ST101 Unit 4: Outliers and Normal Distribution

Contents:

Outliers

Ignoring Data

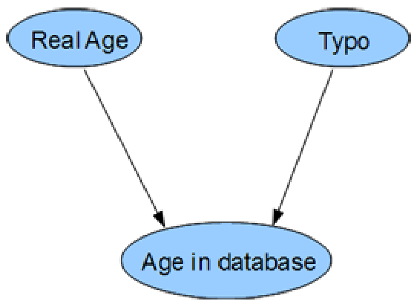
We all know that many politicians and cult leaders happily ignore data when it suits them to do so. But did you know that statisticians also ignore data?

Consider the example of a sports club with the following members:

Name	**Age**
Joseph	22
Maria	21
Susan	24
Marc	20
Tom	211
Jack	23

If we wanted to compute the mean age of this group, should we ignore any of the data?

And in this case the answer is yes. Tom's age is obviously a mistake. It was probably the result of a typo when entering the data onto a computer. The ages in our database can all be explained as either the real ages of members or as typos:



So how do we identify, and so ignore, these outlying data points?

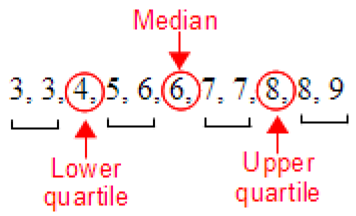
Quartiles

The easiest way to ignore outliers is by using quartiles, or percentiles.

Suppose we have the following dataset of 11 items sorted into numerical order:

3, 3, 4, 5, 6, 6, 7, 7, 8, 8, 9

As the name suggests, quartiles partition the data into four subsets, with a single digit gap between each subset:



The element in the middle of the sorted group is the median, which we met in the last unit. The other two elements are called the lower quartile and the upper quartile. The range between these elements is called the **inter-quartile range**. It is this range that we would use to calculate things like the mean. Data falling outside this range is then ignored.

This is a simple, but often very effective technique for removing outliers. It will remove extreme values that are often attributable to things other than what you are trying to understand.

Now this works well for a dataset with 11 items, and also for a dataset of 15 items, 19 items, 21 items, and so on. Any number of elements that satisfies

$$4N + 3$$

will give a nice symmetrical division into quartiles, since we need four quartiles of N elements, plus the three separating elements. If the dataset contains a slightly different number of values we will end up breaking the symmetry slightly, but this is usually not significant as most data sets are relatively large.

Compute Quartiles Quiz

Here is our age distribution with the frequencies for each age:

Age	19	20	21	22	23	24	25
Frequency	2	1	1	3	2	1	1

What are the lower quartile, the median, and the upper quartile?

What is the “trimmed” mean after removing the outliers (i.e. the mean of the inter-quartile range)?

Compute Quartiles Quiz 2

Let's consider a second example:

33, -99, 17, 13, 1489

What is the mean of this group of numbers? What is the mean after outliers are removed?

Percentiles

Percentiles are similar to quartiles. The k-th percentile is the value that occurs k% of the way through the data. So, for the 10^{th} percentile, we would have:

For our original age data:

Name	Age
Joseph	22
Maria	21
Susan	24
Marc	20
Tom	211
Jack	23

removing the upper 20^{th}-percentile will remove Tom's age from our dataset.

Binomial Distribution

We have looked at ways to calculate the probabilities of events like coin flips, but what happens when we have a large number of events?

For example, how do we calculate the probability that a coin having P(H) = 0.8 will come up heads nine times out of a total of 12 flips?

Suppose we flip two coins. We know that there are two possible outcomes in which we will see the same number of heads and tails (i.e. exactly one head and one tail):

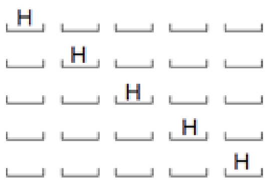
Heads	Heads
Heads	Tails
Tails	Heads
Tails	Tails

If we flip four coins, there are now $2^4 = 16$ possible outcomes, and we will see exactly the same number of heads and tails in six of these outcomes:

Heads	Heads	Heads	Heads
Heads	Heads	Heads	Tails
Heads	Heads	Tails	Heads
Heads	Heads	Tails	Tails
Heads	Tails	Heads	Heads
Heads	Tails	Heads	Tails
Heads	Tails	Tails	Heads
Heads	Tails	Tails	Tails
Tails	Heads	Heads	Heads
Tails	Heads	Heads	Tails
Tails	Heads	Tails	Heads
Tails	Heads	Tails	Tails
Tails	Tails	Heads	Heads
Tails	Tails	Heads	Tails
Tails	Tails	Tails	Heads
Tails	Tails	Tails	Tails

Clearly this can only work for an even number of coin flips. We cannot have the same number of heads and tails if we flip an odd number of coins.

