没有上大学的这两组人之间的简单比较。第二部分是假如两组人都不上大学的话，他们的情况有什么差异，这就是他们的未观察到的异质性问题（unobserved heterogeneity）。上大学的人可能比不上大学的人要好，即使在两组人都不上大学的情况下，他们也可能会有差异。第三部分中， $\delta_1 = E(Y_1' - Y_1^c)$ 是上大学的这组人得益于大学对他们收入的影响， $\delta_0 = E(Y_0' - Y_0^c)$ 是没有上大学的这组人如果上了大学能够增加的收入， $\delta_1 - \delta_0$ 即第一组人上大学得到的好处减去第二组人如果上大学得到的好处。换言之，读大学的好处对这两组人可能是不一样的。

在现实工作中，我们经常把读过大学的人的平均收入减去没有读过大学的人的平均收入，以为得到的结果就是读大学的好处和回报。这一方法靠的是假设，假设这两组人是没有差异的。平时我们计算的是：

$$E(Y_1' - Y_0^c)$$

实际上讲的是：

$$E(Y_1' - Y_0^c) - E(Y_1^c - Y_0^c) - (\delta_1 - \delta_0)q$$

这两者不一定是一样的。需要两个假设使它们相同，我已讲到这两个假设。 $E(Y_1' - Y_0^c)$ 是我们平时做的，可是要可靠地得到你想知道的东西，实际上还要知道两个信息：第一个就是在没有接受大学教育的情况下两组人之间的差异 $E(Y_1^c - Y_0^c)$ ；第二个就是两组人接受大学教育的回报率可能是不一样的 $(\delta_1 - \delta_0)q$ ，这个是经常被忽略的。

另外一种表达方式是这样的：常用的估计值就是 $E(Y_1' - Y_0^c)$ ，即读过大学的人的收入平均值和没有读过大学的人的收入平均值的差。这个估计值有两种偏误：第一种偏误的起因是在没有干预（没有读大学）的情况下，这两组人的平均有差异，就是这两组人在都没有读大学的情况下是不一样的，即一组人能力比较强，或者比较用功，或者比较听老师的话，或者比较能读书、能自学等等，这种偏误叫做异质性偏误（heterogeneity bias）。第二种偏误源于第二个假设，就是大学教育这个因素对两组人的影响是一样的，而实际上大学教育的影响可能是不一样的，由此产生的偏误叫内生性偏误（endogeneity bias）。只有在这两种偏误都不存在的情况下，你才可以用 $E(Y_1' - Y_0^c)$ 来计算。第一

种偏误很简单：有些人天生就聪明，有些人天生就笨；有些人天生就勤劳，有些人天生就懒。这些不需要做干预。我的同事 Robert Willis 有一个模型是用来对付第二种偏误的：^①有的人适合读大学，他们读大学受益多；有的人不适合读大学，他们不读大学反而好。所以说，每一个人对相同的干预的反应是不一样的，这就是内生性偏误。你不能假设这两组人读大学的效果是一样的，这两组人存在着系统性的差异。只有两种情况都没有，才可以作简单的比较。我花了十多年的时间才真正理解了这个道理。

在随机赋值（random assignment）的条件下，两种偏误的平均值都等于零。为什么要做随机赋值呢？这样就保证了在没有干预之前两组之间应该一样，干预之后两组之间也应该一样。这就是为什么在实验的情况下，随机赋值能够解决异质性偏误和内生性偏误的双重问题。

这是一个一元回归模型：

$$Y_i = \alpha + \delta_i D_i + \varepsilon_i$$

在这种情况下， D_i 是指干预或是控制。我认为真正写方程时应该考虑 $\delta_i D_i$ ，因为干预效应是不一样的。两种假设的含义是这样的：

异质性： ε_i ，如果 $\text{Corr}(\varepsilon, D) = 0$ ，则无异质性偏误，就是说被忽略的变量（两组人本质上的差异）和读大学或不读大学没有关系；

内生性： δ_i ，如果 $\text{Corr}(\delta, D) = 0$ ，则无内生性偏误，就是说回报率和读大学或不读大学没有关系。

下面我对以上说的内容做五点注释：

注释一：随机赋值是指不管你接受干预还是不接受干预（即 D 无论取值为 1 还是 0）和任何东西都没有关系，和 α 没有关系，和 ε 也没有关系。 D 本身在研究中是一个随机变量。

注释二：异质性偏误可以由忽略变量偏误（omitted-variable bias）产生，

^① Robert Willis 现为密歇根大学经济学系教授，兼任密歇根大学调查研究中心（Survey Research Center）和人口研究中心（Population Studies Center）教授研究员。文中提到的这一模型，请参见 Willis, R. and S. Rosen. 1979. "Education and Self-Selection," *The Journal of Political Economy* 87 : S7 - S36。

比如
比如
获，
具有
或缩
情况
会的
你的
情况
有限
息。
要么
原理
所说
们可
也就
如说
读大
不是
境比
似的
我们

比如说能力、政治倾向等等。

注释三：内生性偏误可能由理性的期望行为 (anticipatory behavior) 产生，比如说，因为你知道读大学对你有好处，所以你来读大学；你觉得你会有收获，你才来听我的讲座，这就是一个内生性偏误。这是因为人和动物不一样，具有理性。

