1. If \( X \) is normally distributed with mean 2 and variance 1, find \( \Pr[X < 2] \).

\[
\Pr[X < 2] = \Pr[-1 < X - 2 < 1] = \Pr[1 < X < 3] = \Pr[X < 3] - \Pr[X < 1] = \\
\Pr\left[\frac{X - 2}{1} < \frac{3 - 2}{1}\right] - \Pr\left[\frac{X - 2}{1} < \frac{1 - 2}{1}\right] = \Pr[z < 1] - \Pr[z < -1] = \\
0.8413 - 0.1586 = 0.6827
\]

2. A distributor of bean seeds determines from extensive testing that 5% of a large batch of seeds will not germinate. He sells seeds in packages of 200 and guarantees 90% germination. What is the probability that a given package will violate the guarantee?

\( X \) = number of seeds out of 200 that do not germinate; probability of single seed not germinating is 0.05...\( X \) ~ Binomial \((p = 0.05, n = 200)\)

Violate guarantee if more than 90% of 200 (i.e., 20) seeds do not germinate.

\( \Pr[X > 20] = 1 - \Pr[X \leq 20] = 1 - 0.9988 = 0.0012 \)

3. A fair coin is tossed until a head appears. Let \( X \) denote the number of tosses required.

(a) Find the density of \( X \).

(b) Find the mean and variance of \( X \).

Must toss tails \((X - 1)\) times prior to a head.

\( \Pr[\text{Tail}] = \Pr[\text{Head}] = 0.5. \)

\( X \) ~ Geometric with \( p = 0.5 \)

(a) PDF: \( f(x) = p(1 - p)^{x-1}, \ x \in \{1, 2, 3, \ldots\} \)

(b) Mean: \( E(X) = \frac{1}{p} = \frac{1}{0.5} = 2 \)

Variance: \( Var(X) = \frac{1-p}{p^2} = \frac{1-0.5}{0.5^2} = \frac{1}{0.5} = 2 \)
4. \textit{(Hyghens problem)} A and B alternately throw a fair pair of six sided dice. A wins if he scores 6 points before B gets 7 points, in which case B wins. If A starts the game what is the probability she wins?

Tossing pair of dice has following results:

<table>
<thead>
<tr>
<th>d1 \ d2</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
</tbody>
</table>

Distribution of sum of pair of dice:

\[
\begin{array}{ccccccccccc}
\end{array}
\]

A wins with probability \(a = 5/36\)
B wins with probability \(b = 6/36\)

Toss 1:
A wins: \(p = a\)
B wins: \((1 - a)b\)

Next toss: \((1 - a)(1 - b)\)

Toss 2:
A wins: \(a(1 - a)(1 - b)\)
B wins: \((1 - a)^2(1 - b)\)

Next toss: \((1 - a)^3(1 - b)^2\)

\[
A \text{ wins: } \sum_{n=1}^{\infty} a[(1-a)(1-b)]^{n-1} = a \sum_{n=0}^{\infty} [(1-a)(1-b)]^n = \frac{a}{1-(1-a)(1-b)} = \frac{\frac{5}{36}}{1-\frac{31}{36}} = 0.4918
\]

\[
B \text{ wins: } \sum_{n=1}^{\infty} (1-a)^n(1-b)^{n-1}b = (1-a)b \sum_{n=0}^{\infty} [(1-a)(1-b)]^n = \frac{(1-a)b}{1-(1-a)(1-b)} = \frac{\frac{1-\frac{5}{36}}{\frac{6}{36}}}{1-\frac{31}{36}} = 0.5082
\]
5. A student in the Department of Economics takes a multiple-choice test consisting of twenty problems, each of which have four answers to choose from. As he missed most of his classes, he only knows the answers to eight of the questions, can narrow the choice to one of two answers for four of the questions, and has no idea as to the correct answers to the remaining eight questions. Calculate:

(a) The expected number of answers that the student will get right.
(b) The probability that the student will answer 70% of the questions correctly.
(c) The probability that the student will answer 50% of the questions correctly.

