1. A pumpkin farmer produces pumpkins that average 12lbs in weight, with a standard deviation of two pounds. The pumpkins are packed 16 to a crate. What is the chance that a single crate weighs more than 195 pounds?

(a) Less than 1%
(b) About 7%
(c) About 13%
(d) About 34%
(e) About 69%

$16 \times 12 = 192$ lbs

$$z = \frac{195-192}{\sqrt{16} \cdot 2} = \frac{3}{8} \approx +.40$$

The next 5 problems are related to each other and use the information below:

There are 20,000 restaurants in the County of Los Angeles, 50% of them received a letter grade of "A" during inspections, 40% received either a B or a C grade and 10% failed their inspections. Restaurant grades are not normally distributed. My financial adviser, the Oracle, has hired you as a temporary personal assistant. Your job is to schedule his next 16 dinners (Oracle never eats at home). Unfortunately, you didn't know about the rating system and you never eat out because you don't have the money. So you listened to your best friend and picked 16 restaurants at random from an internet database of the 20,000 restaurants in Los Angeles. The Oracle will give you +3 points if you choose "A" restaurants, +1.25 points if you choose "B" or "C" restaurants, and -20 points if you choose a restaurant with a failing grade. Treat your restaurant selections as if they were a simple random sample of restaurants.

2. Construct a box model for this problem

$\boxed{+3}$	$\boxed{+1.25}$	$\boxed{-20}$
.50	.40	.10

Box Average = 0

3. What is the expected value for the total score of the 16 restaurants selected at random?

$$EV_{sum} = 16 * \left((+3 \times .5) + (+1.25 \times .40) + (-20 \times .10)\right)$$
$$= 0$$

4. What is the standard deviation for the box?

$$SD_{BOX} = \sqrt{(.50)(+3-0)^2 + (.40)(+1.25-0)^2 + (.10)(-20-0)^2}$$
$$= \sqrt{45.125}$$
$$= 6.7175$$

5. What is the standard error for the total score of a sample of 16 restaurants?

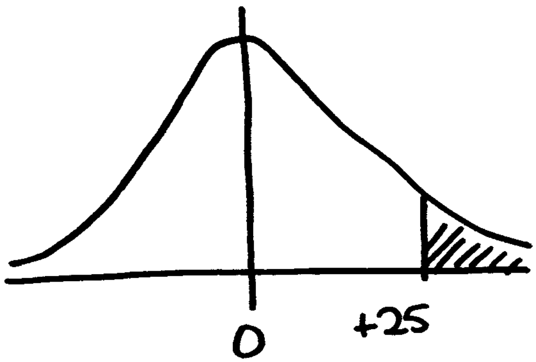
$$SE_{sum} = \sqrt{16} \cdot 6.7175$$

$$= 26.8701$$

6. To convert your temporary job into a permanent job, you must have accumulated a total of at least +25 points from the Oracle after picking 16 restaurants for him. What's your chance of getting a total of at least +25 points after picking 16 restaurants? If it is not possible to calculate the chance, please write "not possible" below and explain why.

The best answer here is to argue that the normal approximation does not apply because 16 draws is NOT REASONABLY LARGE (see Ch. 18 parts 5+6) (and Ch. 18 summary point 3) – therefore it is not possible.

But you'd probably get some credit for doing the work so here is a solution:



$$z = \frac{25-0}{\sqrt{16} \cdot 6.7175} = .93 \approx {}^{+}.95$$

area between $-.95$ and $^{+}.95$ is 65.79%

so chance is $\frac{100-65.79}{2} = 17.105 \approx \boxed{17.1\%}$

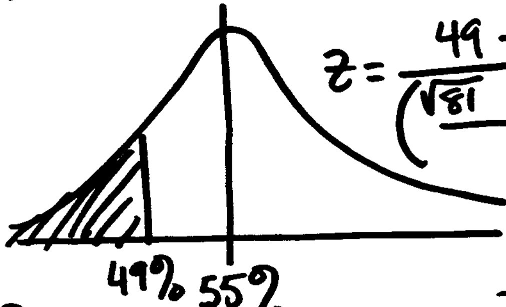
7. Suppose we are psychics and we know that former police chief Bernard Parks will be the next Mayor of Los Angeles with a final winning percentage of 55%. Unfortunately, we don't know Parks and he doesn't return our phone calls or e-mails so he doesn't know he will get 55% of the vote after the next election. In fact, he is spending a lot of money right now on random surveys of votes of size 81 to help him make decisions about the upcoming election.

a. What is the chance that one of his surveys will give a result showing that he will get 49% or less of the vote if it is true that he really has 55%?

