Instructor: Dr. Jennifer Murdock
Duration: 50 minutes. You must stay in the test room for the entire time.
Format: 20 multiple choice questions with answers recorded on SCANTRON form
Point values: Each multiple question worth 5 points. There are 100 total possible points.
Allowed aids: A non-programmable calculator (and attached aid sheets, which you may detach)

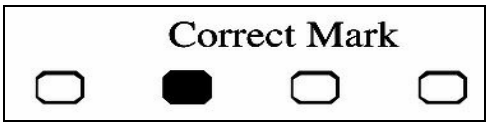
INSTRUCTIONS:

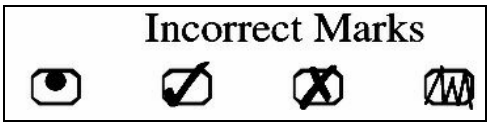
Do NOT write your answers to the multiple choice questions on these test papers
ONLY those answers correctly marked on the SCANTRON form can earn marks
You MAY do scratch work on these pages

SCANTRON INSTRUCTIONS

- Use only a blue or black ball point pen or pencil

- **Pencil strongly recommended**, it can be erased if a mistake is made
- Make dark solid marks that fill the bubble completely





- Select the one best alternative
  - Questions with more than one answer selected will be scored incorrect
- Erase completely any marks you want to change
  - Do not use correction fluid
- Do not make stray marks on the form
- Answer every question: there is no penalty for guessing

1st : Print your **LAST NAME** and **INITIALS** in boxes provided
- Use exact name you are officially registered under
- Darken each letter in the corresponding bracket below each box

2nd : Print your 9 digit **STUDENT NUMBER** in the boxes provided
- Fill in zeros in front of the number if less than 9 digits
- Darken each number in the corresponding bracket below each box

3rd : Print 2 digit **FORM** number in the boxes provided
- Your FORM number is 03 if you are in Section L0301 (11:10 – 12:00)
- Your FORM number is 04 if you are in Section L0401 (12:10 – 1:00)
- Darken each number in the corresponding bracket below each box

4th : Sign your name in the **SIGNATURE** box

For the 20 questions, choose the <u>most correct</u> answer and mark it on the SCANTRON form.

► For Questions **(1) - (2)**: Suppose employees' hourly wages are normally distributed with a mean of $50 and a standard deviation of $1.25.

(1) If an employee were randomly selected, what interval would you be 90% confident would include the employee's wage?

(a) ($42.11, $57.89)
(b) ($47.50, $52.50)
(c) ($47.55, $52.45)
(d) ($47.94, $52.06)
(e) ($48.75, $51.25)

(2) What is the probability that the population mean is greater than $51.25?

(a) 0.00
(b) 0.05
(c) 0.10
(d) 0.16
(e) 0.34

(3) Consider the number of liters of water per month employees consume at work in four different large corporations. Suppose random samples of employees are taken from each corporation with varying sample sizes. For each, the sample mean and sample standard deviation are computed:

Corporation 1: $n_1 = 5;\ \overline{X}_1 = 20;\ s_1 = 5$
Corporation 2: $n_2 = 35;\ \overline{X}_2 = 20;\ s_2 = 10$
Corporation 3: $n_3 = 50;\ \overline{X}_3 = 20;\ s_3 = 15$
Corporation 4: $n_4 = 80;\ \overline{X}_4 = 20;\ s_4 = 20$

For which corporation could water consumption be normally distributed?

(a) Corporation 1
(b) Corporation 2
(c) Corporation 3
(d) Corporation 4
(e) All four corporations

(4) If X is normally distributed with mean 57 and standard deviation 2, for a random sample of 12, what is $P(\overline{X} < 57)$?

(a) 0.25
(b) 0.50
(c) 0.75
(d) 1.00
(e) None of the above

► For Questions **(5) - (6)**, consider a sample of size 36 with a sample mean of 19.9 and a sample standard deviation of 4.1. The research hypothesis is that the population mean is less than 20 and a 5% significance level is chosen.

(5) What is the rejection region for this hypothesis test?

(a) $t < 1.69$
(b) $t > 1.69$
(c) $t < -1.69$
(d) $t > -1.69$
(e) $t < -2.03$ and $t > 2.03$

(6) Which of the following is closest to the p-value of this hypothesis test?

(a) 0
(b) 0.025
(c) 0.05
(d) 0.10
(e) 0.45

(7) Suppose a computer printout for a hypothesis test reports a p-value of 0.0003. What statistical inference could you reasonably make?

