Answer to the question number : 1(A)

A sampling distribution is a probability distribution of a statistic obtained from a larger number of samples drawn from a specific population. The sampling distribution of a given population is the distribution of frequencies of a range of different outcomes that could possibly occur for a statistic of a population.1

In statistics, a population is the entire pool from which a statistical sample is drawn. A population may refer to an entire group of people, objects, events, hospital visits, or measurements. A population can thus be said to be an aggregate observation of subjects grouped together by a common feature.

A lot of data drawn and used by academicians, statisticians, researchers, marketers, analysts, etc. are actually samples, not populations. A sample is a subset of a population. For example, a medical researcher that wanted to compare the average weight of all babies born in North America from 1995 to 2005 to those born in South America within the same time period cannot draw the data for the entire population of over a million childbirths that occurred over the ten-year time frame within a reasonable amount of time. They will instead only use the weight of, say, 100 babies, in each continent to make a conclusion. The weight of 100 babies used is the sample and the average weight calculated is the sample mean.

Now suppose that instead of taking just one sample of 100 newborn weights from each continent, the medical researcher takes repeated random samples from the general population, and computes the sample mean for each sample group. So, for North America, they pull up data for 100 newborn weights recorded in the U.S., Canada, and Mexico as follows: four 100 samples from select hospitals in the U.S., five 70 samples from Canada, and three 150 records from Mexico, for a total of 1,200 weights of newborn babies grouped in 12 sets. They also collect a sample data of 100 birth weights from each of the 12 countries in South America.

Each sample has its own sample mean, and the distribution of the sample means is known as the sample distribution.

The average weight computed for each sample set is the sampling distribution of the mean. Not just the mean can be calculated from a sample. Other statistics, such as the standard deviation, variance, proportion, and range can be calculated from sample data. The standard deviation and variance measure the variability of the sampling distribution

The number of observations in a population, the number of observations in a sample, and the procedure used to draw the sample sets determine the variability of a sampling distribution. The standard deviation of a sampling distribution is called the standard error. While the mean of a sampling distribution is equal to the mean of the population, the standard error depends on the standard deviation of the population, the size of the population, and the size of the sample.4

Knowing how spread apart the mean of each of the sample sets are from each other and from the population mean will give an indication of how close the sample mean is to the population mean. The standard error of the sampling distribution decreases as the sample size increases.

Special Considerations

A population or one sample set of numbers will have a normal distribution. However, because a sampling distribution includes multiple sets of observations, it will not necessarily have a bell-curved shape.

Following our example, the population average weight of babies in North America and in South America has a normal distribution because some babies will be underweight (below the mean) or overweight (above the mean), with most babies falling in between (around the mean). If the average weight of newborns in North America is seven pounds, the sample mean weight in each of the 12 sets of sample observations recorded for North America will be close to seven pounds as well.

However, if you graph each of the averages calculated in each of the 1,200 sample groups, the resulting shape may result in a uniform distribution, but it is difficult to predict with certainty what the actual shape will turn out to be. The more samples the researcher uses from the population of over a million weight figures, the more the graph will start forming a normal distribution.

Answer to the question number : 1(B)

Population density is the average number of people living per square mile/km A high population density implies that the population is high relative to the size of the country. Countries, such as Belgium and the Netherlands have a high population density. Large countries, such as Australia and Canada have very low densities. Though this low density is skewed by the fact large areas of Australia and Canada are considered inhospitable places to live, due to desert / Arctic conditions.

There is little correlation between population density and economic development. Bangladesh and Japan have similar population densities, both Japan has much higher real GDP per capita.

