STATISTICS · STUDY GUIDE

Final Exam Study Guide

The final covers all fourteen chapters. Each chapter here is shorter than in the exam guides: three paragraphs of what to know, the points to lock in, its diagram, one formula table, one worked example, the traps and six new self-test items. If a chapter feels shaky, go back to its exam guide for the full treatment. The long cram sheet at the end pulls every formula and rule into one place. Work two chapters a day and the whole course fits into a week.

WHAT THIS GUIDE COVERS

- Describing data: chapters 1 to 4 (what statistics is, sampling and design, graphs, center and spread).
- Probability: chapters 5 and 6 (rules of probability, discrete random variables).
- The bridge to inference: chapters 7 and 8 (the normal model, sampling distributions and the CLT).
- Inference: chapters 9 to 12 (confidence intervals, testing logic, one-sample and two-sample tests).
- More than two groups and two variables: chapters 13 and 14 (chi-square, ANOVA, correlation and regression).
- Every chapter carries what to know, a diagram, a table, a worked example, the traps and a six-item self-test, and the cram sheet holds every formula.

CHAPTER 1

What Statistics Is

Read it in full: **Chapter 1 in the textbook**. Practice it in the lab: **Describe or Infer** and **Spot the Misleading Claim**.

What to know

Descriptive statistics summarizes the data in hand; inferential statistics generalizes from a sample to a population and must report its uncertainty. The mean ($\Sigma x \div n$) is pulled by extreme values and the median is not, so in the book's commutes the mean is 31.1 and the median 24.

Parameters (μ, σ, p, N) describe populations and are fixed but unknown; statistics ($\bar{x}$, s, $\hat{p}$, n) describe samples and vary. Sampling error shrinks with n; sampling bias does not. Sample

variance divides by n − 1, population variance by N.

Data are qualitative or quantitative, and quantitative data are discrete or continuous. The four levels are nominal, ordinal, interval and ratio; only ratio supports "twice as much." Correlation is not causation, and a relative risk without its baseline can make a tiny effect sound large.

LOCK THESE IN

- Descriptive describes the data in hand; inferential generalizes and must report uncertainty.
- Parameter goes with population and Greek letters (μ, σ, p, N); statistic goes with sample and Roman letters ($\bar{x}$, s, $\hat{p}$, n).
- The mean is pulled by outliers; the median is not. Report both when they disagree.
- Sampling error shrinks with bigger n. Sampling bias needs better design, not more data.
- Sample variance divides by n − 1; population variance divides by N.
- Counted is discrete, measured is continuous; digits used as names are qualitative.
- Nominal, ordinal, interval, ratio: each level adds order, then equal gaps, then a true zero.
- Relative risk reduction = $(r_c - r_t) \div r_c$; absolute = $r_c - r_t$; NNT = 1 ÷ absolute reduction.



One extreme value moves the mean, not the median. Nine commutes have a median of 24 minutes, but the single 88 drags the mean up to 31.1.

Chapter 1 formulas and facts

Quantity	Formula	When to use it
Mean	$\bar{x} = \Sigma x \div n$ (sample); $\mu = \Sigma x \div N$ (population)	Center of quantitative data; totals and long-run averages
Median	Middle sorted value; average of the two middle values if n is even	Typical value; skewed data or outliers
Range	Largest − smallest	Quickest spread; fragile, set by two values
Population SD	$\sigma = \sqrt{[\Sigma(x - \mu)^2 \div N]}$	You have every member of the population
Sample SD	$s = \sqrt{[\Sigma(x - \bar{x})^2 \div (n - 1)]}$	You have a sample and want to estimate σ
Relative risk reduction	$(r_c - r_t) \div r_c$	How large the change is relative to baseline
Absolute risk reduction	$r_c - r_t$	How many fewer cases per person treated
Number needed to treat	NNT = 1 ÷ ARR	People treated for one to benefit
Celsius to Fahrenheit	°F = (9 ÷ 5)°C + 32	Testing whether a ratio survives a unit change
Celsius to kelvin	K = °C + 273.15	Kelvin has a true zero, so its ratios are real

Worked examples

EXAMPLE 1.1 · MEAN, MEDIAN AND RANGE WITH AN OUTLIER

A clinic records patient wait times in minutes: 12, 15, 9, 14, 11, 13, 58, 10. Find the mean, median and range, and say which better describes a typical wait.

STEP 1 · ADD AND DIVIDE

Sum = 142 and n = 8, so $\bar{x} = 142 \div 8 =$ **17.75 minutes**.

STEP 2 · SORT

Sorted: 9, 10, 11, 12, 13, 14, 15, 58.

STEP 3 · FIND THE MIDDLE

With n = 8 the median is the average of the 4th and 5th values: (12 + 13) ÷ 2 = **12.5 minutes**.

STEP 4 · RANGE

58 − 9 = **49 minutes**.

STEP 5 · INTERPRET

Seven of the eight waits are 15 minutes or less, yet the mean is 17.75 because of the single 58. The median of 12.5 describes a typical wait. Without the 58 the mean is 12.00, close to the median.

TRAPS PEOPLE FALL INTO

- **Calling a sample summary inferential.** "These 40 patients averaged 128" is descriptive. It turns inferential only when it claims something about patients not measured.
- **Mixing up the symbols.** μ and σ are population parameters; $\bar{x}$ and s are sample statistics. A question that gives $\bar{x}$ is giving you a statistic.
- **Thinking a huge sample cures bias.** Size shrinks sampling error only. A self-selected sample of 500,000 still estimates the wrong group.
- **Averaging labels.** A spreadsheet will average ZIP codes or 1 = single, 2 = married. The result means nothing.
- **Taking ratios on an interval scale.** 20°C is not twice 10°C, and the year 2000 is not twice the year 1000.
- **Reporting only the mean for skewed data.** One extreme value can move the mean far from a typical case. Give the median too.
- **Reading relative risk alone.** "Halves your risk" may mean 2 in 1,000 down to 1 in 1,000. Ask for the baseline and the NNT.
- **Misreading the p-value.** p = 0.03 is not a 3% chance there is no effect, and p = 0.40 does not prove there is no effect.

Chapter 1 self-test · six items

1. A poll of 1,200 adults finds 58% support a plan and reports "about 58% of all adults support it, give or take 3 points." This claim is:

A. Inferential, because it reaches beyond the 1,200
B. Descriptive, because it gives a single percent
C. A census, because 1,200 is a large number
D. A parameter, because it is about all adults

A. Inferential, because it reaches beyond the 1,200 Any claim about people who were not measured is inferential, and it comes with a statement of uncertainty, here 3 points.

2. Six values are 3, 5, 5, 8, 9 and 40. What is the median?

A. 5
B. 11.7
C. 6.5
D. 8

C. 6.5 With n = 6, average the 3rd and 4th sorted values: (5 + 8) ÷ 2 = 6.5. The mean, 70 ÷ 6 ≈ 11.7, is pulled up by the 40.

3. Which statement about sampling error is correct?

A. It is caused by poor question wording
B. It shrinks as the sample size grows
C. It disappears with random selection
D. It grows when the sample is larger

B. It shrinks as the sample size grows Sampling error is the natural difference between a statistic and the parameter. It cannot be removed, but larger samples make it smaller.

4. Jersey numbers on a soccer team are measured at which level?

A. Ordinal
B. Interval
C. Ratio
D. Nominal

D. Nominal The digits are names. Number 10 is not ranked above or "more than" number 4, so only counts and the mode make sense.

5. A drug lowers risk from 2 in 1,000 to 1 in 1,000. Which pair is correct?

A. Relative reduction 0.1%, absolute 50 points
B. Relative reduction 1%, absolute 0.5 points
C. Relative reduction 50%, absolute 50 points
D. Relative reduction 50%, absolute 0.1 points

D. Relative reduction 50%, absolute 0.1 points Relative: (0.002 − 0.001) ÷ 0.002 = 50%. Absolute: 0.002 − 0.001 = 0.001, or 0.1 percentage points, so NNT = 1,000.

6. A study reports p = 0.40. What may you conclude?

A. Compatible with no effect, but an effect is not ruled out
B. There is a 40% chance that the null hypothesis is true
C. The study has proven that there is no effect at all
D. There is a 60% chance that the effect is real

A. Compatible with no effect, but an effect is not ruled out A large p-value means the data do not give convincing evidence against the null. Absence of evidence is not evidence of absence, and a p-value is never the probability that a hypothesis is true.

Sampling and Data Collection

Read it in full: **Chapter 2 in the textbook**. Practice it in the lab: **Why Sampling Works** and **Design an Experiment**.

What to know

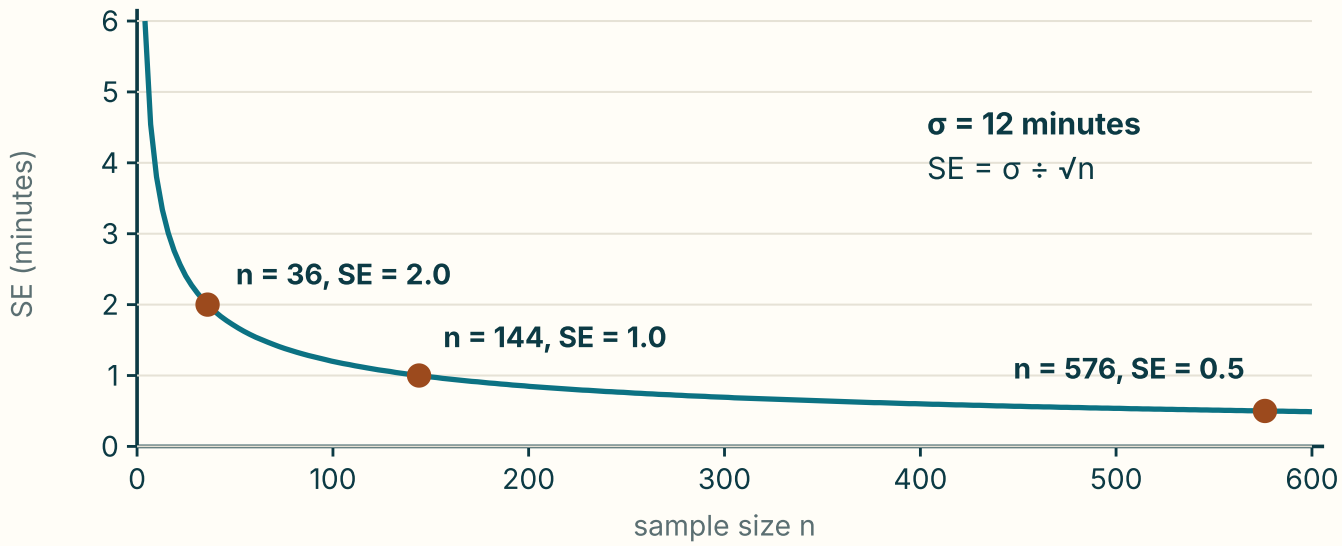
A random sample represents the population on every variable at once, and its wobble is measurable: $SE(\bar{x}) = \sigma \div \sqrt{n}$. Quadruple n to halve the SE; population size barely matters if the sample is a small share of it.

Probability methods: simple random, stratified (some from every group, $n_h = n \times N_h \div N$), cluster (everyone from some groups) and systematic (every k-th, $k = N \div n$). Convenience and voluntary response samples support no inference. Bias, unlike sampling error, never shrinks with n, and a margin of error ignores it.

Observational studies show association and are open to confounding; experiments assign treatments and can show cause. Random sampling gives generalization; random assignment gives cause. Good experiments use control, randomization, replication, blocking, a placebo and double blinding.

LOCK THESE IN

- $SE(\bar{x}) = \sigma \div \sqrt{n}$. Four times the sample gives half the standard error.
- Population size does not appear in the SE formula, as long as the sample is under about 10% of it.
- SRS: every group of size n equally likely. Systematic: every k-th after a random start, $k = N \div n$.
- Stratified takes some from every group; cluster takes everyone from some groups.
- Proportional allocation: $n_h = n \times N_h \div N$.
- Bias is systematic and never shrinks with n. A margin of error covers chance only.
- Random sampling licenses generalization; random assignment licenses cause.
- Experiment pillars: control, randomization, replication, blocking; add a placebo and double blinding.



Quadruple n to halve the standard error. With σ = 12 minutes, the standard error falls from 2.0 at n = 36 to 1.0 at n = 144 and 0.5 at n = 576. Population size never appears in the formula.

Chapter 2 formulas and facts

Quantity	Formula	When to use it
Standard error of $\bar{x}$	SE = σ ÷ √n	Typical miss of a sample mean; random, independent sample
Halving the SE	Multiply n by 4	Planning a sample size
10% condition	n ≤ 0.10N	Checking that draws are close to independent
Proportional allocation	$n_h = n \times (N_h \div N)$	Stratified sample sized to each stratum
Systematic step	k = N ÷ n; picks are start + k(j − 1)	Every k-th unit from a list or stream
Sample proportion	$\hat{p}$ = successes ÷ n	Share of a sample with a trait
Margin of error (proportion)	$1.96 \times \sqrt{[\hat{p}(1 - \hat{p}) \div n]}$	95% chance error only; needs 10 successes and 10 failures
Number of treatments	levels of factor 1 × levels of factor 2	Counting arms in a factorial experiment

Worked examples

EXAMPLE 2.1 · STANDARD ERROR AT THREE SAMPLE SIZES

Systolic blood pressure in a large adult population has σ = 15 mmHg. Find the standard error of $\bar{x}$ for n = 25, 100 and 400.

STEP 1 · N = 25

SE = 15 ÷ √25 = 15 ÷ 5 = **3.0 mmHg**.

STEP 2 · N = 100

SE = 15 ÷ √100 = 15 ÷ 10 = **1.5 mmHg**.

STEP 3 · N = 400

SE = 15 ÷ √400 = 15 ÷ 20 = **0.75 mmHg**.

STEP 4 · READ THE PATTERN

Each step multiplies n by 4 and halves the SE. Sixteen times the data (25 to 400) buys four times the precision.

CHECK THE CONDITION

The population is large, so 400 is far less than 10% of it, and the draws can be treated as independent.

TRAPS PEOPLE FALL INTO

- **Thinking precision grows with n.** It grows with √n. Doubling n cuts the SE by only about 29%; you need four times the n to halve it.
- **Sizing a sample as a share of the population.** N is not in σ ÷ √n. A random 1,000 works for a country as well as a town.
- **Calling systematic sampling an SRS.** Everyone has an equal chance, but most groups of size n can never be chosen, so it is not an SRS.
- **Swapping stratified and cluster.** Stratified: some from every group. Cluster: all from some groups.
- **Trusting the margin of error to cover bias.** It counts chance only. Nonresponse and undercoverage can be many times larger.
- **Using causal verbs for observational findings.** Rewrite "cuts" or "boosts" as "is associated with" unless the treatment was randomly assigned.
- **Mixing up the two randomizations.** Random sampling gives generalization. Random assignment gives cause.
- **Thinking randomization is about fairness.** Its job is to balance every variable, including ones nobody measured.

Chapter 2 self-test · six items

1. With σ = 20, what sample size gives a standard error of 2?

A. 10
B. 40
C. 100
D. 400

C. 100 SE = σ ÷ √n, so √n = 20 ÷ 2 = 10 and n = 100.

2. A list of 4,000 invoices is sampled with k = 25 after a random start. How large is the sample?

A. 100
B. 250
C. 25
D. 160

D. 160 k = N ÷ n, so n = N ÷ k = 4,000 ÷ 25 = 160.

3. Which design takes a random sample from inside every group?

A. Cluster sampling
B. Stratified sampling
C. Systematic sampling
D. Convenience sampling

B. Stratified sampling Stratified sampling samples within every stratum. Cluster sampling picks some whole groups and skips the rest.

4. People asked "Do you floss daily?" tend to say yes more often than they actually floss. Which error is this?

A. Undercoverage
B. Nonresponse bias
C. Sampling error
D. Response bias

D. Response bias The answers themselves are inaccurate because of social desirability. That is response bias, a kind of nonsampling error.

5. A trial uses random assignment but volunteers from one city. What can it conclude?

A. Cause, for the whole national population
B. Association only, for the whole population
C. Cause, but only for people like these volunteers
D. Neither cause nor association

C. Cause, but only for people like these volunteers Random assignment supports cause. Without random sampling, the result generalizes only to subjects similar to those studied.

6. What does a placebo do in an experiment?

A. Makes the belief of being treated equal in both groups
B. Makes the treatment group larger than the control
C. Removes the need to randomize the subjects
D. Lets subjects pick the arm they prefer

A. Makes the belief of being treated equal in both groups The placebo effect is real. Giving the control group an inert look-alike means any difference reflects the treatment, not the belief.

Displaying Data

Read it in full: **Chapter 3 in the textbook**. Practice it in the lab: **Choosing the Bin Width** and **Truncated Axis**.

