ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH

TRƯỜNG ĐẠI HỌC BÁCH KHOA

BÀI TẬP LỚN MÔN HỌC

XÁC SUẤT THỐNG KÊ

Khóa 2020

Đề 21118

NHÓM: 18 - HK211

GVHD: NGUYỄN NGỌC PHÚC DIỄM

SINH VIÊN THỰC HIỆN

STT	MSSV	HỌ	TÊN	% ĐIỂM BTL	ĐIỂM BTL	GHI CHÚ
1	2012846	Trần Ngọc	Duy	34%		Nhóm trưởng
2	2015055	Đinh Quang	Vinh	33%		
3	2015110	Trần Hoàng	Vương	33%		

TP. HỒ CHÍ MINH, NĂM 2021

MỤC LỤC

LỜI MỞ ĐẦU

Giá trị của xác suất thống kê:

Ta thường cho rằng những con số là vô tri, vô giác, là khô khan hay đơn thuần là chán nản với chúng. Nhưng thực tế rằng, chúng có giá trị vô cùng mạch mẽ, chúng là " những con số biết nói", được thể hiện giá trị qua Xác suất Thống kê. Nơi chúng ứng dụng thực tế to lớn trong cuộc sống hàng ngày. Cũng có thể vì lý do đó mà môn học Xác suất Thống kê được dạy cho hầu hết các ngành trong trường đại học. Và cũng như trong thời đại công nghệ thông tin, với số lượng dữ liệu khổng lồ chưa từng có, kiến thức xác suất thống kê càng phát huy và lên tiếng cho chúng ta thấy những điều ngay cạnh từ cuộc sống thường ngày cho đến công việc.

Xác suất là độ đo của toán học để đo tính phi chắc chắn của khả năng xảy ra một sự kiện (biến cố).

Thống kê là một phần toán học của khoa học, gắn liền với tập hợp dữ liệu, phân tích, giải thích hoặc thảo luận về một vấn đề nào đó, và trình bày dữ liệu, hay là một nhánh của toán học.

Định nghĩa thống kê về xác suất có ưu điểm lớn là không đòi hỏi những điều kiện áp dụng như đối với những định nghĩa cổ điển. Nó hoàn toàn dựa trên các quan sát thực tế để làm cơ sở kết luận về xác suất xảy ra của một biến cố.

Dựa vào đó, có thể hiểu thống kê toán học là một phương pháp khoa học phân tích và xử lý dữ liệu có được nhờ các thí nghiệm, các cuộc điều tra nghiên cứu các hiện tượng tự nhiên, các vấn đề kỹ thuật cũng như các vấn đề xã hội. Những dữ liệu ở đây có thể là những đặc tính định tính, cũng có thể là những đặc tính định lượng. Theo đó, từ những dữ liệu thu thập được, dựa vào các quy luật xác suất để đưa ra những quyết định, những đánh giá và các dự báo về những hiện tượng đang được thí nghiệm hoặc đang được quan sát là mục đích của thống kê toán học.

Và để trực tiếp cho ta thấy về giá trị, sức mạnh lẫn những điều mà từng con số "nói" cho chúng ta biết, chúng ta sẽ đi đến với 2 bài toán điển hình cho xác suất và thống kê. Ở đây, ta có bài toán liên quan trực tiếp đến công việc của những kĩ sư điện tương lai và cũng là những người thực hiện đề tài này. Vậy giờ cùng đến với những bài xác suất thống kê tính khoảng ước lượng, các giá trị, thông số và từ đó có những cái nhìn tổng quát nhất cho xác suất thống kê ở khối nghành kỹ thuật.

I. BÀI 1: Xác định khoảng phóng điện đánh thủng của mẫu điện môi:

1.1. Giới thiệu nội dung, ý nghĩa các thông số bài toán, giá trị bài toán.

1.1.1. *Giới thiệu nội dung:*

Ở bài toán thứ nhất, ta thực hiện việc xác định khoảng phóng điện chọc thủng của mẫu điện môi qua giá trị 15 lần đánh thủng và độ tin cậy cho trước. Tức ta đang tìm khoảng ước lượng cho 15 mẫu thử và trên thực tế là tìm ngưỡng an toàn cho giấy cách điện hay khoảng điện áp phù hợp để tránh bị đánh thủng.

1.1.2. *Ý nghĩa các thông số bài toán*

Hiện tượng phóng điện trong điện môi và điện áp đánh thủng:

Khi đặt U lên 2 đầu điện môi, vượt quá một giới hạn nào đó sẽ xảy ra phóng điện chọc thủng điện môi, khi đó bị mất hoàn toàn tính chất cách điện, Hiện tượng đó chính là sự phóng điện chọc thủng của điện môi hay là sự phá huỷ độ bền điện môi.

Phóng điện chọc thủng còn gọi là đánh thủng điện môi hay phóng điện xuyên qua. Trị số điện áp mà ở đó xảy ra đánh thủng điện môi được gọi là điện áp đánh thủng (Uđt) trị số tương ứng của cường độ điện trường là cường độ đánh thủng hay cường độ điện trường cách điện (Eđt).

Cường độ điện trường cách điện "E" = Eđt chính là điện áp đánh thủng điện môi trên 1 mm chiều dày. Khi tính toán để chọn chiều dày điện môi của một thiết bị làm việc ở điện áp định mức nào đó (Uđm), cần tính đến hệ số an

Thực tế có rất nhiều yếu tố ảnh hưởng tới E cách điện của điện môi: dạng điện trường, dạng điện áp, thời gian tác dụng của điện áp, điều kiện môi trường như áp suất, nhiệt độ, độ ẩm,..

Vậy trong cùng điều kiện môi trường và loại giấy cách điện, ta cần xác định khoảng điện áp mà giấy bị đánh thủng, tức tầm trị mà có nguy cơ bị đánh thủng với độ tin cậy 99%...

1.1.3. *Giá trị bài toán*:

Về mặc học thuật, bài toán giúp ta định hình về các phương pháp ước lượng khoảng, các cách thức tính phương sai và kiến thức về thống kê mẫu.

Về mặt ý nghĩa thực tế, đây là là toán tiêu biểu cho việc ước lượng khoảng an toàn cho các dụng cụ kỹ thuật điện. Trong thực tế, đây là phương pháp thử trước khi đưa dụng cụ vào trong mạch điện sử dụng . Ngoài giấy cách điện, ta còn sử dụng hình thức thử này cho các dụng cụ vật liệu khác để tính được tầm hư hại, khoảng an toàn cho mạch.

Bên cạnh đó, việc tính giá trị khoảng đánh thủng(khoảng an toàn hay ngưỡng định mức) giúp ta chọn mạch phù hợp để sử dụng linh kiện, bởi thực tế ta không thể chọn loại dụng cụ có khoảng đánh thủng quá cao hoặc quá thấp bởi các lí do chi phí, mạng điện, trở kháng,...

1.2. Trình bày hướng làm, cách giải.

1.2.1. Lý thuyết xác suất liên quan:

Bảng 1- Tóm tắt một số hàm ước lượng tham số thông dụng:

Tham số θ cần ước lượng	Chọn thống kê $\hat{\theta}$ để ước lượng	E[$\hat{\theta}$]	D[$\hat{\theta}$]	Tính chất của ước lượng
Tỉ lệ p (xác suất)	$F = \frac{m}{n}$	$E(F) = p$	$D(F) = \frac{p(1-p)}{n}$	Không chệch, vững, hiệu quả, đủ; hợp lý cực đại.
Kỳ vọng $a = E(X)$	$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$	$E(\overline{X}) = a$	$D(\overline{X}) = \frac{\sigma^2}{n}$	Không chệch, vững, hiệu quả, đủ; hợp lý cực đại.
Phương sai $\sigma^2 = D(X)$	$\hat{S}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$	$E(\hat{S}^2) = \frac{n-1}{n}\sigma^2$	…	Chệch, vững, đủ; hợp lý cực đại.
	$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$	$E(S^2) = \sigma^2$	…	Không chệch, vững, đủ.

