

Question 1

Integrating to find the median gives

$$\int_0^m \theta e^{-\theta x} dx = [-e^{-\theta x}]_0^m = 1 - e^{-\theta m} = 1/2 \quad (1)$$

Therefore $\theta = \log(2)/m$ and $g(x | m) = \log(2)/(m \cdot 2^{x/m})$.

Question 2

The program Code 2.1, referenced on page 14, produced the exponential random variables x_i from the uniform random variables u_i in Table 1. It also plotted the log likelihood function in Figure 1 using the same data.

i	u_i	x_i
1	0.15237801896922	0.1377670997206
2	0.82581697748955	1.4563738992246
3	0.53834243526006	0.64410988715659
4	0.99613471662689	4.6297669015646
5	0.078175528753184	0.067833709839194
6	0.44267826977545	0.48717716099783

Table 1: Exponential Variables with $n = 6$ and $\theta_0 = 1.2$

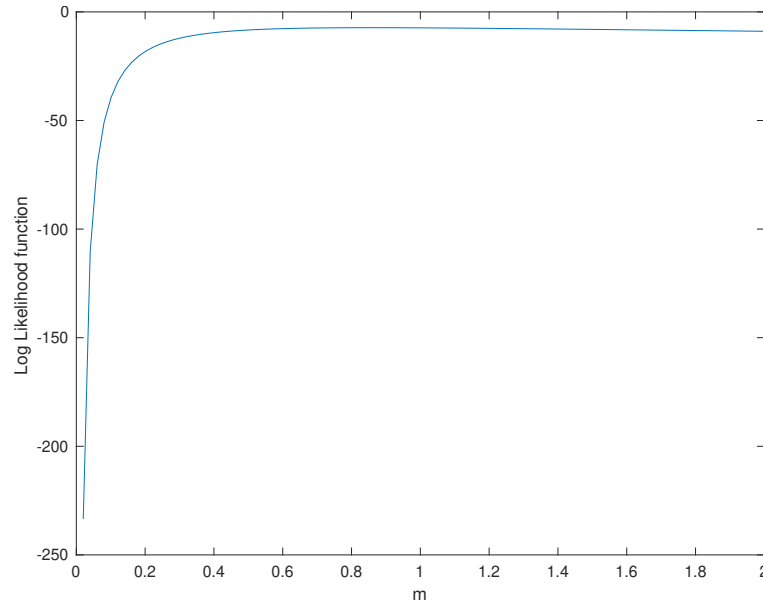


Figure 1: Plot of the log likelihood function against m with $n = 6$

Expanding $\ell_n(m)$ gives

$$\ell_n(m) = n \log(\log(2)) - n \log(m) - \frac{\log(2) S}{m} \quad (2)$$

where $S = \sum_i x_i$. Differentiating with respect to m gives the maximum value $\hat{m}_n = \log(2) \sum_i x_i / n$. (This is indeed a maximum since ℓ_n is continuous, and its endpoints $(0, \infty)$ tend to $-\infty$. Therefore the maximum must occur at the only turning point). For the example given above, we calculate \hat{m}_6 as

0.857541897642894, which is not a good approximation to the actual value $m_0 = 0.5776226505$. This is because the variance in \hat{m}_n is

$$\frac{\log(2)^2}{n^2} \sum_{i=1}^n \text{Var}(x_i) = \frac{\log(2)^2}{n} \times \frac{1}{\theta^2} = \frac{m_0^2}{n} \quad (3)$$

and so when n is small, this value is likely not to be accurate. However as $n \rightarrow \infty$ this variance will tend to 0, and so $\hat{m}_n \rightarrow \mathbb{E}(\hat{m}_n) = \log(2)/\theta_0 = m_0$.¹ We deduce further that this maximum likelihood estimator is unbiased.

Question 3

Code 2.1 was also used to produce Figure 2, a plot of the log likelihood function for different values of n .

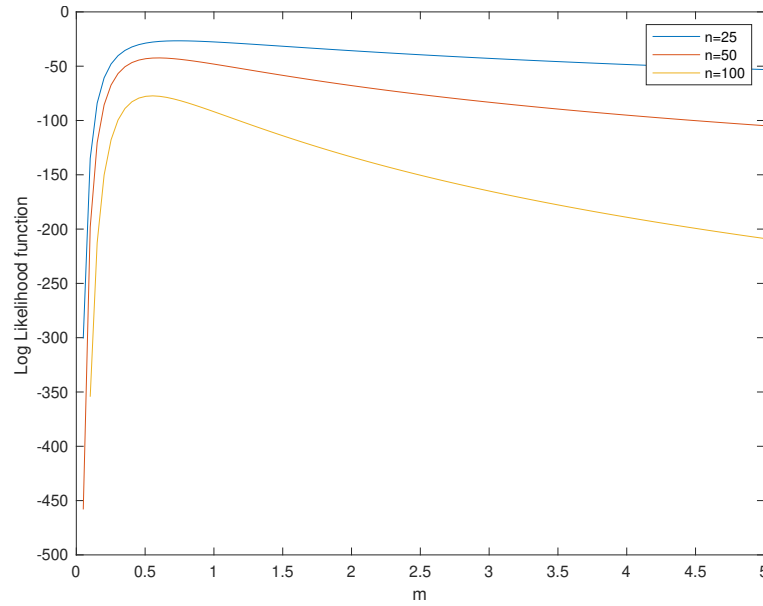


Figure 2: Plot of the log likelihood function against m with different n

n	\hat{m}_n
25	0.73960458432
50	0.59504738029
100	0.55249375490
True Value m_0	0.5776226505

Table 2: \hat{m}_n for data in Figure 2

Firstly note that Table 2 confirms the idea that $\hat{m}_n \rightarrow m_0$. We also observe that as n increases, the height of each maximum is decreasing. The value of the log likelihood function at its maximum is

$$-n \left(1 + \log \left(\frac{S}{n} \right) \right) \quad (4)$$

By the law of large numbers, when n is big we may consider $S/n \approx 1/\theta_0$. So for $\theta_0 < e$ the maximums will decrease, and for $\theta_0 > e$ the maximums will increase. Certainly when $\theta_0 = 1.2$ the values of the maximums are decreasing.

