Statistics

TI-84-focused notes, quick references, and representative examples for an introductory statistics course.

Module 1: Introduction, Data Collection, Sampling, and Experimental Design

Students learn the language of statistics, study design, sampling methods, bias, and experiment design.

Section 1.1 Introduction to the Practice of Statistics

Use the basic vocabulary of statistics and classify variables correctly.

QUICK REFERENCE

- Statistics is the science of collecting, organizing, summarizing, and analyzing information to answer questions.
- Descriptive statistics summarizes data. Inferential statistics uses sample data to draw conclusions about a population.
- A population is the entire group being studied. A sample is a subset of that population.
- A parameter describes a population. A statistic describes a sample.
- Qualitative variables describe categories. Quantitative variables are numerical and may be discrete or continuous.

Example 1. A researcher surveys 420 adults from a city to estimate the percent of all adults in the city who support a new park. Identify the population and sample.

Answer: Population: all adults in the city. Sample: the 420 adults surveyed.

Example 2. A report says 61% of the 420 surveyed adults support the park. Is 61% a statistic or a parameter?

Answer: Statistic, because it summarizes the sample.

Example 3. Classify each variable as qualitative or quantitative: eye color, number of siblings, commute time, and phone brand.

Answer: Eye color: qualitative; number of siblings: quantitative discrete; commute time: quantitative continuous; phone brand: qualitative.

Example 4. A graph summarizes the ages of students in one class. Is this descriptive or inferential statistics?

Answer: Descriptive statistics, because it summarizes the data that were collected.

Section 1.2 Observational Studies versus Designed Experiments

Distinguish observational studies from experiments and avoid unsupported causal claims.

QUICK REFERENCE

- An observational study measures characteristics without trying to change the subjects.
- A designed experiment imposes treatments and observes the response.
- The explanatory variable may help explain changes in the response variable.
- A lurking or confounding variable can make it hard to know what caused an observed result.
- Observational studies can show association, but they do not prove cause and effect by themselves.

Example 1. A researcher records coffee intake and sleep hours for 200 adults without assigning coffee amounts. Is this observational or experimental?

Answer: Observational study.

Example 2. Students are randomly assigned to either use a new study app or not use it, then their exam scores are compared. Is this observational or experimental?

Answer: Designed experiment.

Example 3. In a study of study time and exam score, identify the explanatory variable and response variable.

Answer: Explanatory variable: study time. Response variable: exam score.

Example 4. A study finds that people who exercise more also have lower stress. What wording is safest?

Answer: There is an association between exercise and lower stress. Do not claim exercise caused lower stress from this alone.

Section 1.3 Simple Random Sampling

Recognize and create simple random samples.

QUICK REFERENCE

- A census collects data from every individual in the population.
- A simple random sample of size n gives every group of n individuals the same chance of being selected.
- A sampling frame is the list of individuals from which the sample is chosen.
- Random sampling helps reduce selection bias.
- TI-84 note: `randInt(` can generate random integers when individuals are numbered.

Example 1. A school numbers all 1200 students from 1 to 1200, then uses random integers to choose 80 student numbers. What sampling method is being used?

> **Answer: Simple random sampling, assuming repeated numbers are ignored and each group of 80 is equally likely.**

Example 2. A teacher surveys the first 30 students who enter the cafeteria. Is this a simple random sample?

> **Answer: No. This is a convenience sample.**

Example 3. A city wants a simple random sample of 50 addresses from a list of $18{,}000$ addresses. What should be numbered?

> **Answer: The addresses in the sampling frame should be numbered from 1 to $18{,}000$.**

Example 4. Why is random selection preferred over letting volunteers respond?

> **Answer: Random selection reduces selection bias; volunteers may differ from the population in important ways.**

Section 1.4 Other Effective Sampling Methods

Classify stratified, cluster, and systematic sampling.

QUICK REFERENCE

- Stratified sampling separates the population into groups, then samples from every group.
- Cluster sampling separates the population into groups, randomly selects some groups, then surveys everyone in those selected groups.
- Systematic sampling selects every kth individual after a random starting point.
- Convenience sampling uses individuals who are easy to reach and is not an effective random method.
- Stratified samples preserve representation from each group; cluster samples are often cheaper to collect.

Example 1. A college samples 40 freshmen, 40 sophomores, 40 juniors, and 40 seniors. What method is this?

Answer: Stratified sampling.

Example 2. A researcher randomly selects 8 classrooms and surveys every student in those classrooms. What method is this?

Answer: Cluster sampling.

Example 3. A quality checker inspects every 25th item coming off an assembly line after a random start. What method is this?

Answer: Systematic sampling.

Example 4. A website posts a poll and lets anyone answer. What sampling concern is present?

Answer: Voluntary response bias; this is not a random sampling method.

Section 1.5 Bias in Sampling

Identify common sources of bias and improve sampling plans.

QUICK REFERENCE

- Sampling bias occurs when the sampling method tends to favor certain outcomes.
- Undercoverage happens when some groups in the population are left out or underrepresented.
- Nonresponse bias happens when selected individuals do not respond.
- Response bias happens when answers are inaccurate because of wording, interviewer effects, or other pressure.
- Wording bias occurs when the question wording influences the response.

Example 1. A phone survey only calls landlines to estimate opinions of all adults. What type of bias is likely?

Answer: Undercoverage, because adults without landlines may be missed.

Example 2. A survey asks, "Do you support the wasteful new tax increase?" What type of bias is present?

Answer: Wording bias.

Example 3. A mail survey is sent to 2000 people, but only 180 respond. What bias is a concern?

Answer: Nonresponse bias.

Example 4. How could a mall-intercept survey about city transportation be improved?

Answer: Use a random sampling method from the full city population instead of only people at the mall.

Section 1.6 The Design of Experiments

Identify experimental design features and explain their purpose.

QUICK REFERENCE

- Experimental units are the individuals or objects receiving treatments.
- A treatment is a condition applied in an experiment.
- Random assignment helps create comparable treatment groups.
- A control group provides a comparison, and a placebo can help measure placebo effect.
- Blinding hides treatment information from subjects or evaluators; double-blinding hides it from both.

Example 1. A study randomly assigns 90 patients to a new medication, old medication, or placebo. Identify the experimental units and treatments.

Answer: Experimental units: the 90 patients. Treatments: new medication, old medication, and placebo.

Example 2. Why should patients be randomly assigned to treatments?

Answer: Random assignment helps balance other factors across the treatment groups.

Example 3. A researcher groups subjects by age group first, then randomly assigns treatments within each age group. What design is being used?

Answer: Randomized block design.

Example 4. In a drug study, neither patients nor doctors evaluating outcomes know who received the real drug. What is this called?

Answer: Double-blind experiment.

Module 2: Organizing and Displaying Data

Students organize qualitative and quantitative data using tables and graphs, then identify misleading displays.

Section 2.1 Organizing Qualitative Data

Build and interpret frequency tables and categorical graphs.

