Statistics Questions

Thursday, November 2, 2023 6:38 PM

STATISTICS

Miscellaneous Distributions & CLT

Statistics Questions

SECTION-(A) MULTIPLE CHOICE QUESTIONS (MCQ)

1. An unknown distribution has a mean of 90 and a standard deviation of 15. Samples of size n = 25 are drawn randomly from the population. Find the probability that the **sample mean** is between 85 and 92.

   (A) 0.3546 (B) 0.6997
   (C) 0.1254 (D) 0.4521

2. A produce company claims that the mean weight of peaches in a large shipment is 6.0 oz with a standard deviation of 1.0 oz. Assuming this claim is true, what is the probability that a random sample of 1000 of these peaches would have a mean weight of 5.9 oz or less?

   (A) 0.0008 (B) 0.654
   (C) 3.16 (D) 6.0

3. Let $Y = X_1 + X_2 + \ldots + X_{15}$ be the sum of a random sample of size 15 from the distribution whose density function is

$$f(x) = \begin{cases} \frac{3}{2}x^2 & \text{if } -1 < x < 1 \\ 0 & \text{otherwise.} \end{cases}$$

   What is the approximate value of $P(-0.3 \le Y \le 1.5)$ when one use the central limit theorem?

   (A) 0.2131 (B) 0.5214
   (C) 0.2313 (D) None of the above

4. Let X_1, X_2, ..., X_n be a random sample of size n = 25 from a population that has a mean $\mu = 71.43$ and variance $\sigma^2 = 56.25$. Let $\bar{X}$ be the sample mean. What is the probability that the sample mean is between 68.91 and 71.97?

(A) 0.3654 (B) 0.21465

(C) 0.3654 (D) 0.5941

5. Light bulbs are installed successively into a socket. If we assume that each light bulb has a mean life of 2 months with a standard deviation on 0.25 months, what is the probability that 40 bulbs last at least 7 years?

(A) 0.0057 (B) 0.57

(C) 0.057 (D) None of the above

6. Light bulbs are installed into a socket. Assume that each has a mean life of 2 months with standard deviation of 0.25 month. How many bulbs n should be bought so that one can be 95% sure that the supply of n bulbs will last 5 years?

(A) 31 (B) 31.15

(C) 31.10 (D) 31.5

7. American Airlines claims that the average number of people who pay for in-flight moves, when the plane is fully loaded, is 42 with a standard deviation of 8. A sample of 36 fully loaded planes is taken. What is the probability that fewer than 38 people paid for the in-flight moves?

(A) 0.0013 (B) 0.125

(C) 0.012 (D) 0.2311

8. For a geometric distribution with $f(x) = \frac{1}{2^x}$, x = 1, 2, ..., using Chebyshev's inequality which of the following is correct?

(A) $P(| X + 2 | \le 2) \ge \frac{1}{2}$ (B) $P(| X - 2 | \le 2) \ge \frac{1}{2}$

(C) $P(| X - 2 | > 2) \ge \frac{1}{2}$ (D) $P(| X - 2 | < 2) \ge \frac{1}{2}$

9. Find the least value of probability $P(1 \le X \le 7)$ where X is a random variable with E(X) = 4 and V(X) = 4.

(A) 1/9 (B) 2/9

(C) 4/9 (D) 5/9

10 If X ~ B(100, 0.5), using Chebyshev's Inequality obtain the lower bound for, $P(| X - 50 | < 7.5)$

(A) 0.56 (B) 0.32

(C) 0.21 (D) 0.65

11. The heights of 18-year-old men are approximately normally distributed with mean 68 inches and standard deviation 3 inches. What is the probability that a randomly selected 18-years-old man is between 67 and 69 inches tall.

(A) 0.2322 (B) 0.2365

(C) 0.2586 (D) 0.3212

standard deviation 3 inches. What is the probability that a randomly selected 18-years-old man is between 67 and 69 inches tall.

(A) 0.2322 (B) 0.2365

(C) 0.2586 (D) 0.3212

12. Suppose that taxi and takeoff time for commercial jets is a random variable x with a mean of 8.5 minutes and a standard deviation of 2.5 minutes. What is the probability that for 36 jets on a given runway total taxi and takeoff time will be less than 320 minutes?

