DO NOT REMOVE FROM THE EXAM HALL

Please complete

ZJU Student ID (10 digits):__________________________

Student Roll Number (4 digits): _______________________

浙江大学爱丁堡大学联合学院

ZJU-UoE INSTITUTE

Advanced Mathematics I

MATH1001

Exam Date:	Start time:	End time:	Duration:	Term:	Year:
Jan 6th	9:30	11:00	1.5 Hours	1st	2021/22

Marking Table

Part 1	Part 2	Part 3	Part 4	Part 5		
3 pts	7 pts	8 pts	9 pts	10 pts		
Part 6	Part 7	Part 8	Part 9	Part 10		Total
10 pts	12 pts	12 pts	14 pts	15 pts		

UNTIL INSTRUCTED DO NOT TURN TO THE NEXT PAGE

PLEASE WRITE YOUR UNIQUE IDENTIFYING NUMBER IN THE TOP LEFT HAND CORNER OF EVERY PAGE

PLEASE READ FULL INSTRUCTIONS BEFORE COMMENCING WRITING

Exam paper information

Total number of pages: 10

Number of questions: 10

Special instructions

1. This is a closed book and notes examination.
2. There is a total of 10 questions in the exam paper. You must answer all the questions.
3. Write your answer with black or black blue coloured pen in the blank area below each questions or sub-questions.
4. Stand-alone dictionary, calculators and statistical tables are allowed. To borrow stationery and dictionary in the exam hall is not allowed.

This examination will be marked anonymously

3 digit student identification number:

1. Researchers studied 1,718 persons over age 65 living in North Carolina. They found that those who attended religious services regularly were more likely to have strong immune systems (as determined by the blood levels of the protein interleukin-6) than those who didn't.

Does this mean that attending religious services improves one's health? Why or why not?

3 pts # 1.2.9 (correlation vs causation)

No. This is an observational study, which means that confounding variables are a concern.

In this case, one might expect that relatively healthy persons are more likely to attend religious services than are those with weak immune systems, so it may be that the immune system affects attendance, rather than the other way around. It could also be that attendance at services is beneficial to the immune system, but that the benefit is due to social interaction, not to the religious nature of the services, per se. Other explanations are also possible.

2. In a study of milk production in sheep (for use in making cheese), a researcher measured the 3-month milk yield for each of 11 ewes. The yields (liters) were as follows:

56.5, 89.8, 110.1, 65.6, 63.7, 82.6, 75.1, 91.5, 102.9, 44.4, 108.1

(a) Determine the median and the quartiles.

(b) Determine the interquartile range.

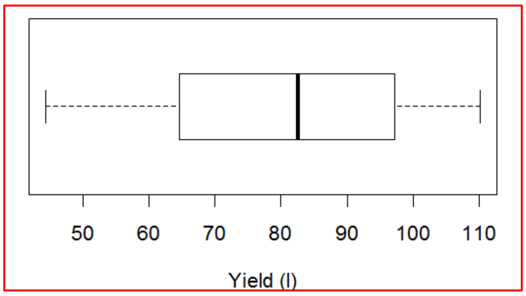
(c) Construct a boxplot of the data.

2.4.3

3pts (a) Median = 82.6, Q1 = 63.7, Q3 = 102.9.

1pt (b) IQR = 39.2.

3pts (c)



3 digit student identification number:

3. Childhood lead poisoning is a public health concern in the United States. In a certain population, 1 child in 8 has a high blood lead level (defined as 30 mg/dl or more). In a randomly chosen group of 16 children from the population, what is the probability that

(a) none has high blood lead?

(b) 1 has high blood lead?

(c) 2 have high blood lead?

(d) 3 or more have high blood lead?

3.6.12

Pr{high blood level} = 1/8 = p. To apply the binomial formula, we arbitrarily identify "success" as "high blood lead." Then n = 16 and p = 1/8.

