(18 marks) In **Question 1** we will make use of the following shiny app:

https://shiny.math.uwaterloo.ca/sas/stat231/teststatistics/

This app can be used to explore test statistics and hypothesis tests. You can first choose a probability distribution and a test statistic. You then specify a value for the model parameter under the null hypothesis. You can then adjust the sample size and set the point estimate of the model parameter resulting from the sample. The right-hand window then displays a plot of the probability distribution corresponding to the test statistic chosen. The plot is then separated into regions based on the value of the resulting test statistic. You should think about how the areas under the probability distribution curves correspond to the resulting p-values.

a) Binomial(n, θ)
On the shiny app:
Select distribution: Binomial distribution
Select test statistic: Asymptotic Gaussian test statistic
Enter H_0 for θ: 0.1
Enter sample size: 25

As you move the slider for MLE of θ, you will see how the value of the test statistic and the corresponding p-value for testing H_0: θ = 0.1 vary as the value of $\hat{\theta}$ varies.

Complete the table below by inserting the Gaussian test statistics value and p-value corresponding to each combination.

a) The Gaussian test statistics value and p-value corresponding to each combination is shown in Tab.1.

Sample Size, n=25			
$\hat{\theta}$	**$\lvert \hat{\theta} - 0.1 \rvert$**	**value of Gaussian test statistic**	**p-value**
0.15	**0.05**	0.833	0.4047
0.20	**0.10**	1.677	0.0956
0.25	**0.15**	2.5	0.0124
0.30	**0.20**	3.333	0.0009
Sample Size, n=40			
0.15	**0.05**	1.054	0.2918
0.20	**0.10**	2.108	0.035
0.25	**0.15**	3.162	0.0016
0.30	**0.20**	4.216	<0.0001

Tab.1 Gaussian test statistics value and p-value corresponding to each combination

i. (2 marks) For the sample size of n=25, how does the p-value for testing H_0: $\theta = 0.1$ change as the quantity $|\hat{\theta} - 0.1|$ increases? Explain why this behaviour makes sense.

ii. (1 mark) For the sample size of n=25, what other value of $\hat{\theta}$ generates the identical test statistic and p-value as when $|\hat{\theta} - 0.1| = 0.15$?

iii. (1 mark) For the sample size of n=25 and $\hat{\theta} = 0.3$, use the information from Table 1 to determine what the p-value is for testing H_0: $\theta = 0.1$ versus the one-sided alternative hypothesis H_A: $\theta > 0.1$.

iv. (2 marks) Now compare your results between sample size n=25 and n=40, How does the p-value for testing H_0: $\theta = 0.1$ change as the sample size increases for a fixed value of $|\hat{\theta} - 0.1|$? Explain why this behaviour makes sense.

i. The p-value decreases as the quantity $|\hat{\theta} - 0.1|$ increases. This is because the p-value is the probability for the population with the hypothesis parameter $\theta = 0.1$ to generate a maximum likelihood estimate $\hat{\theta}$ based on the sample data, the increase of $|\hat{\theta} - 0.1|$ means the increase of the distance between the hypothesis parameter and the maximum likelihood estimate, which decreases the probability for the maximum likelihood estimate based on the sample data to be from a population with the hypothesis parameter $\theta = 0.1$ since the expected value of the maximum likelihood estimator of the hypothesis population is $\theta = 0.1$, which means p-value decreases. Mathematically, since all $y - 0.1n = \hat{\theta}n - 0.1n$, we have

$$p - value \approx P\left(|Z| \geq \frac{|y - 0.1n|}{\sqrt{0.1n(1 - 0.1)}}\right) = 2\left[1 - P\left(Z \leq \frac{|y - 0.1n|}{\sqrt{0.1n(1 - 0.1)}}\right)\right]$$

$$= 2 - 2pnorm(\frac{|\hat{\theta} - 0.1|n}{\sqrt{0.1n(1 - 0.1)}})$$

Note that since when $|\hat{\theta} - 0.1|$ increases, $\frac{|\hat{\theta} - 0.1|n}{\sqrt{0.1n(1 - 0.1)}}$ increases, so $pnorm(\frac{|\hat{\theta} - 0.1|n}{\sqrt{0.1n(1 - 0.1)}})$ increases because pnorm is the cumulative density function of G(0, 1), which is non-decreasing, so $p - value \approx 2 - 2pnorm(\frac{|\hat{\theta} - 0.1|n}{\sqrt{0.1n(1 - 0.1)}})$ decreases as a result.

