Problem 2.1. Find MLE for the following parameters:
(a) Probability of success p in Bernoulli(p) model

Solution (a). Let X be a bernoulli random variable with parameter p. Let X_1,...,X_n be the random sample of X.
PDF for the bernoulli distribution with parameter p is
\[ f(x) = p^x(1 - p)^{1-x} \quad x = 0, 1 \]

Likelihood function of the sample:
\[ = \prod_{i=1}^{n} f(x_i, p) \]
\[ = \prod_{i=1}^{n} p^{x_i}(1 - p)^{1-x_i} \]
\[ = p^{\sum_{i=1}^{n} x_i}(1 - p)^{n - \sum_{i=1}^{n} x_i} \]

Log-likelihood function of the sample:
\[ = \ln \left( p^{\sum_{i=1}^{n} x_i}(1 - p)^{n - \sum_{i=1}^{n} x_i} \right) \]
\[ = \left( \sum_{i=1}^{n} x_i \right) \ln(p) + \left( n - \sum_{i=1}^{n} x_i \right) \ln(1 - p) \]

Maximum value of this function can be found by taking the derivative of the above function with respect to p, and setting it to 0.
\[ = \frac{\sum_{i=1}^{n} x_i}{p} - \left( \frac{n - \sum_{i=1}^{n} x_i}{1 - p} \right) = 0 \]
gives that
\[ p = \frac{\sum_{i=1}^{n} x_i}{n} \]
(b) Probability of success $p$ in Binomial($n,p$) model

**Solution (b).** Let $X$ be a binomial random variable with parameters $n$ and $p$. Let $X_1,...,X_m$ be the random sample of $X$.

PDF for the binomial distribution with parameters $n$ and $p$ is

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x} \quad x = 0, 1, .., n.$$ 

Likelihood function of the sample:

$$= \prod_{i=1}^{m} f(x_i, p)$$

$$= \prod_{i=1}^{m} \binom{n}{x_i} p^{x_i} (1-p)^{n-x_i}$$

Log-likelihood function of the sample:

$$= \ln \left( \prod_{i=1}^{m} \binom{n}{x_i} p^{x_i} (1-p)^{n-x_i} \right)$$

$$= \sum_{i=1}^{m} \ln \left( \binom{n}{x_i} \right) + \left( \sum_{i=1}^{m} x_i \right) \ln(p) + \left( mn - \sum_{i=1}^{m} x_i \right) \ln(1-p)$$

Maximum value of this function can be found by taking the derivative of the above function with respect to $p$, and setting it to 0.

$$= 0 + \sum_{i=1}^{m} \frac{x_i}{p} - \left( \frac{mn - \sum_{i=1}^{m} x_i}{1-p} \right) = 0$$

gives that $p = \frac{\sum_{i=1}^{m} x_i}{mn}$

(c) Probability of success $p$ in Geometric($p$) model

**Solution (c).** Let $X$ be a geometric random variable with parameter $p$. Let $X_1,...,X_n$ be the random sample of $X$.

PDF for the geometric distribution with parameter $p$ is

$$f(x) = p(1-p)^{x-1} \quad x = 1, 2, ...$$
Likelihood function of the sample:

\[ L(p) = \prod_{i=1}^{n} f(x_i, p) \]

\[ = \prod_{i=1}^{n} p(1 - p)^{x_i - 1} \]

\[ = p^n (1 - p)^{\sum x_i - n} \]

Log-likelihood function of the sample:

\[ \ell(p) = \ln \left( p^n (1 - p)^{\sum x_i - n} \right) \]

\[ = n \ln(p) + \left( \sum_{i=1}^{m} x_i - n \right) \ln(1 - p) \]

Maximum value of this function can be found by taking the derivative of the above function with respect to \( p \), and setting it to 0.

\[ \frac{n}{p} + \left( \frac{n - \sum_{i=1}^{m} x_i}{1 - p} \right) = 0 \]

gives that \( p = \frac{n}{\sum_{i=1}^{m} x_i} \).

(d) Intensity \( \lambda \) in Poisson(\( \lambda \)) model

**Solution (d).** Let \( X \) be a Poisson random variable with parameter \( \lambda \). Let \( X_1, \ldots, X_n \) be the random sample of \( X \).

PDF for the poisson distribution with parameter \( \lambda \) is

\[ f(x) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0, 1, \ldots \]

Likelihood function of the sample:

\[ L(\lambda) = \prod_{i=1}^{n} f(x_i, \lambda) \]

\[ = \prod_{i=1}^{n} \frac{e^{-\lambda} \lambda^{x_i}}{x_i!} \]

\[ = e^{-\lambda n} \lambda^{\sum x_i} C \quad \text{where} \quad C = \prod_{i=1}^{n} \frac{1}{x_i!} \]
Log-likelihood function of the sample:

\[
\log \left( e^{-\lambda n} \lambda^{\sum_{i=1}^{m} x_i} C \right)
\]

\[
= -n \lambda + \left( \sum_{i=1}^{m} x_i \right) \ln(\lambda) + \ln(C)
\]

Maximum value of this function can be found by taking the derivative of the above function with respect to \( \lambda \), and setting it to 0.

\[
= -n + \left( \frac{\sum_{i=1}^{m} x_i}{\lambda} \right) + 0 = 0
\]

gives that \( \lambda = \frac{\sum_{i=1}^{m} x_i}{n} \).

