Homework 3

Graded

Student

NATHAN LEUNG

Total Points

30 / 30 pts

Question 1

1 — **5** / 5 pts

✔ **– 0 pts** Correct

– 1 pt Minor issue(s)

– 3 pts Major issue(s)

– 5 pts Completely incorrect

Question 2

2 — **5** / 5 pts

✔ **– 0 pts** Correct

– 1 pt Minor issue(s)

– 3 pts Major issue(s)

– 5 pts Completely incorrect

1 nice recursive function

Question 3

3 — **5** / 5 pts

✔ **– 0 pts** Correct

– 1 pt Minor issue(s)

– 3 pts Major issue(s)

– 5 pts Completely incorrect

Question 4

4 **5** / 5 pts

✔ **− 0 pts** Correct

− 1 pt Minor issue(s)

− 3 pts Major issue(s)

− 5 pts Completely incorrect

2 This is the first time I've ever seen TypeScript in a math problem set

Question 5

5 **5** / 5 pts

✔ **− 0 pts** Correct

− 1 pt Minor issue(s)

− 3 pts Major issue(s)

− 5 pts Completely incorrect

Question 6

6 **5** / 5 pts

✔ **− 0 pts** Correct

− 1 pt Minor issue(s)

− 3 pts Major issue(s)

− 5 pts Completely incorrect

Question assigned to the following page: 1

MATH 170A Homework 3

Nathan Leung, UID 005835316

26 January 2024

Problem 1

Here again are the probabilities from the last problem on HW2:

M	A	E
T	1/16	7/32
N	3/32	1/8

F	A	E
T	1/8	1/8
N	3/16	1/16

Show that F and T are not independent, but are independent conditioned on A. Does this information tend to suggest or to refute discrimination?

Let $\mathbb{Q}(\cdot) = \mathbb{P}(\cdot \mid A)$, the conditional probability measure given A.

By the definition of independence, we need to show the following two statements:

- $\mathbb{P}(F \cap T) \neq \mathbb{P}(F) \cdot \mathbb{P}(T)$
- $\mathbb{Q}(F \cap T) = \mathbb{Q}(F) \cdot \mathbb{Q}(T)$

For the first statement, we have:

$$\begin{aligned}
\mathbb{P}(F \cap T) &= \frac{1}{8} + \frac{1}{8} = \frac{2}{8} = \frac{1}{4} \\
\mathbb{P}(F) &= \frac{1}{8} + \frac{1}{8} + \frac{3}{16} + \frac{1}{16} = \frac{8}{16} = \frac{1}{2} \\
\mathbb{P}(T) &= \frac{1}{16} + \frac{3}{32} + \frac{1}{8} + \frac{1}{8} = \frac{17}{32} \\
\mathbb{P}(F) \cdot \mathbb{P}(T) &= \frac{1}{2} \cdot \frac{17}{32} = \frac{17}{64}
\end{aligned}$$

Since $\mathbb{P}(F) \cdot \mathbb{P}(T) = \frac{17}{64} \neq \frac{16}{64} = \frac{1}{4} = \mathbb{P}(F \cap T)$, the events F and T are not independent.

On the other hand, for the second statement, we have

$$\mathbb{Q}(F \cap T) = \mathbb{P}(F \cap T \mid A) = \frac{\mathbb{P}(F \cap T \cap A)}{\mathbb{P}(A)} = \frac{\frac{1}{8}}{\frac{1}{8} + \frac{3}{16} + \frac{1}{16} + \frac{3}{32}} = \frac{\frac{4}{32}}{\frac{15}{32}} = \frac{4}{15}$$

Questions assigned to the following page: 1 and 2

$$\mathbb{Q}(F) = \mathbb{P}(F \mid A) = \frac{\mathbb{P}(F \cap A)}{\mathbb{P}(A)} = \frac{\frac{1}{8} + \frac{3}{16}}{\frac{15}{32}} = \frac{\frac{10}{32}}{\frac{15}{32}} = \frac{10}{15}$$
$$\mathbb{Q}(T) = \mathbb{P}(T \mid A) = \frac{\mathbb{P}(T \cap A)}{\mathbb{P}(A)} = \frac{\frac{1}{8} + \frac{1}{16}}{\frac{15}{32}} = \frac{\frac{6}{32}}{\frac{15}{32}} = \frac{6}{15}$$
$$\mathbb{Q}(F) \cdot \mathbb{Q}(T) = \frac{10}{15} \cdot \frac{6}{15} = \frac{60}{225} = \frac{4}{15}$$

Since $\mathbb{Q}(F) \cdot \mathbb{Q}(T) = \frac{4}{15} = \mathbb{Q}(F \cap T)$, the events F and T are indeed independent conditioned on A.

This information refutes discrimination within the Arts (A) department; the gender of the professor does not affect their probability of promotion in A. However, we still need to check what is happening in E too.

