Hypothesis Testing

Jerome Dumortier

17 August 2023

Example of Hypothesis Testing

Weight and weight distribution are important parameters for airplanes and the European Aviation Safety Agency (EASA) publishes passenger standard weights:

- Female passenger: 66.5 kilograms (146.6 lbs)
- Male passenger: 84.6 kilograms (186.5 lbs)
- Children under age 12: 30.7 kilograms (67.7 lbs)

In 2017, the airline Finnair asks passengers to step on a scale before boarding flights to better understand their payload. At the time of the BBC article Why Finnair wants to put passengers on the weighing scales, they had collected data from 180 volunteers.

Introduction to Hypothesis Testing

Hypothesis:

- A statement about a parameter taking on a particular value. This is formulated as the null hypothesis H_0.

Hypothesis test:

- A procedure to verify the hypothesis based on a random sample of size N. We never accept H_0 but **fail to reject** H_0. This is similar to guilty versus not guilty in courts. The opposite of the null hypothesis is labeled H_a (sometimes H_1) as the alternative hypothesis.

Overview of Hypothesis Tests

One-sample (or one-group) tests:

- Population mean with unknown variance
- Population proportion

Two-sample (or two-group) tests:

- Population proportions
- Population means
  - Equal versus unequal variance
- Paired difference test

Note:

- Statistics textbooks often include “population mean with known variance.” This is a highly unlikely case and thus, it is skipped for this section.

Hypothesis Testing Procedure

Steps

1. Formulating the null hypothesis H_0 stating that the parameter takes a particular value:
   - One-sided test: H_0: $\mu \geq \mu_0$ or $\mu \leq \mu_0$
   - Two-sided test: H_0: $\mu = \mu_0$
2. Setting the significance level α, e.g., 1%, 5%, or 10%.
3. Test statistic: Value based on the sample used to **reject** or **fail to reject** the null hypothesis.
4. Critical value and p-value:
   - Critical value represents the border point between rejecting and failing to reject H_0.
   - p-Value: Probability of observing the parameter given the null hypothesis. Small p-values represent evidence against H_0.

Note that equality is always part of H_0, i.e., $=$, $\leq$, or $\geq$.

Decisions and Errors in Hypothesis Testing

Null Hypothesis	Fail to reject H_0	Reject H_0
H_0 is true	Correct	Type I Error
H_0 is false	Type II Error	Correct

Type I Error:

- Probability of rejecting H_0 when it is true.
- Also known as the significance level of a test denoted with α.

Type II Error:

- Probability of failing to reject H_0 when it is false.

Interpretation of the p-Value

Each statistical software provides a p-value:

- Lowest level of significance at which the null hypothesis can be rejected.
- Represents the probability of observing the sample given that the hypothesis is true. The lower the p-value the more unlikely is the hypothesis.
- The null hypothesis H_0 is rejected if the p-value is smaller than the significance level.

The smaller the p-value, the stronger the evidence against H_0 being true. This is true for any type of hypothesis test.

Mean with Unknown Variance: Test Statistic

Unknown variance requires the use of the t-distribution given the following test statistic:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

where

- $\bar{x}$ is the sample mean
- μ_0 is the hypothesized mean
- s is the sample standard deviation
- n is the sample size

Two-Sided Test for Water Pressure

You work as a engineer for the local water company and you are concerned about the daily water pressure in the city's pipes. Too much pressure may burst pipes whereas too little pressure causes customer complaints. Regulation requires a water pressure of 50 psi. You collect a sample of 30 daily water pressures. The sample mean and the sample standard deviation are $\bar{x} = 51.788$ and $s = 3.389$.

