Econometrics HW5

Timothy Schwieg

a

$$\sum_{i=1}^{\infty} p_Y(y|\theta) = \sum_{i=1}^{\infty} \frac{-\theta^y}{y\log(1-\theta)} = \frac{1}{\log(1-\theta)} \sum_{i=1}^{\infty} \frac{(-1)^{y+1}(-\theta)^y}{y} = \frac{\log(1-\theta)}{\log(1-\theta)} = 1$$

b

$$\mathbb{E}[Y] = \sum_{i=1}^{\infty} \frac{-y\theta^y}{ylog(1-\theta)} = \sum_{i=1}^{\infty} \frac{-\theta^y}{\log(1-\theta)} = \frac{-1}{\log(1-\theta)} \sum_{i=1}^{\infty} \theta^y = \frac{-\theta}{(1-\theta)\log(1-\theta)}$$

c

$$\mathbb{E}[Y^2] = \sum_{i=1}^{\infty} \frac{-y\theta^y}{\log(1-\theta)} = \frac{-\theta}{\log(1-\theta)} \sum_{i=1}^{\infty} y\theta^{y-1} = \frac{-\theta}{\log(1-\theta)} \frac{\partial}{\partial\theta} \sum_{i=1}^{\infty} \theta^y = \frac{-\theta}{\log(1-\theta)} \frac{\partial}{\partial\theta} \frac{\theta}{1-\theta} = \frac{-\theta}{\log(1-\theta)(1-\theta)^2}$$

$$\mathbb{V}(Y) = \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 = \frac{-\theta}{\log(1-\theta)(1-\theta)^2} - \frac{\theta^2}{\log(1-\theta)^2(1-\theta)^2} = \frac{-\theta\log(1-\theta) - \theta^2}{\log(1-\theta)^2(1-\theta)^2}$$

d

Let $g(y) = \frac{1}{y}, h(\theta) = \frac{-1}{\log(1-\theta)}, \eta(\theta) = \log(\theta), \tau(y) = y$, this clearly implies the log-series is a member of the exponential family.

Its sufficient statistic is given by: $T(y) = \sum_{n=1}^{N} \tau(y_n) = \sum_{n=1}^{N} y_n$.

e

$$L(y_n|\theta) = \prod_{n=1}^{N} \frac{-\theta^{y_n}}{y_n \log(1-\theta)}$$

$$f(\theta) = \log(L(y_n|\theta)) = \sum_{n=1}^{N} y_n \log(\theta) - \log(y_n) - \log(-\log(1-\theta))$$

$$f'(\theta) = \sum_{n=1}^{N} \frac{y_n}{\theta} + \frac{1}{(1-\theta)\log(1-\theta)} = 0$$

$$\sum_{n=1}^{N} y_n + \sum_{n=1}^{N} \frac{\theta}{(1-\theta)\log(1-\theta)} = 0$$

$$\bar{Y_N} + \frac{\theta}{(1-\theta)\log(1-\theta)} = 0$$

The maximum liklihood estimator defined by this relationship can be given by the solution $\hat{\theta}$ to this equation. We shall find it using a Newton Step. Voraciously applying the quotient rule:

$$f''(\theta) = \frac{\log(1-\theta) + \theta}{(1-\theta)^2 \log(1-\theta)^2}$$

f

We wish to show that $f'(\theta)$ is a monotonically decreasing function of θ. We will show that $f''(\theta) < 0$ we may first note that that since the denominator is squared, it is non-negative, so we need only concern ourselves with the numerator. Taking the taylor series expansion of $\log(1-\theta)$ and applying lagrange's remainder theorem, we may note that the numerator is equal to: $\theta - \theta - \frac{\xi^2}{2}$ which is strictly less than zero. This implies that $f'(\theta)$ is a decreasing function.

g

The Newton steps we will take will being at some suitable starting guess θ_0 and continue based on the following rule:

$$\theta_{k+1} = \theta_k - \frac{f'(\theta_k)}{f''(\theta_k)} \theta_{k+1} = \theta_k - \frac{\left(-\theta + \bar{Y_n}\left(\theta - 1\right)\log\left(-\theta + 1\right)\right)\left(\theta - 1\right)\log\left(-\theta + 1\right)}{\theta \log\left(-\theta + 1\right) + \theta - \left(\theta - 1\right)\log\left(-\theta + 1\right)}$$

h

Since it is known that $f''(\theta) < 0$ we may apply the implicit function theorem to the equation $f'(\hat{\theta}) = 0$ which tells us that $\exists l(.) s.t. \hat{\theta} = l(\bar{Y_N})$ which is continous and differentiable in a neighborhood around $\bar{Y_N}$. We may apply Slutsky's theorem to this continous function to reach: $plim\hat{\theta} = plim l(\bar{Y_N}) = l(plim\bar{Y_N}) = l(\mu^0) = \theta$.

