NUEN 647
Uncertainty Quantification for Nuclear Engineering
Assignment 1

Due on Tuesday, October 4, 2016

Dr. McClarren

Paul Mendoza

Contents

Problem 1 **3**

Problem 2 **5**

Problem 3 **7**

Problem 4 **8**

Problem 5 **10**

Complete the exercises in the Chapter 2 notes. Be sure to include discussion of results where appropriate. You may use any tools that are approrpriate to solving the problem.

Problem 1

Show that the transformation in equation 1 results in a standard normal random variable by computing the mean and variance of z.

$$z = \frac{x-\mu}{\sigma} \tag{1}$$

An important special case of the expectation value is the mean which is the expected value of x. It is often denoted as μ,

$$\mu = E[x] = \int_{-\infty}^{\infty} x f(x) dx$$

where x is a realization of a random sample and $f(x)$ is the probability density function (PDF) for the random variable. For a normal distribution,

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

For the sake of the transformation, the value of z substitutes for x, the realization of a random sample (not the PDF because we are transforming that distribution). Therefore, the mean for z is:

$$\mu_z = \int_{-\infty}^{\infty} \frac{x-\mu}{\sigma} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} dx$$

If $u = (x-\mu)^2$ and $\frac{du}{2} = (x-\mu)dx$ (note that the limits change from $(-\infty, \infty)$ to (∞, ∞) - but that seems fishy to me so I will change it back after integration).

$$\mu_z = \int_{\infty}^{\infty} \frac{1}{2\sigma^2\sqrt{2\pi}} e^{\frac{-u}{2\sigma^2}} du = \left| \frac{-1}{\sqrt{2\pi}} e^{\frac{-u}{2\sigma^2}} \right|_{\infty}^{\infty}$$

$$\mu_z = \left| \frac{-1}{\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \right|_{-\infty}^{\infty} = \frac{-1}{\sqrt{2\pi}} (e^{-\infty} - e^{-\infty}) = \boxed{0}$$

The variance is defined as:

$$Var(X) = E[(X-\mu_X)^2]$$

Substituting Eq. 1 for X, (but not for the pdf)

$$Var(Z) = E[(\frac{x-\mu_X}{\sigma_X} - \mu_Z)^2] = E\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2\right] = \frac{1}{\sigma_X^2}(E[x^2] - 2\mu_X E[x] + \mu^2 E[1])$$

Noting that above it was proven that $E[x] = \mu_X$ and given that the definition of $E[1] = 1$ and assuming that $E[x^2] = \sigma_X^2 + \mu_X^2$ (will solve on next page)

$$\frac{1}{\sigma_X^2}(\sigma_X^2 + \mu_X^2 - 2\mu_X^2 + \mu_X^2) = \boxed{1}$$

$$E[x^2] = \int_{-\infty}^{\infty} \frac{x^2}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

If $t = \frac{(x-\mu)}{\sqrt{2}\sigma}$ and $\sqrt{2}\sigma dt = dx$ and $x = t\sqrt{2}\sigma + \mu$ then (limits of integration don't change)

$$E[x^2] = \int_{-\infty}^{\infty} \frac{\left(t\sqrt{2}\sigma + \mu\right)^2}{\sqrt{\pi}} e^{-t^2} dt = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \left(2\sigma^2 \left(t^2 e^{-t^2}\right) + 2\sqrt{2}\sigma\mu \left(te^{-t^2}\right) + \mu^2 \left(e^{-t^2}\right)\right)$$

According to wolfram alpha

$$\int_{-\infty}^{\infty} t^2 e^{-t^2} = \frac{\sqrt{\pi}}{2}$$

$$\int_{-\infty}^{\infty} te^{-t^2} = 0$$

$$\int_{-\infty}^{\infty} e^{-t^2} = \sqrt{\pi}$$

Which simplifies the above to $\sigma^2 + \mu^2$.

Problem 2

Consider the random variables $X \sim U(-1,1)$ and $Y \sim X^2$. Are these independent random variables? What is their covariance?

