# 8 Tests of Hypotheses Based on a Single Sample

## 8.1 Hypotheses and Test Procedures

**Example:** *Legal setting.* Test the claim that Ralph committed armed robbery.

$H_0$: null hypothesis, status quo, conventional wisdom, old idea, accepted idea.

$H_a$: alternative hypothesis, the challenge to the conventional wisdom, new idea, proposed idea.

If we wish to reject $H_0$ (i.e., reject the idea that Ralph is innocent of armed robbery) in favor of $H_a$ (i.e., in favor of the idea that Ralph is guilty of armed robbery), we need overwhelming evidence to support our claim, such as witnesses, videotapes, confession, or DNA evidence.

Otherwise, we fail to prove him guilty (i.e., not guilty).

Innocent until proved guilty.

We **reject** the null hypothesis in favor of the alternative hypothesis, or we **fail to reject** the null hypothesis.

Do **NOT** say "accept $H_0$" (which is equivalent to saying "proved innocent"), as a substitution for "**fail to reject** $H_0$."

What is the goal in hypothesis testing?

Regarding the goal of hypothesis testing, the **researcher** is analogous to whom in the legal setting?

□

**Example:** State the hypotheses for testing the claim that peanut oil **causes** colon cancer.

□

**Example:** State the hypotheses for testing the claim that peanut oil **prevents** colon cancer.

□

## Statistical setting

**Example:** Suppose a particular politician's approval rating last month was 55%. You believe that this approval rating has decreased, due to a scandal. State the appropriate hypotheses. Let $p$ be the *unknown* current population approval rating of this politician.

□

**Example:** Suppose the mean personal income of your community last year was \$41,000. You believe that mean personal income has increased, due to improved infrastructure. State the appropriate hypotheses. Let $\mu$ be the *unknown* population mean personal income this year.

□

**Example:** In a college's handbook, the mean SAT score is listed as 1100. You believe that the information is outdated. State the appropriate hypotheses. Let $\mu$ be the *unknown* population mean SAT score.

**Definition:** A **Type I error** occurs if we reject $H_0$, when $H_0$ is true.

Define $\alpha = P(\text{Type I error}) = P(\text{We reject } H_0 | H_0 \text{ is true})$

**Definition:** A **Type II error** occurs if we fail to reject $H_0$, when $H_a$ is true.

Define $\beta = P(\text{Type II error}) = P(\text{We fail to reject } H_0 | H_a \text{ is true})$

**Example:** Consider the hypotheses:

$H_0$ : Ralph is **innocent** of armed robbery.

$H_a$ : Ralph is **guilty** of armed robbery.

Describe the Type I error.

Describe the Type II error.

Describe $\alpha$ and $\beta$ in terms of probabilities and proportions.

□

What are the possible values of $\alpha$?

What are the possible values of $\beta$?

Do we want $\alpha$ to be *large* or *small*?

Do we want $\beta$ to be *large* or *small*?

For the *defendants-in-court* example, how can $\alpha$ be made small (near zero)?

What happens to $\beta$, as $\alpha$ gets small (near zero)?

For the *defendants-in-court* example, how can $\beta$ be made small (near zero)?

What happens to $\alpha$, as $\beta$ gets small (near zero)?

**Example:** In October 2002, just prior to the Persian Gulf War, Iraqi President Saddam Hussein released most all Iraqi prisoners and detainees. What were Hussein's values of $\alpha$ and $\beta$?

□

**Example:** State the null and alternative hypotheses, when a person is tested for a disease.

Describe the **Type I error** in regular English and in medical terminology.

Describe the **Type II error** in regular English and in medical terminology.

□

In the **statistical** setting, how can we minimize both $\alpha$ and $\beta$?

## Binomial proportions

**Example:** Suppose that the best (i.e., old) therapy for curing some form of cancer has a success rate of **30%**, in the sense that the cancer is gone in 30% of the patients in three months. We are interested in testing if a **new** therapy is better than the **old** therapy.

Let $p$ be the unknown **population** success rate of the **new** therapy.

State the null and alternative hypotheses.

Sample $n = 10$ cancer patients who will be undergoing this **new** therapy.

Let $X$ be the number of these patients who are cured.

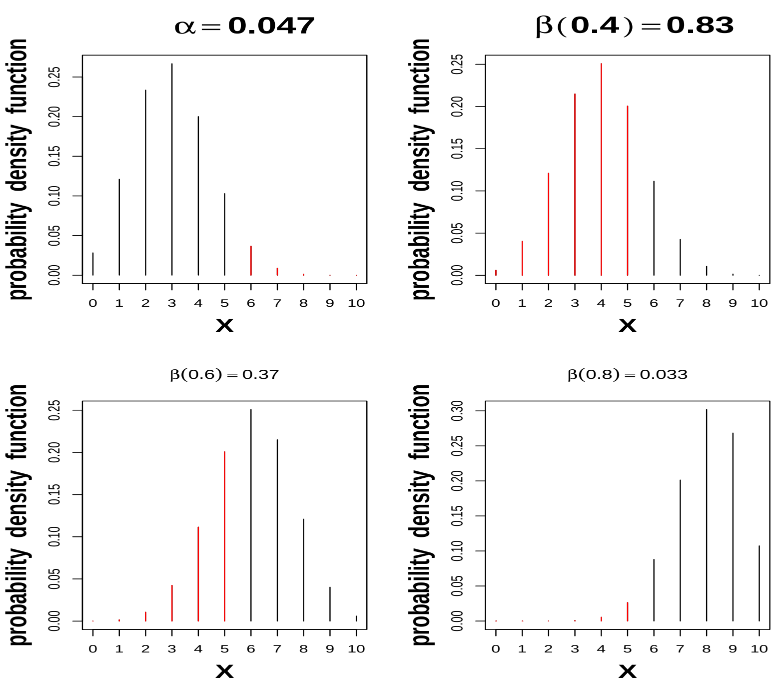
Suppose we decide to reject $H_0$ in favor of $H_a$ **iff** (i.e., if and only if) $\boldsymbol{X \geq 6}$.

$X$ is the **test statistic**.

State the **rejection region**.

Determine $\boldsymbol{\alpha}$.

