Exam Assistance Note

$X_0 =$ Value in the Base Year , $X_n =$ Value in the Recent Year

$\hat{X}_{n+k} =$ Predicted Value in the Year $n+k$

CBR= Crude Brith Rate (out of 1000 population) , CDR= Crude Death Rate (out of 1000 population)

Percentage of Change	$\frac{X_n - X_0}{X_0} \times 100\%$
Average Annual Growth Rate	$\text{AAGR\%} = \left(\frac{X_n}{X_0}\right)^{\left(\frac{1}{n}\right)} \times 100 - 100\%$
Prediction in the year $n+k$ based on AAGR% and X_n	$\widehat{X}_{n+k} = X_n \times \left(\frac{100 + AAGR}{100}\right)^k$
Average Growth Rate Percentage (AGR%) based on CBR and CDR	$\text{AGR\%} = \frac{\text{CBR} - \text{CDR}}{10}\%$
Doubling Time (DT)	$\text{DT} \approx \frac{70}{AAGR\%}$

Related to Human Development Index (HDI)

PcGNI = Percapita Gross National Product

$$\text{General Dimension Index} := \frac{\text{Actual Value} - \text{Min}}{\text{Max} - \text{Min}}$$

$$\text{Income Index} : \text{I}_{\text{Income}} = \frac{\log(\text{PcGNI}) - \log(\text{Minimun PcGNI})}{\log(\text{Maximum PcGni}) - \log(\text{Minimun PcGNI})}$$

$$\text{Education Index} : \text{I}_{\text{Education}} = \frac{\sqrt{I_1 \times I_2}}{0.951}, \text{ where}$$

$I_1 =$ Dimension Index for the Variable 'Mean Years of Adult Education '

$I_2 =$ Dimension Index for the Variable 'Expected Years of Schooling '

$$\text{Human Development Index} : \text{HDI} = (\text{I}_{\text{Life}} \times \text{I}_{\text{Education}} \times \text{I}_{\text{Income}})^{\left(\frac{1}{3}\right)}$$

Name of Attribute	Minimum Value	Maximum Value
Life Expetancy	20	83.2
Mean Years of Adult Education	0	13.2
Expected Years of Schooling	0	20.6
Combined Education Index	0	0.951
Percapita GNI (PcGNI)	163	108211

Related to Regression

Let $(X_1,Y_1),(X_2,Y_2),\ldots,(X_n,Y_n)$ are sijmple random sample of size n of the variabes X,Y. X is the covariate and Y is the response variable.

Attribute	Notation	Formula
Sample Mean of X	$\overline{X}$	$\frac{1}{n}\sum_{i=1}^{n} X_i$
Sample Mean of Y	$\overline{Y}$	$\frac{1}{n}\sum_{i=1}^{n} Y_i$
Sample Standard Deviation of X	S_X	$\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$
Sample Standard Deviation of Y	S_Y	$\frac{1}{n-1}\sum_{i=1}^{n}(Y_i-\overline{Y})^2$
Covariance between X and Y	S_{XY}	$\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y})$
Correlation between X and Y	r_{XY}	$\frac{S_{XY}}{S_X S_Y}$

Let us consider a linear regression line $\hat{Y}=a+bX$, where Y is the response variable, X is the independent variable or Covariate. a is called the **'intercept'** and b is called the **'slope'** of the fitted line. The formula for the estimated value of a and b is given as below:

$$\hat{b}=r_{X,Y}\frac{S_Y}{S_X} \quad \text{and} \quad \hat{a}=\overline{Y}-\hat{b}\overline{X}$$

Note that S_Y and S_X are always a positive number. Where as, the correlation coefficient between X and Y, r_{XY} is a number between -1 to 1.

Gini's Formula

There are several ways to compute Gini coefficient however it is often calculated with Brown Formula given by:

$$G=100\%-\left(\frac{\sum[f_i\times(y_i+y_{i+1})]}{10,000}\times 100\right)\%$$

where f_i = percentage decile shares of world population. y_i = cumulative percentage shares of world product for percentile groupings. An Example of Gini's Table from Slide

Decile	f_i	x_i	y_i	y_i+y_{i-1}	$f_i(y_i+y_{i-1})$
10	10	2.1	2.1	2.21	21
20	10	3.5	5.6	7.7	77
40	20	9.9	15.50	21.1	422
60	20	14.2	29.7	45.2	904
80	20	21.1	50.8	80.5	1610
90	10	15.7	66.5	117.3	1173
100	10	33.5	100	166.52	1665
				Total=	**5872**

$$G=100\%-\left(\frac{5872}{10,000}\times 100\right)=41.28\%$$

Confidence Interval

Let $X_1, X_2, \ldots, X_n$ be observed random **sample of size** n. Let $\overline{X}$ denotes the sample Mean while S_X refers to the sample Standard deviation. Then, assuming the population has a fine variance and the sample size n to be a large number, a formula for confidence interval for the population mean is given as:

$$\left[\overline{X} - z_{\text{Cutoff}} \times \frac{S_X}{\sqrt{n}}, \qquad \overline{X} + z_{\text{Cutoff}} \times \frac{S_X}{\sqrt{n}}\right]$$

z_{Cutoff} is determined based on the confidence-level of the interval. The values of z_{Cutoff} can be obtained from the following table:

Confidence Level	90%	95%	99%
z_{Cutoff}	1.645	1.96	2.576

An Example of Remaining Life Expectancy Table

Complete the following table.

Reverse cumulative L_x.

x	l_x	d_x	$L_x = \frac{l_x + l_{x+1}}{2}$	T_x	$e_x = \frac{T_x}{l_x}$
0	100	30	$\frac{100+70}{2} = 85$	0+20+50+65+85=220	$\frac{220}{100} = 2.2$
1	100-30 = 70	10	$\frac{70+60}{2} = 65$	0+20+50+65 = 135	$\frac{135}{70} = 2.0$
2	70-10 = 60	20	$\frac{60+40}{2} = 50$	0+20+50 = 70	$\frac{70}{60} = 1.167$
3	60-20 = 40	40	$\frac{40+0}{2} = 20$	0+20 = 20	$\frac{20}{40} = 0.5$
4	40-40 = 0	0	0	0 = 0	0 = 0