DIRECTIONS:

- This exam contains 44 multiple choice questions, each worth 2 points.
- Circle the correct response for each question. Make sure that your answer is clearly marked. You will not receive partial credit for any work done.
- The standard normal table appears on Page 11 of this exam.
- This is a closed-book, closed notes examination. You may use a calculator if you wish. However, cell phones are not permitted for use in any way.
- Any discussion or otherwise inappropriate communication between examinees, as well as the appearance of any unnecessary material or cell-phone usage, will be dealt with severely. Violations may result in an "F" for this exam, "F" for the class, suspension, or expulsion.
- This exam is worth a total of 88 points. **Print** your name **at the top of this page in the upper right hand corner**. *Good Luck!!*

HONOR PLEDGE FOR THIS EXAM:

After you have finished the exam, please read the following statement and sign your name below it.

> *I promise that I did not discuss any aspect of this exam with anyone other than the instructor, that I neither gave nor received any unauthorized assistance on this exam, and that the work presented herein is entirely my own.*

1. We have learned that regression models can be used to make predictions. Which statement is **false**?
(a) Prediction outside the range of the available data can be risky.
(b) The prediction depends on the model.
(c) A prediction will always be within the margin of error.
(d) Prediction works best when the model fits the data well.

2. In class, we talked about a particular graph which has been described as **quite possibly the best graphical display ever.** What real-life phenomenon did the graph depict?
(a) the growing debate over teaching creationism in high school science courses
(b) the gravitational pull of Pluto's moons
(c) the history of the HIV-AIDS epidemic
(d) Napoleon's army and its invasion of Russia

3. Which of the following graphical displays uses the **5-Number Summary**?
(a) time plot
(b) boxplot
(c) histogram
(d) pie chart

4. True or False: The **mean** of a density curve is the "equal-areas point," the point that divides the area under the curve in half.
(a) True
(b) False

5. Which statement concerning correlation is **false**?
(a) The correlation measures the strength of the straight-line association between two variables.
(b) The correlation's value can not be less than 0.
(c) The correlation's value can be strongly affected by outliers.
(d) The correlation ignores the distinction between explanatory and response variables.

6. Quiz scores for a large undergraduate class follow a normal distribution with mean 70 and standard deviation 10. Your score is 88. What is your **standardized score**?
(a) 0.04
(b) 0.96
(c) 1.8
(d) -1.8

7. True or False. Probability deals with **short-term behavior**; not long-term regularity.
(a) True
(b) False

8. These data represent the number of accidents (per month) at a busy intersection in Columbia, SC, for $n = 10$ consecutive months.

```
6  4  3  4  6  6  4  6  4  7
```

What is the **sample mean**?
(a) 3
(b) 4
(c) 5
(d) 6

9. Refer to the data in Question 8. What is the **sample standard deviation**?
(a) 0.2
(b) 1.3
(c) 1.8
(d) 2.0

10. Refer to the data in Question 8. What is the **sample range**?
(a) 1
(b) 2
(c) 3
(d) 4

11. If a density curve is **skewed to the right** (the high side), then
(a) the mean will be less than the median.
(b) the mean will be greater than the median.
(c) the mean and median will be equal.

12. Which of the following values is closest to the **20th percentile** for the standard normal distribution?
(a) 2.1
(b) -2.1
(c) -0.8
(d) 0.8

13. A stem and leaf display is shown below. The stem is the tens place and the leaf is the units place.

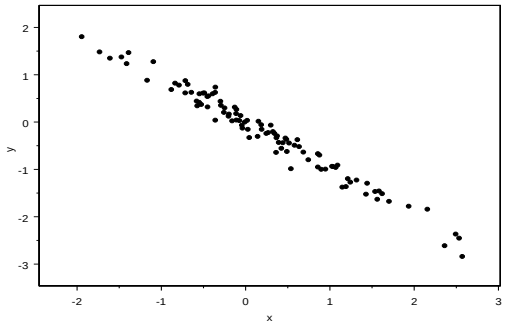
4	1	3				
5	0	5	5	5	9	9
6	2	5	7	9		
7	8					

What is the **median** of this distribution?
(a) 57
(b) 59
(c) 62
(d) 78

14. When examining a quantitative data distribution graphically (e.g., with a histogram, etc.), what **four things** do we try to characterize?
(a) center, spread, shape, outliers
(b) center, spread, shape, margin of error
(c) spread, shape, outliers, randomization
(d) equipoise, shape, outliers, randomization

