Problem 1. Note that $E_0 = E_1 \cup E_2 \cup E_3$.

(a) (1) For disjoint events, $P(E_0) = P(E_1) + P(E_2) + P(E_3)$, so $P(E_0) = 3/4$.

(2) For independent events, $1 - P(E_0)$ is the probability that none of the events occur, which is the product of the probabilities that each one doesn’t occur. Thus $1 - P(E_0) = (3/4)^3$ and $P(E_0) = 37/64$.

(3) If $E_1 = E_2 = E_3$, then $E_0 = E_1$ and $P(E_0) = 1/4$.

(b) (1) From the Venn diagram in Fig. 1, $P(E_0)$ is clearly maximized when the events are disjoint, so $\max P(E_0) = 3/4$.

![Figure 1: Venn Diagram for problem 1 (b)(1)](image)

(2) The intersection of each pair of sets has probability $1/16$. As seen in Fig. 2, $P(E_0)$ is maximized if all these pairwise intersections are identical, in which case $P(E_0) = 3(1/4 - 1/16) + 1/16 = 5/8$. One can also use the formula $P(E_0) = P(E_1) + P(E_2) + P(E_3) - P(E_1 \cap E_2) - P(E_1 \cap E_3) - P(E_2 \cap E_3) + P(E_1 \cap E_2 \cap E_3)$, and notice that all the terms except the last is fixed by the problem, and the last term cannot be made more than $\min_{i,j} P(E_i \cap E_j) = 1/16$.

![Figure 2: Venn Diagram for problem 1 (b)(2)](image)

(c) Same considerations as in (b)(2) yields the upper bound $P(E_0) \leq 3p - 2p^2$ As $P(E_0) = 1$, we find that $p \geq 1/2$. 
Problem 2. Let $L$ be the event that the loaded die is picked and $H$ the event that the honest die is picked. Let $A_i$ be the event that $i$ is turned up on the first roll, and $B_i$ be the event that $i$ is turned up on the second roll. We are given that $P(L) = 1/3$, $P(H) = 2/3$; 
$P(A_i \mid L) = 2/3$, $P(A_i \mid L) = 1/15$ for $2 \leq i \leq 6$; $P(A_i \mid H) = 1/6$ for $1 \leq i \leq 6$. Then

$$P(L \mid A_1) = \frac{P(L, A_1)}{P(A_1)} = \frac{P(A_1 \mid L) P(L)}{P(A_1 \mid L) P(L) + P(A_1 \mid H) P(H)} = \frac{2}{3}.$$ 

This is the probability that the loaded die was picked conditional on the first roll showing a 1. For two rolls we make the assumption from the physical mechanism involved in rolling a die that the outcome on the two successive rolls of a given die are independent. Thus

$$P(A_1, B_1 \mid L) = \left(\frac{2}{3}\right)^2 \text{ and } P(A_1, B_1 \mid H) = \left(\frac{1}{6}\right)^2.$$

It follows as before that

$$P(L \mid A_1, B_1) = \frac{8}{9}.$$ 

Problem 3. Since $A, B, C, D$ form a Markov chain their probability distribution is given by

$$p(a)p(b|a)p(c|b)p(d|c) \quad (1)$$

(a) Yes: Summing (1) over $d$ shows that $A, B, C$ have the probability distribution $p(a)p(b|a)p(c|b)$.

(b) Yes: The reverse of a Markov chain is also a Markov chain. Applying this to $A, B, C, D$ and using part (a) we get that $D, C, B$ is a Markov chain. Reversing again we get the desired result.

(c) Yes: Since $A, B, C, D$ is a Markov chain, given $C, D$ is independent of $B$, and thus $p(d|c) = p(d|(b, c))$. So (1) can be written as

$$p(a, (b, c), d) = p(a)p((b, c)|a)p(d|(b, c)).$$

(d) Yes, by a similar (in fact easier) reasoning as (c).

Problem 4. No. Take for example $A = D$ and let $A$ be independent of the pair $(B, C)$. Then both $A, B, C$ and $B, C, A$ (same as $B, C, D$) are Markov chains. But $A, B, C, D$ is not: $A$ is not independent of $D$ when $B$ and $C$ are given.

Problem 5.

