

Midterm 2 Solutions

Note: These solutions are not necessarily model answers. They are designed to be tutorial in nature, and sometimes contain a little more explanation than an ideal solution. Also, bear in mind that there may be more than one correct solution. The maximum number of points is 60. Comments in italics following the solutions highlight some common errors or give explanations.

1. Euclid's tunnels [7 points]

- (a) The graph has an Euler path starting at A and ending at D iff A and D have odd degree and all other vertices have even degree. The tunnel BC should be built to ensure this condition. 3pts
- (b) The modified graph has an Euler path from B to D iff B and D have odd degree and all other vertices have even degree. The tunnel AB should be removed to ensure this condition. 2pts
- (c) A graph has an Euler tour iff all the vertex degrees are even. Adding the tunnel BD ensures that all vertices have even degree. (Alternatively, we can think of the additional tunnel BD as closing the Euler path in part (b).) 2pts

Comments: Almost all people solved this problem correctly. A few people did not realize that part (b) was equivalent to the existence of an Euler path from B to D . One point was taken off for not explaining the answers.

2. Counting, counting... [12 points]

- (a) Depending on the number of E's appearing in the six-letter word, we have two cases: 3pts
- E appears zero or one times: there are $\binom{6}{6} = 1$ choice of six letters with no E, and $\binom{6}{5} = 6$ choices of six letters with one E; each of these choices contributes $6!$ distinct six-letter words;
 - E appears twice: there are $\binom{6}{4}$ choices of six letters with two E's; each of these choices contributes $6!/2!$ distinct six-letter words (the factor $2!$ accounts for swapping the two E's).

In total, there are $\boxed{\binom{6}{6}6! + \binom{6}{5}6! + \frac{6!}{2!}\binom{6}{4}}$ distinct six-letter words.

Comments: Quite a few students solved this problem correctly. (Many students divided the first case above into two separate cases, for no E's and one E respectively; this gives the same answer and is fine.) Many people simply answered $\frac{6!\binom{8}{6}}{2!}$ (or some variant of it), without realizing that the factor $2!$ is not always present (in the first case above). Partial credit was given to solutions which considered separately the case when both E's are present but did not get the correct answer due to careless mistakes.

- (b) The number of triples (a, b, c) satisfying $0 \leq a \leq b \leq c \leq n$ is exactly the number of ways to sample three unordered objects from $n + 1$ objects with replacement, which is $\binom{(n+1)+3-1}{3} = \boxed{\binom{n+3}{3}}$. 3pts

For example, if $n = 4$ and the three sampled objects are $\{0, 4, 4\}$ (as an unordered set), then $(a, b, c) = (0, 4, 4)$. The formula for counting unordered sets can be found in Lecture Note 10. Note that we are sampling from $\{0, 1, \dots, n\}$, a set of size $n + 1$ (instead of n).

Alternative Solution We count the triples (a, b, c) separately for the following three cases:

- $0 \leq a < b < c \leq n$: there are $\binom{n+1}{3}$ such triples;
- $0 \leq a < b = c \leq n$: there are $\binom{n+1}{2}$ such triples, because we are just counting pairs (a, b) such that $0 \leq a < b \leq n$;
- $0 \leq a = b < c \leq n$: again, there are $\binom{n+1}{2}$ such triples;

- $0 \leq a = b = c \leq n$: there are $\binom{n+1}{1}$ such triples, because we are just counting elements a such that $0 \leq a \leq n$.

In total, there are $\boxed{\binom{n+1}{3} + 2\binom{n+1}{2} + \binom{n+1}{1}}$ such triples.

After simplification (not required), this of course gives the same number $\binom{n+3}{3}$ as in other solutions.

Comments: A common minor mistake was to give n (rather than $n + 1$) as the number of elements from which we are sampling. A much bigger common mistake was to answer $\binom{n+1}{3}$ (or $\binom{n}{3}$), which counts the number of objects without replacement. Solutions involving iterated summation (e.g. $\sum_{0 \leq a \leq n} \sum_{a \leq b \leq n} (n - b + 1)$) or in terms of a, b or c (e.g. $(n - a)(n - b)(n - c)$) were clearly way off base.

