统计的核心是“知行”，知是“认识世界”，行是“改造世界”。知行合一是统计的最高境界。

认识世界，难吗？先不要着急回答这个问题，一起来看看我提出的几个问题。

* 吸烟导致肺癌吗？
* 明天太阳升起的概率是多少？
* 某中药号称治疗Covid-19治愈率高达95%，你相信吗？
* 某春运调查研究报道称“经过在火车站内实地调查显示，春运期间买票虽然困难但大家都买到票最终乘车回家了”，如何看待这一调查？

解答上面几个问题，需要用到概率和统计知识。

* 解读1：怎么判断是否吸烟导致肺癌？反对吸烟的人说，吸烟导致肺癌，你看吸烟的人比不吸烟的人得肺癌的人数要多很多。爱吸烟的人说，你看好多人不吸烟还得了肺癌，好多吸烟得人没得肺癌，所以不能说吸烟导致肺癌。英国著名流行病学家Hill曾经做了病例对照试验，结果显示病例组吸烟人和不吸烟人得比值比非病例组得比值要高，表明吸烟导致肺癌。医学界也开始接受这个“吸烟导致肺癌得观点“，但是由于没有严格得临床试验证据，统计学鼻祖费舍尔持反对观点，他表示”怎么证明不是肺癌导致吸烟的呢，或许是因为得肺癌的人存在未被检测诊断到的chronic inflammation，这种chronic inflammation导致得肺癌的人比不得肺癌的人通过吸烟缓解疼痛“，费舍尔进一步怀疑，如果不是肺癌导致的吸烟，那很可能是一种遗传因素导致既爱吸烟又容易得肺癌。费舍尔坚持随机对照试验才能给出因果证据，而病例对照试验因为无法排除混杂因素干扰只能给出相关证据。而吸烟导致肺癌很难做试验，从伦理层面没有办法让一组人不抽烟，另外一组人抽烟，观察几年对比肺癌发病率。当然，后来科学的进步从生物层面证实了吸烟导致肺癌。但是，如何解读病例对照试验结果，如何从现有的研究和数据中解读并且审视科研结论，需要统计学知识。
* 解读2：答案有三个：1，1/2和(n+1)/(n+2)。如果用过去的经验，那明天太阳升起的概率是1；假如不知道过去太阳升起的情况，对明天不确定是否升起，那么概率为1/2；如果用贝叶斯概率知识，结合过去的经验和对明天的不确定，那么概率为(n+1)/(n+2)。此处，n是过去的天数。到底哪个概率才是明天太阳升起的概率，这涉及到贝叶斯学派，古典概率学派和频率概率学派的区别。
* 解读3：中药这么厉害？那它怎么做到的治愈率高达95%。有做对照吗？有随机双盲吗？是骡子是马拿出来溜溜，没有对比就不知道它是不是真的厉害。对照是做了随机双盲吗？如果没有随机双盲，把一匹马放在坑坑洼洼的跑道，把骡子放到平滑的跑道，那这种对比存在不公平。如何识别这个中药的效果是不是真的如所说的那么好，需要统计知识去解读宣称的结果是通过什么证据得到的，证据的可信度有多高。不管什么药，好不好不是专家说了算，不是嘴上说了算，不是文章说了算，是设计良好的试验说了算，试验设计用到统计知识。解读试验结果，需要用到统计知识。
* 解读4：很明显这个调查研究有选择偏倚，开展调查研究要选择具有代表性的人群。本调查研究旨在研究春运购买火车票人群购买难度，在火车站内选择调查对象显然无法代表研究人群，因为进火车站的人肯定是买到票的，而没买到火车票的人进不来火车站。所以，这个研究没包含没买到火车票的人，所以存在严重偏倚。如何控制偏倚，如何选择具有代表性的人群避免偏倚，是统计学的重要理念。

解读完上面四个问题后，相信读者对统计学如何帮助人类认识世界有了初步认识。如果没有统计知识和知识储备，上面几个问题无法给出理性的回答。统计学帮助人类认识世界的用途还有很多，生活中也处处可见这类例子。比如，大家可能都有过类似的经历，民间某偏方可以治疗癌症，用过这个偏方的人都好了。没治好的人以及他的家人不会说这个偏方不好，大家听到的都是关于治好的消息。这其中存在着严重的选择性偏倚，统计术语称之为“非随机缺失“。2018年全国高考卷二的语文作文题就曾考过这个话题。