Box is KNOWN

for Parks → [1] [0] ← against

.55 .45



$$z = \frac{49-55}{\left(\frac{\sqrt{81}\cdot\sqrt{.55\times.45}}{81}\times 100\right)} = -1.09$$

$z \approx -1.10$

area between ± 1.10 is 72.87% so tail is $\boxed{13.565\%}$

Some bad question → (b) Suppose again that Parks does not know that he will get 55% of the vote and suppose he takes another random survey of size 81 and it shows that 49% will vote for him. Can ~~you~~ he construct a 95% confidence interval for the population percentage of votes for Parks?

(should give you a level of confidence)

(should have read he since you know the parameter)

Circle one: Yes No

If you circled yes, please construct a 95% confidence interval. If you circled no, please explain why ~~you~~ he cannot construct a confidence interval.

Sure, use his sample information

$$49\% \pm 2*\left(\frac{\sqrt{81}\cdot\sqrt{.49\times.51}}{81}\times 100\right)$$

$$\pm 2*(5.5544)$$

$$\boxed{49\% \pm 11.11\%}$$

8. There were a total of 226,324 deaths in California in 1999. A random sample of 242 deaths was selected. Detailed research determined that the deceased was cremated in 99 of the deaths.

a. Determine a 99% confidence interval for the proportion (or percentage) of deaths in California in which the deceased is cremated.

$\frac{99}{242} = .4091$

$\approx 41\%$

$41\% \pm (2.6)\left(\frac{\sqrt{242} \cdot \sqrt{.41 \times .59}}{242} \times 100\right) = 41\% \pm 8.22\%$

use $z = 2.6$ for 99%

(using 3 is OK)

b. Suppose the confidence interval is too narrow, identify 2 things you can do to make the interval wider.

1) increase confidence

2) decrease sample size

c. A classmate comes up to you and says, this is the interpretation of a 99% confidence interval:

"There is a 99% probability that the true parameter is in the interval you gave in part (a)"

Is your classmate's interpretation correct? (circle one) YES NO
And justify your choice in the space below.

It is wrong to talk about the true parameter having a probability.
A parameter is a fixed value.

Instead it is the statistic and the endpoints of the confidence interval which have probabilities (or chances)
99% refers to the percentage of intervals over the long run that contain the parameter.
You have a 99% chance of picking a good interval.

9. You know that every UCLA student will definitely get a job after graduation. The only uncertainty is the salary. Suppose this is what you know about the job prospects of UCLA students after graduation:

There is a 35% chance that the salary will be \$20,000 per year; a 45% chance that it will be \$90,000 per year; and a 20% chance that it will be \$40,000 per year. Suppose you draw a random sample of 9 UCLA students.

a. Draw a reasonable box model for this problem

$\boxed{20}\ \boxed{90}\ \boxed{40}$ ← in thousands

.35 .45 .20

Box Average is 55,500

b. Find the expected value of the total (sum) salary for the 9 UCLA students.

$EV_{SUM} = 9 \times 55{,}500 = \boxed{499{,}500}$

in thousands

c. What is the Standard Deviation of the "box" you drew?

