Stat 710 HW2 Solutions

 # 1 in-class. We are considering *iid* $\mathcal{N}(\mu, \sigma^2)$ r.v.'s X_i, and $\vartheta = (\mu, \sigma)$, with generalized method of moments estimators defined in terms of $\mathbf{e}(\vartheta) = E_\vartheta\Big((I_{[X_1 \le -1]},\, I_{[X_1 \le 1]})\Big) = (\Phi(-(1+\mu)/\sigma), \Phi((1-\mu)/\sigma))$. These moment estimators $(\tilde{\mu},\, \tilde{\sigma})$ are uniquely determined from the equations

$$\frac{-2\tilde{\mu}}{\tilde{\sigma}} = \Phi^{-1}(\frac{1}{n}\sum_{i=1}^n I_{[X_i \le 1]}) + \Phi^{-1}(\frac{1}{n}\sum_{i=1}^n I_{[X_i \le -1]})$$

$$\frac{2}{\tilde{\sigma}} = \Phi^{-1}(\frac{1}{n}\sum_{i=1}^n I_{[X_i \le 1]}) - \Phi^{-1}(\frac{1}{n}\sum_{i=1}^n I_{[X_i \le -1]})$$

The Jacobian derivative matrix $J_{\mathbf{e}}$ is given by

$$\begin{pmatrix} \frac{-1}{\sigma}\,\phi((1+\mu)/\sigma) & \frac{1+\mu}{\sigma^2}\,\phi((1+\mu)/\sigma) \\ \frac{-1}{\sigma}\,\phi((1-\mu)/\sigma) & \frac{\mu-1}{\sigma^2}\,\phi((1-\mu)/\sigma) \end{pmatrix}$$

We learn from Theorem 4.1 that these generalized moment estimators are asymptotically normal with mean ϑ and asymptotic variance (without factor $1/n$) given at $\mu_0 = 0,\, \sigma_0 = 1$ by

$$J_{\mathbf{e}}^{-1}\, E \begin{pmatrix} I_{[X_1 \le -1]} & I_{[X_1 \le -1]} \\ I_{[X_1 \le -1]} & I_{[X_1 \le 1]} \end{pmatrix} J_{\mathbf{e}}^{-1\,tr}$$

$$= \frac{1}{\phi^2(1)} \begin{pmatrix} -1 & 1 \\ -1 & -1 \end{pmatrix}^{-1} \begin{pmatrix} \Phi(-1) & \Phi(-1) \\ \Phi(-1) & \Phi(1) \end{pmatrix} \begin{pmatrix} -1 & -1 \\ 1 & -1 \end{pmatrix}^{-1}$$

$$= \frac{1}{4\phi^2(1)} \begin{pmatrix} \Phi(1) + 3\Phi(-1) & \Phi(1) - \Phi(-1) \\ \Phi(1) - \Phi(-1) & \Phi(1) - \Phi(-1) \end{pmatrix}$$

and $\text{Avar}(\tilde{\mu}) = (3\Phi(-1) + \Phi(1))/(4\phi^2(1)) = 5.62$ versus the corresponding value of 1 for $\text{Avar}(\overline{X})$.

 # 1, p.40. Here $EX_i^2 = \vartheta^2\,\frac{1}{2}\int_{-1}^1 x^2\,dx = \vartheta^2/3$, wich implies the moment-based estimator is $\tilde{\vartheta} = g(\overline{X^2}) = \Big(\frac{3}{n}\sum_{i=1}^n X_i^2\Big)^{1/2}$. Then $g'(g^{-1}(\vartheta)) = \frac{\sqrt{3}}{2}\,(\vartheta^2/3)^{-1/2} = 3/(2\vartheta)$, while $\sqrt{n}\,(n^{-1}\sum_{i=1}^n X_i^2 - \vartheta^2/3) \to^{\mathcal{D}} \mathcal{N}(0, 4\vartheta^4/45)$. This implies $\sqrt{n}\,(\tilde{\vartheta} - \vartheta) \sim \mathcal{N}(0, \vartheta^2/5)$.

