CHAPTER 17: Tests of Significance: The Basics

Basic Practice of Statistics
7th Edition

Lecture PowerPoint Slides

In Chapter 17, we cover …

- The reasoning of tests of significance
- Stating hypotheses
- *P*-value and statistical significance
- Tests for a population mean
- Significance from a table

Statistical inference

- Confidence intervals are one of the two most common types of statistical inference. Use a confidence interval when your goal is to estimate a population parameter.
- The second common type of inference, called *tests of significance*, has a different goal: to assess the evidence provided by data about some claim concerning a population.

  - A test of significance is a formal procedure for comparing observed data with a claim (also called a hypothesis) whose truth we want to assess.
  - Significance tests use an elaborate vocabulary, but the basic idea is simple: *an outcome that would rarely happen if a claim were true is good evidence that the claim is not true*.

What are Tests of Significance?

- Claim: “John gets 80% of free shots”
- Data: results on 1000 free shots.
- Law of large numbers: average number of shots scored should be close to the true scoring percentage.
- Data: average number of shots scored is 60%.

Reasoning of Tests of Significance

- How likely would it be to see the results we saw if the claim were true?

  (if John truly gets 80% of his shots, how likely is he to get only 60% in 1000 shots?!!)
- Do the data give enough evidence against the claim?

  (what if John scored 75%? 20%?)

The reasoning of tests of significance (part I)

- Artificial sweeteners in colas gradually lose their sweetness over time. Manufacturers test for loss of sweetness on a scale of –10 to 10, with negative scores corresponding to a gain in sweetness, positive to a loss of sweetness.
- Suppose we know that for any cola, the sweetness loss scores vary from taster to taster according to a Normal distribution with standard deviation $\sigma = 1$. The mean μ for all tasters measures loss of sweetness and is different for different colas.
- Here are the sweetness losses for a cola currently on the market, as measured by 10 trained tasters:

1.6 0.4 0.5 –2.0 1.5 –1.1 1.3 –0.1 –0.3 1.2

- The average sweetness loss is given by the sample mean $\bar{x} = 0.3$. Most scores were positive. That is, most tasters found a loss of sweetness. But the losses are small, and two tasters (the negative scores) thought the cola gained sweetness. Are these data good evidence that the cola lost sweetness in storage?

The reasoning of tests of significance (part II)

- We make a claim and ask if the data give evidence against it.
- We seek evidence that there is a sweetness loss, so the claim we test is that there is *not* a loss. In that case, the mean loss for the population of all trained testers would be $\mu = 0$.
- If the claim that $\mu = 0$ is true, the sampling distribution of $\bar{x}$ from 10 tasters is Normal, with mean $\mu = 0$ and standard deviation $\frac{\sigma}{\sqrt{n}} = \frac{1}{\sqrt{10}} = 0.316$.

The reasoning of tests of significance (part III)

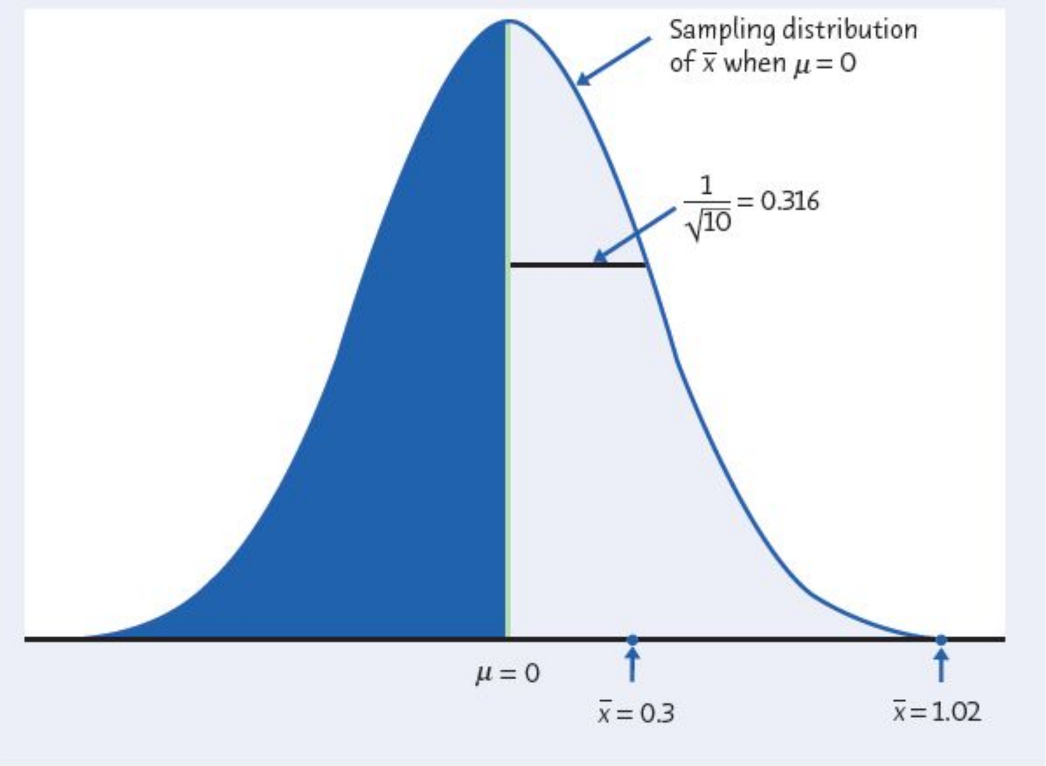
- If the claim that $\mu = 0$ is true (no loss of sweetness, on average), the sampling distribution of $\bar{x}$ from 10 tasters is Normal with $\mu = 0$ and standard deviation

$$\frac{\sigma}{\sqrt{n}} = \frac{1}{\sqrt{10}} \approx 0.316$$

- The data yielded $\bar{x} = 1.02$, which is more than three standard deviations from $\mu = 0$. This is strong evidence that the new cola lost sweetness in storage.
- If the data yielded $\bar{x} = 0.3$, which is less than one standard deviations from $\mu = 0$, there would be no evidence that the new cola lost sweetness in storage.

The reasoning of tests of significance (illustrated)

- This is like calculations we did in Chapter 15—we can locate our $\bar{x}$ of 0.3 in this distribution and comment on whether it is surprising.



Reasoning of Significance Tests (1 of 2)

If you observe an outcome that would __________ happen if a claim were true, you have good evidence that the claim is ________.

a) frequently; true
b) rarely; true
c) frequently; not true
d) rarely; not true

Reasoning of Significance Tests (1 of 2) (answer)

If you observe an outcome that would __________ happen if a claim were true, you have good evidence that the claim is ________.

a) frequently; true
b) rarely; true
c) frequently; not true
d) **rarely; not true**

The correct answer is D.

Reasoning of Significance Tests (2 of 2)

- The manufacturer of a certain toy claims that the mean lead content in this toy is 100 parts per million (ppm). A sample of 25 such toys produced $\bar{x}$ = 105.1 ppm. This would provide evidence against the manufacturer's claim if we can show that:

a) 105.1 could reasonably occur by chance if the claim were true.
b) 105.1 could rarely occur by chance if the claim were true.
c) 100 could rarely occur by chance if the claim were true.
d) 100 could reasonably occur by chance if the claim were true.

Reasoning of Significance Tests (2 of 2) (answer)

- The manufacturer of a certain toy claims that the mean lead content in this toy is 100 parts per million (ppm). A sample of 25 such toys produced $\bar{x}$ = 105.1 ppm. This would provide evidence against the manufacturer's claim if we can show that:

a) 105.1 could reasonably occur by chance if the claim were true.
b) 105.1 could rarely occur by chance if the claim were true.
c) 100 could rarely occur by chance if the claim were true.
d) 100 could reasonably occur by chance if the claim were true.

