Sections 5-1

Probability Distributions

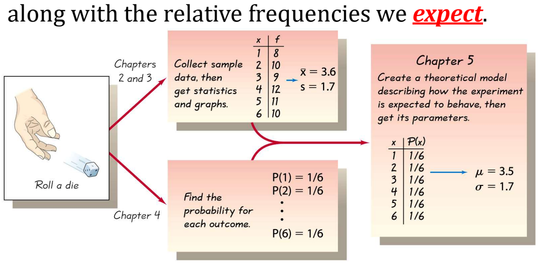
PROBABILITY DISTRIBUTIONS

This chapter will deal with the construction of **probability distributions** by combining the methods of Chapters 2 and 3 with the those of Chapter 4.

Probability Distributions will describe what will ***probably*** happen instead of what actually ***did*** happen.

COMBINING DESCRIPTIVE METHODS AND PROBABILITIES

In this chapter we will construct probability distributions by presenting possible outcomes along with the relative frequencies we ***expect***.



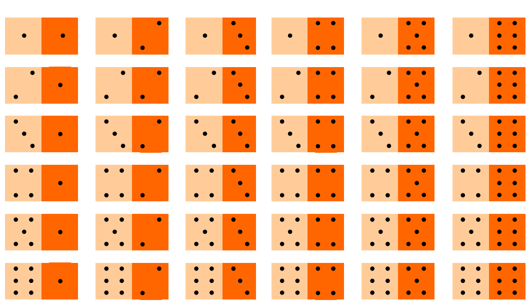
RANDOM VARIABLES AND PROBABILITY DISTRIBUTIONS

- A **random variable** is a variable (typically represented by x) that has a single numerical value, determined by chance, for each outcome of a procedure.
- A **probability distribution** is a description that gives the probability for each value of a random variable. It is often expressed in the format of a table, formula, or graph.

EXAMPLES

1. Suppose you toss a coin three times. Let x be the total number of heads. Make a table for the probability distribution of x.
2. Suppose you throw a pair of dice. Let x be the sum of the numbers on the dice. Make a table for the probability distribution of x.

SAMPLE SPACE FOR ROLLING A PAIR OF DICE



DISCRETE AND CONTINUOUS RANDOM VARIABLES

- A **discrete random variable** has a collection of values that is finite or countable. (If there are infinitely many values, the number of values is countable if it is possible to count them individually, such as the number of tosses of a coin before getting tails.)
- A **continuous random variable** has infinitely many values, and the collection of values is not countable. (That is, it is impossible to count the individual items because at least some of them are on a continuous scale.

EXAMPLES

Determine whether the following are discrete or continuous random variables.

1. Let x be the number of cars that travel through McDonald's drive-through in the next hour.
2. Let x be the speed of the next car that passes a state trooper.
3. Let x be the number of *A*s earned in a section of statistics with 15 students enrolled.

PROBABILITY HISTORGRAM

A **probability histogram** is like a relative frequency histogram with ***probabilities*** instead of relative frequencies.

EXAMPLES

1. Suppose you toss a coin three times. Let x be the total number of heads. Draw a probability histogram for x.
2. Suppose you throw a pair of dice. Let x be the sum of the numbers on the dice. Draw a probability histogram for x.

REQUIREMENTS FOR A PROBABILITY DISTRIBUTION

1. There is a ***numerical*** (not categorical) random variable x, and its values are associated with corresponding probabilities.
2. $\sum P(x) = 1$ where x assumes all possible values. The sum of all probabilities must be 1, but sums near 1 (like 0.999 or 1.002) are acceptable because they result from rounding errors.
3. $0 \leq P(x) \leq 1$ for every individual value of x. That is, each probability value must be between 0 and 1 inclusive.

