COMPUTER PROJECT MAS291

Presented by Group 2

MEMBERS:

Lê Thị Chí Thương
Vũ Ngọc Ánh
Võ Công Huy
Nguyễn Trung Dũng
Nguyễn Ngô Chiến

Q.1

Test a hypothesis and construct a confidence interval for the mean of a population.

- Test hypothesis that the average life expactancy of male is 70 (with confidence level 95%)
- Select a sample has 100 random countries to test this hypothesis.
- Test hypotheses:

  H0: $\mu = 70$

  H1: $\mu \neq 70$

Q.1 Test a hypothesis and construct a confidence interval for the mean of a population.

We find test statistic

$$t_{\text{STAT}} = \frac{\overline{X} - \mu}{\frac{S}{\sqrt{n}}}$$

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{n-1}}$$

Construct a confidence interval

$$\overline{X} - t_{\alpha/2,n-1}\frac{S}{\sqrt{n}} \leq \mu \leq \overline{X} + t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}$$

	Test the hypothesis that the average life expactance of male in random 100 countries is 70 (with α = 0.05)							
		H0	μ0 = 70					
		H1	μ0 ≠ 70					
		α	0.05					
	Sample mean	x	70.162					
	Sample variance	s	7.6694169					
	Sample size		100					
	Test statistic	t0	0.21122857					
	Critical value	t(α/2,n-1)	0.96022313		Fail to reject H0			
	Lower-Confidence Interval		69.42568					
	Upper-Confidence Interval		70.89832					

=> Fail to reject H0 mean we do not have effective evidence to conclude that the average life expectancy of male is different from 70.

Q.2 Test a hypothesis and construct a confidence interval for the proportion of a population.

30% of the countries have a mean household incomes greater than $7,500 (with confidence level is 95%).

1.Test a hypothesis:

- Choose total samples = 189 countries to test hypothesis:
- Create table Sample Statistics with Confidence Level (95%) for Proportion

Q.2

Test a hypothesis and construct a confidence interval for the proportion of a population.

We have:

Total samples	n	189
Sample satisfies the condition	X	66
Signficance level	α	0.05
Critical value	$Z(\alpha/2)$	1.96
Hypothesis probability	$P=X/n$	0.349
	H0	$p0 = 0.3$
	H1	$p0 \neq 0.3$ (Two-tail test)

Q.2 Test a hypothesis and construct a confidence interval for the proportion of a population.

1.Test a hypothesis:

Tradition method: H0: p0 = 0.3; H1: p0 ≠ 0.3

From Sample Statistics, we find test statistic:

$$z_0 = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

$$= \frac{0.349 - 0.3}{\sqrt{\frac{0.3 \times (1-0.3)}{189}}} = 1.47$$

Z0 = 1.47 → −Critical Values ≤ Z0 ≤ Critical Values → Fail to reject H0

Conclussion: At significance level of 0.05, there is not effective evidence that mean household income greater than $7,500 is different from 30% of 189 countries.

Q.2

Test a hypothesis and construct a confidence interval for the proportion of a population.

2.Construct a confidence interval:

Base on:

$$\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq p \leq \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$0.349 - 1.96 \times \sqrt{\frac{0.349 \times (1-0.349)}{189}} \leq p \leq 0.349 + 1.96 \times \sqrt{\frac{0.349 \times (1-0.349)}{189}}$$

$$0.2810439109 \leq p \leq 0.4169560891$$

Q.2 Test a hypothesis and construct a confidence interval for the proportion of a population.

2.Construct a confidence interval:

We have:

Lower - confidence interval	0.2810439109
Upper - confidence interval	0.4169560891

Conclussion: We are 95% confident that the proportion of countries with a mean household income greater than $7,500 is between 0.2810439109 and 0.4169560891.

Q.3 Test a hypothesis and construct a confidence interval for the difference in means of two populations.

The mean household incomes in countries with male's life expectancy greater than or equal to 70 are different from the mean household incomes in countries with male's life expectancy less than 70 (with confidence level is 95%)?

