Group 09: Luu Thu Trang (2695303) and Raminta Povilaityte (2692655)

Exercise 1

(a)

To compute Bias of $\hat{\lambda}$ and $\tilde{\lambda}$ we use the following formula:

$$Bias(\hat{\theta}) = E(\hat{\theta}) - \theta$$

To apply this formula we first find $E(\hat{\lambda})$ for $\hat{\lambda}$:

$$E(\hat{\lambda}) = E(\frac{(\overline{X})^2}{9}) = \frac{1}{9}E((\overline{X})^2) = \frac{1}{9}(V(\overline{X}) + (E(\overline{X}))^2)$$
$$= \frac{1}{9}(\frac{V(X)}{n} + (E(X))^2) = \frac{1}{9}(\frac{3\lambda}{n} + 9\lambda) = \lambda + \frac{\lambda}{3n}$$

We plug in the above expectation of $\hat{\lambda}$ in Bias formula to find :

$$Bias(\hat{\lambda}) = E(\hat{\lambda}) - \lambda = \lambda + \frac{\lambda}{3n} - \lambda = \frac{\lambda}{3n}$$

We can conclude that $\hat{\lambda}$ is an biased estimator of λ

(b)

Continue using the formula from (a), we have:

$$E(\tilde{\lambda}) = E(\frac{\sum_{i=1}^{n}(X_i)^2}{12n}) = \frac{1}{12n}E(\sum_{i=1}^{n}(X_i)^2) = \frac{1}{12n}(E(X_1)^2 + \ldots + E(X_n)^2) = \frac{1}{12n} \times 12\lambda n = \lambda$$

We plug in the above expectation of $\hat{\lambda}$ in Bias formula to find :

$$Bias(\tilde{\lambda}) = E(\tilde{\lambda}) - \lambda = \lambda - \lambda = 0$$

We can conclude that $\tilde{\lambda}$ is an unbiased estimator of λ

(c)

Suppose that:

$$V(\hat{\lambda}) = \frac{2\lambda^2}{n}$$

Then we compute the MSE (Mean Squared Error) for $\hat{\lambda}$ using the MSE formula:

$$MSE_{\hat{\lambda}}(\lambda) = Bias(\hat{\lambda})^2 + V(\hat{\lambda})$$

By using the fact that $\hat{\lambda}$ is an unbiased estimator from part a) and plugging in the given expression for $V(\hat{\lambda})$ we find $MSE_{\hat{\lambda}}(\lambda)$ to be the following :

$$MSE_{\hat{\lambda}}(\lambda) = (\frac{\lambda}{3n})^2 + \frac{2\lambda^2}{n} = \frac{\lambda^2}{9n^2} + \frac{2\lambda^2}{n} = \frac{\lambda^2(1 + 18n)}{9n^2}$$

(d)

To calculate the $MSE_{\tilde{\lambda}}(\lambda)$, we use the formula from (c).

We will first calculate $Var(\lambda) = V_\lambda(\tilde{\lambda})$.

$$V_\lambda(\tilde{\lambda}) = V(\frac{\sum_{i=1}^{n}(X_i)^2}{12n})$$
$$= (\frac{1}{12n})^2(V(X_1)^2 + ... + V(X_n)^2) = (\frac{1}{12n})^2 \times nV(X)^2$$
$$= \frac{1}{12^2 n} \times (EX^4 - (EX^2)^2) = \frac{1}{12^2 n} \times (360\lambda^2 - (12\lambda)^2)$$
$$= \frac{1}{12^2 n} \times (360\lambda^2 - 144\lambda^2) = \frac{216\lambda^2}{144n} = \frac{3\lambda^2}{2n}$$

Then,

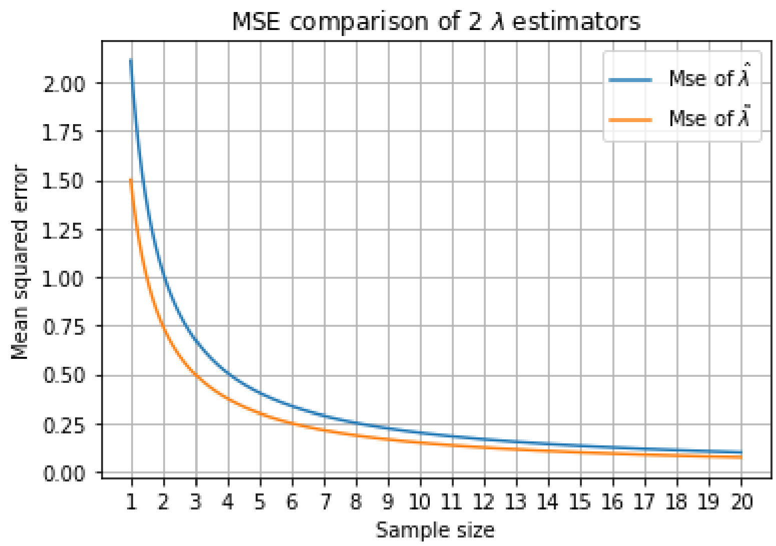
$$MSE_{\tilde{\lambda}}(\lambda) = 0 + \frac{3\lambda^2}{2n} = \frac{3\lambda^2}{2n}$$

