33. **Evaluation of Linear Regression Models for Predicting Dosage of Medication Based on Age**

**Abstract:**

This research explores the relationship between age and the dosage of medication using simple linear regression models. Two visualizations of the regression fit, with confidence intervals and residuals against fitted values, are used to assess the predictive accuracy and reliability of the models. The study's primary objective is to examine whether age serves as a significant predictor for the dosage of medication required by patients. Results show that the linear regression model's fit is poor, as evidenced by the nearly flat regression line, low variability, and an unclear pattern in the residuals. These findings suggest that age alone does not adequately predict medication dosage, highlighting the need to include additional variables to improve the model's predictive performance.

**Introduction:**

Linear regression is a fundamental statistical technique widely used to predict the relationship between a dependent variable and one or more independent variables. In this study, we aim to understand the potential predictive power of age in determining the required dosage of medication for patients. The analysis relies on a dataset containing the age of individuals and their corresponding medication dosages. By employing simple linear regression, this research seeks to investigate whether a linear relationship exists between these variables and how well age alone can predict medication dosage.

**Methods:**

A simple linear regression model was constructed with age as the independent variable and dosage of medication as the dependent variable. The dataset includes ages ranging from 0 to 80 years and corresponding medication dosages ranging from 0 mg to 500 mg. The regression line, along with the confidence intervals, was plotted to visualize the model fit. Additionally, a residuals versus fitted values plot was created to assess the model's assumptions and potential biases.

**Results:**

**1. Simple Linear Regression: Dosage vs. Age**

The regression plot (Figure 1) shows the relationship between age and dosage of medication. The fitted line is nearly horizontal, indicating a lack of a strong linear relationship between age and dosage. The confidence intervals around the regression line are narrow, suggesting that the model has a consistent estimate of the mean dosage across different ages. However, this consistency appears to be misleading since the predicted dosage does not vary significantly with age.

**2. Residuals vs. Fitted Values**

The residual plot (Figure 2) displays the residuals against the fitted values of the dosage prediction. The residuals appear to be randomly scattered around the horizontal axis, with no apparent pattern. This randomness is typical of a model where the independent variable (age) does not significantly contribute to explaining the variance in the dependent variable (dosage). The residuals show a considerable spread, indicating a high level of unexplained variance by the model.

**Discussion:**

The analysis indicates that age is not a significant predictor of the required dosage of medication, as demonstrated by the nearly flat regression line in the linear model. The lack of a discernible trend in the residuals plot further confirms that the variability in medication dosage cannot be attributed to age alone. Given these results, the model's performance is poor, suggesting that additional variables, such as patient weight, height, gender, medical history, and genetic factors, may be necessary to improve the predictive power of the model.

The narrow confidence intervals around the regression line suggest that while the model is consistent in its prediction of a constant mean dosage, this consistency is not beneficial, as it does not capture any real relationship between the variables. The residual plot’s random distribution indicates that the assumptions of homoscedasticity and normality are not violated; however, the high variance implies that the model lacks sufficient explanatory power.

**Conclusion:**

The study demonstrates that using age as the sole predictor of medication dosage is insufficient, as indicated by the weak linear relationship and substantial unexplained variance. Future research should explore multiple regression models incorporating additional relevant variables to enhance the predictive accuracy. This approach could provide a more comprehensive understanding of the factors influencing medication dosage, leading to better healthcare decisions and personalized treatment strategies.

**Future Work:**

Given the limitations identified in this study, future work should focus on including additional patient characteristics and conducting multivariate regression analyses. These analyses could provide a more accurate model for predicting medication dosage by incorporating factors such as patient weight, health conditions, genetic predispositions, and lifestyle habits. Furthermore, exploring non-linear models, such as polynomial regression or machine learning approaches like random forests, could offer insights into more complex relationships that a simple linear model cannot capture.