

Linear Statistical Models

Video 78 - Probit Solution

Shreeja Bhakat, BS2043

29 September 2022

1 Introduction

In an experiment, researchers are interested in predicting the lethal dose of a poison on a mouse. Consider the lethal dose to be X and the applied dose to be d . Here, X is a latent variable, i.e., they are not directly observed but are rather inferred through a mathematical model from other variables that are observed.

2 Theory

In the experiment, n mice are taken and each is given the same amount of dose d . Then it is observed how many of them die. In this case, we can say that if the mouse is dead then $X \leq d$ and if it is alive then $X > d$. Say K many are dead. Then,

$$P[X \leq d] \approx \frac{K}{n}.$$

Here, X is a random variable denoting the lethal dose of the poison. Consider $X \sim N(\mu, \sigma^2)$. Therefore, we can say,

$$\Phi\left(\frac{d-\mu}{\sigma}\right) \approx \frac{K}{n}$$

Except μ and σ , all the other values are known. Φ is the cumulative distribution function of the Standard Normal distribution.

Φ is a smooth function. Hence, we can take Φ^{-1} on both sides, which gives us,

$$\frac{d-\mu}{\sigma} \approx \Phi^{-1}\left(\frac{K}{n}\right)$$

The experiment is carried out for different values of dose d . Therefore, for the doses d_1, d_2, \dots, d_j , we have corresponding values of number of dead mice K_1, K_2, \dots, K_j . Each dose is given to n mice.

| Dose | No. of Mice dead |
|---------|------------------|
| d_1 | K_1 |
| d_2 | K_2 |
| \dots | \dots |
| d_j | K_j |

Next, we plot the points $(d_i, \Phi^{-1}(\frac{K_i}{n}))$. This should give us a more or less straight line. In case it does not, then we can say that the latent variables X_i 's are not coming from a Gaussian distribution.

In case of a straight line, we can fit a least squares line and from the slope and intercept of the fitted line we can estimate μ and σ .

Now, we define a function called Probit as follows,

$$\text{Probit}(x) = \Phi^{-1}(x)$$

Earlier, D.J.Finney defined it as,

$$\text{Probit}(x) = \Phi^{-1}(x) + 5$$

The main reason for doing this was that, most of the calculations were done by hand and negative numbers made it more difficult. Φ^{-1} takes values mostly between -3 and +3. Thus, adding 5 to the numbers, ensured that all of them are positive making calculations easier.

3 Conclusion

Thus, in this video, we introduced to the Probit function and the way of estimating the parameters of a latent variable.