

## MAST90105 Lab and Workshop 8 Solutions

## 1 Lab

In this lab, you will learn how to use R to find a minimum of function and to generate random variables from a given CDF which does not have its inverse in closed form.

1. Function `nlm()` in R can be used to find a minimum of a nonlinear function.

- a. Define a function  $f(x_1, x_2) = \exp(x_1^2 + x_2^2) - 2(\exp(x_1) + \exp(x_2))$  in R and use `nlm()` to find the minimum of this function.

```
func1 = function(x) {
  x1 = x[1]
  x2 = x[2]
  out = exp(x1^2 + x2^2) - 2 * (exp(x1) + exp(x2))
  out
}
mle1 = nlm(f = func1, p = c(1, 1))
mle1

## $minimum
## [1] -5.397726
##
## $estimate
## [1] 0.7235762 0.7235762
##
## $gradient
## [1] 1.646683e-06 1.646683e-06
##
## $code
## [1] 1
##
## $iterations
## [1] 7
```

*Z has CDF  $F_Z$*

*①  $V \stackrel{d}{=} \text{Uniform}$*

*②  $Z = F_Z^{-1}(V)$*

- b. As you have learned, to generate a random variable  $Z$  from a continuous distribution with given CDF  $F_Z(z)$ , one can follow these steps:

- i. Generate  $U \sim U(0, 1)$
- ii. Compute  $Z = F_Z^{-1}(U)$

The quantile function  $F_Z^{-1}(q)$  might not be available in closed form and can be computed numerically. For a strictly increasing CDF  $F_Z(z)$  there exists a unique solution to the equation:  $F_Z(z_q) = q$ . It implies the function  $G_Z(z) = (F_Z(z) - q)^2$  attains its minimum (which is zero) at  $z = z_q$ . The point  $z_q$  at which this minimum

is attained can be found using `nlm()` function in R, and one can follow these steps to generate a random variable  $Z$  in this case:

- i. Generate  $U \sim U(0, 1)$
- ii. Compute  $Z = \operatorname{argmin}_z (F_Z(z) - U)^2$

Use this approach to generate a sample of size  $N = 1000$  from a continuous random variable with the PDF:

$$f_Z(z) = \begin{cases} 0.5 + z^2 + z^5, & 0 < z < 1, \\ 0, & \text{otherwise.} \end{cases}$$

不好算  $F_Z^{-1}(U)$  时,

Draw a histogram of the generated data.

The CDF is

let  $z = \operatorname{argmin}_z ((F_Z(z) - U)^2)$

$$F_Z(z) = \int_0^z (0.5 + x^2 + x^5) dx = \frac{z}{2} + \frac{z^3}{3} + \frac{z^6}{6}.$$

The R code:

