(IJDSSH) 2017, Vol. No. 4, Jul-Dec **e-ISSN: 2455-5142; p-ISSN: 2455-7730**

BUILDING AN ECONOMETRIC PREDICTION MODEL TO TEST THE GIVEN HYPOTHESIS ON THE SPECIMEN DEMOGRAPHICS WITH PROVISIONS FOR CONTROLLED VARIABLES

Nishita Lamba

Satyawati College, University of Delhi

U.S Bureau of labor statistics published an unemployment study and estimated that mean timeframe of joblessness was 14.6 weeks in November 1998.A portion of this study has been provided in the form of 15 observations on Age and unemployment time of residents of Philadelphia. Following is the set of information:

Age	Weeks	Age	Weeks	Age	Weeks
56	22	22	11	25	12
35	19	48	6	25	1
22	7	48	22	59	33
57	37	25	5	49	26
40	18	40	20	33	13

The question that pursues are:

1. Use the elucidating insights to outline the information.

2. Develop a 95% certainty interim gauge of the mean time of joblessness people in Philadelphia.

3. Conduct a theory test to decide if the mean term of joblessness in Philadelphia is more prominent than the national mean span of 14.6 weeks. Use a .01 dimension of centrality. What is your decision?

4. Is there a connection between the age of a jobless individual and the number of long stretches of joblessness? Clarify.

1. Descriptive statistics simply includes Mean and standard of the sample which can be used as a summary for the sample data. So our task is to determine the mean and standard deviation of Age and unemployment time. We let X denote the age of an individual and let Y denote the weeks of unemployment for the sake of notation convenience. Procedure to find mean and standard of X and Y is given below.

X_i (Age)	Y_i (weeks)	X_i^2	Y_i^2
56	22	3136	484
35	19	1225	361
22	7	484	49
57	37	3249	1369
40	18	1600	324
22	11	484	121
48	6	2304	36
48	22	2304	484
25	5	625	25
40	20	1600	400
25	12	625	144
25	1	625	1
59	33	3481	1089
49	26	2401	676
33	13	1089	169
$\sum X_i=584$	$\sum Y_i=252$	$\sum X_i^2=25232$	$\sum Y_i^2=5732$

$\bar{X}= \frac{\sum Xi}{N}$ $s_X=\sqrt{\frac{\sum Xi^2}{N} - \frac{\sum Xi}{N}}$

$\bar{X}$= sample mean age

s_X= sample standard deviation of age

$\sum X_i$=sum of ages of all observations

N=number of observations, i.e. 15

$\bar{Y} = \frac{\sum Yi}{N}$ $s_Y=\sqrt{\frac{\sum Yi^2}{N} - \frac{\sum Yi}{N}}$

$\bar{Y}$=sample mean of length of joblessness

s_Y= sample standard deviation of weeks of unemployment

$\sum Y_i$=sum of all observations of weeks of unemployment

N=number of observations, i.e. 15

$\bar{X}= \frac{584}{15} = 38.93$ $s_X=\sqrt{\frac{25232}{15} - \frac{584}{15}}$ =12.90

$\bar{Y} = \frac{252}{15} = 16.8$ $s_Y=\sqrt{\frac{5732}{15} - \frac{252}{15}}$ =9.99

2. CONFIDENCE INTERVAL

The estimation of the populace standard deviation isn't known and furthermore N<30, thus we will utilize T-appropriation to build certainty interim for certainty dimension of 95% of the mean period of jobless people. Sample standard deviation of age i.e. s_X to estimate population standard deviation.

s_X =estimate of population σ of age=12.90

$\sigma_{\overline{X}}$=standard error of mean=$\frac{Sx}{\sqrt{N}}$ =$\frac{12.90}{\sqrt{15}}$ =3.33

α Significance level =0.5

Degrees of freedom=N-1=14

Confidence interval

$= \mathbf{X} - \sigma_{X}(t_{\alpha/2})$ and $\mathbf{X} + \sigma_{X}(t_{\alpha/2})$

= **38.93-3.33(1.729)** and **38.93+3.33(1.729)**

=**33.17 -44.69**

3. HYPOTHESIS TESTING

To conduct hypothesis testing we will proceed as follows.

H_0: μ=14.6 (mean duration of joblessness is 14.6 weeks.)

H_1 $\mu \neq 14.6$ (mean duration of unemployment is not 14.6 weeks.)