Problem 2

I toss a fair coin many times. Each trial is independent of the others. For each pair of integers $1 \leq k \leq n$, we define

$$\begin{aligned} H_k &= \{ \text{ the } k\text{th throw lands heads } \} \\ Y_{0,n} &= \{ \text{ the first } n \text{ trials contain no tails } \} \\ Y_{k,n} &= \{ \text{ the } k\text{th throw is the last (largest } k\text{) tail among the first } n \text{ trials } \} \\ R_n &= \{ \text{ the first } n \text{ trials contain a run of 4 consecutive heads } \} \end{aligned}$$

Part (a)

Show that all sets $Y_{\ell,n}$ are in the σ-algebra generated by $H_1, H_2, ...$

First, fix an arbitrary $\ell, n \in \mathbb{N}$. If $\ell > n$, the event is empty, which is in the σ-algebra. So, assume $\ell \leq n$. From above, we know that

$$Y_{\ell,n} = \{ \text{ the } \ell\text{th throw is the last (largest } k\text{) tail among the first } n \text{ trials } \}$$

Essentially, this is the set of all n-toss outcomes where we disregard the first $(\ell - 1)$ tosses, the ℓth toss is tails, and the remaining $(\ell + 1)$st to nth tosses are heads. More formally, if we consider the sample space to be

$$\Omega = \{H, T\}^{\mathbb{N}} = \{(H, H, T, T, ...), ...\}$$

where the ith coordinate is the outcome of the ith toss, the event could be written

$$Y_{\ell,n} = \{H, T\}^{\ell-1} \times \{T\} \times \{H\}^{n-\ell} \times \{H, T\}^{\mathbb{N}}$$

To show that $Y_{\ell,n}$ is in the σ-algebra generated by the H_i for $i \in \mathbb{N}$, we need to show that $Y_{\ell,n}$ can be generated from the H_i using the operations of complement and countable union. Indeed,

Question assigned to the following page: 2

Nathan Leung, UID 005835316
MATH 170A Homework 3
26 January 2024

---

note that H_ℓ^{C} is the event that the ℓth toss is tails and $H_{\ell+1}, ..., H_n$ are the events that the $(\ell+1)$st to nth tosses are heads. Thus, we can write

$$Y_{\ell,n} = H_\ell^{\mathsf{C}} \cap H_{\ell+1} \cap ... \cap H_n$$

Applying De Morgan's Laws, we get

$$Y_{\ell,n} = (H_\ell \cup H_{\ell+1}^{\mathsf{C}} \cup ... \cup H_n^{\mathsf{C}})^{\mathsf{C}}$$

which is an expression for $Y_{\ell,n}$ in terms of the H_is, using only countable unions and complements. Thus, $Y_{\ell,n}$ is in the σ-algebra generated by the H_i. Since ℓ and n were arbitrary, this applies to all such ℓ and n.

Part (b)

What are $\mathbb{P}(H_k)$, $\mathbb{P}(Y_{0,n})$, and $\mathbb{P}(Y_{k,n})$ for each $1 \leq k \leq n$?

$\boxed{\mathbb{P}(H_k) = \frac{1}{2}}$, since the kth toss is either heads or tails. The coin is fair, so each outcome is equiprobable.

$\boxed{\mathbb{P}(Y_{0,n}) = \left(\frac{1}{2}\right)^n}$, since if the first n trials contain no tails, then the first n trails must be all heads. That is, $Y_{0,n} = H_1 \cap H_2 \cap ... \cap H_n$. Since the coin is fair and each trial is independent of the others, we have

$$\begin{aligned}
\mathbb{P}(Y_{0,n}) &= \mathbb{P}(H_1 \cap H_2 \cap ... \cap H_n) \\
&= \mathbb{P}(H_1)\mathbb{P}(H_2)...\mathbb{P}(H_n) \\
&= \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\cdots\left(\frac{1}{2}\right) \\
&= \left(\frac{1}{2}\right)^n
\end{aligned}$$

$\boxed{\mathbb{P}(Y_{k,n}) = \left(\frac{1}{2}\right)^{n-k+1}}$. From **Part (a)**, we know that $Y_{\ell,n} = H_\ell^{\mathsf{C}} \cap H_{\ell+1} \cap ... \cap H_n$. Since the coin is fair and each trial is independent of the others, we have

$$\begin{aligned}
\mathbb{P}(\mathbb{P}(Y_{k,n}) &= \mathbb{P}(H_k^{\mathsf{C}} \cap H_{k+1} \cap ... \cap H_n) \\
&= \mathbb{P}(H_k^{\mathsf{C}})\underbrace{\mathbb{P}(H_{k+1})...\mathbb{P}(H_n)}_{n-k \text{ terms}} \\
&= \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)^{n-k} \\
&= \left(\frac{1}{2}\right)^{n-k+1}
\end{aligned}$$

Question assigned to the following page: 2

Nathan Leung, UID 005835316
MATH 170A Homework 3
26 January 2024

Part (c)

Show that H_k and $Y_{\ell,n}$ are independent if $k < \ell \leq n$ but not if $\ell \leq k$.

