Statistics

PAGE
DATE

Q.1 a) To find the probability that no one has same seat for both courses, we can consider each student individually.

For 1^{st} student,
probability of not having the same seat for both courses is 1.

For 2^{nd} student,
there is a 99/100 prob. of not having the same seat for the first course.
- Assuming the first student took a seat in the 1^{st} course,
there is a 98/99 prob. of not having the same seat for the second course.
- overall prob. for 2^{nd} student is
$(99/100) * (98/99) = 98/100$

- For 3^{rd} student,
there is a 98/100 prob. of not having the same seat for the first course.
- Assuming that the 1st two students took seats in the first course, there is a 97/99 prob of not having the same seat for the 2^{nd} course.
The overall probability for the 3^{rd} student
$(98/100) * (97/99) = (98/99) * (97/100)$.

Continuing this pattern, the prob for the n^{th} student is $(n-1)/100 * (n-2)/99 * \ldots * 2/101 * 1/100$.

To find the prob that no one has same seat for both courses, we multiply prob for each student:

P(no one has the same seat) $= (99/100) * (98/99) * (97/100) * \ldots * (2/101) * (1/100)$

(b) To approximate the prob we can use the fact that for small values of x, $(1-x) = e^{\wedge}(-x)$. in our case, x is (1/100).

P(no one has the same seat) =
$\approx (99/100) * (98/99) * (97/100) * \ldots * (2/101) * (1/100)$
$\approx e^{\wedge}(-1/100) * e^{\wedge}(-2/100) * \ldots * e^{\wedge}(-99/100)$
$\approx e^{\wedge}(-1/100) - (2/100) - \ldots - (99/100))$
$\approx e^{\wedge}(-99/100 * (100/2))$
$\approx e^{\wedge}(-99/2)$.

(c) To find the prob. that at least two students have the same seat for both courses, we can subtract the probability of ~~a~~ no one having the same seat from 1:

P(at least two students have same seat)
$= 1 -$ P(no one has the same seat).

Using the approx from part (b), we have:
P (at least two students have the same seat)
$\approx 1 - e^{\wedge}(-99/2)$.

(2) After the first person, neither of the passengers show any preference for the last person's seat nor the seat of the first passenger.

- Once all the passengers except the last passenger occupy the seats, the first passenger would be sitting in the seat allotted to him or in that of the last person.

- Therefore, the probability that the last person to board gets his assigned seat unoccupied is $\frac{1}{2} = 0.5$

(3) A reasonable choice of distribution is Poisson (λt), ~~where $\lambda = 20, 5 = 100$~~
, where $\lambda = 20 \times 5 = 100$ (the avg. no. of raindrops per minute hitting the region.)

Assuming the distribution,

P (no raindrops in 1/20 of a minute)
$= e^{-100/20} (100/20)^0 / 0! = e^{-5}$.

(4) Given that X is a random day of the week, coded so that Monday is 1, Tuesday is 2, etc. (So X takes values 1, 2, ..., 7, with equal probabilities).

Y is the next day after X.
So Y can be written as $X+1$.
But since highest value is 7, we can write
$Y = X+1 \mod 7$

i.e. Whenever $X=7$, $Y=8 = 1$ again

Thus, Y can take values of

Y: 1 2 3 4 5 6 7
$P(Y) = P(X+1) = 1/7$

Y also has the same distribution as X.

- They have the same distribution since Y is also equally likely to represent any day of the week but $P(X<Y) = P(X \neq 7) = 6/7$.

(5) Let A_i be the event that there are no birthdays in the ith season. The probability that all seasons occurs at least once is $1 - P(A_1 \cup A_2 \cup A_3 \cup A_4)$. Note that $A_1 \cap A_2 \cap A_3 \cap A_4 = \phi$. Using the inclusion-exclusion principle & the symmetry of the seasons,

$P(A_1 \cup A_2 \cup A_3 \cup A_4)$

$$= \sum_{i=1}^{4} P(A_i) - \sum_{i=1}^{3} \sum_{j>i} P(A_i \cap A_j)$$

$$+ \sum_{i=1}^{2} \sum_{j>i} \sum_{k>j} P(A_i \cap A_j \cap A_k)$$

$$= 4P(A_1) - 6P(A_1 \cap A_2) + 4P(A_1 \cap A_2 \cap A_3)$$

We have $P(A_1) = (3/4)^7$.
Similarly,

$P(A_1 \cap A_2) = 1/2^7$ and

$P(A_1 \cap A_2 \cap A_3) = 1/4^7$.

