**Name: Nur Ahammad**

**Subject:** STA-220

**Batch No: BBA 52**

**ID NO: 1121520011**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Ans To The Question No. 1.**

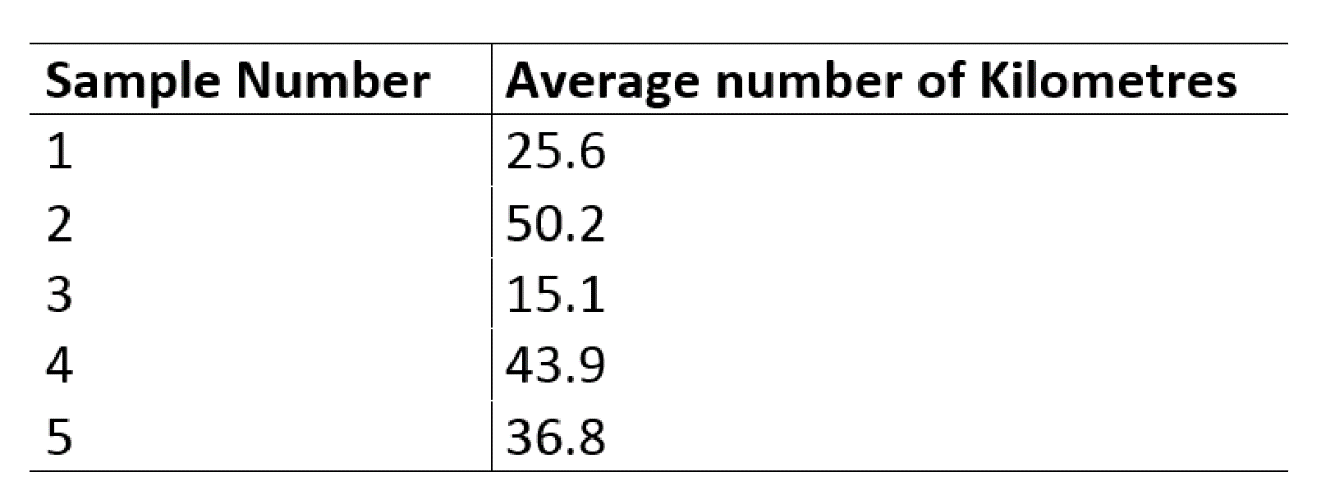
**a.)Sampling Distribution**

The sampling distribution of the mean refers to the pattern of sample means that will occur as samples are drawn from the population at large

**Example**

I want to perform a study to determine the number of kilometers the average person in Australia drives a car in one day. It is not possible to measure the number of kilometresdriven by every person in the population, so I randomly choose a sample of 10 people and record how far they have driven.

I then randomly sample another 10 drivers in Australia and record the same information. I do this a total of 5 times. The results are displayed in the table below.



**b. Population Distribution**: The population is the whole set of values, or individuals, you are interested in. For example, if you want to know the average height of the residents of India, that is your population, ie, the population of India

Population characteristic are mean (μ), Standard deviation (σ) , proportion (P) , median, percentiles etc. The value of a population characteristic is fixed. This characteristics are called population distribution. They are symbolized by Greek characters as they are population parameters.

A population is of course the entire group of individuals that you are interested in studying. This could be anything from all humans to a specific type of cell. The population distribution however, is a bit more narrow with its definition because it is specific to the measure you are interested in. So, if we are studying the heights of adult humans your population would be all adults but your population distribution would be all of the heights of every human in centimeters. Many resources don’t make this hard distinction, but if we think of population distributions in this way it will help we conceptualize exactly what the population parameters are.

**c.ans: Sampling errors**

Sampling errors are statistical errors that arise when a sample does not represent the whole population. They are the difference between the real values of the population and the values derived by using samples from the population. Sampling errors occur when numerical parameters of an entire population are derived from a sample of the entire population. Since the whole population is not included in the sample, the parameters derived from the sample differ from those of the actual population.

They may create distortions in the results, leading users to draw incorrect conclusions. When analysts do not select samples that represent the entire population, the sampling errors are significant.

**d.ans: Non sampling errors:**

Non-sampling error refers to all sources of error that are unrelated to sampling. Non-sampling errors are present in all types of survey, including censuses and administrative data. They arise for a number of reasons: the frame may be incomplete, some respondents may not accurately report data, data may be missing for some respondents, etc.

Non-sampling errors can be classified into two groups: random errors and systematic errors.

Random errors are errors whose effects approximately cancel out if a large enough sample is used, leading to increased variability.

