Probability & Statistics

Dongjiao Ge

djge@cityu.edu.mo

Session 10

Reading list:
https://minerva.it.manchester.ac.uk/~saralees/lecturenotes.pdf
Our slides for this part are made mainly based on chapter 7, 8 and 9 of the above lecture
Notes.

7.2 Single sample procedures

7.2.4 Confidence interval for the unknown mean of a non-normal distribution with either known or unknown variance

Suppose that we now have a '**large**' random sample from a non-normal distribution, and that we wish to use the data to construct a confidence interval for the unknown distribution mean μ.

We can appeal to the **central limit theorem** and construct a 100(1 − α)% CI as follows.

7.2 Single sample procedures

7.2.4 Confidence interval for the unknown mean of a non-normal distribution with either known or unknown variance

- If the **variance σ^2 is known** then, by the central limit theorem, for **large n** the statistic

$$Z_1 = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$$

is approximately distributed as $N\ (0, 1)$. Thus an approximate $100(1\ -\ \alpha)\%$ confidence interval for μ is given by

$$\left[\overline{X} - \frac{z_{1-\frac{\alpha}{2}}\,\sigma}{\sqrt{n}}\ ,\ \overline{X} + \frac{z_{1-\frac{\alpha}{2}}\,\sigma}{\sqrt{n}}\right].$$

7.2 Single sample procedures

7.2.4 Confidence interval for the unknown mean of a non-normal distribution with either known or unknown variance

- If the **variance is unknown**, then we instead plug in the sample standard deviation S for σ to obtain the statistics

$$Z_2 = \frac{\overline{X} - \mu}{S/\sqrt{n}}.$$

It can also be shown that Z_2 is also **distributed approximately** as $N(0,1)$ for **large n**. Thus an approximate $100(1-\alpha)\%$ confidence interval for μ is given by

e.g., n>50

$$\left[\overline{X} - \frac{z_{1-\frac{\alpha}{2}}S}{\sqrt{n}},\ \overline{X} + \frac{z_{1-\frac{\alpha}{2}}S}{\sqrt{n}}\right].$$

7.2 Single sample procedures

7.2.4 Confidence interval for the unknown mean of a non-normal distribution with either known or unknown variance

Example The data set contains $n = 500$ observations and we have that $x = 33.27$ and $s^2 = 503.554$. By the above discussion, the end points

$$33.27 \pm 1.96 \times \sqrt{\frac{503.554}{500}}$$

define a 95% confidence interval for μ, namely $(31.30, 35.24)$. This gives a range of plausible values for the unknown value of μ.

7.2 Single sample procedures

7.2.5 Confidence interval for the unknown variance of a normal distribution, mean also unknown

Let $X_1, \ldots, X_n$ be a random sample from the $N(\mu, \sigma^2)$ distribution where **both μ and σ^2 are unknown**. We would like to construct a $100(1 - \alpha)\%$ confidence interval for σ^2.

Given that

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

is an unbiased estimator of σ^2. Also, we have the distributional result that

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

7.2 Single sample procedures

7.2.5 Confidence interval for the unknown variance of a normal distribution, mean also unknown

It then follows that

$$\mathrm{P}\left(\chi^2_{\frac{\alpha}{2}} < \frac{(n-1)S^2}{\sigma^2} < \chi^2_{1-\frac{\alpha}{2}}\right) = 1-\alpha\,,$$

where $\chi^2_{1-\frac{\alpha}{2}}$ denotes the $(1-\frac{\alpha}{2})$ point of a $\chi^2(n-1)$ distribution, i.e. if $Y \sim \chi^2(n-1)$ then $P(Y \leq \chi^2_{1-\frac{\alpha}{2}}) = 1-\frac{\alpha}{2}$. We can re-arrange the inequalities to give bounds for the parameter σ^2, as follows

$$\mathrm{P}\left(\frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}}\right) = 1-\alpha\,.$$

7.2 Single sample procedures

7.2.5 Confidence interval for the unknown variance of a normal distribution, mean also unknown

Hence the 100(1 – α)% confidence interval for σ^2, based on a sample of size n from a normal population is given by

$$\left[\frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}}, \frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}} \right].$$

The inference is that this random interval contains the true value of σ^2 with probability $1 - \alpha$. A $100(1 - \alpha)\%$ confidence interval for σ can be obtained by taking the square roots of the confidence limits for σ^2.

$$\left[\sqrt{\frac{(n-1)S^2}{\chi^2_{1-\alpha/2}}}, \sqrt{\frac{(n-1)S^2}{\chi^2_{\alpha/2}}} \right]$$

7.2 Single sample procedures

7.2.5 Confidence interval for the unknown variance of a normal distribution, mean also unknown

Example We have data with $n = 50$ and $s^2 = 15.288$ so that a 95% confidence interval for σ^2, assuming normality, is given by

$$\left(\frac{49 \times 15.288}{\chi^2_{0.975}}, \frac{49 \times 15.288}{\chi^2_{0.025}} \right),$$

where the χ^2 values correspond to a χ^2 distribution with 49 degrees of freedom.

The above interval is a 95% confidence interval for σ^2

7.2 Single sample procedures

7.2.6 Confidence interval for an unknown population proportion

Let $X_1, \ldots, X_n$ be a random sample from $Bi(1, p)$, i.e. the Bernoulli distribution, where the value of p is unknown. We have already seen that the estimator, $\hat{p} = \bar{X}$ is an unbiased estimator of p with variance $p(1 - p)/n$.

