Exercise Series 7 (Solutions)

Exercise 1 (a) For $X \sim \text{Bernoulli}(p)$, $\mathbb{E}_p(X) = p$

So, for method of moments, we equate : $p = \bar{X} \Rightarrow \hat{p} = \bar{X}$

Therefore, $\hat{p}_{MoM} = \bar{X}$.

In this example, $L(p) = L(p|\underset{\sim}{x}) = p^{\sum_{i=1}^{n} x_i} (1-p)^{n - \sum_{i=1}^{n} x_i}$, $0 \le p \le 1$

$$\ell(p) = \log L(p) = \sum_{i=1}^{n} x_i \log p + \left(n - \sum_{i=1}^{n} x_i\right) \log(1-p), \quad 0 \le p \le 1.$$

$$\ell'(p) = \frac{\sum_{i=1}^{n} x_i}{p} - \frac{\left(n - \sum_{i=1}^{n} x_i\right)}{(1-p)}$$

$$\ell''(p) = -\frac{\sum_{i=1}^{n} x_i}{p^2} - \frac{\left(n - \sum_{i=1}^{n} x_i\right)}{(1-p)^2} < 0$$

$$\ell'(p) = 0 \iff \frac{\sum_{i=1}^{n} x_i}{n - \sum_{i=1}^{n} x_i} = \frac{p}{1-p} \Rightarrow p = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{x}$$

$\therefore$ $\ell(p)$ is maximized at $p = \bar{x}$

$\therefore$ MLE of p is $\hat{p}_{MLE} = \bar{X}$.

In this example, $\hat{p}_{MoM} = \hat{p}_{MLE}$. Both are functions of $\sum_{i=1}^{n} X_i$, which is minimal sufficient for p.

(b) For $X \sim \text{Poisson}(\lambda)$, $\mathbb{E}_\lambda(X) = \lambda$

So, MoM estimator is obtained by solving $\lambda = \bar{X}$
$\Rightarrow \hat{\lambda}_{MoM} = \bar{X}$.

$$L(\lambda) = e^{-n\lambda} \frac{\lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!}$$

$$\ell(\lambda) = -n\lambda + \sum_{i=1}^{n} x_i \log \lambda - \sum_{i=1}^{n} \log(x_i!)$$

$$\ell'(\lambda) = -n + \frac{\sum_{i=1}^{n} x_i}{\lambda}$$

$$\ell''(\lambda) = -\frac{\sum_{i=1}^{n} x_i}{\lambda^2} < 0$$

$$\ell'(\lambda) = 0 \Rightarrow \lambda = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{x}$$

$\therefore \hat{\lambda}_{MLE} = \bar{X} = \hat{\lambda}_{MoM}$. Both are functions of the minimal sufficient statistic $\sum_{i=1}^{n} X_i$

(c) $X \sim \text{Geometric}(p)$, $f_p(x) = p(1-p)^x$, $x = 0, 1, 2, \ldots$

$$\mathbb{E}_p(X) = p \sum_{x=0}^{\infty} x(1-p)^x = \frac{1-p}{p}$$

So, MoM estimator is found by solving $\frac{1-p}{p} = \bar{X}$

$$\Rightarrow \frac{1}{p} = \bar{X} + 1 \Rightarrow p = \frac{1}{\bar{X}+1}$$

$$\therefore \hat{p}_{MoM} = \frac{1}{\bar{X}+1}.$$

$$L(p) = p^n (1-p)^{\sum_{i=1}^{n} x_i}, \quad \ell(p) = n \log p + \sum_{i=1}^{n} x_i \log(1-p)$$

$$\ell'(p) = \frac{n}{p} - \frac{\sum_{i=1}^{n} x_i}{1-p}$$

$$\ell''(p) = -\frac{n}{p^2} - \frac{\sum_{i=1}^{n} x_i}{(1-p)^2} < 0$$

$$\ell'(p) = 0 \Rightarrow \frac{n}{p} = \frac{\sum_{i=1}^{n} x_i}{1-p} \Rightarrow \frac{1-p}{p} = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{x}$$

$$\Rightarrow \frac{1}{p} = \bar{x} + 1 \Rightarrow p = \frac{1}{\bar{x}+1}$$

$$\therefore \hat{p}_{MLE} = \frac{1}{\bar{X}+1} = \hat{p}_{MoM}$$

Both are functions of the minimal sufficient statistic $\sum_{i=1}^{n} X_i$

(d) For $X \sim \text{Uniform}(\theta, 1)$,

$$\mathbb{E}_\theta(X) = \int_\theta^1 \frac{x}{1-\theta}\,dx = \frac{1}{1-\theta}\left[\frac{x^2}{2}\right]_\theta^1 = \frac{1}{2(1-\theta)}(1-\theta^2)$$
$$= \frac{1+\theta}{2}$$

So, MoM estimator is obtained by solving

$$\frac{1+\theta}{2} = \bar{X} \Rightarrow 1+\theta = 2\bar{X} \Rightarrow \theta = 2\bar{X}-1.$$
$$\therefore \hat{\theta}_{MoM} = 2\bar{X}-1.$$

$$L(\theta) = \underbrace{\frac{1}{(1-\theta)^n}}_{\text{increasing in } \theta} \mathbb{1}\{\theta \le x_{(1)}\}$$

$\therefore L(\theta)$ is max. at the max. value of θ, which is $x_{(1)}$.

So, $\hat{\theta}_{MLE} = X_{(1)}$.

Here, $\hat{\theta}_{MoM}$ and $\hat{\theta}_{MLE}$ are different.

