Introduction: This is the analysis of SENIC, where infection probability is an outcome of different hospital facilities (as the exposure). This report describes and analyzes these variables, and their association.

Description: In a random sample of 113 US hospitals, the authors investigated the relationship between infection and potential facilities and services that are provided by the hospital. Descriptive analysis was performed to describe both probability of infection and of hospital facility. Pearson correlation was calculated to investigate the correlation between rate of infection and potential facilities and services. A linear regression model was design to investigate the association between probability of infection (dependent variable) and hospital facilities and services (independent variable).

Results**:** Among a total of 113 US hospitals, the mean number of facilities/services offered per hospital was 43.1 ± 15.2, and the mean infection probability was 4.3 ± 1.34. There was a significand positive correlation between probability of infection (dependent variable) and hospital facilities and services (B=0.41, P<0.001). On linear regression analysis, every 1 percent increase in the number of facilities offered per hospital led to a 0.036% increase in the probability of infection (P<0.001).

1. Numerical and Graphical Descriptions of the variables.



Briefly, the distribution of facility is seen above as a normal distribution.

This is similar to that of the probability of infection, which is also a normal distribution.



This shows that probability of infection has a relatively normal distribution as well.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variable | N | Mean | Std Dev | Minimum | Maximum |
| facility prob\_inf | 113 113 | 43.2 4.4 | 15.2 1.3 | 5.7 1.3 | 80.0 7.8 |

1. Create a Scatter plot of the IV and DV



As we can see, there is a pattern of positive correlation between facility and infection probability.

1. Correlation between variable

The correlation coefficient (rho) is 0.41260, with an associated p.value of <0.000

1. State the SLR Model

The SLR model is stated below.

*Prob\_inf = B\*facility + B0 + E*

1. SLR Regression coefficients

The regression coefficient for facility is 0.03640.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | DF | Parameter Estimate | Standard Error | P-value |
| **facility** | 1 | 0.03640 | 0.00763 | <.0001 |

1. Test the significance of the SLR model.
2. Describe the relationship between the variables

For both F and G, we can see that there is a pattern in this figure between facility and probabilyt of infection. The pattern is a positive relationship. The value is significant, as described in H. This suggests that the probability of infection increases with high percentage of facilities offered by each hospital.

