ALLAMA IQBAL OPEN UNIVERSITY



AIOU LIBRARY

Name:

Registration No.:

Course Code: 1430

Program: BA/B. Com/AD

Semester: Autumn 2026

ALLAMA IQBAL OPEN UNIVERSITY, ISLAMABAD

Course: Business Statistics (1430) • Level: BA/B. Com/AD

Semester: Autumn 2026

Assignment No. 2 (Units 5–7)

Total Marks: 100 • Pass Marks BA: 40 • Pass Marks (AD): 50

SOLVED ASSIGNMENT GUIDE

(For Study & Reference)

Note: This guide is prepared for study and reference purposes. Students are advised to read the relevant units of the course book, understand the concepts and methods, and write the assignment in their own words.

Question 1: Hypothesis Testing and Confidence Intervals; A Test on the Mean P/E Ratio (10 + 10 Marks)

Q. 1 (a) How does hypothesis testing differ from confidence intervals? (b) From 1980 until 1985, the mean price/earnings (P/E) ratio of the approximately 1800 stocks was 14.35 and the standard deviation was 9.73. In a sample of 30 randomly chosen stocks, the mean P/E ratio in 1986 was 11.77. Does this sample present sufficient evidence to conclude that in 1986 the mean P/E ratio for stocks had changed from its earlier value at the 0.05 level of significance?

Introduction

Hypothesis testing and confidence interval estimation are the two great branches of statistical inference, and students often confuse them because both are built from the same ingredients: a sample statistic, its standard error, and the sampling distribution. Yet they answer different questions. A hypothesis test asks a yes-or-no question about a specific claimed value of a parameter; a confidence interval asks a "how large, and with what precision" question and answers it with a range of plausible values. Part (a) draws the distinction systematically, and part (b) applies a hypothesis test to the price/earnings ratios of stocks — and then confirms the

verdict with the corresponding confidence interval, showing the two methods working together exactly as they do in professional practice.

Part (a): How Hypothesis Testing Differs from Confidence Intervals

- Purpose and question asked: a hypothesis test begins with a definite claim about the parameter, stated as the null hypothesis (for example "the mean P/E ratio is still 14.35"), and asks whether the sample evidence is strong enough to reject that claim. A confidence interval begins with no claim at all; it asks which values of the parameter are consistent with the sample evidence.
- Form of the answer: a test ends in a decision — reject H_0 or do not reject H_0 — at a stated significance level. An interval ends in a range of values, such as "the mean lies between 8.3 and 15.3 with 95 per cent confidence", which carries information about both the size of the effect and the precision of the estimate.
- What is fixed in advance: in testing, the hypothesised value is fixed before the data are seen, and the standard error is often computed under that hypothesis. In interval estimation nothing is hypothesised; the interval is centred on the sample statistic and built entirely from the data.
- Role of the significance level: a test is governed by α, the accepted probability of a Type I error (rejecting a true null hypothesis). An interval is governed by the confidence level (1 − α): a 95 per cent interval is constructed by a method that, in repeated sampling, captures the true parameter 95 per cent of the time.

- One-sided and two-sided forms: tests may be one-sided when only one direction matters. Intervals are normally two-sided, although one-sided confidence bounds exist; a two-sided interval corresponds exactly to a two-sided test at the matching level.
- The duality between them: the two procedures are mathematically two faces of one idea. A two-sided test at level α rejects a hypothesised value precisely when that value lies outside the (1 − α) confidence interval. The interval therefore contains the verdicts of the tests of every possible hypothesised value at once, which is why many statisticians regard the interval as the more informative report.
- What each conceals: a bare test decision conceals magnitude — a rejection may reflect a trivially small but precisely measured difference. An interval displays magnitude and precision together, but does not by itself announce a decision about any particular claimed value until the reader checks whether that value lies inside it.

Good reporting therefore uses both: the test for the formal decision about the claim in dispute, and the interval to show the range of values the data support, so that the reader can judge not only whether an effect exists but whether it is large enough to matter commercially.

Part (b): Has the Mean P/E Ratio Changed in 1986?