What to know

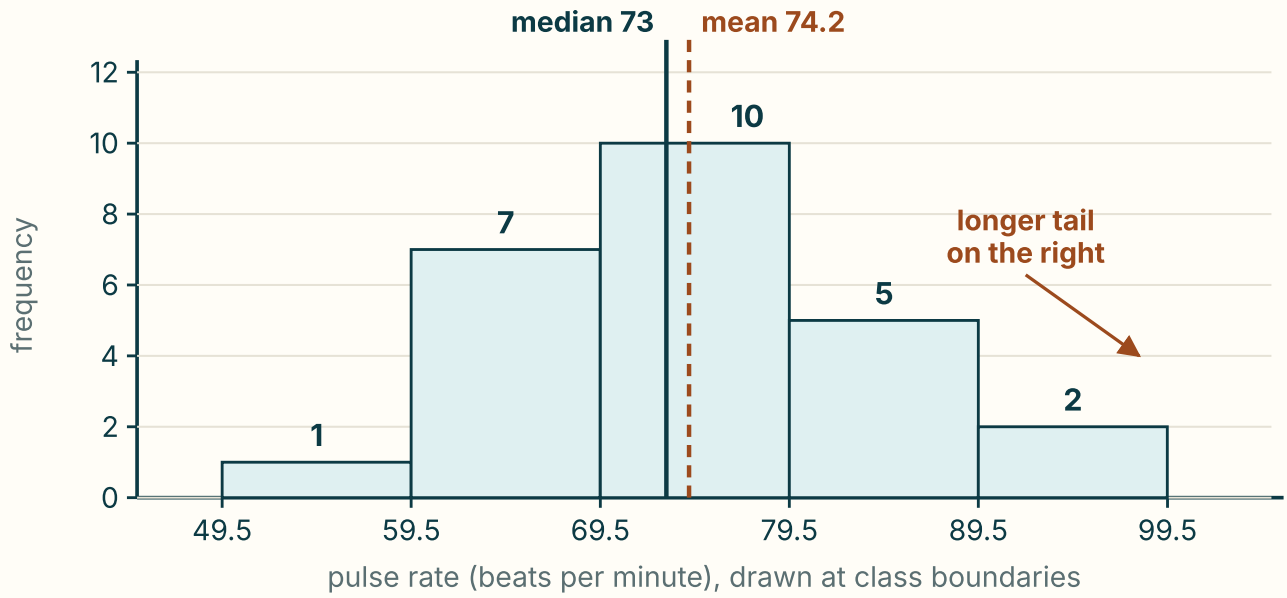
Frequency counts sum to n, relative frequencies (f ÷ n) to 1, and the cumulative column ends at n. Grouped tables use 5 to 10 classes with width = range ÷ classes, rounded up, and no overlaps or gaps.

Bar charts (gaps) and pies (angle = f ÷ n × 360°) are for categories; histograms (touching bars), stemplots (with a key) and dot plots are for measurements. Skew is named for the tail; right skew usually puts the mean above the median.

Bar charts and histograms must start at zero. Lie factor = apparent effect ÷ actual effect, and a truncated axis can turn a 6.1% rise into an apparent 600%.

LOCK THESE IN

- Frequencies sum to n; relative frequencies (f ÷ n) sum to 1; cumulative frequency ends at n.
- Class width = (max − min) ÷ number of classes, rounded up. No overlaps, no gaps.
- Midpoint = (lower limit + upper limit) ÷ 2.
- Pie slice angle = (f ÷ n) × 360°; 1% of a pie is 3.6°.
- Bar chart bars have gaps (categories); histogram bars touch (continuous scale).
- Skew is named for the tail. Right skew: mean usually above median.
- Bimodal usually means two groups mixed together; split them.
- Lie factor = apparent effect ÷ actual effect. Bars must start at zero.



Bars touch because the axis is a continuous scale; bar charts of categories keep gaps.

A histogram of 25 pulse rates. Classes of width 10 give counts 1, 7, 10, 5 and 2. The tail stretches right, so the mean (74.2) sits a little above the median (73).

Chapter 3 formulas and facts

Quantity	Formula	When to use it
Relative frequency	f ÷ n	Comparing groups of different sizes
Percent	(f ÷ n) × 100%	Reporting a share
Cumulative frequency	Running total of f	"How many at or below?"; ends at n
"More than" share	1 − cumulative relative frequency	The complement of "at most"
Class width	(max − min) ÷ classes, rounded up	Grouped table, 5 to 10 classes
Class midpoint	(lower + upper) ÷ 2	One value to stand for a class
Class boundary	Halfway between one upper limit and the next lower limit	Where histogram bars meet
Pie angle	(f ÷ n) × 360°	Exclusive, exhaustive categories only
Lie factor	apparent effect ÷ actual effect	Judging a graph; honest is about 1
Area inflation	scale² for area; scale³ for apparent volume	Resized icons

Worked examples

EXAMPLE 3.1 · A GROUPED FREQUENCY TABLE

Twenty-two clinic waits in minutes, sorted: 4, 6, 9, 11, 12, 14, 15, 15, 17, 18, 19, 21, 22, 24, 25, 27, 29, 31, 34, 38, 41, 46. Build a grouped table with 5 classes and find the share who waited 30 minutes or less.

STEP 1 · RANGE AND WIDTH

Range = 46 − 4 = 42. 42 ÷ 5 = 8.4, rounded up to a width of **9**.

STEP 2 · CLASSES

Starting at the minimum: 4 to 12, 13 to 21, 22 to 30, 31 to 39, 40 to 48. The last class reaches 48, past the maximum of 46.

STEP 3 · COUNT

4 to 12: 5; 13 to 21: 7; 22 to 30: 5; 31 to 39: 3; 40 to 48: 2. Total 22 = n.

STEP 4 · MIDPOINTS

8, 17, 26, 35, 44.

STEP 5 · CUMULATIVE

Running totals: 5, 12, 17, 20, 22. At or below 30 minutes: 17 of 22 = **0.773**, about 77%.

TRAPS PEOPLE FALL INTO

- **Rounding the class width down.** Then the classes stop short of the maximum. Always round up.
- **Overlapping classes such as 10 to 20, 20 to 30.** A value of 20 lands in two classes and the counts no longer sum to n.
- **Naming skew by the pile.** Most data sit on the left in a right-skewed distribution. Follow the tail.
- **Putting multiple-response data in a pie.** If people could pick more than one, the shares pass 100%. Use bars.
- **Gaps in a histogram or touching bars in a bar chart.** The gaps say "categories"; touching bars say "continuous scale." Swap them and the picture lies.
- **Trusting a bar chart that starts above zero.** Length means something only from zero. Read the axis first.
- **Resizing an icon to show a doubling.** Doubling width and height makes 4 times the area. Repeat icons instead.
- **Summarizing bimodal data with one mean.** The mean may fall in the valley where almost nobody is. Split the groups.

Chapter 3 self-test · six items

1. A frequency table has 60 observations. What must the cumulative frequency column end at?

A. 60

B. 1

C. 100

D. 30

A. 60 The cumulative column is a running total of counts, so its last entry equals n = 60. The relative column ends at 1.

2. Which is the midpoint of the class 24 to 31?

A. 27.5
B. 27.0
C. 28.0
D. 7.0

A. 27.5 (24 + 31) ÷ 2 = 55 ÷ 2 = 27.5.

3. A Pareto chart is:

A. A pie chart with the slices placed in order
B. A histogram with equal class widths throughout
C. A cumulative frequency line graph, an ogive
D. A bar chart sorted from tallest to shortest

D. A bar chart sorted from tallest to shortest Sorting bars by height makes the biggest categories stand out.

4. Ten exam scores have a mean of 71 and a median of 80. The likely shape is:

A. Skewed right
B. Uniform
C. Skewed left
D. Symmetric

C. Skewed left Low scores in a left tail drag the mean below the median.

5. An icon is drawn 3 times as wide and 3 times as tall to show a tripling. By what factor does its area grow?

A. 3
B. 6
C. 9
D. 27

C. 9 Area grows by 3 × 3 = 9. If it looks three-dimensional, the implied volume grows by 27.

6. Which display keeps every original value while showing the shape?

A. A histogram with narrow bins
B. A stem-and-leaf plot
C. A pie chart of the values
D. A Pareto chart of the counts

B. A stem-and-leaf plot A stemplot shows the shape like a histogram but keeps each value as a leaf.

CHAPTER 4

Center and Spread

Read it in full: **Chapter 4 in the textbook**. Practice it in the lab: **Z Scores** and **Empirical Rule**.

What to know

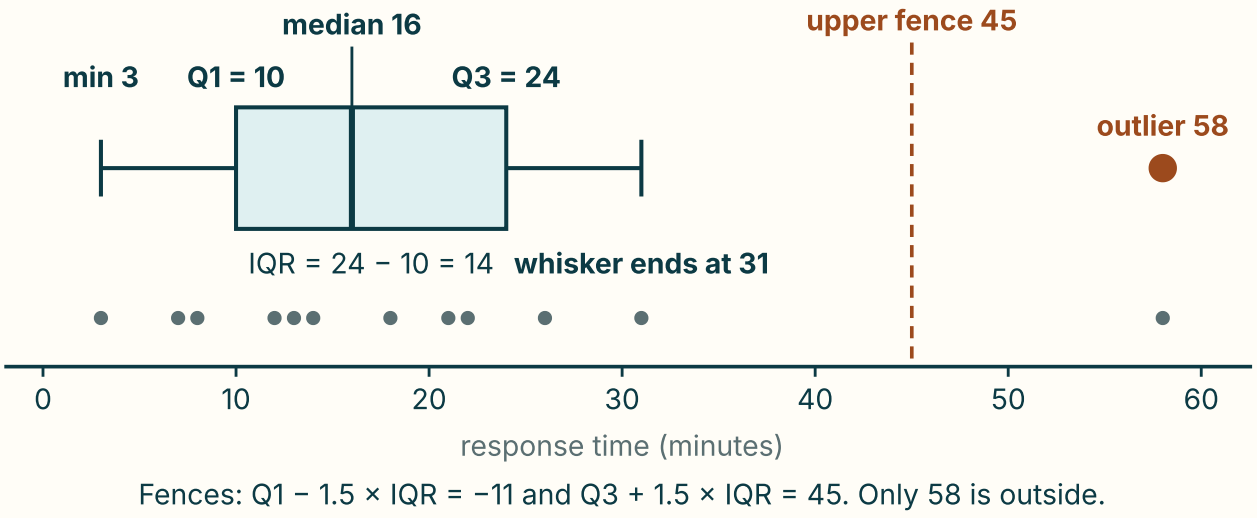
Mean (balance point, not resistant), median (middle, resistant) and mode (most common, the only center for categories). Mean above median suggests right skew.

Sample variance $s^2 = \Sigma(x - \bar{x})^2 \div (n - 1)$; the SD is its square root in the original units. IQR = Q3 − Q1, and the fences Q1 − 1.5 × IQR and Q3 + 1.5 × IQR flag outliers on a boxplot.

$z = (x - \mu) \div \sigma$ measures distance in SDs, and $x = \mu + z\sigma$ reverses it. For bell-shaped data 68, 95 and 99.7% lie within 1, 2 and 3 SD; for any shape Chebyshev guarantees at least $1 - 1 \div k^2$.

LOCK THESE IN

- $\bar{x} = \Sigma x \div n$. Median: position $(n + 1) \div 2$, or the average of the two middle values.
- The median and IQR resist outliers; the mean, range and SD do not.
- $s^2 = \Sigma(x - \bar{x})^2 \div (n - 1)$; $\sigma^2 = \Sigma(x - \mu)^2 \div N$; SD = √variance.
- Deviations from the mean always sum to zero.
- IQR = Q3 − Q1. Fences: Q1 − 1.5 × IQR and Q3 + 1.5 × IQR.
- $z = (x - \mu) \div \sigma$, and $x = \mu + z\sigma$.
- Empirical rule (bell shapes only): 68%, 95%, 99.7% within 1, 2, 3 SD.
- Chebyshev (any shape): at least $1 - 1 \div k^2$; 75% at k = 2, 88.9% at k = 3.



The five-number summary drawn as a boxplot. The box spans the IQR from 10 to 24 with the median at 16. The whisker stops at 31, the last value inside the upper fence of 45, and 58 is flagged as an outlier.

Chapter 4 formulas and facts

Quantity	Formula	When to use it
Mean	$\bar{x} = \Sigma x \div n$	Symmetric quantitative data; later formulas
Median	Middle value; $(n + 1) \div 2$ position	Skewed or ordinal data; "typical"
Range	max − min	Quick look; fragile
Sample variance	$s^2 = \Sigma(x - \bar{x})^2 \div (n - 1)$	Spread of a sample
Population variance	$\sigma^2 = \Sigma(x - \mu)^2 \div N$	Every member measured
Standard deviation	$s = \sqrt{s^2}$; $\sigma = \sqrt{\sigma^2}$	Spread in original units
Coefficient of variation	$CV = (s \div \bar{x}) \times 100\%$	Comparing spread across scales; ratio data
IQR and fences	Q3 − Q1; Q1 − 1.5 IQR, Q3 + 1.5 IQR	Resistant spread; outlier flags
z-score	$z = (x - \mu) \div \sigma$	Comparing positions across scales
Back from z	$x = \mu + z\sigma$	Value at a given number of SDs
Chebyshev	at least $1 - 1 \div k^2$, $k > 1$	Any shape, unknown shape

Worked examples

EXAMPLE 4.1 · SAMPLE VARIANCE AND STANDARD DEVIATION BY HAND

Five patients wait 6, 9, 11, 14, 20 minutes. Find s^2 and s.

STEP 1 · MEAN

$\Sigma x = 60$, so $\bar{x} = 60 \div 5 = \mathbf{12}$.

STEP 2 · DEVIATIONS

−6, −3, −1, 2, 8. Check: they sum to 0.

STEP 3 · SQUARE AND ADD

$36 + 9 + 1 + 4 + 64 = SS = \mathbf{114}$.

STEP 4 · DIVIDE BY N − 1

$s^2 = 114 \div 4 =$ **28.5** square minutes.

STEP 5 · SQUARE ROOT

$s = \sqrt{28.5} \approx$ **5.34 minutes**. As a population the variance would be $114 \div 5 = 22.8$ and $\sigma \approx 4.77$, a little smaller.

TRAPS PEOPLE FALL INTO

- **Calling the mean "typical" for skewed data.** Income, prices and wait times are skewed. The median is the typical value.
- **Dividing by n for a sample.** Sample variance divides by n − 1. Dividing by n underestimates σ^2.
- **Reading the variance as a distance.** It is in squared units. Take the square root to get the SD.
- **Including the median in both halves when n is odd.** The common textbook method leaves it out of both halves for Q1 and Q3.
- **Deleting a point because it is outside a fence.** The 1.5 × IQR rule flags values to investigate, not to throw away.
- **Treating z as a probability.** z is a location in SDs. A percentile needs a known shape.
- **Using 68, 95, 99.7 on skewed data.** If $\mu - 2\sigma$ is negative for a quantity that cannot be negative, the rule does not fit.
- **Applying Chebyshev with k = 1.** It gives 0%, which is true but useless. The theorem needs $k > 1$.

Chapter 4 self-test · six items

1. With n = 8, the median is:

A. The 4th value in sorted order

B. The average of the 4th and 5th sorted values

C. The 5th value in sorted order

D. The mean of all eight of the values

B. The average of the 4th and 5th sorted values For even n there is no single middle value, so average the two middle ones.

2. The sum of squares for a sample of 6 is 90. What is s^2?

A. 15
B. 9.5
C. 3
D. 18

D. 18 $s^2 = SS \div (n - 1) = 90 \div 5 = 18$.

3. Q1 = 14 and Q3 = 26. What is the upper fence?

A. 38
B. 32
C. 44
D. 40

C. 44 IQR = 12, 1.5 × 12 = 18, and 26 + 18 = 44.

4. A score of 88 comes from a test with μ = 80 and σ = 5. What is its z-score?

A. 1.60
B. 0.63
C. 8.00
D. 1.10

A. 1.60 $z = (88 - 80) \div 5 = 8 \div 5 = 1.60$.

5. Chebyshev guarantees what share within 4 SD of the mean?

A. About 99.99%
B. At least 93.75%
C. At least 75%
D. At least 88.9%

B. At least 93.75% $1 - 1 \div 4^2 = 1 - 1/16 = 0.9375$.

6. For bell-shaped data with μ = 50 and σ = 4, about 95% lie between:

A. 46 and 54
B. 38 and 62
C. 42 and 58
D. 48 and 52

C. 42 and 58 Two SDs: 50 ± 8 gives 42 to 58.

CHAPTER 5

Probability Basics

Read it in full: **Chapter 5 in the textbook**. Practice it in the lab: **Addition Rule** and **Conditional Probability**.

