

QuantNeuro HW 1: Bayesian vs Frequentist Statistics

Anna Voss

2023-09-05

Background

“Imagine running an HIV test on A SAMPLE of 1000 persons ...”

“The test has a false positive rate of 5% (0.05)...” i.e., the probability that someone who takes the test gets a POSITIVE result despite the fact that the person does NOT have HIV

“...and no false negative rate.” i.e., The probability that someone who takes the test gets a NEGATIVE result despite the fact that the person DOES have HIV.

Exercise #1:

If someone gets a positive test, is it “statistically significant” at the $p < 0.05$ level? Why or why not?

At the $p < 0.05$ level, a positive test is considered statistically significant according to frequentist statistics. This is based on the alpha value/false positive rate of this test = 0.05. This value is interpreted to mean that there is a 5% chance of a Type 1 error, i.e. that the positive test was a result of chance (false positive). Given this “alpha” significance of value of 0.05, a positive test has a statistical significance of $p = 0.05$, so we reject the null hypothesis that the test results do not indicate an HIV infection and conclude that the test is statistically significant at the 0.05 level.

Exercise #2:

What is the probability that if someone gets a positive test, that person is infected?

Following Bayes’ Rule, we know that:

$$p(\text{Hypothesis} \mid \text{Data}) = p(\text{Data} \mid \text{Hypothesis}) * p(\text{Hypothesis}) / p(\text{Data})$$

For our purposes this translates to:

$$p(\text{HIV Infection} \mid \text{Positive HIV test}) <- p(\text{Positive HIV Test} \mid \text{HIV Infection}) * \text{Infection Rate} / ((\text{Infection rate} * 1 - \text{False positive test rate}) + (1 - \text{Infection Rate}) * \text{False positive test rate})$$

Note that the first half of the denominator is the same as the numerator, and the second half of the denominator is simply the reverse of the numerator to account for the inverse situation.

```
#For this exercise, assume a range of priors (infection rates) from 0 to 1 in steps of 0.1
infection_rate <- seq(from = 0, to = 1, by = 0.1)
```

```
false_positive_rate <- 0.05
```

```

for(i in infection_rate){
  p_positive_test_given_hypothesis <- 1 - false_positive_rate
  probability_of_infection <- p_positive_test_given_hypothesis * i /
  ((i * (1 - false_positive_rate)) + ((1 - i) * false_positive_rate))
  cat("Infection rate =", i, ", Probability of infection given a positive test =",
  probability_of_infection, "\n")
}

```

```

## Infection rate = 0 , Probability of infection given a positive test = 0
## Infection rate = 0.1 , Probability of infection given a positive test = 0.6785714
## Infection rate = 0.2 , Probability of infection given a positive test = 0.826087
## Infection rate = 0.3 , Probability of infection given a positive test = 0.890625
## Infection rate = 0.4 , Probability of infection given a positive test = 0.9268293
## Infection rate = 0.5 , Probability of infection given a positive test = 0.95
## Infection rate = 0.6 , Probability of infection given a positive test = 0.9661017
## Infection rate = 0.7 , Probability of infection given a positive test = 0.9779412
## Infection rate = 0.8 , Probability of infection given a positive test = 0.987013
## Infection rate = 0.9 , Probability of infection given a positive test = 0.994186
## Infection rate = 1 , Probability of infection given a positive test = 1

```