Problem 2.2.5

Using $B$ (for Bad) to denote a miss and $G$ (for Good) to denote a successful free throw, the sample tree for the number of points scored in the 1 and 1 is

From the tree, the PMF of $Y$ is

$$P_Y(y) = \begin{cases} 
1 - p & y = 0 \\
(1 - p)p & y = 1 \\
p^2 & y = 2 \\
0 & \text{otherwise}
\end{cases}$$

Problem 2.2.9

(a) In the setup of a mobile call, the phone will send the “SETUP” message up to six times. Each time the setup message is sent, we have a Bernoulli trial with success probability $p$. Of course, the phone stops trying as soon as there is a success. Using $r$ to denote a successful response, and $n$ a non-response, the sample tree is

(b) We can write the PMF of $K$, the number of “SETUP” messages sent as

$$P_K(k) = \begin{cases} 
(1 - p)^{k-1}p & k = 1, 2, \ldots, 5 \\
(1 - p)^5p + (1 - p)^6 = (1 - p)^5 & k = 6 \\
0 & \text{otherwise}
\end{cases}$$

Note that the expression for $P_K(6)$ is different because $K = 6$ if either there was a success or a failure on the sixth attempt. In fact, $K = 6$ whenever there were failures on the first five attempts which is why $P_K(6)$ simplifies to $(1 - p)^5$.

(c) Let $B$ denote the event that a busy signal is given after six failed setup attempts. The probability of six consecutive failures is $P[B] = (1 - p)^6$. To be sure that $P[B] \leq 0.02$, we need $p \geq 1 - (0.02)^{1/6} = 0.479$. 

1
Problem 2.3.5
Each paging attempt is a Bernoulli trial with success probability \( p \) where a success occurs if the pager receives the paging message.

(a) The paging message is sent again and again until a success occurs. Hence the number of paging messages is \( N = n \) if there are \( n - 1 \) paging failures followed by a paging success. That is, \( N \) has the geometric PMF

\[
P_N(n) = \begin{cases} 
(1 - p)^{n-1}p & n = 1, 2, \ldots \\
0 & \text{otherwise}
\end{cases}
\]

(b) The probability that no more three paging attempts are required is

\[
P[N \leq 3] = 1 - P[N > 3] = 1 - \sum_{n=4}^{\infty} P_N(n) = 1 - (1 - p)^3
\]

This answer can be obtained without calculation since \( N > 3 \) if the first three paging attempts fail and that event occurs with probability \((1 - p)^3\). Hence, we must choose \( p \) to satisfy \( 1 - (1 - p)^3 \geq 0.95 \) or \((1 - p)^3 \leq 0.05\). This implies

\[p \geq 1 - (0.05)^{1/3} \approx 0.6316\]

Problem 2.4.7
From Problem 2.2.9, the PMF of the number of call attempts is

\[
P_N(n) = \begin{cases} 
(1 - p)^{n-1}p & n = 1, 2, \ldots, 5 \\
(1 - p)^{5}p + (1 - p)^{6} = (1 - p)^{5} & k = 6 \\
0 & \text{otherwise}
\end{cases}
\]

For \( p = 1/2 \), the PMF can be simplified to

\[
P_N(n) = \begin{cases} 
(1/2)^n & n = 1, 2, \ldots, 5 \\
(1/2)^5 & n = 6 \\
0 & \text{otherwise}
\end{cases}
\]

The corresponding CDF of \( N \) is

\[
F_N(n) = \begin{cases} 
0 & n < 1 \\
1/2 & 1 \leq n < 2 \\
3/4 & 2 \leq n < 3 \\
7/8 & 3 \leq n < 4 \\
15/16 & 4 \leq n < 5 \\
31/32 & 5 \leq n < 6 \\
1 & n \geq 6
\end{cases}
\]
Problem 2.5.10

By the definition of the expected value,

\[
E[X_n] = \sum_{x=1}^{n} x \left( \frac{n}{x} \right) p^x (1-p)^{n-x}
\]

\[
= np \sum_{x=1}^{n} \frac{(n-1)!}{(x-1)! \cdot (n-1-(x-1))!} p^{x-1} (1-p)^{n-1-(x-1)}
\]

With the substitution \(x' = x - 1\), we have

\[
E[X_n] = np \sum_{x'=0}^{n-1} \left( \frac{n-1}{x'} \right) p^{x'} (1-p)^{n-x'} = np \sum_{x'=0}^{n-1} P_{X_{n-1}}(x) = np
\]

The above sum is 1 because it is the sum of a binomial random variable for \(n-1\) trials over all possible values.

