Contents

List of Figures

1 Introduction

Random number generating algorithms are procedures that can be used to create independent, random observations from particular statistical distributions. This is usually done with the help of computer software packages. Most modern software packages are pre-equipped to produce random samples from a host of statistical distributions. Nevertheless, it is still useful to understand how these software packages produce random samples from these distributions.

Almost all software packages are capable of uniformly generating random real numbers within a fixed interval - a Uniform distribution. This is usually done within the interval 0 and 1. Random number generating algorithms take advantage of this basic capability to produce random observations from a host of statistical distributions, given that the form of the desired distribution's probability density function (pdf) - or a related probability density function - is known. It should be noted that these random observations are actually pseudo random, rather than truly random. [1]

This report investigates 3 common random number generating algorithms. The Inverse Transform Method, the Box Muller Method, as well as the Acceptance Rejection Method. The Inverse Transform Method utilises the inverse Cumulative Distribution Function (CDF) of a distribution to generate random observations from the desired distribution. The Box Muller Method is an algorithm specifically designed to generate random observations from the Normal distribution. While a detailed treatment of the mathematical derivation of this method is beyond the scope of this introductory report, a brief overview of some of the key concepts in the derivation are provided. Lastly, the Acceptance Rejection Method is an algorithm that samples from a target distribution by comparing the pdf of the target distribution to the pdf of an alternative distribution that is easy to sample from.

The structure of the report is as follows. Section 2 provides a detailed overview of the Inverse Transform Method, as well as several detailed examples. Section 3 provides a brief outline and illustration of the Box Muller Method. Section 4 details the Acceptance Rejection Method as well as some additional examples of this algorithm. Section 5 concludes by discussing and comparing the algorithms covered in the report. All examples in this report were simulated in the software package *Mathematica*.

[1]This is because computers are deterministic machines that follow deterministic processes (Lambert, 2018a). This means that these process are initiated by a particular input or cause, and thus are not truly random. Nevertheless, this report is merely an introduction to random number generating algorithms. As a result, no further distinction will be made be pseudo randomness/independence and true randomness/independence in this report.

2 Inverse Transform Method

The Inverse Transform Method (ITM) is one of the simplest random number generating algorithms. The ITM can be used to generate a random number (or a sample of random numbers) from a particular statistical distribution. Suppose that we have some random variable X that follows a known probability function given by $f_X(x)$. Now suppose we want to generate a random sample/number from X's distribution. Lambert (2018a) summaries the ITM algorithm for doing so as follows:

1. We note that X should have a known cumulative density function (CDF) given by:

$$F_X(x) = P(X \leq x) = p \tag{1}$$

   Where p is some percentile. Thus, the input of F_X is x and the output of F_X is p.

2. Furthermore, we note that the percentiles (p) are uniformly distributed between 0 and 1 because $0 \leq p \leq 1$;

$$p_i \sim U(0,1) \tag{2}$$

3. Now we consider the inverse-CDF of X, given by:

$$F_X^{-1}(p) = x \tag{3}$$

   Note that the input of F_X^{-1} is p and the output of F_X^{-1} is x:

4. We know that it is very easy to take a random sample from the uniform distribution. The ITM takes advantage of this fact. We take a random observation from $U \sim UNIF(0,1)$. We plug this into the inverse-CDF of X. This returns a random observation x of the random variable X.

In summary, the Inverse Transform Method takes a random observation from $U \sim UNIF(0,1)$ and *transforms* it into a random observation from X - using the *inverse*-CDF of X. Of course, one significant disadvantage of the ITM is that it cannot be used if the inverse-CDF of X cannot be determined.

2.1 Example 1: Triangular Distribution

We first consider a scenario where we need to draw a sample from the triangular distribution. The general pdf of a triangular distribution over the range [a, b] - and where $a \leq c \leq b$ - is given by:

$$f_X(x; a, b, c) = \begin{cases} \frac{2(x-a)}{(b-a)(c-a)}, & a \leq x < c \\ \frac{2(b-x)}{(b-a)(b-c)}, & c \leq x \leq b \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

As we can see, the pdf is given by a piece-wise function. Because the pdf depends on x as well as a, b and c, we will utilise the notation $X \sim TRI(a, b, c)$ in this report to denote a random variable that follows a Triangular distribution with the parameters a, b and c - where $a \leq c \leq b$. We now consider a random variable X that follows a Triangular distribution with parameter values: $a = -1, c = 0$,and $b = 1$. Therefore, the pdf of X becomes:

$$f_X(x) = \begin{cases} x+1, & -1 \leq x < 0 \\ 1-x, & 0 \leq x \leq 1 \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

Now, we need to find the CDF of this distribution. We can do this for each component of the support of the distribution. Firstly, for $x < 0$:

$$\int_{x'=-1}^{x} (x'+1)dx'$$

$$= [\frac{x'^2}{2}]_{x'=-1}^{x} + [x']_{x'=-1}^{x}$$

$$= [\frac{x^2}{2} - \frac{1}{2}] + [x+1]$$

$$= \frac{x^2}{2} + x + \frac{1}{2}$$

And secondly, for $0 \leq x \leq 1$:

$$\int_{x'=0}^{x} (1-x')dx'$$

$$= [x']_{x'=0}^{x} - [\frac{x'^2}{2}]_{x'=0}^{x}$$

$$= [x - 0] - [\frac{x^2}{2} - 0]$$

$$= -\frac{x^2}{2} + x$$

Therefore, the CDF is given by:

$$F_X(x) = \begin{cases} 0, & x < -1 \\ \frac{x^2}{2} + x + \frac{1}{2}, & -1 \leq x < 0 \\ -\frac{x^2}{2} + x, & 0 \leq x \leq 1 \\ 1 & x > 1 \end{cases} \tag{6}$$

We can now solve for the inverse CDF. Again, we examine $x < 0$ first. We let:

$$y = \frac{x^2}{2} + x + \frac{1}{2}$$

$$x = \frac{y^2}{2} + y + \frac{1}{2}$$

$$x = \frac{1}{2}(1 + y)^2$$

$$2x = (1 + y)^2$$

$$\sqrt{2x} = 1 + y$$

$$y = \sqrt{2x} - 1$$

And secondly, we examine $0 \leq x \leq 1$. We let:

$$y = -\frac{x^2}{2} + x$$

$$x = -\frac{y^2}{2} + y$$

$$x = 1 - \frac{1}{2}(y-1)^2$$

$$x - 1 = -\frac{1}{2}(y-1)^2$$

$$-2x + 2 = (y-1)^2$$

$$\pm\sqrt{2-2x} = y - 1$$

$$y = 1 \pm \sqrt{2-2x}$$

$$y = 1 - \sqrt{2-2x}$$

Thus the inverse CDF is given by:

$$F_X^{-1}(x) = \begin{cases} 0, & x < -1 \\ \sqrt{2x} - 1, & -1 \le x < 0 \\ 1 - \sqrt{2-2x}, & 0 \le x \le 1 \\ 1 & x > 1 \end{cases} \tag{7}$$

We can now generate a random sample from $UNIF(0,1)$ and pass each random observation through this inverse CDF. These transformed values will follow a Triangular distribution. In *Mathematica* we generate a sample of 50 000 observations from $UNIF(0,1)$ and transform each with the inverse CDF above. Note that, because the inverse CDF is also piece wise, we transform all $u_i < \frac{1}{2}$ using $\sqrt{2x} - 1$ and all $u_i \ge \frac{1}{2}$ using $1 - \sqrt{2-2x}$. Plotting the histogram of our transformed X observations in Figure 1 below, we can see that they do indeed follow the pdf of the Triangular distribution with parameters $a = -1, c = 0$,and $b = 1$.

