Stochastic Modeling of Heat Transfer through Heterogeneous Materials

Steven Brill

Department of Aerospace and Mechanical Engineering

University of Notre Dame

Prepared for: AME 48491 Undergraduate Research

Advisors: Dr. Joseph M. Powers, Dr. Samuel Paolucci, Dr. Waad Subber

June 3, 2015

Abstract

Heterogeneous materials contain uncertainty in their degree of mixing; this affects the thermal diffusivity. In order to quantify the consequences of the uncertainty in material properties on quantities of interest, such as temperature, the one dimensional heat equation was modeled using the Monte Carlo method, intrusive polynomial chaos, and non-intrusive polynomial chaos. The Monte Carlo method can be computationally expensive. In order to reduce the computational time, so-called polynomial chaos methods were used. It was shown that the intrusive polynomial chaos method is able to approximate solutions to the stochastic heat equation for Gaussian distributions of the thermal diffusivity. It was also shown that the non-intrusive method produces a solution that converges to the Monte Carlo solution. The Monte Carlo method took 4.27 hours of computation to achieve a final temperature with a mean of 477 and a standard deviation of 40; in contrast the non-intrusive polynomial chaos method achieved nearly identical results in 1.54 seconds.

Contents

1 Introduction **5**

2 Heat Equation **5**

3 Method **6**

3.1 Monte Carlo 7

3.2 Intrusive Polynomial Chaos 8

3.3 Non-Intrusive Polynomial Chaos 9

4 Results **10**

4.1 Monte Carlo Method 10

4.2 Intrusive Polynomial Chaos 12

4.3 Non-Intrusive Polynomial Chaos 14

5 Discussion and Conclusions **15**

6 Future Work **18**

7 Appendix A **19**

8 Appendix B **22**

9 Appendix C **23**

10 Appendix D **26**

10.1 Introduction to Probability 26

10.1.1 Random Variables 26

10.1.2 Discrete and Continuous Random Variables 27

10.1.3 Probability Distributions 28

10.2 Mathematical Expectation 31

10.2.1 Expected Value 31

10.2.2 Moments 32

10.2.3 Chebyshev's Inequality 35

10.2.4 Law of Large Numbers 35

10.2.5 Miscellaneous topics 36

10.3 Special Probability Distributions 37

10.3.1 Gaussian Distribution 37
10.3.2 Poisson Distribution 38
10.3.3 The Central Limit Theorem 40
10.3.4 Uniform Distribution 41
10.4 Sampling Theory 41
10.4.1 Basics 41

1 Introduction

The goal of computational science is to develop models that predict phenomena observed in nature. However, these models are often based on parameters that are uncertain. For example, when making a weather prediction, it is impossible to know the entire temperature field in the atmosphere to predict the creation of clouds. Uncertainty in the initial problem formulation affects future calculations and has an impact on the final solution. Hence, it is necessary to understand how the uncertainty manifests itself in the solution, that is the focus of the field of uncertainty quantification (UQ). In UQ some common metrics are the probability density function of the solution, the mean and variance of the solution, and the confidence interval of the solution [1].

One class of problems that contains uncertainty is the study of heterogeneous materials. A heterogeneous material is a material made of different materials. Heterogeneous materials are commonly found in daily life. For example, Swiss cheese is a heterogeneous material because it is a mixture of cheese and voids. When solving a problem with a heterogeneous material, it is uncertain how the material is mixed. For example, if one were to select a random point in the Swiss cheese, it is uncertain whether one would select cheese or a void. The degree of mixture of the materials will also affect the bulk material properties of the body. The material properties will more closely resemble the material in higher concentration. However, even if two bodies had the same concentrations of materials, it is possible that the materials would be more poorly mixed in one of the bodies that would also affect the material properties. In summary, uncertainty in the heterogeneity in a material affects the material properties, that will affect later predictions. Hence the results of the uncertainty due to material heterogeneity must be quantified. In this report, the effect of material heterogeneity on conduction heat transfer will be studied, because heat transfer through heterogeneous materials has implications in a wide range of fields, like the creation of new materials from heterogeneous mixtures.

2 Heat Equation

Conduction heat transfer through a material is often well modeled by

$$\frac{\partial T}{\partial t} = \alpha \nabla^2 T, \tag{1}$$

where T is the temperature of the material, t is time, α is the thermal diffusivity of the material, and ∇^2 is the Laplacian operator. Equation (1) is a parabolic partial differential equation (PDE) with respect to space and time. It can be solved by a variety of analytic and numerical methods. In this study, the heat

equation was solved using the NDsolve function in Wolfram Mathematica. This software uses standard numerical methods to discretize the equations in space and time and generate an approximate solution [2]. The thermal diffusivity of a pure material material is defined as

$$\alpha = \frac{k}{\rho c_p}, \tag{2}$$

where k is the thermal conductivity, ρ is the density, and c_p is the specific heat capacity of the material. It is assumed that the thermal diffusivity is a constant material property, because in solids k, ρ, and c_p do not vary much with temperature. Also, α must be positive, because physically, a negative thermal diffusivity would violate the second law of thermodynamics by allowing temperature to flow from a cold body to a hot body. Also, mathematically, solutions to Eq. (1) are unstable when α is negative.

3 Method

Equation (1) is deterministic. It does not contain any uncertainty. Because α is a material property, the uncertainty due to the material heterogeneity is added to the heat equation by defining the thermal diffusivity

$$\alpha(\xi) = \mu_\alpha + \sigma_\alpha \xi, \tag{3}$$

where μ_α is the mean of thermal diffusivity, σ_α is the standard deviation of the thermal diffusivity, and $\xi \in (-\infty, \infty)$ is a random variable with a probability distribution function (pdf) of a Gaussian distribution with a mean of 0 and a standard deviation of 1. The Gaussian distribution in described in Appendix D. Hence, $\alpha(\xi)$ has a pdf of a Gaussian with mean of μ_α and a standard deviation σ_α. Each ξ sampled from the Gaussian represents a different thermal diffusivity. However, because there is no spatial component in the formulation of $\alpha(\xi)$, each sample of ξ is equivalent to sampling a new material. For example, if the heterogeneity of a material is described by a Gaussian distribution, like in Eq. (3), each ξ value would represent a different mixture of the materials. However, it assumes the same thermal diffusivity throughout the whole body. In order to simplify the problem, it is assumed that the problem is one dimensional. Hence, the stochastic heat equation is modeled as

$$\frac{\partial T(x,t,\xi)}{\partial t} = \alpha(\xi) \frac{\partial^2 T(x,t,\xi)}{\partial x^2}. \tag{4}$$

The set of initial and boundary conditions is chosen to be

$$T(0,t) = T_0, \tag{5}$$

$$T(1,t) = T_0, \tag{6}$$

$$T(x,0) = T_0 + T_0 \sin(\pi x), \tag{7}$$

for a domain of $x \in (0,1)$ and $T_0 = 300$. Because T is now a function of ξ that is described by a pdf, there is also a pdf that represents T at each point in space and time. Equation (4) is not a deterministic problem, so it cannot be solved using traditional numerical methods. However, if a specific ξ_0 was sampled from the pdf, Eq. (4) could be solved deterministically like Eq. (1) [3]. In order to develop a general solution to the stochastic heat equation, three techniques are used: the Monte Carlo method, the intrusive polynomial chaos method, and the non-intrusive polynomial chaos method.

3.1 Monte Carlo

The first method used to solve the stochastic heat equation is the Monte Carlo method. This method involves sampling a large number of values of ξ and substituting them into Eq. (4). Once a ξ value is substituted, the stochastic heat equation becomes deterministic. Hence for each sampled ξ, the heat equation is solved, and the temperature is evaluated at a specific point (x_0, t_0). Once a large number of ξs are sampled and the resulting values for $T(x_0, t_0)$ are calculated, a histogram of the $T(x_0, t_0)$ is created. As the number of samples increases, the histogram of $T(x_0, t_0)$ approaches the probability density function of $T(x_0, t_0)$. Then the mean and standard deviation of $T(x_0, t_0)$ are calculated to characterize the distribution [4].

The Monte Carlo method is used throughout UQ as a standard to compare other methods against. Because the Monte Carlo method simply involves sampling random variables and solving the PDE numerically, it provides a reliable solution, because the numerical methods used to solve the PDE are chosen so that the errors due to spatial and temporal discretization are negligible. Although the solutions to the PDEs are reliable, an infinite number of ξs must be sampled in order to obtain the true pdf of the temperature at a point. However, a large number of samples will lead to a mean and standard deviation against which other methods can be compared. Although the Monte Carlo method provides reliable solutions, it is computationally expensive because it relies on solving a large number of partial differential equations. Hence, less computationally expensive methods are desirable.

3.2 Intrusive Polynomial Chaos

The uncertainty in the thermal diffusivity is approximated using a $P+1$ term Fourier expansion

$$\alpha(\xi) = \sum_{p=0}^{P} \alpha_p \phi_p(\xi), \tag{8}$$

where α_p is a set of coefficients and $\phi_p(\xi)$ is a set of known basis functions. As ξ varies, the value of $\alpha(\xi)$ varies, that causes the output T to vary. Hence, the output temperature is also approximated with a Fourier expansion

$$T(x,t,\xi) = \sum_{p=0}^{P} T_p(x,t)\phi_p(\xi), \tag{9}$$

where $T_p(x,t)$ is a set of amplitudes that are a function of space and time. It is worth noting that the term amplitude is not a traditional amplitude that usually connotes the difference between the maximum and minimum of a function. With this approximation, a separation of variables technique can be performed so that the uncertainty in ξ is contained in the basis functions and the spatial and temporal dependence is contained in the amplitudes.

The detailed derivation of this method is included in Appendix A. The basis functions are chosen to be the probabilists' Hermite polynomials. The Galerkin method was used to take the inner product of both sides of Eq. (8) to calculate the coefficients. Then Eqs. (8-9) are substituted into Eq. (4). Subsequently the amplitudes of the boundary conditions are found and a system of coupled PDEs is created to solve for the amplitudes.

