Derivation of Bayes Theorem

Andy Grogan-Kaylor

2025-06-24

Table of contents

1 Some Definitions

$$\begin{array}{ccc} & P(A) & P(!A) \\ P(B) & P(A,B) & P(A,!B) \\ P(!B) & P(A,!B) & P(!A,!B) \end{array} \tag{1}$$

- $P(A)$ and $P(B)$ are examples of *marginal* probabilities.
- $P(A,B)$ is an example of a *joint* probability.
- A *conditional* probability, for example, $P(A|B)$ is defined as the appropriate *joint* probability divided by the *marginal* probability: $\frac{P(A,B)}{P(B)}$

2 From The Definition Of Conditional Probability:

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

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

3 Multiply Each Fraction By The Denominator:

$P(A|B)P(B) = P(A,B)$

$P(B|A)P(A) = P(A,B)$

4 Set The Two Expressions To Be Equivalent:

$P(A|B)P(B) = P(B|A)P(A)$

5 Divide by $P(B)$:

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

This is Bayes Theorem.