

HƯỚNG DẪN THỰC HÀNH MÔN XÁC SUẤT THỐNG KÊ

Tài liệu này hướng dẫn cho sinh viên Trường Đại học Xây dựng sử dụng công cụ tính toán hỗ trợ việc học tập và thi môn Xác suất thống kê với 12 lệnh của phần mềm Mathematica và 11 lệnh Excel.

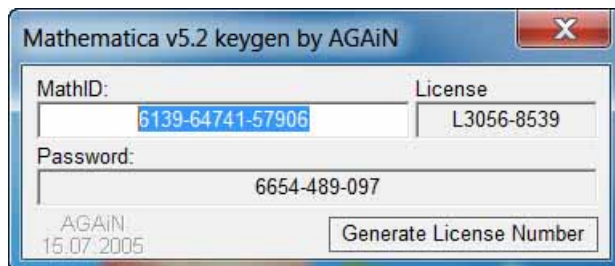
Yêu cầu: tự học 10 tiết, tổng thời gian thực hành làm mẫu trên lớp 3 tiết, thực hành trên phòng máy 3 tiết.

1 Mathematica

Download phần mềm và xem video hướng dẫn thực hành lần lượt tại:

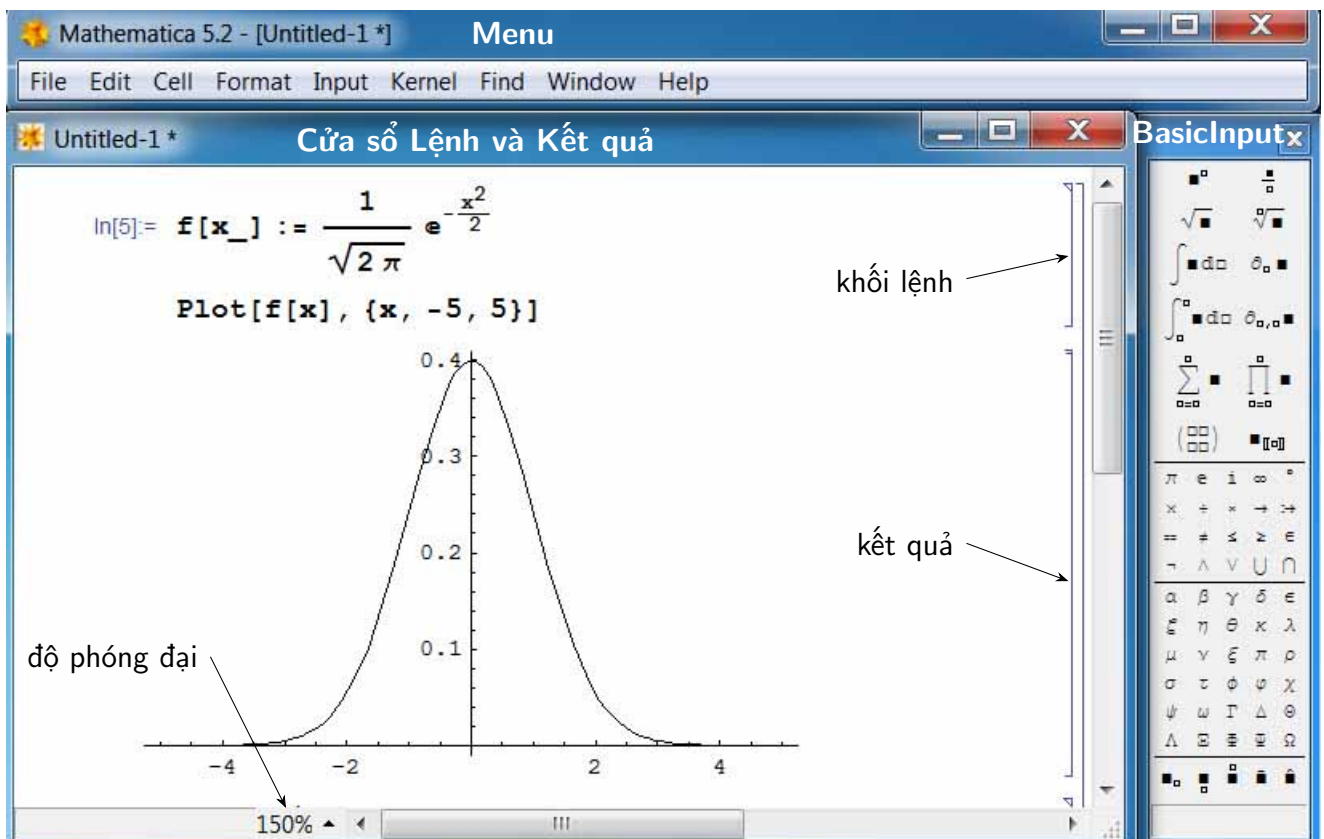
megaurl.in/xstk megaurl.in/XstkMathematicaExcel

Điền thông tin lấy từ file Keygen.exe vào mục tương ứng trong quá trình cài đặt (mỗi máy có MathID riêng):



- * *Mathematica*: Chép License từ Keygen, bấm OK.
- * *Keygen*: Lấy MathID từ Mathematica → Generate License Number.
- * *Mathematica*: bấm Back, chép License mới vào → OK → chép Password vào → OK.

