FOOD SCIENCE AND TECHNOLOGY

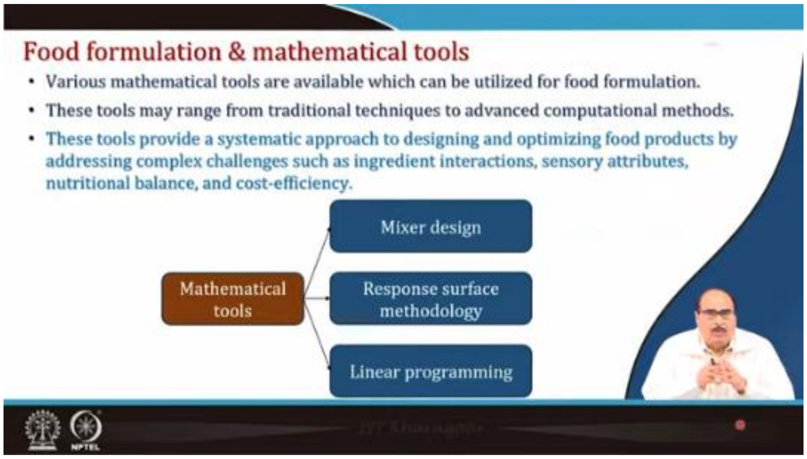
Lecture 48

Lecture 48: Mathematical Tools for Food Formulation

Hello everyone. Namaste.

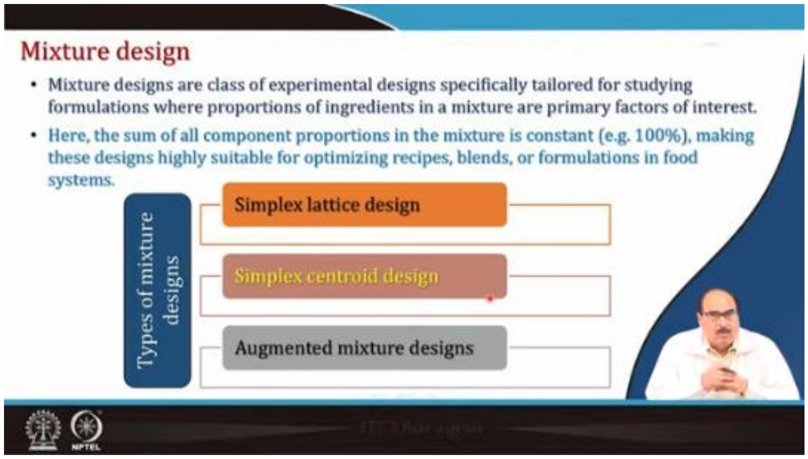
In the earlier class, we talked about food formulation and design. In today's class, we will study the various mathematical tools for food formulation.

The topics we will discuss in today's class include mixture design, response surface methodology, and linear programming. Finally, we will also talk about consumer-based food formulation.

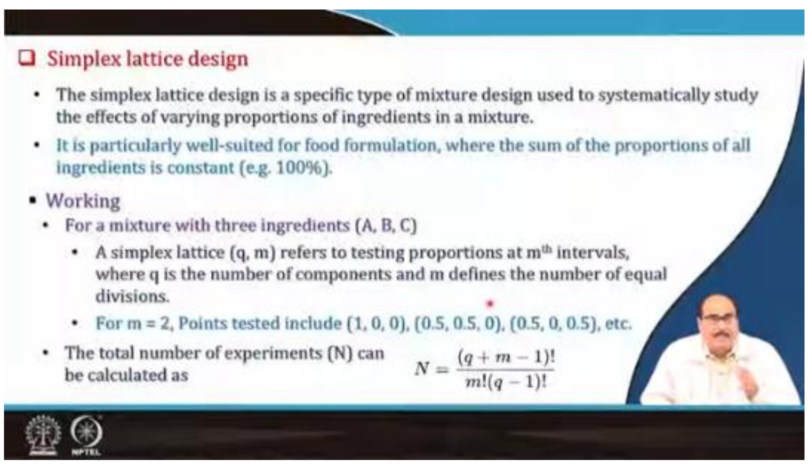


You know, various mathematical tools are available which can be utilized for food formulation. These tools may range from traditional techniques to advanced computational methods. These tools provide a systematic approach to designing and optimizing food products by addressing complex challenges, such as the interaction among ingredients, sensory attributes of various foods, nutritional balance, and, more importantly from a commercial point of view, cost efficiency.

So, the different mathematical tools that are available for designing and formulating food include mixture design, response surface methodology, and linear programming. So, let us see what a mixture design is.



Mixture designs are a class of experimental designs specifically tailored for studying formulations where the proportion of ingredients in a mixture are the primary factors of interest. Here, the sum of all component proportions in the mixture is constant, that is, 100 percent. And this makes the design highly suitable for optimizing recipes, blends, or formulations in food systems. The mixture design again may be of several types. It may be a simplex lattice design, simplex centroid design, or augmented mixture designs. Let us see what a simplex lattice design is.

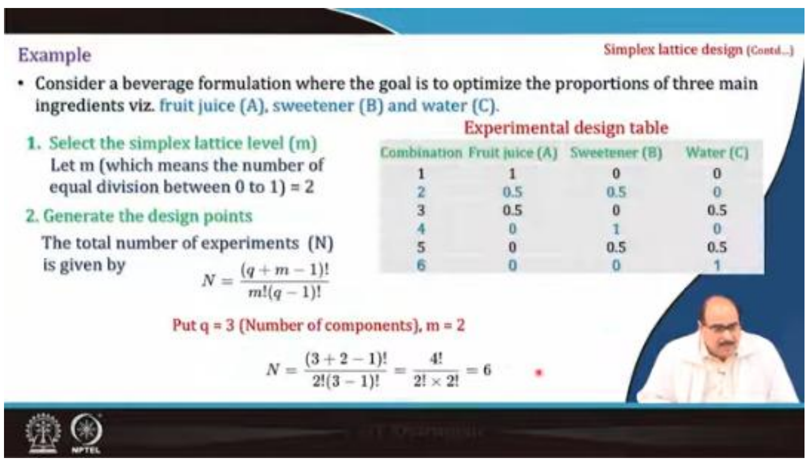


The simplex lattice design is a specific type of mixture design which is used to systematically study the effects of varying proportions of ingredients in a mixture. In a mixture, suppose there are several ingredients. So, what is the contribution of each ingredient if you vary the proportion of each ingredient? What is its contribution, and how will it influence the response variable or responses, etcetera? So, that you can find out in this, and it is particularly well suited for food formulations where the sum of proportions of all ingredients is constant, that is, 100 percent. So, how does it work?