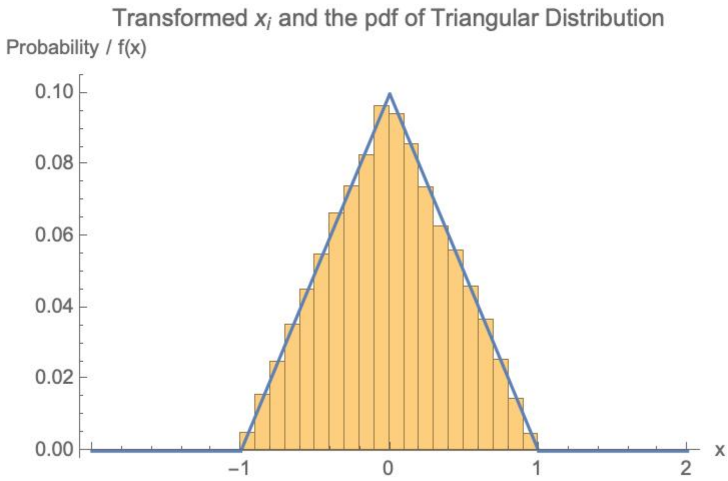

Figure 1: Plot of Transformed X_i Against the pdf of $TRI(-1, 1, 0)$

2.2 Example 2: Exponential Distribution

Now we consider a scenario where it is necessary to draw a random sample from an exponential distribution. In general, the pdf of an Exponentially distributed random variable is given as:

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

We consider a random variable $X \sim EXP(1)$. Thus, the pdf of X is given by:

$$f_X(x) = \begin{cases} e^{-x} & x > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

We first find the CDF of this distribution:

$$\int_{x'=0}^{x} e^{-x'} dx'$$

$$= [-e^{-x'}]_{x'=0}^{x}$$

$$= [-e^{-x} + 1]$$

$$= 1 - e^{-x}$$

Thus the CDF is given by:

$$F_X(x) = \begin{cases} 0 & x \leq 0 \\ 1 - e^{-x} & x > 0 \end{cases} \tag{10}$$

Now we can consider the inverse CDF. Let:

$$y = 1 - e^{-x}$$

$$x = 1 - e^{-y}$$

$$x - 1 = -e^{-y}$$

$$1 - x = e^{-y}$$

$$ln[|1 - x|] = -y$$

$$y = -ln[|1 - x|]$$

Therefore the inverse CDF is given by:

$$F_X^{-1}(x) = \begin{cases} 0 & x \leq 0 \\ -ln[|1-x|] & x > 0 \end{cases} \tag{11}$$

We can now generate a random sample from $UNIF(0,1)$ and pass each observation in the sample through this inverse CDF. In *Mathematica* we generate 50 000 random observations from $UNIF(0,1)$ and transform the using the inverse CDF above. The histogram of these transformed observations should follow the Exponential distribution with rate parameter 1. Figure 2 plots this histogram below:

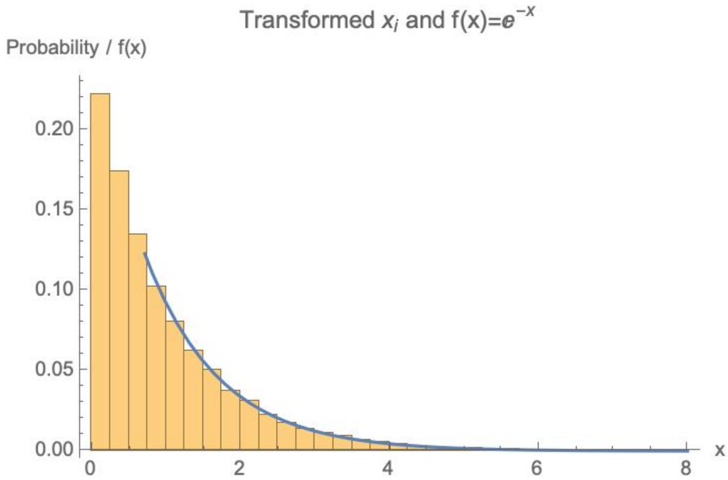

Figure 2: Plot of Transformed X_i Against the pdf of $EXP(1)$

As we can see, the figure confirms that the sample of 50 000 transformed Xs follow an exponential distribution with rate parameter $\lambda = 1$.

2.3 Example 3: Gamma Distribution

Now we consider an example where we want to sample from the Gamma distribution using the ITM. We first consider sampling from the Gamma distribution by using its inverse CDF. We note that if $X \sim GAM(\alpha, \beta)$, then the general from of the pdf of X is given by:

$$f_X(x; \alpha, \beta) = \begin{cases} \dfrac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} & x, \alpha, \beta > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

Where $\Gamma(\alpha)$ is the gamma function with parameter α. Let us briefly consider a Gamma distribution with $\alpha = 10$ and $\beta = 1$:

$$f_X(x) = \begin{cases} \dfrac{x^9 e^{-x}}{9!} & x > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$

Where $\Gamma(10) = (10 - 1)!$. Now we need to consider the CDF of X. This can be calculated by:

$$\int_0^x \frac{x'^9 e^{-x'}}{9!} dx'$$

$$= \frac{1}{9!} \int_0^x x'^9 e^{-x'} dx'$$

This presents us with a problem: the above function cannot be integrated easily. Therefore, it is quite challenging to find the CDF or inverse CDF of X. In fact, this is a common challenge with the Gamma distribution. Furthermore, this highlights one of the most significant weaknesses of using the ITM: it is challenging or impossible to use the ITM when the CDF/inverse-CDF cannot be found or analysed easily.