By the central limit theorem, $\hat{p} \sim N(p, p(1 - p)/n)$ approximately for large n. Thus, for large n,

$$\mathrm{P}\left(-z_{1-\alpha/2} \leq \frac{\widehat{p} - p}{\sqrt{p(1-p)/n}} \leq z_{1-\alpha/2}\right) \approx 1 - \alpha\,,$$

In fact it can also be shown that the above remains true even if $\sqrt{Var\ \hat{p}}$ in the denominator is estimated via $\sqrt{\hat{p}(1 - \hat{p})/n}$, i.e. for large n,

$$\mathrm{P}\left(-z_{1-\alpha/2} \leq \frac{\widehat{p} - p}{\sqrt{\widehat{p}(1-\widehat{p})/n}} \leq z_{1-\alpha/2}\right) \approx 1 - \alpha\,.$$

7.2 Single sample procedures

7.2.6 Confidence interval for an unknown population proportion

Hence we have that for large n

$$\mathrm{P}\left(\widehat{p} - z_{1-\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}} \leq p \leq \widehat{p} + z_{1-\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}}\right) \approx 1-\alpha\,.$$

It then follows that

$$\left[\widehat{p} - z_{1-\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}}\,,\ \widehat{p} + z_{1-\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}}\right]$$

is an approximate $100(1-\alpha)\%$ confidence interval for the parameter p.

7.2 Single sample procedures

7.2.6 Confidence interval for an unknown population proportion

Example The opinion poll data collected from $n = 1000$ voters. We would like to use these data to obtain a 95% CI for the proportion in the population who support Labour, denoted by p_L. The proportion in the sample supporting Labour was found to be 0.314 which is our sample estimate of p_L, i.e. $\widehat{p_L} = 0.314$. From the above, our 95% CI has end points

$$0.314 \pm 1.96 \times \sqrt{\frac{0.314 \times 0.686}{1000}},$$

i.e. the interval is (0.285, 0.343).

8 Hypothesis testing (Part I)

8.1 Introduction

- One of the main aims of statistical analysis is to make inferences about the unknown values of population parameters based on a sample of data from the population.
- We previously considered both point and interval estimation of such parameters.
- Here we instead explore how to test hypotheses about the values of parameters.

8 Hypothesis testing (Part I)

8.1 Introduction

A statistical hypothesis is a conjecture or proposition regarding the distribution of one or more random variables.

In order to specify a statistical hypothesis we need to specify the family of the ***underlying distribution*** (e.g. normal, Poisson, or binomial) as well as the set of ***possible values of any parameters***.

- A ***simple hypothesis*** specifies the distribution and the parameter values uniquely. (e.g., the data arising from $N(\mu, 1^2), with\ \mu = 5$)
- In contrast, a ***composite hypothesis*** specifies several different possibilities for the distribution, most commonly corresponding to different possibilities for the parameter values. (e.g., the data arise from $N(\mu, 1^2), with\ \mu > 5$)

8 Hypothesis testing (Part I)
8.1 Introduction

The elements of a statistical test (first two of five):

i. The **null hypothesis**, denoted by H_0, is the hypothesis to be tested. This is usually a 'conservative' or 'skeptical' hypothesis that we believe by default unless there is significant evidence to the contrary.
ii. The **alternative hypothesis**, denoted by H_1 or H_a , is a hypothesis about the population parameters which we will accept if there is evidence that H_0 should be rejected.

E.g., when assessing a new medical treatment, it is common for the null hypothesis to correspond to the statement that the new treatment is no better (or worse) than the old one. The alternative hypothesis would be that the new treatment is better.

8 Hypothesis testing (Part I)

8.1 Introduction

In this module, the **null hypothesis** will always be **simple**, while the **alternative hypothesis** may either be **simple or composite**. For example, consider the following hypotheses about the value of the mean μ of a normal distribution with known variance σ^2:

$H_0: \mu = \mu_0$, where μ_0 is a specific numerical value, is a **simple null hypothesis**.

- $H_1: \mu = \mu_1$ (with $\mu_1 \neq \mu_0$) is a **simple alternative hypothesis**.
- $H_1: \mu > \mu_0$ is a **one-sided composite alternative** hypothesis.
- $H_1: \mu < \mu_0$ is a **one-sided composite alternative** hypothesis.
- $H_1: \mu \neq \mu_0$ is a **two-sided composite alternative** hypothesis.

8 Hypothesis testing (Part I)

8.1 Introduction

How do we use the sample data to decide between H_0 and H_1?

The elements of a statistical test ("three&four" of five):

iii. **Test statistic.** This is a function of the sample data whose value we will use to decide whether or not to reject H_0 in favour of H_1. Clearly, the test statistic will be a random variable.

iv. **Acceptance and rejection regions**. We consider the set of all possible values that the test statistic may take, i.e. the range space of the statistic, and we examine the distribution of the test statistic under the assumption that H_0 is true. *The range space is then divided into two disjoint subsets called the* ***acceptance region*** *and* ***rejection region****.*

8 Hypothesis testing (Part I)
8.1 Introduction

Notes on (iv) **Acceptance and rejection regions:**

- On observing data, if the *calculated value of the test statistic falls into the rejection region* then we **reject** H_0 in favour of H_1.
- If the value of the *test statistic falls in the acceptance region* then we **do not reject** H_0 .

The **rejection region** is usually defined to be **a set of extreme values** of the test statistic which together have low probability of occuring if H_0 is true. Thus, if we observe such a value then this is taken as evidence that H_0 is in fact false.

8 Hypothesis testing (Part I)

8.1 Introduction

The elements of a statistical test ("five" of five):

v. **Type I and type II errors.** The procedure described in **(iv)** above can lead to two types of possible errors:
   - **Type I error** - this occurs if we reject H_0 when it is in fact true.
   - **Type II error** - this occurs if we fail to reject H_0 when it is in fact false.

