Correlation Between Couples' Ages At The Time Of Marriage

October 2020

Samyam Lamichhane

lcsamyam@gmail.com

Budhanilkantha School, Kathmandu Nepal

Abstract

The age difference is the amount by which ages are different. Often recognized as 'Age disparity' in marriages, the age difference between spouses is a daunting challenge for a developing nation like Nepal. Studies have found that smaller age gaps (in marriages) are correlated with higher socioeconomic status and higher class [2]. Therefore, it is a topic which requires immediate attention. Here, we perform a series of statistical tests on data collected to examine existing age disparity in Nepal. The purpose of this test is to determine whether there exists a correlation between the husband's age and wife's age (at the time of marriage) and list possible sources of errors.

Acknowledgement

The author acknowledges Mr Ganesh Man Chaudhary, Statistics teacher at Budhanilkantha School, for the guidance in the research, and all the individuals for providing the data for analysis.

Introduction

According to the census of 2011 AD, the median age at first marriage of Nepalese women is 17.9 years and of Nepalese men is 22.1 years, while the age gap is 4.2 years [3]. The given is an average value based on the responses from the entire population of Nepal. However, it does not reveal the kind of relationship that exists between husband's and wife's ages at the time of their marriage. Is there a positive correlation or negative? Or does it exist?

Hence, this research is conducted to examine the existing relationship between two ages. The subject of the test is "**Correlation between spouses' ages at the time of marriage**".

Methods

Responses were collected from a pool of **a random population**. The population size was N = 16. The process of couples' selection was random.

Also, the prime term of the agreement with the couples was **confidentiality** to be maintained. Hence, keeping this factor in prime, the analysis was done.

A **survey sheet** was prepared using Google Docs and categorized into two sections – Section 1, consisting of personal details, and Section 2, consisting of three questions – date of birth of both the spouses and the date of marriage.

Then, the form was distributed via online portals. Moreover, a few responses were collected using printed forms.

The next step was **'Data analysis'**, for which three software – **Stata, Python and Excel** - were used. The raw response sheet is categorized using the Pandas library on a new sheet. Then, using Stata, equation of the line of regression, and the values of variables, for instance, Standard error of the slope (SE), SE Coefficient and so forth, are determined.

Afterwards, **a hypothesis test** is performed to determine whether the value of the slope of regression is zero or not.

Analysis

The responses collected were tabulated as below:

X, Husband's age at the time of marriage	Y, Wife's age at the time of marriage
19	17
26	24
27	26
33	21
25	19
22	20
18	17
24	21
35	24
29	26
25	18
25	19
34	27
25	25
25	18
23	26

Table 01: Table showing corresponding values of the independent and dependent variable.

Note: For this research, X is used as independent variable, while Y is the dependent variable.

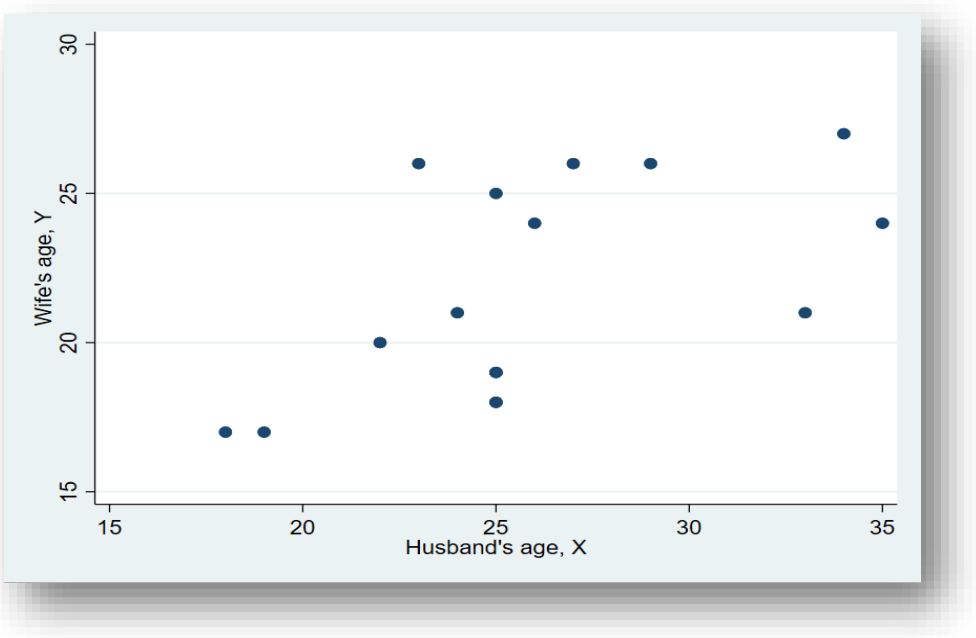

Figure 1: Scatterplot of X & Y

The result obtained after giving commands on Stata looks like the table below:

