

## Hw503

A coin is thrown 10 times. Find the probability density function for X, where X is the random variable representing the number of heads obtained.

In the above example, what is the expected number of heads thrown

A die is thrown repeatedly until a 6 is obtained. Find the probability density function for the number times we throw the die.

Given three real numbers, its arithmetic mean (average), geometric mean and harmonic mean are defined as follows:

$$\begin{aligned}\text{Arithmetic mean} &= \frac{1}{3}(x + y + z) \\ \text{Geometric mean} &= (x \cdot y \cdot z)^{1/3} \\ \text{Harmonic mean} &= \frac{3}{\frac{1}{x} + \frac{1}{y} + \frac{1}{z}}\end{aligned}$$

Write a program to compute and display the means of three **REAL** variables initialized with positive real values.

The radius of a circle is  $x = (3.0 \pm 0.2)$  cm. Find the circumference and its uncertainty.

$$C = 2 \pi x = 18.850 \text{ cm}$$

$$\Delta C = 2 \pi \Delta x = 1.257 \text{ cm (The factors of 2 and } \pi \text{ are exact)}$$

$$C = (18.8 \pm 1.3) \text{ cm}$$

We round the uncertainty to two figures since it starts with a 1, and round the answer to match

:  $x = (2.0 \pm 0.2)$  cm,  $y = (3.0 \pm 0.6)$  cm. Find  $z = x - 2y$  and its uncertainty.

$$z = x - 2y = 2.0 - 2(3.0) = -4.0 \text{ cm}$$

$$\Delta z = \Delta x + 2 \Delta y = 0.2 + 1.2 = 1.4 \text{ cm}$$

$$\text{So } z = (-4.0 \pm 1.4) \text{ cm.}$$

Using Eq 1b,  $z = (-4.0 \pm 0.9)$  cm.

The 0 after the decimal point in 4.0 is significant and must be written in the answer. The uncertainty in this case starts with a 1 and is kept to two significant figures. (More on rounding in Section 7.)

$w = (4.52 \pm 0.02)$  cm,  $x = (2.0 \pm 0.2)$  cm. Find  $z = w x$  and its uncertainty.  
 $z = w x = (4.52)(2.0) = 9.04 \text{ cm}^2$

Round off  $z = 12.0349$  cm and  $\Delta z = 0.153$  cm.

Since  $\Delta z$  begins with a 1, we round off  $\Delta z$  to two significant figures:

$\Delta z = 0.15$  cm. Hence, round  $z$  to have the same number of decimal places:

$$z = (12.03 \pm 0.15) \text{ cm.}$$

Suppose we repeat the dice tossing experiment described in Example 1. This time, we ask what is the probability that the die will land on a number that is smaller than 5?

Suppose we conduct a simple [statistical experiment](#). We flip a coin one time. The coin flip can have one of two outcomes - heads or tails. Together, these outcomes represent the sample space of our experiment. Individually, each outcome represents a sample point in the sample space. What is the probability of each sample point?

1. A student performs an experiment to determine the specific heat of a sample of metal. 212.01 g of the metal at 95.5°C was placed into 150.25 g of 25.2 °C water in the calorimeter. The temperature of the water went to 27.5°C. Given:  $C_{H_2O} = 4.18 \text{ J/g-}^\circ\text{C}$ . The thermometer was marked in 1 °C increments and the balance was digital.
  - a. Calculate the specific heat of the metal  $C_m$  using the following equation:

$$m_{H_2O} C_{H_2O} (T_f - T_i) = -m_m C_m (T_f - T_i)$$

$$150.25(4.18)(27.5 - 25.2) = -212.01(C_m)(27.5 - 95.5)$$

$$= .1002 \text{ J / g}^\circ\text{C}$$

### Example:

Find the sample space of rolling two dices and tossing three coins.

### Example:

Find the sample space of rolling two dices and tossing three coins.

