243 Homework 6

Ariel Sim (50% of work) & Samantha Maillie (50% of work)

6/4/2018

Problem 1

This problem was very straightfoward particularly given the textbook's walk through with the code. We see an estimated correlation coefficent of .5459189. We later find that the bootstrap produces a smaller standard error than the jackknife. The t-interval approach was a little wider than the others and as a result it crossed over zero and contained negative values. This doesn't give us much insight into the time of correlation betweent the two variables.

We ran more than just the normal theory confidence intervals. We also ran basic and percentile. Using these three we get confidence intervals that are entirely positive. This makes intuitive sense given we are examining a relationship bewtween LSAT scores and gpa. However, they do drop down to as low as .1 for a lower bound and rise to around .9 for an upper bound. This indicates a positive correlation but gives little insight into the strength of that relationship. See Hw6 01 code for plots and additional information.

Problem 2

a.

The distribution of $\hat{\theta}_{MLE}$ is the distribution of the maximum order statistic. We know the formula for this is $nf(x)F(x)^{n-1}$ plugging in the appropriate values we get that $\hat{\theta}_{MLE}$ follows the distribution $\frac{n}{\theta^n}x_n^{n-1}$.

b.

$Var(X_n) = E[X_n^2] - E[X_n]^2$

$E[X_n^2] = \int_0^3 \frac{n}{3^n}x_n^{n-1}x_n^2 dx_n$
$E[X_n^2] = \frac{n}{3^n}\int_0^3 x_n^{n+1} dx_n$
$E[X_n^2] = \frac{n}{3^n}\frac{1}{n+2}3^{n+2}$
$E[X_n^2] = \frac{9n}{n+2}$

$E[X_n] = \int_0^3 \frac{n}{3^n}x_n^{n-1}x_n dx_n$
$E[X_n] = \frac{n}{3^n}\int_0^3 x_n^n dx_n$
$E[X_n] = \frac{n}{3^n}\frac{1}{n+1}3^{n+1}$
$E[X_n] = \frac{3n}{n+1}$

$Var(X_n) = \frac{9n}{n+2} - (\frac{3n}{n+1})^2$
If we let $n = 50$ we find the variance to be approximately equal to .003327.

discussion for c - f

We get a consistantly excellent estimation of this using parametric bootstraping methods. The last run we did gave an estimate of 0.003017126 although we have seen it vary a little bit. The histogram of estimated $\hat{\theta}^*$'s matches very nicely with the true MLE function plot.

We see that the nonparametric estimate fails. Since we sample and recalculate the theta value based on the generated values and then regenerate simulation data we think this is causing the issue. The MLE for theta from the uniform distribution with theta as a parameter is always a biased estimator that underestimates theta. We see quickly that as we recalculate theta it goes to zero and hence the method doesn't work. This distribution tends to have difficulties in many applications that work well for most of the other common distributions.

Plots are contained in the hw6 02 r code file.

Problem 3

a

Because the true function (test function 1) in HW#2 is a piecewise constant, using a genetic algorithm on this test function would provide more reasonable regression curve estimates than applying it on the 2nd test function provided in this problem. This is because the 2nd function is smooth, so approximating using a piecewise function will not produce good results. Plots of both functions can be found attached below.

NOTE TO SELF, add plots once the algorithm is done.

b

The confidence bands for both functions are found using bootstrapping residuals and pairs. The resulting plots are shown below.

NOTE TO SELF, add plots once the algorithm is done!!!

From the plots, we noticed that bootstrapping residuals' confidence bands are closer to the estimated piecewise function. This is not surprising since this method only resamples the estimated residuals and not the jump points.

As such, we see that the confidence bands obtained from the bootstrapping pairs are more volatile and in some sense wider than the bootstrap residuals, especially around the jump points. Perhaps this method is not as stable since there is a possibility of resampling the same pair.

As for the test function from HW#2, the true function is a smooth function whereas the approximation is piecewise, so the confidence bands themselves will not contain the true function either, thus a poorer approxmation.

c

The steps are as follows: 1. Create 2000 bootstrap samples 2. Find the best chromosome using the genetic algorithm for each of the bootstrapped samples 3. Count the n_i's that are selected as jump points for each i 4. For confidence level $1-\alpha$, take the smallest set of points where the proportion of the sum of their jump points counts are at least $1-\alpha$.

