SOLUTIONS

1. (a) A model with only an intercept is fitted by Poisson regression (hence the fitted model is the same as the *null model*).

From model `m1` we find the deviance as $D_1 = 15.94$, and $\chi^2_{0.05}(13) = 22.36$. Hence, the fitted model cannot be rejected, we do not reject the hypothesis of independent Poisson distributed observations.

A model including an offset, model `m2`, yields the deviance $D_2 = 16.06$. The usual F test of deviances cannot be performed (degrees of freedom the same for both models), but the difference seems slight (a slightly worse fit for the more complex model). We compare estimates of the dispersion parameter:

$$\text{Model } \texttt{m1}: \qquad \hat{\sigma}^2 = \frac{D_1}{n-1} = \frac{15.94}{13} = 1.24,$$

$$\text{Model } \texttt{m2}: \qquad \hat{\sigma}^2 = \frac{D_2}{n-1} = \frac{16.06}{13} = 1.23.$$

In conclusion, inclusion of the offset does not affect the fit very much. This could be explained from data, since the values for miles are rather similar for the different years.

(b) *Feel free to use a consistent notation as there are many equally good styles. But show the derivation.*

We want to find the ML Estimate $\widehat{\lambda}_n$ of λ and produce a $1-\alpha$ confidence interval for λ.

The MLE can be obtained as follows:

The likelihood function is:

$$L(\lambda) := L(x_1, x_2, \dots, x_n; \lambda) = \prod_{i=1}^{n} f(x_i; \lambda) = \prod_{i=1}^{n} e^{-\lambda} \frac{\lambda_i^x}{x_i!}$$

Hence, the log-likelihood function is:

$$\begin{aligned}
\ell(\theta) := \log(L(\lambda)) &= \log\left(\prod_{i=1}^{n} e^{-\lambda}\frac{\lambda^{x_i}}{x_i!}\right) = \sum_{i=1}^{n} \log\left(e^{-\lambda}\frac{\lambda^{x_i}}{x_i!}\right) = \sum_{i=1}^{n}\left(\log(e^{-\lambda}) + \log(\lambda^{x_i}) - \log(x_i!)\right) \\
&= \sum_{i=1}^{n}\left(-\lambda + x_i\log(\lambda) - \log(x_i!)\right) = \sum_{i=1}^{n} -\lambda + \sum_{i=1}^{n} x_i \log(\lambda) - \sum_{i=1}^{n}\log(x_i!) \\
&= n(-\lambda) + \log(\lambda)\left(\sum_{i=1}^{n} x_i\right) - \sum_{i=1}^{n}\log(x_i!)
\end{aligned}$$

Next, take the derivative of $\ell(\lambda)$:

$$\frac{\partial}{\partial\lambda}\ell(\lambda) = \frac{\partial}{\partial\lambda}\left(n(-\lambda) + \log(\lambda)\left(\sum_{i=1}^{n} x_i\right) - \sum_{i=1}^{n}\log(x_i!)\right) = n(-1) + \frac{1}{\lambda}\left(\sum_{i=1}^{n} x_i\right) + 0$$

and set it equal to 0 to solve for λ, as follows:

$$0 = n(-1) + \frac{1}{\lambda}\left(\sum_{i=1}^{n} x_i\right) + 0 \iff n = \frac{1}{\lambda}\left(\sum_{i=1}^{n} x_i\right) \iff \lambda = \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right) = \overline{x}_n$$

Finally, the ML Estimator of λ is $\widehat{\Lambda}_n = \overline{X}_n$ and the ML estimate is $\widehat{\lambda}_n = \overline{x}_n$.

Now, we want an $1-\alpha$ confidence interval for λ using the $\widehat{\mathsf{se}}_n \approxeq \sqrt{1/I_n(\widehat{\lambda}_n)}$ that is based on the Fisher Information $I_n(\lambda) = nI_1(\lambda)$. We need I_1. Since $X_1, X_2, \ldots, X_n \sim Poisson(\lambda)$, we have discrete RVs:

$$I_1 = -\sum_{x=0}^{\infty}\left(\frac{\partial^2 \log(f(x;\lambda))}{\partial^2\lambda}\right) f(x;\lambda) = -\sum_{x=0}^{\infty}\left(\frac{\partial^2 \log(f(x;\lambda))}{\partial^2\lambda}\right) f(x;\lambda)$$

First find

$$\begin{aligned}
\frac{\partial^2 \log(f(x;\lambda))}{\partial^2\lambda} &= \frac{\partial}{\partial\lambda}\left(\frac{\partial}{\partial\lambda}\log\left(f(x;\lambda)\right)\right) = \frac{\partial}{\partial\lambda}\left(\frac{\partial}{\partial\lambda}\log\left(e^{-\lambda}\frac{\lambda^x}{x!}\right)\right) \\
&= \frac{\partial}{\partial\lambda}\left(\frac{\partial}{\partial\lambda}\left(-\lambda + x\log(\lambda) - \log(x!)\right)\right) = \frac{\partial}{\partial\lambda}\left(-1+\frac{x}{\lambda}-0\right) = -\frac{x}{\lambda^2}
\end{aligned}$$

