Study Guide on Measuring the Variability of Chain-Ladder Reserve Estimates for the Casualty Actuarial Society (CAS) Exam 7 and Society of Actuaries (SOA) Exam GIADV: Advanced Topics in General Insurance

(Based on Thomas Mack's Paper, *"*Measuring the Variability of Chain Ladder Reserve Estimates*"*)

G. Stolyarov II, ASA, ACAS, MAAA, CPCU, ARe, ARC, API, AIS, AIE, AIAF
Study Guide Created in Spring 2011, Updated February 1, 2015

Source: Mack, T., "Measuring the Variability of Chain Ladder Reserve Estimates," Casualty Actuarial Society *Forum*, Spring 1994.

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

Problem S-7-MVCLRE-1. What is the objective of Mack's paper, in terms of a response to the fact that the estimated ultimate claim amount can never be known with certainty? (Mack, p. 103)

Solution S-7-MVCLRE-1. The objective is to enable the construction of *confidence intervals* for the estimated reserves.

Problem S-7-MVCLRE-2. Let $C_{i,k}$ and $C_{i,k+1}$ be the claim amounts for accident year i and development years k and (k+1), respectively. Let f_k be the age-to-age factor for this time period, derived using the chain-ladder method.

(a) Fill in the blanks (Mack, p. 106):
Each increase from $C_{i,k}$ to $C_{i,k+1}$ is considered a ____________ of an expected increase from $C_{i,k}$ to ______, where f_k is an unknown "true" factor of increase which is the same for all accident years and which estimated from the available data.

(b) Using the $C_{i,\_}$ notation, formulate the first assumption of the chain-ladder method, as described by Mack (p. 108). Let I be the year in which all claims have developed to ultimate.

Solution S-7-MVCLRE-2.
(a) Each increase from $C_{i,k}$ to $C_{i,k+1}$ is considered a **random disturbance** of an expected increase from $C_{i,k}$ to $\mathbf{C_{i,k}*f_k}$, where f_k is an unknown "true" factor of increase which is the same for all accident years and which estimated from the available data.

(b) The first assumption of the chain-ladder method is that the information contained in $C_{i,I+1-i}$ in order to project the claims to ultimate cannot be augmented by also using $C_{i,1}$ through $C_{i,I-i}$ or $C_{1,I+1-i}$ through $C_{i-1,I+1-i}$. (That is, past claims for the same accident year or claims of the same maturity for prior accident years are irrelevant. The magnitude of a development factor for a prior maturity should not affect the magnitude of the current estimated development factor.)

Problem S-7-MVCLRE-3.
(a) Is it reasonable to assume, for the chain-ladder method that the variables $\{C_{i,1, \ldots,} C_{i,I}\}$ and $\{C_{j,1, \ldots,} C_{j,I}\}$ for different accident years i and j, are independent?
(b) What, if any, exceptions exist to the assumption in (a)? (Mack, pp. 110-111)

Solution S-7-MVCLRE-3.
(a) Yes. The chain-ladder method explicitly does not take into account any dependency among the accident years.
(b) If there is a change in claim handling, case reserving, or the rate of inflation, this can affect several accident years in the same way and render the assumption of independence dubious.

Problem S-7-MVCLRE-4.
(a) Which of the following is an unbiased estimator of the development factor? (i) The weighted-average chain-ladder factor; (ii) The simple-average chain ladder factor. (Mack, p. 112)
(b) Give a mathematical reason to prefer one of the factors in (a) over the other.
(c) State the *proportionality condition* of a chain-ladder estimate. Let α_k be a proportionality constant. (Mack, p. 113)

Solution S-7-MVCLRE-4.
(a) Both (i) and (ii) are unbiased estimators.
(b) The weighted-average factor is preferable because by choosing weights proportional to claim amounts, we minimize the variance of the weighted average.
(c) Proportionality condition: $Var(C_{j,k+1} \mid C_{j,1}, \ldots, C_{j,k}) = \alpha_k^2 * C_{j,k}$.