QUICK REFERENCE

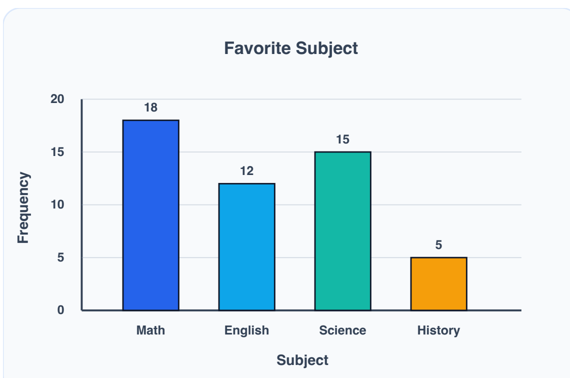
- Frequency is the count in a category.
- Relative frequency is $\frac{\text{category frequency}}{\text{total frequency}}$.
- Bar graphs and Pareto charts are common displays for qualitative data.
- A Pareto chart orders bars from largest frequency to smallest frequency.
- Pie charts show parts of a whole, usually as percentages.

Example 1. Use the table and bar graph to find the relative frequency for each favorite subject.

Subject	Math	English	Science	History
Frequency	18	12	15	5

Favorite Subject Bar Graph

The tallest bar is Math at 18 students, followed by Science, English, and History.

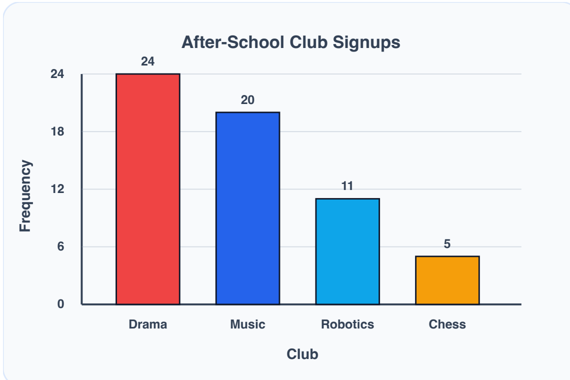


Answer: **Total** $= 50$. **Math:** 0.36; **English:** 0.24; **Science:** 0.30; **History:** 0.10.

Example 2. Use the Pareto chart. Which club has the greatest frequency, and why is this graph a Pareto chart?

Club Signups Pareto Chart

The categories are ordered from largest frequency to smallest frequency.

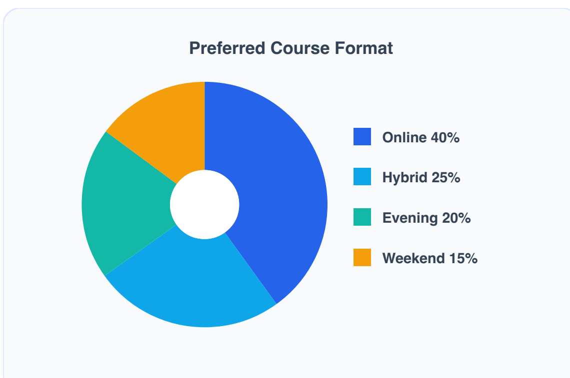


Answer: **Drama has the greatest frequency. It is a Pareto chart because the bars are ordered from largest frequency to smallest frequency.**

Example 3. Use the pie chart. Which format is most popular, and what percent chose either online or hybrid?

Preferred Course Format

Online is the largest slice at 40 percent, and Weekend is the smallest slice at 15 percent.



Answer: **Online is most popular. Online or hybrid is** $40\% + 25\% = 65\%$.

Example 4. Should a histogram or bar graph be used for eye-color categories?

Answer: **Bar graph, because eye color is qualitative.**

Section 2.2 Organizing Quantitative Data: The Popular Displays

Read and create grouped frequency displays for quantitative data.

QUICK REFERENCE

- A frequency distribution groups quantitative data into classes.
- Class width is the distance from one lower class limit to the next lower class limit.
- A histogram uses touching bars because the data are quantitative.
- Distribution shapes may be symmetric, skewed left, skewed right, uniform, or bimodal.
- Relative-frequency histograms show proportions instead of counts.

Example 1. Complete the grouped-frequency idea: the lower class limits are $10, 20, 30, 40$. What is the class width?

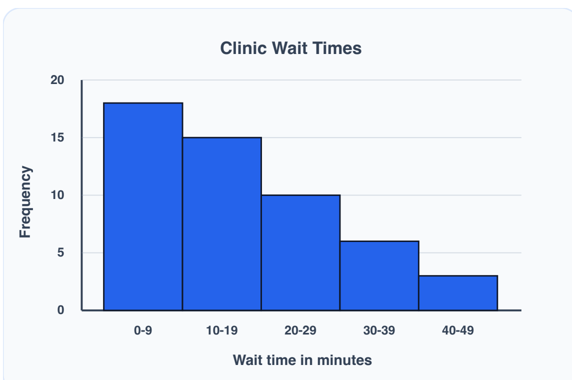
Class	10-19	20-29	30-39	40-49
Frequency	6	11	9	4

Answer: **The class width is $20 - 10 = 10$.**

Example 2. Use the histogram. Describe the shape of the distribution.

Clinic Wait Times

Most wait times are in the lower classes, and the bars taper to the right.

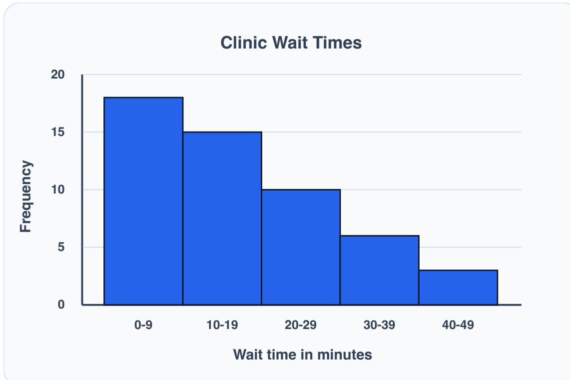


Answer: **Skewed right, because the bars are tallest on the left and taper off to the right.**

Example 3. In the wait-time histogram, which class has the greatest frequency?

Clinic Wait Times

Most wait times are in the lower classes, and the bars taper to the right.



Answer: **The 0-9 minute class has the greatest frequency.**

Example 4. A class has frequency 12 out of 80. What is the relative frequency?

Answer: $\frac{12}{80} = 0.15$, **or** 15%.

Section 2.3

Additional Displays of Quantitative Data

Use dot plots, stem-and-leaf plots, time-series plots, and ogives when appropriate.

