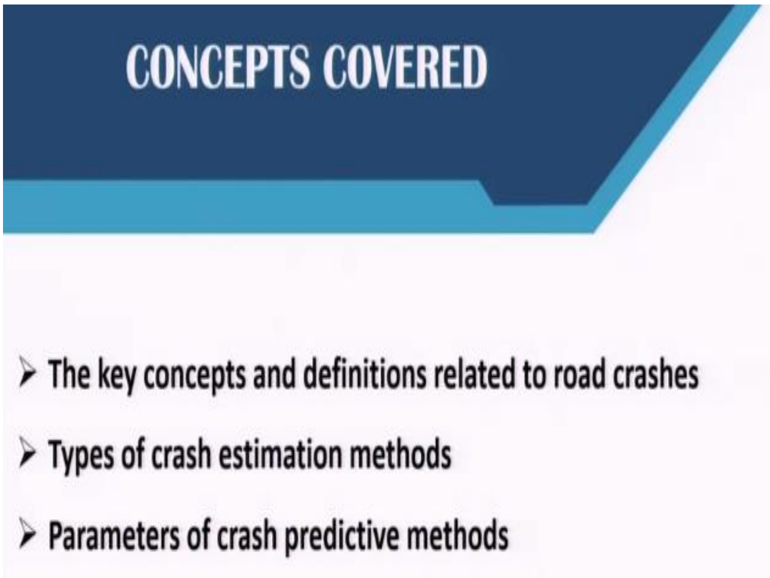
Introduction to Multimodal Urban Transportation System
Prof. Arkopal Kishore Goswami, Ph.D.
Ranbir and Chitra Gupta School of Infrastructure Design and Management
Indian Institute of Technology - Kharagpur

Lecture – 57
Urban Transport and Sustainability:
Road Crash Estimation and Elements of Predictive Methods

(Refer Slide Time: 00:39)



Welcome back friends. Now that we have introduced you to safety and pedestrian safety, in this lecture what we are going to expose you to are the different estimation methods of road crashes in general. Then we will subsequently look at specifically pedestrians and bicycle crashes, but you must first understand the concepts and definitions related to road crashes, how they are estimated and what are the parameters involved in these estimation methods.

(Refer Slide Time: 00:55)

So, you would now notice that in most of the transportation safety literature, we are using the word crash rather than accident. Because more and more research is showing that all of these events may not be an accident. For example, the meaning of the word accident in English means that it happens suddenly or without anybody's intervention or anything's intervention, but what we are noticing is that many of these accidents can be predicted. And if they can be predicted that means they can be avoided or stopped as well. So, hence this word crash seems to be a better fit. So, when we say crash what we mean is that a set of events that result in injury or property damage due to collision of at least one motorized vehicle and may involve collision with another motorized vehicle, a bicyclist, a pedestrian, or an object okay. So, a crash does not always necessarily lead to a fatality or a death.

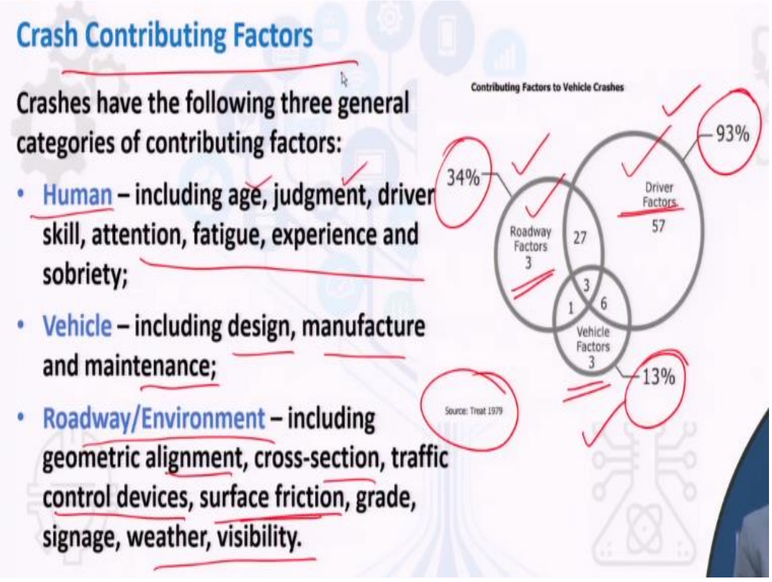
It may be just a property damage kind of a crash. It may be involving multiple vehicles or just one vehicle crashing with an object on the road or it may be a vehicle crashing with either a pedestrian or bicyclist. So, everything inclusive of all these terminologies is what is a crash.

(Refer Slide Time: 02:39)

So, when we want to estimate a crash, one way to know about a crash or an accident is after the fact that the accident happens. Now, that is a very reactive way of understanding or knowing where the crash has happened, but if you had to predict a crash or if you had to estimate that this roadway is likely to have n number of crashes in the next month or in the next year, then it is a proactive measure, and you make some interventions so that these crashes can be prevented. So, that is why it is especially important nowadays in today's transportation arena to estimate the crashes. So, when we say estimate of crashes what we are usually trying to do is we are trying to develop different methodologies to predict crash frequency, how frequently a crash can occur. So, it can be on an existing roadway for existing conditions during the past or in the future. So, if you already have data about the number of accidents or crashes in the past, then for that a roadway you can say for example on state highways in such and such state and in this district, there were n number of crashes per year. So, that is kind of a crash frequency that is quite easy to understand simple terminology of understanding what crash frequencies in the past are.

But can you use the same trend and predict that the crash frequency will be the same in the future or not that is what something we will tell you how to do that okay. So, it can be for an existing roadway, it can be for existing roadway for alternative conditions, maybe different conditions cause different crash frequencies, or it could be for a totally new roadway as well, that a new roadway is being built between two points. So, what is the probability or what is the prediction that how many crashes will occur there or how frequently they will occur. So that is also something that must be considered. So that this is the best time to consider the crash frequencies for a new roadway because vehicles have not yet started playing on that roadway so that you can incorporate safety measures proactively to reduce the probability of those crashes happening.

