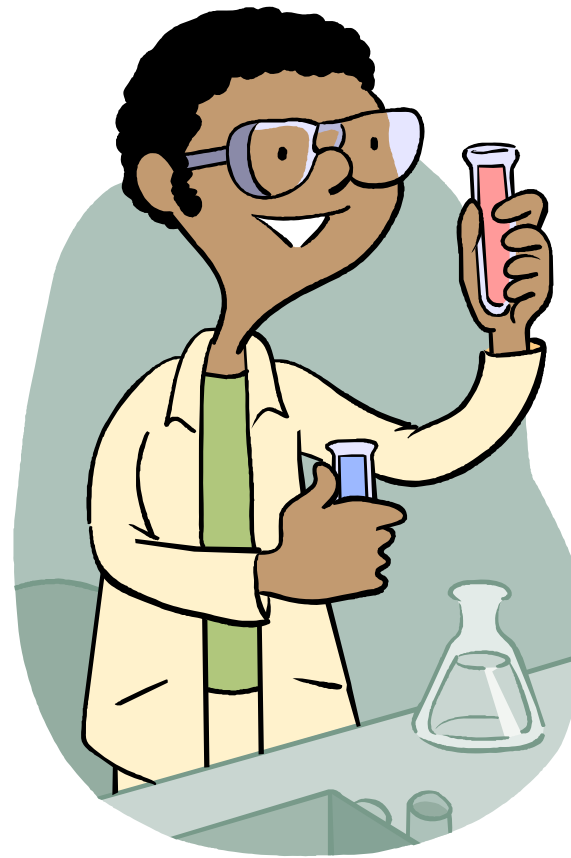# HYPOTHESIS TESTING

Hypothesis testing uses statistics to test the likelihood of a claim made about a population. The claim could be, for example, that a mean, proportion, or standard deviation has a particular value.

We will illustrate first the *classical method* for hypothesis testing.

We will begin our example with a statement that the mean of a population takes on a certain value. This statement is called the *null hypothesis*. In general, we think of a null hypothesis as a statement that nothing unusual is going to happen. That is, we assume the results from our sample are going to support the null hypothesis. Also, we generally denote the null hypothesis by H_0.

Let's do this. We know that IQ is normally distributed with a mean of 100 and a standard deviation of 15. Suppose we take a random sample of 100 college students. We will initially assume that the average IQ of college students is not statistically different from the population mean of 100.

$$H_0 : \mu = 100$$

Next, we need to formulate an *alternative hypothesis* that we will lean toward if the null hypothesis is unlikely to be true. The alternative hypothesis can by denoted by *H₁* or *Hₐ* or *H_A*.

$$H_0 : \mu = 100$$
$$H_1 : \mu \neq 100$$

We know that the distribution of the sample means will be normally distributed with mean equal to *mu* and standard deviation equal to sigma divided by the square root of n. Thus, we can easily convert our sample mean to a z score. We will call the formula for doing this our *test statistic formula* .

$H_0 : \mu = 100$

$H_1 : \mu \neq 100$

Test Statistic Formula: $z = \dfrac{\bar{x} - \mu}{\sigma / \sqrt{n}}$

Our criteria for rejecting the null hypothesis will be if we get a sample mean which is *too far away* from 100 in either direction. By *too far away*, we mean that we get a sample mean that could have occurred by chance only 5 times out of a hundred or less. We call this value the level of significance or alpha.

$H_0 : \mu = 100$

$H_1 : \mu \neq 100$

Test Statistic formula: $z = \dfrac{\bar{x} - \mu}{\sigma / \sqrt{n}}$

Level of Significance: $\alpha = .05$

Our test is called a *two-tailed test* since we reject our null hypothesis if the sample mean is either too large or too small. Also, in a normal distribution, the extreme 5% lies either 1.96 standard deviations below the mean or 1.96 standard deviations above the mean. We call these values our *critical values.*

