1. [Section 3.1 # 4.]

The sum of the two dice is divisible by 5 if and only if it is either 5 or 10. Let $A$ be the event that the sum is divisible by 5 and $B$ that both dice show 5. Taking the sample space to be the pair of all possible outcomes $S = \{(i, j) : i, j = 1, \ldots, 6\}$, we have

$$A = \{(1, 4), (2, 3), (3, 2), (4, 1), (4, 6), (5, 5), (6, 4)\}$$

and, obviously, $AB = \{(5, 5)\}$. Since all outcomes are equally likely,

$$P(B|A) = \frac{P(AB)}{P(A)} = \frac{N(AB)}{N(A)} = \frac{1}{7}.$$

2. [Which of the following statements is always true . . .]

(a) True.

If $P(A|B) > P(A)$, then we have $\frac{P(AB)}{P(B)} > P(A)$, which upon re-arrangement yields $P(AB) > P(A)P(B)$. Dividing both sides by $P(A)$, we get $\frac{P(AB)}{P(A)} > P(B)$, which is equivalent to $P(B|A) > P(B)$.

(b) False.

Here's a simple example where the statement does not hold. In the experiment of a single roll of a die, let $A$ be the event “even number comes up”, and $B$ the event “2 comes up”. As is easily verified, $P(A|B) = 1$ and $P(B) = 1/6$, but $P(B|A) = 1/3$ and $P(A) = 1/2$.

(c) False.

Again consider the experiment of a single roll of a die, but this time, let $A$ be the event “1 comes up”, $B$ the event “2 comes up”, and $C$ the event “even number comes up”. As you can easily verify, $P(A) = P(B) = 1/6$, but $P(A|C) = 0$ while $P(B|C) = 1/3$.

3. [Ghahramani, 3.2, # 6]

Let $R_i$ and $W_i$ denote the events “red chip drawn on $i$th draw” and “white chip drawn on $i$th draw”, respectively. The colors of the first four chips can alternate in two possible ways: $W_1R_2W_3R_4$ or $R_1W_2R_3W_4$. Therefore, the probability that the color of the first four chips alternates is $P(W_1R_2W_3R_4) + P(R_1W_2R_3W_4)$. 

1
To compute \( P(W_1R_2W_3R_4) \) and \( P(R_1W_2R_3W_4) \), we apply the product rule (in its general form):

\[
P(W_1R_2W_3R_4) = P(W_1)P(R_2|W_1)P(W_3|W_1R_2)P(R_4|W_1R_2W_3)
\]

\[
P(R_1W_2R_3W_4) = P(R_1)P(W_2|R_1)P(R_3|R_1W_2)P(W_4|R_1W_2R_3)
\]

It is clear that \( P(W_1) = \frac{5}{8} \) and \( P(R_1) = \frac{3}{8} \). The conditional probabilities are easy enough to compute by keeping track of how the sample space changes with each draw. For example, to compute \( P(W_3|W_1R_2) \), simply note that given that the first draw was a white and the second red, at the end of two draws, the urn contains eight (5 old + 3 new) white chips and five (3 old + 2 new) red chips. Hence, the probability of drawing a white on the third draw is \( \frac{8}{13} \). The other conditional probabilities can be computed similarly.

Thus, we have

\[
P(W_1R_2W_3R_4) = \frac{5}{8} \times \frac{3}{11} \times \frac{8}{13} \times \frac{5}{16} = \frac{75}{2288},
\]

\[
P(R_1W_2R_3W_4) = \frac{3}{8} \times \frac{5}{10} \times \frac{5}{13} \times \frac{8}{15} = \frac{1}{26}.
\]

So, the probability that the first four chips alternate in color is \( \frac{75}{2288} + \frac{1}{26} = \frac{163}{2288} \approx 0.0712 \).

4. [Ghahramani, 3.3, #19]

Let \( E \) be the event “the three balls selected the second time are all new”. Of course, if we knew exactly how many new balls were drawn the first time around, computation of \( P(E) \) would be easy. Since we do not know what exactly happened the first time, we must take into account all the possible scenarios for the first round of ball selection, and put them all together using the law of total probability. The possible scenarios in the first round are the events \( N_0, N_1, N_2 \) and \( N_3 \), where \( N_i \) denotes the event “exactly \( i \) new balls selected the first time”. These are obviously mutually exclusive events which cover all possible outcomes.