The probability of making a **type I error** is denoted by α and is also called the **significance level** or **size of the test**. The value of α is usually ***specified in advance***; the rejection region is chosen in order to achieve this value. A common choice is $\alpha = 0.05$. Note that $\alpha = P(reject\ H_0\ |H_0)$.

The probability of making a **type II error** is $\beta = P(do\ not\ reject\ H_0\ |H_1)$. For a good testing procedure, β should be small for all values of the parameter included in H_1.

8 Hypothesis testing (Part I)

8.1 Introduction

Example *Is a die biased or not?*
It is claimed that a particular die used in a game is biased in favour of the six.

To test this claim the die is rolled 60 times, and each time it is *recorded whether or not a six is obtained*. At the end of the experiment, *the total number of sixes is counted*, and this information is used to decide whether or not the die is biased.

The **null hypothesis** to be tested is that the die is fair, i.e. $P(rolling\ a\ six) = 1/6$.
The **alternative hypothesis** is that the die is biased in favour of the six so that $P(rolling\ a\ six) > 1/6$.
Let the probability of rolling a six be denoted by p. We can write the above hypotheses as:

$$H_0 : p = 1/6$$

$$H_1 : p > 1/6\,.$$

8 Hypothesis testing (Part I)

8.1 Introduction

Let X denote the number of sixes thrown in 60 attempts.

- If H_0 is true then $X \sim Bi(60, 1/6)$, whereas if H_1 is true then X ~ Bi(60, p), with p > 1/6. H_0 is a simple hypothesis, whereas H_1 is a composite hypothesis.

If H_0 were true, **we would expect to see 10 sixes**, since $E(X) = 10$ **under H_0** . However, the actual number observed will vary randomly around this value.

- If we observe a large number of sixes, then this will constitute evidence against H_0 in favour of H_1 . The question is, **how large does the number of sixes need to be so that we should reject H_0 in favour of H_1?**

The **test statistic** here is x and the ***rejection region*** is

$$\{x : x > k\},$$

for some $k \in N$.

8 Hypothesis testing (Part I)

8.1 Introduction

We now choose the smallest value of k that ensures a significance level $\alpha < 0.05$, i.e. the smallest k such that

$$\alpha = \mathrm{P}(X > k \mid H_0) < 0.05\,.$$

Try different values of k, what can be observed?

- $k = 14, P(X > k|H_0) = 0.0648$
- $k = 15, P(X > k|H_0) = 0.0338$

Thus, we select $k = 15$. In this case, the actual significance level of the test is 0.0338.

8 Hypothesis testing (Part I)

8.1 Introduction

Example (continue)

Remarks

- When the test statistic is a discrete random variable, for many choices of significance level there is no corresponding rejection region achieving that significance level exactly (e.g. $\alpha = 0.05$ above).

What can you conclude from the example?

- Under H_0 the probability of observing more than 15 sixes in 60 rolls is 0.0338. This event is sufficiently unlikely under H_0 that if it occurs then we reject H_0 in favour of H_1 . It is possible that by rejecting H_0 we may make a **type I error**, with probability 0.0338 if H_0 is true.
- If 15 or fewer sixes are obtained, then this is within the acceptable bounds of random variation under H_0 . Thus, in this case, we would not reject the null hypothesis that the die is unbiased. However, in making this decision, we may be making a **type II error**, if H_1 is in fact true.

8 Hypothesis testing (Part I)

8.1.1 Probability of correctly rejecting H_0 when it is false

The probability of correctly rejecting H_0 when it is false satisfies

$$\mathrm{P}(\text{reject } H_0 \,|\, p) = 1 - \mathrm{P}(\text{type II error})\,.$$

Ideally, we would like the probability on the left to be high.

It is straightforward to evaluate this probability for particular values of $p > 1/6$. Specifically, $P(reject\ H_0 \mid p) = P(X > 15 \mid p)$, where $X \sim Bi(60, p)$. For example, the following values have been computed:

p	$\mathrm{P}(\text{reject } H_0 \,\|\, p)$
0.2	0.1306
0.25	0.4312
0.3	0.7562

Clearly, the larger the true value of p, the more likely we are to correctly reject H_0.

9 Hypothesis testing (Part 2) --- single sample

9.1 Introduction

We will now discuss specific applications of hypothesis testing where we have a **single sample** of data and wish to test hypotheses regarding the value of a ***population mean parameter***.

The exact case we will consider: the random sample is from a $N(\mu, \sigma^2)$ distribution with **μ unknown and σ^2 known**.

The above ideas are then **extended** to develop hypothesis tests for

(i) the mean of a normal distribution with **unknown variance**,

(ii) the mean of a **non-normal distribution**,

(iii) a population **proportion** p.

*In cases (ii) and (iii) it is not possible to calculate the exact distribution of the test statistic under the null hypothesis, however we can appeal to the **central limit theorem** to find an approximate normal distribution.*

9 Hypothesis testing (Part 2) --- single sample

9.2 Inference about the mean of a normal distribution when the variance is known

Let $X_1, \ldots, X_n$ be a random sample from $N(\mu, \sigma^2)$, where the value of $\boldsymbol{\mu}$ **is unknown** but the value of $\boldsymbol{\sigma^2}$ **is known**.

$$H_0 : \mu = \mu_0 \text{ vs } H_1 : \mu > \mu_0 .$$

- This is a **one-sided** test.
- The **null hypothesis** H_0 posits that the data are sampled from $N(\mu_0, \sigma^2)$,
- The **alternative hypothesis** H_1 posits that the data arise from $N(\mu_1, \sigma^2)$, where $\mu_1 > \mu_0$ is an unspecified value of μ.

