MISCELLANEOUS EXAMPLES (WITH SOLUTIONS)

1. The following table shows the daily income of a random sample of 40 workers of the QMB Manufacturing Company Limited:

Daily income, $	Number of workers
210 and under 220	2
220 and under 230	5
230 and under 240	4
240 and under 250	8
250 and under 260	10
260 and under 270	5
270 and under 280	4
280 and under 290	2

    (a)   Calculate the mean and the standard deviation of the daily income.

    (b)   Determine the median from the above table.

    (c)   Discuss the skewness of the above frequency distribution from results obtained from (a) and (b).

    (d)   Estimate, from the above table, the percentage of workers whose daily income is at least $242 but less than $280.

    (e)   Information relating to the workers employed by a rival firm gives the mean monthly income as $7,000 with a standard deviation $400. Are they, or the QMB workers, more variable with respect to income?

2. A random sample of the accounts of 50 shipping companies gave the following frequency distribution of annual profit.

Annual Profit ($million)	Number of Companies
−5 - under 0	2
0 - under 5	2
5 - under 10	4
10 - under 15	8
15 - under 20	11
20 - under 25	13
25 - under 30	6
30 - under 35	4

    (a)   Calculate the sample mean, median and standard deviation.

    (b)   Construct the cumulative frequency curve.

(c)   Using the result of part (b) or otherwise, estimate the annual profit exceeded by 30% of the shipping companies.

3.   A survey of customer purchasing behaviour, carried out by a supermarket, reveals the following grouped frequency distribution:

Frequency of purchasing (times per annum)	Number of customers
1 – 5	150
6 – 8	287
9 – 11	337
12 – 14	403
15 – 17	393
18 – 20	305
21 – 25	125
Total	2000

(a)   Draw a histogram of the above data.

(b)   Calculate the mean, mode, median, first and third quartiles, standard deviation, skewness and coefficient of variation of the above data.

(c)   Describe the shape of the distribution of customer purchasing behaviour in this supermarket based on part (b).

4.   A construction company has recently built a hotel. The probability that the new hotel will get an award for its design is 0.28, and the probability that it will get an award for the efficient use of materials is 0.13. If the probability that it will get at least one award is 0.36, what is the probability that it will get both awards?

5.   Mr. Chan is a mortgage lending officer at a local bank. Examining his loan files, he classifies 80% of the mortgage loans made by him as good and the remaining 20% as bad. Among the good loans, 70% are long-term and the remaining 30% are short-term. Among the bad loans, 20% are long-term and the remaining 80% are short-term. Suppose that a loan applicant applies for a long-term loan, what is the probability that the loan will turn out to be a good one?

6.   In a group of 16 students, 6 are enrolled in a course in management, 7 are enrolled in a course in accountancy, and 3 are enrolled in a course in marketing. If 4 students are selected at random to form a committee, find the probability that there is at least one student from each course.

7.   A company is planning to introduce a new product. The marketing department estimates that the sales volume is approximately normally distributed with a mean of 10,000 units and a standard deviation of 2,000 units.

   (a)   Estimate the probability that the sales volume will lie between 7,000 and 13,000 units.

   (b)   Determine the probability that the company will at least break-even given that:

            Selling price: $20 per unit
            Variable costs: $16 per unit
            Fixed costs: $30,000

8.   Suppose a box of 20 ball pens contains 1 defective ball pen. Three ball pens are now randomly chosen from the box for testing. What is the probability that the defective ball pen is among the 3 chosen for testing?

9.   Suppose there are 20 boxes of 10 ball pens. Among them, 15 boxes contain 2 defective ball pens each and 5 boxes contain 1 defective ball pen each. Now 1 box is randomly selected and from it 3 ball pens are chosen randomly for testing. What is the probability that 1 defective ball pen is found among the 3 chosen for testing.

10.   The average life of a certain type of small motor is 12 years, with a standard deviation of 3 years. The lives of the motors follow a normal distribution. Find the probability that the life of a randomly selected motor is between 10 years to 15 years.

11. Parts leaving an assembly line are examined by two inspectors. Each inspector detects 85% of the defective parts, and 80% of the defective parts are detected by both inspectors.

    (a) What is the probability that a defective part not detected by the first inspector will be detected by the second inspector?

    (b) What is the probability that a defective part will be detected by a least one of the inspectors?

    (c) A batch of output contains three defective parts. What is the probability that both inspectors will fail to detect at least one of these? Assume that detection of one defective part is independent of detection of another.

12. A restaurant has 50 seats available for those customers who make reservation for lunch by telephone. Experience indicates that 15 percent of those who make a reservation will fail to turn up. Suppose that the manager of the restaurant accepts 55 reservations for lunch one day. Find the probability that the restaurant will be able to accommodate all the customers who turn up on that day.

13. In a lengthy typed manuscript, it is discovered that 14% of the pages contain no typing error. Assuming that the number of errors per page is a random variable having a Poisson distribution, find the percentage of pages each of which has exactly one error.

14. On average a secretary makes 2 errors per page of her typing. What are the probabilities that in her next 2 pages of typing she

    (a) makes less than 3 errors?

    (b) makes more than 5 errors?

15. 42% of employees in a corporation were in favour of a modified health care plan, and 22% of the corporation's employees favoured a proposal to change the work schedule. 34% of those favouring the health plan modification favoured the work schedule change.

    (a) What is the probability that a randomly selected employee is in favour of both the modified health care plan and the changed work schedule?

    (b) What is the probability that a randomly chosen employee is in favour of at least one of the two changes?

    (c) What is the probability that an employee favouring the work schedule change also favours the modified health plan?

