Financial Econometrics 2026
Shanghai Jiao Tong University
Summer School 2026
Quiz 1

Reading time: 10 minutes. **Writing time:** 60 minutes.
Total marks: 100. **Calculators are NOT permitted.** Every calculation is designed to be done by hand.
Instructions: Answer **all** questions. In Section A, circle the letter of the single best answer. In Sections B and C, show your working and write in clear sentences.

Structure:

- **Section A** — Multiple choice: 20 questions, 2.5 marks each (**50 marks**).
- **Section B** — One analytical problem in five parts (**30 marks**).
- **Section C** — One short interpretation (**20 marks**).

Useful facts. For a 95% confidence interval use 1.96 (for 90% use 1.64, for 99% use 2.58). $SE(\bar{Y}) = s_Y/\sqrt{n}$; $t = \dfrac{\text{estimate} - \text{null}}{SE}$; $\widehat{\beta}_1 = \dfrac{s_{XY}}{s_X^2} = r\,\dfrac{s_Y}{s_X}$; $\widehat{\beta}_0 = \bar{Y} - \widehat{\beta}_1\bar{X}$; $r = \dfrac{s_{XY}}{s_X s_Y}$; $R^2 = \dfrac{ESS}{TSS}$ (with one regressor, $R^2 = r^2$).

Section A: Multiple Choice (20 questions $\times$ 2.5 marks = 50 marks)

1. An analyst has the monthly return of the S&P 500 index for every month from January 1990 to December 2025. This data set is an example of:

(A) cross-sectional data.

(B) time-series data.

(C) panel (longitudinal) data.

(D) experimental data.

2. You lend a friend money. With probability 0.95 you are repaid \$110; with probability 0.05 the friend defaults and you receive \$0. The expected repayment is:

(A) \$110.

(B) \$105.

(C) \$104.50.

(D) \$100.

3. In finance, the standard deviation of an asset's returns is used as the standard measure of:

(A) its risk, also called its volatility.

(B) its expected return.

(C) its correlation with the market.

(D) its average price.

4. Monthly returns on a portfolio are approximately normal with mean 0% and standard deviation 5%. About 95% of monthly returns fall between:

(A) -5% and 5%.

(B) -2.5% and 2.5%.

(C) -15% and 15%.

(D) -10% and 10%.

5. Which statement about the *correlation* between two random variables is correct?

(A) A correlation of zero proves that the two variables are independent.

(B) It always lies between -1 and $+1$ and measures the strength of their *linear* relationship.

(C) It has the same units as X multiplied by Y.

(D) It can exceed $+1$ when the relationship is very strong.

6. An investor puts half of her money in each of two stocks that have the same volatility. Combining them lowers the portfolio's risk below that of either stock alone, provided the two returns are:

(A) less than perfectly positively correlated (correlation below $+1$).

(B) perfectly positively correlated (correlation equal to $+1$).

(C) equal in expected return.

(D) both riskier than the market.

7. Individual daily stock returns have fat tails and are clearly not normal. Yet we can still use the value 1.96 to build a confidence interval for an *average* return computed from many observations. This is justified by:

(A) the assumption that individual returns are normal.

(B) the law of large numbers, which makes the sample mean exactly equal to the population mean.

(C) the fact that 1.96 is valid for every distribution and every sample size.

(D) the central limit theorem: the sample average is approximately normal in large samples, whatever the shape of the underlying data.

8. The sample mean $\bar{Y} = \frac{1}{n}\sum_i Y_i$, seen as a rule applied to a random sample, is best described as:

(A) a fixed number, known before the data are drawn.

(B) equal to the population mean μ_Y.

(C) a random variable, because a different sample would give a different value.

(D) the same thing as the standard error.

9. A fund database includes only funds that are still operating today; funds that closed (usually after poor performance) were removed. Using it to estimate the average return of *all funds ever launched* gives an estimate that is:

(A) biased upward, and collecting more surviving funds will not remove the bias.

(B) unbiased, because the sample is large.

(C) biased downward.

(D) unbiased, provided returns are normally distributed.

10. The standard error of the sample mean, $SE(\bar{Y}) = s_Y/\sqrt{n}$, measures:

(A) the spread of the individual observations around the sample mean.

(B) roughly how far the estimate $\bar{Y}$ typically lands from the true mean μ_Y.

(C) the probability that the null hypothesis is true.

(D) the fraction of the variation in Y explained by a model.

11. A two-sided test of $H_0 : \mu = 0$ at the 5% level gives a t-statistic of 1.2. The correct conclusion is:

(A) reject H_0: the mean is not zero.

(B) accept H_0: the mean is exactly zero.

(C) fail to reject H_0: the data are consistent with $\mu = 0$, though this does not prove $\mu = 0$.

(D) the test cannot be carried out with a positive t.

12. A strategy's average monthly return over $n = 64$ months is 0.8%, with a sample standard deviation $s = 3.2\%$. Using $SE = s/\sqrt{n}$, the t-statistic for $H_0 : \mu = 0$ is:

(A) 0.25.

(B) 6.4.

(C) 16.

(D) 2.0.

