**Section 2.2 – Density Curves & Normal Distributions**

**Density Curves**

|  |
| --- |
| **Exploring Quantitative Data**1. Always plot your data: make a graph, usually a dotplot, stemplot or a histogram.2. Look for the overall pattern (shape, center, spread) and for striking departures such as outliers.3. Calculate a numerical summary to briefly describe center and spread.New step:4. Sometimes the overall pattern of a *large* number of observations is so regular that we can describe it with a *smooth curve*. |

This type of *smooth curve* is called a **Density Curve**.

**Definition**: A **density curve** is a curve that

* Is always above the horizontal axis, and
* Has an area of exactly 1 underneath it

A density curve describes the overall pattern of a distribution. The area under the curve and above any interval of values on the horizontal axis is the proportion of all observations that fall in that interval.

**Note**: *no set of real data is exactly described by a density curve. The curve is an approximation that is easy to use and accurate enough for practical use.*

Because the density curve represents a *population* of individuals, the mean is denoted by μ (the Greek letter mu) and the standard deviation is denoted by σ (the Greek letter sigma).

|  |
| --- |
| **Distinguishing the Median and Mean of a Density Curve** (Diagrams on p. 102)* The **median** of a density curve is the *equal-areas* point, the point that divides the area under the curve in half.
* The **mean** of a density curve is the *balance point*, the point at which the curve would balance if made of solid material.
* The median and mean are the same for a perfectly symmetric density curve. The both lie at the center of the curve. The mean of a skewed curve is pulled away from the median in the direction of the long tail.
 |

**Team Work**: Complete Check Your Understanding on p. 107.

Probably the most famous of all *density curves* are **Normal curves**. The distributions they describe are called **Normal distributions**. They play a very large part in statistics.



Normal curves have several properties:

* All Normal curves have the same overall shape: symmetric, single-peaked, bell-shaped.
* Any specific Normal curve is completely described by its mean μ and standard deviation σ.
* The mean is located at the center and is equal to the median. Changing μ without changing σ moves the Normal curve along the horizontal axis without changing its shape.
* The standard deviation σ controls the spread of a Normal curve. Normal curves with larger standard deviations are more spread out.

The points at which the Normal curve changes from *concave down* to *concave up* occurs one standard deviation from the mean. Because of this, the standard deviation can be estimated by the graph.

**Definition:** A **Normal distribution** is described by a Normal density curve. Any particular Normal distribution is completely specified by its mean μ and standard deviation σ. The mean of a Normal distribution is at the center of the symmetric **Normal curve** and equals the median. The standard deviation is the distance from the center to the inflection points (where concavity changes) on either side.

**Notation**: We abbreviate the Normal distribution with mean μ and standard deviation σ as ***N(μ, σ)****.*

**The 68-95-99.7 Rule**

|  |  |
| --- | --- |
| In a Normal distribution with mean μ and standard deviation σ:* Approximately 68% of the observations fall within 1 σ of the mean μ.
* Approximately 95% of the observations fall within 2 σ’s of the mean μ.
* Approximately 99.7% of the observations fall within 3 σ’s of the mean μ.

(Note: this rule does not apply to any distribution – only the Normal. Common error on AP Exam.) |  |

**Example**: The mean batting average for the 432 Major League Baseball players in 2009 was 0.261 with a standard deviation of 0.034. Suppose the distribution is exactly Normal with μ = 0.261 and σ = 0.034.

a. Sketch a Normal density curve for this distribution. Label the points that are 1, 2, and 3 standard deviations from the mean.

b. What percent of batting averages are above 0.329?

c. What percent of batting averages are between 0.193 and 0.295?

**Team Work**: Complete Check Your Understanding on p. 112.

**The Standard Normal Distribution**

**Definition**: The **standard Normal distribution** is the Normal distribution with mean 0 and standard deviation 1. If a variable x has any Normal distribution *N(μ, σ)* with mean μ and standard deviation σ, then the standardized variable $z=\frac{x-μ}{σ}$ has the standard Normal distribution.

|  |  |
| --- | --- |
|  | **68-95-99.7 Rule**: For the standard Normal distribution |

The **standard Normal table** is contained in Table A. It is a table of areas under the Normal curve. The table entry for each value z is the area under the curve to left of z. This is also known as the *lower tail.*



**Example**: Finding areas under the standard Normal curve.

Use *Table A* to find the proportion of observations from the standard Normal distribution given the following z-values. Draw a diagram for each.

a. Less than z = -1.25

b. Less than z = 0.81

c. Greater than z = 0.81

d. Between z = -1.25 and z = 0.81

**Example**: Repeat the previous example using *technology*.

a. Less than z = -1.25

b. Less than z = 0.81

c. Greater than z = 0.81

d. Between z = -1.25 and z = 0.81

**Example**: Working backwards…..

Find the 90th percentile of standard Normal distribution

a. Using *Table A*

b. Using *technology*

**Team Work**: Complete Check Your Understanding on p. 116.

Homework: pp. 128 problems 35, 37, 41, 43, 45, 47, 49, 51