```
begin
    using PlutoUI   , LaTeXStrings   , Latexify   ,Distributions   ,StatsPlots
end
```

贝叶斯定理

给定任意两个事件 A, B, 贝叶斯定理表示为:

$$P(A|B) = \frac{P(A) * P(B|A)}{P(B)}$$

公式符号表示为:

- $P(A|B)$: 当事件B 发生时,事件A 发生的概率
- $P(B|A)$: 当事件A 发生时,事件B 发生的概率
- $P(A)$: 事件A 发生的概率.
- $P(B)$: 事件B 发生的概率

其中$P(A)$ 类型表示事件发生的概率. $P(A|B)$ 表示B 事件发生时A 事件发生的概率. | 的符号是和经典统计学描述不同的地方. $P(A|B)$ 称为条件概率

在上面的的贝叶斯概率公式中, $P(A)$ 事件发生的概率的计算由于增加了B 事件发生的信息, 所以相应事件概率也发生了变化. 这完全就是我们生物学习新知识的过程:随着学习的深入, 了解的信息越多, 对某个知识的理解也会加深. 所以贝叶斯方法有了智能的成分

贝叶斯公式刚开始理解很抽象. 在面对具体实例时,把两个事件分别替换A, B, 逐个实例去理解. 贝叶斯方法在统计学和机器学习中威力巨大, 所以需要花功夫学习, 否则后面会有很多方法理解不了.

Notice

虽然贝叶斯统计方法和经典的统计学有很大不同,但是贝叶斯统计是以经典统计学为基础的, 两者并不对立. 理解贝叶斯首先要对经典统计学有一定的理解.

Example

示例1

首先声明! 现在不要看窗外,然后考虑这个问题: 明天下雨的概率是多少?

Info

示例1 描述的是一个随机事件,之所以称为随机事件,因为对于到底有哪些因素决定明天是否会下雨还不清楚, 如果清楚了解具体因素决定导致降雨,那么我们就能确定知道明天是否会下雨. 随机事件表示事件的不确定性.

- 明天下雨和不下雨的概率合计为 1:

$$P(rain) + P(not\_train) = 1$$

如果我们不看窗外, 那么下雨的概率为:

$$50\% : P(rain) = 0.5$$

由于总的概率合计为1,所以不下雨的概率为:

$$P(not\_train) = 1 - P(rain) = 0.5$$

$P(not\_train), P(rain)$ 称为**互斥事件**, 因为一个发生另一个就不可能发生.

互斥事件也可以用贝叶斯公式计算,意义不大, 但是为了学习, 我们还是计算一下

Example

示例2

现在看窗外下着雨, 问不下雨的概率是所少?

$P(not\_train), P(rain)$ 为互斥事件, 所以$P(rain|not_train)$ 概率都为0

我们套用一下贝叶斯公式:

$$P(not\_train|rain) = \frac{P(rain) * P(rain|not\_train)}{P(not\_train)}$$

$$P(not\_train|rain) = \frac{0.5 * 0}{0.5}$$

$$P(not\_train|rain) = 0$$

注意分子的第二项为0

计算遵守了贝叶斯规则和数学规则. 如果两个事件不是互斥事件, 那么问题就有意思多了.

Example

示例3

问题: 在某地,如果现在看窗外乌云密布, 问下雨的概率是多少?

- 信息1:过去100天的天气做了频数统计:降雨:20,多云:40, 晴天:40,单位为天
- 信息 2: 根据过往统计信息, 该地下雨时有云的概率为85%

过去 100 天的数据进行标准化处理(归一)得到天气的概率:

$$P(rain) = 0.2$$

$$P(cloudy) = 0.4$$

$$P(sunny) = 0.4$$

$$P(rain) + P(cloudy) + P(sunny) = 1$$

以上的概率称为先验概率,在计算完成后, 我们再回来讨论这个问题

此外根据长期的经验和观察知道下雨的时候有云的概率为85%:

$$P(cloudy|rain) = 0.85$$

将上面的数据带入到贝叶斯公式:

$$P(rain|cloudy) = \frac{P(rain) * P(cloudy|rain)}{P(cloudy)}$$

$$P(rain|cloudy) = \frac{0.2 * 0.85}{0.4}$$

$$P(rain|cloudy) = 0.425$$

结论描述: 根据长期观察形成的概率分布, 某天降雨概率为$P(rain) = 0.2$, 当观察到当前为多云的时候, 降雨概率为$P(rain|cloudy) = 0.425$. 根据当天观察, 降雨的概率从0.2 提高到 0.425 差不多提高了一倍的降雨概率.

考虑贝叶斯定理在做什么工作比背公式更为重要.

贝叶斯定理的理论就是:基于观察信息对已有的概率问题进行修正. 基于长期观测,降雨概率会为20%, 如果添加了多云的信息, 降雨的概率修正为42.5% .做出修正的基础是因为下雨时有85% 的概率都是多云.

示例 3 的继续分析

$$P(rain) = 0.2$$

$$P(cloudy) = 0.4$$

$$P(sunny) = 0.4$$

以上的概率称为先验概率, 就是在我们谈论这个问题之前知道的信息.如果在实际中, 不同地方三个事件的概率会有所不同. 基于长期观察的信息我们总会知道某个具体地方的天气大致的分布信息. 例如在撒哈拉沙漠是一种分布, 在热带雨林地区是另一种分布. 在贝叶斯统计中先验概率分布越准确对后面的概率计算越有利.

当然并非绝对如此, 如果我们对一个系统的信息一无所知, 那么我们可以首先做一个假设. 例如在半人马座有一个和地球类似的行星, 我们想了解它的天气变化. 由于信息一无所知, 可以假设三种天气发生的概率都为$1/3$. 做这个假设并不是说就固定不变, 一旦确定了一个均匀分布, 后续根据信息我们会将其修改为多项式分布.

在建立行星天气模型时声明一个均匀分布和我们解决数学问题的思路一样, 首先设一个未知数, 然后求解.

随着信息的添加, 逐步对天气的分布进行修改. 提前确定一个分布为后面的计算提供了入口, 在编程的时候我们就知道要修改什么内容.

$$\text{建立假设} \xrightarrow{\text{收集信息}} \text{修改假设} \Rightarrow \text{确立模型}$$

如果你已经对这个流程很感兴趣, 那么可以先跳转到:**Bayesian Statistics using Julia and Turing** 看看一些实例.

小结

以上通过更新下雨概率的实例介绍了贝叶斯定理. 具体的的计算细节没有列出. 因为不可能在这么一篇短的文章里把这个问题讲清楚. 实际这个过程也符合贝叶斯定理: 你对贝叶斯定理理解的程度会随着添加的相关信息而变化的.

在实际研究中, 贝叶斯定理结合其他概率分布形成了强大的工具. 当你手中有一根枯藤时,借助贝叶斯定律有可能将其雕琢一个十分趁手的登山手杖.

```
md"""

##  小结

以上通过更新下雨概率的实例介绍了贝叶斯定理． 具体的的计算细节没有列出． 因为不可能在这么一篇短的文章里把这个问题讲清楚．  实际这个过程也符合贝叶斯定理： 你对贝叶斯定理理解的程度会随着添加的相关信息而变化的．

在实际研究中， 贝叶斯定理结合其他概率分布形成了强大的工具． 当你手中有一根枯藤时，借助贝叶斯定律有可能将其雕琢一个十分趁手的登山手杖．
"""
```