Problem 1

狼来了：从前有个放羊娃，每天都把羊群带到山上去吃草, 山里有狼出没．第一天，放羊娃觉得无聊，想要作弄山下耕作的村民。他朝着山下大喊“狼来了!狼来了”，村民们信以为真，冲上山来准备帮助他，发现被欺骗了，大家很生气。第二天，放羊娃故技重施，村民们虽然有点迟疑，但还是冲上山来准备打狼，结果又一次发现被欺骗了，大家非常生气. 第三天，狼真的来了，此时放羊娃慌了，哭着向山下大喊“狼来了!狼来了!”，请求村民的帮助．但这一次村民们认为他又在撒谎，无人相信他。最后他所有的羊都被狼吃掉了。

Solution

It is denoted that event A: there come wolves; event B: the boy yells. After the boy plays with farmers, the farmers know that $P(B|A^C)$ is high. With the Bayesian formula, we have

$$P(A^C|B) = \frac{P(B|A^C)P(A^C)}{P(B)}. \tag{1}$$

With $P(A)$ and $P(B)$ fixed, this probability increases with $P(B|A^C)$, so the farmers have a higher probability of not trusting the boy.

Problem 2

A fair die is rolled repeatedly, and a running total is kept (which is, at each time, the total of all the rolls up until that time). Let p_n be the probability that the running total is ever exactly n (assume the die will always be rolled enough times so that the running total will eventually exceed n, but it may or may not ever equal n).

(a) Write down a recursive equation for p_n (relating p_n to earlier terms p_k in a simple way). Your equation should be true for all positive integers n, so give a definition of p_0 and p_k for $k < 0$ so that the recursive equation is true for small values of n.

(b) Find p_7.

(c) Give an intuitive explanation for the fact that $p_n \quad 1/3.5 = 2/7$ as $n \to \infty$.

Solution

(a) For an arbitrary integer n, it can be rolled by rolling a $(n-1)$ and a 1, or a $(n-2)$ and a 2, ..., or a $(n-6)$ and a 6, where temporarily ignore the limitation of positive integers. Thus we have

$$p_n = \frac{1}{6}(p_{n-1} + p_{n-2} + p_{n-3} + p_{n-4} + p_{n-5} + p_{n-6}) \tag{2}$$

On one hand, it is easy to find that $p_1 = 1/6$. On the other hand, $p_1 = 1/6 \cdot (p_0 + p_{-1} + p_{-2} + p_{-3} + p_{-4} + p_{-5})$. Consider practical scenarios, where a 0 can always be rolled without rolling it, and a negative number can never be rolled. Thereby, we have

$$\begin{cases} p_0 = 1, \\ p_i = 0, \quad i = -1, \dots, -5. \end{cases} \tag{3}$$

(b) Since

$$p_7 = \frac{1}{6}(p_6 + p_5 + p_4 + p_3 + p_2 + p_1) \tag{4}$$

We have

$$\begin{aligned}
p_1 &= \frac{1}{6} * p_0 = \frac{1}{6} \\
p_2 &= \frac{1}{6} * (p_0 + p_1) = \frac{1}{6} * (1 + \frac{1}{6}) = \frac{7}{36} \\
p_3 &= \frac{1}{6} * (p_0 + p_1 + p_2) = \frac{1}{6} * (1 + \frac{1}{6} + \frac{7}{36}) = \frac{49}{216} \\
p_4 &= \frac{1}{6} * (p_0 + p_1 + p_2 + p_3) = \frac{1}{6} * (1 + \frac{1}{6} + \frac{7}{36} + \frac{49}{216}) = \frac{343}{1296} \\
p_5 &= \frac{1}{6} * (p_0 + p_1 + p_2 + p_3 + p_4) = \frac{1}{6} * (1 + \frac{1}{6} + \frac{7}{36} + \frac{49}{216} + \frac{343}{1296}) = \frac{2401}{7776} \\
p_6 &= \frac{1}{6} * (p_0 + p_1 + p_2 + p_3 + p_4 + p_5) = \frac{1}{6} * (1 + \frac{1}{6} + \frac{7}{36} + \frac{49}{216} + \frac{343}{1296} + \frac{2401}{7776}) = \frac{16807}{46656}
\end{aligned}$$

