CS215 Fall, 2024: Tutorial 1

Atharva Tambat

September 15, 2024

Table of Contents

Question 1

Using Chebyshev's Inequality:

$$P\{|X-\mu| \geq k\} \leq \frac{\sigma^2}{k^2}$$

Here, $\sigma^2 = 0$. Take $k = \frac{1}{n}$

$$P\left\{|X-\mu| \geq \frac{1}{n}\right\} \leq 0 \implies P\left\{|X-\mu| \geq \frac{1}{n}\right\} = 0$$

Take limit $n \to \infty$,

$$\lim_{n\to\infty} P\left\{|X-\mu| \geq \frac{1}{n}\right\} = P\left\{\lim_{n\to\infty}\left\{|X-\mu| \geq \frac{1}{n}\right\}\right\} = P\{X \neq \mu\} = 0$$

Table of Contents

Question 2 - Independence $\not\Longleftrightarrow$ Uncorrelated

1. Proof: Independence $\implies$ Uncorrelated
   For Independent variables X and Y, E[XY] = E[X]E[Y]
   Proof:

$$
\begin{aligned}
E[XY] = \mu_{xy} &= \sum_i \sum_j x_i y_j p(x_i, y_j) \\
&= \sum_i \sum_j x_i y_j p(x_i) p(y_j) \\
&= \left(\sum_i x_i p(x_i) \right) \left(\sum_j y_j p(y_j) \right) \\
&= E[X]E[Y] = \mu_x \mu_y
\end{aligned}
$$

Question 2 - Independence $\not\Leftrightarrow$ Uncorrelated

1. Proof: Uncorrelated $\not\Rightarrow$ Independence
   Proof: Counterexample: $X \in \{-1, 0, 1\}$

   $P(X = -1) = P(X = 0) = P(X = 1) = 1/3$

$$Y = \begin{cases} 1, & \text{if } X = 0 \\ 0, & \text{otherwise} \end{cases}$$

Table of Contents

Question 3 - Result 1

- First, prove that $Cov(X_1 + X_2, Y_1) = Cov(X_1, Y_1) + Cov(X_2, Y_1)$
  Proof:

$$Cov(X_1 + X_2, Y_1) = E[(X_1 + X_2)Y_1] - E[X_1 + X_2]E[Y]$$

$$= E[X_1 Y_1] + E[X_2 Y_1] - (E[X_1]E[Y_1] + E[X_2]E[Y_1])$$

$$= (E[X_1 Y_1] - E[X_1]E[Y_1]) + (E[X_2 Y_1] - E[X_2]E[Y_1])$$

$$= Cov(X_1, Y_1) + Cov(X_2, Y_1)$$

  - Since, $E[X + Y] = E[X] + E[Y]$ for R.V. X and Y

Question 3 - Result 2

- Next, prove that $Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$
  Proof:

$$
\begin{aligned}
Var(X+Y) &= E[(X+Y)^2] - E[X+Y]^2 \\
&= E[X^2 + Y^2 + 2XY] - (E[X] + E[Y])^2 \\
&= E[X^2 + Y^2 + 2XY] - (E[X]^2 + E[Y]^2 + E[X]E[Y]) \\
&= (E[X^2] - E[X]^2) + (E[Y^2] - E[Y]^2) + 2(E[XY] - E[X]E[Y]) \\
&= Var(X) + Var(Y) + 2Cov(X,Y)
\end{aligned}
$$

Question 3 - Final Step

$$Corr(\sum_i X_i, \sum_i Y_i) = \frac{Cov(\sum_i X_i, \sum_i Y_i)}{\sqrt{Var(\sum_i X_i) Var(\sum_i Y_i)}}$$

Since all the X_is are independent of each other (same for all Y_is), which means $Cov(X_i, X_j) = 0 \ \forall i \neq j$, Using Result 2

$$Corr(\sum_i X_i, \sum_i Y_i) = \frac{Cov(\sum_i X_i, \sum_i Y_i)}{\sqrt{\sum_i Var(X_i) \sum_i Var(Y_i)}}$$

