STA258H5

Al Nosedal
and Alison Weir

Winter 2017

INTRODUCTION TO HYPOTHESIS TESTING

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.

Simulation

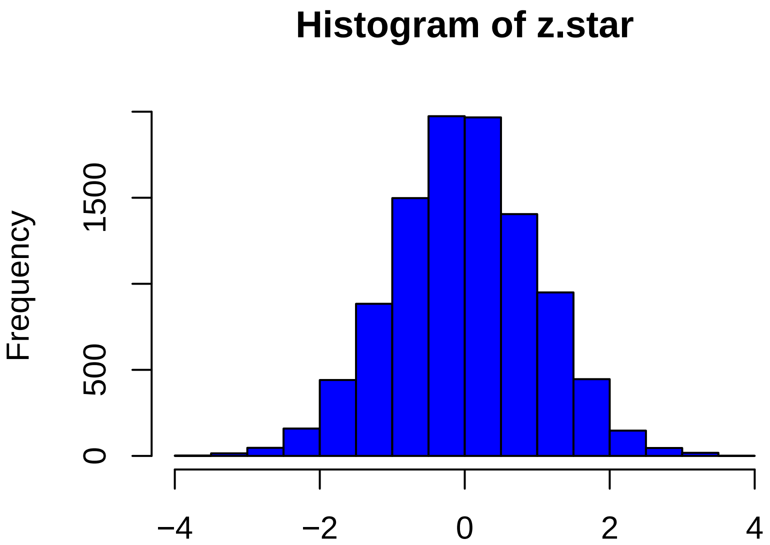
```
# n = sample size;
n<-10;
mu.zero<-0;
sigma<-1;
sigma.xbar<-sigma/sqrt(n);
# x bar = sample mean with 10 obs;
x.bar<-rnorm(1,mean=mu.zero,sd=sigma.xbar);
x.bar;

## [1] 0.398947

# z.star = test statistic;
z.star<-(x.bar-mu.zero)/sigma.xbar;
z.star;

## [1] 1.261581
```

Another Simulation

```
# n = sample size;
n<-10;
mu.zero<-0;
sigma<-1;
sigma.xbar<-sigma/sqrt(n);
# x bar = sample mean with 10 obs;
x.bar<-rnorm(1,mean=mu.zero,sd=sigma.xbar);
x.bar;

## [1] 0.1236054

# z.star = test statistic;
z.star<-(x.bar-mu.zero)/sigma.xbar;
z.star;

## [1] 0.3908747
```

10000 Simulations

```r
n<-10;
mu.zero<-0;
sigma<-1;
sigma.xbar<-sigma/sqrt(n);
# x bar = sample mean with 10 obs;
# m = number of simulations;
m<-10000;
x.bar<-rnorm(m,mean=mu.zero,sd=sigma.xbar);
x.bar;
# z.star = test statistic;
z.star<-(x.bar-mu.zero)/sigma.xbar;
hist(z.star,xlab="differences",col="blue");
```


Histogram of z.star
Frequency
0
500
1500
−4
−2
0
2
4

P-value

```
## P-value

p_value<-length(z.star[z.star>3.23])/m;

p_value

## [1] 4e-04
```

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.

Tests for a population mean

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 $\mu > 25$.

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 $\mu \neq 15$.

Tests from confidence intervals

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 $H_0 : \mu = 15$.

Hypotheses Tests for a Proportion

To test the hypothesis $H_0 : p = p_0$, compute the z_* statistic,
$z_* = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$
In terms of a variable Z having the standard Normal distribution, the approximate P-value for a test of H_0 against
$H_a : p > p_0$: is : $P(Z > z_*)$
$H_a : p < p_0$: is : $P(Z < z_*)$
$H_a : p \neq p_0$: is : $2P(Z > |z_*|)$

Example

Consider the following hypothesis test:

$H_0 : p = 0.75$

$H_a : p < 0.75$

A sample of 300 items was selected. Compute the p-value and state your conclusion for each of the following sample results. Use $\alpha = 0.05$.

a. $\hat{p} = 0.68$

b. $\hat{p} = 0.72$

c. $\hat{p} = 0.70$

d. $\hat{p} = 0.77$

Solution a.