Now, let us see for a mixture with three ingredients: A, B, and C. A simplex lattice (q, m) refers to the testing proportions at the m-th intervals, where q is the number of components and m defines the number of equal divisions. So, if we take m = 2, that is, the m is the interval. So, m = 2. Then, points tested include (1, 0, 0), (0.5, 0.5, 0), (0.5, 0, 0.5), and so on. So, the total number of experiments as per this design, that is n, can be calculated. using the formula:

$$N = \frac{(q+m-1)!}{m!(q-1)!}$$

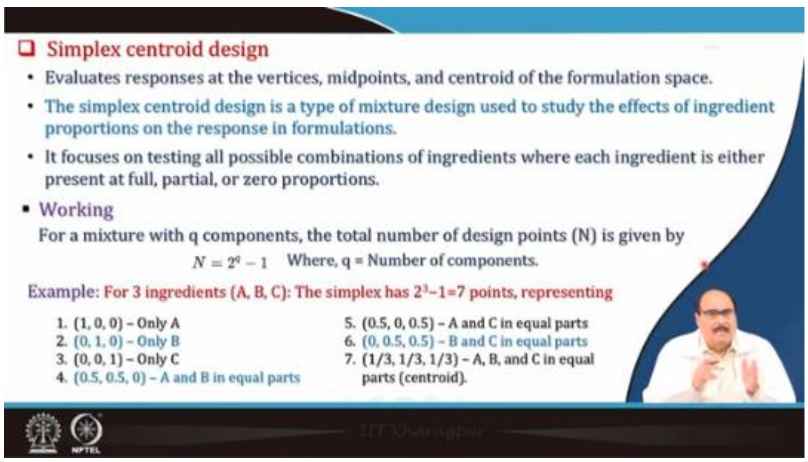
So, let us take one example and understand this design.



Like, consider a beverage which is being formulated, and the objective is to optimize the proportions of the three main ingredients. In the beverage, there is A fruit juice, the second ingredient is sweetener, and the third ingredient is water. So, we want to optimize the proportion of these three ingredients using this design. So, the first thing is that you have to select the simplex lattice level, that is m, as we discussed earlier, which will be the interval. So, let m, which means the number of equal divisions between 0 and 1, be equal to 2. Then, generate the design points. The total number of experiments in the formula we saw in the earlier slide. So, here, for q, that is the number of components, there are 3 components: fruit juice, sweetener, and water. So, q = 3, and m = 2. So, putting these values in the equation,

$$N = \frac{(3+2-1)!}{2!(3-1)!} = \frac{4!}{2! \times 2!} = 6$$

So, in this case, there will be 6 combinations, and accordingly, the experimental design table will have a combination. The first may include 100% fruit juice, 0% sweetener, and 0% water. In the second combination, it may be 0.5 proportion fruit juice, 0.5 sweetener, and 0 water. Then, the third combination may accordingly be 0.5 fruit juice, 0 sweetener, and 0.5 water. The fourth combination may be 0 fruit juice, 1 sweetener, and again 0 water. And accordingly, the fifth combination will have no fruit juice, and the sweetener and water will be 0.5 and 0.5. The sixth combination may be 0, 0, and water becomes 1. So, in this way, you formulate the design, then finally conduct the experiment, evaluate the responses, and proceed with further optimization.



The second is the simplex centroid design. This simplex centroid design evaluates responses at the vertices, midpoints, and centroid of the formulation space. The simplex centroid design is a type of mixture design used to study the effects of ingredient proportions on the responses in the formulations. It focuses on testing all possible combinations of ingredients where each ingredient is either present at full, partial, or 0 proportion.

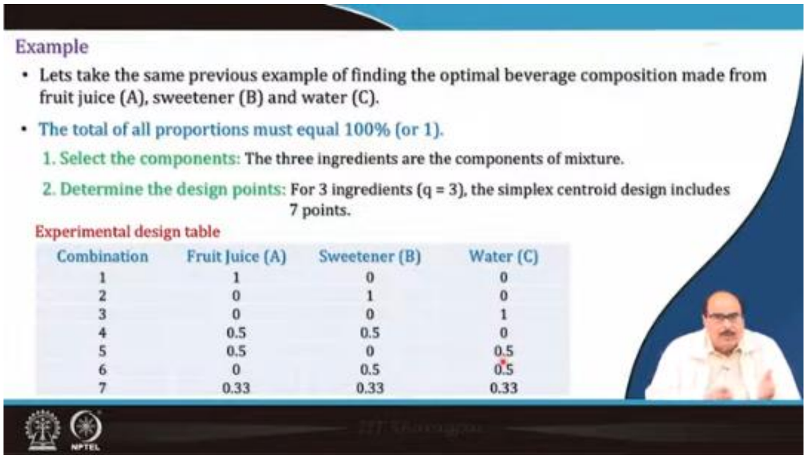
Let us see how it works for a mixture design with q components. The total number of design points (the combinations N) is given by the equation:

$$N = 2^q - 1$$

where q is the number of components. So, for the 3 ingredients which were A, B, and C in the earlier example. So, here as per this simplex centroid design, the number of experiments will be, that is,

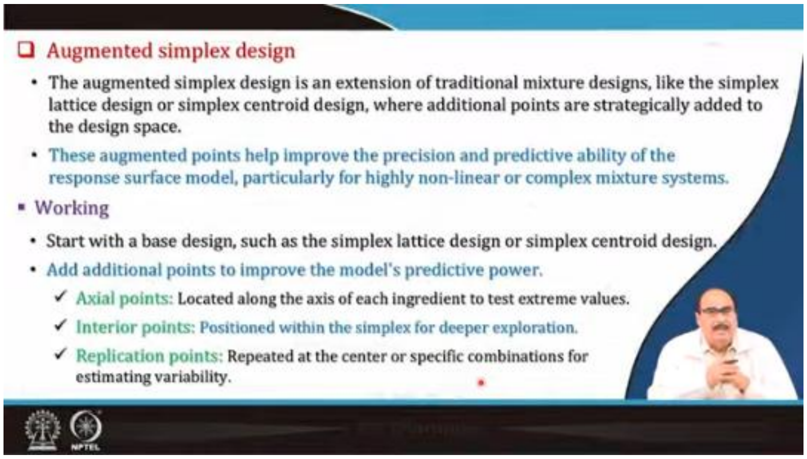
$$N = 2^3 - 1$$

And these 7 points may include (1, 0, 0), where only A is represented. There may be a second combination (0, 1, 0), where only B is represented, and then a third combination (0, 0, 1), which includes only C. And then the fourth combination may be (0.5, 0.5, 0), meaning A and B are in equal parts. In the fifth combination, A and C are in equal parts (0.5, 0, 0.5). In the sixth combination, B and C are in equal parts, like A (0), and B and C are 0.5 each. And in the seventh combination, it will be A, B, and C, all three components in equal parts, that is, 1/3, 1/3, and 1/3.

