Stat 205 Exam I; Feb. 20, 2014

1. For "well-behaved" data sets the empirical rule says that certain percentages of observations are within 1, 2, and 3 standard deviations of the mean. These percentages are
   (a) 65%, 95%, and 99%.
   (b) 68%, 90%, and 99%
   (c) 68%, 95%, and >99%.
   (d) None of the above.

2. Which of the following is correct?
   (a) The mean is pulled further in the direction of skew than the median.
   (b) The median is pulled further in the direction of skew than the mean.
   (c) The median is always larger than the third quartile.
   (d) The mean is a good measure of center for highly skewed data sets.

3. The five number summary for a data set is 3, 7, 8, 9, 10. That is, min $= 3$, $Q_1 = 7$, $\tilde{y} = 8$, $Q_3 = 9$, and max $= 10$. Which of the following is true?
   (a) 3 and 10 are both outliers.
   (b) 3 is an outlier.
   (c) This data set does not have outliers.
   (d) 3, 7, and 10 are all outliers.

4. The U.S. Office of Management and Budget collects data on race, which falls into one of five categories: White, Black or African American, American Indian or Alaska Native, Asian, and Native Hawaiian or Other Pacific Islander. Race is an example of what kind of variable?
   (a) numeric discrete.
   (b) numeric continuous.
   (c) categorical ordinal.
   (d) cateogrical nominal.

5. Data on the number of major seizures suffered by $n = 20$ epilepsy patients over eight weeks are sorted from smallest to largest.

   0 0 0 0 0 0 0 0 0 0 1 4 5 5 5 6 6 7 7 9.

   Which of the following is correct?
   (a) $\bar{y} = 2.75$ and $\tilde{y} = 0.5$.
   (b) These data are skewed to the left.
   (c) The most common outcome was zero seizures.
   (d) Both (a) and (c) are correct.

6. A simple random sample is a sample where
   (a) the mean is pulled larger than the median.
   (b) each experimental unit is chosen in prespecified proportions according to gender, race, etc.
   (c) we can expect bimodality.
   (d) each experimental unit has the same probability of being chosen.

7. A group of college students were surveyed to learn how many times they had visited a denstist in the previous year. Let Y be the number of dentist visits in a year for a randomly selected student; the study found $\Pr\{Y = 0\} = 0.15$, $\Pr\{Y = 1\} = 0.50$, and $\Pr\{Y = 2\} = 0.35$. The mean of Y is
   (a) $\mu_Y = 1$ visit.
   (b) $\mu_Y = 0.33$ visit.
   (c) $\mu_Y = 2$ visits.
   (d) $\mu_Y = 1.2$ visits.

8. Which of the following is correct?
   (a) A normal random variable is continuous.
   (b) A binomial random variable is discrete.
   (c) For $Y \sim \text{bin}(n, p)$, $\mu_Y = np$.
   (d) All of these are correct.

The following table cross-classifies 6549 subjects living in Massachusetts according to health risk (stressed or not stressed) and income (low, medium, or high). Use this table to answer the next seven questions.

	Income			
Stress level	Low	Medium	High	Total
Stressed	526	274	216	1016
Not Stressed	1954	1680	1899	5533
Total	2480	1954	2115	6549

9. What is the probability that someone in this study is stressed?

   (a) 0.298.
   (b) 0.925.
   (c) 0.102.
   (d) 0.155.

10. Given that someone has high income, what is the probability that they are stressed, i.e. Pr{stressed|high income}?

    (a) 0.298.
    (b) 0.925.
    (c) 0.102.
    (d) 0.155.

11. Is income level independent of being stressed?

    (a) Yes.
    (b) No.
    (c) Cannot tell from the table.
    (d) Both (a) and (b).

12. What is the probability of someone having low income or being not stressed?

    (a) 0.298.
    (b) 0.925.
    (c) 0.102.
    (d) 0.155.

13. What is the probability of someone having low income and being not stressed?

    (a) 0.298.
    (b) 0.925.
    (c) 0.102.
    (d) 0.155.