$\min (F_Z(z) - U)^2$

$F_Z(z) = U$

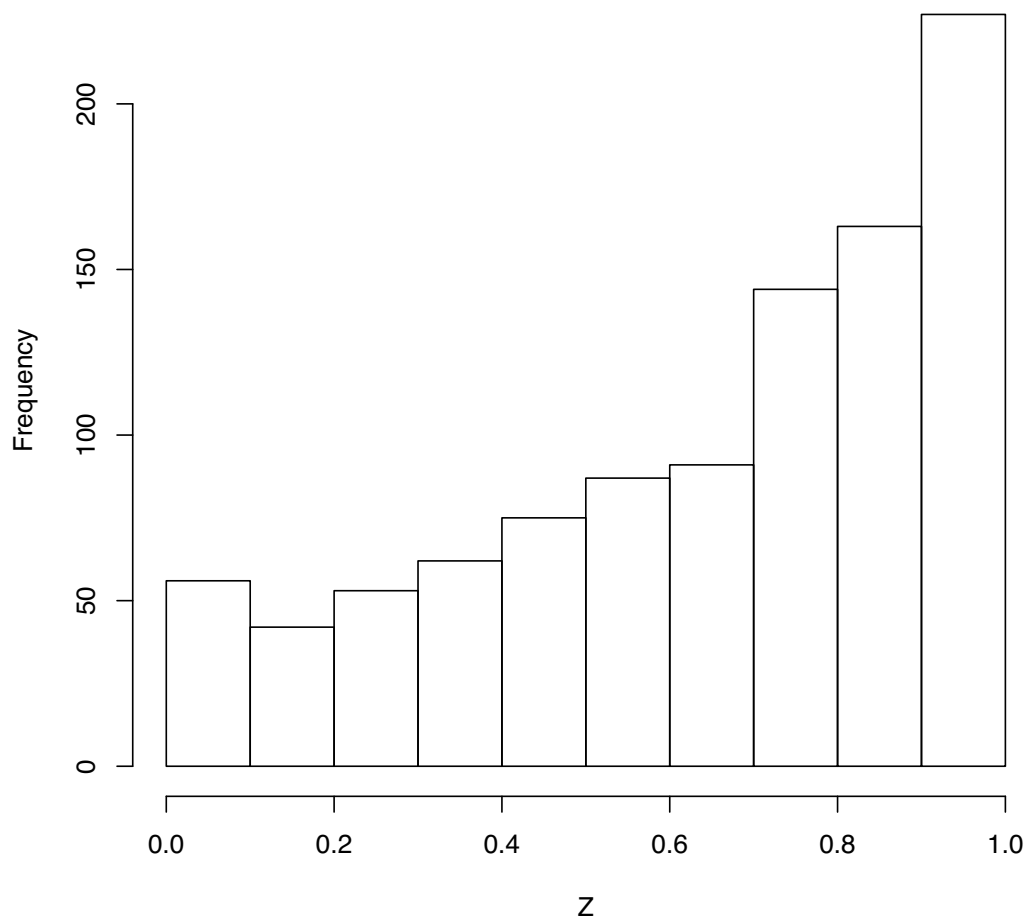
$Z = F_Z^{-1}(U)$

```
func2 = function(z, q) {
  if (z < 0 || z > 1)
    return(1e+100) #to force nlm() search a solution in (0,1) interval
  diff = (z/2 + z^3/3 + z^6/6 - q)^2
  diff
}

genf = function(N) {
  U = runif(N)
  Z = rep(0, N)
  for (i in 1:N) {
    Ui = U[i]
    mle2 = nlm(f = func2, p = 0.5, q = Ui)
    Z[i] = mle2$estimate
  }
  Z
}

set.seed(1234)
Z = genf(1000)
hist(Z, main = "histogram of the generated data")
```

histogram of the generated data



2. A similar approach can be used to generate a random pair  $(X, Y)$  from a continuous bivariate distribution with given joint CDF  $F_{X,Y}(x, y)$ . Define

$$F_{Y|X}(y|x) = \Pr(Y \leq y | X = x) = \frac{\frac{\partial}{\partial x} F_{X,Y}(x, y)}{f_X(x)},$$

where  $F_X(x)$  and  $f_X(x)$  are the marginal CDF and PDF of  $X$ , respectively. Assuming  $F_X(x)$  and  $F_{Y|X}(y|x)$  are strictly increasing and continuously differentiable functions of  $x$  and  $y$ , respectively, one can follow these steps to generate a random pair  $(X, Y)$ :

a. Generate independent  $U_X \sim U(0, 1)$  and  $U_Y \sim U(0, 1)$

b. Compute  $X = F_X^{-1}(U_X)$  *gibb's sampler*

c. Compute  $Y = F_{Y|X}^{-1}(U_Y | X)$

*how to simulate Bivariate: r.v. (= 随机变量)*

*$F_{X,Y}(x,y)$  Find  $F_X(x)$ ,  $F_{Y|X}(y|x)$*

*$U_1, U_2 \sim U(0,1)$  iid.*

*independent*

The inverse CDFs  $F_X^{-1}$  and  $F_{Y|X}^{-1}$  can be computed numerically using `nlm()` function if they are not available in closed form.

- Bonus:** Show that the joint CDF of  $(X, Y)$  generated using the algorithm above is  $F_{X,Y}$ .
- Use this algorithm to generate a sample of size  $N = 1000$  from a continuous bivariate random variable  $(X, Y)$  with the joint PDF:

$$f_{X,Y}(x, y) = \begin{cases} x + y, & 0 < x, y < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Draw a scatter plot of the generated data.