Nevertheless, we are still able to use the ITM to draw random samples from the Gamma distribution. We do this taking advantage of the unique relationship shared between the Exponential distribution and the Gamma distribution. Misic (2014) notes that if we take any random sample of size n from $X \sim EXP(\lambda)$, then the sample total ($\sum_{i=1}^n X_i$) follows a Gamma distribution with parameters $\alpha = n$ and $\beta = \lambda$. Therefore:

$$\sum_{i=1}^{n} X_i = Y \sim GAM(n, \lambda) \tag{14}$$

In fact, Sahoo (2008) notes that the Exponential distribution is simply a special case of the Gamma distribution where $\alpha = 1$. We take advantage of the above relation between the Exponential and Gamma distributions to sample from the Gamma distribution using the ITM. We will generate this sample from the Gamma distribution in *Mathematica* using the following steps:

1. We use the ITM to generate a random sample of 50 000 observations from an Exponential distribution with parameter $\lambda = 1$ as described in the previous section.
2. Then, we will take many random samples (say, 5 000 samples) of size 5 from this Exponential distribution.
3. Then, in all of the 5 000 samples of size 5, we take the sample total. We call the sample total Y_i (where $i = 1, ..., 5000$). Therefore, we have 5 000 sample totals.
4. Collectively, these sample totals make up a random sample of size 5 000 from the Gamma distribution:

$$Y \sim GAM(5, 1)$$

Figure 3 below plots the histogram of these 5 000 sample totals (each of size 5) taken from an Exponential distribution with $\lambda = 1$. This is plotted against the pdf of a Gamma distribution with $\alpha = 5$ and $\beta = 1$. As we can see, these sample totals are distributed according to $GAM(5, 1)$.

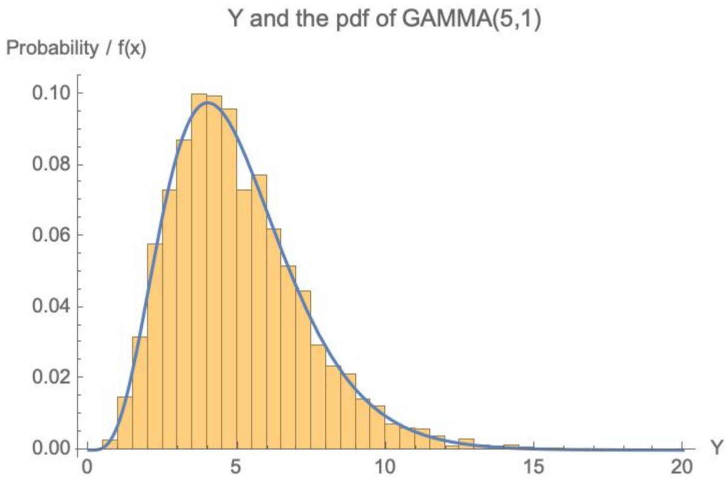

Figure 3: Plot of Y Against the pdf of $GAM(5, 1)$

2.4 Example 4: Beta Distribution

We now consider a scenario where we need to sample from the Beta distribution. As we will see, we run in to problems that are similar to the ones experienced when trying to sample from the Gamma distribution. If a random variable $X \sim BETA(\alpha, \beta)$ - with $\alpha, \beta > 0$ - then the pdf of X is given by:

$$f_X(x) = \begin{cases} \frac{1}{\mathrm{B}(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1} & 0 < x < 1 \\ 0 & \text{otherwise.} \end{cases} \tag{15}$$

Where $\mathrm{B}(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$. Let us briefly consider a Beta distribution with $\alpha = 4.5$ and $\beta = 0.5$. Then , the pdf of X would be given by:

$$f_X(x) = \begin{cases} \frac{1}{\mathrm{B}(4.5,0.5)} x^{3.5}(1-x)^{-0.5} & 0 < x < 1 \\ 0 & \text{otherwise.} \end{cases} \tag{16}$$

Therefore, the CDF of X is given by:

$$\int_0^x \frac{1}{\mathrm{B}(4.5,0.5)} x'^{3.5}(1-x')^{-0.5} dx'$$

$$= \int_0^x \frac{1}{\frac{\Gamma(4.5)\Gamma(0.5)}{\Gamma(4.5+0.5)}} x'^{3.5}(1-x')^{-0.5} dx'$$

$$= \int_0^x \frac{\Gamma(4.5+0.5)}{\Gamma(4.5)\Gamma(0.5)} x'^{3.5}(1-x')^{-0.5} dx'$$

$$= \int_0^x \frac{\Gamma(5)}{\Gamma(4.5)\Gamma(0.5)} x'^{3.5}(1-x')^{-0.5} dx'$$

$$= \int_0^x \frac{4!}{\Gamma(4.5)\Gamma(0.5)} x'^{3.5}(1-x')^{-0.5} dx'$$

Again , we encounter the issue that we cannot easily integrate this function. As a result, it is difficult to find the CDF and inverse-CDF of the Beta distribution. This is a common problem with the Beta distribution, especially for non-integer values of α and β. Nevertheless, we can still sample from the Beta distribution by exploiting one of the remarkable properties of the Gamma distribution. If 2 independent random variables X_1 and X_2 follow:

$$X_1 \sim GAM(\alpha, \theta) \tag{17}$$

$$X_2 \sim GAM(\beta, \theta) \tag{18}$$

Then, as van der Waerden (1969) points out, the random variable $\frac{X_1}{X_1+X_2}$ will be distributed as:

$$\frac{X_1}{X_1 + X_2} \sim BETA(\alpha, \beta) \tag{19}$$

Therefore, if we sample from 2 independent Gamma distributed random variables and combine them as above, we can produce a random sample from a Beta distribution. We will simulate this in *Mathematica* in order to generate a random sample from the Beta distribution. We begin by creating 2 Gamma distributed random variables. In order to do this, we simply employ the method described in the previous section twice. In the first iteration, we create the first Gamma distributed random variable X_1 by doing the following:

1. We use the ITM to generate a random sample of 50 000 observations from an Exponential distribution with parameter $\lambda = 1$ as described in the section on the Exponential distribution.

2. Then, we will take many random samples (say, 5 000 samples) of size 5 from this Exponential distribution.

3. Then, in all of the 5 000 samples of size 5, we take the sample total. We call the sample total $X_{1,i}$ (where $i = 1, ..., 5000$). Therefore, we have 5 000 sample totals.

4. Collectively, these sample totals make up a random sample of size 5 000 from the Gamma distribution:

$$X_1 \sim GAM(5, 1)$$

We then repeat this process, taking a different sample size from the Exponential distribution this time in order to produce our second Gamma distributed random variable X_2:

1. We use the ITM to generate a random sample of 50 000 observations from an

Exponential distribution with parameter $\lambda = 1$ as described in the section on the Exponential distribution.

2. Then, we will take many random samples (say, 5 000 samples) of size 10 from this Exponential distribution.

3. Then, in all of the 5 000 samples of size 10, we take the sample total. We call the sample total $X_{2,i}$ (where $i = 1, ..., 5000$). Therefore, we have 5 000 sample totals.

4. Collectively, these sample totals make up a random sample of size 5 000 from the Gamma distribution:

$$X_2 \sim GAM(10, 1)$$

Based on this, we can construct a sample from the Beta distribution. By computing:

$$\frac{X_1}{X_1 + X_2}$$

we will get a random sample of size 5 000 distributed as $BETA(5, 10)$. Figure 4 below plots the histogram of this sample of size 5 000 against the pdf of BETA(5, 10). As we can see, the histogram follows the desired Beta distribution.