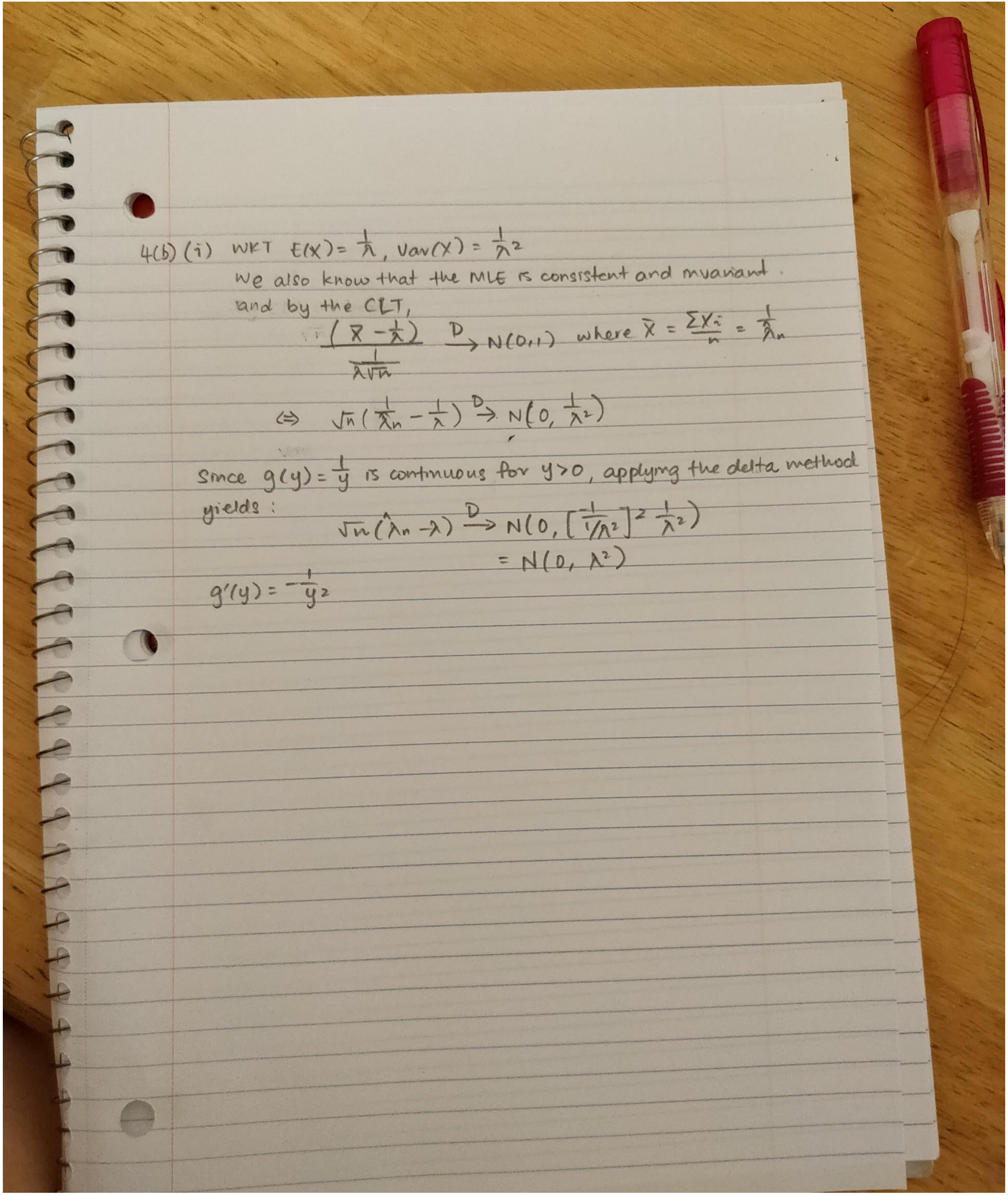
Problem 4

a.

$likelihood = \prod \lambda e^{-\lambda x_i}$
$likelihood = \lambda^n e^{-\lambda \sum x_i}$
$L = log - likelihood = nlog(\lambda) - \lambda \sum x_i$
$L' = \frac{n}{\lambda} - \sum x_i = 0$
$\hat{\lambda}_{MLE} = \frac{n}{\sum x_i} = \frac{1}{\bar{X}}$

b for part 1 see the picture below.

