-St~~a~~t~~i~~s~~t~~ics Bas~~i~~c

3rdZ

Semester

-Louisa Mandy Halim

statistics

Week 1

G jus =

Keroper

Three types of data

· quantitative , consists of numbers weights , heights

· Ranked , -11-relative standing within a group

·qualitative , describ~~i~~ng how they feel , look , wear

Level of measurement

-sorting observations into different

classes or categories

-classy fying

independent & dependent variable

z

Week 2

stem and leaf

·pie chart22 , 24 , 26 , 29 · Bar Chart 50% 13 , 14 , 15 , 16 , 17 , 19

6 , 7

2 2 , 4 , 6 , 9

314 , 5 , 61 7 , 9 - lea~~f~~

1 0 617

stem

Comparing distributi~~o~~ns example

↓ ~~Hi~~s~~t~~ogram

a bar chart of the number of people

playing different card games on graph~~i~~cal me~~th~~od Wednesday and Thursdayd~~i~~splaying ~~th~~e shape of make simple and easy to understand distributi~~o~~n

designs for charts

-~~fr~~equency ~~ta~~ble

-his~~t~~ogram

-

-

-

-

frequency polygons

lower upper count

good cho~~ic~~e to

display frequency d~~i~~stribution

Box plo~~t~~

30 · comparing genders 15 · identifyi~~n~~g outliers and for comparing d~~i~~s~~t~~ribu~~t~~ions

z

15

percentile

80% o~~f~~ the people are shor~~t~~er than you

no cap

Exercise 01 -Sep 11 , 2024 ① create a stem and leaf Display

Data Set :

62 , 65, 68 , 70 , 73 , 75,75, 78, 81 , 83 , 84 , 85 , 87 , 89 , 92 , 95 , 96 , 98 , 100

Stem leaf

6 2 , 5 , 8

70 . 3 , 5 , 5 , 8

8 1 , 3 , 4 , 5 , 7 , 9

92. 5~~,~~o

② construct a box plot

Dataset:↓ ↓ ↓

55 , 60 , 62 , 63, 65 , 66 , 68 ,70 . 72 , 75 ,77 , 78 , 80, 85 , 88 -

a) Determine the five number summary

· minimum value = 55

· maximum value = 88

· Q1 = 03

· Q2 = 70 (middle)

· Q3 = 78

b) Draw the boxplot

Q Q2 Q3

"

55 6370 78 &a

a identify potential outliers

= 78-63 ~~!~~ = 63- (115 X 15)

IQR = Q3-Q 1.5 lower fence = Q1 -(115 X IQR)

= 15 = 63- 22,5

= Y1 , 5

upper fence = Q3 + (1 , 5 x IQR] -

= 70 + (115 X 15)

= 70 +22 ,5 since all data points lie within = 100 : 5 range 41 . 5 to 10015~~.~~There are no outliers

Exercise 02-Sep 25 , 2024 ① calculate the Trimean for a dataset below

Dataset :

↓ ↓ ↓

10 , 12 , 15 , 18 , 21 , 24 , 27 , 30 , 33 , 36 , 39 , 42 , 45 , 48 , 50 Since the dataset is sorted already -find the median :

Q1 =18

Q2 = 30

Q3 = 42 The G~~r~~imean for this datase~~t~~

Calculate the Trimean

Trimean = Q1 +2 x Median + &3 Y

is 30

Trimean formula :

= 10 + 2(30) + 42 ~

+(my 2)+ 93

Y

= 120/y = 30

② Geometric Mean

suppose that the population of a City changes over four years , w~~i~~th the follow~~in~~g annual growth rates :

year 2 : + 10 % -> 11 GM = ~~"~~.72 .. .

> 1,05

Year 1 : + 5 % -

> 0 .97

year 3 : -3 % -

> 1,06

Year y : + 6 % -

Geometric mean formula :

Calculate the geometric mean of the growth rates to find the o M =T1105 ·1 , 1 . 0 ,97 . 1 ,06

average population growth rate over these 4 years

= 1 , 0426

= 1 , 0426 -1 = 0, 0426 x 100

> the growth rate over y years

= 4 , 26 % -

③ TrimmedMean10 % of 10~~=~~ X\* = 1 , remove from 10 - each end

55, 70 , 72, 75 , 80 , 85 , 90 , 95 No

calculate the 20 % trimmed mean (Trim 10 % from both Sides 70+72 + 75 + %0 + 85 + 90 +95~~=~~82 ,375

8 -

add conclusion

AMore examples

mean~~:~~

a figure skating competition produces the following Scores :

6 . 0 ; 8 , 1 ; 8 . 3 ; 9, 1 ; 9 ,9

20 + & , 1 +dis + gl + 9,a mean trimmed by

=~~Ma~~ 8 . 5 US & . 28 , which

40% would equal

= reduced the original

Trim the mean by 40 %& mean by 0 . 22 points

-remove the lowest 20% and the

~~+~~

highest 20% of values· eliminating

the scores : 6 , 0 & g . g

Exercise 03 -Oct 02 , 2024

~~-~~?