**Answer:**

**Step 1: Find the elements of each event.**

elements of each dice = {1, 2, 3, 4, 5, 6}

elements of each coin = {H, T}

**Step 2: Find the sample space by applying the first fundamental counting principle.**

sample space =  $6 \times 6 \times 2 \times 2 \times 2 = 228$

**A zero is said to be significant if:**

(1) it is between two non-zero digits	<b>3001 m, 30.001 m</b>	4 SD, 5 SD
(2) it is at the end of a decimal expression	0.00 <b>310</b> km	3 SD
(3) it is required when expressing the number in scientific notation	<b>3.10</b> $\times 10^6$ m	3 SD

**Otherwise a zero is considered to be only a placeholder.**

150,000 m	all four zeros are placeholders	$1.5 \times 10^5$ m
0.0015 km	all three zeros are placeholders	$1.5 \times 10^{-3}$ km
150. Gm	no zeros are placeholders (note the deliberate inclusion of the decimal)	$1.50 \times 10^{11}$ m

When **multiplying or dividing two measurements**, your answer should be rounded off so that it only has accurate as many significant digits as your least accurate original value.

When **adding or subtracting two measurements**, first convert them to the same unit of measurement, then line up the decimals. Your final answer should be rounded off so that it only has as many decimal places as your least accurate original value.

Numerical constants ( $\pi$ ,  $e$ ,  $\frac{1}{2}$ ) do not have significant digits.

For example, the volume of sphere is calculated with the formula  $V = \frac{4}{3} \pi r^3$ .

Using this formula, a sphere with a measured diameter of 24 cm would have a volume equal to

$$\frac{4}{3} \pi (12)^3 = \frac{4}{3} \pi (1728) = 2304\pi \text{ cm}^3$$

Since 12 only had two significant digits, your final value for the sphere's volume should only have 2 SD.

This means that a calculated value of  $V = 7238.229 \text{ cm}^3$  should be expressed in final form as  $7200 \text{ cm}^3 = 7.2 \times 10^3 \text{ cm}^3$

**Scientific Notation.** Express your value so that it has one digit to the left of the decimal and all other significant digits to the right of the decimal. It should then be multiplied by an appropriate power of 10.

(1) When the absolute value of the original number is greater than one, then moving the decimal point will require the resulting number to be multiplied by 10 raised to a positive exponent.

(2) When the absolute value of the original number is less than one, then moving the decimal point will require the resulting number to be multiplied by 10 raised to a negative exponent.

**where the decimal was moved ....**

0.066 g	$6.6 \times 10^{-2}$ g	$ 0.066  < 1$	two decimal places to the right	0's are placeholders
200.0 g	$2.000 \times 10^2$ g	$ 200.0  > 1$	two decimal places to the left	all three zeros are significant
0.543 g	$5.43 \times 10^{-1}$ g	$ 0.543  < 1$	one decimal place to the right	0 is a placeholder
1600 g	$1.6 \times 10^3$ g	$ 1600  > 1$	three decimal places to the left	both zeros are placeholders
75.2 g	$7.52 \times 10^1$ g	$ 75.2  > 1$	one decimal place to the left	-----

## SIGNIFICANT DIGITS

The number of significant digits in an answer to a calculation will depend on the number of significant digits in the given data, as discussed in the rules below. *Approximate* calculations (order-of-magnitude estimates) always result in answers with only one or two significant digits.

### When are Digits Significant?

Non-zero digits are always significant. Thus, 22 has two significant digits, and 22.3 has three significant digits.

With zeroes, the situation is more complicated:

- a. Zeroes placed before other digits are not significant; 0.046 has two significant digits.
- b. Zeroes placed between other digits are always significant; 4009 kg has four significant digits.
- c. Zeroes placed after other digits but behind a decimal point are significant; 7.90 has three significant digits.
- d. Zeroes at the end of a number are significant only if they are behind a decimal point as in (c). Otherwise, it is impossible to tell if they are significant. For example, in the number 8200, it is not clear if the zeroes are significant or not. The number of significant digits in 8200 is at least two, but could be three or four. To avoid uncertainty, use scientific notation to place significant zeroes behind a decimal point:

$8.200 \times 10^3$  has four significant digits

$8.20 \times 10^3$  has three significant digits

$8.2 \times 10^3$  has two significant digits

### Significant Digits in Multiplication, Division, Trig. functions, etc.

In a calculation involving multiplication, division, trigonometric functions, etc., the number of significant digits in an answer should equal the least number of significant digits in any one of the numbers being multiplied, divided etc.