Sơ lược về hàm Student:

Ứng dụng:

Phân phối T – Student thường được dùng rộng rãi trong việc suy luận phương sai tổng thể khi có giả thiết tổng thể phân phối chuẩn, đặc biệt khi cỡ mẫu càng nhỏ thì độ chính xác càng cao. Ngoài ra, còn được ứng dụng trong kiểm định giả tiết về trung bình khi chưa biết phương sai tổng thể là bao nhiêu.

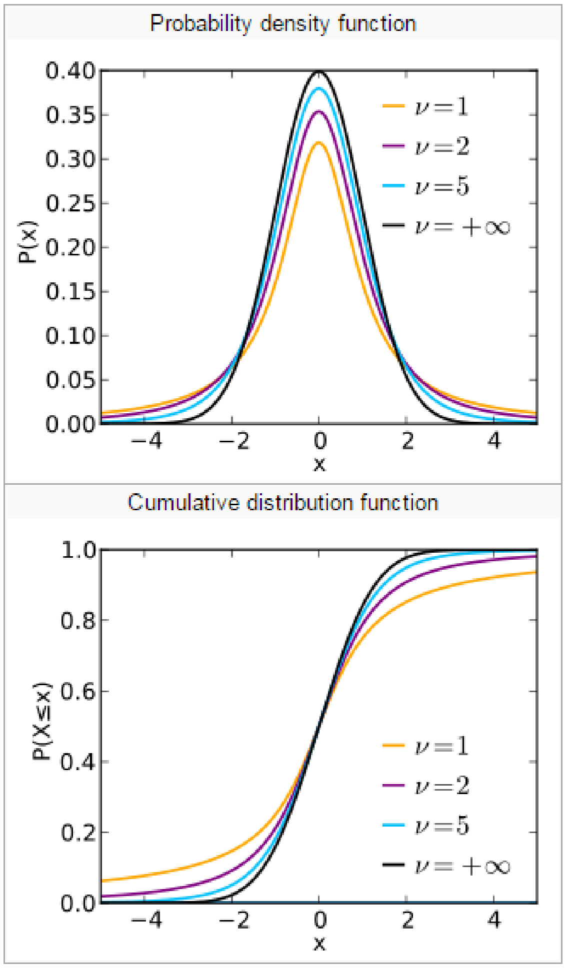
Phân phối này được ứng dụng trong cả xác suất thống kê và kinh tế lượng.

Đặc điểm:

Hình dạng đối xứng gần giống phân phối chuẩn hóa

Khi cỡ mẫu càng lớn càng giống phân phối chuẩn hóa

Cỡ mẫu càng nhỏ, phần đuôi càng nặng và xa hơn



-**Ước lượng bằng khoảng tin cậy** chính là tìm ra khoảng ước lượng $(G_1;G_2)$ cho tham số θ trong tổng thể sao cho ứng với độ tin cậy *(confidence)* bằng $(1-\alpha)$ cho trước, $P(G_1 < \theta < G_2) = 1-\alpha$.

Phương pháp ƯL bằng khoảng tin cậy có ưu thế hơn phương pháp ƯL điểm vì nó làm tăng độ chính xác của ước lượng và còn đánh giá được mức độ tin cậy của ước lượng. Nó chứa đựng khả năng mắc sai lầm là α .

Phương pháp tìm khoảng tin cậy cho tham số θ với độ tin cậy *$1-\alpha$* cho trước:

- Trước tiên ta tìm hàm ước lượng G = f(X1 , X2 , .., Xn , θ) sao cho quy luật phân phối xác suất của G hoàn toàn xác định, không phụ thuộc vào các đối số. Chọn cặp giá trị $\alpha 1$, $\alpha 2 \geq 0$ sao cho $\alpha 1 + \alpha 2 = \alpha$ và tìm $G\alpha 1$, $G\alpha 2$ mà ($G < G\alpha 1$) $=\alpha 1$ và P ($G > G\alpha 2$) = $\alpha 2$, suy ra P($G\alpha 1 < G < G\alpha 2$) = $1 - \alpha$. Biến đổi để tìm được các giá trị G1, G2 sao cho P(G1 < θ < G2) = $1-\alpha$. Khi đó khoảng (G1, G2) chính là một trong các khoảng tin cậy *(confidence interval)* cần tìm.
- Theo nguyên lý xác suất lớn thì với độ tin cậy ($1 -\alpha$) đủ lớn, hầu như chắc chắn biến cố (G1 < θ < G2) sẽ xảy ra trong một phép thử. Vì vậy trong thực tế chỉ cần thực hiện phép thử để có được một mẫu cụ thể w = (x1, x2 , .., xn) rồi tính giá trị của G1 và G2 ứng với mẫu đã cho sẽ cho ta một khoảng ước lượng thỏa yêu cầu.

Lưu ý:

* Đối với mẫu đã xác định, khoảng tin cậy đối xứng có độ dài càng hẹp thì độ tin cậy càng thấp. Nếu chúng ta muốn có được sai số nhỏ (khoảng tin cậy hẹp) và độ tin cậy như mong muốn thì chúng ta phải tăng kích thước mẫu hợp lý.

* Có vô số khoảng ước lượng cho giá trị p của tổng thể tùy theo cách chọn α_1, α_2 sao cho $\alpha_1 + \alpha_2 = \alpha$. Đối với bài toán ƯL tỉ lệ hay ƯL trung bình thì khoảng ƯL được trình bày ở trong bài chính là khoảng ƯL đối xứng và nó có độ dài ngắn nhất.

Bài toán minh họa : *Giả sử tổng thể X có phân phối chuẩn, chưa biết trung bình tổng thể a và phương sai tổng thể σ^2. Từ tổng thể, người ta lấy được mẫu tổng quát với kích thước n, trung bình mẫu $\overline{X}$ và phương sai mẫu hiệu chỉnh S^2.*

Tìm khoảng tin cậy cho trung bình tổng thể a với độ tin cậy $1-\alpha$; trong trường hợp mẫu có kích thước nhỏ.

khi n <30 thì: $Q = \frac{\overline{X} - a}{s}\sqrt{n} \sim T(n-1)$

Chọn khoảng ước lượng đối xứng có dạng $(\overline{X} - \varepsilon; \overline{X} + \varepsilon)$

Dẫn đến bài toán tìm ε để $P(\overline{X} - \varepsilon < a < \overline{X} + \varepsilon) = 1 - \alpha$

$\Rightarrow P(-\frac{\varepsilon}{s}\sqrt{n} < Q = \frac{\overline{X} - a}{s}\sqrt{n} < \frac{\varepsilon}{s}\sqrt{n}) = 1 - \alpha$. Đặt : $T_\alpha = \frac{\varepsilon}{s}\sqrt{n}$

$\Rightarrow$ Dựa vào bảng tra 1 phía trong Phụ lục VII cho hàm Student, ta tìm được giá trị $T_\alpha = t_{\alpha/2}^{(n-1)}$ bằng cách tìm số nằm ở cột $\alpha/2$, dòng thứ (n-1). Từ đó suy ra ε cần tìm.

(Nhắc lại: Khi $n \geq 30$, phân phối Student xấp xỉ phân phối Chuẩn tắc.)

