On Wednesday May 3 we will have our final exam starting at 9AM. The exam will take place in CHEM 102 and in SKIR AUD.

The students whose last name begins with a letter in the range A-N will take the exam in CHEM 102. The students whose last name begins with a letter in the range O-Z will take the exam in SKIR AUD. The exam will last for 120 minutes. Please be in CHEM 102 or SKIR AUD promptly at 9AM so we have time to seat everybody properly. You may want to set up a buddy system by which you make sure your buddy woke up and vice versa.

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

Some of the solutions to the these problems will be posted as early as Sunday evening.

Val will hold a review session on Monday May 1, 6-7:30PM in Heilmeier Hall (TOWNE 100).

The TAs will also hold a number of minireviews, schedule posted on Piazza.

1 Readings

STUDY IN-DEPTH... ...the posted notes for lectures 1–23.

STUDY IN-DEPTH... ...the posted guides for recitations 1–10.

STUDY IN-DEPTH... ...the posted solutions to homeworks 1–10. Compare with your own solutions.

STUDY IN-DEPTH... ...the solutions to the mock exam and the additional problems contained in this document.

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 (100 minutes for 200 points)

1. (40 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) 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?

**Solution:**
FALSE. Any such subgraph must contain every vertex of the graph. Now 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 \( K_{n,n} \). Any such pair of permutations (and there are \((n!)^2\) such pairs) forms a cycle of length \( 2n \): \( r_1 - s_1 - \cdots - r_n - s_n - r_1 \). You can also verify that each such cycle also corresponds to one such pair of permutations. However, each cycle can contribute \( n \) such pairs (we can choose which of the \( n \) red nodes to start with), and we also count both "directions" of the cycle (i.e. reading it clockwise and counterclockwise). Thus, we can divide by \( 2n \) to get a total of \( \frac{(n!)^2}{2n} \) such isomorphic subgraphs.

NOTE: This problem is more difficult than intended, and more difficult than what you should see on the final.

(b) 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?

**Solution:**
TRUE.

\[
\Pr[C_1 = C_2] = \Pr[(C_1 = H) \cap (C_2 = H)] + \Pr[(C_2 = T) \cap (C_2 = T)] = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{2}
\]

Similarly

\[
\Pr[C_2 = C_3] = \frac{1}{2}
\]

Also,

\[
\Pr[(C_1 = C_2) \cap (C_2 = C_3)] = \Pr[(C_1 = H) \cap (C_2 = H) \cap (C_3 = H)] + \Pr[(C_1 = T) \cap (C_2 = T) \cap (C_3 = T)]
\]

\[
= \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2}
\]

\[
= \frac{1}{4}
\]

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

(c) Let \( e_1 \) and \( e_2 \) be two tautologies (boolean expressions that are true for every truth assignment to their variables) and let \( p \) be a boolean variable. Then \((e_1 \Rightarrow p) \land (\neg p \Rightarrow e_2)\) is also a tautology.

**Solution:**
FALSE. \((e_1 \Rightarrow p) \land (\neg p \Rightarrow e_2)\) is not a tautology because there exist truth assignments that make it false. Namely, assign \( p = F \). Then we get \((T \Rightarrow F) \land (T \Rightarrow T) = F \land T = F\).
(d) If we add one edge to a cycle the resulting graph cannot be bipartite.

**Solution:**
FALSE. Consider $C_6$ and add an edge between vertex 1 and vertex 4. Coloring the odd vertices red and the even vertices blue gives a proper 2-coloring of the resulting graph.

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

**Solution:**
TRUE. Suppose for contradiction it has a source $r$. Since there is at least another node $u$ and since the digraph is strongly connected we must have a path $u \rightarrow r$ (also $r \rightarrow u$ but it is not used in our argument). Hence $r$ has at least one incoming edge, contradiction. We can make a similar contradiction argument to show it does not have any sinks.

(f) The number of subsets of $[0..99]$ that do contain 11 but do not contain 88 is $2^{98}$.

**Solution:**
TRUE. Each of these subsets can either contain or not contain each of the other 98 elements in $[0..99]$ aside from 11 and 88. There are 98 such elements, thus there are 98 steps in the process of constructing a subset, with 2 ways to do each step (include or not include the specified element; thus, by the multiplication rule we have $2^{98}$.

(g) Let $n$ be a positive integer. There are $2^n$ surjective functions with domain $[1..n]$ and codomain $\{0,1\}$.

**Solution:**
FALSE. There are $2^n$ functions with domain $[1..n]$ and codomain $\{0,1\}$. However, not all of them are surjective: the function that maps all to 0 is not surjective and neither is the function that maps all to 1.

(h) 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 $G$.

**Solution:**
FALSE. We provide a counterexample. Let the DAG consist of two isolated vertices, $u,v$. Both $uv$ and $vu$ are topological sorts of this DAG but there is no directed path from either $u$ to $v$ or $v$ to $u$. (BTW, the converse is true: if there is path from $u$ to $v$ then $u$ must appear before $v$ in any topological sort.)

2. (20pts) In how many ways can we distribute $2n$ indistinguishable coins to $n \geq 3$ distinguishable kids such that two of them get at least 2 coins each while the rest get at least 1 coin each.

**Solution:**
We begin by distributing 1 coin each to the $n$ children. This ensures that we meet the last requirement. There is exactly 1 way to do this, since the coins are indistinguishable.

Now we need to distribute the remaining $n$ coins such that at least 2 of the children get at least 1 of them. We can do this by counting complementarily. There are $\binom{n}{n} = \binom{2n-1}{n}$ ways to distribute the coins. However, in some of these distributions, only one child gets at least 1 coin - that is, all of the coins go to one child. There are exactly $n$ such distributions, one for each child. Then, we have that there are $\binom{2n-1}{n} - n$ valid distributions.
3. (25pts) All the events in this problem are in the probability space \((\Omega, P)\).

(a) State the inclusion-exclusion formula for two events, \(P(E \cup F) =?\).

**Solution:**
\[
P(E \cup F) = P(E) + P(F) - P(E \cap F)
\]

(b) Show that part (a) implies the following inequality (known as the union bound for two events): \(P(E \cup F) \leq P(E) + P(F)\).

**Solution:**
Since \(P(E \cap F) \geq 0\) we have
\[
P(E \cup F) = P(E) + P(F) - P(E \cap F) \leq P(E) + P(F).
\]

(c) (The union bound for \(n\) events.) Prove by induction on \(n\) that for any \(n \geq 1\) we have
\[
P(A_1 \cup \cdots \cup A_n) \leq P(A_1) + \cdots + P(A_n)
\]
(Only proofs by induction will receive credit. Hint: in the induction step use the inequality in part (b).)

**Solution:**

**BASE CASE:** \(n = 1\). \(P(A_1) \leq P(A_1)\).

**INDUCTION STEP:** Let \(k \geq 1\) arbitrary. Assume (IH) that \(P(A_1 \cup \cdots \cup A_k) \leq P(A_1) + \cdots + P(A_k)\).

Now consider \(A_1, \ldots, A_k, A_{k+1}\). Using part (b) with \(E = A_1 \cup \cdots \cup A_n\) and \(F = A_{k+1}\) we derive
\[
P(A_1 \cup \cdots \cup A_k \cup A_{k+1}) \leq P(A_1 \cup \cdots \cup A_k) + P(A_{k+1})
\]

Now using the IH we have
\[
P(A_1 \cup \cdots \cup A_k) + P(A_{k+1}) \leq P(A_1) + \cdots + P(A_k) + P(A_{k+1})
\]

Consequently
\[
P(A_1 \cup \cdots \cup A_k \cup A_{k+1}) \leq P(A_1) + \cdots + P(A_k) + P(A_{k+1})
\]

and this finishes the induction step.

4. (25pts) Alice and Bob are playing a game of chance. They roll a fair die. If the die shows an even number then Alice pays Bob $1. If the die shows an odd number then Bob pays Alice $1.

(a) What is the probability that Alice wins $1?

**Solution:**
Let \(O\) be the event that the die shows an odd number and \(\overline{O}\) be the event that the die shows an even number. Further, we define \(A_i\) to be the event that the die shows the number \(i\) for \(i \in \mathbb{Z}, 1 \leq i \leq 6\). Alice wins if the die shows an odd number, so we seek \(\Pr[O] = \Pr[A_1 \cup A_3 \cup A_5] = \Pr[A_1] + \Pr[A_3] + \Pr[A_5] = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2}\). Hence the probability that Alice wins $1 is \(\frac{1}{2}\).

(b) They play the game 2 times in a row, independently. What is the probability that Bob wins $2?

**Solution:**
The probability that the die shows an even number is \(\Pr[\overline{O}] = 1 - \Pr[O] = 1 - \frac{1}{2} = \frac{1}{2}\). For Bob to win $2 the die has to roll an even number in both games. By independence, the probability of this happening is \(\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}\).
(c) Alice and Bob start with 100 dollars each and play the game 100 times in a row, independently. Let $X$ be the random variable representing how much money Alice has after all this. Express $X$ as a sum of 101 random variables.

**Solution:**
Let $X_i$ be random variable that returns the gain/loss for Alice in game $i$, for $i = 1, \ldots, 100$. We have $\text{Val}(X_i) = \{1, -1\}$ and by the above, $P(X_i = 1) = 1/2$ and $P(X_i = -1) = 1/2$.

We also consider the initial amount of money with which Alice started. This is also a random variable (albeit a trivial one!) which returns 100 with probability 1. Then

$$X = 100 + X_1 + \cdots + X_{100}$$

(d) Compute $E(X)$.

**Solution:**
$$E(X_i) = (1)(\frac{1}{2}) + (-1)(\frac{1}{2}) = 0.$$ Now, by linearity of expectation

$$E(X) = E(100) + E(X_1) + \cdots + E(X_{100}) = 100 + 0 + \cdots 0 = 100$$

5. (30pts) Let $X, Y, Z$ be three finite nonempty sets such that $X \cap Y = \emptyset$, $Z \cap Y = \emptyset$, $X \cap Z = \emptyset$ and denote $|X| = m$, $|Y| = n$, $|Z| = p$. Assume that $m < n < p$. Let also $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ be two functions. Consider the undirected graph $G = (V,E)$ where $V = X \cup Y \cup Z$ and

$$E = \{ \{x, f(x)\} \mid x \in X \} \cup \{ \{y, g(y)\} \mid y \in Y \}$$

(a) What is $|V|$ and what is $|E|$ (in terms of $m, n, p$)?

**Solution:**
Since $X, Y, Z$ are disjoint, $|V| = |X| + |Y| + |Z| = m + n + p$.

Because $f, g$ are functions there is exactly one set of the form $\{x, f(x)\}$ for each $x \in X$ and exactly one set of the form $\{y, g(y)\}$ for each $y \in Y$. Therefore $|E| = m + n$.

(b) What is the maximum number of nodes of degree 0 that $G$ can have (in terms of $m, n, p$)?

**Solution:**
Because every node in $X$ and $Y$ is the endpoint of an edge only nodes in $Z$ can have degree 0. They correspond to elements outside of the range of $g$ (i.e., the direct image $g(Y)$). To maximize their number we must minimize the size of $g(Y)$. This happens when $g$ maps all the elements of $Y$ to one element of $Z$. Therefore the answer is $p - 1$.

(c) What is the minimum number of nodes of degree 0 that $G$ can have (in terms of $m, n, p$)?

**Solution:**
Reason like in part (b) then maximize the size of $g(Y)$. This happens when $g$ is injective and the size of $g(Y)$ is $n$. Then there are $p - n$ in $Z$ which are not mapped to; hence, $p - n$.

