Problems 1 to 3 refer to the following situation. An insurance company has two group policies. The aggregate claim amounts (in millions of dollars) for the first three policy years are summarized in the table below. Assume that the two groups have the same number of insureds.

<table>
<thead>
<tr>
<th>Group</th>
<th>Policy Year</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>5</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>11</td>
<td>13</td>
<td>12</td>
</tr>
</tbody>
</table>

1. Find the Buhlmann credibility premium for group 1 for the fourth year.
A) 8.0  B) 8.2  C) 8.4  D) 8.6  E) 8.8

2. Using the credibility-weighted average estimate of \( \mu \) (also referred to as the method that preserves total losses), find the Buhlmann credibility premium for group 1 for the fourth year.
A) 8.0  B) 8.2  C) 8.4  D) 8.6  E) 8.8

3. Suppose that the aggregate claim amounts for group 2 for policy years 1, 2 and 3 are 2, 8 and 14 (instead of 11, 13 and 12). Find the estimated Buhlmann credibility premium for group 1 for the fourth year.
A) 8.0  B) 8.2  C) 8.4  D) 8.6  E) 8.8

Problems 4 and 5 refer to the data of Example CR-34 in the notes, with the following modification. Assume that there is a third group with the following experience for the three years:

<table>
<thead>
<tr>
<th>Policy Year</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggregate Claim</td>
<td>10,000</td>
<td>15,000</td>
<td>13,500</td>
</tr>
<tr>
<td>Size of Group</td>
<td>50</td>
<td>60</td>
<td>60</td>
</tr>
</tbody>
</table>

4. Find the credibility premium per exposure unit for policyholder 1 for the fourth year (nearest 1).
A) 203  B) 205  C) 207  D) 209  E) 211

5. Using the credibility-weighted average estimate of \( \mu \), find the Buhlmann credibility premium per unit of exposure for group 1 for the fourth year (nearest 1).
A) 203  B) 205  C) 207  D) 209  E) 211
Problems 6 and 7 refer to the data of Example CR-34 in the notes, with the following modification. For group 1, assume that there is no data for the first policy year (but all three years of data are still available for group 2).

6. Find the credibility premium per exposure unit for policyholder 1 for the fourth year (nearest 1).
   A) 206  B) 208  C) 210  D) 212  E) 214

7. Using the credibility-weighted average estimate of $\mu$, find the Buhlmann credibility premium per unit of exposure for group 1 for the fourth year (nearest 1).
   A) 206  B) 208  C) 210  D) 212  E) 214

8. For a large sample of insureds, the observed relative frequency of claims during an observation period is as follows:

<table>
<thead>
<tr>
<th>Number of Claims</th>
<th>Relative Frequency of Claims</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>61.9%</td>
</tr>
<tr>
<td>1</td>
<td>28.4%</td>
</tr>
<tr>
<td>2</td>
<td>7.8%</td>
</tr>
<tr>
<td>3</td>
<td>1.6%</td>
</tr>
<tr>
<td>4</td>
<td>.3%</td>
</tr>
<tr>
<td>5 or more</td>
<td>0</td>
</tr>
</tbody>
</table>

Assume that for a randomly chosen insured, the underlying conditional distribution of number of claims per period given the parameter $\Theta$ is Poisson with parameter $\Theta$. Given an individual who had $c$ claims in the observation period, use semiparametric empirical Bayesian estimation to find expected number claims that the individual will have in the next period.
   A) .93$c + .465$  B) .93$c + .500$  C) .07$c + .465$  D) .07$c + .500$  E) .93$c + .035$

9. You are given the following:
   - The number of losses arising from 500 individual insureds over a single period of observation is distributed as follows:

<table>
<thead>
<tr>
<th>Number of Losses</th>
<th>Number of Insureds</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>450</td>
</tr>
<tr>
<td>1</td>
<td>30</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>5 or more</td>
<td>0</td>
</tr>
</tbody>
</table>

   - The number of losses for each insured follows a Poisson distribution, but the mean of each distribution may be different for individual insureds. Determine the Buhlmann credibility of the experience of an individual insured over a single period.
10. You are given the following table of data for three policyholders over a three year period.

<table>
<thead>
<tr>
<th>Policy Year</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Policyholder</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>40</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>Number of Claims</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Claim Size</td>
<td>200</td>
<td>220</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>120</td>
<td>120</td>
</tr>
<tr>
<td>Number of Claims</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Claim Size</td>
<td>200</td>
<td>200</td>
<td>150</td>
</tr>
<tr>
<td>3</td>
<td>50</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>Number of Claims</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Claim Size</td>
<td>200</td>
<td>250</td>
<td></td>
</tr>
</tbody>
</table>

Apply the nonparametric empirical Bayes credibility method to find the credibility premium per claim in the 4th year for Policyholder 2 using the standard method for \( \mu \) (not the method that preserves total losses).

11. Semi-parametric empirical Bayesian credibility is being applied in the following situation.

The distribution of annual losses \( X \) on an insurance policy is uniform on the interval \((0, \theta)\), where \( \theta \) has an unknown distribution. A sample of annual losses for 100 separate insurance policies is available. It is found that

\[
\sum_{i=1}^{100} X_i = 200 \quad \text{and} \quad \sum_{i=1}^{100} X_i^2 = 600.
\]

For a particular insurance policy, it is found that the total losses over a 3 year period is 4. Find the semi-parametric estimate of the losses in the 4th year for this policy.

