1. (25 pt) In a genetics experiment, investigators looked at 300 chromosomes of a particular type and counted the number of sister-chromatid exchanges on each. The number of exchanges \(x\) and the observed counts \(y\) are shown in the code below, along with a call to \texttt{glm}.

\[
\begin{align*}
    x & \leftarrow 0:9 \\
    y & \leftarrow c(6,24,42,59,62,44,41,14,6,2) \\
    \text{fit} & \leftarrow \text{glm}(y \sim x + \text{offset}(-\text{lgamma}(y+1)), \text{family=poisson})
\end{align*}
\]

Part of the results are summarized below (\texttt{summary(fit)}).

\begin{verbatim}
Coefficients: Estimate Std. Error z value Pr(>|z|)
(Intercept) 1.78683 0.13084 13.66 <2e-16 ***
x 1.36715 0.03029 45.14 <2e-16 ***
---
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
(Dispersion parameter for poisson family taken to be 1)

Null deviance: 1436.04 on 9 degrees of freedom
Residual deviance: 6.22 on 8 degrees of freedom
\end{verbatim}

(a) Assuming a Poisson distribution for the counts, obtain the \(\lambda\) estimate from the \texttt{glm} fit. (8 pt)

(b) Assuming a Poisson distribution for the counts, obtain the MLE of \(\lambda\) directly from the data. Provide a number with at least 4 figures or a line of \texttt{R} code. (10 pt)

(c) Based on the \texttt{glm} fit output, can you assess the adequacy of a Poisson distribution for the counts? If yes, do so by citing relevant information. If not, explain why not. (7 pt)

2. (30 pt) The admission to a department of UC Berkeley yielded the following data.

\begin{center}
\begin{tabular}{ccc}
& Female & Male & Total \\
Rejected & 299 & 138 & 437 \\
Admitted & 94 & 53 & 147 \\
Total & 393 & 191 & 584
\end{tabular}
\end{center}

Define

\[
    x_1 = \begin{cases} 
      1, & \text{Admitted} \\
      0, & \text{Rejected}
    \end{cases} \quad x_2 = \begin{cases} 
      1, & \text{Female} \\
      0, & \text{Male}
    \end{cases}
\]

Treat the counts as Poisson random variables with mean \(\lambda\), and consider a log linear model of the form

\[
    \log(\lambda) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2
\]

(a) Fit the model to the data. Estimate \(\beta_3\) and obtain the deviance of the fit. (10 pt)

(b) Assuming \(\beta_3 = 0\), fit the additive model to the data. Estimate \(\beta_0, \beta_1, \beta_2\). (10 pt)

(c) Calculate the deviance of the additive model and test for \(H_0 : \beta_3 = 0\). (10 pt)
3. (45 pt.) The survival of a cohort of patients with Acute Myelogenous Leukemia was observed as follows,

\[ \begin{array}{ccccccccccc}
9 & 13 & 13+ & 18 & 23 & 28+ & 31 & 34 & 45+ & 48 & 161+ \\
\end{array} \]

where the “+” indicates censored observation.

(a) Calculate the Nelson estimate of \( \Lambda(30) \) and its standard error; use the Tsiatis formula for standard error. (10 pt)

(b) Construct a 95% confidence interval for \( \Lambda(30) \) based on your calculations in (a). Be sure that your solution is meaningful. (5 pt)

(c) Assuming an exponential distribution, calculate the MLE of \( \Lambda(30) \). (10 pt)

(d) Assuming an exponential distribution, construct a 95% confidence interval for \( \Lambda(30) \). (10 pt)

(e) Test \( H_0 : \Lambda(t) = 0.02t \) using a one-sample logrank test. (10 pt)