CE 605 Applied Statistics

Assignment Report

submitted by

Ajmal Babu M S
(Roll No. 194040007)

Under the supervision of

Prof. Subimal Ghosh

Department of Civil Engineering
INDIAN INSTITUTE OF TECHNOLOGY BOMBAY
October 8, 2019

Contents

Assignment 1

Function of Random Variable

1.1 Problem Statement

Show that the sum of two independent identically distributed exponential random variables is a gamma distribution.

Let X and Y be two independent identically distributed exponential random variables with $\lambda = 0.15$

Let Z = X + Y

Find the pdf of Z, and obtain its parameters by independently generating 1000 values for X and Y

1.2 Analytical Solution

$$X \sim Exp[\lambda] \qquad 0 \leq X \leq \infty$$

$$Y \sim Exp[\lambda] \qquad 0 \leq Y \leq \infty$$

$$Z = X + Y$$

where,

X and Y are independent

$$X = Z - Y \qquad\qquad 0 \leq X \leq Z$$

The distribution function for Z,

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} f_X(x)\ f_Y(z-x)\ dx \\ &= \int_0^z \lambda\ e^{-\lambda\ x}\ \lambda\ e^{-\lambda\ (z-x)}\ dx \\ &= \lambda^2\ z\ e^{-\lambda z} \end{aligned}$$

which is a gamma distribution

$$\begin{aligned} f_Z(z) &= \frac{\nu(\nu z)^{a-1}}{\Gamma(a)}\ e^{-\nu z} && z \geq 0 \\ &= 0 && z < 0 \end{aligned}$$

with $a = 2$ and $\nu = \lambda$

1.3 Numerical Solution

The problem is solved numerically by independently generating 1000 exponential random variables for X and Y with $\lambda = 0.15$ using Python-based *scipy.stats* library.

Obtained parameters for Z are a = 2.0114 and ν = 0.1482

The histogram, and PDF are plotted for X, Y and Z as below.