Giao diện:



Cách sử dụng Mathematica:

* Xuống dòng \leftarrow để nhập lệnh tiếp theo trong cùng khối lệnh.

* Biên dịch cả khối lệnh (Shift + \leftarrow) khi nhập xong mỗi lệnh, xem kết quả, soát và sửa lỗi.

* Lập biểu thức toán học bằng các phép toán + - * / (Ctrl + /), lũy thừa (Ctrl + 6), nhóm biểu thức trong dấu () và các hàm

Hàm lý thuyết	$\sin x$	$\cos x$	$\tan x$	$\cot x$	$\ln x$
Nhập trong Math	Sin[x]	Cos[x]	Tan[x]	Cot[x]	Log[x]
	$\arcsin x$	$\arccos x$	$\arctan x$	$\operatorname{arccot} x$	$\log_a x$
	ArcSin[x]	ArcCos[x]	AcrTan[x]	ArcCot[x]	Log[a,x]

với x là biến hoặc biểu thức.

* Mathematica phân biệt **CHỮ HOA** và **chữ thường**. Các lệnh (hàm) có sẵn của Mathematica có chữ cái đầu mỗi từ viết hoa (lệnh có thể là từ ghép), các tham số của lệnh được đặt trong dấu [] và ngăn cách bởi dấu ,

TênLệnh[tham số 1 , tham số 2 , ...]

* Vào [Menu]Help/Help Browser để tìm hiểu các lệnh.

* Thanh công cụ **BasicInput** hỗ trợ nhập nhanh các kí hiệu và biểu thức toán học theo mẫu, gọi ra bằng thao tác [Menu]File → Palettes

The image shows the BasicInput palette in Mathematica, which provides a visual interface for entering mathematical symbols. The symbols are organized into rows, with labels on the left and right pointing to specific symbols.

- lũy thừa** (power) points to the \square^\square symbol.
- phân số** (fraction) points to the $\frac{\square}{\square}$ symbol.
- căn bậc 2** (square root) points to the $\sqrt{\square}$ symbol.
- căn bậc n** (nth root) points to the $\sqrt[n]{\square}$ symbol.
- nguyên hàm** (integral) points to the $\int \square d\square$ symbol.
- đạo hàm riêng** (partial derivative) points to the $\partial_{\square} \square$ symbol.
- tích phân xác định** (definite integral) points to the $\int_{\square}^{\square} \square d\square$ symbol.
- đạo hàm riêng theo nhiều biến** (partial derivative with respect to multiple variables) points to the $\partial_{\square, \square} \square$ symbol.
- tổng của dãy** (sum of series) points to the $\sum_{\square=\square}^{\square} \square$ symbol.
- tích của dãy** (product of series) points to the $\prod_{\square=\square}^{\square} \square$ symbol.
- ma trận cấp 2** (2x2 matrix) points to the $\begin{pmatrix} \square & \square \\ \square & \square \end{pmatrix}$ symbol.
- phần tử của dãy, ma trận...** (element of a sequence, matrix...) points to the $\square[\square]$ symbol.
- π, e, i, ∞** points to the row containing π , e , i , ∞ , and $^{\circ}$.
- toán tử so sánh** (comparison operators) points to the row containing \times , \div , \ast , \rightarrow , and \Rightarrow .
- toán tử logic** (logical operators) points to the row containing $=$, \neq , \leq , \geq , and \in .
- toán tử logic** (logical operators) points to the row containing \neg , \wedge , \vee , \cup , and \cap .