Thus in evaluating  $\sin(kx)$ , where  $k = 0.097 \text{ m}^{-1}$  (two significant digits) and  $x = 4.73 \text{ m}$  (three significant digits), the answer should have two significant digits.

Note that whole numbers have essentially an unlimited number of significant digits. As an example, if a hair dryer uses 1.2 kW of power, then 2 identical hairdryers use 2.4 kW:

$$1.2 \text{ kW} \{2 \text{ sig. dig.}\} \times 2 \{\text{unlimited sig. dig.}\} = 2.4 \text{ kW} \{2 \text{ sig. dig.}\}$$

### Significant Digits in Addition and Subtraction

When quantities are being added or subtracted, the number of *decimal places* (not significant digits) in the answer should be the same as the least number of decimal places in any of the numbers being added or subtracted.

Example:

5.67 J (two decimal places)

1.1 J (one decimal place)

0.9378 J (four decimal place)

7.7 J (one decimal place)

### Keep One Extra Digit in Intermediate Answers

When doing multi-step calculations, *keep at least one more significant digit in intermediate results* than needed in your final answer.

For instance, if a final answer requires two significant digits, then carry at least three significant digits in calculations. If you round-off all your intermediate answers to only two digits, you are discarding the information contained in the third digit, and as a result the *second* digit in your final answer might be incorrect. (This phenomenon is known as "round-off error.")

### The Two Greatest Sins Regarding Significant Digits

1. Writing more digits in an answer (intermediate or final) than justified by the number of digits in the data.
2. Rounding-off, say, to two digits in an intermediate answer, and then writing three digits in the final answer.

### Try these Exercises:

1.  $e^{kt} = ?$ , where  $k = 0.0189 \text{ yr}^{-1}$ , and  $t = 25 \text{ yr}$ .
  2.  $ab/c = ?$ , where  $a = 483 \text{ J}$ ,  $b = 73.67 \text{ J}$ , and  $c = 15.67$
  3.  $x + y + z = ?$ , where  $x = 48.1$ ,  $y = 77$ , and  $z = 65.789$
  4.  $m - n - p = ?$ , where  $m = 25.6$ ,  $n = 21.1$ , and  $p = 2.43$
- 
1.  $e^{kt} = ?$ , where  $k = 0.0189 \text{ yr}^{-1}$ , and  $t = 25 \text{ yr}$ . [Ans. 1.6]
  2.  $ab/c = ?$ , where  $a = 483 \text{ J}$ ,  $b = 73.67 \text{ J}$ , and  $c = 15.67$  [Ans.  $2.27 \times 10^3 \text{ J}^2$ ]
  3.  $x + y + z = ?$ , where  $x = 48.1$ ,  $y = 77$ , and  $z = 65.789$  [Ans. 191]
  4.  $m - n - p = ?$ , where  $m = 25.6$ ,  $n = 21.1$ , and  $p = 2.43$  [Ans. 2.1]

## Measurements and Significant Digits

The valid digits in a measurement are called significant digits. Measurements are made with measuring tools that have a calibrated scale or digital LED display for determining the quantity to be measured. A meterstick is divided into a thousand divisions called millimeters. Between two adjacent millimeter marks there are no additional calibration marks. One has to estimate the measurement value between the two millimeter marks. The precision of a meterstick is 0.5 millimeters. The length of a block of wood is measured to be 32.47 cm or 324.7 mm. The number seven, the last digit in the measurement is an estimated value found between two adjacent millimeter marks. The estimated value is a significant digit in this measured quantity for the wood's length. Thus, the measurement has four significant digits.

The last digit given for any measurement is the uncertain or estimated digit. It is uncertain because it is estimated. The last digit given for a measurement is always assumed to be estimated and it is significant. **All nonzero digits in a measurement are significant (1,2,3,4,5,6,7,8,9). Zeroes are only sometimes significant when reported in a measurement.**

### When are Zeroes in a measurement Significant?

Not all zeroes are significant. For example, if you reported the measured length of an object to be 0.0030400 m, it would have only five significant digits. The first three zeroes only serve to locate the decimal point and are not significant. The other three zeroes are significant. The last zero is the estimated digit and it is significant. The other two are between two significant digits (3, and the last zero) and will always be significant in this position.

