[Marks]

(3) 1. A teacher surveys their students at the end of the semester and asks each of the questions below. In each case, classify the variable of interest as nominal, ordinal, discrete, or continuous.

(a) How much did you enjoy this course on a scale of 1 (not at all) to 5 (very much)? **ordinal**

(b) Would you take this course again? (Yes or no?) **nominal**

(c) How many classes were you absent in the whole semester? **discrete**

(4) 2. True or False (write down **T** or **F** at the front of each statement)

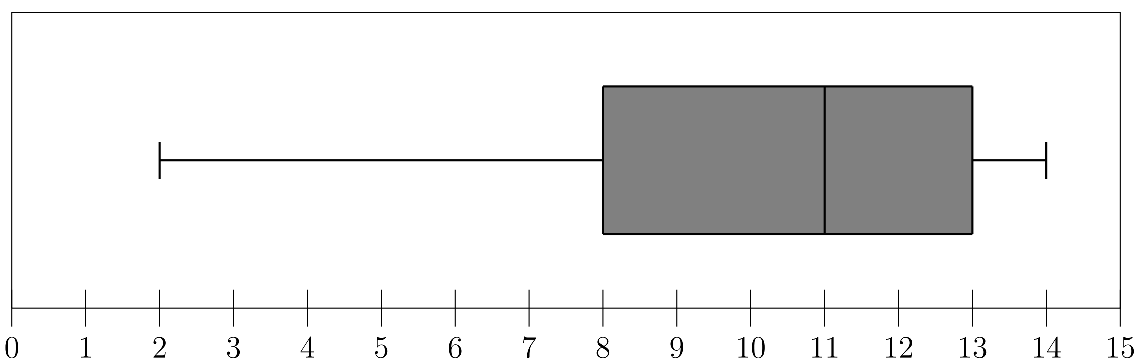
(a) **T** An Ogive is a graph of cumulative relative frequencies

(b) **F** The Poisson distribution is a continuous probability distribution

(c) **F** Every symmetric continuous distribution is normal.

(d) **T** $P(x = a) = 0$ for a continuous random variable x and any real number a.

(5) 3. A set of data is summarized in the following box-and-wisker display:



(a) State the median. $\mathbf{\tilde{x} = Q_2 = 11}$

(b) State the range. $\mathbf{H - L = 14 - 2 = 12}$

(c) State the midrange. $\mathbf{\frac{H+L}{2} = \frac{14+2}{2} = 8}$

(d) State the midquartile. $\mathbf{\frac{Q_1+Q_3}{2} = \frac{8+13}{2} = 10.5}$

(e) State the interquartile range. $\mathbf{Q_3 - Q_1 = 13 - 8 = 5}$

(4) 4. A teacher records student test marks in the following stem-and-leaf display:

4	2							
5	2	8	9					
6	0	0	4	4	4	7		
7	0	1	2	5	6	6	7	8
8	4	5	8	9				
9	3	3	5					

(a) Find the mode.

Solution:

64

(b) Find the median.

Question 4 continues on the next page.

[Marks]

Solution:

$$72$$

(c) Find Q_3. (The third quartile.)

Solution:

$$Q_3 = P_{75}; \quad \frac{nk}{100} = \frac{25 \cdot 75}{100} = 18.75$$

So

$$Q_3 = 84$$

(d) Find P_{80}. (The 80-th percentile.)

Solution:

$$\frac{nk}{100} = \frac{25 \cdot 80}{100} = 20$$

So

$$P_{80} = \frac{85 + 88}{2} = 86.5$$

(4) 5. The mean on a class test was 65 and the standard deviation was 10.

(a) According to Tchebyshev's theorem, at least what percentage of the class earned between 50 and 80 on the test?

Solution: First, we calculate the z-scores of grades between 50 and 80:

$$\frac{50 - 65}{10} \leq z \leq \frac{80 - 65}{10}; \quad -1.5 \leq z \leq 1.5$$

So $k = 1.5$ and according to Tchebyshev's theorem, at least $1 - \frac{1}{1.5^2} = 55.55\%$ of the class earned between 50 and 80 on the test.

(b) If the distribution is not normal, does Tchebyshev's theorem still apply? Yes or no?

Solution: Yes, Tchebyshev's theorem can still be applied because this theorem is valid for any kind of distribution.

