# Statistical Data Analysis

ASSIGNMENT 1

Fall 2020

Instructor: Dr. Mainak Thakur

# Table of contents

Title Page no.

# Procedure for generating custom PMF and PDF (without library functions)

- For generating the PMF or PDF of distributions, we just wrote a function **(that executes PDF or PMF formula)** that takes in value of x and gives the corresponding PDF or PMF value depending upon the distribution
- **For example:**
    - Let us take **Poisson Distribution**
    - The formula of poisson distribution is:

$$f(k;\lambda) = Pr(X = k) = \frac{\lambda^k \exp(-\lambda)}{k!}$$

    - We simply wrote a function that outputs Poisson PMF value of x when x is given, the code snippet is below:

```
def my_poisson_function(x,lambda_value):
    return (np.power(lambda_value,x,dtype=np.float64)*(np.exp(-lambda_value)))/(scipy.special.factorial(x))
```

## Procedure for generating samples from distributions (without library functions)

- Once we are able to generate PMFs or PDFs the next thing to do is to drawing the samples from the distribution.
- We wrote a function to generate samples given a distribution, It takes the range of values Xo and X1 between which we wish to sample, no of samples we want, and the distribution PMFs or PDFs as input.
  - First we sample a random value **x**(using uniform distribution) between **Xo** and **X1** and find the **probability of x .**
  - Then we compare its probability to random value sampled between 0 and 1(uniform distribution between 0 and 1)
  - If **probability of x > random uniform sampling between 0 and 1,** then we keep x in the sample otherwise we don't
  - **We keep on doing this until we get required no of samples.**
  - In this way , we draw samples from a any distribution **(without using library functions)**

**Note: we only used code snippet once in previous slide to explain how we are generating PMFs and PDFs**

# Binomial Distribution

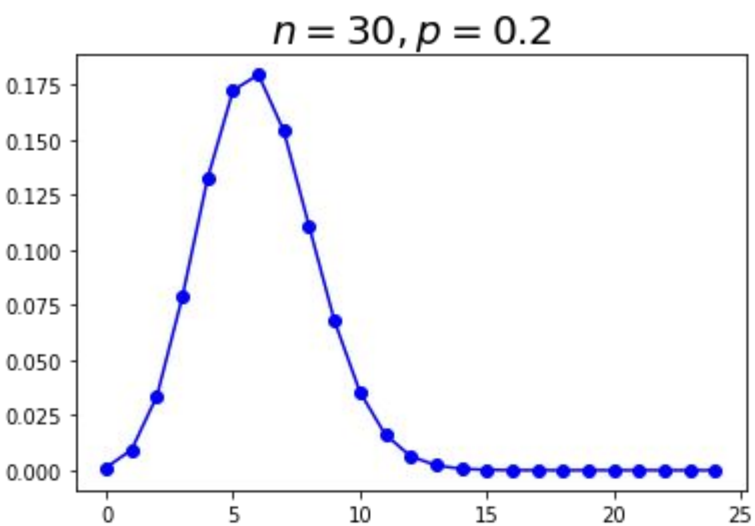
- The PMF(Probability Mass Function) of Binomial distribution is as follows:

$$f(k, n, p) = \Pr(k; n, p) = \Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

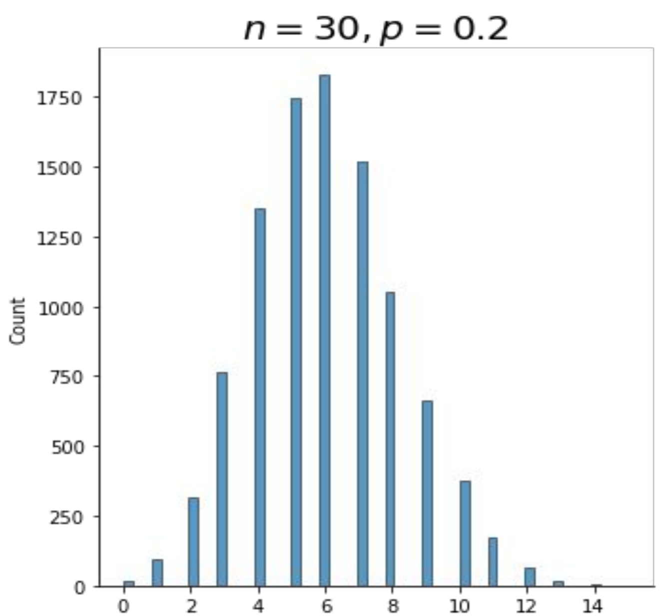
- Here *n*, *p* are the parameters. The mean and variance of the distribution depends on these parameters
- The formula for calculating mean and variance of the distributions are as follows:
    - Mean: np
    - Variance: np(1-p)

# Plots of Binomial distribution for various values of n, p (1)

- Graph of PMF




- Graph of values sampled from distribution(10,000 samples)

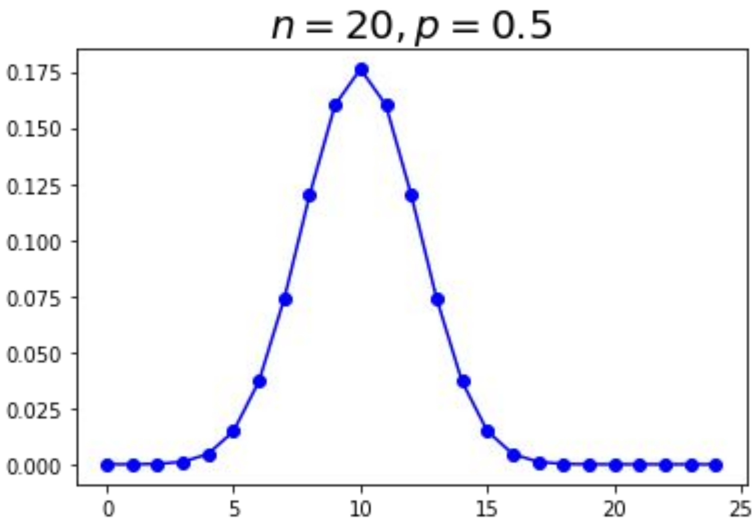



- Theoretical mean = 6.0
- Theoretical variance = 4.8

- Sample mean = 5.999
- Sample variance = 4.78789

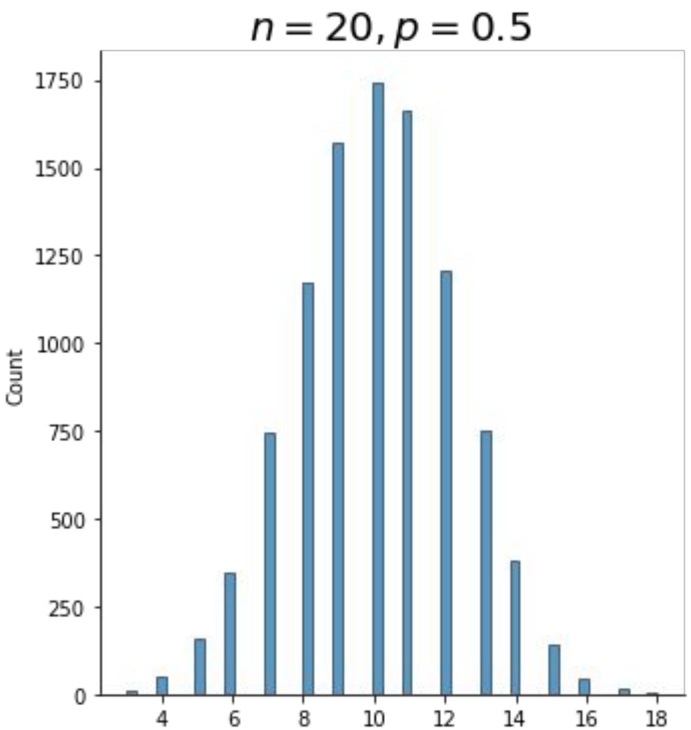
# Plots of Binomial distribution for various values of n, p (2)

- Graph of PMF




- Theoretical mean = 10.0
- Theoretical variance = 5.0

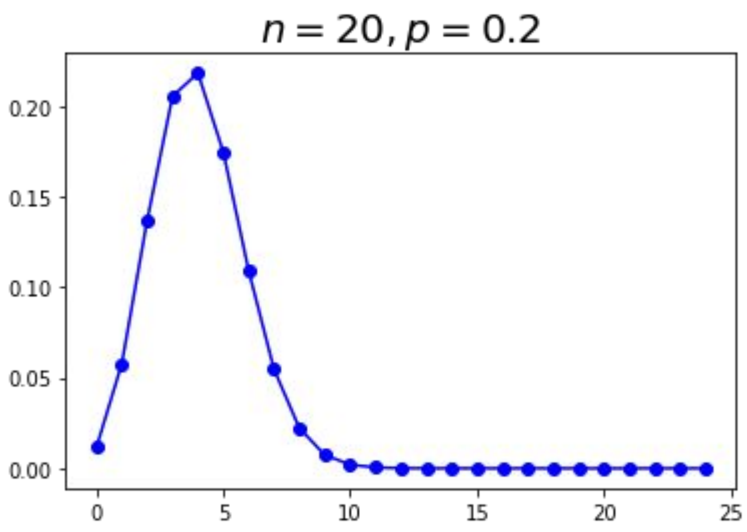
- Graph of values sampled from distribution(10,000 samples)



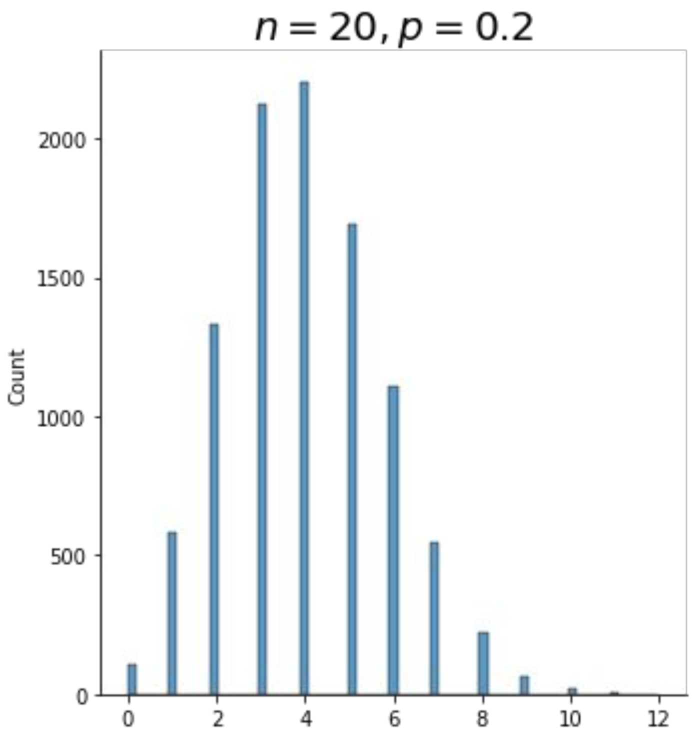

- Sample mean = 10.0241
- Sample variance = 5.02331919

# Plots of Binomial distribution for various values of n, p (3)

- Graph of PMF




- Graph of values sampled from distribution(10,000 samples)




- Theoretical mean = 4.0
- Theoretical variance = 3.2

- Sample mean = 3.996
- Sample variance = 3.1529

# Verifying Central Limit Theorem for Binomial Distribution

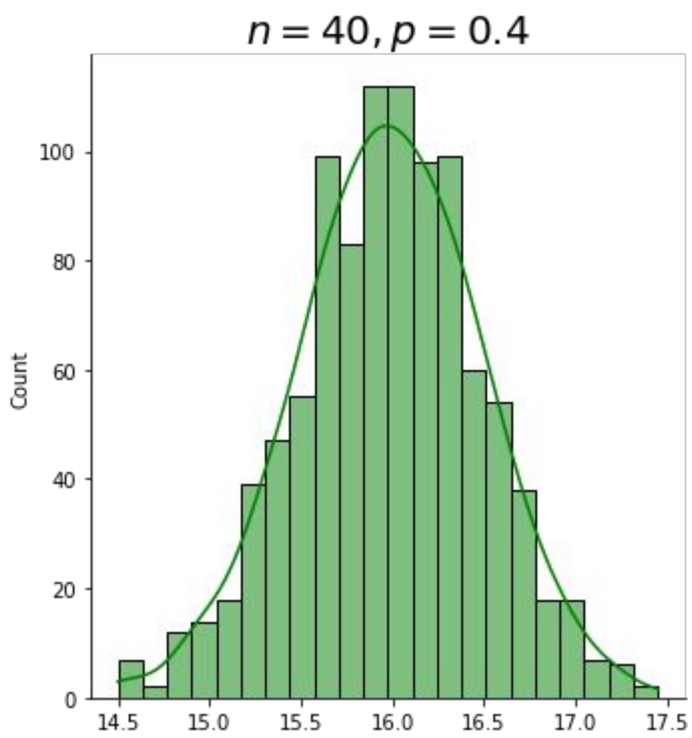
- Parameters:
  - n = 40
  - p = 0. 4
  - **n** = 40 (each sample is average of 40 samples)
  - No of samples = 1000

- Mean of Binomial Distribution: np = 16

- Variance of Binomial Distribution: np(1-p) = 9.6

- If Central limit theorem holds true then for then **mean and Variance** of the expectation of static sample should be approximately
  - Mean : $\mu = \mu_{distribution}$ = 16
  - Variance : $\sigma^2 = \frac{\sigma^2_{distribution}}{n}$ = 0.24

# Plot of Sample Mean  Statistic




- Theoretical mean : 16.0
- Theoretical variance: 0.24

- Expectation of sample mean : 15.982
- Variance of sample mean: 0.2486

- As you can see from the above values that mean and variance of sample mean is approximately equal to theoretical ones.