The significant zeroes in these measurements are colored black and the insignificant zeroes are red. The total number of significant digits is the sum of all the black digits in the measurements

**0.0860 m**                      **1.0030 s**                      **0.000010203 m**

**\$18,000**                      **\$18,000.00**                      **\$18,000.**

**0.10001 cm**

The following rules summarize how to determine the number of significant digits.

- Nonzero digits are always significant
- All final zeroes after a decimal point are significant
- Zeroes between two other significant digits are always significant
- Zeroes used solely as placeholders are NOT significant
- Zeroes between a decimal point and a nonzero digit are significant.

The number of significant digits in a measurement is an indication of the precision with which the measurement was taken.

**Practice Problems:** State the number of significant digits in each measurement. (Record these answers in your notebook under the heading "Significant Digits, Part A".)

- 1) **2804 m**                      2) **2.84 km**                      3) **0.029 m**  
4) **0.003068 m**                      5) **4.6 x 10<sup>5</sup> m**                      6) **4.06 x 10<sup>-5</sup> m**  
7) **750 m**                      8) **75 m**                      9) **75,000 m**  
10) **75,000. m**                      11) **75,000.0 m**                      12) **10 cm**

---

## Arithmetic with Significant Digits

In physics you are required to record all experimental data with the correct number of significant digits. Often, you will be required to add, subtract, divide, and multiply these measurements. **When you perform any arithmetic operation, it is important to remember that the result can never be more precise than the least precise measurement.**

To add or subtract measurements, first perform the operation, then round off the result to correspond to the least precise value involved. For example, add these values:

$$24.686 \text{ m} + 2.343 \text{ m} + 3.21 \text{ m} = \mathbf{30.239 \text{ m}}$$

However, 3.21 m is the least precise value - accurate to the hundredth of a meter. The above answer should be report with the same amount of precision.

This requires you to round-off the value, 30.239 m, to 30.24 m. You will report the correct calculated answer as **30.24 m**.

A different method is used to find the correct number of significant digits when multiplying or dividing measurements. After performing the calculation, note the factor that has the least number of significant digits. Round the product or quotient to this number of digits. For example, multiply

$$3.22 \text{ cm by } 2.1 \text{ cm} = 6.762 \text{ cm}^2 \text{ corrected to } \underline{6.8 \text{ cm}^2}$$

Divide these two measurements and report answer with correct number of significant digits

$$36.5 \text{ m divided by } 3.414 \text{ s} = 10.691 \text{ m/s corrected to } \underline{10.7 \text{ m/s}}$$

### **Practice Problems:**

Solve the following problems and report answers with appropriate number of significant digits. (Write answers in your notebook under the heading "**Arithmetic and Significant Digit**".)

- 1) **6.201 cm + 7.4 cm + 0.68 cm + 12.0 cm = ?**
- 2) **1.6 km + 1.62 m + 1200 cm = ?**
- 3) **8.264 g - 7.8 g = ?**
- 4) **10.4168 m - 6.0 m = ?**
- 5) **12.00 m + 15.001 kg = ?**
- 6) **131 cm x 2.3 cm = ?**
- 7) **5.7621 m x 6.201 m = ?**
- 8) **20.2 cm divided by 7.41 s = ?**
- 9) **40.002 g divided by 13.000005 ml = ?**

---

## **Calculators and Significant Digits**

Many calculators display several additional, meaningless digits, some always display only two. Be sure to record your answer with the correct number of significant digits. **Calculator answers are not round to significant digits - you will have to round-off the answer to the correct number of digits.**

Note that significant digits are only associated with measurements; **there is no uncertainty associated with counting**. If you counted four laps for a runner and measured the time to be 2.34 minutes. The number of laps does not have an uncertainty, but the measured time does.

## SIGNIFICANT FIGURES TEST

The following test, by **Stephen L. Morgan**, will check whether or not you know how to use significant figures.