注释四：内生性的变化 (δ_i) 意味着 Y 最终的变异性可以随干预 D 增大或缩小。上面这个方程，如果得益大就参加，得益小就不参加，这会出现什么情况？社会差距被拉大了。所有的社会中都有教育回报，而去读大学的人的社会回报又特别高，这样教育就会造成社会不平等。这是加强的例子。也有减少的，比如说靠政策，你受益小的我就多给你一些。比如在美国，你的工资高，你的税也就高，这样做就会减少差异性。

注释五：这个模型是不可估计的，它需要约束条件 (constraint)，一般的情况下，我们假定 $\text{Corr}(\varepsilon, D) = 0$, $\delta_i = \delta$ (常数)。

我在评价经济学家 Charles Manski^① 的书中这样写道：“当观察的数据十分有限时，需要很强的假定才能产生显著的结果。在统计学中，没有免费的信息。要么你去收集它，要么你去假定它。”没有其他办法，因为差异这么大，要么你去收集信息，要么你去假设它，假设是不可避免的。那么我们用第二个原理（社会分组原理）怎么做呢？我们就用社会分组来控制异质性。如前面所说的，社会分组总是减少组内异质性，意味着更多的组内同质性。此后，我们可能会得到这样一个假设，即组内的差异是可以忽略的。

假定条件： $\varepsilon \perp D \mid X$ ，组内无异质性偏误，这就比刚才的假设要弱一些。也就是说，我们不假定任何读大学的人和任何不读大学的人是完全一样的。比如说我们测量了人们中学的成绩，我们假定在同样成绩的人中读大学的和没有读大学的没有差异。也就是说，我们控制了可以看到的变量之后再做假设，而不是在控制这些变量前就来做假设。例如家庭背景，能上大学的人相对来说家境比较好，父母收入比较高，那么我们控制了家境这个因素，再来假设家境相似的人无论读没有读大学都没有异质性差异。这就是社会分组的方法，也就是我们常常做的多元分析的方法。通过控制社会分组以后我们再做无异质性偏误

① Charles Manski 现系西北大学经济学系教授，其研究领域主要包括计量经济学、决策判断、社会政策分析。

和无内生性偏误的假设，这样的假设比直接做的假设要弱一些，因而更符合实际。于是我们把这个方程扩展为：

$$Y_i = \alpha + \delta D_i + \beta' X_i + \varepsilon_i$$

这个方程不但有 D 还有 X 。这就是我们为什么要用多元回归分析，因为多元分析可以让我们控制一些和 D 相关的自变量。这个方程式是可以估计的。

下面我对以上说的内容再做两点注释：

注释一：对 X 而言，它需要与 D 相关，这一条件叫做相关条件（correlation condition），并且影响 Y ，这一条件叫做有关条件（relevance condition）。我刚才为什么说要测中学的成绩，因为中学的成绩和读大学有关系，家庭背景和读大学也有关系，没有关系的 X 可以不用去考虑。

注释二： X 应该是发生在干预 D 之前，比如说家庭背景和你读大学以前而不是以后的成绩；而且 X 对 Y 有影响，比如说中学的成绩和能力对你的收入的影响。

这些方法都依赖于较强的、不可检验的假定。没有假设就没法做。归根结底，社会统计可以建立不同的模型，因为你对社会现象做出的假设是不一样的，一种假设用一种模型得到一串数字，另一种假设用一种模型又得到另一串数字。

那么要是效果或者回报不一样，也就是解释的效力有差异怎么办呢？这就涉及自由度（degree of freedom）的问题，很难解决。有一个方法我们用得比较多，比如说一个家庭有几个子女，一个地区有很多人，就假设在同一个社会经济环境下，在一定的时间和空间的情况下，有一种模式具有同质性。我刚才介绍给你们的例子，比如说收入在北京和上海是不一样的，在内地也不一样，那么地区之间的差异就可以用嵌套数据来解决。假设在一个地区里面收入的模型是一样的，要是假设人和人之间不一样的话，就没法去计算。我们可以估算地区之间的差异、家庭之间或是组织之间的差异，但不能把人和人的差异反映出来，因为数据是个人层面的。所以这里需要假设 δ_i 为 δ_k ， k 是社会情境而不是个人。假设 δ 即回报在整个社会环境 k 当中是一致的，那么这个假设就让你得到更多的方法获取更多的数据资料。观察的数据是个人的，但是又有地区之间的差异，就是说一个地区有不同的个人的数据。 k 可以是空间也可以是时

间，这

注
布，这
数据的
注
具有共
注
说，多
model)
curve r
样的。
如果人
括。

现
(
要抽样
(
社会学
东西，
述，我
少小孩
(
实验而
个农村
才能推
(
我也觉
人说美
很糟糕
没有办

间，这就是一个社会情境原理的应用。

注释一：可以在个体水平作一个参数假定，并使它遵守一定的参数化分布，这就是贝叶斯方法（Bayesian approach）。现在这个方法用得比较多，其数据的结构一定要是嵌套的。

注释二：你用嵌套这个方法就可以假设在共同的社会情境中，不同的个体具有共同的特征。

注释三：假如跨越社会情境的变异是系统的，你就可以来建模型。比如说，多分层模型（multi-level model）或多层线性模型（hierarchical linear model）、随机系数模型（random coefficient model）和增长曲线模型（growth curve model），这些模型现在用得比较频繁，比如一个小孩增长的曲线是怎么样。这是同样的道理，因为每一个人都是相似的，他们的成长是有系统的。如果人和人之间的差异不能用系统的方法来概括，那么就可以用饱和模型来概括。