Marginal and Joint PDFs

The PDF for X is:

$$f_X(x) = \frac{1}{(1-(-1))} = 0.5 \quad x \in [-1,1]$$

The PDF for Y is: ^{link}

$$f_Y(y) = \frac{1}{2\sqrt{y}} \quad y \in [0,1]$$

Without using the handy reference, this may be derived from the joint PDF, $f(x,y)$, defined as (between McClarrens Eq. 2.30 and 2.31):

$$f(x,y) = f(y|x)f_X(x)$$

From the definition of Y, $f(y|x)$ is 0 except when $y = x^2$. I think this would be.

$$f(y|x) = \delta(y - x^2)$$

Which means,

$$f(x,y) = 0.5\delta(y - x^2)$$

To calculate the PDF for Y (f(y)), we need to integrate over all other variables (in this case, X).

$$f(y) = \int_{-1}^{1} f(x,y)dx = \int_{-1}^{1} 0.5\delta(y - x^2)dx$$

Wolfram alpha tells me the answer is,

$$f_Y(y) = \frac{1}{2\sqrt{y}} \quad y \in [0,1]$$

The same as above.

Independance

If two random variables, X and Y, are independent, they satisfy the following condition: ^{link}

- $P(Y|X) = P(Y)$, for all values of X and Y.

Because $P(Y|X) = \delta(y - xY2) \neq P(Y) = \frac{1}{2\sqrt{y}}$ (at least not for ALL values of x and y), these two variables are $\boxed{dependent}$.

Covariance

The covariance for two random variables is defined as:

$$\sigma_{XY} = E[(x-\mu_X)(y-\mu_Y)]$$

This simplifies down to:

$$\sigma_{XY} = E(XY) - \mu_X\mu_Y = \int_{-1}^{1} dx \int_{0}^{1} dy \;\; xyf(x,y) - \mu_X\mu_Y$$

Because $\mu_X = 0$ this reduces to

$$\begin{aligned}
\sigma_{XY} = E(XY) &= \int_{-1}^{1} dx \int_{0}^{1} dy \;\; xyf(x,y) \\
&= \int_{-1}^{1} dx \int_{0}^{1} dy \;\; xy0.5\delta(y-x^2) \\
&= \int_{-1}^{1} dx \;\; 0.5x^3 dx \\
&= \boxed{0}
\end{aligned}$$

Wolfram alpha gave the step between the second and third line. These variables are dependant, but have a zero covariance.

Problem 3

Show that a general covariance matrix must be positive definite, i.e. $\vec{x}^T\Sigma\vec{x} > 0$ for any vector $\vec{x}$ that is not all zeros.

Given that $\vec{Y}$ is a vector of random variables and $\vec{\mu}_Y$ is a vector of the mean values for the random variables found in $\vec{Y}$.

$$\begin{aligned}\vec{x}^T\Sigma\vec{x} &= \vec{x}^T E[(\vec{Y}-\vec{\mu}_Y)(\vec{Y}-\vec{\mu}_Y)^T]\vec{x}\\ &= E[\vec{x}^T(\vec{Y}-\vec{\mu}_Y)(\vec{Y}-\vec{\mu}_Y)^T\vec{x}]\end{aligned}$$

The last step above puts a constant inside the expectation value integral. Notice

$$\vec{x}^T(\vec{Y}-\vec{\mu}_Y) = (\vec{Y}-\vec{\mu}_Y)^T\vec{x}$$

and that both are scaler functions of the random variables. Therefore,

$$\begin{aligned}\vec{x}^T\Sigma\vec{x} &= E[(\vec{x}^T(\vec{Y}-\vec{\mu}_Y))^2]\\ &= E[g(\vec{Y})^2] = \sigma_f^2\end{aligned}$$

The expectation value for a multivariate distribution is defined as

$$E[g(\vec{Y})] = \int_{-\infty}^{\infty} dy_1 \int_{-\infty}^{\infty} dy_2 \ ... \int_{-\infty}^{\infty} dy_p \ g(\vec{y})f(\vec{y})$$

Where $f(\vec{y})$ is the multivariate PDF for the random variables of $\vec{Y}$. If a number of samples is given, rather than functions that can be integrated, the expectation value for a multivariate distribution is defined as:

$$E[g(\vec{Y})] = SC$$

To prove that the covariance matrix is positive definite the above integral must be proved to be positive with $g(x) = (\vec{x}^T(\vec{Y}-\vec{\mu}_Y))^2$. Explicitly,

$$\begin{aligned}E[g(Y)] &= \int_{-\infty}^{\infty} dy_1 \int_{-\infty}^{\infty} dy_2 \ ... \int_{-\infty}^{\infty} dy_p \ (\vec{x}^T(\vec{Y}-\vec{\mu}_Y))^2 f(y)\\ &= \int_{-\infty}^{\infty} dy_1 \int_{-\infty}^{\infty} dy_2 \ ... \int_{-\infty}^{\infty} dy_p \ ((y_1-\mu_1)x_1 + (y_2-\mu_2)x_2 + \ ... \ + (y_p-\mu_p)x_p)^2 f(y)\end{aligned}$$