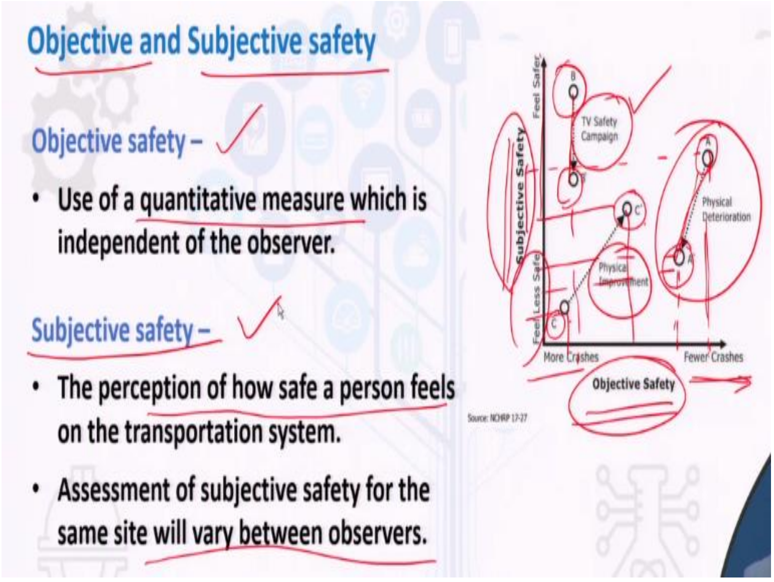
(Refer Slide Time: 05:19)



Now, what are the different contributing factors when it comes to crashes? So, what you will see in this picture on the right-hand side is overall what are the 3 different types of factors that we group them under. One is the roadway factors; the other are the vehicle factors and then there are driver factors. So, you will see this is a little bit of an old data set, but these percentages may differ as we are improving our research day by day. But overall, what we have seen is the crashes may occur due to different driver related factors, different roadway related factors and different vehicle related factors. So, driver related, or human factors include age, judgment, driver skill, attention, fatigue whereas vehicle related factors include the design of the vehicle, the maintenance of the vehicle, the manufacturer, what type of vehicle is it.

And whereas the roadway or the environment include the geometric alignment, cross section, traffic control devices, meaning signals or unsignalized intersection, the surface friction, grade, weather, visibility and all. So, these are broadly all the different types of factors that contribute to crashes on a facility on a highway or urban road.

(Refer Slide Time: 06:50)

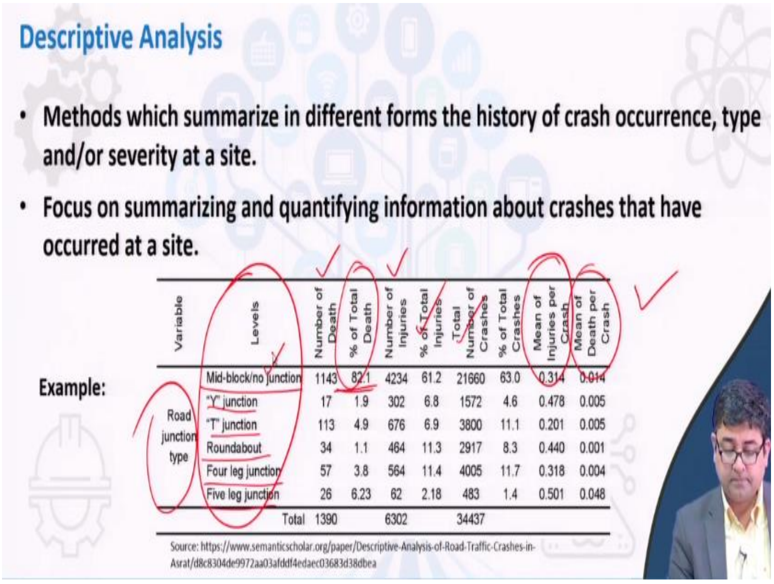


The other important thing to understand is this whole idea of objective and subjective safety. Objectively when we look at it something that can be quantified. So, when we look at objective safety, essentially what we are trying to do is trying to find out what are the number of crashes or what is the crash frequency and how can we reduce that, whereas there is a totally different arena of what is called a subjective frequency. That is the perception of how a person feels safe on the transportation system. This may vary based on the different types of users, vary from location to location, but nonetheless is becoming an especially important aspect when we talk about transportation safety. For example, see if you do a physical improvement on a roadway, what do you expect? You expect that because of that improvement, your crashes, which may be here, and it has led to now fewer number of crashes, i.e. it has gone from here to here and at the same time the subjective safety, so, now people feel that here they were feeling less safe, now they are feeling that they are safer. So, a physical improvement when you do something, take a measure to improve safety a physical measure that it should affect both the objective safety as well as the subjective safety. However, when you say for example do a TV safety campaign, it has been noticed that if the situation was such that you were feeling unsafe or you are feeling safer, the perception of subjective safety may change, but your objective safety is not changing, your objective safety is not moving. This line is parallel to your y axis. So, your objective safety is not changing because of a TV safety campaign.

So, these are different types of means of understanding how you can improve the safety of a facility. Similarly, when there is physical deterioration, so when a road or a bridge or anything is deteriorating in age, so what you start to suddenly feel is, you were feeling safer here, but you suddenly now start to feel less safe because that bridge is getting less and less safe. At the same time, maybe objectively as well there are increasing number of crashes.

Here there are fewer crashes, now there are more crashes, but the slope of the line says that your feeling of safety is reducing by a larger amount than the actual objective safety or the actual number of crashes that are taking place. So different measures give result to different types of objective and subjective safety implications. So, you are not only looking at the objective safety, but at the same time you must look at subjective safety of your highway or the facility in your area.

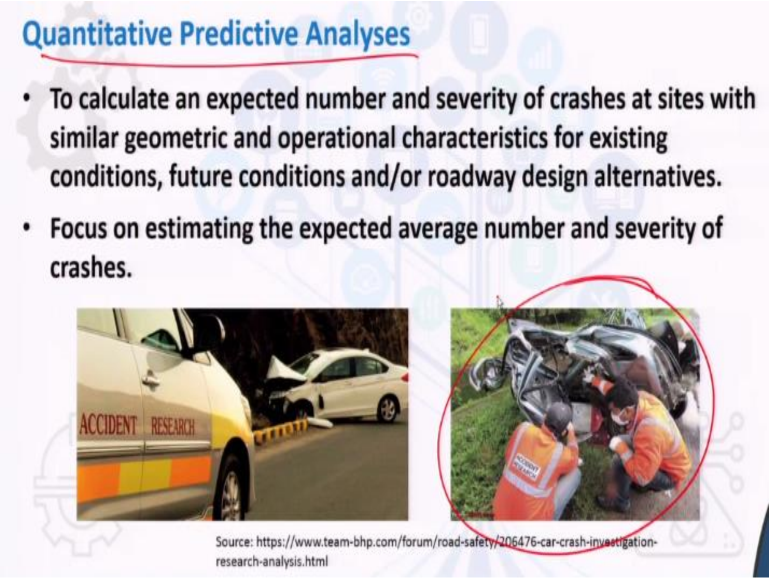
(Refer Slide Time: 10:20)



Then different descriptive analysis or descriptive measures have been developed over the years, which gives you an idea about the number of crashes or the frequency of crashes that are happening. For example, you may want to say you have decided that you want to know the number of accidents because of the different types of intersections that are there in your area. So, it may be at a midblock intersection, it may be at a Y intersection, it may be at a T intersection, it may be a roundabout, four legged or even five-legged junction. So, you can divide it up in that way and understand what the safety implications at different types of intersections in your city are. So, to know different parameters that are involved in understanding the road safety you may include just purely the number of people that have died at each of these or the percentage of death, number of injuries. Remember all crashes do not lead to death, there may be a lot of crashes that lead to minor injuries, major injuries and then maybe they lead to death. So, you may want to know either the total number or the percentage, just the total number of crashes or you want to divide it up into different types of injuries, minor injury or mean of the injuries or mean of the deaths.