现在做几个结论：

（1）抽样很重要。因为我们只能讨论总体特征，我们不能讲个体，所以要抽样。

（2）描述性研究是很重要的。很多人看不起描述性的研究，但是我们搞社会学的、搞量化研究的，在没有很强的假定的条件下，能做的只是描述性的东西，我觉得这是很伟大、很重要的东西，否则我们什么也不知道。没有描述，我们就不知道你能活多久，你的生活状况怎么样，家庭质量怎么样，有多少小孩等等。

（3）随机实验不能完全解决我们的问题。因为我们做实验不是对总体做实验而是对一部分做实验，很难把实验结果推广到总体。比如说你在乡下的一个农村做实验，这个实验结果能够推广到全国吗？不一定，你要做很多的假设才能推广。

（4）统计学虽然不完美，但却是社会科学刻画异质性唯一可靠的工具。我也觉得统计学并不能解决一切问题，但是除了它以外没有更好的办法了。有人说美国要民主很不好，当然民主可能会运作得很不好，有时会造成很荒谬、很糟糕的结果，假如世界上有比民主更好的东西，我们一定不会要民主。这是没有办法的事情。统计也一样，因为没有比它更好的东西。在个体水平上，反

事实结果是不可能被评估和证明的。你看完一个案例，你能说它可以推广吗？不可能的。所以说定量是唯一的科学方法，没有定量的方法就没有科学的东西。

(5) 只有当解释涉及所研究的总体时，统计结果才会有意义。很多人的统计报告并不告诉你报告的结果对哪个总体是适用的。做统计的时候，你就要想到你的总体对象是什么，并不是所有的对象都是适用的。比如有些结论只有北京才能适用，有的只对大学生适用，有的只对男生适用等等。你做事情一定要有一个总体的概念。没有总体就不需要做统计，你可以写感想、作汇报、谈经验等。那统计是什么意思呢？实际上，统计是有一个加权的，即合计结果本质上是加权的。

(6) 因果关系总是概率性的。比如吸烟会导致癌症，但并不是说每个人吸烟都会导致癌症。

(7) 我们只能研究原因的结果 (effects of causes)，而不是结果的原因 (causes of effects)。这就是识别问题 (identification problem)。因为一个事情发生了，它可能是由不同的原因造成的。比如说有十个原因，那到底是一个原因造成的还是由十个原因一起造成的呢？很多人搞不清楚，他们说引起这个事情发生有五个原因或者有十个原因。到底是这些原因都存在了，还是只要其中一部分原因存在就可以导致事件的发生就不清楚了。我们能做的就是解释有了一个特别原因会有什么样的结果，比如说教育对你的生活、婚姻和收入会有什么影响。但是假如说你问为什么你有钱，你就没法解释。

在统计分析的时候，我们总要做假设，所以说理论是很重要的，我没有说理论不重要。理论有两个要点：一个是对以前经验的总结概括，这是一个积累的过程，以前的经验知识可能会变成理论的东西，比如说为什么教育对人会有收入的回报，这就是对以前的经验的概括。另一个要点是，理论是一个逻辑的思维过程。有了理论的思考才能做一个比较好的假设，没有理论就没有假设。

参考文献

Willis, R. and S. Rosen. 1979. "Education and Self-Selection," *The Journal of Political Economy* 87 : S7 - S36.

这是技术——
一些体会
社会科学
当中的知
识的民族
很多有学
手
法。个
见解。
的，个
批判性
么快，
是对自
理。才

？
卡
的
要
有
定
谈
本
人

因
青
发
原
因
事
情
中
一
了
一
什
么
有
说
积
累
会
有
理
辑
的
假
设。

Political



社会科学与自然科学的关系

这个报告强调哲学方面的内容，最重要的是思想上和观念上的内容，而不是技术方面的内容。这些思想上的、观念上的内容——可能和你以前学的不一样——是我多年来对方法论的理解，对社会科学和自然科学的了解和自己的一些体会。

社会科学和自然科学的关系事实上是一个很宽泛的问题。任何一个研究社会科学的人都会接触到这个问题，即社会科学和自然科学的关系是什么？你们当中有不少是学社会学的，很多老师也教过社会学方面的课，了解不少社会学的知识，所以我想你们对社会学有一定的认识，从课程中学到了很多关于社会学的知识，比如中国的现状、中国的人口、中国的家庭、中国的文化、中国的民族等等。但是我想提出这样一个问题：社会学的核心是什么呢？你可能学了很多东西，对社会有很深的了解，掌握很多关于社会的知识，但是你可能还没有学到社会学的核心。

我认为社会学的核心并不是关于社会的知识，而是得到知识的手段和方法。它不是你从一些学者的言论中获得的知識，而是你对他们的言论的批判性见解。这是最重要的。知道什么并不重要，而知道你应该知道什么，什么是对的，什么是衡量知识的标准，才是最重要的。所以，社会学的学习是培养一种批判性的思维。知识是变化的，社会是变化的，尤其是中国社会现在变化得那么快，你今天学的东西到明天可能就没有用了。但是你应该知道什么样的知识是对的，什么样的知识是错的，你要能对它作一个评价，评价它的结论是不是有道理。权威人物讲的并不一定就是正确的，因此不要趋之若鹜。所以我在训练学

