Bayesian Statistics: Applying Baye's Theorem

Tanner Crane

June 2025

1 Medical Test problem

A medical test for a disease has a 99 percent sensitivity and 98 percent specificity. The disease prevalence is 0.5 percent. Calculate the probability that a person has the disease given a positive test result.

Let A denote a person having the disease
Let B denote a positive test result
Let P(A) be the probability of getting the disease at 0.005
$P(B|A) = 0.99$ of people with disease test positive
0.98 of people without the disease test negative

We want to calculate the probability that a person has the disease given a positive test result or $P(A|B)$.

$P(B) = P(B|A) * P(A) + P(B|A') * P(A')$

Where $P(B|A')$ is the probability of a positive test result given that the patient does not have the disease (false positive). This is calculated by $1 - 0.98 = 0.02$

P(A') is given by 1 - 0.005 = 0.995

Hence, $P(B) = (0.99 * 0.005) + (0.02 * 0.995) = 0.02485$

$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{0.99*0.005}{0.02485} = 0.199$

Therefore, the probability that a person will have a disease given a positive test result is 19.9 percent.

2 Survey Problem

Survey Problem: In a survey, 70 percent of respondents said they liked a new product. If the probability of a person liking the product given they are under 30 is 80 percent, and the probability of being under 30 is 50 percent, calculate the probability that a respondent is under 30 given they like the product.

We want to know the probability that a respondent is under 30 given they like the product or, $P(U|L)$ where U is being under 30 and L is liking the product.
$P(L|U)$ is 0.8
$P(L) = 0.7$
$P(U) = 0.5$

Hence, $P(U|L) = \frac{P(U|L)*P(U)}{P(L)} = \frac{0.8*0.5}{0.7} = 0.571$

Therefore, the probability of respondent is under 30 given they like the product is 57.1 percent.

3 Spam Email Problem

An email classifier tags 95 percent of spam emails correctly (sensitivity) and 90 percent of non-spam emails correctly (specificity). If 20 percent of the emails are spam, calculate the probability that an email is spam given it was tagged as spam.

Let P(E) be the probability that an email is spam, we want to calculate the probability that an email is spam given that it was tagged as spam, or $P(E|T)$ Let E denote an email being spam and T denote an email being tagged as spam.

P(E) = 0.2 is the probability of an email being spam
P(E') = 0.8 is the probability of an email not being spam
P(N) = 0.9 is the probability of non-spam emails correctly classified
$P(T|E') = 1 - 0.9 = 0.1$ is the probability that an email is tagged as spam and it's not spam
$P(T|E) = 0.95$ is the probability that an email is tagged as spam and is spam

Hence, $P(T) = P(T|E)*P(E)+P(T|E')*P(E') = (0.95*0.2)+(0.1*0.8) = 0.27$

$P(E|T) = \frac{P(T|E)P(E)}{P(T)} = \frac{0.95*0.20}{0.27} = 0.7037$

We can conclude the probability that an email is spam given it was tagged as spam is 70.37 percent.

4 Drug Test Problem

A drug test is 99 percent accurate for users (true positive rate) and 99 percent accurate for non-users (true negative rate). If 1 percent of the population uses the drug, calculate the probability that a person is a drug user given they tested positive.

We want to know probability that a person is a drug user given they tested positive, denoted as $P(U|+)$. Where:

p(U) = 0.01
P(U') = 1 - 0.01 = 0.99
$P(+|U) = 0.99$
$P(+|U') = 0.01$

Using the law of total probability:
$P(+) = P(U) * P(+|U) + P(U') * P(+|U') = (0.01 * 0.99) + (0.99 * 0.01) = 0.0198$

Hence, $P(U|+) = \frac{P(+|U)P(U)}{P(+)} = \frac{(0.99*0.01)}{0.0198} = 0.5$

Therefore, the probability that a person is a drug user given they tested positive is 50 percent.

5 Marketing Campaign Problem

A marketing campaign targets a specific demographic with a 70 percent probability of purchasing a product if they see an ad. The probability of a person seeing the ad is 40 percent, and the overall purchase rate is 30 percent. Calculate the probability that a person has seen the ad given that they purchased the product.

We want to know the probability that a person has seen the add given that they purchased the product, or $P(S|P)$
Let S denote a person seeing the ad
Let P denote a purchase, then:

$P(P|S) = 0.7$
$P(S) = 0.4$
$P(P) = 0.3$

$P(S|P) = \frac{P(P|S)*P(S)}{P(P)} = \frac{0.7*0.4}{0.3} = 0.9333$

Therefore, the probability that person has seen the add given that they purchased the product is 93.33 percent.