$z_* = \frac{\hat{p}-p_0}{\sqrt{p_0(1-p_0)/n}} = \frac{0.68-0.75}{\sqrt{0.75(1-0.75)/300}} = -2.80$

Using Normal table, P-value $= P(Z < z_*) = P(Z < -2.80) = 0.0026$

P-value$< \alpha = 0.05$, reject H_0.

Solution b.

$z_* = \frac{\hat{p}-p_0}{\sqrt{p_0(1-p_0)/n}} = \frac{0.72-0.75}{\sqrt{0.75(1-0.75)/300}} = -1.20$

Using Normal table, P-value $= P(Z < z_*) = P(Z < -1.20) = 0.1151$

P-value$> \alpha = 0.05$, do not reject H_0.

Solution c.

$$z_* = \frac{\hat{p}-p_0}{\sqrt{p_0(1-p_0)/n}} = \frac{0.70-0.75}{\sqrt{0.75(1-0.75)/300}} = -2.00$$

Using Normal table, P-value $= P(Z < z_*) = P(Z < -2.00) = 0.0228$

P-value$< \alpha = 0.05$, reject H_0.

Example

Consider the following hypothesis test:

$H_0 : p = 0.20$

$H_a : p \neq 0.20$

A sample of 400 provided a sample proportion $\hat{p} = 0.175$.

a. Compute the value of the test statistic.

b. What is the p-value?

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

d. What is the rejection rule using the critical value? What is your conclusion?

Solution

a. $z_* = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.175-0.20}{\sqrt{\frac{(0.20)(0.80)}{400}}} = -1.25$
b. Using Normal table, P-value $=$
$2P(Z > |z_*|) = 2P(Z > |-1.25|) = 2P(Z > 1.25) = 2(0.1056) = 0.2112$
c. P-value $> \alpha = 0.05$, we CAN'T reject H_0.

Problem

A study found that, in 2005, 12.5% of U.S. workers belonged to unions. Suppose a sample of 400 U.S. workers is collected in 2006 to determine whether union efforts to organize have increased union membership.
a. Formulate the hypotheses that can be used to determine whether union membership increased in 2006.
b. If the sample results show that 52 of the workers belonged to unions, what is the p-value for your hypothesis test?
c. At $\alpha = 0.05$, what is your conclusion?

Solution

a. $H_0 : p = 0.125$ vs $H_a : p > 0.125$

b. $\hat{p} = \frac{52}{400} = 0.13$

$z_* = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.13 - 0.125}{\sqrt{\frac{(0.125)(0.875)}{400}}} = 0.30$

Using Normal table, P-value =

$P(Z > z_*) = P(Z > 0.30) = 1 - 0.6179 = 0.3821$

c. P-value $= > 0.05$, do not reject H_0. We cannot conclude that there has been an increase in union membership.

R Code

```
prop.test(52,400,p=0.125,alternative="greater",
correct=FALSE);

##
##  1-sample proportions test without continuity correction
##
## data:  52 out of 400, null probability 0.125
## X-squared = 0.0914, df = 1, p-value = 0.3812
## alternative hypothesis: true p is greater than 0.125
## 95 percent confidence interval:
##  0.1048085 1.0000000
## sample estimates:
##    p
## 0.13
```

Problem

A study by Consumer Reports showed that 64% of supermarket shoppers believe supermarket brands to be as good as national name brands. To investigate whether this result applies to its own product, the manufacturer of a national name-brand ketchup asked a sample of shoppers whether they believed that supermarket ketchup was as good as the national brand ketchup.

a. Formulate the hypotheses that could be used to determine whether the percentage of supermarket shoppers who believe that the supermarket ketchup was as good as the national brand ketchup differed from 64%.
b. If a sample of 100 shoppers showed 52 stating that the supermarket brand was as good as the national brand, what is the p-value?
c. At $\alpha = 0.05$, what is your conclusion?