Problem 4

Use rejection sampling to sample from a Gamma random variable $X \sim \mathscr{G}(\alpha, \beta)$ where

$$f(x) = \frac{\theta^{\alpha-1}e^{-\theta\beta}}{\Gamma(\alpha)\beta^{-\alpha}} \quad \alpha, \beta > 0$$

Let $\alpha = 1$ and $\beta = 0.5$. From rejection sampling with a $N = 10^4$, compute a rejection rate for the sampling procedure. Now draw a triangle around the function and do rejection sampling. Compare the rejection rate from the triangle versus the rectangle. You may consider that the PDF is zero if $f(x) < 10^{-6}$.

Python script for rejection sampling.

Listing 1: Python Script for problem

```
#!/usr/bin/env python3

###############################################################
#################### Import packages ########################
###############################################################

import time
start_time = time.time()
import scipy.special as sps
import numpy as np
import matplotlib.pyplot as plt
import matplotlib
import random as rn
import Functions as fun
import copy

###############################################################
####################### Calculations ########################
###############################################################

#Values go to 10^-6 around 26.245
N=100;alpha=1;beta=0.5;a=0;b=26.245;h=0.5;Nsamples=10**4

theta=np.linspace(a,b,N)
f_x=fun.GammaPDF(alpha,beta,theta)

(fig,ax)=fun.Plot(theta,f_x)

OutsideSquare=0;OutsideTri=0
for i in range(0,Nsamples):
    X=rn.uniform(a,b);Xt=copy.copy(X)
    Y=rn.uniform(0,h);Yt=copy.copy(Y)

    if Y>(-h/b)*X+h: #Triangular
        Xt=b-Xt
        Yt=h-Yt

    H=fun.GammaPDF(alpha,beta,X) #Square
    Ht=fun.GammaPDF(alpha,beta,Xt) #Triangle

```

```
    if(Y<H): #Square
        ax=fun.PlotaxIn(X,Y,i,ax,1)
    else:
        OutsideSquare=OutsideSquare+1
        ax=fun.PlotaxOut(X,Y,i,ax,1)

    if(Yt<Ht): #Triangular
        ax=fun.PlotaxIn(Xt,Yt,i,ax,2)
    else:
        OutsideTri=OutsideTri+1
        ax=fun.PlotaxOut(Xt,Yt,i,ax,2)

ax=fun.Plotlegend(ax,theta,f_x)

RejectionSq=OutsideSquare/Nsamples;RejectionTri=OutsideTri/Nsamples

print("Square = "+str(RejectionSq)+" Triangle = "+str(RejectionTri))
plt.savefig('P4F1.pdf')

##################### Time To execute ################
print("--- %s seconds ---" % (time.time() - start_time))
```