While $\hat{\theta}_{MLE}$ is a function of a minimal sufficient statistic, $\hat{\theta}_{MoM}$ is not. On the other hand, $\hat{\theta}_{MoM}$ is unbiased for θ but $\hat{\theta}_{MLE}$ is not. $\hat{\theta}_{MLE}$ is consistent. In fact, the MSE of $\hat{\theta}_{MLE}$ goes to 0 much faster compared to the variance of $\hat{\theta}_{MoM}$. So, it is better to use the MLE in this case.

(e) For $X \sim \text{Uniform}(\theta, \theta+1)$, $\mathbb{E}_\theta(X) = \theta + \frac{1}{2}$.

$\therefore \hat{\theta}_{MoM}$ satisfies $\hat{\theta}_{MoM} + \frac{1}{2} = \bar{X} \Rightarrow \hat{\theta}_{MoM} = \bar{X} - \frac{1}{2}$.

Now, $$L(\theta) = \begin{cases} 1 & \text{if } \theta < x_{(1)} < x_{(n)} < \theta+1 \\ 0 & \text{o.w.} \end{cases}$$

$$\Leftrightarrow L(\theta) = \begin{cases} 1 & \text{if } x_{(n)}-1 < \theta < x_{(1)} \\ 0 & \text{o.w.} \end{cases}$$

So, $L(\theta)$ is max. for any value of $\theta \in (x_{(n)}-1, x_{(1)})$.

So, any estimator taking values in $(X_{(n)}-1, X_{(1)})$ is an MLE.

Check whether $\hat{\theta}_{MoM} = \bar{X} - \frac{1}{2} \in (X_{(n)}-1, X_{(1)})$.

This would mean that $\hat{\theta}_{MoM}$ is an MLE.

Keep in mind: $\theta \le X_{(1)} \le X_i \le X_{(n)} \le \theta+1$.

$\therefore X_{(n)}-1 \le X_{(1)} \le X_i$ &
$X_i - 1 \le X_{(n)}-1 \le X_{(1)}$ || $\bar{X} - \frac{1}{2} = \frac{1}{2}(\bar{X}+\bar{X}-1)$

(f) $X \sim \text{Normal}(0,\sigma^2)$ $\mathbb{E}_\sigma(X)=0$, $\mathbb{E}_\sigma(X^2)=\sigma^2$.

Although we have one parameter, the first moment is free of the parameter. So, we need to look at the next moment. In general, we keep searching till we have enough equations to get a solution.

The MoM estimator is obtained by solving $\sigma^2 = \frac{1}{n}\sum_{i=1}^n x_i^2$

$\therefore \hat{\sigma}^2_{MoM} = \frac{1}{n}\sum_{i=1}^n x_i^2$.

You may be tempted to use $var(X)=\sigma^2$ and hence $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (x_i-\bar{x})^2$. But then, we would miss the fact that $\mathbb{E}_\sigma(X)=0$, so, $\bar{X}$ should not be considered while calculating the sample variance.

$$L(\sigma) = \frac{1}{(2\pi)^{n/2}\sigma^n} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n x_i^2}$$

$$\therefore \ell(\sigma) = -\frac{n}{2}\log(2\pi) - n\log\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^n x_i^2$$

$$\ell'(\sigma) = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^n x_i^2$$

$$\ell''(\sigma) = \frac{n}{\sigma^2} - \frac{3}{\sigma^4}\sum_{i=1}^n x_i^2$$

$$l'(\sigma) = 0 \Rightarrow \frac{n}{\sigma} = \frac{1}{\sigma^3}\sum_{i=1}^{n} x_i^2 \Rightarrow \sigma^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 = \hat{\sigma}^2 \text{ (say)}$$

$$l''(\hat{\sigma}^2) = \frac{n}{\hat{\sigma}^2} - \frac{3}{\hat{\sigma}^4}\sum_{i=1}^{n} x_i^2 \qquad \frac{\sum x_i^2}{\hat{\sigma}^2} = n$$

$$= \frac{n}{\hat{\sigma}^2} - \frac{3n}{\hat{\sigma}^2} = -\frac{2n}{\hat{\sigma}^2} < 0$$

$$\therefore \hat{\sigma}^2_{MLE} = \frac{1}{n}\sum_{i=1}^{n} X_i^2 .$$

The MoM estimator and the MLE are the same. They are functions of the minimal sufficient statistic $\sum_{i=1}^{n} X_i^2$.

(9) For $X \sim N(\mu, \sigma^2)$, $\mathbb{E}_{\mu,\sigma^2}(X) = \mu$, $\mathbb{E}_{\mu,\sigma^2}(X^2) = \mu^2 + \sigma^2$

$$\therefore \hat{\mu}_{MoM} = \bar{X}, \quad \hat{\mu}^2_{MoM} + \hat{\sigma}^2_{MoM} = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

$$\Rightarrow \hat{\sigma}^2_{MoM} = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

$$L(\mu,\sigma) = \frac{1}{(2\pi)^{n/2}\sigma^n} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

$$l(\mu,\sigma) = -\frac{n}{2}\log(2\pi) - n\log\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

$$\frac{\partial l(\mu,\sigma)}{\partial\mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu),$$

$$\frac{\partial l(\mu,\sigma)}{\partial\sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i-\mu)^2$$

Equating to 0 we get

$$\sum_{i=1}^{n} x_i - n\mu = 0 \Rightarrow \mu = \frac{\sum x_i}{n}$$

$$\frac{n}{\sigma} = \frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i-\mu)^2 \Rightarrow \sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

Calculate $\frac{\partial^2}{\partial\mu^2} l(\mu,\sigma)$, $\frac{\partial^2}{\partial\sigma^2} l(\mu,\sigma)$, $\frac{\partial^2}{\partial\mu\partial\sigma} l(\mu,\sigma)$ and verify that the matrix

$$\begin{pmatrix} \frac{\partial^2}{\partial\mu^2} l(\mu,\sigma) & \frac{\partial^2}{\partial\mu\partial\sigma} l(\mu,\sigma) \\ \frac{\partial^2}{\partial\mu\partial\sigma} l(\mu,\sigma) & \frac{\partial^2}{\partial\sigma^2} l(\mu,\sigma) \end{pmatrix}$$ is negative definite.