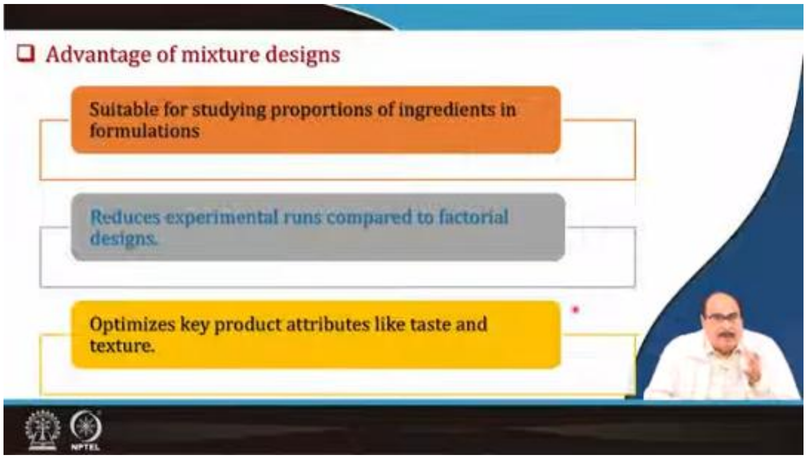


So, again we take the same example we saw earlier, that is, the fruit juice with three components: A the juice, B sweetener, and C water. So, the total of all proportions must obviously be equal to 100 percent or 1. So, to select the components, that is, the 3 ingredients are the components of the mixture, determine the design points. Again, for 3 ingredients, q = 3, then as per the calculation in the last equation we saw, there will be 7 points. And these 7 points will accordingly form the experimental design table, that is, one (1, 0, 0), (0, 1, 0), (0, 0, 1), (0.5, 0.5, 0), (0.5, 0, 0.5), the sixth combination will be (0, 0.5, 0.5), and in the 7th combination, it will be all 0.33, 0.33, and 0.33.

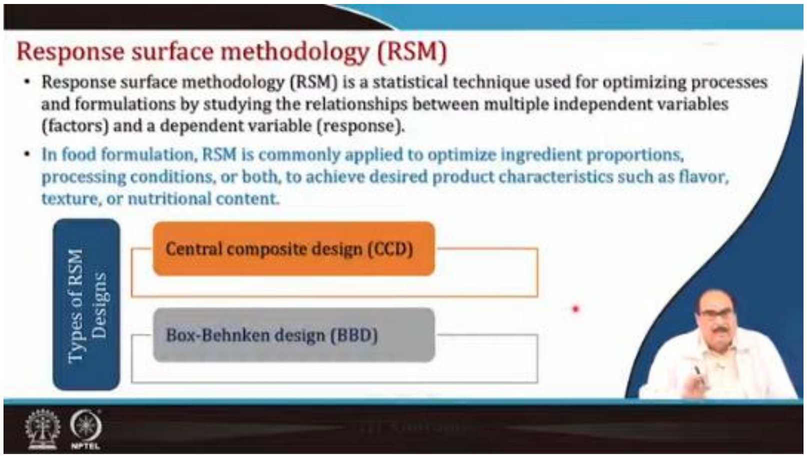
So, here again, now that is needed. One has to conduct experiments varying these proportions, having these proportions in the mixer. Same in the experiment to be conducted, and then you go for the next method of optimization. Fitting the quadratic equation or whatever equation values that I will tell you a little later.



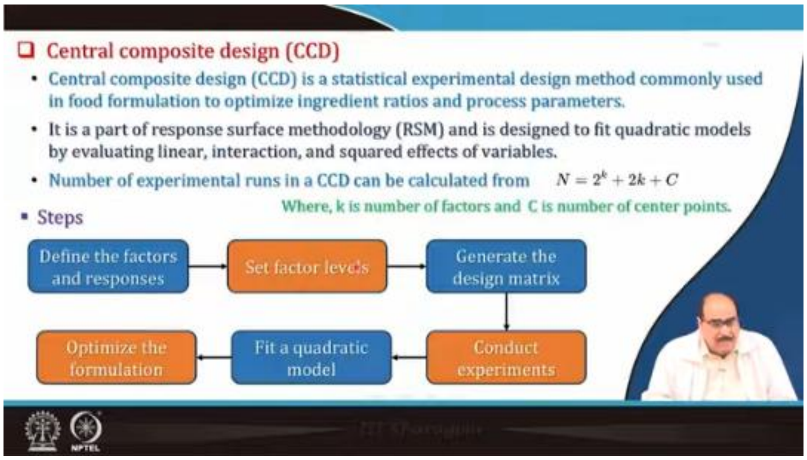
Then, augmented simplex design. The augmented simplex design is an extension of traditional mixture designs like the simplex lattice design or simplex centroid design, where additional points are strategically added to the design space. These augmented points help improve the precision and predictive ability of the response surface models, particularly for highly non-linear or complex mixture systems; these designs become more useful. So, again, the working of this design is to start with a base design such as a simplex lattice design or simplex centroid design. Add additional points to improve the model's predictive power. Additional points may be axial points located along the axis of each ingredient to test extreme values. Interior points positioned within the simplex for deeper exploration, and replication points, that is, repeated at the center or specific combinations for estimating variability. So, this gives the augmented simplex design.



Advantages of using mixture design in food formulations are that, it is suitable for studying proportions of ingredients in formulations. It reduces experimental runs compared to factorial designs and optimizes key product attributes like taste and texture.



Now, let us talk about response surface methodology, more commonly known as RSM. RSM is a statistical technique used for optimizing processes and formulations by studying the relationship between multiple independent variables, which are generally called factors, and a dependent variable, that is, response. So, in food formulation, RSM is commonly applied to optimize ingredient proportions, processing conditions, or both to achieve desired product characteristics such as flavor, texture, or nutritional content. The types of RSM designs may be central composite design (CCD or CCRD) and Box-Behnken design (BBD). So, let us see central composite design, what is this?



