1. Mark-recapture experiments are often used by wildlife biologists to estimate the size of an animal population. In its simplest version, a random sample of k animals are captured, marked, and released. Some time later (long enough to allow the marked animals to mix with the rest of the population, but so short that births, deaths and migration may be ignored) a second random sample of size n is taken, and the number, r, of marked animals in the second sample is recorded. k and n are fixed numbers. Let N denote the unknown population size.

(a) Assuming that N is much larger than k, derive the mle of N.
Note: You may regard N as continuous, even though it is, in fact, discrete.

(b) Derive an expression for the asymptotic variance of the estimator in (a). Is this a good approximation to the true variance?

(c) What would you do if $r = 0$?

2. Let $X_1, \ldots, X_n$ be iid positive random variables with density $f(x;\theta) = \theta^{-1}\exp(-x/\theta)$ and cdf $F(x;\theta)$. Let $Y_1, \ldots, Y_n$ be iid with known density g and cdf G. Suppose that we observe $Z_i = \min(X_i, Y_i)$, and $\delta_i = 1(X_i < Y_i)$. You may want to think of the X's as survival times, and the Y's as censoring times.

(a) Show that the pairs (Z_i, δ_i) are iid with density

$$h(z,\delta;\theta) = [f(z;\theta)\bar{G}(z)]^{\delta}\,[g(z)\bar{F}(z;\theta)]^{1-\delta}$$

where δ =0 or 1 and $z > 0$, $\bar{F} = 1 - F$, and $\bar{G} = 1 - G$.

(b) Find $I(\theta)$ based on observing (Z_i, δ_i), $i = 1, \ldots, n$.

3. Due to the influence of an eccentric (but well-placed) uncle, you have become a rising star at Throgmorton Enterprises. Indeed, Mrs. Throgmorton herself has just given you an important assignment. The company is about to buy some components whose lifetimes are modeled as iid exponential rv's with mean $1/\lambda$. Your task is to estimate λ using the manufacturer's experimental records. When you arrive at the factory, however, you find that the records only show whether a part lasted longer than a fixed time t. In addition, the general untidyness of the plant makes you suspicious that some errors may have crept into the data.

(a) Find the mle $\hat{\lambda}$ of λ based on the manufacturer's records.

(b) Find the approximate distribution of this estimator for large n.

4. Let $X_1, \ldots, X_n$ be iid $f(x;\boldsymbol{\theta})$, and suppose that $Y_1, \ldots, Y_n$ are iid censoring variables, independent of the X's, and with a density that does not depend on $\boldsymbol{\theta}$. Suppose that we observe $Z_i = \min(X_i, Y_i)$ and $\delta_i = 1(X_i \le Y_i)$. Let $S(t;\boldsymbol{\theta}) = \int_t^\infty f(u;\boldsymbol{\theta})du$ be the survival function of the X's.

(a) Show that $l_n(\boldsymbol{\theta}) = \sum_{\delta_i=1} \log f(z_i,\boldsymbol{\theta}) + \sum_{\delta_i=0} \log S(z_i;\boldsymbol{\theta}) + \text{constant}$.

(b) When $f(x;\alpha,\beta) = \alpha\beta(\beta x)^{\alpha-1}\exp(-(\beta x)^{\alpha})$ (the Weibull distribution), show that the mle of β for a given α is

$$\hat{\beta}(\alpha) = \left(\frac{d}{(\sum z_i^\alpha)} \right)^{1/\alpha}$$

where $d = \sum_1^n \delta_i$.

(c) Show that $\hat{\alpha}$ satisfies the equation

$$\frac{d}{\alpha} + \sum_{\delta_i = 1} \log z_i - d \, \frac{\sum z_i^\alpha \log z_i}{\sum z_i^\alpha} = 0.$$

(d) Determine the information matrix for the Weibull case.