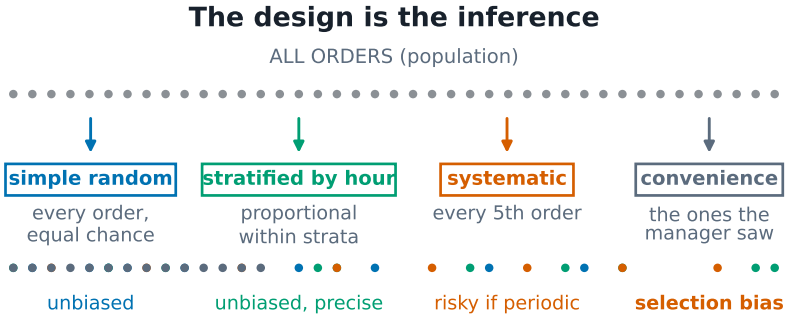
Business Statistics

Curated undergraduate study track (not accredited, no external exam). Original explanations and practice only, following the reviewed course handout.

One dataset, one habit

- Forty takeaway delivery times carry the whole course - every number recomputable.
- The habit is one sentence - state the claim, state the uncertainty, say what would change your mind.

Sampling: the design is the inference



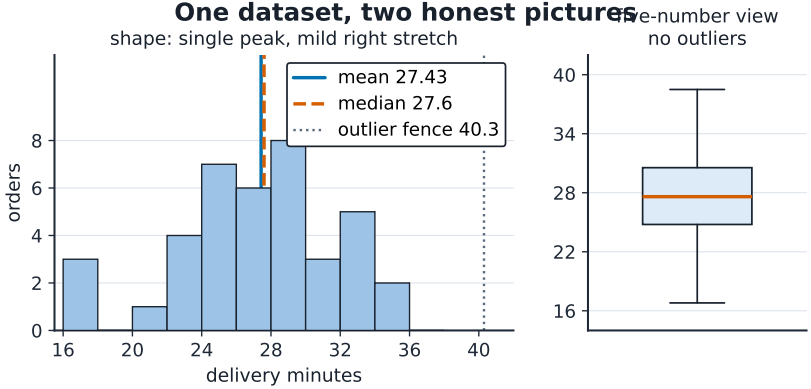
size cannot cure bias: 10,000 convenience orders still say nothing about lunch

Four ways to sample; size cannot cure a biased design.

Sampling in action: match design to claim

1. **Compare** Weekends slower? 20 Saturday + 20 Wednesday orders - **stratified**; 40 Saturday orders answer Saturday only.
2. **Missing** 3 forgotten timers: mean of the 37 complete cases, with the missingness stated - zero-filling invents three instant deliveries.
3. **Verdict** The design matches the claim or the analysis decorates a guess.

Descriptive measures and visualisation: the four questions

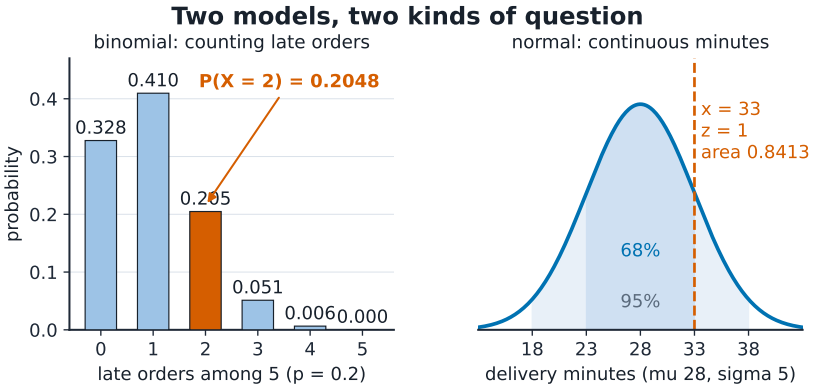


Mean 27.43, **median** 27.6, fence at 40.3: centre, spread and outliers.

Descriptives in action: the typo stress test

1. **Before** Mean 27.43, **median** 27.6, sd 4.91 - a calm dataset, no fence crossings.
2. **After** A 168-minute typo added: mean 30.86 (+3.4), sd 22.5 (x4.6), median 27.7 (+0.1).
3. **Response** The fence, still about 40, flags it; the dictionary explains it; the decision is documented.

Probability: match the model

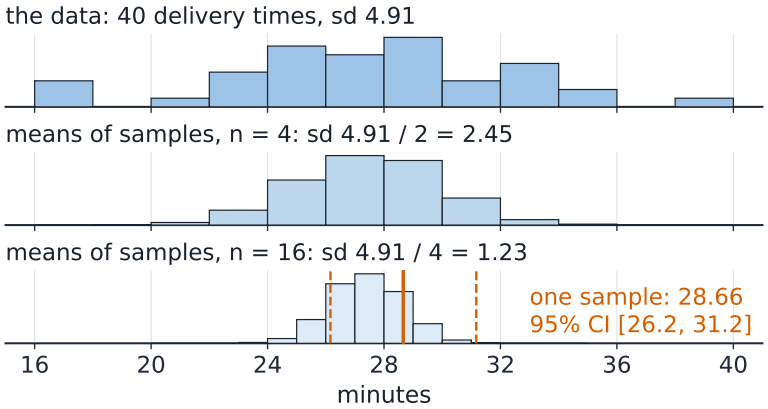


Binomial counts beside the normal curve with its 68/95 bands.

Probability in action: price the promise

1. **Price** mu 28, sigma 5: P(over 30) = 0.345 - a third of orders claim the free dessert.
2. **Fix** mu 26: P(over 30) = 0.212 - better, still one order in five.
3. **Match** Counts of late orders: binomial. One order's minutes: normal. Same kitchen, two models, both justified.