The two-term formulation ($P=1$) was performed. Evaluating the inner products leads to the following coupled PDEs:

$$T_0(0,t) = 300, \qquad T_1(0,t) = 300, \qquad T_0(1,t) = 300, \qquad T_1(1,t) = 300, \tag{10}$$

$$T_0(x,0) = 300 + 300\sin(\pi x), \qquad T_1(x,0) = 0, \tag{11}$$

$$\frac{\partial T_0(x,t)}{\partial t} = \mu \frac{\partial^2 T_0(x,t)}{\partial x^2} + \sigma \frac{\partial^2 T_1(x,t)}{\partial x^2}, \tag{12}$$

$$\frac{\partial T_1(x,t)}{\partial t} = \sigma \frac{\partial^2 T_0(x,t)}{\partial x^2} + \mu \frac{\partial^2 T_1(x,t)}{\partial x^2}. \tag{13}$$

The system of PDEs is solved numerically in Mathematica for the amplitudes $T_0(x,t)$ and $T_1(x,t)$, but these equations may have an analytic solution. Then $T_0(x,t)$ and $T_1(x,t)$ are substituted into Eq. (9) to give the solution

$$T(x,t,\xi) = T_0(x,t) + T_1(x,t)\xi. \tag{14}$$

In a similar manner to the Monte Carlo method, once the expression for $T(x,t,\xi)$ is found, a large number of ξ values are sampled and substituted into it. Then the results at a specific x_0 and t_0 could be calculated and statistically analyzed. The term $T(x_0,t_0,\xi)$ is a polynomial of ξ because the basis functions are Hermite polynomials. Although both the intrusive polynomial chaos and the Monte Carlo methods sample a large number of ξs, the polynomial chaos method involves calculating the values of polynomials for each ξ, and the Monte Carlo method involves solving a PDE for each ξ value. Also, in the polynomial chaos method the PDEs are solved once, which is advantageous because solving the PDEs is the most computationally expensive part of the method. Hence, one would expect that the intrusive polynomial chaos method is a more efficient method of quantifying the uncertainty [4]. This method is called an intrusive method, because it only requires the formulation of a stochastic version of the original model and not multiple solutions of the original model [1]. However, this method only requires one to solve $P+1$ PDEs, so one would expect that it is less computationally expensive than the Monte Carlo method.

3.3 Non-Intrusive Polynomial Chaos

Non-intrusive polynomial chaos is another method that manipulates stochastic problems to solve them more efficiently than the Monte Carlo method. Like in the intrusive polynomial chaos method, the temperature is approximated by a set of linearly independent basis functions and amplitudes to separate the random variable from the deterministic variables:

$$T(x,t,\xi) = \sum_{p=0}^{P} T_p(x,t)\phi_p(\xi). \tag{15}$$

The details of this method are included in Appendix B. Again, the basis functions are chosen to be the probabilists' Hermite polynomials. The Galerkin procedure is used to take the inner product of both sides of Eq. (15). In the end, an expression for the amplitudes is determined:

$$T_p(x,t) = \frac{\int_{-\infty}^{\infty} T(x,t,\xi)\phi_p(\xi)w(\xi)d\xi}{p!} \approx \frac{\sum_{i=0}^{I} W_i T(x,t,\xi_i)\phi_p(\xi_i)w(\xi_i)}{p!}. \tag{16}$$

Gaussian quadrature is used to approximate the integral. The values of $T(x,t,\xi_i)$ are evaluated for all Gauss points ξ_i values by substituting the values into Eq. (4), the stochastic heat equation, and it can be solved deterministically. Hence, for a specific number of terms, P, the heat equation is solved deterministically for each of I ξ_i values and $T_p(x,t)$ is formed by performing a weighted summation. Hence, through Gaussian quadrature, the integral is approximated, and the resulting amplitudes are multiplied by the bases to approximate the temperature distribution like in Eq. (15). Once a specific point in space

and time, (x_0, t_0) is chosen, the resulting approximation is simply a polynomial function of ξ. Like in the intrusive polynomial chaos method, once the polynomial approximation is created, a large number of ξ values are sampled and the resulting temperature values can be statistically analyzed. This method is non-intrusive polynomial chaos, because the solution is dependent on solving the original PDE multiple times for each ξ_i value for Gaussian quadrature [1].

4 Results

In this study, the initial conditions, boundary conditions, and thermal diffusivity approximation were arbitrarily defined. The initial conditions in Eqs. (5-7) were arbitrarily chosen. Also, the thermal diffusivity as defined in Eq. (3) was arbitrarily selected with $\mu = 1$, $\sigma = 0.75$, and ξ as a random variable from a Gaussian distribution with a mean of 0 and a standard deviation of 1. The ratio of standard deviation to mean for α was selected to be 0.75 so there would be a large range of α values. In this formulation, it was assumed that α is described by a Gaussian distribution, but when modeling a real material, one would characterize its heterogeneity to determine the correct distribution to model the material. However, because the Gaussian distribution is fully described by the mean and standard deviation, the heterogeneity in a material could be approximated by a Gaussian by calculating the mean and standard deviation of the thermal diffusivity of the material.

The solution was studied by calculating the temperature at a specific point in space and time. At a specific spatial and temporal point, the temperature is simply a function of ξ. In each method, a large number of ξ values were sampled and the resulting solutions were plotted in a histogram. The mean and standard deviation of the solutions at the specific point were calculated in order to characterize the solution. From these properties, an approximating Gaussian could be created. If this approximation has a large amount of error, other distributions could also be used to approximate the solution by calculating the higher moments of the solution. After an approximate pdf is created, a confidence interval is calculated.

4.1 Monte Carlo Method

It was arbitrarily chosen that the temperature would be analyzed at the point $x_0 = 0.3$, $t_0 = 0.3$. However, because the Monte Carlo method involves sampling a large number of ξ values, there is a chance that a negative thermal diffusivity value is selected. Specifically, when $\xi < -4/3$ is chosen, $\alpha < 0$. As previously stated, a negative thermal diffusivity is both physically impossible and mathematically unstable. When such a value is chosen, a non-physical temperature value is also predicted. For example, with a Gaussian thermal diffusivity, 100 ξ values were sampled and the Monte Carlo method was used to solve the heat equation.

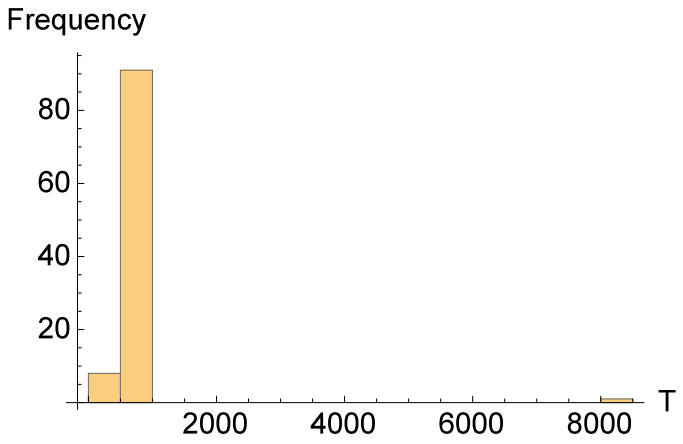

Figure 1: The Monte Carlo method was used to solve the heat equation. The solution at $T(x,t)$ was sampled at $x_0 = 0.3$, $t_0 = 0.3$ for 100 trials with a Gaussian α and it shows the non-physical results of negative thermal diffusivity values.

The resulting histogram in Fig. 1 shows that while the majority of the α values produced physical results, one negative thermal diffusivity causes the non-physical results that greatly skew the results. A temperature of 8000 at a point is non-physical when the maximum initial temperature in the domain was 600. The 100 sample ξ values chosen for this trial were relatively small, so solving the same problem with even more samples would cause more non-physical results. One possible remediation would be to change the values of μ_α and σ_α so that it would be less likely that a negative thermal diffusivity was chosen. This method could work, but because a large sample size is needed to correctly model the pdf of the solution, there is still a chance that a non-physical value of α would be realized. Another remedy is to approximate the thermal diffusivity with a distribution that guarantees a positive value. Because the distribution representing α is arbitrary, a positive Gaussian distribution could still approximate the heterogeneity in a real material. The distribution of the heterogeneity of a physical material is guaranteed to be positive because the thermal diffusivity of a material is always positive, so the issue of non-physical results would not occur with a real material. To this end, the Monte Carlo method was solved with a log-normal distribution so that

$$\alpha = e^{\mu_\alpha + \sigma_\alpha \xi}. \tag{17}$$

Using a log-normal distribution ensures that the thermal diffusivity is always positive.

In order to get physical results, the thermal diffusivity was modeled with a log-normal distribution as in Eq. (17) with $\mu_\alpha = 1$, $\sigma_\alpha = 0.75$ and ξ modeled by a Gaussian with mean 0 and standard deviation 1 as before. In order to evaluate the solution well before it is fully relaxed to a steady state, the temperature was evaluated at $x_0 = 0.3$ and $t_0 = 0.01$. Using the Monte Carlo method, the stochastic heat equation was

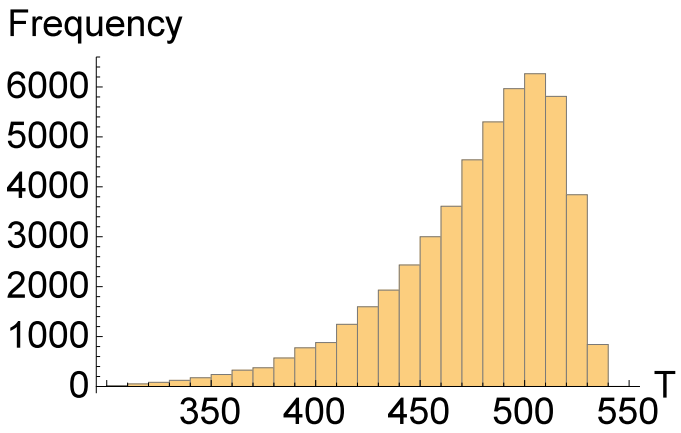

Figure 2: The Monte Carlo method was used to solve the heat equation. The solution at $T(x,t)$ was sampled at $x_0 = 0.3$, $t_0 = 0.01$ for $50,000$ trials with a log-normal α and it shows that the resulting pdf also resembles a log-normal distribution.

solved for $50,000$ random ξ values. The resulting histogram in Fig. 2 shows that the resulting temperature pdf resembles a log-normal distribution. The mean of these results is $\mu = 477$ and the standard deviation is $\sigma = 40$. There is 95% confidence that $397 \leq T(0.3, 0.01, \xi) \leq 557$. The total run time for the Monte Carlo method was 4.27 hours.

A convergence study was also conducted on the Monte Carlo method by measuring the solution for varying N, the number of ξ values sampled. Then the solution for $50,000$ ξ values was treated as the exact solution and the percentage of error of each solution was calculated. Figure 3 shows the results of the convergence study for the mean and standard deviation. As the number of samples increases, the mean and standard deviation both converge. The slope of the log-log plot of the mean is -0.57 and the slope of the log-log plot of the standard deviation is -0.58. Also, Figure 4 shows that as the number of samples increases, the computational time increases. The slope of the log-log plot is about 1, so the computational time varies linearly with the number of samples. This result was expected because the program should solve each PDE in the same amount of time and increasing the number of samples should not affect the time of to solve the PDE once. Hence, increasing the number of terms further would cause additional increase in the computation time. In conclusion, the mean and standard deviation of the temperature converge with increasing the number of samples and the computational time required increases with the number of samples.

