MATH3423 Statistical Inference

Exercise 1 (Introduction)

Questions on the Central Limit Theorem:

1. Let X equal the weight of the soap in a "6-pound" box. Assume that the distribution of X is $N(6.05, 0.0004)$.
   (a) Find $P(X < 6.0171)$.
   (b) If nine boxes of soap are selected at random from the production line, find the probability that at most two boxes weigh less than 6.0171 each.
   (c) Let $\bar{X}$ be the sample mean of the nine boxes. Find $P(\bar{X} \leq 6.035)$.
2. At a heat-treating company, iron castings and steel forges are heat-treated to achieve desired mechanical properties and machinability. One steel forging is annealed to soften the part for each machining. Two lots of this part, made of 1020 steel, are heat-treated in two different furnaces. The specification for this part is 36-66 on the Rockwell G scale. Let X_1 and X_2 are $N(47.88, 2.19)$ and $N(43.04, 14.89)$, respectively. Compute $P(X_1 > X_2)$.
3. Suppose that the distribution of the weight of a prepackaged "1-pound bag" of carrots is $N(1.18, 0.07^2)$ and the distribution of the weight of a prepackaged "3-pound bag" of carrots is $N(3.22, 0.09^2)$. Selecting bags at random, find the probability that the sum of three 1-pound bags exceeds the weight of one 3-pound bag.
4. Suppose that for a particular population of students SAT mathematics scores are $N(529, 5732)$ and SAT verbal scores are $N(474, 6368)$. Select two students at random, and let X equal the first student's math score and Y the second student's verbal score. Find $P(X > Y)$.
5. Approximate $P(39.75 \leq \bar{X} \leq 41.25)$, where $\bar{X}$ is the mean of a random sample of size 32 from a distribution with mean $\mu = 40$ and variance $\sigma^2 = 8$.
6. Let X equal the weight in grams of a miniature candy bar. Assume that $\mu = E(X) = 24.43$ and $\sigma^2 = Var(X) = 2.20$. Let $\bar{X}$ be the sample mean of a random sample of $n = 30$ candy bars. Find
   (a) $E(\bar{X})$.
   (b) $Var(\bar{X})$.
   (c) $P(24.17 \leq \bar{X} \leq 24.82)$, approximately.
7. Let X equal the number out of $n = 48$ mature aster seeds that will germinate when $p = 0.75$ is the probability that a particular seed germinates. Determine $P(35 \leq X \leq 40)$, approximately.
8. Many things once regarded as luxuries are now regarded by most Americans as necessities. For example, the automobile is regarded as a necessity by 90% of American (Public Opinion, 1984). Let X equal the number of Americans in a random sample of $n = 100$ who regard the automobile as a necessity. Assuming that it is still true that $p = 0.90$, determine $P(89 \leq X \leq 94)$, approximately.
9. A candy maker produces mints that have a label weight of 20.4 grams. Assume that the distribution of the weights of these mints is $N(21.37, 0.16)$.
   (a) Let X denote the weight of a single mint selected at random from the production line. Find $P(X < 20.857)$.

(b) During a particular shift 100 mints are selected at random and weighed. Let Y equal the number of these mints that weigh less than 20.857 grams. Find approximately $P(Y \leq 5)$.

(c) Let $\bar{X}$ equal the sample mean of the 100 mints selected and weighed on a particular shift. Find $P(21.31 \leq \bar{X} \leq 21.39)$.

10. Let X equal the number of alpha particles counted by a Geiger counter during 30 seconds. Assume that the distribution of X is Poisson with a mean of 4829. Determine approximately $P(4776 \leq X \leq 4857)$.

11. In the casino game roulette, the probability of winning with a bet on red is $p = 18/38$. Let Y equal the number of winning bets out of 1000 independent bets that are placed. Find $P(Y > 500)$ approximately.

12. Let $X_1, X_2, ..., X_{25}$ and $Y_1, Y_2, ..., Y_{25}$ be two random samples from two independent normal distributions $N(0, 16)$ and $N(1, 9)$, respectively. Let $\bar{X}$ and $\bar{Y}$ denote the corresponding sample means. Compute $Pr(\bar{X} > \bar{Y})$.