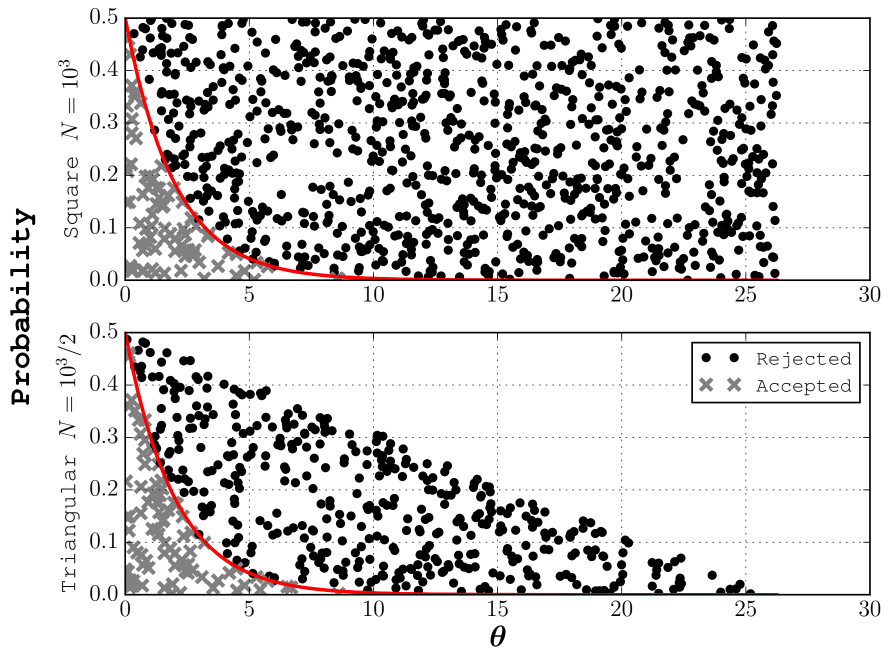

Figure 1: Square (top) and triangular (bottom) rejection sampling for the Gamma random variable.

The rejection rate for the square is $\boxed{92.65\%}$.
The rejection rate for the triangle is $\boxed{85.04\%}$.

The acceptance rate about doubled from the square to the triangle $((1 - 0.9265) * 2 = 0.147 \approx 0.1496 = (1 - 0.8504))$. This is what is expected because we cut the sampling area in half. This could also be used to verify the PDF is properly normalized. $0.1496 * 0.5 * 26.245 * 1/2 = 0.98$.

Problem 5

Consider a random variable, $X > 0$, that has it's logarithm distributed by a normal distribution with mean $\mu = 0$ and variance $\sigma^2 = 1$. Such a distribution is called a log-normal distribution. Compute this distribution's a) mean, b) variance, c) median, d) mode, e) skew, and d) kurtosis.

The PDF for the log-normal distribution, found on wikipedia, is:

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

For a standard log-normal distribution this is

$$f(X) = \frac{1}{x\sqrt{2\pi}} e^{-\frac{ln(x)^2}{2}}$$

Which simplifies to:

$$\begin{aligned} f(X) &= \frac{1}{x\sqrt{2\pi}} \left(e^{-ln(x)}\right)^{\frac{ln(x)}{2}} \\ &= \frac{1}{x\sqrt{2\pi}} \left(\frac{1}{x}\right)^{\frac{ln(x)}{2}} \end{aligned}$$

a) mean

The mean, μ, is defined as:

$$\mu = E[x] = \int_{-\infty}^{\infty} x f(x) dx$$

Using the log-normal distribution the mean is:

$$\begin{aligned} \mu = E[x] &= \int_0^{\infty} \frac{1}{\sqrt{2\pi}} \left(\frac{1}{x}\right)^{\frac{ln(x)}{2}} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_0^{\infty} \left(\frac{1}{x}\right)^{\frac{ln(x)}{2}} dx \end{aligned}$$

Wolfram alpha says

$$\int_0^{\infty} \left(\frac{1}{x}\right)^{\frac{ln(x)}{2}} dx = \sqrt{2e\pi}$$

Therefore the answer is $\boxed{\mu = \sqrt{e}}$

b) Variance

Variance, σ^2, is defined as:

$$\begin{aligned}\sigma^2 &= E[(X-\mu)^2] \\ &= E[X^2] - 2E[\mu X] + E[\mu^2] \\ &= E[X^2] - \mu^2 \\ &= \int_{-\infty}^{\infty} x^2 f(x)dx - \mu^2\end{aligned}$$

Using the log-normal distribution the variance is:

$$\begin{aligned}\sigma^2 = E[x^2] - \mu^2 = &\int_0^{\infty} \frac{x}{\sqrt{2\pi}} \left(\frac{1}{x}\right)^{\frac{ln(x)}{2}} dx - \mu^2 \\ &= \frac{1}{\sqrt{2\pi}} \int_0^{\infty} x \left(\frac{1}{x}\right)^{\frac{ln(x)}{2}} dx - \mu^2\end{aligned}$$

Wolfram alpha says

$$\int_0^{\infty} x \left(\frac{1}{x}\right)^{\frac{ln(x)}{2}} dx = e^2\sqrt{2\pi}$$

Therefore the answer is $\boxed{\sigma^2 = e(e-1)}$

c) Median

The median is defined as the point where the CDF is equal to one-half. The CDF, in terms of the PDF is:

$$F_X(x) = \int_{-\infty}^{x} f(x')dx'$$

For the log-normal distribution

$$\begin{aligned}F_X(x) &= \int_0^x \frac{1}{x'\sigma\sqrt{2\pi}} e^{-\frac{(ln(x')-\mu)^2}{2\sigma^2}} dx' \\ &= \frac{1}{2} + \frac{1}{2} erf\left[\frac{ln(x)-\mu}{\sqrt{2}\sigma}\right]\end{aligned}$$

If $\sigma = 1$ and $\mu = 0$,

$$F_X(x) = \frac{1}{2} + \frac{1}{2} erf\left[\frac{ln(x)}{\sqrt{2}}\right]$$

If $F_X(x) = \frac{1}{2}$ then

$$\begin{aligned}erf^{-1}[0]\sqrt{2} &= ln(x) \\ e^{erf^{-1}[0]\sqrt{2}} &= x\end{aligned}$$

$erf^{-1}[0] = 0$, because wolfram says so...

