STATISTICAL COMPUTATION

WEEK 4 – HYPOTHESIS TESTING (2/2)

Annisa Auliya
I Melda Puspita

GET TO KNOW US

ANNISA AULIYA R.

082334174749

I MELDA PUSPITA L.

085257113961

https://intip.in/KomstatC2023

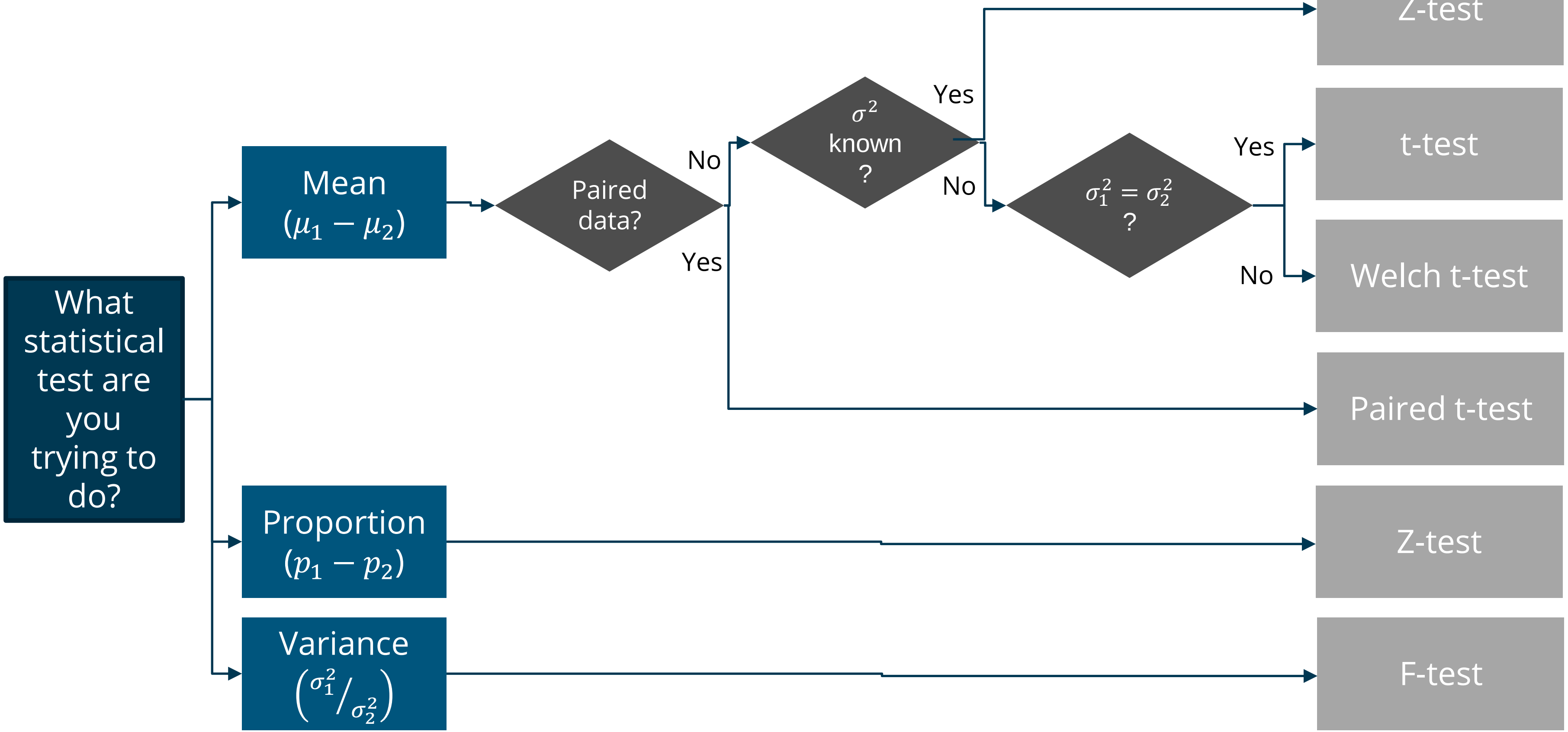
MATERIALS

- 2-Sample Test
- P-value
- Study Case

01

2-SAMPLE TEST

CHEAT SHEET (2 SAMPLE)