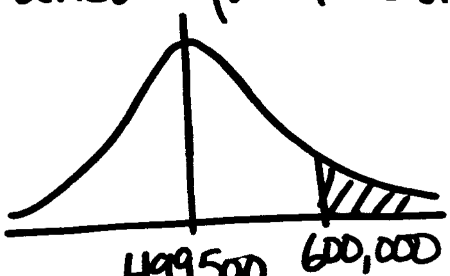
$SD_{BOX} = \sqrt{.35(20-55.5)^2 + .45(90-55.5)^2 + .20(40-55.5)^2}$

$\boxed{= \$32{,}012}$

d. What is the Standard Error of the total (sum) salary for 9 students?

$SE_{SUM} = \sqrt{9} \cdot 32{,}012 \approx \boxed{\$96{,}036}$

e. Suppose you work for me and I tell you to go draw a different random sample of 9 UCLA students and you get a total (sum) salary of \$600,000. What is the chance that you could have gotten a total salary this large or larger? Yes, 9 is too small, but go ahead for this one, I didn't ask.



$\frac{600{,}000 - 499{,}500}{96{,}036} = z = +1.05$

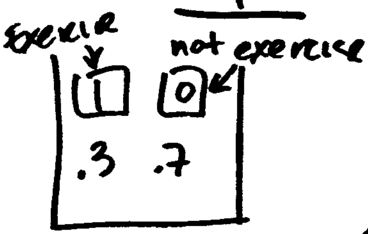
$\frac{100 - 70.63}{2} = \boxed{14.69\%}$

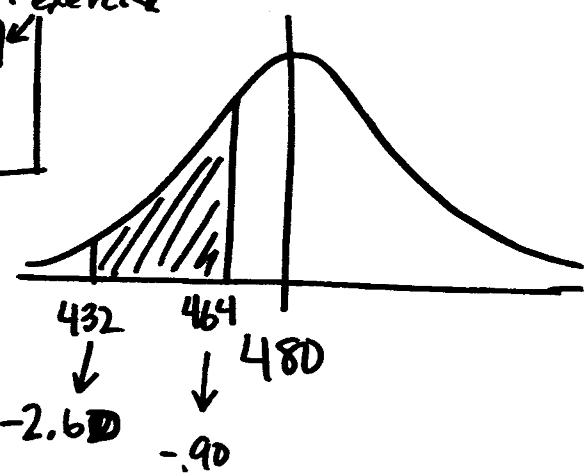
10.Exercise is particularly beneficial for young adults. A study wants examines the exercise habits of college students. A sample of 36 college students is selected by a random process and it is found that a 68% confidence interval for the percentage who exercise weekly is 22.4% to 37.6%. Using this information, please answer the following true/false questions (4 points each)

	True	False	Statement
A		X	There is a 68% chance that for all college students, percentage who exercise weekly is between 22.4% and 37.6%
B	X		If we quadrupled the sample size, we would expect the 68% confidence interval to be one half as wide.
C	X		If we increased the confidence level to 99%, we would expect the confidence interval get wider.
D		X	This 68% confidence interval is valid only when the population variable (percentage who exercise) is normally distributed.
E	X		The sample statistic that generated this confidence interval is 30%, in other words 30% of the sample was found to exercise weekly.

11. Suppose it is know than 30% of all college students exercise weekly. A physiology class needs enough students for a medically valid study and a sample of size 36 is not large enough. So if 1600 college students are chosen at random instead of 36, what is the chance that between 432 and 464 of those chosen do indeed exercise weekly? (8 points)

If 1600 are drawn and 30% exercise, we expect 480 to be exercisers so.





Need 2 z scores

$$z_{464} = \frac{464-480}{\sqrt{1600}\cdot\sqrt{.3\times.7}}$$

$$z_{464} = -.87 \approx -.90$$

$$z_{432} = \frac{432-480}{\sqrt{1600}\cdot\sqrt{.3\times.7}}$$

$$z_{432} = -2.62 \approx -2.60$$

Area between

$z = \pm.90$ is 63.19%

$z = \pm 2.60$ is 99.07%

so $\frac{99.07-63.19}{2} = \boxed{17.94\%}$

12.A study was conducted on the sleep patterns of infants in the United States. A sample of 25 infants was drawn at random with 54% sleeping at least 12 hours a night.