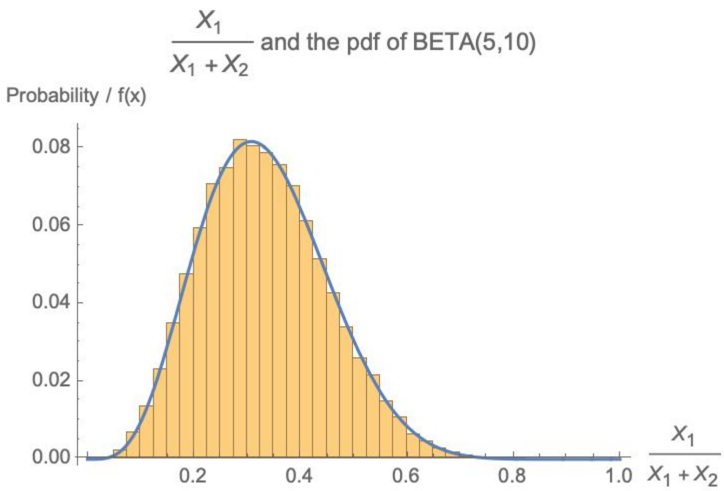

Figure 4: Plot of $\frac{X_1}{X_1+X_2}$ Against the pdf of $BETA(5, 10)$

2.5 Example 5: Chi-Square Distribution

We now consider a case where we need to sample from the Chi-Square distribution. As Sahoo (2008) notes, the Chi-Square distribution is a special case of the gamma distribution when $\alpha = \frac{r}{2}$ and $\beta = \frac{1}{2}$. If $X \sim \chi^2(r)$ then the pdf of X is given by:

$$f_X(x;r) = \begin{cases} \dfrac{x^{\frac{r}{2}-1}e^{-\frac{1}{2}x}}{\Gamma(\frac{r}{2})2^{\frac{r}{2}}} & 0 < x < \infty \\ 0 & \text{otherwise.} \end{cases} \tag{20}$$

Therefore, in order to generate a sample from $\chi^2(r)$ we need to simply generate a sample from $GAM(\alpha, \frac{1}{2})$. We do this by first sampling from an Exponential distribution with $\lambda = \frac{1}{2}$. Then we take a random sample from this distribution of size n and compute the sample total. This sample total will follow $GAM(n, \frac{1}{2})$ - where $2 * n = r$. This will produce the desired sample from a Chi-Squared distribution with $2 * n = r$ degrees of freedom.

We start by considering the random variable $W \sim EXP(\frac{1}{2})$. The pdf of W is thus:

$$f_W(w) = \begin{cases} \dfrac{1}{2}e^{-\frac{1}{2}w} & w > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{21}$$

The CDF is given by:

$$\int_0^w \frac{1}{2}e^{-\frac{1}{2}w'}dw'$$

$$= 1 - e^{-\frac{1}{2}w}$$

Thus the CDF is:

$$F_W(w) = \begin{cases} 0 & w \leq 0 \\ 1 - e^{-\frac{1}{2}w} & w > 0 \end{cases} \tag{22}$$

Therefore, the inverse CDF is given by:

$$F_W^{-1}(w) = \begin{cases} 0 & w \leq 0 \\ -2ln[|1 - w|] & w > 0 \end{cases} \tag{23}$$

We are now in the position to generate a sample from the Chi-Squared distribution in *Mathematica*. We can do this by taking the following stpes:

1. We use the ITM to generate a random sample of 50 000 observations from an Exponential distribution with parameter $\lambda = \frac{1}{2}$ as described in the section on the Exponential distribution. In our case, we use F_W^{-1} above.

2. Then, we will take many random samples (say, 5 000 samples) of size 5 from this Exponential distribution.

3. Then, in all of the 5 000 samples of size 5, we take the sample total. We call the sample total $\sum_{j=1}^{5} W_i = Y_i$ (where $i = 1, ..., 5000$). Therefore, we have 5 000 sample totals.

4. Collectively, these sample totals make up a random sample of size 5 000 from the Gamma distribution:

$$Y \sim GAM(5, \frac{1}{2})$$

5. However, because the Chi-Square distribution is a special case of the Gamma distribution where $\alpha = \frac{r}{2}$ and $\beta = \frac{1}{2}$, we can rewrite this. We note that β is indeed $\frac{1}{2}$ and $5 = \frac{r}{2}$ which means $r = 10$. Therefore, the sample totals make up a random sample of size 5 000 from the Chi-Squared distribution:

$$Y \sim \chi^2(10)$$

Figure 5 below plots the histogram of the aforementioned sample totals against the pdf of $\chi^2(10)$. The figure indicates that the sample totals do indeed follow a Chi-Square distribution with 10 degrees of freedom.

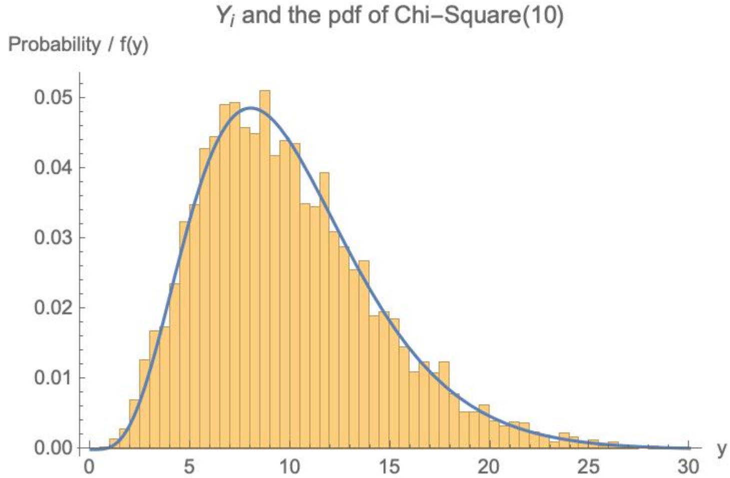

Figure 5: Plot of Sample Totals From $EXP(\frac{1}{2})$ Against the pdf of $\chi^2(10)$

3 Box-Muller Method: The Standard Normal Distribution

The Box-Muller Method (BMM), or the Box-Muller Transformation, is a method specifically devised to sample from the Normal distribution. The simple ITM cannot be used to sample from the normal distribution because the CDF of the normal distribution is notoriously hard to define - it requires the use of the error function (Nguyen, 2020).

