Section 8, Supplementary Exercise 2: Detailed solution

Question: *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.*

Solution: Since we're assuming that the roulette wheel is fair, the only way for you to make the wrong decision is if the number of odd results exceeds 55, which causes you to predict that the wheel is unfair (even when it is actually fair). We want to find the probability that this happens.

Let X_i be an indicator variable that is set to 1 when the i-th round of roulette generates an odd result (and 0 otherwise). Then, the total number of odd results over all 100 rounds is just the sum of all of the X_i's for each of the individual rounds. Since this is the result we're interested in, let random variable Y be $X_1 + X_2 + \cdots + X_{100}$. Y represents the number of odd results among the 100 rounds.

If the roulette wheel is fair, then the X_i's are independent. Since the X_i's are independent and identically distributed, by the Central Limit Theorem, the sum of all of these X_i's can be approximated by a normal distribution. What is the mean and variance of this normal distribution? Well, we should set the mean and variance of the normal so that it is the **same mean and variance as the sum** we are approximating.

Let's calculate the mean and variance of the sum. There are two ways to do this:

Approach 1: Linearity of expectation/variance

Notice that each X_i is Bernoulli(0.5). So, $E[X_i] = p = 0.5$ and $Var(X_i) = p(1-p) = 0.25$. Using linearity of expectation,

$$E[X_1 + X_2 + \cdots + X_{100}] = E[X_1] + E[X_2] + \cdots + E[X_{100}] = 100 \cdot 0.5 = 50$$

Since the variables are independent, we can also use linearity of variance.

$$Var(X_1 + X_2 + \cdots + X_{100}) = Var(X_1) + Var(X_2) + \cdots + Var(100) = 100 \cdot 0.25 = 25$$

Thus, the mean of the sum is 50, and the variance of the sum is 25.

Approach 2: Normal approximation of binomial

Notice that $Y = X_1 + X_2 + \cdots + X_{100}$ is just the sum of 100 independent Bernoulli random variables, so $Y \sim Bin(100, 0.5)$. (This is because Y represents the number of successes (i.e. odd results) among the 100 rounds, and the probability of each round being a success is $\frac{1}{2}$.)

We know the mean and variance of a Binomial random variable is

$$E[Y] = np = 100 \cdot \frac{1}{2} = 50$$

$$Var(Y) = np(1-p) = 100 \cdot \frac{1}{2} \cdot (1 - \frac{1}{2}) = 25$$

Thus, using both approaches, we find that the expected value of the sum is 50, and the variance of the sum is 25. Thus, by the Central Limit Theorem, Y can be approximated by $N(50, 25)$.

Standardizing the normal

Now, we know that Y is approximated by $N(50, 25)$. We want to find $P(Y > 55)$. After applying the continuity correction (since we're using the normal distribution to approximate a discrete distribution, we must round the value returned by the normal to the nearest integer), we are now looking for $P(Y > 55.5)$. To find this, we must find some way to relate Y to the standard normal distribution with mean 0 and variance 1.

Since Y is approimately N(50, 25), if we standardize Y, we find that $\frac{Y-50}{\sqrt{25}}$ is approximately N(0, 1).

Therefore, since we want to find $P(Y > 55.5)$, subtract the mean (50) from both sides of the inequality, and then divide both sides by the standard deviation (5).

$$\begin{aligned} P(Y > 55.5) &= P(\frac{Y-50}{\sqrt{25}} > \frac{55.5-50}{\sqrt{25}}) \\ &= P(\frac{Y-50}{5} > 1.1) \end{aligned}$$

Since $\frac{Y-50}{5}$ is approximated by the standard normal, the probability we're looking for is just the probability that the standard normal is greater than 1.1, which is simply $\boxed{1 - \Phi(1.1)}$. (Recall that $\Phi(1.1)$ is the probability that the standard normal is less than 1.1, so to find the probability that it is greater, we must take 1 minus that.) Now you can look up $\Phi(1.1) = 0.8643$ in the standard normal distribution table to finish the solution.