12. (SOA) An insurer has data on losses for four policyholders for seven years. \( X_{ij} \) is the loss from the \( i^{th} \) policyholder for year \( j \). You are given:

\[
\sum_{i=1}^{4} \sum_{j=1}^{7} (X_{ij} - \overline{X}_j)^2 = 33.60 \quad \text{and} \quad \sum_{i=1}^{4} (\overline{X}_i - \overline{X})^2 = 3.30
\]

Calculate the Buhlmann credibility factor for an individual policyholder using nonparametric empirical Bayes estimation.

(A) Less than 0.74  
(B) At least 0.74, but less than 0.77  
(C) At least 0.77, but less than 0.80  
(D) At least 0.80, but less than 0.83  
(E) At least 0.83
13. (SOA) The number of claims a driver has during the year is assumed to be Poisson distributed with an unknown mean that varies by driver. The experience for 100 drivers is as follows:

<table>
<thead>
<tr>
<th>Number of Claims during the Year</th>
<th>Number of Drivers</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>54</td>
</tr>
<tr>
<td>1</td>
<td>33</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
</tr>
</tbody>
</table>

Determine the credibility of one year’s experience for a single driver using semiparametric empirical Bayes estimation.

(A) 0.046        (B) 0.055        (C) 0.061        (D) 0.068        (E) 0.073

14. (SOA) The following information comes from a study of robberies of convenience stores over the course of a year:

(i) $X_i$ is the number of robberies of the $i^{th}$ store, with $i = 1, 2, \ldots, 500$.

(ii) $\sum X_i = 50$  
(iii) $\sum X_i^2 = 220$

(iv) The number of robberies of a given store during the year is assumed to be Poisson distributed with an unknown mean that varies by store.

Determine the semiparametric empirical Bayes estimate of the expected number of robberies next year of a store that reported no robberies during the studied year.

(A) Less than 0.02        (B) At least 0.02, but less than 0.04
(C) At least 0.04, but less than 0.06       (D) At least 0.06, but less than 0.08       (E) At least 0.08

15. Survival times are available for four insureds, two from Class A and two from Class B. The two from Class A died at times $t = 1$ and $t = 9$. The two from Class B died at times $t = 2$ and $t = 4$. Nonparametric Empirical Bayes estimation is used to estimate the mean survival time for each class. Unbiased estimators of the expected value of the process variance and the variance of the hypothetical means are used. Estimate $Z$, the Buhlmann credibility factor.

(A) 0       (B) 2/19       (C) 4/21       (D) 8/25       (E) 1

16. You are given the following experience for two insured groups:

<table>
<thead>
<tr>
<th>Group</th>
<th>Year</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>25</td>
</tr>
<tr>
<td>Total</td>
<td>Number of members</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Average loss per member</td>
<td>109</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[
\sum_{i=1}^{2} \sum_{j=1}^{3} m_{ij} (x_{ij} - \bar{x}_i)^2 = 2020 \\
\sum_{i=1}^{2} m_i (\bar{x}_i - \bar{x})^2 = 4800
\]

Determine the nonparametric Empirical Bayes credibility premium for group 1, using the method that preserves total losses.

(A) 98       (B) 99       (C) 101       (D) 103       (E) 104
17. The number of claims per month for a given risk is assumed to be Poisson distributed with an unknown mean that varies by risk. It is found that for a risk that has reported no claims for the past month, the semiparametric empirical Bayes estimate of the expected number of claims next month is \( \frac{1}{30} \), and it is found that for a risk that has reported no claims for the past two months, the semiparametric empirical Bayes estimate of the expected number of claims next month is \( \frac{1}{55} \). Find the semiparametric empirical Bayes estimate of the expected number of claims next month for a risk that has reported no claims for the past three months.

(A) \( \frac{1}{70} \)  (B) \( \frac{1}{75} \)  (C) \( \frac{1}{80} \)  (D) \( \frac{1}{85} \)  (E) \( \frac{1}{90} \)

18. You are given the following table of data for three policyholders over a three year period.

Policy Year → 1 2 3
Policyholder ↓

<table>
<thead>
<tr>
<th>Policyholder</th>
<th>Number of Claims</th>
<th>Average Claim Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>40</td>
<td>200</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>220</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>200</td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>200</td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>150</td>
</tr>
<tr>
<td>3</td>
<td>50</td>
<td>200</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>250</td>
</tr>
</tbody>
</table>

Apply the nonparametric empirical Bayes credibility method to find the credibility premium per claim in the 4th year for Policyholder 2.

19. A particular type of individual health insurance policy models the annual loss per policy as an exponential distribution with a mean that varies with individual insured. A sample of 1000 randomly selected policies results in the following data regarding annual loss amounts in interval grouped form.

<table>
<thead>
<tr>
<th>Interval</th>
<th>Number of Losses</th>
</tr>
</thead>
<tbody>
<tr>
<td>[0, 100]</td>
<td>500</td>
</tr>
<tr>
<td>(100, 200]</td>
<td>250</td>
</tr>
<tr>
<td>(200, 500]</td>
<td>150</td>
</tr>
<tr>
<td>(500, 1000]</td>
<td>60</td>
</tr>
<tr>
<td>(1000, 2000]</td>
<td>40</td>
</tr>
</tbody>
</table>

It is assumed that the loss amounts are uniformly distributed within each interval.

Apply semiparametric empirical Bayes credibility to estimate the loss in the 3rd year for a particular individual who had annual policy losses of 150 in the first year and 0 in the second year.
20. You are given the following table of data for three policyholders over a three year period.

