Midterm 3 Review

Thursday April 13

On Thursday April 20 we will have our third midterm exam during the usual class time. The exam will take place in COHN G17 and in COLL 200.

The students whose last name begins with a letter in the range I-O will have to take the exam in COLL 200. (We’ll switch alphabetical range again for the third midterm.) The rest (A-H and P-Z) should come take the exam in our lecture room, COHN G17. Apologies to those few who ended up in COLL 200 for two out of three midterms.

The exam will last for 60 minutes. Please be in COHN G17 or COLL 200 at 1:30PM so we have time to seat everybody properly.

This here is a midterm review document with readings, a mock (practice) exam and more practice problems. You should solve the practice exam while timing yourselves.

Solutions to the practice exam will be posted Sunday evening.

Val will hold a review session on Tuesday April 18, 7:30-9:00PM in Heilmeier Hall (TOWNE 100).

The TAs will also hold a review session (TBA).

1 Readings

STUDY IN-DEPTH... ...the posted notes for lectures 15, 16, 17, 18, 19, 20, 21, and 22.

STUDY IN-DEPTH... ...the posted guides for recitations 7, 8, 9, and 10 (when posted).

STUDY IN-DEPTH... ...the posted solutions to homeworks 7, 8, 9, and 10 (when posted). Compare with your own solutions.

STUDY IN-DEPTH... ...the solutions to the answer mock exam and the additional problems contained in this document, to be posted Sunday evening. Until then, try very hard to solve these on your own.

2 Memorize!

Find and memorize formulas:

- For the sum of a geometric progression.
- For the sum of the integers in [1..n].
- For the sum of the squares of the integers in [1..n].
- For the expectation and the variance of the Bernoulli, binomial, and geometric distributions.
- For the Markov and Chebyshev inequalities.
3 Mock Exam (60 minutes for 120 points)

1. (25 pts) For each statement below, decide whether it is TRUE or FALSE and circle the right one. In each case attach a very brief explanation of your answer.

(a) The complete graph \( K_5 \) has a cycle of length 5, TRUE or FALSE?

(b) For any tree if we add one edge between any two existing vertices the resulting graph is not bipartite, TRUE or FALSE?

(c) Let \( G \) be a DAG with \( n \geq 2 \) vertices and let the sequence \( \sigma \) be a topological sort of \( G \). If \( u \) appears before \( v \) in \( \sigma \) then there exists a directed path from \( u \) to \( v \) in \( G \), TRUE or FALSE?

(d) Let \( X, Y, Z \) be random variables on a probability space \( (\Omega, \Pr) \) such that \( Z = X + Y - XY \). If \( X \) and \( Y \) are Bernoulli random variables, then \( Z \) is also a Bernoulli random variable, TRUE or FALSE?

(e) There exists a random variable \( X \) for which \( \mathbb{E}[X^2] < (\mathbb{E}[X])^2 \), TRUE or FALSE?

Answer

(a) TRUE. In fact, any permutation of 5 vertices allows us to define a cycle, for example 1-2-3-4-5-1.

(b) FALSE. We know \( P_4 \) is a tree, but adding the edge 1-4 gives us \( C_4 \), which is bipartite.

(c) FALSE. We can have a directed graph \( (\{1, 2\}, \emptyset) \), which has a valid topological sort 1, 2. But there is no path from 1 to 2.

(d) TRUE. We observe that in any case, \( Z(\omega) \in \{0, 1\} \).

Case 1: \( X = 0, Y = 0 \)

Then \( Z = 0 + 0 - 0 = 0 \).

Case 2: \( X = 0, Y = 1 \)

Then \( Z = 0 + 1 - 0 = 1 \).

Case 3: \( X = 1, Y = 0 \)

Then \( Z = 1 + 0 - 0 = 1 \).

Case 4: \( X = 1, Y = 1 \)

Then \( Z = 1 + 1 - 1 = 1 \).

So \( Z \) is a Bernoulli random variable with probability

\[ \Pr[Z = 1] = \Pr[(X = 1 \land Y = 0) \lor (X = 0 \land Y = 1)] \]

(e) FALSE. This is equivalent to saying that \( \text{Var}[X] < 0 \), but we know \( \text{Var}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2] \), i.e. it is the expectation of a nonnegative random variable, which cannot be negative.

2. (30pts) In this problem the graphs are undirected.

(a) Draw a connected graph with 6 nodes and exactly 2 cut edges.

(b) Draw an acyclic graph with 5 edges and 7 nodes.

(c) Draw a connected graph in which every node has degree 3.
We include one example for each; there are many other valid examples.

(a) One such example is as follows:

Where the cut edges here are 2–3 and 4–5.

(b) Consider the following graph:

We can just count the vertices and edges to see that the conditions specified above are satisfied.

(c) Consider $K_4$, the complete graph on 4 vertices:

Since there is an edge between every pair of vertices, they are all connected. Additionally, we see that each vertex has degree 3.

3. (30pts) Alice has an urn with three marbles labeled 1, 2, and 3. Each of the marbles is equally likely to be extracted. Bob has a separate, similar urn. They play the following game of chance:

(1) Alice extracts a marble from her urn and obtains $a \in \{1, 2, 3\}$.

(2) Independently, Bob extracts a marble from his urn and obtains $b \in \{1, 2, 3\}$.

(3) If $a > b$ then Alice wins. If $b > a$ then Bob wins. If $a = b$ they flip a fair coin and if the coin shows heads, Alice wins. If the coin shows tails, Bob wins.

In the calculations below, do not spend time on the arithmetic. It’s OK to leave your results as products and fractions.

(a) Draw the “tree of possibilities” diagram for this game, with all the outcomes and their probabilities.

(b) Compute the probability that the game was decided by a coin flip.
(c) Compute the conditional probability that Alice wins, knowing that Bob extracted the marble labeled 2.

(d) Alice and Bob put bets on the game. If Alice wins without a coin flip Bob pays her 2$. If Alice wins with a coin flip then Bob pays her 1$. If Bob wins then Alice pays him 1.5$.

What is Alice’s expected monetary win/loss (wins are positive, losses are negative) after $n$ such games?

**Answer**

Let $A_i$ be the event that Alice removes the marble with value $i$ and $B_j$ be the event that Bob removes the marble with value $j$. Let $C$ be the event that the coin is flipped. Finally, let $W_A$ be the event that Alice wins.

