

Abstract geometric lines in the top-left corner of the slide, consisting of several overlapping, tilted rectangles and polygons drawn with thin black lines.

# HYPOTHESIS TESTING

Dang Cao Thien

# VẤN ĐỀ CHUNG

\* Giả thuyết: giả định về trung bình, phương sai, tỉ lệ.. Của tổng thể

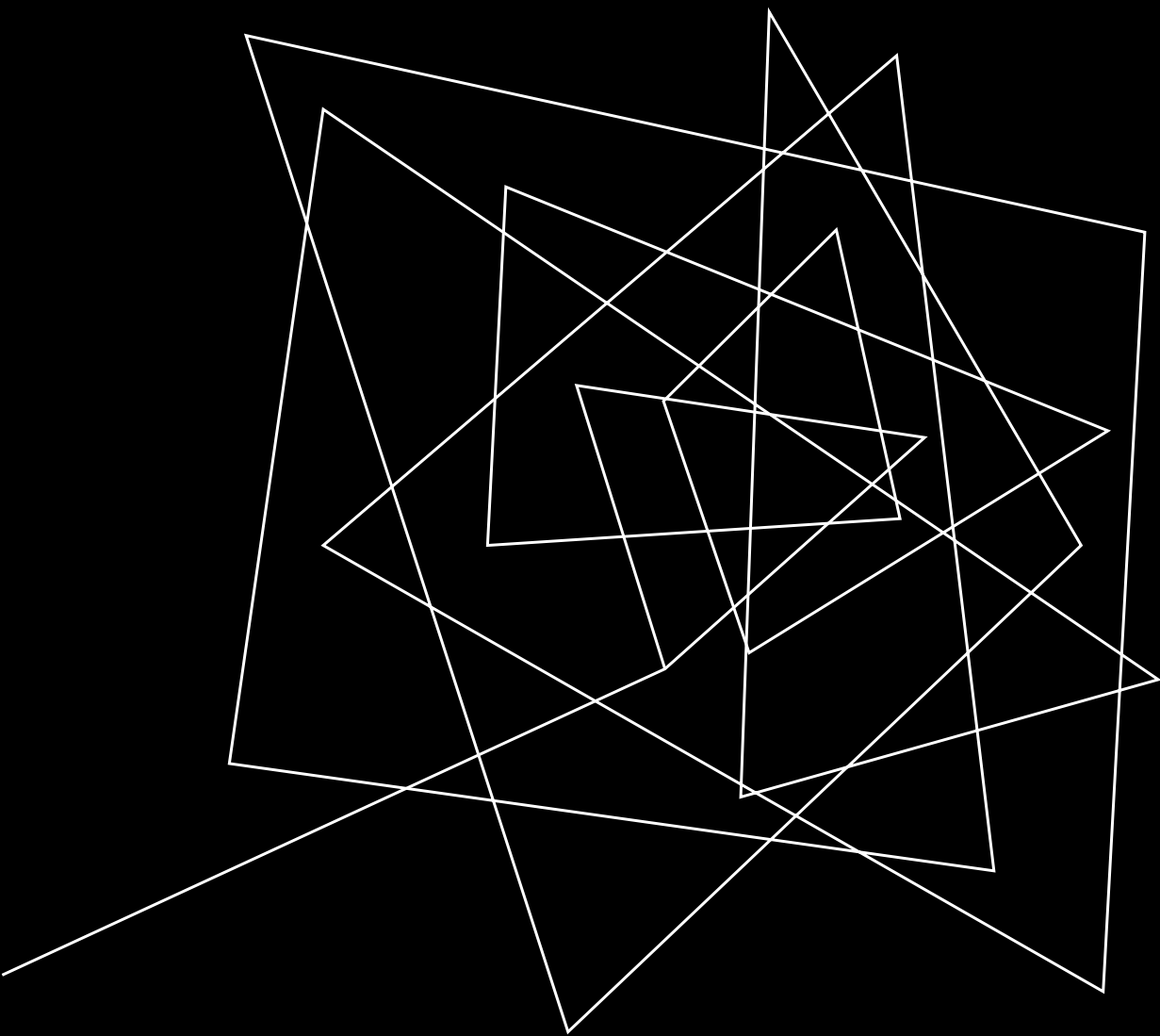
\* Tiêu chuẩn kiểm định: là quy luật phân phối xác suất dùng để kiểm định