(d) What is the maximum length that a path in $G$ can have?

**Solution:**
$X$ is nonempty so either $m = 1$ or $m \geq 2$.

CASE 1 $m = 1$: Let $X = \{a\}$. Since $n > m$ we have $n \geq 2$ so let $b_1, b_2 \in Y$ s.t. $b_1 \neq b_2$. Define $f(a) = b_1$ and $g(b_1) = g(b_2)$. Then we have a path of length 3 in $G$: $a \rightarrow \ldots \rightarrow f(a) = b_1 \rightarrow \ldots \rightarrow g(b_1) = g(b_2) \rightarrow \ldots \rightarrow b_2$.

Can there be paths of length 4? Since $X$ has just one element there is exactly one edge between $X$-nodes and $Y$-nodes. So in any path of length 4 there must be 3 edges between $Y$-nodes and $Z$-nodes. At least two of these edges would have to share a $Y$ node, which cannot happen.

So when $m = 1$ the maximum path length is 3.

CASE 2 $m \geq 2$: Let $a_1, a_2 \in X$ s.t. $a_1 \neq a_2$. Since also $n \geq 2$ let $b_1, b_2 \in Y$ s.t. $b_1 \neq b_2$. Define $f(a_1) = b_1$ and $f(a_2) = b_2$ and $g(b_1) = g(b_2)$. Then we have a path of length 4 in $G$: $a_1 \rightarrow \ldots \rightarrow f(a_1) = b_1 \rightarrow \ldots \rightarrow g(b_1) = g(b_2) \rightarrow \ldots \rightarrow b_2 = f(a_2) \rightarrow \ldots \rightarrow a_2$.

Can there be paths of length 5? By the pigeonhole principle In any path with 5 edges there must be 3 edges between $X$-nodes and $Y$-nodes or three 3 edges between $Y$-nodes and $Z$-nodes. In either case, at least 2 of the 3 edges would have to share an $X$-node (in the first case) or a $Y$-node (in the second case). Neither can happen.

So when $m \geq 2$ the maximum path length is 4.

(e) Prove that $G$ is acyclic.

**Solution:**

Consider a hypothetical cycle in this graph. Can it have an $X$-node? An $X$-node cannot have more than 1 edge incident to it so the cycle cannot have any $X$-nodes. Any $Y$-Node that must be part of this cycle may be incident to an $X$-Node, but since $X$-Nodes cannot be a part of the cycle, this is of no help. For $Y$-Nodes that are incident to some $Z$-Node, there is no possible way to form a cycle since any $Z$-Node can only lead to a different $Y$-Node, and since no two $Y$ or $Z$-Nodes share an edge with other nodes from the same set, there is no way to form a cycle. Therefore no such cycles exist.

6. (20pts) Alice receives 20 candies from her granny.

- On day 1, for each candy that she has, Alice is equally likely to keep it or to give it to Bob.
- On day 2, for each candy that he has (all given previously by Alice), Bob is equally likely to eat it or to give it back to Alice. (Bob is weird.)
- On day 3, for each candy that she has (either kept in day 1 or received from Bob in day 2), Alice is equally likely again to eat it or to throw it in the trash. (Less appetizing by now.)

(a) Given a particular candy, $c$, what is the probability that Alice eats $c$?

**Solution:**

We draw a tree-of-possibilities diagram for the candy $c$. 

---

**6**
The probability that Alice eats $c$ is $\frac{1}{4} + \frac{1}{8} = \frac{3}{8}$.

(b) Calculate the expected number of candies that Alice eats.

**Solution:**

For each candy $c = 1, \ldots, 20$ let $X_c$ be the indicator variable of the event “Alice eats $c$”. From part (a), the probability of this event is $\frac{3}{8}$ therefore $E(X_c) = \frac{3}{8}$.

The random variable that gives the number of candies that Alice eats can be expressed as the sum $X_1 + \cdots + X_{20}$. The expected number of candies that Alice eats is $E(X_1 + \cdots + X_{20})$. By linearity of expectation this equals $20 \cdot \frac{3}{8} = \frac{15}{2} = 7.5$.

7. (20pts) In this problem we use directed graphs, a.k.a. digraphs.

(a) Draw an example of a DAG such that it has exactly 2 sources and 3 sinks, there is a path from every source to every sink, and it has the fewest possible number of edges while fulfilling the above requirements.

**Solution:**
(b) Prove that if $G$ is a DAG with at least two distinct sinks such that there is a path from every source to every sink, then $G$ must have at least one node of outdegree $\geq 2$.

**Solution:**

Assume for the sake of contradiction that $G$ has only nodes of outdegree $\leq 1$. We know that $G$ must have at least 1 source $s$. We know there exist paths from $s$ to 2 distinct sinks $t_1$ and $t_2$. We denote the nodes in these paths thus: $s = a_1a_2...a_i = t_1$, $s = b_1b_2...b_j = t_2$. We know that these two paths share the first node $s$. However, since $t_1 \neq t_2$, there exists some integer $k$ such that for all integers $m < k$, $a_m = b_m$, but $a_k \neq b_k$. We know that $a_k \neq b_k$, but we also know that there is an edge from $a_{k-1} = b_{k-1}$ to both $a_k$ and $b_k$. Thus, $a_{k-1} = b_{k-1}$ must have outdegree of at least 2.

Note that this problem is nearly identical to 6 Midterm 3.

8. (15pts) We have $2n$ distinguishable people, $A_1, \ldots, A_n, B_1 \ldots, B_n$, where $n \geq 1$.

(a) In how many ways can these people be arranged in one row?

**Solution:**

This is simply the number of permutations of $2n$. The answer is $(2n)!$

(b) In how many ways can these people be arranged in two rows, the $A$’s in the front row and the $B$’s in the back row, each $B$ behind some $A$, such that for all $i = 1, \ldots, n$, $B_i$ is not behind $A_i$.

**Solution:**

We produce such an arrangement in 2 steps as follows:

1. Arrange the $A$’s in the first row. This can be done in $n!$ ways.
2. Arrange the $B$’s in the second row so that $B_i$ is not behind $A_i$ for all $i = 1, \ldots, n$. This is known as a derangement (see lecture 6T p.5) and the number of ways in which it can be done is

$$n! \sum_{k=0}^{n} \frac{(-1)^k}{k!}$$

By the multiplication rule the answer is

$$(n!)^2 \sum_{k=0}^{n} \frac{(-1)^k}{k!}$$

*Even if you do not remember the formula for the number of derangements* you can derive it as follows.

Assume as above that we are given a permutation of the $A$’s.

Let $P_i$ be set of all permutations of the $B$’s in which $B_i$ is behind $A_i$. The permutations in $P_1 \cup \cdots \cup P_n$ are not allowed and the number of those who are allowed (the derangements) is $n! - |P_1 \cup \cdots \cup P_n|$. Now we apply the Inclusion-Exclusion formula.

For any $J \subseteq [1..n]$ let $|J| = k$ and observe that

$$| \bigcap_{j \in J} P_j | = (n - k)!$$
because we already know which elements go in the positions from \( J \). There are \( \binom{n}{k} \) such sets \( J \). Hence
\[
\sum_{J \subseteq [1..n], |J| = k} |\bigcap_{j \in J} P_j| = \binom{n}{k} (n - k)!
\]

Therefore, by the Principle of Inclusion-Exclusion
\[
|\bigcup_{i=1}^n P_i| = \sum_{k=1}^n (-1)^{k-1} \binom{n}{k} (n - k)!
\]

Hence, the number of derangements is
\[
n! - |P_1 \cup \cdots \cup P_n| = n! - \sum_{k=1}^n (-1)^{k-1} \binom{n}{k} (n - k)!
\]

You can leave it like this for full credit but if you replace
\[
\binom{n}{k} = \frac{n!}{k!(n-k)!}
\]
and factor out \( n! \) you get the formula we used above.

9. (5pts) Bob searches for the solutions to the homework 1. He’s got a filing cabinet with \( n \geq 2 \) drawers and we assume that the solutions are equally likely to be in any of the \( n \) drawers. However, the drawers are so messy that even if the solutions are in a specific drawer, the probability of Bob finding them in that drawer is only \( 1/2 \). Bob searches in drawer \( D_2 \) and does not find them. Compute the probability that the solutions are in drawer \( D_1 \), conditioned on knowing that Bob searched in drawer \( D_2 \) and did not find them.

**Solution:**

Let \( E_i \) be the event that the solutions are in drawer \( i \) for \( i \in \mathbb{Z}, 1 \leq i \leq n \), \( F \) be the event that Bob finds the solutions. We seek:

\[
\Pr[E_1|F] = \frac{\Pr[F|E_1] \Pr[E_1]}{\Pr[F]} \quad \text{(Bayes’ Rule)}
\]

\[
= \frac{\Pr[F|E_1] \Pr[E_1]}{\Pr[F|E_1] \Pr[E_1] + \Pr[F|\overline{E_1}] \Pr[\overline{E_1}]} \quad \text{(Law of Total Probability)}
\]

Note that \( \Pr[F|E_1] = 1 \), since Bob will never find the solutions if he searches \( D_2 \). Additionally, \( \Pr[E_1] = \frac{1}{n} \), since the solutions are equally likely to be in any drawer. This gives us \( \Pr[\overline{E_1}] = \frac{n-1}{n} \).

We now solve for \( \Pr[F|\overline{E_1}] \). Note that we can split this into two cases - when the solutions are in
and when they are not. More formally:

$$\Pr[F \mid E_1] = \frac{\Pr[F \cap E_1]}{\Pr[E_1]} = \frac{\Pr[F \cap E_1 \cap E_2] + \Pr[F \cap E_1 \cap E_2]}{\Pr[E_1]} = \frac{\Pr[F \cap E_2] + \Pr[F \cap E_1 \cap E_2]}{\Pr[E_1]} = \frac{\Pr[F \mid E_2] \Pr[E_2] + \Pr[F \mid E_1 \cap E_2] \Pr[E_1 \cap E_2]}{\Pr[E_1]}$$

Note that $\Pr[F \mid E_1 \cap E_2] = 1$, since Bob can’t find the solutions if they’re not in $D_2$. Additionally, $\Pr[E_1 \cap E_2] = \frac{n-2}{n}$, since this is the event where the solutions are not in $D_1$ or $D_2$.

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

Finally, plugging everything back in to the original equation, we get:

$$\Pr[E_1 \mid F] = \frac{\Pr[F \mid E_1] \Pr[E_1]}{\Pr[F \mid E_1] \Pr[E_1] + \Pr[F \mid E_1 \cap E_2] \Pr[E_1 \cap E_2]} = \frac{1 \times \frac{1}{n}}{1 \times \frac{1}{n} + \frac{2n-3}{2n-2} \times \frac{n-1}{n}} = \frac{2}{2n + 2n-3} = \frac{2}{2n - 1}$$

### 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) Let $A, B$ be two finite nonempty sets. Then $A \cup B$ has exactly the same number of elements as $(A \setminus B) \times (B \setminus A)$, true or false?

   **Solution:** FALSE. Counterexample: $A = \emptyset$ then $A \setminus B = \emptyset$ then $(A \setminus B) \times (B \setminus A) = \emptyset$. Then take some nonempty $B$; $A \cup B = B$, so $|A \cup B| > 0$.

   (b) If $|A| = n$ then there are $n!$ injective functions with domain $A$ and codomain $A$, true or false?

   **Solution:** TRUE. The injective functions are in one-to-one correspondence with the permutations of the elements of $A$. 