- **Hence, Central Limit Theorem holds for the Binomial Distribution**

# Geometric Distribution

- The PMF(Probability Mass Function) of Geometric distribution is as follows:

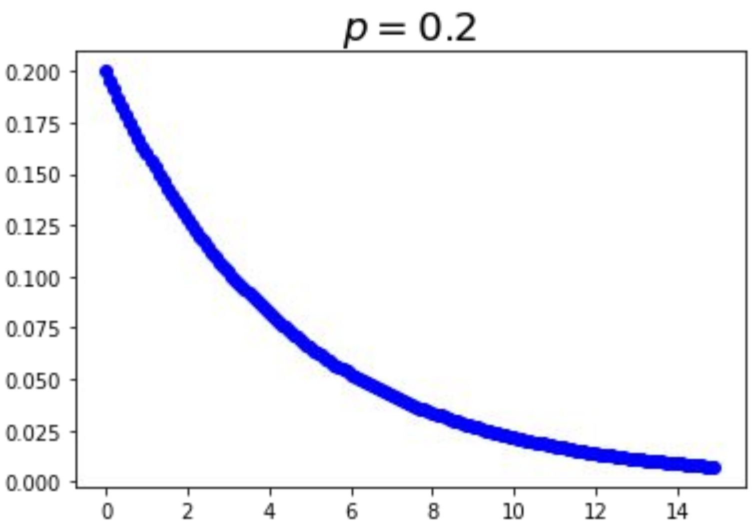
$$\Pr(Y = k) = (1 - p)^k p$$

for $k = 0, 1, 2, 3, ....$

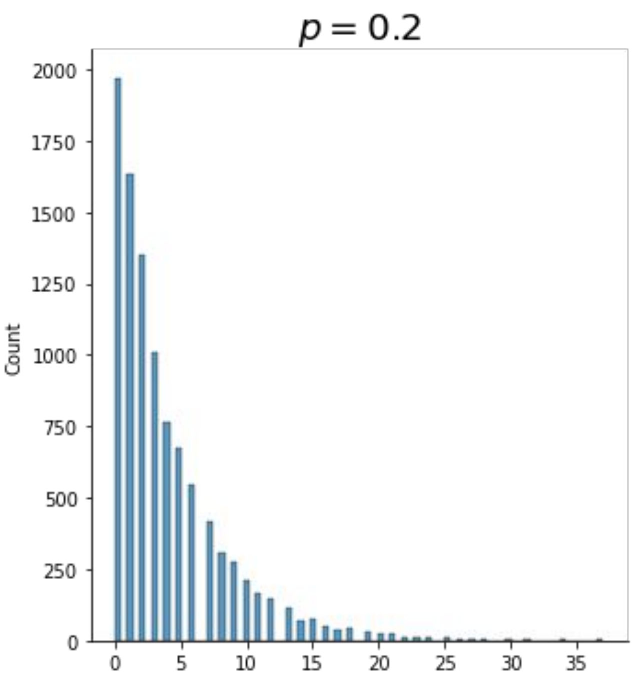
- Here $p$ is the parameter. The mean and variance of the distribution depends on these parameter
- The formula for calculating mean and variance of the distributions are as follows:
  - Mean: $\frac{1-p}{p}$
  - Variance: $\frac{1-p}{p^2}$

# Plots of Geometric distribution for various values of p (1)

- Graph of PMF




- Graph of values sampled from distribution(10,000 samples)

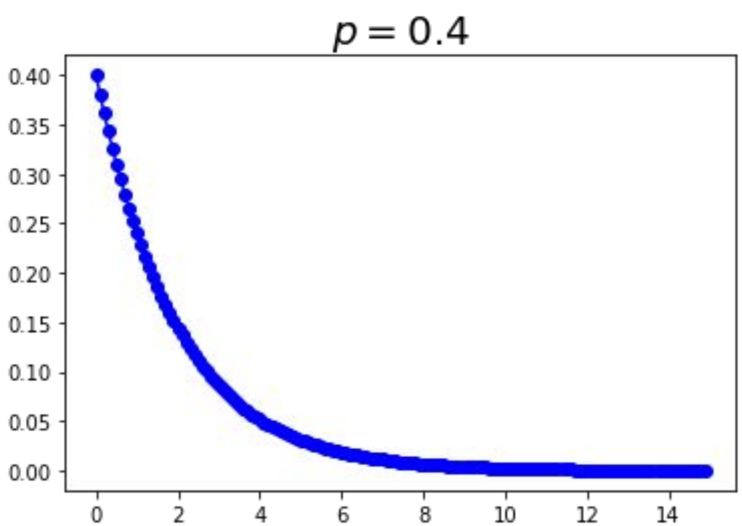



- Theoretical mean = 4.0
- Theoretical variance = 19.999

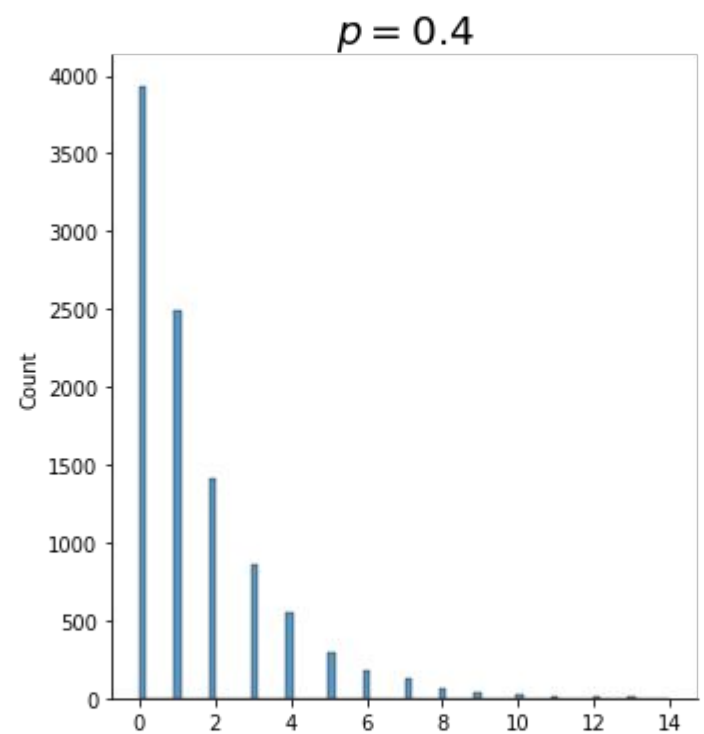
- Sample mean = 3.9788
- Sample variance = 19.87015056

# Plots of Geometric distribution for various values of p (2)

- Graph of PMF




- Graph of values sampled from distribution(10,000 samples)

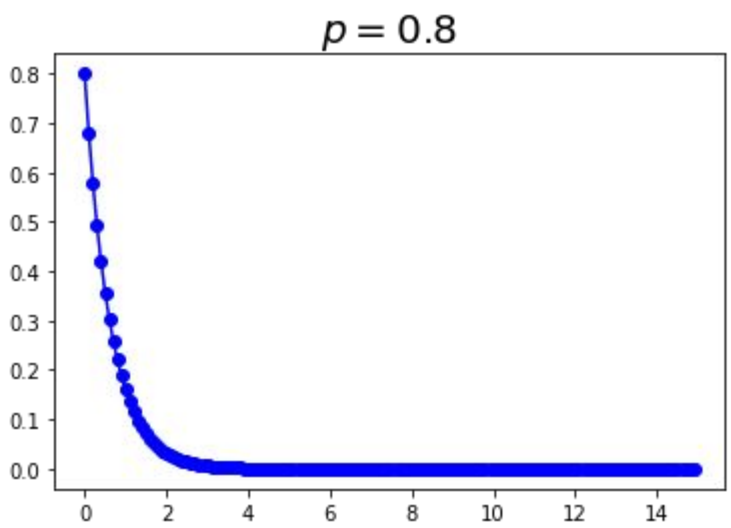



- Theoretical mean = 1.4999
- Theoretical variance = 3.7499

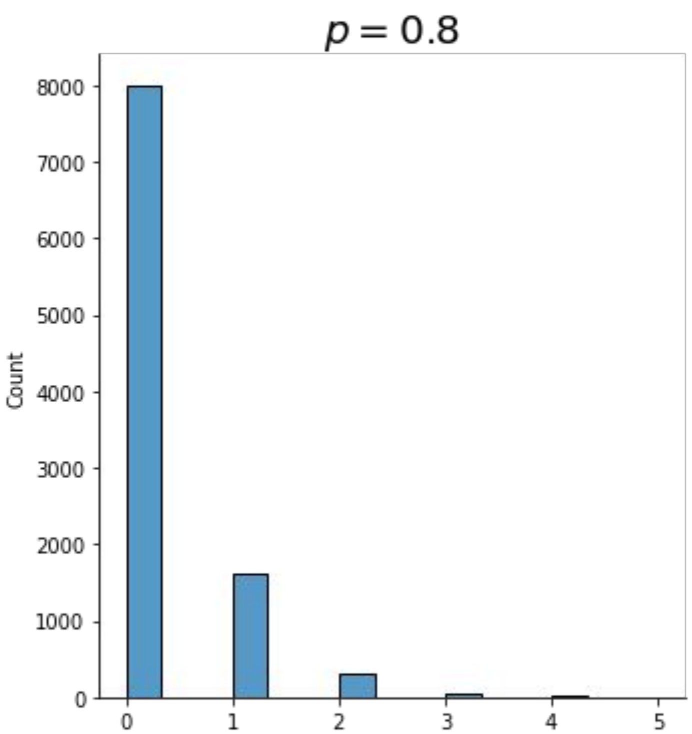
- Sample mean = 1.4968
- Sample variance = 3.62238

# Plots of Geometric distribution for various values of p (3)

- Graph of PMF




- Graph of values sampled from distribution(10,000 samples)




- Theoretical mean = 0.24999
- Theoretical variance = 0.312499

- Sample mean = 0.2476
- Sample variance = 0.30729424

# Verifying Central Limit Theorem for Geometric Distribution

- Parameters:
  - p = 0 .4
  - **n** = 40 (each sample is average of 40 samples)
  - No of samples = 1000

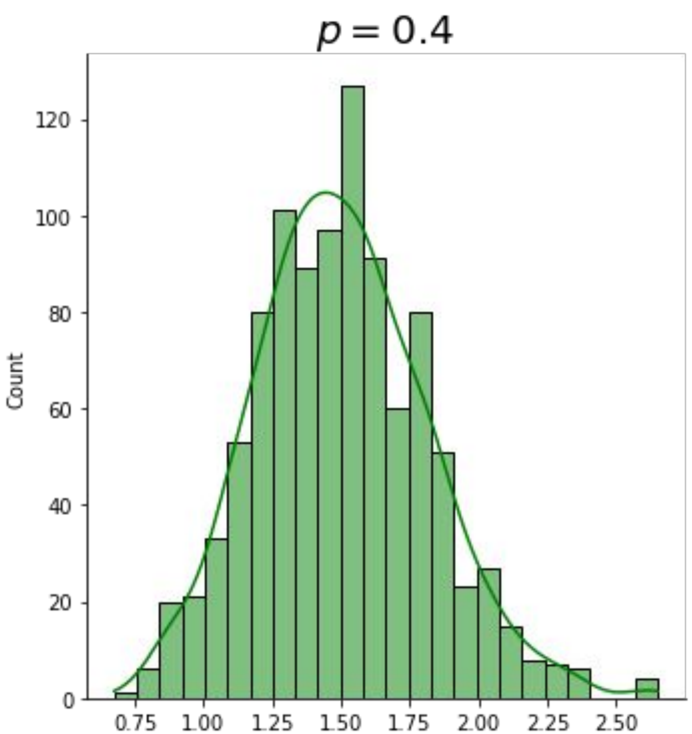
- Mean of Geometric Distribution: $\frac{1-p}{p}$ = 1.4999

- Variance of Geometric Distribution: $\frac{1-p}{p^2}$ = 3.75

- If Central limit theorem holds true then for then **mean and Variance** of the expectation of static sample should be approximately

  - Mean : $\mu = \mu_{distribution}$ = 1.4999
  - Variance : $\sigma^2 = \frac{\sigma^2_{distribution}}{n}$ = 0.0937
  -

# Plot of Sample Mean  Statistic




- Theoretical mean : 1.49999
- Theoretical variance: 0.09374999

- Expectation of sample mean : 1.501225
- Variance of sample mean: 0.096247874

- As you can see from the above values that mean and variance of sample mean is approximately equal to theoretical ones.

- **Hence, Central Limit Theorem holds for the Geometric Distribution**

# Poisson Distribution

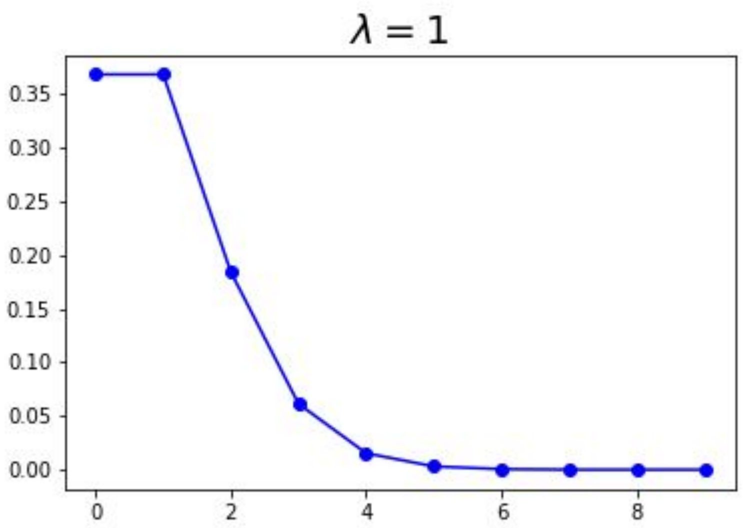
- The PMF(Probability Mass Function) of Poisson distribution is as follows:

$$f(k;\lambda) = Pr(X = k) = \frac{\lambda^k \exp(-\lambda)}{k!}$$

- Here $\lambda$ is the parameter. The mean and variance of the distribution depends on these parameter
- The formula for calculating mean and variance of the distributions are as follows:
  - Mean: $\lambda$
  - Variance: $\lambda$

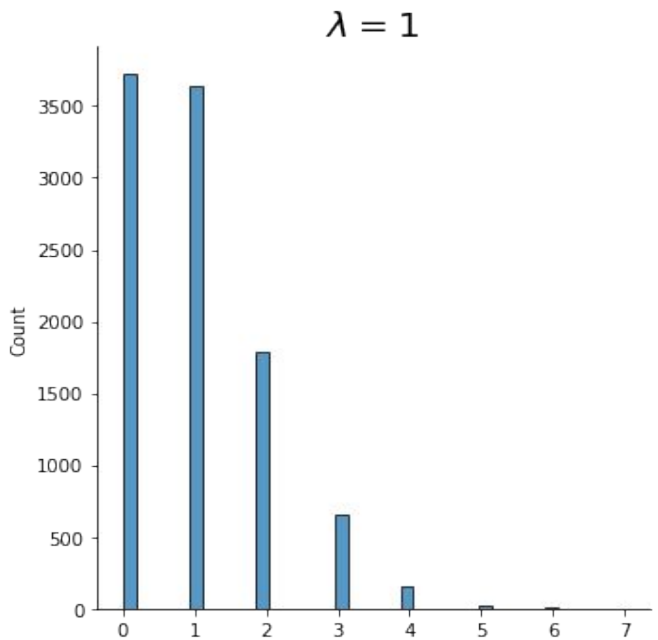
# Plots of Poisson distribution for various values of $\lambda$ (1)

- Graph of PDF




- Theoretical mean = 1
- Theoretical variance = 1

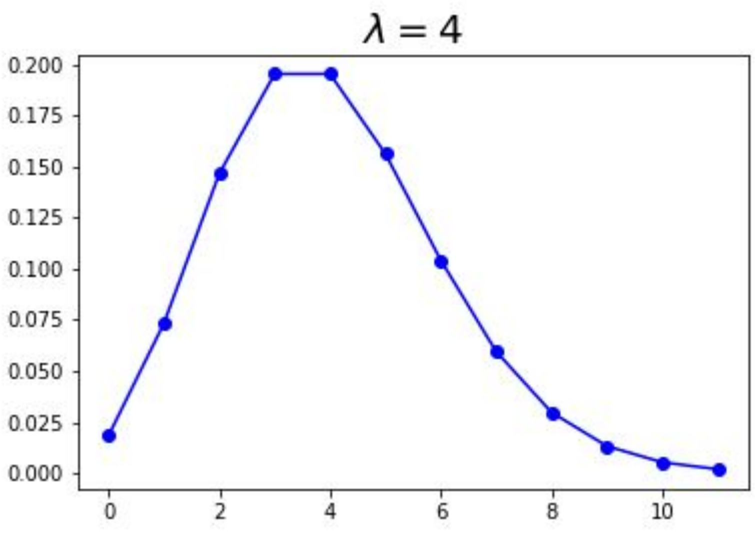
- Graph of values sampled from distribution(10,000 samples)




- Sample mean = 0.9969
- Sample variance = 1.0102

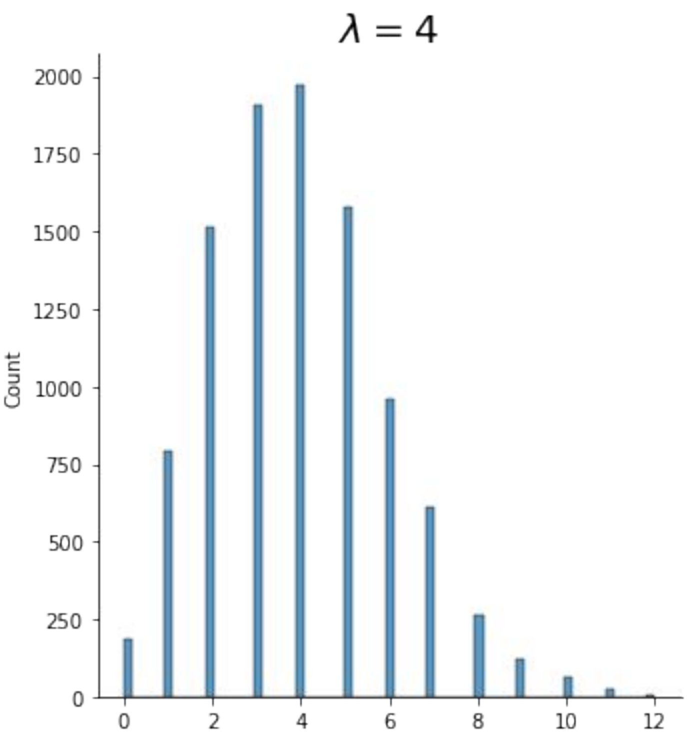
# Plots of Poisson distribution for various values of $\lambda$ (2)

