6 ccs3ain,教程 02 (版本 1.0)

1.从联合分布(见幻灯片)计算:

(一)P(腔)

(b) P(牙痛)。

(c) P(牙痛|腔)。(d) P(catch，不确定符号)(e) P(|牙痛，不确定符号)

回想一下 P 和 P 的区别(幻灯片来自第二周讲座)

2.你做 T 测试来判断你是否患有 d，测试结果是阳性的。你知道这个测试的准确率是 95%(当你确实患病时，检测呈阳性的概率是 0.95，而当你没有患病时，检测呈阴性的概率也是 0.95)。你也知道这种病很少见，只有 1 万人得这种病。

你得这种病的概率是多少?

如果这种疾病更常见，比如每 100 人中就有 1 人患病，这种情况会发生什么变化?

3.考虑两种病毒测试，A 和 B。

测试是 95%有效的识别病毒出现时(这是 95%的时间,该病毒存在,测试检测),但有 10%的假阳性率(10%的时间,它表明病毒存在时)。

B 测试识别病毒的效率为 90%，但有 5%的假阳性率。

这两种测试使用不同的、独立的方法来鉴定病毒。

1%的人感染了这种病毒。

乔使用测试 a 检测出病毒呈阳性，鲍勃使用测试 b 检测出病毒呈阳性，谁更有可能感染病毒?

4.请看第二周讲座中淋巴结上的脑膜炎例子。不使用 P(s)的值进行计算。

这使用了将分母看作标准化常数 α 的技巧。

您需要一个 P(s |mm)的值来完成此操作—使用 P(s |mm) = 0.1。

5.现在是本教程中可选的计算部分。

在济慈上你可以找到内衣裤。py 文件。下面是使用名为石榴的 Python 包编写的注释中的“错误内裤”示例:

https://pomegranate.readthedocs.io/en/latest/index.html

石榴为概率推理提供了工具。你可以安装它使用:康达安装石榴

如果你安装了 Anaconda[1]. 否则，请参阅网页。

安装了石榴后，您可以使用 python underwear.py 运行内衣示例

从命令行。设置好后，这就给了你第一次观察错误内衣的结果。(你可以对照幻灯片上的值。)要对第二次观察到的错误内衣进行计算，需要更改 prior_cheat 的值，即作弊的先验概率。你把它改成什么?对于第二个观测值，先验值是第一个观测值。这样做，并验证您得到了与幻灯片中相同的结果。

现在，尝试为脑膜炎示例构建您自己的概率模型。

1 .如果你没有安装 Anaconda，你可以从 https://www.anaconda.com/download 安装它。如果您计划使用 Python 做很多事情，那么拥有 Anaconda 是非常方便的，因为它使安装新包变得很容易。

第 02 课回答

(版本 1.0)

1.(一)

$$P(腔) = \sum_{\omega \mid =腔} P(\omega)$$

= P(腔内∧捕集∧牙痛)+ P(腔内∧捕集∧牙痛)

+ P(型腔∧* catch∧*牙痛)+ P(型腔∧* catch∧*牙痛)= 0.108 + 0.012 + 0.072 + 0.008

= 0.2

(b)

P(牙痛)= hP(牙痛)，P(牙痛)= hP(牙痛

P(牙痛)= P(腔内∧∧网∧牙痛)+ P(腔内∧网∧网∧牙痛)+ P(腔内∧网∧网∧网∧牙痛)+ P(腔内∧网∧网∧网∧牙痛)= 0.108 + 0.012 + 0.016 + 0.064

= 0.2

P(牙痛)= 0.2 吗? 0.8

第二个概率的计算方法与第一个相同，或者知道两个的和是 1。

(c)

$$P(牙痛|腔) = \frac{P(牙痛∧腔}{P(腔)}$$

P(∧腔)= P(腔内∧捕集∧腔)+ P(腔内∧捕集∧∧腔)

= 0.108 + 0.012

= 0.12

所以:

P(牙痛|腔)= 0.6

类似地(或通过减法):