① you have a people , and you need to select and arrange y of them in a row of photo. ~~n~~e~~=10~~ ~~a~~-

How many different ways can you arrange ?

②You have I books , and you want to choose 4 to take on a trip. How many d~~i~~fferent ways can you select the books

I~~=~~was

③ a bag containslo red balls and 15 blue balls . If you randomly 25 (5~~:~~=~~M~~cN~~:~~

15(2=~~=E~~

select 5 balls without replacement , what is the possibility that exactly 3 Combination formula : nCr=

of the selected balls are red ?

Hypergeometric formula : p

.r!I

= 105)120

~~>~~

X24X23x22x21 - 530

in~~a~~

probabilit~~y~~-

· more

examples

how

ways a committee consisting of 5 men and 3 Women , can be chosen from men and 12 women many

Total number of ways ↳ 126 X 270 = 27728

Exercise on - October 9 , 2024

① find the percentage returns from an investment over 5 > 0, 1 6M = =x0115 x0195x0,8X0, 12

consecutive years , were :- Year 1 : 10 % -

Year 2: 15 % > 0,15

Year 3 : -5 % => 0 . 95 = 1 , 073

Yeary : 8 % -> 0 , 8

=

> 0 , 12 1 , 073 -1 = 0 , 073x100

year 5 : 12 % -

② Box Plot

Dataset :

↓ ↓ ↓

Group A : 7 , 9 , 12 , 13 , 17 , 15 , 16

Group B : 51 7, 8 , 10 , 12 , 15 , 18

↑ ↑ ↑

a) calculate the fire number summary c) compare the distributions of two groups based on the box plots.

· min value = 7

· min value = 5

· max value = 16 o max value = 18 o which group has a higher · Q = 9 L Q1 = 7 median ? -Group A

I

· Q2 = 13 · Q2 = 10

· Q3 = 15 · Q3 = 15·Are there any outliers ? no outliers

Group A Group B

19-

b)~~"~~

13-

11-

g-

~~-~~

③ a card is drawn from a standard deck of 52 cards , and then a coin is flipped. what is the probability of drawing"King" from the deck and flipping a "Tail" ?

1 Deck = 52 cards

Kings = Y

~~\*~~\*\*~~=~~

↑

four kings

I

tail side,

out of theout of the

total cards 2 sides (coin)

④ stem and leaf

Department : 12 , 14 , 17 , 19 , 21 , 24 , 26 , 28 , 30 ,32

Department y : 13 , 16 , 18 , 20 , 23 , 25 , 27 , 29 , 31 , 33

Back to back stem and leaf display

Department c Department y

2 , 4 , 7 , 9 13, 6 , 8

1,4 . 6 , 8 2 0 . 3 1 5 , 7 , 9

0 .23 1 . 3

⑤ The probability of getting exactly 3 heads when

5(3~~:5~~ ~~==~~10~~w~~ay 3 probability

flipping a fa~~i~~r coin 5 times (where getting heads is considered a success) sing the formula ~~!~~(l

after 5 flips - > 25= 32

a

n = 5 (number of trials) = 10X 0. 53 X 0.52

T = 0 .5 (probability of success) =X~~XX~~ = 3 (number of successes)

=~~F~~ -> probabil~~i~~ty

-

⑧ in a basketball game , aplayer has a free throw success rate of 80 %. If the player takes 15 free throws , what is the probability that they make at least 12 Successful free throws ?