(a) The tree of possibilities:
(b) We want to find $\Pr[C]$. Note that we flip a coin exactly when $a = b$, i.e. we are interested in the event $\bigcup_{i=1}^{3} A_i \cap B_i$. Since each of these are disjoint, we apply the Sum Rule:

$$\Pr \left[ \bigcup_{i=1}^{3} A_i \cap B_i \right] = \Pr[A_1 \cap B_1] + \Pr[A_2 \cap B_2] + \Pr[A_3 \cap B_3]$$

Since the marbles are drawn independently, $A_i \perp B_j$ for all $i, j$:

$$= \Pr[A_1] \Pr[B_1] + \Pr[A_2] \Pr[B_2] + \Pr[A_3] \Pr[B_3]$$

Since we know each marble is equally likely to be extracted:

$$= \frac{1}{3} \times \frac{1}{3} + \frac{1}{3} \times \frac{1}{3} + \frac{1}{3} \times \frac{1}{3} = \frac{1}{3}$$

(c) We want to find $\Pr[W_A \mid B_2]$.

$$\Pr[W_A \mid B_2] = \frac{\Pr[W_A \cap B_2]}{\Pr[B_2]} = \frac{\Pr[W_A \cap B_2 \cap C] + \Pr[W_A \cap B_2 \cap \overline{C}]}{\Pr[B_2]} = \frac{\Pr[W_A \mid C \cap B_2] \Pr[C \mid B_2] \Pr[B_2] + \Pr[W_A \mid \overline{C} \cap B_2] \Pr[\overline{C} \mid B_2] \Pr[B_2]}{\Pr[B_2]}$$

Note that $\Pr[W_A \mid \overline{C} \cap B_i] = \frac{1}{2}$, since we are flipping a coin at this point. Additionally, $\Pr[C \mid B_i] = \frac{1}{3}$, since we flip the coin exactly when $a = b$. Finally, $\Pr[W_A \mid C \cap B_2] = \frac{1}{2}$, since Alice is equally likely to have drawn a 1 or a 3.

$$= \frac{1}{2} \times \frac{1}{3} \times \frac{1}{3} + \frac{1}{2} \times \frac{2}{3} \times \frac{1}{3} \times \frac{1}{3} = \frac{1}{2}$$

(d) Let $G$ be the random variable denoting Alice’s monetary gain after a single game. We seek $E[G]$. We express $G$ as a piecewise function:

$$G = \begin{cases} 
2 & W_A \cap \overline{C} \\
1 & W_A \cap C \\
-1.5 & W_A 
\end{cases}$$

Then we can find $E[G]$ thus:

$$E[G] = G(W_A \cap \overline{C}) \cdot \Pr[W_A \cap \overline{C}] + G(W_A \cap C) \cdot \Pr[W_A \cap C] + G(W_A) \cdot \Pr[W_A]$$

$$= (2) \cdot \Pr[W_A \cap \overline{C}] + (1) \cdot \Pr[W_A \cap C] + (-1.5) \cdot \Pr[W_A]$$
We proceed by finding the relevant probabilities. We have $W_A \cap C$ when the coin is flipped and comes up with a head. Then:

$$\Pr[W_A \cap C] = \Pr[W_A | C] \Pr[C]$$

$$= \frac{1}{2} \cdot \frac{1}{3} \quad \text{(from part (b))}$$

$$= \frac{1}{6}$$

We also know the event $W_A \cap \overline{C}$ occurs exactly when Alice draws a value strictly greater than Bob’s. That is:

$$\Pr[W_A \cap \overline{C}] = \Pr[A_3 \cap B_2] + \Pr[A_3 \cap B_1] + \Pr[A_2 \cap B_1]$$

$$= \frac{1}{3} \cdot \frac{1}{3} + \frac{1}{3} \cdot \frac{1}{3} + \frac{1}{3} \cdot \frac{1}{3} \quad \text{(because } A_i \perp B_j \text{ from part (b))}$$

$$= \frac{1}{3}$$

By the Law of Total Probabilities:

$$\Pr[W_A] = 1 - \Pr[W_A \cap C] - \Pr[W_A \cap \overline{C}] = 1 - \frac{1}{6} - \frac{1}{3} = \frac{1}{2}$$

We now plug these values in:

$$E[G] = (2) \cdot \Pr[W_A \cap C] + (1) \cdot \Pr[W_A \cap C] + (-1.5) \cdot \Pr[W_A]$$

$$= 2 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} - 1.5 \cdot \frac{1}{2}$$

$$= \frac{1}{12}$$

Now, if we play $n$ games, we can apply linearity of expectation to see that Alice’s expected winnings are $\frac{n}{12}$.

4. (10pts) Consider $X$ and $Y$, two independent Bernoulli random variables defined on the same probability space. We are given $\Pr[X = 1] = \frac{1}{3}$ and $\Pr[Y = 1] = \frac{1}{4}$. Compute $E[(X + Y)^2]$.

**Answer**

We expand, using linearity of expectation:

$$E[(X + Y)^2] = E[X^2 + 2XY + Y^2]$$


But for Bernoulli random variables, $X = X^2$:

$$= E[X] + 2E[XY] + E[Y]$$
And since $X \perp Y$, $E[XY] = E[X]E[Y]$

$$= E[X] + 2E[X]E[Y] + E[Y]
= (E[X] + E[Y])^2
= (Pr[X = 1] + Pr[Y = 1])^2
= \left(\frac{1}{3} + \frac{1}{4}\right)^2
= \frac{49}{144}$$

5. (10pts) A fair coin is flipped $2n$ times ($n \geq 1$), independently. Let $X_H$ the random variable that returns the number of heads that occurred and $X_T$ the random variable that returns the number of tails that occurred. Compute $P(X_H > X_T)$.

**Answer**

We first observe that we are working in a uniform probability space of size $2^{2n}$, since any series of tosses is equally likely. Additionally, note that there is a bijection between outcomes where $X_H > X_T$ and $X_T > X_H$, by inverting the results of every flip in the sequence. In other words, $Pr[X_H > X_T] = Pr[X_T > X_H]$.

But we know that our sample space can be partitioned into the events $[X_H > X_T]$, $[X_T > X_H]$, and $[X_H = X_T]$. Thus,

$$1 = Pr[X_H > X_T] + Pr[X_T > X_H] + Pr[X_H = X_T] = 2 Pr[X_H > X_T] + Pr[X_H = X_T]$$

We can count $\binom{2n}{n}$ outcomes where $X_H = X_T$ (choose $n$ spots of the $2n$ for the heads to occur).

Since the probability space is uniform, $Pr[X_H = X_T] = \frac{|\{X_H = X_T\}|}{|\Omega|} = \frac{\binom{2n}{n}}{2^{2n}}$.

Plugging this into our equation from above gives us:

$$Pr[X_H > X_T] = 1 - \frac{\binom{2n}{n}}{2^{2n}} = \frac{2^{2n} - \binom{2n}{n}}{2^{2n+1}}$$

6. (15pts) Let’s call a slug a DAG $G = (V, E)$ with at least 4 vertices, ($|V| \geq 4$), and such that $G$ has exactly one source $r$ and exactly one sink $s$.

(a) (9pts) Draw two different slugs, both with 4 vertices, one of them with 3 edges and the other one with 5 edges.

(b) (6pts) Prove that in any slug, for every node $u$ that is not $r$ or $s$, there exists a directed path from $r$ to $s$ that passes through $u$.

**Answer**

(a) This is a slug with 3 edges:
(b) Consider a maximum-length directed path through \( u \), which we call \( v_0, v_1, \ldots, u, \ldots, v_k \). We claim that the endpoints of this path are \( r \) and \( s \), i.e. \( v_0 = r \) and \( v_k = s \).

We know that \( \text{in}(v_0) = 0 \) or \( \text{in}(v_0) > 0 \). If it is the former, then \( v_0 = r \), since there is only one source in the graph. Otherwise, there must be some vertex \( v \) such that \( v \to v_0 \) is present in the graph. Then \( v \) must already be on the path; otherwise, we could extend it and the path does not have maximum length. Then there is a cycle \( v_0 \to v \to v_0 \) (the first part of the cycle is found by following our longest path). Since we know that our graph is a DAG, this is a contradiction. Hence, it must be the case that \( v_0 = r \).

Similarly, we can show that \( v_k = s \) by instead considering the outdegree of \( v_k \). Thus, this path must lead from \( r \) to \( s \) and contain \( u \), as desired.

