Lecture Slides

Essentials of Statistics
5th Edition

and the Triola Statistics Series

by Mario F. Triola

Chapter 5
Probability Distributions

X	f
0	20
1	2
2	3
3	2
4	4
5	18
6	5
7	4
8	6
9	6

X	f
0	7
1	6
2	6
3	7
4	8
5	6
6	7
7	9
8	8
9	6

Key Concept

This section introduces the important concept of a probability distribution, which gives the probability for each value of a variable that is determined by chance.

Give consideration to distinguishing between outcomes that are likely to occur by chance and outcomes that are "unusual" in the sense they are not likely to occur by chance.

Random Variable Probability Distribution

❖ Random Variable
a variable (typically represented by *x*) that has a single numerical value, determined by chance, for each outcome of a procedure

❖ Probability Distribution
a description that gives the probability for each value of the random variable, often expressed in the format of a graph, table, or formula

Discrete and Continuous Random Variables

❖ **Discrete Random Variable**
either a finite number of values or countable number of values, where "countable" refers to the fact that there might be infinitely many values, but that they result from a counting process

❖ **Continuous Random Variable**
has infinitely many values, and those values can be associated with measurements on a continuous scale without gaps or interruptions.

Probability Distribution: Requirements

1. There is a numerical random variable *x* and its values are associated with corresponding probabilities.
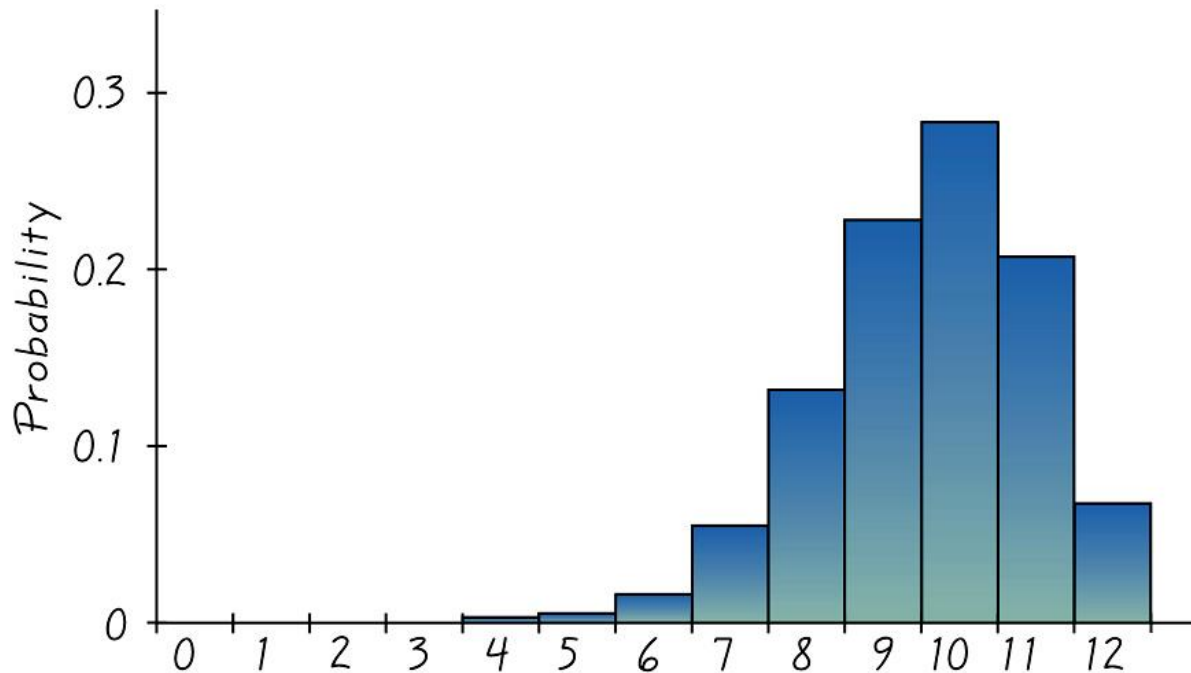
2. The sum of all probabilities must be 1.

$$\sum P(x) = 1$$

3. Each probability value must be between 0 and 1 inclusive.

$$0 \le P(x) \le 1$$

Graphs

The probability histogram is very similar to a relative frequency histogram, but the vertical scale shows probabilities.