Using Result 1,

$$= \frac{\sum_i \sum_j Cov(X_i, Y_j)}{\sqrt{\sum_i Var(X_i) \sum_i Var(Y_i)}}$$

$$= \frac{\sum_{i=j} Cov(X_i, Y_j) + \sum_{i \neq j} Cov(X_i, Y_j)}{\sqrt{(n\sigma_x^2)(n\sigma_y^2)}} = \frac{n\rho\sigma_x\sigma_y + 0}{\sqrt{(n\sigma_x^2)(n\sigma_y^2)}} = \rho$$

Table of Contents

Question 4

Define R.Vs (called indicator RVs - which indicate whether an event has happened or not) as follows

$$X_i = \begin{cases} 1, & \text{if you get the } i^{th} \text{ Ace in the dealt cards} \\ 0, & \text{otherwise} \end{cases}$$

$$Y_i = \begin{cases} 1, & \text{if you get the } i^{th} \text{ Spade in the dealt cards} \\ 0, & \text{otherwise} \end{cases}$$

So you have $X_1, X_2 ... X_4$ and $Y_1, Y_2 ... Y_{13}$ random variables. **Also notice two R.Vs (one from Xs and other from Ys), which denote Ace of Spades (WLOG, we'll call them X_1 and Y_1), will have the same value**. i.e $X_1 = Y_1$.

Question 4

So, number of aces in the cards dealt (X) $= \sum_i X_i$.

Number of spades in the cards dealt (Y) $= \sum_i Y_i$

We'll show Cov(X, Y) $= 0$

Question 4

$$Cov(X, Y) = Cov(\sum_i X_i, \sum_j Y_j)$$

Using the result 1 from Question 3,

$$Cov(\sum_i X_i, \sum_j Y_j) = \sum_i \sum_j Cov(X_i, Y_j)$$

$$= Cov(X_1, Y_1) + \sum_{i \neq 1 \ or \ j \neq 1} Cov(X_i, Y_j)$$

Question 4

- First find $Cov(X_1, Y_1)$.

$$Cov(X_1, Y_1) = E[X_1 Y_1] - E[X_1]E[Y_1] = E[X_1^2] - E[X_1]^2$$

Since $X_1 = Y_1$

$$= Var(X_1)$$

- Notice two things
  - $E[X_1^2] = E[X_1]$ since X_1 is an indicator RV
  - $E[X_1] =$ **probability, p, of the event which is indicated by the indicator variable**. Since $E[X_1] = p.1 + (1-p).0 = p$
  - So, $Var(X_1) = \frac{1}{4}\left(1 - \frac{1}{4}\right) = \frac{3}{16}$

Question 4

Next, calculate $Cov(X_i, Y_j)$, where $i \neq 1$ *or* $j \neq 1$ - i.e. X_i and Y_j represent **different** cards. Note, there are **51 such terms**.

$$Cov(X_i, Y_j) = E[X_i Y_j] - E[X_i]E[Y_j]$$

- $$\begin{aligned} E[X_i Y_j] &= P(X_i = 1, Y_j = 1).1.1 + P(X_i = 1, Y_j = 0).1.0 \\ &+ P(X_i = 0, Y_j = 1).0.1 + P(X_i = 0, Y_j = 0).0.0 \\ &= \frac{50C11}{52C13} = 1/17 \end{aligned}$$
- $E[X_i] = E[Y_j] = 51C12/52C13 = 1/4$
- $Cov(X_i, Y_j) = -1/272$

Question 4 - The final nail in the coffin

$$Cov(X, Y) = \frac{3}{16} + 51(-\frac{1}{272}) = 0$$

Question 4 (ii)

These two events are not independent, even though they are uncorrelated - one more example of Question 2.

$$P(X = 4, Y = 13) = 0 \neq P(X = 4)P(Y = 13)$$

But for X and Y to be independent, $P(X, Y) = P(X).P(Y)$ should hold for all X and Y, therefore, contradiction.

Table of Contents

Question 5

We'll calculate the probability of Mr. Jones winning. Let W be the event of Mr. Jones winning at a game, and let T be the event that in the 10 previous spins, the ball has landed on a black number.

