362 (conditional probability) Bayes defined conditional probability, using his own notation, as follows:

$$P(A \mid B) = \frac{P(A \wedge B)}{P(B)}$$

This is to be read "the probability that A is true given that B is true is equal to the probability that both are true divided by the probability that B is true". How is conditional probability expressed in our notation?

After trying the question, scroll down to the solution.

§ A binary expression B may not be a distribution. It becomes a distribution if we divide it by its sum. For example, let x be the only variable. Then $B / (\Sigma x\cdot B)$ is the distribution proportional to B. Define

$$\updownarrow B \;=\; B / (\Sigma x\cdot B)$$

We pronounce $\updownarrow B$ as "normalize B". We can now express Bayes' conditional probability $P(A \mid B)$ as

$$\updownarrow B'.\,A$$

This describes the situation where we learn that B is true and then ask if A is true. This situation is just one of infinitely many situations for which we may want to calculate a probability. We cannot make a special notation for each one, as Bayes has done for conditional probability. We need to be able to describe the situation using a basic set of connectives, and from that description, calculate probabilities, as we do.

To prove that we have described Bayes' conditional probability, again let x be the only variable, and let n be the size of its domain. Then

$$
\begin{array}{lll}
 & \updownarrow B'.\,A & \text{use definition of } \updownarrow \\
= & B' / (\Sigma x'\cdot B').\;\; A & \text{use definition of } . \\
= & \Sigma x''\cdot\;\; B'' / (\Sigma x'\cdot B') \times A'' & \text{rearrange and rename local variables} \\
= & (\Sigma x\cdot A \times B) / (\Sigma x\cdot B) & \text{divide numerator and denominator each by } n \\
= & \dfrac{(\Sigma x\cdot A \times B) / n}{(\Sigma x\cdot B) / n} & \text{switch to Bayes' probability notation} \\
= & \dfrac{P(A \wedge B)}{P(B)} & \text{use Bayes definition of conditional probability} \\
= & P(A \mid B) &
\end{array}
$$