1. Four candidates were separately evaluated and ranked by a two-person committee. When the two committee members compared their rankings, they found that they agreed on the top two candidates, but reversed the rankings of those two candidates. Otherwise they were in total agreement. What is the Kendall’s $\tau$ correlation of their evaluations?

A) $\frac{1}{12}$  B) $\frac{1}{6}$  C) $\frac{1}{4}$  D) $\frac{1}{3}$  E) $\frac{5}{12}$
F) $\frac{1}{2}$  G) $\frac{7}{12}$  H) $\frac{2}{3}$  I) $\frac{3}{4}$  J) $\frac{5}{6}$

Solution. The four paired points are $(1,1)$, $(2,2)$, $(3,4)$, and $(4,3)$ (where, as usual in statistics, the ranking numbers are lowest to highest). Comparing $(1,1)$ to the three pairs that follow, there are three concordant pairs. Comparing $(2,2)$ to the two pairs that follow, there are two concordant pairs. Comparing $(3,4)$ to the pair that follows, there is one discordant pair. Therefore, $\tau = (3 + 2 - 1)/6 = 2/3$.

Answer: H

2. The exam scores of a class of three calculus students on their first two exams are given in the table. What is Spearman’s $\rho$ correlation for the exam 1 and exam 2 scores?

<table>
<thead>
<tr>
<th>Student</th>
<th>Exam 1 Score</th>
<th>Exam 2 Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rudiger</td>
<td>80</td>
<td>90</td>
</tr>
<tr>
<td>Rufus</td>
<td>85</td>
<td>70</td>
</tr>
<tr>
<td>Rupert</td>
<td>90</td>
<td>100</td>
</tr>
</tbody>
</table>

A) 0.46  B) 0.48  C) 0.50  D) 0.52  E) 0.54
F) 0.56  G) 0.58  H) 0.60  I) 0.62  J) 0.64

Solution. The Spearman’s $\rho$ correlation for the data is the Pearson’s $r$, 0.5, for the rankings:

<table>
<thead>
<tr>
<th>Student</th>
<th>Exam 1 Rank</th>
<th>Exam 2 Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rudiger</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Rufus</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Rupert</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

Answer: C

3. After two exams, a statistics professor noticed that the students at the top of the class on the first exam were at the top on the second exam. The students at the bottom of the class on the first exam were at the bottom on the second exam. A quick glance suggested more mobility in the middle, so he randomly selected five students in that group and compared their class rankings in the usual statistical manner.

<table>
<thead>
<tr>
<th>Student</th>
<th>Exam 1 Rank</th>
<th>Exam 2 Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jo</td>
<td>56</td>
<td>65</td>
</tr>
<tr>
<td>Jy</td>
<td>61</td>
<td>66</td>
</tr>
<tr>
<td>Ja</td>
<td>69</td>
<td>63</td>
</tr>
<tr>
<td>Ju</td>
<td>73</td>
<td>74</td>
</tr>
<tr>
<td>Ji</td>
<td>75</td>
<td>72</td>
</tr>
</tbody>
</table>
What is Spearman’s correlation for the exam 1 and exam 2 rankings in this sample?

A) 0.46  B) 0.48  C) 0.50  D) 0.52  E) 0.54  
F) 0.56  G) 0.58  H) 0.60  I) 0.62  J) 0.64

Solution. For a Spearman’s correlation, we replace the data with the rankings of the data.

<table>
<thead>
<tr>
<th>Student</th>
<th>Exam 1 Rank</th>
<th>Exam 2 Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jo</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Jy</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Ja</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Ju</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Ji</td>
<td>5</td>
<td>4</td>
</tr>
</tbody>
</table>

The mean of each column is 3 and the standard deviations of column 1 and column 2, $s_1$ and $s_2$ are given by

$$s_1 = s_2 = \sqrt{\frac{1}{4} ((1 - 3)^2 + (2 - 3)^2 + (3 - 3)^2 + (4 - 3)^2 + (5 - 3)^2)} = \frac{\sqrt{10}}{2}.$$  

Spearman’s $\rho$ for the data in the given column is Pearson’s $r$ for the data in the table just given:

$$\rho = \frac{1}{(5-1) \cdot s_1 \cdot s_2} \left( (1-3)(2-3)(3-3)(4-3)(5-3)(1-3)(2-3)(3-3)(4-3)(5-3) \right) = \frac{6}{4 \left( \frac{\sqrt{10}}{2} \right)^2} = 0.6.$$

Alternatively, the differences in ranks for the rows are given by $d_1 = -1, d_2 = -1, d_3 = 2, d_4 = -1, d_5 = 1$.