9 Hypothesis testing (Part 2) --- single sample

9.2 Inference about the mean of a normal distribution when the variance is known

- We know that the sample mean, $\bar{X}$, is an unbiased estimator of μ. Hence, if the true value of μ is μ_0, then $E[\bar{X} - \mu_0] = \mu_0 - \mu_0 = 0$.
- In contrast, if H_1 is true, we would have that $E[\bar{X} - \mu_0] = \mu - \mu_0 > 0$. This suggests that we should reject H_0 in favour of H_1 if $\bar{X}$ is 'significantly' larger than μ_0, i.e. if $\bar{X} > k$, for some $k > \mu_0$.

The question is, how much greater than μ_0 should $\bar{x}$ be before we reject H_0? In other words, what value should we choose for k?

9 Hypothesis testing (Part 2) --- single sample

9.2 Inference about the mean of a normal distribution when the variance is known

One way to answer that question is to fix the probability of rejecting H_0 if H_0 is true, i.e. the probability of making a **Type I error**; the critical value k can then be determined on this basis.

I.e., to fixing the significance level of the test. Suppose that we do indeed use $\bar{X}$ as the test statistic, with **rejection region**

$$C = \{\overline{x} > k\}\,,$$

and suppose we wish to find $k > \mu_0$ to ensure that

$$\mathrm{P}(\text{type I error}) = \mathrm{P}(\text{reject } H_0 \mid H_0 \text{ true}) = \alpha\,.$$

9 Hypothesis testing (Part 2) --- single sample

9.2 Inference about the mean of a normal distribution when the variance is known

Hence, we have that

$$\begin{aligned}\alpha &= \mathrm{P}(\text{reject } H_0 \mid H_0 \text{ true}) = \mathrm{P}(\overline{X} > k \mid H_0 \text{ true}) \\ &= \mathrm{P}\left(\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} > \frac{k-\mu_0}{\sigma/\sqrt{n}}\right) \\ &= \mathrm{P}\left(Z > \frac{k-\mu_0}{\sigma/\sqrt{n}}\right),\end{aligned}$$

where $Z = \frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$ under H_0.

Let $z_{1-\alpha}$ denote the α point of $N(0,1)$, i.e. $\mathrm{P}(Z \le z_{1-\alpha}) = 1-\alpha$.

So, $z_{1-\alpha} = \frac{k-\mu_0}{\sigma/\sqrt{n}}$ ⟹ $k = \mu_0 + \frac{z_{1-\alpha}\,\sigma}{\sqrt{n}}$.

Thus, H_0 is rejected in favour of H_1 if the sample mean is greater than μ_0 by $z_{1-\alpha}$ standard errors.

9 Hypothesis testing (Part 2) --- single sample

9.2 Inference about the mean of a normal distribution when the variance is known

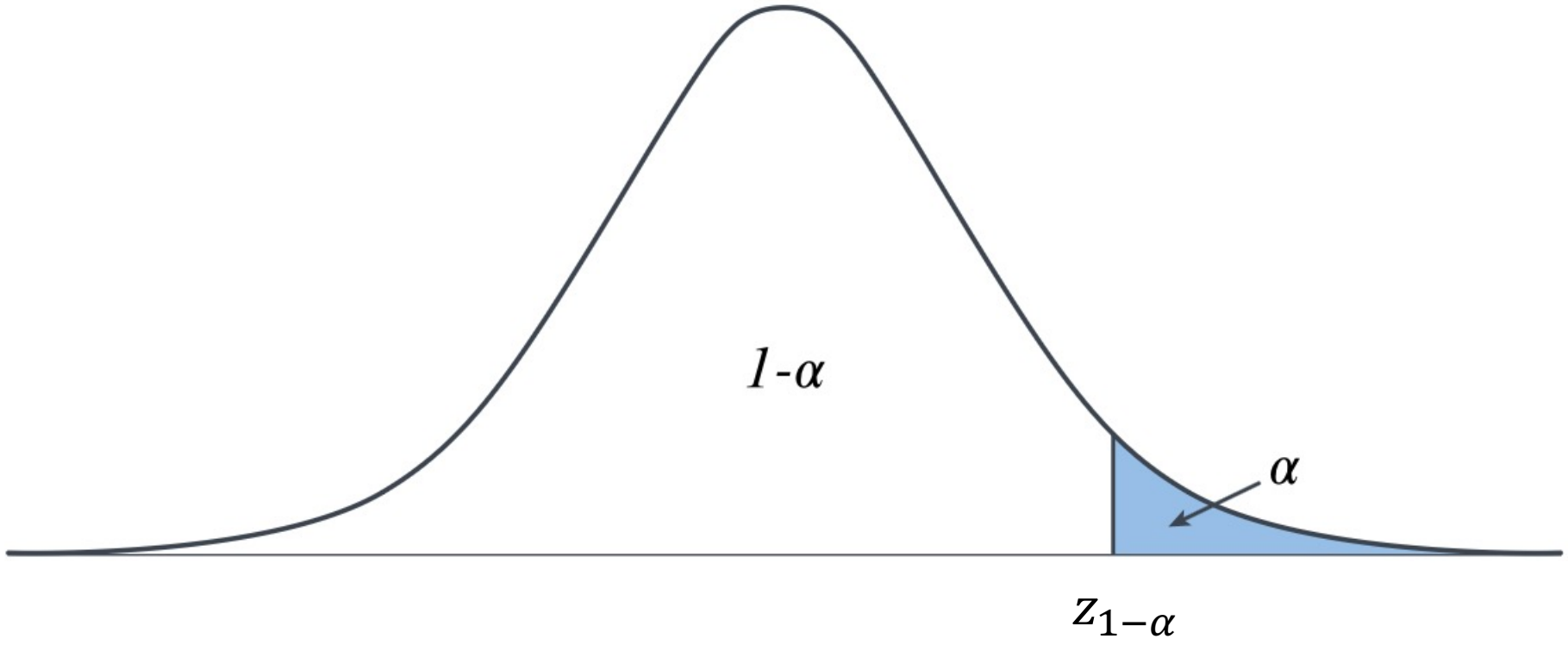
Equivalently, we reject H_0 in favour of H_1 at the $100\alpha\%$ significance level if

$$Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} > z_{1-\alpha}\,.$$

The critical value $z_{1-\alpha}$ can be obtained from standard normal tables. In hypothesis testing it is common to use $\alpha = 0.05$, and in this case $z_{0.95} = 1.645$.

