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

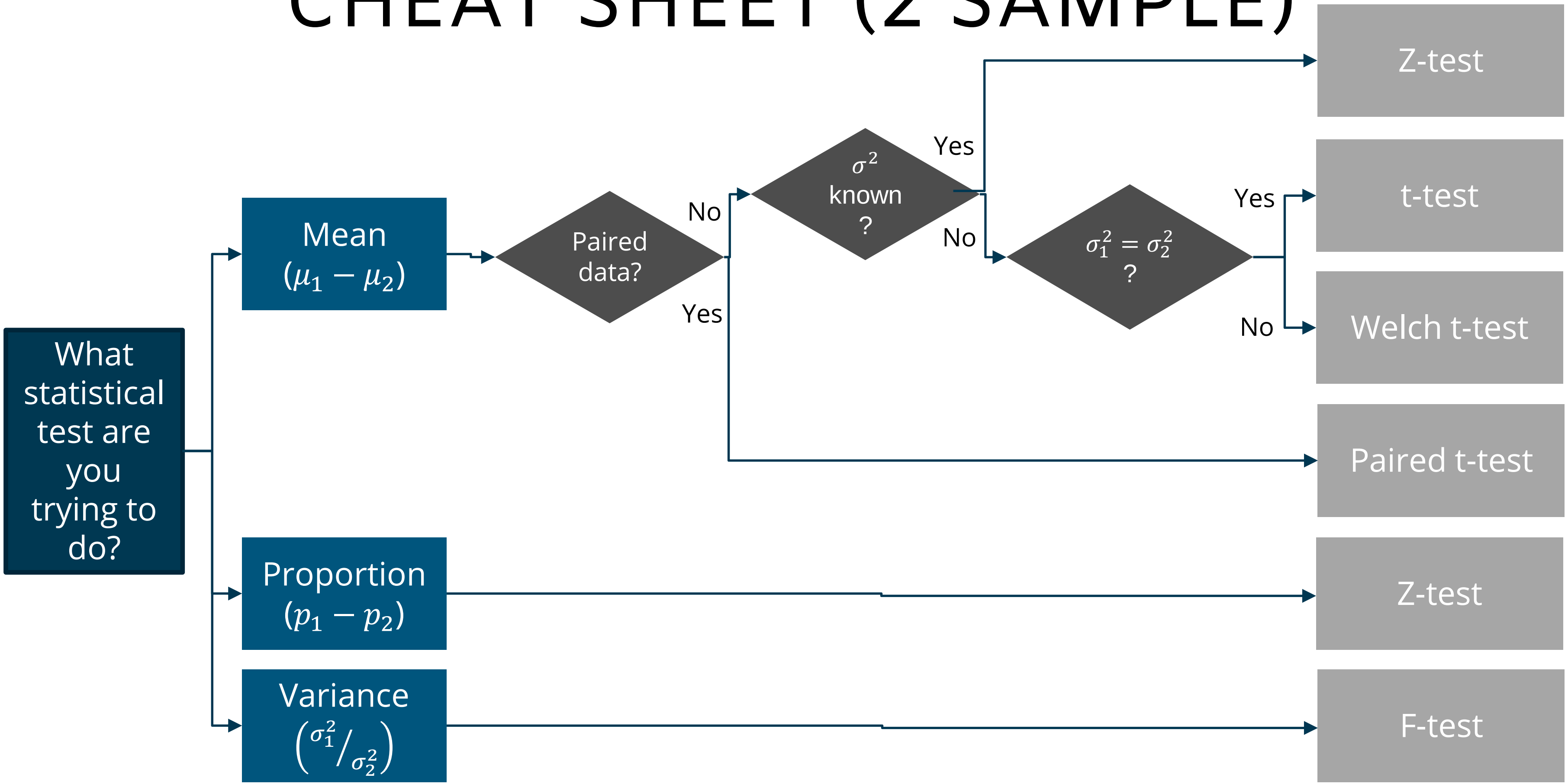
MATERIALS

- 2-Sample Test
- P-value

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

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	$Z_{hit} = \dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q}(1/n_1 + 1/n_2)}}$; $\hat{p}_i = \dfrac{x_i}{n_i}, \hat{p} = \dfrac{x_1 + x_2}{n_1 + n_2}$		
Critical Value	$Z_{hit} > Z_\alpha$	$Z_{hit} < -Z_\alpha$	$Z_{hit} < -Z_{\alpha/2}$ or $Z_{hit} > Z_{\alpha/2}$
Confidence Interval	$(\hat{p}_1 - \hat{p}_2) - Z_{\alpha/2}\sqrt{\dfrac{\hat{p}_1\hat{q}_1}{n_1} + \dfrac{\hat{p}_2\hat{q}_2}{n_2}} < p_1 - p_2 < \hat{p}_1 - \hat{p}_2 + Z_{\alpha/2}\sqrt{\dfrac{\hat{p}_1\hat{q}_1}{n_1} + \dfrac{\hat{p}_2\hat{q}_2}{n_2}}$		

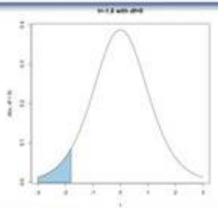
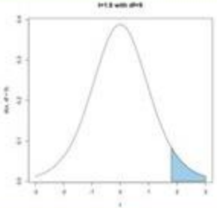
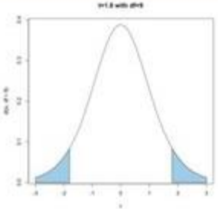
VARIANCES

	Right-side	Left-side	Two-tail
Hypothesis	$H_0 : \sigma_1^2 = \sigma_2^2$ $H_1 : \sigma_1^2 > \sigma_2^2$	$H_0 : \sigma_1^2 = \sigma_2^2$ $H_1 : \sigma_1^2 < \sigma_2^2$	$H_0 : \sigma_1^2 = \sigma_2^2$ $H_1 : \sigma_1^2 \neq \sigma_2^2$
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

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)

R FUNCTIONS

Function	Description	Usage
d*dist*	Density	Find PDF
p*dist*	Distribution function	Find CDF (p-value)
q*dist*	Quantile function	Find inverse CDF (critical value)
r*dist*	Random deviated	Generate random variable

THANKS

https://intip.in/KomstatC2023