Professor Hardin
Math 152 – Fall 2007
warm-up # 22

Name: ________________________

Consider the situation given in the handout, "Patterns in DNA". Now, let's say we would like to know a little bit more about the actual rate (# of palindromes / segment). Assume the count in each segment is distributed according to a Poisson distribution.

Using Fisher's information and what you know about MLEs, give a 95% confidence interval for the rate of palindromes / 4000bp.

Solution:
We know $\hat{\lambda} = \overline{X}$ is the MLE (you should be able to show that.) Also, we know that $\text{Var}(\overline{X}) = \lambda/n$. (Notice this makes it efficient, see Theorem 6.2.1 and below...) For any MLE:

$$\sqrt{nI(\lambda)}(\hat{\lambda} - \lambda) \xrightarrow{D} N(0,1)$$

where

$$\begin{aligned} I(\lambda) &= E\Big[\frac{-\partial^2 \ln f(x_i;\lambda)}{\partial \lambda^2}\Big] \\ &= E\Big[\frac{-\partial}{\partial \lambda}\Big(\frac{X_i}{\lambda} - 1\Big)\Big] = E\Big[-\frac{-X_i}{\lambda^2}\Big] = \frac{1}{\lambda} \end{aligned}$$

An approximate asymptotic 95% CI for λ is:

$$\hat{\lambda} \pm 1.96\sqrt{\frac{1}{nI(\lambda)}} \rightarrow 5.16 \pm 1.96\sqrt{\frac{5.16}{57}}$$

(4.57, 5.75) is a 95% confidence interval for λ. That is, if, in fact, we believe the Poisson model (why is it we have to believe the Poisson model??), we are 95% confident that the true rate of palindromes is between 4.57 and 5.75 per 4000 bp.