

Section #2 Solutions

1 Warmup: A Simple Random Variable

A discrete random variable is defined by its probability mass function. Let $X = Y + Z$ where Y and Z are independent uniformly random integers between 1 and 4 inclusive. Express X 's probability mass function as a table, a graph, and a mathematical function. Verify that linearity of expectation holds (i.e. $E[X] = E[Y] + E[Z]$). As a bonus calculate $Var(X)$. Is variance linear?

To get $P(X = x)$ we can multiply the number of assignments of Y, Z that result in $Y + Z = x$ by the likelihood of each assignment, $1/16$. Since there are 4 ways to get $X = 5$, 3 ways to get each of 4 and 6, 2 ways to get 3/7, one way to get 2/8, we have:

$$P(X = x) = \frac{4 - |x - 5|}{16}, x = 2, \dots, 8$$

We can calculate the expectation of X in two different ways:

$$E[X] = \sum_{x=2}^8 xP(X = x) = \sum_{x=2}^8 \frac{x(4 - |x - 5|)}{16} = 5$$

$$E[X] = E[Y + Z] = E[Y] + E[Z] = 2 \sum_{i=1}^4 \frac{i}{4} = 5$$

verifying linearity of expectation. Finally, X 's variance is:

$$Var(X) = E[X^2] - E[X]^2 = \sum_{x=2}^8 \frac{x^2(4 - |x - 5|)}{16} - 25 = 2.5$$

In general, variance is not linear, but since Y and Z are independent, $Var(Y + Z) = Var(Y) + Var(Z)$. (and indeed $Var(Y) = Var(Z) = 1.25$).

2 Taking Expectation: Breaking Vegas

Preamble: When a random variable fits neatly into a family we've seen before (e.g. Binomial), we get its expectation for free. When it does not, we have to use the definition of expectation.

Problem: If you bet on "Red" in Roulette, there is $p = 18/38$ that you will win $\$Y$ and a $(1 - p)$ probability that you lose $\$Y$. Consider this algorithm for a series of bets:

1. Let $Y = \$1$.
2. Bet Y .
3. If you win, then stop.
4. If you lose, then set Y to be $2Y$ and goto step (2).

What are your expected winnings when you stop? It will help to recall that the sum of a geometric series $a^0 + a^1 + a^2 + \dots = \frac{1}{1-a}$ if $0 < a < 1$. Vegas breaks you: Why doesn't everyone do this?

Let X be the number of dollars that you earn.

The possible values of x are from the outcomes of: winning on your first bet, winning on your second bet, and so on.

$$\begin{aligned}
 E[X] &= \frac{18}{38} + \frac{20}{38} \frac{18}{38} (2 - 1) + \left(\frac{20}{38}\right)^2 \frac{18}{38} (4 - 2 - 1) + \dots \\
 &= \sum_{i=0}^{\infty} \left(\frac{20}{38}\right)^i \left(\frac{18}{38}\right) \left(2^i - \sum_{j=0}^{i-1} 2^j\right) \\
 &= \left(\frac{18}{38}\right) \sum_{i=0}^{\infty} \left(\frac{20}{38}\right)^i \\
 &= \left(\frac{18}{38}\right) \frac{1}{1 - \frac{20}{38}} = 1
 \end{aligned}$$

Real games have maximum bet amounts. You have finite money and casinos can kick you out. But, if you had no betting limits and infinite money, then go for it! (and tell me which planet you are living on).

3 Binomial Distribution: Sending Bits to Space

When sending binary data to satellites (or really over any noisy channel) the bits can be flipped with high probabilities. In 1947 Richard Hamming developed a system to more reliably send data. By using Error Correcting Hamming Codes, you can send a stream of 4 bits with 3 redundant bits. If zero or one of the seven bits are corrupted, using error correcting codes, a receiver can identify the original 4 bits.

Lets consider the case of sending a signal to a satellite where each bit is independently flipped with probability $p = 0.1$

- a. If you send 4 bits, what is the probability that the correct message was received (i.e. none of the bits are flipped).
- b. If you send 4 bits, with 3 Hamming error correcting bits, what is the probability that a correctable message was received?

- c. Instead of using Hamming codes, you decide to send 100 copies of each of the four bits. If for every single bit, more than 50 of the copies are not flipped, the signal will be correctable. What is the probability that a correctable message was received?

Hamming codes are super interesting. It's worth looking up if you haven't seen them before! All these problems could be approached using a binomial distribution (or from first principles).

- a. Let Y be the number of bits corrupted. $Y \sim \text{Bin}(n = 4, p = 0.1)$.

$$P(Y = 0) = \binom{4}{0} 0.9^4 = 0.656$$

- b. Let Z be the number of bits corrupted. $Z \sim \text{Bin}(n = 7, p = 0.1)$. A correctable message is received if Z equals 0 or 1:

$$\begin{aligned} P(\text{correctable}) &= P(Z = 0) + P(Z = 1) \\ &= \binom{7}{0} (0.1)^0 (0.9)^7 + \binom{7}{1} (0.1)^1 (0.9)^6 = 0.850 \end{aligned}$$

That is a 30% improvement!

