PAGE 1 - 9

Statistics 100 Exam 3 November 9, 2016

Question 1 *(6 pts. total)*
Fill in the blanks to make the chances equally likely in both scenarios.
We are looking at flipping a coin 56 times vs. 1400 times.

a) The number of tosses (n) is increasing by a factor of? 25

b) 28 ± 4 heads in 56 tosses is about as likely as 700 ± 20 heads in 1400 tosses.
FILL IN THE FIRST BLANK WITH THE NEW EV and THE SECOND BLANK WITH THE NEW SE THAT WILL MAKE THE CHANCES EQUAL.

Next, we are looking at flipping a coin 100 times vs. 1600 times.

c) The number of tosses (n) is increasing by a factor of? 16

d) 50% ± 8% heads in 100 tosses is about as likely as 50 % ± 2 % heads in 1600 tosses.
FILL IN THE FIRST BLANK WITH THE NEW EV and THE SECOND BLANK WITH THE NEW SE THAT WILL MAKE THE CHANCES EQUAL.

Question 2 *(20 pts. total)*
Fill in the first blank with the number of draws, the second with either "with" or "without" and the third with the letter corresponding to the appropriate box model. Choose from the box models below. Use each box model exactly once.

Box A: 1 -1/2 -1/2 Box B: 0 1 Box C: 1 2 3 4 5 6 Box D: 0 0 0 1 0 0

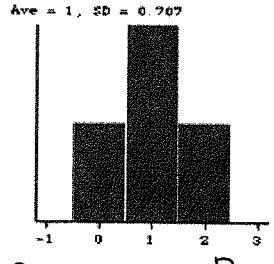
a) A fair coin is tossed 25 times and the number or heads is counted. This corresponds to drawing 25 times with replacement from Box B

b) A pair of dice is rolled once and the total number of spots is counted. This corresponds to drawing 2 times with replacement from Box C

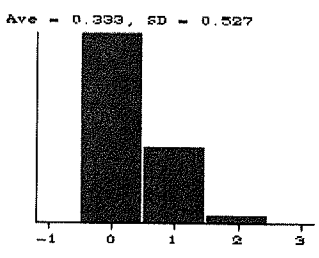
c) A die is rolled 50 times and the number of 4's is counted. This corresponds to drawing 50 times with replacement from Box D

d) A multiple choice test has 100 questions. Each question had 3 options, only one of which is right. Suppose you randomly guess on all 100 questions, if you get a question right you get 1 point and if you get a question wrong you lose half of a point.
This corresponds to drawing 100 times with replacement from Box A

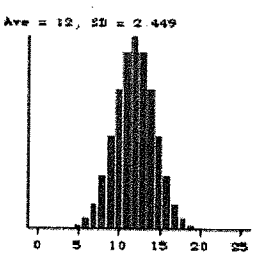
e) Look at Box B and Box D above. The 4 histograms below are the probability histograms for the sum of 2 draws from Box B, 2 draws from Box D, 24 draws from Box B and 24 draws from Box D. Which is which?
Fill in the blanks below



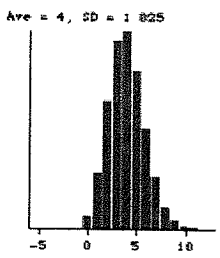
2 draws from Box B



2 draws from Box D



24 draws from Box B



24 draws from Box D

Statistics 100 Exam 3 November 9, 2016

Question 3 *(12 pts. total)*

400 draws are made at random with replacement from the box containing these 5 tickets: 2 3 4 5 6

a) *(2 pts.)* The smallest the sum of the 400 draws could possibly be is 800 and the largest is 2400.
(Fill in the 2 blanks above with the correct numbers)

b) *(2 pts.)* What is the EV for the sum of the draws? **Show work. Circle answer.** $-\frac{1}{2}$ if no work

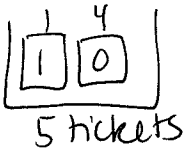
$EV_{sum} = n \times \text{avg of box} = 400 \times 4 = (1600)$

c) *(2 pts.)* What is the SE for the sum of the draws? (The SD of the box is 1.4)
Show work. Circle answer.