We find the joint CDF first:

$$F_{X,Y}(x, y) = \int_0^x \int_0^y f_{X,Y}(x^*, y^*) dx^* dy^* = 0.5(xy^2 + x^2y), \quad 0 < x, y < 1.$$

It implies that  $F_X(x) = F_{X,Y}(x, 1) = 0.5(x + x^2)$ ,  $f_X(x) = F'_X(x) = 0.5 + x$ ,  $0 < x < 1$ , and

$$F_{Y|X}(y|x) = \frac{\frac{\partial}{\partial y} F_{X,Y}(x, y)}{f_X(x)} = \frac{0.5y^2 + xy}{0.5 + x}.$$

The inverse CDFs can be computed in closed form in this case by solving quadratic equations. Note that the range of random variables  $X$  and  $Y$  is  $(0, 1)$ , and therefore a positive root should be selected:

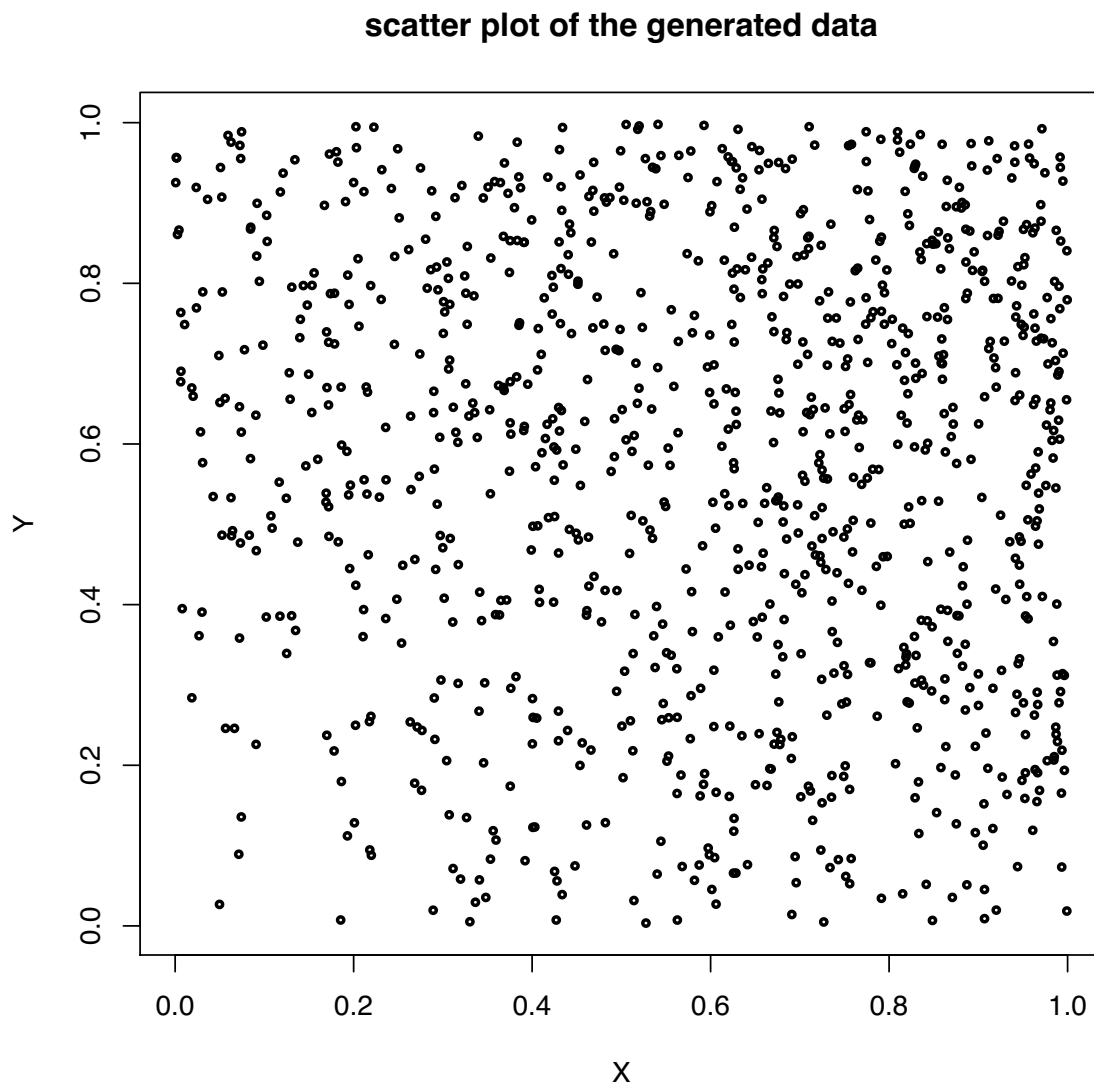
$$\begin{aligned} 0.5(x + x^2) &= q \Rightarrow F_X^{-1}(q) = -0.5 + \sqrt{0.25 + 2q}, \\ \frac{0.5y^2 + xy}{0.5 + x} &= q \Rightarrow F_{Y|X}^{-1}(q|x) = -x + \sqrt{x^2 + q(1 + 2x)}. \end{aligned}$$