The correct answer is B.

Stating hypotheses

- A significance test starts with a careful statement of the claims we want to compare.
- The claim tested by a statistical test is called the null hypothesis (H_0). The test is designed to assess the strength of the evidence against the null hypothesis. Often the null hypothesis is a statement of "no difference."
- The claim about the population that we are trying to find evidence for is the alternative hypothesis (H_a). The alternative is one-sided if it states that a parameter is *larger* or *smaller* than the null hypothesis value. It is two-sided if it states that the parameter is *different from* the null value (it could be either smaller or larger).
- In the sweetness example, our hypotheses are:

$$H_0: \mu = 0$$

$$H_a: \mu > 0$$

- The alternative hypothesis is one-sided because we are interested only in whether the cola lost sweetness.

Stating hypotheses: two-tailed example (part I)

- Does the job satisfaction of assembly workers differ when their work is machine-paced rather than self-paced? Assign workers either to an assembly line moving at a fixed pace or to a self-paced setting. All subjects work in both settings, in random order. After two weeks in each work setting, the workers take a test of job satisfaction.

- This is a matched pairs design.
- The response variable is the difference in satisfaction scores: self-paced minus machine-paced.
- The parameter of interest is the mean μ of the differences in scores in the population of all assembly workers.

Stating hypotheses: two-tailed example (part II)

- Here, our response variable is the difference in satisfaction scores; the parameter of interest is the mean μ of the differences in scores in the population of all assembly workers.
- The null hypothesis says that there is no difference between self-paced work and machine-paced work:

$$H_0\colon \mu = 0$$

- The authors of the study wanted to know if the two work conditions have different levels of job satisfaction. They did not specify the direction of the difference. The alternative hypothesis is therefore *two-sided*:

$$H_a\colon \mu \neq 0$$

- ***Caution!*** The hypotheses should express the hopes or suspicions we have before we see the data. It is cheating to first look at the data and then frame hypotheses to fit what the data show.

Test Statistic

Testing the Mean of a Normal Population

Take an SRS of size n from a Normal population with unknown mean μ and known standard deviation σ. The test statistic for hypotheses about the mean (H_o: $\mu = \mu_o$) of a Normal distribution is the standardized version of $\bar{x}$:
(called the

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

z-score)

Case Study

Sweetening Colas

If the null hypothesis of no average sweetness loss is true, the test statistic would be:

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{1.02 - 0}{1 / \sqrt{10}} \approx 3.23$$

Because the sample result is more than 3 standard deviations above the hypothesized mean 0, it gives strong evidence that the mean sweetness loss is not 0, but positive.

Examples and Exercises(Making Hypotheses

Grading a teaching assistant. The examinations in a large statistics class are scaled after grading so that the mean score is 75. The professor thinks that one teaching assistant is a poor teacher and suspects that his students have a lower mean score than the class as a whole. The TA's students this semester can be considered a sample from the population of all students in the course, so the professor compares their mean score with 75. State the hypotheses H_0 and H_a.

Examples and Exercises(Making Hypotheses

Women's incomes. The average income of American women who work full-time and have only a high school degree is \$31,666. You wonder whether the mean income of female graduates from your local high school who work full-time but have only a high school degree is different from the national average. You obtain income information from an SRS of 62 female graduates who work full-time and have only a high school degree and find that $\bar{x} = \$30{,}052$. What are your null and alternative hypotheses?

Examples and Exercises(Making Hypotheses

Stating hypotheses. In planning a study of the annual consumption of carbonated soft drinks by high school students, a researcher states the hypotheses as

$$H_0: \overline{x} = 60 \text{ gallons per year}$$

$$H_a: \overline{x} > 60 \text{ gallons per year}$$

What's wrong with this?

Stating Hypotheses (1 of 6)

- The manufacturer of a certain toy claims that the mean lead content in this toy is 100 ppm. The Consumer Product Safety Commission takes a random sample of 25 such toys to evaluate the manufacturer's claim. What is the commission's null hypothesis?

a) H_0: $\bar{x} < 100$ ppm
b) H_0: $\bar{x} \neq 100$ ppm
c) H_0: $\bar{x} = 100$ ppm
d) H_0: $\mu < 100$ ppm
e) H_0: $\mu \neq 100$ ppm
f) H_0: $\mu = 100$ ppm

Stating Hypotheses (1 of 6) (answer)

- The manufacturer of a certain toy claims that the mean lead content in this toy is 100 ppm. The Consumer Product Safety Commission takes a random sample of 25 such toys to evaluate the manufacturer's claim. What is the commission's null hypothesis?

a) H_0: $\bar{x} < 100$ ppm
b) H_0: $\bar{x} \neq 100$ ppm
c) H_0: $\bar{x} = 100$ ppm
d) H_0: $\mu < 100$ ppm
e) H_0: $\mu \neq 100$ ppm
***f)* H_0: $\mu = 100$ ppm**

The correct answer is F.

Stating Hypotheses (2 of 6)

- A consumer advocate evaluates the claim that a new granola cereal contains "4 ounces of cashews in every bag." She recognizes that the amount of cashews will vary slightly from bag to bag but suspects that the mean amount of cashews per bag is actually less than 4 ounces. The advocate purchases a random sample of 40 bags of cereal and calculates $\bar{x}$ to be 3.68 ounces. What alternative hypothesis does she want to test?

a) $H_a : \bar{x} < 3.68$ ounces
b) $H_a : \bar{x} \neq 3.68$ ounces
c) $H_a : \bar{x} > 3.68$ ounces
d) $H_a : \mu < 4$ ounces
e) $H_a : \mu \neq 4$ ounces
f) $H_a : \mu > 4$ ounces

Stating Hypotheses (2 of 6) (answer)

- A consumer advocate evaluates the claim that a new granola cereal contains "4 ounces of cashews in every bag." She recognizes that the amount of cashews will vary slightly from bag to bag but suspects that the mean amount of cashews per bag is actually less than 4 ounces. The advocate purchases a random sample of 40 bags of cereal and calculates $\bar{x}$ to be 3.68 ounces. What alternative hypothesis does she want to test?

a) $H_a : \bar{x} < 3.68$ ounces
b) $H_a : \bar{x} \neq 3.68$ ounces
c) $H_a : \bar{x} > 3.68$ ounces
d) $\boldsymbol{H_a : \mu < 4}$ **ounces**
e) $H_a : \mu \neq 4$ ounces
f) $H_a : \mu > 4$ ounces

The correct answer is D.

Stating Hypotheses (3 of 6)

Null and alternative hypotheses always refer to:

a) a sample statistic.
b) a sampling distribution.
c) a population parameter.

Stating Hypotheses (3 of 6) (answer)

Null and alternative hypotheses always refer to:

a) a sample statistic.
b) a sampling distribution.
c) **a population parameter.**

The correct answer is C.

Stating Hypotheses (4 of 6)

- The test H_0: $\mu = 40$ versus H_a: $\mu < 40$ is:

a) one-sided (left tail).
b) one-sided (right tail).
c) two-sided.

Stating Hypotheses (4 of 6) (answer)

- The test H_0: $\mu = 40$ versus H_a: $\mu < 40$ is:

a) **one-sided (left tail).**
b) one-sided (right tail).
c) two-sided.

The correct answer is A.

Stating Hypotheses (5 of 6)

- The test H_0: $\mu = 40$ versus H_a: $\mu \neq 40$ is:

a) one-sided (left tail).

b) one-sided (right tail).

c) two-sided.