The BMM requires us to consider the pdf of a 2-dimensional normal distribution, which is transformed using polar coordinates (Goodman, 2005). Let us consider two independent random variables that follow a standard normal distribution: X and Y. We know that the joint density function (the 2 dimensional distribution) of the independent X and Y is given by:

$$f(x,y) = \frac{1}{\sigma_X\sqrt{2\pi}}exp[-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2] * \frac{1}{\sigma_Y\sqrt{2\pi}}exp[-\frac{1}{2}\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2]$$

$$f(x,y) = \frac{1}{\sqrt{2\pi}}exp[-\frac{1}{2}x^2] * \frac{1}{\sqrt{2\pi}}exp[-\frac{1}{2}y^2]$$

$$f(x,y) = \frac{1}{2\pi}exp[-\frac{1}{2}(x^2+y^2)] \tag{24}$$

Goodman (2005) notes that, because this distribution is radially symmetric, we can rewrite it using polar co-ordinates. We do this using the polar co-ordinate random variables r and θ. We achieve this transformation by defining:

$$x = r * cos(\theta) \tag{25}$$

$$y = r * sin(\theta) \tag{26}$$

Because θ is defined as the angular coordinate in polar form,it is uniformly distributed as follows (Goodman, 2005):

$$\theta \sim UNIF(0, 2\pi) \tag{27}$$

And thus:

$$f_\theta(\theta) = \begin{cases} \frac{1}{2\pi} & 0 \le x \le 2\pi \\ 0 & \text{otherwise.} \end{cases} \tag{28}$$

which is very easy to sample from. Furthermore, r is defined as the radial coordinate in polar form. Using the Theorem of Pythagoras it can be written as:

$$r = \sqrt{x^2 + y^2} \tag{29}$$

Therefore, making the appropriate substitutions, we can rewrite the joint 2-D pdf in terms of r as:

$$f(r) = \frac{1}{2\pi} r * exp[-\frac{1}{2}r^2] \tag{30}$$

This function of r does have a primitive. As a result, we can find its inverse CDF and use the ITM to sample from rs distribution. In summary, the BMM requires us to sample from the densities of r and θ because they are easy to sample from. We then take the samples from both r and θ and transform them into random samples from X and Y using equations (27) and (28). The random samples from both X and Y will each follow a Standard Normal Distribution.

In order to proceed with the BMM, we first need the inverse CDF of r. We can compute the CDF of r as:

$$F(r) = 1 - exp[-\frac{1}{2}r^2] \tag{31}$$

Similarly, we can compute the inverse-CDF. Let:

$$y = 1 - exp[-\frac{1}{2}r^2]$$

$$r = 1 - exp[-\frac{1}{2}y^2]$$

$$1 - r = exp[-\frac{1}{2}y^2]$$

$$ln[1 - r] = -\frac{1}{2}y^2$$

$$2 * ln[\frac{1}{1 - r}] = y^2$$

$$y = \sqrt{2 * ln[\frac{1}{1 - r}]}$$

And thus the inverse CDF is given by:

$$F^{-1}(r) = \sqrt{2 * ln[\frac{1}{1-r}]} \tag{32}$$

We can now simulate this algorithm in *Mathematica* to generate a random sample from $X \sim N(0,1)$. We do this by taking the following steps:

1. We first use the ITM to generate a random sample of size 200 000 from r using $F^{-1}(r)$.

2. We then generate a random sample of size 200 000 from θ by simply sampling from $UNIF(0, 2\pi)$.

3. We then transform these random samples from r and θ into a single random sample from $X \sim N(0,1)$ by using the fact that:

   $$x = r * cos(\theta)$$

   The resulting transformation will produce a random sample of size 200 000 that is distributed $N(0,1)$.

Figure 6 below provides the histogram of the transformed sample from X. This histogram is plotted against the pdf of a Standard Normal Distribution. As we can see, the transformed sample is distributed as $N(0,1)$.

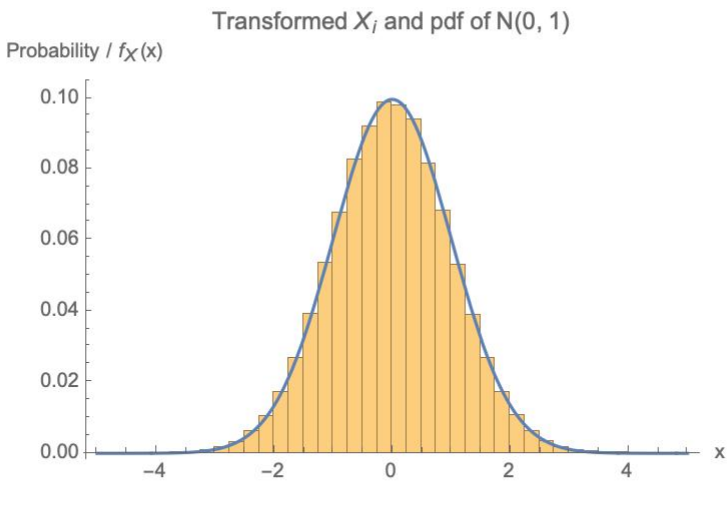

Figure 6: Plot of Transformed X_i Against the pdf of $N(0,1)$

4 Acceptance Rejection Method

The Acceptance Rejection Method (ARM) is another commonly used method of generating a random sample from a particular distribution. It is a type of Monte Carlo random sampling method. Suppose that we have some random variable X that follows a known probability function given by $f_X(x)$. Now suppose we want to generate a random sample from X's distribution (the target distribution). Sigman (2007) summarises the relevant steps of the ARM algorithm as the following:

1. We want to sample from the distribution $f_X(x)$ but it is difficult to do so. We first consider a distribution - given by $g(x)$ - that is easy to sample from. We call $g(x)$ the proposal distribution. We also consider some constant c. We must choose $g(x)$ and c such that:

$$f_X(x) \leq c * g(x) \tag{33}$$

   The implication here is that when our proposal distribution is multiplied by some constant c, then it will completely envelop our target distribution.

2. We are free to choose the appropriate $g(x)$. However, once this is chosen, c must be chosen by:

$$\frac{f_X(x)}{g(x)} \leq c \tag{34}$$

   Therefore c must at least be equal to:

$$Max\left[\frac{f_X(x)}{g(x)}\right] = c \tag{35}$$

3. Now we generate a random variable Y that follows the distribution given by the proposal distribution $g(x)$. We sample n observations from Y and call each observation y_i.

4. Then we generate a random variable $U \sim UNIF(0,1)$. We sample n observations from U and call each observation u_i.

5. We then set up our acceptance-rejection rule. We accept y_i as an observation from our target distribution $f_X(x)$ if:

$$u_i \leq \frac{f_X(y_i)}{c * g(y_i)} \tag{36}$$

Otherwise, we reject y_i as an observation from our target distribution $f_X(x)$ if:

$$u_i > \frac{f_X(y_i)}{c * g(y_i)} \tag{37}$$

6. We repeat this process for all i. All accepted y_i will thus be distributed according to the target distribution $f_X(x)$.

In summary, we want to sample from the target distribution given by $f_X(x)$. However, because this is challenging, we instead sample from an easier proposal distribution $g(x)$ as well as $UNIF(0, 1)$. We accept some observations from $g(x)$ and reject other observations from $g(x)$ based on our acceptance-rejection rule. The accepted observations from $g(x)$ will be distributed according to $f_X(x)$.