2 pts (a) To find the probability that none has high blood lead, we set j = 0, so n - j = 16. The binomial formula gives

Pr{none has high blood lead} = ${}_{16}C_0\ (1/8)^0\ (7/8)^{16}$ = $(1)(1)\ (7/8)^{16}$ = 0.1181.

2 pts (b) To find the probability that one has high blood lead, we set j = 1, so n - j = 15. The binomial formula gives

Pr{one has high blood lead} = ${}_{16}C_1\ (1/8)^1\ (7/8)^{15}$ = $(16)\ (1/8)^1\ (7/8)^{15}$ = 0.2699.

2 pts (c) To find the probability that two have high blood lead, we set j = 2, so n - j = 14. The binomial formula gives

Pr{two have high blood lead} = ${}_{16}C_2\ (1/8)^2\ (7/8)^{14}$ = $(120)\ (1/8)^2\ (7/8)^{14}$ = 0.2891.

2 pts (d) Pr{three or more have high blood lead} = 1 - Pr{2 or fewer have high blood lead} = 1 - [0.1181 + 0.2699 + 0.2891] = 0.3229.

3 digit student identification number:

4. The bill lengths of a population of male Blue Jays follow approximately a normal distribution with mean equal to 25.4 mm and standard deviation equal to 0.8 mm. Find the 95th percentile of the bill length distribution.

4.S.2

2 pts In Table 3, the areas closest to 0.95 are 0.9495, corresponding to $z = 1.64$, and 0.9505, corresponding to $z = 1.65$.

5 pts Using 1.64, the 95th percentile y^* satisfies the equation

$$1.64 = (y^* - 25.4) / 0.8$$

2 pts which yields $y^* = (0.8)(1.64) + 25.4 =$ 26.7 mm.

Alternatively, using 1.65, we find $y^* = (0.8)(1.65) + 25.4 = 26.72$ mm, which rounds to 26.7 mm.

3 digit student identification number:

5. In a certain lab population of mice, the weights at 20 days of age follow approximately a normal distribution with mean weight = 8.3 gm and standard deviation = 1.7 gm. Suppose many litters of 10 mice each are to be weighed. If each litter can be regarded as a random sample from the population, what percentage of the litters will have a total weight of 80 gm or more?

5.S.9 μ = 8.3; σ = 1.7.

2 pts If the total weight of 10 mice is 80 gm, then their mean weight is

80/10= 8.0 gm.

Thus, we wish to find the percentage of litters for which $\bar{y} \geq 8.0$ gm. We are concerned with the sampling distribution of $\bar{Y}$ for n = 10.

2 pts From Theorem 5.2.1, the mean of the sampling distribution of $\bar{Y}$ is

$$\mu_{\bar{Y}} = \mu = 8.3,$$

2 pts and the standard deviation is

$$\sigma_{\bar{Y}} = \sigma/\sqrt{n} = 1.7/\sqrt{10} = 0.538,$$

2 pts We need to find the Pr{ $\bar{Y}$ > 8.0 gm}.

For $\bar{Y}$ = 8.0, $z = \frac{\bar{Y}-\mu_{\bar{Y}}}{\sigma_{\bar{Y}}} = \frac{8.0-8.3}{0.538} = -0.56$

From Table 3, the area below –0.56 is 0.2877.

2 pts Thus, the percentage with $\bar{Y} \geq 8.0$ is 1 – 0.2877 = 0.7123, or 71.23%

3 digit student identification number:

6. As part of a study of the development of the thymus gland, researchers weighed the glands of five chick embryos after 14 days of incubation. The thymus weights (mg) were as follows:

29.6   21.5   28.0   34.6   44.9

For these data, the mean is 31.7 and the standard deviation is 8.7.

(a) Calculate the standard error of the mean.

(b) Construct a 90% confidence interval for the population mean.

6.3.3

3 pts (a) $\bar{y} = 31.720$ mg; $s = 8.729$ mg; $n = 5$.

2 pts The standard error of the mean is $SE_{\bar{Y}} = s/\sqrt{n} = 8.729/\sqrt{5} = 3.89 \approx 3.9$ mg.

