Hypothesis Testing

Dr Wan Nor Arifin
Unit of Biostatistics and Research Methodology, Universiti Sains Malaysia.
wnarifin@usm.my

Last update: 3 October, 2018

Outlines

Introduction

- Recall – it is one of the approaches of statistical inference.

Hypothesis

Types of hypothesis

- Research hypothesis "is the conjecture or supposition that motivates the research" (Daniel, 1995).
- Statistical hypothesis is the hypothesis that is stated in a way that is possible to evaluate by appropriate statistical analysis.

Types of statistical hypothesis

- Null hypothesis (H_0) – It states a hypothesis of no difference/agreement with the population of interest.
- Alternative hypothesis (H_A) – The inverse of H_0. It states a hypothesis of disagreement with the population of interest. In most situations, it reflect the proposition in the research hypothesis.

Statistical test and hypothesis

- For a test statistic $W(\mathbf{X})$, where **X** consists of sample values and $W(\mathbf{X})$ is a function of the sample, the sample space is divided into two regions:

  1. **Acceptance region**, the subset of the test statistic values for which H_0 is accepted.

  2. **Rejection region** or **critical region**, the complement of the acceptance region, for which H_0 is rejected and H_A is accepted.

- This is called **Neyman-Pearson paradigm** for this approach in decision making.

Relationship between confidence interval and hypothesis testing

- The region covered by a confidence interval is the region or area covered by the H_0, regarded as the **acceptance region**. If a test statistic value falls in this region, we accept H_0.
- The region outside the confidence interval at both tails for a two-tailed test, or at either tail for a one-tailed test is regarded as the **rejection region**. If a test statistic value that falls in this region, we reject H_0 and accept H_A.

Types of errors related to testing statistical hypothesis

- Type I error (α)
  - Erratically rejects True H_0.
  - Erratically fail to reject False H_A .
  - False Positive (significant) result.
- Type II error (β)
  - Erratically fail to reject False H_0.
  - Erratically rejects True H_A.
  - False Negative (insignificant) result.

Relationship between statistical test and hypothesis

- The relationship between **statistical test** and **hypothesis** can be viewed as relationship between diagnostic test and its gold standard test as shown in table below:

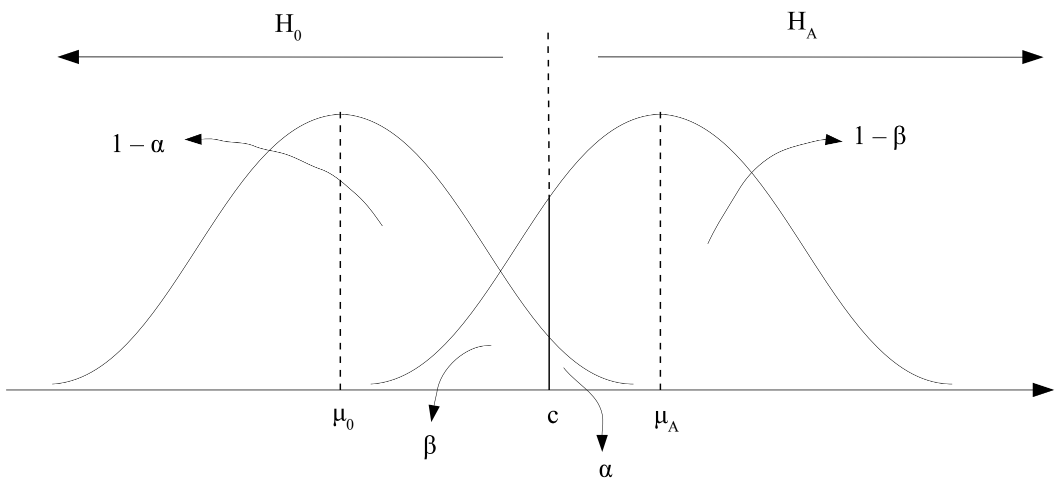
		Hypothesis state (Gold standard)	
		True H_A False H_0 (Difference +)	True H_0 False H_A (Difference –)
Statistical test result (diagnostic test)	Significant (Test +)	True Positive 1 – β Sensitivity/Power	False Positive α Type I Error
	No significant (Test –)	False Negative β Type II Error	True Negative 1 – α Specificity

* Familiarize yourself with diagnostic test.