So, you can develop different types of descriptive statistics, which will give you a clearer picture based on where you want to know about the safety of your facility in the city. So, similarly, you can do it for different types of roads, you can do it for daytime, nighttime, so you can do it for multiple parameters, and you can calculate all these. These are all done so that the measure that you implement to improve, it should be specific. So, in this type of situation, you must look at possibly midblock section is the most vulnerable section in this case. So, you must improve the safety of the pedestrians crossing at midblock for example maybe, so, what type of intervention do you put? You should not go for an intervention that is more effective at a signalized intersection versus now you have to think of something that is more practical for midblock sections. So that is the reason why you develop all different types of descriptive statistics.

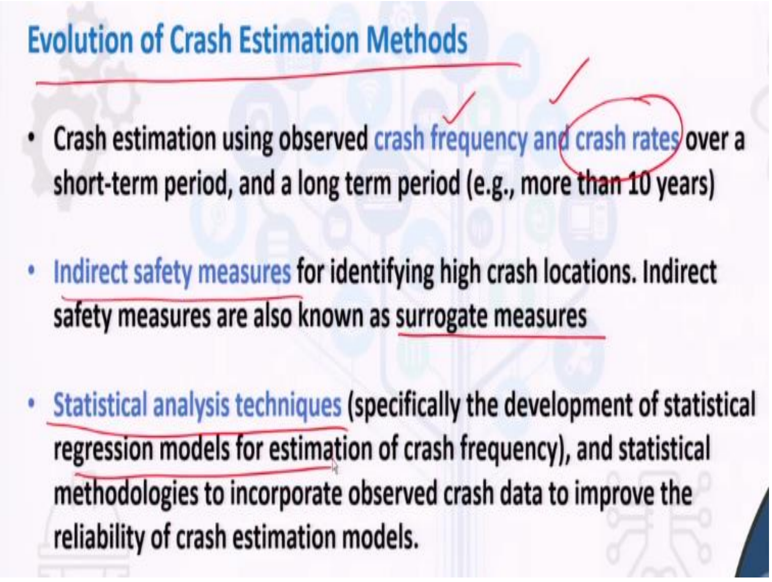
(Refer Slide Time: 12:56)



You also must know quantitatively how well to predict. Prediction always has a lot of errors. Now, when you are trying to predict anything, you need essentially lot of input values. The better the input values, the more precise the input values, the better will be your prediction. So, a system that we are moving towards is to capture all incidents or all crashes that happen at a roadway and write down the parameters then and there. So, when an accident happens, you would see that not only the police arrives, maybe the ambulance arrives at that spot, we are also trying to have an engineer along with them arrive at the spot so that the engineer can help point out some deficiencies in the roadway, design, or the intersection design or whatever it is. So that all the parameters are captured at the same time of the crash.

By collecting all those parameters, then you can hope to develop a predictive model which will give you better results. So, it is important for us to do such kind of accident investigation right at the spot and you must have different types of people come there, so not only the ambulance, not only the police, but also maybe engineers on the spot so that they can verify how different parameters or what are the different parameters that were involved for that crash.

(Refer Slide Time: 14:38)

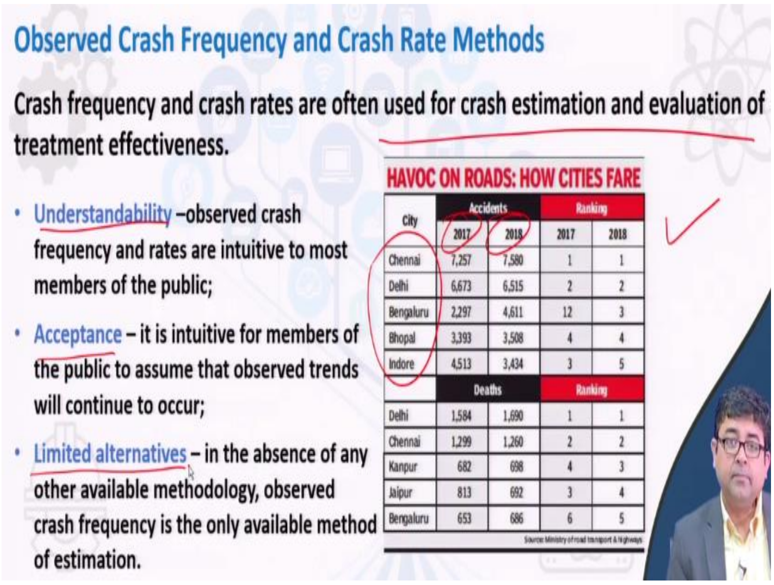


Now, how have the different crash estimation methods evolved? We have already talked about crash frequency. The other thing that people now look at rather than just crash frequency is something called a crash rate. That tells you that how many crashes are happening per n number of miles of vehicles driven on that road. So, the point is the more the number of vehicles, the likelihood that the number of crashes will increase is higher.

So, it is not just the frequency or how many accidents are happening per year or per month, it is also how many vehicles are on that road or how many vehicles have been driven on that road. So that is something is called a crash rate and then there are of course a lot of indirect safety measures known as surrogate measures that have been developed. Surrogate measures usually mean for example if you do not wear a helmet, you are more likely to be involved in an injury or fatality related crash if you are on a two-wheeler. So, people are trying to relate the wearing of helmet or the wearing of a seatbelt when you are in a car to the probability of you being involved in a major crash or a major injury crash. So different types of crashes you can categorize that into. So, these are called surrogate measures or indirect measures. So, when you do not have good information, good data on your roadway segment direct number of accidents or information that is available firsthand, if you do not have that kind of data, then what you do usually is try to predict an accident using different types of indirect measures and those indirect measures are usually called surrogate measures. And then there are of course statistical techniques, which looks at what is the best model that fits the existing number of accidents. Usually, people will see that can regression models estimate crash frequency or can Poisson models or negative exponential models fit the crash frequency that is happening.

So, those are all statistical tools, then again if those can also help in predicting future crashes if the models fit is very good. If the models fit is very good, then you can predict these crashes in the future as well.

