

Math 340 HW 3

Asa Royal (ajr74) [collaborators: none]

February 2, 2024

1. (i) In this setup, we are essentially modelling 120 flips of a p-coin. Each $\omega \in \Omega$ is a sequence of 120 true/false values. Thus,

$$\begin{aligned}\mathbb{P}(< 3 \text{ calls}) &= \mathbb{P}(0 \text{ calls}) + \mathbb{P}(1 \text{ calls}) + \mathbb{P}(2 \text{ calls}) \\ &= \binom{120}{0}(0.05)^0(0.95)^{120-0} + \binom{120}{1}(0.05)^1(0.95)^{120-1} + \binom{120}{2}(0.05)^2(0.95)^{120-2}\end{aligned}$$

- (ii) In this setup, $\omega \in \mathbb{W}$, and we can use a Poisson distribution with parameter $\lambda = 6$ to model the nubmer of arrivals in an hour. We choose $\lambda = 6$ so that $p = \lambda/n = 6/120 = 0.05$, which the last problem stated was the probability of an arrival during an interval. We have proved earlier that as $n \rightarrow \infty$, the $\mathbb{P}(A_k)$ follows the probability mass function stated in * if $\lambda = pn$. Once again, the probability of seeing < 3 calls during a 1 hour interval can be calculated by summing the probabilites of seeing zero, one, or two calls during the hour:

$$\mathbb{P}(< 3 \text{ calls}) = \mathbb{P}(0 \text{ calls}) + \mathbb{P}(1 \text{ calls}) + \mathbb{P}(2 \text{ calls})$$

* Per the probability mass function of the Poisson distribution, $\mathbb{P}(A_k) = \frac{\lambda^k}{k!}e^{-\lambda}$, so

$$\begin{aligned}\mathbb{P}(< 3 \text{ calls}) &= \frac{6^0}{0!}e^{-6} + \frac{6^1}{1!}e^{-6} + \frac{6^2}{2!}e^{-6} \\ &= e^{-6} + 6e^{-6} + 18e^{-6} \\ &= 25e^{-6} \\ &\approx 0.062\end{aligned}$$

2. (i) What is the probability that after n tosses of a p-coin, you have not seen heads?
This is the probability of seeing n tails in a row. Each p-coin flip is independent, so

$$\mathbb{P}(n \text{ tails}) = \prod_{k=1}^n \rho(\{T\}) = \prod_{k=1}^n (1-p) = \boxed{(1-p)^n}$$

- (ii) What is the probability that you toss $n-1$ tails and then heads occurs for the first time on the n th toss.
The first $n-1$ tosses are independent of the n th toss, so $\mathbb{P}(A)$, where A is the event described above, is $\mathbb{P}(n-1 \text{ tails}) * \mathbb{P}(\text{head})$. Per 2a, $\mathbb{P}(n-1 \text{ tails}) = (1-p)^{n-1}$. $\mathbb{P}(\text{head}) = p$ by construction. Thus,

$$\boxed{\mathbb{P}(A) = (1-p)^{n-1} * p}$$

3. (i) Let B_1 and B_2 be the events that your chosen coin lands heads on the 1st and 2nd tosses, respectively. Are these events independent?

Intuitively, we have reason to believe B_1 and B_2 are not independent. Assuming we have picked a coin with $p \neq 0.5$, seeing heads on the first toss slightly increases our expectation that $p > 0.5$, and makes us a bit more likely to expect heads on the next toss. Below, we mathematically show B_1 and B_2 are not independent by showing $\mathbb{P}(B_1 \cap B_2) \neq \mathbb{P}(B_1)\mathbb{P}(B_2)$.

We first calculate the probability that the chosen coin lands heads by summing the "probability paths" that lead to heads. Let X_i represent the event that we draw coin i .

$$\mathbb{P}(B_1) = \sum_{i=1}^k \mathbb{P}(B_1|X_i)\mathbb{P}(X_i)$$

Here, $\mathbb{P}(X_i) = 1/M$, since all coins are equally likely to be drawn. $\mathbb{P}(B_1|X_i)$ is given by p_i , per the problem setup. Thus

$$\mathbb{P}(B_1) = \mathbb{P}(B_2) = \frac{1}{M} \sum_{i=1}^k p_i$$

We calculate similarly $\mathbb{P}(B_1 \cap B_2)$, the chance that for any given coin, we flip a heads on the first and second toss. Since tosses are independent, $\mathbb{P}((B_1 \cap B_2)|X_i) = (p_i)^2$. Thus,

$$\begin{aligned} \mathbb{P}(B_1 \cap B_2) &= \sum_{i=1}^M \mathbb{P}((B_1 \cap B_2)|X_i) \mathbb{P}(X_i) \\ &= \sum_{i=1}^M (p_i)^2 \frac{1}{M} \end{aligned}$$

We now compare $\mathbb{P}(B_1 \cap B_2)$ to $\mathbb{P}(B_1) * \mathbb{P}(B_2)$.

$$\sum_{i=1}^M (p_i)^2 \frac{1}{M} \neq \frac{1}{M^2} \sum_{i=1}^M p_i \sum_{i=1}^M p_i$$

So $\boxed{B_1 \text{ and } B_2 \text{ are not independent events.}}$

- (ii) Suppose your coin lands heads k times (out of the n tosses). What is the probability that your coin is the j th coin.

