Section (circle): 12:30 pm 2:30 pm 3:30 pm

Name: ___KEY___________________________________

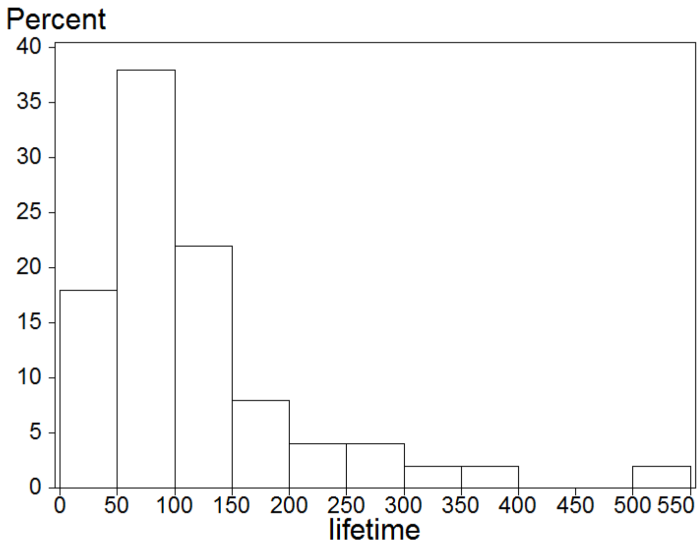
INSTRUCTIONS

1. Calculators (using the programmable features will not help you since ALL work including integration steps are required to receive full credit) and your cheat sheet are allowed. Exam is closed book/closed notes.
2. There are 5 pages front and back for the exam in addition to the cover sheet. You may tear off the last page (table) to make your work easier. However, this sheet must be returned.
3. All numeric answers must have at least 3 decimal places.
4. Work is required to receive credit. Partial credit will be given for work that is partially correct. Points will be deducted for incorrect work even if the final answer is correct.
5. Please attach your cheat sheet, table and any scrap paper used to the exam.
6. Good Luck!

Question	Possible	Score
1	21	
2	12	
3	15	
4	17	
5	17	
6	12	
7	14	
Total	108	

(21 pts.) 1. Multiple choice questions. Please circle your answer. There is only one correct answer for each choice. (3 pts. each)

a) The following is a histogram of drill lifetimes (number of holes that the drill can make before it breaks) when holes were drilled in a certain brass alloy.



i) Which of the following codes was used to create the histogram?

A)
```
proc univariate data=1a;
histogram lifetime;
run;
```
B)
```
proc univariate data=1a;
qqplot lifetime;
run;
```
C)
```
proc univariate data=1a;
histogram lifetime/endpoints=0 to 550 by 50;
run;
```

ii) What is the shape of the histogram?

A) negatively skewed B) symmetric C) positively skewed

b) In the following cases, indicate whether the situations provided are examples of experiments or observational studies.

i) In a study of the relationship between birthweight and race, birth records of babies born in Illinois were examined. The researchers found that the percentage of low birthweight babies among babies born to U.S.-born white women was much lower than the percentage of low birthweight babies among babies born to U.S.-born black women.

A) Experiment B) Observational Study

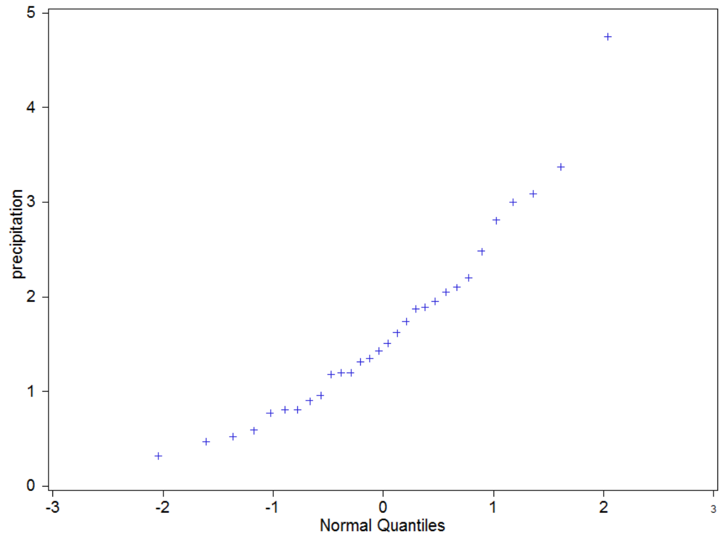
ii) In an article (Exercise 11.4 in the book), there is a scatter plot where x = rainfall volume (m^3) and y = runoff value (m^3) for a particular location in Austin ,Texas.

A) Experiment B) Observational Study

c) If you have performed an observational study, can you determine the causality directly from your results?

