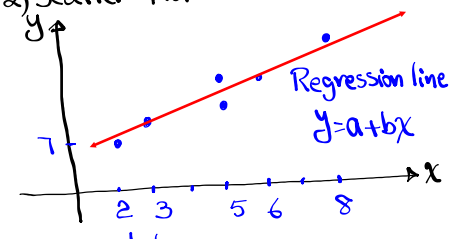
Statistics
Winter 2022
Lecture 5

Consider the chart below

x	y
2	7
3	10
5	12
5	15
6	15
8	20

1) Sample Size $n=6$

2) Scatter Plot



x → L1
y → L2

Use 2-Var stats

$\Sigma x = 29$ $\quad \Sigma y = 79$

$\Sigma x^2 = 163$ $\quad \Sigma y^2 = 1143$

$n = 6$ $\quad \Sigma xy = 429$

Now use LinReg(a+bx)

STAT CALC

8: LinReg(a+bx)

$a = 3.182$

$b = 2.066$ $\Rightarrow \boxed{y \approx 3.2 + 2.1x}$

$r^2 = .947 \Rightarrow \boxed{r^2 \approx 95\%}$

$r = .973$

How to find a & b using formulas:

$$a = \frac{\sum y \sum x^2 - \sum x \sum xy}{n\sum x^2 - (\sum x)^2} = \frac{79 \cdot 163 - 29 \cdot 429}{6 \cdot 163 - 29^2}$$

$$= \frac{436}{137} \approx \boxed{3.182}$$

$$b = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2} = \frac{6 \cdot 429 - 29 \cdot 79}{6 \cdot 163 - 29^2}$$

$$= \frac{283}{137} \approx \boxed{2.066}$$

$r \rightarrow$ Linear Correlation Coef.

$$-1 \le r \le 1$$

when r is close to 1 or -1

$\Rightarrow$ Linear Correlation is Significant

when r is close to 0

$\Rightarrow$ Linear Correlation is not Significant

In the last example $r = .973$

r is close to 1

$\Rightarrow$ Linear Correlation is Significant.

How to find r using formula:

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

$$= \frac{6 \cdot 429 - 29 \cdot 79}{\sqrt{6 \cdot 163 - 29^2}\ \sqrt{6 \cdot 1143 - 79^2}} = \frac{283}{\sqrt{137}\ \sqrt{617}}$$

$r = 283 \div (\sqrt{137} * \sqrt{617})$ Enter $\boxed{r = .973}$

$r^2 \Rightarrow$ Coef. of determination

Express r^2 in whole %.

r^2 tells us what % of Y-values are explained by x-values.

Using last example

$r^2 = .973^2 = .946729 \approx \boxed{95\%}$

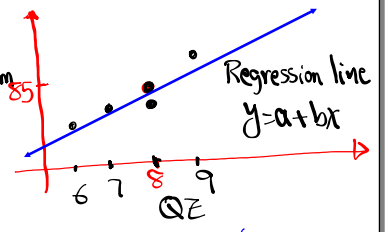
95% of Y-values are explained by x-values

5% are unexplained.

I randomly Selected 5 students, chart below shows QZ Score & Exam result.

QZ Score	Exam Result
8	85
7	80
6	70
9	100
8	80

$n=5$



QZ Score → X → L1

Exam Score → y → L2

Use LinReg(a+bx)

$y=16+9x$

$a=15.769$ $a\approx16$

$\Rightarrow$

$b=8.846$ $b\approx9$

$r^2=.848$ $r^2\approx85\%$

$r=.921$

85% of exam Scores are explained by QZ Scores

$r=.921$ → r is close to 1

⇒ Linear Correlation is Significant.

How to make predictions using Linear regression:

If r is Significant,

⇒ Plug in the x-value into $y=a+bx$, and predict y

If r is not Significant,

⇒ use $\bar{y}$ as the prediction value

$\bar{y}=\frac{\Sigma y}{n}$, VARS 5: Statistics 5: $\bar{y}$ Enter

$\bar{y}=83$

Let's Say Evie got 8 on QZ, assume r is Significant, Predict exam Score

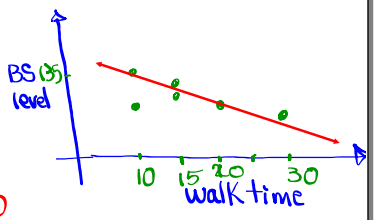
Use the regression line

$y=16+9x=16+9(8)=88$

Walk time	Blood Sugar level
10	135
15	125
15	130
10	120
20	115
30	100

1) $n = 6$

2) Scatter Plot



Clear all lists [2nd] [+] [4:clear lists] [Enter]

[2nd] [0] ↓↓↓···↓ ▸DiagnosticOn [Enter] [Enter]

walk time → X → L1

BS level → Y → L2

use LinReg(a+bx)

$a \approx 144.9$

$b \approx -1.4$

$\boxed{y = 144.9 - 1.4x}$

$a = 144.853$

$b = -1.441$

$r^2 = .763$

$r = -.874$

$r^2 \approx 76\%$

76% of my BS level are explained by my walking time.