Let X_j be the event that we have the j th coin. Let A_k be the event that we flip k heads. Per Bayes' Rule

$$\begin{aligned} \mathbb{P}(X_j|A_k) &= \frac{\mathbb{P}(A_k|X_j) * \mathbb{P}(X_j)}{\sum_{i=1}^M (\mathbb{P}(A_k|X_i) * \mathbb{P}(X_i))} \\ &= \frac{\left[\binom{n}{k} (p_j)^k (1 - p_j)^{n-k} \right] \frac{1}{M}}{\sum_{i=1}^M \left[\binom{n}{k} (p_i)^k (1 - p_i)^{n-k} \right] \frac{1}{M}} && \text{expand } \mathbb{P}(A_k) \text{ for p-coins. Sub } \mathbb{P}(X_j) = \frac{1}{M} \\ &= \frac{\binom{n}{k} (p_j)^k (1 - p_j)^{n-k}}{\sum_{i=1}^M \binom{n}{k} (p_i)^k (1 - p_i)^{n-k}} && \text{Cancel } \frac{1}{M} \text{ terms} \\ &= \frac{(p_j)^k (1 - p_j)^{n-k}}{\sum_{i=1}^M (p_i)^k (1 - p_i)^{n-k}} && \text{Cancel } \binom{n}{k} \text{ terms} \end{aligned}$$

- (iii) Suppose your coin lands heads n times in a row. You suspect it is the most biased coin (i.e. with the largest $p = p_M$). How large would n have to be in order that the probability of it being the most biased coin is at least ten times as large as the probability of it being the next most biased (i.e. the one with $p = p_{M-1}$)?

Let A_n be the event that we flip n heads in no tosses. Let X_j be the event we have selected the j th coin. Note that

$$\mathbb{P}(A_n|X_m) = \binom{n}{n} (p_m)^n (1 - p_m)^0 = (p_m)^n \quad (1)$$

$$\mathbb{P}(A_n|X_{m-1}) = \binom{n}{n} (p_{m-1})^n (1 - p_{m-1})^0 = (p_{m-1})^n \quad (2)$$

If we are at least ten times likelier to have coin m vs coin $m-1$, we can set $\mathbb{P}(X_m|A_n) \geq 10 * \mathbb{P}(X_{m-1}|A_n)$ and

solve to find the minimum satisfying n .

$$\begin{aligned}
& \mathbb{P}(X_m|A_n) \geq 10 * \mathbb{P}(X_{m-1}|A_n) && \text{problem setup} \\
& \frac{\mathbb{P}(A_n|X_m) * \mathbb{P}(X_m)}{\mathbb{P}(A_n)} \geq 10 \left(\frac{\mathbb{P}(A_n|X_{m-1}) * \mathbb{P}(X_{m-1})}{\mathbb{P}(A_n)} \right) && \text{Bayes' Rule} \\
& \mathbb{P}(A_n|X_m) \geq 10 * \mathbb{P}(A_n|X_{m-1}) && \text{cancel terms. } \mathbb{P}(X_m) = \mathbb{P}(X_{m-1}) = 1/m \\
& (p_m)^n \geq 10(p_{m-1})^n && \text{plugging in (1) and (2)} \\
& n \ln(p_m) \geq \ln(10) + n \ln(p_{m-1}) && \text{natural log both sides} \\
& n \ln(p_m) - n \ln(p_{m-1}) \geq \ln(10) \\
& n(\ln(p_m) - \ln(p_{m-1})) \geq \ln(10) && \text{factor out } n \\
& n \geq \frac{\ln(10)}{\ln(p_m) - \ln(p_{m-1})} \\
& \boxed{\geq \frac{\ln(10)}{\ln\left(\frac{p_m}{p_{m-1}}\right)}} && \text{diff of logs}
\end{aligned}$$

4. Meester 1.7.36

A pack contains m cards, labelled $1, 2, \dots, m$. The cards are dealt out in a random order, one by one. Given that the label of the k th card dealt is the largest of the first k cards, what is the probability that it is also the largest in the whole pack?