6. Following the hint: suppose $\vartheta_1,\ \vartheta_2$ are distinct zero-points for $\mathbf{e}$ on the convex set Θ, and $g(\lambda) = (\vartheta_1 - vt_2)^{tr}\, \mathbf{e}(\lambda\vartheta_1 + (1-\lambda)\vartheta_2)$ for $\lambda \in [0,1]$. Then $g(0) = g(1) = 0$, but $g'(\lambda) = (\vartheta_1 - vt_2)^{tr}\, J_{\mathbf{e}}(\lambda\vartheta_1 + (1-\lambda)\,(\vartheta_1 - \vartheta_2) > 0$ on $(0,1)$, which is not possible since $g(1) - g(0) = \int_0^1 g(\lambda)\, d\lambda$.

 # 5, p.83. $L(\mathbf{X}, \vartheta) = \prod_i f_{X_i}(X_i, 1/\vartheta) = \prod_i (1/X_i!)\, \vartheta^{-n\overline{X}}\, e^{-n/\vartheta}$, which is maximized in ϑ at $1/\overline{X}$. Since $\sqrt{n}\,(\overline{X} - 1/\vartheta) \sim \mathcal{N}(0, 1/\vartheta)$ and $g(z) = 1/z$ satisfies $g'(1/\vartheta) = -\vartheta^2$ in the Δ-method, we get $\sqrt{n}\,(1/\overline{X} - \vartheta) \sim \mathcal{N}(0,\ (1/\vartheta)\,(-\vartheta)^2) = \mathcal{N}(0, \vartheta^3)$.

 # 9. $L(\mathbf{X}, \vartheta) \ = \ \vartheta^{-n} \prod_i I_{[X_i \le \vartheta]} \ = \ \vartheta^{-n}\, I_{[\max X_i \le \vartheta]}$, and this is maximized in ϑ at $\hat{\vartheta} = \max_i X_i$. To show consistent and not asymptotically normal, calculate

$$P(\hat{\vartheta} \le \vartheta - \frac{t}{n}) \ = \ (P(X_1 \le \vartheta - \frac{t}{n}))^n \ = \ (1 - \frac{t}{n\vartheta})^n \ = \ \exp(-t/\vartheta)$$

So the asymptotic distribution of $n(\vartheta - \hat{\vartheta})$ is *Expon*$(1/\vartheta)$.

 # 14. The ingredients of this problem, like the Example 5.26 on which it is based by analogy, are: (i) an unobserved component Z_i associated with each observation (X_i, Y_i), (ii) a sufficient statistic $U_i(\vartheta)$ for Z_i if the observation (X_i, Y_i) and ϑ are given; and (iii) a function arising from factorization of the likelihood for (X_i, Y_i, Z_i) which, after correcting for conditional expectation given $U_i(\vartheta)$ under the model with parameter ϑ, can serve as estimating function ψ.

In this example, using h to denote the density of Z_i, we have log-likelihood and factorization as follows:

$$\log f_{X,Y,Z}(x,y,z) \ = \ \log \frac{h(z)}{2\pi\sigma^2} - \frac{1}{2\sigma^2}\,(x^2 + (1+\beta^2)z^2 + (y-\alpha)^2 - 2z(x + \beta(y-\alpha)))$$

$$= \ \log \frac{h(z)}{2\sigma^2\pi} \ - \ \frac{1}{2(1+\beta^2)\sigma^2}\ \Big\{(u - z(1+\beta^2))^2 + (y - \alpha - \beta x)^2\Big\}$$

where $u \equiv x + \beta(y - \alpha)$, and we define the sufficient 'statistic', for known $(\alpha, \beta, \sigma^2)$, by $U_i(\alpha, \beta) \ = \ X_i + \beta(Y_i - \alpha)$. The likelihood factorization (or the hint given in class to think about the ML estimating function when Z_i is absent) suggests to take $\psi(X_i, Y_i, \alpha, \beta)$ as