13. A regression coefficient has a p-value of 0.03. This means:

(A) there is a 3% probability that the null hypothesis is true.

(B) if the null hypothesis were true, a result at least this extreme would occur about 3 times in 100 by chance alone.

(C) the coefficient explains 3% of the variation in the outcome.

(D) there is a 97% probability that the estimate equals the true value.

14. A 95% confidence interval for a fund's mean monthly return is [0.1%, 0.9%]. From this alone, a two-sided 5% test of $H_0 : \mu = 0$ would:

(A) reject H_0, because 0 lies outside the interval.

(B) fail to reject H_0, because the interval is narrow.

(C) fail to reject H_0, because the interval contains only positive values.

(D) be impossible to decide from the interval alone.

15. Ordinary least squares (OLS) chooses the intercept and slope that make which quantity as small as possible?

(A) the sum of the residuals.

(B) the largest single residual.

(C) the sum of the absolute values of the residuals.

(D) the sum of the *squared* residuals.

16. In a regression of a stock's return Y on the market's return X, the sample covariance of X and Y is 12 and the sample variance of X is 8. The estimated slope (the stock's beta) is:

(A) 0.67.

(B) 3.

(C) 1.5.

(D) 96.

17. A fitted regression line is $\widehat{Y} = 2 + 0.5X$. For an observation with $X = 10$ and an actual value $Y = 9$, the residual (actual minus predicted) is:

(A) +9.

(B) +2.

(C) -2.

(D) $+7$.

18. In a regression of test scores on class size across many school districts, $R^2 = 0.05$. The best interpretation is:

(A) class size accounts for about 5% of the variation in scores across districts; other factors account for the rest.

(B) the regression has failed, because a good model needs an R^2 close to 1.

(C) 5% of the slope estimate is due to chance.

(D) the slope must be statistically insignificant.

19. OLS gives a slope of -2.28 for a regression of test scores on class size. For this slope to be read as the *causal* effect of class size, rather than a mere association, the key requirement is that:

(A) R^2 is close to 1.

(B) the sample is large enough that the t-statistic exceeds 1.96.

(C) the two variables are strongly correlated.

(D) other factors in the error term (such as district income) do not move systematically with class size, that is, $\mathbb{E}(u \mid X) = 0$.

20. A stock's return is regressed on the market's return and the estimated slope (its beta) is 1.4. On average, when the market rises by 1%, this stock:

(A) rises by exactly 1%.

(B) falls by 1.4%.

(C) rises by about 1.4%, so it swings more than the market.

(D) is unaffected by the market.

Section B: Analytical Problem (30 marks)

You have $n = 60$ months of data on a stock. For each month you observe the stock's excess return Y (in %) and the market's excess return X (in %). The summary statistics are:

$$\bar{X} = 0.5, \qquad \bar{Y} = 1.0, \qquad s_X^2 = 16, \qquad s_{XY} = 24, \qquad s_Y^2 = 100.$$

You will fit the "market model" regression $Y = \beta_0 + \beta_1 X + u$ by OLS. (Here s_X^2 and s_Y^2 are the sample variances and s_{XY} the sample covariance.)

Useful roots: $\sqrt{16} = 4$ *and* $\sqrt{100} = 10$. *No calculator is needed anywhere in this question.*

(a) **[6 marks]** Compute the OLS slope $\widehat{\beta}_1$ (the stock's *beta*) and the intercept $\widehat{\beta}_0$ (its *alpha*). In one sentence, interpret the slope.

(b) **[6 marks]** Compute the correlation r between X and Y, and the R^2 of the regression. What fraction of the variation in the stock's return does the market explain, and what does the rest represent?

(c) **[6 marks]** Use the fitted line to predict the stock's excess return in a month when the market's excess return is $X = 2.5\%$. If the stock actually returned 5.0% that month, compute the residual.

(d) **[6 marks]** The standard error of the slope is $SE(\widehat{\beta}_1) = 0.25$. Test $H_0 : \beta_1 = 1$ (the stock is exactly as risky as the market) against a two-sided alternative at the 5% level, and give the 95% confidence interval for beta.

(e) **[6 marks]** The estimated alpha is $\widehat{\beta}_0 = 0.25\%$ per month. Explain what a *positive* alpha would mean for an investor, and why you must test whether it is statistically different from zero before concluding that the manager has genuine "skill".

Section C: Interpretation (20 marks)

Read the following and answer in **no more than five sentences**. Use the concepts from the course; you do not need any calculation.

A financial analyst collects one year of data on 200 companies. For each firm she records its stock return over the year (Y) and the number of times the firm was mentioned in the financial press that year (X). She regresses Y on X, finds a positive slope that is highly statistically significant ($t = 4.5$, $p < 0.001$), and concludes that generating more press coverage ***causes*** *a company's stock to rise. She advises firms to buy more media attention in order to lift their share price.*

Question [20 marks]. Explain why this significant positive slope does *not* establish that press coverage causes higher returns. Name the key least-squares assumption a causal interpretation requires, and give at least one concrete reason it is likely to fail here. Answer in at most five sentences.