Consider the random variable $y$ with density
\[
f(y) = (y + 1)(2\lambda)^y(1 + 2\lambda)^{-(y+0.5)}, \quad y = 0, 1, 2, \ldots, \lambda > 0.
\]
The first two moments of $y$ can be shown to be
\[
E[y] = \lambda, \quad V[y] = \lambda(1 + 2\lambda).
\]
We form a regression model by specifying that for the d.g.p.
\[
E[y_i|x_i] = \exp(x_i'\beta_0)
\]
where $\beta$ is a $k \times 1$ parameter vector and $x_i$ is a $k \times 1$ nonstochastic regressor vector. Independence over $i$ is assumed.
In the d.g.p. the regressors $x_i$ are nonstochastic and $y_i|x_i$ is independent over $i$ with the preceding density for $y_i|x_i$ and $\beta = \beta_0$.
If you need to make other assumptions, state them as you go along.
You can apply laws of large numbers and central limit theorems without formally verifying necessary conditions for their use (except 1(d) requires further detail).
You need not verify any second-order conditions.
(a) Show that the MLE for $\beta$ maximizes (upon scaling by $N^{-1}$)
\[
Q_N(\beta) = \frac{1}{N} \sum_{i=1}^{N} \{ \ln(y_i + 1) + y_i \ln 2 + y_i x_i'\beta - (y_i + 0.5) \ln(1 + 2 \exp(x_i'\beta)) \}.
\]
(b) Obtain $Q_0(\beta) = \plim Q_N(\beta)$.
(c) Using (b) prove that the local max to $Q_N(\beta)$ is consistent for $\beta_0$. State any assumptions made.
(d) Now state what LLN you would use to verify part (b) and what additional information, if any, is needed to apply this law. A brief answer will do. There is no need for a formal proof.
(e) If the density $f(y)$ is misspecified will the MLE in this example be consistent? Explain.
2. Consider the nonlinear least squares estimator that minimizes

\[ S_N(\beta) = \frac{1}{2N} \sum_{i} (y_i - g(x'_i\beta))^2, \]

where without loss of generality we use \(1/2N\) rather than \(1/N\) as it simplifies the results, and in the d.g.p.

\[ E[y_i|x_i] = g(x'_i\beta_0). \]

In the d.g.p. the regressors \(x_i\) are nonstochastic and \(y_i|x_i\) is independent over \(i\).

You can assume that the NLS estimator is consistent.

(a) Give the limit distribution of \(\sqrt{N} \frac{\partial S_N}{\partial \beta} \bigg|_{\beta_0}\), assuming that the usual asymptotic result applies.

(b) Obtain the probability limit of \(\frac{\partial^2 S_N}{\partial \beta \partial \beta} \bigg|_{\beta_0}\), assuming that the usual asymptotic result applies.

(c) Combine the above to obtain the limit distribution of \(\sqrt{N} (\hat{\beta} - \beta_0)\).

(d) Specialize this result to nonlinear least squares estimator when the d.g.p. is the density given in question 1.

(e) Hence provide a test of \(H_0 : \beta_j = 0\) against \(H_a : \beta_j \neq 0\) at level 0.05.

3. (a) Give an example of a law of large numbers in the scalar case, stating clearly the assumptions of the law of large numbers and the conclusion.

(b) For estimator \(\hat{\beta}\) that maximizes \(Q(\beta) = \sum_{i=1}^{N} \{y_i x'_i \beta - \exp(x'_i\beta) - \ln y_i\}\), give the exact algebraic formula for an iterative method to calculate \(\beta\).
4. Consider the following Stata code and output, where the command

```
nl (y = exp(fxg*x + fintercept))
```

Performs nonlinear least squares regression of $y$ on $\exp(\beta_1 + \beta_2 x)$ with coefficients stored in _b[/intercept] and _b[/x] and the default for command `nl` is to assume homoskedastic errors.

```
. program mystery, rclass
   1. version 10.1
   2. drop _all
   3. set obs 100
   4. generate x = rnormal(0,1)
   5. generate y = rpoisson(exp(1+1*x))
   6. nl (y = exp({x}*x + {intercept}))
   7. return scalar b2 =_b[/x]
   8. return scalar se2 =_se[/x]
   9. test _b[/x]=1
  10. return scalar r2 = (r(p)<.05)
  11. end
.
```

```
. simulate b2f=r(b2) se2f=r(se2) reject2f=r(r2), reps(1000) nodots: mystery
```

Command: mystery
b2f: r(b2)
se2f: r(se2)
reject2f: r(r2)
.
```
. summarize b2f se2f reject2f
```

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>b2f</td>
<td>1000</td>
<td>1.003304</td>
<td>.0751338</td>
<td>.795578</td>
<td>1.295699</td>
</tr>
<tr>
<td>se2f</td>
<td>1000</td>
<td>.0413251</td>
<td>.0132319</td>
<td>.0101686</td>
<td>.0850081</td>
</tr>
<tr>
<td>reject2f</td>
<td>1000</td>
<td>.272</td>
<td>.4452125</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Are the results of the simulation what you expect given the d.g.p. for this example and estimation by NLS?
Explain your answer.