2. (2.75 points)

An actuary is modeling claim frequency for a portfolio with the following distribution of exposures.

| | Vehicle Class | | | |
|---|---|---|---|---|
| Territory | Car | Van | Truck | Other |
| A | 10,000 | 2,000 | 0 | 0 |
| B | 2,000 | 5,000 | 0 | 0 |
| C | 0 | 0 | 5,000 | 0 |
| D | 30 | 0 | 0 | 3,000 |

The actuary proposes a generalized linear model (GLM) with the following parameterization.

| Territory | |
|---|---|
| Factor level | Parameter |
| A | $\beta_1$ |
| B | $\beta_2$ |
| C | $\beta_3$ |
| D | $\beta_4$ |

| Vehicle Class | |
|---|---|
| Factor level | Parameter |
| Car | $\beta_5$ |
| Van | $\beta_6$ |
| Truck | $\beta_7$ |
| Other | $\beta_8$ |

a. (1.0 points)

Briefly discuss how intrinsic and extrinsic aliasing are present in this analysis using examples from the data. For each type of aliasing briefly explain the potential impact on the results.

b. (0.5 point)

Provide one example of near aliasing in this analysis and briefly describe any potential impact on the modeling results.

c. (1.25 points)

Propose an alternative GLM approach to avoid extrinsic, intrinsic, and near aliasing. Describe how many covariates would be required.

CONTINUED ON NEXT PAGE

**QUESTION 2**

**Total Point Value: 2.75** **Learning Objective: A3**

**Sample Answers**

**Part a:** 1.00 points

*Sample 1*

Intrinsic aliasing is occurring because there are covariates for each level of each variable used. By definition Beta 4 = 1 – Beta 1 – Beta 2 – Beta 3 and Beta 8 = 1 – Beta 5 – Beta 6 – Beta 7.

Extrinsic aliasing is occurring for vehicle class truck and territory C because all observations in C are trucks and vice versa. Based on nature of data, Beta 3 = Beta 7.

These can lead to convergence issues or confusing results. Alternatively, modern GLM software will usually automatically correct for these.

**Part b:** 0.50 points

*Sample 1*

Near aliasing occurs when there is strong correlation (but not perfect) between covariates. In our case, Territory D is highly correlated with Other type vehicle class. This will cause convergence issues.

**Part c:** 1.25 points

*Sample 1*

- First I would eliminate all observations with Territory = D and Vehicle Class = Car.
- We can eliminate Truck (Beta 7) and Other (Beta 8) because they are aliased with Territory C and D, respectively.
- From there I would eliminate Vehicle Class = Van so the model is uniquely defined.
- So we have Beta 1, Beta 2, Beta 3, Beta 4, and Beta 5 for a total of 5 covariates.

*Sample 2*

- We can have a base term for Territory A and class Car (eliminates intrinsic aliasing)
- We can keep Territory C and eliminate class Truck (eliminates extrinsic aliasing)
- We can eliminate the 30 cars in Territory D (eliminates near aliasing) and get rid of Other class

- Covariates are Beta 0 (A and car), Beta 2 (Territory B), Beta 3 (Territory C), Beta 4 (Territory D), Beta 6 (Class Van).
- 5 covariates needed.

**Examiners Report**

**Part a:**

Candidates did reasonably well on this part of the problem, generally receiving a majority of the possible points. The candidate was expected to identify that intrinsic aliasing was present as the fourth territory (or vehicle class) could be expressed as a linear combination of the other territories (or classes). Extrinsic aliasing is present because Territory 'C' and vehicle class 'Trucks' are perfectly correlated – that is, all trucks are in territory C and territory C is comprised of only trucks.

Both types of aliasing can result in convergence issues as the model will not be uniquely defined. Another acceptable response is that for both types of aliasing, modern GLM software will make the necessary corrections.

When only partial credit was given, common mistakes included:

- Simply stating that intrinsic aliasing was caused because there were 8 parameters without identifying the linear relationships between them.
- Failing to provide examples of intrinsic and extrinsic aliasing using the data provided.
- Failing to describe the impact of the intrinsic or extrinsic aliasing.

**Part b:**

In general candidates did well on this part of the problem, with a majority of candidates receiving full credit. The candidate was expected to highlight that Territory D and the 'Other' vehicle class were nearly perfectly correlated and that this near aliasing would create convergence issues, unstable parameter estimates, or confusing results. The most common mistake in this part of the problem was failing to provide both the example and impact as requested in the problem.

**Part c:**

This subpart proved to be the most difficult as few candidates received full credit. This part was challenging in that multiple instances of aliasing needed to be addressed to receive full credit. To receive full credit the candidate needed to:

- Remove either Beta 3 or Beta 7 to address the extrinsic aliasing between Territory 'C' and vehicle class 'Truck'
- Address the rogue data causing the near aliasing (the 30 cars in territory D) by reclassifying or removing the observations
- Remove either Beta 4 or Beta 8 as Territory 'D' and vehicle class 'Other' are now extrinsically aliased after removing the 30 cars in Territory 'D'

- To address the intrinsic aliasing, the candidate needed to remove one additional parameter or remove two parameters (one territory and one vehicle type) and introduce an intercept.
- This will result in a total of 5 covariates.

Common mistakes included:

- Failing to address the rogue observations causing the near aliasing
- Including too many parameters in the intercept (e.g., including all of Territory 'C', Territory 'D', and classes 'Truck' and 'Other' in the intercept)