So, $\hat{\mu}_{MLE} = \bar{X}$, $\hat{\sigma}^2_{MLE} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$.

The MOM estimator and the MLE are the same and they are functions of the minimal suff. statistic $\left(\sum_{i=1}^{n} X_i, \sum_{i=1}^{n} X_i^2\right)$.

(h) For $X \sim \text{Laplace}(\mu, 1)$, $\mathbb{E}_\mu(X) = \mu$

$\therefore \hat{\mu}_{MOM} = \bar{X}$

$$L(\mu) = \frac{1}{2^n} e^{-\sum_{i=1}^{n}|x_i - \mu|}$$

$$l(\mu) = -n\log 2 - \sum_{i=1}^{n}|x_i - \mu|$$

Maximizing $l(\mu)$ w.r.t. μ is equivalent to minimizing $\sum_{i=1}^{n}|x_i - \mu|$ w.r.t. μ, which happens if μ is the median of $x_1, \ldots x_n$.

So, $\hat{\theta}_{MLE} = \tilde{X}_{me} = \text{Median}(X_1, \ldots, X_n)$

Both the estimators are functions of the minimal suff. statistic $(X_{(1)}, \ldots, X_{(n)})$.

(i) $X \sim \text{Laplace}(0, \sigma)$

$$f_\sigma(x) = \frac{1}{2\sigma} e^{-\frac{|x|}{\sigma}}, \quad x \in \mathbb{R}$$

$$\mathbb{E}_\sigma(X) = 0, \quad \mathbb{E}_\sigma(X^2) = \frac{1}{2\sigma} \int_{-\infty}^{\infty} x^2 e^{-\frac{|x|}{\sigma}} dx$$

$$= \frac{1}{\cancel{2}\sigma} \cancel{2} \int_0^\infty x^2 e^{-\frac{x}{\sigma}} dx$$

$$= \sigma^2 \int_0^\infty z^2 e^{-z} dz \qquad \Big| \ \frac{x}{\sigma} = z$$

$$= \sigma^2 \Gamma 3 = 2\sigma^2$$

MoM estimator is obtained by solving

$$2\sigma^2 = \frac{1}{n}\sum_{i=1}^n X_i^2 \Rightarrow \hat{\sigma}_{MOM} = \sqrt{\frac{1}{2n}\sum_{i=1}^n X_i^2}$$

$$L(\sigma) = \frac{1}{(2\sigma)^n} e^{-\frac{1}{\sigma}\sum_{i=1}^n |x_i|}$$

$$\ell(\sigma) = -n\log 2 - n\log\sigma - \frac{1}{\sigma}\sum_{i=1}^n |x_i|$$

$$\ell'(\sigma) = -\frac{n}{\sigma} + \frac{1}{\sigma^2}\sum_{i=1}^n |x_i|$$

$$\ell''(\sigma) = \frac{n}{\sigma^2} - \frac{2}{\sigma^3}\sum_{i=1}^n |x_i|$$

$$\ell'(\sigma) = 0 \Rightarrow \sigma = \frac{1}{n}\sum_{i=1}^n |x_i| =: \hat{\sigma}, \text{ say}$$

$$\ell''(\hat{\sigma}) = -\frac{n}{\hat{\sigma}^2}\sum_{i=1}^n |x_i| < 0 \qquad \therefore \hat{\sigma}_{MLE} = \frac{1}{n}\sum_{i=1}^n |X_i|$$

MLE is a function of a minimal suff. stat. but MoM is not

(j) $X \sim \text{Laplace}(\mu, \sigma)$. $\mathbb{E}(X) = \mu$, $Var(X) = 2\sigma^2$

$\therefore$ The MoM estimators are obtained by solving

$$\left.\begin{aligned} \mu &= \bar{X} \\ 2\sigma^2 &= \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 \end{aligned}\right\} \Rightarrow \begin{aligned} \hat{\mu}_{MoM} &= \bar{X} \\ \hat{\sigma}_{MoM} &= \sqrt{\frac{1}{2n}\sum_{i=1}^{n}(X_i - \bar{X})^2} \end{aligned}$$

$$L(\mu, \sigma) = \frac{1}{(2\sigma)^n} e^{-\frac{1}{\sigma}\sum_{i=1}^{n}|x_i - \mu|}$$

$$\ell(\mu, \sigma) = -n\log 2 - n\log\sigma - \frac{1}{\sigma}\sum_{i=1}^{n}|x_i - \mu|$$

For any fixed σ, $\ell(\mu, \sigma)$ is minimized at $\mu = \tilde{X}_{Me}$

For a fixed μ, $\ell(\mu, \sigma)$ is minimized at $\hat{\sigma} = \frac{1}{n}\sum_{i=1}^{n}|x_i - \mu|$

$\therefore$ MLE of (μ, σ) are

$$\hat{\mu}_{MLE} = \tilde{X}_{Me} = \text{Median}(X_1 \dots X_n)$$

$$\hat{\sigma}_{MLE} = \frac{1}{n}\sum_{i=1}^{n}|X_i - \tilde{X}_{Me}|$$

(k) For $X \sim \text{Normal}(\theta, \theta^2)$, $\mathbb{E}_\theta(X) = \theta$ & $Var_\theta(X) = \theta^2$

