Math 263, Section 5 **Name**____
Second Test Spring 2014
75 minutes; total 125 points.

1. (**12 points**) The Physicians' Health Study of 22,000 male physicians attempted to determine whether aspirin prevented heart attacks. A randomly selected group of 11,000 physicians took an aspirin every day, while the other 11,000 took a placebo. At the end of the study, the physicians who had taken aspirins had had significantly fewer heart attacks.

(a) Was this study an experiment _X___ or an observational study____? (*Check one*)

(b) Which of the following statements explain why we *cannot* conclude from this study that *all* adults should take an aspirin every day?

Mark each statement **T** (True) if it provides a valid explanation, or **F** (False) if it does not:

__T__The study included only physicians and different results might occur in individuals in other occupations

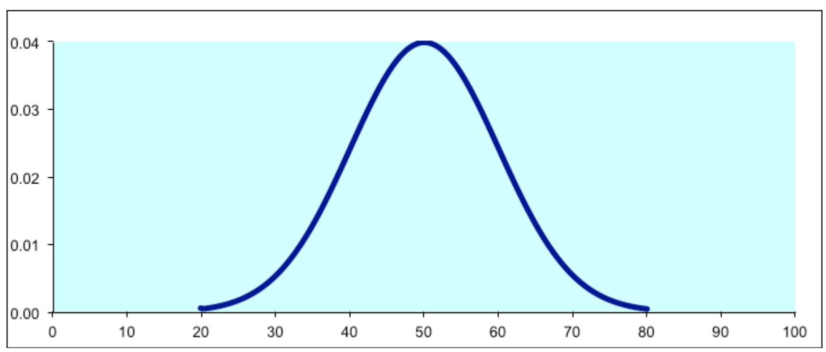
__T__The study included only males and different results might occur in females.

__T__Although aspirin may reduce heart attacks, it maybe harmful to health in other ways.

__F__The physicians given aspirin were selected randomly.

2. (**5 points**) Which of the following is the best estimate of the standard deviation of the following distribution?[1] (*Circle one answer*)

(a) 5
(b) 10
(c) 30
(d) 40
(e) 50
(f) 60



3. (**5 points**) A school has 20 classrooms with 30 students in each. The School Board asks for a random sample of 40 students in the school. The principal directed a teacher in each of the 20 classrooms to select randomly two of the students in that class; these 40 students constitute the sample.[2] Does this procedure give a simple random sample of students from the school? (*Circle one answer*)

(a) No, because not every sample of 40 had the same chance of being selected.

(b) No, because not all the students had the same chance of being selected.

(c) No, because the teachers were not randomly selected.

(d) Yes, because each student had the same chance of being selected.

(e) Yes, because each student was randomly selected in the classroom.

[1] CB, 1997
[2] Based on CB, 1998.

4.

(a) The probability that a shipment is accepted is the probability that there are no defective items in the sample of 10. Since the probability of any item being non-defective is 98%,

$$\text{Probability all acceptable} = (0.98)^{10} = 0.8171 = 81.71\%.$$

(b) Since a shipment is rejected if it is not accepted,

$$\text{Probability} = 1 - 0.8171 = 0.1829 = 18.29\%.$$

(c) The effect is:

(i) A new program is instituted at the suppliers manufacturing plant, thereby reducing the proportion of defective parts in the original shipment. (All other values keep their original values.) What is effect on the probability of acceptance you found part (a)?

Increases: __X__ Decreases: ____ Remains same:____ (check one; no reasons needed.)

(ii) You hire a new quality control engineer, who increases the number of parts sampled. (All other values keep their original values.) What is the effect on the probability of acceptance you found part (a)?

Increases: ____ Decreases: __X__ Remains same:____ (check one; no reasons needed.)

