

CHỦ ĐỀ 3.1 KIỂM ĐỊNH GIẢ THUYẾT MỘT MẪU

Trong chủ đề này, các em phải nắm và thực hiện được các nội dung như sau

- 1) Nắm được 5 bước cơ bản trong bài toán kiểm định gồm
 - ☑ B1. Phát biểu giả thuyết kiểm định.
 - ☑ B2. Xác định mức ý nghĩa α .
 - ☑ B3. Tính giá trị thống kê kiểm định.
 - ☑ B4. Xác định miền bác bỏ hoặc tính p -giá trị.
 - ☑ B5. So sánh và kết luận.
- 2) Thực hiện được kiểm định giả thuyết cho trung bình khi phương sai đã biết.
- 3) Thực hiện được kiểm định giả thuyết cho trung bình khi phương sai chưa biết.
- 4) Thực hiện được kiểm định giả thuyết cho tỷ lệ.

1. KIỂM ĐỊNH MỘT MẪU CHO TRUNG BÌNH

1.1. KIỂM ĐỊNH TRUNG BÌNH (SO SÁNH TRUNG BÌNH VỚI MỘT SỐ) TRƯỜNG HỢP PHƯƠNG SAI ĐÃ BIẾT

- ☑ B1. Phát biểu giả thuyết kiểm định.
- ☑ B2. Xác định mức ý nghĩa α .
- ☑ B3. Tính giá trị thống kê kiểm định:

$$Z_0 = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}.$$

- ☑ B4. Xác định miền bác bỏ.

Đối thuyết	Miền bác bỏ
$H_1 : \mu \neq \mu_0$	$ z_0 > z_{1-\alpha/2}$
$H_1 : \mu < \mu_0$	$z_0 < -z_{1-\alpha}$
$H_1 : \mu > \mu_0$	$z_0 > z_{1-\alpha}$

Ngược lại, chưa đủ cơ sở bác bỏ H_0 .

- ☑ B5. So sánh và kết luận.

Ở **Bước 4**, ta có thể sử dụng p -giá trị thay thế bằng cách tính p -giá trị theo bảng dưới đây

Đối thuyết	p -giá trị
$H_1 : \mu \neq \mu_0$	$p\text{-giá trị} = 2 [1 - \Phi(z_0)]$
$H_1 : \mu < \mu_0$	$p\text{-giá trị} = \Phi(z_0)$
$H_1 : \mu > \mu_0$	$p\text{-giá trị} = 1 - \Phi(z_0)$

Bác bỏ H_0 khi $p\text{-giá trị} < \alpha$. Ngược lại, chưa đủ cơ sở bác bỏ H_0 .

BÀI 3.1 (Câu 2 - Đề 1 HKI 22-23). Một hệ thống tên lửa phản lực sử dụng động cơ đẩy nhiên liệu rắn. Tốc độ cháy của nhiên liệu rắn là một đặc trưng quan trọng của động cơ. Các kỹ sư biết rằng độ lệch chuẩn của tốc độ cháy là 2 cm/s. Các kỹ sư chọn cỡ mẫu là $n = 25$ thu được trung bình mẫu tốc độ cháy là $\bar{x} = 51,3$ cm/s. Biết rằng tốc độ cháy tuân theo phân phối chuẩn.

- Tìm khoảng tin cậy 95% cho tốc độ cháy trung bình của nhiên liệu.
- Thông số kỹ thuật yêu cầu tốc độ cháy trung bình của thanh nhiên liệu là 50 cm/s. Dựa trên dữ liệu đã thu thập được, hãy kiểm định xem thông số kỹ thuật này có được đáp ứng hay không với mức ý nghĩa 1%.

BÀI 3.2 (Câu 2 - Đề 2 HKI 22-23). Nhiệt độ nước trung bình hạ lưu từ ống thoát xả giải nhiệt của nhà máy điện không được lớn hơn 100°F. Kinh nghiệm quá khứ đã chỉ ra rằng độ lệch chuẩn của nhiệt độ là 2°F. Nhiệt độ nước được đo trên chín ngày được lựa chọn ngẫu nhiên, và nhiệt độ trung bình được tìm thấy là 98°F. Biết rằng nhiệt độ nước tuân theo phân phối chuẩn.

- Tìm khoảng tin cậy 96% cho nhiệt độ nước trung bình.
- Có bằng chứng gì cho ta thấy nhiệt độ nước có thể chấp nhận được hay không với mức ý nghĩa 2%?

1.2. KIỂM ĐỊNH TRUNG BÌNH (SO SÁNH TRUNG BÌNH VỚI MỘT SỐ) TRƯỜNG HỢP PHƯƠNG SAI CHƯA BIẾT

☒ B1. Phát biểu giả thuyết kiểm định.

☒ B2. Xác định mức ý nghĩa α .

☒ B3. Tính giá trị thống kê kiểm định

$$T_0 = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$$

☒ B4. Xác định miền bác bỏ.

Đối thuyết	Miền bác bỏ
$H_1 : \mu \neq \mu_0$	$ t_0 > t_{\alpha/2;n-1}$
$H_1 : \mu < \mu_0$	$t_0 < -t_{\alpha;n-1}$
$H_1 : \mu > \mu_0$	$t_0 > t_{\alpha;n-1}$

Ngược lại, chưa đủ cơ sở bác bỏ H_0 .