CCD is a statistical experimental design method commonly used in formulation to optimize ingredient ratios and process parameters. It is part of response surface methodology and is designed to fit quadratic models by evaluating linear, interaction, and squared effects of variables.

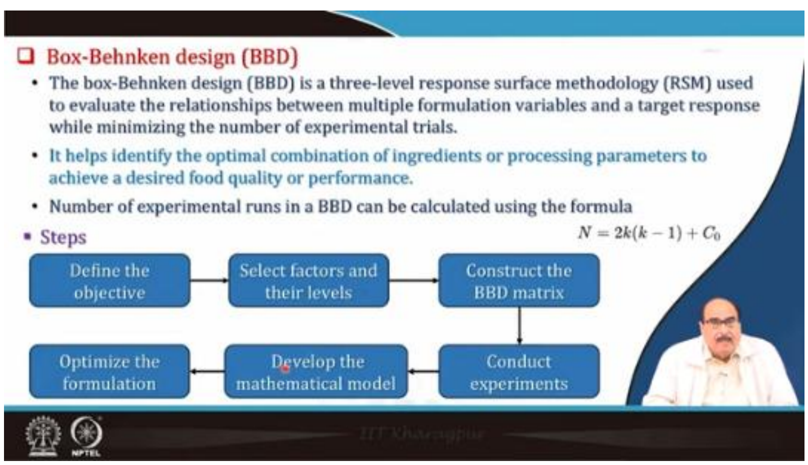
The number of experimental runs in a CCD or central composite design can be calculated from the equation

$$N = 2^k + 2k + C$$

where k is the number of factors. Suppose it may be 5 factors, 3 factors, or 4 factors. So, k is the number of factors, and c is the number of central points, how many central point’s one is taking.

So, the steps in this CCD, what are the steps to achieve this? Like number 1, it depends upon your requirement of the process, the requirement of the formulations, etcetera. You define the factors as well as identify the responses. Then, set factor levels, suppose there are 3 levels, 5 levels, when you are varying the factors (independent parameters). Then, the third step may be to generate the design matrix.

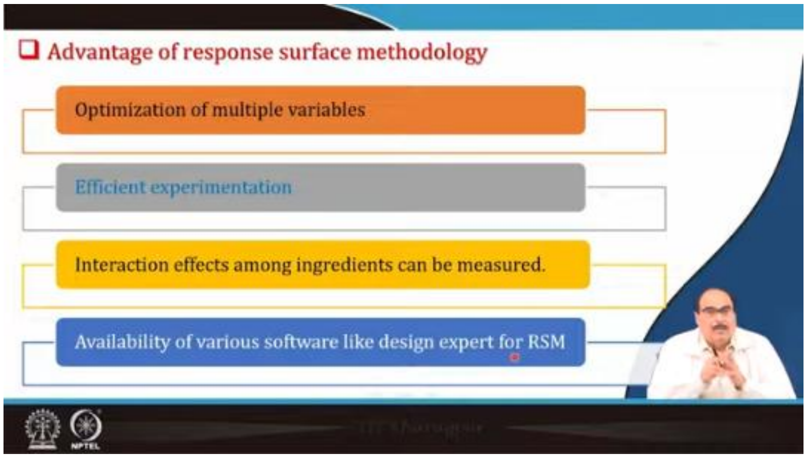
And once the design matrix is generated, you need to conduct experiments as per the design matrix, then fit a quadratic model, and finally, optimize the formulation. We will take an example again in this case as well, a little later.



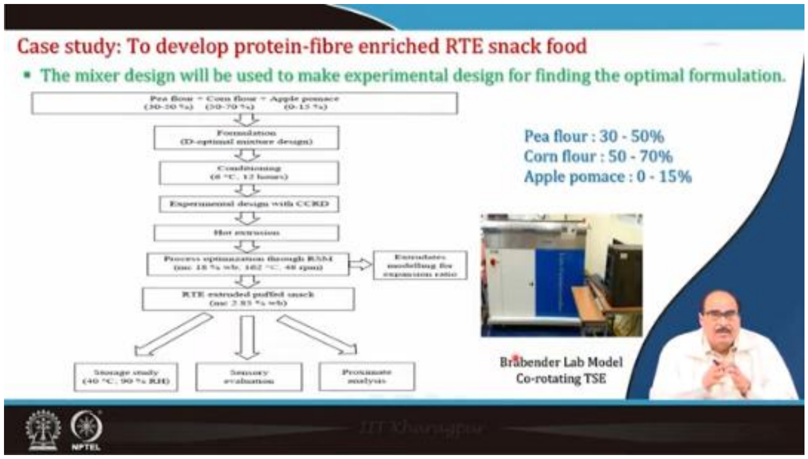
Box-Behnken design (BBD) is again a three-level response surface methodology used to evaluate the relationship between multiple formulation variables and a target response while minimizing the number of experimental trials. It helps identify the optimal combination of ingredients or processing parameters to achieve desired food quality or performance. The number of experimental runs in the BBD can be calculated using the formula,

$$N = 2k(k-1) + C_0$$

Now, the steps in BBD again: the first thing is to define the objective, what is the objective of your experiment? Select the factors and their levels, construct the BBD matrix, and conduct experiments as per the matrix. Develop the mathematical model and optimize the formulation. So, one has to go step by step in this process.



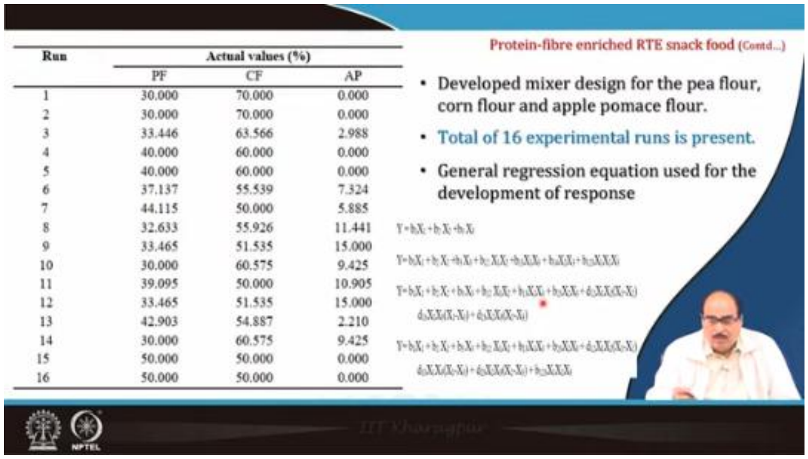
So, the advantages of the response surface methodology, you can optimize multiple variables; it gives an efficient experimentation technique, interaction effects among ingredients can be measured easily and the availability of various software like Design-Expert, etc., for the use of RSM. Once the variables, levels, etc., are decided, variables and factors and their labels, then there is various software available where you can input the variable number, it will design the matrix and tell you how many experiments one has to perform, and accordingly, once the experiments are done, then again there is software that will help you fit the quadratic models, etc., and develop your formulation.