Let A be the event that the k th card is the largest of the first k cards. Let B be the event that the k th card is the largest of all M cards. We wish to calculate $\mathbb{P}(B|A)$. Per the the definition of conditional probability,

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)} \quad (3)$$

Note that $\mathbb{P}(A) = 1/k$, since each of the first k cards is equally likely to have the highest value. For the same reason, $\mathbb{P}(B) = 1/M$.

Note further that $\mathbb{P}(A|B) = 1$, because if a card has a larger value than all other $M - 1$ cards, it certainly has the largest value of the first k . Now, since the intersection operation is commutative,

$$\mathbb{P}(B \cap A) = \mathbb{P}(A \cap B) = \mathbb{P}(A|B)\mathbb{P}(B) = (1)\frac{1}{M} \quad (4)$$

Plugging (4) into (3), we see

$$\mathbb{P}(B|A) = \frac{\frac{1}{M}}{\frac{1}{k}} = \boxed{\frac{k}{M}}$$

5. Meester 1.7.38.

Let A_1, A_2, \dots be events. Show that

$$\mathbb{P}\left(\bigcap_{i=1}^n A_i\right) \geq \sum_{i=1}^n \mathbb{P}(A_i) - (n-1)$$

Proof.

$$\begin{aligned}
\mathbb{P}\left(\bigcap_{i=1}^n A_i\right) &= 1 - \mathbb{P}\left[\left(\bigcap_{i=1}^n A_i\right)^c\right] && \mathbb{P}(A) = 1 - \mathbb{P}(A^c) \\
&= 1 - \mathbb{P}\left(\bigcup_{i=1}^n (A_i)^c\right) && \text{DeMorgan's} \\
&\geq 1 - \sum_{i=1}^n \mathbb{P}(A_i)^c && \text{Union probability bound} \\
&= 1 - \sum_{i=1}^n (1 - \mathbb{P}(A_i)) && \mathbb{P}(A^c) = 1 - \mathbb{P}(A) \\
&= 1 - n + \sum_{i=1}^n \mathbb{P}(A_i) && \text{Split summation terms. } \sum_{i=1}^n 1 = n \\
&= \sum_{i=1}^n \mathbb{P}(A_i) - (n-1)
\end{aligned}$$

□

6. **Meester 1.7.23.** A fair coin is tossed n times. Let H_n and T_n denote the number of heads and tails among these n tosses. Show that, for any $\varepsilon > 0$ we have

$$P_n \left(-\varepsilon \leq \frac{H_n - T_n}{n} \leq \varepsilon \right) \rightarrow 1 \quad (5)$$

Proof. We first note that $H_n/n + T_n/n = 1$, so $T_n/n = 1 - H_n/n$. Thus $H_n/n - T_n/n = H_n/n - (1 - H_n/n) = 2H_n/n - 1$. Accordingly, we can rewrite the event in 5 as

$$\begin{aligned} & P_n \left(-\varepsilon \leq \frac{2H_n}{n} - 1 \leq \varepsilon \right) \\ &= P_n \left(\left(\frac{2H_n}{n} - 1 \geq -\varepsilon \right) \cap \left(\frac{2H_n}{n} - 1 \leq \varepsilon \right) \right) \\ &= P_n \left(\left(\frac{2H_n}{n} \geq 1 - \varepsilon \right) \cap \left(\frac{2H_n}{n} \leq 1 + \varepsilon \right) \right) \\ &= P_n \left(\left(H_n \geq \left(\frac{1}{2} - \frac{\varepsilon}{2} \right) n \right) \cap \left(H_n \leq \left(\frac{1}{2} + \frac{\varepsilon}{2} \right) n \right) \right) \\ &= P_n \left(\left(H_n \geq \left(\frac{1}{2} - \varepsilon' \right) n \right) \cap \left(H_n \leq \left(\frac{1}{2} + \varepsilon' \right) n \right) \right) \quad \text{where } \varepsilon' = \varepsilon/2 > 0 \end{aligned}$$

But this is the probability of the union of all events where we see $H_n \geq (1/2 - \varepsilon')n$ and $H_n \leq (1/2 + \varepsilon')n$. That is,

$$\mathbb{P} \left(\bigcup_{(1/2 - \varepsilon')n < k < (1/2 + \varepsilon')n} A_k \right)$$

And by the law of large numbers,

$$\mathbb{P} \left(\bigcup_{(1/2 - \varepsilon')n < k < (1/2 + \varepsilon')n} A_k \right) \geq 1 - 2e^{-\varepsilon'^2 n}$$

