FORMULA STATISTICS

<table>
<tr><td>Formula 1</td><td>Class Boundary
<table>
<tr><td>Mutually Exclusive Classification</td><td>UCB = UCL and LCB = LCL</td></tr>
<tr><td>Mutually Inclusive Classification</td><td>UCB = UCL + 0.5 and LCB = LCL - 0.5</td></tr>
</table></td></tr>
<tr><td>Formula 2</td><td>Mid-Point / Class Mark of Class Interval: $\frac{LCL+UCL}{2}$ or $\frac{LCB+UCB}{2}$</td></tr>
<tr><td>Formula 3</td><td>Class Length / Width of Class / Size of Class: $UCB-LCB$</td></tr>
<tr><td>Formula 4</td><td>Frequency Density of a Class: $\frac{\text{Frequency of the class}}{\text{Class length of the class}}$</td></tr>
<tr><td>Formula 5</td><td>Relative Frequency: $\frac{\text{Frequency of the class}}{\text{Total Frequency of distribution}}$
Percentage Frequency: $\frac{\text{Frequency of the class}}{\text{Total Frequency of distribution}}\times 100$</td></tr>
<tr><td>Formula 6</td><td>AM of Discrete Distribution/Series: $\bar{x}=\frac{x_1+x_2+x_3+...+x_n}{n}$ in short $\bar{x}=\frac{\sum x}{n}$</td></tr>
<tr><td>Formula 7</td><td>AM of Frequency Distribution: $\bar{x}=\frac{\sum fx}{N}$
<table>
<tr><td>In case of ungrouped distribution</td><td>x = individual value</td></tr>
<tr><td>In case of grouped frequency distribution</td><td>x = mid-point of class interval</td></tr>
</table></td></tr>
<tr><td>Formula 8</td><td>AM using assumed mean / step deviation method
$\bar{x}=A+\frac{\sum fd}{N}\times C$ where $d=\frac{x-A}{C}$, A is assumed mean, C is class length</td></tr>
<tr><td>Formula 9</td><td>The algebraic sum of deviations of a set of observations from their AM is zero
$\sum(x-\bar{x})=0$</td></tr>
<tr><td>Formula 10</td><td>Combined AM: $\bar{x}_c=\frac{n_1\bar{x}_1+n_2\bar{x}_2}{n_1+n_2}$</td></tr>
<tr><td>Formula 11</td><td>Median in case of discrete distribution
<table>
<tr><td>If number of observations are odd</td><td>Median is middle term</td></tr>
</table></td></tr>
</table>

<table>
<tr><td></td><td>
<table>
<tr><td>If number of observations are even</td><td>AM of two middle terms</td></tr>
</table>
Same formula is used for ungrouped frequency distribution
</td></tr>
<tr><td>Formula 12</td><td>
Median in case of grouped frequency distribution
<table>
<tr><td>Step 1</td><td>Prepare a less than type cumulative frequency distribution</td></tr>
<tr><td>Step 2</td><td>Calculate $\frac{N}{2}$ and check between which class boundaries it falls and call it as Median Class</td></tr>
<tr><td>Step 3</td><td>
<table>
<tr><td>l_1</td><td>N_u</td><td>N_l</td><td>C</td></tr>
<tr><td>LCB of Median Class</td><td>Cum Freq. of Median Class</td><td>Cum. Freq. of Pre-Median Class</td><td>Class length of Median Class</td></tr>
</table>
</td></tr>
<tr><td>Step 4</td><td>Appy Formula
$Me = l_1 + \left(\frac{\frac{N}{2} - N_l}{N_u - N_l}\right) \times C$</td></tr>
</table>
</td></tr>
<tr><td>Formula 13</td><td>For a set of observations, the sum of absolute deviations is minimum, when the deviations are taken from the median. $\sum(x-\bar{x})=0$ is minimum</td></tr>
<tr><td>Formula 14</td><td>
Quartiles in case of discrete observations:
<table>
<tr><th>First Quartile</th><th>Second Quartile</th><th>Third Quartile</th></tr>
<tr><td>$Q_1 = \left((n+1) \times \frac{1}{4}\right)^{th}$ term</td><td>$Q_2 = \left((n+1) \times \frac{2}{4}\right)^{th}$ term</td><td>$Q_3 = \left((n+1) \times \frac{3}{4}\right)^{th}$ term</td></tr>
</table>
Note: above formula gives the term. Final value to be calculated based on the term
</td></tr>
<tr><td>Formula 15</td><td>
Deciles in case of discrete observations:
<table>
<tr><th>First Decile</th><th>Second Decile</th><th>Ninth Decile</th></tr>
<tr><td>$D_1 = \left((n+1) \times \frac{1}{10}\right)^{th}$ term</td><td>$D_2 = \left((n+1) \times \frac{2}{10}\right)^{th}$ term</td><td>$D_9 = \left((n+1) \times \frac{9}{10}\right)^{th}$ term</td></tr>
</table>
Note: above formula gives the term. Final value to be calculated based on the term
</td></tr>
<tr><td>Formula 16</td><td>
Percentiles in case of discrete observations:
<table>
<tr><th>First Percentile</th><th>Second Percentile</th><th>99^{th} Percentile</th></tr>
<tr><td>$P_1 = \left((n+1) \times \frac{1}{100}\right)^{th}$ term</td><td>$P_2 = \left((n+1) \times \frac{2}{100}\right)^{th}$ term</td><td>$P_{99} = \left((n+1) \times \frac{99}{100}\right)^{th}$ term</td></tr>
</table>
Note: above formula gives the term. Final value to be calculated based on the term
</td></tr>
</table>