- (c) Assume that we name the three committees A, B and C . Every citizen has exactly six choices $\{A, B, C, AB, AC, BC\}$ to serve on at least one and not all three committees. Since there are n citizens, each having six possible choices, the required number is $\boxed{6^n}$. 3pts

Comments: Many people got this right. Some people made mistakes and got instead $5^n, 7^n$ or 3^n ; they received partial credit.

- (d) **Solution 1** The number of ways to seat all n students is $n!$. The number of ways to seat all n students such that the two friends are next to each other is $2(n - 1)!$: to see this, note that we can then treat the two friends A and B as a “super-student” (AB or BA), giving $n - 1$ “students” and thus $(n - 1)!$ ways of seating them; the factor of 2 comes from the ordering of A and B . The number of seatings in which they are not next to each other is therefore $\boxed{n! - 2(n - 1)!}$. 3pts

Solution 2 Let the two friends be A and B , and the seats be numbered 1 through n . Depending on the seating of A , we count the number of ways of seating B not next to A :

- A takes the first seat 1: B can sit in seats 3 through n , giving a total of $n - 2$ choices;
- A takes the last seat n : this is symmetric to the previous case, giving a total of $n - 2$ choices;
- A takes a seat from 2 to $n - 1$: B cannot take the two seats next to A (nor the seat taken by A), and thus has a total of $n - 3$ choices.

Hence the number of possible ways of seating A and B is $2(n - 2) + (n - 2)(n - 3) = (n - 2)(n - 1)$. For each of these choices, there are $(n - 2)!$ ways to seat the remaining $n - 2$ students (occupying seats not taken by A or B). Therefore, the total number of ways to seat all n students is $(n - 2)(n - 1)(n - 2)! = \boxed{(n - 2)(n - 1)!}$.

The answers in both solutions of course agree with each other. For Solution 2, it is also possible to count the number of seatings of just A and B (i.e. $(n - 2)(n - 1)$) using ideas related to sampling with replacement (See part (o) of Question 1, Homework 7).

Comments: Many people got this right. Less compact answers (e.g. $[2(n - 2) + (n - 2)(n - 3)](n - 2)!$) were also accepted. Partial credit was given to solutions which demonstrated a valid counting argument, but did not arrive at the correct answer due to careless mistakes.

3. A random number of dice throws [12 points]

- (a) This part asks for the probability of scoring a total of 3 points in two throws of a fair die. This is a uniform probability space with $6 \times 6 = 36$ outcomes. Exactly two of these outcomes give a score of 3 (namely, $(1, 2)$ and $(2, 1)$), so we have $\Pr[X = 3 \mid N = 2] = \frac{2}{36} = \frac{1}{18}$. 3pts

Comments: The most common error here was to confuse the **conditional** probability $\Pr[X = 3 \mid N = 2]$ with the **joint** probability $\Pr[X = 3 \wedge N = 2]$. Many students thought that the above value, $\frac{2}{36}$, is actually the joint probability, so to get the conditional probability they divided by $\Pr[N = 2] = \frac{1}{4}$, giving the incorrect solution $\frac{2/36}{1/4} = \frac{2}{9}$. Note that the correct value for the joint probability is actually $\Pr[X = 3 \wedge N = 2] = \Pr[N = 2] \times \Pr[X = 3 \mid N = 2] = \frac{1}{4} \times \frac{2}{36} = \frac{1}{72}$.

(b) Using the total probability rule, we have

3pts

$$\Pr[X = 3] = \sum_{i=1}^4 \Pr[X = 3 \mid N = i] \times \Pr[N = i]. \quad (1)$$

(This follows from the fact that the events $\{N = i\}$ for $1 \leq i \leq 4$ *partition* the probability space: i.e., they are disjoint and cover the whole space.) From part (a) we have $\Pr[X = 3 \mid N = 2] = \frac{1}{18}$. Also, $\Pr[X = 3 \mid N = 1] = \frac{1}{6}$ (there is only one way to score 3 on a single throw); and $\Pr[X = 3 \mid N = 3] = \frac{1}{6^3} = \frac{1}{216}$ (there is only one way to score 3 on three throws, and there are 6^3 total outcomes); and $\Pr[X = 3 \mid N = 4] = 0$ (it is not possible to score 3 on four throws). Plugging these values into (1) gives

$$\Pr[X = 3] = \frac{1}{6} \times \frac{1}{2} + \frac{1}{18} \times \frac{1}{4} + \frac{1}{216} \times \frac{1}{8} + 0.$$