Since only one parameter is missing, MoM estimator is obtained as

$\hat{\theta}_{MoM} = \bar{X}$. [Also note here that $\theta \in \mathbb{R}$]

$$L(\theta) = \left(\frac{1}{\sqrt{2\pi}\,\theta}\right)^n e^{-\frac{1}{2\theta^2}\sum_{i=1}^{n}(x_i - \theta)^2}$$

$$\therefore\ \ell(\theta) = -\frac{n}{2}\log(2\pi) - n\log\theta - \frac{1}{2\theta^2}\sum_{i=1}^{n}(x_i-\theta)^2$$

$$= -\frac{n}{2}\log(2\pi) - n\log\theta - \frac{1}{2\theta^2}\sum_{i=1}^{n}x_i^2 + \frac{1}{\theta}\sum_{i=1}^{n}x_i - \frac{n}{2}$$

$$\ell'(\theta) = -\frac{n}{\theta} + \frac{1}{\theta^3}\sum_{i=1}^{n}x_i^2 - \frac{1}{\theta^2}\sum_{i=1}^{n}x_i$$

$$\ell''(\theta) = \frac{n}{\theta^2} - \frac{3}{\theta^4}\sum_{i=1}^{n}x_i^2 + \frac{2}{\theta^3}\sum_{i=1}^{n}x_i$$

$$\ell'(\theta) = 0 \implies n\theta^2 + \theta\sum_{i=1}^{n}x_i - \sum_{i=1}^{n}x_i^2 = 0$$

$$\iff \theta^2 + \theta\bar{x} - S_{xx} = 0 \qquad \left[S_{xx} = \frac{1}{n}\sum_{i=1}^{n}x_i^2\right]$$

$$\therefore\ \theta = \frac{-\bar{x} \pm \sqrt{\bar{x}^2 + 4S_{xx}}}{2} \longleftarrow \text{two roots}$$

Notice that $\dfrac{\theta^4\ell''(\theta)}{n} = \theta^2 - 3S_{xx} + 2\theta\bar{x}$

$$= \underbrace{\theta^2 + \theta\bar{x} - S_{xx}}_{0} + \underbrace{\theta\bar{x} - S_{xx}}_{-\theta^2} - S_{xx}$$

at the solution to $\ell'(\theta) = 0$

$$= -\theta^2 - S_{xx} < 0$$

$$\therefore\ \ell''(\hat{\theta}) < 0.$$

So, there are two MLE^s given by

$$\frac{-\bar{X} + \sqrt{\bar{X}^2 + 4S_{xx}}}{2} \quad \& \quad \frac{-\bar{X} - \sqrt{\bar{X}^2 + 4S_{xx}}}{2}$$

(c) $X \sim \text{Normal}(\theta, \theta)$, $\theta > 0$

Since $\mathbb{E}_\theta(X) = \theta$ & only one parameter is unknown, $\hat{\theta}_{MoM} = \bar{X}$. But, $\bar{X}$ may be negative.

Therefore, we can also use $\text{Var}_\theta(X) = \theta$ to set

$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$, which is always non-negative

Strictly speaking, this is not the MoM estimator. But, it is more sensible to use in this case.

$$L(\theta) = \frac{1}{(2\pi\theta)^{n/2}} e^{-\frac{1}{2\theta}\sum_{i=1}^{n}(x_i-\theta)^2}$$

$$\ell(\theta) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\theta - \frac{1}{2\theta}\sum_{i=1}^{n}(x_i-\theta)^2$$

$$= -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\theta - \frac{1}{2\theta}\sum_{i=1}^{n}x_i^2 + \sum_{i=1}^{n}x_i - \frac{n\theta}{2}$$

$$\ell'(\theta) = -\frac{n}{2\theta} + \frac{1}{2\theta^2}\sum_{i=1}^{n}x_i^2 - \frac{n}{2}$$

$$\ell''(\theta) = \frac{n}{2\theta^2} - \frac{1}{\theta^3}\sum_{i=1}^{n}x_i^2$$

$$\ell'(\theta) = 0 \Rightarrow \frac{1}{\theta} - \frac{1}{\theta^2}\frac{1}{n}\sum_{i=1}^{n}x_i^2 + 1 = 0$$

$$\Rightarrow \theta^2 + \theta - S_{xx} = 0$$

$$\Rightarrow \theta = \frac{-1 \pm \sqrt{4S_{xx}+1}}{2}$$

Since $\theta > 0$, we take $\hat{\theta} = \frac{\sqrt{4S_{xx}+1} - 1}{2} > 0$

$\therefore \hat{\theta}_{MLE} = \frac{\sqrt{4S_{xx}+1}-1}{2}$

Also, $\frac{2\theta^3}{n}\ell''(\theta) = \theta - 2S_{xx} = \theta - S_{xx} - S_{xx} = -\theta^2 - S_{xx} < 0$

(m) For $X \sim \text{Exponential}(\text{scale} = \theta)$, $\mathbb{E}_\theta(X) = \frac{1}{\theta}$

$\therefore \hat{\theta}_{MOM}$ is obtained by solving $\frac{1}{\theta} = \bar{X} \Rightarrow \theta = \frac{1}{\bar{X}}$

$\therefore \hat{\theta}_{MOM} = \frac{1}{\bar{X}}$.

$$L(\theta) = \theta^n e^{-\theta \sum_{i=1}^{n} x_i}$$

$$\ell(\theta) = n \log \theta - \theta \sum_{i=1}^{n} x_i$$

$$\ell'(\theta) = \frac{n}{\theta} - \sum_{i=1}^{n} x_i$$

$$\ell''(\theta) = -\frac{n}{\theta^2} < 0$$

$$\ell'(\theta) = 0 \Rightarrow \theta = \frac{n}{\sum_{i=1}^{n} x_i} = \frac{1}{\bar{x}}$$

$\therefore \hat{\theta}_{MLE} = \frac{1}{\bar{X}}$.

