CHAPTER 5

Solution to Problem 5.1. (a) We have $\sigma_{M_n} = 1/\sqrt{n}$, so in order that $\sigma_{M_n} \leq 0.01$, we must have $n \geq 10,000$.

(b) We want to have

$$\mathbf{P}\Big(|M_n - h| \leq 0.05\Big) \geq 0.99.$$

Using the facts $h = \mathbf{E}[M_n]$, $\sigma^2_{M_n} = 1/n$, and the Chebyshev inequality, we have

$$\begin{aligned}\mathbf{P}\Big(|M_n - h| \leq 0.05\Big) &= \mathbf{P}\Big(|M_n - \mathbf{E}[M_n]| \leq 0.05\Big)\\ &= 1 - \mathbf{P}\Big(|M_n - \mathbf{E}[M_n]| \geq 0.05\Big)\\ &\geq 1 - \frac{1/n}{(0.05)^2}.\end{aligned}$$

Thus, we must have

$$1 - \frac{1/n}{(0.05)^2} \geq 0.99,$$

which yields $n \geq 40,000$.

(c) Based on Example 5.3, $\sigma^2_{X_i} \leq (0.6)^2/4$, so he should use 0.3 meters in place of 1.0 meters as the estimate of the standard deviation of the samples X_i in the calculations of parts (a) and (b). In the case of part (a), we have $\sigma_{M_n} = 0.3/\sqrt{n}$, so in order that $\sigma_{M_n} \leq 0.01$, we must have $n \geq 900$. In the case of part (b), we have $\sigma_{M_n} = 0.3/\sqrt{n}$, so in order that $\sigma_{M_n} \leq 0.01$, we must have $n \geq 900$. In the case of part (a), we must have

$$1 - \frac{0.09/n}{(0.05)^2} \geq 0.99,$$

which yields $n \geq 3,600$.

Solution to Problem 5.4. Proceeding as in Example 5.5, the best guarantee that can be obtained from the Chebyshev inequality is

$$\mathbf{P}\Big(|M_n - f| \geq \epsilon\Big) \leq \frac{1}{4n\epsilon^2}.$$

(a) If ϵ is reduced to half its original value, and in order to keep the bound $1/(4n\epsilon^2)$ constant, the sample size n must be made four times larger.

(b) If the error probability δ is to be reduced to $\delta/2$, while keeping ϵ the same, the sample size has to be doubled.

Solution to Problem 5.5. In cases (a), (b), and (c), we show that Y_n converges to 0 in probability. In case (d), we show that Y_n converges to 1 in probability.

(a) For any $\epsilon > 0$, we have

$$\mathbf{P}\Big(|Y_n| \geq \epsilon\Big) = 0,$$

for all n with $1/n < \epsilon$, so $\mathbf{P}\big(|Y_n| \geq \epsilon\big) \to 0$.

(b) For all $\epsilon \in (0,1)$, we have

$$\mathbf{P}\big(|Y_n| \geq \epsilon\big) = \mathbf{P}\big(|X_n|^n \geq \epsilon\big) = \mathbf{P}\big(X_n \geq \epsilon^{1/n}\big) + \mathbf{P}\big(X_n \leq -\epsilon^{1/n}\big) = 1 - \epsilon^{1/n},$$

and the two terms in the right-hand side converge to 0, since $\epsilon^{1/n} \to 1$.

(c) Since $X_1, X_2, \ldots$ are independent random variables, we have

$$\mathbf{E}[Y_n] = \mathbf{E}[X_1] \cdots \mathbf{E}[X_n] = 0.$$

Also

$$\text{var}(Y_n) = \mathbf{E}[Y_n^2] = \mathbf{E}[X_1^2] \cdots \mathbf{E}[X_n^2] = \text{var}(X_1)^n = \left(\frac{4}{12}\right)^n,$$

so $\text{var}(Y_n) \to 0$. Since all Y_n have 0 as a common mean, from Chebyshev's inequality it follows that Y_n converges to 0 in probability.