☒ B5. Kết luận.

BÀI 3.3 (Câu 3 - Đề 2 CKI 19-20). Hàm lượng natri của hai mươi hộp bắp hữu cơ 300 gram được xác định. Dữ liệu (tính bằng miligam) như sau:

131.15 130.69 130.91 129.54 129.64 128.77 130.72 128.33 128.24 129.65
130.14 129.29 128.71 129.00 129.39 130.42 129.53 130.12 129.78 130.92

Giả sử rằng hàm lượng natri có phân phối chuẩn. Bạn hãy kiểm định xem giá trị trung bình có khác 130 milligram với $\alpha = 0.05$?

.....

.....

.....

.....

.....

.....

.....

.....

BÀI 3.4 (Câu 3 - Đề 3 CKII 19-20). Một máy đóng gói các sản phẩm có khối lượng 1 kg. Nghi ngờ máy hoạt động không bình thường, người ta chọn ra một mẫu ngẫu nhiên gồm 100 sản phẩm thì thấy như sau

Khối lượng	0.95	0.97	0.99	1.01	1.03	1.05
Số gói	9	31	40	15	3	2

Với mức ý nghĩa 5%, hãy kết luận về nghi ngờ trên.

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

BÀI 3.5 (Câu 2 - Đề 2 CKII 19-20). Một trung tâm khám chữa bệnh tuyên bố rằng thời gian trung bình một bệnh nhân chờ khám không quá 20 phút. Một cuộc khảo sát ngẫu nhiên 35 bệnh nhân cho thấy thời gian chờ khám trung bình là 24.77 phút và độ lệch chuẩn là 7.26 phút. Giả sử thời gian chờ khám là đại lượng ngẫu nhiên có phân phối chuẩn. Dựa vào dữ liệu khảo sát, hãy kiểm tra tuyên bố của phòng khám đó có đúng không với mức ý nghĩa 3%.

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

2. KIỂM ĐỊNH MỘT MẪU CHO TỶ LỆ

Với Y là số phần tử có tính chất A trong mẫu gồm n phần tử.

Mẫu: $n = \dots; y = \dots \Rightarrow$ tỷ lệ mẫu $\hat{P} = \frac{y}{n}$.

- ☑ B1. Phát biểu giả thuyết kiểm định.
- ☑ B2. Xác định mức ý nghĩa α .
- ☑ B3. Tính giá trị thống kê kiểm định

$$Z_0 = \frac{\hat{P} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}}$$

- ☑ B4. Xác định miền bác bỏ.

Đối thuyết	Miền bác bỏ H_0
$H_1 : p \neq p_0$	$ z_0 > z_{1-\alpha/2}$
$H_1 : p < p_0$	$z_0 < -z_{1-\alpha}$
$H_1 : p > p_0$	$z_0 > z_{1-\alpha}$

Ngược lại, chưa đủ cơ sở bác bỏ H_0 .

- ☑ B5. Kết luận.

Ở **Bước 4**, ta có thể sử dụng p -giá trị thay thế bằng cách tính p -giá trị theo bảng dưới đây

Đối thuyết	p -giá trị
$H_1 : p \neq p_0$	$p\text{-giá trị} = 2 [1 - \Phi(z_0)]$
$H_1 : p < p_0$	$p\text{-giá trị} = \Phi(z_0)$
$H_1 : p > p_0$	$p\text{-giá trị} = 1 - \Phi(z_0)$

Bác bỏ H_0 khi p -giá trị $< \alpha$. Ngược lại, chưa đủ cơ sở bác bỏ H_0 .

BÀI 3.6 (Câu 2 - Đề 1 CKI 20-21). Giả sử rằng trong 1000 khách hàng được khảo sát có 850 người hài lòng hoặc rất hài lòng với các sản phẩm và dịch vụ của công ty. Gọi p là tỷ lệ người hài lòng hoặc rất hài lòng trong tất cả khách hàng. Kiểm định giả thuyết $H_0 : p = 0.9$ với đối thuyết $H_1 : p \neq 0.9$ với $\alpha = 0.05$. Tìm p -giá trị.

BÀI 3.7 (Câu 2 - Đề 2 CKI 20-21). Giả sử người ta kiểm tra 500 thành phần máy móc do một nhà máy sản xuất và thấy có 10 thành phần bị loại bỏ. Gọi p là tỷ lệ thành phần bị loại bỏ do nhà máy sản xuất. Kiểm định giả thuyết $H_0 : p = 0,03$ với đối thuyết $H_1 : p < 0,03$, sử dụng $\alpha = 0,05$. Tìm p -giá trị.

BÀI 3.8 (Câu 4 - Đề 3 CKII 19-20). Một bài báo trên tạp chí y khoa Anh “So sánh điều trị sỏi thận bằng phẫu thuật phẫu thuật, cắt bỏ sỏi thận, và Lithotripsy sóng nổ xung,” (1986, Vol. 292, pp. 879–882)] thấy rằng tác động qua da (PN) có tỷ lệ thành công trong việc loại bỏ sỏi thận của 289 trong số 350 bệnh nhân. Phương pháp truyền thống đạt hiệu quả 78%. Có bằng chứng gì cho thấy tỉ lệ thành công của PN lớn hơn so với truyền thống với mức ý nghĩa $\alpha = 1\%$? Tìm p -giá trị.