36

Tham số cần ước lượng	Phân bố của tổng thể	Thông tin bổ sung	Khoảng tin cậy khi chọn $\alpha_1 = \alpha_2 = \alpha/2$
Trung bình a (μ)	Bất kỳ	Mẫu lớn ($n \geq 30$)	$(\overline{X} \pm \varepsilon)$; $\varepsilon = z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$
		Nếu chưa biết σ^2 thì dùng s^2 thay thế	
	Chuẩn $N(a,\sigma^2)$	σ^2 đã biết	$(\overline{X} \pm \varepsilon)$; $\varepsilon = z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$
	Chuẩn $N(a,\sigma^2)$	σ^2 chưa biết	$(\overline{X} \pm \varepsilon)$ $\varepsilon = t_{\frac{\alpha}{2}}(n-1) \cdot \frac{s}{\sqrt{n}}$
		Nếu mẫu lớn thì có thể dùng $z_{\frac{\alpha}{2}}$ thay thế cho $t_{\frac{\alpha}{2}}(n-1)$	

1.2.2. Giải toán: (đề 18)

16	21116	Upd(kV)	2,888	2,774	3,002	3,306	2,584	3,306	2,622	2,622	2,85	2,926	2,66	2,85	2,812	2,698	2,85	95%
17	21117	Upd(kV)	2,812	2,698	2,622	2,584	2,584	3,192	2,774	2,774	2,66	2,888	3,002	3,04	2,058	2,242	2,812	95%
18	21118	Upd(kV)	2,774	2,66	2,85	3,306	2,66	2,926	2,812	3,154	2,622	2,622	2,812	2,926	2,85	3,002	2,774	99%
19	21119	Upd(kV)	2,812	2,66	2,622	2,812	2,736	3,002	2,812	2,774	2,698	2,698	2,774	2,546	2,508	2,698	2,66	98%

Vậy quanh trở lại bài toán ban đầu, ta có độ tin cậy 0.99, số lượng 15 mẫu thử(<30) nên ta sử dụng giá trị student với T(0.005 ; 14)= 2.947

Giá trị trung bình là: $\frac{U1+U2+\cdots Un}{n}$(với n=15) = 2.85 (kV)

Sử dụng công thức với bình phương phương sai là $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$

Suy ra phương sai S=0.193763

Ngưỡng ước lượng so với trung bình là: $\varepsilon = t_{\frac{\alpha}{2}}(n-1).\frac{s}{\sqrt{n}}$ =0.147436

Khoảng ước lượng là $\overline{(\overline{X} \pm \varepsilon)}$= (2.702564 ; 2.997436) kV

1.3. Trình bày code và ý nghĩa. *(chi tiết trong file đính kèm)*

1.3.1. Định dạng:

Bài code đươc thực hiện trên ứng dụng Dev C.

Viết dưới dạng ngôn ngữ lập trình C.

1.3.2. Ý tưởng thực hiện và trình bày:

Bài code được viết nhằm kết hợp môn Xác suất thống kê và kiến thức về môn Ngôn ngữ lập trình đang học để tạo kiến thức liên môn.

Đồng thời việc tự tạo nên các hàm tính tổng, trung bình, phương sai cũng như các kiến thức để tính α, ε giúp chúng ta hiểu rõ và nắm vững cách thức tính cũng như việc ghi nhớ, đọc hiểu quá trình các công thức và ý nghĩa của chúng.

Đồng thời đoạn code thích hợp việc điều chỉnh các giá trị mẫu thử khi thay đổi và tích hợp một số giá trị student thường gặp để điều chỉnh phù hợp với các độ tin cậy khác nhau.

```c
#include <stdio.h>
#include <stdlib.h>
#include <math.h>
// giai bai toan uoc luong khoang voi 15 mau
main ()
{
int i=0, j=0, k=0, l= 1;
double phuongsai = 0 , trungbinh = 0, mau = 0 ,student=0, tong , nguongsaiso = 0;
double biennho1 = 0, biennho2 = 0 , trunggian1 = 0 , trunggian2 = 0 ;
double khoangtincay = 0, min = 0, max =0 ;
double array[15];
    printf("nhom 18: tran ngoc duy");
    printf("bai toan uoc luong khoang\n");
    printf("moi nhap khoang tin cay, vidu 0.99\n");
    scanf("%lf", &khoangtincay);
     if (khoangtincay == 0.990)
     {student=2.947;}
     else if (khoangtincay==0.980)
     {student=2.602;}
     else if (khoangtincay==0.950)
     {student=2.131;}
     else if (khoangtincay==0.900)
     {student=1.753;}
     else if (khoangtincay==0.850)
     {student=1.517;}
     else { printf("moi nhap lai khoang tin cay\n");
     return 0;}
for(i=0 ; i<15 ; i++)
    {
    printf("moi nhap gia tri mau thu %d \n",l);
    l++;
    scanf("%lf",&mau);
    array[i]= mau ;
    }
for(j=0 ; j<15 ; j++)
    {
    biennho1 = array[j];
    tong +=biennho1;
    }
    trungbinh= tong/15;
for(k=0 ; k<15 ; k++)
    {
    biennho2 = array[k];
    trunggian1 += (trungbinh-biennho2)*(trungbinh-biennho2);
    }
phuongsai= sqrt(trunggian1/14);
nguongsaiso = student*phuongsai/sqrt(15);
min= trungbinh-nguongsaiso;
max= trungbinh+nguongsaiso;
printf("gia tri student là %f :\n", student);
printf("gia tri trung binh là %f :\n", trungbinh);
printf("gia tri phuong sai là %f :\n", phuongsai);
printf("gia tri nguong sai so là %f :\n", nguongsaiso);
printf("ket qua cuoi cung cua bai la\n");
printf("gia tri khoanng uoc luong la: ( %f , %f )\n", min, max);
printf("cam on cac ban va thay co da theo doi\n");
printf("xin chao va hen gap lai !!!\n");
    return 0;
}
```

```
C:\Users\ACER\Documents\NNLT\test.exe
nhom 18: tran ngoc duy
bai toan uoc luong khoang
moi nhap khoang tin cay, vidu 0.99
0.99
moi nhap gia tri mau thu 1
2.774
moi nhap gia tri mau thu 2
2.66
moi nhap gia tri mau thu 3
2.85
moi nhap gia tri mau thu 4
3.306
moi nhap gia tri mau thu 5
2.66
moi nhap gia tri mau thu 6
2.926
moi nhap gia tri mau thu 7
2.812
moi nhap gia tri mau thu 8
3.154
moi nhap gia tri mau thu 9
2.622
moi nhap gia tri mau thu 10
2.622
moi nhap gia tri mau thu 11
2.812
moi nhap gia tri mau thu 12
2.926
moi nhap gia tri mau thu 13
2.85
moi nhap gia tri mau thu 14
3.002
moi nhap gia tri mau thu 15
2.774
gia tri student la 2.947000 :
gia tri trung binh la 2.850000 :
gia tri phuong sai la 0.193763 :
gia tri nguong sai so la 0.147436 :
ket qua cuoi cung cua bai la
gia tri khoanng uoc luong la: ( 2.702564 , 2.997436 ) kV
cam on cac ban va thay co da theo doi
xin chao va hen gap lai !!!

--------------------------------
Process exited after 181.8 seconds with return value 0
Press any key to continue . . .
```