Comments: The most common error here was to write

$$\Pr[X = 3] = \sum_{i=1}^4 \Pr[X = 3 \mid N = i]$$

instead of equation (1) above—i.e., to forget to weight the conditional probabilities by the probabilities $\Pr[N = i]$. Another common error was further confusion between conditional and joint probabilities, as in part (a). It is OK to work with joint probabilities here, in which case (1) would be replaced by

$$\Pr[X = 3] = \sum_{i=1}^4 \Pr[X = 3 \wedge N = i].$$

But this is actually equivalent to (1) because of course $\Pr[X = 3 \wedge N = i] = \Pr[X = 3 \mid N = i] \times \Pr[N = i]$.

(c) We have

3pts

$$\begin{aligned} \Pr[X = 3 \mid N \text{ even}] &= \frac{\Pr[X = 3 \wedge N \text{ even}]}{\Pr[N \text{ even}]} = \frac{\Pr[X = 3 \wedge N = 2] + \Pr[X = 3 \wedge N = 4]}{\Pr[N = 2] + \Pr[N = 4]} \\ &= \frac{\Pr[X = 3 \mid N = 2] \Pr[N = 2] + 0}{\Pr[N = 2] + \Pr[N = 4]} \\ &= \frac{(1/18) \times (1/4) + 0}{1/2 + 1/4} = \frac{1}{27}. \end{aligned}$$

Comments: A lot of students were confused by this part. Most of the confusion stemmed from mixing up conditional and joint probabilities (similar to parts (a) and (b) above). Another very common error was to write

$$\Pr[X = 3 \mid N \text{ even}] = \Pr[X = 3 \mid N = 2] + \Pr[X = 3 \mid N = 4],$$

which is simply false. A correct version of this statement is

$$\Pr[X = 3 \mid N \text{ even}] = \Pr[X = 3 \mid N = 2] \Pr[N = 2 \mid N \text{ even}] + \Pr[X = 3 \mid N = 4] \Pr[N = 4 \mid N \text{ even}],$$

which leads to the same correct answer as above.

(d) By Bayes' Rule, we have

3pts

$$\Pr[N = 2 \mid X = 3] = \frac{\Pr[X = 3 \mid N = 2] \Pr[N = 2]}{\Pr[X = 3]} = \frac{(1/18) \times (1/4)}{\text{part (b)}},$$

where the first factor in the numerator comes from part (a).

Comments: Most people got this right, even if they had trouble with some of the previous parts. Full credit was given for a correct use of Bayes' Rule even if the values substituted from parts (a) and/or (b) were incorrect.

4. Locks and keys [10 points]

(a) We open one box iff box 1 contains key 1. The key inside box 1 is chosen uniformly at random, hence key 1 is found inside box 1 with probability $1/n$. 2pts

(b) Let A be the event that box 1 does not contain key 1 and B be the event that the second box opened contains key 1. Two boxes will be opened iff the event $A \cap B$ occurs: 2pts

$$\Pr[A \cap B] = \Pr[A] \times \Pr[B|A] = \frac{n-1}{n} \times \frac{1}{n-1} = \frac{1}{n}.$$

(Note that $\Pr[B|A]$ is equal to $1/(n-1)$ because, given that key 1 is not in box 1, it is equally likely to be contained in any of the remaining $n-1$ boxes.)

(c) Let A_i denote the event that the i -th box opened does not contain the key to box 1. Exactly k boxes will be opened iff the event $A_1 \cap A_2 \cap \dots \cap A_{k-1} \cap \overline{A_k}$ occurs. We compute using the definition of conditional probability: 3pts

$$\begin{aligned} \Pr[A_1 \cap A_2 \cap \dots \cap A_{k-1} \cap \overline{A_k}] &= \Pr[A_1] \times \Pr[A_2|A_1] \times \dots \times \Pr[A_{k-1}|\bigcap_{j<k-1} A_j] \times \Pr[\overline{A_k}|\bigcap_{j<k} A_j] \\ &= \frac{n-1}{n} \times \frac{n-2}{n-1} \times \dots \times \frac{n-k+1}{n-k} \times \frac{1}{n-k+1} = \frac{1}{n}. \end{aligned}$$

(Note that, if the first $i-1$ boxes opened do not contain key 1, key 1 is equally likely to be found among the remaining $n-i+1$ boxes, hence $\Pr[A_i|\bigcap_{j<i} A_j] = (n-i)/(n-i+1)$.)

