Probability and Statistics: MA6.101

Homework 2

Topics Covered: Sigma Algebra, Probability spaces, Conditional Probability, and Total Probability.

Q1: A four-sided die is rolled repeatedly, until the first time (if ever) that an even number is obtained. What is the sample space for this experiment? Also find the sigma algebra.

A: The sample space depends on what we choose to observe from the experiment. Let $O = \{1, 3\}$ be the odd faces and $E = \{2, 4\}$ be the even faces. We roll the die repeatedly until the first even number appears.

(a) **Observation: Full outcome sequence (exact faces seen).**

$$\Omega_1 = \bigcup_{n=1}^{\infty} \{(x_1, x_2, \dots, x_n) \,|\, x_1, \dots, x_{n-1} \in O,\ x_n \in E\}$$

The σ-algebra is:

$$\mathcal{F}_1 = \mathcal{P}(\Omega_1)$$

(b) **Observation: Only parity pattern observed.**

$$\Omega_2 = \bigcup_{n=1}^{\infty} O^{n-1} \times E$$

Here we distinguish odd vs even, but not exact face values. The σ-algebra is:

$$\mathcal{F}_2 = \mathcal{P}(\Omega_2)$$

(c) **Observation: Only stopping time recorded.**

$$\Omega_3 = \mathbb{N} = \{1, 2, 3, \dots\}$$

where outcome k means the first even number appears on the k-th roll. The σ-algebra is:

$$\mathcal{F}_3 = \mathcal{P}(\mathbb{N})$$

Remark: The physical process is the same, but the sample space changes depending on the level of detail in the observations.

Q2: A vendor is heading to the iconic Narendra Modi Stadium Mortera for the IPL final between CSK and RCB. He carries two bags of jerseys:

- Bag 1: 100 CSK jerseys
- Bag 2: 100 RCB jerseys

On the way, an accident causes a *swap of k jerseys each way*:

- k RCB jerseys move into Bag 1

- k CSK jerseys move into Bag 2

After the swap, each bag still has exactly 100 jerseys. A spectator will first choose a bag at random with $\Pr(\text{Bag 1}) = p$ and $\Pr(\text{Bag 2}) = 1 - p$, then draw a jersey from the chosen bag. The overall probability of drawing an RCB jersey is known to be 0.4.

(a) Find p in terms of k.

(b) Determine which integer values for $k \in \{0, 1, \ldots, 100\}$ are valid.

(c) Suppose if $k = 30$, check if any p can make the overall probability of RCB draw equal to 0.4. If so, find it.

(d) For any valid k from part (b), if a bag is chosen and the jersey drawn is CSK, what is the probability that the jersey is from Bag 1? Substitute the values from part (c) and return the final answer.

A: Let k be the number of jerseys swapped each way. After the swap:

$$\Pr(\text{RCB} \mid \text{Bag 1}) = \frac{k}{100}, \qquad \Pr(\text{RCB} \mid \text{Bag 2}) = \frac{100 - k}{100}.$$

Let $\Pr(\text{Bag 1}) = p$ and $\Pr(\text{Bag 2}) = 1 - p$.

(a) Find p in terms of k. Using the law of total probability and the given overall $\Pr(\text{RCB}) = 0.4$,

$$0.4 \;=\; p \cdot \frac{k}{100} + (1 - p) \cdot \frac{100 - k}{100} \quad\Longleftrightarrow\quad 40 = pk + (1 - p)(100 - k).$$

Solving for p:

$$p(100 - 2k) = 60 - k \quad\Longrightarrow\quad \boxed{p = \frac{60 - k}{100 - 2k}} \qquad (k \neq 50).$$

(b) Feasibility: for which integer $k \in \{0, 1, \ldots, 100\}$ does a valid $p \in [0, 1]$ exist? From (a), $p \in [0, 1]$ iff

$$0 \leq \frac{60 - k}{100 - 2k} \leq 1.$$

Self exercise

$$\boxed{k \in \{0, 1, \ldots, 40\} \;\cup\; \{60, 61, \ldots, 100\}}$$

(while $k = 50$ is impossible since the overall RCB rate would be fixed at 0.5 regardless of p; for $41 \leq k \leq 59$, the required p falls outside $[0, 1]$).