Exercise 2 For $(X, Y) \sim BVN(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$

$$\mathbb{E}(X) = \mu_x, \quad \mathbb{E}(Y) = \mu_y, \quad Var(X) = \sigma_x^2, \quad Var(Y) = \sigma_y^2$$

$$\mathbb{E}(XY) = Cov(X, Y) + \mathbb{E}(X)\mathbb{E}(Y)$$

$$= \sigma_x \sigma_y \rho + \mu_x \mu_y$$

$\therefore$ MoM estimators are

$$\hat{\mu}_{x,MoM} = \bar{X}, \quad \hat{\mu}_{y,MoM} = \bar{Y}, \quad \hat{\sigma}^2_{x,MoM} = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2$$

$$\hat{\sigma}^2_{y,MoM} = \frac{1}{n}\sum_{i=1}^n (Y_i - \bar{Y})^2$$

$$\rho \hat{\sigma}_x \hat{\sigma}_y + \hat{\mu}_x \hat{\mu}_y = \frac{1}{n}\sum_{i=1}^n X_i Y_i$$

$$\Rightarrow \rho \hat{\sigma}_x \hat{\sigma}_y = \frac{1}{n}\sum_{i=1}^n X_i Y_i - \bar{X}\bar{Y} = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})$$

$$\therefore \hat{\rho}_{MoM} = \frac{\frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2 \frac{1}{n}\sum_{i=1}^n (Y_i - \bar{Y})^2}}.$$

Now, $L(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho) = \left(\frac{1}{2\pi \sigma_x \sigma_y \sqrt{1-\rho^2}}\right)^n$

$$e^{-\frac{1}{2(1-\rho^2)}\sum_{i=1}^n \left\{ \left(\frac{x_i - \mu_x}{\sigma_x}\right)^2 + \left(\frac{y_i - \mu_y}{\sigma_y}\right)^2 - 2\rho\left(\frac{x_i - \mu_x}{\sigma_x}\right)\left(\frac{y_i - \mu_y}{\sigma_y}\right)\right\}}$$

$$\therefore \ell(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho) = -n\log(2\pi) - n\log\sigma_x - n\log\sigma_y - \frac{n}{2}\log(1-\rho^2)$$

$$-\frac{1}{2(1-\rho^2)}\left\{\sum_{i=1}^n \left(\frac{x_i - \mu_x}{\sigma_x}\right)^2 + \sum_{i=1}^n \left(\frac{y_i - \mu_y}{\sigma_y}\right)^2 - 2\rho\sum_{i=1}^n \left(\frac{x_i - \mu_x}{\sigma_x}\right)\left(\frac{y_i - \mu_y}{\sigma_y}\right)\right\}$$

Take derivative, equate to zero to show that the MLE is same as the MoM estimator.

Parts (b) & (c) are similar. Formulate the likelihood equations and find the solutions.

Exercise 3 (a) $L(\theta) = \begin{cases} \{f_0(x)\}^n & , \text{ if } \theta = 0 \\ \{f_1(x)\}^n & , \text{ if } \theta = 1 \end{cases}$

$$\therefore \quad \left.\begin{aligned} L(0) &= L(0|\underline{x}) = 1 \quad , \quad 0 < x_i < 1 \\ L(1) &= L(1|\underline{x}) = \frac{1}{2^n} \prod_{i=1}^{n} \frac{1}{\sqrt{x_i}} \quad , \quad 0 < x_i < 1 \end{aligned}\right] \text{likelihood function}$$

(b) Maximum likelihood estimator

$$\hat{\theta}(\underline{x}) = \underset{\theta \in \{0,1\}}{\arg\max}\ L(\theta|\underline{x}) = \begin{cases} 0 & \text{if } L(0|\underline{x}) > L(1|\underline{x}) \\ 1 & \text{if } L(1|\underline{x}) > L(0|\underline{x}) \end{cases}$$

$$L(0|\underline{x}) > L(1|\underline{x}) \iff 1 > \frac{1}{2^n} \prod_{i=1}^{n} \frac{1}{\sqrt{x_i}}$$

$$\iff \prod_{i=1}^{n} \sqrt{x_i} > \frac{1}{2^n}$$

$$\iff \prod_{i=1}^{n} x_i > \frac{1}{4^n}$$

$$\iff \underbrace{\left(\prod_{i=1}^{n} x_i \right)^{1/n}}_{\text{geometric mean (GM) of } x_1 \dots x_n} > \frac{1}{4}$$

So, the MLE is

$$\hat{\theta}_{MLE} = \begin{cases} 0 & \text{if } GM(X_1 \dots X_n) > \frac{1}{4} \\ 1 & \text{if } GM(X_1 \dots X_n) < \frac{1}{4} \end{cases}$$

Notice that we have not specified what happens if $GM(X_1 \dots X_n) = \frac{1}{4}$. This is referred as the "boundary case", which is nondiscriminating. In this case, the boundary occurs with prob 0, so we don't care.

Exercise 4 (a) $L(\theta) = L(\theta|x) = \mathbb{P}_\theta(X=x)$

$$\therefore L(\theta|x) = \begin{cases} 6\theta^2 - 4\theta + 1 & \text{if } x = 0 \\ \theta - 2\theta^2 & \text{if } x = 1 \\ 3\theta - 4\theta^2 & \text{if } x = 2 \end{cases}$$

(b) MLE is obtained by $\hat{\theta}(x) = \underset{\theta \in \Theta}{\operatorname{argmax}}\ L(\theta|x)$

For $x=0$, $L(\theta) = 6\theta^2 - 4\theta + 1$

$$\left.\begin{aligned} L'(\theta) &= 12\theta - 4 \\ L''(\theta) &= 12 > 0 \end{aligned}\right\} \therefore \underline{L \text{ is minimized at } \theta = \tfrac{1}{3}}$$

does not give us the MLE.

