

**Statistics 300 : Spring 2013**

**Instructor : L. C. Larsen**

**Student name & ID#:**

---

**Student signature:**

---

**Exam : Unit 1**

**Time allowed : 2 hours and 5 minutes**

**Resources allowed:**

- == > Textbook (Author: Triola)**
- == > Notes/helps written by the student**
- == > Quiz and exam solutions written by instructor**
- == > Quiz and exam solutions written by the student**
- == > Calculator/laptop of choice (no outside messages)**
- == > Instructor at 916-346-6324**

**Resources not allowed:**

- == > Consultants other than the instructor**
- == > No phones, unless used as a calculator only**

(4 points and 2 points; 5 minutes)

1. A quiz has two questions. The first question is multiple choice with 6 possible choices: a, b, c, d, e, and f. The second question is "true" (T) or "false" (F). What is the sample space for the two questions together?

If an outcome is selected at random (all equally likely) from your sample space, what is the probability that your selection is wrong for the first problem and right for the second?

(4 points; 4 minutes)

2. A lock uses a touch pad with the letters A B and C and the digits 1 2 3 4 5 6 7 and 8. The unlock code begins with 2 of the letters and ends with 4 of the digits. Each letter and each digit can only be used one time, and every different arrangement is a different unlock code. How many different unlock codes are possible?

(3 points; 3 minutes)

3. If only the letters A, B, C, D, E, F, G, and H are used, how many sets of five letters can be selected if no letter is allowed to be used more than one time?  
Example: the sets {A F C B} and {F B A C} are the same.

4. Give a short definition of statistics (1 point; 1 minute):

\_\_\_\_\_

(4 points; 4 minutes)

5. For each situation below, select the appropriate statistical term from the list provided and write it in the blank next to the description or situation. Choose the term that is best connected to the underlined text in the description or situation.

|        |                  |             |                      |
|--------|------------------|-------------|----------------------|
| Terms: | 1. randomization | 4. blinding | 7. experimental unit |
|        | 2. replication   | 5. placebo  | 8. treatment         |
|        | 3. confounding   | 6. block    |                      |

\_\_\_\_\_ Doctors studied three different diets for children that can't eat dairy (milk) products. A total of 210 children were studied with equal numbers chosen from 5 different age groups. Each of the diets was assigned to 70 of the children selected at random from the 210 children in the experiment. No effort was made to control for ethnicity or gender. Experts observed the children's behavior in school and health experts monitored the children's condition.

\_\_\_\_\_ Doctors studied three different diets for children that can't eat dairy (milk) products. A total of 210 children were studied with equal numbers chosen from 5 different age groups. Each of the diets was assigned to 70 of the children selected at random from the 210 children in the experiment. No effort was made to control for ethnicity or gender. Experts observed the children's behavior in school and health experts monitored the children's condition.

\_\_\_\_\_ Doctors studied three different diets for children that can't eat dairy (milk) products. A total of 210 children were studied with equal numbers chosen from 5 different age groups. Each of the diets was assigned to 70 of the children selected at random from the 210 children in the experiment. No effort was made to control for ethnicity or gender. Experts observed the children's behavior in school and health experts monitored the children's condition.

\_\_\_\_\_ Doctors studied three different diets for children that can't eat dairy (milk) products. A total of 210 children were studied with equal numbers chosen from 5 different age groups. Each of the diets was assigned to 70 of the children selected at random from the 210 children in the experiment. No effort was made to control for ethnicity or gender. Experts observed the children's behavior in school and health experts monitored the children's condition.