a. Please compute a 95% confidence interval for the percentage of all U.S. infants who sleep at least 12 hours per night.

$$54\% \pm 2\left(\frac{\sqrt{25}\cdot\sqrt{.54\times.46}}{25}\times 100\right)$$

$$54\% \pm 19.9\%$$

b. Suppose the sample size was increased to 100 infants, what effect would this have on the confidence interval? Assume that 54% of them slept at least 12 hours per night.

it would decrease the width (divide in $\frac{1}{2}$)

or $54\% \pm 9.97\%$

c. Please calculate a 90% confidence interval for the percentage of all U.S. Infants who sleep at least 12 hours per night. Again, please use the sample of size 25 and assume the percentage in the sample was 54% sleep at least 12 hours per night.

for 90% use $z = 1.65$

$$54\% \pm 1.65\times\left(\frac{\sqrt{25}\cdot\sqrt{.54\times.46}}{25}\times 100\right)$$

$$\boxed{54\% \pm 16.4\%}$$

13. It's a family tradition: your professor goes to Las Vegas every year for Thanksgiving. A new casino has opened and they are playing a modified roulette game that has 40 possible numbers that can be spun on a wheel. To play you bet \$4 and you get to choose 4 numbers. If the wheel lands on any number that you chose, you win \$10. If the wheel does not land on a number you choose but on one of 8 "special numbers" you don't win or lose anything. If the wheel lands on any of the remaining numbers (not the ones you choose or the special numbers), you lose your bet of \$4. Suppose the typical person plays 25 times.

a. This game of modified roulette can be represented by a box model, please construct a reasonable one in the space below.

$\boxed{+10}$	$\boxed{0}$	$\boxed{-4}$
.1	.2	.7

.1 because $4/40 = .1$

.2 because $8/40 = .2$

and $1.0 - .2 - .1 = .7$

or $28/40 = .70$

b. The 25 plays can be treated like a random sample of size 25. Find the expected value of this game.

$EV = 25 \times -1.8 = -45$

(box average is $(10 \times .1) + (0 \times .2) + (-4 \times .7)$)

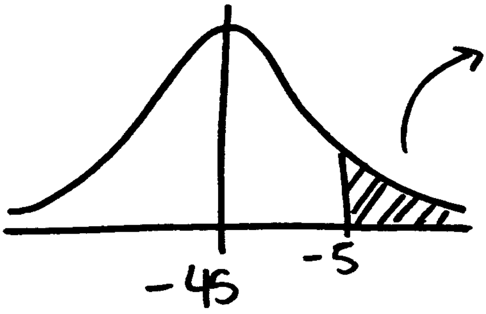
c. Find the standard error of this same game.

$SD_{BOX} = \sqrt{(.1)(10 - -1.8)^2 + (.2)(0 - -1.8)^2 + (.7)(-4 - -1.8)^2}$

$SE = \sqrt{25} * 4.2379$

$SE = 21.1896$

d. Suppose the professor decides to spend \$100 playing a total of 25 times and she lost \$5. Calculate the chance that your professor could lose 5 or fewer dollars playing this game. Show all of your work and answer this question – based on your calculations is she lucky? (let's suppose lucky means the chance of losing 5 or fewer dollars is less than 5%)?



find z

$z = \dfrac{-5 - (-45)}{21.1896} \approx +1.90$

area for ± 1.90 is 94.26

so $\dfrac{100 - 94.26}{2} = \boxed{2.87\%}$

LUCKY

~~14. A survey of Los Angeles is conducted and 3,000 households are drawn at random. Suppose 30% of all households in Los Angeles have only one person living there. What is the chance that the number of households with only one person will be in the range 740 to 850?~~

BAD

15. A multiple-choice quiz has 25 questions. Each question has 5 possible answers, one of which is correct. Four points are given for each correct answer, but a point is taken off for a wrong answer. A passing score is 30.

$\boxed{+4}$	$\boxed{-1}$
.2	.8

a. If a student answers all the questions at random, what is the chance of passing?

