

Section 6 – 1A: Continuous Probability Distributions

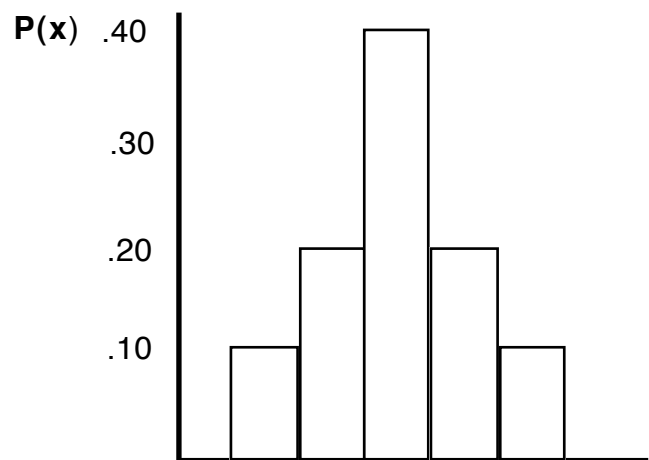
Discrete Probability Distributions

Chapter 5 developed the mathematics of **Discrete** Probability Distributions. Discrete Probability Distributions have a **fixed number n** of values for the **discrete random variable x**. These probability distributions are represented by a table or a histogram graph with rectangles. The width of each rectangle is one. The height of each rectangle is $P(x)$ and represents the probability for each of the different values of the random variable x . The area of each rectangle is the width of the rectangle (1) times the height of the rectangle $P(x)$. **The area of each rectangle is $P(x)$** . **The total of all the $P(x)$ values is 1 so the area of all the rectangles in a Discrete Probability Histogram must also total 1.**

A Discrete Probability Distribution Table

x	P(x)
0	.10
1	.20
2	.40
3	.20
4	.10

A Discrete Probability Distribution Histogram



The total of all the $P(x)$ values total 1

$$\sum P(x) = .10 + .20 + .40 + .20 + .10 = 1$$

The area of each rectangle is $P(x)$.

The area of all the rectangles must total 1

Continuous Probability Distributions

This chapter will introduce the concept of a **Continuous Probability Distribution**. These distributions will **not be represented by probability tables or histograms**. These distributions will be represented by **smooth continuous curves**. The area in yellow represents the total of all the separate values of $P(x)$ so the total area must be 1.



The area under the curve is used to find $P(x)$
The entire area under the curve must total 1

Continuous Probability Distributions

Discrete Probability Distributions have a discrete random variable x . The discrete random variable x is represented by the **whole numbers** on the interval that x is defined for

Continuous Probability Distributions have a **continuous random variable x** . The continuous random variable x takes on **every real number on the interval** that x is defined for. Since the real numbers are continuous on any interval the random variable x is continuous. Continuous random variables are used to represent data that is measured on a continuous scale: weight, time, length, etc..

Discrete Probability Distributions are represented by a histogram graph of rectangles. The area of each rectangle is $P(x)$ and the total area of all the rectangles is 1.

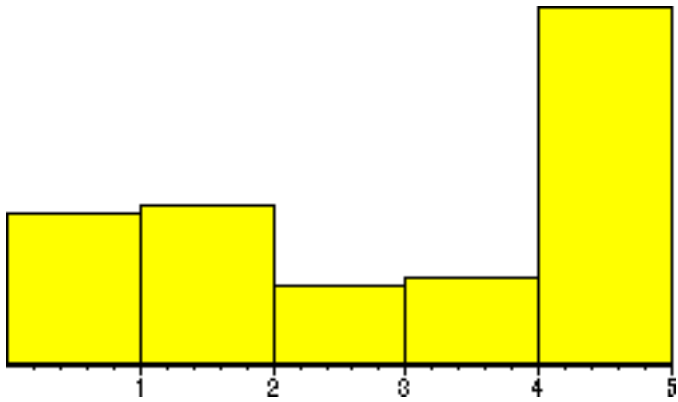
Continuous Probability Distributions are far more difficult to describe. The graph of a Continuous Probability Distributions is an **unbroken curve above an x axis**. The x axis represents the continuous real values for x . No rectangle can be drawn to represent a single value of x as it would have to be the width of a single point which is zero. The point on the curve above a value for x is not $P(x)$. In fact, for any single value of x , $P(x)$ will be infinitely small with a round off value of $0+$. The total area under the curve represents the total of all the infinite number of values of $P(x)$ and that total is 1.