- |   |                   |
|---|-------------------|
| 1. $37.76 + 3.907 + 226.4 =$  | 268.1             |
| 2. $319.15 - 32.614 =$  | 286.54            |
| 3. $104.630 + 27.08362 + 0.61 =$  | 132.32            |
| 4. $125 - 0.23 + 4.109 =$   | 129               |
| 5. $2.02 \times 2.5 =$  | 5.0               |
| 6. $600.0 / 5.2302 =$   | 114.7             |
| 7. $0.0032 \times 273 =$  | 0.87              |
| 8. $(5.5)^3 =$  | $1.7 \times 10^2$ |
| 9. $0.556 \times (40 - 32.5) =$   | 4                 |
| 10. $45 \times 3.00 =$  | $1.4 \times 10^2$ |
| 11. What is the average of 0.1707, 0.1713, 0.1720, 0.1704, and 0.1715? Answer = | 0.1712            |
12. Calculate the sum of the squares of the deviations from the mean for the five numbers given in Question 11 above, in two different ways: (a) carrying all digits through all the calculations; (b) rounding all intermediate results to 2 significant figures after subtracting the mean from each (on your way to calculating the sum of the squares about the mean). Compare the two results. Answer: They are drastically

### Significant Figures Rules

*modified from a web page by Robert Zinn, Yale U.*

#### Quick Reference Section

This section presents the basic rules for significant figures. Read later sections to gain a complete understanding of what these rules really mean!

#### How does one tell how many significant digits there are in a given number?

- The left most digit (&larr) which is not a zero is the *most* significant digit.
- If the number *does not* have a decimal point, the right most digit (&rarr) which is not a zero is the *least* significant digit.
- If the number *does* have a decimal point, the right most significant digit ( $\rightarrow$ ) is the *least* significant digit, even if it's a zero.
- **Every digit between the least and most significant digits should be counted as a significant digit.**

**For example,**  
according to these rules, all of these numbers have **three significant digits**:

123,000  
 12.3  
 1.23 x 10<sup>6</sup>  
 1.00  
 0.000123

How many significant figures should one retain in the final answer to a problem?

- For results obtained using addition or subtraction, **the number of places after the decimal point in the result should be less than or equal to the number of decimal places in every term.**
- For results obtained using multiplication or division, **the number of significant figures in the result should be equal to the number of significant digits in the least precise number** (the number with the fewest significant digits given).

It is sometimes good practice to give one more significant figure than is required by these rules; this helps prevent rounding errors if the number is used in later calculations. This extra digit should be in a smaller font to indicate its lesser significance.

### Addition and Subtraction

When adding and subtracting numbers, the rules of significant figures require that *the number of places after the decimal point in the answer is less than or equal to the number of decimal places in every term in the sum.* (Treat subtraction as adding the same number with a negative sign in front of it.) If some of the numbers have no digits after the decimal point, use the same basic rule, but don't record any digits to the right of the last digit in the least significant number. Clarify these rules, are

some examples:

2355.2342	15600.00	15600	13.7	137000
+ 23.24	+ 172.49	+ 172.49	+ 1.3	+ 1330
-----	-----	-----	-----	-----
2378.47	15772.49	15800	15.0	138000

Note it is not unusual for a sum to have more significant figures than the measurements added. This is why finding an average gives greater information than a single measurement.

### Multiplication and Division

When multiplying and dividing numbers, *the number of significant digits you use is simply the same number of significant figures as is the number with the fewest significant figures.*

Some examples:

13.1	13.10	13.100	1500	15310	1.00
x 2.25	x 2.25	x 2.2500	x 2.315	x 2.3	x 10.04
-----	-----	-----	-----	-----	-----
29.5	29.5	29.475	3400	35000	10.0

Someday calculators may be able to do significant figures; but in the meantime the operators need that wisdom.

### Why do Multiplication and Addition Have Different Rules?

When you add two numbers, you add their uncertainties, more or less. If one of the numbers is smaller than the uncertainty of the other, it doesn't make much of a difference to the value (and hence, uncertainty) of the final result. Thus it is the location of the digits, not the amount of digits that is important.