5. We are interested in the random variable $M = A - J$, where both A and J are normally distributed.

(a) The mean is

$$E(M) = E(A - J) = E(A) - E(J) = 47 - 43 = 4 \text{ pounds.}$$

(b) The variance is given by $V(M) = V(A - J) = V(A) + V(J) = 6^2 + 5^2 = 61$. Thus

$$SD(M) = \sqrt{V(M)} = \sqrt{61} = 7.81 \text{ pounds.}$$

(c) The probability that the Jersey cow gives more milk than the Ayrshire is $P(M < 0)$. Finding the z-value of $M = 0$ gives

$$z = \frac{0 - 4}{7.81} = -0.51.$$

The probability $P(M < 0) = P(Z < -0.51) = 0.3050 = 30.5\%$.

6. Let Z be the event that a case is caused by the Ebo-Z virus and let D be the event that an Ebola patient dies. Completing the two-way table, we have

Ebola cases, 1976-2012	Ebola subtypes		
	Ebo-Z	Other subtype	Totals
Deaths	1092	469	1561
Lived	299	459	758
Total number of cases	1391	928	2319

(a) Probability an Ebola case is caused by Ebo-Z is

$$P(Z) = \frac{1391}{2319}.$$

(b) Larger, because the fact that the patient died makes it more likely that the Ebola was caused by the most lethal variety, Ebo-Z.

(c) We have,

(i) This is the probability that an Ebola case died.

$$P(D) = \frac{1561}{2319} = 0.673 = 67.3\%.$$

(ii) This is the probability that an Ebola death was caused by Ebo-Z.

$$P(Z|D) = \frac{1092}{1561} = 0.70 = 70\%.$$

(iii) This is the death rate for the Ebo-Z strain.

$$P(D|Z) = \frac{1092}{1391} = 0.785 = 78.5\%$$

(d) Death and Ebo-Z are not independent, because for example $P(D) = 67.3\%$ and $P(D|Z) = 78.5\%$

7.

(a) By the CLT, the proportions, $\hat{p}$, in the sample are

(i) Normally distributed

(ii) Mean $p = 0.139$

(iii) Standard error

$$SE = \sqrt{\frac{0.139(1 - 0.139)}{1100}} = 0.0104.$$

(b) Thus to find the probability that $\hat{p} < 12\%$, we calculate the z-value

$$z = \frac{0.12 - 0.139}{0.0104} = -1.82.$$

From the table, $P(\hat{p} < 0.12) = P(Z < -1.82) = 0.0344 = 3.44\%$. Thus, there is a 3.44% change of getting a sample with 12% or lower.

(c) To make sure that Hispanic views are fully represented, the polling company can use stratified sampling, where a proportionately larger sample is taken from Hispanics.

8. Using the data given, we have

(a) The expected value, or long run mean, is

$$\text{Mean} = 50(0.50) + 100(0.30) + 250(0.15) + 500(0.04) + 1000(0.01) = \$122.50.$$

Thus, in the long run, the average donation is \$122.50.

(b) The standard deviation is

$$SD = \sqrt{\text{Variance}} = \sqrt{35520} = 188.48$$

Thus the standard deviation is \$188.48.

(c) No, because the distribution is not symmetric. It is skewed right. Standard deviation is too large for normality.

(d) No. A random sample of individual donations will generally be distributed like the original distribution, which is not normal.
Mean is approximately \$122.50.
SD is approximately \$188.48

(e) Yes, by the Central Limit Theorem.
Mean is \$122.50.
SD is $188.48/\sqrt{100} = 18.848.$

(f) A total of \$10,000 from 100 members is equivalent to a mean donation of $10{,}000/100 = 100$ dollars. Since the means are distributed normally, to find the probability, we find the z-value:

$$z = \frac{100 - 122.50}{188.48/\sqrt{100}} = \frac{100 - 122.50}{18.848} = -1.19.$$

The probability that the mean is greater than this value is $1 - 0.1170 = 0.8830 = 88.3\%.$

(g) Because the samples are all size 100, the largest mean corresponds to the largest total. The top 10% of means have $z = 1.28$ (looking up 0.9 in the body of the table). Thus, the mean, $\bar{x}$, is given by

$$1.28 = \frac{\bar{x} - 122.50}{18.848}$$

$$\bar{x} = 122.50 + 1.28(18.848) = 146.63.$$

So certificates will be awarded to volunteers who obtained a mean of more that \$153.50, which corresponds to a total from 100 members of \$14,663.