Observe that $L'(\theta) \gtrless 0 \iff \theta \gtrless \frac{1}{3}$

So, L is $\downarrow$ in $[0, \frac{1}{3})$

and L is $\uparrow$ in $(\frac{1}{3}, \frac{1}{2}]$

$L(0) = 1$, $L(\frac{1}{2}) = \frac{1}{2}$, $L(\frac{1}{3}) = \frac{1}{3}$

$\therefore L$ is max. at 0. $\therefore \hat{\theta}(0) = 0$.

This can be also seen by plotting $L(\theta|0)$, which is a quadratic function.

(*) This shows that solving the likelihood equations may not always be enough to find the MLE.

For $x=1$, $L(\theta) = \theta - 2\theta^2$

$$\left.\begin{aligned} L'(\theta) &= 1 - 4\theta \\ L''(\theta) &= -4 < 0 \end{aligned}\right\} L \text{ is max. at } \theta = \tfrac{1}{4}$$

$\therefore \hat{\theta}(1) = \frac{1}{4}$.

For $x=2$, $L(\theta) = 3\theta - 4\theta^2$

$$\left.\begin{aligned} L'(\theta) &= 3 - 8\theta \\ L''(\theta) &= -8 < 0 \end{aligned}\right\} L \text{ is max. at } \tfrac{3}{8}$$

$\therefore \hat{\theta}(2) = \frac{1}{4}$.

So, $\hat{\theta}_{MLE} = \begin{cases} 0 & \text{if } X=0 \\ \frac{1}{4} & \text{if } X=1 \\ \frac{3}{8} & \text{if } X=2 \end{cases}$

(c) For MoM estimator, we find

$$\mathbb{E}_\theta(X) = 0 \cdot \mathbb{P}_\theta(X=0) + 1 \cdot \mathbb{P}_\theta(X=1) + 2 \cdot \mathbb{P}_\theta(X=2)$$
$$= 1 \cdot (\theta - 2\theta^2) + 2(3\theta - 4\theta^2)$$
$$= \theta - 2\theta^2 + 6\theta - 8\theta^2$$
$$= 7\theta - 10\theta^2$$

With a single observation, $\bar{X} = X$

So, we equate

$$7\theta - 10\theta^2 = X$$
$$\Rightarrow 10\theta^2 - 7\theta + X = 0$$
$$\Rightarrow \theta = \frac{7 \pm \sqrt{49 - 40X}}{20}$$

For $X=0$, $\theta = \frac{7 \pm 7}{20} = 0$ or $\frac{7}{10}$. Since $\Theta = [0, \frac{1}{2}]$ we take $\hat{\theta}_{MoM} = 0$.

For $X=1$, $\theta = \frac{7 \pm 3}{20} = \frac{1}{2}$ or $\frac{1}{5}$. Both are valid estimates.

For $X=2$, $\theta = \frac{7 \pm \sqrt{-31}}{20}$, which is not a real number.

So, $\hat{\theta}_{MoM} = \begin{cases} 0 & \text{if } X=0 \\ \frac{1}{2} \text{ or } \frac{1}{5} & \text{if } X=1 \\ \text{undefined} & \text{if } X=2 \end{cases}$

Exercise 5 (a) $x_1, \dots, x_n$ are fixed constants

$$Y_i = \beta x_i + \varepsilon_i, \quad i = 1, \dots, n, \quad \varepsilon_1, \dots, \varepsilon_n \overset{iid}{\sim} N(0, \sigma^2).$$

Since ε_i's are independent, therefore Y_i's are indep.

Also, $Z \sim N(\mu, \sigma^2) \Rightarrow a + bZ \sim N(a + b\mu, b^2\sigma^2)$

This implies $Y_i \sim N(\beta x_i, \sigma^2)$, $i = 1, \dots, n$

Putting together, we get $Y_i \overset{indep}{\sim} N(\beta x_i, \sigma^2)$, $i = 1, \dots, n$

$$\therefore L(\beta, \sigma^2) = L(\beta, \sigma^2 \mid \underset{\sim}{y}, \underset{\sim}{x}) = \prod_{i=1}^{n} f_{Y_i}(y_i)$$

$$= \prod_{i=1}^{n} \left\{ \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2\sigma^2}(y_i - \beta x_i)^2} \right\}$$

$$= \frac{1}{(2\pi)^{n/2} \sigma^n} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \beta x_i)^2}$$

$$\ell(\beta, \sigma^2) = \log L(\beta, \sigma^2) = -\frac{n}{2} \log(2\pi) - n \log \sigma - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \beta x_i)^2$$

(b) The unknown parameters are β & σ.

MLEs of these parameters are obtained as

$$(\hat{\beta}, \hat{\sigma}) = \underset{\beta, \sigma}{\operatorname{argmax}}\ \ell(\beta, \sigma^2).$$

These can be obtained directly by solving the likelihood equation:

$$\frac{\partial \ell(\beta, \sigma^2)}{\partial \beta} = 0 \iff \sum_{i=1}^{n} (y_i - \beta x_i) x_i = 0 \Rightarrow \beta = \frac{\sum_{i=1}^{n} y_i x_i}{\sum_{i=1}^{n} x_i^2}.$$

$$\frac{\partial \ell(\beta, \sigma^2)}{\partial \sigma} = 0 \iff \frac{n}{\sigma} = \frac{1}{\sigma^3} \sum_{i=1}^{n} (y_i - \beta x_i)^2 \Rightarrow \sigma^2 = \frac{1}{n} \sum_{i=1}^{n} (y_i - \beta x_i)^2$$

Therefore, the MLEs are

$$\hat{\beta} = \frac{\sum_{i=1}^{n} y_i x_i}{\sum_{i=1}^{n} x_i^2}, \quad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{\beta} x_i)^2.$$

It can be verified that these indeed maximize $\ell(\beta, \sigma^2)$.