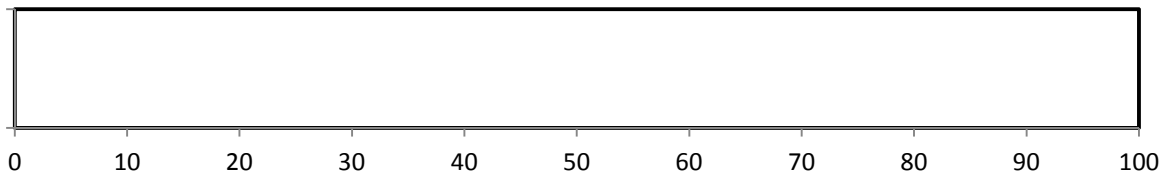
\_\_\_\_\_ Doctors studied three different diets for children that can't eat dairy (milk) products. A total of 210 children were studied with equal numbers chosen from 5 different age groups. Each of the diets was assigned to 70 of the children selected at random from the 210 children in the experiment. No effort was made to control for ethnicity or gender. Experts observed the children's behavior in school and health experts monitored the children's condition.

(4 points; 4 minutes)

6. A subway train has three safety systems that will stop the train in an emergency. The first safety system will work 90% of the time, the second will work 80% of the time, and the third will work 70% of the time. What is the probability that the train will not stop safely the next time an emergency happens? (Assume that the three systems operate independently of each other.)

(5 points and 4 points; 8 minutes)

7. (a) For the set of 53 values shown below in sorted order, prepare a Boxplot inside the rectangle that is above the number line.



|      |      |      |      |      |      |      |      |      |      |
|------|------|------|------|------|------|------|------|------|------|
| 1.9  | 2.0  | 2.2  | 2.8  | 3.3  | 3.6  | 3.8  | 4.0  | 4.9  | 5.9  |
| 6.9  | 8.7  | 9.3  | 9.5  | 9.7  | 13.3 | 13.4 | 13.9 | 14.7 | 15.7 |
| 20.1 | 21.1 | 21.2 | 21.3 | 22.2 | 24.5 | 28.7 | 29.5 | 29.9 | 30.8 |
| 34.7 | 35.1 | 40.7 | 46.6 | 48.3 | 48.3 | 50.5 | 53.1 | 53.6 | 54.3 |
| 54.8 | 59.1 | 71.2 | 71.3 | 71.8 | 76.4 | 77.3 | 82.9 | 87.2 | 92.7 |
| 93.0 | 95.0 | 96.2 |      |      |      |      |      |      |      |

- (b) For the set of 53 values shown above in sorted order, what percentile is represented by the value 48.3?

(5 points; 5 minutes)

8. The scores of two different tests have bell-shaped distributions. The scores for Test A have a mean of 110 points and a standard deviation of 42 points. The scores for Test B have a mean of 330 and a standard deviation of 68 points. A student takes Test A and earns a score of 91 points. Another student takes Test B and earns a score of 350 points. Answer each of the questions below, including the "Why".

|      |       |   |
|------|-------|---|
| TRUE | FALSE | The score on Test B is more unusual than the score on Test A.<br>Why? |
|------|-------|---|

|      |       |  |
|------|-------|--|
| TRUE | FALSE | The score on Test B is relatively higher than the score on Test A.<br>Why? |
|------|-------|--|

(8 points; 7 minutes)

9. Complete the columns in the "Frequency Distribution" table using the data values given below.

Frequency Distribution

| Class Limits |       | Tally | Frequency | Relative Frequency | Cumulative Frequency | Cumulative Relative Frequency |
|--------------|-------|-------|-----------|--------------------|----------------------|-------------------------------|
| Lower        | Upper |       |           |                    |                      |                               |
| 1            | 40    |       |           |                    |                      |                               |
| 41           | 80    |       |           |                    |                      |                               |
| 81           | 120   |       |           |                    |                      |                               |

| Class Midpoint | Class Boundary |
|----------------|----------------|
|                |                |
|                |                |
|                |                |
|                |                |

|       |      |      |      |       |      |
|-------|------|------|------|-------|------|
| Data: | 91.8 | 10.2 | 58.8 | 108.0 | 80.4 |
|       | 14.6 | 69.6 | 13.2 |       |      |

|             |
|-------------|
| Class Width |
|             |

(5 points; 6 minutes)

10. The Army wanted to know the 95<sup>th</sup> percentile ( $P_{95}$ ) of the scores of ALL its soldiers in a test of physical endurance that simulates combat. Because it would be too expensive to put all the soldiers through the test, the Army selected 800 soldiers at random to do the test and then they used the 95<sup>th</sup> percentile of the test scores as their estimate of the population parameter they want to know.