formula :(~~A)~~ ·px .qn ->~~M~~p. a

n = 15

7 = 12 , 13 , 14 , 15

p = 0,d

q = 012

· for 15 C12

~~↳~~Xaxo = 02

3 Total = 0,635

· for 15 (3 ↓ 4) ~~x0~~x02 = 02~~t~~he

· for 15 (14 3 probability

↳ada = 07

·for 3 16~~5~~ X 08 ~~x~~0120 = 00s

Hours of Height (cm)

sunlight

⑦ x Y (x5) (y - Tj) (x-5c)(y - ij) (7) -5t (y -j)2

2 10 -Y -18 40 16 100

Y 15 -2 -5 18 Y 25

6 200000 O

& 15 25 10Y 25

18 30 Y 18 YO 10 100

Total = 30 100 8 8 100 40250

Mean = 6 20 O j

formula ~~:~~=

Calculate the Pearson correlation Coefficient

↓

mean of > = 2 , 4 , 6 , 8 , 10

↓

mean of y =10 , 15 , 20 , 25 , 30

Exercise 05-october 23 Oct , 2014

① Find the standard deviation from those data

Scores : 70, 85 , 75 , go , 88

mean -> 70~~-~~

+ 85 + 75+ 90 +80 = 82, 2

5

(70 - 82 ,572 + (05 -82 ,5)2 + (75-82,572 + (98-825)" + 100 -02,5)" =(- 12 ,2) + (2,8) + (-4,2)" + (718) + (5,07 5 5

variance = 53,76

Standard deviation = 55576-

> 7 1 33

=

② suppose a survey indicates that 30 % of people prefer coffee over tea .If you randomly select 100 people , what is the probability that fewer than 25 people prefer coffee ? use z-table

because 25 -) 25-0. 5 = 24. 5

n = 100

p = 30 % - 0,3 9 = 0170

M = nxP

2~~=45~~0-~~5~~-

=100 X7 = 30from z table , -1 ,20- 0 . 1151

0 = /oox 0 ,3 X 0 , 7The probability that fewer than 25 people prefer coffee = E2)4, 50 is approximately 0 . 1151 or 11 . 51 %

③ you are conducting an experiment with 100 trials (n =100) and the probability of success in each trial is p = 0,Y , find the probability

that at least 45 success will occur

n = 100 M = nxp o : Moxo, y x0 1 6

p = 0 . Y = 100 X0, 7

2 =~~1445-40~~ -

9 = 0 . 6 = 40 = 25o -

> Y,89

from 2 table+ > 0 , 8212

1- 0 , 8212 = 0 , 1788 ->11%

Exercise of -November 20, 2024

① a company claims their light bulblast 1000 hours on average.

formula :~~to~~ a

950 ,960 , 970 , g00 , 1020 , 1030 , 990 , 1010 , 1000 , 955 -> mean = ggoi5 Test whether the mean l~~i~~fespan differs significantly from 1000 hours using X = 0 1 05

S = 25 187

t statistic mean =99015

r = 1000

compare t Statistic with Critical value

Degrees of freedom : n - 1 -> 10-1 = 9

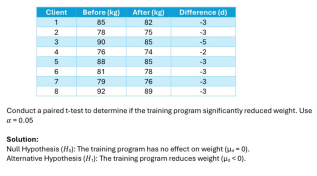
at a = 0,05 (two tailed)

the critical value is approximately 12. 262 (from table

Since t = -1 , 16 falls within the range

[-2.262 ; 2 .262] , we fail to reject the null hypothesis

② a fitness coach measures the weight of 8 clients before and after a 6 week

training program . 

sa~~=~~= 0 .03 Sd or s~~ta~~ndard deviation formula :

Conclusion : (x = 0 .05)

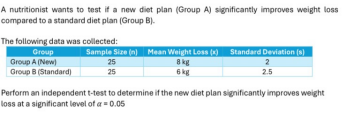
Degrees of freedom = d -1-)7 t crit = 2,365

t~~ormul~~a :+35~~/~~lig significant difference

Since -10 .594-2 ,365

reject null hypothesis

standard=~~Ed~~ + -~~3~~3 - 5 -2-3 - 3 =3 -3 = -3 , 125 deviation

③ 

~~t~~t+ 31 3 calculate + statistic

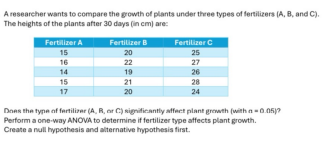
Degrees of freedom (df) = na + nB-z = Yo

from EtableCone ta~~i~~l) = 1 , 679

sincevalue) Critical value

3113) 1 :679

reject hull hyp~~ot~~hes~~i~~s

Exercise 07 - December 11 , 2024 O ~~+~~

·Fertilizer

A mean

16 +14 + 15 + 17 == - 15 ,4(J~~A)~~

· Fertilizer B mean

22+19+ 21 + 20 =~~1~~+ 204(B)

· Fertilizer c mean

~~1~~527 +20 + 20 + m = 26 (~~x~~)

it = 20~~+~~

· Overall means

2 + 19 + 21 +20+ 25 + 27 + 26 + 20 += 2016 ~~T~~

Sum of squares

· Sum of squares (SST)