\* P-value: mức xác suất max dùng để bác bỏ  $H_0$

Gia thuyet	Chap nhan	Bac bo
$H_0$ đúng	đúng	sll2
$H_0$ sai	sll1	đúng

# MÔ TẢ DATA

Tên biến	Giải thích	Mã giá trị
Code	Mã khách hàng	
C1	Khu vực sinh sống	1 'Miền Bắc' 2 'Miền Nam' 3 'Miền Trung'
C2	Giới tính	1 "Nam" 2 Nữ
C3	Năm sinh	
C4	Số năm đi học	
C5	Loại công việc	1 'Văn thư' 2 'Nhân viên kinh doanh' 3 'Quản lý'
C6_1	Biết tiếng Anh	
C6_2	Biết tiếng Pháp	
C6_3	Biết tiếng Trung	
C7	Tiền lương	
C8	Ngày đặt vé	
C9	Ngày bay	
C10	Tình trạng bay	1 'Có' 2 'Không'
C11_1	Nhân viên hướng dẫn tận tình, chuyên ng	1 'Rất không đồng ý' 2 'Không đồng ý' 3 'Bình thường' 4 'Đồng ý' 5 'Rất đồng ý'
C11_2	Có dịch vụ chăm sóc khách hàng tốt	1 'Rất không đồng ý' 2 'Không đồng ý' 3 'Bình thường' 4 'Đồng ý' 5 'Rất đồng ý'
C11_3	Sản phẩm đa dạng, phong phú	1 'Rất không đồng ý' 2 'Không đồng ý' 3 'Bình thường' 4 'Đồng ý' 5 'Rất đồng ý'
C11_4	Giá dịch vụ phù hợp	1 'Rất không đồng ý' 2 'Không đồng ý' 3 'Bình thường' 4 'Đồng ý' 5 'Rất đồng ý'
C11_5	Giao dịch thuận tiện	1 'Rất không đồng ý' 2 'Không đồng ý' 3 'Bình thường' 4 'Đồng ý' 5 'Rất đồng ý'
C11_6	Hình thức nhân viên ưu nhìn	1 'Rất không đồng ý' 2 'Không đồng ý' 3 'Bình thường' 4 'Đồng ý' 5 'Rất đồng ý'
C11_7	Đa dạng về thời gian bay	1 'Rất không đồng ý' 2 'Không đồng ý' 3 'Bình thường' 4 'Đồng ý' 5 'Rất đồng ý'
C11_8	Hạ tầng tốt	1 'Rất không đồng ý' 2 'Không đồng ý' 3 'Bình thường' 4 'Đồng ý' 5 'Rất đồng ý'



KIỂM ĐỊNH TRUNG  
BÌNH 1 TỔNG THỂ

# VỚI ĐỘ TIN CẬY 95% TIỀN LƯƠNG BÌNH QUÂN CỦA KHÁCH HÀNG CÓ LÀ 13000 HAY KHÔNG

Kiểm định trung bình

$$H_0 : \mu_1 = 13000$$

$$H_1 : \mu_1 \neq 13000$$

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Tiền lương	500	13859.62	7144.299	319.503

One-Sample Test

	Test Value = 13000					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Tiền lương	2.690	499	.007	859.620	231.88	1487.36

$$T_{qs} = 2,69; sig = .007$$

ĐỦ CƠ SỞ ĐỂ BÁC BỎ  $H_0$  ->  
GIẢ THUYẾT ĐÚNG

# CÓ Ý KIẾN CHO RẰNG TIỀN LƯƠNG BỘ CỦA KH NAM Ở MIỀN BẮC LỚN HƠN 11000 (NGHÌN ĐỒNG). HÃY KẾT LUẬN VỚI ĐỘ TIN CẬY 95%

$$H_0 : \mu = 11000$$

$$H_1 : \mu > 11000$$

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
cau2c4	91	11043.3407	757.35229	79.39209

One-Sample Test

Test Value = 11000						
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
cau2c4	.546	90	.586	43.34066	-114.3856	201.0669

CHƯA ĐỦ CƠ SỞ BÁC BỎ

$H_0 \rightarrow$  KẾT LUẬN SAI

$$T_{qs} = .546$$

$$Sig.(1 - \text{tailed}) = Sig.(2 - \text{tailed})/2 = 0.586/2 > 0.05$$



# KIỂM ĐỊNH TRUNG BÌNH NHIỀU TỔNG THỂ

KIỂM TRA 2 MAU LA ĐỘC LẬP HAY PHỤ THUỘC

# CÓ Ý KIẾN CHO RẰNG TIỀN LƯƠNG TRUNG BÌNH CỦA NAM KHÁC NỮ

Independent Samples Test									
		Levene's Test for Equality of Variances		t-test for Equality of Means					
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference
Tiền lương	Equal variances assumed	128.862	.000	11.113	498	.000	6385.818	574.605	5256.870 7514.766
	Equal variances not assumed			11.817	364.874	.000	6385.818	540.405	5323.120 7448.517

Kiểm định phương sai

$$H_0 : \sigma_1^2 = \sigma_2^2$$

$$H_1 : \sigma_1^2 \neq \sigma_2^2$$

⇒ F ⇒ Sig ⇒ BB H0 ⇒ PS khác nhau

Kiểm định trung bình

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

⇒ Tqs ⇒ Sig ⇒ BB H0 → Tiền lương khác nhau có ý nghĩa thống kê

→ Ý kiến đúng



# KIEM DINH TRUNG BINH NHIEU TONG THE

## \*\* Dieu kien

Các tổng thể phải độc lập và được chọn ngẫu nhiên

Các tổng thể phân phối chuẩn

Phương sai đồng nhất

Với độ tin cậy 95% có sự khác biệt TLTB giữa các khu vực hay không, nếu có thì khác biệt ở khu vực nào?