$H_0 : \mu = 100$

$H_1 : \mu \neq 100$

Test Statistic Formula: $z = \dfrac{\bar{x} - \mu}{\sigma / \sqrt{n}}$

Level of Significance: $\alpha = .05$

Critical Values: $z = \pm 1.96$

Our test is called a *two-tailed test* since we reject our null hypothesis if the sample mean is either too large or too small. Also, in a normal distribution, the extreme 5% lies either 1.96 standard deviations below the mean or 1.96 standard deviations above the mean. We call these values our *critical values* and we call the shaded region the *critical region.*
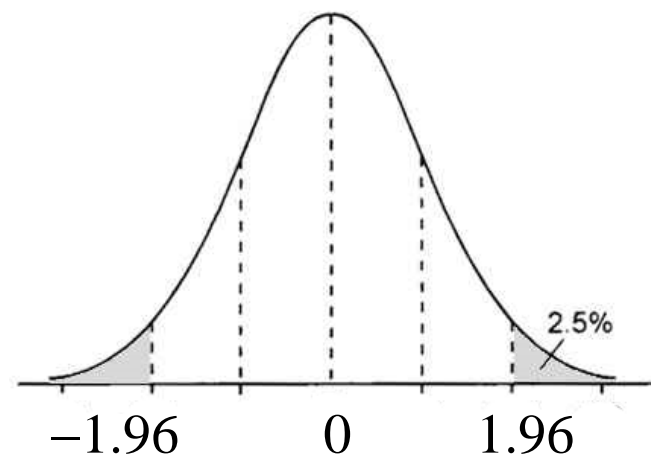
$H_0 : \mu = 100$

$H_1 : \mu \neq 100$

Test Statistic Formula: $z = \dfrac{\overline{x} - \mu}{\sigma / \sqrt{n}}$

2.5%

-1.96 0   1.96

Level of Significance: $\alpha = .05$

Critical Values: $z = \pm 1.96$

Suppose in our sample of 100 college students, the average IQ is 103. We now use our *test statistic formula* to convert this value to a z score. We'll call this our *calculated value.*

$H_0 : \mu = 100$

$H_1 : \mu \neq 100$

Test Statistic Formula: $z = \dfrac{\bar{x} - \mu}{\sigma / \sqrt{n}}$

Level of Significance: $\alpha = .05$

Critical Values: $z = \pm 1.96$

Calculated Value: $z = \dfrac{103 - 100}{15 / \sqrt{100}} = 2$

Our calculated value falls within the critical region, and so we will reject the null hypothesis.

$H_0 : \mu = 100$

$H_1 : \mu \neq 100$

Test Statistic Formula: $z = \dfrac{\bar{x} - \mu}{\sigma / \sqrt{n}}$

Level of Significance: $\alpha = .05$

Critical Values: $z = \pm 1.96$

Calculated Value: $z = \dfrac{103 - 100}{15 / \sqrt{100}} = 2$

Decision: Reject H_0

A variation in this method would have been to have suspected at the start that college students are smarter than the average person. In this case, our alternative hypothesis would have been that *mu is greater than 100.* We call this an example of a *one-tail test.*
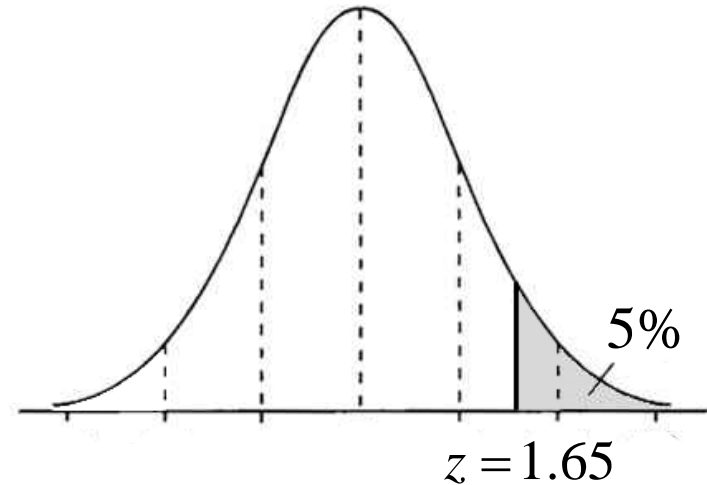
$H_0 : \mu = 100$ (or $\mu \leq 100$)

$H_1 : \mu > 100$

Test Statistic Formula: $z = \dfrac{\bar{x} - \mu}{\sigma / \sqrt{n}}$

Level of Significance: $\alpha = .05$

Critical Value: $z = 1.65$

Calculated Value: $z = \dfrac{103 - 100}{15 / \sqrt{100}} = 2$

Decision: Reject H_0

5%

$z = 1.65$

We now illustrate the *P-value method* for hypothesis testing. [*P-value = probability value*]

We now illustrate the *P-value method* for hypothesis testing. [*P-value = probability value*]

This method is similar to what we just did, but it is easier for modern technology to calculate directly the probability of getting whatever result we have obtained. We then compare this to our specified level of significance.

