6.1 Discrete and Continuous Random Variables

**Random Variable and Probability Distribution**

A **random variable** takes numerical values that describe the outcomes of some chance process. The **probability distribution** of a random variable gives its possible values and their probabilities.

Consider tossing a fair coin 3 times.

Define X = the number of heads obtained

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Value** | **0** | **1** | **2** | **3** |
| **Probability** |   |   |   |   |



**Discrete Random Variables**

There are two main types of random variables: *discrete* and *continuous*. If we can find a way to list all possible outcomes for a random variable and assign probabilities to each one, we have a **discrete random variable**.

A **discrete random variable** *X* takes a fixed set of possible values with gaps between. The probability distribution of a discrete random variable *X* lists the values *xi* and their probabilities *pi*:

 **Value**: *x*1 *x*2 *x*3 …

 **Probability**: *p*1 *p*2 *p*3 …

The probabilities *pi* must satisfy two requirements:

1. Every probability *pi* is a number between 0 and 1.
2. The sum of the probabilities is 1.

To find the probability of any event, add the probabilities *pi* of the particular values *xi* that make up the event.

Example: Babies’ Health at Birth

The Apgar scale is used to rate a baby’s health at birth. It is based on 5 criteria: skin color, heart rate, muscle tone, breathing and response to stimulation. Each criteria is scored as a 0, 1 or 2 and the Apgar score is the total, with a range of 0 to 10.

The following table represents the scores of over 2 million babies:

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Value:** | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** |
| **Probability:** | 0.001 | 0.006 | 0.007 | 0.008 | 0.012 | 0.020 | 0.038 | 0.099 | 0.319 | 0.437 | 0.053 |

1. Show that the probability distribution for *X* is legitimate.
2. Make a histogram of the probability distribution. Describe what you see.
3. Apgar scores of 7 or higher indicate a healthy baby. What is *P*(*X* ≥ 7)?

**Mean of a Discrete Random Variable**

Suppose that *X* is a discrete random variable whose probability distribution is

 **Value**: *x*1 *x*2 *x*3 …

 **Probability**: *p*1 *p*2 *p*3 …

To find the **mean (expected value)** of *X*, multiply each possible value by its probability, then add all the products:



Example: **Probability distribution for the sales department of a manufacturing company**



Compute the mean of the random variable Number of Orders per Week.

**Standard Deviation of a Discrete Random Variable**

Since we use the mean as the measure of center for a discrete random variable, we’ll use the standard deviation as our measure of spread. The definition of the **variance of a random variable** is similar to the definition of the variance for a set of quantitative data.

Suppose that *X* is a discrete random variable whose probability distribution is

 **Value**: *x*1 *x*2 *x*3 …

 **Probability**: *p*1 *p*2 *p*3 …

and that  is the mean of *X*. The **variance** of *X* is



To get the **standard deviation of a random variable**, take the square root of the variance.

Example:

Compute the standard deviation of the random variable Number of Orders per Week.*X*

**Using the calculator**

In Lists & Spreadsheets, put the X values in Column A and the corresponding probabilities in column B.

Label both columns

Menu – Statistics – Stat Calculations – One Variable Statistics – OK (for 1 list)

X1 List – Label for column A; Frequency List – Label for column B; 1st Result Column – c[]; OK

**Continuous Random Variables**

A **continuous random variable** *X* takes on all values in an interval of numbers. The probability distribution of *X* is described by a **density curve**. The probability of any event is the area under the density curve and above the values of *X* that make up the event.

A continuous random variable *Y* has *infinitely many* possible values. All continuous probability models assign probability 0 to every individual outcome. Only *intervals* of values have positive probability.

Example:

Define *Y* as the height of a randomly chosen young woman. *Y* is a continuous random variable whose probability distribution is *N*(64, 2.7).

What is the probability that a randomly chosen young woman has height between 68 and 70 inches?



***P*(68 ≤ *Y* ≤ 70) = ???**