

# Critical thinking on chest sizes of scottish militiamen

[Business](#)



- Construct a relative frequency histogram for the chest circumference data, using classes based on a single value.

Looking at the table we first have to get the total frequency by summing the values in that specific column. The table is filled according to the following procedure. You can check out the excel file attached alongside that show numerical workings.

Arrange the value in the order presented and then multiply each frequency and chest size per row to get the product of frequency and chest size. When you add the entire column of the results and divide by the total frequency, you will get the mean of the data presented.

$$\text{Total frequency} = f = 3 + 19 + 18 + 3 + 1 = 5732$$

we define the relative frequency as

$$\text{Relative frequency} = \frac{\text{Frequency of a given chest size}}{\text{Total frequency}} * 100\%$$

$$\text{Relative frequency for Chest size 33} = \frac{3}{5732} * 100 = 0.05\%$$

$$\text{Relative frequency for Chest size 39} = \frac{106}{5732} * 100 = 18.53\%$$

This formula is applied to the rest of the chest sizes to fill the table appropriately. Obviously you are going to use a suitable scale for the y-axis of your histogram.

The x-axis scale is made of intervals of 1, that is chest size move 1 an incremental form by 1 unit. The y-axis which basically represents the relative frequency calculated for each chest size.

- Our mean for the table data can be given by using the following formula

$$\text{Mean} = \mu = \frac{\text{frequency} * \text{chest size}}{\text{frequency}} = \frac{228414}{5732} = 39.85$$

## Variance is given by the following formula

$$\text{Var}_x = \sum (x_i^2 \cdot f_i) - (\sum x_i \cdot f_i)^2 / n$$

$$\sum (x_i^2 \cdot f_i) = \text{Chest size}^2 \cdot \text{Frequency} / \text{frequency}$$

$$\sum (x_i^2 \cdot f_i) = 91266885732 = 1592.23$$

## Therefore our Variance

$$\text{Var}_x = 1592.23 - 39.85^2 = 4.298$$

$$\text{Standard deviation} = \sqrt{4.298} = 2.07$$

Let us assume that the distribution is converted to a standard normal distribution usually given or expressed as  $Z \sim N(0, 1)$ , that is a normal distribution with a mean of zero and standard deviation of 1

$X - \mu \sim N(0, 1)$  this is the formula we will use in calculating probabilities based on Z-values, X values are the chest sizes values while the mean  $\mu$  is 39.85 and the standard

deviation  $\sigma$  is 2.07, while n has a value of 5732

The curve above represents the normal curve obtained from the relative frequency histogram. The y-axis can be taken to represent the probability value and the x-axis remains as the chest size. This curve is basically obtained by drawing a smooth line passing through the top midpoints of the relative frequency bars. The larger values are at the top while the lower chest sizes values are on the sides subside to form an almost symmetrical curve. This is what makes the normal curve with no realizable skew on the sides of the curve. The curve displays the use of central limit theorem that is usually in practice for the purpose of normal approximation of particular distribution. With the number observation less more than 30, to be exact 5732, the shape of the curve in my view is almost symmetrical and has

<https://assignbuster.com/critical-thinking-on-chest-sizes-of-scottish-militiamen/>

almost a skewed value of zero which statistically makes the curve normally distributed.

The central limit theorem applied in this situation will help us get appropriate probabilities of interval involved in chest size. You can say the intervals are the class size between a given chest size to another. For example Prob35 will represent the probability of percentage of militia falling in that class of chest

size between chest size 35 and chest size 42. This is illustrated in the successive question.

- We use the values we got in the table in part (a), obviously we are applying normal approximation as applied to either a binomial distribution or a Poisson distribution.

prob36 and  $X_1 = 36$  while  $X_2 = 41$  while  $\mu = 39.85$ ,  $\sigma = 2.07$  and  $n = 5732$   
 $= \text{prob}_{36-39.85} \cdot 2.07 \cdot 5732$  This can also be proven by using the cumulative sum of the product of frequency and chest size between 36 and 41. The cumulative summation to chest size of 36 is shown on the table as 3580, while the cumulative summation for chest size of 41 is shown on the table as 177164. Express this as a percentage of total summation.

Percentage of militia with chest size between 36 and 41 =  $\frac{177164 - 3580}{228414} \cdot 100 = 75.99\%$

**The cumulative summation is given below or in the excel file: we use it to calculate the percentages**

- We now have to use the normal curve to approximate the percentage of militia with a chest size between 35.5 and 41.5.

## Using normal approximation

prob $35.5$  and  $X_1 = 35.5$  while  $X_2 = 41.5$  while  $\mu = 39.85$ ,  $\sigma = 2.07$  and  $n = 5732$

$= \text{prob}35.5 - 39.85 \frac{2.07}{5732}$  We predict the relative frequency of  $35.5$  and  $41.5$  using an extrapolation method. For  $35.5$

For  $35.5 \Rightarrow 36 - 35.5 = 0.5$   $36 - 35 = 1$   $3.3 - a$   $3.3 - 1.41 = 0.53$   $3.3 - 1.41 = 3.3 - a = 0.$

$945 = 3.3 - a$

$\Rightarrow a = 2.36\%$

For  $41.5 \Rightarrow 42 - 41.5 = 0.5$   $42 - 41 = 1$   $11.27 - b$   $11.27 - 16.31 = 0.51$   $11.27 - 16.31 = -2.$

$52 = 11.27 - b \Rightarrow b = 13.79\%$  the relative frequency in percentage for  $35.5$  chest sizes is  $2.36\%$  forecasted while relative frequency for  $41.5$  chest sizes

is calculated as  $13.79\%$  forecasted. We can now use the percentages to get the frequency by substituting the formula for relative frequency and then calculate the cumulative summation of the product of chest size and frequency.

Frequency for any given relative frequency = relative frequency for a given chest size  $100 * \text{Total frequency}$

Frequency for  $2.36\% = 2.36100 * 5732 = 135$  and for  $13.79\% = 13.$

$79100 * 5732 = 790$

With these values we can get the cumulative summation of the product and frequency at the respective chest size of  $35.5$  and  $41.5$ . This is elaborately displayed below by step by step calculation and getting the percentage by applying the same formula applied to the former question.

Cumulative sum at  $35.5$  chest size =  $135 * 35.5 + 3580 = 7832$

Cumulative sum at  $41.5 = 790 * 41.5 + 177164 = 209967$

**The approximate hence can be calculated and we can compare the value to the one we calculated in previous question.**

Percentage of militia with chest size between 35.5 and 41.5 =  $\frac{209967-7832228414}{100} = 88.5\%$

Comparing this value to the former which is a value of 76%, we see there is a difference of  $(88.5 - 76) = 12.5\%$ . This is attributed to the continuity correction experienced between the two questions for normal approximation. In the former the integer values were rounded off to the nearest number while in the latter case the values were treated to be continuous rather than discrete.