Stating Hypotheses (5 of 6) (answer)

- The test H_0: $\mu = 40$ versus H_a: $\mu \neq 40$ is:

a) one-sided (left tail).
b) one-sided (right tail).
c) **two-sided.**

The correct answer is C.

Stating Hypotheses (6 of 6)

The hypotheses should express the hopes or suspicions we have ________ we see the data.

a) after
b) before
c) at the same time

Stating Hypotheses (6 of 6) (answer)

The hypotheses should express the hopes or suspicions we have ________ we see the data.

a) after
b) before
c) at the same time

The correct answer is B.

P-value and statistical significance (part I)

- The null hypothesis H_0 states the claim that we are seeking evidence against. The probability that measures the strength of the evidence against a null hypothesis is called a ***P*-value**.

- A test statistic calculated from the sample data measures how far the data diverge from what we would expect if the null hypothesis H_0 were true. Large values of the statistic show that the data are not consistent with H_0.
- The probability, computed assuming H_0 is true, that the statistic would take a value as extreme as or more extreme than the one actually observed is called the *P*-value of the test. The smaller the *P*-value, the stronger the evidence against H_0 provided by the data.

- Small *P*-values are evidence against H_0, because they say that the observed result is unlikely to occur when H_0 is true.
- Large *P*-values fail to give convincing evidence against H_0, because they say that the observed result could have occurred by chance if H_0 were true.

μ,

where $z = \dfrac{\overline{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$

H_a: $\mu > \mu_0$

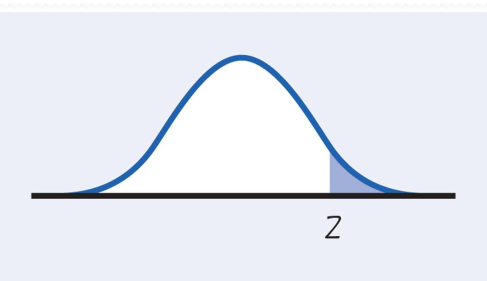
- *P*-value is the probability of getting a value as large or larger than the observed z-score.

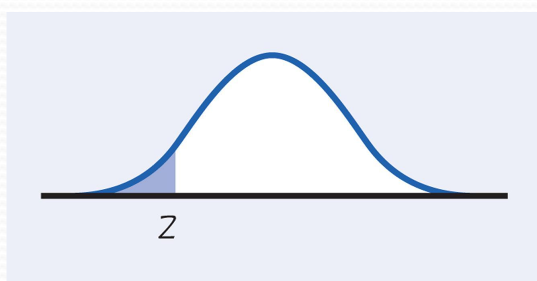
H_a: $\mu < \mu_0$

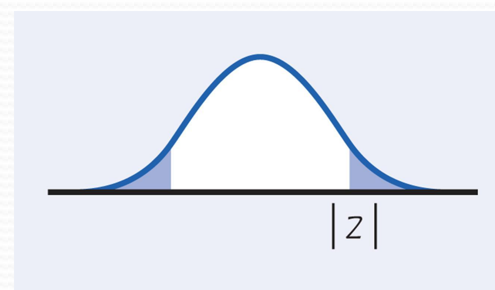
- *P*-value is the probability of getting a value as small or smaller than the observed z-score.

H_a: $\mu \neq \mu_0$

- *P*-value is *two times* the probability of getting a value as large or larger than the absolute value of the observed z-score.







Case Study

Sweetening Colas

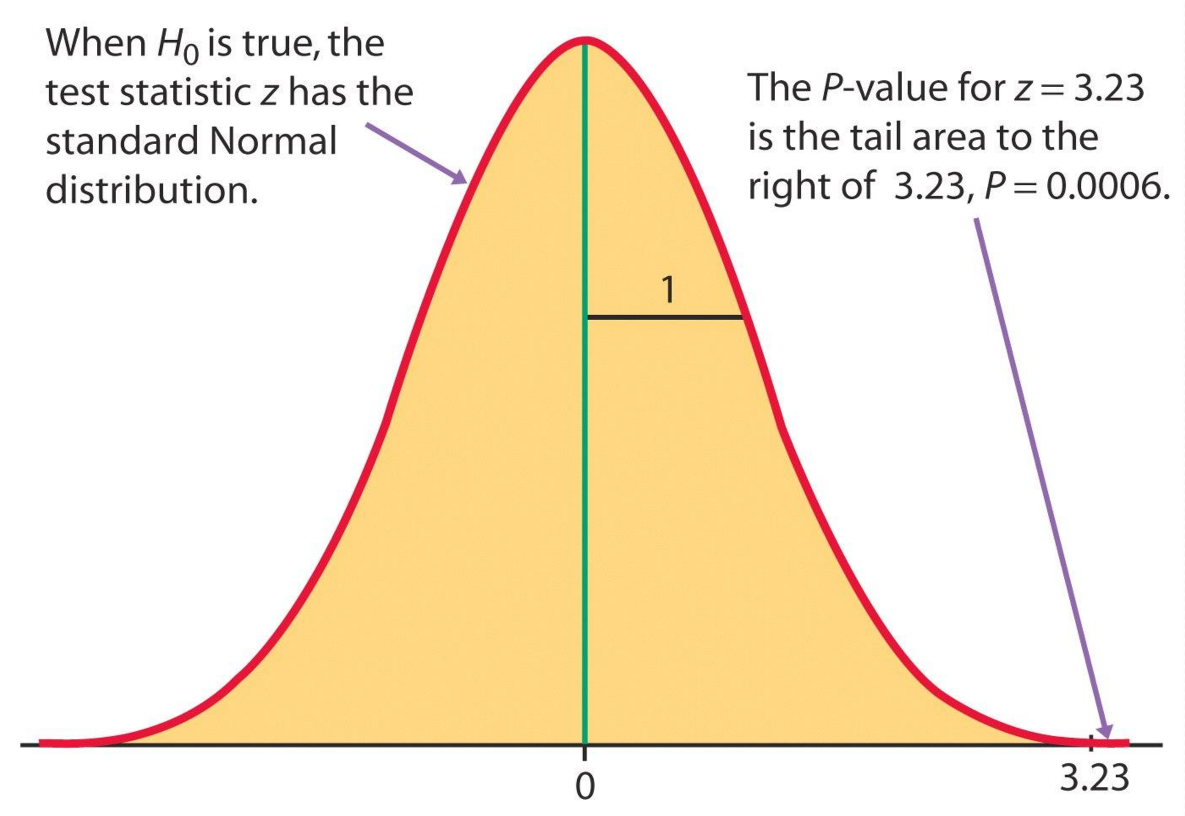
For test statistic $z = 3.23$ and alternative hypothesis H_a: $\mu > 0$, the P-value would be:

$$P\text{-value} = P(Z > 3.23) = 1 - 0.9994 = 0.0006$$

If H_0 is true, there is only a 0.0006 (0.06%) chance that we would see results at least as extreme as those in the sample; thus, since we saw results that are unlikely if H_0 is true, we therefore have evidence against H_0 and in favor of H_a.