ii. To generate the identical test statistic and p-value, let $|\hat{\theta} - 0.1| = 0.15$, we have $\hat{\theta} - 0.1 = 0.15$ or $\hat{\theta} - 0.1 = -0.15$, so $\hat{\theta} = -0.15 + 0.1 = -0.05$ gives the same test statistic as when $\hat{\theta} = 0.25$. However, note that the maximum likelihood estimate for a binomial model is the sample proportion, which is non-negative, so we do not have other value generates the identical test statistic and p-value.

iii. According to Tab.1, the p-value for testing H_0 vs. the one-sided upper-tail alternative hypothesis H_A : $\theta > 0.1$ is 0.0009 / 2 = 0.00045. This is because the p-value for two-sided test is 0.0009.

iv. The p-value for testing H_0 decreases as the sample size n increases for each value of $|\hat{\theta} - 0.1|$. This is because as the sample size becomes larger, we know more information about the population parameter, so we have smaller probability for observing the sample data from the hypothesis population. Mathematically, since all y $- 0.1n = \hat{\theta}n - 0.1n \geqslant 0$, we have

$$p - value \approx P\left(|Z| \geq \frac{|y - 0.1n|}{\sqrt{0.1n(1 - 0.1)}}\right) = 2\left[1 - P\left(Z \leq \frac{|y - 0.1n|}{\sqrt{0.1n(1 - 0.1)}}\right)\right]$$

$$= 2 - 2pnorm\left(\frac{\hat{\theta}n - 0.1n}{\sqrt{0.1n(1 - 0.1)}}\right)$$

$$= 2 - 2pnorm\left(\frac{(\hat{\theta} - 0.1)n}{\sqrt{n}\sqrt{0.1(1 - 0.1)}}\right)$$

$$= 2 - 2pnorm\left(\frac{(\hat{\theta} - 0.1)\sqrt{n}}{\sqrt{0.1(1 - 0.1)}}\right)$$

Since pnorm is the cumulative density function of G(0, 1), which is non-decreasing, and when n increases, $\frac{(\hat{\theta} - 0.1)\sqrt{n}}{\sqrt{0.1(1-0.1)}}$ increases, so $pnorm\left(\frac{(\hat{\theta} - 0.1)\sqrt{n}}{\sqrt{0.1(1-0.1)}}\right)$ increases, so $p - value \approx 2 - 2pnorm\left(\frac{(\hat{\theta} - 0.1)\sqrt{n}}{\sqrt{0.1(1-0.1)}}\right)$ decreases.

b) The Likelihood ratio test statistics value and p-value corresponding to each combination is in Tab.2.

$\hat{\theta}$	$\lvert \hat{\theta} - \theta_0 \rvert$	value of LR test statistic	p-value
0.15	0.05	0. 612	0. 4341
0.20	0.10	2. 22	0. 1362
0.25	0.15	4. 617	0. 0317
0.30	0.20	7. 683	0. 0056

Tab.2 Likelihood ratio test statistics value and p-value corresponding to each combination

According to Tab.1 and Tab.2, the p-values in Tab.2 with the p-values in Tab.1 for the sample size n = 25 are close, but each p-value from the Gaussian test in Tab.1 is slightly smaller than the p-value from the Likelihood ratio test in Tab.2. Different conclusions based on different tests: According to Table 5.1 in the lecture note, the p-value for $\hat{\theta} = 0.20$ from Tab.1 is 0.0956, which is smaller than 0.1 and larger than 0.05, so it has weak evidence against H_0 based on the observed data. Whereas the p-value from Tab.2 is 0.1362, which is larger than 0.1, so it has no evidence against H_0 based on the observed data. The evidence for the p-value for $\hat{\theta} = 0.20$ in Tab.1 is stronger; The p-value for $\hat{\theta} = 0.30$ form Tab.1 is 0.0009, which is smaller than 0.001, so it has very strong evidence against H_0. But the p-value in Tab.2 is 0.0056, which is larger than 0.001 and smaller than 0.01, so it has strong evidence against H_0. The p-value in Tab.1 for $\hat{\theta} = 0.30$ is stronger.

c)

i. The test statistics value and p-value corresponding to each sample size for the μ of the Gaussian model is in Tab.3.

