Problem 7.1.5
Since each $X_i$ has zero mean, the mean of $Y_n$ is

$$E[Y_n] = E[X_n + X_{n-1} + X_{n-2}]/3 = 0$$

Since $Y_n$ has zero mean, the variance of $Y_n$ is

$$\text{Var}[Y_n] = E[Y_n^2] = E[(X_n + X_{n-1} + X_{n-2})^2]/9$$

$$= E[X_n^2 + X_{n-1}^2 + X_{n-2}^2 + 2X_nX_{n-1} + 2X_nX_{n-2} + 2X_{n-1}X_{n-2}]/9$$

$$= (1 + 1 + 2/4 + 0 + 2/4)/9 = 4/9$$

Problem 7.2.2

$$f_{X,Y}(x,y) = \begin{cases} 
1 & 0 \leq x, y \leq 1 \\
0 & \text{otherwise}
\end{cases}$$

Proceeding as in Problem 7.2.1, we must first find $F_W(w)$ by integrating over the square defined by $0 \leq x, y \leq 1$. Again we are forced to find $F_W(w)$ in parts as we did in Problem 7.2.1 resulting in the following integrals for their appropriate regions. For $0 \leq w \leq 1$,

$$F_W(w) = \int_0^w \int_0^{w-x} dx dy = w^2/2$$

For $1 \leq w \leq 2$,

$$F_W(w) = \int_0^{w-1} \int_0^1 dx dy + \int_1^1 \int_{w-1}^{w-y} dx dy = 2w - 1 - w^2/2$$

The complete expression for the CDF of $W$ is

$$F_W(w) = \begin{cases} 
0 & w < 0 \\
w^2/2 & 0 \leq w \leq 1 \\
2w - 1 - w^2/2 & 1 \leq w \leq 2 \\
1 & \text{otherwise}
\end{cases}$$

With the CDF, we can find $f_W(w)$ by differentiating with respect to $w$.

$$f_W(w) = \begin{cases} 
w & 0 \leq w \leq 1 \\
2 - w & 1 \leq w \leq 2 \\
0 & \text{otherwise}
\end{cases}$$
Problem 7.6.3
straightforward random sum of random variables because $N$ and the $X_i$’s are dependent. In particular, given $N = n$, then we know that there were exactly 100 heads in $N$ flips. Hence, given $N$, $X_1 + \cdots + X_N = 100$ no matter what is the actual value of $N$. Hence $Y = 100$ every time and the PMF of $Y$ is

$$P_Y(y) = \begin{cases} 1 & y = 100 \\ 0 & \text{otherwise} \end{cases}$$

Problem 7.7.1
We know that the waiting time, $W$ is uniformly distributed on $[0,10]$ and therefore has the following PDF.

$$f_W(w) = \begin{cases} 1/10 & 0 \leq w \leq 10 \\ 0 & \text{otherwise} \end{cases}$$

We also know that the total time is 3 milliseconds plus the waiting time, that is $X = W + 3$.


(b) The variance of $X$ is $\text{Var}[X] = \text{Var}[W + 3] = \text{Var}[W] = 25/3$.

(c) The expected value of $A$ is $E[A] = 12E[X] = 96$.

(d) The standard deviation of $A$ is $\sigma_A = \sqrt{\text{Var}[A]} = \sqrt{12(25/3)} = 10$.

(e) $P[A > 116] = 1 - \Phi\left(\frac{116 - 96}{10}\right) = 1 - \Phi(2) = 0.02275$.

(f) $P[A < 86] = \Phi\left(\frac{86 - 96}{10}\right) = \Phi(-1) = 1 - \Phi(1) = 0.1587$

Problem 8.3.1

$$P_X(x) = \begin{cases} 0.1 & x = 0 \\ 0.9 & x = 1 \\ 0 & \text{otherwise} \end{cases}$$

(a) $E[X]$ is in fact the same as $P_X(1)$ because $X$ is a Bernoulli random variable.

(b) We can use the Chebyshev inequality to find

$$P[|M_{90}(X) - P_X(1)| \geq .05] = P[|M_{90}(X) - E[X]| \geq .05] \leq \alpha$$

In particular, the Chebyshev inequality states that

$$\alpha = \frac{\sigma^2_X}{90(.05)^2} = \frac{.09}{90(.05)^2} = 0.4$$

(c) Now we wish to find the value of $n$ such that $P[|M_{n}(X) - P_X(1)| \geq .03] \leq .01$. From the Chebyshev inequality, we write $0.01 = \sigma^2_X/[n(.03)^2]$. Solving for $n$ yields $n = 1000$. 
Problem 8.3.2

$X_1, X_2, \ldots$ are iid random variables each with mean 75 and standard deviation 15.

(a) We would like to find the value of $n$ such that

$$P[74 \leq M_n(X) \leq 76] = 0.99$$

When we know only the mean and variance of $X_i$, our only real tool is the Chebyshev inequality which says that

$$P[74 \leq M_n(X) \leq 76] = 1 - P[|M_n(X) - E[X]| \geq 1] \geq 1 - \frac{\text{Var}[X]}{n} = 1 - \frac{225}{n} \geq 0.99$$

This yields $n \geq 22500$.