Các lệnh trong môn học:

Lệnh	Kết quả
1 Binomial[n, k]	C_n^k
2 var = expr	gán biểu thức expr cho biến var
3 f[x_] := expr f[x_, y_] := expr	khai báo hàm số 1 hay nhiều biến
4 f[x] f'''[x]	$f(x)$ $f'''(x)$
5 Plot[f[x], {x,a,b}]	đồ thị hàm số $y = f(x)$ trên (a, b)
6 expr/N	đưa các số trong biểu thức về số thập phân
7 expr/.x->a expr/.{x->a, y->b}	giá trị của biểu thức khi thay x bởi a, \dots
8 Expand[expr] Simplify[expr]	khai triển, rút gọn biểu thức
9 If[biểu thức logic, kq1, kq2]	$\begin{cases} \text{kq1} & \text{nếu biểu thức logic đúng} \\ \text{kq2} & \text{ngược lại} \end{cases}$
10 Thao tác [RClick] → Create Table/Matrix/Palette	mẫu điền ma trận
11 Table[expr, {i,a,b}]	dãy, ma trận theo công thức
12 Table[expr, {i,a,b}, {j,c,d}]	

Trong một lệnh, ta thường dùng biến tạm % để gọi kết quả của lệnh vừa thực thi ngay trước đó.

VD1:

Kết quả	Lệnh
$C_{10}^4 = 210$	Binomial[10, 4]
$\sum_{k=0}^{10} C_{800}^k 0.005^k \cdot 0.995^{800-k} = 0.9972$	$\sum_{k=0}^{10} \text{Binomial}[800, k] * 0.005^k * 0.995^{800-k}$
$\sum_{k=0}^{\infty} k \frac{\lambda^k e^{-\lambda}}{k!} = \lambda$	$\sum_{k=0}^{\infty} k \frac{\lambda^k * E^{-\lambda}}{k!}$

VD2: Đại lượng ngẫu nhiên X có hàm mật độ $f(x) = ae^{4x-x^2}$

a) Tìm a .

b) Tìm EX, DX .

c) Tìm $P(1 < X < 3.5)$.

HD	Lệnh
<p>a) $\int_{-\infty}^{\infty} f(x) dx = 1 \Rightarrow ae^4\sqrt{\pi} = 1 \Rightarrow a = \frac{1}{e^4\sqrt{\pi}} = 0.01033$</p>	<p>$f[x_] := a * E^{4x-x^2}$ $\int_{-\infty}^{\infty} f[x] dx$ $a = \frac{1}{E^4\sqrt{\pi}}$ a / N</p>
<p>b) $EX = \int_{-\infty}^{\infty} xf(x) dx = 2$ $E(X^2) = \int_{-\infty}^{\infty} x^2 f(x) dx = \frac{9}{2}$ $DX = E(X^2) - (EX)^2 = \frac{1}{2}$</p>	<p>$ex = \int_{-\infty}^{\infty} x * f[x] dx$ $ex2 = \int_{-\infty}^{\infty} x^2 * f[x] dx$ $ex2 - ex^2$</p>
<p>c) $P(1 < X < 3.5) = \int_1^{3.5} f(x) dx = 0.9044$</p>	<p>$\int_1^{3.5} f[x] dx$</p>

VD3: Đại lượng ngẫu nhiên X có hàm mật độ $f(x) = \begin{cases} kx, & x \in [0, 2] \\ 0, & x \notin [0, 2] \end{cases}$

a) Xác định k .

b) Tìm hàm phân bố $F(x)$.

c) Tính $P(0 < X < 1)$.

HD	Lệnh
<p>a) $\int_{-\infty}^{\infty} f(x) dx = 1 \Rightarrow 2k = 1 \Rightarrow k = \frac{1}{2}$</p>	<p>$f[x_] := \text{If}[0 \leq x \leq 2, k * x, 0]$ $\int_{-\infty}^{\infty} f[x] dx$ $k = \frac{1}{2}$</p>
<p>b) $F(x) = \int_{-\infty}^x f(t) dt = \begin{cases} 0 & \text{nếu } x \leq 0 \\ \frac{x^2}{4} & \text{nếu } 0 < x \leq 2 \\ 1 & \text{nếu } x > 2. \end{cases}$ (trường hợp = 0 trong Math bị ẩn)</p>	<p>$g[t_] := \text{If}[t < x, f[t], 0]$ $\int_{-\infty}^{\infty} g[t] dt$</p>
<p>c) $P(0 < X < 1) = \int_0^1 f(x) dx = \frac{1}{4}$</p>	<p>$\int_0^1 f[x] dx$</p>

Chú ý: Mathematica có thể tính nhanh $I = \int_D f(x) dx$, khi $f(x)$ xác định phân nhánh, $D \subset \mathbb{R}^n$ là miền phức tạp hoặc phụ thuộc tham số, qua hai bước:

$$g(x) = \begin{cases} f(x), & \text{nếu } x \in D \\ 0 & \text{nếu } x \notin D \end{cases} \Rightarrow I = \int_{\mathbb{R}^n} g(x) dx.$$

