Chapter 6

Estimation

parameters

μ

σ^2

P

population

مجتمع

تقدير

estimation →

Estimators

$\bar{X}$

S^2

$\hat{P}$

Sample

عينة

→ Representive

→ Random

Reference   target   study

---

* Estimation of the mean

$\mu \longrightarrow \bar{X}$ most unbiased point estimate for population mean

بدون أخطاء

* point estimate: single value estimate for a population parameter

* Central limit theorem

$\bar{X}$: Sample mean , n is large ($n \geq 30$)

NOTE : Small n is < 30

$$\bar{X} \dot{\sim} n\left(\mu, \frac{\sigma^2}{n}\right)$$

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim n(0,1)$$

Example suppose that random sample of size 100 is drawn from a population with mean 70 and SD of 20. What is the probability that SAMPLE MEAN will be:

$$\bar{X} \dot{\sim} n\left(70, \frac{20^2}{100}\right)$$

a) more than 70

$$P(\bar{X} > 70) \Rightarrow P\left(Z > \frac{70-70}{20/\sqrt{100}}\right)$$

$$= P(Z > 0) = 0.5$$

b) less than 73

$$P(\bar{X} < 73) = P\left(Z < \frac{73-70}{20/\sqrt{100}}\right)$$

$$= P(Z < 1.5)$$

$$= 0.9332$$

Example The length of pregnancies are normally distributed with mean of 268 and a standard deviation of 15 days.

a) If one pregnant women is randomly selected, find the probability that her length of pregnancy is less than 260 days

$$P(X < 260) \Rightarrow P\left(Z < \frac{260-268}{15}\right)$$

$\vdots$

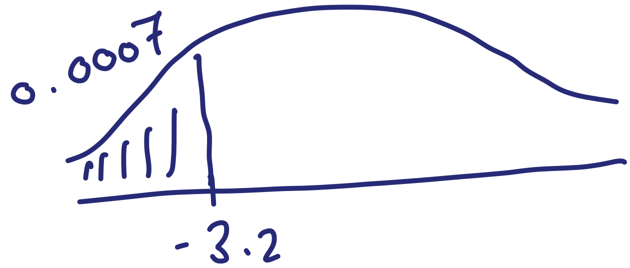
b) If 36 pregnant women are put on a special diet just before they become pregnant, find the probability that their lengths of pregnancy have a mean of that is less than 260 days

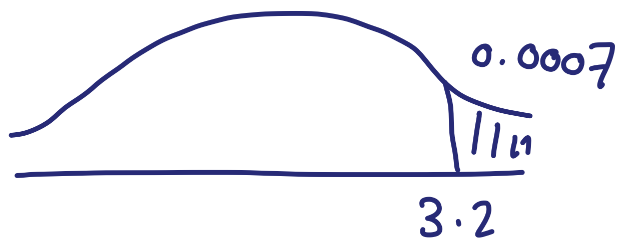
$$\bar{X} \sim n\left(268, \frac{15^2}{36}\right)$$

$$p(\bar{X} < 260)$$

$$= p\left(z < \frac{260 - 268}{15/\sqrt{36}}\right) = p(z < -3.2)$$

$$= 0.0007$$





Example The length of time, in hours, it takes an "over 40" group of people to play one soccer match is normally distributed with mean of 2 hours and standard deviation of 0.5 hours. A sample of size $n = 50$ is Randomly selected. Find the probability that the sample mean is between 1.8 hours and 2.3 hours

$$\bar{X} \sim n\left(2, \frac{0.5^2}{50}\right)$$

$$= P(1.8 < \bar{X} < 2.3)$$

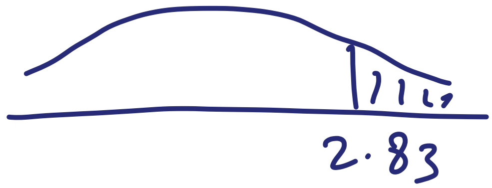
$$= P(\bar{X} < 2.3) - P(\bar{X} < 1.8)$$

$$= P\left(Z < \frac{2.3 - 2}{0.5/\sqrt{50}}\right) - P\left(Z < \frac{1.8 - 2}{0.5/\sqrt{50}}\right)$$

$$= P(Z < 4.24) - P(Z < -2.83)$$

$$= 1 - 0.0023$$

$$= 0.9977$$



* Standard error

$$\bar{X} \dot{\sim} N\left(\mu, \frac{\sigma^2}{n}\right)$$

Standard deviation = Standard error = $\frac{\sigma}{\sqrt{n}}$

كتاب

Example Compute the standard error of the mean for the following birth weights:

(97, 125, 62, 120, 132, 135, 118, 137, 126, 118)

الحل

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{22.4}{\sqrt{10}}$$

كتاب

Example Suppose that for certain large group of indivisuals in Jordan, the mean hemoglobin level in the blood is 21 grams per ml, and the standard deviation is 2 g/ml. If randomly selected sample with $n = 25$ indivisuals. what is the probability that the sample mean will be greater than 21.3 g/ml assuming that the hemoglobin level is normally distributed?

$$\bar{X} \sim n\left(21, \frac{2^2}{25}\right)$$

$$P(\bar{X} > 21.3) = P\left(Z > \frac{21.3 - 21}{2/\sqrt{25}}\right)$$

$$= P(Z > 0.75)$$

$$= 0.2266$$

Example Compute the probability that the mean birth weight from a sample of 10 infants from the Boston City Hospital population will fall between 98 and 126 oz. If the birthweights is normally distributed with a population mean of 112 and standard deviation of 20.6 (Assuming Central limit theorem is applied)

$$\bar{x} \sim n\left(112, \frac{20.6^2}{10}\right)$$

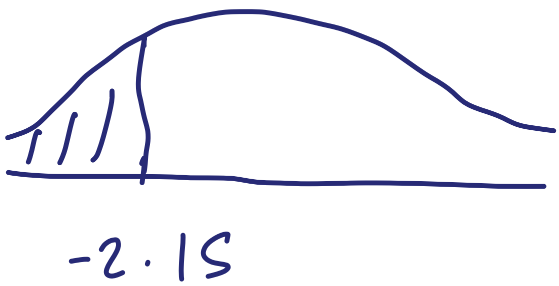
$$= P(98 < \bar{x} < 126)$$

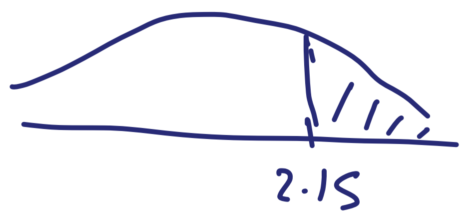
$$= P(\bar{x} < 126) - P(\bar{x} < 98)$$

$$= P\left(z < \frac{126-112}{20.6/\sqrt{10}}\right) - P\left(z < \frac{98-112}{20.6/\sqrt{10}}\right)$$

$$= P(z < 2.15) - P(z < -2.15)$$

$$= 0.9842 \quad - 0.0158 = -$$





Example Suppose that a large dairy production company in Jordan, the mean age of employees is 36.2 year, and the standard deviation is 3.7 years. Assume that the variable is normally distributed, then answer the following:

a) If an employee from the company is Randomly selected, find the probability that his/her age will be between 36 and 37.5 years?

$$= P(36 < X < 37.5)$$

0.1567
Ans.

$$= P(X < 37.5) - P(X < 36)$$

$$= P\left(Z < \frac{37.5 - 36.2}{3.7}\right) - P\left(Z < \frac{36 - 36.2}{3.7}\right)$$

b) If a random sample of 15 employees is selected, find the probability that the mean age will be between 36 and 37.5 years?

$$\bar{X} \sim n\left(36.2, \frac{3.7^2}{15}\right)$$

$$= P(36 < \bar{X} < 37.5)$$

$$= P(\bar{x} < 37.5) - P(\bar{x} < 36)$$

$$= P\left(z < \frac{37.5 - 36.2}{3.7/\sqrt{15}}\right) - P\left(z < \frac{36 - 36.2}{3.7/\sqrt{15}}\right)$$

$$= P(z < 1.36) - P(z < -0.21)$$

$$= \underline{\underline{0.4963}}$$

Example Suppose that the weights of Certain population are normally distributed with mean $\mu = 70$ KG. and $SD = 10$ KG

If a sample size $n = 25$ persons is to be drawn, what is the probability :

i) the average weight will be less than 75 KGs. $P(\bar{X} < 75) \Rightarrow P\left(Z < \frac{75-70}{10/\sqrt{25}}\right)$

$= P(Z < 2.5)$

$= 0.9938$

ii) the total weights exceeds 1800 KGs

$$P\left(\frac{\sum x}{n} > \frac{1800}{25}\right)$$

$$P(\bar{X} > 72) \Rightarrow P\left(Z > \frac{72-70}{10/\sqrt{25}}\right)$$

⋮

Example: Suppose that the mean and SD of orange boxes are 10 and 2 Respectively. If 100 boxes are to be selected in a car with threshold 1000 KG what is the probability that the car will break down?

$$\bar{X} \sim n\left(10, \frac{2^2}{100}\right)$$

$$= P\left(\sum X > 1000\right)$$

$$= P(\bar{X} > 10) = P\left(Z > \frac{10-10}{2/\sqrt{100}}\right)$$

$$= P(Z > 0.5) = 0.50$$

* Confidence interval

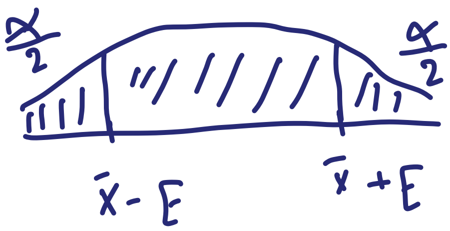
فترة الثقة

① for mean

$$\mu \Rightarrow (\bar{X} \pm E)$$

$$E = Z_{\frac{\alpha}{2}} * \frac{\sigma}{\sqrt{n}} \quad SE$$

$$CI = (1-\alpha)$$



سؤال قديم

Example A college admissions director wishes to estimate the mean age of all students currently enrolled. In a random sample of 20 students, the mean age is found to be 22.9 years. From past studies, the standard deviation is known to be 1.5 years, and the population is normally distributed. Construct a 95% confidence interval for the population <u>mean</u> age.

