# DSE 210: Worksheet #9 - Sampling

*Professor: A. Enis Çetin*

*Teaching Assistant: Shivani Agrawal*

**Joshua Wilson** **A53228518**

# Problem 1

For a single draw, $p(\text{red}) = \frac{9}{10}$. We can model this as the tossing of a coin with $p = \frac{9}{10}$, so that the distribution of the observed number of positive outcomes in $n$ tosses is $N(np, np(1-p))$.

The distribution of red marbles in 900 draws with replacement is roughly $N(900 \times \frac{9}{10}, 900 \times \frac{9}{10} \times \frac{1}{10}) = \boxed{N(810, 81)}$.

# Problem 5

(a) If $n = 100$, then the standard deviation will be on the order of $\dfrac{1}{\sqrt{n}} = \dfrac{1}{10}$. Since $p = 0.5$, the specific standard deviation in this case would be roughly $\sqrt{\dfrac{p(1-p)}{n}} = \dfrac{\sqrt{0.5 \times 0.5}}{\sqrt{100}} = \dfrac{0.5}{10} = 0.05$.

(b) If $n = 2500$, then the standard deviation will be on the order of $\dfrac{1}{\sqrt{n}} = \dfrac{1}{50}$. Since $p = 0.5$, the specific standard deviation in this case would be roughly $\sqrt{\dfrac{p(1-p)}{n}} = \dfrac{\sqrt{0.5 \times 0.5}}{\sqrt{2500}} = \dfrac{0.5}{50} = 0.01$.

# Problem 9

Sample 1,000 people, since the demographics in Dallas are the same as in Austin.

# Problem 11

(a) We can model the sum as the sum of 100 independent draws. If $X_i$ is the distribution of numbers in the box, then the sum of 100 draws has the distribution $N(100 \times \hat{\mu_x}, 100 \times \hat{\sigma_x}^2)$. Since the sum of the 100 draws is 297, $\hat{\mu_x} = \boxed{\dfrac{297}{100}}$.

(b) No; in order to provide a confidence interval we would need to know the sample standard deviation.