Formula 17

Quartiles in case of Grouped Frequency Distribution: Steps are like median with few modifications.

1^{st} Quartile	3^{rd} Quartile
Find Q_1 class using $\frac{N}{4}$	Find Q_3 class using $\frac{3N}{4}$
$Q_1 = l_1 + \left(\frac{\frac{N}{4} - N_l}{N_u - N_l}\right) \times C$	$Q_3 = l_1 + \left(\frac{\frac{3N}{4} - N_l}{N_u - N_l}\right) \times C$

Formula 18

Deciles in case of Grouped Frequency Distribution: Steps are like median with few modifications.

1^{st} Decile	9^{th} Decile
Find D_1 class using $\frac{N}{10}$	Find D_9 class using $\frac{9N}{10}$
$D_1 = l_1 + \left(\frac{\frac{N}{10} - N_l}{N_u - N_l}\right) \times C$	$D_9 = l_1 + \left(\frac{\frac{9N}{10} - N_l}{N_u - N_l}\right) \times C$

Formula 19

Percentiles in case of Grouped Frequency Distribution: Steps are like median with few modifications.

1^{st} Percentile	99^{th} Percentile
Find P_1 class using $\frac{N}{100}$	Find P_{99} class using $\frac{99N}{10}$
$P_1 = l_1 + \left(\frac{\frac{N}{100} - N_l}{N_u - N_l}\right) \times C$	$P_{99} = l_1 + \left(\frac{\frac{99N}{10} - N_l}{N_u - N_l}\right) \times C$

Formula 20

Mode in case of discrete observation: observation repeating for maximum no. of times or observation with highest frequency

Note: There can be multiple modes also. If all observations are having same frequency, then there is no mode.

Formula 21

Mode in case of grouped frequency distribution:

Find Modal Class (Class with highest frequency) then apply below formula

$$Mo = l_1 + \left(\frac{f_0 - f_{-1}}{2f_0 - f_{-1} - f_1}\right) \times C$$

where, l_1 = LCB of modal class f_0 = frequency of modal class, f_{-1} = frequency of pre-modal class, f_1 = frequency of post modal class, C = class length of modal class