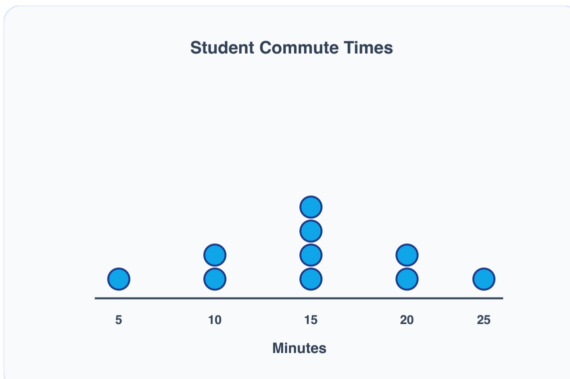
QUICK REFERENCE

- A dot plot stacks or places dots to show each data value.
- A stem-and-leaf plot keeps the original data values visible while organizing them.
- A time-series plot shows data values over time.
- An ogive displays cumulative frequency or cumulative relative frequency.
- Choose a display based on the data type and what pattern you need to show.

Example 1. Use the dot plot. What commute time occurs most often?

Commute Time Dot Plot

Each dot represents one student. The most common commute time is 15 minutes.



Answer: 15 **minutes occurs most often.**

Example 2. Use the stem-and-leaf plot. What original data value is represented by stem 6 with leaf 2?

Key: 4|7 represents 47.

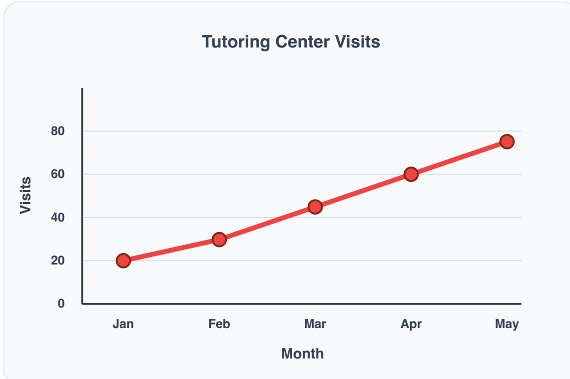
Stem	Leaves
4	2 7 9
5	0 3 8
6	2 2 5
7	1

Answer: **It represents** 62**.**

Example 3. Use the time-series plot. What trend is shown from January through May?

Tutoring Center Visits

The line rises over time, showing an upward trend from January to May.

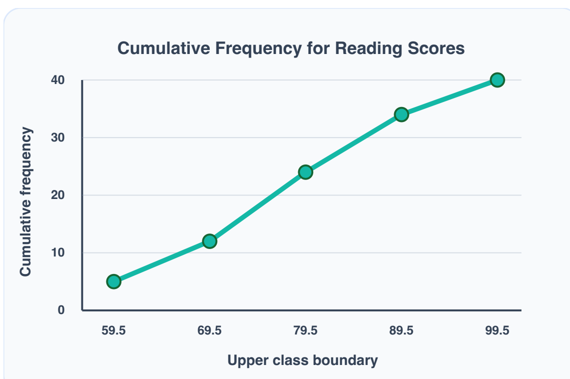


Answer: The number of visits increases over time, so the trend is upward.

Example 4. Use the ogive. What does the point at upper class boundary 79.5 represent?

Reading Score Ogive

The cumulative frequency increases from left to right because each point includes all earlier classes.



Answer: It represents the cumulative frequency at or below 79.5. In this graph, that cumulative frequency is about 24.

Section 2.4 Graphical Misrepresentations of Data

Identify graphs that distort or exaggerate data.

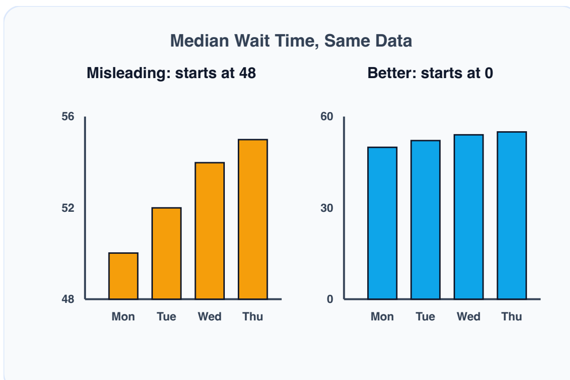
QUICK REFERENCE

- A graph can mislead if the vertical axis is truncated or scaled oddly.
- Pictographs can mislead when pictures change area rather than just height.
- Different class widths can distort histograms.
- Missing labels, units, or context make graphs harder to interpret correctly.
- A good graph should show the data honestly and clearly.

Example 1. Compare the two bar graphs. Why is the left graph misleading?

Same Data, Different Vertical Scales

The left graph starts at 48 and exaggerates the change. The right graph starts at 0 and shows the change more honestly.

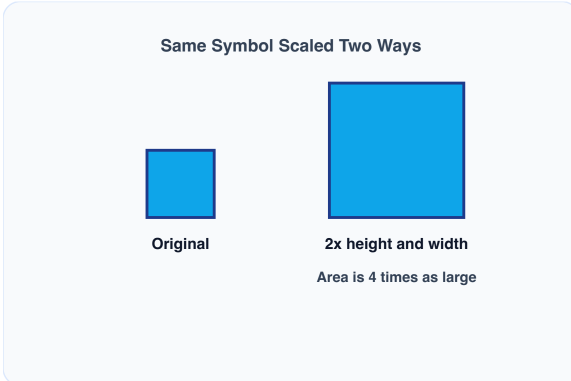


Answer: The left graph uses a truncated vertical axis, so a small change looks much larger than it really is.

Example 2. A pictograph doubles both height and width of a symbol. Why is that misleading?

Pictograph Area Distortion

Doubling both height and width makes the area four times as large.



Answer: **The area becomes four times as large, exaggerating the increase.**

Example 3. A graph has no axis labels or units. What is the problem?

Answer: **The viewer cannot tell what the values represent.**

Example 4. How can a misleading graph usually be improved?

Answer: **Use a clear scale, label axes/units, and avoid visual distortions.**

Module 3: Descriptive Statistics

Students compute and interpret center, spread, position, outliers, five-number summaries, and boxplots.

Section 3.1 Measures of Central Tendency

Find and interpret mean, median, and mode.

QUICK REFERENCE

- The mean is the arithmetic average.
- The median is the middle value after data are ordered.
- The mode is the value that occurs most often.
- The median is resistant to outliers; the mean is not.
- TI-84 note: enter data in `STAT` lists and use `1-Var Stats` for summary statistics.

Example 1. Find the mean, median, and mode of $4, 7, 7, 7, 10, 12$.

Answer: Mean $= 8$; median $= 7$; mode $= 7$.

Example 2. Which measure of center is more resistant to the outlier in $8, 9, 10, 11, 52$?

Answer: Median.

Example 3. A data set has no repeated values. What is the mode?

Answer: There is no mode.

Example 4. A TI-84 `1-Var Stats` output is shown in the answer. Which entries give the mean and median?

Answer: The mean is $\bar{x} = 82.4$. The median is `Med=84`.

$\bar{x}$	82.4	Sx	6.7
n	10	Med	84

Section 3.2 Measures of Dispersion

Use range, variance, and standard deviation to describe spread.

QUICK REFERENCE

- Range is maximum minus minimum.
- Standard deviation measures a typical distance from the mean.
- Sample standard deviation uses s; population standard deviation uses σ.
- The long standard-deviation formulas explain the idea, but calculator output is usually used for computation.
- TI-84 warning: `Sx` is sample standard deviation and `sigma x` is population standard deviation.