Case Study

Sweetening Colas



Case Study II

Studying Job Satisfaction

Suppose job satisfaction scores follow a Normal distribution with standard deviation $\sigma = 60$. Data from 18 workers gave a sample mean score of 17. If the null hypothesis of no average difference in job satisfaction is true, the test statistic would be:

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{17 - 0}{60 / \sqrt{18}} \approx 1.20$$

Case Study II

Studying Job Satisfaction

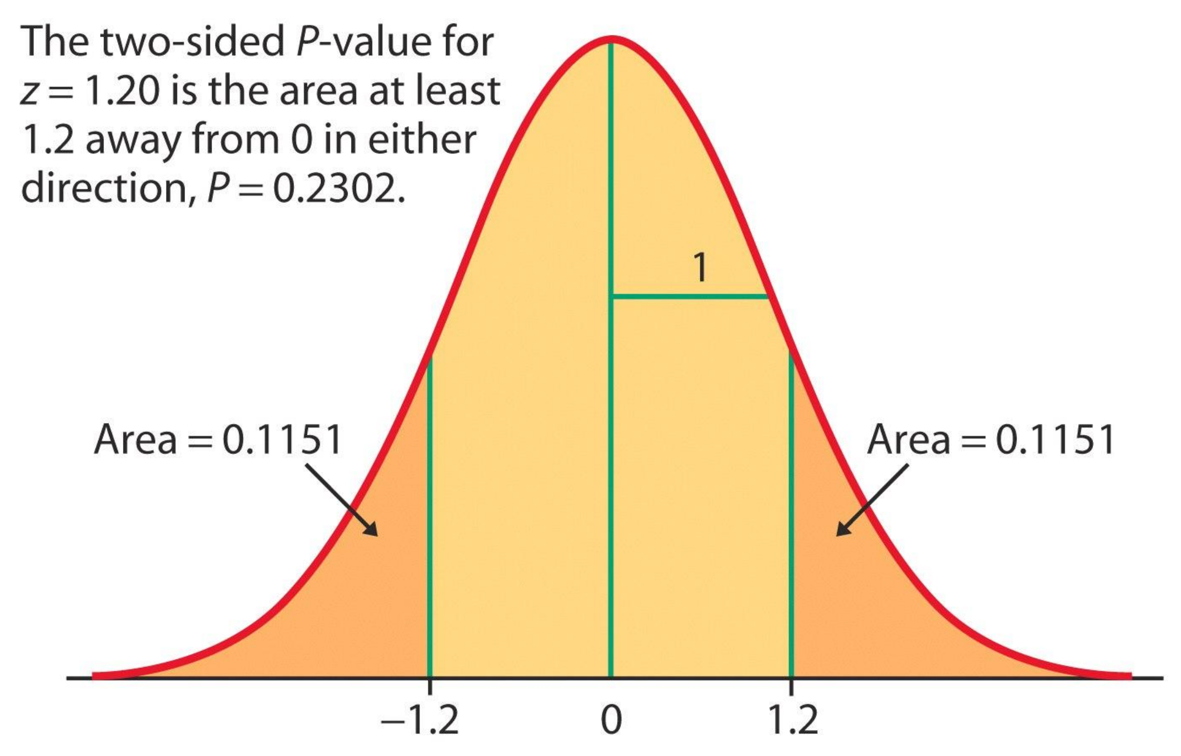
For test statistic $z = 1.20$ and alternative hypothesis H_a: $\mu \neq 0$, the *P*-value would be:

$$
\begin{aligned}
P\text{-value} &= P(Z < -1.20 \text{ or } Z > 1.20) \\
&= 2\,P(Z < -1.20) = 2\,P(Z > 1.20) \\
&= (2)(0.1151) = 0.2302
\end{aligned}
$$

If H_0 is true, there is a 0.2302 (23.02%) chance that we would see results at least as extreme as those in the sample; thus, since we saw results that are likely if H_0 is true, we therefore do not have good evidence against H_0 and in favor of H_a.

Case Study II

Studying Job Satisfaction



P-value and statistical significance (part II)

- Tests of significance assess the evidence against H_0. If the evidence is strong, we can confidently reject H_0 in favor of the alternative.
- Our conclusion in a significance test comes down to:

  P-value small → reject H_0 → conclude H_a (in context)

  P-value large → fail to reject H_0 → cannot conclude H_a (in context)
- There is no rule for how small a *P*-value we should require in order to reject H_0—it's a matter of judgment and depends on the specific circumstances. But we can compare the *P*-value with a fixed value that we regard as decisive, called the significance level. We write it as α, the Greek letter alpha. When our *P*-value is less than the chosen α, we say that the result is statistically significant.
- If the *P*-value is smaller than alpha, we say that the data are statistically significant at level α. The quantity α is called the significance level or the level of significance.

Exercises and Examples (**Finding** P−**value**)

Example For a large-sample z test of $H_0 : \mu = 0$ versus $H_a : \mu \neq 0$, $H_a : \mu > 0$, or $H_a : \mu < 0$, find the P-value associated with $\overline{x} = 0.3$,$\sigma = 1$ and $n = 10$ respectively.

- **Exercise 1**

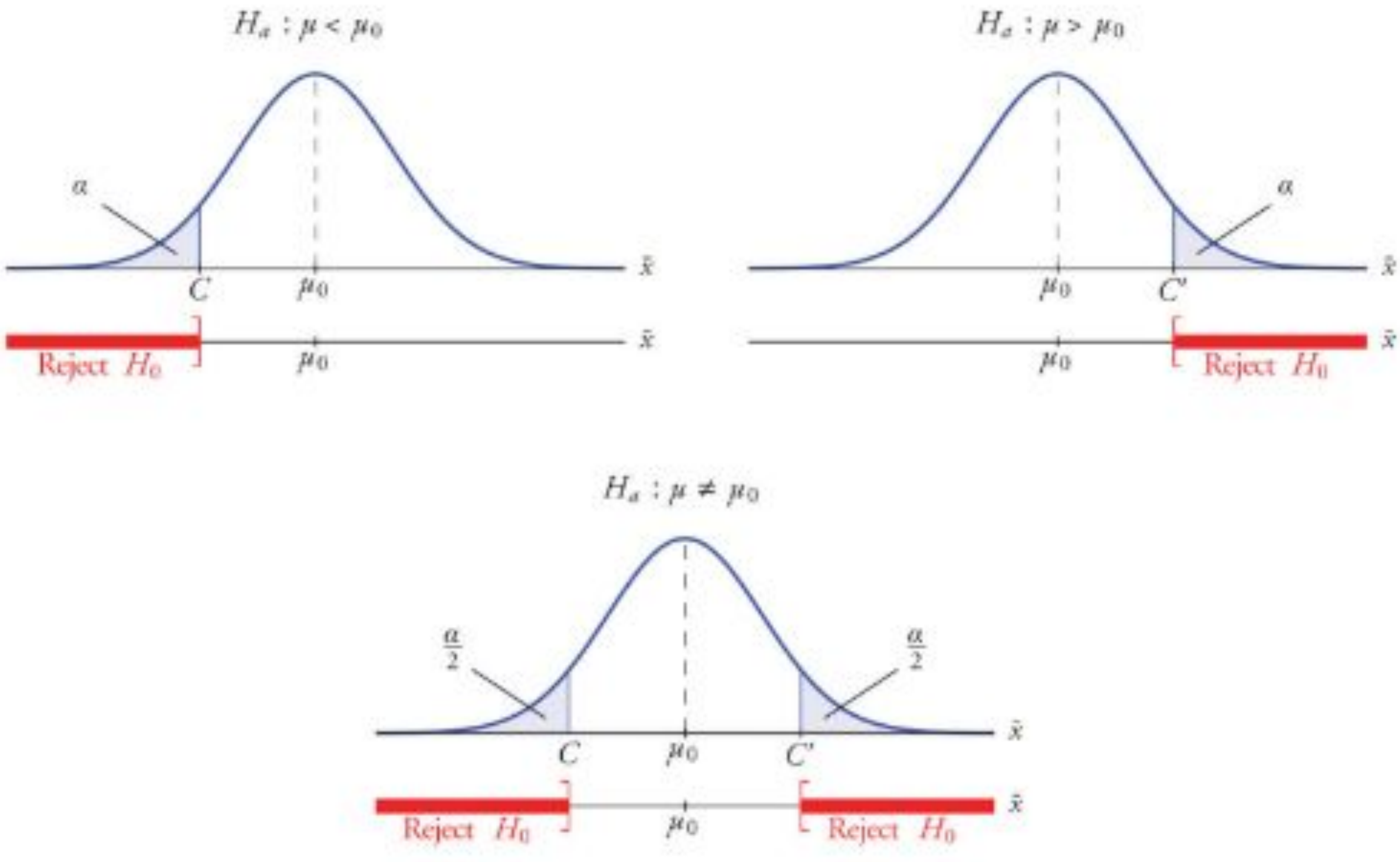
Let μ denote the mean reaction time to a certain stimulus. For a large-sample z test of H_0: $\mu = 5$ versus H_a: $\mu > 5$, find the P-value associated with each of the given values of the z test statistic.

