

# Project 3: NFL Salary Distribution

*Jonathan Fogel*

*11/20/2020*

Author: Jonathan Fogel

## Introduction

When compared to the median salary in the US, \$50,000, salaries for professional athletes are often unfathomably large. American Football, in particular, is full of players making upwards of \$10,000,000 yearly, like Carson Wentz, who averages about \$32,000,000. With many players' contracts being available to the public, one may wonder what the distribution of salaries among NFL players may look like, and if it fits any model particularly well. In this paper, contract data compiled by Spotrac, an independent media company, will be analyzed to determine if it follows a normal distribution. A normal distribution seems appropriate because salary is a continuous variable, and it is a reasonable possibility that the salaries tend to be mostly lie near an average value.

## NFL Player Salary Distribution

Contracts for all current players for the Philadelphia Eagles, the Pittsburgh Steelers, and the New England Patriots were compiled, and yearly salary was calculated by dividing the total contract payment by its length. These 3 teams are being used as a sample because they are all near the average team salary for the NFL, so they will presumably be an accurate representation of the league as a whole.

First, a mean and standard deviation is computed for yearly average salary.

```
mean(~AvgSalary, data=Wages)
```

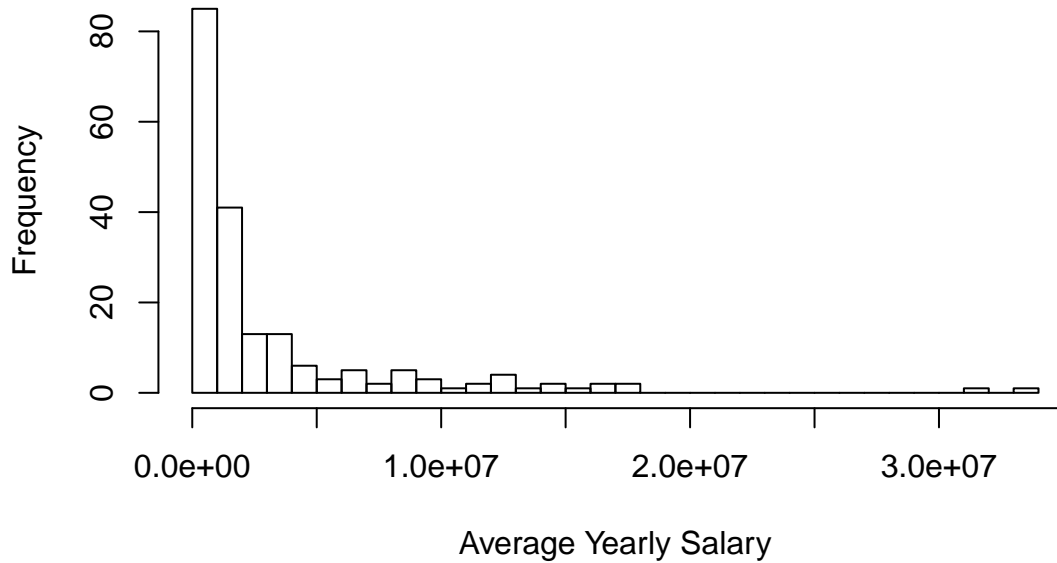
```
## [1] 3301960
```

```
sd(~AvgSalary, data=Wages)
```

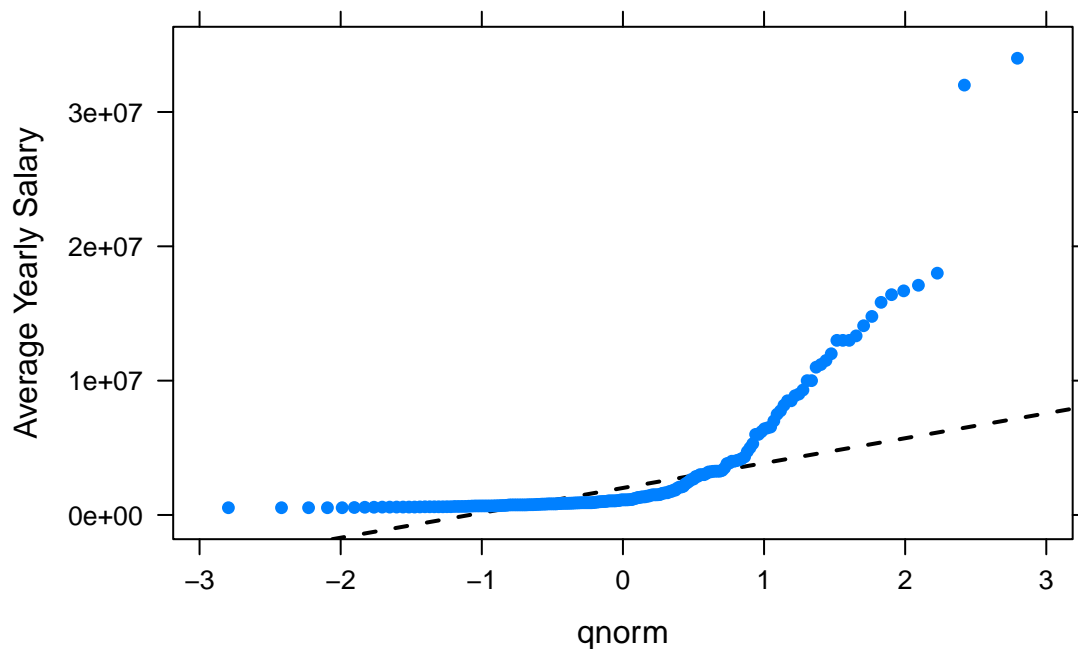
```
## [1] 4941805
```

So the sample mean salary for the is \$3,301,960, and the associated standard deviation is \$4,941,805. Now, the salary information is binned into equasized bins spanning the entire range. Each bin will cover \$1,000,000, starting at \$0 and ending at \$34,000,000. A histogram showing the observations in each bin is shown. To illustrate the correspondance between this and a normal distribution with appropriate mean and standard deviation, a quantile-quantile plot is also shown.

### Density of NFL Player Yearly Salaries



### Q-Q Plot of NFL Player Yearly Salaries



Clearly, the data does not appear normally distributed. There is a very dense section near \$800,000, and the density trails off in the positive direction. Looking at the Q-Q Plot, the data strays far away from the line, and is not linear. This also indicates that the data is not normally distributed.

To quantitatively test how normal the data is, a chi-squared test is used as well. The probability density for

each bin is recorded, and compared to the probability density of each range of salaries. In this test, the Null hypothesis is that the data is, in fact normally distributed, and the alternate hypothesis is that the data is not normally distributed. The results of this test are shown below.

```
chisq.test(x=Wages2$Counts2, rescale.p = TRUE, p=Wages2$NormProb)
```

```
## Warning in chisq.test(x = Wages2$Counts2, rescale.p = TRUE, p =  
## Wages2$NormProb): Chi-squared approximation may be incorrect  
  
##  
## Chi-squared test for given probabilities  
##  
## data: Wages2$Counts2  
## X-squared = 43868, df = 33, p-value < 2.2e-16
```

The p-value is extremely small, and well under 0.05. This means we can confidently reject the null hypothesis at a 95 percent confidence level. So we are confident that reject the idea that the data is normally distributed.

## Conclusion

Using a chi-squared test, there is evidence that the salaries of NFL Players are not normally distributed.