- Graph of PMF




- Theoretical mean = 4.0
- Theoretical variance = 4.0

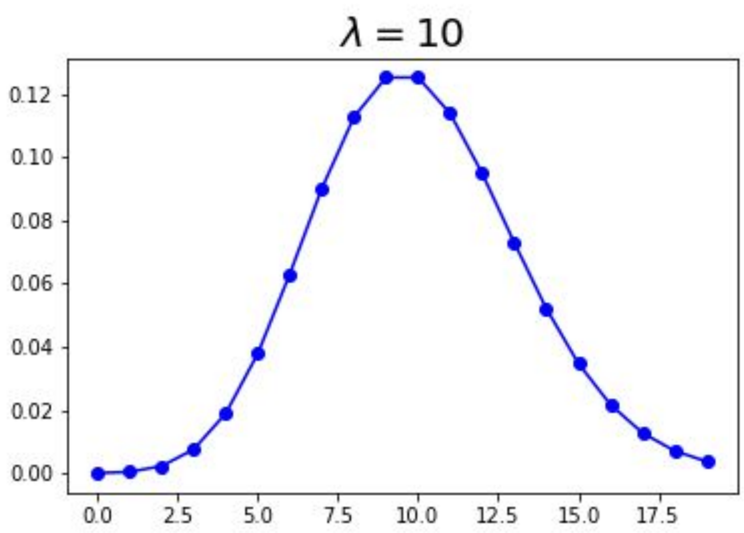
- Graph of values sampled from distribution(10,000 samples)




- Sample mean = 3.9510
- Sample variance = 4.004

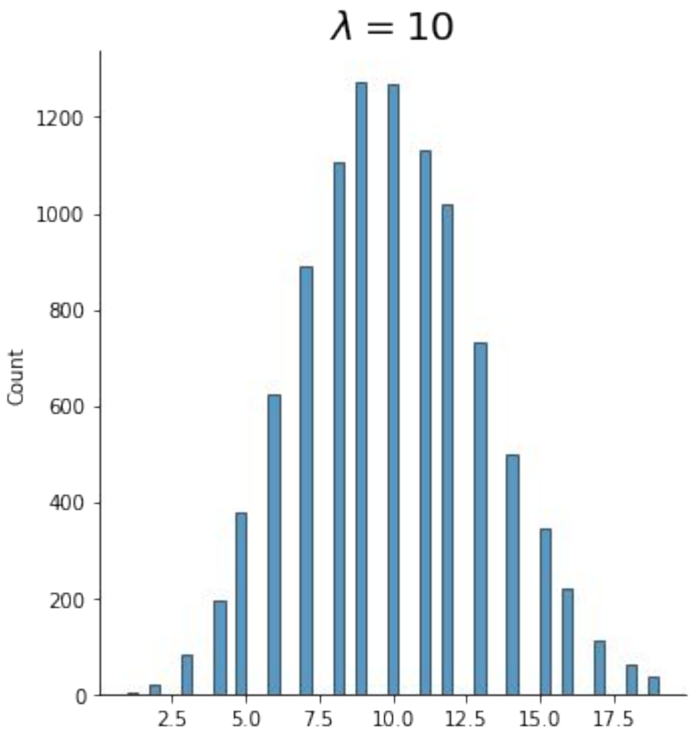
# Plots of Poisson distribution for various values of $\lambda$ (3)

- Graph of PMF




- Theoretical mean = 10.0
- Theoretical variance = 10.0

- Graph of values sampled from distribution(10,000 samples)




- Sample mean = 9.9513
- Sample variance = 9.4629

# Verifying Central Limit Theorem for Poisson Distribution

- Parameters:
  - $\lambda$ = 10
  - **n** = 40 (each sample is average of 40 samples)
  - No of samples = 1000

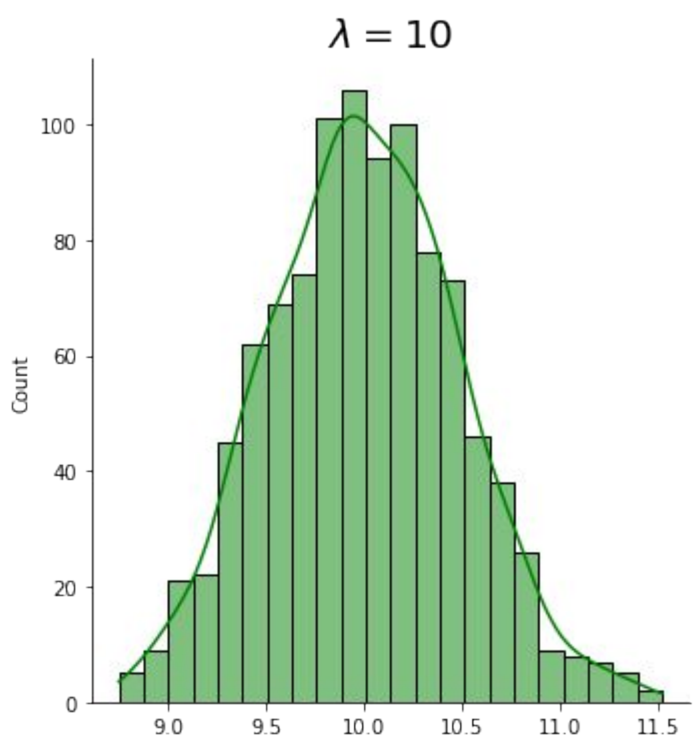
- Mean of Poisson Distribution: $\lambda$ = 10

- Variance of Poisson Distribution: $\lambda$ = 10

- If Central limit theorem holds true then for then **mean and Variance** of the expectation of static sample should be approximately

  - Mean : $\mu = \mu_{distribution}$ = 10
  - Variance : $\sigma^2 = \dfrac{\sigma^2_{distribution}}{n}$ = 0.25

# Plot of Sample Mean Statistic




- Theoretical mean : 10.0
- Theoretical variance: 0.25

- Expectation of sample mean:10.003
- Variance of sample mean: 0.230

- As you can see from the above values that mean and variance of sample mean is approximately equal to theoretical ones.

- **Hence, Central Limit Theorem holds for the Poisson Distribution**

# Negative Binomial Distribution

- The PMF(Probability Mass Function) of Negative Binomial distribution is as follows:

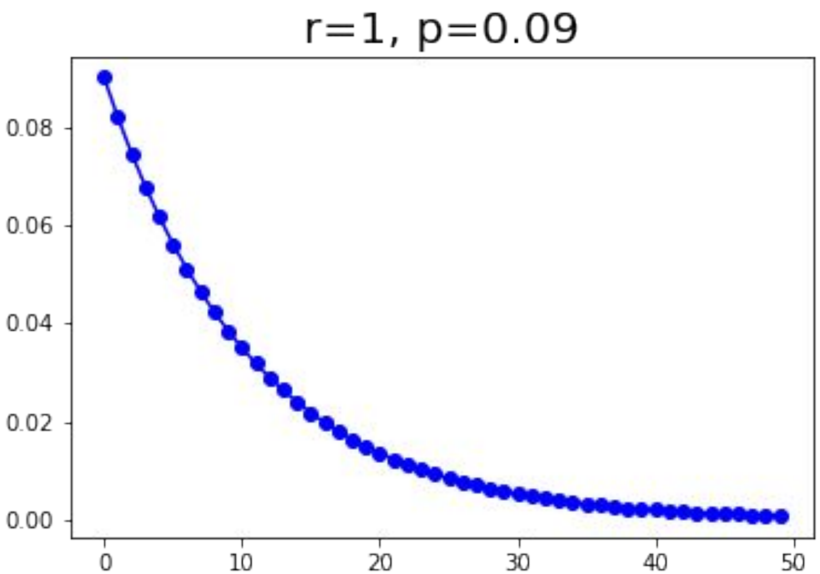
$$Pr(X = k) = \binom{k+r-1}{r-1}(1-p)^k p^r$$

- Here r, p are the parameters. The mean and variance of the distribution depends on these parameters

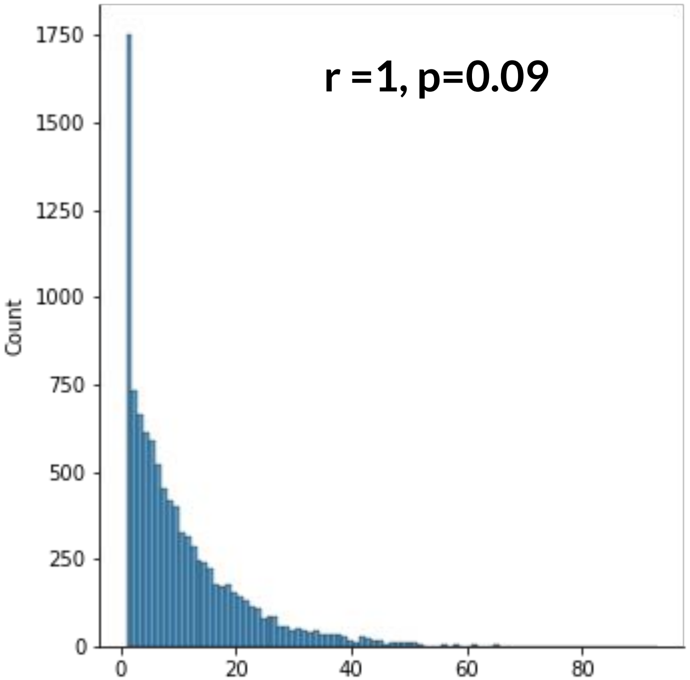
- The formula for calculating mean and variance of the distributions are as follows:
  - Mean: $$\frac{r*(1-p)}{p}$$
  - Variance: $$\frac{r*(1-p)}{p^2}$$

# Plots of Negative Binomial distribution for values: r=1, p=0.09

- Graph of PMF




- Graph of values sampled from distribution(10,000 samples)

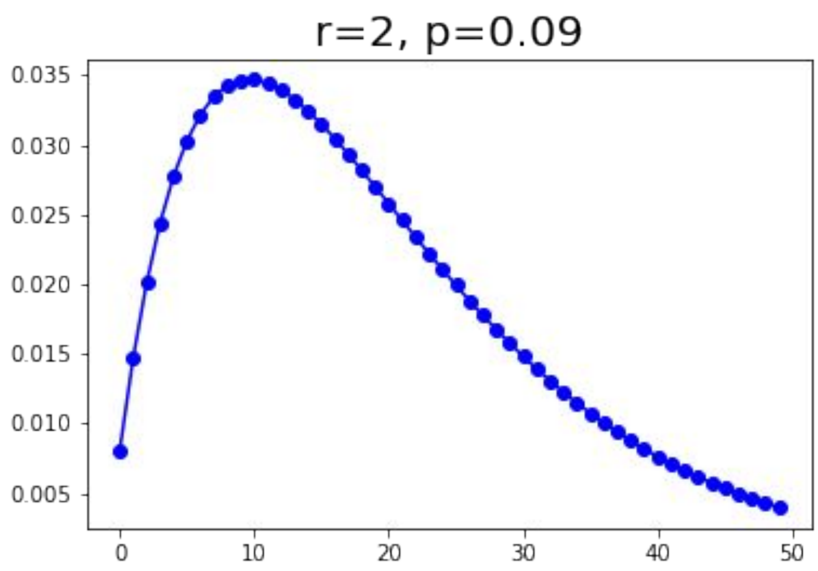



- Theoretical mean = 10.12
- Theoretical variance = 112.34

- Sample mean = 11.11
- Sample variance = 112.89

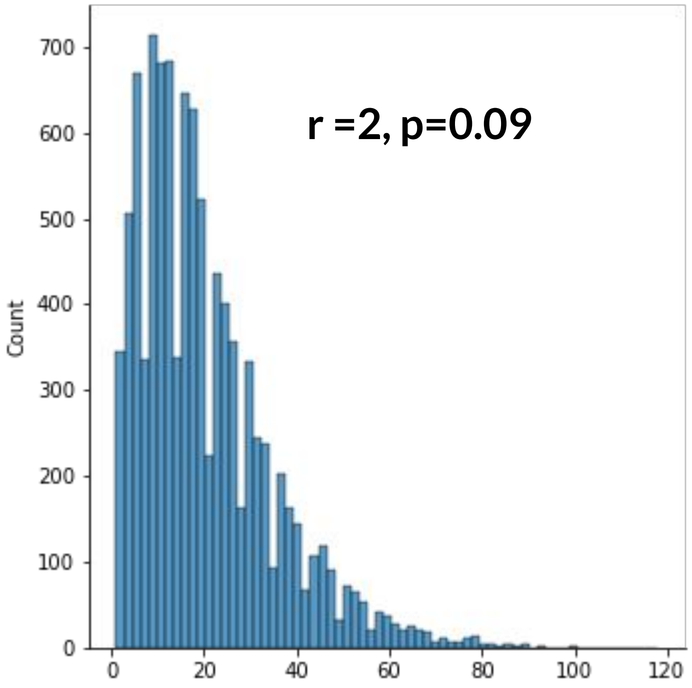
# Plots of Negative Binomial distribution for values: r=2, p=0.09

- Graph of PMF




- Theoretical mean = 20.22
- Theoretical variance = 224.69

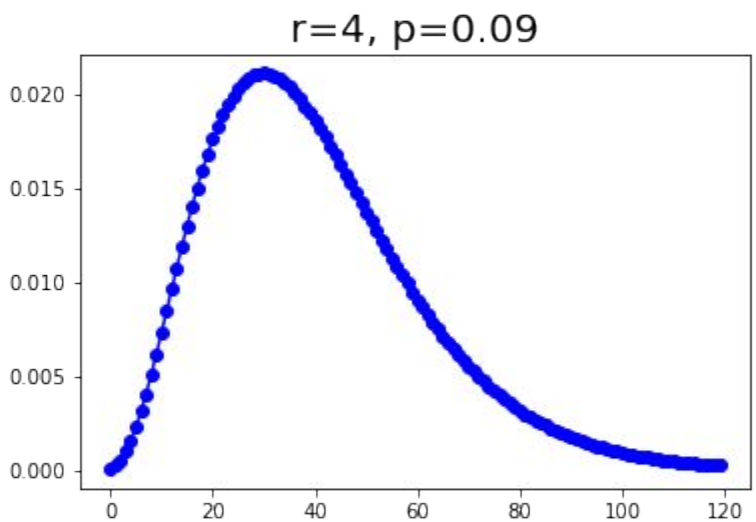
- Graph of values sampled from distribution(10,000 samples)




- Sample mean = 20.34
- Sample variance = 225.43

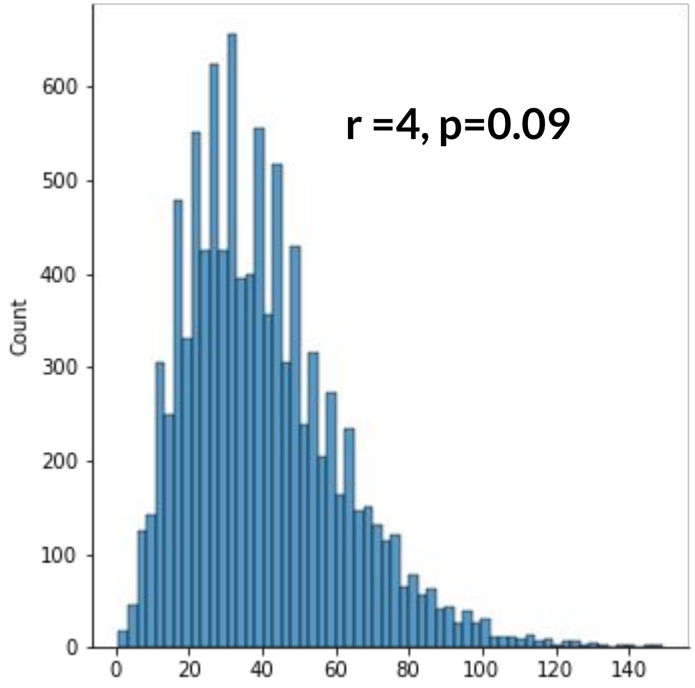
# Plots of Negative Binomial distribution for values: r=4, p=0.09

