Last Name:

First Name:

Student Number:

Instructions:

1. Print your name and student number above AND on the MC answer sheet. A test or MC answer sheet without a name and student number won't be marked. A page (except the front page) without a name at the top right, where shown, won't be marked.
2. Use only a pencil when filling in the MC answer sheet for the multiple choice questions. Circle the correct answers on your question paper first and only when you are certain of your answer fill in the MC answer sheet. Only the answers on the MC answer sheet will be used when marking the multiple choice questions.
3. Check that you have all 5 PAGES before beginning the exam.
4. Pace yourself – you have ~ 50 minutes.
5. Use the blank spaces on exam pages for rough work. No scrap paper is permitted.
6. Simple non-programmable calculators (not cell phones or tablets) are allowed.
7. If you have a cell phone or any electronic device (other than a pacemaker) with you – be sure it is turned off now, and stored in a safe place away from your desk.
8. Hand in BOTH this exam booklet AND the MC answer sheet. Taking an exam booklet from the exam room will result in an automatic grade of 'F' for this course.
9. This test is worth 17.5% of your final mark.

I have read, understood, and will comply with all of the above instructions:

Formulas are at the end. Choose the **best** answer in each of the following multiple-choice questions.
1 mark each

1. What would be the purpose of writing a Maclaurin series approximation to a function in your code?
a. reduce roundoff errors b. increase the overflow limit c. reduce the underflow limit
d. increase execution speed with a small number of terms e. increase precision with a small number of terms

The following data is a summary of weekly hours Internet usage for a sample of 50 users. It is used in questions 2, 3 and 4. The frequency and relative frequency for 20-24 hours are blanked out.

Estimated Hours on Internet	*Frequency*	*Relative Frequency (%)*
0–4	2	4
5–9	9	18
10–14	19	38
15–19	11	22
20–24	[illegible]	[illegible]
25–29	3	6
Total	50	100

2. Each of the entries in the "*Estimated Hours on Internet*" column is called a
a. standard deviation b. variance c. mean d. residual e. class

3. What is the relative frequency (%) for internet hours 20-24
a. 3 b. 6 c. 12 d. 24 e. none of the previous

4. What is the frequency for internet hours 20-24
a. 3 b. 6 c. 12 d. 24 e. none of the previous

5. The number corresponding to the height of each rectangle in a histogram is called the
a. class mark b. class width c. population d. sample e. frequency

6. If f_i represents the frequency of class i, which of the following is the value of the expression $\sum f_i$?
a. 0 b. 0.5 c. 1.0 d. size of the sample e. size of the population

7. For a set of statistical data x_i with relative frequencies f_i, the expression $\sum x_i f_i$ has the meaning
a. total population b. standard deviation c. total sample size d. arithmetic mean e. it has no special meaning

8. What is true about the standard deviation of a sample of data with a small random error?
a. it gets larger as the error increases b. it gets smaller as the error increases
c. it gets larger as the mean increases d. it gets larger as the sample size increases
e. it doesn't depend on the error

9. For a set of statistical data what is true about the expression $\sum (x - \bar{x})^2$?
a. its value is zero b. it's either negative or positive c. it has fewer terms for smaller samples
d. it's used in the calculation of the mean e. wrong, it's not used in statistics

10. Samples of size n are taken from a population of size N. What is true about the standard deviation of the means of these samples?
a. It gets smaller as n increases b. it gets larger as n increases
c. It gets smaller as N increases d. it gets larger as N increases e. it doesn't depend on n or N

The following set of 9 numbers x_i is used in questions 11 and 12:

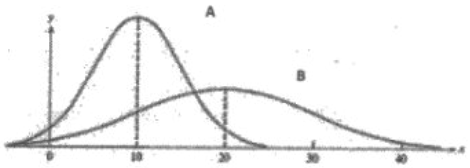
i	1	2	3	4	5	6	7	8	9
x_i =	-4	-3	-2	-1	0	1	2	3	4

11. The mean value of x is closest to
a. 0 b. 3 c. 4 d. 4.5 e. 5

12. The standard deviation of x is closest to
a. 0 b. 3 c. 4 d. 4.5 e. 5

13. In describing random data distributed normally which of the following is correct?
a. 68% lie within 1, 95% lie within 2 and 100% lie within 3 standard deviations of the mean
b. 95% lie within 1, 99.7% lie within 2 and 100% lie within 3 standard deviations of the mean
c. 99.7% lie within 1, 95% lie within 2 and 68% lie within 3 standard deviations of the mean
d. 68% lie within 1, 95% lie within 2 and 99.7% lie within 3 standard deviations of the mean
e. none of the above are correct

