Analysis of variance (ANOVA)

A key statistical test in research fields including biology, economics and psychology, analysis of variance (ANOVA) is very useful for analyzing datasets. It allows comparisons to be made between three or more groups of data. Here, we summarize the key differences between these two tests, including the assumptions and hypotheses that must be made about each type of test.

There are two types of ANOVA that are commonly used, the one-way ANOVA and the two-way ANOVA. This article will explore this important statistical test and the difference between these two types of ANOVA.

What is a one-way ANOVA?

A one-way ANOVA is a type of statistical test that compares the variance in the group means within a sample whilst considering only one independent variable or factor. It is a hypothesis-based test, meaning that it aims to evaluate multiple mutually exclusive theories about our data.

Before we can generate a hypothesis, we need to have a question about our data that we want an answer to. For example, adventurous researchers studying a population of walruses might ask "Do our walruses weigh more in early or late mating season?"

Here, the independent variable or factor (the two terms mean the same thing) is "month of mating season". In an ANOVA, our independent variables are organized in categorical groups. For example, if the researchers looked at walrus weight in December, January, February and March, there would be four months analyzed, and therefore four groups to the analysis.

A one-way ANOVA compares three or more than three categorical groups to establish whether there is a difference between them. Within each group there should be three or more observations (here, this means walruses), and the means of the samples are compared.

What are the hypotheses of a one-way ANOVA?

In a one-way ANOVA there are two possible hypotheses.

- ❖ **The null hypothesis (H0)** is that there is no difference between the groups and equality between means (walruses weigh the same in different months).
- ❖ **The alternative hypothesis (H1)** is that there is a difference between the means and groups (walruses have different weights in different months) .

What are the assumptions and limitations of a one-way ANOVA?

1. **Normality –** that each sample is taken from a normally distributed population
2. **Sample independence** – that each sample has been drawn independently of the other samples
3. **Variance equality –** that the variance of data in the different groups should be the same
4. **Your dependent variable** – here, "weight", should be continuous – that is, measured on a scale which can be subdivided using increments (i.e. grams, milligrams)

What is a two-way ANOVA?

A two-way ANOVA is, like a one-way ANOVA, a hypothesis-based test. However, in the two-way ANOVA each sample is defined in two ways, and resultantly put into two categorical

groups. Thinking again of our walruses, researchers might use a two-way ANOVA if their question is:

"Are walruses heavier in early or late mating season and does that depend on the sex of the walrus?" In this example, both "month in mating season" and "sex of walrus" are factors – meaning in total, there are two factors. Once again, each factor's number of groups must be considered – for "sex" there will only two groups "male" and "female".

The two-way ANOVA therefore examines the effect of two factors (month and sex) on a dependent variable – in this case weight, and also examines whether the two factors affect each other to influence the continuous variable.

What are the assumptions and limitations of a two-way ANOVA?

Your dependent variable – here, "weight", should be continuous – that is, measured on a scale which can be subdivided using increments (i.e. grams, milligrams)

Your two independent variables – here, "month" and "sex", should be in categorical, independent groups.

1. **Sample independence** – that each sample has been drawn independently of the other samples
2. **Variance Equality** – That the variance of data in the different groups should be the same
3. **Normality** – That each sample is taken from a normally distributed population

What are the hypotheses of a two-way ANOVA?

Because the two-way ANOVA considers the effect of two categorical factors, and the effect of the categorical factors on each other, there are three pairs of null or alternative hypotheses for the two-way ANOVA. Here, we present them for our walrus experiment, where month of mating season and sexare the two independent variables.

H0: The means of all month groups are equal
H1: The mean of at least one month group is different

H0: The means of the sex groups are equal
H1: The means of the sex groups are different

H0: There is no interaction between the month and gender

H1: There is interaction between the month and gender

Differences between one-way and two-way ANOVA

The key differences between one-way and two-way ANOVA are summarized clearly below.