Using these rank differences in accordance with the strange formula for $\rho$ when there are no ties, we have

$$\rho = 1 - \frac{6}{5 (5^2 - 1)} \left( (-1)^2 + (-1)^2 + (2)^2 + (-1)^2 + (1)^2 \right) = 1 - \frac{6}{5 \cdot 24} \times 8 = 1 - \frac{8}{20} = \frac{3}{5} = 0.6.$$

Answer: H

4. In the U.S., there are many one-issue voters. Such a voter supports one political party because of that party’s stance on the issue that is of greatest interest to the voter. The voter may not care very much about the party’s platform on other issues. It can therefore be interesting to correlate the positions found in different groups of voters (such as abortion foe Republicans and fiscal hawk Republicans, environmental Democrats and organized labor Democrats). In a survey of young members of the party, two statements were posited, A) The federal government should do more to encourage the development of renewable sources of energy and B) The federal government should begin to decriminalize recreational drug use. Surveyees were to respond with standard Likert scale levels of agreement: Strongly disagree, Disagree somewhat, Neither disagree nor agree, Agree somewhat, or Strongly agree. The responses of six surveyees are tabulated.

<table>
<thead>
<tr>
<th>Party Member</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan</td>
<td>Disagree somewhat</td>
<td>Strongly disagree</td>
</tr>
<tr>
<td>Jane</td>
<td>Neither disagree nor agree</td>
<td>Disagree somewhat</td>
</tr>
<tr>
<td>Jim</td>
<td>Neither disagree nor agree</td>
<td>Neither disagree nor agree</td>
</tr>
<tr>
<td>Joe</td>
<td>Disagree somewhat</td>
<td>Strongly agree</td>
</tr>
<tr>
<td>Jon</td>
<td>Agree somewhat</td>
<td>Strongly agree</td>
</tr>
<tr>
<td>Jos</td>
<td>Strongly disagree</td>
<td>Agree somewhat</td>
</tr>
</tbody>
</table>

Answer: H
Kendall $\tau$ correlate them.

A) 0.07 B) 0.10 C) 0.13 D) 0.17 E) 0.20
F) 0.25 G) 0.33 H) 0.40 I) 0.47 J) 0.50

**Solution.** We first convert the responses to Likert scale numbers from 1 to 5, with 1 representing the lowest level of agreement.

<table>
<thead>
<tr>
<th>Party Member</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Jane</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Jim</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Joe</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Jon</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Jos</td>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>

There are 15 ordered pairs of (2,1), (3,2), (3,3), (2,5), (4,5), and (1,4). Of these, 7 are concordant, 5 are discordant, and three are ties. Therefore, $\tau = (7 - 5)/15 = 0.13$. Answer: C

5. The regression line for $x_1, x_2, \ldots x_N$ and $y_1, y_2, \ldots y_N$ is $y = 2x + 5$. If $\bar{x} = 8$, what is $\bar{y}$?

A) 20 B) 21 C) 22 D) 23 E) 24
F) 25 G) 26 H) 27 I) 28 J) 29

**Solution.** If $y = mx + b$ is a regression line for the data $x_1, x_2, \ldots x_N$ and $y_1, y_2, \ldots y_N$, then $b = \bar{y} - m\bar{x}$, or $\bar{y} = b + m\bar{x}$. Thus, $\bar{y} = 5 + 2 \cdot 8 = 21$.