II. BÀI 2: Đánh giá độ tin cậy của hệ thống nguồn điện:

2.1. Giới thiệu nội dung, ý nghĩa các thông số, giá trị bài toán.

2.1.1. *Giới thiệu nội dung:*

Bài toán thứ hai đưa ta đến với công việc tương lai cho các sinh viên khoa điện hiện giờ. Ở đây, ta sẽ là người tham gia như một kĩ sư cho hệ thống truyền tải điện, có trách nhiệm quản lí, tính toán thời gian kỳ vọng thiếu hụt công suất nguồn LOLE (Loss of Load Expectation) cũng như lượng điện năng kỳ vọng bị thiếu LOEE (Loss of Energy Expectation) trong năm. Từ các điều kiện như đặc điểm nhà máy phát công suất và nơi tiêu thụ với các đặc tính tải, ta cần xác định các giá trị và rút ra các ý nghĩa để tạo nên phương án phù hợp

2.1.2. *Ý nghĩa các thông số*

-Hệ số cưỡng bức FOR:

Trong kỹ thuật điện , mất điện cưỡng bức là tình trạng ngừng hoạt động của trạm điện , đường dây tải điện hoặc đường dây phân phối khi tổ máy phát điện không thể sản xuất được do sự cố bất ngờ

Sự cố mất điện cưỡng bức có thể do lỗi thiết bị, gián đoạn dây chuyền cung cấp nhiên liệu của nhà máy điện, lỗi của người vận hành, v.v.

Tỷ lệ mất điện cưỡng bức (FOR hoặc FOAR) của một đơn vị trạm điện là xác suất mà thiết bị đó sẽ không sẵn sàng để phục vụ khi được yêu cầu.

FOR được định nghĩa là số giờ đơn vị bị mất điện cưỡng bức trên tổng số giờ trong một năm (là tổng số giờ nhà máy điện sẵn sàng phục vụ và số giờ nhà máy điện bị cắt điện cưỡng bức).

-Phụ tải:

Phụ tải điện là đại lượng đo bằng tổng công suất tiêu thụ của các thiết bị điện trong một thời điểm, đây là hàm số của nhiều yếu tố theo thời gian, không tuân thủ theo một quy luật nhất định và là một thông số quan trọng để lựa chọn các thiết bị của hệ thống điện.

Xác định đúng phụ tải điện (tính toán thường được xác định bằng xác suất và thống kê) có vai trò rất quan trọng trong thiết kế và vận hành hệ thống cung cấp điện.

Xác định phụ tải điện (phụ tải tính toán) xảy ra hai trường hợp: Nhỏ hơn phụ tải thực tế thường dẫn đến các sự cố hoặc làm giảm tuổi thọ các thiết bị, là nguy cơ tiềm ẩn cho các sự cố tai nạn sau này. Lớn hơn phụ tải thực tế sẽ gây lãng phí do các thiết bị không được khai thác, sử dụng hết công suất

-Phụ tải đỉnh:

Phụ tải đỉnh là khoảng giá trị công suất tiêu thụ tăng nhanh, đạt trị số cực đại và tồn tại trong thời gian ngắn. Tuỳ theo khoảng thời gian ta đang nghiên cứu mà sẽ định hướng thời điểm phụ tải đỉnh khác nhau. Ví dụ là phụ tải đỉnh trong ngày sẽ là các giờ cao điểm 17h đến 20h, trong năm sẽ là các khoảng thời gian mùa hè cần tiêu thụ lượng điện lớn.

Việc xác định phụ tải đỉnh có vai trò quan trong trong điện tử công suất để có biện pháp bù công suất, tăng công suất hệ thống điện truyền tải, đồng thời có các biện pháp phân bố điện phù hợp.

-Đường cong đặc tính tải:

Là đồ thị hàm công suất tiêu thụ theo biến thời gian , tuỳ theo muc đích mà ta có các đặc tính tải theo tháng, mùa, năm hay thời gian lớn hơn để theo dõi mức độ tiêu thụ có tốc độ phát triển như thế nào. Từ mỗi đường cong mà ta đọc được các giá trị và có phương án tức thời hoặc lâu dài phù hợp cho nhà máy điện.

Lưu ý: Ở bài toán này, để đơn giản hoá đồ thị và với mục đích giới thiệu, ta mô hình hoá đường cong đặc tính tải ở dạng phương trình bậc nhất y=ax+b.

2.1.3. *Giá trị bài toán:*

Về mặt học thuật, bài toán cho ta tổng ôn lại các kiến thức của xác suất lẫn thống kê từ các giá trị độ lệch chuẩn đến các cách tính xác suất đầy đủ, xác suất Bernoulli lẫn các hình thức phân phối chuẩn, phân phối nhị thức.

Về mặt thực tế, bài toán giúp ta có cái nhìn tổng quan về hệ thống truyền tải công suất và các vấn đề phát sinh trong quá trình này. Từ việc tồn tại hệ số FOR của các tổ máy công suất cho đến đặc tuyến tải trong nắm, các con số và giá trị này tạo nên giá trị, nói lên vai trò và nhiệm vụ cho người kỹ sư quản lí hệ thống truyền tải công suất.

Ở đây, các con số LOLE, LOEE tạo cho người kỹ sư phải xác định được các biện pháp hạn chế thiếu hụt, truyền tải công suất phụ đáp ứng các phụ tải một cách phù hợp đảm bảo ổn định cho cả hệ thống. Bên cạnh đó là vai trò người kỹ sư khi tính toán cân bằng giữa việc xây dựng bao nhiêu tổ máy và nhu cầu điện năng của vùng. Thực tế là ta không thể xây quá nhiều tổ máy bởi các lí do chi phí, đường truyền cũng như lãng phí tài nguyên, bên cạnh đó không được quá ít tổ máy gây thiếu hụt truyền tải cho khu vực. Đồng thời bài toán đặc vấn đề dự trù, thiết lập các trường hợp có xác suất tổ máy hư hoặc thời điểm trong năm thiếu hụt công suất có tỉ lê nhất.

2.2. Trình bày hướng làm, cách giải.

2.2.1. *Lý thuyết xác suất liên quan:*

Xác suất đầy đủ:

Giả sử ta có nhóm đầy đủ các biến cố $B_1, B_2, \ldots, B_n$ (nghĩa là 2 biến cố bất kỳ đều xung khắc với nhau và tổng của n biến cố này tương đương với biến cố chắc chắn:

$B_1 \cup B_2 \cup ... \cup B_n = \Omega$.. Biến cố A cần tìm xác suất quan hệ với nhóm đầy đủ như sau: Biến cố A xảy ra thì suy ra xảy ra biến cố Bi nào đó; còn ngược lại nếu xảy ra biến cố Bi nào đó thì chưa khẳng định biến cố A xảy ra. Khi đó P(A) được tính như sau:

$$P(A) = \sum_{i=1}^{n} P(B_i)P(A/B_i)$$

Hay một cách dễ hiểu là xác suất đầy đủ là chia nhỏ bài toán lớn thành các trường hợp biến cố nhỏ, tính chúng và cộng lại. Cũng giống như chia nhỏ phân công công việc, thực hiện và tổng kết lại.

-Phân phối nhị thức:

Phân phối nhị thức (binomial distribution) là một phân phối xác suất tóm tắt khả năng để một giá trị lấy một trong hai giá trị độc lập trong một tập hợp các tham số hoặc giả định nhất định. Giả định cơ sở của phân phối nhị thức là chỉ có một kết quả cho mỗi phép thử, mỗi phép thử có xác suất thành công giống nhau, và những phép thử này xung khắc hay độc lập với nhau.