Sample Standard deviation 2		
sample size	**value of test statistic for mean**	**p-value**
25	1. 25	0. 2113
35	1. 479	0. 1391
45	1. 677	0. 0935
55	1. 854	0. 0637

Tab.3 The test statistics value and p-value corresponding to each sample size for the μ of the Gaussian model

As the sample size increases, the p-value decreases. This makes sense because when we have a larger sample size, we know more information about the population where the sample data are from, which decreases the probability for the observed data with a different sample standard deviation to be from the hypothesis population. Mathematically,

$$p - value = 2P\left(T \geq \frac{|0.5 - 0|}{\frac{2}{\sqrt{n}}}\right) = 2\left[1 - pt\left(\frac{\sqrt{n}}{4}\right)\right] = 2 - 2\,pt\left(\frac{\sqrt{n}}{4}, n - 1\right)$$

Note that since when the sample size n increases, $\frac{\sqrt{n}}{4}$ increases. Also, since n is large enough, $pt\left(\frac{\sqrt{n}}{4}, n-1\right) \approx pnorm(\frac{\sqrt{n}}{4})$, which is a non-decreasing function. Thus, $2\,pt\left(\frac{\sqrt{n}}{4}, n-1\right)$ increases as n increases, which means $p - value = 2 - 2\,pt\left(\frac{\sqrt{n}}{4}, n-1\right)$ decreases.

ii. The test statistics value and p-value corresponding to each sample standard deviation for the μ of the Gaussian model is in Tab.4.

sample standard deviation	value of test statistic for mean	p-value
0.8	3.125	0.0018
1.2	2.083	0.0372
1.6	1.562	0.1182
2.0	1.25	0.2113

Tab.4 The test statistics value and p-value corresponding to each sample standard deviation for the μ of the Gaussian model

The p-value for testing H_0: μ =0 increases when the sample standard deviation increases. This makes sense because a larger sample deviation indicate possibly bigger variability for the population from which the sample data are from, which increases the probability for the observed data to be from the hypothesis model. Mathematically,

$$p - value = 2P\left(T \geq \frac{|0.5 - 0|}{\frac{s}{\sqrt{25}}}\right) = 2 - 2\,pt\left(\frac{5}{2s}, 24\right)$$

Since when s increases, $\frac{5}{2s}$ decreases. Note that since $\mathrm{pt}(x, 24)$ is the cumulative density function for T~t(24), which is non-decreasing, so $pt\left(\frac{5}{2s}, 24\right)$ decreases when s increases, thus $2 - 2\,pt\left(\frac{5}{2s}, 24\right)$ increases.

iii. The test statistics value and p-value corresponding to each sample standard deviation for the σ of the Gaussian model is in Tab.5.

sample standard deviation	value of test statistic for standard deviation	p-value
4.4	29.04	0.4373
4.7	33.135	0.2025
5.2	40.56	0.0372
5.5	45.375	0.0105

Tab.5 The test statistics value and p-value corresponding to each sample standard deviation for the σ of the Gaussian model

The p-value for testing H_0: $\sigma = 4$ decreases as the sample standard deviation increases. This makes sense because when the sample standard deviation is larger, it indicates that the population from which the observed data are from may have a significant difference with the hypothesis population, so the probability for the observed data to be from the hypothesis population is smaller.

(15 marks) In **Question 2**, we will analyse the **Reached on time variate** from the E-commerce shipping dataset. The purpose of this question is to test a hypothesis of interest.

Recall in Assignment 3 and the midterm we assumed that the Reached on time variate can be modeled as a Binomial(n, θ) distribution. We wish to test the Hypothesis $H_0: \theta = \theta_0$.

a) (1 mark) Quote the maximum likelihood estimate for θ based on your dataset that you obtained from Question 2 of Assignment 3. In addition, give the value of n for your data (missing values should be removed from your count).

b) (This part can be handwritten- 2 marks) Determine d, the observed value of the test statistic

$$D = \frac{|\tilde{\theta}-\theta_0|}{\sqrt{\theta_0(1-\theta_0)/n}}$$

for your data for testing $H_0: \theta = \theta_0$ where $\theta_0 = 0.6$.