<table>
<thead>
<tr>
<th>Policy Year</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Policyholder</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Claims</td>
<td>40</td>
<td>50</td>
<td>40</td>
</tr>
<tr>
<td>Average Claim Size</td>
<td>200</td>
<td>220</td>
<td>250</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Claims</td>
<td>100</td>
<td>120</td>
<td>120</td>
</tr>
<tr>
<td>Average Claim Size</td>
<td>200</td>
<td>200</td>
<td>150</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Claims</td>
<td>50</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>Average Claim Size</td>
<td>200</td>
<td>250</td>
<td></td>
</tr>
</tbody>
</table>

Apply the nonparametric empirical Bayes credibility method to find the credibility premium per claim in the 4th year for Policyholder 2 using the standard method for \( \mu \) (not the method that preserves total losses).

21. Semi-parametric empirical Bayesian credibility is being applied in the following situation. The distribution of annual losses \( X \) on an insurance policy has pdf \( f(x|\theta) = \frac{2x}{\theta^2} \) for \( 0 < x < \theta \), where \( \theta \) has an unknown distribution. A sample of annual losses for 100 separate insurance policies is available. It is found that \( \sum_{i=1}^{100} X_i = 200 \) and \( \sum_{i=1}^{100} X_i^2 = 600 \).

For a particular insurance policy, it is found that the total losses over a 3 year period is 4. Find the semi-parametric estimate of the losses in the 4th year for this policy.

22. The parameter \( \theta \) has a uniform distribution on the interval \( (1, 2) \).

The model distribution \( S \) is a compound distribution with frequency \( N \) that has a Poisson distribution with mean \( \theta \) and with severity \( Y \) that has an exponential distribution with mean \( \theta \).

One observed value of \( S \) is available. Find the Buhlmann credibility premium in terms of \( S \). Identify the components of the Buhlmann method \( (\mu, v, a) \).

(The usual independence assumptions are made for the compound distribution).
CREDIBILITY - PROBLEM SET 7 SOLUTIONS

1. \( r = 2 \) policyholders (groups) and \( n_1 = n_2 = n_3 = 3 \) exposure periods (years) for each group, and \( m_{ij} = 1 \) exposure unit for combination of group and year.

\[
X_1 = \frac{5+8+11}{3} = 8, \quad X_2 = \frac{11+13+12}{3} = 12, \quad X = \frac{8+12}{2} = 10 = \bar{\mu}.
\]

\[
\hat{\nu} = \frac{1}{r(n-1)} \sum_{i=1}^{r} \sum_{j=1}^{n} (X_{ij} - \bar{X}_i)^2 = \frac{1}{2(2)} \left( [(5 - 8)^2 + (8 - 8)^2 + (11 - 8)^2] + [(11 - 12)^2 + (13 - 12)^2 + (12 - 12)^2] \right) = 5.0, \text{ and}
\]

\[
\hat{a} = \frac{1}{r-1} \sum_{i=1}^{r} (\bar{X}_i - X)^2 - \frac{\hat{\nu}}{n} = \frac{1}{3} [(8 - 10)^2 + (12 - 10)^2] - \frac{5}{3} = \frac{19}{3}.
\]

Then \( \hat{k} = \frac{\hat{\nu}}{\hat{a}} = \frac{5}{19/3} = .7895 \), and the estimated credibility factor for group 1 is

\[
\hat{Z}_1 = \frac{m_1}{m_1 + k} = \frac{3}{3 + .7895} = .7917. \quad \text{The credibility premium for group 1 for the fourth year is}
\]

\[
\hat{Z}_1 X_1 + (1 - \hat{Z}_1) \bar{\mu} = (.7917)(8) + (.2083)(10) = 8.42. \quad \text{Answer: C}
\]

2. The credibility-weighted average estimate of \( \mu \)

\[
\text{From Problem 1, we have } \hat{Z}_1 = .7917. \quad \text{We find } \hat{Z}_2 = \frac{m_2}{m_2 + k} = \frac{3}{3 + .7895} = .7917.
\]

The credibility-weighted average estimate of \( \mu \) is

\[
\frac{(.7917)(8) + (.7917)(12)}{.7917 + .7917} = 10.
\]

This is the same as the original sample mean estimate of \( \mu \), so the resulting credibility premium will be the same as in Problem 1. It is not a coincidence that the credibility-weighted estimate of \( \mu \) is the same as the sample mean estimate. When we have an "equal sample size" data set \((n_1 = n_2 = \cdots = n_r = n \quad \text{and} \quad m_{ij} = 1 \quad \text{for all} \quad i, j)\) then the two will always be equal.

Answer: C

3. \( X_1 = \frac{5+8+11}{3} = 8, \quad X_2 = \frac{2+8+14}{3} = 8, \quad X = \frac{8+8}{2} = 8 = \bar{\mu}, \)

\[
\hat{\nu} = \frac{1}{(2)(2)} \left( [(5 - 8)^2 + (8 - 8)^2 + (11 - 8)^2] + [(8 - 8)^2 + (8 - 8)^2 + (14 - 8)^2] \right) = 22.5, \text{ and}
\]

\[
\hat{a} = \frac{1}{r-1} \sum_{i=1}^{r} (\bar{X}_i - X)^2 - \frac{\hat{\nu}}{n} = \frac{1}{3} [(8 - 8)^2 + (8 - 8)^2] - \frac{22.5}{3} = - \frac{22.5}{3} < 0
\]