$n = 20$

$\bar{x} = 22.9$

$\sigma = 1.5$

$CI = 0.95$

$\mu \Rightarrow \bar{x} \pm E$

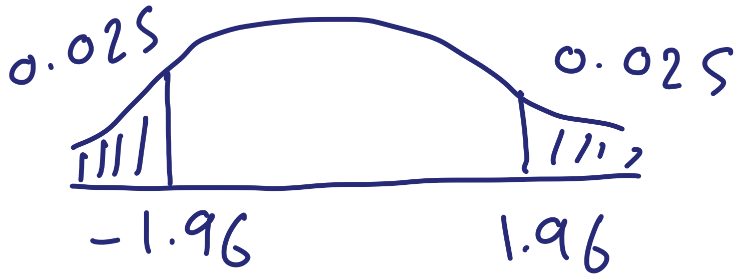
$$E = z_{\frac{\alpha}{2}} * \frac{\sigma}{\sqrt{n}} = 1.96 * \frac{1.5}{\sqrt{20}} = 0.657$$

الحل

$$0.95 = 1 - \alpha$$

$$0.05 = \alpha$$

$$\boxed{0.025 = \frac{\alpha}{2}}$$



$$\boxed{z_{\frac{\alpha}{2}} = \pm 1.96}$$

$$(22.9 - 0.657,\ 22.9 + 0.657)$$

$$(22.24,\ 23.55)$$

Example Suppose average DoorDash delivery times are normally distributed with a population standard deviation of 6 minutes. A random sample of $n=40$ restaurants is taken and has a sample mean delivary time of $\bar{x} = 42$ minutes. Find a 90% confidence interval for the population mean delivary time

$\sigma = 6$

$n = 40$

$\bar{x} = 42$

$$\mu \Rightarrow \bar{x} \pm E$$

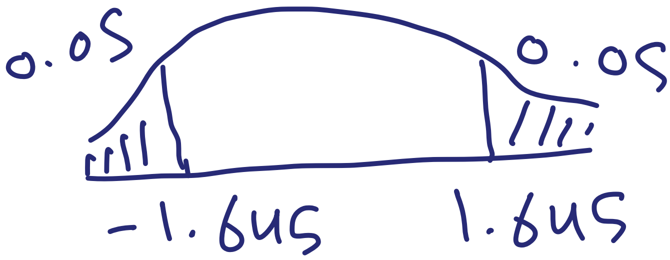
$$E = z_{\frac{\alpha}{2}} * \frac{\sigma}{\sqrt{n}} = 1.645 * \frac{6}{\sqrt{40}}$$
$$= 1.5606$$

CI = 0.90

$$1-\alpha = 0.90$$

$$\alpha = 0.10$$

$$\boxed{\frac{\alpha}{2} = 0.05}$$



$$\boxed{Z_{\frac{\alpha}{2}} = \pm 1.645}$$

$$(42 - 1.5606,\ 42 + 1.5606)$$

$$(40.439,\ 43.561)$$

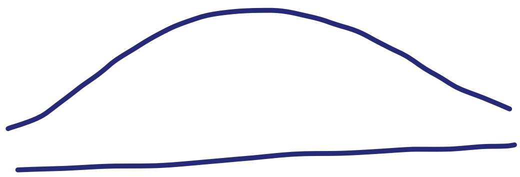
CI for mean

- σ known
  $(\bar{X} \pm E)$
  $$E = Z_{\frac{\alpha}{2}} * \frac{\sigma}{\sqrt{n}}$$
  ✓
- σ unknown, $n < 30$
  $(\bar{X} \pm E)$
  $$E = t_{\frac{\alpha}{2}} * \frac{S}{\sqrt{n}}$$
- σ unknown, $n > 30$
  $(\bar{X} \pm E)$
  $$E = Z_{\frac{\alpha}{2}} * \frac{S}{\sqrt{n}}$$

---

* T - distribution :

(1) σ unknown (2) $n < 30$



(1) μ = mode = Q_2 (2) Symmetric

(3) Total area = 1

جديدة

(4) Tails are thicker (أسمك) than standard normal

distribution

$\bar{X}$

σ : population standard deviation

S : sample standard deviation

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t\,(n-1)$$

degree of freedom

d.f.

Example: suppose that the weights of newborn babies are normally distributed with mean 3 KGs. A Random sample of size 10 is taken and showed that its standard deviation is 2.

$\mu = 3$

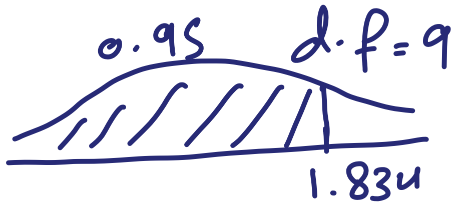
$n = 10$

$S = 2$

A) find the probability that the sample average will be less than 4.16 KGs?

$$P(\bar{X} < 4.16) = P\left(T < \frac{4.16 - 3}{2/\sqrt{10}}\right)$$

$$= P\,(T < 1.834) = 0.95$$

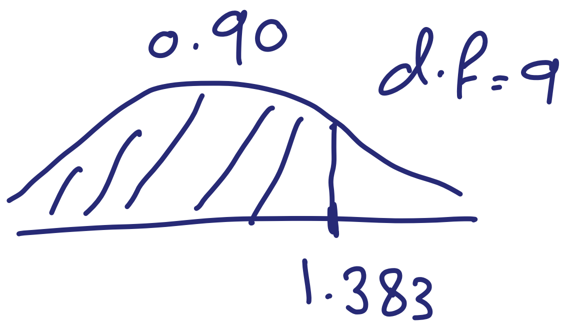


B) what is 90^{th} percentile of the distribution of $\bar{X}$?

$$P(\bar{X} < P_{90}) = 0.90$$

$$= P\left(T < \frac{P_{90} - 3}{2/\sqrt{10}}\right) = 0.90$$

$$\frac{P_{90} - 3}{2/\sqrt{10}} = 1.383$$



$$\boxed{P_{90} = 3.874}$$

Example If $T \sim t(10)$, find:

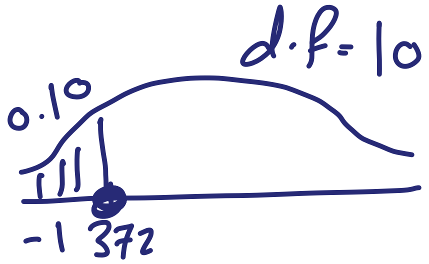
a) a 95^{th} percentile of T

$$P(T < P_{95}) = 0.95$$

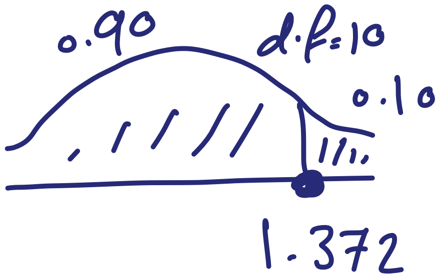
$$P_{95} = 1.812$$

B) a 10^{th} percentile of T

$$P(T < P_{10}) = 0.10$$



$$\boxed{P_{10} = -1.372}$$

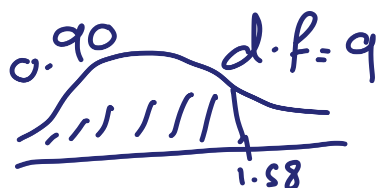


Example If $X_1, X_2, \ldots, X_{10} \sim n(5, \sigma^2)$

$S = \underline{2}$, find:

1) $$P(\bar{X} < 6) = P\left(T < \frac{6-5}{2/\sqrt{10}}\right)$$

$$= P(T < 1.58) = 0.90$$

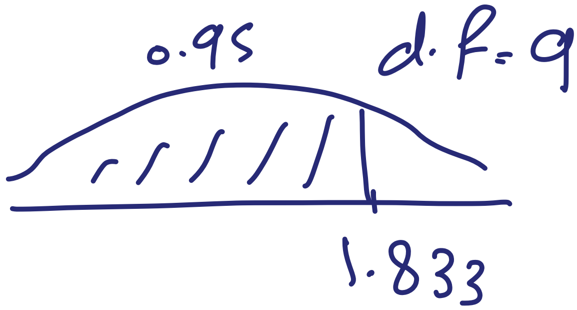


ii) the 95^{th} percentile of $\bar{X}$

$$P(\bar{X} < P_{95}) = 0.95$$

$$= P\left(T < \frac{P_{95} - 5}{2/\sqrt{10}}\right) = 0.95$$

$$\frac{P_{95} - 5}{2/\sqrt{10}} = 1.833$$



$$\boxed{P_{95} = 6.16}$$

★ CI for mean (σ unknown)

$n < 30 \Rightarrow T$ $\boxed{E = t_{\frac{\alpha}{2}} * \frac{S}{\sqrt{n}}}$

$n > 30 \Rightarrow n$ $\boxed{E = z_{\frac{\alpha}{2}} * \frac{S}{\sqrt{n}}}$

Example A Sample of 10 orthopaedic Surgeon had a mean weight of 240 pounds and standard deviation of 25. Construct a 95% Confidence interval for the population mean weight.