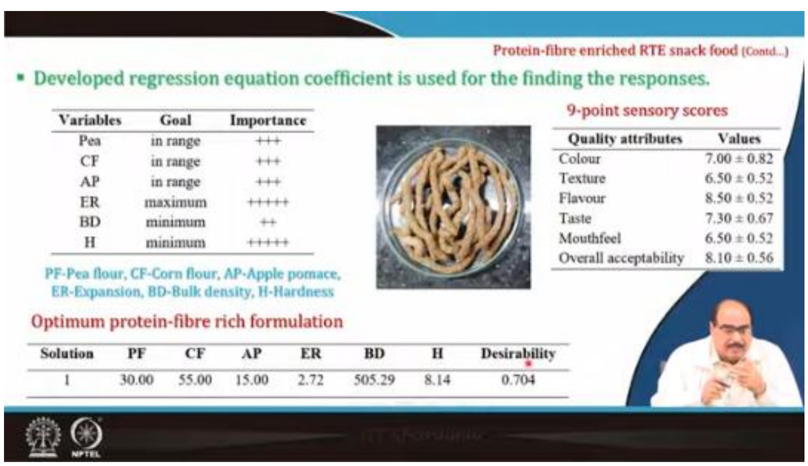
So, let us see; we will take one case study where our objective is to develop a protein-fiber-enriched, ready-to-eat snack food. And there are three ingredients in this: pea flour (30 to 50 percent), corn flour (50 to 70 percent), ranges that we have decided, the ingredients, that is, the factors and their variable ranges, and apple pomace, which may vary from 0 to 15 percent.

So, the formulations we decided that that is the D-optimal mixture design. Conditioning that it has to be done, then the experimental design was considered. After the formulation was finalized by the experimental design, the optimal mixture design was selected. Then, the experimental design was prepared using CCRD, and the experiments were conducted using a hot extruder. The process optimization parameters were varied during the process.

So, the extrusion process conditions were optimized, and we obtained ready-to-eat extruded food. Experiments were conducted in the Brabender lab model co-rotating twin-screw extruder. So, here the mixture design was used to create an experimental design for finding the optimal formulation. And then, finally, the CCRD was used to optimize the experimental parameters, the process parameters. So, the developed mixture design for the pea flour, corn flour, and apple pomace flour was this.



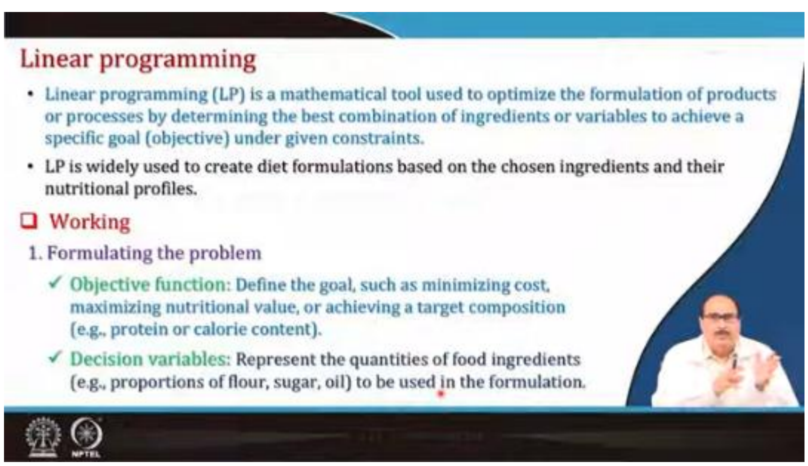
You can see that there were 16 experimental runs found. And the protein pea flour, corn flour, and apple pomace were varied from (30, 70, 0), (30, 70, 0), (33.446, 63.556, 2.9), and finally, even (50, 50, 0), the repeat point. So, these 16 experiments were decided. Now, next is that these data generated the regression equations used for the development of responses, fit the various equations quadratic models, different models are here, which is having the individual effect parameters, interaction effects all the things. So, there are various, even it is done in the computer, there are software. You can put this data, and it will fit that is a which model is getting it, that it will fit the model and then it will give the result.



Like developed regression equation coefficient which is generated in the as per the model through software, and these coefficients are used for finding the actual response and then we go for the optimization. And then for optimization, we give our conditions like in this case, instant case, that is you want that pea flour, this CF and AP. That is corn flour and apple pomace, all these ingredients, they should be in the range, that is we which we decided that is yes, that is the range is 30 to 50, this CF for 0 to 50 percent, AP 15 percent or so whatever range we give in the earlier slide I told you. And, then the further thing in the optimization, we want that expansion ratio should be maximum. So, that the snack you get good it's chewability, eatability, bulk density should be minimum, because if the bulk density is more it will become hard and then hardness also should be minimum. And, accordingly the importance how much we are giving to each of these variables also we decided.

And, once we decide these parameters, etcetera. Then again, in the program, you run the program, and it will give the optimum formulation, like the solution. Here it gives that protein pea flour 30%, corn flour 55%, and apple pomace 15%. In this proportion, we prepare and we get a product that has an expansion ratio of 2.72, bulk density of 505.29, and hardness of 8.14. So, this is the thing, and then finally, what we need to do is validate this.