(e)

In [6]:

```
# assume data was generated with given pdf and \lambda = 1
# then we plot MSE's for both \tilde{\lambda} and \hat{\lambda} for n from 1 to 20.

import numpy as np
import matplotlib.pyplot as plt
lamda = 1
n = np.linspace(1,20,2000)
mseHat = (lamda**2*(1+18*n))/(9*n**2)
mseTilde = (3/2*(lamda**2))/n
plt.xlabel('Sample size')
plt.xticks(np.arange(min(n), max(n)+1, 1.0))
plt.ylabel('Mean squared error')
plt.title(r'MSE comparison of 2 $\lambda$ estimators')
plt.plot(n, mseHat, label= r'Mse of $\hat{\lambda}$')
plt.plot(n, mseTilde, label= r'Mse of $\tilde{\lambda}$')
plt.legend()
plt.grid()
plt.show()
```



Based on the plot we would prefer $\tilde{\lambda}$ as it has lower MSE compared to $\hat{\lambda}$

(f)

Based on the two MSE's, we can say that $\tilde{\lambda}$ is a better estimator for λ as its MSE is smaller.

This can be deducted from MSE's equations :

It follows that $MSE_{\tilde{\lambda}} = \frac{3\lambda^2}{2n} \leq \frac{2\lambda^2}{n} \leq \frac{\lambda^2(1+18n)}{9n^2} = MSE_{\hat{\lambda}}$

Since both estimators depend on sample size, we see that they perform better as n increases. In particular both MSE's converge to 0 as n goes to ∞.

Exercise 2

(a)

To compute Maximum Likelihood Estimator (MLE) for β we first find the likelihood function (pdf of the sample) as follows:

$$L(\beta; X_1, \ldots, X_n) = \beta e^{-\beta x_1} \cdot \beta e^{-\beta x_2} \cdot \ldots \cdot \beta e^{-\beta x_n} = \beta^n e^{-\beta \sum_{i=1}^n x_i} = \beta^n e^{-\beta n \overline{X}}$$

Then we find a log-likelihood function :

$$l(\beta; X_1, \ldots, X_n) = n \log \beta - \beta n \overline{X}$$

To find an MLE(β) we take the derivative of log-likelihood in terms of β :

$$\frac{\partial l}{\partial \beta} = \frac{n}{\beta} - n\overline{X} = 0$$

Find MLE we solve the above equation for β to find:

$$\beta = \frac{n}{n\overline{X}} = \frac{1}{\overline{X}} = \hat{\beta}$$

where $\hat{\beta}$ is the MLE of β.

(b)

To calculate the Method of Moments Estimator for β based on the first moment of X. We first calculate the first moment of X. This is exactly the expectation of X.

$$EX = \frac{1}{\beta} = g_1(\beta)$$

Now we solve the equation:

$$\overline{X} = g_1(\tilde{\beta}) \leftrightarrow \overline{X} = \frac{1}{\tilde{\beta}} \leftrightarrow \tilde{\beta} = \frac{1}{\overline{X}}$$

So $\tilde{\beta} = \frac{1}{X}$ is a MM estimator for β based on the first moment of X

(c)

Second moment of X is found as follows:

$$EX^2 = V(X) + (E(X))^2 = \frac{1}{\beta^2} + \frac{1}{\beta^2} = \frac{2}{\beta^2}$$

We solve:

$$\overline{X^2} = \frac{2}{\beta^2} \leftrightarrow \beta^2 = \frac{2}{\overline{X^2}} \leftrightarrow$$

$$\check{\beta} = \sqrt{\frac{2}{\overline{X^2}}}$$

where $\tilde{\beta}$ is the MME for β based on the second moment of X

Exercise 3

(a)

We know that the prior on p has a $Beta(2, 3)$ distribution.