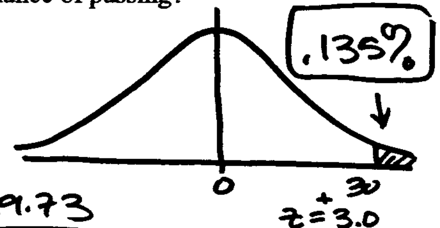
Box Average = 0 so E.V. = 0

Box SD = $(4 - -1)\sqrt{.2 \times .8} = 2$

so SE = $\sqrt{25} \times 2 = 10$

$z = \frac{30-0}{10} = +3$ area between $\pm z$ is 99.73

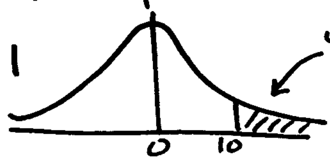
so 99.73/2 = $\boxed{.135\%}$



b. Could a student score 10 points or more just by guessing? Answer yes or no and explain.

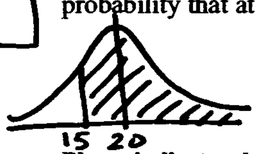
Yes, if we use the info above:

$z = \frac{10-0}{10} = +1$



error	no errors
$\boxed{1}$	$\boxed{0}$
.2	.8

16. From their extensive records, the IRS knows that 20% of federal income tax returns have an arithmetic error serious enough to warrant an audit. If 100 tax returns are randomly selected, what is the probability that at least 15 of them have arithmetic errors? 20% of 100 = expect 20 out of 100



$z = \frac{15-20}{\sqrt{100} \times \sqrt{.2 \times .8}} = -1.25$ area is 78.87 in the middle and 10.565 in one tail so

78.87 + 10.565 = $\boxed{89.435\%}$

Please indicate whether each statement is true or false (3 points each)

	True	False	Statement
17A		X	The Central Limit Theorem suggests that all populations are normally distributed
17B		X	The Central Limit Theorem only applies when the number of draws (sample size) is reasonably large and the population is normal.
17C	X		The variability of a probability histogram is estimated by the standard error.
17D		X	For a population that is not normally distributed, the distribution of sample percentages will have the same shape as the population even when the sample is reasonably large.
17E	X		The Central Limit Theorem implies that as the number of draws (sample size) increases, the probability histogram for a sum becomes more and more normal in its appearance

18. In 2001, a survey organization takes a simple random sample of 1,600 adults in Los Angeles, California, a large American city. Among this sample of adults, it was found that 975 support the death penalty, 525 support life imprisonment with no parole and the rest did not believe in penalties for homicide. It was noted that support for the death penalty had changed from a survey taken in 1991 when approximately 80% of adults in Los Angeles supported the death penalty.

a. Is it possible to construct a 95% confidence interval for the population percentage of Los Angeles adults who support the death penalty in 2001. (circle one) (7 points)

YES NO

If you circled YES, please construct a 95% confidence interval in the space below. If you circled NO, please use the space to explain why it is not possible to construct a 95% confidence interval.

$\frac{975}{1600} = .6094$ $SD = \sqrt{.6094 \times .3906} = .487$ $\rightarrow 2 \times 1.22\% = 2.44\%$

$1 - .6094 = .3906$ $SE = \left(\frac{\sqrt{1600} \times .487}{1600} \times 100\right) = 1.22\%$

so $\boxed{60.94\% \pm 2.44\%}$

b. If the sample size were 400 instead of 1600 it would (circle one to fill in the blank) the width of any confidence interval constructed from the sample information (4 points)

Increase Decrease Not Affect

c. If the level of confidence were 99% instead of 95% it would (circle one to fill in the blank) the width of any confidence interval from the sample information (4 points)

Increase Decrease Not Affect

d. Suppose it was known that actually 60% of all adults in Los Angeles support the death penalty. So if a simple random sample of 625 adults were to be taken, the SE for the sample percentage of death penalty supporters is calculated to be about 2%. You should assume these numbers are correct.

