

1 LaTeX formatting

Reproduce the formatting of the following equation. You might want to use the following commands:

`\align;` `\intertext;` `\left;` `\right;` `\sum;` `\prod` `\dots;` `\frac;` `\log;`

$$\frac{\sum_{i=1}^n \log(a_i)}{n} = \frac{\log(a_1) + \log(a_2) + \dots + \log(a_n)}{n} \tag{1}$$

$$= \frac{\log(a_1 a_2 \dots a_n)}{n} \tag{2}$$

$$= \log \left[\left(\prod_{i=1}^n a_i \right)^{\frac{1}{n}} \right] \tag{3}$$

because $\log(MN) = \log(M) + \log(N)$ and $\log(M^k) = k \log(M)$

$$= \log \left(\prod_{i=1}^n \sqrt[n]{a_i} \right) \tag{4}$$

2 Crime scene

Suppose a crime has been committed. Blood is found at the scene for which there is no innocent explanation. It is of a type which is present in 1% of the population.

1. The prosecutor claims: “There is a 1% chance that the defendant would have the crime blood type if he were innocent. Thus there is a 99% chance that he is guilty”. This is known as the prosecutor’s fallacy. Express the prosecutor’s claim symbolically (e.g. “ $P(A|B)$ ”) and explain what is wrong with the argument.

What is wrong with this argument?

2. The defender claims: “The crime occurred in a city of 800,000 people. The blood type would be found in approximately 8000 people. The evidence has provided a probability of just 1 in 8000 that the defendant is guilty, and thus has **absolutely no** bearing on the investigation.” This is known as the defender’s fallacy. Express the defender’s statement symbolically and explain what is wrong with this argument.

3 Conditional independence iff joint factorizes

In the text we said that $X \perp Y | Z$ iff

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

for all x, y, z such that $p(z) > 0$. Now prove the following alternative definition: $X \perp Y | Z$ iff there exist functions g and h such that

$$p(x, y | z) = g(x, z)h(y, z)$$

for all x, y, z such that $p(z) > 0$.

Hint: What are $p(x | z)$ and $p(y | z)$ expressed using g and h ?