(c) Suppose $k = 30$. Find p (if it exists). From (a),

$$p = \frac{60 - 30}{100 - 2 \cdot 30} = \frac{30}{40} = \boxed{0.75}.$$

(d) For any valid k from (b), given that the drawn jersey is CSK, compute $\Pr($**Bag 1** $\mid$ **CSK**$)$**. Then substitute the values from (c).** Bayes' rule gives

$$\Pr(\text{Bag 1} \mid \text{CSK}) = \frac{\Pr(\text{Bag 1})\Pr(\text{CSK} \mid \text{Bag 1})}{\Pr(\text{Bag 1})\Pr(\text{CSK} \mid \text{Bag 1}) + \Pr(\text{Bag 2})\Pr(\text{CSK} \mid \text{Bag 2})}$$

$$= \frac{p(100-k)}{p(100-k)+(1-p)k}.$$

For $k = 30$ and $p = 0.75$ (from (c)):

$$\Pr(\text{Bag 1} \mid \text{CSK}) = \frac{0.75 \cdot 70}{0.75 \cdot 70 + 0.25 \cdot 30} = \frac{52.5}{52.5 + 7.5} = \boxed{\frac{7}{8} = 0.875}.$$

Q3: Alice searches for her term paper in her filing cabinet. which has several drawers. She knows that she left her term paper in drawer j with probability $p_i > 0$. The drawers are so messy that even if she correctly guesses that the term paper is in drawer i. the probability that she finds it is only d_i. Alice searches in a particular drawer say drawer i. but the search is unsuccessful. Conditioned on this event, find the probability that her paper is in drawer j in terms of p_i, p_j, d_i and d_j.

A: Let A_j be the event "the paper is in drawer j", and let U_i be the event "Alice searches drawer i and the search is unsuccessful."

We are given $P(A_j) = p_j$ and that if the paper is in drawer i, the probability she finds it when searching i is d_i. Hence

$$P(U_i \mid A_j) = \begin{cases} 1 - d_i, & j = i, \\ 1, & j \neq i. \end{cases}$$

By Bayes' theorem,

$$P(A_j \mid U_i) = \frac{P(U_i \mid A_j)P(A_j)}{\sum_k P(U_i \mid A_k)P(A_k)}.$$

The denominator simplifies because $P(U_i \mid A_k) = 1$ for all $k \neq i$:

$$\sum_k P(U_i \mid A_k)P(A_k) = p_i(1-d_i) + \sum_{k \neq i} p_k = p_i(1-d_i) + (1-p_i) = 1 - p_i d_i.$$

Thus,

$$P(A_j \mid U_i) = \begin{cases} \dfrac{p_i(1-d_i)}{1-p_i d_i}, & j = i, \\ \dfrac{p_j}{1-p_i d_i}, & j \neq i. \end{cases}$$

One can verify that these probabilities sum to 1:

$$\frac{p_i(1-d_i)}{1-p_i d_i} + \sum_{j \neq i} \frac{p_j}{1-p_i d_i} = \frac{1-p_i d_i}{1-p_i d_i} = 1.$$

Q4: You have a 10 minute break between class and the tutorial during which you want to buy a cup of coffee from VC. It will rain during the break with probability 0.3, VC will be crowded with probability 0.5. Answer the following:

(a) Assuming that VC cannot be crowded when it is raining. If it rains you will be late to the tutorial with probability 0.5 and if it is crowded the probability is 0.3. What is the probability that you will be late to the tutorial?

(b) In the scenario from (a), given that you are late, which is more probable, rain or VC being crowded?

(c) Assuming that VC being crowded and rain are independent events. If both occur you will be late with probability 0.75, if it rains but VC is not crowded probability of being late to the tutorial is 0.5 and if it is crowded but not raining the probability is 0.3.