When you multiply two numbers, you more or less multiply the uncertainties. Thus it is the percentage by which you are uncertain that is important -- the uncertainty in the number divided by the number itself. This is given roughly by the number of digits, regardless of their placement in terms of powers of ten. Hence the number of digits is what is important.

### Which Digits are "Significant?"

In order to figure out how many significant figures to put into your final answer you must figure out how many significant figures are in each of the numbers you are working with. The rules are best explained separately for fundamental constants, physical constants, numbers not ending in 0 and numbers ending in 0.

Now suppose you want to generate a Monte Carlo Simulation of a fair coin with probability about 0.5 to get head and 0.5 to get tail, you can use the one of following formulas

**Formula :**  $R = \text{int}(U + \frac{1}{2})$

Random number  $U$  that uniformly distributed is generated according to previous page. To put this formula in action, I have made a worksheet in MS Excel. You may download it here. The explanation of how to use the worksheet is given below.

Suppose you want to make card game that consists of 52 decks. Each of the cards has equal chance to be selected. Then you need to generate 52 random integers with uniform probability  $\{1, 2, \dots, 51, 52\}$ . Another example, you want to make a program that behaves as a dice with six faces  $\{1, 2, 3, 4, 5, 6\}$ . Each face has equal chance to be selected. In this section, you will learn general formula to make Monte Carlo Simulation by generating  $N$  integers from lower boundary  $L$  until upper boundary  $B$  with equal chance to be selected.

If you want to generate  $N = B - L + 1$  random integers  $\{L, L+1, L+2, \dots, B-2, B-1, B\}$  that will distribute according to uniform probability, the formula is

$$R = \text{int}((B - L + 1) \cdot U) + L$$

Random number  $U$  that uniformly distributed is generated according to previous page. To put this formula in action, I have made a worksheet in MS Excel. You may download it here. The explanation of how to use the worksheet is given below.

The expected mean is  $\mu_R = \frac{1}{N} (L + (L + 1) + (L + 2) + \dots + (L + N - 1))$

$$\text{Or, } \mu_R = \frac{1}{N} (L \cdot N + \frac{N(N-1)}{2}) = L + \frac{N-1}{2}$$

The expected variance is

$$\sigma_R^2 = \frac{1}{N} \sum_{k=L}^B (k - \mu_R)^2 = \frac{1}{N} \sum_{k=L}^B (k - L - \frac{N-1}{2})^2 = \frac{1}{N} \sum_{k=0}^{N-1} (k - \frac{N-1}{2})^2 \quad \text{or}$$

$$\sigma_R^2 = \left( \frac{N-1}{2} \right) \left( \frac{N+1}{6} \right) = \frac{1}{12} (N^2 - 1)$$

## The Binomial Distribution

If a **discrete random variable**  $X$  has the following **probability density function** (p.d.f.), it is said to have a **binomial distribution**:

- $P(X = x) = {}^n C_x q^{(n-x)} p^x$ , where  $q = 1 - p$

$p$  can be considered as the probability of a success, and  $q$  the probability of a failure.

Note:  ${}^n C_r$  (“**n choose r**”) is more commonly written  $\binom{n}{r}$ , but I shall use the former because it is easier to write on a computer. It means the number of ways of choosing  $r$  objects from a collection of  $n$  objects (see **permutations and combinations**).

If a random variable  $X$  has a binomial distribution, we write  $X \sim B(n, p)$  ( $\sim$  means ‘has distribution...’).

$n$  and  $p$  are known as the **parameters** of the distribution ( $n$  can be any integer greater than 0 and  $p$  can be any number between 0 and 1). All random variables with a binomial distribution have the above p.d.f., but may have different parameters (different values for  $n$  and  $p$ ).

## Example

A coin is thrown 10 times. Find the probability density function for  $X$ , where  $X$  is the random variable representing the number of heads obtained.

The probability of throwing a head is  $\frac{1}{2}$  and the probability of throwing a tail is  $\frac{1}{2}$ . Therefore, the probability of throwing 8 tails is  $(\frac{1}{2})^8$

If we throw 2 heads and 8 tails, we could have thrown them HTTTTTHTT, or TTHHTTTTTT, or in a number of other ways. In fact, the total number of ways of throwing 2 heads and 8 tails is  ${}^{10}C_2$  (see the [permutations and combinations](#) section).

