Study Guide on Testing the Assumptions of Age-to-Age Factors for the Casualty Actuarial Society (CAS) Exam 7 and Society of Actuaries (SOA) Exam GIADV: Advanced Topics in General Insurance

(Based on Gary Venter's Paper, "Testing the Assumptions of Age-to-Age Factors")

G. Stolyarov II, ASA, ACAS, MAAA, CPCU, ARe, ARC, API, AIS, AIE, AIAF

Study Guide Created in Spring 2011, Updated February 1, 2015

Source: Venter, G.G., "Testing the Assumptions of Age-to-Age Factors," *PCAS* LXXXV, 1998, pp. 807-847.

This is an open-source study guide and may be revised pursuant to suggestions.

***Venter's Notation* (Venter, p. 808)**

c(w,d): cumulative loss from accident year w as of age d
c(w,∞): total loss from accident year w when end of triangle reached
q(w,d): incremental loss for accident year w from (d - 1) to d
f(d): factor applied to c(w,d) to estimate q(w, d+1)
F(d): factor applied to c(w,d) to estimate c(w,∞)

Problem S7-TAAF-1. State Mack's three assumptions pertaining to the chain-ladder method in terms of Venter's notation (Venter, p. 809).

Solution S7-TAAF-1.
Assumption 1: **E[q(w, d+1) │data to (w+d)] = f(d)*c(w,d)**. (Expected-value assumption)
Assumption 2: **Unless v = w, c(w, d) and c(v, g) are independent for all v, w, d, and g.** (Assumption of independence of accident years)
Assumption 3: **Var[q(w, d+1) │data to (w+d)] = a[d, c(w, d)]**. (Proportionality assumption for the variance). In this case, the function a(_, _) can be any function that does not vary by accident year.

Problem S7-TAAF-2. Under Mack's three assumptions, what sort of estimator of future emergence does the chain-ladder method produce? (Venter, p. 810)

Solution S7-TAAF-2. Under Mack's three assumptions, the chain-ladder method produces the **minimum-variance unbiased linear estimator** of future emergence.

Problem S7-TAAF-3. What did Mack assume about the function a(_, _) as used by Venter in expressing Mack's third assumption pertaining to the chain-ladder method? (Venter, p. 810)

Solution S7-TAAF-3. Mack assumed that **a[d, c(w, d)] = k(d)*c(w, d)**, where k(d) is a constant for a given development year d. That is, the variance is proportional to the previous cumulative loss.

Problem S7-TAAF-4. What is the expression generally minimized in order to find the minimum-variance unbiased estimator of f(d)? (Venter, p. 810)

Solution S7-TAAF-4. The expression generally minimized in order to find the minimum-variance unbiased estimator is $_w\sum$**[f(d)*c(w,d) - q(w, d+1)]^2*k(d)/a[d, c(w, d)]**.

Problem S7-TAAF-5. Venter (pp. 812-813) lists six testable implications of the hypothesis that the chain-ladder method is appropriate in a given situation. What are these implications?

Solution S7-TAAF-5.
1. Significance of the factor f(d)
2. Superiority of factor assumption to alternative emergence patterns
3. Linearity of the model (Look at residuals as a function of c(w, d).)
4. Stability of the factor (Look at residuals as a function of time.)
5. No correlation among columns
6. No particularly high or low diagonals

Problem S7-TAAF-6. Venter (p. 812) lists three possible alternative emergence patterns to Mack's first assumption that E[q(w, d+1) | data to (w+d)] = f(d)*c(w,d). Describe these three alternative assumptions mathematically.

Solution S7-TAAF-6.
Alternative assumption 1: Linearity with constant: **E[q(w, d+1) | data to (w+d)] = f(d)*c(w,d) + g(d)**, where g(d) is a constant.
Alternative assumption 2: Factor times parameter: **E[q(w, d+1) | data to (w+d)] = f(d)*h(w)**, where h(w) does not depend on already emerged experience.
Alternative assumption 3: Including calendar-year effect:**E[q(w, d+1) | data to (w+d)] = f(d)*h(w)*g(w+d)**.

Problem S7-TAAF-7.
(a) Fill in the blanks (Venter, p. 813): A usual test is that the absolute value of a factor is required to be at least ____ (how many?) times its standard deviation in order for the factor to be regarded as significantly different from zero. If a development triangle fails this test, then the chain-ladder method ______ (is or is not?) optimal for those triangles.
(b) Fill in the blanks (Venter, p. 813): For a Normal distribution, the requirement in (a) provides that there is only a ___% probability of getting a factor of this absolute value or greater if the true probability is zero. Many analysts have sufficient comfort with a factor with absolute value ___ (how many?) times its standard deviation, which could happen about ___% of the time by chance alone.
(c) Fill in the blank (Venter, p. 814): If a factor is not significant under a Normal distribution, then it is probably _____ (more or less?) significant under other distributions.

