Probability Exercise, May 11th

1. Consider a Bernouilli random variable X defined by the parameter p: $p(X = 1) = p$ and $p(X = 0) = 1 - p$
   - The Shannon entropy of a discrete random variable X that takes n possible values i for $i = 1, \ldots, n$ each with probability p_i , is defined as
     $$H = -\sum_{i=1}^{n} p_i \log p_i$$
     This entropy is supposed to measure the "randomness" of a discrete random variable.
     Calculate the Shannon entropy of $b(x)$ as a function of p.
   - What is the value of p that maximizes (resp. minimizes) the Shannon entropy for a Bernouilli distribution?
   - Interpret this result.

2. The Shannon entropy of a continuous random variable associated with the probability density function $p(x)$ is defined as:
   $$H = -\int_{-\infty}^{+\infty} p(x) \log p(x) \mathrm{dx}$$
   - This entropy is supposed to measure the "randomness" of a continuous random variable.
     Calculate the Shannon entropy of a Gaussian distribution.
   - When is this entropy maximum?
   - Interpret this result

3. Calculate the Kullback-Leibner (KL) divergence between two Gaussians $N(x; \mu_1, \sigma_1^2)$ and $N(x; \mu_2, \sigma_2^2)$.

4. Suppose that z is an n-dimensional standardized multivariate Gaussian variable: $N(z; 0, I)$.
   Suppose that the n-dimensional positive definite matrix Σ has the Cholesky decomposition: $\Sigma = LL^T$, where L is a lower-triangular matrix. Show that the random vector $y = \mu + Lz$ follows a multivariate Gaussian distribution. What is its mean and what is its variance-covariance matrix? This is a commonly-used algorithm to generate samples of a multivariate Gaussian!