Since \( \hat{a} < 0 \), we assign a value of 0 to \( \hat{Z}_1 \). The credibility premium is

\[
\hat{Z}_1 X_1 + (1 - \hat{Z}_1) \bar{\mu} = \bar{\mu} = 8. \quad \text{Answer: A}
\]
4. There are \( r = 3 \) policyholders, with \( n_1 = n_2 = n_3 = 3 \) exposure periods for each policyholder. Exposure units are \( m_{11} = 40, m_{12} = 50, m_{13} = 70 \), so that \( m_1 = 160 \), and \( m_{21} = 100, m_{22} = 120, m_{23} = 115 \), so that \( m_2 = 335 \) and \( m_{31} = 50, m_{32} = 60, m_{33} = 60 \), so that \( m_3 = 170 \). As found in Example CR-34,
\[
X_{11} = \frac{8000}{40} = 200, \, X_{12} = \frac{8000+11,000}{50} = 220, \, X_{13} = \frac{15,000}{70} = 214.29, \quad \text{and} \\
X_1 = \frac{8000+11,000+15,000}{40+50+70} = 212.50, \quad \text{and} \quad X_{21} = \frac{20,000}{100} = 200, \, X_{22} = \frac{24,000}{220} = 200, \quad \text{and} \\
X_{23} = \frac{19,000}{115} = 165.22, \quad \text{and} \quad X_2 = \frac{20,000+24,000+19,000}{100+120+115} = 188.06, \quad \text{and with the additional} \\
policyholder, we have \( X_{31} = \frac{10,000}{50} = 200, \, X_{32} = \frac{15,000}{60} = 250, \, X_{33} = \frac{13,500}{60} = 225, \quad \text{and} \\
\bar{X}_3 = \frac{10,000+15,000+13,500}{50+60+60} = 226.47. \\
\]

Overall number of exposure units is \( m = m_1 + m_2 + m_3 = 665 \). The estimate of the overall mean is
\[
\bar{\mu} = \bar{X} = \frac{m_1 \bar{X}_1 + m_2 \bar{X}_2 + m_3 \bar{X}_3}{m} = \frac{(160)(212.50)+(335)(188.06)+(170)(226.47)}{160+335+170} = 203.76. \\
\]
\[
\hat{\nu} = \frac{1}{\sum_{i=1}^{m} (n_i-1)} \cdot \left[ \sum_{i=1}^{m} \sum_{j=1}^{r} m_{ij} (X_{ij} - \bar{X}_i)^2 \right] = \frac{1}{(3-1)+(3-1)+(3-1)} = \frac{1}{6}. \\
\]
\[
\left((40)(200 - 212.50)^2 + (50)(220 - 212.50)^2 + (70)(214.29 - 212.50)^2 \right] \\
+ \left( (100)(200 - 188.06)^2 + (120)(200 - 188.06)^2 + (115)(165.22 - 188.06)^2 \right) \\
+ \left( (50)(200 - 226.47)^2 + (60)(250 - 226.47)^2 + (60)(225 - 226.47)^2 \right) = 28,171. \\
\]

The estimated variance of the hypothetical means is
\[
\hat{\alpha} = \frac{1}{m-\frac{1}{m} \sum_{i=1}^{m} n_i^2} \cdot \left[ \sum_{i=1}^{m} m_i (\bar{X}_i - \bar{X})^2 - \hat{\nu} (r-1) \right] = \frac{1}{665-\frac{1}{665}[160^2+335^2+170^2]} = \frac{1}{665} \cdot \frac{1}{160^2+335^2+170^2}. \\
\]
\[
\left[ (160)(212.50 - 203.76)^2 + (335)(188.06 - 203.76)^2 + (170)(226.47 - 203.76)^2 \right. \\
\left. - (28,171)(2) \right] = 304.45. \\
\]

The estimated value of \( k \) is \( \hat{k} = \frac{\hat{\nu}}{\hat{\alpha}} = \frac{28,171}{304.45} = 92.53 \), and the credibility factors for the three policyholders are \( \hat{Z}_1 = \frac{m_1}{m_1+k} = \frac{160}{160+92.53} = .634 \), \( \hat{Z}_2 = \frac{m_2}{m_2+k} = \frac{335}{335+92.53} = .784 \), and \( \hat{Z}_3 = \frac{m_3}{m_3+k} = \frac{170}{170+92.53} = .648 \). 

The credibility premium per exposure unit for policyholder 1 for the fourth year is
\[
\hat{Z}_1 \bar{X}_1 + (1 - \hat{Z}_1) \bar{\mu} = (.634)(212.50) + (.366)(203.76) = 209.30. \quad \text{Answer: D} \\
\]

5. From Problem 4, we have \( \hat{Z}_1 = .634, \hat{Z}_2 = .784, \hat{Z}_3 = .648 \). The credibility-weighted estimate of \( \mu \) is
\[
\frac{\sum_{i=1}^{r} \hat{Z}_i \bar{X}_i}{\sum_{i=1}^{r} \hat{Z}_i} = \frac{\hat{Z}_1 \bar{X}_1 + \hat{Z}_2 \bar{X}_2 + \hat{Z}_3 \bar{X}_3}{\hat{Z}_1 + \hat{Z}_2 + \hat{Z}_3} \\
= \frac{(.634)(212.50) + (.784)(188.06) + (.648)(226.47)}{.634 + .784 + .648} = 207.61. \\
\]