¹Note that the same result can also be proved using the law of large numbers.

We can see that $\ell_n(m)$ tends to $-\infty$ as θ grows very small. This makes sense, since $f(x, |m=0) = 0$ and so taking the logarithm means this will occur for any values of x_i . Notice that as n increases, the log-likelihood function decays faster. This is because when m is large, $\ell_n(m) \approx -n \log(m)$, and so visibly decays quicker.

Question 4

Expanding $M_X(\lambda)$ gives

$$\mathbb{E}(e^{\lambda X}) = \int_0^\infty e^{\lambda x} f(x|\theta) dx = \theta \int_0^\infty e^{(\lambda-\theta)x} dx = \frac{\theta}{\lambda-\theta} \left[e^{(\lambda-\theta)x} \right]_0^\infty \quad (5)$$

Therefore $M_X(\lambda)$ equals $\theta/(\lambda-\theta)$ for $\lambda < \theta$ and does not exist for $\lambda \geq \theta$.

For the second part we will aim to prove a stronger statement; if $X \sim \Gamma(p, \theta)$ and $Y \sim \Gamma(q, \theta)$ are independent, then $X + Y \sim \Gamma(p + q, \theta)$. The moment generating function $M_{X+Y}(t)$ is equal to $\mathbb{E}(\exp(t(X + Y))) = \mathbb{E}(\exp(tX)) \mathbb{E}(\exp(tY))$ (since independent) $= (1 - t/\theta)^{-(p+q)} = M_Z(t)$ where $Z \sim \Gamma(p + q, \theta)$. We now use the fact that if two random variables have the same moment generating function in some non-zero interval around 0 then they are identically distributed.² We are therefore done, since these moment generating functions are identical in the interval $t \in [-\theta/2, \theta/2]$. In the case of Question 4, notice that the exponential distribution is identical to a $\Gamma(1, \theta)$ distribution, and so the result follows immediately.

Question 5

Integrating by parts twice gives

$$F(x) = \int_0^x \theta^2 t e^{-\theta t} dt = \theta^2 \left[-\frac{t e^{-\theta t}}{\theta} - \frac{e^{-\theta t}}{\theta^2} \right]_0^x = 1 - (\theta x + 1)e^{-\theta x} \quad (6)$$

Let $\omega = -\theta x - 1$, so if $F(x) = k$, then $(k - 1)/e = \omega e^\omega$, or $\omega = W\left(\frac{k-1}{e}\right)$, where W represents the Lambert W function. Assume that we could solve for x in closed form. This would imply we could solve for ω , and hence W , in closed form too. However we know that the Lambert W function cannot be solved in closed form, and so we conclude that neither can x .³

Question 6

The log likelihood function is now equal to

$$\ell_n = \log \left(\theta^{2n} e^{-\theta S} \cdot \prod_{i=1}^n x_i \right) = 2n \log(\theta) - \theta S + \sum_{i=1}^n \log(x_i) \quad (7)$$

Differentiating with respect to θ gives

$$0 = \frac{d\ell_n}{d\theta} = \frac{2n}{\hat{\theta}_n} - S \quad (8)$$

or that a maximum occurs at $\hat{\theta}_n = 2n/S$ (where S is the same as in Question 2). This is indeed a maximum since $\ell_n(\theta) \rightarrow -\infty$ at $\theta = 0, \infty$. By the law of large numbers, for large n , $S/n \rightarrow \mathbb{E}(X)$ (where X has p.d.f $f(x|\theta)$). $\mathbb{E}(X) = \left[(-\theta x^2 - 2x - 2/\theta) e^{-\theta x} \right]_0^\infty = 2/\theta_0$, and so $\hat{\theta}_n \rightarrow \theta_0$.

²The proof of this is laid out in Geng, *The Mathematical Theory of Moment Generating Functions*

³The proof that the Lambert W function cannot be written in closed form is laid out in Rosenlicht, "On the explicit solvability of certain transcendental equations".

Question 7

The program Code 7.1, referenced on page 15, was used to produce Figure 3 and the maximum likelihood estimators in Table 3.

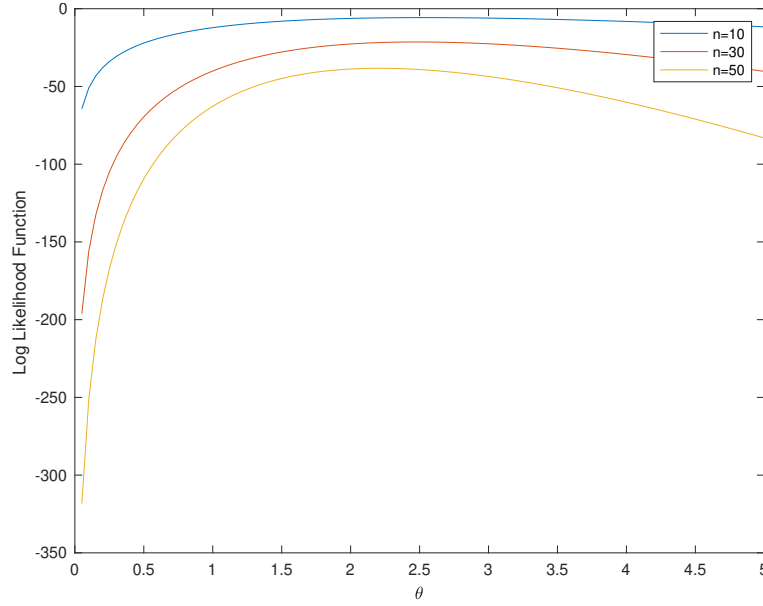


Figure 3: $\ell_n(\theta)$ against θ with $n = 10, 30$ and 50

n	$\hat{\theta}_n$
10	2.52903145291339
30	2.4787073067176
50	2.20599743507643
Actual Value θ_0	2.2

Table 3: $\hat{\theta}_n$ for data in Figure 3

Firstly we can verify that the m.l.e. tends towards θ_0 . Secondly in Figure 3 the log likelihood function appears strictly concave, whereas in Figure 2 this is not the case. We deduce from equation (7) that $\ell_n'' = -2n/\theta^2 < 0$ for all θ .⁴ What's interesting about this is that no matter the values of x_i , ℓ_n will always be strictly concave, unlike the first case where it is only strictly concave in the region $(0, \log(4)S/n)$.