Source	SS	df	MS	Number of obs	=	16
				F(1, 14)	=	6.96
Model	64.7567231	1	64.7567231	Prob > F	=	0.0195
Residual	130.243277	14	9.30309121	R-squared	=	0.3321
				Adj R-squared	=	0.2844
Total	195	15	13	Root MSE	=	3.0501

WifesageY	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
HusbandsageX	.4295637	.1628166	2.64	0.019	.0803567	.7787706
_cons	10.60819	4.291346	2.47	0.027	1.404171	19.81221

Table 02: Table showing regression values of X & Y.

Let us assume the equation of the line of regression [#] is

$$y = \beta_0 + \beta_1 x + \varepsilon$$

Where,

β_0 is a constant (Y-intercept),

β_0 is the slope (also called the regression coefficient)

ε is the value for other variables

x is the independent variable

y is the dependent variable.

From the *Table 02*,

We know,

$$\beta_0 = 10.6082\ (4\ d.p.)$$

$$\beta_1 = 0.4296\ (4\ d.p.)$$

Hence, substituting the values of β_0 and β_1 in the equation [#], we get,

The equation of the line of regression is:

$$y = 10.6082 + 0.4296\,x + \varepsilon$$

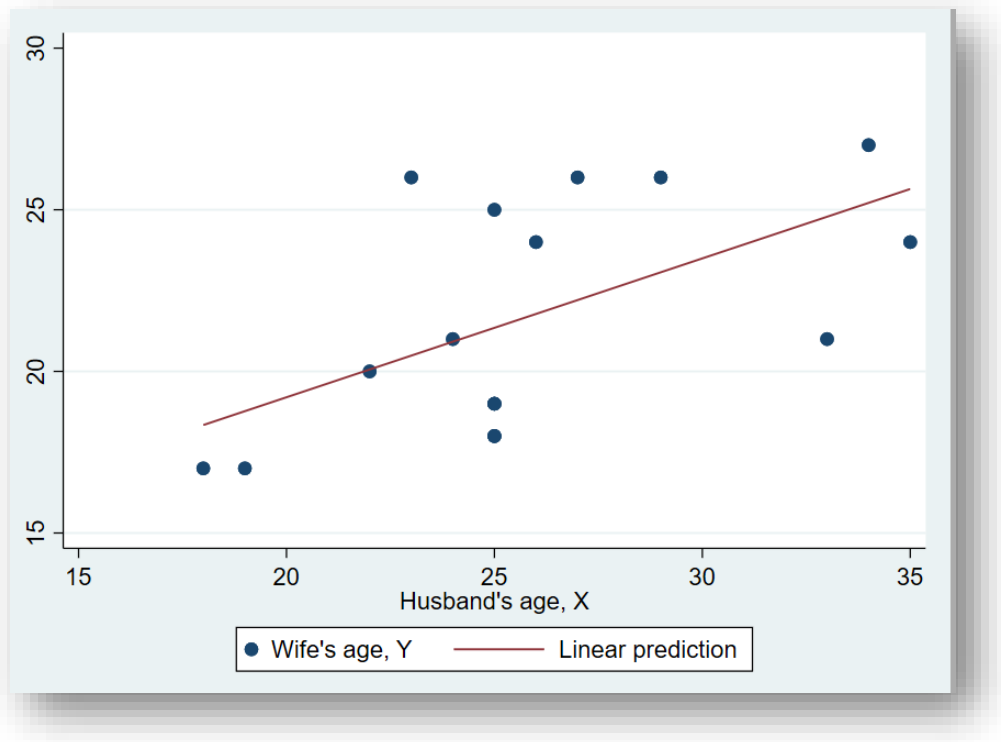

Figure 2: Line of regression of the data

Now,

Let us determine the value of **Pearson's correlation coefficient, r,** using Stata. Using the appropriate code gives:

	Wifesa~Y	Husban~X
WifesageY	**1.0000**	
HusbandsageX	**0.5763**	**1.0000**

Table 03: Table showing correlations and covariances of X & Y.