- Graph of PMF




- Theoretical mean = 40.44
- Theoretical variance = 449.38

- Graph of values sampled from distribution(10,000 samples)



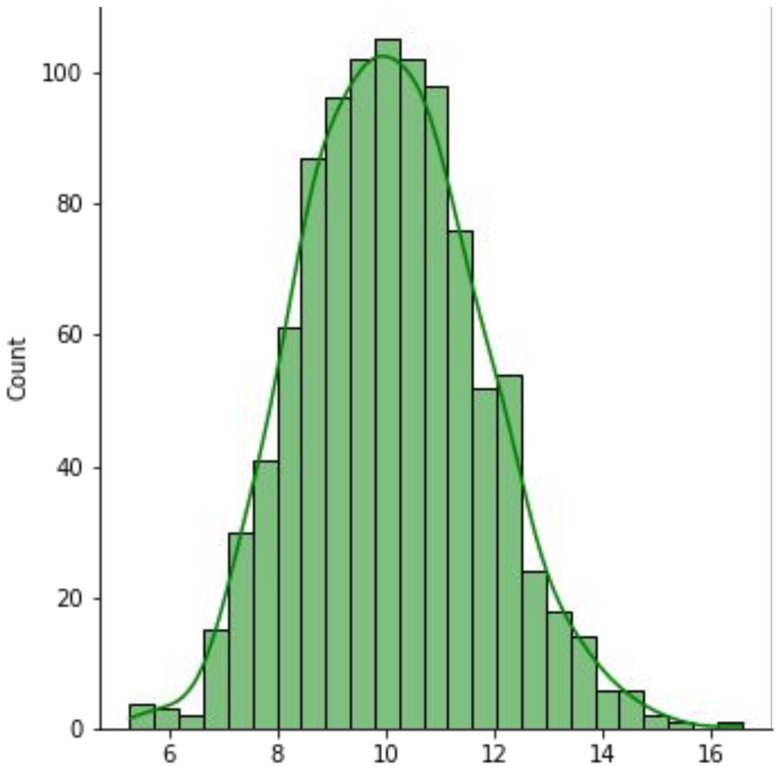

- Sample mean = 40.27
- Sample variance = 452.89

# Verifying Central Limit Theorem for Negative Binomial

- Parameters:
  - **r** = 1
  - **p** = 0.09
  - **n** = 40 (each sample is average of 40 samples)
  - No of samples = 1000
- Mean of Negative Binomial: $\frac{r * (1-p)}{p}$ = 10.1111
- Variance of Negative Binomial: $\frac{r * (1-p)}{p^2}$ = 112.3456
- If Central limit theorem holds true then for then **mean and Variance** of the expectation of static sample should be approximately
  - Mean : $\mu = \mu_{distribution}$ = 10.1111
  - Variance : $\sigma^2 = \frac{\sigma^2_{distribution}}{n}$ = 2.8086

# Plot of Sample Mean  Statistic

**r =1, p=0.09**




- Theoretical mean : 10.1111
- Theoretical variance: 2.8086

- Expectation of sample mean: 10.10
- Variance of sample mean: 2.75

- As you can see from the above values that mean and variance of sample mean is approximately equal to theoretical ones.

- **Hence, Central Limit Theorem holds for the Negative Binomial Distribution**

# Discrete Uniform Distribution

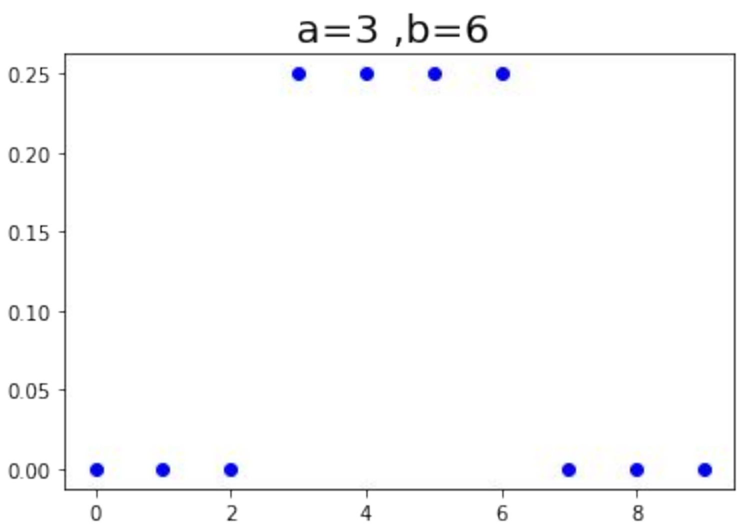
- The PMF(Probability Mass Function) of **Discrete Uniform distribution** is as follows here both **a** and **b** are inclusive:

$$f(x; a, b) = Pr(x = k) = \frac{1}{b - a + 1}$$

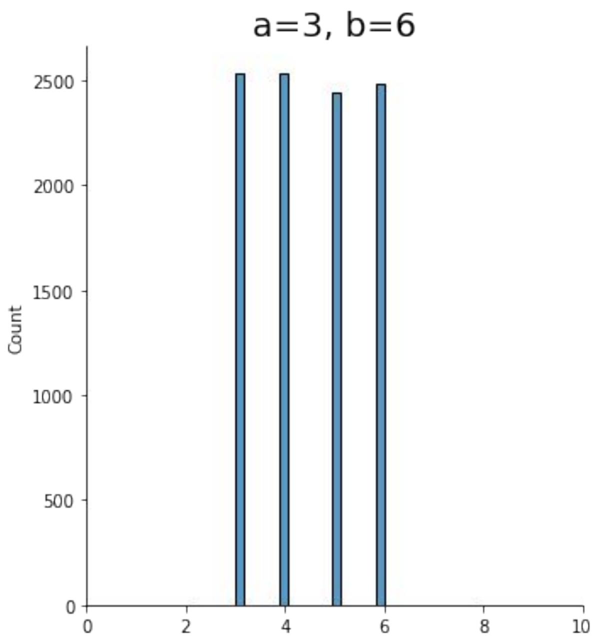
- Here **a** and **b** are parameters. The mean and variance of the distribution depends on these parameter
- The formula for calculating mean and variance of the distributions are as follows:
  - Mean: $\frac{a+b}{2}$
  - Variance: $\frac{(b-a+1)^2 - 1}{12}$

# Plots of Discrete Uniform distribution for various values of **a** and **b** (1)

- Graph of PMF




- Graph of values sampled from distribution(10,000 samples)

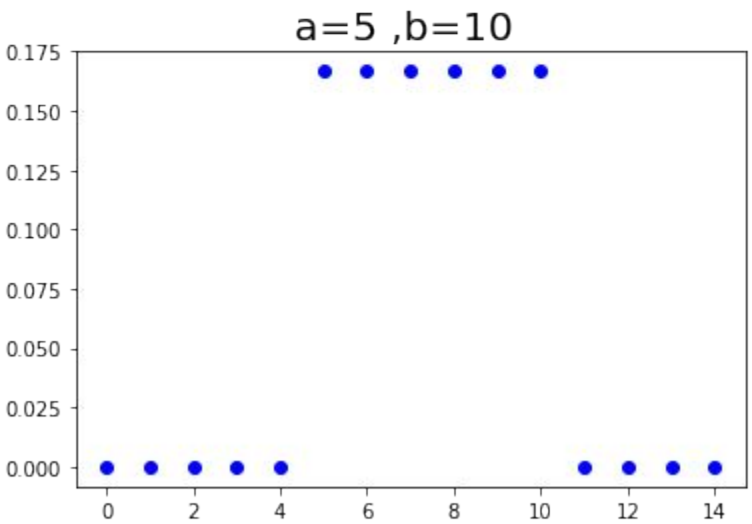



- Theoretical mean = 4.5
- Theoretical variance = 1.25

- Sample mean = 4.4881
- Sample variance = 1.25445839

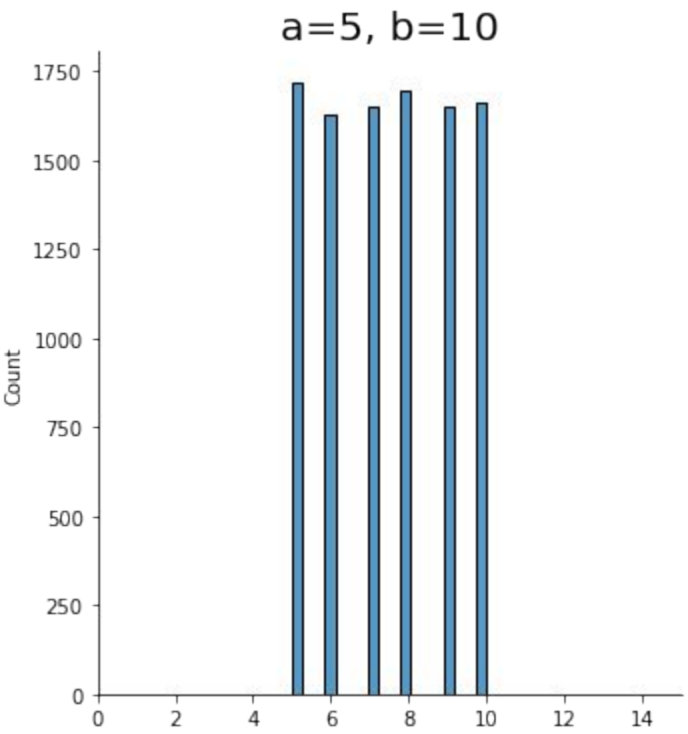
# Plots of Discrete Uniform distribution for various values of **a** and **b** (2)

- Graph of PMF




- Theoretical mean = 7.5
- Theoretical variance = 2.9166

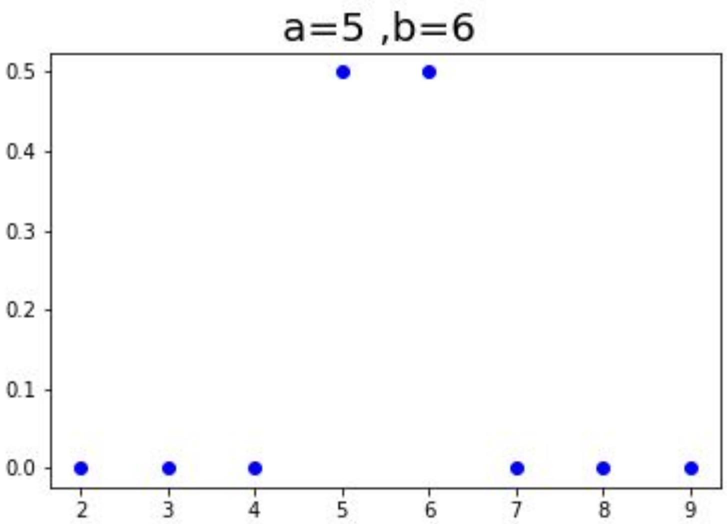
- Graph of values sampled from distribution(10,000 samples)



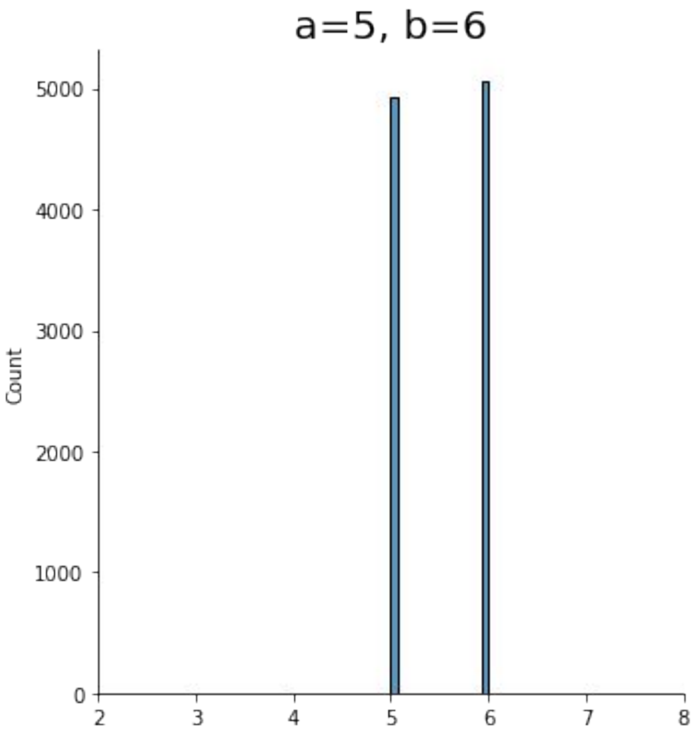

- Sample mean = 7.4902
- Sample variance = 2.9329

# Plots of Discrete Uniform distribution for various values of **a** and **b** (3)

- Graph of PMF




- Graph of values sampled from distribution(10,000 samples)



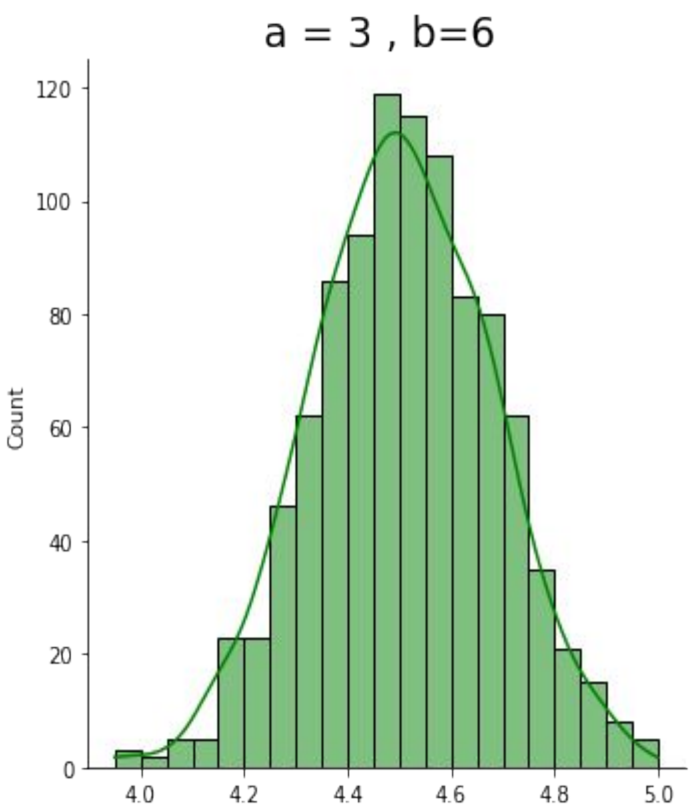

- Theoretical mean = 5.5
- Theoretical variance = 0.25

- Sample mean = 5.5066
- Sample variance = 0.2499

# Verifying Central Limit Theorem for Poisson Distribution

- Parameters:
  - **a** = 3 and **b** = 6
  - **n** = 40 (each sample is average of 40 samples)
  - No of samples = 1000
- Mean of Discrete Uniform Distribution: $\frac{a+b}{2}$ = 4.5
- Variance of Discrete Uniform Distribution: $\frac{(b-a+1)^2-1}{12}$ = 1.25
- If Central limit theorem holds true then for then **mean and Variance** of the expectation of static sample should be approximately
  - Mean : $\mu = \mu_{distribution}$ = 4.5
  - Variance : $\sigma^2 = \frac{\sigma^2_{distribution}}{n}$ = 0.03125

# Plot of Sample Mean  Statistic




- Theoretical mean : 4.5
- Theoretical variance: 0.03125

- Expectation of sample mean: 4.5024
- Variance of sample mean : 0.0299

- As you can see from the above values that mean and variance of sample mean is approximately equal to theoretical ones.