改造世界，难吗？先不要着急回答这个问题，一起来看看我提出的几个问题。

* 怎样证明某药物治疗Covid-19有效？
* 怎样预测房价？
* 海量数据中怎样提取少量的关键信息？
* 怎样做好产品的质量控制？

解答上面几个问题，需要用到概率和统计知识。

* 解读1：需要随机对照双盲试验验证某药物治疗Covid-19的安全有效性。
* 解读2：从网上爬虫房价数据，构建统计模型预测房价；为了避免过拟合效应，把数据切割成两部分，一部分用来训练模型，一部分用来测试模型效果。未来的数据，再用模型去验证效果如何，再根据未来数据不断调整模型。
* 解读3：我最近在购买婴儿车，在网上搜集了婴儿车的选择标准大概有30个列了出来，大致可以分为几类（轻便灵巧，安全稳定，舒适，美观，尺寸，其他）。如果我们面对的是海量信息，假如变量信息有成百上千个，怎么提取汇总关键的信息，此时需要统计降维方法，譬如主成分分析和因子分析等。
* 解读4：产品的质量控制需要用到统计抽样等方法。

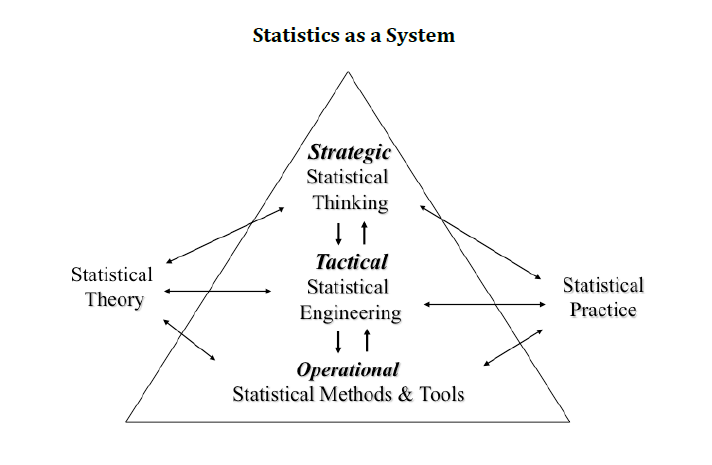
知道怎么用统计知识解读日常生活和工作中碰到的一些现象，并且能用统计方法去指导生活和工作，做到知行合一，才能达到统计的最高境界。如果知道临床试验应该用随机对照的方法，但是再具体操作时却投机取巧，那这种情况属于统计操守有问题。知性合一是统计的最高境界，那什么是统计的最低境界呢，中间有没有其他境界？

有篇微信公众号文章《你们看到的都是招数，不是内功》里直言：鸠摩智上少林寺挑衅，使遍七十二绝技，方丈群僧无不骇然。这时，小和尚虚竹跑过来，只瞅了一眼，就说：“这位大师用的明明是小无相功嘛。”鸠摩智慌了。**一般人看到的是招数，厉害的人看到的是内功。你在一个行当里面学到的都是招数，是行业规矩，它保证你有资格进入这个领域。但是，在这个领域你能达到多深的造诣，这是内功，内功的训练，可能往往要超出这个领域之外。**所以，少林寺诸高僧抱着七十二绝技的招数啃，一人啃一门绝技，啃到头最多一人身兼十三门绝技，根本无法和番僧鸠摩智相抗衡。要身兼七十二门绝技，那就要用到内功心法，内功心法在哪呢？不在武学书里，在佛经里。

很多统计应用者（社会科学、医学、心理学等领域研究人员）看到的是统计的招数，一个问题对应多种统计方法。用到统计的时候，打开软件，输入数据，选择方法，然后软件就给自动产生结果了，如果结果是阳性就开始写文章了，结果是阴性就试试其他方法，直到结果呈现阳性也就是P值小于0.05为止。一切看起来都是这么自然，这么合理，统计不就是为了分析出阳性结果的吗？这就是统计的最低境界，学会了五花八门的招数，却不懂得统计的心法。

