CHAPTER
TWO

# Mathematics for Microeconomics

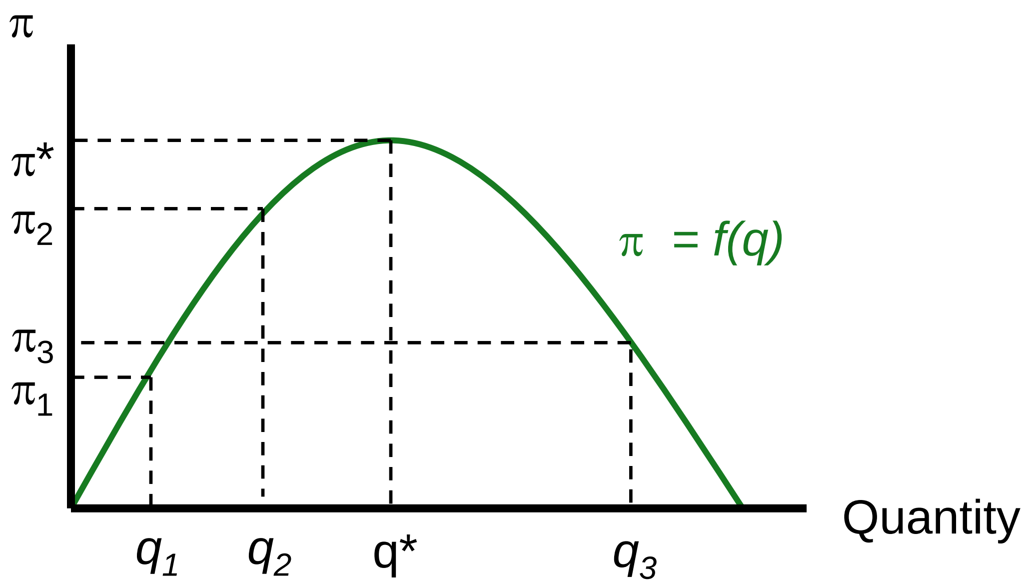
PowerPoint Slides prepared by:
Andreea CHIRITESCU
Eastern Illinois University

# Maximization of a Function of One Variable

- Economic theories assume that
  - An economic agent is seeking to find the optimal value of some function
    - Consumers seek to maximize utility
    - Firms seek to maximize profit
- Simple example, $\pi = f(q)$
  - Manager of a firm wants to maximize profits, $\pi$
    - The profits ($\pi$) received depend only on the quantity ($q$) of the good sold

FIGURE 2.1

# Hypothetical Relationship between Quantity Produced and Profits




If a manager wishes to produce the level of output that maximizes profits, then $q^*$ should be produced. Notice that at $q^*$, $d\pi/dq = 0$.

# Maximization of a Function of One Variable

- Vary $q$ to see where maximum profit occurs
  - An increase from $q_1$ to $q_2$ leads to a rise in $\pi$

$$\frac{\Delta \pi}{\Delta q} > 0$$

# Maximization of a Function of One Variable

- If output is increased beyond $q^*$, profit will decline
  - An increase from $q^*$ to $q_3$ leads to a drop in $\pi$

$$\frac{\Delta \pi}{\Delta q} < 0$$

# Maximization of a Function of One Variable

- Derivatives
  - The derivative of $\pi = f(q)$ is the limit of $\Delta\pi/\Delta q$ for very small changes in $q$
  - Is the slope of the curve
  - The value depends on the value of $q_1$

$$\frac{d\pi}{dq} = \frac{df}{dq} = \lim_{h \to 0} \frac{f(q_1 + h) - f(q_1)}{h}$$

# Maximization of a Function of One Variable

- Value of a derivative at a point
  - The evaluation of the derivative at the point q = $q_1$ can be denoted

$$\left.\frac{d\pi}{dq}\right|_{q=q_1}$$

- In our previous example,

$$\left.\frac{d\pi}{dq}\right|_{q=q_1} > 0$$

$$\left.\frac{d\pi}{dq}\right|_{q=q_3} < 0$$

$$\left.\frac{d\pi}{dq}\right|_{q=q^*} = 0$$

# Maximization of a Function of One Variable

- First-order condition for a maximum
  - For a function of one variable to attain its maximum value at some point, the derivative at that point must be zero

$$\left.\frac{df}{dq}\right|_{q=q^*} = 0$$

# Maximization of a Function of One Variable

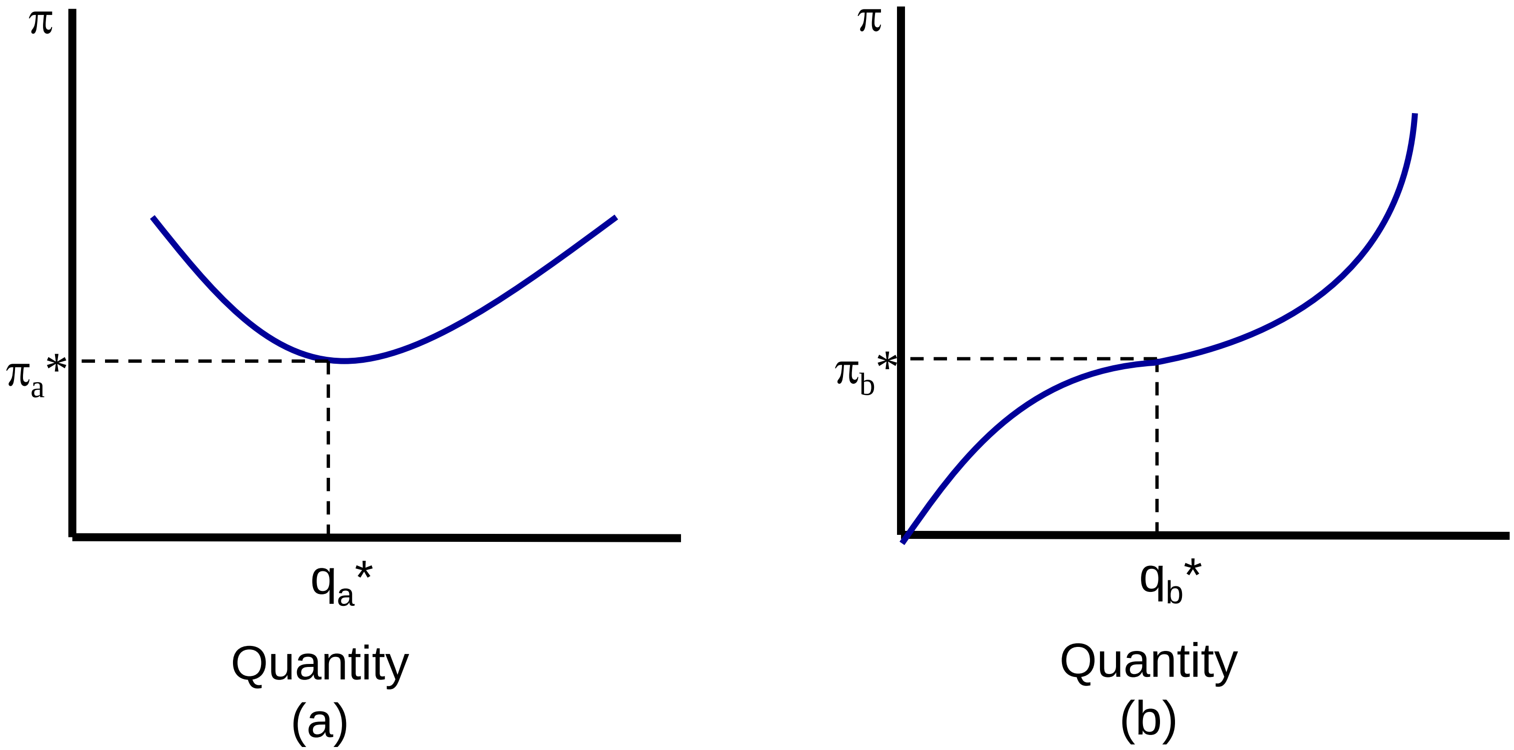
- The first order condition ($d\pi/dq$)
  - Is a necessary condition for a maximum
  - But it is not a sufficient condition
- The second order condition
  - In order for $q^*$ to be the optimum,

$$\frac{d\pi}{dq} > 0 \text{ for } q < q^* \quad \text{and} \quad \frac{d\pi}{dq} < 0 \text{ for } q > q^*$$

  - At $q^*$, $d\pi/dq$ must be decreasing
    - The derivative of $d\pi/dq$ must be negative at $q^*$

FIGURE 2.2

# Two Profit Functions That Give Misleading Results If the First Derivative Rule Is Applied Uncritically




In (a), the application of the first derivative rule would result in point $q_a$* being chosen. This point is in fact a point of minimum profits. Similarly, in (b), output level $q_b$* would be recommended by the first derivative rule, but this point is inferior to all outputs greater than $q_b$* . This demonstrates graphically that finding a point at which the derivative is equal to 0 is a necessary, but not a sufficient, condition for a function to attain its maximum value.

# Maximization of a Function of One Variable

- Second derivative
  - The derivative of a derivative
  - Can be denoted by:

$$\frac{d^2\pi}{dq^2} \text{ or } \frac{d^2 f}{dq^2} \text{ or } f''(q)$$

# Maximization of a Function of One Variable

- The second order condition
  - To represent a (local) maximum is:

$$\left.\frac{d^2\pi}{dq^2}\right|_{q=q^*} = f''(q)\Big|_{q=q^*} < 0$$

# Rules for Finding Derivatives

1. If $a$ is a constant, then $\dfrac{da}{dx} = 0$

2. If $a$ is a constant, then $\dfrac{d[af(x)]}{dx} = af'(x)$

3. If $a$ is a constant, then $\dfrac{dx^a}{dx} = ax^{a-1}$

4. $\dfrac{d \ln x}{dx} = \dfrac{1}{x}$

5. $\dfrac{da^x}{dx} = a^x \ln a$ for any constant $a$

   - special case: $\dfrac{de^x}{dx} = e^x$

# Rules for Finding Derivatives

- Suppose that $f(x)$ and $g(x)$ are two functions of $x$ and $f'(x)$ and $g'(x)$ exist
- Then

$$6.\ \frac{d[f(x)+g(x)]}{dx} = f'(x)+g'(x)$$

$$7.\ \frac{d[f(x)\cdot g(x)]}{dx} = f(x)g'(x)+f'(x)g(x)$$

$$8.\ \frac{d\left(\dfrac{f(x)}{g(x)}\right)}{dx} = \frac{f'(x)g(x)-f(x)g'(x)}{[g(x)]^2} \text{ provided that } g(x)\neq 0$$

# Rules for Finding Derivatives

- If $y = f(x)$ and $x = g(z)$ and if both $f'(x)$ and $g'(x)$ exist, then:

$$9. \quad \frac{dy}{dz} = \frac{dy}{dx} \cdot \frac{dx}{dz} = \frac{df}{dx} \cdot \frac{dg}{dz}$$