P(¬牙痛|腔)= 0.4

和:

P(牙痛|腔)= ? 0.6 ?0 . 4

(d)

P(抓住∨腔)= P(腔∧抓住∧牙痛)+ P(腔∧抓住∧¬牙痛)+ P(¬腔∧抓住∧牙痛)+ P(¬腔∧抓住∧¬牙痛)+ P(腔∧¬抓住∧牙痛)+ P(腔∧¬抓住∧¬牙痛)= 0.108 + 0.072 + 0.016 + 0.144 + 0.012 + 0.008

= 0.36

(e)

$$痛∨伤处) = \frac{P(型腔∧(牙痛∨伤处))P(型腔|牙}{P(牙痛∨接住)}$$

和前面的问题类似，我们可以计算:

P(牙痛，∨接住)= 0.416

(在这种情况下，我们把第 32 页上所有的数值加起来，除了最后一列中我们提到的“蛀牙”和“牙痛”)。然后:

$$P(腔\wedge(牙痛\vee捕捉))= 0.108 + 0.012 + 0.072$$

$$= 0.192$$

$$P(腔|牙痛\vee接住)= 0.462$$

$$\mathbf{P}(Cavity|toothache \vee catch) = \begin{pmatrix} 0.462 \\ 0.538 \end{pmatrix}$$

2.问题告诉我们:

$$P(t|d) = 0.95$$
$$P(\neg t|\neg d) = 0.95$$
$$P(d) = 0.0001$$

和:

$$P(t|\neg d) = 1 - P(\neg t|\neg d)$$
$$= 0.05$$

现在,

$$\begin{aligned} P(d|t) &= \frac{P(d \wedge t)}{P(t)} \\ &= \frac{P(t|d)P(d)}{P(t|d)P(d) + P(t|\neg d)P(\neg d)} \\ &= \frac{0.95 \cdot 0.0001}{0.95 \cdot 0.0001 + 0.05 \cdot 0.9999} \\ &= 0.00189 \end{aligned}$$

因此，一旦你有阳性测试，患病的概率是 0.00189，这不是很大，尽管测试的准确性。原因是这种疾病非常不可能发生，这意味着先验概率很低。

重复 P(d) = 0.01 的计算，得到的患病概率为 0.16，这显然要大得多(尽管仍然远低于 0.95，许多人在没有计算的情况下会说，如果检测结果为阳性，患病的概率为 0.95)。

3.我们有:

$$P(a|v) = 0.95 \qquad P(b|v) = 0.9$$
$$P(a|\neg v) = 0.1 \qquad P(b|\neg v) = 0.05$$

现在，对于乔，我们想要:

$$\begin{aligned} P(v|a) &= \frac{P(a|v)P(v)}{P(a|v)P(v) + P(a|\neg v)P(\neg v)} \\ &= \frac{0.95 \cdot 0.01}{0.95 \cdot 0.01 + 0.99 \cdot 0.1} \\ &= 0.0876 \end{aligned}$$

而且，和以前一样，因为即使经过准确的检测，也不太可能出现这种病毒，因此感染这种病毒的几率也不是很大。

对 Bob 进行类似的计算，我们发现:

$$P(v|b) = 0.154$$

所以即使第二次测试不太准确，但较低的假阳性率意味着鲍勃感染病毒的可能性是乔的两倍。

4.在第二周讲座的幻灯片中

$$P(s|m) = 0.0001$$

我们还知道 P(s| mm) = 0.1。然后，用来表示阿达玛乘积(你知道它是什么并不重要，重要的是你要按行相乘;你也可以把它看成是两个方程的集合)

P(M|s) = α P(s|M) P(M)

P(m |s) = α ? P (s | m) ?? P(m) ?

P (s |¬m) P(¬米)

0.8 吗?0.0001

吗?= α

0.1 - 0.9999

0.8 *

0.0001 ?= α

0.1 • 0.9999 ?0.

00008?

= α

0.09999

5.如果您完成了本教程的这一部分，您应该能够通过检查上述计算的答案来检查您的代码是否工作。