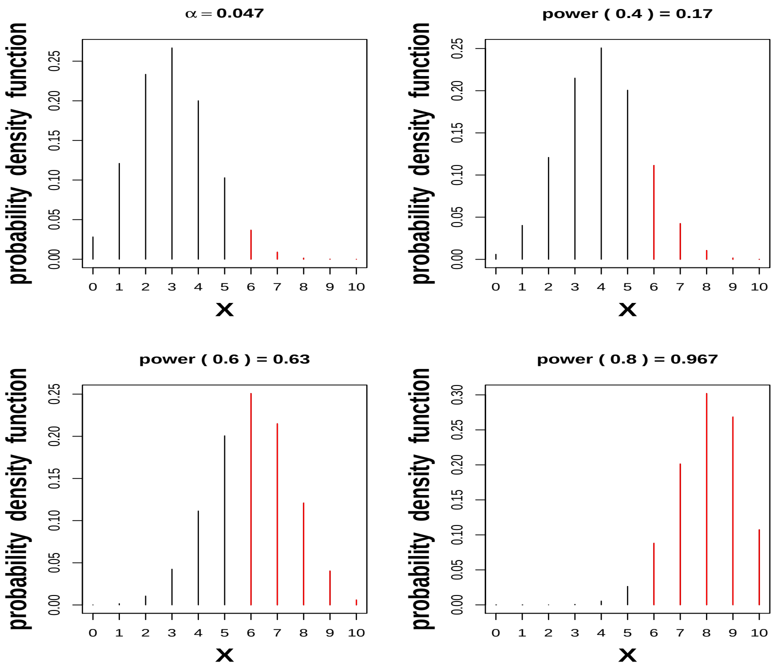
Determine $\boldsymbol{\beta}$.



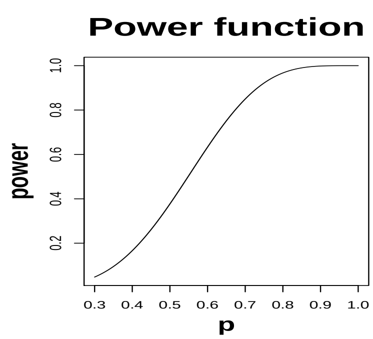
| $p$ | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1 |
|---|---|---|---|---|---|---|---|
| $\beta(p)$ | 0.83 | 0.62 | 0.37 | 0.15 | 0.033 | 0.0016 | |

Define **power** $= \mathbf{1} - \boldsymbol{\beta}$

Determine **power** for our *cancer* example.




| $p$ | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1 |
|---|---|---|---|---|---|---|---|
| power$(p) = 1 - \beta(p)$ | 0.17 | 0.38 | 0.63 | 0.85 | 0.967 | 0.9984 | |




$\square$

**Example:** Consider the same above *cancer* example.

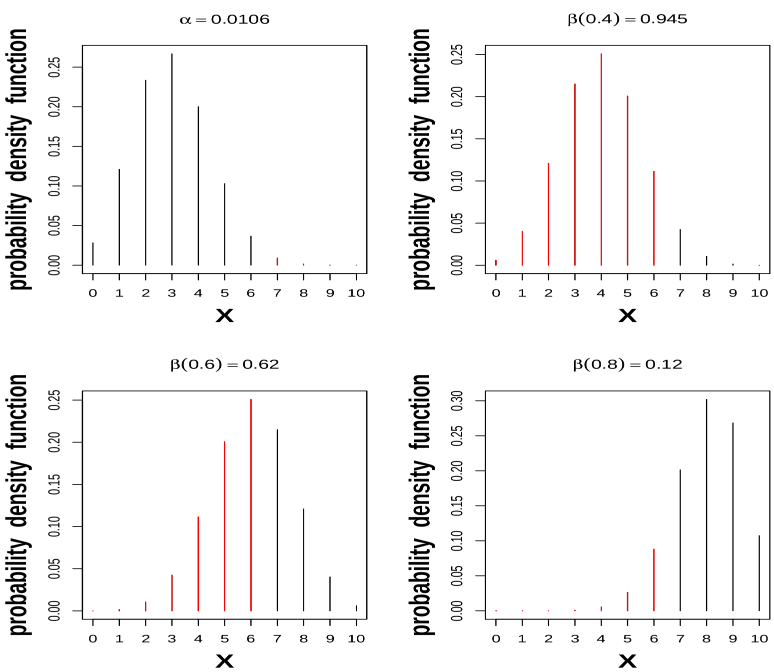
However, this time, suppose we decide to reject $H_0$ in favor of $H_a$ **iff $\boldsymbol{X \geq 7}$**.

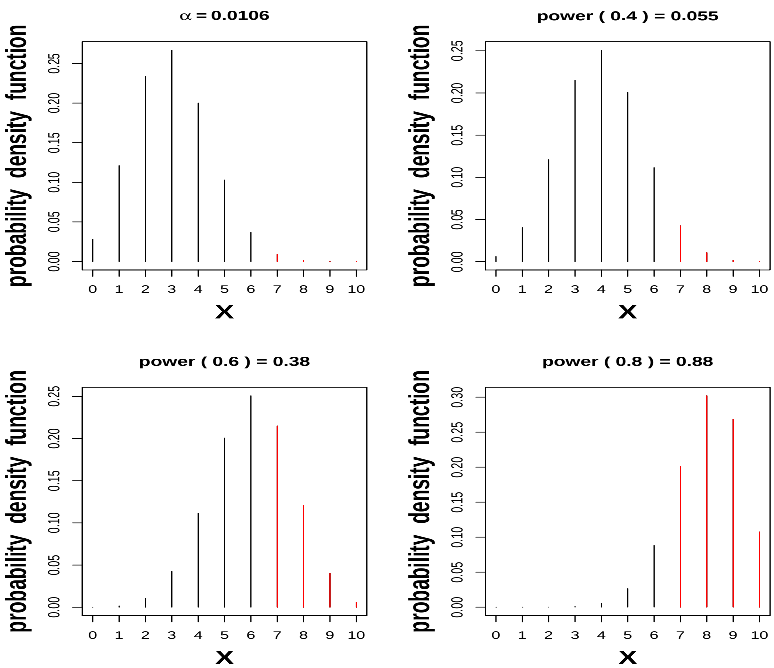
State the **rejection region**.

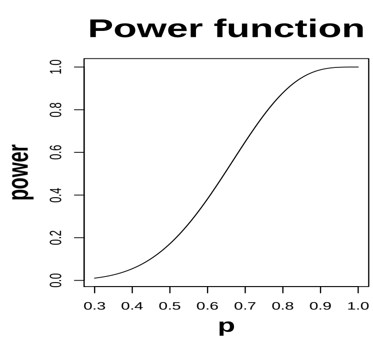
Determine $\boldsymbol{\alpha}$.

