STATISTICAL MODEL FOR GENDER INEQUALITY

V.V. Patil[1], N. K. Dhane[2]
[1,2]Tuljaram Chaturchand College, Baramati.
[1]vaishutcc@gmail.com, [2]neetadhane@gmail.com

ABSTRACT

Gender disparity stymies a country's development efforts around the world, particularly in developing countries. Gender equality means that men and women are both free to develop their strengths and make their own decisions. Gender inequality research is critical for society to improve its attitudes toward women. To test the effects of social issues on gender inequality, we used logistic regression to model. Using stepwise regression, we pick a collection of key issues. We were able to properly verify the model's sufficiency. We conclude that gender equality is unaffected by location, social class, or appreciation of women's role in religious activities. Age, nature of marriage, religious belief, dowry concerns, and equal voice in decision-making largely effects on gender inequality, according to the logistic regression. These factors can help you modify your social situation.

Keywords: *Gender equality, Logistic regression, Stepwise regression, Naïve-Bayes' classification, Decision tree.*

Introduction

In our country, from the ancient culture women is worshipped as a goddess. Women play important role in our society in every era. Afterward, traditional patriarchal customs and norms have relegated women to a secondary status within the household and workplace.

In the world, women face gender discrimination everywhere. But less gender equality creates problems for women as well as society. Women must have equal rights as men in the family as well as in society. The effect of gender discrimination on the different sectors and also social growth of any nation is the motivation of the study. In this paper, we are interested to find out the reasons behind gender inequality.

Gender equality refers to equitable treatment of men and women based on their individual needs.Gender inequality impacts on:

- Women's health over their lifetime.
- Women's educational attainment and economic conditions
- Women's Political empowerment

Although the constitution of India grants men and women equal rights, gender inequalities remain. Different inequalities occurred in various areas like Economic Inequalities, Educational Inequalities, Political Inequalities, Religious Participation, etc.

Literature Review

This research study looked into the reasons that contributed to gender disparity in Kyebi, Ashanti Region of Ghana.

Mavis Dako-Gyeke and Prince Owusu (2013) explored factors that perpetuated gender inequality in Kyebi, in the Ashanti Region of Ghana. Four causes were discovered to be at the root of gender inequality in the Kyebi community. These were cultural and traditional practices, gender socialization, poverty and discrimination in access to land.

The impact of relative socioeconomic position and perceptions of gender inequality in the marital relationship, in combination with domestic labour, on psychological discomfort was explored by L. Harryson et. al. (2012) using logistic regression analysis.

Sumanjeet S. (2016) explores numerous mechanisms that help countries close gender gaps as they grow. Why has the sex ratio gotten increasingly male-skewed with development, according to the author of the study? In addition, the author outlines several legislative options for addressing gender inequality.

Objectives

- To find out the causes of gender inequality in the Baramati Region.
- To study which factors most impact gender inequality.

- Develop a statistical model for gender inequality.

Methodology

In the Baramati region, the study depictsgender inequality in rural and urban areas. As a result, we chose four places for pilot research, two of which are rural regions and the other two are urban areas.

Malegaon and Jalgaon Supe are two villages in the rural area. Rui-paati and Kasba are both located inthe urban area. Then, for a pilot study, we create a questionnaire. We take a sample of 25 women from each area for the pilot study and analyse the data to develop the questionnaire.

We chose 40 questions for the final study, and the response question is "Is there gender equality in your house or your village?". These 40 questions are the 40 variables, and these variables are coded as Q1 to Q40. Out of which Q40 is the response variable. The primary data were collected with the approval of the Baramati municipality and the village Sarpanch from 371 women.

Our response is dichotomous, which means it has only two options: "Yes" or "No," therefore we used Logistic Regression to determine which factors are important in determining gender inequality.

Logistic Regression (LR):

Linear regression is typically employed when the response or dependent variable occurs on a continuous basis and the residual errors are normally distributed. We perform logistic regression when the dependent or response variable is not continuous. This regression shows a categorical or dichotomous variable regression model. The association between several independent variables and categorical dependent variables is investigated using logistic regression, often known as the logistic or logit model.

