1. [Section 11.3, #7.]

Let $X$ denote the waiting period from the time the book is ordered until it is received, in days. Then $E[X] = 7$ and $\text{Var}(X) = 4$. Helen wishes to choose $k$ so that if she orders the book $k$ days before the given date, she gets $P(X \leq k) \geq .95$. We may apply Chebyshev’s inequality as follows:

$$P(X \geq k + 1) = P(X - 7 \geq k + 1 - 7) \leq P(|X - 7| \geq k - 6) \leq \frac{\text{Var}(X)}{(k - 6)^2},$$

or

$$P(X \leq k) = P(X < k + 1) \geq 1 - \frac{\text{Var}(X)}{(k - 6)^2} = 1 - \frac{4}{(k - 6)^2}.$$  

Setting $1 - 4/(k - 6)^2 \geq .95$ we get $(k - 6)^2 \geq 4/.05 = 80$, or $|k - 6| \geq 8.9$. Thus, we can use $k = 15$. Note that in the above we assumed that the waiting period, $X$, was discrete with a unit of 1 day. If we assume that $X$ is continuous then we should apply Chebyshev’s inequality to $P(X \geq k)$ which leads to $k = 16$. Either $k = 15$ or $k = 16$ is acceptable as a final answer.

2. [Section 11.3, #20.]

Note that $X$ has a gamma distribution with parameters $r = n + 1$ and $\lambda = 1$. Hence, the mean and variance of $X$ are given by

$$E[X] = \frac{r}{\lambda} = n + 1 \quad \text{and} \quad \text{Var}(X) = \frac{r}{\lambda^2} = n + 1.$$  

Using Chebyshev’s inequality, we have

$$P(0 < X < 2n + 1) = P(|X - (n + 1)| < n + 1) = 1 - P(|X - (n + 1)| \geq n + 1) \geq 1 - \frac{\text{Var}(X)}{(n + 1)^2} = 1 - \frac{n + 1}{(n + 1)^2} = 1 - \frac{1}{n + 1} = \frac{n}{n + 1}.$$
3. [Section 11.4, #3.]

To show that $Y_n$ converges to 0 in probability we must show that, if $\epsilon > 0$ is given, then $P(|Y_n| > \epsilon) \to 0$ as $n \to \infty$. But, assuming $\epsilon < 1$,

$$P(|Y_n| > \epsilon) = P(Y_n = 1) = P(X > n) = 1 - \int_0^n f(x)dx.$$

But

$$\lim_{n \to \infty} \int_0^n f(x)dx = \int_0^\infty f(x)dx = 1,$$

and so

$$\lim_{n \to \infty} P(|Y_n| > \epsilon) = 1 - 1 = 0,$$

as required.

4. [Let $X_1, X_2, \ldots$ be a sequence of independent...]

(a) By the strong law of large numbers, $\overline{X}_n$ converges to $E[X_1] = 1/\lambda$ almost surely. Letting $g(x) = 1/x$ we have that $g(x)$ is a continuous function on $(0, \infty)$. Therefore, for any $\omega$ such that $\overline{X}_n(\omega) \to 1/\lambda$, we also have $1/\overline{X}_n(\omega) \to \lambda$ and we conclude that $Y_n = 1/\overline{X}_n \to \lambda$ almost surely. Since almost sure convergence implies convergence in probability, we also have that $Y_n$ converges to $\lambda$ in probability.

(b) We need to show that $E[(\overline{X}_n - \mu)^2] \to 0$ as $n \to \infty$. But

$$E[(\overline{X}_n - \mu)^2] = \text{Var}(\overline{X}_n) = \frac{\sigma^2}{n}$$

which clearly goes to 0 as $n \to \infty$. Therefore, $\overline{X}_n$ converges to $\mu$ in mean square.

5. [Let $X_1, X_2, \ldots$ be a sequence of independent...]