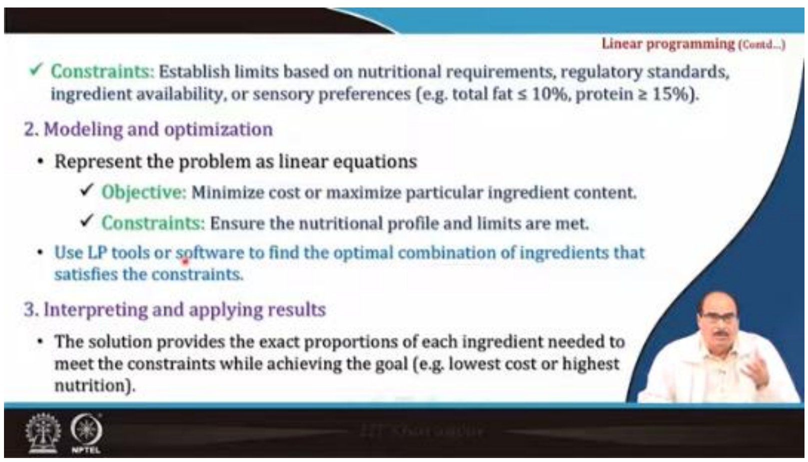
That is, the optimum solution is to prepare the product as per the optimum solution and analyze it. Maybe sensory characteristics, chemical parameters, and so, depending upon the type and nature of the product, more commonly we go for the sensory characteristics, determine and then find whether this product is acceptable or not, good in quality or not. So, that was it. Now next is the linear programming.



Linear programming again is a mathematical tool used to optimize the formulations of products or processes by determining the best combination of ingredients or variables to achieve a specific goal that is the objective under the given constraints. Linear programming is widely used to create diet formulations based on the chosen ingredients and their nutritional profile. So, how does it work?

That is first thing you have to formulate the problem. Where you take, decide your objective function that is define the goal such as minimizing the cost, maximizing nutritional value or achieving a target composition. For example, you want a particular amount of protein content, calorie content and whatever. So, that becomes your objective function.

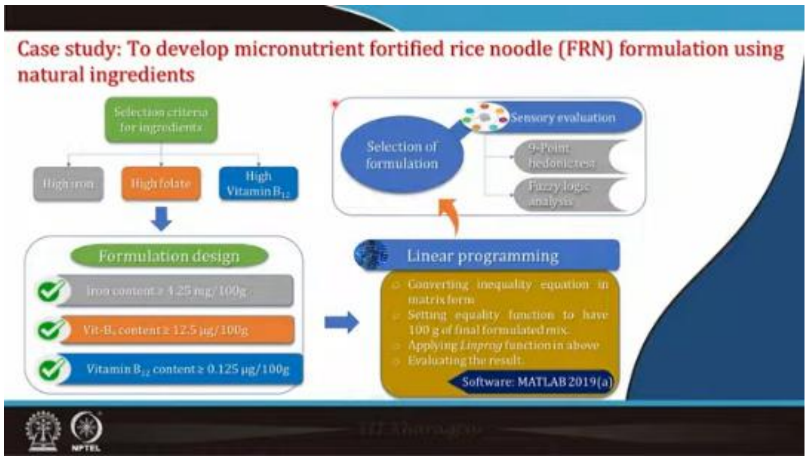
Similarly, you find out the decision variables, what are the decision variables which represent the quantities of the food ingredients like proportion of the flour, sugar, oil etcetera to be used in the formulation.



So, once you have decided the objective function as well as decision variables, then you also find the fix the constraints that is establish the limits based on the nutritional requirement, regulatory standards, ingredients availability or sensory preparations etcetera. For example, in a particular product you say that fat ≤ 10 percent, protein ≥ 15 percent or whatever the case may be. Then, once you have decided that, fixed it, then you go for modeling and optimization that is the modeling and optimization represents the problem as linear equation.

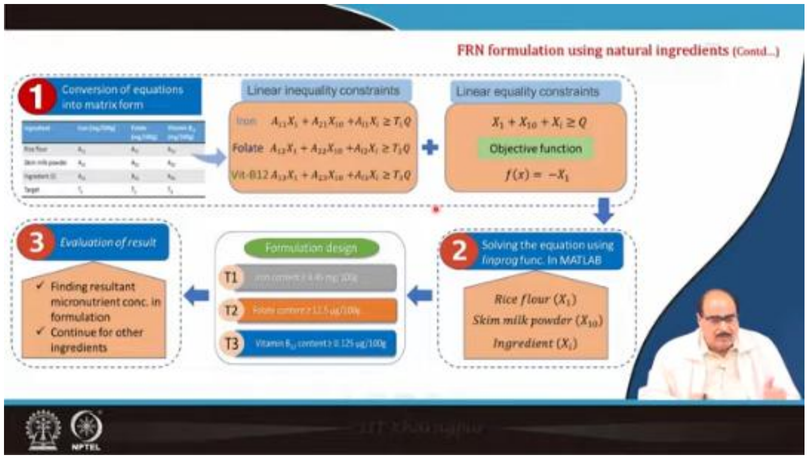
In this case you have to focus, you have to make the problem in the form of a linear equations, where objective is to minimize cost or maximize maximize practical ingredient, particular ingredient content. The constraints of your equation may be ensure that the nutritional profile and limits are met. Then use the linear programming tool or software to find the optimal combination of ingredients and that satisfies the decided constraints.

And then finally, interpreting the data and applying the results. The solution provides the exact proportions of each ingredient needed to meet the constraints while achieving the goal that is the lowest cost or highest nutrition whatever goal you have set.



So, here again we will take one case study that where we want to develop a micronutrient fortified rice noodles formulation using natural ingredients. The ingredients are that is we want that is ingredient selection criteria is that product should be of high iron content, high folate and high vitamin B_{12}. These are the three ingredients which want to optimize. Then we are going for the formulating design, and where the set of, that is, the limit of all the three ingredients iron, vitamin B_9 and vitamin B_{12} is decided it is given into this. Then once the formulation has been decided that is what is the objective.

Then you go for the linear programming software MATLAB that converts inequality equations into mixture form, setting quality functions to have 100 grams of the final formulated mix. Apply the linear programming function or linprog functions above and evaluate the result. So, it gives, that is, it will now provide the formulation or criteria, and then again you go and conduct experiments and evaluate this process.