Example 1. Find the range of $12, 18, 21, 25, 30$.

Answer: $30 - 12 = 18$.

Example 2. A teacher records scores from one class and treats them as a sample of possible students. Should the TI-84 output use Sx or σx?

Answer: **Use Sx, the sample standard deviation.**

Example 3. A data set includes every employee at a small company. Should the standard deviation be s or σ?

Answer: **σ, because the whole population is included.**

Example 4. Two classes have the same mean score. Class A has standard deviation 4, and Class B has standard deviation 12. Which class has more spread?

Answer: **Class B.**

Section 3.3 Measures of Central Tendency and Dispersion from Tables

Use frequency tables and grouped data to estimate summary measures.

QUICK REFERENCE

- A frequency table can be used like repeated data values.
- For a frequency table, multiply each value by its frequency before adding.
- Grouped-data summaries using class midpoints are approximations.
- TI-84 can use one list for values and another list for frequencies.
- Always interpret summaries in the units of the original variable.

Example 1. Use the frequency table to find the mean.

Score	6	8	10
Frequency	2	3	5

Answer: $\dfrac{6(2)+8(3)+10(5)}{2+3+5}=\dfrac{86}{10}=8.6.$

Example 2. For grouped data, why do we use class midpoints?

Answer: **The exact data values are not known, so midpoints estimate the values in each class.**

Example 3. On the TI-84, what does the frequency list represent in `1-Var Stats`?

Answer: **How many times each corresponding value occurs.**

Section 3.4 Measures of Position and Outliers

Use z-scores, percentiles, quartiles, and outlier fences.

QUICK REFERENCE

- A z-score tells how many standard deviations a value is from the mean: $z = \dfrac{x - \mu}{\sigma}$ or $z = \dfrac{x - \bar{x}}{s}$.
- Positive z-scores are above the mean; negative z-scores are below the mean.
- A percentile describes the percent of data at or below a value.
- IQR is $Q_3 - Q_1$.
- Outlier fences are $Q_1 - 1.5(IQR)$ and $Q_3 + 1.5(IQR)$.

Example 1. A score is 84, the mean is 70, and the standard deviation is 7. Find the z-score.

Answer: $z = \dfrac{84 - 70}{7} = 2.$

Example 2. A student's score is at the 80th percentile. What does that mean?

Answer: About 80% of scores are at or below that student's score.

Example 3. If $Q_1 = 12$ and $Q_3 = 28$, find the IQR.

Answer: $IQR = 28 - 12 = 16.$

Example 4. Using $Q_1 = 12$, $Q_3 = 28$, and $IQR = 16$, find the outlier fences.

Answer: Lower fence $= 12 - 24 = -12$. Upper fence $= 28 + 24 = 52$.

Section 3.5
The Five-Number Summary and Boxplots

Summarize data with quartiles, IQR, outliers, and boxplots.

QUICK REFERENCE

- The five-number summary is minimum, Q_1, median, Q_3, and maximum.
- A boxplot displays the five-number summary visually.
- The box runs from Q_1 to Q_3, and the median is marked inside the box.
- Whiskers extend to non-outlier values.
- Outliers should be plotted separately in a modified boxplot.

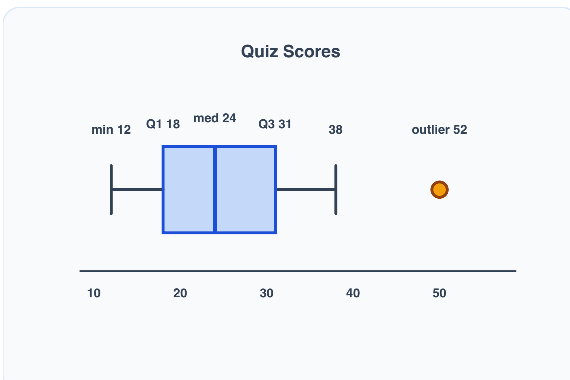
Example 1. Find the five-number summary for $2, 4, 5, 8, 10, 12, 15$.

Answer: **Minimum** 2, $Q_1 = 4$, **median** 8, $Q_3 = 12$, **maximum** 15.

Example 2. Use the modified boxplot. What does the separate point at 52 represent?

Modified Boxplot

The box runs from Q1 to Q3, the median is marked inside the box, and the outlier is plotted separately.



Answer: **It represents an outlier plotted separately from the whisker.**

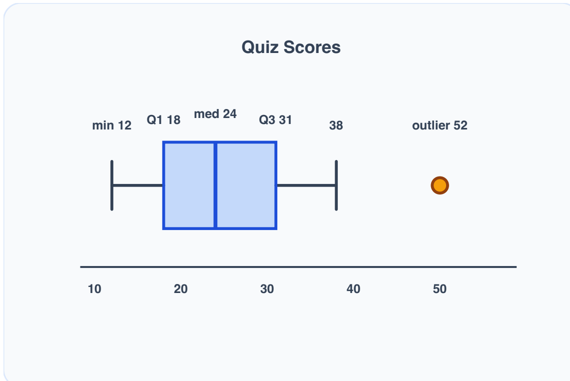
Example 3. If $Q_1 = 20$ and $Q_3 = 35$, what is the IQR?

Answer: 15.

Example 4. In the modified boxplot, what does the line inside the box represent?

Modified Boxplot

The box runs from Q1 to Q3, the median is marked inside the box, and the outlier is plotted separately.



Answer: **The line inside the box represents the median.**

Module 4: Correlation and Regression

Students describe relationships between two quantitative variables and use linear regression carefully.

Section 4.1 Scatter Diagrams and Correlation

Describe scatterplots and interpret the correlation coefficient.

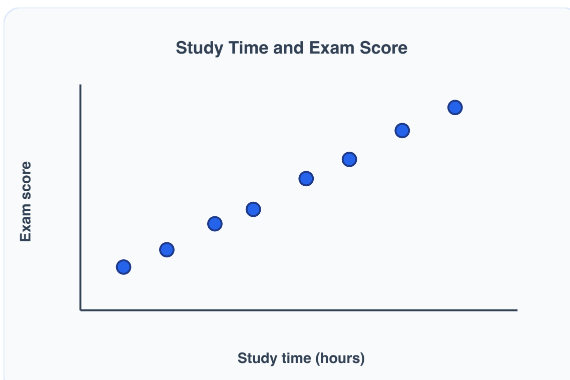
QUICK REFERENCE

- A scatterplot displays paired quantitative data.
- Describe direction, form, and strength when discussing a scatterplot.
- The correlation coefficient r measures strength and direction of a linear relationship.
- The value of r is between -1 and 1.
- Correlation does not imply causation.

Example 1. Use the scatterplot. Describe the direction, form, and strength of the association.

Scatterplot: Study Time and Score

The points rise from left to right and stay close to a line.



Answer: The association is positive, roughly linear, and strong.

Example 2. Which correlation is stronger: $r = -0.91$ or $r = 0.45$?