One way to try to understand a Continuous Probability Distribution is to start with a Discrete Probability Distribution and begin to transform it into a Continuous Probability Distribution

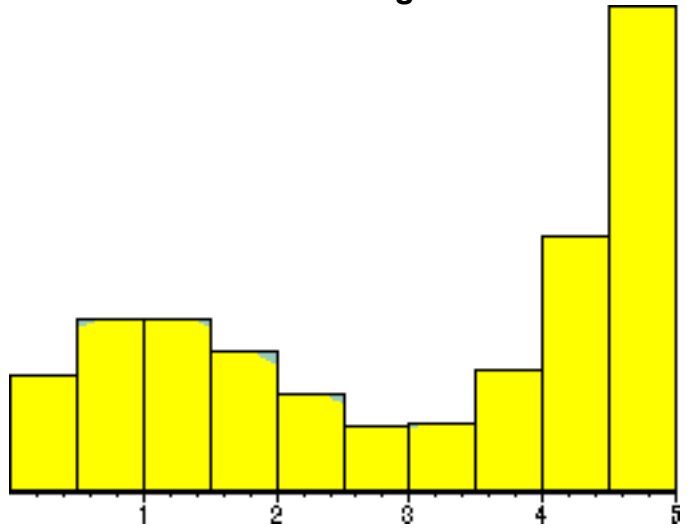
One way to look at the transition from discrete distributions to continuous distributions is to think of the following process. Start with a histogram of n rectangles. The width of each rectangle is one, the height of each rectangle is $P(x)$ and the total area of all of the rectangles is 1. Now split each rectangle into 2 equal widths so that the width of each rectangle is now $.5$. Continue splitting the rectangles over and over. This process will lead to the width of each rectangle becoming very small. As the rectangles become thinner and thinner the graph starts to look like one continuous set of thin lines. The final graph is called a Continuous Probability Distribution.

Discrete to Continuous Example

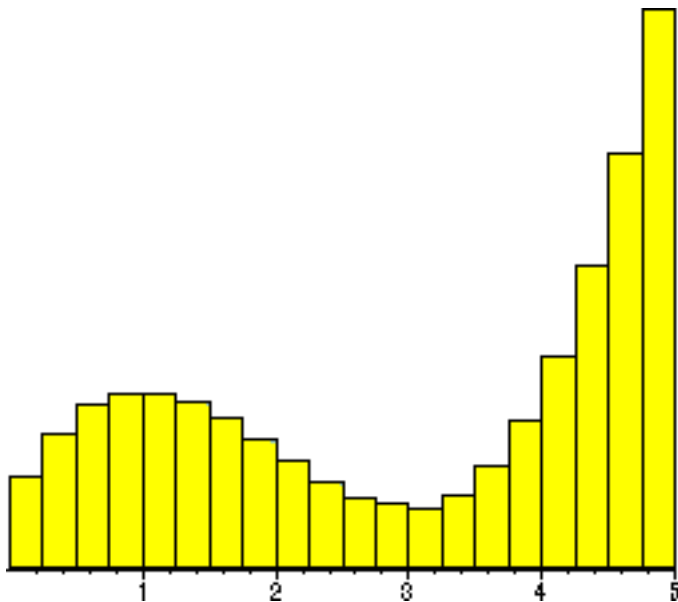
**A Discrete Probability Distribution
with 5 values for x between 0 and x
The area of all 5 rectangles = 1**