$\sqrt{n}(\hat{\lambda}_n - \lambda) \rightarrow^d N(0, \lambda^2)$



4(b) (i) WKT $E(X) = \frac{1}{\lambda}$, $Var(X) = \frac{1}{\lambda^2}$

We also know that the MLE is consistent and invariant.

and by the CLT,

$$\frac{(\bar{X} - \frac{1}{\lambda})}{\frac{1}{\lambda\sqrt{n}}} \xrightarrow{D} N(0,1) \quad \text{where } \bar{X} = \frac{\sum X_i}{n} = \frac{1}{\hat{\lambda}_n}$$

$$\Leftrightarrow \quad \sqrt{n}\left(\frac{1}{\hat{\lambda}_n} - \frac{1}{\lambda}\right) \xrightarrow{D} N\left(0, \frac{1}{\lambda^2}\right)$$

Since $g(y) = \frac{1}{y}$ is continuous for $y > 0$, applying the delta method yields :

$$\sqrt{n}(\hat{\lambda}_n - \lambda) \xrightarrow{D} N\left(0, \left[\frac{-1}{1/\lambda^2}\right]^2 \frac{1}{\lambda^2}\right)$$
$$= N(0, \lambda^2)$$

$$g'(y) = -\frac{1}{y^2}$$

$\sqrt{n}(log(\hat{\lambda}_n) - log(\lambda)) \rightarrow^d N(0, 1)$

From the above we know that applying the log function will result in the mean remaining as zero. We also know from the delta method that our new variance will be $(g'(\lambda))^2 Var(\lambda)$ where $g(\lambda) = log(\lambda)$ and therefore $g'(\lambda) = \frac{1}{\lambda}$ Putting it all together we get the $g'(\lambda)^2$ and $Var(\lambda)$ cancelling each other out so that we have a variance of 1.

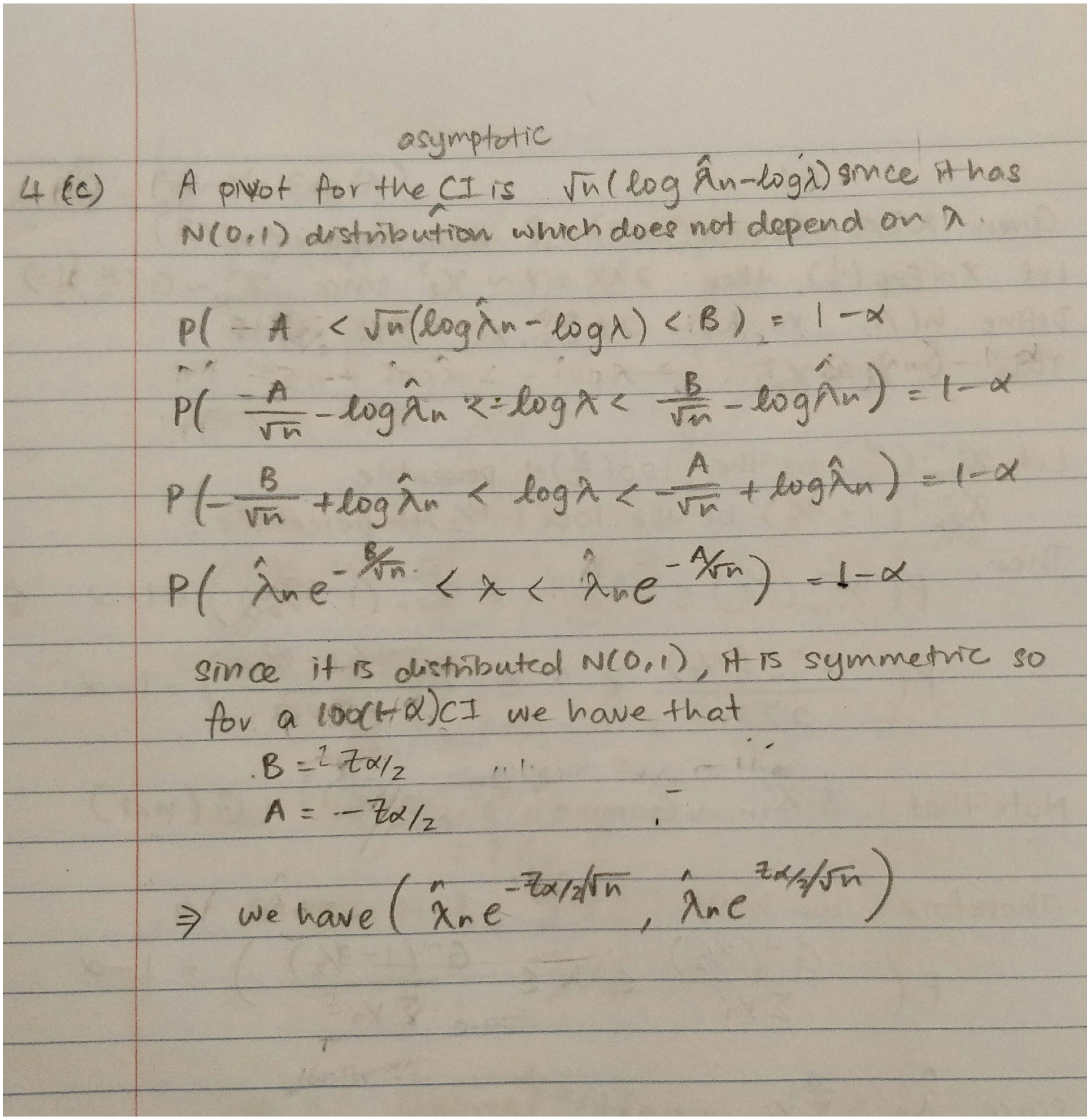
c See the picture below.