c) (2 marks) Use the value of d determined in (b) and the Normal approximation to the Binomial distribution to determine the approximate p-value for testing $H_0: \theta = 0.6$. (A continuity correction is not needed.) State your conclusion regarding $H_0: \theta = 0.6$ based on the approximate p-value.

d) (2 marks) Is the value $\theta = 0.6$ an element of an approximate 95% confidence interval for θ based on the asymptotic Gaussian pivotal quantity? Explain why or why not using only the p-value determined in (c).

e) (This part can be handwritten – 2 marks) Determine the observed value of the likelihood ratio statistic $\lambda(\theta_0)$ where $\theta_0 = 0.6$ for your data.

f) (2 marks) Use the value of $\lambda(0.6)$ and the approximate distribution of the likelihood ratio statistic to determine the approximate p-value for testing $H_0: \theta = 0.6$. State your conclusion regarding $H_0: \theta = 0.6$ based on the approximate p-value.

g) (2 marks) Is your conclusion in (f) the same as your conclusion in (d)? Explain briefly why you would expect this result.

h) (2 marks) Is the value $\theta = 0.6$ an element of a 15% likelihood interval for θ? Explain why or why not without determining the interval.

a) Since the maximum likelihood estimate for θ is the sample proportion of the purchases reached on time, assuming the variate follows a binomial distribution. Then, the sample proportion of the purchases reached on time is 0.5777778 according to the R script output.

```
> theta_hat <- length(dataset$Reached_on_Time[which(dataset$Reached_on_Time == 1)]) / n
> theta_hat
[1] 0.5777778
```

b) Following the given D, we have

$$d = \frac{|\hat{\theta} - \theta_0|}{\sqrt{\frac{\theta_0(1-\theta_0)}{n}}} = \frac{|0.5777778 - 0.6|}{\sqrt{\frac{0.6(1-0.6)}{495}}} \approx 1.009217$$

c) Since $d \approx 1.278019$, we have

$$p - value \approx P(|Z| \geq d) \approx 2 - 2pnorm(1.009217) \approx 0.3128706$$

The output is from R script:

```
> 2-2*pnorm(1.009217)
[1] 0.3128706
```

Because the p-value is about 0.3128706, which is larger than 0.1, we say there is no evidence against H_0: $\theta = 0.6$.

d) Because the p-value for testing H_0: $\theta = 0.6$ is greater than or equal to 0.05 if and only if the value $\theta = 0.6$ is inside the 95% confidence interval, and the p-value $\approx$ 0.2012427 is indeed greater than 0.05, we say $\theta = 0.6$ is inside the 95% confidence interval.

e) First, we calculate the relative maximum likelihood function

$$R(\theta_0) = \left(\frac{\theta_0}{\hat{\theta}}\right)^y \left(\frac{1-\theta_0}{1-\hat{\theta}}\right)^{n-y}$$
$$= \left(\frac{0.6}{0.5777778}\right)^{0.5777778\times495} \left(\frac{1-0.6}{1-0.5777778}\right)^{495-0.5777778\times495}$$
$$\approx 0.6027122$$

Then, we get the likelihood ratio statistic

$$\lambda(\theta_0) = -2log(R(\theta_0)) \approx -2\log(0.6027122) = 1.012631$$

Thus, $\lambda(\theta_0)$ is approximately 1.012631.

f) The p-value is calculated by

$$p - value \approx 2[1 - P(Z \leq \sqrt{\lambda(\theta_0)})] \approx 2 - 2pnorm(\sqrt{1.012631})$$
$$\approx 0.3142734$$

Thus, the p-value is approximately 0.3142734. Since 0.3142734 is larger than 0.1, we say there is no evidence against H_0: $\theta = 0.6$.

g) My conclusion remains the same. This is because mathematically, the likelihood ratio test statistic and the test statistic we used in question (c) are asymptotically equivalent, and since n = 495 is large enough, the p-values generated using the two methods should be very close, and thus give approximately equivalent conclusions.

h) The value $\theta = 0.6$ is not in any 15% likelihood interval for θ. This is because $\theta = 0.6$ is inside a 95% confidence interval as claimed in question (d), thus it is inside a 15% likelihood interval since a 15% interval is approximately a 95% confidence interval. This can be proved by:

$$2P\left(Z \leq \sqrt{-2\log(0.15)}\right) - 1 \approx 0.95$$

(8 marks) In **<u>Question 3</u>**, we will analyse the **Customer care calls** variate from the E-commerce shipping dataset. The purpose of this question is to test a hypothesis of interest.