A student, looking at the numbers in part d, interprets them as follows: this means that there is about a 95% chance for the percentage of death penalty supporters in the sample to be in the range 60% ± 4%. (circle one)

The student is correct The student is not correct

Please explain your choice below: (5 points)

If the S.E. is 2% then theory tells us that the sampling distribution (or probability histogram) for samples of size 625 will have 95% of all possible sample %s falling within ± 2 S.E. or 4%

19. Does it pay to sue for damages for work related injuries? Suppose this is what is known about people who have taken their employers to court: 10% have won $100,000, 30% have won $10,000 and the rest have won nothing. And suppose it costs $10,000 in legal fees to take a case to court regardless of whether a person wins or loses.

a. The net award (money won minus legal fees) for work related damages can be represented by a box model. Please construct a reasonable model in the space below.

100	10	0
.10	.30	.60

OR

90	0	-10
.10	.30	.60

b. Suppose a large employer gets sued for work related injuries 121 times per year. The 121 lawsuits can be treated like a random sample of size 121. Find the expected value of the total net award.

BOX AVERAGE = $3,000 (if used $13,000 need to "net out" the answer)

$121 * 3{,}000 = \boxed{\$363{,}000} = EV_{NET}$

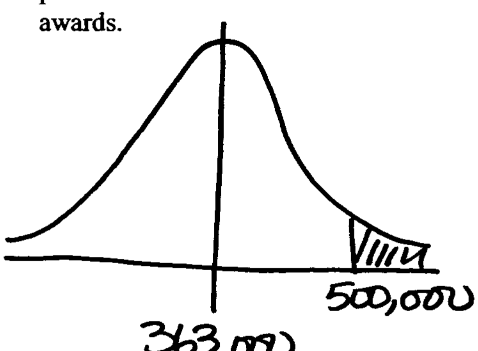
c. Find the standard error of the net award for the 121 lawsuits.

$$\sqrt{121} * \sqrt{(.10)(87)^2 + (.30)(-3)^2 + (.60)(-13)^2}$$

$$\sqrt{121} * \sqrt{861} = \boxed{\$322{,}771}$$

d. Suppose a large employer knows it will get sued 121 times and has set aside $500,000 to pay potential awards. Calculate the chance that the employer has not set aside enough money to pay for the awards.

> 500,000



$$z = \frac{500{,}000 - 363{,}000}{322{,}771}$$

$= +.42 \sim +.40$

$$\frac{100 - 31.08}{2} = \boxed{34.46\%}$$

20. The most recent census of Los Angeles, California revealed that 63% of the residents in the city identified their race/ethnic background as "Hispanic", 15% identified their race/ethnic background as "Non-Hispanic White", 12% as "Asian", 8% as "Black" or "African American and 2% as "Other". Next month, a research group at the UCLA Medical School plans to take a simple random sample of 300 residents in Los Angeles

1

everyone else

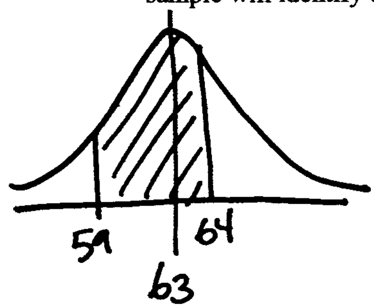
1	0
.15	.85

a. Can you calculate the standard error for the total number of residents in the sample identifying themselves as "Non-Hispanic White"? If it is possible please write "possible" below and justify your response. If it is not possible, please write "not possible below" and justify your response.

$\sqrt{300} \times \sqrt{.15 \times .85} = 6.18$

Possible. It's like a "box" w/ NHW = 1 and all others = 0.

b. What is the chance that between 59% and 64% of residents in the UCLA Medical School sample will identify themselves as "Hispanic"?



$$z_{64} = \frac{64-63}{\left(\frac{\sqrt{300} \times \sqrt{.63 \times .37}}{300} \times 100\right)} = \frac{1}{2.7875} \approx +.35$$

$$z_{59} = \frac{59-63}{\left(\frac{\sqrt{300} \cdot \sqrt{.63 \times .37}}{300} \times 100\right)} = \frac{-4}{2.7875} \approx -1.45$$

$$\frac{85.29}{2} + \frac{27.37}{2} = \boxed{56.33\%}$$

c. If the UCLA Medical School increases the sample size to 900, the expected percentage of residents in the sample identifying themselves as "Other" is expected to: (3 points)