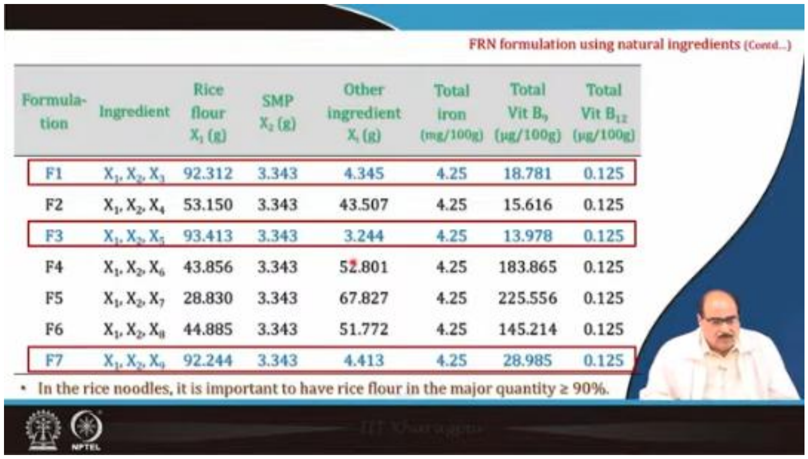
So, the same thing is shown here again: the first step is the conversion of equations into matrix form like rice flour, skim milk powder, ingredient i, or other what targets you get, and then you have linear inequality constraints for iron, folate, and vitamin B_{12}, as well as linear equality constraints, and the objective function is

$$f(x) = -X_1$$

and

$$X_1 + X_{10} + X_i \geq Q$$

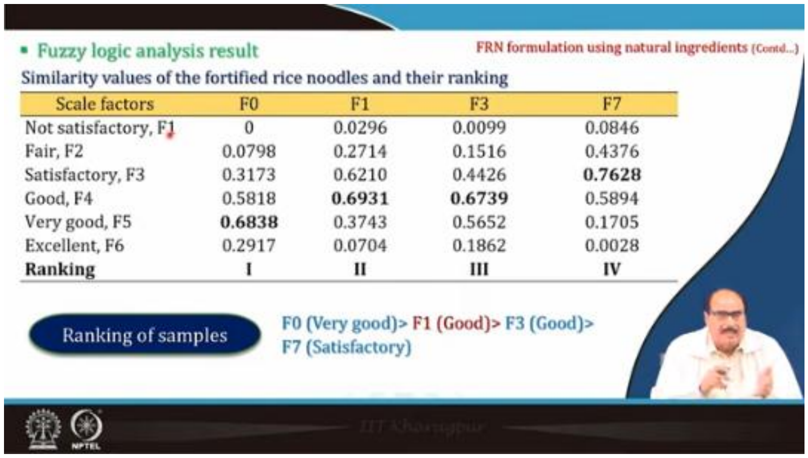
So, these are your linear inequality constraints, etc., in the form of equations. And then solve these equations using the linprog function in MATLAB. And it will give the combination of rice flour, skim milk powder in this case, or other ingredient i, that is X_1, X_{10}, or X_i, etc. And then you get the formulation design and finally, evaluate the result. So, as per this formulation, which we have given, the experiment that we conducted now gave 3 formulations.



That is F1, as you can see here, F3 and F7. It is provided that this is rice flour because one of the criteria, we have set is that the rice flour in the noodle should be more than 90 percent or equal to 90 percent. Now, in these three formulations, it is meeting that requirement, and other SMP and ingredients are also meeting it. So, it has given three formulations. So, you can go with the three, but we normally decide which one is best among these three.

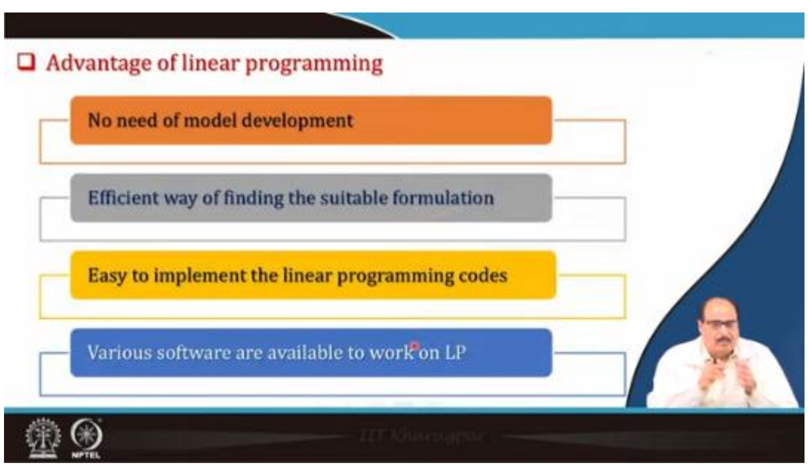
So, for that, what you do is take and conduct the experiment by the standard procedures. Like here, put water in this case. We are making noodles. So, in the flour combinations, they are mixed, water is added, then dough is made, and then extrusion. It is a continuous process or batch process, whatever is available with you, you go with this, dry, and you get the noodles.

And these noodles, it is also recommended that one control, F0, that is the control, is made, and these noodles now they are evaluated for their various characteristics, sensory responses.



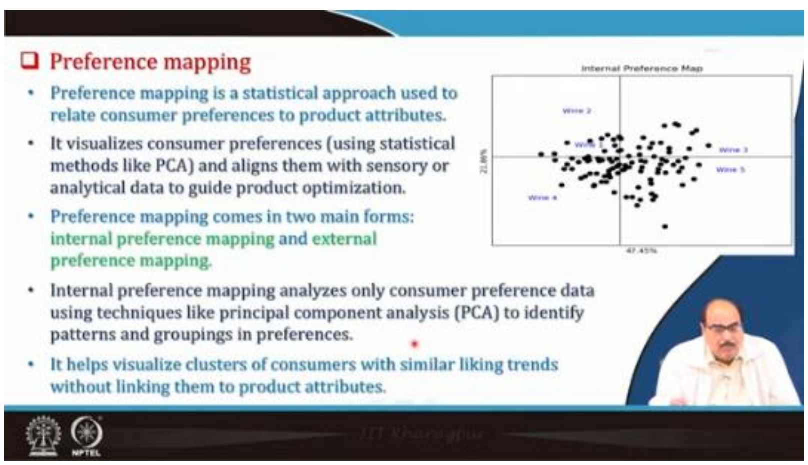
And then using the fuzzy logic approach, that is, these four products which were made that is F1, F2, F3, and F4 (factors) like this, there were the F0 is the control, F1, F3, and F7, these were the given by the design. And finally, we found that, yes, that is once using fuzzy logic, relative values were calculated, and it was found that, yes, obviously, the F0 control sample was best, then among the F1, F3, and F7,

F1 was closer to F0, that is, F0 is 0.6838 and it gives 0.6931, even more, that is a value of this goodness of fit, okay. It is very good, and so, then accordingly, this becomes the F1 value. So, in this way, you can find out, that is, conduct the experiment and then validate the experiment. And then finally, you get your optimum value.



So, the advantages of linear programming are: there is no need for model development, an efficient way of finding the suitable formulation, it is easy to implement the linear programming course, and various software tools are available to work on the linear programming model.