First, suppose $k < \ell \leq n$. We want to show that H_k and $Y_{\ell,n}$ are independent. That is, we want to show

$$\mathbb{P}(H_k) \cdot \mathbb{P}(Y_{\ell,n}) = \mathbb{P}(H_k \cap Y_{\ell,n})$$

In words, we know that

$$\mathbb{P}(H_k \cap Y_{\ell,n}) = \{ \text{ } k\text{th toss is heads, and } \ell\text{th toss is last tail among first } n \text{ tosses } \}$$

$k < \ell$ by assumption, so this event occurs when the kth toss is heads, the ℓth toss is tails, and the remaining $n - \ell$ tosses are all heads. It does not matter how the first $(k-1)$ tosses turn out. The probability of the event occurring is thus

$$\mathbb{P}(H_k \cap Y_{\ell,n}) = \underbrace{\frac{1}{2}}_{k\text{th toss is heads}} \cdot \underbrace{\frac{1}{2}}_{\ell\text{th toss is heads}} \cdot \underbrace{\left(\frac{1}{2}\right)^{n-\ell}}_{\text{last } (n-\ell) \text{ tosses are heads}} = \left(\frac{1}{2}\right)^{n-\ell+2x}$$

At the same time, from **Part (b)** we have

$$\mathbb{P}(H_k) \cdot \mathbb{P}(Y_{\ell,n}) = \frac{1}{2} \cdot \left(\frac{1}{2}\right)^{n-\ell+1} = \left(\frac{1}{2}\right)^{n-\ell+2}$$

Thus $\mathbb{P}(H_k \cap Y_{\ell,n}) = \mathbb{P}(H_k) \cdot \mathbb{P}(Y_{\ell,n})$ and so the events are indeed independent.

Next, suppose $\ell \leq k$. Again, in words, we know that

$$\mathbb{P}(H_k \cap Y_{\ell,n}) = \{ \text{ } k\text{th toss is heads, and } \ell\text{th toss is last tail among first } n \text{ tosses } \}$$

We have two cases. If $\ell = k$, then the $(k = \ell)$th toss must be both heads and tails. This is impossible! Thus, $\mathbb{P}(H_k \cap Y_{\ell,n}) = 0$. From **Part (b)**, we know the individual events have nonzero probability, so their product must have nonzero probability. Thus, $\mathbb{P}(H_k \cap Y_{\ell,n}) \neq \mathbb{P}(H_k) \cdot \mathbb{P}(Y_{\ell,n})$ in this case and so the events are not independent.

In the other case, $\ell < k$. But if the ℓth toss is the last tail among the first n tosses, and the kth toss comes afterwards, the kth toss must necessarily be heads. That is,

$$H_k \cap Y_{\ell,n} = Y_{\ell,n}$$

but

$$\mathbb{P}(H_k \cap Y_{\ell,n}) = \mathbb{P}(Y_{\ell,n}) = \left(\frac{1}{2}\right)^{n-\ell+1}$$
$$\mathbb{P}(H_k) \cdot \mathbb{P}(Y_{\ell,n}) = \frac{1}{2} \cdot \left(\frac{1}{2}\right)^{n-\ell+1} = \left(\frac{1}{2}\right)^{n-\ell+2}$$

i.e. $\mathbb{P}(H_k \cap Y_{\ell,n}) \neq \mathbb{P}(H_k) \cdot \mathbb{P}(Y_{\ell,n})$ and so the events are not independent.

Note above that we implicitly assumed that $k \leq n$, but this is fine since it is given in the problem that the events are only defined for $1 \leq k \leq n$.

In both cases when $\ell \leq k$, $\mathbb{P}(H_k \cap Y_{\ell,n}) \neq \mathbb{P}(H_k) \cdot \mathbb{P}(Y_{\ell,n})$ so the events are not independent.

Question assigned to the following page: 2

Nathan Leung, UID 005835316
MATH 170A Homework 3
26 January 2024

Part (d)

Provide a formula for $\mathbb{P}(R_n^{\complement})$ by using a partition built from the Y-sets and the Partition Theorem (that helps with what follows).

For a fixed $n \in \mathbb{N}$, we claim that the Y-sets $Y_{0,n}, Y_{1,n}, Y_{2,n}, ..., Y_{n,n}$ form a partition of $\Omega = \{H, T\}^{\mathbb{N}}$. To prove this, we need to show that the $Y_{i,n}$ for $0 \leq i \leq n$ are disjoint and form the entire sample space when their union is taken.

Disjointness is straightforward: suppose for contradiction there was an elementary event ω in two sets $Y_{i,n}$ and $Y_{j,n}$ for $i \neq j$ (without loss of generality, $i < j$). By the definition of $Y_{i,n}$, it must be the case that ω's last tail occurs on toss i. Since $j > i$, toss j must be a head. But by the definition of $Y_{j,n}$, ω's last tail occurs on toss j. So toss j must be a tail as well, a contradiction. Thus, no such elementary event ω can exist. Since i and j were arbitrary, this applies to all such Y-sets, and hence they are all disjoint.

To show that their union covers the sample space, let $\omega \in \Omega$ be an arbitrary elementary event. There are two cases: it either has a tail in the first n trials, or it does not have a tail in the first n trials. If it has no tails in the first n trials, then $\omega \in Y_{0,n}$. If it has a tail in the first n trials, then consider the position of the most recent (i.e. last) tail among the first n trials. If it is in position i (with $1 \leq i < n$, since there exists at least one tail), then $\omega \in Y_{i,n}$. Since any $\omega \in \Omega$ is in one of the Y-sets, they indeed cover the sample space.