14. The following graph shows two Normal Distributions A and B.
Which of the following is the best description?



a. A has a mean of 10 and a Standard deviation of ~ 5; B has a mean of 20 and a Standard deviation of ~10
b. A has a mean of 10 and a Standard deviation of ~ 5; B has a mean of 20 and a Standard deviation of ~ 5
c. A has a mean of 10 and a Standard deviation of ~ 10; B has a mean of 20 and a Standard deviation of ~ 10
d. A has a mean of 10 and a Standard deviation of ~ 20; B has a mean of 20 and a Standard deviation of ~ 20
e. A has a mean of 10 and a Standard deviation of ~ 10; B has a mean of 20 and a Standard deviation of ~ 5

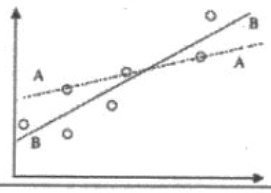
15. In the definition of the standard score, or z score, defined by the relation on the right, the symbols A, B and C represent:

$z = (B - A)/C$

a. A is data value, B is mean, C is standard deviation
b. B is data value, C is mean, A is standard deviation
c. C is data value, A is mean, B is standard deviation
d. B is data value, A is mean, C is standard deviation
e. A is data value, C is mean, B is standard deviation

The following xy diagram is used in question 16.
It shows two lines that follow the equation

$y = mx + c$



16. What is different between the lines AA and BB?
a. AA has the smaller sum of squares of residuals and BB has the larger sum of squares of residuals
b. AA has the larger sum of squares of residuals and BB has the smaller sum of squares of residuals
c. AA has the smaller c and BB has the larger c
d. AA has the larger m and BB has the smaller m
e. none of the previous give the difference between AA and BB

17. For the following straight line sum of squares of residuals: $SS = (y_1 - mx_1 - c)^2 + (y_2 - mx_2 - c)^2 + \ldots + (y_N - mx_N - c)^2$ which of the following is the condition to find the minimum slope?

a. $\sum_{i=1}^{i=n}(y_i - mx_i - c) = 0$
b. $\sum_{i=1}^{i=N}(y_i - mx_i - c)x_i = 0$
c. $\sum_{i=1}^{i=N}(y_i + mx_i - c)x_i = 0$
d. $\sum_{i=1}^{i=N}(y_i + mx_i - c) = 0$
e. $\sum_{i=1}^{i=N}(y_i - mx_i + c)x_i = 0$

18. For the following straight line sum of squares of residuals: $SS = (y_1 - mx_1 - c)^2 + (y_2 - mx_2 - c)^2 + \ldots + (y_N - mx_N - c)2$ which of the following is the condition to find the minimum intercept?

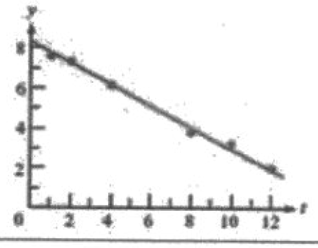
a. $\sum_{i=1}^{i=n}(y_i - mx_i - c) = 0$
b. $\sum_{i=1}^{i=N}(y_i - mx_i - c)x_i = 0$
c. $\sum_{i=1}^{i=N}(y_i + mx_i - c)x_i = 0$
d. $\sum_{i=1}^{i=N}(y_i + mx_i - c) = 0$
e. $\sum_{i=1}^{i=N}(y_i - mx_i + c)x_i = 0$

19. In the formulas at the end, the expression $(\sum x)^2$ can alternatively be written for a set of x and y data as

a. $(x_1x_1)(x_2x_2)\ldots(x_nx_n)$
b. $(x_1x_1 + x_2x_2 + x_3x_3 + \ldots\ldots)$
c. $(x_1+x_1)(x_2+x_2)\ldots(x_n+x_n)$
d. $x_1(x_1 x_2 \ldots x_n) + x_2(x_1 x_2 \ldots x_n) + \ldots + x_n(x_1 x_2 \ldots x_n)$
e. $(x_1 + x_2 + \ldots + x_n)(x_1 + x_2 + \ldots + x_n)$

20.
In the following graph the solid line is a linear regression fit to the equation $y = mx + b$.
Which one of the following is close to likely values of m and b?

a. m = 0.5, b = 8
b. m = -0.5, b = 8
c. m = 2, b = 8
d. m = -2, b = 8
e. none of the previous is correct



Short Questions

21. Recordings of lifetimes of cellphone batteries are distributed normally with $\mu = 30$ months and $\sigma = 2$ months.

21a. [5 marks] In a sample of 1000 what is closest to the number expected to last between 28 months and 34 months?

For 28 months z = (28-30)/2 = -1.0 A = 0.3413
for 34 months z = (34-30)/2 = 2.0 A = 0.4772 total A = 0.8185
Num = 0.8185*1000 = 818.5

21b. [5 marks] In a sample of 1000 what is closest to the number expected to last less than 27 months?

