PRACTICAL ISSUES

- How to handle issues which can derail the bootstrap techniques. Common issues include negative incremental values, extreme outcomes and needing to use different averages.

NEGATIVE INCREMENTAL VALUES

- These can break the GLM process when they occur in the observed data or arise from the fitted values. They can also produce extreme outcomes.
- Look at the column sums of the incremental values. If some of the q(w,d) are negative but the column sums are positive then we can set

$$\ln(q(w,d)) = \begin{cases} \ln(q(w,d)) & \text{when } q(w,d)>0 \\ 0 & \text{when } q(w,d)=0 \\ -\ln(q(w,d)) & \text{when } q(w,d)<0 \end{cases}$$

and proceed as before.

- However, if we have a negative column sum this doesn't work. In that case, you need an adjusted log-link function

$\ln(q^{+}(w,d)) = \ln(q(w,d) - \Psi)$ where Ψ can be chosen in two ways:

1. Set Ψ equal to the largest negative incremental value found in the entire triangle. OR
2. Set Ψ equal to the largest negative column sum.

- In the second approach, you may still have some negative incremental values but the column sums are no longer negative so you can use the earlier adjustment.
- Once you've solved the GLM equation, you must shift the fitted triangle of incremental values $m_{w,d}$, back by adding Ψ
- A negative incremental value in the fitted results may be problematic because the unscaled residual contains $\sqrt{m_{w,d}^{z}}$
  - Some z values (such as z=1) leave this undefined.
- Shapland recommends using $r_{w,d} = \dfrac{q(w,d) - m_{w,d}}{\sqrt{|m_{w,d}|^{z}}}$ setting $q^*(w,d) = r^*\sqrt{|m_{w,d}|^{z}} + m_{w,d}$ *r* denotes a sampled residual*

- Negative incremental values in the fitted results can also break how we incorporate future process variance because we cannot sample from a Gamma distribution with negative variance.
- There are 2 ways to address this. Sample from $\Gamma(|m_{w,d}|)$ then:

1. Multiply by -1 to turn into a negative increment
   a. A drawback of #1 is it flips the distribution about the y-axis so we have a left skewed distribution instead of a right skewed. This can lead to understating results.
2. Add $2m_{w,d}$ to the result.

my notes:

MANAGING EXTREME OUTCOMES

1. Remove the simulation results from the bootstrap
2. Recalibrate the model. Omit an accident year, model gross of reinsurance and reinsurance recoveries separately, etc.
3. Limit the incremental losses. Replace negative incremental losses with 0 or even implement or arbitrarily lower the bound on the incremental loss.

N-YEAR WEIGHTED AVG ODP BOOTSTRAP

This uses the latest N-year average instead of the standard all-year average. We now have a trapezoid of interest.

Follow the standard ODP Bootstrap process but:

1. Calculate N-year averages.
2. Perform the reverse chain-ladder across the entire triangle.
3. Calculate residuals only for the trapezoid
4. Resample with replacement from the trapezoid but perform the sampling for each cell in the triangle.

PRACTICAL ISSUES (cont)

Shapland

Homoscedasticity is when the residuals all have the same variance. If some development periods have greater variability in their residuals then others we have **heteroscedasticity**.

Heteroscedasticity can be addressed by:

- *Stratified sampling: sample from subsets of the residuals which have similar variance*
- *Apply hetero-adjustment parameters*
- *Apply non-constant scale parameters*

NON-CONSTANT SCALE PARAMETERS

1. Group together development periods whose standardized residuals have homogeneous variance. Let n_i be the number of residuals in the i^{th} heterogroup.
2. Calculate the number of parameters in the model. The heterogroups add # heterogroup - 1 extra parameters.
3. Calculate the standard scale parameter $\Phi = \dfrac{\sum r_{w,d}^2}{N-p}$
4. Calculate $\Phi_i = \sum_{i=1}^{n}\left(r_{w,d}\sqrt{\dfrac{N}{N-p}}\right)^2 \Big/ n_i$
5. Calculate the heteroadjustment factor for each hetero group $h_i = \sqrt{\Phi/\Phi_i}$
6. Set $r_{w,d}^{iH} = \dfrac{q(w,d) - m_{w,d}}{\sqrt{|m_{w,d}|^z}} \cdot f_{w,d}^H \cdot h_i$
7. Sample with replacement from the set of all $r_{w,d}^{iH}$
8. Calculate $q^{i*}(w,d) = \dfrac{r^*}{h_i}\sqrt{m_{w,d}^{\ z}} + m_{w,d}$
9. Incorporate future process variance by sampling from $\Gamma(q^{i*}(w,d),\ \Phi_i q^{i*}(w,d))$

my notes:

HETERO-ADJUSTMENT PARAMETERS

1. Group development periods whose standardized residuals have homogeneous variance
2. Calculate the standard deviation for each group
3. Calculate the hetero-adjustment parameter for each group:
$$h_i = \frac{\text{stdev}(\bigcup_i^j r_{w,d}^H)}{\text{stdev}(\bigcup_i r_{w,d}^H)}$$
   a. The numerator is the standard deviation across all (groups of) residuals. It doesn't vary by residual group. The denominator is the standard deviation of the i^{th} residual group.
4. The number of additional parameters is the number of hetero-adjustment groups minus 1. Calculate the scale parameter and hat matrix adjustment factor based on the number of parameters we now have.
5. For each group i, multiply the standardized residuals within group i by h_i. $r_{w,d}^{iH} = \dfrac{q(w,d) - m_{w,d}}{\sqrt{m_{w,d}^{\ z}}} \cdot f_{w,d}^H \cdot h_i$
6. Since all residuals now have the same standard deviation, we can resample with replacement from across the entire triangle.
7. Calculate $q^{i*}(w,d) = \dfrac{r^*}{h_i} \cdot \sqrt{|m_{w,d}|^z} + m_{w,d}$

   to reinstate the heteroscedasticity present in the triangle.
8. Incorporate future process variance by sampling from $\Gamma\left(q^{i*}(w,d),\ \dfrac{\Phi}{h_i^2} q^{i*}(w,d)\right)$

HETEROECTHESIOUS DATA

This refers to incomplete or non-uniform exposures.

Partial First Development Period: the first development column has a different exposure period. *Example: Accident years evaluated at 6mo, 18mo, 30mo, ...*	Deterministic and bootstrap methods automatically control for this. Except: the latest accident year is incomplete and is annualized by the link ratios. Do not assume the first half and second half are the same, so halve the future projected incremental loss to match the exposure base.
Partial Last Development Period: The latest diagonal is at a different cadence to the rest of the triangle *Example: Accident years evaluated at 12mo, 24mo, 36mo, 48mo and 54mo. The last is 6 months after, not 12.*	The issue is the latest accident year is only 6mo, not 12mo so cannot readily apply the link ratios. Deterministic approach: Exclude latest diagonal & apply interpolation. Bootstrapping approach: Annualize the exposure by doubling the incremental exposure. Once complete, reduce the latest diagonal by half.