Statistics

Lecture 5

Feb 19-8:47 AM

Class Qz 1 (open Notes)

Consider the Sample below

5	8	8	10
6	7	5	9

clear all lists

2nd [+] 4:ClearAllLists Enter

STAT Edit 1:Edit

L1
5
8
8
10
6
7
5
9

Find

1) $\bar{x} = 7.25 \approx 7.3$

2) $S = 1.832 \approx 1.8$

} Round to 1-decimal

3) $S^2 = \frac{47}{14}$ } Reduced fraction

STAT ▶ CALC 1: 1-Var Stats

VARS 5: Statistics 3: S_x x^2

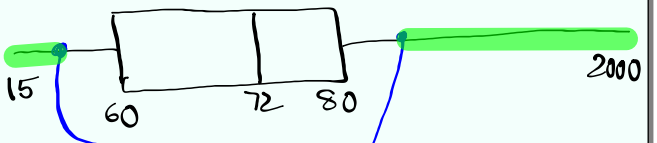
MATH 1: ▶Frac Enter

Jul 1-6:57 PM

A data set has the following 5-number Summary.

Min	Q_1	MD	Q_3	Max
15	60	72	80	2000

1) Draw box Plot



2) $IQR = Q_3 - Q_1 = 20$

3) upper fence

$Q_3 + 1.5(IQR) =$

$80 + 1.5(20) = 110$

4) Lower fence

$Q_1 - 1.5(IQR) =$

$60 - 1.5(20) = 30$

5) Discuss outliers.

$15 - 30$ & $110 - 2000$

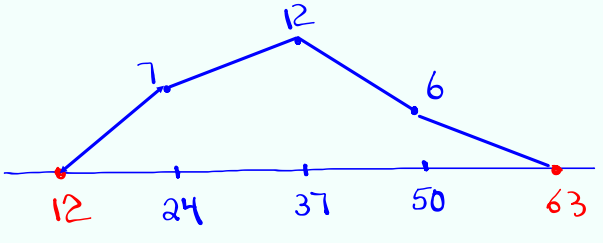
Jul 2-4:36 PM

Complete the Chart below

class limits	class MP (L1)	class F (L2)	Cum. F	Rel. F	% F
18 – 30	24	7	7	.28	28%
31 – 43	37	12	19	.48	48%
44 – 56	50	6	25	.24	24%

$n = 25$ Rel. F $= \frac{f}{25}$ CW = 13

Draw Freq. Polygon



Jul 2-4:41 PM

Find $\bar{x}$ & s of that grouped data

Clear All lists

class MP → L1

class F → L2

L1	L2
24	7
37	12
50	6

STAT → CALC → 1:1-Var STAT

with Menu: List: L1, FreqList: L2, Calculate

No Menu: L1, L2, Enter

$\bar{x} = 36.48$

$S = S_x = 9.553$

Now S^2 in reduced fraction

VARS → 5: Statistics → 3: S_x → χ^2 → MATH → 1: ▶Frac → Enter

$$S^2 = \frac{4563}{50}$$

Jul 2-4:48 PM

Consider the chart below

Class MP	Class F
17	8
30	12
43	15
56	10
69	5

1) How many classes? 5

2) Class width = 30 − 17 = 13

3) Sample Size

$n = \sum f = 50$

class MP → L1

class F → L2

$\bar{x} = 40.92$

$S \approx 15.837$

STAT → CALC → 1:1-VAR STATS

Menu: List: L1, FreqList: L2, Calculate

No Menu: L1, L2, Enter

VARS → 5: Statistics → 3: S_x

χ^2 → MATH → 1: ▶Frac → Enter

$S^2 = \frac{307242}{1225}$

Jul 2-4:59 PM

Consider the Stem plot below

1	8
2	0 3 5 8
3	0 2 2 5 5 7 8
4	2 5 6 8 8
5	0 3 8

1) $n = 20$

2) Range $= 58 - 18 = 40$

3) Midrange $= \frac{58+18}{2} = 38$

4) Estimate $S \approx \frac{\text{Range}}{4} = \frac{40}{4} = 10$

5) Store in a list, use 1-Var Stats to find $\bar{x}$ & S.

STAT → CALC
1:1-Var Stats
List: L1
FreqList: clear
Calculate

$\bar{x} = 37.15$
$S = 11.371$

6) Find S^2 in reduced fraction.

