Solutions for HW1, Stat 710, F07

#5.12. The condition is: there exists a unique pth quantile x_p, i.e. a number such that $P(X_1 \le x_p) = p$ and for all $\delta > 0$,

$$P(X_1 \le x_p - \delta) < p, \quad P(X_1 \ge x_p + \delta) < 1 - p$$

The Lemma needed to prove consistency under this condition is Lemma 5.10 on p.47. (We cannot use Theorem 5.23 for this purpose because it has consistency as a *hypothesis* not a conclusion.)

#5.14. It seems that by inspection, there is a naive estimating equation for (α, β) given by:

$$\sum_{i=1}^{n} (1, X_i)' \, (Y_i - \alpha - \beta X_i) = 0$$

However, the second equation does not actually have expectation 0 so we cannot use this as written !

To follow the method of Example 5.26, we can factor the conditional density of (X_i, Y_i) given Z_i (treating the parameters as known) to find that $X_i + \beta(Y_i - \alpha) \sim \mathcal{N}((1+\beta^2)Z_i, (1+\beta^2)\sigma^2)$ is 'sufficient' for Z_i. As in Example 5.26, we factorize the conditional joint density, finding in this case $f_{Y|X+\beta(Y-\alpha)}$ is

$$\mathcal{N}((\alpha + (\beta/(1+\beta^2))(X + \beta(Y-\alpha))), \sigma^2/(1+\beta^2))$$

Thus, we find the log-conditional-likelihood contribution for the i'th observation (given the sufficient statistic value $X_i + \beta(Y_i - \alpha)$) equal to

$$-\frac{1}{2}\log\frac{2\pi\sigma^2}{1+\beta^2} - \frac{1+\beta^2}{2\sigma^2}\Big(Y_i - (\alpha + \frac{\beta}{1+\beta^2}(X_i + \beta(Y_i - \alpha)))\Big)^2$$

$$= -\frac{1}{2}\log\frac{2\pi\sigma^2}{1+\beta^2} - \frac{1}{2(1+\beta^2)\sigma^2}(Y_i - \alpha - \beta X_i)^2$$

Finally, the inference will be based on the 'score' for this last conditional likelihood, obtained by differentiating the estimating equation only with respect to $\alpha, \beta)$ and multiplying through by $(1+\beta^2))$:

$$\sum_{i=1}^{n} \binom{1}{X_i}(Y_i - \alpha - \beta X_i) + \frac{\beta}{1+\beta^2}\sum_{i=1}^{n}\binom{0}{1}(Y_i - \alpha - \beta X_i)^2 + \binom{0}{n\beta}$$

This displayed expression must according to the idea of Example 5.26 be bias-corrected, i.e. the expectation calculated conditionally given $\{X_i + \beta(Y_i - \alpha)\}_i$ must be subtracted. Note that *all* terms should have conditional expectations taken in this way. (In the first version of this solution, I erroneously omitted the expectation of the first displayed term.) Note that since $Y_i - \alpha - \beta X_i = f_i - \beta e_i$ and $X_i + \beta(Y_i - \alpha) = (1+\beta^2) Z_i + e_i + \beta f_i$ are uncorrelated, $Y_i - \alpha - \beta X_i = f_i - \beta e_i$ is indenpendent of $(Z_i, X_i + \beta(Y_i - \alpha))$, and the expectation of the last display is

$$\binom{0}{n\beta} + E\Big(\sum_{i=1}^{n} (Y_i - \alpha - \beta X_i) \begin{pmatrix} 1 \\ (X_i + \beta(Y_i - \alpha))/(1+\beta^2) \end{pmatrix} \Big) = \binom{0}{n\beta}$$

Thus the conditionally bias-corrected estimating equation becomes

$$\sum_{i=1}^{n} (Y_i - \alpha - \beta X_i) \begin{pmatrix} 1 \\ X_i + \beta(Y_i - \alpha) \end{pmatrix} = \binom{0}{0} \qquad (*)$$