(c) Let $P(n)$ be a predicate defined on natural numbers. The negation of the statement $\forall k P(k) \Rightarrow P(k + 1)$ is the statement $\exists k (\neg P(k)) \land (\neg P(k + 1))$, true or false?

Solution:
FALSE. The negation is $\exists k P(k) \land (\neg P(k + 1))$.

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

Solution:
FALSE. Counterexample: $(\{a, b, c\}, \{a \rightarrow b, b \rightarrow c, c \rightarrow b\})$. $a$ is not reachable from $b$ (or $c$).

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 $v \rightarrow w$ whenever we have the edge $v \rightarrow w$ or the edge $w \rightarrow v$ in $G$.

(e) If $G$ is strongly connected then $G^u$ is connected, true or false?

Solution:
TRUE. Let $v, w$ be two vertices in $G^u$ (hence in $G$). 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 $v \rightarrow \cdots \rightarrow w$ in $G^u$.

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

Solution:
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.

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

Solution:
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.

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

Solution:
TRUE. Let $\text{Pr}[X = 1] = p$. $E[X] = (-1) \cdot \text{Pr}[X = -1] + (1) \cdot \text{Pr}[X = 1] = -(1 - p) + p = 2p - 1$. Setting $E[X] = 0$ we find $p = 1/2$.

2. (15pts) Consider the recurrence relation

$$a_{n+1} = a_n + 3 \quad (n \geq 0) \quad \text{and} \quad a_0 = 4$$

Prove that $\forall n \geq 4 \ a_n \leq 2^n$.

Solution:
BASE CASE $n = 4$. We compute successively $a_1 = 4 + 3 = 7$, $a_2 = 7 + 3 = 10$, $a_3 = 10 + 3 = 13$, $a_4 = 13 + 3 = 16$. Now $a_4 = 16 \leq 2^4$. Check.
INDUCTION STEP  Let \( k \geq 4 \) arbitrary. Assume (IH) that \( a_k \leq 2^k \). We want to show that \( a_{k+1} \leq 2^{k+1} \).

Since \( a_{k+1} = a_k + 3 \) and \( 2^{k+1} = 2 \cdot 2^k = 2^k + 2^k \) the desired inequality would follow if we had \( a_k \leq 2^k \) and \( 3 \leq 2^k \).

Indeed, the first inequality is the IH and the second one is true for \( k \leq 4 \) because \( 2^k \) is an increasing function and already \( 3 \leq 2^4 \).

Note that I did not write that \( a_k + 3 \leq 2^{k+1} = 2^k + 2^k \) is the same as \( a_k \leq 2^k \) and \( 3 \leq 2^k \) (it is not) but rather that it would follow from \( a_k \leq 2^k \) and \( 3 \leq 2^k \).

3. In this problem the graphs are undirected.

(a) Draw a connected graph with 4 nodes and exactly 1 cut edge.

**Solution:**

\[ G = (\{a, b, c, d\}, \{a-b, b-c, c-a, c-d\}) \]

The cut edge is \( c-d \).

(b) Let \( G \) be a graph (not necessarily connected). Prove that if an edge in \( G \) is not a cut edge then it is part of some cycle in \( G \).

**Solution:**

Let \( u-v \) be an edge in \( G \). Since there is a path between \( u \) and \( v \) it must be the case that \( u \) and \( v \) are in the same CC of \( G \).

Now, assume \( u-v \) is not a cut edge. Delete \( u-v \) resulting in a graph \( G' \) with the same vertices and one less edge. Since \( u-v \) is not a cut edge, the number of CCs of \( G' \) is the same as that of \( G \). Since we are not adding any edges or deleting any other edges, this means that, as blocks of vertices, the CCs of \( G' \) are the same as the CCs of \( G \).

Therefore \( u \) and \( v \) are also in the same CC of \( G' \). Hence there is a path \( p \) from \( u \) to \( v \) in \( G' \).

This path does not contain the edge that we deleted hence it forms a cycle in \( G \) together with that edge.

4. Consider the following set 

\[ W = \{(x, y) \mid x \in [1..100] \land y \in [1..50] \land x = 3y\} \]

(a) Compute \( |W| \).

**Solution:**

We count the distinct ordered pairs \((x, y) \in W\). For all such \((x, y)\), \( x \) must be between 1 and 100, it must be divisible by 3 and \( y = \frac{x}{3} \) must be at most 50. Therefore \( y = \frac{34}{3} \) can be any of the numbers in \([1..33]\) because \( 33 \cdot 3 = 99 < 100 \), but it cannot be any other number as \( 34 \cdot 3 = 102 > 100 \).

It follows that \( |W| = 33 \).
(b) How many bijective functions \( f : W \to W \) are there?

**Solution:**

You can create a bijection in 33 steps: in step \( i \) for \( i \in \mathbb{Z}, 1 \leq i \leq 33 \), we choose \( f((3i,i)) \) from the remaining elements of \( W \) which have not been chosen in previous steps, of which there are \( 33 - i \). Thus, we have \( 33 \cdot 32 \cdot \ldots \cdot 1 = 33! \).

We can also see that these bijections are in one-to-one correspondence with the permutations of \( W \) - interpret each permutation as a bijection by considering the \( i^{th} \) element of the permutation to be \( f((3i,i)) \) - so the answer is \( 33! \).

(c) How many injective functions \( f : W \to W \) are there?

**Solution:**

For any injective function the size of its range is the same as that of its domain. But the range is also a subset of \( W \) so it cannot be a proper subset. It follows that \( f : W \to W \) is also surjective hence also bijective.

Thus counting the injections \( W \to W \) is the same as counting the bijections \( W \to W \) so the answer is the same \( 33! \).

5. (20pts) For each of the questions below give the answer and an explanation of how you derived it.

(a) Let \( n \) be an even natural number. Count the number of distinct sequences of bits of length \( n \) in which the number of 0’s exactly equals the number of 1’s.

**Solution:**

Let \( n = 2m \). Exactly half are 1’s. Their position can be chosen in \( \binom{2m}{m} \) ways. The rest are 0’s.

The answer is \( \binom{2m}{m} \).

(b) Let \( n \geq 2 \) be a natural number. Count the number of distinct sequences of bits of length \( n \) with exactly two 1’s.

**Solution:**

The positions of the two 1’s can be chosen in \( \binom{n}{2} \) ways. The other positions are 0. The answer is \( \binom{n}{2} \).

(c) Let \( n \geq 5 \) be a natural number. Count the number of distinct sequences of bits of length \( n \) with

- at least two 1’s and
- at least three 0’s.

**Solution:**

We count instead the number of distinct sequences that do not satisfy the condition “at least two 1’s and at least three 0’s” and then we subtract this from the total number of sequences which is \( 2^n \).

For the strings that do not satisfy the condition we have five cases:

- strings with no 1’s; count is 1
- strings with exactly one 1; count is \( n \)
- strings with no 0’s; count is 1
- strings with exactly one 0; count is $n$
- strings with exactly two 0’s; count is $\binom{n}{2}$

Because $n \geq 5$ the cases do not “overlap”, that is, they describe pairwise disjoint sets of strings. Therefore the sum rule applies.

The answer is $2^n - \binom{n}{2} - 2n - 2$.

6. (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 $A \sim B$.

**Solution:**

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

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

**Solution:**

All the subsets in an equivalence class have the same min therefore the classes correspond to the elements of $[1..n]$. There $n$ such hence $n$ equivalence classes.

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

**Solution:**

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

**Solution:**

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

**Solution:**

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 $2^{n-k}$.

7. 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.

Solution:
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) Prove that a tree with at least three leaves must have at least one node of degree $\geq 3$.

Solution:
Let $T$ be a tree with at least three leaves and let these leaves be $a, b, c$. Since $T$ is a tree there is a path $p_1 \equiv a \cdots b$ and also a path $p_2 \equiv a \cdots c$.
Consider the subgraph $S$ consisting of the path $p_1$ together with the path $p_2$.
The following claim will end the proof.

CLAIM: There exists a vertex in $S$ that has degree $\geq 3$ (in $S$, therefore also in $T$)

PROOF OF CLAIM: $S$ is connected and acyclic therefore, like $T$, it is a tree.
Suppose toward a contradiction that all vertices in $S$ have degree $\leq 2$. Then the tree $S$ is a path.
But $a, b, c$ are distinct and are in this path. By PHP one of them cannot be an endpoint in this path. Hence it has degree 2, contradicting the fact that it is a leaf.

8. (15pts) Give a combinatorial proof of the identity

$$\binom{100}{70} = \binom{99}{30} + \binom{99}{29}$$

Specifically, you have to identify a counting problem such that counting one way you get $\binom{100}{70}$ and counting another way you get $\binom{99}{30} + \binom{99}{29}$. You are not allowed to rearrange the sides of the identity and you are not allowed to use any other identities about binomial coefficients, or to express them in terms of factorials.

Solution:
A course involves 100 people: 99 distinguishable students and 1 instructor. Counting problem: what is $W$, the number of different ways in which we can put 70 of these people in classroom A and the remaining 30 in classroom B.
The LHS just counts \( W \) by counting the number of ways of choosing the 70 out of 100 that go in classroom A (the instructor may of may not go in A). That’s \( \binom{100}{70} \) ways.

The RHS also counts \( W \), by considering two exhaustive and disjoint cases (so we add the numbers by the sum rule): in the first case the instructor goes in classroom A so we only have to choose the 30 students that go in classroom B, and this is done in \( \binom{99}{30} \) ways; in the second case the instructor goes to classroom B so we only have to choose the 29 students that go with the instructor to classroom B, and that is done in \( \binom{99}{29} \) ways.

9. (15pts) 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)

Solution:

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

(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.

Solution:

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.

10. (30pts) Here is a list of facts about probability \((X,Y)\) are arbitrary events in some probability space \((\Omega, P)\) and \(X = \Omega \setminus X\):

\[(F1) \quad P(\emptyset) = 0 \text{ and } P(\Omega) = 1\]

\[(F2) \quad X \cap Y = \emptyset \text{ implies } P(X \cup Y) = P(X) + P(Y)\]

\[(F3) \quad P(X) = 1 - P(X)\]

\[(F4) \quad X \subseteq Y \text{ implies } P(X) \leq P(Y)\]

\[(F5) \quad P(X \cup Y) = P(X) + P(Y) - P(X \cap Y)\]

\[(F6) \quad P(X) = P(X \mid Y) \cdot P(Y) + P(X \mid Y^c) \cdot P(Y^c) \quad \text{ (law of total probability)}\]

\[(F7) \quad P(X \mid Y) = P(Y \mid X) \cdot P(X) / P(Y) \quad \text{ (Bayes' law)}\]

