Question 1

The probability density function f for the χ^2_ν distribution (the chi-squared distribution with ν degrees of freedom) is

$$f(x) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{\nu/2-1} e^{-x/2},$$

where the support is $x \in \mathbb{R}$ and $x > 0$, and the degrees of freedom $\nu \in \{1, 2, \ldots\}$.

(a) Let $Y \sim \chi^2_\nu$. Assuming that we know $\mathrm{E}(Y) = \nu$ and $\mathrm{E}\left(Y^2\right) = \nu(\nu + 2)$, find $\mathrm{Var}(Y)$.

(b) Assume that $X_1, X_2, \ldots, X_n \sim \mathrm{N}\left(\mu, \sigma^2\right)$. Show that

$$\mathrm{Var}\left(S^2\right) = \frac{2\sigma^4}{n-1},$$

where S^2 is the sample variance, i.e. $S^2 = \frac{1}{n-1}\sum_{i=1}^n \left(X_i - \overline{X}\right)^2$, as usual.

Solution to Question 1

Part (a):

$$\mathrm{Var}(Y) = \mathrm{E}\left(Y^2\right) - (\mathrm{E}(X))^2 = \nu(\nu+2) - (\nu)^2 = 2\nu$$

Part (b):

Theorem 1.7.2 proves that

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}.$$

Therefore,

$$\begin{aligned}
\mathrm{Var}\left(\frac{(n-1)S^2}{\sigma^2}\right) &= 2(n-1) \\
\Rightarrow \left(\frac{n-1}{\sigma^2}\right)^2 \mathrm{Var}\left(S^2\right) &= 2(n-1) \\
\Rightarrow \mathrm{Var}\left(S^2\right) &= 2(n-1) \cdot \frac{\sigma^4}{(n-1)^2} \\
\Rightarrow \mathrm{Var}\left(S^2\right) &= \frac{2\sigma^4}{n-1}.
\end{aligned}$$

Question 2

Suppose you are on a gameshow and win a prize: there are two containers, labelled A and B, that are each filled with 1000 banknotes, where each banknote is either a £10 note, a £20 note or a £50 note. The prize is that you can choose to keep one of the containers. Before you choose a container, you are told the distributions of the banknotes in each of the two containers is different. Writing the sample space as $\Omega = \{10, 20, 50\}$, and considering the parameter θ as identifying the distribution, $\theta \in \{1, 2\}$, you are given the information that the two distributions are summarised by the following table, where f_θ is the probability mass function of the distribution when the parameter value is θ:

	$\omega = 10$	$\omega = 20$	$\omega = 50$
$f_1(\omega)$	0.3	0.4	0.3
$f_2(\omega)$	0.3	0.1	0.6

Furthermore, you are allowed to sample from one of the containers before making our choice: you can pull exactly one banknote out of exactly one of the containers, and then choose to keep either that container or the other container. We plan to use maximum likelihood estimation to aid us in choosing which container to keep (and you would like to keep the container with the most money).

(a) What is the maximum likelihood estimate (MLE) for θ if you sample $\omega = 50$?

(b) What is expected amount of money in each container?

(c) If you chose to sample from container A and pulled out a £50, would you prefer to keep container A or container B? (Supposing you want to keep the container with the most money.)

(d) Again, suppose you choose to sample from container A and pull out a £50. How many times more (or less) plausible/likely is it that container A is the container with the most money?

(e) What is the MLE for θ if you sample $\omega = 20$?

(f) Suppose you choose to sample from container A and pull out a £20. Would you choose to keep container A or container B?

(g) Again, suppose you choose to sample from container A and pull out a £20. How many times more (or less) plausible/likely is it that container A is the container with the most money?

(h) What is the MLE for θ if you sample $\omega = 10$?

(i) If you chose to sample from container A and pulled out a £10, would you choose to keep container A or container B?

(j) Is the MLE always unique?

Solution to Question 2

Part (a):

Since

$$L(\theta = 2|\omega = 50) = 0.6 > 0.3 = L(\theta = 1|\omega = 50),$$

the MLE is given by $\widehat{\theta} = 2$.

