

# Tests of Hypotheses: Means

We often use inferential statistics to make decisions or judgments about the value of a parameter, such as a population mean. For example, we might need to decide whether the mean weight,  $\mu$ , of all bags of pretzels packaged by a particular company differs from the advertised weight of 454 grams (g), or we might want to determine whether the mean age,  $\mu$ , of all cars in use has increased from the year 2000 mean of 9.0 years.

One of the most commonly used methods for making such decisions or judgments is to perform a *hypothesis test*. A **hypothesis** is a statement that something is true. For example, the statement “the mean weight of all bags of pretzels packaged differs from the advertised weight of 454 g” is a hypothesis.

To develop procedures for testing statistical hypotheses, we must always know exactly what to expect when a hypothesis is true, and it is for this reason that we often hypothesize the opposite of what we hope to prove. Suppose, for instance, that we suspect that a dice game is not honest. If we formulate the null hypothesis that the dice are crooked, everything would depend on how crooked they are, but if we assume that they are perfectly balanced, we could calculate all the necessary probabilities and take it from there.

The idea of setting up a null hypothesis is common in nonstatistical thinking. It is precisely what we do in criminal proceedings, where an accused is presumed to be innocent until his guilt has been established beyond a reasonable doubt. The presumption of innocence is a null hypothesis.

Typically, a hypothesis test involves two hypotheses: the *null hypothesis* and the *alternative hypothesis*, which we define as follows.

**Null hypothesis:** A hypothesis to be tested. We use the symbol  $H_0$  to represent the null hypothesis.

**Alternative hypothesis:** A hypothesis that we use as an alternative to the null hypothesis, namely, the hypothesis that we accept when the null hypothesis is rejected. We use the symbol  $H_A$  to represent the alternative hypothesis.

For instance, in the pretzel packaging example, the null hypothesis might be “the mean weight of all bags of pretzels packaged equals the advertised weight of 454 g,” and the alternative hypothesis might be “the mean weight of all bags of pretzels packaged differs from the advertised weight of 454 g.”

## Choosing the Hypotheses

The first step in setting up a hypothesis test is to decide on the null hypothesis and the alternative hypothesis. The following are some guidelines for choosing these two hypotheses. Although the guidelines refer specifically to hypothesis tests for one population mean,  $\mu$ , they apply to any hypothesis test concerning one parameter.

**Null hypothesis:** The null hypothesis for a hypothesis test concerning a population mean,  $\mu$ , always specifies a single value for that parameter. Hence we can express the null hypothesis as

$$H_0 : \mu = \mu_0$$

where  $\mu_0$  is some number.

**Alternative hypothesis:** The alternative hypothesis must always be formulated together with the null hypothesis, for otherwise we would not know when to reject  $H_0$ . The choice of the alternative hypothesis depends on and should reflect the purpose of the hypothesis test. Three choices are possible for the alternative hypothesis.

- If the primary concern is deciding whether a population mean,  $\mu$ , is different from a specified value  $\mu_0$ , we express the alternative hypothesis as

$$H_A : \mu \neq \mu_0$$

- If the primary concern is deciding whether a population mean,  $\mu$ , is *less than* a specified value  $\mu_0$ , we express the alternative hypothesis as

$$H_A : \mu < \mu_0$$

- If the primary concern is deciding whether a population mean,  $\mu$ , is *greater than* a specified value  $\mu_0$ , we express the alternative hypothesis as

$$H_A : \mu > \mu_0$$

EXAMPLE: A snack-food company produces a 454-g bag of pretzels. Although the actual net weights deviate slightly from 454 g and vary from one bag to another, the company insists that the mean net weight of the bags be 454 g. As part of its program, the quality assurance department periodically performs a hypothesis test to decide whether the packaging machine is working properly, that is, to decide whether the mean net weight of all bags packaged is 454 g.

- Determine the null hypothesis for the hypothesis test.
- Determine the alternative hypothesis for the hypothesis test.

Solution: Let  $\mu$  denote the mean net weight of all bags packaged.

- The null hypothesis is that the packaging machine is working properly, that is, that the mean net weight,  $\mu$ , of all bags packaged *equals* 454 g. In symbols,  $H_0 : \mu = 454$  g.
- The alternative hypothesis is that the packaging machine is not working properly, that is, that the mean net weight,  $\mu$ , of all bags packaged is *different from* 454 g. In symbols,  $H_A : \mu \neq 454$  g.

EXAMPLE: The R. R. Bowker Company collects information on the retail prices of books and publishes the data in The Bowker Annual Library and Book Trade Almanac. In 2005, the mean retail price of history books was \$78.01. Suppose that we want to perform a hypothesis test to decide whether this year's mean retail price of history books has increased from the 2005 mean.

- Determine the null hypothesis for the hypothesis test.
- Determine the alternative hypothesis for the hypothesis test.