$r = -.874$

r is close to −1

⇒ Linear Correlation is Significant.

Predict my BS level if I walk 20 minutes

a) Assume r is Significant,

use regression line

$y = 144.9 - 1.4x = 144.9 - 1.4(20) = 116.9 \approx \boxed{117}$

b) Assume r is not Significant.

⇒ use $\bar{y}$

[VARS] [5:Statistics] [5:$\bar{y}$]

[Enter] $120.83 \approx \boxed{121}$

SE 9 ✓

Intro. to Probabilities: SG 10 - 13

E → Desired event (outcome)

P(E) is the prob. that event E happens.

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

Final Ans:
1) Reduced Fraction
2) Round to 3-decimals
3) Scientific Notation

A class has 8 males & 12 Females.

If we randomly select one person,

Find the prob. that we select

1) one male

$$P(\text{Male}) = \frac{\text{\# Males}}{\text{\# People}} = \frac{8}{20} = \boxed{\frac{2}{5}} = \boxed{.4}$$

2) one female

$$P(\text{Female}) = \frac{\text{\# Females}}{\text{\# people}} = \frac{12}{20} = \boxed{\frac{3}{5}} = \boxed{.6}$$

A standard deck of playing cards has 52 cards, 26 red, 12 face, 4 Aces. Let's randomly select one card,

$P(\text{Red}) = \frac{26}{52} = \frac{1}{2} = .5$

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

12 ÷ 52 MATH 1:▶Frac Enter

MATH E:▶Dec Enter

$P(\text{Ace}) = \frac{4}{52} = \frac{1}{13}$

$E \rightarrow$ Desired event

$\bar{E} \rightarrow$ E-bar, E-complement, Not E

$P(E) + P(\bar{E}) = 1$ Complement Rule

$P(\bar{E}) = 1 - P(E)$

Suppose $P(E) = .12$

1) Convert to reduced fraction

.12 MATH 1:▶Frac Enter $P(E) = \frac{3}{25}$

2) Convert $P(E)$ to % notation

$P(E) = .12(100)\% = 12\%$

3) Find $P(\bar{E})$ $P(\bar{E}) = 1 - P(E)$

$= 1 - .12 = .88$

Some Prob. rules:

1) $0 \le P(E) \le 1$

2) Sum of all prob. is 1.

3) $P(E) = 1 \iff$ Sure event

4) $P(E) = 0 \iff$ Impossible event

5) $0 < P(E) \le .05 \iff$ Rare event

6) $P(\bar{E}) = 1 - P(E)$ Complement Rule

Suppose $P(E) = \frac{2}{15}$

1) Write in decimals. 2 ÷ 15 enter

.133

2) Write in %. $P(E) = .133(100)\% = 13.3\%$

3) Find $P(\bar{E})$

$P(\bar{E}) = 1 - P(E) = 1 - \frac{2}{15} = \frac{13}{15}$

1 − 2 ÷ 15 MATH 1:▶Frac Enter

I surveyed 100 people. I asked them for Yes, NO, or No Comment answers "Are You Vaccinated?"

	Yes	No	No Comment	Total
Males	24	10	16	50
Females	12	18	20	50
Total	36	28	36	100

If we randomly Select 1 person,

$P(\text{Male}) = \frac{50}{100} = .5$ $P(\text{Yes}) = \frac{36}{100} = .36$

$P(\text{male and Yes}) = \frac{24}{100} = .24$

$P(\text{Male or Yes}) = \frac{62}{100} = .62$

Ex:

If we randomly Select one person, what is the prob. that he/she has birthday

1) Today

$\frac{1}{365 \text{ Days}} = .003$ Rare event

2) This week

$\frac{1}{52 \text{ weeks}}$

3) This month.