Part (b):

Consider the container cointaining banknotes with distribution $\theta = 1$. Let X_1 be the value of a banknote in this container, and let Y_1 be the amount of money in this container. Then

$$\begin{aligned}
\mathrm{E}(Y_1) = \mathrm{E}(1000X_1) = 1000\mathrm{E}(X_1) &= 1000 \sum_{\omega \in \{10,20,50\}} \omega \cdot f_1(\omega) \\
&= 1000\,(10 \cdot 0.3 + 20 \cdot 0.4 + 50 \cdot 0.3) = 1000(3 + 8 + 15) \\
&= 26000
\end{aligned}$$

So the expected amount in the container with $\theta = 1$ is £23,000.

Now consider the container cointaining banknotes with distribution $\theta = 2$. Let X_2 be the value of a banknote in this container, and let Y_2 be the amount of money in this container. Then

$$\begin{aligned}
\mathrm{E}(Y_2) = \mathrm{E}(1000X_2) = 1000\mathrm{E}(X_2) &= 1000 \sum_{\omega \in \{10,20,50\}} \omega \cdot f_1(\omega) \\
&= 1000\,(10 \cdot 0.3 + 20 \cdot 0.1 + 50 \cdot 0.6) = 1000(3 + 2 + 30) \\
&= 35000
\end{aligned}$$

So the expected value of the amount of money in the container with distribution $\theta = 2$ is £35,000, and so is the most valuable container.

Part (c):

From Part (a) the MLE is $\widehat{\theta} = 2$. In other words container A, which we have sampled from, is most likely to be the container with distribution $\theta = 2$, i.e. the container with the most money, so we would prefer to keep container A.

Part (d):

Since

$$L(\theta = 2|\omega = 50) = 0.6 = 2 \times 0.3 = 2L(\theta = 1|\omega = 50),$$

we could say that it is **twice as likely** that container A is the container with the most money.

Part (e):

$$L(\theta = 1|\omega = 20) = 0.4 > 0.1 = L(\theta = 2|\omega = 20),$$

the MLE is given by $\widehat{\theta} = 1$.

Part (f):

From Part (e) the MLE is $\widehat{\theta} = 1$, so container A is most likely to be the container with distribution $\theta = 1$, i.e. the **not** container with the most money (and the most), so we would prefer to take container B.

Part (g):

Since

$$L(\theta = 1|\omega = 20) = 0.4 = 4 \times 0.1 = 4L(\theta = 2|\omega = 20),$$

we could say that it is **four times less likely** that container A is the container with the most money, as opposed to container B.

Part (h):

$$L(\theta = 1|\omega = 10) = 0.3 = L(\theta = 2|\omega = 10)$$

Therefore, in this case the MLE is given by both $\widehat{\theta} = 1$ and $\widehat{\theta} = 2$.

Part (i):

From Part (h), pulling out a £10 is equally likely to be from the container with distribution $\widehat{\theta} = 1$as it is to be from the container with distribution $\widehat{\theta} = 2$. Therefore, statistically, given the information we have, deciding to keep A or rather take B are equally good decisions.

Part (j):

As Part (i) shows, there can be situations where the MLE is not unique.

Question 3

Suppose that the lifetime of a particular lightbulb is known to be distributed as $\text{Exp}(\theta)$, i.e. an exponential distribution with parameter θ. Suppose that a group of n friends bought a multipack containing n of these lightbulbs and each person keeps one lightbulb to use at home. In a few years' time, they get together and share how long their lightbulbs lasted, and these measurements (lifetimes) are written down as $x_1, x_2, \ldots, x_n$.

(a) Write down the probability density function of the $\text{Exp}(\theta)$ distribution.

(b) Given the sample of measurements $x_1, x_2, \ldots, x_n$, write down the likeihood function for θ based on these measurements.

(c) Write down the log-likelihood of θ given the measurements $x_1, x_2, \ldots, x_n$.

(d) Find the maximum likelihood estimate $\widehat{\theta}$.

(e) Are you sure that $\widehat{\theta}$ is a maximum, or could it be a minimum? If you have not already done so in (d), provide proof that $\widehat{\theta}$ is a maximum/minimum.