$n = 10$

$\bar{X} = 240$

$S = 25$

$CI = 0.95$

$\mu \Rightarrow \bar{X} \pm E$

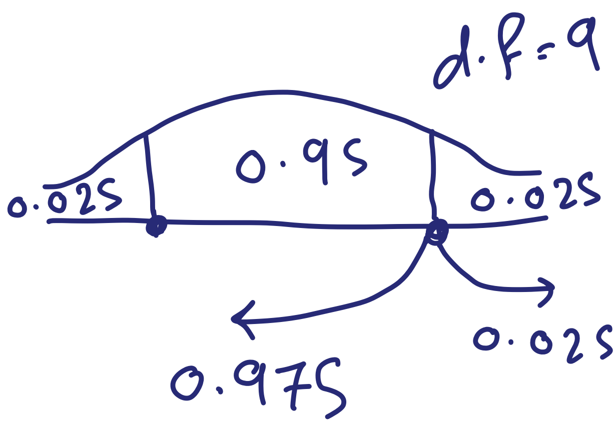
$$E = t_{\frac{\alpha}{2}} * \frac{S}{\sqrt{n}} = 2.262 * \frac{25}{\sqrt{10}}$$

$$= 17.88$$

$$1 - \alpha = 0.95$$

$$\alpha = 0.05$$

$$\frac{\alpha}{2} = 0.025$$



$$t_{\frac{\alpha}{2}} = \pm 2.262$$

$$(240 - 17.88,\ 240 + 17.88)$$

$$(222.117,\ 257.88)$$

Example you randomly select and weigh 16 samples of an allergy medicine. The sample standard deviation is 1.20 mg, with mean 20. Assuming the weights are normally distributed. Construct a 99% confidence interval for the population mean.

$n = 16$

$S = 1.2$

$\bar{X} = 20$

$CI = 0.99$

$$\mu \Rightarrow \bar{X} \pm E$$

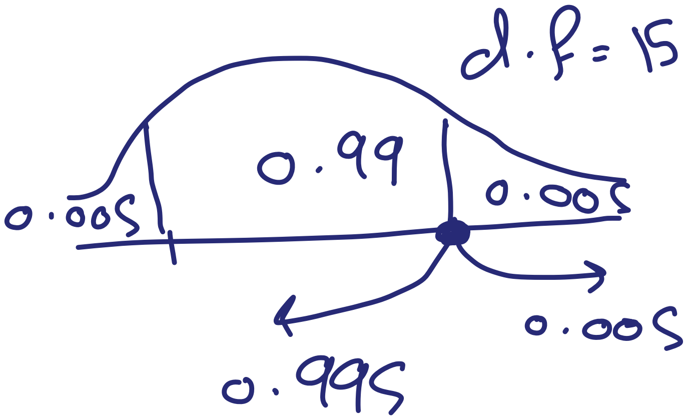
$$E = t_{\frac{\alpha}{2}} * \frac{S}{\sqrt{n}} = 2.947 * \frac{1.2}{\sqrt{16}}$$

$$= 0.8841$$

$$1 - \alpha = 0.99$$

$$\alpha = 0.01$$

$$\frac{\alpha}{2} = 0.005$$



$$\boxed{t_{\frac{\alpha}{2}} = \pm 2.947}$$

$$(20 - 0.8841 , 20 + 0.8841)$$

$$(19.1159 , 20.8841)$$

Example A local physical therapist and her acupuncturist are interested in the use of acupuncture for pain relief. They decide to conduct an informal study and measure sensory rate for n = 15 randomly selected patients. The population is assumed to be normally distributed. The sensory rate results reveal a mean pain level of $\bar{x} = 8.23$ and $S = 1.67$. Use the sample data to construct a 90% confidence interval for the true mean sensory rate for the population.

$n = 15$

$\bar{x} = 8.23$

$S = 1.67$

$CI = 0.90$

$$\mu \ni \bar{x} \pm E$$

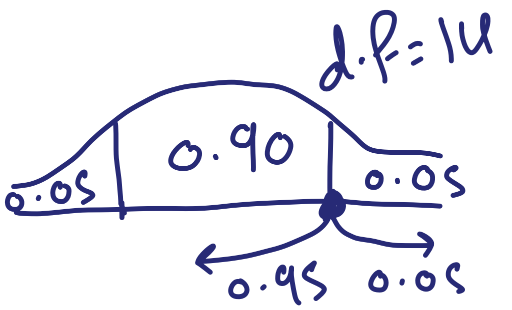
$$E = t_{\frac{\alpha}{2}} * \frac{S}{\sqrt{n}} = 1.761 * \frac{1.67}{\sqrt{15}}$$

$$= 0.7593$$

$$1 - \alpha = 0.90$$

$$\alpha = 0.10$$

$$\frac{\alpha}{2} = 0.05$$



$$\boxed{t_{\frac{\alpha}{2}} = \pm 1.761}$$

$$(8.23 - 0.7593,\ 8.23 + 0.7593)$$

$$(7.471,\ 8.989)$$

Example A survey in 2020 found that a sample of 67 respondents spent an average of 2.5 hours each day on social media, with standard deviation of 0.35 hours. Construct a 99% confidence interval for the average daily use of social media for the population.

$n = 67$

$\bar{x} = 2.5$

$S = 0.35$

$CI = 0.99$

$$\mu \Rightarrow \bar{x} \pm E$$

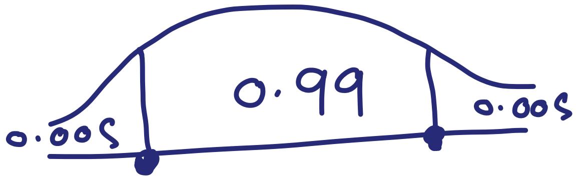
$$E = Z_{\frac{\alpha}{2}} * \frac{S}{\sqrt{n}} = 2.575 * \frac{0.35}{\sqrt{67}}$$

$$= 0.11$$

$$1 - \alpha = 0.99$$

$$\alpha = 0.01$$

$$\boxed{\frac{\alpha}{2} = 0.005}$$



$$\boxed{Z_{\frac{\alpha}{2}} = \pm 2.575}$$

$$(2.5 - 0.11 \, , \, 2.5 + 0.11)$$

$$(2.39 \, , \, 2.61)$$

---

* Estimation of the binomial Proportion :

النسبة

$$\boxed{P \Rightarrow \hat{P} \pm E}$$

$$\boxed{\hat{p} = \frac{x}{n} = \frac{\text{المراد دراسته}}{\text{العدد الكلي}}}$$

$$\boxed{E = Z_{\frac{\alpha}{2}} * \sqrt{\frac{\hat{p}\,\hat{q}}{n}}} \qquad \boxed{\hat{p} + \hat{q} = 1}$$

$$\hat{p} \dot{\sim} n\left(p, \frac{pq}{n}\right)$$

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} \sim n(0,1)$$

Example suppose that 10% of certain population are defective. If 400 items are drawn from the production, what is the probability that the sampe proportion will be:

$P = 0.10$
$n = 400$

$$\hat{p} \dot{\sim} n\left(0.10, \frac{0.10 * 0.90}{400}\right)$$

A) more than 12%

$$P(\hat{p} > 0.12) = P\left(Z > \frac{0.12 - 0.10}{\sqrt{\frac{0.10 * 0.90}{400}}}\right)$$

$$= P(Z > 1.33)$$

$$= 0.0918$$

B) Between 9% and 11%

$$= P(0.09 < \hat{p} < 0.11)$$

$$= P(\hat{P} < 0.11) - P(\hat{P} < 0.09)$$

$$= P\left(Z < \frac{0.11 - 0.10}{\sqrt{\frac{0.10 * 0.90}{400}}}\right) - P\left(Z < \frac{0.09 - 0.10}{\sqrt{\frac{0.10 * 0.90}{400}}}\right)$$

$$\vdots$$

سنوات

Example suppose that 90% of the university of Jordan students pass calculas 101. In a sample of 200 students taking calculas 101, what is the probability that the proportion of those who will pass is less than 85%.

$P = 0.90$

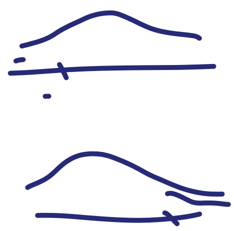
$n = 200$

$$= P(\hat{P} < 0.85)$$

$$= P\left(Z < \frac{0.85 - 0.90}{\sqrt{\frac{0.90 * 0.10}{200}}}\right)$$

$$= P(Z < -2.36)$$

$$= P(Z > 2.36)$$



$$= 0.0091$$

#

Example It was believed in the Arab world that 50% of persons are Smoking. During the year 2000, a sample of 1000 persons showed that the number of smokers is 620. Establish 95% Confidence interval for proportion of smoker.