  - This is called the chain rule
  - Allows us to study how one variable ($z$) affects another variable ($y$) through its influence on some intermediate variable ($x$)

# Rules for Finding Derivatives

- Some examples of the chain rule include:

$$10.\quad \frac{de^{ax}}{dx} = \frac{de^{ax}}{d(ax)} \cdot \frac{d(ax)}{dx} = e^{ax} \cdot a = ae^{ax}$$

$$11.\quad \frac{d\left[\ln(ax)\right]}{dx} = \frac{d\left[\ln(ax)\right]}{d(ax)} \cdot \frac{d(ax)}{dx} = \frac{1}{ax} \cdot a = \frac{1}{x}$$

$$12.\quad \frac{d[\ln(x^2)]}{dx} = \frac{d[\ln(x^2)]}{d(x^2)} \cdot \frac{d(x^2)}{dx} = \frac{1}{x^2} \cdot 2x = \frac{2}{x}$$

- Suppose that the relationship between profit and output is

$$\pi = 1{,}000q - 5q^2$$

- The first order condition for a maximum is

$$d\pi/dq = 1{,}000 - 10q = 0$$

$$q^* = 100$$

- Since the second derivative is always -10, then $q = 100$ is a global maximum

# Functions of Several Variables

- Most goals of economic agents depend on several variables
  - Trade-offs must be made
- The dependence of one variable ($y$) on a series of other variables ($x_1, x_2, \ldots, x_n$) is denoted by

$$y = f(x_1, x_2, \ldots, x_n)$$

# Functions of Several Variables

- Partial derivatives
  - Partial derivative of *y* with respect to $x_1$:

$$\frac{\partial y}{\partial x_1} \text{ or } \frac{\partial f}{\partial x_1} \text{ or } f_{x_1} \text{ or } f_1$$

    - All of the other *x*'s are held constant

  - A more formal definition is

$$\left.\frac{\partial f}{\partial x_1}\right|_{\bar{x}_2,\ldots,\bar{x}_n} = \lim_{h\to 0} \frac{f(x_1+h,\bar{x}_2,\ldots,\bar{x}_n) - f(x_1,\bar{x}_2,\ldots,\bar{x}_n)}{h}$$

# Calculating Partial Derivatives

1. If $y = f(x_1, x_2) = ax_1^2 + bx_1x_2 + cx_2^2$, then

$$\frac{\partial f}{\partial x_1} = f_1 = 2ax_1 + bx_2 \quad \text{and} \quad \frac{\partial f}{\partial x_2} = f_2 = bx_1 + 2cx_2$$

2. If $y = f(x_1, x_2) = e^{ax_1 + bx_2}$, then

$$\frac{\partial f}{\partial x_1} = f_1 = ae^{ax_1 + bx_2} \quad \text{and} \quad \frac{\partial f}{\partial x_2} = f_2 = be^{ax_1 + bx_2}$$

3. If $y = f(x_1, x_2) = a \ln x_1 + b \ln x_2$, then

$$\frac{\partial f}{\partial x_1} = f_1 = \frac{a}{x_1} \quad \text{and} \quad \frac{\partial f}{\partial x_2} = f_2 = \frac{b}{x_2}$$

# Functions of Several Variables

- Partial derivatives
  - Are the mathematical expression of the *ceteris paribus* assumption
  - Show how changes in one variable affect some outcome when other influences are held constant
- We must be concerned with units of measurement

# Functions of Several Variables

- Elasticity
  - Measures the proportional effect of a change in one variable on another
  - Unit free
  - Of *y* with respect to *x* is

$$e_{y,x} = \frac{\frac{\Delta y}{y}}{\frac{\Delta x}{x}} = \frac{\Delta y}{\Delta x} \cdot \frac{x}{y} = \frac{\partial y}{\partial x} \cdot \frac{x}{y}$$

# EXAMPLE 2.2 Elasticity and Functional Form

- For: $y = a + bx + other\ terms$
  - The elasticity is:

$$e_{y,x} = \frac{\partial y}{\partial x} \cdot \frac{x}{y} = b \cdot \frac{x}{y} = b \cdot \frac{x}{a + bx + \cdots}$$

  - $e_{y,x}$ is not constant
    - It is important to note the point at which the elasticity is to be computed

- For $y = ax^b$
  - The elasticity is a constant:

$$e_{y,x} = \frac{\partial y}{\partial x} \cdot \frac{x}{y} = abx^{b-1} \cdot \frac{x}{ax^b} = b$$

- For $\ln y = \ln a + b \ln x$
  - The elasticity is:

$$e_{y,x} = \frac{\partial y}{\partial x} \cdot \frac{x}{y} = b = \frac{\partial \ln y}{\partial \ln x}$$

- Elasticities can be calculated through logarithmic differentiation

# Functions of Several Variables

- Second-order partial derivatives
  - The partial derivative of a partial derivative

$$\frac{\partial(\partial f / \partial x_i)}{\partial x_j} = \frac{\partial^2 f}{\partial x_j x_i} = f_{ij}$$

# Second-order partial derivatives

$$1.\ y = f(x_1, x_2) = ax_1^2 + bx_1x_2 + cx_2^2,\ \mathit{then}$$

$$f_{11} = 2a;\ f_{12} = b;\ \ f_{21} = b;\ \ f_{22} = 2c$$

$$2.\ y = f(x_1, x_2) = e^{ax_1 + bx_2},\ \mathit{then}$$

$$f_{11} = a^2 e^{ax_1 + bx_2};\ f_{12} = abe^{ax_1 + bx_2};$$

$$f_{21} = abe^{ax_1 + bx_2};\ f_{22} = b^2 e^{ax_1 + bx_2}$$

$$3.\ \mathit{If}\ \ y = f(x_1, x_2) = a\ln x_1 + b\ln x_2,\ \mathit{then}$$

$$f_{11} = -ax_1^{-2};\ f_{12} = 0;\ f_{21} = 0;\ \ f_{22} = -bx_2^{-2}$$

# Functions of Several Variables

- Young's theorem
  - Under general conditions
  - The order in which partial differentiation is conducted to evaluate second-order partial derivatives does not matter

$$f_{ij} = f_{ji}$$

# Functions of Several Variables

- Second-order partials
  - Play an important role in many economic theories
  - A variable's own second-order partial, $f_{ii}$
    - Shows how $\partial y/\partial x_i$ changes as the value of $x_i$ increases
    - $f_{ii} < 0$ indicates diminishing marginal effectiveness

# Functions of Several Variables

- The chain rule with many variables
  - $y = f(x_1,x_2,x_3)$
    - Each of these *x*'s is itself a function of a single parameter, *a*
  - $y = f[x_1(a),x_2(a),x_3(a)]$
  - How a change in *a* affects the value of *y*:

$$\frac{dy}{da} = \frac{\partial f}{\partial x_1}\cdot\frac{dx_1}{da} + \frac{\partial f}{\partial x_2}\cdot\frac{dx_2}{da} + \frac{\partial f}{\partial x_3}\cdot\frac{dx_3}{da}$$

# Functions of Several Variables

- If $x_3 = a$, then: $y = f[x_1(a), x_2(a), a]$
  - The effect of a on y:
    - A direct effect (which is given by $f_a$
    - An indirect effect that operates only through the ways in which a affects the x's

$$\frac{dy}{da} = \frac{\partial f}{\partial x_1} \cdot \frac{dx_1}{da} + \frac{\partial f}{\partial x_2} \cdot \frac{dx_2}{da} + \frac{\partial f}{\partial a}$$

# Functions of Several Variables

- Implicit functions
  - If the value of a function is held constant
    - An implicit relationship is created among the independent variables that enter into the function
    - The independent variables can no longer take on any values
      - But must instead take on only that set of values that result in the function's retaining the required value

# Functions of Several Variables

- Implicit functions
  - Ability to quantify the trade-offs inherent in most economic models
- $y = f(x_1,x_2)$; Implicit function: $x_2=g(x_1)$

$$y = 0 = f(x_1, x_2) = f(x_1, g(x_1))$$

Differentiate with respect to $x_1$: $0 = f_1 + f_2 \cdot \dfrac{dg(x_1)}{dx_1}$

Rearranging terms: $\dfrac{dg(x_1)}{dx_1} = \dfrac{dx_2}{dx_1} = -\dfrac{f_1}{f_2}$

# EXAMPLE 2.3 Using the Chain Rule

- A pizza fanatic
  - Each week, he consumes three kinds of pizza, denoted by $x_1$, $x_2$, and $x_3$
    - Cost of type 1 pizza is *p* per pie
    - Cost of type 2 pizza is *2p*
    - Cost of type 3 pizza is *3p*
  - Allocates \$30 each week to each type of pizza
  - How the total number of pizzas purchased is affected by the underlying price p

# EXAMPLE 2.3 Using the Chain Rule

- Quantity purchased:
  - $x_1=30/p$; $x_2=30/2p$; $x_3=30/3p$
- Total pizza purchases:
  - $y = f[x_1(p), x_2(p), x_3(p)] = x_1(p) + x_2(p) + x_3(p)$
- Applying the chain rule:

$$\frac{dy}{dp} = f_1 \cdot \frac{dx_1}{dp} + f_2 \cdot \frac{dx_2}{dp} + f_3 \cdot \frac{dx_3}{dp} =$$

$$= -30p^{-2} - 15p^{-2} - 10p^{-2} = -55p^{-2}$$

- A production possibility frontier for two goods of the form

  $x^2+0.25y^2=200$

- The implicit function:

$$\frac{dy}{dx}=\frac{-f_x}{f_y}=\frac{-2x}{0.5y}=\frac{-4x}{y}$$

# Maximization of Functions of Several Variables

- Suppose an agent wishes to maximize

$$y = f(x_1, x_2, \ldots, x_n)$$

  - The change in $y$ from a change in $x_1$ (holding all other $x$'s constant) is
    - Equal to the change in $x_1$ times the slope (measured in the $x_1$ direction)

$$dy = \frac{\partial f}{\partial x_1} dx_1 = f_1 dx_1$$

# Maximization of Functions of Several Variables

- First-order conditions for a maximum
  - Necessary condition for a maximum of the function $f(x_1,x_2,\ldots,x_n)$ is that $dy = 0$ for any combination of small changes in the $x$'s:

    $$f_1=f_2=\ldots=f_n=0$$

    - Critical point of the function
  - Not <u>sufficient</u> to ensure a maximum
- Second-order conditions, $f_{ii} < 0$
  - Second partial derivatives must be negative