Use the information in the "story" to answer the following:

(a) What is the population of interest?

---

---

---

(c) What parameter was important for the Army to know?

---

---

(d) What statistic did the Army use instead?

---

---

(e) If all the 800 test scores were sorted from smallest to largest, at what location in the list would you find the statistic that the Army used?

Location = the \_\_\_\_th position in the list.

(8 points; 10 minutes)

11. For each discrete probability distribution, calculate the mean, variance, and standard deviation. Use the columns in the tables as you wish to use them.

| x   | P(x) |  |  |  |  |
|-----|------|--|--|--|--|
| 168 | 0.26 |  |  |  |  |
| 185 | 0.18 |  |  |  |  |
| 196 | 0.29 |  |  |  |  |
| 216 | 0.27 |  |  |  |  |

| x | P(x) |  |  |  |  |
|---|------|--|--|--|--|
| 0 | 0.25 |  |  |  |  |
| 1 | 0.50 |  |  |  |  |
| 2 | 0.15 |  |  |  |  |

Write the formulas for the mean, the variance, and the standard deviation of a discrete probability distribution.

$\mu =$  \_\_\_\_\_

$\sigma^2 =$  \_\_\_\_\_

$\sigma =$  \_\_\_\_\_

(14 points; 8 minutes)

12. Use the data below to determine the value of each statistic. Write an expression for each statistic or describe how it is calculated in principle (do NOT describe how to use the calculator to determine the result).

| Data   | Expression or Description | Value of statistic |
|--|---------------------------|--------------------|
| 35<br>32<br>31<br>30<br>40<br>32<br>30<br>32 | median                    |                    |
|  | mode                      |                    |
|  | variance                  |                    |
|  | range                     |                    |
|  | standard deviation        |                    |
|  | midrange                  |                    |
|  | mean                      |                    |



(6 points; 5 minutes)

13. In a major city, the rate of illicit drug use by high school students is 12%. If a random sample of 10 high school students is tested, what is the probability that exactly three of the 10 students will test positive for illegal drug use?

(6 points; 6 minutes)

14. A large study found that 13% of the high school students in a major city have used illegal drugs two or more times during the last year. If 390 of 20,000 high school students in the city are selected at random and tested for illegal drug use, would it be unusual to find 65 positive tests in the sample of 390? Circle "YES" or "NO" and show "Why". The "Why" is worth 5 of the 6 points!

|     |      |
|-----|------|
| YES | Why? |
| NO  |      |

(5 points; 5 minutes)

15. Based on many years of experience, an investment advisor tells you that your investments will earn \$70,000 with 20% probability, earn \$0 with 50% probability, and lose \$12,000 with 30% probability. If your advisor is correct, what is the expected value of your income from investments this year? Showing "why" is worth 4 of the 5 points.

Expected value = \_\_\_\_\_

Why?

(6 points; 6 minutes)

16. Circle the correct answer for each situation. If a part is underlined, focus on that part only.

A store manager wants to survey customers in a representative way. The survey is designed to have 25 days selected randomly next year. On each of the 25 days, all customers will answer three questions and receive \$5 in store credit.

|                   |             |
|-------------------|-------------|
| Simple Random     | Systemmatic |
| Stratified Random | Cluster     |
| Convenience       | Census      |

A store manager wants to survey customers, so a bell is installed that will ring at a random time in each hour. When the bell sounds, each cashier records the next customer's gender and the total cost of the items purchased.

|                   |             |
|-------------------|-------------|
| Simple Random     | Systemmatic |
| Stratified Random | Cluster     |
| Convenience       | Census      |

A store manager wants to survey customers in a representative way. The survey is designed to make it easy for all customers to take part, so a questionnaire is given to every customer along with their receipt. Questionnaires can be returned on the customer's next visit.