<table>
<tr><td>Formula 22</td><td>Relationship between Mean, Median and Mode in case of Symmetrical Distribution:
Mean = Median = Mode</td></tr>
<tr><td>Formula 23</td><td>Relationship between Mean, Median and Mode in case of moderately skewed distribution: Mean – Mode = 3 (Mean – Median)
Mode = 3 Median – 2 Mean</td></tr>
<tr><td>Formula 24</td><td>Geometric Mean in case of discrete positive observations:
$G=\left(x_1 \times x_2 \times \ldots \times x_n\right)^{1/n}$</td></tr>
<tr><td>Formula 25</td><td>Geometric Mean in case of frequency distribution:
$G=\left(x_1^{f_1} \times x_2^{f_2} \times \ldots \times x_n^{f_n}\right)^{1/N}$</td></tr>
<tr><td>Formula 26</td><td>Harmonic Mean in case of discrete observations: $H=\frac{n}{\Sigma(\frac{1}{x})}$</td></tr>
<tr><td>Formula 27</td><td>Harmonic Mean in case of frequency distribution: $H=\frac{N}{\Sigma(\frac{f}{x})}$</td></tr>
<tr><td>Formula 28</td><td>Combined HM $=\frac{n_1+n_2}{\frac{n_1}{H_1}+\frac{n_2}{H_2}}$</td></tr>
<tr><td>Formula 29</td><td>Relationship between AM, GM and HM
<table>
<tr><th>Situation</th><th>Relationship</th></tr>
<tr><td>When all the observations are identical / same</td><td>$AM=GM=HM$</td></tr>
<tr><td>When all the observations are distinct / different</td><td>$AM>GM>HM$</td></tr>
<tr><td>In General</td><td>$AM\geq GM\geq HM$</td></tr>
</table></td></tr>
<tr><td>Formula 30</td><td>Range in case of discrete observations: L – S
where L = Largest Observation, S = Smallest Observation</td></tr>
<tr><td>Formula 31</td><td>Range in case of Grouped Frequency Distribution: L – S
L = UCB of last class interval, S = LCB of first-class interval</td></tr>
<tr><td>Formula 32</td><td>Coefficient of Range $\frac{L-S}{L+S}\times 100$</td></tr>
<tr><td>Formula 33</td><td>Mean Deviation in case of discrete observations
$MD_A=\frac{1}{n}\Sigma\lvert x-A\rvert$ where A is any appropriate central tendency (as given)</td></tr>
<tr><td>Formula 34</td><td>Mean Deviation (in case of grouped frequency distributions)
$MD_A=\frac{1}{N}\Sigma f\lvert x-A\rvert$ where A is any appropriate central tendency (as given)</td></tr>
<tr><td>Formula 35</td><td>Coefficient of Mean Deviation: $\frac{\text{Mean Deviation about A}}{A}\times 100$</td></tr>
</table>

Formula 36	Standard Deviation in case of discrete observations: $\sigma_x = SD_x = \sqrt{\frac{\sum(x-\bar{x})^2}{n}}$ or shorter formula $\sigma_x = SD_x = \sqrt{\frac{\sum x^2}{n} - (\bar{x})^2}$
Formula 37	Standard Deviation in case of grouped frequency observations $\sigma_x = SD_x = \sqrt{\frac{\sum f(x-\bar{x})^2}{N}}$ or shorter formula $\sigma_x = SD_x = \sqrt{\frac{\sum fx^2}{N} - (\bar{x})^2}$
Formula 38	Coefficient of Variation: $\frac{SD_x}{\bar{x}} \times 100$
Formula 39	If there are only two observations, then SD is half of range $SD = \frac{\lvert a-b \rvert}{2}$
Formula 40	Standard Deviation of first n natural numbers: $s = \sqrt{\frac{n^2-1}{12}}$
Formula 41	Combined SD: $SD_c = \sqrt{\frac{n_1 s_1{}^2 + n_2 s_2{}^2 + n_1 d_1{}^2 + n_2 d_2{}^2}{n_1 + n_2}}$ $d_1 = \bar{x}_c - \bar{x}_1$ and $d_2 = \bar{x}_c - \bar{x}_2$
Formula 42	If all the observations are constant, then SD/ MD/ Range is ZERO
Formula 43	Change of Origin and Scale: No effect of change of origin but affected by change of scale in the magnitude (ignore sign) $SD_y = \lvert b \rvert SD_x$ Note: same thing will apply to all the measures of dispersion
Formula 44	Quartile Deviation: $QD_x = \frac{Q_3 - Q_1}{2}$
Formula 45	Coefficient of Quartile Deviation: $\frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100$
Formula 46	Relationship between SD, MD and QD $4SD = 5MD = 6QD$ or $SD:MD:QD = 15:12:10$
Formula 47	Basic Formula of Probability: $P(A) = \frac{\text{No. of favorable events to A}}{\text{Total no. of events}}$
Formula 48	Odds in favour of Event A: $\frac{\text{no. of favorable events}}{\text{no. of unfavorable events}}$
Formula 49	Odds against an Event A: $\frac{\text{no. of unfavorable events}}{\text{no. of favorable events}}$
Formula 50	Number of total outcomes of a random experiment: If an experiment results in p outcomes and if it is repeated q times, then Total number of outcomes is p^q
Formula 51	Relative Frequency Probability $\frac{\text{no. of times the event occurred during experimental trials}}{\text{total no. of trials}} = \frac{f_A}{n}$

