



TEXT HERE

决策 or 判断
or 数学概率

料万事于
此数中

PB20070312





什么是概率

What is probability?



生活中有哪些概率

Where does probability live?



概率的作用

Why do we need probability?



贝叶斯原理及其案例分析

More concrete examples

目录

PB20010312

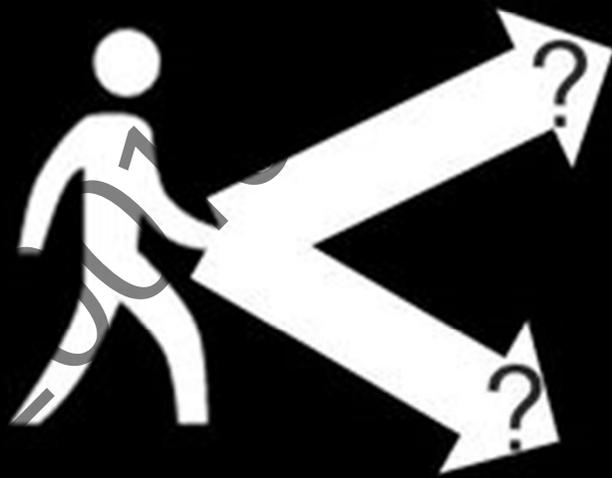
1

什么是概率

PART ONE



如何考虑到不同的**可能性**
并做出最好的选择？



- 腾讯股票明天上涨的可能性
- 明天腾讯股票上涨的可能性
- 后天下雨的可能性很小

蜀山区

中国天气 上次更新时间: 08:54

| 09:00 | 10:00 | 11:00 | 12:00 | 13:00 | 14:00 |
|-------|-------|-------|-------|-------|-------|
| | | | | | |
| 27°C | 29°C | 31°C | 31°C | 31°C | 31°C |

| | | |
|---------|--|-------------|
| 6月1日昨天 | | 32°C / 21°C |
| 6月2日今天 | | 31°C / 22°C |
| 6月3日明天 | | 28°C / 17°C |
| 6月4日星期五 | | 30°C / 17°C |
| 6月5日星期六 | | 34°C / 17°C |
| 6月6日星期日 | | 34°C / 22°C |
| 6月7日星期一 | | 35°C / 23°C |
| 6月8日星期二 | | 35°C / 22°C |



