A study to evaluate Vitamin D status among schoolchildren in Thailand

1. Compute numerical summaries of the distribution of serum 25(OH)D level.

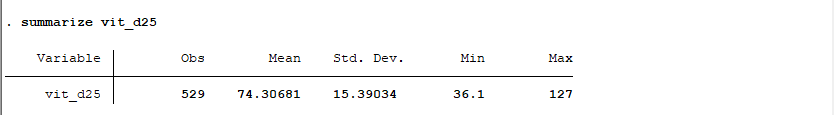


Table 1: Numerical summaries of the distribution of serum 25(OH)D level.

1. Create graphical summaries of the distribution of serum 25(OH)D level

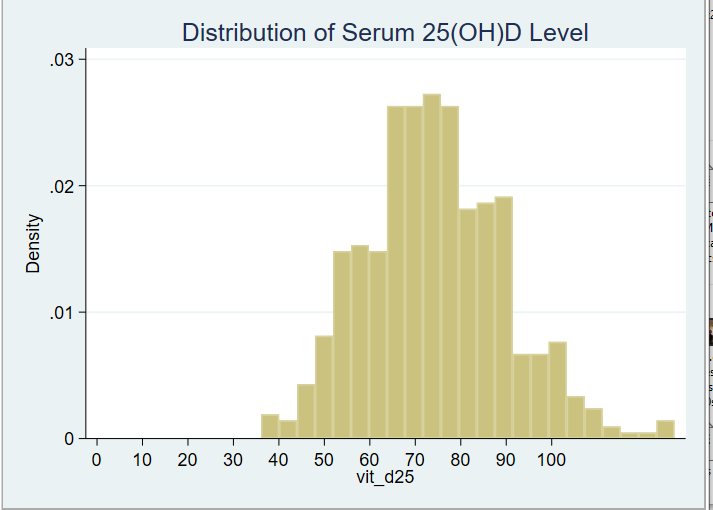


Figure 1: Histogram of the distribution of serum 25(OH)D level

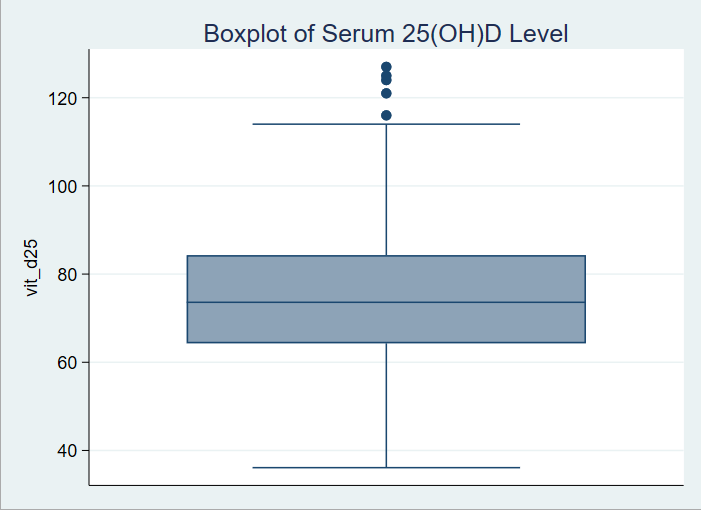


Figure 2: Box Plot of the distribution of serum 25(OH)D level

1. With reference to the summaries from parts i. and ii. as needed, describe the distribution of serum 25(OH)D level in no more than five sentences.

The average serum 25(OH)D level is approximately 74.31 nmol/L, shows majority of the sampled children have adequate vitamin D.

The standard deviation is 15.39 nmol/L, suggesting that: there is moderate variability in serum 25(OH)D levels among the children.

The serum 25(OH)D levels range from a minimum of 36.1 nmol/L to a maximum of 127 nmol/L, showing a wide spread of Vitamin D levels in the population.

Based on the histogram, the distribution of serum 25(OH)D levels is almost normal but slightly right-skewed, with most values concentrated around the mean and a few higher values extending towards the upper end of the range

Based on the box plot, the distribution of serum 25(OH)D is mainly concentrated between 60 nmol/L and 85 nmol/L.