We now illustrate the *p-value method* for hypothesis testing. [*p-value = probability value*]

This method is similar to what we just did, but it is easier for modern technology to calculate directly the probability of getting whatever result we have obtained. We then compare this to our specified level of significance.

Here's how we can use our TI calculator to solve the problem we just did.

Suppose a sample of 100 college students has an average IQ of 103, and we hypothesize that they come from a population whose true mean is 100. Is our result significant at the .05 level?

Suppose a sample of 100 college students has an average IQ of 103, and we hypothesize that they come from a population whose true mean is 100. Is our result significant at the .05 level?

```
EDIT CALC TESTS
1:Edit…
2:SortA(
3:SortD(
4:ClrList
5:SetUpEditor
```

```
EDIT CALC TESTS
1:Z-Test…
2:T-Test…
3:2-SampZTest…
4:2-SampTTest…
5:1-PropZTest…
6:2-PropZTest…
7↓ZInterval…
```

```
Z-Test
Inpt:Data Stats
μ₀:0
σ:0
x̄:0
n:0
μ:≠μ₀ <μ₀ >μ₀
Calculate Draw
```

Suppose a sample of 100 college students has an average IQ of 103, and we hypothesize that they come from a population whose true mean is 100. Is our result significant at the .05 level?

```
EDIT CALC TESTS
1:Edit…
2:SortA(
3:SortD(
4:ClrList
5:SetUpEditor
```

```
EDIT CALC TESTS
1:Z-Test…
2:T-Test…
3:2-SampZTest…
4:2-SampTTest…
5:1-PropZTest…
6:2-PropZTest…
7↓ZInterval…
```
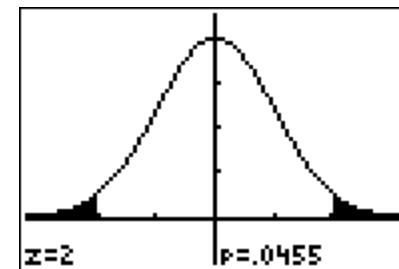
```
Z-Test
 Inpt:Data Stats
 μ0:0
 σ:0
 x̄:0
 n:0
 μ:≠μ0 <μ0 >μ0
Calculate Draw
```

```
Z-Test
 Inpt:Data Stats
 μ0:100
 σ:15
 x̄:103
 n:100
 μ:≠μ0 <μ0 >μ0
Calculate Draw
```

```
Z-Test
 μ≠100
 z=2
 p=.045500124
 x̄=103
 n=100
```


z=2   p=.0455

Suppose a sample of 100 college students has an average IQ of 103, and we hypothesize that they come from a population whose true mean is 100. Is our result significant at the .05 level?

$H_0 : \mu = 100$

$H_1 : \mu \neq 100$

Test Statistic Formula: $z = \dfrac{\overline{x} - \mu}{\sigma / \sqrt{n}}$
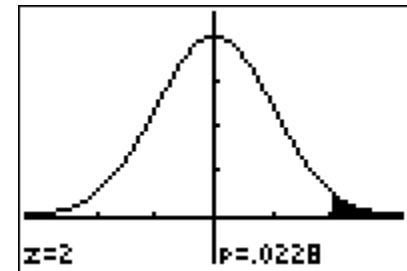
Level of Significance: $\alpha = .05$

P-Value: $P = 0.0455$

Decision: Reject H_0

Suppose a sample of 100 college students has an average IQ of 103, and we hypothesize that they come from a population whose true mean is 100. At the .05 level, is our result significantly better than 100?

```
Z-Test
 Inpt:Data Stats
 μ₀:100
 σ:15
 x̄:103
 n:100
 μ:≠μ₀ <μ₀ >μ₀
Calculate Draw
```

```
Z-Test
 μ>100
 z=2
 p=.022750062
 x̄=103
 n=100
```

```
z=2        p=.0228
```

Suppose a sample of 100 college students has an average IQ of 103, and we hypothesize that they come from a population whose true mean is 100. At the .05 level, is our result significantly better than 100?

