6. (2.5 points)

An actuary has constructed a three-variable Tweedie GLM with a log-link function to estimate loss ratios for commercial property new business. The actuary wants to create a second model for renewal business that will include all of the variables from the new business model, plus a variable for the prior year claim count. The actuary requires that the coefficients of the variables (Average Building Age, log(Manual Premium), and Location Count) are consistent between the new and renewal models. The fitted new business model parameters are as follows:

Variable	Name	Estimate
	intercept	0.910
Average Building Age (Years)	age	0.013
log(Manual Premium)	logprem	-0.187
Location Count	loccnt	0.062

a. (0.75 point)

Calculate the modeled loss ratio for a new business policy with a manual premium of $25,000, an average building age of four years, and having eight locations.

b. (0.75 point)

Briefly describe how to produce the renewal business model, and specify the resulting equation for the renewal business modeled loss ratio.

c. (1 point)

Identify and briefly describe two techniques that the actuary can use to assess the stability of the new variable in the renewal business model.

CONTINUED ON NEXT PAGE

EXAM 8 FALL 2016 SAMPLE ANSWERS AND EXAMINER'S REPORT

QUESTION: 6	
TOTAL POINT VALUE: 2.5	**LEARNING OBJECTIVE(S): A3, A4**

SAMPLE ANSWERS

Part a: 0.75 point

Sample 1 (using natural log)

$\ln(\mu) = 0.910 + 4 * (0.013) + \ln(25{,}000) * (-0.187) + 8 * 0.062 = -0.43568$

$\mu = e^{-.43567} = 64.7\%$

Sample 2 (using log base 10)

$\ln(\mu) = 0.910 + 4 * (0.013) + \log(25{,}000) * (-0.187) + 8 * 0.062 = .63559$

$\mu = e^{.63559} = 188.8\%$

Part b: 0.75 point

Sample 1

The result from the new business model can be added into the renewal model as an offset. The resulting equation is:

$g(\mu) = \ln(\mu) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4$

β_0 = Intercept, to be re-estimated

β_1= parameter for average building age = 0.013

β_2 = parameter for log(manual premium) = -0.187

β_3 = parameter for location count = 0.062

β_4 = parameter for prior year claim count, to be estimated

Part c: 1 point

- **Cross-Validation** – Split the data into k parts and run the model on the (k-1) parts, then validate the result on the last part. Compare how similar the estimates are from the k iterations to assess variable stability.
- **Bootstrapping** – Create multiple datasets from the initial train dataset by sampling with replacement. Run the model (with same specs) of each sampled dataset. Assess stability of estimates of coefficients by comparing the results from each run. You can compute standard errors, means and confidence intervals for the variable.
- **Cook's Distance** – Sort the observations based on their Cook's Distance value (higher distance = more influence on the model.) Remove some of the most influential observations and re-run the model on this new set of data to see the effect on estimated parameters.
- **Validation on Holdout Dataset** – Split the data into train and test. Run the model on the train and validate on test dataset by comparing variable fit. The models should produce similar results.
- **Time- Consistency** – Review the results of model by accident year, to assess stability over time.

EXAMINER'S REPORT

Candidates were expected to know the components of a GLM formula, GLM technical foundation and model refinement to get full credit for this question.

Generally, candidates understood the components of a GLM formula, but struggled with the technical foundation of how to offset a model and with the difference between variable significance and variable stability.

Part a

The candidate was expected to use the components of the GLM formula to produce the modeled loss ratio. Both the log and ln of the manual premium were accepted as correct answers.

Common mistakes include:

- Using the natural log of average building age or location count.
- Not converting to a loss ratio.
- Using the incorrect link function.

Part b

Candidates were expected to produce the renewal business loss ratio, while keeping the coefficients from the new business model the same.

It was important to recognize that the new business model was also modeling a loss ratio, not whether a policy would renew. Candidates that gave the correct formula but with no description of offset or how to apply it were not given full credit.

Common mistakes include:

- Describing how to model a new/renew indicator or probability of renewal (not the loss ratio).
- Not keeping the coefficients of average building age, ln(manual premium) and location count the same between models.
- Not recognizing that the intercept is different between models.

Part c

Candidates were expected to assess the stability of the new variable (parameter estimate).

Common mistakes include:

- Giving a definition of the technique without fully describing how the technique can be used to assess variable stability.
- Giving answers that described how to assess variable significance with no tie-in to the concept of variable stability or giving answer on penalized measures of fit (e.g. AIC, BIC). Answers that described how to assess variable significance that did tie into the concept of variable stability were given full credit.