VD4: Vectơ ngẫu nhiên (X, Y) có hàm mật độ xác suất đồng thời

$$f(x, y) = \begin{cases} A(2x^2 + xy + y^2) & \text{nếu } (x, y) \in [0, 1] \times [0, 1] \\ 0 & \text{nếu } (x, y) \notin [0, 1] \times [0, 1] \end{cases}. \text{ Tìm}$$

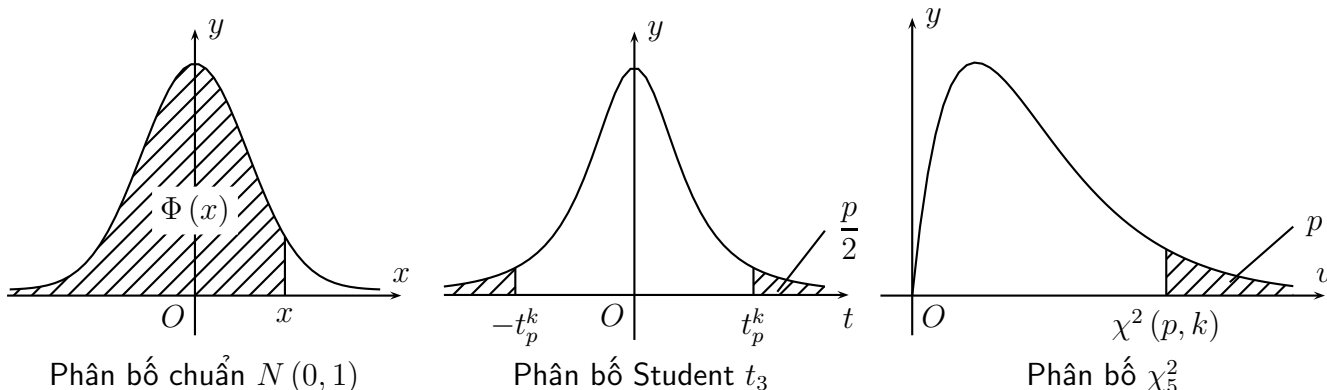
a) Tìm A .

b) Tính EX .

c) Tính $P(X < 0.5 | Y > 0.5)$.

HD	Lệnh
	<code>f[x_, y_] := If[0 ≤ x ≤ 1 & 0 ≤ y ≤ 1, A(2x² + x * y + y²), 0]</code>
a) $\boxed{\iint_{\mathbb{R}^2} f(x, y) dx dy = 1 \Rightarrow \frac{5A}{4} = 1 \Rightarrow A = \frac{4}{5}}$	<code>\int_{-\infty}^{\infty} (\int_{-\infty}^{\infty} f[x, y] dy) dx</code> <code>A = $\frac{4}{5}$</code>
b) $EX = \iint_{\mathbb{R}^2} x f(x, y) dx dy = \frac{2}{3}$.	<code>\int_{-\infty}^{\infty} (\int_{-\infty}^{\infty} x * f[x, y] dy) dx</code>
c) $P(X < 0.5 Y > 0.5) = \frac{P(X < 0.5, Y > 0.5)}{P(Y > 0.5)} = \frac{t}{m}$.	
$\boxed{t = \iint_{x < 0.5, y > 0.5} f(x, y) dx dy = 0.1875}$	<code>t = \int_{-\infty}^{0.5} (\int_{0.5}^{\infty} f[x, y] dy) dx</code>
$\boxed{m = \iint_{y > 0.5} f(x, y) dx dy = 0.65}$	<code>m = \int_{-\infty}^{\infty} (\int_{0.5}^{\infty} f[x, y] dy) dx</code>
$P(X < 0.5 Y > 0.5) = 0.2885$.	<code>$\frac{t}{m}$</code>

2 Excel



Ngoài cách gõ hàm trực tiếp, có thể nhập hàm bằng thao tác trên giao diện, bằng một trong ba cách:

- Phím tắt Shift + F3, hoặc
- Nhập vào nút f_x ngay phía dưới thanh Menu, hoặc
- [Menu]Formulas → Insert Function.

Trên giao diện, nên chọn chủ đề Statistical để lọc các hàm sử dụng trong môn học:

Giá trị	Tham số	Lệnh	Ví dụ & chú thích
13 $\Phi(x)$	$x \in \mathbb{R}$	=norm.dist(x,0,1,true)	$\Phi(2) = 0.9772$
14 z_0 s/c: $\Phi(z_0) = p$	$0 < p < 1$	=norm.inv(p,0,1)	$\Phi(z_0) = 0.95$ $\Rightarrow z_0 = 1.6449$
15 t_p^k	$k \in \mathbb{N}^*$	=t.inv.2t(p,k)	$t_{0.1}^{29} = 1.6991$
16 $\chi^2(p, k)$		=chisq.inv.rt(p,k)	$\chi^2(0.05, 25) = 37.6525$
17 \bar{x}	x_1, x_2, \dots, x_n	=average(X)	Nhập x_1, x_2, \dots, x_n liên tiếp vào 1 cột (n ô đầu của cột A) \Rightarrow X là A1:An
18 s'^2		=var.s(X)	
19 s'		=sqrt(ô chứa s'^2)	
20 s^2		=var.p(X)	
21 r	x_1, x_2, \dots, x_n	=correl(X, Y)	X là A1:An, Y là B1:Bn. Đường hồi quy tuyến tính thực nghiệm của Y theo X là $y = ax + b$.
22 a	y_1, y_2, \dots, y_n	=slope(Y, X)	
23 b		=intercept(Y, X)	

Lệnh Excel không phân biệt chữ hoa hay chữ thường, khi dấu chấm thập phân là dấu ".", có dạng

= tên_lệnh(tham số 1 , tham số 2 , ...)

và khi dấu chấm thập phân là dấu ",", là

= tên_lệnh(tham số 1 ; tham số 2 ; ...)

VD5: Mẫu cỡ 50 từ $X \in N(a, \sigma^2)$ với $\sigma = 2$ cho số liệu theo bảng sau:

x_i	10 – 12	12 – 14	14 – 16	16 – 18
n_i	9	18	17	6

- Tìm ƯLKC của a .
- Tìm KTC của a với độ tin cậy 97%.
- Kiểm định ở mức ý nghĩa 7% xem $EX = 13$ hay $EX > 13$.
- Kiểm định ở mức ý nghĩa 6% xem có phải $EX = 13$ hay không.

HD	Lệnh
a) $X \in N(a, \sigma^2) \Rightarrow a = EX$ có ƯLKC $\boxed{\bar{x} = 13.8}$	Nhập x_i vào cột A:
b) $\sigma = 2 \Rightarrow$ KTC của a : $(\bar{x} - z_0 \frac{\sigma}{\sqrt{n}}, \bar{x} + z_0 \frac{\sigma}{\sqrt{n}})$.	11 vào ô A1 \rightarrow đưa con trỏ về góc dưới bên phải ô A1 để thấy dấu + \rightarrow kéo tới dòng 9 (copy công thức)
$\Phi(z_0) = \frac{1+\gamma}{2} = \frac{1+0.97}{2} = 0.985 \Rightarrow \boxed{z_0 = 2.1701}$	13 vào A10, copy tới dòng 9 + 18 = 27
KTC: (13.1862, 14.4138).	15 vào A28 tới dòng 27 + 17 = 44
c) $z_{qs} = \frac{\bar{x} - a_0}{\sigma} \sqrt{n} = 2.8284$.	17 tới dòng 50
$\Phi(z_0) = 1 - \alpha = 0.93 \Rightarrow \boxed{z_0 = 1.4758}$	=average(A1:A50)
$z_{qs} > z_0 \Rightarrow$ bác bỏ $EX = 13$ (chấp nhận $EX > 13$).	=norm.inv(0.985, 0, 1)
d) $H : EX = 13, K : EX \neq 13, \alpha = 0.06$.	=norm.inv(0.93, 0, 1)
$z_{qs} = 2.8284$ (như ý (c)).	=norm.inv(0.97, 0, 1)
$\Phi(z_0) = 1 - \frac{\alpha}{2} = 0.97 \Rightarrow \boxed{z_0 = 1.8808}$	
$ z_{qs} > z_0 \Rightarrow$ bác bỏ H (chấp nhận $EX \neq 13$).	

VD6: Mẫu cỡ $n = 31$ từ $X \in N(a, \sigma^2)$ cho số liệu theo bảng sau

x_i	58 – 60	60 – 62	62 – 64	64 – 66
n_i	3	12	13	3

- Tìm ƯLKC của a và σ^2 .
- Tìm KTC của a với độ tin cậy 92%.
- Kiểm định ở mức ý nghĩa 4% xem $EX = 64$ hay $EX < 64$.

HD	Lệnh
a) $X \in N(a, \sigma^2) \Rightarrow EX = a, DX = \sigma^2$. ƯLKC của a là $\boxed{\bar{x} = 62.0322}$, của σ^2 là $\boxed{s'^2 = 2.6323}$	Nhập x_i vào cột A: 59 (vào 3 ô), 61 (12 ô), 63 (13 ô), 65 (3 ô)

b) KTC của a : $(\bar{x} - t_0 \frac{s'}{\sqrt{n}}, \bar{x} + t_0 \frac{s'}{\sqrt{n}})$.

$$s' = \sqrt{s'^2} = 1.6224$$

$$t_0 = t_{1-\gamma}^{n-1} = t_{1-0.92}^{31-1} = 1.812$$

KTC: (61.5042, 62.5603).

c) $t_{qs} = \frac{\bar{x} - a_0}{s'} \sqrt{n} = -6.7528$.

$$t_0 = t_{2\alpha}^{n-1} = t_{0.08}^{30} = 1.812.$$

$t_{qs} < -t_0 \Rightarrow$ bác bỏ $EX = 64$ (chấp nhận $EX < 64$).

