## Section 5.1 - Discrete Probability Distributions

A **random variable** is a variable (often represented by $x$) that has a single numerical value, determined by chance, for each outcome of a procedure.

A **probability distribution** is a description that gives the probability for each value of a random variable. A probability distribution is often expressed in graph, table, or formula.

Examples...

1. Roll a fair six-sided die. Let $x =$ the number rolled.
2. Randomly select a single quiz from those returned. Let $x =$ score on the quiz.
3. Randomly select a full-term, newborn baby from St. James Hospital. Let $x =$ the weight (in lbs) of the baby.

A **discrete random variable** is a variable from a discrete data set.

A **continuous random variable** is a random variable from a continuous data set.

Requirements for a Probability Distribution

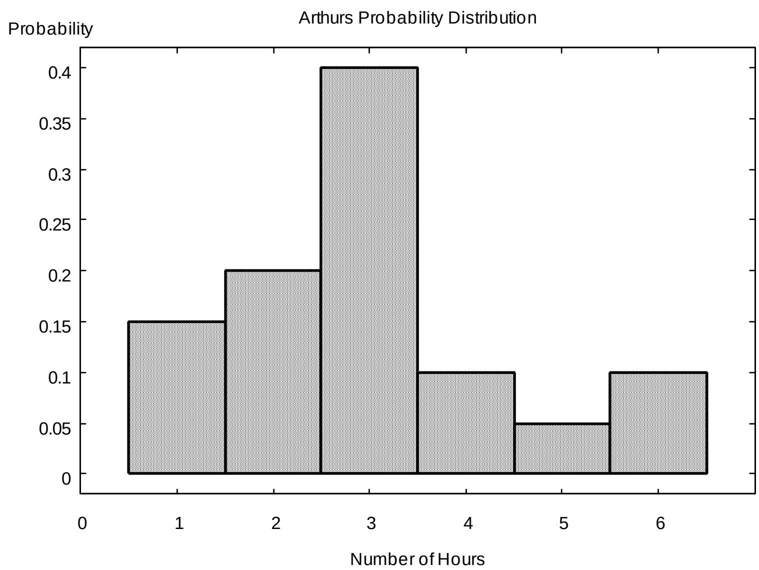
1. $\sum P(x) = 1$, or at least very close to 1 if you are rounding.
2. $0 \leq P(x) \leq 1$

**Example:** The following table shows the probabilities assigned by Arthur to the number of hours spent on homework on any given night. Show that the table defines a probability distribution.

| Hours | Probability |
|---|---|
| 1 | 0.15 |
| 2 | 0.20 |
| 3 | 0.40 |
| 4 | 0.10 |
| 5 | 0.05 |
| 6 | 0.10 |

There are a number of ways to graph a probability distribution. We will normally use a **probability histogram**.

Here is the probability histogram for Arthur's probability distribution.



<u>Mean, Variance, and Standard Deviation</u>

(for Discrete Random Variables)

- Mean for a probability distribution (Expected Value)

$$\mu = \sum [x \times P(x)]$$

- Variance for a probability distribution

$$\sigma^2 = \sum [(x - \mu)^2 \times P(x)]$$

or

$$\sigma^2 = \sum [x^2 \times P(x)] - \mu^2$$

- Standard deviation for a probability distribution

$$\sigma = \sqrt{\sigma^2}$$

Recall that we used the standard deviation to identify the minimum and maximum "usual" values:

minimum "usual" value = $\mu - 2\sigma$

maximum "usual" value = $\mu + 2\sigma$

Rare Event Rule for Inferential Statistics

If, under a given assumption, the probability of a particular *observed* event is extremely small, we conclude that the assumption is probably not correct.

Using probabilities to determine when results are unusual...

- $x$ is an unusually high number if the probability of $x$ or more is 5% or less
- $x$ is an unusually low number if the probability of $x$ or fewer is 5% or less