Alternative solution: The total number of assignments of n keys to n boxes is equal to $n!$. We count the number of assignments for which exactly k boxes are opened. The k boxes $b_1 = 1, b_2, \dots, b_k$ are opened iff box b_i contains key b_{i+1} for $1 \leq i < k$ and box b_k contains key 1, while the remaining boxes can contain arbitrary keys.

The order in which the k boxes are opened can be chosen in $\binom{n-1}{k-1} \cdot (k-1)!$ ways, and the key assignment to these k boxes is fixed given the order. There are $(n-k)!$ ways to assign keys to the remaining boxes. Hence

$$\Pr[\text{exactly } k \text{ boxes opened}] = \frac{\binom{n-1}{k-1} \cdot (k-1)! \cdot (n-k)!}{n!} = \frac{(n-1)!}{n!} = \frac{1}{n}.$$

(d) Let X be the random variable counting the number of boxes opened. Then from part (c) we have $\Pr[X = k] = 1/n$ for all $k \leq n$. Hence by definition of expectation, we have 3pts

$$E[X] = \sum_{k=1}^n k \times \Pr[X = k] = \sum_{k=1}^n \frac{k}{n} = \frac{n(n+1)}{2n} = \frac{n+1}{2}.$$

Comments: Some students did not calculate the conditional probabilities correctly in part (b). Many students did not generalize the argument in part (b) to solve part (c). One point was taken off for not simplifying the product in part (c) or for inadequate explanations.

5. Random graphs [9 points]

(a) Let X be the number of Hamiltonian cycles in the random graph G with n vertices. We want to express X as a sum of simple indicator random variables and apply linearity of expectation. A Hamiltonian cycle in G is specified by a sequence of edges $\{(v_1, v_2), (v_2, v_3), (v_3, v_4), \dots, (v_{n-1}, v_n), (v_n, v_1)\}$, where $\{v_1, v_2, \dots, v_n\}$ is a permutation of $\{1, 2, \dots, n\}$, the n vertices of G . For each permutation $\{v_1, v_2, \dots, v_n\}$ of $\{1, 2, \dots, n\}$, we define the (indicator) random variable $X_{\{v_1, v_2, \dots, v_n\}}$ to be 1 if the Hamiltonian cycle corresponding to the permutation $\{v_1, v_2, \dots, v_n\}$ is present; namely, if the n

edges $\{(v_1, v_2), (v_2, v_3), \dots, (v_{n-1}, v_n), (v_n, v_1)\}$ are all present. We may therefore write the number of Hamiltonian cycles as

$$X = \sum_{\{v_1, v_2, \dots, v_n\}} X_{\{v_1, v_2, \dots, v_n\}}. \quad (2)$$

Now we know that

$$\begin{aligned} E[X_{\{v_1, v_2, \dots, v_n\}}] &= \Pr[X_{\{v_1, v_2, \dots, v_n\}} = 1] \\ &= \Pr[(v_1, v_2), (v_2, v_3), \dots, (v_{n-1}, v_n), (v_n, v_1) \text{ are all present}] \\ &= p^n, \end{aligned}$$

because each edge is present with probability p , independently of others. Now by linearity of expectation, from (2) we compute

$$E[X] = E\left[\sum_{\{v_1, v_2, \dots, v_n\}} X_{\{v_1, v_2, \dots, v_n\}}\right] = \sum_{\{v_1, v_2, \dots, v_n\}} E[X_{\{v_1, v_2, \dots, v_n\}}] = \sum_{\{v_1, v_2, \dots, v_n\}} p^n = n!p^n,$$

because there are $n!$ different permutations of $\{1, 2, \dots, n\}$.