15. True or False. We know that smoking causes cancer, because numerous experiments have demonstrated this fact.
(a) True
(b) False

16. In the scatterplot below, which value is closest to the **correlation**?



(a) $r = 0.70$
(b) $r = 1.00$
(c) $r = -0.98$
(d) $r = -0.08$

17. True or False. In regression, a value of r^2 **close to 0** is often taken as evidence that the predictions made are going to be precise.
(a) True
(b) False

18. Which statement concerning density curves is **false**?
(a) The curve can never fall below 0.
(b) The area under the curve is 1 (or 100 percent).
(c) The area under the curve between two values gives the proportion of observations that will fall in that range.
(d) All density curves we see in real life will be symmetric.

19. Which plot gives an impression of how a **single quantitative variable** is distributed?
(a) histogram
(b) pie chart
(c) scatterplot
(d) two-way table

20. Daily high temperatures (in Fahrenheit) recorded in Nome, AK, follow a normal distribution with mean 40 and standard deviation 5. Approximately what proportion of days will have a high temperature **exceeding** 45 degrees?
(a) 0.95
(b) 0.68
(c) 0.16
(d) 0.02

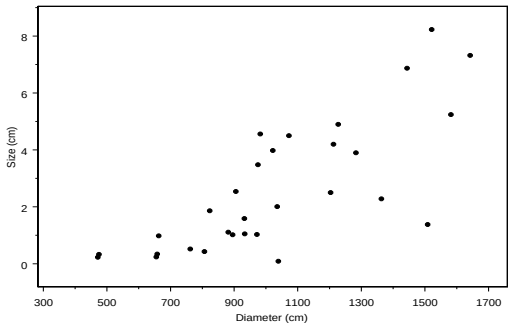
21. Which statement concerning the sample standard deviation is **true**?
(a) The standard deviation is always between -1 and 1.
(b) The standard deviation is measured in original units of the data.
(c) The larger the standard deviation, the smaller the spread.
(d) Outliers can not affect the value of the standard deviation.

22. For a scatterplot of data, the **least squares regression line** is the line that
(a) goes through the largest number of points.
(b) minimizes the sum of the vertical distances from the points to the line.
(c) gives the best predictions.
(d) None of the above answers are true.

23. True or False. We use sample statistics to estimate population parameters.
(a) True
(b) False

24. Which of the following variables is **quantitative**?
(a) the amount of fertilizer (in pounds) applied in an agricultural experiment
(b) political affiliation
(c) the NFL team that Terrell Owens plays for next year
(d) marital status of USC professors

25. An entomological experiment was conducted to study the survivability of stalk borer larvae. It was of interest to develop a model relating the mean size of larvae (cm) as a function of the stalk head diameter (cm). The scatter plot of the data is below.



The **least squares regression line** for these data has a _____ slope and a _____ y-intercept.
(a) positive, positive
(b) positive, negative
(c) negative, positive
(d) negative, negative

26. In class, we talked about the "birthday problem." What was the **main point** of that example?
(a) Unusual events sometimes are not that unusual.
(b) Subject identifiers could violate data confidentiality.
(c) Most birthdays, surprisingly, occur in odd-numbered months (e.g., January, March, May, etc.).
(d) None of the above.

27. Consider the following straight-line regression model, obtained from an SRS of first-year students at a small school in Menomonie, WI.

$$\text{GPA} = -1.2 + 0.045(\text{IQ})$$

What GPA would you **predict** for a student with an IQ equal to 100?
(a) 1.3
(b) 2.3
(c) 3.3
(d) 4.3

28. Which statement concerning personal probability assignments is **true**?
(a) A personal probability is subjective and is sometimes based on opinion.
(b) Personal probabilities are based on the notion of repeated trials.
(c) Personal probabilities are often based on scientific evidence.
(d) Each of these statements is true.