4.2 Intrusive Polynomial Chaos

The intrusive polynomial chaos method was used to solve the heat equation with Gaussian thermal diffusivity. Although the Monte Carlo method is unable to solve the problem with negative α values, intrusive polynomial

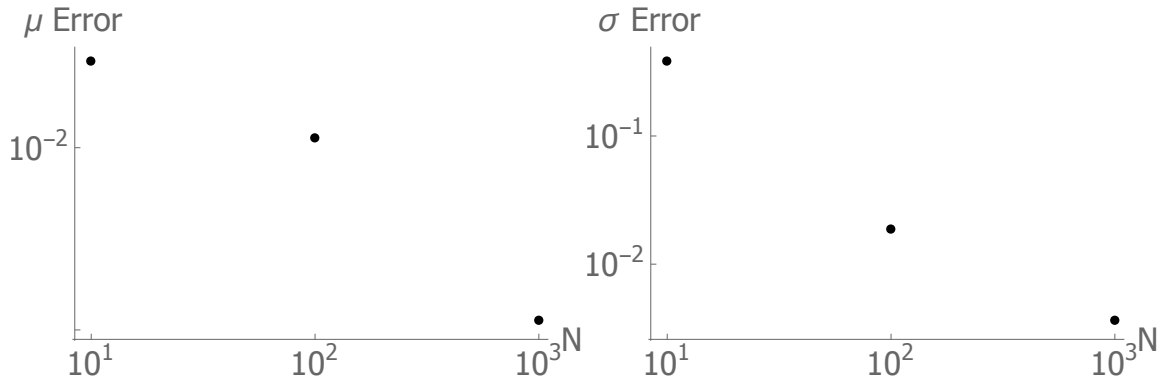

Figure 3: The convergence study shows that the slope of the best fit lines of the mean and standard deviation are -0.57 and -0.58 respectively.

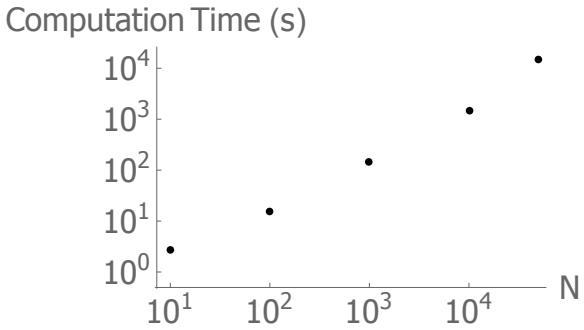

Figure 4: The Monte Carlo convergence study shows that increasing the number of ξ values sampled causes an exponential increase in the computational time. The slope is approximately unity.

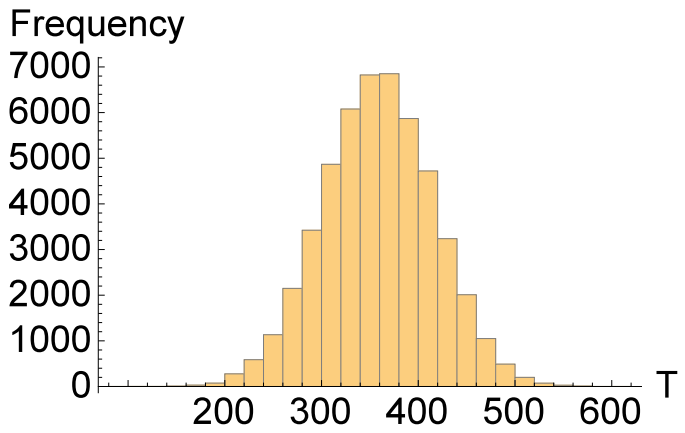

Figure 5: Intrusive polynomial chaos with 2 terms was used to solve the heat equation. The solution at $T(x,t)$ was sampled at $x_0 = 0.3$, $t_0 = 0.3$ for $50,000$ trials with a Gaussian α and it shows that the resulting pdf also resembles a normal distribution.

chaos has no issues. The Gaussian thermal diffusivity must be used for this method because α is represented by a series of polynomial basis functions as in Eq. (8). Because the Gaussian representation of α is also a polynomial, the series in Eq. (8) has a finite number of terms. However, if a log-normal distribution were used, an infinite number of terms in the series would be required to approximate the thermal diffusivity. Unless a large number of terms were used, the solution to the stochastic heat equation would not be accurate. Hence, an α that can be approximated by a finite number of Hermite polynomials is desirable. However, because the representation of the thermal diffusivity is different in the intrusive polynomial chaos method than in the Monte Carlo method, there is no solution to which to compare the results.

A Gaussian thermal diffusivity was used with a mean of 1 and a standard deviation of 0.75 was used to solve the stochastic heat equation using intrusive polynomial chaos. The arbitrary point $x_0 = 0.3$ and $t_0 = 0.3$ was chosen so that the temperature had not reached steady state. A two term approximation of the solution was used. Once the amplitudes were evaluated as described previously, $50,000$ values of ξ were sampled and substituted into the resulting polynomial determine the temperature at the (x_0, t_0). Figure 5 shows the resulting histogram form the trials. The pdf resembles a normal distribution as expected. The mean and standard deviation were 359 and 57 respectively. There is 95% confidence that $244 \leq T(0.3, 0.3, \xi) \leq 473$. The total run time was 2.28 seconds.

4.3 Non-Intrusive Polynomial Chaos

Because the non-intrusive polynomial chaos method relies upon solving the stochastic heat equation deterministically for certain ξ_i values, the thermal diffusivity could be approximated with a Gaussian distribution

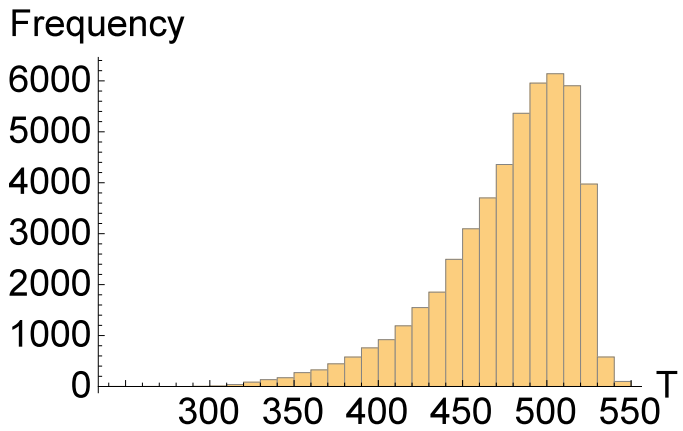

Figure 6: Non-intrusive polynomial chaos with 10 terms was used to solve the heat equation. The solution at $T(x,t)$ was sampled at $x_0 = 0.3$, $t_0 = 0.01$ for $50,000$ trials with a log-normal α and it shows that the resulting pdf also resembles a log-normal distribution.

assuming the $\xi_i \not< -4/3$. However, using a log-normal distribution lifts these restrictions and ensures that the results are directly comparable to the Monte-Carlo results. Hence, the log-normal distribution as described in Eq. (17) was used with $\mu_\alpha = 1$, $\sigma_\alpha = 0.75$, and ξ as a random variable from at Gaussian distribution with a mean of 0 and a standard deviation of 1. The temperature was evaluated at $x_0 = 0.3$ and $t_0 = 0.01$. The solutions were calculated with $I = P$. $50,000$ ξ values were sampled and a histogram of the resulting temperature for a ten-term approximation is shown in Figure 6. The resulting pdf appears to be a log-normal distribution. It has a mean of 476.7, a standard deviation of 40, and 95% confidence that $397 \leq T(0.3, 0.01, \xi) \leq 557$. The run time was 1.54 seconds.

The convergence of the non-intrusive polynomial chaos method was studied with increasing the number of terms in the approximation. Figure 7 shows that both the mean and standard deviation converge as the number of terms in the approximation is increased. The error of the solutions relative to the ten term solution. Figure 8 shows that the slope of the error of the mean is -0.41 and the slope of the error of the standard deviation is -0.33. The mean has less than 0.5% error after five terms and the standard deviation has less than 0.5% error after nine terms. Hence the mean converges more quickly than the standard deviation. One would expect the mean to converge more quickly because the higher moments of a distribution, like the standard deviation, are finer details that require more resolution to resolve.

5 Discussion and Conclusions

This study has shown that the Monte Carlo method does provide a solution to the stochastic heat equation. Because the numerical methods used to solve the deterministic PDE are chosen so that errors due to spatial

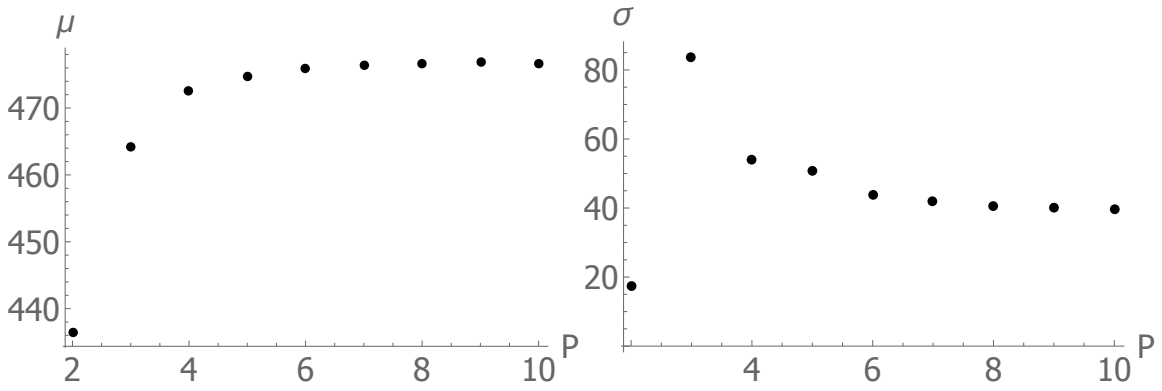

Figure 7: The convergence study shows that the mean converges more quickly than the standard deviation when increasing the number of terms in the approximation

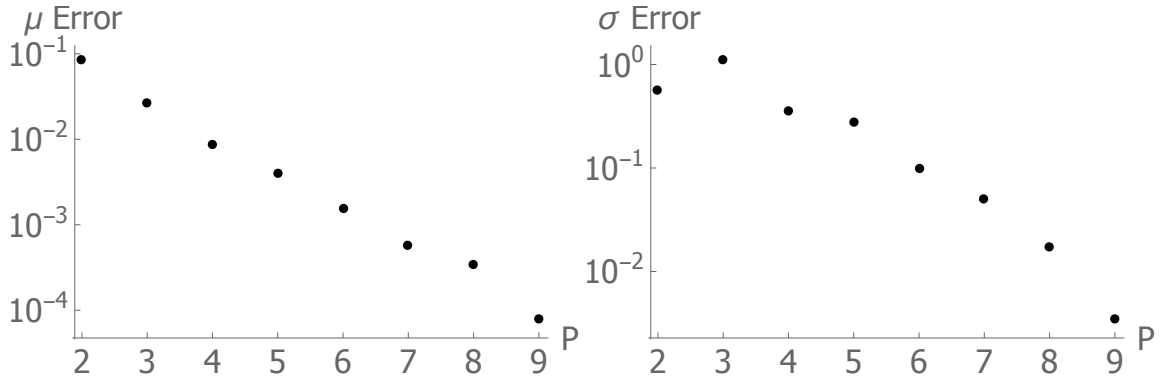

Figure 8: The slope of the log linear plot of the mean and standard deviation errors are -0.41 and -0.33 respectively.

and temporal discretization are negligible, the solutions can be trusted and used for comparison. However, in order to get solutions for $50,000$ samples of ξ, the Monte Carlo method took 4.27 hours. In order to obtain more accurate solutions with the method, additional samples must be taken that further increases the computational expense. However, the Monte Carlo method could be sped up by using parallel programming. Because the method is based on solving a large number of computationally intense problems, and then combining the results at the end, it lends well to parallel programming because separate processors could sample ξ values and solve the problem independently and then combine the results in a histogram at the end. Hence, if $1,000$ processors were used, each processor would only have to solve 50 PDEs to reach $50,000$ total samples, that would greatly reduce the wall time used.