Another difference is that $\ell_n(\theta)$ decays constantly with gradient $-S$ as $\theta \rightarrow \infty$, unlike the $-n \log(m)$ decay seen in Figure 2. We can still see ℓ_n tending to $-\infty$ for large and small θ . This is similar to Figure 2, and still would be the case if we parametrised ℓ_n in terms of m . This is because from equation (6) we can deduce that $\theta m = K$ for some constant K . Therefore if $\theta \rightarrow \infty$, $m \rightarrow 0$, and likewise the other way around, so this property is preserved when parametrisising in terms of θ or m .

Question 8

We can analytically derive the p.d.f. for $\hat{\theta}_n = 2n/\sum_i x_i$. We know each $x_i \sim \Gamma(2, \theta)$, so $\hat{\theta}_n \sim 2n/Y$ where $Y \sim \Gamma(2n, \theta)$ (using Question 4). Defining the 1-1 function $Z = g(Y) = 2n/Y$ means the p.d.f.

⁴Here ℓ_n'' represents the second derivative of the log likelihood function with respect to θ .

of Z , $f_Z(y)$ is

$$\begin{aligned}
 f_Z(y) &= f_Y(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| \\
 &= \frac{\theta^{2n}}{\Gamma(2n)} \left(\frac{2n}{y} \right)^{2n-1} \exp\left(\frac{-2n\theta}{y}\right) \times \frac{2n}{y^2} \\
 &= \frac{(2n\theta)^{2n}}{\Gamma(2n) y^{2n+1}} \exp\left(\frac{-2n\theta}{y}\right)
 \end{aligned} \tag{9}$$

Code 8.1, referenced on page 16, was used to produce Figures 4 to 6 to investigate the distribution of $\hat{\theta}_n$. (The exact p.d.f was superimposed, along with the true value θ_0).

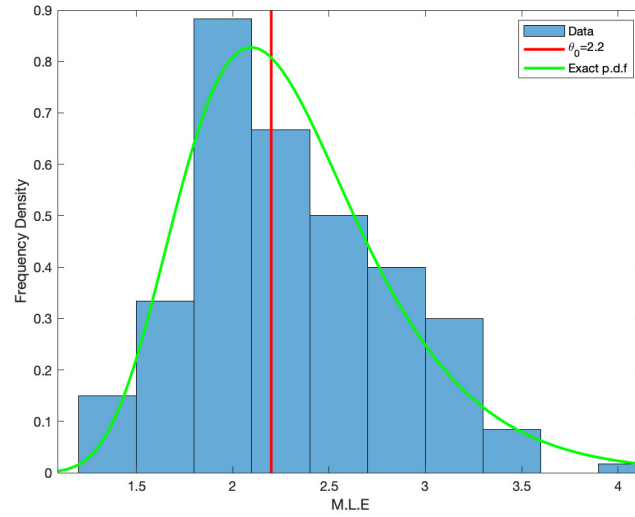


Figure 4: P.d.f. of m.l.e with $n = 10$

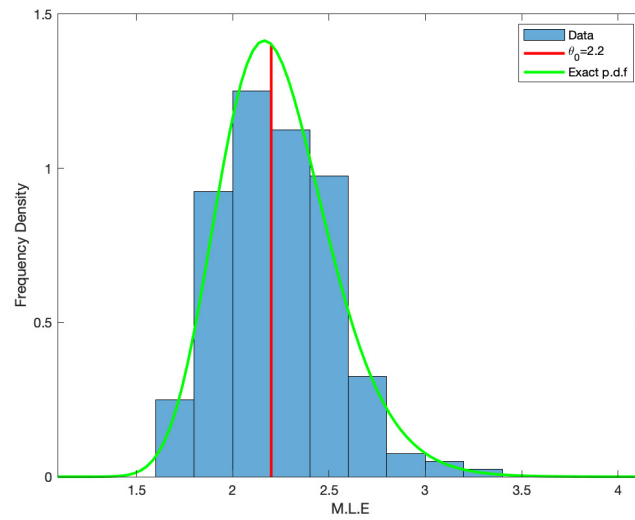


Figure 5: P.d.f. of m.l.e with $n = 30$

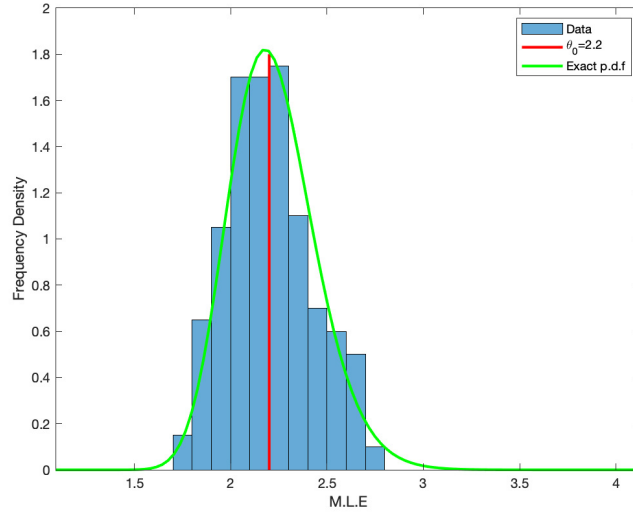


Figure 6: P.d.f. of m.l.e with $n = 50$

n	Average m.l.e	Variance from θ_0
10	2.31165782061807	0.27443154548354
30	2.23204239273873	0.0864587928435113
50	2.20292403340839	0.0482363078542326