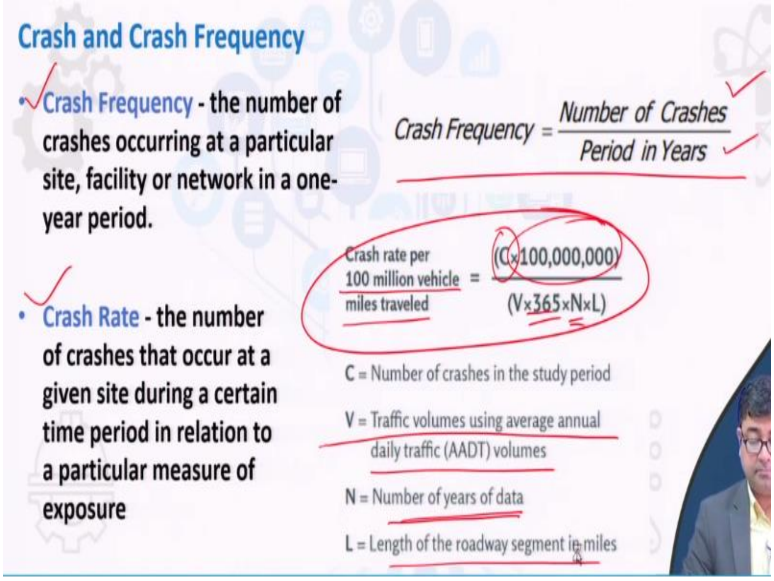
(Refer Slide Time: 17:08)



So, crash frequency or crash rates are often used for crash estimation and evaluation, they are simplistic. So, for example, you many times see such statistics in the newspapers or in the digital media or in the TV showing the number of accidents that have happened per city per year and how well they rank. So, these are easy to understand and since these crashes have already happened, so the validity, the acceptance of these numbers is high.

People do believe in these numbers, but you know it does not provide a means of predicting crashes, it is just a very reactive method. So just know about the crash after the crashes happen. That is an old way of trying to deal with reducing the number of crashes. So, it does not give you alternatives, it does not give you different methods of prediction or anything like that.

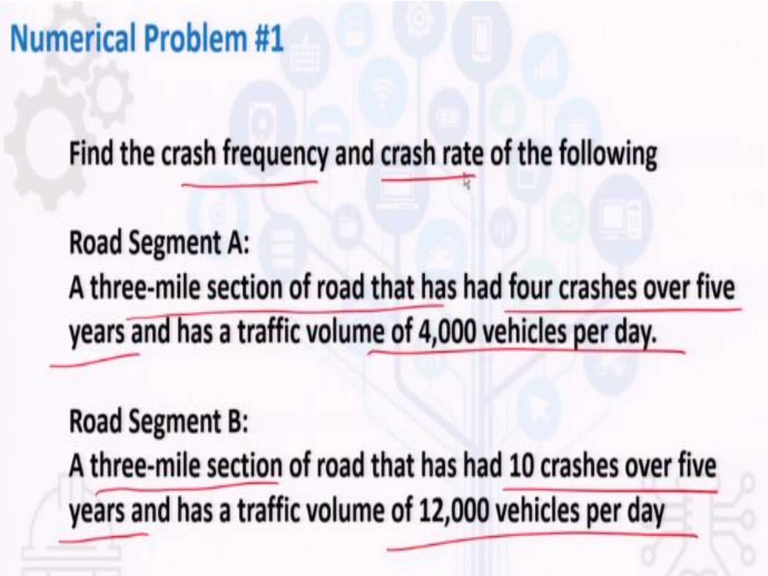
(Refer Slide Time: 18:11)



However, if you look at each of those methods the crash frequencies, simply the number of crashes divided by the period in either years or months or whatever you are looking at and crash rate, which is a better measure than crash frequency. What it says is that a crash rate per 100 million vehicles miles traveled. So, if more the number of vehicle miles VMT or VKT traveled on the road, the number of crashes is likely to be higher.

So, number of crashes into 100 million divided by the volume on that road every day, number of years you want it for the total length segment that you are looking at. So that is essentially how you determine the crash rate.

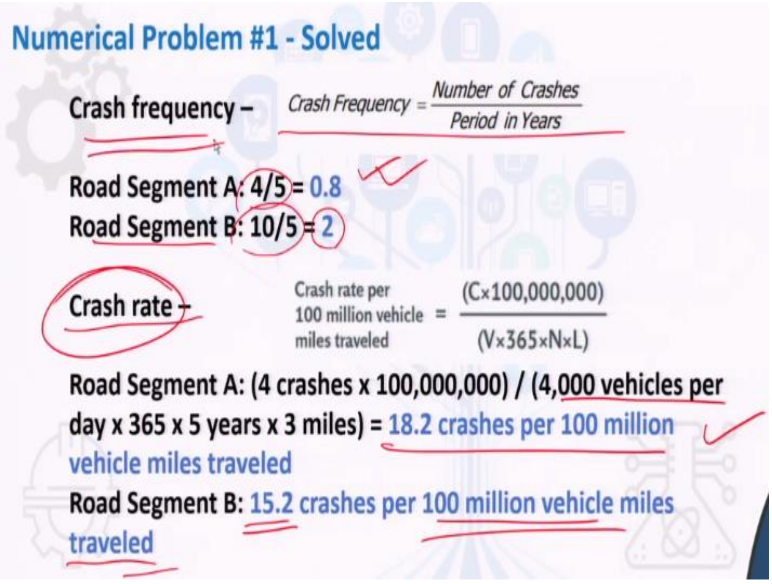
(Refer Slide Time: 19:00)



So, if you are given simple problems like this, say for example on road segment A, which is a three-mile section, it has had 4 crashes over 5 years and has a traffic volume of 4000 vehicles per day, whereas the next segment B is also a three-mile section, but it has had 10 crashes over 5 years and has a traffic volume of 12,000 vehicles per day. So, if you determine crash frequency and crash rate, what do you expect to see?

Which is the worst segment, or which is the better segment? That is what you essentially want to compare these two segments to see which is the worst segment and then maybe prioritize. Usually, you want to help both segments, but if you must only prioritize, then you would have to pick which segment to improve.