The thermal diffusivity was modeled with a log-normal distribution when using the Monte Carlo and non-intrusive polynomial chaos methods in order to ensure positive values for α. If the heterogeneity of a physical material were characterized, the thermal diffusivity would always be positive. Therefore, all sampled thermal diffusivity values would be positive, so regardless of the type of distribution the values would not cause mathematical issues in the model.

The intrusive polynomial chaos method provided reasonable solutions, but the solution could not be compared to the Monte Carlo method because the Monte Carlo method was unable to solve the heat equation with a Gaussian thermal diffusivity. However, although the solution could not be confirmed, it could be used to model a physical material, because a physical material could have a Gaussian distribution. If a material were characterized as a Gaussian, this method could be compared to the Monte Carlo method and validated by experiment. Intrusive polynomial chaos also showed large performance improvements relative to the Monte Carlo method. Without any parallelization, the intrusive polynomial chaos method took only a matter of seconds as opposed to the hours the Monte Carlo method took.

The non-intrusive polynomial chaos method provided a solution that could be compared with the Monte Carlo method. From a visual comparison of Figs. 2 and 6, it is clear that the methods produced similar solutions. The non-intrusive polynomial chaos mean was 0.04% different from the Monte Carlo method. The standard deviations were 0.4% different. The larger difference in standard deviations is due to the slower convergence of the higher moments. The small difference in the solutions from these methods verifies the polynomial chaos method. One may also notice that it is possible for the non-intrusive polynomial chaos method to produce non-physical results. For example, Fig. 6 shows temperatures lower than 300, which is not possible for the given boundary conditions. These cases can occur for ξ values far from the mean. However, because the log-normal distribution has compact support, the frequency of cases at the extremes are of the pdf are 0 [3]. Hence, the non-physical solutions do not interfere with the accuracy of the model. Also, non-intrusive polynomial chaos produced similar results to the Monte Carlo method in much shorter

time.

Overall, all three methods are able to solve the stochastic heat equation. The Monte Carlo method produces reliable results but takes a large amount of time. Also, it cannot solve the problem with a Gaussian thermal diffusivity. The intrusive polynomial chaos method provides a solution for a Gaussian α, but could not be compared with the Monte Carlo results. The non-intrusive polynomial chaos solution converges to the same solution as the Monte Carlo method. Hence it can be trusted.

6 Future Work

There are multiple avenues of future works after this study. First, the study could be refined by studying additional methods or increasing the complexity of the problem. The polynomial chaos method could be used with Laguerre basis functions instead of the Hermite polynomials, because the Laguerre basis functions dwell on the domain of $(0, \infty)$, which is the domain of α. Also, the heat equation could be modeled in two or three dimensions to add complexity to the model.

Another area of additional study could be modeling the material heterogeneity as spatially. This model assumes that the material has uniform thermal diffusivity. However, in physical materials, this is not true. Modeling the spatial heterogeneity would increase the accuracy of the model, but would require a different approach than polynomial chaos models in this study [5].

Finally, the models in this study could be validated by experiments. First, a heterogeneous material must be characterized. This process would involve fitting a pdf to experimental data. From the heterogeneity, the thermal diffusivity could be calculated and input into the model. Then the boundary conditions for the experiment would be input into the model. Next the results from the numerical models would be compared with the experimental results to validate that the physics is accurately modeled. Until the method is validated, it simply represents a mathematical method to solve a problem, but the problem being solved is arbitrary and may not represent actual physics. In summary, potential avenues of future research are increasing the detail in the model, modeling the spatial heterogeneity in the material, and validating the model with experiments.

References

[1] G. Iaccarino, M. Eldred, A. Doostan, and O. Ghattas. Introduction to uncertainty quantification. *Mechanical Engineering and Institute for Computational Mathematical Engineering, Stanford University (course)*, 2011. See also URL http://web.stanford.edu/ jops/UQsiam09_files/iaccarino.pdf.

[2] Numerical solution of partial differential equations. `http://reference.wolfram.com/language/tutorial/NDSolvePDE.html`. Accessed: 2015-05-25.

[3] C.W. Gardiner. *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences.* Springer-Verlag, Berlin, 2nd edition, 1985.

[4] J.M. Powers and M. Sen. *Mathematical Methods in Engineering.* Cambridge University Press, New York, NY, 2015.

[5] M.R. Baer, D.K. Gartling, and P.E. DesJardin. Probabilistic models for reactive behaviour in heterogeneous condensed phase media. *Combustion Theory and Modelling*, 16(1):75–106, 2012.

[6] S. Hosder, R.W. Walters, and R. Perez. A non-intrusive polynomial chaos method for uncertainty propagation in cfd simulations. 44th AIAA Aerospace Sciences Meeting and Exhibit, AIAA-2006-0891, January 2006. Reno, Nevada.

[7] S. Paolucci. Numerical methods. University of Notre Dame AME 60614, Lecture Notes, 2011.

[8] J. Schiller, Srinivasan R.A., and Spiegel M.R. *Schaum's Outline of Probability and Statistics.* McGraw-Hill, New York, NY, 4th edition, 2013.

7 Appendix A

The derivation of the intrusive polynomial chaos method is continued from Eq. (9). First, the basis functions are chosen to be orthogonal so that

$$\langle \phi_p(\xi), \phi_m(\xi) \rangle = 0, \qquad p \neq m. \tag{18}$$

The inner product is defined to be the Gaussian inner product

$$\langle \phi_p(\xi), \phi_m(\xi) \rangle = \int_{-\infty}^{\infty} \phi_p(\xi)\phi_m(\xi)w(\xi)d\xi, \tag{19}$$

where $w(\xi)$ is a weighting function. For the Gaussian inner product the weighting function is

$$w(\xi) = \frac{1}{\sqrt{2\pi}} e^{-\xi^2/2} d\xi. \tag{20}$$

The basis functions are chosen to be the probabilists' Hermite polynomials because ξ dwells on the domain from negative to positive infinity and they are orthogonal with respect to the Gaussian inner product. The

probabilists' Hermite polynomials are

$$
\begin{aligned}
He_0(\xi) &= 1, && (21)\\
He_1(\xi) &= \xi, && (22)\\
He_2(\xi) &= \xi^2 - 1, && (23)\\
He_3(\xi) &= \xi^3 - 3\xi, && (24)\\
&\vdots && (25)\\
He_p(\xi) &= (-1)^p e^{\xi^2/2} \frac{d^p e^{-\xi^2/2}}{d\xi^p}. && (26)
\end{aligned}
$$

Hence, the inner product between two basis functions is

$$
\begin{aligned}
\langle \phi_p(\xi), \phi_m(\xi) \rangle &= \int_{-\infty}^{\infty} He_p(\xi) He_m(\xi) \frac{1}{\sqrt{2\pi}} e^{-\xi^2/2} d\xi, && (27)\\
&= p! \delta_{pm}, && (28)
\end{aligned}
$$

where δ_{pm} is the Kronecker delta [4].

In order to calculate the coefficients α_p, the Galerkin method is performed by taking the inner product of both sides of Eq. (8) with respect to the basis functions.

$$
\begin{aligned}
\alpha(\xi) &= \sum_{p=0}^{P} \alpha_p \phi_p(\xi), && (29)\\
\langle \phi_m(\xi), \alpha(\xi) \rangle &= \langle \phi_m(\xi), \sum_{p=0}^{P} \alpha_p \phi_p(\xi) \rangle \qquad m = 0, 1, .., P, && (30)\\
&= \sum_{p=0}^{P} \alpha_p \langle \phi_m(\xi), \phi_p(\xi) \rangle \qquad m = 0, 1, .., P, && (31)\\
&= \sum_{p=0}^{P} \alpha_p p! \delta_{pm} \qquad m = 1, 2, .., P, && (32)\\
&= m! \alpha_m \qquad m = 0, 1, .., P. && (33)
\end{aligned}
$$

Then, algebraic manipulation and substituting Eq. (3) gives

$$
\begin{aligned}
\alpha_p &= \frac{\langle \phi_p(\xi), \alpha(\xi) \rangle}{p!} \qquad p = 0, 1, .., P, && (34)\\
&= \frac{\langle \phi_p(\xi), \mu_\alpha + \sigma_\alpha \xi \rangle}{p!} \qquad p = 0, 1, .., P, && (35)\\
&= \frac{1}{\sqrt{2\pi} p!} \int_{-\infty}^{\infty} He_p(\xi)(\mu_\alpha + \sigma_\alpha \xi) e^{-\xi^2/2} d\xi \qquad p = 0, 1, .., P. && (36)
\end{aligned}
$$

Because Eq. (36) is an integral of a known function, it can be calculated analytically or numerically. However, because $\mu_\alpha + \sigma_\alpha \xi$ is a polynomial and can be rewritten:

$$\alpha(\xi) = \mu_\alpha + \sigma_\alpha \xi = \mu_\alpha(1) + \sigma_\alpha(\xi) = \mu_\alpha He_0(\xi) + \sigma_\alpha He_1(\xi), \tag{37}$$

because $He_0(\xi) = 1$ and $He_1(\xi) = \xi$. It is clear that $\alpha(\xi)$ is orthogonal to all Hermite basis functions with $n > 1$, because the Hermite basis functions are orthogonal. It is also clear from Eq. (37) that $\alpha_0 = \mu$ and $\alpha_1 = \sigma$. Hence, the integral did not have to be calculated for this case and actually could have been determined by inspection [4].