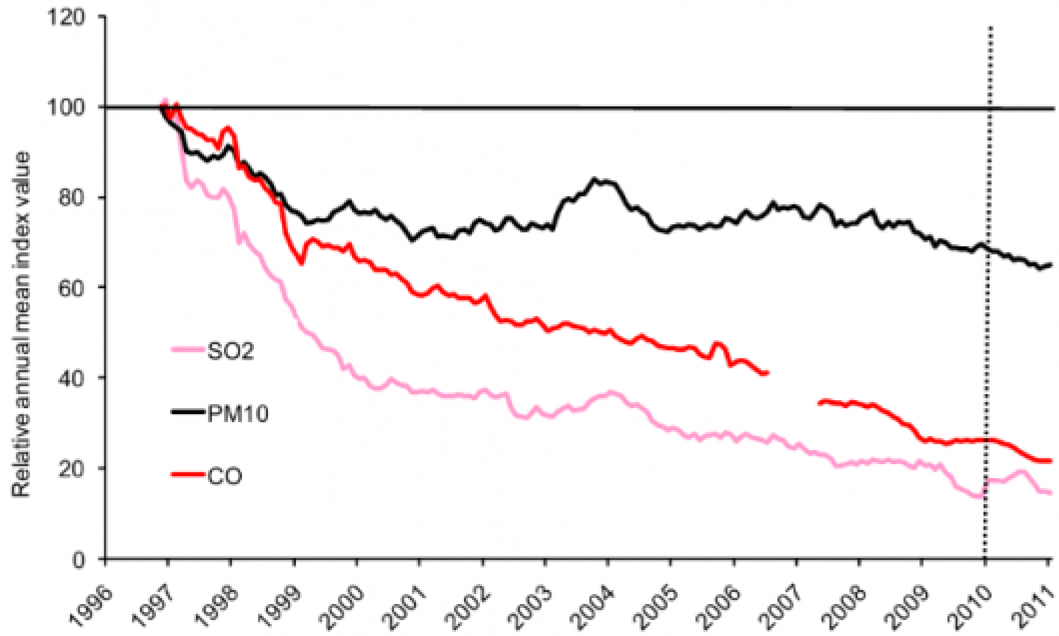
Most countries have seen a very significant rise in population densities over the past two hundred years. This rise in population density has been consistent with rising living standards and a better quality of life. However, others are concerned that a rising global population density could lead to a strain on resources, leading to food shortages, congestion and loss of the environment.

Benefits of a rising population density

- Economies of scale in national infrastructure. If you have a motorway network/train system connecting different parts of the country, a larger population density will help reduce the average costs of the transport network.
- Urban areas tend to be more energy efficient. Rural areas have a higher per-capita energy composition. For example, in remote areas, people will have to drive a long way to shops. In heavily populated urban areas, shops and facilities are likely to be within walking distance. Living in heavily built areas will benefit from your neighbour’s heat. However, if you look at suburbs, the situation is different with higher incomes leading to more energy consumption without the density of close housing in urban areas.

- Urban areas also make efficient city-wide public transport networks more efficient. As population density rises, the city is almost forced to switch from cars to more space efficient forms of transport, such as underground railway systems.
- Greater intellectual capital. Rising population leads to greater scope society will produce entrepreneurs and innovators, who come up with improved technology and business which helps to improve living standards.
- Technology has mitigated the negative forecasts of many who feared rising population. Malthus predicted in the Nineteenth Century, a rising population would lead to food shortages. The population has risen considerably, but areas of high population density have not seen the food shortages because of improved yields from agriculture and the ability to trade food.
- Technology enables higher living standards with lower levels of population. For example, in the 1950s, UK houses were often using coal fires to keep warm, leading to smogs and high levels of pollution in cities, but these smogs have been cleaned up. In theory, there is the scope for meeting energy needs from renewable energy, such as solar power.

Figure 3 – London Air Quality Network Index (Average) for PM_{10}, SO_2 and CO



Source: London clean air

- Despite rising population, London has seen some improvement in air quality standards in past decades. Though it should be noted O3 has been rising and air quality is still amongst the worst recorded in the UK.
- Rising population creates demand for labour and increased supply. Some fear, a rise in population (e.g. due to net migration) may lead to unemployment as migrants take jobs from native workers. However, a rising population both increases supply of labour and creates new demand from higher economic growth.
- Globalisation means countries are increasingly interdependent and don't need to be self-sufficient. A fear of rising population density is that it requires a country to import food. However, the UK has been a net importer of food for many years, and apart from the blockades of the World wars has not seen any food shortages. It doesn't make sense for any country to be self-sufficient in all goods and services. Greater specialisation enables increased living standards. Countries with higher population densities are likely to see increased specialisation in industries

suited to urban environments, such as the financial services of London.