Solution: Let  $\mu$  denote this year's mean retail price of history books.

- The null hypothesis is that this year's mean retail price of history books *equals* the 2005 mean of \$78.01; that is,  $H_0 : \mu = \$78.01$ .
- The alternative hypothesis is that this year's mean retail price of history books is *greater than* the 2005 mean of \$78.01; that is,  $H_A : \mu > \$78.01$ .

EXAMPLE: Calcium is the most abundant mineral in the human body and has several important functions. Most body calcium is stored in the bones and teeth, where it functions to support their structure. Recommendations for calcium are provided in Dietary Reference Intakes, developed by the Institute of Medicine of the National Academy of Sciences. The recommended adequate intake (RAI) of calcium for adults (ages 19-50 years) is 1000 milligrams (mg) per day. Suppose that we want to perform a hypothesis test to decide whether the average adult with an income below the poverty level gets less than the RAI of 1000 mg.

- (a) Determine the null hypothesis for the hypothesis test.
- (b) Determine the alternative hypothesis for the hypothesis test.

Solution: Let  $\mu$  denote the mean calcium intake (per day) of all adults with incomes below the poverty level.

- (a) The null hypothesis is that the mean calcium intake of all adults with incomes below the poverty level *equals* the RAI of 1000 mg per day; that is,  $H_0 : \mu = 1000$  mg.
- (b) The alternative hypothesis is that the mean calcium intake of all adults with incomes below the poverty level is *less than* the RAI of 1000 mg per day; that is,  $H_A : \mu < 1000$  mg.

EXAMPLE: The average drying time of a manufacturer's paint is 20 minutes. Investigating the effectiveness of a modification in the chemical composition of the paint, the manufacturer wants to test the null hypothesis that  $\mu = 20$  against a suitable alternative, where  $\mu$  is the mean drying time of the modified paint,

- (a) What alternative hypothesis should the manufacturer use if she does not want to make the modification unless it actually decreases the drying time of the paint?
- (b) What alternative hypothesis should the manufacturer use if she wants to make the modification unless it actually increases the drying time of the paint?

Solution:

- (a) He should use the alternative hypothesis  $\mu < 20$  and make the modification only if the null hypothesis can be rejected.
- (b) He should use the alternative hypothesis  $\mu > 20$  and make the modification unless the null hypothesis is rejected.

## Type I and Type II Errors

Any decision we make based on a hypothesis test may be incorrect because we have used partial information obtained from a sample to draw conclusions about the entire population. There are two types of incorrect decisions — *Type I error* and *Type II error*.

|                | Accept $H_0$     | Reject $H_0$     |
|----------------|------------------|------------------|
| $H_0$ is true  | Correct decision | Type I error     |
| $H_0$ is false | Type II error    | Correct decision |

EXAMPLE: A psychologist wants to test the hypothesis that it takes an adult 0.38 second to react to a visual stimulus. He uses the null hypothesis

$$H_0 : \mu = 0.38 \text{ second}$$

against the alternating hypothesis

$$H_A : \mu \neq 0.38 \text{ second}$$

To perform this test, the psychologist decides to take a random sample of  $n = 40$  adults with the intention of accepting the null hypothesis if the mean of the sample falls anywhere between 0.36 second and 0.40 second; otherwise he will reject it. Assume that it is known from similar studies that  $\sigma = 0.08$  second for this kind of data.

- Find the probability of Type I error;
- Find the probability of Type II error when  $\mu = 0.41$ ;
- Suppose that the psychologist has actually taken a sample and obtained  $\bar{x} = 0.408$ . What will he decide and will it be an error if  $\mu = 0.38$  second? What will he decide and will it be an error if  $\mu = 0.42$  second?

Solution:

- (a) We have

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{0.08}{\sqrt{40}} \approx 0.0126$$

therefore

$$z = \frac{0.36 - 0.38}{0.0126} \approx -1.59 \quad \text{and} \quad z = \frac{0.40 - 0.38}{0.0126} \approx 1.59$$

It follows from Table I that the area in each tail of the sampling distribution is  $0.5 - 0.4441 = 0.0559$  and hence the probability of getting a value in either tail of the sampling distribution is  $2(0.0559) = 0.1118$ .

- (b) We have

$$z = \frac{0.36 - 0.41}{0.0126} \approx -3.97 \quad \text{and} \quad z = \frac{0.40 - 0.41}{0.0126} \approx -0.79$$

Since the area under the curve to the left of  $-3.97$  is negligible, it follows from Table I that the area of the region to the left of 0.4 is  $0.5 - 0.2852 = 0.2148$ .