Probability Histogram for Number of Mexican-American Jurors Among 12

Mean, Variance and Standard Deviation of a Probability Distribution

$$\mu = \Sigma[x \cdot P(x)] \qquad \text{Mean}$$

$$\sigma^2 = \Sigma[(x - \mu)^2 \cdot P(x)] \qquad \text{Variance}$$

$$\sigma^2 = \Sigma[(x^2 \cdot P(x)] - \mu^2 \qquad \text{Variance (shortcut)}$$

$$\sigma = \sqrt{\Sigma[(x^2 \cdot P(x)] - \mu^2} \qquad \text{Standard Deviation}$$

Round-off Rule for μ, σ, and σ² from a Probability Distribution

- **Round results at the END of the calculations, rounding in the middle can/will introduce errors.**

- **Round results by carrying one more decimal place than the number used for the random variable *x*.**

- **If the values of *x* are integers then round μ, σ, and σ² to one decimal place.**

Expected Value

The expected value of a discrete random variable is denoted by *E*, and it represents the mean value of the outcomes. It is obtained by finding the value of

$$\Sigma[x \cdot P(x)].$$

$$E = \Sigma[x \cdot P(x)]$$

Example

The following table describes the probability distribution for the number of girls in two births.

Find the mean, variance, and standard deviation.

x	$P(x)$		
0	0.25		
1	0.50		
2	0.25		
Total			

PEARSON

Example

The following table describes the probability distribution for the number of girls in two births.

Find the mean, variance, and standard deviation.

x	$P(x)$	$x \cdot P(x)$	$(x-\mu)^2 \cdot P(x)$
0	0.25	0.00	0.25
1	0.50	0.50	0.00
2	0.25	0.50	0.25
Total		1.00	0.50

Example

The following table describes the probability distribution for the number of girls in two births.

Find the mean, variance, and standard deviation.

Mean:

$$\mu = \sum[x \cdot P(x)] = 1.0$$

Variance:

$$\sigma^2 = \sum[(x - \mu)^2 \cdot P(x)] = 0.5$$

Standard Deviation: $\sqrt{\sigma^2} = 0.707107 \approx 0.7$

Identifying *Unusual* Results
Range Rule of Thumb

According to the range rule of thumb, most values should lie within 2 standard deviations of the mean.

We can therefore identify "unusual" values by determining if they lie outside these limits:

$$\text{Maximum usual value} = \mu + 2\sigma$$

$$\text{Minimum usual value} = \mu - 2\sigma$$

Example – continued

We found for families with two children, the mean number of girls is 1.0 and the standard deviation is 0.7 girls.

Use those values to find the maximum and minimum usual values for the number of girls.

$$\text{maximum usual value} = \mu + 2\sigma = 1.0 + 2(0.7) = 2.4$$

$$\text{minimum usual value} = \mu - 2\sigma = 1.0 - 2(0.7) = -0.4$$

Identifying Unusual Results Probabilities

Rare Event Rule for Inferential Statistics

If, under a given assumption (such as the assumption that a coin is fair), the probability of a particular observed event (such as 992 heads in 1000 tosses of a coin) is extremely small, we conclude that the assumption is probably not correct.

Identifying Unusual Results Probabilities

Using Probabilities to Determine When Results Are Unusual

❖ Unusually high: *x* successes among *n* trials is an unusually high number of successes if $P(x \text{ or more}) \leq 0.05$.

❖ Unusually low: *x* successes among *n* trials is an unusually low number of successes if $P(x \text{ or fewer}) \leq 0.05$.

Expected Value in Decision Theory

You get to bet $5 in the Golden Nugget casino in Las Vegas. You have two choices:

- **Roulette: bet on 7**
- **Craps: bet on the "pass" line.**

Which bet is better?

Roulette	X	P(x)	x·P(x)
Lose	-$5	37/38	-$4.8684
Win(net)	$175	1/38	$4.6053
Total			≈-$0.26

Craps	X	P(x)	x·P(x)
Lose	-$5	251/495	-$2.5354
Win(net)	$5	244/495	$2.4646
Total			≈-$0.07