Problems of a rising population density

- Pollution. Although we have cleaned up the visible smogs. Air quality is often still poor in heavily built-up areas. Rising population makes it harder to reduce pollution and improve air quality.
- Limit to improvement in agricultural productivity. Malthus massively underestimated potential for agricultural productivity to improve. However, increased use of fertilizers and pesticides has diminishing returns. Past improvements in agricultural yields is no guarantee, yields will keep rising. More turbulent weather, e.g. water shortages could place strain on global food production to meet the demands from rising global population. Countries with high population densities are more likely to have to import food.
- Global rise in population is so marked that it is having uncertain impacts of global environments. The rise in number of people in recent decades is unprecedented. Technology is unable to mitigate the impact on use of raw materials and pollution. For example, burning of fossil fuels is leading to higher levels of CO_2, which could have potential dire consequences for the world's environment and global weather.
- Congestion. With higher population densities, we can see road and transport congestion, unless there are satisfactory solutions in forms of pedestrian areas, good transport and new roads
- The biggest problem of higher population density is the potential loss of 'green-belt' land impacting on quality of life. Many people value green spaces as an important factor in the quality of life. If we lose all the countryside to roads and housing, then this reduces the quality of life.
- Limit to new roads. Rising population leads to more demand for transport, but with increased building, there is limited space for new road.

Should we be worried about a rising population density?

- Areas of high population density are often because they are seen as a desirable place to live. People move to live and work in London, despite high house prices, air pollution and congestion because there are many desirable amenities. People living in more rural areas may be likely to resist moves to increase population density because they are attracted to a quiet village life.
- Higher population density has definitely enabled economic and social development. But, at the same time the growth in the overall population of the planet is threatening to exacerbate many environmental and economic population, such as over-fishing, higher pollution, loss of habitat and stress on water.

Population density in Europe

Name	Population Density (/km2)	Area (km2)	Population
Malta	1,260	316	397,499
Netherlands	393	41,526	16,932,500
Belgium	337	30,510	11,007,020
United Kingdom	269	243,610	65,542,579
Germany	233	357,021	81,799,600
Liechtenstein	205	160	32,842
Italy	192	301,230	59,715,625

Population density in Europe			
Name	**Population Density (/km2)**	**Area (km2)**	**Population**
Switzerland	207	41,290	7,301,994
Luxembourg	173	2,586	512,000
Andorra	146	468	68,403
Moldova	131	33,843	4,434,547
Czech Republic	130	78,866	10,674,947
Denmark	125	43,094	5,368,854
Poland	124	312,685	38,625,478
Albania	123	28,748	3,544,841
Cyprus	117	9,248	803,147
Slovakia	111	48,845	5,422,366
France	111	547,030	63,601,002
Portugal	109	92,391	10,617,999
Armenia	108	29,743	3,262,200

Population density in Europe			
Name	**Population Density (/km2)**	**Area (km2)**	**Population**
Hungary	108	93,030	10,075,034
Serbia	97	88,361	7,498,001
Austria	97	83,858	8,169,929
Slovenia	95	20,273	2,048,847
Spain	92	505,782	46,777,373
Greece	81	131,940	11,606,813
Macedonia	81	25,713	2,054,800
Romania	80	238,391	19,043,767
Croatia	78	56,542	4,490,751
Bosnia and Herzegovina	78	51,129	3,964,388
Ukraine	76	603,700	45,396,470
Georgia	71	69,700	4,960,951
Bulgaria	69	110,910	7,621,337

Population density in Europe			
Name	**Population Density (/km2)**	**Area (km2)**	**Population**
Ireland	65	70,280	4,581,269
Belarus	50	207,600	10,335,382
Montenegro	48	13,812	626,000
Lithuania	44	65,200	2,881,020
Latvia	37	64,589	1,973,127
Estonia	28	45,226	1,294,236
Sweden	20	449,964	9,515,744
Finland	16	338,424	5,410,233
Norway	13	385,252	5,033,675
Russia	8	17,075,400	143,100,000
Iceland	3	103,001	320,060
Total	**31.56**	**26,680,676**	**842,033,572**

Answer to the question number : 1(C)

How do you make sure that the sample you use to gather your data is representative of the population you are researching? By taking the time to choose a sampling strategy. Choosing a sampling strategy is an essential step in the capture phase of the data journey and will ensure that your data is reliable and reflects the characteristics of your target group. In this blog, we'll take you step by step through the process by outlining the ways in which primary data is collected using an example in which a survey on characteristics (tax, education levels, etc) is collected on residents in five towns. The towns are of different sizes and have a total of 3,200 households. These 3,200 households make up the target population for your survey.

Step one: Define your sample and target population

At times, your survey may require you to cover the entire target population, as is the case in mapping or population studies. That's usually referred to as a census survey. However, target populations are generally large and expensive to survey. In our example, it may not be feasible to visit all 3,200 households of the five towns. Instead, you'd want to choose a smaller sample that would be representative of the population and reflect its characteristics.