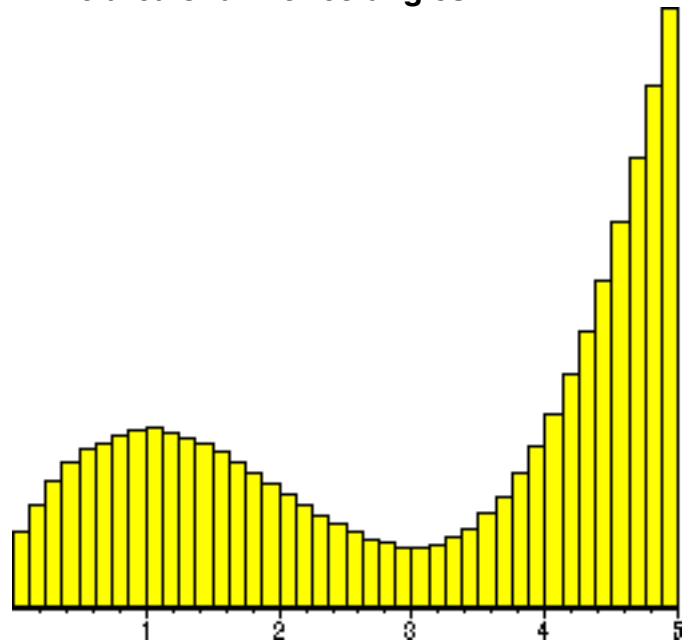
**The same Probability Distribution with
with 10 values for x between 0 and x
The area of all 10 rectangles = 1**



**The same Probability Distribution with
with 20 values for x between 0 and x
The area of all 20 rectangles = 1**



**The same Probability Distribution with
with 40 values for x between 0 and x
The area of all 40 rectangles = 1**

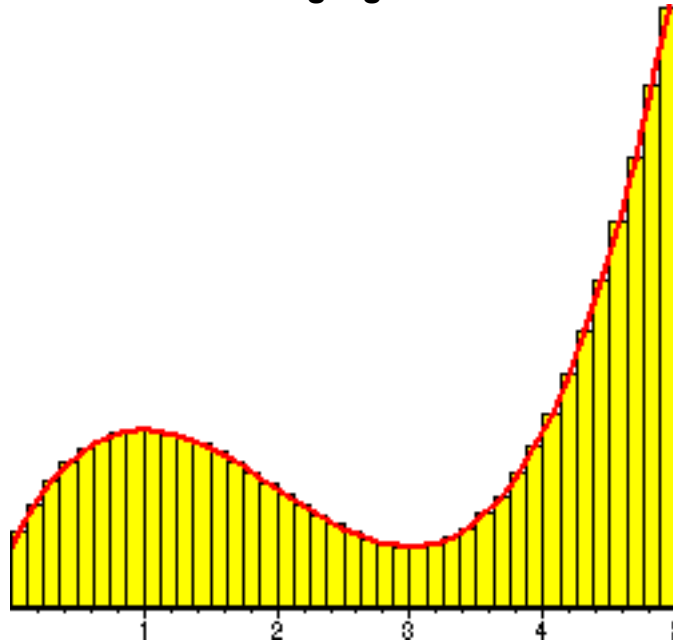


If we continue to split each rectangle over and over we would have an infinite number of very thin rectangles from $x = 0$ to $x = 5$. The area of each rectangle will get closer and closer to zero as we add more and more thinner and thinner rectangles to fill the area from $x = 0$ to $x = 5$. While the area of each rectangle gets closer and closer to zero the total of all of the rectangles will still remain 1.