If we were to flip five coins, there are just five possible outcomes where we would see exactly one head (or exactly one tail):



How many outcomes will give us exactly two heads?

Well, the first head could come up in any of the positions shown above, so there are five possibilities for the first head. This leaves four possible positions for the second head, giving us 4 x 5 = 20 outcomes. But this over-counts by exactly a factor of two. The reason is that we have counted the two cases below as different outcomes, depending on which coin came up heads first:

In reality of course, these should be counted as just one outcome, and the number of outcomes where we see exactly two heads is:

$$\frac{5 \times 4}{2} = 10$$

For three heads, we can work out the number of outcomes in two ways. Firstly, seeing three heads from five coin flips is exactly the same as seeing two tails, so we can use the same calculation we used to establish the number of outcomes having exactly two heads above.

The second way is to recognise that there are five places where the first head can appear, four for the second, and three for the third. But this is over-counting again. In each arrangement of three heads, there are three possible positions where we could have placed the first head, two for the second, and just one for the third. i.e. there are six arrangements, and our original calculation counted each of these as a different outcome, so the actual number of outcomes where we will see exactly three heads out of our five coin flips is:

$$\frac{5\times 4\times 3}{3\times 2\times 1}=10$$

So, following this logic, if we have ten coins, and we want to know the number of outcomes where we will see exactly four heads, we can just use:

$$\frac{10\times 9\times 8\times 7}{4\times 3\times 2\times 1}=\frac{5040}{24}=210$$

Combinatorics Quiz

We flip ten coins. How many possible outcomes have exactly five heads showing?

Formulae

There is a special notation that will help us express these examples in more general terms. This is called factorials. The factorial of any number, n is written as n!, and evaluates as:

$$n! \;=\; n\times(n-1)\times(n-2)\times\ldots\times 1$$

So, $10! \;=\; 10\times 9\times 8\times 7\times 6\times 5\times 4\times 3\times 2\times 1$

Now, we can actually write:

$$10\times 9\times 8\times 7\times 6=\frac{10!}{5!}$$

So, to calculate how many outcomes show k heads out of n coin flips, our formula becomes:

$$\frac{n!}{k! \cdot (n-k)!}$$

Arrangements Quiz

We flip 125 coins. How many outcomes end with exactly 3 heads showing?

Binomial

What happens when we introduce probabilities? Let's say that we have a fair coin (i.e. P(H)=0.5), and we flip the coin five times. What is the probability that we will get exactly one head?

Well, we know that there are exactly five ways that we can observe exactly one head in five flips:

$$\frac{5!}{1! \cdot (5-1)!} = \frac{5!}{1! \cdot 4!} = 5$$

We also know that there are $2^5 = 32$ possible outcomes (this is the size of the truth table).

So the probability of seeing exactly one head when we flip a fair coin five times is:

P(#HEADS = 1) = 5/32 = 0.15625

If we wanted to know the probability that we would see three heads out of five flips of a fair coin, we can calculate it using:

$$P(\#HEADS = 3) = \frac{\frac{5!}{3!(5-3)!}}{2^5} = \frac{10}{32} = 0.3125$$

What if the coin is loaded? Does that make a difference?

Let's say we have a loaded coin where P(H) = 08. We want to calculate the probability that we will see heads exactly once when we flip the coin three times.

We can answer this using a truth table:

Heads	Heads	Heads
Heads	Heads	Tails
Heads	Tails	Heads
Heads	Tails	Tails
Tails	Heads	Heads
Tails	Heads	Tails
Tails	Tails	Heads
Tails	Tails	Tails

Because the coin is loaded, not all outcomes in the truth table are equally likely. For the highlighted rows, where we have exactly one head, and two tails, the probability for that outcome will be:

$$P(H)\times(1-P(H))\times(1-P(H))=0.8\times0.2\times0.2=0.032$$

There are three outcomes, so the total probability of seeing exactly one head from three flips is:

$$3\times0.032=0.096$$

Binomial Quiz

Let's say that we flip our loaded coin with P(H) = 0.8 five times. What is the probability that we will see exactly four heads?