Again,

Let us use determine the value of **Standard error of the slope, SE** or $\boldsymbol{S_{b_1}}$, from the Table-02.

We get,

$$S_{b_1} = 0.1628\ (4\ d.p.)$$

Since, we have the equation of the line of regression and the value of β_1, let's perform a test to confirm whether the value of $\beta_1$is zero or not.

Let us assume our Null Hypothesis as H_0 and Alternate Hypothesis as H_1 .

$H_0: \beta_1 = 0$

$H_1: \beta_1 \neq 0$

Using the level of significance [*] as $\alpha = 5\%$.

To calculate the value of test statistic,

$$t_{.cal} = \frac{\beta_1}{S_{b_1}}$$

$$or, t_{.cal} = \frac{0.4296}{0.1628}$$

$$or, t_{.cal} = 2.6384\ (4\ d.p.)$$

Note: From *Table-02*, $t_{.cal} = t = 2.64$.

To calculate $t_{.tab}$, we need the value of 'Degree of freedom, Df' and 'α':

Degree of freedom,

$$\text{Df} = (N - k - 1)$$

$$or, \quad \text{Df} = (16 - 1 - 1)$$

$$or, \quad Df = 14$$

Now,

Using the values of Df, α, and the fact that it is a two-tail test, we have,

$$t_{.tab} = t_{14,95\%,two-tail}$$

$$t_{.tab} = 2.145$$

Discussion

Here,

$$| t_{.tab} | < |t_{.cal}|$$

Which means that $H_0$is rejected.

Therefore, at 5% significance level, it can be concluded that the value of $\boldsymbol{\beta_1 \neq 0}$ for the data collected.

Also,

From *Table 03,*

Pearson's correlation coefficient, r, between Y and X $= 0.5763\ (4\ d.p.)$

So, we have,

$$r^2 = 0.3318\ (4\ d.p.)$$

Which means that the **coefficient of determination** is 33.2%.

Note:

Based on following **assumptions**, *t*-test has been used for the hypothesis test:

i. The scale of measurement of the data collected follows ordinal scale.
ii. The data, when plotted, follow normal distribution.
iii. The data collected is simple random sample.
iv. The data collected follow homogeneity of variance.

Limitations of the research [*]

i. N = 16 was used for the research; however larger sample might have yielded better results.
ii. The research was done to determine 'Simple Linear Regression'. Including other variables would have resulted in a better set of results.
iii. Lack of previous studies in the subject of study caused problem in narrowing questions of study.

Conclusion

The study concludes that there exists a strong positive association between two variables. It means that as the age of husband (at the time of marriage) increases, wife's age (at the time of marriage) also increases [as shown by *figure 2*]. Since people who are married to much older or younger spouses tend to have larger declines in marital satisfaction [4]. This is a positive indicator of the quality of life. Still, the test was performed on a specific sample population, so it is not a representative for the overall Nepalese population. However, it can be used by population experts and Mathematicians for study.

Let's hope, the age gap will even decrease in the days to come, and all of the Nepalese couples will have an impressive quality of life.

References

1. Age Difference. (n.d.). Retrieved October 14, 2020, from https://www.merriam-webster.com/dictionary/age difference
2. Written by Katharina Buchholz, D. (2019, September). This chart shows the age that people get married across the world. Retrieved October 14, 2020, from https://www.weforum.org/agenda/2019/09/when-people-get-married-around-the-world/
3. List of countries by age at first marriage. (2020, October 12). Retrieved October 15, 2020, from https://en.wikipedia.org/wiki/List_of_countries_by_age_at_first_marriage
4. Rathore, V. (2019, March 18). Too much age gap posses negative impact on marriage: Study. Retrieved October 15, 2020, from https://english.newstracklive.com/news/too-much-age-gap-posses-negative-impact-on-marriage-study-sc23-nu-1004255-1.html