Next, Eqs. (9,8) are substituted into the Eq. (4):

$$\frac{\partial}{\partial t}\left(\sum_{p=0}^{P} T_p(x,t)\phi_p(\xi)\right) = \left(\sum_{p=0}^{P} \alpha_p \phi_p(\xi)\right) \frac{\partial^2}{\partial x^2}\left(\sum_{m=0}^{P} T_m(x,t)\phi_m(\xi)\right). \tag{38}$$

The boundary conditions must also be approximated using a series, so

$$T(0,t,\xi) = 300 = \sum_{p=0}^{P} T_p(0,t)\phi_p(\xi), \tag{39}$$

$$T(1,t,\xi) = 300 = \sum_{p=0}^{P} T_p(1,t)\phi_p(\xi), \tag{40}$$

$$T(x,0,\xi) = 300 + 300\sin(\pi x) = \sum_{p=0}^{P} T_p(x,0)\phi_p(\xi). \tag{41}$$

The initial amplitudes, $T_p(x,t)$ could be calculated using the same process as Eq. (36), but by inspection because $\phi_0(\xi) = 1$, it is clear by inspection that

$$T_0(0,t) = 300, \qquad T_1(0,t) = 0, \qquad T_2(0,t) = 0, \qquad ..., \qquad T_P(0,t) = 0, \tag{42}$$

$$T_0(1,t) = 300, \qquad T_1(1,t) = 0, \qquad T_2(1,t) = 0, \qquad ..., \qquad T_P(1,t) = 0, \tag{43}$$

$$T_0(x,0) = 300 + 300\sin(\pi x), \qquad T_1(x,0) = 0, \qquad T_2(x,0) = 0, \qquad ..., \qquad T_P(x,0) = 0. \tag{44}$$

Now, manipulating Eq. (38) gives

$$\sum_{p=0}^{P} \frac{\partial T_p(x,t)}{\partial t}\phi_p(\xi) = \sum_{p=0}^{P}\sum_{m=0}^{P} \alpha_p \frac{\partial^2 T_m(x,t)}{\partial x^2}\phi_p(\xi)\phi_m(\xi). \tag{45}$$

Once again, the dependence on the random variable, ξ, must be removed in order to calculate the amplitudes $T_p(x,t)$ to approximate the temperature. Once again, the Galerkin method is used by taking the inner

product of both sides of Eq. (45) with respect to the set of the basis functions, $\phi_l(\xi)$:

$$\langle\phi_l(\xi),\sum_{p=0}^{P}\frac{\partial T_p(x,t)}{\partial t}\phi_p(\xi)\rangle = \langle\phi_l(\xi),\sum_{p=0}^{P}\sum_{m=0}^{P}\alpha_p\frac{\partial^2 T_m(x,t)}{\partial x^2}\phi_p(\xi)\phi_m(\xi)\rangle \qquad l=0,1,..,P, \tag{46}$$

$$\sum_{p=0}^{P}\frac{\partial T_p(x,t)}{\partial t}\langle\phi_l(\xi),\phi_p(\xi)\rangle = \sum_{p=0}^{P}\sum_{m=0}^{P}\alpha_p\frac{\partial^2 T_m(x,t)}{\partial x^2}\langle\phi_l(\xi),\phi_p(\xi)\phi_m(\xi)\rangle \qquad l=0,1,...,P, \tag{47}$$

$$\sum_{p=0}^{P}\frac{\partial T_p(x,t)}{\partial t}(l!\delta_{lp}) = \sum_{p=0}^{P}\sum_{m=0}^{P}\alpha_p\frac{\partial^2 T_m(x,t)}{\partial x^2}\langle\phi_l(\xi),\phi_p(\xi)\phi_m(\xi)\rangle \qquad l=0,1,...,P, \tag{48}$$

$$l!\frac{\partial T_l(x,t)}{\partial t} = \sum_{p=0}^{P}\sum_{m=0}^{P}\alpha_p\frac{\partial^2 T_m(x,t)}{\partial x^2}\langle\phi_l(\xi),\phi_p(\xi)\phi_m(\xi)\rangle \qquad l=0,1,...,P, \tag{49}$$

$$\frac{\partial T_l(x,t)}{\partial t} = \frac{1}{l!}\sum_{p=0}^{P}\sum_{m=0}^{P}\alpha_p\frac{\partial^2 T_m(x,t)}{\partial x^2}\langle\phi_l(\xi),\phi_p(\xi)\phi_m(\xi)\rangle \qquad l=0,1,...,P. \tag{50}$$

The dependence on ξ has been successfully removed because inner product expression is simply the integral

$$\langle\phi_l(\xi),\phi_p(\xi)\phi_m(\xi)\rangle = \int_{-\infty}^{\infty}\phi_l(\xi)\phi_p(\xi)\phi_m(\xi)\frac{1}{\sqrt{2\pi}}e^{-\xi^2/2}d\xi, \tag{51}$$

that can be calculated and tabulated. Because there is no longer dependence on ξ, Eq. (50) is a system of $P+1$ coupled partial differential equations [4]. The system can be solved using numerical schemes, like NDSolve in Mathematica.

8 Appendix B

The derivation of the non-intrusive polynomial chaos method is continued from Eq. (15). Once again, the basis functions were chosen to be the probabilists' Hermite polynomials because they are orthogonal with respect to the Gaussian weighting function. Then a Galerkin formulation is used, and the inner product of both sides of Eq. (15) is taken with respect to the basis functions. Then the orthogonality of the basis functions is used to solve for the amplitudes [6]:

$$\langle T(x,t,\xi),\phi_m(\xi)\rangle = \langle\sum_{p=0}^{P}T_p(x,t)\phi_p(\xi),\phi_m(\xi)\rangle \qquad m=1,2,..,P, \tag{52}$$

$$\langle T(x,t,\xi),\phi_m(\xi)\rangle = \sum_{p=0}^{P}T_p(x,t)\langle\phi_p(\xi),\phi_m(\xi)\rangle \qquad m=1,2,..,P, \tag{53}$$

$$\langle T(x,t,\xi),\phi_m(\xi)\rangle = \sum_{p=0}^{P}T_p(x,t)m!\delta_{pm} \qquad m=1,2,..,P, \tag{54}$$

$$\frac{\langle T(x,t,\xi),\phi_m(\xi)\rangle}{m!} = T_m(x,t) \qquad m=1,2,..,P. \tag{55}$$

Because the amplitudes $T_m(x,t)$ are simply equal to an inner product, the inner product can be expanded to show that the amplitudes are equal to integrals:

$$T_p(x,t) = \frac{\int_{-\infty}^{\infty} T(x,t,\xi)\phi_p(\xi)w(\xi)d\xi}{p!} \qquad p = 1,2,..,P, \tag{56}$$

$$w(\xi) = \frac{1}{\sqrt{2\pi}} e^{-\xi^2/2}. \tag{57}$$

In the integral in Eq. (56), $\phi_p(\xi)$ and $w(\xi)$ are known, but the integral cannot be evaluated outright because $T(x,t,\xi)$ is unknown. However, the integral can be approximated using Gaussian quadrature. Gaussian quadrature is a method of approximating an integral as a series of weights, W_i and function values at specific points, ξ_i:

$$\int_{-\infty}^{\infty} f(\xi)d\xi \approx \sum_{i=0}^{I} W_i f(\xi_i). \tag{58}$$

For the $I+1$ term approximation of the integral, the values for W_i and ξ_i are taken from Appendix C. For this specific problem, the Gaussian quadrature approximation of the integral is

$$T_p(x,t) = \frac{\int_{-\infty}^{\infty} T(x,t,\xi)\phi_p(\xi)w(\xi)d\xi}{p!} \approx \frac{\sum_{i=0}^{I} W_i T(x,t,\xi_i)\phi_p(\xi_i)w(\xi_i)}{p!}. \tag{59}$$

9 Appendix C

Gaussian quadrature is a method to approximate an integral as a sum, similar to the trapezoidal rule. In general the quadrature formula is

$$I = \int_a^b f(\xi)d\xi \approx \sum_{i=1}^{I} W_i f(\xi_i), \tag{60}$$

where $f(\xi)$ is a general function, W_i are weights, and ξ_i are specific abscissas. The weights and abscissas are chosen based on the method. For the trapezoidal rule these values are chosen so that a line is integrated between the two adjacent points in the approximation. However, in Gaussian quadrature, the W_i and ξ_i values are chosen for the most accurate approximation for a specific number of terms. On an infinite interval, $(a,b) \to (-\infty,\infty)$, the abscissa values are chosen to be the zeros of the nth physicists' Hermite polynomial. The physicists' Hermite polynomials are different than the probabilists' Hermite polynomials discussed in

this report:

$$H_0(\xi) = 1, \tag{61}$$

$$H_1(\xi) = 2\xi, \tag{62}$$

$$H_2(\xi) = 4\xi^2 - 2, \tag{63}$$

$$H_3(\xi) = 8\xi^3 - 12\xi, \tag{64}$$

$$\vdots \tag{65}$$

$$H_n(\xi) = (-1)^n e^{\xi^2} \frac{d^n e^{-\xi^2}}{d\xi^n}. \tag{66}$$

However, with a small effort, one could also formulate the Gaussian quadrature for the probabilists' Hermite polynomials. The weights for an infinite domain are

$$W_i = \frac{2^{n-1} n! \sqrt{\pi}}{n^2 [H_{n-1}(\xi_i)]^2}. \tag{67}$$

The abscissas and weights for Gaussian quadrature for up to $I = 10$ are tabulated in Tables 1-9 [7].

Table 1: Gaussian quadrature ξ_i and W_i values for $I = 2$

$\pm\xi_i$	W_i
0.70710 67811 86548	1.46114 11826 611

Table 2: Gaussian quadrature ξ_i and W_i values for $I = 3$

$\pm\xi_i$	W_i
0.00000 00000 00000	1.18163 59006 037
1.22474 48713 31589	1.32393 11752 136

Table 3: Gaussian quadrature ξ_i and W_i values for $I = 4$

$\pm\xi_i$	W_i
0.52464 76232 75290	1.05996 44828 950
1.65068 01238 85785	1.24022 58176 958

Table 4: Gaussian quadrature ξ_i and W_i values for $I = 5$

$\pm\xi_i$	W_i
0.00000 00000 00000	0.94530 87204 829
0.95857 24646 13819	0.98658 09967 514
2.020184 28704 56086	1.18148 86255 360

Table 5: Gaussian quadrature ξ_i and W_i values for $I = 6$

$\pm\xi_i$	W_i
0.43607 74119 27617	0.87640 13344 362
1.33584 90740 13697	0.93558 05576 312
2.35060 49736 74492	1.13690 83326 745

Table 6: Gaussian quadrature ξ_i and W_i values for $I = 7$

$\pm\xi_i$	W_i
0.00000 00000 00000	0.81026 46175 568
0.81628 78828 58965	0.82868 73032 836
1.67355 16287 67471	0.89718 46002 252
2.65196 13568 35233	1.10133 07296 103

Table 7: Gaussian quadrature ξ_i and W_i values for $I = 8$

$\pm\xi_i$	W_i
0.38118 69902 07322	0.76454 41286 517
1.15719 37124 46780	0.79289 00483 864
1.98165 67566 95843	0.86675 26065 634
2.93063 74202 57244	1.07193 01442 480

Table 8: Gaussian quadrature ξ_i and W_i values for $I = 9$

$\pm\xi_i$	W_i
0.00000 00000 00000	0.72023 52156 061
0.72355 10187 52838	0.73030 24527 451
1.46855 32892 16668	0.76460 81250 946
2.26658 05845 31843	0.84175 27014 787
3.19099 32017 81528	1.04700 35809 767

Table 9: Gaussian quadrature ξ_i and W_i values for $I = 10$

$\pm\xi_i$	W_i
0.34290 13272 23705	0.68708 18539 513
1.03661 08297 89514	0.70329 63231 049
1.75668 36492 99882	0.74144 19319 436
2.53273 16742 32790	0.82066 61264 048
3.43615 91188 37738	1.02545 16913 657

10 Appendix D

Here is an overview of some basic concepts of probability theory.