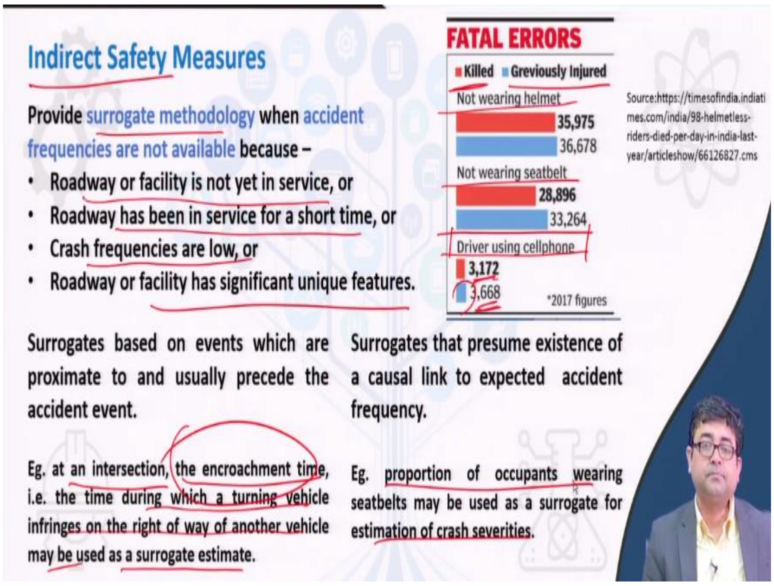
(Refer Slide Time: 19:48)



So, if you purely looked at crash frequency, then you would have seen that segment A where 4 crashes in 5 years 0.8, whereas 10 crashes in 5 years is 2. So, you would have thought that segment B is a worse segment. However, if you start now looking at the crash rate which depends upon how many vehicles are traveling on that road, then you would see that segment A, the crash rate in segment A is 18.2 crashes per 100 million vehicles traveled.

Whereas segment B is only 15.2 per 100 million vehicles traveled that gives you an idea of maybe road segment A is a more serious situation and we must focus more there because the crash rate is higher on segment A, although crash frequency is lower. So that is why more and more people are looking at crash rates rather than just the frequency of crashes.

(Refer Slide Time: 20:50)



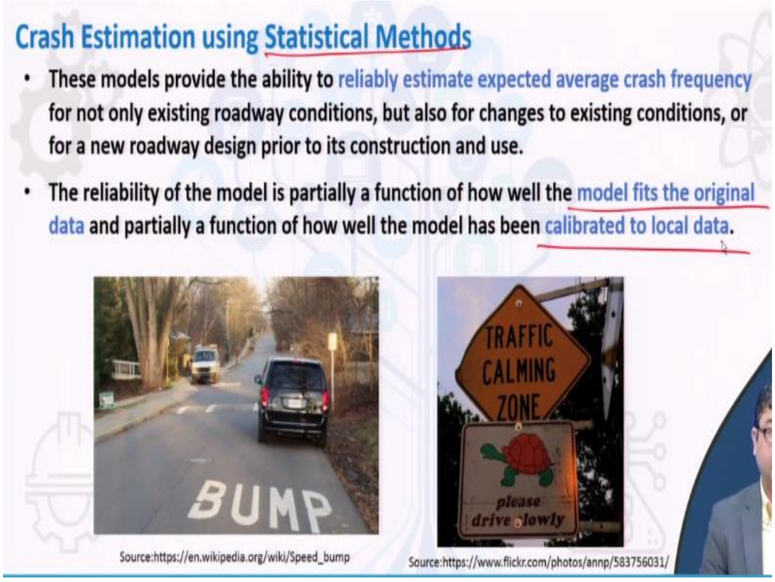
What are the indirect means of understanding safety, like we said the surrogate methodology is used primarily because maybe a road or a facility is not yet in service, so it is a road that is just being built, so you do not have any previous data on the road, so how can you predict what the crashes will be on that road or it has been only just open for the last few months or a few weeks, so you do not have a lot of data on it.

So, how can you then predict, or maybe crash frequencies are low. Frequencies are low; however, people are complaining that there are a lot of accidents that are leading to fatalities maybe. So, you do not know then how you deal with it or if it has some unique features like it has lots of curves, horizontal curves, or vertical curves more than the average amount number of curves.

So, such unique features may also lead to developing a unique surrogate measures which tells you about the safety of that area. So, here you have seen that people are, like we were talking, not wearing helmet versus not wearing a seatbelt and drivers using cellphone. So, it has been noticed that if you are using a cellphone, then the number of grievous injuries is higher and even the number of killed is that much. So, you can then use these surrogate measures to predict how crash frequency will happen and what type of crash will happen. For example, say for example at an intersection the encroachment time that is the time during which a turning vehicle infringes on the right of way of another vehicle can be used as a surrogate estimate. So, for example if you have you know free left turns and many intersections you have.

So, the free left turn time and through vehicle is coming on the main road, so there is a possibility that these two vehicles may collide or during the free left turn, maybe somebody is trying to cross, a pedestrian is trying to cross that road. So, they may interact or collide with each other. So, the encroachment time may be a surrogate measure for safety at an intersection and we have already looked at or talked about wearing seatbelts or not wearing helmets being a surrogate measure.

(Refer Slide Time: 23:25)

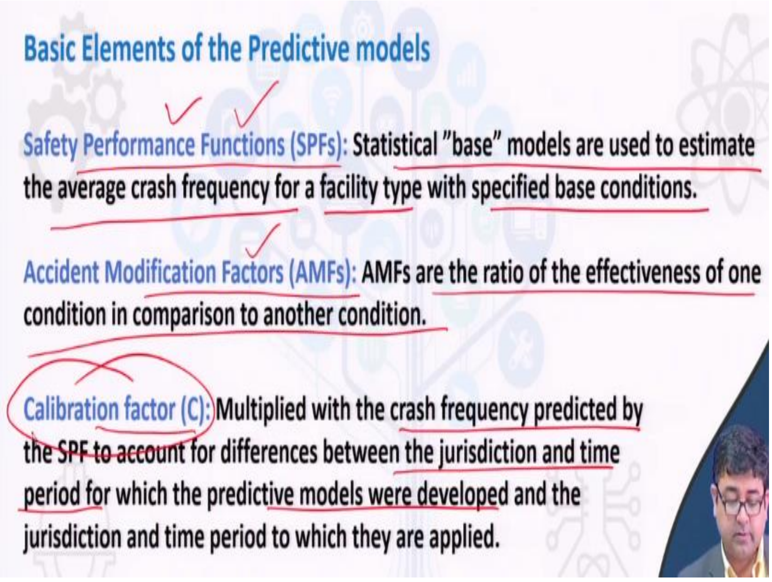


When we start looking at the statistical method, we have already told you that these may be either simple regression equations or they may also be Poisson or negative binomial. People are tending to look at negative binomial and Poisson because usually when we look at the number of crashes that happen on one roadway for example over the year or two years, they are usually termed as rare events.

I mean rare events as in you would see that there are a lot of accidents happening, but in terms of count or in terms of if you count each crash as an event, so for over 365 days you may find you may find only there are 50 accidents that have happened. So, 50 out of 365 days seems like a rare event. So that is why some researchers are modeling crashes as rare events. However, it is imperative that these statistical models have a good fit.