Problem 2.5.11

We write the sum as a double sum in the following way:

\[
\sum_{i=0}^{\infty} P[X > i] = \sum_{i=0}^{\infty} \sum_{j=i+1}^{\infty} P_X(j)
\]

At this point, the key step is to reverse the order of summation. You may need to make a sketch of the feasible values for \(i\) and \(j\) to see how this reversal occurs. In this case,

\[
\sum_{i=0}^{\infty} P[X > i] = \sum_{j=1}^{\infty} \sum_{i=0}^{j-1} P_X(j) = \sum_{j=1}^{\infty} j P_X(j) = E[X]
\]

Problem 2.7.7

Let \(W\) denote the event that a circuit works. The circuit works and generates revenue of \(k\) dollars if all of its 10 constituent devices work. For each implementation, standard or ultra-reliable, let \(R\) denote the profit on a device. We can express the expected profit as

\[
\]

Let’s first consider the case when only standard devices are used. In this case, a circuit works with probability \(P[W] = (1-q)^{10}\). The profit made on a working device is \(k - 10\) dollars while a nonworking circuit has a profit of -10 dollars. That is, \(E[R|W] = k - 10\) and \(E[R|W'] = -10\). Of course, a negative profit is actually a loss. Using \(R_s\) to denote the profit using standard circuits, the expected profit is

\[
E[R_s] = (1-q)^{10}(k-10) + (1-(1-q)^{10})(-10) = (0.9)^{10}(k-10)
\]
And for the ultra-reliable case, the circuit works with probability \( P = (1 - q/2)^{10} \). The profit per working circuit is \( E[R|W] = k - 30 \) dollars while the profit for a nonworking circuit is \( E[R|W^c] = -30 \) dollars. The expected profit is

\[
E[R_u] = (1 - q/2)^{10}(k - 30) + (1 - (1 - q/2)^{10})(-30) = (0.95)^{10}k - 30
\]

Now we wish to determine which implementation will generate the most profit. Realizing that both profit functions are linear functions of \( k \), we can plot them versus \( k \) and find for which values of \( k \) each plan is preferable. The two lines intersect at a value of \( k = 80 \) dollars. So for values of \( k < 80 \) using all standard devices results in greater revenue, and for values of \( k > 80 \) more revenue will be generated by implementing all ultra-reliable devices. So we can see that when the price commanded for each working circuit is sufficiently high it is worthwhile to spend the extra money to ensure that more working circuits can be produced.

**Problem 2.8.6**

(a) The expected value of \( X \) is

\[
E[X] = \sum_{x=0}^{5} xP_X (x)
\]

\[
= 0 \binom{5}{0} \frac{1}{2^5} + 1 \binom{5}{1} \frac{1}{2^5} + 2 \binom{5}{2} \frac{1}{2^5} + 3 \binom{5}{3} \frac{1}{2^5} + 4 \binom{5}{4} \frac{1}{2^5} + 5 \binom{5}{5} \frac{1}{2^5} = \frac{5}{2}
\]

The expected value of \( X^2 \) is

\[
E[X^2] = \sum_{x=0}^{5} x^2P_X (x)
\]

\[
= 0^2 \binom{5}{0} \frac{1}{2^5} + 1^2 \binom{5}{1} \frac{1}{2^5} + 2^2 \binom{5}{2} \frac{1}{2^5} + 3^2 \binom{5}{3} \frac{1}{2^5} + 4^2 \binom{5}{4} \frac{1}{2^5} + 5^2 \binom{5}{5} \frac{1}{2^5} = \frac{240}{32} = \frac{15}{2}
\]

The variance of \( X \) is

\[
\text{Var}[X] = E[X^2] - (E[X])^2 = \frac{15}{2} - \frac{25}{4} = \frac{5}{4}
\]

By taking the square root of the variance, the standard deviation of \( X \) is \( \sigma_X = \sqrt{\frac{5}{4}} \approx 1.12 \).