❖ **A one-way ANOVA is primarily designed to enable the equality testing between three or more means. A two-way ANOVA is designed to assess the interrelationship of two independent variables on a dependent variable.**

❖ **A one-way ANOVA only involves one factor or independent variable, whereas there are two independent variables in a two-way ANOVA.**

❖ **In a one-way ANOVA, the one factor or independent variable analyzed has three or more categorical groups. A two-way ANOVA instead compares multiple groups of two factors.**

❖ **One-way ANOVA needs to satisfy only two principles of design of experiments, i.e. replication and randomization. As opposed to two-way ANOVA, which meets all three principles of design of experiments which are replication, randomization and local control.**

Linear Programming problem

- Linear Programming is a problem solving approach that has been developed to help managers to make decisions.
- Linear Programming is a mathematical technique for determining the optimum allocation of resources and obtaining a particular objective when there are alternative uses of the resources, money, manpower, material, machine and other facilities

THE NATURE OF LINEAR PROGRAMMING PROBLEM

Two of the most common are:

1. The product-mix problem

2. The blending Problem

In the product- mix problem there are two or more products also called candidates or activities competing for limited resources. The problem is to find out which products to
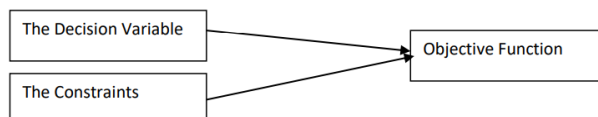
include in the production plan and in what quantities these should be produced or sold in order to maximize profit, market share or sales revenue.

The blending problem involves the determination of the best blend of available ingredients to form a certain quantity of a product under strict specifications. The best blend means the least cost blend of the required inputs.

FORMULATION OF THE LINEAR PROGRAMMING MODEL

Three components are:

1. The decision variable

2. The environment (uncontrollable) parameters

3. The result (dependent) variable

Linear Programming Model is composed of the same components

TERMINOLOGY USED IN LINEAR PROGRAMMING PROBLEM

1. **Components of LP Problem**: Every LPP is composed of a. Decision Variable, b. Objective Function, c. Constraints.

2. **Optimization:** Linear Programming attempts to either maximize or minimize the values of the objective function.

3. **Profit of Cost Coefficient**: The coefficient of the variable in the objective function expresses the rate at which the value of the objective function increases or decreases by including in the solution one unit of each of the decision variables.

4. **Constraints:** The maximization (or minimisation) is performed subject to a set of constraints. Therefore LP can be defined as a constrained optimization problem. They reflect the limitations of the resources.

5. I**nput-Output coefficients:** The coefficients of constraint variables are called the InputOutput Coefficients. They indicate the rate at which a given resource is unitized or depleted. They appear on the left side of the constraints.

6. **Capacities:** The capacities or availability of the various resources are given on the right hand side of the constraints.

ASSUMPTIONS

Linear programs are constrained optimization models that satisfy three requirements.

1. The decision variables must be continuous; they can take on any value within some restricted range.

2. The objective function must be a linear function.

3. The left-hand sides of the constraints must be linear functions

Proportionality. With linear programs, we assume that the contribution of individual variables in the objective function and constraints is proportional to their value. That is, if we double the value of a variable, we double the contribution of that variable to the objective function and each constraint in where the variable appears. The contribution per unit of the variable is constant. For example, suppose the variable xj is the number of units of product j produced and cj is the cost per unit to produce product j. If doubling the amount of product j produced doubles its cost, the per unit cost is constant and the proportionality assumption is satisfied.

2. Additivity. Additivity means that the total value of the objective function and each constraint function is obtained by adding up the individual contributions from each variable.

3. Divisibility. The decision variables are allowed to take on any real numerical values within some range specified by the constraints. That is, the variables are not restricted to integer values. When fractional values do not make a sensible solution, such as the number of flights an airline should have each day between two cities, the problem should be formulated and solved as an integer program.

4. Certainty. We assume that the parameter values in the model are known with certainty or are at least treated that way. The optimal solution obtained is optimal for the specific problem formulated. If the parameter values are wrong, then the resulting solution is of little value

Formulation of Linear Programming Model

Step 1

From the study of the situation, find the key-decision to be made. In the given situation key decision is to decide the extent of products 1, 2 and 3, as the extents are permitted to vary.