# EXAMPLE 2.5 Finding a Maximum

- Suppose that $y$ is a function of $x_1$ and $x_2$

$$y = -(x_1 - 1)^2 - (x_2 - 2)^2 + 10$$

$$y = -x_1^2 + 2x_1 - x_2^2 + 4x_2 + 5$$

- First-order conditions imply that

$$\frac{\partial y}{\partial x_1} = -2x_1 + 2 = 0$$

$$\frac{\partial y}{\partial x_2} = -2x_2 + 4 = 0$$

OR

$$x_1^* = 1$$

$$x_2^* = 2$$

# The Envelope Theorem

- The envelope theorem
  - How the optimal value for a function changes when a parameter of the function changes
- A specific example: $y = -x^2 + ax$
  - Represents a family of inverted parabolas
    - For different values of *a*
  - Is a function of *x* only
    - If *a* is assigned a specific value
    - Can calculate the value of *x* that maximizes *y*

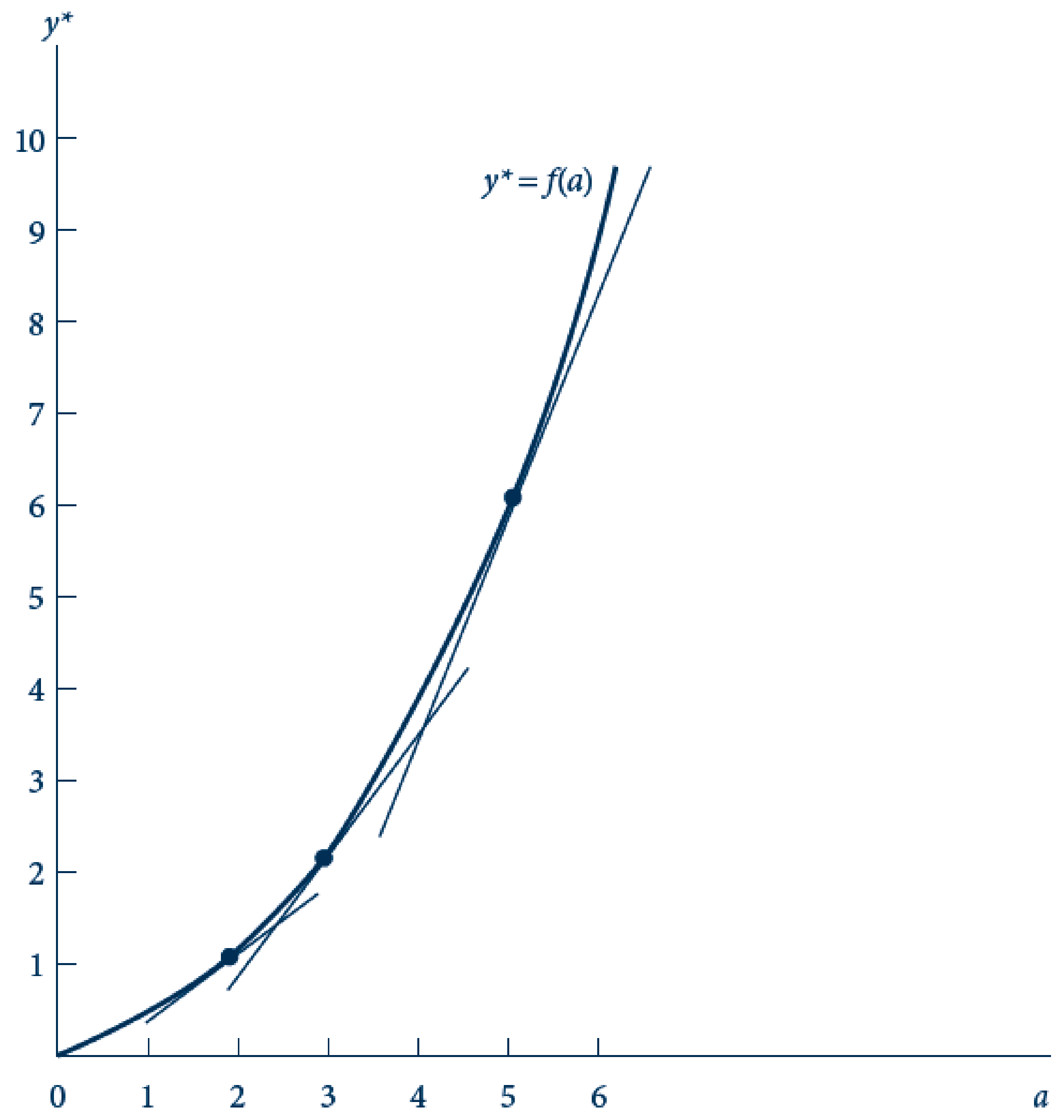
## TABLE 2.1

# Optimal values of y and x for alternative values of a in $y=-x^2+ax$

| Value of $a$ | Value of $x^*$ | Value of $y^*$ |
|---|---|---|
| 0 | 0 | 0 |
| 1 | $\frac{1}{2}$ | $\frac{1}{4}$ |
| 2 | 1 | 1 |
| 3 | $\frac{3}{2}$ | $\frac{9}{4}$ |
| 4 | 2 | 4 |
| 5 | $\frac{5}{2}$ | $\frac{25}{4}$ |
| 6 | 3 | 9 |

FIGURE 2.3

# Illustration of the Envelope Theorem




The envelope theorem states that the slope of the relationship between y (the maximum value of y) and the parameter a can be found by calculating the slope of the auxiliary relationship found by substituting the respective optimal values for x into the objective function and calculating ∂y/∂a.

# The Envelope Theorem

- If we are interested in how $y^*$ changes as $a$ changes
  - Calculate the slope of $y$ directly
  - Hold $x$ constant at its optimal value and calculate $\partial y/\partial a$ directly (the envelope theorem)

# The Envelope Theorem

- Calculate the slope of *y* directly
  - Must solve for the optimal value of *x* for any value of *a*

$$dy/dx = -2x + a = 0; \quad x^* = a/2$$

  - Substituting, we get

$$y^* = -(x^*)^2 + a(x^*) = -(a/2)^2 + a(a/2);$$

$$y^* = -a^2/4 + a^2/2 = a^2/4$$

- Therefore, $dy^*/da = 2a/4 = a/2$

# The Envelope Theorem

- Using the envelope theorem
  - For small changes in *a*, $dy^*/da$ can be computed by holding *x* at $x^*$ and calculating $\partial y/\partial a$ directly from *y*

$$\partial y/\partial a = x$$

  - Holding $x = x^*$

$$\partial y/\partial a = x^* = a/2$$

# The Envelope Theorem

- The envelope theorem
  - The change in the optimal value of a function with respect to a parameter of that function
  - Can be found by partially differentiating the objective function while holding *x* (or several *x*'s) at its optimal value

$$\frac{dy^*}{da} = \frac{\partial y}{\partial a}\{x = x^*(a)\}$$

# The Envelope Theorem

- Many-variable case
  - y is a function of several variables

    $y = f(x_1,\ldots x_n,a)$
  - Finding an optimal value for *y:* solve *n* first-order equations:

    $\partial y/\partial x_i = 0 \quad (i = 1,\ldots,n)$
  - Optimal values for these *x*'s would be a function of *a*

    $x_1^* = x_1^*(a);\ x_2^* = x_2^*(a);\ \ldots;\ x_n^* = x_n^*(a)$

# The Envelope Theorem

- Many-variable case
  - Substituting into the original objective function gives us the optimal value of *y* (*y**)

$$y^* = f[x_1^*(a), x_2^*(a), \ldots, x_n^*(a), a]$$

  - Differentiating yields

$$\frac{dy^*}{da} = \frac{\partial f}{\partial x_1} \cdot \frac{dx_1}{da} + \frac{\partial f}{\partial x_2} \cdot \frac{dx_2}{da} + \ldots + \frac{\partial f}{\partial x_n} \cdot \frac{dx_n}{da} + \frac{\partial f}{\partial a}$$

$$\frac{dy^*}{da} = \frac{\partial f}{\partial a}$$

# EXAMPLE 2.6 The Envelope Theorem: Health Status Revisited

- $y = -(x_1 - 1)^2 - (x_2 - 2)^2 + 10$
  - We found: $x_1^*=1$, $x_2^*=2$, and $y^*=10$
- For $y = -(x_1 - 1)^2 - (x_2 - 2)^2 + a$
  - $x_1^*=1$, $x_2^*=2$
  - $y^*=a$ and $dy^*/da = 1$
- Using the envelope theorem:

$$\frac{dy^*}{da} = \frac{\partial f}{\partial a} = 1$$

# Constrained Maximization

- What if all values for the *x*'s are not feasible?
  - The values of *x* may all have to be > 0
  - A consumer's choices are limited by the amount of purchasing power available
- Lagrange multiplier method
  - One method used to solve constrained maximization problems

# Lagrange Multiplier Method

- Lagrange multiplier method
  - Suppose that we wish to find the values of $x_1, x_2, \ldots, x_n$ that maximize:

    $y = f(x_1, x_2, \ldots, x_n)$
  - Subject to a constraint: $g(x_1, x_2, \ldots, x_n) = 0$
- The Lagrangian expression

  $\mathscr{L} = f(x_1, x_2, \ldots, x_n) + \lambda g(x_1, x_2, \ldots, x_n)$
  - $\lambda$ is called the Lagrange multiplier
  - $\mathscr{L} = f$, because $g(x_1, x_2, \ldots, x_n) = 0$

# Lagrange Multiplier Method

- First-order conditions
  - Conditions for a critical point for the function $\mathscr{L}$

$$\partial\mathscr{L}/\partial x_1 = f_1 + \lambda g_1 = 0$$

$$\partial\mathscr{L}/\partial x_2 = f_2 + \lambda g_2 = 0$$

$$\ldots$$

$$\partial\mathscr{L}/\partial x_n = f_n + \lambda g_n = 0$$

$$\partial\mathscr{L}/\partial \lambda = g(x_1, x_2, \ldots, x_n) = 0$$

# Lagrange Multiplier Method

- First-order conditions
  - Can generally be solved for $x_1, x_2, \ldots, x_n$ and $\lambda$
  - The solution will have two properties:
    - The $x$'s will obey the constraint
    - These $x$'s will make the value of $\mathscr{L}$ (and therefore $f$) as large as possible

# Lagrange Multiplier Method

- The Lagrangian multiplier ($\lambda$)
  - Important economic interpretation
  - The first-order conditions imply that

$$f_1/\text{-}g_1 = f_2/\text{-}g_2 = \ldots = f_n/\text{-}g_n = \lambda$$

    - The numerators measure the marginal benefit of one more unit of $x_i$
    - The denominators reflect the added burden on the constraint of using more $x_i$

# Lagrange Multiplier Method

- The Lagrangian multiplier ($\lambda$)
  - At the optimal $x_i$'s, the ratio of the marginal benefit to the marginal cost of $x_i$ should be the same for every $x_i$
  - $\lambda$ is the common cost-benefit ratio for all $x_i$

$$\lambda = \frac{\text{marginal benefit of } x_i}{\text{marginal cost of } x_i}$$