PB200705

概率就是这条直线上的某个数值



不可能

不太可能



可能性相等

很可能



必然发生



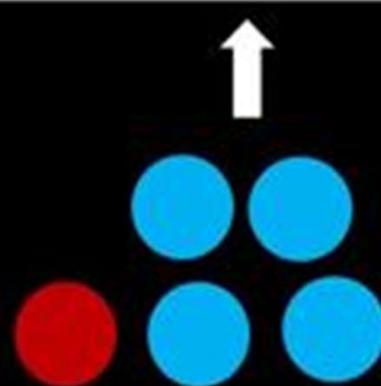
0



1



1/6的机会



4/5的机会

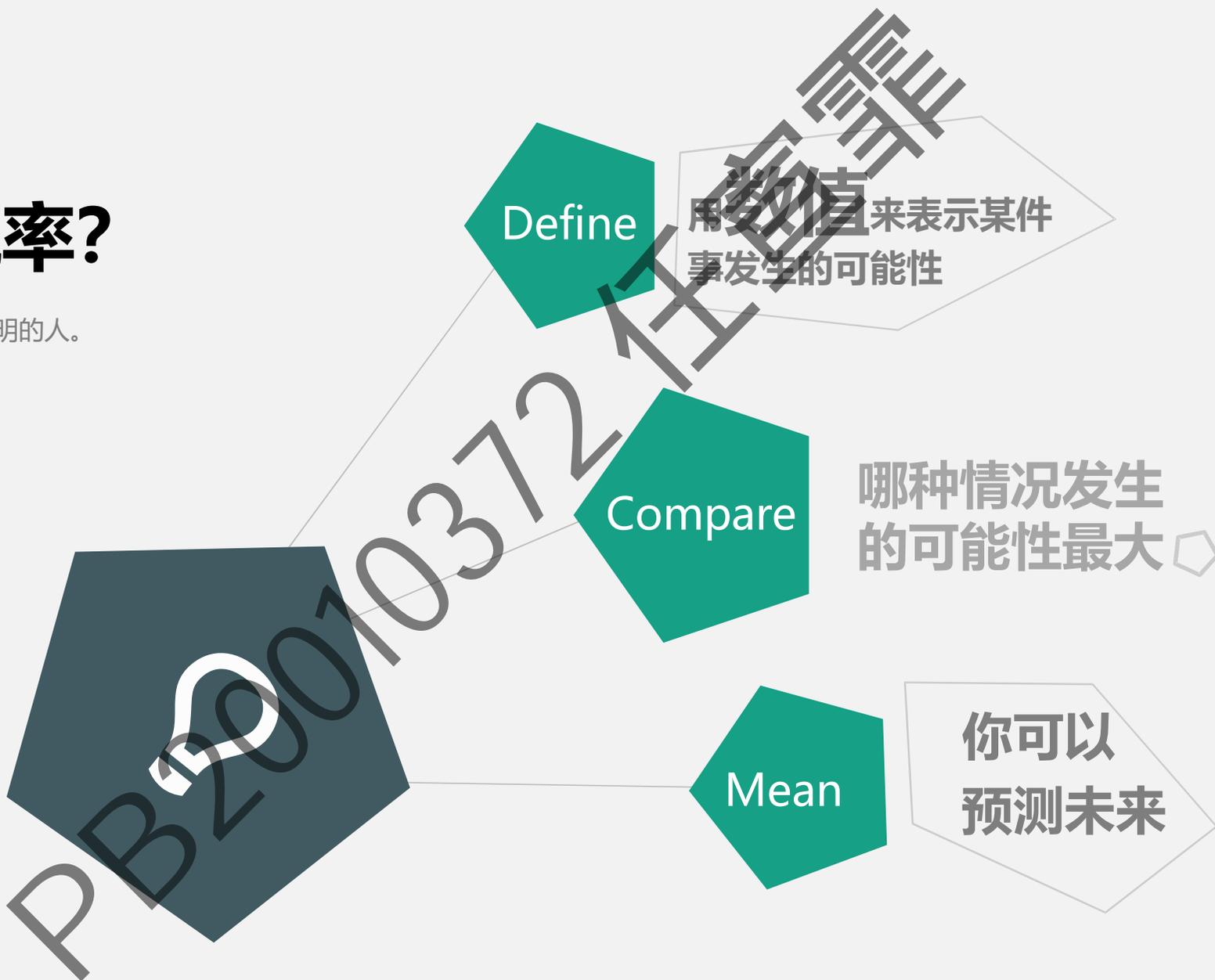
“这个粒子衰变的概率是50%；”
“掷硬币正面朝上的概率是50%；”
“这个手术成功的概率是60%；”
“Trump被弹劾的概率是5%；”





什么是概率?