Step 2

Assume symbols for variable quantities noticed in step 1. Let the extents (amounts) of products 1, 2 and 3 manufactured daily be x1, x2 and x3 units respectively.

Step 3

Express the feasible alternatives mathematically in terms of variables. Feasible alternatives are those which are physically, economically and financially possible. In the given situation feasible alternatives are sets of values of x1, x2 and x3 units respectively.where x1, x2 and x3 ≥ 0.

since negative production has no meaning and is not feasible.

Step 4

Mention the objective function quantitatively and express it as a linear function of variables.

In the present situation, the objective is to maximize the profit.

i.e., $Z = 4x_1 + 3x_2 + 6x_3$

Step 5

Put into words the influencing factors or constraints. These occur generally because of constraints on availability (resources) or requirements (demands). Express these constraints

also as linear equations/inequalities in terms of variables.

Here, constraints are on the machine capacities and can be mathematically expressed as

$2x_1 + 3x_2 + 2x_3 \leq 440$

$4x_1 + 0x_2 + 3x_3 \leq 470$

$2x_1 + 5x_2 + 0x_3 \leq 430$

MERITS OF LPP

1. Helps management to make efficient use of resources.

2. Provides quality in decision making.

3. Excellent tools for adjusting to meet changing demands.

4. Fast determination of the solution if a computer is used.

5. Provides a natural sensitivity analysis.

6. Finds solutions to problems with a very large or infinite number of possible solutions.

DEMERITS OF LPP

1. Existence of nonlinear equation: The primary requirements of Linear Programming is The objective function and constraint function should be linear. Practically linear relationships do not exist in all cases.

2. Interaction between variables: LP fails in a situation where non-linearity in the equation emerges due to joint interaction between some of the activities like total effectiveness.

3. Fractional Value: In LPP fractional values are permitted for the decision variable.

4. Knowledge of Coefficients of the equation: It may not be possible to state all coefficients in the objective function and constraints with certainty

LINEAR PROGRAMMING SOLUTION – GRAPHICAL METHOD

There are two methods available to find optimal solutions to a Linear Programming Problem.

One is the graphical method and the other is the simplex method.

Graphical method can be used only for a two variables problem i.e. a problem which involves two decision variables. The two axes of the graph (X & Y axis) represent the two decision variables X1 & X2.

METHODOLOGY OF GRAPHICAL METHOD

Step 1: Formulation of LPP (Linear Programming Problem)

Use the given data to formulate the LPP.

Maximization

Step 2: Determination of each axis

Horizontal (X) axis: Product A (X1)

Vertical (Y) axis: Product B (X2)

Step 3: Finding coordinates of constraint lines to represent constraint lines on the graph.

The constraints are presently in the form of inequality ($\leq$). We should convert them into equality to obtain coordinates.

Step 4:

Representing constraint lines on graph

To mark the points on the graph, we need to select the appropriate scale. Which scale to take will depend on the maximum value of X1 & X2 from coordinates.

Step 5: Identification of Feasible Region

The feasible region is the region bounded by constraint lines. All points inside the feasible region or on the boundary of the feasible region or at the corner of the feasible region satisfy all constraints.

