1. In a cybersecurity scenario, if two security breaches A and B are dependent, then

   (a) $P(A \cap B) = P(A) + P(B) - P(A \cup B)$
   (b) $P(A \cap B) = P(A) + P(B)$
   (c) $P(A \cap B) = P(A) \times P(B \mid A)$
   (d) $P(A \cap B) = P(B) \times P(A \mid B)$

   Solution: The correct formula for the probability of the intersection of two dependent events A and B is $P(A \cap B) = P(A) \times P(B \mid A)$.

2. Which of the following statements are correct in the context of digital privacy laws?

   (a) If the data privacy laws are independent then they are also mutually exclusive.
   (b) If the data privacy laws are mutually exclusive then they are dependent.
   (c) If the data privacy laws are independent, then they cannot be mutually exclusive.
   (d) If the data privacy laws are mutually exclusive then they are independent.

   Solution: The correct statement is: If the data privacy laws are independent, then they cannot be mutually exclusive.

3. In an IT company, the number of software bugs reported hourly follows the probability distribution in Table 1.

Table 1: Hourly Occurrence of Software Bugs

Number of bugs (X)	0	1	2	3	4	5	6
Probability [P(X)]	0.15	0.20	0.25	0.20	0.15	0.04	0.01

   Calculate

   (a) the expected number of bugs X reported per hour,
   (b) the variance, and
   (c) the standard deviation for the discrete random variable.

   Solution:

(a) Expected value:

$$\begin{aligned}
E(X) &= \sum_{i=0}^{6} X_i \cdot P(X_i) \\
&= (0)(0.15) + (1)(0.20) + (2)(0.25) + (3)(0.20) + (4)(0.15) + (5)(0.04) + (6)(0.01) \\
&= 0 + 0.20 + 0.50 + 0.60 + 0.60 + 0.20 + 0.06 \\
&= 2.16 \text{ bugs per hour}
\end{aligned}$$

(b) Variance:

$$\begin{aligned}
\mathrm{Var}(X) &= \sum_{i=0}^{6} (X_i - E(X))^2 \cdot P(X_i) \\
&= (0 - 2.16)^2(0.15) + (1 - 2.16)^2(0.20) + \ldots + (6 - 2.16)^2(0.01) \\
&= 0.7064
\end{aligned}$$

(c) Standard deviation:

$$\mathrm{SD}(X) = \sqrt{\mathrm{Var}(X)} \approx 0.8408$$

4. According to a survey on cybercrime activities, the time taken by IT experts to resolve a security breach follows a normal distribution with a mean of 45 minutes and a standard deviation of 8 minutes. Assume normality.

   (a) What proportion of IT experts take more than 60 minutes to resolve a security breach?

   (b) What proportion of security breaches are resolved in less than 30 minutes?

   (c) What proportion of security breaches are resolved between 40 minutes and 50 minutes?

   Solution: Let X be the time taken to resolve a security breach.

   (a) $P(X > 60) = 1 - P(X \leq 60) = 1 - \Phi\left(\frac{60-45}{8}\right) = 1 - \Phi(1.875) \approx 0.0301$

   (b) $P(X < 30) = \Phi\left(\frac{30-45}{8}\right) = \Phi(-1.875) \approx 0.0301$

   (c) $P(40 < X < 50) = P(X < 50) - P(X < 40) = \Phi\left(\frac{50-45}{8}\right) - \Phi\left(\frac{40-45}{8}\right) = \Phi(0.625) - \Phi(-0.625) \approx 0.2659$

5. Estimated value of parameter from a sample in IT context is called ...

   (a) Sample statistic

   (b) Population parameter

   (c) Unbiased estimator

   (d) Sampling

Solution: Sample statistic

6. In a random sampling for digital evidence, the sample subjects are selected from the dataset using ...

   (a) Normal distribution
   (b) Uniform distribution
   (c) Binomial distribution
   (d) Random distribution

   Solution: Random distribution

7. An estimate of population parameter in IT context is unbiased when ...

   (a) It has minimum variance
   (b) It has minimum error
   (c) Its expected value is same as the population parameter
   (d) It has minimum standard deviation

   Solution: Its expected value is same as the population parameter

8. In an IT security audit, the time taken by employees to log into the system (in seconds) is given in Table 2. Calculate the 95% confidence interval for the average login time.

Table 2: Login Times (in seconds)

12	16	20	9	25	18	22	27	14	8
32	28	21	17	19	23	26	30	30	24
11	10	40	9	12	38	8	21	25	26

Solution: The sample mean, $\bar{x} = 20.45$ seconds
The sample standard deviation, $s = 8.847$ seconds
The sample size, $n = 30$
The 95% confidence interval for the population mean, μ, is given by:

$$\bar{x} \pm t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)$$

From the t-distribution table with $df = n - 1 = 29$ and $\alpha/2 = 0.025$, we find $t_{\alpha/2} \approx 2.045$.

So, the confidence interval is approximately:

$$20.45 \pm 2.045\left(\frac{8.847}{\sqrt{30}}\right)$$

$$20.45 \pm 3.196$$

Thus, the 95% confidence interval for the average login time is approximately $[17.254, 23.646]$ seconds.