- In form of probability statement *that you love so much

$$P(\text{Test} + \mid H_0) = \alpha$$
$$P(\text{Test} - \mid H_0) = 1 - \alpha$$
$$P(\text{Test} - \mid H_A) = \beta$$
$$P(\text{Test} + \mid H_A) = 1 - \beta$$

- In form of a plot,



Test statistic

- It is the *statistic* obtained from our calculation using appropriate formula e.g. value of z or t.
- In general formula for hypothesis testing is given by,

$$\text{test statistic } W(\boldsymbol{X}) = \frac{\text{relevant statistic} - \text{hypothesized parameter}}{\text{standard error of relevant statistic}}$$

- It is then compared against a preset value e.g. preset value of z or t (i.e. based on significance level).
- Or convert it into probability value called *P-value*.

Significance level, α

- It is "the probability of rejecting a true null hypothesis" (Daniel, 1995).
- A preset value of acceptable significance level (0.05, 0.01, 0.001).
- Determines the cutoff *c* in the plot above.
- Acceptable level to commit type I error.
- Compared to a test's *P*-value or converted to value of *z* or *t* for comparison to the preset value.

P-value

- It is the probability of getting a test statistic as **extreme** as the one that we computed.
- If we get very low *P*-value, thus it means that the probability of getting a value similar or better than the one that we have is very unlikely.
- Think of anyone beating Usain Bolt for his world record of 9.58secs in 100m sprint, would you think anyone else can easily get similar or even better record than what he did?
- We reject such thinking/hypothesis when *P*-value < α → significant result i.e. such thing would not happen easily by chance/very unlikely.

Hypothesis testing

General steps

1. Data description
   Nature of data: numerical/categorical.
   Present basic information: *n*, mean, variance/standard deviation.

2. Assumptions
   Distribution: normal/not normal.
   Random samples.
   Independence of samples – independent/paired.
   Population variance/standard deviation: known/not known.

3. Hypothesis statements
   H_0 and H_A can be stated in three ways based on H_A (for example mean systolic blood pressure, SBP):

   Difference (two-tailed test)
   Population mean SBP is not 120mmHg,

$$H_0: \mu = 120$$
$$H_A: \mu_A \neq 120$$

   *Most common in health sciences.

   More (one-tailed test)
   Population mean SBP is more than 120mmHg,

$$H_0: \mu_A \leq 120$$
$$H_A: \mu_A > 120$$

Less (one-tailed test)
Population mean SBP is less 120mmHg,

$$H_0: \mu_A \geq 120$$
$$H_A: \mu_A < 120$$

4. Test statistic used
Depending on the data and assumptions, state an appropriate statistic to use.
For example, for normally distributed data, when population variance is unknown → *t* statistic.

5. Decision rule
Set the significance level, usually $\alpha = 0.05$.
Determine the cut-off value of the test statistic at the significance level.
Specify rejection/acceptance region.
One-tailed or two-tailed test based on the hypothesis statement → commonly two-tailed.
If test statistic is outside the acceptance region/within rejection region → significant test, reject H_0. Accept H_A.
If test statistic is within the acceptance region → test not significant, accept H_0. Reject H_A.

6. Calculation of test statistic
Based on the chosen test statistic.

7. Statistical decision
Reject/accept H_0.
Significant/not significant given the preset significance level, α.
State the *P*-value. For the two-tailed *P*-value, multiply the probability obtained from the probability distribution by two.

8. Conclusion
Conclude in relation to the research hypothesis.

One Population Mean

- Test statistic for one population mean, in general is given by

$$\text{test statistic} = \frac{\text{sample mean} - \text{hypothesized mean}}{\text{standard error of mean}}$$

Population variance known: Standard normal distribution, z

- The test statistic for one population mean, when population variance σ^2 is known, is given by

$$z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{(n)}}}$$

given H_0, is distributed as z.
- Data are normally distributed.