Answer: $r = -0.91$, because it is closer to -1 or 1.

Example 3. Can $r = 1.24$ be a correlation coefficient?

Answer: No. r must be between -1 and 1.

Example 4. A study finds a strong correlation between ice cream sales and drowning incidents. What causal wording should be avoided?

Answer: Do not say ice cream sales cause drowning incidents.

Section 4.2 Least-Squares Regression

Use regression equations for prediction and interpret slope, intercept, residuals, and extrapolation.

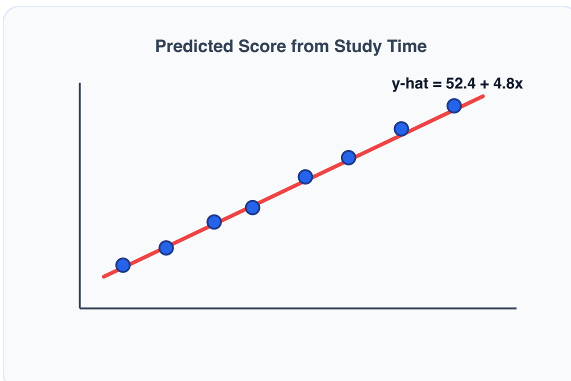
QUICK REFERENCE

- A least-squares regression line predicts a response variable from an explanatory variable.
- The slope gives the predicted change in y for a one-unit increase in x.
- A residual is observed value minus predicted value.
- The coefficient of determination r^2 describes the percent of variation in y explained by the linear model.
- Interpolation predicts inside the data range; extrapolation predicts outside the data range.
- TI-84 note: use `LinReg` for regression and turn diagnostics on when r and r^2 are needed.

Example 1. Use the regression graph. Interpret the slope in context.

Least-Squares Regression Line

The line models the upward linear trend in the paired data.



Answer: **For each additional hour of study time, the predicted exam score increases by about 4.8 points.**

Example 2. A TI-84 `LinReg` output gives $a = 52.4$, $b = 4.8$, $r = 0.94$, and $r^2 = 0.884$. Write the regression equation and interpret r^2.

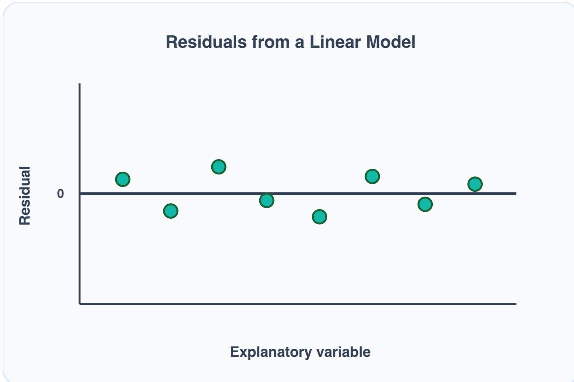
Answer: **$\hat{y} = 52.4 + 4.8x$. About 88.4% of the variation in exam score is explained by the linear model with study time.**

LinReg	$y = a + bx$	a	52.4
b	4.8	r	0.94
r^2	0.884		

Example 3. Use the residual plot. Does the residual plot show an obvious curved pattern?

Residual Plot

Residuals are scattered around zero without a clear curve, which supports a linear model.



Answer: **No. The residuals are scattered around 0, so this plot does not show an obvious curved pattern.**

Example 4. Data were collected for x-values from 2 to 15. Is predicting at $x = 40$ interpolation or extrapolation?

Answer: **Extrapolation.**

Module 5: Probability

Students compute probabilities using complements, addition rules, multiplication rules, conditional probability, and counting techniques.

Section 5.1 Probability Rules

Use basic probability language and simple probability rules.

QUICK REFERENCE

- A probability is always between 0 and 1, inclusive.
- The sample space contains all possible outcomes.
- An event is a collection of outcomes from the sample space.
- The complement of event A is the event that A does not occur.
- The complement rule is $P(A^c) = 1 - P(A)$.

Example 1. If $P(A) = 0.37$, find $P(A^c)$.

Answer: $1 - 0.37 = 0.63$.

Example 2. Can a probability equal 1.08?

Answer: **No. Probabilities cannot be greater than 1.**

Example 3. A fair die is rolled. What is the sample space?

Answer: $\{1, 2, 3, 4, 5, 6\}$.

Example 4. A fair die is rolled. Find the probability of rolling an even number.

Answer: $\frac{3}{6} = \frac{1}{2}$.

Section 5.2 The Addition Rule and Complements

Compute probabilities involving "or" statements.

QUICK REFERENCE

- Mutually exclusive events cannot happen at the same time.
- For mutually exclusive events, $P(A \text{ or } B) = P(A) + P(B)$.
- General addition rule: $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$.
- Use complements when it is easier to find the probability that something does not happen.
- In probability, "or" usually means one event, the other event, or both.

Example 1. If $P(A) = 0.30$, $P(B) = 0.25$, and A and B are mutually exclusive, find $P(A \text{ or } B)$.

Answer: $0.30 + 0.25 = 0.55$.

Example 2. If $P(A) = 0.40$, $P(B) = 0.35$, and $P(A \text{ and } B) = 0.10$, find $P(A \text{ or } B)$.

Answer: $0.40 + 0.35 - 0.10 = 0.65$.

Example 3. A card is drawn from a deck. Are "heart" and "club" mutually exclusive?

Answer: **Yes. One card cannot be both a heart and a club.**

Example 4. If the probability of at least one success is needed, what complement is often easier?

Answer: **Find the probability of no successes, then subtract from 1.**

Section 5.3 Independence and the Multiplication Rule

Use multiplication rules for independent and dependent events.

QUICK REFERENCE

- Independent events do not affect each other's probabilities.
- For independent events, $P(A \text{ and } B) = P(A)P(B)$.
- Without replacement usually creates dependent events.
- With replacement often keeps probabilities the same.
- For dependent events, update the probability after the first event.

Example 1. A coin is flipped twice. Find the probability of two heads.

Answer: $\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$.

Example 2. A bag has 5 red and 7 blue marbles. Two marbles are drawn without replacement. Find the probability both are red.

Answer: $\frac{5}{12} \cdot \frac{4}{11} = \frac{5}{33}$.

Example 3. Are two draws from a deck independent if the first card is not replaced?

Answer: **No. The first draw changes the deck.**

Example 4. Are two rolls of a fair die independent?

Answer: **Yes. The first roll does not affect the second roll.**

Section 5.4 Conditional Probability and the General Multiplication Rule

Compute and interpret conditional probabilities.

QUICK REFERENCE

- Conditional probability $P(A \mid B)$ means the probability of A given that B has occurred.
- $P(A \mid B) = \dfrac{P(A \text{ and } B)}{P(B)}$.
- $P(A \mid B)$ and $P(B \mid A)$ are usually different.
- Two-way tables are common sources for conditional probabilities.
- General multiplication rule: $P(A \text{ and } B) = P(B)P(A \mid B)$.

