3.1 On considère trois modèles de régression emboîtés afin de modéliser le nombre d'accidents de voiture selon la région (`region`). La variable catégorielle niveaux de risque (`risque`) a 3 niveaux et le nombre d'années d'expérience (`expe`) au volant est un facteur à 4 niveaux.

Modèle	variables	$p+1$	$\ell(\widehat{\boldsymbol{\beta}})$	AIC	BIC
M_1	`risque`	3	-244.566	495.132	510.362
M_2	`risque` + `region`	$\star$	-151.620	$\star$	$\star$
M_3	`risque` + `region` + `expe`	10	-139.734	299.468	350.235

Table 1: Mesures d'adéquation de trois modèles emboîtés avec les covariables, le nombre de coefficients du modèle ($p+1$), la valeur de la log-vraisemblance évaluée au maximum de vraisemblance ($\ell(\widehat{\boldsymbol{\beta}})$) et les critères d'information.

Quelle est la différence entre le AIC et le BIC du modèle M_2 (en valeur absolue)?

3.2 Une variable aléatoire X suit une loi géométrique de paramètre p si sa fonction de masse est

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

(a) Écrivez la vraisemblance et la log-vraisemblance d'un échantillon aléatoire de taille n si les observations sont indépendantes.

(b) Dérivez l'estimateur du maximum de vraisemblance pour le paramètre p.

(c) Calculez l'information observée.

(d) Supposons qu'on a un échantillon de 15 observations, $\{5, 6, 3, 7, 1, 2, 11, 8, 7, 34, 1, 7, 10, 1, 0\}$, dont la somme est 216. Calculez l'estimé du maximum de vraisemblance et son erreur-type approximative.

(e) Calculez la statistique du rapport de vraisemblance et la statistique de Wald pour un test à niveau 5% de l'hypothèse $\mathscr{H}_0: p_0 = 0.1$ contre l'alternative bilatérale $\mathscr{H}_a: p_0 \neq 0.1$.