University of New South Wales

ACTL2131-Probability and Math Statistics

2131 Assignment 2 Report 2025 T1

Xing Chen | z5588124

Task 1

1. Some interesting observations were that the variables: Initial claims, first payments and final payments all had extreme levels of leptokurtosis (see Figure 1) and positive skewness, suggesting that extreme outliers are present in the dataset, possibly largely caused by periods of eligibility for payment and surges in claims during crises like COVID or the GFC.

   An example could be the final payment variable, where a kurtosis of 171 and positive skewness of 11.68 indicates that there are a few cases with massive payout events at the end of unemployment claim processes (see Figure 1).

   The only variable that remotely exhibited "normal" behaviour was that of the Average weekly benefit, which appears to be less affected by shocks from crisis events, which makes sense as this is set by policymakers.

```
##                              Variable      Mean            Variance Skewness
## Initial.Claims         Initial.Claims    245069         18402118247    9.241
## First.Payments         First.Payments     99712          7726298159   17.829
## Weeks.Claimed           Weeks.Claimed   1826786       1379582938800    6.434
## Weeks.Compensated Weeks.Compensated   1693128       1227221643600    7.292
## Avg..Wkly.Benefit Avg..Wkly.Benefit       196               10315    0.207
## Benefits.Paid           Benefits.Paid 345876987 127502141976783952    5.539
## Final.Payments         Final.Payments     43990          1865636317   11.677
##                   Kurtosis
## Initial.Claims       118.1
## First.Payments       376.2
## Weeks.Claimed         56.7
## Weeks.Compensated     66.9
## Avg..Wkly.Benefit      1.5
## Benefits.Paid         46.4
## Final.Payments       171.0
```

Figure 1: Summary Data

2. For first payments (see Figure 2), we see that between 2006 and mid 2007 in a pre-GFC period, initial claims seldom exceed 100,000. The largest increase in first payments can be seen during the peak of the Global Financial Crisis during mid-2008 and 2009 where figures rose above 175,000. This could be interpreted as claimants requiring an increasingly larger benefit package to offset long-term losses due to the mass-layoffs that took place during this period of the GFC.

   Benefits Paid (see Figure 3) followed a similar trend, which had over $100 million paid over the time period of 2009-2010. However, it should be noted that this increase was more gradual, reflecting a slower economic deterioration compared to that of a crisis event like COVID. Overall growth trajectories were less abrupt, depicting a progressive rather than shock-driven response to economic stress.

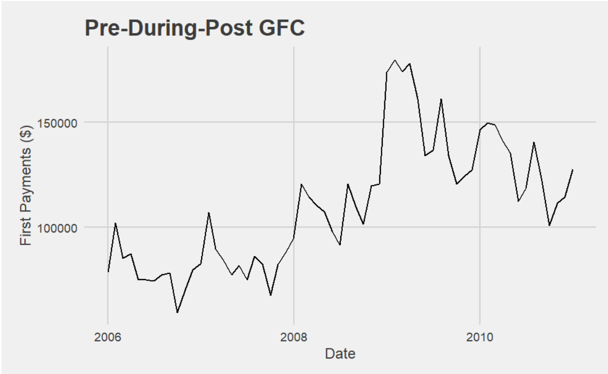

Figure 2: First Payments GFC

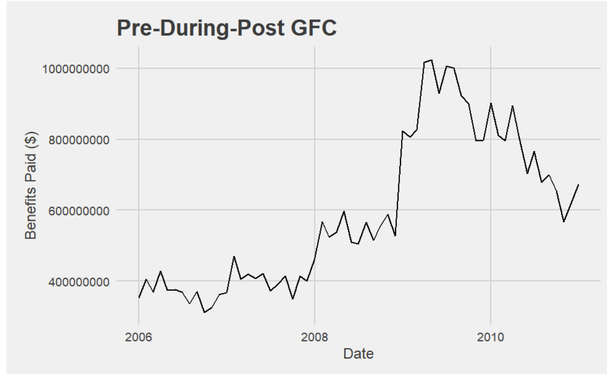

Figure 3: Benefits Paid GFC

Compared to the GFC which experienced several incremental increases in first payments, a sharp spike in first payments is seen in mid-2020 (see Figure 4), corresponding with the start of the COVID pandemic, reflecting mass layoffs, industry shutdowns, surges in unemployment claims due to permanent or temporary closure of businesses.

Benefits paid also portraying similar traits, spiking to $4B in 2020 (see figure 5). However, an interesting observation recorded was the fact that following the sharp increase in the two variables was a rapid decline back down to near pre-COVID levels during 2018 to early 2020. This highlights that the impact of COVID was concentrated and shock based, reflective of the simultaneous layoffs across multiple industries due to nationwide lockdowns.