VARS → 5: Statistics → 3: S_x → x^2

MATH → 1: ▶Frac → Enter

$S^2 = \frac{49131}{380}$

Jul 2-5:11 PM

1	8
2	0 3 5 8
3	0 2 2 5 5 7 8
4	2 5 6 8 8
5	0 3 8

1) P_{12}

$L = \frac{12}{100} \cdot 20 = 2.4$

$L = 3$

P_{12} = 3rd element = 23

12% | 88%

$P_{12} = 23$

2) P_{80}

$L = \frac{80}{100} \cdot 20 = 16$

$L = 16$

$P_{80} = \frac{16\text{th element} + \text{Next one}}{2} = \frac{48+48}{2} = 48$

3) Find K such that $P_K = 40$

$K = \frac{B}{n} \cdot 100 = \frac{12}{20} \cdot 100 = 60$

$P_{60} = 40$

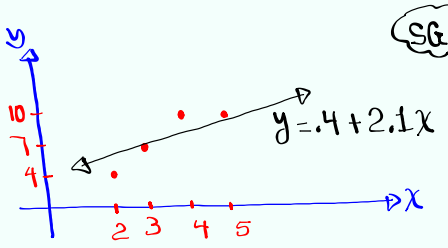
60% | 40%

$P_{60} = 40$

Jul 2-5:23 PM

SG 9

x	y
3	7
2	4
4	10
5	10



x → L1
y → L2

Use 2-VAR Stats with L1 & L2

$\sum x = 14$ $\sum y = 31$
$\sum x^2 = 54$ $\sum y^2 = 265$
$n = 4$ $\sum xy = 119$

STAT → CALC
8: LinReg(a+bx)
xlist: L1
ylist: L2
Clear
Calculate

a = .4
b = 2.1
$r^2 = .891$
$r = .944$ very close to 1

⇒ Regression line
$y = a + bx$
$y = .4 + 2.1x$

when r & r^2 missing
2nd 0 ↓↓↓···↓ ▸Diagnostic On Enter Enter

Jul 2-5:29 PM

r → Linear Correlation Coef.

$-1 \le r \le 1$

when r is close to ±1,
Linear Correlation is Significant

when r is close to 0,
Linear correlation is not Significant

when r	use
is Significant	Regression line to make Predictions
is not Significant	$\bar{y} = \frac{\sum y}{n}$

Jul 2-5:42 PM

r^2 → Coef. of determination

→ Always express as whole %.

from Last example

$r^2 = .891 \approx 89\%$

what does it tell us?

89% of Y-Values are explained by X-Values.

Jul 2-5:46 PM

Complete the Chart below

x	y	x^2	y^2	xy
1	3	1	9	3
2	5	4	25	10
3	8	9	64	24
3	6	9	36	18
5	10	25	100	50

$\Sigma x = 14$ $\quad \Sigma y = 32$

$\Sigma x^2 = 48$ $\quad \Sigma y^2 = 234$

$n = 5$ $\quad \Sigma xy = 105$

$$a = \frac{\Sigma y \, \Sigma x^2 - \Sigma x \, \Sigma xy}{n \Sigma x^2 - (\Sigma x)^2} = \frac{32 \cdot 48 - 14 \cdot 105}{5 \cdot 48 - 14^2} = \frac{66}{44} = 1.5$$

$$b = \frac{n \Sigma xy - \Sigma x \Sigma y}{n \Sigma x^2 - (\Sigma x)^2} = \frac{5 \cdot 105 - 14 \cdot 32}{5 \cdot 48 - 14^2} = \frac{77}{44} = 1.75$$

