(1) Suppose that a discrete, integer-valued random variable $y$ has a Poisson distribution with density function:

$$f(y|\theta) = e^{-\theta} \frac{\theta^y}{y!}, \quad y = 0, 1, 2, 3, \cdots$$

where “!” denotes the factorial symbol. (Note that the Poisson distribution is widely used for modeling count outcomes). Finally, note that if $y$ has a Poisson distribution, as described above, then $E(y) = \text{Var}(y) = \theta$.

(1a) Verify that $E(y) = \theta$.

*Hint:* Recall the infinite series representation of the exponential function:

$$\exp(x) = \sum_{j=0}^{\infty} \frac{x^j}{j!}.$$ 

(1b) Suppose that you obtain a random sample of $n$ observations from the Poisson distribution, \{y_1, y_2, \cdots, y_n\}. Derive the log likelihood function.

(1c) Obtain the maximum likelihood estimator of $\theta$.

(1d) Prove for this particular model that the zero expected score property holds.

(1e) The data on the course website are taken from the study of Riphahn, Wambach and Million (2003 *Journal of Applied Econometrics*). We use an abbreviated version of their data to study the number of physician visits during the past 3 months as a function of age and gender. The data set “doctorvisits.txt” on the course website consists of 3 different variables. The first column is an indicator denoting if the respondent is female. The second column is age and the final column is the number of physician visits during the past 3 months.

Suppose a Poisson model is applied to this data in the following way:
Pr(Visits_i = j|X_i) = \frac{\exp^{-\theta_i} \theta_i^j}{j!}, \quad j = 0, 1, 2, \ldots

where

\theta_i = \exp(\beta_0 + \beta_1 female_i + \beta_2 age_i).

Fit the Poisson model to this sample of data. Obtain the MLE estimate of the vector \( \beta \).

(1f) Based on your estimation results, what is the probability that a 35 year old female will have had more than 2 physician visits during the past 3 months?

(1g) Plot the estimated probability of the number of physician visits for each element of the set: \{0, 1, \ldots, 10\}, first assuming the individual is male and 40 years of age. Alongside this graph, do the same thing, but this time for a female respondent who is also 40 years of age.

(1h) Suppose we are interested in the following marginal effect: The change in the expected number of physician visits associated with a one year increase in age. (For this purpose, we can treat age as if it were a continuous variable). Provide a point estimate of this effect for both a 50 year old male and a 50 year old female.

(1i) In the raw data, what is the overall probability that an individual (not differentiating by gender) will have exactly 0 physician visits? What is the predicted probability of your model that a 25 year old male and a 25 year old female will have 0 physician visits? (Note that 25 is the youngest age observed in the sample).

Suppose, instead, we model the count outcome \( Y \) as follows:

\[ Y \mid u \sim \begin{cases} 0 & \text{if } u \leq \lambda \\ \text{Poisson}(\theta) & \text{if } u > \lambda \end{cases} \]

where \( \sim \) means “is distributed as,” \( u \) is random variable that is uniformly distributed on the unit interval, \( \lambda \) is a parameter \( \in (0, 1) \) and \( \text{Poisson}(\theta) \) is simply the Poisson distribution with parameter \( \theta \) that we have been considering in (2a)-(2f). When \( u \leq \lambda \), the above implies that \( Y = 0 \) whereas when \( u > \lambda \), \( Y \) is generated from a Poisson \( (\theta) \) distribution.

Show, in this model, that

\[ \Pr(Y = 0) = \lambda + (1 - \lambda) \exp(-\theta) \]

\[ \Pr(Y = j) = (1 - \lambda) \exp(-\theta) \frac{\theta^j}{j!}, \quad j = 1, 2, \ldots \]
(1j) Now, consider applying a variant of the mixture model discussed in (1i) to the doctor visit data. Specifically, consider the model:

\[
\Pr(\text{Visits}_i = 0|X_i) = \Phi(\delta_0) + (1 - \Phi(\delta_0)) \exp(-\theta_i)
\]

\[
\Pr(\text{Visits}_i = j|X_i) = (1 - \Phi(\delta_0)) \exp(-\theta_i) \frac{\theta_i^j}{j!}, \quad j = 1, 2, \ldots,
\]

\[
\theta_i = \exp(\delta_1 + \delta_2\text{female}_i + \delta_3\text{age}_i).
\]

Note that the above has incorporated the constraint that \( \lambda \in (0, 1) \). That is, we have parameterized \( \lambda \) as \( \Phi(\delta_0) \), where \( \Phi(\cdot) \) denotes the standard normal cdf. Note that the command \texttt{normcdf} in matlab will calculate this normal cumulative distribution function.

Write a MATLAB program that calculates the MLE of the vector \( \delta \). Then, use your estimated results to recalculate the quantities in (1i): the probabilities of zero visits for both a 25 year old male and female.