a. 1.42 **b.** .90 **c.** 1.96 **d.** 2.48 **e.** $-.11$

Newly purchased tires of a certain type are supposed to be filled to a pressure of 30 lb/in^2. Let μ denote the true average pressure. Find the P-value associated with each given z statistic value for testing H_0: $\mu = 30$ versus H_a: $\mu \neq 30$.

a. 2.10 **b.** -1.75 **c.** $-.55$ **d.** 1.41 **e.** -5.3

α, rejection region, and critical value C = −z∗ ,or C′ = z∗ in the graphs below



The number α is the total area of a tail or a pair of tails.

Note: C = −z∗ ,or C′ = z∗ can be gained in the tables according to α.

z TEST FOR A POPULATION MEAN

Draw an SRS of size n from a Normal population that has unknown mean μ and known standard deviation σ. To test the null hypothesis that μ has a specified value,

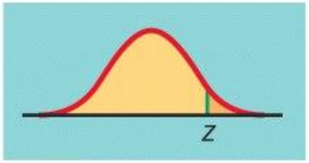
$$H_0\colon \mu = \mu_0$$

calculate the **one-sample z statistic**

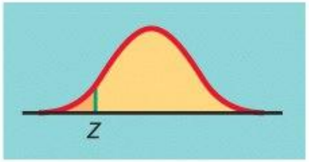
$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

In terms of a variable Z having the standard Normal distribution, the P-value for a test of H_0 against

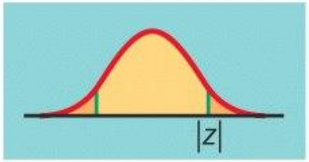
$H_a\colon \mu > \mu_0$ is $P(Z \geq z)$



$H_a\colon \mu < \mu_0$ is $P(Z \leq z)$



$H_a\colon \mu \neq \mu_0$ is $2P(Z \geq |z|)$



Tests of significance

THE FOUR-STEP PROCESS

- **STATE**: What is the practical question that requires a statistical test?
- **PLAN**: Identify the parameter, state null and alternative hypotheses, and choose the type of test that fits your situation.
- **SOLVE**: Carry out the test in three phases:

  1. Check the conditions for the test you plan to use.

  2. Calculate the test statistic.

  3. Find the *P*-value.

- **CONCLUDE**: Return to the practical question to describe your results in this setting.

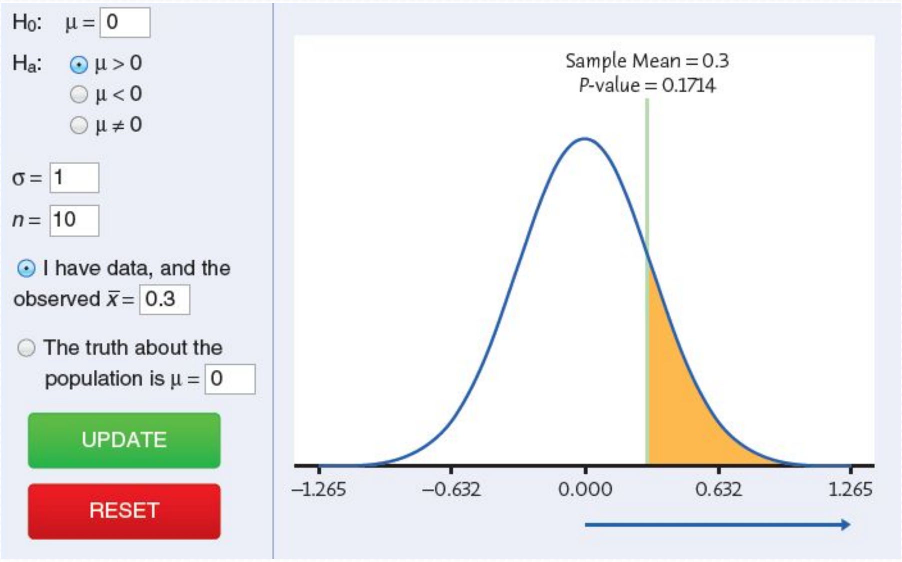
Sweetening colas—one-sided *P*-value

- **Cola sweeteners**

STATE: In a matched-pairs study of 10 trained tasters, the researchers found that $\bar{x} = 0.3,$ with values of $\bar{x} > 0$ favoring H_a over H_0. Is this strong enough evidence to decide that the sweetener is losing sweetness?

PLAN: Take μ to be the mean sweetness loss rating difference. We want to test the hypotheses:

$$H_0: \mu = 0$$
$$H_a: \mu > 0$$



SOLVE: The illustration shows the calculation of the P-value; the $P(\bar{x} \geq 0.3) = 0.1714$ if we assume $\mu = 0$, in other words, the probability of a value of the test statistic at least as extreme, if the null hypothesis were true.

CONCLUDE: More than 17% of the time, an SRS of size 10 of trained testers would have a mean sweetness difference at least as great as that observed. The observed $\bar{x} = 0.3$ is therefore not good evidence that this cola experienced loss of sweetness.

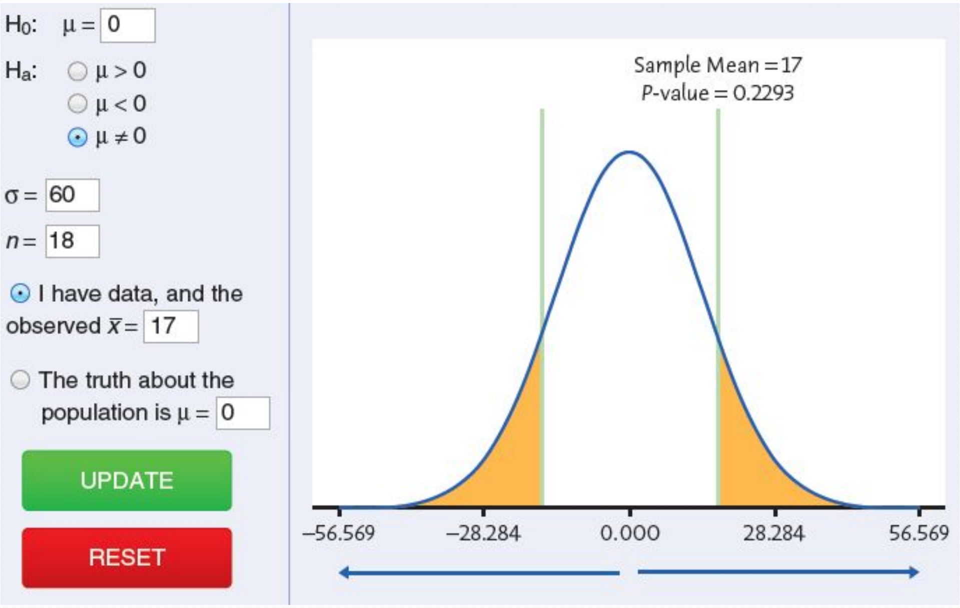
Job satisfaction—two-sided *P*-value

- **Job satisfaction**

STATE: In a matched-pairs study of 18 workers, the analysis found: $\bar{x} = 17$, with values of $\bar{x} > 0$ favoring H_a over H_0; that is, workers preferring the self-paced over the machine-paced environment.