1. Create a plot to graphically show the association between sex and serum 25(OH)D level. Describe what you see.

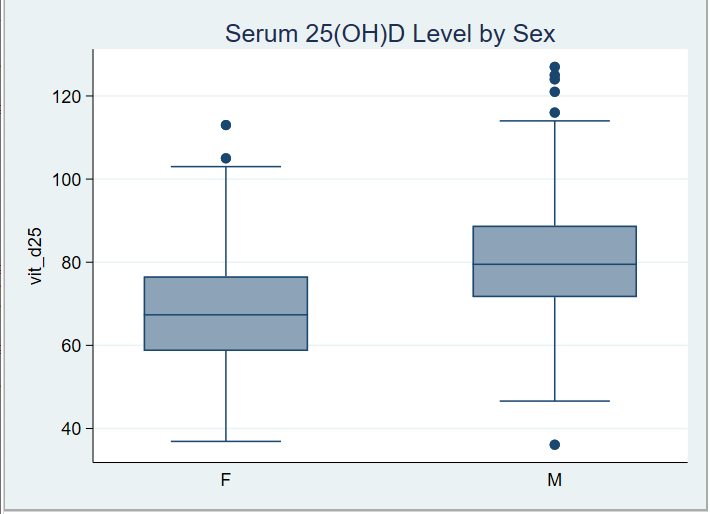


Figure 3: Box plot of the association between sex and serum 25(OH)D level.

The Box Plot indicates:

Males tend to have higher serum 25(OH)D levels than females, as shown by a higher median and a wider IQR.

The distribution of serum 25(OH)D levels is more spread out for males, indicating greater variability.

Both sexes have some outliers, with males having more high outliers and both sexes having a similar overall range.

1. The investigators were mostly interested in estimating the prevalence of Vitamin D deficiency in the population under study. Using these data, calculate an estimate of the prevalence of Vitamin D deficiency. Hint: Make sure you only include individuals with a recorded value of serum 25(OH)D level.

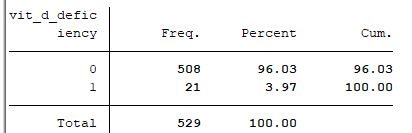


Table 2: Prevalence

**Frequency**: Out of the total 529 observations, 508 children (96.03%) do not have Vitamin D deficiency going by the stated level of deficiency. 21 children (3.97%) have Vitamin D deficiency

The prevalence of Vitamin D deficiency among the studied children is 3.97%.

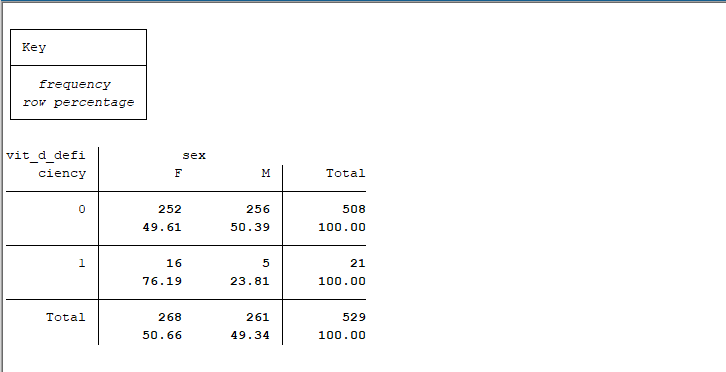


Table 3: Prevalence

For those who do not have Vitamin D deficiency, the distribution is roughly equal between females (49.61%) and males (50.39%).

Among those who have Vitamin D deficiency, a higher proportion are females (76.19%) compared to males (23.81%).

Generally, total population of the study appears evenly split between females (50.66%) and males (49.34%).

1. Calculate the relative risk of Vitamin D deficiency, comparing females to males.

Interpret the calculated value.

**Relative Risk for Males vs. Females**

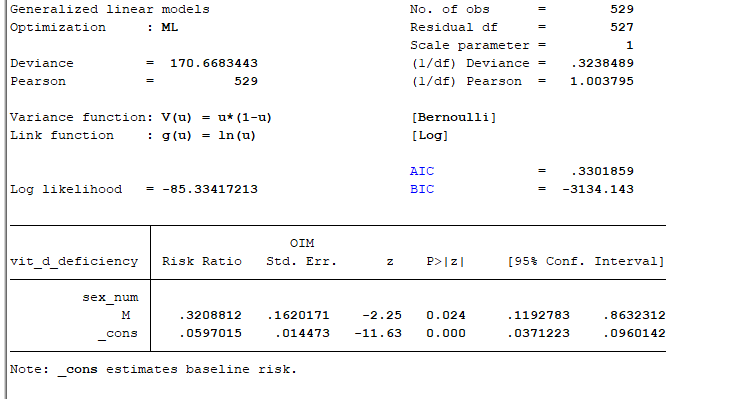




Table 4: Relative Risk

The risk ratio for males is 0.3208812, indicates that males have about 32.1% of the risk of Vitamin D deficiency compared to females.

The confidence interval [0.1192783, 0.8632312] does not include 1, indicating that this result is statistically significant.

The p-value of 0.024 suggests that there is a statistically significant difference in the risk of Vitamin D deficiency between males and females at the 5% significance level.

The baseline risk /intercept (for females, the reference category) is 0.0597015. This means that the estimated probability of Vitamin D deficiency for females is approximately 5.97%.