(d) We have for all $\epsilon \in (0,1)$, using the independence of $X_1, X_2, \ldots$,

$$\begin{aligned}
\mathbf{P}\big(|Y_n - 1| \geq \epsilon\big) &= \mathbf{P}\big(\max\{X_1, \ldots, X_n\} \geq 1 + \epsilon\big) + \mathbf{P}\big(\max\{X_1, \ldots, X_n\} \leq 1 - \epsilon\big) \\
&= \mathbf{P}(X_1 \leq 1 - \epsilon, \ldots, X_n \leq 1 - \epsilon) \\
&= \big(\mathbf{P}(X_1 \leq 1 - \epsilon)\big)^n \\
&= \left(1 - \frac{\epsilon}{2}\right)^n.
\end{aligned}$$

Hence $\mathbf{P}\big(|Y_n - 1| \geq \epsilon\big) \to 0$.

Solution to Problem 5.8. Let S be the number of times that the result was odd, which is a binomial random variable, with parameters $n = 100$ and $p = 0.5$, so that $\mathbf{E}[X] = 100 \cdot 0.5 = 50$ and $\sigma_S = \sqrt{100 \cdot 0.5 \cdot 0.5} = \sqrt{25} = 5$. Using the normal approximation to the binomial, we find

$$\mathbf{P}(S > 55) = \mathbf{P}\left(\frac{S - 50}{5} > \frac{55 - 50}{5}\right) \approx 1 - \Phi(1) = 1 - 0.8413 = 0.1587.$$

A better approximation can be obtained by using the de Moivre-Laplace approximation, which yields

$$\begin{aligned}
\mathbf{P}(S > 55) = \mathbf{P}(S \geq 55.5) &= \mathbf{P}\left(\frac{S - 50}{5} > \frac{55.5 - 50}{5}\right) \\
&\approx 1 - \Phi(1.1) = 1 - 0.8643 = 0.1357.
\end{aligned}$$

Solution to Problem 5.9. (a) Let S be the number of crash-free days, which is a binomial random variable with parameters $n = 50$ and $p = 0.95$, so that $\mathbf{E}[X] = 50 \cdot 0.95 = 47.5$ and $\sigma_S = \sqrt{50 \cdot 0.95 \cdot 0.05} = 1.54$. Using the normal approximation to the binomial, we find

$$\mathbf{P}(S \geq 45) = \mathbf{P}\left(\frac{S - 47.5}{1.54} \geq \frac{45 - 47.5}{1.54}\right) \approx 1 - \Phi(-1.62) = \Phi(1.62) = 0.9474.$$

A better approximation can be obtained by using the de Moivre-Laplace approximation, which yields

$$\mathbf{P}(S \geq 45) = \mathbf{P}(S > 44.5) = \mathbf{P}\left(\frac{S - 47.5}{1.54} \geq \frac{44.5 - 47.5}{1.54}\right)$$
$$\approx 1 - \Phi(-1.95) = \Phi(1.95) = 0.9744.$$

(b) The random variable S is binomial with parameter $p = 0.95$. However, the random variable $50 - S$ (the number of crashes) is also binomial with parameter $p = 0.05$. Since the Poisson approximation is exact in the limit of small p and large n, it will give more accurate results if applied to $50 - S$. We will therefore approximate $50 - S$ by a Poisson random variable with parameter $\lambda = 50 \cdot 0.05 = 2.5$. Thus,

$$\begin{aligned}
\mathbf{P}(S \geq 45) &= \mathbf{P}(50 - S \leq 5) \\
&= \sum_{k=0}^{5} \mathbf{P}(n - S = k) \\
&= \sum_{k=0}^{5} e^{-\lambda} \frac{\lambda^k}{k!} \\
&= 0.958.
\end{aligned}$$

It is instructive to compare with the exact probability which is

$$\sum_{k=0}^{5} \binom{50}{k} 0.05^k \cdot 0.95^{50-k} = 0.962.$$

Thus, the Poisson approximation is closer. This is consistent with the intuition that the normal approximation to the binomial works well when p is close to 0.5 or n is very large, which is not the case here. On the other hand, the calculations based on the normal approximation are generally less tedious.