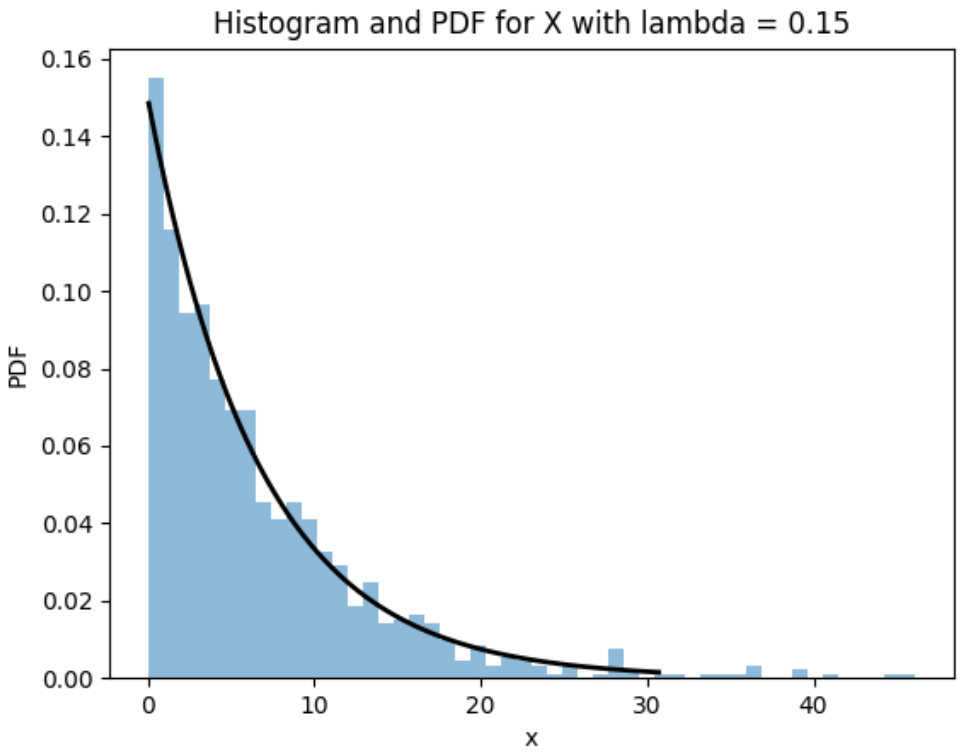

Figure 1.1: Histogram and PDF for X with $\lambda = 0.15$

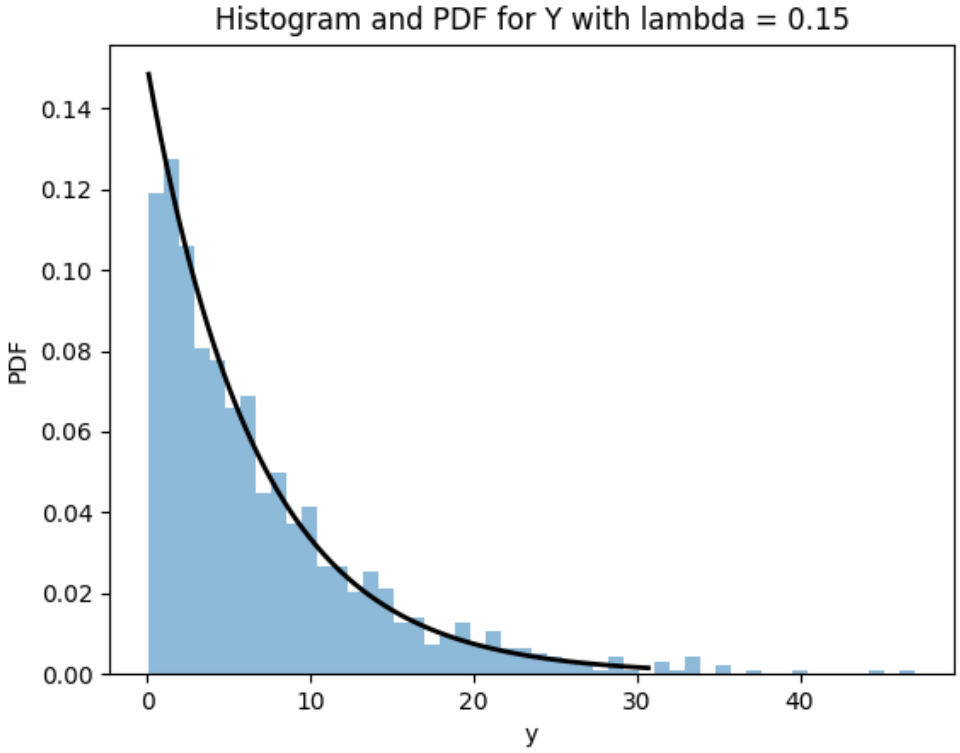

Figure 1.2: Histogram and PDF for Y with $\lambda = 0.15$

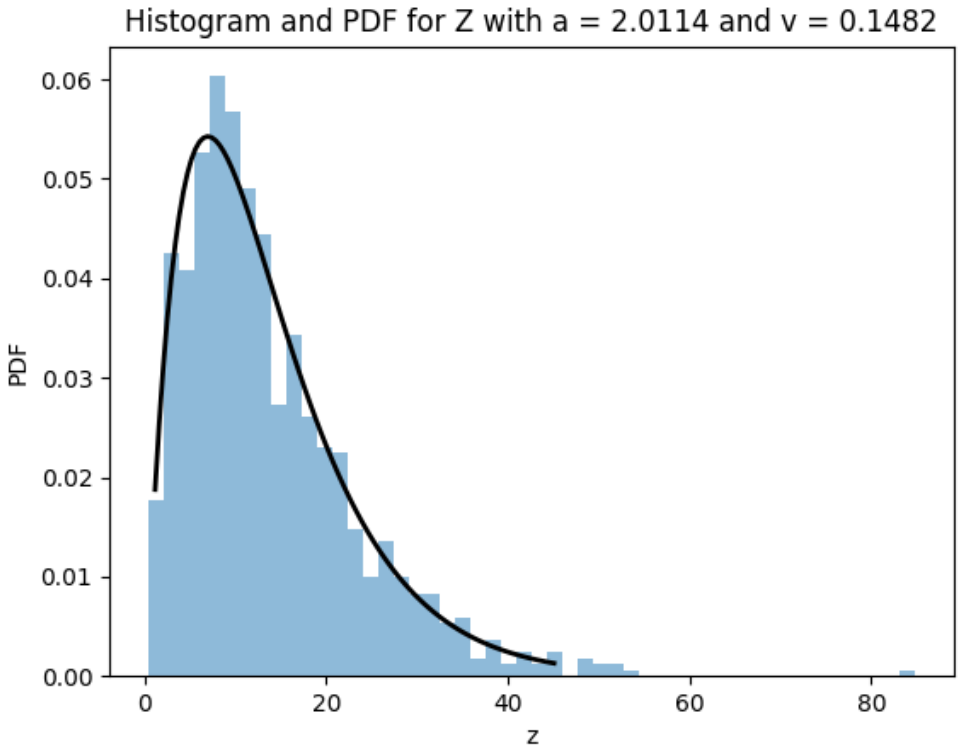

Figure 1.3: Histogram and PDF for Z with a = 2.0114 and $\nu = 0.1482$

The *Python* code used for this computation is given in 1.3 Appendix

The source codes are made available in github.com/ajmalbabums/ce605assignments

Appendix

Assignment 1

The following python code is made for verifying the sum of two independent identically distributed exponential random variables results a gamma distribution. The input values are the parameter for exponential distribution(λ) and number of random numbers to be generated for the computation. In the current code, λ is taken as 0.15 and 1000 random variables are generated.

```
"""
Assignment 1
Let X and Y be two independent identically distributed exponential random variables with
\lambda = 0.15
Let Z = X + Y
Find the pdf of Z, and obtain its parameters by independently generating 1000 values for X and Y
"""

import numpy as np
from scipy.stats import expon, gamma
import matplotlib.pyplot as plt

# Inputs
exp_lambda = 0.15
size = 1000


# Generating random variables for X and Y and obtaining Z
x = expon.rvs(scale=1/exp_lambda, size=size)
y = expon.rvs(scale=1/exp_lambda, size=size)
z = x + y


# Defining Subplots
fig1, ax1 = plt.subplots(1, 1)
plt.title("Histogram and PDF for X with lambda = {0}".format(exp_lambda))