## 4 Additional Problems

1. For each statement below, decide whether it is TRUE or FALSE and circle the right one. In each case attach a very brief explanation of your answer.

   (a) The complete bipartite graph \( K_{5,5} \) has a cycle of length 5.

   (b) Let \( A, B \) be events in the same probability space and let \( I_A, I_B \) be their indicator random variables. If \( E(I_A + I_B) = 1 \) then \( P(A) = P(B) \).

   (c) A strongly connected digraph with at least two nodes can have neither sources nor sinks.

   (d) Let \((\Omega, \Pr)\) be a probability space such that \(|\Omega| \geq 2\). Assume that there exists \( u \in \Omega \) such that \( \Pr[u] > 1/2 \). Then, there exists \( v \in \Omega \) such that \( \Pr[v] < 1/2 \).

   (e) Let \( A, B, C \) be three events of non-zero probability in a probability space \((\Omega, \Pr)\). If \( A \cap B = B \cap C, A \perp B, \) and \( B \perp C \) then \( \Pr[A] = \Pr[C] \).
(f) There exists an undirected graph with 23 vertices such that 11 of them have degree 11 and 12 of them have degree 12.

(g) There exists a connected undirected graph with 100 vertices and 50 edges.

**Answer**

(a) FALSE. Bipartite graphs do not have odd cycles; hence $K_{5,5}$ cannot have a 5-cycle.

(b) TRUE. By LOE and the definition of a random variable, we have $E[I_A + I_B] = E[I_A] + E[I_B] = \Pr[A] + \Pr[B] = 1$. So $\Pr[A] = 1 - \Pr[B] = \Pr[\overline{B}]$.

(c) TRUE. Assume for contradiction that there is source $v$. Then there must be no edges leading to $v$, and therefore no paths to $v$. The graph cannot be strongly connected, so we have a contradiction. By similar logic, there cannot be any sinks.

(d) TRUE. Since $|\Omega| \geq 2$, there is at least one $e \in \Omega$ such that $u \neq e$. Assume for contradiction that there is no $v \in \Omega$ such that $\Pr[v] < \frac{1}{2}$. Then $\Pr[e] \geq \frac{1}{2}$. $\Pr[e] + \Pr[u] > 1$, but we know that $\Pr[e] + \Pr[u] \leq 1$ by the definition of a probability space, so we have a contradiction.

(e) TRUE. If $A \cap B = B \cap C$, then $\Pr[A \cap B] = \Pr[B \cap C]$. Since $A \perp B$ and $B \perp C$, we have $\Pr[A] \Pr[B] = \Pr[B] \Pr[C]$. Since we also have $\Pr[B] \neq 0$, we divide both sides by $\Pr[B]$ to have $\Pr[A] = \Pr[C]$.

(f) FALSE. By the Handshaking Lemma, such a graph would have

$$\frac{11 \cdot 11 + 12 \cdot 12}{2} = \frac{265}{2} = 132.5$$

edges. However, a graph cannot have a non-integer number of edges, so such a graph cannot exist.

(g) FALSE. Trees are minimally connected, and for trees we have $|V| = |E| + 1$. Therefore, a minimally connected graph with 100 vertices must have 99 edges; a graph with 100 vertices and 50 edges cannot be connected.

2. (15pts) Let $A, B, C$ be three events in the same probability space such that $A \subseteq B$, $A \subseteq C$, $B \perp C$, and $\Pr[A] = 1$. Prove that $\Pr[A \cap B \cap C] = \Pr[A] \Pr[B] \Pr[C]$.

**Answer**

Since $A \subseteq B$ and $A \subseteq C$ we have $A \subseteq B \cap C$. Therefore, $A \cap B \cap C = A$.

Moreover, by monotonicity of probability, $A \subseteq B$ implies $\Pr[A] \leq \Pr[B]$. Since $1 = \Pr[A] = \Pr[B] \leq 1$ we have $\Pr[B] = 1$.

Similarly, $\Pr[C] = 1$.

Therefore, $\Pr[A \cap B \cap C] = \Pr[A] = 1 = 1 \cdot 1 = \Pr[A] \Pr[B] \Pr[C]$.

3. Let $A, B, C$ be three events of non-zero probability in a probability space $(\Omega, \Pr)$. Prove or disprove: $A \cap B = B \cap C$, $A \perp B$, and $B \perp C$ then $\Pr[A] = \Pr[C]$.

**Answer**

TRUE. $\Pr[A] \Pr[B] = \Pr[A \cap B] = \Pr[B \cap C] = \Pr[B] \Pr[C]$. Since $\Pr[B] \neq 0$, it follows that $\Pr[A] = \Pr[C]$.

4. For each statement below, decide whether it is true or false. In each case attach a very brief explanation of your answer.


(a) There exists an undirected acyclic graph with 5 vertices, 3 connected components, and 2 edges, true or false?

(b) Let $A, B$ be events in a probability space such that $\Pr[A] = 0$ and $\Pr[B] \neq 0$. Then, $\Pr[A \mid B] = 0$, true or false?

(c) For any probability space $(\Omega, \Pr)$ and any event $A \subseteq \Omega$ such that $\Pr[A] \neq 0$ we have $\Pr[\Omega \mid A] = \Pr[A \mid \Omega]$, true or false?

**Answer**

(a) **TRUE.** Take $V = \{1, 2, 3, 4, 5\}$ and $E = \{\{1, 2\}, \{3, 4\}\}$. The three connected components are $\{1, 2\}$, $\{3, 4\}$, and $\{5\}$.

(b) **TRUE.** Since $A \cap B \subset A$ it follows by monotonicity of probability that $0 \leq \Pr[A \cap B] \leq \Pr[A] = 0$ so $\Pr[A \cap B] = 0$.

Therefore $\Pr[A \mid B] = \frac{\Pr[A \cap B]}{\Pr[B]} = 0$.

(c) **FALSE.** Counterexample: let $(\Omega, \Pr)$ be the probability space of one flip of a fair coin and $A$ be the event that the coin shows heads. Then $\Pr[\Omega \mid A] = \frac{\Pr[\Omega \cap A]}{\Pr[A]} = \frac{\Pr[A]}{\Pr[A]} = 1$ and $\Pr[A \mid \Omega] = \frac{\Pr[A \cap \Omega]}{\Pr[\Omega]} = \frac{\Pr[A]}{\frac{1}{2}} = \frac{1}{2} \neq 1$

5. For each statement below, decide whether it is true or false. In each case attach a very brief explanation of your answer.

(a) Let $G$ be an undirected graph in which all vertices have degree 3. Then $G$ has an even number of vertices, true or false?

(b) Let $(\Omega, \Pr)$ be a probability space with 2 or more outcomes and $X : \Omega \to \mathbb{R}$ a random variable such that $\text{Val}(X) = \{-1, 1\}$. If $\mathbb{E}[X] = 0$ then $\Pr[X = 1] = 1/2$, true or false?

(c) Let $E, F$ be two events in a finite probability space. If $|E| = |F|$ then $\Pr[E] = \Pr[F]$, true or false?

(d) If $E, F$ are two events in a finite probability space such that $\Pr[E \cap F] > 0$ then $E$ and $F$ can be disjoint, true or false?

**Answer**

(a) **TRUE.** Let $n$ be the number of vertices and $m$ be the number of edges of $G$. The sum of the degrees of all the vertices is $3n$. By the handshake lemma this equals $2m$. Thus $n$ must be even.