A) Yes B) No

d) The following is a QQplot for the precipitation in March for a 30-year period in Minneapolis-St. Paul.



i) Which of the following codes was used to generate the QQplot?

A)
```
proc univariate data=1c;
qqplot precipitation*group;
run;
```

This above code is not correct

B)
```
proc univariate data=1c;
qqplot precipitation;
run;
```

C)
```
proc univariate data=1c;
histogram precipitation;
run;
```

ii) Is it plausible that this data has a normal distribution?

A) Yes   B) No

This looks like a curve to me

(12 pts.) 2. The current in a certain circuit as measured by an ammeter has the following density function:

$$f(x) = \begin{cases} 0.075x + 0.2 & 3 \le x \le 5 \\ 0 & otherwise \end{cases}$$

(7 pts.) a) What is the mean of the current in this circuit?

$$\mu = \int_{-\infty}^{\infty} xf(x)dx = \int_{3}^{5} x(0.075x + 0.2)dx = \int_{3}^{5} (0.075x^2 + 0.2x)dx = \frac{0.75x^3}{3} + \frac{0.2x^2}{2}\Big|_3^5$$
$$= 3.125 + 2.5 - 0.675 - 0.9 = 4.05$$

Note: points were removed if you did not put in 'dx'

(5 pts.) b) Without solving it, write the equation (including the integral) that you would use to find the median of X

$$\int_{3}^{\tilde{\mu}} (0.075x + 0.2)dx$$

(15 pts.) 3. Suppose that 17% of students on a campus are mathematics majors and 19% of the students are computer science. In addition, no students are both mathematics majors and computer science majors.

(5 pts.) a) What is the probability that a particular student is a mathematics major or a computer science major.

P(M) = 0.17 P(CS) = 0.19
P(M or CS) = P(M) + P(CS) = 0.17 + 0.19 = 0.36

(5 pts.) b) If 25% of the students have a car on campus and a student's major has no affect on whether on or not they own a car, what is the probability that a student has a car and is a mathematics major?

P(M) = 0.17 P(C) = 0.25
P(M and C) = P(M) P(C) = (0.17) (0.25) = 0.0425

(5 pts.) c) If 8% of the students are mathematics majors and female, given that student is a mathematics major, what is the probability that the student is female?

P(M) = 0.17 P(F and M) = 0.08

$$P(F|M) = \frac{P(F \text{ and } M)}{P(M)} = \frac{0.08}{0.17} = 0.4706$$

(17 pts.) 4. The weekly amount spent for maintenance and repairs in a certain company has a normal distribution with μ = \$415 and σ = \$21.

(5 pts.) a) What is the probability that the cost with exceed \$450 which is the budgeted amount?

$$P(X > 450) = P\left(Z > \frac{450 - 415}{21}\right) = P(Z > 1.67) = 1 - P(Z \leq 1.67) = 1 - 0.9525 = 0.0475$$

Note: Points were removed if incorrect terminology was used. We are looking at probabilities in the z table.

(5 pts.) b) How large must a value be to be among the top 85.08% of the weekly costs?

a) P(Z > c*) = 0.8505 ==> P(Z ≤ c*) = 0.1495 ==> c* = -1.04

$$c^* = \frac{c - \mu}{\sigma} \Rightarrow c = \mu + c^*\sigma = 415 + (-1.04)(21) = 428.16$$

OR

b) P(Z ≤ c*) = 0.8505 ==> c* = 1.04

$$c^* = \frac{c - \mu}{\sigma} \Rightarrow c = \mu + c^*\sigma = 415 + (1.04)(21) = 436.84$$

(7 pts.) c) If the budget officer looks at the average of the amount spend for maintenance in the four weeks in a certain month, what is the probability that the average cost will exceed $450?

This is a sampling distribution problem.

$\mu_{\bar{x}} = \mu = 415$

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{21}{\sqrt{4}} = 10.5$$

$$P(\bar{X} > 450) = P\left(Z > \frac{450-415}{10.5}\right) = P(Z > 3.33) = 1 - P(Z \leq 3.33) = 1 - 0.9996 = 0.0004$$

(17 pts.) 5. The following data is the oxygen consumption (ml/kg/min) of 10 firefighters performing a fire-suppression simulation (sorted).