(a) Complete (replace the question mark with the correct formula) the definition of conditional probability below:

\[
P(X \mid Y) = ?
\]
Solution:
For \( Y \) such that \( P(Y) \neq 0 \):
\[
P(X \mid Y) = \frac{P(X \cap Y)}{P(Y)}
\]
(b) Complete (replace the question mark with one of the correct properties) the definition of event independence below:
\[
X \perp Y \hspace{1em} \text{iff} \hspace{1em} ?
\]
Solution:
\( X \perp Y \) iff \( P(X \cap Y) = P(X)P(Y) \) iff \( P(Y) \neq 0 \Rightarrow P(X \mid Y) = P(X) \).
(c) Now let \( A, B \subseteq \Omega \) be two events in a probability space \((\Omega, P)\) such that \( 0 < P(A) < 1 \) and \( 0 < P(B) < 1 \), i.e., neither \( P(A) \) nor \( P(B) \) can be 0 or 1. As usual, let \( \overline{A} = \Omega \setminus A \).
Denote \( p = P(A) \), as well as \( q = P(B) \) and \( r = P(A \cap B) \).
Using only the facts about probability in the list (F1)-(F7) above (you must indicate when you use them), compute \( P(\overline{A} \mid B) \) in terms of \( p, q, r \).
(You can also use set algebra identities that seem obvious to you, however you must state and circle these identities and write “we assume” next to them.)
Solution:
By (F7) \( P(\overline{A} \mid B) = P(B \mid \overline{A})P(\overline{A})/P(B) \). Therefore by (F3) \( P(\overline{A} \mid B) = P(B \mid \overline{A})(1 - P(A))/P(B) \). Plugging in the notation \( P(\overline{A} \mid B) = P(B \mid \overline{A})(1 - p)/q \). It remains to compute \( P(B \mid \overline{A}) \).
By (F6) \( P(B) = P(B \mid A)P(A) + P(B \mid \overline{A})P(\overline{A}) \) therefore by def of cond prob and by (F3) \( P(A) = P(B \cap A) + P(B \mid \overline{A})(1 - P(A)) \). Plugging in the notation \( q = r + P(B \mid \overline{A})(1 - p) \).
Therefore \( P(B \mid \overline{A}) = (q-r)/(1-p) \).
Replacing, we get
\[
P(\overline{A} \mid B) = P(B \mid \overline{A})(1-p)/q = [(q-r)/(1-p)] \cdot [(1-p)/q] = (q-r)/q
\]
Alternative Solution By the definition of conditional probability \( P(\overline{A} \mid B) = P(\overline{A} \cap B)/P(B) \). We assume that \( \overline{A} \cap B = B \setminus (A \cap B) \). Therefore \( P(\overline{A} \mid B) = P(B \setminus (A \cap B))/P(B) \).
There is no fact about the probability of event difference in the list but we can derive one:
CLAIM \( X \subseteq Y \) implies \( P(Y \setminus X) = P(Y) - P(X) \)
PROOF OF CLAIM We assume that \( (Y \setminus X) \cap X = \emptyset \) and that \( X \subseteq Y \) implies \( (Y \setminus X) \cup X = Y \). Therefore we can apply (F2) and we get \( P(Y) = P(Y \setminus X) + P(X) \). Hence \( P(Y \setminus X) = P(Y) - P(X) \).
Now, since \( A \cap B \subseteq B \) (we assume), we can use the claim, and we get \( P(\overline{A} \mid B) = P(B \setminus (A \cap B))/P(B) = (P(B) - P(A \cap B))/P(B) = (q-r)/q \), the same answer.
(d) With \( A, B \) as in part (10c), prove that \( A \perp B \) implies \( \overline{A} \perp B \).
(Again you can use only the facts about probability in the list (F1)-(F7)— you must indicate when you use them— and you can also use set algebra identities that seem obvious to you, however you must state and circle these identities and write “we assume” next to them.)
**Solution:**

We assume $A \perp B$ and we want to show $\overline{A} \perp B$.

By definition, $A \perp B$ means $P(A \cap B) = P(A)P(B)$, i.e., $r = pq$.

From part (c) we have $P(\overline{A} \mid B) = (q - r)/q$. Replacing $r = pq$ we get (by F3) $P(\overline{A} \mid B) = (q - pq)/q = 1 - p = P(\overline{A})$. Since $P(B) \neq 0$ this implies, by the alternative definition of independence, that $\overline{A} \perp B$.

**Alternative Solution** We show instead that $P(\overline{A} \cap B) = P(\overline{A})P(B)$.

We assume that $\overline{A} \cap B = B \setminus (A \cap B)$. Therefore $P(\overline{A} \cap B) = P(B \setminus (A \cap B))$. Using the claim that we proved in the alternative solution of part (c) we get $P(\overline{A} \cap B) = P(B) - P(A \cap B) = q - r$. Since $r = pq$ we further have $P(\overline{A} \cap B) = q - r = q - qp = (1 - p)q$. By (F3), $P(\overline{A} \cap B) = P(\overline{A})P(B)$.

Done

11. 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 $t$.

**Solution:**

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.

![DAG with 3n+1 vertices](image)

12. Let $n \geq 2$ be a natural number. Draw (use dot dot dot) an example of a DAG with exactly $3n - 1$ vertices such that it has exactly $2^n$ distinct topological sorts.

**Solution:**

For each $i \in [2..n]$, the graph contains the 3 nodes $i$, $a_i$ and $b_i$, and it also contains nodes $a_1$ and $b_1$. These are the only nodes in the graph, so there are $3(n - 1) + 2 = 3n - 1$ nodes. In any topological sort, there are exactly $2^n$ distinct topological sorts.
sort of this graph, we must have that the nodes 2, 3, ...n appear in sequential order, as there are paths from 2 to 3, from 3 to 4, etc. Further, nodes \( a_i \) and \( b_i \) must come before node \( i + 1 \) for integers \( i \) s.t. \( 1 \leq i \leq n - 1 \) and after node \( i \) for integers \( i \) s.t. \( 2 \leq i \leq n \). Then we can create a topological sort in the following way:

1. **Step 1**: Choose either \( a_1 \) or \( b_1 \) to come first. (2 ways)
2. **Step 2**: Choose the other to come second. (1 way)
3. **Step 3**: Choose node 2 to come next. (1 way)
4. **Step 4**: Choose either \( a_2 \) or \( b_2 \) to come next. (2 ways)
   :
5. **Step 3n-3**: Choose node \( n \) to come next. (1 way)
6. **Step 3n-2**: Choose either \( a_n \) or \( b_n \) to come next. (2 ways)
7. **Step 3n-1**: Choose the other to come last. (1 way)

\( n \) of these steps can be done in 2 ways; the rest can be done in 1 way. By the multiplication rule, there are \( 2^n \) different topological sorts.

13. (10pts) Let \( G = (V, E) \) be a connected graph with at least two distinct spanning trees.

(a) Prove that \( |E| \geq |V| \).

**Solution:**
Since \( G \) is connected we have \( |E| \geq |V| - 1 \).
We argue by contradiction that \( |E| \neq |V| - 1 \). Indeed, if \( |E| = |V| - 1 \) then \( G \) is a tree and therefore it has exactly one spanning tree, itself. Contradiction.

(b) Prove that the graph has at least three distinct spanning trees.

**Solution:**
Consider \( T \), one of the spanning trees of \( G \). \( T \) has exactly \( n - 1 \) edges, and since \( G \) has at least \( n \) edges (shown in part (a)), we know that there exists an edge \( e = (u, v) \) in \( G \) that is not in \( T \). Thus, consider the graph \( G' \) created from adding the edge \( e \) into \( T \), and so \( G' \) has \( n \) edges. Because \( T \) was a tree, there was a unique path from \( u \) to \( v \) before adding in \( e \), and so the addition of \( e \) creates exactly one cycle, \( C \). Because \( G' \) is undirected, we know \( C \) has at least 3 edges. Now, consider the graphs \( T_1, T_2, \) and \( T_3 \) created by removing any one of 3 distinct edges within \( C \), and so each of these graphs is distinct. Because we removed an edge belonging to a cycle, these graphs remain connected. Furthermore, since \( G' \) had \( n \) edges, each of \( T_1, T_2, \) and \( T_3 \) has \( n - 1 \) edges. Thus, since each of these 3 graphs is connected and contains \( n - 1 \) edges, they form 3 distinct spanning trees.

**Alternative Solution 1**
Let \( T_1 \) and \( T_2 \) be the two distinct spanning trees. Since they are distinct there must exist some \( T_1 \)-edge that is not a \( T_2 \)-edge. Call it \( u-v \).

Since \( T_2 \) is a spanning tree there must exist a path \( P \equiv u \cdots v \) in \( T_2 \). At least one edge in \( P \) cannot be a \( T_1 \)-edge. Indeed, if all of them were, then together with \( u-v \) they would form a cycle in \( T_1 \) which is impossible. Call this \( T_2 \)-edge in the path \( P \) that is not a \( T_1 \)-edge \( x-y \).

To the path \( P \) we add the edge \( v-u \) to form a cycle (in \( G \), call this cycle \( C \). It contains the two distinct edges \( u-v \) and \( x-y \). But cycles have at least 3 edges so there must be a third edge
in $C$ distinct from both $u-v$ and $x-y$. Call it $s-t$. (By the way saying that these three edges are distinct does imply that their endpoints are distinct. Some of these edges may be adjacent.)

To recap, in the cycle $C$ we have $v-u$ which is a $T_1$-edge but not a $T_2$-edge, $x-y$ which is a $T_2$-edge but not a $T_1$-edge, and $s-t$ which is a $T_2$-edge (and may or may not be a $T_1$ edge).

Now let $T_3 = T_2 \cup \{u-v\} \setminus \{s-t\}$ (these are the edges; the vertices are $V$). $T_3$ differs from $T_1$ because it still has $x-y$. $T_3$ also differs from $T_2$ because it does not have $s-t$. So if we show $T_3$ is a spanning tree we have our third distinct spanning tree and we are done.

$T_3$ has $|V| - 1$ edges because although we deleted an edge from $T_2$ we added another one. $T_3$ is connected because $T_2$ is connected and any path in $T_2$ that used the edge we deleted, $s-t$ can be “repaired” to a path in $T_3$ using the rest of the cycle $C$. Therefore $T_3$ is a tree, and a spanning because it has all the vertices.

**Alternative Solution 2**

Let $T_1$ and $T_2$ be the two distinct spanning trees. Since they are distinct and they have the same number of edges ($|V| - 1$) there must exist some $T_1$-edge that is not a $T_2$-edge, call it $u-v$, and a $T_2$-edge that is not a $T_1$-edge, call it $x-y$.

Since $T_2$ is a spanning tree, it must have a path $P = u \cdots x$ and a path $Q = v \cdots y$. We claim that $P$ and $Q$ are node-disjoint. Indeed, if $P$ and $Q$ have a node in common then that would give a path in $T_2$ between $x$ and $y$ that is not the same as the edge $x-y$. Both paths in a tree are unique, contradiction.

By concatenating $u-v$, $Q$, $y-x$, and the reverse of $P$ we get a cycle, $C$.

From here on we proceed as in the alternative solution 1.

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

(a) If $f : A \to B$, where $A$ is nonempty and $|B| = 1$, then $f$ is surjective, true or false?

**Solution:**

TRUE. Let $y_0$ be the one element of $B$. Since $A$ is nonempty there exists some $x_0 \in A$. Then $f(x_0)$ must equal $y_0$ because there is no other choice. Thus $f$ is surjective.

(b) Suppose that $A, B, C$ are each finite nonempty sets, each with an even number of elements, and that $A$ and $B$ are disjoint. Then $|(A \cup B) \times C|$ is divisible by 4, true or false?

**Solution:**

TRUE. Since $A$ and $B$ are disjoint, $|A \cup B| = |A| + |B|$ is even, and so $|(A \cup B) \times C| = (|A| + |B|)|C|$ is divisible by 4.

(c) Suppose that $A, B$ are finite nonempty sets such that $A \cup B$ has 4 elements. Then, $|A \setminus B| = |B \setminus A| = |A \cap B| = 1$, true or false?

**Solution:**

FALSE. For a counterexample consider $A = \{0, 1\}$ and $B = \{2, 3\}$.

(d) For any function $f : \mathbb{N} \to \mathbb{N}$ the following holds: $\exists m \forall n \ f(n) \leq m$, true or false?

**Solution:**

FALSE. As a counterexample, consider the function $f(n) = n$. Then for every $m$ there is some $n$ (namely $n = m + 1$) such that $f(n) = f(m + 1) = m + 1 \leq m$. 