Answer: B

6. The regression line for $x_1, x_2, \ldots x_N$ and $y_1, y_2, \ldots y_N$ is $y = 2x + 5$. If $s_x = 4$ and $s_Y = 11$, then what is Pearson’s correlation $r$ for the data?

A) 0.41 B) 0.45 C) 0.49 D) 0.53 E) 0.57
F) 0.61 G) 0.65 H) 0.69 I) 0.73 J) 0.77

**Solution.** If $y = mx + b$ is a regression line for the data $x_1, x_2, \ldots x_N$ and $y_1, y_2, \ldots y_N$, then $m = r \cdot s_Y/s_x$, or $r = m \cdot s_X/s_Y$. Thus, $r = 2 \cdot 4/11 = 0.73$.

Answer: I

7. The regression line for the explanatory variable $X$ and the response variable $Y$ is calculated from the observations 1, 2, 3, 6 of $X$ and 3, 5, 6, 14 of $Y$. What is the prediction $\hat{y}$ for the value of $Y$ that corresponds to the $X$-value 5? (The next four questions pertain to the data of this problem, so you might want to look ahead to see if you will need to organize any calculations that are related to the questions that follow.)

A) 11.32 B) 11.37 C) 11.43 D) 11.51 E) 11.60
F) 11.67 G) 11.73 H) 11.80 I) 11.87 J) 11.94
Solution. The regression line is \( y = (31/14) \cdot x + (5/14) \). Substituting \( x = 5 \) yields \( \hat{y} = 80/7 \approx 11.43 \).

Answer: C

8. For the linear model of problem 7, calculate the absolute values of the residuals. What is the smallest of these numbers?

A) 0.06  B) 0.11  C) 0.16  D) 0.21  E) 0.26
F) 0.31  G) 0.36  H) 0.41  I) 0.46  J) 0.51

Solution. Using the regression line stated in the solution of the preceding problem, we have \( \hat{y}_1 = (31/14) \cdot 1 + (5/14) = 18/7, \hat{y}_2 = (31/14) \cdot 2 + (5/14) = 67/14, \hat{y}_3 = (31/14) \cdot 3 + (5/14) = 7, \) and \( \hat{y}_4 = (31/14) \cdot 6 + (5/14) = 191/14 \). The residuals are \( e_1 = y_1 - \hat{y}_1 = 3 - 18/7 = 3/7, \) \( e_2 = y_2 - \hat{y}_2 = 5 - 67/14 = 3/14, \) \( e_3 = y_3 - \hat{y}_3 = 6 - 7 = -1, \) and \( e_4 = y_4 - \hat{y}_4 = 14 - 191/14 = 5/14 \). The absolute residuals are, in order, 0.42856, 0.21429, 1, 0.35714.

Answer: D

9. For the linear model of problem 7, what is SSE?

A) 1.36  B) 1.41  C) 1.46  D) 1.51  E) 1.56
F) 1.61  G) 1.66  H) 1.71  I) 1.76  J) 1.81

Solution. Using the residuals found in the preceding problem, we have \( SSE = (3/7)^2 + (3/14)^2 + (-1)^2 + (5/14)^2 = 19/14 \approx 1.357. \)

Answer: A

10. For the linear model of problem 7, what is SSR?

A) 68.40  B) 68.46  C) 68.52  D) 68.58  E) 68.64
F) 68.70  G) 68.76  H) 68.82  I) 68.88  J) 68.94

Solution. The mean of the Y values is 7. In problem 8 we calculated \( \hat{y}_1 = 18/7, \hat{y}_2 = 67/14, \hat{y}_3 = 7, \) and \( \hat{y}_4 = 191/14 \). Therefore, \( SSR = (18/7 - 7)^2 + (67/14 - 7)^2 + (7 - 7)^2 + (191/14 - 7)^2 = 961/14 \approx 68.64. \)

Answer: E

11. For the linear model of problem 7, what is the coefficient of determination?

A) 0.62  B) 0.66  C) 0.70  D) 0.74  E) 0.78
F) 0.82  G) 0.86  H) 0.90  I) 0.94  J) 0.98

Solution. The coefficient of determination is \( r^2, \) which equals 0.9806.

