STAT 111: Mathematical Statistics II

David Yang

Spring 2024

Abstract

These notes arise from my studies in STAT 111: Mathematical Statistics II, taught by Professor Phil Everson, at Swarthmore College, following the material of Munkre's *Topology*. I am responsible for all faults in this document, mathematical or otherwise. Feel free to message me with any suggestions or corrections at dyang5@swarthmore.edu.

Contents

1 Discrete Probability Distributions

1.2 Indicator Variables

a) **Define $I(A)$ to be an indicator variable for the event A, meaning $I_{(A)} = 1$ if A occurs and $I_{(A)} = 1$ if A^c occurs. Relate this to a Bernoulli random variable. Show how an indicator variable is the *fundamental bridge* between probability and expected value, in that $P(A) = E(I_{(A)})$. Use this to prove Boole's inequality: $P(A_1 \cup A_2 \cup \cdots \cup A_n) \leq P(A_1) + P(A_2) + \cdots + P(A_n)$. Try to think of a non-trivial example.**

We can write $I(A)$ as

$$I(A) = \begin{cases} 1 & \text{if } A \\ 0 & \text{if } A^c \end{cases}$$

where A occurs with probability of $P(A)$, and A^c occurs with probability $1 - P(A)$. This is equivalent to a Bern($P(A)$) random variable.

For the fundamental bridge, note that $E[I(A)] = P(A) \cdot 1 + (1 - P(A)) \cdot 0 = P(A)$.

To prove Boole's Inequality, we will show that $P(A_1 \cup A_2 \cup \cdots \cup A_n) \leq P(A_1) + P(A_2) + \cdots + P(A_n)$. Rewriting the probability values on the left and right hand sides as expected values of indicator variables, we know that

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = E[I_{A_1 \cup A_2 \cup \cdots \cup A_n}].$$

On the other hand,

$$P(A_1) + P(A_2) + \cdots + P(A_n) = E[I_{A_1}] + E[I_{A_2}] + \cdots + E[I_{A_n}] = E[I_{A_1} + I_{A_2} + \cdots + I_{A_n}]$$

Thus, to prove Boole's inequality, it suffices to show that

$$E[I_{A_1 \cup A_2 \cup \cdots \cup A_n}] \leq E[I_{A_1} + I_{A_2} + \cdots + I_{A_n}].$$

Note that $I_{A_1 \cup A_2 \cup \cdots \cup A_n}$ can only be either 0 or 1, since it is an indicator. In the former case, none of A_1 to A_n have occurred, so $I_{A_1} + I_{A_2} + \cdots + I_{A_n} = 0$. In the latter case, if $I_{A_1 \cup A_2 \cup \cdots \cup A_n} = 1$, then at least one of A_1 to A_n has occurred, so $I_{A_1} + I_{A_2} + \cdots + I_{A_n} \geq 1$.

It follows that

$$I_{A_1 \cup A_2 \cup \cdots \cup A_n} \leq I_{A_1} + I_{A_2} + \cdots + I_{A_n}.$$

Taking the expected value of both sides, we arrive at Boole's Inequality.

Example of Boole's Inequality: Let A represent the event of a fair coin flip; $A_i = 1$ if flip i is heads, and $A_i = 0$ if it is tails. Boole's Inequality tells us that if we flip the coin 5 times, the probability we flip at least one heads is less than or equal to 5 times the probability we flip a heads on any single coin flip. Thus, the probability we flip at least one heads is at most $5 \cdot \frac{1}{2} = \frac{5}{2}$; this is trivial as this is greater than 1.

b) **A special case of Boole's inequality occurs when the events all have the same probability. Suppose n graduates all throw their caps in the air and then retrieve a cap at random. Find an expression for the probability that none of the students retrieve their own cap (a derangement). Find the limit of this probability if the number of caps $n \to \infty$. *Hint: For $i = 1, ..., n$, let A_i represent the event that person i retrieves their own cap. Then $(A_1 \cup A_2 \cup ... \cup A_n)^c$ is the event that nobody ends up with their own cap.***