Therefore, it is clear that using the ARM instead of the ITM can be advantageous, as we do not need to know the CDF of the distribution we are sampling from. However, the ARM has its own drawbacks. When using the ARM, we have to discard (reject) many of the observations that we generate because they do not satisfy our acceptance-rejection rule. This means that the ARM is a relatively inefficient sampling method when compared to the ITM - as the ITM utilises 100% of the observations it generates.

4.1 Example 1: Exponential Distribution

We begin by considering a scenario where we need to sample from the Exponential distribution. We know if $X \sim EXP(\lambda)$ then the pdf of X is given by:

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{38}$$

Again, we will consider the case where $\lambda = 1$:

$$f_X(x) = \begin{cases} e^{-x} & x > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{39}$$

We now need to consider a proposal distribution that - when multiplied by some constant c - envelops our Exponential distribution with parameter 1. It is difficult to consider a proposal distribution that has a range of $x > 0$. However, we can take advantage of the fact that larger values of x are extremely unlikely if $X \sim EXP(1)$. We can see this in Figure 7 below:

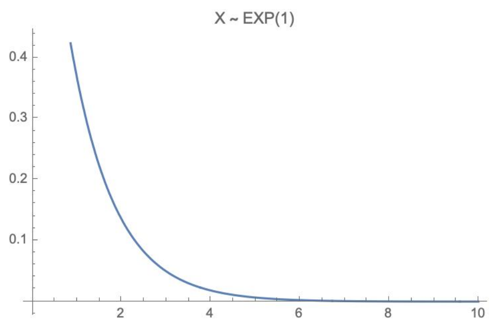

Figure 7: Plot of $f(x) = e^{-x}$

Based on this plot, it can be seen that the probability of obtaining a value of x that is greater than 8 is extremely unlikely. Therefore, for the purposes of our simulation, we can restrict out analysis to the x-range of $0 < x \leq 8$.[2]

It should be noted that because we now have a finite range of x values that we are considering, the uniform distribution can serve as our proposal distribution! This is because the Uniform distribution can envelop any distribution with finite (in our case pseudo-finite) support. Therefore, the proposal distribution that will envelop $f_X(x)$

[2] This decision is based on the tutorial provided by Lambert (2018b).

across the support $0 < x \leq 8$ (once multiplied by a constant) is a Uniform distribution between 0 and 8. In other words:

$$g_X(x) = \begin{cases} \frac{1}{8} & 0 < x < 8 \\ 0 & \text{otherwise.} \end{cases} \tag{40}$$

$$Y \sim UNIF(0, 8) \tag{41}$$

Now that we have chosen a $g_X(x)$, we can turn our attention to determining the appropriate value for c. Recall that the condition that c needs to satisfy is given by:

$$\frac{f_X(x)}{g(x)} \leq c$$

$$\frac{e^{-x}}{\frac{1}{8}} \leq c$$

$$8e^{-x} \leq c$$

Therefore, c must be at least equal to:

$$Max[8e^{-x}] = c$$

We know that, because x must be greater than 0 in our case, the maximum of our Exponential function across this range for x occurs at $x \approx 0$:

$$\text{when } x \approx 0 : \ 8e^{-0} = 8 = c$$

We can now sample from $U \sim UNIF(0, 1)$ as well as $Y \sim UNIF(0, 8)$ (i.e. Y is distributed according to $g(x)$) and utilise our acceptance-rejection rule. We know that our acceptance-rejection rule will be to accept all y_i that satisfy:

$$u_i \leq \frac{f_X(y_i)}{8 * g(y_i)} \tag{42}$$

All accepted y_i will be distributed according to an Exponential Distribution with parameter $\lambda = 1$. We simulate this exact procedure for 50 000 u_i and y_i using *Mathematica*. We accept only the y_i that satisfy our acceptance-rejection rule. Figure 8 below plots the histogram of accepted y_i against the pdf of $EXP(1)$. We can see that the accepted y_i do indeed follow an Exponential distribution with parameter $\lambda = 1$.

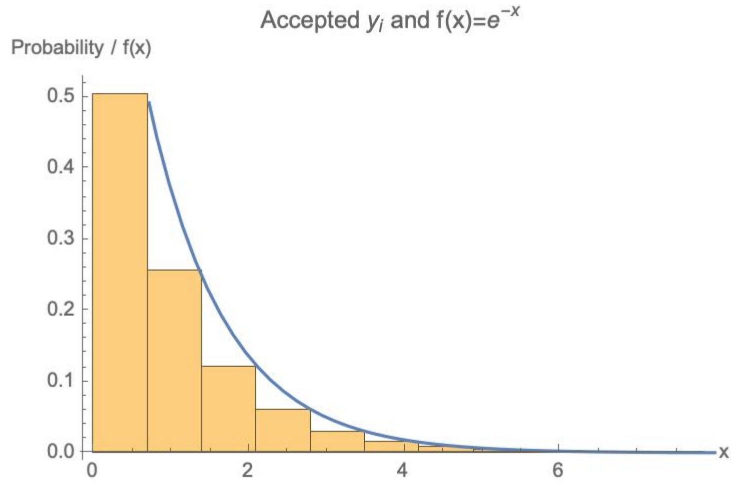

Figure 8: Plot of Accepted y_i Against an Exponential Distribution With $\lambda = 1$

4.2 Example 2: Polynomial Distribution

Now we consider a scenario where we need to sample from some arbitrary polynomial distribution[3]. We assume that the pdf of some random variable X is given by the polynomial:

$$f_X(x) = \frac{3}{2}x^3 + \frac{11}{8}x^2 + \frac{1}{6}x + \frac{1}{12}; for : 0 \leq x \leq 1 \tag{43}$$

This distribution is illustrated graphically in Figure 9 below:

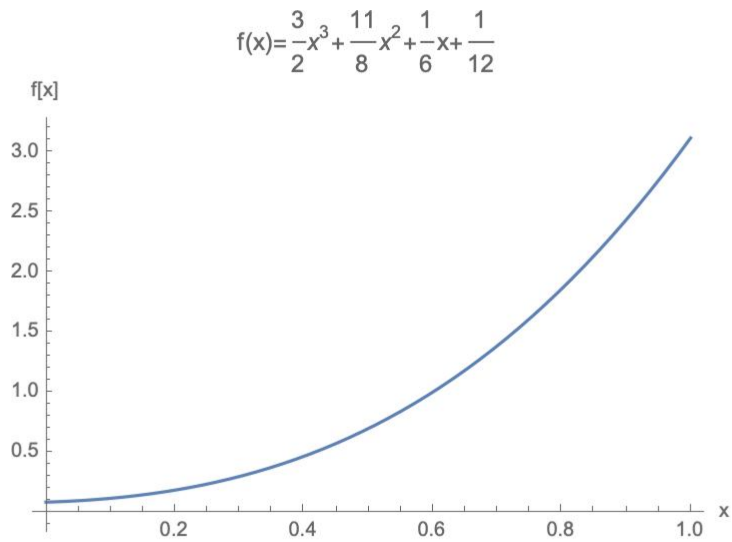

Figure 9: Plot of Polynomial Distribution for $0 \leq x \leq 1$

It can be shown that this is indeed a valid pdf over the support $0 \leq x \leq 1$ as it is normalised:

$$\int_0^1 \frac{3}{2}x^3 + \frac{11}{8}x^2 + \frac{1}{6}x + \frac{1}{12} = 1 \tag{44}$$