Binomial Quiz 2

Let's say that we flip our loaded coin with P(H) = 0.8 five times. What is the probability that we will see exactly three heads?

Binomial Quiz 3

Let's say that we flip our loaded coin with P(H) = 0.8 twelve times. What is the probability that we will see exactly nine heads?

Conclusion

So, for any coin having P(H) = p, that we flip n times, the probability of seeing k heads is given by:

$$\text{P(\#HEADS = k)} = \frac{n!}{(n-k)!\cdot k!}\cdot p^k \cdot (1-p)^{n-k}$$

This formula gives the probability of what is known as the binomial distribution. This is the cumulative outcome of many identical coin flips.

As you can see, we can take very large experiments, with very large numbers of coin flips, and compute the probability that heads will appear a specified number of times using the relatively simple formula above.

Central Limit Theorem Programming (Optional)

In this section, we will take some steps towards one of the deepest insights in all of statistics. It is called the Central Limit Theorem. The way that we will get to this insight is through a series of programming exercises.

Now, all of the programming on this course is optional, and it is fine to skip this section, but this is probably the most interesting way to understand the central limit theorem and statistics involving large numbers.

Programming Flips

Write a function, flip(N), that simulates flipping a coin 1000 times.

Having done this, compute the mean and the standard deviation of your results.

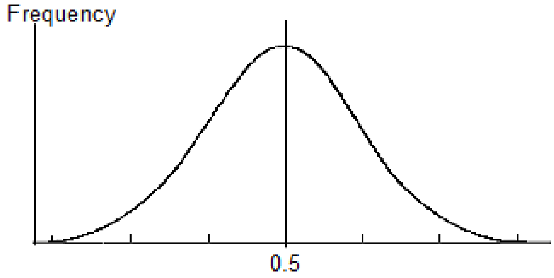
Hint:

The function random.random() gives a random output that sits between 0 and 1.

Sets of Flips

Write a function sample(N), that stores the means of N iterations of the function flip(N) in a list. Print the resulting list as a histogram using 30 bins.

The interesting thing here is that in this binomial distribution the frequency of outcomes seems to be centred around the expected outcome (in this case 0.5), and falls off according to this characteristic curve.



This curve is often known as a **bell curve**.

The significance of these bell curves, and their relationship to the central limit theorem, will be discussed in the next section.

The Normal Distribution

This section describes one of the most transformative things in modern statistics. We will start with the binomial distribution which we met earlier, and then move on to the central limit theorem. This effectively means that we are taking the number of coin flips to infinity. From that, we will arrive at the normal distribution, which is the basis for so much in statistics.

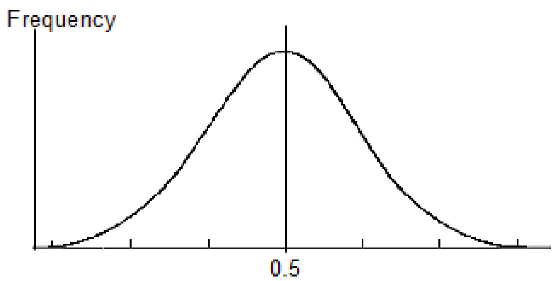
The reason why this is important is that experiments often generate thousands of data points (e.g. the thousands of patients in a drug trial), rather than the one or two coin flips we considered earlier. In order to analyse this amount of data, it is often more practical to start from the normal distribution as an approximation to the binomial distribution.

We start with our well-established formula for binomial distributions:

$$\frac{n!}{(n-k)! \cdot k!} \cdot p^k \cdot (1-p)^{n-k}$$

- n is the number of coin flips
- k is the number of times it comes up heads
- p is P(HEADS)

This expression is maximised when k = n/2. The other interesting thing is that the frequency of outcomes falls off as we move away from the maximum value according to this characteristic curve:



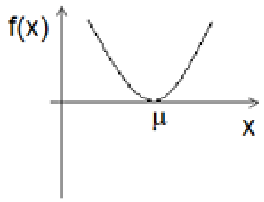
This is the bell curve we discussed earlier, and the question arises: can we identify a better formula to describe this curve? It turns out that there is a better formula, and that it can be applied to almost any distribution that is sampled many times.

