STATS ACADEMY

GCSE Statistics (9-1) Formulae

Sampling

$$\textbf{Stratified Sample} = \frac{strata}{total} \times sample\ size$$

Capture – Recapture:

$$\frac{First\ Capture}{Total\ (N)} = \frac{Tagged}{Second\ Capture}$$

Representing Data

Angle in a Pie Chart

$$\frac{360}{total\ frequency} \times frequency\ for\ each\ class$$

Comparative Pie Chart:

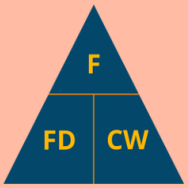
$$\frac{r_1}{r_2} = \sqrt{\frac{F_1}{F_2}} \qquad r_2 = r_1 \times \sqrt{\frac{F_1}{F_2}}$$

r = radius F = Frequency

Histograms

$$\textbf{Frequency Density} = \frac{Frequency}{Class\ Width}$$

$$\textbf{Class Width} = \frac{Frequency}{Frequency\ Denisty}$$



$$\textbf{Frequency} = Frequency\ Density \times Class\ Width$$

Averages

$$\textbf{Mode} = most\ common$$
$$\textbf{Modal Class} = Class\ with\ highest\ frequency$$

$$\textbf{Median (discrete data)} = \frac{1}{2}(n+1)th\ value$$
$$\textbf{Median (grouped data)} = \frac{1}{2}nth\ value$$

Linear Interpolation

$$\textbf{Estimated Median} = L + \frac{\frac{n}{2} - F}{f} \times w$$

- L = lower bound of class containing median
- n = total number of values
- F = cumulative frequency of intervals before one containing median
- f = frequency of median class interval
- w = width of median class interval

Mean, $\bar{x}$ (discrete data):

$$\bar{x} = \frac{\sum x}{n}$$

$\sum x$ = sum of all values n = number of values

Mean, $\bar{x}$ (frequency table – not grouped):

$$\bar{x} = \frac{\sum fx}{\sum f}$$

$\sum fx$ = frequency × value in first column
$\sum f$ = total frequency

Mean, $\bar{x}$ (grouped frequency table):

$$\frac{\sum(f \times midpoint)}{\sum f}$$

$$\textbf{Weighted Mean} = \frac{\sum(weight \times value)}{\sum weights}$$

$$\textbf{Geometric Mean} = \sqrt[n]{value_1 \times value_2 \times \cdots \times value_n}$$

Foundation Tier Formula

Higher Tier Formula

www.statsacademy.co.uk

STATS ACADEMY

GCSE Statistics (9-1) Formulae

Measures of Dispersion

$$\textbf{Range} = largest\ value\ -\ smallest\ value$$

Interquartile Range (IQR)

$$Upper\ Quartile\ (UQ) - Lower\ Quartile\ (LQ)$$

$$\textbf{LQ (discrete data)} = \frac{1}{4}(n+1)th\ value$$

$$\textbf{LQ (grouped data)} = \frac{1}{4}nth\ value$$

$$\textbf{UQ (discrete data)} = \frac{3}{4}(n+1)th\ value$$

$$\textbf{UQ (grouped data)} = \frac{3}{4}nth\ value$$

$$\textbf{Decile} = \frac{n+1}{10}th\ value$$

$$\textbf{Percentile} = \frac{n+1}{100}th\ value$$

Interdecile Range

$difference\ between\ 2\ deciles$

Interpercentile Range

$difference\ between\ 2\ percentiles$

Standard Deviation, σ

Discrete Data:

The formulae for discrete data are given in exams

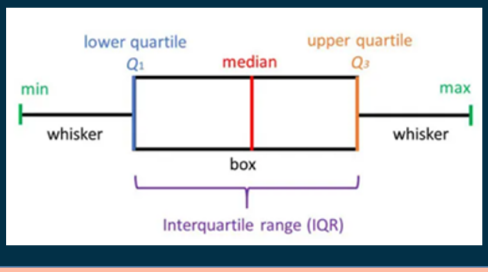
$$\sigma = \sqrt{\frac{1}{n}\sum(x-\bar{x})^2} \quad or\ \sigma = \sqrt{\frac{\sum x^2}{n} - (\frac{\sum x}{n})^2}$$

Frequency Table or Grouped Data:

$$\sigma = \sqrt{\frac{\sum f(x-\bar{x})^2}{\sum f}} \quad or\ \sigma = \sqrt{\frac{\sum fx^2}{\sum f} - (\frac{\sum fx}{\sum f})^2}$$

$For\ grouped\ data\ use\ the\ midpoint\ for\ x$

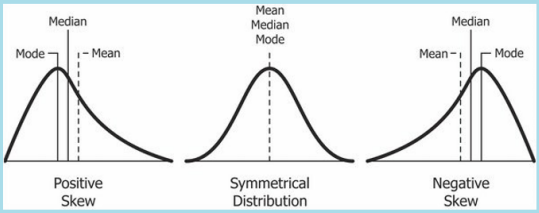
Box Plots



Outliers

$$values > UQ + (1.5 \times IQR) \quad or \quad < LQ - (1.5 \times IQR)$$

Skewness (by inspection)



