

# Bayesian Statistics

Quan Zhang

- Bayes Theorem

$$p(H | E) = \frac{p(H) \times p(E | H)}{p(E)}$$

- Influence of prior distributions

Suppose a person voted for Trump in 2020. Where is he/she most likely from?

A. CA

B. NY

C. NE

- Strong Evidence

3% of a country's population has a particular disease.

A test has 98% true positive rate, i.e.,  $p(+|\text{disease})=0.98$ ,  
and 4% false positive rate, i.e.,  $p(+|\text{no disease})=0.04$ .

If a person simultaneously take the test twice and have both positive result. What is the probability that he/she indeed has the disease?

(Assume the two simultaneous tests are independent)

- In frequentist statistics, a parameter is an unknown constant number that needs to be estimated.
- In Bayesian statistics, an unknown parameter or a parameter of interest is a random variable and thus has a distribution. We estimate the distribution of the parameter. By assuming some prior distribution (hypothesis) and observing some data (evidence), we find the posterior distribution by Bayes' theorem.