Therefore the answer is $\boxed{Median = 1}$

d) Mode

The mode is the point where the PDF takes its maximum value, the most likely value of the distribution. This could also be the point where its derivative goes to 0.

$$f(X) = \frac{1}{x\sqrt{2\pi}} e^{-\frac{ln(x)^2}{2}}$$

Derivatives:

$$e^{-\frac{ln(x)^2}{2}} \frac{d}{dx} = \frac{-ln(x)}{x} e^{-\frac{ln(x)^2}{2}}$$

$$\frac{1}{x\sqrt{2\pi}} \frac{d}{dx} = -\frac{1}{x^2\sqrt{2\pi}}$$

Combining:

$$f'(X) = \left(\frac{-ln(x) - 1}{x^2\sqrt{2\pi}}\right) e^{-\frac{ln(x)^2}{2}}$$

Setting Equal to Zero

$$0 = \left(\frac{-ln(x) - 1}{x^2\sqrt{2\pi}}\right) e^{-\frac{ln(x)^2}{2}}$$
$$0 = (-ln(x) - 1)e^{-\frac{ln(x)^2}{2}}$$
$$0 = (-ln(x) - 1)$$
$$-1 = ln(x)$$
$$x \boxed{= e^{-1}}$$

e) Skew

The skewness, γ_1, is related to the third moment of $f(x)$, that is the expected value of X^3.

$$\gamma_1 = \frac{E[(X-\mu)^3]}{Var(X)^{3/2}}$$

Earlier it was proven that $Var(X) = e^2 - e$. Expanding the numerator...

$$\begin{aligned}(X-\mu)^3 &= (X-\mu)(X-\mu)(X-\mu) \\ &= X^3 - 3X^2\mu - X\mu^2 + \mu^3 \\ &= X^3 - 3X^2\mu\end{aligned}$$

The last two terms cancel because $E[X] = \mu$. Recall, that earlier, we discovered $E[X^2] = e^2$. Plugging in what we know,

$$\gamma_1 = \frac{E[X^3] - 3e^2\sqrt{e}}{(e^2-e)^{3/2}}$$

To determine $E[X^3]$,

$$\begin{aligned}E[X^3] &= \int_0^\infty \frac{x^2}{\sqrt{2\pi}} \left(\frac{1}{x}\right)^{\frac{ln(x)}{2}} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_0^\infty x^2 \left(\frac{1}{x}\right)^{\frac{ln(x)}{2}} dx\end{aligned}$$

Wolfram alpha says

$$\int_0^\infty x^2 \left(\frac{1}{x}\right)^{\frac{ln(x)}{2}} dx = e^{9/2}\sqrt{2\pi}$$

Which Means:

$$E[X^3] = e^{9/2}$$

Plugging this in,

$$\gamma_1 = \frac{e^{9/2} - 3e^{5/2}}{(e^2-e)^{3/2}}$$

This is an example citation [1].

References

[1] E. T. Tatro, S. Hefler, S. Shumaker-Armstrong, B. Soontornniyomkij, M. Yang, A. Yermanos, N. Wren, D. J. Moore, and C. L. Achim. Modulation of bk channel by microrna-9 in neurons after exposure to hiv and methamphetamine. *J Neuroimmune Pharmacol*, 2013. Tatro, Erick T Hefler, Shannon Shumaker-Armstrong, Stephanie Soontornniyomkij, Benchawanna Yang, Michael Yermanos, Alex Wren, Nina Moore, David J Achim, Cristian L R03 DA031591/DA/NIDA NIH HHS/United States U19 AI096113/AI/NIAID NIH HHS/United States Journal article Journal of neuroimmune pharmacology : the official journal of the Society on NeuroImmune Pharmacology J Neuroimmune Pharmacol. 2013 Mar 19.