.....

.....

.....

.....

.....

.....

.....

.....

.....

BÀI 3.9 (Câu 2 - Đề 1 HKI 22-23). Một nhà sản xuất chất bán dẫn sản xuất bộ điều khiển sử dụng trong công nghệ động cơ ô tô. Nhà sản xuất bán dẫn lấy mẫu ngẫu nhiên gồm 200 sản phẩm và thấy rằng có bốn sản phẩm bị lỗi.

- a) Tìm khoảng tin cậy 95% cho tỷ lệ lỗi của sản phẩm.
 - b) Khách hàng yêu cầu tỷ lệ lỗi của sản phẩm phải dưới 0,05. Hỏi nhà sản xuất có thể chứng minh khả năng đáp ứng ở mức chất lượng này cho khách hàng không? Sử dụng mức ý nghĩa 1%.
-
-
-
-
-
-
-
-
-

.....

BÀI 3.10 (Câu 2 - Đề 2 HKI 22-23). Trong một mẫu ngẫu nhiên gồm 85 vòng bi trục khuỷu động cơ ô tô trong đó có 10 vòng bi độ nhám bề mặt hoàn thiện vượt quá các thông số kỹ thuật (gọi tắt là vượt chuẩn).

- a) Tìm khoảng tin cậy 96% cho tỷ lệ vòng bi vượt chuẩn.
- b) Dữ liệu này có cho thấy rằng tỷ lệ vòng bi vượt chuẩn là cao hơn 0.10 hay không với mức ý nghĩa 2%.

BÀI 3.11 (Câu 1 - Đề 2 CKI 18-19). Thực hiện một khảo sát xã hội về số tiền chi trả cho các hoạt động vui chơi giải trí trong 1 tháng của 400 thanh niên tại TP.HCM người ta thu được bảng sau:

Số tiền (USD)	50 – 80	80 – 120	120 – 160	160 – 200	200 – 220	220 – 250
Số người	50	80	100	80	60	30

Giả thiết số tiền phải bỏ ra cho các hoạt động vui chơi giải trí trong một tháng của một thanh niên tại TP.HCM là một đại lượng ngẫu nhiên phân phối theo qui luật chuẩn.

- Ước lượng số tiền trung bình một thanh niên phải bỏ ra với độ tin cậy 95%. (1.5đ)
- Những thanh niên bỏ ra trên 200 USD/tháng cho các hoạt động vui chơi là những thanh niên khá giả. Hãy ước lượng tỉ lệ những thanh niên khá giả với độ tin cậy 97%. Nếu muốn sai số ≤ 0.01 thì phải khảo sát thêm bao nhiêu thanh niên? (2.5đ)
- Một nhà nghiên cứu xã hội cho rằng cứ 100 thanh niên ở TP.HCM thì có 30 người thuộc diện khá giả, trong khi nhà thống kê lại tỏ ra nghi ngờ và họ cho rằng con số này thực sự phải nhỏ hơn con số thống kê do nhà nghiên cứu ngày đưa ra. Vậy theo các bạn, ý kiến nào đúng dẫn với mức ý nghĩa $\alpha = 5\%$. (1.5đ)
- Lời khẳng định: “Tỷ lệ thanh niên có thu nhập hạn chế là 50%” có được chấp nhận hay không, mức ý nghĩa 1%. Biết rằng một thanh niên được gọi là có thu nhập hạn chế nếu số tiền bỏ ra cho hoạt động vui chơi dưới 120 USD/tháng. (1.5đ)

BÀI 3.12 (Câu 1 - Đề 1 CKII 21-22). Biết trọng lượng X (g/quả) của mỗi quả trứng có phân phối chuẩn. Dem cân 100 quả trứng ta có kết quả sau:

x_i	155	160	165	170	175	180	185
n_i	5	12	14	25	24	14	6

Cho biết trứng có trọng lượng **lớn hơn** 170 g là trứng loại một.

- Tìm khoảng tin cậy 97% cho trọng lượng trứng trung bình.
- Tìm khoảng tin cậy 98% cho tỷ lệ trứng loại một. Nếu ta muốn sai số ước lượng không quá 0.1 g thì cần khảo sát thêm bao nhiêu trứng?
- Có ý kiến cho rằng trọng lượng trứng trung bình lớn hơn 170 g/quả. Hãy kiểm định ý kiến trên ứng với mức ý nghĩa 1%.
- Có ý kiến cho rằng 50% số trứng thuộc loại một. Hãy kiểm định ý kiến trên với mức ý nghĩa 1%.