A:

(a) Events "rain" (R) and "crowded" (C) are disjoint, with $\mathbb{P}(R) = 0.3$, $\mathbb{P}(C) = 0.5$. Assume you are not late if neither occurs. Then

$$\mathbb{P}(\text{late}) = \mathbb{P}(R)\,\mathbb{P}(\text{late} \mid R) + \mathbb{P}(C)\,\mathbb{P}(\text{late} \mid C) = 0.3 \cdot 0.5 + 0.5 \cdot 0.3 = 0.30.$$

(b) By Bayes' rule,

$$\mathbb{P}(R \mid \text{late}) = \frac{\mathbb{P}(\text{late} \cap R)}{\mathbb{P}(\text{late})} = \frac{0.3 \cdot 0.5}{0.30} = 0.5, \qquad \mathbb{P}(C \mid \text{late}) = \frac{0.5 \cdot 0.3}{0.30} = 0.5.$$

They are equally likely given that you are late.

(c) Now R and C are independent with $\mathbb{P}(R) = 0.3$, $\mathbb{P}(C) = 0.5$. Partition on (R, C):

$$\begin{aligned}\mathbb{P}(\text{late}) &= \mathbb{P}(R \cap C) \cdot 0.75 + \mathbb{P}(R \cap C^c) \cdot 0.5 + \mathbb{P}(R^c \cap C) \cdot 0.3 + \mathbb{P}(R^c \cap C^c) \cdot 0 \\ &= (0.3 \cdot 0.5) \cdot 0.75 + (0.3 \cdot 0.5) \cdot 0.5 + (0.7 \cdot 0.5) \cdot 0.3 \\ &= 0.1125 + 0.075 + 0.105 \\ &= 0.2925.\end{aligned}$$

Q5: Given set A:

$$A\{(-\infty, a] : a \in \mathbb{R}\}$$

Prove that the sigma algebra generated by A is the Borel σ-algebra.

A:

(i) $\mathbb{B}(\mathbb{R}) \subseteq \sigma(\mathcal{C})$ will follow if we show that $(a, b) \in \sigma(\mathcal{C})$ for all real numbers $a < b$. This is true because

$$(a, b) = (a, \infty) \cap (-\infty, b) = (-\infty, a]^c \cap \bigcup_{n=1}^{\infty} \left(-\infty, b - \frac{1}{n}\right].$$

(ii) $\mathbb{B}(\mathbb{R}) \supseteq \sigma(\mathcal{C})$ will follow if we show that $(-\infty, a] \in \mathbb{B}(\mathbb{R})$ for every real number a. This follows from the observation that

$$(-\infty, a] = \bigcup_{m=1}^{\infty} (a - m, a] = \bigcup_{m=1}^{\infty} \bigcap_{n=1}^{\infty} \left(a - m, a + \frac{1}{n}\right).$$

Q6: Let $\{A_1\}, \{A_2\}, \ldots$ be a finite or countable partition of a non-empty set Ω (i.e., A_i are pairwise disjoint and their union is Ω). What is the σ-algebra generated by the collection of subsets $\{A_n\}$?

A: The requirements for a set to be a σ-algebra $\mathcal{F}$ are:

(a) $\emptyset \in \mathcal{F}$ (empty set must belong to the σ-algebra)

(b) If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$ (closed under complementation)

(c) If $A_1, A_2, \ldots \in \mathcal{F}$, then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$ (closed under countable unions)

Given the partition $\{A_1, A_2, A_3, \ldots\}$ of Ω, the σ-algebra generated by this collection is:

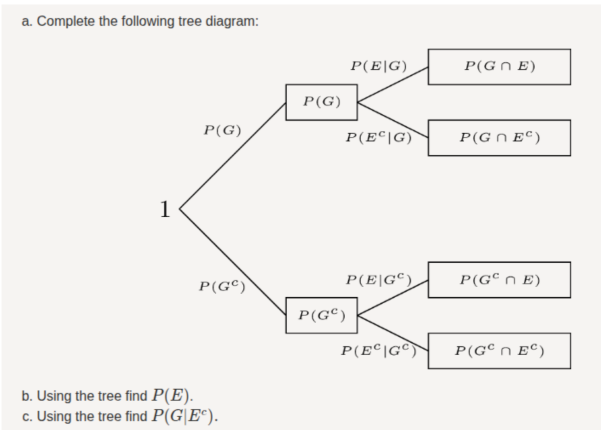
$$\sigma(\{A_n\}) = \left\{ \bigcup_{i \in I} A_i : I \subseteq \mathbb{N} \right\} \cup \{\emptyset\}$$