懂得概率的人，才是真正聪明的人。



2

生活中有

哪些概率



PB200703
任意選擇

身边的例子

02

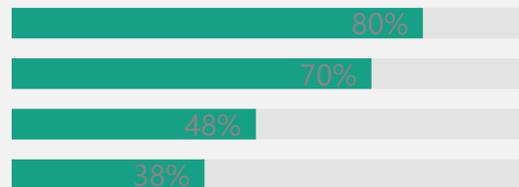
旁边的那队
永远比你更
快？

01

7:30醒来以
后，到不同餐
厅吃饭耽误早
课的可能性

03

哪支球队
更有可能
夺冠？



概率对你我

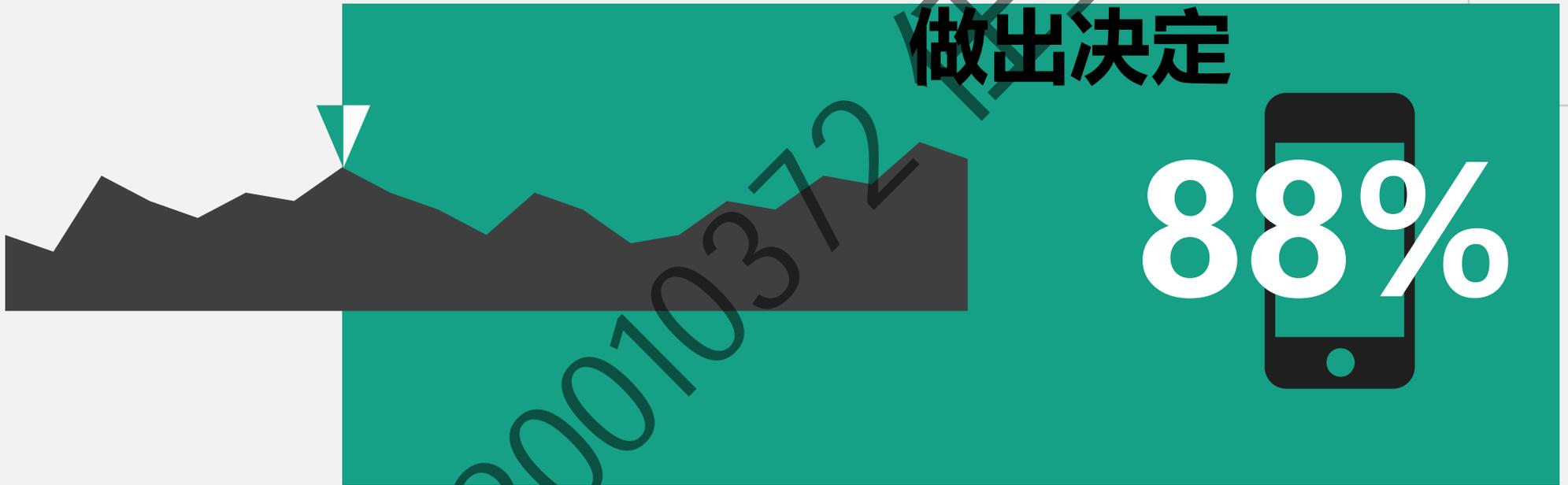
3

有什么用?



应用

**决策者往往能够看清风险，
做出决定**



■ 飞机安全吗？

- 自20世纪60年代起就再没有发生过一起商业航空致死事故。
- 商业航空旅行的风险是微乎其微的。
- 航空旅行每一亿公里的死亡率基本为0
- 汽车每一亿公里旅行的死亡率为0.5
- 而摩托车高出35倍！

其它死亡 VS 空难

相对于其它死亡，空难绝对属于小概率事件，不信瞧瞧：



热水烫死的概率是空难的2.2倍



毒蛇咬死的概率是空难的0.14倍



食物毒死的概率是空难的3.66倍



坠梯摔死的概率是空难的4.78倍



掉床摔死的概率是空难的5.5倍



洗澡呛死的概率是空难的16倍



火车撞死的概率是空难的22倍



因公殉职的概率是空难的252倍



车祸死亡的概率是空难的1375倍



死于癌症的概率是空难的220万倍



中风、心脏病突发的死亡概率是空难的440万倍

■ 飞机安全吗？

康奈尔大学的3位研究人员在一篇引人入胜的论文中这样说道，可能有数以千计的美国人在“9·11”恐怖袭击事件发生之后由于害怕坐飞机而死于非命。我们永远都不知道遭受恐怖袭击的真正风险到底有多大，但我们知道开车确实是一件危险的事。

在“9·11”恐怖袭击事件发生之后，越来越多的美国人选择自驾出行，而不选择乘坐飞机。

据统计，在考虑平均死亡率和天气等导致路面交通事故因素的前提下，2001年10~12月，平均每个月因交通事故致死的人数比以往多了344人。该效应随着时间的推移逐渐减弱，这是因为大家对恐怖主义的恐惧在慢慢消退。但这项研究的作者认为，“9·11”恐怖袭击事件导致的驾车死亡人数或已超过2000人。