Otherwise, you may not be able to predict the future crashes very accurately and calibration to local data is important. So, once you have your statistical model, you then calibrate it. Maybe the model was developed in another city, in another town, in another country, you have to calibrate it to local conditions that validate it with the actual number of accidents that have happened. Only then statistical models can be used with confidence.

(Refer Slide Time: 25:08)



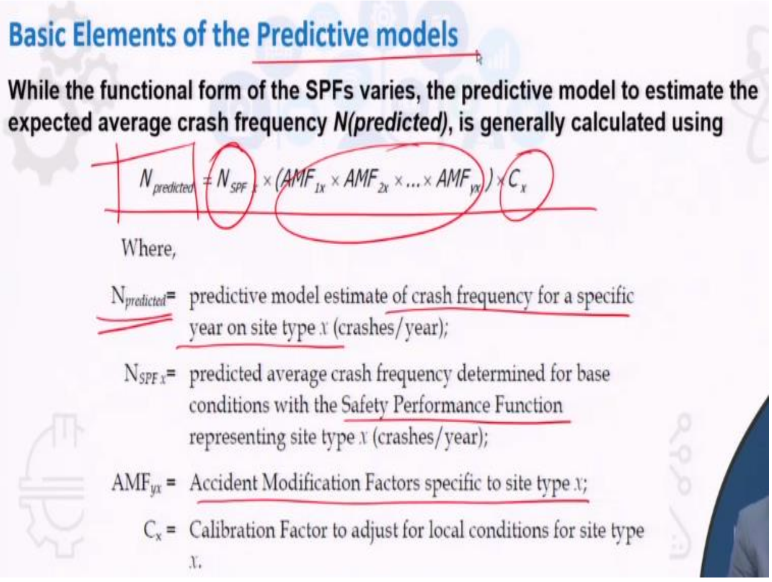
So, in these predictive models, what people use in general are these are what are called safety performance functions. So, if you are really a person that is working in transportation safety arena, then you will be dealing with a lot of these safety performance functions, SPF. These are statistical base models used to estimate average crash frequency for a particular facility with specified base conditions.

These are equations most likely regression equations, sometimes they are negative exponential equations that have been developed for a certain base condition. Now, if you take that to another condition then like we said you must calibrate it to those conditions and then use it in order to predict. So, we will show you what safety performance functions are. Then accident modification functions or AMFs are the ratio of the effectiveness of one condition in comparison to the other condition.

So, now you have developed a countermeasure, maybe you have implemented a countermeasure to reduce safety. So, what is the modification factor because of that implementation of that safety measures, maybe you have increased police patrolling. So because of increased police patrolling on that roadway how have accidents changed. So, the ratio of before and after gives you the actual modification factor.

Then somebody else in some other neighborhood or in some other city can also think that well if I have more police patrolling this roadway, maybe my accidents would also go down. So, those are called accident modification factors. Finally, the calibration factors are these are multiplied with the crash frequency predicted by the SPF to account for differences between the jurisdiction and the time-period for which they were developed. So, maybe this was developed 5 years ago for a different city. However, now you want to use it 5 years later for a different city. So, you must calibrate it to your area, to the time period. So, all those calibrations are involved in this.

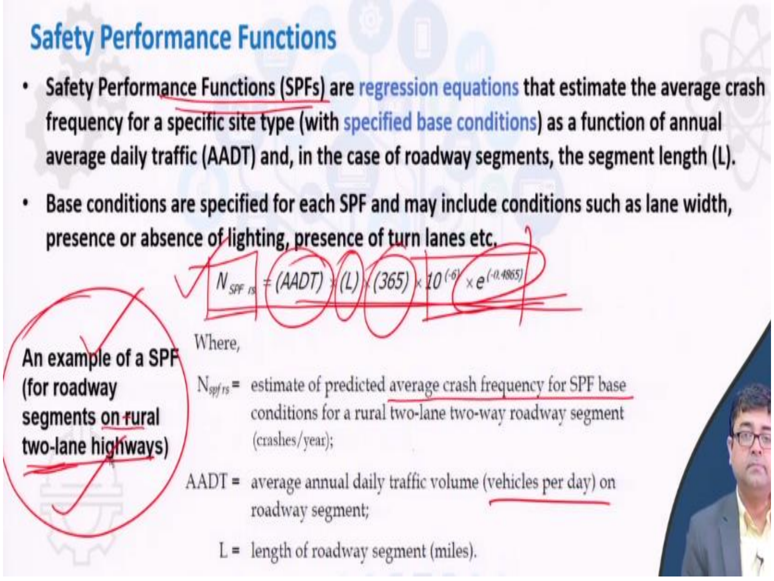
(Refer Slide Time: 27:28)



Usually, this is how you will see a predictive statistical model. So, N predicted or the number of crash or the crash frequency for a specific year on a specific type will be given by your SPF function, safety performance function, which is then multiplied by your accident modification factors and your calibration factor. So, this is a basic gross level model.

If you are trying to develop a predictive model for your area, it will have all these elements, but different factors will be there for your case calibration, the value of calibration may be less, maybe high. So, this is basically but this is how your model should look like.

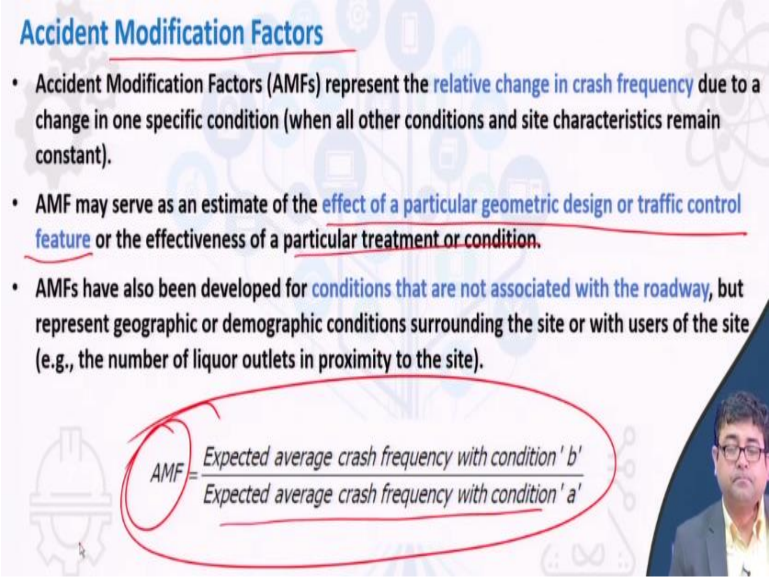
(Refer Slide Time: 28:19)



Then, if you have looking at developing these safety performance functions, usually they will look like something like this. So, for example a safety performance function for roadway segment on rural two-lane highways. This is as specific as that it must be for a particular type of facility. So, if you had to develop a SPF for a two-lane rural highway, maybe this will look like this.