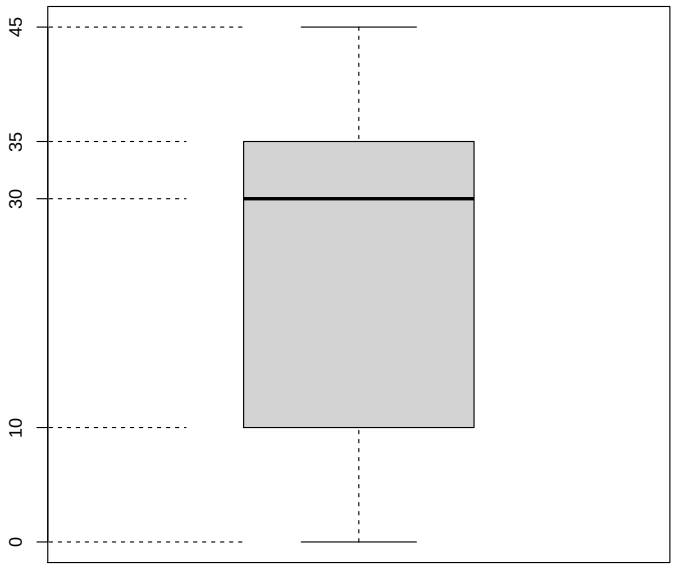
14. Given that someone is not stressed, what is the probability that they have high income?

    (a) 0.323.
    (b) 0.213.
    (c) 0.343.
    (d) 0.845.

15. In this study, income is what type of variable?

    (a) numeric discrete.
    (b) numeric continuous.
    (c) categorical ordinal.
    (d) cateogrical nominal.

An experiment was carried out to see how long it takes toddlers aged 2–3 years to knock over a pile of blocks (in seconds). Use the boxplot for this data set, below, to answer the next five questions.



16. The interquartile range for these data is

    (a) 35 seconds.
    (b) 45 seconds.
    (c) 25 seconds.
    (d) not computable from the boxplot.

17. 75% of the observations are less than

    (a) 35 seconds.
    (b) 45 seconds.
    (c) 25 seconds.
    (d) not computable from the boxplot.

18. 75% of the observations are greater than

    (a) 0 seconds.
    (b) 10 seconds.
    (c) 30 seconds.
    (d) not computable from the boxplot.

19. The upper fence for these data is

    (a) 45 seconds.
    (b) 25 seconds.
    (c) -27.5 seconds.
    (d) 72.5 seconds.

20. Which of the following is true?

    (a) There was at least one child who knocked the blocks over immediately.
    (b) There was at least one child who took over a minute to knock over the blocks.
    (c) The children typically took 10 seconds to knock over the blocks.
    (d) None of these are correct.

The brain weights of a population of adult Swedish males is normal with mean 1400 gm and standard deviation 100 gm. Use the following R code to answer the next three questions.

```
> pnorm(1325,1400,100)
[1] 0.2266274
> pnorm(1475,1400,100)
[1] 0.7733726
> qnorm(0.1,1400,100)
[1] 1271.845
> qnorm(0.9,1400,100)
[1] 1528.155
```

21. What proportion of brain weights are greater than 1325 grams?

    (a) 0.227.
    (b) 0.773.
    (c) 0.546.
    (d) 1271.8 gm.

22. What proportion of brain weights are between 1325 and 1475 grams?

    (a) 0.227.
    (b) 0.773.
    (c) 0.546.
    (d) 1271.8 gm.

23. 10% of brain weights are greater than

    (a) 0.462.
    (b) 0.125.
    (c) 1271.8 grams.
    (d) 1528.2 grams.

Bell (2007) notes that the percentage of eggs cracked after sizing and packing is 1.2% in a certain population, or $p = 0.012$. Say you buy a dozen eggs, $n = 12$ and let Y count the number of eggs (out of 12) that are cracked. Use the following R code to answer the next four questions.

```
> dbinom(0,12,0.012)
[1] 0.8651339
> dbinom(1,12,0.012)
[1] 0.1260924
> dbinom(2,12,0.012)
[1] 0.008423176
```

24. What is the probability that no eggs are cracked, $\Pr\{Y = 0\}$?

    (a) 0.865.
    (b) 0.126.
    (c) 0.008.
    (d) 0.135.

25. What is the probability that *at least* one egg is cracked?

    (a) 0.865.
    (b) 0.126.
    (c) 0.008.
    (d) 0.135.

26. The mean number of cracked eggs μ_Y is

    (a) 0.144 egg.
    (b) 1 egg.
    (c) 6 eggs.
    (d) none of the above.

27. Y is an example of a

    (a) normal random variable.
    (b) geometric random variable.
    (c) Poisson random variable.
    (d) binomial random variable.