```

```
plt.xlabel("x")
plt.ylabel("PDF")

fig2, ax2 = plt.subplots(1, 1)
plt.title("Histogram and PDF for Y with lambda = {0}".format(exp_lambda))
plt.xlabel("y")
plt.ylabel("PDF")

fig3, ax3 = plt.subplots(1, 1)
plt.title("Histogram and PDF for Z")
plt.xlabel("z")
plt.ylabel("PDF")

# Plotting histograms for X, Y and Z
ax1.hist(x, bins=50, density=True, histtype='stepfilled', alpha=0.5)
ax2.hist(y, bins=50, density=True, histtype='stepfilled', alpha=0.5)
ax3.hist(z, bins=50, density=True, histtype='stepfilled', alpha=0.5)


# Finding parameters for gamma distribution, Z
fit_gamma_a, fit_loc, fit_beta = gamma.fit(z)
g_mean, g_var, g_skew, g_kurt = gamma.stats(fit_gamma_a, moments='mvsk')

# Plotting PDF over histograms for X, Y and Z
var1 = np.linspace(expon.ppf(0.01, scale=1/exp_lambda), expon.ppf(0.99, scale=1/exp_lambda), 100)
var2 = np.linspace(gamma.ppf(0.01, a=fit_gamma_a, loc=fit_loc, scale=fit_beta),
                   gamma.ppf(0.99, a=fit_gamma_a, loc=fit_loc, scale=fit_beta), 100)

rv1 = expon(scale=1/exp_lambda)
rv2 = gamma(a=fit_gamma_a, loc=fit_loc, scale=fit_beta)
ax1.plot(var1, rv1.pdf(var1), 'k-', lw=2, label='pdf')
ax2.plot(var1, rv1.pdf(var1), 'k-', lw=2, label='pdf')
ax3.plot(var2, rv2.pdf(var2), 'k-', lw=2, label='pdf')

fit_gamma_a = round(fit_gamma_a, 4)
gamma_v = round(1/fit_beta, 4)
print("Obtained parameters for Z, a = {0} and v = {1}".format(fit_gamma_a, gamma_v))
plt.title("Histogram and PDF for Z with a = {0} and v = {1}".format(fit_gamma_a, gamma_v))

plt.show()
```