Math 362: Mathematical Statistics II

Le Chen
le.chen@emory.edu
chenle02@gmail.com

Emory University
Atlanta, GA

Last updated on Spring 2021
Last compiled on January 15, 2023

2021 Spring

Chapter 7. Inference Based on The Normal Distribution

§ 7.1 Introduction

§ 7.2 Comparing $\frac{\overline{Y}-\mu}{\sigma/\sqrt{n}}$ and $\frac{\overline{Y}-\mu}{S/\sqrt{n}}$

§ 7.3 Deriving the Distribution of $\frac{\overline{Y}-\mu}{S/\sqrt{n}}$

§ 7.4 Drawing Inferences About μ

§ 7.5 Drawing Inferences About σ^2

Plan

Chapter 7. Inference Based on The Normal Distribution

§ 7.1 Introduction

§ 7.2 Comparing $\frac{\overline{Y}-\mu}{\sigma/\sqrt{n}}$ and $\frac{\overline{Y}-\mu}{S/\sqrt{n}}$

§ 7.3 Deriving the Distribution of $\frac{\overline{Y}-\mu}{S/\sqrt{n}}$

§ 7.4 Drawing Inferences About μ

§ 7.5 Drawing Inferences About σ^2

Let $Y_1, \cdots, Y_n$ be a random sample from $N(\mu, \sigma^2)$.

Question: Find a test statistic Λ in order to test $H_0 : \mu = \mu_0$ v.s. $H_1 : \mu \neq \mu_0$.

Case I. σ^2 is known: $\Lambda = \dfrac{\overline{Y} - \mu_0}{\sigma/\sqrt{n}}$

Case II. σ^2 is unknown: $\Lambda = ?$ $\quad \Lambda \stackrel{?}{=} \dfrac{\overline{Y} - \mu_0}{s/\sqrt{n}} \sim ?$

Let $Y_1, \cdots, Y_n$ be a random sample from $N(\mu, \sigma^2)$.

Question Find a test statistic Λ in order to test $H_0 : \mu = \mu_0$ v.s. $H_1 : \mu \neq \mu_0$.

Case I. σ^2 is known: $\Lambda = \dfrac{\overline{Y} - \mu_0}{\sigma/\sqrt{n}}$

Case II. σ^2 is unknown: $\Lambda = ?$ $\Lambda \overset{?}{=} \dfrac{\overline{Y} - \mu_0}{s/\sqrt{n}} \sim ?$

Let $Y_1, \cdots, Y_n$ be a random sample from $N(\mu, \sigma^2)$.

Question Find a test statistic Λ in order to test $H_0 : \mu = \mu_0$ v.s. $H_1 : \mu \neq \mu_0$.

Case I. σ^2 is known:
$$\Lambda = \frac{\overline{Y} - \mu_0}{\sigma/\sqrt{n}}$$

Case II. σ^2 is unknown:
$$\Lambda = ? \qquad \Lambda \stackrel{?}{=} \frac{\overline{Y} - \mu_0}{s/\sqrt{n}} \sim ?$$

Let $Y_1, \cdots, Y_n$ be a random sample from $N(\mu, \sigma^2)$.

Question Find a test statistic Λ in order to test $H_0 : \mu = \mu_0$ v.s. $H_1 : \mu \neq \mu_0$.

Case I. σ^2 is known:

$$\Lambda = \frac{\overline{Y} - \mu_0}{\sigma/\sqrt{n}}$$

Case II. σ^2 is unknown:

$$\Lambda = ? \qquad \Lambda \stackrel{?}{=} \frac{\overline{Y} - \mu_0}{s/\sqrt{n}} \sim ?$$

Let $Y_1, \cdots, Y_n$ be a random sample from $N(\mu, \sigma^2)$.

Question Find a test statistic Λ in order to test $H_0 : \mu = \mu_0$ v.s. $H_1 : \mu \neq \mu_0$.

Case I. σ^2 is known: $$\Lambda = \frac{\overline{Y} - \mu_0}{\sigma/\sqrt{n}}$$

Case II. σ^2 is unknown: $\Lambda = ?$ $$\Lambda \stackrel{?}{=} \frac{\overline{Y} - \mu_0}{s/\sqrt{n}} \sim ?$$

Let $Y_1, \cdots, Y_n$ be a random sample from $N(\mu, \sigma^2)$.