i

A consistent estimator of the variance of $\hat{\theta}$ using Fisher's Information matrix is the negative inverse Hessian.

$$V(\hat{\theta}) \sim -[f''(\theta)]^{-1} = -\frac{(1-\theta)^2 \log(1-\theta)^2}{\log(1-\theta)+\theta}$$

j

Since we know that our function l(.) was obtained from a sufficient statistic which is a sum, we may apply the delta method to reach the asymptotic distribution of $\hat{\theta} \rightarrow N(0, V(\hat{\theta})) = N(0, -\frac{(1-\theta)^2 \log(1-\theta)^2}{\log(1-\theta)+\theta})$

k

```
LogData <- c( 710, 175, 74, 23, 10, 4, 2, 1, 1)
barY <- 0
numMeasures <- 0
for( i in 1:9){
 barY <- barY + LogData[i]*i
 numMeasures <- numMeasures + LogData[i]
}
barY <- barY / numMeasures

datapoints <- c()
for( i in 1:9 ){
    datapoints <- c( datapoints, rep( i, LogData[i] ) )
}

#Should make the epsilon a variable or something
hatTheta <- .5
fprime <- 1.0

while( abs( fprime ) > 1e-9) {
 fprime <- (barY / hatTheta) - (1 / ((hatTheta-1)*log( 1- hatTheta)))
 fdubprime <- ( -barY / hatTheta^2 )+ 1 / ((1-hatTheta)^2 * log( 1 - hatTheta )
^2 ) + 1 / ((1-hatTheta)^2 * log( 1- hatTheta ) )
 hatTheta <- hatTheta - fprime / fdubprime
}
barY

## [1] 1.479

hatTheta
```

```
## [1] 0.5217389

vTheta <- -1 / fdubprime
vTheta

## [1] 0.3007821
```

As we can see, the Maximum liklihood estimate for $\hat{\theta}$ is: 0.5217389

L

```
nullTheta <- .5
1000*(hatTheta - nullTheta)^2 / vTheta

## [1] 1.571174

1- pchisq( 1000*(hatTheta - nullTheta)^2 / vTheta, 1 )

## [1] 0.2100366
```

Since we obtain a p-value of only around 0.2100366, we fail to reject the null hypothesis that $\theta \neq .5$

M

Since it is known that the liklihood ratio test is invariant to nonlinear transformations, applying it to $\theta = \exp(-.7)$ is equivalent to testing: $\log(\theta) = -.7$.

```
#Since we're looking at a non-linear function we should use the liklihood ratio test instead of the
nullTheta <- exp( -.7 )
liklihoodThetaHat <- 0
liklihoodNULL <- 0

#Whats vectorization?
 for( i in 1:1000 ){
     liklihoodThetaHat <- liklihoodThetaHat + datapoints[i]*log( hatTheta) -
         log( datapoints[i]) - log( - log( 1 - hatTheta ) )
     liklihoodNULL <- liklihoodNULL + datapoints[i]*log( nullTheta) -
         log( datapoints[i]) - log( - log( 1 - nullTheta ) )
}

1 - pchisq(2*( liklihoodThetaHat - liklihoodNULL ),1)

## [1] 0.1462737
```

As we can see from the p-value coming from the liklihood ratio test, we fail to reject the null hypothesis at a reasonable confidence level.

N

Since the data is already in a relatively easily binned state, we may apply the chi-squared test to test if our estimates are consistent with the log-series law.

```
Probs <- numeric( 9 )
Probs[9] <- 1
for( i in 1:8) {
    Probs[i] <- -hatTheta^i / ( i*log( 1- hatTheta))
    Probs[9] <- Probs[9] - Probs[i]
}
Expected <- Probs*1000
Expected
```

```
## [1] 707.348120 184.525526  64.182767  25.114986  10.482773   4.557726
## [7]   2.038237   0.930499   0.819366
```

```
ChiSum <- 0
for( i in 1:9 ){
    ChiSum <- ChiSum + ((LogData[i]-Expected[i])^2 / Expected[i])
}
ChiSum
```

```
## [1] 2.317605
```

```
1- pchisq(ChiSum,8)
```

```
## [1] 0.9696944
```

Again, we fail to reject the null hypothesis that the data is inconsistent with the log-series distribution, and are left to conclude that it is possible that it was generated from the distribution.