Determine $\boldsymbol{\beta}$.

| $p$ | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1 |
|---|---|---|---|---|---|---|---|
| $\beta(p)$ | 0.945 | 0.83 | 0.62 | 0.35 | 0.12 | 0.013 | |
| power$(p)$ | 0.055 | 0.17 | 0.38 | 0.65 | 0.88 | 0.987 | |





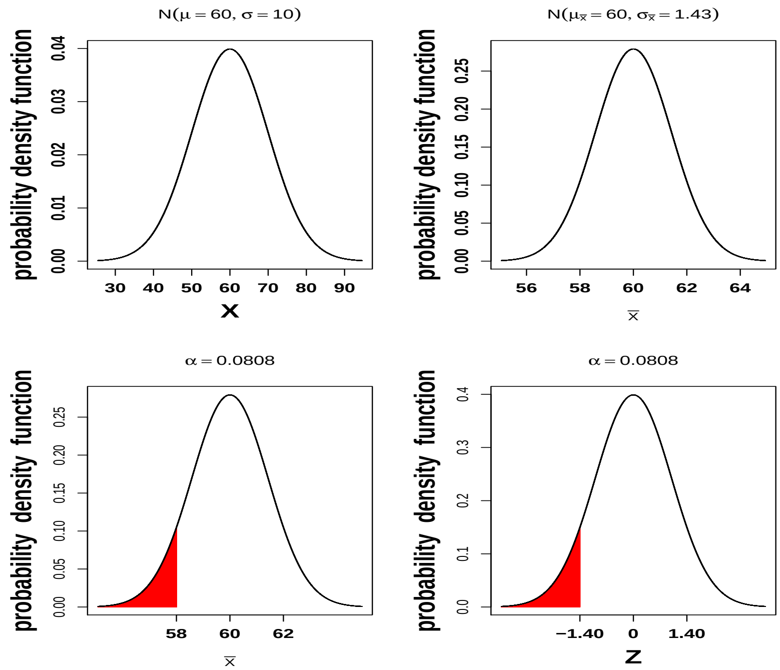

# Normal distribution

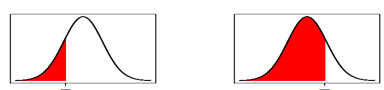
**Example:** (unrealistic) Let $X_1, \ldots, X_{49}$ be independent observations from a $N(\mu,\ \sigma = 10)$ distribution, where $\mu$ is unknown.

We are interested in testing if $\mu$ is **60** against the alternative that $\mu$ is less than **60**. We will reject the null hypothesis in favor of the alternative hypothesis if and only if the **sample mean** is **smaller** than **58**.

**(a)** State the null and alternative hypotheses.

**(b)** Determine $\boldsymbol{\alpha}$.



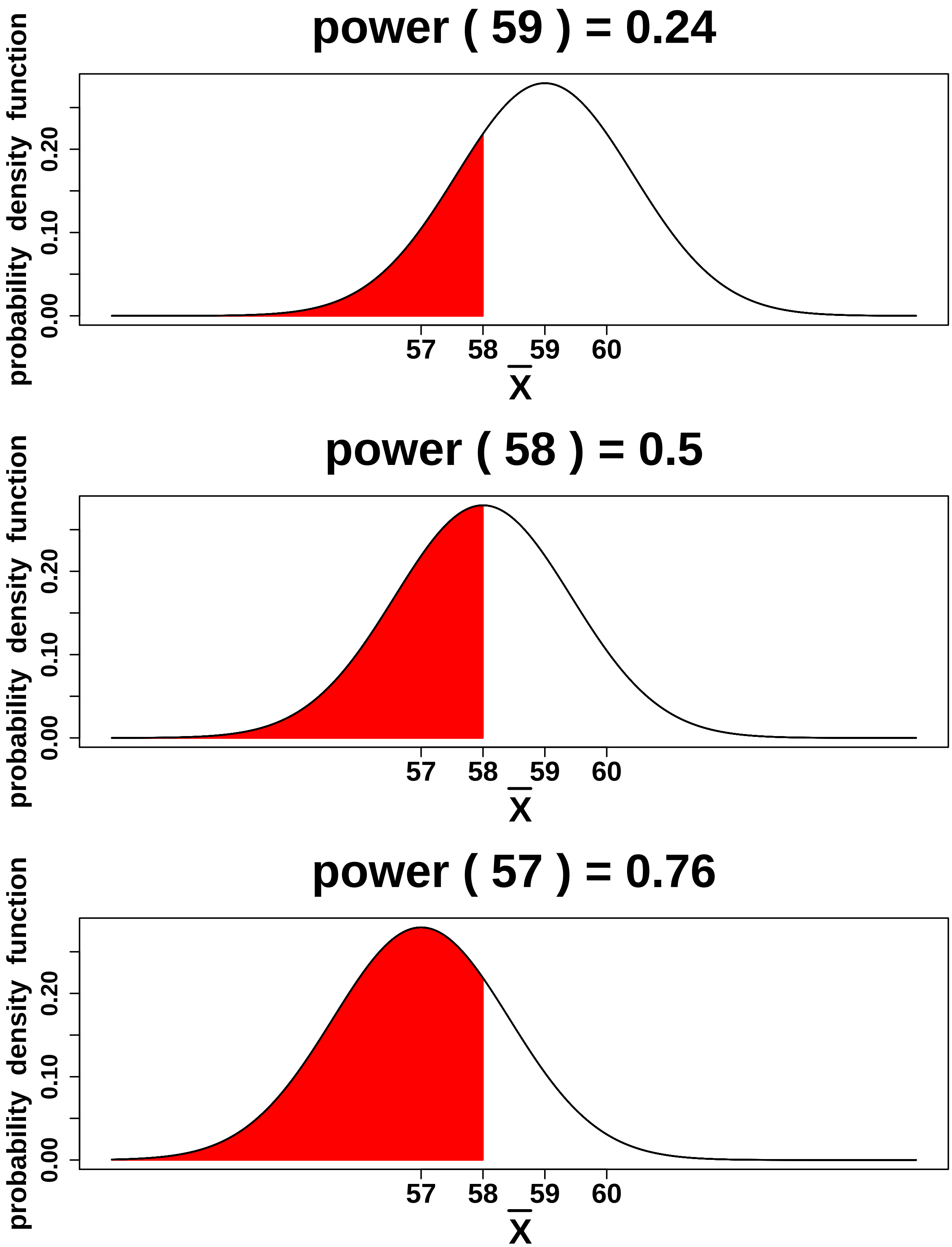



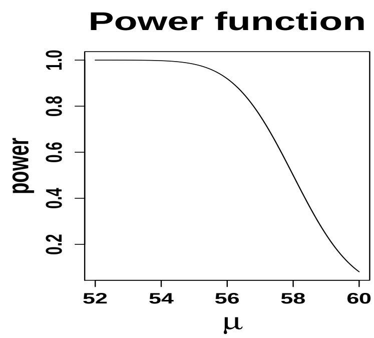

Standard normal table, pp. 722–723

| $z$ | **.00** | **.01** | **.02** | **.03** | **.04** | **.05** | **.06** | **.07** | **.08** | **.09** |
|---|---|---|---|---|---|---|---|---|---|---|
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| −1.5 | .0668 | .0655 | .0643 | .0630 | .0618 | .0606 | .0594 | .0582 | .0571 | .0559 |
| −1.4 | .0808 | .0793 | .0778 | .0764 | .0749 | .0735 | .0721 | .0708 | .0694 | .0681 |
| −1.3 | .0968 | .0951 | .0934 | .0918 | .0901 | .0885 | .0869 | .0853 | .0838 | .0823 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |

**(c)** Determine $\boldsymbol{\beta}$ and **power** under $\mu = 57, 57.5, 58, 58.5, 59$, and $59.5$.

| $\mu$ | 57 | 57.5 | 58 | 58.5 | 59 | 59.5 |
|---|---|---|---|---|---|---|
| $\beta(\mu)$ | 0.24 | 0.36 | 0.50 | 0.64 | 0.76 | 0.85 |
| power$(\mu)$ | 0.76 | 0.64 | 0.50 | 0.36 | 0.24 | 0.15 |


power ( 59 ) = 0.24
probability density function
0.00
0.10
0.20
57
58
59
60
X̄
power ( 58 ) = 0.5
probability density function
0.00
0.10
0.20
57
58
59
60
X̄
power ( 57 ) = 0.76
probability density function
0.00
0.10
0.20
57
58
59
60
X̄




# 8.2 Tests About a Population Mean

Set the level of significance, $\boldsymbol{\alpha}$, to some value in advance.

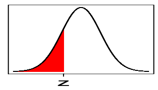
Find the **rejection region**.

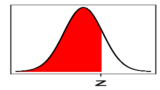
## Case I: Normal distribution with $\boldsymbol{\mu}$ unknown but $\boldsymbol{\sigma}$ known (unrealistic)

**Example:** *Revisit previous example.* (unrealistic) Let $X_1, \ldots, X_{49}$ be independent observations from a $N(\mu,\ \sigma = 10)$ distribution, where $\mu$ is unknown.

We are interested in testing if $\mu$ is **60** against the alternative that $\mu$ is less than **60** at **significance level** $\alpha = 0.05$.

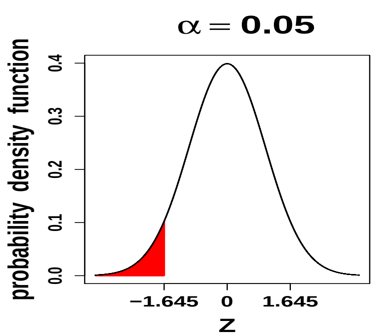
**(a)** State the **rejection region** in terms of $\boldsymbol{z}$, the **standard normal random variable**.



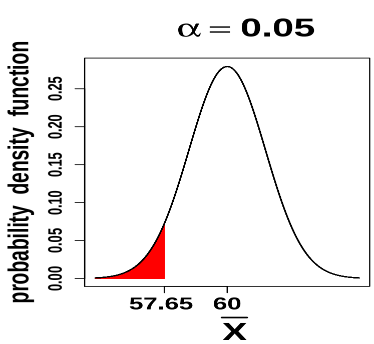




Standard normal table, pp. 722–723

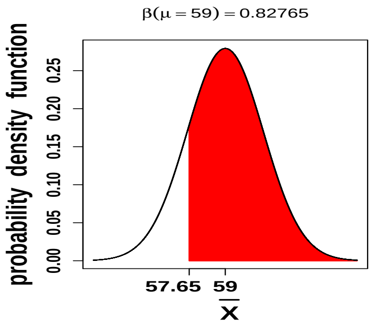
| $z$ | **.00** | **.01** | **.02** | **.03** | **.04** | **.05** | **.06** | **.07** | **.08** | **.09** |
|---|---|---|---|---|---|---|---|---|---|---|
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| −1.7 | .0446 | .0436 | .0427 | .0418 | .0409 | .0401 | .0392 | .0384 | .0375 | .0367 |
| −1.6 | .0548 | .0537 | .0526 | .0516 | .0505 | .0495 | .0485 | .0475 | .0465 | .0455 |
| −1.5 | .0668 | .0655 | .0643 | .0630 | .0618 | .0606 | .0594 | .0582 | .0571 | .0559 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |

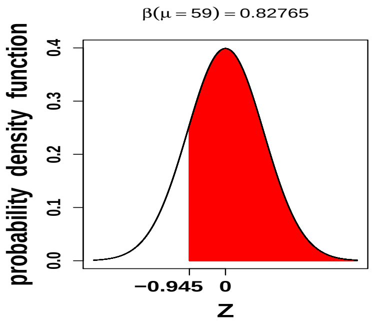


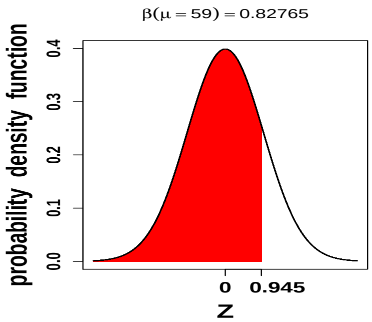

**(b)** State the **rejection region** in terms of $\bar{x}$, the **sample mean**.

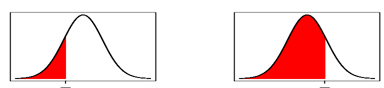



**(c)** Compute $\beta$ under $\mu = 59$, when $\alpha = 0.05$



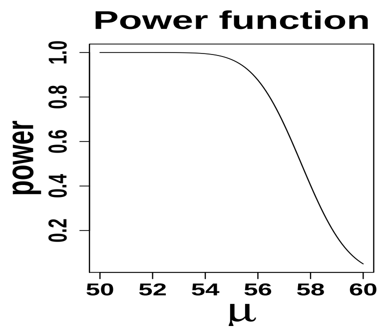




Standard normal table, pp. 722–723

| $z$ | .00 | .01 | .02 | .03 | .04 | .05 | .06 | .07 | .08 | .09 |
|---|---|---|---|---|---|---|---|---|---|---|
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 0.8 | .7881 | .7910 | .7939 | .7967 | .7995 | .8023 | .8051 | .8078 | .8106 | .8133 |
| 0.9 | .8159 | .8186 | .8212 | .8238 | .8264 | .8289 | .8315 | .8340 | .8365 | .8389 |
| 1.0 | .8413 | .8438 | .8461 | .8485 | .8508 | .8531 | .8554 | .8577 | .8599 | .8621 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |

**(d)** Suppose we want to set $\boldsymbol{\beta(\mu = 59) = 0.2}$ at **significance level** $\boldsymbol{\alpha = 0.05}$. What do we need to do?