23.5	26.3	28.0	28.2	29.4	29.5	30.6	31.6	33.9	49.3
1	2	3	4	5	6	7	8	9	10

(6 pts.) a) Determine Q1, median and Q3. No work is required for full credit. However, no partial credit will be given if there is no work.

$$median = \frac{29.4 + 29.5}{2} = 29.45$$

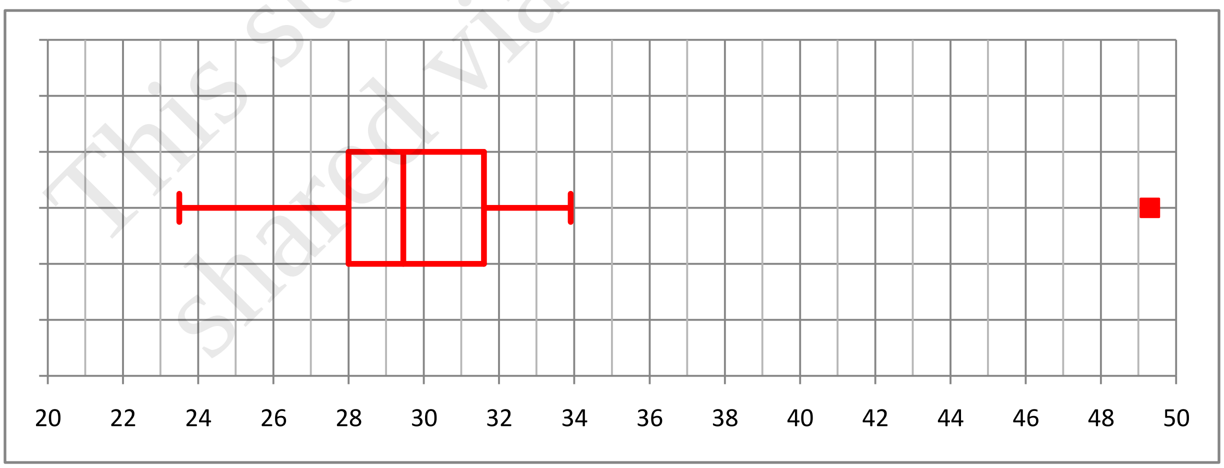
Q1 = 28.0 Q3 = 31.6

(6 pts.) b) Are there any outliers? Why or why not?

IQR = Q3 – Q1 = 31.6 – 28.0 = 3.6 1.5(IQR) = 1.5(3.6) = 5.4
lower: Q1 – 1.5(IQR) = 28.0 – 5.4 = 22.6 ==> no outliers
upper: Q3 + 1.5(IQR) = 31.6 + 5.4 = 37 ==> yes, one outlier, 49.3

(5 pts.) c) Draw an accurate boxplot for this data.

min = 23.6, Q1 = 28.0, median = Q2 = 29.45, Q3 = 31.6, max = 33.9, outlier: 49.3
Note: I graded on consistency here.



(12 pts.) 6. The College Board reports that 2% of the High School students who take the SAT each year receive special accommodations because of documented disabilities. Assume that we have a random sample of 25 students. Note: This is a binomial distribution. Hint: $\frac{n!}{1!(n-1)!} = n$.

(7 pts.) a) What is the probability that at most one of the students receive special accommodations?

n = 25 π = 0.02

$$P(X \leq 1) = p(0) + p(1) = \frac{25!}{0!\,25!} 0.02^0 0.98^{25} + \frac{25!}{1!\,24!} 0.02^1 0.98^{24} = 0.98^{25} + (25)(0.02)0.98^{24}$$
$$= 0.603 + 0.308 = 0.911$$

(5 pts.) b) What are the mean and the standard deviation of the number of students that receive special accommodations?

For a binomial distribution

μ = n π = (25)(0.02) = 0.5

$$\sigma = \sqrt{n\pi(1-\pi)} = \sqrt{(25)(0.02)(0.98)} = \sqrt{0.49} = 0.7$$

(14 pts.) 7. The number of hours that it rains at the Toronto airport has an exponential distribution with a mean of 3. Hint: What is the parameter for an exponential distribution?

(7 pts.) a) What is the probability that it will rain between 2 and 3 hours on one particular night?

For an exponential distribution, $\mu = \frac{1}{\lambda} \Longrightarrow \lambda = \frac{1}{\mu} = \frac{1}{3}$, $f(x) = \lambda e^{-\lambda x} = \frac{1}{3} e^{-\frac{x}{3}}$

$$P(2 < X < 3) = \int_2^3 \frac{1}{3} e^{-\frac{x}{3}} dx = -e^{-\frac{x}{3}} \Big|_2^3 = -e^{-1} + e^{-\frac{2}{3}} = 0.1455$$

(7 pts.) b) What is the probability that the average time that it will rain in the next 60 days is between 2 and 3 hours?

This is a sampling distribution problem.

$\mu_{\bar{x}} = \mu = 3$

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{\mu}{\sqrt{n}} = \frac{3}{\sqrt{60}} = 0.387$$

$$P(2 < \bar{X} < 453) = P\left(\frac{2-3}{0.387} < Z < \frac{3-3}{0.387}\right) = P(-2.58 < Z < 0) = P(Z < 0) - P(Z < -2.58)$$
$$= 0.5000 - 0.0049 = 0.4951$$