The established position, from the approximately 1,800 stocks over 1980–1985, is a population mean of $\mu = 14.35$ with a known population standard deviation $\sigma = 9.73$. The 1986 sample of $n = 30$ stocks gives a sample mean $\bar{x} = 11.77$. Because σ is known, the z-test is used. Step 1 — Hypotheses: the question asks whether the mean has changed, in either direction, so H_0: $\mu = 14.35$ and H_1: $\mu \neq 14.35$. Step 2 — Significance level: $\alpha = 0.05$, as given. Step 3 — Decision rule: for a two-sided z-test at 0.05 the critical values are ±1.96; reject H_0 if the computed z is below

−1.96 or above +1.96. Step 4 — Computation: the standard error of the mean is σ/√n = 9.73/√30 = 9.73/5.477 = 1.78, and the standardised statistic is z = (11.77 − 14.35)/1.78 = −2.58/1.78 = −1.45. Step 5 — Decision: −1.45 does not cross −1.96, and the two-sided p-value is about 0.146, well above 0.05, so we do not reject H_0.

The confidence-interval check tells the same story and adds information. The 95 per cent interval for the 1986 mean is $\bar{x} \pm 1.96 \times 1.78 = 11.77 \pm 3.48$, that is, from 8.29 to 15.25. The old value 14.35 lies comfortably inside this interval, confirming the test decision: a mean as low as 11.77 in a sample of only 30 stocks is entirely compatible with an unchanged population mean of 14.35, because P/E ratios are highly variable (σ = 9.73) and a sample of 30 leaves a wide margin of sampling error. The practical conclusion for an investor or analyst is that the 1986 sample does not present sufficient evidence, at the 5 per cent level, that the mean P/E ratio had changed from its 1980–1985 value; the observed dip may well be sampling noise, and a much larger sample would be needed to establish a genuine fall of this size.

Conclusion

To conclude, hypothesis testing and confidence intervals differ in the question they ask, the form of their answer, and the role of the hypothesised value, yet they are dual procedures that must agree when applied consistently. In the P/E problem they do agree: the test statistic z = −1.45 fails to reach the critical value 1.96 at the 0.05 level, and the 95 per cent confidence interval (8.29, 15.25) contains the earlier mean of 14.35. There is no sufficient evidence that the mean P/E ratio changed in 1986.

Question 2: Tests for the Difference Between Two Means; The Business and A&S GPA Comparison (10 + 10 Marks)

Q. 2 (a) Write the procedures for tests for the difference between means when population variance is known and unknown. (b) Connie Rodrigues, the Dean of Students at Midstate College, is wondering about grade distributions at the school. She has heard grumblings that the GPAs in the Business School are about 0.25 lower than those in the College of Arts and Sciences (A&S). A quick random sampling produced the following GPAs. Business: 2.86, 2.77, 3.18, 2.80, 3.14, 2.87, 3.19, 3.24, 2.91, 3.00, 2.83. A&S: 3.35, 3.32, 3.36, 3.63, 3.41, 3.37, 3.45, 3.43, 3.44, 3.17, 3.26. Do these data indicate there is a factual basis for the grumblings? State and test appropriate hypotheses at α = 0.02.

Introduction

Comparing two groups — two schools, two machines, two training methods, two hospitals — is the commonest task in applied statistics. The logic is always the same: the difference between the two sample means is compared with the size of difference that sampling variation alone would typically produce, measured by the standard error of the difference. What changes from case to case is how that standard error is obtained, and that depends on whether the population variances are known or must be estimated from the samples. Part (a) gives both procedures in full; part (b) then applies the unknown-variance procedure to the Dean's GPA data, in which the claim to be tested is unusually specific: not merely that two means differ, but that they differ by about 0.25 grade points.

Part (a): Procedures — Variance Known and Variance Unknown

Case 1: Population Variances Known (the Two-Sample z-Test)

Step 1 — State the hypotheses in terms of the difference $\mu_1 - \mu_2$: for a two-sided test, H_0: $\mu_1 - \mu_2 = D$ (often $D = 0$) against H_1: $\mu_1 - \mu_2 \neq D$, with one-sided alternatives used when a direction is specified in advance. Step 2 — Fix the significance level α and take independent random samples of sizes n_1 and n_2, assumed drawn from normal populations (or large enough for the central limit theorem to apply). Step 3 — Compute the standardised statistic $z = [(\bar{x}_1 - \bar{x}_2) - D] / \sqrt{(\sigma_1^2/n_1 + \sigma_2^2/n_2)}$, in which the denominator is the standard error of the difference built from the known variances σ_1^2 and σ_2^2. Step 4 — Compare z with the standard normal critical values (±1.96 at 0.05, ±2.58 at 0.01, two-sided) and reject H_0 if z falls in the rejection region. Because the variances are known exactly, no estimation error enters, and the normal distribution is the exact reference distribution for large or normal samples.