生的时候，特别强调学生要独立地思考，批判地思考，质疑他人观点是否正确，结论是否有道理。

我为什么这样讲呢？我做过美国最著名的社会学刊物的副主编，还做过一些其他的工作。假如你仔细读社会学的文章，很多文章在方法上是靠不住的，错误层出不穷。有的人认为社会学就是技术性的东西，只是作简单描述；有的人认为社会学只是对政策表示赞同意见。其实，社会学是一门严格的科学，要做好并不容易。很多人在方法论上犯错误，他们的研究在方法论上站不住脚。一个好的学者，不但要发表自己的言论和观点，还要知道自己的言论和观点的缺陷和不足在什么地方。比如说现在有一个假设，但是由于他的证据不足，并不能证明他的假设是成立的，因此，研究者就应该意识到自己论点的缺陷，尤其是在方法论上的缺陷。假如不能认识到这点缺陷，研究就可能出问题。

我现在讲一个很简单的例子。假如我有一个研究生，他提出了这样一个假设，说一个人的数学成绩取决于他的语言能力，语言能力强，数学成绩就高。这就是一个假设的理论模型。他有了这个理论模型之后，想做一个验证，他就制定了这样一个研究方案。他去了两个班：一个是数学 101 班，这个班是基础班，有 100 个学生；另外一个班是数学 501 班，是给研究生上的高级班，这是一个小班，只有 10 个学生。这位研究生到这两个班上收集了这样一些数据：他问这两个班上的同学，“你认为你的语言能力怎么样”，“数学成绩怎么样”，“强还是不强”，强就用“1”表示，不强就用“0”表示。在 101 班上，100 个学生中有 90 个学生说他的语言能力强，有 94 个学生说他的数学能力强。在 501 班上，10 个学生中有 8 个说他语言能力强，有 9 个学生说他的数学能力强。这个研究生收集好数据之后，第二天就很高兴地跑到我的办公室来，说他收集的数据证明了他的假设，因为一个班上的同学语言能力强，数学能力也强；另一个班上的学生语言能力低，数学能力也低，这就证明了他的理论假设。假设你是我，你会给他提出什么样的批评？这样的数据存在什么样的问题？我想听听你们的意见。

学生一：1 和 0 是定类变量，不足以客观地描述具体的强弱状态；另外，他是通过一种主观测评来获得数据的，是被测者自己说的，因此数据的客观性不强。

学生二：我认为他在抽样上存在问题，他得出的数据不具有统计上的

面来
上的
会科
三个
在本
多人
见。

意义。

学生三：两种能力的相关可能是虚假的。

学生四：他是通过集体的语言能力和数学能力来判断个人在这两种能力上的相关性。

学生五：样本太小，调查的数量不够。

学生六：他调查的学生是两个不同层次的，他们之间不具有可比性。

谢宇：非常好。这个数据存在很多问题：一是样本规模太小，其中的差异在统计学上未必是显著（significance）的。二是这两种能力并不存在因果关系，这种相关是一种伪相关。三是天花板效应（ceiling effect），这是一个测量的问题，在做“好”与“差”两种选择时，大家都自我感觉良好——这是美国社会的一个普遍现象——这就显示不出其中的差异性。四是选择性偏差，他在收集数据时以班级为单位，而他选择的班级存在级别上的差异，它们之间不具有可比性。五是测度问题，他的数据是由被测者自己报的，学的简单的就自我感觉好，自我感觉好的自然就报好，这是由主观因素主导的，缺乏客观性。六是他的验证存在一个生态学谬误，他的理论假设是以个人为单位的，是说个人的语言能力高会导致数学成绩好，这是对于个人而言的，而他收集的数据和得出的结论是以班级为单位的。在同一班级中有些人可能数学成绩好，但是语言能力差；有些人数学成绩差，但是语言能力强。它们的关系可能是正的，可能是负的，也可能不存在关系。这就是说，当你的数据层次和你的理论层次不吻合的时候，你的数据可能说明不了你的理论，甚至有可能和你的理论是背道而驰的。这就是为什么以团体为单位来研究问题经常会出错误，因为这反映不了个人层次上的现象。

下面，进入今天的主题：社会科学与自然科学的本质区别。我将从三个方面来讲社会科学与自然科学的差异：第一个是讲社会科学与自然科学在本体论上的差异；第二个是讲社会科学与自然科学在认识论上的差异；第三个是讲社会科学与自然科学在方法论上的差异。讲述前面两个差异的目的是为了说明第三个差异，因为我侧重的是对方法论的研究。我只有先把社会科学与自然科学在本体论和认识论上的差异讲清楚，才能讲它们在方法论上的差异。当然有很多人并不同意我的观点，我也讲过这是我自己总结的观点，大家可以各持己见。有的人认为社会科学就是自然科学的延伸，社会科学就要按照自然科学的