$P = 0.50$

$n = 1000$

$X = 620$

$\hat{p} = \frac{620}{1000} = \underline{0.62}$

$CI = 0.95$

$P \Rightarrow \hat{p} \pm E$

$$E = Z_{\frac{\alpha}{2}} * \sqrt{\frac{\hat{p}\,\hat{q}}{n}}$$

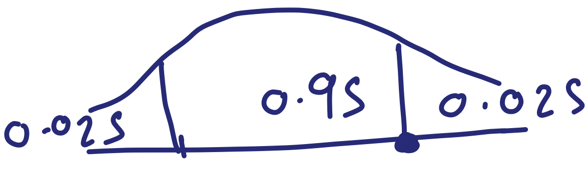
$$= 1.96 * \sqrt{\frac{0.62 * 0.38}{1000}} = 0.03$$

$1 - \alpha = 0.95$

$\alpha = 0.05$

$\frac{\alpha}{2} = 0.025$

$$\boxed{Z_{\frac{\alpha}{2}} = \pm 1.96}$$



$$(0.62 - 0.03,\ 0.62 + 0.03)$$

$$(0.59,\ 0.65)$$

Example A random sample of 200 students from the UJ had 45 smoker. Set a 90% CI for proportion of smoker in the UJ.

$$P \Rightarrow \hat{P} \pm E$$

$n = 200$

$X = 45$

$CI = 0.90$

$\hat{P} = \frac{45}{200} = 0.225$

$$E = Z_{\frac{\alpha}{2}} * \sqrt{\frac{\hat{p}\,\hat{q}}{n}}$$

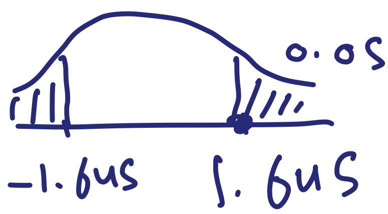
$$= 1.645 * \sqrt{\frac{0.225 * 0.775}{200}} = 0.05$$

$$1 - \alpha = 0.90$$

$$\alpha = 0.10$$

$$\frac{\alpha}{2} = 0.05$$

$$\boxed{Z_{\frac{\alpha}{2}} = \pm 1.645}$$



$$(0.225 - 0.05 \, , \, 0.225 + 0.05)$$

$$(0.175 \, , \, 0.275)$$

Example For a class project a political Science student at a UJ wants to estimate the percent of students who are registered voters. He surveys 500 students and find that 300 are registered voters. Construct a 90% CI for the percent of students who are registered voters.

$n = 500$

$X = 300$

$CI = 0.90$

$\hat{p} = \frac{300}{500} = 0.6$

$$P \Rightarrow \hat{p} \pm E$$

$$E = z_{\frac{\alpha}{2}} * \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$= 1.645 * \sqrt{\frac{0.6 * 0.4}{500}}$$

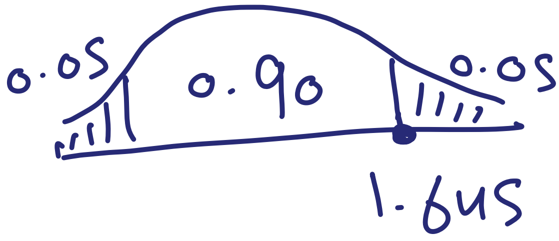
$$= 0.04$$

$$1 - \alpha = 0.90$$

$$\alpha = 0.10$$

$$\frac{\alpha}{2} = 0.05$$

$$\boxed{Z_{\frac{\alpha}{2}} = \pm 1.645}$$



$$(0.6 - 0.04,\ 0.6 + 0.04)$$

$$(0.5639,\ 0.64)$$

---

* Finding the sample size. (n)

1. For mean (σ known)
2. For proportion

1. For the mean (σ known)

$$E = Z_{\frac{\alpha}{2}} * \frac{\sigma}{\sqrt{n}} \Rightarrow n = \left(\frac{Z_{\frac{\alpha}{2}}}{E}\right)^2 * \sigma^2$$

Example A researcher wants to estimate the average weight loss of people who are in a new diet plan. In previous study, the population standard deviation σ of a weight losses is about 5 KGs. How large a sample should be to estimate the mean weight loss by a 95% CI to within 1.5 KGs?

$\sigma = 5$

$n = ?$

$CI = 0.95$

$$n = \left(\frac{Z_{\frac{\alpha}{2}}}{E}\right)^2 * \sigma^2$$

$$= \left(\frac{1.96}{1.5}\right) * 5^2 = 42.68 \simeq 43$$

$E = 1.5$

$1 - \alpha = 0.95$

$\alpha = 0.05$

$\boxed{\frac{\alpha}{2} = 0.025}$

$\boxed{Z_{\frac{\alpha}{2}} = \pm 1.96}$

NOTES

★ كلما زاد n يقل E

$$n = \left(\frac{Z_{\frac{\alpha}{2}}}{E}\right)^2 * \sigma^2$$

level of significance : Significance level : α ★

level of confidence : Confidence level : $1 - \alpha$ ★

$$\uparrow CI = \uparrow (1-\alpha)$$

Example which of the following Confidence level produce the widest Confidence interval

a) 90% b) 95%

c) 98% D) 99%

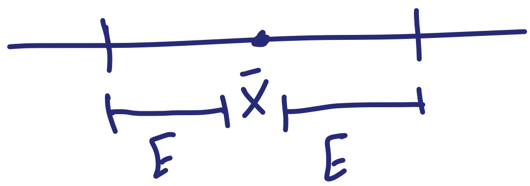
$$\uparrow (1-\alpha) = \uparrow CI$$

Ans. D

كتاب

NOTE

1. As sample size n increase, the length of CI decrease



$\uparrow n \quad \downarrow E$

② As the Standard deviation increase the length of CI increase

$$\uparrow E = Z_{\frac{\alpha}{2}} * \frac{\uparrow \sigma}{\sqrt{n}}$$

③ As α decrease, the length of CI increase

---

② For proportion

$\hat{p}$ مش معروفة

$\hat{p} = 0.5$

$$\boxed{n = \left(\frac{Z_{\frac{\alpha}{2}}}{E}\right)^2 * \hat{p}\hat{q}}$$

$\hat{p}$ معروفة

نعوضها نفسها

Example You are running a political campaign and wish to estimate, with 95% CI, the proportion of Registered voters who will vote for your Candidate. your estimate must be accurate within 3% of the population proportion. Find the sample size when: $CI = 0.95$, $E = 0.03$

① no preliminary estimate is available

$\hat{p} = 0.5$, $\hat{q} = 0.5$

$$n = \left(\frac{Z_{\frac{\alpha}{2}}}{E}\right)^2 * \hat{p}\hat{q} = \left(\frac{1.96}{0.03}\right)^2 * 0.5 * 0.5 = 1067$$

② a preliminary estimate gives $\hat{p} = 0.31$

$\hat{p} = 0.31$, $\hat{q} = 0.69$

$$n = \left(\frac{1.96}{0.03}\right)^2 * 0.31 * 0.69 = 913$$

chapter 7 Test of hypothesis

اختبار الفرضيات

⇒ for one sample ⇐

* null hypothesis (H_0): is a statistical

الفرضية الأصلية

hypothesis that contains a statement of

المساواة

equality

$\mu > 30$
$\mu < 30$
$\mu \neq 30$

$\mu = 30$

الفرضية البديلة

* Alternative hypothesis (H_1) (H_a): is a complement of null hypothesis. It is true if H_0 is false

* Null hypothesis table (H_0)

H_0	True	False
accept	✓	Type 2 error (β)
Reject	Type 1 error (α)	✓

α : $P(\text{Type 1 error})$

$= P(\text{reject } H_0 \mid H_0 \text{ is True})$

$=$ Significance level

β : $P(\text{Type 2 error})$

$= P(\text{accept } H_0 \mid H_0 \text{ is false})$

$1 - \beta$: power of the test

كيف نختبر الفرضيات ؟

ضريبه
العفاى

H_0 : Vs. H_1 :

<u>test statistic</u>

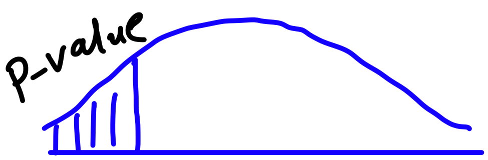
P-value

Rejection Region

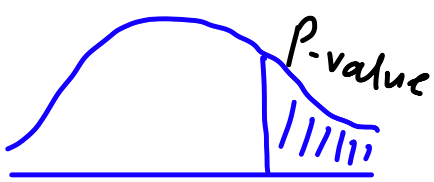
الدول
الحسافى
إليها

---

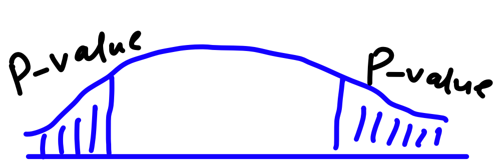
① P-value عبارة عن مساحة



left tailed test



Right tailed test



2-tailed test

① Left tailed test $\Rightarrow$ P(area on the left)

② Right tailed test $\Rightarrow$ P(area on the Right)

③ 2-tailed test $\Rightarrow$ 2 * P(area on one side)

NOTE

more than $>$ Right tailed test

Less than $<$ Left tailed test

different $\neq$ 2 tailed test

* متى نحل على P-value ؟

⇐ إذا السؤال طلب الحل على P-value

NOTE

P-value $> \alpha$ accept H_0

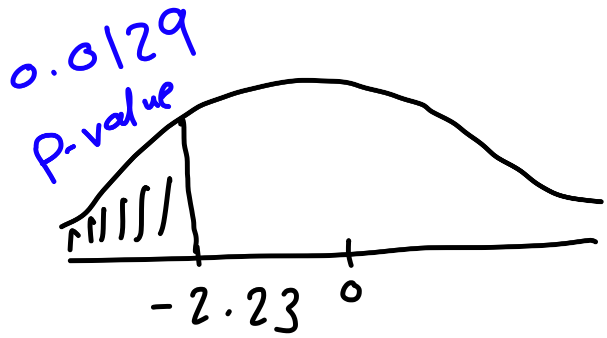
P-value $\leq \alpha$ reject H_0

NOTE

Cases of test of hypothesis:

1. accept H_0 (H_0 is true)
2. reject H_0, accept H_1

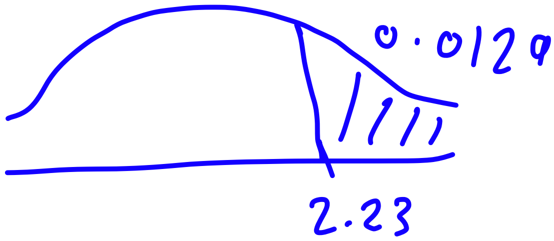
Example Find the p-value for left tailed hypothesis test with a standarized test statistics $z = -2.23$. Decide whether to reject H_0 when the level of significance is $= 0.01$



$\alpha = 0.01$

P-value $= 0.0129$

$\alpha = 0.0100$

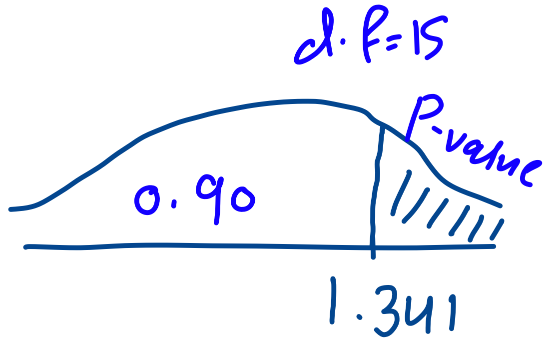


P-value $> \alpha$

∴ we accept H_0 and reject H_1

Example A General surgery research about the effect of dexamethasone injection in treating appendicitis, If you know that the test statistics given are $t = 1.341$, calculate the P-value for a right tailed test using d.f of 15

لذا P-value = 0.10



∴ ∴ ∴

* test statistic (test stat)

① for mean

σ known

$$Z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

σ unknown $\boxed{n < 30}$

$$t = \frac{\bar{x} - \mu}{S / \sqrt{n}}$$

σ unknown $\boxed{n \geq 30}$

$$Z = \frac{\bar{x} - \mu}{S / \sqrt{n}}$$

Example: In auto racing, a pit stop is where a racing vehicle stops for new tires, fuel, repairs, and other mechanical adjustments. The efficiency of a pit crew that makes these adjustments can affect the outcome of a race. A pit crew claims that its mean pit stop time (for 4 new tires and fuel) is less than 13 seconds. A random sample of 32 pit stop times has a sample mean of 12.9 seconds. Assume the population standard deviation is 0.19 second. Is there enough evidence to support the claim at a = 0.01? Use a P-value.

$n = 32$

$\bar{x} = 12.9$

$\sigma = 0.19$

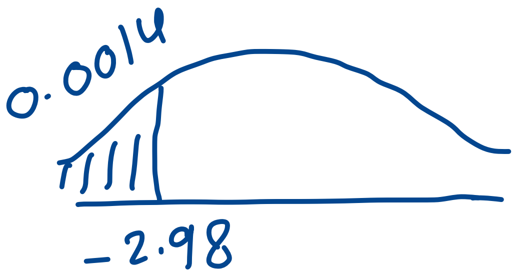
$\alpha = 0.01$

$H_0: \mu \geq 13$ Vs. $H_1: \mu < 13$

<u>test stat</u>

$$z = \frac{12.9 - 13}{0.19/\sqrt{32}}$$

$$= -2.98$$



$$P\text{-value} = 0.0014$$
$$\alpha = 0.0100$$

$P\text{-value} < \alpha$

$\therefore$ we reject H_0, and accept H_1

Example: Homeowners claim that the mean speed of automobiles traveling on their street is greater than the speed limit of 35 miles per hour. A random sample of 100 automobiles has a mean speed of 36 miles per hour. Assume the population standard deviation is 4 miles per hour. Is there enough evidence to support the claim at a = 0.05? use p - value

$n = 100$

$\bar{x} = 36$

$\sigma = 4$

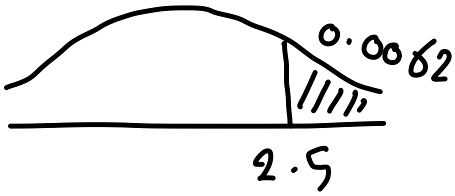
$\alpha = 0.05$

$H_0: \mu \le 35$ Vs. $H_1: \mu > 35$

test stat

$$z = \frac{36 - 35}{4/\sqrt{100}}$$

$$= 2.5$$



$$P\text{-value} = 0.0062$$
$$\alpha = 0.0500$$

$$P\text{-value} < \alpha$$

$\therefore$ we reject Ho and accept H_1

Example The average weight of a dumbbel in a gym is $\overset{\mu}{90}$ lbs. However, a physical trainer believes that the average weight might be heigher

A random sample of 5 dumbbells with an average weight of 110 lbs and a standard deviation of 18 lbs. Using hypothesis testing, check if the physical trainer's claim can be supported for 5% significance level? (use p-value)

$n = 5$

$\bar{x} = 110$

$S = 18$

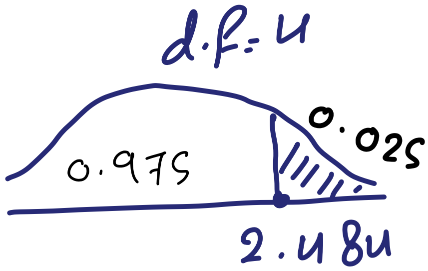
$\alpha = 0.05$

$H_0 : \mu \leq 90$ Vs. $H_1 : \mu > 90$

<u>test stat</u>

$$t = \frac{110 - 90}{18/\sqrt{5}}$$

$$= 2.484$$



$$\text{P-value} = 0.025$$

$$\alpha = 0.050$$

$$\text{P-value} < \alpha$$

$\therefore$ we reject Ho and accept H_1

② Rejection Region
منطقة الرفض

Example: the mean cholesterol levels in general examination are normally distributed. A sample of 16 persons is taken under a test with mean $\bar{x}$ = 220 mg/ dL and standard deviation s = 25 mg/ dL. Test of 1% significance level that the mean cholesterol level is less than 230 mg / dL.

$n = 16$

$\bar{x} = 220$

$S = 230$

$\alpha = 0.01$

مهمة

$H_0 : \mu \geq 230$ Vs. $H_1 : \mu < 230$

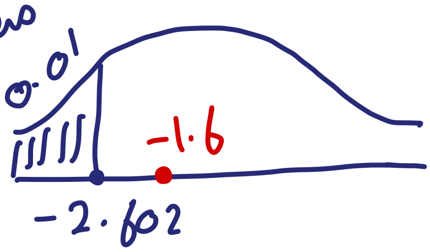
Test stat

$$t = \frac{220 - 230}{220/\sqrt{16}}$$

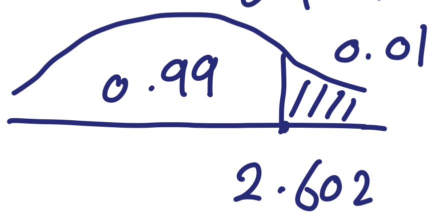
$$= -1.6$$

$df = 15$

منطقة الرفض

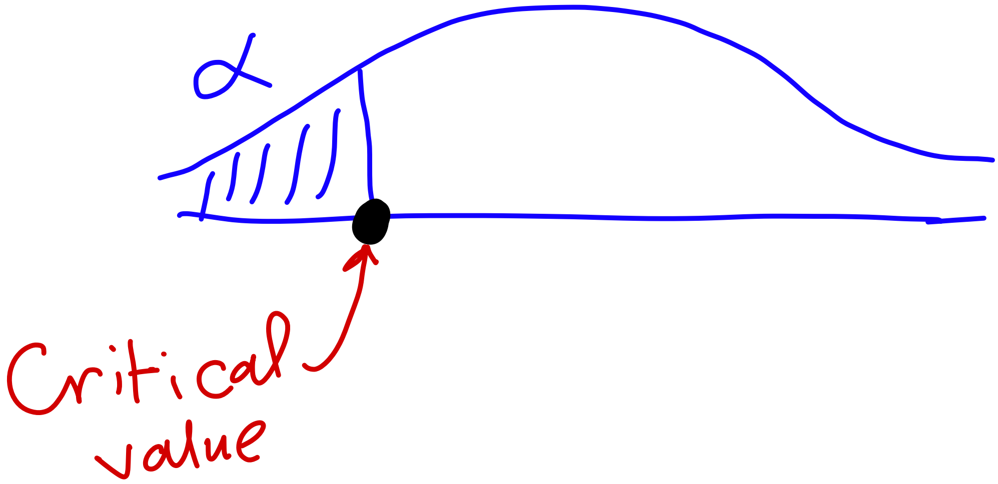


$d.f = 15$



$\therefore$ we accept H_0 and reject H_1

NOTE in Rejection Region method



Example A Random sample of 400 people with a proffisional degree taken showed that their mean monthly salary is 450, whith standard deviation of 100 JDs. test of 5% significance level that the mean monthly salary is different from 460 JDs