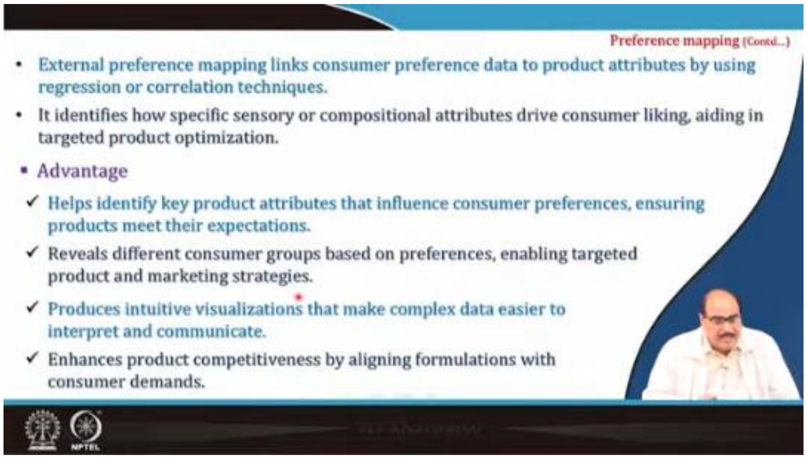
Now, let us talk about consumer-based food formulation. Consumer-based food formulation focuses on creating food products that help understand the preferences of target consumers. This process combines sensory science, consumer research, and mathematical tools to optimize food formulations based on feedback from consumers. Among the mathematical techniques employed, preference mapping and multinomial regression are used to analyze and predict consumer preferences.



So, preference mapping, what is that? You can see here in this figure with the PCA. Preference mapping is a statistical approach used to rate consumer preferences to predict attributes. And it visualizes consumer preferences using statistical methods like PCA, other tools, etc., and aligns them with sensory or other analytical data to guide product optimization, that is, similar types of products are grouped together.

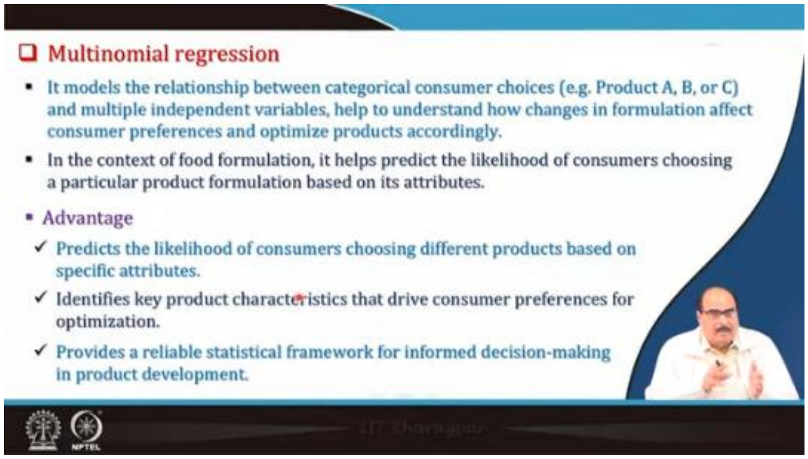
So, preference mapping comes in two main forms: internal preference mapping or external preference mapping. Internal preference mapping analyzes only consumer preferences data using techniques like principal component analysis to identify patterns and groupings in preferences.

Suppose there are various numbers of formulations where you find that, the flavor of the products had to, and there are 5-6 types of different juices, and you have identified that you like a particular flavor. Then, you conduct an experiment and, on the basis of using PCA tools and etcetera, it will show that the products with similar flavors will be grouped together and those which have bad flavor, foul flavor, or other different flavors will be grouped together. So, that is principal component preference mapping. So, it helps visualize clusters of consumers with similar liking trends without linking them to product attributes.



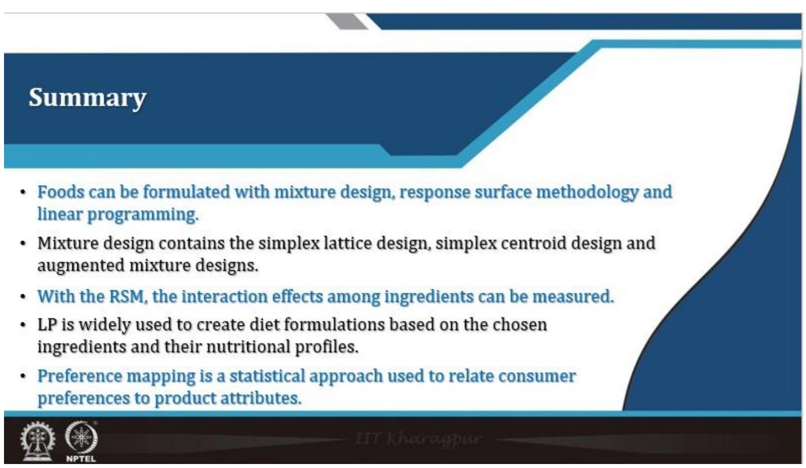
External preference mapping links consumer preference data to product attributes by using regression or correlation techniques. It identifies how specific sensory or computational attributes drive consumer liking, aiding in the target product optimization.

The advantages of this are that it helps identify key product attributes that influence consumer preferences, ensuring the product meets their expectations. It reveals different consumer groups based on preferences, enabling targeted product and marketing strategies. It produces intuitive visualizations that make complex data easier to interpret and communicate. It enhances product competitiveness by aligning formulations with consumer demands.



The multinomial regression models the relationship between categorical consumer choices. For example, product A, B, or C, and multiple independent variables help to understand how changes in formulation affect consumer preferences and optimize the product accordingly. In the context of food formulation, it helps predict the likelihood of consumers choosing a particular product formulation based on its attributes.

The advantage of multinomial regression analysis is that it predicts the likelihood of consumers choosing different products based on specific attributes. It identifies key product characteristics that drive consumer preferences for optimization, and it provides a reliable statistical framework for informed decision-making in product development.



Finally, I summarize this lecture by saying that foods can be formulated using appropriate designs like mixture design, response surface methodology, and linear programming, etcetera. Mixture design includes the simplex lattice design, simplex centroid design, and augmented mixture design with RSM. The interaction effects among the ingredients can be easily measured. Linear programming is widely used to create diet formulations based on chosen ingredients and their nutritional profiles. Preference mapping is a statistical approach used to relate consumer preferences to product attributes.

These were the references that were used in preparing this lecture.

Thank you very much for your patience here. Thanks.