HW 5

Sonya Eason

1.

A Pareto Type 1 disribution has the following pdf:

$$f_x = \frac{\alpha x_m^\alpha}{x^{\alpha+1}} \mathbf{1}_{x \geq x_m}$$

To determine if the Pareto Type I distribution is a member of the exponential family, I beg by exponentiating the log of the given PDF.

$$\exp\left(\log\left(\frac{\alpha x_m^\alpha}{x^{\alpha+1}}\right)\right) \mathbf{1}_{x \geq x_m}$$

$$\exp\left(\log(\alpha) + \alpha \log\left(\frac{x_m}{x}\right) - \log(x)\right) \mathbf{1}_{x \geq x_m}$$

$$\frac{1}{x} \exp\left(\log(\alpha) + \alpha \log\left(\frac{x_m}{x}\right)\right) \mathbf{1}_{x \geq x_m}$$

Yes, this is a member of the exponential family because we can write in the form:

$$f(x) = h(x) \exp\left(\eta(\theta) T(x) - \psi(\theta)\right)$$

or

$$f(x) = h(x) \exp\left(\eta(\theta) T(x) - A(\eta)\right)$$

$$h(x) = \frac{1}{x} \mathbf{1}_{x \geq x_m}$$

$$\eta(\alpha) = \alpha$$

$$T(x) = \log\left(\frac{x_m}{x}\right)$$

$$\psi(\alpha) = -\log(\alpha)$$

$$A(\eta) = -\log(\eta)$$

2.

a)

$$\log\left(\frac{p}{1-p}\right) = X_{n\times p}\beta_{p\times 1},$$

Note that (X_i) is a row vector where its dimension is (1 ×p). Provide the log-likelihood function as a function of the regression parameters. That is, show that

The log-likelihood function for logistic regression is

$$\mathcal{L}(p) = \prod_{i=1}^{n} f(y_i)$$

$$= \prod_{i=1}^{n} p^{y_i}(1-p)^{1-y_i},$$

$$\log \mathcal{L}(p) = \sum_{i=1}^{n} \left[y_i \log(p) + \log(1-p) - y_i \log(1-p)\right]$$

$$= \sum_{i=1}^{n} \left[y_i \log\left(\frac{p}{1-p}\right) + \log(1-p)\right].$$

Recall that

$$\log\left(\frac{p}{1-p}\right) = X\beta,$$

so the log-likelihood becomes

$$\log \mathcal{L}(p) = \sum_{i=1}^{n} \left[y_i X_i \beta - \log(1 + \exp(X_i\beta))\right].$$

In matrix notation, we can write the expression as

$$\log \mathcal{L}(\beta) = y^T X\beta - \mathbf{1}^T \log(1 + \exp(X\beta)).$$

b)

$$\nabla_\beta \log \mathcal{L}(p) = \sum_{i=1}^{n} \left[y_i \mathbf{X}_i - \frac{\exp(\mathbf{X}_i^T \beta)}{1 + \exp(\mathbf{X}_i^T \beta)} \mathbf{X}_i \right].$$

You may also write the expression in matrix notation as follows:

Define

$$\pi_i = \frac{\exp(\mathbf{X}_i^T \beta)}{1 + \exp(\mathbf{X}_i^T \beta)}.$$

Then,

$$\nabla_\beta \log \mathcal{L}(p) = \mathbf{X}^T (\mathbf{y} - \pi).$$

c)

$$\nabla_\beta^2 \log \mathcal{L}(p) = -\sum_{i=1}^{n} \left[\frac{1}{1 + \exp(\mathbf{X}_i^T \beta)} \cdot \frac{\exp(\mathbf{X}_i^T \beta)}{1 + \exp(\mathbf{X}_i^T \beta)} \mathbf{X}_i \mathbf{X}_i^T \right].$$

Let $W = diag(\pi_1(1 - \pi_1), \dots, \pi_n(1 - \pi_n))$
It follows that the Hessian can be written as

$$-\mathbf{X}^T W \mathbf{X}.$$

d)

Logistic Newton Raphson Algorithm

```
logistic_Newton_Raphson <- function(x, y, b_init = rep(0, ncol(x)),
                                    tol = 1*10^(-8))
{
  change <- Inf
  b_old <- b_init

  while(change>tol)
  {
    eta <- x %*% b_old
```

```
    pie <- exp(eta) / (1 + exp(eta))
    w <- diag(as.vector(pie * (1 - pie)), nrow = length(pie), ncol = length(pie))

    b_new <- b_old + solve(t(x)%*% w %*% x) %*% t(x) %*% (y-pie)

    change <- sqrt(sum((b_new-b_old)^2))

    b_old <- b_new
  }

  b_new
}
```

e)

```
df <- read.csv('https://stats.idre.ucla.edu/stat/data/binary.csv')
```

```
#pre-processing

y <- df$admit

X <- as.matrix(df[, c("gre", "gpa", "rank")])
int <- rep(1, nrow(df))
X <- cbind(int, X)
```

```
logistic_Newton_Raphson(X, y)
```

```
            [,1]
int  -3.44954840
gre   0.00229396
gpa   0.77701357
rank -0.56003139
```

This matches the glm.

3.

Proposed question

The Rayleigh distribution can be used to model wave displacement.

X_i represents the displacement of a wave, which is i.i.d. and follows the following distribution

$$f_x = \frac{x}{\sigma^2} e^{-x^2/2\sigma^2}$$

Find the sufficient statistic of a sample of n displacement waves. Use the canonical parameterization to calculate the MLE for σ^2.

Citation: Utilized the MIT open courseware notes at https://ocw.mit.edu/courses/18-655-mathematical-statistics-spring-2016/resources/mit18_655s16_lecnote7/ for the canonical parameterization.

Solution below:

$$f_x = \frac{x}{\sigma^2} e^{-x^2/2\sigma^2}$$

We exponentiate the log of this to gain try to find the exponential family formulation:

$$exp(log(\frac{x}{\sigma^2} e^{-x^2/2\sigma^2}))$$

We are able to rewrite the density to see the exponential family formulation:

$$f_x = \frac{1}{x} e^{-log(\sigma^2) - \frac{x^2}{2\sigma^2}}$$

and recover the parameters as:

$$h(x) = \frac{1}{x}$$

$$T(x) = x^2$$

$$\eta = -\frac{1}{2\sigma^2}$$

$$A(\eta) = log(-\frac{1}{2\eta})$$

The sufficient statistic of the sample would be

$$T(X) = \sum_{i=1}^{n} {X_i}^2$$

We can recover the MLE through

$$\nabla_\eta A(\eta) = \frac{1}{n} \sum_{i=1}^{n} T(X_i)$$

We can take the gradient of the log partition function:

$$\nabla_\eta A(\eta) = \frac{\frac{1}{2\eta^2}}{\frac{-1}{2\eta}} = \frac{-1}{\eta} = 2\sigma^2$$

We can utilize the sufficient statistic and method of moments:

$$\frac{1}{n}\sum_{i=1}^{n} T(X_i) = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

Then we set the gradient of log partition function to the empirical mean of the sufficient statistic.

$$2\sigma^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

This is the MLE:

$$\hat{\sigma}^2_{MLE} = \frac{1}{2n}\sum_{i=1}^{n} X_i^2$$

Rubric (1) finds the sufficient statistic. (1) shows evidence of correct exponential family understanding (i.e., may have messed up the fact that we need the sample sufficient statistic, but still used exponential family concept to attempt to recover sufficient statistic). (1) finds the correct MLE. (1) utilizes the canonical parameterization to find the MLE.