Question Find a test statistic Λ in order to test $H_0 : \mu = \mu_0$ v.s. $H_1 : \mu \neq \mu_0$.

Case I. σ^2 is known:

$$\Lambda = \frac{\overline{Y} - \mu_0}{\sigma/\sqrt{n}}$$

Case II. σ^2 is unknown:

$$\Lambda = ? \qquad \Lambda \overset{?}{=} \frac{\overline{Y} - \mu_0}{s/\sqrt{n}} \sim ?$$

Summary

A random sample of size n from
a normal distribution $N(\mu, \sigma^2)$

	σ^2 known	σ^2 unknown
Statistic	$Z = \frac{\overline{Y}-\mu}{\sigma/\sqrt{n}}$	$T_{n-1} = \frac{\overline{Y}-\mu}{S/\sqrt{n}}$
Score	$z = \frac{\bar{y}-\mu}{\sigma/\sqrt{n}}$	$t = \frac{\bar{y}-\mu}{s/\sqrt{n}}$
Table	z_α	$t_{\alpha,n-1}$
$100(1-\alpha)\%$ C.I.	$\left(\bar{y} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{y} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$	$\left(\bar{y} - t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}, \bar{y} + t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}\right)$
Test $H_0 : \mu = \mu_0$		
$H_1 : \mu > \mu_0$	Reject H_0 if $z \geq z_\alpha$	Reject H_0 if $t \geq t_{\alpha,n-1}$
$H_1 : \mu < \mu_0$	Reject H_0 if $z \leq z_\alpha$	Reject H_0 if $t \leq t_{\alpha,n-1}$
$H_1 : \mu \neq \mu_0$	Reject H_0 if $\lvert z\rvert \geq z_{\alpha/2}$	Reject H_0 if $\lvert t\rvert \geq t_{\alpha/2,n-1}$

Computing s from data

Step 1 $a = \sum_{i=1}^{n} y_i$

Step 2. $b = \sum_{i=1}^{n} y_i^2$

Step 3. $s = \sqrt{\dfrac{nb - a^2}{n(n-1)}}$

Proof.

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2 = \frac{n\left(\sum_{i=1}^{n} y_i^2\right) - \left(\sum_{i=1}^{n} y_i\right)^2}{n(n-1)}$$

$\square$

Computing s from data

Step 1 $a = \sum_{i=1}^{n} y_i$

Step 2. $b = \sum_{i=1}^{n} y_i^2$

Step 3. $s = \sqrt{\dfrac{nb - a^2}{n(n-1)}}$

Proof.

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2 = \frac{n\left(\sum_{i=1}^{n} y_i^2\right) - \left(\sum_{i=1}^{n} y_i\right)^2}{n(n-1)}$$

$\square$

Computing s from data

Step 1 $a = \sum_{i=1}^{n} y_i$

Step 2. $b = \sum_{i=1}^{n} y_i^2$

Step 3. $s = \sqrt{\dfrac{nb - a^2}{n(n-1)}}$

Proof.

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})^2 = \frac{n\left(\sum_{i=1}^{n} y_i^2\right) - \left(\sum_{i=1}^{n} y_i\right)^2}{n(n-1)}$$

$\square$

Computing s from data

Step 1 $a = \sum_{i=1}^{n} y_i$

Step 2. $b = \sum_{i=1}^{n} y_i^2$

Step 3. $s = \sqrt{\dfrac{nb - a^2}{n(n-1)}}$

Proof.

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2 = \frac{n\left(\sum_{i=1}^{n} y_i^2\right) - \left(\sum_{i=1}^{n} y_i\right)^2}{n(n-1)}$$

$\square$

Case 7.4.1 How far apart are the bat and the insect when the bat first senses that insect is there?

Or, what is the effective range of a bat's echolocation system?

Answer the question by contruct a 95% C.I.

Sol. ... □

Case 7.4.1 How far apart are the bat and the insect when the bat first senses that insect is there?

Or, what is the effective range of a bat's echolocation system?

Answer the question by contruct a 95% C.I.

Sol. ... □

Case 7.4.1 How far apart are the bat and the insect when the bat first senses that insect is there?

Or, what is the effective range of a bat's echolocation system?

Table 7.4.1

Catch Number	Detection Distance (cm)
1	62
2	52
3	68
4	23
5	34
6	45
7	27
8	42
9	83
10	56
11	40

Answer the question by contruct a 95% C.I.