Two-Sided Test for Water Pressure: Step-by-Step

Formulating the null hypothesis H_0 and the alternative hypothesis H_a

- H_0: $\mu = 50$
- H_a: $\mu \neq 50$

Setting the significance level:

- $\alpha = 0.05$ or 5%

Calculating the test statistic

$$t = \frac{51.788 - 50}{3.389/\sqrt{30}} = 2.8895$$

The critical value for $t_{0.05/2,29}$ is 2.045 (`qt(c(0.025,0.975),df=29)`). Thus, we reject H_0 and there is evidence that the water pressure is different from 50 psi. To calculate the p-value in R: `(1-pt(tstatistic,29))*2`

Two-Sided Test for Water Pressure: R

Note that the data in `waterpressure` is randomly generated and value differ from previous slide.

```r
t.test(waterpressure$psi,mu=50)
```

```
##
##  One Sample t-test
##
## data:  waterpressure$psi
## t = 10.637, df = 29, p-value = 1.595e-11
## alternative hypothesis: true mean is not equal to 50
## 95 percent confidence interval:
##  51.81673 52.68163
## sample estimates:
## mean of x
##  52.24918
```

One-Sided Test for MPA Scores

Consider the scores from a graduate MPA class which has eighteen students in `mpa`.

- Sample mean: $\bar{x} = 69$
- Sample standard deviation: $s = 21.14933$.

We are interested in the null hypothesis H_0: $\mu \geq 80$. We can compute the t-statistic as follows:

$$t = \frac{69 - 80}{21.14933/\sqrt{18}} = -2.206644$$

What is the critical value in this case?

One-Sided Test for MPA Scores

```
t.test(mpa$scores,mu=80,alternative="less")
```

```
##
##  One Sample t-test
##
## data:  mpa$scores
## t = -2.2066, df = 17, p-value = 0.02069
## alternative hypothesis: true mean is less than 80
## 95 percent confidence interval:
##      -Inf 77.67184
## sample estimates:
## mean of x
##        69
```

Hypothesis Test about Population Proportion

Test statistic for a proportion

$$z = \frac{\bar{p} - p_0}{\sqrt{p_0 \cdot (1 - p_0)/n}}$$

where p_0 is the hypothesized population proportion. Recall that the sampling distribution of a sample proportion has mean p and standard error $\sqrt{p \cdot (1 - p)/n}$.
Example:

- Use of Instagram by GSS respondents
- $\bar{p} = 0.3097$ and $n = 1366$

Let us calculate two possible hypothesis tests manually and with R.

Hypothesis Test about Population Proportion

Example: Assume $\bar{p} = 0.3097$ and $n = 1366$. Under H_0: $p_0 = 0.33$, the test statistic is

$$z = \frac{0.3097 - 0.333}{\sqrt{\frac{0.333(1-0.333)}{1366}}} = -1.8914$$

For a two-sided hypothesis test at the $\alpha = 0.05$ level, we fail to reject the hypothesis because $1.8914 < 1.96$.

Hypothesis Test about Population Proportion: Social Media

```r
gss$instagram    = NA
gss$instagram[which(gss$instagrm=="no")]     = 0
gss$instagram[which(gss$instagrm=="yes")]    = 1
t.test(gss$instagram,mu=0.33,na.action=na.omit)
```

```
##
##  One Sample t-test
##
## data:  gss$instagram
## t = -1.7392, df = 1371, p-value = 0.08222
## alternative hypothesis: true mean is not equal to 0.33
## 95 percent confidence interval:
##  0.2838431 0.3327750
## sample estimates:
## mean of x
```

Hypothesis Test about Population Proportion: Social Media

```
t.test(gss$instagram,mu=0.33,
       alternative="less",na.action=na.omit)
```

```
##
##  One Sample t-test
##
## data:  gss$instagram
## t = -1.7392, df = 1371, p-value = 0.04111
## alternative hypothesis: true mean is less than 0.33
## 95 percent confidence interval:
##       -Inf 0.3288373
## sample estimates:
## mean of x
##  0.308309
```

Two Sample Test: Overview

Difference between two mean:

$$\bar{x}_1 - \bar{x}_2$$

Means of two dependent populations:

- Assumption of equal variance, i.e., $\sigma_1^2 = \sigma_2^2$
- Example: Pre- and post-test
- Pooled-Variance t-test: One estimate of unknown σ^2, i.e., s_p

Means of two independent populations:

- Assumption of unequal variance, i.e., $\sigma_1^2 \neq \sigma_2^2$
- Samples from two different populations
- Separate-Variance t-test: Two estimates for unknown σ_1^2 and σ_2^2.

