Problem 1. $Y = X_1 + X_2 + \ldots + X_n$

Method 1: The maximum value taken by $Y$ is $n$. 
$Y$ takes on the value $k$, $(0 \leq k \leq n)$, when $k$ of the $X_i$'s are 1 and the rest are 0. 
Hence $P(Y = k) = \binom{n}{k}p^k(1-p)^{n-k}$

Method 2: $G_{X_i}(z) = E[z^{X_i}] = z^1p + z^0(1-p) = zp + 1 - p$
$G_Y(z) = E[z^Y] = E[z^{\sum_{i=1}^n X_i}]$

Since all $X_i$ are independent, 
$G_Y(z) = E[z^Y] = \prod_{i=1}^n E[z^{X_i}] = (zp + 1 - p)^n$

which is the same as the probability generating function of a Binomial Random variable with parameters $n$ and $p$. $B(n,p)$.

Method 3: Define $Y_1 = X_1$, 
$Y_2 = X_1 + X_2 = Y_1 + X_1$
$Y_3 = Y_2 + X_2$ and so on 
then $Y_n = Y_{n-1} + X_n$

$P(Y_2 = y) = \sum_{k=0}^1 P(Y_1 = n - k)P(X_2 = k)$

The above is convolution of the pmf of $Y_1$ and $X_2$.

$$P(Y_2 = y) = \sum_{k=0}^1 P(Y_1 = y - k)P(X_2 = k)$$
$$= \sum_{k=0}^1 P(Y_1 = y)P(X_2 = 0) + P(Y_1 = y - 1)P(X_2 = 1)$$
$$= P(X_1 = y)P(X_2 = 0) + P(X_1 = y - 1)P(X_2 = 1)$$
$$= \binom{n}{y}p^y(1-p)^{n-y}$$

Next we will try to prove that $Y_n$ is $B(n,p)$ by induction i.e. $P(Y_n = y) = \binom{n}{y}p^y(1-p)^{n-y}, (0 \leq y \leq n)$
Assume that $Y_{n-1}$ is $B(n-1, p)$ i.e. $P(Y_{n-1} = y) = \binom{n-1}{y}p^y(1-p)^{n-1-y}(0 \leq y \leq n-1)$.

$$P(Y_n = y) = \sum_{k=0}^{1} P(Y_{n-1} = n-y)P(X_n = k)$$
$$= P(Y_{n-1} = y)P(X_n = 0) + P(Y_{n-1} = y-1)P(X_n = 1)$$
$$= \binom{n-1}{y}p^y(1-p)^{n-1-y}(1-p) + \binom{n-1}{y-1}p^{y-1}(1-p)^{n-y}p$$
$$= \binom{n}{y}p^y(1-p)^{n-y}$$

**Problem 2.**

$$\mathcal{X} = \{X_1, X_2, \ldots, X_m\} \quad \mathcal{Y} = \{X_{m+1}, X_{m+2}, \ldots, X_n\}$$

We can easily derive $P(\mathcal{X} \in S_x, \mathcal{Y} \in S_y) = P(\mathcal{X} \in S_x)P(\mathcal{Y} \in S_y)$

If $g(\mathcal{X}) \in G_a$, then $g^{-1}(G_a)$ could be an arbitrary region in the m-dimensional probability space. Such a space can be approximated via a countable number (possibly infinitely many) of disjoint events.

For example consider a two dimensional space spanned by two RVs as shown in Figure 1.