(A) 0.2315 (B) 0.8238

(C) 0.3256 (D) 0.3155

full $5t = 7.5$, $t = 1.5$ $P(|x-50| \le 5t) \ge 1 - 1/t^2$

$P|x-50| < 5(1.5) > 1 - \frac{1}{(1.5)^2}$

$P|x-50| < 7.5 \ge 0.56$

Lower Bound is 0.56

$n \ge 30$ CLT ∴ $S_{\bar{x}} = \sigma/\sqrt{36} = 0.4167$

Convert 320 → $\frac{320}{36} = 8.89$

Then $P(\bar{x} \le 8.89) = P\left(z \le \frac{8.89 - 8.5}{0.4167}\right) = P(z \le 0.93)$

$= 0.8238$

13. The Central Limit Theorem states that :

(A) if n is large then the distribution of the sample can be approximated closely by a normal curve

(B) if n is large, and if the population is normal, then the variance of the sample mean must be small

(C) if n is large, then the sampling distribution of the sample mean can be approximately closed by a normal curve

(D) if n is large, and if the population is normal, then the sampling distribution of the sample mean can be approximated closely by a normal curve

14. The central limit theorem tells us that the sampling distribution of the sample mean is approximately normal. Which of the following conditions are necessary for the theorem to be valid ?

(A) The sample size has to be large

(B) We have to be sampling from a normal population

(C) The population has to be symmetric

(D) Population variance has to be small

15. The Central Limit Theorem is important in Statistics because it allows us to use the normal distribution to make inferences concerning the population mean:

(A) provided that the population is normally distributed and the sample size is reasonably large

(B) provided that the population is normally distributed (for any sample size)

(C) provided that the sample size is reasonably large (for any sample size)

(D) None of these

16. The Central Limit Theorem is important in Statistics because :

(A) it tells us that large samples do not need to be selected

(B) it guarantees that, when it applies, the sample that are drawn are always randomly selected

(C) it enables reasonably accurate probabilities to be determined for events involving the sample average when the sample size is large regardless of the distribution of the variable

(D) it tells us that if several samples have product sample averages which seem to be different than expected, the next sample average will likely be close to its expected value

17. A symmetric die is thrown 600 times. Find the lower bound for the probability of getting 80 to 120 sixes.

(A) 19/20 (B) 19/23

(C) 19/24 (D) 24/19

18. Use Chebychev's inequality to determine how many times a fair coin must be tossed in order that the probability will be at least 0·90 that the ratio of the observed number of heads to the number of tosses will lie between 0·4 and 0·6.

(A) 300 (B) 250

(C) 245 (D) 255

19. Let $X_1, X_2, \ldots, X_n$ be i.i.d. variables with mean m and variance σ^2 and as $n \to \infty$,

$$(X_1^2 + X_2^2 + \ldots + X_n^2)/n \xrightarrow{P} c,$$

for some constant c; $(0 \le c \le \infty)$. Find c.

(A) $c = \sigma^2 + \mu^2$ (B) $c = \sigma^2$

(C) $c = \sigma^2 - \mu^2$ (D) $c = \mu^2$

20. How large a sample must be taken in order that the probability will be at least 0·95 that $\overline{X}_n$ will be lie within 0·5 of μ.μ is unknown and σ = 1.

(A) n = 80 (B) n < 80

(C) n ≥ 80 (D) n > 70

21. For which of the following the law of large numbers can be applied to the independent variables $X_1, X_2, \ldots$, i.e., X_i's.

(A) if X_i assume that values i and –i with equal probabilities

(B) If X_i can have only two values with equal probabilities i^{α} and $-i^{\alpha}$, if $\alpha < \frac{1}{2}$

(C) if X_i assume that values i and –i with unequal probabilities

(D) None of the above

22. Let $\{X_k\}$ be mutually independent and identically distributed random variables with mean m and finite variance. If $S_n = X_1 + X_2 + \ldots + X_n$ then which of the following holds?

(A) the law of large numbers does not hold for the sequence $\{S_n\}$

(B) the law of large numbers holds for the sequence $\{S_n\}$

(C) it satisfies the central limit theorem

(D) None of the above

23. The sequence $\{X_k\}$ of independent random variables defined as follows:

$$P[X_k = \pm 2^k] = 2^{-(2k+1)}$$

$$P[X_k = 0] = 1 - 2^{-2k}$$

then which of the following statement holds?

$P[X_k = \pm 2^k] = 2^{-(2k+1)}$

$P[X_k = 0] = 1 - 2^{-2k}$

then which of the following statement holds?