# Lagrange Multiplier Method

- The Lagrangian multiplier ($\lambda$)
  - A high value of $\lambda$ indicates that each $x_i$ has a high cost-benefit ratio
  - A low value of $\lambda$ indicates that each $x_i$ has a low cost-benefit ratio
  - $\lambda = 0$ implies that the constraint is not binding

# Constrained Maximization

- Duality
  - Any constrained maximization problem has a dual problem in constrained minimization
    - Focuses attention on the constraints in the original problem

# Constrained Maximization

- Individuals maximize utility subject to a budget constraint
  - Dual problem: individuals minimize the expenditure needed to achieve a given level of utility
- Firms minimize the cost of inputs to produce a given level of output
  - Dual problem: firms maximize output for a given cost of inputs purchased

# EXAMPLE 2.7 Constrained Maximization: Health status yet again

- Individual's goal is to maximize
  - $y = -x_1^2 + 2x_1 - x_2^2 + 4x_2 + 5$
  - With the constraint: $x_1 + x_2 = 1$ or $1 - x_1 - x_2 = 0$
  - Set up the Lagrangian expression:
    - $\mathscr{L} = = -x_1^2 + 2x_1 - x_2^2 + 4x_2 + 5 + \lambda(1 - x_1 - x_2)$
  - First-order conditions:

$$\partial\mathscr{L}/\partial x_1 = -2x_1 + 2 - \lambda = 0$$

$$\partial\mathscr{L}/\partial x_2 = -2x_2 + 4 - \lambda = 0$$

$$\partial\mathscr{L}/\partial\lambda = 1 - x_1 - x_2 = 0$$

  - Solution: $x_1=0$, $x_2=1$, $\lambda=2$, $y=8$

## EXAMPLE 2.8 Optimal Fences and Constrained Maximization

- Suppose a farmer had a certain length of fence ($P$)
  - Wished to enclose the largest possible rectangular area – with $x$ and $y$ the lengths of the sides
  - Choose $x$ and $y$ to maximize the area ($A = x \cdot y$)
  - Subject to the constraint that the perimeter is fixed at $P = 2x + 2y$

- The Lagrangian expression:

$$\mathscr{L} = x \cdot y + \lambda(P - 2x - 2y)$$

- First-order conditions

$$\partial\mathscr{L}/\partial x = y - 2\lambda = 0$$

$$\partial\mathscr{L}/\partial y = x - 2\lambda = 0$$

$$\partial\mathscr{L}/\partial\lambda = P - 2x - 2y = 0$$

  - $y/2 = x/2 = \lambda$, then $x=y$, the field should be square
  - $x = y$ and $y = 2\lambda$, then

$$x = y = P/4 \text{ and } \lambda = P/8$$

# EXAMPLE 2.8 Optimal Fences and Constrained Maximization

- Interpretation of the Lagrange multiplier
  - $\lambda$ suggests that an extra yard of fencing would add $P/8$ to the area
  - Provides information about the implicit value of the constraint
- Dual problem
  - Choose $x$ and $y$ to minimize the amount of fence required to surround the field

    minimize $P = 2x + 2y$ subject to $A = x \cdot y$
  - Setting up the Lagrangian:

    $\mathcal{L}^D = 2x + 2y + \lambda^D(A - x \cdot y)$

- Dual problem
  - First-order conditions:

$$\partial \mathscr{L}^D/\partial x = 2 - \lambda^D \cdot y = 0$$

$$\partial \mathscr{L}^D/\partial y = 2 - \lambda^D \cdot x = 0$$

$$\partial \mathscr{L}^D/\partial \lambda^D = A - x \cdot y = 0$$

  - Solving, we get: $x = y = A^{1/2}$
  - The Lagrangian multiplier $\lambda^D = 2A^{-1/2}$

# Envelope Theorem in Constrained Maximization Problems

- Suppose that we want to maximize

  $y = f(x_1,\ldots,x_n;a)$

  – Subject to the constraint: $g(x_1,\ldots,x_n;a) = 0$

- One way to solve
  – Set up the Lagrangian expression
  – Solve the first-order conditions

- Alternatively, it can be shown that

  $dy^*/da = \partial\mathscr{L}/\partial a(x_1^*,\ldots,x_n^*;a)$

# Inequality Constraints

- Maximize $y = f(x_1,x_2)$ subject to $g(x_1,x_2) \geq 0$, $x_1 \geq 0$, and $x_2 \geq 0$
- Slack variables
  - Introduce three new variables ($a$, $b$, and $c$) that convert the inequalities into equalities
  - Square these new variables

$$g(x_1,x_2) - a^2 = 0; \; x_1 - b^2 = 0; \text{ and } x_2 - c^2 = 0$$

  - Any solution that obeys these three equality constraints will also obey the inequality constraints

# Inequality Constraints

- Maximize $y = f(x_1,x_2)$ subject to $g(x_1,x_2) \geq 0$, $x_1 \geq 0$, and $x_2 \geq 0$
- Lagrange multipliers

$$\mathcal{L} = f(x_1,x_2) + \lambda_1[g(x_1,x_2) - a^2] + \lambda_2[x_1 - b^2] + \lambda_3[x_2 - c^2]$$

  – There will be 8 first-order conditions

$$\partial\mathcal{L}/\partial x_1 = f_1 + \lambda_1 g_1 + \lambda_2 = 0$$

$$\partial\mathcal{L}/\partial x_2 = f_1 + \lambda_1 g_2 + \lambda_3 = 0$$

$$\partial\mathcal{L}/\partial a = -2a\lambda_1 = 0$$

$$\partial\mathcal{L}/\partial b = -2b\lambda_2 = 0$$

$$\partial\mathcal{L}/\partial c = -2c\lambda_3 = 0$$

$$\partial\mathcal{L}/\partial \lambda_1 = g(x_1,x_2) - a^2 = 0$$

$$\partial\mathcal{L}/\partial \lambda_2 = x_1 - b^2 = 0$$

$$\partial\mathcal{L}/\partial \lambda_3 = x_2 - c^2 = 0$$

# Inequality Constraints

- Complementary slackness
  - According to the third condition, either *a* or $\lambda_1 = 0$
    - If $a = 0$, the constraint $g(x_1,x_2)$ holds exactly
    - If $\lambda_1 = 0$, the availability of some slackness of the constraint implies that its value to the objective function is 0
  - Similar complementary slackness relationships also hold for $x_1$ and $x_2$

# Inequality Constraints

- Complementary slackness
  - These results are sometimes called <u>Kuhn-Tucker conditions</u>
    - Show that solutions to problems involving inequality constraints will differ from those involving equality constraints in rather simple ways
      - Allows us to work primarily with constraints involving equalities

# Second-Order Conditions and Curvature

- Functions of one variable, $y = f(x)$
  - A necessary condition for a maximum: $dy/dx = f'(x) = 0$
    - $y$ must be decreasing for movements away from it
  - The total differential measures the change in $y$: $dy = f'(x)\,dx$
    - To be at a maximum, $dy$ must be decreasing for small increases in $x$

# Second-Order Conditions and Curvature

- Functions of one variable, $y = f(x)$
  - To see the changes in $dy$, we must use the second derivative of $y$

$$d(dy) = d^2 y = \frac{d[f'(x)dx]}{dx} \cdot dx = f''(x)dx \cdot dx = f''(x)dx^2$$

    - Since $d^2y < 0$ , $f''(x)dx^2 < 0$
    - Since $dx^2$ must be $> 0$, $f''(x) < 0$
    - This means that the function $f$ must have a concave shape at the critical point

- Finding the maximum of: $\pi = 1{,}000q - 5q^2$
  - First-order condition:
    - $d\pi/dq = 1{,}000 - 10q = 0$, so $q^* = 100$
  - Second derivative of the function
    - $d^2\pi/dq^2 = -10 < 0$
  - Hence the point $q^*=100$ obeys the sufficient conditions for a local maximum

# Second-Order Conditions and Curvature

- Functions of two variables, $y = f(x_1, x_2)$
  - First order conditions for a maximum:

$$\partial y/\partial x_1 = f_1 = 0$$

$$\partial y/\partial x_2 = f_2 = 0$$

  - $f_1$ and $f_2$ must be diminishing at the critical point
  - Conditions must also be placed on the cross-partial derivative ($f_{12} = f_{21}$)

# Second-Order Conditions and Curvature

- The total differential of $y$: $dy = f_1\, dx_1 + f_2\, dx_2$
  - The differential:

$$d^2y = (f_{11}dx_1 + f_{12}dx_2)dx_1 + (f_{21}dx_1 + f_{22}dx_2)dx_2$$

$$d^2y = f_{11}dx_1^2 + f_{12}dx_2dx_1 + f_{21}dx_1\, dx_2 + f_{22}dx_2^2$$

  - By Young's theorem, $f_{12} = f_{21}$ and

$$d^2y = f_{11}dx_1^2 + 2f_{12}dx_1dx_2 + f_{22}dx_2^2$$

$$d^2y = f_{11}dx_1^2 + 2f_{12}dx_1dx_2 + f_{22}dx_2^2$$

    – $d^2y < 0$ for any $dx_1$ and $dx_2$, if $f_{11}<0$ and $f_{22}<0$

    – If neither $dx_1$ nor $dx_2$ is zero, then $d^2y < 0$ only if $f_{11}\, f_{22} - f_{12}^2 > 0$

# EXAMPLE 2.10 Second-Order Conditions: Health status

- $y = f(x_1, x_2) = -x_1^2 + 2x_1 - x_2^2 + 4x_2 + 5$
  - First-order conditions
    - $f_1 = -2x_1 + 2 = 0$ and $f_2 = -2x_2 + 4 = 0$
    - Or: $x_1^* = 1$, $x_2^* = 2$
  - Second-order partial derivatives
    - $f_{11} = -2$
    - $f_{22} = -2$
    - $f_{12} = 0$

# Second-Order Conditions and Curvature

- Concave functions
  - $f_{11}\,f_{22} - f_{12}^2 > 0$
  - Have the property that they always lie below any plane that is tangent to them
    - The plane defined by the maximum value of the function is simply a special case of this property