不懂概率的结果

虽说上述提到的恐怖袭击和意外导致的是不同原因的空难，但仅仅因为“感觉”飞机更危险去选择其他方式而忽视概率，是不理智的。

在不懂概率的情况下，选择交通方式都会导致死亡率的提升。

虽然概率并不会确凿地告诉我们将会发生什么，但我们通过概率能够知道很有可能发生什么，不太可能发生什么。聪明的人会使用这些数据为自己的事业和生活指明方向。



知其所以然

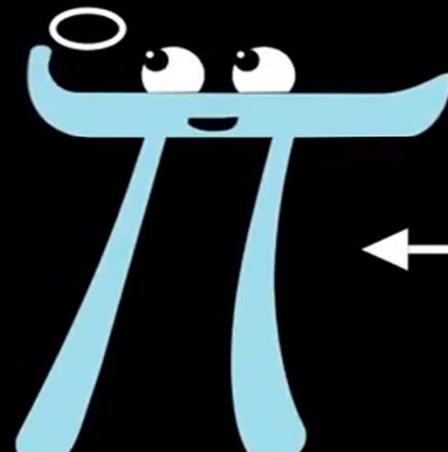
4

贝叶斯原理



PB200703
任宣

$$P(H|E) = \frac{P(H) \cdot P(E|H)}{P(E)}$$



← you

具体的例子

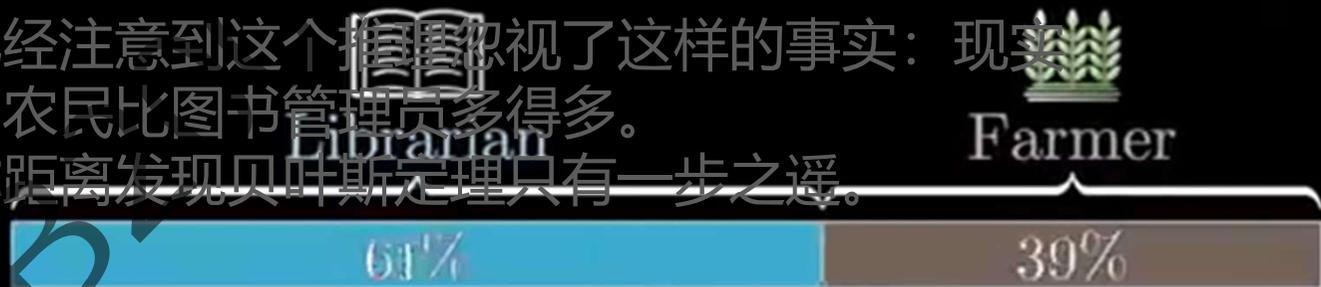
你在路上遇到一个人，他叫史蒂夫。谈吐优雅，做事有条理，那么他是图书管理员的可能性大，还是他是农民的可能性大？

Steve is very shy and withdrawn, invariably helpful but with very little interest in people or in the world of reality. A meek and tidy soul, he has a need for order and structure, and a passion for detail.

你的脑海里可能出现了这样的推导：把上述特性（谈吐优雅，做事有条理），记作A，你觉得大概有40%的图书管理员是这样的，而只有10%的农民这样，所以这个人更可能是图书管理员。



可能你已经注意到这个推理忽视了这样的事实：现实生活中，农民比图书管理员多得多。那么，你距离发现贝叶斯定理只有一步之遥。



???

具体的例子

那么我们就来具体看一看究竟哪一种情况发生的可能性更大些。
假定图书管理员中有40%符合特性A，农民中只有10%符合这一特性。

就总量来说，图书管理员与农民之比为1：20。



具体的例子

任取一个人，E表示他满足特性A的概率，H表示他是图书管理员的概率，那么前面遇到的人是图书管理员的概率仅仅为

$$P(H|E) = \frac{P(H)P(E|H)}{P(E)} = \frac{\frac{1}{21} \times \frac{4}{10}}{\frac{24}{210}} = \frac{4}{24} \approx 16.7\%$$

与人们最初的判断大相径庭！

贝叶斯原理

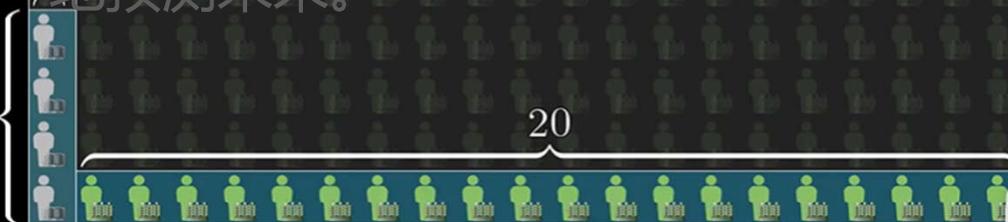
其中P(H|E)称为后验概率，P(H)称为先验概率。正是因为人的直觉上对先验概率的忽视，导致了直觉上的结果与实际结果的较大偏差。

$$P(H|E) = \frac{P(H) \cdot P(E|H)}{P(E)}$$

而恰恰是对概率的研究，使人拥有了超乎直觉的预判，仿佛拥有了“神的力量”，可以较为准确地预测未来。

40%

20



应用:以医疗行业为例



99%=1?