Both the constraints are 'less than or equal to' ($\leq$) type. Hence, the feasible region should be inside both constraint lines.
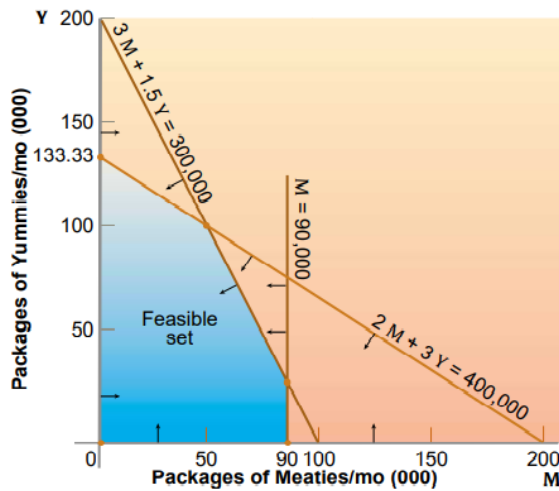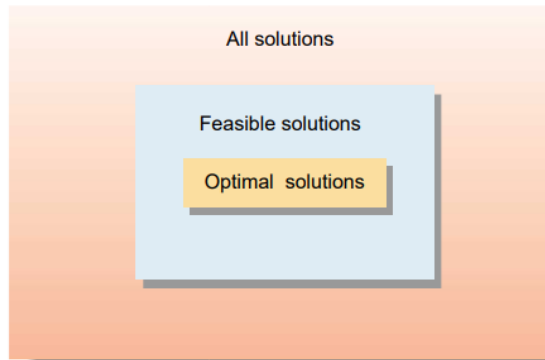
Step 6: Finding the optimal Solution

The optimal solution always lies at one of the vertices or corners of the feasible region.To find optimal solution:

We use corner point method. We find coordinates (X1, X2 Values) for each vertex or corner point. From this we find the 'Z' value for each corner point.

THE GRAPHICAL SOLUTION APPROACH

Maximize $\quad z = 0.65M + 0.45Y$

Subject to $\begin{aligned} 2M + \quad 3Y &\le 400{,}000 \\ 3M + 1.5Y &\le 300{,}000 \\ M \quad &\le 90{,}000 \\ M, Y \quad &\ge 0 \end{aligned}$

All solutions

Feasible solutions

Optimal solutions

We begin the solution process by finding the feasible set. The geometric representation of a linear equality is the set of points that lie on and to one side of the line obtained by replacing the inequality sign with an equality sign.

The constraint M ≥ 0 restricts us to the points on or to the right of the vertical axis (the line M 0). The constraint Y ≥ 0 restricts us to the points on or above the horizontal axis. Next, we draw the constraint 2M + 3Y =400,000.

To find the points that satisfy this inequality, we construct the line 2M   3Y 400,000 by finding two points that lie on the line and then constructing a line

through these points. The easiest points to find on the line are the ones that lie on the two axes.

First, set M = 0 and solve for Y. This yields the point (M = 0, Y= 133,333.22). We then set Y= 0 and solve for M.

This yields the point (M = 200,000 and Y= 0). This line is plotted on Figure B-2.

We now determine on which side of the line the points satisfy the constraint. If one point satisfies the constraint, then all points on the same side of the line satisfy the constraint. If one point does not satisfy the constraint, then no point on that side of the line satisfies the constraint, but all the points on the opposite side of the line do satisfy the constraint. It makes sense to select a simple point with which to work, such as (M = 0, Y= 0). This point satisfies the constraint 2M + 3Y ≤ 400,000. Therefore,all points to the lower left do also. The points to the upper right of the line represent product mixes that require more than 400,000 pounds of cereal each month and can be eliminated from consideration.

We do the same thing for the meat constraint: 3M +1.5Y ≤ 300,000. We find two points on the line 3M + 1.5Y = 300,000. We first set M 0 and solve for Y, and then set Y 0 and solve for M, yielding the points (M = 0, Y= 200,000) and (M 100,000, Y= 0).

 Checking a point on one side of the line shows that the points on or to the lower left of the line are the ones that satisfy the constraint.

The final constraint, M ≤90,000, is satisfied by the points that lie on or to the left of the vertical line M = 90,000.

The feasible set is the set of points in the five sided shaded area in Figure B-2. The feasible set for a linear program will always have a shape like the one in this problem, with edges that are straight lines and corners where the edges meet. The corners of the feasible set are called extreme points. Note that each extreme point is formed by the intersection of two or more constraints.

 The fundamental theorem of linear programming is: If a finite optimal solution exists, then at least one extreme point is optimal. To find the exact coordinate values for the optimum from the graph:

- We identify the constraints that intersect to form the extreme point.
- . We solve simultaneously the equations corresponding to the constraints to
- find the point that lies on both lines (the extreme point).In this example, the optimal extreme point is formed by the intersection of the

lines 2M + 3Y = 400,000 and 3M+ 1.5Y =300,000.