It is clear that for \( i = 0, 1, 2, 3 \),

\[
P(N_i) = \binom{8}{i} \cdot \binom{10}{3-i} \binom{18}{3}
\]

as this is the probability of choosing \( i \) new balls and \( 3 - i \) old balls out of 8 new and 10 old balls originally in the box. Also,

\[
P(E|N_i) = \binom{8-i}{3} \cdot \binom{10+i}{0} \binom{18}{3} = \binom{8-i}{3} \binom{18}{3}
\]

since given that \( i \) new balls were chosen in the first round, there are \( 8 - i \) new and \( 10 + i \) old balls in the box when the second round of selection is done.
Therefore, by the law of total probability,

\[ P(E) = \sum_{i=0}^{3} P(E|N_i)P(N_i) = \sum_{i=0}^{3} \frac{\binom{8-i}{3}}{\binom{18}{3}} \frac{\binom{8}{i}}{\binom{18}{3}} \approx 0.0383. \]

[You have to compute each term in the summation separately, and add everything up; I don’t know of any magic formula for the sum.]

5. [Acme Insurance . . .]

This kind of a problem can obviously be tackled using a tree diagram, but we give the formal mathematical solution.

Let LR, AR and HR denote the events that a randomly selected person from the general population is classified as low-risk, average-risk and high-risk, respectively. We are given that \( P(LR) = 0.2, P(AR) = 0.5 \) and \( P(HR) = 0.3 \). Note that LR, AR and HR constitute a partition of the sample space, which here is the general population.

Next, let \( F \) be the event that a randomly selected person from the general population is involved in an accident during any given one-year period. The information at hand tells us that \( P(F|LR) = 0.05, P(F|AR) = 0.15 \) and \( P(F|HR) = 0.3 \).

(a) By the law of total probability,

\[
P(F) = P(F|LR)P(LR) + P(F|AR)P(AR) + P(F|HR)P(HR)
\]

\[
= (0.05 \times 0.2) + (0.15 \times 0.5) + (0.3 \times 0.3) = 0.175.
\]

Thus, 17.5% of the general population is involved in an accident in any given year.

(b) We assume here that the distribution of low-risk, average-risk and high-risk individuals among policy-holders at Acme is the same as their distribution in the general population. We will also assume that John Smith is just a randomly selected person. So, the event that John Smith is not involved in an accident in 2005 is simply \( F^c \). We are asked to find \( P(LR|F^c) \), so we apply Bayes’ formula.

\[
P(LR|F^c) = \frac{P(F^c|LR)P(LR)}{P(F^c|LR)P(LR) + P(F^c|AR)P(AR) + P(F^c|HR)P(HR)}.
\]

Now, observe that \( P(F|LR) + P(F^c|LR) = 1 \) (a low-risk person is either involved in an accident or not.) Hence, \( P(F^c|LR) = 1 - P(F|LR) = 0.95 \). Similarly, \( P(F^c|AR) = 0.85 \) and \( P(F^c|HR) = 0.7 \). Plugging these probabilities into Bayes’ formula, we obtain

\[
P(LR|F^c) = \frac{0.95 \times 0.2}{0.95 \times 0.2 + 0.85 \times 0.5 + 0.7 \times 0.3} = \frac{38}{175} \approx 0.22.
\]
6. [Urn I originally contains . . . ]

For \( i = 0, 1, 2, 3 \), let \( B_i \) denote the event that \( i \) black balls and \( 3 - i \) red balls were transferred from Urn I to Urn II. It is clear that these constitute all the possibilities for the three balls transferred between the urns. Elementary counting tells us that

\[
P(B_i) = \frac{3 \choose i}{{8 \choose 3}} \frac{{5 \choose 3-i}}{{8 \choose 3}}
\]

Let \( A \) be the event that the ball drawn from Urn II is black.