For example, Y can represent values such as "success" or "failure," "Yes" or "No," "Like" or "Dislike," all of which can be represented by the numbers 0 and 1.

The logistic regression model is:

$$\log i\ t(y)=\ln(\text{odds})=\ln\ [\![p/(1-p)]\!] = \alpha+\beta\_1\ X\_1+\beta\_2\ X\_2.+.....+\beta\_k\ X\_k$$

Where p is the probability of desirable outcome, $X_1,X_2,.....,X_k$is independent variable, α is intercept and $\beta_1,\beta_2,.....,\beta_k$ are regression coefficients.

The assumptions of Logistic regression as follows:

Errors are independent but not normally distributed.

Binary logistic regression is used if dependent or response variable is in binary in nature.

It can hold non-linear relationships between all variables, including independent and dependent variables, and it can also convert non-linear logs to LR.

In comparison to linear regression, this regression requires a large sample size because maximum likelihood approximations have low power for small samples.

Stepwise Regression:

The forward and backward selection approaches are combined in Stepwise Regression. Stepwise Regression is a kind modification of forward selection in which all regressors previously entered into the model are appraised using their partial F (or t) statistics at each step.Because of the relationships between it and other regressors in the equation, a regressor introduced earlier in the process may be redundant. If a variable's partial F (or t) statistic is less than F OUT (t OUT), it is removed from the model.

Stepwise regression requires the use of two cutoff values: one for entering variables and the other for deleting them.Some analysis prefers to choose F IN (or t IN) = F OUT (t OUT), although this is not necessary. Frequently we choose F IN (or t IN) > F OUT (t OUT), making it relatively more difficult to add a regressor than to delete one.

Confusion Matrix

A confusion matrix is a table that shows how well a classification model (or "classifier") performs on a set of test data for which the true values are known. It is useful to calculate Sensitivity and Specificity.

True Positives (TP): These are cases in which we predicted YES (they have the disease), and they do have the disease. Sensitivity can be calculated by formula as,

Sensitivity = TP/actual yes

True Negatives (TN): We classify NO, and they don't have the disease.

Specificity can be calculated by formula asSpecificity = TN/actual no

False Positives (FP): We classify YES, but they don't actually have the disease. (Also known as a "Type I error.")

False Negatives (FN): We classify NO, but they actually do have the disease. (Also known as a "Type II error.")

Statistical Analysis

Logistic Regression

This is suitable regression analysis used where response variable is occurs binary in nature. We apply this on our data.

Initial Model: fit1=glm(Q40 ~., family = binomial(link = "logit"), data = d1)

Table 1: Table for Coefficients of initial Logistic Regression model

(Intercept)	Q1	Q2	Q3	Q4	Q5	Q6	Q7
14.618378	0.008588	0.264677	-0.121454	-0.721143	0.117977	0.018066	-0.012936
Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15
-0.051111	0.082232	-0.179252	0.745625	-0.091411	-0.10271	-0.015852	-0.141884
Q16	Q17	Q18	Q19	Q20	Q21	Q22	Q23
0.725171	0.247032	0.564258	-0.145317	0.01547	-0.207072	0.460606	-2.212649
Q24	Q25	Q26	Q27	Q28	Q29	Q30	Q31
-0.086223	-14.07962	1.286363	-0.924711	0.05523	-0.484673	0.72801	1.121051
Q32	Q33	Q34	Q35	Q36	Q37	Q38	Q39
0.399346	-0.610032	0.537779	-0.016188	1.108892	0.387874	-0.133588	-0.139423

Residual Deviance: 336.8 AIC: 416.8

From the analysis of deviance, we observe that the Q11(Do you think that both men and women should share household works equally?), Q23(What is your opinion about dowry system?), Q26(Do you agree that both boys and girls are getting quality education equally from school?), Q36(Would women in decision making improve delivery of basic services?) are significantly affected on the response factor “Is there gender equality in your house or in your village?

