Math 263, Section 6 **Name____**

Second Test Spring 2010 **If you didn't give one before: (4 figs) Pin________________**

For Questions 1-4, circle *one* answer. (5 points each)

1. A set of test scores are normally distributed. Their mean is 100 and standard deviation is 20. These scores are converted to standard normal Z scores. What are the mean and standard deviation of the new distribution?

   (a) $\mu = 0 ; \sigma = 1$

   (b) $\mu = 1 ; \sigma = 0$

   (c) $\mu = 1 ; \sigma = 5$

   (d) $\mu = 100 ; \sigma = 20$

   (e) $\mu = 0 ; \sigma = 20$

2. If the occurrence of one event does not influence the outcome of another event, then two events are:

   (a) conditional

   (b) disjoint

   (c) independent

   (d) interdependent

   (e) mutually exclusive

3. A bowl contains 100 well-mixed candies. 20 yellow, 50 red, and 30 blue. Without looking, Carmen pulls out 10 candies, and counts the number of reds. Then she puts the candies back into the bowl, and mixes them up. She repeats this four more times. The number of reds in Carmen's five pulls is most likely to be:

   (a) 8,9,7,10,9

   (b) 3,7,5,8,5

   (c) 5,5,5,5,5

   (d) 2,4,3,4,3

   (e) 3,0,9,2,8.

4. A student population is 30% from Phoenix. For samples of 100 students, the sampling distribution of the proportion of students from Phoenix is approximately

   (a) Binomially distributed, with mean 30 and standard deviation 4.6

   (b) Binomially distributed, with mean 0.3 and standard deviation 0.46

   (c) Binomially distributed, with mean 0.3 and standard deviation 0.046

   (d) Normally distributed, with mean 30 and standard deviation 4.6

   (e) Normally distributed, with mean 0.3 and standard deviation 0.46

   (f) Normally distributed, with mean 0.3 and standard deviation 0.046

5. (**10 points**) A random sample is selected from a population with mean μ and standard deviation, σ. The Central Limit Theorem tells us that (mark each statement T (true) or false (F)):

__F__The sample mean $\bar{x}$ gets closer to the population mean μ as the sample size increases.

__F__If the sample size n is sufficiently large, the sample will be approximately normal.

__F__The sample mean $\bar{x}$ will be μ if the sample size n is sufficiently large.

__T__If the sample size is sufficiently large, the distribution of $\bar{x}$ will be approximately normal with mean μ and standard deviation, $\sigma/\sqrt{n}$.

__T__ The distribution of $\bar{x}$ will be normal for samples of size 2 if the population from which the samples are selected is normal.

6. (**18 points**) An architect may have to show a drawing to a client up to five times before the client accepts it. Based on past experience, an architect knows the probability distribution for X, the total number of times a drawing is shown to the client, has the distribution given in the table:

Number of times shown, X	1	2	3	4	5
Probability has to be shown X times	0.1	0.2	0.3	0.2	0.2

(a) Is X discrete__X___ or continuous_____? (check one)

(b) What is the probability that a drawing is shown only once?

0.1

(c) Is the function shown the pdf__X___ or the cdf_____? (check one)

(d) Put the values of the other function (cdf or pdf) in this table:

Number of times shown, X	1	2	3	4	5
	0.1	0.3	0.6	0.8	1

(e) What is the average number of times a drawing must be shown to a client before it is accepted?

$$\text{Mean} = 3.2 \text{ times}$$

(f) If the architect changes his method of work, making drawings more likely to be accepted in one to three showings, does the mean increase______ or decrease___X___? (check one)
Reason:
Since the mean is above 3, increasing the probability of the number of showings being 3 or less will decrease the mean.

(g) Mr. Klein is a client of this architect. Mr. Klein has already inspected his drawing twice and not accepted it.. What is the probability that he will accept the drawing at the third showing?
This is a conditional probability, with the "given that" being 2 showings. Since the probability of more than 2 showings is $0.3 + 0.2 + 0.2 = 0.7$,

$$\text{Conditional probability that exactly 3 showings are needed} = 0.3/0.7 = 0.4286$$

7. (**20 points**) The temperatures of healthy humans are approximately normally distributed with mean 98.6°Fahrenheit and standard deviation 0.8°Fahrenheit.

(a) What is the probability that a randomly selected healthy person has a temperature of over 99°Fahrenheit?