The credibility premium for group 1 per unit of exposure using this estimate of \( \mu \) is
\[
\hat{Z}_1 \bar{X}_1 + (1 - \hat{Z}_1) \bar{\mu} = (.634)(212.50) + (.366)(207.61) = 210.71. \quad \text{Answer: E} \\
\]
6. There are \( r = 2 \) policyholders, with \( n_1 = 2 \), \( n_2 = 3 \) exposure periods for the two policyholders. Exposure units are \( m_{11} = 50 \) (second policy year) and \( m_{12} = 70 \) (third policy year) so that \( m_1 = 120 \), and as in the original Example CR-34, \( m_{21} = 100 \), \( m_{22} = 120 \), \( m_{23} = 115 \), so that \( m_2 = 335 \). The average claim amounts per exposure unit are

\[
X_{11} = \frac{11,000}{50} = 220, \quad X_{12} = \frac{15,000}{70} = 214.29, \quad \text{and} \quad X_3 = \frac{11,000 + 15,000}{50 + 70} = 216.67, \text{ and}
\]

\[
X_{21} = \frac{20,000}{100} = 200, \quad X_{22} = \frac{24,000}{220} = 200, \quad X_{23} = \frac{19,000}{115} = 165.22, \quad \text{and}
\]

\[
X_2 = \frac{20,000 + 24,000 + 19,000}{100 + 220 + 115} = 188.06.
\]

The overall number of exposure units is \( m = m_1 + m_2 = 120 + 335 = 455 \).

\[
\hat{\mu} = \bar{X} = \frac{(120)(216.67) + (335)(188.06)}{120 + 335} = 195.61,
\]

\[
\hat{v} = \frac{1}{(2-1) + (3-1)} \cdot \left( [(50)(220 - 216.67)^2 + (70)(214.29 - 216.67)^2] \\
+ [(100)(200 - 188.06)^2 + (120)(200 - 188.06)^2 + (115)(165.22 - 188.06)^2] \right) \\
= 30,769,
\]

\[
\hat{a} = \frac{1}{455 - \frac{1}{35}[120^2 + 335^2]} \times \left[(120)(216.67 - 195.61)^2 + (335)(188.06 - 195.61)^2 - (30,769)(1)\right] = 235.1.
\]

The estimated value of \( k \) is \( \hat{k} = \frac{\hat{v}}{\hat{a}} = \frac{30,769}{235.1} = 130.9 \), and the credibility factors for the two policyholders are \( \hat{Z}_1 = \frac{m_1}{m_1 + \hat{k}} = \frac{120}{120 + 130.9} = .478 \), \( \hat{Z}_2 = \frac{m_2}{m_2 + \hat{k}} = \frac{335}{335 + 130.9} = .719 \).

The credibility premium for group 1 per unit of exposure using this estimate of \( \mu \) is

\[
\hat{Z}_1 \bar{X}_1 + (1 - \hat{Z}_1)\hat{\mu} = (.478)(216.67) + (.522)(195.61) = 205.68. \quad \text{Answer: A}
\]

7. From Problem 6, we have \( \hat{Z}_1 = .478 \) and \( \hat{Z}_2 = .719 \). The credibility- weighted estimate of \( \mu \) is

\[
\hat{\mu} = \frac{\sum_{i=1}^{2} \hat{Z}_i X_i}{\sum_{i=1}^{2} \hat{Z}_i} = \frac{\hat{Z}_1 X_1 + \hat{Z}_2 X_2}{\hat{Z}_1 + \hat{Z}_2} = \frac{(.478)(216.67) + (.719)(188.06)}{.478 + .719} = 199.48.
\]

The credibility premium for group 1 per unit of exposure using this estimate of \( \mu \) is

\[
\hat{Z}_1 \bar{X}_1 + (1 - \hat{Z}_1)\hat{\mu} = (.478)(216.67) + (.522)(199.48) = 207.70. \quad \text{Answer: B}
\]

8. The average number of claims per insured is

\[
\bar{X} = (0)(.619) + (1)(.284) + (2)(.078) + (3)(.016) + (4)(.003) = .500.
\]

This is the estimate of \( E[X] = \mu \). Since the conditional distribution of \( X \) given \( \Theta \) is Poisson with parameter \( \Theta \), we have \( E[X|\Theta] = Var[X|\Theta] = \Theta \). Then, since \( E[E[X|\Theta]] = E[X] \), our estimate for \( v = E[Var[X|\Theta]] \) is also .50 (since \( Var[X|\Theta] = E[X|\Theta] \)). We estimate \( Var[X] \), the variance of the relative claim frequency per insured; the relative frequency of claims forms the estimated probability distribution of \( X \).

\[
Var[X] = (0 - .5)^2(.619) + (1 - .5)^2(.284) + (2 - .5)^2(.078) + (3 - .5)^2(.016) + (4 - .5)^2(.003) = .538.
\]
8. continued
Since \( \text{Var}[X] = v + a \), we use the estimated variance of \( X \) along with the estimate of \( v \) to get an estimate of \( a \): \( \hat{a} = \text{Var}[X] - \hat{v} = .538 - .5 = .038 \).

The estimate of \( k \) is \( \hat{k} = \frac{\hat{v}}{\hat{a}} = \frac{.50}{.038} = 13.2 \), and the estimated credibility factor for one individual is \( \hat{Z} = \frac{1}{1+13.2} = .070 \). The credibility premium for the next period for an individual who had \( c \) claims in the current period is \( \hat{Z}c + (1 - \hat{Z})\hat{\mu} = .07c + .465 \).

Answer: C

9. Since each insured has a Poisson claim number, \( E[N|\Theta] = \text{Var}[N|\Theta] = \Theta \).

We use the semiparametric Buhlmann credibility estimate for one observation, \( Z = \frac{1}{1+\hat{v}} \).