What to know

For equally likely outcomes, P(A) = outcomes in A ÷ outcomes in S. Probabilities lie in [0, 1], sum to 1 over S, and $P(A^c) = 1 - P(A)$. Nothing is ever "due."

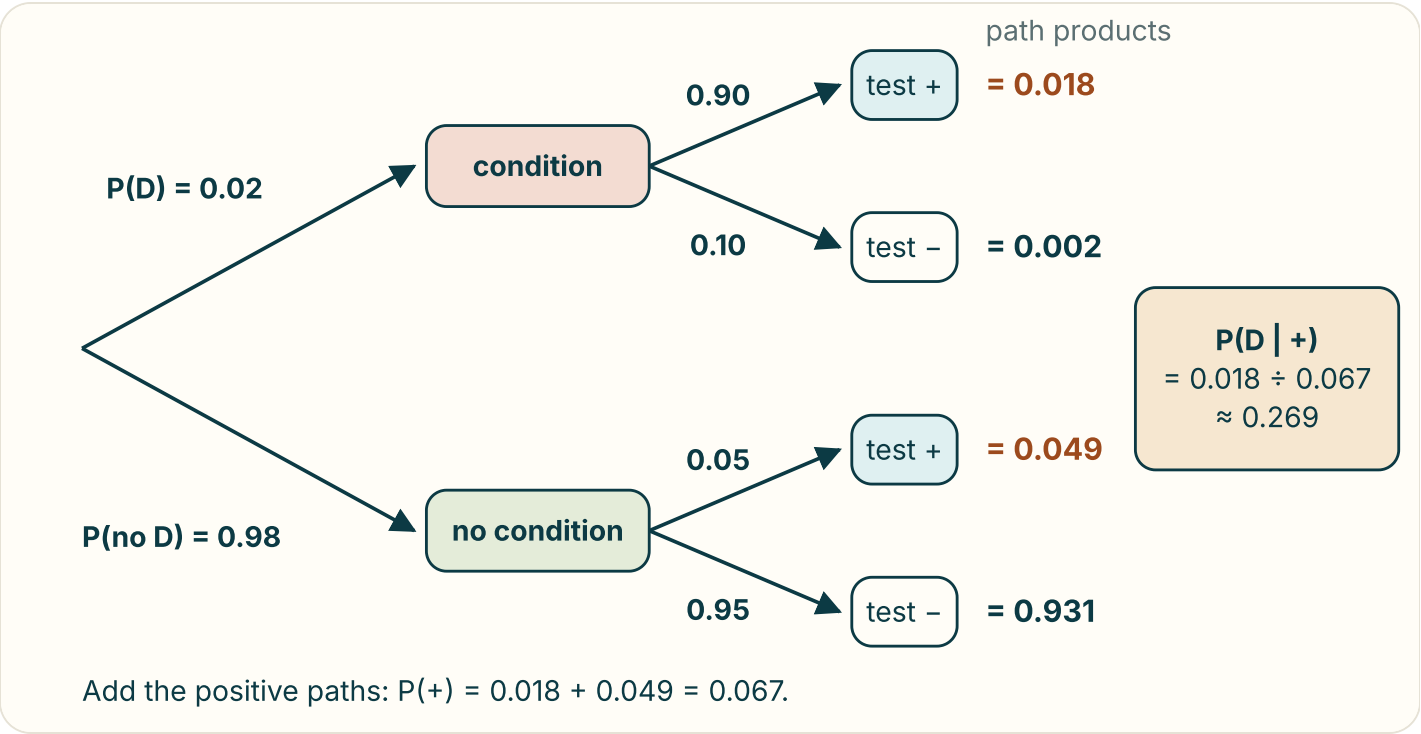
Addition: P(A or B) = P(A) + P(B) − P(A and B). Multiplication: P(A and B) = P(A)P(B | A), which becomes P(A)P(B) only for independent events. Mutually exclusive events with nonzero probability are dependent. "At least one" = 1 − P(none).

P(A | B) = P(A and B) ÷ P(B), and P(A | B) is not P(B | A). In a two-way table joint, marginal and conditional probabilities differ only in the denominator; in a tree, multiply along paths and add

across them.

LOCK THESE IN

- P(A) = outcomes in A ÷ outcomes in S, for equally likely outcomes only.
- Complement: $P(A^c) = 1 - P(A)$. "At least one" = 1 − P(none).
- Addition: P(A or B) = P(A) + P(B) − P(A and B).
- Multiplication: P(A and B) = P(A) × P(B | A); independent means P(B | A) = P(B).
- Mutually exclusive means P(A and B) = 0, and that makes the events dependent.
- Conditional: P(A | B) = P(A and B) ÷ P(B). Reversing the bar changes the answer.
- Joint ÷ grand total; marginal = margin ÷ grand total; conditional = cell ÷ its row or column.
- Trees: multiply along, add across, and the ends sum to 1.



Multiply along paths, add across paths. With 2% prevalence, a test that catches 90% of cases and gives 5% false positives yields P(+) = 0.067, and only about 27% of positives have the condition.

Chapter 5 formulas and facts

Rule	Formula	When to use it
Classical probability	$P(A) = n(A) \div n(S)$	Equally likely outcomes
Complement	$P(A^c) = 1 - P(A)$	"Not" and "at least one" questions
General addition	$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$	Any "or" question
Disjoint addition	$P(A \text{ or } B) = P(A) + P(B)$	Events that cannot co-occur
General multiplication	$P(A \text{ and } B) = P(A) \times P(B \mid A)$	Any "and" question, including without replacement
Independent multiplication	$P(A \text{ and } B) = P(A) \times P(B)$	Independent events; also the test for independence
Conditional	$P(A \mid B) = P(A \text{ and } B) \div P(B)$	"Given" questions
At least one in n independent tries	$1 - (1 - p)^n$	Repeated trials
Total probability	$P(B) = P(A)P(B \mid A) + P(A^c)P(B \mid A^c)$	Adding tree paths
Reverse a condition	$P(A \mid B) = \text{path}(A \text{ then } B) \div P(B)$	Screening, Bayes style questions

Worked examples

EXAMPLE 5.1 · COUNTING WITH TWO DICE

Roll two fair dice. Find P(sum = 8), P(sum ≤ 4) and P(sum is not 8).

STEP 1 · THE SAMPLE SPACE

Keep the dice distinct: $6 \times 6 = 36$ equally likely ordered pairs.

STEP 2 · SUM OF 8

(2,6), (3,5), (4,4), (5,3), (6,2): 5 pairs, so $P = 5/36 \approx$ **0.139**.

STEP 3 · SUM AT MOST 4

Sum 2: 1 pair; sum 3: 2 pairs; sum 4: 3 pairs; total 6. $P = 6/36 = 1/6 \approx$ **0.167**.

STEP 4 · COMPLEMENT

P(not 8) = 1 − 5/36 = 31/36 ≈ **0.861**.

TRAPS PEOPLE FALL INTO

- **Listing three outcomes for two coins.** HT and TH are different outcomes. S has four.
- **Thinking an outcome is "due."** Independent trials have no memory. The law of large numbers works by dilution, not correction.
- **Forgetting to subtract the overlap.** P(A or B) needs − P(A and B) unless the events are disjoint.
- **Calling disjoint events independent.** Disjoint events with nonzero probability are dependent.
- **Multiplying when events are dependent.** Use P(B | A), especially when sampling without replacement.
- **Adding probabilities across repeated tries.** Use 1 − P(none) for "at least one."
- **Swapping P(A | B) and P(B | A).** Ask "what is my denominator now?" before you divide.
- **Reading cause from a two-way table.** A table shows association. Cause needs random assignment.

Chapter 5 self-test · six items

1. P(rain) = 0.35. What is P(no rain)?

A. 0.35

B. 1.35

C. 0.50

D. 0.65

D. 0.65 Complement rule: 1 − 0.35 = 0.65.

2. One card is drawn. What is P(heart or face card)?

A. 25/52

B. 22/52

C. 3/52

D. 13/52

B. 22/52 13/52 + 12/52 − 3/52 = 22/52; the three face hearts are counted only once.

3. P(A) = 0.6 and P(B | A) = 0.5. What is P(A and B)?

A. 0.30
B. 1.10
C. 0.83
D. 0.10

A. 0.30 General multiplication rule: 0.6 × 0.5 = 0.30.

4. P(A and B) = 0.12 and P(B) = 0.40. What is P(A | B)?

A. 0.048
B. 0.52
C. 0.30
D. 0.28

C. 0.30 0.12 ÷ 0.40 = 0.30.

5. Which statement shows the confusion of the inverse?

A. Using 1 − P(none) for an "at least one" question
B. Treating P(positive | disease) as P(disease | positive)
C. Subtracting the overlap in an "or" question
D. Multiplying the probabilities along a tree path

B. Treating P(positive | disease) as P(disease | positive) P(A | B) and P(B | A) usually differ, sometimes by a lot. Screening results are the classic case.

6. A coin has landed heads 6 times in a row. What is P(heads) on the next fair flip?

A. Less than 0.5
B. More than 0.5
C. 1/64
D. 0.5

D. 0.5 Flips are independent. The coin has no memory, so the probability stays 0.5.

CHAPTER 6

Discrete Random Variables

Read it in full: **Chapter 6 in the textbook**. Practice it in the lab: **Valid Probability Distribution** and **Poisson and Geometric**.

What to know

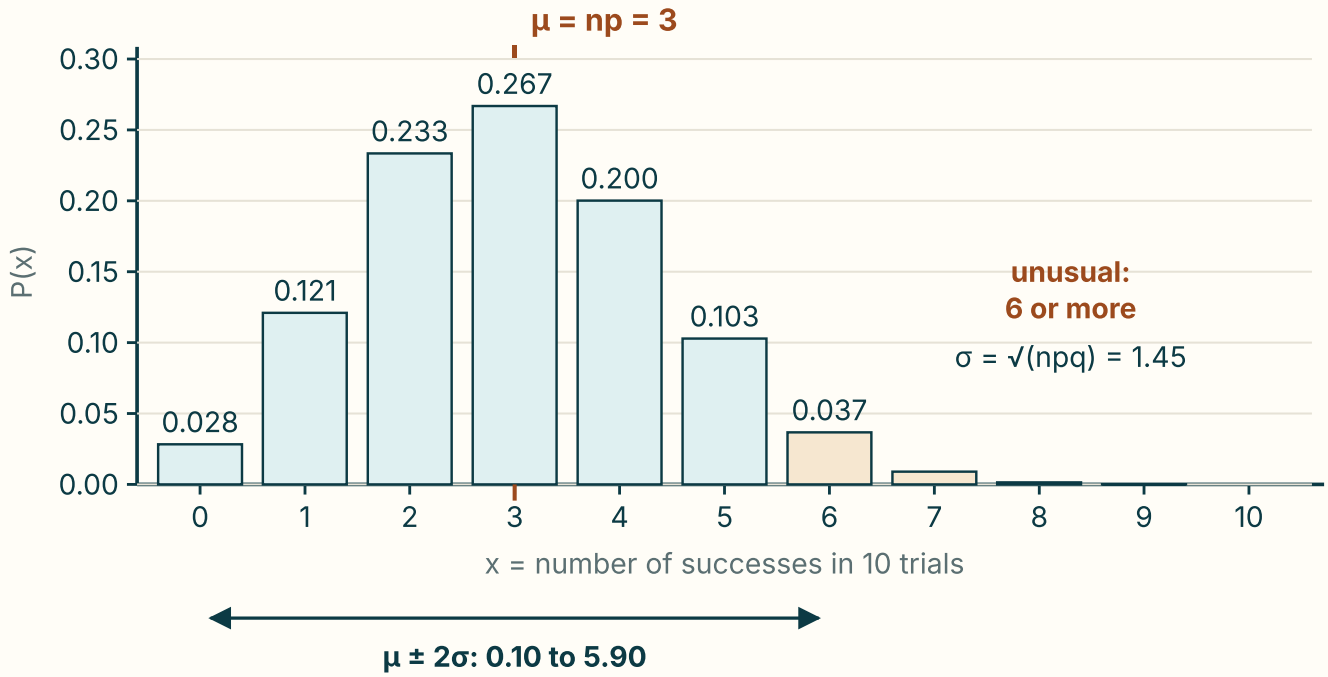
A random variable is discrete (counts, even an endless list) or continuous (measurements, intervals only). A distribution is valid when $\Sigma P(x) = 1$ and every $P(x)$ is between 0 and 1.

$\mu = \Sigma x \cdot P(x)$ is the long-run average; $\sigma^2 = \Sigma x^2 P(x) - \mu^2$. Values outside $\mu \pm 2\sigma$ are unusual. Expected value prices a bet: a $3 raffle ticket with one $1,000 prize in 500 is worth −$1 on average.

Binomial (fixed n, two outcomes, independent, constant p): $P(x) = {}_nC_x p^x q^{n-x}$, $\mu = np$, $\sigma = \sqrt{(npq)}$. Poisson: $\lambda^x e^{-\lambda}/x!$, mean = variance = λ. Geometric: $q^{x-1}p$, $\mu = 1/p$.

LOCK THESE IN

- Valid distribution: $\Sigma P(x) = 1$ and $0 \le P(x) \le 1$ for every x.
- $\mu = E(X) = \Sigma x \cdot P(x)$; $\sigma^2 = \Sigma x^2 \cdot P(x) - \mu^2$; $\sigma = \sqrt{\sigma^2}$.
- Expected value is the long-run average, not the next result.
- Binomial needs fixed n, two outcomes, independence and constant p.
- $P(X = x) = {}_nC_x\, p^x\, q^{n-x}$; $\mu = np$; $\sigma = \sqrt{(npq)}$.
- Poisson: $P(x) = \lambda^x e^{-\lambda} \div x!$; mean = variance = λ.
- Geometric: $P(x) = q^{x-1}p$; $\mu = 1/p$; $P(X > k) = q^k$.
- Unusual: outside $\mu \pm 2\sigma$.



A binomial distribution drawn to scale. With n = 10 and p = 0.3 the bars peak at 3, the mean np. The standard deviation is √(10 × 0.3 × 0.7) ≈ 1.45, so counts of 6 or more fall outside μ ± 2σ.

Chapter 6 formulas and facts

Quantity	Formula	When to use it
Validity	$\Sigma P(x) = 1$; $0 \le P(x) \le 1$	Checking any discrete table
Mean	$\mu = \Sigma[x \cdot P(x)]$	Long-run average of a random variable
Variance	$\sigma^2 = \Sigma[x^2 \cdot P(x)] - \mu^2$	Spread; take √ for σ
Unusual values	Outside $\mu \pm 2\sigma$	Range rule of thumb
Binomial probability	${}_nC_x\, p^x\, q^{n-x}$	Fixed n, two outcomes, independent, constant p
Combinations	${}_nC_x = n! \div [x!(n-x)!]$	Counting arrangements
Binomial mean and SD	$\mu = np$; $\sigma = \sqrt{(npq)}$	Exact shortcuts
Poisson probability	$\lambda^x e^{-\lambda} \div x!$	Events in a fixed interval at rate λ
Poisson mean and SD	$\mu = \sigma^2 = \lambda$; $\sigma = \sqrt{\lambda}$	Variance far above the mean means Poisson fits poorly
Geometric probability	$q^{x-1}\, p$	Trial of the first success
Geometric mean and SD	$\mu = 1/p$; $\sigma = \sqrt{q} \div p$	Average wait for the first success

Worked examples

EXAMPLE 6.1 · A MISSING PROBABILITY, THE MEAN AND THE SD

X is the number of streaming services a household pays for, with P(0) = 0.10, P(1) = 0.25, P(2) = k, P(3) = 0.20, P(4) = 0.10. Find k, μ and σ.

STEP 1 · FIND K

0.10 + 0.25 + k + 0.20 + 0.10 = 1, so 0.65 + k = 1 and **k = 0.35**, a legal probability.

STEP 2 · MEAN

(0)(0.10) + (1)(0.25) + (2)(0.35) + (3)(0.20) + (4)(0.10) = **μ = 1.95**.

STEP 3 · $\Sigma X^2 P(X)$

$(0)(0.10) + (1)(0.25) + (4)(0.35) + (9)(0.20) + (16)(0.10) = 5.05$.

STEP 4 · VARIANCE AND SD

$\sigma^2 = 5.05 - 1.95^2 = 5.05 - 3.8025 = 1.2475$, so $\sigma \approx$ **1.12**.

STEP 5 · UNUSUAL?

$\mu \pm 2\sigma = -0.28$ to 4.18. Every possible value lies inside, so none is unusual.

TRAPS PEOPLE FALL INTO

- **Thinking discrete means "few values."** An endless list 0, 1, 2, ... is still discrete.
- **Checking only the sum.** A table that sums to 1 with a negative entry is invalid. Check both rules.
- **Dividing by n after $\Sigma x \cdot P(x)$.** The weights already sum to 1. No final division.
- **Reading E(X) as the next result.** It is the long-run average and may be impossible on one trial.
- **Using the binomial without fixed n.** "Until the first success" is geometric, and "per hour" with no cap is Poisson.
- **Forgetting the coefficient.** $p^x q^{n-x}$ is one arrangement. Multiply by ${}_nC_x$.
- **Mismatching p and "success."** If you count defects, p is the chance of a defect.
- **Believing a success is overdue.** The geometric is memoryless; past failures change nothing.