PLAN: Take μ to be the mean difference for all workers; because the researchers didn't know which environment the workers would prefer, we have to test the hypotheses:

$$H_0: \mu = 0$$
$$H_a: \mu \neq 0$$



SOLVE: The illustration shows the calculation of the P-value; the $P(|\bar{x}| \geq 17) = 0.2293$ if we assume $\mu = 0$, in words, the probability of a value of the test statistic at least as extreme, if the null hypothesis were true.

CONCLUDE: More than 22% of the time, an SRS of size 18 of workers would have a mean difference at least as great as that observed. The observed $\bar{x} = 17$ is therefore not good evidence that workers preferring the self-paced over the machine-paced environment or other way around.

step 1 (**H**)	Make Hypothesis	$H_0 : \mu = \mu_0$ $H_a : \mu \neq \mu_0,$ $\mu > \mu_0,$ $\mu < \mu_0$	$H_0 : \mu = \mu_0$ $H_a : \mu \neq \mu_0,$ $\mu > \mu_0,$ $\mu < \mu_0$	$H_0 : \mu_1 = \mu_2$ $H_a : \mu_1 \neq \mu_2,$ $\mu_1 > \mu_2,$ $\mu_1 < \mu_2$	$H_0 : p = p_0$ $H_a : p \neq p_0,$ $p > p_0,$ $p < p_0$	$H_0 : p_1 = p_2$ $H_a : p_1 \neq p_2,$ $p_1 > p_2,$ $p_1 < p_2$
step 2 (**A**)	Critical value of z or t by using tables and α	z-table (z^*)	t-table (t^*)	t-table (t^*)	z-table (z^*)	z-table (z^*)
step 3:(T) The hard-est step	Calculate test statistics or P-value	$z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$ **(P–value)** $2P(Z \geq \lvert z \rvert),$ $P(Z \geq z),$ $P(Z \leq z)$ respectively	$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$ **(P–value)** $2P(T \geq \lvert t \rvert),$ $P(T \geq t),$ $P(T \leq t)$ respectively	$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ **(P–value)** $2P(T \geq \lvert t \rvert),$ $P(T \geq t),$ $P(T \leq t)$ respectively	$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$ **(P–value)** $2P(Z \geq \lvert z \rvert),$ $P(Z \geq z),$ $P(Z \leq z)$ respectively	$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})(\frac{1}{n_1} + \frac{1}{n_2})}}$ **(P–value)** $2P(Z \geq \lvert z \rvert),$ $P(Z \geq z),$ $P(Z \leq z)$ respectively
step 4(**C**): Con-clu-sion	Reject H_0 based on Method1) c.v vs. t.s (z^* (or t^*) is positive.) Method2) α vs. P-value	M1) $\lvert z \rvert \geq z^*,$ $z \geq z^*,$ $z \leq -z^*$ respectively ——— M2) $\alpha \geq P-$value	M1) $\lvert t \rvert \geq t^*,$ $t \geq t^*,$ $t \leq -t^*$ respectively ——— M2) $\alpha \geq P-$value	M1) $\lvert t \rvert \geq t^*,$ $t \geq t^*,$ $t \leq -t^*$ respectively ——— M2) $\alpha \geq P-$value	M1) $\lvert z \rvert \geq z^*,$ $z \geq z^*,$ $z \leq -z^*$ respectively ——— M2) $\alpha \geq P-$value	M1) $\lvert z \rvert \geq z^*,$ $z \geq z^*,$ $z \leq -z^*$ respectively ——— M2) $\alpha \geq P-$value

Four steps: "HAT C"

1. **Step 1 (H): Make Hypotheses**
   - H_0: $\mu = \mu_0$ and H_a: $\mu > \mu_0$, $\mu < \mu_0$, or $\mu \neq \mu_0$
2. **Step 2 (A):** α ⇒ **Critical value of z=**z∗ **and Rejection region**
   - z∗ and the rejection region can be gained in z-table according to α. But when n is very large, we can use t-table (Table C) to get z∗.
3. **Step 3 (T): Calculate Test statistic or** P−**value**
   - See for the previous slides how to calculate p-value

4. **Step 4 (C): Conclusion to reject** H_0 **or not.**
   H_0 is rejected when the test statistic is in the rejection region.
   - Method 1) Comparing critical value with test statistic
     We reject H_0 when $|z| \geq z^*$, $z \geq z^*$,and $z \leq -z^*$ respectively where z^* is positive.
   - Method 2) Comparing α with P −value
     We reject H_0 when α ≥ P −value

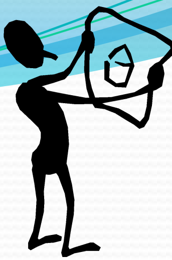
Significance from a table

- Statistics in practice uses technology to get *P*-values quickly and accurately. In the absence of suitable technology, you can get approximate *P*-values by comparing your test statistic with critical values from a table.

SIGNIFICANCE FROM A TABLE OF CRITICAL VALUES

- To find the approximate *P*-value for any z statistic, compare z (ignoring its sign) with the critical values z^* at the bottom of Table C. If z falls between two values of z^*, the *P*-value falls between the two corresponding values of *P* in the "One-sided *P*" or the "Two-sided *P*" row of Table C.

Case Study I



Sweetening Colas

1. ***Hypotheses:*** $H_0: \mu = 0$
   $H_a: \mu > 0$
2. ***Test Statistic:***
   $$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{1.02 - 0}{1 / \sqrt{10}} \approx 3.23$$
3. ***P-value:*** P-value = $P(Z > 3.23) = 1 - 0.9994 = 0.0006$
4. ***CONCLUSION:***

   Since the P-value is smaller than $\alpha = 0.01$, **there is very strong evidence that the new cola loses sweetness on average during storage at room temperature.**

Case Study II

Studying Job Satisfaction

1. ***Hypotheses:*** H_0: $\mu = 0$
   H_a: $\mu \neq 0$
2. ***Test Statistic:***
$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{17 - 0}{60 / \sqrt{18}} \approx 1.20$$
3. ***P-value:*** P-value $= 2P(Z > 1.20) = (2)(1 - 0.8849) = 0.2302$
4. ***CONCLUSION:***

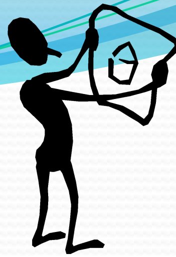
   Since the P-value is larger than $\alpha = 0.10$, **there is not sufficient evidence that mean job satisfaction of assembly workers differs when their work is machine-paced rather than self-paced.**

Confidence Intervals & Two-Sided Tests

A level α two-sided significance test rejects the null hypothesis H_0: $\mu = \mu_0$ exactly when the value μ_0 falls outside a level $(1 - \alpha)$ confidence interval for μ.