Recall in the midterm we assumed that the Customer care calls variate can be modeled as a Poisson(θ). We wish to test the Hypothesis $H_0: \theta = \theta_0$.

For purposes of this question you will each generate your own unique value of θ_0 by running the R code below:

set.seed(XXXXXXXX) #replace XXXXXXXX with your student ID
generate a random value of the hypothesised theta from a Uniform(3,4) distribution

theta0<-round(runif(1,3,4),digits=1)   # generate value for theta0

a) (2 marks) Give your values of θ_0, $\hat{\theta}$ and the sample size n for your data (missing values should be removed from your count).
b) (This part can be handwritten – 2 marks) Determine d, the observed value of the test statistic

$$D = \frac{|\tilde{\theta} - \theta_0|}{\sqrt{\theta_0/n}}$$

your generated value of θ_0 and your solved value of $\hat{\theta}$.

c) (2 marks) Use the value of d determined in (b) and the Normal approximation to the Poisson distribution to determine the approximate p-value for testing H_0: $\theta = \theta_0$. (A continuity correction is not needed.) State your conclusion regarding H_0: $\theta = \theta_0$ based on the approximate p-value.
d) (2 marks) Is the value $\theta = \theta_0$ an element of an approximate 95% confidence interval for θ based on the asymptotic Gaussian pivotal quantity? Explain why or why not using only the p-value determined in (c).

a) My $\theta_0 = 3.3$, my $\hat{\theta} \approx$ 3.977778, and my 0sample size n is 495 according to the result of R script:

```
> calls_thetahat
[1] 3.977778
> # Q3a
> theta0
[1] 3.3
> calls_thetahat <- mean(dataset_omitted$Customer_ca
re_calls)
> calls_thetahat
[1] 3.977778
> n <- length(dataset_omitted$Customer_care_calls)
> n
[1] 495
```

b) Following the test statistic, we have

$$d = \frac{|\hat{\theta} - \theta_0|}{\sqrt{\theta_0/n}} = \frac{|3.977778 - 3.3|}{\sqrt{3.3/495}} \approx 8.301049$$

Thus, the observed value of the test statistic is $d \approx 8.301049$.

c) We calculate the p-value by

$$p - value = 2[1 - P(Z \leq d)] = 2 - 2pnorm(8.301049) \approx 0 < 0.001$$

Because the p-value is approximately 0, which is smaller than 0.001, we say there is very strong evidence against H_0: $\theta_0 = 3.3$.

d) Because the p-value for testing H_0: $\theta_0 = 3.3$ is greater than or equal to 0.05 if and only if the value $\theta = 3.3$ is inside the 95% confidence interval, and the p-value is less than 0.001, which means it is not greater than or equal to 0.05, we say $\theta = 3.3$ is not inside the 95% confidence interval.

(7 marks) In **Question 4** we will analyse the **Discount offered variate** from the E-commerce shipping dataset.

Recall in Assignment 1 and the midterm we assumed that the Discount offered variate can be modeled as an Exponential(θ). We wish to test the Hypothesis $H_0: \theta = \theta_0$.

a) (1 mark) Give the values of n and $\hat{\theta}$ for your data. (missing values should be removed from your count).

b) (This part can be handwritten – 2 marks) Determine d, the observed value of the test statistic

$$D = \frac{|\tilde{\theta} - \theta_0|}{\theta_0/\sqrt{n}}$$

for your data for testing H_0: $\theta = \theta_0$ where $\theta_0 = 13$.

c) (2 marks) Use the value of d determined in (b) and the Normal approximation to the Exponential distribution to determine the approximate p-value for testing H_0: $\theta = 13$. State your conclusion regarding H_0: $\theta = 13$ based on the approximate p-value.

d) (2 marks) Is the value $\theta = 13$ an element of an approximate 90% confidence interval for θ based on the asymptotic Gaussian pivotal quantity? Explain why or why not using only the p-value determined in (c).