# Second-Order Conditions and Curvature

- Constrained maximization
  - Choose $x_1$ and $x_2$ to maximize: $y = f(x_1, x_2)$
  - Linear constraint: $c - b_1x_1 - b_2x_2 = 0$
  - The Lagrangian: $\mathscr{L} = f(x_1, x_2) + \lambda(c - b_1x_1 - b_2x_2)$
  - The first-order conditions:

$$f_1 - \lambda b_1 = 0,\ f_2 - \lambda b_2 = 0,$$

$$\text{and } c - b_1x_1 - b_2x_2 = 0$$

# Second-Order Conditions and Curvature

- Constrained maximization
  - Use the "second" total differential:

$$d^{2}y = f_{11}dx_1^{2} + 2f_{12}dx_1dx_2 + f_{22}dx_2^{2}$$

    - Only values of $x_1$ and $x_2$ that satisfy the constraint can be considered valid alternatives to the critical point
  - Total differential of the constraint

$$-b_1\, dx_1 - b_2\, dx_2 = 0,\ dx_2 = -(b_1/b_2)dx_1$$

    - Allowable relative changes in $x_1$ and $x_2$

# Second-Order Conditions and Curvature

- Constrained maximization
  - First-order conditions imply that $f_1/f_2 = b_1/b_2$, we get: $dx_2 = -(f_1/f_2)\, dx_1$
  - Since: $d^2y = f_{11}dx_1^2 + 2f_{12}dx_1dx_2 + f_{22}dx_2^2$
  - Substitute for $dx_2$ and get

$$d^2y = f_{11}dx_1^2 - 2f_{12}(f_1/f_2)dx_1^2 + f_{22}(f_1^2/f_2^2)dx_1^2$$

  - Combining terms and rearranging, we get

$$d^2y = f_{11}\, f_2^2 - 2f_{12}f_1f_2 + f_{22}f_1^2\, [dx_1^2/\, f_2^2]$$

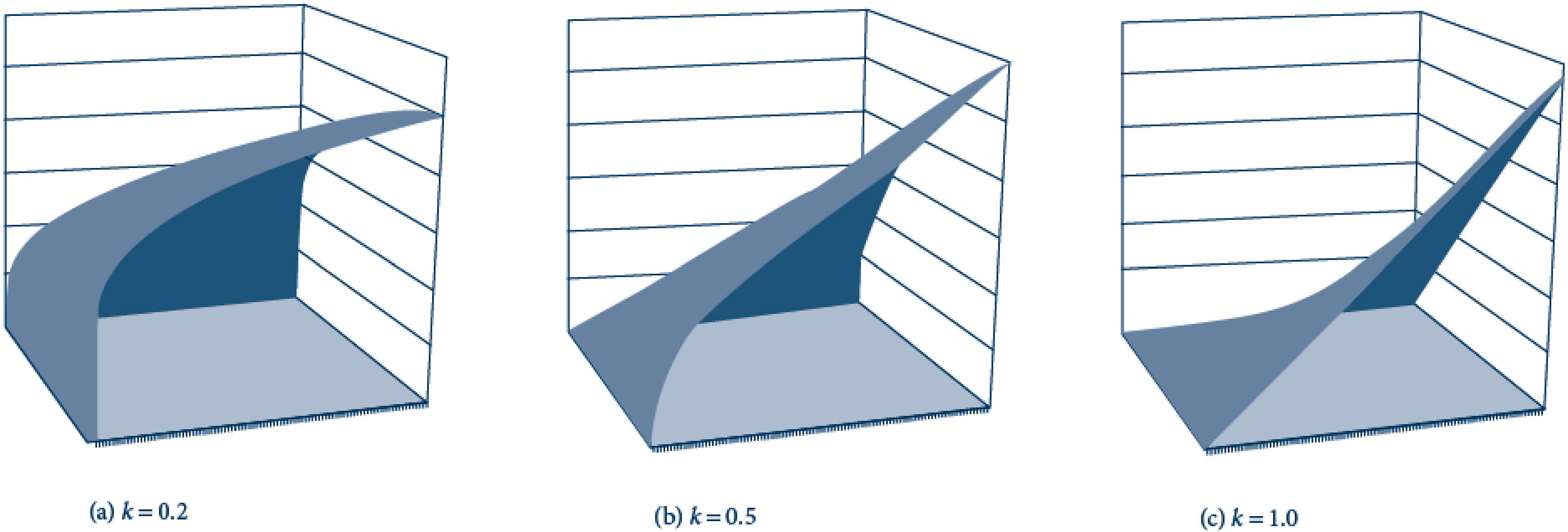
# Second-Order Conditions and Curvature

- Constrained maximization
  - Therefore, for $d^2y < 0$, it must be true that
    $$f_{11} f_2^2 - 2f_{12}f_1f_2 + f_{22}f_1^2 < 0$$
    - This equation characterizes a set of functions termed quasi-concave functions
- Quasi-concave functions
  - Any two points within the set can be joined by a line contained completely in the set

# EXAMPLE 2.11 Concave and Quasi-Concave Functions

- $y = f(x_1,x_2) = (x_1 \cdot x_2)^k$
  - Where $x_1 > 0$, $x_2 > 0$, and $k > 0$
  - No matter what value $k$ takes, this function is quasi-concave
- Whether or not the function is concave depends on the value of $k$
  - If $k < 0.5$, the function is concave
  - If $k > 0.5$, the function is convex

# Concave and Quasi-Concave Functions




In all three cases these functions are quasi-concave. For a fixed y, their level curves are convex. But only for k =0.2 is the function strictly concave. The case k = 1.0 clearly shows nonconcavity because the function is not below its tangent plane.

# Homogeneous Functions

- A function $f(x_1,x_2,\ldots x_n)$ is said to be homogeneous of degree $k$ if

$$f(tx_1,tx_2,\ldots tx_n) = t^k f(x_1,x_2,\ldots x_n)$$

  - When $k = 1$, a doubling of all of its arguments doubles the value of the function itself
  - When $k = 0$, a doubling of all of its arguments leaves the value of the function unchanged

# Homogeneous Functions

- If a function is homogeneous of degree *k*
  - The partial derivatives of the function will be homogeneous of degree *k*-1
- Euler's theorem, homogeneous function
  - Differentiate the definition for homogeneity with respect to the proportionality factor *t*

$$kt^{k-1}f(x_1,\ldots,x_n) = x_1 f_1(tx_1,\ldots,tx_n) + \ldots + x_n f_n(x_1,\ldots,x_n)$$

    - There is a definite relationship between the value of the function and the values of its partial derivatives

# Homogeneous Functions

- A homothetic function
  - Is one that is formed by taking a monotonic transformation of a homogeneous function
  - They generally do not possess the homogeneity properties of their underlying functions

# Homogeneous Functions

- Homogeneous and homothetic functions
  - The implicit trade-offs among the variables in the function
  - Depend only on the ratios of those variables, not on their absolute values
- Two-variable function, $y=f(x_1,x_2)$
  - The implicit trade-off between $x_1$ and $x_2$ is: $dx_2/dx_1 = -f_1/f_2$
  - $f$ is homogeneous of degree $k$

# Homogeneous Functions

- Two-variable function, $y=f(x_1,x_2)$
  - Its partial derivatives will be homogeneous of degree *k-1*
  - The implicit trade-off between $x_1$ and $x_2$ is

$$\frac{dx_2}{dx_1} = -\frac{t^{k-1} f_1(tx_1, tx_2)}{t^{k-1} f_2(tx_1, tx_2)} = -\frac{f_1(tx_1, tx_2)}{f_2(tx_1, tx_2)}$$

$$\text{Let } t = 1/x_2$$

$$\frac{dx_2}{dx_1} = -\frac{f_1(x_1/x_2, 1)}{f_2(x_1/x_2, 1)}$$

# EXAMPLE 2.12 Cardinal and Ordinal Properties

- Function $f(x_1,x_2)=(x_1x_2)^k$
  - Quasi-concavity [an ordinal property] - preserved for all values of $k$
  - Is concave [a cardinal property] - only for a narrow range of values of $k$
    - Many monotonic transformations destroy the concavity of $f$
  - A proportional increase in the two arguments: $f(tx_1,tx_2)=t^{2k}x_1x_2 = t^{2k}f(x_1,x_2)$
  - Degree of homogeneity - depends on $k$
  - Is homothetic because $\frac{dx_2}{dx_1} = -\frac{f_1}{f_2} = -\frac{kx_1^{k-1}x_2^k}{kx_1^k x_2^{k-1}} = -\frac{x_2}{x_1}$

# Integration

- Integration is the inverse of differentiation
  - Let $F(x)$ be the integral of $f(x)$
  - Then $f(x)$ is the derivative of $F(x)$

$$\frac{dF(x)}{dx} = F'(x) = f(x)$$

$$F(x) = \int f(x)dx$$

- If f(x) = x then

$$F(x) = \int f(x)dx = \int xdx = \frac{x^2}{2} + C$$

# Integration

- Calculation of antiderivatives
  1. Creative guesswork
     - What function will yield $f(x)$ as its derivative?
     - Use differentiation to check your answer
  2. Change of variable
     - Redefine variables to make the function easier to integrate
  3. Integration by parts

# Integration

- Integration by parts: *duv = udv + vdu*
  - For any two functions *u* and *v*

$$\int duv = uv = \int udv + \int vdu$$

$$\int udv = uv - \int vdu$$

# Integration

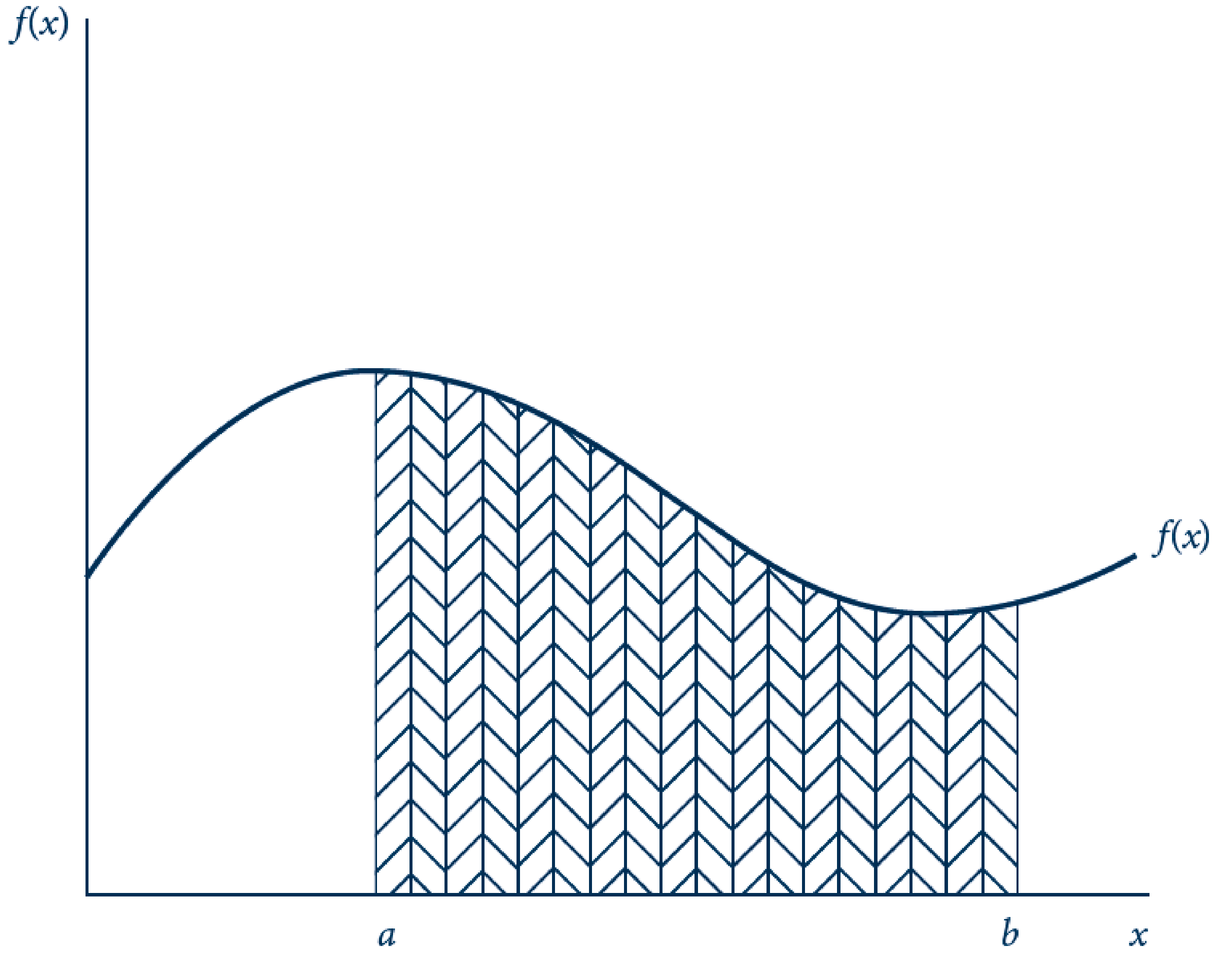
- Definite integrals
  - To sum up the area under a graph of a function over some defined interval
- Area under *f(x)* from *x* = *a* to *x* = *b*