Therefore, $P(A_1 \cup A_2 \cup A_3 \cup A_4) = 4(3/4)^7 - 6/2^7 + 4/4^7$.

So, the probability that all 4 seasons occur at least once is

$$1 - \left(4(3/4)^7 - 6/2^7 + 4/4^7\right)$$

$$\approx 0.513.$$

⑥ Direct Method :-

There are two general ways that Alice can have class every day: either she has 2 days with 2 classes & 3 days with 1 class, or she has 1 day with 3 classes, & has 1 class on each of other 4 days. The number of possibilities for the former is $\binom{5}{2}\binom{6}{2}^2 6^3$ (to choose the 2 days when she has 2 classes, & then select 2 classes on those days & 1 class for the other days.).

The number of possibilities for the latter is $\binom{5}{1}\binom{6}{3}6^4$. So the probability is

$$= \frac{\binom{5}{2}\binom{6}{2}^2 6^3 + \binom{5}{1}\binom{6}{3}6^4}{\binom{30}{7}}$$

$$= \frac{114}{377}$$

$$\approx 0.302$$

⑦ Suppose that A & B are events.
If $P(A)=0$ or $P(A)=1$, then A & B are independent. A is independent of itself if & only if $P(A)=0$ or $P(A)=1$.

- The ans. It is not possible for an event to be said to be independent of itself. The concept of independence in probability theory relies on the notion that the occurrence or non-occurrence of one event does not affect the probability of another event.

- If an event is truly independent, it means that the outcome of one event provides no information or influence on the outcome to the event.

- Therefore, by definition, an event cannot be independent of itself. Independence requires distinct events or variables to be considered.

Statistics

8) Is it always true that if A and B are independent events, then Ac and Bc are independent events? Show that it is, or give a counterexample?

→ No, it is not always true that if events A & B are independent, then their complements (A_C & B_C) are also independent events. This statement does not hold in general.

To demonstrate this, let is provide a counterexample:

Consider rolling a fair six-sided die. Let A be the event "rolling an ~~prime~~ even number" & B be the event "rolling an prime number" The probabilities of A & B are as follows:

$P(A) = 3/6 = 1/2$ (since there are 3 even no: 2,4,6)
$P(B) = 3/6 = 1/2$ (since there are 3 prime no: 2,3,5)

A & B are independent events because the probability of both events occurring is equal to the product of their individual probabilities

$P(A \cap B) = P(A) * P(B) = (1/2) * (1/2) = 1/4$

Now, let's consider the complements A_c & B_c

A_c is the event "not rolling an even number". It consists of the odd no: 1, 3 & 5.

B_c is the event "not rolling a prime no". It consists of the composite no: 4 & 6.

The probabilities of A_c & B_c are as follows

$P(A_c) = 3/6 = 1/2$ (since there are 3 odd no: 1, 3, 5)
$P(B_c) = 2/6 = 1/3$ (since there are 2 composite no: 4, 6)

Now, let's examine the probability of the intersection of A_c & B_c:

$P(A_c \cap B_c) = P(\text{rolling a composite no})$
$= 2/6 = 1/3.$

The prob. of A_c & B_c occurring together is not equal to the product of their individual prob. Hence, A_c & B_c are not independent event.

- Therefore, we have shown a counterexample where A & B are independent event "If A & B are independent events, then A_c & B_c are independent event" is not universally true.

(a) Consider two fair independent coin tosses, & let A be the event that the first toss is Head,
B be the event that the second toss is Head, & C be the event that the two tosses have the same result.