Now, substitute the above expression into the right-hand side of I_1 to obtain:

$$I_1 = -\sum_{x=0}^{\infty}\left(-\frac{x}{\lambda^2}\right) f(x;\lambda) = \frac{1}{\lambda^2}\sum_{x=0}^{\infty}(x)\, f(x;\lambda) = \frac{1}{\lambda^2}\sum_{x=0}^{\infty}(x)\, e^{-\lambda}\frac{\lambda^x}{x!} = \frac{1}{\lambda^2}E_\lambda(X) = \frac{1}{\lambda^2}\lambda = \frac{1}{\lambda}$$

In the third-to-last step above, we recognise the sum as the expectation of the $Poisson(\lambda)$ RV X, namely $E_\lambda(X) = \lambda$. Therefore, the estimated standard error is:

$$\widehat{\mathsf{se}}_n \approxeq \sqrt{1/I_n(\widehat{\lambda}_n)} = \sqrt{1/(nI_1(\widehat{\lambda}_n))} = \sqrt{1/(n(1/\widehat{\lambda}_n))} = \sqrt{\widehat{\lambda}_n/n}$$

and the approximate $1-\alpha$ confidence interval is

$$\widehat{\lambda}_n \pm z_{\alpha/2}\widehat{\mathsf{se}}_n = \widehat{\lambda}_n \pm z_{\alpha/2}\sqrt{\widehat{\lambda}_n/n}$$

(c) From the Table the MLE of *lambda* is the sample mean of the collissions:

$$\widehat{\lambda}_n = \bar{x}_n = \frac{1}{14}(3+6+4+7+6+2+2+4+1+7+3+5+6+1) = \frac{57}{14} = 4.07143$$

(d) The approximate 95% confidence interval for λ is

$$\widehat{\lambda}_n \pm z_{\alpha/2}\sqrt{\widehat{\lambda}_n/n} = 4.07 \pm 1.96 * \sqrt{4.07/14} = 4.07 \pm 1.0568 = (3.0132, 5.1267).$$

(e) The estimated intercept term is merely $\log(\widehat{\lambda}_n)$:

$$\log(57/4) = 1.404$$

2. (a) We can write $Y_i \sim \text{Po}(\mu_i)$, $i = 1, \ldots, 10$, where $\mu_i = r_i t_i$. Here r_i is the rate (expected fatals per passenger miles) and t_i the observation distance (passenger miles). Using the canonical link we find

$$\ln r_i = \beta_0 + \beta_1 x_i$$

where x_i is the regressor of year i. Using the offset, we find the formulation

$$\ln \mu_i = \ln(r_i t_i) = \ln t_i + \ln r_i = \ln t_i + \beta_0 + \beta_1 x_i.$$

In the R code related to model `m1`, we have

Y	`fatals`
x_i	`time`
t_i	`miles`

(b) In both cases the time trend is highly significant judging from the estimated coefficient for `time`. Moreover, the estimate is negative in both cases (-0.10 and -0.06, respectively), indicating that air travel was becoming safer over the period.

(c) Model 1: $D_1 = 5.4551 < 15.51 = \chi^2_{0.05}(8)$. Hence, we have a good fit. Model 2: $D_2 = 1051.4 > 15.51$. Here we have a bad fit, due to overdispersion. This large degree of overdispersion is due to compounding with the aircraft size, as each fatal accident leads to some multiple number of deaths. So, while the *fatal accident* data is consistent with an underlying Poisson process giving Poisson counts for the number of fatal accidents each year, the number of *passenger deaths* each year is a Poisson sum of random variables from the aircraft size distribution, which is not Poisson distributed.

3. (a) Check the residual deviance: $D_1(y;\mu(\hat{\beta})) = 226.52 > \chi^2_{0.05}(8) = 15.50$. This is a significant result, the fit is not adequate. (P-value: $4.2 \cdot 10^{-4}$.)

(b) We have nested models and perform a F test of deviances. Consider $F = [(D_1 - D_2)/(f_1 - f_2)/[D_2/f_2] = [(226.52 - 226.44)/(8-7)]/[226.44/7] = 0.0025$. Compare with $F_{0.05}(1,7) = 5.59$. Do not reject the simpler model.

(c) Estimate of dispersion parameter for model 1: $D_1(y;\mu(\hat{\beta}))/(n-k) = 226.52/8 = 28.315$. Might try a quasi-binomial approach.

