STATISTICS 100 EXAM 3 **Spring 2016**

PRINT NAME Version 1 Key _____
(Last name) (First name)

***NETID**_____

CIRCLE SECTION: **Laska MWF L1** **Laska Tues/Thurs L2** **Robin Tu**

Write answers in appropriate blanks. When no blanks are provided **CIRCLE** your answers.
SHOW WORK when requested, otherwise no credit.
Do NOT use scrap paper.

Make sure you have all 7 pages including the normal table (13 problems).

DO NOT WRITE BELOW THIS LINE

The numbers written in each blank below indicate how many points you missed on each page.
The numbers printed to the right of each blank indicate how many points each page is worth.

Page 1 _____20

Page 2 _____27

Page 3 _____19

Page 4 _____19

Page 5 _____ 15

Extra Credit_____

Total Score _____

*Karle Laska's Sections: There is no class tomorrow and Friday! Have a good weekend!

Scores will be posted in Compass early Friday morning ☺

Question 1 *(12 points total)*

Look at the 2 boxes and 6 probability histograms below. Each box has 3 probability histograms associated with it. One is the probability histogram for 1 draw from the box, one is the sum of 3 draws made at random with replacement, and one is the probability histogram for the sum of 30 draws made at random with replacement. Assume there are 3 tickets in box A and 4 tickets in Box B.

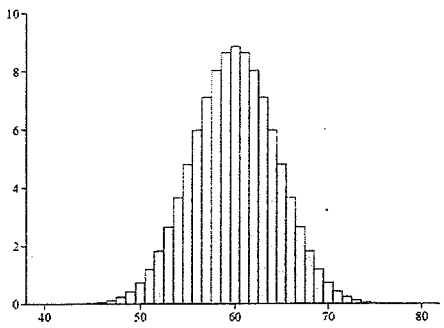
Box A

1	2	3

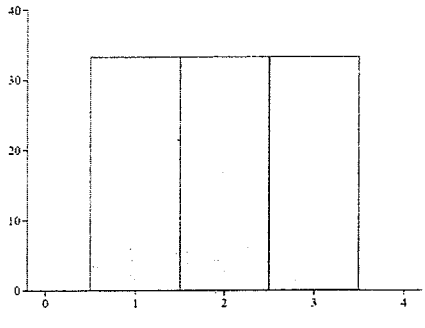
Box B

0	1	9	9

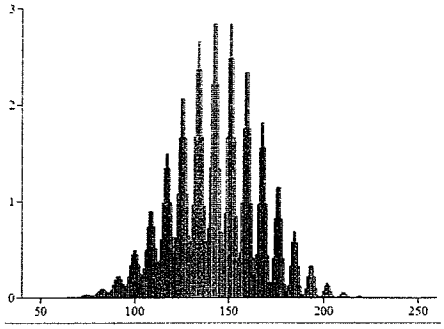
Under each of the 6 histograms, fill in the first blank with either 1, 3, or 30 and the second blank with either A or B.



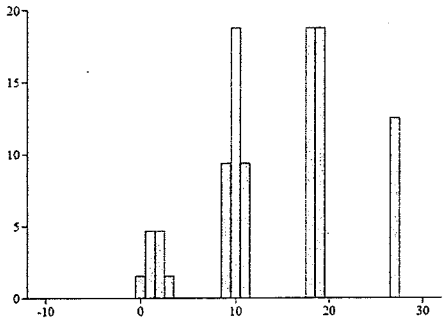
30 draws from Box A



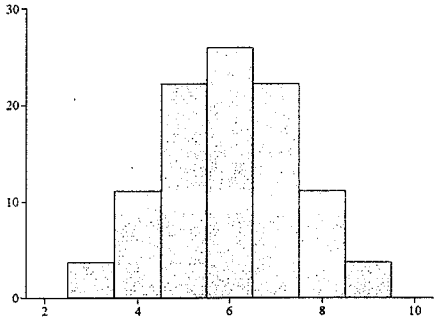
1 draws from Box A



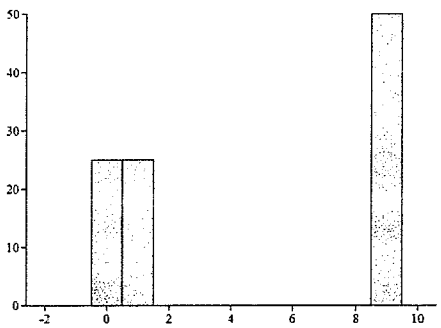
30 draws from Box B



3 draws from Box B



3 draws from Box A



1 draws from Box B

Question 2 *(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% +/- 9% heads in 30 tosses is about as likely as getting 50% +/- **?** heads in 270 tosses.