$Beta$ distribution has the following probability density function:

$$f_X(t) = \frac{t^{\alpha-1}(1-t)^{\beta-1}}{B(\alpha, \beta)}$$

Then,

$$\pi(p) \propto \frac{p^{\alpha-1}(1-p)^{\beta-1}}{B(\alpha, \beta)}$$

$$\pi(p) \propto p^{2-1}(1-p)^{2-1} = p(1-p)^2$$

To find the mode of the prior, we use the following formula:

Mode of a $Beta(\alpha, \beta)$ distribution is $\frac{\alpha-1}{\alpha+\beta-2}$

So the most likely value of p is:

$$\frac{2-1}{2+3-2} = \frac{1}{3}$$

(b)

We are given the sample size $n = 20$ with $x = 13$ successes and $n - x = 7$ failures.

$$\pi(p|X) = \frac{f_p(x) \cdot \pi(p)}{f(x)} \propto f_p(x) \cdot \pi(p)$$

$$f_p(x) \cdot \pi(p) = p(p-1)^2 \cdot \binom{n}{x} p^x (1-p)^{n-x} \propto p(p-1)^2 \cdot p^x(1-p)^{n-x} = p^{x+1}(1-p)^{n-x+2}$$

The latter expression suggests that the posterior $\pi(p|X)$ would be $Beta(\alpha, \beta)$ distribution. Next we find the posterior parameters:

$$\alpha - 1 = x + 1 \leftrightarrow \boldsymbol{\alpha = x + 2}$$

$$\beta - 1 = n - x + 2 \leftrightarrow \boldsymbol{\beta = n - x + 3}$$

Therefore we find the posterior distribution with plugged in values for n and x to be the following:

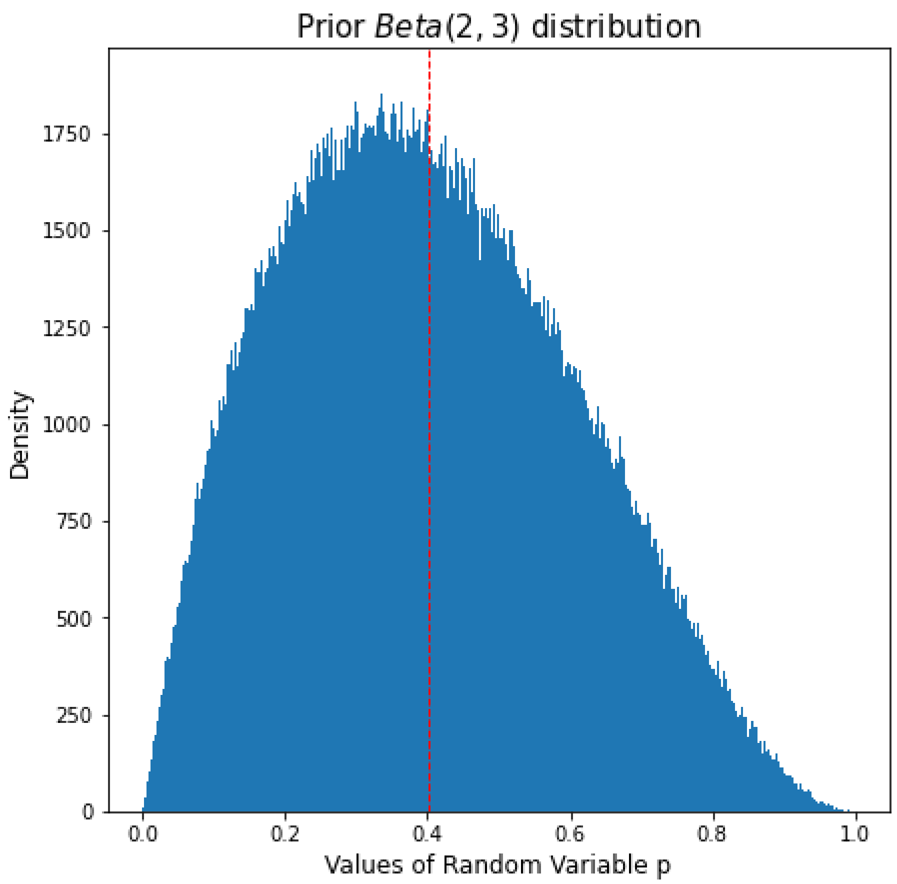
$$\pi(p|X) \sim Beta(x + 2, n - x + 3) = Beta(15, 10)$$

(c)

In [50]:

```
import numpy as np
import matplotlib.pyplot as plt
import matplotlib.mlab as mlab
# Set the shape paremeters
a, b = 2, 3
# Generate the value between
p1 = np.random.beta(a,b,2000000)

plt.figure(figsize=(7,7))
bins = np.linspace(0,1,2000)
plt.hist(p1, bins = bins)
plt.title('Prior $Beta(2,3)$ distribution', fontsize='15')
plt.xlabel('Values of Random Variable p', fontsize='12')
plt.ylabel('Density', fontsize='12')
plt.axvline(p1.mean(), color='r', linestyle='dashed', linewidth=1)
plt.show()
```