In other words, the σ-algebra consists of:

- The empty set $\emptyset$
- All possible unions of the partition elements A_i
- Since $\Omega = \bigcup_{i=1}^{\infty} A_i$, we also have Ω in the σ-algebra

This σ-algebra contains exactly 2^n elements if the partition is finite with n elements, or is countably infinite if the partition is countably infinite.

Q7: Consider a communication system. At any given time, the communication channel is in good condition with probability 0.8, and is in bad condition with probability 0.2. An error occurs in a transmission with probability 0.1 if the channel is in good condition, and with probability 0.3 if the channel is in bad condition. Let G be the event that the channel is in good condition and E be the event that there is an error in transmission.



A: Given information:

- $\mathbb{P}(G) = 0.8$ (channel in good condition)
- $\mathbb{P}(G^c) = 0.2$ (channel in bad condition)
- $\mathbb{P}(E|G) = 0.1$ (error given good condition)

- $\mathbb{P}(E|G^c) = 0.3$ (error given bad condition)

(a) Complete the following tree diagram:

From the given information, we can complete the tree diagram by calculating the missing probabilities:

- $\mathbb{P}(E^c|G) = 1 - \mathbb{P}(E|G) = 1 - 0.1 = 0.9$
- $\mathbb{P}(E^c|G^c) = 1 - \mathbb{P}(E|G^c) = 1 - 0.3 = 0.7$

The joint probabilities at the end of each branch are:

- $\mathbb{P}(G \cap E) = \mathbb{P}(G) \cdot \mathbb{P}(E|G) = 0.8 \times 0.1 = 0.08$
- $\mathbb{P}(G \cap E^c) = \mathbb{P}(G) \cdot \mathbb{P}(E^c|G) = 0.8 \times 0.9 = 0.72$
- $\mathbb{P}(G^c \cap E) = \mathbb{P}(G^c) \cdot \mathbb{P}(E|G^c) = 0.2 \times 0.3 = 0.06$
- $\mathbb{P}(G^c \cap E^c) = \mathbb{P}(G^c) \cdot \mathbb{P}(E^c|G^c) = 0.2 \times 0.7 = 0.14$

(b) Using the tree find $\mathbb{P}(E)$:

Using the law of total probability (summing the probabilities of all paths that lead to event E): $\mathbb{P}(E) = \mathbb{P}(G \cap E) + \mathbb{P}(G^c \cap E) = 0.08 + 0.06 = 0.14$

(c) Using the tree find $\mathbb{P}(G|E^c)$:

First, we find $\mathbb{P}(E^c)$ using the law of total probability: $\mathbb{P}(E^c) = \mathbb{P}(G \cap E^c) + \mathbb{P}(G^c \cap E^c) = 0.72 + 0.14 = 0.86$

Then using Bayes' theorem: $\mathbb{P}(G|E^c) = \frac{\mathbb{P}(G \cap E^c)}{\mathbb{P}(E^c)} = \frac{0.72}{0.86} = \frac{36}{43} \approx 0.837$

Q8: $E_1, E_2, \ldots, E_n$ be n events, each with positive probability. Prove that

$$\mathbb{P}\left(\bigcap_{i=1}^{n} E_i\right) = \mathbb{P}\{E_1\} \cdot \mathbb{P}\{E_2 \mid E_1\} \cdot \mathbb{P}\{E_3 \mid E_1 \cap E_2\} \cdots \mathbb{P}\left\{E_n \mid \bigcap_{i=1}^{n-1} E_i\right\}.$$

A: By expanding the right-hand side using the definition of conditional probability, we get:

$$\mathbb{P}\{E_1\} \cdot \frac{\mathbb{P}(E_1 \cap E_2)}{\mathbb{P}(E_1)} \cdot \frac{\mathbb{P}(E_1 \cap E_2 \cap E_3)}{\mathbb{P}(E_1 \cap E_2)} \cdots \frac{\mathbb{P}(\cap_{i=1}^{n} E_i)}{\mathbb{P}(\cap_{i=1}^{n-1} E_i)}.$$

After cancelling terms, we are left with only the numerator of the last fraction, which is equal to the left-hand side.