(A) (Weak) Law of large numbers, does not holds for the sequence of independent r.v.'s $\{X_k\}$

(B) (Weak) Law of large numbers, holds for the sequence of independent r.v.'s $\{X_k\}$

(C) (Strong) of large numbers, holds for the sequence of independent r.v.'s $\{X_k\}$

(D) None of the above

24. Let $X_1, X_2, \ldots, X_n$ be jointly normal with $E(X_i) = 0$ and $E(X_i^2) = 1$ for all i and $Cov(X_i, X_j) = \rho$ if $|j - i| = 1$ and 0 otherwise, then which of the following statement holds ?

(A) WLLN does not hold for the sequence $|X_n|$

(B) SLLN holds for the sequence $|X_n|$

(C) WLLN holds for the sequence $|X_n|$

(D) None of the above

25. Suppose that some distribution with an unknown mean has variance equal to 1. How large a sample must be taken in order that the probability will be at least .95 that the sample mean $\bar{X}_n$ will lie within .5 of the population mean?

(A) 20 (B) 63

(C) 51 (D) 80

26. Let $\{X_n, n \geq 1\}$ be i.i.d. N(0, 1) then

$$\cos\left(\frac{X_1^4 + X_2^4 + \ldots + X_n^4}{n}\right)$$

as $n \to \infty$?

(A) cos(2) (B) cos(5)

(C) cos(3) (D) None of the above

27. Let $\{X_n, n \geq 1\}$ be i.i.d. with mean μ and finite variance σ^2. Let

$$\mu_n := \frac{X_1 + \ldots + X_n}{n} \text{ and } \sigma_n^2 = \frac{(X_1 - \mu_n)^2 + \ldots + (X_n - \mu_n)^2}{n}.$$

then

(A) $\sigma_n^2 \to_{a.s.} \sigma$ as $n \to \infty$ (B) $\sigma_n^2 \to_{a.s.} \sigma^2$ as $n \to \infty$

(C) $\sigma_n \to_{a.s.} \sigma^2$ as $n \to \infty$ (D) None of the above

28. Let $\{X_n, n \geq 1\}$ be a sequence of independent Bernoulli random variables with mean $p \in (0, 1)$. We construct an estimate $\hat{p}_n$ of p from $\{X_1, ..., X_n\}$. We know that $p \in (0.4, 0.6)$. Find the smallest value of n so that

$$P\left(\frac{|\hat{p}_n - p|}{p} \leq 5\%\right) \geq 95\%.$$

(A) 2000 (B) 2400

(C) 2600 (D) 2500

29. Let $\{\varepsilon_n\}$ be an iid sequence. A generic term of the sequence has mean μ and variance σ^2. Let $\{X_n\}$ be a covariance stationary sequence such that a generic term of the sequence satisfies

$$X_n = \rho X_{n-1} + \varepsilon_n$$

where $-1 < \rho < 1$. Denote by

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_n$$

the sample mean of the sequence. Verify whether the sequence $\{X_n\}$ satisfies the conditions that are required by Chebyshev's Weak Law of Large Numbers. In affirmative case, find its probability limit.

(A) $\frac{\mu}{1-\rho}$ (B) $\frac{\mu}{1+\rho}$

(C) $\frac{2\mu}{1-\rho}$ (D) None of the above

30. Suppose X is B(100; 0.3). Compute $P(X > 35)$.

(A) 0.2254 (B) 0.1379

(C) 0.3214 (D) 0.1236

SECTION-(B) MULTIPLE SELECT QUESTIONS (MSQ)

1. Suppose the number X of items produced by a factory in a week is a random variable with mean 50, and variance 25.

Then which of the following are correct ?

(A) an upper bound for $P(X \geq 25) = 0.75$

(B) an upper bound for $P(X \geq 75) = \frac{2}{3}$

(C) a lower bound for $P(40 < X < 60) = 0.75$

(D) a lower bound for $P(40 < X < 60) = \frac{2}{3}$

2. Let X be a r.v. with cdf $F_X(x)$ and pdf $f_X(x)$. Let $Y = aX + b$, where a and b are real contents and $a \neq 0$.

Then which of the following holds ?