(a)

$$E[X + Y] = \sum_{x,y} (x + y)P_{XY}(x, y) = \sum_{x} xP_{XY}(x,y) + \sum_{y} yP_{XY}(x,y) = \sum_{x} xP_{X}(x) + \sum_{y} yP_{Y}(y) = E[X] + E[Y].$$

Note that independence is not necessary here and that the argument extends to non-discrete variables if the expectation exists.
\[
E[XY] = \sum_{x,y} xyP_{XY}(x,y) = \sum_{x,y} xP_X(x)yP_Y(y) = E[X]E[Y].
\]

Note that the statistical independence was used on the second line. Let \( X \) and \( Y \) take on only the values \( \pm 1 \) and 0. An example of uncorrelated but dependent variables is

\[
P_{XY}(1,0) = P_{XY}(0,1) = P_{XY}(-1,0) = P_{XY}(0,-1) = \frac{1}{4}.
\]

An example of correlated and dependent variables is

\[
P_{XY}(1,1) = P_{XY}(-1,-1) = \frac{1}{2}.
\]

(c) Using (a), we have

\[
\]

The middle term, from (a), is \(2(E[XY] - E[X]E[Y])\). For uncorrelated variables that is zero, leaving us with \(\sigma^2_{X+Y} = \sigma^2_X + \sigma^2_Y\).

**Problem 6.** We solve the problem for a general vehicle with \( n \) wheels.

(a) Out of \( n! \) possible orderings \((n - 1)!\) has the tyre 1 in its original place. Thus tyre 1 is installed in its original position with probability \(1/n\).

(b) All tyres end up in their original position in only 1 of the \( n! \) orders. Thus the probability of this event is \(1/n!\).

(c) Let \( X_i \) be the indicator random variable that tyre \( i \) is installed in its original position, so that the number of tyres installed in their original positions is \( N = \sum_{i=1}^{n} X_i \). By (a), \( E[X_i] = 1/n \). By the linearity of expectation, \( E[N] = n(1/n) = 1 \). Note that the linearity of the expectation holds even if the \( X_i \)'s are not independent (as it is in this case).

(e) Let \( A_i \) be the event that the \( i \)th tyre remains in its original position. Then, the event we are interested in is the complement of the event \( \bigcap_{i} A_i \) and thus has probability \(1 - \Pr(\bigcap_{i} A_i)\). Furthermore, by the inclusion/exclusion formula,

\[
\Pr(\bigcap_{i} A_i) = \Pr(A_i) - \Pr(A_i \cap A_{i1}) + \Pr(A_i \cap A_{i1} \cap A_{i2}) - \ldots
\]

The \( j \)th sum above consists of \( \binom{n}{j} \) terms, each term having the value \( P(A_1 \cap \ldots \cap A_j) \). Note that this is the probability of the event that tyres 1 through \( j \) have remained in their original positions, and equals \((n-j)!/n!\). Consequently,

\[
\Pr(\bigcap_{i} A_i) = \prod_{j=1}^{n} X^j_{i} (-1)^{j-1} \frac{(n-j)!}{n!} = \prod_{j=1}^{n} (-1)^{j-1} 1/j!.
\]
and the event that no tyre remains in its original position has probability

\[ 1 - \Pr \left[ \bigcap_{i} A_i \right] = X^n \sum_{j=0}^{n} \frac{(-1)^j}{j!} \cdot \]

(For the case \( n = 4 \), the value is \( 3/8 \).)

**Problem 7.**

(a) Let \( A_i \) denote the event that \( X_i \neq X \). The event that \( X \) does not appear in the inventory is thus

\[ A = A_1 \cap \ldots A_n. \]

Note that the events \( A_1, \ldots, A_n \) are not independent—because they involve the common random variable \( X \). However, they become independent when conditioned on the value of \( X \), with \( P(A_i|X = x) = 1 - p(x) \). Thus,

\[ P(A|X = x) = (1 - p(x))^n. \]

Consequently \( P(A) = \sum_x p(x)(1 - p(x))^n \).

(b) With \( p \) the uniform distribution on \( n \) items, the above value for \( P(A) \) equals \( (1 - 1/n)^n \).

(c) For \( n \) large, \( (1 - 1/n)^n \) approaches \( 1/e \approx 37\% \). 