Common Chunks

Sums of Squares

Often used to quantify some sort of error

Use in variance and covariance formulas

Normalizing Constants

Often nightmarish, but if you look closely, you can usually ignore them

Sample size, sample size correction

N-1, n-1

Mean

Text

Description automatically generated

Notation Types

Set Notation  
Capital/lowercase Notation  
Summation Notation  
Bar Notation  
Normalizing and Sample Size Notation

Sum of Squared Errors

The sum of the squared differences between each value and the mean

Diagram, text, schematic

Description automatically generated with medium confidence

Variance and Covariance

Measures of *spread* or *dispersion*  
They are sample statistics – we can use them to estimate population parameters  
Variance is a univariate statistic  
Covariance measures an association between two variable: this is directly analogous to correlation

Text, letter

Description automatically generated

Text

Description automatically generated with medium confidence

Text

Description automatically generated with low confidence

Positive Covariance: High x-values tend to co-occur with high y-values

Negative Covariacnce: High x-values tend to occur with low y-values

Positive Covariance

Chart, scatter chart

Description automatically generated

Negative Covariance

Chart, scatter chart

Description automatically generated

*Bernoulli Distribution*

Sample space with only 2 elements:  
 True/False  
 Success/Failure  
 Presence/Absence

Realizations of a *Bernoulli process* produces *binary* outcome

It has one parameter: the probability of *success*

Binomial Distributions

Binomial process is a collection of *n* independent Bernoulli trials

Each trial must have the same probability of success  
 Binomial has 2 parametes: *n* and *p*  
 *n* is number of trials  
 *p* is the probability of success in an individual trial

Sample space has *n* + 1 elements

Poisson Distribution

Single parameter: lambda  
 Describes counts:  
 a count, or census  
 sample space is 0 to infinity  
 Infinite sample space

*Theoretical* distribution

Often appropriate for things that occur *randomly* but at a certain *constant rate*