(c) Since  $\bar{x} = 0.408$  exceeds 0.40, the psychologist will reject the null hypothesis  $\mu = 0.38$  second. Since the null hypothesis is true and rejected, the psychologist will be *making a Type I error* if  $\mu = 0.38$ . Since the null hypothesis is false and rejected, the psychologist will *not be making an error* if  $\mu = 0.42$ .

EXAMPLE: For a given population with  $\sigma = \$12$ , we want to test the null hypothesis  $\mu = \$75$  on the basis of random sample of size  $n = 100$ . If the null hypothesis is rejected when  $\bar{x}$  is greater than or equal to \$76.50 and otherwise it is accepted, find

- (a) the probability of Type I error;
- (b) the probability of Type II error when  $\mu = \$75.3$ ;
- (c) the probability of Type II error when  $\mu = \$77.22$ .

Solution: We have

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{12}{\sqrt{100}} = 1.2$$

(a) We have

$$z = \frac{76.5 - 75}{1.2} = \frac{1.5}{1.2} = 1.25$$

and the probability is  $0.5 - 0.3944 = 0.1056$ .

(b) We have

$$z = \frac{76.5 - 75.3}{1.2} = \frac{1.2}{1.2} = 1$$

and the probability is  $0.5 + 0.3413 = 0.8413$ .

(c) We have

$$z = \frac{76.5 - 77.22}{1.2} = \frac{-0.72}{1.2} = -0.6$$

and the probability is  $0.5 - 0.2257 = 0.2743$ .

