Hi. I hope you are feeling up to the task. Relax. This is just a test, not an actual emergency. Show your work (i.e. function symbols, calculations) in order to communicate with me what you are thinking. Some of you may not want to show me your work for the same reason. But I digress. If you are unsure about what I am asking, come see me for clarification. Good luck. As usual the number in the [ ] brackets is the number of points the problem is worth. Round your probabilities to four decimal places.

1. [4] A game warden takes a SRS of 8 deer that were checked in during hunting season. He finds the average weight, of the eight deer, to be 155.0lbs. Assuming the standard deviation of deer weights is 20lbs, find the a 95% confidence interval for the mean weight of deer. Round each number to the nearest tenth of a pound.

\[
155.0 \pm 1.96 \frac{20}{\sqrt{8}} = 155.0 \pm 13.9 \\
= (141.1, 168.9)
\]

2. [2] In order to calculate the above 95% confidence interval what did you have to assume was true about the distribution of \( \bar{X} \)?

We must assume that the distribution of \( \bar{X} \) is normal or nearly normal.

3. [4] After being absent from his History class for three weeks, Mr. M walks in to his class to discover that a multiple choice exam is being given. The exam is multiple-choice, with each question having four choices. There are 40 questions in all. If Mr. M guesses at each question on that exam what is the probability of getting exactly 10 questions correctly?

Let the random variable \( Y \) count the number of times Mr. M guesses a question correctly.

\[
\begin{align*}
I & \text{ want to calculate } P(Y = 10) = \frac{40!}{30!10!} (0.25)^{10} (0.75)^{30} \\
& = 0.1436 \\
& = \text{binomdist}(10,40,0.25,\text{false})
\end{align*}
\]

4. [4] When skyscrapers are constructed in New York city there is a 0.5% chance that a worker will be killed. We will assume that the probability of any one worker dying during construction is independent of any other worker surviving or perishing. If it takes 1000 workers to construct a skyscraper in New York city what is the expected average number of deaths?

\[\mu = 0.005(1000) = 5 \text{ deaths.}\]
5. I want to estimate my ability to leap up into the air by measuring how high I can jump. I am interested in my average height of a jump. I will measure each jump height in inches. I have noticed that the standard deviation associated with my distribution of jump heights is 2 inches, \( \sigma = 2 \) inches. If I want to estimate my mean jump height, so the margin of error is 0.5 inches, with 95% confidence, how many jumps should I measure?

\[
n = \left( \frac{1.96(2)}{0.5} \right)^2
\]

\[n = 62\] jumps

6. You are testing a new medication for relief of depression. You are going to give the new medication to subjects suffering from depression, and after a month see if their symptoms have lessened. You have twelve subjects available. Half the subjects are to be given the new medication and the other half a placebo. The names of the eight subjects are given below.

(01) Tanner      (02) Yee     (03) Daniels (04) Chavez (05) Duvall (06) Harrison
   (07) Arroyo      (08) Costello (09) Dimitri (10) Regan (11) Smith (12) Gunther

Using the list of random digits, the first six people chosen will receive the medication. Who are the first two people chosen?

81 50 72 71 02 56

First person chosen is Yee, second person chosen is Harrison.

7. (Circle the correct response) Suppose that the distribution of a population is not normal. Then the sampling distribution of the mean, for a large enough sample size,

a. is binomially distributed.
   b. is the same as the original population
   c. is exactly normally distributed.
   d. is approximately normally distributed

8. What is the purpose of creating a confidence interval, as presented in section 6.1?

To estimate the population parameter \( \mu \).
9. [4] A fair coin has a 50% chance of landing heads. As mentioned in your text (section 4.1) the French naturalist Count Buffon (1707 – 1788) tossed a coin 4040 times of which the coin landed heads 2048 times. If the coin is fair, how likely is it to observe a head outcome of 2048 times or more out of 4040 tosses? Use a normal approximation with out continuity correction.

\[ P(X \geq 2048) = P \left( Z > \frac{2048 - 2020}{\sqrt{(0.5)(0.5)(4040)}} \right) \]

\[ = P(Z > 0.8810) \]

\[ = 0.1894 \text{ using the tables for } Z = 0.88. \]

10. [4] I create a 99% confidence interval with a set of data. My supervisor walks in and says to change it to a 90% confidence interval using the exact same set of data. Will my new interval range for the 90% confidence interval be larger or smaller compared to the 99% confidence interval?

The 90% confidence interval range will be smaller compared to the 99% confidence interval.

11. [4] A 95% confidence interval is supposed to contain the population mean, \( \mu \), 95% of the time. If we gather 12 separate samples from the same population, and calculate a 95% confidence interval for each, what is the probability that only 7 or less of the 12 confidence intervals contain the population mean, \( \mu \)?