(A) the cdf of Y in terms of $F_X(x)$, is $F_Y(y) = 1 - F_X\left(\frac{y-b}{a}\right)$ $a < 0$

(B) the pdf of Y in terms of $f_X(x)$, is $F_Y(y) = 1 - F_X\left(\frac{y-b}{a}\right)$ $a < 0$

(C) the pdf of Y in terms of $f_X(x)$, is $f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)$

(D) the cdf of Y in terms of $F_X(x)$, is $f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)$

3. Let Y = X^2 then the pdf of Y if X = N(0; 1). will be

(A) $\frac{1}{\sqrt{y}} f_X(\sqrt{y}) = \frac{1}{\sqrt{2\pi y}} e^{-y/2}$ $y > 0$

(B) 0 $y < 0$

(C) $\frac{1}{\sqrt{y}} f_X(\sqrt{y}) = \frac{1}{\sqrt{2\pi y}} e^{-y/2}$ $y < 0$

(D) $\frac{1}{\sqrt{y}} f_X(\sqrt{y}) = \frac{1}{\sqrt{2\pi y}} e^{-y/2}$ $y = 0$

4. Let the moment of a discrete r.v. X be given by

$$E(X^k) = 0.8 \quad k = 1, 2, \ldots$$

(A) P(X = 1) = 0.2

(B) P(X = 0) = 0.2

(C) P(X = 1) = 0.8

(D) P(X = 0) = 0.8

5. Light bulbs are installed into a socket. Assume that each has a mean life of 2 months with standard deviation of 0.25 month. How many bulbs n should be bought so that one can be 95% sure that the supply of n bulbs will last 5 years ?

(A) 36

(B) 32

(C) 30

(D) 31

6. The ounce of fill from the bottling machine are assumed to have a normal distribution with $\sigma^2 = 1$. Suppose that we plan to select a random sample of ten bottles and measure the amount of fill in each bottle. If these ten observations are used to calculate S^2, it might be useful to specify an interval of values that will include S^2 with a high probability. Then which of the following are correct?

(A) $b_1 = 0.369$ (B) $b_2 = 1.880$

(C) $b_1 = 1.369$ (D) $b_2 = 1.008$

7. Let X_1 and X_2 be a random sample of size 2 from a distribution with probability density function

$$f(x) = \begin{cases} 6x(1-x) & \text{if } 0 < x < 1 \\ 0 & \text{otherwise.} \end{cases}$$

for the sample sum $Y = X_1 + X_2$? which of the following are correct?

(A) mean (Y) = 0.5 (B) Var (Y) = 0.05

(C) Var (Y) = 0.1 (D) mean (Y) = 1

8. If $X_1, X_2, \ldots, X_n$ are mutually independent random variables with respective means $\mu_1, \mu_2, \ldots, \mu_n$ and variances $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$, then the mean and variance of $Y = \Sigma_{i=1}^{n} a_i X_i$, where $a_1, a_2, \ldots, a_n$ are real constants, are given by

(A) $\sigma_Y^2 = \sum_{i=1}^{n} a_i \sigma_i^2$ (B) $\mu_Y = \sum_{i=1}^{n} a_i \mu_i$

(C) $\sigma_Y^2 = \sum_{i=1}^{n} a_i^2 \sigma_i^2$ (D) $\mu_Y = \sum_{i=1}^{n} a_i^2 \mu_i$

9. Let the independent random variables X_1 and X_2 have means $\mu_1 = -4$ and $\mu_2 = 3$, respectively and variances $\sigma_1^2 = 4$ and $\sigma_2^2 = 9$. Then which of the following are correct?

(A) mean = –18
(B) var = 27
(C) mean = 72
(D) var = 72

10. Let $X_1, X_2, \ldots, X_{50}$ be a random sample of size 50 from a distribution with density

$$f(x) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}} & \text{for } 0 \le x < \infty \\ 0 & \text{otherwise.} \end{cases}$$

which of the following are correct?

(A) mean = θ
(B) var = $\frac{\theta^2}{50}$
(C) mean = $\frac{\theta^2}{50}$
(D) var = θ

SECTION-(C) NUMERICAL ANSWER TYPE QUESTIONS (NAT)

1. There are 100 men on a plane. Let X_i be the weight (in pounds) of the ith man on the plane. Suppose that the X_i's are i.i.d., and $EX_i = \mu = 170$ and $\sigma_{X_i} = \sigma = 30$. The probability that the total weight of the men on the plane exceeds 18,000 pounds equals to ________.

2. Let X_1, X_2, ..., X_{25} be i.i.d. with the following PMF

$$P_X(k) = \begin{cases} 0.6 & k = 1 \\ 0.4 & k = -1 \\ 0 & \text{otherwise} \end{cases}$$

And let

$$Y = X_1 + X_2 + \ldots + X_n.$$

Using the CLT and continuity correction, $P(4 \leq Y \leq 6)$ = __________.

3. You have invited 64 guests to a party. You need to make sandwiches for the guests. You believe that a guest might need 0, 1 or 2 sandwiches with probabilities $\frac{1}{4}, \frac{1}{2}$, and $\frac{1}{4}$ respectively. You assume that the number of sandwiches each guest needs is independent from other guests. Number of sandwiches you should make so that you are 95% sure that there is no shortage equals to ________.

