Binary Variables

Definition. If x is a random variable and $x \in \{0, 1\}$ then x is a binary variable.

Definition. Let $P(x = 1) = \mu$. Then the *Bernoulli* distribution is given by

$$\mathrm{Bern}(x|\mu) = \mu^x (1 - \mu)^{1-x}.$$

Let $\mathcal{D} = x_1, ..., x_N$, then

$$p(\mathcal{D}|\mu) = \prod_{n=1}^{N} p(x_n|\mu)$$

which forms the likelihood function for $p(\mu|\mathcal{D})$. Then,

$$\ln p(\mathcal{D}|\mu) = (\ln \mu - (\ln(1 - \mu))) \left(\sum_{n=1}^{N} x_n \right) + N \ln(1 - \mu).$$

Since the above function depends on $\mathcal{D}$ only through the sum, $\sum_{n=1}^{N} x_n$ is called a *sufficient statistic*. Maximising the log will result in the sample mean as follows

$$\mu_{\mathrm{ML}} = \frac{1}{N} \sum_{n=1}^{N} x_n.$$

If there is a μ chance of a binary variable being 1, then we can use a binomial distribution to compute the probability of drawing m instances of 1 out of a data set of N elements.

Definition. Given a size N and m, and a mean μ, the *binomial distribution* is defined as follows

$$\mathrm{Bin}(m|N, \mu) = \binom{N}{m} \mu^m (1 - \mu)^{N-m}$$

where

$$\binom{N}{m} = \frac{N!}{m!(N - m)!}.$$

For the binomial distribution, we have the following properties

$$\begin{aligned} \mathbb{E}[m] &= N\mu \\ \mathrm{var}[m] &= N\mu(1 - \mu). \end{aligned}$$