20
(e) There exists a tree with exactly 3 leaves, in which the length of a longest path is 1000, true or false?

Solution:
TRUE. Consider the graph which consists of the distinct vertices $a, b, v_1, v_2, \ldots, v_{999}, v_{1000}$ and the edges $a-v_1, b-v_1, v_1-v_2, \ldots, v_{999}-v_{1000}$. This a tree with leaves $a, b, v_{1000}$, and there are two longest paths, both of length 1000.

(f) There exist DAGs in which all vertices have outdegree 2, true or false?

Solution:
FALSE. We proved in class that in any DAG there exists at least one node of outdegree 0.

(g) In any DAG the reachability relationship, $\rightarrow$, has the property $x \rightarrow y \land y \rightarrow x \Rightarrow x = y$ (this is called “antisymmetry”), true or false?

Solution:
TRUE. Let $u, v$ be vertices such that $u \rightarrow v$ and $v \rightarrow u$. To show $u = v$ we argue by contradiction. If $u \neq v$ then there exists a walk of length non-zero from $u$ to $u$. Then there exists a cycle. Contradiction.

15. Let $(\Omega, P)$ where $P : \Omega \rightarrow [0, 1]$ be a probability space, $A, B \subseteq \Omega$ be two events, and $\overline{A} = \Omega \setminus A$ and $\overline{B} = \Omega \setminus B$ be their complements. Here is a list of facts about probability:

(F1) $P(\emptyset) = 0$
(F2) $P(X \cup Y) = P(X) + P(Y) - P(X \cap Y)$
(F2') $P(X \cup Y) = P(X) + P(Y)$ when $X \cap Y = \emptyset$
(F3) $P(\overline{X}) = 1 - P(X)$

Using only the facts about probability in the list above prove that if $A \perp B$ then $\overline{A} \perp \overline{B}$.

(You can also use set algebra identities that seem obvious to you, however you must state and circle these identities and write “we assume” next to them.)

Solution:
We know that $P(A \cap B) = P(A) \cdot P(B)$ and we want to show that $P(\overline{A} \cap \overline{B}) = P(\overline{A}) \cdot P(\overline{B})$.

By De Morgan, $\overline{A} \cap \overline{B} = \overline{A \cup B}$.

Then, by (F3) and (F2), $P(\overline{A} \cap \overline{B}) = 1 - P(A \cup B) = 1 - (P(A) + P(B) - P(A \cap B))$.

Further, using (F3) again, $1 - (P(A) + P(B) - P(A \cap B)) = (1 - P(A)) \cdot (1 - P(B)) = P(\overline{A}) \cdot P(\overline{B})$.

Done.

16. Consider a connected graph $G = (V, E)$ such that $|E| = |V|$. Prove that $G$ contains exactly one cycle.

Solution:
First we argue by contradiction that the graph must contain at least one cycle. Suppose not. Then the graph is connected and acyclic hence a tree. Therefore $|E| = |V| - 1$, contradiction.

Now we argue, again by contradiction, that it cannot contain two distinct cycles. Suppose it contains distinct cycles $C_1$ and $C_2$. Since the cycles are distinct there must exist some edge that is in one of the cycles but not in the other. WLOG suppose $u-v \in C_1$ but $u-v \notin C_2$. Now remove the edge $u-v$
producing a graph $G = (V, E)$, where $E = E \setminus \{u-v\}$. Note also that $C_2$ remains a cycle in $G$. We argue that $G$ is still connected. Indeed, $u$ and $v$ remain connected in $G$, by the path $P$ that remains from the cycle $C_1$ after removal of the edge $u-v$. Therefore any walk in $G$ that used the edge $u-v$ can be “repaired” to produce a walk in $G$ by replacing $u-v$ with $P$. It follows that $G$ is also connected.

But in $G$ we have $|E| = |E| - 1 = |V| - 1$ and since $G$ is connected it must be a tree, thus acyclic. This contradicts the existence of the cycle $C_2$ in $G$.

17. Consider a set $A$ with $n \geq 1$ elements. We color independently each of the elements of $A$ red with probability $1/3$ and blue with probability $2/3$. This determines a partition on $A$: in one block (the "blue block") we have all the blue elements and in the other block (the "red block") we have all the red elements. Let $\rho$ be the equivalence relation that corresponds to this partition. Calculate the expected value of $|\rho|$. No proof necessary but explain your calculation.

Solution:
The elements of $\rho$ are ordered pairs $(x, y)$ where $x, y \in A$.

First we examine the case $x \neq y$. For the pair $(x, y)$ to be in $\rho$ both $x$ and $y$ must be colored with the same color. Using the independence, the probability that they are both red is $(1/3)(1/3) = 1/9$. Similarly, the probability that they are both blue is $(2/3)(2/3) = 4/9$. Since these are disjoint, the probability that they are colored with the same color is $(1/9) + (4/9) = 5/9$.

Next we examine the case $x = y$. In this case $(x, y) \in \rho$ by reflexivity. Now define for each $(x, y) \in A \times A$ an indicator random variable $X_{x,y}$ that is 1 when $(x, y) \in \rho$ and 0 otherwise. We have $E(X_{x,y}) = P((x, y) \in \rho)$ which equals $5/9$ when $x \neq y$ and 1 when $x = y$. By linearity of expectation, where we let the sum range over all $(x, y) \in A \times A$:

$$E[|\rho|] = E\left[ \sum_{x,y} X_{x,y} \right] = \sum_{x,y} E[X_{x,y}] = \sum_{x \neq y} E[X_{x,y}] + \sum_x E[X_{x,x}] = \frac{5}{9} \cdot n(n-1) + 1 \cdot n = \frac{n(5n+4)}{9}$$

18. Let $X = \{1, 2, \ldots, 2n\}$ where $n \geq 2$. How many nonempty subsets of $X$ contain at most 2 odd numbers?

Solution:
We have three disjoint cases for such subsets. We will then, by the sum rule, add up the numbers.

CASE 1: nonempty subsets containing zero odd numbers. There are $n$ even numbers in $X$ and therefore $2^n$ subsets of $X$ consisting of just even numbers, $2^n - 1$ of which are nonempty.

CASE 2: nonempty subsets containing exactly one odd number. First pick an odd number $p$. This can be done in $n$ ways. For each $p$ pick a subset (maybe empty) $S$ of the even numbers, producing $\{p\} \cup S$ and it can be done in $n \cdot 2^n$ ways.
CASE 3: nonempty subsets containing exactly two distinct odd numbers. First pick two distinct odd numbers $p_1, p_2$. This can be done in $\binom{n}{2}$ ways. For each $\{p_1, p_2\}$ pick a subset (maybe empty) $S$ of the even numbers, which can be done in $2^n$ ways. This produces $\{p_1, p_2\} \cup S$ and it can be done in $\binom{n}{2} \cdot 2^n$ ways.

The answer to the question is

$$2^n - 1 + n \cdot 2^n + \binom{n}{2} 2^n = (n^2 + n + 2)2^{n-1} - 1$$

19. Prove that in any DAG we have $2 \cdot |E| \leq |V|(|V| - 1)$.

Solution:

We prove that for any $n \geq 1$, for any DAG $G = (V, E)$ such that $|V| = n$ we have $2 |E| \leq |V|(|V| - 1)$.

BASE CASE: $n = 1$. Let $u$ be the unique vertex. Any directed edge would have to be a self-loop on $u$ but we cannot have these in a DAG. Hence $|E| = 0$ and the inequality becomes $0 \leq 0$. Check.

INDUCTION STEP: Let $k \geq 1$. Assume (IH) that for any DAG with $k$ vertices the inequality holds. Now consider $G = (V, E)$ an arbitrary DAG with $k + 1$ vertices and we want to show that $2 |E| \leq (k + 1)(k + 1 - 1) = k(k + 1)$.

We have shown in class that $G$ must have at least one vertex of indegree 0 (a source). Let $v$ be that vertex and let $G' = (V', E')$ be the graph obtained by deleting $v$ and all the edges that start in $v$ from $G$. Since $G'$ has $k$ vertices the IH applies to it. Therefore we have $2 |E'| \leq k(k - 1)$.

The maximum number of edges that we could have deleted when we went from $G$ to $G'$ is $k$. This is also the maximum outdegree of any vertex in $G$. Indeed, because the indegree of $v$ is 0, there could have only been edges from $v$ to any vertex other than $v$. Therefore $|E| \leq |E'| + k$. Now

$$2|E| \leq 2(|E'| + k) = 2|E'| + 2k \leq k(k - 1) + 2k = k^2 + k = (k + 1)k$$

and we are done.

20. In a bizarre parallel universe they are able to manufacture fair $n$-dice for various $n \geq 1$. Such a die has $n$ faces showing the numbers $\{1, 2, \ldots, n\}$. Because the dice are fair, when you roll them, each of these $n$ numbers shows up with equal probability, $1/n$. A citizen of this universe, Bizz, plays the following game: it rolls independently two $n$-dice, a red one, and we denote by Red the number the red die shows, and a blue one and we denote by Blue the number the blue die shows. If Red $<$ Blue then Bizz wins, otherwise it loses. For the following questions give the answer and a short explanation of how you obtained it. No proofs required. (Hint: first think that $n = 6$ so these are normal dice, then generalize.)

(a) How many outcomes does the prob. space associated with Bizz’s game have? What is the probability of each outcome?

Solution:

The outcomes of the probability space correspond to the results of the two rolls, not the result of the game. Each die has $n$ possible outcomes, so by the multiplication rule, there are $n \cdot n = n^2$
possible outcomes. For any outcome \((\text{Red, Blue})\), the probability that the red die rolls \text{Red} is \(\frac{1}{n}\), and the probability that the blue die rolls \text{Blue} is \(\frac{1}{n}\); since the two rolls are independent, the probability of this (any) outcome is \(\frac{1}{n} \cdot \frac{1}{n} = \frac{1}{n^2}\).

(b) Compute the probability that Bizz wins. (Hint: for \(n = 6\) this probability should be \(\frac{5}{12}\) so you can check what you got.)

Solution:

Solution One  
Note that any choice of two distinct (unordered) integers from 1 to \(n\) corresponds to exactly one winning outcome, in particular, when \text{Red} is the smaller integer and \text{Blue} is the larger. Similarly, every winning outcome is a pair \((\text{Red, Blue})\) where \text{Red} < \text{Blue} and thus corresponds to exactly one choice of two distinct (unordered) integers from 1 to \(n\). Thus, there is a bijection, so exactly \(\binom{n}{2}\) outcomes correspond to wins. Each has probability \(\frac{1}{n^2}\), so the probability that Bizz wins is

\[
\binom{n}{2} \cdot \frac{1}{n^2} = \frac{1}{2} - \frac{1}{2n}.
\]

Solution Two  
Let us partition the sample space into three disjoint events: \text{Red} < \text{Blue}, \text{Red} = \text{Blue}, and \text{Red} > \text{Blue}. We can find \(P(\text{Red} = \text{Blue})\) by noticing that once the red die is rolled, no matter what its result is, the probability that the blue die rolls to be the same is \(\frac{1}{n}\). Thus

\[P(\text{Red} = \text{Blue}) = 1 \cdot \frac{1}{n} = \frac{1}{n}.\]

Since there is no difference between the red and blue dice, we must have

\[P(\text{Red} < \text{Blue}) = P(\text{Red} > \text{Blue}).\]

We know that the union of these three disjoint events is the sample space, so

\[
1 = P(\text{Red} < \text{Blue}) + P(\text{Red} = \text{Blue}) + P(\text{Red} > \text{Blue})
\]

\[= P(\text{Red} < \text{Blue}) + \frac{1}{n} + P(\text{Red} < \text{Blue})
\]

\[= 2P(\text{Red} < \text{Blue}) + \frac{1}{n},\]

and thus

\[P(\text{Bizz wins}) = P(\text{Red} < \text{Blue}) = \frac{1}{2} \left(1 - \frac{1}{n}\right) = \frac{1}{2} - \frac{1}{2n}.\]

(c) Consider a random variable \(X\) defined on the probability space of Bizz’s game representing Bizz’z monetary gains/losses: When Bizz wins, it receives \$2, when it loses, it pays \$1. Compute the expectation of \(X\). (Hint: for \(n = 6\) this expectation should be 0.25 so you can check what you got.)