Example 1:

A study was done in a sample of 30 adults to investigate the systolic blood pressure (SBP) in a population. It was hypothesized the mean SBP is 120mmHg. The data was normally distributed, with the mean SBP of 125mmHg and standard deviation s of 15mmHg. The sample s is treated as the population σ. Following the general hypothesis testing steps, based on the sample data, decide on the following hypotheses:

1. H_A: $\mu \neq 120$

Answer:

Remember, for the two-tailed P-value, multiply the probability obtained from the probability distribution by two.

2. H_A: $\mu > 120$

Answer:

3. H_A: $\mu < 120$

Answer:

Population variance unknown: Student's *t* distribution

- The test statistic for one population mean, when population variance σ^2 is unknown, is given by

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

given H_0, is distributed as t with $n - 1$ degrees of freedom.

Example 2:

A study was done in a sample of 20 adults to investigate the systolic blood pressure (SBP) in a population. It was hypothesized the mean SBP is 120mmHg. The data was normally distributed, with the mean SBP of 125mmHg and standard deviation *s* of 15mmHg. The population σ is not known, thus sample *s* is used. Following the general hypothesis testing steps, based on the sample data, decide on the following hypotheses:

1. H_A: $\mu \neq 120$

Answer:

Remember, for the two-tailed *P*-value, multiply the probability obtained from the probability distribution by two.

2. H_A: μ > 120

Answer:

3. H_A: μ < 120

Answer:

One Population Proportion

- Test statistic for one population proportion,

$$\text{test statistic} = \frac{\text{sample proportion} - \text{hypothesized proportion}}{\text{standard error of proportion}}$$

- The test statistics is given by,

$$z = \frac{\hat{p} - p}{\sqrt{\left(\frac{p(1-p)}{n}\right)}}$$

given H_0, is distributed as z (provided np and $n(1 - p) > 5$).

Example 3:

A study was done in a sample of 60 adults to investigate the proportion of diabetics in a population.

It was hypothesized the proportion of diabetics in the population is 0.3. Based on the data, 25 participants were found to be diabetics. Following the general hypothesis testing steps, based on the sample data, decide on the following hypotheses:

1. H_A: $p \neq 0.3$

Answer:

Remember, for the two-tailed *P*-value, multiply the probability obtained from the probability distribution by two.

2. H_A: $p > 0.3$

Answer:

3. H_A: $p < 0.3$

Answer:

One Population Variance

- Test statistic for one population variance,

$$\text{test statistic} = \frac{\text{sample variance}}{\text{hypothesized population variance}}$$

- The test statistic is given by,

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

given H_0, is distributed as χ^2 with n – 1 degrees of freedom.

Example 4:

In a study measuring SBP involving 30 subjects, it was found the sample standard deviation was 18.5mmHg. The population standard deviation was hypothesized to be 15mmHg. Following the general hypothesis testing steps, based on the sample data, decide on the following hypotheses for the variance:

1. H_A: $\sigma^2 \neq 15^2$

Answer:

For the two-tailed *P*-value of asymmetric distribution like X^2, we cannot multiply the *P*-value by two like we did for symmetrical distributions like *z* and *t*. We must report the one-sided *P*-values or just state $P > \alpha$, e.g. $P > 0.05$.

2. H_A: $\sigma^2 > 15^2$

Answer:

3. H_A: $\sigma^2 < 15^2$

Answer:

Other Topics

Hypothesis testing approach for the

- difference between two population means → independent t-test lecture.
- population mean difference for paired comparison → paired t-test lecture.
- difference between two population proportions → chi-squared test lecture.
- ratio between two population variances → equality of variance test.

Topics for Self-study

Likelihood ratio test
Score test
Wald test
Articles:

- [Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations]

- [The ASA's Statement on p-Values: Context, Process, and Purpose]

References

Casella, G., & Berger, R. L. (2002). Statistical inference. Delhi, India: Cengage Learning.

Daniel, W. W. (1995). *Biostatistics: A foundation for analysis in the health sciences* (6th ed.). USA: John Wiley & Sons.

Gaddis, G. M., & Gaddis, M. L. (1990). Introduction to biostatistics: Part 3, sensitivity, specificity, predictive value, and hypothesis testing. *Annals of Emergency Medicine*, **19(5)**, 145-151.

Rice, J. A. (1995). *Mathematical statistics and data analysis* (2nd ed.). USA: Duxbury Press.