13. Let Y be $b(72, \frac{1}{3})$. Approximate $Pr(22 \leq Y \leq 28)$.

14. Let Y be $b(400, \frac{1}{5})$. Compute an approximate value of $Pr(0.25 < Y/400)$.

15. If Y is $b(100, \frac{1}{2})$, approximate the value of $Pr(Y = 50)$.

16. Let $Y = X_1 + X_2 + ... + X_{15}$ be the sum of a random sample of size 15 from the distribution whose p.d.f. is $f(x) = (3/2)x^2, -1 < x < 1$. Approximate

$$P(-0.3 \leq Y \leq 1.5).$$

17. A random sample of size $n = 18$ is taken from the distribution with p.d.f. $f(x) = 1 - x/2, 0 \leq x \leq 2$.

    (a) Find μ and σ^2.

    (b) Find, approximately, $P(2/3 \leq \bar{X} \leq 5/6)$.

18. Let $X_1, X_2, ..., X_{36}$ be a random sample of size 36 from the geometric distribution with p.d.f. $f(x) = (1/4)^{x-1}(3/4), x = 1, 2, 3, ...$. Approximate

    (a) $P\left(46 \leq \sum_{i=1}^{36} X_i \leq 49\right)$.

    (b) $P(1.25 \leq \bar{X} \leq 1.50)$.

19. Let $\bar{X}$ denote the mean of a random sample of size 100 from a distribution which is $\chi^2(50)$. Compute an approximate value of $Pr(49 < \bar{X} < 51)$.

20. Let $\bar{X}$ denote the mean of a random sample of size 128 from a gamma distribution with $\alpha = 2$ and $\beta = 4$. Approximate $Pr(7 < \bar{X} < 9)$.

21. Compute an approximate probability that the mean of a random sample of size 15 from a distribution having p.d.f. $f(x) = 3x^2, 0 < x < 1$, zero elsewhere, is between $\frac{3}{5}$ and $\frac{4}{5}$.

22. Let Y denote the sum of the items of a random sample of size 12 from a distribution having p.d.f. $f(x) = \frac{1}{6}, x = 1, 2, 3, 4, 5, 6$, zero elsewhere. Compute an approximate value of $Pr(36 \leq Y \leq 48)$.

23. Let $f(x) = 1/x^2, 1 < x < \infty$, zero elsewhere, be the p.d.f. of a random variable X. Consider a random sample of size 72 from the distribution having this p.d.f. Compute approximately the probability that more than 50 of the items of the random sample are less than 3.

24. Forty-eight measurements are recorded to several decimal places. Each of these 48 numbers is rounded off to the nearest integer. The sum of the original 48 numbers is approximated by the sum of these integers. If we assume that the errors made by rounding off are stochastically independent and have uniform distributions over the interval $(-\frac{1}{2}, \frac{1}{2})$, compute approximately the probability that the sum of the integers is within 2 units of the true sum.

Questions on Confidence Interval:

25. A random sample of size $n = 36$ from $N(\mu, 25)$ has mean $\bar{x} = 49.2$. Find a 90 percent confidence interval for μ.

26. The mean and standard deviation of $n = 42$ mathematics SAT test scores (selected at random from the entering freshman class of a large private university) are $\bar{x} = 680$ and $s = 35$. Find an approximate 99 percent confidence interval for the population mean μ. Calculate 95 percent confidence intervals for σ^2 and σ.

27. In comparing the times until failure (in hours) of two different types of light bulbs, we obtain the sample characteristics $n_1 = 45,\ \bar{x} = 984,\ s_x^2 = 8742$ and $n_2 = 52,\ \bar{y} = 1121,\ s_y^2 = 9411$. Find an approximate 90 percent confidence interval for the difference of the two population means. Interpret the result and explain why we can use the normal table here despite the fact that the distribution of individual failure times is probably exponential or Weibull.

28. The mean μ of the tear strength of a certain paper is under consideration. The $n = 22$ determinations (taken at random) yielded $\bar{x} = 2.4$ pounds. State your assumption.

    (a) If the standard deviation of an individual measurement is known to be $\sigma = 0.2$, find an approximate 95 percent confidence interval for μ.

    (b) If the standard deviation σ is unknown but the sample standard deviation is $s = 0.2$, determine a 95 percent confidence interval for μ.

29. Let μ be the mean mileage of a certain brand of tire. A sample of $n = 14$ tires was taken at random, resulting in $\bar{x} = 32,132$ and $s = 2596$ miles. Find a 99 percent confidence interval for μ.