Step 1: Compare the number of tosses in both cases. The number of tosses (n) is increasing by a factor of 9

Step 2: This means that we are going to: Multiply or (Divide) by 3 **(Fill in the blank with a number)**

(Circle one)

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

b) 50 +/- 5 heads in 100 tosses is about as likely as getting 2450 +/- **?** heads in 4900 tosses.

Step 1: Compare the number of tosses in both cases. The number of tosses (n) is increasing by a factor of 49

Step 2: This means that we are going to: (Multiply) or Divide by 7 **(Fill in the blank with a number)**

(Circle one)

Step 3: Your new SE is 35 **(Fill in the blank with a number)**

Question 3 *(12 points total)*
Fill in the first blank with the number of draws and 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	Box B	Box C	Box D
	12 26	1 5	
1 2 3 4 5 6	2 -1	1 0	0 0 0 1 0

a) A die is rolled 2 times and the sum of the spots is counted.
This corresponds to drawing 2 times with replacement from Box A.

b) A die is rolled 10 times and the total number of 6's are counted.
This corresponds to drawing 10 times with replacement from Box C.

c) Suppose you play roulette 48 times and your total winnings are calculated. There are 38 slots on the roulette wheel and if your ball lands on 1-12, you win $2, otherwise you lose a dollar.
This corresponds to drawing 48 times with replacement from Box B.

d) Suppose you guess on a multiple choice test that has 100 questions and your number of correct answers is counted. Each question has 5 choices- one right answer and four wrong answers.
This corresponds to drawing 100 times with replacement from Box D.

Question 4 *(15 points total)*
25 draws are made at random with replacement from the box containing these 5 tickets: 0 2 3 4 6

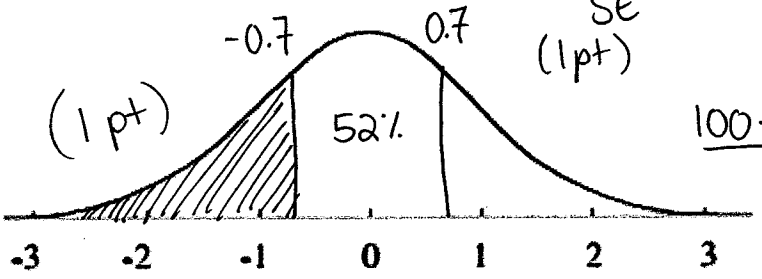
a) *(2 points)* The smallest the sum of the 25 draws could possibly be is 0 and the largest is 150.
(Fill in the 2 blanks above with the correct numbers)

b) *(2 points)* What is the EV for the sum of the draws? **Show work below and circle your answer.**
$EV_{sum} = n \times \text{ave of box} = 25 \times 3 = (75)$

c) *(2 points)* What is the SE for the sum of the draws? (The SD of the box is 2)
Show work below and circle your answer.
$SE_{sum} = \sqrt{n} \times \text{SD of box} = \sqrt{25} \times 2 = (10)$

d) *(4 points)* Use the normal approximation to estimate the *chance* that the sum of the draws will be less than 68? **Show ALL work— 1 point for each: calculate the z-score, mark it on the curve, shade correctly, and find the chance.**

$z = \frac{\text{val} - EV}{SE} = \frac{68-75}{10} = -0.7$ (1pt)



$\frac{100-52}{2} = 24\%$

Chance= 24 % or 24.5 (1pt)

e) Now suppose you draw at random with replacement from the same box above, but this time you're only interested in counting how many 2's you get. What is the EV and the SE of the number of 2's in 25 draws?
(Hint: draw a new box)

i) *(2 points)* What is the expected value of the number of "2"s in 25 draws? 5

ii) *(3 points)* What is the SE for the number of "2"s in 25 draws? **Show work and circle answer.**

1 4
1 0
5 tickets

$EV_{sum} = n \times \text{ave of box} = 25 \times \frac{1}{5} = 5$

$SE_{sum} = \sqrt{n} \times SD \text{ of box} = \sqrt{25} \times 0.4 = (2)$

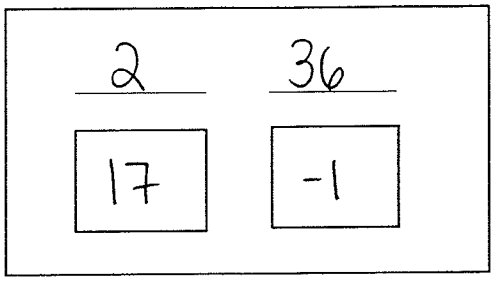
$SD = |1-0|\sqrt{\frac{1}{5} \times \frac{4}{5}} = 0.4$

if EV + SE for ave or percent are calculated -2 total

Question 5 *(15 points total)*

A gambler plays roulette 100 times betting \$1 on two numbers, 7 and 11, each time. If the ball lands on either 7 or 11, the gambler wins \$17, if the ball lands on any of the other 36 numbers, the gambler loses \$1. The roulette wheel has 38 slots numbered 1-36, 0 and 00.