Case 2: Population Variances Unknown (the Two-Sample t-Test)

In practice the population variances are almost never known, so they are estimated by the sample variances s_1^2 and s_2^2, and the reference distribution becomes Student's t. Where the two population variances may reasonably be assumed equal — the usual assumption for similarly constituted groups, checked by comparing the sample variances — the estimates are pooled into a single pooled variance $s_p^2 = [(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2] / (n_1 + n_2 - 2)$, and the statistic is $t = [(\bar{x}_1 - \bar{x}_2) - D] / \sqrt{[s_p^2(1/n_1 + 1/n_2)]}$, which follows the t-distribution with $n_1 + n_2 - 2$ degrees of freedom. Where the equal-variance assumption is not defensible, Welch's form is used instead: the denominator becomes $\sqrt{(s_1^2/n_1 + s_2^2/n_2)}$ and the degrees of freedom are reduced by Welch's approximation formula. The remaining steps mirror the z-procedure: state hypotheses,

fix α, read the t critical value for the appropriate degrees of freedom and tail structure, compute, decide, and interpret in context. The price of not knowing the variances is visible in the t-table: with small samples the critical values are larger than the corresponding z values, honestly reflecting the extra uncertainty.

Part (b): Are Business GPAs Really About 0.25 Lower?

The samples each contain n = 11 GPAs. From the Business sample, $\bar{x}_1$ (Business) = 2.98 with s_1 = 0.18 (more precisely, mean 2.981, s = 0.176), from the A&S sample, $\bar{x}_2$ (A&S) = 3.38 with s_2 = 0.12 (mean 3.381, s = 0.117). The observed difference is therefore $\bar{x}_2 - \bar{x}_1 = 3.381 - 2.981 = 0.40$ grade points — larger than the rumoured 0.25, but based on small samples, so it must be tested rather than taken at face value. The two sample variances (0.031 and 0.014) are of the same order, so the pooled-variance t-procedure of part (a) is appropriate.

Table 1: Summary statistics for the two GPA samples

School	n	Mean GPA	Sample standard deviation
Business	11	2.981	0.176
Arts & Sciences (A&S)	11	3.381	0.117
Difference (A&S − Business)	—	0.400	—

The pooled variance is $s_p^2 = [10 \times 0.0308 + 10 \times 0.0138] / 20 = 0.0223$, so $s_p = 0.149$, and the standard error of the difference is $\sqrt{[0.0223 \times (1/11 + 1/11)]} = \sqrt{(0.00405)} = 0.064$. The grumbling makes a specific claim, so the appropriate hypotheses are H_0: $\mu_2 - \mu_1 = 0.25$ against the two-sided alternative H_1: $\mu_2 - \mu_1 \neq 0.25$, tested at $\alpha = 0.02$. With 11 + 11 − 2 = 20 degrees of freedom, the two-sided critical value at the 0.02 level is t = ±2.528. The test statistic is t = (0.40

− 0.25) / 0.064 = 0.15 / 0.064 = 2.36. Since 2.36 is less than 2.528, we do not reject H_0 at the 2 per cent level: the observed difference of 0.40 is statistically consistent with a true difference of 0.25, given samples of only eleven students each.

It is equally important to record what the data do establish beyond doubt. Tested against the hypothesis of no difference at all, H_0: $\mu_2 - \mu_1 = 0$, the statistic is $t = 0.40 / 0.064 = 6.28$, which far exceeds any conventional critical value; the GPA gap between the two schools is unquestionably real and not a sampling accident. Combining the two findings gives the Dean the complete picture: yes, the grumblings have a firm factual basis — Business School GPAs are genuinely lower, by an estimated 0.40 grade points in these samples — and the size of the gap is consistent, at the strict 0.02 level, with the "about 0.25" that is being claimed, the word "about" being essential. If the college wishes to pin the gap down more precisely for policy purposes — for example before reviewing grading standards — it should sample substantially more students from each school, since the standard error, and with it the width of the plausible range for the true difference, shrinks only with the square root of the sample size.

Conclusion

To conclude, the difference between two means is tested by a z-statistic when the population variances are known and by a t-statistic — pooled when the variances may be assumed equal, Welch's form when they may not — when they must be estimated. Applied to Midstate College, the procedure shows a real and highly significant GPA difference between the schools ($t = 6.28$ against zero), estimated at 0.40 points, which does not differ significantly from the claimed 0.25 at the 0.02 level ($t = 2.36$ against a critical value of 2.528). The grumblings, in short, are factually justified.