Kiểm định Levene (phương sai không bằng nhau → bác bỏ H0)

### Test of Homogeneity of Variances

Tiền lương

Levene Statistic	df1	df2	Sig.
205.132	2	497	.000

### Robust Tests of Equality of Means

Tiền lương

	Statistic <sup>a</sup>	df1	df2	Sig.
Welch	441.891	2	283.473	.000
Brown-Forsythe	310.727	2	217.658	.000

a. Asymptotically F distributed.

sig < 0.05 → bác bỏ H0 trung bình của các nhóm khác nhau

### Multiple Comparisons

Dependent Variable: Tiền lương

	(I) Khu vực sinh sống	(J) Khu vực sinh sống	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	Miền Bắc	Miền Nam	-8835.675*	526.240	.009	-10097.59	-7573.76
		Miền Trung	2540.607*	629.866	.009	1030.19	4051.02
	Miền Nam	Miền Bắc	8835.675*	526.240	.009	7573.76	10097.59
		Miền Trung	11376.282*	617.405	.009	9895.75	12856.81
	Miền Trung	Miền Bắc	-2540.607*	629.866	.009	-4051.02	-1030.19
		Miền Nam	-11376.282*	617.405	.009	-12856.81	-9895.75
Games-Howell	Miền Bắc	Miền Nam	-8835.675*	560.508	.009	-10185.65	-7485.70
		Miền Trung	2540.607*	105.243	.009	2287.20	2794.02
	Miền Nam	Miền Bắc	8835.675*	560.508	.009	7485.70	10185.65
		Miền Trung	11376.282*	563.373	.009	10019.60	12732.96
	Miền Trung	Miền Bắc	-2540.607*	105.243	.009	-2794.02	-2287.20
		Miền Nam	-11376.282*	563.373	.009	-12732.96	-10019.60

\*. The mean difference is significant at the 0.05 level.

Bảng kép so sánh giữa các nhóm

Phương sai không bằng nhau → Games-Howell (ở dưới)

→ Sig < 0.05 hết → có thể kết luận tất cả các khu vực đều khác nhau hết

# KIỂM ĐỊNH KHİ BÌNH PHƯƠNG

Có ý kiến cho rằng tỷ lệ khách là nhân viên kinh doanh = văn thư và gấp 5 lần quản lý

H0: tỷ lệ thực tế = tỷ lệ lý thuyết (kỳ vọng)

H1: tỷ lệ thực tế khác tỷ lệ lý thuyết (kỳ vọng)

Loại công việc

	Observed N	Expected N	Residual
Văn thư	239	227.3	11.7
Nhân viên kinh doanh	217	227.3	-10.3
Quản lý	44	45.5	-1.5
Total	500		

Test Statistics

	Loại công việc
Chi-Square	1.116 <sup>a</sup>
df	2
Asymp. Sig.	.572

a. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 45.5.

⇒ Chưa đủ cơ sở bác bỏ H0 → phát biểu trên là hợp lý

Có ý kiến cho rằng trình độ và loại hình công việc của KH có mối liên hệ với nhau. Hãy đưa ra kết luận với mức ý nghĩa 5%

H0: trình độ và loại hình công việc của khách hàng là độc lập

H1: trình độ và loại hình công việc của khách hàng là độc lập là phụ thuộc

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	41.248 <sup>a</sup>	4	.000
Likelihood Ratio	45.875	4	.000
Linear-by-Linear Association	21.016	1	.000
N of Valid Cases	500		

a. 1 cells (11.1%) have expected count less than 5. The minimum expected count is 2.46.

⇒ KL: Hệ số Sig = 0 < 0.05 → đủ cơ sở để bác bỏ H0

# KIỂM ĐỊNH VỀ DẠNG PHÂN PHỐI

/KIỂM ĐỊNH PHÂN PHỐI CỦA SỐ NĂM ĐI HỌC/

One-Sample Kolmogorov-Smirnov Test

		Số năm đi học
N		500
Normal Parameters <sup>a,b</sup>	Mean	13.47
	Std. Deviation	2.903
Most Extreme Differences	Absolute	.208
	Positive	.208
	Negative	-.190
Kolmogorov-Smirnov Z		4.644
Asymp. Sig. (2-tailed)		.000

a. Test distribution is Normal.

b. Calculated from data.

⇒ Hệ số Sig = 0 < 0.05 ⇒ Đủ cơ sở để bác bỏ  $H_0$

⇒ Phân phối không chuẩn

A series of white, thin, overlapping geometric lines on a light beige background, creating a complex, abstract pattern on the left side of the slide.

# THANK YOU FOR READING