\( \hat{v} \) is the estimate of \( v = E[\text{Var}[X|\Theta]] = E[\Theta] \), whose estimate from the data is
\[
\hat{v} = \frac{\text{Var}(X) + 30(1)+10(2)+5(3)+5(4)}{450(0)+30(1)+10(2)+5(3)+5(4)} = .17 = \hat{v}.
\]

We can also estimate \( \text{Var}[X] \) from the data
\[
\text{Var}(X) = \frac{\text{Var}(X) + 400 + 300 + 100 + 50}{400 + 300 + 100 + 50} = 86.08.
\]

Since \( \text{Var}[X] = \text{Var}[E[X|\Theta]] + E[\text{Var}[X|\Theta]] \), we can estimate \( a = \text{Var}[E[X|\Theta]] \) as the estimate of \( \text{Var}[X] - E[\text{Var}[X|\Theta]] \). This estimate is \( .362 - .17 = .192 = \hat{a} \).

Then \( \hat{k} = \frac{\hat{v}}{\hat{a}} = \frac{.17}{.192} = .89 \) and \( \hat{Z} = \frac{1}{1+\hat{k}} = \frac{1}{1.89} = .53 \).

10. \( \bar{X}_1 = \frac{(40)(200)+(50)(220)}{40+50} = 211.11 \), \( m_1 = 90 \)
\( \bar{X}_2 = \frac{(100)(200)+(120)(200)+(120)(150)}{100+120+120} = 182.35 \), \( m_2 = 340 \)
\( \bar{X}_3 = \frac{(50)(200)+(60)(250)}{50+60} = 227.27 \), \( m_3 = 110 \)
\( \hat{\mu} = X = \frac{(40)(200)+(50)(220)+(100)(200)+(120)(200)+(120)(150)+(50)(200)+(60)(250)}{90+340+110} = 196.30 \)
\( m = 540 \).

\( \hat{v}_1 = \frac{1}{3-1} \cdot [40(200 - 211.11)^2 + 50(220 - 211.11)^2] = 8888.89 \),
\( \hat{v}_2 = \frac{1}{3-1} \cdot [100(200 - 182.35)^2 + 120(200 - 182.35)^2 + 120(150 - 182.35)^2] = 97,058.82 \),
\( \hat{v}_3 = \frac{1}{3-1} \cdot [50(200 - 227.27)^2 + 60(250 - 227.27)^2] = 68,181.82 \)
\( \hat{v} = \frac{\hat{v}_1+\hat{v}_2+\hat{v}_3}{1+2+1} = 67,797.1 \).

\( \hat{a} = \frac{1}{m-\frac{1}{n}} \cdot \left[ \sum \frac{m_i(X_i - \bar{X})^2}{\hat{v}(r-1)} \right] \)
\( = \frac{1}{540-\frac{1}{10}(90^2+340^2+110^2)} \times \left[ [90(211.11 - 196.3)^2 + 340(182.35 - 196.3)^2 + 110(227.27 - 196.3)^2] - 67,797.1(2) \right] \)
\( = 193.5 \).

The credibility premium for policyholder 2 is
\( \hat{Z}\bar{X}_2 + (1 - \hat{Z})\hat{\mu} = (\frac{340}{340+67.797.1}) \cdot (182.35) + (1 - \frac{340}{340+67.797.1})(196.3) = 189.4 \).
11. Hypothetical mean is \( E(X|\theta) = \frac{\theta}{2} \).
Process variance is \( Var(X|\theta) = \frac{\theta^2}{12} \).
Expected hypothetical mean is \( \mu = E[X] = E[E(X|\theta)] = E\left(\frac{\theta}{2}\right) = \frac{1}{2} E[\theta] \),
Expected process variance = \( v = E[Var(X|\theta)] = E\left[\frac{\theta^2}{12}\right] = \frac{1}{12} E[\theta^2] \).
Variance of hypothetical mean = \( a = Var[E(X|\theta)] = Var\left(\frac{\theta}{2}\right) = \frac{1}{4} Var(\theta) = \frac{1}{4} [E(\theta^2) - (E(\theta))^2] \).

From the sample, we can estimate \( E(X) \) as \( \bar{X} = 2 \), so this is also the estimate of \( \frac{1}{2} E[\theta] \).
The estimate of \( E[\theta] \) is 4.
From the sample we can estimate \( Var(X) \) using the unbiased sample estimate,
\[
\frac{1}{n-1} \sum (X_i - \bar{X})^2 = \frac{1}{99} [\sum X_i^2 - 100 \bar{X}^2] = \frac{1}{99} [600 - 100(2^2)] = 2.02 .
\]
But \( Var(X) = a + v = \frac{1}{12} E[\theta^2] + \frac{1}{4} [E(\theta^2) - (E(\theta))^2] = \frac{1}{4} E[\theta^2] - \frac{1}{4} (E(\theta))^2 \).
Using the estimated variance of \( X \) and the estimated mean of \( \theta \), we have
\[
2.02 = \frac{1}{3} E[\theta^2] - \frac{1}{4} (4^2) ,
\]
so that the estimate of \( E[\theta^2] \) is 18.06 .
Then, \( v = \frac{1}{12} E[\theta^2] \) is estimated to be 1.505 , and
\( a = \frac{1}{4} [E(\theta^2) - (E(\theta))^2] \) is estimated to be .515 .

The estimate of losses in the 4th year is \( \hat{Z} Y + (1 - \hat{Z})\hat{\mu} \)
where \( \hat{Z} = \frac{3}{3 + \frac{a}{v}} = \frac{3 + \frac{1800}{45}}{3 + \frac{1}{3 + \frac{1800}{45}}} \right) = .507 \), and \( \hat{\mu} = \bar{X} = 2 \),
so that \( \hat{Z} Y + (1 - \hat{Z})\hat{\mu} = (.507)(\frac{1}{3}) + (.493)(2) = 1.66 \).

