MAST90105: Lab and Workshop 8

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.

   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 $Z = \text{argmin}_z (F_Z(z) - U)^2$

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

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

      Draw a 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 $X = F_X^{-1}(U_X)$

$L(\theta) = \prod_{i=1}^{n} \frac{\pi^{x_i} e^{-\pi}}{x_i!} = e^{-n\pi} \prod_{i=1}^{n} \frac{\pi^{x_i}}{x_i!}$ $\quad \ell(L(\theta)) = -n\pi + \sum \ln(\pi^{x_i}/x_i!)$

$= -n\pi + \sum(\ln \pi^{x_i} - \ln x_i!)$

$= -n\lambda + \Sigma (x_i \ln \lambda - \ln x_i!)$

$$\ell'(\theta) = -n + \frac{\Sigma x_i}{\lambda} = 0$$

$$\lambda = \bar{x}$$

$$f(x;\theta) = \frac{1}{2} \exp(-|x-\theta|) = \frac{1}{2} e^{-|x-\theta|}$$

$$L(\theta) = \prod \frac{1}{2} e^{-|x_i-\theta|} = \left(\frac{1}{2}\right)^n e^{-\Sigma|x_i-\theta|} = \left(\frac{1}{2}\right)^n \frac{1}{e^{\Sigma|x_i-\theta|}}$$

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

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.

$$\ell(\theta) = n\ln\frac{1}{2} + (-\Sigma|x_i-\theta|) = -n\ln 2 - \Sigma|x_i-\theta|$$

$$|x| \begin{cases} x, & x>0 \\ -x, & x<0 \end{cases}$$

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

$$|x_i-\theta| \begin{cases} |x_i-\theta) & x_i>\theta>0 \\ -(x_i-\theta) & x_i-\theta<0 \\ 0 & x_i-\theta=0 \end{cases} \Rightarrow \begin{cases} -1 & x_i-\theta>0 \\ 1 & x_i-\theta<0 \\ 0 & x_i=\theta \end{cases} = -\text{signum}(x_i-\theta)$$

$$\ell'(\theta) = \sum_{i=1}^{n} \text{signum}(x_i-\theta) = 0$$

$-1, -1, 0, \cdots +1, +, +1,$

正负数量相同

$$\text{signum} \begin{cases} 1 & x>0 \\ -1 & x<0 \\ 0 & x=0 \end{cases}$$

$x_1 \quad x_1-\theta \quad -1$

$x_2 \quad x_2-\theta \quad -1$

$\vdots \qquad\qquad \vdots$

$x_n \qquad\qquad +1$

θ = median of the data

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

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

a. **Bonus:** Show that the joint CDF of (X, Y) generated using the algorithm above is $F_{X,Y}$.

b. 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.

2 Workshop

$\prod_{i=1}^{n} \left[\frac{e^{-\lambda}\lambda^{x}}{x!} \right] =$

1. a. A random sample $X_1, \dots, X_n$ of size n is taken from a Poisson distribution with mean $\lambda > 0$.

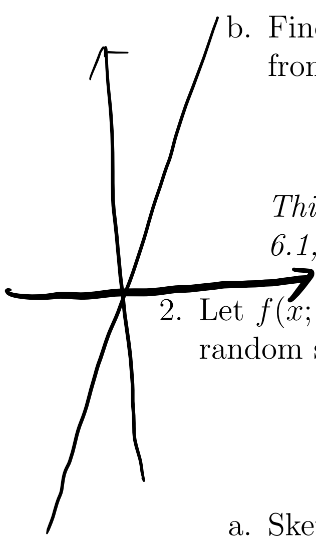
   i. Show the maximum likelihood estimator of λ is $\hat{\lambda} = \bar{X}$.

   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 λ.

   b. Find the maximum likelihood estimator, $\hat{\theta}$, if $X_1, \cdots, X_n$ 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.



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

$\frac{1}{2}x^{\frac{1}{2}}$ $2x$

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

$\prod_{i=1}^{n} \theta x_i^{\theta-1} = n\theta^{n} \prod x_i^{\theta-1}$

$n \ln n\theta \prod x_i^{\theta-1}$

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

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

   c. 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

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

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

   $\prod_{i=1}^{n} \frac{1}{\theta} \exp(-x_i/\theta)$

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

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

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

   is an unbiased estimator of σ^2. HINT: Write

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

   and compute $E(S^2)$. $E(S^2) =$

5. Let $X_1, \ldots, X_n$ 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 $E(X) = 1/(1+\theta)$.

b. Show the maximum likelihood estimator of θ is

$$\hat{\theta} = -\frac{1}{n}\sum_{i=1}^{n} \ln X_i = E\left(-\frac{1}{n}\sum \ln X_i\right)$$

$$= -\frac{1}{n}E\left(\sum \ln X_i\right)$$

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

d. Show the method of moments estimator of θ is

$$= -E(\ln(X_i))$$

$$\widetilde{\theta} = \frac{1-\bar{X}}{\bar{X}}. \qquad = -\int_0^1 \ln x_i \cdot f(x)$$

1. a. A random sample $X_1, \ldots, X_n$ of size n is taken from a Poisson distribution with mean $\lambda > 0$.

   i. Show the maximum likelihood estimator of λ is $\hat{\lambda} = \bar{X}$. $\prod_{i=1}^{n} \frac{\lambda^x e^{-\lambda}}{x!}$

   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 λ.

   b. Find the maximum likelihood estimator, $\hat{\theta}$, if $X_1, \cdots, X_n$ 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.