$n = 400$

$\bar{X} = 450$

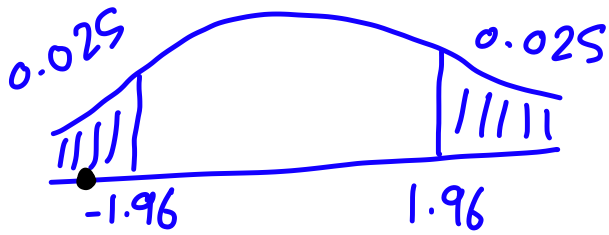
$S = 100$

$\alpha = 0.05$

$H_0: \mu = 460$ Vs. $H_1: \mu \neq 460$

test stat

$$Z = \frac{450 - 460}{100/\sqrt{400}} = -2$$



$\alpha = 0.05$

$\boxed{\frac{\alpha}{2} = 0.025}$

$\therefore$ So, we reject H_0 and accept H_1

Example: Blood Glucose level for obese patients have a mean of 100 (μ) with standard deviation of 15. A researcher thinks that the diet high in Raw Cornstarch will have a positive or negative effect on blood Glucose levels. A sample of 30 patients who have tried the raw Cornstarch diet have a mean Glucose levels of 140. Test the hypothesis that the raw corn starch had an effect using 10% significance level

$n = 30$

$\bar{x} = 140$

$\alpha = 0.10$

$H_0: \mu = 100$ Vs. $H_1: \mu \neq 100$

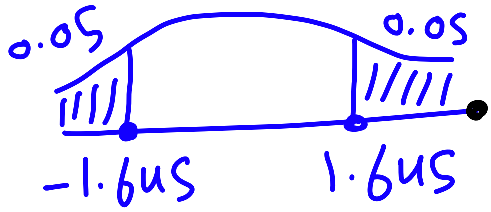
test stat:

$\sigma = 15$

$$Z = \frac{140 - 100}{15/\sqrt{30}} = 14.60$$

$\alpha = 0.10$

$\frac{\alpha}{2} = 0.05$



$\therefore$ we reject H_0 and accept H_1

Example You work in the HR department at a large company and you are currently working in the expenses department. You want to test whether you have set your employee monthly allowances correctly. In the past, it was believed that the average claim was μ 500 \$ with a standard deviation of σ 150 \$. However, you believe this may have increased due to inflation. You want to test if the monthly allownecces should be increased. A random sample is taken of 40 employees and give a mean monthly claim of 640 $ using 1 % significance level

$n = 40$

$\bar{X} = 640$

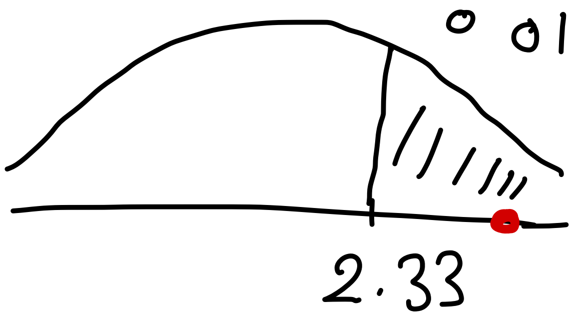
$\sigma = 150$

$\alpha = 0.01$

$H_0: \mu \leq 500$ Vs. $H_1: \mu > 500$

test stat

$$Z = \frac{640 - 500}{150/\sqrt{40}} = 5.903$$



$\therefore$ we reject H_0 and accept H_1

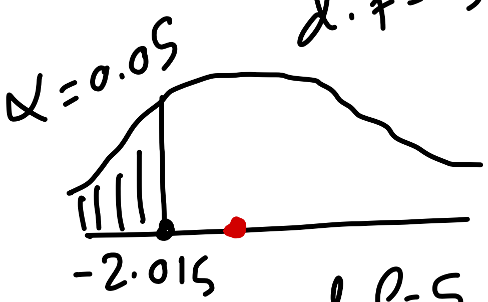
Example The average score of a class is 90 (μ). However, a teacher believes that the average score might be lower. The scores of 6 studendents was randomly measured and showed mean of 82 with standard deviation of 18, with 0.05 significance level. use the Hypothesis testing to check if the claim is true.

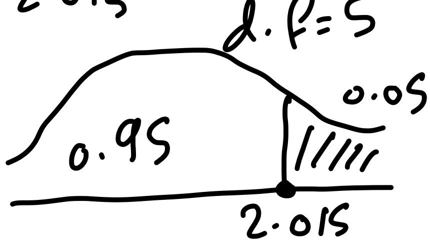
$n = 6$
$\bar{x} = 82$
$s = 18$
$\alpha = 0.05$

$H_0: \mu \geq 90$ Vs. $H_1: \mu < 90$

test stat

$$t = \frac{82 - 90}{18/\sqrt{6}} = -1.088$$





∴ we accept H_0 and reject H_1

Guidelines for Judging the Significance of a *p*-Value

If $.01 \le p < .05$, then the results are *significant*.
If $.001 \le p < .01$, then the results are *highly significant*.
If $p < .001$, then the results are *very highly significant*.
If $p > .05$, then the results are considered *not statistically significant* (sometimes denoted by NS).
However, if $.05 < p < .10$, then a trend toward statistical significance is sometimes noted.

$$\underset{\leftarrow \text{VHS}}{}\ 0.001\ \underbrace{\qquad}_{\text{HS}}\ 0.01\ \underbrace{\qquad}_{\text{S}}\ 0.05\ \xrightarrow[\text{NS}]{}\ 0.10$$

* Estimation of the sample size

for mean : $n = \left(\dfrac{z_{\frac{\alpha}{2}}}{E}\right)^2 * \sigma^2$

CI

$$\underbrace{\bar{x}-E \quad \bar{x} \quad \bar{x}+E}_{}$$

Length $= 2 * E$

$$\boxed{L = 2 * E}$$

$$n = \left(\frac{2 Z_{\frac{\alpha}{2}}}{L}\right)^2 * \sigma^2$$

$$\boxed{n = 4 * \left(Z_{\frac{\alpha}{2}} * \sigma\right)^2 / L^2}$$

---

* Test for proportion

$P \Rightarrow \hat{P}$

$H_0 : P = P_0$ Vs. $H_1 : P \neq P_0$

<u>test stat</u>

<u>Continuity correction</u>

$$Z_{corr} = \frac{|\hat{p} - p| - \frac{1}{2n}}{\sqrt{\frac{pq}{n}}}$$

$$\boxed{npq \geq 5}$$

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

$$\boxed{npq < 5}$$

مثال

Example Consider a breast Cancer problem were we interested in the effect of having a family history of breast Cancer on the incidence of $\boxed{\hat{p} = 0.04}$ women ages (50-54) sampled whose mother had breast Cancer at some time in their lives. Given large study with $\boxed{n = 10000}$ assuming the prevelance rate of breast Cancer for US women in this age group is about 2% P Test 2-sided alternative hypothesis that it differ using $\alpha = 0.05$?

$n = 10\,000$

$\hat{p} = 0.04$

$\alpha = 0.05$

$H_0: P = 0.02$ Vs. $H_1: P \neq 0.02$

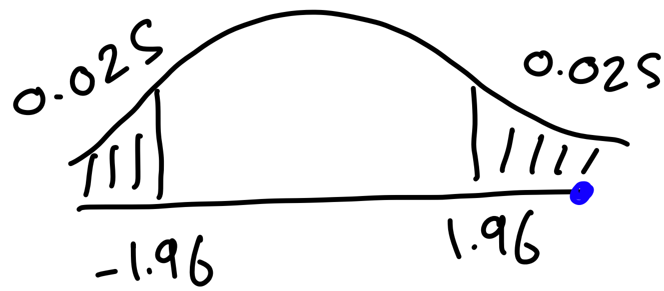
test stat

$$Z_{corr} = \frac{|\hat{p} - p| - \frac{1}{2n}}{\sqrt{\frac{pq}{n}}}$$

$$= \frac{|0.04 - 0.02| - \frac{1}{2 * 10\,000}}{\sqrt{\frac{0.02 * 0.98}{10\,000}}} = 14.3$$

$\alpha = 0.05$

$\frac{\alpha}{2} = 0.025$



$\therefore$ we reject H_0 and accept H_1

Example A researcher from Jordan Claimed that the national unemployment rate is 8%. (P)
In a random sample of size 200 (n) residents show that 22 residents who were unemployed
At $\alpha = 0.05$ and assuming a normal distribution test whether the researcher Claim is true.

$n = 200$

$\hat{P} = \frac{22}{200}$

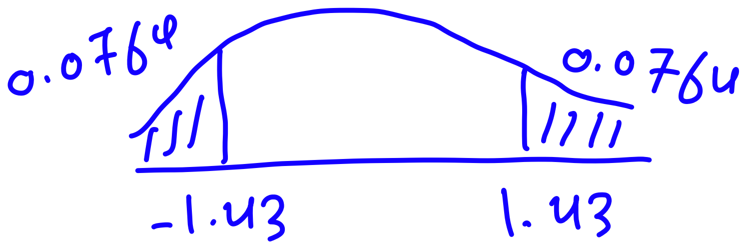
$\alpha = 0.05$

$H_0 : P = 0.08$ Vs. $H_1 : P \neq 0.08$

test stat

$$Z_{Corr} = \frac{|0.11 - 0.08| - \frac{1}{2 * 200}}{\sqrt{\frac{0.08 * 0.92}{200}}} = \underline{\underline{1.43}}$$

* P-value



$P = 2 * 0.0764 = 0.1528$

$P = 0.1528$ $\qquad P > \alpha$

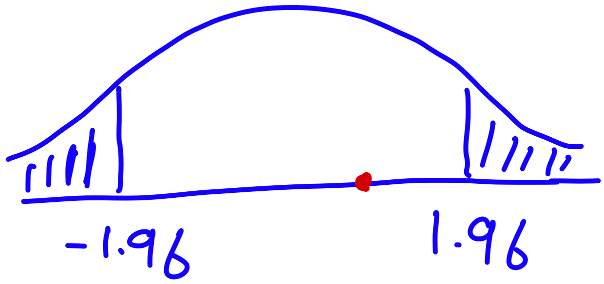
$\alpha = 0.0500$

$\therefore$ we accept H_0 and reject H_1

* Rejection region

$\alpha = 0.05$

$\frac{\alpha}{2} = 0.025$



So we accept H_0 and reject H_1

Example suppose the incidence rate of myocardial infarction (MI) was 5 per 1000 men in the year 2000. To look at changes in incidence over time, 5000 men in this group were followed after 1 year starting in 2010, fifteen new cases of MI were found.