A survey that is done on a smaller number of the target population is referred to as a sample survey. You can infer your findings for the entire population based on this representative sample. In the following sections, we'll describe the different terminologies that are associated with sample surveys, such as sample size and sampling technique. These concepts will enable you to determine the number of surveys needed to accurately reflect the true characteristics of a population and to choose the best method of selecting a sample from that population.

Step two: Define your sample size

The first step in your sampling exercise will be to decide on an appropriate sample size. There are no strict rules for selecting a sample size. You can make a decision based on the objectives of the project, time available, budget, and the necessary degree of precision.

In order to select the appropriate sample size, you will need to determine the degree of accuracy that you want to achieve. For this, you'll need to establish the confidence interval and confidence level of your sample.

The confidence interval, also called the margin of error, is a plus or minus figure. It is the range within which the likelihood of a response occurs. The most commonly used confidence interval is +/- 5. If you wish to increase the precision level of your data, you would further reduce the error margin or confidence interval to a +/- 2. For example, if your survey question is "does the household pay tax?" and 65% of your sampled households say "yes," then using a confidence interval of +/- 5, you can state with confidence that if you had asked the question to all 3,200 households, between 60% (i.e. 65-5) and 70% (i.e. 65+5) would have also responded "yes."

The confidence level tells you how sure you want to be and is expressed as a percentage. It represents how often the responses from your selected sample reflect the responses of the total population. Thus, a 95% confidence level means you can be 95% certain. The lower the confidence level, the less certain you will be.

Most surveys use the 95% confidence level and a +/- 5 confidence interval. When you put the confidence level and the confidence interval together, you can say that you are 95% sure that, if you had surveyed all (3,200) households, between 60% and 70% of the households of the target population would have answered "yes," to the question "does the household pay tax?".

The size of your sample may be determined using any standard sample size calculator such as Survey Monkey or Raosoft. Using a standard sample size calculator (as can be seen in table one below) for our example of 3,200 households in five towns, we can examine the difference in sample sizes based on different confidence levels and intervals.

Option A
If you decide on a 5% confidence interval and want to achieve a 95% confidence level, the sample size will be 345 households.

Option B
If you wish to have higher accuracy and increase the confidence level to 99%, the recommended sample size would be 551.

Option C
For an even higher accuracy you could choose a 2% confidence interval and 99% confidence level and arrive at a sample size of 1807.

If time and resources permit, you could opt for larger samples and choose option C, to survey 1807 households. However, the quality of your findings are likely to only be marginally better than with option A or B, as the rate of improvement of accuracy gradually diminishes with the increase in sample size. The size of your sample should therefore be decided by the objectives of the study and resources available.

Table 1: Calculate your sample size

Factors	Factors description	Option A	Option B	Option C
Population (no.)	The total population that your sample will present	3,200	3,200	3,200
Confidence level (%)	The probability that your sample accurately represents the characteristics of your population	95%	99%	99%
Confidence interval (%)	The range that your population's responses may deviate from your samples	5	5	2
	Sample size calculated	*345*	*551*	*1807*

Step three: Define your sampling technique

Once you've chosen the sample size for your survey, you'll need to define which sampling technique you'll use to select your sample from the target population. The sampling technique that's right for you depends on the nature and objectives of your project. Sampling techniques can be broadly divided into two types: random sampling and non-random sampling.

Random sampling

As the name suggests, random sampling literally means selection of the sample randomly from a population, without any specific conditions. This may be done by selecting the sample from a list, such as a directory, or physically at the location of the survey. If you want to ensure that a particular household does not get selected more than once, you can remove it from the list. This type of sampling is called simple random sampling without replacement. If you choose not to remove duplicate households from the list, you would do a simple random sampling with replacement.

Systematic sampling is the most commonly used method of random sampling, whereby you divide the total population by the sample size and arrive at a figure which becomes the sampling interval for selection. For example, if you need to choose 20 samples from a total population of 100, your sampling interval would be five. Systematic sampling works best when the population is homogeneous, i.e. most people share the same characteristics. In our example, the sampling interval would be nine (3200/345= 9 for a 95% confidence level and 5% confidence interval). You would thus select every ninth household in a town.