=average(A1:A31)
=var.s(A1:A31)
=sqrt(ô chứa s'^2)
=t.inv.2t(0.08, 30)

VD7: Phương pháp thứ nhất cho tỷ lệ sản phẩm tốt là 85%. Kiểm tra 300 sản phẩm sản xuất theo phương pháp thứ hai thì thấy có 30 phế phẩm. Hãy kiểm định ở mức ý nghĩa 8% xem có phải phương pháp thứ hai tốt hơn phương pháp thứ nhất không.

HD: p = tỷ lệ sản phẩm tốt sản xuất theo phương pháp thứ hai. Bài toán

$$H : p = 0.85, K : p > 0.85, \alpha = 0.08.$$

m = số sản phẩm tốt sản xuất theo phương pháp thứ hai = 270.

$$z_{qs} = \frac{\frac{m}{n} - p_0}{\sqrt{p_0(1-p_0)}} \sqrt{n} = \frac{\frac{270}{300} - 0.85}{\sqrt{0.85 \cdot 0.15}} \sqrt{300} = 2.4254.$$

$$\Phi(z_0) = 1 - \alpha = 0.92 \Rightarrow z_0 = 1.4051$$

$z_{qs} > z_0 \Rightarrow$ bác bỏ H tức là phương pháp thứ hai tốt hơn.

Lệnh Excel: =(270/300-0.85)/sqrt(0.85*0.15)*sqrt(300) và =norm.inv(0.92, 0, 1)

VD8: Với mức ý nghĩa 0.06 hãy kiểm định $H : EX = EY$ với đối thuyết $K : EX > EY$ trong đó X, Y là 2 đại lượng ngẫu nhiên có phân bố chuẩn. Biết rằng 2 mẫu độc lập cỡ $n = 17$ và từ X và $m = 13$ từ Y cho ta số liệu sau:

x_i	21.7	21.9	22.1	22.2	22.5	y_i	21.8	22	22.2
n_i	1	4	6	5	1	m_i	5	7	1

Cho biết $DX = 0.03, DY = 0.02$.

$$HD: \bar{x} = 22.0824, \bar{y} = 21.9385 \Rightarrow z_{qs} = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{DX}{n} + \frac{DY}{m}}} = 2.5036.$$

$$\Phi(z_0) = 1 - \alpha = 0.94 \Rightarrow z_0 = 1.5548$$

$z_{qs} > z_0 \Rightarrow$ bác bỏ H .

Excel: Nhập x_i vào cột A, y_i vào cột B. Địa chỉ mẫu của X là A1:A17 và Y là B1:B13

VD9: Khảo sát thu nhập (triệu đồng) trong 1 tháng của 10 công nhân ngành A và 15 công nhân ngành B ta thu được số liệu sau:

Ngành A	Ngành B
1, 1.2, 1.3, 1.3, 1.2, 1.3, 1.4, 1.2, 1.3, 1.4	1.1, 1.3, 1.3, 1.3, 1.4, 1.4, 1.4, 1.3, 1.4, 1.2, 1.5, 1.5, 1.5, 1.2, 1.2

Giả sử thu nhập trong 1 tháng của 1 công nhân ngành A và B là những đại lượng ngẫu nhiên X, Y có phân bố chuẩn với phương sai bằng nhau. Hãy kiểm định ở mức ý nghĩa 3% giả thuyết nói rằng thu nhập trung bình của công nhân 2 ngành trên như nhau với đối thuyết cho rằng thu nhập trung bình của công nhân ngành B cao hơn ngành A.

HD: $H : EX = EY, K : EX < EY, \alpha = 0.03$.

$$\bar{x} = 1.26, \bar{y} = 1.3333, s_X^2 = 0.0124, s_Y^2 = 0.01422$$

$$t_{qs} = \frac{\bar{x} - \bar{y}}{\sqrt{ns_X^2 + ms_Y^2}} \sqrt{\frac{nm(n+m-2)}{n+m}} = -1.4832.$$

$$t_0 = t_{2\alpha}^{n+m-2} = t_{0.06}^{23} = 1.9782$$

$t_{qs} > -t_0 \Rightarrow$ chấp nhận H , tức là thu nhập trung bình của công nhân 2 ngành trên như nhau.

