**Indian Statistical Institute**

**BSDS: 2024-26**

**First Year: Semester – II**

**Economics-II**

**Practical Exercise 6**

1. The following table gives the size distribution of monthly income of 460 individuals collected in a survey.

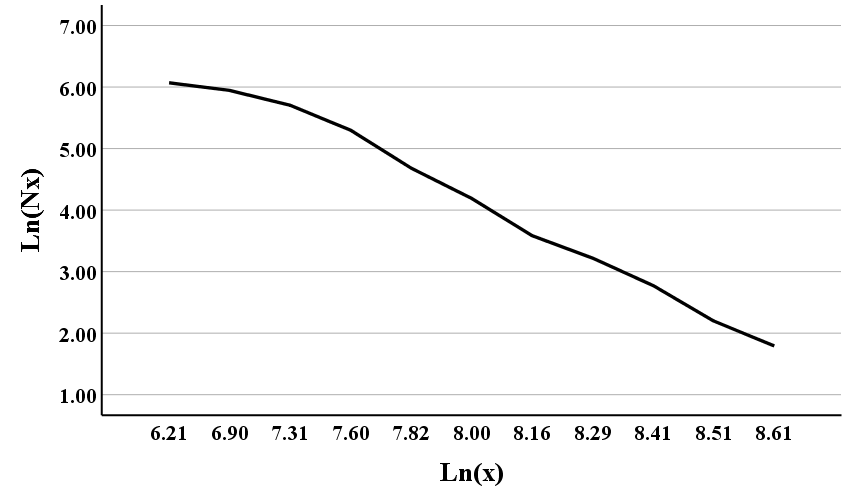
Table 1: Size Distribution of monthly income of 460 individuals

|  |  |
| --- | --- |
| Income Class  (Rs.) | No. of persons |
| 0 – 500 | 28 |
| 500 – 1000 | 50 |
| 1000 – 1500 | 82 |
| 1500 – 2000 | 100 |
| 2000 – 2500 | 92 |
| 2500 – 3000 | 42 |
| 3000 – 3500 | 30 |
| 3500 – 4000 | 11 |
| 4000 – 4500 | 9 |
| 4500 – 5000 | 7 |
| 5000 – 5500 | 3 |
| 5500 & above | 6 |

Find graphically the appropriate range of income over which Pareto law holds. Also fit the distribution in the appropriate range and perform a χ2 test of goodness of fit.

**Solution to Qn.1:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Income Class  () | f | Lower boundary  (x) | Cum. Freq. | Ln(x) | Ln( |
| 0 – 500 | 28 |  |  |  |  |
| 500 – 1000 | 50 | 500 | 432 | 6.2146 | 6.0684 |
| 1000 – 1500 | 82 | 1000 | 382 | 6.9078 | 5.9454 |
| 1500 – 2000 | 100 | 1500 | 300 | 7.3132 | 5.7038 |
| 2000 – 2500 | 92 | 2000 | 200 | 7.6009 | 5.2983 |
| 2500 – 3000 | 42 | 2500 | 108 | 7.8240 | 4.6821 |
| 3000 – 3500 | 30 | 3000 | 66 | 8.0064 | 4.1897 |
| 3500 – 4000 | 11 | 3500 | 36 | 8.1605 | 3.5835 |
| 4000 – 4500 | 9 | 4000 | 25 | 8.2940 | 3.2189 |
| 4500 – 5000 | 7 | 4500 | 16 | 8.4118 | 2.7726 |
| 5000 – 5500 | 3 | 5000 | 9 | 8.5172 | 2.1972 |
| 5500 & above | 6 | 5500 | 6 | 8.6125 | 1.7918 |



It is also known that for Pareto distribution

If c is known then take only one quantile, preferably median to estimate α. If it is not known, then take two quantiles. We shall only demonstrate it by taking a single quantile, i.e., assuming c to be known as 1560.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Income Class  () | f | Cum. Freq.  (≤ type) |  |  | Expected frequency  (e) |
| 1560 – 2000 | 88 | 88 | 0.3569 | 0.3569 | 102.79 |
| 2000 – 2500 | 92 | 180 | 0.5674 | 0.2105 | 60.62 |
| 2500 – 3000 | 42 | 222 | 0.6871 | 0.1197 | 34.47 |
| 3000 – 3500 | 30 | 252 | 0.7621 | 0.0750 | 21.60 |
| 3500 – 4000 | 11 | 263 | 0.8124 | 0.0503 | 14.49 |
| 4000 – 4500 | 9 | 272 | 0.8478 | 0.0354 | 10.20 |
| 4500 – 5000 | 7 | 279 | 0.8738 | 0.0260 | 7.49 |
| 5000 – 5500 | 3 | 282 | 0.8934 | 0.0196 | 5.64 |
| 5500 & above | 6 | 288 | 1 | 0.1066 | 30.70 |
| Total | 288 | --- | --- | 1 |  |

Since 288/2 = 144 is in the interval 2000 – 2500, the median must be in that interval. Hence,

Now using the formula of quantile for Pareto distribution, we get

The calculated value (45.4) is greater than the tabulated value (14.07) of Chi square at 5% level with 7 degrees of freedom. Hence, the data does not come from Pareto distribution.