Question 3: Testing Differences Between Means with Dependent Samples; The Sales-Training Question (10 + 10 Marks)

Q. 3 (a) Explain the process for testing differences between means with dependent samples. (b) A credit-insurance organization has developed a new high-tech method of training new sales personnel. The company sampled 16 employees who were trained the original way and found average daily sales to be $688 and the sample standard deviation was $32.63. They also sampled 11 employees who were trained using the new method and found average daily sales to be $706 and the sample standard deviation was $24.84. At α = 0.05, can the company conclude that average daily sales have increased under the new plan?

Introduction

Not every comparison of two means involves two independent groups of people. Very often the same subjects are measured twice — before and after a training course, before and after a medicine, on a left hand and a right hand — or subjects are deliberately matched in pairs. Such data are called dependent, or paired, samples, and they need a different test from the one used for independent groups, because the pairing itself removes a large part of the variation that has nothing to do with the treatment. Part (a) explains that paired process carefully. Part (b) then examines the sales-training data printed in the question — and notes an important feature of those data: the two groups consist of different employees in different numbers (16 trained the original way, 11 trained by the new method), so they are in fact independent samples, and the correct calculation is the independent-samples pooled t-test of Question 2. That calculation is carried out in full, and the difference the pairing would have made is explained.

Part (a): The Paired (Dependent-Samples) Test — Process

Step 1 — Recognise the design. Samples are dependent when each observation in one set is linked to a specific observation in the other: the same person measured twice, twins, matched machines, the same shop in two periods, the same student's marks in two examinations. The link, not the subject matter, is what makes the samples dependent. Step 2 — Reduce the pairs to differences. For each of the n pairs form the difference d = (second measurement − first measurement), keeping the direction consistent throughout. The two-sample problem has now become a one-sample problem about a single new variable, the difference. Step 3 — State the hypotheses about the mean difference μ_d: H_0: $\mu_d = 0$ (no average change) against a one-sided alternative ($\mu_d > 0$ for an improvement, $\mu_d < 0$ for a decline) or a two-sided alternative ($\mu_d \neq 0$), according to the question asked, and fix α. Step 4 — Compute the mean difference $\bar{d}$ and the standard deviation of the differences s_d from the n differences, and form the statistic $t = \bar{d} / (s_d / \sqrt{n})$, which, if the differences are approximately normally distributed, follows the t-distribution with n − 1 degrees of freedom. Step 5 — Compare with the critical t value for n − 1 degrees of freedom, decide, and interpret the conclusion as a statement about the average change within the pairs.

The power of the design lies in what the differencing removes. Two employees may differ enormously in natural selling ability, experience and territory, and in an independent-samples test all of that person-to-person variation inflates the standard error and can hide a genuine training effect. In a paired test, each person acts as his or her own control: the differences strip out the personal level entirely, and only the change produced by the treatment remains, so a much smaller true effect becomes detectable with the same number of subjects. The costs and

cautions are also real: pairing requires the double measurement to be feasible and free of carry-over effects (a person cannot be "untrained" and retrained), the test assumes the differences, not the original measurements, are roughly normally distributed, and careless or forced matching on irrelevant characteristics wastes degrees of freedom without reducing variation. Used where a natural pairing exists, however, the dependent-samples test is one of the most efficient designs in all of applied statistics.

Part (b): Have Daily Sales Increased Under the New Training Plan?

Read carefully, the data of part (b) describe two separate groups of different employees, 16 trained the original way and 11 trained by the new method; no employee appears in both groups, and the group sizes are unequal, which pairing would not permit. The honest and correct procedure is therefore the independent-samples pooled t-test (the sample standard deviations, 32.63 and 24.84, are similar enough for pooling to be reasonable). The claim to be tested is directional — sales have increased — so the hypotheses are H_0: $\mu_2 \le \mu_1$ (new method no better than original) against H_1: $\mu_2 > \mu_1$, tested one-sided at $\alpha = 0.05$, where group 1 is the original method ($n_1 = 16$, $\bar{x}_1 = \$688$, $s_1 = \$32.63$) and group 2 is the new method ($n_2 = 11$, $\bar{x}_2 = \$706$, $s_2 = \$24.84$).