R code:

```
genXY = function(N) {
  Ux = runif(N)
  Uy = runif(N)
  X = -0.5 + sqrt(0.25 + 2 * Ux)
  Y = -X + sqrt(X^2 + Uy * (1 + 2 * X))
  cbind(X, Y)
}

set.seed(1234)
Z = genXY(1000)

plot(Z, cex = 0.5, lwd = 2, main = "scatter plot of the generated data")
```



## 2 Workshop

1. a. A random sample  $X_1, \dots, X_n$  of size  $n$  is taken from a Poisson distribution with mean  $\lambda > 0$ .
  - i. Show the maximum likelihood estimator of  $\lambda$  is  $\hat{\lambda} = \bar{X}$ .
  - ii. Suppose with  $n = 40$  we observe 5 zeros, 7 ones, 12 twos, 9 threes, 5 fours, 1 five, and 1 six. What is the maximum likelihood estimate of  $\lambda$ .

•

$$L(\lambda) = \prod_{i=1}^n \frac{\exp(-\lambda) \lambda^{x_i}}{x_i!} = e^{-n\lambda} \lambda^{\sum_{i=1}^n x_i} \frac{1}{\prod_{i=1}^n x_i!}$$

so that

$$\log L(\lambda) = \ell(\lambda) = -n\lambda + \ln \lambda \sum_{i=1}^n x_i - \ln \prod_{i=1}^n x_i!$$

and

$$\ell'(\lambda) = -n + \frac{\sum_{i=1}^n x_i}{\lambda}$$

so that setting  $\ell'(\lambda) = 0$  yields  $\hat{\lambda} = \bar{X}$ .

- Since  $\ell''(\lambda) = -\frac{\sum_{i=1}^n x_i}{\lambda^2} < 0$ , so  $\bar{X}$  is the mle.
- $\bar{x} = (7 + 24 + 27 + 20 + 5 + 6)/40 = 2.225$ .

- b. Find the maximum likelihood estimator,  $\hat{\theta}$ , if  $X_1, \dots, X_n$  is a random sample from the following probability density function:

$$f(x; \theta) = (1/2) \exp(-|x - \theta|), -\infty < x < \infty, 0 < \theta < \infty$$

This involves minimizing  $\sum_{i=1}^n |x_i - \theta|$ , which is difficult. Try  $n = 5$  and a sample 6.1, -1.1, 3.2, 0.7, 1.7. Then deduce the MLE.

•

$$L(\theta) = \left(\frac{1}{2}\right)^n \prod_{i=1}^n \exp(-|x_i - \theta|)$$

$$\log L(\theta) = \ell(\theta) = -n \log 2 + \sum_{i=1}^n |x_i - \theta|$$

$$\ell'(\theta) = \sum_{i=1}^n \text{signum}(x_i - \theta)$$

where  $\text{signum}(x) = 1$  if  $x > 0$ ,  $\text{signum}(x) = 0$  if  $x = 0$ ,  $\text{signum}(x) < 0$  if  $x < 0$ . Note that  $\ell(\theta)$  is piecewise linear and nondifferentiable when  $\theta$  equals any  $x_i$ , so  $\ell'(\theta)$  is not defined at those points. If  $n$  is odd,  $\ell'(\theta) > 0$  for  $\theta < x_{[n/2]+1}$  and  $\ell'(\theta) < 0$  for  $\theta > x_{[n/2]+1}$ .  $\hat{\theta}$  must equal the middle value,  $x_{[n/2]+1}$ . If  $n$  is even, a similar argument shows that  $\hat{\theta}$  can be taken anywhere between the middle two ordered values, for example at the average of the middle two values. So  $\hat{\theta}$  can be taken to be the median in either case.

2. Let  $f(x; \theta) = \theta x^{\theta-1}$ ,  $0 < x < 1$ ,  $\theta \in \Omega = \{\theta : 0 < \theta < \infty\}$  and let  $X_1, \dots, X_n$  denote a random sample from this distribution. Note that

$$\int_0^1 x \theta x^{\theta-1} dx = \frac{\theta}{\theta + 1}$$

- a. Sketch the p.d.f. of  $X$  for  $\theta = 1/2$  and  $\theta = 2$ .