Table 4: Average and Variance in Histogram Data

Firstly note that these histograms are approximations of the p.d.f. for $\hat{\theta}_n$, as they plot the (sampled) frequency density.⁵ Observe that as n increases, it appears the variance from θ_0 shrinks to 0 and the expectation tends to θ_0 . We hypothesise that the p.d.f $f_Z(y)$ approaches $\delta(y - \theta_0)$ for large n (δ represents the Dirac delta function). To better understand how $\hat{\theta}_n$ behaves, we can calculate the exact expectation and variance using equation (9):

$$\begin{aligned}
\mathbb{E}(\hat{\theta}_n) &= \frac{(2n\theta)^{2n}}{\Gamma(2n)} \int_0^\infty \frac{y}{y^{2n+1}} \exp\left(\frac{-2n\theta}{y}\right) dy \\
&= \frac{2n\Gamma(2n-1)}{\Gamma(2n)} \theta_0 \\
&= \frac{2n\theta_0}{2n-1}
\end{aligned} \tag{10}$$

and

$$\begin{aligned}
\text{Var}(\hat{\theta}_n) &= \frac{(2n\theta)^{2n}}{\Gamma(2n)} \int_0^\infty \frac{y^2}{y^{2n+1}} \exp\left(\frac{-2n\theta}{y}\right) dy - \left(\frac{2n\theta_0}{2n-1}\right)^2 \\
&= \frac{4n^2\Gamma(2n-2)\theta_0^2}{\Gamma(2n)} - \left(\frac{2n\theta_0}{2n-1}\right)^2 \\
&= \frac{4n^2\theta_0^2}{(2n-1)^2(2n-2)}
\end{aligned} \tag{11}$$

This confirms that the variance does indeed tend to 0 as $n \rightarrow \infty$. We also find that this estimator is positively biased (e.g. $\mathbb{E}(\hat{\theta}_n) > \theta_0$), however it is asymptotically unbiased. (as $n \rightarrow \infty$, $\mathbb{E}(\hat{\theta}_n) = \theta_0$). This explains that all three estimates of the m.l.e. in Table 4 are greater than the true value. Using

⁵These histograms have been normalised so that they match with the exact p.d.f.

these equations we can also show that the mean squared error equals

$$\frac{2(n+1)\theta_0^2}{(2n-1)(2n-2)} \quad (12)$$

It's interesting to see that if we sample one point ($n = 1$) the mean squared error is ∞ , implying that we must sample 2 or more points to infer any information about θ_0 .

Finally we can also deduce that since the log-likelihood function is always strictly convex, no matter the values of x_i or θ_0 , then $\hat{\theta}_n$ will always be the *unique* m.l.e. This means that $\hat{\theta}_n$ is minimal sufficient.⁶

Question 9

Note that the map $(V, \Phi) \rightarrow (X, Y)$ defined in the project booklet is 1-1 since we have restricted $0 \leq \Phi < 2\pi$, and so we may use the formula given. We deduce

$$\begin{aligned} V &= \left(\frac{X - \mu_1}{\sigma} \right)^2 + \left(\frac{Y - \mu_2}{\sigma} \right)^2 \\ \tan(\Phi) &= \frac{Y - \mu_2}{X - \mu_1} \end{aligned} \quad (13)$$

Although the second of these equations is not 1-1, it is a necessary condition on Φ and sufficient to deduce the Jacobian:

$$\begin{aligned} \left| \frac{\partial(\phi, v)}{\partial(x, y)} \right| &= \left| \frac{-(y - \mu_2)}{(x - \mu_1)^2 + (y - \mu_2)^2} \times \frac{2(y - \mu_2)}{\sigma^2} - \frac{(x - \mu_1)}{(x - \mu_1)^2 + (y - \mu_2)^2} \times \frac{2(x - \mu_1)}{\sigma^2} \right| \\ &= \frac{2}{\sigma^2} \end{aligned} \quad (14)$$

We therefore get

$$g(x, y) = \frac{1}{4\pi} \exp \left(-\frac{(x - \mu_1)^2}{2\sigma^2} - \frac{(y - \mu_2)^2}{2\sigma^2} \right) \times \frac{2}{\sigma^2} \quad (15)$$

and hence the desired result. Notice that since $g(x, y)$ is the product of two normal distributions we can conclude that these variables are independent.

Question 10

Code 10.1, referenced on page 18, was used to produce 1000 normally distributed variables. Plotting a histogram of them in Figure 7 verifies that the code has worked.

⁶This is because if $T(X)$ is a sufficient statistic for θ then we may write the p.d.f. $f(x|\theta) = g(T(x), \theta) h(x)$. Maximising over θ will yield a result only dependant on $T(x)$, and since there is one m.l.e, $\hat{\theta}_n$ will always be a function of T .

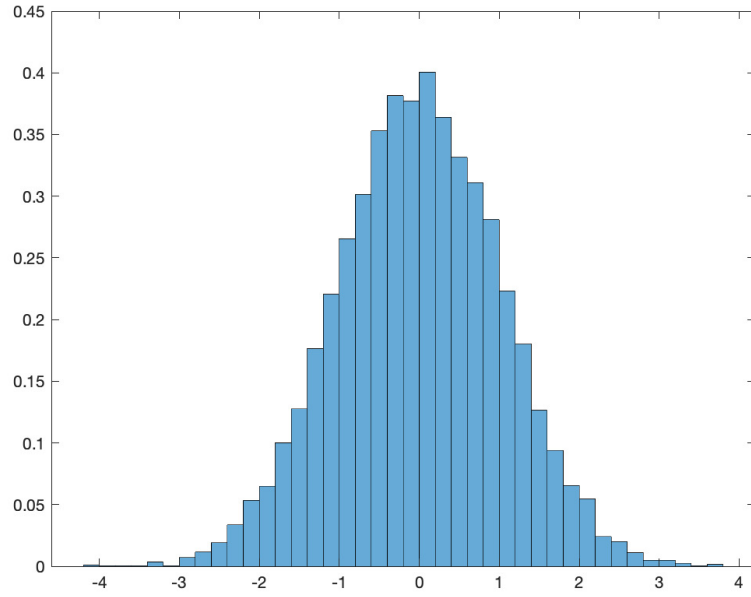


Figure 7: Test Data for Normal Distribution

To find a $100\alpha\%$ confidence interval we first must construct a pivot quantity. Suppose $X_1, X_2, \dots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ are independent and σ^2 is known. Note that $\bar{X} = \sum_i X_i/n \sim \mathcal{N}(\mu, \sigma^2/n)$ and so $Z = \sqrt{n/\sigma^2}(\bar{X} - \mu) \sim \mathcal{N}(0, 1)$ for all μ and σ^2 . If we take γ_α to equal the $(1 + \alpha)/2$ upper quantile of a $\mathcal{N}(0, 1)$ distribution then

$$\begin{aligned} \alpha &= \mathbb{P}\left(-\gamma_\alpha \leq \sqrt{n/\sigma^2}(\bar{X} - \mu) \leq \gamma_\alpha\right) \\ &= \mathbb{P}\left(\bar{X} - \gamma_\alpha \sqrt{\sigma^2/n} \leq \mu \leq \bar{X} + \gamma_\alpha \sqrt{\sigma^2/n}\right) \end{aligned} \quad (16)$$