**(e)** Graph the **power function** for $\mu$ between 50 and 60 at significance level $\alpha = 0.05$.




## Case II: Large-Sample Tests

Suppose $X_1, \ldots, X_n$ are independent observations sampled from **any** population with mean $\mu$ and positive finite standard deviation $\sigma$. By the **Central Limit Theorem**,

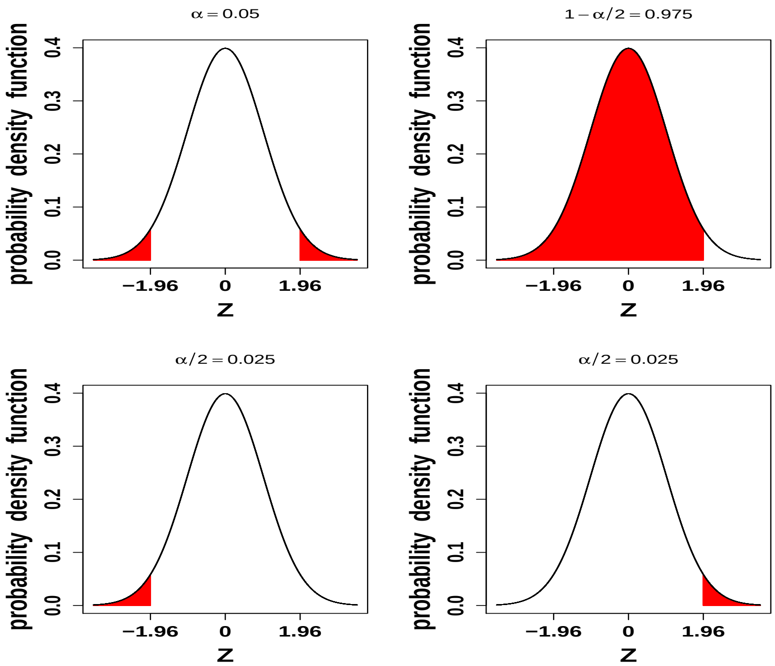
$$\bar{X} \overset{approx.}{\sim} N(\mu,\ \sigma/\sqrt{n}) \quad \text{for large } n.$$

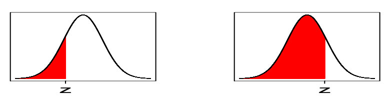
**Example:** The mean body temperature of healthy adults is typically assumed to be **98.6°F**. In a study at the University of Maryland of **106** healthy adults at midnight, the average temperature was found to be **98.20°F**, and the standard deviation was found to be **0.62°F**. Test the claim that the population mean body temperature differs from **98.20°F**, at significance level $\boldsymbol{\alpha = 0.05}$.

**(a)** State the null and alternative hypotheses, and define any notation used.

Let $\mu$ be the **population mean** body temperature of healthy adults.

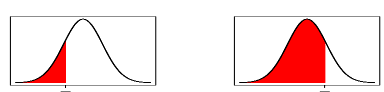
**(b)** Determine the **rejection region** in terms of the **standardized test statistic**.



Standard normal table, pp. 722–723

| $z$ | .00 | .01 | .02 | .03 | .04 | .05 | .06 | .07 | .08 | .09 |
|---|---|---|---|---|---|---|---|---|---|---|
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| −2.0 | .0228 | .0222 | .0217 | .0212 | .0207 | .0202 | .0197 | .0192 | .0188 | .0183 |
| −1.9 | .0287 | .0281 | .0274 | .0268 | .0262 | .0256 | .0250 | .0244 | .0239 | .0233 |
| −1.8 | .0359 | .0351 | .0344 | .0336 | .0329 | .0322 | .0314 | .0307 | .0301 | .0294 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |



Standard normal table, pp. 722–723

| $z$ | .00 | .01 | .02 | .03 | .04 | .05 | .06 | .07 | .08 | .09 |
|---|---|---|---|---|---|---|---|---|---|---|
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 1.8 | .9641 | .9649 | .9656 | .9664 | .9671 | .9678 | .9686 | .9693 | .9699 | .9706 |
| 1.9 | .9713 | .9719 | .9726 | .9732 | .9738 | .9744 | .9750 | .9756 | .9761 | .9767 |
| 2.0 | .9772 | .9778 | .9783 | .9788 | .9793 | .9798 | .9803 | .9808 | .9812 | .9817 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |

**(c)** Determine the **value** of the **standardized test statistic**.

**(d)** State the conclusion of the hypothesis test in **statistical terms** and in **regular English**.

## Case III: A Normal Population Distribution

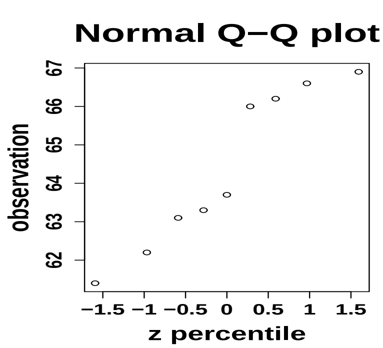
Suppose $X_1, \ldots, X_n$ are independent observations sampled from a $N(\mu,\ \sigma)$ distribution, where $\mu$ and $\sigma$ are **both unknown**.

**Example:** *Chicken Feed.* Suppose that the weight of a newborn chicken when the mother is fed enriched feed is approximately **normally** distributed. The **mean** weight of a newborn chicken is known to be 62.2 ounces when regular feed is fed to the mother. Does enriched feed increase the **mean** weight in comparison to regular feed? Test that claim, at significance level $\boldsymbol{\alpha = 0.01}$. In a sample, **9** chickens were born to mothers who were fed enriched feed, and their weights in ounces are the following: {61.4, 62.2, 66.9, 63.3, 66.2, 66.0, 63.1, 63.7, 66.6}.

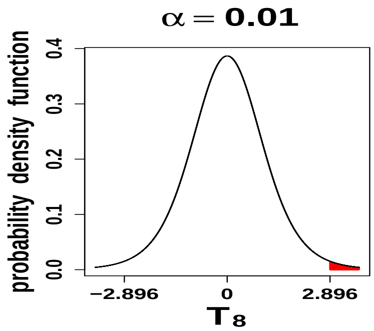
**(a)** State the null and alternative hypotheses, and define any notation used.

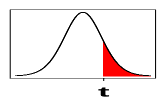
Let $\mu$ be the **population mean** weight of chickens whose mothers were given enriched feed.

**(b)** What assumption needs to be verified?



(c) Determine the **rejection region** in terms of the **standardized test statistic**.




*t*-table, p. 725

| | $\alpha$ | | | | | | |
|---|---|---|---|---|---|---|---|
| $\nu$ | **.10** | **.05** | **.025** | **.01** | **.005** | **.001** | **.0005** |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 7 | 1.415 | 1.895 | 2.365 | 2.998 | 3.499 | 4.785 | 5.408 |
| 8 | 1.397 | 1.860 | 2.306 | 2.896 | 3.355 | 4.501 | 5.041 |
| 9 | 1.383 | 1.833 | 2.262 | 2.821 | 3.250 | 4.297 | 4.781 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |

(d) Determine the **value** of the **standardized test statistic**.

(e) State the conclusion of the hypothesis test in **statistical terms** and in **regular English**.

