Chris Piech
CS109

Section #2
Oct 6th, 2021

Section 2: Random Variables

Overview of Section Materials

The warmup questions provided will help students practice concepts introduced in lectures. The section problems are meant to apply these concepts in more complex scenarios similar to what you will see in problem sets and quizzes.

1 Warmups

1. Definitions: Cite Bayes' Theorem. Can you explain to your partner why $P(A|B)$ is different than $P(B|A)$?

2. True or False. Note that true means true for ALL cases.

   (a) In general, $P(AB|C) = P(B|C)P(A|BC)$

   (b) If A and B are independent, so are A and B^C.

1. Bayes' Theorem: $P(E|F) = \frac{P(F|E)P(E)}{P(F)}$

2. (a) True

   (b) True

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 with 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 $a^0 + a^1 + a^2 + \cdots = \frac{1}{1-a}$ if $0 < a < 1$. Vegas breaks you: Why doesn't everyone do this?

Let X be the number of dollars that your earn.
The possible values of x are from the outcomes of: winning on your first bet, winning on your second bet, and so on.

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

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 Sending Bits to Space

Preamble: When sending binary data to satellites (or really over any noisy channel), the bits can be flipped with high probability. 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 along 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.

Problem: 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 an interpretable message (i.e. a message with zero or one errors) 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!

a. Let Y be the number of 4 bits corrupted. Then $P(Y = k)$ is given as:

$$P(Y = 0) = \binom{4}{0}(0.1)^0(0.9)^4 = 0.656$$

b. Let Z be the number of 7 bits corrupted. 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 random variable as we did in parts a and b.

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

But now you need to send 400 bits, 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.

a. What is the probability that a user said they will vote for Ayesha and that they responded to the poll $P(A \text{ and } M^C)$?

b. Which formula from class would you use to calculate $P(A)$? Your formula should rely on the context that voters for Ayesha are in one of two (mutually exclusive) groups: those that missed the poll, and those that did not.

c. Calculate the $P(A)$. You estimate that the probability that a voter is missing, given that they were going to vote for Ayesha is $P(M|A) = \frac{1}{5}$.

a. $P(A \text{ and } M^C) = \frac{40}{100}$

b. The law of total probability. It breaks down $P(A)$ into two parts, the part which intersects with M and the part that intersections with M^C.

$$P(A) = P(A \text{ and } M) + P(A \text{ and } M^C)$$

c.

$$
\begin{aligned}
P(A) &= P(A \text{ and } M^C) + P(A \text{ and } M) && \text{Law of total probability} \\
&= \frac{40}{100} + P(A \text{ and } M) && \text{From part a} \\
&= \frac{40}{100} + P(M|A)P(A) && \text{Chain rule} \\
P(A) - P(M|A)P(A) &= \frac{40}{100} && \text{The rest is algebra} \\
P(A) \cdot [1 - P(M|A)] &= \frac{40}{100} \\
P(A) \cdot \frac{4}{5} &= \frac{40}{100} \\
P(A) &= \frac{40}{100} \cdot \frac{5}{4} \\
P(A) &= \frac{1}{2}
\end{aligned}
$$