Name	SOLUTION
PID #	
Instructor	Sergey Kirshner

STAT 598G Spring 2011
Quiz #2
March 22, 2011

You are not allowed to use books or notes. Please read the directions carefully. The quiz is graded out of 3 points. You have 15 minutes to complete it. Please show all your work. Use the back of the page if you need more space.

A log-normal density is given by

$$f\left(x|\mu,\sigma^2\right)=\frac{1}{\sqrt{2\pi\sigma^2}x}e^{-\frac{1}{2\sigma^2}(\ln x-\mu)^2},\ x>0.$$

Suppose a data set $\mathcal{D}=\{x_1,\ldots,x_n\}$ is obtained by drawing n i.i.d. samples from f with the same unknown $\boldsymbol{\theta}=\left(\mu,\sigma^2\right)$, $\mu,\sigma>0$.

1. Write down the log-likelihood $l(\boldsymbol{\theta})$.

   Solution:

$$\begin{aligned} l(\boldsymbol{\theta}) &= \ln P(\mathcal{D}|\boldsymbol{\theta}) = \sum_{i=1}^{n}\ln f\left(x_i|\mu,\sigma^2\right)=\sum_{i=1}^{n}\left[-\frac{1}{2}\ln(2\pi)-\frac{n}{2}\ln\sigma^2-\ln x_i-\frac{1}{2\sigma^2}(\ln x_i-\mu)^2\right] \\ &= -\frac{n}{2}\ln(2\pi)-\frac{n}{2}\ln\sigma^2-\sum_{i=1}^{n}\ln x_i-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(\ln x_i-\mu)^2. \qquad (1)\end{aligned}$$

2. Derive MLE $\left(\hat{\mu},\hat{\sigma^2}\right)$.

   Solution: We'll find the critical points by finding sets of parameters making the gradient vanish:

$$\left(\hat{\mu},\hat{\sigma^2}\right)=\operatorname*{argmax}_{\mu,\sigma^2} l\left(\mu,\sigma^2\right)=\operatorname*{argmin}_{\mu,\sigma^2}\frac{n}{2}\ln\sigma^2+\frac{1}{2\sigma^2}\sum_{i=1}^{n}(\ln x_i-\mu)^2;$$

$$\frac{\partial l\left(\mu,\sigma^2\right)}{\partial\mu}=\frac{1}{\sigma^2}\sum_{i=1}^{n}(\ln x_i-\mu)=0, \qquad (2)$$

$$\frac{\partial l\left(\mu,\sigma^2\right)}{\partial\sigma^2}=-\frac{n}{2\sigma^2}+\frac{1}{2\sigma^4}\sum_{i=1}^{n}(\ln x_i-\mu)^2=0. \qquad (3)$$

   Solving (2) yields $\hat{\mu}=\frac{1}{n}\sum_{i=1}^{n}\ln x_i$. Plugging $\hat{\mu}$ into (3) yields $\hat{\sigma^2}=\frac{1}{n}(\ln x_i-\hat{\mu})^2$.

3. Prove that your estimate indeed maximizes the likelihood/log-likelihood.

Solution: There are several ways to do this part. One is to notice that log-normal distribution is a member of the exponential family:

$$f(x|\mu,\sigma) = \frac{1}{x}\exp\left(-\frac{1}{2\sigma^2}\ln^2 x + \frac{\mu}{\sigma^2}\ln x - \frac{\mu^2}{2\sigma^2} - \frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln\sigma^2\right). \tag{4}$$

Set $\theta_1 = \frac{1}{\sigma^2}$, $\theta_2 = \frac{\mu}{\sigma^2}$, $\phi_1(x) = -\frac{1}{2}\ln^2 x$, $\phi_2(x) = \ln x$, $g(\theta_1,\theta_2) = \frac{\mu^2}{2\sigma^2} + \frac{1}{2}\ln(2\pi) + \frac{1}{2}\ln\sigma^2 = \frac{1}{2}\left(\frac{\theta_2^2}{\theta_1} - \ln\theta_1 + \ln(2\pi)\right)$. Thus (4) becomes

$$f(x|\mu,\sigma) = f(x|\theta_1,\theta_2) = \frac{1}{x}\exp\left(\theta_1\phi_1(x) + \theta_2\phi_2(x) - g(\theta_1,\theta_2)\right)$$

which is strictly log-concave in θ_1,θ_2 and thus having a unique global maximum. $g(\theta_1,\theta_2)$ has the same form as for normal distribution, with the same maximum likelihood solution $\left(\hat{\theta}_1,\hat{\theta}_2\right)$ satisfying:

$$\begin{aligned}
0 &= \frac{\partial}{\partial\theta_1}\sum_{i=1}^{n}\left(\theta_1\phi_1(x_i) + \theta_2\phi_2(x_i) - g(\theta_1,\theta_2)\right) = \sum_{i=1}^{n}\phi_1(x_i) + n\frac{\theta_2^2}{2\theta_1^2} + n\frac{1}{2\theta_1}, \\
0 &= \frac{\partial}{\partial\theta_2}\sum_{i=1}^{n}\left(\theta_1\phi_1(x_i) + \theta_2\phi_2(x_i) - g(\theta_1,\theta_2)\right) = \sum_{i=1}^{n}\phi_2(x_i) - n\frac{\theta_2}{\theta_1}.
\end{aligned}$$

From the second equation, $\frac{\hat{\theta}_2}{\hat{\theta}_1} = \frac{1}{n}\sum_{i=1}^{n}\phi_2(x) = \hat{\mu}$. From the first equation, $\frac{1}{\hat{\theta}_1} = \frac{2}{n}\sum_{i=1}^{n}\phi_1(x_i) + \frac{1}{n}\left(\sum_{i=1}^{n}\phi_1(x_i)\right)^2 = \hat{\sigma^2}$. So $\left(\hat{\mu},\hat{\sigma_2^2}\right)$ indeed correspond to the maximum value of $l(\boldsymbol{\theta})$.