3.1 We consider three nested regression models for modelling the number of car accidents according to region (`region`). The variables risk class (`risk`) has 3 categories and the number of years of driving experience (`exp`) is broken down into 4 categories.

Model	variables	$p+1$	$\ell(\widehat{\boldsymbol{\beta}})$	AIC	BIC
M_1	`risk`	3	−244.566	495.132	510.362
M_2	`risk + region`	⋆	−151.620	⋆	⋆
M_3	`risk + region + exp`	10	−139.734	299.468	350.235

Table 1: Goodness-of-fit measures for three nested regression models with the number of parameters in each model ($p+1$), the value of the log-likelihood function evaluated at the maximum likelihood estimate ($\ell(\widehat{\boldsymbol{\beta}})$) and information criteria.

What is the difference between AIC and BIC of Model M_2 (in absolute value)?

3.2 A random variable X follows a geometric distribution with parameter p if its probability mass function is

$$\mathsf{P}(X=x) = (1-p)^{x-1}p, \qquad x = 1,2,\ldots$$

(a) Write the likelihood and the log-likelihood of the n sample.
(b) Derive the maximum likelihood estimator for the parameter p.
(c) Compute the observed information matrix.
(d) Suppose we have a sample of 15 observations, $\{5,6,3,7,1,2,11,8,7,34,1,7,10,1,0\}$, whose sum is 103. Compute the maximum likelihood estimate and its approximate standard error.
(e) Compute the likelihood ratio and the Wald test statistics. Perform a test at level 5% for $\mathscr{H}_0: p_0 = 0.1$ against the two-sided alternative $\mathscr{H}_a: p_0 \neq 0.1$.

3.3 We consider the failure time of an engine based on its level of corrosion w. Failure time T is modelled with an exponential distribution with density $f(t) = \lambda\exp(-\lambda t)$, but whose rate parameter $\lambda = aw^b$; if $b=0$, the failure time is a^{-1}. Assume we have an n sample of independent observations with w_i assumed known. [Coles (2001)]
(a) Write down the log-likelihood of the model.
(b) Derive the observed and the Fisher information matrices.
(c) Show that the profile log-likelihood for b is

$$\ell_{\mathsf{p}}(b) = n\ln(\widehat{a}_b) + b\sum_{i=1}^n \ln(w_i) - \widehat{a}_b\sum_{i=1}^n w_i^b t_i,$$

and give an explicit formula for the partial maximum likelihood estimator $\widehat{a}_b$.

3.4 We consider a simple Poisson model for the number of daily sales in a store, which are assumed independent from one another. Your manager tells you the latter depends on whether the store is holding sales or not. The mass function of the Poisson distribution is

$$\mathsf{P}\left(Y_i = y_i \mid \texttt{sales}_i\right) = \frac{\exp(-\lambda_i)\lambda_i^{y_i}}{y_i!}, \qquad y_i = 0,1,\ldots$$

and we model $\lambda_i = \exp(\beta_0 + \beta_1\texttt{sales}_i)$, where $\texttt{sales}_i$ is a binary indicator equal to unity during sales and zero otherwise.
(a) Derive the maximum likelihood estimator of (β_0,β_1). *Hint: maximum likelihood estimators are invariant to reparametrization.*
(b) Calculate the maximum likelihood estimates for a sample of size 12, where the number of transactions outside sales is $\{2;5;9;3;6;7;11\}$, and during sales, $\{12;9;10;9;7\}$.

(c) Calculate the observed information matrix and use the latter to derive standard errors for $\widehat{\beta}_0$ and $\widehat{\beta}_1$ and a 95% confidence interval for the parameters.

(d) Your manager wants to know if the daily profits during sales period are different from those outside of the sales period. She calculates that the average profit during sales is \$20 per transaction, compared to \$25 normally. Test this hypothesis using a likelihood ratio test. *Hint: write the null hypothesis of equal profit in terms of the model parameters β_0 and β_1.* (**difficult**)