We can define a normal distribution with a specific mean, μ, and variance, σ^2.

Now, for any outcome x, we can write the function:

$$f(x) = (x-\mu)^2$$

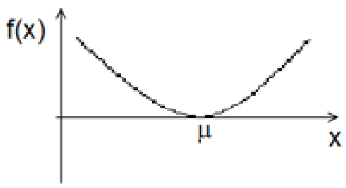
This function produces a characteristic quadratic curve with a minima where x = μ:



Now we divide our function f(x) by σ^2:

$$f(x) = \frac{(x-\mu)^2}{\sigma^2}$$

This scales down the output, f(x), by a factor of σ^2 and has the effect of 'widening' or flattening the quadratic curve:

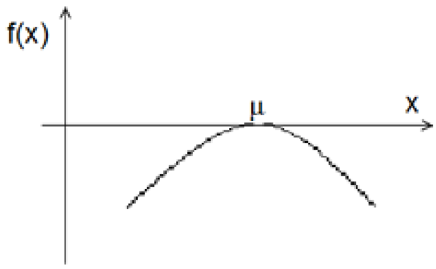


Note that large variances will result in very wide quadratic curves and small variances will result in relatively narrow or 'sharp' quadratic curves.

Next, we multiply our function by -½ to get:

$$f(x) = -\frac{1}{2} \cdot \frac{(x-\mu)^2}{\sigma^2}$$

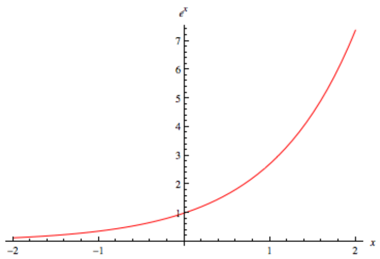
This flattens and inverts our quadratic curve so we get the following curve that has a maximum of zero when x = μ, but is otherwise strictly negative:



Finally, we make our function the exponent of e:

$$f(x) = e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$$

Now the exponential function $f(x) = e^x$ has a well known characteristic curve:



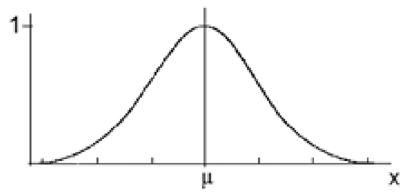
The exponential function is monotonic, that is, the larger the value of the exponent, the larger the value of the function.

The maximum value of our exponent is zero, and this occurs when x = μ. So our function f(x) is maximised when x = μ.

When x = μ, $(x-\mu) = 0$ and $f(x) = e^0 = 1$

The value of the function will be minimised when x = ± ∞. At these values of x we obtain $f(x) = e^{-\infty} = 0$

So we now have a function f(x) that has the value 1 when x = μ, and decays to 0 when x approaches ± ∞. This function also looks like the bell curve (although that is not entirely obvious!):

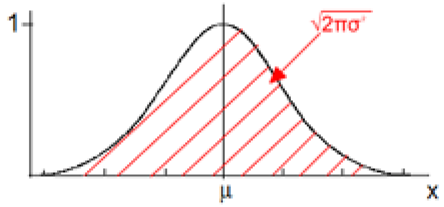


So this relatively simple formula describes the limit of computing the mean over any set of experiments, such as making infinitely many coin flips. No matter what kinds of experiment we do, when we drive n to very large numbers, we will obtain a bell curve like this.

The only flaw with this curve as it stands is that the area under the curve doesn't always add up to 1. It turns out that we do need the area under the curve to add up to 1 (just as we wanted the coin flip and its compliment to add up to one).