Solution:
We apply the definition of expectation:

\[ E(X) = 2P(X = 2) - 1P(X = -1) \]

\[ = 2P(\text{Bizz wins}) - P(\text{Bizz loses}) \]

\[ = 2P(\text{Bizz wins}) - P(\overline{\text{Bizz wins}}) \]

\[ = 2P(\text{Bizz wins}) - (1 - P(\text{Bizz wins})) \]

\[ = 3P(\text{Bizz wins}) - 1 \]

\[ = 3 \left( \frac{1}{2} - \frac{1}{2n} \right) - 1 \]

\[ = \frac{3}{2} - \frac{3}{2n} - 1 \]

\[ = \frac{1}{2} - \frac{3}{2n} \]

(d) Let \( E \) be the event “Bizz wins”. Fix \( i \) such that \( 1 \leq i \leq n \). Let \( F_i \) be the event “\( \text{Red} = i \)”, Compute the conditional probability \( P(E|F_i) \).

Solution:

\( F_i \) contains \( n \) outcomes—\( (\text{Red}, \text{Blue}) = (i,1), (i,2), \ldots, (i,n) \)—each with probability \( \frac{1}{n^2} \). Of these outcomes, the outcomes that are in the event “Bizz wins” are exactly those for which \( \text{Red} = i < \text{Blue} \), that is, \( (\text{Red}, \text{Blue}) = (i,i+1), (i,i+2), \ldots, (i,n) \). There are \( n - i \) such outcomes, so we have

\[ P(E|F_i) = \frac{P(E \cap F_i)}{P(F_i)} = \frac{(n-i)\frac{1}{n^2}}{n \frac{1}{n^2}} = \frac{n-i}{n} = 1 - \frac{i}{n} \]

21. Consider strings (sequences) over the 3-letter alphabet \( \{a,b,c\} \) and the set \( W \) of these strings such that

- the total number of \( a \)'s in the string is exactly \( m \), where \( m \geq 1 \);
- the total number of \( b \)'s in the string is exactly \( n \), where \( n \geq 1 \);
- there are exactly 3 \( c \)'s in the string.

For the following questions just give the answer. No explanation and no proofs required.

(a) Let \( L \) be the length of a string in \( W \). Compute \( L \) in terms of \( m \) and \( n \).

Solution:

The total number of letters is \( \left\lfloor \frac{m+n+3}{1} \right\rfloor \), since we use all the \( a \)'s, \( b \)'s, and \( c \)'s.

(b) In how many different ways can the \( c \)'s be placed in a string of length \( L \)?

Solution:

Since we have to choose 3 positions for 3 indistinguishable \( c \)'s, there are \( \binom{L}{3} \) different ways the \( c \)'s can be placed.

(c) In terms of \( m \) and \( n \), how many different strings are there in \( W \)?

Solution:
We need to determine how many different arrangements of \( m \) \( a \)'s, \( n \) \( b \)'s, and 3 \( c \)'s we can have in a string of length \( m + n + 3 \). There are \( \binom{m+n+3}{m} \) ways to place the \( m \) \( a \)'s. For each of these placements, there are \( n + 3 \) positions remaining for the \( n \) \( b \)'s and 3 \( c \)'s. There are \( \binom{n+3}{n} \) ways to place the \( b \)'s in \( n \) of these positions. This leaves 3 positions for the 3 \( c \)'s. We can place these in \( \binom{3}{3} \) way. Thus, by the Multiplication Rule, there are \( \binom{m+n+3}{m} \binom{n+3}{n} \binom{3}{3} \) different arrangements of the letters.

If we place the \( b \)'s or \( c \)'s first, we can get different forms for the same answer. The following are some equivalent solutions (there are more):

\[
\binom{m+n+3}{n} \binom{m+3}{3} = \binom{m+n+3}{3} \binom{m+n}{m} = \frac{(m+n+3)!}{m!n!3!}
\]

Note that we can also simply use the formula for permutations of a multiset to obtain the same answer.

22. Prove by ordinary induction that for any \( n \in \mathbb{N} \) we have

\[
\sum_{i=0}^{n} 2^{2i} = \frac{2^{2n+2} - 1}{3}
\]

You must prove this by induction; you are not allowed to use the formula for the sum of a geometric progression.

**Solution:**

**Base case:** \( n = 0 \).

\[
\sum_{i=0}^{n} 2^{2i} = \sum_{i=0}^{0} 2^{2i} = 2^0 = 1 = \frac{4 - 1}{3} = \frac{2^2 - 1}{3} = \frac{2^{0+2} - 1}{3} = \frac{2^{2-1}}{3} = \frac{2^{2} - 1}{3},
\]

so the base case holds.

**Inductive step:** Let \( k \in \mathbb{N} \) be arbitrary, and assume (IH) the proposition holds for \( n = k \), that is, assume

\[
\sum_{i=0}^{k} 2^{2i} = \frac{2^{2k+2} - 1}{3}.
\]

We can see that

\[
\sum_{i=0}^{k+1} 2^{2i} = 2^{2(k+1)} + \sum_{i=0}^{k} 2^{2i} = 2^{2k+2} + \frac{2^{2k+2} - 1}{3} = \frac{3 \cdot 2^{2k+2} + 2^{2k+2} - 1}{3} = \frac{4 \cdot 2^{2k+2} - 1}{3} = \frac{2^{2k+2} - 1}{3} = \frac{2^{2(k+1)+2} - 1}{3},
\]

(\text{by IH})
so the proposition holds for \( n = k + 1 \).

23. 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.

**Solution:**

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 \( r_1 - b_1 - r_2 - b_2 - \ldots - r_k - b_k - r_1 \) where \( k = \min(m, n) \).

24. Let \( A, B, C \) be three finite nonempty sets such that \( |A \cup B \cup C| = |A| + |B| + |C| \). Prove that the sets \( A, B, C \) must be pairwise disjoint, that is, \( A \cap B = B \cap C = C \cap A = \emptyset \).

**Solution:**

**Solution One**  By the given equality and the inclusion–exclusion principle,

\[
|A| + |B| + |C| = |A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|,
\]

and subtracting \( |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| \) from both sides, we obtain

\[
n := |A \cap B| + |A \cap C| + |B \cap C| = |A \cap B \cap C|.
\]

Since each pairwise intersection is a superset of \( A \cap B \cap C \),

\[
|A \cap B| \geq |A \cap B \cap C| \\
|A \cap C| \geq |A \cap B \cap C| \\
|B \cap C| \geq |A \cap B \cap C|.
\]

However, adding these three inequalities shows that

\[
n \geq 3n,
\]

and by algebra we can conclude that \( 0 \geq n \). Thus, \( |A \cap B|, |A \cap C|, \) and \( |B \cap C| \) must all be 0, so \( A, B, \) and \( C \) are pairwise disjoint.

**Solution Two**  Assume, for the sake of contradiction, that some two of the sets are not disjoint, and assume without loss of generality that these two sets are \( A \) and \( B \). Since \( A \) and \( (B \cup C) \setminus A \) are disjoint sets whose union is \( A \cup (B \cup C) \), we know that

\[
n := |A \cup (B \cup C)| = |A| + |(B \cup C) \setminus A|.
\]

Since we assumed \( A \) and \( B \) are not disjoint, we know \( |A \cap B| > 0 \), and since \( A \cap B \subseteq (B \cup C) \cap A \), \( |(B \cup C) \cap A| > 0 \). Thus,

\[
n < |A| + |(B \cup C) \setminus A| + |(B \cup C) \cap A|.
\]
However, since \((B \cup C) \setminus A\) and \((B \cup C) \cap A\) are two disjoint sets whose union is \(B \cup C\), the RHS reduces to \(|A| + |B \cup C|\), and by the union bound, this is no greater than \(|A| + |B| + |C|\). Thus,

\[
n < |A| + |B| + |C|,
\]

which is a contradiction since we know \(n = |A| + |B| + |C|\). Thus, any two of the sets must be disjoint.

25. 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.

Solution:

Solution One

We proceed by induction on \(n\).

Base case: \(n = 4\). If our graph is

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

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, ..., \(n - 3\), \(a\), \(b\), \(c\) (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, ..., \(n - 3\) 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 \(a \to b \to c \to a\), 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.

26. Let \(X\) be a finite nonempty set. Consider the digraph \(G = (V, E)\) such that the vertices are all the subsets of \(X\), i.e., \(V = 2^X\) and the edges are defined as follows. Given two vertices, \(A\) and \(B\), there is a directed edge from \(A\) to \(B\) iff \(B\) has all the elements of \(A\) plus some other element, i.e., there exists some \(x \in X\) such that \(x \notin A\) and \(B = A \cup \{x\}\).

(a) Draw the digraph \(G\) for the case \(X = \{1, 2, 3\}\).

(b) No doubt you noticed that the digraph you drew in part 26a is a DAG. Give a topological sort for it.

Solution:
There are many topological sorts for \(G\). Here is one of them:

\[
\emptyset < \{1\} < \{2\} < \{3\} < \{1, 2\} < \{1, 3\} < \{2, 3\} < \{1, 2, 3\}.
\]

(c) Now \(X\) is an arbitrary finite nonempty set. Prove by induction on directed walk length that if there is a directed walk of length \(\geq 1\) from \(A\) to \(B\) in \(G\) then \(A \subset B\). (Here, \(A \subset B\) stands for “strict” subset, i.e., \(A \subseteq B\) but \(A \neq B\).)

Solution:
We will prove by induction on \(n \geq 1\) that for any directed walk \(A_0 \rightarrow A_1 \rightarrow \cdots \rightarrow A_n\) of length \(n\), we have \(A_0 \subset A_n\).

Base Case \((n = 1)\): In this case the path is a single edge from \(A_0\) to \(A_1\). By the definition of the digraph, \(A_1 = A_0 \cup \{x\}\), where \(x \notin A_0\). Thus \(A_0 \subseteq A_1\), but \(A_0 \neq A_1\), meaning \(A_0 \subset A_1\).
**Inductive Case:** Assume the statement holds for \( n = k \) and suppose we have a walk

\[ A_0 \rightarrow A_1 \rightarrow \cdots \rightarrow A_k \rightarrow A_{k+1} \]

of length \( k + 1 \). By the inductive hypothesis we know \( A_0 \subset A_k \), and by the base case we know \( A_k \subset A_{k+1} \). Since \( \subset \) is a transitive relation, we can conclude \( A_0 \subset A_{k+1} \).

(d) Now \( X \) is again an arbitrary finite nonempty set. Let \( n = |X| \). What is the biggest length that a directed path can have in \( G \)? Give the answer and a short explanation of how you obtained it. No proofs required.

**Solution:**
We can enumerate the elements of \( X \) as \( X = \{ x_1, \ldots, x_n \} \). Define \( A_0 = \emptyset \) and \( A_{i+1} = A_i \cup \{ x_i \} \).

