1. Let $\mathcal{X}, \mathcal{Y}$ be finite sets and X a $\mathcal{X}$-valued random variable.

   (a) Show that for any instantaneous code $c : \mathcal{X} \to \mathcal{Y}^\star$, only finitely many instantaneous codes $c' : \mathcal{X} \to \mathcal{Y}^\star$ exist such that $\mathbb{E}[|c'(X)|] \leq \mathbb{E}[|c(X)|]$.

   (b) Conclude that an optimal code always exists.

2. **(Fano's inequality)** Let X,Y be discrete random variables that take values in a finite state space $\mathcal{X}$.

   (a) Show that $H(X|Y) \leq H(1_{X \neq Y}) + \mathbb{P}(X \neq Y)(\log|\mathcal{X}| - 1)$.
   [Hint: $H(X|Y) = H(X|Y) + H(1_{X\neq Y}|X,Y) = H(X, 1_{X\neq Y}|Y) = H(1_{X\neq Y}|Y) + H(X|Y, 1_{X\neq Y}) \leq \ldots$].

   (b) Show that $H(X|Y) < 1 + \mathbb{P}(X \neq Y)\log|\mathcal{X}|$,

   (c) Use the above (as in the proof of Shannon's NCT) to derive a lower bound for the arithmetic error $\overline{\epsilon}$ of a channel code (c, d) with rate $\rho(c, d) > C$. Plot how this bound varies with the rate.

3. 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.

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 $\boldsymbol{X} = (X_1, \ldots, X_n)$, $\boldsymbol{Y} = (Y_1, \ldots, Y_n)$ with conditional pmf $p(\boldsymbol{y}|\boldsymbol{x}) = \prod_{i=1}^n p_i(y_i|x_i)$ 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 $\max_{p_{\boldsymbol{X}}} \mathrm{I}(\boldsymbol{X}; \boldsymbol{Y})$ (subject to the usual constraint that $\boldsymbol{Y}|\boldsymbol{X} \sim p(\boldsymbol{y}|\boldsymbol{x})$).

5. **(Hamming code)** Consider the binary noisy 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 $s_1 s_2 s_3 s_4$ be the binary expansion of i and $p_1 := s_1 \oplus s_2 \oplus s_3$, $p_2 := s_2 \oplus s_3 \oplus s_4$, $p_3 := s_1 \oplus s_3 \oplus s_4$ where $\oplus : \{0, 1\} \to \{0, 1\}$ denotes exclusive OR ($a \oplus b = 1$ iff $a \neq b$; e.g. $c(1) = 0001011$ since $s_1 s_2 s_3 s_3 = 0001$, $c(4) = 0100110$ since $s_1 s_2 s_3 s_4 = 0100$). We call p_1, p_2, p_3 parity bits (they show if the sum of bits is even or odd).

   (a) Visualize this by drawing three intersecting circles [Hint: put the first four bits into regions intersecting at least two of these circles. Put the parity bits in the remaining regions]. Use this, to find a good decoder $d : \mathcal{Y}^7 \to \{1, \ldots, 16\}$,

   (b) Decode the outputs 1100101, 1000001,

   (c) Calculate the rate of this channel code.

6. **(Hamming code and finite fields)** Let $\mathbb{F}_2 = \{0,1\}$ and define the usual modulo 2 arithmetic on $\mathbb{F}_2$ $(0+0=1+1=0,\ 0+1=1+0=1,\ 0\cdot 0=0\cdot 1=1\cdot 0=0,\ 1\cdot 1=0)$.

   (a) Show that $(\mathbb{F}_2,+,\cdot)$ is a field, and describe how $\mathbb{F}_2^n = \{0,1\}^n$ can be seen as a vector space over the field $\mathbb{F}_2$,

   (b) 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*

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

   Use G to show that the Hamming code is a linear code [Hint: multiply with 0000, 0001, 0010,...]. Define P as $\begin{pmatrix} I_4 \\ P \end{pmatrix} := G^T$ 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. We call H the *parity matrix.*

7. 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 $\boldsymbol{b} = (b(1), \ldots, b(m))$, $\boldsymbol{p} = (p_1, \ldots, p_m)$ the wealth grows exponentially, i.e. $\lim_{n\to\infty} \frac{1}{n} \log \frac{S_n}{2^{nW(\boldsymbol{b},\boldsymbol{p})}} = 0$, where $W(\boldsymbol{b},\boldsymbol{p})$ is to be determined. [Hint: Weak law of large numbers]

   (b) Define $W^\star(\boldsymbol{p}) := \max_{\boldsymbol{b}:\sum b(i)=1, b(i)\geq 0} W(\boldsymbol{b},\boldsymbol{p})$ and find $\boldsymbol{b}$ that achieves this maximum. [Hint: Once you found an extremum, express $W(\boldsymbol{b},\boldsymbol{p})$ using $H(\boldsymbol{p})$ and $D(\boldsymbol{p} \parallel \boldsymbol{b})$ to verify that it is a maximum]

   (c) We can regard $q(i) := \frac{1}{o(i)}$ as the "probabilities" the bookmaker implicitily assigns to outcomes. Consider the cases $\sum q_i = 1$, $\sum q_i < 1$ and $\sum q_i > 1$ and argue which is a fair game, which favours the gambler, which favours the bookmaker?

8. A stock market is represented as $\boldsymbol{X} = (X_1, \ldots, X_m)$ 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 3percent). A portfolio $\boldsymbol{b} = (b(1), \ldots, b(m))$ 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 $\boldsymbol{b}$ on the stock market $\boldsymbol{X}$, leads to a relative wealth change of $S = \boldsymbol{b}^T\boldsymbol{X} = \sum_{i=1}^m b_i X_i$. The wealth change after n trading days using the same portfolio $\boldsymbol{b}$ is therefore $S_n = \prod_{i=1}^n \boldsymbol{b}^T\boldsymbol{X}_i$.

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

(b) Show that $W(\boldsymbol{b}, F)$ is concave in $\boldsymbol{b}$ and linear in F. Show that $W^\star(F) = \max_{\boldsymbol{b}} W(\boldsymbol{b}, F)$ is convex in F. [$\boldsymbol{b}$ that achieves this maximum is called a *growth optimal portfolio*].

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