STA218
Introduction to Hypothesis Testing

Al Nosedal.
University of Toronto.
Fall 2018

October 24, 2018

Let's say that one of you is invited to this popular show. As you probably know, you have to answer a series of multiple choice questions and there are four possible answers to each question. Perhaps you also have seen that if you don't know the answer to a question you could either "jump the question" or you could "ask the audience".
Suppose that you run into a question for which you don't know the answer with certainty and you decide to "ask the audience". Let's say that you initially believe that the right answer is **A**. Then you ask the audience and only 2% of the audience shares your opinion. What would you do? Change your initial answer or keep it?

Sweetening colas

Diet colas use artificial sweeteners to avoid sugar. These sweeteners gradually lose their sweetness over time. Manufacturers therefore test new colas for loss of sweetness before marketing them. Trained tasters sip the cola along with drinks of standard sweetness and score the cola on a "sweetness score " of 1 to 10. The cola is then stored for a month at high temperature to imitate the effect of four months' storage at room temperature. Each taster scores the cola again after storage. This is a matched pairs experiment. Our data are the differences (score before storage minus score after storage) in the tasters' scores. The bigger these differences, the bigger the loss of sweetness.

Sweetening colas (cont.)

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.
The following are the sweetness losses for a new cola, as measured by 10 trained tasters: 2.0 0.4 0.7 2.0 -0.4 2.2 -1.3 1.2 1.1 2.3.
Are these data good evidence that the cola lost sweetness in storage?

Solution

$\mu=$ mean sweetness loss for the population of **all** tasters.

1. State hypotheses. $H_0 : \mu = 0$ vs $H_a : \mu > 0$
2. Test statistic. $z_* = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \frac{1.02-0}{1/\sqrt{10}} = 3.23$
3. P-value. $P(Z > z_*) = P(Z > 3.23) = 0.0006$
4. Conclusion. We would very rarely observe a sample sweetness loss as large as 1.02 if H_0 were true. The small P-value provides strong evidence against H_0 and in favor of the alternative $H_a : \mu > 0$, i.e., it gives good evidence that the mean sweetness loss is not 0, but positive.

Executives' blood pressures

The National Center for Health Statistics reports that the systolic blood pressure for males 35 to 44 years of age has mean 128 and standard deviation 15. The medical director of a large company looks at the medical records of 72 executives in this age group and finds that the mean systolic blood pressure in this sample is $\bar{x} = 126.07$. Is this evidence that the company's executives have a different mean blood pressure from the general population? Suppose we know that executives' blood pressures follow a Normal distribution with standard deviation $\sigma = 15$.

Solution

$\mu=$ mean of the **executive population**.

1. State hypotheses. $H_0 : \mu = 128$ vs $H_a : \mu \neq 128$

2. Test statistic. $z_* = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \frac{126.07-128}{15/\sqrt{72}} = -1.09$

3. P-value. $2P(Z > |z_*|) = 2P(Z > |-1.09|) = 2P(Z > 1.09) = 2(1 - 0.8621) = 0.2758$

4. Conclusion. More than 27% of the time, a simple random sample of size 72 from the general male population would have a mean blood pressure at least as far from 128 as that of the executive sample. The observed $\bar{x} = 126.07$ is therefore not good evidence that executives differ from other men.

There are four steps in carrying out a significance test:

1. State the hypotheses.
2. Calculate the test statistic.
3. Find the P-value.
4. State your conclusion in the context of your specific setting.

Once you have stated your hypotheses and identified the proper test, you or your computer can do Steps 2 and 3 by following a recipe. Here is the recipe for the test we have used in our examples.

Z test for a population mean μ

Draw a simple random sample 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, $H_0 : \mu = \mu_0$ calculate the **one-sample z statistic**

$$z_* = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

In terms of a variable Z having the standard Normal distribution, the P-value for a test of H_0 against
$H_a : \mu > \mu_0$ is $P(Z > z_*)$
$H_a : \mu < \mu_0$ is $P(Z < z_*)$
$H_a : \mu \neq \mu_0$ is $2P(Z > |z_*|)$

Example 1

Consider the following hypothesis test:

$H_0 : \mu = 20$

$H_a : \mu < 20$

A sample of 50 provided a sample mean of 19.4. The population standard deviation is 2.

a. Compute the value of the test statistic.

b. What is the p-value?

c. Using $\alpha = 0.05$, what is your conclusion?

Solution

a. Test statistic.

$z_* = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \frac{19.4-20}{2/\sqrt{50}} = -2.1213$

b. P-value.

$P(Z < z_*) = P(Z < -2.1213) = 0.0169$

c. Conclusion.

Since P-value $= 0.0169 < \alpha = 0.05$, we reject $H_0 : \mu = 20$. We conclude that $\mu < 20$.

Example 2

Consider the following hypothesis test:

$H_0 : \mu = 25$

$H_a : \mu > 25$

A sample of 40 provided a sample mean of 26.4. The population standard deviation is 6.

a. Compute the value of the test statistic.

b. What is the p-value?

c. Using $\alpha = 0.01$, what is your conclusion?

Solution

a. Test statistic.

$z_* = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{26.4 - 25}{6/\sqrt{40}} = 1.4757$

b. P-value.

$P(Z > z_*) = P(Z > 1.4757) = 0.0700$

c. Conclusion.

Since P-value $= 0.0700 > \alpha = 0.01$, we CAN'T reject $H_0 : \mu = 25$. We conclude that we don't have enough evidence to claim that $\mu > 25$. (Some of us would say that we accept that $\mu = 25$).

Example 3

Consider the following hypothesis test:
$H_0 : \mu = 15$
$H_a : \mu \neq 15$
A sample of 50 provided a sample mean of 14.15. The population standard deviation is 3.

a. Compute the value of the test statistic.
b. What is the p-value?
c. Using $\alpha = 0.05$, what is your conclusion?

Solution

a. Test statistic.
$z_* = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \frac{14.15-15}{3/\sqrt{50}} = -2.0034$
b. P-value.
$2P(Z > |z_*|) = 2P(Z > |-2.0034|) = 2P(Z > 2.0034) = 0.0451$
c. Conclusion.
Since P-value $= 0.0451 < \alpha = 0.05$, we reject $H_0 : \mu = 15$. We conclude that $\mu \neq 15$.

CONFIDENCE INTERVALS AND TWO-SIDED TESTS.

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

Example 3 (again...)

The 95% confidence interval for μ in example 3 is:

$$\bar{x} \pm z_*(\frac{\sigma}{\sqrt{n}})$$

$$14.15 \pm 1.96(\frac{3}{\sqrt{50}})$$

$$(13.3184, 14.9815)$$

The hypothesized value $\mu_0 = 15$ in example 3 falls outside this confidence interval, so we reject $H_0 : \mu = 15$.