Chapter Four Review

Spring Semester, 2023

Conditional Probability

- $P(A|B) = \frac{P(A \cap B)}{P(B)}$, when $P(B) > 0$.
- Conditional probability defines a probability measure. That is, conditional probabilities satisfy the assumptions of a probability measure.
- An important identity is that

$$P(A \cap B) = P(A|B)P(B)$$

Law of Total Probability

- We start with the sample space S.
- Next, we define a collection of covering sets:

$$S = C_1 \cup C_2 \cup C_3 \text{ where } C_1 \cap C_2 = \emptyset,$$
$$C_1 \cap C_3 = \emptyset, \text{ and } C_2 \cap C_3 = \emptyset.$$

- Now we seek to calculate $P(E)$:

$$P(E) = P(E \cap S) = P(E \cap (C_1 \cup C_2 \cup C_3)) =$$
$$P[(E \cap C_1) \cup (E \cap C_2) \cup (E \cap C_3)]$$

Law of Total Probability (cont)

$$P(E) = P[(E \cap C_1) \cup (E \cap C_2) \cup (E \cap C_3)]$$

$$= P(E \cap C_1) + P(E \cap C_2) + P(E \cap C_3)$$

$$= P(E|C_1)P(C_1) + P(E|C_2)P(C_2) + P(E|C_3)P(C_3)$$

- This is the law of total probability when there are three covering sets.

Bayes' Theorem

$$P(C_1|E) = P(C_1 \cap E)/P(E)$$

$$= \frac{P(E|C_1)P(C_1)}{P(E)}.$$

- Use the law of total probability to find $P(E)$.

Chapter 4 Guide, Problem 1

An individual has one of three genotypes called A, B, and C, respectively, for a gene associated with disease X. The probability that an individual has genotype A is 0.64; the probability that an individual has genotype B is 0.32; and the probability that an individual has genotype C is 0.04. The probability that an individual with the A genotype is affected with disease X is 0.05. The probability that an individual with the B genotype is affected with disease X is 0.80. The probability that an individual with the C genotype is affected with disease X is 0.99.

a. What is the probability that an individual is affected with disease X?
b. Given that an individual has disease X, what is the probability that the individual is genotype B?

Random Variables

- Discrete
  - Bernoulli: $P(X = 0) = 1 - p$ and $P(X = 1) = p$
  - Binomial: number of successes in *n* independent trials each with probability of success *p*
  - Poisson: number of “rare” events
- Continuous
  - Normal: $N\ (\mu, \sigma^2)$

Normal Distribution

Standard score form of the normal distribution:

$$Z = \frac{X - \mu}{\sigma}$$

Any probability calculation about a normal distribution can be transformed to a calculation with a standard normal:

$$P(X \leq a) = P\left(\frac{X - \mu}{\sigma} \leq \frac{a - \mu}{\sigma}\right) = ø\left(\frac{a - \mu}{\sigma}\right)$$

Key Percentiles of the Standard Normal

- $P(Z \leq -2.576) = 0.005$
- $P(Z \leq -2.326) = 0.01$
- $P(Z \leq -1.960) = 0.025$
- $P(Z \leq -1.645) = 0.05$
- $P(Z \leq -1.282) = 0.10$
- $P(Z \leq -0.6745) = 0.25$

Expected Value of a Random Variable

- Example expectation of a Bernoulli rv:

$$E(X) = \sum_{APV} xP(X = x)$$

$$= 0(1 - p) + 1(p) = p$$

- The expected value of an indicator variable is the probability that the indicator variable is on.
- Expectation is a linear operator

Variance of a Random Variable

- Definition of variance: $var(X) = E((X - EX)^2)$
- Bernoulli random variable variance:
  $(0 - p)^2(1 - p) + (1 - p)^2 p = p(1 - p)$
- Important identity:
  - $var(X) = E((X - EX)^2) = E(X^2) - (EX)^2$

Chapter 4 Guide, Problem 7

The random variables W_1 and W_2 are a random sample of 2 drawn from the random variable W

which has expected value μ_W and standard deviation σ_W. Find $E(W_1 - W_2)$ and $E((W_1 - W_2)^2)$.