BÀI 3.13 (Câu 1 - Đề 2 CKII 20-21). Kết quả quan sát về hàm lượng Vitamin C của một loại trái cây cho bởi bảng sau

Hàm lượng Vitamin C (mg)	6	8	10	12	14	16
Số trái	5	10	20	35	25	5

- Ước lượng hàm lượng Vitamin C trung bình trong một trái cây với độ tin cậy 90%.
- Những trái có hàm lượng Vitamin C trên 11 mg trở lên là trái loại I. Ước lượng tỷ lệ trái loại I với độ tin cậy 92%.
- Có ý kiến rằng hàm lượng Vitamin C trung bình là 10 mg. Với mức ý nghĩa 6%, hãy kiểm tra ý kiến trên.
- Lời khẳng định: “Tỷ lệ trái cây loại I là 60%” có được chấp nhận hay không với mức ý nghĩa 4%.

CHỦ ĐỀ 3.2 KIỂM ĐỊNH GIẢ THUYẾT HAI MẪU

Trong chủ đề này, các em phải nắm và thực hiện được các nội dung như sau

- 1) Nắm được 5 bước cơ bản trong bài toán kiểm định gồm
 - ☑ B1. Phát biểu giả thuyết kiểm định.
 - ☑ B2. Xác định mức ý nghĩa α .
 - ☑ B3. Tính giá trị thống kê kiểm định.
 - ☑ B4. Xác định miền bác bỏ hoặc tính p -giá trị.
 - ☑ B5. So sánh và kết luận.
- 2) Thực hiện được kiểm định giả thuyết so sánh hai trung bình khi phương sai đã biết.
- 3) Thực hiện được kiểm định giả thuyết so sánh hai trung bình khi phương sai bằng nhau chưa biết.
- 4) Thực hiện được kiểm định giả thuyết so sánh hai trung bình khi phương sai khác nhau chưa biết.
- 5) Thực hiện được kiểm định giả thuyết so sánh hai tỷ lệ.

3. KIỂM ĐỊNH SO SÁNH HAI TRUNG BÌNH Ở HAI MẪU ĐỘC LẬP

3.1. KIỂM ĐỊNH GIẢ THUYẾT SO SÁNH HAI TRUNG BÌNH KHI PHƯƠNG SAI ĐÃ BIẾT

Mẫu 1: $n_1 = \dots; \bar{x}_1 = \dots; \sigma_1^2 = \dots$: đã biết.

Mẫu 2: $n_2 = \dots; \bar{x}_2 = \dots; \sigma_2^2 = \dots$: đã biết.

- ☑ B1. Phát biểu giả thuyết kiểm định $\Rightarrow \Delta_0 = \mu_1 - \mu_2 = \dots$
- ☑ B2. Xác định mức ý nghĩa α .
- ☑ B3. Tính giá trị thống kê kiểm định (TKKD)

$$Z_0 = \frac{\bar{X}_1 - \bar{X}_2 - \Delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

- ☑ B4. Xác định miền bác bỏ hoặc tính p - giá trị

Đối thuyết	Miền bác bỏ H_0	p -giá trị
$H_1 : \mu_1 - \mu_2 \neq \Delta_0$	$ z_0 > z_{1-\alpha/2}$	$p\text{-giá trị} = 2[1 - \Phi(z_0)]$
$H_1 : \mu_1 - \mu_2 < \Delta_0$	$z_0 < -z_{1-\alpha}$	$p\text{-giá trị} = \Phi(z_0)$
$H_1 : \mu_1 - \mu_2 > \Delta_0$	$z_0 > z_{1-\alpha}$	$p\text{-giá trị} = 1 - \Phi(z_0)$

Ngược lại, chưa đủ cơ sở bác bỏ H_0 .

☑ **B5.** So sánh và kết luận.

3.2. KIỂM ĐỊNH GIẢ THUYẾT SO SÁNH HAI TRUNG BÌNH KHI PHƯƠNG SAI BẰNG NHAU CHƯA BIẾT

Mẫu 1: $n_1 = \dots; \bar{x}_1 = \dots; s_1^2 = \dots$

Mẫu 2: $n_2 = \dots; \bar{x}_2 = \dots; s_2^2 = \dots$

Phương sai không biết giá trị nhưng biết rằng $\sigma_1^2 = \sigma_2^2$.

☑ **B1.** Phát biểu giả thuyết kiểm định $\Rightarrow \Delta_0 = \mu_1 - \mu_2 = \dots$

☑ **B2.** Xác định mức ý nghĩa α .

☑ **B3.** Tính giá trị thống kê kiểm định (TKKD)

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

$$T_0 = \frac{\bar{X}_1 - \bar{X}_2 - \Delta_0}{\sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}}$$

☑ **B4.** Xác định miền bác bỏ hoặc tính p -giá trị

Đối thuyết	Miền bác bỏ	p -giá trị
$H_1 : \mu_1 - \mu_2 \neq \Delta_0$	$ t_0 > t_{\alpha/2, n_1+n_2-2}$	p -giá trị $= 2\mathbb{P}(T \geq t_0)$
$H_1 : \mu_1 - \mu_2 < \Delta_0$	$t_0 < -t_{\alpha, n_1+n_2-2}$	p -giá trị $= \mathbb{P}(T \leq t_0)$
$H_1 : \mu_1 - \mu_2 > \Delta_0$	$t_0 > t_{\alpha, n_1+n_2-2}$	p -giá trị $= \mathbb{P}(T \geq t_0)$

Ngược lại, chưa đủ cơ sở bác bỏ H_0 .