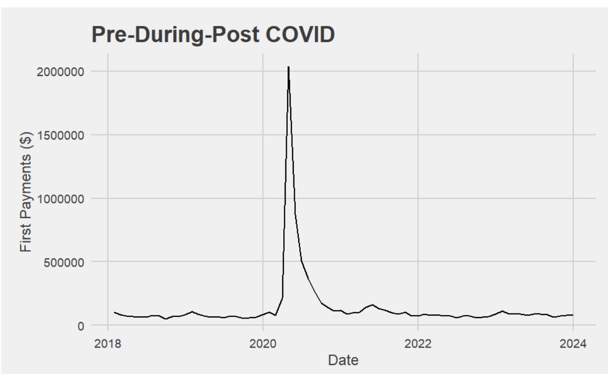

Figure 4: First Payments COVID

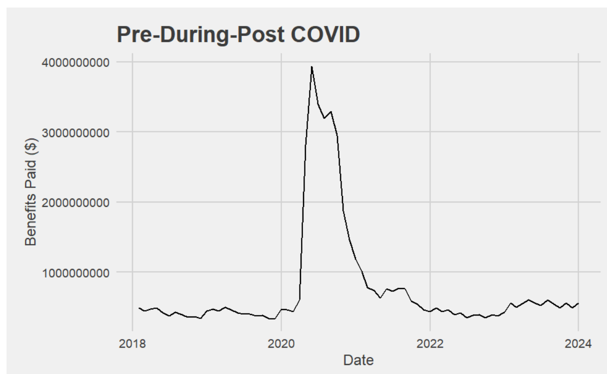

Figure 5: Benefits Paid COVID

3. Chosen Distribution: Log-Normal Distribution
   Reasons:
   (a) Insurance data often have non-normal distributions because they are highly skewed and have heavy tails, Log-Normal distributions can capture positive-skews as implied by the mean of Initial Claims being significantly higher than the median of 225,827 and confirmed with the skewness of 9.24.
   (b) Log-Normal is defined for positive continuous data
   (c) Is frequently utilised to model aggregated insurance claims and forecast the size and freq1uency of claims in industries such as property or casualty insurance. This is in line with the data in the Initial Claims, where most values are moderate but large claim events do occasionally occur.

   We can utilise the MLE for μ and σ^2 (refer to R code for parameter values):

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} \log(x_i), \quad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} (\log(x_i) - \hat{\mu})^2$$

Task 2

1. Due to high skewness and heavy tails (high kurtosis), it is not appropriate to describe the variable "Initial Claims" as normally distributed.
2. The histogram (see Figure 6) depicts a distribution that remains positively skewed even after the standardisation of the initial claims. The distribution of the initial claims indicates a clear misfit when compared to the standard normal distribution in red overlay with extreme outliers such as a Z-score of 15 still occurring. Hence this begins to imply that the variable in question is fundamentally non-normal.

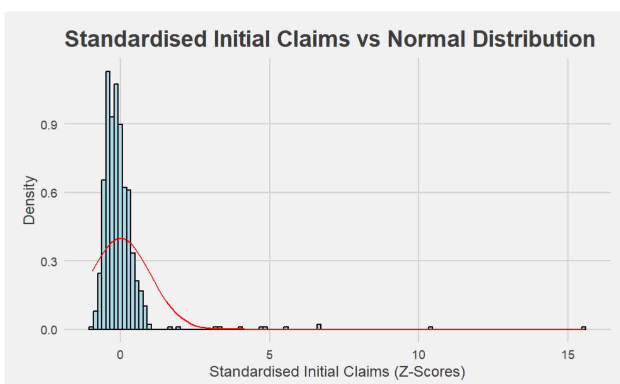

Figure 6: Standardised Initial Claims vs Normal Distribution

3. Additions:
   Utilising Gamma or Log-Normal distributions would serve as a better alternative as the two distributions accounts for high positive skews (accounts for asymmetry of data along with how $x > 0$ and it accounts for extreme outliers).

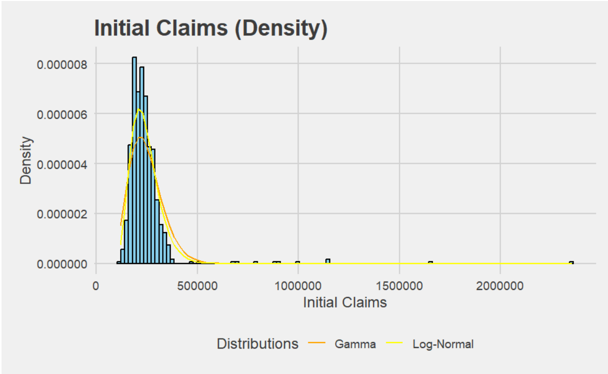

Figure 7: Initial Claims vs Gamma vs Log-Normal

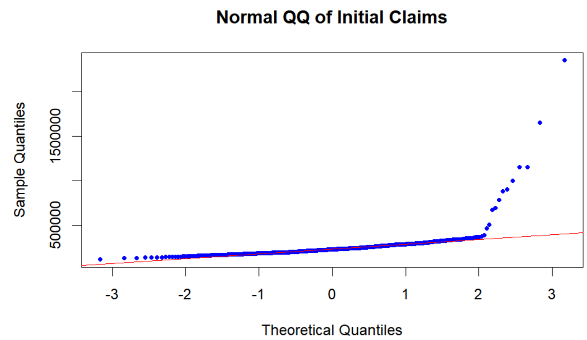

Figure 8: QQ Plot of Initial Claim

Based on the overlay (see Figure 7) of both the Gamma and Log-Normal distributions against the density graph for the 'Initial Claims' variable, the Log-Normal appears to have a better fit. This could be due to the Log-Normal having heavier tails along with a higher positive skew.