The pooled variance is $s_p^2 = [(16 - 1)(32.63)^2 + (11 - 1)(24.84)^2] / (16 + 11 - 2) = [15 \times 1064.72 + 10 \times 617.03] / 25 = (15{,}970.8 + 6{,}170.3) / 25 = 22{,}141.1 / 25 = 885.64$, giving $s_p = \$29.76$. The standard error of the difference is $\sqrt{[885.64 \times (1/16 + 1/11)]} = \sqrt{(885.64 \times 0.1534)} = \sqrt{135.86} = \$11.66$. The test statistic is therefore $t = (706 - 688) / 11.66 = 18 / 11.66 = 1.54$, on 25 degrees of freedom. The one-sided critical value at $\alpha = 0.05$ with 25 degrees of freedom is 1.708. Since

1.54 < 1.708, we do not reject H_0. (Welch's unpooled version gives t = 1.63 against a similar critical value, so the conclusion does not depend on the pooling assumption.)

The company therefore cannot conclude, at the 5 per cent level, that average daily sales have increased under the new plan. The new-method group did sell $18 a day more on average, but with day-to-day standard deviations of around $25–$33 and samples of only 16 and 11, an $18 gap is well within what sampling variation alone produces; the one-sided p-value is about 0.07. This is precisely the situation in which the design of part (a) would have served the company better: had the same salespeople been measured before and after the new training, their personal selling levels would have cancelled out in the differences, the standard error would have been far smaller, and an $18 genuine improvement could very likely have been detected. As the study stands, the sensible managerial decision is not to abandon the new method — the point estimate favours it — but to run a larger, or preferably paired, trial before committing the whole sales force to an expensive high-tech programme on the strength of inconclusive evidence.

Conclusion

To conclude, dependent samples are tested by reducing each pair to a single difference and applying a one-sample t-test to those differences with n − 1 degrees of freedom — a design that removes between-subject variation and is correspondingly powerful. The training data printed in part (b), however, come from two independent groups, and the correct pooled t-test gives t = 1.54 against a critical value of 1.708 at α = 0.05 on 25 degrees of freedom. The increase in average daily sales under the new plan is not statistically established, and a larger or paired study is the proper next step.

Question 4: Regression of Overhead on Units Produced (10 + 5 + 5 Marks)

Q. 4 Cost accountants often estimate overhead based on the level of production. At the Standard Knitting Co., they have collected information on overhead expenses and units produced at different plants, and want to estimate a regression equation to predict future overhead. Overhead: 191, 170, 272, 155, 280, 173, 234, 110, 153, 178. Units: 40, 42, 53, 35, 56, 39, 48, 30, 37, 40. (a) Develop the regression equation for the cost accountants. (b) Predict overhead when 50 units are produced. (c) Calculate the standard error of the estimate.

Introduction

Overhead — rent, supervision, power, maintenance and other indirect costs — is the hardest element of product cost to plan, because it is not traced to individual units the way materials and direct labour are. Cost accountants therefore look for a statistical relationship between overhead and the level of activity, so that the overhead of a future period can be budgeted as soon as its planned production is known. Simple linear regression supplies exactly that relationship: a straight line, fitted by the method of least squares to past observations of the two variables, which can then be used for prediction and whose accuracy can be measured. This answer fits the line of overhead (the dependent variable, y) on units produced (the independent variable, x) for the ten plants of the Standard Knitting Co., uses it to predict overhead at 50 units, and measures the reliability of the fit by the standard error of the estimate.

Part (a): Developing the Regression Equation

The least-squares line $\hat{y} = a + bx$ is chosen to make the sum of the squared vertical deviations of the observed points from the line as small as possible; its slope is $b = \Sigma(x - \bar{x})(y - \bar{y}) / \Sigma(x - \bar{x})^2$ and its intercept is $a = \bar{y} - b\,\bar{x}$, which guarantees that the line passes through the centre of the data, the point $(\bar{x}, \bar{y})$. From the ten observations, $\Sigma x = 420$ so $\bar{x} = 42.0$ units, and $\Sigma y = 1{,}922$ so $\bar{y} = 192.2$. The deviation sums are $\Sigma(x - \bar{x})^2 = 588.0$ and $\Sigma(x - \bar{x})(y - \bar{y}) = 3{,}817.0$ (equivalently, from the raw sums, $\Sigma xy = 84{,}541$ and $\Sigma x^2 = 18{,}228$). Hence the slope is $b = 3{,}817.0 / 588.0 = 6.49$, and the intercept is $a = 192.2 - 6.49 \times 42.0 = 192.2 - 272.64 = -80.44$. The regression equation is therefore $\hat{y} = -80.44 + 6.49x$.