d See the picture below.

e

References

1.Rizzo, Maria L. Statistical Computing with R. Chapman & Hall/CRC, 2017.



4 (c) A pivot for the asymptotic CI is $\sqrt{n}(\log \hat{\lambda}_n - \log \lambda)$ since it has $N(0,1)$ distribution which does not depend on λ.

$$P(-A < \sqrt{n}(\log \hat{\lambda}_n - \log \lambda) < B) = 1-\alpha$$

$$P\left(\frac{-A}{\sqrt{n}} - \log \hat{\lambda}_n < \log \lambda < \frac{B}{\sqrt{n}} - \log \hat{\lambda}_n\right) = 1-\alpha$$

$$P\left(-\frac{B}{\sqrt{n}} + \log \hat{\lambda}_n < \log \lambda < -\frac{A}{\sqrt{n}} + \log \hat{\lambda}_n\right) = 1-\alpha$$

$$P(\hat{\lambda}_n e^{-B/\sqrt{n}} < \lambda < \hat{\lambda}_n e^{-A/\sqrt{n}}) = 1-\alpha$$

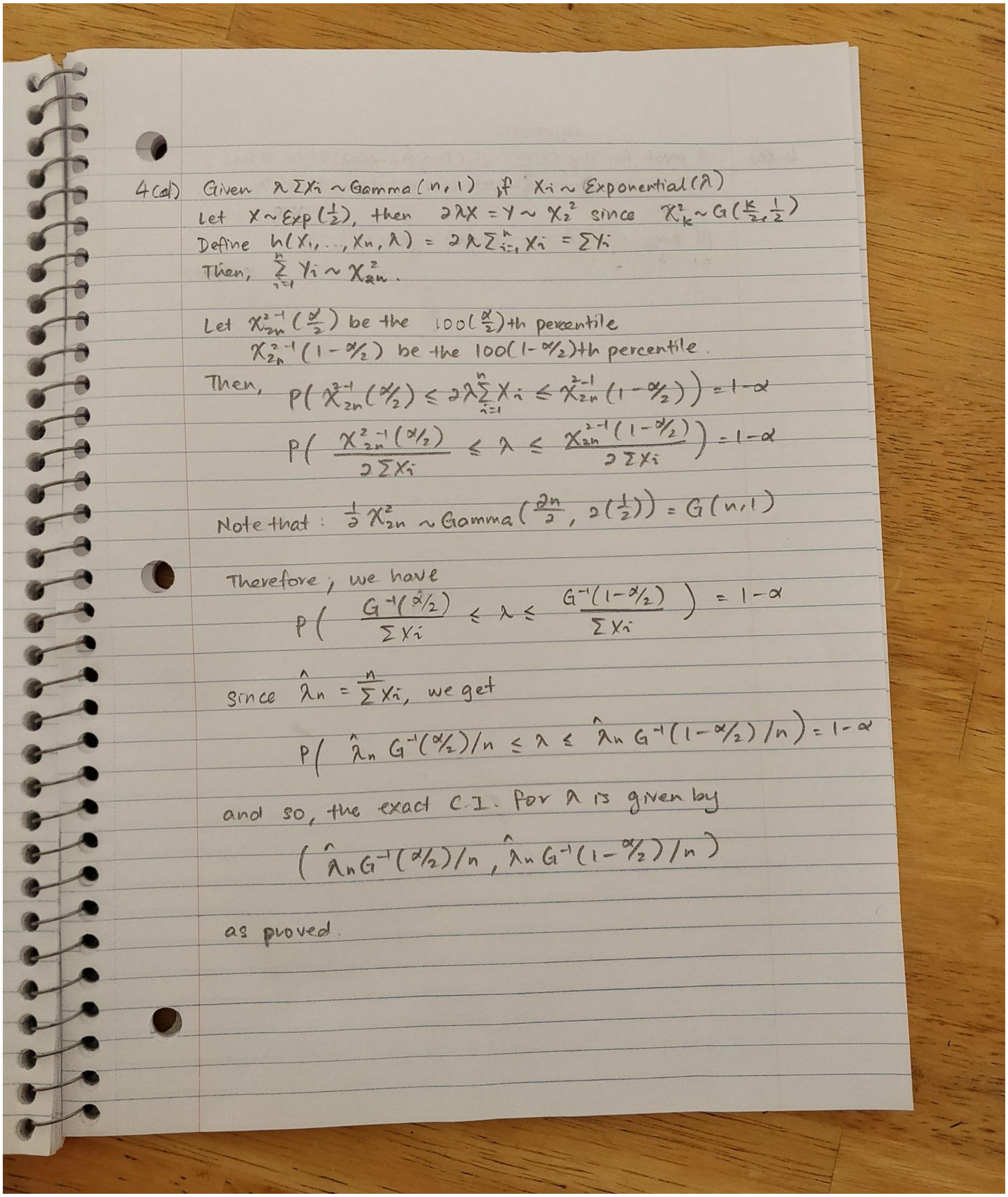
since it is distributed $N(0,1)$, it is symmetric so for a $100(1-\alpha)$CI we have that

