1. Plants are randomly distributed throughout an (effectively infinite) area. In other words, the number of plants in a set A is distributed Poisson($\mu|A|$), where $|A|$ is the area of A, and μ is unknown.

(a) Suppose that a plot B is extensively surveyed and k plants are found. Find the mle of μ, and the variance of this estimator.

(b) Alternatively, a surveyor walks along a transect line for a distance L, making observations to both sides (but not ahead or behind). A plant at perpendicular distance x from the transect line has probability $\exp(-\lambda x)$ of being observed. Then the number of plants found is again Poisson, but now with parameter $2\mu L/\lambda$, and the density of distance of an observed plant from the transect line is $\lambda \exp(-\lambda x)$, independently from plant to plant (you may take this as known). Suppose that the surveyor observes k plants at distances $x_1, \ldots, x_k$ from the transect line. Show that the likelihood for (μ, λ) is proportional to

$$\mu^k \exp(-\frac{2\mu L}{\lambda} - \lambda \sum_{i=1}^{k} x_i)$$

(c) Find the mle of (μ, λ), and determine the asymptotic variance of the estimator (you need not invert the matrix).

(d) Suppose that the expected number of plants are the same in both observational schemes, so that $\mu|B| = 2\mu L/\lambda$. In estimating μ, how much precision is lost, compared to the survey method, due to the necessity of estimating λ in the transect method (use 1/variance as a measure of precision).

2. (a) Suppose that $\hat{\theta}$ solves the likelihood equation. Let $\eta = q(\theta)$ be a differentiable one-to-one reparametrization of the problem. Show that $\hat{\eta} = q(\hat{\theta})$ is the mle of η.

(b) Let, again, $\eta = q(\theta)$ be a reparametrization, but only assume that q is a function (not necessarily one-to-one or differentiable). Assume that the likelihood function $L(\theta)$ is bounded for $\theta \in \Theta$. Define the likelihood function **induced by** q as

$$L^q(\eta) = \sup_{\theta: q(\theta)=\eta} L(\theta).$$

Show that $\hat{\eta} = q(\hat{\theta})$ maximizes $L^q(\eta)$ over $\eta \in H = \{q(\theta): \theta \in \Theta\}$, where $\hat{\theta}$ maximizes $L(\theta)$.

(c) Suppose that $X_1, \ldots, X_n$ are iid Bern(p). Show that the mle of $\mathbf{Var} X_1$ is $\bar{X}(1 - \bar{X})$.

3. Let $X(u) \sim \mathrm{N}(\alpha + \beta u, \sigma^2)$, where $u \in [0, 1]$, σ^2 is known, and α, β are unknown. A **design** for this problem consists of choosing numbers $u_1, \ldots, u_n$, in order to observe $X(u_1), \ldots, X(u_n)$. The statistical analysis of the data is based on the mle/least squares estimate $\hat{\beta} = \sum(u_i - \bar{u})X(u_i)/\sum(u_i - \bar{u})^2$.

(a) Determine the information matrix for (α, β) based on the sample.

(b) Assume for simplicity that n is even. Show that a design that puts $n/2$ points at 0 and $n/2$ at 1 maximizes the determinant of the information matrix, and minimizes the variance of $\hat{\beta}$.

(c) A competing design has $u_i = (i - 1)/(n - 1)$. Find the relative efficiency (in terms of a ratio of variances) of the estimate $\hat{\beta}$ from this design with respect to the design in (b).
Hint: You may want to use that $\sum_1^n i = n(n + 1)/2$, and that $\sum_1^n i^2 = n(n + 1)(2n + 1)/6$.

(d) Which design would you prefer to use, and why?