<table>
<tr><td>Formula 52</td><td>Set Based Probability: $P(A)=\frac{\text{no.of sample points in A}}{\text{no.of sample points in S}}=\frac{n(A)}{n(S)}$
here A is Event Set and S is Sample Space</td></tr>
<tr><td>Formula 53</td><td>Addition Theorem 1: In case of two mutually exclusive events A and B
$P(A \cup B) = P(A+B) = P(A \text{ or } B) = P(A) + P(B)$</td></tr>
<tr><td>Formula 54</td><td>Addition Theorem 2: In case of two or more mutually exclusive events
$P(A_1 \cup A_2 \cup A_3 \cup \ldots) = P(A_1) + P(A_2) + P(A_3) + \ldots$</td></tr>
<tr><td>Formula 55</td><td>Addition Theorem 3: For any two events
$P(A \cup B) = P(A) + P(B) - P(A \cap B)$</td></tr>
<tr><td>Formula 56</td><td>Addition Theorem 4: In case of any three events
$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(A \cap C) + P(A \cap B \cap C)$</td></tr>
<tr><td>Formula 57</td><td>Conditional Probability of Event B when Event A is already occurred
$P(B/A)=\frac{P(B \cap A)}{P(A)}$ provided $P(A) \neq 0$</td></tr>
<tr><td>Formula 58</td><td>Conditional Probability of Event A when Event B is already occurred
$P(A/B)=\frac{P(B \cap A)}{P(B)}$ provided $P(B) \neq 0$</td></tr>
<tr><td>Formula 59</td><td>Compound Theorem: In case of two dependent events
$P(A \cap B) = P(B) \times P(A/B)$ or $P(A \cap B) = P(A) \times P(B/A)$</td></tr>
<tr><td>Formula 60</td><td>Compound Theorem: In case of two independent events
$P(A \cap B) = P(A) \times P(B)$</td></tr>
<tr><td>Formula 61</td><td>Expected value of a Probability Distribution: $E(x)=\sum p_i x_i$
Also, $E(x)=\mu$ (here μ means mean of probability distribution)</td></tr>
<tr><td>Formula 62</td><td>Variance of Probability Distribution: $V(x)=E(x-\mu)^2=E(x^2)-[E(x)]^2$</td></tr>
<tr><td>Formula 63</td><td>Probability Mass Function in case of Binomial Distribution:
$f(x)=P(X=x)={}^{n}C_x p^x q^{n-x}$</td></tr>
<tr><td>Formula 64</td><td>Mean of Binomial Distribution: $\mu=np$
Variance of Binomial Distribution: $\sigma^2=npq$</td></tr>
<tr><td>Formula 65</td><td>Mode in case of Binomial Distribution:
<table>
<tr><td>Step 1</td><td>Calculate $(n+1)p$</td></tr>
<tr><td>Step 2A</td><td>If $(n+1)p$ is an integer, there will be two modes:
$\mu_0=(n+1)p$ & $[(n+1)p-1]$</td></tr>
<tr><td>Step 2B</td><td>If $(n+1)p$ is a non-integer, there will be only one mode:
μ_0 = largest integer contained in $(n+1)p$</td></tr>
</table></td></tr>
<tr><td>Formula 66</td><td>Probability Mass Function in case of Poisson Distribution:
$f(x)=P(X=x)=\frac{e^{-m}m^x}{x!}$</td></tr>
</table>