$$\text{area under } f(x) \approx \sum_i f(x_i)\Delta x_i$$

$$\text{area under } f(x) = \int_{x=a}^{x=b} f(x)dx$$

FIGURE 2.5

# Definite Integrals Show the Areas Under the Graph of a Function




Definite integrals measure the area under a curve by summing rectangular areas as shown in the graph. The dimension of each rectangle is f(x)dx.

# Integration

- Fundamental theorem of calculus
  - Directly ties together the two principal tools of calculus: derivatives and integrals
  - Used to illustrate the distinction between "stocks" and "flows

$$\text{area under } f(x) = \int_{x=a}^{x=b} f(x)dx = F(b) - F(a)$$

- Net population increase, $f(t)=1{,}000e^{0.02t}$
  - "Flow" concept
  - Net population change - is growing at the rate of 2 percent per year
  - How much in total the population ("stock" concept) will increase within 50 years:

$$\text{increase in population} = \int_{t=0}^{t=50} f(t)dt = \int_{t=0}^{t=50} 1{,}000e^{0.02t}dt =$$

$$= \quad F(t)\Big|_0^{50} = \frac{1{,}000e^{0.02t}}{0.02}\Big|_0^{50} = \frac{1{,}000e}{0.02} - 50{,}000 = 85{,}914$$

- Total costs: $C(q)=0.1q^2+500$
  - q – output during some period
  - Variable costs: $0.1q^2$
  - Fixed costs: 500
  - Marginal costs MC = $dC(q)/dq=0.2q$
  - Total costs for q=100
    - Fixed cost (500) + Variable cost

$$\text{variable cost} = \int_{q=0}^{q=100} 0.2q\,dq = \left. 0.1q^2 \right|_0^{100} = 1{,}000 - 0 = 1{,}000$$

# Differentiating a Definite Integral

1. Differentiation with respect to the variable of integration
   - A definite integral has a constant value
   - Hence its derivative is zero

$$\frac{d\int_a^b f(x)dx}{dx} = 0$$

# Differentiating a Definite Integral

2. Differentiation with respect to the upper bound of integration

– Changing the upper bound of integration will change the value of a definite integral

$$\frac{d\int_a^x f(t)dt}{dx} = \frac{d\left[F(x)-F(a)\right]}{dx} = f(x)-0 = f(x)$$

# Differentiating a Definite Integral

2. Differentiation with respect to the upper bound of integration

   – If the upper bound of integration is a function of x,

$$\frac{d\int_a^{g(x)} f(t)dt}{dx} = \frac{d\left[F(g(x)) - F(a)\right]}{dx} =$$

$$= \frac{d\left[F(g(x))\right]}{dx} = f\,\frac{dg(x)}{dx} = f(g(x))g'(x)$$

# Differentiating a Definite Integral

3. Differentiation with respect to another relevant variable
   - Suppose we want to integrate $f(x,y)$ with respect to $x$
     - How will this be affected by changes in $y$?

$$\frac{d\int_a^b f(x,y)dx}{dy} = \int_a^b f_y(x,y)dx$$

# Dynamic Optimization

- Some optimization problems involve multiple periods
  - Need to find the optimal time path for a variable that succeeds in optimizing some goal
  - Decisions made in one period affect outcomes in later periods

# Dynamic Optimization

- Find the optimal path for $x(t)$
  - Over a specified time interval $[t_0,t_1]$
  - Changes in $x$ are governed by

$$\frac{dx(t)}{dt} = g\left[x(t),c(t),t\right]$$

    - c(t) is used to "control" the change in x(t)
  - Each period: derive value from x and c from f [x(t),c(t),t]

# Dynamic Optimization

- Find the optimal path for $x(t)$
  - Each period: derive value from x and c from f [x(t),c(t),t]
  - Optimize

$$\int_{t_0}^{t_1} f\left[x(t),c(t),t\right]dt$$

- There may also be endpoint constraints: $x(t_0) = x_0$ and $x(t_1) = x_1$

# Dynamic Optimization

- The maximum principle
  - At a single point in time, the decision maker must be concerned with
    - The current value of the objective function
    - The implied change in the value of $x(t)$ from its current value of $\lambda(t)x(t)$ given by

$$\frac{d\left[\lambda(t)x(t)\right]}{dt} = \lambda(t)\frac{dx(t)}{dt} + x(t)\frac{d\lambda(t)}{dt}$$

# Dynamic Optimization

- The maximum principle
  - At any time $t$, a comprehensive measure of the value of concern to the decision maker is:

$$H = f\left[x(t), c(t), t\right] + \lambda(t)\, g\left[x(t), c(t), t\right] + x(t)\frac{d\lambda(t)}{dt}$$

    - Represents both the current benefits being received and the instantaneous change in the value of x

# Dynamic Optimization

- The maximum principle
  - The two optimality conditions

$$1st: \frac{\partial H}{\partial c} = f_c + \lambda g_c = 0, \text{ or } f_c = -\lambda g_c$$

$$2nd: \frac{\partial H}{\partial x} = f_x + \lambda g_x + \frac{\partial \lambda(t)}{\partial t} = 0,$$

$$\text{or } f_x + \lambda g_x = -\frac{\partial \lambda(t)}{\partial t}$$

# Dynamic Optimization

- The maximum principle
  - The 1st condition:
    - Present gains from *c* must be balanced against future costs
  - The 2nd condition:
    - The current gain from more *x* must be weighed against the declining future value of *x*

# EXAMPLE 2.14 Allocating a Fixed Supply

- Inherited 1,000 bottles of wine
  - Drink them bottles over the next 20 years
  - Maximize the utility
  - Utility function for wine is given by $u[c(t)] = \ln c(t)$
    - Diminishing marginal utility: $u' > 0$, $u'' < 0$
  - Maximize

$$\int_0^{20} u[c(t)]dt = \int_0^{20} \ln c(t)dt$$

- Let x(t) = the number of bottles of wine remaining at time t
  - Constrained by x(0) = 1,000 and x(20) = 0
  - The differential equation determining the evolution of x(t): $dx(t)/dt=-c(t)$
  - The current value Hamiltonian expression

$$H = \ln c(t) + \lambda[-c(t)] + x(t)\frac{d\lambda}{dt}$$

First-order conditions:

$$\frac{\partial H}{\partial c} = \frac{1}{c} - \lambda = 0, \text{ and } \frac{\partial H}{\partial x} = \frac{d\lambda}{dt} = 0$$

- For the utility function:

$$u[c(t)] = \begin{cases} c(t)^{\gamma} / \gamma \text{ if } \gamma \neq 0, \gamma < 1; \\ \ln c(t) \text{ if } \gamma = 0 \end{cases}$$

$$\text{Maximize: } \int_0^{20} u[c(t)]dt = \int_0^{20} e^{-\delta t} \frac{c(t)^{\gamma}}{\gamma} dt$$

$$\text{Constraints: } \frac{dx(t)}{dt} = -c(t);$$

$$x(0) = 1,000; \quad \text{and} \quad x(20) = 0$$

Hamiltonian: $H = e^{-\delta t}\dfrac{c(t)^{\gamma}}{\gamma} + \lambda(-c) + x(t)\dfrac{d\lambda(t)}{dt}$

The maximum principle:

$$\frac{\partial H}{\partial c} = e^{-\delta t}[c(t)]^{\gamma-1} - \lambda = 0,$$

and $\dfrac{\partial H}{\partial x} = 0 + 0 + \dfrac{d\lambda}{dt} = 0$

$\lambda = k$ (a constant)

$e^{-\delta t}[c(t)]^{\gamma-1} = k$, or $c(t) = k^{1/(\gamma-1)} e^{\delta t/(\gamma-1)}$

# Mathematical Statistics

- A random variable
  - Describes the outcomes from an experiment subject to chance
  - Discrete (roll of a die)
  - Continuous (outside temperature)
- e.g., flipping a coin

$$x = \begin{cases} 1 & \text{if coin is heads} \\ 0 & \text{if coin is tails} \end{cases}$$

# Mathematical Statistics

- Probability density function (PDF)
  - For any random variable
  - Shows the probability that each outcome will occur
  - The probabilities specified by the PDF must sum to 1