Let us now consider what distribution to choose for our proposal distribution $g(x)$. Because $f_X(x)$ has a finite support $(0 \leq x \leq 1)$, we can use the Uniform distribution over 0 and 1. Therefore:

[3]This example is based on the tutorial given in: https://www.youtube.com/watch?v=kMb4JlvuGlw&t=446s - (no author provided).

$$g_X(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{otherwise.} \end{cases} \tag{45}$$

$$Y \sim UNIF(0,1) \tag{46}$$

Now that $g(x)$ has been selected, we can choose our c based on:

$$\frac{f_X(x)}{g(x)} \leq c$$

$$\frac{\frac{3}{2}x^3 + \frac{11}{8}x^2 + \frac{1}{6}x + \frac{1}{12}}{1} \leq c$$

$$\frac{3}{2}x^3 + \frac{11}{8}x^2 + \frac{1}{6}x + \frac{1}{12} \leq c$$

Thus, c must at least be equal to:

$$Max\left[\frac{3}{2}x^3 + \frac{11}{8}x^2 + \frac{1}{6}x + \frac{1}{12}\right] = c$$

We know from Figure 9 that $g(x)$ is only a valid pdf for x between 0 and 1 and it is increasing over this interval. Therefore, the maximum will be given when $x = 1$:

$$g(1) = 3.125 = c$$

We can now sample from $U \sim UNIF(0,1)$ as well as $Y \sim UNIF(0,8)$ (i.e. Y is distributed according to $g(x)$). Thereafter, we utilise our acceptance-rejection rule. We know that our acceptance-rejection rule will be to accept all y_i that satisfy:

$$u_i \leq \frac{f_X(x_i)}{3.125 * g(x_i)} \tag{47}$$

We simulate this example using *Mathematica* for 50 000 u_i and y_i. We only accept the y_i that satisfy the acceptance-rejection rule set out above. In Figure 10 below, we plot the histogram of all these accepted y_i. In the figure, we can see that the accepted y_i do indeed follow the the desired polynomial distribution set out above.

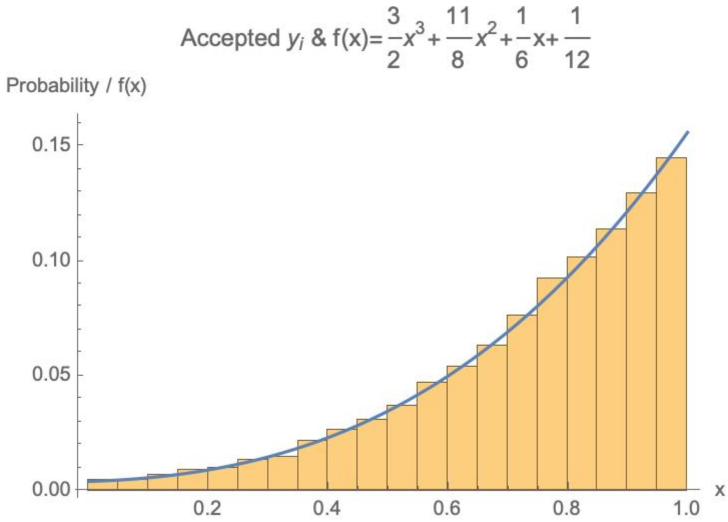

Figure 10: Plot of Accepted y_i Against Polynomial PDF

4.3 Example 3: Beta Distribution

Now, let let us suppose that we want to sample from the random variable X that follows a Beta distribution with $\alpha = \beta = 1.5$:

$$f_X(x) = \begin{cases} \dfrac{1}{\mathrm{B}(1.5, 1.5)} x^{0.5}(1-x)^{0.5} & 0 < x < 1 \\ 0 & \text{otherwise.} \end{cases} \tag{48}$$

We noted when trying to sample from the Beta distribution using the ITM that it was very difficult to do when α and β took on non-integer values. In this example, we have non-integer values for both α and β. We will see that it is much easier to sample from the Beta distribution using the ARM when α and β are not integers.

Because the graphical form of the Beta distribution can vary drastically as the parameters α and β vary, it is useful to visualise what this distribution might look like. Figure 11 provides this illustration below:

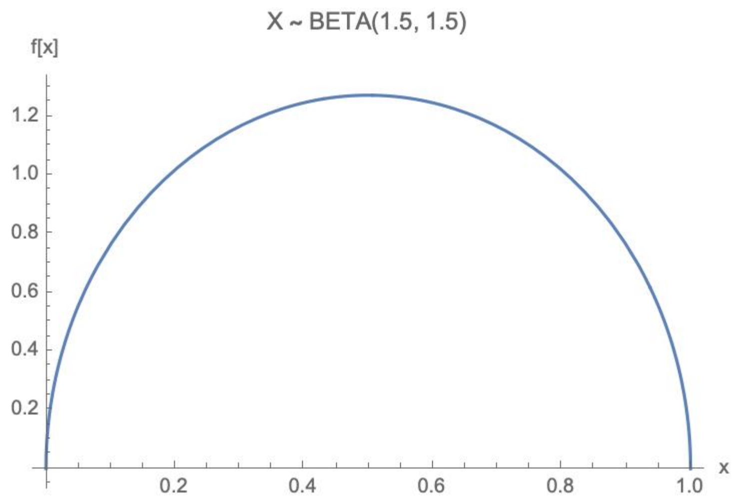

Figure 11: Plot of the pdf of $X \sim BETA(1.5, 1.5)$

We must first determine what distribution to choose for our proposal distribution $g(x)$. Because $f_X(x)$ again has a finite domain, we can use the Uniform distribution over 0 and 1. This means that:

$$g_X(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{otherwise.} \end{cases} \tag{49}$$

$$Y \sim UNIF(0,1) \tag{50}$$

Therefore, we can now choose c based on:

$$\frac{f_X(x)}{g(x)} \leq c$$

$$\frac{\frac{1}{\mathrm{B}(1.5,1.5)}x^{0.5}(1-x)^{0.5}}{1} \leq c$$

$$\frac{1}{\mathrm{B}(1.5,1.5)}x^{0.5}(1-x)^{0.5} \leq c$$

Thus, c must at least be equal to:

$$Max\left[\frac{1}{\mathrm{B}(1.5,1.5)}x^{0.5}(1-x)^{0.5}\right] = c$$

Therefore, we can attempt to maximise f(x) by:

$$\frac{d}{dx}\left[\frac{1}{\mathrm{B}(1.5,1.5)}x^{0.5}(1-x)^{0.5}\right] = 0$$

$$\frac{d}{dx}\left[\frac{1}{\mathrm{B}(1.5,1.5)}x^{0.5}(1-x)^{0.5}\right] = 0$$

$$\frac{1}{\mathrm{B}(1.5,1.5)}\frac{d}{dx}[x-x^2]^{0.5} = 0$$

$$\frac{1}{\mathrm{B}(1.5,1.5)}0.5[x-x^2]^{-0.5}[1-2x] = 0$$

$$\frac{1}{\mathrm{B}(1.5,1.5)}\frac{1-2x}{2\sqrt{x-x^2}} = 0$$

$$1-2x = 0$$

$$1 = 2x$$

$$x = \frac{1}{2}$$

Therefore, at $x = \frac{1}{2}$ we have:

$$f_X(\frac{1}{2}) = \frac{1}{\mathrm{B}(1.5, 1.5)}(0.5)^{0.5}(1-(0.5))^{0.5}$$

$$f_X(\frac{1}{2}) = 2.54648...(0.5)$$

$$f_X(\frac{1}{2}) = 1.27324...$$

Thus c must be at least:

$$c = 1.27324...$$

We can now sample from $U \sim UNIF(0, 1)$ as well as $Y \sim UNIF(0, 8)$ (i.e. Y is distributed according to $g(x)$). Thereafter, we utilise our acceptance-rejection rule. We know that our acceptance-rejection rule will be to accept all y_i that satisfy:

$$u_i \leq \frac{f_X(x_i)}{1.27324... * g(x_i)} \tag{51}$$

We again simulate this problem in *Mathematica* for 50 000 u_i and y_i - accepting only the y_i that satisfy our acceptance-rejection rule. Figure 12 below plots the histogram of all accepted y_i against the pdf of X. The figure indicates that all y_i are distributed according to $BETA(1.5, 1.5)$.

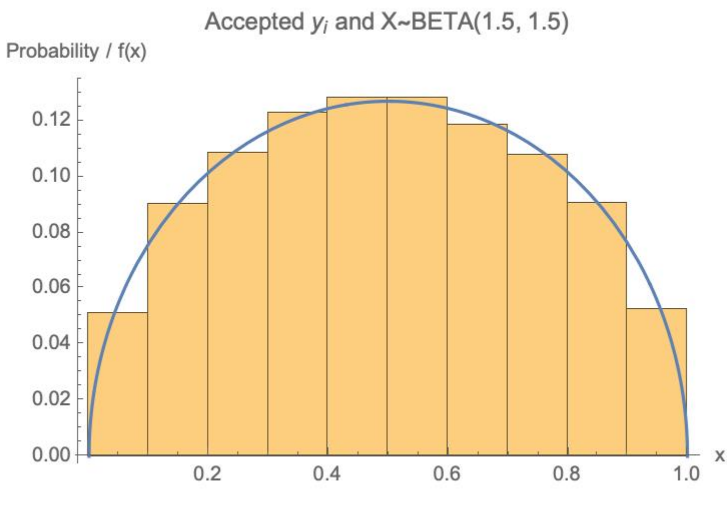

Figure 12: Plot of Accepted y_i Against pdf of $X \sim BETA(1.5, 1.5)$

5 Discussion & Conclusion

As we have seen in the examples above, there are various methods that can be used to generate random samples from particular statistical distributions. Each method takes advantage of the fact that it is relatively easy to draw a random sample from a Uniform distribution.

The ITM allowed us to generate a random sample from a particular density by using the inverse of its CDF. While this method proved to be highly accurate, it is a strenuous requirement to always be able to use the inverse CDF. As the examples illustrated, this can limit the direct application of the ITM to certain distributions where the inverse CDF is not easily obtainable. In these cases, it requires one to be able to take advantage of special relationships between different distributions - which may not always exist.

The BMM is a particular example of how the short comings of the ITM can be circumvented in order to sample from the Normal distribution - which has a notoriously ill-defined CDF. However, while we saw that the implementation of this method is straightforward, its derivation is not.

Furthermore, we considered the ARM. This algorithm allows one to sample from a wide range of distributions without having to utilise a CDF. However, this method becomes relatively more difficult to use when the support of the target distribution is not finite. Furthermore, the ARM has the added disadvantage of discarding many of the sample values generated by the algorithm.

This brings us to the topic of the efficiency of the methods discussed in the report. It is self evident that the ARM is the least efficient algorithm discussed above. This is because it generates a relatively large amount of values that do not belong to the desired target distribution. The ITM and BMM on the other hand are 100% efficient (Lambert, 2018b), as all of the values generated by these methods belong to the desired target distribution. In running the simulations for the examples provided above, I can note the simulations of the ATM and BMM algorithms required relatively less computational time when compared to the ARM simulations. The ARM examples often required relatively more time to complete simulations with comparable sample sizes.

This report has explained and demonstrated a wide range of random number generating algorithms. The report has explained and illustrated some of the most prominent advantages and disadvantages of utilising the 3 methods discussed herein. However, this report merely serves as an introductory tutorial for understanding basic random number generating algorithms. There exists a wide range of random number generating algorithms that were not discussed here, such as Importance Sampling. Many of these algorithms fall outside the scope of the Mathematics of statistics and probability distributions - such as von Neumann's Middle Square Method. Similarly, a treatment of the distinction between true random number generation and pseudo random number generation did not fall within the scope of this report. Nevertheless, the reader should feel comfortable with utilising the random number generating algorithms discussed above, while also feeling well equipped to delve deeper into the field of random number generation.

6 References

Goodman, J. (2005). Lecture Notes on Monte Carlo Methods - Chapter 2: Simple Sampling of Gaussians. [PDF]. Courant Institute of Mathematical Sciences, New York University. Retrieved 28 May 2021, from https://www.math.nyu.edu/ goodman/teaching/MonteCarlo2005/notes/GaussianSampling.pdf.

Lambert, B. (2018). An Introduction to Inverse Transform Sampling. youtube.com. Retrieved 28 May 2021, from https://www.youtube.com/watch?v=rnBbYsysPaU.

Lambert, B. (2018). An Introduction to Rejection Sampling. youtube.com. Retrieved 28 May 2021, from https://www.youtube.com/watch?v=kYWHfgkRc9s.

Misic, B. (2014). Gamma as a Sum of Exponentials.. Researchgate. Retrieved 28 May 2021, from https://www.researchgate.net/figure/Gamma-as-a-sum-of-exponentials-A-gamma-distributed-random-variable-with-shape-parameter_fig4_261405464.

Nguyen, K. (2020). How to Generate Gaussian Samples. Medium. Retrieved 28 May 2021, from https://medium.com/mti-technology/how-to-generate-gaussian-samples-3951f2203ab0.

Sahoo, P. (2008). Probability and Mathematical Statistics (1st ed.). University of Louisville.

Sigman, K. (2007). The Acceptance-Rejection Method [PDF]. Columbia University. Retrieved 28 May 2021, from http://www.columbia.edu/ ks20/4703-Sigman/4703-07-Notes-ARM.pdf.

van der Waerden, B. (1969). Mathematical statistics. Springer-Verlag.