MEANS (σ_1^2 and σ_2^2 known)

	Right-side	Left-side	Two-tail
Hypothesis	$\mathrm{H}_0 : \mu_1 - \mu_2 = d_0$ $\mathrm{H}_1 : \mu_1 - \mu_2 > d_0$	$\mathrm{H}_0 : \mu_1 - \mu_2 = d_0$ $\mathrm{H}_1 : \mu_1 - \mu_2 < d_0$	$\mathrm{H}_0 : \mu_1 - \mu_2 = d_0$ $\mathrm{H}_1 : \mu_1 - \mu_2 \neq d_0$
Statistics	$Z_{hit} = \dfrac{(\bar{X}_1 - \bar{X}_2) - d_0}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$		
Critical Value	$Z_{hit} > Z_\alpha$	$Z_{hit} < -Z_\alpha$	$Z_{hit} < -Z_{\alpha/2}$ or $Z_{hit} > Z_{\alpha/2}$
Confidence Interval	$(\bar{X}_1 - \bar{X}_2) - Z_{\alpha/2}\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{X}_1 - \bar{X}_2) + Z_{\alpha/2}\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$		

MEANS (σ_1^2 and σ_2^2 unknown)
$\sigma_1^2 = \sigma_2^2$

	Right-side	Left-side	Two-tail
Hypothesis	$\mathrm{H}_0 : \mu_1 - \mu_2 = d_0$ $\mathrm{H}_1 : \mu_1 - \mu_2 > d_0$	$\mathrm{H}_0 : \mu_1 - \mu_2 = d_0$ $\mathrm{H}_1 : \mu_1 - \mu_2 < d_0$	$\mathrm{H}_0 : \mu_1 - \mu_2 = d_0$ $\mathrm{H}_1 : \mu_1 - \mu_2 \neq d_0$
Statistics	$T_{hit} = \dfrac{(\bar{X}_1 - \bar{X}_2) - \mathrm{d}_0}{s_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}; \; s_p^2 = \dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$		
Critical Value	$T_{hit} > t_{\alpha;v}$	$T_{hit} < -t_{\alpha;\mathrm{v}}$	$T_{hit} < -t_{\alpha/2;v}$ or $T_{hit} > t_{\alpha/2;v}$
Degree of Freedom	$v = n_1 + n_2 - 2$		
Confidence Interval	$(\bar{X}_1 - \bar{X}_2) - t_{\alpha/2}s_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}} < \mu_1 - \mu_2 < (\bar{X}_1 - \bar{X}_2) + t_{\alpha/2}s_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}$		

MEAN (σ_1^2 and σ_2^2 unknown) $\sigma_1^2 \neq \sigma_2^2$

	Right-side	Left-side	Two-tail
Hypothesis	$H_0 : \mu_1 - \mu_2 = d_0$ $H_1 : \mu_1 - \mu_2 > d_0$	$H_0 : \mu_1 - \mu_2 = d_0$ $H_1 : \mu_1 - \mu_2 < d_0$	$H_0 : \mu_1 - \mu_2 = d_0$ $H_1 : \mu_1 - \mu_2 \neq d_0$
Statistics	$T_{hit} = \dfrac{(\bar{X}_1 - \bar{X}_2) - \mathrm{d}_0}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$		
Critical Value	$T_{hit} > t_{\alpha;v}$	$T_{hit} < -t_{\alpha;\mathrm{v}}$	$T_{hit} < -t_{\alpha/2;v}$ or $T_{hit} > t_{\alpha/2;v}$
Degree of Freedom	$v = \dfrac{(s_1^2/n_1 + s_2^2/n_2)^2}{[(s_1^2/n_1)^2/(n_1 - 1)] + [(s_2^2/n_2)^2/(n_2 - 1)]}$		
Confidence Interval	$(\bar{X}_1 - \bar{X}_2) - t_{\alpha/2}\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{X}_1 - \bar{X}_2) + t_{\alpha/2}\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}$		