$\frac{1}{12 \text{ Month}}$

SG 10✓

Addition Rule:

Keyword: OR

Single Action event

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

overlap (both)

Given $P(A)=.4$, $P(B)=.7$, $P(A \text{ and } B)=.25$

$P(\overline{A}) = 1 - P(A) = 1 - .4 = \boxed{.6}$

$P(\overline{B}) = 1 - P(B) = 1 - .7 = \boxed{.3}$

$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$

$= .4 + .7 - .25$

$= \boxed{.85}$

$P(HB) = .65$

$P(FF) = .45$

$P(HB \text{ and } FF) = .3$

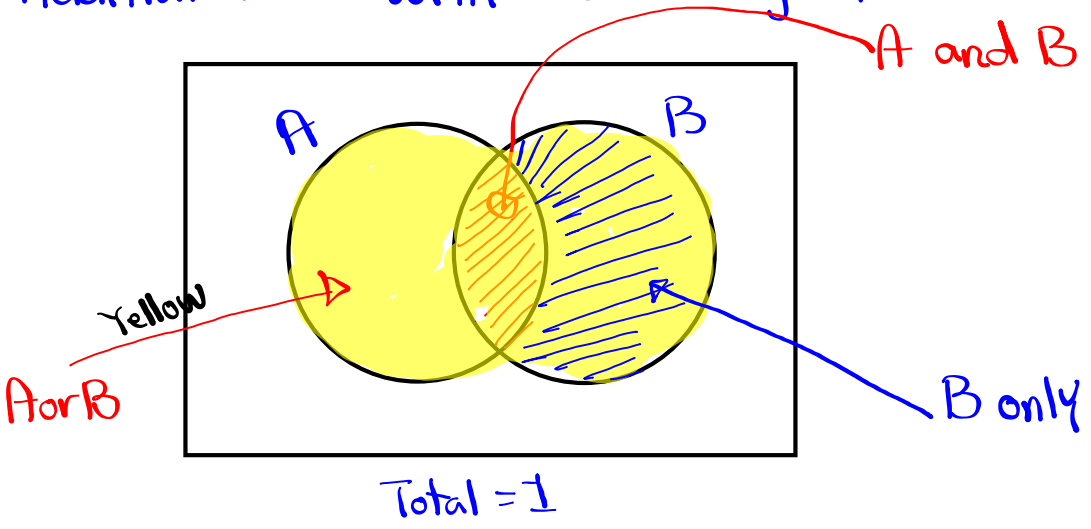
$P(\overline{HB}) = 1 - .65 = \boxed{.35}$

$P(\overline{FF}) = 1 - .45 = \boxed{.55}$

$P(HB \text{ or } FF) = P(HB) + P(FF) - P(HB \text{ and } FF)$

$= .65 + .45 - .3 = \boxed{.8}$

Addition Rule with Venn Diagram



$P(A) = .8$

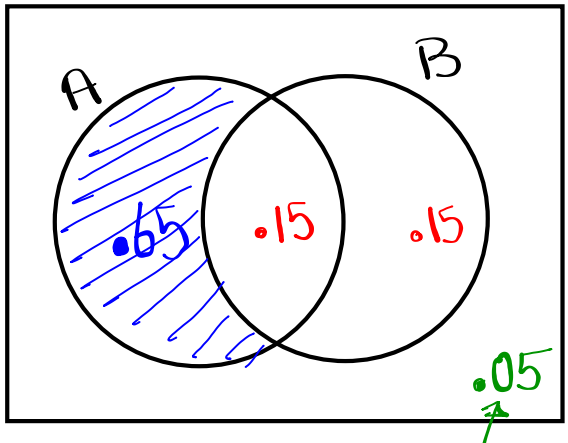
$P(B) = .3$

$P(A \text{ and } B) = .15$

overlap

$P(B \text{ only}) = .3 - .15$

$= .15$



$P(A \text{ only}) = .8 - .15 = .65$

Total = 1

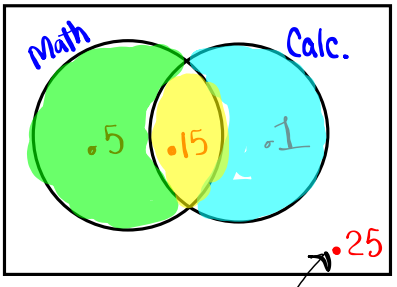
$P(\overline{A \text{ or } B}) = .05$

$P(\text{Math}) = .65$

$P(\text{Calculator}) = .25$

$P(\text{Math and Calculator}) = .15$

Construct Venn Diagram



$P(\text{Math only}) = .65 - .15$

$= .5$

$P(\text{Calc. only}) = .25 - .15 = .1$

Total $= 1$

$P(\overline{\text{Math or Calc}}) = .25$