Regression line $y = a + bx$

$$y = 1.5 + 1.75x$$

Jul 2-5:59 PM

$$r = \frac{n\Sigma xy - \Sigma x \Sigma y}{\sqrt{n\Sigma x^2 - (\Sigma x)^2}\ \sqrt{n\Sigma y^2 - (\Sigma y)^2}}$$

$$= \frac{5 \cdot 105 - 14 \cdot 32}{\sqrt{5 \cdot 48 - 14^2}\ \sqrt{5 \cdot 234 - 32^2}} = \frac{77}{\sqrt{44}\ \sqrt{146}}$$

$$= \frac{77}{\sqrt{6424}} \approx .961$$

$$r^2 = .961^2 = .924 \approx 92\%$$

Jul 2-6:12 PM

$x \to L1$

$y \to L2$

L1	L2
1	3
2	5
3	8
3	6
5	10

Use 2-VAR Stats with L1 & L2

$\Sigma x = 14$ $\Sigma y = 32$

$\Sigma x^2 = 48$ $\Sigma y^2 = 234$

$n = 5$ $\Sigma xy = 105$

Use 8: LinReg(a+bx) with L1 & L2

$a = 1.5$

$b = 1.75$

$r^2 = .923$

$r = .961$

Predict y when x is 4

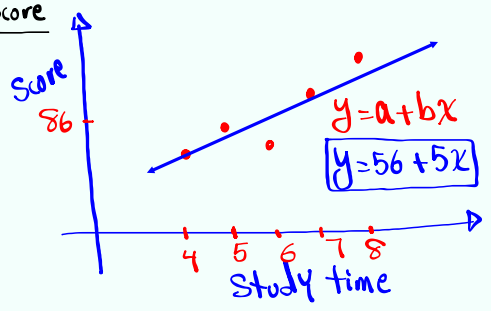
1) Assume r is Significant.
Use regression line $y = 1.5 + 1.75(4) = 8.5$

2) Assume r is not Significant.
Use $\bar{y}$ 6.4
VARS | 5: Statistics | 5: $\bar{y}$ | Enter

$$\bar{y} = \frac{\Sigma y}{n}$$

Jul 2-6:18 PM

Study time	Exam Score
5	86
7	90
8	98
6	80
4	75



Study time → x → L1

Exam Score → y → L2

$r^2 \approx 79\%$

79% of exam Scores are explained by Study time.

r is close to 1

Linear Correlation seems to be Significant.

use LinReg(a+bx) with L1 & L2

$a = 55.8 \approx 56$

$b = 5$

$r^2 = .789$

$r = .888$

Jul 2-6:27 PM

Predict exam score for 7 hrs study time

1) Assume r is Significant.

Use regression line

$y \approx 56 + 5(7) = \boxed{91}$

2) Assume r is not Significant.

Use $\bar{y}$ $\quad \bar{y} = 85.8 \approx \boxed{86}$

SG 9

Jul 2-6:36 PM

Introduction to Probability

$E \rightarrow$ Desired event or outcome

$P(E) \rightarrow$ Prob. that E happens

$$P(E) = \frac{\text{Total \# of all desired outcomes}}{\text{Total \# of all outcomes}}$$

15 Female students, 10 male students

Randomly Select one student,

$$P(\text{Select a female}) = \frac{15}{25} = \boxed{\frac{3}{5}} = \boxed{.6}$$

Jul 2-6:40 PM

Full deck of playing cards

52 Cards, 26 Red, 12 face, 4 Aces

Draw one Card,

$$P(\text{select a red Card}) = \frac{26}{52} = \frac{1}{2} = .5$$

$$P(\text{Select a face Card}) = \frac{12}{52} = \frac{3}{13} = .231$$

$$P(\text{Select a red ace}) = \frac{2}{52} = \frac{1}{26} = .038$$

12 ÷ 52 Math 1:▶Frac Enter

Math 2:▶Dec enter

Jul 2-6:44 PM

I did a Survey, and here is the result.

	Yes	NO	Total
Females	38	7	45
Males	15	40	55
Total	53	47	100

If one person is randomly Selected,

1) $P(\text{Male}) = \frac{55}{100} = \frac{11}{20} = .55$

2) $P(\text{Yes}) = \frac{53}{100} = .53$

3) $P(\text{Male and Yes}) = \frac{15}{100} = \frac{3}{20} = .15$

4) $P(\text{Male or Yes}) = \frac{93}{100} = .93$

Jul 2-6:50 PM

Class QZ 2 (open Notes)

Complete the chart below

class limits	class MP	class F
12 – 20	16	6
21 – 29	25	12
30 – 38	34	2

Find

1) $\bar{x} = 23.2$

2) $s = 5.5$

} Round to 1-dec

3) $s^2 = \frac{2916}{45}$ } Reduced fraction

Jul 2-6:59 PM