Consistency and asymptotic normality (possibly degenerate) of the solution $(\hat\alpha, \hat\beta)'$ is immediate from the delta-method and CLT (or from Theorem 5.23) once we verify that the (expected) Jacobian of the left-hand side of (*) is non-singular. This Jacobian is easily calculated to be

$$n \begin{pmatrix} -1 & -\bar X \\ 2\beta(\alpha - \bar Y) - (1-\beta^2)\bar X & \sum_{i=1}^{n} ((Y_i - \alpha)^2 - X_i^2 - 2\beta X_i (Y_i - \alpha)) \end{pmatrix}$$

which after substitution of the obvious relation $\hat\alpha = \bar Y - \hat\beta \bar X$ becomes

$$n \begin{pmatrix} 1 & \bar X \\ (1+\beta^2)\bar X & (1+\beta^2)\bar X^2 + \frac{1}{n}\sum_{i=1}^{n} ((X_i - \bar X)^2 - (Y_i - \bar Y)^2 + 2\beta X_i (Y_i - \bar Y)) \end{pmatrix}$$

and which in turn is easily seen to be asymptotic for large n by the Law of Large Numbers to

$$n \begin{pmatrix} 1 & (EZ)^2 \\ (1+\beta^2)\,(EZ)^2 & (1+\beta^2)\,E(Z^2) \end{pmatrix}$$

In case one or both moments of Z are infinite, we find that the order of precision of one or both of the estimators is actually better than $1/\sqrt{n}$.

 # 5.18. Here

$$p_\vartheta(x) = \frac{\vartheta^x e^{-\vartheta}}{x!} \quad , \quad \dot l_\vartheta = \frac{x}{\vartheta} - 1 \quad , \quad \ddot l_\vartheta = -\,\vartheta^{-2}$$

and the assertion in question is that $-E\,\ddot l_\vartheta(X) = \vartheta^{-2} = E\,\dot l_\vartheta^2(X)$. To get this using general principles via Dominated Convergence (with respect to counting measure on the nonnegative integers), dominate $p_\vartheta(x)$ for $\vartheta \in (\vartheta - \delta, \vartheta + \delta)$ by $(\vartheta + \delta)^x\, e^{-\vartheta+\delta}/x!$, and switch orders of differentiation and infinite-summation because the differentiated series are dominated by an absolutely convergent one.

5.25. Here $m_\theta(x) = -\frac{1}{2}\log(2\pi\sigma^2) - (x_\mu)^2/(2\sigma^2)$.

(i) Consider $E(\sup_{\theta \in K} m_\theta(X_1) = -\frac{1}{2}\log(2\pi\sigma_*^2)$ where $\sigma_*^2 = \inf_{\theta\in K} \theta_2$.

(ii) Using a *compactification* means allowing $K^* \subset \mathbf{R}\times[0,\infty)$ compact with $m_\theta(x)$ at $\theta^* = (\mu, 0)$ given as limit of log-densities $\lim_{\theta\to\theta^*} m_\theta(x)$, but for $\mu \neq 0$ this limit is $-\infty$ for a.e. x, while $E\sup_{\theta\in K^*} m_{\theta(X_1)} = +\infty$.

(iii). Now the unit of data is (X_1, X_2) and $m_\theta(X)$ is replaced by $m_\theta(\underline{X}) = m_\theta(X_1) + m_\theta(X_2)$. For a.e. value $X_1 \neq X_2$, so if $K^* \subset \mathbf{R}\times[0,\infty)$ is allowed in the maximization and $\theta_2^* = 0$, then $\lim_{\theta\to\theta^*} m_\theta(\underline{X}) = -\infty$ and

$$\sup_{\theta\in[0,\infty)} m_\theta(\underline{X}) = \max_{\theta\in[0,\infty)} \Big\{ -\log(2\pi\sigma^2) - (X_1 - X_2)^2/(8\sigma^2)\Big\}$$

occurs at $\sigma^2 = (X_1 - X_2)^2/8$ and has finite expectation.

Extra Problem. As hinted in class, the result here is a Corollary of the Theorem 19.4 proved in class, that if for all $\epsilon > 0$ there is a finite number of ϵ-bracketing functions for a class $\mathcal{F}$ of functions in L_1, then that class is Glivenko-Cantelli, i.e. satisfies the uniform law of large numbers. In this problem, the bracketing functions are found:

(i) by surrounding each $\theta \in K \in \mathbf{R}$ by a small open ball U_θ such that

$$E\Big(\sup_{\theta'\in U_\theta} m_{\theta'}(X_1) - \inf_{\theta'\in U_\theta} m_{\theta'}(X_1)\Big) \leq \epsilon$$

(ii) by finding a finite cover of the compact parameter set K by balls $U_{\theta_1}, \ldots, U_{\theta_k}$; and

(iii) by bracketing all functions m_θ for $\theta \in K$ using only finite linear combinations of functions which are piecewise constant on the partition of K induced by (intersections of) the balls U_{θ_j}, using coefficients which can be taken from a finite set with spacing depending on ϵ.

The result of this problem is taken up again in the book in essentially the same way in Example 19.8.