When $n \rightarrow \infty$, $1 - 2e^{-\varepsilon'^2 n} = 1 - \frac{2}{e^{\varepsilon'^2 n}} \rightarrow 1 - 0 = 1$, so the probability in (6) ≥ 1 . But we also know that the probability of the union of events in (6) must be ≤ 1 , since no event can have a probability > 1 . Thus, when $n \rightarrow \infty$

$$\begin{aligned} 1 &\leq \mathbb{P} \left(\bigcup_{(1/2 - \varepsilon')n < k < (1/2 + \varepsilon')n} A_k \right) \leq 1, \text{ and therefore,} \\ \mathbb{P} \left(\bigcup_{(1/2 - \varepsilon')n < k < (1/2 + \varepsilon')n} A_k \right) &\rightarrow 1 \end{aligned}$$

□

7. **Meester 2.1.15.** Prove the following: Let $B_1 \supseteq B_2 \supseteq B_3 \supseteq \dots$, and let $B = \bigcap_{i=1}^{\infty} B_i$. Then $\mathbb{P}(B) = \lim_{i \rightarrow \infty} \mathbb{P}(B_i)$

Proof. Since $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$ for an event A ,

$$\mathbb{P} \left(\bigcap_{i=1}^{\infty} B_i \right) = 1 - \mathbb{P} \left[\left(\bigcap_{i=1}^{\infty} B_i \right)^c \right] \quad (1)$$

$$= 1 - \mathbb{P} \left(\bigcup_{i=1}^{\infty} B_i^c \right) \quad (2)$$

Note that since $B_1 \supseteq B_2 \supseteq B_3 \supseteq \dots$, it follows that

$$B_1^c \subseteq B_2^c \subseteq B_3^c \subseteq \dots \quad (3)$$

We can thus apply Lemma 2.1.14 to (2) under the condition of (3), deducing that

$$\mathbb{P} \left(\bigcup_{i=1}^{\infty} B_i^c \right) = \lim_{i \rightarrow \infty} \mathbb{P}(B_i^c) \quad (4)$$

Plugging (4) into (1)/(2), we find

$$\begin{aligned}
\mathbb{P}(B) &= \mathbb{P}\left(\bigcap_{i=1}^{\infty} B_i\right) \\
&= 1 - \lim_{i \rightarrow \infty} \mathbb{P}(B_i^c) \\
&= 1 - \lim_{i \rightarrow \infty} (1 - \mathbb{P}(B_i)) \\
&= \lim_{i \rightarrow \infty} \mathbb{P}(B_i)
\end{aligned}$$

□

8. Suppose you are going to toss a fair coin 1000 times. What is the smallest value of m such that the probability of tossing fewer than m heads is at least 0.99.

Consider one form of the law of large numbers,

$$\mathbb{P}\left(\bigcup_{k \geq (\frac{1}{2} + \varepsilon)n} A_k\right) \leq e^{-\varepsilon^2 n} \quad (5)$$

Now substitute $m = (\frac{1}{2} + \varepsilon)n$ into (5), since we know we want to set up a calculation where \bigcup_{A_k} ranges up to $k = m$.

$$\mathbb{P}\left(\bigcup_{k \geq m} A_k\right) \leq e^{-\varepsilon^2 n} \quad (6)$$

Note that the complement of the set of event described on the LHS of (6), where $A_k \geq m$ is the set of events where $A_k < m$. Thus, (6) can be re-written as

$$\mathbb{P}\left(\bigcup_{k < m} A_k\right) \geq 1 - e^{-\varepsilon^2 n} \quad (7)$$

We now have on the LHS, the event (or union of events) described in the prompt. We wish to find m when the probability of that union is ≥ 0.99 , so we set $1 - e^{-\varepsilon^2 n} = 0.99$ and solve for ε .

$$\begin{aligned}
1 - e^{-\varepsilon^2 n} &= 0.99 \\
e^{-\varepsilon^2 n} &= 0.01 && \text{algebra} \\
-1000\varepsilon^2 &= \ln(0.01) && \text{natural log both sides} \\
\varepsilon^2 &= -\frac{\ln(0.01)}{1000} && \text{algebra} \\
\varepsilon &= \sqrt{-\frac{\ln(0.01)}{1000}} && \varepsilon > 0 \text{ so ignore negative root of } \varepsilon^2
\end{aligned}$$

We can use this ε value to solve for m , which we set equal to $(\frac{1}{2} + \varepsilon)n$:

$$\begin{aligned}
m &= \left(\frac{1}{2} + \varepsilon\right)n \\
&= \left(\frac{1}{2} + \sqrt{\frac{-\ln(0.01)}{1000}}\right)1000 \\
&= 500 + \sqrt{\frac{-\ln(0.01)}{1000}}1000 \\
&\approx 567.8
\end{aligned}$$

We conclude that the smallest value of m such that $\mathbb{P}(A_k < m) \geq 0.99$ is 568.