(a) By the law of total probability,

\[
P(A) = \sum_{i=0}^{3} P(A|B_i)P(B_i)
\]

To compute \( P(A|B_i) \), we only have to observe that Urn II originally contained 2 black balls, and so, after the transfer of \( i \) black balls and \( 3 - i \) red balls from Urn I to Urn II, Urn II contains 5 balls, of which \( 2 + i \) are black. Hence, \( P(A|B_i) = \frac{2+i}{5} \). Thus,

\[
P(A) = \sum_{i=0}^{3} \left( \frac{2+i}{5} \right) \frac{3 \choose i}{{8 \choose 3}} \frac{5 \choose 3-i}{{8 \choose 3}}
\]

\[
= \frac{1}{5} \left( \frac{5}{8} \right) \left[ 2 \left( \frac{3}{0} \right) \left( \frac{5}{3} \right) + 3 \left( \frac{3}{1} \right) \left( \frac{5}{2} \right) + 4 \left( \frac{3}{2} \right) \left( \frac{5}{1} \right) + 5 \left( \frac{3}{3} \right) \left( \frac{5}{0} \right) \right]
\]

\[
= \frac{1}{5} \left( \frac{5}{8} \right) [20 + 90 + 60 + 5] = \frac{5}{8} = 0.625.
\]

(b) \[
P(B_3|A) = \frac{P(B_3 \cap A)}{P(A)} = \frac{P(A|B_3)P(B_3)}{P(A)} = \frac{1 \times \frac{1}{35}}{5/8} = \frac{1}{35}.
\]

Bonus question

(a) We use induction on \( n \). Let \( n = 2 \). From class we know that for any two events \( A \) and \( B \),

\[
P(A \cup B) = P(A) + P(B) - P(AB).
\]

Therefore, if \( P(AB) = 0 \), we obtain \( P(A \cup B) = P(A) + P(B) \) as claimed.

Now let \( n > 2 \) and assume that for any collection \( B_1, \ldots, B_{n-1} \) of \( n-1 \) events with \( P(B_iB_j) = 0 \) for \( i \neq j \), we have

\[
P\left( \bigcup_{i=1}^{n-1} B_i \right) = \sum_{i=1}^{n-1} P(B_i).
\]

\[
= 1
\]
Let $A_1, \ldots, A_n$ be events such that $P(A_iA_j) = 0$ for $i \neq j$. Then, letting $A = A_1$ and $B = A_2 \cup \cdots \cup A_n$, we have

$$\bigcup_{i=1}^{n} A_i = A \cup B$$

and

$$AB = A_1(A_2 \cup \cdots \cup A_n) = A_1A_2 \cup A_1A_3 \cup \cdots \cup A_1A_n$$

so that by the union bound (Problem 4)

$$P(AB) \leq \sum_{i=2}^{n} P(A_1A_i) = 0.$$ 

Thus from (*) we obtain

$$P\left( \bigcup_{i=1}^{n} A_i \right) = P(A \cup B) = P(A) + P(B) = P(A_1) + P(A_2 \cup \cdots \cup A_n).$$

Letting $B_i = A_{i+1}$ for $i = 1, \ldots, n - 1$, the induction hypothesis (**) gives

$$P(A_2 \cup \cdots \cup A_n) = P(A_2) + \cdots + P(A_n).$$

Combing the last two equations, we get

$$P\left( \bigcup_{i=1}^{n} A_i \right) = \sum_{i=1}^{n} P(A_i)$$

as claimed.

(b) Now define the events $A_i$ by $A_i = AB_i$, $i = 1, \ldots, n$. Then $A_iA_j = (AB_i)(AB_j) = AB_iB_j \subset B_iB_j$, implying $P(A_iA_j) = 0$ for $i \neq j$ since $P(A_iA_j) \leq P(B_iB_j) = 0$ if $i \neq j$. This means that the events $AB_1, \ldots, AB_n$ satisfy therequirement of the Assignment 1 Bonus Problem, so

$$P\left( \bigcup_{i=1}^{n} AB_i \right) = \sum_{i=1}^{n} P(AB_i). \quad (*)$$

Then, on the one hand, the multiplication rule implies

$$\sum_{i=1}^{n} P(AB_i) = \sum_{i=1}^{n} P(A|B_i)P(B_i). \quad (**)$$

On the other hand, notice that $\bigcup_{i=1}^{n} AB_i = A(\bigcup_{i=1}^{n} B_i)$. Letting $B = \bigcup_{i=1}^{n} B_i$, we have

$$P(A \cup B) = P(A) + P(B) - P(AB).$$
But $P(A \cup B) \geq P(B) = 1$, so $P(A \cup B) = 1$, which, when plugged into the above equation, gives

$$P(A) = P(AB) = P\left(\bigcup_{i=1}^{n} AB_i\right).$$

Combining this with (*) and (**) proves that claim that

$$P(A) = \sum_{i=1}^{n} P(A|B_i)P(B_i).$$