a) Draw the appropriate box model. *(4 points)*



b) How many draws are made from this box? 100

c) The draws are made… **circle one**

i) with replacement
ii) without replacement

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

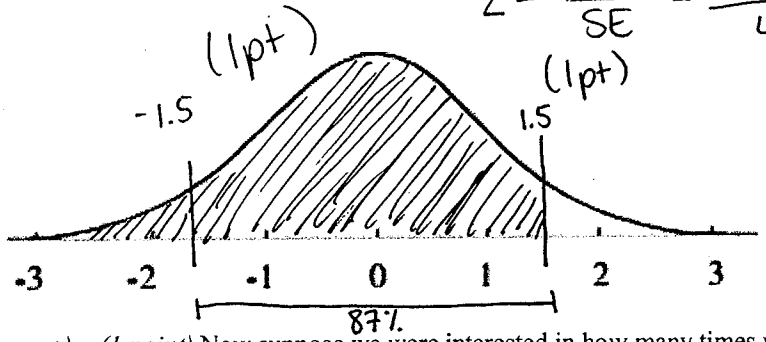
$$\text{ave} = \frac{2(17) + 36(-1)}{38} = \left(\frac{-2}{38}\right)$$

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

$$SD = |17-(-1)|\sqrt{\frac{2}{38} \times \frac{36}{38}} = (4.02)$$

f) *(4 points)* Use the normal approximation and the fact that the EV= \$ -5 and the SE = \$40 to figure the chance that the gambler will win less than \$56 in 100 plays? **Show all work- calculate the z-score, mark it on the normal curve, shade the appropriate region, and calculate the chance.**

$$Z = \frac{\text{val}-EV}{SE} = \frac{56-(-5)}{40} = 1.5 \text{ or } 1.55 \quad (1pt)$$



Chance=93.5 % (1pt)

accept 93-94

$87 + \frac{100-87}{2} = 93.5$

g) *(1 point)* 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: 2 marked "1" and 36 marked "-1"
ii) The box has 38 tickets: one each of 1, 2, 3, ..., 36, 0, and 00.
iii) iii) The box has 38 tickets: 2 marked "17" and 36 marked "0"
iv) iv) The box has 38 tickets: 2 marked "1" and 36 marked "0"

Question 6 *(4 points total)*

A recent Fox News Poll asked a random sample of 600 adults nationwide the following question: "Would you rather eat at McDonald's or Burger King?" At the same time, I asked my classes the same question on Bonus Survey 4. Anyone in the class could participate and earn extra credit. Here are the results of both surveys:

	McDonald's	**Burger King**	**n**
Fox News Poll	67%	33%	600
Bonus Survey 4	75.65%	24.35%	772

(2pts) *a)* As you can see, the results of the 2 polls are a bit different. Which survey gives a better estimate of the percentage of all US adults who would prefer McDonald's over Burger King? **Choose one:**

i) Bonus Survey 4 because it has more people.
ii) Fox News Survey because Fox claims to be "fair and balanced."
iii) The Fox News survey because the people were randomly drawn from all adults nation-wide.

(2 pts) *b)* For which poll is it appropriate to calculate the margin of error? **Choose one:**

i) Only Bonus Survey 4 because their results are more in favor of McDonald's.
ii) Only the Fox News Poll because the people were randomly chosen.
iii) Both polls
iv) Neither poll

Question 7 *(2 points total)*

a) Suppose a government survey organization is planning to take a simple random sample of 900 people in each state in order to estimate the percentage of college graduates in that state. Other things being equal, the accuracy to be expected in New York (population =20 million) is _____ the accuracy to be expected in West Virginia (population = 2 million). Choose one: (i) about the same as ii) quite a bit lower than iii) quite a bit higher than

Question 8 *(9 points total)*

A Harris Poll asked a random sample of 1,015 adults nationwide the following question: "Do you believe in life after love?" 10% of the people in the sample answered "YES". The SE of the sample % is about 1%.

a) *(2 points)* An approximate 95% confidence interval for the percentage of all American adults who believe in life after love is closest to: **Choose one:**

i) (8.7%-11.3%) ii) (7%-13%) (iii) (8%-12%) iv) (93%-97%)