The area under the curve is given by:

$$\sqrt{2\pi \cdot \sigma^2}$$



We can normalise our function to ensure that the area under the curve always equals 1 by multiplying our function by the inverse of this to give the true normal distribution:

$$f(x) = \frac{1}{\sqrt{2\pi \cdot \sigma^2}} \cdot e^{-\frac{1}{2}\cdot\frac{(x-\mu)^2}{\sigma^2}}$$

This is the normal distribution for any value x, indexed by the parameters μ and σ^2. We can write this using mathematical notation as follows:

$$N(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi \cdot \sigma^2}} \cdot e^{-\frac{1}{2}\cdot\frac{(x-\mu)^2}{\sigma^2}}$$

Formula Summary

So, we have our normal distribution, which can also be written as:

$$N(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi \cdot \sigma^2}} \cdot \exp\left\{-\frac{1}{2}\cdot\frac{(x-\mu)^2}{\sigma^2}\right\}$$

Now, if you are new to this, then this expression will look really cryptic. In time, if you continue working with statistics, this formula will become second nature to you. For now, it is important that you just understand how the formula is constructed, with the quadratic penalty term for deviations from the expected mean of the expression.

We can extract values from the normal distribution just as we did with flipping coins before. The way to do this is to recognise that any value of x has a probability given by the equation above. Therefore, a value x, that has twice the value as some other value x' is twice as likely to be drawn.

P(x) = 2.P(x')

Now, the normal distribution has an entire continuous range of outcomes. Obviously, this renders each individual outcome to be probability 0, but in essence, we can think of the height of the curve at point x as being proportional to the probability of that value being drawn.

Central Limit Theorem

There are a number of types of experiment we can carry out. We might characterise them as:

1. single coin flip
2. many coin flips - mean
3. infinitely many coin flips - mean

These examples have parallels in other fields. A medical doctor with a single patient can treat them much like the single coin-flip. If they have 10 patients, then it is more like the binomial distribution in case 2. Alternatively, if they have many thousands of patients - as they do on many drug-trials, for example - then they will be treated as a normal distribution.

In each case, there is an appropriate formula to determine the probability distribution:

1. p
2. $\frac{n!}{(n-k)! \cdot k!} \cdot p^k \cdot (1-p)^{n-k}$
3. $\frac{1}{\sqrt{2\pi \cdot \sigma^2}} \cdot e^{-\frac{1}{2} \cdot \frac{(x-p)^2}{\sigma^2}}$ with $\sigma^2 = \frac{p(1-p)}{n}$

To clarify:

For 1 coin flip: This is the formula for figuring out the probability of the flip being heads?

For many coin flips: This is the formula for figuring out the probability of a given number of flips (k) being heads?

For infinitely many flips: What is the formula for figuring out the probability of a given proportion of flips being heads?

For formula 2:

$$\sigma^2 = \frac{p(1-p)}{n}$$

What is interesting is that it turns out that it is the central limit theorem that governs the transition from a single coin-flip to many coin-flips, and right through to infinitely many coin-flips.

The exponential function in the third expression captures the distribution of a possible mean if it transitions from a discrete space of many, but finite, outcomes to a space that is continuous and which has infinitely many outcomes.

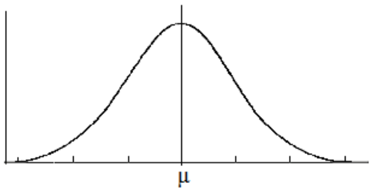
It turns out that the transition to a normal distribution works not just for coin-flips, but also for many other distributions that fall outside the scope of this class.

Manipulating Normals

Now that you understand something of the theory of normal distributions, we're going to return to something intuitive, that is how to manipulate normal distributions.

As we said in the last section, you can draw values from the normal distribution in just the same way as you would by flipping coins. So perhaps we should try to understand what happens when we manipulate normal distributions?

Suppose all the salaries at Udacity are normally distributed. That is to say, if we were to look at the salaries, there will be a well-defined mean and the salary distribution would approximate to the normal distribution curve, with fewer and fewer people having salaries as we get further and further from the mean:



Let's say that the mean is \$60,000 per year and the standard deviation is 10,000:

$\mu = \$60,000$
$\sigma = 10,000$

Now, we already know from previous units that if we give everybody a raise of a fixed amount, say \$10,000, the new mean and standard deviation will be:

$\mu' = \$70,000$
$\sigma' = 10,000$

In the context of the normal distribution, if we ignore the normalisation constant, we know that all the salaries are drawn from a distribution that looks like this:

$$\exp\left\{-\frac{1}{2}\cdot\frac{(x-60000)^2}{10000^2}\right\}$$

Now, the new salary, $x' = x + 10,000$

Or, alternatively, $x = x' - 10,000$

We can substitute this into our exponent to give:

$$\exp\left\{-\frac{1}{2}\cdot\frac{(x'-10000-60000)^2}{10000^2}\right\} = \exp\left\{-\frac{1}{2}\cdot\frac{(x'-70000)^2}{10000^2}\right\}$$

So the mean has increased to \$70,000, but the standard deviation remained unchanged.