9. The data shown below describes the processing times for legal documents between 2010 and 2020. The times measured in hours were obtained as follows:

Table 3: Document Processing Times (in hours)

3.5	4.2	5.1	6.2	4.8	5.3	4.9	4.7	4.6	5.5	6.1

Assume that the standard deviation is known to be $\sigma = 0.7$.

(a) Construct a 99% two-sided confidence interval on the mean processing time.

(b) Construct a 95% lower-confidence bound on the mean processing time.

(c) Suppose that we wanted to be 95% confident that the error in estimating the mean processing time is less than 1 hour. What sample size should be used?

Solution: Given: $n = 11$, $\sigma = 0.7$.

(a) 99% two-sided confidence interval:

$$\text{CI} = \bar{x} \pm z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

From the normal distribution table, for 99% confidence level, $z_{\alpha/2} = 2.576$.

$$\text{CI} = \bar{x} \pm 2.576\left(\frac{0.7}{\sqrt{11}}\right)$$

$$\text{CI} = \bar{x} \pm 0.593$$

So, the 99% confidence interval is approximately $[4.907, 5.740]$ hours.

(b) 95% lower-confidence bound:

$$\text{Lower bound} = \bar{x} - z_{\alpha}\left(\frac{\sigma}{\sqrt{n}}\right)$$

From the normal distribution table, for 95% confidence level, $z_\alpha = 1.645$.

$$\text{Lower bound} = \bar{x} - 1.645\left(\frac{0.7}{\sqrt{11}}\right)$$

$$\text{Lower bound} = \bar{x} - 0.383$$

So, the 95% lower-confidence bound is approximately $[4.117, \infty)$ hours.

(c) Sample size needed for 95% confidence level with error less than 1 hour:

$$n = \left(\frac{z_{\alpha/2} \cdot \sigma}{\text{error}}\right)^2$$

Given: $\sigma = 0.7$, error $= 1$. From the normal distribution table, for 95% confidence level, $z_{\alpha/2} = 1.960$.

$$n = \left(\frac{1.960 \cdot 0.7}{1}\right)^2$$

$$n \approx 11.38$$

So, a sample size of at least 12 should be used.

10. In a legal case, the number of hours spent in court by lawyers follows a Poisson distribution with a mean of 10 hours per day. Calculate the probability of:

    (a) Exactly 8 hours spent in court by lawyers in a day.

    (b) At most 5 hours spent in court by lawyers in a day.

    (c) More than 12 hours spent in court by lawyers in a day.

    Solution: Let X be the number of hours spent in court by lawyers.

    (a) $P(X = 8) = \frac{e^{-10} \cdot 10^8}{8!} \approx 0.1126$

    (b) $P(X \leq 5) = \sum_{x=0}^{5} \frac{e^{-10} \cdot 10^x}{x!} \approx 0.0671$

    (c) $P(X > 12) = 1 - P(X \leq 12) = 1 - \sum_{x=0}^{12} \frac{e^{-10} \cdot 10^x}{x!} \approx 0.2034$

11. In an IT department, the number of successful login attempts per hour follows a binomial distribution with $n = 20$ and $p = 0.8$. Calculate:

    (a) The probability of exactly 15 successful login attempts in an hour.

    (b) The probability of at least 18 successful login attempts in an hour.

    (c) The expected number of successful login attempts in an hour.

    (d) The variance and standard deviation of successful login attempts in an hour.

    Solution: Let X be the number of successful login attempts per hour.

    (a) $P(X = 15) = \binom{20}{15}(0.8)^{15}(0.2)^5 \approx 0.1026$

    (b) $P(X \geq 18) = 1 - P(X < 18) = 1 - \sum_{x=0}^{17} \binom{20}{x}(0.8)^x(0.2)^{20-x} \approx 0.8222$

    (c) Expected value: $E(X) = np = 20 \times 0.8 = 16$

    (d) Variance: $\text{Var}(X) = np(1-p) = 20 \times 0.8 \times 0.2 = 3.2$
    Standard deviation: $\text{SD}(X) = \sqrt{\text{Var}(X)} = \sqrt{3.2} \approx 1.79$

12. According to IT regulations, the proportion of websites compliant with security standards is estimated to be 0.75. In a random sample of 100 websites, calculate:

    (a) The probability that exactly 80 websites are compliant with security standards.

    (b) The probability that at most 70 websites are compliant with security standards.

    (c) The average number of compliant websites in the sample.

    (d) The variance and standard deviation of the number of compliant websites in the sample.

    Solution: Let X be the number of compliant websites in the sample.

    (a) $P(X = 80) = \binom{100}{80}(0.75)^{80}(0.25)^{20} \approx 0.0235$

    (b) $P(X \leq 70) = \sum_{x=0}^{70} \binom{100}{x}(0.75)^{x}(0.25)^{100-x} \approx 0.0000$

    (c) Expected value: $E(X) = np = 100 \times 0.75 = 75$

    (d) Variance: $\text{Var}(X) = np(1-p) = 100 \times 0.75 \times 0.25 = 18.75$
    Standard deviation: $\text{SD}(X) = \sqrt{\text{Var}(X)} = \sqrt{18.75} \approx 4.33$

13. In an e-discovery process, the time taken to review documents follows a normal distribution with a mean of 120 minutes and a standard deviation of 25 minutes. Assuming normality:

    (a) What proportion of reviewers spend more than 150 minutes on document review?

