Quiz Section #8: Supplementary Exercise Answers

CSE 312: Foundations of Computing II

1. Let $X \sim N(50, 5)$. What is the probability that $X$ is greater than 45 and less than 52? The $\Phi$ table is on the next page.

$$P(45 < X < 52) = P\left(\frac{45 - 50}{\sqrt{5}} < \frac{X - 50}{\sqrt{5}} < \frac{52 - 50}{\sqrt{5}}\right) \approx P\left(-2.24 < \frac{X - 50}{\sqrt{5}} < 0.89\right) = \Phi(0.89) - \Phi(-2.24)$$

$$= \Phi(0.89) - (1 - \Phi(2.24)) = \Phi(0.89) + \Phi(2.24) - 1 \approx 0.8133 + 0.9875 - 1 = 0.8008$$

2. Before putting any bets down on roulette, you watch 100 rounds, each of which results in an integer between 1 and 36. You count how many rounds have a result that is odd and, if the count exceeds 55, you decide the roulette wheel is unfair. Assuming the roulette wheel is fair, approximate the probability that you make the wrong decision.

Let $X$ be the number of rounds whose result is odd. If the roulette wheel is fair, then $X \sim \text{Bin}(100, 0.5)$. $E[X] = 100 \times 0.5 = 50$ and $\text{Var}(X) = 100 \times 0.5 \times (1 - 0.5) = 25$. You will decide the roulette wheel is unfair if and only if $X > 55$.

$$P(X > 55) = P(X > 55.5) = P\left(\frac{X - 50}{5} > \frac{55.5 - 50}{5}\right) \approx 1 - \Phi\left(\frac{55.5 - 50}{5}\right) = 1 - \Phi(1.1) \approx 1 - 0.8643 = 0.1357.$$  

3. A factory produces $X_i$ gadgets on day $i$, where the $X_i$ are independent and identically distributed random variables, each with mean 5 and variance 9.

Let $Y_n = X_1 + X_2 + \cdots + X_n$. By the Central Limit Theorem, if $n$ is large enough then $Y_n$ is distributed approximately as $N(5n, 9n)$. (Prove that these are the correct parameters for the normal distribution.)

(a) Approximate the probability that the total number of gadgets produced in 100 days is less than 440.

$$P(Y_{100} < 440) = P(Y_{100} < 439.5) = P\left(\frac{Y_{100} - 500}{30} < \frac{439.5 - 500}{30}\right) \approx \Phi\left(\frac{439.5 - 500}{30}\right)$$

$$= \Phi(-2.02) = 1 - \Phi(2.02) \approx 1 - 0.9783 = 0.0217$$

(b) Approximate the greatest value of $n$ such that $P(X_1 + X_2 + \cdots + X_n \geq 5n + 200) \leq 0.05$.

$$P(Y_n \geq 5n + 200) = P\left(Y_n \geq 5n + 199.5\right) = P\left(\frac{Y_n - 5n}{3\sqrt{n}} \geq \frac{199.5}{3\sqrt{n}}\right) = 1 - \Phi\left(\frac{199.5}{3\sqrt{n}}\right) \leq 0.05$$

Therefore, $\Phi\left(\frac{199.5}{3\sqrt{n}}\right) \geq 0.95$. What value of $z$ has $\Phi(z) = 0.95$ in the standard normal table? It is between $z = 1.64$ and $z = 1.65$. Choosing $z = 1.645$, then, we need to solve the equation $\frac{199.5}{3\sqrt{n}} = 1.645$. The solution is $n \approx 1634$. 
4. (a) A fair coin is tossed 50 times. Use the Central Limit Theorem to estimate the probability that fewer than 20 of those tosses come up heads.

Let \( X \sim \text{Bin}(50, 0.5) \). Then \( E[X] = 50 \times 0.5 = 25 \) and \( \text{Var}(X) = 50 \times 0.5 \times 0.5 = 12.5 \).

\[
P(X < 20) = P(X < 19.5) = P\left(\frac{X - 25}{12.5} < \frac{19.5 - 25}{12.5}\right) \approx P\left(\frac{X - 25}{12.5} < -1.56\right) \approx \Phi(-1.56)
\]

\[
= 1 - \Phi(1.56) = 0.0594
\]

(b) A fair coin is tossed until it comes up heads for the 20th time. Use the Central Limit Theorem to estimate the probability that more than 50 tosses are needed. (Hint: you will need the mean and variance of a geometric random variable, which you can find in Example 2.15 of the text.)