Solution S7-TAAF-7.
(a) A usual test is that the absolute value of a factor is required to be at least **2** times its standard deviation in order for the factor to be regarded as significantly different from zero. If a development triangle fails this test, then the chain-ladder method **is not** optimal for those triangles.
(b) For a Normal distribution, the requirement in (a) provides that there is only a **4.5**% probability of getting a factor of this absolute value or greater if the true probability is zero. Many analysts have

sufficient comfort with a factor with absolute value **1.65** times its standard deviation, which could happen about **10**% of the time by chance alone.
(c) If a factor is not significant under a Normal distribution, then it is probably **less** significant under other distributions.

Problem S7-TAAF-8.
(a) Mathematically describe three adjustments suggested by Venter (pp. 814-815) to the sum of squared errors (SSE), using the notation SSE = sum of squared errors, n = number of observations, and p = number of parameters.
(b) Compute the relevant measures using each of those adjustments for a model where the SSE is 5520, the number of observations is 43, and the number of parameters is 5.

Solution S7-TAAF-8.
(a) Adjustment 1: $SSE/(n-p)^2$.
Adjustment 2: Akaike Information Criterion = AIC ≈ $SSE*e^{2p/n}$.
Adjustment 3: Bayesian Information Criterion = BIC ≈ $SSE*n^{p/n}$.
(b) Adjustment 1: $SSE/(n-p)^2 = 5520/(43-5)^2$ = **3.822714681**.
Adjustment 2: AIC ≈ $SSE*e^{2p/n}$ = 5520*exp(2*5/43) = **6965.26741**.
Adjustment 3: BIC ≈ $SSE*n^{p/n} = 5520*43^{5/43}$ = **8548.251088**.

Problem S7-TAAF-9.
(a) Fill in the blanks (Venter, pp. 815-816): If one tests for the linear-with-constant alternative emergence pattern and finds that the constant term is more statistically significant than the development factor (or the development factor is not significant at all), it is important to _______ the development triangle as much as possible. One way to do this is by ________________. For these adjustments, use a _______ method instead of a purely multiplicative method.
(b) Fill in the blanks (Venter, pp. 815-816): For the linear-with-constant alternative emergence pattern, the constant is often significant in the age range of ______, especially for _________ lines of business such as ________.

Solution S7-TAAF-9.
(a) If one tests for the linear-with-constant alternative emergence pattern and finds that the constant term is more statistically significant than the development factor (or the development factor is not significant at all), it is important to **normalize** the development triangle as much as possible. One way to do this is by **adjusting for differences among accident years in exposure and cost levels (trend)**. For these adjustments, use a **purely additive** method instead of a purely multiplicative method.
(b) For the linear-with-constant alternative emergence pattern, the constant is often significant in the age range of **0 to 1 years**, especially for **highly variable and slowly reporting** lines of business such as **excess reinsurance**.

Problem S7-TAAF-10.
(a) What method is an example of the assumption of the alternative emergence pattern of a factor times a parameter? (Venter, p. 816)
(b) For a complete development triangle with n accident years, the method in (a) has how many parameters? How does this number compare to the number of parameters involved in the chain-ladder method? (Venter, p. 817)
(c) Fill in the blank: The method in (a) has to produce much ____ (higher or lower?) fit errors than the

chain-ladder method in order to give a better fit statistic.
(d) What method is, according to Venter, an "important special case" of the parameterized method in part (a)? How is it described using Venter's notation with regard to the parameter h(w)?

Solution S7-TAAF-10. (a) The **Bornhuetter-Ferguson (BF) Method**
(b) The BF method has **2n-2** parameters. This is **twice** the number of chain-ladder parameters (n-1).
(c) The method in (a) has to produce much **lower** fit errors than the chain-ladder method in order to give a better fit statistic.
(d) The **Cape Cod (CC) Method:** Sets h(w) equal to a single h for all accident years w.

Problem S7-TAAF-11.
(a) According to Venter (pp. 818-819), what action could improve the predictions of the Bornhuetter-Ferguson model, especially for the most recent years?
(b) How might the Cape Cod model be considered a reduced-parameter version of the Bornhuetter-Ferguson model? (Venter, p. 819)
(c) What two other approaches does Venter (p. 819) describe for reducing the number of parameters in a Bornhuetter-Ferguson model?

