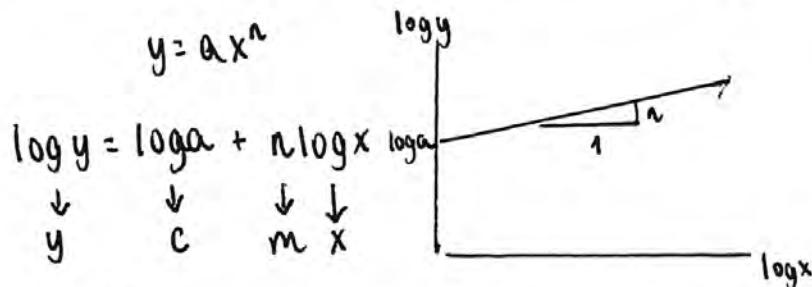


STATISTICS

1) REGRESSION, CORRELATION AND HYPOTHESIS TESTING

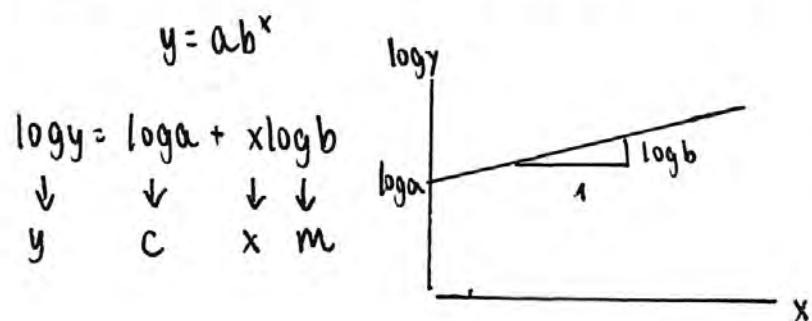
1.1) Exponential models → logarithms and coding used to examine trends in non-linear data.

$y = ax^n$ - power ($a, n \rightarrow \text{constant}$) → polynomial base unknown



graph: $\log y$ against $\log x$, straight line with gradient n and vertical intercept $\log a$.

$y = ab^x$ - exponential ($a, b \rightarrow \text{constant}$) → exponential power unknown (exponent)



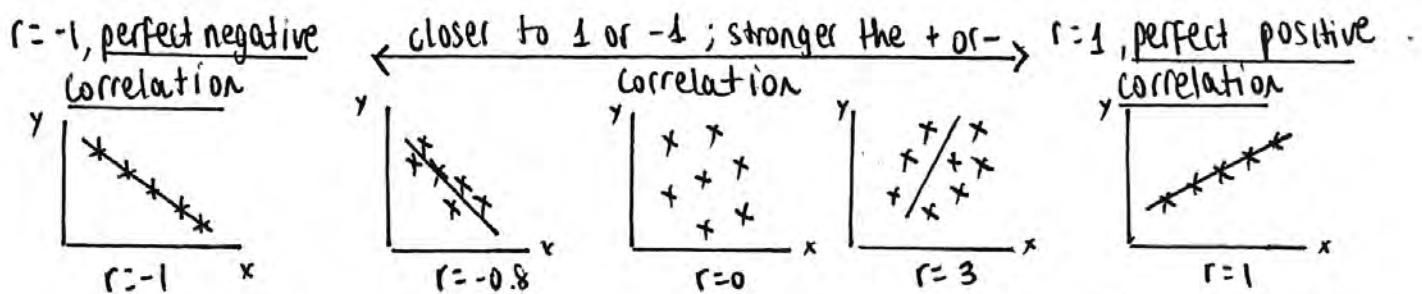
graph: $\log y$ against x , straight line with gradient of $\log b$ and vertical intercept $\log a$.

1.2 - measuring correlation

$$r = \pm 1 \rightarrow \text{straight line}$$

The product moment correlation coefficient describes the linear correlation between two variables. It can take any value between -1 and 1 .

→ PMCC measures the type and strength of , linear correlation ←



- $-0.2 < x < 0.2 \Rightarrow$ No significant correlation
- $-0.4 < x < -0.2 / 0.2 < x < 0.4 \Rightarrow$ weak correlation
- $-0.7 < x < -0.4 / 0.4 < x < 0.7 \Rightarrow$ moderate correlation
- $x < 0.7 / x > 0.7 \Rightarrow$ strong correlation
- close to ± 1 or $\pm -1 \Rightarrow$ very strong correlation

As [name of x] increases, [name of y] increases / decreases on average.
 ↴
 always say on average unless it is perfect.