$P = 0.005$

$$\frac{15}{5000} = \hat{p}$$

1) using critical value method with $\alpha = 0.05$ test the hypothesis that incidence rates of MI changed from 2000 to 2010

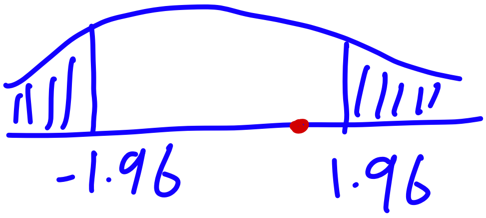
$H_0: P = 0.005$ Vs. $H_1: P \neq 0.005$

test stat

$$Z_{corr} = \frac{|0.003 - 0.005| - \frac{1}{2*5000}}{\sqrt{\frac{0.005*0.995}{5000}}} = 1.90$$

$\alpha = 0.05$

$\frac{\alpha}{2} = 0.025$

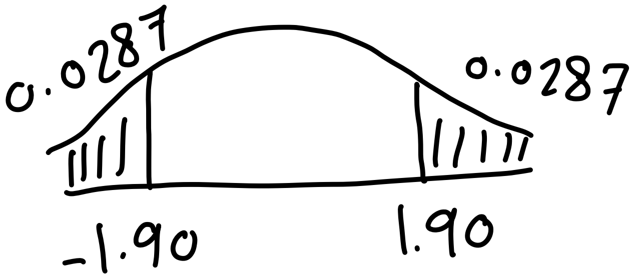


∴ we accept H_0 and reject H_1

(2) Report P-value to correspond to your answer in (1)

P-value = $2 * 0.0287$

$= 0.0574$



(1) $\bar{d} = \frac{\sum d}{n}$

(2) $S_d^2 = \frac{\sum d^2}{n-1} - \frac{(\sum d)^2}{n(n-1)}$

(3) $\mu_d \Rightarrow \bar{d} \pm E$

$E = t_{\frac{\alpha}{2}} * \frac{S_d}{\sqrt{n}}$

CI for paired data

test stat $= t = \frac{\bar{d} - \cancel{\mu_d}}{S/\sqrt{n}}$ 0

P-value $= 0.0574$ $\quad P > \alpha$

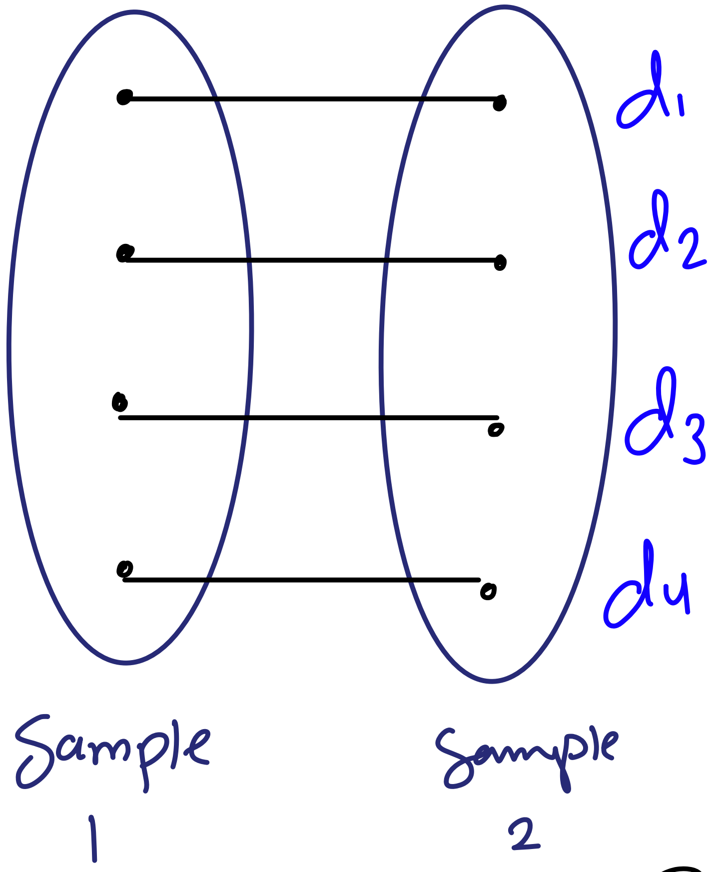
$\alpha = 0.0500$

Chapter 8

Test of hypothesis

⇒ for 2 samples ⇐

① For paired data



Dependent samples

① $\bar{d} = \frac{\Sigma d}{n}$

② $S_d^2 = \frac{\Sigma d^2}{n-1} - \frac{(\Sigma d)^2}{n(n-1)}$

③ $\mu_d \Rightarrow \bar{d} \pm E$

$E = t_{\frac{\alpha}{2}} * \frac{S_d}{\sqrt{n}}$

CI for paired data

test stat $= t = \frac{\bar{d} - \cancel{\mu_d}^{\,0}}{S/\sqrt{n}}$

Example Construct a 95% Confidence interval for the difference between SBP before and after using of oral contraceptions in a sample of 10 women using OCs, given the following data:

i	SBP level while not using OCs (x_{i1})	SBP level while using OCs (x_{i2})	d_i^*
1	115	128	13
2	112	115	3
3	107	106	−1
4	119	128	9
5	115	122	7
6	138	145	7
7	126	132	6
8	105	109	4
9	104	102	−2
10	115	117	2

$^*d_i = x_{i2} - x_{i1}$

$\bar{d} = 4.8$

$CI = 0.95$

$n = 10$

$\mu_d \Rightarrow \bar{d} \pm E$

$E = t_{\frac{\alpha}{2}} * \frac{S}{\sqrt{n}}$

$S^2 = \frac{\Sigma d^2}{n-1} - \frac{(\Sigma d)^2}{n(n-1)}$

$= 20.844$

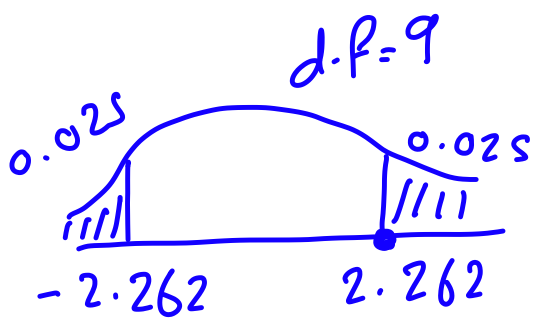
$S = 4.566$

$= 2.262 * \frac{4.566}{\sqrt{10}} = 3.266$

$1 - \alpha = 0.95$

$\alpha = 0.05$

$\boxed{\frac{\alpha}{2} = 0.025}$



$\boxed{t_{\frac{\alpha}{2}} = \pm 2.262}$

$$(\bar{d} - E, \ \bar{d} + E)$$

$$(4.8 - 3.266, \ 4.8 + 3.266)$$

Example: the sleep hours of 5 patients before and after taking a medication are given by the following table:

	1	2	3	4	5
Before	6	5	7	4	5
After	9	4	9	7	6
d	3	-1	2	3	1
d^2	9	1	4	9	1

$\sum d = 8$

$\sum d^2 = 24$

1) Construct 95% confidence interval for the mean difference.