Thus, the $Y_{i,n}$ for $0 \leq i \leq n$ are indeed disjoint and form the entire sample space when their union is taken. Hence, they form a partition. Applying the Partition Theorem, we get the following expression for $\mathbb{P}(R_n^{\complement})$:

$$
\begin{aligned}
\mathbb{P}(R_n^{\complement}) = {} & \mathbb{P}(R_n^{\complement} \mid Y_{0,n})\mathbb{P}(Y_{0,n}) + \mathbb{P}(R_n^{\complement} \mid Y_{1,n})\mathbb{P}(Y_{1,n}) \\
& + \mathbb{P}(R_n^{\complement} \mid Y_{2,n})\mathbb{P}(Y_{2,n}) + \mathbb{P}(R_n^{\complement} \mid Y_{3,n})\mathbb{P}(Y_{3,n}) \\
& + \cdots \\
& + \mathbb{P}(R_n^{\complement} \mid Y_{n-3,n})\mathbb{P}(Y_{n-3,n}) + \mathbb{P}(R_n^{\complement} \mid Y_{n-2,n})\mathbb{P}(Y_{n-2,n}) \\
& + \mathbb{P}(R_n^{\complement} \mid Y_{n-1,n})\mathbb{P}(Y_{n-1,n}) + \mathbb{P}(R_n^{\complement} \mid Y_{n,n})\mathbb{P}(Y_{n,n})
\end{aligned}
$$

Part (e)

Deduce the recursion formula

$$\mathbb{P}(R_n^{\complement}) = \frac{1}{2}\mathbb{P}(R_{n-1}^{\complement}) + \frac{1}{4}\mathbb{P}(R_{n-2}^{\complement}) + \frac{1}{8}\mathbb{P}(R_{n-3}^{\complement}) + \frac{1}{16}\mathbb{P}(R_{n-4}^{\complement})$$

at least if $n \geq 5$.

Note that for $k = 0, ..., n-4$, the event $Y_{k,n}$ means that the last $n, ..., 4$ tosses (respectively) are all heads. So $\mathbb{P}(R_n^{\complement} \mid Y_{k,n})$, the probability that there is no consecutive run of four heads given $Y_{k,n}$, must be 0. We must have a run of at least four heads if the last $4, ..., n$ tosses must all be heads! Thus, those terms disappear from the expression from **Part (d)** and we are left with

$$
\begin{aligned}
\mathbb{P}(R_n^{\complement}) = {} & \mathbb{P}(R_n^{\complement} \mid Y_{n-3,n})\mathbb{P}(Y_{n-3,n}) + \mathbb{P}(R_n^{\complement} \mid Y_{n-2,n})\mathbb{P}(Y_{n-2,n}) \\
& + \mathbb{P}(R_n^{\complement} \mid Y_{n-1,n})\mathbb{P}(Y_{n-1,n}) + \mathbb{P}(R_n^{\complement} \mid Y_{n,n})\mathbb{P}(Y_{n,n})
\end{aligned}
$$

Question assigned to the following page: 2

Looking more closely at each term, note that when $Y_{n-3,n}$, $Y_{n-2,n}$, $Y_{n-1,n}$, and $Y_{n,n}$ occur, they fix the last $4, 3, 2$, and 1 toss(es) (respectively) out of the first n trials. That is, if $Y_{n-3,n}$ occurs, the the fourth-to-last toss (of the first n trials — i.e. the $(n-3)$rd toss) must be tails while the last 3 tosses of the first n trials must be heads. Similarly, if $Y_{n-2,n}$ occurs, the $(n-2)$nd toss must be tails and the $(n-1)$st and nth tosses must be heads; if $Y_{n-1,n}$ occurs, the $(n-1)$st toss must be tails and the nth toss must be heads; if $Y_{n,n}$ occurs, the nth toss must be tails.

Thus, $\mathbb{P}(R_n^{\mathsf{C}} \mid Y_{n-3,n})$ is really just the probability that there is not a run of four heads in the first $n-4$ tosses; the $(n-3)$rd, $(n-2)$nd, $(n-1)$st, and nth tosses are all determined by $Y_{n-3,n}$. That is, $\mathbb{P}(R_n^{\mathsf{C}} \mid Y_{n-3,n}) = R_{n-4}^{\mathsf{C}}$. Similarly, $\mathbb{P}(R_n^{\mathsf{C}} \mid Y_{n-2,n})$ is the probability that there is not a run of four heads in the first $n-3$ tosses and so $\mathbb{P}(R_n^{\mathsf{C}} \mid Y_{n-2,n}) = R_{n-3}^{\mathsf{C}}$. Likewise, $\mathbb{P}(R_n^{\mathsf{C}} \mid Y_{n-1,n}) = R_{n-2}^{\mathsf{C}}$ and $\mathbb{P}(R_n^{\mathsf{C}} \mid Y_{n,n}) = R_{n-1}^{\mathsf{C}}$.