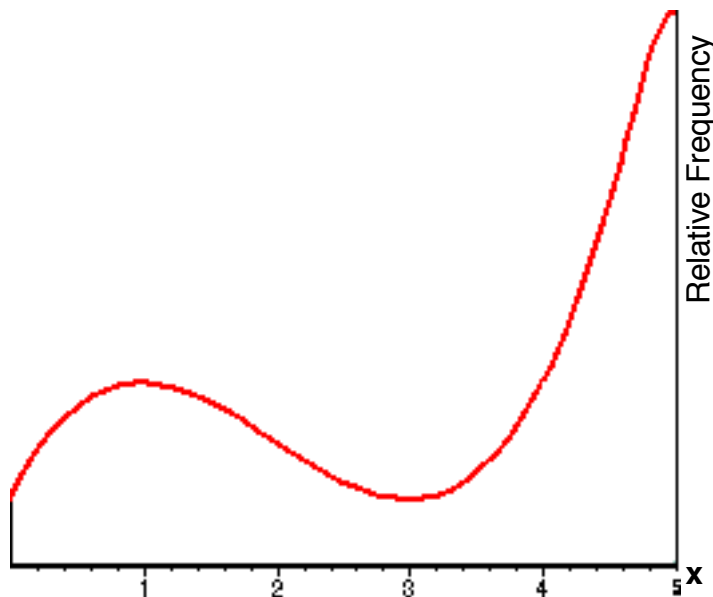
Changing An infinite Number of **Discrete** Rectangles into a **Continuous** Curve

If we continue to split each rectangle over and over again we would have an infinite number of very, very, very thin rectangles from $x = 0$ to $x = 5$. In fact the width of each rectangle would get closer and closer to 0. If we place one point at the top of each of the infinitely many thin rectangles and erase the rectangles themselves we would get a **continuous curve above a horizontal x axis from 0 to 5**. The points on the x axis stand for the values of the **continuous random variable x**. The red curve is a **continuous curve** and the area under the curve represents a **Continuous Probability Distribution**. Another name for a **Continuous Probability Distribution** is a **Probability Density Function** or a **PDF**.

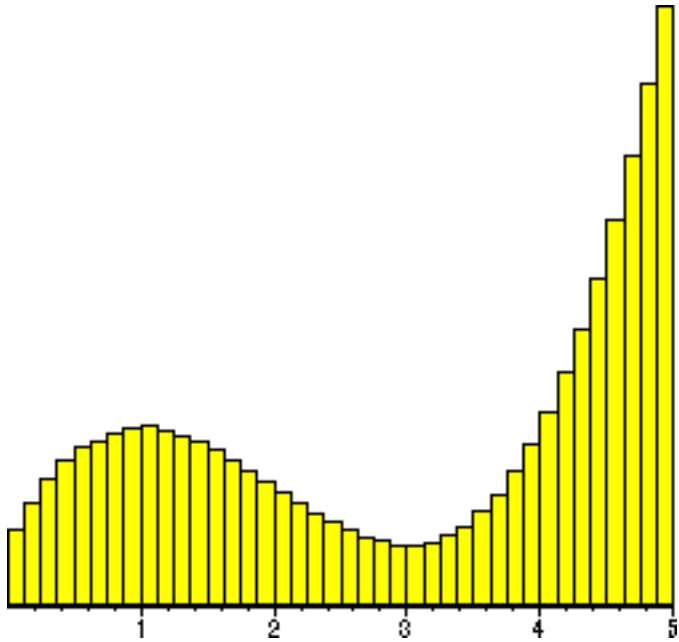
As the number of rectangles between 0 and 5 get larger and larger the width of each rectangle gets smaller and smaller.



If we place one point at the top of each of the infinitely many rectangles and erase the rectangles themselves we would get a **continuous curve above a horizontal x axis from 0 to 5**.