- c. Let X_i be the number of copies of bit i which are not corrupted. We can represent each as a Binomial Random Variable: $X_i \sim \text{Bin}(n = 100, p = 0.9)$.

$$\begin{aligned} P(\text{correctable}) &= \prod_{i=1}^4 P(X_i > 50) \\ &= \prod_{i=1}^4 \sum_{j=51}^{100} P(X_i = j) \\ &= \prod_{i=1}^4 \sum_{j=51}^{100} \binom{100}{j} (0.9)^j (0.1)^{100-j} \\ &= \left(\sum_{j=51}^{100} \binom{100}{j} (0.9)^j (0.1)^{100-j} \right)^4 > 0.999 \end{aligned}$$

But now you need to send 400 bites, instead of the 7 required by hamming codes :-).

4 Conditional Probabilities: Missing Not at Random

Preamble: We have three big tools for manipulating conditional probabilities:

- Definition of conditional probability: $P(EF) = P(E|F)P(F)$
- Law of Total Probability: $P(E) = P(EF) + P(EF^C) = P(E|F)P(F) + P(E|F^C)P(F^C)$
- Bayes Rule: $P(E|F) = \frac{P(F|E)P(E)}{P(F)} = \frac{P(F|E)P(E)}{P(F|E)P(E) + P(F|E^C)P(E^C)}$

This is a good time to commit these three to memory and start thinking about when each of them is useful.

Problem: You collect data on whether or not people intend to vote for Ayesha, a candidate in an upcoming election. You send an electronic poll to 100 randomly chosen people. You assume all 100 responses are IID.

User Response	Count
Responded that they will vote for Ayesha	40
Responded that they will not vote for Ayesha	45
Did not respond	15

Let A be the event that a person says they will vote for Ayesha. Let M be the event that a user did not respond to the poll. We are interested in estimating $P(A)$, however that is hard given the 15 users who did not respond.

- What is the fraction of users that responded to the poll, $P(M^C)$?
- What is the probability that a user said they will vote for Ayesha, given that they responded to the poll $P(A|M^C)$?
- You estimate that $P(M|A) = \frac{1}{5}$ and that $P(M|A^C) = \frac{1}{10}$. What is $P(A|M)$? You may leave your answer in terms of $P(A)$. To make your future life easier, simplify your answer as much as possible.
- Using the same assumption as the previous part, calculate $P(A)$. You can make use of a function `root(a, b, c)` which returns the value of x such that $ax^2 + bx + c = 0$.

a. $P(M^C) = \frac{100-15}{100} = \frac{85}{100}$

b. $P(A|M^C) = \frac{40}{40+45} = \frac{40}{85}$

c.

$$\begin{aligned}
 P(A|M) &= \frac{P(M|A)P(A)}{P(M|A)P(A) + P(M|A^C)P(A^C)} \\
 &= \frac{P(M|A)P(A)}{P(M|A)P(A) + P(M|A^C)(1 - P(A))} \\
 &= \frac{\frac{1}{5}P(A)}{\frac{1}{5}P(A) + \frac{1}{10}(1 - P(A))} = \frac{2P(A)}{2P(A) + 1 - P(A)} \\
 &= \frac{2P(A)}{P(A) + 1}
 \end{aligned}$$

d.

$$\begin{aligned}
 P(A) &= P(A|M)P(M) + P(A|M^C)P(M^C) \\
 &= \frac{2P(A)}{P(A) + 1} \cdot \frac{15}{100} + \frac{40}{85} \frac{85}{100} \\
 100P(A) &= \frac{30P(A)}{P(A) + 1} + 40 \\
 [100P(A) - 40][P(A) + 1] &= 30P(A) \\
 100P(A)^2 + 100P(A) - 40P(A) - 40 &= 30P(A) \\
 0 &= 10P(A)^2 + 3P(A) - 4 \\
 P(A) &= \text{root}(10, 3, -4)
 \end{aligned}$$

Continuous Random Variable: Let X be a continuous random variable with the following probability density function:

$$f_X(x) = \begin{cases} c(e^{x-1} + e^{-x}) & \text{if } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

a. Find the value of c that makes f_X a valid probability distribution.

We need $\int_{-\infty}^{\infty} dx f_X(x) = 1$.

$$\begin{aligned}
 1 &= \int_{-\infty}^{\infty} f_X(x)dx = \int_0^1 c(e^{x-1} + e^{-x})dx \\
 &= c [e^{x-1} - e^{-x}]_{x=0}^1 \\
 1 &= c(e^{1-1} - e^{-1} - (e^{0-1} - e^{-0})) \\
 c &= \frac{1}{1 - e^{-1} - (e^{-1} - 1)} = \frac{1}{2 - \frac{2}{e}}
 \end{aligned}$$

b. What is $P(X > 0.75)$?

$$\begin{aligned}
 P(X > 0.75) &= \int_{0.75}^1 dx \, c(e^{x-1} + e^{-x}) \\
 &= -c \left[e^{x-1} - e^{-x} \right]_{x=0.75}^1 \\
 &= \boxed{-c \left(e^{1-1} - e^{-1} - (e^{0.75-1} - e^{-0.75}) \right)} \\
 &= -c \left(1 - e^{-1} - e^{-0.25} + e^{-0.75} \right) = \frac{1 - e^{-1} - e^{-0.25} + e^{-0.75}}{2 - \frac{2}{e}}
 \end{aligned}$$