Then, we have

$$p_7 = \frac{70993}{279936}$$

(c) As $n \to +\infty$, the gap of rolling different numbers is negligible, where the expectation for each "increment" is given by $1/6 \cdot (1+2+3+4+5+6) = 7/2$, so the probability for rolling a certain number is given by $p_n \approx 1/(7/2) = 2/7$.

Problem 3

A sequence of $n \geq 1$ independent trials is performed, where each trial ends in “success” or “failure” (but not both). Let p_i be the probability of success in the i^{th} trial, $q_i = 1 - p_i$, and $b_i = q_i - 1/2$, for $i = 1, 2, ..., n$. Let A_n be the event that the number of successful trials is even.
(a) Show that for $n = 2$, $P(A_2) = 1/2 + 2b_1b_2$.
(b) Show by induction that $P(A_n) = 1/2 + 2^{n-1}b_1b_2...b_n$ (This result is very useful in cryptography. Also, note that it implies that if n coins are flipped, then the probability of an even number of Heads is 1/2 if and only if at least one of the coins is fair.) Hint: Group some trials into a super-trial.
(c) Check directly that the result of (b) is true in the following simple cases: $p_i = 1/2$ for some i; $p_i = 0$ for all i; $p_i = 1$ for all i.
solution:

(a) Suppose, S is the trail ends ”success”, F is the trail ends ”fail”. We have $P(A_2) = P(S, S) + P(F, F)$.

$$P(A_2) = \ q_1 \cdot q_2 + p_1 \cdot p_2.$$

According to the equation, it is easy to get $q_i = b_i + 1/2$, $p_i = 1/2 - b_i$.Thus,

$$\begin{aligned} P(A_2) = & \ (b_1 + 1/2) \cdot (b_2 + 1/2) + (1/2 - b_1) \cdot (1/2 - b_2) \\ = & \ 1/2 + 2b_1b_2. \end{aligned}$$

(b)

$$P(A_1) = q_1 = 1/2 + b_1$$

$$P(A_2) = 1/2 + 2b_1b_2$$

Assume $P(A_{n-1}) = 1/2 + 2^{n-2}b_1b_2...b_{n-1}$, $n \geq 2$.

$$\begin{aligned} P(A_n) &= P(A_{n-1})q_n + [1 - P(A_{n-1})]p_n \\ &= (1/2 + 2^{n-2}b_1 \cdots b_{n-1})(1/2 + b_n) \\ &\ + (1/2 - 2^{n-2}b_1 \cdots b_{n-1})(1/2 - b_n) \\ &= 1/2 + 2^{n-1}b_1 \cdots b_n \end{aligned}$$

(c)

- $p_i = 1/2 = q_i$, $b_i = 0$, for all i, which means the probability of success or failure is the same. So, $P(A_n) = 1/2 = 1/2 + 0$, True.

- $p_i = 0$, $b_i = 1/2$ for all i, which means the probability of success is 0, $A_n = 0$. So, $P(A_n) = 1 = 1/2 + 2^{n-1}(1/2)^n$, True.

- $p_i = 1$, $b_i = -1/2$, for all i, which means the probability of success is 1, $A_n = n$. So, when n is even, $P(A_n) = 1 = 1/2 + 2^{n-1}(-1/2)^n$. When n is odd, $P(A_n) = 0 = 1/2 + 2^{n-1}(-1/2)^n$, True.