Two Population Means: Equal Variance

Pooled variance:

$$s_p^2 = \frac{(n_1 - 1) \cdot s_1^2 + (n_2 - 1) \cdot s_2^2}{n_1 + n_2 - 2}$$

Test statistic:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s_p^2 \cdot \left(\frac{1}{n_1} + \frac{1}{n_1}\right)}}$$

Confidence interval for $\mu_1 - \mu_2$:

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2}\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_1}\right)}$$

Two Population Means: Unequal Variance

Test statistic:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{(s_1^2/n_1 + s_2^2/n_2)}}$$

Degrees of Freedom:

$$d_1 = \left(s_1^2/n_1 + s_2^2/n_2\right)^2$$

$$d_2 = \left(s_1^2/n_1\right)^2$$

$$d_3 = \left(s_2^2/n_2\right)^2$$

Then the degrees of freedom are $d.f. = d_1/(d_2 + d_3)$.

Two Sample Test (Equal and Unequal Variance): R

Separating schools into three different income groups

```
oh   = merge(ohioincome,ohioscore,by="irn")
oh_s = subset(oh,enrollment<1000)
oh_l = subset(oh,enrollment>3000)
oh_m = subset(oh,enrollment>1000 & enrollment <3000)
t.test(oh_s$score,oh_l$score,var.equal = TRUE)
t.test(oh_s$score,oh_l$score,var.equal = FALSE)
```

Paired Difference Test: Related Populations

Example: Textbook prices:

- Online vs. bookstore because prices exist for both purchase options.

Difference between paired (!) values:

$$D_i = x_{1,i} - x_{2,i}$$

Elimination of variation among subjects. Point estimate for paired difference

$$\bar{D} = \frac{1}{n} \sum_{i=1}^{n} D_i$$

Sample standard deviation

$$S_d = \sqrt{\frac{\sum_{i=1}^{n} (D_i - \bar{D})^2}{n-1}}$$

Paired Difference Test: Related Populations

Test statistic:

$$t_p = \frac{\bar{D} - \mu_D}{S_d / \sqrt{n}}$$

Confidence interval:

$$\bar{D} \pm t_{\alpha/2} \frac{S_D}{\sqrt{n}}$$

t_p has $n-1$ degrees of freedom

Textbook Example

Book	Online	Bookstore	Difference
History 1	10.2	11.4	-1.2
History 2	18.95	19	-0.05
Economics 1	184.53	200.75	-16.22
Business 1	236.75	247.2	-10.45
Business 2	67.41	71.25	-3.48

Note that $\sum D_i = -31.76$, $\bar{D} = -6.352$, and $s_D = 6.833$.

Textbook Example

```
online       = c(10.20,18.95,184.53,236.75,67.41)
bookstore    = c(11.40,19,200.75,247.20,71.25)
t.test(online,bookstore,paired=TRUE)
```

Hypothesis Tests for Population Proportions

Confidence interval and hypothesis test for difference between two population proportions. Point estimate for difference:

$$p_1 - p_2$$

Pooled estimate for overall proportion

$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

Test statistic

$$z_p = \frac{(p_1 - p_2) - (\eta_1 - \eta_2)}{\bar{p}(1 - \bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

Hypothesis Tests for Population Proportions

Gun ownership among males and females from the General Social Survey

- Females: 495 no, 207 yes, $n_1 = 702$
- Males: 334 no, 231 yes, $n_2 = 565$

Is there a statically significant difference in gun ownership between women and men?

$$\bar{p} = (207 + 231)/(702 + 565) = 0.3457$$

$$p_1 = 207/702 = 0.2949$$

$$p_2 = 231/565 = 0.4088$$