(b) **TRUE.**

$$\mathbb{E}[X] = (-1) \cdot \Pr[X = -1] + (1) \cdot \Pr[X = 1]$$
$$= \Pr[X = 1] - \Pr[X = -1]$$
$$= \Pr[X = 1] - (1 - \Pr[X = 1])$$
$$= 2 \cdot \Pr[X = 1] - 1$$

When $\mathbb{E}[X] = 0$ we compute $\Pr[X = 1] = \frac{1}{2}$. 

10
(c) FALSE. We can find a counterexample by taking $E$ and $F$ to be sets of single outcomes, but each with a different probability. (BTW, this would be a non-uniform prob. space. In a uniform probability space it’s true.)

(d) FALSE. If $\Pr[E \cap F] > 0$, then there is some common, positive-probability outcome in $E$ and $F$, so $E \cap F \neq \emptyset$.

6. Let $E, F$ be two events in a finite probability space such that $\Pr[E \cap F] > 0$. Prove that $\Pr[E \setminus F] + \Pr[F \setminus E] < \Pr[E \cup F]$.

Answer By potato diagram, $(E \setminus F) \cup (E \cap F) \cup (F \setminus E) = E \cup F$. By the Sum Rule, since the LHS sets are pairwise disjoint:

$$\Pr[E \setminus F] + \Pr[E \cap F] + \Pr[F \setminus E] = \Pr[E \cup F]$$

Therefore $\Pr[E \setminus F] + \Pr[F \setminus E] = \Pr[E \cup F] - \Pr[E \cap F] < \Pr[E \cup F]$.

7. For each statement below, decide whether it is true or false. In each case attach a very brief explanation of your answer.

(a) Let $A, B$ be events in a finite probability space such that $\Pr[A] = 1/4$ and $\Pr[A \cup B] = 1/2$. Then, $1/4 \leq \Pr[B] \leq 1/2$, true or false?

Answer TRUE. By Inclusion-Exclusion $1/2 = 1/4 + \Pr[B] - \Pr[A \cap B]$. Hence $\Pr[B] = 1/4 + \Pr[A \cap B]$. Thus $\Pr[B] \geq 1/4$. We still have to show that $\Pr[B] \leq 1/2$. This follows from monotonicity of probability because $B \subseteq A \cup B$.

8. For each statement below, decide whether it is TRUE or FALSE. In each case attach a very brief explanation of your answer.

(a) If $P(A \cap B) \neq 0$ and $P(A|B) = P(B|A)$, then $P(A) = P(B)$, TRUE or FALSE?

(b) Let $X$ and $Y$ be two r.v.s such that $\text{Val}(X)$ and $\text{Val}(Y)$ are subsets of $[0, \infty)$ and such that $E(X) = E(Y) = 10$. Then $P(X < 100 - Y) = 0$, TRUE or FALSE?

(c) In a graph with $n \geq 3$ vertices and 1 (one) edge, the number of connected components is $n - 2$, TRUE or FALSE?

(d) If $|A| = 2$ then there exists exactly two equivalence relation on $A$, TRUE or FALSE?

(e) Let $n \geq 3$. There are exactly $\binom{n}{2}$ paths of length $\geq 1$ in $C_n$, TRUE or FALSE?

Answer

(a) TRUE. $\Pr[A \mid B] = \frac{\Pr[A \cap B]}{\Pr[B]}$ and $\Pr[B \mid A] = \frac{\Pr[A \cap B]}{\Pr[A]}$. Equating these, dividing both sides by $\Pr[A \cap B]$, and multiplying both sides by $\Pr[A] \cdot \Pr[B]$ we obtain $\Pr[A] = \Pr[B]$. (Not required, but observe that since $A \cap B \subseteq A$ we have $\Pr[A] \neq 0$. Similarly $\Pr[B] \neq 0$. So the two conditional probabilities are well-defined.)

(b) FALSE. Suppose, toward a contradiction, that $\Pr[X < 100 - Y] = 0$. Then $0 = \Pr[X < 100 - Y] = \Pr[X + Y < 100] = 1 - \Pr[X + Y \geq 100]$. Hence $\Pr[X + Y \geq 100] = 1$. 

11
However, by Linearity of Expectation $E[X + Y] = E[X] + E[Y] = 20$ and by Markov’s Inequality $\Pr[X + Y \geq 100] = \Pr[X + Y \geq 5 \cdot 20] = \Pr X + Y \geq 5 \cdot E[X + Y] \leq \frac{1}{5}$. Contradiction. Note that we prove here a stronger statement, that it is in fact impossible for this to be the case.

(c) FALSE. Suppose, toward a contradiction, that the graph has $n$ vertices, 1 edge and $n - 2$ connected components. We have shown in class that in any graph $|E| \geq |V| - |CC|$. Plugging in we get $1 \geq n - (n - 2) = 2$, which is impossible.

(d) TRUE. Let $\rho$ be an equivalence relation on $A$. We always have $a_0 \rho a_0$ and $a_1 \rho a_1$. Then, we may have $a_0 \rho a_1$ or not. Thus, there are only two possible equivalence relations, $\{(a_0, a_0), (a_1, a_1)\}$ and $A \times A$.

(e) FALSE. Recall that $C_n$ has $n$ vertices and $n$ edges. Each path of length $\geq 1$ has two distinct endpoints. Therefore, there are at least $\binom{n}{2}$ paths. But there are more path than that. Indeed, let $u, v$ be two distinct vertices in $C_n$. Because $n \geq 3$ there exists at least one more distinct vertex, $w$ such that $w \neq u$ and $w \neq v$. On $C_n$, $w$ is encountered when going clockwise from $u$ to $v$ and or when going counterclockwise from $u$ to $v$, but not in both cases. Thus $C_n$ has at least two paths with endpoints $u$ and $v$, one with $w$ and one without.

9. Let $(\Omega, P)$ be a probability space and let $X$ be a random variable defined on $\Omega$ such that $\text{Val}(X) = \{a, b\}$ where $a < b$. We also denote $\mu = E(X)$.

(a) Express $P(X \leq (a + b)/2)$ in terms of $a, b$ and $\mu$.

(b) Let $a = -1$ and $b = 1$. Show that if $E(X) = 0$ then there exists an event $A \subseteq \Omega$ such that $P(A) = 1/2$.

**Answer**

(a) Since $a < b$ we have $a < \frac{a+b}{2} < b$. Then, $\{\omega \mid X(\omega) \leq \frac{a+b}{2}\} = \{\omega \mid X(\omega) = a\}$. Therefore $\Pr\left[X \leq \frac{a+b}{2}\right] = \Pr[X = a]$ so it suffices to express $\Pr[X = a]$ in terms of $\mu, a, b$.

$$\mu = E[X] = a \Pr[X = a] + b \Pr[X = b]$$

But we also have $\Pr[X = a] + \Pr[X = b] = 1$:

$$= a \Pr[X = a] + b(1 - \Pr[X = a])$$

$$\mu = (a - b) \Pr[X = a] + b$$

Hence

$$\Pr[X = a] = \frac{\mu - b}{a - b} = \frac{b - \mu}{b - a}$$

We conclude that

$$\Pr\left[X \leq \frac{a+b}{2}\right] = \frac{b - \mu}{b - a}$$
(b) Plugging \( a = -1, b = 1, \mu = 0 \) in to the result of part (a), we get

\[
\Pr \left[ X \leq \frac{a + b}{2} \right] = \frac{1 - 0}{1 - (-1)} = \frac{1}{2}
\]

Therefore we can take \( A = \{ w \mid X(w) \leq \frac{-1+1}{2} \} = \{ w \mid X(w) \leq 0 \} \)

10. In this problem the graphs are undirected.

(a) Draw an example of a tree that must have exactly three leaves and has the fewest possible number of edges.

