Financial Econometrics | Lecture 2

Review of Statistics: Estimation, Testing and Confidence Intervals

In Class Group Activity

· SJTU, Summer School 2026 · Fatih Kansoy

How this block runs. You work in your home group of three or four, textbook open. Roles rotate: **Scribe**, **Reporter**, **Sceptic**, **Timekeeper**. All money today is in US dollars ($) and pounds sterling (£).

- **Phase 1, Concept checks (~10 min).** Vote on your own, talk for 90 seconds, then vote again. Do not talk before the first vote.
- **Phase 2, Desk problem (~20 min).** One problem in parts, on a fund's returns. We solve it together, part by part.
- **Phase 3, Case (~25 min).** One real story, ending hidden. You answer four questions and take a position, then we reveal what happened.

A short word list. The **standard error** $SE = s/\sqrt{n}$ is the estimated standard deviation of a sample mean: how far your estimate typically lands from the truth. A **confidence interval** is estimate $\pm\, 1.96\, SE$ at the 95% level. The **t-statistic** is $t = (\text{estimate} - \text{null})/SE$; you reject the null at 5% when $|t| > 1.96$. The **p-value** is the probability of data at least this extreme if the null were true. The **Student t-distribution** replaces the normal for small n: same centre, fatter tails, so the critical value is a little larger.

	large n (z)	$n = 16$ (t_{15})	$n = 10$ (t_9)	one-sided z
5% critical value	1.96	2.131	2.262	1.645

Phase 1 – Concept checks ~10 MIN

ConcepTest 1. You test whether a fund's average monthly return differs from zero, and you get a p-value of 0.03. Which statement is correct?

(a) There is a 3% probability that the fund's true average return is zero.

(b) There is a 3% probability that you have made the wrong decision.

(c) If the true average return were exactly zero, there is a 3% chance of seeing a sample gap at least this far from zero by luck alone.

(d) There is a 97% probability that the fund really beats zero.

ConcepTest 2. Two studies test whether a strategy's average monthly return is above zero. Study 1: average +0.04%, n = 50,000 trades, $t = 6.0$. Study 2: average +1.8%, n = 12 trades, $t = 1.3$. Which would you rather run with real money?

(a) Study 1, because $t = 6.0$ means a much bigger and more reliable edge.

(b) Study 1, because it is significant and Study 2 is not, so Study 2 has no edge.

(c) Study 2 looks far more useful (+1.8% a month), but with only 12 trades it is too noisy to trust yet; Study 1 is significant but +0.04% is almost worthless after costs.

(d) Study 1, because a larger sample always means a larger true effect.

ConcepTest 3. From your sample, a 95% confidence interval for a fund's mean monthly return is [0.2%, 1.4%]. Which statement is correct?

(a) There is a 95% probability that the true mean lies between 0.2% and 1.4%.

(b) 95% of the fund's monthly returns fall between 0.2% and 1.4%.

(c) If you repeated the sampling many times, about 95% of the intervals built this way would contain the true mean, and this is one of them.

(d) We are 95% sure the sample mean lies between 0.2% and 1.4%.

Phase 2 – The desk problem ~20 MIN

The Meridian Alpha Fund desk

The setting. You are on the risk desk. The Meridian Alpha Fund reports monthly returns in per cent. Over the last $n = 10$ **months** the sample mean is $\overline{Y} = 1.20\%$ and the sample standard deviation is $s = 1.50\%$. Returns are roughly normal; treat the ten months as an i.i.d. sample.

(a) Standard error and confidence interval. Compute $SE = s/\sqrt{n}$, then the 95% confidence interval for the mean monthly return using 1.96. State in one sentence what the half-width means.

(b) Is the edge real? Test H_0 : mean = 0 against a two-sided alternative at 5%. Compute $t = (\overline{Y} - 0)/SE$, give the decision using the 1.96 rule, and say roughly what the p-value is. Check your answer against part (a): is zero inside or outside the interval?

(c) Two desks compared. Now compare two trading desks over the last **36 months each**: Helix has mean 1.50% and SD 3.0%; Pinnacle has mean 0.60% and SD 2.4%. Test whether their average returns differ, two-sided, at 5%. Use

$$SE_{\text{diff}} = \sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}, \qquad t = \frac{\overline{Y}_A - \overline{Y}_B}{SE_{\text{diff}}}.$$

Is the gap real, or could it be sampling noise? Be careful what "fail to reject" means.

(d) The small-sample twist. Go back to the Meridian fund in (a) and (b) with $n = 10$. Because n is small and returns are roughly normal, the honest critical value is the Student t with 9 degrees of freedom, **2.262**, not 1.96.

(1) Re-do the 95% confidence interval with 2.262. Is it wider or narrower than in (a)?

(2) Your t-statistic from (b) is still the same. Do you still reject at 5% using 2.262?

(3) **The flip.** Suppose instead the fund had $\overline{Y} = 1.05\%$ with the same SE, so $t = 2.214$. What does the lazy 1.96 rule say? What does the correct $t_9 = 2.262$ rule say? Explain the difference in one sentence.

Phase 3 – The case ~25 MIN

Momentum Edge: the newsletter that beat the market

How to run this. Read the story with your group. The ending is hidden on purpose. Your group must take a position and hold it. Do not search the internet. I will reveal what happened only after every group has answered.

The story. A London research shop, Momentum Edge, sells a stock-picking signal to retail investors. In their pitch deck they show a backtest on UK equities, 2014 to 2019 (about 72 months). The signal's portfolio earned an average of +0.95% per month above the market, with a monthly tracking standard deviation of about 3.6%.

They run the obvious test. The standard error of the mean excess return is $SE = 3.6/\sqrt{72} = 0.424$, so $t = 0.95/0.424 = 2.24$, two-sided p-value about 0.025. The deck says, in bold: "Statistically significant outperformance at the 5% level ($t = 2.2$, $p = 0.025$). The edge is real." In small print, they mention that this signal was the best of **20 candidate signals** they screened on the same 2014 to 2019 data, and that they "selected the winner".

A pension fund is deciding whether to allocate £50 million.

The story stops here.

Four discussion questions (answer all four; take a position):

1. **Read the number.** With $t = 2.24$ and $p = 0.025$, is the in-sample outperformance significant at 5%?
2. **Size the effect.** +0.95% a month is about +12% a year before costs. Is that economically large, and what one cost would you want to see before believing it?
3. **The multiple-testing trap.** If you test 20 unrelated signals at the 5% level and none truly works, what is the chance at least one looks significant by luck? Compute $1 - 0.95^{20}$.
4. **Decide.** Knowing the winner was the best of 20, would you allocate the £50m on this evidence alone? What single piece of evidence would change your mind?

End of activity handout. Solutions are distributed after the block.