1. (a) Let $X_1$ and $X_2$ be independent random variables, both with a geometric distribution. Let $Y = X_1 + X_2$. Give an expression for $\mathbb{P}(Y = k)$ and deduce that $Y$ has a negative binomial distribution with $r = 2$.

(b) Let $X_1, \ldots, X_r$ be independent random variables, all with a common geometric distribution with parameter $p$. It can be shown that $Y = X_1 + \cdots + X_r$ has a negative binomial distribution with parameters $r$ and $p$. It can also be shown that $\text{Var} X_1 = \frac{(1-p)}{p^2}$. Compute expectation and variance of $Y$.

(c) Let $f(k|p)$ be the probability mass function, $f(k|p) = \mathbb{P}(Y = k|p)$, and $\dot{l}(p) = \frac{\partial}{\partial p} \log f(Y|p)$. Show that $\mathbb{E} \dot{l}(p) = 0$. Let $I(p) = \mathbb{E} \dot{l}(p)^2$. Compute $I(p)$.

Let $Y_1, \ldots, Y_n$ be a sample from a negative binomial distribution with parameters $p$ (unknown) and $r$.

(d) Compute the maximum likelihood estimator $\hat{p}_n$ of $p$ and show that it is equal to the moment estimator.

(e) What is the asymptotic distribution of $\sqrt{n}(\hat{p}_n - p)$?

(f) Suppose that $n = 100$, $r = 4$, and that the sample is such that $\hat{p}_{100} = 0.75$. Give an approximate 95% confidence interval for $p$.

2. Let $X_1, \ldots, X_n$ be sample from a Poisson distribution with parameter $\lambda$ (unknown). Let $T = X_1 + \cdots + X_n$.

(a) Consider the simple hypothesis testing problem $H_0 : \lambda = \lambda_0$ against $H_A : \lambda = \lambda_1$, where $\lambda_0 > \lambda_1$. Show that the Neyman-Pearson test rejects $H_0$ for ‘small values’ of $T$, $T \leq c_n$ say, for some integer $c_n$.

(b) Show that the function $\lambda \mapsto \mathbb{P}(T \leq c_n|\lambda)$ is decreasing. Hint: Let $\lambda_1 < \lambda_2$, and let $U$ have a Poisson distribution with parameter $\lambda_1$ and $V$, independent of $U$, have a Poisson distribution with parameter $\lambda_2 - \lambda_1$. Use the trivial inequality $U + V \geq U$. 

(c) Let $\alpha = \mathbb{P}(T \leq c_n|\lambda_0)$. Consider the composite testing problem $H_0: \lambda \geq \lambda_0$ against $H_A: \lambda < \lambda_0$. We use (again) the test that rejects $H_0$ if $T \leq c_n$. Compute $\sup_{\lambda \geq \lambda_0} \mathbb{P}(T_n \leq c_n|\lambda)$, and deduce that this test has significance level $\alpha$.

(d) Is the above test uniformly most powerful for the testing problem under consideration?

(e) Replace $c_n$ with $\xi_n := n\lambda_0 + \xi\sqrt{n\lambda_0}$ for some (negative) real number $\xi$. Compute, by using the Central Limit Theorem (CLT) and in terms of the cdf $\Phi$, $\mathbb{P}(T \leq \xi_n|\lambda_0)$. How should one choose $\xi$ to have the last probability (approximately) equal to $\alpha$?

(f) Suppose that $n = 100$, $\lambda_0 = 1$. Give a numerical value for $\xi_n$, if $\alpha = 0.0202$. If $T = 90$ is observed, should one reject $H_0$?

(g) Fix some $\lambda_1 < \lambda_0$. Use the CLT again to show that the (asymptotic) power of the test, $\mathbb{P}(T \leq \xi_n|\lambda_1)$ is equal to $\Phi(\xi\sqrt{\lambda_0/\lambda_1} + (\lambda_0 - \lambda_1)\sqrt{n/\lambda_1})$. What happens with this probability as $n \to \infty$?

3. In this exercise we consider quadratic regression, we assume a model of the form $y_i = \beta_0 + \beta_1x_i + \beta_2x_i^2 + e_i$, $i = 1, \ldots, n$. The $e_i$ are assumed to be iid with a common normal $N(0, \sigma^2)$ distribution. In matrix notation, we summarize the model by writing

\[ Y = X\beta + e, \]

following the usual conventions.

(a) How would you cast this model as an ordinary linear regression model by choosing the right independent variables?

(b) We know that it is important that $X$ has rank 3. Show that this is the case if at least three of the $x_i$ ($x_1, x_2, x_3$ for instance) are different. Hint: compute the determinant

\[
\begin{vmatrix}
1 & x_1 & x_1^2 \\
1 & x_2 & x_2^2 \\
1 & x_3 & x_3^2 \\
\end{vmatrix}.
\]

(c) Show that the rank of $X$ is at most 2, if all the $x_i$ assume at most two different values.
(d) Let \( \hat{\beta} \) be the least squares estimator of \( \beta \) and \( \hat{Y} = X\hat{\beta} \). Show that \( Y - \hat{Y} = Qe \), where \( Q = I - X(X^\top X)^{-1}X^\top \). Show also that \( Q^2 = Q \).

(e) Determine a matrix \( M \) such that

\[
U := \begin{pmatrix} Y - \hat{Y} \\ \hat{\beta} - \beta \end{pmatrix} = Me.
\]

(f) Show that \( U \) has covariance matrix equal to \( \sigma^2 \begin{pmatrix} Q & 0 \\ 0 & (X^\top X)^{-1} \end{pmatrix} \).

Are \( Y - \hat{Y} \) and \( \hat{\beta} \) independent?

(g) Let \( S^2 = (Y - \hat{Y})^\top(Y - \hat{Y}) \). It is known that \( \frac{S^2}{\sigma^2} \) has a \( \chi^2 \)-distribution with \( n - 3 \) degrees of freedom. Use this to deduce that

\[
\frac{\hat{\beta}_i - \beta_i}{s_{\hat{\beta}_i}}
\]

has a \( t \)-distribution with \( n - 3 \) degrees of freedom, where \( s_{\hat{\beta}_i} = S \sqrt{(X^\top X)_{ii}^{-1}} \).

(h) Suppose that \( n = 20 \) and that computations with the data result in \( \hat{\beta}_2 = s_{\hat{\beta}_2} = 0.42 \). Give a 95% confidence interval for \( \beta_2 \).

(i) Suppose that one wants to test the hypothesis that the regression is linear in one variable. Formulate this as a testing problem on the coefficients \( \beta_i \). Should one reject this hypothesis in the situation of the previous part?