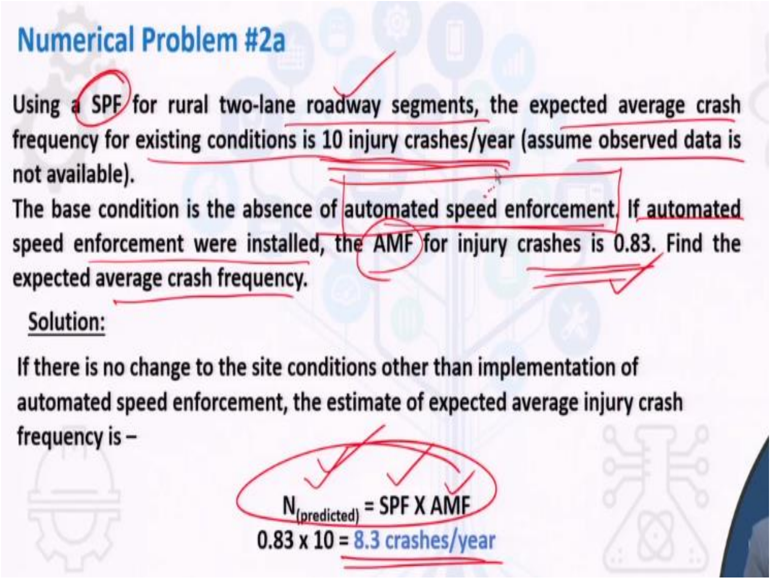
So, it has the vehicles per day, AADT, length of stretch, it is for all 365 days and then these have been developed based on all the data that has been collected. So, it looks like it is some form of exponential function that you are using to predict the average crash frequency for the base conditions. This is for base conditions for rural two-lane highways.

(Refer Slide Time: 29:28)



Now, like we said those are all for basic conditions. If you must now improve those conditions and you have taken certain strategies, maybe you must estimate the effect of geometric design or traffic control feature or the effectiveness of the particular treatment, so essentially you have to have accident modification factors that is the expected average crash frequency with condition b. Condition b is the improved condition, divided by the expected average crash frequency with condition a. So, that will give you by how much, how effective was your implementation on that roadway.

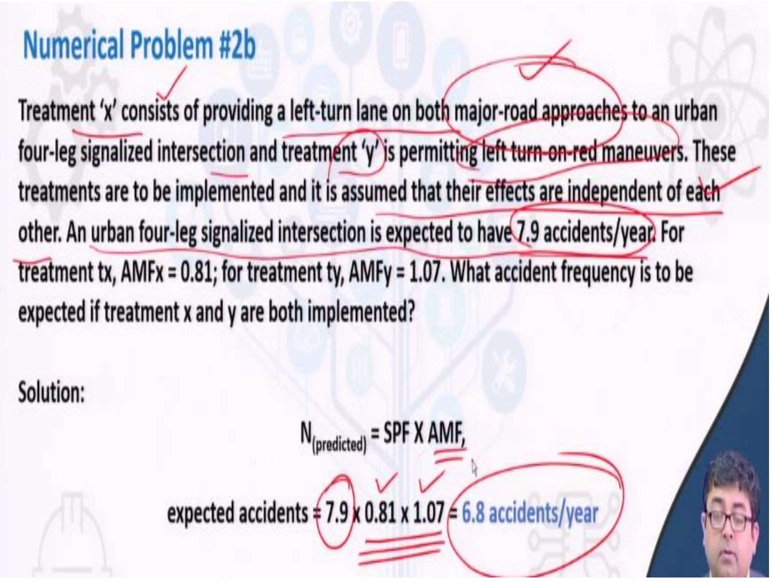
(Refer Slide Time: 30:13)



Simple examples using SPF for two-lane roadways, the expected average crash frequency of existing conditions is 10 injury crashes per year, assume observed data is not available. This is just using an SPF. The base condition in the absence of automated speed enforcement. So, this is the base conditions where automated speed enforcement was not available. However, if automated speed enforcement was installed, the AMF for the injury crashes is 0.83.

It has been noticed that it is 0.83. So, what is the average crash frequency that you can expect? You just have to you know that this is the formula which will give you the expected or the predicted average frequency. You know your SPF, and you know your SPF is 10 and you know your accident modification factor for this kind of intervention. This is the kind of intervention for which accident modifying factor of 0.83. So, now you can expect that there will be 8.3 crashes per year, which are down from 10 crashes per year. So, that is quite simple.

(Refer Slide Time: 31:24)

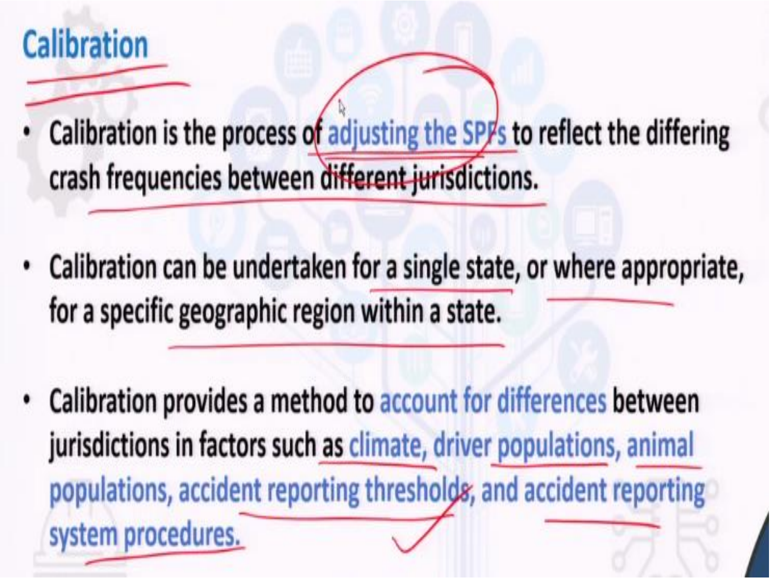


Similarly, if there is a treatment x which consists of providing a left-turn lane on both major-road approaches to an urban four-leg signalized intersection and treatment y is permitting left turn on red maneuvers, so you have free left turns that you can do. Both treatments are to be implemented and it is assumed that their effects are independent of each other. So, if both are implemented what will happen and if you assume that their effects are independent.

So, an urban four-leg signalized intersection is expected to have, so you have seen from other data somewhere else then it is expected to have 7.9 accidents in year, but at your location you have implemented these two treatments. So, how much do you expect it to change by? So, this is what is the base condition. You have these two accident modification factors, since they are independent of each other, you can multiply both.