The QQ plot (see Figure 8) depicts points mostly hugging the red line but bends are apparent in the left tail of the QQ Plot. Additionally, on the right tail of the plot, points rise far above the red line which indicates extreme outliers were present possibly due to crises like COVID or the GFC.

The QQ plot implies that a normal distribution would underestimate extreme claim spikes and confirms that Normal distributions would not describe 'Initial Claims' very well, which challenges the manager's conjecture that "any variable can be modelled with a normal distribution" especially when stress tested under crisis events.

Task 3

1. The Chi-Squared Goodness of Fit test for normality provided a p-value way below the 5% level of 2.2e-16. Hence there is enough statistical evidence at the 5% level of significance to suggest that the variable Initial Claims does not follow a normal distribution. This result coincides with earlier conclusions of the actual distribution being similar to that of a Gamma or Log-Normal distribution rather than a Normal distribution.

   The second Chi Squared goodness of fit test yields the same p-value when attempting to fit the standardised data set, suggesting that the standardised data set does not fit into an alternative distribution like that of a Student-t distribution. This makes sense because a t-distribution has similar characteristics to that of a normal distribution.

2. The Chi-Squared Goodness of Fit test for normality revealed a p-value of 2.2e-16, which at a 5% significance level, enables us again to reject the notion that the difference of the two logged variables are normally distributed.

   The Wilcoxon signed-rank test is a useful test for a dataset that is not normally distributed and compares the 2 related samples using non-parametric statistics. We assume dependent samples (same subjects under two conditions, i.e. weeks claimed and compensated), independence within pairs (selection or result of one pair does not interfere with another pair's), continuity, ordinal measurement (differences between pairs can be ascertained, i.e. greater than, equal to, less than) and scale compatibility (measurements of the samples are on a similar scale).

The Wilcoxon signed-rank test also yielded a p-value $< 2.2e\text{-}16$, leading us to reject the null hypothesis at the 5% level of significance. At this significance level, there is sufficient evidence to suggest that the number of weeks claimed and weeks compensated differ on average.

This is consistent with intuition, with most weeks claimed likely being either paid out or are considered initial filing cases. Whereas for weeks compensated there can be cases of an eligibility criteria that need to be met when considering the claim and compensation processes that filters some of the claims, resulting in differences between the two.

3. The one-sample t-test was used to test whether the average value of $\log(Benefits.Paid)$ exceeded its 60th percentile of 19.72. The resulting critical value of 19.36 meant that we fail to reject the null hypothesis at a 5% level of significance and we can infer that the mean of $\log(Benefitspaid)$ is not greater than its own 60th percentile.

   A important assumption worth mentioning is that normality is assumed for the underlying data of the one-sample t-test. In terms of skewness and kurtosis, skewness was -0.10, which is very close to a normal distribution's skewness of 0. Additionally, the kurtosis of 3.075, which only has an excess kurtosis of 0.075, which is very close to the kurtosis level of a normal distribution. Hence the sample t-test is an appropriate test for the hypothesis test.

4. Throughout this assignment, we assumed that the data analysed were composed of i.i.d. random variables. This is not a realistic assumption as crises can strongly influence bias in datasets where observations could be correlated which violates the assumption of independence. Furthermore, large macroeconomic events like the GFC and COVID introduce high variability of variables that occur due to factors such as market/economic sentiments (benefits paid and initial claims) which breaks the assumption of identical data. Task 3 assumes that each observation is not influenced by previous observations, which leads to a risk where variance is underestimated as events could very well be correlated. Observations would portray artificially small and unreliable p-values in our t-tests and Wilcoxon tests because of this assumption which increases the risk of reaching insignificant conclusions because of the increased probability for Type I error. Properties such as the mean and variance were assumed constant over time but factors like inflation, benefit caps, and job market shocks make these parameters time-varying, meaning that they could change through time and crisis events, hence potentially losing its inferential meaning. This is seen in task 3.3, where the 60th percentile may reflect one part of the distribution (i.e. post-COVID stability), while the mean is heavily affected by earlier shocks which leads to inaccurate interpretations of the benefits paid. I.i.d. simplifies and justifies the use traditional statistical inferential techniques which in practice may not be reliable for meaningful data analysis.