(b) Prove that a tree with at least three leaves must have at least one node of degree \( \geq 3 \). (Hint: it is possible to prove this by induction but the proof is longish. Try to find a shorter proof first.)

**Answer**

(a) Consider the following graph:

```
1 -- 2
|   |
4 --- 3
```

We can bound the number of edges by looking at the number of vertices (for trees, \( |V| = |E| + 1 \)). There must be at least 3 vertices, since there are 3 leaves, but we cannot have a 3-vertex tree where every vertex is a leaf (this would violate the Handshaking Lemma). Thus, we know that any such tree must have at least 4 vertices, which means it needs at least 3 edges.

(b) Assume for the sake of contradiction that a tree \( T \) with vertex set \( V \) and edge set \( E \) exists such that it has \( \ell \geq 3 \) leaves, but all nodes have degree \( \leq 2 \). By the Handshaking Lemma:

\[
|E| = \ell \cdot 1 + (|V| - \ell) \cdot 2
\]

\[
= |V| - \frac{\ell}{2}
\]

\[
\leq |V| - \frac{3}{2}
\]

However, we know by the definition of a tree that \( |V| = |E| + 1 \). Thus, we have a contradiction, and there must be at least one node in the tree with degree \( \geq 3 \).

11. Weird Al (WAI) is playing with his coins. The game uses two *fair coins* and one *urn*. The result of the game is one of \( H \) (heads) or \( T \) (tails) and is determined as follows:

- WAI places both coins in the urn.
- WAI reaches inside the urn and (a) with probability 2/3 WAI grabs one of the coins and tosses it, OR (b) with probability 1/3 WAI grabs both coins, then tosses them separately in some order (doesn’t matter which order).
If WAl has tossed just one coin then whatever that coin shows is the result of the game. If WAl has tossed both coins then applying the weird $\otimes$ operation to what the two coins show is the result of the game, where $T \otimes T = T$, $T \otimes H = H$, $H \otimes T = H$, and $H \otimes H = T$.

(a) Draw the “tree of possibilities” diagram for WAl’s game.
(b) Calculate the probability that the result of the game is $H$.
(c) What simpler game could Weird Al play that would give him exactly the same odds?

**Answer**

We define the following events: $C_1$ indicating the event WAl flips one coin, $H_1$ indicating the event that flip is a head, $C_2$ indicating the event he flips 2 coins, $HH, HT, TH, TT$ indicating the sequence of flips he gets.

(a) The tree of possibilities:

(b) The outcome of the game is $H$ if 1) WAl flips one coin (event $C_1$) and it lands on $H$ (event $H_1$) or 2) WAl flips two coins (event $C_2$) and gets one $H$, one $T$ (event $HT$). If WAl flips two coins, he will get one $H$, one $T$ with probability $\frac{1}{2}$, since both HT and TH will yield a head, and both occur with probability $\frac{1}{3} \cdot \frac{1}{2} = \frac{1}{4}$.

$$Pr[H] = Pr[C_1 \cap H_1] + Pr[C_2 \cap HT]$$

$$= Pr[C_1]Pr[H_1|C_1] + Pr[C_2]Pr[HT|C_2]$$

$$= \frac{2}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot \frac{1}{2}$$

$$= \frac{1}{2}$$

(c) Since the result of WAl’s game is $H$ with probability $\frac{1}{2}$ and $T$ otherwise, this is equivalent to simply flipping a fair coin once.
12. Let $S$ be the probability space $(\Omega, P)$ with $\Omega = \{f : \mathbb{R} \to \mathbb{R} \mid f(x) = ax + b\}$ such that $a, b \in \mathbb{N}$, $1 \leq a \leq 10$, and $1 \leq b \leq 10$, and $P$ is the uniform probability distribution on $\Omega$. For each natural number $k$, let $X_k$ be the random variable that takes on the value of $f(k)$.

(a) What is $E[X_5]$?
(b) What is $P(X_5 > 1)$?
(c) What is $E[X_k]$ in terms of $k$?
(d) Now consider the probability space $S' = (d\Omega dx, P)$, where $d\Omega dx = \{\frac{df}{dx} \mid f \in \Omega\}$. Let $Y$ be the random variable that takes on $f(5)$. What is $E[Y]$?

**Answer**

(a) We apply the definition of expectation:

\[
E [X_5] = \sum_{\omega \in \Omega} X_5(\omega) \Pr[\omega]
\]

Since the probability distribution is uniform, we can count the outcomes and divide by $|\Omega|:

\[
= \sum_{i=1}^{10} \sum_{j=1}^{10} \frac{5i + j}{100} \\
= \sum_{i=1}^{10} \frac{50i + 55}{100} \\
= \sum_{i=1}^{10} \frac{i}{2} + \sum_{i=1}^{10} \frac{55}{100} \\
= \frac{1}{2} \left( \sum_{i=1}^{10} i \right) + \frac{11}{2} \\
= \frac{11}{2} \left( \frac{100}{2} \right) + \frac{11}{2} \\
= \frac{33}{2}
\]

(b) Since the minimum value of $X_5$ is $1 \cdot 5 + 1 = 6$, $\Pr [X_5 > 1] = 1$.

(c) We again proceed by using the definition of expectation:

\[
E [X_k] = \sum_{\omega \in \Omega} X_k(\omega) \Pr[\omega]
\]

\[
= \sum_{i=1}^{10} \sum_{j=1}^{10} \frac{ki + j}{100} \\
= \sum_{i=1}^{10} \frac{10ki + 55}{100} \\
= \frac{k}{10} \left( \sum_{i=1}^{10} i \right) + \frac{11}{2} \\
= \frac{11}{2} (k + 1)
\]
(d) Since this new probability consists of constant functions \( f(x) = c \) for \( 1 \leq c \leq 10 \), and each has the same probability, the average value of \( f(x) \) for any \( x \) must be 5.5.

13. You have a standard deck of 52 cards, from which you draw 13 cards, without replacement.

a) Given that you drew the 4 of spades, what is the probability that all the other cards that you drew are aces, twos, threes, or fours?

b) Define \( S \) to be the number of spades you draw. What is \( E[S] \)?

c) For what \( s \in \text{Val}(S) \) do we have the maximum value of \( \Pr[S = s] \)?

d) Suppose that the number of spades you have in your hand is equal to the number you found in part (c). What is the probability that the sum of their numerical values (letting J = 11, Q = 12, K = 13, A = 1) is odd?

