**Activity 7.5.1 – Discrete Probability Distributions**

**Number of Days Absent from School**

The following frequency histogram displays the number of days absent from school for a small 10th grade class last school year. The class consists of 90 students.



This frequency histogram shows the distribution of a *discrete random variable*. A random variable is a numerical outcome to a random circumstance. It is considered discrete if it has a finite number of outcomes that can be listed. The random variable, number of days absent from school, takes on the values {0, 1, 2, 3, 4, 6, 7}. By carefully examining a random variable’s distribution we can find probabilities of individual outcomes.

1. Suppose one student is selected at random from this sample. What is the probability the student was absent for 4 days?
2. If one student is selected at random from this sample, what is the probability the student was absent for at least 4 days?

The previous questions illustrate how relative frequencies correspond to probabilities. If we find the probability for each individual outcome, we have a *discrete probability distribution*.

A discrete probability distribution is a list of possible values that a random variable can assume along with each value’s probability. Each probability must be between 0 and 1 and the sum of the probabilities must be 1.

1. Use the frequency histogram on page 1 to construct a discrete probability distribution for the number of days absent. Fill in the table below. Round probabilities to the three decimal places.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| $$x$$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| $$P(x)$$ |  |  |  |  |  |  |  |  |

1. What is more likely: randomly selecting a student who was absent for 5 or more days, or randomly selecting a student who was absent for 2 or less days?

**Mean of a Discrete Random Variable**

Suppose we randomly select a student many, many times, and find the average of the outcomes we observe. The value we obtain would be the *mean* or *expected value* of the discrete random variable. The mean is a weighted average and is found by the following formula.

$$μ=\sum\_{}^{}xP(x)$$

1. Find and interpret the mean of the discrete random variable.

**Variance and Standard Deviation of a Discrete Random Variable**

The standard deviation of a random variable describes the typical deviation between random variable outcomes and the mean. The smaller the standard deviation, the closer outcomes are to the mean. The greater the standard deviation, the farther outcomes are from the mean.

$$σ=\sqrt{\sum\_{}^{}\left(x-μ\right)^{2}P(x)}$$

1. Complete the table below to calculate the standard deviation of the discrete probability distribution.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| $$x$$ | $$P(x)$$ | $$x-μ$$ | $$\left(x-μ\right)^{2}$$ | $$\left(x-μ\right)^{2}P(x)$$ |
| 0 | 0.056 |  |  |  |
| 1 | 0.089 |  |  |  |
| 2 | 0.200 |  |  |  |
| 3 | 0.256 |  |  |  |
| 4 | 0.189 |  |  |  |
| 5 | 0.111 |  |  |  |
| 6 | 0.078 |  |  |  |
| 7 | 0.022 |  |  |  |

1. The variance is the sum of the values in the last column. Find the variance.
2. The standard deviation is the square root of the variance. Find the standard deviation.
3. Are any of the outcomes unusual? Explain.

**Constructing a Probability Histogram**

A *probability histogram* of a discrete random variable displays the probability associated with each random variable outcome. The horizontal axis contains the outcomes of the random variable and the vertical axis contains probabilities.

1. Lisa scored 0, 1, 2 and 3 goals in soccer games she played last season. Suppose the random variable *x* represents the number of goals she scored in a game. The probability distribution below details the probability for each value of *x*.

|  |  |
| --- | --- |
| $$x$$ | $$P(x)$$ |
| 0 | 0.40 |
| 1 | 0.35 |
| 2 | 0.19 |
| 3 | 0.06 |

Construct a probability histogram. Center each bar at a random variable value and make the height of the bar the random variable value’s probability.

1. Find the mean and standard deviation of the random variable.