Case Study II

Studying Job Satisfaction



A 90% confidence interval for μ is:

$$\bar{x} \pm z^* \frac{\sigma}{\sqrt{n}} = 17 \pm 1.645 \frac{60}{\sqrt{18}} = 17 \pm 23.26$$

$$= -6.26 \text{ to } 40.26$$

Since $\mu_0 = 0$ is in this confidence interval, it is plausible that the true value of μ is 0; thus, there is not sufficient evidence (at $\alpha = 0.10$) that the mean job satisfaction of assembly workers differs when their work is machine-paced rather than self-paced.

Examples and exercises(z testing for μ)

Significance from a table. A test of H_0: $\mu = 0$ against H_a: $\mu > 0$ has test statistic $z = 1.876$. Is this test significant at the 5% level ($\alpha = 0.05$)? Is it significant at the 1% level ($\alpha = 0.01$)?

Examples and exercises(z testing for μ)

Significance from a table. A test of H_0: $\mu = 0$ against H_a: $\mu \neq 0$ has test statistic $z = 1.876$. Is this test significant at the 5% level ($\alpha = 0.05$)? Is it significant at the 1% level ($\alpha = 0.01$)?

Examples and exercises(z testing for μ)

Testing a random number generator. A random number generator is supposed to produce random numbers that are uniformly distributed on the interval from 0 to 1. If this is true, the numbers generated come from a population with $\mu = 0.5$ and $\sigma = 0.2887$. A command to generate 100 random numbers gives outcomes with mean $\bar{x} = 0.4365$. Assume that the population σ remains fixed. We want to test

$$H_0: \mu = 0.5$$
$$H_a: \mu \neq 0.5$$

(a) Calculate the value of the z test statistic.

(b) Use Table C: is z significant at the 5% level ($\alpha = 0.05$)?

(c) Use Table C: is z significant at the 1% level ($\alpha = 0.01$)?

(d) Between which two Normal critical values z^* in the bottom row of Table C does z lie? Between what two numbers does the P-value lie? Does the test give good evidence against the null hypothesis?

Comprehensive Example

STATE: To investigate water quality, on August 8, 2010, the *Columbus Dispatch* took water samples at 20 Ohio State Park swimming areas. Those samples were taken to laboratories and tested for fecal coliform, which are bacteria found in human and animal feces. An unsafe level of fecal coliform means there's a higher chance that disease-causing bacteria are present and more risk that a swimmer will become ill. Ohio considers it unsafe if a 100-milliliter sample (about 3.3 ounces) of water contains more than 400 coliform bacteria. Here are the fecal coliform levels found by the laboratories:[6]

160	40	2800	80	2000	2000	1500	400	150	500
3000	2200	15	80	2000	2000	2600	600	1000	1500

Are these data good evidence that, on average, the fecal coliform levels in these swimming areas were unsafe?

Assume that the population distribution of fecal coliform levels is normal. $\sigma = 1070$ is given. Use $\alpha = 0.01$ as a level of significance.

STEP 1 (**H**)) $H_0 : \mu = 400$ vs. $H_a : \mu > 400$.

STEP 2 (**A**)) Critical value of z (z^*)= 2.325.

Draw the rejection region.

STEP 3 (**T**))

z statistic: $z = \frac{\overline{x}-\mu}{\frac{\sigma}{\sqrt{n}}} = 3.47$

$\overline{x} = \frac{160+40+\cdots+1500}{20} = 1231,\ \mu = 400$

$\sigma = 1070,\ n = 20$

STEP 4 (**C**))

- Method 1) z-statistic > $z*$.
  It means that z-statistic is in the rejection region.
- Method 2) P−value=0.0003 by using Table A. So P−value < α = 0.01.
- Hence we reject H0. We conclude that there is sufficient evidence to support the alternative hypothesis that , on average, fecal coliform levels in these Ohio State Park swimming areas are unsafe.

Comprehensive Exercise 1

Glycerol is a major by-product of ethanol fermentation in wine production and contributes to the sweetness, body, and fullness of wines. The article "A Rapid and Simple Method for Simultaneous Determination of Glycerol, Fructose, and Glucose in Wine" (American J. of Enology and Viticulture, 2007: 279–283) includes the following observations on glycerol concentration (mg/mL) for samples of standard-quality (uncertified) white wines: 2.67, 4.62, 4.14, 3.81, 3.83 assuming that the population distribution of glycerol concentration is normal. Suppose the desired concentration value is 4. Does the sample data suggest that true average concentration is something other than the desired value at $\alpha = 0.05$? $\sigma = 0.75$ is given.

z Procedures

- If we know the standard deviation σ of the population, a confidence interval for the mean μ is:

$$\bar{x} \pm z^* \frac{\sigma}{\sqrt{n}}$$

- To test a hypothesis $H_o: \mu = \mu_o$ we use the one-sample z statistic:

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

- These are called **z procedures** because they both involve a one-sample *z*-score and use the standard Normal distribution.

P-Value, Statistical Significance (1 of 12)

The probability (computed assuming that H_o is true) that the test statistic would take a value as extreme or more extreme than that actually observed is called the ___________ of the test.

a) *P*-value
b) confidence interval
c) conditional probability
d) null hypothesis

P-Value, Statistical Significance (1 of 12) (answer)

The probability (computed assuming that H_o is true) that the test statistic would take a value as extreme or more extreme than that actually observed is called the ___________ of the test.

a) ***P*-value**
b) confidence interval
c) conditional probability
d) null hypothesis

The correct answer is A.

P-Value, Statistical Significance (2 of 12)

- A consumer advocate evaluates the claim that a new granola cereal contains “4 ounces of cashews in every bag.” She recognizes that the amount of cashews will vary slightly from bag to bag but suspects that the mean amount of cashews per bag is actually less than 4 ounces. The advocate purchases a random sample of 40 bags of cereal and calculates $\bar{x}$ to be 3.68 ounces. What is the definition of the *P*-value applied to this problem?

a) the probability of obtaining an $\bar{x}$ as large as 3.68 ounces
b) the probability, if the mean for all bags was 3.68 ounces, of obtaining an $\bar{x}$ equal to 3.68
c) the probability, if the mean for all bags was 4 ounces, of obtaining an $\bar{x}$ of 3.68 or smaller
d) the probability, if the mean for all bags was 4 ounces, of obtaining an $\bar{x}$ equal to 3.68

P-Value, Statistical Significance (2 of 12) (answer)

- A consumer advocate evaluates the claim that a new granola cereal contains "4 ounces of cashews in every bag." She recognizes that the amount of cashews will vary slightly from bag to bag but suspects that the mean amount of cashews per bag is actually less than 4 ounces. The advocate purchases a random sample of 40 bags of cereal and calculates $\bar{x}$ to be 3.68 ounces. What is the definition of the *P*-value applied to this problem?

a) the probability of obtaining an $\bar{x}$ as large as 3.68 ounces

b) the probability, if the mean for all bags was 3.68 ounces, of obtaining an $\bar{x}$ equal to 3.68

c) the probability, if the mean for all bags was 4 ounces, of obtaining an $\bar{x}$ of 3.68 or smaller

d) the probability, if the mean for all bags was 4 ounces, of obtaining an $\bar{x}$ equal to 3.68

The correct answer is C.

P-Value, Statistical Significance (3 of 12)

- To calculate the *P*-value for a significance test, we use information about the:

a) sample distribution.

b) population distribution.

c) sampling distribution of $\bar{x}$.

P-Value, Statistical Significance (3 of 12) (answer)

- To calculate the *P*-value for a significance test, we use information about the:

a) sample distribution.
b) population distribution.
c) **sampling distribution of $\bar{x}$.**

The correct answer is C.

P-Value, Statistical Significance (4 of 12)

There is evidence against the null hypothesis whenever the *P*-value is:

a) small.
b) large.
c) 0.5.
d) 1.