(c) If Annie has a z-score of -0.3, what was her mark on the test?

Solution: Since $z = \frac{x - \mu}{\sigma}$,

$$-0.3 = \frac{x - 65}{10}; \quad x - 65 = -3; \quad x = 62$$

[Marks]

(5) 6. Consider the following frequency distribution for a sample of data:

Data x	Frequency f
0	9
1	5
2	10
3	7
4	6
5	3

(a) Compute the mean.

Solution:

Data x	Frequency f	xf	x^2	$x^2 f$
0	9	0	0	0
1	5	5	1	5
2	10	20	4	40
3	7	21	9	63
4	6	24	16	96
5	3	15	25	75
	$\sum f = 40$	$\sum xf = 85$		$\sum x^2 f = 279$

$$\overline{x} = \frac{\sum xf}{\sum f} = \frac{85}{40} = 2.125$$

(b) Compute the sample standard deviation.

Solution: From the table given above,

$$s^2 = \frac{\sum x^2 f - \frac{(\sum xf)^2}{\sum f}}{\sum f - 1} = \frac{279 - \frac{85^2}{40}}{40 - 1} = 2.522$$

Therefore $s = \sqrt{2.522} = 1.588$

(c) Suppose one data value is selected at random, compute the probability that $x = 2$.

Solution:

$$\frac{10}{40} = 0.25$$

(d) Suppose one data value is selected at random, compute the conditional probability that $x = 2$ *given* that x is even.

Solution:

$$P(x = 2|x \text{ is even}) = \frac{P(x = 2 \text{ and } x \text{ is even})}{P(x \text{ is even})} = \frac{\frac{10}{40}}{\frac{9+10+6}{40}} = \frac{10}{25} = 0.4$$

[Marks]

(5) 7. Harry the hacker has discovered seven passwords that Simon uses for all of his online accounts.

(a) Find the *odds in favor* of Harry guessing the password of one account correctly in one try.

Solution:

$$1 : 6$$

(b) Suppose Simon has three bank accounts, what is the probability that Harry can hack into all three on the first try?

Solution:

$$\left(\frac{1}{7}\right)^3 = \frac{1}{343} = 0.002915$$

(c) Suppose Simon has three bank accounts, what is the probability that Simon uses the same password on all three?

Solution:

$$\frac{7}{7^3} = \frac{1}{49} = 0.0204$$

(d) Suppose Simon has three bank accounts, what is the probability that Simon uses different password on all three?

Solution:

$$\frac{7 \cdot 6 \cdot 5}{7^3} = 0.6122$$

(e) Suppose Simon has three bank accounts, what is the probability that Simon uses the same password on at least two?

Solution:

$$\frac{7 + 7 \cdot 6}{7^3} = \frac{1}{7} = 0.1429$$

(6) 8. Suppose A and B are events that satisfy $P(A) = 0.3$ and $P(B) = 0.4$ and that A and B are *mutually exclusive*. Find the following probabilities:

(a) $P(A \cap B)$

Solution:

$$P(A \cap B) = 0$$

because A and B are mutually exclusive.

(b) $P(A \cup B)$

Question 8 continues on the next page.

[Marks]

Solution:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = P(A) + P(B) = 0.7$$

(c) $P(\overline{B})$

Solution:

$$P(\overline{B}) = 1 - P(B) = 1 - 0.4 = 0.6$$

(d) $P(A \cap \overline{B})$

Solution: Since A and B are mutually exclusive, $P(A \cap \overline{B}) = P(A) = 0.3$.

(e) $P(A|\overline{B})$

Solution:

$$P(A|\overline{B}) = \frac{P(A \cap \overline{B})}{P(\overline{B})} = \frac{P(A)}{1 - P(B)} = \frac{0.3}{1 - 0.4} = \frac{1}{2}$$

(f) $P(\overline{B}|A)$

Solution:

$$P(\overline{B}|A) = \frac{P(A \cap \overline{B})}{P(A)} = \frac{P(A)}{P(A)} = 1$$

(3) 9. Suppose we have 10 adults and 8 kids as an audience of a certain show. Find the number of ways the host can select four persons from the audience to volunteer in such a way that the choice must contain at most three kids and at most two adults.