Answer: J

12. Constants \( C \) and \( p \) are to be found so that the generalized parabola \( y = C \cdot x^p \) fits the data 0.6, 0.7, 0.9, 1 of the explanatory variable \( X \) and 0.14, 0.29, 1.6, 3.1 of the response variable \( Y \). The procedure is to find the regression line for the data sets that are obtained by applying the natural logarithm to the original data sets. Then the generalized parabola is obtained from the regression line by inverting the transformation. What power \( p \) is obtained for the parabolic model?

A) 5.89  B) 5.94  C) 5.99  D) 6.04  E) 6.09
F) 6.14  G) 6.19  H) 6.24  I) 6.29  J) 6.34
Solution. We calculate \( \ln(X) = -0.5108256238, -0.3566749439, -0.1053605157, 0 \) and \( \ln(Y) = -1.966112856, -1.237874356, 0.4700036292, 1.131402111 \). We then calculate the regression line using these lists instead of \( X \) and \( Y \). Doing so, we find that \( m = 6.188701113 \) and \( b = 1.104541250 \). The regression line in a loglog plot is 
\[ \ln(y) = 6.188701113 \ln(x) + 1.104541250. \]
The equation obtained by equating the exponential of the left side with the exponential of the right side is 
\[ y = 3.01784 x^{6.1887}. \]

Answer: \( G \)

13. In the Monte Hall Problem, a valuable prize is behind one of three doors. A contestant picks one door. The host opens a different door, but never one that reveals the prize. (If there is a choice of two such doors to open, the choice does not matter.) The contestant is then allowed, but not required, to reject the first choice made and select the other unopened door, winning the prize if it is behind the final chosen door. Theory shows that the best strategy is to always switch. In the table that follows, forty pairs of random digits selected from \{1, 2, 3\} are given.

| 13 | 13 | 23 | 33 | 22 | 21 | 13 | 32 | 11 | 13 |
| 11 | 12 | 12 | 13 | 33 | 31 | 33 | 13 | 31 | 13 |
| 13 | 22 | 12 | 11 | 23 | 31 | 22 | 12 | 11 | 23 |
| 11 | 11 | 21 | 11 | 23 | 13 | 11 | 33 | 32 | 31 |

Each pair is used to simulate the strategy of switching doors in the Monte Hall problem. In each pair, the first digit represents the door initially chosen. The second digit represents the door behind which the prize is located. For the forty simulations, what proportion of wins results when the switching strategy is employed?

A) 21/40  B) 22/40  C) 23/40  D) 24/40  E) 25/40
F) 26/40  G) 27/40  H) 28/40  I) 29/40  J) 30/40

Solution. The strategy wins when the two numbers in the pair are different. There are 25 such pairs. (Note: Humans are not very good with random number simulations. The last line does not appear random at all to most humans. It begins with two consecutive pairs of ones and is followed by two more pairs of ones. From time to time in the past, scientists, even highly distinguished ones, have been caught forging data thanks to their inability to properly fake random variation.)

Answer: \( E \)
14. In the accompanying figure, a scatter plot and a regression line for the plotted data are shown. The point $(\bar{x}, \bar{y})$ has also been added to the plot.

![Scatter plot with regression line and points labeled α, β, γ, δ.]

Four data points are highlighted and labeled $\alpha, \beta, \gamma, \delta$. From among these four points, count the number $m$ of regression outliers and the number $n$ of high leverage points. What is $(m, n)$?

A) (1,1)  B) (1,2)  C) (1,3)  D) (2,1)  E) (2,2)  
F) (2,3)  G) (3,1)  H) (3,2)  I) (3,3)  J) (4,4)

**Solution.** There are three regression line outliers: $\alpha, \beta, \gamma$. Of these, $\alpha$ is high leverage, and $\delta$ is also.