Q9: The alarm system at a nuclear power plant is not completely reliable. If there is something wrong with the reactor, the probability that the alarm goes off is 0.99. On the other hand, when nothing is actually wrong the alarm still goes off on 0.02 of the days. Suppose that something is wrong with the reactor only one day out of 100.

(a) What is the probability that the alarm goes off on a randomly chosen day?

(b) Given that the alarm goes off on a particular day, what is the probability that there is actually something wrong with the reactor on that day?

A:

Define: Let W denote the event "something is wrong with the reactor" and A denote the event "the alarm goes off."

Given:

$$P(W) = 0.01, \qquad P(W^c) = 0.99, \qquad P(A \mid W) = 0.99, \qquad P(A \mid W^c) = 0.02.$$

(a) **Find $P(A)$.**
Using the law of total probability:

$$P(A) = P(A \mid W)P(W) + P(A \mid W^c)P(W^c).$$

Substitute values:

$$P(A) = 0.99 \times 0.01 + 0.02 \times 0.99.$$

$$P(A) = 0.0099 + 0.0198 = 0.0297.$$

(b) **Find $P(W \mid A)$.**
Use Bayes' theorem:

$$P(W \mid A) = \frac{P(A \mid W)P(W)}{P(A)}.$$

Substitute:

$$P(W \mid A) = \frac{0.99 \times 0.01}{0.0297} = \frac{0.0099}{0.0297} = \frac{1}{3}.$$

Q10: A survey shows 56% of all American workers have a workplace retirement plan, 68% have health insurance, and 49% have both benefits. We select a worker at random.

(a) What is the probability he has neither health insurance nor a retirement plan?

(b) What is the probability he has health insurance, if he has a retirement plan?

(c) Are having health insurance and a retirement plan independent events? Are these two benefits mutually exclusive?

A:

Define: Let $R =$ the event that the worker "has a workplace retirement plan", and $H =$ the event that the worker "has health insurance".

Given: $P(R) = 0.56,\ P(H) = 0.68,\ P(R \cap H) = 0.49.$

(a) Use De Morgan and union formula:

$$P(\text{neither}) = P((R \cup H)^c) = 1 - P(R \cup H)$$

and

$$P(R \cup H) = P(R) + P(H) - P(R \cap H).$$

Compute:

$$P(R \cup H) = 0.56 + 0.68 - 0.49 = 1.24 - 0.49 = 0.75.$$

Hence

$$P(\text{neither}) = 1 - 0.75 = 0.25.$$

(b) **Find $P(H \mid R)$.**

$$P(H \mid R) = \frac{P(H \cap R)}{P(R)} = \frac{0.49}{0.56} = \frac{7}{8} = 0.875.$$

(c) - Events H and R are independent iff $P(H \cap R) = P(H)P(R)$. $P(H)P(R) = 0.68 \times 0.56$.

$$0.68 \times 0.56 = 0.3808.$$

But $P(H \cap R) = 0.49 \neq 0.3808$. Therefore they are *not independent.*

- Here $P(H \cap R) = 0.49 > 0$, so they are *not mutually exclusive* as $P(H \cap R)$ would be 0 if they were mutually exclusive.

Q11: IIITH is managing two independent government projects. The SERC lab is responsible for the first project, and the CSTAR lab is responsible for the second. Let A be the event that the SERC lab's project is successful, and let B be the event that the CSTAR lab's project is successful. The probabilities of success are $P(A) = 0.4$ and $P(B) = 0.7$.

(a) If the SERC lab's project is unsuccessful, what is the probability that the CSTAR lab's project is also unsuccessful?

(b) What is the probability that at least one of the two projects is successful?