模式去做，这个我不赞同。我认为社会科学与自然科学是有根本区别的，这一点稍后我会讲到。

先从本体论上来讲社会科学与自然科学的差异。大家都知道什么是本体论，本体论是研究世界本质的一个哲学分支，研究什么是世界的本质，什么是世界上最重要的东西。换句话说，哲学也好，科学也好，我们到底要研究什么。我本科是学工科的，研究生最开始是学科学史的，当时一个很热的话题就是科学为什么起源于西方而不是起源于东方？古代中国的技术很发达，但为什么没有科学？很多人都考虑过这个问题，提出过见解。我认为这个问题提得并不好。为什么呢？因为科学是偶然产生的，科学的产生受到很多因素的影响；而技术是人人都需要的，要生存、要发展都需要技术。科学和技术在本质上是不一样的，有技术不一定有科学。科学很重要的一个来源就是柏拉图对真正知识的定义。柏拉图对自然哲学提出了很高的要求，他认为真正的知识并不是你对看得见、摸得到的现实社会的认识，而是对抽象的、超现实的理念世界的了解，这样的知识才是真正的知识。因此，他把世界分成两种：本质的世界和形成的世界。柏拉图讲的这两个世界是有很大区别的。我们一般人接触的都是形成的世界，是现实的东西，是具体的东西。而柏拉图说，你要做一个哲学家，就不应该研究这些具体的东西，应该抛开你接触到的东西，提出一种永恒的、放之四海而皆准的真理性的理论。举个例子，我们日常生活中会接触到很多圆，但是柏拉图说，如果你要研究圆，你就不应该研究你所接触到的圆，因为你接触到的只是圆的一种具体形态，并不是真正的圆，不是完美的圆。完美的圆不存在于现实中，只存在于哲学家的思想中。你要研究圆就必须抛弃具体的圆，要想像一个完美的圆，这个圆是没有形状，不占空间的，既画不出来，也得不到。科学就是要理解这个圆，你知道了这个圆的特征，就可以把它推广到任何具体的圆，这就是科学的来源。它不是要你去寻找一种具体的东西，而是要找到一种永恒的、处处适用的真理性的东西。柏拉图讲的形成的世界是什么呢？就是现实的东西，是我们人为的仿制品。所以，如果你要得到知识，就要透过现象看本质，要看到那个抽象的圆。柏拉图的理论是科学产生的一个很重要的基础，追求真理是科学家和科研活动者的一个永恒的动力。而这些东西在中国古代都是没有的，中国古代的东西都是很实用的，都是为了生存而设计的，并没有这种对真理的追求。这是中国古代没有科学的一个很重要的原因。所以，我认为自然科学是以“发掘”本质的世界中的真理为最终目的，这也是其精华所在。而社会科学是以“了解”形成的世界为最终目的。历史上很

多人想
方面的
要去了
社
识论是
道什么
然它追
象。你
到很多
个典型
重点并
社会科
没有太
的了解
一个辩
社
都希望
他们长
影响，
素的景
只能运
来因秀
它有所
所得，
你的发
来因秀
象，至
我的类
和分科
道这个
类
差异。

多人想在社会科学领域找到一种真理，能够适用于各个方面，并且做过许多这方面的尝试。我认为社会科学不应该是这样的。在社会科学中，我们的目的是要去了解现实社会，而不是去挖掘永恒的真理。这可能和你们想像的不一样。

社会科学与自然科学本体论上的区别也导致了它们在认识论上的差异。认识论是探索人类知识的起源、本质、方法及局限的一个哲学分支。你到底能知道什么，你怎样认识世界，这是认识论的命题。自然科学的关注点是什么？既然它追求的是永恒的、真理性的东西，那么它的关注点就在于一些典型的现象。你知道了一个圆就可以把它应用到很多地方，你知道一个分子也可以应用到很多地方，甚至知道一个疾病也可以应用到很多地方。所以，你需要的是一个典型的例子，通过这个典型的例子可以使你了解一个类别的东西。因此你的重点并不在于了解具体的现象，而在于了解典型的现象。而社会科学则不同，社会科学的关注点在于所有个案组成的总体的状况。你了解一个典型的例子并没有太大的用途，你最好能够了解所有的个体，对所有个体的了解就是对总体的了解。你最终要了解总体，但你是通过了解每一个个体来了解总体的。这是一个辩证的关系，我后面会讲到为什么会是这样，怎样才能做到这一点。

社会科学与自然科学在方法论上也有差异。自然科学家在可能的情况下，都希望用实验的方法来证明他的结论。虽然有些情况也是不可能的，但是至少他们是希望用实验的方法来隔离外来因素的影响。比如，你想知道温度变化的影响，在实验时你就可以对压力等其他因素进行控制。通过实验，隔离外来因素的影响，就可以得到一个很纯的规律。而社会科学不一样，我们社会科学家只能运用一定社会环境下的数据（被称为观察数据），而观察数据必然受到外来因素的影响。我们也可以做一些实验，但是大多数情况下这是无法实现的，它有很大的局限性。社会科学之所以复杂，是因为我们运用的数据是通过观察所得，而观察所得的数据必然受到外来因素的影响，这些外来因素都可能解释你的数据。你可以用统计的方法排除一些外来因素，但不可能排除所有的外来因素。我下面讲的因果关系就是讲为什么这些外来因素会给你一种假的现象，致使因果关系不一定成立。所以说我们的数据都有问题。如果一个人说，我的数据非常好，要么他在说假话，要么他根本就没有去收集数据。收集数据和分析数据的人都知道数据来之不易，但还是不能回答所有的问题。你应该知道这个局限性，因为它来自现实生活，受到其他因素的干扰太多。

