STA255 Week-9 (Day-2)

Shahriar Shams

09/03/2020

Review of day-1

- Confidence interval for population proportion
- Sample size calculation

Learning goals

- Idea of test of hypothesis
- Null vs Alternative hypothesis
- Type-I error
- Critical Region
- Testing $\mu = \mu_0$ vs $\mu \neq \mu_0$ under Normal distribution
- Two sided vs one sided test
- Type-II error and power of a test.

Idea of test of hypothesis

- Suppose we are interested in θ
- In point and interval estimation we try to guess the value of θ based on the sample observations.
- In test of hypothesis we start with a hypothetical statement like $\theta = \theta_0$
- We call this null hypothesis, H_0
- The idea is to check whether our observed data supports H_0 or not.

A numerical example

- Suppose, we are interested in the average income of all Canadians (μ)
- We want to test $H_0 : \mu = \$35,000$
- We collect 10K (representative samples) individuals and get their income data.
- We calculate the sample mean ($\bar{x}$) and here are few scenarios:
  - scenario-1: $\bar{x} = 35,100$
  - scenario-2: $\bar{x} = 35,500$
  - scenario-3: $\bar{x} = 36,000$ ...
  - scenario-10: $\bar{x} = 50,000$

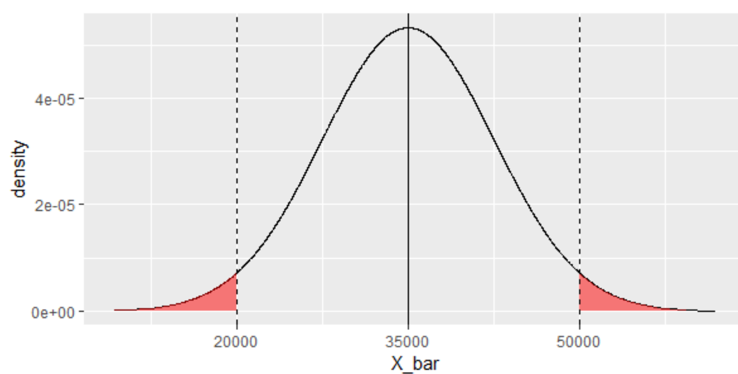
- In which scenario you will reject H_0? (rejecting means the observed data doesn't support our hypothesis)
- In other words: in which scenario the sample mean looks surprising to you if you believe the H_0 to be true?
- Though at which value of $\bar{x}$ you will reject H_0 may seem subjective at this point, but we all agree that at some point the difference between $\bar{x}$ and μ will seem too big.

Type-I error

- In the previous example though we may say $\bar{x} = 50000$ is an indication that the true mean may not be 35000
- But it is totally possible to observe a sample mean of 50000 or higher even though the true population mean is 35000.
- So after observing 50000 as the sample mean if we decide to reject our hypothesis (even though say it is a true hypothesis) we are making a mistake.
- This mistake is called **Type-I** error.
- We denote the probability of making this error using α

$$\alpha = P[rejecting\ H_0 | H_0\ true]$$

- Often we use "level of significance" to refer to type-I error.



Test statistic

- A test statistic is summary of the sample observations.
- In theory it's a random variable whose distribution is known given the null hypothesis.
- In our example, we can say $\bar{X}$ is a test statistic.
- we can also use $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ as the test statistic.
- We use $T(X)$ or simply T to denote a test statistic.

Critical region

- A region of the distribution of the test statistic such that we will reject H_0 if $T(X)$ falls in the rejection region (RR).
- Example: for the numerical example of average income of all Canadians, we can reject the hypothesis $H_0 : \mu = \$35,000$ if $\bar{x} < 20000$ or $\bar{x} > 50000$ (these are made up numbers)
- Here, $\bar{x} < 20000$ and $\bar{x} > 50000$ constitutes the rejection region.
- We need to make sure that

$$P[T(X) \in RR | H_0 \, true] = \alpha$$

- In our example,

$$\alpha = P[\bar{X} < 20000 \; or \; \bar{X} > 50000 | \mu = 35000]$$

Testing $\mu = \mu_0$ vs $\mu \neq \mu_0$ under Normal distribution

Population variance(σ^2) known

- When σ^2 is known, under H_0,

$$\bar{X} \sim N(\mu_0, \sigma^2/n)$$

- We can also write,

$$\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0, 1)$$

- Though it is possible to use $\bar{X}$ as the test statistic, we will use

$$T = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

- Suppose we want to set the probability of type-I error to α.
- The rejection region will then be

$$(-\infty, z_{\alpha/2}) \cup (z_{1-\alpha/2}), \infty$$

- For a given α (say 0.05), it is possible to find out $Z_{\alpha/2}$ and $z_{1-\alpha/2}$
- After calculating the value T, if we find out this value to be in the rejection region, we conclude by saying, "at $\alpha = 0.05$, we reject H_0".

Example: (4.7, 5.5, 4.4, 3.3, 4.6, 5.3, 5.2, 4.8, 5.7, 5.3) $\overset{iid}{\sim} N(\mu, \sigma_0^2)$ with $\sigma_0^2 = 0.5$

Test $H_0 : \mu = 5$ vs $H_a : \mu \neq 5$ at level of significance, $\alpha = 0.05$

1. $\bar{x} = \frac{1}{10}(4.7 + 5.5 + ... + 5.3) = 4.88$
2. test statistic, $T(X) = \frac{4.88-5}{\frac{\sqrt{0.5}}{\sqrt{10}}} = -0.537$
3. given level of significance, $\alpha = 0.05$
4. Using the standard normal distribution, Rejection region $\implies (-\infty, -1.96) \cup (1.96, \infty)$
5. Since, test statistic value -0.537 does not fall in to the rejection area, we fail to reject H_0

Note: We never say we accept H_0. We failed to prove that H_0 is wrong $\nRightarrow$ H_0 is right!