$H_0 : \mu = 100$ (or $\mu \leq 100$)

$H_1 : \mu > 100$

Test Statistic Formula: $z = \dfrac{\bar{x} - \mu}{\sigma / \sqrt{n}}$

Level of Significance: $\alpha = .05$

P-Value: $P = .0228$

Decision: Reject H_0

Notice that if we had used a different level of significance such as .01, then we would not have arrived at the same conclusion.

Suppose a sample of 100 college students has an average IQ of 103, and we hypothesize that they come from a population whose true mean is 100. Is our result significant at the .01 level?

$H_0 : \mu = 100$

$H_1 : \mu \neq 100$

Test Statistic Formula: $z = \dfrac{\bar{x} - \mu}{\sigma / \sqrt{n}}$

Level of Significance: $\alpha = .01$

P-Value: $P = 0.0455$

Decision: Accept H_0 (Fail to reject H_0)

In all of the work we've done so far, there is the possibility of having made an error due to chance.

In all of the work we've done so far, there is the possibility of having made an error due to chance.

For example, suppose our students really do come from a population with a mean IQ of 100.

In all of the work we've done so far, there is the possibility of having made an error due to chance.

For example, suppose our students really do come from a population with a mean IQ of 100.

Then at the .05 level of significance, we would still get an extreme result 5% of the time as the result of chance.

In all of the work we've done so far, there is the possibility of having made an error due to chance.

For example, suppose our students really do come from a population with a mean IQ of 100.

Then at the .05 level of significance, we would still get an extreme result 5% of the time as the result of chance.

That means that 5% of the time we would erroneously reject the null hypothesis.

In all of the work we've done so far, there is the possibility of having made an error due to chance.

For example, suppose our students really do come from a population with a mean IQ of 100.

Then at the .05 level of significance, we would still get an extreme result 5% of the time as the result of chance.

That means that 5% of the time we would erroneously reject the null hypothesis.

We call this a *Type I error*, and it's probability is the same as your level of significance.

Here is the bottom line.

Table 8-1 **Type I and Type II Errors**

		True State of Nature	
		The null hypothesis is true	The null hypothesis is false
Decision	We decide to reject the null hypothesis	**Type I error** (rejecting a true null hypothesis) $P(\text{type I error}) = \alpha$	Correct decision
	We fail to reject the null hypothesis	Correct decision	**Type II error** (failing to reject a false null hypothesis) $P(\text{type II error}) = \beta$

- A *Type I error* occurs when we reject a true null hypothesis

Table 8-1   Type I and Type II Errors

		True State of Nature	
		The null hypothesis is true	The null hypothesis is false
Decision	We decide to reject the null hypothesis	**Type I error** (rejecting a true null hypothesis) $P(\text{type I error}) = \alpha$	Correct decision
	We fail to reject the null hypothesis	Correct decision	**Type II error** (failing to reject a false null hypothesis) $P(\text{type II error}) = \beta$

- A *Type I error* occurs when we reject a true null hypothesis
- A *Type II error* occurs when we accept (fail to reject) a false null hypothesis

Table 8-1   Type I and Type II Errors

		True State of Nature	
		The null hypothesis is true	The null hypothesis is false
Decision	We decide to reject the null hypothesis	Type I error (rejecting a true null hypothesis) $P(\text{type I error}) = \alpha$	Correct decision
	We fail to reject the null hypothesis	Correct decision	Type II error (failing to reject a false null hypothesis) $P(\text{type II error}) = \beta$

- $P(\text{Type I Error}) = \alpha$

- $P(\text{Type II Error}) = \beta$

- Power of the test $= 1 - \beta$

Table 8-1   Type I and Type II Errors

		True State of Nature	
		The null hypothesis is true	The null hypothesis is false
Decision	We decide to reject the null hypothesis	**Type I error** (rejecting a true null hypothesis) $P(\text{type I error}) = \alpha$	Correct decision
	We fail to reject the null hypothesis	Correct decision	**Type II error** (failing to reject a false null hypothesis) $P(\text{type II error}) = \beta$