Let the binomial random variable \( Y \), count the number of times a confidence interval contains the population mean, \( \mu \).

\[ P(Y \leq 7). \text{ I used the Excel command } \text{binomdist}(7,12,0.95,\text{true}) \]

\[ P(Y \leq 7) = 0.000184 \]
12. [4] Two friends decide to go fishing at Lake Shasta, since the have heard from other fisherman about the large fish caught there. Suppose that the average size of a fish at that lake is 26 inches with a standard deviation of 8 inches. The distribution is approximately normal. Let the random variable $X$ measure the length of the fish in inches. What is the probability that the first fish caught exceeds 28 inches?

\[
P(X > 28) = P\left(Z > \frac{28 - 26}{8}\right)
= P(Z > 0.25)
= 0.4013
\]

13. [4] The time it takes a 12 year old child to run half a mile is normally distributed with a mean time of 7.45 minutes, and a standard deviation of 1.44 minutes. Suppose 10 children were gathered at random from the general population. What is the probability that the average time of the ten children is less than 6 minutes? Let the random variable $Y$ equal the number of minutes it takes to run half a mile.

\[
P(\bar{Y} < 6) = P\left(Z < \frac{6 - 7.45}{\frac{1.44}{\sqrt{10}}}\right)
= P(Z < -3.18)
= 0.0007
\]
14. A simple, two-person game of chance, in which both persons have an equal probability of winning goes like this: A standard deck of 52 cards is used. One person is assigned aces, and the other person is assigned the rest of the 48 cards. A card is drawn at random from the deck of 52, if an ace is selected then the other person must pay the person assigned aces $120, otherwise, if any of the other cards appear, the person assigned aces must pay the other person $10. The card is placed back and the deck is reshuffled for another round.

Let the random variable $A$ equal the amount of money won or lost by the person assigned the aces.

<table>
<thead>
<tr>
<th>$A$</th>
<th>$120$</th>
<th>$-10$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(A)$</td>
<td>$\frac{4}{52}$</td>
<td>$\frac{48}{52}$</td>
</tr>
</tbody>
</table>

a. [4] In the long run, what is the average amount of money the person assigned the aces can expect to win/lose?

$$\mu = 120 \left(\frac{4}{52}\right) + -10 \left(\frac{48}{52}\right)$$

$$\mu = 0$$

b. [4] Two people play the game 169 times. Approximate the probability that the person assigned aces, on average, wins $2.00 or more?

$$\sigma = \sqrt{\frac{4}{52} (120 - 0)^2 + \frac{48}{52} (-10 - 0)^2}$$

$$\sigma = 34.641$$

$$P(\bar{A} > 2) \approx P \left( Z > \frac{2.00 - 0}{34.641} \right)$$

$$\approx P(Z > 0.75)$$

$$\approx 0.2266$$

15. [4] A six question multiple choice quiz has five selections per question. What is the probability that if a person merely guesses at each question, gets all six answers wrong? Let the random variable $G$ equal the number of incorrect answers on the quiz.

$$P(G = 6) = \left(\frac{4}{5}\right)^6 = 0.2621$$
The Central Limit Theorem is a crucial concept in understanding how we estimate parameters. We don't offer a proof of it in our book, since the method used to prove such a theorem is not appropriate at this level. The question in problem 14 is from section 5.2; the average amount won by playing 169 games. The question asked for the probability of winning $2 or more on average, for the player who is assigned aces. If he wins exactly $2 per game on average, this would mean that the person won a total of $2(169) = $338. As it turns out it is impossible to win exactly $2 per game on average - the reason is that no total combination (169 numbers) of -$10, and $120 will yield an average of $2.

The Central Limit theorem, says that a probability calculation involving averages of numbers due to sampling can be approximated by assuming that the sampling distribution of averages is normally distributed - for a large enough sample. This means we need to calculate $\mu$, and $\sigma$, whose calculation depends on the population's original distribution. For problem 14 we needed to use the calculation for discrete distributions.

Here is another big concept. The distribution presented in problem 14 is not binomial! While there are only two outcomes, the random variable $A$ defined in problem 14, gives the amount of money lost or won by the player assigned the aces; a binomial random variable counts the number of successes.

How good is our approximation as found in problem 14? I calculated that the probability that a player assigned aces wins on average $2 or more, on 169 games, is 22.66% by using a normal approximation. I calculated the exact value, by calculating the actual sampling distribution of the mean (averages) for the situation described. The actual value is 22.91%. I chose a sample size of 169 because, due to experience, I felt the sample size was large enough given the situation.

How did I calculate 22.91%?

As I was grading the exam, I got curious about the actual probability value. I felt that since there were only two outcomes, I could exploit the binomial random variable (this means I need to redefine my random variable).

Step 1 - I wondered, if I calculated every possible average using a combination of -$10, and $120, 169 times how many actual averages are possible? After some thought, I understood that only 169 distinct values are possible; at one extreme is 169 “-$10” numbers which would give me an average of -$10. At the other extreme, all the numbers are $120, to yields an average of $120. That leaves only 167 other possible values for $A$, between -$10 and $120.

Step 2- But what are these values? The same impetus that gave me the answer to the first question, would then help me answer the second question. I imagined starting at one of the extremes, like 169 values of -$10. Then replace one of the numbers with 120 and calculating the new average. Then replace two of the -$10, with $120 and then recalculating the average, and so on until I reached the other extreme.