P-Value, Statistical Significance (4 of 12) (answer)

There is evidence against the null hypothesis whenever the *P*-value is:

a) **small.**
b) large.
c) 0.5.
d) 1.

The correct answer is A.

P-Value, Statistical Significance (5 of 12)

The data are statistically significant whenever the P-value:

a) $> \alpha$.

b) $\leq \alpha$.

c) $\neq \alpha$.

d) $\neq \mu$.

P-Value, Statistical Significance (5 of 12) (answer)

The data are statistically significant whenever the P-value:

a) $> \alpha$.

b) $\leq \alpha$.

c) $\neq \alpha$.

d) $\neq \mu$.

The correct answer is B.

P-Value, Statistical Significance (6 of 12)

- A consumer advocate evaluates the claim that a new granola cereal contains "4 ounces of cashews in every bag." She recognizes that the amount of cashews will vary slightly from bag to bag but suspects that the mean amount of cashews per bag is actually less than 4 ounces. The advocate purchases a random sample of 40 bags of cereal and calculates $\bar{x}$ to be 3.68 ounces. Suppose the *P*-value is calculated to be 0.098 and α is set at 0.05. Her result is:

a) statistically significant.

b) not statistically significant.

P-Value, Statistical Significance (6 of 12) (answer)

- A consumer advocate evaluates the claim that a new granola cereal contains "4 ounces of cashews in every bag." She recognizes that the amount of cashews will vary slightly from bag to bag but suspects that the mean amount of cashews per bag is actually less than 4 ounces. The advocate purchases a random sample of 40 bags of cereal and calculates $\bar{x}$ to be 3.68 ounces. Suppose the *P*-value is calculated to be 0.098 and α is set at 0.05. Her result is:

a) statistically significant.

b) not statistically significant.

The correct answer is B.

P-Value, Statistical Significance (7 of 12)

- Suppose the *P*-value for a hypothesis test is 0.0304. Using $\alpha = 0.05$, what is the appropriate conclusion?

a) Reject the null hypothesis.
b) Reject the alternative hypothesis.
c) Do not reject the null hypothesis.
d) Do not reject the alternative hypothesis.

P-Value, Statistical Significance (7 of 12) (answer)

- Suppose the *P*-value for a hypothesis test is 0.0304. Using $\alpha = 0.05$, what is the appropriate conclusion?

a) **Reject the null hypothesis.**
b) Reject the alternative hypothesis.
c) Do not reject the null hypothesis.
d) Do not reject the alternative hypothesis.

The correct answer is A.

P-Value, Statistical Significance (8 of 12)

- Suppose the *P*-value for a hypothesis test is 0.304. Using $\alpha = 0.05$, what is the appropriate conclusion?

a) Reject the null hypothesis.
b) Reject the alternative hypothesis.
c) Do not reject the null hypothesis.
d) Do not reject the alternative hypothesis.

P-Value, Statistical Significance (8 of 12) (answer)

- Suppose the *P*-value for a hypothesis test is 0.304. Using $\alpha = 0.05$, what is the appropriate conclusion?

a) Reject the null hypothesis.
b) Reject the alternative hypothesis.
c) Do not reject the null hypothesis.
d) Do not reject the alternative hypothesis.

The correct answer is C.

P-Value, Statistical Significance (9 of 12)

The significance level is denoted by:

a) μ.

b) σ .

c) α.

d) the *P*-value.

P-Value, Statistical Significance (9 of 12) (answer)

The significance level is denoted by:

a) μ.

b) σ.

***c)* α.**

d) the P-value.

The correct answer is C.

P-Value, Statistical Significance (10 of 12)

"Significant" in the statistical sense does not mean "important"; it means simply "not likely to happen just by chance."

a) true

b) false

P-Value, Statistical Significance (10 of 12) (answer)

"Significant" in the statistical sense does not mean "important"; it means simply "not likely to happen just by chance."

a) **true**
b) false

The correct answer is A.

P-Value, Statistical Significance (11 of 12)

Suppose a significance test is being conducted using a significance level of 0.10. If a student calculates a *P*-value of 1.9, the student:

a) should reject the null hypothesis.
b) should not reject the null hypothesis.
c) made a mistake in calculating the *P*-value.

P-Value, Statistical Significance (11 of 12) (answer)

Suppose a significance test is being conducted using a significance level of 0.10. If a student calculates a *P*-value of 1.9, the student:

a) should reject the null hypothesis.
b) should not reject the null hypothesis.
c) **made a mistake in calculating the *P*-value.**

The correct answer is C.

P-Value, Statistical Significance (12 of 12)

A small *P*-value means that there is a strong association between variables under investigation.

a) true
b) false

P-Value, Statistical Significance (12 of 12) (answer)

A small *P*-value means that there is a strong association between variables under investigation.

a) true
b) false

The correct answer is B.

Tests for a Population Mean (1 of 3)

- Which of the following null hypotheses is tested by the statistic
- $z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$?

$a)\ H_0\!:\ \bar{x} = \mu_0$

$b)\ H_0\!:\ \bar{x} = \mu$

$c)\ H_0\!:\ \bar{x} = 0$

$d)\ H_0\!:\ \mu = \mu_0$

$e)\ H_0\!:\ \mu = 0$

Tests for a Population Mean (1 of 3) (answer)

- Which of the following null hypotheses is tested by the statistic
- $z = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$?

a) $H_0\text{:}\ \bar{x} = \mu_0$

b) $H_0\text{:}\ \bar{x} = \mu$

c) $H_0\text{:}\ \bar{x} = 0$

d) $H_0\text{:}\ \mu = \mu_0$

e) $H_0\text{:}\ \mu = 0$

The correct answer is D.

Tests for a Population Mean (2 of 3)

- If the null hypothesis is true, what is the distribution of the test statistic
- $z = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$?

a) Normal with mean μ_0 and standard deviation 1
b) Normal with mean 0 and standard deviation 1
c) Normal with mean 1 and standard deviation 0
d) Normal with mean μ_0 and standard deviation σ

Tests for a Population Mean (2 of 3) (answer)

- If the null hypothesis is true, what is the distribution of the test statistic
- $z = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$?

a) Normal with mean μ_0 and standard deviation 1

b) Normal with mean 0 and standard deviation 1

c) Normal with mean 1 and standard deviation 0

d) Normal with mean μ_0 and standard deviation σ

The correct answer is B.

Tests for a Population Mean (3 of 3)

- A company wants to see if its bottling machine is filling correctly. The bottles should be filled to 20 ounces. It is known the process has a standard deviation of 0.5 ounces. A sample of $n = 25$ is taken and the sample mean is found to be 19.7. The correct test statistic is:

$$a)\ z = \frac{19.7 - 20}{0.5/5}$$

$$b)\ z = \frac{20 - 19.7}{0.5/5}$$

$$c)\ z = \frac{19.7 - 20}{0.5/25}$$

$$d)\ z = \frac{20 - 19.7}{0.5/25}$$

Tests for a Population Mean (3 of 3) (answer)

- A company wants to see if its bottling machine is filling correctly. The bottles should be filled to 20 ounces. It is known the process has a standard deviation of 0.5 ounces. A sample of $n = 25$ is taken and the sample mean is found to be 19.7. The correct test statistic is:

a) $z = \dfrac{19.7 - 20}{0.5/5}$

b) $z = \dfrac{20 - 19.7}{0.5/5}$

c) $z = \dfrac{19.7 - 20}{0.5/25}$

d) $z = \dfrac{20 - 19.7}{0.5/25}$

The correct answer is A.