And you can see that your accidents might reduce to 6.8 accidents per year. So that is quite simple understanding of how to carry out the statistical modeling.

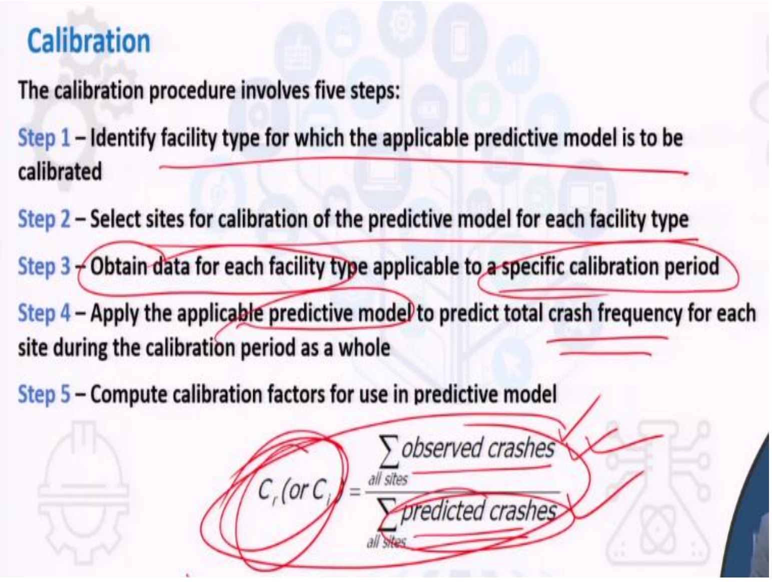
(Refer Slide Time: 32:30)



Calibration. The last thing in your understanding how to predict crash frequencies is the process of adjusting the SPF to reflect different crash frequencies of different jurisdictions. Calibration can be undertaken for a single state or where appropriate within different geographical areas within the status as well. It accounts for factors such as climate, driver population, animal population, accident reporting thresholds, and accident reporting system procedure.

So, all of these may vary from place to place. So, in order to take into account all of these you have to adjust your SPF or your safety performance functions, only then you can use it for your jurisdiction. So, accident modification factors are interventions that you take, whereas calibration is you calibrate it to your jurisdiction. You have not implemented anything yet. But because the SPF was developed for a different jurisdiction, you had to bring it to your jurisdiction, so you must calibrate it, calibrate that function, or calibrate that equation or model to your jurisdiction. So that is kind of what the calibration is.

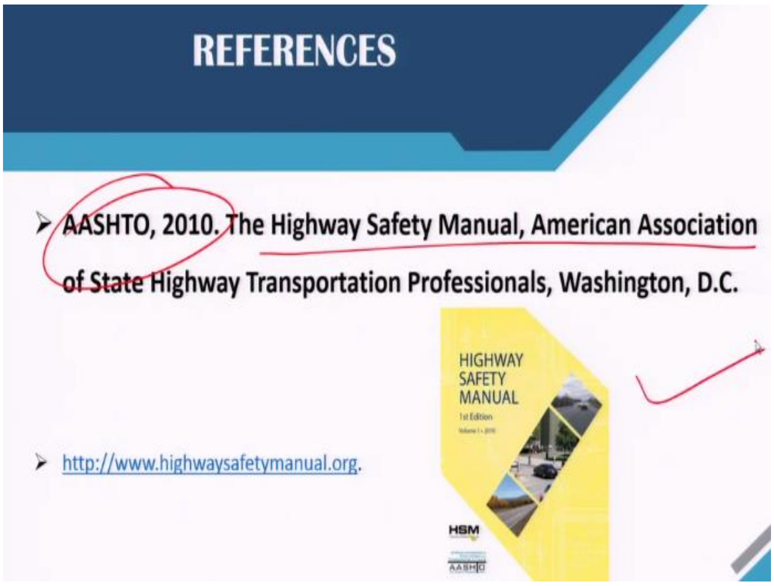
(Refer Slide Time: 33:50)



Simply the calibration factors are again summation of all observed crashes divided by summation of all predicted crashes. So, once you identify the facility type, select sites for calibration, obtain data for each facility for the specific calibration period, then you apply your predictive model to predict the crash frequency. Once you get the crash frequency, you also know what the observed crashes in your area are.

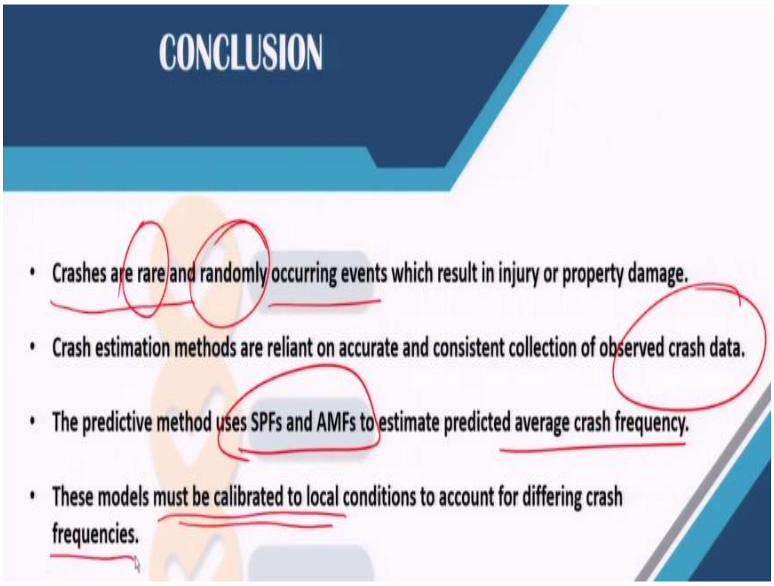
So, while you are calibrating you are not yet predicting it, you are calibrating so that your model should be as close as possible to the actual crashes that have happened. So, if this value is close to 1, then you would say that your observed and predicted crashes are very close to each other, your equations are predicting, or your model is predicting the crashes very well. So, then your calibration factor is almost close to 1.

(Refer Slide Time: 34:53)



So, that is the end of this lecture. Most of the materials here have been taken from the highway safety manual that has been developed by AASHTO in 2010 in USA. This is a particularly good resource when you are talking about safety in the highway arena in the transportation arena. So, I recommend that you guys go through this manual very carefully.

(Refer Slide Time: 35:21)



In conclusion, we have introduced all the different types of crashes and how they occur. They are very random and rare, so it is important to have prediction tools so that you can then implement certain safety measures to reduce the occurrence of these crashes. We have introduced you to different types of safety performance functions and accident modification functions that allow you to use statistical tools to predict the average crash frequency in your area.

We have told you how to determine crash rates and finally it is important to know that these safety performance functions must be calibrated to local conditions in order to account for different crash frequencies. Thank you.