Problem S-7-MVCLRE-5.
(a) Give the formula for mean square error $MSE(c_{i,I})$ where D is the set of observed data: $D = \{c_{i,k} \mid i+k \leq I + 1\}$. Express the MSE in terms of the random variable $C_{i,I}$, the specific estimated value $c_{i,I}$, and D.
(b) The formula in (a) involves conditionality. Why is the conditionality important here? (Mack, p. 114)

(c) Reformulate the equation in (a) such that a variance expression is one of the terms.
(d) What does this MSE *not* take into account?
(e) What is the square root of MSE called?
(Mack, p. 115)

Solution S-7-MVCLRE-5.
(a) $MSE(c_{i,I}) = E[(C_{i,I} - c_{i,I})^2 \mid D]$.
(b) The conditionality on D is important because we only want to estimate the error due to *future* random variations. Thus, we assume we already know D, which is the set of past data. Without the conditionality on D, we would be calculating MSE over both the past and the future, which not assist us in predicting the future on the basis of a particular development triangle.
(c) $MSE(c_{i,I}) = Var(C_{i,I} \mid D) + (E(C_{i,I} \mid D) - c_{i,I})^2$.
(d) The MSE does not take into account future changes in the underlying model, such as the emergence of previously unanticipated claim types.
(e) The square root of MSE is **standard error**.

Problem S-7-MVCLRE-6.
Let $R_i = C_{i,I} – C_{i,I+1-i}$ be the outstanding claim reserve for accident year i. Let $r_i = c_{i,I} – C_{i,I+1-i}$ be the estimate of the outstanding claim reserve.
(a) What is the MSE of r_i? Give the formula in terms of r_i, R_i, and D.
(b) To what other MSE is the MSE of r_i equal? What is the verbal meaning of this?
(Mack, p. 116)

Solution S-7-MVCLRE-6.
(a) $MSE(r_i) = E[(r_i - R_i)^2 \mid D]$.
(b) $MSE(r_i) = MSE(c_{i,I})$. This means that the mean square error of the reserve is the same as the mean square error of the ultimate-loss estimate.

Problem S-7-MVCLRE-7.
(a) Given the ultimate claim estimate $c_{i,I}$, known claim data points $C_{j,k}$, estimated development factors f_k, and estimators $ά_k^2$ of the proportionality constants $α_k^2$, what is the formula for estimating $MSE(c_{i,I})$ solely from known data?
(b) In the formula in (a), how is $ά_k^2$ determined (also solely from known data)? This is itself a rather involved equation.
(c) Give the special formula for the latest of the $ά_k^2$ estimators: $ά_{I-1}^2$. Conceptually, why is a special formula needed? (Mack, pp. 116-117)

Solution S-7-MVCLRE-7.
(a) $MSE(c_{i,I}) = (c_{i,I})^2 *({}_{(k = I + 1 – i)}{}^{(I-1)}\sum[(ά_k^2/ f_k^2)*(1/c_{i,k} + 1/[{}_{(j=1)}{}^{(I-k)}\sum(C_{j,k})])])$.
(b) $ά_k^2 = (1/(I-k-1))*[{}_{(j=1)}{}^{(I-k)}\sum((C_{j,k})*(C_{j,k+1}/C_{j,k} - f_k)^2)$, for $1 \le k \le I – 2$.
(c) $ά_{I-1}^2 = \min[ά_{I-2}^4/ά_{I-3}^2, \min(ά_{I-3}^2 , ά_{I-2}^2)]$. A special formula for this constant is needed, because the estimate is based on only a single observation for development years (I-1) and I: $C_{1,I}/C_{1,I-1}$. This is the LDF calculated using the rightmost top two entries of the

loss-development triangle. It is impossible to use the standard formulas to estimate both $ά_{I-1}$ and f_{I-1} from this observation.

Problem S-7-MVCLRE-8.
(a) Give the expression for the symmetric 95%-confidence interval for the reserve R_i.
(b) What distributional assumption may lead the expression in (a) to not reflect reality?
(c) What solution does Mack recommend for the problem in (b)? What useful property does this approach have?
(d) What mathematical formulas are used to obtain the estimates in the solution in part (c)?
(Mack, pp. 118-119)

Solution S-7-MVCLRE-8.
(a) The interval is **$(R_i – 2*se(R_i), R_i + 2*se(R_i))$** where $se(R_i) = se(c_{i,I})$ is the standard error of both the reserve and the ultimate-loss estimate.
(b) The expression in (a) relies on a symmetric Normal distribution of the possible reserve amounts. However, the real-world distribution may be skewed and may be highly volatile, with large standard errors. If the standard error exceeds 50% of R_i, then the lower bound of the interval will be negative, which is not always possible in reality.
(c) Mack recommends using a Lognormal distribution. Using a Lognormal distribution prevents negative boundaries for confidence intervals.
(d) The parameters of the Lognormal distribution are μ_i and σ_i^2. Then the estimates are
$R_i = \exp(\mu_i + \sigma_i^2/2)$;
$(se(R_i))^2 = \exp(2\mu_i + \sigma_i^2)*(\exp(\sigma_i^2) - 1)$; and thus
$\sigma_i^2 = \ln(1 + (se(R_i))^2/R_i^2)$; and
$\mu_i = \ln(R_i) - \sigma_i^2/2$.