$$-\nabla_{\alpha,\beta} \left[\frac{(Y_i - \alpha - \beta X_i)^2}{(1+\beta^2)} \right] \ = \ \frac{Y_i - \alpha - \beta X_i}{1 + \beta^2} \begin{pmatrix} 1 \\ U_i/(1+\beta^2) \end{pmatrix} \qquad (1)$$

corrrected if necessary by its conditional expectation given U_i . To find the correction-term, we need to calculate the conditional distribution (for given (α, β)) of X_i given U_i which by sufficiency is the same as the conditional distribution of X_i given U_i, Z_i. Recall that Z_i (with mean and variance which we denote μ_Z, σ_Z^2) is not itself assumed to be a normal random variable, so we find the conditional distribution from first principles: first integrating out the x variable from $f_{X,U,Z}(x, u, z)$, we find that $f_{U,Z}(u, z)$ is proportional to $\exp(-(u/\beta - (1+\beta^2)x)^2/(2\sigma^2(1+\beta^2)))$. Therefore,

$$f_{X|U,Z}(x|u,z) \propto \exp\Big(-\frac{(u-(1+\beta^2)x/\beta)^2}{2\sigma^2(1+\beta^2)}\Big) \quad , \quad \text{hence} \quad \sim \mathcal{N}\Big(\frac{u}{1+\beta^2}\, , \, \frac{\sigma^2\,\beta^2}{1+\beta^2}\Big)$$

Now

$$E(X_i \,|\, U_i) \;=\; \frac{U_i}{1+\beta^2} \quad , \quad E(Y_i - \alpha - \beta X_i \,|\, U_i) \;=\; E(\frac{1}{\beta}(U_i - (1+\beta^2)X_i) \,|\, U_i) \;=\; 0$$

and it follows that expressions (1) have expectation 0. Now we will calculate the V, W, matrices arising in the asymptotic distribution theory of estimating-equation solutions, and will find that both of these matrices are proportional to the same matrix

$$\Sigma \;\equiv\; E\begin{pmatrix} 1 \\ U_i/(1+\beta^2) \end{pmatrix}^{\otimes 2} \;=\; \begin{pmatrix} 1 \\ \mu_Z \end{pmatrix}^{\otimes 2} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \frac{(1+\beta^2)^2(\sigma_Z^2+\sigma^2)+\beta^2\sigma^2}{(1+\beta^2)^2}$$

where in the last term we have inserted the unconditional variance of $U_i = (1+\beta^2)(Z_i + e_i) + \beta f_i$. Then, repeatedly using the preceding conditional expectation of $Y_i - \alpha - \beta X_i$ equal to 0 in order to simplify expressions,

$$V \;=\; E_{\alpha,\beta,\sigma^2}\Big(-\nabla^{tr}_{\alpha,\beta}\,\psi(X_1, Y_1, \alpha, \beta)\Big) \;=\; \frac{1}{1+\beta^2}\, E\begin{pmatrix} 1 \\ U_i/(1+\beta^2) \end{pmatrix}\begin{pmatrix} 1 \\ X_i \end{pmatrix}^{tr}$$

which implies, after substituting $E(X_i|U_i) = U_i/(1+\beta^2)$,

$$V = \frac{1}{1+\beta^2}\,\Sigma$$

Next,

$$\begin{aligned} W \;&=\; E_{\alpha,\beta,\sigma^2}\Big(\psi(X_i, Y_i, \alpha, \beta)^{\otimes 2}\Big) \\ &=\; (1+\beta^2)^{-2}\, E\Big(\big(\frac{U_i}{\beta} - \frac{1+\beta^2}{\beta}X_i\big)^2 \begin{pmatrix} 1 \\ U_i/(1+\beta^2) \end{pmatrix}^{\otimes 2}\Big) \end{aligned}$$

which by substitution of the conditional variance for X_i given U_i yields

$$W \;=\; \frac{1}{\beta^2}\, E\Big(\frac{\sigma^2\beta^2}{1+\beta^2} \begin{pmatrix} 1 \\ U_i/(1+\beta^2) \end{pmatrix}^{\otimes 2} \Big) \;=\; \sigma^2\, V$$