29. What is the **median** of the standard normal distribution?
(a) 0
(b) 1
(c) This can not be determined.

30. What value is closest to the **interquartile range** for the standard normal distribution?
(a) 0
(b) 0.5
(c) 1.3
(d) 3.0

31. **Bias** is a property that concerns the _____ of a sampling distribution.
(a) center
(b) spread
(c) shape

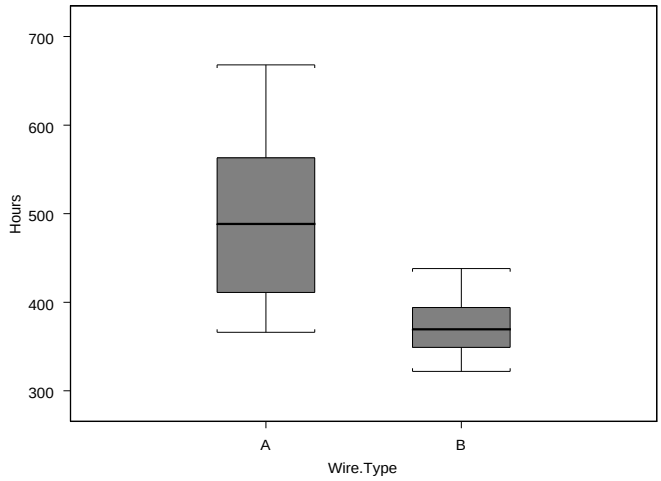
32. What characteristics would we like a statistic (estimator) to have?
(a) small variability, small bias
(b) small variability, large bias
(c) large variability, small bias
(d) large variability, large bias

33. Historical evidence suggests that SAT scores are normally distributed with mean 1000 and standard deviation 180. What score do you have to make to be in the **top 1 percent**?
(a) 586
(b) 1180
(c) 1234
(d) 1414

34. We have computed the probability of an event to be 0.999. Which of the following statements is **correct**?
(a) If the event did occur, we would consider it to be unusual.
(b) The event will always occur.
(c) The event is very likely to occur.
(d) We would expect the event to occur about 9.99 percent of the time.

35. In class, we talked about the Chicago-fire truck example. What was the **main point**?
(a) Fires are often set in Chicago because fans are disgusted with the Cubs' shameful playoff record.
(b) Standard deviation is often misinterpreted by news reporters.
(c) Correlation does not necessarily imply that the variables involved have a causal link between them.
(d) None of the above.

The performance of two types of wires is compared in the production of light bulbs (TYPE A and TYPE B wire types). To compare the two types, 40 production runs with TYPE A and 40 production runs with TYPE B were performed. For each run, we record the number of operational hours. Use this information to answer Questions 36-40.



36. Which wire distribution has a smaller **interquartile range**?
(a) Type A
(b) Type B

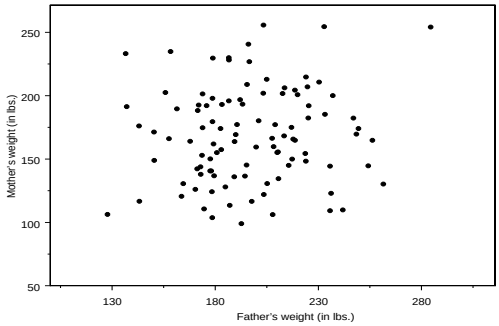
37. Which wire distribution has a smaller **standard deviation**?
(a) Type A
(b) Type B

38. The standard deviation for the **Type B** data distribution is closest to
(a) 2
(b) 20
(c) 200

39. The median for the **Type B** data distribution is closest to
(a) 300
(b) 375
(c) 500

40. For the **Type A** data distribution, approximately what percentage of the observations are larger than 400?
(a) 20 percent
(b) 40 percent
(c) 60 percent
(d) 80 percent