Example 1. If $P(A \text{ and } B) = 0.18$ and $P(B) = 0.30$, find $P(A \mid B)$.

Answer: $\dfrac{0.18}{0.30} = 0.60$.

Example 2. Out of 80 students, 30 take statistics and 18 of those also take biology. Find the probability a student takes biology given they take statistics.

Answer: $\dfrac{18}{30} = 0.60$.

Example 3. Does $P(A \mid B)$ always equal $P(B \mid A)$?

Answer: No.

Example 4. If $P(B) = 0.40$ and $P(A \mid B) = 0.25$, find $P(A \text{ and } B)$.

Answer: $0.40(0.25) = 0.10$.

Section 5.5 Counting Techniques

Use the fundamental counting principle, permutations, and combinations.

QUICK REFERENCE

- The fundamental counting principle multiplies choices across stages.
- A factorial $n!$ multiplies $n(n-1)(n-2)\cdots 1$.
- Use permutations when order matters.
- Use combinations when order does not matter.
- TI-84 note: use nPr for permutations and nCr for combinations.

Example 1. A code uses 3 letters followed by 2 digits. If repetition is allowed, how many codes are possible?

Answer: $26^3 \cdot 10^2 = 1{,}757{,}600.$

Example 2. How many ways can 4 students be arranged in a line from a group of 10?

Answer: $10P4 = 5040.$

Example 3. How many committees of 4 can be chosen from 10 students?

Answer: $10C4 = 210.$

Example 4. Should order matter when choosing a committee?

Answer: No, so use a combination.

Module 6: Discrete Random Variables and Binomial Distributions

Students work with discrete probability distributions, expected value, and binomial probabilities.

Section 6.1 Discrete Random Variables

Determine whether a distribution is valid and compute expected values.

QUICK REFERENCE

- A random variable assigns a numerical value to outcomes of a probability experiment.
- A discrete random variable has countable possible values.
- A probability distribution must have probabilities between 0 and 1, and the probabilities must sum to 1.
- Expected value is the long-run average: $\mu = \sum xP(x)$.
- The standard deviation of a random variable describes typical distance from its expected value.
- TI-84 can compute weighted statistics using values in one list and probabilities or frequencies in another list.

Example 1. Is this a valid probability distribution?

x	0	1	2	3
$P(x)$	0.20	0.35	0.25	0.20

Answer: **Yes. All probabilities are between 0 and 1, and they sum to 1.**

Example 2. Find the expected value for $x = 0, 1, 2$ with probabilities $0.2, 0.5, 0.3$.

Answer: $\mu = 0(0.2) + 1(0.5) + 2(0.3) = 1.1$.

Example 3. A TI-84 weighted `1-Var Stats` output gives $\bar{x} = 1.1$ and $\sigma x = 0.7$ for a probability distribution. Interpret the standard deviation.

Answer: **Values of the random variable typically vary about 0.7 from the expected value, in the long-run distribution sense.**

Example 4. A probability table sums to 1.08. Is it valid?

Answer: **No. The probabilities must sum to 1.**

Section 6.2 The Binomial Probability Distribution

Recognize binomial settings and compute binomial probabilities with technology.

QUICK REFERENCE

- A binomial experiment has fixed trials, independent trials, two outcomes, and constant success probability.
- Use n for number of trials, p for success probability, $q = 1 - p$, and x for number of successes.
- The binomial formula $P(X = x) = \binom{n}{x} p^x q^{n-x}$ shows the structure.
- For a binomial distribution, $\mu = np$ and $\sigma = \sqrt{npq}$.
- TI-84 `binompdf(n,p,x)` finds $P(X = x)$.
- TI-84 `binomcdf(n,p,x)` finds $P(X \leq x)$.

Example 1. A basketball player makes 70% of free throws and shoots 8 free throws. Identify n and p.

Answer: $n = 8$, $p = 0.70$.

Example 2. Use binomial notation to compute the probability of exactly 3 successes in 10 trials when $p = 0.40$.

Answer: **Use `binompdf(10,0.40,3)`; approximately** 0.215.

Example 3. For $X \sim Binomial(12, 0.25)$, what TI-84 command gives $P(X \leq 4)$?

Answer: `binomcdf(12,0.25,4)`.

Example 4. For $X \sim Binomial(12, 0.25)$, how can $P(X \geq 5)$ be found using the complement?

Answer: $P(X \geq 5) = 1 - P(X \leq 4)$, **so use** $1-$`binomcdf(12,0.25,4)`.

Example 5. For $X \sim Binomial(12, 0.25)$, find the mean and standard deviation.

Answer: $\mu = np = 12(0.25) = 3$. **Since** $q = 0.75$, $\sigma = \sqrt{12(0.25)(0.75)} = 1.5$.

Module 7: Normal Distributions

Students use normal distribution properties, normal probabilities, inverse normal values, and normality checks.

Section 7.1 Properties of the Normal Distribution

Understand the normal curve, z-scores, and the empirical rule.

QUICK REFERENCE

- A normal distribution is bell-shaped and symmetric.
- For a normal distribution, mean, median, and mode are equal.
- The total area under a normal curve is 1.
- The z-score formula is $z = \dfrac{x - \mu}{\sigma}$.
- The empirical rule gives about 68%, 95%, and 99.7% within 1, 2, and 3 standard deviations of the mean.

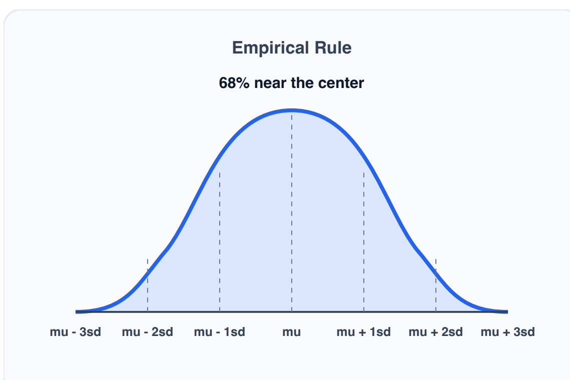
Example 1. A normal distribution has mean 50 and standard deviation 6. Find the z-score for $x = 62$.

Answer: $z = \dfrac{62 - 50}{6} = 2$.

Example 2. Use the normal curve. Using the empirical rule, about what percent of values lie within 2 standard deviations of the mean?

Empirical Rule on a Normal Curve

About 68 percent is within one standard deviation, 95 percent within two, and 99.7 percent within three.



Answer: About 95%.

Example 3. A normal distribution has mean 100 and standard deviation 15. What interval is within 1 standard deviation?

Answer: 85 to 115.

Example 4. What is the total area under a normal curve?

Answer: 1, or 100%.

Section 7.2 Applications of the Normal Distribution

Use TI-84 normalcdf and invNorm to compute probabilities and cutoff values.

QUICK REFERENCE

- Use `normalcdf(lower, upper, mean, standard deviation)` for normal probabilities.
- Use a very large negative lower bound for "less than" probabilities.
- Use a very large positive upper bound for "greater than" probabilities.
- Use `invNorm(area, mean, standard deviation)` to find a value from a left-tail area.
- Always sketch or name the shaded region before choosing the calculator command.