案例分析



1%=0?

尽管你觉得它有99%可能为真，但事实上它很可能为假。

PB200703

贝叶斯定理在医疗行业的应用

每一个医学检测，都存在假阳性率和假阴性率。假阳性，就是没病，但是检测结果显示有病。假阴性正好相反，有病但是检测结果正常。即使检测准确率是99%，如果医生完全依赖检测结果，也会误诊。也就是说假阳性的情况，根据检测结果显示有病，但是你实际并没有得病。举个更具体的例子，因为艾滋病潜伏期很长，所以即便感染了也可能在很长的一段时间，身体没有任何感觉，所以艾滋病检测的假阳性会导致被测人非常大的心理压力。

**你可能会觉得，检测准确率都99%了，误测几乎可以忽略不计了吧？
所以你觉得这人肯定患艾滋病了对不对？**

贝叶斯定理在医疗行业的应用

让我们用贝叶斯定理算一下，就会发现你的直觉是错误的。

可以建立如下的数学模型：

假设某种疾病的发病率是0.001，即1000人中会有1个人得病。现在有一种试剂可以检验患者是否得病，它的准确率是0.99，即在患者确实得病的情况下，它有99%的可能呈现阳性。它的误报率是5%，即在患者没有得病的情况下，它有5%的可能呈现阳性。

现在有一个病人的检验结果为阳性，请问他确实得病的可能性有多大？

设阳性事件为E，得病事件为H：

$$P(H|E) = \frac{P(H)P(E|H)}{P(E)} = \frac{\frac{1}{1000} \times 99\%}{0.001 \times 99\% + 5\% \times 0.999} \approx 1.94\%$$

我们得到了一个惊人的结果， $P(H|E)$ 等于1.94%。
也就是说，筛查的准确率都到了99%了，通过体检结果有病（阳性）确实得病的概率也只有1.94%

Bayes' theorem

$$P(H|E) = \frac{P(H)P(E|H)}{P(E)} = \frac{P(H)P(E|H)}{P(H)P(E|H) + P(\neg H)P(E|\neg H)}$$

贝叶斯定理在医疗行业的应用

你可能会说，再也不相信那些吹得天花乱坠的技术了，说好了筛查准确率那么高，结果筛查的结果对于确诊疾病一点用都没有，这还要医学技术干什么？

没错，这就是贝叶斯分析告诉我们的。

当我们对**一大群人**做艾滋病筛查时，虽说准确率有99%，但仍然会有相当一部分人因为误测而被诊断为艾滋病，这一部分人在人群中的数目甚至比真正艾滋病患者的数目还要高。

造成这么不靠谱的误诊的原因，是**无差别地给一大群人**做筛查，而不论测量准确率有多高，因为正常人的数目远大于实际的患者，所以误测造成的干扰就非常大了。

贝叶斯定理在医疗行业的应用

根据贝叶斯定理，我们知道提高先验概率，可以有效地提高后验概率。

所以解决的办法倒也很简单，就是先锁定可疑的人群，比如10000人中检查出现问题的那10个人，再独立重复检测一次。

因为正常人连续两次体检都出现误测的概率极低，这时筛选出真正患者的准确率就很高了

这也是为什么艾滋病检测第一次呈阳性的人，还需要做第二次检测，第二次依然是阳性的还需要送交国家实验室做第三次检测。

贝叶斯定理在医疗行业的应用

在《医学的真相》这本书里举了个例子，假设检测艾滋病毒，对于每一个呈阳性的检测结果，只有50%的概率能证明这位患者确实感染了病毒。但是如果医生具备先验知识，先筛选出一些高风险的病人，然后再让这些病人进行艾滋病检查，检查的准确率就能提升到95%。

概率影响生活，改变生活，从来都不是一句空话。



研讨心得

未来的我们，
步履不停，用数学去实现曾
经难以想象的价值

参考文献

✓
《概率
(Probability)
的本质是什么?》

——猴子 (来源知乎)

✓
《贝叶斯定理, 使概率论直觉化》

——3Blue1Brown

✓
《心理研究引入贝叶斯统计推断的必要性_应用思路与领域》

——钟建军

✓
《概率统计数学模型在投资决策中的应用》

——冯燕茹

✓
《优化模型在数学建模中的应用》

——杨依晨



THANKS

PB 20070312 任道賢