Estimation: the band around the number



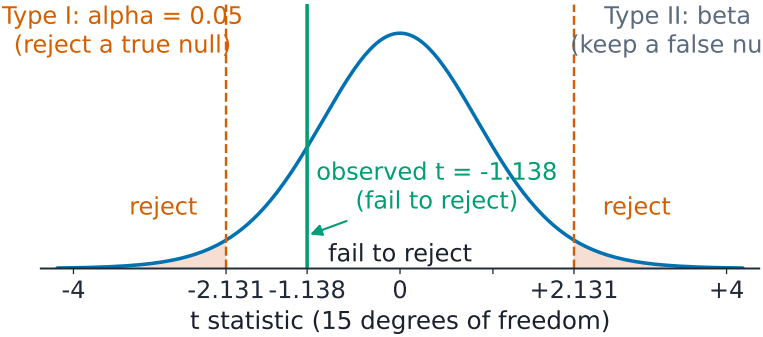
Sample means narrow by root n: the **central limit theorem** at work.

Estimation in action: read the band

1. **Build** Seed-11 sample: 28.66 +/- 2.131 x 1.175 = [26.2, 31.2] at 95%.
2. **Read** The METHOD captures the truth 95% of the time - not the band, not 95% of orders.
3. **Size** +/-0.5 minutes needs n >= 343 - five times the precision costs over twenty times the data.

Testing: surprise, measured

The evidence fell short of the threshold - a result, not a failure

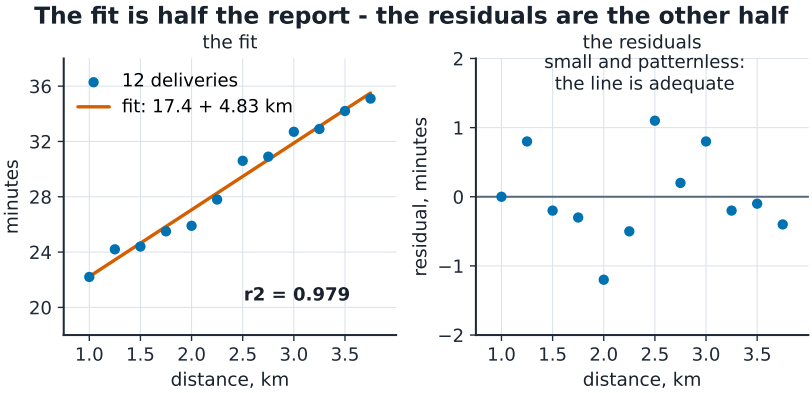


t = -1.138 falls inside +/-2.131: fail to reject is a result.

Testing in action: the rider question

1. **Compute** t = (26.9 - 28.7)/1.0 = -1.8; p about 0.09 two-sided.
2. **Read** Not significant at 0.05 - and not nothing: report the effect with its interval.
3. **Decide** If a wrong re-train is costly, keep collecting; if desserts ride on it, 0.10 may be the right budget.

Regression: the line and its police

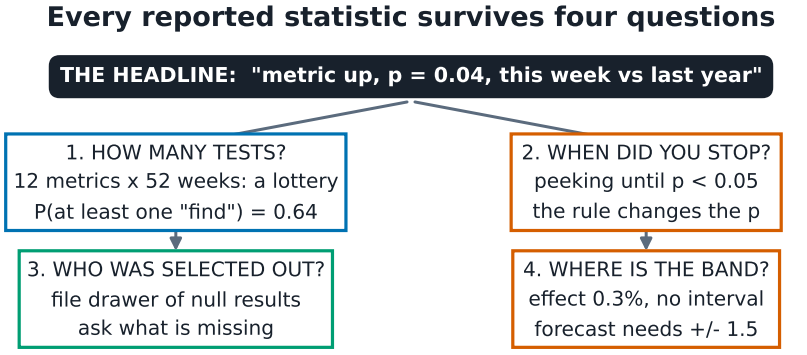


Time = 17.4 + 4.83 km, r2 = 0.979, with patternless residuals.

Regression in action: what the slope may say

1. **Predict** 2 km: 27.1 minutes; 3 km: 31.9 - inside the data's vouched range.
2. **Refuse** 7 km: 51 minutes is a story - no delivery beyond 3.75 km vouches for it.
3. **Police** Distance predicts time here; a causal rate needs route assignment held fixed.

Limitations: the four gates



Challenge a headline through four gates: tests, stopping, selection, error.

Limitations in action: the dashboard audit

1. **Retire** "Store up, p = 0.04": one of 12 x 52 comparisons - selected, not declared.
2. **Repair** Pre-register the hypothesis, confirm on fresh data, report the effect size.
3. **Band** Next Saturday's mean: 27.4 +/- 1.5 at 95% - staff to 28.9, the band's upper end.

The dataset page, reproducible

1. **Verify** n 40, mean 27.43, sd 4.91, quartiles 24.5 / 27.6 / 30.9, no fence crossings.
2. **Reuse** The seed-11 sample (mean 28.66, SE 1.175) and the 12 km-minutes pairs live on the same page.
3. **Audit** A reader with a calculator can check any number in this course - that is the standard.

Check yourself

- The design for "weekends slower"? (**Stratified**: both days sampled.)
- Mean 27.43, sd 4.91, n = 40: the SE? (0.78.)
- n = 5, p = 0.2: P(exactly 2)? (0.2048.)
- The fit: minutes per km, and what r2 = 0.979 does not mean? (4.83; not accuracy, not causation.)