Alternatively, note that for any fixed σ, maximizing $\ell(\beta,\sigma^2)$ w.r.t. β is equivalent to minimizing $\sum_{i=1}^{n}(y_i-\beta x_i)^2$ w.r.t. β.

Now, $\sum_{i=1}^{n}(y_i-\beta x_i)^2 = \sum_{i=1}^{n} x_i^2\left(\frac{y_i}{x_i}-\beta\right)^2$

Recall that $\underbrace{\sum_{i=1}^{n}(z_i-c)^2}_{\text{sum of squares}}$ is min. w.r.t. c if $c = \underbrace{\frac{1}{n}\sum_{i=1}^{n} z_i}_{\text{average}}$.

Similarly, $\underbrace{\sum_{i=1}^{n} \underset{\text{weights}}{w_i}(z_i-c)^2}_{\text{weighted sum of squares}}$ is min. w.r.t. c if $c = \underbrace{\frac{\sum_{i=1}^{n} w_i z_i}{\sum_{i=1}^{n} w_i}}_{\text{weighted average}}$ ← weights

Therefore $\sum_{i=1}^{n} x_i^2\left(\frac{y_i}{x_i}-\beta\right)^2$ is min. if $\beta = \dfrac{\sum_{i=1}^{n} x_i^2 \frac{y_i}{x_i}}{\sum_{i=1}^{n} x_i^2}$

$$= \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}.$$

Again for a fixed β, $\ell(\beta,\sigma^2)$ is min. w.r.t σ if $\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i-\beta x_i)^2$.

So, the MLEs are $\hat\beta = \dfrac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}$, $\hat\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i-\hat\beta x_i)^2$.

(c) If $\epsilon_1,\ldots,\epsilon_n \overset{iid}{\sim} \text{Laplace}(0,\sigma)$, then $Y_i \overset{indep}{\sim} \text{Laplace}(\beta x_i,\sigma)$. So, the likelihood becomes

$$L(\beta,\sigma) = \frac{1}{2^n\sigma^n} e^{-\frac{1}{\sigma}\sum_{i=1}^{n}|y_i-\beta x_i|}.$$

$$\ell(\beta,\sigma) = -n\log 2 - n\log\sigma - \frac{1}{\sigma}\sum_{i=1}^{n}|y_i-\beta x_i|.$$

The usual way of setting the derivatives to 0 does not work for β. But, it is still true that for a fixed σ, $l(\beta,\sigma)$ is max. if $\sum_{i=1}^{n}|y_i - \beta x_i|$ is min.

Now, $\sum_{i=1}^{n}|y_i - \beta x_i| = \underbrace{\sum_{i=1}^{n} \underset{\uparrow \text{weight}}{|x_i|} \left|\frac{y_i}{x_i} - \beta\right|}_{\text{weighted sum of absolute deviations}}$.

This is min. if β is the <u>weighted median</u> of $\frac{y_1}{x_1}, \dots, \frac{y_n}{x_n}$ with weights $|x_1|, \dots, |x_n|$, resp.

More precisely, we calculate the median from the following freq. table.

Value	Freq. (rel.
y_1/x_1	$\lvert x_1 \rvert$
y_2/x_2	$\lvert x_2 \rvert$
$\vdots$	
y_n/x_n	$\lvert x_n \rvert$

Now, for fixed β, $l(\beta,\sigma)$ is max. w.r.t. σ if

$$\frac{n}{\sigma} = \frac{1}{\sigma^2}\sum_{i=1}^{n}|y_i - \beta x_i| \Rightarrow \sigma = \frac{1}{n}\sum_{i=1}^{n}|y_i - \beta x_i|.$$

So, the MLE's are obtained as

$\hat{\beta}$ = wt. median of $\left\{\frac{y_1}{x_1}, \dots, \frac{y_n}{x_n}\right\}$ with wts. $\{|x_1|, \dots, |x_n|\}$

$$\hat{\sigma} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{\beta} x_i|.$$

(d) When $Y_i = \alpha + \beta x_i + \varepsilon_i$, $i = 1, \ldots, n$, $\varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$,

then $Y_i \overset{indep.}{\sim} N(\alpha + \beta x_i, \sigma^2)$, $i = 1, \ldots, n$. So,

$$L(\alpha, \beta, \sigma^2) = \frac{1}{(2\pi)^{n/2} \sigma^n} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \alpha - \beta x_i)^2}$$

$$\ell(\alpha, \beta, \sigma^2) = -\frac{n}{2} \log(2\pi) - n \log \sigma - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \alpha - \beta x_i)^2$$

For a fixed σ, this is max. w.r.t. (α, β) if

$\sum_{i=1}^{n} (y_i - \alpha - \beta x_i)^2$ is min.

Check that this happens if $\alpha = \frac{1}{n} \sum_{i=1}^{n} y_i - \beta \frac{1}{n} \sum_{i=1}^{n} x_i$

$= \bar{y} - \beta \bar{x}$

and $\beta = \dfrac{\sum_{i=1}^{n} (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$

Finally, max. w.r.t. σ happens if $\sigma^2 = \frac{1}{n} \sum_{i=1}^{n} (y_i - \alpha - \beta x_i)^2$

So, the MLEs are

$$\hat{\alpha} = \bar{y} - \hat{\beta} \bar{x}, \quad \hat{\beta} = \frac{\sum_{i=1}^{n} (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{\alpha} - \hat{\beta} x_i)^2$$

Note: These models are known as linear models. You will learn more about these models in a separate course.