SST= (15 -20 . 672 + (16 -201 672 + + (24= 20.6) 2 = 89.2

Sum of Squares between group (SSB)

SSB = 5 [(15,4 -20,67 + (20,4 -20,67"+ (26 -20 16)]

= 281 ,04

in this case.p < 0 , 001 < 0, 05

· Sum of squares within group CSSW) so we reject the null hypothesis

SSW = SS total + SS between

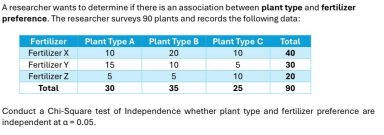
the pvalue is extremely small (P( 0, 001),

which indicates very strong evidence against the -M null hypothesis . Therefore , we conclude that the

type feltilizer has a significant effect on plant growth.

conclusion -> The F-static (F = 24.35) is significant effect on plant growth. At least one fertilizer produces a different mean plant height.

②

calculate the expected Frequencies

formula=

Now totalColumnois

· plant type A :0x~~40~~ =~~120~~ 13. Fertilizer z

Fertilizer

· plant type A :

· plant type B :35x4 ~~-~~ - 15156/15155 ~~b==~~6. · plant type C :~~ZX~~40~~=~~I , weEX20 ~~=~~ -7~~7~~7

· plant type B :

Fertilizer Y

I · Plant type C :

· ~~plan~~t~~a~~~~=5~~- 5,55 · plant type C :~~Z~~ Fertilizer y :Fer~~t~~ilizer Z Fertilizer plant~~A~~ plant B Plant c ~~to~~t~~al~~= 1~~61~~66 = l

·Plant type A: · plant type A

~~·~~Fertilizer · plant type (

x 13,33 15.56 11 ,Il

Y 10 11 . 06 8 , 33 38 · plant type B :· plant type B Z 6 , 66 7 ,77 5155 20 ~~166~~ =0~~13~~ , Total 29 ,99-30 34,997 3524,99725·plant type C :

compute chi-square statistic : ·~~P~~l~~an~~t type = 1 .2~~6~~·plant type C = 0 1~~03~~ summing all given v~~alt=~~

~~Degrees~~ offreedom-1x (number of Columns -1)compare tes~~t~~ sta~~t~~istics

to critical value

= (3-1)x(3-1) · if (21 xcritical , fail to rejec~~t~~ Ho = 2X2 · if <" >"Intical .Reject Ho df = Y

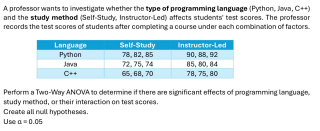
Here :

DetermineCritical valuex2 = 11 , 21

from Chi square table<critical =g , 48 for df = Y , 4 = 0105

the Critical value is 9, 488 Since 11 ,21) 9 , 488 reject the

hull hy~~poth~~esis

③ 

-

Python self study :~~[~~~~+~~2+5 = 16 python instructor - led

· find Grand/Group mean

Java Self Study :~~EE5~~+7 = 73,6~~6d~~ = go(++7~~075~~0 = 77 Java instructor

+70 = 67 , 6 ~~-~~= 13

(++ Self Study=

65 +

60

15 + 20 + 84

Grand mean (~~)~~ =~~1166~~ + 73,66 + ~~67~~ ,66+ 90 + 83 + 77, 66 =~~7~~8194

Group means Python & self study 81 , 66 Python (01 . 66+ 90)= 2 = 85 ,83 python & instructor go Java (73,66 + 83) = 2 = 70 , 33 Java and self 73, 66 C+ + (67 , 66 + 77, 66) = 2 = 72, 06 Java & instructor 83 self total all= g = 74 ,33 < + + & self 67 , 66 instructor -11- & 3 , 55 C++ & instructor 7 7 . 06

Compute Sum of squares :

To~~t~~al = E (xi~~j~~ -j~~)~~ ?