$SE_{sum} = \sqrt{n} \times SD \text{ of box} = \sqrt{400} \times 1.4 = (28)$

d) Now suppose you draw at random with replacement from the same box above, but this time you're only interested in the percent of 3's that you get. What is the EV and the SE of the percent of 3's in 400 draws?
(Hint: draw a new box)

i) *(2 pts.)* What is the expected value of percent of 3's in 400 draws? 20 %



ii) *(2 pts.)* *What is the SD of your new box?* 0.4.
Show work. $SD = |1-0|\sqrt{\frac{1}{5} \times \frac{4}{5}} = 0.4$

iii) *(2 pts.)* What is the SE for the percent of 3's in 400 draws? ***Choose one.***

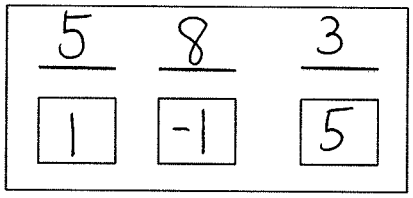
a) 33% (b) 2% c) 4.67% d) 7% e) 20% f) 40%

$SE_{\%} = \frac{SD}{\sqrt{n}} \times 100 = \frac{0.4}{\sqrt{400}} \times 100 = 2$

Question 4 *(9 pts. total)*

Suppose you randomly draw 81 marbles with replacement from a bag that contains 5 red marbles, 8 blue marbles, and 3 green marbles. If a red is drawn you win \$1, if a blue marble is drawn you lose \$1, if a green marble is drawn you win \$5.

a) *(3 pts.)* Draw the appropriate box model by labeling the 3 tickets inside the box on the right with the correct numbers and writing how many of each ticket in the blanks above them.
½ pt for each #



b) *(2 pts.)* What is the average of the box? 0.75
Show work.

$\text{avg} = \frac{5(1) + 8(-1) + 3(5)}{16} = \frac{12}{16} = 0.75$

c) *(2 pts.)* What is the expected value of the sum of 81 draws (your net gain/loss)? \$ 60.75 continued error
Show work.

$EV_{sum} = n \times \text{avg} = 81 \times 0.75 = 60.75$

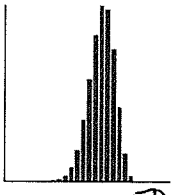
d) *(2 pts.)* The SD of the box is \$2.22. What is the SE of the sum of your winnings? \$ 19.98
Show work.

$SE_{sum} = \sqrt{n} \times SD = \sqrt{81} \times 2.22 = 19.98$

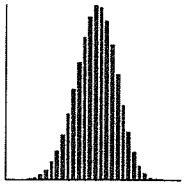
Question 5 *(3pts. total)*

The 3 histograms below (in scrambled order) are the probability histograms for the sum of **25, 50 and 150** random draws with replacement from a box that has 10 tickets 1 marked "0" and 9 marked "1". Which histogram depicts **25** draws, which **50** draws and which **150**? ***Fill in each blank below with the correct number of draws.***

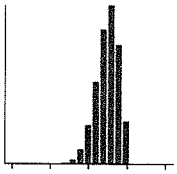
Histogram A



Histogram B



Histogram C



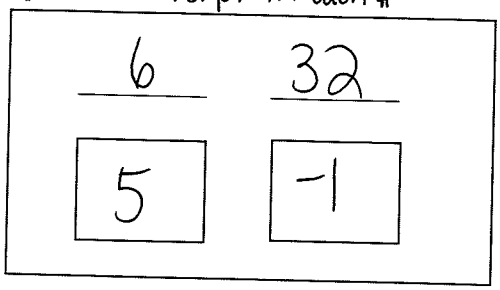
a) Histogram A is the sum of 50 draws **b)** Histogram B is the sum of 150 draws **c)** Histogram C is the sum of 25 draws

Statistics 100 Exam 3 November 9, 2016

Question 6 *(14 pts. total)*

A gambler plays roulette 100 times betting \$1 on 6 numbers, 1-6, each time. If the ball lands on any of those 6 numbers, th gambler wins \$5, if the ball lands on any of the other 32 numbers, the gambler loses \$1. The roulette wheel has 38 slots numbered 1-36, 0 and 00.

a) Draw the appropriate box model.
(2 pts.) Write a number in each blank and each square below. ½ pt for each #



b) *(1 pt.)* How many draws are made from this box? 100

c) *(1 pt.)* The draws are made… **circle one**
(i) with replacement
ii) without replacement

d) *(2 pts.)* What is the average of this box? **Show work below, write your answer as a fraction, and circle the answer.**

$$\text{avg} = \frac{6(5)+32(-1)}{38} = \left(\frac{-2}{38}\right)$$

-½ for no work

e) *(2 pts.)* What is the SD of the box? **Show work below and circle answer.**

$$SD = |5-(-1)|\sqrt{\frac{6}{38}\times\frac{32}{38}} = (2.19)$$

f) Use the normal approximation and the fact that the EV = \$ - 5 and the SE = \$22 (approximately) to figure the chance that the gambler will win more than \$39 in 100 plays?