Solution to Question 3

Part (a):

Usually it is written with parameter $\lambda > 0$:

$$f(x) = \lambda e^{-\lambda x}$$

But we shall write it with parameter $\theta > 0$:

$$f_\theta(x) = \theta e^{-\theta x}$$

Part (b):

We can assume that each of the lightbulbs have an independent lifetime, and therefore the joint p.d.f. (and therefore joint likelihood) is a product of the individual p.d.f.'s (and likelihoods). Since the likelihood for θ given the sample point x_i ($i \in \{1, 2, \ldots, n\}$) is

$$L(\theta|x_i) = \theta e^{-\theta x_i} = \theta \exp(-\theta x_i),$$

(it will be more useful to use the notation using the function $\exp(\cdot)$), the joint likelihood is

$$L(\theta|\mathbf{x}) = L(\theta|x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} \theta \exp(-\theta x_i) = \theta^n \exp\left(-\theta \sum_{i=1}^{n} x_i\right) = \theta^n \exp(-\theta n \overline{x})$$

Part (c):

Taking logs of both sides of the equation above, we have

$$\begin{aligned}
\log L(\theta|\mathbf{x}) &= \log(\theta^n \exp(-\theta n \overline{x})) \\
&= \log(\theta^n) + \log(\exp(-\theta n \overline{x})) \\
&= n \log \theta - \theta n \overline{x}
\end{aligned}$$

Part (d):

Maximising the log-likelihood is equivalent to maximising the likelihood, and the log-likelihood is in a simpler form. Therefore, taking the derivative with respect to θ,

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\log L(\theta|\mathbf{x}) = \frac{n}{\theta} - n\overline{x}$$

Setting the derivative to 0,

$$\begin{aligned}
&\frac{n}{\theta} - n\overline{x} = 0 \\
\Rightarrow &\frac{n}{\theta} = n\overline{x} \\
\Rightarrow &\frac{1}{\theta} = \overline{x} \\
\Rightarrow &\theta = (\overline{x})^{-1}
\end{aligned}$$

At this point, we do not know if this value for θ is a maximum or a minimum. We therefore need to take the second derivative:

$$\frac{\mathrm{d}^2}{\mathrm{d}\theta^2}\log L(\theta|\mathbf{x}) = -\frac{n}{\theta^2}$$

which is negative for all values of $\mathbf{x}$, because it does not even depend on $\mathbf{x}$. Therefore, $\widehat{\theta} = (\overline{x})^{-1}$ is the maximum likelihood estimate.

Part (e)

It is a maximum; in (d) it was shown that the second derivative of the log-likelihood is negative.

Question 4 (This question is an exercise in calculus)

Returning to, Question 1, the probability density function f for the χ^2_ν distribution is

$$f(x) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{\nu/2-1} e^{-x/2},$$

where, as usual, $\Gamma(z)$ is the gamma function:

$$\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} \mathrm{d}x.$$

(a) Show that the gamma function has the property $\Gamma(z+1) = z\Gamma(z)$ (Hint: use integration by parts).

(b) Show that if $Y \sim \chi^2_\nu$ then $\mathrm{E}(Y) = \nu$ (Hint: try to get the integral into a form that is a constant times 'something' that integrates to 1).

(c) Show that if $Y \sim \chi^2_\nu$ then $\mathrm{E}\left(Y^2\right) = \nu(\nu+2)$.

Solution to Question 4 (This question is an exercise in calculus)

Part (a)

When the argument is $z+1$, the gamma function is:

$$\Gamma(z+1) = \int_0^\infty x^z e^{-x} \mathrm{d}x.$$

Let's use integration by parts, and we choose

$$\begin{aligned} u &= x^z && \Rightarrow \mathrm{d}u = zx^{z-1}\mathrm{d}x \\ dv &= e^{-x}\mathrm{d}x && \Rightarrow v = -e^{-x} \end{aligned}$$