12. Under the nonparametric empirical Bayes method applied to the Buhlmann credibility model,
the estimated expected process variance is \( \hat{\sigma} = \frac{1}{r(n-1)} \sum_{i=1}^{r} \sum_{j=1}^{n} (X_{ij} - X_i)^2 \),
where \( r \) is the number of policyholders and \( n \) is the number of exposures per policyholder,
and the estimated variance of the hypothetical means: \( \hat{a} = \frac{1}{r-1} \sum_{i=1}^{r} (X_i - \bar{X})^2 - \frac{\hat{\sigma}}{n} \).

The estimated credibility factor for each policyholder is \( \hat{Z} = \frac{n}{n+\frac{a}{v}} \).
In this problem we have \( r = 4 \) policyholders and \( n = 7 \) exposure periods (years) per policyholder. From the given values we get
\( \hat{v} = \frac{1}{(4)(7-1)} \cdot (33.60) = 1.4 \) and
\( \hat{a} = \frac{1}{4-1} \cdot (3.30) - \frac{1.4}{7} = .90 . \) Then, \( \hat{Z} = \frac{7}{7+\frac{7.30}{.90}} = .82 . \) Answer: D

13. "Credibility" refers to the factor \( Z \) found in the semiparametric empirical Bayes approach.
The following comments review semiparametric estimation. The model for the portfolio may have a parametric distribution for \( X \) given \( \Theta = \theta \), but an unspecified non-parametric distribution for \( \Theta \). In this case, we may be able to use the fact that relationships linking \( \mu(\Theta) = E[X|\Theta] \) and \( v(\Theta) = Var[X|\Theta] \) and the fact that \( Var[X] = v + a \) in order to get estimates for \( \mu, v \) and \( a \) to use in the credibility premium formulation.
In this problem (the typical example) the conditional distribution of claim number \( X \) given \( \Theta \) is Poisson with parameter \( \Theta \). Then \( E[X|\Theta] = Var[X|\Theta] = \Theta \), so that 
\[
\mu = E[E[X|\Theta]] = E[Var[X|\Theta]] = v .
\]
We then use \( \bar{X} \) to estimate \( \mu \), and also use this as the estimate of \( v \). We find the (unbiased) sample variance of the \( X \)'s and set that equal to \( v + a \), so that the estimate of \( a \) is 
\[
V\hat{\text{ar}}[X] - \hat{v} = V\hat{\text{ar}}[X] - \bar{X}. \]
Then, as usual 
\[
Z = \frac{n}{\frac{n}{a} + \frac{a}{\bar{v}}}.
\]
From the given data, we have 
\[
\hat{\bar{v}} = \bar{X} = \frac{(54)(0) + (33)(1) + (10)(2) + (3)(3) + (1)(4)}{100} = 0.63 ,
\]
and 
\[
s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{1}{100}[(54)(0 - 0.63)^2 + (33)(1 - 0.63)^2 + (10)(2 - 0.63)^2 + (3)(3 - 0.63)^2 + (1)(4 - 0.63)^2] = 0.680 .
\]
Then 
\[
a = s^2 - v = 0.05 \rightarrow k = \frac{v}{\hat{v}} = \frac{0.63}{0.05}
\]
We are asked for the "credibility of one year’s experience for a single driver". This is the value of \( Z \) when \( n = 1 \) (one driver's experience for one year). 
\[
\bar{Z} = \frac{1}{1 + \frac{0.05}{0.63}} = 0.73 . \quad \text{Answer: E}
\]

14. The semiparametric estimate is 
\[
\hat{Z} = (\hat{\bar{Y}} - 1)(\hat{\mu}) , \quad \text{where} \quad \hat{\mu} = \frac{1}{N}\sum_i X_i = \frac{(50)}{50} = .1 , \quad \text{and} \quad \hat{\bar{Y}} = \text{sample mean of the claims for the risk being considered}. \]
In this case there is a single observation of \( Y \) (the number of robberies in one year) and \( Y = 0 \), since there were no robberies in the year for the store being considered. Then 
\[
\bar{Z} = \frac{n}{\frac{n}{\hat{a}} + \frac{\hat{a}}{\bar{v}}} = .768 , \quad \text{and finally, the semiparametric estimate is} \quad \hat{Z} = \frac{\bar{Y} - 1}{1 + \frac{0.06}{0.768}} = 0.0232 . \quad \text{Answer: B}
\]

15. For empirical Bayes estimation in the equal sample size case, the estimated credibility factor for each \( i \) is 
\[
\hat{Z}_i = \frac{n}{\frac{n}{\hat{a}} + \frac{\hat{a}}{\hat{v}_i}} . \quad \text{In this formulation,} \quad n = 2 \quad \text{is the number of observations for each group (or policyholder, or sample), and} \quad \hat{v}_i = \frac{1}{r(n-1)}\sum_{j=1}^{n}\sum_{i=1}^{r}(X_{ij} - \bar{X_i})^2 = \frac{1}{r}\sum_{j=1}^{n}\hat{v}_i , \quad \text{where} \quad \hat{v}_i = \frac{1}{n-1}\sum_{j=1}^{n}(X_{ij} - \bar{X_i})^2 , \quad \text{and} \quad r \quad \text{is the number of policyholders}.
\]
Also, 
\[
\hat{a} = \frac{1}{r-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 - \frac{\hat{v}}{n}.
\]
In this example \( r = 2 \), \( n = 2 \), \( X_{11} = 1 \), \( X_{12} = 9 \) (Class A death times), \( X_{21} = 2 \), \( X_{22} = 4 \) (Class B death times), \( X_1 = 5 \), \( X_2 = 3 \), \( \bar{X} = 4 \).
\[
\hat{v} = \frac{1}{2(1)} \cdot [(1 - 5)^2 + (9 - 5)^2 + (2 - 3)^2 + (4 - 3)^2] = 17 ,
\]
\[
\hat{a} = \frac{1}{1} \cdot [(5 - 4)^2 + (3 - 4)^2] - \frac{17}{2} = -6.5 < 0 .
\]
When \( \hat{a} < 0 \), the credibility factor \( \bar{Z} \) is set equal to 0. \quad \text{Answer: A}
16. Since the numbers of exposures differs among exposure periods, we use non-parametric empirical Bayes estimation for the Buhlmann-Straub model. Under the Buhlmann-Straub model, there are \( n_i \) exposure periods for policy holder (group) \( i \), for \( i = 1, 2, \ldots, r \). For policyholder \( i \) and exposure period (year) \( j \), there are \( m_{ij} \) exposure units (8 members for group 1 in year 1, etc.), and \( X_{ij} \) represents the observed average claim per exposure unit (member) (for "cell" \( i, j \)) (96 for group 1 in year 1). In this case, \( r = 2 \) policyholders (groups 1 and 2), and \( n_1 = n_2 = 3 \) exposure periods (years) for each policyholder. The usual unbiased estimates that are used for the structural parameters are 