However, populations are generally mixed and heterogeneous. To ensure sufficient inclusion of all categories of the population, you will need to identify the different strata or characteristics and their actual representation (i.e. proportion) in the population. In such cases, you can use the stratified random sampling technique, whereby you first calculate the proportion of each strata within the population and then select the sample in the same proportion, randomly or systematically, from all the strata.

If we take our earlier example of five towns, to calculate a stratified random sample, you will need to calculate the proportion of each town within the sample size of 345 as shown in table two below. Column three gives the proportion of each town of the total population (3,200). In column four, the sample size (345) is proportionately divided across the five towns. For example, town three, which is 25% of the total population, will select 86 households with a sampling interval of nine (i.e. 800/86) in the same manner as was done for systematic sampling.

Table 2: Calculate stratified random sample

Location	**Population size**	**Proportion (%) of population**	**Stratified sample size**
Town 1	1200	38%	129
Town 2	900	28%	97
Town 3	800	25%	86
Town 4	180	6%	19
Town 5	120	4%	13
Total	*3200*		*345*

Non-random sampling

In non-random sampling, the sample selection follows a particular set of conditions and is generally used in studies where the sample needs to be collected based on a specific characteristic of the population. For example, you may need to select only households which own a car, or have children under six years of age. For this, you would consciously select only the 345 or 551 households that have those characteristics. Also termed purposive or subjective sampling, non-random sampling methods include convenience, judgment, quota and snowball sampling.

Step four: Minimise sampling error

It's normal to make mistakes during sample selection. Your efforts should therefore always be to reduce the sampling error and make your chosen sample as representative of the population as possible. The robustness of your sample depends on how you minimise the sampling error. The extent of errors during sampling vary according to the technique or method you choose for sample selection.

For samples selected randomly from a target population, the results are generally prefixed with the +/- sampling error, which is the degree to which the sample differs from the population. If your study requires you to know the extent of sampling error that is acceptable for the survey, you can select a random sampling technique. In random sampling, you will be able to regulate the survey design to arrive at an acceptable level of error. In a non-random sample selection, the sampling error remains unknown.

Thus, when your sample survey needs to infer the proportion of a certain characteristics of the target population, you can select a random sampling method. But if you want to know the perceptions of residents regarding taxation laws or the school curriculum, you would want to capture as many perceptions as possible, and therefore select a non-random method in situations where sampling errors or sampling for proportionality are not of concern. Non-random sampling techniques can be very useful in situations when you need to reach a targeted sample with specified characteristics very quickly.

If you don't have a sampling strategy in place, you may collect data which is biased or not representative, rendering your data invalid.

Answer to the question number : 1(D)

What Is a Non-Sampling Error?

A non-sampling error is a statistical term that refers to an error that results during data collection, causing the data to differ from the true values. A non-sampling error differs from a sampling error. A sampling error is limited to any differences between sample values and universe values that arise because the sample size was limited. (The entire universe cannot be sampled in a survey or a census.)

A sampling error can result even when no mistakes of any kind are made. The "errors" result from the mere fact that data in a sample is unlikely to perfectly match data in the universe from which the sample is taken. This "error" can be minimized by increasing the sample size.

Non-sampling errors cover all other discrepancies, including those that arise from a poor sampling technique.

How a Non-Sampling Error Works

Non-sampling errors may be present in both samples and censuses in which an entire population is surveyed. Non-sampling errors fall under two categories: random and systematic.

Random errors are believed to offset each other and therefore, most often, are of little concern. Systematic errors, on the other hand, affect the entire sample and therefore present a more significant issue. Random errors, generally, will not result in scrapping a sample or a census, whereas a systematic error will most likely render the data collected unusable.

Non-sampling errors are caused by external factors rather than an issue within a survey, study, or census.

There are many ways non-sampling errors can occur. For example, non-sampling errors can include but are not limited to, data entry errors, biased survey questions, biased processing/decision making, non-responses, inappropriate analysis conclusions, and false information provided by respondents.

Special Considerations

While increasing sample size can help minimize sampling errors, it will not have any effect on reducing non-sampling errors. This is because non-sampling errors are often difficult to detect, and it is virtually impossible to eliminate them.

Non-sampling errors include non-response errors, coverage errors, interview errors, and processing errors. A coverage error would occur, for example, if a person were counted twice in a survey, or their answers were duplicated on the survey. If an interviewer is biased in their sampling, the non-sampling error would be considered an interviewer error.