Following the hint, we first write

$$\frac{1}{n} \sum_{i=1}^n (X_i - \overline{X}_n)^2 = \frac{1}{n} \left( \sum_{i=1}^n X_i^2 - 2\overline{X}_n \sum_{i=1}^n X_i + \sum_{i=1}^n \overline{X}_n^2 \right)$$

$$= \frac{1}{n} \sum_{i=1}^n X_i^2 - 2\overline{X}_n^2 + \overline{X}_n^2$$

$$= \frac{1}{n} \sum_{i=1}^n X_i^2 - \overline{X}_n^2.$$
The sequence $X_1^2, X_2^2, X_3^2, \ldots$ is a sequence of independent and identically distributed random variables. By the strong law of large numbers

$$\frac{1}{n} \sum_{i=1}^{n} X_i^2 \to E[X_1^2] \quad \text{almost surely as } n \to \infty,$$

where

$$E[X_1^2] = \text{Var}(X_1) + E[X_1]^2 = \sigma^2 + \mu^2.$$

Again by the strong law of large numbers, $\bar{X}_n \to \mu$ almost surely as $n \to \infty$. Since the function $g(x) = x^2$ is continuous, for all $\omega$ such that $\bar{X}_n(\omega) \to \mu$, we also have $(\bar{X}_n(\omega))^2 \to \mu^2$, so

$$\bar{X}_n^2 \to \mu^2 \quad \text{almost surely as } n \to \infty.$$

Finally, let $A = \{ \omega : \frac{1}{n} \sum_{i=1}^{n} (X_i(\omega))^2 \to E[X_1^2] \}$ and $B = \{ \omega : (\bar{X}_n(\omega))^2 \to \mu^2 \}$. Then $P(A) = P(B) = 1$; therefore $P(A \cap B) = 1$. But for all $\omega \in A \cap B$, we have

$$\frac{1}{n} \sum_{i=1}^{n} (X_i(\omega))^2 - (\bar{X}_n)^2 \to (\sigma^2 + \mu^2) = \sigma^2.$$

Thus

$$\frac{1}{n} \sum_{i=1}^{n} X_i^2 - \bar{X}_n^2 \to \sigma^2 \quad \text{almost surely as } n \to \infty.$$

6. [Section 11.5, #11.]

The probability of obtaining at least 25 heads before 50 tails is the probability that there are at least 25 heads in the first 74 flips. If we let $X_i = 1$ if the $i$th flip is a heads and $X_i = 0$ if the $i$th flip is a tails, then $X = X_1 + \ldots + X_{74}$ is the number of heads in the first 74 flips. The mean of each $X_i$ is 1/2 and the variance of each $X_i$ is 1/4, and the $X_i$’s are independent (assuming independent flips). By the central limit theorem, $(X - (74)(1/2))/\sqrt{(74)(1/4)} = (X - 37)/\sqrt{18.5}$ has approximately a $N(0, 1)$ distribution. Since we wish to compute $P(X \geq 25)$, we have

$$P(X \geq 25) = P \left( \frac{X - 37}{\sqrt{18.5}} \geq \frac{25 - 37}{\sqrt{18.5}} \right) \approx P(Z \geq -2.79) = \Phi(2.79) = 0.9974.$$

An alternative solution is to let $Y_1$ be the number of flips until the first tails, $Y_2$ the number of additional flips until the 2nd tails, and in general $Y_k$ the number of additional flips since the $(k-1)$st tail to get the $k$th tail. Then the $Y_i$’s are independent Geometric(1/2) random variables and the probability of obtaining at least 25 heads before 50 tails is the probability that the number of flips required to get 50 tails is at least 75: $P(Y_1 + \ldots + Y_{50} \geq 75)$. 

3
Each $Y_i$ has mean 2 and variance $(1/2)/(1 - 1/2)^2 = 2$. Then, by the central limit theorem,

$$
\frac{Y_1 + \ldots + Y_{50} - (50)(2)}{\sqrt{(50)(2)}} = \frac{Y_1 + \ldots + Y_{50} - 100}{10}
$$

has approximately a $N(0, 1)$ distribution. So

$$
P(Y_1 + \ldots + Y_{50} \geq 75) = P\left(\frac{Y_1 + \ldots + Y_{50} - 100}{10} \geq \frac{75 - 100}{10}\right)
\approx P(Z \geq -2.5) = \Phi(2.5) = 0.9938.
$$

Note that this answer differs slightly from the first solution. Either solution is acceptable. This goes to show that expressing the same probability in terms of two different sets of random variables can lead to different answers when applying the central limit theorem, depending on how well the normal distribution approximates the sum of the random variables in the first set compared to how well it approximates the sum of the random variables in the second set. We can expect that our first solution is more accurate because it involves the sum of more random variables (74 versus 50) and each of the random variables in the first solution is more symmetric to begin with (Bernoulli(1/2) versus Geometric(1/2)).