Hence we have a $100\alpha\%$ confidence interval for μ : $[\bar{X} \pm \gamma_\alpha \sqrt{\sigma^2/n}]$. In the case of Question 10, we set $\sigma = 1$ and $\gamma_{0.8} \approx 1.2815516$ (the 90% quantile of a standard normal variable). Our confidence interval then becomes

$$\left[\bar{X} - \frac{1.2815516}{\sqrt{n}}, \bar{X} + \frac{1.2815516}{\sqrt{n}}\right] \quad (17)$$

Question 11

When we set $n = 100$, our confidence interval for μ becomes

$$[\bar{X} - 0.12815516, \bar{X} + 0.12815516] \quad (18)$$

Code 11.1 referenced on page 19 was used to generate Table 5 below.

\bar{X}	Lower Bound for C.I.	Upper Bound for C.I.	True/False for interval
-0.11141056385008	-0.23956572040454	0.016744592704381	1
-0.097442634600331	-0.22559779115479	0.030712521954129	1
-0.049409489376418	-0.17756464593088	0.078745667178042	1
-0.065699270002723	-0.19385442655718	0.062455886551737	1
0.081887773925232	-0.046267382629228	0.21004293047969	1
-0.023318872858958	-0.15147402941342	0.1048362836955	1
-0.0038378606508644	-0.13199301720532	0.1243172959036	1
-0.17261859581236	-0.30077375236682	-0.044463439257898	0
0.053820604769822	-0.074334551784638	0.18197576132428	1
-0.2112695160693	-0.33942467262376	-0.083114359514837	0
-0.0024082818682331	-0.13056343842269	0.12574687468623	1
0.090217072305278	-0.037938084249182	0.21837222885974	1
0.027318990321538	-0.10083616623292	0.155474146876	1
0.096422038106622	-0.031733118447838	0.22457719466108	1
-0.091935809382862	-0.22009096593732	0.036219347171598	1
0.022751653259254	-0.10540350329521	0.15090680981371	1
-0.11744730176472	-0.24560245831918	0.01070785478974	1
0.0022436660630844	-0.12591149049138	0.13039882261754	1
-0.045793939331109	-0.17394909588557	0.082361217223351	1
-0.027257261678682	-0.15541241823314	0.10089789487578	1
-0.014931886796702	-0.14308704335116	0.11322326975776	1
0.17818514704297	0.050029990488512	0.30634030359743	0
0.15136875657566	0.023213600021196	0.27952391313012	0
-0.13501257434764	-0.2631677309021	-0.0068574177931814	0
0.18822715308006	0.060071996525596	0.31638230963452	0
Total			19

Table 5: Confidence Intervals for Normal Variable

There were 6 cases where the mean did not lie in the confidence interval. The expected number of times the mean would not lie in the confidence interval is $(1 - 0.8) \times 25 = 5$, so the value of 6 is reasonable.

Question 12

Using the general form of a $100\alpha\%$ confidence interval in Question 10 we see that changing μ has no effect, however changing n does. If we keep the same interval $[\bar{X} \pm 0.12815516]$ (the confidence interval for Question 11), when $n = 50$ and $\mu = 4$, then the new upper quantile $\gamma_{\tilde{\alpha}}$ would equal $0.12815516 \times \sqrt{50} = 0.90619$. We calculate $\Phi(\gamma_{\tilde{\alpha}}) = (1 + \tilde{\alpha})/2 = 0.81758$ (where Φ is the cumulative distribution function for a standard normal variable) and so $\tilde{\alpha} = 0.635167$. This implies that on average $(1 - 0.635) \times 25 \approx 9$ cases will lie outside this confidence interval. This conclusion was tested by Code 11.1⁷, however lines 21 and 22:

```
Data(Counter,2)=sum(Variables)/n-norminv((1+alpha)/2)*sqrt(SigmaSquared/n);
Data(Counter,3)=sum(Variables)/n+norminv((1+alpha)/2)*sqrt(SigmaSquared/n);
```

were replaced with:

```
Data(Counter,2)=sum(Variables)/n-0.12815516;
Data(Counter,3)=sum(Variables)/n+0.12815516;
```

Using this modified code, the number of cases that lied outside the interval was calculated as 11, which is close to 9. This code was run 50 times and the average number of cases outside the interval was recorded as 9.36. This confirms the conclusion.

⁷The function `NormGenerate`, defined in Code 10.1, was used implicitly and has not been referenced at the end of this project.