统计的心法是什么？1.从嘈杂的世界中找出有用的信号；2. 给不确定赋予概率；3. 量化变异性大小；4. 熟知不同统计方法的优缺（等等）。世界是随机的，不确定的，存在变异的，所以需要统计。如果世界是固定的，确定的，同质的，那统计就没有用武之地了。正是因为世界是随机不确定存在变异的，所以很多现象是随机的，是偶然的现象，有了统计的工具，正确使用统计工具，才能判断什么现象是嘈杂的偶然现象，什么现象是主要发现的信号。譬如做临床试验，统计不是用来证明药物的临床效果，而是用来从概率上保证药物的临床效果不是随机偶然现象而是重要的信号。掌握了统计心法，才能准确的把现实问题转化为统计问题，然后通过设计控制噪音，搜集到相应的数据，在统计百宝箱中选择出合适的分析方法，最后给出合理的解释。

统计心法到统计百宝箱工具之间还有一层，统计工程。如下图所示，统计心法（统计思维）是战略层，统计工程是战术层，统计工具和方法是实施层。什么是统计工程，我觉得下面的解释非常好：The study of how to best utilize statistical concepts, methods, and tools,and integrate them with information technology and other relevant disciplines,to achieve enhanced results. 统计工程这个词是美国通用电气的统计学家Roger W. Hoerl和Snee Associates公司的统计学家Ronald D. Snee一起提出来的。平常我们常常听到统计科学，统计思维和统计理论，但是很少听到统计工程。统计工程思维很重要，它是衔接统计思维和统计诸多方法之间的利器。正如作者所言：Engineers—statistical or otherwise—do not focus on advancement of the fundamental laws of science, but rather how these laws might be best used for practical benefit. This is not to say engineers do not research or develop theory. Rather, it suggests engineers’ theoretical developments tend to be oriented toward the question of how to best use known science to benefit society. We’re not suggesting that society no longer needs research in new statistical techniques for improvement; it does. The balance needed at this time, however, is perhaps 80% for statistics as an engineering discipline and 20% for statistics as a pure science.可见，统计科学是不断的创造出更多的统计方法，统计理论是为这些方法奠定理论基础，统计思维是创造性的用好统计方法贯彻好统计思维，系统化的解决现实问题。



我认为，再上面还需要加一层，是统计哲学层面，这一层面考虑统计的逻辑基础问题，譬如统计有几大互相不兼容的学派：贝叶斯学派，频率学派，费舍尔似然学派，这几个学派之间的关系就好像是物理学里的经典力学学派和量子力学学派之间的关系，代表了不同的范式。在不同的范式下，方法有不同的解释。去年美国统计哲学家出版了一本《超越统计之争》的书，里面详尽分析了统计史上的学派之争，理清了这些争论的焦点在哪儿，各自学派的优缺。

**统计五境界**

|  |  |  |
| --- | --- | --- |
|  | **统计境界** | **具体解释** |
| **第一境界：实施层** | 统计方法 | 五花八门的招数 |
| **第二境界：战术层** | 统计工程 | 理解统计心法，明悉各种统计方法的优缺, 在统计心法和统计方法之间搭建桥梁，系统科学的使用招数， 解决实际问题 |
| **第三境界：战略层** | 统计心法 | 为统计方法和统计设计提供理论基础。如果没有统计心法，统计就变成了没有根基的玩弄数据的杂技。 |
| **第四境界：哲学层** | 统计哲学 | 理解不同统计学派和范式之间的区别和联系；  追求百花齐放，百家争鸣，求同存异，共同繁荣 |
| **第五境界：知行层** | 知行合一 | 认识世界，改造世界 |

统计从根本上说是为了通过合理的设计旨在从数据中辨别出有价值的信息从而认识世界和改造世界。无论是贝叶斯还是皮尔逊、费舍尔、尼曼等统计学家，都希望通过统计方法从归纳的角度认识世界，改造世界，只是他们理解“认识”的角度不一样，由此产生了不同的统计学派。站在知行合一的角度，从哲学层面理解不同学派的分歧，求同存异，深入掌握统计心法，采用统计工程思维，系统科学的应用统计方法，最终更好的认识世界，改造世界，做到真正的知行合一。

**参考文献：**

1. <https://priceonomics.com/why-the-father-of-modern-statistics-didnt-believe/>
2. <http://bff-stat.org/>
3. Hoerl,R.W.and Snee,R.D.(2010a)“Moving the Statistics Profession Forward to the Next Level,”The American Statistician,64,1,10-14.
4. Hoerl, R.W. and Snee, R.D. (2010b) “Closing the Gap; Statistical Engineering Can

Bridge Statistical Thinking With Methods and Tools,” Quality Progress, May,52-53.

1. Hoerl, R.W. and Snee, R.D. (2017) “Statistical Engineering: An Idea Whose Time Has Come,” The American Statistician, 71, 209-219.