Remark. We here give, for instructional reasons, the data frame (called `caesar`) and related R calls. Note that weights play an important role in this problem: each element in the vector `weight` gives the number of observations in the data set for the combination of other variables. They are included in the calls to `glm`.

```
caesar = data.frame(planned=factor(c(rep(1,8),rep(0,8))),
antibio = factor(c(rep(1,4),rep(0,4),rep(1,4),rep(0,4))),
risk = factor( rep(c(1,1,0,0),4)),
infection = factor(rep(c(1,0),8)),
weight=c(1,17,0,2,28,30,8,32,11,87,0,0,23,3,0,9))

m1 = glm(infection ~ antibio +
 planned+risk,family=binomial,weight=weight,data=caesar)

m2 = glm(infection ~ antibio + planned+risk +
 planned*antibio,family=binomial,weight=weight,data=caesar)
```

4. (a) We have that $Y_i \sim \text{Po}(\mu_i)$ with $\ln\mu_i = a + bx_i$, $i = 1,2,3$. and hence

$$\mathsf{P}(Y_i = y_i) = \exp(-\mu_i)\frac{\mu_i^{y_i}}{y_i!}$$

so the likelihood function is

$$L(a,b) = \prod_{i=1}^{n} \mathsf{P}(Y_i = y_i) = \exp(-\sum_{i=1}^{n}\mu_i)\prod_{i=1}^{n}\frac{\mu_i^{y_i}}{y_i!}$$

The log-likelihood function is then found as

$$\ell(a,b) = \sum_{i=1}^{3}\left[y_i(a+bx_i) - \exp(a+bx_i) - \ln(y_i!)\right]$$

(b) The derivatives of the log-likelihood are

$$\begin{aligned}\frac{\partial \ell}{\partial a} &= \sum_{i=1}^{3}(y_i - e^{a+bx_i}) \\ \frac{\partial \ell}{\partial b} &= \sum_{i=1}^{3}(y_i x_i - x_i e^{a+bx_i})\end{aligned}$$

The first score equation is equal to 0 when

$$\sum_{i=1}^{3} y_i = \sum_{i=1}^{3} e^{a+bx_i} = e^a \sum_{i=1}^{3} e^{bx_i}$$

which gives the relation

$$e^a = \frac{\sum_{i=1}^{3} y_i}{\sum_{i=1}^{3} e^{bx_i}} \qquad (*)$$

The second score equation is equal to 0 when (substituting from Eq. (*))

$$\begin{aligned}\sum_{i=1}^{3} y_i x_i &= \sum_{i=1}^{3} x_i e^{a+bx_i} = e^a \sum_{i=1}^{3} x_i e^{bx_i} \\ &= \frac{\sum_{i=1}^{3} y_i}{\sum_{i=1}^{3} e^{bx_i}} \sum_{i=1}^{3} x_i e^{bx_i}\end{aligned}$$

from which we find

$$\frac{\sum_{i=1}^{3} y_i x_i}{\sum_{i=1}^{3} y_i} = \frac{\sum_{i=1}^{3} x_i e^{bx_i}}{e^{bx_i}} \qquad (**)$$

Now insert numbers:

$$\begin{aligned}\sum_{i=1}^{3} y_i &= 3 + 4 + 13 = 20, \quad \sum_{i=1}^{3} y_i x_i = -3 + 13 = 10 \\ \sum_{i=1}^{3} e^{bx_i} &= e^{-b} + 1 + e^b, \quad \sum_{i=1}^{3} x_i e^{bx_i} = -e^{-b} + e^b\end{aligned}$$

Substituting into Eq. (**) gives

$$\frac{1}{2} = \frac{-e^{-b} + e^b}{e^{-b} + 1 + e^b} = \frac{-1 + e^{2b}}{1 + e^b + e^{2b}}$$

which can be rewritten as

$$1 + e^b + e^{2b} = 2(-1 + e^{2b}) \quad \Leftrightarrow \quad e^{2b} - e^b - 3 = 0,$$

i.e. a quadratic equation in e^b. The valid solution is

$$e^b = \frac{1 + \sqrt{13}}{2} \quad \Rightarrow \quad \hat{b} = \ln\Big(\frac{1 + \sqrt{13}}{2}\Big) \doteq 0.83.$$

Turning to Eq. (*), we may find the estimate of a, using that

$$\sum_{i=1}^{3} e^{\hat{b}x_i} = \frac{4 + 2\sqrt{13}}{3}$$

and hence

$$e^a = \frac{\sum_{i=1}^{3} y_i}{\sum_{i=1}^{3} e^{\hat{b}x_i}} = -\frac{10(2 - \sqrt{13})}{3} \quad \Rightarrow \quad \hat{a} = \ln\Big(\frac{10(\sqrt{13} - 2)}{3}\Big) \doteq 1.68.$$