Answer

(a) You know that one of the cards you drew is the 4 of spades. You know that 12 cards remain in your hand, and there are exactly 51 cards you could have chosen from, of which 15 are aces, twos, threes, and fours. Any combination of 12 cards is equally likely; that is, our probability distribution is uniform. Thus, defining \( A \) as the event that you draw only aces, twos, threes and fours, given that one of your cards is the 4 of spades:

\[
P[A] = \frac{\binom{15}{12}}{\binom{51}{12}} = \frac{15!39!}{3!51!}
\]

(b) We define indicator variables \( s_i \) for \( i = 1, 2... 13 \). \( s_i = 1 \) if the \( i \)th card is a spade; otherwise, \( s_i = 0 \). We know that

\[
E[S] = E[s_1 + s_2 + ... s_{13}]
= \sum_{i=1}^{13} E[s_i]
= \sum_{i=1}^{13} \Pr[s_i = 1]
\]

Each card has a \( \frac{13}{52} = \frac{1}{4} \) probability of being a spade. Thus,

\[
E[S] = \sum_{i=1}^{13} \frac{1}{4} = \frac{13}{4}
\]

Alternatively, you could note that, letting \( H, C, \) and \( D \) be the number of hearts, clubs, and diamonds respectively,

\[
E[S] = E[H] = E[C] = E[D]
\]

and, since \( S + H + C + D = 13 \),

\[
\]

\[
4 \cdot E[S] = 13
E[S] = \frac{13}{4}
\]
(c) For \( s = 0, 1, \ldots, 13 \), we can express the probability of choosing \( s \) spades by determining how many ways there are to choose 13 cards at random (our outcome space) and how many ways there are to choose \( s \) spades and \( 13 - s \) other cards at random (our desired event). Note that our probability distribution is uniform.

\[
Pr[S = s] = \frac{\binom{13}{s} \binom{39}{13-s}}{\binom{52}{13}}
\]

Since all of these probabilities have the same denominator - the same outcome space - we only compare the numerator to determine the maximum \( Pr[S = s] \) value.

\[
\binom{13}{s} \binom{39}{13-s} = \frac{13!}{s!(13-s)!}(26+s)!
\]

To find where the maximum probability lies, we find the ratio

\[
\frac{Pr[S = s]}{Pr[S = s+1]} = \frac{(s+1)!(12-s)!(12-s)!(27+s)!(13-s)!}{s!(13-s)!((13-s)!(26+s)!}
\]

By calculating the values for which this ratio is \( > 1 \) or \( < 1 \), we can determine where the maximum lies:

\[
Pr[S = s] > Pr[S = s + 1]
\]

\[
\frac{(s+1)(27+s)}{(13-s)(13-s)} > 1
\]

\[
(s+1)(27+s) > (13-s)(13-s)
\]

\[
s^2 + 28s + 27 > s^2 - 26s + 169
\]

\[
s > \frac{142}{54} \approx 2.63
\]

Clearly, we can simply reverse the sign in all of the above inequalities to have that \( Pr[S = s] < Pr[S = s + 1] \) if \( s < \frac{142}{54} \approx 2.63 \). Thus, the value of \( Pr[S = s] \) is increasing from 0 to 2 and decreasing after 3. Either \( s = 2 \) or \( s = 3 \) yields the greatest \( Pr[S = s] \) value. Plugging in, we find

\[
Pr[S = 2] = \frac{3 \cdot 29}{11 \cdot 11} = \frac{87}{121} < 1
\]

\[
Pr[S = 2] < Pr[S = 3]
\]

\( s = 3 \) yields the greatest \( Pr[S = s] \) value; it is most likely that we draw 3 spades. In fact, plugging in our values, we will find

\[
Pr[S = 3] \approx .286
\]

(d) Since you have three cards, the only ways to have an odd sum are with 3 odd cards (call this event \( O_1 \)) or 1 odd and 2 even cards (call this event \( O_2 \)). Our outcome space \( \Omega \) is the possible ways of choosing 3 spades out of the 13 available. Note again that our probability distribution is uniform, so we can simply sum over the two cases and divide by total number of possible outcomes. There are 7 odd spades and 6 even spades. We have:

\[
\frac{|O_1| + |O_2|}{|\Omega|} = \frac{\binom{7}{3} + \binom{7}{1} \binom{6}{2}}{\binom{13}{3}} = \frac{140}{286} \approx .49
\]
14. (35pts) Let \( n \geq 2 \) be a natural number. Consider the set \( W \) of nonempty subsets of \([1..n]\). By the Well-Ordering Principle, each \( A \in W \) has a unique least element. Let’s denote this element by \( \min(A) \). Define a binary relation \( \sim \) on \( W \) as follows: for any \( A, B \in W \)

\[
A \sim B \quad \text{iff} \quad \min(A) = \min(B)
\]

It is easy to see that \( \sim \) is an equivalence relation on \( W \) (you don’t have to prove this). In answering the parts of this question make sure your answer works for any \( n \geq 2 \).

(For each of the questions below give the answer and an explanation of how you derived it. No proofs required.)

(a) Give an example of \( A, B \in W \) such that \( A \neq B \) but \( A \sim B \).

(b) Express in terms of \( n \) the number of distinct equivalence classes determined by \( \sim \).

(c) What is the size of the smallest equivalence class determined by \( \sim \)?

(d) Express in terms of \( n \) the size of a largest equivalence class determined by \( \sim \).

(e) Let \( A \in W \). Express in terms of \( n \) and of \( k = \min(A) \) the size of the equivalence class of \( A \).

Answer

(a) Example: \( A = \{1\}, B = \{1, 2\} \). Clearly \( A \neq B \) but \( \min(A) = 1 = \min(B) \) hence \( A \sim B \).

(b) All the subsets in an equivalence class have the same \( \min \) therefore the classes correspond to the elements of \([1..n]\). There are therefore \( n \) such equivalence classes.

(c) The equivalence class with just one element \( \{n\} \) must be the smallest, all the other equivalence classes (since \( n \geq 2 \) there are other equivalence classes) have at least two elements.

(d) We claim that the largest equivalence class is the one corresponding to \( \min(A) = 1 \). Observe that this class contains \( 2^{n-1} \) elements, since it contains exactly the elements of the form \( 1 \cup B \) where \( B \subseteq [2..n] \). But this is strictly more than half of the \( 2^n - 1 \) total elements of \( W \), so this must be the largest equivalence class.

(e) A subset \( B \in W \) is such that \( B \sim A \) iff \( \min(B) = k \). Therefore iff \( B \) consists of \( k \) and possibly some elements of \([k+1..n]\). There are \( 2^{n-k} \) subsets of \([k+1..n]\). Therefore the size of the equivalence class of \( A \) is \( 2^{n-k} \).

15. (20pts) A biased coin shows heads with probability \( 1/3 \) and tails with probability \( 2/3 \). The coin is flipped \( n \geq 3 \) times, independently.

(a) Compute the expected number of occurrences of consecutive heads, tails, tails.

(b) Recall Markov’s Inequality: \( P(X \geq cE(X)) \leq 1/c \) where \( c > 0 \), \( X \) returns only nonnegative values, and \( E(X) \neq 0 \). Using Markov’s Inequality find an upper bound for the probability that the number of occurrences of consecutive heads, tails, occurrences is bigger than or equal to \( 8n/9 \).