- **Hence, Central Limit Theorem holds for the Discrete Uniform Distribution**

# Normal Distribution

- The PDF(Probability Density Function) of Normal distribution is as follows:

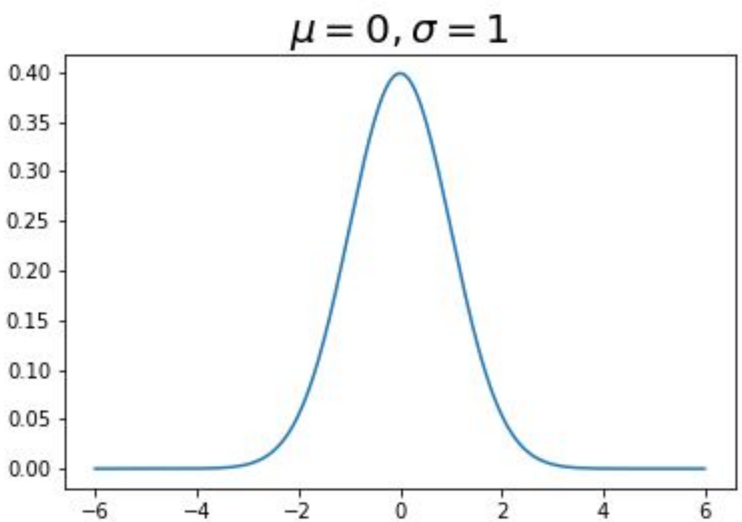
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

- Here $\mu$ and $\sigma$ are the parameters. The mean and variance of the distribution depends on these parameters

- The formula for calculating mean and variance of the distributions are as follows:
  - Mean: $\mu$
  - Variance: $\sigma^2$

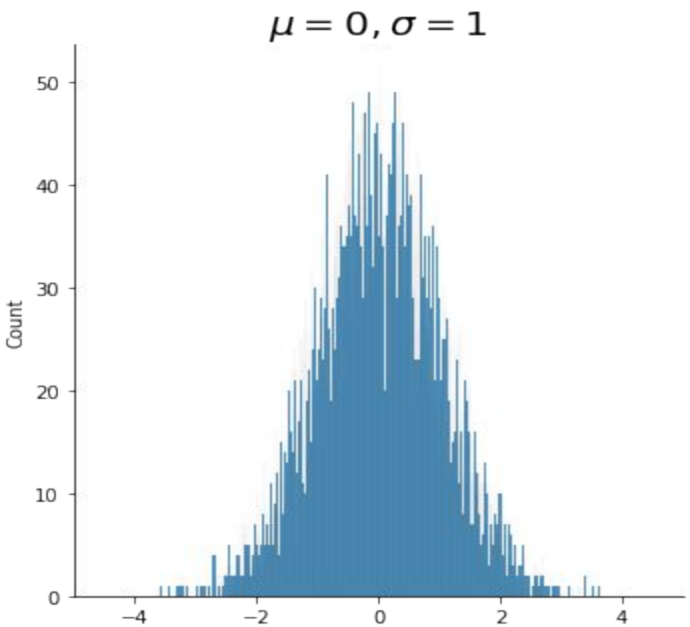
# Plots of Normal distribution for value: $\mu$=0 , **σ**= 1 (1)

- Graph of PDF




- Theoretical mean = 0.00
- Theoretical variance = 1.00

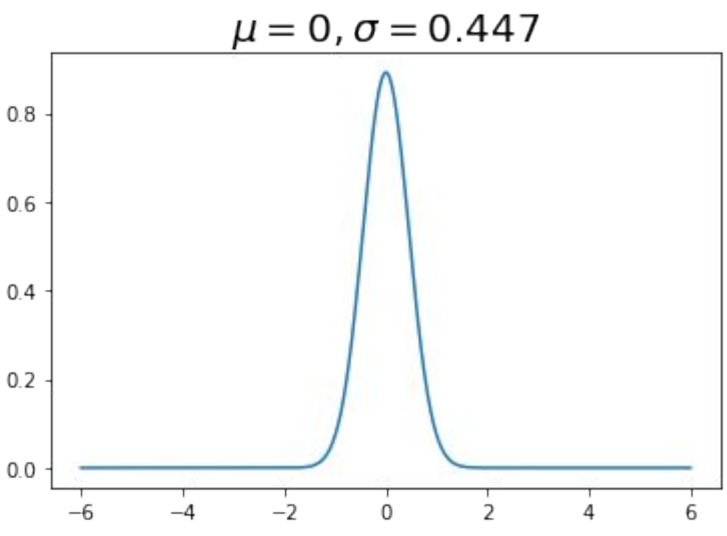
- Graph of values sampled from distribution(10,000 samples)




- Sample mean = 0.004
- Sample variance = 1.004

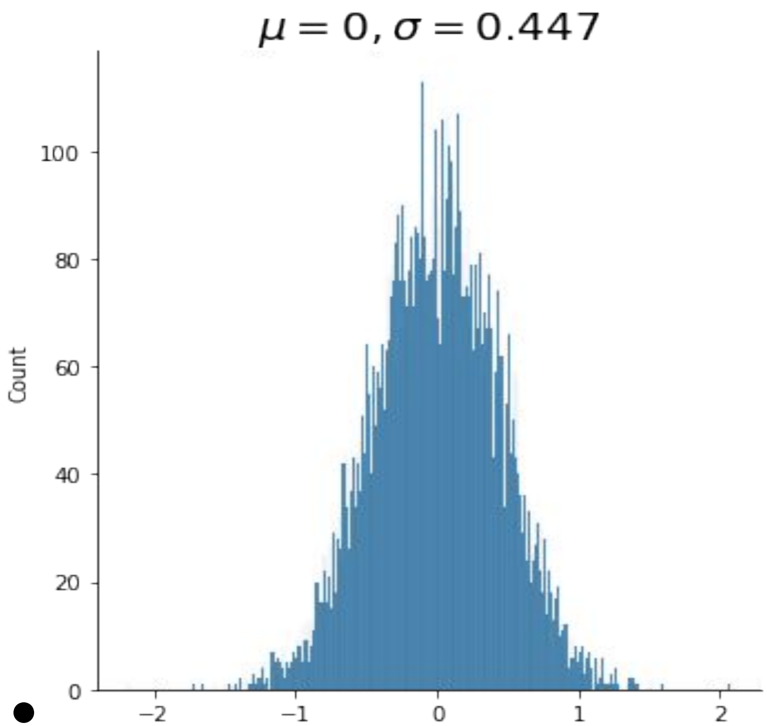
# Plots of Normal distribution for value: $\mu$=0 , **σ** = 0.447 (2)

- Graph of PDF




- Theoretical mean = 0.00
- Theoretical variance = 0.19

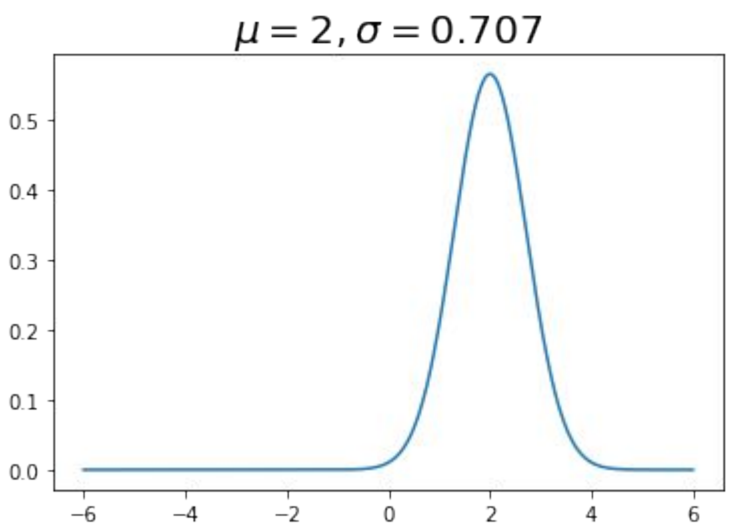
- Graph of values sampled from distribution(10,000 samples)




- Sample mean = 0.015
- Sample variance = 0.20

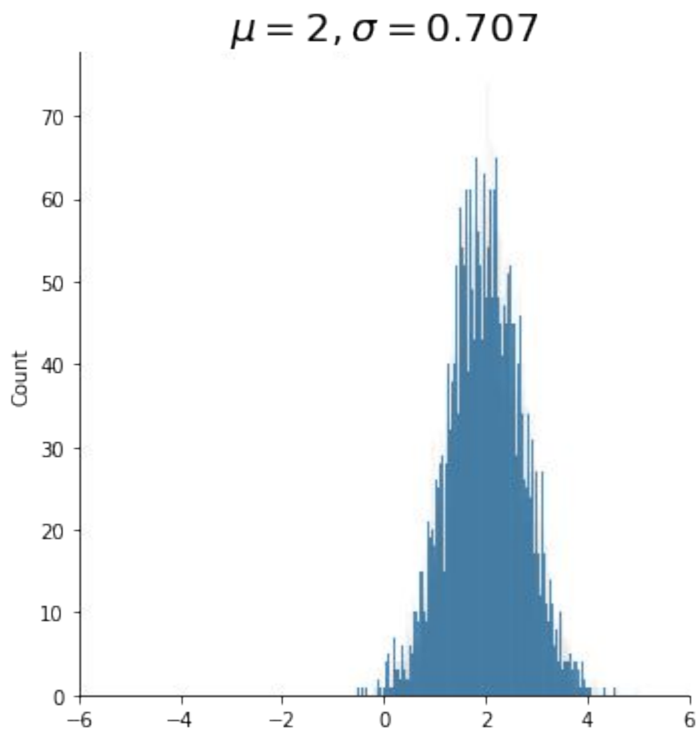
# Plots of Normal distribution for value: $\mu$=2 , **σ**= 0.707 (1)

- Graph of PDF




- Theoretical mean = 2.00
- Theoretical variance = 0.499

- Graph of values sampled from distribution(10,000 samples)




- Sample mean = 2.01
- Sample variance = 0.5

# Verifying Central Limit Theorem for Normal Distribution

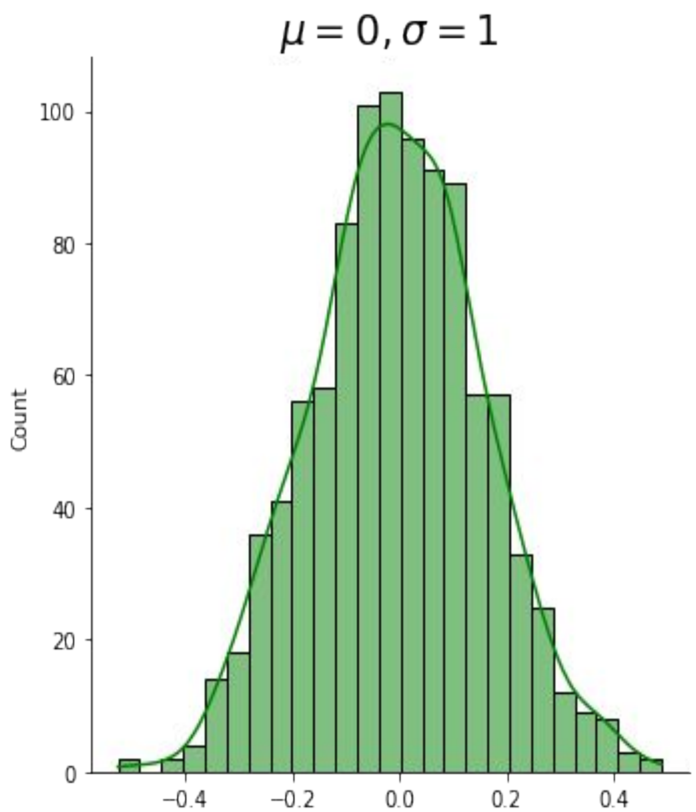
- Parameter:
  - $\mu$ = 0,$\boldsymbol{\sigma}$ =1
  - **n** = 40 (each sample is average of 40 samples)
  - No of samples = 1000

- Mean of Normal Distribution: $\mu$ = 0.0

- Variance of Normal Distribution: $\boldsymbol{\sigma}^2$ = 1.0

- If Central limit theorem holds true then for then **mean and Variance** of the expectation of static sample should be approximately
  - Mean : $\mu = \mu_{distribution}$ = 0.00
  - Variance : $\sigma^2 = \frac{\sigma^2_{distribution}}{n}$ = 0.025

# Plot of Sample Mean Statistic




- Theoretical mean : 0.0
- Theoretical variance: 0.025

- Expectation of sample mean : 0.0
- Variance of sample mean : 0.023

- As you can see from the above values that mean and variance of sample mean is approximately equal to theoretical ones.

- **Hence, Central Limit Theorem holds for the Normal Distribution**

# Exponential Distribution

- The PDF(Probability Density Function) of Exponential distribution  is as follows:

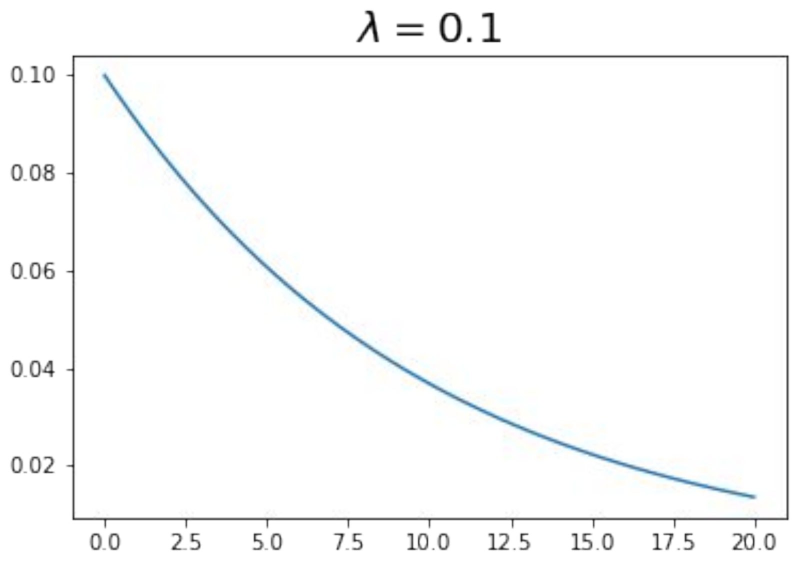
$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x} \; x \geq 0 \\ 0 \quad x < 0 \end{cases}$$

- Here $\lambda$ is the parameter. The mean and variance of the distribution depends on these parameters

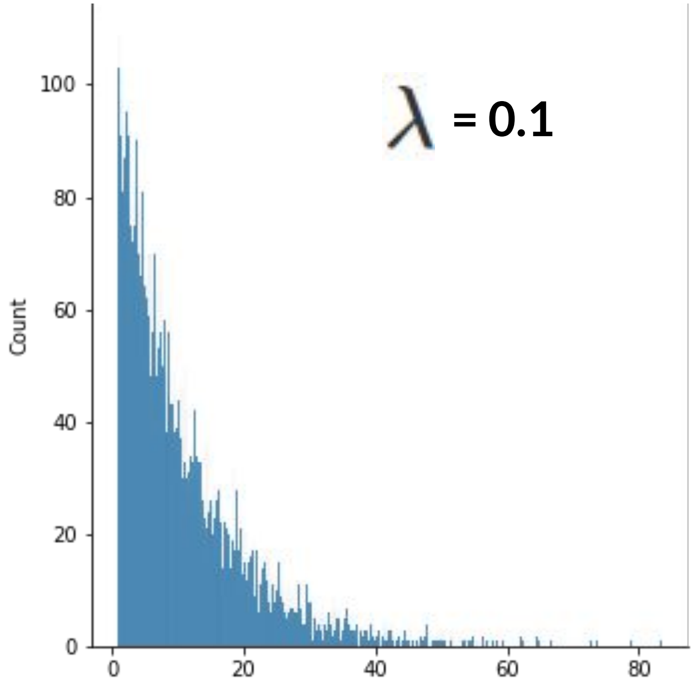
- The formula for calculating mean and variance of the distributions are  as follows:
  - Mean:    $1/\lambda$
  - Variance:  $1/\lambda^2$