In addition, it is difficult to prove that respondents in a survey are providing false information—either by mistake or on purpose. Either way, misinformation provided by respondents count as non-sampling errors and they are described as response errors.

Technical errors exist in a different category. If there are any data-related entries—such as coding, collection, entry, or editing—they are considered processing errors.

Answer to the question number : 2(OR)

The binomial distribution is a probability distribution used in statistics that summarizes the likelihood that a value will take one of two independent values under a given set of parameters or assumptions.

The underlying assumptions of the binomial distribution are that there is only one outcome for each trial, that each trial has the same probability of success, and that each trial is mutually exclusive, or independent of one another.

Understanding Binomial Distribution

To start, the “binomial” in binomial distribution means two terms. We’re interested not just in the number of successes, nor just the number of attempts, but in both. Each is useless to us without the other.

The binomial distribution is a common discrete distribution used in statistics, as opposed to a continuous distribution, such as the normal distribution. This is because the binomial distribution only counts two states, typically represented as 1 (for a success) or 0 (for a failure) given a number of trials in the data. The binomial distribution thus represents the probability for x successes in n trials, given a success probability p for each trial.

Binomial distribution summarizes the number of trials, or observations when each trial has the same probability of attaining one particular value. The binomial distribution determines the probability of observing a specified number of successful outcomes in a specified number of trials.[1]

The binomial distribution is often used in social science statistics as a building block for models for dichotomous outcome variables, such as whether a Republican or Democrat will win an upcoming election or whether an individual will die within a specified period of time, etc. It also has applications in finance, banking, and insurance, among other industries.

Analyzing Binomial Distribution

The expected value, or mean, of a binomial distribution is calculated by multiplying the number of trials (n) by the probability of successes (p), or n x p.

For example, the expected value of the number of heads in 100 trials of heads or tales is 50, or (100 * 0.5). Another common example of the binomial distribution is by estimating the chances of success for a free-throw shooter in basketball where 1 = a basket is made and 0 = a miss.

The binomial distribution formula is calculated as:

$P_{(x:n,p)} = {}_{n}C_{x} \times p^{x}(1-p)^{n-x}$
where:

- n is the number of trials (occurrences)
- X is the number of successful trials
- p is probability of success in a single trial
- nCx is the combination of n and x. A combination is the number of ways to choose a sample of x elements from a set of n distinct objects where order does not matter and replacements are not allowed. Note that nCx=n!/(r!(n−r)!), where ! is factorial (so, 4! = 4 x 3 x 2 x 1)

The mean of the binomial distribution is np, and the variance of the binomial distribution is np (1 − p). When p = 0.5, the distribution is symmetric around the mean. When p > 0.5, the distribution is skewed to the left. When p < 0.5, the distribution is skewed to the right.

The binomial distribution is the sum of a series of multiple independent and identically distributed Bernoulli trials. In a Bernoulli trial, the experiment is said to be random and can only have two possible outcomes: success or failure.

For instance, flipping a coin is considered to be a Bernoulli trial; each trial can only take one of two values (heads or tails), each success has the same probability (the probability of flipping a head is 0.5), and the results of one trial do not influence the results of another.2 The Bernoulli distribution is a special case of the binomial distribution where the number of trials n = 1.

Example of Binomial Distribution

The binomial distribution is calculated by multiplying the probability of success raised to the power of the number of successes and the probability of failure raised to the power of the difference between the number of successes and the number of trials. Then, multiply the product by the combination between the number of trials and the number of successes.

For example, assume that a casino created a new game in which participants are able to place bets on the number of heads or tails in a specified number of coin flips. Assume a participant wants to place a $10 bet that there will be exactly six heads in 20 coin flips. The participant wants to calculate the probability of this occurring, and therefore, they use the calculation for the binomial distribution.

The probability was calculated as: (20! / (6! * (20 - 6)!)) * (0.50)^(6) * (1 - 0.50) ^ (20 - 6). Consequently, the probability of exactly six heads occurring in 20 coin flips is 0.037, or 3.7%. The expected value was 10 heads in this case, so the participant made a poor bet.

So how can this be used in finance? One example: Let’s say you’re a bank, a lender, that wants to know within three decimal places the likelihood of a particular borrower defaulting. What are the chances of so many borrowers defaulting that they’d render the bank insolvent? Once you use the binomial distribution function to calculate that number, you have a better idea of how to price insurance, and ultimately how much money to lend out and how much to keep in reserve.