Substituting with our work above, the expression becomes

$$\begin{aligned}\mathbb{P}(R_n^{\mathsf{C}}) &= R_{n-4}^{\mathsf{C}}\mathbb{P}(Y_{n-3,n}) + R_{n-3}^{\mathsf{C}}\mathbb{P}(Y_{n-2,n}) \\ &\quad + R_{n-2}^{\mathsf{C}}\mathbb{P}(Y_{n-1,n}) + R_{n-1}^{\mathsf{C}}\mathbb{P}(Y_{n,n})\end{aligned}$$

Since $\mathbb{P}(Y_{n-3,n})$ fixes the last 4 tosses, it occurs $\left(\frac{1}{2}\right)^4 = \frac{1}{16}$ of the time. Similarly,

$$\begin{aligned}\mathbb{P}(Y_{n-2,n}) &= \left(\frac{1}{2}\right)^3 = \frac{1}{8} \\ \mathbb{P}(Y_{n-1,n}) &= \left(\frac{1}{2}\right)^2 = \frac{1}{4} \\ \mathbb{P}(Y_{n,n}) &= \left(\frac{1}{2}\right)^1 = \frac{1}{2}\end{aligned}$$

Thus, we get

$$\begin{aligned}\mathbb{P}(R_n^{\mathsf{C}}) &= R_{n-4}^{\mathsf{C}}\left(\frac{1}{16}\right) + R_{n-3}^{\mathsf{C}}\left(\frac{1}{8}\right) \\ &\quad + R_{n-2}^{\mathsf{C}}\left(\frac{1}{4}\right) + R_{n-1}^{\mathsf{C}}\left(\frac{1}{2}\right)\end{aligned}$$

Reordering gives us

$$\mathbb{P}(R_n^{\mathsf{C}}) = \frac{1}{2}R_{n-1}^{\mathsf{C}} + \frac{1}{4}R_{n-2}^{\mathsf{C}} + \frac{1}{8}R_{n-3}^{\mathsf{C}} + \frac{1}{16}R_{n-4}^{\mathsf{C}}$$

which is the recurrence we wanted to show.

Question assigned to the following page: 2

Nathan Leung, UID 005835316
MATH 170A Homework 3
26 January 2024

Part (f)

Tabulate enough such probabilities to determine the smallest n for which $\mathbb{P}(R_n^{\mathsf{C}}) \leq \frac{1}{2}$.

According to the following code, after $\boxed{n = 22 \text{ tosses}}$ the probability $\mathbb{P}(R_n^{\mathsf{C}})$ will be less than or equal to $\frac{1}{2}$.

```typescript
import _ from 'lodash';

function probabilityRc(nTosses: number): number {
  if (nTosses === 4) {
    // 4 heads occurs only 1/16 of the time
    return 15 / 16;
  }
  if (nTosses <= 3) {
    // 4 heads can never occur
    return 1;
  }
  return (
    (1 / 2) * probabilityRc(nTosses - 1) +
    (1 / 4) * probabilityRc(nTosses - 2) +
    (1 / 8) * probabilityRc(nTosses - 3) +
    (1 / 16) * probabilityRc(nTosses - 4)
  );
}

const memoizedProbabilityRc = _.memoize(probabilityRc);

// Calculate smallest number of tosses for which
// the probability of not getting any runs of four
// heads is less than half (as we toss more, the
// probability of not getting any runs of four heads
// should go down)
let leastTosses = 0;
while (memoizedProbabilityRc(leastTosses) > 1 / 2) {
  console.debug(leastTosses, memoizedProbabilityRc(leastTosses));
  leastTosses += 1;
}
console.log(
  `You need at least ${leastTosses} tosses for the ` +
  `probability of not getting any runs of four ` +
  `heads to be less than half.`
);
```

Runnable TypeScript Code

Question assigned to the following page: 3

Nathan Leung, UID 005835316
MATH 170A Homework 3
26 January 2024

Problem 3

Suppose my knowledge/ignorance of the number of branches of a certain store (in my city) is given by the following probability law:

$$\mathbb{P}(k \text{ branches}) = (1-p)p^k \qquad \text{where } 0 < p < 1 \text{ is fixed and } k = 0, 1, 2, 3, ...$$

Part (a)

If I subsequently discover that they have at least 7 branches (e.g. I walk into a store and it says "Branch #7") what new probability law describes my revised knowledge?

We can apply the definition of conditional probability:

$$\mathbb{P}(k \text{ branches} \mid \text{at least 7 branches}) = \frac{\mathbb{P}(k \text{ branches} \cap \text{at least 7 branches})}{\mathbb{P}(\text{at least 7 branches})}$$

In the denominator, we have

$$\begin{aligned}
\mathbb{P}(\text{at least 7 branches}) &= \mathbb{P}(7 \text{ branches}) + \mathbb{P}(8 \text{ branches}) + \mathbb{P}(9 \text{ branches}) + ... \\
&= \sum_{i=7}^{\infty} \mathbb{P}(i \text{ branches}) \\
&= \sum_{i=7}^{\infty} (1-p)p^i \\
&= (1-p)\sum_{i=7}^{\infty} p^i \\
&= (1-p)\left(\frac{p^7}{1-p}\right) && \text{Geometric series sum} \\
&= p^7
\end{aligned}$$