\[ X_1 = \text{# correct for 8 certain questions... } E(X_1) = 8 \]
\[ X_2 = \text{# correct for 4 50-50 questions; } X_2 \sim \text{Binomial (} p = 0.5, n = 4); E(X_2) = 4(0.5) = 2 \]
\[ X_3 = \text{# correct for 8 clueless questions; } X_3 \sim \text{Binomial (} p = 0.25, n = 8); E(X_3) = 8(0.25) = 2 \]

(a) \( E(X) = 8 + 2 + 2 = 12 \)

(b) 70% = 14 questions correct... needs 6 more

\[
\begin{array}{c|cccc|c}
 X_1 \times X_2 & 0 & 1 & 2 & 3 & 4 \\
\hline
 0 & 0.00626 & 0.02503 & 0.03754 & 0.02503 & 0.00626 & 0.10011 \\
 1 & 0.01669 & 0.06674 & 0.10011 & 0.06674 & 0.01669 & 0.26697 \\
 2 & 0.01947 & 0.07787 & 0.11680 & 0.07787 & 0.01947 & 0.31146 \\
 3 & 0.01298 & 0.05191 & 0.07787 & 0.05191 & 0.01298 & 0.20764 \\
 4 & 0.00541 & 0.02163 & 0.03244 & 0.02163 & 0.00541 & 0.08652 \\
 5 & 0.00144 & 0.00577 & 0.00865 & 0.00577 & 0.00144 & 0.02307 \\
 6 & 0.00024 & 0.00096 & 0.00144 & 0.00096 & 0.00024 & 0.00385 \\
 7 & 0.00002 & 0.00009 & 0.00014 & 0.00009 & 0.00002 & 0.00037 \\
 8 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00002 \\
\end{array}
\]

\[ Pr[X_2 = 2] = \frac{1}{10} \]
\[ Pr[X_3 = 7] = \frac{1}{8} \]
\[ Pr[X_2 = 4] = \frac{1}{40} \]

\[ Pr(X = 70\%) = 0.10983 \] (add blue cells)
\[ Pr(X \geq 70\%) = 0.16969 \] (add blue cells and all cells below the line in table)

(c) 50% = 10 questions correct... needs 2 more

\[ Pr(X = 50\%) = 0.12375 \] (add blue cells)
\[ Pr(X \geq 50\%) = 0.95203 \] (add blue cells and all cells below them in table)

6. If \( E(X) = E(X^2) = 0 \), show that \( Pr(X = 0) = 1 \).

\[
Var(X) = E(X^2) - [E(X)]^2 = 0
\]
\[
Var(X) = \sum (x - E(x))^2 f(x) = \int (x - E(x))^2 f(x) dx
\]
\[ \Rightarrow \text{ all } x = 0 \quad (\text{or } f(x) = 0 \text{ which is the same thing}) \]

Either way, \( Pr(X = 0) = 1 \)
7. Let $X_i$ be random variables distributed $N(i, i^2)$, $i = 1, 2, 3$. Assume that $X_1$, $X_2$, and $X_3$ are independent. Using only these three random variables:
(a) Give an example of a statistic that has a chi-square distribution with three degrees of freedom.
(b) Give an example of a statistic that has an $F$ distribution with one, two degrees of freedom.
(c) Give an example of a statistic that has a $t$ distribution with two degrees of freedom.

(a) $z_i = \frac{X_i - i}{i} \sim N(0,1)$
$z_i^2 \sim \chi^2_1$
$X = z_1^2 + z_2^2 + z_3^2 \sim \chi^2_3$

(b) $F = \frac{z_1^2 / 1}{(z_2^2 + z_3^2) / 2} \sim F_{1,2}$

(c) $t = \frac{z_1}{\sqrt{(z_2^2 + z_3^2) / 2}} \sim t_2$

Documentation

Most of the work was done using course notes. Prof Werner answered various questions in class. The two I remember were those confirming problem 2 used a binomial distribution and problem 3 used a geometric distribution.

I checked my answers with Josh Kneifel, Jon Parker, and Guille Sabbioni.

Formulas for $E(X)$ and $Var(X)$ in problem 3 were taken from *Introduction to Probability and Its Applications* by Richard L. Scheaffer, p.86 (Duxbory, 1990)