Chapter 6 self-test · six items

1. P(0) = 0.3, P(1) = 0.3, P(2) = k. What is k?

A. 0.6

B. 0.4

C. 0.3

D. 1.0

B. 0.4 The probabilities must sum to 1: $1 - 0.6 = 0.4$.

2. E(X) for a raffle ticket is −$1.00. What does that mean?

A. Every buyer loses exactly $1 on each ticket
B. The prize is worth $1 less than the ticket price
C. On average, buyers lose about $1 per ticket
D. One ticket in every 100 wins a prize

C. On average, buyers lose about $1 per ticket Expected value is the long-run average result, not what happens on any single ticket.

3. n = 4, p = 0.5. What is P(X = 4)?

A. 0.5000
B. 0.2500
C. 0.0625
D. 0.1250

C. 0.0625 ${}_4C_4 = 1$ and $0.5^4 = 0.0625$.

4. A binomial has n = 100 and p = 0.1. What is σ?

A. 9
B. 10
C. 0.3
D. 3

D. 3 $\sigma = \sqrt{(100 \times 0.1 \times 0.9)} = \sqrt{9} = 3$.

5. For a Poisson count with λ = 9, what is σ?

A. 9
B. 81
C. 4.5
D. 3

D. 3 Poisson variance equals λ, so σ = √9 = 3.

6. With p = 0.3, what is P(the first success comes on trial 2)?

A. 0.21
B. 0.09
C. 0.30
D. 0.49

A. 0.21 One failure then a success: 0.7 × 0.3 = 0.21.

CHAPTER 7

The Normal Distribution

Read it in full: **Chapter 7 in the textbook**. Practice it in the lab: **Normal Curve** and **Standardizing**.

What to know

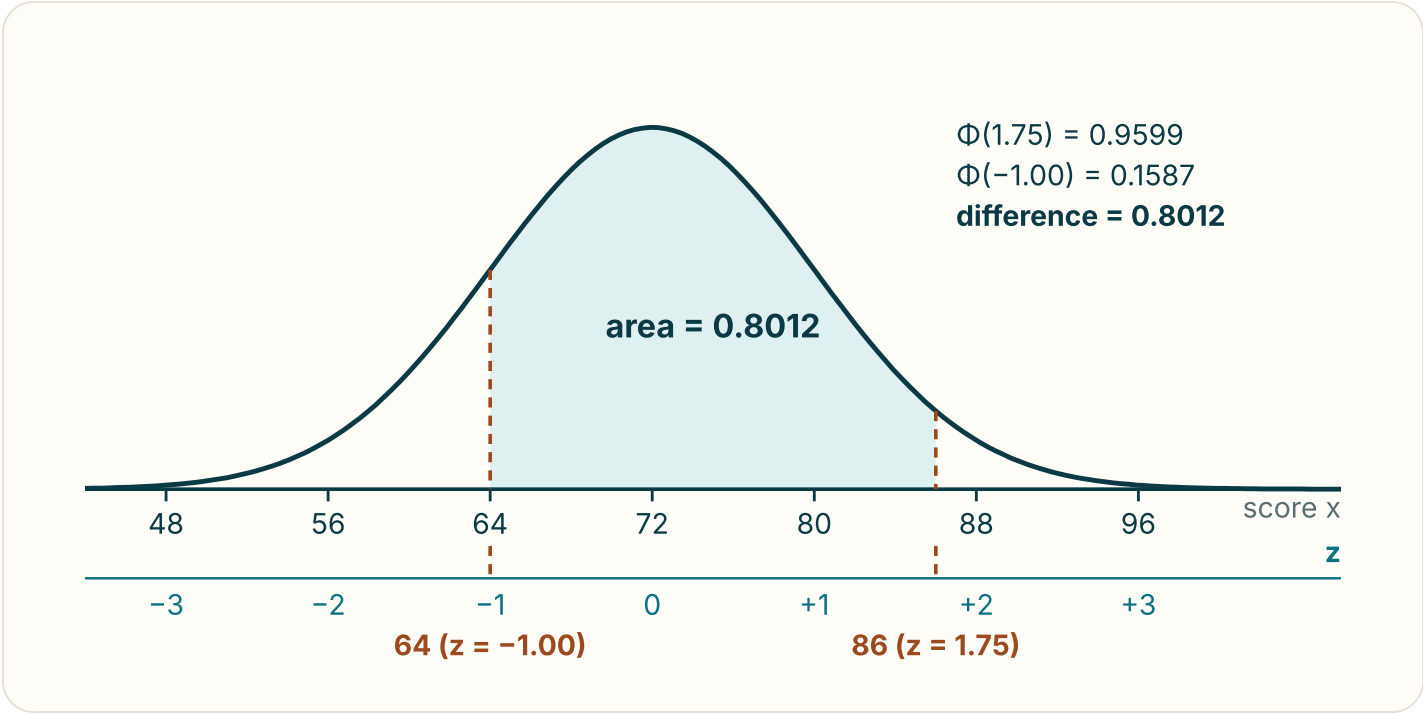
A normal curve is symmetric and fixed by μ and σ. Area is probability, the total is 1, and a single exact value has probability 0. The empirical rule gives 68, 95, 99.7% within 1, 2, 3 SD, with blocks of 34, 13.5, 2.35 and 0.15% on each side.

Standardize with $z = (x - \mu) \div \sigma$. The table gives left areas $\Phi(z)$: right tail = $1 - \Phi(z)$; between = $\Phi(z_b) - \Phi(z_a)$; $\Phi(-z) = 1 - \Phi(z)$.

Backward: area → z → $x = \mu + z\sigma$. "Top 10%" is a left area of 0.90, and percentiles below 50 need a negative z. The middle 95% uses ±1.96.

LOCK THESE IN

- Normal is fixed by μ and σ; area under the curve is probability.
- 68, 95, 99.7 within 1, 2, 3 SD; blocks 34, 13.5, 2.35, 0.15 each side.
- $z = (x - \mu) \div \sigma$; $x = \mu + z\sigma$.
- Left: $\Phi(z)$. Right: $1 - \Phi(z)$. Between: $\Phi(z_b) - \Phi(z_a)$.
- $\Phi(-z) = 1 - \Phi(z)$, and $P(X < a) = P(X \leq a)$ for a continuous variable.
- Key z values: 1.645 (top 5%), 1.96 (middle 95%), 2.576 (middle 99%), 1.28 (90th percentile).
- Top k% means a left-tail area of $1 - k/100$.
- Standardizing never makes skewed data normal.



Area is probability. For scores with μ = 72 and σ = 8, the share between 64 and 86 is $\Phi(1.75) - \Phi(-1.00) = 0.9599 - 0.1587 = 0.8012$.

Chapter 7 formulas and facts

Quantity	Formula	When to use it
z-score	$z = (x - \mu) \div \sigma$	Convert a value to standard units
Back to x	$x = \mu + z\sigma$	Percentiles and cutoffs
Left tail	$P(X < a) = \Phi(z_a)$	"Below," "less than," "at most"
Right tail	$P(X > a) = 1 - \Phi(z_a)$	"Above," "more than," "at least"
Between	$\Phi(z_b) - \Phi(z_a)$	"Between a and b"
Symmetry	$\Phi(-z) = 1 - \Phi(z)$	Tables with only positive z
Middle C%	$z = \pm z^*$ with $(1 - C)/2$ in each tail	Symmetric intervals
Chebyshev	at least $1 - 1 \div k^2$	Shape unknown or not normal

Worked examples

EXAMPLE 7.1 · THE EMPIRICAL RULE WITH NO TABLE

Newborn weights are about normal with $\mu = 3{,}400$ g and $\sigma = 500$ g. What percent weigh more than 4,400 g, and what percent weigh between 2,900 and 4,400 g?

STEP 1 · BUILD THE RULER

$\mu \pm 1\sigma$: 2,900 to 3,900. $\mu \pm 2\sigma$: 2,400 to 4,400. $\mu \pm 3\sigma$: 1,900 to 4,900.

STEP 2 · ABOVE 4,400

4,400 is $\mu + 2\sigma$. Outside $\pm 2\sigma$ lies 5%, half on each side: **2.5%**.

STEP 3 · SPLIT AT THE MEAN

From 2,900 ($\mu - 1\sigma$) to 3,400 is 34%. From 3,400 to 4,400 ($\mu + 2\sigma$) is 47.5%.

STEP 4 · ADD

34% + 47.5% = **81.5%**.

CHECK

The answer is above the 68% inside $\pm 1\sigma$ and below the 95% inside $\pm 2\sigma$, as it should be.

TRAPS PEOPLE FALL INTO

- **Treating curve height as probability.** Only area is probability. A single exact value has probability 0.
- **Using the empirical rule on skewed data.** It holds for approximately normal data only.
- **Reading a right tail straight from the table.** The table gives left areas. Subtract from 1.
- **Turning "top 10%" into an area of 0.10.** The cutoff has 90% below it. Look up 0.90.
- **Dropping the minus sign.** A percentile below 50 has a negative z and lands below the mean.
- **Using 2 for the exact 95% middle.** The exact value is 1.96. 2 gives 95.45%.
- **Believing standardizing fixes shape.** z-scores of skewed data are still skewed.
- **Mixing "middle 95%" with "95% confidence."** One describes individuals with known μ and σ; the other is about estimating an unknown parameter.

Chapter 7 self-test · six items

1. About what percent of a normal distribution lies above μ + 1σ?

A. 34%

B. 32%

C. 84%

D. 16%

D. 16% 68% lies within ±1σ, leaving 32% split evenly: 16% above.

2. A value has x = 85 with μ = 70 and σ = 6. What is z?

A. 2.50

B. 1.50

C. 15.00

D. 0.40

A. 2.50 (85 − 70) ÷ 6 = 15 ÷ 6 = 2.50.

3. Φ(0.85) = 0.8023. What share lies above z = 0.85?

A. 0.8023
B. 0.3023
C. 0.6046
D. 0.1977

D. 0.1977 Right tail = 1 − 0.8023 = 0.1977.

4. The 75th percentile of N(60, 10) is about:

A. 53.3
B. 66.7
C. 67.5
D. 70.0

B. 66.7 z ≈ 0.674, so x = 60 + 0.674(10) ≈ 66.7.

5. Which z values bound the middle 99%?

A. ±1.96
B. ±3.00
C. ±2.576
D. ±2.326

C. ±2.576 0.5% in each tail gives Φ(2.576) = 0.9950.

6. Which statement about the normal model is correct?

A. It fits any data with a mean
B. Its median is above its mean
C. It is fixed entirely by μ and σ
D. Its total area depends on σ

C. It is fixed entirely by μ and σ Two parameters pin down the whole curve; the total area is always 1 and the mean equals the median.

CHAPTER 8

Sampling Distributions and the Central Limit Theorem

Read it in full: **Chapter 8 in the textbook**. Practice it in the lab: **Distribution of the Sample Mean** and **Sampling Distribution of a Proportion**.

What to know

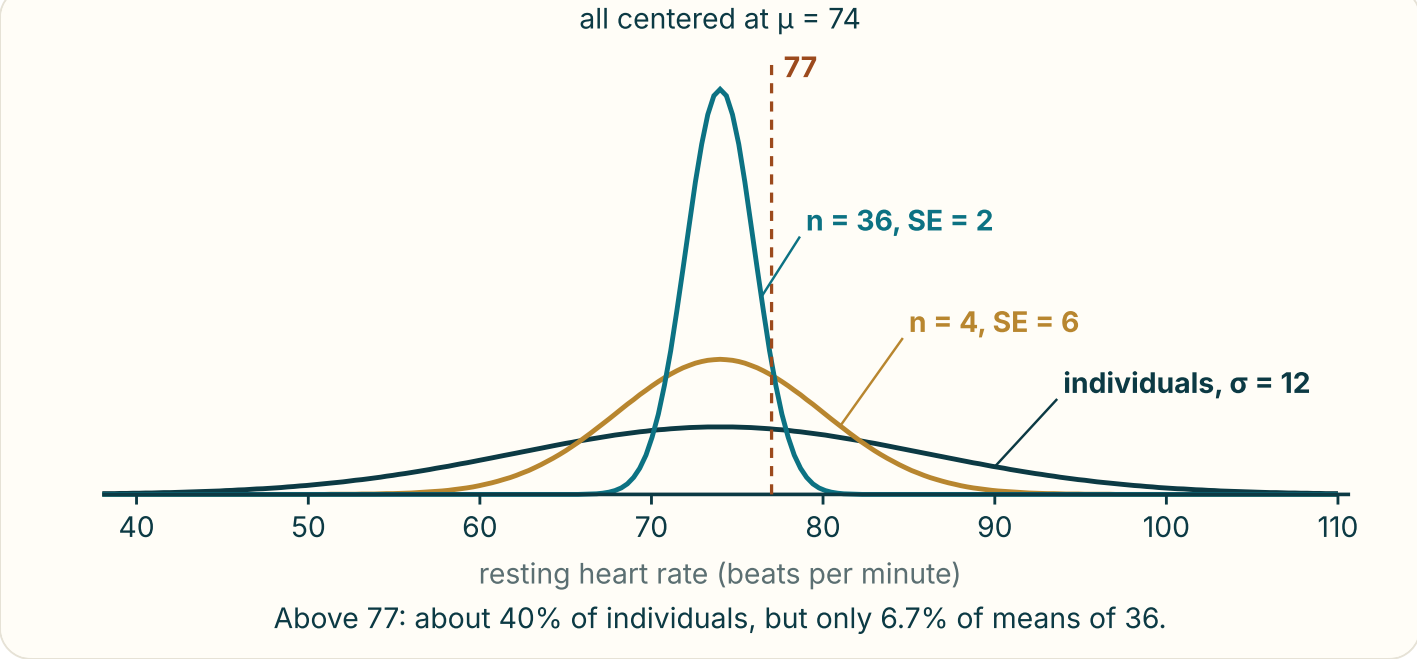
Population, sample data and sampling distribution are different: one dot in a sampling distribution is a whole sample. $\mu_{\bar{x}} = \mu$ (unbiased) and $\sigma_{\bar{x}} = \sigma \div \sqrt{n}$ with independence.

The CLT: $\bar{x}$ is approximately Normal(μ, $\sigma \div \sqrt{n}$) for large n from any population with finite σ. It changes only the shape of the sampling distribution, never the population or your data. $n \geq 30$ is a rule of thumb; more skew needs more n.

SE describes an estimate and shrinks as $1 \div \sqrt{n}$; SD describes individuals. For proportions, SE = $\sqrt{[p(1 - p) \div n]}$, normal when $np \geq 10$ and $n(1 - p) \geq 10$.

LOCK THESE IN

- One dot in a sampling distribution is one whole sample.
- $\mu_{\bar{x}} = \mu$ and $\sigma_{\bar{x}} = \sigma \div \sqrt{n}$ (independence: n under 10% of N).
- One person: divide by σ. A sample mean: divide by $\sigma \div \sqrt{n}$.
- CLT: $\bar{x} \approx \text{Normal}(\mu, \sigma \div \sqrt{n})$ for large n, any shape with finite σ.
- The CLT never makes the population or the data normal.
- SE shrinks as $1 \div \sqrt{n}$; SD does not shrink with n.
- Proportions: $\mu_{\hat{p}} = p$; $\text{SE} = \sqrt{[p(1 - p) \div n]}$; need $np \geq 10$ and $n(1 - p) \geq 10$.
- Margin of error ≈ 2 × SE and ignores bias.



Averages vary less than individuals. The sampling distribution of $\bar{x}$ keeps the center μ = 74 but narrows to $\sigma \div \sqrt{n}$: 6 for n = 4 and 2 for n = 36. A threshold that 40% of people pass is cleared by only 6.7% of sample means.

Chapter 8 formulas and facts

Quantity	Formula	When to use it
Center of $\bar{x}$	$\mu_{\bar{x}} = \mu$	Always; $\bar{x}$ is unbiased
Spread of $\bar{x}$	$\sigma_{\bar{x}} = \sigma \div \sqrt{n}$	Independent draws
z for a mean	$z = (\bar{x} - \mu) \div (\sigma \div \sqrt{n})$	Questions about a sample mean
z for one value	$z = (x - \mu) \div \sigma$	Questions about one individual
Sum of n draws	Normal($n\mu$, $\sigma\sqrt{n}$)	Totals, by the CLT
Estimated SE	$s \div \sqrt{n}$	σ unknown
n for a target SE	$n = (s \div SE)^2$, round up	Planning precision
Finite population correction	$\times \sqrt{[(N - n) \div (N - 1)]}$	Sample over 10% of N
Proportion center and SE	$\mu_{\hat{p}} = p$; $\sqrt{[p(1 - p) \div n]}$	Categorical data
Success-failure	$np \geq 10$ and $n(1 - p) \geq 10$	Before using a normal model for $\hat{p}$
Margin of error (rough)	$\approx 2 \times SE$	Random error only

Worked examples

EXAMPLE 8.1 · A SAMPLING DISTRIBUTION BUILT BY HAND

A population holds the values 1, 3, 5. List every sample of size 2 drawn with replacement, and find the mean and variance of $\bar{x}$.

