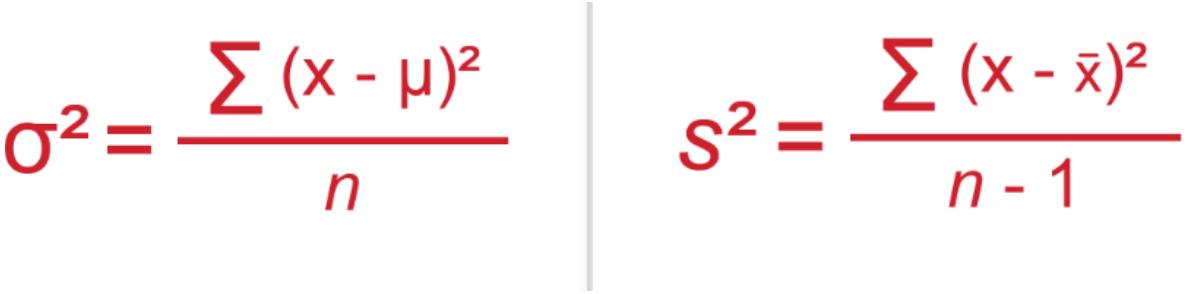
Unit	Content	Sources for help
1	Types/sources of data, methods of statistical measurement Mean, Median, Mode, Range (interquartile range e.g. 75^{th} -25^{th} Quartile percentiles) Variance – differentiation from around the mean Standard Deviation – Square root of the variance	Medium Blog
2	Unit was more R focused, but small maths topic of vectors and Matrices introduced. Vectors – Magnitude and Size + Direction Moved box 5 (or 2.5m/s) Metres to the right. Scalars - Magnitude and Size Moved 5 (or 2.5m/s) metres. Understanding how to read a matrix (2x3 means two rows, three columns) – Always Row x Column	Vectors/scalars Matrices
3	Again, R was heavily used in this unit. Reading suggested was to cover maths covered in units 1&2.	
4	Probabilities maths was covered in this unit. Discrete vs. Continuous probability Discrete = Binomial, Poisson, Hypergeometric (Card Deck, Poker) Continuous = Uniform distribution, Normal distribution There was a brief mention of Stratified Sampling, and sampling in general.	Numerical Analysis, Statistical Probability Notes Probability multiplication and addition rule video Sampling (Stratified sampling vid)
5	Calculus was introduced in the unit. With a focus on differentiation and integration. Differentiation – Multiply coefficient by the power and reduce the power by 1. Integration – Add to the power by 1, divide coefficient by the new power. Differentiation can be used for finding the rate of change, or gradient of a curve.	The essence of calculus video

	Integration can be used for area under the curve.	
6	Hypothesis Testing was topic of discussion. Analysis of Variance, or ANOVA, is a commonly used statistical method for testing a hypothesis. Put simply, ANOVA is used to compare the mean of three or more independent groups, to determine if there are any statistical differences. Assessing the **p-value** and **confidence interval**: • The p-value proves or disproves the null hypothesis based on its significance. (P-Value less than 0.05, significant at level 5%) • For the 95 % CI, if the range doesn't cross 1, it is significant. 5% is considered the minimum level of significance. A large p-value of 0.9(90%) means your results have a 90% probability of being completely random and not due to anything in your experiment.	Online article on hypothesis tests Another article linked, on parametric test
7	Unit focussed on Parametric Tests. Unit document was supplied. Assumptions to perform Parametric test; • Normality- normal distribution following • Homogeneity of variance- sample variance same as population variance • Independence – independent observations, that are randomly sampled. • Outliers – no extreme outliers	Parametric Test notes, with added comments PowerPoint 3

Population variance on the left, and sample population on the right-hand side of images below



$$\sigma^2 = \frac{\sum (x - \mu)^2}{n} \qquad s^2 = \frac{\sum (x - \bar{x})^2}{n - 1}$$

For sample variance of sequence 1,2,4,5 you would find the mean (left image), and then find the difference between each value and the mean (image right), and insert into the relevant formula to get 3.33.

