Brock Pinagel

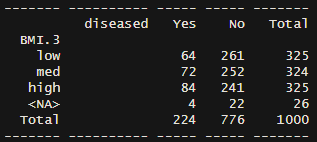
Dr. Colin M. Wasiloff

MBA 6103 Statistics for Data Analytics and Visualization

2 June 2025

Module 5 Assignment

Use the tobacco data frame from the summarytools package, recode BMI into three equal sized groups (low, med, high).

* 1. Paste a contingency table of Diseased x BMI.3
  2. What is the probability that a person is Diseased?

P(diseased) = 224/1000 = **22.4%**

* 1. What is the probability that a person is Diseased or High BMI?

P(diseased or high BMI) = P(diseased) + P(high BMI) – P(high BMI and diseased)

= 224/1000 + 325/1000 – 84/1000 = **46.5%**

* 1. What is the probability that a person is a Diseased if a person is High BMI?

P(diseased | high BMI) = (84/1000)/(325/1000) = **25.85%**

* 1. What is the probability that a person is Diseased and is Low BMI?

P(diseased and low BMI) = P(diseased)P(low BMI | diseased) = (224/1000)(64/224) = **6.4%**

* 1. Use the Bayes theorem function for an event in R to answer the following question:

Given the following statistics, what is a patient’s probability of having liver disease if they are an alcoholic?

* 10% of patients have liver disease
* 5% of patients are alcoholics
* Among those patients diagnosed with liver disease, 7% are alcoholics

P(A) = P(Diseased) = 0.10

P(B) = P(Alcoholic) = 0.05

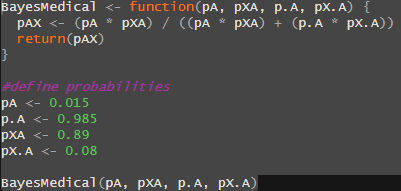
P(Alcoholic | Diseased) = 0.07

P(Diseased | Alcoholic) = 0.10 \* 0.07 / 0.05 = 0.14 or **14%**

* 1. Use the Bayes theorem function for a medical test in R to answer the following question?

Given the following statistics, what is the probability that a woman has cancer if she has a positive mammogram result?

* 1.5 percent of women over 50 have breast cancer
* 89 percent of women who have breast cancer test positive on mammograms
* Eight percent of women will have false positives

Answer: **14.49%**