(b) The probability that \( X \) is within one standard deviation of its mean is

\[
P[\mu_X - \sigma_X \leq X \leq \mu_X + \sigma_X] = P[2.5 - 1.12 \leq X \leq 2.5 + 1.12] = P[1.38 \leq X \leq 3.62] = P[2 \leq X \leq 3]
\]
By summing the PMF over the desired range, we obtain

\[ P[2 \leq X \leq 3] = P_X(2) + P_X(3) = 10/32 + 10/32 = 5/8 \]

**Problem 2.8.8**

Given the following description of the random variable \( Y \),

\[ Y = \frac{1}{\sigma_X} (X - \mu_X) \]

we can use the linearity property of the expectation operator to find the mean value

\[ E[Y] = \frac{E[X] - \mu_X}{\sigma_X} = \frac{E[X] - E[X]}{\sigma_X} = 0 \]

Using the fact that \( \text{Var}[aX + b] = a^2 \text{Var}[X] \), the variance of \( Y \) is found to be

\[ \text{Var}[Y] = \frac{1}{\sigma_X^2} \text{Var}[X] = 1 \]

**Problem 2.8.10**

We wish to minimize the function

\[ e(\hat{x}) = E[(X - \hat{x})^2] \]

with respect to \( \hat{x} \). We can expand the square and take the expectation while treating \( \hat{x} \) as a constant. This yields

\[ e(\hat{x}) = E[X^2 - 2\hat{x}X + \hat{x}^2] = E[X^2] - 2\hat{x}E[X] + \hat{x}^2 \]

Solving for the value of \( \hat{x} \) that makes the derivative \( de(\hat{x}) / d\hat{x} \) equal to zero results in the value of \( \hat{x} \) that minimizes \( e(\hat{x}) \). Note that when we take the derivative with respect to \( \hat{x} \), both \( E[X^2] \) and \( E[X] \) are simply constants.

\[ \frac{d}{d\hat{x}} (E[X^2] - 2\hat{x}E[X] + \hat{x}^2) = 2E[X] - 2\hat{x} = 0 \]

Hence we see that \( \hat{x} = E[X] \). In the sense of mean squared error, the best guess for a random variable is the mean value. In Chapter 9 this idea is extended to develop minimum mean squared error estimation.

**Problem 2.9.5**

The probability of the event \( B \) is

\[ P[B] = P[X \geq \mu_X] = P[X \geq 3] = P_X(3) + P_X(4) + P_X(5) = \frac{\binom{3}{3} + \binom{3}{4} + \binom{3}{5}}{32} = 21/32 \]
The conditional PMF of $X$ given $B$ is

$$P_{X|B}(x) = \begin{cases} \frac{P_X(x)}{P(B)} & x \in B \\ 0 & \text{otherwise} \end{cases}$$

for $x = 3, 4, 5$.

The conditional first and second moments of $X$ are

$$E[X|B] = \sum_{x=3}^5 x P_{X|B}(x) = 3 \left( \frac{5}{3} \right) \frac{1}{21} + 4 \left( \frac{5}{4} \right) \frac{1}{21} + 5 \left( \frac{5}{5} \right) \frac{1}{21} = [30 + 20 + 5]/21 = 55/21$$

$$E[X^2|B] = \sum_{x=3}^5 x^2 P_{X|B}(x) = 3^2 \left( \frac{5}{3} \right) \frac{1}{21} + 4^2 \left( \frac{5}{4} \right) \frac{1}{21} + 5^2 \left( \frac{5}{5} \right) \frac{1}{21} = [90 + 80 + 25]/21 = 195/21 = 65/7$$

