

# Sketch of a salary regression model



## 1 *Construct an analysis dataset.*

After company level data is received, a dataset that is organized into one table composed of uniquely defined rows and columns is constructed from the raw electronic data. The most common datasets are constructed as employee data snapshots. In these types of datasets, referred to as cross-sectional datasets, each row of the dataset contains employment factors for one employee at a given point in time such as the end of the year. Each column contains information about one specific employment factor such as the salary grade at a given point in time.

## 2 *Estimate mathematical model.*

The analysis dataset constructed in the first step of the analysis, is used to construct a mathematical representation of the employment factors that are correlated with an employee's salary. The mathematical representation of these factors is referred to as a salary regression model. In its purest form, the salary regression model is nothing more than a simple equation that describes what (and how much) different employment factors affect a person's salary.

## 3 *Make inferences.*

The mathematical representation of the salary process (salary regression model) is typically used to make two types of inferences. First, the salary regression model can be used to predict an employee's salary given the employee's salary grade, job location, and other relevant factors. Second, and more common in litigation settings, the salary regression model is used to make inferences about the importance of any disparity in salary levels between protected group and non-protected group members. The significance of the salary disparity is inferred by determining how likely the salary disparity would have been generated by sheer random chance.

The fundamentals of statistical analysis are not difficult to understand. Many statistical analysis reports are heavy with mathematical symbols, but do not take much more than basic arithmetic to gain a basic understanding of a statistical analysis, including more advanced techniques such as linear regression.



### **Questions or comments:**

Dwight Steward, Ph.D.  
Economist and Director

[dsteward@employstats.com](mailto:dsteward@employstats.com)  
[www.employstats.com](http://www.employstats.com)