|                   |             |
|-------------------|-------------|
| Simple Random     | Systemmatic |
| Stratified Random | Cluster     |
| Convenience       | Census      |

A store manager designs a representative survey to learn about the store's customers. A systematic sample of 800 customers requests information on the customer's gender, age (in years), annual income, height, weight, zipcode, and number of family members that live with them.

|                                  |
|----------------------------------|
| Qualitative Nominal              |
| Qualitative Ordinal              |
| Quantitative Discrete Interval   |
| Quantitative Discrete Ratio      |
| Quantitative Continuous Interval |
| Quantitative Continuous Ratio    |

A store manager designs a representative survey to learn about the store's customers. A systematic sample of 800 customers requests information on the customer's gender, age (in years), annual income, height, weight, zipcode, and number of family members that live with them.

|                                  |
|----------------------------------|
| Qualitative Nominal              |
| Qualitative Ordinal              |
| Quantitative Discrete Interval   |
| Quantitative Discrete Ratio      |
| Quantitative Continuous Interval |
| Quantitative Continuous Ratio    |

A store manager designs a representative survey to learn about the store's customers. A systematic sample of 800 customers requests information on the customer's gender, age (in years), annual income, height, weight, zipcode, and number of family members that live with them.

|                                  |
|----------------------------------|
| Qualitative Nominal              |
| Qualitative Ordinal              |
| Quantitative Discrete Interval   |
| Quantitative Discrete Ratio      |
| Quantitative Continuous Interval |
| Quantitative Continuous Ratio    |

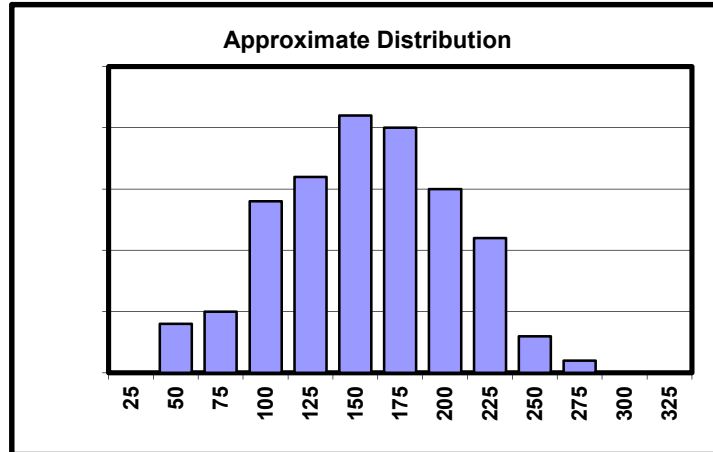
(6 points; 5 minutes)

17. Use the information provided to determine two numbers between which you expect to find 68% of the data.

Mean = 144.6  
Standard Deviation = 49.8

Lower Value = \_\_\_\_\_

Upper Value = \_\_\_\_\_



18. Use the data in the following table to answer parts (a) through (d).

(1 point, 1 minute)

(a) What is the probability that a randomly selected person from this sample will be from Georgia?

| Employment Type for People in Sample |      |         |          |         |       |
|--------------------------------------|------|---------|----------|---------|-------|
| State                                | Farm | Factory | Gov'ment | Service | Total |
| Alabama                              | 296  | 217     | 289      | 262     | 1064  |
| Florida                              | 265  | 266     | 232      | 209     | 972   |
| Georgia                              | 205  | 279     | 217      | 286     | 987   |
| Louisiana                            | 297  | 261     | 290      | 298     | 1146  |
| Mississippi                          | 242  | 272     | 265      | 229     | 1008  |
| Total                                | 1305 | 1295    | 1293     | 1284    | 5177  |

(3 points, 3 minutes)

(b) What is the probability that a person randomly selected from this sample will be a farm worker given that the person is from Alabama?

(3 points, 3 minutes)

(c) What is the probability that a person randomly selected from this sample will be from Mississippi or work in the service sector of the economy?