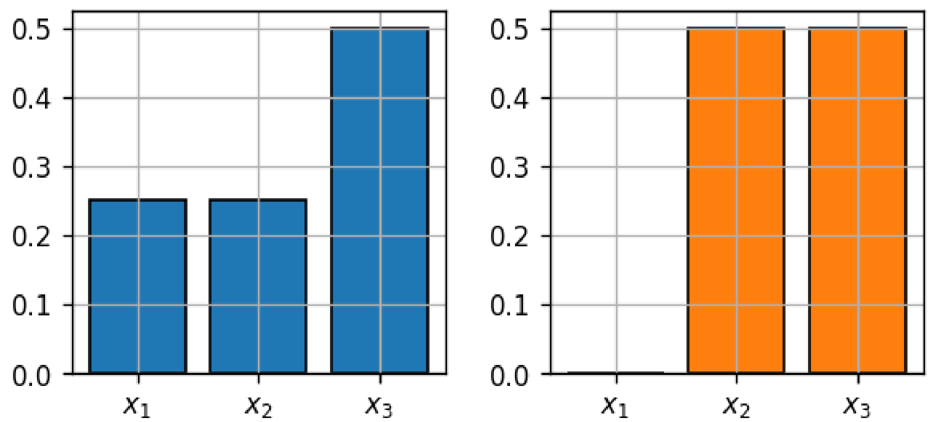
# KL Divergence

***Submit a PDF of your answers to Canvas***

**1.** Consider two discrete distributions $p$ and $q$ over $\mathscr{X} = \{x_1, x_2, x_3\}$. Distribution $p$ is shown in blue, while $q$ is shown in orange.




base2 base e

**a)** Compute $D(p||q)$ in bits and nats. not defined ($q=0$ but $p \neq 0$ for $x_1$)

**b)** Compute $D(q||p)$ in bits and nats. $= \frac{1}{2}\log\frac{1/2}{1/4} + \frac{1}{2}\log\frac{1/2}{1/2} = 0.5\log 2$

**c)** Consider a simple hypothesis test in which you observe $n$ i.i.d. samples from either $p$ or $q$, and try to discern if the samples came from distribution $p$ or $q$. You design a test that has a fixed probability of deciding $q$ if $p$ was true. What is the best *error exponent* you can expect for the probability of deciding $p$ when $q$ is true?

**2.** Consider two pdfs $p$ and $q$, with

$$p(x) = \begin{cases} x + 1/2 & \text{for } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

and

$$q(x) = \begin{cases} 1 & \text{for } 0 \leq x \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Compute $D(p||q)$ and $D(q||p)$ with the Wolfram Integrator and then with the numerical integration functions in the included .ipynb file ***(focus on the second and third cells, which show how to use the functions; don't worry about first cell that defines the functions)***.

$$D(p||q) = \int_0^1 p(x)\log\left(\frac{p(x)}{q(x)}\right)dx = \int_0^1 \left(x+\frac{1}{2}\right)\log\left(x+\frac{1}{2}\right)dx \approx 0.042$$

$$D(q||p) = \int_0^1 q(x)\log\frac{q(x)}{p(x)}dx = \int_0^1 1\log\left(\frac{1}{x+\frac{1}{2}}\right)dx \approx 0.045$$

**3.** In this problem we consider the KL divergence between both Bernoulli and binomial random variables.

**a)** Find an expression for the Kullback-Leibler divergence between two Bernoulli random variables with respective bias $\theta$ and $\theta'$.

**b)** Recall the Binomial pmf, which gives the probability of $k$ heads in $n$ coin flips:

$$p(k) = \binom{n}{k} \theta^k (1-\theta)^{n-k}$$

where $\theta$ is the bias of the coin. Find an expression for the Kullback-Leibler divergence between two binomial random variables. Simplify your expression as much as possible.

**c)** How do the expressions compare?

a) $D(P_\theta \| P_{\theta'}) = \theta \log \frac{\theta}{\theta'} + (1-\theta) \log \frac{1-\theta}{1-\theta'}$

b) $D(P_\theta \| P_{\theta'}) = \sum_k P_\theta(k) \log\left(\frac{P_\theta(k)}{P_{\theta'}(k)}\right)$

$$= \sum \binom{n}{k} \theta^k (1-\theta)^k \log\left(\left(\frac{\theta}{\theta'}\right)^k \left(\frac{1-\theta}{1-\theta'}\right)^{n-k}\right)$$

c) Bernoulli is a special case of binomial $(n = 1)$