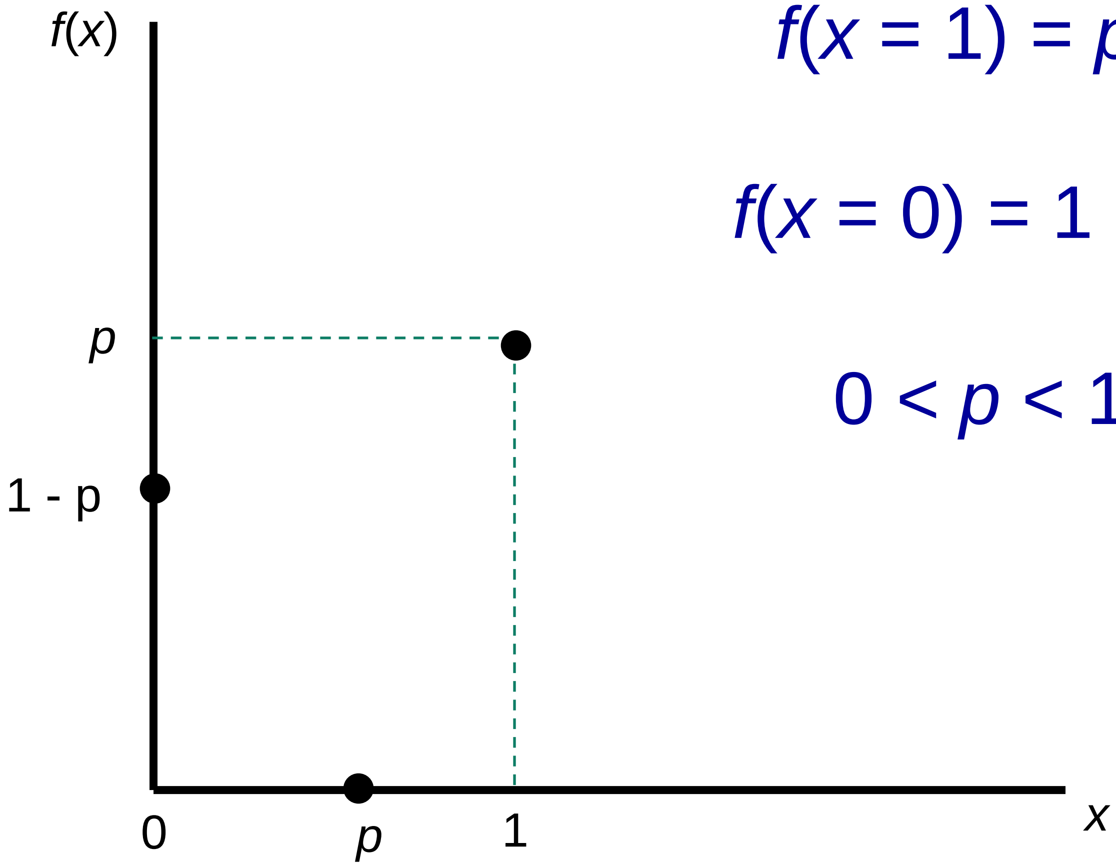
Discrete case:

$$\sum_{i=1}^{n} f(x_i) = 1$$

Continuous case:

$$\int_{-\infty}^{+\infty} f(x)\,dx = 1$$

# Four Common Probability Density Functions

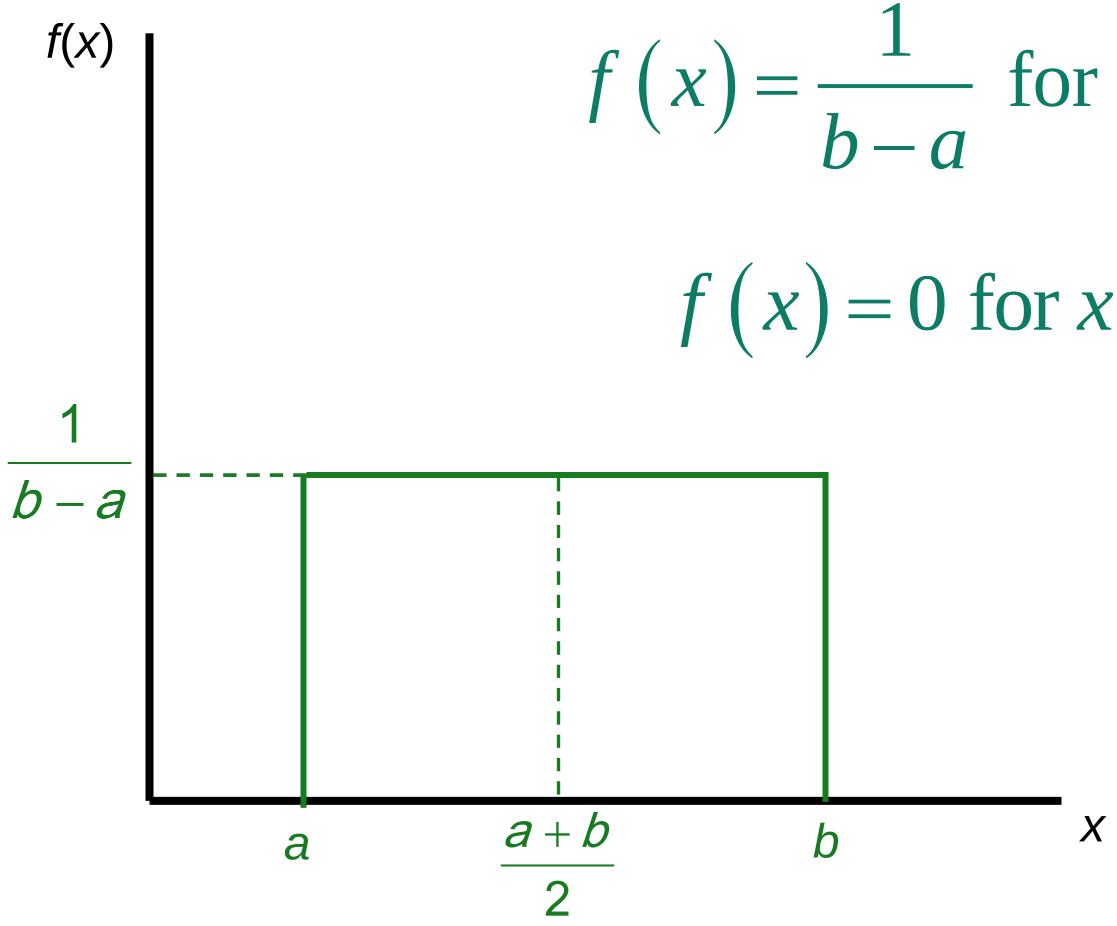



$$f(x = 1) = p$$

$$f(x = 0) = 1 - p$$

$$0 < p < 1$$

# Four Common Probability Density Functions




$$f(x) = \frac{1}{b-a} \text{ for } a \le x \le b$$

$$f(x) = 0 \text{ for } x < a \text{ or } x > b$$

# Four Common Probability Density Functions

$$f(x) = \lambda e^{-\lambda x} \text{ if } x > 0$$

$$f(x) = 0 \text{ if } x \le 0$$

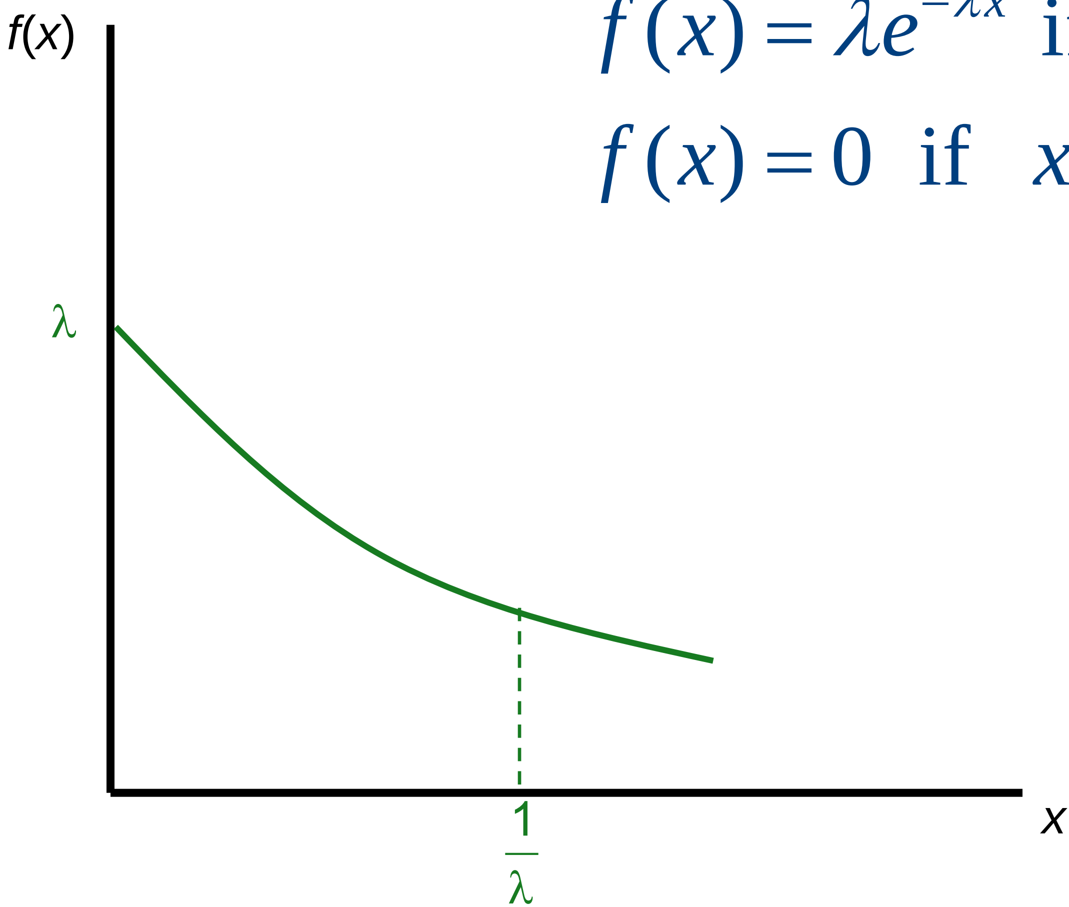
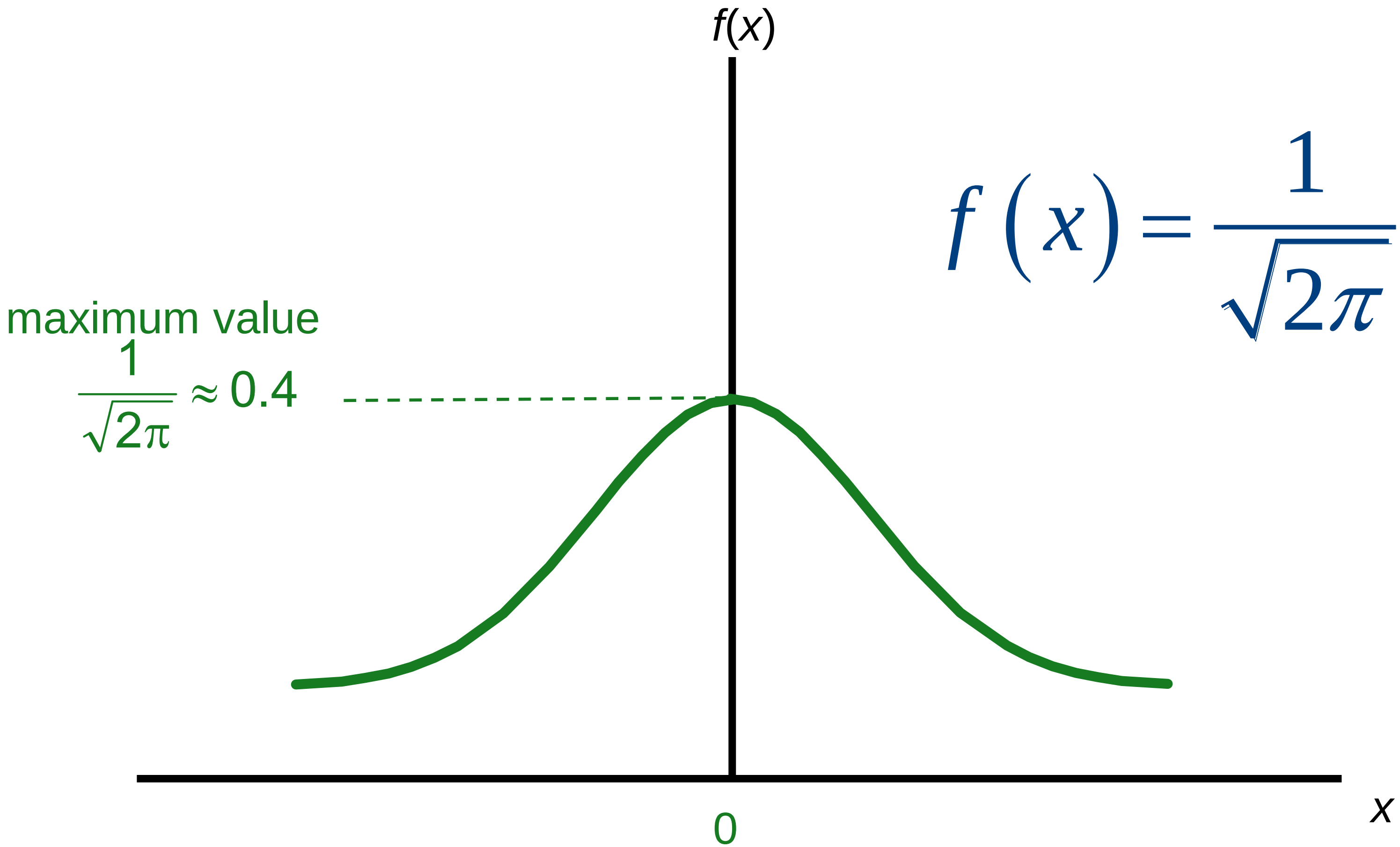