Solution: There are following possibilities: two kids two adults; three kids one adult. So the number of ways to select is

$$C_8^2 C_{10}^2 + C_8^3 C_{10}^1 = 28 \cdot 45 + 56 \cdot 10 = 1820$$

(5) 10. Given the following table,

x	a	-11
$P(x)$	$\frac{b}{8}$	$\frac{b}{8}$

(a) find the values of a and b so that the above table determines a probability distribution with $\mu = -5$.

Question 10 continues on the next page.

[Marks]

Solution: Since $\sum P(x) = 1$, $\frac{b}{8} + \frac{b}{8} = 1$, hence $b = 4$.

Since $\mu = \sum xP(x)$,

$$-5 = a \cdot \frac{4}{8} - 11 \cdot \frac{4}{8}; \quad -5 = \frac{1}{2}a - \frac{11}{2}; \quad -5 + \frac{11}{2} = \frac{1}{2}a; \quad a = 1$$

(b) find σ given that $\mu = -5$.

Solution:

$$\sigma^2 = \left(\sum x^2 P(x)\right) - \mu^2 = \left(1^2 \cdot \frac{1}{2} + (-11)^2 \cdot \frac{1}{2}\right) - (-5)^2 = 36$$

So $\sigma = 6$.

(4) 11. According to the *Applied Ecology and Environmental Research*(Vol.1, 2003) study of beech trees damaged by fungi, it is found that 25% of the beech trees in east central Europe had been damaged by fungi. Consider a random sample of 10 beech trees from this area.

(a) What is the expected number of trees damaged by fungi?

Solution: This is $np = 10 \cdot 0.25 = 2.5$.

(b) What is the probability that at least two trees are damaged by fungi?

Solution: This is a binomial distribution $B(10, 0.25)$ with "success " = a beech tree damaged.

$$\begin{aligned} P(x \geq 2) &= 1 - P(x < 2) = 1 - (P(x = 0) + P(x = 1)) \\ &= 1 - \left(\binom{10}{0} 0.25^0 \cdot 0.75^{10} + \binom{10}{1} 0.25^1 \cdot 0.75^9\right) \\ &= 1 - 0.244 = 0.756 \end{aligned}$$

(6) 12. VoIP(Voice over IP) module is installed in a LAN (Local Area Network). Calls are blocked if all LAN lines are "busy". Let x equal the number of calls blocked during a given hour. It is known x follows the Poisson distribution and on average 5 calls are blocked during the given hour.

(a) What is the probability that no calls are blocked in the given hour?

Solution:

$$P(x = 0) = \frac{5^0 e^{-5}}{0!} = 0.00674 = 0.674\%$$

(b) Find the probability that at least 3 calls are blocked in the given hour.

Question 12 continues on the next page.

[Marks]

Solution: In the Poisson distribution, $\lambda = 5$.

$$\begin{aligned} P(x \geq 3) &= 1 - P(x < 3) = 1 - (P(0) + P(1) + P(2)) \\ &= 1 - \left(\frac{5^0 e^{-5}}{0!} + \frac{5^1 e^{-5}}{1!} + \frac{5^2 e^{-5}}{2!} \right) \\ &= 1 - (0.00674 + 0.0337 + 0.0842) \\ &= 1 - 0.1246 = 0.8754 \end{aligned}$$

(c) What is the probability to have 3 calls blocked in 12 minutes?

Solution: Since on average there are 5 calls blocked in one hour, there will be on average one call blocked in 12 minutes. Hence here $\lambda = 1$ and so

$$P(x = 3) = \frac{1^3 e^{-1}}{3!} = 0.0613$$

(4) 13. Evaluate the following probabilities from the standard normal distribution.

(a) $P(z > 1.22)$

Solution:

$$0.5 - 0.3888 = 0.1112$$

(b) $P(-1.11 < z < 1.22)$

Solution:

$$0.3665 + 0.3888 = 0.7553$$

(6) 14. Solve the following questions.

(a) Find a such that $P(-a < z < 1.1) = 0.5011$

Solution:

$$a = 0.35$$

(b) Find the z-scores that bound the middle 60% area of a normal distribution.

Solution:

$$-0.84 < z < 0.84$$

(c) Find the z-score of P_{40}.

Solution:

$$-0.25$$

[Marks]

(3) 15. Suppose a random variable x is normally distributed with $\mu = -3$ and $\sigma = 1$. Find a such that $P(x < a) = \frac{2}{5}$.