$X_i \overset{iid}{=} \text{Possion}(\lambda)$ where $i = 1, 2, \cdots, n$

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

求参数估计

$$\ln L(\theta) = \ln\left(\frac{e^{-n\lambda}\lambda^{\sum x_i}}{\prod_{i=1}^{n} x_i!}\right)$$

求导

$$= \ln e^{-n\lambda} + \ln \lambda^{\sum x_i} - \ln \prod_{i=1}^{n} x_i!$$

$$= -n\lambda + \sum x_i \ln\lambda - \ln \prod_{i=1}^{n} x_i!$$

求λ导

$$-n + \frac{\sum x_i}{\lambda} - 0 = 0$$

ANS:

$$\lambda = \bar{x} = \frac{\sum x_i}{n}$$

$\boxed{\text{The MLE is } \hat{\lambda} = \bar{x}}$

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

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

$l(\theta) = \frac{1}{2}^n e^{-\sum_{i=1}^{n}|x_i-\theta|}$

$\ln l(\theta) = \ln \frac{1}{2}^n + \ln e^{-\sum_{i=1}^{n}|x_i-\theta|} = -n\ln 2 - \sum_{i=1}^{n}|x_i-\theta|$

$X_i \sim U(0,\theta)$ $\quad f(x_i;\theta) = \frac{1}{\theta}$ joint PMF $\rightarrow f(x_i,\theta) = \frac{1}{\theta} \times 1 \quad (0 < x_i \leq \theta)$

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

$\cancel{\ln' l(\theta) = -\frac{n}{\theta} = 0\,?}$

$l(\theta) = \prod^{n} \frac{1}{\theta} 1 (0 < x_i \leq \theta)$

$= \frac{1}{\theta^n} \prod^{n} 1 (0 < x_i \leq \theta)$

$\hat{\theta} = max(X_1, \ldots, X_n)$

5. Let $X_1, \ldots, X_n$ 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 $E(X) = 1/(1+\theta)$.

   b. Show the maximum likelihood estimator of θ is

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

$\frac{1+\theta}{\theta}$

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

   d. Show the method of moments estimator of θ is

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

$E(X) = \int_0^1 x f(x)dx = \int_0^1 \frac{x}{\theta} x^{\frac{1-\theta}{\theta}} dx = \frac{1}{\theta}\int_0^1 x^{\frac{1}{\theta}-1+1} dx = \frac{1}{\theta}\int_0^1 x^{\frac{1}{\theta}} = \frac{1}{\theta}\frac{\theta}{1+\theta} x^{\frac{1}{\theta}+1} dx$

$= \frac{1}{1+\theta}$

$\prod_{i=1}^{n} \frac{1}{\theta} x_i^{\frac{1-\theta}{\theta}} = \theta^{-n} \prod_{i=1}^{n} x_i^{(\frac{1-\theta}{\theta})} = \theta^{-n} (\prod_{i=1}^{n} x_i)^{\frac{1-\theta}{\theta}}$ $\quad -n\ln\theta + \frac{1-\theta}{\theta}\ln(\prod_{i=1}^{n} x_i) = -n\ln\theta + \frac{1-\theta}{\theta}\sum \ln x_i$

$\ln'(\theta) = -\frac{n}{\theta} - \frac{1}{\theta^2}\sum \ln x_i \Rightarrow -\frac{1}{\theta^2}\sum_{i=1}^{n} \log(x_i) = \frac{n}{\theta}$

$-\frac{1}{\theta}\sum_{i=1}^{n} \log(x_i) \quad n$

$\theta = \frac{-\sum_{i=1}^{n} \log(x_i)}{n}$ ✓ $\hat{\theta} =$

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

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

is an unbiased estimator of σ^2. HINT: Write

?

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

and compute $E(S^2)$.

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

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

$$= \frac{1}{n-1} E\left[\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right]$$

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

$$= \frac{1}{n-1}\left(\sum_{i=1}^{n} E(X_i^2) - nE(\bar{X}^2)\right) \to \mathrm{Var}(\bar{X}) + E(\bar{X})^2$$

$$\frac{\sigma^2}{n} \quad \mu^2$$

$$\sum_{i=1}^{n} \left(\mathrm{Var}(X_i) + E(X_i)^2\right) -$$

$$\sigma^2 \quad \mu^2$$

$$= \frac{1}{n-1}\left\{n(\sigma^2 + \mu^2) - n\left(\frac{\sigma^2}{n} + \mu^2\right)\right\}$$

$$= \sigma^2$$

验证

$l'(\theta) = -\frac{n}{\theta} - \frac{1}{\theta^2}\sum_{i=1}^{n}\log(x_i)$ （接第5题） 验证部分写了步不了.

$l''(\theta) = \frac{n}{\theta^2} + \frac{2}{\theta^3}\sum_{i=1}^{n}\log(x_i)$

把 $\hat{\theta} = \frac{-\sum\log(x_i)}{n}$ 代入看 $l''(\hat{\theta})$ 是否<0

$\log x_i < 0$ when $0 < x_i < 1$

$l''(\hat{\theta})$ is negative $\therefore$ $\hat{\theta}$ is MLE

$e^{-(x_i-\eta)}$ $x_i \geq 0$ get MLE for η.

$= e^{-(x_i-\eta)}\mathbb{1}(x_i \geq \eta)$

$\hat{\eta} = \min(x_1, \dots, x_n)$