(a) Increase
(b) Decrease
(c) Stay the same
(d) Double
(e) Triple

d. What is the chance that in a sample of 300 residents, 27 or more will identify themselves as "Black" or "African American"? If this is calculable, please show how to calculate the chance below, if it is not, please write "not calculable" and justify your response (5 points).

calculable

$9\% = \frac{27}{300}$

$$\frac{27-24}{\sqrt{300} * \sqrt{.08 \times .92}} = .64 \approx .65 = z$$

OR

8% of 300 = 24 expected number

OR

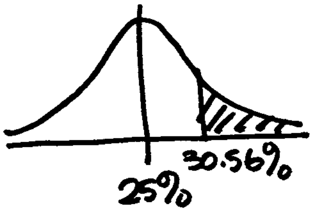
$$\frac{9\% - 8\%}{\left(\frac{\sqrt{300} * \sqrt{.08 \times .92}}{300} \times 100\right)} = .64 \approx .65 = z$$

$$\frac{100 - 48.43}{2} = \boxed{25.785\%}$$

21. The Dull Computer Company makes its own computers and delivers them directly to customers who order them via the Internet. The CEO of Dull has stated publicly that if customers make unassisted online purchases of their computers, they will have a mean delivery time of 46 hours from time of purchase with a standard deviation of 16 hours. He also went on to state that at least 25% of the computers ordered online are delivered in less than 72 hours. It was noted that delivery times are not normally distributed.

A consumer research organization decided to test the CEO's claim by purchasing 36 computers at random from Dull as unassisted online purchases. Treat this sample as if it were a simple random sample (SRS). The mean delivery time of the 36 was 51 hours with a standard deviation of 18 hours. They also had a mean cost of $1,588 with a standard deviation of $276. 33% of the 36 computers were delivered in less than 72 hours. It was noted that the cost variable appeared to be normally distributed.

(a) What is the probability that at least 11 of the computers purchased by the consumer research organization will be delivered in less than 72 hours? (10 points)



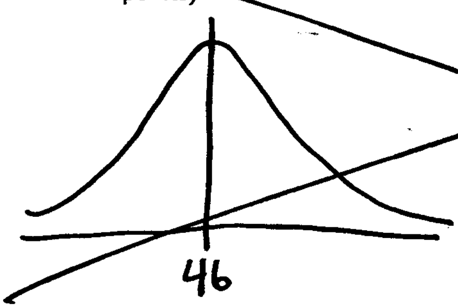
$\frac{11}{36} = .3056$ or 30.56%

$$z = \frac{30.56 - 25}{\left(\frac{\sqrt{36} \cdot \sqrt{.25 \times .75}}{36} \times 100\right)} = .77 \sim .75$$

Area is 54.67

So answer is $\frac{100 - 54.67}{2} = \boxed{22.67\%}$

(b) What percentage of purchases of 36 computers will have delivery times of 48 hours or more? (5 points)



Ch. 23 question

BAD QUESTION

(c) Can you construct a valid 95% confidence interval for the mean cost of a Dull Computer purchased online without assistance?

(circle one) YES NO (1 point)

If yes, please construct the confidence interval below using information from the consumer research organization's purchase. If no, do not construct a confidence interval, instead justify your "no" answer using English in the space below (be brief) (4 points).

Ch. 23 question

BAD QUESTION

22. A box has average = 5 and SD = 2. A hundred draws are made with replacement.

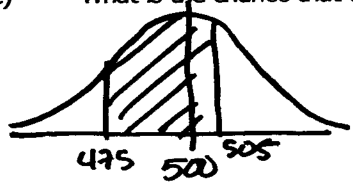
a) What is the expected value for the sum of 100 draws?

