演讲题目

林诗杰

linshijie@whu.edu.cn

2019 年 4 月 23 日

conditional probabilities:

$$p(x|y) := \frac{p(x,y)}{p(y)}$$

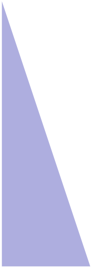
the joint probalility of x and y:

$$p(x,y) = p(x|y)p(y) = p(y|x)p(x)$$

Theorem: Bayes Rule

Denote by X and Y random variables then the following holds

$$p(y|x) = \frac{p(x|y)p(y)}{p(x)}$$



$p(t\|x)$	$X = \texttt{HIV}-$	$X = \texttt{HIV}+$
$T = \texttt{HIV}-$	0.99	0
$T = \texttt{HIV}+$	0.01	1

$$p(X = \texttt{HIV}+) = 0.0015$$

By Bayes rule we may write

$$p(X = \mathtt{HIV+} | T = \mathtt{HIV+}) = \frac{p(T = \mathtt{HIV+} | X = \mathtt{HIV+}) p(X = \mathtt{HI}}{p(T = \mathtt{HIV+})}$$

While we know all terms in the numerator, $p(T = \mathtt{HIV+})$itself is unknown. That said, it can be computed via

$$\begin{aligned} p(T = \mathtt{HIV+}) &= \sum_{x \in \{\mathtt{HIV+}, \mathtt{HIV-}\}} p(T = \mathtt{HIV+}, x) \\ &= \sum_{x \in \{\mathtt{HIV+}, \mathtt{HIV-}\}} p(T = \mathtt{HIV+} | x) p(x) \\ &= 1.0 \cdot 0.0015 + 0.01 \cdot 0.9985 \end{aligned}$$

Substituting back into the conditional expression yields

$$p(X = \mathtt{HIV+} | T = \mathtt{HIV+}) = \frac{1.0 \cdot 0.0015}{1.0 \cdot 0.0015 + 0.01 \cdot 0.9985} = 0.1$$

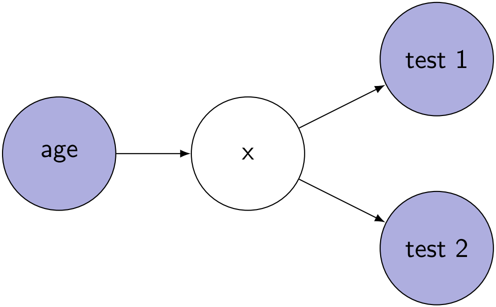
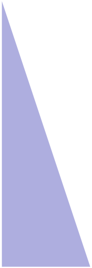


Figure: A graphical description of our HIV testing scenario. Knowing the age of the patient influences our prior on whether the patient is HIV positive (the random variable X). The outcomes of the tests 1 and 2 are independent of each other given the status X. We observe the shaded random variables (age, test 1, test 2) and would like to infer the un-shaded random variable X.



The corresponding expression yields:

$$\frac{p(T = \texttt{HIV+} | X = \texttt{HIV+}, A) p(X = \texttt{HIV+} | A)}{p(T = \texttt{HIV+} | A)}$$

We may assume that the test is independent of the age of the patient, i.e.

$$p(t|x, a) = p(t|x)$$

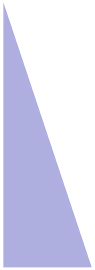
What remains therefore is $p(X = \mathtt{HIV+}|A)$. Recent US census data pegs this number at approximately 0.9%.

$$p(X = \mathtt{H+}|T = \mathtt{H+}, A) = \frac{p(T = \mathtt{H+}|X = \mathtt{H+}, A)p(X = \mathtt{H+}|A)}{p(T = \mathtt{H+}|A)}$$

$$= \frac{p(T = \mathtt{H+}|X = \mathtt{H+}, A)p(X = \mathtt{H+}|A)}{p(T = \mathtt{H+}|X = \mathtt{H+}, A)p(X = \mathtt{H+}|A) + p(T = \mathtt{H+}|X = \mathtt{H-},}$$

$$= \frac{p(T = \mathtt{H+}|X = \mathtt{H+})p(X = \mathtt{H+}|A)}{p(T = \mathtt{H+}|X = \mathtt{H+})p(X = \mathtt{H+}|A) + p(T = \mathtt{H+}|X = \mathtt{H-})p(}$$

$$= \frac{1 \cdot 0.009}{1 \cdot 0.009 + 0.01 \cdot 0.991} = 0.48$$



we assume that the statistics for T_2 are given by

$p(t_2 \vert x)$	$X = \mathtt{HIV}-$	$X = \mathtt{HIV}+$
$T_2 = \mathtt{HIV}-$	0.95	0.01
$T_2 = \mathtt{HIV}+$	0.05	0.99

for $t_1 = t_2 = \mathtt{HIV}+$ we have

$$p(X = \mathtt{H}+ | T_1 = \mathtt{H}+, T_2 = \mathtt{H}+)$$

$$= \frac{p(T_1 = \mathtt{H}+, T_2 = \mathtt{H}+ | X = \mathtt{H}+) p(X = \mathtt{H}+ | A)}{p(T_1 = \mathtt{H}+, T_2 = \mathtt{H}+ | A)}$$

$$= p(T_1 = \mathtt{H}+ | X = \mathtt{H}+) p(T_2 = \mathtt{H}+ | X = \mathtt{H}+) p(X = \mathtt{H}+ | A) \; /$$

$$p(T_1 = \mathtt{H}+ | X = \mathtt{H}+) p(T_2 = \mathtt{H}+ | X = \mathtt{H}+) p(X = \mathtt{H}+ | A)$$

$$+ p(T_1 = \mathtt{H}+ | X = \mathtt{H}-) p(T_2 = \mathtt{H}+ | X = \mathtt{H}-) p(X = \mathtt{H}- | A)$$

$$= \frac{1 \cdot 0.99 \cdot 0.009}{1 \cdot 0.99 \cdot 0.009 + 0.01 \cdot 0.05 \cdot 0.991} = 0.95$$