Question 13

We aim to prove that a χ_n^2 distribution with n degrees of freedom is the same as a $\Gamma(n/2, 1/2)$ distribution. It is sufficient to prove the two moment generating functions of these distributions are the same in some non-zero interval around 0. From Question 4 we already know the m.g.f. of $Z \sim \Gamma(\alpha, \beta)$ is $\mathbb{E}(e^{-tZ}) = (1-t/\beta)^{-\alpha}$ for $t < \alpha$ and does not exist for $t \geq \alpha$. In the case $X \sim \chi_n^2$ we find $\mathbb{E}(e^{tX}) = \mathbb{E}(e^{tX_1^2 + \dots + tX_n^2})$ where each $X_i \sim \mathcal{N}(0, 1)$ are independent. Hence

$$\mathbb{E}(e^{tX}) = \prod_{i=1}^n \mathbb{E}(e^{tX_i^2}) = \left(\mathbb{E}(e^{tX_1^2})\right)^n \quad (19)$$

by independence. We know

$$\begin{aligned} \mathbb{E}(e^{tX_1^2}) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(1-2t)x^2\right) dx \\ &= \sqrt{\frac{1}{1-2t}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sqrt{(1-2t)^{-1}}} \exp\left(-\frac{1}{2(1-2t)^{-1}}x^2\right) dx \\ &= \sqrt{\frac{1}{1-2t}} \end{aligned} \quad (20)$$

provided $t < 1/2$ (and does not exist for $t \geq 1/2$). Hence we deduce that the m.g.f. of a χ_n^2 distribution is

$$\mathbb{E}(e^{tX}) = \left(\frac{1}{1-2t}\right)^{n/2} \quad (21)$$

for $t < 1/2$. This is identical to the m.g.f of the gamma distribution $\Gamma(n/2, 1/2)$ in the interval $t \in (-1/2, 1/2)$. We conclude that these two distributions are the same and that the p.d.f of a χ_n^2 distribution is

$$\frac{x^{n/2-1} e^{-x/2}}{2^{n/2} \Gamma(n/2)} \quad (22)$$

This exact solution has been plotted against the histograms below to better portray the shape of this distribution.

The following figures until the end of this report were produced by Code 13.1, referenced on page 20.⁸ Below are three plots of the estimated χ_1^2 distribution with $n = 100, 300$ and 500 .

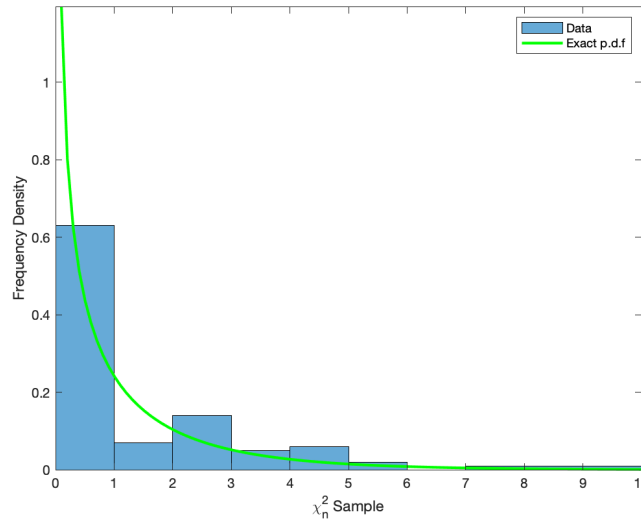


Figure 8: P.d.f. of χ_1^2 distribution with $n = 100$

⁸The function `NormGenerate`, defined in Code 10.1, was used implicitly here and has not been included in the reference.

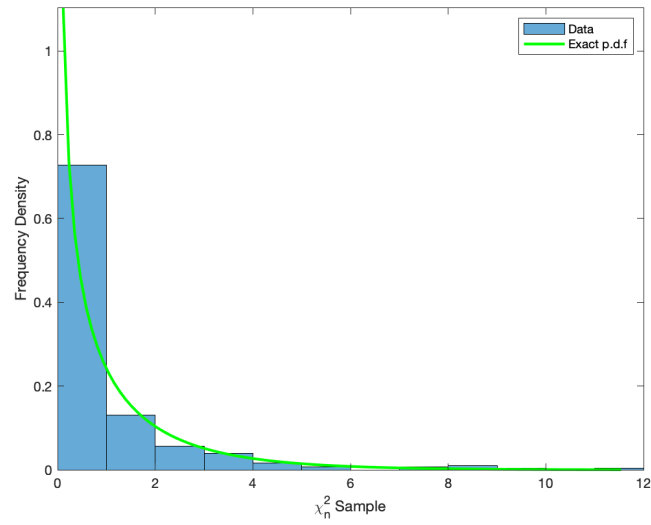


Figure 9: P.d.f. of χ_1^2 distribution with $n = 300$

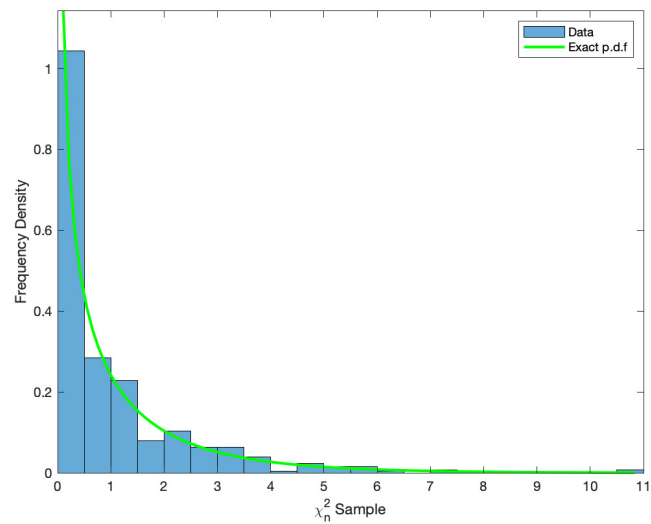


Figure 10: P.d.f. of χ_1^2 distribution with $n = 500$

We can see that as n grows the approximation tends towards the analytic p.d.f. This is because for a larger sample size, outliers do not effect the overall conclusion as much, and so our sample tends towards the exact distribution. Therefore when comparing how this distribution evolves with time we should set $n = 500$ to be most accurate. The following two figures are plots of the χ_5^2 and χ_{40}^2 when $n = 500$.