☑ **B5.** So sánh và kết luận.

3.3. KIỂM ĐỊNH GIẢ THUYẾT SO SÁNH HAI TRUNG BÌNH KHI PHƯƠNG SAI KHÁC NHAU CHƯA BIẾT

Mẫu 1: $n_1 = \dots; \bar{x}_1 = \dots; s_1^2 = \dots$

Mẫu 2: $n_2 = \dots; \bar{x}_2 = \dots; s_2^2 = \dots$

Phương sai không biết giá trị nhưng biết rằng $\sigma_1^2 \neq \sigma_2^2$.

☑ **B1.** Phát biểu giả thuyết kiểm định $\Rightarrow \Delta_0 = \mu_1 - \mu_2 = \dots$

☑ **B2.** Xác định mức ý nghĩa α .

☑ **B3.** Tính giá trị thống kê kiểm định (TKKD)

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

☑ **B4.** Xác định miền bác bỏ hoặc tính p -giá trị Với

$$v = \left\lfloor \frac{[(S_1^2/n_1) + (S_2^2/n_2)]^2}{\frac{(S_1^2/n_1)^2}{n_1 - 1} + \frac{(S_2^2/n_2)^2}{n_2 - 1}} \right\rfloor$$

trong đó $\lfloor a \rfloor$ là phép lấy phần nguyên và nhỏ hơn a , ví dụ $\lfloor 38.89 \rfloor = 38$.

Đối thuyết	Miền bác bỏ	p -giá trị
$H_1 : \mu_1 - \mu_2 \neq \Delta_0$	$ t_0 > t_{\alpha/2, v}$	p -giá trị $= 2\mathbb{P}(T \geq t_0)$
$H_1 : \mu_1 - \mu_2 < \Delta_0$	$t_0 < -t_{\alpha, v}$	p -giá trị $= \mathbb{P}(T \leq t_0)$
$H_1 : \mu_1 - \mu_2 > \Delta_0$	$t_0 > t_{\alpha, v}$	p -giá trị $= \mathbb{P}(T \geq t_0)$

Ngược lại, chưa đủ cơ sở bác bỏ H_0 .

☑ **B5.** So sánh và kết luận.

BÀI 3.14 (Câu 4 - Đề 1 CKI 20-21). Hai máy được sử dụng để rót đầy các chai nhựa. Khối lượng được rót từ hai máy được giả định có phân phối chuẩn với độ lệch chuẩn lần lượt là $\sigma_1 = 0.020$ và $\sigma_2 = 0.025$ ounce. Một nhân viên kỹ thuật cho rằng khối lượng trung bình được rót hai máy là bằng nhau. Hai mẫu ngẫu nhiên từ mỗi máy như sau

Máy 1	16.03	16.01	16.04	15.96	16.05	15.98	16.05	16.02	16.02	15.99
Máy 2	16.02	16.03	15.97	16.04	15.96	16.02	16.01	16.01	15.99	16.00

Kiểm định ý kiến của nhân viên kỹ thuật trên. Sử dụng $\alpha = 0.05$. Tìm p -giá trị.

.....

.....

.....

.....

.....

.....

.....

BÀI 3.16 (Câu 4 - Đề 1 - CKI 19-20). Hai loại nhựa phù hợp để sử dụng cho một nhà sản xuất linh kiện điện tử. Sức chịu phá hủy của loại nhựa này là quan trọng và được giả sử có phân phối chuẩn. Được biết, $\sigma_1 = \sigma_2 = 1$ psi. Từ một mẫu ngẫu nhiên có kích thước $n_1 = 10$ và $n_2 = 12$, ta có được $\bar{x}_1 = 162.5$ và $\bar{x}_2 = 155.0$. Công ty sẽ không áp dụng nhựa loại 1 trừ khi sức chịu phá hủy trung bình của nó vượt quá nhựa loại 2 ít nhất 10 psi.

Trên cơ sở thông tin đó, ta có nên sử dụng nhựa loại 1? Sử dụng $\alpha = 0.05$ đưa ra câu trả lời.

BÀI 3.19 (Câu 5 - Đề 1 CKI 21-22). Trong sản xuất chất bán dẫn, khắc hóa chất ướt thường được sử dụng để loại bỏ silic từ mặt sau của tấm wafer trước khi kim loại hóa. Tỷ lệ ăn mòn (etch) là một đặc tính quan trọng trong quá trình này và được biết là tuân theo phân phối chuẩn. Hai phương pháp khắc khác nhau đã được so sánh bằng cách sử dụng hai mẫu ngẫu nhiên gồm 8 tấm wafer cho mỗi dung dịch. Gọi X và Y lần lượt là tỷ lệ ăn mòn tương ứng với phương pháp khắc I và II. Kết quả quan sát được như sau

Phương pháp I (x_i)	11.1	11.2	11.4	11.1	11.7	11.4	11.2	11.1
Phương pháp II (y_i)	9.1	8.9	8.9	9.1	8.9	9.2	9	9.3

18 | Lý thuyết Xác suất và Thống kê

BÀI 3.20 (Câu 3 - Đề 2 CKII 20-21). Để tìm ra liệu một loại huyết thanh mới có kìm hãm được bệnh bạch cầu hay không, 9 con chuột, tất cả các con đều trong giai đoạn tiến triển của bệnh, được chọn. Năm con chuột nhận trị liệu và 4 con không. Thời gian sống, theo năm, từ thời điểm thí nghiệm bắt đầu là như sau

Trị liệu	2.1	5.3	1.4	4.6	0.9
Không trị liệu	1.9	0.5	2.8	3.1	

Tại mức ý nghĩa 5%, huyết thanh có thể được nói là có hiệu quả hay không? Giả sử hai tổng thể có phân phối chuẩn với các phương sai bằng nhau.