16. An investment portfolio contains stocks of a large number of corporations. Over the last year the rates of return on these corporate stocks follows a normal distribution, with mean 12.2% and standard deviation 7.2%.

    (a) For what proportion of these corporations was the rate of return higher than 20%?
    (b) For what proportion of these corporations was the rate of return negative?

17. Frequently, human populations are surveyed by mail questionnaires, and investigators are eager to obtain as high a response rate as possible. A questionnaire, printed as a single sheet, front and back, was sent to a random sample of 220 households, of which 36% responded. The same questionnaire, printed on two sheets, front only, was sent to an independent random sample of 220 households, and the achieved response rate was 30%. Find a 95% confidence interval for the difference between the two population proportions.

18. In a random sample of 225 purchasers of paper tissues, 67 sample members indicated cheapness as the major reason for brand selection. Construct a 95% confidence interval to estimate the population proportion.

19. A printer is interested in examining the relationship between the number of printing errors and the type size used. He selects 3 different books recently printed by his company, each using a different type size. From each book he records the number of pages with printing errors and the number of error-free pages. The results are shown in the table below. Do the data indicate at 0.05 level of significance a dependence between type size and printing errors?

	Type Size		
	A	B	C
Pages with Errors	23	17	41
Pages without Errors	241	183	210

20. (a) In a study of short-term absenteeism from work of ex-smokers, a random sample of 34 recent ex-smokers found a mean absenteeism of 2.21 days per month and a sample standard deviation of 2.21 days per month. For an independent random sample of 68 long-term ex-smokers, mean absenteeism was 1.47 days per month and the sample standard deviation was 1.69 days per month. Is there a significant difference between the mean monthly absenteeism of recent ex-smokers and long-term ex-smokers? Use a 0.05 level of significance.

    (b) Describe briefly how to set the level of significance of a significance test.

21. Of a random sample of 545 accountants engaged in preparing city operating budgets for use in planning and control, 117 indicated that estimates of cash flow were the most difficult element of the budget to derive.

    (a) Test the null hypothesis that at least 25% of all accountants find cash flow the most difficult estimates to derive at 5% level of significance.

    (b) What is the probability that the null hypothesis would be rejected at 5% level of significance if the true percentage of those finding cash flow estimates most difficult was 28%?

22. The following table below presents data for a random sample of $n = 8$ students in regard to the hours of study outside of class during a 3-week period and the grade earned on an examination at the end of first semester for a class in statistics.

Student	Hours of study	Examination grade
1	20	64
2	16	61
3	34	84
4	23	70
5	27	88
6	32	92
7	18	72
8	22	77

    (a) Determine the regression equation for estimating the examination grade given the hours of study.

    (b) Use the regression equation to estimate the examination grade for a student who devotes 30 hours to study of the course materials.

    (c) Compute the correlation coefficient and interpret this number.

    (d) How good is the regression equation in the light of data surveyed?

23. The following table shows the Price/Earning (P/E) ratios and Research & Development Expenditure/Sales (R/S) ratios of ten companies in the same industry.

Company	P/E Ratio (y)	R/S Ratio (x)
1	5.6	0.028
2	8.1	0.092
3	6.0	0.020
4	10.0	0.038
5	8.5	0.045
6	13.2	0.053
7	11.5	0.075
8	7.0	0.062
9	11.8	0.058
10	15.0	0.089

(a) By the method of least squares find the linear regression line of $\hat{y} = a + bx$.

(b) Interpret the estimated slope, b, of the linear regression line.

(c) Estimate the P/E ratio for a company with an R/S ratio equal to 0.06. Discuss the reliability of this estimate.

24. A marketing analyst for a major shoe manufacturer would like to predict the long-term ability to absorb shock after repeated impact test (ltimp), using the forefoot shock-absorbing capability (Foreimp) and the change in impact properties over time (Midsole). A random sample of 15 types of currently manufactured running shoes was selected for testing. The following outputs are produced using a statistical package:

ANOVA table

Source	df	SS	MS	F	Sig.
Regression	2	12.61020	6.30510	97.69	0.0001
Residual	12	0.77453	0.06454		
Total	14	13.38473			

Variable	Coefficients	Standard error	t	Sig.
Intercept	-0.02686	0.06905	-0.39	
Foreimp	0.79116	0.06295	12.57	0.0000
Midsole	0.60484	0.07174	8.43	0.0000

(a) Write down the least-squares regression equation using the above outputs.

(b) Interpret the meaning of the slopes in this problem.
(c) Compute the coefficient of multiple determination R^2 and interpret its meaning.

25. The president of a company that manufactures drywall wants to analyze the factors that affect demand for his product. Drywall is used to construct walls in houses and offices. Consequently, the president decides to develop a regression model in which the dependent variable is monthly sales of drywall in hundreds of 4×8 sheets (Drywall) and the independent variables are number of building permits issued (Permits), five-year mortgage rates in percentages (Mortgage), vacancy rate in apartments in percentages (A-vacant) and vacancy rate in office buildings in percentage (O-vacant). Monthly observations from the past two years are used and the following outputs are obtained.

ANOVA table

Source	df	SS	MS	F	Sig.
Regression	4	256793	64198	39.86	0.000
Residual	19	30602	1611		
Total	23	287395			

Variable	Coefficients	Standard error	t	Sig.
Intercept	-111.8	134.3	-0.83	0.416
Permits	4.7631	0.3950	12.06	0.000
Mortgage	16.99	15.16	1.12	0.276
A-vacant	-10.528	6.394	-1.65	0.116
O-vacant	1.308	2.791	0.47	0.645

(a) Test for the significance of the overall regression equation at 5% level of significance.

(b) Interpret each of the regression coefficients.

(c) Test whether each of the independent variables is linearly related to drywall demand at 5% level of significance.