

# November Page 1

Statistics 100

Exam 3

November 11<sup>th</sup>, 2015

## Question 1 (8 points total)

In tossing a fair coin, follow the steps to find the standard error (SE) that makes the chances equally likely in both scenarios.

- a) 50%  $\pm$  12% heads in 16 tosses is about as likely as getting 50%  $\pm$  ? heads in 256 tosses.

Step 1: Compare the number of tosses in both cases. The number of tosses (n) is increasing by a factor of 16

Step 2: This means that we are going to: Multiply or Divide by 4 (Fill in the blank with a number)  
(Circle one)

Step 3: Your new SE is 3 (Fill in the blank with a number)

- b) 10  $\pm$  2 heads in 20 tosses is about as likely as getting 810  $\pm$  ? heads in 1620 tosses.

Step 1: Compare the number of tosses in both cases. The number of tosses (n) is increasing by a factor of 81

Step 2: This means that we are going to: Multiply or Divide by 9 (Fill in the blank with a number)  
(Circle one)

Step 3: Your new SE is 18 (Fill in the blank with a number)

## Question 2 (12 points total)

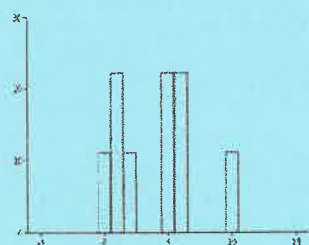
Look at the 3 boxes and 6 probability histograms below. Each box has 2 probability histograms associated with it. One is the probability histogram for the sum of 2 draws made at random with replacement and one is the probability histogram for the sum of 15 draws made at random with replacement. Assume there are 3 tickets in each box.

Box A  
0 1 2

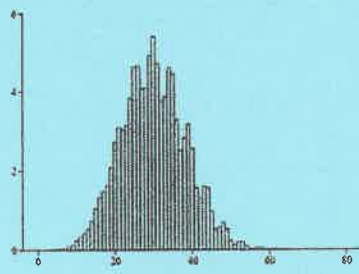
Box B  
0 1 5

Box C  
0 1 15

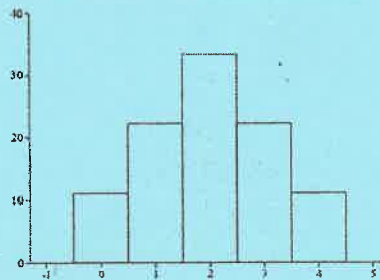
Under each of the 6 histograms, fill in the first blank with either 2 or 15 and the second blank with either A, B, or C.



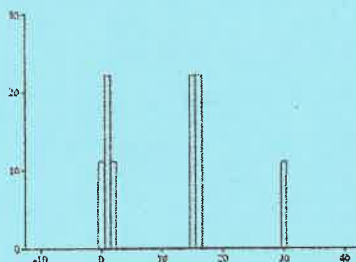
2 draws from Box B



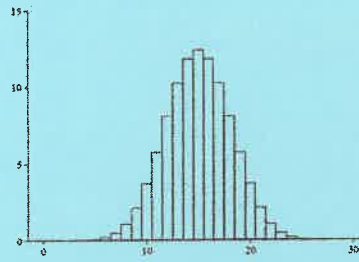
15 draws from Box B



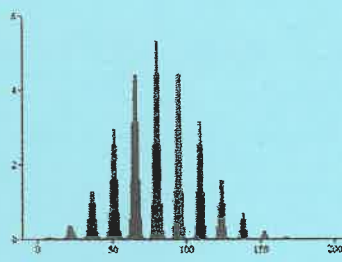
2 draws from Box A



2 draws from Box C



15 draws from Box A



15 draws from Box C

# November Page 2

Statistics 100

Exam 3

November 11<sup>th</sup>, 2015

## Question 3 (14 points total-one for each blank)

Fill in the first blank with the number of draws, the second blank with the word "with" or "without", and the third with the letter corresponding to the appropriate box model. Choose from the box models below and use each box exactly once.