Hence the probability of throwing 2 heads and 8 tails is  ${}^{10}C_2 \times (\frac{1}{2})^2 \times (\frac{1}{2})^8$ . As you can see this has a Binomial distribution, where  $n = 10$ ,  $p = \frac{1}{2}$ .

You can see, therefore, that the p.d.f. is going to be:  $P(X = x) = {}^{10}C_x (\frac{1}{2})^{(10-x)} (\frac{1}{2})^x$ .

From this, we can work out the probability of throwing, for example, 3 heads (put  $x = 3$ ).

## Expectation and Variance

If  $X \sim B(n,p)$ , then the [expectation and variance](#) is given by:

- $E(X) = np$
- $\text{Var}(X) = npq$

## Example

In the above example, what is the expected number of heads thrown?

$$E(X) = np$$

Now in the above example,  $p =$  probability of throwing a head  $= \frac{1}{2}$ .  $n =$  number of throws  $= 10$

Hence expected number of heads  $= 5$ .

This is what you would expect: if you throw a coin 10 times you would expect 5 heads and 5 tails on average.

## Mean and Variance of the Binomial Distribution

The binomial distribution for a random variable  $X$  with parameters  $n$  and  $p$  represents the sum of  $n$  independent variables  $Z$  which may assume the values 0 or 1. If the probability that each  $Z$  variable assumes the value 1 is equal to  $p$ , then the [mean](#) of

each variable is equal to  $1 \cdot p + 0 \cdot (1-p) = p$ , and the **variance** is equal to  $p(1-p)$ . By the addition properties for independent random variables, the mean and variance of the binomial distribution are equal to the sum of the means and variances of the  $n$

$$\mu_X = np$$

$$\sigma_X^2 = np(1-p)$$

independent  $Z$  variables, so

These definitions are intuitively logical. Imagine, for example, 8 flips of a coin. If the coin is fair, then  $p = 0.5$ . One would expect the mean number of heads to be half the flips, or  $np = 8 \cdot 0.5 = 4$ . The variance is equal to  $np(1-p) = 8 \cdot 0.5 \cdot 0.5 = 2$ .

### **Example**

A die is thrown repeatedly until a 6 is obtained. Find the probability density function for the number times we throw the die.

Let  $X$  be the random variable representing the number of times we throw the die.

$P(X = 1) = 1/6$  (if we only throw the die once, we get a 6 on our first throw. The probability of this is  $1/6$ ).

$P(X = 2) = (5/6) \times (1/6)$  (if we throw the die twice before getting a 6, we must throw something that isn't a 6 with our first throw, the probability of which is  $5/6$  and we must throw a 6 on our second throw, the probability of which is  $1/6$ )

etc

In general,  $P(X = x) = (5/6)^{(x-1)} \times (1/6)$

### ***LISTING ALL POSSIBLE OUTCOMES (THE SAMPLE SPACE)***

---

Suppose that we will observe some process or experiment in which the outcome is not known in advance. For example, suppose we plan to roll two dice and we're interested in the sum of the two numbers appearing on the top faces. Before we can talk about probabilities of various sums, say 3 or 7, we have to understand what outcomes are possible in this experiment.

If we roll two dice, each die could show 1, 2, 3, 4, 5, 6. So the sum of the two faces could be any whole number from 2 to 12. We call this set of possible outcomes in the random experiment the *sample space*. Here the sample space can be written as

$$\text{Sample space} = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$$

Let's consider the set of all possible outcomes for other basic random experiments. Suppose we plan to toss a coin 3 times and the outcome of interest is the number of heads. The sample space in this case is the different numbers of heads you could get if you toss a coin three times. Here you could get 0 heads, 1 heads, 2 heads or 3 heads, so we write the sample space as

$$\text{Sample space} = \{0, 1, 2, 3\}$$

Don't forget to include the outcome 0 -- if we toss a coin three times and get all tails, then the number of heads is equal to 0.