学过统计的人就知道，统计上有两个东西很重要：一个是平均值，一个是差异。我在这里把柏拉图的本质的世界和形成的世界这两个观念用统计学的东

西来进行对照。这是科学史上发展演变过来的东西。关于平均值，我举一个例子，对于一根一米长的棍子，在测量时就会碰到问题，你让几个人去测量，如果要求数字很精确的话，有几个人测量就会得到几种结果。你需要得到的长度只有一个，而你得到的数据却有一串，这时候怎么办？取平均值。平均值在统计上的理论基础是大数定律和中心极限定理。大数定律就是讲，测量的平均值会随着测量次数的增加而越来越可靠和稳定。它可能永远不能达到真正的值，但它存在这种趋向。所以要得到一个比较准确的平均值，你就要增加测量次数。这是平均值在测量上的一个很好的依据。另外一个就是中心极限定理，随着样本数量的增加，其平均值越来越趋向于呈正态分布（normal distribution）。大数定理和中心极限定理对于平均值而言具有十分重要的意义，给我们的测量提供了帮助。但这两个定理也有条件，就是差异的来源必须是独立的、非系统的微小因素。比如，由你的疲劳、气候等微小因素造成的差异就可以通过取平均值的方法基本得到消除。如果没有这个条件，这两个定理并不适用。这两个定理对社会科学很重要。有一位数学家、天文学家叫 Quetelet，他认为社会科学很简单。柏拉图认为科学是要找到永恒的、真理性的东西。但是柏拉图讲的科学不包括社会科学，因为社会科学研究的是现实的东西，而现实的东西是不断变化的，因此柏拉图的科学中没有社会科学，只有自然科学和哲学。但是 Quetelet 说社会科学中也可以得到永恒的、一成不变的东西。怎么办呢？我们就取平均值。当样本足够大的时候，我们就能很准确地通过平均值来预测一些现象。由此，他提出了社会物理学。他认为要获得永恒性的东西就需要取平均值，这就形成了他提出的普通人的概念。社会物理学是研究普通人的，普通人是永恒不变的，是符合柏拉图对科学的定义的。下面是他的一段话：“我们人类多悲哀啊！我们可以知道每年将有多少人的双手会沾满他们同胞的鲜血，每年有多少人编造谎言，每年有多少人死于毒药……”这些都是可以通过普通人来预测的。因此他就用统计的方法，用取平均值的方法来研究社会学，他的社会学就是取平均值的社会学。自然科学研究的是典型现象，而社会科学的普通人也是一个典型，但这个典型不是一种个别现象，而是一个平均数。这就是 Quetelet 的观点。他这种观点是将柏拉图的传统意义上的科学观念用于社会科学的一个尝试，就是通过统计的方法得到一个不变的东西，这个东西是抽象的，不是具体的。但这种思想在 19 世纪受到很大的冲击。我下面要讲的是达尔文的总体逻辑思维对我们社会科学的影响。

我认为社会科学起源于达尔文，虽然我们今天记住达尔文主要是因为他的

自然选择当
然选择学
或母亲是
差异。同
的家庭环
现象。这
异是世界
个简单的
能光讲平
地怎么样
为造成的
是说你知
此，今天
总体逻辑
对比这两
要求你得
得到所有
达尔
西介绍到
会科学。
值并没有
进冰箱，
刚好。这
Galton 还
现在不
均值，不
忽略了。
有的人
变异，
东西一
高的关
后发现

自然选择学说，但是他对社会科学的影响是非常深远的。请考虑一下，他的自然选择学说是哪里来的？自然选择的基础是差异。假如每一个孩子和他的父亲或母亲是一模一样的，这个社会就不会有变化，更不会有进化。进化的前提是差异。同样的原因会出现不一样的结果，相同的父母会生出不同的孩子，同样的家庭环境会造就孩子不同的性格。这里就有一个随机的因素，这是一种自然现象。这种差异对于达尔文来讲就是世界的本质，而不是人为的差错。既然差异是世界的本质，那差异就应该是研究的对象。我们研究的对象不应该只是一个简单的平均数，虽然这对我们也很重要，但我们不能仅仅局限于此，我们不能光讲平均，而更要讲差异。比如经济上，北京和上海等大城市很发达，但内地怎么样呢？这里就存在一个地域差异。对于达尔文而言，差异并不是一种人为造成的错误，而是一种现实的必然，是精华所在，是值得研究的东西，而不是说你知道了了一个平均值，差异就可以扔掉。这是思想上一个很大的转折。因此，今天我讲的最重要的东西就是差异的重要性。我把达尔文的这种思想叫做总体逻辑思维，而前面讲的柏拉图的思想就叫类型逻辑思维，我在后面会详细对比这两种思维。总体逻辑思维是对科学的传统定义的一个挑战。传统科学是要求你得到一个典型的规律，而达尔文的总体逻辑思维是要你得到一种分布，得到所有由个体组成的一个整体。

