1. Fix $p > 0$. We are given a fair coin and want to generate independent samples from a Bernoulli random variable $\mathbb{P}(X = 1) = p$, $\mathbb{P}(X = 0) = 1 - p$. Find an algorithm that does this, such that the expected number of needed coin flips to generate one sample of X is less or equal than 2.

2. Let $q \in [0,1]$, $n \in \mathbb{N}$ such that nq is an integer in the range $[0,n]$. Show that

$$\frac{2^{nH(q)}}{n+1} \leq \binom{n}{nq} \leq 2^{nH(q)}$$

where $H(q) := -q \log q - (1-q)\log(1-q)$ is the entropy of a Bernoulli distributed random variable.

3. Let X_1 be a $\mathcal{X}_1 = \{1,\ldots,m\}$ valued random variable and X_2 be a $\mathcal{X}_2 = \{m+1,\ldots,n\}$-valued random variable. Further assume X_1 and X_2 to be independent. Define a random variable X as

$$X = X_\theta$$

where θ is random variable such that $\mathbb{P}(\theta = 1) = \alpha, \mathbb{P}(\theta = 2) = 1 - \alpha$ for some $\alpha \in [0,1]$ and θ is independent of X_1 and independent of X_2.

   (a) Express $H(X)$ as a function of $H(X_1)$,$H(X_2)$,$H(\theta)$ and α.

   (b) Show that $2^{H(X)} \leq 2^{H(X_1)} + 2^{H(X_2)}$. For which α does this become an equality?

4. The *differential entropy* of a $\mathbb{R}^n$-valued random variable X with density f is defined as

$$h(X) := -\int f(x) \log f(x)\, dx$$

(with the integration over the support of f). Calculate $h(X)$ when

   (a) X is uniformly distributed on $[0,1]$,

   (b) X is standard normal distributed,

   (c) X is exponential distributed with parameter λ.

5. Let X be a $\mathbb{R}^n$-valued random variable with zero mean and covariance matrix Σ. Show that

$$h(X) \leq \frac{1}{2} \log (2\pi e)^n |\Sigma|$$

with equality iff X is multivariate normal.

6. A Markov chain is a sequence of discrete random variables $(X_n)_{n\geq 1}$ such that for all $x_1,\ldots,x_{n+1} \in \mathcal{X}$

$$\mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n, \ldots, X_1 = x) = \mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n).$$

The chain is called homogenous if $p_n(x,y) := \mathbb{P}(X_{n+1} = y | X_n = x)$ does not dependend on n (for every $x, y \in \mathcal{X}$). In this case we call $(p(x,y))_{x,y\in\mathcal{X}}$ the transition matrix of (X_n). A fair die is rolled repeatedly. Which of the following are Markov chains? For those that are, give the transition matrix.

   (a) X_n is the largest roll up to the nth roll,

   (b) X_n is the number of sixes in n rolls,

   (c) X_n is the number of rolls since the most recent six,

(d) X_n is the time until the next six.

7. Let (X_n) be a Markov chain. Which of the following are Markov chains?

   (a) $(X_{m+n})_{n\geq 1}$ for a fixed $m \geq 0$,

   (b) $(X_{2n})_{n\geq 1}$,

   (c) $(Y_n)_{n\geq 1}$ with $Y_n := (X_n, X_{n+1})$.

8. Prove the strong AEP: denote with $\mathcal{S}^n_\epsilon$ the smallest subset of $\mathcal{X}^n$ such that $\mathbb{P}\left(\boldsymbol{X} \in \mathcal{S}^\epsilon_n\right) \geq 1 - \epsilon$ where $\boldsymbol{X} = (X_1, \ldots, X_n)$ are iid copies of a $\mathcal{X}$-valued rv X. Then for any sequence (ϵ_n) with $\lim_{n\to\infty} \epsilon_n = 0$ we have

$$\lim_{n\to\infty} \frac{1}{n} \log \frac{\left|\mathcal{S}^{\epsilon_n}_n\right|}{\left|\mathcal{T}^{\epsilon_n}_n\right|} = 0.$$

   [Hint: show that $\mathbb{P}(A \cap B) > 1 - \epsilon_1 - \epsilon_2$ for any sets with $\mathbb{P}(\boldsymbol{X} \in A) > 1 - \epsilon_1$,$\mathbb{P}(\boldsymbol{X} \in B) > 1 - \epsilon_2$ and use this to estimate $\mathbb{P}\left(\mathcal{S}^\epsilon_n \cap \mathcal{T}^\epsilon_n\right)$]