1. Set $Y = (X + Z) \mod 11$, Z is independent of X and has pmf $p_Z(i) = 3^{-1}$ for $i \in \{1, 2, 3\}$. Consider a DMC with $\mathcal{X} = \mathcal{Y} = \{0, 1, \ldots, 10\}$ and $M = (\mathbb{P}(Y = y|X = x))_{x\in\mathcal{X}, y\in\mathcal{Y}}$. Find the capacity of this channel and the distribution of X that achieves capacity.

2. Consider a DMC $|\mathcal{X}| = |\mathcal{Y}| = 3$ with stochastic matrix

$$M = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{2}{3} \end{pmatrix}.$$

   (a) Calculate the capacity of this channel,

   (b) Give an intuitive argument why the capacity is achieved with a distribution that places zero probablity on an input symbol.

3. **(Two independent looks at Y)** Let $\mathcal{X}$ and $\mathcal{Y}$ be finite sets, X be any r.v. on $\mathcal{X}$, and Y_1 and Y_2 be r.v. on $\mathcal{Y}$ which are, conditionally on X, i.i.d.

   (a) Show that $I(X; Y_1, Y_2) = 2I(X; Y_1) - I(Y_1; Y_2)$.

   (b) Consider two DMCs of which (X, Y_1) and $\Big(X, (Y_1, Y_2)\Big)$ are realisations. Prove that the capacity of the second DMC is at most twice that of the first.

4. **(Time varying channel)** Let $\mathcal{X} = \mathcal{Y} = \{0, 1\}$ and for each time $i \in \{1, \ldots, n\}$ we can use a DMC

$\mathcal{X}\backslash\mathcal{Y}$	0	1
0	$1 - q_i$	q_i
1	q_i	$1 - q_i$

   to transmit a symbol. This is an example of a time-varying discrete memoryless channel. Let $\mathbf{X} = (X_1, \ldots, X_n)$, $\mathbf{Y} = (Y_1, \ldots, Y_n)$ with conditional pmf $p(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^n p_i(y_i|x_i)$ where p_i is the conditional distribution of above symmetric binary noisy channel ($p_i(0|0) = p_i(1|1) = 1 - q_i$). Calculate $\max_{p_\mathbf{X}} I(\mathbf{X}; \mathbf{Y})$ (subject to the usual constraint that $\mathbf{Y}|\mathbf{X} \sim p(\mathbf{y}|\mathbf{x})$).

5. **(Hamming code)** Consider the binary symmetric channel, i.e. $\mathcal{X} = \mathcal{Y} = \{0, 1\}$ and

$\mathcal{X}\backslash\mathcal{Y}$	0	1
0	$1 - q$	q
1	q	$1 - q$

   Let $i \in \{1, \ldots, 16\}$ define an encoder $c(i) = (s_1, s_2, s_3, s_4, p_1, p_2, p_3) \in \mathcal{Y}^7$ by letting $s_1s_2s_3s_4$ be the binary expansion of $i - 1$ and p_1, p_2, p_3 be parity bits defined by $p_1 := s_1 \oplus s_2 \oplus s_3$, $p_2 := s_2 \oplus s_3 \oplus s_4$, $p_3 := s_1 \oplus s_3 \oplus s_4$ where $\oplus : \{0, 1\} \to \{0, 1\}$ denotes sum modulo 2. Examples: $c(2) = 0001011$ since $s_1s_2s_3s_3 = 0001$, $c(5) = 0100110$ since $s_1s_2s_3s_4 = 0100$).

   (a) Visualize this by drawing three intersecting circles. Put the first four bits into the regions intersecting at least two of these circles, and the parity bits in the remaining regions. Arrange the positions such that the sum of the four bits within each circle is even. Use this to find a good decoder $d : \mathcal{Y}^7 \to \{1, \ldots, 16\}$, which will flip the minimal amount of bits to restore even parity within each circle.

(b) Decode the outputs 1100101, 1000001.

(c) Calculate the error probabilities of this channel code.

(d) Calculate the rate of this channel code.

6. **(Information theory and gambling)** m horses run a race, the ith horse wins with probability p_i. An investment of one pound returns $o(i)$ pounds if horse i wins, otherwise the investment is lost. A gambler distributes all of his wealth across the horses: $b(i) \geq 0$ denotes the fraction of the gambler's wealth that he bets on horse i and $\sum_{i=1}^m b(i) = 1$. We now consider repeating this game over and over. If S_n denotes the gambler's wealth after the nth race, then

$$S_n = \prod_{i=1}^n b(X_i)\, o(X_i)$$

where X_i is the horse that wins the i-th race and $s_0 \in \mathbb{R}$ is the start capital.