Systematic errors are errors that tend to go in the same direction, and thus accumulate over the entire sample leading to a bias in the final results. Unlike random errors, this bias is not reduced by increasing the sample size. Systematic errors are the principal cause of concern in terms of a survey’s data quality. Unfortunately, non-sampling errors are often extremely difficult, if not impossible, to measure.

**Ans To The Question No. 2.**

Binomial distributions : Many instances of binomial distributions can be found in real life. For example, if a new drug is introduced to cure a disease, it either cures the disease (it’s successful) or it doesn’t cure the disease (it’s a failure). If you purchase a lottery ticket, you’re either going to win money, or you aren’t. Basically, anything you can think of that can only be a success or a failure can be represented by a binomial distribution.

In binomial probability distribution, the number of ‘Success’ in a sequence of n experiments, where each time a question is asked for yes-no, then the boolean-valued outcome is represented either with success/yes/true/one (probability p) or failure/no/false/zero (probability q = 1 − p). A single success/failure test is also called a Bernoulli trial or Bernoulli experiment, and a series of outcomes is called a Bernoulli process. For n = 1, i.e. a single experiment, the binomial distribution is a Bernoulli distribution. The binomial distribution is the base for the famous binomial test of statistical importance.

Four conditions are necessary for a binomial distribution.

1. A fixed number of trials; e.g. roll a die three times or flip a coin five times.
2. each trial has two possible outcomes: success or failure. Note the die example seems like it should have six outcomes. We need to consider something like “get a 6” (success) or “not get a 6” (failure).
3. The probability of success must be constant (the same for each trial. In the die example, P(6) =1/6 for each roll.
4. The trials are independent. The outcome of one does not affect the probability of success on succeeding trials.

(The fourth one has always seemed a little redundant to me. If the trials were not independent I’d think you couldn’t have a constant probability of success. Maybe someone more knowledgeable than I can elaborate on that.)

**Ans To The Question No. 3.**

**Definition**

The **union of events** A

and B, denoted A∪B, is the collection of all outcomes that are elements of one or the other of the sets A and B

, or of both of them. It corresponds to combining descriptions of the two events using the word "or".

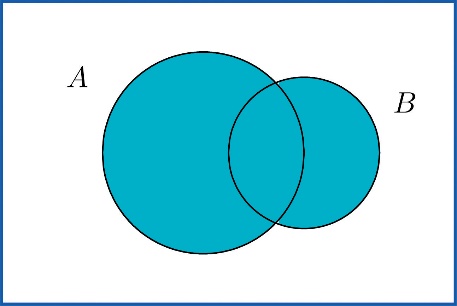
To say that the event *A*∪*B*

occurred means that on a particular trial of the experiment either *A* or *B* occurred (or both did). A visual representation of the union of events *A* and *B* in a sample space *S* is given in Figure 3.5

"The Union of Events". The union corresponds to the shaded region.

Figure 3.5 The Union of Events *A*

*and B*



**EXAMPLE 1**

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 , *E*∪*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*∪*T*

given above.

**EXAMPLE 2**

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*∪*D* and *B*∪*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 *D*

is already in *B*, so the outcomes that are in at least one or the other of the sets *B* and *D* is just the set *B* itself: *B*∪*D*={*bb*,*bg*,*gb*}=*B*

.

Every outcome in the whole sample space *S*

is in at least one or the other of the sets *B* and *M*, so *B*∪*M*={*bb*,*bg*,*gb*,*gg*}=*S*

The following **Additive Rule of Probability** is a useful formula for calculating the probability of *A*∪*B*

**Additive Rule of Probability**

*P*(*A*∪*B*)=*P*(*A*)+*P*(*B*)−*P*(*A*∩*B*)

The next example, in which we compute the probability of a union both by counting and by using the formula, shows why the last term in the formula is needed.

**EXAMPLE 3**

Two fair dice are thrown. Find the probabilities of the following events:

a. both dice show a four

b. at least one die shows a four

Solution:

As was the case with tossing two identical coins, actual experience dictates that for the sample space to have equally likely outcomes we should list outcomes as if we could distinguish the two dice. We could imagine that one of them is red and the other is green. Then any outcome can be labeled as a pair of numbers as in the following display, where the first number in the pair is the number of dots on the top face of the green die and the second number in the pair is the number of dots on the top face of the red die.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 11 | 12 | 13 | 14 | 15 | 16 |
| 21 | 22 | 23 | 24 | 25 | 26 |
| 31 | 32 | 33 | 34 | 35 | 36 |
| 41 | 42 | 43 | 44 | 45 | 46 |
| 51 | 52 | 53 | 54 | 55 | 56 |
| 61 | 62 | 63 | 64 | 65 | 66 |

a. There are 36 equally likely outcomes, of which exactly one corresponds to two fours, so the probability of a pair of fours is 1/36.