To solve equations simultaneously, we use the following property. We can either

(1) multiply any constraint by a nonzero constant or (2) add or subtract a multiple of any equation to or from any other equation, without changing the set of solutions that the equations have simultaneously.

Therefore, to solve simultaneously the equations

$2M + 3Y = 400,000$

$3M + 1.5Y = 300,000$

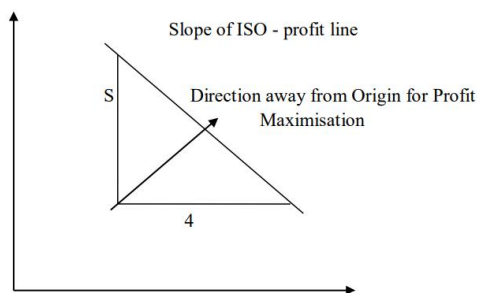we can subtract two times the second equation from the first equation, leaving

$-4M = 200,000$

$3M + 1.5Y = 300,000$

The first equation is easily solved: M 50,000. Substituting this value into the second equation gives 3(50,000) 1.5Y 300,000, or Y 100,000. This method can be used for any set of linear equations as long as the number of variables and equations is equal.

ISO Profit line:

ISO profit line is the line which passes through the points of optimal solution (Maximum Profit). The slope of the iso-profit line depends on the objective function.

THE SIMPLEX ALGORITHM

In 1949, George Dantzig developed an efficient procedure for solving linear programs called the simplex method or simplex algorithm. This is the most widely used method in instructional and commercial computer packages. A method developed by Narendra Karmarkar in 1984 is gaining popularity, but since it requires more sophisticated mathematics, it is not presented here.

The fundamental theorem of linear programming reduces to a finite value the number of feasible solutions that need to be evaluated. One solution strategy might be to identify the coordinates of every extreme point and then evaluate the objective

Nonnegative Right-Hand Sides.

Suppose a constraint has a negative right-hand side, such as

The right-hand side can be made nonnegative by multiplying through the entire constraint by 1, so the constraint becomes

$$3x_1 + 4x_2 - 2x_3 \le -6$$

The right-hand side can be made nonnegative by multiplying through the entire traint by -1, so the constraint becomes

$$-3x_1 - 4x_2 + 2x_3 \ge 6$$

Note that the direction of the inequality reverses when the constraint is multi- by a negative constant.

Note that the direction of the inequality reverses when the constraint is multiplied by a negative constant.

All Constraints Must Be Equalities. Any inequality constraint can be rewritten as an equivalent equality by introducing a nonnegative variable called a slack variable. Let's look at constraints first.

Let's look at $\le$ constraints first.

$$3x_1 + 4x_2 \le 12$$

If we define a new variable, s_1, as $s_1 = 12 - (3x_1 + 4x_2)$, then this constraint can be written as

$$3x_1 + 4x_2 + s_1 = 12$$

For any solution, the variable s_1 measures the "slack" that exists in the constraint. That is, s_1 measures by how much a solution satisfies a constraint.

We use a similar approach to convert $\ge$ constraints into equalities. For example, consider

$$2x_1 + 4x_2 \ge 10$$

If we define a new variable, s_2, so that $s_2 = (2x_1 + 4x_2) - 10$, then we can rewrite the constraint as

$$2x_1 + 4x_2 - s_2 = 10$$
$$s_2 \ge 0$$

In summary, for $\le$ constraints, we add a nonnegative slack variable to the left-hand side of the constraint, and for $\ge$ constraints we subtract a nonnegative slack variable from the left-hand side of the constraint.

For any solution, the variable s1 measures the "slack" that exists in the constraint. That is, s1 measures by how much a solution satisfies a constraint.