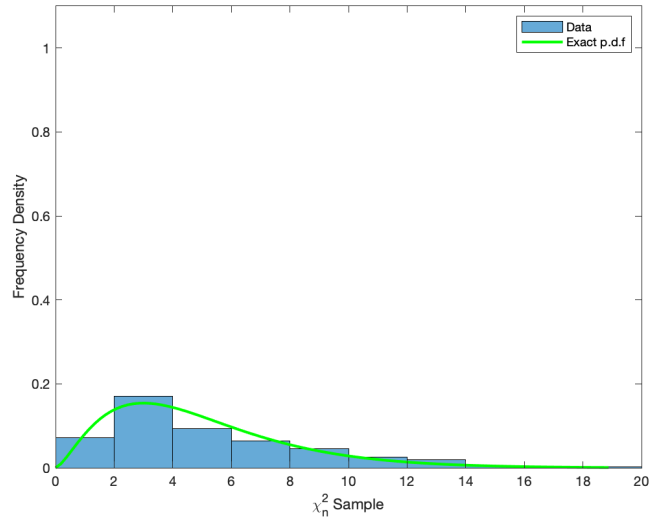


Figure 11: P.d.f. of χ_5^2 distribution with $n = 500$

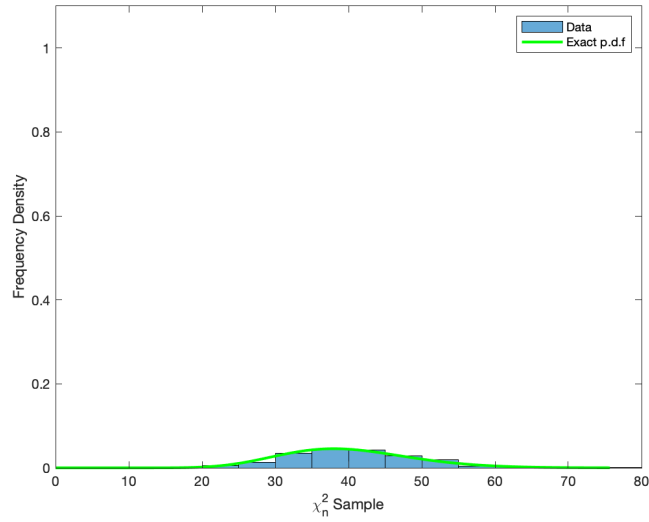


Figure 12: P.d.f. of χ_{40}^2 distribution with $n = 500$

Using the p.d.f. in equation (22), we can deduce the expectation and variance of $X_n \sim \chi_n^2$ distribution:

$$\begin{aligned}\mathbb{E}(X_n) &= \frac{1}{2^{n/2}\Gamma(n/2)} \int_0^\infty x^{n/2} e^{-x/2} dx \\ &= \frac{2\Gamma(n/2 + 1)}{\Gamma(n/2)} = n\end{aligned}\tag{23}$$

and

$$\begin{aligned}\text{Var}(X_n) &= \frac{1}{2^{n/2}\Gamma(n/2)} \int_0^\infty x^{n/2+1} e^{-x/2} dx - n^2 \\ &= \frac{2\Gamma(n/2 + 2)}{\Gamma(n/2)} - n^2 = 2n\end{aligned}\tag{24}$$

This makes sense in the figures above, since, as n increases, we can see the distribution spreading out and shifting to the right. It also appears to be positively skewed, however it looks like it becomes more

symmetrical as n grows large. In fact we claim that the p.d.f approaches a $\mathcal{N}(n, 2n)$ distribution. The proof comes from the central limit theorem; if Y_i are independently distributed as the squares of standard normal variables with $\mu = \mathbb{E}(Y_i)$ and $\sigma^2 = \text{Var}(Y_i)$, then

$$\lim_{n \rightarrow \infty} \left(\sqrt{n} \frac{\frac{X_n}{n} - \mu}{\sqrt{\sigma^2}} \right) \xrightarrow{d} Z \quad (25)$$

where $Z \sim \mathcal{N}(0, 1)$ and $X = \sum_i Y_i \sim \chi_n^2$. From equations (23) and (24) we know $\mu = 1$ and $\sigma^2 = 2$, so $X \xrightarrow{d} \sqrt{n\sigma^2}Z + n\mu \sim \mathcal{N}(n, 2n)$. This proves that the skewness and kurtosis (the measure of how heavily the tails of a distribution differ from that of a normal distribution) both tends to 0.

Also when $n = 1$ or 2 , the mode of this distribution is equal to 0 , however for all other n it is strictly positive. Also the only value of n for which the p.d.f. is unbounded is $n = 1$ (using equation (22)).

References

- Geng, Xi. *The Mathematical Theory of Moment Generating Functions*. URL: <https://researchers.ms.unimelb.edu.au/~xgge@unimelb/Files/Notes/The%20Mathematical%20Theory%20of%20Moment%20Generating%20Functions.pdf>.
- Rosenlight, Maxwell. "On the explicit solvability of certain transcendental equations". eng. In: *Publications Mathématiques de l'IHÉS* 36 (1969), pp. 15–22. URL: <http://eudml.org/doc/103891>.

Code

Code 2.1

```
n=[25,50,100];
dimN=size(n,2);
theta=1.2;
Legend=cell(dimN,1);

Counter=1;
figure
while Counter<=dimN
    Variables=zeros(n(Counter),3);
    Variables(:,1)=1:n(Counter);
    Variables(:,2)=rand(n(Counter),1);
    Variables(:,3)=-log(1-Variables(:,2))/theta;
    % disp(Variables)
    %Produces a latex table
    %latex(sym(vpa(Variables)))
    Max=log(2)*sum(Variables(:,3))/n(Counter);
    disp(Max)

    %Plots the log likelihood function
    m=linspace(0,5);
    y=loglike(Variables(:,3),m,n(Counter));
    plot(m,y)
    hold on
    Legend{Counter,1}=strcat('n=',num2str(n(Counter)));
    Counter=Counter+1;
end
xlabel("m")
ylabel("Log Likelihood function")
legend(Legend)
print('Image_3_4','-depsc')

%This calculates the log likelihood using the data x plotted against m
function answer=loglike(x,m,n)
    count=1;
    answer=1;
    while count<=n
        answer=answer.*g(x(count),m);
        count=count+1;
    end
    answer=log(answer);
end

%This is the p.d.f paramterised in terms of m
function answer=g(x,m)
    answer=log(2)./(m.*2.^(x./m));
end
```

