Recitation Session 5 Solutions

1. *CLT and LLN.*

   CLT says,

   - Distribution of the total number of successes looks like a normal distribution centered at np, whose standard deviation is a constant multiple of $\sqrt{n}$.
   - Distribution of the proportion of successes looks like a normal distribution centered at p, whose standard deviation is a constant multiple of $1/\sqrt{n}$.

   LLN says,

   - The proportion of successes is close to p.

2. *Random walk.* Let X_i be the ith step. $\mathrm{E}(X_i) = (-1)\frac{1}{3} + (0)\frac{1}{3} + (+1)\frac{1}{3} = 0$, and $\mathrm{Var}(X_i) = \mathrm{E}(X^2) - \mathrm{E}(X)^2 = (-1)^2\frac{1}{3} + (0)^2\frac{1}{3} + (+1)^2\frac{1}{3} = 2/3$ hence its standard deviation is 0.8165. The position at the 10,000th step is given by $S_{10{,}000} = \sum_{i=1}^{10{,}000} X_i$. By the normal approximation $S_{10{,}000}$ is approximately normal with mean zero and standard deviation 81.65. Then $\mathrm{P}(S_{10{,}000} > 100) = \mathrm{P}(S_{10{,}000}/81.65 > 100/81.65) = Q(100/81.65) \approx 11\%$.

3. *Biased die.* Since we count number of sixes let X_i be 1 if the ith roll is a six and 0 otherwise. Then $\sum X_i = 180.000$ is observed, so the estimator is $\hat{p} = 0.18$, $n = 10^6$, $\sigma = \sqrt{np(1-p)}$. Now plug them into equation 38 in the lecture notes. This interval depends on p, however we can slightly enlarge the interval to a more conservative value so the confidence is not broken but is independent of p. This is done by the bound $p(1-p) < 1/4$. We arrive at $(0.178, 0.182)$.

4. *Biased coin.* Note that $Q^{-1}(0.5) \approx 2.575$ so we have $\frac{\sqrt{p(1-p)}}{\sqrt{n}} 2.575 = 0.1$ gives $n = 25.75^2 p(1-p) < 25.75^2(0.5)^2 = 165.77$ To decrease the length of the interval 10 times, increase the number of samples by 100 times: 16,577.

5. *Two surveys.* The larger sample will have a shorter interval. Since each interval scales with $1/\sqrt{n}$ the shorter interval will be $\sqrt{\frac{350}{1000}} = 0.59$ times the size of the longer one.