We use a similar approach to convert ≥ constraints into equalities. For example,

Consider If we define a new variable, s2, so that s2 (2x1 4x2) 10, then we can rewrite the constraint as

In summary, for constraints, we add a nonnegative slack variable to the left hand side of the constraint, and for constraints we subtract a nonnegative slack variable from the left-hand side of the constraint.

SUPPLEMENT B LINEAR PROGRAMMING

Canonical form is needed to use the simplex method.

Slack variables have physical meaning. For example, slack variables may measure how much of an available resource is not used for a given solution or how much of a nutrient an animal receives above some minimum required amount. The slack variables in each constraint measure something different. Therefore, each one should have a different designation. A standard convention uses si for a slack variable that has been incorporated into constraint i.

All Variables Must Be Nonnegative. Most linear programming computer packages automatically convert right-hand-side constants into nonnegative values and change constraints into equalities. They assume all variables are nonnegative. If they are not,the problem must be rewritten so that all variables are nonnegative.

Each Constraint Must Have a Unique Variable with a +1 Coefficient. Using the Healthy Pet Food problem, we perform the needed steps, giving us the following:

$$\text{Maximize} \qquad z = 0.65M + 0.45Y$$

$$
\begin{aligned}
\text{Subject to} \qquad & 2M + 3Y + s_1 && = 400{,}000 \\
& 3M + 1.5Y \;\;\; + s_2 && = 300{,}000 \\
& M && + s_3 = 90{,}000
\end{aligned}
$$

All variables ≥ 0

All variables 0

We check to see whether each constraint has a variable that appears only in that constraint and has a coefficient of 1. In the first constraint, s1 satisfies this requirement; in the second constraint, s2 satisfies it; and in the third constraint, s3 satisfies it.

This problem is in canonical form and ready to be solved.

function at each. The one that produces the best objective function value is the optimum. In practice, this approach is not efficient because the number of extreme points can be very large for real problems with hundreds or thousands of variables and constraints.

The simplex algorithm begins by identifying an initial extreme point of the feasible set. The algorithm then looks along each edge intersecting at the extreme point and computes the net effect on the objective function if we were to move along the edge. If the objective function value does not improve by moving along at least one of these edges, it can be proved that the extreme point is optimal. If movement along one or more of the edges improves the objective function value, we move along one of these edges until we reach a new extreme point. We repeat the previous steps: checking along each edge intersecting at the extreme point, and then either stopping or sliding along another edge that improves the objective function value.

This algorithm has many desirable features in practice.

1. It only moves from one extreme point to a better or equally good extreme point, thereby skipping large numbers of suboptimal extreme points without explicitly identifying them. Thus it usually only has to check a small subset of the extreme points to find an optimum.

2. When it finds an optimum, it identifies this fact and stops.

3. The algorithm detects whether the problem is infeasible, is unbounded, or has multiple optima.

4. The algorithm uses very simple mathematics that are easy to implement on a computer

The Big-M Method

In the Big-M method, we make the objective function coefficients for the artificial variables either infinity for maximization problems or infinity for minimization problems. The name comes from the fact that the letter M represents some large number that replaces infinity. This method encourages the algorithm to drive the artificial variables to zero by kicking them out of the basis. Rather than use infinity, in our example M 1000. The first tableau is in Table B-8.

Table B-8 First Tableau: Big-*M* Method

c_B	Basic Variable	A 20	B 6	C 9	s_1 0	s_2 0	A_c −1000	A_B −1000	b	Ratio	
0	s_3	4	3	1	1	0	0	0	24	(24/1) = 24	
−1,000	A_2	2	4	12	0	−1	1	0	30	(30/12) = 2.5	←Pivot row
−1,000	A_3	2	3	1	0	0	0	1	10	(10/1) = 10	
	z_j	−4,000	−7,000	−13,000	0	1,000	−1,000	−1,000	−4,000		
	$c_j - z_j$	4,020	7,006	13,009	0	−1,000	0	0			

↑
Pivot column

Variable C has the most positive cj zj , so C enters the basis. The second row

has the smallest ratio, so the basis variable in that row, A2, will leave the basis. Performing the pivot yields Table B-9.