STEP 1 · PARAMETERS

$\mu = 9 \div 3 = 3$. Squared deviations 4, 0, 4 sum to 8, so $\sigma^2 = 8 \div 3 \approx 2.667$.

STEP 2 · THE 9 SAMPLES

(1,1), (1,3), (1,5), (3,1), (3,3), (3,5), (5,1), (5,3), (5,5), each with probability 1/9.

STEP 3 · THEIR MEANS

$\bar{x}$ values: 1 (1 way), 2 (2 ways), 3 (3 ways), 4 (2 ways), 5 (1 way).

STEP 4 · MEAN OF $\bar{X}$

$\Sigma\bar{x} \div 9 = 27 \div 9 =$ **3**, exactly μ.

STEP 5 · VARIANCE OF $\bar{X}$

$\Sigma(\bar{x} - 3)^2 \div 9 = 12 \div 9 \approx$ **1.333**, which is $\sigma^2 \div 2 = 2.667 \div 2$. The square-root law falls out of counting.

TRAPS PEOPLE FALL INTO

- **Calling the sample histogram the sampling distribution.** Your data show individuals. The sampling distribution shows statistics from many samples.
- **Dividing by √n for one individual.** One person uses σ. Only a sample mean uses $\sigma \div \sqrt{n}$.
- **Saying the CLT makes the data normal.** Your 5,000 incomes stay skewed. Only the distribution of $\bar{x}$ goes normal.
- **Treating n ≥ 30 as a law.** Heavier skew needs more; normal populations need none.
- **Mixing up SD and SE.** Reporting mean ± SE to describe people makes the data look far tidier than they are.
- **Thinking SE depends on population size.** With a small sampling fraction, N drops out.
- **Skipping the success-failure check.** With p near 0 or 1 and small n, the normal model for $\hat{p}$ fails.
- **Trusting the margin of error to cover bias.** It counts random sampling error only.

Chapter 8 self-test · six items

1. Which is true of $\bar{x}$ as an estimator of μ?

A. It always equals μ exactly for n ≥ 30

B. Its sampling distribution centers on μ

C. It is biased low for skewed populations

D. It is centered at μ only if n is large

B. Its sampling distribution centers on μ $\mu_{\bar{x}} = \mu$ for every n and every population shape.

2. $\sigma = 30$ and $n = 9$. What is $\sigma_{\bar{x}}$?

A. 3.33
B. 30
C. 90
D. 10

D. 10 $30 \div \sqrt{9} = 30 \div 3 = 10$.

3. When does $\bar{x}$ have an exactly normal distribution for any n?

A. When n is at least 30
B. When the population is normal
C. When σ is known
D. When the sample is random

B. When the population is normal A normal population gives a normal $\bar{x}$ at every sample size.

4. $p = 0.2$ and $n = 400$. What is $SE(\hat{p})$?

A. 0.04
B. 0.0004
C. 0.02
D. 0.16

C. 0.02 $\sqrt{(0.2 \times 0.8 \div 400)} = \sqrt{0.0004} = 0.02$.

5. Which value of p gives the largest SE(p̂) for a fixed n?

A. 0.5
B. 0.1
C. 0.9
D. 0.99

A. 0.5 p(1 − p) peaks at p = 0.5.

6. A margin of error of ±3 points means:

A. The estimate is certainly within 3 points of the truth
B. Bias in the survey is under 3 points
C. There is a 3% chance the poll is wrong
D. Random sampling error is about ±3 points if the method is clean

D. Random sampling error is about ±3 points if the method is clean A margin of error budgets chance only and assumes no bias.

CHAPTER 9

Confidence Intervals

Read it in full: **Chapter 9 in the textbook**. Practice it in the lab: **Z Interval or T Interval** and **Margin of Error, n and Confidence**.

What to know

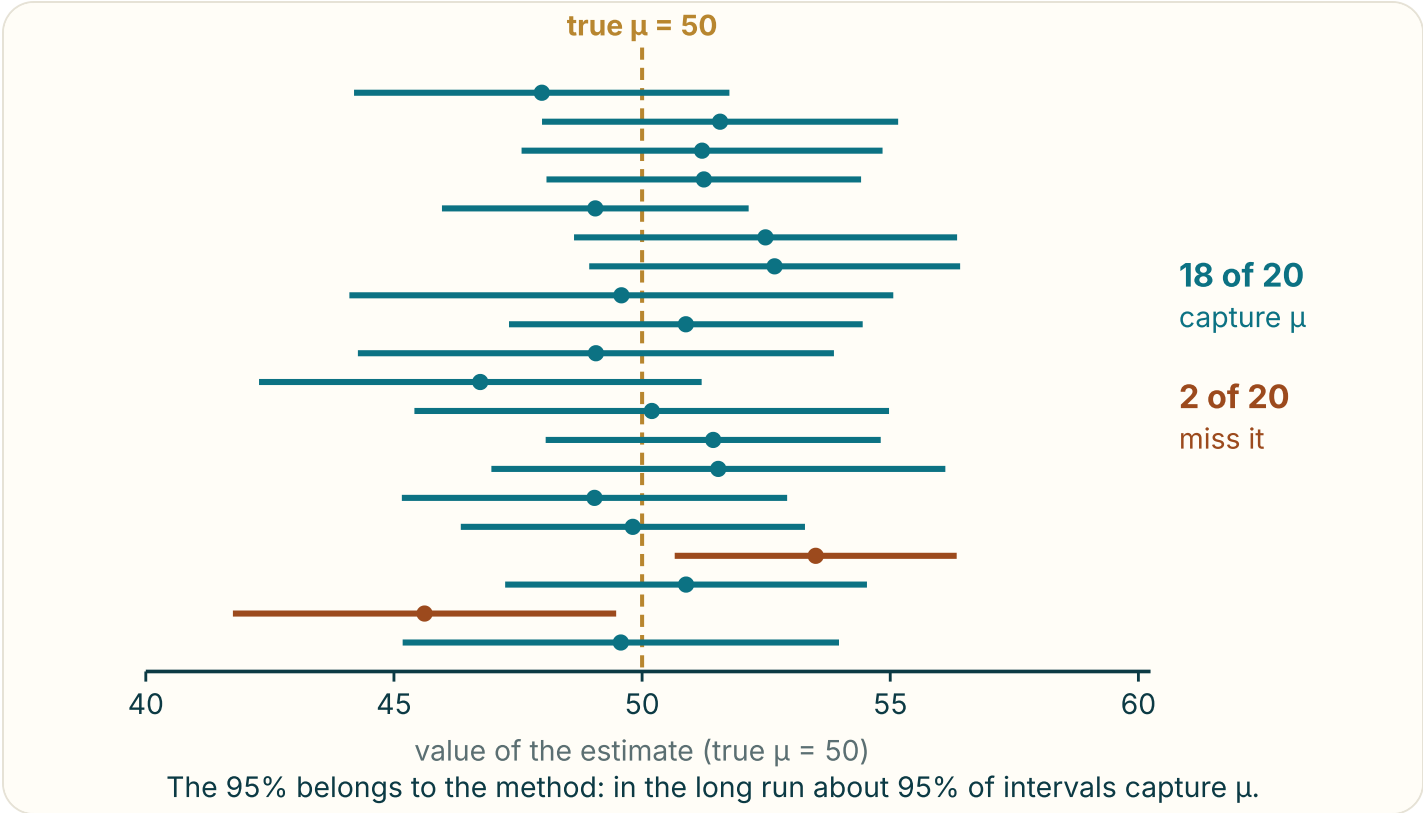
Every interval is estimate ± critical value × SE, and its width is twice the margin of error. It estimates a parameter, not individuals.

Means: $\bar{x} \pm z^*(\sigma \div \sqrt{n})$ if σ is known; $\bar{x} \pm t^*(s \div \sqrt{n})$, df = n − 1, if you computed s. Proportions: $\hat{p} \pm z^*\sqrt{[\hat{p}(1 - \hat{p}) \div n]}$ with $n\hat{p} \geq 10$ and $n(1 - \hat{p}) \geq 10$. Higher confidence widens the interval; quadrupling n halves it.

Plan with $n = (z^*\sigma \div E)^2$ or $n = z^{*2}p(1 - p) \div E^2$ ($p = 0.5$ if unknown), rounding up. "95% confident" describes the long-run procedure, not a probability for this interval, and the margin covers chance only.

LOCK THESE IN

- CI = estimate ± critical value × SE; width = 2 × margin of error.
- σ known: $\bar{x} \pm z^*\sigma/\sqrt{n}$. σ unknown: $\bar{x} \pm t^*s/\sqrt{n}$ with df = n − 1.
- Proportion: $\hat{p} \pm z^*\sqrt{[\hat{p}(1 - \hat{p})/n]}$; need $n\hat{p} \geq 10$ and $n(1 - \hat{p}) \geq 10$.
- z^* = 1.645 (90%), 1.96 (95%), 2.576 (99%).
- Quadruple n to halve the margin. Higher confidence, wider interval.
- $n = (z^*\sigma/E)^2$ for a mean; $n = z^{*2}p(1 - p)/E^2$ for a proportion; round up.
- "95% confident" describes the method; this interval either contains μ or not.
- A CI excluding a value matches a two-sided test rejecting it at $\alpha = 1 - C$.



Confidence is about the procedure. Twenty samples of n = 25 from a population with μ = 50 give twenty different 95% intervals. 18 capture μ and 2 miss; you never know which kind yours is.

Chapter 9 formulas and facts

Interval or plan	Formula	When to use it
Skeleton	estimate ± (critical value)(SE)	Every interval
Mean, σ known	$\bar{x} \pm z^*(\sigma \div \sqrt{n})$	σ given from outside the data
Mean, σ unknown	$\bar{x} \pm t^*(s \div \sqrt{n})$, df = n − 1	You computed s from the sample
Proportion	$\hat{p} \pm z^*\sqrt{[\hat{p}(1-\hat{p}) \div n]}$	$n\hat{p} \geq 10$ and $n(1-\hat{p}) \geq 10$
Margin from an interval	E = (upper − lower) ÷ 2	Reading a reported CI
Estimate from an interval	center = (upper + lower) ÷ 2	Reading a reported CI
n for a mean	$n = (z^*\sigma \div E)^2$	Planning; round up
n for a proportion	$n = z^{*2}\, p(1-p) \div E^2$	Use p = 0.5 if unknown; round up

Worked examples

EXAMPLE 9.1 · A Z INTERVAL WITH Σ KNOWN

A scale's long-run σ is 5 g. A random sample of n = 49 packages has $\bar{x}$ = 68.3 g. Build a 95% interval for μ.

STEP 1 · SE

$\sigma \div \sqrt{n} = 5 \div 7 = 0.7143$.

STEP 2 · MARGIN

$1.96 \times 0.7143 =$ **1.40**.

STEP 3 · INTERVAL

68.3 ± 1.40: **66.90 to 69.70 g**.

STEP 4 · SAY IT

We are 95% confident that the interval from 66.90 to 69.70 g contains the true mean package weight.

TRAPS PEOPLE FALL INTO

- **Saying "95% chance μ is in this interval."** μ is fixed. The 95% describes how often the method works.

- **Reading the CI as the range of the data.** An interval for a mean is far narrower than the spread of individuals.
- **Using z after computing s.** If s came from the sample, use t with df = n − 1.
- **Using t for a proportion.** The proportion interval uses z.
- **Calling the width the margin of error.** The width is twice the margin.
- **Mixing percent and percentage points.** A 4-point margin on 37% runs 33% to 41%.
- **Rounding a planned n down.** Always round up to keep the promised margin.
- **Thinking a bigger n removes bias.** The margin of error covers chance only.

Chapter 9 self-test · six items

1. A 95% interval runs from 20.4 to 26.0. What are the estimate and margin?

A. 23.2 and 5.6

B. 20.4 and 5.6

C. 26.0 and 2.8

D. 23.2 and 2.8

D. 23.2 and 2.8 Center = (20.4 + 26.0) ÷ 2 = 23.2; margin = (26.0 − 20.4) ÷ 2 = 2.8.

2. Which critical value goes with a 99% z interval?

A. 1.96

B. 2.576

C. 1.645

D. 3.00

B. 2.576 0.5% in each tail gives z* = 2.576.

3. Why is t* larger than z* for small samples?

A. Small samples are always biased
B. Estimating σ with s adds uncertainty
C. The t curve has a higher peak
D. The SE is smaller for small n

B. Estimating σ with s adds uncertainty Heavier t tails widen the interval to pay for estimating σ.

4. σ ≈ 20, E = 4, 95% confidence. What n is needed?

A. 96
B. 10
C. 97
D. 49

C. 97 $(1.96 \times 20 \div 4)^2 = 9.8^2 = 96.04$, rounded up to 97.

5. $\hat{p} = 0.08$ with n = 100. Which condition fails?

A. $n\hat{p} \geq 10$
B. $n(1 - \hat{p}) \geq 10$
C. The 10% rule
D. None of them

A. $n\hat{p} \geq 10$ $n\hat{p} = 8$, under 10, so the normal-based interval is not trustworthy.

6. A survey used volunteers who clicked a link. What does its margin of error cover?

A. Nothing useful, as the sample is not random
B. Both chance error and the selection bias
C. Only the selection bias in who responded
D. The full error of the estimate, bias included

A. Nothing useful, as the sample is not random The formula assumes random sampling. Self-selected samples make the margin meaningless.

CHAPTER 10

The Logic of Hypothesis Testing

Read it in full: **Chapter 10 in the textbook**. Practice it in the lab: **What a P-Value Is** and **Statistical vs Practical Significance**.

What to know

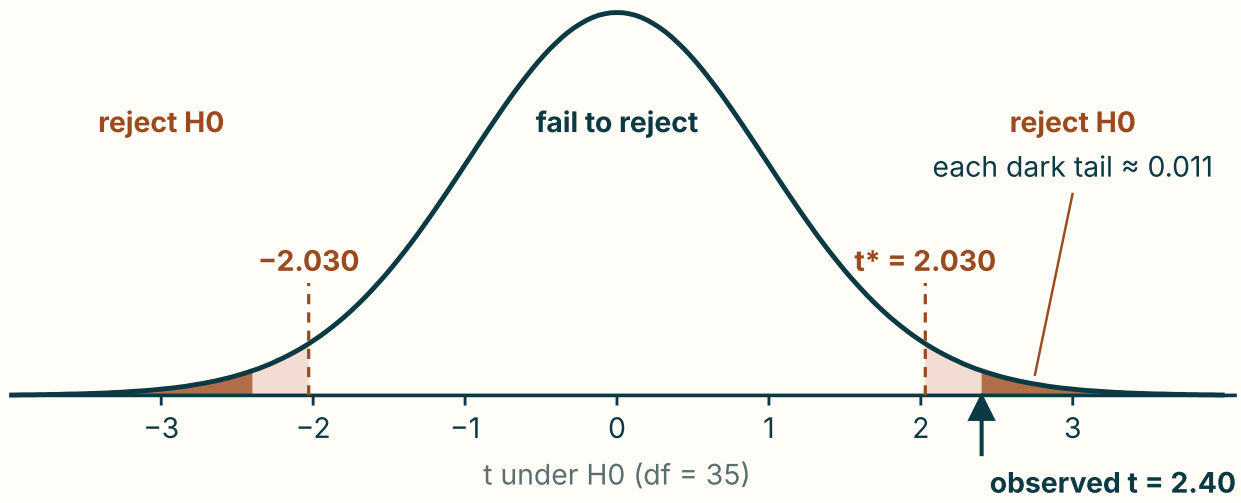
H0 is the no-effect claim with the equals sign; Ha carries the burden. Both are about parameters, and the tails are chosen before looking. The p-value is P(data this extreme or more | H0 and assumptions), found from SE, test statistic and tail area.

Set α first. p ≤ α: reject H0; p > α: fail to reject, never accept. Type I error (false alarm) has probability α; Type II (miss) has probability β; power = 1 − β against a stated effect size.

Large n makes trivial differences significant, and small n can miss big ones. Report the estimate, the CI and an effect size such as d = difference ÷ s, and remember significance never establishes cause.