Solution to Problem 5.10. (a) Let $S_n = X_1 + \cdots + X_n$ be the total number of gadgets produced in n days. Note that the mean, variance, and standard deviation of S_n is $5n$, $9n$, and $3\sqrt{n}$, respectively. Thus,

$$\begin{aligned}
\mathbf{P}(S_{100} < 440) &= \mathbf{P}(S_{100} \leq 439.5) \\
&= \mathbf{P}\left(\frac{S_{100} - 500}{30} < \frac{439.5 - 500}{30}\right) \\
&\approx \Phi(\frac{439.5 - 500}{30}) \\
&= \Phi(-2.02) \\
&= 1 - \Phi(2.02) \\
&= 1 - 0.9783 \\
&= 0.0217.
\end{aligned}$$

(b) The requirement $\mathbf{P}(S_n \geq 200 + 5n) \leq 0.05$ translates to

$$\mathbf{P}\left(\frac{S_n - 5n}{3\sqrt{n}} \geq \frac{200}{3\sqrt{n}}\right) \leq 0.05,$$

or, using a normal approximation,

$$1 - \Phi\left(\frac{200}{3\sqrt{n}}\right) \leq 0.05,$$

and

$$\Phi\left(\frac{200}{3\sqrt{n}}\right) \geq 0.95.$$

From the normal tables, we obtain $\Phi(1.65) \approx 0.95$, and therefore,

$$\frac{200}{3\sqrt{n}} \geq 1.65,$$

which finally yields $n \leq 1632$.

(c) The event $N \geq 220$ (it takes at least 220 days to exceed 1000 gadgets) is the same as the event $S_{219} \leq 1000$ (no more than 1000 gadgets produced in the first 219 days). Thus,

$$\begin{aligned}
\mathbf{P}(N \geq 220) &= \mathbf{P}(S_{219} \leq 1000) \\
&= \mathbf{P}\left(\frac{S_{219} - 5 \cdot 219}{3\sqrt{219}} \leq \frac{1000 - 5 \cdot 219}{3\sqrt{219}}\right) \\
&= 1 - \Phi(2.14) \\
&= 1 - 0.9838 \\
&= 0.0162.
\end{aligned}$$

Solution to Problem 5.11. Note that W is the sample mean of 16 independent identically distributed random variables of the form $X_i - Y_i$, and a normal approximation is appropriate. The random variables $X_i - Y_i$ have zero mean, and variance equal to $2/12$. Therefore, the mean of W is zero, and its variance is $(2/12)/16 = 1/96$. Thus,

$$\mathbf{P}\big(|W| < 0.001\big) = \mathbf{P}\left(\frac{|W|}{\sqrt{1/96}} < \frac{0.001}{\sqrt{1/96}}\right) \approx \Phi\big(0.001\sqrt{96}\big) - \Phi\big(-0.001\sqrt{96}\big)$$

$$= 2\Phi\big(0.001\sqrt{96}\big) - 1 = 2\Phi(0.0098) - 1 \approx 2 \cdot 0.504 - 1 = 0.008.$$

Let us also point out a somewhat different approach that bypasses the need for the normal table. Let Z be a normal random variable with zero mean and standard deviation equal to $1/\sqrt{96}$. The standard deviation of Z, which is about 0.1, is much larger than 0.001. Thus, within the interval $[-0.001, 0.001]$, the PDF of Z is approximately constant. Using the formula $\mathbf{P}(z - \delta \leq Z \leq z + \delta) \approx f_Z(z) \cdot 2\delta$, with $z = 0$ and $\delta = 0.001$, we obtain

$$\mathbf{P}\big(|W| < 0.001\big) \approx \mathbf{P}(-0.001 \leq Z \leq 0.001) \approx f_Z(0) \cdot 0.002 = \frac{0.002}{\sqrt{2\pi}(1/\sqrt{96})} = 0.0078.$$