Table 2: Observations, fitted values and residuals for the overhead regression

Units (x)	Overhead (y)	Predicted $\hat{y}$	Residual $(y - \hat{y})$	$(y - \hat{y})^2$
40	191	179.22	11.78	138.77
42	170	192.20	−22.20	492.84
53	272	263.61	8.39	70.39
35	155	146.76	8.24	67.90
56	280	283.08	−3.08	9.49
39	173	172.73	0.27	0.07
48	234	231.15	2.85	8.12
30	116	114.30	1.70	2.89
37	153	159.74	−6.74	45.43
40	178	179.22	−1.22	1.49
Total: $\Sigma x = 420$	$\Sigma y = 1{,}922$	—	≈ 0	837.56

The two coefficients have direct managerial meanings. The slope, 6.49, is the estimated variable overhead rate: each additional unit produced is associated with about 6.49 additional units of overhead cost (in the money units of the data), so it functions exactly like the variable-cost-per-

unit of cost-volume-profit analysis. The intercept, −80.44, would formally represent fixed overhead at zero production, but here it must not be read literally: it is negative, which no real fixed cost can be, and zero production lies far outside the observed range of 30 to 56 units. The intercept is simply the mathematical anchor that positions the line correctly within the observed range; the negative value is a warning against extrapolating the equation to activity levels the data never covered. Within the observed range the line performs strongly, as part (c) quantifies, and the residuals in Table 2 show no systematic pattern — they are positive and negative in a balanced way and their sum is (apart from rounding) zero, as least squares guarantees.

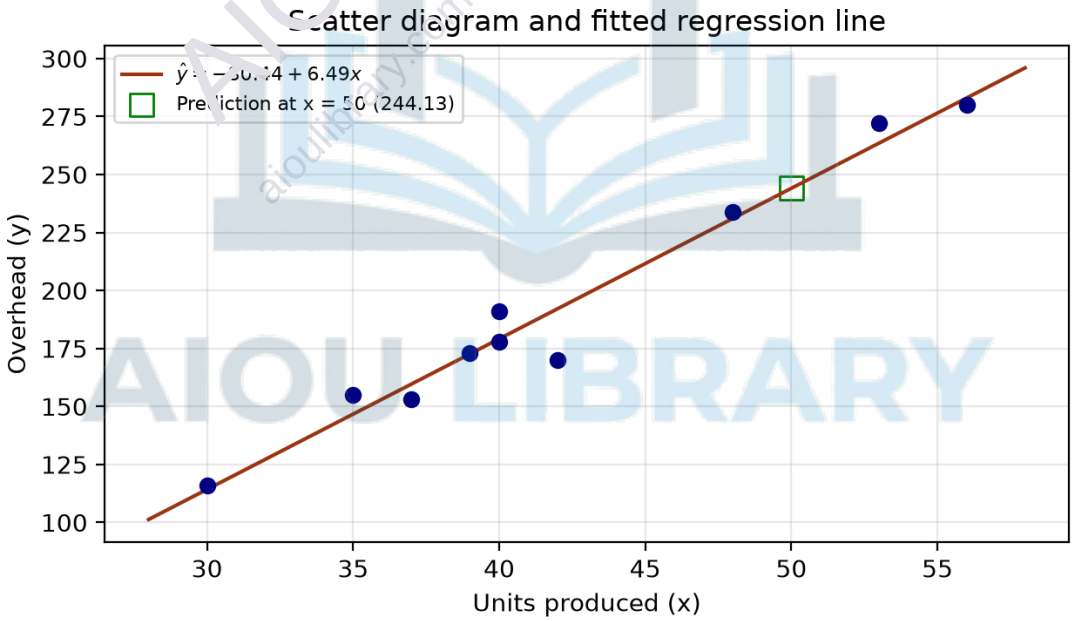


Figure 1: Scatter diagram of overhead against units produced, the fitted regression line, and the prediction at 50 units.

Part (b): Predicting Overhead When 50 Units Are Produced

Fifty units lies inside the observed range (30–56), so prediction by substitution is legitimate interpolation rather than risky extrapolation. Substituting x = 50 in the equation: ŷ = −80.44 +

6.49 × 50 = −80.44 + 324.57 = 244.13. The predicted overhead when 50 units are produced is therefore approximately 244 (in the money units of the data, for example $244 or Rs 244 as the company's accounts express them). This is the figure the cost accountants should build into the budget for a 50-unit period as the expected overhead, while remembering from part (c) that the actual outcome will typically differ from it by roughly the standard error of the estimate.