PAIRED MEAN

	Right-side	Left-side	Two-tail
Hypothesis	$H_0 : \mu_d = d_0$ $H_1 : \mu_d > d_0$	$H_0 : \mu_d = d_0$ $H_1 : \mu_d < d_0$	$H_0 : \mu_d = d_0$ $H_1 : \mu_d \neq d_0$
Statistics	$T_{hit} = \dfrac{\bar{d} - d_0}{S_D/\sqrt{n}}$		
Critical Value	$T_{hit} > t_{\alpha;v}$	$T_{hit} < -t_{\alpha;\mathrm{v}}$	$T_{hit} < -t_{\alpha/2;v}$ or $T_{hit} > t_{\alpha/2;v}$
Degree of Freedom	$v = n - 1$		
Confidence Interval	$\bar{d} - t_{\alpha/2}\dfrac{s_d}{\sqrt{n}} < \mathrm{D} < \bar{d} + t_{\alpha/2}\dfrac{s_d}{\sqrt{n}}$		

PROPORTIONS

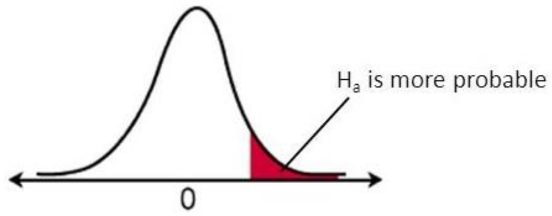
	Right-side	Left-side	Two-tail
Hypothesis	$\mathrm{H}_0 : p_1 = p_2$ $\mathrm{H}_1 : p_1 > p_2$	$\mathrm{H}_0 : p_1 = p_2$ $\mathrm{H}_1 : p_1 < p_2$	$\mathrm{H}_0 : p_1 = p_2$ $\mathrm{H}_1 : p_1 \neq p_2$
Statistics	$Z_{hit} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q}(1/n_1 + 1/n_2)}};\ \hat{p}_i = \frac{x_i}{n_i}, \hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$		
Critical Value	$Z_{hit} > Z_\alpha$	$Z_{hit} < -Z_\alpha$	$Z_{hit} < -Z_{\alpha/2}$ or $Z_{hit} > Z_{\alpha/2}$
Confidence Interval	$(\hat{p}_1 - \hat{p}_2) - Z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} < p_1 - p_2 < \hat{p}_1 - \hat{p}_2 + Z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$		

VARIANCES

	Right-side	Left-side	Two-tail
Hypothesis	$\mathrm{H}_0 : \sigma_1^2 = \sigma_2^2$ $\mathrm{H}_1 : \sigma_1^2 > \sigma_2^2$	$\mathrm{H}_0 : \sigma_1^2 = \sigma_2^2$ $\mathrm{H}_1 : \sigma_1^2 < \sigma_2^2$	$\mathrm{H}_0 : \sigma_1^2 = \sigma_2^2$ $\mathrm{H}_1 : \sigma_1^2 \neq \sigma_2^2$
Statistics	$F_{hit} = \frac{s_1^2}{s_2^2}$		
Critical Value	$F_{hit} > f_{\alpha,v_1,v_2}$	$F_{hit} < f_{1-\alpha,v_1,v_2}$	$F_{hit} < f_{1-\frac{\alpha}{2},v_1,v_2}$ or $F_{hit} > f_{\frac{\alpha}{2},v_1,v_2}$
Degree of Freedom	$v_1 = n_1 - 1$ $v_2 = n_2 - 1$		
Confidence Interval	$\frac{s_1^2}{s_2^2} \cdot \frac{1}{f_{\frac{\alpha}{2},v_1,v_2}} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{s_1^2}{s_2^2} \cdot f_{\frac{\alpha}{2},v_2,v_1}$		

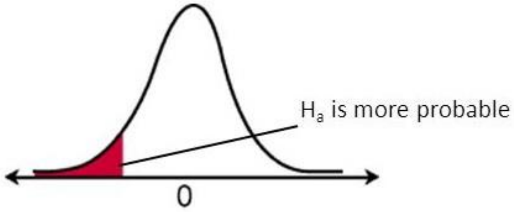
02
P-VALUE

TAIL OF TEST



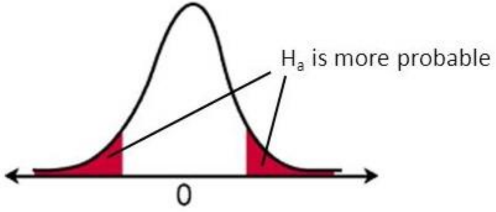
Right-tail test

H_a: $\mu >$ value



Left-tail test

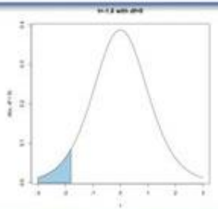
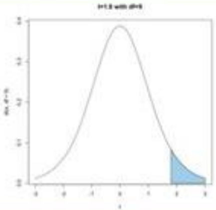
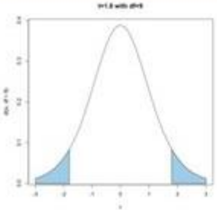
H_a: $\mu <$ value