9 Hypothesis testing (Part 2) --- single sample

9.2 Inference about the mean of a normal distribution when the variance is known



9 Hypothesis testing (Part 2) --- single sample

9.2 Inference about the mean of a normal distribution when the variance is known

Suppose now that we wish to use our sample to test the hypotheses

$$H_0 : \mu = \mu_0 \quad \text{vs} \quad H_1 : \mu < \mu_0 \,. \qquad \text{(one-sided test)}$$

In this case we will reject H_0 in favour of H_1 if $\overline{X} < k$ where $k < \mu_0$.

we will reject H_0 in favour of H_1 at the $100\alpha\%$ significance level if

$$\overline{X} < \mu_0 - \frac{z_{1-\alpha}\,\sigma}{\sqrt{n}}\,,$$

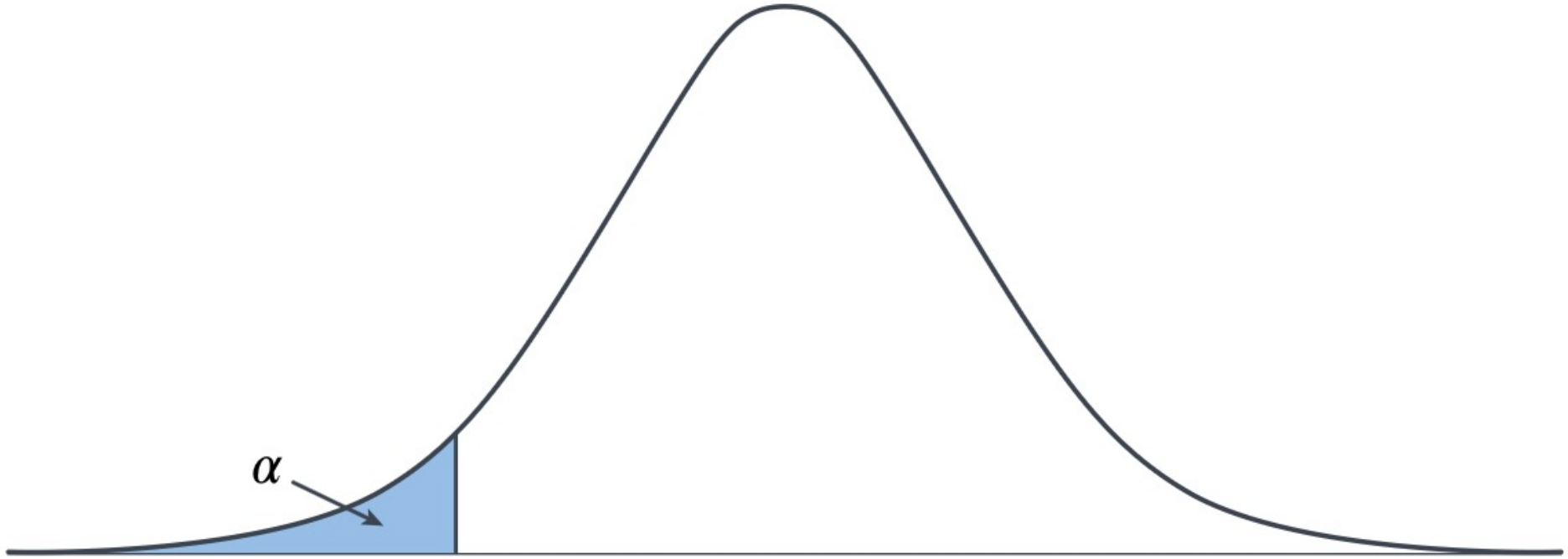
or, equivalently, if

$$Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} < -z_{1-\alpha}\,.$$

For a test having a 5% significance level the critical value is $-z_{0.95} = -1.645$.

9 Hypothesis testing (Part 2) --- single sample

9.2 Inference about the mean of a normal distribution when the variance is known



$$z_\alpha = -z_{1-\alpha}$$

9 Hypothesis testing (Part 2) --- single sample

9.2 Inference about the mean of a normal distribution when the variance is known

If in fact our interest is in testing

$$H_0 : \mu = \mu_0 \quad \text{vs} \quad H_1 : \mu \neq \mu_0 \,, \qquad \text{(two-sided test)}$$

We will reject H_0 in favour of H_1 if $\bar{X}$ is either significantly greater or significantly less than μ_0, i.e. if

$$\overline{X} < k_1 \;\text{ or }\; \overline{X} > k_2 \,,$$

The critical values $k_1 < \mu_0$ and $k_2 > \mu_0$ are chosen so that the significance level is equal to α, i.e.

$$\begin{aligned} \alpha &= \mathrm{P}(\overline{X} < k_1 \text{ or } \overline{X} > k_2 \mid H_0 \text{ true}) \\ &= \mathrm{P}(\overline{X} < k_1 \,|\, H_0) + \mathrm{P}(\overline{X} > k_2 \,|\, H_0) \,. \end{aligned}$$