2 pts (b) The degrees of freedom are $df = n - 1 = 5 - 1 = 4$. The critical value is

$$t_{0.05} = 2.132.$$

3 pts The 90% confidence interval for μ is

$$\bar{y} \pm t_{0.05} \times s/\sqrt{n}$$

$$31.7 \pm 2.132\ (8.729/\sqrt{5})$$

$$(23.4, 40.0) \text{ or } 23.4 < \mu < 40.0 \text{ mg}.$$

3 digit student identification number:

7. Nutritional researchers conducted an investigation of two high-fiber diets intended to reduce serum cholesterol level. Twenty men with high serum cholesterol were randomly allocated to receive an “oat” diet or a “bean” diet for 21 days. The table summarizes the fall (before minus after) in serum cholesterol levels.

	Fall in cholesterol (MG/DL)		
Diet	*n*	Mean	SD
Oat	10	53.6	31.1
Bean	10	55.5	29.4

Use a t test to compare the diets at the 5% significance level.

[Note: Formula (6.7.1) yields 17.9 df.]

7.2.13

2 pts H_0: Mean fall in cholesterol is the same on both diets ($\mu_1 = \mu_2$)
2 pts H_A: Mean fall in cholesterol is not the same on both diets ($\mu_1 \neq \mu_2$)

2 pts SE $_{(\bar{Y}_1 - \bar{Y}_2)} = \sqrt{\frac{31.1^2}{10} + \frac{29.4^2}{10}} = 13.53.$

2 pts $t_s = (53.6 - 55.5)/13.53 = -0.14$; df = $n_1 + n_2 - 2 = 18$ and $t_{18,0.20} = 0.862$.

2 pts Thus, P-value > 0.40, so we do not reject H_0.

2 pts There is insufficient evidence (P > 0.40) to conclude that the two diets differ in their effects on cholesterol.

3 digit student identification number:

8. Each of 36 men was asked to touch the foreheads of three women, one of whom was their romantic partner, while blindfolded. The two "decoy" women were the same age, height, and weight as the man's partner. Of the 36 men tested, 18 were able to correctly identify their partner.
Do the data provide sufficient evidence to conclude that men can do better than they would do by merely guessing? Conduct an appropriate test.

9.4.11
The hypotheses are

2 pts H_0: The men are guessing (Pr{correct} = 1/3)
2 pts H_A: The men have some ability to detect their partners (Pr{correct} > 1/3)

4 pts

Correct	Wrong
18 (12)	18 (24)

2 pts $\chi^2_s = 4.5$.

2 pts With df = 1, Table 9 gives $\chi^2_{0.05} = 3.84$ and $\chi^2_{0.02} = 5.41$, so $0.01 < P < 0.025$ and we reject H_0. (Note that no α level was specified, but a P-value less than 0.025 is generally considered to be small.)

There is sufficient evidence ($0.01 < P < 0.025$) to conclude that the men have some ability to detect their partners by touching them on the forehead.

3 digit student identification number:

9. A researcher collected daffodils from four sides of a building and from an open area nearby. She wondered whether the average stem length of a daffodil depends on the side of the building on which it is growing. Summary statistics are given in the following table.

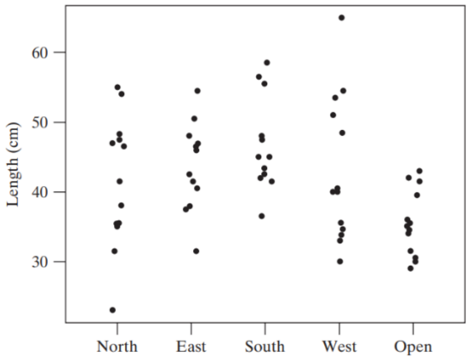
	North	East	South	West	Open
Mean	41.4	43.8	46.5	43.2	35.5
SD	9.3	6.1	6.6	10.4	4.7
n	13	13	13	13	13

The ANOVA SS(between) is 871.408 and the SS(within) is 3588.54.