4. Let X_1, X_2, ..., X_n be i.i.d. **Exponential(λ)** random variables with $\lambda = 1$. Let

$$\bar{X} = \frac{X_1 + X_2 + \ldots + X_n}{n}.$$

If $P(0.9 \leq \bar{X} \leq 1.1) \geq 0.95$ then n should be ________.

5. A bank teller serves customers standing in the queue one by one. Suppose that the service time X_i for customer i has mean EX_i = 2(minutes) and $Var(X_i) = 1$. We assume that service times for different bank customers are independent. Let Y be the total time the bank teller spends serving 50 customers then $P(90 < Y < 110)$ = _________.

6. In a communication system each data packet consists of 1000 bits. Due to the noise, each bit may be received in error with probability 0.1. It is assumed bit errors occur independently. The probability that there are more than 120 errors in a certain data packet equals to _______.

7. The probability of exactly 55 heads in 100 tosses of a coin equals to _______.

8. A coin is tossed 100 times. The probability that the number of heads lies between 40 and 60 (the word \between" in mathematics means inclusive of the endpoints) equals to ________.

9. A die is rolled 420 times. The probability that the sum of the rolls lies between 1400 and 1550 equals to _________.

10. The value that is 2 standard deviations above the expected value (it is 90) of the sample mean equals to ________.

11. Customers arrive at a mall in accordance with a Poisson process with rate 4,000 persons per day. The approximate value for the probability that tomorrow at least 3850 customers will enter the mall equals to ________.

12. If a can of paint covers on the average 5/3.3 square feet with a standard deviation of 31.5 square feet, then the probability that the mean area covered by a sample of 40 of these cans will be anywhere from 510.0 to 520.0 square feet equals to__________.

13. A distribution with unknown mean μ has variance equal to 1.5. If we use central limit theorem to find how large a sample should be taken from the distribution in order that the probability will be at least 0.95 that the sample mean will be within 0.5 of the population mean.

14. If $X_1, X_2, \ldots, X_n$ are poisson variates with parameter $\lambda = 2$, use the central limit theorem to estimate $P(120 \leq S_n \leq 160)$, where $S_n = X_1 + X_2 + \ldots + X_n$ and $n = 75$.

15. The new Endeavor SUV has been recalled because 5% of the cars experience brake failure. The Tahoe dealership has sold 200 of these cars. The probability that fewer than 4% of the cars from Tahoe experience brake failure equals to _________.

16. Take a random sample of size n = 15 from a distribution whose probability density function is:

$$f(x) = \frac{3}{2}x^2$$

for $-1 < x < 1$. The probability that the sample mean falls between $-2/5$ and $1/5$ is ________.

17. The average time between Infection with the AIDS virus and developing AIDS has been estimated to be 8yr with a standard deviation of about 2yr. Approximately, what fraction of people develop ADS with in 4yr of infection ?

18. If X_1, X_2, ... are iid RVs with common mean μ and finite fourth moment, then

$$P\left\{\lim_{n\to\infty} \frac{S_n}{n} = \mu\right\} = \_\_\_\_\_\_$$

19. Let X_1, X_2, ..., be iid $P(\lambda)$ RVs. and S_n has approximately in $N(n\lambda, n\lambda)$ distribution for large n then $P(S_n = 10) =$ ______

20. Let X be a r.v. with pdf $f_X(x)$. Let $Y = X^2$. the pdf of Y for $y \le 0$ equals to _________.

ANSWER KEY

SECTION-(A) MULTIPLE CHOICE QUESTIONS (MCQ)

1	2	3	4	5	6	7	8	9	10
B	A	C	D	A	B	A	B	D	A
11	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
C	B	C	A	C	C	C	B	A	C
21	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**
B	A	B	C	D	C	B	B	A	B

SECTION-(B) MULTIPLE SELECT QUESTIONS (MSQ)

1	2	3	4	5	6	7	8	9	10
B,C	A,C	A,B	B,C	B	A,B	C,D	B,C	A,D	A,B

SECTION-(C) NUMERICAL ANSWER TYPE QUESTIONS (NAT)

1	2	3	4	5	6	7	8	9	10
0.0004	0.2405	74	385	0.8427	0.0175	0.484	0.9642	0.9670	96
11	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
0.9911	0.6553	16	0.7952	0.2090	0.8185	12	1	0.1009	0