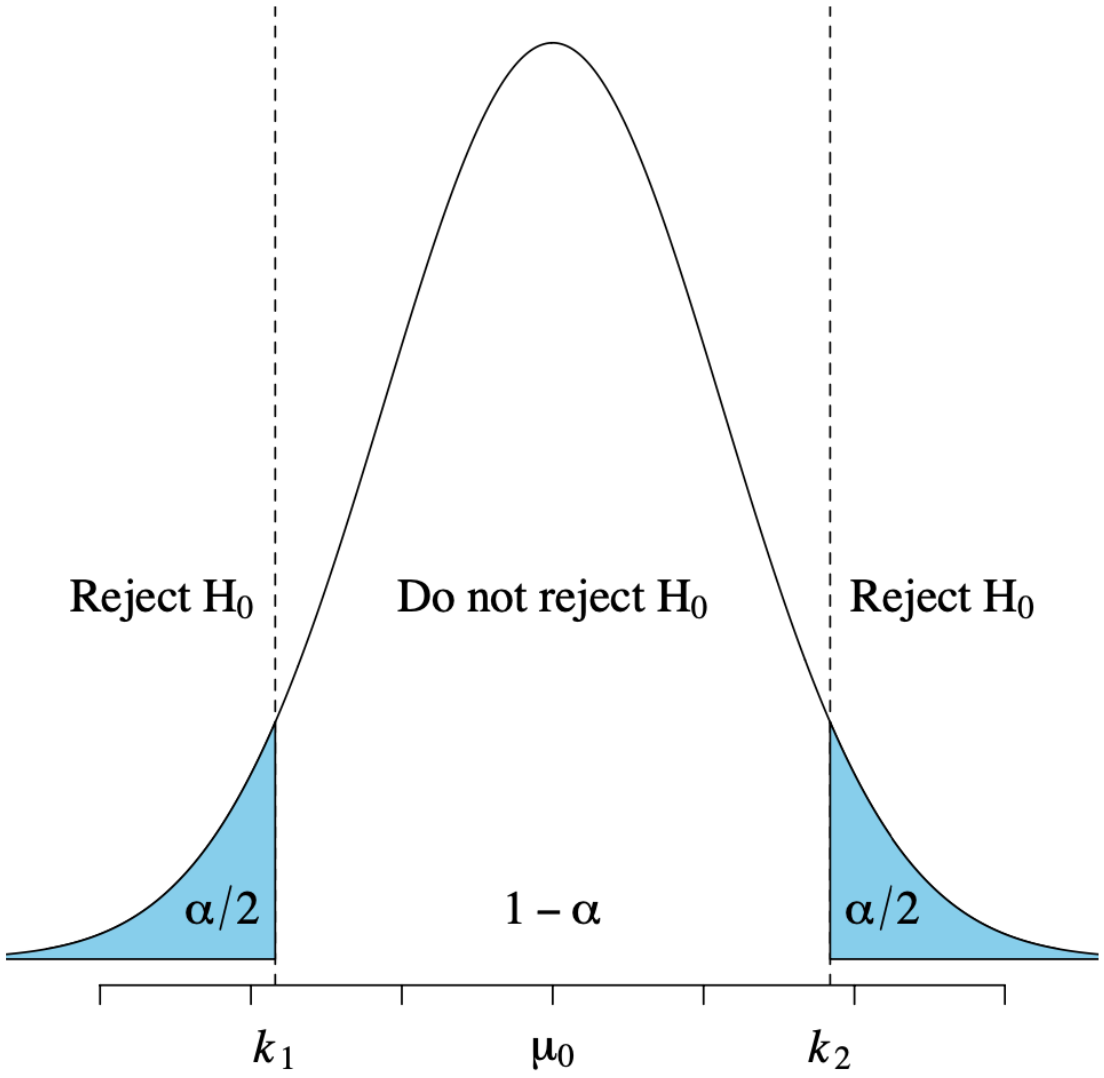
We choose k_1 and k_2 such that

$$\mathrm{P}(\overline{X} < k_1 \,|\, H_0) = \mathrm{P}(\overline{X} > k_2 \,|\, H_0) = \alpha/2 \,.$$

9 Hypothesis testing (Part 2) --- single sample

9.2 Inference about the mean of a normal distribution when the variance is known

See p.d.f of $\bar{X}$



9 Hypothesis testing (Part 2) --- single sample

9.2 Inference about the mean of a normal distribution when the variance is known

We now find appropriate values of k_1 and k_2 satisfying this property. We begin with k_2. Note that

$$\alpha/2 = \mathrm{P}(\overline{X} > k_2 \mid H_0 \text{ true }) = \mathrm{P}\left(\frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} > \frac{k_2 - \mu_0}{\sigma/\sqrt{n}}\right)$$
$$= \mathrm{P}\left(Z > \frac{k_2 - \mu_0}{\sigma/\sqrt{n}}\right), \text{ with } Z \sim N(0,1).$$

However, we know that $z_{1-\alpha/2}$ satisfies $\mathrm{P}(Z \leq z_{1-\alpha/2}) = 1 - \alpha/2$. Hence,

$$\frac{k_2 - \mu_0}{\sigma/\sqrt{n}} = z_{1-\alpha/2},$$

and so we have that

$$k_2 = \mu_0 + \frac{z_{1-\alpha/2}\,\sigma}{\sqrt{n}}.$$

9 Hypothesis testing (Part 2) --- single sample

9.2 Inference about the mean of a normal distribution when the variance is known

For k_1, observe that

$$\alpha/2 = \mathrm{P}(\overline{X} < k_1 \mid H_0 \text{ true }) = \mathrm{P}\left(\frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} < \frac{k_1 - \mu_0}{\sigma/\sqrt{n}}\right)$$
$$= \mathrm{P}\left(Z < \frac{k_1 - \mu_0}{\sigma/\sqrt{n}}\right), \text{ with } Z \sim N(0,1).$$