Code 7.1

```
n=[10,30,50];
theta=2.2;

Counter=1;
mle=zeros(3,1);

figure
while Counter<=3
    u=rand(n(Counter),2);
    x=-log(1-u(:,1))/theta-log(1-u(:,2))/theta;
    %Produces a latex table
    %latex(sym(vpa(x)))
    mle(Counter)=2*n(Counter)/sum(x);

    %Plots the log likelihood function
    t=linspace(0,5);
    y=loglike(x,t,n(Counter));
    plot(t,y)
    hold on
    Counter=Counter+1;
end

xlabel("\theta")
ylabel("Log Likelihood Function")
legend('n=10','n=30','n=50')
print('Image_7_1','-depsc')

disp(mle)

function answer=loglike(x,t,n)
    count=1;
    answer=1;
    while count<=n
        answer=answer.*g(x(count),t);
        count=count+1;
    end
    answer=log(answer);
end

function answer=g(x,t)
    answer=t.^2*x.*exp(-t*x);
end
```

Code 8.1

```
n=[10,30,50];
theta=2.2;
Counter=1;
dimN=size(n,2);

MLEStats=zeros(dimN,3);

while Counter<=dimN
    %Calculates random variables for different sample sizes
    data=zeros(n(Counter),200);
    mle_data=zeros(1,200);
    count=1;
    while count<=200
        u=rand(n(Counter),2);
        x=-log(1-u(:,1))/theta-log(1-u(:,2))/theta;
        data(:,count)=x;
        mle_data(1,count)=2*n(Counter)/sum(x);
        count=count+1;
    end

    %Calculating Average and Variance in data
    MLEStats(Counter,1)=sum(mle_data)/200;
    var_mle=0;
    weight_var_mle=0;
    count=1;
    while count<=200
        var_mle=var_mle+(mle_data(1,count)-MLEStats(Counter,1))^2;
        weight_var_mle=weight_var_mle+(mle_data(1,count)-theta)^2;
        count=count+1;
    end
    MLEStats(Counter,2)=var_mle/200;
    MLEStats(Counter,3)=weight_var_mle/200;

    if Counter==1
        XLower=min(mle_data)-0.2;
        XHigher=max(mle_data)+0.2;
    end

    %Plots histogram with appropriate data
    figure
    histogram(mle_data,'Normalization','pdf')
    xlim([XLower,XHigher])
    hold on
    line([theta,theta],ylim,'LineWidth', 2,...
        'Color', 'r');
    hold on
    x=linspace(XLower,XHigher);
    y=pdf(x,2*n(Counter),theta);
    line(x,y,'LineWidth',2,'Color','g')

    legend('Data','\theta_{0}=2.2','Exact p.d.f','Location','northeast')
    xlabel('M.L.E')
    ylabel('Frequency Density')
    print(strcat('Image_8_',num2str(Counter)),'-depsc')
```



```

        Counter=Counter+1;
    end
    %MLE denotes average mle | Variance from average mle | Variance from 2.2
    disp(MLEStats)
    latex(sym(vpa(MLEStats)))

function answer = pdf(x,a,b)
    answer=(a*b)^(a)/((gamma(a)).*(x.^(a+1))).*exp(-b.*a./x);
end

```

Code 10.1

```
format shortg

Variables=NormGenerate(10000,0,1);
histogram(Variables,'Normalization','pdf')

function Variables = NormGenerate(n,mu,SigmaSquared)
    A=rand((n-mod(n,2))/2+mod(n,2),1);
    B=rand((n-mod(n,2))/2+mod(n,2),1);
    Phi=2*pi*A;
    V=-2*log(1-B);
    X=mu+SigmaSquared*sqrt(V).*cos(Phi);
    Y=mu+SigmaSquared*sqrt(V).*sin(Phi);
    Variables=[X;Y(1:n-(n-mod(n,2))/2-mod(n,2),1)];
end
```

Code 11.1

```
format shortg

%Set parameters
n=100;
mu=0;
SigmaSquared=1;
alpha=0.8;
SampleSize=25;

Counter=1;
Data=zeros(SampleSize,4);

disp([num2str(alpha),'% confidence interval is XBar plus or minus ',...
      num2str(norminv((1+alpha)/2)*sqrt(SigmaSquared/n))])

%Calculates random variables and fills table
while Counter<=SampleSize
    Variables=NormGenerate(n,mu,SigmaSquared);

    Data(Counter,1)=sum(Variables)/n;
    Data(Counter,2)=sum(Variables)/n-norminv((1+alpha)/2)*sqrt(SigmaSquared/n);
    Data(Counter,3)=sum(Variables)/n+norminv((1+alpha)/2)*sqrt(SigmaSquared/n);
    Data(Counter,4)=and(mu>=Data(Counter,2),mu<=Data(Counter,3));
    Counter=Counter+1;
end
disp(Data)
% latex(sym(vpa(Data)))

Out=SampleSize-sum(Data(:,4));
disp(['Number of cases outside interval: ',num2str(Out)])
```

Code 13.1

```
DoF=[1,5,40];
mu=0;
SigmaSquared=1;
SampleSize=100;

ChiData=zeros(SampleSize,1);
dimDoF=size(DoF,2);
Counter=1;
while Counter<=dimDoF

    %Calculates random variables by summing and squaring normal distributed
    %ones
    count=1;
    while count<= SampleSize
        n=DoF(Counter);
        Variables=NormGenerate(n,mu,SigmaSquared);
        ChiData(count,1)=dot(Variables,Variables);
        count=count+1;
    end
    %Plots histogram with distribution imposed
    figure
    histogram(ChiData,'Normalization','pdf')
    hold on
    x=linspace(0,max(ChiData));
    y=pdf(x,DoF(Counter)/2,1/2);
    line(x,y,'LineWidth',2,'Color','g')
    xlim([0,inf])
    ylim([0,1.1])
    legend('Data','Exact p.d.f','Location','northeast')
    xlabel('\chi^2_{n} Sample')
    ylabel('Frequency Density')
    print(strcat('Image_11_',num2str(Counter)),'-depsc')
    Counter=Counter+1;
end

function answer = pdf(x,a,b)
    answer=b^(a).*x.^(a-1).*exp(-b.*x)/gamma(a);
end
```