Note that the numerator is piecewise-defined:

$$\mathbb{P}(k \text{ branches} \cap \text{at least 7 branches}) = \begin{cases} 0 & k < 7 \\ (1-p)p^k & k \geq 7 \end{cases}$$

Thus, our new probability law must also be piecewise-defined. In the case where $k < 7$, we get

$$\mathbb{P}(k \text{ branches} \mid \text{at least 7 branches}) = \frac{0}{p^7} = 0$$

In the case where $k \geq 7$, we get

$$\mathbb{P}(k \text{ branches} \mid \text{at least 7 branches}) = \frac{(1-p)p^k}{p^7}$$

Question assigned to the following page: 3

$$= (1-p)p^{k-7}$$

Thus our final new probability law is

$$\mathbb{P}(k \text{ branches} \mid \text{at least 7 branches}) = \begin{cases} 0 & k < 7 \\ (1-p)p^{k-7} & k \geq 7 \end{cases}$$

Part (b)

What is the probability of this observation if there are k stores? What number of stores makes the observation most likely?

We have $\Omega = \{(k, \ell) \mid k \in \mathbb{N}, 0 \leq \ell \leq k\}$ where k is the number of branches and ℓ is the number of the branch we walk into. We can define events

$$E_{k \text{ branches}} = \{(k, \ell) \mid 0 \leq \ell \leq k\}$$
$$E_{\text{walk into branch } \ell} = \{(k, \ell) \mid k \geq \ell\}$$

Applying the definition of conditional probability, we have

$$\mathbb{P}(E_{\text{walk into branch 7}} \mid E_{k \text{ branches}}) = \frac{\mathbb{P}(E_{\text{walk into branch 7}} \cap E_{k \text{ branches}})}{\mathbb{P}(E_{k \text{ branches}})}$$

When $0 \leq k < 7$, the event $(E_{\text{walk into branch 7}} \mid E_{k \text{ branches}})$ is the empty set and the probability is 0.

For $k \geq 7$, the event $(E_{\text{walk into branch 7}} \mid E_{k \text{ branches}})$ corresponds to the singleton $\{(k, 7)\}$; $E_{k \text{ branches}})$ corresponds to the k-element set $\{(k, 1), (k, 2), ..., (k, k)\}$. Assuming (reasonably) that the probability of walking into each branch is equal, we get

$$\mathbb{P}(E_{\text{walk into branch 7}} \mid E_{k \text{ branches}}) = \boxed{\frac{1}{k}}$$

for the probability of the observation if there are k stores.

This value decreases as k increases, so it is maximized when $\boxed{k = 7}$ (the minimum allowable choice).

Part (c)

What number of stores has the greatest posterior probability?

In **Part (a)**, we determined the posterior probability law:

$$\mathbb{P}(k \text{ branches} \mid \text{at least 7 branches}) = \begin{cases} 0 & k < 7 \\ (1-p)p^{k-7} & k \geq 7 \end{cases}$$

We can eliminate all $k < 7$ since for $0 < p < 1$, $(1-p)p^{k-7}$ will be nonzero and hence greater.

Questions assigned to the following page: 3 and 4

We claim that $(1-p)p^{k-7}$ is a decreasing function on $\mathbb{N}$. Indeed, proceeding by induction, in the base case $k=7$, we have

$$\begin{aligned}(1-p)p^{7-7} &= (1-p)p^0 = 1-p \\ (1-p)p^{8-7} &= (1-p)p^1 = (1-p)p < 1-p && \text{Since } 0<p<1\end{aligned}$$

Assume the function is decreasing for all steps $n<N$. We want to show it is decreasing from N to $N+1$. We have

$$(1-p)p^{N+1-7} = (1-p)p^{N-7}p < (1-p)p^{N-7} \qquad \text{Since } 0<p<1$$

Thus, $(1-p)p^{k-7}$ is indeed a decreasing function on $\mathbb{N}$, and hence is maximized at the smallest allowable value we can choose, which is $\boxed{k=7}$ stores.

Part (d)

Which method for estimating the total number of stores is influenced by the choice of prior distribution: maximum likelihood (b) or maximum a posteriori probability (c)?

The prior distribution $\mathbb{P}(k \text{ branches}) = (1-p)p^k$ does not show up at all in the maximum likelihood (b) esimation; it only shows up in the maximum a posteriori probability (c) estimation.

Thus, it is the $\boxed{\text{maximum a posteriori probability}}$ method which is influenced by the choice of prior distribution.

Problem 4

Each of n people are randomly and independently assigned a number from the set $\{1,2,3,...,365\}$. We will call this number their birthday.

Part (a)

What is the probability that no two people share a birthday?

We can reframe this problem as calculating the probability of choosing n distinct (i.e. no shared birthdays) numbers from the set $\{1,2,3,...,365\}$ when we are choosing with replacement.

Once we pick the first number, there is a $\frac{364}{365}$ probability we pick a distinct second number. Subsequently, for the third number, there is a $\frac{363}{365}$ probability we pick a distinct number. We can continue on in this fashion until we reach the 365th choice, for which there is a $\frac{1}{365}$ probability we pick a number distinct from the other 364 numbers. Once we have to pick 366 or more numbers out of the set, the Pigeonhole Principle tells us that the probability of choosing all distinct numbers is 0.