Part (c): The Standard Error of the Estimate

The standard error of the estimate measures the typical size of the prediction error — the average distance, in the units of y, between the observed overheads and the values fitted by the line. It is computed as $s_{px} = \sqrt{[\,\Sigma(y - \hat{y})^2 / (n - 2)\,]}$, where the divisor is n − 2 = 8 because two constants, a and b, were estimated from the data. From the last column of Table 2, $\Sigma(y - \hat{y})^2 = 837.56$, so $s_{px} = \sqrt{(837.56 / 8)} = \sqrt{104.69} = 10.23$. The standard error of the estimate is therefore about 10.23 overhead units. In practical terms, predictions from this equation are typically accurate to within roughly ±10, and (the errors being approximately normal) about two-thirds of actual overheads will fall within ±10.23 of the predicted value and roughly 95 per cent within about ±20.5. Judged against an average overhead of 192.2, an error scale of about 10 — roughly 5 per cent — is small, and the correlation confirms the strength of the relationship: r = 3,817 / √(588 × 25,615.6) = 0.98, so $r^2 = 0.967$, meaning that about 96.7 per cent of the variation in overhead between plants is explained by the variation in units produced.

Conclusion

To conclude, the least-squares regression of overhead on units produced for the Standard Knitting Co. is $\hat{y} = -80.44 + 6.49x$, with a slope that reads as a variable overhead rate of about 6.49 per unit and an intercept that is a mathematical anchor only, not a real fixed cost. At a

production level of 50 units the predicted overhead is 244.13, and the standard error of the estimate, 10.23, with r = 0.98, shows that the equation predicts closely enough to be genuinely useful for budgeting — provided it is used within, or very near, the observed production range of 30 to 56 units.

Question 5: Laspeyres, Paasche and Fixed-Weight Index Numbers; Constructing a Wholesale Price Index (10 + 10 Marks)

Q. 5 (a) Differentiate between Laspeyres, Paasche and fixed-weight aggregates methods. (b) Write the main steps in the construction of wholesale price index numbers.

Introduction

An index number is a single figure that summarises the change in a whole group of related variables — most commonly prices — between two periods. Because the items in the group change by different amounts, they must be combined, and the way they are weighted determines the character of the index. The three classical weighting schemes, those of Laspeyres, Paasche and the fixed-weight aggregates method, answer the same question with three different baskets of goods, and each has distinct merits, costs and biases. Part (a) distinguishes them carefully. Part (b) then shows how these ideas are put to work in the most commercially important of all price indices, the wholesale price index, by setting out the main steps in its construction.

Part (a): Laspeyres, Paasche and Fixed-Weight Aggregates Compared

All three methods are weighted aggregative indices of the general form: (weighted sum of current-period prices ÷ weighted sum of base-period prices) × 100. They differ only in the quantities used as weights. The Laspeyres index, proposed by Étienne Laspeyres in 1864, uses

base-period quantities q_o as weights throughout its life: $L = [\Sigma p_1 q_o / \Sigma p_o q_o] \times 100$. It answers the question, "how much more (or less) would the original base-period basket cost at today's prices?" The Paasche index, proposed by Hermann Paasche in 1874, uses current-period quantities q_1 as weights: $P = [\Sigma p_1 q_1 / \Sigma p_o q_1] \times 100$. It answers the mirror-image question, "how does the cost of today's basket compare with what that same basket would have cost at base-period prices?" The fixed-weight aggregates method uses quantities q^* drawn from some fixed, typical period — neither the base nor the current period, but often an average of several periods — held constant for the whole series: $F = [\Sigma p_1 q^* / \Sigma p_o q^*] \times 100$.

- Weights: Laspeyres — base-period quantities; Paasche — current-period quantities, changing every period; fixed-weight — quantities of a fixed typical period (or multi-period average), constant across the series.
- Data burden and cost: Laspeyres needs quantity data only once, for the base period, so it is cheap and quick to compute and publish — the practical reason most official consumer price indices are of Laspeyres type. Paasche needs fresh quantity data for every period, which is slow and expensive, and its figures are often revised when the quantities are finalised. Fixed-weight needs the quantity data of its chosen period only.
- Comparability over time: Laspeyres compares every period against the same basket, so its values are directly comparable across the whole series. Because Paasche changes its basket every period, part of any movement in a Paasche series reflects changing quantities rather than changing prices, weakening comparability. Fixed-weight preserves comparability, like Laspeyres, because its basket too is constant.