The estimating function ψ leads to uniquely defined estimators as follows:

$$\sum_{i=1}^n (Y_i - \alpha - \beta X_i) \;=\; 0 \qquad \Longrightarrow \qquad \overline{Y} - \alpha - \beta \overline{X} \;=\; 0$$

and after substituting the last equation we find

$$\sum_{i=1}^n (Y_i - \alpha - \beta X_i) U_i \;=\; 0 \;\;\Rightarrow\;\; \frac{1}{n-1} \sum_{i=1}^n (Y_i - \overline{Y} - \beta(X_i - \overline{X}))\, (Y_i - X_i/\beta) \;=\; 0$$

which yields a unique solution using $s_y^2 - s_x^2 + (\beta^{-1} - \beta) s_{xy} \;=\; 0.$ The hypotheses of our asymptotic theorem on estimating equations are readily checked, and the net result of that theorem is that

$$\sqrt{n}\, \Big(\begin{pmatrix} \tilde{\alpha} - \alpha \\ \tilde{\beta} - \beta \end{pmatrix} \;\xrightarrow{\;\mathcal{D}\;}\; \mathcal{N}(\mathbf{0}\,,\, \sigma^2\,(1+\beta^2)\,\Sigma)$$

 # 16. The estimator under consideration is $\tilde{\vartheta} \;=\; \arg\min_\vartheta \sum_{i=1}^n |Y_i - f_\vartheta(X_i)|$, which you would expect to converge in probability to a minimizer ϑ_* of $E|Y_1 - f_\vartheta(X_1)|$. In fact, Theorem 5.14 could easily be used to prove that. Or, with some further assumptions on f_ϑ, one could guarantee that the minimizer ϑ_* would be unique, in which case a Theorem like 5.7 would establish consistency. The asymptotic distribution of $\sqrt{n}\,(\tilde{\vartheta} - \vartheta_*)$ would then be normal with mean zero and an asymptotic variance as in Theorem 5.23. All of this depends on appropriate smoothness of $E|Y_1 - f_\vartheta(X_1)|$ in ϑ, which we must verify, as follows, using the same idea as in Example 5.24, p. 55.

$$E|Y_1 - f_\vartheta(X_1)| \;=\; E|e_1 + f_{\vartheta_*}(X_1) - f_\vartheta(X_1)| \;=\; \int\!\!\int \Big\{ (-z + f_\vartheta(x) - f_{\vartheta_*}(x))^+ \\ + (-z + f_\vartheta(x) - f_{\vartheta_*}(x))^- \Big\} dz dx$$

Therefore, using the Fundamental Theorem to remove the terms arising from differentiation of limits of integration (in z, from $-\infty$ to $f_\vartheta(x) - f_{\vartheta_*}(x)$ for $(\cdot)^+$, and from $f_\vartheta(x) - f_{\vartheta_*}(x)$ to ∞ for $(\cdot)^-$, we obtain

$$\nabla_\vartheta\, E|Y_1 - f_\vartheta(X_1)| \;=\; \int\!\!\int \Big\{ I_{[z \le f_\vartheta(x) - f_{\vartheta_*}(x)]} - I_{[z > f_\vartheta(x) - f_{\vartheta_*}(x)]} \Big\}\, \nabla_\vartheta\, f_\vartheta(x)\, f_e(z)\, dz\, dx \\ =\; \int\!\!\int (2F_e(f_\vartheta(x) - f_{\vartheta_*}(x)) - 1)\; \nabla_\vartheta\, f_\vartheta(x)\, dx$$