1.Test a hypothesis:

- Choose random 75 samples of Mean household incomes of male's life expectancy >= 70 is N1 and 75 samples of Mean household incomes of male's life expectancy < 70 is N2 to test hypothesis:

N1	
Mean	13208.66667
Standard Error	932.3796028
Median	11200
Mode	20200
Standard Deviation	8074.64422
Sample Variance	65199879.28
Kurtosis	0.489120481
Skewness	0.850209213
Range	37670
Minimum	1130
Maximum	38800
Sum	990650
Count	75
Confidence Level(95.0%)	1857.806913

N2	
Mean	3017.826667
Standard Error	342.6662366
Median	2300
Mode	1430
Standard Deviation	2967.576659
Sample Variance	8806511.226
Kurtosis	9.287168528
Skewness	2.793058589
Range	15789
Minimum	411
Maximum	16200
Sum	226337
Count	75
Confidence Level(95.0%)	682.7773807

Q.3 Test a hypothesis and construct a confidence interval for the difference in means of two populations.

II. Create table Sample Statistics with Confidence Level (95.0%) for Mean

+Tradition method: I choose H0: u1 -u2 = 0, H1: u1 - u2 ≠ 0

From Sample Statistics: We find test statistic T0

Based on

$$T_0^* = \frac{\overline{X}_1 - \overline{X}_2 - \Delta_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

$$\nu = \left\lfloor \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}} \right\rfloor.$$

To	10.25902299
v	93.63211377

Because we choose Confidence Level (95.0%) for Mean:

=T.INV.2T(2*probability,deg_freedom)

1.985523442

T0 is in rejection region, then we **reject H0.**

Q.3 Test a hypothesis and construct a confidence interval for the difference in means of two populations.

Construct a confidence interval

A $100(1-\alpha)\%$ confidence interval for $\mu_1 - \mu_2$ is

$$(\overline{x}_1 - \overline{x}_2) - t_{\alpha/2,\nu}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \leq \mu_1 - \mu_2 \leq (\overline{x}_1 - \overline{x}_2) + t_{\alpha/2,\nu}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}.$$

Lower Confidence Interval	10188.85448
Upper Confidence Interval	10192.82552

Q.4 Test a hypothesis and construct a confidence interval for the difference in proportions of two populations.

Sample 1 with 89 countries

In this, there are 48 countries that have male ratio >= 70

p1 = 48/89

Sample 2 with 100 countries

In this, there are 41 countries that have male ratio >= 70

p2 = 41/100

Pooled proportion p = (48+41) /(89+100)

Q.4 Test a hypothesis and construct a confidence interval for the difference in proportions of two populations.

Construct confidence interval:

A $100(1-\alpha)\%$ confidence interval for $p_1 - p_2$ is

$$\hat{p}_1 - \hat{p}_2 - z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} \leq p_1 - p_2 \leq \hat{p}_1 - \hat{p}_2 + z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

where $z_{\alpha/2}$ is the upper $\alpha/2$ percentage point of the standard normal distribution.

From this: Lower confidence interval:-0.012155914
Upper confidence interval: 0.270807599

Q.4 Test a hypothesis and construct a confidence interval for the difference in proportions of two populations.

Test a hypothesis: Ho : p1 - p2 = 0; H1 : p1 - p2 ≠ 0

Test statistic:

$$z_0 = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\bar{p}(1-\bar{p})}{n_1} + \frac{\bar{p}(1-\bar{p})}{n_2}}}.$$

Zo = 1,78

Identify the acceptance region (95% confidence interval)