b. From the table we can see that there are 11 pairs that correspond to the event in question: the six pairs in the fourth row (the green die shows a four) plus the additional five pairs other than the pair 44, already counted, in the fourth column (the red die is four), so the answer is 11/36

. To see how the formula gives the same number, let *AG* denote the event that the green die is a four and let *AR* denote the event that the red die is a four. Then clearly by counting we get *P*(*AG*)=6/36 and *P*(*AR*)=6/36. Since *AG*∩*AR*={44},*P*(*AG*∩*AR*)=1/36;

this is the computation in part (a), of course. Thus by the Additive Rule of Probability,

*P*(*AG*∪*AR*)=*P*(*AG*)+*P*(*AR*)−*P*(*AG*−*AR*)=636+636−136=1136

**EXAMPLE 4**

A tutoring service specializes in preparing adults for high school equivalence tests. Among all the students seeking help from the service, 63% need help in mathematics, 34% need help in English, and 27% need help in both mathematics and English. What is the percentage of students who need help in either mathematics or English?

Solution:

Imagine selecting a student at random, that is, in such a way that every student has the same chance of being selected. Let *M*

denote the event "the student needs help in mathematics" and let *E* denote the event "the student needs help in English". The information given is that *P*(*M*)=0.63, *P*(*E*)=0.34, and *P*(*M*∩*E*)=0.27

. The Additive Rule of Probability gives

*P*(*M*∪*E*)=*P*(*M*)+*P*(*E*)−*P*(*M*∩*E*)=0.63+0.34−0.27=0.70

Note how the naive reasoning that if 63%

need help in mathematics and 34% need help in English then 63 plus 34 or 97% need help in one or the other gives a number that is too large. The percentage that need help in both subjects must be subtracted off, else the people needing help in both are counted twice, once for needing help in mathematics and once again for needing help in English. The simple sum of the probabilities would work if the events in question were mutually exclusive, for then *P*(*A*∩*B*)

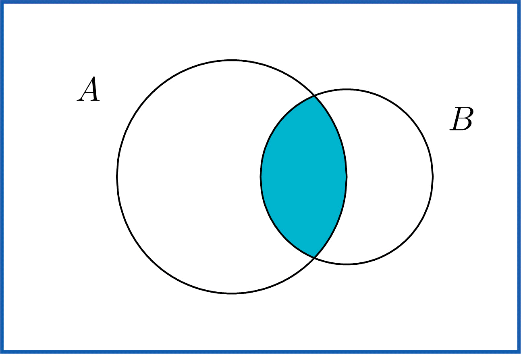
is zero, and makes no difference.

**Ans To The Question No. 4.**

**The intersection of events:**

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

To say that the event A ∩ B occurred means that on a particular trial of the experiment both A and B occurred. A visual representation of the intersection of events A and B in a sample space S is given in Figure 3.4 "The Intersection of Events ". The intersection corresponds to the shaded lens-shaped region that lies within both ovals.



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

**Example 1:**

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

**Solution:**

**The sample space is S={1,2,3,4,5,6}.**

**Since the outcomes that are common to E={2,4,6} and T={3,4,5,6} are 4 and 6, E∩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 ∩ T given above.

**Example 2:**

**A single die is rolled.**

Suppose the die is fair. Find the probability that the number rolled is both even and greater than two.

Suppose the die has been “loaded” so that P(1)=1∕12

, P(6)=3∕12

, and the remaining four outcomes are equally likely with one another. Now find the probability that the number rolled is both even and greater than two.

Solution:

In both cases the sample space is S={1,2,3,4,5,6}

and the event in question is the intersection E∩T={4,6}

of the previous example.

Since the die is fair, all outcomes are equally likely, so by counting we have P(E∩T)=2∕6.

The information on the probabilities of the six outcomes that we have so far is

OutcomeProbablity11122p3p4p5p6312

Since P(1)+P(6)=4∕12=1∕3

and the probabilities of all six outcomes add up to 1,

P(2)+P(3)+P(4)+P(5)=1−13=23

Thus 4p=2∕3

, so p=1∕6. In particular P(4)=1∕6.

Therefore

P(E∩T)=P(4)+P(6)=16+312=512