Box A

1 2 3 4 5 6

Box B

0 1

Box C

3 3  
1 -1

Box D

0 0 0 1 1 0

- A die is rolled 30 times and the sum of the spots is counted.  
This corresponds to drawing 30 times with replacement from Box A.
- A die is rolled 6 times and the total number of 4's and 5's are counted.  
This corresponds to drawing 6 times with replacement from Box D.
- A die is rolled 90 times and you win \$1 if you roll an even number, but lose \$1 if you roll an odd number.  
This corresponds to drawing 90 times with replacement from Box C.
- Suppose you guess on a multiple choice test has 45 true/false questions and your number of correct answers is counted.  
This corresponds to drawing 45 times with replacement from Box B.
- What's the SD of Box B? 0.5 What's the SD of Box C? 1

## Question 4 (14 points total)

100 draws are made at random with replacement from the box containing these 5 tickets: 1 2 3 4 5

- (2 points) The smallest the sum of the 100 draws could possibly be is 100 and the largest is 500.  
(Fill in the 2 blanks above with the correct numbers)
- (2 points) What is the EV for the sum of the draws? Show work below and circle your answer.

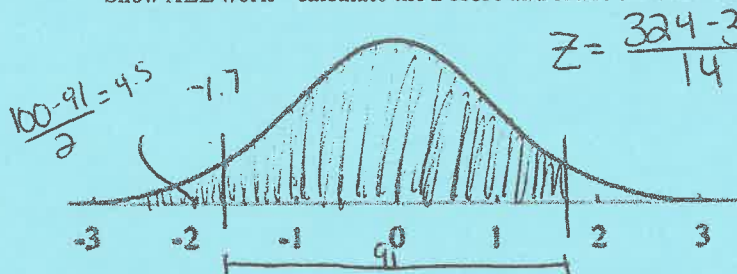
$$EV_{\text{sum}} = n \times \text{ave of box} = 100 \times 3 = \textcircled{300}$$

- (2 points) What is the SE for the sum of the draws? (The SD of the box is 1.4)  
Show work below and circle your answer.

$$SE_{\text{sum}} = \sqrt{n} \times SD \text{ of box} = \sqrt{100} \times 1.4 = \textcircled{14}$$

- (3 points) Use the normal approximation to estimate the chance that the sum of the draws will be less than 324?

Show ALL work—calculate the z-score and shade the correct area on the curve



$$z = \frac{324 - 300}{14} = 1.7$$

Chance = 95.5 %

$$91 + 4.5 = 95.5$$

- Now suppose you draw 100 times at random with replacement from the same box above, but this time you're only interested in the percent of 2's you get. What is the EV and SE of the percent of 2's in 100 draws? (Hint: draw a new box)

- (2 points) What is the expected value of the percent of "2"s in 100 draws? 20%

- (3 points) What is the SE for the percent of "2"s in 100 draws? Show work. 4%

1 point

1 4  
1 0

5 tickets

$$SD = |1 - 0| \sqrt{\frac{1}{5} \times \frac{4}{5}} = 0.4$$

1 point

$$SE\% = \frac{SD}{\sqrt{n}} \times 100$$

$$= \frac{0.4}{\sqrt{100}} \times 100 = \textcircled{4\%}$$

3 of 7 pages (9 problems)

1 point



# November Page 3

Statistics 100

Exam 3

November 11<sup>th</sup>, 2015

Question 5 pertains to some of the important concepts we've covered in Chapters 13-19. (11 points total)

- a) (6 points total) Suppose you play a game that involves a bag of marbles. There are 17 blue marbles, 4 green marbles, and 1 black marble. If you get a blue marble you lose \$1, if you get a green marble you win \$5, and if you get the black marble you win \$8. Suppose you randomly draw from the bag 20 times with replacement and your winnings are counted. Draw the appropriate box model to represent this scenario below.

17	4	1
-1	5	8

- b) (2 points) What is the average of this box to the left?