LOCK THESE IN

- H0 has the equals sign; hypotheses are about parameters.
- Choose one or two tails before seeing the data; default to two.
- p-value = P(result this extreme or more | H0 true).
- $p \le \alpha$: reject H0. $p > \alpha$: fail to reject. Never "accept."
- Type I = false alarm (α). Type II = miss (β). Power = $1 - \beta$.
- Power rises with n, effect size and α, and falls with σ.
- Test statistic = (estimate − null value) ÷ SE.
- Statistical significance is not practical importance, and not causation.



p-value = 2 × 0.0109 ≈ 0.022 ≤ α = 0.05, so reject H0.

Two routes to one verdict. For the commute test, t = 2.40 lies beyond the critical value 2.030, and the two dark tails add to p ≈ 0.022, below α = 0.05.

Chapter 10 formulas and facts

Quantity	Formula	When to use it
Test statistic (mean)	$t = (\bar{x} - \mu_0) \div (s \div \sqrt{n})$, $df = n - 1$	σ unknown
Test statistic (σ known)	$z = (\bar{x} - \mu_0) \div (\sigma \div \sqrt{n})$	σ given
Two-sided p-value	2 × (tail area beyond $\lvert t \rvert$)	Ha: ≠
One-sided p-value	One tail in the direction of Ha	Ha: < or >, chosen in advance
Decision	Reject H0 if $p \le \alpha$	α fixed beforehand
Rejection cutoffs (power)	$\mu_0 \pm z^*(\sigma \div \sqrt{n})$	Power calculations
Power	P(reject H0 | specific true μ)	Planning; aim for 0.80 or more
Cohen's d	$d = (\bar{x} - \mu_0) \div s$	Effect size beside the p-value

Worked examples

EXAMPLE 10.1 · A ONE-SAMPLE T TEST FROM START TO FINISH

A coffee chain says its drinks average 50 mg of caffeine per cup. A lab tests a random sample of n = 25 cups: $\bar{x} = 53.1$ mg, s = 7.5. Test at $\alpha = 0.05$, two-tailed.

STEP 1 · HYPOTHESES

H0: μ = 50; Ha: μ ≠ 50. Two-tailed, chosen before the data.

STEP 2 · SE

$s \div \sqrt{n} = 7.5 \div 5 = 1.50$.

STEP 3 · TEST STATISTIC

$t = (53.1 - 50) \div 1.50 = 3.1 \div 1.50 =$ **2.067**, df = 24.

STEP 4 · P-VALUE

One tail beyond 2.07 is 0.0249; doubled, **p ≈ 0.050**.

STEP 5 · DECIDE

0.050 is at or below 0.05, so **reject H0**. Critical-value route: $t^* = 2.064$, and 2.067 is beyond it.

STEP 6 · IN CONTEXT

If the true mean really were 50 mg, a sample mean this far away would happen only about 5.0% of the time. The data suggest the mean differs from 50 mg.

TRAPS PEOPLE FALL INTO

- **Writing hypotheses about $\bar{x}$ or $\hat{p}$.** You can see the statistic. Hypotheses are about μ or p.
- **Putting the inequality in H0.** The equals sign belongs in the null.
- **Choosing one tail after peeking.** It doubles the false-alarm rate. Decide first.
- **Reading p as P(H0 true).** p is P(data | H0), not P(H0 | data).
- **Saying "accept H0."** Say "fail to reject." The test cannot confirm the null.
- **Treating p = 0.051 as "no effect."** It is weak evidence; look at the interval.
- **Ignoring power.** A null result from a low-power study tells you little.
- **Equating significant with important or causal.** Check the effect size, and remember cause comes from the design.

Chapter 10 self-test · six items

1. Where does the equals sign go?

A. In the alternative hypothesis
B. In both hypotheses
C. In the null hypothesis
D. In neither hypothesis

C. In the null hypothesis H0 pins the parameter to one value so the probability model can be built.

2. t = 1.50 with a two-sided Ha. One tail beyond 1.50 has area 0.07. What is p?

A. 0.07
B. 0.93
C. 0.14
D. 0.035

C. 0.14 Two-sided p doubles the tail: 2 × 0.07 = 0.14.

3. α is best described as:

A. The Type I error rate, set in advance
B. The probability that H0 is true
C. The power of the test being run
D. The p-value of your one sample

A. The Type I error rate, set in advance α is the false-alarm budget, fixed before the data.

4. Lowering α from 0.05 to 0.01 with n fixed does what to β?

A. Raises it
B. Lowers it
C. Leaves it unchanged
D. Sets it to 0.01

A. Raises it A smaller rejection region makes misses more likely.

5. A mean difference of 3 with s = 12 gives Cohen's d of:

A. 4.00
B. 0.25
C. 0.36
D. 36.0

B. 0.25 d = 3 ÷ 12 = 0.25.

6. Which statement about significance is correct?

A. A small p-value proves a causal effect
B. Significant results will always replicate
C. Large effects are always statistically significant
D. Significance alone never establishes causation

D. Significance alone never establishes causation Causal claims come from randomized, controlled designs, not from the size of a test statistic.

CHAPTER 11

One-Sample Tests

Read it in full: **Chapter 11 in the textbook**. Practice it in the lab: **The T Distribution** and **Check the Assumptions**.

What to know

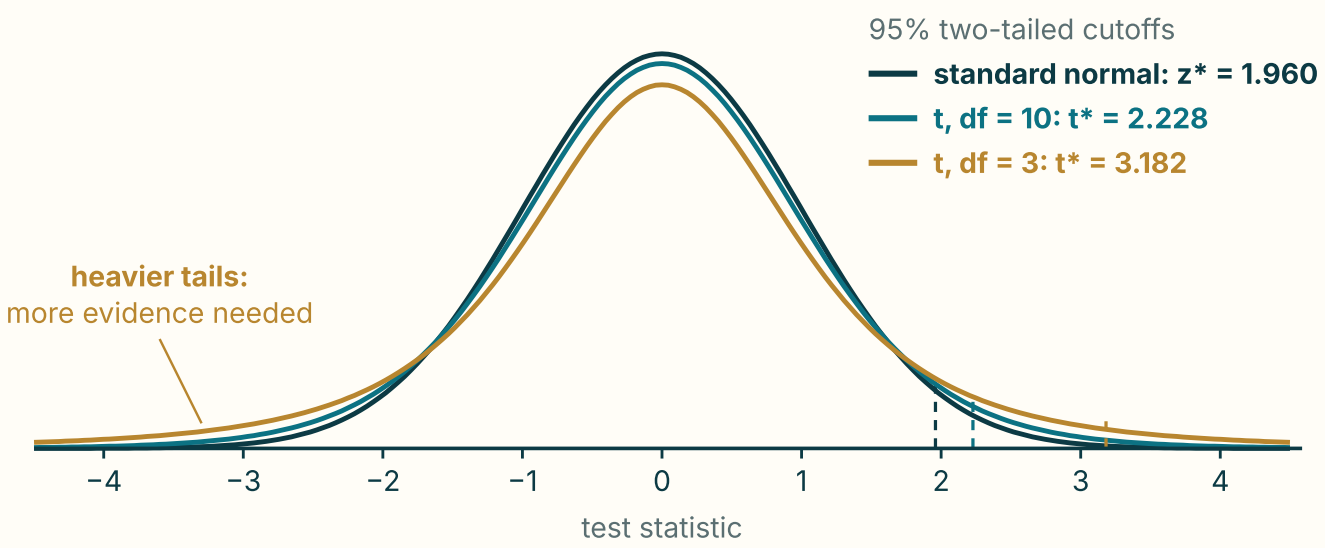
Five steps: hypotheses, α and conditions, test statistic (estimate − claim) ÷ SE, p-value, conclusion in context. Fail to reject, never accept.

Means: $z = (\bar{x} - \mu_0) \div (\sigma/\sqrt{n})$ only with σ known; otherwise $t = (\bar{x} - \mu_0) \div (s/\sqrt{n})$, df = n − 1, with heavier tails. A two-sided test at α matches the 100(1 − α)% CI.

Proportions: $z = (\hat{p} - p_0) \div \sqrt{[p_0(1 - p_0)/n]}$, with np_0 and $n(1 - p_0)$ at least 10; the interval uses $\hat{p}$ instead. Randomness is the condition nothing can repair.

LOCK THESE IN

- Five steps: hypotheses, α and conditions, statistic, p-value, conclusion in context.
- Test statistic = (estimate − claim) ÷ SE.
- z-test only when σ is truly known; otherwise t with df = n − 1.
- t has heavier tails than z; t* shrinks toward 1.96 as df grows.
- Proportion test SE uses p_0; the CI uses $\hat{p}$.
- Check $np_0 \geq 10$ and $n(1 - p_0) \geq 10$.
- Two-sided test rejects at α exactly when the $100(1 - \alpha)\%$ CI excludes the claim.
- Random sampling, independence, normal enough: randomness cannot be repaired.



As df grows, t* shrinks toward 1.96 and the t curve becomes the normal curve.

Estimating σ costs heavier tails. With df = 3 a t statistic must pass 3.182 to reject at α = 0.05, two-tailed; with df = 10 it must pass 2.228; with σ known the cutoff is 1.96.

Chapter 11 formulas and facts

Test	Statistic	When to use it
z-test for μ	$z = (\bar{x} - \mu_0) \div (\sigma \div \sqrt{n})$	σ known from long history
t-test for μ	$t = (\bar{x} - \mu_0) \div (s \div \sqrt{n})$, $df = n - 1$	σ estimated by s
One-proportion z-test	$z = (\hat{p} - p_0) \div \sqrt{[p_0(1 - p_0) \div n]}$	Yes/no outcome; np_0, $n(1 - p_0) \geq 10$
Proportion CI	$\hat{p} \pm z^* \sqrt{[\hat{p}(1 - \hat{p}) \div n]}$	Estimating p after the test
p-value, two-sided	2 × P(beyond |statistic|)	Ha: ≠
p-value, one-sided	One tail toward Ha	Direction set in advance
Chance of a false positive in k tests	$1 - (1 - \alpha)^k$	Multiple testing

Worked examples

EXAMPLE 11.1 · A Z-TEST WITH Σ KNOWN

A standardized test has σ = 15 and a published mean of 100. A random sample of 36 students at one district averages 104.5. Test H0: μ = 100 against Ha: μ ≠ 100 at α = 0.05.

STEP 1 · HYPOTHESES

H0: μ = 100; Ha: μ ≠ 100.

STEP 2 · CONDITIONS

Random sample, independent students, σ known, and n = 36 is large enough for the CLT.

STEP 3 · STATISTIC

SE = 15 ÷ 6 = 2.50; z = (104.5 − 100) ÷ 2.50 = **1.80**.

STEP 4 · P-VALUE

Upper tail beyond 1.80 = 1 − 0.9641 = 0.0359; doubled, **p = 0.0718**.

STEP 5 · CONCLUDE

0.0718 > 0.05, so **fail to reject H0**. The 95% interval 104.5 ± 1.96(2.50) = 99.60 to 109.40 contains 100, as it must. The district may differ, but this sample cannot show it.

TRAPS PEOPLE FALL INTO

- **Picking the direction from the data.** A one-sided Ha chosen after seeing the sample doubles the real error rate.
- **Using z when s came from the data.** Estimated σ means t with df = n − 1.
- **Using a t-test for a proportion.** Proportions use z; p_0 fixes the variability.
- **Using p̂ in the test's SE.** The test lives in the null world: use p_0. The interval uses p̂.
- **Entering percents.** Use 0.54, not 54.
- **Reading p = 0.051 as "no effect."** It is almost the same evidence as 0.049. Give the interval and n.
- **Checking normality on the wrong thing.** It is the sampling distribution of x̄ that must be normal; the CLT helps as n grows.
- **Expecting large n to fix a bad sample.** The CLT fixes shape, never bias.

Chapter 11 self-test · six items

1. When is the z-test for a mean appropriate?

A. Whenever n is at least 30
B. When σ is truly known from long history
C. Whenever the data are skewed
D. When s is computed from the sample

B. When σ is truly known from long history With σ estimated, the honest choice is t.

2. Two-tailed t* at α = 0.05 for df = 15 is 2.131. t = 2.125. Decision?

A. Reject, since 2.125 is close
B. Accept H0
C. Fail to reject, just barely
D. Reject, since t is positive

C. Fail to reject, just barely 2.125 does not reach 2.131, so p is just above 0.05.

3. Which SE belongs in a one-proportion test?

A. $\sqrt{[\hat{p}(1 - \hat{p})/n]}$
B. $s/\sqrt{n}$
C. σ/n
D. $\sqrt{[p_0(1 - p_0)/n]}$

D. $\sqrt{[p_0(1 - p_0)/n]}$ The test is computed assuming H0, so it uses p_0.

4. Same $\bar{x}$ and s, but n grows from 16 to 64. What happens to t?

A. It doubles
B. It halves
C. It stays the same
D. It quadruples

A. It doubles SE = $s/\sqrt{n}$ halves when n quadruples, so t doubles.

5. The t-test is least robust to:

A. Mild skew with n = 50
B. A large sample size of 200
C. Outliers in small samples
D. Using a two-sided alternative

C. Outliers in small samples One wild value can move both $\bar{x}$ and s in a small sample.

6. Twenty independent tests at α = 0.05 with no real effects give about how many false positives?

A. 1

B. 0

C. 5

D. 20

A. 1 20 × 0.05 = 1 expected by chance alone.

CHAPTER 12

Two-Sample Tests

Read it in full: **Chapter 12 in the textbook**. Practice it in the lab: **Comparing Two Means** and **Which Test**.

What to know

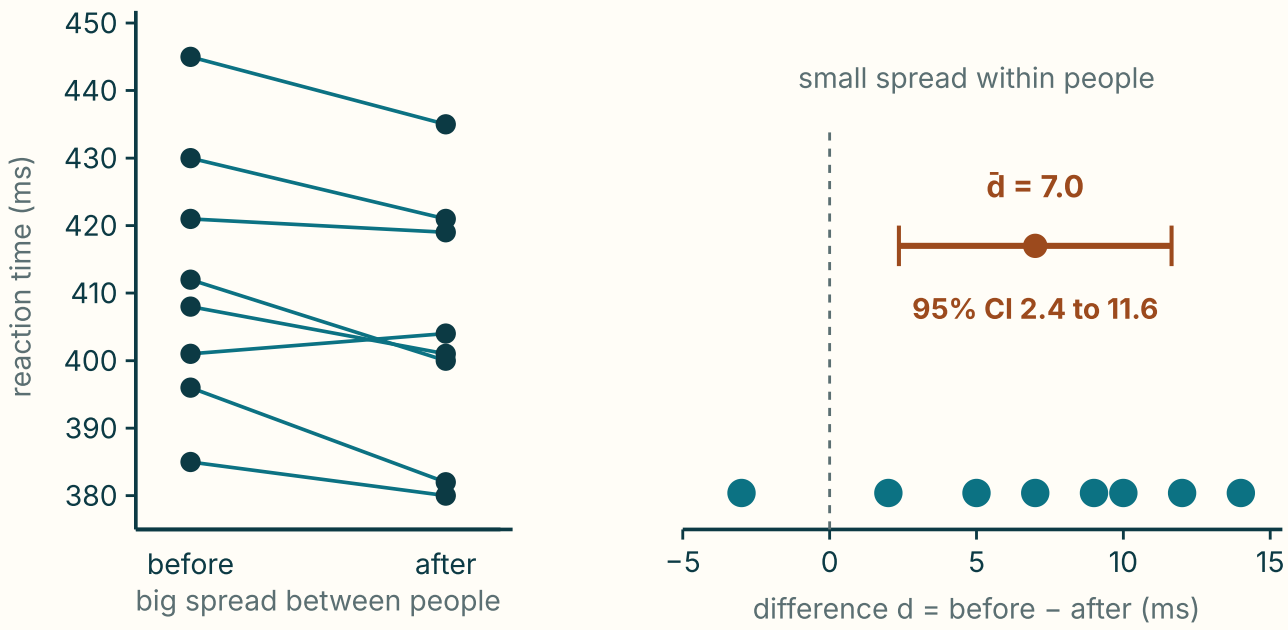
Independent means: SE = $\sqrt{(s_1^2/n_1 + s_2^2/n_2)}$, t = $(\bar{x}_1 - \bar{x}_2)/SE$ with Welch df. Variances add; SDs never do.

Paired: take differences, then t = $\bar{d}/(s_d/\sqrt{n})$ with df = pairs − 1. Pairing removes person-to-person noise; the book's 7 ms effect gave p ≈ 0.009 paired and 0.48 when wrongly treated as independent.

Two proportions: z with the pooled proportion in the test and the unpooled SE in the interval. The design picks the test: numeric or yes/no, independent or paired.

LOCK THESE IN

- Independent: SE = $\sqrt{(s_1^2/n_1 + s_2^2/n_2)}$; variances add.
- Use Welch df by default; rounding df down by hand is conservative.
- Paired: differences first, then a one-sample t with df = number of pairs − 1.
- Pairing is a design fact, not something you read off the numbers.
- Two proportions: z, pooled $\hat{p}$ in the test, unpooled SE in the CI.
- Pooled $\hat{p} = (x_1 + x_2)/(n_1 + n_2)$; check four expected counts ≥ 10.
- Numeric or yes/no × independent or paired picks the test.
- Percentage points are not percent change.