Phân phối nhị thức là một dạng phân phối rời rạc thường dùng trong thống kê, ngược lại của các dạng phân phối liên tục như phân phối chuẩn. Điều này là vì phân phối nhị thức chỉ tính đến hai trường hợp, thường được thể hiện là 1 (cho thành công) hoặc 0 (cho thất bại) trong một số lượng lần thử. Ví dụ điển hình là tính xác suất số lần mặt sấp (hoặc ngửa) khi tung đồng xu n lần.

Phân phối nhị thức thể hiện xác suất để x thành công trong n phép thử, với xác suất thành công p của mỗi phép thử.

Phân phối nhị thức được tính bằng cách nhân xác suất thành công p lũy thừa số lần thành công k với xác suất thất bại lũy thừa chênh lệch giữa số lần thử n và số lần thành công. Sau đó, nhân với tổ hợp giữa số lần thử và số lần thành công vì số lần thành công có thể được phân bố bất kì trong số lần thử.

QUY LUẬT NHỊ THỨC - $B(n,p)$

$$P(X = k) = C_n^k p^k (1-p)^{n-k}$$

$$E(X) = np, \quad Var(X) = np(1-p), \quad Mod(X)$$

-Phân phối chuẩn:

Phân phối chuẩn (Normal distribution) được nêu ra bởi một người Anh gốc Pháp tên là Abraham de Moivre (1733). Sau đó Gauss, một nhà toán học người Đức, đã dùng luật phân phối chuẩn để nghiên cứu các dữ liệu về thiên văn học (1809) và do vậy cũng được gọi là phân phối Gauss.

Hai thông số quan trọng trong một phân phối là giá trị trung tâm hay gọi là trung bình μ và phương sai σ^2 (hoặc độ lệch chuẩn σ) và thường biểu thị bằng $X \sim N(\mu, \sigma^2)$ (N viết tắt từ normal).

Nếu phân phối chuẩn được chuẩn hóa với trung bình $\mu = 0$ và độ lệch chuẩn $\sigma = 1$, được viết tắt là: $Z \sim N(\mu = 0, \sigma = 1)$, được gọi là phân phối chuẩn chuẩn hóa (standardized normal distribution)

Nói chung, các đặc tính sinh trắc học của người khỏe mạnh (cân nặng, chiều cao, trị số mạch, huyết áp, đường máu, số lượng hồng cầu), thường tuân theo luật phân phối chuẩn. Và quan trọng có ứng dụng lớn là một số các qui luật phân phối khác khi đủ lớn và thích hợp sẽ hội tụ và qui luật chuẩn. Đồng thời qui luật chuẩn có giá trị lớn trong kĩ thuật để ước lượng xác suất và phân chia tỉ lệ phù hợp.

-Một số tính chất với phân phối chuẩn:

Hàm mật độ là đối xứng qua giá trị trung bình (giá trị kì vọng).

Khoảng giá trị ước lượng:

68.26894921371% của diện tích dưới đường cong là nằm trong khoảng 1 lần độ lệch chuẩn tính từ trị trung bình (tức là khoảng $(\mu - \sigma ; \mu + \sigma)$,

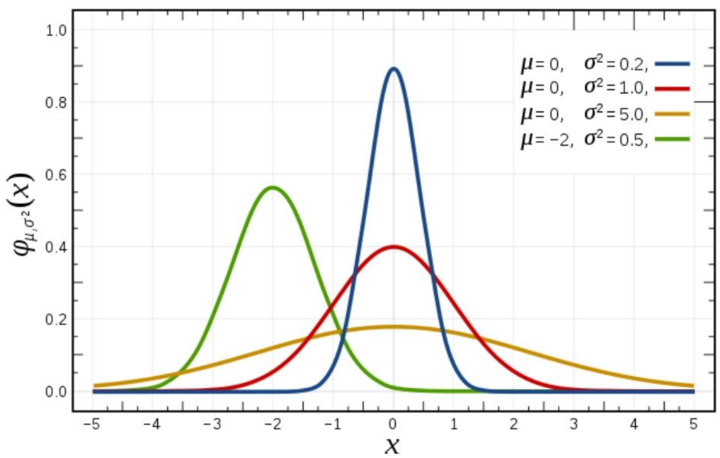
95.44997361036% của diện tích dưới đường cong là nằm trong khoảng 2 lần độ lệch chuẩn $(\mu - 2\sigma ; \mu + 2\sigma)$

99.73002039367% của diện tích dưới đường cong là nằm trong khoảng 3 lần độ lệch chuẩn $(\mu - 3\sigma ; \mu + 3\sigma)$.

Vậy nên trong các bài toán cần giá định giá trị phân vùng xác suất, ta thường chia 7 đoạn với các đoạn chủ yếu lấy từ $(\mu - 3\sigma ; \mu + 3\sigma)$. Điều kiện này đảm bảo chính xác cho bài toán đồng dễ thực hiện phân chia 7 đoạn cho các trường hợp cụ thể.

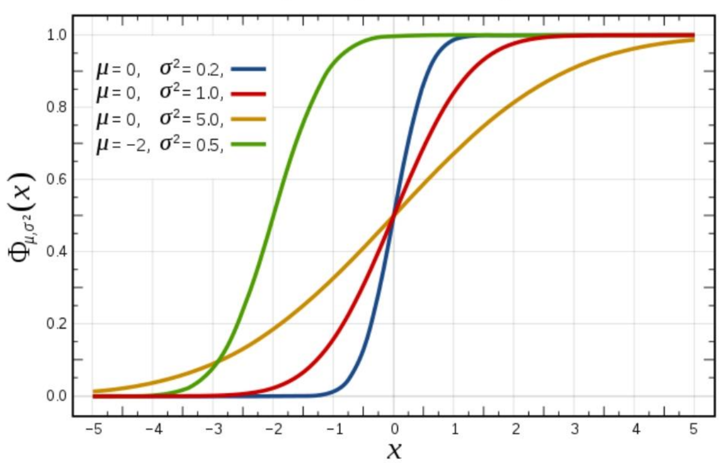
Phân phối chuẩn

Hàm mật độ xác suất



Đường màu đỏ là phân phối chuẩn tắc

Hàm phân phối tích lũy



Màu sắc tương ứng với hình trên

Tham số	μ cho biết vị trí (thực) $\sigma^2 > 0$ bình phương tỉ lệ (thực)
Giá	$x \in (-\infty; +\infty)$
Hàm mật độ xác suất	$\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$
Hàm phân phối tích lũy	$\frac{1}{2}\left(1 + \operatorname{erf}\frac{x-\mu}{\sigma\sqrt{2}}\right)$

2.2.2. Giải toán:

Mô tả bài toán:

Hệ thống nguồn điện gồm 12 tổ máy 6 MW, mỗi tổ máy có hệ số FOR = 0,09, dự báo phụ tải đỉnh là 59 MW với độ lệch chuẩn σ = 0,03. Đường cong đặc tính tải trong năm là đường thẳng nối từ 100% đến 40% so với tải đỉnh.

a, Xác định thời gian kỳ vọng thiếu hụt công suất nguồn LOLE (Loss of load expectation) trong năm.

b, Xác định lượng điện năng kỳ vọng bị thiếu LOEE (Loss of energy expectation) trong năm.

