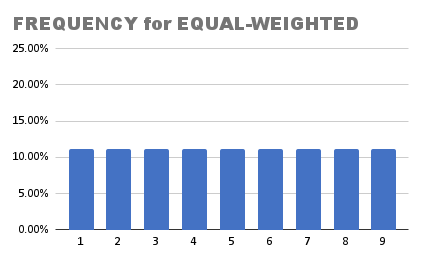
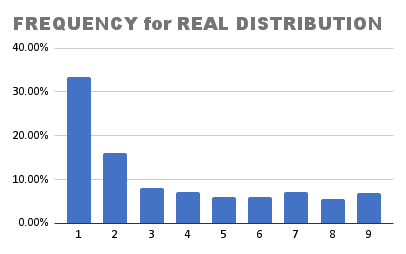
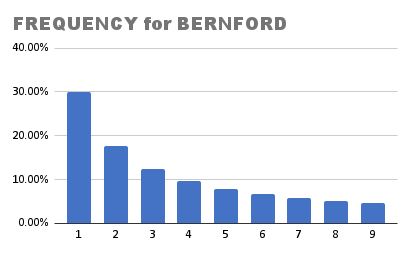
**Assignment 4 Bernford Law**

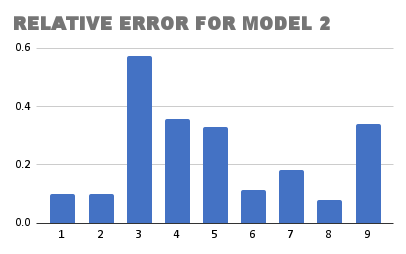
### **1. Plot 3 histograms for the frequencies for real distribution, equal-weight and Bernford (for each digit)**





### **2. plot 2 histograms for the relative errors for Models 1 and 2 (for each digit)**

### 



### **3. Compute RMSE (root mean squared error) for model 1 and 2. Which model is closer to the real distribution?**

RMSE for Model 1 = 0.08611697019

RMSE for Model 2 = 0.02299826213

Because RMSE for Model 2 is smaller than RMSE for Model 1, Model 2 is closer to the real distribution

### 

### **4. Take 3 categories of your choice For each of these categories do the following: (a) compute F, P and π. (b) using RMSE as a ”distance” metric, for which of these chosen three countries is the distribution ”closest” to equal weight P?**

### Jewellery-Women

| F |
| --- |
| 24.74% |
| 18.90% |
| 16.72% |
| 10.31% |
| 7.75% |
| 4.12% |
| 5.18% |
| 3.85% |
| 4.51% |

| P |
| --- |
| 11.11% |
| 11.11% |
| 11.11% |
| 11.11% |
| 11.11% |
| 11.11% |
| 11.11% |
| 11.11% |
| 11.11% |

| 𝝿 |
| --- |
| 30.10% |
| 17.60% |
| 12.50% |
| 9.70% |
| 7.90% |
| 6.70% |
| 5.80% |
| 5.10% |
| 4.60% |

RMSE for Model 1: 0.07227688273

RMSE for Model 2: 0.02521771864

Lingerie&Nightwear-Women

| F |
| --- |
| 24.53% |
| 16.10% |
| 9.51% |
| 8.82% |
| 10.07% |
| 7.81% |
| 8.35% |
| 5.33% |
| 2.60% |

| P |
| --- |
| 11.11% |
| 11.11% |
| 11.11% |
| 11.11% |
| 11.11% |
| 11.11% |
| 11.11% |
| 11.11% |
| 11.11% |

| 𝝿 |
| --- |
| 30.10% |
| 17.60% |
| 12.50% |
| 9.70% |
| 7.90% |
| 6.70% |
| 5.80% |
| 5.10% |
| 4.60% |

RMSE for Model 1: 0.06130237443

RMSE for Model 2: 0.02570901269

Watches-Women

| F |
| --- |
| 26.30% |
| 17.71% |
| 10.47% |
| 9.58% |
| 7.86% |
| 8.58% |
| 6.57% |
| 4.35% |
| 4.57% |

| P |
| --- |
| 11.11% |
| 11.11% |
| 11.11% |
| 11.11% |
| 11.11% |
| 11.11% |
| 11.11% |
| 11.11% |
| 11.11% |

| 𝝿 |
| --- |
| 30.10% |
| 17.60% |
| 12.50% |
| 9.70% |
| 7.90% |
| 6.70% |
| 5.80% |
| 5.10% |
| 4.60% |

RMSE for Model 1: 0.06692373769

RMSE for Model 2: 0.01608301381

### **5. Discuss your findings**

Compared to Model 1, model 2 which is from the Bernnford law is much closer to the real distribution.