ACTL 2131/5101 - Assignment 2 (practical)

Submission deadline:

Friday, 25th April, 5 pm sharp via Turnitin on Moodle

This assessment activity aims to reinforce your learning by providing you with an opportunity to apply the concepts you have learned in the course to practical problems. They will help you solidify the relevance of the course to the real world.

Instructions:

This assignment is worth 10% of the course mark.

1. Please upload a maximum of 5 pages report which provides answers to all questions.
2. Submit your R code.
3. Please pay attention to plagiarism [Turnitin will pick it up].

Background/Data preparation

You are an actuarial analyst working for the insurance provider. Your manager has requested you to download, prepare and analyse data set which contains of observations for unemployment insurance (UI) obtained from the United States Department of Labour available on

`https://oui.doleta.gov/unemploy/claimssum.asp`

Data should be downloaded on a monthly basis for a particular state of your choice (it could be easier to just copy-paste data into the Excel/csv file). The sample should cover from January 1, 1971 (or the earliest available date) until the most recent date available. The data to download should consist of 7 variables for the state of your choice. These variables include:

- Initial claims
- First payments
- Weeks claimed
- Weeks compensated
- Average weekly benefit
- Benefits paid
- Final payments

There should be no missing data. The definition of all variable is provided in the Appendix.

For the purpose of this assignment, assume that each monthly observation are independent and identically distributed (i.i.d.) random variables.

Task 1

Perform explanatory data analysis, which should include:

1. Presenting summary statistics (including mean, variance, skewness, kurtosis) on the 7 variables and comment on the findings you find interesting.

2. Using two variables of your choice identify trends by plotting data over various times frames (think of pre-during-post Global Financial Crisis, pre-during-post COVID). Discuss your findings.

3. As part of an exploratory analysis of the distribution of Initial Claims, select any distribution covered in the course that you believe may be a reasonable candidate. You are not expected to choose the "correct" distribution: the goal is to critically analyse your choice based on the characteristics of the data. Justify why the distribution you selected could be appropriate (or not), and write down the Maximum Likelihood Estimator (MLE) for its parameters. Calculate MLE(s) using the data.

   [*Note: You do not need to perform any statistical goodness-of-fit tests for this task.*]

Task 2

Your manager suggests that *any* variable can be well described by a normal distribution. Remember, if a random variable $Y \sim N(\mu, \sigma^2)$, then we know $\frac{Y-\mu}{\sigma} \sim N(0,1)$. This process is called standardisation. Doing this creates a random variable that has mean 0 and variance 1.

You are asked to critically assess the manager's conjecture. Select an appropriate variable from your dataset, standardise it, and investigate whether it can reasonably be described by a normal distribution. When responding, please address the following:

1. Briefly comment on whether a normal distribution seems appropriate for your standardised variable of choice from the UI dataset. [Brief discussion would be sufficient to address this question.]

2. Present a histogram for the standardised variable of your choice. In the same graph, draw a density of the standard normal distribution. Comment on your results.

3. Use some additional numerical and graphical illustrations (in addition to the histogram from point two) to support your conclusions. Can you think of any other alternative distribution that could be used instead of a normal distribution and why? [Brief discussion would be sufficient to address the last question.]

Task 3

1. Select the same variable that was used in Task 2, standardise it and use an appropriate *statistical test* at a 5% significance level to investigate whether your data follows Normal distribution. Is your conclusion inline with findings documented in Task 2? Can you now investigate using a formal test if your data follows a Student-t or any other distribution (rather than normal)?

2. In addition, the manager is also interested in whether the number of weeks claimed and compensated are equal on average. Assist the manager in completing this task by comparing the means for the $log(Weeksclaimed)$ and $log(Weekscompensated)$. Comment on the choice of the test and the results. Use a 5% significance level. Is this result consistent with your intuition? [Note: You might have to first check the assumptions underlying the chosen test.]

3. The manager is interested in the magnitude of the Benefits paid, and has requested you to test the hypothesis that the average value of $log(Benefitspaid)$ is greater than the value corresponding to the 60% quantile of this variable. You are asked to comment on the choice of the test and the results. Use a 5% significance level. Comment on the result of this test in terms of skewness and kurtosis of the data.

4. In this assignment we allowed you to assume the data is i.i.d. random variables. Discuss the implications of this assumption in the context of your analysis. Is this assumption reasonable? What are the implications of all the previous tests in Task 3 if the data is not i.i.d.?

Plagiarism awareness

Please ensure that your submission is your own work. Discussion of assignments is allowed, but submitted material must be independently created (both, the code and the report), which should be developed on your own without copying from peers. Significant similarities in code, even with altered variable names or comments, will be flagged by Turnitin as plagiarism. Familiarise yourself with what constitutes plagiarism, as detection is highly likely and not diminished by collective work or varied assignment markers.

Appendix

These are the definitions of variables above (generated by ChatGPT). Please be aware of the limitations of AI, e.g., the last paragraph is just repeated for each definition, even when not prompted or asked about how it's impacted. It demonstrates at this current stage AI still needs our input to make sense of it.