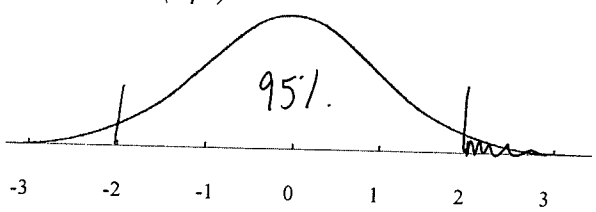
i) *(2 pts.)* First calculate the Z score. Show work. Circle answer.

$$Z = \frac{39-(-5)}{22} = 2$$

ii) *(2 pts. total)* Now accurately mark z on the normal curve below and shade the areas corresponding to the chance that the gambler will win more that \$39 in 100 plays. *(1 pt.)*

Chance = 2.5 % *(1 pt.)* continued error

$$\frac{100-95}{2} = 2.5$$



g) *(2 pts.)* Now suppose we were interested in how many times we'd expect the gambler to win playing 100 times (instead of how many dollars we'd expect him to win).
Which is the appropriate box model?
Circle one:

i) The box has 38 tickets: 6 marked "5" and 32 Marked "0"
ii) The box has 38 tickets: one each of 1, 2, 3, ..., 36, 0, and 00.
iii) The box has 38 tickets: 6 marked "1" and 32 marked "0"
iv) The box has 38 tickets: 6 marked "1" and 32 marked "-1"

Statistics 100 Exam 3 November 9, 2016

Question 7 *(14 pts. total)*

At the last grading meeting we graded a random sample of **64 of the 1600** exams to see how students would do. The average of the 64 exams was 84 with a SD of 10.

a) *(2 pts.)* Which most closely resembles the relevant box model?
Circle one:
i) The box has 1600 tickets marked with "1"s and "0"s
ii) The box has 64 tickets marked with "1"s and "0"s. The exact percentages are unknown, but estimated from the sample.
iii) The box has 1600 tickets, marked with numbers ranging from 0 to 100, the exact average and SD of the box are unknown and are estimated from the sample.
iv) The box has 1600 tickets with an average of 84 and a SD of 10.

b) *(1 pts.)* How many draws from the box? 64

c) *(1 pts.)* ***Circle one:*** i) with replacement ii) without replacement

d) *(2 pts.)* The best estimate for the **average** exam scores of all 1600 students is 84.

e) *(2 pts.)* What is the **SE** of the **sample average**? ***Show work. Circle answer.*** -1 if they divide by $\sqrt{1600}$

$SE_{avg} = \frac{SD}{\sqrt{n}} = \frac{10}{\sqrt{64}} = 1.25$

f) *(2 pts.)* A 95% CI for the average of all 1600 exams is about (81.5 to 86.5)

95% CI = sample avg $\pm$ 2 SE_{avg} = $84 \pm 2(1.25)$ Continued error

g) *(2 pts.)* Suppose we also computed a 68% CI for the average of all 1600 exams. Which interval would be smaller?
Choose one: i) the 68% CI ii) the 95% CI iii) they'd be the same width iv) impossible to tell

h) *(2 pts.)* We were worried that the true average of the class was less than 80 and we might have to curve the exam. Let's say the true average of all 1600 exams was only 79 with an SD of 10. What is the chance that we'd randomly draw a sample of 64 exams and get an average as high as 84 or more? ***Choose one:*** $Z = \frac{84-79}{1.25} = 4$

i) Not enough information to calculate since we don't know if the exam scores follow the normal curve.

ii) There's less than a 1/10,000 chance of getting an average of 84 or more since 84 converts to a Z score = 4.

iii) There's more than a 30% chance of getting an average of 84 or more, since 84 converts to a Z score of 0.5.

iv) We know that 95% of the 1600 scores are between 59 and 79. Since 84 is well within that range we can be 95% confident that our sample average reflects the true population average.