OK, so let's say that Udacity is doing really well and they decide to double everybody's salary. Once again, we met this situation in earlier units, but let's now look at it in the context of the normal distribution.

Say the mean salary is now \$70,000, with a standard deviation of 10,000, so:

μ = \$70,000
σ = 10,000
x' = 2x

Substituting into the formula gives:

$$\exp\left\{-\frac{1}{2}\cdot\frac{\left(\frac{1}{2}x'-70000\right)^2}{10000^2}\right\} = \exp\left\{-\frac{1}{2}\cdot\frac{\left(\frac{1}{2}x'-\frac{1}{2}140000\right)^2}{10000^2}\right\}$$

$$= \exp\left\{-\frac{1}{2}\cdot\frac{1}{4}\frac{(x'-140000)^2}{10000^2}\right\} = \exp\left\{-\frac{1}{2}\cdot\frac{(x'-140000)^2}{(2\times 10000)^2}\right\}$$

$$= \exp\left\{-\frac{1}{2}\cdot\frac{(x'-140000)^2}{(20000)^2}\right\}$$

So we can see that:

μ = \$140,000
σ = 20,000

Throwing Quiz

Suppose that we are out on a playing field learning to throw a ball. On average we are able to throw it 30m, but because there is an element of randomness in the ball we have a standard deviation of 5m.

After training we have improved our performance by 10% and we learned to step forward more before we throw the ball, giving a further improvement of 2m on the distance we are able to throw the ball.

For a Gaussian outcome like this, how does this improvement change the mean, μ', and standard deviation σ'?

Golfer Quiz 1

Suppose that a golfer hits a golf ball down the fairway. The average distance is 100m with a variance of $30m^2$. The second shot down the fairway also has an average distance of 100m and a variance of $30m^2$.

For the combined strokes, what would you expect the mean distance and variance to be?

Golfer Quiz 2

Consider the same situation, but with the problem expressed in terms of standard deviation rather than variance:

$\mu = 100m$
$\sigma = 10m$

We know that the combined μ will be 200m, but what about the new standard deviation?

> So, when we add Gaussian variables, the means and variances add up, but the standard deviations **do not** add up.

Constants Quiz

Let's say that we draw the value A from a normal distribution:

$N(\mu, \sigma^2)$ μ is the mean and σ^2 is the variance.

Calculate the mean and variance of **aA + b** where a and b are constants.

Adding Normals Quiz

Let's say we have a normal distribution with μ and σ^2 from which we draw A, and we do the same with B from a distribution having the same μ and σ^2.

$A \sim N(\mu, \sigma^2)$ $B \sim N(\mu, \sigma^2)$

We add A + B. What are the new μ and σ^2?

Subtracting Normals Quiz

What happens to μ and σ^2 if we subtract A - B?

Summary

We have covered a lot in this section. We have seen that if X is normal, having the parameters μ and σ^2 then,

$$\mathbf{aX + b}$$

will always have the parameters:

$$\mathbf{a\mu + b} \quad \text{and} \quad \mathbf{a^2\sigma^2}$$

We also say that if we add X and Y (both normals), the result has a mean that is the sum of the means of X and Y, and has a variance that is the sum of the variances of X and Y.

You have now practiced some basic math on normals, and should be beginning to get a feel for how they change as we manipulate them.

Statistical Mythbusters

It is a known fact that most drivers believe that they drive better than the average driver. Is it possible that they could be right?

Most people also believe that they have a higher IQ than the average person. Could they also be right about this?

It is even said that most people believe that they can run faster than the average person. They can't possibly be right, can they?

In fact, people could also be correct in all of these cases. Let's consider a very simple example to explain why.

Here is a hand.

It has five fingers. No real surprise there.