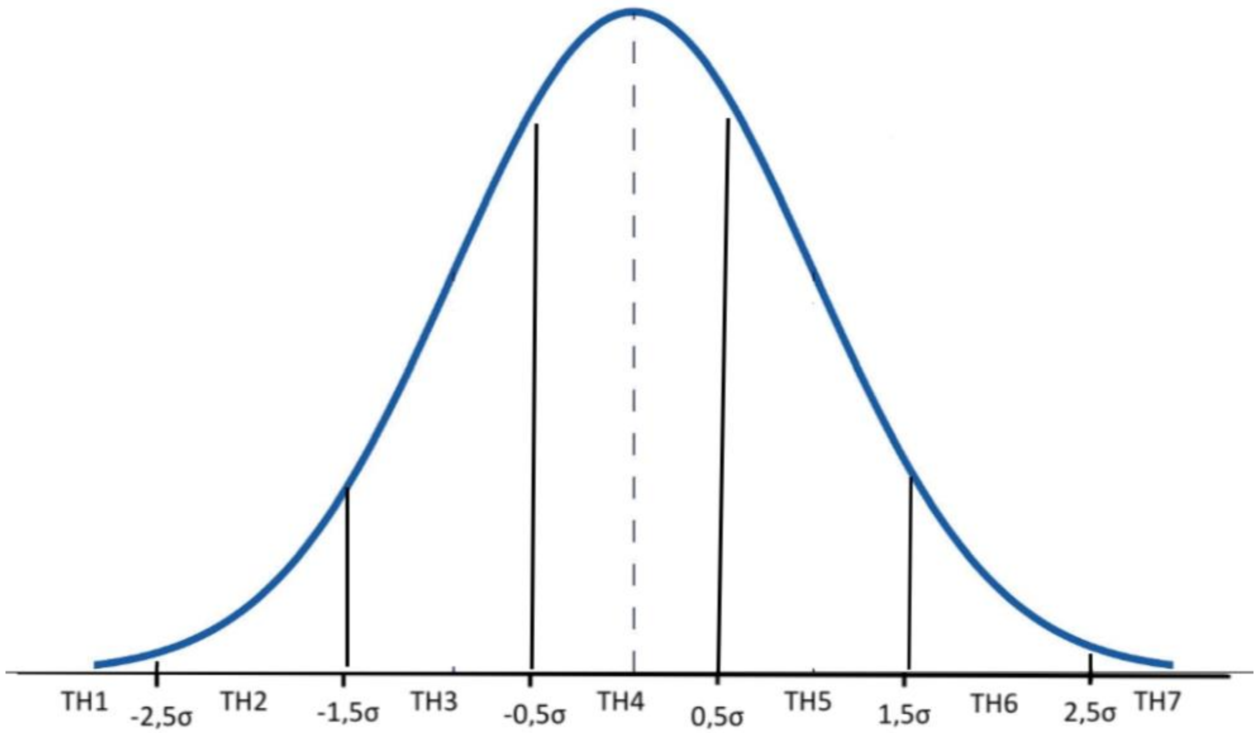
Theo đề bài

Số lượng tổ máy	n	12
Công suất của một tổ máy	P (MW)	6
Tải đỉnh	Pload (MW)	59
Đặc tính của tải trong năm	Px	40%
Hệ số cưỡng bức	FOR	0.09

Độ lệch chuẩn	σ	3%

Có 7 trường hợp có thể xảy ra:

$$\sigma = 1$$



Ta có bảng xác suất các trường hợp như sau

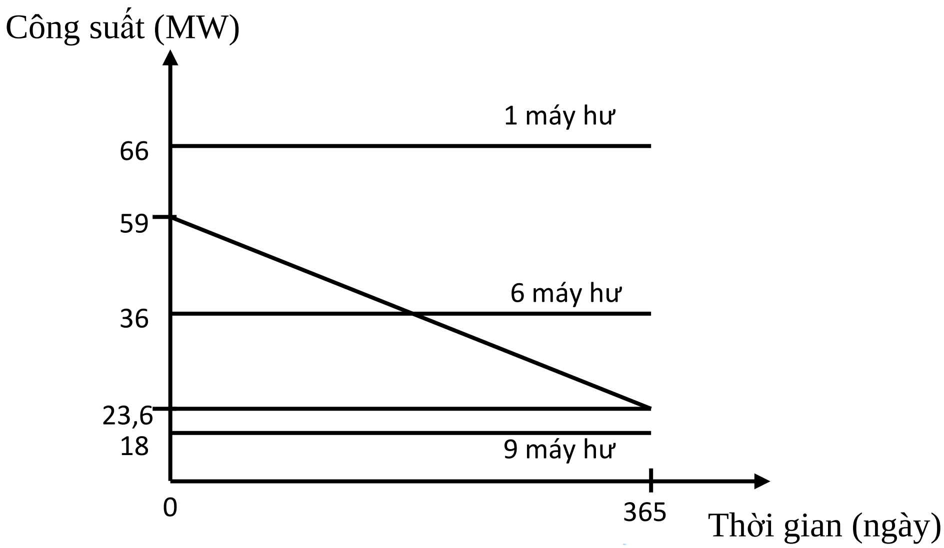
Trường hợp	Lệch -3σ	Lệch -2σ	Lệch -1σ	Lệch 0σ	Lệch 1σ	Lệch 2σ	Lệch 3σ
Xác suất	0,006	0,061	0,242	0,382	0,242	0,062	0,006

Pmax = (Tải đỉnh) *(1± α *σ) với α = -3 → 3

Pmin = Pmax * 40%

Trường hợp	P_{max}	P_{min}
Lệch -3σ	53,69	21,476
Lệch -2σ	55,46	22,184
Lệch -1σ	57,23	22,892
Lệch 0σ	59	23,6
Lệch 1σ	60,77	24,308
Lệch 2σ	62,54	25,016
Lệch 3σ	64,31	25,724

a, Ta giải cho 1 trường hợp lệch 0σ



Phương trình đặc tuyến tải có dạng: y = ax + b

→ a = $\frac{59*0,4-59}{365}$ ≈ - 0,097 ; b = 59

Xác suất 1 tổ máy hư là FOR = 0,09

→ Xác suất để 1 tổ máy không hư là (1- FOR)= 1-0,09= 0,91

Xác suất tổ máy hỏng trong từng trường hợp: p= $C_n^k\ (FOR)^k(1-FOR)^{n-k}$

Liệt kê các trường hợp các máy bị hỏng: 0 tổ máy đến 12 tổ máy

Từ số máy hỏng ta suy ra công suất thiếu hụt bằng công thức: số máy hỏng * 6 (MW)