10.1 Introduction to Probability

10.1.1 Random Variables

A random variable is a function of all of the points in a sample space. Random variables are usually denoted with capital letters, like X. A random variable can also be thought of as a method of assigning values to outcomes of a random process.

Example 1 Define a random variable, X, that is the value of the roll of a die.

Solution:

$$X = \begin{cases} 1, & \text{if the value of the roll is 1} \\ 2, & \text{if the value of the roll is 2} \\ 3, & \text{if the value of the roll is 3} \\ 4, & \text{if the value of the roll is 4} \\ 5, & \text{if the value of the roll is 5} \\ 6, & \text{if the value of the roll is 6.} \end{cases}$$

Another way to write the solution is:

$$X = \begin{cases} x, \text{where } x \text{ is the value of the roll.} \end{cases}$$

Random variables are not used like traditional variables because they do not have a fixed value. Hence the expression

$$7 = X + 2, \tag{68}$$

where X is a random variable is not meaningful, because it limits the possibilities of the random variable. Instead, random variables are described with probabilities. For example, the probability that the random variable X from Example 1 is greater than 3 could be written $P(X > 3)$. This example show cases the notational convenience of random variables because without the random variable one would have to write that probability as $P(\text{The roll of a die} > 3)$. Hence, random variables are not used in equations like algebraic variables, but are used as specific functions.

10.1.2 Discrete and Continuous Random Variables

A random variable is called a discrete random variable if it takes on a finite or countable number of values. A random variable is called a continuous random variable if it takes on an infinite number of values.

Example 2 Are the following random variables discrete or infinite?

1. X =The value of the roll of a die
2. W =The weight of an elephant

3. Y =The exact time of the winning horse in the Kentucky Derby

4. O =The number of oranges on an orange tree

5. P =The stock market value of a share of the Dow Jones

Solution:

1. X is a discrete random variable because there are only 6 possible values.

2. W is a continuous random variable because the exact weight of an elephant could be any number between $12,000$ lbf and $12,001$ lbf.

3. Y is a continuous random variable because the exact winning time could be measured in $2:03.66$ or it could be $2:03.65764$ or it could be $2:03.6576498012$.

4. O is a discrete random variable because the number of oranges on a tree is countable.

5. P is a discrete random variable because the stock market value of the Dow Jones is rounded to the nearest cent. Therefore, there are a countable number of values it can take.

10.1.3 Probability Distributions

For a discrete random variable X, let $x_1, x_2, ...x_n$ be the n possible values of X. Each of these values have probabilities

$$P(X = x_i) = f(x_i). \qquad i = 1, 2, ..., n \tag{69}$$

The probability function or probability distribution is defined as

$$P(X = x) = f(x). \tag{70}$$

It is clear that for $x = x_i$ Eq. (70) becomes Eq. (69) and for other x values, $f(x) = 0$. For a discrete random variable, the probability distribution will be a piecewise function. Two properties of discrete probability distributions are:

$$f(x) \geq 0, \tag{71}$$

$$\sum_{i=1}^{n} f(x_i) = 1. \tag{72}$$

These properties are derived from the fact that a probability is always greater than or equal to zero and the fact that all of the discrete values are included in the distribution.

For a continuous random variable, the resulting probability distribution will be continuous because there are an infinite number of possible values. Hence, the properties of a continuous probability distribution are:

$$f(x) \geq 0, \tag{73}$$

$$\int_{-\infty}^{\infty} f(x) = 1. \tag{74}$$

Example 3 What is the probability distribution of the value of the sum of two dice rolls?

Solution: When rolling two dice, there are 36 possible outcomes:

$1,1$	$1,2$	$1,3$	$1,4$	$1,5$	$1,6$
$2,1$	$2,2$	$2,3$	$2,4$	$2,5$	$2,6$
$3,1$	$3,2$	$3,3$	$3,4$	$3,5$	$3,6$
$4,1$	$4,2$	$4,3$	$4,4$	$4,5$	$4,6$
$5,1$	$5,2$	$5,3$	$5,4$	$5,5$	$5,6$
$6,1$	$6,2$	$6,3$	$6,4$	$6,5$	$6,6$

When these results are summed they become

2	3	4	5	6	7
3	4	5	6	7	8
4	5	6	7	8	9
5	6	7	8	9	10
6	7	8	9	10	11
7	8	9	10	11	12

Hence, there are 11 possible results. Let Y be the discrete random variable of the sum of the two dice rolls.

$$
\begin{aligned}
f(2) = P(Y = 2) &= \frac{1}{11} \\
f(3) = P(Y = 3) &= \frac{2}{11} \\
f(4) = P(Y = 4) &= \frac{3}{11} \\
f(5) = P(Y = 5) &= \frac{4}{11} \\
f(6) = P(Y = 6) &= \frac{5}{11} \\
f(7) = P(Y = 7) &= \frac{6}{11} \\
f(8) = P(Y = 8) &= \frac{5}{11} \\
f(9) = P(Y = 9) &= \frac{4}{11} \\
f(10) = P(Y = 10) &= \frac{3}{11} \\
f(11) = P(Y = 11) &= \frac{2}{11} \\
f(12) = P(Y = 12) &= \frac{1}{11}
\end{aligned}
$$

Hence, $f(y)$ can be plotted as shown in Figure 10.1.3.

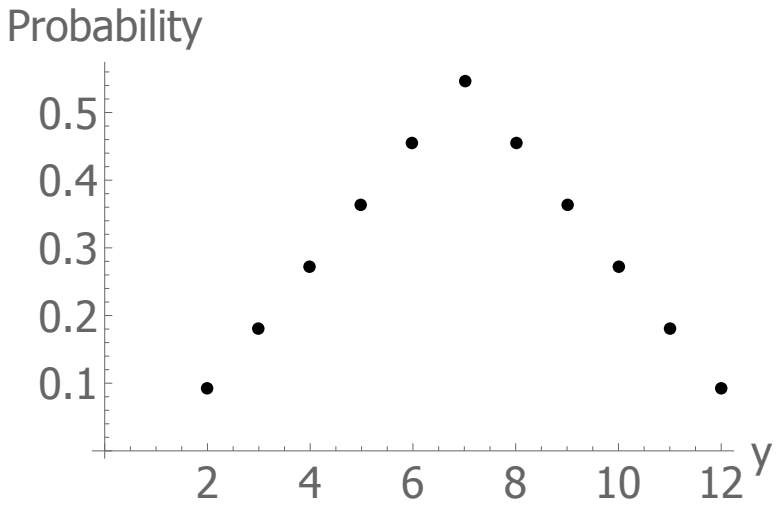

Figure 9: Example of a probability distribution.

10.2 Mathematical Expectation

10.2.1 Expected Value

The expected value or mean of a discrete random variable is

$$E(X) = \sum_{i=1}^{n} x_i f(x_i), \tag{75}$$

where n is the number of possibilities and the expected value for a continuous random variable is

$$E(X) = \int_{-\infty}^{\infty} x f(x)\, dx. \tag{76}$$

The expected value for a function of a random variable is

$$E[g(x)] = \sum_{i=1}^{n} g(x_i) f(x_i), \tag{77}$$

$$E[g(x)] = \int_{-\infty}^{\infty} g(x) f(x)\, dx. \tag{78}$$

Some theorems about expected values are

$$E(cX) = cE(X), \tag{79}$$

$$E(X+Y) = E(X) + E(Y), \tag{80}$$

$$E(XY) = E(X)E(Y). \tag{81}$$

The mean is generally denoted by μ [8].

Example 4 What is the expected value of the sum of five fair dice?

Solution: Because a die has 6 discrete values, we will use Equation 75. If U, V, W, X, and Y are the values of the dice, for one die,

$$E(U) = E(V) = E(W) = E(X) = E(Y) = 1\left(\frac{1}{6}\right) + 2\left(\frac{1}{6}\right) + 3\left(\frac{1}{6}\right) + 4\left(\frac{1}{6}\right) + 5\left(\frac{1}{6}\right) + 6\left(\frac{1}{6}\right) = \frac{7}{2}.$$

Then the expected value of the sum can be written:

$$E(U+V+W+X+Y) = E(U) + E(V) + E(W) + E(X) + E(Y) = 5\left(\frac{7}{2}\right) = \frac{35}{2}.$$

Example 5 What is the expected value in minutes of clock that stops a random time during an hour?

Solution: The probability distribution function for this problem is:

$$f(x) = \begin{cases} \frac{1}{60}, & 0 \leq x \leq 60 \\ 0, & \text{otherwise.} \end{cases}$$

Because it could stop at any fraction between the minute values, this pdf is continuous. Therefore we will use Equation (76).

$$E(X) = \int_{-\infty}^{\infty} x f(x)\, dx = \int_{0}^{60} x \left(\frac{1}{60}\right)\, dx = \left(\frac{1}{120}\right)\left(60^2 - 0\right) = 30. \tag{82}$$

10.2.2 Moments

The variance is defined as

$$\mathrm{Var}(X) = E[(X - \mu)^2] = \sigma_X^2, \tag{83}$$

where σ_X is the standard deviation.

$$\sigma^2 = \sum_{i=1}^{n} (x_i - \mu)^2 f(x_i), \tag{84}$$

$$\sigma^2 = \int_{-\infty}^{\infty} (x - \mu)^2 f(x). \tag{85}$$

The variance is a measure of the dispersion. If the values are concentrated near the mean, the variance is small.

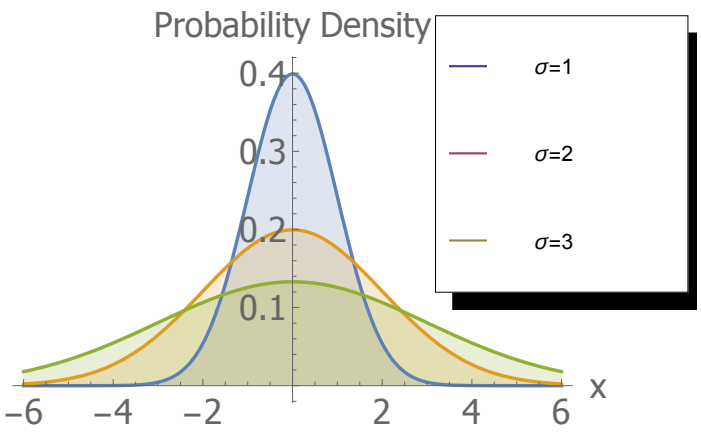

Figure 10: Example of a normal distribution with different variances.