Comments: Depending on how one interprets when two Hamiltonian cycles are different, we may get different answers. For example, if we treat $\{1, 2, 3, \dots, n-1, n\}$ as the same Hamiltonian cycle as $\{2, 3, \dots, n-1, n, 1\}$ (i.e., we treat a shifted cycle as the same as the original), then we get the answer $E[X] = (n-1)!p^n$. If we further treat $\{1, 2, 3, \dots, n-1, n\}$ as the same Hamiltonian cycle as $\{n, n-1, \dots, 3, 2, 1\}$ (i.e., we further treat a reversed cycle as the same as the original), then we get the answer $E[X] = (n-1)!p^n/2$. All such answers received full credit. Few people were able to solve this problem correctly using linearity of expectation. Partial credit was given for the expectation of the individual indicator random variables $E[X_{\{v_1, v_2, \dots, v_n\}}] = p^n$.

3pts

- (b) Let X be the number of Hamiltonian paths from 1 to n in the random graph G . A Hamiltonian path from 1 to n is specified by a sequence of edges $\{(v_1, v_2), (v_2, v_3), \dots, (v_{n-1}, v_n)\}$, where $\{v_1, v_2, v_3, \dots, v_n\}$ is a permutation of $\{1, 2, \dots, n\}$ such that $v_1 = 1$ and $v_n = n$. For each permutation $\{v_2, v_3, \dots, v_{n-1}\}$ of $\{2, 3, \dots, n-1\}$, we define the (indicator) random variable $X_{\{v_2, v_3, \dots, v_{n-1}\}}$ to be 1 if the Hamiltonian path $\{1, v_2, v_3, \dots, v_{n-1}, n\}$ is present; namely, if the $n-1$ edges $\{(1, v_2), (v_2, v_3), (v_3, v_4), \dots, (v_{n-2}, v_{n-1}), (v_{n-1}, n)\}$ are all present. Similar to part (a) we have

$$\begin{aligned} E[X_{\{v_2, v_3, \dots, v_{n-1}\}}] &= \Pr[X_{\{v_2, v_3, \dots, v_{n-1}\}} = 1] \\ &= \Pr[(1, v_2), (v_2, v_3), \dots, (v_{n-2}, v_{n-1}), (v_{n-1}, n) \text{ are all present}] \\ &= p^{n-1}, \end{aligned}$$

by the independence of the edges. Now by linearity of expectation we compute

$$\begin{aligned} E[X] &= E\left[\sum_{\{v_2, v_3, \dots, v_{n-1}\}} X_{\{v_2, v_3, \dots, v_{n-1}\}}\right] = \sum_{\{v_2, v_3, \dots, v_{n-1}\}} E[X_{\{v_2, v_3, \dots, v_{n-1}\}}] \\ &= \sum_{\{v_2, v_3, \dots, v_{n-1}\}} p^{n-1} = (n-2)!p^{n-1}, \end{aligned}$$

because there are $(n-2)!$ different permutations of $\{2, 3, \dots, n-1\}$.

Few students were able to solve this problem correctly.

3pts

- (c) Let X be the number of simple paths of length k in G . Any such path in G is specified by a sequence of k edges $\{(v_0, v_1), (v_1, v_2), \dots, (v_{k-1}, v_k)\}$, where the $k+1$ vertices $\{v_0, v_1, \dots, v_k\}$ are distinct (because the path is simple). For each enumeration $\{v_0, v_1, \dots, v_k\}$ of $k+1$ distinct vertices of $\{1, 2, \dots, n\}$, we define the (indicator) random variable $X_{\{v_0, v_1, \dots, v_k\}}$ to be 1 if the simple path

$\{v_0, v_1, \dots, v_k\}$ is present; namely, if the k edges $\{(v_0, v_1), (v_1, v_2), \dots, (v_{k-1}, v_k)\}$ are all present. We have

$$\begin{aligned} E[X_{\{v_0, v_1, \dots, v_k\}}] &= \Pr[X_{\{v_0, v_1, \dots, v_k\}} = 1] \\ &= \Pr[(v_0, v_1), (v_1, v_2), \dots, (v_{k-1}, v_k) \text{ are all present}] \\ &= p^k, \end{aligned}$$

because each edge is present with probability p , independently of others. Now we compute

$$\begin{aligned} E[X] &= E\left[\sum_{\{v_0, v_1, \dots, v_k\}} X_{\{v_0, v_1, \dots, v_k\}}\right] = \sum_{\{v_0, v_1, \dots, v_k\}} E[X_{\{v_0, v_1, \dots, v_k\}}] \\ &= \sum_{\{v_0, v_1, \dots, v_k\}} p^k = \binom{n}{k+1} (k+1)! p^k, \end{aligned}$$

because there are $\binom{n}{k+1}$ choices of $k+1$ distinct vertices out of n vertices, and each contributing $(k+1)!$ enumerations.