VD10: Mẫu cỡ $n = 100$ từ đại lượng ngẫu nhiên X cho ta số liệu sau:

x_i	1 – 2	2 – 3	3 – 4	4 – 5	5 – 6	6 – 7	7 – 8
n_i	4	12	27	30	20	5	2

Hãy kiểm định ở mức ý nghĩa 7% xem có phải X có phân bố chuẩn $N(4, 1.3^2)$.

HD: $H : X \in N(4, 1.3^2), K : X \notin N(4, 1.3^2), \alpha = 0.07$

S_i	n_i	p_{i0}	E_i	$\frac{(n_i - E_i)^2}{E_i}$
$(-\infty, 2]$	4	0.06197	6.1968	0.7788
$(2, 3]$	12	0.1589	15.891	0.9527
$(3, 4]$	27	0.2791	27.9122	0.02981
$(4, 5]$	30	0.2791	27.9122	0.1561
$(5, 6]$	20	0.1589	15.891	1.0625
$(6, 7]$	5	0.05146	5.146	0.004141
$(7, \infty)$	2	0.01051	1.0508	0.8574
χ_{qs}^2				3.8415

$\chi_0^2 = \chi^2(0.07, 7 - 1) = 11.6599. \chi_{qs}^2 < \chi_0^2 \Rightarrow$ chấp nhận H .

Mathematica:

$f[x_] := \frac{1}{1.3\sqrt{2\pi}} E^{-\frac{(x-4)^2}{2*1.3^2}}$	hàm mật độ của $X \in N(4, 1.3^2)$
$S = \{-\infty, 2, 3, 4, 5, 6, 7, \infty\}$	cận trên của khoảng này là cận dưới của khoảng sau
$n = \{4, 12, 27, 30, 20, 5, 2\}$	
$p = \text{Table}[\int_{S[[i]]}^{S[[i+1]]} f[x] dx, \{i, 1, 7\}]$	
$e = 100 * p$	
$t = \frac{(n - e)^2}{e}$	tính theo vị trí tương ứng
$\sum_{i=1}^7 t[[i]]$	

Excel:

$p_{10} = P(X \leq 2 H)$	=norm.dist(2,4,1.3,true)
$p_{20} = P(2 < X \leq 3 H)$	=norm.dist(3,4,1.3,true) - norm.dist(2,4,1.3,true)
$E_1 = np_{10}$	=100*ô chứa p_{10}
$\frac{(n_1 - E_1)^2}{E_1}$	=(ô chứa n_1 - ô chứa E_1)^2/ô chứa E_1
$E_i, \frac{(n_i - E_i)^2}{E_i}, i = \overline{2, 7}$	copy công thức từ ô chứa E_1 và $\frac{(n_1 - E_1)^2}{E_1}$
χ_{qs}^2	=sum(địa chỉ của dãy $\frac{(n_i - E_i)^2}{E_i}$)

$$\chi_0^2 \quad \Bigg| \quad = \text{chisq.inv.rt}(0.07, 6)$$

VD11: Mẫu cỡ $n = 60$ từ đại lượng ngẫu nhiên X cho ta số liệu dưới đây:

x_i	10 – 11	11 – 12	12 – 13	13 – 14	14 – 15	15 – 16	16 – 18
n_i	9	6	7	8	6	7	17

Kiểm định ở mức ý nghĩa 8% xem X có phân bố đều không?

HD: $*H : X \in U[a, b], K : X \in U[a, b], \alpha = 0.08.$

$\bar{x} = 14.0583, s^2 = 5.5924, s = 2.3648.$ Ước lượng theo phương pháp bình phương tối thiểu của a là $a^* = \bar{x} - s\sqrt{3} = 9.9623$, của b là $b^* = \bar{x} + s\sqrt{3} = 18.1543$ (số tham số chưa biết $r = 2$).

$*H' : X \in U[9.9623, 18.1543], K' : X \notin U[9.9623, 18.1543], \alpha = 0.08.$

S_i	n_i	p_{i0}	E_i	$\frac{(n_i - E_i)^2}{E_i}$
$(-\infty, 11]$	9	0.1267	7.6001	0.2578
$(11, 12]$	6	0.1221	7.3242	0.2394
$(12, 13]$	7	0.1221	7.3242	0.01435
$(13, 14]$	8	0.1221	7.3242	0.06236
$(14, 15]$	6	0.1221	7.3242	0.2394
$(15, 16]$	7	0.1221	7.3242	0.01435
$(16, \infty)$	17	0.263	15.7788	0.09451
			χ_{qs}^2	0.9222

$$\chi_0^2 = \chi^2(\alpha, h - r - 1) = \chi^2(0.08, 7 - 2 - 1) = 8.3365.$$

$$\chi_{qs}^2 < \chi_0^2 \Rightarrow \text{chấp nhận } H \text{ (} X \text{ có phân bố đều)}.$$

Excel: tính a^*, b^* .