Calculate the Mean

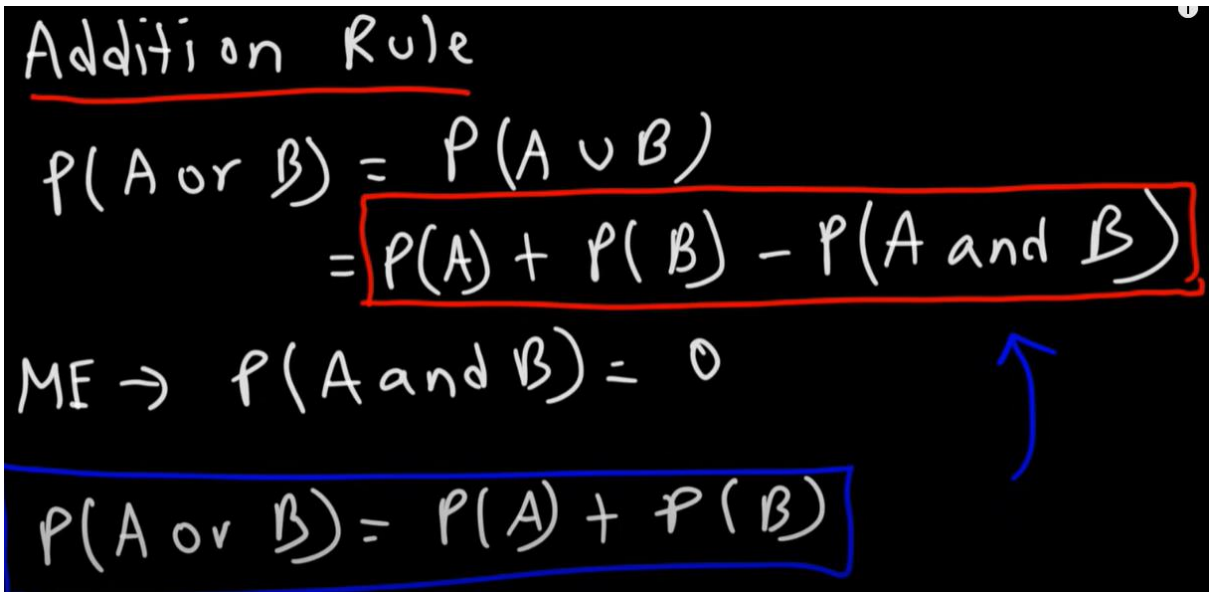
$$\bar{x} = (1+2+4+5)/4 = 12/4 = 3$$

First you calculate the mean which turns out to be 3.