Exercise 6 (a) The non-full-rank likelihood is

$$L(p_A, p_B, p_{AB}, p_0) = p_A^{N_A}\, p_B^{N_B}\, p_{AB}^{N_{AB}}\, p_0^{N_0}, \quad \left(\begin{array}{l}\text{See the sol. to Ex 5}\\ \text{in Series 2}\end{array}\right)$$

$$0 \le p_A, p_B, p_{AB}, p_0 \le 1, \quad p_A + p_B + p_{AB} + p_0 = 1.$$

Also, $N_A + N_B + N_{AB} + N_0 = 500$

One full-rank likelihood is obtained by writing $p_0 = 1 - p_A - p_B - p_{AB}$ and $N_0 = 500 - N_A - N_B - N_{AB}$.

$$L(p_A, p_B, p_{AB}) = p_A^{N_A}\, p_B^{N_B}\, p_{AB}^{N_{AB}} (1 - p_A - p_B - p_{AB})^{500 - N_A - N_B - N_{AB}},$$

$$0 \le p_A, p_B, p_{AB} \le 1,$$
$$p_A + p_B + p_{AB} \le 1$$

$$\ell(p_A, p_B, p_{AB}) = N_A \log p_A + N_B \log p_B + N_{AB} \log p_{AB} + (500 - N_A - N_B - N_{AB}) \log(1 - p_A - p_B - p_{AB})$$

The MLE is obtained by solving the likelihood equations:

$$\frac{\partial \ell}{\partial p_A} = 0 \iff \frac{N_A}{p_A} = \frac{500 - N_A - N_B - N_{AB}}{1 - p_A - p_B - p_{AB}}$$

$$\frac{\partial \ell}{\partial p_B} = 0 \iff \frac{N_B}{p_B} = \frac{500 - N_A - N_B - N_{AB}}{1 - p_A - p_B - p_{AB}}$$

$$\frac{\partial \ell}{\partial p_{AB}} = 0 \iff \frac{N_{AB}}{p_{AB}} = \frac{500 - N_A - N_B - N_{AB}}{1 - p_A - p_B - p_{AB}}$$

$$\therefore \frac{N_A}{p_A} = \frac{N_B}{p_B} = \frac{N_{AB}}{p_{AB}} = \frac{500 - N_A - N_B - N_{AB}}{1 - p_A - p_B - p_{AB}} = c, \text{ say.}$$

$$\therefore p_A = \frac{N_A}{c}, \quad p_B = \frac{N_B}{c}, \quad p_{AB} = \frac{N_{AB}}{c}$$

$$p_A + p_B + p_{AB} = \frac{N_A + N_B + N_{AB}}{c}$$

$$c = \frac{500 - N_A - N_B - N_{AB}}{1 - p_A - p_B - p_{AB}} = \frac{500 - N_A - N_B - N_{AB}}{c - N_A - N_B - N_{AB}} \cdot c \implies c = 500$$

$$\therefore \hat{p}_A = \frac{N_A}{500}, \quad \hat{p}_B = \frac{N_B}{500}, \quad \hat{p}_{AB} = \frac{N_{AB}}{500}. \quad \hat{p}_0 = 1 - \hat{p}_A - \hat{p}_B - \hat{p}_{AB} = \frac{N_0}{500}.$$

(b) With the additional information,

$$p_A = q_A(1-q_B), \quad p_B = (1-q_A)q_B, \quad p_{AB} = q_A q_B,$$

$$1-p_A-p_B-p_{AB} = p_O = (1-q_A)(1-q_B).$$

The updated likelihood becomes

$$\tilde{L}(q_A, q_B) = q_A^{N_A+N_{AB}}(1-q_A)^{500-N_A-N_{AB}}\, q_B^{N_B+N_{AB}}(1-q_A)^{500-N_B-N_{AB}}$$

$$\tilde{\ell}(q_A, q_B) = (N_A+N_{AB})\log q_A + (500-N_A-N_{AB})\log(1-q_A)$$
$$+ (N_B+N_{AB})\log q_B + (500-N_B-N_{AB})\log(1-q_B)$$

The likelihood equations are

$$\frac{N_A+N_{AB}}{q_A} = \frac{500-N_A-N_{AB}}{1-q_A} \Rightarrow \frac{q_A}{1-q_A} = \frac{N_A+N_{AB}}{500-N_A-N_{AB}} \Rightarrow q_A = \frac{N_A+N_{AB}}{500}.$$

$$\frac{N_B+N_{AB}}{q_B} = \frac{500-N_B-N_{AB}}{1-q_B} \Rightarrow \frac{q_B}{1-q_B} = \frac{N_B+N_{AB}}{500-N_B-N_{AB}} \Rightarrow q_B = \frac{N_B+N_{AB}}{500}.$$

For both the models, the solutions to the likelihood equations are indeed the maximizers. This can be shown by verifying that the hessian matrices are negative definite.

Exercise 7. As discussed in the class, technically, the MLE may fail to exist if the parameter space is restricted. If the parameter space is finite, then we simply evaluate the likelihood at the finitely many parameter values and take the one for which the likelihood is maximum. The problem arises when the parameter space is not the usual one, but still infinite (countable or uncountable). In this case, one strategy is to first find the MLE disregarding the restriction and then incorporating the restriction.

For example, in (a) we use $\hat{p} = \begin{cases} \bar{X} & \text{if } \frac{1}{4} \le \bar{X} \le \frac{3}{4} \\ \frac{1}{4} & \text{if } \bar{X} < \frac{1}{4} \\ \frac{3}{4} & \text{if } \bar{X} > \frac{3}{4} \end{cases}$

In (b), we use $\hat{\lambda} = \lceil \bar{X} \rceil$, where $\lceil t \rceil$ is the ceiling of t, which is the smallest integer greater than or equal to t.

In (c), we take the integer value that falls within $(X_{(n)}-1, X_{(1)})$, or the one which is closest to this interval.

In (d), we take $\hat{\mu} = \begin{cases} \bar{X} & \text{if } \bar{X} > 0 \\ 0 & \text{o.w.} \end{cases}$