Chapter 3 page 140 (DHS): 1, 4, 15.

Section 3.2

1. Let x have an exponential density

$$p(x|\theta) = \begin{cases} \theta e^{-\theta x} & x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Plot $p(x|\theta)$ versus x for $\theta = 1$. Plot $p(x|\theta)$ versus θ, $(0 \leq \theta \leq 5)$, for $x = 2$.

(b) Suppose that n samples $x_1, \ldots, x_n$ are drawn independently according to $p(x|\theta)$. Show that the maximum-likelihood estimate for θ is given by

$$\hat{\theta} = \frac{1}{\frac{1}{n}\sum_{k=1}^{n} x_k}.$$

(c) On your graph generated with $\theta = 1$ in part (a), mark the maximum-likelihood estimate $\hat{\theta}$ for large n.

4. Let $\mathbf{x}$ be a d-dimensional binary (0 or 1) vector with a multivariate Bernoulli distribution

$$P(\mathbf{x}|\boldsymbol{\theta}) = \prod_{i=1}^{d} \theta_i^{x_i}(1-\theta_i)^{1-x_i},$$

where $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_d)^t$ is an unknown parameter vector, θ_i being the probability that $x_i = 1$. Show that the maximum-likelihood estimate for $\boldsymbol{\theta}$ is

$$\hat{\boldsymbol{\theta}} = \frac{1}{n}\sum_{k=1}^{n} \mathbf{x}_k.$$

15. Consider the problem of learning the mean of a univariate normal distribution. Let $n_0 = \sigma^2/\sigma_0^2$ be the dogmatism, and imagine that μ_0 is formed by averaging n_0 fictitious samples x_k, $k = -n_0 + 1, -n_0 + 2, \ldots, 0$.

(a) Show that Eqs. ~~31~~ 34 and ~~32~~ 35 for μ_n and σ_n^2 yield

$$\mu_n = \frac{1}{n + n_0} \sum_{k=-n_0+1}^{n} x_k$$

and

$$\sigma_n^2 = \frac{\sigma^2}{n + n_0}.$$

(b) Use this result to give an interpretation of the prior density $p(\mu) \sim N(\mu_0, \sigma_0^2)$.