30. The effectiveness of two methods of teaching statistics is compared. A class of 24 students is randomly divided into two groups and each group is taught according to a different method. Their test scores at the end of the semester show the following characteristics:

$$n_1 = 13, \quad \bar{x} = 74.5, \quad S_x^2 = 82.6$$

and

$$n_2 = 11, \quad \bar{y} = 71.8, \quad S_y^2 = 112.6.$$

Assuming underlying normal distributions with $\sigma_1^2 = \sigma_2^2$, find a 95 percent confidence interval for $\mu_1 - \mu_2$. Compute 90 percent confidence intervals for σ_1^2/σ_2^2, the ratio of the population variances.

31. One tire manufacturer found that after 5000 miles, $y = 32$ of $n = 200$ steel belted tires selected at random were defective. Find an approximate 99 percent confidence interval for p, the proportion of defective tires in the total production.

32. To test two different training methods, 200 workers were divided at random into two groups of 100 each. At the end of the training program there were $y_1 = 62$ and $y_2 = 74$ successes, respectively. Find an approximate 90 percent confidence interval for $P_1 - P_2$, the difference of the true proportions of success.

33. A sample of $n = 21$ observations leads to $\bar{x} = 74.2$ and $s^2 = 562.8$. Determine a 90 percent confidence interval for σ^2.

Questions on Distribution of Function of r.v.s:

34. Let X_1 and X_2 denote a random sample of size 2 from a distribution which is $N(0, 1)$. Find the p.d.f. of $Y = X_1^2 + X_2^2$.

35. Let X_1 and X_2 be stochastically independent random variables. Let X_1 and $Y = X_1 + X_2$ have Poisson distributions with parameters μ_1 and μ, respectively. Here $\mu_1 < \mu$. Show that X_2 has a Poisson distribution with parameter $\mu - \mu_1$.

36. Let X have the p.d.f. $f(x) = (\frac{1}{2})^x$, $x = 1, 2, 3, ...$, zero elsewhere. Find the p.d.f. of $Y = X^3$.

37. Let X have the p.d.f. $f(x) = x^2/9$, $0 < x < 3$, zero elsewhere. Find the p.d.f. of $Y = X^3$.

38. If the p.d.f. of X is $f(x) = 2xe^{-x^2}$, $0 < x < \infty$, zero elsewhere, determine the p.d.f. of $Y = X^2$.

39. Show that
$$Y = \frac{1}{1 + (r_1/r_2)F},$$
F has an F distribution with parameters r_1 and r_2, has a beta distribution.

40. If $f(x) = \frac{1}{2}$, $-1 < x < 1$, zero elsewhere, is the p.d.f. of the random variable X, find the p.d.f. of $Y = X^2$.

41. If $f(x_1, x_2) = (\frac{2}{3})^{x_1+x_2}(\frac{1}{3})^{2-x_1-x_2}$, $(x_1, x_2) = (0, 0), (0, 1), (1, 0), (1, 1)$, zero elsewhere, is the joint p.d.f. of X_1 and X_2, find the joint p.d.f. of $Y_1 = X_1 - X_2$ and $Y_2 = X_1 + X_2$.

42. Let X_1 and X_2 denote a random sample of size 2 from a distribution which is $n(\mu, \sigma^2)$. Let $Y_1 = X_1 + X_2$ and $Y_2 = X_1 - X_2$. Find the joint p.d.f. of Y_1 and Y_2 and show that these random variables are stochastically independent.

43. If X_1, X_2 is a random sample from a distribution which is $n(0, 1)$, find the joint p.d.f. of $Y_1 = X_1^2 + X_2^2$ and $Y_2 = X_2$ and the marginal p.d.f. of Y_1.

Questions on Order Statistic:

44. Find the probability that the range of a random sample of size 4 from the uniform distribution having the p.d.f. $f(x) = 1, 0 < x < 1$, zero elsewhere, is less than $\frac{1}{2}$.

45. Let $Y_1 < Y_2 < Y_3 < Y_4$ be the order statistics of a random sample of size 4 from the distribution having p.d.f. $f(x) = e^{-x}, 0 < x < \infty$, zero elsewhere. Find $\Pr(3 \leq Y_4)$.

46. Let $f(x) = \frac{1}{6}, x = 1, 2, 3, 4, 5, 6$, zero elsewhere, be the p.d.f. of a distribution of the discrete type. Show that the p.d.f. of the smallest item of a random sample of size 5 from this distribution is
$$g_1(y_1) = \left(\frac{7 - y_1}{6}\right)^5 - \left(\frac{6 - y_1}{6}\right)^5, \quad y_1 = 1, 2, \ldots, 6.$$
zero elsewhere. Note that the random sample is from a distribution of the discrete type.