In general, for $1 \le n \le 365$ people, the probability is given by the product

$$\frac{364}{365} \cdot \frac{363}{365} \cdot \dots \cdot \frac{365-n+1}{365}$$

And for $n>365$, the probability is 0.

Questions assigned to the following page: 4 and 5

Nathan Leung, UID 005835316
MATH 170A Homework 3
26 January 2024

Part (b)

Use a computer or calculator to evaluate your answer as a decimal for $n = 22$ and $n = 23$.

According to the following code, when $n = 22$, we get $\mathbb{P}(\text{no two people share a birthday}) = 0.5243046923374497$; when $n = 23$, we get $\mathbb{P}(\text{no two people share a birthday}) = 0.4927027656760144$.

```
function probabilityNoSharedBirthday(nPeople: number): number {
  if (nPeople > 365) {
    // Pigeonhole principle
    return 0;
  }

  let probability = 1;
  for (let i = 1; i < nPeople; i += 1) {
    probability *= (365 - i) / 365
  }
  return probability;
}

console.log(probabilityNoSharedBirthday(22));
console.log(probabilityNoSharedBirthday(23));
```

Runnable TypeScript Code

Problem 5

I repeatedly attempt the same task. My probability of success on the kth attempt is $p \in (0, 1)$, independent of the outcome of all previous attempts.

Part (a)

Explain why each trial is also independent of all future attempts.

Each trial is also independent of all future attempts because independence is symmetric. More precisely, let A and B be the events corresponding to success on a trial in the past and success on a trial in the future, respectively. Since attempts are independent of the outcome of previous attempts, the definition of independence tells us that

$$\mathbb{P}(B \mid A) = \mathbb{P}(B)$$

i.e.

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

so

$$\mathbb{P}(B)\mathbb{P}(A) = \mathbb{P}(B \cap A)$$

But intersection and multiplication are symmetric, i.e. we can write

Question assigned to the following page: 5

$$\mathbb{P}(A)\mathbb{P}(B) = \mathbb{P}(A \cap B)$$

so

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

which by the definition of conditional probability means

$$\mathbb{P}(A) = \mathbb{P}(A \mid B)$$

and so A (the past attempt) is also independent of B (the future attempt).

Part (b)

What is the probability that my first success occurs on the kth trial?

For the first success to occur on the kth trial, the first $(k-1)$ trials must fail and the kth must succeed. The probability of this occurring is

$$\underbrace{(1-p)^{k-1}}_{k-1 \text{ failures}} \cdot p$$

Part (c)

Given that my first seven attempts failed, what is the probability that I will need to make k more attempts before succeeding?

Since probability of success is independent of previous attemptss, the probability is the same as in **Part (b)**, i.e. $(1-p)^{k-1} \cdot p$.

Part (d)

Given that my second success occurred on the nth trial ($n \geq 2$), what is the probability that my first success occurred on the kth trial?

In symbols, we are being asked to calculate

$$\mathbb{P}(\text{first success occurred on the } k\text{th trial} \mid \text{second success occurred on the } n\text{th trial})$$

Applying the definition of conditional probability, we have

$$\begin{aligned}&\mathbb{P}(\text{first success occurred on the } k\text{th trial} \mid \text{second success occurred on the } n\text{th trial}) \\ &\quad = \frac{\mathbb{P}(\text{first success occurred on the } k\text{th trial} \cap \text{second success occurred on the } n\text{th trial})}{\mathbb{P}(\text{second success occurred on the } n\text{th trial})}\end{aligned}$$

In the numerator, we have

Question assigned to the following page: 5

$$\mathbb{P}(\text{first success occurred on the } k\text{th trial} \cap \text{second success occurred on the } n\text{th trial})$$

This occurs exactly when the first $(k-1)$ trials are failures, the kth trial is a success, the next $(n-k-1)$ trials are failures, and the nth trial is a success. Thus, the probability of this occurring is

$$\underbrace{(1-p)^{k-1}}_{k-1 \text{ failures}} \cdot p \cdot \underbrace{(1-p)^{n-k-1}}_{n-k-1 \text{ failures}} \cdot p = p^2(1-p)^{n-2}$$

In the denominator, we have

$$\mathbb{P}(\text{second success occurred on the } n\text{th trial})$$

This occurs when there is a single success within the first $(n-1)$ trials, and the nth trial is also a success. The probability of this occurring is

$$\underbrace{\binom{n-1}{1}}_{\text{choose single success from first } n-1} \cdot p \cdot \underbrace{(1-p)^{n-2}}_{n-2 \text{ failures}} \cdot p = (n-1)p^2(1-p)^{n-2}$$

Plugging in these probabilities into the expression for conditional probability we had above, we get

$$\begin{aligned}
&\mathbb{P}(\text{first success occurred on the } k\text{th trial} \mid \text{second success occurred on the } n\text{th trial}) \\
&= \frac{p^2(1-p)^{n-2}}{(n-1)p^2(1-p)^{n-2}} \\
&= \boxed{\frac{1}{n-1}}
\end{aligned}$$

Part (e)

What is the probability that my first success occurs on an odd numbered attempt?