(a) Research hypothesis is true
(b) Research hypothesis is false
(c) Null hypothesis is true
(d) Population proportion is 0.0003
(e) None because fail to reject the null hypothesis

(8) Consider the problem of filtering spam (unwanted e-mail) from users' inboxes. It is estimated that over 60% of e-mail traffic is spam. Suppose that a company designs spam filtering software based on the following hypothesis test:

H_0: An e-mail is legitimate
H_1: An e-mail is spam

Which of the following is the Type II error?

(a) The significance level: α
(b) One minus the significance level: 1 - α
(c) Legitimate e-mail blocked as spam
(d) Spam that is not blocked by the filter
(e) The percent of e-mail that is spam: 60%

(9) Considering households' decisions about whether to buy home delivery of a daily newspaper, suppose an analyst contends that less than half of the population of households has it. Suppose the point estimate of the population proportion is 0.2. Which is the correct specification of H_0 and H_1?

(a) H_0: $\mu = 0.2$; H_1: $\mu < 0.2$
(b) H_0: $\mu = 0.5$; H_1: $\mu < 0.5$
(c) H_0: $\mu = 0.5$; H_1: $\mu > 0.5$
(d) H_0: $p = 0.2$; H_1: $p < 0.2$
(e) H_0: $p = 0.5$; H_1: $p < 0.5$

► For Question **(10)** consider the following quote from a February 12, 2006 New York Times newspaper article discussing scientific studies investigating the safety of aspartame, which is an chemical sweetener sold under the brand-names NutraSweet and Equal and used in diet soft-drinks and many other low-calorie products:

Michael F. Jacobson, executive director of the Center for Science in the Public Interest, quoted as saying: "For a chemical that is used by hundreds of millions of people around the world, it should be absolutely safe. There shouldn't be a cloud of doubt."

(10) Which of the following specifications of hypotheses and significance levels would best reflect Mr. Jacobson's views?

(a) H_0: Drug is not safe; H_1: Drug is safe; $\alpha = 0.05$
(b) H_0: Drug is not safe; H_1: Drug is safe; $\alpha = 0.01$
(c) H_0: Drug is safe; H_1: Drug is not safe; $\alpha = 0.10$
(d) H_0: Drug is safe; H_1: Drug is not safe; $\alpha = 0.05$
(e) H_0: Drug is safe; H_1: Drug is not safe; $\alpha = 0.001$

(11) Which of the following statements about the relationship among economic significance, statistical significance, and sample sizes is TRUE?

(a) With a small sample it is not possible to find a statistically significant effect but finding an economically significant effect is possible
(b) With a large sample if you find a statistically significant effect this implies an economically significant effect
(c) With a small sample if you find an economically significant effect this implies a statistically significant effect
(d) Increasing the sample size causes an increased likelihood of finding that an economically significant effect is statistically significant
(e) All of the above are true

(12) With conventional significance levels, which is the <u>most powerful</u> statistical test?

(a) H_0: $p = 0.5$; H_1: $p < 0.5$; n = 1; population proportion = 0.505
(b) H_0: $p = 0.5$; H_1: $p < 0.5$; n = 10; population proportion = 0.005
(c) H_0: $p = 0.5$; H_1: $p < 0.5$; n = 100; population proportion = 0.005
(d) H_0: $p = 0.5$; H_1: $p < 0.5$; n = 1000; population proportion = 0.495
(e) H_0: $p = 0.5$; H_1: $p \neq 0.5$; n = 1000; population proportion = 0.495

► For Questions **(13) - (15)**, consider the effect of monetary incentives on primary and secondary school attendance rates. Currently, some schools have made the decision to use local taxpayer dollars to fund incentives to increase student attendance rates. Consider the following two data collection plans:

> Data Collection Plan #1: Collect a random sample of 30 schools that currently have monetary incentive programs and a random sample of 30 schools that do not currently have monetary incentive programs.
>
> Data Collection Plan #2: Start a new federal program that selects a random sample of 60 schools. Of the 60, the federal government randomly selects 30 schools and requires each to use a monetary incentive program. The remaining 30 are forbidden to use any monetary incentive program. Schools required to have an incentive program are given federal funding.

For either data collection plan, the following null hypothesis is proposed, where μ_1 is the average attendance rate across schools if a monetary incentive program is in place and μ_2 is the average attendance rate across schools if a monetary incentive program is not in place:

H_0: $(\mu_1 - \mu_2) = 0$

(13) Which of the data collection plans will provide observational data about attendance rates and monetary incentive programs?

