MME for Binomial(n,p) Derivations for 3.4:

$E[Bin(n,p)] = np$ (1)

$\hat{E}[X] = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$ (2)

$Var[Bin(n,p)] = np(1-p)$ (3)

$\hat{Var}[X] = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}^2 = S_X$ (4)

Equating (1) and (2) yields $\hat{n}\hat{p} = \overline{X} \rightarrow n = \frac{\overline{X}}{\hat{p}}$

Equating (3) and (4) yields $np(1-p) = S_X$

plugging in $n = \frac{\overline{X}}{p}$ yields $\overline{X}(1-p) = S_X$

$1-p = \frac{S_X}{\overline{X}} \rightarrow 1 - \frac{S_X}{\overline{X}} = \hat{p}_{MME}$

Similarly, $\hat{n} = \frac{\overline{X}}{\hat{p}} \rightarrow \hat{n}_{MME} = \frac{\overline{X}}{1 - \frac{S_X}{\overline{X}}}$ ☒

MME for Poisson(λ):

$E[Poisson(\lambda)] = \lambda$ (1) Equating (1) & (2) yields:

$\hat{E}(X) = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$ (2) $\lambda_{MME} = \overline{X}$ ☒

MME for Geometric(p): Equating (1) & (2) yields:

$E[Geometric(p)] = \frac{1}{p}$ (1)

$\hat{E}[X] = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$ (2) $\frac{1}{p} = \overline{X} \rightarrow 1 = \overline{X}p \rightarrow \hat{p}_{MME} = \frac{1}{\overline{X}}$

☒

3.4

- 1-Sample KS with Binomial assumption $H_0: F_D \equiv F_X$ vs $H_1: F_D \not\equiv F_X$ has a statistic of 0.97. Since $0.97 > 0.05$ we reject H_0 and conclude the sample data do not follow a $Bin(\hat{n}_{mme}, \hat{p}_{mme})$ distribution.

- 1-Sample with Poisson assumption. Same hypothesis. Again, since $0.72 > 0.05$, we reject H_0 and conclude the sample data do not follow a $Pois(\hat{\lambda}_{mme})$ distribution

- 1-Sample with Geometric assumption. Same hypothesis. Again, since $0.97 > 0.05$, we reject H_0 and conclude the sample data do not follow a $Geo(\hat{p}_{mme})$ distribution.

- For the 2-sample KS test with

$$H_0: F_{\text{Monthly Housing cost per change}} \equiv F_{\text{Rent of Primary Residence per change}}$$

vs

$$H_1: F_{\text{Monthly Housing cost per change}} \not\equiv F_{\text{Rent of Primary Residence per change}}$$

the test statistic is 0.36. Since $0.36 > 0.05$, we reject the null hypothesis and conclude the distributions are not the same.

- For the Permutation test with the same hypothesis as the 2-Sample KS test, the test statistic is 0.228. Since $0.228 > 0.05$, we again reject the null hypothesis and conclude the samples do not follow the same distribution.

3.5.3

- For CPI Energy per change, the sample quantiles do not align with the theoretical quantiles based on the QQ-plot. The Shapiro-Wilk's W statistic confirms the conclusion from the QQ-plot. The p-value of this statistic is 3.6×10^{-15} which is much smaller than 0.05, thus allowing us to reject the null hypothesis that the data are drawn from a normal distribution.

- For CPI All Items per change, the sample quantiles do not align with the theoretical quantiles based on the QQ-plot. The Shapiro-Wilk's W statistic confirms the conclusion from the QQ-plot. The p-value of this statistic is 1.2×10^{-12} which is much smaller than 0.05, thus allowing us to reject the null hypothesis that the data are drawn from a normal distribution.

3.5.4

Experiment 1: SSE = 0.00904
Experiment 2: SSE = 0.00898
Experiment 3: SSE = 0.00888
Experiment 4: SSE = 0.00888

All of the SSEs are around the same value. Based soley on the SSEs calculated, adding the variable Rent of Primary Residence per change does not improve the model as the SSE of experiments 3 and 4 are the same. Since experiment 3 has a very marginally smaller SSE than experiments 1 and 2, I would suggest this model is most relevant.

$$y = \beta_0 + \beta_1 \underset{\text{per change}}{\text{Dollar Purchasing}} + \beta_2 \underset{\text{per change}}{\text{CPI Energy}} + \beta_3 \underset{\text{per change}}{\underset{\text{Housing}}{\text{Monthly}}} + \varepsilon_i$$