Solution: Using z-scores, we find

$$\frac{a-\mu}{\sigma} = -0.25$$

so $a = -0.25 \cdot \sigma + \mu = -0.25 \cdot 1 - 3 = -3.25$.

(3) 16. Suppose a random variable x is normally distributed with mean 100 and standard deviation 8. Find $P(76 \le x \le 124)$.

Solution:

$$\begin{aligned} P(76 \le x \le 124) &= P\left(\frac{76-\mu}{\sigma} \le z \le \frac{124-\mu}{\sigma}\right) \\ &= P\left(\frac{76-100}{8} \le z \le \frac{124-100}{8}\right) \\ &= P(-3 \le z \le 3) = 0.4987 + 0.4987 = 0.9974 \end{aligned}$$

(4) 17. Suppose a random variable x has a normal distribution with standard deviation 25. It is known that the probability that x exceeds 150 is 0.90. Find μ.

Solution:

$$\begin{aligned} P(x > 150) &= 0.9 \\ P\left(z > \frac{150-\mu}{\sigma}\right) &= 0.9 \\ P\left(z > \frac{150-\mu}{25}\right) &= 0.9 \end{aligned}$$

We find

$$\frac{150-\mu}{25} = -1.28$$

so

$$\mu = 150 + 25 \cdot 1.28 = 182$$

(4) 18. Based on past experience, 7% of all luncheon vouchers are in error. If a random sample of 400 vouchers is selected, Use a normal distribution to estimate the probability that fewer than 25 vouchers are in error?

[Marks]

Solution: We approximate the $B(400, 0.07)$ of random variable x using a normal distribution with mean $\mu = (400)(0.07) = 28$ and standard deviation $\sigma = \sqrt{npq} = \sqrt{(400)(0.07)(0.93)} = 5.103$. The probability calculations are thus

$$\begin{aligned} P(0 < x < 25) &= P(-0.5 < x < 25.5) \\ &\approx P\left(\frac{-0.5-28}{5.103} < x < \frac{25.5-28}{5.103}\right) \\ &= P(-5.58 < z < -0.49) = 0.5 - 0.1878 = 0.3121 \end{aligned}$$

(4) 19. According to the American Automobile Association, the average cost of a gallon of regular unleaded fuel at gas station in April 2007 was \$2.835. Assume the standard deviation of such costs is \$0.15. Suppose that a random sample of 100 gas stations is selected and the April 2007 cost per gallon of regular unleaded fuel is determined for each. Assume the cost per gallon of regular unleaded fuel in gas stations follows a normal distribution. Consider $\overline{x}$, the sample mean cost per gallon.

(a) Calculate $\mu_{\overline{x}}$ and $\sigma_{\overline{x}}$.

Solution:

$$\mu_{\overline{x}} = \mu = 2.835; \quad \sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}} = 0.015$$

(b) What is the probability that the sample has a mean fuel cost between \$2.84 and \$2.86?

Solution:

$$\begin{aligned} P(2.84 < \overline{x} < 2.86) &= P\left(\frac{2.84-\mu_{\overline{x}}}{\sigma_{\overline{x}}} < z < \frac{2.86-\mu_{\overline{x}}}{\sigma_{\overline{x}}}\right) \\ &= P\left(\frac{2.84-2.835}{0.015} < z < \frac{2.86-2.835}{0.015}\right) \\ &= P(0.33 < z < 1.67) \\ &= 0.4525 - 0.1293 = 0.3232 \end{aligned}$$

(4) 20. According to a study about language skills of 15 low-income children, their sentence complexity scores had a mean of 7.62 and a standard deviation of 8.91. Assume the population is normally distributed. Construct a 99% confidence interval for the mean sentence complexity score of all low-income children.

Solution: Here $\alpha = 0.01$ and the point estimation of μ is $\overline{x} = 7.62$. Since we only know the standard the standard deviation s of the sample, t-distribution is used.

$$E = t\,(df, \alpha/2)\left(\frac{s}{\sqrt{n}}\right) = t(14, 0.005)\left(\frac{8.91}{\sqrt{15}}\right) = 2.98 \cdot \frac{8.91}{3.873} = 6.86$$

[Marks]

Hence with 99% confidence level, the mean sentence complexity score of all low-income children should fall in the following interval:

$$(\overline{x} - E, \overline{x} + E) = (7.62 - 6.86, 7.62 + 6.86) = (0.76, 14.48)$$

(4) 21. Marine scientists categorize whistles of dolphins by types — type A, type B, type C, etc. In one study of 185 whistles emitted from dolphins, 97 were categorized as type A whistles. Estimate the proportion of dolphin whistles that are type A whistles. Use a 99% confidence interval.