47. If a random sample of size 2 is taken from a distribution having p.d.f. $f(x) = 2(1-x), 0 < x < 1$, zero elsewhere, compute the probability that one sample item is at least twice as large as the other.

48. Let $Y_1 < Y_2 < Y_3$ be the order statistics of a random sample of size 3 from a distribution having the p.d.f. $f(x) = 2x, 0 < x < 1$, zero elsewhere. Show that $Z_1 = Y_1/Y_2, Z_2 = Y_2/Y_3$, and $Z_3 = Y_3$ are mutually stochastically independent.

49. Let $Y_1 < Y_2 < Y_3$ denote the order statistics of a random sample of size 3 from a distribution with p.d.f. $f(x) = 1, 0 < x < 1$, zero elsewhere. Let $Z = (Y_1 + Y_3)/2$ be the midrange of the sample. Find the p.d.f. of Z.

50. Let $Y_1 < Y_2$ denote the order statistics of a random sample of size 2 from $N(0, \sigma^2)$. Show that $E(Y_1) = -\sigma/\sqrt{\pi}$.

51. Let X_1, X_2, X_3 be a random sample of size 3 from a distribution that is $N(6, 4)$. Determine the probability that the largest sample item exceeds 8.

<u>Miscellaneous:</u>

52. Find the probability that exactly four items of a random sample of size 5 from the distribution having p.d.f. $f(x) = (x + 1)/2, -1 < x < 1$, zero elsewhere, exceed zero.

53. Suppose that an archer is aiming at a target which consists of four concentric circles of radii 1, 2, 3, 4. Assume that each shooting is independent and that the horizontal and vertical distances of the landing point of the arrow from the centre are independent, each having a standard normal distribution. Find the probability that an arrow shot by the archer will

    (i) miss the target entirely,
    (ii) land in the outermost ring, and
    (iii) land in the innermost circle (the bull's-eye)

    Suppose the bull's-eye carries the score 20, the second and the third rings 10 and the outer ring 5. He shoots four arrows and needs a score of 60 or more to qualify for the next round. What is the probability that Robin qualifies? (Note: Work to 4 significant figures.) Hint: Find the distribution of the sum of squares of horizontal and vertical distances first.

54. Suppose that Mark is aiming at a target and the random variables X_1 and X_2 represent the distances of the landing points of his arrows from the vertical and the horizontal axes respectively. Assume that $\underset{\sim}{X} = (X_1, X_2)$ is a bivariate normal distribution with mean $\underset{\sim}{\mu} = \begin{pmatrix} 1.6 \\ 1.3 \end{pmatrix}$ and covariance matrix $\underset{\sim}{\Sigma} = \begin{pmatrix} 1.0 & 0.7 \\ 0.7 & 1.2 \end{pmatrix}$. Suppose that Mark consider that his aim has improved when $|X_1 + X_2| < 1.0$. What proportion of his arrows satisfy this inequality without any improvement having taken place? Suggest another criterion which is more sensible than $|X_1 + X_2| < 1.0$.

55. Let $X_1, \ldots X_n$ be a random sample from $N(0, 1)$. Define

$$\bar{X}_k = \frac{1}{k} \sum_{i=1}^{k} X_i \quad \text{and} \quad \bar{X}_{n-k} = \frac{1}{n-k} \sum_{i=k+1}^{n} X_i$$

    Answer the following:

    (a) What is the distribution of $\frac{1}{2}(\bar{X}_k + \bar{X}_{n-k})$?
    (b) What is the distribution of $k\bar{X}_k^2 + (n-k)\bar{X}_{n-k}^2$?
    (c) What is the distribution of X_1^2/X_2^2?

56. Let Z_1, Z_2 be a random sample of size 2 from $N(0,1)$ and X_1, X_2 a random sample of size 2 from $N(1,1)$. Suppose the Z_i's are independent of the X_j's. Answer the following:

(a) What is the distribution of $\bar{X} + \bar{Z}$?

(b) What is the distribution of $[(X_2 - X_1)^2 + (Z_2 - Z_1)^2]/2$?

(c) What is the distribution of $(Z_1 + Z_2)/\sqrt{[(X_2 - X_1)^2 + (Z_2 - Z_1)^2]/2}$?

(d) What is the distribution of $(X_2 + X_1 - 2)^2/(X_2 - X_1)^2$?