BÀI 3.22 (Câu 3 - Đề 1 CKII 20-21). Một nghiên cứu được thực hiện để xem việc tăng nồng độ cơ chất có tác dụng đáng kể đến tốc độ của một phản ứng hóa học hay không. Với một nồng độ cơ chất 1.5 mol/l, phản ứng được chạy 15 lần, với tốc độ trung bình 7.5 micromoles mỗi 30 phút và độ lệch chuẩn 1.5. Với một nồng độ cơ chất 2.0 mol/l, 12 lần chạy được thực hiện, thu được tốc độ trung bình 8.8 micromoles mỗi 30 phút và độ lệch chuẩn 1.2. Có lý do để tin rằng việc tăng nồng độ cơ chất làm tăng tốc độ trung bình của phản ứng 0.5 micromoles mỗi 30 phút hay không? Sử dụng mức ý nghĩa 0.01 và giả sử rằng các tổng thể là xấp xỉ chuẩn với các phương sai bằng nhau.

BÀI 3.23 (Câu 4 - Đề 2 CKII 21-22 (CNTT hệ CLC)). Một công ty sản xuất xe ô tô chế tạo ra một dòng xe ô tô **Z**, công ty kiểm tra xe bằng cách cho 30 xe chạy trên cùng 1 quãng đường 100 km và đo lượng xăng sử dụng X (Đv: lít), kết quả cho ở bảng sau

X (lít)	7.0	7.1	7.2	7.3	7.4	7.5
Số xe	5	7	6	6	3	3

Giả sử lượng xăng tiêu hao là đại lượng ngẫu nhiên có phân phối chuẩn.

- Lời khẳng định: “Lượng xăng tiêu hao trung bình của dòng xe này là 7.2 lít (trên 100 km)” có được chấp nhận hay không? (Mức ý nghĩa $\alpha = 5\%$)
- Những xe có lượng xăng tiêu hao từ 7.3 lít trở lên được xếp vào loại không đạt tiêu chuẩn về nhiên liệu. Giám đốc công ty khẳng định rằng tỷ lệ xe không đạt tiêu chuẩn về nhiên liệu tối đa bằng 10%. Với dữ liệu khảo sát đã cho, ta có đủ bằng chứng để bác bỏ ý kiến trên không? $\alpha = 1\%$.
- Khảo sát lượng xăng tiêu hao Y (Đv: lít/100 km) trên một mẫu gồm 25 xe của dòng xe **W** cùng hãng, tính được $\bar{y} = 7.35$ (lít/100 km) và $s_Y = 0.2$. Có ý kiến cho rằng dòng xe **Z** tiêu hao ít nhiên liệu hơn dòng xe **W**. Với mức ý nghĩa $\alpha = 1\%$, hãy kiểm định ý kiến trên. Giả sử phương sai của X và Y bằng nhau.

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

BÀI 3.24 (Câu 4 - Đề 2 CKI 21-22). Trong một nhà máy sản xuất linh kiện điện tử, các kỹ sư cân nhắc lựa chọn giữa hai loại nhựa để tiến hành sản xuất linh kiện. Một yếu tố quan trọng được cân nhắc là sức chịu lực phá hủy (đơn vị: psi) của từng loại nhựa. Đối với loại nhựa 1, các kỹ sư kiểm tra trên một mẫu cỡ $m = 13$, tính được sức chịu lực phá hủy trung bình là $\bar{x} = 163.25$ và độ lệch chuẩn mẫu $s_x = 1.1$. Thử nghiệm trên một mẫu cỡ $n = 14$ đối với loại nhựa 2, tính được sức chịu lực phá hủy trung bình là $\bar{y} = 163.5$ và độ lệch chuẩn mẫu $s_y = 1.95$. Với mức ý nghĩa $\alpha = 3\%$, ta có thể khẳng định rằng sức chịu lực phá hủy của hai loại nhựa này là như nhau hay không? Giả sử rằng sức chịu lực phá hủy của hai loại nhựa tuân theo phân phối chuẩn và có phương sai khác nhau.

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

4. KIỂM ĐỊNH GIẢ THUYẾT SO SÁNH HAI TỶ LỆ

Gọi Y_1, Y_2 lần lượt là số phần tử có tính chất A trong mẫu 1 (có n_1 phần tử), mẫu 2 (có n_2 phần tử) tương ứng.

Mẫu 1: $n_1 = \dots; y_1 = \dots \Rightarrow$ tỷ lệ mẫu $\hat{P}_1 = \frac{y_1}{n_1}$.