(c) Given that at least one of the projects succeeds, what is the probability that only the SERC lab's project is successful?

A:

Define: A = SERC project successful; B = CSTAR project successful.

Given: $P(A) = 0.4$, $P(B) = 0.7$. The two projects are independent.

(a) **Find $P(B^c \mid A^c)$.**

Independence means $P(A \cap B) = P(A)P(B)$. Using complement algebra and inclusion–exclusion,

$$P(A^c \cap B^c) = 1 - P(A \cup B) = 1 - \big(P(A) + P(B) - P(A \cap B)\big).$$

Substitute $P(A \cap B) = P(A)P(B)$ to get

$$P(A^c \cap B^c) = 1 - P(A) - P(B) + P(A)P(B) = (1 - P(A))(1 - P(B)) = P(A^c)P(B^c).$$

Thus complements are independent and

$$P(B^c \mid A^c) = \frac{P(A^c \cap B^c)}{P(A^c)} = \frac{P(A^c)P(B^c)}{P(A^c)} = P(B^c).$$

Compute:

$$P(B^c) = 1 - P(B) = 1 - 0.7 = 0.3.$$

(b) **Find** $P(A \cup B)$**.**

Use inclusion–exclusion:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Independence gives $P(A \cap B) = P(A)P(B) = 0.4 \times 0.7 = 0.28$. Therefore

$$P(A \cup B) = 0.4 + 0.7 - 0.28 = 1.1 - 0.28 = 0.82.$$

(c) **Find** $P\big(A \cap B^c \mid A \cup B\big)$**.**

We compute $P(A \cap B^c)$ using independence. By set difference,

$$P(A \cap B^c) = P(A) - P(A \cap B).$$

Using independence,

$$P(A \cap B) = P(A)P(B).$$

$$P(A \cap B^c) = P(A) - P(A)P(B) = P(A)\big(1 - P(B)\big) = P(A)P(B^c).$$

$$P(A \cap B^c) = 0.4 \times (1 - 0.7) = 0.4 \times 0.3 = 0.12.$$

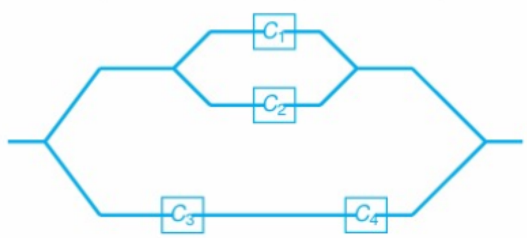
Observe that

$$A \cap B^c \subseteq A \subseteq A \cup B,$$

From part (b), $P(A \cup B) = 0.82$. Therefore,

$$P\big(A \cap B^c \mid A \cup B\big) = \frac{P(A \cap B^c)}{P(A \cup B)} = \frac{0.12}{0.82} = \frac{12}{82} = \frac{6}{41}.$$

Q12: Consider the system of components connected as in the figure given below. Components C_1 and C_2 are connected in parallel, so that subsystem S_1 works if either C_1 or C_2 works; since C_3 and C_4 are connected in series, that subsystem S_2 works if both C_3 and C_4 work. If components work independently of one another, with $P(\text{component works}) = 0.9$, calculate $P(\text{system works})$.



A:

Parallel (two components) — S_1: A parallel subsystem works if at least one component works. If a component works by probability $q = 0.9$, the probability both fail is $(1-q)^2 = 0.1^2 = 0.01$. Thus

$$P(S_1) = 1 - 0.01 = 0.99.$$

Series (two components) — S_2: A series subsystem works if both components work and, as the components are independent,

$$P(S_2) = q \times q = 0.9 \times 0.9 = 0.81.$$

Since components across subsystems are independent, S_1 and S_2 are **independent events**. The overall system is a parallel connection of S_1 and S_2: it works if S_1 or S_2 works. Hence

$$P(\text{system works}) = P(S_1 \cup S_2) = P(S_1) + P(S_2) - P(S_1)P(S_2).$$

$$P(\text{system works}) = 0.99 + 0.81 - 0.99 \times 0.81 = 0.9981.$$