# Plots of Exponential distribution for value: $\lambda$= 0.1

- Graph of PDF




- Graph of values sampled from distribution(10,000 samples)

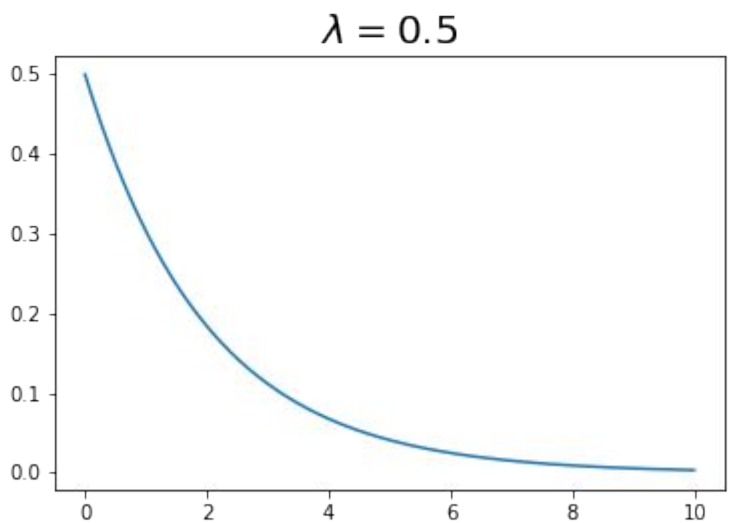



- Theoretical mean = 10.0
- Theoretical variance = 99.99

- Sample mean = 11.14
- Sample variance = 99.67

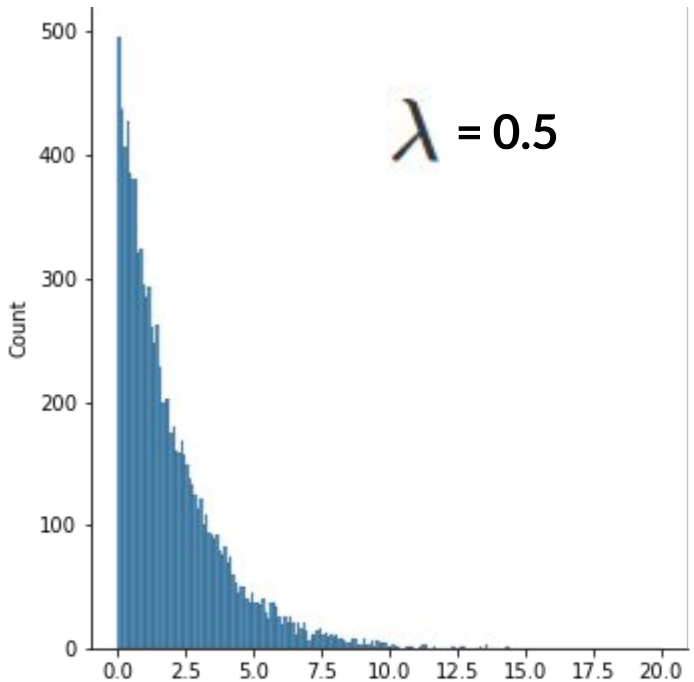
# Plots of Exponential distribution for value: $\lambda = 0.5$

- Graph of PDF




- Theoretical mean = 2.0
- Theoretical variance = 4.0

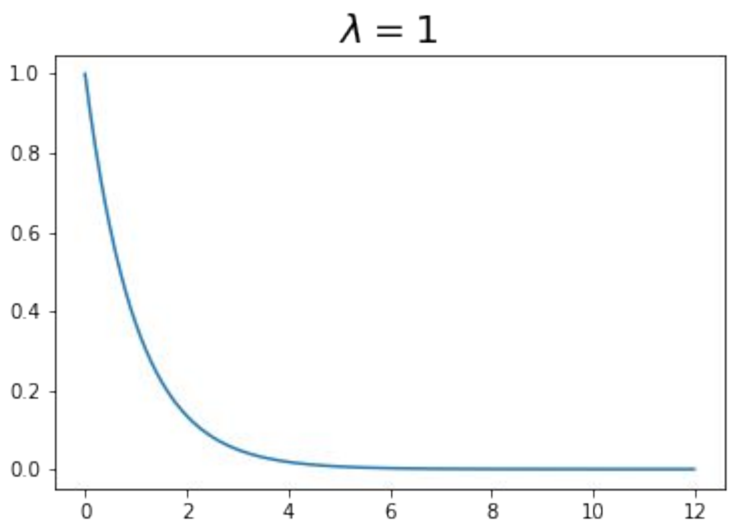
- Graph of values sampled from distribution(10,000 samples)



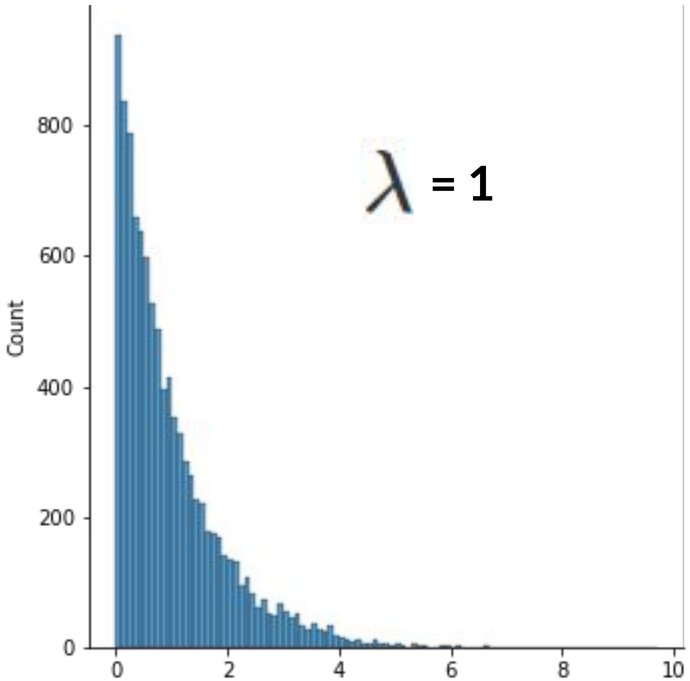

- Sample mean = 2.02
- Sample variance = 3.96

# Plots of Exponential distribution for value: $\lambda$= 1

- Graph of PDF




- Graph of values sampled from distribution(10,000 samples)



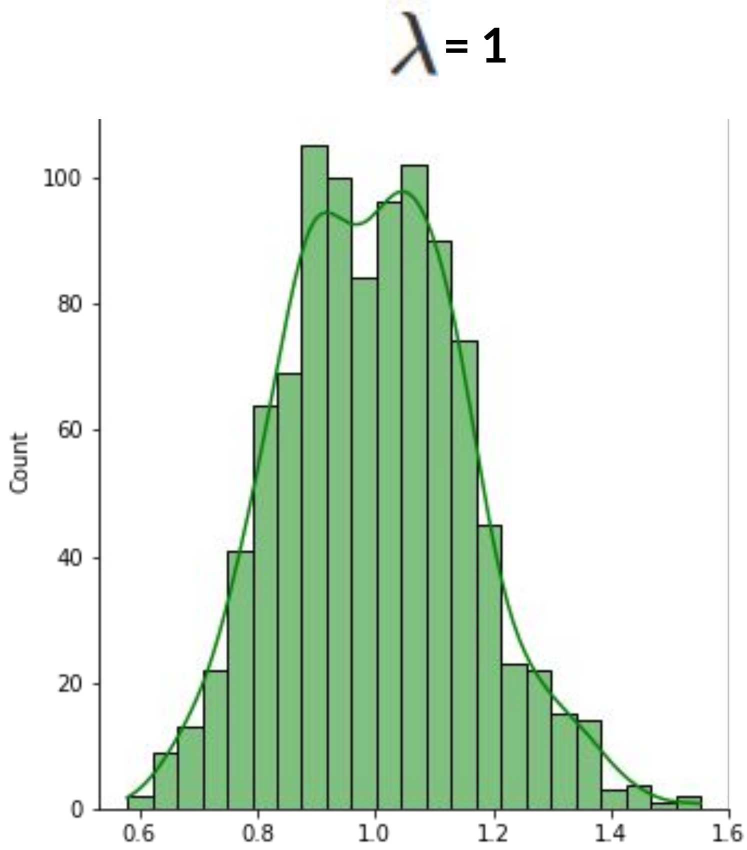

- Theoretical mean = 1.0
- Theoretical variance = 1.0

- Sample mean = 1.008
- Sample variance = 0.99

# Verifying Central Limit Theorem for Exponential

- Parameter:
  - $\lambda$ = 1
  - **n** = 40 (each sample is average of 40 samples)
  - No of samples = 1000
- Mean of Negative Binomial: $1/\lambda$ = 1.0
- Variance of Negative Binomial: $1/\lambda^2$ = 1.0
- If Central limit theorem holds true then for then **mean and Variance** of the expectation of static sample should be approximately
  - Mean : $\mu = \mu_{distribution}$ = 1.00
  - Variance : $\sigma^2 = \frac{\sigma^2_{distribution}}{n}$ = 0.025

# Plot of Sample Mean  Statistic




- Theoretical mean        :  1.0
- Theoretical variance:   0.025

- Expectation of sample mean   : 1.0
- Variance of sample mean          : 0.0248

- As you can see from the above values that mean and variance of sample mean is approximately equal to theoretical ones.

- **Hence, Central Limit Theorem holds for the Exponential Distribution**

# Beta Distribution

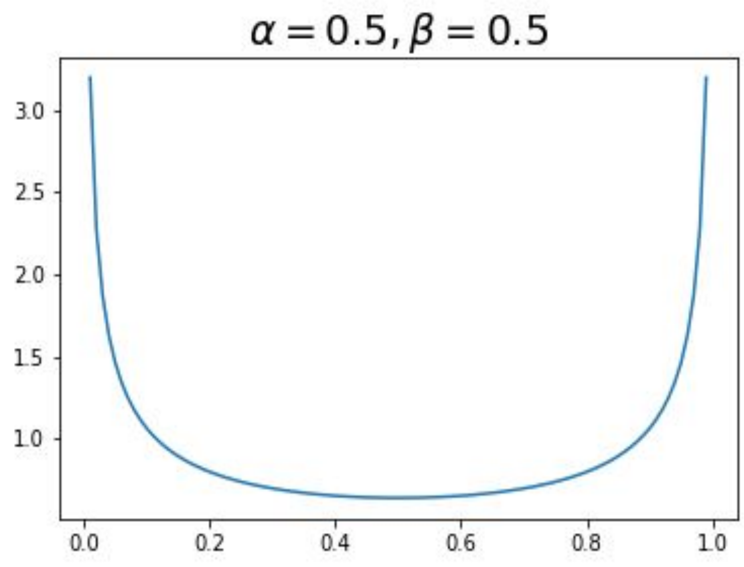
- The PDF(Probability Density Function) of Beta distribution f(x) is as follows:

$$f(x;\alpha,\beta) = \frac{\Gamma(\alpha+\beta)x^{(\alpha-1)}(1-x)^{(\beta-1)}}{\Gamma(\alpha)\Gamma(\beta)}$$

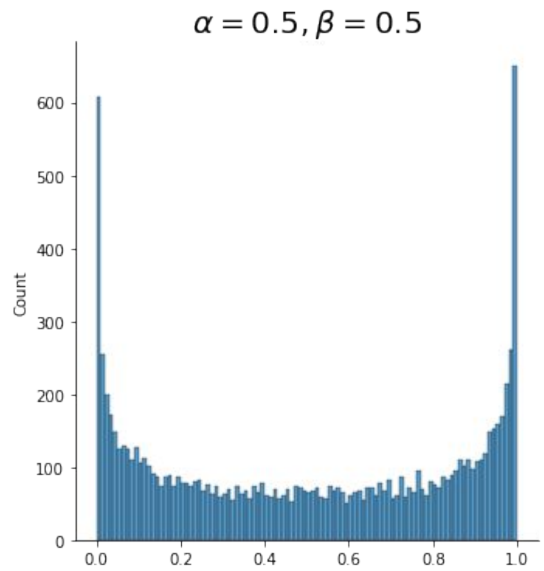
- Here ***α*** and **β** are shape parameters. The mean and variance of the distribution depends on these parameters
- The formula for calculating mean and variance of the distributions are as follows:
  - Mean: $\frac{\alpha}{\alpha+\beta}$
  - Variance: $\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$

# Plots of Beta distribution for various values of $\alpha$ and $\beta$ (1)

- Graph of PDF




- Graph of values sampled from distribution(10,000 samples)

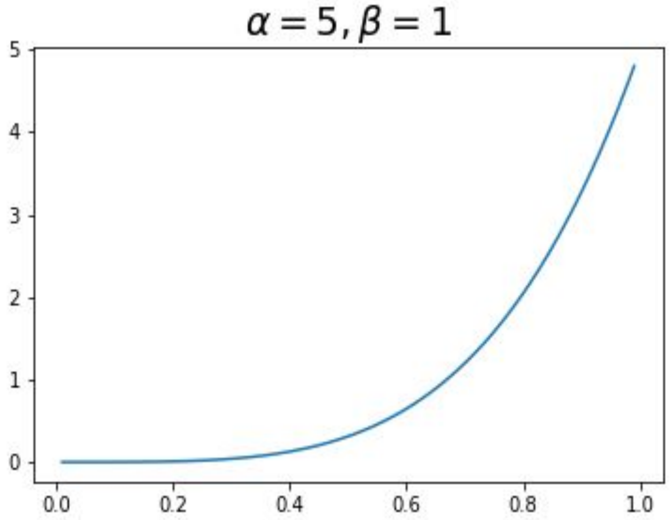



- Theoretical mean = 0.5
- Theoretical variance = 0.125

- Sample mean = 0.505
- Sample variance = 0.125

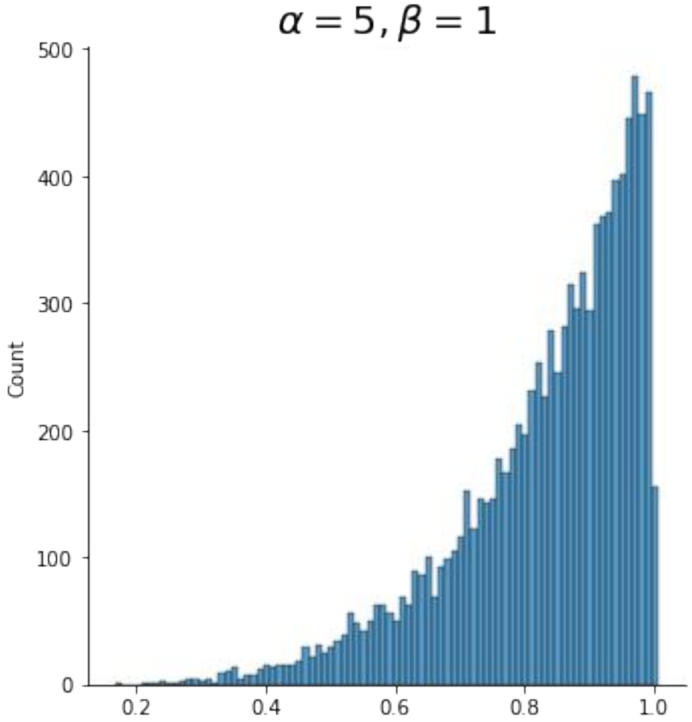
# Plots of Beta distribution for various values of $\alpha$ and $\beta$ (2)