- Direction of bias: Laspeyres tends to overstate a rise in the cost of living (an upward bias), because buyers substitute away from goods whose prices rise most, while the fixed base basket keeps the old, now-expensive pattern; in technical terms it ignores substitution. Paasche tends to understate it (a downward bias), because its current basket already embodies that substitution. The two indices therefore usually bracket the truth, which is why Fisher's "ideal" index is defined as the geometric mean of the two, √(L × P). A fixed-weight index with a typical, averaged basket generally falls between the two and avoids the extremes of either single period.
- Interpretation: Laspeyres measures the changing cost of a fixed standard of the base period; Paasche measures it for the standard actually consumed now; fixed-weight measures it for a representative standard chosen to be typical of the period the index is meant to serve.

A further practical point completes the comparison: every fixed-basket index grows stale. Tastes change, products disappear and new ones appear, so whichever of the three schemes is adopted, the basket and weights must be revised periodically — typically every five to ten years — and the old and new series linked, or the index gradually ceases to describe the economy it claims to measure.

Part (b): Main Steps in Constructing a Wholesale Price Index

Step 1 — Define the purpose and scope. Decide precisely what the index is to measure — the general movement of prices at the wholesale (first bulk transaction) stage — and fix its coverage: the sectors to be included, such as food, fuel and power, agricultural raw materials, and manufactured goods, and the geographical coverage of the country or region. Step 2 — Select the commodities. Choose a representative basket of wholesale commodities that covers

the important items of wholesale trade without becoming unmanageable; each commodity should have a significant volume of trade, a specifiable standard quality, and a price that can be collected regularly. Related items are arranged into groups and sub-groups (for example food articles, non-food articles, minerals, and manufactures) so that group indices can be published alongside the general index. Step 3 — Select the markets and price sources. Prices must be collected from the principal wholesale markets (mandis and trading centres) where the bulk of each commodity is actually traded, from identified and reliable reporting sources, on a fixed day and basis (for example the price of a stated standard grade, per stated unit, at the wholesale stage), so that the quotations are strictly comparable over time.

Step 4 — Choose the base period. The base should be a recent, economically normal period — free of war, drought, controls or abnormal speculation — and often a full year or an average of years is taken to smooth out seasonal and accidental fluctuations; its index is set to 100. Step 5 — Assign the weights. Each commodity is weighted according to its relative importance in total wholesale trade, normally the value of the quantity marketed or transacted in the base period (value weights); weights are distributed first among the major groups and then within each group among its commodities, so that a price change in a major commodity such as wheat, cotton or petroleum moves the index far more than an equal change in a minor item. Step 6 — Choose the formula and compute. The standard practice is a weighted aggregative (Laspeyres-type) formula, equivalently computed as the weighted arithmetic mean of price relatives: for each commodity the price relative (current price ÷ base price × 100) is calculated, multiplied by its weight, and the index is the sum of these products divided by the sum of the weights. Group

indices are computed first and then combined with group weights into the all-commodities index.

Step 7 — Publish, maintain and revise. The index is published at a regular frequency (weekly or monthly) with its group detail, so that users can see which groups are driving the general movement; quotations are continuously checked, missing prices are handled by stated rules (carrying forward or imputing from similar markets), and, at intervals of years, the whole structure — commodities, markets, weights and base — is revised and the new series is linked to the old so that long-run comparisons remain possible. In Pakistan this task is performed by the Pakistan Bureau of Statistics, whose Wholesale Price Index, built on exactly these principles, is watched by the State Bank, the finance ministry, businesses and index-linked contract writers as a leading indicator of inflationary pressure, because wholesale prices move before the retail prices of the consumer index.

Conclusion

To conclude, Laspeyres, Paasche and the fixed-weight aggregates methods are one family of weighted price indices distinguished solely by their weights — base-period, current-period and fixed typical-period quantities respectively — with predictable consequences: Laspeyres is cheap and comparable but biased upward, Paasche is current but data-hungry, hard to compare and biased downward, and the fixed-weight method stands between them. The wholesale price index puts these principles into practice through seven disciplined steps — defining scope, selecting commodities and markets, collecting comparable prices, choosing a normal base, weighting by value of trade, computing by the weighted formula, and publishing and periodically revising the series.