Solution: This is to find confidence interval for binomial distribution with “success” = type A and “failure” = not type A.

One has $\alpha = 0.01$ and $p' = \frac{97}{185} = 0.524$. So

$$E = z(\alpha/2)\sqrt{\frac{p'q'}{n}} = z(0.005)\sqrt{\frac{0.524 \cdot 0.476}{185}} = 2.58 \cdot 0.0367 = 0.095$$

Hence the 99% confidence interval is

$$(0.524 - 0.095, 0.524 + 0.095) = (0.429, 0.619)$$

(5) 22. National Institute for Standards and Technology(NIST) wanted to study the accuracy of checkout scanners at Wal-Mart stores in California. Suppose NIST wanted to estimate the proportion of Wal-Mart stores in California that violated the NIST standard.

(a) Determine the minimum number of Wal-Marts stores that must be sampled in order to estimate the proportion within 0.05 with 90% confidence.

Solution: Here we do not have an estimation of p, so we choose $p = q = 0.5$. So

$$E = z(\alpha/2)\sqrt{\frac{pq}{n}};\quad 0.05 = z(0.05)\sqrt{\frac{0.5 \cdot 0.5}{n}}$$
$$0.05 = 1.65 \cdot \sqrt{\frac{0.25}{n}};\quad \left(\frac{0.05}{1.65}\right)^2 = \frac{0.25}{n}$$
$$9.18 \times 10^{-4} = \frac{0.25}{n};\quad 1089 = \frac{n}{0.25}$$
$$1089 \cdot 0.25 = n;\quad n = 272.25 \approx 273$$

(b) Suppose we know further that NIST found in a sample of 60 Wal-Mart stores, 52 violated the NIST scanner accuracy standard. Determine the minimum number of Wal-Marts stores that must be sampled in order to estimate the proportion within 0.05 with 90% confidence.

[Marks]

Solution: Here we have $p' = \frac{52}{60} = 0.87$ and so $q' = 0.13$. Therefore

$$E = z(\alpha/2)\sqrt{\frac{p'q'}{n}}; \quad 0.05 = z(0.05)\sqrt{\frac{0.87 \cdot 0.13}{n}}$$
$$0.05 = 1.65 \cdot \sqrt{\frac{0.113}{n}}; \quad \left(\frac{0.05}{1.65}\right)^2 = \frac{0.113}{n}$$
$$9.18 \times 10^{-4} = \frac{0.113}{n}; \quad 1089 = \frac{n}{0.113}$$
$$1089 \cdot 0.113 = n; \quad n = 123.057 \approx 124$$

(5) 23. Do teams playing in their home city have a better chance of winning than when playing elsewhere? In a random sample of 32 hockey games, the home team won 20 times, and the visiting team won the other 12. You would like to test the claim that the home team wins more than 50% of the time with 98% confidence.

(a) State the null and alternative hypotheses. State clearly which is H_0 and which is H_1.

Solution:

$$H_0 : p = 0.5, \quad H_1 : p > 0.5$$

(b) Find the test statistic.

Solution: Since

$$p' = \frac{20}{32} = 0.625, \quad \sigma = \sqrt{\frac{pq}{n}} = \sqrt{\frac{0.5 \cdot 0.5}{32}} = 0.0884$$

the test score is

$$z = \frac{p' - p}{\sigma} = \frac{0.625 - 0.5}{0.0884} = 1.414$$

(c) Find the critical value.

Solution: Since $\alpha = 0.02$, the critical value is

$$z(\alpha) = z(0.02) = 2.05$$

(d) Determine whether or not to reject the null hypothesis.

Solution: Since this is a right-tailed test, we can see the test score $z = 1.414$ is not in the critical region, therefore it is failed to reject the null hypothesis.

(e) State the conclusion of the test in a complete sentence considering the context.

Question 23 continues on the next page.

[Marks]

Solution: There is not enough evidence to conclude that the home team wins more than 50% of the time.

This is the end of the examination.