a) The $\hat{\theta}$ is the sample mean of discount offered, which is about 12.22424, and the sample size n is 495. The R script output is here:

```
> Discount_theta_hat <- mean(dataset_omitted1$Discount_offered)
> Discount_theta_hat
[1] 12.22424
> n <- length(dataset_omitted1$Discount_offered)
> n
[1] 495
```

b) Following the test statistic, we have

$$d = \frac{|\hat{\theta} - \theta_0|}{\theta_0/\sqrt{n}} = \frac{|12.22424 - 13|}{13/\sqrt{495}} \approx 1.327655$$

Thus, the observed value of the test statistic is $d \approx 1.327655$.

c) We calculate the p-value by

$$p - value = 2[1 - P(Z \leq d)] = 2 - 2pnorm(1.327655) \approx 0.1842921$$

Because the p-value is approximately 0.1842921, which is larger than 0.1, we say there is no evidence against H_0: $\theta_0 = 13$.

d) Because the p-value for testing H_0: $\theta_0 = 13$ is greater than or equal to 0.1 if and only if the value $\theta = 13$ is inside the 90% confidence interval, and the p-value is about 0.1842921, which means it is indeed greater than 0.1, we say $\theta = 13$ is inside the 90% confidence interval.

(15 marks) In **<u>Question 5</u>** we will analyse the **Weight in grams variate** from the E-commerce shipping dataset.

Recall in the midterm we assumed that the Weight in grams variate can be modeled as a $G(\mu_i, \sigma_i)$ for i=1,2,3 where 1=Ship, 2=Flight, 3=Road. We wish to test the hypotheses $H_0 : \mu = \mu_0$ and $H_0 : \sigma = \sigma_0$.

a) (3 marks) For the products shipped by **Flight** use the R function t.test to test the hypothesis that the average weight in grams of the products is 3600 grams, i.e $H_0 : \mu_2 = 3600$ for your data. (See Chapter 5, Problem 3 for an example.)

   Be sure to state the observed value of the test statistic

$$T = \frac{|\bar{Y}-\mu_0|}{S/\sqrt{n}}$$

   the corresponding p-value, and your conclusion based on this p-value.

b) (2 marks) Is the value $\mu_2 = 3600$ an element of a 95% confidence interval for μ_2? Explain why or why not using only the p-value determined in (a).

c) (This part can be handwritten – 3 marks) Test the hypothesis $H_0 : \sigma_2 = 1600$ for your data.

   Be sure to state the observed value of the test statistic

$$U = \frac{(n-1)S^2}{\sigma_0^2}$$

   the corresponding p-value, and your conclusion based on this p-value.

d) (2 marks) Is the value $\sigma_2 = 1600$ an element of a 95% confidence interval for σ? Explain why or why not using only the p-value determined in (c).

e) (5 marks) Use the R function t.test to solve for the 95% confidence intervals for $\mu_i, i = 1,2,3$, i.e. the average weight in grams for each mode of shipment. Based on the confidence intervals discuss whether your data suggests a significant difference in average weight between modes of

   shipment. Do these findings agree or disagree with your findings in Question 4j of the midterm? Explain.

a) According to the R script output:

```
data:  weight_flight
t = -0.16143, df = 66, p-value = 0.8722
alternative hypothesis: true mean is not equal to 3600
95 percent confidence interval:
 3166.633 3968.531
sample estimates:
mean of x
 3567.582
```

The observed value of the test statistic is t = -0.16143, the p-value is 0.8722. Since the p-value is larger than 0.1, we say there is no evidence against H_0: True mean is equal to 3600.

b) Because the p-value for testing H_0: $\mu_{2,0} = 3600$ is greater than or equal to 0.05 if and only if the value $\mu_2 = 3600$ is inside the 95% confidence interval, and since the p-value is about 0.8722, which means it is indeed greater than 0.1, we say $\mu_2 = 3600$ is inside the 95% confidence interval for μ.

c) First, we get the sample standard deviation s = 1643.778 and sample size n = 67 using R script:

```
> sd(weight_flight)
[1] 1643.778
> length(weight_flight)
[1] 67
```

Then, following the test statistic, we have

$$u = \frac{(n-1)s^2}{{\sigma_0}^2} = \frac{(67-1) \times 1643.778^2}{1600^2} \approx 69.6611$$