The conditional variance of $X$ is

$$\text{Var}[X|B] = E[X^2|B] - (E[X|B])^2 = 65/7 - (55/21)^2 = 1070/441 = 2.43$$

**Problem 3.1.2**

On the $X,Y$ plane, the joint PMF is

(a) To find $c$, we sum the PMF over all possible values of $X$ and $Y$. We choose $c$ so the sum equals one.

$$\sum_{x} \sum_{y} P_{X,Y}(x,y) = \sum_{x=-2,0,2,y=-1,0,1} c \cdot |x+y| = 6c + 2c + 6c = 14c$$

Thus $c = 1/14$.

(b) $P[Y < X] = P_{X,Y}(0,-1) + P_{X,Y}(2,-1) + P_{X,Y}(2,0) + P_{X,Y}(2,1)$

$$= c + c + 2c + 3c = 7c = 1/2$$
(c) 
\[ P[Y > X] = P_{X,Y}(-2,-1) + P_{X,Y}(-2,0) + P_{X,Y}(-2,1) + P_{X,Y}(0,1) \]
\[ = 3c + 2c + c + c = 7c = 1/2 \]

(d) From the sketch of \( P_{X,Y}(x,y) \) given above, \( P[X = Y] = 0 \).

(e) 
\[ P[X < 1] = P_{X,Y}(-2,-1) + P_{X,Y}(-2,0) + P_{X,Y}(-2,1) + P_{X,Y}(0,-1) + P_{X,Y}(0,1) \]
\[ = 8c = 8/14 \]

Problem 3.1.6
The joint PMF of \( X \) and \( N \) is \( P_{N,X}(n,x) = P[N = n, X = x] \), which is the probability that \( N = n \) and \( X = x \). This means that both events must be satisfied. The approach we use is similar to that used in finding the Pascal PMF in Example 2.15. Since \( X \) can take on only the two values 0 and 1, let’s consider each in turn. When \( X = 0 \) that means that a rejection occurred on the last test and that the other \( n - 1 \) rejections must have occurred in the previous \( r - 1 \) tests. Thus,
\[ P_{N,X}(n,0) = \binom{r-1}{n-1} (1-p)^{n-1} p^{r-1-(n-1)} (1-p) \quad n = 1, \ldots, r \]
When \( X = 1 \) the last test was acceptable and therefore we know that the \( N = n \leq r - 1 \) tails must have occurred in the previous \( r - 1 \) tests. In this case,
\[ P_{N,X}(n,1) = \binom{r-1}{n} (1-p)^n p^{r-n} \quad n = 0, \ldots, r - 1 \]
We can combine these cases into a single complete expression for the joint PMF.
\[ P_{X,N}(x,n) = \begin{cases} 
\binom{r-1}{n-1} (1-p)^{n-1} p^{r-n} & x = 0, n = 1, 2, \ldots, r \\
\binom{r-1}{n} (1-p)^n p^{r-n} & x = 1, n = 0, 1, \ldots, r - 1 \\
0 & \text{otherwise}
\end{cases} \]

Problem 3.3.5
The \( x,y \) pairs with nonzero probability are shown in the figure at right. From the figure, we observe that for \( w = 0, 1, \ldots, 10, \)
\[ P[W > w] = P[\min(X,Y) > w] \]
\[ = P[X > w, Y > w] \]
\[ = 0.01(10 - w)^2 \]
To find the PMF of $W$, we observe that for $w = 1, \ldots, 10,$

$$P_W(w) = P[W > w - 1] - P[W > w]$$

$$= 0.01[(10 - w - 1)^2 - (10 - w)^2]$$

$$= 0.01(21 - 2w)$$

The complete expression for the PMF of $W$ is

$$P_W(w) = \begin{cases} 
0.01(21 - 2w) & w = 1, 2, \ldots, 10 \\
0 & \text{otherwise}
\end{cases}$$