We find the Z-value:

$$z = \frac{99 - 98.6}{0.8} = 0.5.$$

Then from the table $P(\text{Temperature} > 99) = P(Z > 0.5) = 1 - 0.6915 = 0.3085.$

(b) What is the probability that a random sample of 16 healthy people have an average temperature of over 99°Fahrenheit?

By the Central Limit Theorem, the standard deviation of $\bar{x}$, the mean, is $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{0.8}{\sqrt{16}} = 0.2$. The Z-value is now given by:

$$z = \frac{99 - 98.6}{0.2} = 2.$$

Then from the table $P(\bar{x} > 99) = P(Z > 2) = 1 - 0.9772 = 0.0228.$

(c) If the standard deviation of the population were larger, how would your answer to part (b) change?
It would Increase__X__Decrease_____Remain the same_____ (check one)
Explain.
Increasing the standard deviation, σ, would increase the standard deviation, $\sigma_{\bar{x}} = \sigma/\sqrt{n}$, of $\bar{x}$. This would make the *Z*-value smaller and the probability larger.
Alternatively, increasing the standard deviation makes the sample means more variable and a larger probability that $\bar{x}$ is above 99°.

(d) If the sample size were larger, how would your answer to part (b) change?
It would Increase____Decrease___X__Remain the same_____ (check one)
Explain.

Increasing the sample size, n, would decrease, $\sigma_{\bar{x}} = \sigma/\sqrt{n}$, the standard deviation of $\bar{x}$. This would make the *Z*-value larger and the probability smaller.
Alternatively, increasing n makes the sample means more scrunched around 98.6° and a smaller probability that $\bar{x}$ is above 99°.

8. (**19 points**) A 2005 article reported on an anti-seizure drug, Valproate, given to bi-polar alcoholics.[1] Some participants were given the drug, some a placebo, and all were questioned six months later to see if they were drinking heavily (five drinks a day for me, four for women); the results are in the table:

	Heavy drinking	No heavy drinking	Total
Valproate (anti-seizure drug)	14	18	32
Placebo	15	7	22
Total	29	25	54

Using the data in the table, answer the questions. *Leave numerical answers as fractions.*

(a) What proportion of the people who remained heavy drinkers had taken Valproate?

$$\frac{14}{29} = 0.483 = 48.3\%$$

(b) What proportion of the people who did not continue heavy drinking had taken Valproate?

$$\frac{18}{25} = 0.720 = 72.0\%$$

(c) What is the probability that someone who took Valproate continued to be a heavy drinker?

$$\frac{14}{32} = 0.438 = 43.8\%$$

(d) What is the probability that someone who took the placebo continued to be a heavy drinker?

$$\frac{15}{22} = 0.682 = 68.2\%$$

(e) Are your answers to parts (a) and (c) the same? Explain whether you would expect them to be the same or different, and why.

Same_____ Different__X____ (check one)

Explain in words:

They are different, 48.3% and 43.8%. They are likely to be different as we are looking at the 14 people who both took the drug and remained heavy drinkers as a fraction of two different groups: the heavy drinkers in part (a) and those who took Valproate in part (c). There are different numbers in each group (29 in one and 32 in the other), so we would expect the fractions to be different.

(f) From the data given, are having taken Valproate and continuing to be a heavy drinker independent?

Yes____ No __X___ (check one)

Explain in words:

Not independent, because $P(\text{Drinking}) = \frac{29}{54} = 53.7\%$ and $P(\text{Drinking}|\text{Valproate}) = 48.3\%$.

Alternatively, the following are not equal:

$P(\text{Drinking and Valproate}) = \frac{14}{54} = 25.6\%$ and $P(\text{Drinking}) \cdot P(\text{Valproate}) = \frac{29}{54} \cdot \frac{32}{54} = 31.8\%$

In addition, $P(\text{Drinking}|\text{Valproate}) = 43.8\%$ and $P(\text{Drinking}|\text{No Valproate}) = 68.2\%$ are different.

(g) In general, if a drug is effective in alleviating a condition, would you expect taking the drug and having the condition to be independent?

Yes____ No ___X__ (check one)

Explain in words:

Not independent. If the drug is effective, having taken it changes the likelihood of the condition, so $P(\text{Condition}|\text{Drug})$ and $P(\text{Condition})$ should not be the same if the drug has an effect.