E.V.SUM = 100 × 5 = 500

b) What is the standard error for the sum?

SE.SUM = $\sqrt{100} * 2 = 20$

c) What is the chance that the sum of these draws is between 475 and 505?



$z_{505} = \frac{505-500}{20} = \frac{5}{20} = +.25$ area is 19.74

$z_{475} = \frac{475-500}{20} = \frac{-25}{20} = -1.25$ area is 78.87

$\frac{19.74}{2} + \frac{78.87}{2} = \boxed{49.31\%}$

23. In a certain precinct, 80% of the voters are Republican. A simple random sample of size 400 is drawn with replacement. Each person in the sample is polled and the percentage of Republicans in the sample is calculated.

a) What is the expected value for the percentage of Republicans in the sample?

80% = box %

REP

all others

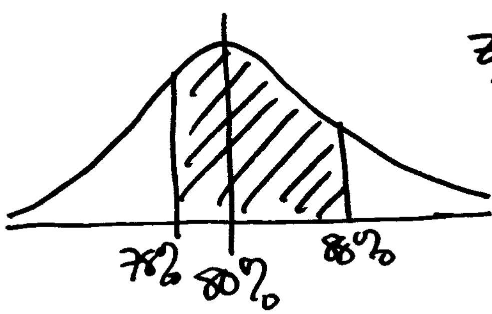
$\boxed{1}$ $\boxed{0}$

.80 .20

b) What is the standard error for the percentage of Republicans?

$\left(\frac{\sqrt{400} \times \sqrt{.8 \times .2}}{400} \times 100 \right) = 2\%$

c) What is the chance that this percentage is between 78% and 88%?



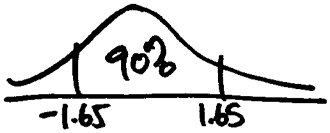
$z_{88} = \frac{88-80}{2} = +4$ 99.9937%

$z_{78} = \frac{78-80}{2} = -1$ 68.27%

$\frac{99.9937}{2} + \frac{68.27}{2} = \boxed{84.132\%}$

24. In a simple random sample of 400 Los Angeles residents taken by a polling organization, only 30% expressed support for the mayor. Find a 90% confidence level for the corresponding percentage in the whole population of Los Angeles

support | no support
$\boxed{1}$ $\boxed{0}$
.3 .7



so

$$30\% \pm (1.65)(2.29)$$

$$SE_{\%} = \left(\frac{\sqrt{400} * \sqrt{.3 \times .7}}{400} \times 100 \right) = 2.29\%$$

$$\boxed{30\% \pm 3.78\%}$$

25. You pay \$1 to roll a pair of dice. If you roll a sum of 7 you receive \$5. For all other rolls you receive nothing. Suppose the chance of rolling a sum of 7 is 1/6. You plan to roll the die 16 times as a game.

$\boxed{+5}$ $\boxed{-1}$
$\frac{1}{6}$ $\frac{5}{6}$

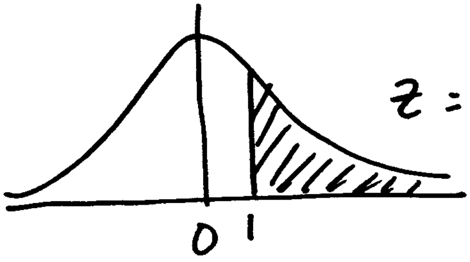
a) The expected value of this game is approximately:

\$0 $\quad EV_{SUM} = 16 \times \left(5 \times \frac{1}{6} + -1 \times \frac{5}{6}\right)$

b) The standard error of this game is approximately:

$$SE_{SUM} = \sqrt{16} \times \left((5 - -1)\sqrt{\frac{1}{6} \times \frac{5}{6}}\right)$$
$$= 8.9443$$

c) The chance that you will win at least one dollar playing this game is approximately:



$$z = \frac{1-0}{8.9443} \stackrel{+}{=} .11 \leadsto \stackrel{+}{.}10$$

area is

$$\frac{100 - 7.97}{2} = \boxed{46.015\%}$$