**TABLE A.1 AREAS OF THE NORMAL DISTRIBUTION**

| <b>z</b> | <b>0.00</b> | <b>0.01</b> | <b>0.02</b> | <b>0.03</b> | <b>0.04</b> | <b>0.05</b> | <b>0.06</b> | <b>0.07</b> | <b>0.08</b> | <b>0.09</b> |
|----------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| 0.0      | 0.0000      | 0.0040      | 0.0080      | 0.0120      | 0.0160      | 0.0199      | 0.0239      | 0.0279      | 0.0319      | 0.0359      |
| 0.1      | 0.0398      | 0.0438      | 0.0478      | 0.0517      | 0.0557      | 0.0596      | 0.0636      | 0.0675      | 0.0714      | 0.0754      |
| 0.2      | 0.0793      | 0.0832      | 0.0871      | 0.0910      | 0.0948      | 0.0987      | 0.1026      | 0.1064      | 0.1103      | 0.1141      |
| 0.3      | 0.1179      | 0.1217      | 0.1255      | 0.1293      | 0.1331      | 0.1368      | 0.1406      | 0.1443      | 0.1480      | 0.1517      |
| 0.4      | 0.1554      | 0.1591      | 0.1628      | 0.1664      | 0.1700      | 0.1736      | 0.1772      | 0.1808      | 0.1844      | 0.1879      |
| 0.5      | 0.1915      | 0.1950      | 0.1985      | 0.2019      | 0.2054      | 0.2088      | 0.2123      | 0.2157      | 0.2190      | 0.2224      |
| 0.6      | 0.2258      | 0.2291      | 0.2324      | 0.2357      | 0.2389      | 0.2422      | 0.2454      | 0.2486      | 0.2518      | 0.2549      |
| 0.7      | 0.2580      | 0.2612      | 0.2642      | 0.2673      | 0.2704      | 0.2734      | 0.2764      | 0.2794      | 0.2823      | 0.2852      |
| 0.8      | 0.2881      | 0.2910      | 0.2939      | 0.2967      | 0.2996      | 0.3023      | 0.3051      | 0.3078      | 0.3106      | 0.3133      |
| 0.9      | 0.3159      | 0.3186      | 0.3212      | 0.3238      | 0.3264      | 0.3289      | 0.3315      | 0.3340      | 0.3365      | 0.3389      |
| 1.0      | 0.3413      | 0.3438      | 0.3461      | 0.3485      | 0.3508      | 0.3531      | 0.3554      | 0.3577      | 0.3599      | 0.3621      |
| 1.1      | 0.3643      | 0.3665      | 0.3686      | 0.3708      | 0.3729      | 0.3749      | 0.3770      | 0.3790      | 0.3810      | 0.3830      |
| 1.2      | 0.3849      | 0.3869      | 0.3888      | 0.3907      | 0.3925      | 0.3944      | 0.3962      | 0.3980      | 0.3997      | 0.4015      |
| 1.3      | 0.4032      | 0.4049      | 0.4066      | 0.4082      | 0.4099      | 0.4115      | 0.4131      | 0.4147      | 0.4162      | 0.4177      |
| 1.4      | 0.4192      | 0.4207      | 0.4222      | 0.4236      | 0.4251      | 0.4265      | 0.4279      | 0.4292      | 0.4306      | 0.4319      |
| 1.5      | 0.4332      | 0.4345      | 0.4357      | 0.4370      | 0.4382      | 0.4394      | 0.4406      | 0.4418      | 0.4429      | 0.4441      |
| 1.6      | 0.4452      | 0.4463      | 0.4474      | 0.4484      | 0.4495      | 0.4505      | 0.4515      | 0.4525      | 0.4535      | 0.4545      |
| 1.7      | 0.4554      | 0.4564      | 0.4573      | 0.4582      | 0.4591      | 0.4599      | 0.4608      | 0.4616      | 0.4625      | 0.4633      |
| 1.8      | 0.4641      | 0.4649      | 0.4656      | 0.4664      | 0.4671      | 0.4678      | 0.4686      | 0.4693      | 0.4699      | 0.4706      |
| 1.9      | 0.4713      | 0.4719      | 0.4726      | 0.4732      | 0.4738      | 0.4744      | 0.4750      | 0.4756      | 0.4761      | 0.4767      |
| 2.0      | 0.4772      | 0.4778      | 0.4783      | 0.4788      | 0.4793      | 0.4798      | 0.4803      | 0.4808      | 0.4812      | 0.4817      |
| 2.1      | 0.4821      | 0.4826      | 0.4830      | 0.4834      | 0.4838      | 0.4842      | 0.4846      | 0.4850      | 0.4854      | 0.4857      |
| 2.2      | 0.4861      | 0.4864      | 0.4868      | 0.4871      | 0.4875      | 0.4878      | 0.4881      | 0.4884      | 0.4887      | 0.4890      |
| 2.3      | 0.4893      | 0.4896      | 0.4898      | 0.4901      | 0.4904      | 0.4906      | 0.4909      | 0.4911      | 0.4913      | 0.4916      |
| 2.4      | 0.4918      | 0.4920      | 0.4922      | 0.4925      | 0.4927      | 0.4929      | 0.4931      | 0.4932      | 0.4934      | 0.4936      |
| 2.5      | 0.4938      | 0.4940      | 0.4941      | 0.4943      | 0.4945      | 0.4946      | 0.4948      | 0.4949      | 0.4951      | 0.4952      |
| 2.6      | 0.4953      | 0.4955      | 0.4956      | 0.4957      | 0.4959      | 0.4960      | 0.4961      | 0.4962      | 0.4963      | 0.4964      |
| 2.7      | 0.4965      | 0.4966      | 0.4967      | 0.4968      | 0.4969      | 0.4970      | 0.4971      | 0.4972      | 0.4973      | 0.4974      |
| 2.8      | 0.4974      | 0.4975      | 0.4976      | 0.4977      | 0.4977      | 0.4978      | 0.4979      | 0.4979      | 0.4980      | 0.4981      |
| 2.9      | 0.4981      | 0.4982      | 0.4982      | 0.4983      | 0.4984      | 0.4984      | 0.4985      | 0.4985      | 0.4986      | 0.4986      |
| 3.0      | 0.4987      | 0.4987      | 0.4987      | 0.4988      | 0.4988      | 0.4989      | 0.4989      | 0.4989      | 0.4990      | 0.4990      |
| 3.1      | 0.4990      | 0.4991      | 0.4991      | 0.4991      | 0.4992      | 0.4992      | 0.4992      | 0.4992      | 0.4993      | 0.4993      |
| 3.2      | 0.4993      | 0.4993      | 0.4994      | 0.4994      | 0.4994      | 0.4994      | 0.4994      | 0.4995      | 0.4995      | 0.4995      |
| 3.3      | 0.4995      | 0.4995      | 0.4995      | 0.4996      | 0.4996      | 0.4996      | 0.4996      | 0.4996      | 0.4996      | 0.4997      |
| 3.4      | 0.4997      | 0.4997      | 0.4997      | 0.4997      | 0.4997      | 0.4997      | 0.4997      | 0.4997      | 0.4997      | 0.4998      |
| 3.5      | 0.4998      | 0.4998      | 0.4998      | 0.4998      | 0.4998      | 0.4998      | 0.4998      | 0.4998      | 0.4998      | 0.4998      |
| 3.6      | 0.4998      | 0.4998      | 0.4999      | 0.4999      | 0.4999      | 0.4999      | 0.4999      | 0.4999      | 0.4999      | 0.4999      |
| 3.7      | 0.4999      | 0.4999      | 0.4999      | 0.4999      | 0.4999      | 0.4999      | 0.4999      | 0.4999      | 0.4999      | 0.4999      |
| 3.8      | 0.4999      | 0.4999      | 0.4999      | 0.4999      | 0.4999      | 0.4999      | 0.4999      | 0.4999      | 0.4999      | 0.4999      |
| 3.9      | 0.49995     | 0.49995     | 0.49996     | 0.49996     | 0.49996     | 0.49996     | 0.49996     | 0.49996     | 0.49997     | 0.49997     |