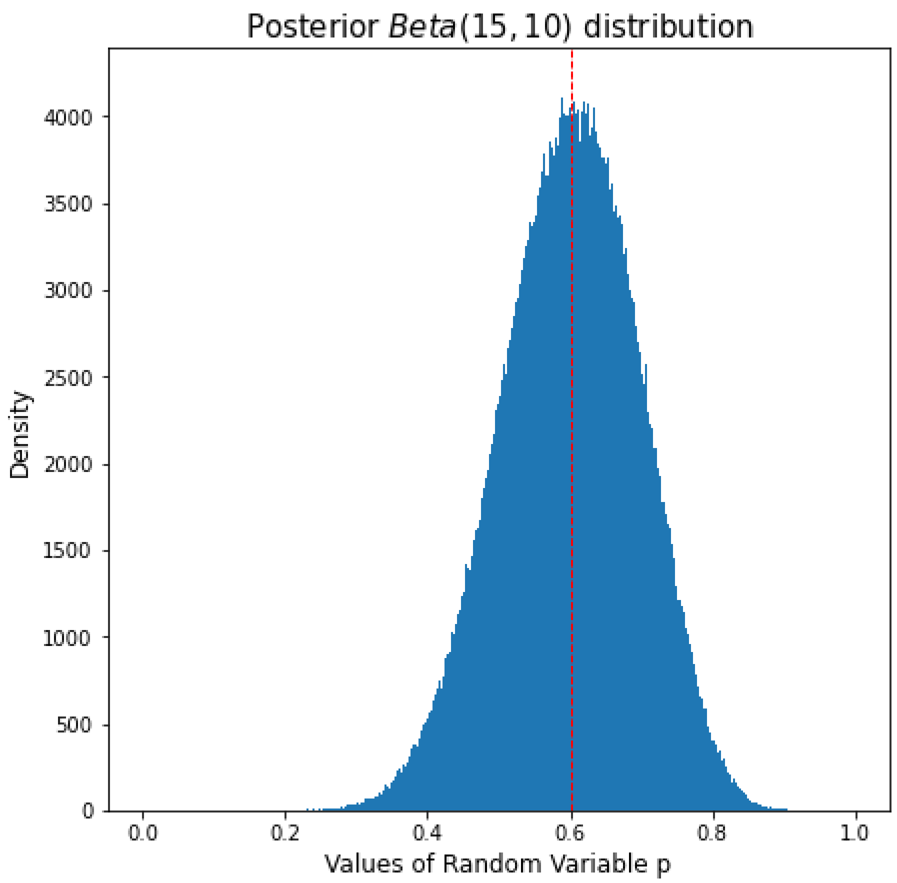


In [49]:

```
import numpy as np
import matplotlib.pyplot as plt

# Set the shape paremeters
a, b = 15, 10
# Generate the value between
p2 = np.random.beta(a,b,2000000)

plt.figure(figsize=(7,7))
bins = np.linspace(0,1,2000)
plt.hist(p2, bins = bins)
plt.title('Posterior $Beta(15,10)$ distribution', fontsize='15')
plt.xlabel('Values of Random Variable p', fontsize='12')
plt.ylabel('Density', fontsize='12')
plt.axvline(p2.mean(), color='r', linestyle='dashed', linewidth=1)
plt.show()
```



(d)

In order to compare the prior and posterior distributions in terms of location and dispersion (spread), we use the following:

- **Mean** to signify the location (red dotted line in the graphs);
- **Visual analysis** for the skewness;

- **Standard Deviation** to quantify the spread of distribution

In [48]:

```
pPrior= p1.mean()
pPost = p2.mean()
print(r'Prior Beta(2,3) distribution estimates p = %.4f'% pPrior)
print(r'Posterior Beta(15,10) distribution estimates p = %.4f'% pPost)
```

```
Prior Beta(2,3) distribution estimates p = 0.3998
Posterior Beta(15,10) distribution estimates p = 0.5998
```

So we find that the **mean of prior is 2/3 that of the posterior**.

From visual inspection of the plots we can state that the **prior distribution presents with some positive skewness**, whilst the **posterior is centered around the mean**. Additionally we can see that prior distribution seems to have a larger spread. This can be quantified by contrasting the standard deviation parameters of both distributions:

In [52]:

```
priorStD = np.std(p1)
postStD = np.std(p2)
print(r'Prior Beta(2,3) distribution StDev = %.4f'% priorStD)
print(r'Posterior Beta(15,10) distribution StDev = %.4f'% postStD)
```

```
Prior Beta(2,3) distribution StDev = 0.2002
Posterior Beta(15,10) distribution StDev = 0.0962
```

Where we find the prior distribution to have more than twice as large spread as the posterior.