Let \( X_1, X_2, \ldots, X_{20} \sim \text{geo}(0.5) \) be independent and let \( X = \sum_{i=1}^{20} X_i \). Then \( E[X] = \sum_{i=1}^{20} E[X_i] = 40 \) and \( \text{Var}(X) = \sum_{i=1}^{20} \text{Var}(X_i) = \sum_{i=1}^{20} \frac{1-0.5}{(0.5)^2} = 40 \).

\[
P(X > 50) = P(X > 50.5) = P\left(\frac{X - 40}{\sqrt{40}} > \frac{50.5 - 40}{\sqrt{40}}\right) \approx P\left(\frac{X - 40}{\sqrt{40}} > 1.66\right) \approx 1 - \Phi(1.66) \approx 0.0485
\]

(c) Compare your answers from parts (a) and (b). Why are they close but not exactly equal?

Suppose the coin is tossed until it has been tossed at least 50 times and heads has come up at least 20 times. If the first 50 tosses have fewer than 20 heads, then the 20th head requires more than 50 tosses. Conversely, if the 20th head requires more than 50 tosses, then the first 50 tosses have fewer than 20 heads. Thus, the exact probabilities in parts (a) and (b) must be identical. The reason we did not get the exact same answers in parts (a) and (b) is that the Central Limit Theorem says that the Normal distribution is only an approximation (except in the limit).

5. Suppose 59 percent of voters favor Proposition 666. Use the Normal approximation to estimate the probability that a random sample of 100 voters will contain:

Let \( X \) be the number of voters in the random sample that favor Proposition 666. Then \( X \sim \text{Bin}(100, 0.59) \), with \( E[X] = 59 \) and \( \text{Var}(X) = 24.19 \).

(a) at most 50 in favor.

\[
P(X \leq 50) = P(X < 50.5) = P\left(\frac{X - 59}{24.19} < \frac{50.5 - 59}{24.19}\right) \approx P\left(\frac{X - 59}{24.19} < -1.73\right) \approx 1 - \Phi(1.73) \approx 0.0418
\]

(b) between 54 and 64 (inclusive) in favor.

\[
P(54 \leq X \leq 64) = P(53.5 < X < 64.5) = P\left(\frac{53.5 - 59}{24.19} < \frac{X - 59}{24.19} < \frac{64.5 - 59}{24.19}\right)
\]

\[
\approx P\left(-1.12 < \frac{X - 59}{24.19} < 1.12\right) \approx 2\Phi(1.12) - 1 \approx 0.7372
\]
(c) fewer than 72 in favor.

\[ P(X < 72) = P(X < 71.5) = P\left( \frac{X - 59}{\sqrt{24.19}} < \frac{71.5 - 59}{\sqrt{24.19}} \right) \approx P\left( \frac{X - 59}{\sqrt{24.19}} < 2.54 \right) \approx \Phi(2.54) \approx 0.9945 \]

6. Each day, the probability your computer crashes is 10%, independent of every other day. Approximate the probability of at least 87 crash-free days out of the next 100 days.

Let \( X \) be the number of crash-free days in the next 100 days. Then \( X \sim \text{Bin}(100, 0.9) \), with \( E[X] = 90 \) and \( \text{Var}(X) = 9 \).

\[ P(X \geq 87) = P(86.5 < X < 100.5) = P\left( \frac{86.5 - 90}{3} < \frac{X - 90}{3} < \frac{100.5 - 90}{3} \right) \]
\[ \approx P\left( -1.17 < \frac{X - 90}{3} < 3.5 \right) \approx \Phi(3.5) + \Phi(1.17) - 1 \approx 0.9998 + 0.8790 - 1 = 0.8788 \]

Notice that, if you had used \( 86.5 < X \) in place of \( 86.5 < X < 100.5 \), your answer would have been nearly the same, because \( \Phi(3.5) \) is so close to 1.

7. Let \( X \sim \text{Exp}(\lambda) \) and \( k > 1/\lambda \).

(a) Use Markov’s inequality to bound \( P(X \geq k) \).

\[ P(X \geq k) \leq \frac{1}{\lambda k} \]

(b) Use Chebyshev’s inequality to bound \( P(X \geq k) \).

\[ P(X \geq k) = P\left( X - \frac{1}{\lambda} \geq k - \frac{1}{\lambda} \right) \leq P\left( \left| X - \frac{1}{\lambda} \right| \geq k - \frac{1}{\lambda} \right) \leq \frac{1}{\lambda^2(k - 1/\lambda)^2} = \frac{1}{(\lambda k - 1)^2} \]

(c) What is the exact formula for \( P(X \geq k) \)?

\[ P(X \geq k) = e^{-\lambda k} \]

(d) For \( \lambda k \geq 3 \), how do the bounds given in parts (a), (b), and (c) compare?

\[ e^{-\lambda k} < \frac{1}{(\lambda k - 1)^2} < \frac{1}{\lambda k} \]

so Markov’s inequality gives the worst bound.