(b) If each $X_i$ is a Gaussian, the sample mean, $M_n(X)$ will also be Gaussian with mean and variance

$$E[M_n(X)] = E[X] = 75$$
$$\text{Var}[M_n(X)] = \text{Var}[X]/n = 225/n$$

In this case,

$$P[74 \leq M_n(X) \leq 76] = \Phi\left(\frac{76-\mu}{\sigma}\right) - \Phi\left(\frac{74-\mu}{\sigma}\right) = \Phi(\sqrt{n}/15) - \Phi(-\sqrt{n}/15) = 2\Phi(\sqrt{n}/15) - 1 = 0.99$$

Thus, $n' \approx 1521$.

Since even under the Gaussian assumption, the number of samples $n'$ is so large that even if the $X_i$ are not Gaussian, the sample mean may be approximated by a Gaussian. Hence, about 1500 samples probably is about right. However, in the absence of any information about the PDF of $X_i$ beyond the mean and variance, we cannot make any guarantees stronger than that given by the Chebyshev inequality.

Problem 8.4.1

The solution to this problem parallels the solution of Example 8.13. First we find the CDF of $V_n = \min\{U_1, \ldots, U_n\}$.

$$P[V_n \leq v] = 1 - P[V_n > v] = 1 - P[U_1 > v, U_2 > v, \ldots, U_n > v]$$

Since $U_1, U_2, \ldots$ is an iid sequence,

$$P[U_1 > y, U_2 > y, \ldots, U_n > y] = P[U_1 > y] \cdots P[U_n > y] = (1 - F_U(v))^n$$

The CDF of each $U_n$ is

$$F_U(u) = \begin{cases} 
0 & u < 0 \\
(u-a)/(b-a) & a \leq u < b \\
1 & u \geq b
\end{cases}$$
This implies that for \( a \leq u \leq b \),

\[
P[V_n \leq v] = 1 - [1 - F_U(v)]^n = 1 - [(b - v)/(b - a)]^n
\]

Now we use Theorem 8.7 to prove w.p.1 convergence. Since each \( V_i \) is nonnegative,

\[
S_n(\varepsilon) = \{|V_i - a| < \varepsilon\} = \{V_i < a + \varepsilon\}
\]

Since \( V_n = \min\{U_1, \ldots, U_n\} \), we observe that \( V_n = \min\{U_n, V_{n-1}\} \). Hence \( V_n \leq V_{n-1} \) for all \( n \) so that if \( V_n \leq a + \varepsilon \), then \( V_k \leq a + \varepsilon \) for all \( k \geq n \). Hence,

\[
P[\bigcap_{k \geq n} S_k(\varepsilon)] = P[\{V_n \leq a + \varepsilon, V_{n+1} \leq a + \varepsilon, \ldots\}] = P[\{V_n \leq a + \varepsilon\}] = 1 - \rho^n
\]

where \( \rho = (b - a - \varepsilon)/(b - a) \). For all \( \varepsilon > 0 \), \( \rho < 1 \) and

\[
\lim_{n \to \infty} P[\bigcap_{k \geq n} S_k(\varepsilon)] = \lim_{n \to \infty} 1 - \rho^n = 1
\]

which proves that \( V_1, V_2, \ldots \) converges to \( a \) with probability 1.

**Problem 8.4.2**

(a) From Theorem 7.2, we have

\[
\text{Var} [X_1 + \cdots + X_n] = \sum_{i=1}^{n} \text{Var} [X_i] + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \text{Cov} [X_i, X_j]
\]

Note that \( \text{Var} [X_i] = \sigma^2 \) and for \( j > i \), \( \sigma_{X_iX_j} = \sigma^2 a^{j-i} \). This implies

\[
\text{Var} [X_1 + \cdots + X_n] = n\sigma^2 + 2\sigma^2 \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} a^{j-i}
\]

\[
= n\sigma^2 + 2\sigma^2 \sum_{i=1}^{n-1} (a + a^2 + \cdots + a^{n-i})
\]

\[
= n\sigma^2 + 2\frac{a \sigma^2}{1-a} \sum_{i=1}^{n-1} (1 - a^{n-i})
\]

\[
= n\sigma^2 + 2\frac{a \sigma^2}{1-a} (n - 1) - 2\frac{a}{1-a} (a + a^2 + \cdots + a^{n-1})
\]

\[
= \left( \frac{n(1+a) \sigma^2}{1-a} \right) - 2\frac{a \sigma^2}{1-a} \left( \frac{a}{1-a} \right)^2 (1 - a^{n-1})
\]

Since \( a/(1-a) \) and \( 1-a^{n-1} \) are both nonnegative,

\[
\text{Var} [X_1 + \cdots + X_n] \leq n\sigma^2 \left( \frac{1+a}{1-a} \right)
\]
(b) Since the expected value of a sum equals the sum of the expected values,

\[ E[M(X_1, \ldots, X_n)] = \frac{E[X_1] + \cdots + E[X_n]}{n} = \mu \]