- Sketch above.

- b. Show that  $\hat{\theta} = -n / \ln(\prod_{i=1}^n X_i)$  is the maximum likelihood estimator of  $\theta$ .

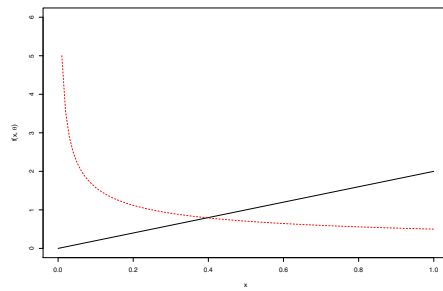


Figure 1: Sketch for Question 2(a)

- The likelihood is

$$L(\theta) = \theta^n \prod_{i=1}^n x_i^{\theta-1}.$$

Hence the log-likelihood is

$$\ell(\theta) = n \ln \theta + (\theta - 1) \sum_{i=1}^n \ln(x_i),$$

and the derivative is

$$\ell'(\theta) = \frac{n}{\theta} + \sum_{i=1}^n \ln(x_i),$$

so that setting  $\ell'(\theta) = 0$  yields

$$\hat{\theta} = \frac{-n}{\sum_{i=1}^n \ln(X_i)} = \frac{-n}{\ln(\prod_{i=1}^n X_i)}$$

- Differentiating again shows

$$\ell''(\theta) = -\frac{n}{\theta^2} < 0$$

so  $\hat{\theta}$  is the mle.

- For each of the following three sets of observations from this distribution compute the maximum likelihood estimates and the methods of moments estimates.

$X$	$Y$	$Z$
0.0256	0.9960	0.4698
0.3051	0.3125	0.3675
0.0278	0.4374	0.5991
0.8971	0.7464	0.9513
0.0739	0.8278	0.6049
0.3191	0.9518	0.9917
0.7379	0.9924	0.1551
0.3671	0.7112	0.0710
0.9763	0.2228	0.2110
0.0102	0.8609	0.2154

( $\sum_{i=1}^n \ln(x_i) = -18.2063$ ,  $\sum_{i=1}^n \ln(y_i) = -4.5246$ ,  $\sum_{i=1}^n \ln(z_i) = -10.42968$ ,  
 $\sum_{i=1}^n x_i = 3.7401$ ,  $\sum_{i=1}^n y_i = 7.0592$ ,  $\sum_{i=1}^n z_i = 4.6368$ .)

- Substituting gives  $\hat{\theta}_X = 0.549$ ,  $\hat{\theta}_Y = 2.210$ ,  $\hat{\theta}_Z = 0.959$ . To find the method of moments estimators solve  $\bar{x} = \theta/(\theta + 1)$  which yields  $\tilde{\theta} = \bar{x}/(1 - \bar{x})$  and hence  $\tilde{\theta}_X = 0.598$ ,  $\tilde{\theta}_Y = 2.400$  and  $\tilde{\theta}_Z = 0.865$ .

3. Let  $X_1, \dots, X_n$  be a random sample from the exponential distribution whose p.d.f. is  $f(x; \theta) = (1/\theta) \exp(-x/\theta)$ ,  $0 < x < \infty$ ,  $0 < \theta < \infty$ .

- a. Show that  $\bar{X}$  is an unbiased estimator of  $\theta$ .

- Recall that for a random sample  $X_1, \dots, X_n$ ,

$$E(\bar{X}) = \frac{1}{n} \sum_{i=1}^n E(X_i) = E(X_1)$$

So  $\bar{X}$  is unbiased for  $E(X_1) = \theta$ .

- b. Show that the variance of  $\bar{X}$  is  $\theta^2/n$ . What is a good estimate of  $\theta$  if a random sample of size 5 yielded the values 3.5, 8.1, 0.9, 4.4 and 0.5?

- Recall that

$$\text{Var}(\bar{X}) = \frac{1}{n^2} \sum_{i=1}^n \text{Var}(X_i) = \frac{\text{Var}(X_1)}{n}$$



and  $\sigma^2 = \text{Var}(X_i) = \theta^2$ . Hence  $\text{Var}(\bar{X}) = \sigma^2/n = \theta^2/n$  as required. A good estimator of  $\theta$  is  $\hat{\theta} = \bar{x} = 3.48$ . This is a good estimator as it is unbiased and its variance tends to zero as  $n \rightarrow \infty$ , and we showed it is the maximum likelihood estimator in class.