As a matter of fact the formula for the average would be

$$average = \frac{-10\text{(number of -10)} + 120\text{(number of 120)}}{169}$$

Since the total number of –10s and 120s is 169 I could reduce the above formula to only one variable, by using substitution - what I have is a system of linear equations in two variables.

Step3 - Great. Now I have a quick way of finding any of the possible 169 averages, by the use of the formula as described above, but how do I calculate the probability of any of the possible averages? Here is where I redefine my random variable. What I realize is that there is only one way I can get an average of -$10, out of 169 tries, and that is for every single game resulting in the player assigned aces loosing every hand. Another situation is that the player assigned
aces loses 168 games and wins one game. This means only one $120 payoff in the 169 games, and there are 169 ways that can happen - imagine moving the position of the $120 as described in step 2. If I kept this going I could see that this would be tedious. But I recognize now that what I have here is the same situation as tossing a coin 169 times. I just need to redefine my random variable. So let the random variable X be the number of times the person assigned aces gets $120. So, the player either gets $120 (success) or not. The probability from play to play is always the same and it is independent. How does this tie in with the averages? If X = 0, then this corresponds to 169, -$10. If X = 1 this corresponds to one $120, and 168, -$10. And so on.

What is not clear is that there is a third random variable, I will call it \( \overline{A} \), which is defined as the actual averages out of a situation involving 169 plays, sampling from the population with random variable A. It turns out, the probability assignment for \( \overline{A} \), (notice it is discrete- only 169 possible values) is the same as for the binomial random variable X.

So I can use \( P(X = k) = \text{Binomdist}(k, 169, 4/52, \text{false}) \) to find all 169 probabilities for the random variable \( \overline{A} \).

\[
P(\overline{A} = -$10) = P(X = 0) \\
P(\overline{A} = $120) = P(X = 169)
\]

Below you can see a scatter plot of the first 27 outcomes, along with a table of the first 30 outcomes, along with the associated probabilities. As you can see the distribution has that normal shape for the first 27 outcomes. Warning the distribution of \( \overline{A} \) is not normal; it is discrete. All the central limit theorem really says is that we can approximate probability questions involving the distribution of \( \overline{A} \) using a normal distribution.

Notice as was mentioned earlier, an average of $2 is not possible. The only averages close to $2 are $1.54 and $2.31. If I sum the probabilities for the averages $2.31 and greater, I get a probability of approximately 22.91%
16. [4] The probability tossing a pair of dice and having the sum equal 7 on the first throw is 1/6. Suppose I want to calculate the probability that out of 30 tosses 8 or less of the tosses results on a sum of 7. Is it appropriate to use a normal approximation to calculate this result? A yes or no response will receive zero credit. You must provide proof/explanation for your decision.

Since the condition \( np \geq 10 \) is not met, \( 30(1/6) \not\geq 10 \), we can not use a normal approximation.

17. [4] State whether each bold face number is a parameter or a statistic. A manufacturer calibrates a liquid dispensing machine so that it dispenses on average 55 mL of liquid. A sample of 10 cups finds that the average liquid dispensed, for the ten cups, is 54 mL.

55 mL __Parameter__________________________ 54 mL __Statistics__________________________

18. [2 each] Match the word with its definition

| a. Bias | 4 |
| b. Parameter | 6 |
| c. Statistic | 8 |
| d. Undercoverage | 10 |
| e. Nonresponse | 11 |
| f. A Block | 1 |
| g. Blind Experiment | 5 |
| h. Statistically Significant | 9 |
| i. Placebo | 2 |
| j. Observational Study | 7 |
| k. Experiment | 3 |

1. A group of experimental units or subjects that are known before the experiment to be similar in some way that is expected to affect the response to the treatments
2. A fake drug (sugar pill) given to a subject to test the affect of the real drug.
3. Deliberately imposes some treatment on individuals in order to observe their responses.
4. Systematically favors a certain outcome
5. The experimenter does not know who is receiving the actual drug being tested.
6. A number that describes a population.
7. A situation in which the researcher merely watches individuals and measures variables of interest but does not attempt to influence the responses.
8. A number used to estimate a quantity that describes a population’s distribution.
9. A result that could not occur by chance alone.
10. Occurs when some groups in the population are left out of the process of choosing a sample.
11. Occurs when an individual chosen for the sample can not be contacted or refuses to cooperate.
19. [4] The probability that a person is born male is approximately 50%. Suppose a couple decides to have children until a son is born. The random variable $X$ counts the number of children the couple has. Is this scenario a reasonable probability model for the **binomial** random variable $X$? Explain.

I did not grade this question since the key word “binomial” was omitted.

The situation is not binomial since the random variable $X$, has no upper value, that is, we do not have a fixed number of observations.

20. [4] On KGW News Channel 8, the announcer gives a phone number so people can voice their opinion on the War in Iraq. The next day the announcer gives the result of the 1068 people that phoned in, as 48% of the public believe more troops are needed in Iraq, with an a sampling error of $\pm 3\%$. Is this result trustworthy? Explain your answer.

The use a confidence interval, as presented in section 6.1, requires that the data be selected using a simple random sample. If you have people calling in, this constitutes a voluntary response sample which violates the main requirement for creating a confidence interval. Thus, this result is not trustworthy.