Since, Mr. Jones does not bet if in the 10 previous turns, the ball hasn't landed on a black number, we only care about $P(W|T)$

Since, W is independent of T, therefore the probability of him winning the match is

$$P(W|T) = P(W)$$

Which is equal to the probability of winning **any individual** game (not necessarily the one after getting 10 blacks)

Table of Contents

Question 6 - Law of Total Probability

Applying the law of total probability:

$$P(E|E \cup F) = P(E|E \cup F, F)P(F) + P(E|E \cup F, \bar{F})P(\bar{F})$$

1. $P(A|C, D)$ - Probability of event A, given events C **and** D have occurred
2. $P(E|E \cup F, F) = \frac{P(E \cap (E \cup F) \cap F)}{P((E \cup F) \cap F)} = P(E \cap F)/P(F) = P(E|F)$
3. $P(E|E \cup F, \bar{F}) = 1.$ Why?

Question 6 - Law of Total Probability

Thus,

$$
\begin{aligned}
P(E|E \cup F) &= P(E|F)P(F) + (1 - P(F)) \\
&\geq P(E|F)P(F) + P(E|F)(1 - P(F)) \quad \textbf{Why?}
\end{aligned}
$$

So,

$$P(E|E \cup F) \geq P(E|F)(P(F) + 1 - P(F)) = P(E|F)$$

Table of Contents

Optimal Strategy

Ans: **Any strategy** has a winning probability of 1/52!!

Proof: Let us prove a stronger point - in a game with n cards, the winning probability is always 1/n regardless of the strategy.

1. $p \rightarrow$ probability that the strategy chooses the first card
2. $G \rightarrow$ Event that the first card is guessed

Two cases of winning:

- 1^{st} card is ace of spades - happens with probability 1/n
- 1^{st} card is not ace of spades: What is the probability of win, **given** we skip the first chance?

Optimal Strategy

H = first card is not ace of spades

$$P(H).P(\{win\}|H)$$

But $P(\{win\}|H)$ = probability of winning with n-1 cards = $\frac{1}{n-1}$ by induction hypothesis

$$P(H).P(\{win\}|H) = \frac{n-1}{n}\frac{1}{n-1} = \frac{1}{n}$$

Using the Law of total probability:

$$\begin{aligned} P(\{win\}) &= P(\{win\}|G)P(G) + P(\{win\}|\bar{G})(1 - P(G)) \\ &= \frac{1}{n}p + \frac{1}{n}(1-p) \\ &= \frac{1}{n} \end{aligned}$$

Table of Contents

Coin Flip

- $C_i \rightarrow$ Event that coin i is chosen
- $F_n \rightarrow$ Event that first n tosses are heads
- $H \rightarrow$ Event that the $(n+1)^{th}$ is a head

$$P(H|F_n) = ?$$

$$P(H|F_n) = \sum_{i=0}^{k} P(H|F_n C_i) P(C_i|F_n)$$

Coin Flip

$P(H|F_n C_i)$ - means given i^{th} coin is selected and first n tosses are heads, what is the probability $(n+1)^{th}$ toss is a head.

$$P(H|F_n C_i) = P(H|C_i) = \frac{i}{k} \quad \textbf{Why?}$$

Also,

$$P(C_i|F_n) = \frac{P(C_i F_n)}{P(F_n)} = \frac{P(F_n|C_i)P(C_i)}{\sum_{j=0}^{k} P(F_n|C_j)P(C_j)}$$

Coin Flip

$$P(C_i|F_n) = \frac{(i/k)^n[1/(k+1)]}{\sum_{j=0}^{n}(j/k)^n[1/(k+1)]}$$

Thus,

$$P(H|F_n) = \sum_{i=0}^{k} \frac{(i/k)^{n+1}}{\sum_{j=0}^{k}(j/k)^n} = \frac{\sum_{i=0}^{k}(i/k)^{n+1}}{\sum_{j=0}^{k}(j/k)^n}$$

Now, use the approximation to simplify the numerator and denominator