<table>
<tr><td>Formula 67</td><td>Mean of Poisson Distribution: $\mu = m$
Variance of Poisson Distribution: $\sigma^2 = m$
SD of Poisson Distribution: $\sigma = \sqrt{m}$</td></tr>
<tr><td>Formula 68</td><td>Mode in case of Poisson Distribution:
<table>
<tr><td>If m is an integer</td><td>there will be two modes: $\mu_0 = m \& m-1$</td></tr>
<tr><td>If m is a non-integer</td><td>there will be only one mode: largest integer contained in m</td></tr>
</table></td></tr>
<tr><td>Formula 69</td><td>Probability Density Function in case of Normal Distribution
$f(x) = \frac{1}{\sigma\sqrt{2\pi}}.e^{\left(\frac{x-\mu}{\sigma}\right)^2 \frac{1}{2}}$</td></tr>
<tr><td>Formula 70</td><td>Mean Deviation in case of Normal Distribution: $MD = 0.8\sigma$</td></tr>
<tr><td>Formula 71</td><td>Quartiles in case of Normal Distribution: $Q_1 = \mu - 0.675\sigma$ & $Q_3 = \mu + 0.675\sigma$</td></tr>
<tr><td>Formula 72</td><td>Quartile Deviation in case of Normal Distribution: $QD = 0.675\sigma$</td></tr>
<tr><td>Formula 73</td><td>Points of Inflex of Normal Curve: $\mu - \sigma$ & $\mu + \sigma$</td></tr>
<tr><td>Formula 74</td><td>In case of Normal Distribution, Ratio between QD: MD: SD = 10:12:15</td></tr>
<tr><td>Formula 75</td><td>Conditions of Standard Normal Distribution: Mean = 0, SD = 1</td></tr>
<tr><td>Formula 76</td><td>Z Score: $Z = \frac{(x-\mu)}{\sigma}$</td></tr>
<tr><td>Formula 77</td><td>Area under Normal Curve (Popular Intervals)
<table>
<tr><th>From</th><th>To</th><th>Area under Normal Curve Probability</th></tr>
<tr><td>μ</td><td>$\mu+\sigma$</td><td>34.135%</td></tr>
<tr><td>$\mu+\sigma$</td><td>$\mu+2\sigma$</td><td>13.59%</td></tr>
<tr><td>$\mu+2\sigma$</td><td>$\mu+3\sigma$</td><td>2.14%</td></tr>
<tr><td>$\mu+3\sigma$</td><td>$+\infty$</td><td>0.135%</td></tr>
</table></td></tr>
<tr><td>Formula 78</td><td>For a $p \times q$ bivariate frequency distribution:
<table>
<tr><td>Number of cells</td><td>pq</td></tr>
<tr><td>Number of marginal distributions</td><td>2</td></tr>
<tr><td>Number of conditional distributions</td><td>p+q</td></tr>
</table></td></tr>
<tr><td>Formula 79</td><td>Karl Pearson’s Product Moment Correlation Coefficient:
$r_{xy} = \frac{Cov(x,y)}{(\sigma_x \times \sigma_y)}$</td></tr>
</table>