The confidence interval [0.0371223, 0.0960142] indicates that we can be 95% confident that the true baseline risk lies within this range

In summary:

**Relative Risk**: Males have a significantly lower risk of Vitamin D deficiency compared to females, as shown by relative risk of approximately 0.32. This means that males are about 68% less likely to have Vitamin D deficiency compared to females.

**Baseline Risk**: The baseline risk of Vitamin D deficiency for females is approximately 5.97%

1. Children in the study were randomly sampled from a rural subregion of Thailand that has a high rate of poverty. Using numerical and graphical summaries, investigate whether the data suggest that growth in children from this subregion might be affected by the impoverished conditions. Summarize your findings. Hint: What do the standardized height-for-age z-scores tell you?

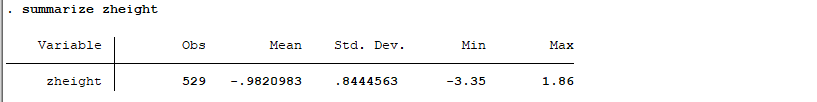


Table 5: zheight

**Mean and Distribution**: The mean z-score for height-for-age is approximately -0.98, which is below zero. This indicates that, on average, the children in this study are shorter than the reference population for their age.

**Spread and Skewness**: The standard deviation is 0.84, suggesting noticeable variability in the standardized height-for-age z-scores.

**Range**: The z-scores range from -3.35 to 1.86. The minimum value of -3.35 shows that some children are severely stunted. The maximum value of 1.86 shows that a few children have heights above the reference median

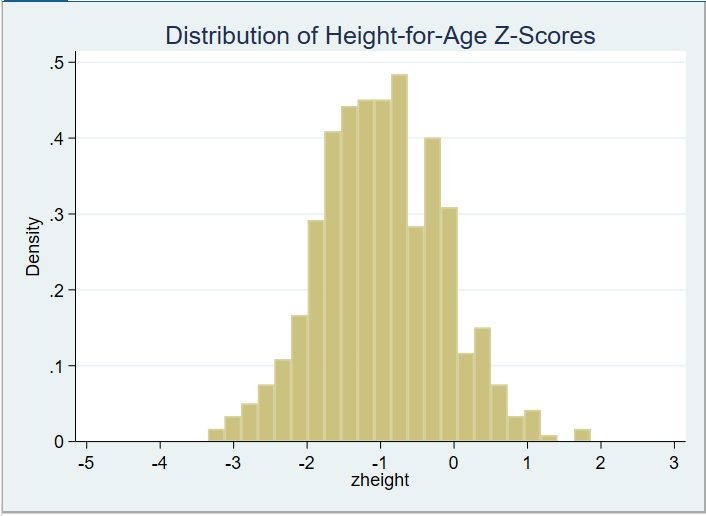


Figure 4: Z-scores

The histogram shows a distribution that is slightly skewed to the left, indicating that more children have z-scores below the mean.

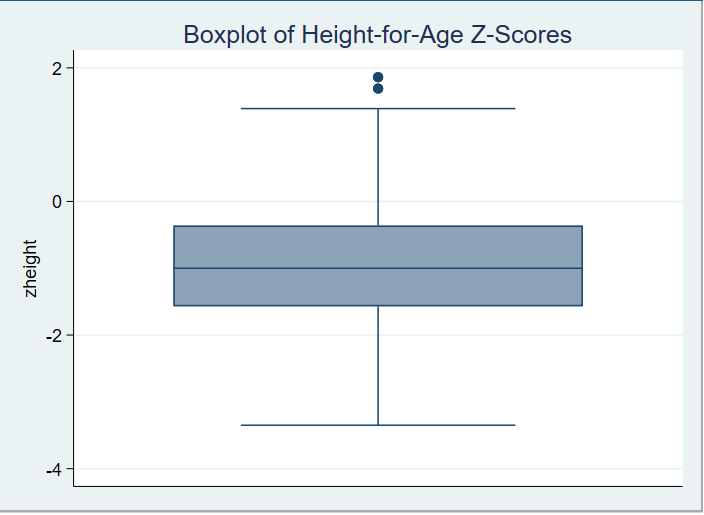


Figure 5: Z-scores Box plot

**Boxplot Analysis**: The boxplot further confirms that the median z-score is below zero. The interquartile range (IQR) extends from around -1.5 to -0.3. There are also a few outliers above the upper whisker, indicating some children are taller than the typical range for this population.

**Growth Implications**: The negative mean z-score and the skewness toward lower z-scores suggest that many children in this subregion are experiencing stunted growth, which could be a result of impoverished conditions. This may include poor nutrition and inadequate health care.

The high rate of stunting is indicating a chronic malnutrition, which is common in regions with high poverty rates.