Then,

$$P(f(X_1, X_2) \in F_a) = P(\{X_1 \in X_{1a}, X_2 \in X_{2a}\} \cup \{X_1 \in X_{1b}, X_2 \in X_{2b}\} \cup \{X_1 \in X_{1b}, X_2 \in X_{2a}\})$$
Hence we can write,
\[ P(g(X) \in G_a) = P(\bigcup_i X_i \in S_{X_i}) \]
and
\[ P(g(Y) \in G_b) = P(\bigcup_j Y_j \in S_{Y_j}) \]

\[ P(g(X) \in G_a, h(Y) \in H_b) = P(\{\bigcup_i X_i \in S_{X_i}\} \cap \{\bigcup_i Y_i \in S_{Y_i}\}) \quad (1) \]
\[ = P(\bigcup_{i,j} \{X_i \in S_{X_i} \cap Y_j \in S_{Y_j}\}) \quad (2) \]
\[ = \sum_{i,j} P(X_i \in S_{X_i}) P(Y_j \in S_{Y_j}) \quad (3) \]
\[ = \sum_{i,j} P(X_i \in S_{X_i}) \sum_j P(Y_j \in S_{Y_j}) \quad (4) \]
\[ = P(\{\bigcup_i X_i \in S_{X_i}\}) P(\{\bigcup_j Y_j \in S_{Y_j}\}) \quad (5) \]
\[ = P(g(X) \in G_a, h(Y) \in H_b) \quad (6) \]
\[ = P(g(X) \in G_a) P(h(Y) \in H_b) \quad (7) \]

A simpler way to solve the problem is to go back to the equation
\[ P(X \in S_x, Y \in S_y) = P(X \in S_x) P(Y \in S_y) \]
Since this is true for all sets \(S_x\) and \(S_y\), we choose \(S_x = g^{-1}(G_a)\) and \(S_y = h^{-1}(H_b)\). In that case we have:

\[ P(g(X) \in G_a, h(Y) \in H_b) = P(X \in g^{-1}(G_a), Y \in h^{-1}(H_b)) \quad (8) \]
\[ = P(X \in g^{-1}(G_a)) P(Y \in h^{-1}(H_b)) \quad (9) \]
\[ = P(g(X) \in G_a) P(h(Y) \in H_b) \quad (10) \]

**Problem 3.** \( Y = X_1 + X_2 \)

\[ f_Y(y) = \int_{-\infty}^{+\infty} f_{X_1}(y - x) f_{X_1}(x) dx \]
Since \(X_1\) and \(X_2\) are \(U[0,1]\), we have the following constraints: \(0 \leq x \leq 1\) and \(0 \leq y - x \leq 1\) which give the following limits \(x \leq y\) and \(y - 1 \leq x\)

When \(y > 1\) we have the following constraints: \(x \leq 1\) and \(y - 1 \leq x\)
Similarly, when \(y < 1\) we have the following constraints: \(x \leq y\) and \(0 \leq x\)

When \(y > 1\), \(f_Y(y) = f_{y-1}^1 dx = 2 - y\)
When \(y < 1\), \(f_Y(y) = f_0^y dx = y\)

Hence,

\[ f_Y(y) = \begin{cases} 
0 & y < 0 \\
\frac{y}{1-y} & 1 \leq y < 1 \\
2-y & 1 \leq y < 2 \\
0 & y \geq 2 
\end{cases} \]
Problem 4.

\[ f_{X_1X_2}(x_1, x_2) = \begin{cases} 1 & 0 \leq x_1, x_2 \leq 1 \\ 0 & \text{otherwise} \end{cases} \]

\[ Y = (Y_1, Y_2)^T = A(X_1, X_2)^T \]

\[ A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \]

The jpdf of \( Y_1 \) and \( Y_2 \) exists only when the determinant of the matrix \( A \) exists i.e. \( |A| \neq 0 \).