-1.96 <= z0 <= 1.96

Make decision: Fail to reject Ho

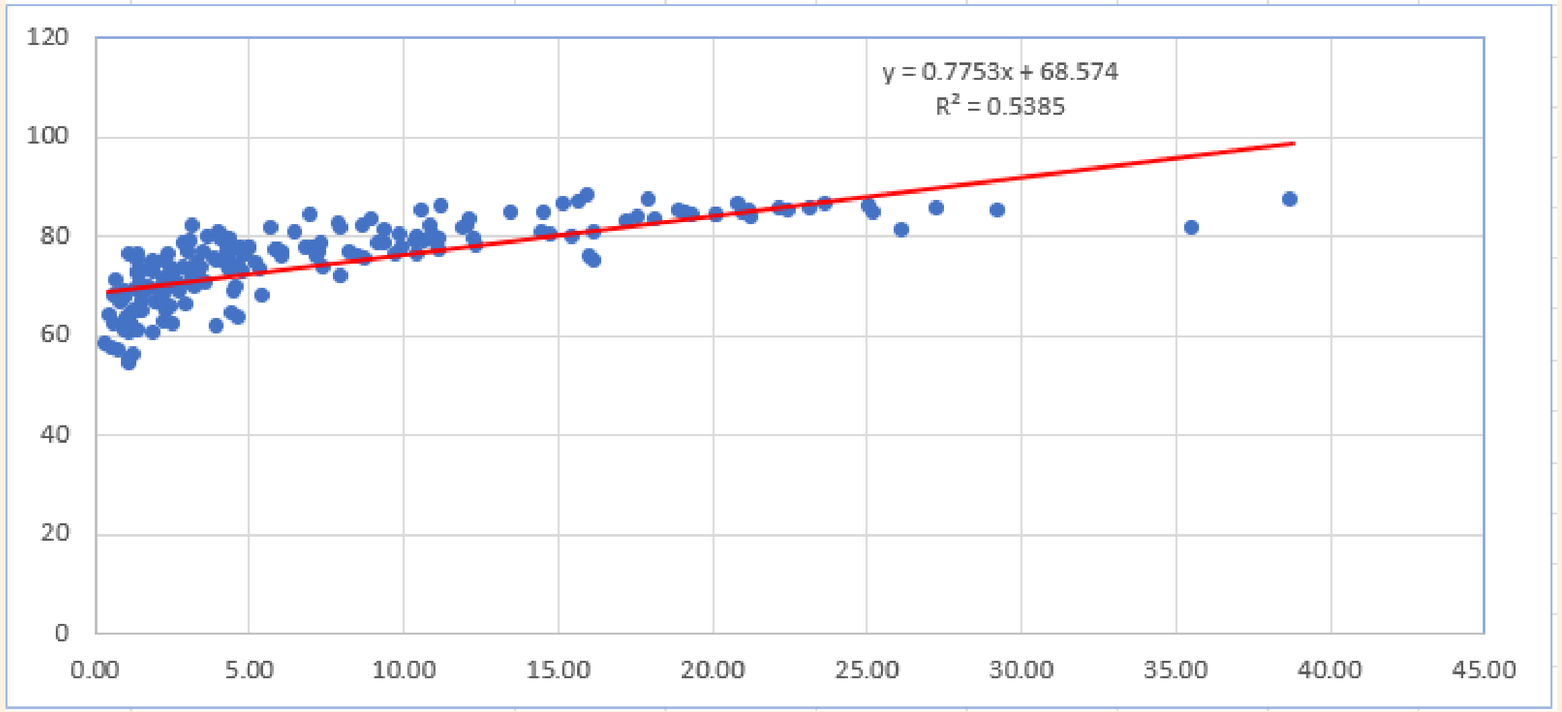
Q.5

Regression analysis

Identify two random variables X and Y	
X	Mean household income
Y	Female

Q.5

Regression analysis



Have a linear realationship here

Q.5

Regression analysis

Regression Statistics	
Multiple R	0.733813276
R Square	0.538481924
Adjusted R Square	0.536013913
Standard Error	5.35529824
Observations	189

r^2 = 0.5385

=> r = ± 0.7338

Q.5

Regression analysis

	Coefficients
Intercept	68.57400253
Mean household income (thousand $)	0.775292911

Slope = 0.7753

=> r = 0.7338

=> Strong positive correlation

Q.5 Regression analysis

	Coefficients
Intercept	68.57400253
Mean household income (thousand $)	0.775292911

=> Female = 68.5740 + 0.7753 * Mean household income

Q.5 Regression analysis

Estimate Line

=> Female = 68.5740 + 0.7753 * Mean household income

Mean household income	68.5740 + 0.7753 * Mean household income	Female
23.7	86.94844451	86.3
15.5	80.59104264	79.5
16.1	81.05621839	75.7
12.4	78.18763462	77.8
8.81	75.40433307	75.1
8.22	74.94691025	
1.63	69.83772997	

A WARM THANK YOU TO ALL OF YOU!