Problem S-7-MVCLRE-9.
(a) If R_i is the reserve for the accident year i, provide the formula for the overall reserve R of the accident years 1 through I represented in a loss-development triangle.
(b) In order to obtain the variance of R, why is it not possible to simply add the squares of the standard errors of each R_i?
(c) Give the formula for $(se(R))^2$, the square of the standard error of R. (Mack, p. 120)

Solution S-7-MVCLRE-9.
(a) $R = R_2 + … + R_I$. (There is no R_1 term, since, presumably, the latest known value of the claim amount for the earliest displayed accident year is already at ultimate.)
(b) For each i, the estimators R_i are not independent of one another. They are positively correlated, because they are all influenced by the same age-to-age factors f_k.
(c) $(se(R))^2 = {}_{i=2}{}^{I}\sum[(se(R_i))^2 + c_{i,I}*({}_{j=i+1}{}^{I}\sum(c_{j,I}))*{}_{k=I+1-i}{}^{I-1}\sum([2ά_k^2/f_k^2]/[{}_{n=1}{}^{I-k}\sum(C_{n,k})])]$.

Problem S-7-MVCLRE-10.
Mack (pp. 122-124) describes three additional estimators for development factors: $f_{k,0}$ (the $c_{i,k}^2$-weighted average), $f_{k,1}$ (the $c_{i,k}$-weighted average), and $f_{k,2}$ (the ordinary unweighted average). Give formulas for each estimator.

Solution S-7-MVCLRE-10.
$\mathbf{f_{k,0} = {}_{i=1}^{I-k}\sum[(C_{i,k}*C_{i,k+1})/({}_{i=1}^{I-k}\sum[C_{i,k}^2])].}$
$\mathbf{f_{k,1} = ({}_{i=1}^{I-k}\sum[C_{i,k+1}])/({}_{i=1}^{I-k}\sum[C_{i,k}]).}$ (Note: This is the same as the weighted-average chain-ladder factor f_k.)
$\mathbf{f_{k,2} = (1/[I-k])({}_{i=1}^{I-k}\sum[C_{i,k+1}/C_{i,k}]).}$ (Note: This is the straight-average chain-ladder factor.)

Problem S-7-MVCLRE-11.
(a) Mack (p. 124) recommends analyzing what plot to check for a linear relationship?
(b) Mack (p. 125) recommends analyzing what three plots to check for random behavior (and to test whether the variance assumption is met)?

Solution S-7-MVCLRE-11.
(a) To check for a linear relationship, analyze the plot of $\mathbf{C_{i,k+1}}$ **against** $\mathbf{C_{i,k}}$.
(b) To test the variance assumptions, analyze the following plots:
(i) For $f_{k,0}$: $\mathbf{C_{i,k+1} - C_{i,k}*f_{k,0}}$ **against** $\mathbf{C_{i,k}}$.
(ii) For $f_{k,1}$: $\mathbf{(C_{i,k+1} - C_{i,k}*f_{k,1})/\sqrt{}(C_{i,k})}$ **against** $\mathbf{C_{i,k}}$.
(iii) For $f_{k,2}$: $\mathbf{(C_{i,k+1} - C_{i,k}*f_{k,2})/(C_{i,k})}$ **against** $\mathbf{C_{i,k}}$.

Problem S-7-MVCLRE-12. Review. What are the three basic implicit chain-ladder assumptions, as described by Mack (p. 121). (Note that this would be a reasonable exam question.)

Solution S-7-MVCLRE-12.
1. $\mathbf{E(C_{i,k+1} \mid C_{i,1}, \ldots, C_{i,k}) = C_{i,k}*f_k}$. (Given observed data, the expected value of the next development period's claims is the current development period's claims, multiplied by the "true" development factor f_k.)
2. **Independence of accident years**.
3. $\mathbf{Var(C_{i,k+1} \mid C_{i,1}, \ldots, C_{i,k}) = \alpha_k^2*C_{j,k}}$ (Proportionality condition)