Mẫu 2: $n_2 = \dots; y_2 = \dots \Rightarrow$ tỷ lệ mẫu $\hat{P}_2 = \frac{y_2}{n_2}$.

$$\Rightarrow \hat{P} = \frac{y_1 + y_2}{n_1 + n_2}$$

☑ B1. Phát biểu giả thuyết kiểm định $\Rightarrow \Delta_0 = p_1 - p_2$.

☑ B2. Xác định mức ý nghĩa α .

☑ B3. Tính giá trị thống kê kiểm định (TKKD)

$$Z_0 = \frac{\hat{P}_1 - \hat{P}_2 - \Delta_0}{\sqrt{\hat{P}(1 - \hat{P}) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

B4. Xác định miền bác bỏ hoặc tính p -giá trị.

Đối thuyết	Miền bác bỏ H_0	p -giá trị
$H_1 : p_1 - p_2 \neq \Delta_0$	$ z_0 > z_{1-\alpha/2}$	$p\text{-giá trị} = 2[1 - \Phi(z_0)]$
$H_1 : p_1 - p_2 < \Delta_0$	$z_0 < -z_{1-\alpha}$	$p\text{-giá trị} = \Phi(z_0)$
$H_1 : p_1 - p_2 > \Delta_0$	$z_0 > z_{1-\alpha}$	$p\text{-giá trị} = 1 - \Phi(z_0)$

Ngược lại, chưa đủ cơ sở bác bỏ H_0 .

☑ B5. So sánh và kết luận.

BÀI 3.25 (Câu 4 - Đề 2 HKI 22-23). Trong một nghiên cứu để ước tính tỷ lệ cư dân trong một thành phố nào đó và các vùng ngoại ô của nó ủng hộ việc xây dựng nhà máy năng lượng hạt nhân, người ta thấy rằng 63 trong 100 cư dân thành thị ủng hộ việc xây dựng, trong khi chỉ 59 trong 125 cư dân ngoại ô là ủng hộ. Hỏi có sự khác biệt giữa tỷ lệ cư dân thành thị và ngoại ô ủng hộ việc xây dựng nhà máy hạt nhân hay không với mức ý nghĩa 2%? Sử dụng p -giá trị.

📖 LỜI GIẢI.

Mẫu 1: $y_1 = 63; n_1 = 100 \Rightarrow \hat{p}_1 = \frac{63}{100} = 0.63$.

Mẫu 2: $y_2 = 59; n_2 = 125 \Rightarrow \hat{p}_2 = \frac{59}{125} = 0.427$.

$$\Rightarrow \hat{p} = \frac{y_1 + y_2}{n_1 + n_2} = \frac{63 + 59}{100 + 125} = \frac{122}{225}.$$

☑ Giả thuyết kiểm định: $\begin{cases} H_0 : p_1 = p_2; \\ H_1 : p_1 \neq p_2 \end{cases}; \quad \Rightarrow \Delta_0 = p_1 - p_2 = 0.$

☑ Giá trị thống kê kiểm định

$$z_0 = \frac{\hat{p}_1 - \hat{p}_2 - \Delta_0}{\sqrt{\hat{p}(1 - \hat{p}) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} = \frac{0.63 - 0.427 - 0}{\sqrt{\frac{122}{225} \left(1 - \frac{122}{225} \right) \left(\frac{1}{100} + \frac{1}{125} \right)}} \approx 3.03699.$$

☑ p -giá trị $= 2(1 - \Phi(|3.03699|)) = 2(1 - 0.99881) = 0.00238$.

Ta có p -giá trị $= 0.00238 < 0.02 = \alpha$

\Rightarrow bác bỏ $H_0 : p_1 = p_2$.

Kết luận: Với mức ý nghĩa 2% thì có sự khác biệt có ý nghĩa giữa tỷ lệ cư dân thành thị và ngoại ô trong ủng hộ việc xây dựng nhà máy hạt nhân.

□

BÀI 3.26 (Câu 3 - Đề 1 CKII 21-22). Hai loại giải pháp khác nhau để đánh bóng thấu kính nội nhãn (được dùng trong mắt người sau phẫu thuật đục thủy tinh thể) đang được đánh giá để sử dụng. Trong 300 thấu kính đã được đánh bóng bằng giải pháp 1 thì có 253 thấu kính không có khuyết tật do đánh bóng. Trong 300 thấu kính khác được đánh bóng bằng giải pháp 2 thì có 196 thấu kính không có khuyết tật do đánh bóng. Có lý do nào để tin rằng hai giải pháp đánh bóng là khác nhau không? Sử dụng $\alpha = 0.05$. p -giá trị cho kiểm định này là bao nhiêu?

