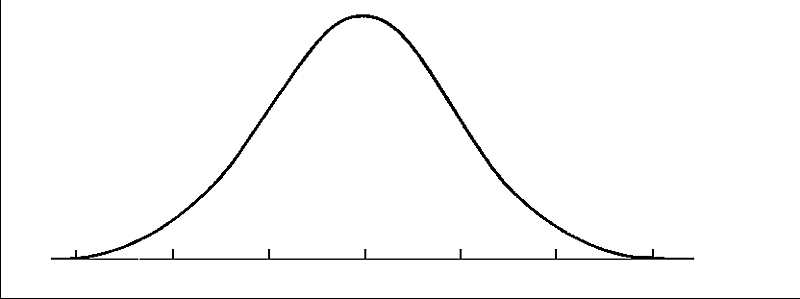
# Lab 3: Continuous Distributions and Discrete Distributions (Chapter 3)

## Objectives

* Explore the various continuous distribution types
* Use JMP and Excel to solve problems with normal distributions
* Solve problems regarding discrete distributions (binomial and Poisson)

## Normal Continuous Distribution

The most commonly used distribution to describe possible values for a random variable is the normal distribution. When the possible values for a normally distributed random variable are graphed, they form a bell curve, shown below.



**The peak of the curve corresponds to the mean or expected value of the variable X, . The curve also takes into account the variance of the variable X, denoted by .**

To easily check if a variable is normally distributed, you can create a histogram in JMP (using Analyze > Distribution), then add a continuous fit line by clicking the red triangle then selecting Continuous Fit > Normal. You can then qualitatively analyze if the fit is normal.



To solve problems having to do with probability of normally distributed random variables, we can use Z charts (located in Appendix A, Table I of your book) or computer software, namely Excel. To efficiently solve these problems:

1. Standardize the variable X using the equation below:

This allows us to use the Z charts in the textbook. If you are using Excel, it is not necessary to standardize X before proceeding. Note that from now on, if we refer to variable Z, that is the standardized variable. If we refer to variable X, it is a non-standardized variable.

1. Now we must rearrange the probability expression so that you can use the Z charts or Excel’s built-in functions. You must rearrange the expression, using the probability rules we learned about in Lab 2, so that all probabilities are in the less-than-or-equal-to form below.

Example 3-9 in the textbook shows several good examples of rearranging probability expressions.

1. Either use the Z charts to solve for (example 3-8 in the textbook describes how to use the charts) or use Excel. Below is a table of helpful Excel functions to use for normal distributions:

|  |  |  |
| --- | --- | --- |
| **Function** | **Function Name** | **Description** |
| normsdist(z) | Normal standard distribution | Find the probability of |
| normdist(x,µ,σ,true) | Normal distribution | If you haven’t standardized your variable, use this function to find .   * µ is the mean * σ is the standard deviation * “true” refers to when you want the probability (corresponding to the CDF) and “false” is used if you just want the value of f(x) at x (corresponding to the PDF) |
| normsinv(P) | Inverse of normal standard distribution | Use this function if you have the probability and want to find the z it corresponds to |
| norminv(x,µ,σ) | Inverse of normal distribution | Use this function if you have the probability of a non-standardized variable and want to find the x it corresponds to |

## Other Continuous Distributions

Other distributions you may use in problem solving are outlined below. Similar to with a normal distribution, you can add their fits to histograms in JMP and use Excel to solve problems with probability. It is not necessary to know these as well as you know normal distributions, so I am pointing out the sections of the book describing these distributions if you’d like to read further.

|  |  |  |
| --- | --- | --- |
| **Distribution** | **Description** | **Excel functions** |
| Lognormal | W has a normal distribution with mean θ and variance ω2 . The logarithm of W follows a lognormal distribution, or:  Explained further in equations 3-6 and 3-7 in the textbook | lognormdist  loginv |
| Gamma | Explained in section 3-5.3 in the textbook | gammadist  gammainv |
| Weibull | Explained in section 3-5.4 | weibull |
| Beta | Explained in section 3-5.5 | betadist  betainv |

## Discrete Distributions

The distributions above all apply to continuous random variables. There are also distributions that apply to discrete random variables.

1. A binomial distribution can be used for discrete random variables that have only two possible values (for example, heads/tails, off/on, high/low, etc.) A binomial distribution can be represented with the PDF:

Where n is the number of trials, p is the probability of success and x is the number of successes. The mean and variance of a binomial random variable are:

To use Excel to solve binomial distribution problems, use the function binomdist(x,n,p,true) where, as in the normdist functions, “true” refers to finding the probability (CDF) and “false” refers to finding the f(x) (PDF).

1. A Poisson distribution can be used to calculate the number of events that occur randomly over an interval. A Poisson distribution is represented by:

Where λ is the mean and variance of X. To use Excel for Poisson distribution problems, use the function poisson(x,λ,true) where, as in the normdist functions, “true” refers to finding the probability (CDF) and “false” refers to finding the f(x) (PDF).

## Lab 3 Exercises

You use JMP or Excel as needed for today’s problems, or if you prefer, you may do them by hand using the charts in your textbook.

*Continuous Distributions*

1. Using the data for Exercise 3-87 (from the [textbook site](http://bcs.wiley.com/he-bcs/Books?action=index&bcsId=6235&itemId=0470631473" \t "_blank)), from your choices of normal, lognormal, or Weibull, which distribution fits the data the best and why?
2. Find the following probabilities:
   1. P(0.02 < Z < 4.8)
   2. P(X < 13.8 mm) where μ = 12.8 mm and σ = 5.2 mm
   3. P(X > 100 hr) where μ = 110 hr and σ = 12 hr
   4. P(0.0050 mg/mL < X < 0.0110 mg/mL) where μ = 0.0099 mg/mL and σ = 0.0021 mg/mL
3. What is the z value that corresponds to a probability of 95%?
4. If a process creates a part with a diameter that follows a normal distribution with a mean value of 106.7 mm and a standard deviation of 1.8 mm, in what range of diameters will 95% of the parts produced occur? Give your answer in the form of 106.7 ± *x* mm.
5. For a lognormal distribution with θ = 5 and ω = 1.5, what proportion of the distribution is greater than 100?

*Discrete Distributions*

1. A game sells tokens for $1 a pack. Each pack has a 1:100 chance of having a special token included.
   1. What are the odds of getting a single special token when purchasing 100 packs?
   2. What are the odds of getting two or more special tokens when purchasing 100 packs?
2. You are monitoring the output of a reactor for catalyst particles escaping by counting the particles over a period of 10 minutes. Historically, the average number of particles per 10 minutes is 4.3 particles.
   1. What is the probability of finding 8 or more particles escaping the reactor in 10 minutes?
   2. What is the probability of finding 1, 2, or 3 particles over a 10 minute span?