□

**Remark:** For hypothesis tests on $\mu$, one-sided or two-sided tests may be performed.

**Remark:** Appropriate sample size may be determined for specified $\alpha$ and $\beta(\mu)$ for some $\mu$ in $H_a$.

# 8.3 Tests Concerning a Population Proportion

Let $\boldsymbol{p}$ be the probability of a Bernoulli response.

Let $X \sim \text{Binomial}(n, p)$.

$\hat{p} = X/n$

Previously we constructed the **confidence interval** on $p$: $\hat{p} \pm z \sqrt{\hat{p}\ (1-\hat{p})/n}$ for large $n\hat{p}$ and $n(1-\hat{p})$.

Perform **hypothesis testing** on $p$.

## Large-Sample Tests

For large $np$ and $n(1-p)$ $(\geq 10)$,

$$\frac{\hat{p}-p}{\sqrt{p(1-p)/n}} \overset{approx.}{\sim} N(0,\ 1)$$

by the Central Limit Theorem.

**Example:** *Coke and Pepsi.* Test at level $\alpha = 0.05$ if Coke or Pepsi is preferred over the other, using a blind taste test. Suppose that in a sample of 100 adults, 43 of them indicated a preference for Coke over Pepsi. *Assume that "no preference" is NOT a valid response.*

**(a)** State the null and alternative hypotheses, and define any notation used.

Let $p$ be the **population proportion** of people who prefer Coke over Pepsi.

**(b)** Check the **rule-of-thumb**.

**(c)** Determine the **rejection region** in terms of the **standardized test statistic**.

(d) Determine the **value** of the **standardized test statistic**.

(e) State the conclusion of the hypothesis test in **statistical terms** and in **regular English**.

□

**Improved approach:** Use normal approximation to binomial distribution with **continuity correction.**

**Example:** *Revisit Coke and Pepsi.* Test at level $\alpha = 0.05$ if Coke or Pepsi is preferred over the other, using a blind taste test. Suppose that in a sample of 100 adults, 43 of them indicated a preference for Coke over Pepsi. *Assume that "no preference" is NOT a valid response.*

(a) Determine the **value** of the **standardized test statistic**, using the normal approximation to the binomial distribution with **continuity correction**.

(b) Determine $\alpha$.

(c) Determine $\beta(p = 0.4)$.

(d) Determine power$(p = 0.4)$.

□

*Formulas for* **sample size determination** *are given on p. 308 regarding tests on p.*

## Small-Sample Tests

If $np$ or $n(1 - p)$ is not large (i.e., less than 10) under $H_0$, then do NOT use the normal approximation; just use the **binomial** distribution directly.

**Example:** *Revisit Coke and Pepsi.* Test if Coke or Pepsi is preferred over the other, using a blind taste test in a sample of 10 adults.

Let $X$ be the number of adults who prefer Coke, and define the **rejection region** for $X$ to be $R = \{0, 1, 9, 10\}$. *Assume that "no preference" is NOT a valid response.*

**(a)** State the null and alternative hypotheses, and define any notation used.

Let $p$ be the **population proportion** of people who prefer Coke over Pepsi.

**(b)** Check the **rule-of-thumb**.

**(c)** Determine $\alpha$.

**(d)** Determine $\beta(p = 0.4)$.

**(e)** Determine power$(p = 0.4)$.

□

# 8.4 *P*-values

## Hypothesis tests about proportions

Let $p$ = unknown population proportion.

Let $\hat{p}$ = sample proportion.

We make inferences on $p$ using the point estimate $\hat{p}$.

We use large samples and apply the Central Limit Theorem; i.e., $\hat{p}$ is approximately normal for large $n$.

### One-sample $Z$-test on a population proportion, $p$

**Example:** Suppose that the National Safety Council believes that more than 20% of all automobile accidents involve pedestrians. Test this claim at **significance level** $\alpha = 0.05$. Suppose a simple random sample of $n = 200$ automobile accidents results in $X = 46$ involving pedestrians.

**(a)** Define your notation.

Let $\boldsymbol{p}$ be the unknown **population** proportion of automobile accidents which involve pedestrians.

Let $\hat{\boldsymbol{p}}$ be the **sample** proportion of automobile accidents which involve pedestrians.

**(b)** State the hypotheses.

**(c)** Check the rule of thumb **under the null hypothesis**.

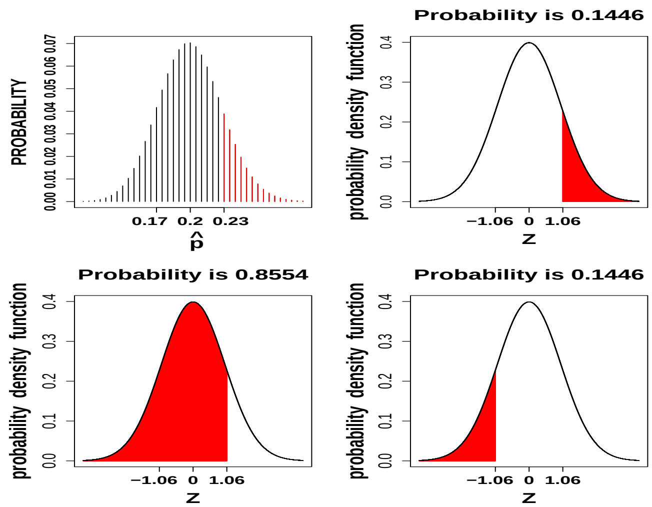
**(d)** Determine our specific value of $\hat{\boldsymbol{p}}$, the point estimate of $\boldsymbol{p}$.

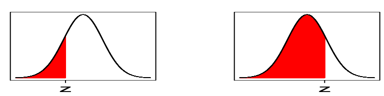
**(e)** What is the approximate distribution of $\hat{\boldsymbol{p}}$ under $H_0$?

**(f)** Find the value of the **standardized test statistic**.

**(g)** Find the $P$-value.

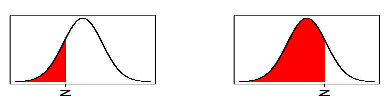
How can we obtain more accuracy?




Standard normal table, pp. 722–723

| $z$ | .00 | .01 | .02 | .03 | .04 | .05 | .06 | .07 | .08 | .09 |
|---|---|---|---|---|---|---|---|---|---|---|
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| −1.1 | .1357 | .1335 | .1314 | .1292 | .1271 | .1251 | .1230 | .1210 | .1190 | .1170 |
| −1.0 | .1587 | .1562 | .1539 | .1515 | .1492 | .1469 | .1446 | .1423 | .1401 | .1379 |
| −0.9 | .1841 | .1814 | .1788 | .1762 | .1736 | .1711 | .1685 | .1660 | .1635 | .1611 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |




Standard normal table, pp. 722–723

| $z$ | .00 | .01 | .02 | .03 | .04 | .05 | .06 | .07 | .08 | .09 |
|---|---|---|---|---|---|---|---|---|---|---|
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 0.9 | .8159 | .8186 | .8212 | .8238 | .8264 | .8289 | .8315 | .8340 | .8365 | .8389 |
| 1.0 | .8413 | .8438 | .8461 | .8485 | .8508 | .8531 | .8554 | .8577 | .8599 | .8621 |
| 1.1 | .8643 | .8665 | .8686 | .8708 | .8729 | .8749 | .8770 | .8790 | .8810 | .8830 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |

The ***P*-value** is the probability of obtaining a value of the standardized test statistic at least as extreme as the observed value, based on the assumption that $H_0$ is true.