Adequacy of the fitted model: p-value = 0.02488.

So, from the adequacy of the fitted model we observe that the p-value = 0.02488, is less than the level of significance(α). Hence, for getting better adequacy we use stepwise regression technique.

Stepwise Regression:

Final model of stepwise regression:

Q40 ~ Q2+Q4+Q11+Q18+Q21+Q23+Q26+ Q29+Q30+Q31+Q34+Q36

Table 2 : Table for Coefficients of Final model of stepwise regression.

(Intercept)	Q2	Q4	Q11	Q18	Q21	Q26
0.05615	0.26938	-0.77296	0.81079	0.65894	-0.29246	1.24195
Q29	Q30	Q31	Q34	Q36	Q23	
-0.48863	0.66106	0.96636	0.56877	1.21725	-2.0661	

Null Deviance: 413.4

Residual Deviance: 347 AIC: 373

Results and Discussion

The AIC of initial model is 416.8, but the AIC of final model of stepwise regression is 373. It shows that final model given by stepwise regression is better. The difference between the null and residual deviances indicates how well our model performs in comparison to the null model (a model with only the intercept).The wider this gap, the better.

Confusion Matrix

Response	FALSE	TRUE
0 (NO)	30	61
1 (YES)	12	268

True Positive Rate:

TP/ (actual yes) = (268/280) = 0.9571

It is also known as "Sensitivity" = 95%.

True Negative Rate:

TN/ (actual no) = (30/91) = 0.3296

It is also known as "Specificity" ≈ 33%

Adequacy of the Final model:

X-squared = 5.4227, df = 8, p-value = 0.7116

Hence, from the p-value = 0.7116 we can say that model is adequate.

Accuracy of the final model is 0.8032345 i.e. 80.32%. It means stepwise logistic regression, is very much useful for identifying the variates, that relates to gender inequality.

Conclusion

Gender inequality is still a reality in the Baramati region, if not the entire world. This situation will change soon if society's perspective on domestic chores, religious belief, dowry system, quality education, and the position of the decision maker in the home and society changes. In the Baramati region, it is hoped that by expanding education and opportunity for women, gender disparity would be eradicated.

References

1. Ashraf wani, Zulfqar, and Ruchi Ghosh Dastidar. 2018. "Gender Inequality-A Global Issue." International Journal of Research Available 5(19): 383–95.
2. L. Harryson ,M. Strandh, A. Hammarström. 2012 "Domestic Work and Psychological Distress−What Is the Importance of Relative Socioeconomic Position and Gender Inequality in the Couple Relationship?" PLoS ONE: 1-11.
3. Sajeda Amin and Anne R. Pebley (1994) "Gender Inequality within Households :The Impact of a Women's Development Programme in 36 Bangladeshi Villages." The Bangladesh Development Studies 22(2):121-154.
4. Sumanjeet, Singh. 2016. "The State of Gender Inequality in India." Gender Studies 15(1): 139–57.
5. Jayachandran, Seema. 2015. "The Roots of Gender Inequality in Developing Countries." Annual Review of Economics 7(1): 63–88.
6. Yerkes, Mara A. et al. 2020. "'Intelligent' Lockdown, Intelligent Effects? Results from a Survey on Gender (in)Equality in Paid Work, the Division of Childcare and Household Work, and Quality of Life among Parents in the Netherlands during the Covid-19 Lockdown." PLoS ONE 15-22.
7. Dako-Gyeke, Mavis, and Prince Owusu. 2013. "A Qualitative Study Exploring Factors Contributing to Gender Inequality in Rural Ghana." Mediterranean Journal of Social Sciences 4(1): 481–89.
8. Laura Ann McCloskey, Corrine Williams and Ulla Larsen. 2005. "Gender Inequality and Intimate Partner Violence among Women in Moshi. Tanzania." International Family Planning Perspectives, : 124-130.