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`	$\star$	-151.620	$\star$	$\star$
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 216. 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 $\widehat{a}_b$.