Formula 80	Covariance between two variables: $\text{Cov}(x,y)=\frac{\Sigma(x-\bar{x})(y-\bar{y})}{n}$ or $\frac{\Sigma xy}{n}-\bar{x}.\bar{y}$
Formula 81	Spearman's Rank Correlation Coefficient: $r_R=1-\frac{6\Sigma d^2}{n(n^2-1)}$ here d means difference in ranks of both variables
Formula 82	Spearman's Rank Correlation Coefficient (in case of tied values) $r_R=1-\frac{6\left(\Sigma d^2+A\right)}{n(n^2-1)}$ here A is adjustment value $A=\frac{\Sigma\left(t^3-t\right)}{12}$ where t = tie length (calculate t value for each of the ties)
Formula 83	Coefficient of Concurrent Deviations $r_c=\pm\sqrt{\pm\left(\frac{2c-m}{m}\right)}$ where c is number of concurrent deviations (same direction) m is number of pairs compared (equals to n-1)
Formula 84	Regression Coefficients: Y on X: $b_{yx}=r.\frac{SD_y}{SD_x}$ or $b_{yx}=\frac{\text{cov}(x,y)}{\left(SD_x\right)^2}$ X on Y: $b_{xy}=r.\frac{SD_x}{SD_y}$ or $b_{xy}=\frac{\text{cov}(x,y)}{\left(SD_y\right)^2}$
Formula 85	Correlation Coefficient is the GM of regression coefficients: $r_{xy}=\pm\sqrt{b_{xy}\times b_{yx}}$ Note: r_{xy}, b_{xy}, b_{yx} all will have same sign
Formula 86	Change of Origin/ Scale for Regression Coefficients: Origin no impact, Scale impact of both magnitude and sign. $b_{vu}=b_{yx}\times\frac{\text{change of scale of y}}{\text{change of scale of x}}$ $b_{uv}=b_{xy}\times\frac{\text{change of scale of x}}{\text{change of scale of y}}$
Formula 87	Two regression lines (if not identical) will intersect at the point $(\bar{x},\bar{y})$
Formula 88	Coefficient of Determination/ Explained Variance/ Accounted Variance: $\left(r_{xy}\right)^2$
Formula 89	Coefficient of Non-determination/ Un-explained Variance/ Un-accounted Variance: $1-\left(r_{xy}\right)^2$

Formula 90	Probable Error in correlation: $0.6745 \times \frac{1-r^2}{\sqrt{N}}$
Formula 91	Error Limits of Population Correlation Coefficient: $r \pm PE$
Formula 92	Price Relatives: $\frac{P_n}{P_0}$, Quantity Relatives: $\frac{Q_n}{Q_0}$, Value Relatives: $\frac{V_n}{V_0}$
Formula 93	Simple Aggregative Index: $\frac{\Sigma P_n}{\Sigma P_0} \times 100$
Formula 94	Simple Average of Relatives – Method Index: $\frac{\Sigma \frac{P_n}{P_0}}{n}$
Formula 95	Laspeyres Index (weight – base year quantity weight) $\frac{\Sigma P_n Q_0}{\Sigma P_0 Q_0} \times 100$
Formula 96	Paasche's Index (weight – current year quantity weight) $\frac{\Sigma P_n Q_n}{\Sigma P_0 Q_n} \times 100$
Formula 97	Marshall-Edgeworth Index (weight – sum of both current and base quantity) $\frac{\Sigma P_n (Q_0 + Q_n)}{\Sigma P_0 (Q_0 + Q_n)} \times 100$
Formula 98	Fisher's Ideal Index: GM of Laspeyres Index and Paasche's Index $\sqrt{\frac{\Sigma P_n Q_0}{\Sigma P_0 Q_0} \times \frac{\Sigma P_n Q_n}{\Sigma P_0 Q_n}} \times 100$
Formula 99	Bowley's Index: AM of Laspeyres Index and Paasche's Index $\frac{\frac{\Sigma P_n Q_0}{\Sigma P_0 Q_0} + \frac{\Sigma P_n Q_n}{\Sigma P_0 Q_n}}{2} \times 100$

About CA. Pranav Popat Sir

- He is a Chartered Accountant (Inter and Final Both Groups in First Attempt) with 8+ years of experience.
- He is an Educator by Passion and his Choice (Dil Se ❤️)
- Taught lakhs of students in last 6 years
- He teaches subjects of QA - Maths, LR and Stats (Paper 3) at CA Foundation Level and Cost & Management Accounting (Paper 4) at CA Intermediate Level.

Hope this formula book helps you in revising all formulas and become helpful to you during exam time, I made this with my whole heart, make best use of it and I just want one thing in return - share these notes to every student who really needs this.

Wishing you ALL THE BEST for upcoming examinations, see you soon in Inter Costing!!!

Ab mushkil nahi kuch bhi, nahi kuch bhi!!!

With Lots of Love

CA. Pranav Popat (P^2 SIR)