Clearly there is an edge between \( A_i \) and \( A_{i+1} \) for each \( i \), so the following sequence is a path of length \( n \) in \( G \):

\[ A_0 \rightarrow A_1 \rightarrow \cdots \rightarrow A_n. \]

There can be no path of length longer than \( n \); we add one element to the set for every edge we traverse, and no subset of \( X \) can have size strictly greater than \( n \).

(e) Now \( X \) is again an arbitrary finite nonempty set. Prove that \( G \) is a DAG.

**Solution:**
Suppose for the sake of contradiction that \( G \) has a cycle. Then there is some \( A \) with a nonempty walk from \( A \) to \( A \). By part (c) this means \( A \subset A \), and in particular, \( A \neq A \), which is a contradiction. Thus \( G \) is acyclic.

(f) Now \( X \) is again an arbitrary finite nonempty set. Again let \( n = |X| \). Prove that \( G \) has \( n \, 2^{n-1} \) edges. (Hint: you can prove this by induction on \( n \) or you can give a combinatorial proof.)

**Solution:**
**Solution One**  
Proof by induction on \( n \):

**Base case \((n = 0)\):** If \( |X| = 0 \) then \( X = \emptyset \), so \( G \) has \( 0 \cdot 2^{-1} = 0 \) edges.

**Inductive case \((n = k + 1)\):** We can write \( X = X' \cup \{ x \} \) where \( x \notin X' \), and so \( |X'| = k \). Let \( G' \) be the graph associated with \( X' \). We can partition the edges \( (A, A \cup \{ y \}) \) of \( G \) into the following three areas:

i. \( x = y \)
ii. \( x \neq y \) and \( x \notin A \)
iii. \( x \in A \)

The edges in part 26(f)i are exactly

\[ \{(A, A \cup \{ x \}) \mid A \subseteq X' \}, \]

so there are \( 2^k \) such edges. The edges in part 26(f)ii are exactly the edges in \( G' \), so by the inductive hypothesis there are \( k \, 2^{k-1} \) of these. Finally, the edges in part 26(f)iii are

\[ \{(A, A \cup \{ y \}) \mid A = A' \cup \{ x \} \text{ where } A' \subseteq X' \}. \]
In other words, every edge \((A', A' \cup \{y\})\) in \(G'\) corresponds to an edge \((A' \cup \{x\}, A' \cup \{x\} \cup \{y\})\) in part 26(f)iii. Therefore the number of edges in part 26(f)iii is again \(k2^{k-1}\).

Since these three groups form a partition of the set of edges, the total number of edges in \(G'\) is

\[
2^k + k2^{k-1} + k2^{k-1} = 2^k + 2k2^{k-1} = 2^k + k2^k = (k+1)2^k.
\]

**Solution Two**  Combinatorial proof.

By the definition of \(G\), we can see that each edge is of the form \((A, A \cup \{x\})\) for some \(x \in X\), where \(A \subseteq X - \{x\}\). We need to show that each choice of \(x \in X\) and an \(A \subseteq X - \{x\}\) defines a distinct edge in \(G\). Since all edges in \(G\) are of that form, we can then conclude that the number of edges is equal to the number of ways to pick \(x\) and \(A\) according to these two restrictions.

Let \(x_1, x_2 \in X\), \(A_1 \subseteq X - \{x_1\}\), and \(A_2 \subseteq X - \{x_2\}\) be arbitrary. First, we check that \((A, B) := (A_1, A_1 \cup \{x_1\})\) is indeed an edge in \(G\), that is, that \(B = A_1 \cup \{x_1\} \subseteq X\), and that there exists some \(x \in X\), in this case, \(x_1\), such that \(x \notin A\) and \(B = A \cup \{x\}\). Since \(x_1\) and \(A\) were arbitrary objects of the specified form, all selections of \(x_1\) and \(A\) correspond to an edge in \(G\). Now we check that the edge defined by \(x_1\) and \(A_1\) is distinct from the edge defined by \(x_2\) and \(A_2\) if \(x_1 \neq x_2\) or \(A_1 \neq A_2\). We have two cases

Case 1: \(A_1 \neq A_2\).

The edges' source vertices are different, so they are different edges.

Case 2: \(A_1 = A_2\) and \(x_1 \neq x_2\).

If \(A_1 = A_2\), then \(A_1 = A_2 \subseteq X - \{x_2\}\), so \(x_2 \notin A_1\). However, since \(x_1 \neq x_2\), then \(x_2 \notin A_1 \cup \{x_1\}\), so \(A_1 \cup \{x_1\} \neq A_2 \cup \{x_2\}\). This means that the destination vertices of the two edges are different, so they are different edges. In both cases, the two edges are different. Thus, two distinct selections of \(x\) and \(A\) correspond to distinct edges.

The number of ways to select an \(x \in X\) is \(n\). For each choice of \(x\), there are \(2^{n-1}\) ways to choose \(A \subseteq X - \{x\}\), since there are \(2^{|X-\{x\}|} = 2^{|X|} - 1 = 2^n - 1\) subsets of \(X - \{x\}\). Thus, by the multiplication rule, we find that \(n2^{n-1}\) is the number of ways to select \(x\) and \(A\), which we have already shown to be equal to the number of edges in \(G\).

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

(a) The function \(f : \mathbb{N} \rightarrow \mathbb{N}\) \(f(x) = 2^x - 1\) is a bijection, true or false?

**Solution:**

FALSE. The function \(f\) is not a bijection because it is not a surjection: for example there is no \(x\) such that \(f(x) = 2\).

(b) 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?

**Solution:**

TRUE. We know that \(\Pr[B] \leq \Pr[A \cup B] \leq \Pr[A] + \Pr[B]\). So \(\Pr[B] \leq 1/2 \leq 1/4 + \Pr[B]\), and hence \(\Pr[B] \geq 1/4\).
(c) There exists a directed graph with 3 vertices, 2 strongly connected components, and 1 edge, true or false?

Solution:
FALSE. If there are 2 strongly connected components (SCCs) then one of them must have one vertex and the other one two vertices. However, one edge does not suffice to link two vertices into an SCC, you need at least two edges.

(d) Consider an undirected graph with 3 or more vertices and with exactly 3 connected components. In order to make this into a connected graph we must add at least 2 edges, true or false?

Solution:
TRUE. Indeed, the addition of each edge can decrease the number of CCs by at most 1. So we need at least two edges to decrease the number of CCs to 1, i.e. to make the graph connected.

28. Consider strings over the 2-letter alphabet \{a, b\} and the set W of these strings such that

- each string starts with an a and end with two b's;
- there are exactly \(m\) total number of a's in the string (including the first one) where \(m \geq 1\);
- there are exactly \(n\) total number of b's in the string (including the last two) where \(n \geq 2\).

How many such strings are there (in other words, what is \(|W|\))?

Solution:
These strings have length \(m + n\). Three of their positions have fixed elements in them. That leaves \(m + n - 3\) positions in which we can put \(m - 1\) a's and \(n - 2\) b's in every possible way. Each way is determined by choosing (say) the positions in which to put the a's. So the answer is \(\binom{m+n-3}{m-1}\).

29. Let \(X\) be a finite nonempty set and \(f : X \rightarrow X\). Let \(x \in X\) arbitrary and consider the sequence

\[x, f(x), f(f(x)), \ldots, f(\cdots f(x)\cdots), \ldots\]

Prove that for any \(k \geq 2\) there must exist \(k\) distinct positions in this sequence in which the same element of \(X\) occurs.

Solution:
Let \(k \geq 2\) arbitrary and let us denote with \(n\) the number of elements of \(X\). Consider the first \(n(k - 1) + 1\) positions in the sequence (the sequence is of infinite length so these positions exist). Each of these positions has in it an element of \(X\).

CLAIM Among these \(n(k - 1) + 1\) positions there exist \(k\) positions that have the same element of \(X\) in them.

PROOF OF CLAIM Suppose not and therefore each element of \(X\) is repeated at most \(k - 1\) times. This can happen for at most \(n(k - 1)\) positions but we have \(n(k - 1) + 1\) here. Contradiction. (This kind of argument is what proves the generalized pigeonhole principle; the positions are the pigeons and the elements of \(X\) are the pigeonholes; at least one pigeonhole must receive at least \(k\) pigeons.)

REMARKS In fact at least one element of \(X\) must be repeated infinitely many times but we tried to avoid infinite versions of the pigeonhole principle (PHP) in this class. If we want to be more formal about applying the PHP we can proceed as follows.
Let \( W \) be the set of functions of domain \( X \) and codomain \( X \) (i.e., \( W = X^X \)). Define \( \varphi : \mathbb{N} \setminus \{0\} \to W \) as follows: \( \varphi(1) = f \) and \( \varphi(i + 1) = f \circ \varphi(i) \). Now position \( i \geq 1 \) in the sequence holds element \( (\varphi(i))(x) \) of \( X \). And we apply the (generalized) PHP to the function

\[
p : \{1, \ldots, n(k - 1) + 1\} \to X \quad p(i) = (\varphi(i))(x)
\]

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

(a) \( G = (V, E) \) be a (finite) connected undirected graph. Then \( G \) can have at most \(|E| + 1 \) vertices, true or false?

Solution:
TRUE. \( G \) is connected so it has exactly one connected component. By the theorem in the textbook/class \( |E| \geq |V| - 1 \) so \( |V| \leq |E| + 1 \).

(b) Assume that \( B \) is a set with 7 elements and that \( A \) is a set with 15 elements. Then, for any function \( f : A \to B \) there exist at least 3 distinct elements of \( A \) that are mapped by \( f \) to the same element of \( B \), true or false?

Solution:
TRUE. By the generalized pigeonhole principle since \( 15 > 2 \cdot 7 \).

(c) Assume that \( A, B \) are finite nonempty sets and \( f : A \to B \) is a function such that there exist at least 3 distinct elements of \( A \) that are mapped by \( f \) to the same element of \( B \). Then \( |A| > 2 \cdot |B| \), true or false?

Solution:
FALSE. For example \( A = \{a_1, a_2, a_3\} \) and \( B = \{b_1, b_2\} \) and \( f(a_1) = f(a_2) = f(a_3) = b_1 \) but \( |A| = 3 \leq 4 = 2 \cdot |B| \).

(d) The number of sequences of bits such that three of the bits are 0, three of the bits are 1, and are such that they start with a 1 and end with two 0’s is 8, true or false?

Solution:
FALSE. Such sequences look like this: \( 1xyzz00 \) where two of the bits \( x, y, z \) are 1 and the third is 0. This is possible in exactly 3 ways and \( 3 \neq 8 \).

Solution:

(e) The composition (look it up!) \( g \circ f \) of an injection \( f \) with a surjection \( g \) is always a bijection, true or false?

FALSE. Counterexample; consider the functions:

\[
f : \{1, 2\} \to \{a, b, c\}, f(1) = a, f(2) = b
\]

\[
g : \{a, b, c\} \to \{1, 2\}, g(a) = 1, g(b) = 1, g(c) = 2
\]

Then \( g \circ f : \{1, 2\} \to \{1, 2\} \) satisfies \( g \circ f(1) = 1, g \circ f(2) = 1 \) so \( g \circ f \) is not a bijection (in fact it is neither injective nor surjective!).

33
(f) Let $X$ be a finite set with 6 or more elements, and let $p$ be the number of subsets of $X$ of size 2. Then $p > 2|X|$, true or false?

Solution:
TRUE. Let $|X| = n$. Then $p = \binom{n}{2} = n(n - 1)/2$. So is $n(n - 1)/2 > 2n$? This is equivalent to $n^2 > 5n$, which is true when $n \geq 6$.