4. Let  $X_1, \dots, X_n$  be a random sample from a distribution having finite variance  $\sigma^2$ . Show that

$$S^2 = \sum_{i=1}^n \frac{(X_i - \bar{X})^2}{n-1}$$

is an unbiased estimator of  $\sigma^2$ . HINT: Write

$$S^2 = \frac{1}{n-1} \left( \sum_{i=1}^n X_i^2 - n\bar{X}^2 \right)$$

and compute  $E(S^2)$ .

•

$$\begin{aligned} S^2 &= \sum_{i=1}^n \frac{(X_i - \bar{X})^2}{n-1} \\ &= \frac{1}{n-1} \sum_{i=1}^n (X_i^2 - 2X_i\bar{X} + \bar{X}^2) \\ &= \frac{1}{n-1} \left( \sum_{i=1}^n X_i^2 - 2n\bar{X}^2 + n\bar{X}^2 \right) \\ &= \frac{1}{n-1} \left( \sum_{i=1}^n X_i^2 - n\bar{X}^2 \right) \end{aligned}$$

as required. Now in general,  $E(X^2) = \sigma^2 + \mu^2$ , so that  $E(X_i^2) = \sigma^2 + \mu^2$  and  $E(\bar{X}^2) = \sigma^2/n + \mu^2$  and hence

$$\begin{aligned} E(S^2) &= \frac{1}{n-1} \left\{ n(\sigma^2 + \mu^2) - n \left( \frac{\sigma^2}{n} + \mu^2 \right) \right\} \\ &= \frac{(n-1)\sigma^2}{n-1} = \sigma^2 \end{aligned}$$

5. Let  $X_1, \dots, X_n$  be a random sample of size  $n$  from the distribution with p.d.f.  $f(x; \theta) = (1/\theta)x^{(1-\theta)/\theta}$ ,  $0 < x < 1$ .

a. Show the mean of  $X$  is  $E(X) = 1/(1 + \theta)$ .

•

$$E(X) = \int_0^1 x \times \frac{1}{\theta} x^{(1-\theta)/\theta} dx = \left[ \frac{\theta}{\theta(\theta+1)} x^{(\theta+1)/\theta} \right]_0^1 = \frac{1}{\theta+1}$$

b. Show the maximum likelihood estimator of  $\theta$  is

$$\hat{\theta} = -\frac{1}{n} \sum_{i=1}^n \ln X_i$$

- The log likelihood is

$$\ell(\theta) = \ln L(\theta) = -n \log(\theta) + \sum_{i=1}^n \frac{1-\theta}{\theta} \ln(x_i),$$

so

$$\ell'(\theta) = \ln L(\theta) = \frac{-n}{\theta} - \frac{1}{\theta^2} \sum_{i=1}^n \ln(x_i).$$

Equating this to zero gives the required expression for  $\hat{\theta}$ . Writing  $t = \sum_{i=1}^n \ln(x_i)$  and differentiating gives

$$\ell''(\hat{\theta}) = \frac{n}{\hat{\theta}^2} + \frac{2t}{\hat{\theta}^3} = -\frac{n^3}{t^2} < 0,$$

so  $\hat{\theta}$  is the MLE.

c. Is the MLE unbiased? (You can do this using integration by parts combined with rules about expectation of the sample mean.)

- Yes, the MLE is unbiased because

$$E(\hat{\theta}) = E(-\ln(X_1)),$$

being the sample average of the logarithms of the sample values and

$$E(-\ln(X_1)) = \int_0^1 -\ln(x) f(x; \theta) dx = [-\ln(x) x^{1/\theta}]_0^1 + \int_0^1 \frac{1}{x} x^{1/\theta} dx = \theta.$$

d. Show the method of moments estimator of  $\theta$  is

$$\tilde{\theta} = \frac{1 - \bar{X}}{\bar{X}}.$$

- Equating  $\bar{X}$  to  $\frac{1}{\theta+1}$  is the same as  $\theta = \frac{1}{\bar{X}} - 1 = \frac{1-\bar{X}}{\bar{X}}$ , as required.