Two-tail test

H_a: $\mu \neq$ value

Calculating P-Values Cheat Sheet

	H_a Description	Figure	Example t values	R code to get p-values
One-tailed	Population mean is below the standard		t = -1.8	pt(-1.8, df=9)
One-tailed	Population mean is above the standard		t = 1.8	1-pt(1.8, df=9) '1 - ' because otherwise you will get the probability for t less than 1.8
Two-tailed	Population mean is different from the standard		t = 1.8 t = -1.8	2*(1-pt(1.8, df=9)) 2*pt(-1.8, df=9)

MACRO MINITAB FUNCTIONS

Function	Description	Usage
PDF	Density	Find PDF
CDF	Distribution function	Find CDF (p-value)
InvCDF	Quantile function	Find inverse CDF (critical value)
Random	Random deviated	Generate random variable

03

STUDY CASE

STUDY CASE 1

Does a Pill Incidentally Reduce Blood Pressure?

A medical researcher wishes to determine if a pill has the undesirable side effect of reducing the blood pressure of the user. The study involves recording the initial blood pressures of 15 college-age women. After they use the pill regularly for six months, their blood pressures are again recorded. The researcher wishes to draw inferences about the effect of the pill on blood pressure from the observations given in Table 1.

(a) Calculate a ~~95%~~ confidence interval for the mean reduction in blood pressure. 99 %

(b) Do the data substantiate the claim that use of the pill reduces blood pressure? Test at $\alpha = .01$.

	Subject														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Before (x)	70	80	72	76	76	76	72	78	82	64	74	92	74	68	84
After (y)	68	72	62	70	58	66	68	52	64	72	74	60	74	72	74
$d = x - y$	2	8	10	6	18	10	4	26	18	−8	0	32	0	−4	10

STUDY CASE 2

Testing Equality of Prevalence of a Virus

A study (courtesy of R. Golubjatnikov) is undertaken to compare the rates of prevalence of CF antibody to parainfluenza I virus among boys and girls in the age group 5 to 9 years. Among 113 boys tested, 34 are found to have the antibody; among 139 girls tested, 54 have the antibody. Do the data provide strong evidence that the rate of prevalence of the antibody is significantly higher in girls than boys? Use $\alpha = .05$. Also, find the P–value.

STUDY CASE 3

Testing Equality of Green Gas Mean Yields

One process of making green gasoline, not just a gasoline additive, takes biomass in the form of sucrose and converts it into gasoline using catalytic reactions. This research is still at the pilot plant stage. At one step in a pilot plant process, the product volume (liters) consists of carbon chains of length 3. Nine runs were made with each of two catalysts and the product volumes measured.

catalyst 1	1.86	2.05	2.06	1.88	1.75	1.64	1.86	1.75	2.13
catalyst 2	.32	1.32	.93	.84	.55	.84	.37	.52	.34

The sample sizes $n_1 = n_2 = 9$ and the summary statistics are

$$\bar{x} = 1.887, \quad s_1^2 = .0269 \quad \bar{y} = .670 \quad s_2^2 = .1133$$

Is the mean yield with catalyst 1 more than .80 liters higher than the yield with catalyst 2? Test with $\alpha = 0.05$

$\sigma_1^2 = \sigma_2^2$ (?)

TASK

- Continue the solution for study case 3 in the previous slide
- Run in macro minitab with critical value, p-value, and CI
  - Provide macro minitab for each test (T test for equal & unequal) — 2
  - Provide 1 macro minitab included all tests (F test, T test for equal & unequal) — 1
- Format file name : NRP_Nama.txt (Ex : 6003221023_Annisa Auliya Rahman)
- Deadline : Tuesday, May 23rd, 2023 (23.59 WIB)

Clue : use if or call another macro

THANKS

https://intip.in/KomstatC2023