$10 \pm 2(1)$

b) *(2 points)* An approximate 80% confidence interval for the percentage of all American adults who believe in life after love is closest to: **Choose one:** $10 \pm 1.3(1)$

(i) (8.7%-11.3%) ii) (78.7%-81.3%) iii) (8%-12%) iv) (70%-90%)

c) *(2 points)* If the researcher sampled 4 times as many people (4,060) and still got 10% answering "Yes", the SE for the sample % would change from about 1% to about ….. **Choose one:**

i) 0.25% (ii) 0.5% iii) 1% iv) 2% v) 4% vi) not possible to calculate

d) *(2 points)* Suppose 100 pollsters each randomly sampled 1,000 adults nationwide asking whether believe in life after love. All 100 pollsters computed 95% confidence intervals to estimate the percentage of all US adults who believe in life after love. Would all 100 intervals correctly include the true population percent? **Choose one:**

i) Yes, all 100 would assuming no errors were made
(ii) No, only about 95 of them would
iii) No, only about half of them would.

e) *(1 point)* Would the Harris Poll described above have selection bias?

i) Impossible to tell because we aren't the researchers
(ii) No because the people were randomly selected/chosen.
iii) Yes because there are some biases we can't avoid.

Question 9 *(4 points total)*

Suppose you wanted to find out what every person in the nation thinks about some issue, like: Which is better, over or under? (I'm talking about toilet paper here!)

a) *(2 points)* How many people would you have to poll to be 95% sure that your sample percent is right to within 5%. In other words, how many people do you need to poll to get a margin of error of 5%? Assume the SD is 0.45. **(Show work and circle answer)**

if they use 0.05 in denominator -1 $n = \left(\frac{200 \times 0.45}{5}\right)^2 = (324)$

b) *(2 points)* Estimate how many people you'd need to poll to get a 95% confidence interval with only a 3% margin of error. **(Assume the SD of the population is around 0.48. Show work and circle answer)**

if they use 0.03 in denominator -1 $n = \left(\frac{200 \times 0.48}{3}\right)^2 = (1024)$

Question 10: *(4 points total)* Consider the following box with 5 tickets: 0 1 1 0 0

a) *(1 point)* What is the expected value for the percent of 1's in the box? EV% of 1's= 40%.

b) *(1 point)* What is the expected value for the percent of 0's in the box? EV% of 0's= 60%.

c) *(2 points)* What is the SD of this box? **Show work and circle answer.**

$SD = |1-0| \sqrt{\frac{2}{5} \times \frac{3}{5}} = (0.49)$

Question 11 *(9 points total)*
Suppose I wanted to study students' attention span during my 50 minute Stat 100 lectures so I chose a random sample of 64 students out of the 900 students enrolled in class to follow in detail. The average number of minutes the 64 students reported not listening during the 50 minute lecture was 20 minutes with a SD of 10 minutes.

a) *(2 points)* Which most closely resembles the relevant box model? **Circle one:**
i) The box has 900 tickets marked with "1"s and "0"s. The exact percentages are unknown, but can be estimated from our random sample statistics.
ii) The box has 64 tickets marked with "1"s and "0"s
(iii) The box has 900 tickets, marked with numbers , the exact average and SD of the box are unknown and are estimated from the sample.
iv) The box has 900 tickets with an average of 20 and a SD of 10.

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

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

d) *(2 points)* The best estimate for the average number of minutes all 900 students would report not listening is 20.

e) *(2 points)* What is the SE of the sample average? **Show work. Circle answer.**
$SE_{ave} = SD \text{ of box}/\sqrt{n} = 10/\sqrt{64} = (1.25)$

f) *(1 point)* Maybe I just had bad luck and happened to randomly choose 64 of the least attentive students. Maybe the true average of the whole class would report not listening is only 10 minutes with a SD of 10 minutes. If the true average of all 900 students was really 10 minutes with an SD of 10, what is the chance that I'd randomly draw a sample of 64 students with an average as high as 20 or more? **Choose one:**
i) Not enough information to calculate
(ii) Getting a sample average of 20 or more would be nearly impossible since 20 converts to a Z score = 8.
iii) There is about a 2.5% chance of getting an average of 20 or more, since 20 converts to a Z score of 2.
iv) People's attention spans are enormously variable, it wouldn't be at all surprising to get a sample average of 20 or higher if the true average was 10; the Z score is very close to 0, so the chance is very close to 50%.