Problem 4

A message is sent over a noisy channel.The message is a sequence $x_1, x_2, ..., x_n$ of n bits ($x_i \in \{0, 1\}$). Since the channel is noisy, there is a chance that any bit might be corrupted, resulting in an error ($a_0$becomes a_1 or vice versa). Assume that the error events are independent. Let p be the probability that an individual bit has an error($0 < p < 1/2$).Let $y_1, y_2, ..., y_n$ be the received message (so $y_i = x_i$ if there is no error in that bit,but $y_i = 1 - x_i$ if there is an error there).
To help detect errors, the n th bit is reserved for a parity check:x_n is defined to be 0 if $x_1 + x_2 + ... + x_{n-1}$ is even, and 1 if $x_1 + x_2 + ... + x_{n-1}$ is odd. When the message is received, the recipient checks whether y_n has the same parity as $y_1 + y_2 + ... + y_{n_1}$. If the parity is wrong, the recipient knows that at least one error occurred; otherwise, the recipient assumes that there were no errors.
(a) For $n = 5, p = 0.1$, what is the probability that the received message has errors which go undetected?
(b) For general n and p, write down an expression (as a sum) for the probability that the received message has errors which go undetected.
(c) Give a simplified expression, not involving a sum of a large number of terms, for the probability that the received message has errors which go undetected.

solution:

(a) If the error message is undetected, the total number of error bits should be even: if the number of error bits in $x_1, x_2, \ldots, x_{n-1}$ is even(> 0), then the last bit x_n should be right; if the number of error bits in $x_1, x_2, \ldots, x_{n-1}$ is odd, then the last bit x_n should be wrong.

$$\begin{aligned} P(\text{undetected error message}) &= \binom{5}{4} p^4(1-p) + \binom{5}{2} p^2(1-p)^3 \\ &= 0.07335 \end{aligned}$$

(b) According to (a), we have

$$P(\text{undetected error message}) = \sum_{k \text{ is even}}^{n} \binom{n}{k} p^k (1-p)^{n-k}$$

(c)

$$\begin{aligned} P(\text{undetected error message}) &= \frac{(p + (1-p))^n + (-1)^n (p - (1-p))^n}{2} - (1-p)^n \\ &= \frac{1 + (1-2p)^n}{2} - (1-p)^n \end{aligned}$$

Problem 5

For x and y binary digits (0 or 1), let $x \oplus y$ be 0 if $x = y$ and 1 if $x \neq y$ (this operation is called exclusive or (often abbreviated to XOR), or addition mod 2).
(a) Let $X \sim Bern(p)$ and $Y \sim Bern(1/2)$, independently. What is the distribution of $X \oplus Y$
(b) With notation as in sub-problem(a), is $X \oplus Y$ independent of X? Is $X \oplus Y$ independent of Y? Be sure to consider both the case $p = 1/2$ and the case $p \neq 1/2$.
(c) Let $X_1, ..., X_n$ be i.i.d. (i.e., independent and identically distributed) Bern(1/2) R.V.s. For each nonempty subset J of $\{1, 2, ..., n\}$, let

$$Y_J = \oplus_{Y \in J} X_J.$$

Show that Y_J $Bern(1/2)$ and that these $2^n - 1$ R.V.s are pairwise independent, but not independent.
Solution:

1. Let $Z = X \oplus Y$. When $Z = 1$ is the same as $X = 1, Y = 0$ or $X = 0, Y = 1$, also $X \sim Bern(p)$ and $Y \sim Bern(1/2)$, thus we can get:

$$\begin{aligned} p(Z = 1) = & \ p(X = 1, Y = 0) + p(X = 0, Y = 1) \\ = & \ p(X = 1)p(Y = 0) + +p(X = 0)p(Y = 1) \\ = & \ p * 1/2 + (1 - p) * 1/2 \\ = & \ 1/2. \end{aligned}$$

   For the same reason, we can get $p(Z = 0) = 1/2$. Therefor, $Z \sim Bern(1/2)$.