$CI = 0.95$

$n = 5$

$\mu_d \Rightarrow \bar{d} \pm E$

$$E = t_{\frac{\alpha}{2}} * \frac{S}{\sqrt{n}}$$

$$\bar{d} = \frac{8}{5} = 1.6$$

$$S^2 = \frac{\sum d^2}{n-1} - \frac{(\sum d)^2}{n(n-1)}$$

$$= 2.8$$

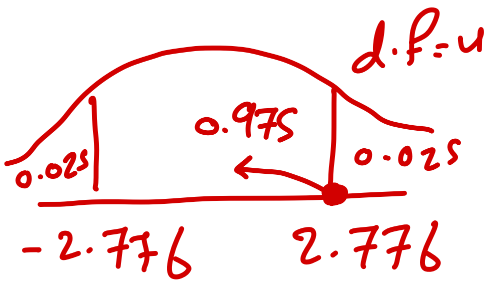
$$S = 1.67$$

$$E = 2.776 * \frac{1.67}{\sqrt{5}} = 2.07$$

$$1 - \alpha = 0.95$$

$$\alpha = 0.05$$

$$\frac{\alpha}{2} = 0.025$$



$t_{\frac{\alpha}{2}} = \pm 2.776$

$$(\bar{d} - E, \ \bar{d} + E)$$

$$(1.6 - 2.07, \ 1.6 + 2.07)$$

علاوة

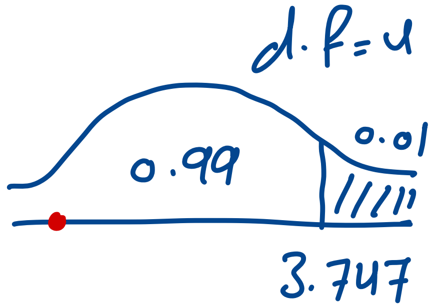
2) Can you conclud that the drug is effective in increasing the sleep hours (use $\alpha = 0.01$)

$H_0 : \mu_d \leq 0$ Vs. $H_1 : \mu_d > 0$

test stat

$$t = \frac{\bar{d} - \cancel{\mu_d}^{\,0}}{S/\sqrt{n}}$$

$$= \frac{1.6 - 0}{1.67/\sqrt{5}} = -2.14$$



$\therefore$ we accept H_0 and reject H_1

كتاب

Gynecology A topic of recent clinical interest is the effect of different contraceptive methods on fertility. Suppose we wish to compare how long it takes users of either OCs or diaphragms to become pregnant after stopping contraception. A study group of 20 OC users is formed, and diaphragm users who match each OC user with regard to age (within 5 years), race, parity (number of previous pregnancies), and socioeconomic status (SES) are found. The investigators compute the differences in time to fertility between previous OC and diaphragm users and find that the mean difference $\bar{d}$ (OC minus diaphragm) in time to fertility is 4 months with a standard deviation (s_d) of 8 months. What can we conclude from these data?

$n = 20$

$\bar{d} = 4$

$S_d = 8$

$\alpha = 0.05$

$H_0: \mu_d = 0$ Vs. $H_1: \mu_d \neq 0$

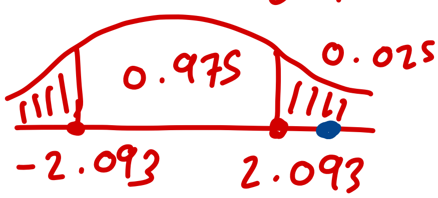
test stat

$$t = \frac{4 - 0}{8/\sqrt{20}} = 2.24$$

$\alpha = 0.05$

$\frac{\alpha}{2} = 0.025$

$d.f = 19$



∴ we reject H_0 and accept H_1

② Independent Samples

① Confidence interval

$(\mu_1 - \mu_2) \Rightarrow (\bar{x} - \bar{y}) \pm E$

CI for mean
(2 Samples)

σ_1, σ_2 known

$$\boxed{E = Z_{\frac{\alpha}{2}} * \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}}$$

n: Sample 1 size

m: sample 2 size

σ_1, σ_2 unknown

$\sigma_1^2 = \sigma_2^2$

$$E = t_{\frac{\alpha}{2}} * \sqrt{\frac{S_1^2}{n} + \frac{S_2^2}{m}}$$

$$= t_{\frac{\alpha}{2}} * \sqrt{S^2} * \sqrt{\frac{1}{n} + \frac{1}{m}}$$

$$\boxed{E = t_{\frac{\alpha}{2}} * S * \sqrt{\frac{1}{n} + \frac{1}{m}}}$$

<u>NOTE</u> S: pooled standard deviation

S^2: pooled variance

$$\boxed{S^2 = \frac{(n-1) * S_1^2 + (m-1) S_2^2}{n+m-2}}$$

NOTE degree of freedom

$$\boxed{d.f = n+m-2}$$

test of Hypothesis
(2 samples)

σ_1, σ_2 known

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}}$$

σ_1, σ_2 unknown

$$\boxed{\sigma_1^2 = \sigma_2^2}$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{Sp * \sqrt{\frac{1}{n} + \frac{1}{m}}}$$

$$Sp = \sqrt{\frac{(n-1)s_1^2 + (m-1)s_2^2}{n+m-2}}$$

مثال

8

Example Suppose a sample of eight (35-39) year old nonpregnants, premenopausal OC users who work in a company have a mean SBP of $\bar{X}_1$ 132.86 mmHg and sample standard deviation of 15.34 S_1 mmHg are identified. A sample of 21 n_2 non-pregnant, premenopausal non-OC users in the same age group are similarly identified who have mean SBP of $\bar{X}_2$ 127.44 mmHg and a sample standard deviation of S_2 18.23 mmHg. test the hypothesis that they have different population mean assuming SBP is normally distributed between 2 groups and Both have the same population variance

$\bar{x}_1 : 132.86$
$S_1 = 15.34$
$n_1 = 8$

$\bar{x}_2 = 127.44$
$S_2 = 18.23$
$n_2 = 21$

$H_0: \mu_1 - \mu_2 = 0$ Vs. $H_1: \mu_1 - \mu_2 \neq 0$

test stat

$$t = \frac{(132.86 - 127.44) - \cancel{(\mu_1 - \mu_2)}^{\,0}}{\cancel{Sp} * \sqrt{\frac{1}{8} + \frac{1}{21}}} = 0.74$$

17.527

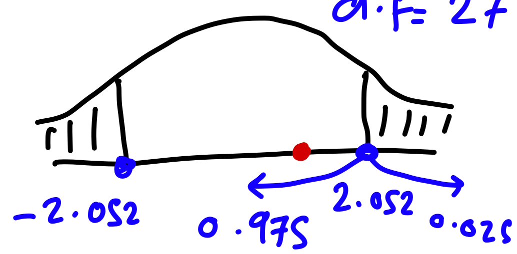
$\boxed{\alpha = 0.05}$

$\frac{\alpha}{2} = 0.025$

$$Sp = \sqrt{\frac{(8-1) * 15.34^2 + (21-1) * 18.23^2}{8 + 21 - 2}}$$

$= 17.527$

d.f = 27



∴ we accept Ho and reject H_1

NOTE when α is unknown ⇒ Assume $\alpha = 0.05$

2) Compute a 95 CI for the true mean difference in SBP between 35-39 y/o OC users and non OC users

$$(\mu_1 - \mu_2) \Rightarrow (\bar{x}_1 - \bar{x}_2) \pm E$$

$$E = t_{\frac{\alpha}{2}} * Sp * \sqrt{\frac{1}{n} + \frac{1}{m}}$$

$$= 2.052 * 17.527 * \sqrt{\frac{1}{8} + \frac{1}{21}} = 14.942$$

$$\left((132.86 - 127.44) - 14.942,\ (132.86 - 127.44) + 14.942\right)$$

$$(-9.52,\ 20.36)$$

Example To test whether males and females IQ mean differ, we selected a random sample of size 15 from adult males and another sample of size 16 from adult females and showed the following info:

Sample	Size	mean	SD
males	15	105	28
Females	16	109	24

Assume the normality of 2 populations and an equal variances.

$$H_0: \mu_1 - \mu_2 = 0 \quad Vs. \quad H_1: \mu_1 - \mu_2 \neq 0$$

<u>test stat</u>

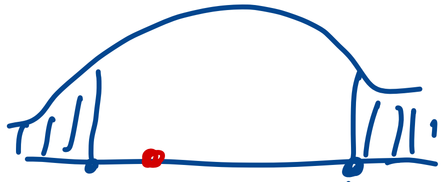
$$t = \frac{(105 - 109) - 0}{\cancel{S_p}\ 26.0079 * \sqrt{\frac{1}{15} + \frac{1}{16}}} = -0.556$$

$$SP = \sqrt{\frac{(15-1) * 28^2 + (16-1) * 24^2}{15 + 16 - 2}} = 26.0079$$

$\boxed{\alpha = 0.05}$

$\boxed{\frac{\alpha}{2} = 0.025}$

d.f = 29



-2.045 2.045

$\therefore$ we accept H_0 and Reject H_1

Example: Construct a 95 CI for $\mu_1 - \mu_2$ with the sample statistics for mean calorie content of two bakeries speciality pies as following:

$\bar{X}_1 = 448$ cal $\bar{X}_2 = 388$ cal

$S_1 = 6.1$ cal $S_2 = 7.8$ cal

$n_1 = 13$ $n_2 = 7$

$$(\mu_1 - \mu_2) \Rightarrow (\bar{x} - \bar{y}) \pm E$$

$$E = t_{\frac{\alpha}{2}} * SP * \sqrt{\frac{1}{n} + \frac{1}{m}}$$

$$= 2.101 * 6.71 * \sqrt{\frac{1}{13} + \frac{1}{7}} = 6.609$$

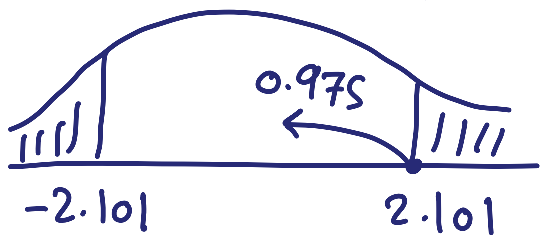
$$SP = \sqrt{\frac{(13-1) * 6.1^2 + (7-1) * 7.8^2}{13 + 7 - 2}} = 6.71$$

$$1 - \alpha = 0.95$$

$$\alpha = 0.05$$

$$\boxed{\frac{\alpha}{2} = 0.025}$$

d.f = 18



$$\left((448 - 388) - 6.609, (448 - 388) + 6.609\right)$$

NOTE

$$E = t_{\frac{\alpha}{2}} * SP * \sqrt{\frac{1}{n} + \frac{1}{m}}$$

$$SE = SP * \sqrt{\frac{1}{n} + \frac{1}{m}}$$