The variance of \( M(X_1, \ldots, X_n) \) is

\[ \text{Var}[M(X_1, \ldots, X_n)] = \frac{\text{Var}[X_1 + \cdots + X_n]}{n^2} \leq \frac{\sigma^2(1 + a)}{n(1 - a)} \]

Applying the Chebyshev inequality to \( M(X_1, \ldots, X_n) \) yields

\[ P[|M(X_1, \ldots, X_n) - \mu| \geq c] \leq \frac{\text{Var}[M(X_1, \ldots, X_n)]}{c^2} \leq \frac{\sigma^2(1 + a)}{n(1 - a)c^2} \]

(c) Taking the limit as \( n \) approaches infinity of the bound derived in part (b) yields

\[ \lim_{n \to \infty} P[|M(X_1, \ldots, X_n) - \mu| \geq c] \leq \lim_{n \to \infty} \frac{\sigma^2(1 + a)}{n(1 - a)c^2} = 0 \]

Thus

\[ \lim_{n \to \infty} P[|M(X_1, \ldots, X_n) - \mu| \geq c] = 0 \]

**Problem 9.2.1**

For the MAP test, we must choose acceptance regions \( A_0 \) and \( A_1 \) for the two hypotheses \( H_0 \) and \( H_1 \). From Theorem 9.2, the MAP rule is

- \( n \in A_0 \) if

\[ \frac{P_{N|H_0}(n)}{P_{N|H_1}(n)} \geq \frac{P[H_1]}{P[H_0]} \]

- \( n \in A_1 \) if

\[ \frac{P_{N|H_0}(n)}{P_{N|H_1}(n)} < \frac{P[H_1]}{P[H_0]} \]

Since \( P_{N|H_i}(n) = \lambda_i^n e^{-\lambda_i}/n! \), the MAP rule becomes

- \( n \in A_0 \) if

\[ \left( \frac{\lambda_0}{\lambda_1} \right)^n e^{-(\lambda_0 - \lambda_1)} \geq \frac{P[H_1]}{P[H_0]} \]

- \( n \in A_1 \) if

\[ \left( \frac{\lambda_0}{\lambda_1} \right)^n e^{-(\lambda_0 - \lambda_1)} < \frac{P[H_1]}{P[H_0]} \]
By taking logarithms and assuming $\lambda_1 > \lambda_0$ yields the final form of the MAP rule

- $n \in A_0$ if
  \[
  n \leq n^* = \frac{\lambda_1 - \lambda_0 + \ln(P[H_0]/P[H_1])}{\ln(\lambda_1/\lambda_0)}
  \]

- $n \in A_1$ if $n > n^*$.

From the MAP rule, we can get the ML rule by setting the a priori probabilities to be equal. This yields the ML rule

- $n \in A_0$ if
  \[
  n \leq n^* = \frac{\lambda_1 - \lambda_0}{\ln(\lambda_1/\lambda_0)}
  \]

- $n \in A_1$ if $n > n^*$

**Problem 9.3.1**

Since the three hypotheses $H_0$, $H_1$, and $H_2$ are equally likely, the MAP and ML hypothesis tests are the same. From Theorem 9.7, the MAP rule is

- $x \in A_m$ if $f_{X|H_m}(x) \geq f_{X|H_j}(x)$ for all $j$

Since $N$ is Gaussian with zero mean and variance $\sigma_N^2$, the conditional PDF of $X$ given $H_i$ is

\[
f_{X|H_i}(x) = \frac{1}{\sqrt{2\pi\sigma_N^2}} e^{-\frac{(x-a(i-1))^2}{2\sigma_N^2}}
\]

Thus, the MAP rule is

- $x \in A_m$ if $(x - a(m - 1))^2 \leq (x - a(j - 1))^2$ for all $j$

This implies that the rule for membership in $A_0$ is

- $x \in A_0$ if $(x + a)^2 \leq x^2$ and $(x + a)^2 \leq (x - a)^2$.

This rule simplifies to

- $x \in A_0$ if $x \leq -a/2$

Similar rules can be developed for $A_1$ and $A_2$. These are:

- $x \in A_1$ if $-a/2 \leq x \leq a/2$

- $x \in A_2$ if $x \geq a/2$

To summarize, the three acceptance regions are

\[
A_0 = \{x|x \leq -a/2\} \quad A_1 = \{x| -a/2 < x \leq a/2\} \quad A_2 = \{x|x > a/2\}
\]

Graphically, the signal space is one dimensional and the acceptance regions are

\[\begin{array}{c|c|c}
\text{Region} & s_0 & s_1 \\
\hline
A_0 & -a & s_0 \\
A_1 & s_0 & s_1 \\
A_2 & s_1 & a \\
\end{array}\]

Just as in the QPSK system of Example 9.9, the additive Gaussian noise dictates that the acceptance region $A_i$ is the set of observations $x$ that are closer to $s_i = (i - 1)a$ than any other $s_j$. 