Then, the p-value is calculated by

$$p - value = \min\big(2P(W \le u), 2P(W \ge u)\big) = 0.7108207$$

According to the R script output:

```
> 2*pchisq(69.6611, 66)
[1] 1.289179
> 2*(1-pchisq(69.6611, 66))
[1] 0.7108207
```

Thus, the p-value for testing H_0: $\sigma_{2,0} = 1600$ is 0.7108207. Since the p-value is larger than 0.1, we say there is no evidence against H_0: $\sigma_2 = 1600$.

d) Because the p-value for testing H_0: $\sigma_{2,0} = 1600$ is greater than or equal to 0.05 if and only if the value $\sigma_2 = 1600$ is inside the 95% confidence interval, and since the p-value is about 0.7108207, which means it is indeed greater than 0.1, we say $\sigma_2 = 1600$ is inside the 95% confidence interval for σ.

e) The t-test result for the weight in grams for each mode of shipment is shown in Fig.1.

```
> t.test(Weight_flight, mu = 3600, conf.level = 0.95)

	One Sample t-test

data:  Weight_flight
t = -0.16143, df = 66, p-value = 0.8722
alternative hypothesis: true mean is not equal to 3600
95 percent confidence interval:
 3166.633 3968.531
sample estimates:
mean of x
 3567.582

> # Q5e
> Weight_ship <- dataset_omitted2$Weight_in_gms[which(dataset_omitted2$Mode_of_Shipment == "Ship")]
> t.test(Weight_ship, mu = 3600, conf.level = 0.95)

	One Sample t-test

data:  Weight_ship
t = 0.7084, df = 343, p-value = 0.4792
alternative hypothesis: true mean is not equal to 3600
95 percent confidence interval:
 3487.340 3839.492
sample estimates:
mean of x
 3663.416

> Weight_road <- dataset_omitted2$Weight_in_gms[which(dataset_omitted2$Mode_of_Shipment == "Road")]
> t.test(Weight_road, mu = 3600, conf.level = 0.95)

	One Sample t-test

data:  Weight_road
t = 1.1036, df = 83, p-value = 0.2729
alternative hypothesis: true mean is not equal to 3600
95 percent confidence interval:
 3439.666 4160.072
sample estimates:
mean of x
 3799.869
```

Fig.1 t-test result for the weight in grams for each mode of shipment

According to Fig.1, the average weight in grams for the purchases shipped by flight is [3166.633, 3968.531], the average weight in grams for the purchases shipped by ship is [3487.340, 3839.492], the average weight in grams for the purchases shipped by road is [3439.666, 4160.072].

The plot from midterm-4j is in Fig.2.