**A Discrete Probability Distribution
with 40 values of x from 0 to 5**



The total area for the 40 rectangles is 1

**A Continuous Probability Distribution
with continuous x values from 0 to 5**



The total area under the curve is 1

A Discrete Probability Distribution	A Continuous Probability Distribution
The values on the x axis are discrete (often the whole numbers) with a gap between each number)	The values for x on the x axis are continuous (no gaps between the real numbers)
A table can display each x and P(x)	No table can represent the infinite values of x
The graph is a histogram of n rectangles	The graph is an unbroken curve above the x axis
The value of each P(x) can be calculated	The value of each P(x) is 0+
The total area of all n rectangles is 1	The total of the entire area under the curve is 1
The probability for a range of discrete x values from x = a to x = b is found by adding up the P(x) values for each x in the range from a to b $P(a) + P(a+1) + P(a+1) + \dots + P(b)$	The probability for a range of continuous x values from x = a to x = b is found by finding the area under the curve from a to b by integral calculus. $\int_a^b f(x)$

**What is a Continuous Probability Distribution good for
if $P(x) = 0+$ for every value of x**

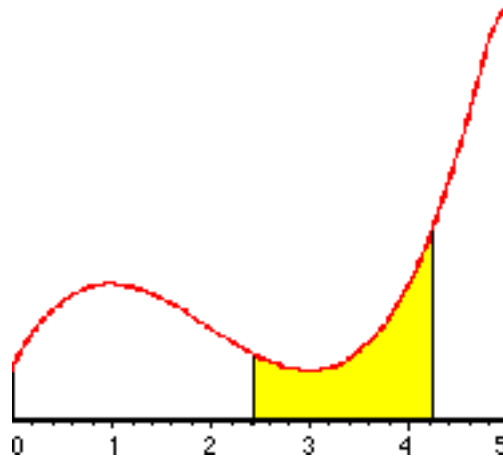
The area under the curve is used to find the Probability for a range of x values ?

Since continuous probability functions are defined for an infinite number of points over a continuous interval, the value of $P(x)$ for any single value of x is always $0+$. Probabilities for continuous probability functions are measured over **intervals**, not single points. That is, **the area under the continuous curve between two distinct points defines the probability for that interval.**

$$P(2.4 < x < 4.2)$$

A **Continuous Probability Distribution that represents a relative frequency curve for an infinite data set of values from $x = 0$ to $x = 5$** is displayed below. If we select **one data bit** from the data set the probability that it will be a value **between** 2.4 and 4.2 is written as $P(2.4 < x < 4.2)$.

To find this probability we would need to **find the area under the curve between $x = 2.4$ to $x = 4.2$**



We would use integral calculus

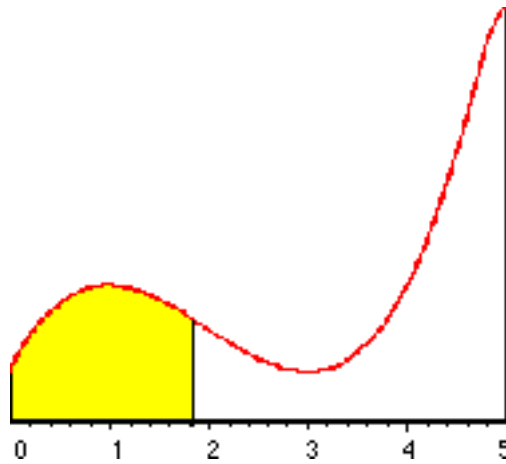
$$\int_{2.4}^{4.2} f(x)$$

to find the area that represents

$$P(2.4 < x < 4.2)$$

$P(x < 1.8)$

If we **select one data bit** from the data set the probability that it will be a value **less than 1.8** is written as $P(x < 1.8)$. To find this probability we would need to **find the area under the curve to the left of 1.8**