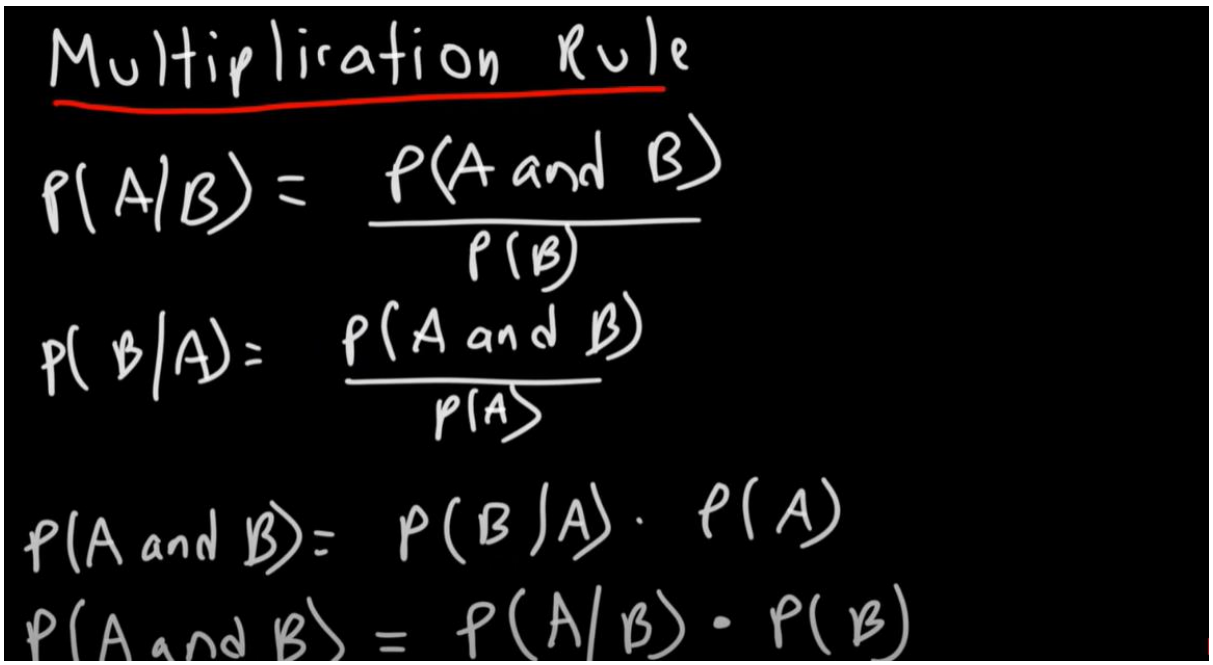
Subtract the Mean

$$s^2 = [(1-3)^2+(2-3)^2+(4-3)^2+(5-3)^2]/(4-1)$$

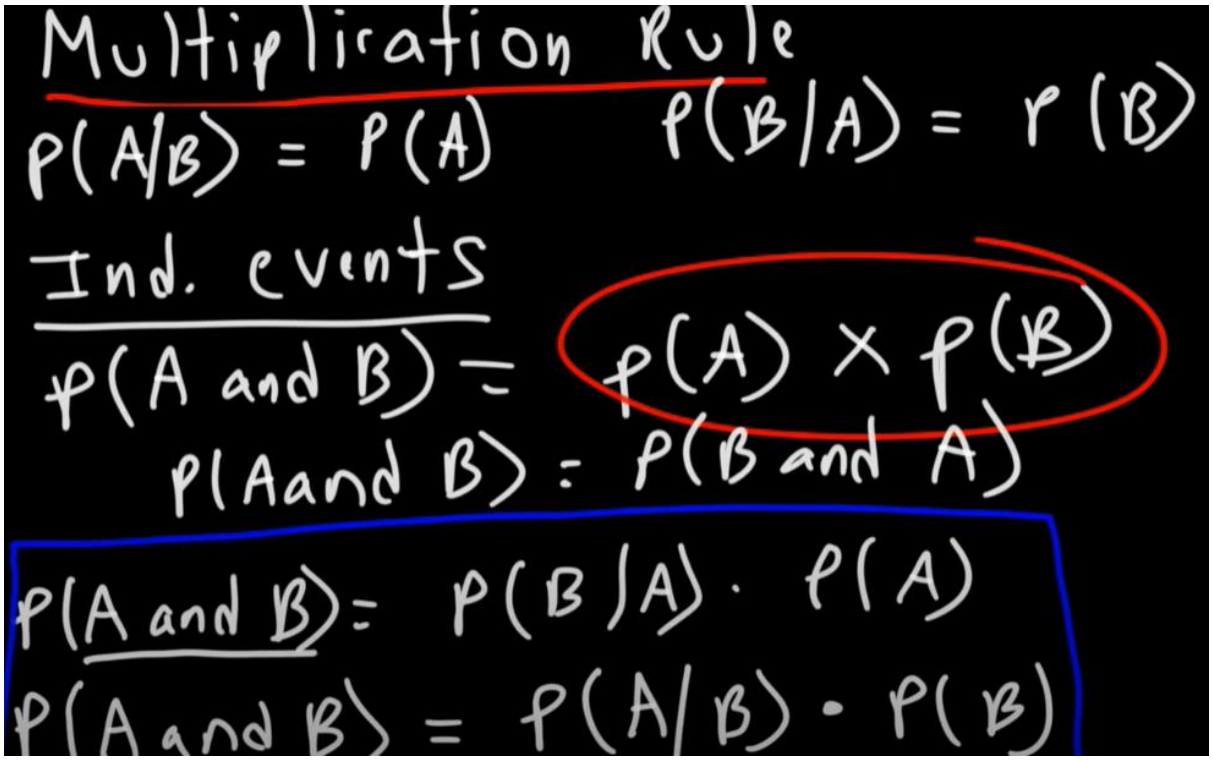
Addition rule – the blue box is if A/B are mutually exclusive, but Red formula always work (reverts to same answer as blue box).



Multiplication rule for **dependent** events…



Multiplication rule for **independent** events…



$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

z : z-score (test statistic)

$\bar{x}$: sample mean

μ : population mean

σ : population standard deviation

n : sample size

Categorical Vs Quantitative var	Quantitative var	
	Normal	Non-normal
Categorical var has two levels (e.g., disease present, disease absent)	Independent samples t-test	Mann-Whitney U test
Categorical variable has >2 levels (e.g., Muslim, Hindu, Christian, Others)	ANOVA	Kruskal Wallis test
Quantitative measurements before and after (related sample)	Paired t-test	Wilcoxon signed rank test