达尔文是一个生物学家，他并没有做什么社会科学的研究。是谁把这些东西介绍到社会学来的呢？是他的一个表弟，叫 Galton，将总体逻辑思维引入社会科学。Galton 去过的地方多，见过世面，他认为有的人聪明有的人笨，平均值并没有多大意义。美国有一个笑话：有人问一个统计学家，如果把一只手放进冰箱，一只手放进火炉里，他感觉如何？他说从统计上来讲两者平均一下刚刚好。这就是平均值的毛病，平均值在很多情况下并不能反映真实的现象。Galton 还认为 Quetelet 的社会物理学的用处不大，普通人并不是万能的。我想现在不论是中国社会学还是美国社会学都还保留着这种传统，就是对数据做平均值，不管什么都平均一下。但这只能反映很小的一个方面，还有很多方面被忽略了。Galton 说科学的探索必须关注变异与共变。变异有个体之间的，比如有的人聪明有的人笨，有的人勤奋有的人懒惰等；而且不仅仅是一个变量存在变异，多个变量之间也有变异，我这里把它翻译成共变，就是什么东西和什么东西一起变化。他说这才是重要的东西。Galton 是通过父亲的身高与儿子的身高的关系发现两个变量的相关性的。他把一些父亲与儿子的身高数据画成图表后发现这些点在一个椭圆之内，但这个椭圆不会变成直线，因为直线就意味着

完全相关。这里，相关性的意义就是父亲的特征会影响到儿子的特征，比如父亲聪明，儿子也偏向于聪明，父亲高，儿子也偏向于高，这和达尔文的自然选择也是相对应的。这里除了相关还有一个关系是回归，回归就是说这种相关关系不是决定性的，不是说父亲高，儿子就一定更高。回归就是一种往平均值方向的倒退，比如你父亲很高，但你可能没有你父亲那么高，但这种倒退一般不会倒退到平均值以下。如果倒退到平均值以下，就意味着没有相关关系了。因此相关关系并不是决定性的，而是偶然性地起着非常重要的作用。社会变迁有随机的因素，再好的人家也会出败类，再不好的人家也会出凤凰。所以，相关与回归似乎是矛盾的。因为相关是讲前者对后者的影响，可能是一种因果关系；而回归是把这种关系往回拉，实际上是弱化这种关系。相关和回归是一个问题的两个方面，相关说明可能有某种关系，但这种关系不是决定性的；这是因为回归弱化了这种关系。相关和回归的关键在于差异，没有差异，就不会有相关，也不会有回归。相关和回归是 Galton 发现并命名的。Galton 是一位很优秀的科学家，但由于他叙述不清，致使相关系数（correlation coefficient）的发明权落入他人手中。Galton 发起过一个优生运动，因为通过统计，他发现有的人聪明有的人笨，他认为通过科学的方法可以促进优生，以改进人口的质量。

下面举一个反例，这是一个数据太“完美”的例子。社会科学用的是观察数据，而观察数据是有规律的，如果有人伪造数据的话，我们就可以分辨出来，因为伪造的数据一般可信度比较低。下面是一个真实的例子，讲的是一个英国著名科学家 Burt 爵士（Sir Burt）。Sir 是爵士的意思，这也反映了西方文化和东方文化对科学的不同看法。在中国，书读得好，学问做得好的人要给官，而西方是给荣誉。这个 Sir 就完全是一个荣誉，非常优秀的科学家都会得到这样一个称号。Burt 是一位心理学家，其主要贡献在于对智商代际遗传的研究。为了得到更科学的数据，他增加了样本数量，但是得到的相关系数的精确度却是不变的，不变到第四位小数都是相同的。按统计学规律，随着样本的增加，相关系数的精确度也会随之发生适当变化，而他的数据太“完美”了，真像柏拉图讲的那样永恒不变，这就好得让人难以置信。但是由于当时他是非常有名的科学家，因此没有人提出质疑，直到 20 世纪 80 年代初，才有人提出质疑。现在我们知道他的数据是伪造的，这是科学界一个很大的伪造事件，他发表的文章中的合作作者都是虚构的，他做的访问也是捏造的，得出的数据都是自己编造的。因此我们说任何学科都要有批判性，没有一种批判性的反思，这门学科就很难发展。

总的本质
会物理
的方法
比较多
法还是
有什么
的，其
面会讲
思维应
有差异
用到统
整体下
但是我
析能力
就要运
相似的
和牛顿
简单的
托勒密
我是实
证为自
然的方法
：一
提高，
受到社
自然科
会科学
证主义
主义、
手

总结一下类型逻辑思维和总体逻辑思维的区别。类型逻辑思维源于柏拉图的本质的世界，而这种类型逻辑思维在社会科学中的运用，就是 Quetelet 的社会物理学，用统计的方法来满足柏拉图对科学的定义。再具体而言就是用实验的方法，把外来因素排除，通过控制变量来达到目的。在这方面，心理学做得比较多，经济学也有，最近经济学新生一个分支叫实验经济学，但是实验的方法还是有很多缺陷的，我在最后一讲会讲到实验方法的局限性。总体逻辑思维有什么特点呢？它源于达尔文的进化论。这种总体逻辑思维并不是社会学独有的，其他很多学科也用到这种思维。但在社会学里，这种思维特别重要。我下面会讲到为什么它对社会学特别重要，我有自己的一些看法。把这种总体逻辑思维应用于社会学的主要是 Galton，他提出了变异和共变。社会科学之间虽然有差异，但是他们都会用到达尔文和 Galton 的总体逻辑思维，并且他们都要用到统计的方法。我们一方面要把事实描述得准确、全面，另一方面又要重视整体下面的每一个个体。把整个过程、事实、变量描述得准确是我们的责任，但是我们在做这件事情的过程中又会遇到一个问题，就是我们的思考能力、分析能力是有限的，当样本数量增加时，为了使我们对事实的描述更准确，我们就要运用到统计方法和统计工具。真正的科学是简单的，经得起考验的理论是相似的，越简单的越可信。很多有名的科学家都信教，最典型的例子是开普勒和牛顿，他们都相信上帝。他们认为上帝一定是一位很聪明的数学家，因此越简单的东西越可信。哥白尼的日心说之所以能够说服后人，是因为他的解释比托勒密的地心说简单很多，省去了很多繁琐的公式。

