**Statistics 582, Problem Set 1**

Wellner; 1/3/2018

**Reading:**

- Chapter 6, Sections 4-6 and 10; Lehmann and Casella, TPE, pages 456-487 and 515-519.
- Course Notes, sections 4.4, 4.5, and 4.6.
- Groeneboom's notes on the EM algorithm.

**Due:** Wednesday, January 10, 2018.

1. Consider the Weibull family of distributions as given in Example 3.2.5 of Chapter 3 of the course notes and as in Example 5.43 of van der Vaart's *Asymptotic Statistics* page 70. Rewrite van der Vaart's example in terms of the parametrization given in Example 3.2.5 of the course notes, including calculation of third derivatives and computation of dominating functions for the 3rd derivatives. (Are the computations simpler in one or the other of the two parametrizations?)

2. This is a continuation of the previous problem concerning ML estimation in the Weibull family. (a) Using the parametrization of the course notes, find the maximizer of the log-likelihood over $\alpha$ for each fixed value of $\beta$,
$$\widehat{\alpha}_n(\beta) \equiv \text{argmax}_{\alpha>0} l_n(\alpha, \beta | \underline{X}_n).$$
(b) The *profile likelihood function* for $\beta$ is defined by
$$l_n^{prof}(\beta | \underline{X}) \equiv l_n(\widehat{\alpha}_n(\beta), \beta | \underline{X}).$$
Compute $l_n^{prof}(\beta | \underline{X})$ for the Weibull family; show that it has a unique maximizer $\widehat{\beta}_n$ and that the MLE $\widehat{(\alpha, \beta)} = (\widehat{\alpha}_n(\widehat{\beta}_n), \widehat{\beta}_n)$. (This is related to Example 6.1 on page 468 of Lehmann and Casella and their problems 6.1-6.3 on page 509. For nice plots to accompany this exercise, see pages 41 - 43 of Cox, D. R. and Oakes, D. (1984); *Analysis of Survival Data*, Chapman and Hall.)
(c) Why does profile likelihood work so nicely in this example?

3. van der Vaart (1998), Problem 25, page 84.

4. Lehmann and Casella, TPE, Problem 4.9, page 504: Consider the following 12 observations from a bivariate normal distribution with parameters $\mu_1 = \mu_2 = 0$, $\sigma_1^2$, $\sigma_2^2$, $\rho$: where $*$ represents a missing value.

Table 1:

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $x_1$ | 1 | 1 | -1 | -1 | 2 | 2 | -2 | -2 | * | * | * | * |
| $x_2$ | 1 | -1 | 1 | -1 | * | * | * | * | 2 | 2 | -2 | -2 |

(a) Show that the likelihood function has global maxima at $\rho = \pm 1/2$, $\sigma_1^2 = \sigma_2^2 = 8/3$ and a saddlepoint at $\rho = 0$, $\sigma_1^2 = \sigma_2^2 = 5/2$.
(b) Show that if an EM sequence starts with $\rho = 0$, then it remains with $\rho = 0$ for all subsequent iterations.
(c) Show that if an EM sequence starts with $\rho$ bounded away from zero, it will converge to a maximum.

5. **Optional bonus problem 1.** Lehmann and Casella, problem 3.15, page 502: note that part (a) is false in the modern understanding of "log-concave" since all uniform distributions $f(x) = (b - a)^{-1} 1_{[a,b]}(x)$ are log-concave, but the mode is not unique. (See Dharmadhikari and Jogdeo (1988), *Unimodality, Convexity, and Applications*, Section 1.4 and page 23.)

6. **Optional bonus problem 2.** Compare the explanation of the EM algorithm in Lehmann and Casella TPE, pages 458-459 with the explanation given in Groeneboom's notes, pages 1 - 3 and 10 - 12. Correct the expressions given in (4.24) of TPE, page 459.