2. To show whether Z is independent of X, it is the same to verify whether:

$$p(Z = z, X = x) = p(Z = z)p(X = x)$$

   Let's consider the left side respectively:

$$p(Z = 0, X = 1) = p(Y = 1, X = 1) = \frac{1}{2}p,$$

$$p(Z = 0, X = 0) = p(Y = 0, X = 0) = \frac{1}{2}(1 - p),$$

$$p(Z = 1, X = 1) = p(Y = 0, X = 1) = \frac{1}{2}p,$$

$$p(Z = 1, X = 0) = p(Y = 1, X = 0) = \frac{1}{2}(1 - p).$$

   Then consider the right side:

$$p(Z = 0)p(X = 1) = \frac{1}{2}p,$$

$$p(Z = 0)p(X = 0) = \frac{1}{2}(1 - p),$$

$$p(Z = 1)p(X = 1) = \frac{1}{2}p,$$

$$p(Z = 1)p(X = 0) = \frac{1}{2}(1 - p),$$

   No matter the value of p, the equation above is always true. Thus, Z is independent of X for all the case.

   To show whether Z is independent of Y, it is very similar to the above. For the left side:

$$p(Z = 0, Y = 1) = \frac{1}{2}p,$$

Problem 5 continued on next page. . .

$$p(Z=0,Y=0)=\frac{1}{2}(1-p),$$

$$p(Z=1,Y=1)=\frac{1}{2}(1-p),$$

$$p(Z=1,Y=0)=\frac{1}{2}p.$$

For the right side:

$$p(Z=z)p(Y=Y)=\frac{1}{2}*\frac{1}{2}=\frac{1}{4}.$$

To make sure the equation is always true, we should guarantee $\frac{1}{2}p=\frac{1}{2}(1-p)=\frac{1}{4}$ for all time. This is true only when $p=\frac{1}{2}$.Thus, only when $p=\frac{1}{2}$, Z is independent of Y.

3. Let l denotes the length of subset J. Then use Mathematical induction to prove the equation. As we know:
When $l=1$, $p(Y_J=1)=p(X_J=1)=\frac{1}{2}$, $Y_J\sim Bern(1/2)$.
Suppose $l=k$, $Y_J\sim Bern(1/2)$.
Then when $l=k+1$, let $\hat{J}=J\cup\{\hat{j}\}$, where the length of J is k, Thus:

$$\begin{aligned}p(Y_{\hat{J}}=1)=&\ p(Y_J=1,X_{\hat{j}}=0)+p(Y_J=0,X_{\hat{j}}=1)\\=&\ \frac{1}{2}\cdot\frac{1}{2}+\frac{1}{2}\cdot\frac{1}{2}\\=&\ 1/2.\end{aligned}$$

Therefore, $Y_J\sim Bern(1/2)$.
To show that they are pairwise independent,but not independent, is equal to verify:

$$p(Y_mY_n)=p(Y_m)p(Y_n),\forall m,n,$$

$$p(Y_1Y_2...Y_{2^n-1})\neq p(Y_1)p(Y_2)...p(Y_{2^n-1}).$$

For the first equation, there are two occasions.
$Y_m\cap Y_n=\emptyset$:
as $X_1,X_2...,X_n$ are IID distribution, Y_m and Y_n are obviouly independent.
$Y_m\cap Y_n\neq\emptyset$:
let $p=Y_m\cap Y_n$, $s=Y_m-Y_m\cap Y_n$, $q=Y_n-Y_m\cap Y_n$,

$$\begin{aligned}P(Y_m=1,Y_n=1)=&\ =P(p=1)p(s=0)p(q=0)+P(p=0)p(s=1)p(q=1)\\=&\ \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}+\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\\=&\ 1/4\\=&\ P(Y_m=1)P(Y_n=1).\end{aligned}$$

Thus, they are pairwise independent.

For the second equation, we can list a counterexample:

$$P(Y_1,Y_2,Y_3)=\frac{1}{4},$$

$$P(Y_1)P(Y_2)P(Y_3)=\frac{1}{8},$$

where $Y_1=x_1\oplus x_2$, $Y_2=x_2\oplus x_3$, $Y_3=x_3\oplus x_1$. Therefore, they are not independent.