Answer
(a) Denote by $H_i$ the event of flipping a head on the $i^{th}$ flip and by $T_i$ the event of flipping a tail on the $i^{th}$ flip. We also define a random variable $A$ denoting the number of consecutive $H, T, T$ sequences, and indicator random variables $A_i$, $1 \leq i \leq n - 2$ such that $A_i = 1$ if we have $H_i \cap T_{i+1} \cap T_{i+2}$; otherwise, $A_i = 0$. Then it is clear that

$$A = \sum_{i=1}^{n-2} A_i$$

By LOE, we have

$$E[A] = \sum_{i=1}^{n-2} E[A_i] = \sum_{i=1}^{n-2} \Pr[A_i = 1] = \sum_{i=1}^{n-2} \Pr[H_i \cap T_{i+1} \cap T_{i+2}]$$

Given the independence of individual coin tosses:

$$= \sum_{i=1}^{n-2} \Pr[H_i] \Pr[T_{i+1}] \Pr[T_{i+2}] = \sum_{i=1}^{n-2} \frac{1}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} = \sum_{i=1}^{n-2} \frac{4}{27} = \frac{4(n-2)}{27}$$

(b) We rearrange so that we can apply Markov’s inequality:

$$\Pr\left[A \geq \frac{8n}{9}\right] = \Pr\left[A \geq \frac{8n}{9} \cdot \frac{27}{4(n-2)} \cdot \frac{4(n-2)}{27}\right] = \Pr\left[A \geq \frac{6n}{n-2} \cdot \frac{4(n-2)}{27}\right] = \Pr\left[A \geq \frac{6n}{n-2} \cdot E[A]\right] \leq \frac{1}{\frac{6n}{n-2}} = \frac{n-2}{6n} \quad \text{(By Markov’s Inequality)}$$

16. Let $G$ be a bipartite graph in which every connected component is a cycle.

(a) Draw the smallest such $G$. (Just the drawing, no need for explanation)
(b) Prove that, not just in the smallest, but in any such \(G\) the number of red nodes is equal to the number of blue nodes.

**Answer**

(a) \(G = (\{r_1, r_2, b_1, b_2\}, \{r_1-b_1, b_1-r_2, r_2-b_2, b_2-r_1\})\)

Any cycle has at least 3 vertices, but we can’t have odd cycles in a bipartite graph. Hence, the smallest such graph has at least 4 vertices.

(b) It suffices to show that in each connected component there are as many red as blue nodes. Adding over all components we get the result.

Each connected component is a cycle. In a bipartite graph cycles must have even length and they must alternate the red and blue nodes, therefore they have as many red as blue nodes. Done.

17. Find the minimum value for \(k\) that ensures that any graph in which every vertex has degree \(k\) or more has at least one cycle of even length. Prove your answer.

**Answer** \(k = 3\).

Let \(G\) be an arbitrary graph where every vertex has degree 3 or more. Consider a maximal path (a longest path), \(P\), in \(G\). Suppose \(P = v_1-v_2-\cdots-v_p\). All of \(v_1\)’s neighbors must lie on \(P\), otherwise, \(P\) is not a maximal path. (If \(u \) is a neighbor of \(v_1\) and \(u \) is not in \(P\), then \(u-v_1-v_2-\cdots-v_n\) is a strictly longer path.) So \(v_1\) has at least 3 neighbors in \(P\), call them \(v_i, v_j, v_k\), where \(i < j < k\). Consider the following 3 paths in the graph:

- \(v_1-v_j\)
- \(v_1-v_i-\cdots-v_j\) (traversing \(P\) between \(v_i\) and \(v_j\))
- \(v_1-v_k-\cdots-v_j\) (traversing the reverse of \(P\) between \(v_k\) and \(v_j\))

Notice that concatenating any two of these paths results in a cycle. By the Pigeonhole Principle, the lengths of 2 of these 3 paths must have the same parity (both odd or both even). So concatenate those 2 paths for a cycle of even length.

Now we show that \(k = 3\) is the minimum value of \(k\).

Consider \(K_3\), the complete graph on 3 vertices. It has just one cycle of length 3. The degree of each vertex is 2, yet there are no even length cycles in the graph. Therefore \(k\) must be strictly bigger than 2.

18. (10pts) Let’s call Peano-digraph a digraph in which every vertex has outdegree 1.
(a) Prove that any Peano-digraph that is strongly connected is, in fact, a directed cycle.
(b) Count the number of different Peano-digraphs whose set of vertices is \([1..n]\), where \(n\) is a positive integer?

**Answer**

(a) Additional problems 1(c) gives us that there can be no sources or sinks in a strongly connected graph. Since every vertex has outdegree 1, we know that there are \(|V|\) edges total. We omit here a simple proof by contradiction and assert that we must have that each vertex also has in-degree 1.

(b) There are \(n!\) different possible permutations of \(n\), representing the possible orders in which a cycle can visit the \(n\) vertices. However, this overcounts by a factor of \(n\), as count each cycle \(n\) times - once for each of the \(n\) vertices in it from which we can start our ordering. We divide by \(n\) to have \((n−1)!\) possible Peano-digraphs total.

19. (50 pts) For each statement below, decide whether it is TRUE or FALSE and circle the right one. In each case attach a very brief explanation of your answer.

(a) We toss 3 fair coins \(C_1, C_2, C_3\) independently of each other. Then the event \((C_1 = C_2)\) is independent of the event \((C_2 = C_3)\), true or false?

(b) A digraph in which all vertices have outdegree 1 is strongly connected, true or false?

(c) Let \(n \geq 2\). The complete bipartite graph \(K_{n,n}\) has at least \((n!)^2\) distinct subgraphs isomorphic to \(C_{2n}\), true or false?

For the two parts below, use the following definition: for any digraph \(G = (V, E)\) without self-loops and without cycles of length 2 define an undirected graph \(G^u = (V, E^u)\) that has the same vertices as \(G\) and moreover in \(G^u\) we have an edge \(v\to w\) whenever we have the edge \(v\to w\) or the edge \(w\to v\) in \(G\).

(d) If \(G\) is strongly connected then \(G^u\) is connected, true or false?

(e) If \(G\) is a DAG then \(G^u\) is acyclic, true or false?

(f) Let \(G\) be an undirected graph in which all vertices have degree 3. Then \(G\) has an even number of vertices, true or false?

(g) Let \((\Omega, \Pr)\) be a probability space with 2 or more outcomes and \(X : \Omega \to \mathbb{R}\) a random variable such that \(\text{Val}(X) = \{-1, 1\}\). If \(E(X) = 0\) then \(P(X = 1) = 1/2\), true or false?

**Answer**

(a) TRUE. \(\Pr [C_1 = C_2] = \frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{2} = \frac{1}{2}\). Similarly \(\Pr [C_2 = C_3] = \frac{1}{2}\). Also,

\[
\Pr [(C_1 = C_2) \land (C_2 = C_3)] = \Pr [C_1 = C_2 = C_3] = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}
\]

Since \(\frac{1}{4} = \frac{1}{2} \times \frac{1}{2}\), the two events are independent.

(b) FALSE. Counterexample: \(\{a, b, c\}, \{a\to b, b\to c, c\to b\}\). But \(a\) is not reachable from \(b\) (or \(c\)).
(c) TRUE. Indeed consider a permutation \( r_1, \ldots, r_n \) of the red nodes of \( K_{n,n} \) and a permutation \( b_1, \ldots, b_n \) of the blue nodes of \( K_{n,n} \). From any such pair of permutations (and there are \((n!)^2\) such pairs) form a distinct cycle of length \( 2n: r_1 - s_1 - \ldots - r_n - s_n - r_1 \).

(d) TRUE. Let \( v, w \) be two vertices in \( G \) (hence in \( G^u \)). Since \( G \) is strongly connected, there exists a directed walk \( v \rightarrow \cdots \rightarrow w \) in \( G \). Erasing the direction of the edges in this walk gives a walk \( v - \cdots - w \) in \( G^u \).

(e) FALSE. Counterexample: \( G = (\{a,b,c\}, \{a \rightarrow b, a \rightarrow c, c \rightarrow b\}) \) is a DAG but \( G^u \) is a cycle of length 3.

