Problem Set #4

1b.

$$\theta = \frac{\int\int (x-\mu_x)(y-\mu_y)dF(x,y)}{\sqrt{\int (x-\mu_x)^2 dF(x) \int (u-\mu_y)^2 dF(y)}}$$

$$\hat{\theta}_n = \frac{\frac{1}{n}\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\frac{1}{n}\sum (X_i - \bar{X})^2 * \frac{1}{n}\sum (Y_i - \bar{Y})^2}}$$

2a. We can use the first two moments:

$$\int_\alpha^\beta x f(x) dx = \int_\alpha^\beta x \left(\frac{1}{\beta-\alpha}\right) dx = \left[\frac{x^2}{2(\beta-\alpha)}\right]_\alpha^\beta$$
$$= \frac{\beta^2-\alpha^2}{2(\beta-\alpha)} = \hat{m}_1$$

$$\int_\alpha^\beta x^2 f(x) dx = \int_\alpha^\beta x^2 \left(\frac{1}{\beta-\alpha}\right) dx = \left[\frac{x^3}{3(\beta-\alpha)}\right]_\alpha^\beta$$
$$= \frac{\beta^3-\alpha^3}{3(\beta-\alpha)} = \hat{m}_2$$

Solving this system of equations:

$$\alpha = \hat{m}_1 - \sqrt{3(\hat{m}_2 - \hat{m}_1^2)}$$

$$\beta = \hat{m}_1 + \sqrt{3(\hat{m}_2 - \hat{m}_1^2)}$$

Where

$$\hat{m}_1 = \frac{1}{n}\sum_{i=1}^n X_i \text{ and } \hat{m}_2 = \frac{1}{n}\sum_{i=1}^n X_i^2$$

2b. We have the likelihood function:

$$L(X_1, \ldots, X_n; \alpha, \beta) = \left(\frac{1}{\beta-\alpha}\right)^n$$

Under the logic of MLE's, we want to find the combination of α and β that will maximum this equation, so we want to minimize the denominator. However, we must pick values that will allow us to explain the entire dataset. All data values must be in the range of α to β or else the likelihood of getting a data value could be zero. Taking these attributes into account we get:

$$\alpha = \min(X_1, \dots, X_n)$$

$$\beta = \max(X_1, \dots, X_n)$$

3a. We know that the mean of the Uniform Distribution gives us

$$\mu = \frac{\alpha + \beta}{2}$$

Therefore, we can use the equivariance property (Lecture 10, equation 32) to find the MLE of μ using our estimates from 2b:

$$\mu_{MLE} = \frac{\min(X_1, \dots, X_n) + \max(X_1, \dots, X_n)}{2}$$

3b. Monte Carlo simulation is attached as code. Analytically:

$$MSE = bias^2 + se^2$$

$$bias = E[\bar{X}] - \mu = \mu - \mu = 0$$

$$se^2 = \frac{Var(\hat{\mu}_n)}{n} = \frac{(3-1)^2}{12(10)} = \frac{1}{30}$$

Since the population size is very large.

$$MSE = 0 + \frac{1}{30} = \frac{1}{30}$$

The results are different.

4a. The expected value will be the area under the curve from zero to infinity:

$$\psi = \int_0^\infty N(\theta, 1)dx$$

$$\psi = \int_0^\infty \frac{1}{\sqrt{2\pi}} e^{\frac{-(x-\theta)^2}{2}} dx$$

To find the MLE of ψ, we can find the MLE of θ and plug this into our equation, since the equivariance property (Lecture 10, equation 32) states this is true. As shown in the notes (Lecture 9, equation 12), we know that the MLE of θ would be $\bar{X}$. Therefore,

$$\psi_{MLE} = \int_0^\infty \frac{1}{\sqrt{2\pi}} e^{\frac{-(x-\bar{X})^2}{2}} dx$$

4b. Finding a CI for ψ, we can find a 95% CI for θ and plug this into ψ:

$$I_n = \hat{\theta}_n \pm z_{\frac{\alpha}{2}} \widehat{se} = \bar{X}_n \pm 1.96 * \frac{1}{\sqrt{n}}$$

Since $N \to \infty$. Plugging this in:

$$I_\psi = \int_0^\infty \frac{1}{\sqrt{2\pi}} e^{\frac{-\left(x-\left(\bar{X}_n \pm 1.96 * \frac{1}{\sqrt{n}}\right)\right)^2}{2}} dx$$

5a. Finding the PDF:

$$CDF = P(\max\{x_1, \dots, x_n\} \leq x) = P(x_1 \leq x, x_2 \leq x, \dots, x_n \leq x)$$

$$\prod_{i=1}^{n} P(x_i \leq x)$$

Which is any value between zero and x:

$$\prod_{i=1}^{n} \frac{x-0}{\theta} = \left(\frac{x}{\theta}\right)^n$$

We differentiate to get the PDF:

$$PDF = \frac{nx^{n-1}}{\theta^n}$$