(a) If X_i are iid, show that for given $\mathbf{b} = (b(1), \ldots, b(m))$, $\mathbf{p} = (p_1, \ldots, p_m)$ the wealth evolves exponentially, i.e. $\lim_{n\to\infty} \frac{1}{n} \log \frac{S_n}{2^{nW(\mathbf{b},\mathbf{p})}} = 0$ almost surely, where $W(\mathbf{b}, \mathbf{p})$ is to be determined. [Hint: Strong law of large numbers]

(b) Define $W^\star(\mathbf{p}) := \max_{\mathbf{b}:\sum b(i)=1, b(i)\geq 0} W(\mathbf{b}, \mathbf{p})$ and find $\mathbf{b}$ that achieves this maximum. [Hint: You can find a candidate by using Lagrange multipliers.]

(c) (Informal.) We can regard $q_i := \frac{1}{o(i)}$ as the "probabilities" the bookmaker implicitly assigns to outcomes. Considering the cases $\sum q_i = 1$, $\sum q_i < 1$ and $\sum q_i > 1$ discuss the fairness of the game.

7. * **(Optional. Information theory and finance)** A stock market is represented as $\mathbf{X} = (X_1, \ldots, X_m)$ where each random variable X_i is non-negative and represents the ratio of prices for stock at i at the end of the day to the beginning of the day (e.g. $\{X_i = 1.03\}$ is the event that stock i went up 3 percent). A portfolio $\mathbf{b} = (b(1), \ldots, b(m))$ consists of numbers $b(i) \geq 0$, $\sum_{i=1}^m b(i) = 1$, where $b(i)$ denotes the fraction the investor's wealth that is invested in stock i. Hence, using a portfolio $\mathbf{b}$ on the stock market $\mathbf{X}$, leads to a relative wealth change of $S = \mathbf{b}^T \mathbf{X} = \sum_{i=1}^m b_i X_i$. The wealth change after n trading days using the same portfolio $\mathbf{b}$ is therefore $S_n = \prod_{i=1}^n \mathbf{b}^T \mathbf{X}_i$.

(a) If $\mathbf{X}_1, \ldots, \mathbf{X}_n$ are iid with cdf F, show that for given $\mathbf{b}$, $\lim_{n\to\infty} \frac{1}{n} \log \frac{S_n}{2^{nW(\mathbf{b},F)}} = 0$, where $W(\mathbf{b}, F)$ is to be determined.

(b) Show that $W(\mathbf{b}, F)$ is concave in $\mathbf{b}$: for $t \in [0, 1]$ and two portfolios $\mathbf{b}_1$ and $\mathbf{b}_2$,

$$W(t\mathbf{b}_1 + (1-t)\mathbf{b}_2, F) \geq tW(\mathbf{b}_1, F) + (1-t)W(\mathbf{b}_2, F).$$

Show that it is "linear" in F, in the sense that for $t \in [0, 1]$ and two cdf F_1 and F_2,

$$W(\mathbf{b}, tF_1 + (1-t)F_2) = tW(\mathbf{b}, F_1) + (1-t)W(\mathbf{b}, F_2).$$

Finally, show that $W^\star(F) = \max_{\mathbf{b}} W(\mathbf{b}, F)$ is convex in F:

$$W^\star(tF_1 + (1-t)F_2) \leq tW^\star(F_1) + (1-t)W^\star(F_2).$$

($\mathbf{b}$ that achieves this maximum is called a *growth optimal portfolio*).

(c) Show that the set of growth optimal portfolios (with respect to F) is a convex set.

8. * **(Optional. Hamming code and finite fields)** Let $\mathbb{F}_2 = \{0,1\}$ and define the usual modulo 2 arithmetic on $\mathbb{F}_2$ ($0+0 = 1+1 = 0$, $0+1 = 1+0 = 1$, $0\cdot 0 = 0\cdot 1 = 1\cdot 0 = 0$, $1\cdot 1 = 0$). We recall that this makes $(\mathbb{F}_2, +, \cdot)$ into a field, and that $\mathbb{F}_2^n = \{0,1\}^n$ is the canonical n-dimensional vector space over this field.

   (a) A linear code is a channel code with a codebook that is a linear subspace $\mathbb{F}_2^n$. Consider the Hamming code from Example 5 and the *generator matrix*

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix}.$$

   Use G to show that the Hamming code is a linear code [Hint: multiply with 0000, 0001, 0010,...].

   (b) Define P as $\begin{pmatrix} I_4 \\ P \end{pmatrix} := G$ and set $H = (P, I_3)$ (I_n is the $n \times n$ identity matrix over $\mathbb{F}_2$). Show that all codewords are in the kernel of H (reminder: the kernel is the set of all column vectors x such that Hx is the zero vector). We call H the *parity matrix*.