Solution S7-TAAF-11.
(a) Reducing the number of parameters (thus using more of the information in the development triangle) could improve the predictions of the Bornhuetter-Ferguson model, especially for the most recent years.
(b) The Cape Cod model has a single ultimate value for all accident years, whereas the BF model has a separate ultimate value for each year. This reduces the number of parameters.
(c) 1. Use a trend line through the BF ultimate-loss parameters. This reduces the number of accident-year parameters to two.
2. Group years using apparent jumps in loss levels; fit a parameter (h) separately to each group.

Problem S7-TAAF-12.
(a) For attempting to fit data to a BF model, what is the expression (using Venter's notation) for the individual residual, the sum of whose squares it is the goal to minimize? (Venter, p. 822)
(b) Venter discusses an iterative regression procedure. What is the starting point of the procedure? (Venter, p. 822)
(c) What is the standard linear regression formula that gives the best values of h for the initial set of values of f? (Venter, p. 822)
(d) After the set of best values of h has been found through the formula in (c), what is the regression formula that gives the best values of f for those values of h? (Venter, p. 822)
(e) For the Cape Cod method, what is the regression formula used to determine h? (Venter, p. 825)
(f) What other method gives the same fit accuracy as the Cape Cod method? Why? (Venter, p. 826)

Solution S7-TAAF-12.
(a) The expression for the individual residual is **[q(w, d) - f(d)*h(w)]**. (This is the difference between the incremental loss and the estimated incremental loss using the Bornhuetter-Ferguson assumption, based on some prior selection of h(w) that is a percentage of expected ultimate losses.)
(b) The **implied incremental development factors f(d) from the chain-ladder method** are the starting point of the iterative regression procedure.
(c) $\mathbf{h(w) = {}_d\sum[f(d)*q(w,d)]/{}_d\sum[f(d)^2]}$**.**

(d) $f(d) = {}_w\sum[h(w)*q*(w,d)]/{}_w\sum[h(w)^2]$.
(e) $h = {}_{w,d}\sum[f(d)*q(w,d)]/{}_{w,d}\sum[f(d)^2]$.
(f) The **additive chain-ladder method** gives the same fit accuracy as the Cape Cod method, since both methods estimate the loss levels for each age with a single value.

Problem S7-TAAF-13. The chain-ladder method assumes that the incremental losses at each age are a linear function of previous cumulative losses. What pattern in a plot of residuals against previous cumulative losses be used to identify a departure from this assumption? (Venter, pp. 829-830)

Solution S7-TAAF-13. If there are consecutive strings of positive or negative residuals, this means that one is attempting to fit a line to a nonlinear pattern of emergence. This would indicate that the chain-ladder method is not optimal.

Problem S7-TAAF-14.
(a) If a plot of residuals over time does *not* show the residuals to exhibit a stable pattern, what two corrections are possible in order for a chain-ladder method to still be usable? (Venter, pp. 830-831)
(b) If the plot of residuals shows a lot of variability around a fixed level of data, how should the data be adjusted (e.g., what data elements should be used)? (Venter, p. 832)

Solution S7-TAAF-14. (a) Two possible corrections are as follows:
1. Use a weighted average of available factors, with more weight going to the more recent years.
2. Adjust the development triangle for measurable instability (e.g., use a Berquist-Sherman adjustment for the settlement rate of claims).
(b) All the data elements should be used if the plot of residuals shows a lot of variability around a fixed level of data. This means that the average factors are actually not changing over time. Adjustments should only be made if the average itself fluctuates with time.

Problem S7-TAAF-15.
(a) Let r be the same correlation coefficient. Let n be the number of observations. What is formula for the test statistic T, as given by Venter on page 833?
(b) Fill in the blanks: According to Venter (p. 833), if we assume the T statistic to follow a _______ distribution with ____ degrees of freedom, then if T > the ____ statistic for ___, then ___ is statistically significant at the 10% level, which would ______ (support or undermine?) the chain-ladder hypothesis.

Solution S7-TAAF-15.
(a) $T = r*[(n-2)/(1-r^2)]^{1/2}$.
(b) According to Venter (p. 833), if we assume the T statistic to follow a **t-**distribution with **(n-2)** degrees of freedom, then if T > the **t-**statistic for **0.9**, then **r** is statistically significant at the 10% level, which would **undermine** the chain-ladder hypothesis.

Problem S7-TAAF-16.
(a) Briefly describe Mack's high-low diagonal test.
(b) Briefly describe the conceptual essence of the high-low diagonal test discussed by Venter. (Venter, p. 834)

Solution S7-TAAF-16.
(a) Mack's test is to count the number of high and low factors along each diagonal and to see whether

these observations are likely to be due to chance.
(b) Venter's test is to use regression to see if any dummy variables are significant and to provide alternatives in case the chain-ladder method is rejected.