Sol. ...

Case 7.4.1 How far apart are the bat and the insect when the bat first senses that insect is there?

Or, what is the effective range of a bat's echolocation system?

Table 7.4.1

Catch Number	Detection Distance (cm)
1	62
2	52
3	68
4	23
5	34
6	45
7	27
8	42
9	83
10	56
11	40

Answer the question by contruct a 95% C.I.

Sol. ... □

```python
# Case7_4_1.py
import numpy as np
import scipy.stats as st


# returns confidence interval of mean
def confIntMean(a, conf=0.95):
    mean, sem, m = np.mean(a), st.sem(a), st.t.ppf((1+conf)/2., len(a)-1)
    return mean - m*sem, mean + m*sem


def main():
    alpha = 5
    data = np.array([62, 52, 68, 23, 34, 45, 27, 42, 83, 56, 40])
    lower, upper = confIntMean(data, 1-alpha/100)
    print ( """ \

    The {alpha}% confidence interval is ({lower:.2f},{upper:.2f})

        """ .format(**locals () ) )


if __name__ == "__main__":
    main()
```

```
In [83]: run Case7_4_1.py

    The 95% confidence interval is (36.21,60.51)
```

Eg. 7.4.2 Bank approval rates for inner-city residents v.s. rural ones.

Approval rate for rural residents is 62%.

Do bank treat two groups equally? $\alpha = 0.05$

Sol.

$$H_0 : \mu = 62 \qquad v.s. \qquad H_1 : \mu \neq 62.$$

Eg. 7.4.2 Bank approval rates for inner-city residents v.s. rural ones.

Approval rate for rural residents is 62%.

Do bank treat two groups equally? $\alpha = 0.05$

Sol.

$$H_0 : \mu = 62 \qquad v.s. \qquad H_1 : \mu \neq 62.$$

Eg. 7.4.2 Bank approval rates for inner-city residents v.s. rural ones.

Approval rate for rural residents is 62%.

Do bank treat two groups equally? $\alpha = 0.05$

Sol.

$$H_0 : \mu = 62 \qquad v.s. \qquad H_1 : \mu \neq 62.$$

Eg. 7.4.2 Bank approval rates for inner-city residents v.s. rural ones.

Approval rate for rural residents is 62%.

Do bank treat two groups equally? $\alpha = 0.05$

Table 7.4.3

Bank	Location	Affiliation	Percent Approved
1	3rd & Morgan	AU	59
2	Jefferson Pike	TU	65
3	East 150th & Clark	TU	69
4	Midway Mall	FT	53
5	N. Charter Highway	FT	60
6	Lewis & Abbot	AU	53
7	West 10th & Lorain	FT	58
8	Highway 70	FT	64
9	Parkway Northwest	AU	46
10	Lanier & Tower	TU	67
11	King & Tara Court	AU	51
12	Bluedot Corners	FT	59

Sol.

$$H_0 : \mu = 62 \qquad v.s. \qquad H_1 : \mu \neq 62.$$

Table 7.4.4

Banks	n	$\bar{y}$	s	t Ratio	Critical Value	Reject H_0?
All	12	58.667	6.946	−1.66	±2.2010	No

Table 7.4.5

Banks	n	$\bar{y}$	s	t Ratio	Critical Value	Reject H_0?
American United	4	52.25	5.38	−3.63	±3.1825	Yes
Federal Trust	5	58.80	3.96	−1.81	±2.7764	No
Third Union	3	67.00	2.00	+4.33	±4.3027	Yes

```python
# Eg7_4_2.py
import numpy as np
import scipy.stats as st

data = np.array([59, 65, 69, 53, 60, 53, 58, 64, 46, 67, 51, 59])
alpha = 5
mean, sem = np.mean(data), st.sem(data)
n = len(data)
s = sem * np.sqrt(n)
cv = st.t.ppf(1-alpha/200., len(data)-1)
tRatio = (mean-62)/sem


 print ( """ \

        n={n}, sample mean={mean:.3f}, s={s:.3f}, t Ratio={tRatio:.2f}, Critical values={cv:.4f}
        """ .format(**locals()))
```

```
In [113]: run Eg7_4_2.py

        n=12, sample mean=58.667, s=6.946, t Ratio=-1.66, Critical values=2.2010
```