1. A certain number of tickets will be drawn at random, with replacement from the box

$\boxed{\boxed{1}\ \boxed{1}\ \boxed{3}\ \boxed{4}\ \boxed{6}}$

Comment: All the answer below rely on the 'law of averages':

> *As the number of draws from the box grows larger, the chance that the percentage of* $\boxed{?}$ *in the sample is close to the percentage of* $\boxed{?}$ *in the box grows larger too.*

(a) (2 pts) You win \$1000 if the percentage of $\boxed{1}$s drawn is more than 44%. Which gives a better chance of winning, 50 draws or 100 draws? Explain briefly.

The percentage of $\boxed{1}$ *s in the box is* 40%*, so the more you draw from the box, the more likely it is that the percentage of* $\boxed{1}$ *in the sample will be close to* 40%*, which is **less** than* 44%*. If you are betting on **more** than* 44%*, you are more likely to win with* 50 *draws than with* 100 *draws.*

(b) (2 pts) You win \$2000 if the percentage of $\boxed{6}$s is more than 17%. Which gives a better chance of winning, 40 draws or 120 draws? Explain briefly.

The percentage of $\boxed{6}$ *s in the box is* 20%*, so the more you draw from the box, the more likely it is that the percentage of* $\boxed{6}$ *in the sample will be close to* 20%*, which is **more** than* 17%*. If you are betting on **more** than* 17%*, you are more likely to win with* 120 *draws than with* 40 *draws.*

(c) (2 pts) You win \$3000 if more $\boxed{3}$s are drawn than $\boxed{1}$s. Which gives a better chance of winning, 25 draws or 80 draws? Explain briefly.

The percentage of $\boxed{1}$ *s in the box is* 40% *and the percentage of* $\boxed{3}$ *s in the box is* 20%*. The more you draw from the box, the more likely it is that the percentage of* $\boxed{1}$ *s in the sample will be close to* 40% *and the percentage of* $\boxed{3}$ *s will be close to* 20%*. If you are betting on **more** than* $\boxed{3}$ *s than* $\boxed{1}$ *s, then **fewer** draws is better — you are more likely to win with* 25 *draws than with* 80 *draws.*

2. (4 pts) Continuing with the box from the previous problem: 1600 tickets will be drawn at random with replacement. Use the normal approximation to find the (approximate) probability that the ***sum*** of the draws is between 4777 and 4891.

(1) *The average of the box is*

$$Avg = \frac{1+1+3+4+6}{5} = 3$$

and the standard deviation of the box is

$$SD = \sqrt{\frac{(1-3)^2 + (1-3)^2 + (3-3)^2 + (4-3)^2 + (6-3)^2}{5}} \approx 1.897.$$

(2) *The expected sum (of* $n = 1600$ *draws) is*

$$EV(sum) = n \cdot Avg = 1600 \cdot 3 = 4800$$

and the standard error for the sum is

$$SE = \sqrt{n} \cdot SD \approx \sqrt{1600} \cdot 1.897 \approx 75.88.$$

(3) *By the* ***normal approximation****, the probability that the sum of the* 1600 *draws is between* 4777 *and* 4891 *is approximately equal to the area under the normal curve between*

$$\frac{4777 - 4800}{75.88} \approx -0.3 \quad \text{and} \quad \frac{4891 - 4800}{75.88} \approx 1.2,$$

which is

$$\frac{1}{2}T(0.3) + \frac{1}{2}T(1.2) = 50.285\%.$$

3. A simple random sample of 1225 California adults was surveyed — 490 of those surveyed reported watching the wedding of Meghan Markle and Prince Harry and 784 of those surveyed own at least two television sets.

(a) (3 pts) Find a 95%-confidence interval for the percentage of all California adults who watched the wedding.

The sample percentage of wedding-watchers is

$$\frac{490}{1225} \cdot 100\% = 40\%$$

and the standard error for percentage in this sample is

$$SE_{\%} = \frac{\textit{population SD}}{\sqrt{\textit{sample size}}} \times 100\% \approx \frac{\textit{sample SD}}{\sqrt{\textit{sample size}}} \times 100\% = \frac{\sqrt{0.4 \cdot 0.6}}{\sqrt{1225}} \times 100\% \approx 1.4\%$$

$\Longrightarrow$ *A 95%-confidence interval or the population percentage of wedding-watchers is*

$$(\textit{sample percentage} \pm 2SE_{\%}) \approx (40\% \pm 2.8\%).$$

(b) (3 pts) Find a 99.7%-confidence interval for the percentage of all California adults who own at least two television sets.

The sample percentage of two-TV-owners is

$$\frac{784}{1225} \cdot 100\% = 64\%$$

and the standard error for percentage in this sample is

$$SE_{\%} = \frac{\textit{population SD}}{\sqrt{\textit{sample size}}} \times 100\% \approx \frac{\textit{sample SD}}{\sqrt{\textit{sample size}}} \times 100\% = \frac{\sqrt{0.64 \cdot 0.36}}{\sqrt{1225}} \times 100\% \approx 1.37\%$$

$\Longrightarrow$ *A 99.7%-confidence interval or the population percentage of two-TV-owners is*

$$(\textit{sample percentage} \pm 3SE_{\%}) \approx (64\% \pm 4.11\%).$$

4. (4 pts) Big State College has $12,000$ undergraduates, $6,500$ women and $5,500$ men. One student, Joe, surveyed a sample of 400 other students and found that 41% of those surveyed prefer Pepsi to Coke, 73% eat pizza at least once a week and 67% were men.

Joe thinks that his sample is a simple random sample, but his friend Sally disagrees. What do you think? Explain your answer.

Joe's sample is almost certainly ***not*** *a simple random sample because the percentage of men in the sample,* 67%, *is* ***very far*** *from the expected percentage (the population percentage) of men, which is known to be* $5500/12000 \approx 46\%$.

Comment: *The three lines above constitute a complete answer to the question (i.e., full credit). On the other hand, we can justify the 'very far' assessment with more precision, as follows.*

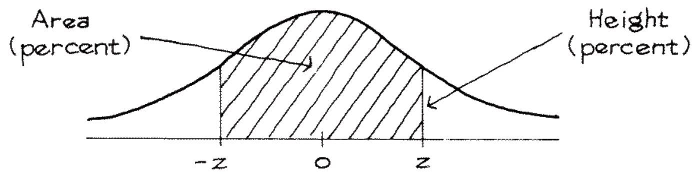
The standard error for percentage in this case is

$$SE_{\%} \approx \frac{\sqrt{0.46 \cdot 0.54}}{\sqrt{400}} \times 100\% \approx 2.5\%,$$

so the observed sample percentage is

$$\frac{67\% - 46\%}{2.5\%} = 8.4$$

standard errors away from expected. The probability of a result like this with a simple random sample is virtually 0.



A NORMAL TABLE

z	*Height*	*Area*	*z*	*Height*	*Area*	*z*	*Height*	*Area*
0.00	39.89	0	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	39.69	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	9.69	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