A random variable X with mean μ and standard deviation σ has an associated standardized random variable

$$X^* = \frac{X - \mu}{\sigma}. \tag{86}$$

Standardized random variables are used for comparing different distributions.

The moments of a distribution are used to determine similarity between two distributions. The r^{th} central moment is defined

$$\mu_r = E[(X - \mu)^r], \tag{87}$$

$$\mu_r = \sum_{i=1}^{n} (x_i - \mu)^r f(x_i), \tag{88}$$

$$\mu_r = \int_{-\infty}^{\infty} (x - \mu)^r f(x)\, dx, \tag{89}$$

that connotes $\mu_0 = 1$, $\mu_1 = \mu$, and $\mu_2 = \sigma^2$, because μ_0 is the total area under the probability distribution function, μ_1 is the mean and, μ_2 is the variance. The r^{th} raw moment is

$$\mu'_r = E(X^r). \tag{90}$$

The moment generating function of X is defined by

$$M_X(t) = E\left(e^{tX}\right). \tag{91}$$

The convergent discrete form is

$$M_X(t) = \sum_{i=1}^{n} e^{tX} f(x_i), \tag{92}$$

and the continuous form is

$$M_X(t) = \int_{-\infty}^{\infty} e^{tX} f(x). \tag{93}$$

$M_X(t)$ generates moments because the rth derivative evaluated at $t = 0$ is the rth moment, μ_r'. When the moment generating function is written as a Taylor series, the coefficient of the rth term is the rth moment [8].

The characteristic function is

$$\phi_X(\omega) = M_X(i\omega) = E\left(e^{i\omega X}\right). \tag{94}$$

for characteristic functions the series and integral converge absolutely. The characteristic function can be related to the function using Fourier transforms.

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega x} \phi_X(\omega)\, d\omega. \tag{95}$$

Example 6 For a fair die, find (a) the moment generating function, (b) the first three moments.

Solution: Because it is a die, there are 6 discrete possibilities, so we will use Equation (92)

$$\begin{aligned} M_X(t) &= \sum e^{tX} f(x) = e^{t(1)}\left(\frac{1}{6}\right) + e^{t(2)}\left(\frac{1}{6}\right) + e^{t(3)}\left(\frac{1}{6}\right) + e^{t(4)}\left(\frac{1}{6}\right) + e^{t(5)}\left(\frac{1}{6}\right) + e^{t(6)}\left(\frac{1}{6}\right) \\ M_X(t) &= \frac{e^t}{6} + \frac{e^{2t}}{6} + \frac{e^{3t}}{6} + \frac{e^{4t}}{6} + \frac{e^{5t}}{6} + \frac{e^{6t}}{6}. \end{aligned}$$

Then to evaluate the moments we must find $M_X(0)$, $M_X'(0)$, $M_X''(0)$, and $M_X'''(0)$.

$$\begin{aligned} M_X'(t) &= \frac{e^t}{6} + \frac{e^{2t}}{3} + \frac{e^{3t}}{2} + \frac{2e^{4t}}{3} + \frac{5e^{5t}}{6} + e^{6t} \\ M_X''(t) &= \frac{e^t}{6} + \frac{2e^{2t}}{3} + \frac{3e^{3t}}{2} + \frac{8e^{4t}}{3} + \frac{25e^{5t}}{6} + 6e^{6t} \\ M_X'''(t) &= \frac{e^t}{6} + \frac{4e^{2t}}{3} + \frac{9e^{3t}}{2} + \frac{32e^{4t}}{3} + \frac{125e^{5t}}{6} + 36e^{6t} \\ \mu_0 = M_X(0) &= \left(\frac{1}{6}\right) + \left(\frac{1}{6}\right) + \left(\frac{1}{6}\right) + \left(\frac{1}{6}\right) + \left(\frac{1}{6}\right) + \left(\frac{1}{6}\right) = 1 \\ \mu_1 = M_X'(0) &= \left(\frac{1}{6}\right) + \left(\frac{1}{3}\right) + \left(\frac{1}{2}\right) + \left(\frac{2}{3}\right) + \left(\frac{5}{6}\right) + 1 = \frac{7}{2} \\ \mu_2 = M_X''(0) &= \left(\frac{1}{6}\right) + \left(\frac{2}{3}\right) + \left(\frac{3}{2}\right) + \left(\frac{8}{3}\right) + \left(\frac{25}{6}\right) + 6 = \frac{91}{6} \\ \mu_3 = M_X'''(0) &= \left(\frac{1}{6}\right) + \left(\frac{4}{3}\right) + \left(\frac{9}{2}\right) + \left(\frac{32}{3}\right) + \left(\frac{125}{6}\right) + 36 = \frac{473}{6} \end{aligned}$$

Note that the $\mu_0 = 1$ as expected and the first moment, μ_1 is the mean, as shown in Example 4. Now let's confirm that μ_2 is the variance. Using Equation (84) we find

$$\sigma^2 = E[X^2] = 1^2\left(\frac{1}{6}\right) + 2^2\left(\frac{1}{6}\right) + 3^2\left(\frac{1}{6}\right) + 4^2\left(\frac{1}{6}\right) + 5^2\left(\frac{1}{6}\right) + 6^2\left(\frac{1}{6}\right) = \frac{91}{6}$$

10.2.3 Chebyshev's Inequality

Chebyshev's Inequality states that if X is a random variable having mean μ and variance σ^2 that are finite then if ϵ is any positive number

$$P\left(|X - \mu| \geq \epsilon\right) \leq \frac{\sigma^2}{\epsilon^2}, \tag{96}$$

or, with $\epsilon = k\sigma$,

$$P\left(|X - \mu| \geq k\sigma\right) \leq \frac{1}{k^2}. \tag{97}$$

Chebyshev's Inequality tells us the probability that X will differ by more than k standard deviations from the mean. This is useful because we have not specified the probability distribution [8].

Example 7 For example, what is the probability that random variable X (with any distribution) differs from its mean by more than 4 standard deviations. Using Equation (97) where $k = 4$,

$$P\left(|X - \mu|\right) \geq 4\sigma) \leq \frac{1}{4^2} = 0.0625$$

10.2.4 Law of Large Numbers

As a result, the Law of Large Numbers states that if $P_n = X_1, X_2, ..., X_n$ are mutually independent random variables with a finite mean of μ and variance σ^2 then

$$\lim_{n\to\infty} P\left(\left|\frac{P_n}{n} - \mu\right| \geq \epsilon\right) = 0, \tag{98}$$

that means that the probability that the arithmetic mean of $X_1, X_2, ..., X_n$ differs from its expected value μ by more than any positive number approaches zero as $n \to \infty$. The strong law of large numbers states that $\lim_{n\to\infty}(X_1 + ... + X_n)/n = \mu$ with probability one.

Example 8 We will test the law of large numbers for a fair die. As shown in Example 4, the mean of one roll of a die is $\mu = \frac{7}{2} = 3.5$. Table 10 shows that as $n \to \infty$, $\frac{S_n}{n} \to 3.5$ as expected. Hence because, $|\frac{S_n}{n} - \mu|$ approaches 0 then the probability that that quantity is greater than any positive number approaches zero as Equation (98) states.

Table 10: The law of large numbers was tested to simulate a fair die for a range of n.

n	5	10	20	50	100	1000	10000	100000	1000000
$\frac{S_n}{n}$	2.8	4.3	3.15	3.34	3.53	3.569	3.4677	3.5034	3.4996
$\lvert\frac{S_n}{n} - \mu\rvert$	0.7	0.8	0.35	0.16	0.03	0.069	0.0323	0.0034	0.0004

10.2.5 Miscellaneous topics

The mean deviation (M.D.) is defined as

$$M.D.(X) = E[|X - \mu|] = \sum_{i=1}^{n} |x_i - \mu| f(x_i), \tag{99}$$

for discrete variables or

$$M.D.(X) = E[|X - \mu|] = \int_{-\infty}^{\infty} |x - \mu| f(x)\, dx, \tag{100}$$

for continuous variables. The mean deviation is a metric used to describe the dispersion of the pdf. Intuitively, it is the mean distance from the mean of pdf.

Skewness can be measured with coefficients of skewness, like

$$\alpha_3 = \frac{E[(X-\mu)^3]}{\sigma^3} = \frac{\mu_3}{\sigma^3}, \tag{101}$$

where σ_3 will be positive or negative depending on if the distribution is skewed right or left. $\sigma_3 = 0$ for a symmetric distribution.

Kurtosis describes how large or small the peak is. A normal curve has a coefficient of kurtosis equal to 3 [8].

$$\alpha_4 = \frac{E[(X-\mu)^4]}{\sigma^4} = \frac{\mu_4}{\sigma^4}. \tag{102}$$

Example 9 When random variable X is one of the set $12, 13, 13, 7, 6, 20, 11, 10$ calculate the mean deviation.

Solution: First we must calculate the mean

$$\mu = \frac{12 + 13 + 13 + 7 + 6 + 20 + 11 + 10}{8} = 11.5.$$

Then using Equation (99) we calculate the mean deviation to be

$$M.D.(X) = \frac{|12-11.5|+|13-11.5|+|13-11.5|+|7-11.5|+|6-11.5|+|20-11.5|+|11-11.5|}{8} + \frac{|10-11.5|}{8}$$

$M.D.(X) = 3.$

10.3 Special Probability Distributions

10.3.1 Gaussian Distribution

A binomial distribution is given by the probability density function

$$f(x) = P(X = x) = \binom{n}{x} p^x q^{n-x} = \frac{n!}{x!(n-x)!} p^x q^{n-x}, \tag{103}$$

where p is the probability of success, $q = 1 - p$ is the probability of failure, X is a random variable that denotes the number of successes of n trials, and x is the number of successes. If $n = 1$, this is called a Bernoulli distribution.

The Gaussian or normal distribution that is given the probability density function

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}, \tag{104}$$

where μ is the mean and σ is the standard deviation. The distribution function is

$$F(x) = P(X \leq x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-(v-\mu)^2/2\sigma^2} dv. \tag{105}$$

We define

$$Z = \frac{X - \mu}{\sigma}, \tag{106}$$

where Z is the standardized variable or standard score. If the mean of Z is 0 and the variance is 1, then the density function is

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \tag{107}$$

that is known as the standard normal density function. Figure 10.2.2 shows an example of a normal distribution with varying standard deviations. If

$$Z = \frac{X - np}{\sqrt{npq}}, \tag{108}$$

then the Gaussian distribution closely approximates the binomial distribution where n is the number of trials, p is the probability of success, and q is the probability of failure. The general rule is that if $np > 5$ and $qp > 5$ then it is a good approximation of the binomial distribution and as n approached infinity, the binomial distribution approaches the Gaussian distribution. Written another way,

$$\lim_{n->\infty} P\left(a \leq \frac{X - np}{\sqrt{npq}} \leq b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-u^2/2} du. \tag{109}$$

The normal distribution is often called a "bell curve" and is seen a variety of fields. For example the distribution of human heights is well approximated by a normal distribution. Table 11 shows the mean and variance of the distribution [8].

