Two of the Seven Pillars

Austin Pesina

2021-09-26

P-Value

A p-value is the probability that the results obtained will be "at least as favorable to the alternative hypothesis as our current data set, if the null hypothesis were true"[1].

Experimental Design

The p-value is related to experimental design because it determines a relationship with the hypothesis we will test. In relation to the p-value, we often determine an alpha value of 0.05, which basically says there is a less that 5% chance that our results are due to randomness. The smaller the p-value, the more likely our results were not obtained due to randomness.

Residual

Calculations of the p-value test the risk of making a Type I error, where we mistakenly reject the null hypothesis when we should have accepted it. By choosing as small a p-value as possible, we lessen the risk of making a Type I error.

R^2

The R^2 value "explain[s] the strength of a linear fit". In a regression, we use the R^2 to explain the percentage that variance is explained by each predictor. The R^2 value will fall between 0 and 1. The closer to 1, the more the model explains.

Experimental Design

The R^2 is used in regression analysis to determine if the predictors in our model accurately explain the results. The R^2 of a model can be applied to a prediction model, and the higher the value of the R^2, the more likely the model is to correctly predict the estimated values.

Likelihood

The R^2 serves as an estimator of how likely a model is to be correct. The likelihood increases with the more certainty of explanation. The risk of increased likelihood is that the model runs the risk of overfitting the data it is being tested on and will not explain other data sets with similar predictors as well.

References

[1] Diez, D. M., Barr, C. D., & Cetinkaya-Rundel Mine. (2019). OpenIntro statistics (Fourth). OpenIntro, Inc.