Correlation and regression on calculator:

[2] Statistics → X's List 1
 Y's List 2 $\begin{bmatrix} \log x's \\ \log y's \end{bmatrix}$ ↓
 If $y = ax^n$
 (power)
 (polynomial)

[F2] CALC $\begin{bmatrix} x's \\ \log y's \end{bmatrix}$ ↓
 If $y = ab^x$
 exponential

[F6]-SET (check values are in the correct list)

2Var X List : List 1
 2Var Y List : List 2
 2Var Freq : 1

Linear Reg (a+bx)
 a: Y-intercept
 b: gradient
 r: PMCC

[EXIT] → [F3] REG → [F1] X → [F2] a+bx

1.3) Hypothesis testing for zero correlation

Hypothesis used to see if there is any correlation

→ Used to determine if the PMCC, r , for a particular sample, indicates that there is likely to be a linear relationship within the whole population.

For a one-tailed test:

- $H_0: p=0, H_1: p > 0$ → hypothesized + correlation
- $H_0: p=0, H_1: p < 0$ → hypothesized - correlation



For a two-tailed test:

- $H_0: p=0, H_1: p \neq 0$ → used to see if there is any correlation (+ or -)
- (>) halve the α (significance level)

Sample level → n , number of x's (or number of y's) → how many columns of data

Use tables (pg 191) → They show the critical value for each sample at each

one-tailed	critical value	our value	different significant level	outside critical region
	$ x > r$		H_0	Not enough evidence to suggest a +/- correlation
	$ x < r$		H_0	Enough evidence to suggest a +/- correlation

* Reject H_0 and accept H_1 , if r falls inside of the critical region, thus, is greater than the critical value

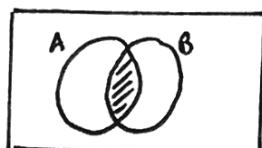
critical value $< r \rightarrow$ reject H_0 .

two tailed → Enough not enough evidence to suggest any correlation

- We measure from distance from 0 e.g. $r > |x| \rightarrow r$ bigger num → reject $H_0 \rightarrow$ inside of critical region

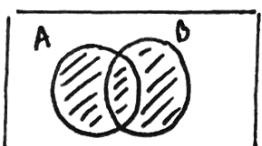
2) CONDITIONAL PROBABILITY

2.1) Set notation



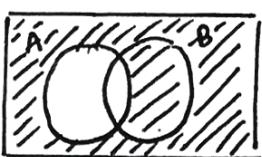
$A \cap B \rightarrow$ intersection (and)

- if A and B are independent $P(A) \times P(B) = P(A \cap B)$



$A \cup B \rightarrow$ union (or)

- if A and B are mutually exclusive, $P(A) + P(B) = P(A \cup B)$
 $\boxed{P(A \cap B) = 0}$



$A' \rightarrow$ complement (not A)

- event A and A' are always mutually exclusive

Venn diagrams → can show the number of outcomes in each event $n(R)$ or a probability. $P(R)$

$P(\emptyset) = 0 \rightarrow$ empty set

E : whole set

2.2) Conditional probability & 2.3 Venn diagrams

Probability B occurs given A $\rightarrow P(B|A)$

→ for independent events and $P(A|B) = P(A|B') = P(A)$
 $P(B|A) = P(B|A') = P(B)$

two-way tables			
	A	A'	Total
B			
B'			
Total			

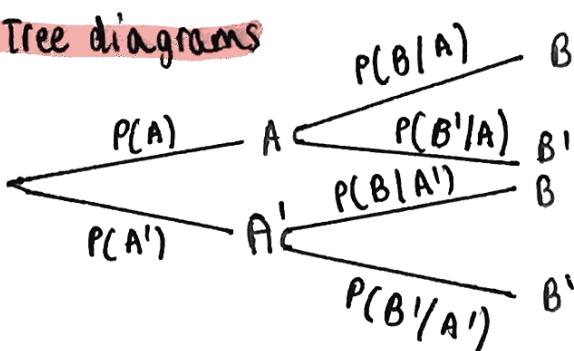
2.4) Probability formulae

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \Rightarrow \text{addition formula}$$

$$P(B|A) = \frac{P(B \cap A)}{P(A)} \therefore P(B \cap A) = P(B|A) \times P(A) \Rightarrow \text{multiplication formula}$$

*Careful brackets = $P(K|J' \cap L') = P(K|(J' \cap L'))$

2.5) Tree diagrams



conditional probabilities

→ replacement & w/o replacement