The smaller the $P$-value, the stronger the evidence against $H_0$.

When the $P$-value is small, we say that the data are **statistically significant**.

In this example, the $P$-value is 0.1446.

**(h)** State the conclusion in statistical terms and in regular English.

*Is the $P$-value small enough that we should reject $H_0$?*

□

*Rule:* Reject $H_0$ in favor of $H_a$ if $P$-value $\leq \alpha$; otherwise, fail to reject $H_0$.

When $P$-value $= 0.1446$, would $H_0$ be rejected if $\alpha = 0.05$, $\alpha = 0.1$, $\alpha = 0.2$, $\alpha = 0.15$, and $\alpha = 0.1446$?

*A mathematically rigorous definition:* The ***P*-value** is the smallest value of $\alpha$ for which the null hypothesis would be rejected.

The ***P*-value** is also called the **observed** significance level.

Do researchers typically prefer **small** or **large** $P$-values?

The $P$-value is NOT $P(H_0$ is true).

For example, if the DNA match is 0.01 for the defendant, does this imply that there is a 1% chance that the defendant is innocent?

$P$-value is $P$(Such strong evidence would exist against the defendant, given that the defendant is innocent).

$P$-value is NOT $P$(Defendant is innocent).

**Example:** You visit a foreign country on vacation and get thrown into jail for no apparent reason. The blood at some crime scene is type $A$, and you (who unluckily have type $A$ blood) become the defendant. Suppose that 42% of all people have type $A$ blood. Is there a 42% chance that you are innocent?

Is there an 58% chance that you are guilty?

□

**Example:** Suppose that two summers ago 60% of *then-recent* high school graduates enrolled in college. We are interested in whether or not the college enrollment rate changed since two summers ago.

Test the claim at **significance level** $\alpha = 0.1$.

Suppose a simple random sample of 500 most recent high school graduates results in 275 enrolled in college.

**(a)** Define your notation.

Let $\boldsymbol{p}$ be the unknown **population** proportion of most recent high school graduates who are enrolled in college.

Let $\hat{\boldsymbol{p}}$ be the **sample** proportion of most recent high school graduates who are enrolled in college.

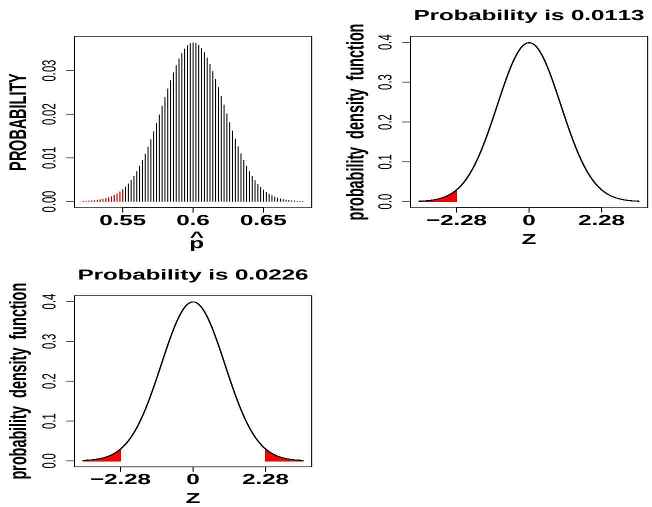
**(b)** State the null and alternative hypotheses.

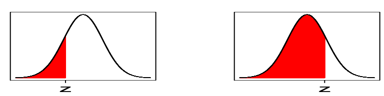
**(c)** Check the rule of thumb **under the null hypothesis**.

**(d)** Under $H_0$, what is the approximate distribution of $\hat{p}$, the point estimate of $p$?

**(e)** Find the value of the **standardized test statistic**.

**(f)** Find the $P$-value.



Standard normal table, pp. 722–723

| $z$ | .00 | .01 | .02 | .03 | .04 | .05 | .06 | .07 | .08 | .09 |
|---|---|---|---|---|---|---|---|---|---|---|
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| −2.3 | .0107 | .0104 | .0102 | .0099 | .0096 | .0094 | .0091 | .0089 | .0087 | .0084 |
| −2.2 | .0139 | .0136 | .0132 | .0129 | .0125 | .0122 | .0119 | .0116 | .0113 | .0110 |
| −2.1 | .0179 | .0174 | .0170 | .0166 | .0162 | .0158 | .0154 | .0150 | .0146 | .0143 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |

**(g)** State the conclusion in statistical terms and in regular English.

We conclude that the **population** proportion of most recent high school graduates who are enrolled in college DIFFERS from 60%.

□

**Remark:** Commonly used values of $\alpha$ are 0.01, **0.05**, and 0.1.

# Hypothesis tests about means

Let $\mu$ = unknown population mean.

Let $\bar{X}$ = sample mean.

We make inferences on $\mu$ using the **point estimate** $\bar{X}$.

We use large samples and apply Central Limit Theorem; i.e., $\bar{X}$ is approximately normal for **large** $n$. Alternatively, we start with an approximately normal population, in which case $\bar{X}$ is approximately normal for **any** $n$.

## One-sample $t$-test on a population mean, $\mu$

*Recall:* For independent or nearly independent observations (and positive finite $\sigma$), if **the original population is approximately normal OR $n$ is large**, then

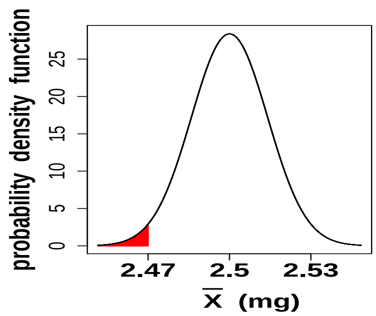
$$T = \frac{\bar{X} - \mu}{s/\sqrt{n}} \overset{approx.}{\sim} t_{n-1}.$$

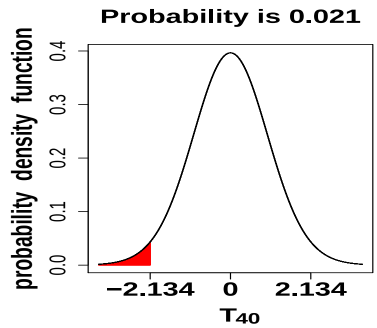
**Example:** The manufacturer of Cardio-delight Inc. lists the mean saturated fat content as being 2.5 grams per chocolate chip cookie. Sugarland Inc. produces a new brand of chocolate chip cookies with the same great taste (and weight/mass) as Cardio-delight, but claims that this new brand has a lower mean saturated fat content. To test the claim of Sugarland at significance level $\alpha = 0.1$, a researcher samples the saturated fat content of 41 Sugarland cookies and finds $\bar{X} = 2.47$ g and $s = 0.09$ g. Let $\mu$ = population mean saturated fat content of Sugarland cookies.