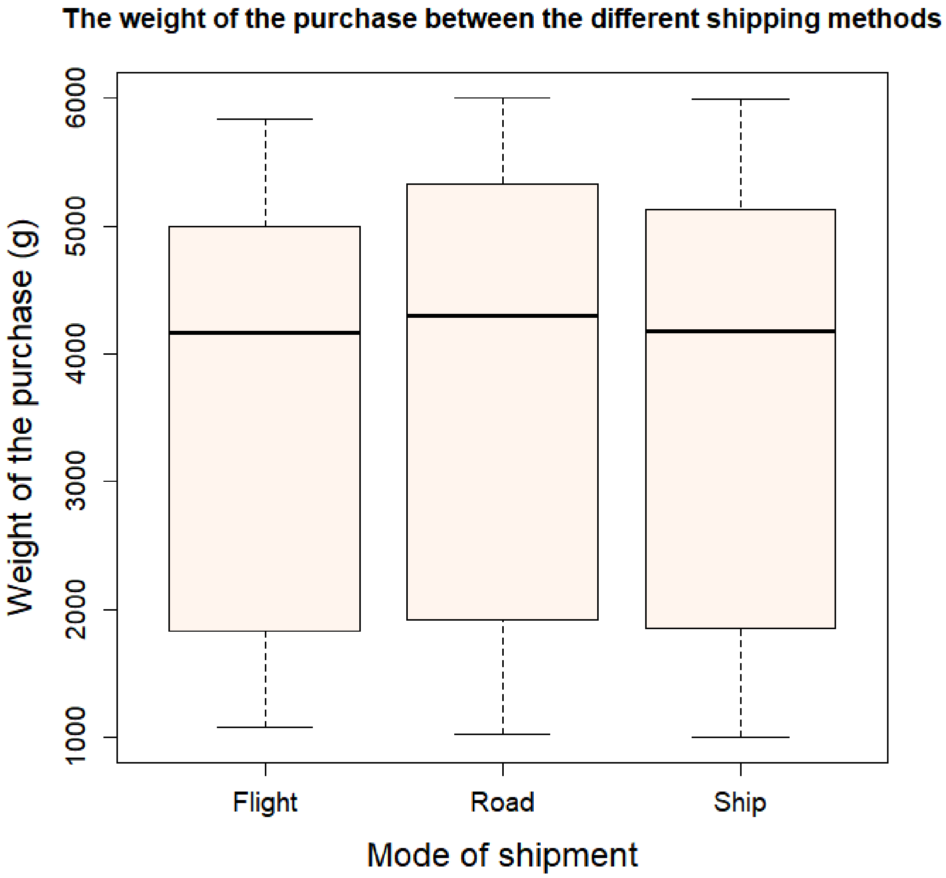

Fig.2 Side-by-side boxplots for the weight of products by mode of shipment

And the findings are:

The symmetry of the three data sets is very close, since they all have more units below the median, and less units above the median, which means they are not symmetric.

From the information above, they all have a longer left tail and shorter right tail since more sample data is concentrated at the region larger than median (i.e. the largest 50% of data has very narrow range).

The ranges of the data sets are very close, but the flight mode has narrower range than other 2 modes slightly.

The weight of the purchases by Ship has a similar median with the weight of the purchases by Flight, which are all smaller than the weight of the purchases by Road.

The data sets do not include any outliers.

Based on the confidence interval, the average weight in grams for purchases shipped by road has the largest confidence interval lower bound and upper bound, which means basically, the population mean weight in grams for purchases shipped by road is expected to be the highest among all the other modes. This corresponds to the finding in midterm that purchases shipped by road has the highest median, since for the data assumed to follow a Gaussian model, the population median is close to the population mean. However, the since the confidence interval for the mean weight in grams for purchases shipped by flight is wider than the confidence interval for the mean weight in grams for purchases shipped by ship, the average weight in grams for purchases shipped by ship is expected to have larger variability, which is against the finding in midterm that the range of weight in grams for purchases shipped by flight is narrower than the range of weight in grams for purchases shipped by ship. This is probably because the sample size of purchases shipped by flight is much smaller than the sample size of purchases shipped by ship, so we are less certain about the population mean weight ingrams for the purchases shipped by flight, which gives the mean weight in grams for purchases shipped by flight a wider confidence interval.

(5 marks) **<u>Question 6</u>** is intended to help you check that you have identified and internalized the important things to learn from this assignment.

Write an approximately 1-2 paragraph reflection (in full sentences) that summarizes how you achieved (or not if you're still not confident with them) the intended learning outcomes by completing this assignment.

I have achieved all the learning outcoimes through the assignment.
First, I tried to calculate the likelihood ratio test statistic and the corrsponding p-value, and drew proper conclusion based on the p-value for several models in question Q2e and Q2f. Then, I observed the asymptotical behavior of likelihooid ratio test for different sizes of sample data, and compared it with the pivotal-quantity test statistic and p-value in Q1. It turns out that although the two methods can be very close if the sample size is large, they still disgree with each other in some cases. Then, using the calculated test statistics for different families or model, I drew proper conclusions based on the p-value for several models in Q1, Q2c, Q2f, Q2g, Q3c, Q4c, Q5a, and Q5c. Basically, when the sample deviation is unknown, and when the sample size is small, t-distribution can be very helpful for testing the population mean; When we want to test the population standard deviation or variance, chi-square distribution can be useful. Then, I visualized and solved for one-sided and two-sided p-value corrsponding to the hypothesis using R and formulas in almost every question. Basically, we conduct one-sided test when the alternative hypothesis is "less than" or "more than" rather than just "different", which is either upper-tail ro lower-tail correspondingly; We conduct two-sided test otherwise. Finally, through the last questions from Q2 to Q5, I discovered the connection between confidence intervals, likelihood intervals, and the hypothesis tests using different families of models. We can know if the null-hypothesis-parameter is in any specific confidence interval by reading its p-value, vice versa.