**FIGURE** 2.6 d Normal

# Four Common Probability Density Functions




$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

# Mathematical Statistics

- Expected value of a random variable
  - The numerical value that the random variable might be expected to have, on average
  - Measure of central tendency

Discrete case:

$$E(x) = \sum_{i=1}^{n} x_i f(x_i)$$

Continuous case:

$$E(x) = \int_{-\infty}^{+\infty} x f(x) dx$$

# Mathematical Statistics

- Expected value of a random variable
  - Extended to function of random variables

$$E[g(x)] = \int_{-\infty}^{+\infty} g(x) f(x) dx$$

Linear function: $y = ax + b$

$$E(y) = E(ax+b) = \int_{-\infty}^{+\infty} (ax+b) f(x) dx = aE(x) + b$$

# Mathematical Statistics

- Expected value of a random variable
  - Phrased in terms of the cumulative distribution function (CDF) F(x)
    - F(x) represents the probability that the random variable t is less than or equal to x

$$F(x) = \int_{-\infty}^{x} f(t)\,dt$$

Expected value of x: $$E(x) = \int_{-\infty}^{+\infty} x\,dF(x)$$

# EXAMPLE 2.15 Expected Values of a Few Random Variables

1. Binomial: $E(x) = 1?f(x = \quad) + 0?f(x = \quad) = p$

2. Uniform: $E(x) = \int_a^b \frac{x}{b-a}dx = \frac{b+a}{2}$

3. Exponential: $E(x) = \int_0^\infty x\lambda e^{-\lambda x}dx = \frac{1}{\lambda}$

4. Normal $E(x) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} xe^{-x^2/2}dx = 0$

# Mathematical Statistics

- Variance
  - A measure of dispersion
  - The expected squared deviation of a random variable from its expected value

$$Var(x) = \sigma_x^2 = E\left[\left(x - E(x)\right)^2\right] = \int_{-\infty}^{+\infty} \left(x - E(x)\right)^2 f(x)\,dx$$

# Mathematical Statistics

- Variance, *Var(x)*
  - A measure of dispersion
  - The expected squared deviation of a random variable from its expected value
- Standard deviation, σ
  - The square root of the variance

$$Var(x) = \sigma_x^2 = E\left[\left(x - E(x)\right)^2\right] = \int_{-\infty}^{+\infty} \left(x - E(x)\right)^2 f(x)\,dx$$

$$\sigma_x = \sqrt{Var(x)} = \sqrt{\sigma_x^2}$$

# EXAMPLE 2.16 Variances and Standard Deviations for Simple Random Variables

1. Binomial: $\sigma_x^2 = \sum_{i=1}^{n} (x_i - E(x))^2 f(x_i)$

$$\sigma_x^2 = (1-p)^2 \cdot p + (0-p)^2 \cdot (1-p) = p \cdot (1-p)$$

$$\sigma_x = \sqrt{p \cdot (1-p)}$$

2. Uniform: $\sigma_x^2 = \int_a^b \left( x - \frac{a+b}{2} \right)^2 \frac{1}{b-a} dx = \frac{(b-a)^2}{12}$

3. Exponential: $\sigma_x^2 = 1/\lambda^2$ and $\sigma_x = 1/\lambda$

4. Normal: $\sigma_x^2 = \sigma_x = 1$

# Standardizing the Normal

- If the random variable x has a standard Normal PDF
  - It will have an expected value of 0
  - And a standard deviation of 1
- Linear transformation y =σx + μ
  - Used to give this random variable any desired expected value (μ) and standard deviation (σ)

$$E(y) = \sigma E(x) + \mu = \mu$$

$$Var(y) = \sigma_y^2 = \sigma^2 Var(x) = \sigma^2$$

# Mathematical Statistics

- Covariance
  - Between two random variables (*x* and *y*)
  - Measures the direction of association between them

$$Cov(x,y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left[x-E(x)\right]\left[y-E(y)\right]f(x,y)\,dxdy$$

# Mathematical Statistics

- Two random variables are <u>independent</u>
  - If the probability of any particular value of one is not affected by the particular value of the other than may occur
  - This means that the PDF must have the property that $f(x,y)=g(x)\cdot h(y)$
  - $Cov(x,y) = 0$
    - Not sufficient to guarantee the two variables are statistically independent

# Mathematical Statistics

- If x and y are independent

$$Cov(x,y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}[x-E(x)][y-E(y)]g(x)h(y)dxdy=0$$

- Sum of two random variables

$$E(x+y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x+y)f(x,y)dxdy=E(x)+E(y)$$

$$Var(x+y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}[x+y-E(x+y)]^2 f(x,y)dxdy$$

$$Var(x+y)=Var(x)+Var(y)+2Cov(x,y)$$

# Matrix algebra background

- An n×k matrix is a rectangular array of terms
  - With i=1,n
  - With j=1,k

$$A = [a_{ij}] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \cdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nk} \end{bmatrix}$$

# Matrix algebra background

- If n=k, A is a square matrix: $a_{ij}=a_{ji}$
- Identity matrix, $I_n$ , is a square matrix where
  - $a_{ij}=1$ if $i=j$ and $a_{ij}=0$ if $i \neq j$
- The determinant of a square matrix, |A|
  - Is a scalar found by suitably multiplying together all the terms in the matrix
- The inverse of an n×n matrix, A,
  - Is another n×n matrix, $A^{-1}$,
  - Such that: $A \times A^{-1}=I_n$

# Matrix algebra background

- A necessary and sufficient condition for the existence of $A^{-1}$
  - $|A| \neq 0$
- The leading principal minors of an n × n square matrix A
  - Are the series of determinants of the first p rows and columns of A
  - Where p=1,n

# Matrix algebra background

- An n × n square matrix, A,
  - Is positive definite if all its leading principal minors are positive
  - Is negative definite if its principal minors alternate in sign starting with a minus
- Hessian matrix
  - Formed by all the second-order partial derivatives of a function

# Matrix algebra background

- Hessian of *f*
  - If *f* is a continuous and twice differentiable function of *n* variables

$$H(f) = \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1n} \\ f_{21} & f_{22} & \cdots & f_{2n} \\ \cdots & & & \\ f_{n1} & f_{n2} & \cdots & f_{nn} \end{bmatrix}$$

# Concave and convex functions

- A concave function
  - Is always below (or on) any tangent to it
  - $f''(x_0) \leq 0$
  - The Hessian matrix - negative definite
- A convex function
  - Is always above (or on) any tangent
  - $f''(x_0) \geq 0$
  - The Hessian matrix - positive definite

# Maximization

- First-order conditions
  - For an unconstrained maximum of a function of many variables
  - Requires finding a point at which the partial derivatives are zero
    - If the function is concave it will be below its tangent plane at this point
      - True maximum

# Constrained maxima

- Maximize $f(x_1,\ldots,x_n)$ subject to the constraint $g(x_1,\ldots,x_n)=0$
  - First-order conditions for a maximum: $f_i+\lambda g_i=0$
    - Where λ is the Lagrange multiplier
  - Second-order conditions for a maximum
    - Augmented (“bordered”) Hessian, $H_b$
    - $(-1)H_b$ must be negative definite

# Constrained maxima

- Augmented (“bordered”) Hessian, $H_b$

$$H_b = \begin{bmatrix} 0 & g_1 & g_2 & \cdots & g_n \\ g_1 & f_{11} & f_{12} & & f_{1n} \\ g_2 & f_{21} & f_{22} & & f_{2n} \\ \cdots & & & & \\ g_n & f_{n1} & f_{n2} & \cdots & f_{nn} \end{bmatrix}$$

# Quasi-concavity

- If the constraint, g, is linear;

  $g(x_1,\ldots,x_n)=c-b_1x_1-b_2x_2-\ldots-b_nx_n=0$

  - First-order conditions for a maximum: $f_i=\lambda b_i$ ; $i=1,\ldots,n$

  – Quasi-concave function

  - The bordered Hessian $H_b$ and the matrix H' have the same leading principal minors except for a (positive) constant of proportionality

    – H' follows the same sign conventions as $H_b$

    » (-1)H' must be negative definite

# Quasi-concavity

- The matrix H'

$$H' = \begin{bmatrix} 0 & f_1 & f_2 & \cdots & f_n \\ f_1 & f_{11} & f_{12} & & f_{1n} \\ f_2 & f_{21} & f_{22} & & f_{2n} \\ \cdots & & & & \\ f_n & f_{n1} & f_{n2} & \cdots & f_{nn} \end{bmatrix}$$