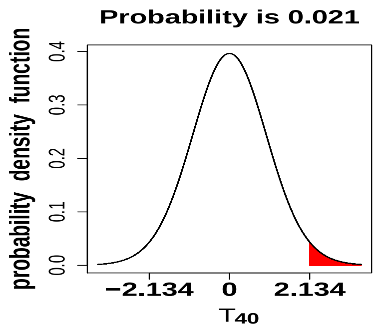
**(a)** State the null and alternative hypotheses.

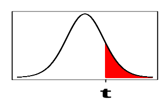
**(b)** Find the value of the **standardized test statistic**.

**(c)** Find the $P$-value.




$t$-table, pp. 728–729

| $t \setminus \nu$ | 29 | 30 | 35 | 40 | 60 | 120 | $\infty(=z)$ |
|---|---|---|---|---|---|---|---|
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| **1.9** | .034 | .034 | .033 | .032 | .031 | .030 | .029 |
| **2.0** | .027 | .027 | .027 | .026 | .025 | .024 | .023 |
| **2.1** | .022 | .022 | .022 | .021 | .020 | .019 | .018 |
| **2.2** | .018 | .018 | .017 | .017 | .016 | .015 | .014 |
| **2.3** | .014 | .014 | .014 | .013 | .012 | .012 | .011 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |

**(d)** State the conclusion in statistical terms and in regular English.

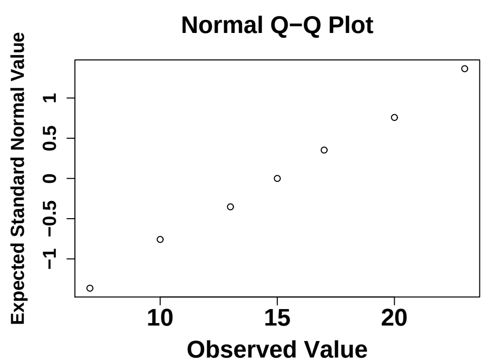
We conclude that the **population** mean saturated fat content of Sugarland cookies is less than 2.5 grams.

□

**Example:** Suppose the mean lifetime of mice is 26 months. Test at level 0.01 if a new strain of mice has mean lifetime different from 26 months. Seven mice are independently sampled, and their lifetimes in months are $\{20,\ 23,\ 13,\ 7,\ 17,\ 15,\ 10\}$.

**(a)** Define your notation.

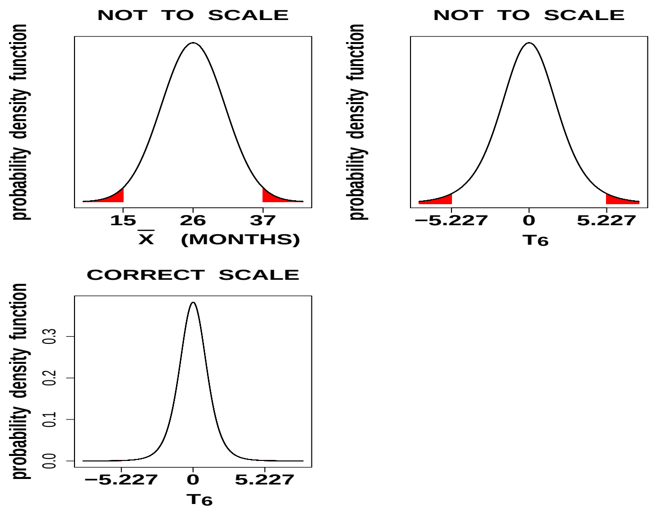
**(b)** Is the original population approximately normal, or is the sample size large?

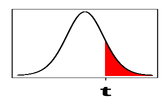



**(c)** State the null and alternative hypotheses.

**(d)** Find the value of the **standardized test statistic**.

**(e)** Find the $P$-value.



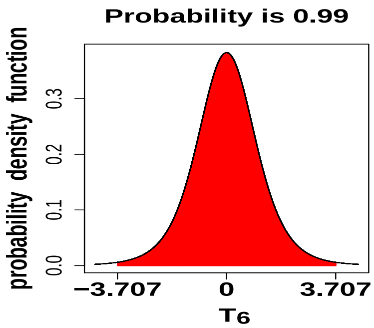

$t$-table, pp. 728–729

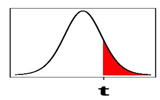
| $t \setminus \nu$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| **3.7** | .084 | .033 | .017 | .010 | .007 | .005 | .004 |
| **3.8** | .082 | .031 | .016 | .010 | .006 | .004 | .003 |
| **3.9** | .080 | .030 | .015 | .009 | .006 | .004 | .003 |
| **4.0** | .078 | .029 | .014 | .008 | .005 | .004 | .003 |

**(f)** State the conclusion in statistical terms and in regular English.

We conclude that this new strain of mice has **population** mean lifetime **different** from 26 months.

**(g)** Now, construct a 99% confidence interval on $\mu$.




$t$-table, p. 725

| | $\alpha$ | | | | | | |
|---|---|---|---|---|---|---|---|
| $\nu$ | **.10** | **.05** | **.025** | **.01** | **.005** | **.001** | **.0005** |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 5 | 1.476 | 2.015 | 2.571 | 3.365 | 4.032 | 5.893 | 6.869 |
| 6 | 1.440 | 1.943 | 2.447 | 3.143 | 3.707 | 5.208 | 5.959 |
| 7 | 1.415 | 1.895 | 2.365 | 2.998 | 3.499 | 4.785 | 5.408 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |

**Layman's interpretation:** We are 99% confident that the population mean lifetime of this new strain of mice is between 7.20 months and 22.80 months.

**Mathematically rigorous interpretation:** If we repeat the sampling procedure many times to construct many 99% confidence intervals on $\mu$, the population mean lifetime of this new strain of mice, then approximately 99% of these 99% confidence intervals will contain the true value of $\mu$.

**(h)** Is our 99% confidence interval consistent with the conclusion of our 2-sided hypothesis test of level $\alpha = 0.01$?

**(i)** Suppose we had tested $H_0 : \mu = 20$ months against $H_a : \mu \neq 20$ months at level 0.01.

□