In most cases people have five fingers on their right hand. If you were to survey a group of 20 people, and count the number of fingers on their hand, you might get a data set that looks something like this:

5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 4, 5, 5, 5, 5, 5, 5, 5

Now, you will find people who have fewer fingers, say 3 or 4 fingers, but that is quite rare (and even the occasional person with an *extra* finger, though that is even more rare!).

If we take the mean of our data sample in this case, we find that the average number of fingers for the group is 4.95.

So 95% of people in our sample group have more than the average number of fingers! 19 out of 20 hands have more than 4.95 fingers. In reality, with a larger sample size, the percentage would probably be very much larger.

So the three statements that we began with are all entirely possible from a statistical point of view, even though none of them may be based on any actual scientific evidence.

Answers

Compute Quartiles Quiz

Lower quartile = 20
Median = 22
Upper quartile = 23

Trimmed mean = 21.857

Compute Quartiles Quiz 2

Mean = 290.6
Trimmed mean = 21

Combinatorics Quiz

$$\frac{10\times 9\times 8\times 7\times 6}{5\times 4\times 3\times 2\times 1}=\frac{30240}{120}=252$$

Arrangements Quiz

$$\frac{125!}{3!\cdot(125-3)!}=\frac{125\times 124\times 123}{3\times 2}=317750$$

Binomial Quiz

The number of outcomes with exactly four heads is given by:

$$\frac{5!}{4!\cdot(5-4)!}=\frac{5!}{4!}=5$$

So, P(#HEADS = 4) = $5\times 0.8\times 0.8\times 0.8\times 0.2=0.4096$

Binomial Quiz 2

The number of outcomes with exactly four heads is given by:

$$\frac{5!}{3! \cdot (5-3)!} = \frac{5!}{3! \cdot 2!} = 10$$

So, P(#HEADS = 3) = $10 \times (0.8)^3 \times (0.2)^2$ = 0.2048

Binomial Quiz 3

The number of outcomes with exactly nine heads is given by:

$$\frac{12!}{9! \cdot (12-9)!} = \frac{12 \times 11 \times 10}{3!} = 220$$

So, P(#HEADS = 3) = $220 \times (0.8)^9 \times (0.2)^3$ = 0.236

Programming Flips

```
import random
from math import sqrt

def mean(data):
    return float(sum(data))/len(data)

def variance(data):
    mu=mean(data)
    return sum([(x-mu)**2 for x in data])/len(data)

def stddev(data):
    return sqrt(variance(data))


def flip(N):
    fred = []
    i = 0
    while i < N:
        if random.random() > 0.5:
            head = 1
        else:
            head = 0
        fred.append(head)
        i += 1
    return fred
```

Sets of Flips

```
import random
from math import sqrt
from plotting import *

def mean(data):
   return float(sum(data))/len(data)

def variance(data):
   mu=mean(data)
   return sum([(x-mu)**2 for x in data])/len(data)

def stddev(data):
   return sqrt(variance(data))


def flip(N):
   return [random.random()>0.5 for x in range(N)]

def sample(N):
   return [mean(flip(N)) for x in range(N)]


N=1000
outcomes=sample(N)
histplot(outcomes,nbins=30)

print mean(outcomes)
print stddev(outcomes)
```

Throwing Quiz

$\mu' = (1.1 \text{ x } \mu) + 2\text{m} = 35\text{m}$
$\sigma' = 1.1 \text{ x } \sigma = 5.5\text{m}$

Golfer Quiz 1

$\mu' = 200\text{m}$
$\sigma^{2\prime} = 60\text{m}^2$

Golfer Quiz 2

$\sigma^2 = 100\text{m}^2$
$2\sigma^2 = 200\text{m}^2$
so $\sigma' = 14.14\text{m}$

Constants Quiz

We have seen that multiplying values by a constant, a, results in a mean multiplied by that constant, $a\mu$. Adding a constant, b, to the values results in a mean $\mu+b$. So:

$$\mu = a\mu + b$$

We also saw that when we multiplied values by a constant, a, the standard deviation, σ, also increased to $a\sigma$. So:

$$\sigma^2 = a^2\sigma^2$$

Adding Normals Quiz

$$\mu = 2\,\mu$$
$$\sigma^2 = 2\,\sigma^2$$

Subtracting Normals Quiz

$$\mu = 0$$
$$\sigma^2 = 2\,\sigma^2$$