Show work below. Circle Answer.

$$\frac{17(-1) + 4(5) + 1(8)}{22} = \frac{11}{22} = 0.5$$

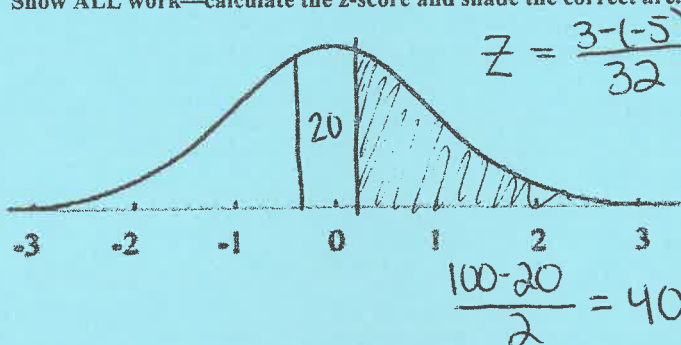
- c) (1 point) The concept that says that as  $n$  increases, the probability histograms for all sums, averages, or percentages will approach the normal curve is called?
- a. The Law of Averages    ii) Confidence Intervals    iii) The Central Limit Theorem
- d) (1 point) If we drew 15 times out of Box A: 1, 2, 3, 4 and 15 times out of Box B: 1, 16, 33, 33 and drew the corresponding probability histograms, which histogram would look more like the normal curve?
- a. Box A    ii) Box B    iii) Impossible to tell
- e) (1 point) A parameter is a numerical fact about the sample.
- a. True    ii) False

Question 6 (11 points total)

A gambler plays roulette 100 times betting \$1 on the numbers 1, 2, and 3 each time. If the ball lands on 1, 2, or 3, the gambler wins \$11 and if the ball lands on any of the other numbers, the gambler loses \$1. The roulette wheel has 38 slots number 1-36, 0, and 00.

- a) (2 points) Which is the appropriate box model?
- a. The box has 38 tickets: 1 marked "1", 1 marked "2", 1 marked "3", and 35 marked "-1"
- b. The box has 100 tickets, some marked "11" and some marked "-1", but the exact percentages are unknown
- c. The box has 100 tickets, half marked "1" and half marked "0"
- d. The box has 38 tickets, 3 marked "11" and 35 marked "-1"
- b) (2 points) This corresponds to taking 100 draws with replacement.  
Fill in the 1<sup>st</sup> blank with a # and the second blank with either with or without.
- c) (2 points) What is the average of the box? Show work for credit and circle your answer. Leave your answer as a fraction.
- $$\frac{3(11) + 35(-1)}{38} = -\frac{2}{38}$$
- d) (2 points) What is the SD of the box? Show work for full credit and circle your answer.
- $$\sqrt{11(-1) \left| \sqrt{\frac{3}{38} \times \frac{35}{38}} \right|} = 3.24$$
- Don't take off for rounding!
- e) (3 points) Use the normal approximation and the fact that the EV is about -\$5 and the SE is about \$32 to figure out the chance that the gambler will win more than \$3 in 100 plays.

Show ALL work—calculate the z-score and shade the correct area on the curve



Chance = 40 %

# November Page 4

Statistics 100

Exam 3

November 11<sup>th</sup>, 2015

## Question 7 (10 points total)

A recent poll was posted on gretwire.com asking the following question: "If the election were tomorrow, would you vote for Donald Trump or another candidate?" Anyone who visited the website could vote. On Bonus Survey 3, we asked the class who they would vote for as well. Anyone in the class could take the survey. Newsweek recently took a random sample of 1015 adults nationwide and asked them the same question. Here are the results of all 3 polls.

	Yes for Trump	Another Candidate	n
Gretawire.com	61%	39%	4,066
Bonus Survey 3	8%	92%	707
Newsweek	35%	65%	1,015