Population variance(σ^2) unknown

- When σ^2 is unknown the frame work remains exactly the same
- the only thing that changes is the test statistic and its distribution.
- We use

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t_{(n-1)}$$

- And the rejection region is then defined as

$$(-\infty, t_{\alpha/2,(n-1)}) \cup (t_{1-\alpha/2,(n-1)}, \infty)$$

Example:

(4.7, 5.5, 4.4, 3.3, 4.6, 5.3, 5.2, 4.8, 5.7, 5.3) $\overset{iid}{\sim} N(\mu, \sigma^2)$ with both μ and σ^2 unknown

Test $H_0 : \mu = 5$ vs $H_a : \mu \neq 5$ at level of significance, $\alpha = 0.05$

1. $\bar{x} = 4.88$ and $s = 0.696$
2. Test statistic, $T = \frac{4.88-5}{0.696/\sqrt{10}} = -0.545$
3. Using a $t_{(9)}$ distribution, Rejection regions=$(-\infty, -2.262) \cup (2.262, \infty)$
4. Fail to reject H_0

Two sided vs one sided test

- In the previous two cases we tested $H_0 : \mu = \mu_0$ vs $H_a : \mu \neq \mu_0$
- Since $\mu \neq \mu_0$ covers both $\mu < \mu_0$ and $\mu > \mu_0$, it's known as two sided test.
- If we interested in an alternative hypothesis $H_a : \mu < \mu_0$ we put the entire rejection region on the left the distribution of the test statistic.
  - RR: $(-\infty, z_\alpha)$ or $(-\infty, t_{\alpha,(n-1)})$
- On the other side, if we interested in an alternative hypothesis $H_a : \mu > \mu_0$ we put the entire rejection region on the right the distribution of the test statistic.
  - RR: $(z_{1-\alpha}, \infty)$ or $(t_{1-\alpha,(n-1)}, \infty)$

Type-II error and power of a test

- When we are conducting a test of hypothesis one of these four cases can happen

	fail to reject H_0	**reject** H_0
H_0 **true**	Correct decision	type-I error
H_0 **false**	type-II error	Correct decision

- We use β to represent the probability of type-II error.
- $1 - \beta$ represents the power of the test.

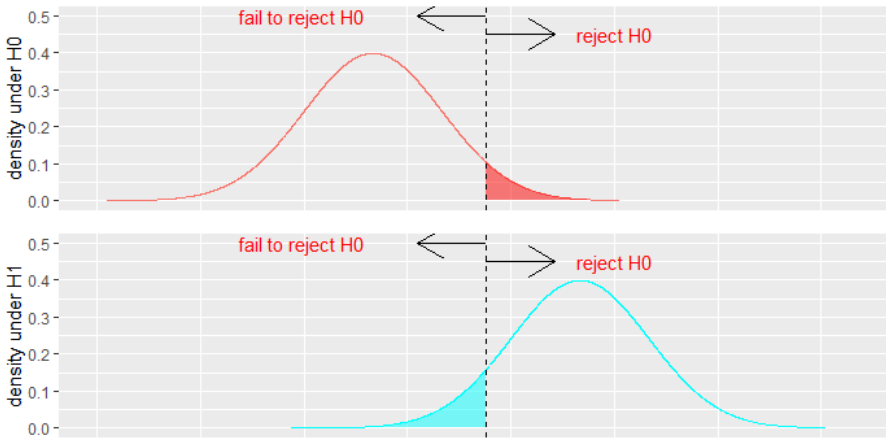
- P[Type-1 error]= $\alpha = P[\text{reject } H_0 | H_0 \text{ true}]$
- P[Type-2 error]= $\beta = P[\text{fail to reject } H_0 | H_0 \text{ false}]$
- Power of a test= $1 - \beta = P[\text{reject } H_0 | H_0 \text{ false}]$

Example:

- Suppose we are testing two simple hypotheses:

$$H_0 : \mu = 1 \ \ vs. \ \ H_a : \mu = 4$$

- Only one of them can be true (and there are no other options)



- Type-1 error: The area shaded in red on the left figure
- Type-2 error: The area shaded in cyan on the right figure
- **Note:** For a given sample size, decreasing one type will increase the other!

Suppose we have $N(\mu, \sigma^2)$ populations with unknown μ and $\sigma = 3$

We want to test $H_0 : \mu = 1 \ vs. \ H_a : \mu = 4$ at $\alpha = 0.05$

we decide to take $n = 9$ observations. Calculate P[type-2 error] and the power.

1. $var[\bar{X}] = \frac{\sigma^2}{n} = \frac{3^2}{9} = 1$
2. Under H_0: $\bar{X} \sim N(1, 1)$
3. Under H_a: $\bar{X} \sim N(4, 1)$
4. $RR = \bar{X}$ satisfying $\frac{\bar{X}-1}{1} > z_{0.95} \implies \bar{X} > 1 + z_{0.95} \implies \bar{X} > 2.645$
5. Power $= P[\bar{X} > 2.645$ Under the $H_a] \implies P[Z > \frac{2.645-4}{1}] = 0.912$
6. P[type-II error] $= 1 - 0.912 = 0.088$

Homework: change the H_a, try $\mu = 3, 5, 6, 7$ etc... and calculate the power in each case.

We will cover chapter 9.3 and 9.4 next week

Chapter 9.5 not needed.

Homework

Chapter 9.1

7, 10

Chapter 9.2

15, 16, 20, 29