你在何种程度上同意“所有的模型都是错误的，但有些是有用的”这一说法（认为是出自乔

治·博克斯）？请参考数学和另外一个和另外一个知识领域展开你的应答。

Outline:

在某种程度上/不/完全程度上，我同意这个说法。

模型：为了预测，模拟，帮助。模型的重点不是找到确切的事实，而是“最有可能”的事实。

（乔治·博克斯（George Box）原话语境：他的观点是，我们应该关注某件事是否能以有用的方式应用于日常生活，而不是无休止地争论一个答案是否在所有情况下都是正确的。）

统计意味的model意味着什么 统计意义上的模型 非统计意义上的模型 意味着什么思想的产生

Utility: Yuval Noah Harari – “Scientists generally agree that no theory is 100 percent correct. Thus, the real test of knowledge is not truth, but utility. Science gives us power. The more useful that power, the better the science.”

Something is useful -> something is good.

其他不是统计意义上的model也可以

Wrong 有假设

Useful 假设范围内

Math:

Example 1: Useful: Paul Lazarsfeld在20世纪40年代使用的线性回归模型。他分析了1940年美国总统选举中社会阶层和投票行为之间的关系

他开发了一个线性回归模型，根据社会阶层和其他人口变量来预测选民的行为。他假设这些变量与投票给特定候选人的可能性之间存在线性关系。然而，这个模型有明显的局限性：

* 过度简化：该模型过度简化了人类行为的复杂本质，假设投票偏好可以纯粹通过线性关系来解释。
* 因果关系：拉扎斯菲尔德的模型暗示了社会阶层和投票行为之间的直接因果关系，而在现实中，许多其他因素（如种族、教育和地区影响）也起着至关重要的作用。

Useful：引入了“多元分析”的概念：Lazarsfeld的工作强调了同时考虑多个变量的重要性，为更复杂和准确的统计方法铺平了道路。影响了政治民意调查和竞选策略：尽管该模型本身存在缺陷，但从中获得的见解帮助政治分析家和竞选者理解了不同人口因素在选举中的重要性。

Example 2: Useless: 20世纪初，包括统计学家罗纳德·a·费雪爵士在内的一些研究人员参与了吸烟对健康影响的研究。费雪的统计模型表明吸烟和肺癌之间存在微弱的相关性。这些模型主要依赖于观察性研究，不能有效地控制诸如遗传易感性和其他生活方式选择等混杂因素。（混淆变量：模型没有考虑到重要的混淆变量。例如，他们没有充分考虑空气污染、职业暴露和个人遗传易感性等因素，这些因素可以独立于吸烟影响健康结果。）

到20世纪中期，随着更全面的研究的出现（比如理查德·多尔和奥斯汀·布拉德福德·希尔进行的具有里程碑意义的英国医生研究），很明显，费雪的模型是不充分的。这些早期模型中使用的统计方法误导关于吸烟对健康的影响，他们提供了一种虚假的安全感。有缺陷的结论阻碍了旨在防治与吸烟有关疾病的公共卫生倡议。

人文科学

Example 1: Useful: Black-Scholes Model 布莱克-斯科尔斯模型假设，股票或期货合约等工具的价格将遵循对数正态分布，并伴随着不断的漂移和波动。该方程利用这一假设和其他重要变量的因素推导出欧式看涨期权的价格。然而其并不包含所有的情况，只能为欧洲期权定价。它没有考虑到美国期权可以在到期日之前行使。其他假设也包括：没有交易成本或税收，所有期限的无风险利率都是恒定的，允许使用收益进行证券卖空，以及没有风险就没有套利机会。这些假设中的每一个都可能导致价格偏离实际结果… 即使在一定程度上该模型是“错误”的，它仍然为金融界创造了估计的可能性。

只有正确才有用

当一些基于的假设被误解和外推时，问题就出现了

Example 2: Useless: 经济理论认为，在一个成熟的经济体中，放松管制和私有化会产生力量，但在一个缺乏强大法律基础的国家引入这些政策，可能会导致由少数人统治的盗贼统治。从本质上讲，经济模型为决策提供了强大的基础，但我们必须意识到现实世界的假设、模型的错误规范以及动态的社会经济和政治环境。

<https://www.imperial.ac.uk/business-school/news/all-models-are-wrong-some-are-useful/>