From **Part (b)**, we know that $\mathbb{P}(\text{first success on } k\text{th trial}) = (1-p)^{k-1} \cdot p$. Thus, we can calculate the first success on an odd numbered attempt by taking the sum

$$\begin{aligned}
&\mathbb{P}(\text{first success on odd-numbered trial}) \\
&= \mathbb{P}(\text{first success on 1st trial}) + \mathbb{P}(\text{first success on 3rd trial}) \\
&+ \mathbb{P}(\text{first success on 5th trial}) + ... \\
&= \sum_{i=0}^{\infty} \mathbb{P}(\text{first success on } (2i+1)\text{th trial}) \\
&= \sum_{i=0}^{\infty} (1-p)^{2i+1-1} \cdot p
\end{aligned}$$

Questions assigned to the following page: 6 and 5

$$
\begin{aligned}
&= \sum_{i=0}^{\infty} (1-p)^{2i} \cdot p \\
&= p \sum_{i=0}^{\infty} ((1-p)^2)^i \\
&= p \cdot \frac{1}{1-(1-p)^2} && \text{Geometric series sum} \\
&= p \cdot \frac{1}{1-(1-2p+p^2)} \\
&= p \cdot \frac{1}{1-1+2p-p^2} \\
&= \frac{p}{2p-p^2} \\
&= \boxed{\frac{1}{2-p}}
\end{aligned}
$$

Note that as p approaches 1, the probability that the first success is on an odd-numbered attempt approaches 1 (since the probability of success on the first trial increases, and the first trial is odd-numbered), and as p approaches 0, the probability approaches $\frac{1}{2}$.

Problem 6

Consider the roll of a single fair die. Define three events A, B, and C with all the following properties, which you must verify:

1. A and B are independent
2. A and C are independent
3. A and $B \cup C$ are *not* independent
4. A and $B \cap C$ are *not* independent

We have $\Omega = \{1, 2, 3, 4, 5, 6\}$ our sample space, corresponding to the number of pips that show face-up when the die is rolled. Since the die is fair, each elementary event is equiprobable with probability $\frac{1}{6}$. Consider the events

$$
\begin{aligned}
A &= \{2, 4, 6\} \\
B &= \{3, 6\} \\
C &= \{3, 4\}
\end{aligned}
$$

We claim that these events satisfy all four properties given above. First, we calculate the probabilities of each event and various combinations:

$$\mathbb{P}(A) = \frac{1}{2}$$

Question assigned to the following page: 6

$$\mathbb{P}(B) = \frac{1}{3}$$
$$\mathbb{P}(C) = \frac{1}{3}$$

$$\mathbb{P}(A \cap B) = \mathbb{P}(\{6\}) = \frac{1}{6}$$
$$\mathbb{P}(A \cap C) = \mathbb{P}(\{4\}) = \frac{1}{6}$$
$$\mathbb{P}(B \cup C) = \mathbb{P}(\{3, 4, 6\}) = \frac{1}{2}$$
$$\mathbb{P}(B \cap C) = \mathbb{P}(\{3\}) = \frac{1}{6}$$

$$\mathbb{P}(A \cap (B \cup C)) = \mathbb{P}(\{4, 6\}) = \frac{1}{3}$$
$$\mathbb{P}(A \cap (B \cap C)) = \mathbb{P}(\varnothing) = 0$$

Next, we verify the properties. To show A and B are independent, we need to show $\mathbb{P}(A)\mathbb{P}(B) = \mathbb{P}(A \cap B)$. Indeed, we have

$$\mathbb{P}(A)\mathbb{P}(B) = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$$
$$\mathbb{P}(A \cap B) = \frac{1}{6}$$

so A and B are independent. To show A and C are independent, we need to show $\mathbb{P}(A)\mathbb{P}(C) = \mathbb{P}(A \cap C)$. We have

$$\mathbb{P}(A)\mathbb{P}(C) = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$$
$$\mathbb{P}(A \cap C) = \frac{1}{6}$$

so A and V are independent also. To show A and $B \cup C$ are not independent, we need to show $\mathbb{P}(A)\mathbb{P}(B \cup C) \neq \mathbb{P}(A \cap (B \cup C))$. We have

$$\mathbb{P}(A)\mathbb{P}(B \cup C) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$
$$\mathbb{P}(A \cap (B \cup C)) = \frac{1}{3}$$

$\frac{1}{4} \neq \frac{1}{3}$, so A and $B \cup C$ are not independent. Finally, to show A and $B \cap C$ are not independent, we need to show $\mathbb{P}(A)\mathbb{P}(B \cap C) \neq \mathbb{P}(A \cap (B \cap C))$. We have

$$\mathbb{P}(A)\mathbb{P}(B \cap C) = \frac{1}{2} \cdot \frac{1}{6} = \frac{1}{12}$$

Question assigned to the following page: 6

Nathan Leung, UID 005835316
MATH 170A Homework 3
26 January 2024

$$\mathbb{P}(A \cap (B \cap C)) = 0$$

$\frac{1}{12} \neq 0$, so A and $B \cap C$ are not independent.
Thus, all four properties hold for A, B, and C.