1. Suppose that $x_i \sim$ uniform(0,1), and $y_i = 1 - x_i^2 + \varepsilon_i$, where $\varepsilon_i$ is iid(0,$\sigma^2$). Suppose we estimate the misspecified model $y_i = \alpha + \beta x_i + \eta_i$ by OLS. Find the numeric values of $\alpha^0$ and $\beta^0$ that are the probability limits of $\hat{\alpha}$ and $\hat{\beta}$

2. Verify your results using Octave by generating data that follows the above model, and calculating the OLS estimator. When the sample size is very large the estimator should be very close to the analytical results you obtained in question 1.

3. Use the asymptotic normality theorem to find the asymptotic distribution of the ML estimator of $\beta^0$ for the model $y = x\beta^0 + \varepsilon$, where $\varepsilon \sim N(0,1)$ and is independent of $x$. This means finding $\frac{\partial^2}{\partial\beta\partial\beta'}s_n(\beta)$, $\mathcal{J}(\beta^0)$, $\left.\frac{\partial s_n(\beta)}{\partial\beta}\right|$, and $\mathcal{I}(\beta^0)$. The expressions may involve the unspecified density of $x$.

4. Assume a d.g.p. follows the logit model: $\Pr(y = 1|x) = \left(1 + exp(-\beta^0 x)\right)^{-1}$.

   (a) Assume that $x \sim$ uniform(-a,a). Find the asymptotic distribution of the ML estimator of $\beta^0$ (this is a scalar parameter).

   (b) Now assume that $x \sim$ uniform(-2a,2a). Again find the asymptotic distribution of the ML estimator of $\beta^0$.

   (c) Comment on the results

5. Show how to cast the generalized IV estimator presented in section 11.4 as a GMM estimator. Identify what are the moment conditions, $m_t(\theta)$, what is the form of the the matrix $D_n$, what is the efficient weight matrix, and show that the covariance matrix formula given previously corresponds to the GMM covariance matrix formula.

6. Using Octave, generate data from the logit dgp . Recall that $E(y_t|\mathbf{x}_t) = \mathbf{p}(\mathbf{x}_t,\theta) = [1 + \exp(-\mathbf{x}_t\prime\theta)]^{-1}$. Consider the moment condtions (exactly identified) $m_t(\theta) = [y_t - p(\mathbf{x}_t,\theta)]\mathbf{x}_t$

   (a) Estimate by GMM, using these moments.

   (b) Estimate by MLE.

   (c) The two estimators should coincide. Prove analytically that the estimators coicide.

7. Considering the MEPS data, for the OBDV ($y$) measure, let $\eta$ be a latent index of health status that has expectation equal to unity.[1] We suspect that $\eta$ and $PRIV$ may be correlated,

[1] A restriction of this sort is necessary for identification.

but we assume that $\eta$ is uncorrelated with the other regressors. We assume that

$$E(y|PUB, PRIV, AGE, EDUC, INC, \eta)$$
$$= \exp(\beta_1 + \beta_2 PUB + \beta_3 PRIV + \beta_4 AGE + \beta_5 EDUC + \beta_6 INC)\eta.$$

We use the Poisson QML estimator of the model

$$\begin{aligned} y &\sim \text{Poisson}(\lambda) \\ \lambda &= \exp(\beta_1 + \beta_2 PUB + \beta_3 PRIV + \beta_4 AGE + \beta_5 EDUC + \beta_6 INC). \end{aligned} \quad (1)$$

Since much previous evidence indicates that health care services usage is overdispersed[2], this is almost certainly not an ML estimator, and thus is not efficient. However, when $\eta$ and $PRIV$ are uncorrelated, this estimator is consistent for the $\beta_i$ parameters, since the conditional mean is correctly specified in that case. When $\eta$ and $PRIV$ are correlated, Mullahy's (1997) NLIV estimator that uses the residual function

$$\varepsilon = \frac{y}{\lambda} - 1,$$

where $\lambda$ is defined in equation 1, with appropriate instruments, is consistent. As instruments we use all the exogenous regressors, as well as the cross products of $PUB$ with the variables in $Z = \{AGE, EDUC, INC\}$. That is, the full set of instruments is

$$W = \{1 \quad PUB \quad Z \quad PUB \times Z\}.$$

(a) Calculate the Poisson QML estimates.

(b) Calculate the generalized IV estimates (do it using a GMM formulation - see the portfolio example for hints how to do this).

(c) Calculate the Hausman test statistic to test the exogeneity of PRIV.

(d) comment on the results

[2] Overdispersion exists when the conditional variance is greater than the conditional mean. If this is the case, the Poisson specification is not correct.