~~7~~ = 70 , 94

-Sum of squares for programming languag~~e~~ ~~(A~~)

SSA = 6 .(85 , 83 -78 .94)2 + 6 . (78 . 33 -78 .gy)" + 6 .(72, 66 -78,97) Python Java C + +

SSA = 523 , 6956

-Sum of squares for factor study Method (B)

SSB = 9 (74 ,33 -70 .94)2 + 9 (83, 55 -70.94)2

self study instructor

SSB = 382 , 5378

-Sum of Squares Within (error) Single numbers from tables -Group means ·SSpython & self study = (70 -01 , 66(2 + (82 -01166) + (85 -81,06" = 24, 6668 · SS Python & instructor-led =190-go (2 + (80~~-~~go( + (g2-~~g~~o)"= &

· SS Java & self Study = (72-73,66)2+ (75 -73,672 + (74 -73,6672= 4, 6668 I

· SSJava & instructor-led = (15-83)2+ (00 - 8372 + (04 -&3)== 14 · SS (++ & self study = (65 =67,6672 + (60 -67.06) + (70 -67,0)"= 12 , 9512 · SS ( + + & instructor - Led = (70-77,6672 + (75 -77,66)2+ (80 -77,60) = 12, 6668

SSE = 24 , 1668 + & + 4 , 6668 ~~+~~ 14 + 12 , 9512 + 12 , 668

SSE = 76 , 9528

-Total Sum of Squares all single numbers from table -It

· SSTotal = (7~~8~~ -78,gy)" + (82-70,94)" + (85-78 ,gy) + (72 -78 ,g47 + (75-78 , 94)2+ (74 -78 , 94)2 + (65 -78 ,94)2 + (Ed -78 , 94)2 +

(70-7~~0~~lgy)2 + (90 -70, 94)2+ 100 - 70 ,94)2+ 192-78,94)'t

(85-70.94) + 100 -70 .94) + (84~~-~~70.94) + (70-70,94)"+

(75 -78 , 94)2 + (80 -78,94) = 984 , 9444

SS interaction = SSTotal- SSA-SSB~~-S~~SE

= 904 , 9444 -523 , 6956 -382 , 5378 -76 , 9528 = 1,7582

Degrees of freedom

d f A = 2

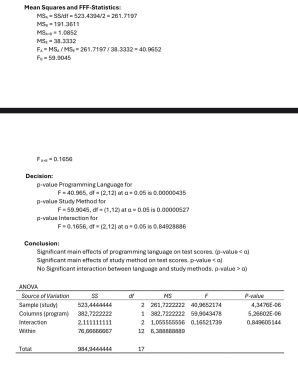
df B = 1

of interaction = 2

If within = 12

of total = 17

Mean Squared and ~~FF-~~Statistics



- -Y

Week

Probability=

Probabi~~li~~ty of a single event

Possible outcomes

I probability to shuffle an ace in a one deck card

A the chance of Dices

↳ probability of two (or more independent events

A flipping a coin

Exercise 1 Louisa Mandy Halim -2702325552 ① K .Y xh ye ~~i~~ · 3 . 481

2. 704

186 3.844

256 16 4 .096

360 25 5.184

788 36 6400

518 49 5476

664

d 1964 6889

al & 28/

ogo 100 7921

2 = 55 2 = 726 2 = 43432 = 3853 =54.276

bi = (n ([xy) - (Ex)([y))

① bo =(((y) (Ex2) -(Ex (Exy)) (n(Ex) =(Ex)) (n([x2) -(Ex))

do =~~(71~~6 x 30~~5~~) -(~~55~~x4343) b1 = (10x 4343) =~~(5~~x7~~2~~6) (10 x55) -(55)2 (rox 395) -(5~~5~~)

b: = 43430 -39930 =) 4 , 24

bo = 279 .518 -238 .86~~5~~ 3850 -3025

825 y' = bo + b , k

bo = 40.645 =) 49 . 3 y = bx + A

& 25 y =-

y-

, 24x + 49. 3

② to predict the student who studied for 7 hours

y = y ,24x + 49,3 y = 78 .98

x= 7 ->hoursgeneral trend and minimizes errors , but

The regression equation captures the

may not precisely align with each individual

y = y , 24(77 + 49 . 3data point~~.~~

③ predict after studying for 11 hours

y = y ,24x + 49. 3 y =95 ,94

> hours

x= 11 -

y = y , 24(117 + 49 . 3

Exercise 2

64 160 1096 10 .240 ~~III~~ -

Ox y <2x.Y ② y' = 2,~~7~~5(7~~0~~) -22 002 145y = 1~~70~~ 1 ~~5~~

3844 aggo -

140 3600 & 400

66 1704356 11 .220

68 1554624 10.548

2 = 32077020.520 49.390

bo =nExy -(Ex) (Ey)

nEx2~~-~~(Exp

bo = 5(49.390) -(320~~) (770~~)

5(20.520) -(370) ?

550

bo = = ) 2 ,75

200

bi = (Ey · <x2) -(Ex .Exy)

n(Ex2) -(Ex)

b11770 x 20.520) -1320x49.390)

200

bi = -22

y = bx + A

y = bo + b, x

-

~~-~~-22

y = 2 ,75x)