Instructions to Teachers: Big Groups Have Loud Voices

Hello and thank you for considering to use this video lesson in your class. This lesson is designed to show how mathematics - and specifically statistics - can be valuable to us in our daily lives. It is meant to be a fun exposure to statistics and probability.

Below are the activities we suggest for your students during the video breaks.

First Break

- First, count how many students are in the class. If the number is odd, add yourself to the group.

- Now divide the class into two groups of equal size. Have people note how many are in each group, and how many are in the class in total.

- Ask one person in each group how many are in her group, and then have the students double that number. Ask whether that answer is the total number of students.

- Point out that Arnie's rule is correct, and get the students to the general observation that, *when the two groups are of equal size, doubling the number in one of them gives the total for the two.*

- Now divide the class into two groups, with one of them roughly twice the size of the other. If there are 30 students, for example, create one group of size 20 and the other of size 10. If the total number of students is not a multiple of three, get close to a 2:1 ratio even if it can't be achieved perfectly. With 28 group members, the groups could be of sizes 19 and 9.

- Once again, ask someone in each group his group size, and then have the students double the two answers.

- Note that the doubling rule does not yield the total number of students. If there are 30 students, doubling the answer from a student from the group of size 20 gives us 40. Doubling the answer from a student in the smaller group of size 10 yields 20.

- Remind the students that the rule is *bound to fail* if the two groups are of unequal size. Assuming that the students have seen algebra, let x be the size of the first group and y the size of the second. Then the total number of people is $x + y$. If $x = y$, then $x + y = 2x = 2y$. But if x y, then $2x$ cannot equal $x + y$, and neither can $2y$. Except when $x = y$, the doubling rule that Arnie proposed is certain to be wrong!

- Have the students summarize the conclusion:

Only when two groups are of equal size does doubling the answer from a randomly-chosen person in one of the groups give the total number of people in both groups. That sometimes happens, but certainly not always.

Instructions to Teachers: Second Break

- Keep the classroom in the same two groups as before where one group is essentially twice the size of the other.

- Say that you will randomly choose six people from the class list and ask them how many people are in his or her group. In practice, *always choose four from the larger group and two from the smaller group*. Double the individual answers and then average the six numbers together. Write it down. Repeat this experiment a few times.

- Note that the six-number average is always larger than the number of people in the room.

- Lead a discussion about why this happens. Remind people that Arnie said that there would be three people chosen at random from the larger group and three from the smaller group. Yet the actual split was consistently four and two. Get people to recognize that, when there are twice as many people in the larger group as the smaller, that 2:1 ratio should prevail for people picked at random. (True: there could be some temporary variations around 2:1, but they should not persist in the long run.) Arnie was wrong to assume a three/three split: more than half the people will come from the bigger group.

- Get the students to recognize that, in the 100/200 example that Arnie and Anna discussed, Arnie's method should lead to an estimate of 333 people in total rather than 300.

Instructions for Teachers: Third Break

In preparation for the next module, teacher asks the class some questions:

- If twelve buses per hour reach a certain bus stop, what is the average time between buses? (60/12 = five minutes)
- If a person arrives at the bus stop at a random time, how long might you expect that person to wait for the next bus? (2.5 minutes, because the person arrives on average halfway between two buses)

- But suppose that the buses arrive as follows between 7 PM and 8 PM:

7:00	7:02
7:10	7:12
7:20	7:22
7:30	7:32
7:40	7:42
7:50	7:52
(8:00)	

Do we still think the average wait is 2.5 minutes?

7:00 7:02 7:10 7:12 7:20 7:22

(No exact answer; just some doubt that 2.5 is correct because very high chance of arriving in an eight-minute interval.)

Instructions for Teachers: Fourth Break

Ask the students to imagine that the buses arrive every five minutes, but are exactly five minutes apart

5 5 5 5

If people arrive at two per minute, how long will they wait on average for the next bus? (2.5 minutes)

But suppose instead the buses still arrive at four per twenty minutes, but the times between buses are irregular

1 9 1 9

Do you think that people will wait longer now? Most of the people arrive in the red intervals and they have long waits until the next bus. (Some discussion.)

Let's work out how long they wait on average. Over the 20 minutes we have four arrivals during the two black periods, and 36 during the two red periods. The lucky ones in the black periods wait on average ½ minute apiece, but the unlucky ones in the red period wait on average 9/2 = 4.5 minutes apiece. How much total waiting was there?

- Four people waited ½ minute apiece, which adds up to two minutes in total

- 36 people waited 4.5 minutes apiece, which is 36*4.5 = 162 minutes in total

The 40 people waited 164 minutes, so their average wait was 164/40 = 4.1 minutes.

Summarize: Four buses in 20 minutes in both cases. But when there are long intervals and short intervals, most people arrive in the long ones and they suffer because of it. Those who arrive in short intervals do well, but only a few people arrive in the short intervals. When all the intervals are equal, no one arrives in a long interval, so there is no extra suffering. Average wait is as low as it can be.

Wrap-Up Session with Class

At the very end of the video, we tell the students that sometimes it is better barely to miss the bus than to arrive at a random time. To illustrate that point, ask the students to imagine that the bus schedule looks like the following:

4: 00 PM	4:02	4:04	4:30
(4:30)	4:32	4:34	5:00

Suppose the passengers arrive at the bus stop at a constant rate of one per minute (which means that, somehow, they do not know the bus schedule). Then:

- For people who just miss the bus, how long on average do they wait?
  (We assume that those who arrive within one minute of a bus departure just missed it.)

- What is the average waiting time for all passengers?

Over the period 4-4:30, there are three people who just missed the bus: the arrival between 4: 00 and 4:01, the arrival between 4:02 and 4:03, and the arrival between 4:04 and 4:05. Assume that each of these people misses the bus by 30 seconds. Then the first person waits 1.5 minutes for the next bus, as does the second. The third is in much worse shape: he must wait 25.5 minutes. But the average wait for the three of them is (1.5+1.5+25.5)/3 = **9.5** minutes.

But what of all 30 people who arrive between 4:00 and 4:30? Fully 26 of the 30 arrive during the 26-minute interval between 4:04 and 4:30. They people wait on average half the long interval, which is 13 minutes. Thus, the 26 people wait a total of 26*13 = 338 minutes. Thus, the average wait for all 30 people must exceed 338/30, which is *more than ten minutes*. The people who just missed the bus—with average wait of 9.5—actually did better! (The same argument would hold true for those who arrive between 4:30 and 5 PM.)

If there is time, you might present a general formula. Suppose that there are two groups of people, and let f be the fraction in the first group. Suppose that X is the average "experience" in the first group, where experience might relate to waiting time for a bus, or how crowded a theater is. Let Y be the average experience in the second group. Then, if we pick a few people at random, the average experience A they will report follows:

$$A = f X + (1-f) Y$$

Perhaps you might do an example or two with (say) $f = 3/4$.

However, the simple average B of the experiences of the two groups follows:

$$B = (X + Y)/2$$

When $f = 1/2$ (meaning the two groups are of equal size), A and B are the same. Otherwise, they differ. When $f > 1/2$, the quantity A pays more attention to the first group than the second. That is what we mean when we say "big groups have loud voices."

Which quantity, A or B, better describes the overall experience? It varies from case to case. In terms of waiting for a bus, we should pay more attention to longer intervals because more passengers get caught in them (i.e. we should favor A). But if we want the total number of people who attended two shows, then finding B (and doubling it) is better than working with A. The key point is that, for good or ill, the bigger group will have the louder voice. If the students recognize that point, they will have will have mastered the main idea of the video.

Thank you so much for participating in this exercise. How can we improve it?