Comments: Similarly to part (a), we may get different answers under different interpretations. For example, if we treat $\{0, 1, 2, \dots, k\}$ as the same path as $\{k, k-1, \dots, 1, 0\}$ (i.e., we treat a reversed path as the same as the original path), then we get the answer $E[X] = \binom{n}{k+1} (k+1)! p^k / 2$. All such variants received full credit. Few students were able to solve this problem fully. There were partial deductions for miscalculation of the factors $\binom{n}{k+1} (k+1)!$ or p^k .

6. A Declaration of (In)dependence [10 points]

(a) Given that $N = 100$, the distribution of X is just the number of Heads that occur in 100 independent coin tosses, each of whose probability of landing Heads is p . This is simply the Binomial distribution $Bin(100, p)$. 2pts

(b) From the law of total probability, $\Pr[X = k] = \sum_{i=0}^{\infty} \Pr[X = k | N = i] \times \Pr[N = i]$. If $i < k$, $\Pr[X = k | N = i] = 0$, and otherwise, from part (a), the probability is given by $Bin(i, p)$. Additionally, $\Pr[N = i]$ is given by the Poisson distribution. Thus the above summation is 2pts

$$\Pr[X = k] = \sum_{i=k}^{\infty} \binom{i}{k} p^k (1-p)^{i-k} \frac{\lambda^i e^{-\lambda}}{i!}.$$

(c) We are given that X is distributed according to a Poisson distribution, thus all we need to do is figure out $E[X]$, since that will be the parameter of the Poisson distribution. Since $E[N] = \lambda$, and, in expectation, a fraction p of the coins will be Heads, we get $E[X] = \lambda p$. (This can be made rigorous using linearity of expectation.) Thus 2pts

$$\Pr[X = k] = \frac{(\lambda p)^k e^{-\lambda p}}{k!}.$$

Comments: Quite a few people gave Np as the parameter of the Poisson distribution. This is the right idea, but N is a random variable, not a parameter.

(d) From the definition of conditional probability, $\Pr[X = k | Y = i] = \frac{\Pr[X=k \text{ and } Y=i]}{\Pr[Y=i]}$. First observe that by part (c), $Y \sim Pois(\lambda(1-p))$. Next, observe that $\Pr[X = k \text{ and } Y = i] = \Pr[X = k | N =$ 3pts

$k + i] \times \Pr[N = k + i]$. Plugging these values in and simplifying, we get:

$$\begin{aligned}
 \Pr[X = k|Y = i] &= \frac{\Pr[X = k \text{ and } Y = i]}{\Pr[Y = i]} = \frac{\Pr[X = k|N = k + i] \times \Pr[N = k + i]}{\frac{(\lambda(1-p))^i e^{-\lambda(1-p)}}{i!}} \\
 &= \frac{\binom{k+i}{i} p^k (1-p)^i \frac{\lambda^{k+i} e^{-\lambda}}{(k+i)!}}{\frac{(\lambda(1-p))^i e^{-\lambda(1-p)}}{i!}} \\
 &= \frac{(k+i)! i! p^k (1-p)^i \lambda^{k+i} e^{-\lambda}}{i! k! (k+i)! \lambda^i (1-p)^i e^{-\lambda} e^{\lambda p}} \\
 &= \frac{p^k \lambda^k e^{-\lambda p}}{k!} = \frac{(\lambda p)^k e^{-\lambda p}}{k!}.
 \end{aligned}$$

Comments: Several people got the correct initial expression, but made no effort to simplify it, and thus failed to realize that this conditional probability is the same as the unconditional probability of part (c)—as is required for part (e).

- (e) From parts (c) and (d), we have $\Pr[X = j] = \frac{(\lambda p)^j e^{-\lambda p}}{j!} = \Pr[X = j|Y = i]$. Thus by definition the event $X = j$ is independent from the event $Y = i$. 1pt

Comments: Many people claimed that they were dependent because $Y = N - X$. This would be correct if N were fixed to be some value; however, N is NOT fixed, and is explicitly given to be a random variable.