Mathematica: tính p_{i0} với khai báo

$$f[x_] := \text{If}[9.9623 \leq x \leq 18.1543, \frac{1}{18.1543 - 9.9623}, 0]$$

rồi thực hiện tiếp như ví dụ 10.

VD12: Mẫu cỡ 200 từ VTNN (X, Y) cho ta số liệu sau:

$X \backslash Y$	T	N	M
C	26	48	24
K	51	43	8

Kiểm định ở mức ý nghĩa 5% xem X và Y độc lập nhau không?

HD: (Bảng sau vừa tính toán và trình bày, vừa mô tả một trang Excel)

	A	B	C	D	E	Excel
1	Bảng n_{ij}, n_{i*}, n_{*j} :					
2	$X \backslash Y$	T	N	M	Σ	
3	C	26	48	24	98	=sum(B3:D3)
4	K	51	43	8	102	
5	Σ	77	91	32	200	=sum(B3:B4)
6	Bảng E_{ij} :					
7		37.73	44.59	15.68		=\$E3*B\$5/\$E\$5
8		39.27	46.41	16.32		
9	Bảng $\frac{(n_{ij} - E_{ij})^2}{E_{ij}}$:					
10		3.6468	0.2608	4.4147		=(B3-B7)^2/B7
11		3.5038	0.2506	4.2416		
	$\chi^2_{qs} =$	16.3181				=sum(B10:D11)
	$\chi^2_0 = \chi^2[\alpha, (h-1)(k-1)] = \chi^2[0.05, (2-1)(3-1)] =$	5.9915				=chisq.inv.rt(0.05,2)
	$\chi^2_{qs} > \chi^2_0 \Rightarrow X, Y$ không độc lập.					

Mathematica:

$$n = \begin{pmatrix} 26 & 48 & 24 \\ 51 & 43 & 8 \end{pmatrix}$$

$$nx = \text{Table}[\sum_{j=1}^3 n_{[i,j]}, \{i, 1, 2\}]$$

$$ny = \text{Table}[\sum_{i=1}^2 n_{[i,j]}, \{j, 1, 3\}]$$

$$e = \text{Table}[\frac{nx_{[i]} * ny_{[j]}}{200.0}, \{i, 1, 2\}, \{j, 1, 3\}]$$

$$t = \frac{(n - e)^2}{e}$$

$$\sum_{i=1}^2 \sum_{j=1}^3 t_{[i,j]}$$

trang 3, bảng lệnh Mathematica, dòng 10

tổng từng hàng

VD13: Mẫu cỡ $n = 12$ từ véctơ ngẫu nhiên (X, Y) cho ta số liệu sau:

x_i	3	3.5	2.5	3.5	3	3.1	2	4	4.5	3	3.5	3.1
y_i	13	14	10	14.5	14	13.5	9.5	16	18	12.5	15	14.5

- Tìm hệ số tương quan mẫu giữa X và Y . Có thể dùng hồi quy bình phương trung bình tuyến tính của Y đối với X để dự báo giá trị của Y được không, vì sao?
- Tìm hàm hồi quy bình phương trung bình tuyến tính thực nghiệm của Y đối với X và ước lượng sai số bình phương trung bình.
- Dự báo giá trị của Y khi biết $X = 2.3$.

HD	Lệnh
a) $r = 0.957$. $ r $ khá lớn (≥ 0.8) \Rightarrow có thể dùng hồi quy bình phương trung bình tuyến tính của Y theo X để dự báo giá trị của Y .	Nhập (3, 13) vào hai ô A1 và B1, (3.5, 14) vào A2, B2... (3.1, 14.5) vào A12, B12 =correl(A1:A12,B1:B12)
b) Hàm hồi quy $Y = aX + b$: $a = 3.4397$, $b = 2.6154$	=slope(B1:B12,A1:A12) =intercept(B1:B12,A1:A12)
$s_Y^2 = 5.0191 \Rightarrow$ sai số $s_{Y/X}^2 = s_Y^2 (1 - r^2) = 0.4221$.	=var.p(B1:B12)
c) Khi $X = 2.3$ ta dự báo $Y = aX + b = 10.5266$.	

Thông kê thời gian thực hành:

VD	1	2	3	4	5	6	7	8	9	10	11	12	13	Σ
Time	1'2"	1'21"	1'1"	1'37"	3'9"	2'33"	47"	2'5"	2'24"	3'9"	4'26"	1'55"	2'45"	31'56"
										1'49"		1'53"		