- a) (2 points) As you can see, the results of the 3 polls are quite different. Which survey gives a better estimate of the percentage of all US adults who would answer "Yes" to voting for Donald Trump?
- The gretawire poll since the sample size is the largest.
  - Bonus Survey 3 because the responses are the most honest.
  - iii The newsweek poll since the sample was randomly selected.
  - All 3 of them will give pretty accurate results.
- b) (2 points) For which poll is it appropriate to calculate the SE of the sample percent?
- i The newsweek poll since the people were randomly selected
  - Bonus Survey 3 since the results were anonymous
  - All 3 polls
  - None of the 3 polls
- c) Now suppose that CNN/Time also conducted a random poll of 888 likely Republican voters and asked them the same question: "If the election were tomorrow, would you vote for Donald Trump?" 53% of the sample said yes they would vote for Donald Trump.
- i) (2 points) What is the SE of the sample percent? Choose one:
- 0.499%
  - 0.0167%
  - iii 1.67%
  - 3.34%
- ii) (2 points) A 90% confidence interval for the percent of all Republican voters who would vote for Donald Trump is
- 50.24, 55.76 (Put the lower number first. Show work for full credit.  
Round to 2 decimal places)
- 90% CI =  $53 \pm 1.65(1.67)$
- iii) (1 point) If 100 pollsters all took samples of 888 likely Republican voters and created 90% confidence intervals, what percent of the confidence intervals would capture the true population percentage?
- 100
  - ii 90
  - 10
  - Impossible to tell
- iv) (1 point) Would CNN/Time Poll described above have selection bias?
- i No because the people were randomly selected/chosen.
  - Yes because there are some biases we can't avoid.

## Question 8 (4 points total)

Say that my fiancé, Steve is starting a new campaign running for the mayor of Urbana. For a pre-election poll in a close race, we may want a 95% confidence interval with a small margin of error.

- a) (2 points) Estimate how many people you'd need to poll to get a 95% confidence interval with only a 2% margin of error. (Assume the SD of the population is around 0.41. Show work, circle answer, and round to the nearest whole number)
- $$n = \left( \frac{200 \times 0.41}{2} \right)^2 = \textcircled{1681}$$
- b) (2 points) Estimate how many people you'd need to poll to get a 95% confidence interval with only a 6% margin of error. (Assume the SD of the population is around 0.5. Show work, circle answer, and round to the nearest whole number)
- $$n = \left( \frac{200 \times .5}{6} \right)^2 = \textcircled{278}$$



# November Page 5

Statistics 100

Exam 3

November 11<sup>th</sup>, 2015

## Question 9 (16 points total)

A CBS Poll conducted last year asked a random sample of 1,048 Illinois adults the following question: "On average, how much do you spend per month on coffee?" The sample average was \$35 with an SD of \$7.

- a) (2 points) What most closely resembles the relevant box model? Circle one.
- i) It has 1048 tickets. The average is 35 and the SD is 7.
  - ii) It has millions of tickets. Exactly 35% are marked "1" and 65% are marked "0"
  - iii) It has millions of tickets marked "1" and "0" and the exact percentages are unknown, but estimated from the sample
  - ☒ iv) It has millions of tickets, each with a dollar amount on them. The exact average and SD are unknown, but estimated from the sample.

- b) (2 points) The 1,048 draws are made without replacement.  
(Fill in the first blank with the number of draws and the second blank with either "with" or "without")

- c) (2 points) We'd expect the true average dollar amount all Illinois adults spend on coffee to be? \$35

- d) (2 points) What is the SE of the sample average? 0.22  
Show work for credit.

$$SE_{ave} = SD / \sqrt{n} = 7 / \sqrt{1,048} = 0.22$$

- e) (2 points) Suppose we created a 95% CI for the average amount of money Illinois adults spend on coffee each month. To which of the following populations can we apply that 95% CI?

- i) All US adults
- ii) All adults in the state of Wisconsin
- iii) All females in the US
- ☒ iv) None of the above
- v) All of the above

- f) (2 points) Suppose we take another poll in the entire US instead of just Illinois. In order to obtain the same accuracy as in the first poll only in IL, the sample size of the second poll in the US should:

- ☒ a) Stay the same
- b) Increase
- c) Decrease
- d) need more info to answer

- g) (2 points) If the study asked the 1048 Illinois adults whether or not they drink coffee, the relevant box model would contain tickets with:

- ☒ i) Only "1"s and "0"s
- ii) Numbers ranging from about 0 to 100
- iii) not enough info

- h) (1 point) If the researcher increased the sample size (n) by a factor of 100, the SE for the average would...

- i) be multiplied by 10
- ii) be multiplied by 100
- ☒ iii) be divided by 10
- iv) be divided by 100
- v) stay the same

- i) (1 point) If the researcher increased the sample size (n) by a factor of 100, the length of a 95% confidence interval would...

- i) be multiplied by 10
- ii) be multiplied by 100
- ☒ iii) be divided by 10
- iv) be divided by 100
- v) stay the same