**Problem 2.2.** Find the Fisher information for the models of problem 2.1

(a) Probability of success \( p \) in Bernoulli(\( p \)) model

**Solution (a).** The Fisher Information for parameter \( \hat{p} \) can be written two ways: \( I(\hat{p}) = E[(\frac{\partial}{\partial p} \ln f(x,p))^2 | p] \) or \( I(\hat{p}) = -E[\frac{\partial^2}{\partial p^2} \ln f(x,p) | p] \).

From problem 2.1 (a) we know

\[
\frac{\partial}{\partial p} \ln f(x,p) = \frac{\partial}{\partial p} [xp + (1-x) \ln (1-p)]
\]

\[
= \frac{x}{p} - \frac{1-x}{1-p}
\]

Recall that \( E(X) = p \) where \( X \) is the bernoulli random variable.

Therefore,

\[
I(\hat{p}) = -E \left[ \frac{\partial^2}{\partial p^2} \ln f(x,p) \right]
\]

\[
= -E \left[ -\frac{X}{p^2} - \left( \frac{1-X}{q^2} \right) \right]
\]

\[
= \frac{E(X)}{p^2} + \frac{E(1-X)}{(1-p)^2}
\]

\[
= \frac{1}{p} + \frac{1}{1-p}
\]

gives \( \frac{1}{p(1-p)} \).
(b) Probability of success $p$ in Binomial$(n, p)$ model

**Solution (b).** The Fisher Information for parameter $\hat{p}$ can be written two ways: $I(\hat{p}) = E\left[\left(\frac{\partial}{\partial p} \ln f(x, p)\right)^2\right]$ or $I(\hat{p}) = -E\left[\frac{\partial^2}{\partial p^2} \ln f(x, p)\right]$. From problem 2.1 (b) we know

$$\frac{\partial}{\partial p} \ln f(x, p) = \frac{\partial}{\partial p} \left[\ln \left(\frac{n}{x}\right) + x \ln p + (n - x) \ln (1 - p)\right]$$

$$= \frac{x}{p} - \frac{n - x}{1 - q}$$

$$\left(\frac{x}{p} - \frac{n - x}{1 - q}\right)^2 = \frac{x^2}{p^2 q^2} - 2 \frac{x}{q^2 p} + \frac{n^2}{q^2}$$

where $q = (1 - p)$

Thus,

$$E\left[\frac{X^2}{p^2 q^2} - 2 \frac{x}{q^2 p} + \frac{n^2}{q^2}\right]$$

Recall that $E(X) = np$ and $E(X^2) = npq + n^2 p^2$ where $X$ is the binomial random variable. Therefore,

$$= \frac{npq}{p^2 q^2} + \frac{n^2 p^2}{p^2 q^2} - 2 \frac{n^2 p}{q^2 p} + \frac{n^2}{q^2}$$

$$= \frac{n}{pq} + \frac{n^2}{q^2} - 2 \frac{n^2}{q^2} + \frac{n^2}{q^2}$$

$$= \frac{n}{pq}$$

(c) Probability of success $p$ in Geometric$(p)$ model

**Solution (c).** The Fisher Information for parameter $\hat{p}$ can be written two ways: $I(\hat{p}) = E\left[\left(\frac{\partial}{\partial p} \ln f(x, p)\right)^2\right]$ or $I(\hat{p}) = -E\left[\frac{\partial^2}{\partial p^2} \ln f(x, p)\right]$. From problem 2.1 (c) we know

$$\frac{\partial}{\partial p} \ln f(x, p) = \frac{\partial}{\partial p} \left[\ln p + (x - 1) \ln (1 - p)\right]$$

$$= \frac{1}{p} - \left(\frac{x}{1 - p}\right)$$

Recall that $E(X) = \frac{1}{p}$ where $X$ is the geometric random variable.
Therefore,

\[ I(\hat{p}) = -E\left[ \frac{\partial^2}{\partial p^2} \ln f(x, p) \right] \]
\[ = -E\left[ -\frac{1}{p^2} - \left( \frac{X}{1-p^2} \right) \right] \]
\[ = \frac{1}{p^2} + \frac{E(X)}{(1-p)^2} \]
\[ = \frac{1}{p^2} + \frac{1}{p(1-p)^2} \]
\[ = \frac{1-p}{p^2(1-p)^2} \]

gives \( \frac{1}{p^2(1-p)} \)

(d) Intensity \( \lambda \) in Poisson(\( \lambda \)) model

**Solution (d).** The Fisher Information for parameter \( \hat{\lambda} \) can be written two ways: \( I(\hat{\lambda}) = E\left[ \left( \frac{\partial}{\partial \lambda} \ln f(x, \lambda) \right)^2|\lambda \right] \) or \( I(\hat{\lambda}) = -E\left[ \frac{\partial^2}{\partial \lambda^2} \ln f(x, \lambda) \right] \).