(a) #1
(b) #2
(c) Both #1 and #2
(d) Neither #1 nor #2

(14) In which of the following cases, which specify a data collection plan, a research hypothesis, and a p-value, would the conclusion that the monetary incentive programs are effective in increasing attendance be most justified?

(a) Data Collection Plan #1; H_1: $(\mu_1 - \mu_2) > 0$; p-value = 0.01
(b) Data Collection Plan #1; H_1: $(\mu_1 - \mu_2) < 0$; p-value = 0.98
(c) Data Collection Plan #1; H_1: $(\mu_1 - \mu_2) \neq 0$; p-value = 0.04
(d) Data Collection Plan #2; H_1: $(\mu_1 - \mu_2) > 0$; p-value = 0.03
(e) Data Collection Plan #2; H_1: $(\mu_1 - \mu_2) \neq 0$; p-value = 0.02

(15) If a non-directional research hypothesis is specified and a 5% significance level used, which could be the rejection region?

(a) $z < -1.96$ and $z > 1.96$
(b) $z < -1.65$ and $z > 1.65$
(c) $t < -1.30$ and $t > 1.30$
(d) $t < -1.67$ and $t > 1.67$
(e) $t < -2.00$ and $t > 2.00$

► For Questions **(16) - (18)**, suppose a survey is administered to a random sample of cellular telephone users. Here are questions 5 (q5) and 6 (q6) from that survey and data summaries:

5) On a scale from 1 to 5, with 1 being "not at all likely" and 5 being "extremely likely", how likely are you to switch cellular telephone service providers in the next six months?

6) For how many months have you been with your current cellular telephone service provider?

Variable	Obs	Mean	Std. Dev.	Min	Max
q5	111	2.873874	1.414966	1	5
q6	111	10.47748	5.97091	1	36

q5	Freq.	Percent	Cum.
1	30	27.03	27.03
2	14	12.61	39.64
3	21	18.92	58.56
4	32	28.83	87.39
5	14	12.61	100.00
Total	111	100.00	

q6	Freq.	Percent	Cum.
1	6	5.41	5.41
5	6	5.41	10.81
6	27	24.32	35.14
7	8	7.21	42.34
8	2	1.80	44.14
9	3	2.70	46.85
10	3	2.70	49.55
12	34	30.63	80.18
14	4	3.60	83.78
15	2	1.80	85.59
16	2	1.80	87.39
17	2	1.80	89.19
18	5	4.50	93.69
24	5	4.50	98.20
30	1	0.90	99.10
36	1	0.90	100.00
Total	111	100.00	

(16) What is the 95% confidence interval estimate of the fraction of cellular telephone users that are extremely likely to switch in the next 6 months?

(a) (0.0010, 0.0565)
(b) (0.0261, 0.0313)
(c) (0.0643, 0.1879)
(d) (0.0984, 0.1538)
(e) (0.3237, 0.3791)

(17) What is the 95% confidence interval estimate of the average number of months cellular telephone users have been with their service provider?

(a) (1, 21)
(b) (1, 24)
(c) (9.36, 11.60)
(d) (9.54, 11.42)
(e) (10.37, 10.59)

(18) Suppose telephone service providers believe that more than a fifth of subscribers are not at all likely to switch in the next six months. What is the p-value for the implied hypothesis test?

(a) 0.033
(b) 0.048
(c) 0.050
(d) 0.064
(e) 0.096

► For Questions **(19) - (20)**, consider conducting a hypothesis test regarding the difference between two population means using matched-pair data. Define the difference as $X_D = X_1 – X_2$.

(19) The following sample statistics are found: $s_1 = 8$ and $s_2 = 6$ and $s_D = 8$. What is the covariance between X_1 and X_2?

(a) 0
(b) 2
(c) 14
(d) 18
(e) 46

(20) In obtaining confidence interval (CI) estimates of the difference between population means, what, if anything, does the coefficient of correlation between X_1 and X_2 imply about the width of the CI estimates when treating these data as matched-pair versus independent samples?

(a) Small coefficient of correlation: expect matched-pair CI to be substantially wider than independent samples CI
(b) Small coefficient of correlation: expect matched-pair CI to be substantially narrower than independent samples CI
(c) Large coefficient of correlation: expect matched-pair CI to be substantially wider than independent samples CI
(d) Large coefficient of correlation: expect matched-pair CI to be substantially narrower than independent samples CI
(e) There is no relationship between the coefficient of correlation and the width of the matched-pair CI versus the independent samples CI