EXAMPLES

Determine if the following are probability distributions

(a)

x	$P(x)$
1	0.20
2	0.35
3	0.12
4	0.40
5	−0.07

(b)

x	$P(x)$
1	0.20
2	0.25
3	0.10
4	0.14
5	0.49

(c)

x	$P(x)$
1	0.20
2	0.25
3	0.10
4	0.14
5	0.31

MEAN, VARIANCE, AND STANDARD DEVIATION

Mean of a Prob. Dist. $\mu = \sum[x \cdot P(x)]$

Variance of a Prob. Dist. $\sigma^2 = \sum[(x-\mu)^2 \cdot P(x)]$

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

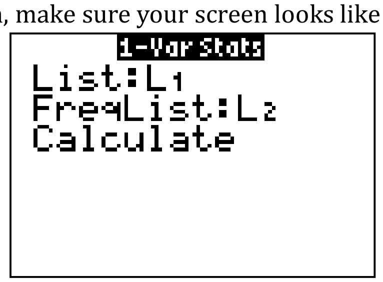
Standard Deviation of a Prob. Dist. $\sigma = \sqrt{\sum[x^2 \cdot P(x)] - \mu^2}$

FINDIND MEAN, VARIANCE, AND STANDARD DEVIATION WITH TI-83/84 CALCULATOR

1. Enter values for random variable in L_1.
2. Enter the probabilities for the random variables in L_2.
3. Run “1-VarStat L_1, L_2”
4. The mean will be $\bar{x}$. The standard deviation will be σx. To get the variance, square σx.

TI-84 WITH NEW OPERATING SYSTEM

If you have the TI-84 with the newest operating system, make sure your screen looks like this:



ROUND-OFF RULE FOR μ, σ, AND σ^2

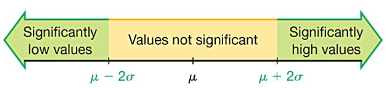
Round results by carrying one more decimal place than the number of decimal places used for the random variable x.

IDENTIFYING SIGNIFICANT VALUES

Significantly low values are $(\mu - 2\sigma)$ or lower.

Significantly high values are $(\mu + 2\sigma)$ or higher.

Values that are not significant are between $(\mu - 2\sigma)$ and $(\mu + 2\sigma)$.



EXAMPLE

Use the range rule of thumb to determine the significant values for rolling a pair of dice.

IDENTIFYING SIGNIFICANT RESULTS USING PROBABILITIES

- **Significantly high number of successes**: x successes among n trials is a ***significantly high*** number of successes if the probability of x or more successes is 0.05 or less. That is, $P(x \text{ or more}) \leq 0.05$.
- **Significantly low number of successes**: x successes among n trials is a ***significantly low*** number of successes if the probability of x or fewer successes is 0.05 or less. That is, $P(x \text{ or fewer}) \leq 0.05$.
- NOTE: The value 0.05 is not absolutely rigid. Other values, such as 0.01, could be used to distinguish between results that are significant and those that are not significant.

EXAMPLE

Consider the procedure of rolling a pair of dice five times and letting x be the number of times that "7" occurs. The table below describes the probability distribution.

x	$P(x)$
0	0.402
1	0.402
2	?
3	0.032
4	0.003
5	0+

(a) Find the value of the missing probability.

(b) Would 3 be a significantly high number of "7s" when a pair of dice is rolled?

RARE EVENT RULE FOR INFERENTIAL STATISTICS

If, under a given assumption, the probability of a particular outcome is very small and the outcome occurs *significantly less* than or *significantly greater* than what we expect with that assumption, we conclude that the assumption is probably not correct.

EXPECTED VALUE

The **expected value** of a discrete random variable x is denoted by E, and it is the mean value of the outcomes, so $E = \mu$ and E can also be found by evaluating $\sum[x \cdot P(x)]$.

That is,

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

EXAMPLE

When you give the Venetian casino in Las Vegas $5 for a bet on the number 7 in roulette, you have 37/38 probability of losing $5 and you have a 1/38 probability of making a net gain of $175. (The prize in $180, including you $5 bet, so the net gain is $175.) If you bet $5 that the outcome is an odd number the probability of losing $5 is 20/38 and probability of making a net gain of $5 is 18/38. (If you bet $5 on an odd number and win, you are given $10 that included your bet, so the net gain is $5.)

(a) If you bet $5 on the number 7, what is your expected value?
(b) If you bet $5 that the outcome is an odd number, what is your expected value?
(c) Which of these options is best: bet on 7, bet on an odd number, or don't bet? Why?