When the above is true, we can find \( f_{Y_1Y_2}(y_1, y_2) \) as follows:

\[ X = (X_1, X_2)^T = A^{-1}(Y_1, Y_2)^T \]

\[ A^{-1} = \frac{1}{|\text{mathbf{A}}|} \begin{pmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{pmatrix} \]

\[ |A| = A_{11}A_{22} - A_{12}A_{21} \]

\[ X_1 = \frac{1}{|A|}(A_{22}Y_1 - A_{12}Y_2) \]

\[ X_2 = \frac{1}{|A|}(-A_{21}Y_1 + A_{12}Y_2) \]

\[ f_{Y_1Y_2}(y_1, y_2) = \frac{1}{|A|}f_{X_1X_2}(\frac{1}{|A|}(A_{22}Y_1 - A_{12}Y_2), \frac{1}{|A|}(-A_{21}Y_1 + A_{12}Y_2)) \]

Remember that \( 0 \leq \frac{1}{|A|}(A_{22}Y_1 - A_{12}Y_2) \leq 1 \)

and \( 0 \leq \frac{1}{|A|}(-A_{21}Y_1 + A_{12}Y_2) \leq 1 \)

If \( |A| = 0 \), then \( A_{11}A_{22} = A_{12}A_{21} \)

Let \( K = \frac{A_{12}}{A_{22}} = \frac{A_{11}}{A_{21}} \)

Then \( Y_1 = K A_{21}X_1 + K A_{22}X_2 = KY_2 \)

\( Y_1 \) is a multiple of \( Y_2 \) and hence the jpdf of \( Y_1 \) and \( Y_2 \) does not exist. A plot of \( Y_2 \) vs \( Y_1 \) is a straight line with a slope of \( K \).

If the joint pdf does not exist we can use the joint probability distribution function (joint cdf) to characterize the the distribution of \( Y \).

Problem 5.

Let us first compute, for \( x < y \),

\[ P[X_2 \in (y, y + dy) \mid X_1 \land X_2 = x] = \frac{P(X_2 \in (y, y + dy), X_1 \in (x, x + dx))}{P(X_1 \land X_2 \in (x, x + dx))} = \frac{\lambda_1 e^{-\lambda_1 x}dx \lambda_2 e^{-\lambda_2 y}dy}{(\lambda_1 + \lambda_2)e^{-(\lambda_1 + \lambda_2)x}dx} = \frac{\lambda_1 \lambda_2}{\lambda_1 + \lambda_2}e^{-\lambda_2(y-x)}dy. \]
We used the easy fact that $X_1 \land X_2$ is exponentially distributed with rate $\lambda_1 + \lambda_2$. Integrating this expression over $y \in [x, \infty)$ we find that

$$P[X_2 \geq X_1 \mid X_1 \land X_2 = x] = \frac{\lambda_1}{\lambda_1 + \lambda_2}.$$

**Problem 6.** Note that $E[(X_i - \mu)(X_j - \mu)] = E[(X_i - \mu)]E[(X_j - \mu)]$ since $X_i$ and $X_j$, $i \neq j$ are independent. Further $E[X_i - \mu] = 0$.

$Y = X_1 + 2X_2 + X_3^2$

$E[X_i] = \int_0^1 xdx = \frac{1}{2}$

$E[X_i^2] = \int_0^1 x^2dx = \frac{1}{3}$

$E[X_i^3] = \int_0^1 x^3dx = \frac{1}{4}$

$E[X_i^4] = \int_0^1 x^4dx = \frac{1}{5}$

$E[Y] = E[X_1] + 2E[X_2] + E[X_3^2] = \frac{1}{2} + 2 \times \frac{1}{2} + \frac{1}{3} = \frac{11}{6}$

$E[Y^2] = E[X_1^2] + 4E[X_2^2] + E[X_3^4] + 4E[X_1X_2] + 4E[X_2X_3^2] + 2E[X_1X_3^2]$ Since $X_1$, $X_2$ and $X_3$ are independent. $E[X_1X_2] = E[X_1]E[X_2]$ and so on.