BÀI 3.28 (Câu 4 - Đề 2 CKI 21-22). Các tác giả của bài báo “Adjuvant Radiotherapy an Chemotherapy in Node-Positive Premenopausal Women with Breast Cancer” (New Engl. J. of Med., 1997: 956-962) báo cáo các kết quả nghiên cứu của họ về một thí nghiệm được thiết kế để so sánh hiệu quả của 2 phương pháp điều trị bệnh ung thư: chỉ điều trị bằng hóa trị (phương pháp 1) và điều trị kết hợp giữa hóa trị và xạ trị (phương pháp 2). Trong số 183 bệnh nhân được điều trị bằng phương pháp 1, có 90 bệnh nhân sống sót ít nhất 15 năm, trong khi 108 trong số 173 bệnh nhân mà được điều trị bằng phương pháp 2 sống sót ít nhất 15 năm. Có ý kiến cho rằng phương pháp điều trị kết hợp giữa hóa trị và xạ trị có hiệu quả hơn là chỉ điều trị bằng hóa trị. Dựa trên kết quả thí nghiệm, ta có đủ bằng chứng ủng hộ ý kiến trên không với mức ý nghĩa $\alpha = 2\%$?

- Tìm khoảng tin cậy 95% cho tỷ lệ bệnh nhân sống sót ít nhất 15 năm khi điều trị bằng phương pháp I.
- Có ý kiến cho rằng phương pháp điều trị kết hợp giữa hóa trị và xạ trị có hiệu quả hơn là chỉ điều trị bằng hóa trị. Dựa trên kết quả thí nghiệm, ta có đủ bằng chứng ủng hộ ý kiến trên không với mức ý nghĩa $\alpha = 1\%$?

BÀI 3.31 (Câu 1 - Đề 1 CKI 18-19). Gọi X (giờ) là thời gian tự học hàng ngày của sinh viên, khảo sát 120 sinh viên trường Đại học KHXHNV. Kết quả cho bởi bảng sau:

BÀI 3.32 (Câu 2 - Đề 3 CKI 17-18). Một khảo sát về chiều cao X (cm) của một giống cây trồng người ta quan sát một mẫu và có kết quả như sau:

Chiều cao (cm)	100	110	120	130	140	150	160
Số cây	10	10	15	30	10	10	15

Giả sử chiều cao X có phân phối chuẩn.

- Ước lượng chiều cao trung bình của giống cây trồng trên với độ tin cậy 95%.
- Những cây trồng có chiều cao từ 135 cm trở lên được gọi là những cây “cao”. Hãy ước lượng tỉ lệ những cây cao với độ tin cậy 95%.
- Người ta áp dụng phương pháp mới trong việc trồng và chăm sóc cây. Sau một thời gian, khảo sát 100 cây trồng theo phương pháp mới được bảng số liệu sau:

Chiều cao (cm)	100	110	120	130	140	150	160
Số cây	6	10	20	34	12	7	11

Với mức ý nghĩa 5%, hãy kiểm định xem phương pháp mới có làm tăng chiều cao trung bình của cây hay không?

- Có ý kiến cho rằng phương pháp mới làm tăng tỉ lệ cây “cao”. Với mức ý nghĩa 5%, hãy kiểm tra ý kiến này.

BÀI 3.33 (Câu 3 - Đề 4 CKI 17-18). Để xem xét tình hình học tập môn xác suất thống kê (XSTK), một giảng viên (GV) đã tiến hành lấy mẫu (2 lần). Biết rằng điểm số sinh viên tuân theo phân phối chuẩn. Điểm sinh viên nam và sinh viên nữ độc lập nhau.

- a) Ở lần lấy mẫu thứ nhất, GV này thu thập được thông tin của 20 sinh viên với điểm trung bình 5.5 và độ lệch chuẩn 0.75. Sử dụng thông tin này để ước lượng điểm trung bình của sinh viên với độ tin cậy 95%.

GV này tiếp tục thực hiện lấy mẫu lần thứ hai và thu thập được bảng thông tin sau:

Điểm	[0, 1)	[1, 2)	[2, 3)	[3, 4)	[4, 5)	[5, 6)	[6, 7)	[7, 8)	[8, 9)	[9, 10]
Nam	4	6	4	6	2	6	5	2	6	4
Nữ	5	10	8	8	5	4	3	2	3	2

(Bảng dữ liệu chứa thông tin: có 4 sinh viên nam và 5 sinh viên nữ có điểm số nằm trong $[0, 1)$, có 6 sinh viên nam và 10 sinh viên nữ có điểm số nằm trong $[1, 2)$,.... Hãy sử dụng thông tin ở lần lấy mẫu thứ hai để giải các câu hỏi (b,c,d)

- b) Một sinh viên được gọi là có điểm cao nếu điểm lớn hơn hoặc bằng 7.0. Ước lượng khoảng tin cậy cho tỉ lệ học sinh có điểm cao với độ tin cậy 99%.
- c) Dựa theo số liệu điểm môn XSTK trong học kì trước, điểm trung bình của sinh viên **nam** là 5.0. Hãy cho biết giá trị trên có phù hợp với dữ liệu quan sát (điểm sinh viên nam đã thu thập được) hay không với mức ý nghĩa 2%?
- d) Có nhiều người cho rằng các bạn nữ chăm học hơn các bạn nam nên điểm trung bình của các bạn nữ cao hơn điểm trung bình của các bạn nam. Tuy nhiên, GV này cho rằng điểm trung bình của các bạn nữ **không cao hơn** điểm trung bình của các bạn nam. Hãy sử dụng dữ liệu quan sát để kiểm tra nhận định về hai điểm trung bình của GV với mức ý nghĩa 2%.