我是做统计的，而统计是要求实证的，因此很多不了解我的同事就认为我是实证主义者。实际上，我不是实证主义者。实证主义是一种哲学观点，它认为自然科学与社会科学没有本质的区别，换言之，社会科学应该照着自然科学的方法去做，而且社会科学可以做得和自然科学一样严格。这里就有两种提法：一种是说社会科学还不够完善，还不够科学化，需要通过实证的方法把它提高，使它向自然科学靠近；另外一种是要把自然科学降低，因为自然科学也受到社会的影响，受到阶级的影响，在这方面它和社会科学一样，因此可以把自然科学降低，从而向社会科学靠拢。这两种提法在本质上是一样的，就是社会科学和自然科学没有本质的差异。但是我不同意这种观点，所以认为我是实证主义者对我而言是很不公平的。实证主义受到很多方面的挑战，包括马克思主义、后现代主义等方面的挑战和冲击。

我虽然不属于实证主义者，但是我承认我的思想属于一个流派，这个流派

就是人口学流派。这个流派在美国社会学界占很大的比重，有很大的影响力，它是以 Duncan 为代表的。人口学是一个包容性很强的综合性学科，它和社会学、人类学、经济学、心理学都是有关系的。中国的人口学太闭塞，还没有发展为一个综合性的学科。Duncan 对人口学的定义是：人口学是对人口数量、人口构成及其变化的研究。这个定义是很广的，人口数量就包括出生、死亡和迁移，而人口结构则包括家庭、学校、教育、职业等等。所以 Duncan 对人口学的定义囊括了很多内容。很多人问我是研究人口学的还是研究社会学的，我说我既是研究人口学的，也是研究社会学的，我是用人口学的观点研究人口学和社会学的。我还给人口学流派取了一个名字，叫做实用经验主义（pragmatic empiricism）。这里的“经验”是以你从现实世界收集的资料为基础，而不是你的主观经验。“实用”是说你收集的资料可能使你的想法很有说服力，但是它并不是十全十美的，只是它对你的研究有可用性。我要提到的是 Lieberman^①，一位很优秀的社会科学家。他讲道：“社会学的重要贡献之一，就在于它提供信息的能力。这些信息包容了社会普遍关心的内容，有关种族平等、强奸、贫困、无家可归者、代际流动的统计。我选择这些问题是因为它们的答案提供了有关社会的有用信息，即使它们有时候并不一定为某种理论服务。”社会学的重要贡献是它能够提供信息，社会学在这方面的能力比其他学科要强，比如说人类学、经济学、心理学、政治学。因此，在提供信息方面，我们的责任更大。在美国，几乎所有对社会信息的描述都是社会学家提供的。对于家庭的组成、家庭的破裂、生育、小孩教育和就业等比较概括性的、描述性的东西，都是我们提供的。因为其他的学科往往重视了一方面而忽视了全面的东西。美国社会的发展趋势是怎么样，家庭是怎么样，父母离婚对小孩的影响怎么样，贫富差距怎么样，上代人对下代人的职业影响怎么样，犯罪率是增加还是减少，离婚率是增加还是减少，这些东西都是由社会学家来做，所以我们的学科和其他学科是有很大差别的。社会学家的使命非常伟大。有时我们讲社会学家就是历史学家，只不过我们描述的历史发生在现在，而不是发生在过去。我们是把现在发生的过程和现象系统地表述出来。

Lieberman 的这段话其实是对纯理论学家提出的批评，他批评理论学家常常忽略了经验性的结果，空谈理论而不重视实际。理论是知识的全部内容，理论应该超越事实，但它必须包含已知的信息。一个忽略了已知事实的理论，只

能是一
理论是
空谈理
本关于
情况。
社会参
厄（Pi
缺乏论
理论要
果，最
种人，

① Stanley Lieberman 现系哈佛大学（Harvard University）社会学系教授。

能是一个片面的、错误的理论。他认为爱因斯坦的理论是好的理论，因为他的理论是尊重过去的经验事实的，而我们一些社会学家却不去了解社会事实，而空谈理论。他批判了三个例子：一个是吉登斯（Anthony Giddens），他写了一本关于美国社会分层的书，书里对美国分层的描述根本不符合美国社会的真实情况。第二个是研究文化社会学的一本书，书名为《心灵的习性》，该书认为社会参与性在降低，但这个结论是建立在猜想的基础上。另外一个就是布尔迪厄（Pierre Bourdieu），他提出一个品位阶级差异理论。Lieberson说他们的理论缺乏论证的严密性，我赞同 Lieberson 的观点。我讲的总体逻辑思维就是强调理论要建立在事实的基础上。做社会学有两种顺序：一种是先写方法，再写结果，最后再写结论；还有一种是先写结论，再写方法，最后写开头。我是第一种人，习惯性地先摆出我使用的方法，然后描述结果，最后总结得出结论。