For 27 months z = (27 – 30)/2 = -1.5 A = 0.4332
Num = (0.5 – 0.4332)*1000 = 66.8

22 [10 marks]
The altitude h (in km) of a rocket was measured at three positions of a horizontal distance s (in km) from the launch site as shown in the table.

s (km)	0	1.0	2.0
h (km)	0	2.25	4.5

Using the least-squares formulas:

22a. [1 mark]. Calculate the value of S_x 3

22d. [1 mark]. Calculate the value of S_y 6.75

22b. [1 mark]. Calculate the value of S_{xx} 5

22c. [1 mark]. Calculate the value of S_{xy} 11.25

22e. [3 marks]. Calculate the value of m in a fit to a line y = mx + c

$\Delta = SS_{xx} - (S_x)^2$ = 3*5 – 9 = 6,

$m = \dfrac{SS_{xy} - S_x S_y}{\Delta}$ = (3*11.25 – 3*6.75)/6 = (33.75 – 20.25)/6 = 2.25

22f. [3 marks]. Calculate the value of c in a fit to a line y = mx + c

$c = \dfrac{S_{xx} S_y - S_x S_{xy}}{\Delta}$ = (5*6.75 – 3*11.25)/6 = (33.75 – 33.75)/6 = 0

Formulas

Statistics

Arithmetic mean

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + \cdots + x_n f_n}{f_1 + f_2 + \cdots + f_n}$$

Standard deviation

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$$

$$s = \sqrt{\frac{n(\sum x^2) - (\sum x)^2}{n(n-1)}}$$

Normal distribution

$$y = \frac{e^{-(x-\mu)^2/2\sigma^2}}{\sigma\sqrt{2\pi}}$$

Linear Regression

$$S = \sum_{i=1}^{N} = N\,; \quad S_x = \sum x_i\,; \quad S_y = \sum y_i\,; \quad S_{xy} = \sum x_i y_i\,; \quad S_{xx} = \sum x_i^2\,; \quad \Delta = SS_{xx} - (S_x)^2\,;$$

The solutions for m and c in the straight line y = mx + c are:

$$m = \frac{SS_{xy} - S_x S_y}{\Delta} \qquad c = \frac{S_{xx} S_y - S_x S_{xy}}{\Delta}$$

Area under a Standard Normal Distribution

the standard score z (or z-score), which is defined as

$$z = \frac{x - \mu}{\sigma}$$

Table 22.1 Standard Normal (z) Distribution

z	Area	z	Area	z	Area
0.0	0.0000	1.0	0.3413	2.0	0.4772
0.1	0.0398	1.1	0.3643	2.1	0.4821
0.2	0.0793	1.2	0.3849	2.2	0.4861
0.3	0.1179	1.3	0.4032	2.3	0.4893
0.4	0.1554	1.4	0.4192	2.4	0.4918
0.5	0.1915	1.5	0.4332	2.5	0.4938
0.6	0.2257	1.6	0.4452	2.6	0.4953
0.7	0.2580	1.7	0.4554	2.7	0.4965
0.8	0.2881	1.8	0.4641	2.8	0.4974
0.9	0.3159	1.9	0.4713	2.9	0.4981
1.0	0.3413	2.0	0.4772	3.0	0.4987

Test 1 CST8233 W18 Answers

1. d 2. c 3. d 4. d 5. a 6. d 7. b 9. b 10. d 11. d 12. b
13. e 14. b 15. c 16. d 17. a,b 18. b 20. a 21. c 22. b 23. c

24. An example in hybrid 2.
$3EC00000_{16}$ = 0011 1110 1100 0000 0000 0000 0000 0000_2
so sign bit = 0 => +ve,
exponent = 01111101 = 125_{10} = 127 - 2 therefore two left shifts occurred
mantissa = [1].100 0000 0000 0000 0000 0000_2 normalised
= 0.0110_2 before normalization
= 0.25 +.125 = 0.375_{10}
Answer is $+0.375_{10}$

25a. *Answer: From the formulas*

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots$$

Then replacing x by x/2 which was from hybrid 1 Exercise 2 #2

sin x/2 =	$\frac{x}{2} - \frac{x^3}{2^3 3!} + \frac{x^5}{2^5 5!} - \frac{x^7}{2^7 7!} + \dots$

Finally, multiplying by ½ we get $\frac{1}{2}sin\left(\frac{x}{2}\right) = \frac{x}{4} - \frac{x^3}{16*3!} + \frac{x^5}{64*5!}$ $= x/4 - x^3/96 + x^5/7680$ *for the first three terms.*

25b. *With 2 terms at x = 1 the % fractional error is:*

$\sim \frac{\frac{1^5}{64*5!}}{\frac{1}{4} - \frac{1^3}{16*3!}} \sim 100 * 4/(7680) = 0.05\%$