    (b) What proportion of reviewers spend less than 90 minutes on document review?

    (c) What proportion of reviewers spend between 100 minutes and 140 minutes on document review?

    Solution: Let X be the time taken to review documents.

    (a) $P(X > 150) = 1 - P(X \leq 150) = 1 - \Phi\left(\frac{150-120}{25}\right) = 1 - \Phi(1.2) \approx 0.1151$

    (b) $P(X < 90) = \Phi\left(\frac{90-120}{25}\right) = \Phi(-1.2) \approx 0.1151$

    (c) $P(100 < X < 140) = P(X < 140) - P(X < 100) = \Phi\left(\frac{140-120}{25}\right) - \Phi\left(\frac{100-120}{25}\right) = \Phi(0.8) - \Phi(-0.8) \approx 0.4683$

14. Estimated value of parameter from a sample in legal context is called ...

    (a) Sample statistic

    (b) Population parameter

    (c) Unbiased estimator

    (d) Sampling

Solution: Sample statistic

15. In a random sampling for legal research, the sample cases are selected from the court database using ...

    (a) Normal distribution
    (b) Uniform distribution
    (c) Binomial distribution
    (d) Random distribution

    Solution: Random distribution

16. An estimate of population parameter in legal context is unbiased when ...

    (a) It has minimum variance
    (b) It has minimum error
    (c) Its expected value is same as the population parameter
    (d) It has minimum standard deviation

    Solution: Its expected value is same as the population parameter

17. In a litigation case, the time taken by different parties to submit evidence (in days) is given in Table 4. Calculate the 95% confidence interval for the average time taken to submit evidence.

Table 4: Time Taken to Submit Evidence (in days)

14	20	22	9	25	18	26	30	16	21
32	28	21	17	19	23	26	30	30	24
11	10	40	9	12	38	8	21	25	26

Solution: The sample mean, $\bar{x} = 21.75$ days
The sample standard deviation, $s = 7.671$ days
The sample size, $n = 30$
The 95% confidence interval for the population mean, μ, is given by:

$$\bar{x} \pm t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)$$

From the t-distribution table with $df = n - 1 = 29$ and $\alpha/2 = 0.025$, we find $t_{\alpha/2} \approx 2.045$.

So, the confidence interval is approximately:

$$21.75 \pm 2.045\left(\frac{7.671}{\sqrt{30}}\right)$$

$$21.75 \pm 2.051$$

Thus, the 95% confidence interval for the average time taken to submit evidence is approximately $[19.699, 23.801]$ days.

18. The data shown below describe temperature readings for sorghum grown at Mantey Farms between 1982 and 1993. The temperatures measured in June were obtained as follows:

Table 5: June temperature readings

15.2	14.2	14.0	12.2	14.4	12.5	14.3	14.2	13.5	11.8	15.2

Assume that the standard deviation is known to be $\sigma = 0.5$.

(a) Construct a 99 % two-sided confidence interval on the mean temperature.

(b) Construct a 95 % lower-confidence bound on the mean temperature.

(c) Suppose that we wanted to be 95 % confident that the error in estimating the mean temperature is less than 2 degrees Celsius. What sample size should be used?

Solution: Given: $n = 11$, $\sigma = 0.5$.

(a) 99% two-sided confidence interval:

$$\text{CI} = \bar{x} \pm z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

From the normal distribution table, for 99% confidence level, $z_{\alpha/2} = 2.576$.

$$\text{CI} = \bar{x} \pm 2.576\left(\frac{0.5}{\sqrt{11}}\right)$$

$$\text{CI} = \bar{x} \pm 0.387$$

So, the 99% confidence interval is approximately $[13.483, 14.907]$ degrees Celsius.

(b) 95% lower-confidence bound:

$$\text{Lower bound} = \bar{x} - z_{\alpha}\left(\frac{\sigma}{\sqrt{n}}\right)$$

From the normal distribution table, for 95% confidence level, $z_{\alpha} = 1.645$.

$$\text{Lower bound} = \bar{x} - 1.645\left(\frac{0.5}{\sqrt{11}}\right)$$

$$\text{Lower bound} = \bar{x} - 0.246$$

So, the 95% lower-confidence bound is approximately $[13.054, \infty)$ degrees Celsius.

(c) Sample size needed for 95% confidence level with error less than 2 degrees Celsius:

$$n = \left(\frac{z_{\alpha/2} \cdot \sigma}{\text{error}}\right)^2$$

Given: $\sigma = 0.5$, error $= 2$. From the normal distribution table, for 95% confidence level, $z_{\alpha/2} = 1.960$.

$$n = \left(\frac{1.960 \cdot 0.5}{2}\right)^2$$

$$n \approx 24.01$$

So, a sample
size of at least 25 should be used.