STATISTICAL COMPUTATION

WEEK 3 – HYPOTHESIS TESTING (1/2)

Annisa Auliya
I Melda Puspita

GET TO KNOW US

ANNISA AULIYA R.

082334174749

I MELDA PUSPITA L.

085257113961

https://intip.in/KomstatC2023

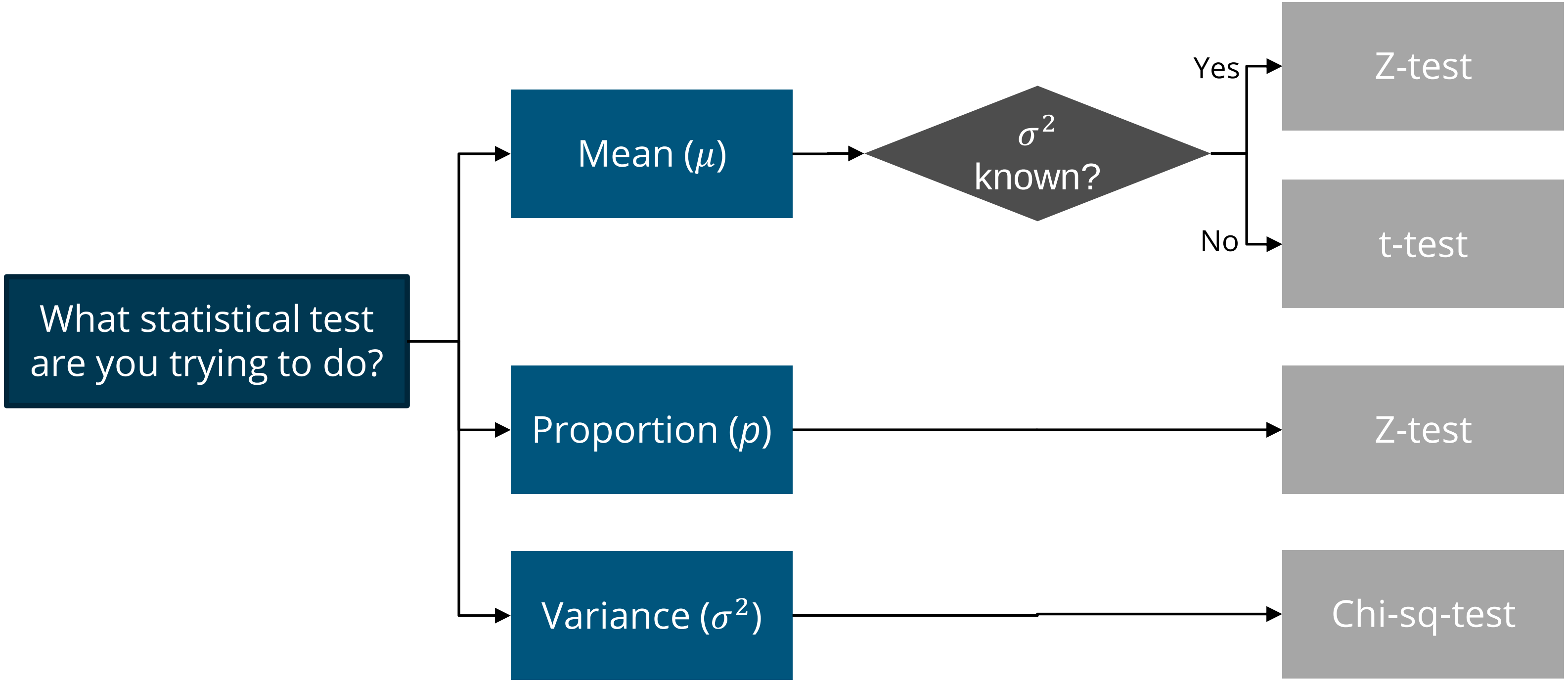
MATERIALS

- 1-Sample Test
- P-value
- Study Case

01

1-SAMPLE TEST

CHEAT SHEET (1 SAMPLE)



MEAN (σ^2 known)

	Right-side	Left-side	Two-tail
Hypothesis	$\mathrm{H}_0 : \mu = \mu_0$ $\mathrm{H}_1 : \mu > \mu_0$	$\mathrm{H}_0 : \mu = \mu_0$ $\mathrm{H}_1 : \mu < \mu_0$	$\mathrm{H}_0 : \mu = \mu_0$ $\mathrm{H}_1 : \mu \neq \mu_0$
Statistics	$Z_{hit} = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$		
Critical Value	$Z_{hit} > Z_\alpha$	$Z_{hit} < -Z_\alpha$	$Z_{hit} < -Z_{\alpha/2}$ or $Z_{hit} > Z_{\alpha/2}$
Confidence Interval	$\bar{X} - Z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}} < \mu < \bar{X} + Z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$		

MEAN (σ^2 unknown)

	Right-side	Left-side	Two-tail
Hypothesis	$\mathrm{H}_0 : \mu = \mu_0$ $\mathrm{H}_1 : \mu > \mu_0$	$\mathrm{H}_0 : \mu = \mu_0$ $\mathrm{H}_1 : \mu < \mu_0$	$\mathrm{H}_0 : \mu = \mu_0$ $\mathrm{H}_1 : \mu \neq \mu_0$
Statistics	$T_{hit} = \dfrac{\bar{X} - \mu_0}{S/\sqrt{n}}$		
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	$\bar{X} - t_{\alpha/2}\dfrac{S}{\sqrt{n}} < \mu < \bar{X} + t_{\alpha/2}\dfrac{S}{\sqrt{n}}$		