Example 1. Let X be normal with $\mu = 70$, $\sigma = 8$. What TI-84 command finds $P(X < 82)$?

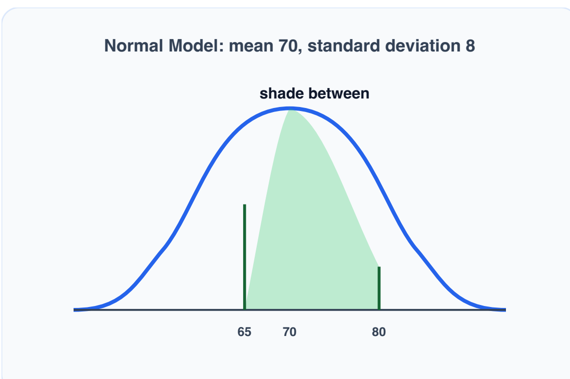
Answer: `normalcdf(-1E99,82,70,8)`.

Example 2. Use the shaded normal curve. What TI-84 command finds \(P(65

Answer: `normalcdf(65,80,70,8)`.

Shaded Normal Probability

The shaded area represents P(65 less than X less than 80).



Example 3. Let X be normal with $\mu = 70$, $\sigma = 8$. What TI-84 command finds the 90th percentile?

Answer: `invNorm(0.90,70,8)`.

Example 4. If the calculator gives $P(X > 90) = 0.0228$, interpret the answer.

Answer: About 2.28% of values are greater than 90.

Section 7.3 Assessing Normality

Use graphs and normal probability plots to decide whether normal methods are reasonable.

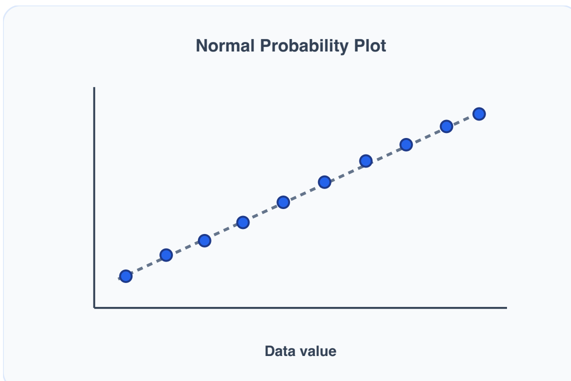
QUICK REFERENCE

- A histogram that is roughly bell-shaped supports normality.
- Strong skewness or outliers are warnings against normality.
- A normal probability plot that is roughly linear supports normality.
- Normality checks are about whether normal-based methods are reasonable, not whether data are perfectly normal.
- StatCrunch is useful for normal probability plots and larger data sets.

Example 1. Use the normal probability plot. What does this suggest about normality?

Normal Probability Plot

The points are close to a straight line, so a normal model appears reasonable.



Answer: The data appear approximately normal because the points are close to a straight line.

Example 2. A histogram is strongly skewed right with several outliers. Are normal methods automatically reasonable?

Answer: No. Strong skewness and outliers are concerns.

Example 3. What graph is specifically designed to assess whether data are approximately normal?

Answer: Normal probability plot.

Example 4. Does a data set have to be perfectly normal for normal methods to be useful?

Answer: No. The data need to be approximately normal enough for the method being used.

Module 8: Sampling Distributions

Students study sampling distributions for sample means and sample proportions.

Section 8.1 Distribution of the Sample Mean

Use the sampling distribution of $\bar{x}$ and the Central Limit Theorem.

QUICK REFERENCE

- The sampling distribution of $\bar{x}$ describes sample means from repeated samples of the same size.
- The mean of $\bar{x}$ is $\mu_{\bar{x}} = \mu$.
- The standard deviation of $\bar{x}$ is $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$, also called standard error.
- The Central Limit Theorem says $\bar{x}$ is approximately normal for large samples under the usual conditions.
- Use TI-84 `normalcdf` with the standard error for probabilities involving sample means.

Example 1. A population has $\mu = 80$, $\sigma = 12$, and samples of size 36. Find $\mu_{\bar{x}}$ and $\sigma_{\bar{x}}$.

Answer: $\mu_{\bar{x}} = 80$, $\sigma_{\bar{x}} = \dfrac{12}{\sqrt{36}} = 2$.

Example 2. For the same setting, what TI-84 command finds $P(\bar{x} < 83)$?

Answer: `normalcdf(-1E99,83,80,2)`.

Example 3. What happens to the standard error as sample size increases?

Answer: **It decreases.**

Example 4. Why is the Central Limit Theorem useful?

Answer: **It lets sample means be modeled approximately normally in many large-sample situations.**

Section 8.2 Distribution of the Sample Proportion

Use the sampling distribution of $\hat{p}$ and check normal approximation conditions.

QUICK REFERENCE

- The sampling distribution of $\hat{p}$ describes sample proportions from repeated samples of the same size.
- The mean is $\mu_{\hat{p}} = p$.
- The standard deviation is $\sigma_{\hat{p}} = \sqrt{\dfrac{p(1-p)}{n}}$.
- The normal approximation usually requires $np \geq 10$ and $n(1-p) \geq 10$.
- Use TI-84 `normalcdf` with p and the standard error for probabilities involving $\hat{p}$.

Example 1. Let $p = 0.40$ and $n = 100$. Find the mean and standard deviation of $\hat{p}$.

Answer: Mean $= 0.40$. Standard deviation $= \sqrt{\dfrac{0.40(0.60)}{100}} \approx 0.049$.

Example 2. Check the normal approximation conditions for $p = 0.40$, $n = 100$.

Answer: $np = 40$ and $n(1-p) = 60$, so both are at least 10.

Example 3. What TI-84 command finds $P(\hat{p} < 0.45)$ when $p = 0.40$, $n = 100$, and standard error is 0.049?

Answer: `normalcdf(-1E99,0.45,0.40,0.049)`.

Example 4. If $np = 7$, what concern should be noted?

Answer: The normal approximation condition is not met.

Module 9: Confidence Intervals

Students estimate population proportions and means using confidence intervals.

Section 9.1 Estimating a Population Proportion

Use one-proportion confidence intervals and interpret them correctly.

QUICK REFERENCE

- A point estimate for a population proportion is $\hat{p} = \dfrac{x}{n}$.
- A confidence interval estimates a population parameter with a margin of error.
- For a one-proportion interval, check large-count conditions before using normal methods.
- TI-84 note: use `1-PropZInt` for one-proportion confidence intervals.
- Correct wording: we are confident the interval captures the population proportion.

Example 1. In a sample of 250 adults, 142 support a proposal. Find $\hat{p}$.

Answer: $\hat{p} = \dfrac{142}{250} = 0.568.$

Example 2. A sample has $x = 142$ supporters out of $n = 250$. Use TI-84 `1-PropZInt` for a 95% confidence interval and interpret the output.