[1] "Study: Anti-Seizure Drug Reduces Drinking in Bi-Polar Alcoholics" in *Bipolar Central* (online), Jan 6, 2005. Reported by N. Pfenning, *Elementary Statistics*, Brookes/Cole, 2011.

9. (**10 points**) Mr. Aaron is about to interview four candidates for a particular job. He hopes all the candidates turn off their cell phones, because he does not like to be interrupted by a ringing cell phone. Suppose all four candidates own cell phones and none turn his/her cell phone off. Each candidate has a 35% chance of receiving a call during the interview.

(a) If there is only one candidate whose cell phone rings during the interview, Mr. Aaron will immediately not hire this candidate. What is the probability of this? Show work.
Since the probability of the phone ringing is 0.3 (a success) and the probability of it not ringing is 0.7, by the Binomial distribution

$$\text{Probability} = 4(0.3)^1(0.7)^3 = 0.4116$$

(b) If there is only one candidate whose cell phone does not ring during the interview, he will immediately hire this candidate. What is the probability of this? *Show work.*

Similarly,

$$\text{Probability} = 4(0.3)^3(0.7)^1 = 0.0756$$

10. (**12 points**) The circumferences of women's upper thighs are normally distributed with mean 22.30 inches and standard deviation 1.88 inches; for men they are normally distributed with mean 22.00 inches and standard deviation 1.88 inches.[2] A woman and a man are randomly selected.

(a) For the randomly selected man and woman, what is

(i) The mean of the difference in their upper thigh sizes?

Let M be the man's thigh size and F be the woman's. Taking the man's length from the woman's (but could have been the other way).

The mean of difference is given by

$$E(F - M) = E(F) - E(M) = 22.30 - 22.00 = 0.3 \text{ inches}$$

(ii) The standard deviation of the difference in their upper thigh sizes?

Since the man and the woman are randomly selected,

$$V(F - M) = V(F) + V(M) = (1.88)^2 + (1.88)^2 = 7.0688.$$

Thus

$$SD(F - M) = \sqrt{7.0688} = 2.66 \text{ inches.}$$

(b) Assuming that the difference in upper thigh sizes is normally distributed (which it is), what is the probability that the man's upper thigh is larger than the woman's?

We are looking for the probability that $F - M$ is negative. Since the difference in lengths is normally distributed with mean 0.30 and standard deviation 2.66 inches, the z-value of 0 is

$$z = \frac{0 - 0.30}{2.66} = -0.11.$$

From the table, the probability that $z \leq -0.11$ is $0.4562 = 45.62\%$.

[2] Data from *Elem Statistics* by N. Pfenning, Brooks/Cole 2011

11. (**16 points**) On a previous Excel assignment, you simulated the results on the New Haven firefighter's lieutenants' exam. We now look at the captains' exam. For the captains' exam, there were 41 candidates: 25 White, 8 Black and 8 Hispanic. Consider the following two distributions:

I. The distribution of the proportion of minorities (Blacks and Hispanics) in a simulation of 500 random samples of 5 firefighters drawn from this group.

II. The distribution of the proportions of minorities in all possible samples of 5 firefighters drawn from this group.

In addition, you are told that the mean proportion of *one* of these distributions is 0.382 and its standard deviation is 0.233.

For this question, assume the Central Limit Theorem (CLT) applies even though the sample size is less than 30.

(a) Which of these two distributions is the sampling distribution? How do you know?

The second one: The sampling distribution is for *all* samples.

(b) What is the mean and standard deviation of the sampling distribution?

The population proportion is

$$p = \frac{8+8}{41} = 0.39.$$

The CLT tells us that the mean of the sampling distribution is 0.39. The CLT tells us the standard deviation of the sampling distribution is

$$\text{Standard Error} = \sqrt{\frac{0.39(1-0.39)}{5}} = 0.218.$$

(c) What is the mean and standard deviation of the other distribution?

The first distribution (the simulation) is an approximation to the sampling distribution. It should have a mean that is close to 0.39, but it may not be equal to 0.39; in this case it is 0.382. Similarly, we expect its standard deviation to be close to, but not equal to 0.218; in this case it is 0.233.

(d) In this simulation, the proportions of the 500 samples with (i) all minority candidates and (ii) no minority candidates is 0.01 and 0.11, not necessarily in that order. Which number is which? Justify your choice.

Since there are fewer minorities than whites (16 versus 25), a sample with all minorities is less likely than a sample with no minorities.
Thus all minorities has probability 0.01 and no minorities has probability 0.11