- Then, A, B, C are dependent since

$P(A \cap B \cap C) = P(A \cap B)$
$= P(A)\,P(B)$
$= 1/4 \neq 1/8$
$\neq P(A)\,P(B)\,P(C)$,

but they are pairwise independent:
A and B are independent by definition.
A & C are independent since,

$P(A \cap C) = P(A \cap B)$
$= 1/4$
$= P(A)\,P(C)$

& Similarly B & C are independent.

(10) Let E_1 & E_2 be the events that marble is green & blue respectively in the bag. Let A be the event of picking up a green marble

Then $P(E_1) = P(E_2) = 1/2$

$P(A|E_1) = 1$

$P(A|E_2) = 1/2$

Now, if the marble taken out is green then probability that remaining marble is also green is $P(E_1/A)$

$$P(E_1/A) = \frac{P(E_1)P(A|E_1)}{P(E_1)P(A|E_1) + P(E_2)P(A|E_2)}$$

$$= \frac{1/2 \times 1}{1/2 \times 1 + 1/2 \times 1/2} = \frac{1/2}{1/2 + 1/4}$$

$$= 1/2 \times 4/3$$

$$= 2/3$$

(11) (a) Find the joint PMF of X, Y, Z

The joint distribution of X, Y, Z is

$$P(X=a, Y=b, Z=c) = \frac{n!}{a!\,b!\,c!}\left(\frac{1}{3}\right)^{a+b+c}$$

Where a, b, c are any nonnegative integers with $a+b+c=n$, since $(1/3)^{a+b+c}$ is the probability of any specific configuration of choice for each players with the right numbers in each category, & the coefficient in front counts the number of distinct ways to permute such a configuration

Alternatively, we can write the joint PMF as

$$P(X=a, Y=b, Z=c) = P(X=a)P(Y=b|X=a)\cdot P(Z=c|X=a, Y=b),$$

where for $a+b+c=n$, $P(X=a)$ can be found from the Bin$(n, 1/3)$ PMF, $P(Y=b|X=a)$ can be found from the Bin$(n-a, 1/2)$ PMF, & $P(Z=c|X=a, Y=b) = 1$.

This is a Multinomial $(n, (1/3, 1/3, 1/3))$ distribution

(b) Find the probability that the game is decisive.

- The game is decisive if & only if exactly one of X, Y, Z is 0. These cases are disjoint & so by symmetry, the probability is 3 times the probability that X is zero & Y & Z are nonzero. Note that if $X=0$ & $Y=k$, then $Z=n-k$.

This gives

$$P(\text{decisive}) = 3 \sum_{k=1}^{n-1} \frac{n!}{0!\,k!\,(n-k)!} \left(\frac{1}{3}\right)^n$$

$$= 3\left(\frac{1}{3}\right)^n \sum_{k=1}^{n-1} \binom{n}{k}$$

$$= \frac{2^n - 2}{3^{n-1}}$$

since, $\sum_{k=1}^{n-1} \binom{n}{k} = -1 - 1 + \sum_{k=0}^{n} \binom{n}{k}$

$$= 2^n - 2$$

As a check, when $n=2$ this reduces to $2/3$, which makes sense since for 2 players, the game is decisive if & only if the two players do not pick the same choice.

(c) What is the probability that the game is decisive for $n=5$? What is the limiting probability that a game is decisive as $n \to \infty$? Explain briefly why.

For $n=5$, the probability is $(2^5 - 2)/3^4$

$= 30/81$

≈ 0.37

As $n \to \infty$, $(2^n - 2)/3^{n-1} \to 0$, which make sense since if the number of players is very large, it is very likely that there will be at least one of each of rock, paper & scissors.