Answer: Use `1-PropZInt` with $x = 142$, $n = 250$, and `C-Level=.95`. We are 95% confident the population proportion is between about 0.507 and 0.629.

Input	$x = 142$	$n = 250$	C-Level $= 0.95$
Output	$(0.507,\ 0.629)$, $\hat{p} = 0.568$		

Example 3. A 95% confidence interval is $(0.51, 0.63)$. Interpret it in context.

Answer: We are 95% confident the population proportion is between 0.51 and 0.63.

Example 4. Is it correct to say there is a 95% probability the fixed population proportion is in this computed interval?

Answer: No. Use confidence wording, not probability wording about the fixed parameter.

Section 9.2 Estimating a Population Mean

Use one-sample t-intervals and interpret them correctly.

QUICK REFERENCE

- A point estimate for a population mean is $\bar{x}$.
- Use a t-interval for a population mean when σ is unknown.
- Degrees of freedom for one sample are $df = n - 1$.
- TI-84 note: use `TInterval` with data or summary statistics.
- Check normality/large-sample conditions before using the interval.

Example 1. A sample has $n = 18$. What are the degrees of freedom for a one-sample t-interval?

Answer: $df = 17$.

Example 2. Use TI-84 `TInterval` with summary statistics $n = 18$, $\bar{x} = 14.6$, $s = 3.8$, and confidence level 90%. Interpret the interval.

Answer: Use `TInterval` with Stats input. The interval is approximately $(13.0, 16.2)$. We are 90% confident the population mean is between 13.0 and 16.2.

Input	$\bar{x} = 14.6$	$Sx = 3.8$	$n = 18$	C-Level $= 0.90$
Output	$(13.0,\ 16.2)$			

Example 3. A 90% confidence interval for mean wait time is $(12.4, 16.8)$ minutes. Interpret it.

Answer: We are 90% confident the population mean wait time is between 12.4 and 16.8 minutes.

Example 4. Why should outliers be checked for a small-sample t-interval?

Answer: Outliers can make the t-interval unreliable when the sample size is small.

Module 10: Hypothesis Testing

Students set up and run hypothesis tests for population proportions and means using correct statistical wording.

Section 10.1 The Language of Hypothesis Testing

Use hypothesis-test vocabulary and write correct conclusions.

QUICK REFERENCE

- The null hypothesis H_0 is the starting claim tested by the procedure.
- The alternative hypothesis H_a describes what the evidence is trying to support.
- The significance level α is the cutoff for deciding whether the evidence is statistically significant.
- The P-value is the probability of observing results as extreme or more extreme, assuming H_0 is true.
- A Type I error rejects a true H_0; a Type II error fails to reject a false H_0.
- Decisions are reject H_0 or fail to reject H_0. Do not say accept H_0.

Example 1. A claim says the population proportion is greater than 0.40. Write the alternative hypothesis.

Answer: $H_a : p > 0.40$.

Example 2. If $H_a : \mu < 25$, what type of test is this?

Answer: **Left-tailed test.**

Example 3. If $P\text{-value} = 0.032$ and $\alpha = 0.05$, what is the decision?

Answer: **Reject H_0, because $0.032 < 0.05$.**

Example 4. If the P-value is large, should we say we accept H_0?

Answer: **No. Say fail to reject H_0.**

Example 5. In a medical screening study, H_0 says the test has the advertised accuracy. What kind of error is rejecting H_0 when the advertised accuracy is actually true?

Answer: **Type I error, because a true null hypothesis was rejected.**

Section 10.2 Hypothesis Tests for a Population Proportion

Use one-proportion z-tests with TI-84 and state conclusions correctly.

QUICK REFERENCE

- Use a one-proportion z-test when testing a claim about one population proportion.
- Set hypotheses using the population proportion p, not the sample proportion $\hat{p}$.
- Check large-count conditions using the null value p_0.
- TI-84 note: use `1-PropZTest` for a one-proportion hypothesis test.
- Conclusion should include the decision and context tied to the original claim.

Example 1. A claim says fewer than 30% of students work full time. Write H_0 and H_a.

Answer: $H_0 : p = 0.30$; $H_a : p < 0.30$.

Example 2. For $H_0 : p = 0.30$, $n = 120$, check the large-count conditions.

Answer: $np_0 = 36$ **and** $n(1 - p_0) = 84$**, so both are at least** 10**.**

Example 3. A sample of 120 students has 27 who work full time. Test the claim $p < 0.30$ at $\alpha = 0.05$ using TI-84 `1-PropZTest`.

Answer: Use $p_0 = 0.30$, $x = 27$, $n = 120$, **and** $p < p_0$. **Since** $P\text{-value} \approx 0.037 < 0.05$, **reject** H_0. **There is sufficient evidence to support the claim that the population proportion is less than** 0.30.

Input	$p_0 = 0.30$	$x = 27$	$n = 120$	**prop \\(**
Output	$z \approx -1.79$	$p \approx 0.037$	$\hat{p} = 0.225$	

Example 4. A test gives $P\text{-value} = 0.081$ with $\alpha = 0.05$. What is the decision?

Answer: Fail to reject H_0.

Example 5. State the conclusion for a failed rejection when the claim was $p < 0.30$.

Answer: There is not sufficient evidence to support the claim that the population proportion is less than 0.30.

Section 10.3 Hypothesis Tests for a Population Mean

Use one-sample t-tests with TI-84 and state conclusions correctly.

QUICK REFERENCE

- Use a one-sample t-test when testing a claim about one population mean and σ is unknown.
- Set hypotheses using the population mean μ, not the sample mean $\bar{x}$.
- Check normality or large-sample conditions before trusting the test.
- TI-84 note: use `T-Test` with data or summary statistics.
- Reject H_0 or fail to reject H_0, then state the conclusion in context.

Example 1. A claim says the population mean wait time is different from 15 minutes. Write H_0 and H_a.

Answer: $H_0 : \mu = 15$; $H_a : \mu \neq 15$.

Example 2. If $H_a : \mu > 42$, what type of test is this?

Answer: **Right-tailed test.**

Example 3. A sample has $n = 16$, $\bar{x} = 43.8$, and $s = 3.2$. Test the claim $\mu > 42$ at $\alpha = 0.05$ using TI-84 `T-Test`.

Answer: **Use Stats input with $\mu_0 = 42$, $\bar{x} = 43.8$, $Sx = 3.2$, $n = 16$, and $\mu > \mu_0$. Since P-value $\approx 0.020 < 0.05$, reject H_0. There is sufficient evidence to support the claim that the population mean is greater than 42.**

Input	$\mu_0 = 42$	$\bar{x} = 43.8$	$Sx = 3.2$	$n = 16$
Output	$t = 2.25$	$p \approx 0.020$	df $= 15$	

Example 4. A t-test gives $P\text{-value} = 0.014$ with $\alpha = 0.05$. What is the decision?

Answer: **Reject H_0.**

Example 5. State the conclusion for rejecting H_0 when the claim was $\mu > 42$.

Answer: **There is sufficient evidence to support the claim that the population mean is greater than 42.**