Econometrics II, third problem set.

Remember "Matlab" can be any programming language that allows you to get the work done.

1. **ML.** Consider coin tossing with a single possibly biased coin. The density function for the random variable $y = 1(heads)$ is

$$\begin{aligned} f_Y(y, p_0) &= p_0^y (1-p_0)^{1-y}, y \in \{0,1\} \\ &= 0, y \notin \{0,1\} \end{aligned}$$

Suppose that we have a sample of size n. We know from theory that

$$\sqrt{n}(\bar{y} - p_0) \overset{a}{\sim} N\left[0, \mathcal{J}_\infty(p_0)^{-1}\mathcal{I}_\infty(p_0)\mathcal{J}_\infty(p_0)^{-1}\right]$$

   (a) find the expression for the maximum likelihood estimator of p_0

   (b) find the analytic expression for $g_t(\theta)$ and show that $\mathcal{E}_\theta[g_t(\theta)] = 0$

   (c) find the analytical expressions for $\mathcal{J}_\infty(p_0)$ and $\mathcal{I}_\infty(p_0)$ for this problem

   (d) Show that $\lim Var\sqrt{n}(\hat{p} - p) = p_0(1-p_0)$, a.s., and that this is equal to $\mathcal{J}_\infty(p_0)^{-1}\mathcal{I}_\infty(p_0)\mathcal{J}_\infty(p_0)^{-1}$

   (e) Write a Matlab program that does a Monte Carlo study that shows that $\sqrt{n}(\bar{y} - p_0)$ is approximately normally distributed when n is large. Set $p_0 = 0.3$. Please give me histograms that show the sampling frequency of $\sqrt{n}(\bar{y} - p_0)$ for $n \in \{5, 20, 100\}$.

2. **ML.** Consider the model $y_t = x_t'\beta + \alpha\epsilon_t$

   (a) Assume that the errors follow the Cauchy (Student-t with 1 degree of freedom) density. So

$$f(\epsilon_t) = \frac{1}{\pi(1+\epsilon_t^2)}, -\infty < \epsilon_t < \infty$$

   The Cauchy density has a shape similar to a normal density, but with much thicker tails. Thus, extremely small and large errors occur much more frequently with this density than would happen if the errors were normally distributed. Find the score function $g_n(\theta)$ where $\theta = \begin{pmatrix} \beta' & \alpha \end{pmatrix}'$.

   (b) Now assume that the errors are IID Gaussian: $\epsilon_t \sim IIN(0,1)$. Find the score function $g_n(\theta)$ where $\theta = \begin{pmatrix} \beta' & \alpha \end{pmatrix}'$.

   (c) Compare the first order conditions that define the ML estimators in the two cases, and interpret the differences.

      i. *Why* are the first order conditions that define an efficient estimator different in the two cases?

      ii. What is the intuitive explanation for the difference?

3. **GMM.** In the GRETL User Guide, in the chapter on GMM estimation, there is a section *"A real example: the Consumption Based Asset Pricing Model"*, where a consumption based CAPM similar to the Hansen-Singleton (1982) model is discussed. Write Matlab code (or some other programming language, but don't use Gretl) to estimate this model, using the hall.gdt data set that is provided with Gretl (file->open data->sample file->Gretl->hall. You will need to export the data to use it with Matlab.

   (a) try to replicate some or all of the estimation results found in the Gretl User Guide. At a minimum, you should provide estimated coefficients and estimated standard errors for the two step estimator.

   (b) interpret the results

   (c) discuss in general terms the usefulness of this GMM procedure for estimating the parameters of this model. What, if anything, is going wrong here?

4. **GMM.** Prove that the GMM estimator based upon the g moment conditions $m_n(\theta) = \left[\begin{array}{cc} p_n'(\theta) & q_n'(\theta) \end{array} \right]'$ and the corresponding true optimal weight matrix is asymptotically efficient with respect to the GMM estimator based upon the $h < g$ moment conditions $p_n(\theta)$ and the corresponding true optimal weight matrix.

   (a) Interpret the result

   (b) Discuss the importance of the result from an empirical point of view. Are there any cautions one should observe when doing applied GMM work? Describe any problems you can imagine.

5. **GMM.** In the context of the Hansen-Sargan test for correct specification of moments, prove that the matrix $P_\infty = I_g - \Omega_\infty^{-1/2} D_\infty' \left(D_\infty \Omega_\infty^{-1} D_\infty' \right)^{-1} D_\infty \Omega_\infty^{-1/2}$ is idempotent and that its rank is $g - K$, where g is the number of moment conditions and K is the number of parameters.

6. **GMM/IV.** Suppose we have two equations

$$\begin{aligned} y_{t1} &= \alpha_1 + \alpha_2 y_{t2} + \epsilon_{t1} \\ y_{t2} &= \beta_1 + \beta_2 x_t + \epsilon_{t2} \end{aligned}$$

   where $V(\epsilon_{t1}) = \sigma_1^2 > 0$, $V(\epsilon_{t2}) = \sigma_2^2 > 0$, $E(\epsilon_{t1}\epsilon_{t2}) = \sigma_{12} \neq 0$. The observations are independent over time. The variable x_t is strictly exogenous: it is uncorrelated with the two epsilons at all time periods.

   (a) Is the OLS estimator of the parameters of the first equation consistent or not? Explain.

   (b) Is the OLS estimator of the parameters of the second equation consistent or not? Explain.

   (c) If the OLS estimator of the parameters of the first equation is not consistent, propose a consistent estimator of the parameters of the first equation and explain why the proposed estimator is consistent.

   (d) If the OLS estimator of the parameters of the second equation is not consistent, propose a consistent estimator of the parameters of the second equation and explain why the proposed estimator is consistent.

7. **GMM.** Given the 10 independent data points

y	0	0	0	1	1	5	8	16	20	30
x	-1	-1	1	0	-1	-1	1	1	1	2

   .

   For the Poisson model, the density $f_Y(y|x) = \frac{\exp(-\lambda)\lambda^y}{y!}$, $y = 0, 1, 2, ...$ To make the model depend on conditioning variables, use the parameterization $\lambda(x) = \exp(\theta_1 + \theta_2 x)$.

   (a) the mean of a Poisson distribution with parameter λ is equal to λ, and so is the variance. Propose moment conditions to define an overidentified $(g > k)$ GMM estimator of θ_1 and θ_2, using the proposed parameterization of the conditional mean.

   (b) Estimate the parameters using two-step efficient GMM, using the moment conditions you have proposed.

   (c) compute the Hansen-Sargan test for correct specification of the moment conditions, and interpret.

   (d) discuss the results.