Skewness (by calculation)

This formula is given in the exam

$$\textbf{Skewness} = \frac{3(mean - median)}{standard\ deviation}$$

Foundation Tier Formula

Higher Tier Formula

www.statsacademy.co.uk

GCSE Statistics (9-1) Formulae

Time Series

Moving Averages

Calculate mean of values

Seasonal Variation

$$Actual\ Value - Trend\ Value$$

Estimated Mean Seasonal Variation

Mean of all the seasonal variations for that season

Predicted Value

$$Trend\ Line\ Value\ (from\ graph) + EMSV$$

Scatter Diagrams

Mean Point $(\bar{x}, \bar{y})$

$$(Mean\ of\ x\ values, Mean\ of\ y\ values)$$

Equation of Line of Best Fit

$$y = ax + b$$

Gradient, a

$$a = \frac{y_2 - y_1}{x_2 - x_1}$$

Y − intercept, b

$$b = y_1 - ax_1$$

Spearman's Rank Correlation Coefficient (SRCC)

$$SRCC, r_s = 1 - \frac{6\sum d^2}{n(n^2 - 1)}$$

This formula is given in the exam

d = difference between ranks n = number of values

Index Numbers

$$Index\ Number = \frac{Price}{Base\ Year\ Price} \times 100$$

Weighted Index Number

$$\frac{\sum(index\ number \times weight)}{\sum weights}$$

Chain Base Index Numbers

$$\frac{price}{last\ year's\ price} \times 100$$

Rates of Change

Crude Rate

$$\frac{number\ of\ births/deaths}{total\ population} \times 1000$$

This formula is given in the exam

Standard Population

$$\frac{number\ in\ age\ group}{total\ population} \times 1000$$

Standardised Rate

$$\frac{Crude\ Rate}{1000} \times Standard\ Population$$

Foundation Tier Formula

Higher Tier Formula

www.statsacademy.co.uk

GCSE Statistics (9-1) Formulae

Probability

$$P(event) = \frac{Number\ of\ successful\ outcomes}{Total\ number\ of\ outcomes}$$

Expected Frequency of Event A

$$P(A) \times number\ of\ trials$$

Estimated Probability

$$\frac{Number\ of\ trials\ with\ successful\ outcomes}{Total\ number\ of\ trials}$$

Risk

$$\frac{Number\ of\ trials\ in\ which\ event\ happens}{Total\ number\ of\ trials}$$

Relative Risk

$$\frac{Risk\ for\ those\ in\ the\ group}{Risk\ for\ those\ not\ in\ the\ group}$$

Mutually Exclusive Events

$$P(A\ or\ B) = P(A) + P(B)$$

Mutually Exclusive, Exhaustive Events

$$P(A) + P(not\ A) = 1$$

$$P(not\ A) = 1 - P(A)$$

General Addition Law

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

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Independent Events

$$P(A\ and\ B) = P(A) \times P(B)$$

$$P(A\ and\ B\ and\ C) = P(A) \times P(B) \times P(C)$$

$$P(at\ least\ 1) = 1 - P(none)$$

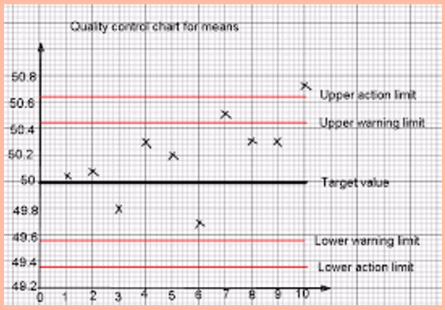
Conditional Probability

$$P(B|A) = \frac{P(A\ and\ B)}{P(B)}$$

$$P(A\ and\ B) = P(B|A) \times P(A)$$

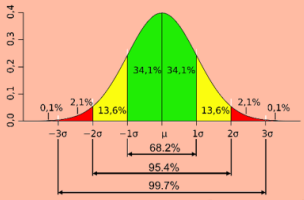
Probability Distributions

Quality Assurance Charts



$\mu + 3\sigma$
$\mu + 2\sigma$
μ
$\mu - 2\sigma$
$\mu - 3\sigma$

Normal Distribution



Conditions for Normal Distribution:

1. *The data is continuous (heights, weights, time)*
2. *The distribution is symmetrical.*
3. *Mode, median and mean are all approximately equal.*

Binomial Distribution

1. *Fixed number of trials (n)*
2. *Each trial has 2 outcomes, success (p) or failure*
3. *All the trials are independent of each other*
4. *Probability of success is constant*

$$Standardised\ Score = \frac{Score - Mean}{Standard\ Deviation}$$

Foundation Tier Formula

Higher Tier Formula

www.statsacademy.co.uk