What Is Binomial Distribution?

The binomial distribution is a probability distribution used in statistics that states the likelihood that a value will take one of two independent values under a given set of parameters or assumptions.

How Is Binomial Distribution Used?

This distribution pattern is used in statistics but has implications in finance and other fields as well. Banks may use it to estimate the likelihood of a particular borrower defaulting or how much money to lend and the amount to keep in reserve. It's also used in the insurance industry to determine policy pricing and to assess risk.

Why Is Binomial Distribution Important?

Binomial distribution is used to figure the likelihood of a pass or fail outcome in a survey or experiment replicated numerous times. There are only two potential outcomes for this type of distribution. More broadly, distribution is an important part of analyzing data sets to estimate all the potential outcomes of the data, and how frequently they occur. Forecasting and understanding the success or failure of outcomes is essential to business development.

Answer to the question number : 3(OR)

Definition:
Union of Events

The

union of events
AA and B,B, denoted A∪BA∪B, is the collection of all outcomes that are elements of one or the other of the sets AA and BB, or of both of them. It corresponds to combining descriptions of the two events using the word "or."

To say that the

event
A∪BA∪B occurred means that on a particular trial of the experiment either AA or BB occurred (or both did). A visual representation of the union of events
AA and BB in a
sample
space SS is given in Figure 3.2.23.2.2. The union corresponds to the shaded region.

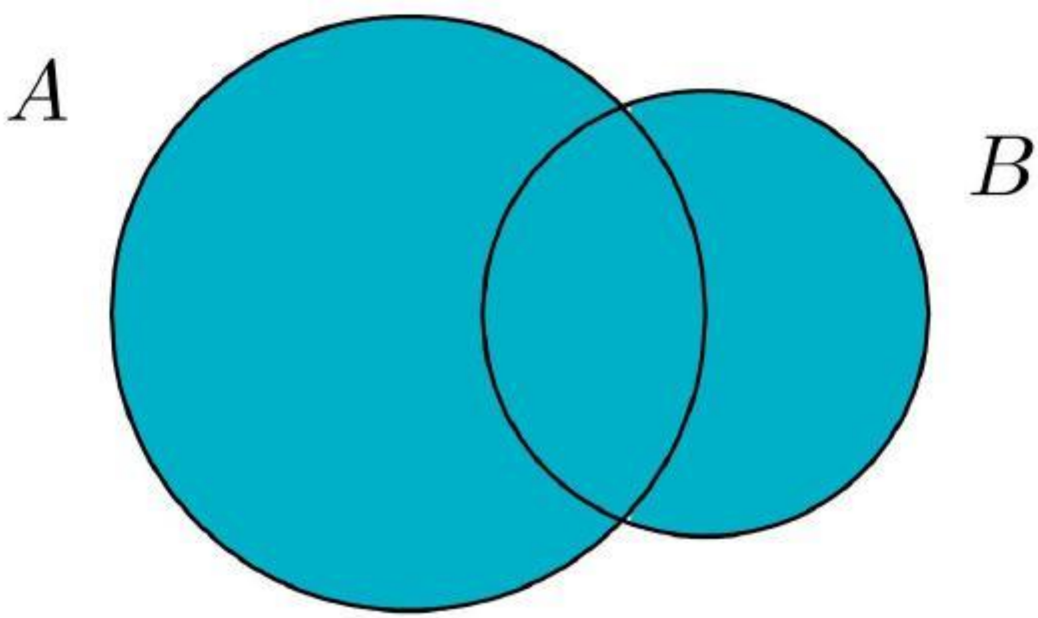


Figure 3.2.23.2.2: *The*

Union of Events

A and B

Example

In the experiment of rolling a single die, find the union of the events E**: "the number rolled is even"** and T**: "the number rolled is greater than two."**

Solution:

Since the outcomes that are in either $E=\{2,4,6\}$ or $T=\{3,4,5,6\}$ (or both) are $2,3,4,5$, and 6, that means $E\cup T=\{2,3,4,5,6\}$.

Note that an outcome such as 4 that is in both sets is still listed only once (although strictly speaking it is not incorrect to list it twice).

In words the union is described by "the number rolled is even or is greater than two." Every number between one and six except the number one is either even or is greater than two, corresponding to $E\cup T$ given above.