→ Công suất hoạt động: P = 12*6 – số máy hỏng *6

Thời gian thiếu hụt công suất nguồn

. t=0 khi P≥59

. t = $\frac{(công\ suất\ hoạt\ động)-59}{-0,097}$ khi 59≥P≥23,6

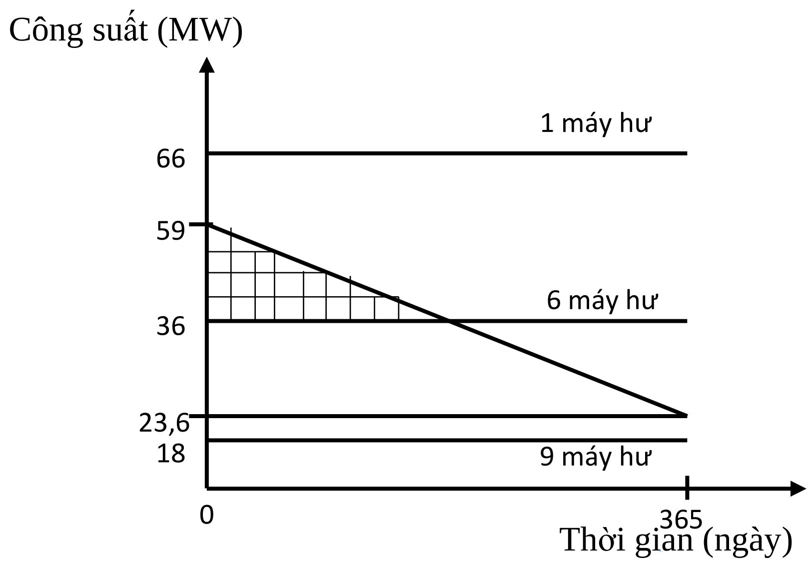
. t= 365 khi P≤ 23,6

→ thời gian kỳ vọng thiếu hụt công suất nguồn là: p*t

Số tổ máy hỏng	Xác suất tổ máy hỏng	Công suất thiếu hụt	Công suất hoạt động	Thời gian thiếu hụt công suất nguồn	Thời gian kỳ vọng thiếu hụt công suất nguồn
0	0,3224754874	0	72	0	0
1	0,3827181609	6	66	0	0
2	0,2081818567	12	60	0	0
3	0,06863138134	18	54	51,54	3,537
4	0,01527236783	24	48	113,4	1,724
5	0,002416726337	30	42	175,26	0,422
6	0,0002788530389	36	36	237,11	0,066
7	0,00002363903313	42	30	298,97	0,00706
8	0,000001461203971	48	24	360,82	0,0000527
9	0,00000006422874597	54	18	365	0,00000234
10	0,000000001905688067	60	12	365	0,0000000695
11	0,00000000003426811709	66	6	365	0,0000000125
12	0,0000000000002824295365	72	0	365	0,000000000103

→ Thời gian kỳ vọng thiếu hụt công suất nguồn trong trường hợp lệch 0σ là:
$\sum Thời\ gian\ kỳ\ vọng\ thiếu\ hụt\ công\ suất\ nguồn$ = 5,767721101

b, Tính lượng điện năng kỳ vọng bị thiếu hụt



Điện năng bị thiếu trong năm là

. E = 0 khi P ≥ 59

. E = $\frac{(59-P)*t}{2}$ khi 59 > P > 23,6

. E = $\int_0^{365}(-0{,}097x + 59) - P$ khi P < 23,6

→ Điện năng kỳ vọng bị thiếu trong năm là: E*p

Số tổ máy hỏng	Xác suất tổ máy hỏng	Công suất thiếu hụt	Công suất hoạt động	Điện năng bị thiếu trong năm	Điện năng kỳ vọng bị thiếu
0	0,3224754874	0	72	0	0
1	0,3827181609	6	66	0	0
2	0,2081818567	12	60	0	0
3	0,06863138134	18	54	128,85	8,843

4	0,01527236783	24	48	623,7	9,526
5	0,002416726337	30	42	1489,71	3,6
6	0,0002788530389	36	36	2726.765	0,7605
7	0,00002363903313	42	30	4335,065	0,1025
8	0,000001461203971	48	24	6314,35	0,009
9	0,00000006422874597	54	18	8503,588	0,000546
10	0,000000001905688067	60	12	10693,588	0,0000204
11	0,00000000003426811709	66	6	12883,588	0,000000441
12	0,0000000000002824295365	72	0	15073,588	0,00000000426

→ Điện năng kỳ vọng bị thiếu trong trường hợp lệch 0σ là: $\sum Điện\ năng\ kỳ\ vọng\ bị\ thiếu$ = 22,84584804

Các trường hợp còn lại chúng ta sẽ làm tương tự

Các trường hợp	Xác suất xuất hiện	Thời gian kỳ vọng thiếu hụt công suất nguồn	Điện năng kỳ vọng bị thiếu
Lệch -3σ	0,006	1,367472347	5,249535887
Lệch -2σ	0,061	2,772263831	8,536682741
Lệch -1σ	0,242	4,31631397	14,4701765

Lệch 0σ	0,382	5,767721101	22,84584804
Lệch 1σ	0,242	8,739246652	34,06810079
Lệch 2σ	0,062	13,56757436	52,62656712
Lệch 3σ	0,006	18,13012256	78,90014016

Thời gian kỳ vọng thiếu hụt công suất nguồn LOLE (Loss of load expectation) trong năm là: $\sum(Thời\ gian\ kỳ\ vọng\ thiếu\ hụt) * (Xác\ suất)$ = 8,49998405

Lượng điện năng kỳ vọng bị thiếu LOEE (Loss of load energy expectation) trong năm là: $\sum(lượng\ điện\ năng\ kỳ\ vọng\ bị\ thiếu) * (Xác\ suất)$ = 24,76185992

2.3. Trình bày code và ý nghĩa. *(chi tiết trong file đính kèm)*

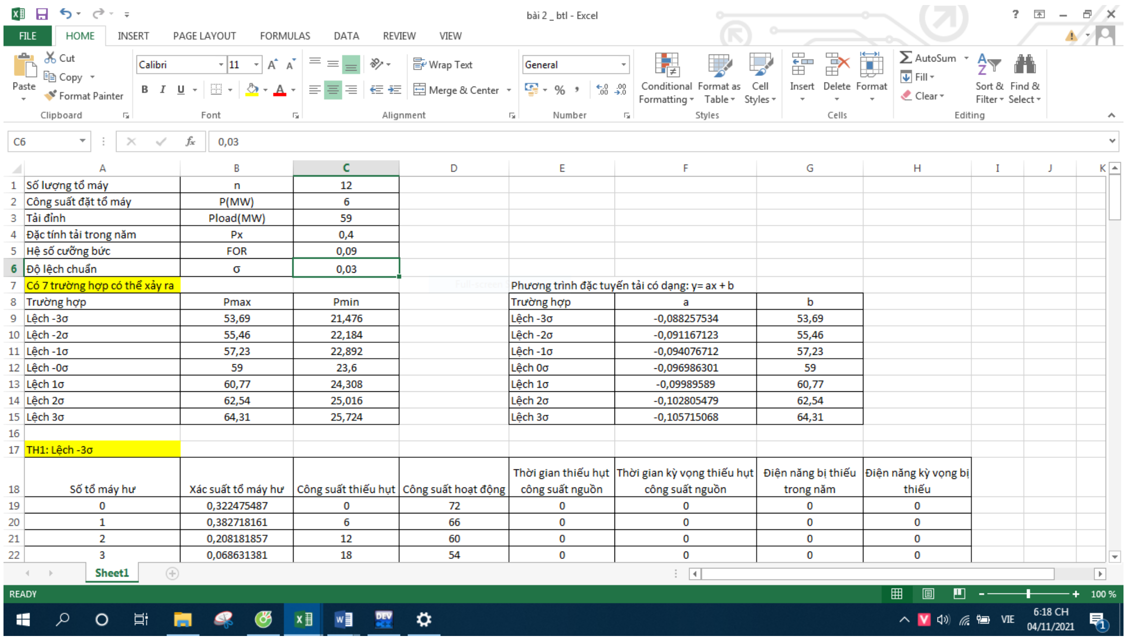
2.3.1. Định dạng:

Bài code được thực hiện trên ứng dụng Excel

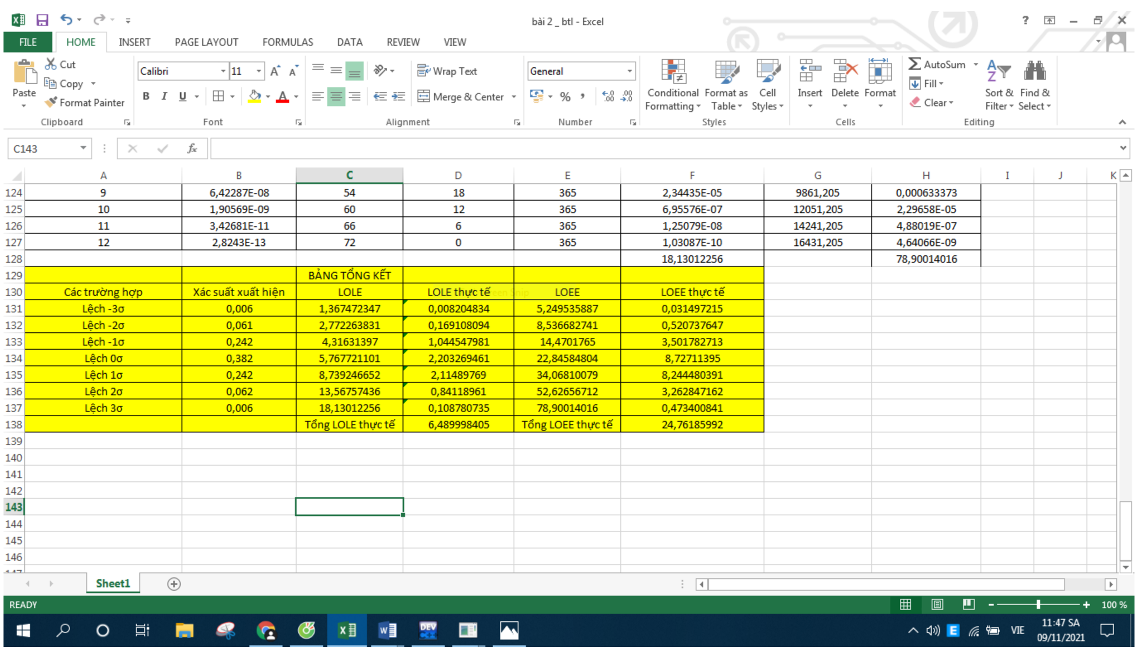
File định dạng: xls

2.3.2. Ý tưởng thực hiện và trình bày:

Để có thể trực quan nhìn vào các số liệu đồng nhất, bài code được trình bày trên ứng dụng Excel. Vì số liệu và các trường hợp khá lớn, ta biểu diễn trên Excel có lợi thế phân chia các vùng số liệu dễ dàng và so sánh các trường hợp trực quan. Bên cạnh đó tạo ưu thế khi điều chỉnh số liệu khi bổ sung, tiếp cận dễ dàng, miễn phí và có khả năng sử dụng lớn, là công cụ có khả năng mạnh trong việc tính toán và sử lí số liệu.



Số lượng tổ máy	n	12					
Công suất đặt tổ máy	P(MW)	6					
Tải đỉnh	Pload(MW)	59					
Đặc tính tải trong năm	Px	0,4					
Hệ số cưỡng bức	FOR	0,09					
Độ lệch chuẩn	σ	0,03					
Có 7 trường hợp có thể xảy ra				Phương trình đặc tuyến tải có dạng: y= ax + b			
Trường hợp	Pmax	Pmin		Trường hợp	a	b	
Lệch -3σ	53,69	21,476		Lệch -3σ	-0,088257534	53,69	
Lệch -2σ	55,46	22,184		Lệch -2σ	-0,091167123	55,46	
Lệch -1σ	57,23	22,892		Lệch -1σ	-0,094076712	57,23	
Lệch -0σ	59	23,6		Lệch 0σ	-0,096986301	59	
Lệch 1σ	60,77	24,308		Lệch 1σ	-0,09989589	60,77	
Lệch 2σ	62,54	25,016		Lệch 2σ	-0,102805479	62,54	
Lệch 3σ	64,31	25,724		Lệch 3σ	-0,105715068	64,31	
TH1: Lệch -3σ							
Số tổ máy hư	Xác suất tổ máy hư	Công suất thiếu hụt	Công suất hoạt động	Thời gian thiếu hụt công suất nguồn	Thời gian kỳ vọng thiếu hụt công suất nguồn	Điện năng bị thiếu trong năm	Điện năng kỳ vọng bị thiếu
0	0,322475487	0	72	0	0	0	0
1	0,382718161	6	66	0	0	0	0
2	0,208181857	12	60	0	0	0	0
3	0,068631381	18	54	0	0	0	0



9	6,42287E-08	54	18	365	2,34435E-05	9861,205	0,000633373
10	1,90569E-09	60	12	365	6,95576E-07	12051,205	2,29658E-05
11	3,42681E-11	66	6	365	1,25079E-08	14241,205	4,88019E-07
12	2,8243E-13	72	0	365	1,03087E-10	16431,205	4,64066E-09
					18,13012256		78,90014016

		BẢNG TỔNG KẾT			
Các trường hợp	Xác suất xuất hiện	LOLE	LOLE thực tế	LOEE	LOEE thực tế
Lệch -3σ	0,006	1,367472347	0,008204834	5,249535887	0,031497215
Lệch -2σ	0,061	2,772263831	0,169108094	8,536682741	0,520737647
Lệch -1σ	0,242	4,31631397	1,044547981	14,4701765	3,501782713
Lệch 0σ	0,382	5,767721101	2,203269461	22,84584804	8,72711395
Lệch 1σ	0,242	8,739246652	2,11489769	34,06810079	8,244480391
Lệch 2σ	0,062	13,56757436	0,84118961	52,62656712	3,262847162
Lệch 3σ	0,006	18,13012256	0,108780735	78,90014016	0,473400841
		Tổng LOLE thực tế	6,489998405	Tổng LOEE thực tế	24,76185992

III. Kết luận

Trải qua thời gian phát triển, xác suất và thống kê từ một nhánh của Toán học đã hình thành và phát triển mạnh mẽ. Chúng tạo nên những ứng dụng vô cùng lớn xung quanh chúng ta ở thời đại 4.0, thời đại mà những tổ hợp những có số sức mạnh vô cùng lớn. Nếu chúng ta biết sử dụng chúng, ta có thể đọc được, chọn được những cách thức xử lí phù hợp hay nói xa hơn, xác suất và thống kê định hình cho chúng ta biết về tương lai và xây dựng cho ta những dự trù phù hợp.

Trực tiếp đến với kĩ thuật, nơi đặc biệt gắn liền với những số liệu, xác suất và thống kê tạo nên những yêu cầu cho những kĩ sư tương lai về nền tảng cơ bản, cách thức sử dụng những con số vào công việc của mình. Tiêu biểu là qua 2 bài toán trên, đã giúp ta định hình cách thức xử lí, vận dụng những con số này vào thực tiễn.

IV. Tài liệu tham khảo

1. Nguyễn Kiều Dung(2021). Slide bài giảng môn Xác suất thống kê/ đại học Bách Khoa tp.HCM
2. https://www.youtube.com/watch?v=HFdjtV24xV4&list=PLsEmKKF4H46mAkcfBDDu6Qo1EY7vZwO5k
3. https://vietnambiz.vn/phan-phoi-nhi-thuc-binomial-distribution-la-gi-nhung-dac-diem-can-luu-y-20191118140522351.htm
4. https://vi.wikipedia.org/wiki/Ph%E1%BB%A5_t%E1%BA%A3i_%C4%91i%E1%BB%87n
5. Và một số tài liệu tham khảo nội dung liên quan, tài liệu code khác.

BÁO CÁO KẾT QUẢ LÀM VIỆC NHÓM

STT	**Mã số SV**	**Họ**	**Tên**	**Nhiệm vụ được phân công**	**Ký tên**	**Phần trăm**
1	2012846	Trần Ngọc	Duy	Nhóm trưởng Giải, code bài 1 Lên nội dung, giải , soạn thảo word.		**34%**
2	2015055	Đinh Quang	Vinh	Trình bày , code bài 2.		**33%**

3	2015110	Trần Hoàng	Vương	Soạn thảo powpoint.		**33%**