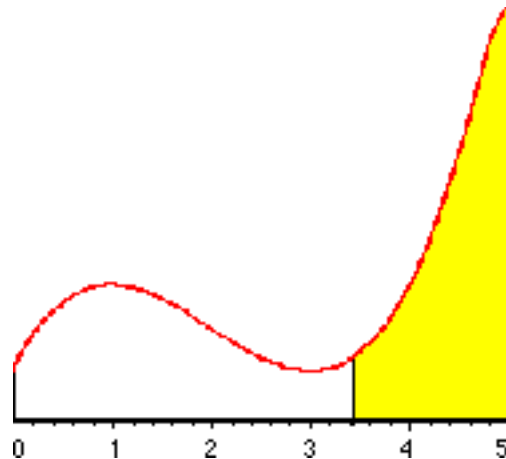
We would use integral calculus

$$\int_0^{1.8} f(x)$$

to find the area
that represents $P(x < 1.8)$

$P(x > 3.4)$

If we select one data bit from the data set the probability that it will be a value **more than 3.4** is written as $P(x > 3.4)$. To find this probability we would need to **find the area under the curve to the right of 3.4**



We would use integral calculus

$$\int_{3.4}^5 f(x)$$

to find the area
that represents $P(x > 3.4)$

What is a Continuous Probability Distribution good for if we need to use integral Calculus to find the Probability for a range of x values ?

That is a great question.

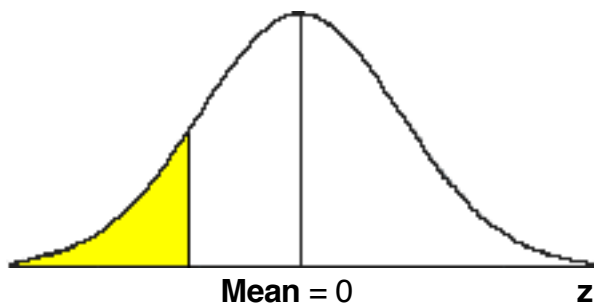
While integral calculus is the basis for finding the area under a curve, it is not required that you know this level of math to find the probabilities of **certain** Continuous Probability Distributions. A small number of different continuous distributions can be used to answer the probability questions that we will be asked in the rest of this course. The area under the curve for a range of x values has been calculated and put into a table for each of these 4 distributions. You will learn to **read the area under the curve for a given range of values from these tables** and use that information to answer the probability questions that a student will encounter in an general education level statistics course.

We will focus on the following 4 Continuous Probability Distributions for the rest of this course. There are many other continuous probability distributions but these 4 form the basis for almost all introductory statistics courses. They allow us to answer most of the important probability questions that a student will encounter.

The 4 basic Continuous Probability Distributions for this course

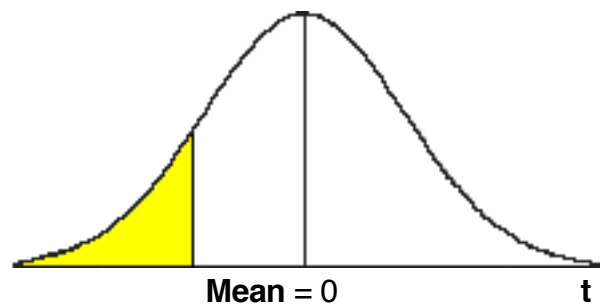
Each Distribution has a table to help you find the areas in yellow

The Standard Normal Probability Distribution (z)



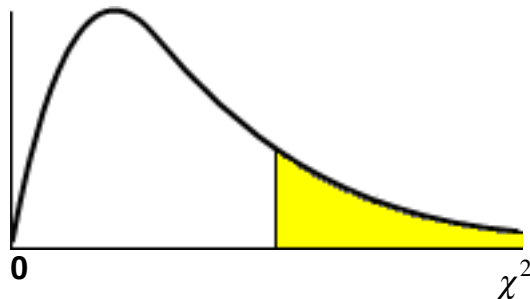
Standard Deviation = 1

The t Probability Distribution



Standard Deviation = varies

The Chi Squared Probability Distribution χ^2



The F Probability Distribution