(2) Let S be the event that an email is spam & F be the event that an email has the "free money" phrase

By Bayes' Rule,

$$P(S|F) = \frac{P(F|S)\,P(S)}{P(F)}$$

$$= \frac{0.1 \cdot 0.8}{0.1 \cdot 0.8 + 0.01 \cdot 0.2} = \frac{80/1000}{82/1000}$$

$$= \frac{80}{82} \approx 0.9756$$

(13) Let M be the event that A's blood type matches the guilty party's & for brevity, write A for "A is guilty" & B for "B is guilty". By Bayes Rule,

$$P(A|M) = \frac{P(M|A)P(A)}{P(M|A)P(A) + P(M|B)P(B)}$$

$$= \frac{1/2}{1/2 + (1/10)(1/2)}$$

$$= \frac{10}{11}$$

(We have $P(M|B) = 1/10$ since, given that B is guilty, the prob that A's blood type matches the guilty party's the same prob as for general population).

(b) Given this new information, what is the prob that B's blood type matches that found at the crime scene?

- let C be the event that B's blood type matches, & condition on whether B is guilty. This gives

$$P(C|M) = P(C|M, A)\,P(A|M) + P(C|M, B)\,P(B|M)$$

$$= \frac{1}{10} \cdot \frac{10}{11} + \frac{1}{11}$$

$$= \frac{2}{11}$$

(14) (a) What is your probability of winning the first game?

Let W_i be the event of winning the i^{th} game. By the law of total probability,

$$P(W_1) = (0.9 + 0.5 + 0.3)/3$$
$$= 17/30.$$

(b) Congratulations: you won the first game...

We have $P(W_2|W_1) = P(W_2, W_1)/P(W_1)$.

The denominator is known from (a), while the numerator can be found by conditioning on the skill level of the opponent:

$P(W_1, W_2) = \frac{1}{3} P(W_1, W_2 \mid \text{beginner})$

$+ \frac{1}{3} P(W_1, W_2 \mid \text{intermediate})$

$+ \frac{1}{3} P(W_1, W_2 \mid \text{expert})$.

Since W_1 & W_2 are conditionally independent given the skill level of the opponent, this becomes

$$P(W_1, W_2) = \left((0.9)^2 + (0.5)^2 + (0.3)^2\right) / 3$$

$$= 23/60$$

$$P(W_2 \mid W_1) = \frac{23/60}{17/30}$$

$$= \frac{23}{34}$$

(c)

Independence here means that knowing one game's outcome gives no information about the other game's outcome, while conditional independence is the same statement where all probabilities are conditional on the opponent's skill level. Conditional independence given the opponent's skill level is a more reasonable assumption here.

- This is because winning the first game gives information about the opponent's skill level, which in turn gives information about the result of the second game.

- That is, if the opponent's skill level is treated as fixed & known, then it may be reasonable to assume independence of games given this information; with the opponent's skill level random, earlier games can be used to help infer the opponent's skill level, which affects the probs. for future games.

(15) Marginally we have $X \sim Bin(n, p_1)$ as shown on a previous homework problem using story proof.

Here X & Y are not independent, unlike in the chicken-egg problem from class. This follows immediately from thinking about an extreme case: if $X = n$, then clearly $Y = 0$. So they are not independent: $P(Y=0) < 1$, while $P(Y=0 \mid X=n) = 1$.

To find the joint distribution condition on N and note that only the $N = i + j$ term is nonzero: for any nonnegative integers i, j with $i+j \leq n$,

$$P(X=i, Y=j) = P(X=i, Y=j \mid N=i+j)\, P(N=i+j)$$

$$= P(X=i \mid N=i+j)\, P(N=i+j)$$

$$= \binom{i+j}{i} s^i (1-s)^j \binom{n}{i+j} p^{i+j} (1-p)^{n-i-j}$$

$$= \frac{n!}{i!\,j!\,(n-i-j)!} (ps)^i (p(1-s))^j (1-p)^{n-i-j}.$$

If we let Z be the no. of eggs which don't hatch, then from the above we have that (X, Y, Z) has a multinomial $(n, (ps, p(1-s), (1-p)))$ distribution, which makes sense ~~activity~~ intuitively since each egg independently falls into 1 of 3 categories: hatch-and-survive, hatch & don't survive, & don't hatch, with probs. ps, $p(1-s)$, $1-p$ respectively.