Problem 6

(Optional Challenging Problem) By LOTP for problems with recursive structure, we generate many difference equations.

(a) Solve the following difference equation:

$$p \cdot f_{i+1} - f_i + q \cdot f_{i-1} = -1, 1 \leq i \leq N-1$$

where $0 < p < 1, q = 1-p, N$ is a constant, $f_0 = 0, f_N = 0$.

(b) Solve the following difference equation:

$$f_{i+1} = b \cdot f_i + a \cdot f_{i-1} + h, i \geq 1.$$

where h is a constant.

(c) Solve the following difference equation:

$$f_{i+1} = b \cdot f_i + a \cdot f_{i-1} + g(i), i \geq 1.$$

where $g(i)$ is a function of i.

Solution

This problem is under the topic of inhomogeneous linear equations, which has the standard method by first getting the general solution of the homogeneous part and then linearly combining the particular solution of the inhomogeneous equation. To illustrate, we first go through the solving process of (a) as follows. Given that $q = 1-p$, the equation can be rewritten as:

$$p \cdot f_{i+1} - f_i + (1-p) \cdot f_{i-1} = -1.$$

First, we solve the homogeneous part of the equation, ignoring the constant term on the right side:

$$p \cdot f_{i+1} - f_i + (1-p) \cdot f_{i-1} = 0.$$

Assuming a solution of the form $f_i = r^i$, we substitute this into the homogeneous equation to get:

$$p \cdot r^{i+1} - r^i + (1-p) \cdot r^{i-1} = 0.$$

Dividing by r^{i-1}, we obtain a quadratic equation in terms of r:

$$p \cdot r^2 - r + (1-p) = 0.$$

The quadratic equation is:

$$pr^2 - r + (1-p) = 0.$$

Solving this quadratic equation for r gives us the characteristic roots. The solutions r of this equation are given by:

$$r = \frac{1 \pm \sqrt{1-4p(1-p)}}{2p}.$$

However, since $0 < p < 1$, the term under the square root becomes $1 - 4p(1-p) = 1 - 4p + 4p^2$, which simplifies to $(2p-1)^2$, giving us real and identical roots, meaning $r = 1/p$ or $r = 1$.

Therefore, for the homogeneous part, we have the solution in the form of

$$f_i = C_1(1/p)^i + C_2,$$

Problem 6 continued on next page...

where C_1 and C_2 are constants.
We then pick the particular solution to the inhomogeneous equation in the form of $f_i = Xi$. By plugging back this value into the difference equation, we may solve for

$$X = -\frac{1}{2p-1}$$

assuming $p \neq 0.5$. Now, we have

$$f_i = C_1(1/p)^i + C_2 - \frac{i}{2p-1}.$$

Since we have $f_0 = 0, f_N = 0$, we finally have the solution as

$$f_i = -\frac{Np^{N-i}}{(2p-1)(p^N-1)} + \frac{Np^N}{(2p-1)(p^N-1)} - \frac{i}{2p-1}.$$

For (b)(c), we have their homogeneous part as follows:

$$f_{i+1} = b \cdot f_i + a \cdot f_{i-1}, i \geq 1.$$

where a and b are constants, we turn to the so-called characteristic equation:

$$x^2 = bx + a.$$

If such equation has two distinct roots r_1 and r_2, then the general form of f_i is

$$f_i = C_1 \cdot r_1^i + C_2 \cdot r_2^i,$$

If there is only one distinct root r, then the general form of f_i is

$$f_i = C_1 \cdot r^i + C_2 \cdot i \cdot r^i.$$

For (b), we pick the particular solution in the form of $f_i = X$. By plugging back this value into the difference equation, we may solve for

$$X = \frac{h}{1-a-b}.$$

For (c), we must assume specific forms of $g(i)$ to work. For example, suppose $g(i) = e^i$ and we pick the particular solution in the form of $f_i = Xe^i$. Therefore, we again solve for

$$X = \frac{e}{e^2 - be - a}.$$