$B = z_{\alpha/2}$

$A = -z_{\alpha/2}$

$\Rightarrow$ we have $\left(\hat{\lambda}_n e^{-z_{\alpha/2}/\sqrt{n}}, \hat{\lambda}_n e^{z_{\alpha/2}/\sqrt{n}}\right)$

Figure 1: Proof for 4c.



4(d) Given $\lambda \Sigma X_i \sim \text{Gamma}(n, 1)$ if $X_i \sim \text{Exponential}(\lambda)$

Let $X \sim \text{Exp}(\frac{1}{2})$, then $2\lambda X = Y \sim \chi_2^2$ since $\chi_k^2 \sim G(\frac{k}{2}, \frac{1}{2})$

Define $h(X_1, \ldots, X_n, \lambda) = 2\lambda \sum_{i=1}^n X_i = \Sigma Y_i$

Then, $\sum_{i=1}^n Y_i \sim \chi_{2n}^2$.

Let $\chi_{2n}^{2^{-1}}(\frac{\alpha}{2})$ be the $100(\frac{\alpha}{2})$th percentile

$\chi_{2n}^{2^{-1}}(1-\alpha/2)$ be the $100(1-\alpha/2)$th percentile.

Then,
$$P\left(\chi_{2n}^{2^{-1}}(\alpha/2) \leq 2\lambda \sum_{i=1}^n X_i \leq \chi_{2n}^{2^{-1}}(1-\alpha/2)\right) = 1-\alpha$$

$$P\left(\frac{\chi_{2n}^{2^{-1}}(\alpha/2)}{2\Sigma X_i} \leq \lambda \leq \frac{\chi_{2n}^{2^{-1}}(1-\alpha/2)}{2\Sigma X_i}\right) = 1-\alpha$$

Note that: $\frac{1}{2}\chi_{2n}^2 \sim \text{Gamma}\left(\frac{2n}{2}, 2(\frac{1}{2})\right) = G(n, 1)$

Therefore, we have
$$P\left(\frac{G^{-1}(\alpha/2)}{\Sigma X_i} \leq \lambda \leq \frac{G^{-1}(1-\alpha/2)}{\Sigma X_i}\right) = 1-\alpha$$

since $\hat{\lambda}_n = \frac{n}{\Sigma X_i}$, we get
$$P\left(\hat{\lambda}_n G^{-1}(\alpha/2)/n \leq \lambda \leq \hat{\lambda}_n G^{-1}(1-\alpha/2)/n\right) = 1-\alpha$$

and so, the exact C.I. for λ is given by
$$\left(\hat{\lambda}_n G^{-1}(\alpha/2)/n, \hat{\lambda}_n G^{-1}(1-\alpha/2)/n\right)$$

as proved.

Figure 2: Proof for 4d.