Example

A two-child family is selected at random. Let B denote the event that at least one child is a boy, let D denote the event that the genders of the two children differ, and let M denote the event that the genders of the two children match. Find $B\cup D$ and $B\cup M$.

Solution:

A sample space for this experiment is $S=\{bb,bg,gb,gg\}$, where the first letter denotes the gender of the firstborn child and the second letter denotes the gender of the second child. The events B,D, and M are $B=\{bb,bg,gb\}$, $D=\{bg,gb\}$, $M=\{bb,gg\}$.

Each outcome in DD is already in BB, so the outcomes that are in at least one or the other of the sets BB and DD is just the set BB itself: $B∪D=\{bb,bg,gb\}=BB∪D=\{bb,bg,gb\}=B$.

Every outcome in the whole

sample space SS is in at least one or the other of the sets BB and MM, so $B∪M=\{bb,bg,gb,gg\}=SB∪M=\{bb,bg,gb,gg\}=S$.

Answer to the question number : 4

Definition: intersections

The

intersection of events
AA and BB, denoted A∩BA∩B, is the collection of all outcomes that are elements of both of the sets AA and BB. It corresponds to combining descriptions of the two events using the word "and."

To say that the

event
A∩BA∩B occurred means that on a particular trial of the experiment both AA and BB occurred. A visual representation of the intersection of events
AA and BB in a
sample
space SS is given in Figure 3.2.13.2.1. The intersection corresponds to the shaded lens-shaped region that lies within both ovals.

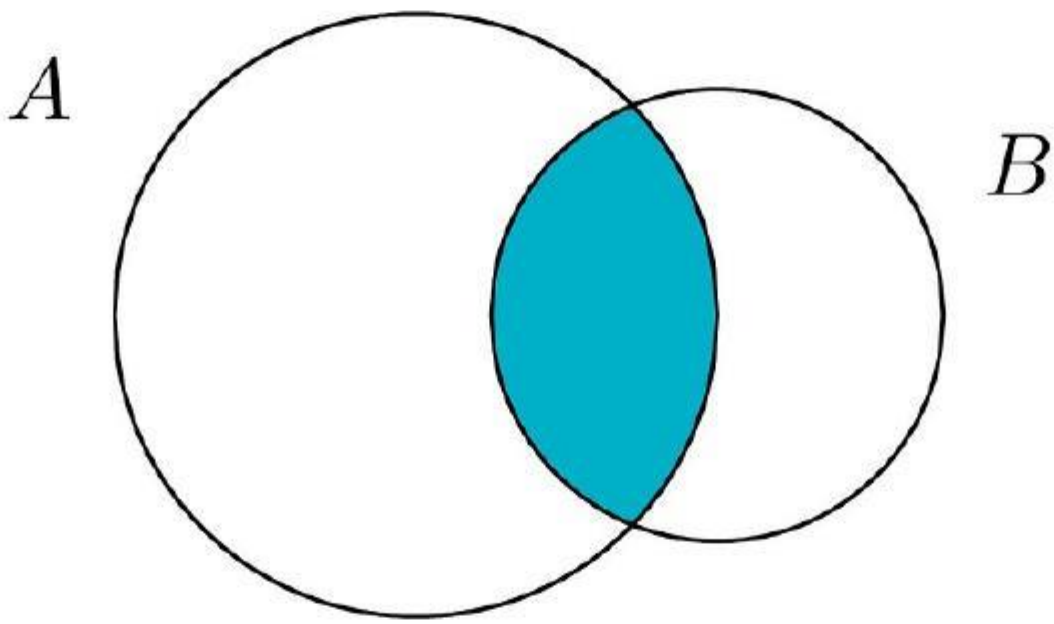


Figure 3.2.13.2.1*: The Intersection of Events*
A and B

Example

In the experiment of rolling a single die, find the intersection $E \cap TE \cap T$ of the events EE: *"the number rolled is even"* and TT: *"the number rolled is greater than two."*

Solution:

The

sample

space is $S=\{1,2,3,4,5,6\}S=\{1,2,3,4,5,6\}$. Since the outcomes that are common to $E=\{2,4,6\}E=\{2,4,6\}$ and $T=\{3,4,5,6\}T=\{3,4,5,6\}$ are 44 and 66, $E \cap T=\{4,6\}E \cap T=\{4,6\}$.

In words the intersection is described by "the number rolled is even and is greater than two." The only numbers between one and six that are both even and greater than two are four and six, corresponding to $E \cap TE \cap T$ given above.