(a) Dotplots of these data follow. Based on the dotplots, does it appear that the null hypothesis is true? Why or why not?

(b) State the null hypothesis in symbols.

(c) Construct the ANOVA table and test the null hypothesis. Let α = 0.10



11.4.5

2 pts (a) The dotplots give the impression of differences between the five groups. In particular, the Open group appears to have a smaller mean than the others.

2 pts (b) In symbols, the null hypothesis is H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$

2 pts (c) I = 5, $n_\bullet$ = 65.

4 pts

Source	df	SS	MS
Between groups	4	871.408	217.852
Within groups	60	3588.54	59.809
Total	64	4459.948	

2 pts The test statistic is F_s = 217.852/59.809 = 3.64. With df = 4 and 60, Table 10 gives $F_{0.02}$ = 3.16 and $F_{0.01}$ = 3.65.

2 pts Thus, we have 0.01 < P < 0.02, so we reject H_0. There is sufficient evidence (0.01 < P < 0.02) to conclude that mean stem lengths of daffodils is NOT the same in all five locations.

3 digit student identification number:

10. To investigate the dependence of energy expenditure on body build, researchers used underwater weighing techniques to determine the fat-free body mass for each of seven men. They also measured the total 24-hour energy expenditure for each man during conditions of quiet sedentary activity. The results are shown in the table.

Subject	Fat-free mass X (kg)	Energy expenditure Y (kcal)
1	49.3	1,894
2	59.3	2,050
3	68.3	2,353
4	48.1	1,838
5	57.6	1,948
6	78.1	2,528
7	76.1	2,568
Mean	62.400	2,168.429
SD	12.095	308.254
	$r = 0.98139$	

(a) The correlation between energy expenditure and fat-free mass is very large (near 1). It is 0.98139, but the sample size is quite small, only 7. Is there enough evidence to claim the correlation is different from zero? Carry out an appropriate test using α = 0.05.

(b) Is this study an observational study or an experiment?

(c) Persons who exercise could increase their fat-free mass. Could these data be used to claim that their energy expenditure would also increase? If not, what could be said? Briefly explain.

(d) For these data, SS(resid) = 21,026.1. Construct a 95% confidence interval for β_1.

12.2.9 + 12.5.5

(a) The hypotheses are
1 pts H_0: $\rho = 0$ (no linear relationship between fat-free mass and energy expenditure)
1 pts H_A: $\rho \neq 0$ (a linear relationship between fat-free mass and energy expenditure)

2 pts The test statistic is

$$t_s = = r\sqrt{\frac{n-2}{1-r2}} = 0.98139\sqrt{\frac{7-2}{1-0.98139^2}} = 11.428$$

2 pts Consulting Table 4 with df = 7 - 2 = 5, we find that $t_{0.0005} = 6.869$.
Thus, $P < 0.001$, so we reject H_0 (with $\alpha = 0.05$).

3 digit student identification number:

There is strong evidence that the population correlation is different from zero, that fat-free mass and energy expenditure are linearly related.

2 pts (b) This study was an observational study. The researchers simply measured both fat-free mass and energy expenditure of each subject without manipulating either variable.

2 pts (c) No. While there is significant evidence that fat-free mass and energy expenditure are linearly related, because this was an observational study, we cannot establish a cause and effect relationship.

(d) The sample slope is $b_1 = 0.98139\ (308.254/12.095) = 25.011$ and

$s_e = \sqrt{21026.1/5} = 64.85$. The standard error of the slope is

2 pts $SE_{b1} = \frac{s_e}{s_x\sqrt{n-1}} = \frac{64.85}{12.095\sqrt{7-1}} = 2.189$.

1 pts To construct a 95% confidence interval we consult Table 4 with df = n - 2 = 7 - 2 = 5; the critical value is $t_{5,0.025} = 2.571$. The confidence interval is

$b_1 \pm t_{0.025}\, SE_{b1}$

$25.011 \pm (2.571)(2.189)$

2 pts $19.4 < \beta_1 < 30.6$ kcal/kg.