- **Initial Claims** refers to the number of new unemployment insurance applications filed in a given week by individuals who have lost their jobs or had their hours reduced. These claims are an indicator of the state of the labor market and the economy, as a sudden increase in initial claims can signal a weakening in employment conditions. In the United States, the Initial Claims data is released on a weekly basis by the Department of Labor and is closely watched by economists and financial market participants. The data provides valuable insight into the current economic climate and can be used to predict future trends in the labor market and the economy as a whole.

- **First payments** refer to the first disbursement of unemployment insurance benefits to individuals who have filed a claim and have been approved for benefits. In the United States, the process of receiving first payments can vary by state, but typically involves a waiting period of one to two weeks after the initial claim is filed, during which time the claim is processed and eligibility is determined.

  Once an individual has been approved for benefits, they will typically receive their first payment within a few days to a week, depending on the state. The amount of the first payment and subsequent weekly payments will depend on the individual's prior earnings and the specific unemployment insurance program in the state where they reside.

  It's important to note that the amount and timing of first payments may be impacted by factors such as the backlog of claims, processing delays, and changes to eligibility requirements due to the ongoing COVID-19 pandemic.

- **Weeks claimed** refer to the number of weeks for which an individual has filed for and received unemployment insurance benefits. In the United States, the maximum number of weeks for which an individual can receive benefits varies by state, but is typically 26 weeks. However, during periods of high unemployment, the federal government may provide additional funding to states to extend the maximum number of weeks that individuals can receive benefits.

  The number of weeks claimed is important because it indicates how long an individual has been unemployed and how long they have been receiving unemployment insurance benefits. It can also provide insight into the current state of the labor market and the economy, as a sudden increase in the number of weeks claimed can signal a weakening in employment conditions and an increase in long-term unemployment.

  It's important to note that the number of weeks claimed can be impacted by factors such as changes to eligibility requirements, the backlog of claims, and processing delays, particularly during times of high unemployment, such as the ongoing COVID-19 pandemic.

- **Weeks compensated** refer to the number of weeks for which an individual has received unemployment insurance benefits after filing a claim and being approved for benefits. This number may be different from the number of weeks claimed, as some individuals may have their claims

denied or may exhaust their benefits before the maximum number of weeks for which they are eligible.

In the United States, the amount and duration of unemployment insurance benefits are determined by the state where the individual resides and are based on factors such as the individual's prior earnings and the specific unemployment insurance program in that state.

The number of weeks compensated provides valuable information about the state of the labor market and the economy, as a sudden increase in the number of weeks compensated can signal a weakening in employment conditions and an increase in long-term unemployment.

It's important to note that the number of weeks compensated can be impacted by factors such as changes to eligibility requirements, the backlog of claims, and processing delays, particularly during times of high unemployment, such as the ongoing COVID-19 pandemic.

- **Average weekly benefits** refer to the average amount of unemployment insurance benefits paid out to individuals on a weekly basis. This statistic provides valuable information about the average financial support being provided to individuals who have lost their jobs or had their hours reduced, and can provide insight into the state of the labor market and the economy.

  In the United States, the average weekly benefit is determined by dividing the total amount of benefits paid by the number of individuals receiving benefits and the number of weeks for which they have received benefits. The average weekly benefit is intended to provide a rough estimate of the typical amount of benefits being received by individuals and can be used to compare the financial support provided by different unemployment insurance programs and states.

  The average weekly benefit can also be used to measure the effectiveness of government policies aimed at supporting individuals who are unemployed and to track the overall impact of unemployment on the economy.

  It's important to note that the average weekly benefit can be impacted by factors such as changes to eligibility requirements, the backlog of claims, and processing delays, particularly during times of high unemployment, such as the ongoing COVID-19 pandemic.

- **Benefits paid** refer to the total amount of unemployment insurance benefits that have been disbursed to individuals who have filed claims and been approved for benefits. This amount is an indicator of the financial support being provided to individuals who have lost their jobs or had their hours reduced, and can provide valuable insight into the state of the labor market and the economy.

  In the United States, the amount of benefits paid is determined by the state where the individual resides and is based on factors such as the individual's prior earnings and the specific unemployment insurance program in that state. The benefits paid are intended to provide a temporary source of income for individuals who are unemployed and are designed to help them meet their basic needs while they search for new employment.

  The amount of benefits paid can also be used to measure the effectiveness of government policies aimed at supporting individuals who are unemployed and to track the overall impact of unemployment on the economy.

  It's important to note that the amount of benefits paid can be impacted by factors such as changes to eligibility requirements, the backlog of claims, and processing delays, particularly during times of high unemployment, such as the ongoing COVID-19 pandemic.

- **Final payments** refer to the last disbursement of unemployment insurance benefits to an individual who has received benefits for the maximum number of weeks for which they are eligible. In the United States, the maximum number of weeks for which an individual can receive benefits varies by state, but is typically 26 weeks.

  Final payments mark the end of an individual's eligibility for unemployment insurance benefits and signal that they have either found new employment or exhausted their benefits. The amount of the final payment will depend on the individual's prior earnings and the specific unemployment insurance program in the state where they reside.

  It's important to note that the timing and amount of final payments may be impacted by factors such as changes to eligibility requirements, the backlog of claims, and processing delays, particularly during times of high unemployment, such as the ongoing COVID-19 pandemic.