The probability of a derangement is

$$P(A_1 \cup A_2 \cup ... \cup A_n)^c = 1 - P(A_1 \cup A_2 \cup ... \cup A_n).$$

Taking the limit of this expression as $n \to \infty$ and using the Principle of Inclusion Exclusion, we know that this

$$\begin{aligned}\lim_{n\to\infty} 1 - P(A_1 \cup A_2 \cup ... \cup A_n) &= 1 - \left[\sum_{k=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k) - \dots\right] \\ &= \lim_{n\to\infty} 1 - \left[n \cdot \frac{1}{n} - \binom{n}{2}\left(\frac{1}{n} \cdot \frac{1}{n-1}\right) + \binom{n}{3}\left(\frac{1}{n} \cdot \frac{1}{n-1} \cdot \frac{1}{n-2}\right)\right]\end{aligned}$$

Simplifying, this becomes

$$\lim_{n\to\infty} 1 - \left[1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \dots\right] = \lim_{n\to\infty} \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \dots$$

Recall the Taylor Expansion of e^x: $e^x = \sum\limits_{k=0}^{\infty} \frac{x^k}{k!}$. We recognize the above expression as e^{-1}:

$$\begin{aligned}e^{-1} = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} &= \lim_{n\to\infty} \left(1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \dots\right) \\ &= \lim_{n\to\infty} \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \dots\end{aligned}$$

Thus,

$$\lim_{n\to\infty} P(A_1 \cup A_2 \cup ... \cup A_n)^c = \boxed{\frac{1}{e}}.$$

c) **In the caps example, let X represent the number of graduates who retrieve their own cap. Explain why the distribution of X is approximately Poisson(1) when n is large (see 1(c)). Compare the exact probabilities of $X = 0$ and $X = 1$ to the corresponding Poisson probabilities when $n = 5$.**

Note that the probability that a given graduate receives their own cap (assuming nothing about all other graduate cap arrangements) is $p = \frac{1}{n}$. Though we are essentially sampling without replacement when throwing all the graduate caps in the air and retrieving them at random, we know from problem 1(b) that the distribution X for the number of graduates who receive their own cap converges to Binom(n, p) when n is large.

Furthermore, note that $\lambda = np = n\frac{1}{n} = 1$ is fixed. Consequently, the limit of $P(X = x)$ as n is large, by problem 1(c), is simply Poisson(λ) = Poisson(1). We conclude that the distribution of X is approximately Poisson(1) when n is large.

Using the approximation $X \sim$ Poisson(1) for large n, we can approximate the probability of a derangement

$$f_x(x = 0, \lambda = 1) = \frac{1^0 e^{-1}}{0!} = \frac{1}{e}$$

and the probability that exactly one graduate gets their own cap:

$$f_x(x = 1, \lambda = 1) = \frac{1^1 e^{-1}}{1!} = \frac{1}{e}.$$

We can also compare these approximations with the exact probabilities of $X = 0$ and $X = 1$ for $n = 5$. Note that the probability of a derangement for $n = 5$ can be calculated using the formula from 2(b):

$$\begin{aligned} P(X = 0) &= 1 - \left[5 \cdot \frac{1}{5} - \binom{5}{2}\frac{1}{5} \cdot \frac{1}{4} + \binom{5}{3}\frac{1}{5} \cdot \frac{1}{4} \cdot \frac{1}{3} - \binom{5}{4}\frac{1}{5} \cdot \frac{1}{4} \cdot \frac{1}{3} \cdot \frac{1}{2} + \binom{5}{5}\frac{1}{5} \cdot \frac{1}{4} \cdot \frac{1}{3} \cdot \frac{1}{2} \cdot \frac{1}{1}\right] \\ &= \frac{11}{30} \approx 0.3\overline{6} \end{aligned}$$

Note that the probability that $X = 1$ can be calculated by "picking" the graduate to get their own cap (of which there are 5 possibilites), multiplying this by the probability that the chosen graduate gets their own hat ($\frac{1}{5}$), and then multiplying this by the probability of a derangement with $n = 4$ (each of the remaining four graduates does not get their own cap):

$$\begin{aligned} P(X = 1) &= 5 \cdot \frac{1}{5} \cdot P(\text{derangement for } n = 4 \text{ students}) \\ &= 1 - \left[4 \cdot \frac{1}{4} - \binom{4}{2}\frac{1}{4} \cdot \frac{1}{3} + \binom{4}{3}\frac{1}{4} \cdot \frac{1}{3} \cdot \frac{1}{2} - \binom{4}{4}\frac{1}{4} \cdot \frac{1}{3} \cdot \frac{1}{2}\frac{1}{1}\right] \\ &= \frac{3}{8} = 0.375 \end{aligned}$$

We can see that even for small n (in our case, $n = 5$), the expected probabilities approach $\frac{1}{e} \approx 0.36788$.