- Graph of PDF




- Theoretical mean = 0.8334
- Theoretical variance = 0.0198

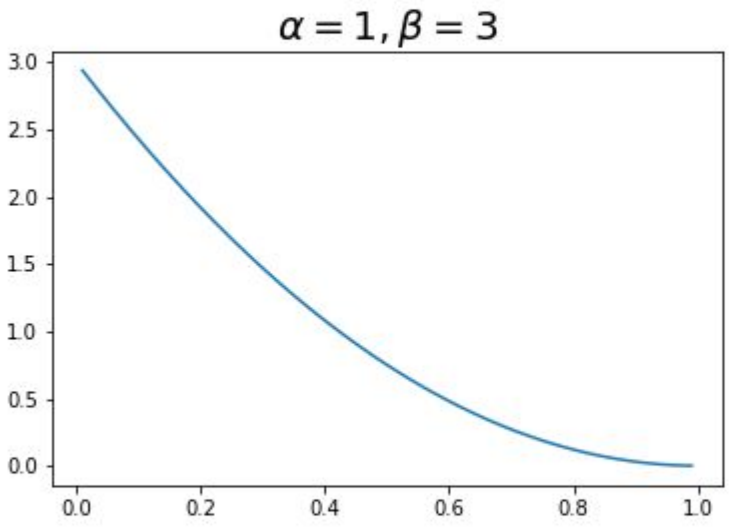
- Graph of values sampled from distribution(10,000 samples)



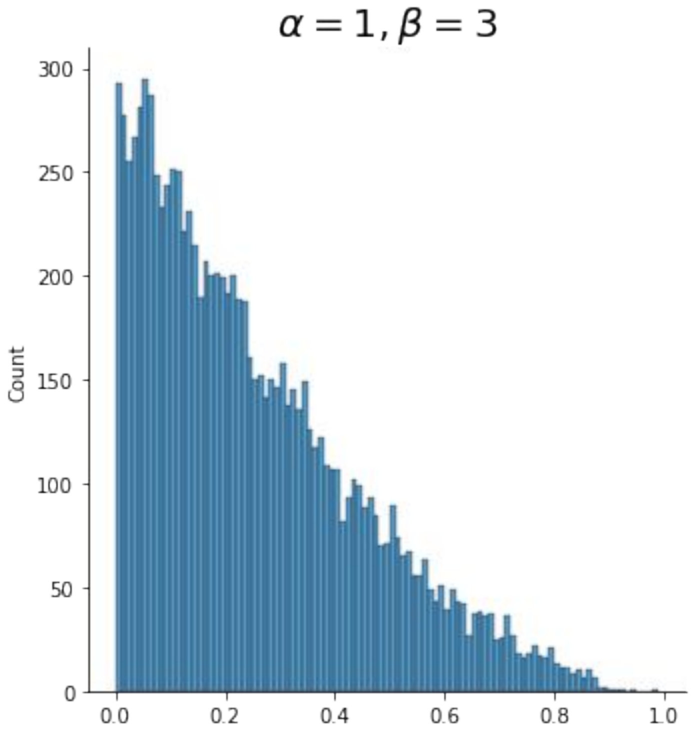

- Sample mean = 0.8338
- Sample variance = 0.0198

# Plots of Beta distribution for various values of $\alpha$ and $\beta$ (3)

- Graph of PDF




- Graph of values sampled from distribution(10,000 samples)



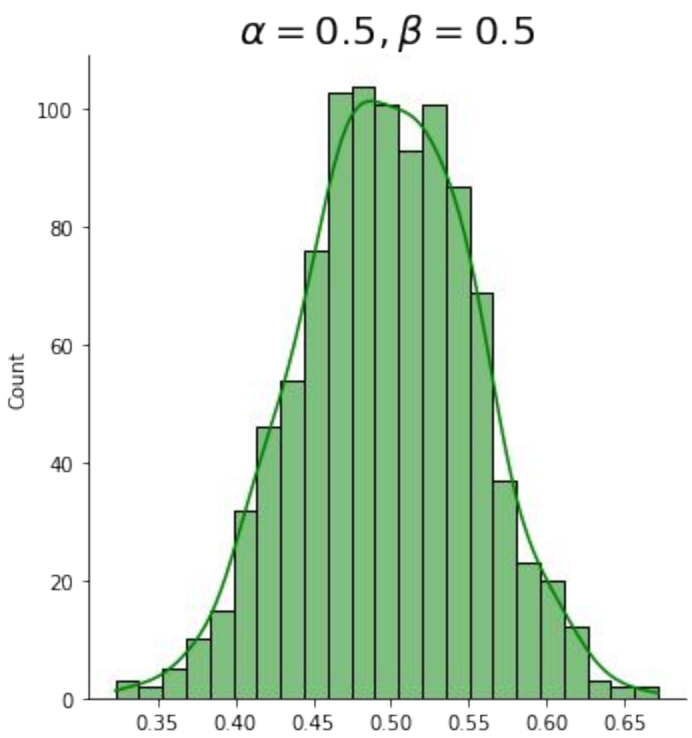

- Theoretical mean = 0.25
- Theoretical variance = 0.0375

- Sample mean = 0.2522
- Sample variance = 0.0381

# Verifying Central Limit Theorem for Beta Distribution

- Parameters:
    - $\boldsymbol{\alpha}$ = 0.5 , $\boldsymbol{\beta}$ =0.5
    - **n** = 40 (each sample is average of 40 samples)
    - No of samples = 1000
- Mean of Beta Distribution: $\frac{\alpha}{\alpha + \beta}$ = 0.5
- Variance of Beta Distribution: $\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$ = 0.125
- If Central limit theorem holds true then for then **mean and Variance** of the expectation of static sample should be approximately
    - Mean : $\mu = \mu_{distribution}$ = 0.5
    - Variance : $\sigma^2 = \frac{\sigma^2_{distribution}}{n}$ = 0.003125

# Plot of Sample Mean  Statistic




- Theoretical mean : 0.5
- Theoretical variance: 0.003125

- Expectation of sample mean:0 .5008
- Variance of sample mean: 0.00327

- As you can see from the above values that mean and variance of sample mean is approximately equal to theoretical ones.

- **Hence, Central Limit Theorem holds for the Beta Distribution**

# Gamma Distribution

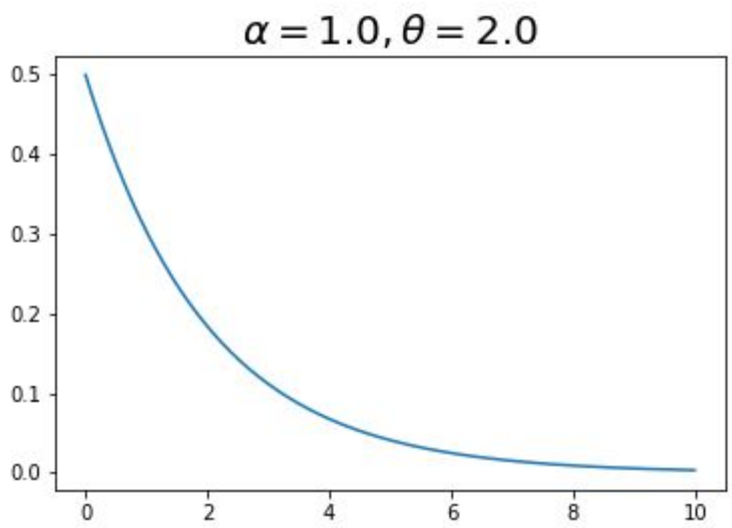
- The PDF(Probability Density Function) of Gamma distribution f(x) is as follows:

$$f(x;\alpha,\theta) = \frac{x^{\alpha-1}e^{\frac{-x}{\theta}}}{\theta^{\alpha}\Gamma(\alpha)}$$

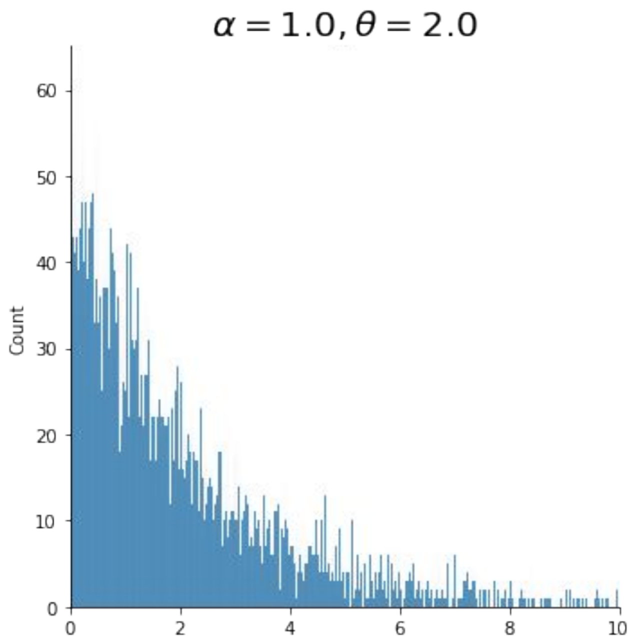
- Here $\alpha$ is shape parameter and $\theta$ is scale parameter. The mean and variance of the distribution depends on these parameters
- The formula for calculating mean and variance of the distributions are as follows:
  - Mean: $\alpha\theta$
  - Variance: $\alpha\theta^2$

# Plots of Gamma distribution for various values of $\alpha$ and $\theta$ (1)

- Graph of PDF




- Graph of values sampled from distribution(10,000 samples)

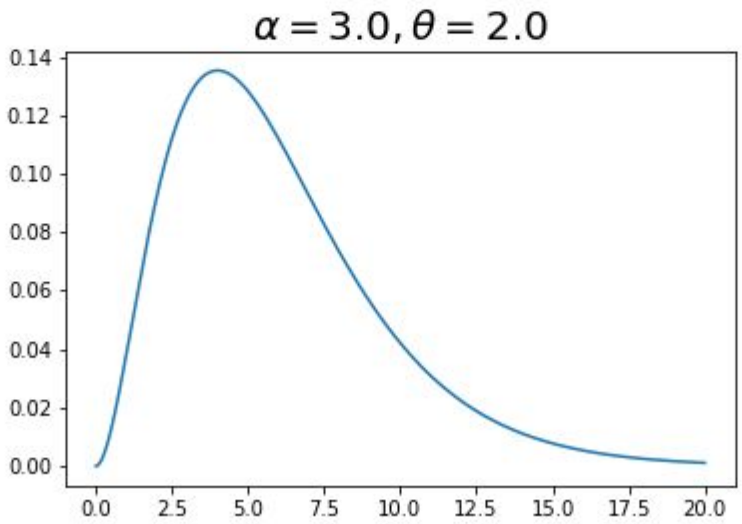



- Theoretical mean = 2.0
- Theoretical variance = 4.0

- Sample mean = 1.9894
- Sample variance = 3.9023

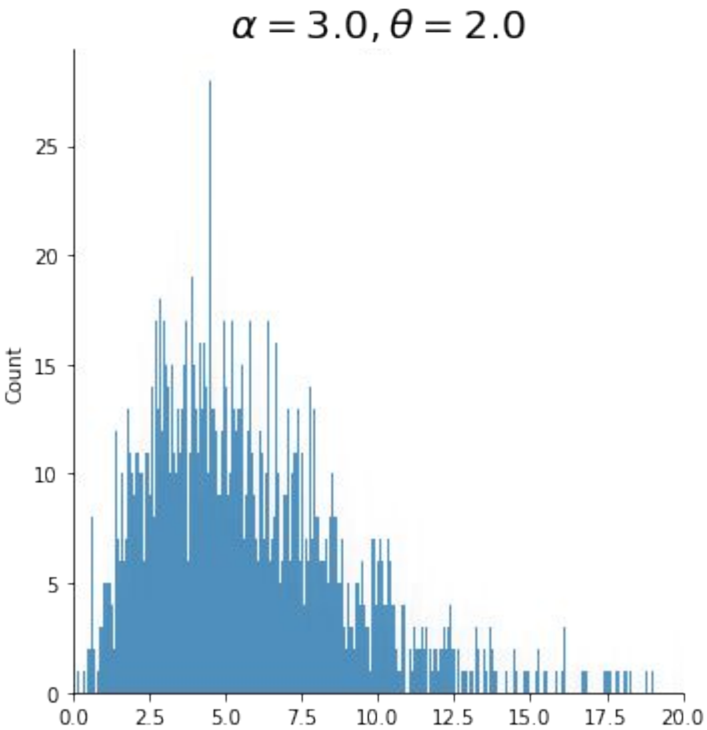
# Plots of Gamma distribution for various values of $\boldsymbol{\alpha}$ and $\boldsymbol{\theta}$ (2)

- Graph of PDF




- Theoretical mean = 6.0
- Theoretical variance = 12.0

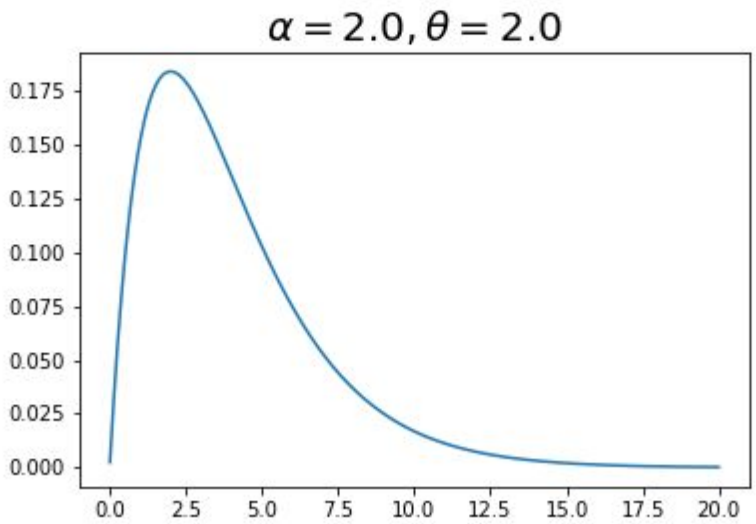
- Graph of values sampled from distribution(10,000 samples)



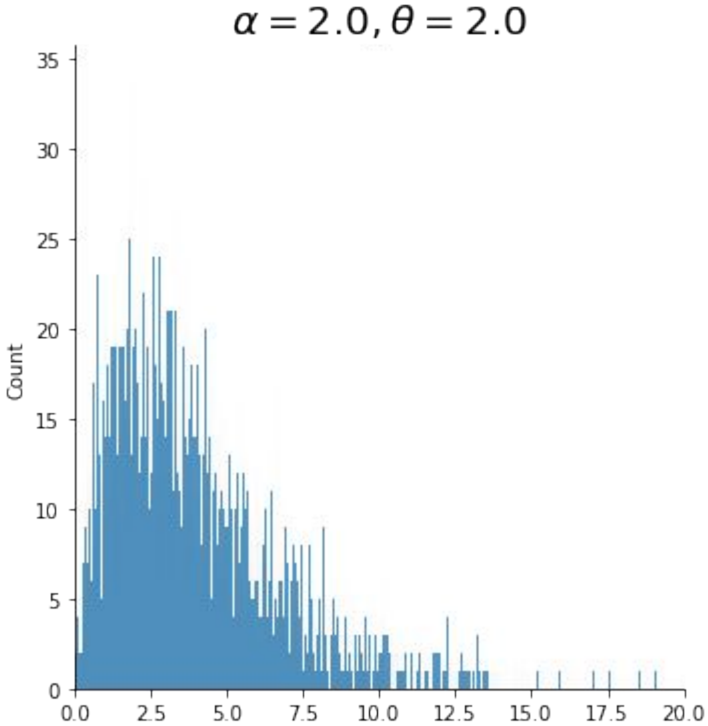

- Sample mean = 5.9831
- Sample variance = 12.2188

# Plots of Gamma distribution for various values of $\boldsymbol{\alpha}$ and $\boldsymbol{\theta}$ (3)

- Graph of PDF




- Graph of values sampled from distribution(10,000 samples)




- Theoretical mean = 4.0
- Theoretical variance = 8.0

- Sample mean = 4.0253
- Sample variance = 8.0941

# Verifying Central Limit Theorem for Gamma Distribution

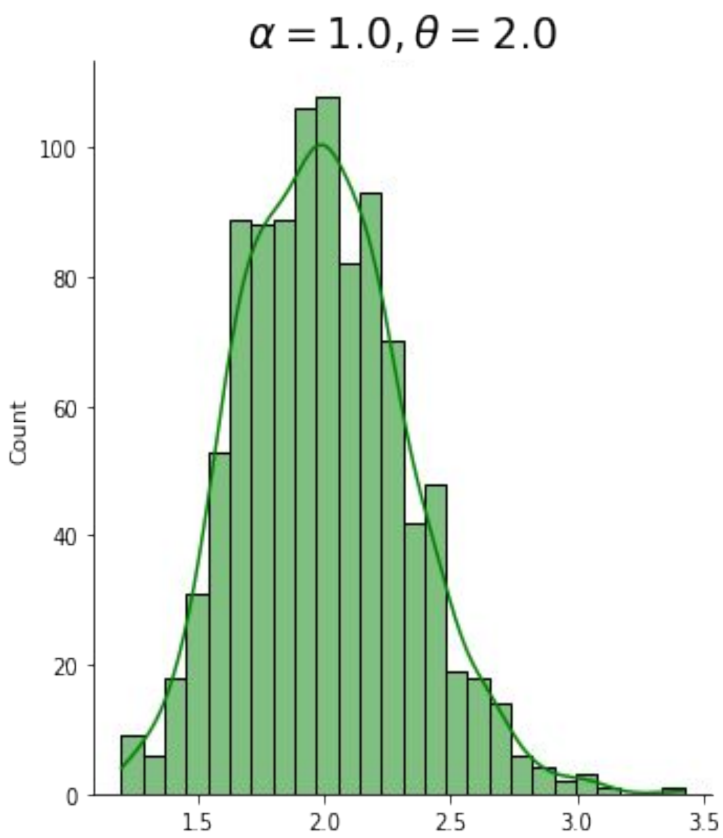
- Parameters:
  - $\alpha$ = 1.0 , $\theta$ = 2.0
  - **n** = 40 (each sample is average of 40 samples)
  - No of samples = 1000

- Mean of Gamma Distribution: $\alpha\theta$ = 2.0

- Variance of Gamma Distribution: $\alpha\theta^2$ = 4.0

- If Central limit theorem holds true then for then **mean and Variance** of the expectation of static sample should be approximately
  - Mean : $\mu = \mu_{distribution}$ = 2.0
  - Variance : $\sigma^2 = \dfrac{\sigma^2_{distribution}}{n}$ = 0.10

# Plot of Sample Mean Statistic




- Theoretical mean : 2.0
- Theoretical variance: 0.10

- Expectation of sample mean: 1.9937
- Variance of sample mean: 0.1049

- As you can see from the above values that mean and variance of sample mean is approximately equal to theoretical ones.