**Answer: H**

15. This problem and the two that follow it concern three events, $E, F, G$, that are constructed from two events $A$ and $B$ in a sample space. Let $E$ be the event that exactly one of the two events $A$ and $B$ occurs, let $F$ be the event that at least one of the two events $A$ and $B$ occurs, and let $G$ be the event that neither of the events $A$ and $B$ occurs. Suppose that $P(A) = 0.4$, that $P(B) = 0.5$, and that $P(A \cap B) = 0.1$. You are to calculate $P(E)$, $P(F)$, and $P(G)$. In this problem, respond with $P(E)$.

A) 0.0  B) 0.1  C) 0.2  D) 0.3  E) 0.4  
F) 0.5  G) 0.6  H) 0.7  I) 0.8  J) 0.9

**Solution.** We have $P(E) = P(A) + P(B) - 2P(A \cap B) = 0.4 + 0.5 - 2 \times 0.1 = 0.7$.

**Answer: H**
16. Refer to problem 15. What is \( P(F) \)?

A) 0.0  B) 0.1  C) 0.2  D) 0.3  E) 0.4  
F) 0.5  G) 0.6  H) 0.7  I) 0.8  J) 0.9

**Solution.** We have

\[
P(F) = P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.4 + 0.5 - 0.1 = 0.8.
\]

**Answer:** I

17. Refer to problem 15. What is \( P(G) \)?

A) 0.0  B) 0.1  C) 0.2  D) 0.3  E) 0.4  
F) 0.5  G) 0.6  H) 0.7  I) 0.8  J) 0.9

**Solution.** We have

\[
P(G) = P((A \cup B)^c) = 1 - P(A \cup B) = 0.2.
\]

**Answer:** C

18. A faculty member has 10 advisees who are graduating seniors with a major in statistics. Seven have taken a course in Bayesian statistics and three have not. If the faculty member selects two (different) students from the group of 10 and examines their academic records, what is the probability that neither student has taken a course in Bayesian statistics?

A) 1/6  B) 1/8  C) 1/9  D) 1/10  E) 1/12  
F) 1/15  G) 1/16  H) 1/18  I) 1/20  J) 1/25

**Solution.** Let \( A_1 \) be the event that the first student chosen has not studied Bayesian statistics. Let \( A_2 \) be the event that the second student chosen has not studied Bayesian statistics. Then \( P(A_1) = \frac{3}{10} \) and \( P(A_2|A_1) = \frac{2}{9} \). Therefore

\[
P(A_1 \cap A_2) = P(A_2|A_1) \cdot P(A_1) = \frac{2}{9} \cdot \frac{3}{10} = \frac{1}{15}.
\]

**Answer:** F

19. Refer to the preceding problem. What is the probability that exactly one of the two students has taken a course in Bayesian statistics?

A) 1/4  B) 2/5  C) 3/8  D) 4/9  E) 5/12  
F) 7/15  G) 9/20  H) 10/21  I) 11/24  J) 12/25

**Solution.** Let \( B_1 \) and \( B_2 \) be the events that the first and, respectively, second, student has studied Bayesian statistics. Then

\[
P(A_1 B_2 \cup B_1 A_2) = \frac{3}{10} \cdot \frac{7}{9} + \frac{7}{9} \cdot \frac{3}{9} = \frac{42}{90} = \frac{7}{15}.
\]

**Answer:** F

20. The following relative frequency histogram displays the probability at birth of a Newfoundland (a large, intelligent Canadian dog) dying at age 0 (leftmost bar), at age 1 (next bar to the right), and so on. The relative frequencies are written above the bar. For example, the probability at birth that a Newfoundland will be alive for its first birthday but not for its second is 0.072.
What is the probability at birth that a Newfoundland will live at least 13 years?

A) 0.010  B) 0.015  C) 0.020  D) 0.025  E) 0.030  
F) 0.035  G) 0.040  H) 0.045  I) 0.050  J) 0.055

Solution. Let \( E \) be the event that a Newfoundland lives to 13. Let \( E_j \) be the event that a Newfoundland dies at age \( j \). Then \( E = E_{13} \cup E_{14} \cup E_{15} \cup E_{16} \). Because the sets on the right side are disjoint, we have

\[
P(E) = P(E_{13}) + P(E_{14}) + P(E_{15}) + P(E_{16}),
\]

or

\[
P(E) = 0.015 + 0.005 + 0.004 + 0.001 = 0.025.
\]

Answer: \( D \)

21. Refer to the data in the preceding problem. Given that a Newfoundland has just turned 10 years old (Happy birthday, Max!), what is the probability that it will live at least 13 years?