Pairing removes the person-to-person noise. The raw times differ by up to 60 ms between people, but each person's change is small and consistent. The differences have mean 7.0 ms and SE 1.964, so the 95% interval sits above zero.

Chapter 12 formulas and facts

Comparison	Statistic	Interval
Two independent means (Welch)	$t = (\bar{x}_1 - \bar{x}_2) \div \sqrt{(s_1^2/n_1 + s_2^2/n_2)}$	$(\bar{x}_1 - \bar{x}_2) \pm t^*$ SE, Welch df
Welch df	$(a_1 + a_2)^2 \div [a_1^2/(n_1 - 1) + a_2^2/(n_2 - 1)]$, $a = s^2/n$	Round down by hand
Paired means	$t = \bar{d} \div (s_d/\sqrt{n})$, $df = n - 1$	$\bar{d} \pm t^* s_d/\sqrt{n}$
Pooled proportion	$\hat{p}_{pool} = (x_1 + x_2)/(n_1 + n_2)$	Test only
Two proportions	$z = (\hat{p}_1 - \hat{p}_2) \div \sqrt{[\hat{p}_{pool}(1 - \hat{p}_{pool})(1/n_1 + 1/n_2)]}$	$(\hat{p}_1 - \hat{p}_2) \pm z^*\sqrt{[\hat{p}_1\hat{q}_1/n_1 + \hat{p}_2\hat{q}_2/n_2]}$
Relative change	$(\hat{p}_1 - \hat{p}_2) \div \hat{p}_2$	Not the same as percentage points

Worked examples

EXAMPLE 12.1 · A WELCH TWO-SAMPLE T-TEST WITH ITS INTERVAL

Students randomly assigned to a flipped class ($n = 20$, $\bar{x} = 24.6$, $s = 5.1$) or a lecture class ($n = 18$, $\bar{x} = 21.2$, $s = 6.3$) take the same quiz. Test for a difference at $\alpha = 0.05$ and build a 95% interval.

STEP 1 · DIFFERENCE

$24.6 - 21.2 = 3.4$ points.

STEP 2 · SE

$5.1^2/20 = 1.3005$; $6.3^2/18 = 2.2050$; SE $= \sqrt{3.5055} =$ **1.8723**.

STEP 3 · T

$t = 3.4 \div 1.8723 =$ **1.816**.

STEP 4 · WELCH DF

$(3.5055)^2 \div [1.3005^2/19 + 2.2050^2/17] = 12.2885 \div 0.3750 =$ **32.8**.

STEP 5 · P AND DECISION

Two-sided $p \approx 0.079$. Fail to reject H0 at $\alpha = 0.05$.

STEP 6 · INTERVAL

With df rounded down to 32, t* = 2.037: 3.4 ± 2.037(1.8723) = 3.4 ± 3.81, so **−0.41 to 7.21**. It contains 0 and is wide: the study is too small to say much.

TRAPS PEOPLE FALL INTO

- **Adding standard deviations.** Variances add: $\sqrt{(s_1^2/n_1 + s_2^2/n_2)}$.
- **Analyzing paired data as independent.** Person-to-person noise swamps the effect; p went from 0.009 to 0.48.
- **Counting measurements instead of pairs.** Eight people measured twice give n = 8 and df = 7.
- **Forcing a pairing on independent groups.** Lining up rows by order invents a relationship.
- **Using the pooled SE in the interval.** Pool only for the test, where H0 says the proportions are equal.
- **Saying the means are equal after failing to reject.** A wide interval that includes 0 means the study was too small.
- **Mixing percentage points and percent.** A rise from 18.8% to 30.0% is 11.3 points and 60%.
- **Switching tests after seeing p.** The design picks the test before any computing.

Chapter 12 self-test · six items

1. What does H0: $\mu_1 - \mu_2 = 0$ claim?

A. The two sample means are equal
B. The two samples are paired
C. The variances are equal
D. The two population means are equal

D. The two population means are equal Hypotheses concern the populations; the samples nearly always differ.

2. Why use Welch's test by default?

A. It always gives a smaller p-value
B. It does not assume equal variances
C. It needs no conditions at all
D. It works only for paired data

B. It does not assume equal variances Welch protects you when variances differ and costs little when they do not.

3. Differences 3, 5, 4 from three pairs. What is $\bar{d}$?

A. 12
B. 3
C. 5
D. 4

D. 4 $(3 + 5 + 4) \div 3 = 4$.

4. In the two-proportion test, the pooled proportion is used because:

A. H0 says the two proportions are equal
B. It always gives a larger z value
C. The two samples are paired
D. The interval requires a pooled SE

A. H0 says the two proportions are equal Under H0 there is one common proportion, best estimated from all the data.

5. Numeric outcome, two independent groups of 6, data badly skewed. Which test fits?

A. Wilcoxon rank-sum
B. Paired t
C. McNemar's test
D. Two-proportion z

A. Wilcoxon rank-sum Tiny, skewed, independent samples call for a rank-based test.

6. A 95% CI for $p_1 - p_2$ is 0.02 to 0.15. What does it say?

A. The two rates are shown to be equal
B. Group 2's rate is plausibly the higher one
C. Group 1's rate is plausibly 2 to 15 points higher
D. The difference is 15 points exactly

C. Group 1's rate is plausibly 2 to 15 points higher The whole interval is above zero, so group 1 is higher by somewhere between 2 and 15 points.

CHAPTER 13

Chi-Square and ANOVA

Read it in full: **Chapter 13 in the textbook**. Practice it in the lab: **Goodness of Fit** and **Which Groups Differ**.

What to know

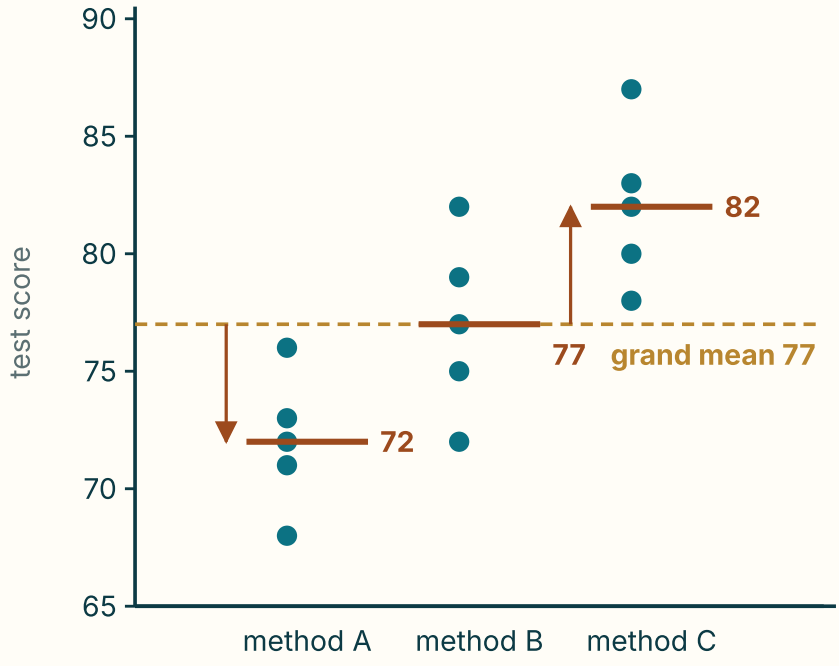
$\chi^2 = \Sigma(O - E)^2/E$, never negative and always right-tailed. Goodness of fit: $E = np$, $df = k - 1$. Independence: E = row × column ÷ total, $df = (r - 1)(c - 1)$, with Cramér's V for size. Needs counts and all expected counts at least 5.

ANOVA avoids the inflated error of many t-tests: $F = MSB/MSW$ with $(k - 1, N - k)$ df, $SS_{total} = SS_{between} + SS_{within}$, $\eta^2 = SS_{between}/SS_{total}$. It assumes independence, rough normality and equal variances (largest s under twice the smallest).

A significant F means at least one mean differs. Tukey's HSD = q√(MSW/n) or Bonferroni (α/m) find which pairs; in the book only A versus C separated.

LOCK THESE IN

- $\chi^2 = \Sigma(O - E)^2/E$; right-tailed; never negative.
- Goodness of fit: E = np, df = k − 1. Never round E.
- Independence: E = (row × column)/total, df = (r − 1)(c − 1).
- Chi-square needs counts and every expected count ≥ 5.
- F = MSB/MSW; df = k − 1 and N − k; $SS_T = SS_B + SS_W$.
- Ha for ANOVA: at least one mean differs, not all.
- Equal variances: largest s < 2 × smallest.
- Post-hoc: Tukey HSD = q√(MSW/n); Bonferroni tests at α/m.



between (signal)
SS = 250, df = 2
MSB = 125.00

within (noise)
SS = 138, df = 12
MSW = 11.50

F = 125.00 ÷ 11.50 = 10.87

Arrows: group means sit far from 77 compared with the scatter inside each group.

F compares signal with noise. The three method means (72, 77, 82) spread far from the grand mean compared with the scatter inside each group, so F = 10.87 on (2, 12) df.

Chapter 13 formulas and facts

Quantity	Formula	When to use it
Chi-square statistic	$\chi^2 = \Sigma(O - E)^2 \div E$	Both chi-square tests
Expected, goodness of fit	$E = n \times p_i$	One variable versus a claim
Expected, independence	E = (row total × column total) ÷ grand total	Two-way table
df	$k - 1$; $(r - 1)(c - 1)$	GOF; independence
Standardized residual	$(O - E) \div \sqrt{E}$	Which cells drive the result
Cramér's V	$\sqrt{[\chi^2 \div (n \times \min(r - 1, c - 1))]}$	Size of an association, 0 to 1
Between SS	$\Sigma\, n_j(\bar{x}_j - \bar{x})^2$, df = k − 1	ANOVA signal
Within SS	$\Sigma\Sigma(x - \bar{x}_j)^2$, df = N − k	ANOVA noise
F	MSB ÷ MSW, MS = SS ÷ df	Equal means?
Eta squared	$\eta^2 = SS_{between} \div SS_{total}$	Share of variation explained
Tukey HSD	$q_{\alpha,\, k,\, N-k} \times \sqrt{(MSW \div n)}$	All pairs, equal n
Bonferroni	Each test at $\alpha \div m$	A few planned comparisons

Worked examples

EXAMPLE 13.1 · GOODNESS OF FIT FOR A DIE

A die is rolled 120 times: faces 1 to 6 come up 15, 22, 18, 26, 17, 22 times. Is it fair? Use $\alpha = 0.05$.

STEP 1 · HYPOTHESES AND E

H0: each face has probability 1/6. $E = 120 \times 1/6 = 20$ for every face, summing to 120. All $E \geq 5$.

STEP 2 · TERMS

$(15 - 20)^2/20 = 1.25$; $(22 - 20)^2/20 = 0.20$; $(18 - 20)^2/20 = 0.20$; $(26 - 20)^2/20 = 1.80$; $(17 - 20)^2/20 = 0.45$; $(22 - 20)^2/20 = 0.20$.

STEP 3 · X^2

Sum = **4.10**, df = 6 − 1 = 5.

STEP 4 · DECIDE

Critical value 11.070; p ≈ 0.535. 4.10 < 11.070, so **fail to reject**.

STEP 5 · SAY IT

These rolls give no convincing evidence that the die is unfair. That is not proof of fairness; 120 rolls cannot detect a small bias. Face 4, with the largest term, is the one to watch.

TRAPS PEOPLE FALL INTO

- **Feeding percentages into χ^2.** Chi-square needs raw counts.
- **Rounding expected counts.** Keep E = 47.95 as 47.95.
- **Checking observed counts for the 5 rule.** The rule is about expected counts; an observed 0 is fine.
- **Reading a significant χ^2 as cause.** Survey tables show association only.
- **Running pairwise t-tests instead of ANOVA.** Each extra test adds false-alarm risk.
- **Writing Ha as "all means differ."** It is "at least one differs."
- **Pooling very unequal variances.** If the largest s is over twice the smallest, use Welch's ANOVA.
- **Fishing after a non-significant F.** Post-hoc tests follow a significant F, or planned contrasts set in advance.

Chapter 13 self-test · six items

1. Why is a chi-square test always right-tailed?

A. χ^2 values are often negative numbers
B. Only large χ^2 values show a mismatch with H0
C. The χ^2 curve is symmetric about 0
D. Small χ^2 values are what reject H0

B. Only large χ^2 values show a mismatch with H0 Each term is a square over a positive number, so bigger means worse fit.

2. Expected counts must be at least 5 for:

A. Every cell, using the observed counts
B. Only the largest cell in the table
C. The row and column totals only
D. Every cell, using expected not observed counts

D. Every cell, using expected not observed counts The approximation depends on expected counts.

3. The same people are asked a yes or no question twice. Which test fits?

A. McNemar's test
B. Chi-square test of independence
C. One-way ANOVA
D. Goodness of fit

A. McNemar's test Paired categorical data break the independence chi-square needs.

4. $SS_{between} = 60$ and $SS_{total} = 200$. What is η^2?

A. 0.43
B. 3.33
C. 0.60
D. 0.30

D. 0.30 60 ÷ 200 = 0.30: 30% of the variation is explained by group.

5. Five groups compared with Bonferroni across all 10 pairs at $\alpha = 0.05$ test each pair at:

A. 0.05
B. 0.01
C. 0.005
D. 0.025

C. 0.005 0.05 ÷ 10 = 0.005.

6. With exactly two groups, how do ANOVA and the pooled t-test relate?

A. $F = t^2$, with the same p-value
B. $F = 2t$
C. They give opposite decisions
D. ANOVA needs paired data

A. $F = t^2$, with the same p-value ANOVA generalizes the pooled two-sample t-test.

CHAPTER 14

Correlation and Regression

Read it in full: **Chapter 14 in the textbook**. Practice it in the lab: **R-Squared and Residual Plots** and **The Danger of Extrapolation**.

What to know

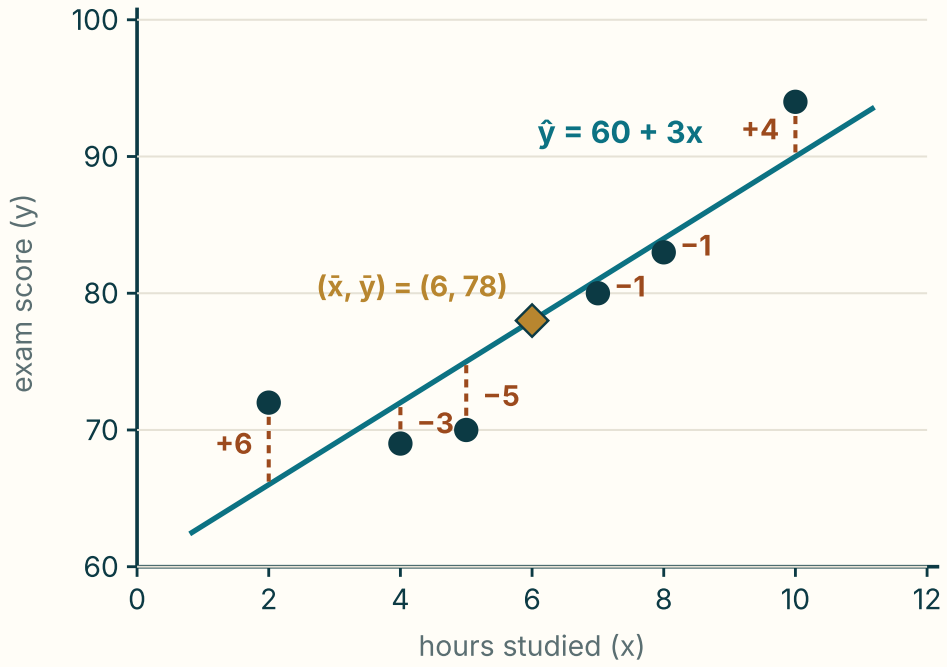
Describe a scatterplot by form, direction, strength and outliers, with x explanatory. $r = \Sigma(x - \bar{x})(y - \bar{y})/\sqrt{[\Sigma(x - \bar{x})^2\Sigma(y - \bar{y})^2]}$ lies in $[-1, 1]$, is unit-free and symmetric, measures only linear association, is not resistant and never shows cause.

Least squares: $b = Sxy/Sxx = r\, s_y/s_x$, $a = \bar{y} - b\bar{x}$, through $(\bar{x}, \bar{y})$. The book's students give $\hat{y} = 60 + 3x$; the slope is a predicted average rate, not a causal effect.

Residuals $y - \hat{y}$ sum to 0; $r^2 = 1 - SSE/SST$ is the share of variation explained (0.81 here); $s_e = \sqrt{[SSE/(n - 2)]}$ is the typical miss. Check the residual plot, interpolate only, and never extrapolate.