- **Hence, Central Limit Theorem holds for the Gamma Distribution**

# Lognormal Distribution

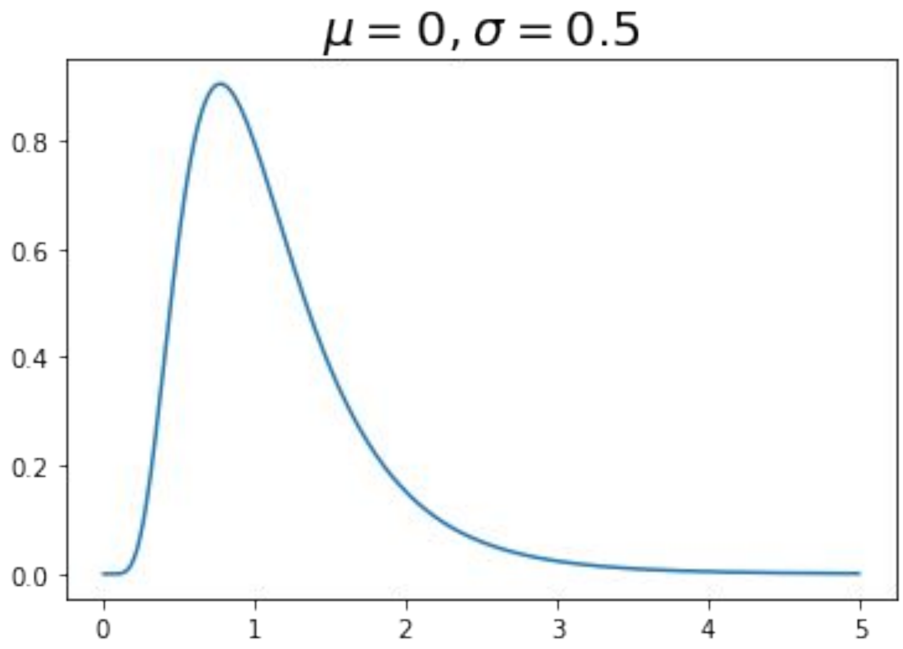
- The PDF(Probability Density Function) of Lognormal distribution f(x) is as follows:

$$f(x;\mu,\sigma) = \frac{1}{x}\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{\ln x - \mu^2}{2\sigma^2}\right)$$

- Here $\mu$ and $\sigma$ are parameters. The mean and variance of the distribution depends on these parameters
- The formula for calculating mean and variance of the distributions are as follows:
  - Mean: $\exp(\mu + \frac{\sigma^2}{2})$
  - Variance: $(\exp(\sigma^2) - 1)(\exp(2\mu + \sigma^2))$

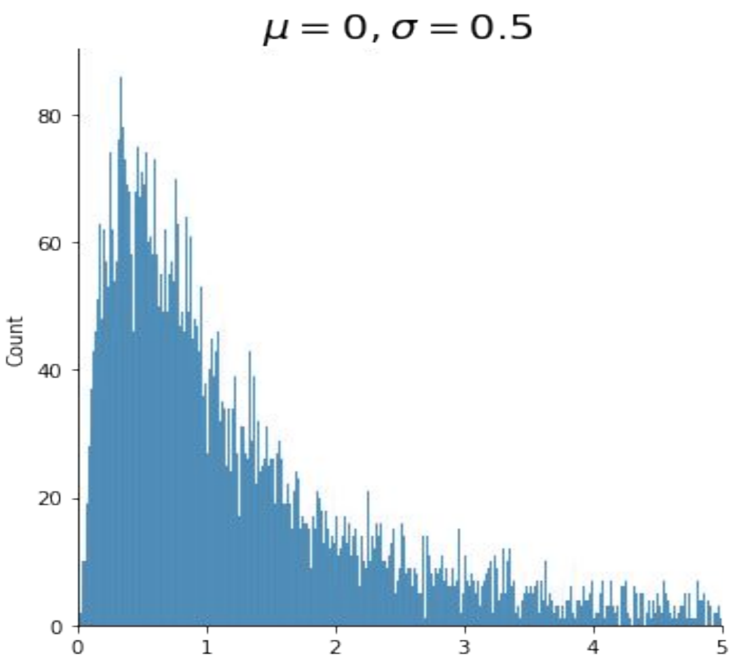
# Plots of Lognormal distribution for various values of $\mu$ and $\sigma$ (1)

- Graph of PDF




- Theoretical mean = 1.1331
- Theoretical variance = 0.3646

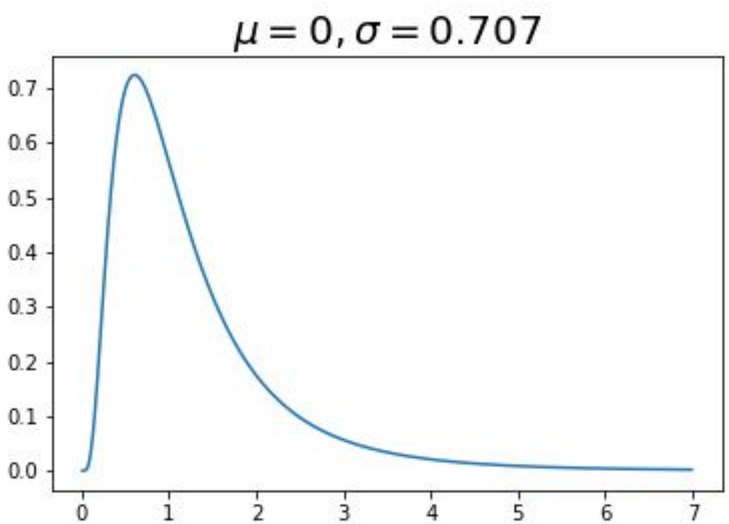
- Graph of values sampled from distribution(10,000 samples)



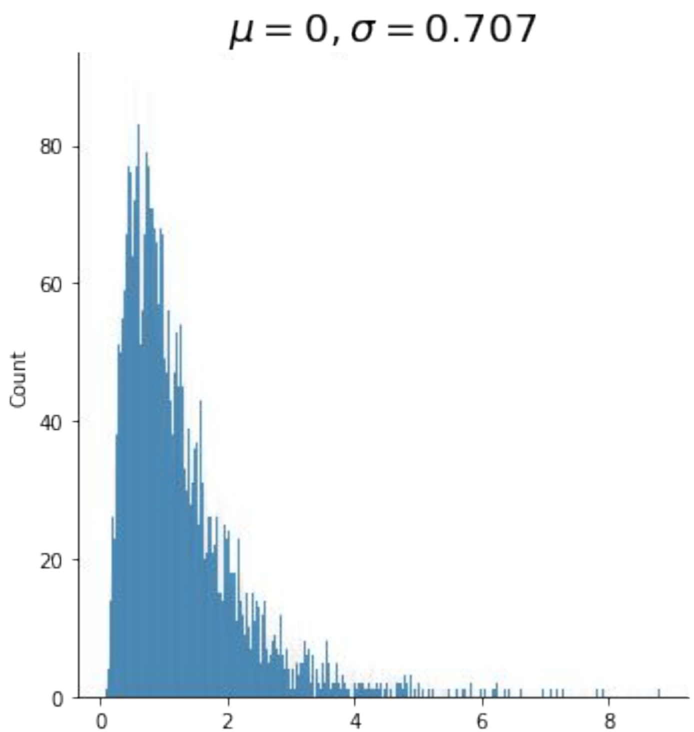

- Sample mean = 1.1296
- Sample variance = 0.3714

# Plots of Lognormal distribution for various values of $\mu$ and $\sigma$ (2)

- Graph of PDF




- Graph of values sampled from distribution(10,000 samples)

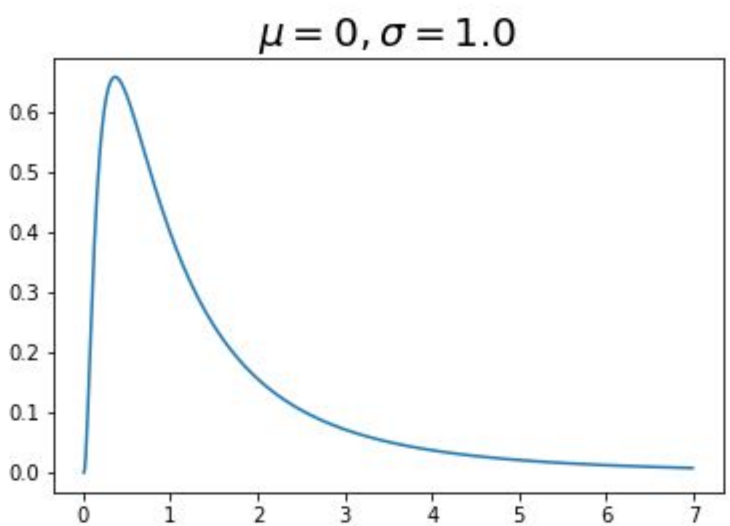



- Theoretical mean = 1.2840
- Theoretical variance = 1.069

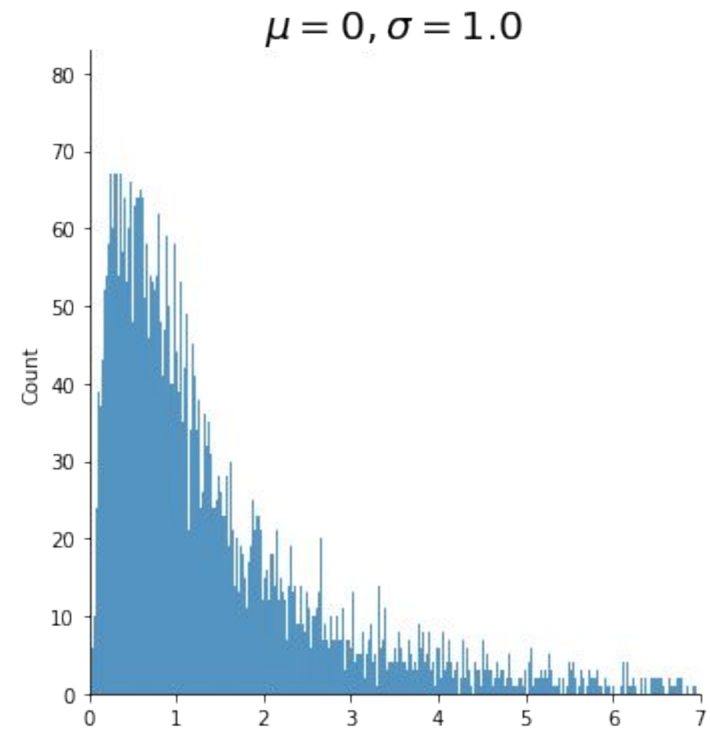
- Sample mean = 1.2863
- Sample variance = 1.0212

# Plots of Lognormal distribution for various values of $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ (3)

- Graph of PDF




- Graph of values sampled from distribution(10,000 samples)



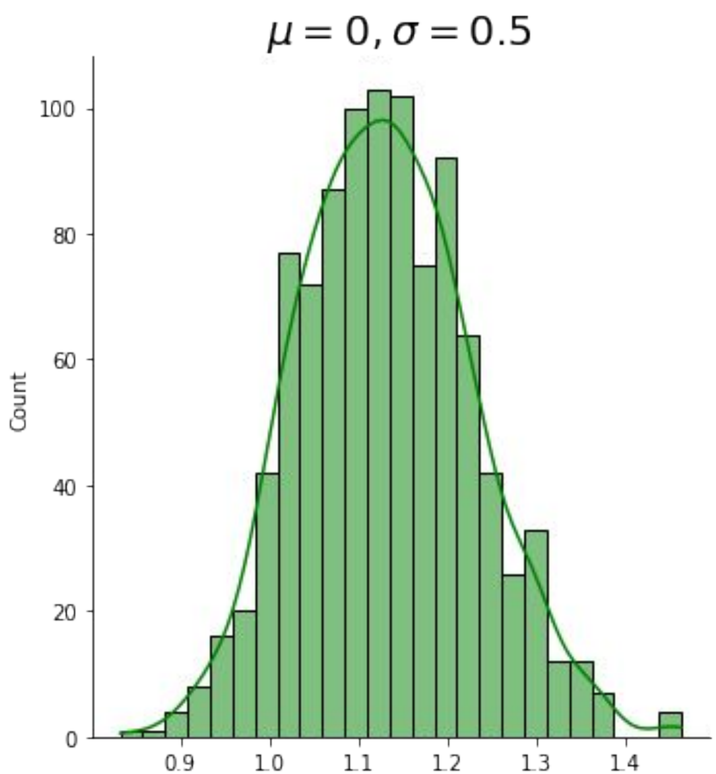

- Theoretical mean = 1.6487
- Theoretical variance = 4.6707

- Sample mean = 1.6358
- Sample variance = 4.3357

# Verifying Central Limit Theorem for Lognormal Distribution

- Parameters:
  - $\mu$ = 0 and $\sigma$=0.5
  - **n** = 40 (each sample is average of 40 samples)
  - No of samples = 1000
- Mean of Lognormal Distribution: $\exp(\mu + \frac{\sigma^2}{2})$ = 1.13
- Variance of Lognormal Distribution: $(\exp(\sigma^2) - 1)(\exp(2\mu + \sigma^2))$ = 0.36
- If Central limit theorem holds true then for then **mean and Variance** of the expectation of static sample should be approximately
  - Mean : $\mu = \mu_{distribution}$ = 1.1331
  - Variance : $\sigma^2 = \frac{\sigma^2_{distribution}}{n}$ = 0.0091

# Plot of Sample Mean Statistic




- Theoretical mean : 1.1331
- Theoretical variance: 0.0091

- Expectation of sample mean: 1.1306
- Variance of sample mean: 0.0085

- As you can see from the above values that mean and variance of sample mean is approximately equal to theoretical ones.

- **Hence, Central Limit Theorem holds for the Lognormal Distribution**

# Thank You