Note that the A2 column has been deleted in TableB-9. As soon as an artificial variable leaves the basis, it is dropped from the problem.

(Do not confuse this with other types of variables.) The resulting solution is not optimalso variable B enters the basis and replaces A3, giving Table B-10. Variable A then enters the basis and B leaves, giving Table B-11.

The optimal solution is A 45/11 4.091, B 0, C 20/11 1.818, s1

64/11 5.818 , s2 0, and z 1080/11 98.181. Although the artificial variables

usually leave the basis quickly, they are not always the first variables to leave as they were in this example.

Table B-9 Second Tableau: Big-*M* Method

c_B	Basic Variable	A 20	B 6	C 9	s_1 0	s_2 0	A_c −1000	A_B −1000	b	Ratio	
0	s_3	23/6	8/3	0	1	1/12	—	0	21.5	(21.5/(8/3)) = 8.06	
9	C	1/6	1/3	1	0	−1/12	—	0	2.5	(2.5/(1/3)) = 7.5	
−1000	A_3	11/6	8/3	0	0	1/12	—	1	7.5	(7.5/(8/3)) = 2.81	←Pivot row
	z_j	−1831.8	−2663.7	9	0	−84.1	—	−1000	−7477.5		
	$c_j - z_j$	1851.8	2669.7	0	0	84.1	—	0			

↑
Pivot column

Table B-10 Third Tableau: Big-*M* Method

c_B	Basic Variable	A 20	B 6	C 9	s_1 0	s_2 0	A_c −1000	A_B −1000	b	Ratio	
0	s_3	2	0	0	1	0	—	—	14	(14/2) = 7	
9	C	−1/16	0	1	0	−3/32	—	—	25/16	no ratio	
6	B	11/16	1	0	0	1/32	—	—	45/16	(45/11) = 4.09	←Pivot row
	z_j	57/16	6	9	0	−21/32	—	—	495/16		
	$c_j - z_j$	263/16	0	0	0	21/32	—	—			

↑
Pivot column

Pivot column

Table B-11 Optimal Tableau: Big-*M* Method

c_B	Basic Variable	A 20	B 6	C 9	s_1 0	s_2 0	A_2 −1000	A_3 −1000	b
0	s_3	0	−32/11	0	1	−1/11	—	—	64/11
9	C	0	1/11	1	0	−1/11	—	—	20/11
20	A	1	16/11	0	0	1/22	—	—	45/11
	z_j	20	329/11	9	0	1/11	—	—	1080/11
	$c_j − z_j$	0	−263/11	0	0	−1/11	—	—	

Duality

It is a unifying theory that develops the relationship between a given linear program and another related linear program stated in terms of variables with this shadow-price interpretation.

This unified theory is important

1. Because it allows fully understanding the shadow-price interpretation of the optimal simplex multipliers, which can prove very useful in understanding the implications of a particular linear-programming model.

2. Because it is often possible to solve the related linear program with the shadow prices as the variables in place of, or in conjunction with, the original linear program, thereby taking advantage of some computational efficiencies.

Shadow prices

optimal simplex multipliers

Marginal worth of an additional unit of resource

Opportunity costs of resource allocation when pricing out a new activity

4.2 DEFINITION OF THE DUAL PROBLEM

The duality principles we have illustrated in the previous sections can be stated formally in general terms. Let the primal problem be:

Primal

$$\text{Maximize } z = \sum_{j=1}^{n} c_j x_j,$$

subject to:

$$\sum_{j=1}^{n} a_{ij} x_j \le b_i \qquad (i = 1, 2, \ldots, m), \tag{3}$$

$$x_j \ge 0 \qquad (j = 1, 2, \ldots, n).$$

Associated with this primal problem there is a corresponding dual problem given by:

Dual

$$\text{Minimize } v = \sum_{i=1}^{m} b_i y_i,$$

subject to:

$$\sum_{i=1}^{m} a_{ij} y_i \ge c_j \qquad (j = 1, 2, \ldots, n), \tag{4}$$

$$y_i \ge 0 \qquad (i = 1, 2, \ldots, m).$$

Generalization

$$\text{Maximize } z = \sum_{j=1}^{n} c_j x_j, \quad \text{Minimize } v = \sum_{i=1}^{m'} b_i y_i + \sum_{i=m'+1}^{m''} b_i y_i' + \sum_{i=m''+1}^{m} b_i y_i'',$$

$$\sum_{j=1}^{n} a_{ij} x_j \leq b_i$$

$$\sum_{j=1}^{n} a_{ij} x_j \geq b_i$$

$$\sum_{j=1}^{n} a_{ij} x_j = b_i$$

$$\sum_{i=1}^{m'} a_{ij} y_i + \sum_{i=m'+1}^{m''} a_{ij} y_i' + \sum_{i=m''+1}^{m} a_{ij} y_i'' \geq c_j$$

$$y_i \qquad \geq 0,$$

$$-y_i' \qquad \geq 0,$$

Generalization

Primal (Maximize)	Dual (Minimize)
ith constraint $\leq$	ith variable ≥ 0
ith constraint $\geq$	ith variable ≤ 0
ith constraint $=$	ith variable unrestricted
jth variable ≥ 0	jth constraint $\geq$
jth variable ≤ 0	jth constraint $\leq$
jth variable unrestricted	jth constraint $=$

$$\text{Maximize } z = \sum_{j=1}^{n} c_j x_j,$$

subject to:

$$\sum_{j=1}^{n} a_{ij} x_j \leq b_i \qquad (i = 1, 2, \ldots, m), \tag{13}$$

$$x_j \geq 0 \qquad (j = 1, 2, \ldots, n).$$

Dual

$$\text{Minimize } v = \sum_{i=1}^{m} b_i y_i,$$

subject to:

$$\sum_{i=1}^{m} a_{ij} y_i \geq c_j \qquad (j = 1, 2, \ldots, n), \tag{14}$$

$$y_i \geq 0 \qquad (i = 1, 2, \ldots, m).$$

The first property is referred to as ''weak duality'' and provides a bound on the optimal value of the objective function of either the primal or the dual. Simply stated, the value of the objective function for any feasible solution to the primal maximization problem is bounded from above by the value of the objective function for any feasible solution to its dual. Similarly, the value of the objective function for its dual is bounded from below by the value of the objective function of the primal. Pictorially, we might represent the situation as follows:

Dual feasible | v decreasing

Primal feasible | z increasing

The Dual Simplex Method

• Solving the primal problem,

moving through solutions (simplex tableaus)
that are dual feasible but primal unfeasible.(1)

– **Primal feasible:**

$$\overline{b}_i \geq 0$$
$$\overline{c}_j \leq 0$$

– **Dual feasible:**

• An optimal solution is a solution that is both
primal and dual feasible.

Except for the selection of the variable to leave and
enter the basis.

• At each iteration of the dual simplex method, we require
That:

$$\overline{c}_j = c_j - \sum_{i=1}^{m} y_i a_{ij} \leq 0;$$

$y_i \geq 0$ for $i = 1, 2, \ldots, m,$

and since these variables are a dual feasible solution. Further, at each iteration of the dual simplex method, the most negative **b$_i$** is chosen to determine the pivot row, corresponding to choosing the most positive to determine the pivot column in the primal simplex Method.

The following observations were made while forming the dual linear programming problem:

➢ The primal or original linear programming problem is of the maximization type while the dual problem is of minimization type.

➢ The constraint values 100 and 150 of the primal problem have become the coefficient of dual variables y1 and y2 in the objective function of a dual problem and while the coefficient of the variables in the objective function of a primal problem has become the constraint value in the dual problem.

➢ The first column in the constraint inequality of the primal problem has become the first row in a dual problem and similarly the second column of constraint has become the second row in the dual problem.

➢ The directions of inequalities have also changed, i.e. in the dual problem, the sign is the reverse of a primal problem. Such that in the primal problem, the inequality sign was "≤" but in the dual problem, the sign of inequality becomes "≥".