Then

$$\begin{aligned} \int_a^b u\mathrm{d}v &= [uv]_a^b - \int_a^b v\mathrm{d}u \\ \Rightarrow \Gamma(z+1) = \int_0^\infty x^z e^{-x}\mathrm{d}x &= \left[-x^z e^{-x}\right]_0^\infty - \int_0^\infty \left(-e^{-x}\right) zx^{z-1}\mathrm{d}x \\ &= (0-0) + z\int_0^\infty x^{z-1}e^{-x}\mathrm{d}x \\ &= z\Gamma(z) \end{aligned}$$

Part (b)

$$\mathrm{E}\left(Y\right) = \int_0^\infty y f(y)\mathrm{d}y = \int_0^\infty y \cdot \frac{1}{2^{\nu/2}\Gamma(\nu/2)} y^{\nu/2-1} e^{-y/2}\mathrm{d}y = \int_0^\infty \frac{1}{2^{\nu/2}\Gamma(\nu/2)} y^{(\nu/2+1-1)} e^{-y/2}\mathrm{d}y$$

Now, we note three things:

1. $\int_0^\infty \frac{1}{2^{\nu/2}\Gamma(\nu/2)} y^{\nu/2-1}e^{-y/2}\mathrm{d}y = 1$, for any value ν.
2. $2^{\nu/2+1} = 2^{\nu/2} \cdot 2 \Rightarrow 2^{\nu/2} = \frac{1}{2} \cdot 2^{\nu/2+1}$
3. $\Gamma(\nu/2+1) = (\nu/2)\Gamma(\nu/2) \Rightarrow \Gamma(\nu/2) = \frac{2}{\nu} \cdot \Gamma(\nu/2+1)$

Then, using these three things, we manipulate the integral above:

$$\begin{aligned}
\mathrm{E}\left(Y\right) &= \int_0^\infty \frac{1}{2^{\nu/2}\Gamma(\nu/2)} y^{(\nu/2+1-1)} e^{-y/2}\mathrm{d}y \\
&= \int_0^\infty \frac{1}{\left(\frac{1}{2}\cdot 2^{\nu/2+1}\right)\frac{2}{\nu}\cdot\Gamma(\nu/2+1)} y^{(\nu/2+1-1)} e^{-y/2}\mathrm{d}y \\
&= \nu \int_0^\infty \frac{1}{2^{\nu/2+1}\Gamma(\nu/2+1)} y^{(\nu/2+1-1)} e^{-y/2}\mathrm{d}y \\
&= \nu \cdot 1 \\
\Rightarrow \mathrm{E}\left(Y\right) &= 1
\end{aligned}$$

where the integral is 1 because the parameter is $\nu+2$, since $(\nu+2)/2 = \nu/2+1$.

Part (c)

The same idea as for Part (b), but we first note $2^{\nu/2} = 2^{\nu/2+2} \cdot \frac{1}{4}$, and

$$\Gamma(\nu/2+2) = \left(\frac{\nu}{2}+1\right)\Gamma(\nu/2+1) = \left(\frac{\nu}{2}+1\right)\left(\frac{\nu}{2}\right)\Gamma(\nu/2) = \left(\frac{\nu(\nu+2)}{4}\right)\Gamma(\nu/2)$$

Then,

$$\begin{aligned}
\mathrm{E}\left(Y^2\right) &= \int_0^\infty y^2 f(y)\mathrm{d}y = \int_0^\infty y^2 \cdot \frac{1}{2^{\nu/2}\Gamma(\nu/2)} y^{\nu/2-1}e^{-y/2}\mathrm{d}y = \int_0^\infty \frac{1}{2^{\nu/2}\Gamma(\nu/2)} y^{\nu/2+2-1}e^{-y/2}\mathrm{d}y \\
&= \int_0^\infty \frac{1}{\left(\frac{1}{4}\cdot 2^{\nu/2+2-1}\right)\frac{4}{\nu(\nu+2)}\Gamma(\nu/2+2)} y^{\nu/2+2} e^{-y/2}\mathrm{d}y \\
&= \nu(\nu+2)\int_0^\infty \frac{1}{2^{\nu/2+2}\Gamma(\nu/2+2)} y^{\nu/2+2-1}e^{-y/2}\mathrm{d}y \\
&= \nu(\nu+2)\cdot 1 \\
&= \nu(\nu+2)
\end{aligned}$$