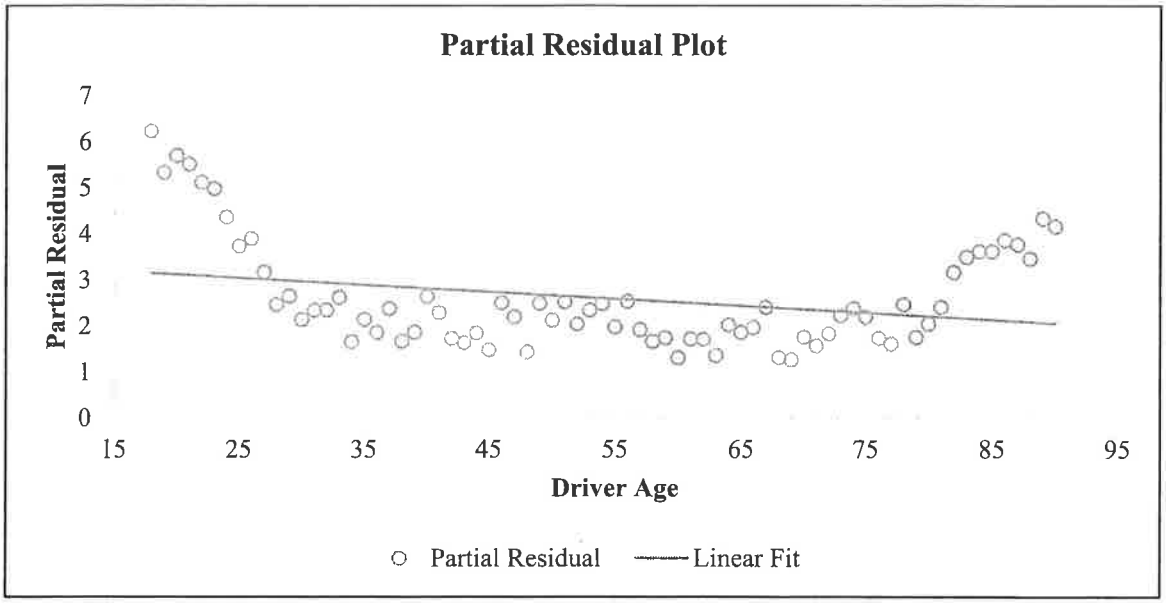
5. (1.5 points)

An actuary is analyzing a partial residual plot of the driver age variable, which is shown below:



a. (1 point)

Adding polynomial terms is one approach to address the non-linearity in the driver age variable.

Briefly describe two other alternative approaches and how they can be used to improve the fit of the driver age variable shown above.

b. (0.5 point)

Briefly describe a downside to each of the two alternative approaches recommended in part a. above.

CONTINUED ON NEXT PAGE

SAMPLE ANSWERS AND EXAMINER'S REPORT

QUESTION 5	
TOTAL POINT VALUE: 1.5	**LEARNING OBJECTIVE(S): A3**

SAMPLE ANSWERS

Part a: 1 point

Sample 1

- We could bin the variable of drive age, create categorical variables. For this example, binning 16-25, 26-75, 76-96 together would cause the residuals to be more in line with the linear fit.
- We could add a piecewise linear function with hinge points at 25 & 75. This would transform the variables to fit the linear fit.

Sample 2

- Binning the age:
  - bin every 5 years into a group
  - can help differentiate the difference in residuals by age
- Piecewise terms
  - Put into 3 pieces of linear curves
  - "15-25", "25-80", "80+"
  - Can track the different slopes of residuals

Sample 3

- Binning: we can bucket the variables into groups, turn the continuous variable into categorical variables. Each group has a predicted rate & residuals would/should improve
  - Ex. Group age 15-25, 26-80, and 80+
- Hinge functions: Allows us to fit a custom shape to the data. For example, we could have a slope from 15-25, then it would hinge at 25 until 80, again at 80 and above.
-

Part b: 0.5 point

Sample 1

- Binning: this adds variables and thus increase the degrees of freedom
- Piecewise linear functions: the hinge points need to be manually selected

Sample 2

- Does not account for variations within the bins
- Cut-off point selection is judgmental

Sample 3

- With binning variables, we may not have enough credibility in each of the buckets to have stable predictions
- For hinge function, these would have to be assigned by visual inspection

Sample 4

- For binning, this could lead to non-intuitive results, such as reversals.
- *(For piecewise)* Increases the degrees of freedom.

SAMPLE ANSWERS AND EXAMINER'S REPORT

EXAMINER'S REPORT

Candidates were expected to understand ways to address non-linearities within a GLM, and how to apply them to the example given, as well as its related downsides.

Part a

Candidates were expected to understand ways to address non-linearity within a GLM, and how to apply them based on the given plot.

Common mistakes included:

- Listing ways to address the non-linearity but not specifying how they can be applied to the specific example
- Not explaining how the approaches can be applied to the given partial residual plot
- Listing "adding interaction terms", "log transformations", or "adding offset terms" as the approach
- For the hinge function, the candidate using an incorrect function form by including log transformation, for example, max (log(age) - 25, 0)

Part b

Candidates were expected to understand the downsides of the approaches in part a.

Common mistakes included:

- Being too vague or too broad, for example:
    - "Binning will lose information."
    - For piecewise, "it is difficult to select break point."