We know that $\mathrm{P}(Z < -z_{1-\alpha/2}) = \alpha/2$ and so $\frac{k_1-\mu_0}{\sigma/\sqrt{n}} = -z_{1-\alpha/2}$. Hence

$$k_1 = \mu_0 - \frac{z_{1-\alpha/2}\,\sigma}{\sqrt{n}}.$$

9 Hypothesis testing (Part 2) --- single sample

9.2 Inference about the mean of a normal distribution when the variance is known

To summarize the two-tailed test here, we reject H_0 at significance level α if

$$\overline{X} > \mu_0 + \frac{z_{1-\alpha/2}\,\sigma}{\sqrt{n}} \quad \text{or if}$$

$$\overline{X} < \mu_0 - \frac{z_{1-\alpha/2}\,\sigma}{\sqrt{n}}\,.$$

Equivalently, we reject H_0 at significance level α if

$$Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} > z_{1-\alpha/2} \quad \text{or if}$$

$$Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} < -z_{1-\alpha/2}\,.$$

9 Hypothesis testing (Part 2) --- single sample

9.2.1 Connection between the two-tailed test and a confidence interval for the mean when the variance is known

Let $X_1, \ldots X_n$ be a random sample from $N(\mu, \sigma^2)$ with μ unknown and σ^2 known.

Recall that a 100(1 − α)% confidence interval for μ is given by

$$\left[\overline{X} - \frac{z_{1-\alpha/2}\,\sigma}{\sqrt{n}},\ \overline{X} + \frac{z_{1-\alpha/2}\,\sigma}{\sqrt{n}}\right].$$

If we are testing the hypotheses

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu \neq \mu_0\,,$$

then we will 'accept' H_0 at the $100\alpha\%$ significance level if

$$\mu_0 - \frac{z_{1-\alpha/2}\,\sigma}{\sqrt{n}} \leq \overline{X} \leq \mu_0 + \frac{z_{1-\alpha/2}\,\sigma}{\sqrt{n}}\,,$$

or, equivalently, if

$$\overline{X} - \frac{z_{1-\alpha/2}\sigma}{\sqrt{n}} \leq \mu_0 \leq \overline{X} + \frac{z_{1-\alpha/2}\sigma}{\sqrt{n}}\,.$$

9 Hypothesis testing (Part 2) --- single sample

9.2.1 Connection between the two-tailed test and a confidence interval for the mean when the variance is known

- Thus, the values of μ in the confidence interval correspond to values of μ_0 for which the corresponding null hypothesis H_0 would not be rejected.
- In other words, informally, the 100(1 − α)% confidence interval is a set of values of μ which would 'pass a hypothesis test at significance level α'.
- It is in this sense that we can regard the confidence interval as a set of plausible values of μ given the data.

9 Hypothesis testing (Part 2) --- single sample

9.2.1 Connection between the two-tailed test and a confidence interval for the mean when the variance is known

Example. Suppose now that we have a random sample of $n = 50$ observations from a normal distribution with unknown mean and known variance $\sigma^2 = 36$. It is found that $\bar{x} = 30.8$.

(i) Test $H_0 : \mu = 30$ vs $H_1 : \mu \neq 30$ at the 5% significance level.

Solution: here the test statistic is

$$Z = \frac{30.8 - 30.0}{\sqrt{36/50}} = 0.943\,.$$

Using a 5% significance level the critical values are $-z_{0.975} = -1.96$ and $z_{0.975}$ = 1.96. The observed value of Z lies between the two critical values, thus H_0 is not rejected at the 5% significance level. We conclude that there is insufficient evidence to reject the claim that the normal distribution from which the data arise has mean 30.

9 Hypothesis testing (Part 2) --- single sample

9.3 Inference about the mean of a normal distribution when the variance is unknown

Let $X_1, \ldots, X_n$ be a random sample from $N(\mu, \sigma^2)$, where the value of $\boldsymbol{\mu}$ **is unknown** but the value of $\boldsymbol{\sigma^2}$ **is unknown**. We want to test the following hypotheses:

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu > \mu_0$$

at significance level α.

An appropriate test statistic that measures the discrepancy between μ_0 and the sample estimator $\bar{X}$ is given by

$$T = \frac{\overline{X} - \mu_0}{S/\sqrt{n}}$$

where S is the sample standard deviation.

Under the assumption that H_0 is true, i.e. T has a Student t-distribution with n − 1 degrees of freedom, i.e., $T \sim t(n-1)$

9 Hypothesis testing (Part 2) --- single sample

9.3 Inference about the mean of a normal distribution when the variance is unknown

- For the one-sided alternative hypothesis $H_1 : \mu > \mu_0$,

$$\text{reject } H_0 \text{ if } T > t_{1-\alpha}\,,$$

where $t_{1-\alpha}$ is the $1-\alpha$ point of a $t(n-1)$ distribution, i.e. $\mathrm{P}(T \leq t_{1-\alpha}) = 1-\alpha$.

- For the one-sided alternative hypothesis $H_1 : \mu < \mu_0$,

$$\text{reject } H_0 \text{ if } T < -t_{1-\alpha}\,.$$

- For the two-sided alternative hypothesis $H_1 : \mu \neq \mu_0$,

$$\text{reject } H_0 \text{ if } T < -t_{1-\alpha/2} \text{ or } T > t_{1-\alpha/2}\,.$$

9 Hypothesis testing (Part 2) --- single sample

9.3 Inference about the mean of a normal distribution when the variance is unknown

Example. The drug 6-mP is used to treat leukaemia. A random sample of 21 patients using 6-mP were found to have an average remission time of $\bar{x} = 17.1$ weeks with a sample standard deviation of $s = 10.00$ weeks. A previously used drug treatment had a known mean remission time of $\mu_0 = 12.5$ weeks. Assuming that the remission times of patients taking 6-mP are normally distributed with both the mean μ and variance σ^2 being unknown, test at the 5% significance level whether the mean remission time of patients taking 6-mP is greater than $\mu_0 = 12.5$ weeks.