A) 0.11  B) 0.14  C) 0.17  D) 0.20  E) 0.23  
F) 0.26  G) 0.29  H) 0.32  I) 0.35  J) 0.38

Solution. Let \( F \) be the event that a Newfoundland has just turned 10. Let \( E \) and \( E_j \) be as in the preceding solution. Then

\[
P(F) = P(E_{10}) + P(E_{11}) + P(E_{12}) + P(E_{13}) + P(E_{14}) + P(E_{15}) + P(E_{16}),
\]

or

\[
P(F) = 0.041 + 0.027 + 0.016 + 0.015 + 0.005 + 0.004 + 0.001 = 0.109.
\]

Next, observe that \( E \subset F \), so \( E \cap F = E \). Therefore, \( P(E \cap F) = P(F \cap E) = 0.025 \). Finally, the problem asks for \( P(E|F) \), which is

\[
P(E \cap F)/P(F), \text{ or } 0.025/0.109, \text{ or } 0.229.
\]

Answer: \( E \)

22. In a Math 2200 class, 37% of the students have studied Bio 296 and 59% have studied Math 233. If a student is randomly selected from the group of students who have studied Math 233, the probability that that student has studied Bio 296 is 0.39. Here is the question you are to answer: If a student is selected at random from the entire class, what is the probability that the student has studied Bio 296 but not Math 233?

A) 0.14  B) 0.16  C) 0.18  D) 0.20  E) 0.22  
F) 0.24  G) 0.26  H) 0.28  I) 0.30  J) 0.32
Solution. Let $B$ be the event that a randomly selected student class has taken Biology 296. Let $M$ be the event that a randomly selected student has taken Math 233. We are given $P(B) = 0.37$ and $P(M) = 0.59$. We are also given $P(B \mid M) = 0.39$. It follows that

$$P(B \cap M) = P(B \mid M)P(M) = (0.39)(0.59) = 0.2301,$$

and

$$P(B \setminus M) = P(B) - P(B \cap M) = 0.37 - 0.2301 = 0.14.$$

Answer: A

23. Air traffic controllers are subject to random drug testing. An inexpensive, easily performed urine test is used as the initial screening for performance-impairing drugs such as amphetamines, barbiturates, cannabinoids, and opiates. The sensitivity of this test is 0.96 and its specificity is 0.93. The prevalence of drug use among air traffic controllers is 0.7%. What is the probability that an air traffic controller who tests positive actually has used performance-impairing drugs?

A) 0.04  B) 0.07  C) 0.09  D) 0.14  E) 0.17  
F) 0.79  G) 0.84  H) 0.89  I) 0.93  J) 0.96

Solution. Let $S$ denote the event that an air traffic controller has used a performance-impairing drug. We have

$$P(S \mid POS) = \frac{\text{sensitivity} \cdot \text{prevalence}}{\text{sensitivity} \cdot \text{prevalence} + (1 - \text{specificity})(1 - \text{prevalence})} = \frac{0.96 \cdot 0.007}{0.96 \cdot 0.007 + (1 - 0.93)(1 - 0.007)} = 0.08815.$$

Answer: C

24. A fair coin is tossed twice. If the result of each toss was a head, then a pair of dice is rolled. If the sum of the dots is 10 or more, then $X = 10$, but if the sum of the dots is less than 10, then $X = 1$. In the event that one or two tails resulted from the coin tosses, a random digit $k$ is selected from a table of random digits and $X = k$. What is $E(X)$?

