Financial Econometrics | Lecture 1

Economic Questions and Data, and a Review of Probability

In Class Group Activity

· SJTU, Summer School 2026 · Fatih Kansoy

Lesson of the day. Every number you compute from data is one draw from a distribution of numbers you could have got. A finance analyst always asks not just “what is the estimate?” but “how much would it move with a different sample, and does this correlation actually mean that one thing causes the other?”

Phase 1 – Concept checks

ANSWERS

ConcepTest 1. **Answer.** Option **(b)**. The ideal way to answer “does higher pay cause higher returns?” would be a randomised experiment, randomly assigning pay levels to firms, which is impossible. With 200 observational firms, better-run firms both pay more and perform better, so pay and returns are correlated without pay causing returns. Option (a) reads a correlation as causal; (c) confuses a large sample with random assignment; (d) mislabels a causal question as a forecasting one.

A large sample does not fix confounding. Only a design that holds other factors constant can.

ConcepTest 2. **Answer.** Option **(b)**. Forecasting exploits a stable pattern; it does not need a cause. The umbrella forecasts rain without causing it, and the term spread can forecast GDP whether or not it causes anything. Option (a) treats a good forecaster as a cause; (c) wrongly demands a causal mechanism before any forecast is allowed.

Predict on a correlation; intervene only on a cause.

ConcepTest 3. **Answer.** Option **(b)**. Drawing a sample makes the average a random variable with its own sampling distribution. A different 30 months would give a different number, so 0.8% is one estimate, not the truth. Next week we measure how far it can sit from the truth (the standard error). Option (a) treats the estimate as the truth; (c) wrongly ties the randomness of the mean to normality; (d) confuses units, since the variance has *squared* units and cannot equal 0.8%.

Phase 2 – The desk problem

WORKED SOLUTIONS

(a) Asset A. The completed table:

Month	A	$A-\mu_A$	$(A-\mu_A)^2$
1	$+4$	$+2$	4
2	-2	-4	16
3	$+6$	$+4$	16
4	$+1$	-1	1
5	-3	-5	25
6	$+6$	$+4$	16
Sum	12	0	78

Mean $\mu_A = \dfrac{12}{6} = 2.0\%$. Variance $\sigma_A^2 = \dfrac{78}{6} = 13.0$. Standard deviation $\sigma_A = \sqrt{13} = 3.61\%$.

The variance, 13, is in "per cent squared", which has no plain meaning. The standard deviation, 3.61%, is back in per cent, the units of the return, which is why risk is quoted as a standard deviation.

(b) Covariance and correlation. Multiply the two deviation columns month by month:

Month	$A-\mu_A$	$B-\mu_B$	product
1	$+2$	-1	-2
2	-4	$+1$	-4
3	$+4$	-3	-12
4	-1	0	0
5	-5	$+2$	-10
6	$+4$	$+1$	$+4$
Sum			-24

$\operatorname{cov}(A,B) = \dfrac{-24}{6} = -4.0$. Correlation $\rho = \dfrac{-4}{3.6056 \times 1.6330} = \dfrac{-4}{5.8878} = -0.68$.

The covariance is negative, so the two stocks tend to move in opposite directions: when the airline falls, the utility tends to rise. The correlation -0.68 puts that on the fixed -1 to $+1$ scale, and unlike the covariance it has no units.

(c) The normal table.

(1) $z = \dfrac{-7-1}{4} = -2.00$, so $\mathbb{P}(\text{return} < -7) = \Phi(-2) = 0.0228$, about **2.3**%.

(2) $z = \dfrac{9-1}{4} = +2.00$, so $\mathbb{P}(\text{return} > 9) = 1 - \Phi(2) = 0.0228$, about **2.3**% (the same, by symmetry).

(3) $\mu \pm 1.96\sigma = 1 \pm 1.96 \times 4 = 1 \pm 7.84$, that is from $-\mathbf{6.84}\%$ to $+\mathbf{8.84}\%$. So 5% of the time the index sits outside this band, 2.5% in each tail.

The 1.96 and 5% pair is the single most used fact in this course.

(d) The portfolio twist. With $a = b = 0.5$:

$$\mu_P = 0.5 \times 2 + 0.5 \times 2 = 2.0\%.$$

$$\sigma_P^2 = (0.25)(13) + 2(0.5)(0.5)(-4) + (0.25)(2.667) = 3.25 - 2.0 + 0.667 = 1.92,$$

so $\sigma_P = \sqrt{1.92} = 1.38\%$. Direct check: the portfolio returns are $2.5, 0.5, 2.5, 1.5, 0.5, 4.5$, mean 2.0, squared deviations summing to 11.5, giving $11.5/6 = 1.92$. They match.

If the two stocks had been uncorrelated, the portfolio variance would be $3.25+0.667 = 3.92$*. Because the covariance is negative, the cross term subtracts* 2.0 *and pulls the variance down to* 1.92*. The portfolio is far less risky than either stock alone (*$\sigma_A = 3.61\%$*,* $\sigma_B = 1.63\%$*,* $\sigma_P = 1.38\%$*), with the same expected return of* 2%*. That is diversification: while correlation is below* 1*, mixing assets lowers risk for free.*

Phase 3 – The case

WORKED ANSWERS

1. $z = \dfrac{-22.6 - 0.04}{1.12} \approx \mathbf{-20.2}$ standard deviations.

2. Essentially never. The normal probability of a move at least this large is about 5.5×10^{-89}, a decimal point followed by 88 zeros. For scale, the chance of picking one specific second at random from the whole 14-billion-year age of the universe is about 2×10^{-18}, vastly more likely than this.

3. Option **(b)**: fat tails. It is not one freak. Stock and Watson's own Table 2.5 lists ten days between 1980 and 2012 that each exceed 6.4 standard deviations, all astronomically unlikely under a normal model, yet all real. Real stock returns have heavier tails (more kurtosis) than the normal.

4. A -22.6% day is about -20.2 standard deviations, an event the normal model rates at roughly 5.5×10^{-89}. The model did not say "unlikely"; it said "will not happen before the universe ends". Yet it happened on an ordinary Monday. The profession did not conclude that 1987 was a one-off. The ten largest Dow days from 1980 to 2012, including October 2008 and September 2001, each sit beyond 6.4 standard deviations and are all "impossible" under a normal model. Two fixes followed: keep the normal but let the variance change over time, so crises are high-volatility regimes (the GARCH-style models later in the course), or drop the normal for a fat-tailed distribution, the idea Nassim Taleb popularised as the "Black Swan". Tie it back to this morning: the sample you hold, six months or thirty years, is one draw, and a risk number built on the prettiest possible distribution will fail you exactly when it matters most.

End of worked solutions.