Solution: We want to test $H_0 : \mu = 12.5$ vs $H_1 : \mu > 12.5$ at the 5% significance level.

The test statistic is

$$T = \frac{\overline{x} - \mu_0}{s/\sqrt{n}} = \frac{17.1 - 12.5}{10/\sqrt{21}} = 2.108$$

Under H_0, $T \sim t(20)$. For a one-tailed test at the 5% significance level we will reject H_0 if $T > 1.725$ (from tables). Our observed value of T is greater than 1.725 and so we reject the null hypothesis that $\mu = 12.5$ at the 5% significance level and conclude that $\mu > 12.5$, i.e. the drug 6-mP improves remission times compared to the previous drug treatment.

9 Hypothesis testing (Part 2) --- single sample
9.4 Using central limit theorem

(i) **Inference about the mean of a non-normal distribution.**

Let $X_1, \ldots, X_n$ be a random sample from a non-normal distribution, where the value of the mean μ is *unknown* and that of the variance σ^2 is also *unknown*. We want to test the following hypotheses:

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu > \mu_0$$

at significance level α. We can again use the test statistic

$$Y = \frac{\overline{X} - \mu_0}{S/\sqrt{n}}$$

defined above which, by asymptotic (large n) results, has an approximate $N(0,1)$ distribution when H_0 is true ($n \geq 30$). Aside from the choice of test statistic, the rejection regions for the various versions of H_1 are otherwise identical to those defined in the case of normal data with a known variance.

9 Hypothesis testing (Part 2) --- single sample
9.4 Using the central limit theorem

(ii) **Inference about the population proportion** p.

Let $X_1, \ldots, X_n$ be a random sample of $\text{Bi}(1, p)$ random variables, where the value of p is *unknown*. We want to test the following hypotheses:

$$H_0 : p = p_0$$
$$H_1 : p > p_0$$

at significance level α. As we have seen earlier in this module, an unbiased sample estimator of the parameter p is given by

$$\widehat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}_n \, .$$

By the central limit theorem, $\widehat{p} \sim N(p, p(1-p)/n)$ approximately for large n. As a rule of thumb, $n \geq 9\max\{p/(1-p), (1-p)/p\}$ guarantees this approximation has a good degree of accuracy. A suitable test statistic is

$$Y = \frac{\widehat{p} - p_0}{\sqrt{p_0(1-p_0)/n}}$$

9 Hypothesis testing (Part 2) --- single sample
9.4 Using the central limit theorem

Here we have estimated the standard error of $\widehat{p}$ by $\sqrt{p_0(1-p_0)/n}$ which uses the value of p specified under H_0. If H_0 is true then Y has an approximate $N(0,1)$ distribution for large n. Thus, to achieve an approximate significance level of α, we reject H_0 in favour of the above H_1 if $Y > z_{1-\alpha}$.

- For the one-sided alternative hypothesis $H_1 : p < p_0$, to achieve an approximate significance level of α, we reject H_0 if $Y < -z_{1-\alpha}$.
- For the two-sided alternative hypothesis $H_1 : p \neq p_0$, to achieve an approximate significance level of α, we reject H_0 if

$$Y < -z_{1-\alpha/2} \quad \text{or} \quad Y > z_{1-\alpha/2}\,.$$

9 Hypothesis testing (Part 2) --- single sample

9.4 Using the central limit theorem

Example. A team of eye surgeons has developed a new technique for an eye operation to restore the sight of patients blinded by a particular disease. It is known that 30% of patients who undergo an operation using the old method recover their eyesight.
A total of 225 operations are performed by surgeons in various hospitals using the new method and it is found that 88 of them are successful in that the patients recover their sight. Can we justify the claim that the new method is better than the old one? (Use a 1% level of significance).

Solution: Let p be the probability that a patient recovers their eyesight following an operation using the new technique. We wish to test $H_0 : p = 0.30$ vs $H_1 : p > 0.30$ at the 1% significance level.

Our test statistic is

$$Y = \frac{\frac{88}{225} - 0.30}{\sqrt{\frac{0.30 \times 0.70}{225}}} = 2.9823$$

As a check for the approximate normality of the distribution of Y under H_0, we require $n > 9 \max\{0.429, 2.333\} = 20.997$ which is true since $n = 225$.

The approximate 1% critical value, taken from standard normal tables, is 2.3263 which is less than the observed value of Y. Hence, we reject the null hypothesis at the 1% significance level and conclude that $p > 0.30$.

Exercises:

1. *Assume we have 4 sample values 3, 6, 6, 10 that independent draw from a normal distribution* $N(\mu, \sigma^2)$. *We have null hypothesis* $H_0 : \mu = 0$, *and alternative hypothesis* $H_a : \mu \neq 0$ *(or* $H_1 : \mu \neq 0$*). We further assume* $\sigma^2 = 16$ *is known, please test* H_0 *against* H_a

2. *Assume we have 4 sample values 3, 6, 6, 10 that independent draw from a normal distribution* $N(\mu, \sigma^2)$. *We have null hypothesis* $H_0 : \mu = 0$, *and alternative hypothesis* $H_a : \mu \neq 0$ *(or* $H_1 : \mu \neq 0$*). We further assume* σ^2 *is unknown, please test* H_0 *against* H_a