

## 6.25: Putting It Together- Probability and Probability Distribution

### Let's Summarize

Here is a summary of the key concepts developed in this module:

- The *probability* of an event is a measure of the likelihood that the event occurs. Probabilities are always between 0 and 1. The closer the probability is to 0, the less likely the event is to occur. The closer the probability is to 1, the more likely the event is to occur.
- The two ways of determining probabilities are empirical and theoretical.
  - *Empirical* methods are based on data. The probability of an event is approximated by the relative frequency of the event.
  - *Theoretical* methods use the nature of the situation to determine probabilities.
- Following are some common probability rules:
  - $P(\text{not } A) = 1 - P(A)$ .
  - When two events have no outcomes in common, they are disjoint. If  $A$  and  $B$  are disjoint events,  $P(A \text{ or } B) = P(A) + P(B)$ .
  - When the knowledge of the occurrence of one event  $A$  does not affect the probability of another event  $B$ , we say the events are independent. If  $A$  and  $B$  are **independent events**,  $P(A \text{ and } B) = P(A) \cdot P(B)$ .
- When we have a quantitative variable with outcomes that occur as a result of some random process (e.g., rolling a die, choosing a person at random), we call it a *random variable*. There are two types of random variables:
  - *Discrete* random variables have numeric values that can be listed and often can be counted. We find probabilities using areas in a probability histogram.
  - *Continuous* random variables can take any value in an interval and are often measurements. We use a density curve to assign probabilities to intervals of  $x$ -values. We use the *area under the density curve to find probabilities*.
- We use a *normal density curve* to model the probability distribution for many variables, such as weight, shoe sizes, foot lengths, and other physical characteristics. For a normal curve, the empirical rule for normal curves tells us that 68% of the observations fall within 1 standard deviation of the mean, 95% within 2 standard deviations of the mean, and 99.7% within 3 standard deviations of the mean.
- To compare  $x$ -values from different distributions, we standardize the values by finding a **z-score**:  $Z = \frac{x - \mu}{\sigma}$
- A z-score measures how far  $X$  is from the mean in standard deviations. In other words, the z-score is the number of standard deviations  $X$  is from the mean of the distribution. For example,  $Z = 1$  means the  $x$ -value is one standard deviation above the mean.
- If we convert the  $x$ -values into z-scores, the distribution of z-scores is also a normal density curve. This curve is called the *standard normal distribution*. We use a simulation with the standard normal curve to find probabilities for any normal distribution.

CC licensed content, Shared previously

- Concepts in Statistics. **Provided by:** Open Learning Initiative. **Located at:** <http://oli.cmu.edu>. **License:** [CC BY: Attribution](#)

This page titled [6.25: Putting It Together- Probability and Probability Distribution](#) is shared under a [CC BY 4.0](#) license and was authored, remixed, and/or curated by [Bill Pelz](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.