41. Below is a scatterplot displaying height data from $n = 120$ mother-father couples. The correlation is closest to



(a) $r = 0.00$
(b) $r = -2.0$
(c) $r = -0.90$
(d) $r = 0.90$

42. The distribution of male birthweights is normal in shape, with mean 7.6 pounds and standard deviation 1.2 pounds. Approximately what percentage of birthweights will be **between** 5.8 and 9.4 pounds?
(a) 68 percent
(b) 86 percent
(c) 95 percent
(d) 97 percent

43. What does the **square of the correlation** r^2 measure?
(a) the proportion of the total variation in the response data explained by the straight line relationship with the explanatory variable.
(b) the amount of the variability explained by the standard deviations of the response and explanatory variables.
(c) the amount of error that is incurred from using a regression model to make predictions.
(d) None of the above.

44. A class has four exams, each worth 100 points. After the first three exams, you have an average of 85. What score do you need to make on the last exam to have a four-exam average of **exactly** 80?
(a) 65
(b) 75
(c) 85
(d) 95

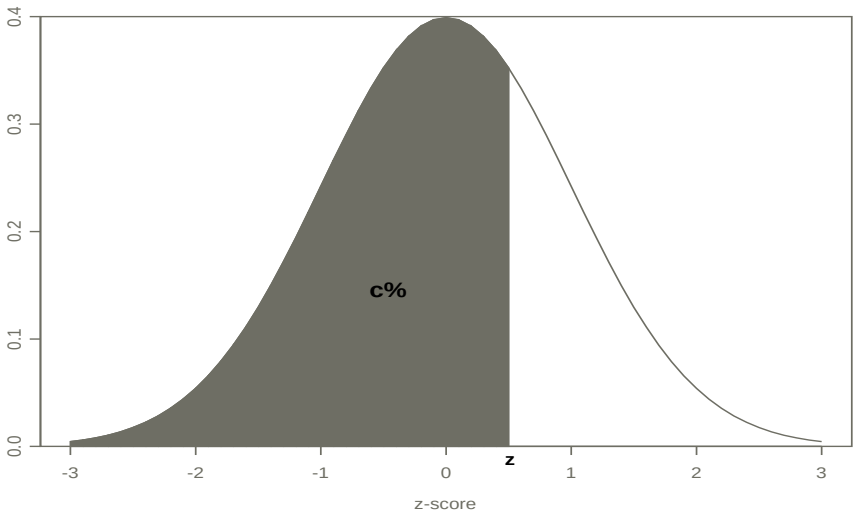


Table 1: *Percentiles of the standard normal distribution.*

Standard score (z)	Percentile (c)	Standard score (z)	Percentile (c)	Standard score (z)	Percentile (c)
-3.4	0.03	-1.1	13.57	1.2	88.49
-3.3	0.05	-1.0	15.87	1.3	90.32
-3.2	0.07	-0.9	18.41	1.4	91.92
-3.1	0.10	-0.8	21.19	1.5	93.32
-3.0	0.13	-0.7	24.20	1.6	94.52
-2.9	0.19	-0.6	27.42	1.7	95.54
-2.8	0.26	-0.5	30.85	1.8	96.41
-2.7	0.35	-0.4	34.46	1.9	97.13
-2.6	0.47	-0.3	38.21	2.0	97.73
-2.5	0.62	-0.2	42.07	2.1	98.21
-2.4	0.82	-0.1	46.02	2.2	98.61
-2.3	1.07	0.0	50.00	2.3	98.93
-2.2	1.39	0.1	53.98	2.4	99.18
-2.1	1.79	0.2	57.93	2.5	99.38
-2.0	2.27	0.3	61.79	2.6	99.53
-1.9	2.87	0.4	65.54	2.7	99.65
-1.8	3.59	0.5	69.15	2.8	99.74
-1.7	4.46	0.6	72.58	2.9	99.81
-1.6	5.48	0.7	75.80	3.0	99.87
-1.5	6.68	0.8	78.81	3.1	99.90
-1.4	8.08	0.9	81.59	3.2	99.93
-1.3	9.68	1.0	84.13	3.3	99.95
-1.2	11.51	1.1	86.43	3.4	99.97