A) 3.75  B) 3.79  C) 3.83  D) 3.875  E) 3.92  
F) 3.96  G) 4.00  H) 4.04  I) 4.08  J) 4.12

Solution. We begin by calculating the probability function $f_X$. The values that $X$ can assume are 0, 1 (in two ways), 2, 3, 4, 5, 6, 7, 8, 9, and 10. Let $E$ be the event of two heads on the coin toss, $E^c$ the event of one or two tails, $F$ the event of a 10 or more on the dice roll, and $F^c$ the event of less than a 10. Then $P(E) = (1/2)^2 = 1/4$, $P(E^c) = 1 - 1/4 = 3/4$, $P(F) = P(\{ (4,6), (5,5), (6,4), (5,6), (6,5), (6,6) \}) = 6/36 = 1/6$, and $P(F^c) = 1 - 1/6 = 5/6$. Finally, if $G_k$ is the event of a random digit $k$, then $P(G_k) = 1/10$. From all this, it follows that

$$f_X(0) = P(E^c)P(G_9) = \frac{3}{4} \cdot \frac{1}{10} = \frac{1}{40},$$

$$f_X(1) = P(E) \cdot P(F) + P(E^c) \cdot P(G_1) = \frac{1}{4} \cdot \frac{1}{6} + \frac{3}{4} \cdot \frac{1}{10} = \frac{17}{60},$$

$$f_X(2) = f_X(3) = f_X(4) = f_X(5) = f_X(6) = f_X(7) = f_X(8) = f_X(9) = \frac{3}{4} \cdot \frac{1}{10} = \frac{3}{40},$$

and

$$f_X(10) = P(E) \cdot P(F) = \frac{1}{4} \cdot \frac{1}{6} = \frac{1}{24}.$$
As a check, we add all these numbers—we better get 1:
\[
\frac{3}{40} + \frac{17}{60} + 8 \cdot \frac{3}{40} + \frac{1}{24} = 1.
\]

Good. We are ready to calculate \(E(X)\):
\[
E(X) = \sum_{k=0}^{10} k \cdot f_X(k) = \frac{3}{40}(0 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9) + \frac{17}{60} \cdot 1 + \frac{1}{24} \cdot 10 = 4.
\]

Answer: \(G\)

25. One die is rolled. It appears to be an ordinary six-faced die with the usual number of dots, but it is “loaded”. The outcomes of the roll that are less than 6 are equally likely to occur, but the probability of a 6 is 1/2. If \(X\) is the number of dots of the roll, then what is the variance of \(X\)?

A) 1.00  B) 1.25  C) 1.50  D) 1.75  E) 2.00
F) 2.25  G) 2.50  H) 2.75  I) 3.00  J) 3.25

Solution. Because \(P(X \in \{1, 2, 3, 4, 5\}) = 1 - P(X = 6) = 1 - 1/2 = 1/2\), and because \(P(X = 1) = P(X = 2) = P(X = 3) = P(X = 4) = P(X = 5)\), each of these probabilities must be \((1/2)/5\), or 1/10. Therefore, the probability function for \(X\) is
\[
f_X(k) = \frac{1}{10}\text{ for } 1 \leq k \leq 5 \text{ and } f_X(6) = \frac{1}{2}.
\]
The mean is
\[
\mu_X = 1 \cdot f_X(1) + 2 \cdot f_X(2) + 3 \cdot f_X(3) + 4 \cdot f_X(4) + 5 \cdot f_X(5) + 6 \cdot f_X(6) = \frac{1}{10}(1 + 2 + 3 + 4 + 5) + \frac{1}{2} \cdot 6 = \frac{9}{2}.
\]
The variance is
\[
\sigma^2_X = \left(1 - \frac{9}{2}\right)^2 \cdot f_X(1) + \left(2 - \frac{9}{2}\right)^2 \cdot f_X(2) + \left(3 - \frac{9}{2}\right)^2 \cdot f_X(3) + \left(4 - \frac{9}{2}\right)^2 \cdot f_X(4) + \left(5 - \frac{9}{2}\right)^2 \cdot f_X(5) + \left(6 - \frac{9}{2}\right)^2 \cdot f_X(6)
\]
\[
= \frac{13}{4}
\]
\[
= 3.25.
\]

Answer: \(J\)