(f) TRUE. Let \( n \) be the number of vertices and \( m \) be the number of edges of \( G \). The sum of the degrees of all the vertices is \( 3n \). By the handshake lemma this equals \( 2m \). Thus \( n \) must be even.

(g) TRUE.

\[
E[X] = (-1) \cdot \Pr[X = -1] + (1) \cdot \Pr[X = 1]
= \Pr[X = 1] - \Pr[X = -1]
= \Pr[X = 1] - (1 - \Pr[X = 1])
= 2 \cdot \Pr[X = 1] - 1
\]

When \( E[X] = 0 \) we compute \( \Pr[X = 1] = \frac{1}{2} \).

20. (35pts) Let \( n \geq 2 \) be a natural number. Consider the set \( W \) of nonempty subsets of \([1..n]\). By the Well-Ordering Principle, each \( A \in W \) has a unique least element. Let’s denote this element by \( \min(A) \). Define a binary relation \( \sim \) on \( W \) as follows: for any \( A, B \in W \)

\[
A \sim B \iff \min(A) = \min(B)
\]

It is easy to see that \( \sim \) is an equivalence relation on \( W \) (you don’t have to prove this). In answering the parts of this question make sure your answer works for any \( n \geq 2 \).

(For each of the questions below give the answer and an explanation of how you derived it. No proofs required.)

(a) Give an example of \( A, B \in W \) such that \( A \neq B \) but \( A \sim B \).

(b) Express in terms of \( n \) the number of distinct equivalence classes determined by \( \sim \).

(c) What is the size of the smallest equivalence class determined by \( \sim \)?

(d) Express in terms of \( n \) the size of a largest equivalence class determined by \( \sim \).

(e) Let \( A \in W \). Express in terms of \( n \) and of \( k = \min(A) \) the size of the equivalence class of \( A \).

Answer This is a duplicate of 14.

21. Let \( G \) be a bipartite graph in which every connected component is a cycle.

(a) Draw the smallest such \( G \). (Just the drawing, no need for explanation)

(b) Prove that, not just in the smallest, but in any such \( G \) the number of red nodes is equal to the number of blue nodes.
22. Recall that $K_{m,n}$, $m, n \geq 1$ is the complete bipartite undirected graph with $m$ vertices colored (say) red, and $n$ vertices colored (say) blue and with edges between any red node and any blue node.

(a) Draw $K_{3,2}$. Name the red vertices $r_1, r_2, r_3$ and the blue vertices $b_1, b_2$.

(b) Draw a spanning tree of $K_{3,2}$ using the same vertex names as in part 22a.

(c) How many edges do we have to delete from $K_{m,n}$ so we are left with a spanning tree? Give the answer in terms of $m$ and $n$ and a short explanation of how you obtained it.

(d) Assume that $m, n \geq 2$. What is the biggest length that a cycle can have in $K_{m,n}$? Give the answer in terms of $m$ and $n$ and a short explanation of how you obtained it.

**Answer**

(a) The graph is as follows:

```
   r1
   |
   v
   r2
   |
   v
   r3

   b1
   |
   v
   b2
```

(b) One spanning tree appears below (there are other valid examples):

```
   r1
   |
   v
   r2
   |
   v
   r3

   b1
   |
   v
   b2
```

(c) There are $mn$ edges in $K_{m,n}$ since there is exactly one edge for each pairing of a red with a blue vertex, and there are $mn$ such pairs. For any tree, the number of edges is one less than the number of vertices, so since any spanning tree of $K_{m,n}$ has $m + n$ vertices, it must have $m + n - 1$ edges. Thus, we must delete $mn - (m + n - 1) = (m - 1)(n - 1)$ edges.
(d) Any cycle in $K_{m,n}$ must alternate between red and blue vertices. Since we can’t repeat vertices (except the first and last), the length cannot be more than double the number of red vertices or more than double the number of blue vertices, so we can’t have a length that is more than $\min(2m, 2n) = 2 \min(m, n)$.

To show that this length is indeed attainable, we consider the cycle $r_1-b_1-r_2-b_2-\cdots-r_k-b_k-r_1$ where $k = \min(m, n)$.

23. Recall the complete undirected graph on $n$ vertices, $K_n$. Prove that for any $n \geq 4$ it is possible to assign direction to each of the edges of $K_n$ such that the resulting digraph has exactly $n-2$ strongly connected components.

**Answer**

**Solution One**

We proceed by induction on $n$.

**Base case:** $n = 4$. If our graph is

\[ \{(a, b, c, 1), (a, b, c), (c, a), (a, 1), (b, 1), (c, 1)\} \]

We can see that the $n-2=2$ strongly connected components are $\{a, b, c\}$ and $\{1\}$ since the subgraph induced on $\{a, b, c\}$ is a cycle and since $a$, $b$, and $c$ are not reachable from 1.

**Induction step:** Let $k$ be an arbitrary integer strictly greater than 3, and assume that there is a digraph $G = (V, E)$ with exactly $k-2$ strongly connected components that is the result of assigning direction to each of the edges of $K_k$ (our induction hypothesis). Now consider the graph $G'$ formed by adding a vertex $v$ to $G$ and directing edges to $v$ from every other vertex. Since no other vertex is reachable from $v$, it forms its own strongly connected component, so $G'$ has $(k+1) - 2$ strongly connected components, and the induction step holds.

**Solution Two**

Let’s name the vertices $1, 2, \ldots, n-3, a, b, c$ (check that there are indeed $n$ vertices). We have three types of edges in $K_n$: edges between two numbered vertices, edges between two vertices labeled with letters, and edges between one vertex labeled with a number and one with a letter. For the edges with numbers, we orient the edge so that it points from the lower number to the higher number. For the edges with letters, we use the edges $(a, b)$, $(b, c)$, and $(c, a)$. For the edges between letters and numbers, we orient the edge so that it points from the letter to the number.

We claim that each vertex $1, 2, \ldots, n-3$ forms a distinct strongly connected component and that $\{a, b, c\}$ forms another. To prove this, we show that no two vertices from different strongly connected components are mutually reachable from each other.
Since every edge that begins in a numbered vertex ends in a numbered vertex, we cannot have a path from a numbered vertex to a non-numbered vertex, so the lettered vertices are not reachable from the numbered vertices and thus cannot be in the same strongly connected component. In addition, since the only edge from any numbered vertex is to a higher-numbered vertex, we cannot have any paths from a higher-numbered vertex to a lower-numbered vertex, and thus the lower-numbered vertices are not reachable from the higher-numbered vertices; therefore no two numbered vertices are in the same strongly connected component, and since they are all separate from the letters, they each constitute their own strongly connected component. Finally, since $a$, $b$, and $c$ are part of the cycle $a \rightarrow b \rightarrow c \rightarrow a$, they are in the same strongly connected component.

We have shown that the strongly connected components are $\{1\}, \{2\}, \ldots, \{n-3\}, \{a,b,c\}$. Thus there are exactly $n - 2$, and we are done.

24. Let $n \geq 1$ be a natural number. Draw (use dot dot dot) an example of a DAG with exactly $3n + 1$ vertices among which there is exactly one source $s$ and exactly one sink $t$ such that there are exactly $3^n$ distinct directed paths from $s$ to $t$.

**Answer** The source is node 1 and the sink is node 0. There are 3 paths from 1 to 2, three paths from 2 to 3, etc., and finally three paths from $n$ to 0. By the product rule there are $3^n$ paths from 1 to 0.