Solution

a. $H_0 : p = 0.64$ vs $H_a : p \neq 0.64$
b. $\hat{p} = \frac{52}{100} = 0.52$
$z_* = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.52-0.64}{\sqrt{\frac{(0.64)(0.36)}{100}}} = -2.50$
Using Normal table, P-value =
$2P(Z > |z_*|) = 2P(Z > |-2.50|) = 2P(Z > 2.50) = 2(0.0062) = 0.0124$
c. P-value $= < 0.05$, reject H_0. Proportion differs from the reported 0.64.

R Code

```
prop.test(52,100,p=0.64,alternative="two.sided",
correct=FALSE);

##
##  1-sample proportions test without continuity correction
##
## data:  52 out of 100, null probability 0.64
## X-squared = 6.25, df = 1, p-value = 0.01242
## alternative hypothesis: true p is not equal to 0.64
## 95 percent confidence interval:
##  0.4231658 0.6153545
## sample estimates:
##    p
## 0.52
```

Problem

The National Center for Health Statistics released a report that stated 70% of adults do not exercise regularly. A researcher decided to conduct a study to see whether the claim made by the National Center for Health Statistics differed on a state-by-state basis.
a. State the null and alternative hypotheses assuming the intent of the researcher is to identify states that differ from 70% reported by the National Center for Health Statistics.
b. At $\alpha = 0.05$, what is the research conclusion for the following states:
Wisconsin: 252 of 350 adults did not exercise regularly.
California: 189 of 300 adults did not exercise
regularly.

Solution (Wisconsin)

a. $H_0 : p = 0.70$ vs $H_a : p \neq 0.70$

b. Wisconsin $\hat{p} = \frac{252}{350} = 0.72$

$z_* = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.72 - 0.70}{\sqrt{\frac{(0.70)(0.30)}{350}}} = 0.82$

Using Normal table, P-value =

$2P(Z > |z_*|) = 2P(Z > |0.82|) = 2P(Z > 0.82) = 2(0.2061) = 0.4122$

c. P-value > 0.05, we don't have enough evidence to reject H_0. There is not enough evidence against the claim made by the National Center for Health Statistics.

Power

The probability that a fixed level α significance test will reject H_0 when a particular alternative value of the parameter is true is called the **power** of the test against that alternative.

Sweetening colas: power

The cola maker of our example determines that a sweetness loss is too large to accept if the mean response for all tasters is $\mu = 1.1$. Will a 5% significance test of the hypotheses

$H_0 : \mu = 0$

$H_a : \mu > 0$

based on a sample of 10 tasters usually detect a change this great?

We want the power of the test against the alternative $\mu = 1.1$. This is the probability that the test rejects H_0 when $\mu = 1.1$ is true.

Step 1. Write the rule for rejecting H_0 in terms of $\bar{x}$.
We know that $\sigma = 1$, so the z test rejects H_0 at the $\alpha = 0.05$ level when

$$z = \frac{\bar{x} - 0}{1/\sqrt{10}} \geq 1.645$$

This is the same as

$$\bar{x} \geq 0 + 1.645 \frac{1}{\sqrt{10}}$$

or

$$\text{Reject } H_0 \text{ when } \bar{x} \geq 0.520$$

Step 2.

Step 2. The power is the probability of this event under the condition that the alternative $\mu = 1.1$ is true.

To calculate this probability, standardize $\bar{x}$ using $\mu = 1.1$.

$$\begin{aligned} \text{power } &= P(\bar{x} \geq 0.520 \text{ when } \mu = 1.1) \\ &= P\left(\frac{\bar{x}-1.1}{1/\sqrt{10}} \geq \frac{0.520-1.1}{1/\sqrt{10}}\right) \\ &= P(Z \geq -1.83) = 1 - 0.0336 = 0.9664 \end{aligned}$$

The cola maker is satisfied. The test will declare that the cola loses sweetness only 5% of the time when it actually does not ($\alpha = 0.05$) and 97% of the time when the true mean sweetness loss is 1.1 (power $= 0.97$).