LOCK THESE IN

- Form, direction, strength, outliers; x is explanatory.
- r lies in [−1, 1], has no units and measures linear association only.
- $b = \Sigma(x - \bar{x})(y - \bar{y}) \div \Sigma(x - \bar{x})^2 = r\, s_y/s_x$; $a = \bar{y} - b\bar{x}$.
- The line passes through $(\bar{x}, \bar{y})$; residuals sum to 0.
- $r^2 = 1 - SSE/SST$ = share of variation in y explained by the line.
- $s_e = \sqrt{[SSE \div (n - 2)]}$.
- Interpolate, do not extrapolate. Refit to predict x from y.
- Correlation is not causation; strength is not steepness.



r = 0.90
r² = 0.81
SSE = 88
SST = 466
se = √(88 ÷ 4) = 4.69
residuals sum to 0

dashed: residual
y − ŷ

Least squares minimizes the squared vertical misses. The line $\hat{y} = 60 + 3x$ passes through (6, 78). Its residuals +6, −3, −5, −1, −1 and +4 square and sum to SSE = 88, the smallest possible, giving $r^2 = 0.81$.

Chapter 14 formulas and facts

Quantity	Formula	When to use it
Correlation	$r = Sxy \div \sqrt{(Sxx \times Syy)}$	Strength and direction of a linear pattern
Correlation (z form)	$r = \Sigma z_x z_y \div (n - 1)$	Why r has no units
Slope	$b = Sxy \div Sxx = r(s_y \div s_x)$	Predicted change in y per unit x
Intercept	$a = \bar{y} - b\bar{x}$	Anchors the line
Prediction	$\hat{y} = a + bx$	Average y at x, inside the data range
Residual	$e = y - \hat{y}$	What the line missed
Sums of squares	$SST = \Sigma(y - \bar{y})^2$; $SSE = \Sigma(y - \hat{y})^2$	Variation before and after the line
r^2	$1 - SSE \div SST = r^2$	Share of variation explained
Std. error of estimate	$s_e = \sqrt{[SSE \div (n - 2)]}$	Typical miss

Worked examples

EXAMPLE 14.1 · COMPUTING R FROM A DEVIATION TABLE

Six weeks of ad spend x (hundreds of dollars) and sales y (thousands): x = 3, 5, 6, 8, 9, 11; y = 20, 25, 24, 31, 33, 37. Find r.

STEP 1 · MEANS

$\bar{x} = 42 \div 6 = 7.00$; $\bar{y} = 170 \div 6 = 28.33$.

STEP 2 · DEVIATIONS

$x - \bar{x}$: −4.00, −2.00, −1.00, 1.00, 2.00, 4.00. $y - \bar{y}$: −8.33, −3.33, −4.33, 2.67, 4.67, 8.67. Both sum to 0.

STEP 3 · SUMS

$Sxy = \Sigma(x - \bar{x})(y - \bar{y}) = 91.00$; $Sxx = 42.00$; $Syy = 203.33$.

STEP 4 · R

$r = 91.00 \div \sqrt{(42.00 \times 203.33)} = 91.00 \div 92.41 =$ **0.985**.

STEP 5 · READ IT

Strong and positive, and only after the scatterplot shows a straight band. These are observational weeks, so r says nothing about whether ads caused the sales.

TRAPS PEOPLE FALL INTO

- **Computing r before looking.** r is blind to curves, clusters and single influential points.
- **Treating strength as steepness.** A shallow tight band is strong; a steep fuzzy one is weak.
- **Calling −0.87 weaker than +0.60.** Strength is distance from zero.
- **Reading the slope as cause.** Say "predicted, on average" for observational data.
- **Interpreting an intercept outside the data.** If x = 0 was never observed, the intercept only anchors the line.
- **Saying r^2 is the share of points on the line.** It is the share of variation in y explained.
- **Extrapolating.** The line predicted an impossible 105 at 15 hours.
- **Inverting the line to predict x.** Refit with the roles reversed.

Chapter 14 self-test · six items

1. r = 0.05 for two variables. What can you conclude?

A. The variables are unrelated in every way
B. The slope of the line is large
C. Little linear pattern, though a curve is possible
D. x causes y, but only weakly

C. Little linear pattern, though a curve is possible r measures linear association only.

2. Converting x from inches to centimeters changes r how?

A. It multiplies r by 2.54
B. It divides r by 2.54
C. It changes its sign
D. Not at all

D. Not at all r is unit-free because it is built from z-scores.

3. The least-squares line always passes through:

A. (0, 0)
B. The first data point
C. The largest residual
D. $(\bar{x}, \bar{y})$

D. $(\bar{x}, \bar{y})$ $a = \bar{y} - b\bar{x}$ rearranges to $\bar{y} = a + b\bar{x}$.

4. SSE = 50 with n = 12. What is s_e?

A. 2.13
B. 5.00
C. 2.24
D. 4.17

C. 2.24 $\sqrt{(50 \div 10)} = \sqrt{5} \approx 2.24$.

5. Least-squares residuals always sum to:

A. 1
B. 0
C. SSE
D. r^2

B. 0 That is a guaranteed property and a free arithmetic check.

6. A strong positive r between ice cream sales and drownings means:

A. Eating ice cream causes people to drown
B. They move together, likely via a lurking variable
C. Drownings raise sales of ice cream
D. The data must contain an error

B. They move together, likely via a lurking variable Hot weather drives both. Correlation is not causation.

ONE PAGE

Cram sheet

Every formula and rule in the course, sorted by topic. Cover the right column and rebuild each row. If a row will not come back, go to that chapter's exam guide.

Data, sampling and design (chapters 1 and 2)

Name	Formula	Note
Mean	$\bar{x} = \Sigma x \div n$; $\mu = \Sigma x \div N$	Pulled by outliers
Median	Middle sorted value	Average the two middle values when n is even
Range	Max − min	Set by the two extremes
Variance	$\sigma^2 = \Sigma(x - \mu)^2 \div N$; $s^2 = \Sigma(x - \bar{x})^2 \div (n - 1)$	SD is the square root
Risk	$RRR = (r_c - r_t) \div r_c$; $ARR = r_c - r_t$	NNT = 1 ÷ ARR
Standard error	$SE(\bar{x}) = \sigma \div \sqrt{n}$	4 × n halves it
Allocation	$n_h = n \times N_h \div N$	Proportional stratified sample
Systematic	$k = N \div n$	Random start from 1 to k
Margin of error	$1.96 \times \sqrt{[\hat{p}(1 - \hat{p}) \div n]}$	Chance error only

Displaying and summarizing data (chapters 3 and 4)

Name	Formula	Note
Relative frequency	f ÷ n	Column sums to 1
Class width	(max − min) ÷ classes, round up	5 to 10 classes
Pie angle	(f ÷ n) × 360°	1% = 3.6°
Lie factor	apparent ÷ actual effect	Honest is about 1
Mean and median	$\Sigma x \div n$; middle sorted value	Mean not resistant
Variance	$s^2 = SS \div (n - 1)$; $\sigma^2 = SS \div N$	$SS = \Sigma(x - \bar{x})^2$
CV	$(s \div \bar{x}) \times 100\%$	Ratio data only
IQR and fences	Q3 − Q1; Q1 − 1.5 IQR; Q3 + 1.5 IQR	Beyond a fence = flag
z-score	$z = (x - \mu) \div \sigma$; $x = \mu + z\sigma$	Unitless
Chebyshev	$\geq 1 - 1 \div k^2$	75% at k = 2; 88.9% at k = 3

Probability and random variables (chapters 5 and 6)

Name	Formula	Note
Classical	$P(A) = n(A) \div n(S)$	Equally likely outcomes
Complement	$P(A^c) = 1 - P(A)$	At least one = 1 − P(none)
Addition	$P(A) + P(B) - P(A \text{ and } B)$	Drop the last term if disjoint
Multiplication	$P(A) \times P(B \mid A)$	P(A)P(B) if independent
Conditional	$P(A \text{ and } B) \div P(B)$	New denominator
Expected value	$\Sigma x \cdot P(x)$	Long-run average
Variance	$\Sigma x^2 P(x) - \mu^2$	$\sigma = \sqrt{\sigma^2}$
Binomial	${}_nC_x p^x q^{n-x}$; $\mu = np$; $\sigma = \sqrt{(npq)}$	Four conditions
Poisson	$\lambda^x e^{-\lambda} \div x!$; $\mu = \sigma^2 = \lambda$	Fixed interval
Geometric	$q^{x-1}p$; $\mu = 1/p$	First success

Normal model and sampling distributions (chapters 7 and 8)

Name	Formula	Note
z-score	$z = (x - \mu) \div \sigma$	$x = \mu + z\sigma$
Areas	Left $\Phi(z)$; right $1 - \Phi(z)$; between $\Phi(z_b) - \Phi(z_a)$	$\Phi(-z) = 1 - \Phi(z)$
Key z	1.28, 1.645, 1.96, 2.326, 2.576	90th, 95th, 97.5th, 99th, 99.5th percentiles
Mean of $\bar{x}$	$\mu_{\bar{x}} = \mu$	Unbiased
SE of $\bar{x}$	$\sigma \div \sqrt{n}$ (or $s \div \sqrt{n}$)	4 × n halves it
z for $\bar{x}$	$(\bar{x} - \mu) \div (\sigma \div \sqrt{n})$	Sample mean questions
Sum	Normal($n\mu$, $\sigma\sqrt{n}$)	CLT for totals
Proportion	$\mu_{\hat{p}} = p$; SE = $\sqrt{[p(1 - p) \div n]}$	$np \geq 10$, $n(1 - p) \geq 10$
Planning n	$n = (s \div SE)^2$	Round up

Confidence intervals and testing logic (chapters 9 and 10)

Name	Formula	Note
CI skeleton	estimate ± critical value × SE	Width = 2 × margin
Mean, σ known	$\bar{x} \pm z^* \sigma/\sqrt{n}$	z* 1.645, 1.96, 2.576
Mean, σ unknown	$\bar{x} \pm t^* s/\sqrt{n}$	df = n − 1
Proportion	$\hat{p} \pm z^* \sqrt{[\hat{p}(1-\hat{p})/n]}$	$n\hat{p}$, $n(1-\hat{p}) \geq 10$
Sample size	$n = (z^* \sigma/E)^2$; $n = z^{*2} p(1-p)/E^2$	Round up; p = 0.5 if unknown
Test statistic	(estimate − null value) ÷ SE	t with df = n − 1 for means
Two-sided p	2 × one tail	One-sided only if chosen first
Power	1 − β	Aim for 0.80 or more
Cohen's d	difference ÷ s	Effect size

One- and two-sample tests (chapters 11 and 12)

Name	Formula	Note
z for μ	$(\bar{x} - \mu_0) \div (\sigma/\sqrt{n})$	σ known only
t for μ	$(\bar{x} - \mu_0) \div (s/\sqrt{n})$, df = n − 1	σ estimated
One proportion	$(\hat{p} - p_0) \div \sqrt{[p_0(1-p_0)/n]}$	CI uses $\hat{p}$
Welch t	$(\bar{x}_1 - \bar{x}_2) \div \sqrt{(s_1^2/n_1 + s_2^2/n_2)}$	Welch df
Paired t	$\bar{d} \div (s_d/\sqrt{n})$	df = pairs − 1
Pooled $\hat{p}$	$(x_1 + x_2)/(n_1 + n_2)$	Test only
Two proportions	$(\hat{p}_1 - \hat{p}_2) \div SE_{pool}$	CI uses unpooled SE
False positives	$1 - (1 - \alpha)^k$	k tests

Chi-square, ANOVA and regression (chapters 13 and 14)

Name	Formula	Note
Chi-square	$\Sigma(O - E)^2/E$	Right-tailed
Expected counts	np; row × column ÷ total	All ≥ 5; never round
df	$k - 1$; $(r - 1)(c - 1)$	GOF; independence
Cramér's V	$\sqrt{[\chi^2/(n \times \min(r - 1, c - 1))]}$	Size of association
ANOVA	$F = MSB/MSW$; df $k - 1$, $N - k$	$SS_T = SS_B + SS_W$
Effect size	$\eta^2 = SS_B/SS_T$	Share explained
Tukey	$HSD = q\sqrt{(MSW/n)}$	Bonferroni: α/m
Correlation	$r = Sxy/\sqrt{(Sxx\ Syy)}$	$[-1, 1]$, unit-free
Line	$b = Sxy/Sxx = r\ s_y/s_x$; $a = \bar{y} - b\bar{x}$	Through $(\bar{x}, \bar{y})$
Fit	$r^2 = 1 - SSE/SST$; $s_e = \sqrt{[SSE/(n - 2)]}$	Residuals sum to 0

Rules that are not formulas, all chapters

Topic	The rule
Symbols	Greek for populations (μ, σ, p, N); Roman for samples ($\bar{x}$, s, $\hat{p}$, n).
Error versus bias	Sampling error shrinks with n; bias needs better design.
Data types	Qualitative or quantitative; quantitative is discrete (counted) or continuous (measured).
Levels	Nominal (mode), ordinal (median), interval (mean, no ratios), ratio (all, ratios valid).
Correlation	Four explanations: cause, reverse cause, confounder, coincidence.
Sampling methods	Stratified: some from every group. Cluster: all from some groups.
No inference	Convenience and voluntary response samples.
Two randomizations	Random sampling generalizes; random assignment shows cause.
Experiment	Control, randomization, replication, blocking, placebo, double blinding.
P-value and CI	Neither is the probability that a hypothesis or parameter value is true.
Checks	Frequencies sum to n; relative to 1; cumulative ends at n.
Classes	No overlaps, no gaps; round the width up.
Bars	Bar chart gaps for categories; histogram bars touch; both start at zero.
Pie	Only for exclusive, exhaustive categories.
Skew	Named for the tail; right skew pulls the mean above the median.
Bimodal	Two groups mixed; split before summarizing.
Choosing center	Mean for symmetric; median for skewed or ordinal; mode for categories.
Quartiles	Median of each half; leave the median out when n is odd.
Empirical rule	68, 95, 99.7 within 1, 2, 3 SD, bell shapes only.
Sample space	List ordered outcomes; HT and TH differ.

Topic	The rule
Long run	Proportions settle toward P; nothing is due.
Disjoint vs independent	Disjoint events with nonzero P are dependent.
Two-way table	Joint over total; marginal over total; conditional over its row or column.
Trees	Multiply along, add across; ends sum to 1.
Inverse	P(A | B) is not P(B | A).
Valid table	Sum to 1 and every entry in [0, 1].
Unusual	Outside μ ± 2σ.
Model choice	Fixed trials: binomial. Fixed interval: Poisson. One success: geometric.
Empirical rule	68, 95, 99.7; blocks 34, 13.5, 2.35, 0.15.
Area	Area is probability; a single point has probability 0.
Backward	Area → z → x; "top 10%" is a left area of 0.90.
Shape	Standardizing never fixes skew.
Three distributions	Population, sample data, sampling distribution.
Noun check	One person: σ. Sample mean: σ ÷ √n.
CLT	Any shape with finite σ; n ≥ 30 is a rule of thumb.
SD vs SE	SD for people; SE for estimates.
z or t	If you computed s, use t. Proportions use z.
Dials	More confidence widens; 4 × n halves the margin.
CI meaning	About C% of intervals from this method capture the parameter.
CI and tests	A value outside a 95% CI is rejected at α = 0.05, two-sided.
Hypotheses	About parameters; equals sign in H0; tails chosen first.

Topic	The rule
p-value	P(data this extreme | H0); not P(H0 true).
Decision	p ≤ α reject; p > α fail to reject; never accept.
Errors	Type I false alarm (α); Type II miss (β).
Significance	Not size, not importance, not cause.
Five steps	Hypotheses, α and conditions, statistic, p-value, conclusion in context.
Direction	Chosen from the question before the data.
z or t	Estimated σ means t; proportions use z.
Duality	Two-sided test at α rejects exactly when the 100(1 − α)% CI excludes the claim.
Conditions	Random (unfixable), independent (10%), normal enough (by n).
Pairing	Decided by design; n counts pairs.
Test choice	Numeric or yes/no × independent or paired.
Zero in the CI	Fail to reject; width tells how much you learned.
Chi-square data	Counts only; paired yes/no uses McNemar.
Small counts	Combine categories or use Fisher's exact test.
ANOVA Ha	At least one mean differs.
ANOVA conditions	Independence, rough normality, largest s < 2 × smallest.
After F	Post-hoc only after a significant F (or planned contrasts).
Scatterplot	Form, direction, strength, outliers.
r limits	Linear only, not resistant, shrinks with restricted range, not causal.
Residual plot	Want no pattern; U = curve; fan = changing spread.
Prediction	Interpolate only; refit to predict x from y.

Dr. Garg Learning · CC BY-SA 4.0