PROPORTION

	Right-side	Left-side	Two-tail
Hypothesis	$\mathrm{H_0} : \mathrm{p} = p_0$ $\mathrm{H_1} : \mathrm{p} > p_0$	$\mathrm{H_0} : \mathrm{p} = p_0$ $\mathrm{H_1} : \mathrm{p} < p_0$	$\mathrm{H_0} : \mathrm{p} = p_0$ $\mathrm{H_1} : p \neq p_0$
Statistics	$Z_{hit} = \dfrac{\hat{p} - p_0}{\sqrt{\hat{p}\hat{q}/n}}$		
Critical Value	$Z_{hit} > Z_\alpha$	$Z_{hit} < -Z_\alpha$	$Z_{hit} < -Z_{\alpha/2}$ or $Z_{hit} > Z_{\alpha/2}$
Confidence Interval	$\hat{p} - Z_{\alpha/2}\sqrt{\hat{p}\hat{q}/n} < \mathrm{p} < \hat{p} + Z_{\alpha/2}\sqrt{\hat{p}\hat{q}/n}$		

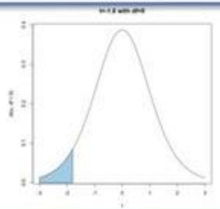
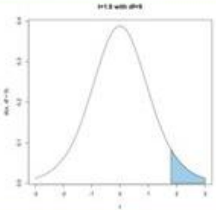
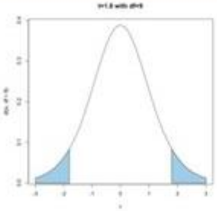
VARIANCE

	Right-side	Left-side	Two-tail
Hypothesis	$H_0 : \sigma^2 = \sigma_0^2$ $H_1 : \sigma^2 > \sigma_0^2$	$H_0 : \sigma^2 = \sigma_0^2$ $H_1 : \sigma^2 < \sigma_0^2$	$H_0 : \sigma^2 = \sigma_0^2$ $H_1 : \sigma^2 \neq \sigma_0^2$
Statistics	$\chi^2_{hit} = (n-1)\dfrac{s^2}{\sigma^2}$		
Critical Value	$\chi^2_{hit} > \chi^2_{\alpha;\mathrm{v}}$	$\chi^2_{hit} < \chi^2_{(1-\alpha);\mathrm{v}}$	$\chi^2_{hit} < \chi^2_{(1-\frac{\alpha}{2});\mathrm{v}}$ or $\chi^2_{hit} > \chi^2_{(\frac{\alpha}{2});\mathrm{v}}$
Degree of Freedom	$v = n - 1$		
Confidence Interval	$(n-1)\dfrac{s^2}{\chi^2_{(\frac{\alpha}{2});\mathrm{v}}} < \sigma^2 < (n-1)\dfrac{s^2}{\chi^2_{(1-\frac{\alpha}{2});\mathrm{v}}}$		

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

03

STUDY CASE 1-SAMPLE

STUDY CASE 1

Andhika, as an eight-year old, established a **mean time of 16.43 seconds** for swimming the 25-yard freestyle, with a **standard deviation of 0.8 seconds**. His dad thought that Andhika could swim the 25-yard freestyle faster using goggles. His dad bought Andhika a new pair of expensive goggles and timed Andhika for 25-yard freestyle swims. **For the 15 swims, Andhika's mean time was 16 seconds**. His dad thought that the goggles helped Andhika to swim **faster than the 16.43 seconds**. Conduct a hypothesis test using a preset $\alpha = 0.05$.

STUDY CASE 2

A manufacturer of salad dressings uses machines to dispense liquid ingredients into bottles that move along a filling line. The machine that dispenses salad dressings is **working properly when 8 ounces are dispensed**. Suppose that the **average amount dispensed in a particular sample of 35 bottles is 7.91 ounces** with a **variance of 0.03 ounces squared, s2**. Is there evidence that the machine should be stopped and production wait for repairs? The lost production from a shutdown is potentially so great that management feels that the **level of significance in the analysis should be 99%**

STUDY CASE 3

According to the N.Y. Times Almanac the **mean family size in the U.S. is 3.18**. A sample of a college math class resulted in the following family sizes:

545443643355633274522232

At **α = 0.05** level, is the class' **mean family size greater than the national average?** Does the Almanac result remain valid? Why?

STUDY CASE 4

The mortgage department of a large bank is interested in the nature of loans of first-time borrowers. This information will be used to tailor their marketing strategy. They believe that 50% of first-time borrowers take out smaller loans than other borrowers. They perform a hypothesis test to determine if the percentage is the **same or different from 50%.** They sample **100 first-time borrowers and find 53 of these loans are smaller than the other borrowers**. For the hypothesis test, they choose a **5% level of significance**.

→ 2-tails, 1-sample, proportion test (Z-test), a=0.05, p0=0.5, pest=53/100

STUDY CASE 5

A manufacturer of ca batteries claims that the life of the company's batteries is approximately normally distribution with **standard deviation equal to 0.9 year**. If a **random sample of 10 of these batteries has a standard deviation of 1.2 years**, do you think that $\sigma \neq 0.9$ year? Use a 0.05 level of significance.

→ 2-tails, 1-sample, var (chi test), a=0.05, var0=0.9, sd=1.2, n=9

THANKS

https://intip.in/KomstatC2023