$$= \frac{1}{3} + 4 \times \frac{1}{3} + \frac{1}{5} + 4 \times \frac{1}{4} + 4 \times \frac{1}{6} + 2 \times \frac{1}{6}$$

$$= \frac{1}{3} + \frac{4}{3} + \frac{1}{5} + \frac{1}{2} + \frac{2}{3} + \frac{1}{3}$$

$$= \frac{58}{15}$$

$$\text{var}(Y) = E[Y^2] - E[Y]^2 = \frac{58}{15} - \frac{121}{36} = 0.5$$

**Problem 7.** By Chebyshev’s inequality we have,

$$P\left( \frac{X_1 + X_2 + \cdots + X_n}{n} - \mu \geq \epsilon \right) \leq \frac{E[(\frac{1}{n} \sum_{i=1}^{n} X_i - \mu)^2]}{\epsilon^2}$$

$$= \frac{1}{\epsilon^2 n^2} E\left[ \sum_{i=1}^{n} (X_i - \mu)^2 + 2 \sum_{1 \leq i,j \leq n, i \neq j} (X_i - \mu)(X_j - \mu) \right]$$

$$= \frac{1}{\epsilon^2 n^2} \sum_{i=1}^{n} E[(X_i - \mu)^2] + 2 \sum_{1 \leq i,j \leq n, i \neq j} E[(X_i - \mu)]E[(X_j - \mu)]$$

$$= \frac{n \sigma^2}{\epsilon^2 n^2} + \frac{\sigma^2}{\epsilon^2 n}$$
Problem 8. First lets find the pmf of $Y$.

If $x \geq m$, 
$$P(Y = m \mid X = x) = \binom{x}{m} p^m (1 - p)^{x-m}$$
else 
$$P(Y = m \mid X = x) = 0$$

$$P(Y = m) = \sum_{x=m}^{\infty} P(Y = m \mid X = x) P(X = x)$$
$$= \sum_{x=m}^{\infty} \binom{x}{m} p^m (1 - p)^{x-m} \frac{\lambda^x e^{-\lambda}}{x!}$$
$$= \frac{p^m (1 - p)^{-m} e^{-\lambda}}{m!} \sum_{x=m}^{\infty} \frac{(1 - p)^x \lambda^x}{(x - m)!}$$
$$= \frac{p^m (1 - p)^{-m} \lambda^m e^{-\lambda} (1 - p)^m}{m!} \sum_{y=0}^{\infty} \frac{((1 - p)\lambda)^y}{y!}$$
$$= \frac{(\lambda p^m e^{-\lambda} e^{-\lambda} (1 - p))^{m+n}}{m!}$$
$$= \frac{(\lambda p)^m e^{-\lambda} e^{-\lambda} (1 - p)}{m!}$$

Note that \( \sum_{n=0}^{\infty} x^n = e^x \)

Hence $Y$ is $P(\lambda p)$

Similarly we can prove that, $Z$ is $P(\lambda(1 - p))$

Now to prove independence. Consider the general case when we paint red balls with probability $p$ and blue balls with probability $q$, where $p + q$ may be less that 1. In that case.

$$P(Y = m, Z = n) = \sum_{x=m+n}^{\infty} P(Y = m, Z = n, X = x)$$
$$P(Y = m, Z = n) = \sum_{x=m+n}^{\infty} P(Y = m, Z = n \mid X = x) P(X = x)$$

In our case $p + q = 1$ so $P(Y = m, Z = n \mid X = x) = 0$ if $x \neq m + n$.

Hence we get.

$$P(Y = m, Z = n) = P(Y = m, Z = n \mid X = m + n) P(X = m + n)$$
$$= \binom{m+n}{m} p^m (1 - p)^n \frac{\lambda^{m+n} e^{-\lambda}}{(m+n)!}$$
$$= \frac{m+n}{m} p^m (1 - p)^n \frac{\lambda^{m+n} e^{-\lambda} e^{-\lambda(1 - p)}}{(m+n)!}$$
$$= \frac{(\lambda p)^m e^{-\lambda} (\lambda(1 - p))^{m+n} e^{-\lambda(1 - p)}}{n!}$$
$$= P(Y = m) P(Z = n)$$