Table 11: The properties of the normal distribution as described in Equation (104).

Mean	μ
Variance	σ^2
Moment Generating Function	$M(t) = e^{ut+(\sigma^2 t^2/2)}$
Characteristic Function	$\phi(\omega) = e^{i\mu\omega-(\sigma^2\omega^2/2}$

Example 10 Show that the binomial distribution can be closely approximated by the normal distribution for increasing n for flipping a fair coin.

Solution: If a fair coin is flipped, $p = 0.5$. Figure 11 shows that as n increases, the Guassian distributions becomes more like the binomial distribution.

10.3.2 Poisson Distribution

The Poisson distribution has a pdf

$$f(x) = P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}, \qquad x = 0, 1, 2, ... \tag{110}$$

where λ is a positive constant and X is a random variable that can be a non-negative integer. The Poisson distribution approaches the normal distribution as $\lambda \to \infty$ where $(X - \lambda)/\sqrt{\lambda}$ is the standardized random variable. The Poisson distribution describes many situations like the number of mutations in a strands of DNA or the number of cars going through a traffic light. Figure 10.3.2 shows an example of a Poisson distribution for a variety of λ values. Table 12 shows the mean and variance of the distribution [8].

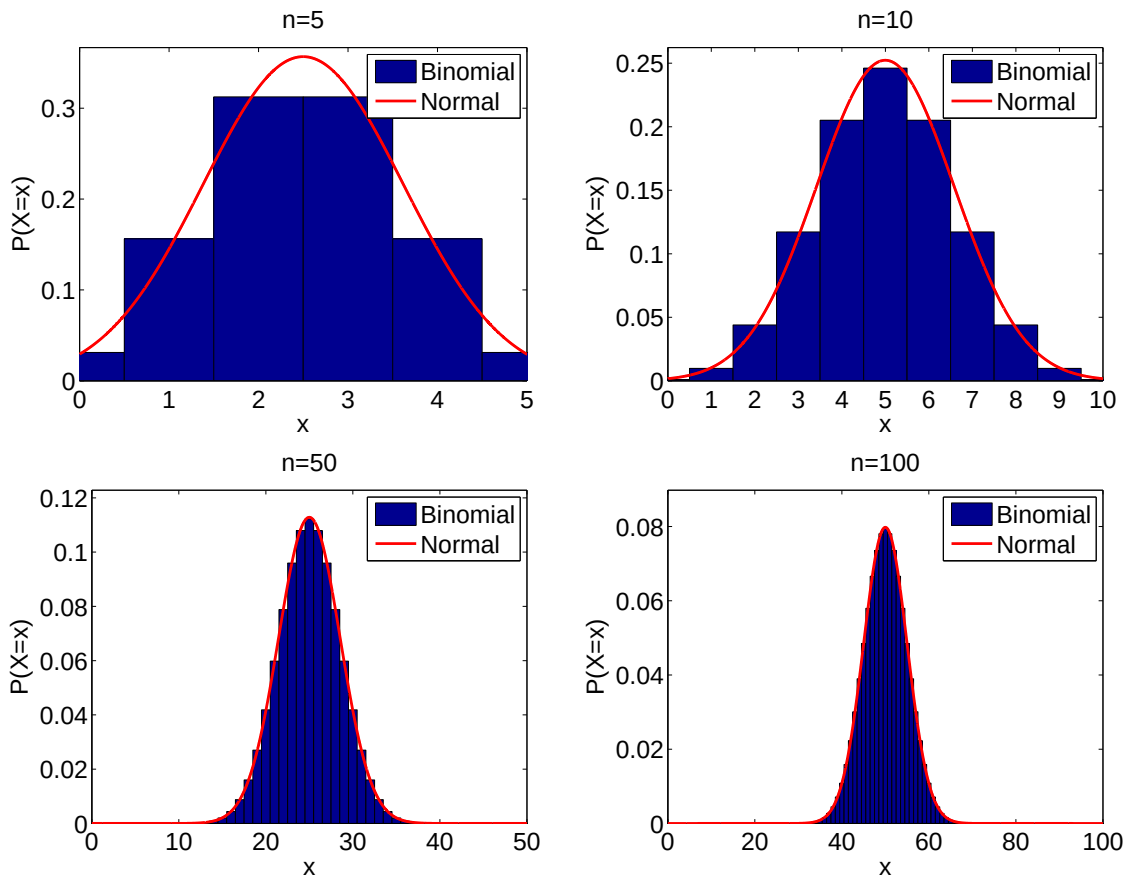

Figure 11: The figures show that the Gaussian distribution more closely approximates the binomial distribution as $n \to \infty$ with $p = 0.5$.

Table 12: The properties of the Poisson distribution as described in Equation (110).

Mean	$\mu = \lambda$
Variance	$\sigma^2 = \lambda$
Moment Generating Function	$M(t) = e^{\lambda(e^t - 1)}$
Characteristic Function	$\phi(\omega) = e^{\lambda(e^{i\omega} - 1)}$

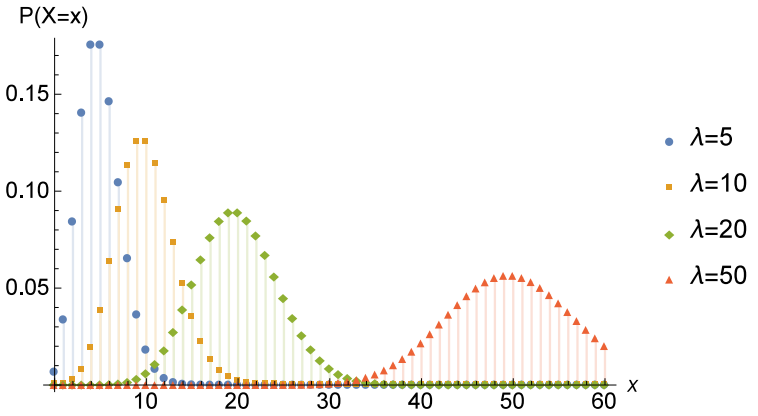

Figure 12: Example of Poisson distributions with different means.

Example 11 Demonstrate that the Poisson distribution approaches the Gaussian distribution as $\lambda \to \infty$.

Solution: As shown in Figure 10.3.2, the normal distribution more closely approximates the Poisson distribution as $\lambda \to \infty$.

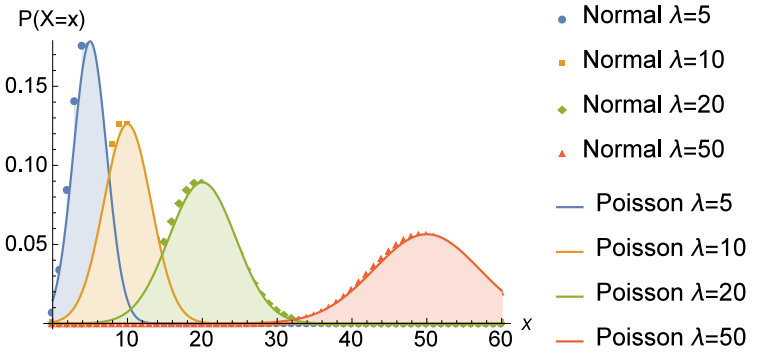

Figure 13: As λ increases, the Gaussian distribution with $\mu = \lambda$ and $\sigma = \sqrt{\lambda}$ more closely approximates the Poisson distribution.

10.3.3 The Central Limit Theorem

The Central Limit Theorem states that if $X_1, X_2, ..., X_n$ are independent random variables with the same probability distribution function that has a finite mean μ and a finite variance σ^2, and $S_n = \sum_{i=1}^{n} X_i$ then,

$$\lim_{n\to\infty} P\left(a \leq \frac{S_n - n\mu}{\sigma\sqrt{n}} \leq b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{u^2/2} du. \tag{111}$$

This theorem shows that any distribution with a finite mean and variance can be approximated with the Gaussian distribution as $n \to \infty$.

10.3.4 Uniform Distribution

The uniform distribution for a random variable X is given by the density function

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{otherwise.} \end{cases} \tag{112}$$

Although the uniform distribution seems trivial, it is important to notice that as show in Figure 10.3.4 because the area under the function must always be 1, the distribution is larger when the difference between the boundaries is smaller. Table 13 shows the mean and variance of the distribution.

Table 13: The properties of the uniform distribution as described in Equation (112).

Mean	$\mu = \frac{1}{2}(a+b)$
Variance	$\sigma^2 = \frac{1}{12}(b-a)^2$

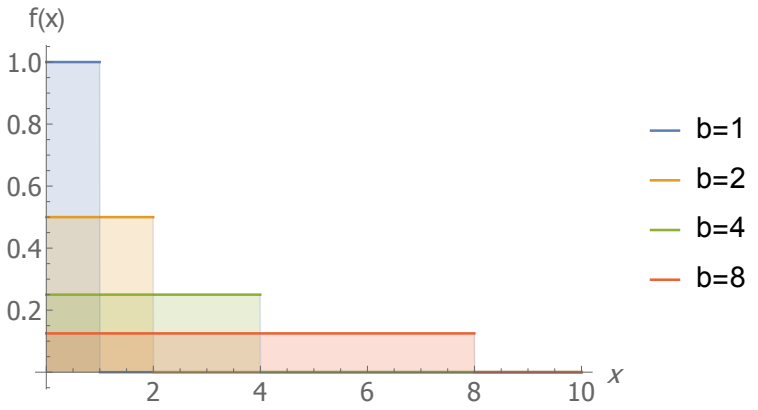

Figure 14: The probability distribution function is shown for a uniform distribution with $a = 0$.

10.4 Sampling Theory

10.4.1 Basics

Sampling is choosing one sample from a population that is defined by a probability distribution. Sampling with replacement is when a sample can be chosen multiple times. Sampling without replacement does not

allow a member to be chosen multiple times. A sample statistic is a number obtained from a sample in order to estimate a parameter of the population. Sample statistics for an n sized sample can be written as a function of the random variables $X_1, X_2, ..., X_n$. Usually, Greek letters are used for population parameters and Roman letters are used for the related sample statistics. The sampling distribution is the probability distribution of the sample statistic [8].

Example 12 If there is a bag containing only one red marble, one green marble, and one blue marble and you take a sample marble and then put it back in the bag, what type of sampling did you do?

Solution: This situation is an example of sampling with replacement because the marble was returned to the bag, so it could be chosen again.