Question 8

Suppose a government survey organization took a simple random sample of 1000 people in Illinois and computed the $SE_{\%}$ and a 95% Confidence Interval in order to estimate the percentage of all adults who favor the death penalty in that state.

a) (2 pts.) If they decided to increase the sample size to 4000, the new $SE_{\%}$ would...
Choose one: i) stay the same ii) be multiplied by 2 iii) be multiplied by 4 iv) be divided by 2 v) be divided by 4

b) *(2 pts.)* If the sample size was increased to 4000, the width of the new 95% confidence interval would ...
Choose one: i) stay the same ii) be multiplied by 2 iii) be multiplied by 4 iv) be divided by 2 v) be divided by 4

c) *(2 pts.)* Suppose 100 pollsters each randomly sampled 1,000 Illinois adults asking the same question. All 100 pollsters computed 90% confidence intervals to estimate the percentage of all US adults who favor the death penalty.

About how many of the 100 confidence intervals would miss the true population percentage? 10

Question 9 pertains to the following situation: *(6 pts. total)*
After the 3rd presidential debate, 3 polls were taken asking who won: Hillary Clinton or Donald Trump. The Washington T Breitbart Polls posted the question on their websites, allowing anyone who visited to cast their vote. The CNN poll was bas responses of 1,000 randomly selected adults nation-wide who watched the debate. Here are the results of each of the polls:

	Clinton	Trump	Sample Size
Washington Times	17%	83%	25,000
Breitbart Poll	60%	40%	83,000
CNN	57%	43%	1,000

a) *(2 pts.)* As you can see, the results of the 3 polls are quite different. Which poll gives the best estimate of the percentage of all adults who watched the debate who thought Donald Trump won? **Choose one:**
i) The Breitbart Poll because it has the largest sample size.
ii) The CNN Poll because the people were randomly chosen from all adult viewers nation-wide.
iii) Both the Breitbart Poll and the CNN Poll can be trusted since they are within 3% pts. of with each other.
iv) The Washington Times poll because it gives the most definitive result of who won the debate.

b) *(2 pts.)* What is SE of the sample percent for the Washington Times Poll? **Choose one:**
i) It's not possible to calculate a SE for this sample because we don't know the SD of the sample.
ii) It's not possible to calculate a SE for this sample because this poll was not randomly selected.
iii) The SE of the sample percent is approximately 0.23%
iv) The SE of the sample percent is approximately 23%

c) *(2 pts.)* A fourth poll was done by Gallup asking the same question as above. The results were from 1,300 randomly sampled adults nationwide who watched the debate. If we computed a 95% confidence interval for the percentage of people who think Hillary Clinton won the debate, to which of the following populations can we apply that interval?
Choose one.
i) all US adults who plan to vote **ii)** all US adults who watched the debate **iii)** all US adults **iv)** all Hillary supporters

Question 10 *(4 pts. total)*
The following question pertains to a box containing tickets only marked with 1's and 0's. ***Fill in blanks with correct numbers. If there is more than one correct number for any of the blanks just write one of them.***

The smallest the SD of a 0-1 box can be is 0. This would happen when the box has 100 % 1's. or 0%.
The largest the SD of a 0-1 box can be is 0.5. This would happen when the box has 50 % 1's.

Question 11 *(6 pts. total)*
Suppose UCLA and Stanford both decide to do a poll of their undergraduates. Both universities want a margin of error of 5%. The undergrad population at UCLA is about 9 times larger than the undergrad population at Stanford.

a) *(2 pts.)* Other things being equal, in order to obtain the same accuracy in the two polls, the number of people you'd have to poll at UCLA is _____ the number of people you'd have to poll at Stanford.
Choose one:
i) 9 times larger **ii)** 3 times larger **iii)** the same as **iv)** 3 times less than **v)** 9 times less than.

b) *(2 pts.)* How many people would you have to poll get a Margin of Error of 5% at UCLA? (Assume SD=0.5) (Show work)

$$n = \left(\frac{200 \times 0.5}{5}\right)^2 = 400$$

-1 for using 0.05 in denominator

c) *(1 pt.)* Say Stanford changed their minds and wants a smaller margin of error, only 3% , how should they adjust their sample size to get a smaller margin of error? (Assume SD=0.5)

Choose one: i) increase it ii) decrease it iii) keep it the same

d) *(1 pt.)* How many people would you have to poll at Stanford to get a Margin of Error of only 3%? (Assume SD=0.5) (Show work)

$$n = \left(\frac{200 \times 0.5}{3}\right)^2 = 1111$$

-½ for using 0.03 in denominator