Since the alternative hypothesis contains a sign of inequality it is a two-tailed test. Significance level (α) is 0.01. To test the hypothesis we will construct the confidence level. The two-tailed test at the 1 percent level of significance is equivalent to a two-sided 90 percent confidence interval. If the confidence interval does not contain μ=14.6, we reject H_0. For the duration of unemployment I, the sample mean is 16.8 with a sample standard deviation of 9.99. Using t-distribution table $t_{0.05}$ is 2.977, so the 99 percent confidence interval for μ is

$$\overline{Y}+ t\frac{Sy}{\sqrt{n}} =16.8+(2.977)\frac{9.99}{\sqrt{15}} = 16.8 +7.68$$

$$9.12 < \mu < 24.48$$

Since μ=14.6 lies within the 99 percent confidence interval as established above, we cannot reject the hypothesis H_0:μ=14.6 at $\boldsymbol{\alpha}$=0.01 in a two-tailed test. This test can also be seen graphically by showing acceptance and rejection region. In figure nest page confidence limit shown above is the acceptance region is shown by the area under t curve from 9.12 to 24.48, and the outside this region is the rejection area. Rejection area is shown as the dark area under the curve at both the tails.

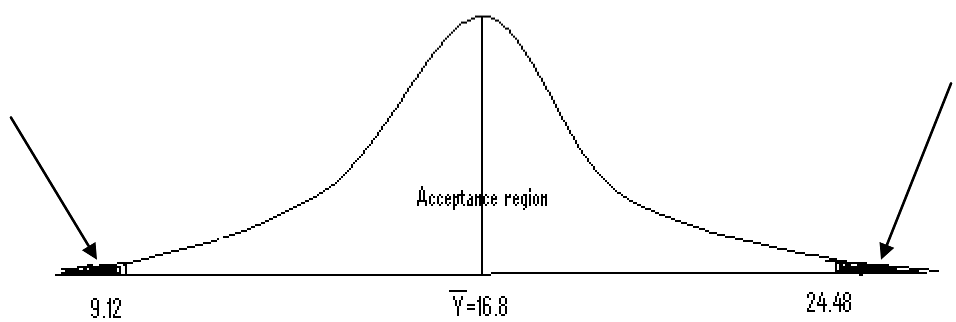

Figure 1

Since the estimation of μ=14.6 lies in the acknowledgment we would not dismiss the invalid hypothesis. Therefore there is no proof to trust the case that the normal length of joblessness is unique in relation to 14.6 weeks.

4. The relationship between age of an individual and number of weeks of unemployment. To see if there is any connection between the age of an individual and number of long stretches of unemployment we can make utilization of dissipating diagram with estimating age on x-hub and number of weeks on joblessness on y-pivot.

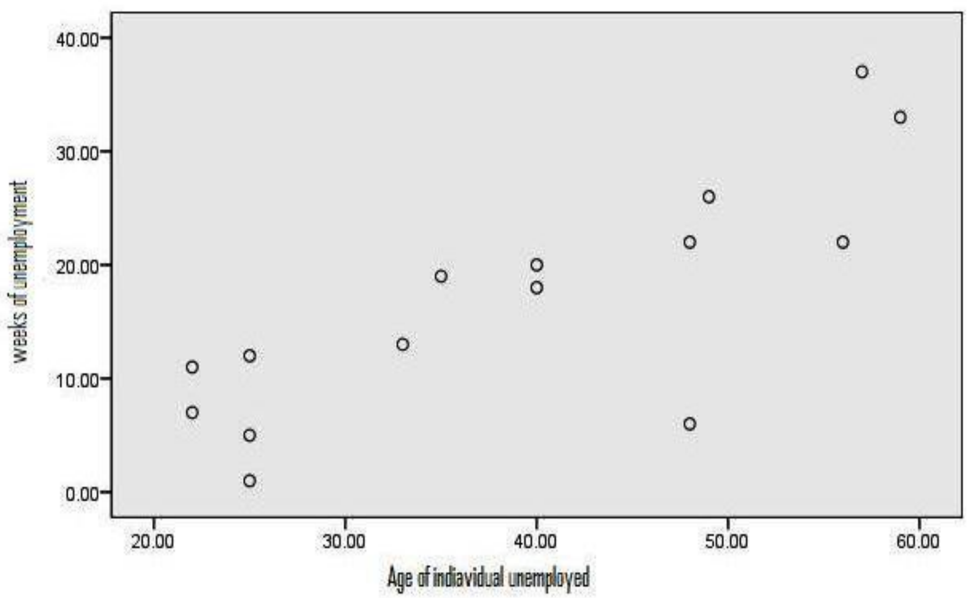

Figure 2

It very well may be obviously observed from the disperse outline that there is a positive relationship between's age of an individual and the span of joblessness. Moreover, we also can use correlation coefficient to show the positive linear association between age and duration of unemployment. Coefficient correlation can be calculated with the help of following formula.

$$r=\frac{\sum xy}{n.Sy.Sx}$$

where r= coefficient of correlation

$\sum xy$= sum of the product of deviation of x and y form their respective means

n= number of observations 15

Sx=sample standard deviation of x

Sy= sample standard deviation of y