(g) Let $X$ be a set with $m$ elements and $Y$ be a set with $n$ elements such that $m > n$. Then, there exist at least $n$ distinct surjective function with domain $X$ and codomain $Y$, true or false?

Solution:
TRUE. If $n = 0$ the statement holds because “there exist at least 0 objects such that blah” is vacuously true.
Otherwise, let $X = \{a_1, \ldots, a_m\}$ and $Y = \{b_1, \ldots, b_n\}$. Consider the function from $a_1, \ldots, a_n$ to $Y$ that maps $a_i$ to $b_i$, $i = 1, \ldots, n$. This is clearly surjective. To extend this to functions of domain $X$, map the elements $a_{n+1}, \ldots, a_m$ to any element of $Y$. There are $m - n$ of these “extra” elements in $X$, and each can be mapped to one of $n$ elements. Thus, there are $n^{m-n}$ ways to do this and each such way yields a different surjective function from $X$ to $Y$. Since $m > n$ we have $n^{m-n} > n$, so we have defined at least $n$ distinct surjective functions (there are many more, of course).

31. Sophie is playing the following game:

- First she chooses with equal probability one of the numbers 0, 1, …, 4, call it $a$.
- Then she chooses with equal probability one of the remaining numbers, call it $b$.
- Then she adds $a + b = c$ and $c$ is the result of her game.

(a) What possible results can Sophie’s game have?

Solution:
Any natural number between 1 and 7 can be expressed as the sum of two distinct numbers between 0 and 4:

\[
1 = 0 + 1, 2 = 0 + 2, 3 = 0 + 3, 4 = 0 + 4, 5 = 1 + 4, 6 = 2 + 4, 7 = 3 + 4
\]

but 8 (or bigger numbers cannot). So the possible results are 1, 2, 3, 4, 5, 6, 7.

(b) Draw the “tree of possibilities” diagram for Sophie’s game.

Solution:
See Figure 1.

(c) We denote with $R(c)$ the event that the result of the game is $c$. What is the probability of $R(2)$?

Solution:
Here are all of the probabilities of interest that we could obtain by looking at the tree of possibilities: $P(1) = 2/20$

$P(2) = 2/20$

$P(3) = 4/20$

$P(4) = 4/20$
Figure 1: The “tree of possibilities” for Sophie’s Game.
\[ P(5) = \frac{4}{20} \]
\[ P(6) = \frac{2}{20} \]
\[ P(7) = \frac{2}{20} \]
The probability of \( R(2) \) is \( \frac{2}{20} \).

(d) What is the probability of \( R(2) \cup R(3) \)?

**Solution:**
Since these events are disjoint, we can apply the Sum Rule:
\[ \frac{2}{20} + \frac{4}{20} = \frac{6}{20} \]

(e) What is the conditional probability \( P(R(2) \mid R(2) \cup R(3)) \)?

**Solution:**
We proceed as follows:
\[
P(R(2) \mid R(2) \cup R(3)) = \frac{P(R(2) \cap (R(2) \cup R(3)))}{P(R(2) \cup R(3))} = \frac{P(R(2))}{P(R(2) \cup R(3))} = \frac{2/20}{6/20} = \frac{1}{3}
\]

(f) Find two events \( E,F \) in the probability space of Sophie’s game such that neither is empty, neither equals the whole sample space and \( E \perp F \).

**Solution:**
\( E = \{1,2\} \) and \( F = \{2,3,4\} \).
\[ P(\{1,2\}) = \frac{4}{20} \]
\[ P(\{2,3,4\}) = \frac{10}{20} \]
\[ P(\{1,2\}) \cdot P(\{2,3,4\}) = \left( \frac{4}{20} \right) \cdot \left( \frac{10}{20} \right) = \frac{2}{20} = P(\{1,2\} \cap \{2,3,4\}) \]

32. For any digraph \( G = (V,E) \) without self-loops and without cycles of length 2 we 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 \( u-v \) whenever we have the edge \( u \rightarrow v \) or the edge \( v \rightarrow u \) (or both) in \( G \).

(Note: Going from \( G \) to \( G^u \) is sometimes called ”erasing direction”. You can see that if \( G \) had self-loops or cycles of length 2 then erasing direction naively would produce features not allowed in undirected graphs.)

(a) Prove that if \( G^u \) is acyclic then \( G \) is a DAG. Then give a counterexample that shows that the converse of this statement is false.

**Solution:**
We prove the contrapositive. Suppose \( G \) is not a DAG and thus it has a cycle. By the restrictions on \( G \) this cycle cannot have length 1 or 2 so it must have length 3 or more: \( v_1 \rightarrow v_2 \rightarrow v_3 \rightarrow \cdots \rightarrow v_1 \).
Then \( v_1 \rightarrow v_2 \rightarrow v_3 \rightarrow \cdots \rightarrow v_1 \) is an undirected cycle and therefore \( G^u \) is not acyclic.
The converse is not true. Let \( G = (\{a,b,c\}, \{a \rightarrow b, b \rightarrow c, a \rightarrow c\}) \). This is a DAG but \( G^u \) has a cycle.
(b) Prove that if $G$ is strongly connected then $G^u$ is connected. Then give a counterexample that shows that the converse of this statement is false.

**Solution:**
Again we prove the contrapositive. Suppose that there exist two nodes $v, w$ which are not connected by a walk in $G^u$. Then there is no directed walk $v \rightarrow \cdots \rightarrow w$ in $G$ because erasing directions along this walk would give us a walk in $G^u$. Therefore $G$ is not strongly connected because $w$ is not reachable from $v$ (neither is $v$ reachable from $w$, by the way).

The converse is not true. Take $G = (\{a, b\}, \{a \rightarrow b\})$. $G^u$ is connected but $G$ is not strongly connected.

(c) Prove that if $G$ is a DAG in which every sink is reachable from every source then $G^u$ is connected.

**Solution:**
Assume $G$ is such a DAG and let’s prove that $G^u$ is connected.

Let $v, w$ be two distinct vertices. We want to show that there is a walk $v \cdots w$ in $G^u$.

CLAIM For all vertices $v \in V$, there exists a source $s$ in $G$ such that $v$ is reachable from $s$.

PROOF OF CLAIM Consider the set $S$ of all paths in $G$ from some vertex to $v$. There is at least one such path, the path of length 0. By the Well-Ordering Principle at least one of these paths has maximum length among the paths in $S$, we call it $P$.

If $P$ has length 0 then $v$ itself must be a source, otherwise we can extend the length of the path by including one of $v$’s predecessors. So $v$ is reachable from a source (itself).

Suppose $P$ is not of length 0, let it be $s \rightarrow \cdots \rightarrow v$. Then $s$ must be a source in $G$. Indeed, if $s$ has a predecessor $p$ then either $p$ is among the vertices of $P$ (and then $G$ has a cycle, contradiction) or $p$ is not among the path’s vertices and then $P$ does not have maximum length (also a contradiction) because we can extend it with the edge $p \rightarrow s$. So $v$ is reachable from the source $s$. This ends the proof of the claim.

Similarly, we prove the claim:

CLAIM For all vertices $w \in V$, there exists a sink $t$ in $G$ such that $t$ is reachable from $w$.

So we have a source $s$ and a sink $t$ such that $s \rightarrow v$ and $w \rightarrow t$. In addition we know that in $G$ every sink is reachable from every source therefore $s \rightarrow t$. Now we erase direction on the walks that give $s \rightarrow v, w \rightarrow t$ and $s \rightarrow t$. This gives a walk $v \cdots s \cdots t \cdots w$ in $G^u$. Done.

33. Prove that a connected undirected graph in which every node has degree 2 is a cycle.

**Solution:**
A graph in which every node has degree 2 is called 2-regular. Let $G = (V, E)$ be such a graph. By the handshake lemma $2|E| = 2|V|$ therefore $|E| = |V|$. We will make use of this observation later.

(Since $|E| = |V|$ you might be tempted to invoke Problem 3 on the mock exam above. But that only gives us that the graph contains exactly one cycle, not that it is a cycle!)

So we need to show that any connected 2-regular graph is a cycle. We shall do so by induction on the number $n$ of vertices.

For $n = 1$ and $n = 2$ the statement holds vacuously because graphs with one or two nodes cannot be 2-regular.
BASE CASE  \( n = 3 \). Since it also has 3 edges the only possibility is \( K_3 \), a triangle, hence a cycle.

INDUCTION STEP  Let \( k \geq 3 \) arbitrary. Assume (IH) that any 2-regular graph with \( k \) nodes is a cycle. Now let \( G = (V, E) \) be a connected 2-regular graph with \( k + 1 \) nodes. We want to show that \( G \) is a cycle.

Pick a node \( u \in V \). We have exactly two nodes adjacent to \( u \) in \( G \) call them \( v \) and \( w \). Since \( |V| \geq 4 \) we must have yet another node in \( G \), call it \( z \).

CLAIM  \( v-w \notin E \)

PROOF OF CLAIM  By contradiction. If \( v-w \in E \) then \( u, v, w \) are all adjacent to each other and not adjacent to any other node in \( V \) (because the graph is 2-regular). Therefore there cannot exist a walk from \( z \) to, say, \( v \) and thus \( G \) is not connected, contradiction.

Now we construct from \( G \) a new graph \( G' \) by deleting the node \( u \), the edges \( u-v \) and \( u-w \), and then adding an edge \( v-w \). Since \( v \) and \( w \) still have degree 2 in \( G' \), \( G' \) must be 2-regular. Moreover, \( G' \) is connected. Indeed, any walk in \( G \) that passed through \( u \) must have used the edges \( v-u \) and \( u-w \). These edges, as well as \( u \) are gone, but we can ”repair” the path in \( G' \) by using the edge \( v-w \).

It follows that \( G' \) is a connected 2-regular graph with \( k \) nodes. By IH it is a cycle. Now we go back to \( G \) from the cycle \( G' \) by deleting the edge \( v-w \) and adding instead the vertex \( u \) and edges \( v-u \) and \( u-w \). Clearly this constructs a cycle.

34. Let \( G \) be a graph with \( n \) vertices and exactly \( n-1 \) edges. Prove that \( G \) has either a vertex of degree 1 or an isolated vertex.

**Solution:**

Denote \( G = (V, E) \). If \( G \) were a tree we would be done because we know (textbook, lecture) that every tree with two or more vertices has at least one leaf and that in a tree with one vertex that vertex is isolated.

Now we consider the case where \( G \) is not a tree. Then \( G \) is not connected, since we have that \( |V| = |E| + 1 \) and all connected graphs with \( |V| = |E| + 1 \) are trees. Then \( G \) has 2 or more connected components. We will show that at least one of its connected components must be a tree.

Let \( V_1, \ldots, V_k \) be the connected components of \( G \) and let \( E_1, \ldots, E_k \) be the sets of edges of these connected components.

CLAIM  For some \( 1 \leq i \leq k \) we have \( |E_i| < |V_i| \).

PROOF OF CLAIM  By contradiction. Suppose \( |E_i| \geq |V_i| \) for all \( i = 1, \ldots, k \). Adding the left hand sides and the right hand sides of these \( k \) inequalities, we conclude \( |E| \geq |V| \) because the connected components are both vertex-disjoint and edge-disjoint. Thus \( |V|-1 \geq |V| \), contradiction, and this ends the proof of the claim.

WLOG let \( V_j, E_j \) be a connected component for which \( |E_j| < |V_j| \). As a connected subgraph this component also satisfies \( |E_j| \geq |V_j| - 1 \). Therefore \( |E_i| = |V_i| - 1 \) and since it is connected, this component must be a tree. This component must have a leaf or is an isolated vertex (if the tree has no edges).