$Z = \frac{20-10}{1.25} = 8$

Question 12 *(4 points total)*
According to the World Health Organization, 33% of the adults aged 15-49 in Swaziland (a Sub-Saharan African country with a population of 1.1 million) are infected with HIV, the virus that causes AIDs. Assume this percentage is calculated from taking a random sample of 500 adults aged 15-49 in Swaziland.

a) Circle whether each of the statements below is true or false: *(3 points)*
i) The expected value for the percent of HIV infected adults in all of Africa is about 33%.
True (False)
ii) The expected value for the percent of HIV infected sex workers in Swaziland is about 33%.
True (False)
iii) The expected value for the percent of HIV infected adults aged 15-49 in Swaziland is about 33%.
(True) False

b) *(1 point)* Is it possible to compute a 95% confidence interval for the percent of all **adults worldwide** who are infected with HIV virus from the information given?
i) Yes.
ii) No, because we're not given the SD of the sample
(iii) No, because the data from the Swaziland is not a worldwide random sample

Question 13 *(2 points)*
The smallest the SD of a 0-1 box can be is 0 and the largest the SD of a 0-1 box can be is 0.5.

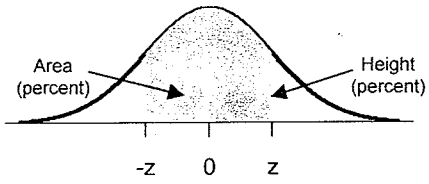
Extra Credit: You get 1 point per question (2 points total- all or nothing!)
1) Write the name any Stat 100 TA (first name is fine): _____
2) How many people would you need to poll to get an 84% confidence interval with a margin of error of only 8%? **Assume that SD=0.5. Show work below and circle your answer. Round to the nearest whole number.**

$n = \left(\frac{100 \times 0.5 \times 1.4}{8}\right)^2 = \boxed{76 \text{ or } 77}$

5.3 5.2

central limit thm.

STANDARD NORMAL TABLE



z	*Height*	*Area*	*z*	*Height*	*Area*	*z*	*Height*	*Area*
0.00	39.89	0.00	1.50	12.95	86.64	3.00	0.443	99.730
0.05	39.84	3.99	1.55	12.00	87.89	3.05	0.381	99.771
0.10	39.70	7.97	1.60	11.09	89.04	3.10	0.327	99.806
0.15	39.45	11.92	1.65	10.23	90.11	3.15	0.279	99.837
0.20	39.10	15.85	1.70	9.40	91.09	3.20	0.238	99.863
0.25	38.67	19.74	1.75	8.63	91.99	3.25	0.203	99.885
0.30	38.14	23.58	1.80	7.90	92.81	3.30	0.172	99.903
0.35	37.52	27.37	1.85	7.21	93.57	3.35	0.146	99.919
0.40	36.83	31.08	1.90	6.56	94.26	3.40	0.123	99.933
0.45	36.05	34.73	1.95	5.96	94.88	3.45	0.104	99.944
0.50	35.21	38.29	2.00	5.40	95.45	3.50	0.087	99.953
0.55	34.29	41.77	2.05	4.88	95.96	3.55	0.073	99.961
0.60	33.32	45.15	2.10	4.40	96.43	3.60	0.061	99.968
0.65	32.30	48.43	2.15	3.96	96.84	3.65	0.051	99.974
0.70	31.23	51.61	2.20	3.55	97.22	3.70	0.042	99.978
0.75	30.11	54.67	2.25	3.17	97.56	3.75	0.035	99.982
0.80	28.97	57.63	2.30	2.83	97.86	3.80	0.029	99.986
0.85	27.80	60.47	2.35	2.52	98.12	3.85	0.024	99.988
0.90	26.61	63.19	2.40	2.24	98.36	3.90	0.020	99.990
0.95	25.41	65.79	2.45	1.98	98.57	3.95	0.016	99.992
1.00	24.20	68.27	2.50	1.75	98.76	4.00	0.013	99.9937
1.05	22.99	70.63	2.55	1.54	98.92	4.05	0.011	99.9949
1.10	21.79	72.87	2.60	1.36	99.07	4.10	0.009	99.9959
1.15	20.59	74.99	2.65	1.19	99.20	4.15	0.007	99.9967
1.20	19.42	76.99	2.70	1.04	99.31	4.20	0.006	99.9973
1.25	18.26	78.87	2.75	0.91	99.40	4.25	0.005	99.9979
1.30	17.14	80.64	2.80	0.79	99.49	4.30	0.004	99.9983
1.35	16.04	82.30	2.85	0.69	99.56	4.35	0.003	99.9986
1.40	14.97	83.85	2.90	0.60	99.63	4.40	0.002	99.9989
1.45	13.94	85.29	2.95	0.51	99.68	4.45	0.002	99.9991