From problem 2.1 (d) we know

\[ \frac{\partial}{\partial p} \ln f(x, \lambda) = \frac{\partial}{\partial p} \left[ -\lambda + x\ln\lambda + \ln \frac{x}{x!} \right] \]
\[ = \frac{x}{\lambda} - 1 \]
\[ \frac{\partial^2}{\partial \lambda^2} \ln f(x, \lambda) = -\frac{x}{\lambda^2} \]

Recall that \( E(X) = \lambda \) where \( X \) is the poisson random variable with parameter \( \lambda \).

Therefore,

\[ I(\lambda) = -E\left[ \frac{\partial^2}{\partial p^2} \ln f(x, p) \right] \]
\[ = -E\left( \frac{-X}{\lambda^2} \right) \]
\[ = \frac{1}{\lambda^2} E(X) \]
\[ = \frac{\lambda}{\lambda^2} \]

gives \( \frac{1}{\lambda} \)
Problem 2.3. Assess the efficiency of the MLE estimators for the models of problem 2.1.

(a) Probability of success \( p \) in Bernoulli(\( p \)) model

**Solution (a).** According to Cramer-Rao inequality \( \text{Var}(\hat{p}) \geq \frac{1}{nI(\hat{p})} \)

It states that if there is an estimator that achieves this lower bound, then it is the efficient estimator for that parameter.

From problem 2.1 (a) we have that \( \hat{p} = \frac{\sum x_i}{n} \). So, the variance of this parameter is

\[
\text{var}(\hat{p}) = \text{var} \left( \frac{\sum_{i=1}^{n} x_i}{n} \right) = \frac{\text{var} \sum_{i=1}^{n} x_i}{n^2} = \frac{n p (1 - p)}{n^2} = \frac{p(1 - p)}{n}
\]

Also, from problem 2.2 (a), \( I(\hat{p}) = \frac{1}{p(1-p)} \). Hence \( \hat{p} \) is an efficient estimator.

(b) Probability of success \( p \) in Binomial(\( n, p \)) model

**Solution (b).** From problem 2.1 (b) we have that \( \hat{p} = \frac{\sum_{i=1}^{m} x_i}{mn} \). So, the variance of this parameter is

\[
\text{var}(\hat{p}) = \text{var} \left( \frac{\sum_{i=1}^{m} x_i}{mn} \right) = \frac{\text{var} \sum_{i=1}^{m} x_i}{m^2n^2} = \frac{mnp(1 - p)}{m^2n^2} = \frac{p(1 - p)}{mn}
\]

Also, from problem 2.2 (b), \( I(\hat{p}) = \frac{n}{p(1-p)} \). Hence \( \hat{p} \) is an efficient estimator.

(c) Probability of success \( p \) in Geometric(\( p \)) model

**Solution (c).** From problem 2.1 (c) we have that \( \hat{p} = \frac{n}{\sum_{i=1}^{m} x_i} = \frac{1}{n} \). Variance of this estimator can be calculated using delta method.
Delta method says that: if $\theta$ follows normal distribution asymptotically with mean $= \theta$ and variance $= \sigma^2$ then $f(\theta)$ follows normal distribution asymptotically with mean $f(\theta)$ and variance $f'(\theta)^2 \sigma^2$.

From CLT we know that $X$ follows normal distribution if $n$ is large. Hence $X$ follows normal distribution asymptotically with mean $= \frac{1}{p}$ and variance $= \frac{1-p}{n p^2}$. Here the function is $f(\theta) = \frac{1}{\bar{X}}$ and $f'(\theta) = -\frac{1}{\bar{X}^2}$. From the delta method, mean of $f(\theta)$ is $p$ and variance of $f(\theta)$ is $p^4 \frac{1-p}{n p^2} = \frac{(1-p)p^2}{n}$. So, the variance of this parameter is

$$\text{var}(\hat{p}) = \frac{(1-p)p^2}{n}$$

Also, from problem 2.2 (c), $I(\hat{p}) = \frac{1}{p^2(1-p)}$. Hence $\hat{p}$ is an efficient estimator.

**Solution (d).** From problem 2.1 (d) we have that $\hat{\lambda} = \frac{\sum_{i=1}^{n} x_i}{n}$. So, the variance of this parameter is

$$\text{var}(\hat{\lambda}) = \text{var} \left( \frac{\sum_{i=1}^{n} x_i}{n} \right)$$

$$= \text{var} \sum_{i=1}^{n} x_i$$

$$= \frac{\text{var} \sum_{i=1}^{n} x_i}{n^2}$$

$$= \frac{n \lambda}{n^2}$$

$$= \frac{\lambda}{n}$$

Also, from problem 2.2 (d), $I(\hat{\lambda}) = \frac{1}{\lambda}$. Hence $\hat{\lambda}$ is an efficient estimator.