\[
\hat{\mu} = \bar{X} \quad \text{and} \quad \hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^n (X_{ij} - \bar{X})^2
\]

(estimated mean of the process variances), and 

\[
\hat{\alpha} = \frac{1}{m \sum_{i=1}^r m_i^2} \cdot \left[ \sum_{i=1}^r m_i (\bar{X}_i - \bar{X})^2 - \hat{\sigma}^2 (r-1) \right]
\]

(estimated variance of the hyp. means).

From the given values we get 

\[
\hat{\sigma} = \frac{1}{(3-1)+(3-1)} \cdot (2020) = 505.6
\]

\[
m = m_1 + m_2 = (8 + 12 + 5) + (25 + 30 + 20) = 100, \quad \text{and then}
\]

\[
\hat{\alpha} = \frac{1}{100} \left[ \frac{1}{(25^2 + 75^2)} \cdot [4800 - 505(1)] \right] = 114.5.
\]

The estimated credibility factor for group \( i \) is 

\[
\hat{Z}_i = \frac{m_i}{m_i + \hat{\alpha}} = \frac{m_i}{m_i + \hat{\alpha}};
\]

\[
\hat{Z}_1 = \frac{25}{25 + \hat{\alpha}} = \frac{25}{25 + 100} = .85, \quad \hat{Z}_2 = \frac{75}{75 + 100} = .94.
\]

An alternative to the unbiased estimate \( \hat{\mu} = \bar{X} \) just described for the Buhlmann-Straub model is the credibility-weighted average estimate of \( \hat{\mu} \), which is found by first estimating \( \hat{Z}_1, \hat{Z}_2, \ldots, \hat{Z}_r \) in the way just described for the Buhlmann-Straub model, and then 

\[
\hat{\mu} = \frac{\sum_{i=1}^r \hat{Z}_i \bar{X}_i}{\sum_{i=1}^r \hat{Z}_i}.
\]

If we use this estimate of \( \mu \) to find the credibility premiums for groups 1, ..., \( r \), and then calculate total credibility premiums for past exposures, that total will equal to the actual total past claims. This is what is meant by using the method that "preserves total losses". In this case,

\[
\hat{\mu} = \frac{\sum_{i=1}^r \hat{Z}_i \bar{X}_i}{\sum_{i=1}^r \hat{Z}_i} = \frac{(.85)(97) + (.94)(113)}{.85 + .94} = 105.4.
\]

Credibility premium for the group 1 is 

\[
\hat{Z}_1 \bar{X}_1 + (1 - \hat{Z}_1)\hat{\mu} = (.85)(97) + (.15)(105.4) = 98.3.
\]

Answer: A

17. For a risk with sample mean \( \bar{Y} \) of claims for the past \( n \) months, the semiparametric empirical Bayes estimate of the expected number of claims next month is 

\[
\hat{Z} \bar{Y} + (1 - \hat{Z})\hat{\mu}, \quad \text{where} \quad \hat{Z} = \frac{n}{n+k}.
\]

We are given that with \( n = 1 \) and \( \bar{Y} = 0 \), we have

\[
\hat{Z} \cdot (0) + (1 - \hat{Z})\hat{\mu} = (1 - \frac{1}{1+k})\hat{\mu} = \frac{k}{1+k} \cdot \hat{\mu} = \frac{1}{30} \quad \text{and with} \quad n = 2 \quad \text{and} \quad \bar{Y} = 0, \text{we have}
\]

\[
\hat{Z} \cdot (0) + (1 - \hat{Z})\hat{\mu} = (1 - \frac{2}{2+k})\hat{\mu} = \frac{k}{2+k} \cdot \hat{\mu} = \frac{1}{55}.
\]

It follows that

\[
(\frac{k}{1+k} \cdot \hat{\mu})/(\frac{k}{2+k} \cdot \hat{\mu}) = \frac{2+k}{1+k} = \frac{55}{30}, \quad \text{so that} \quad \hat{k} = .2. \quad \text{Then, with} \quad n = 3 \quad \text{and} \quad \bar{Y} = 0, \text{we have}
\]

\[
\hat{Z} \cdot (0) + (1 - \hat{Z})\hat{\mu} = (1 - \frac{3}{3+k})\hat{\mu} = \frac{k}{3+k} \cdot \hat{\mu} = (\frac{2+k}{3+k})(\frac{k}{2+k} \cdot \hat{\mu}) = (\frac{3}{3+k})(\frac{2+k}{3+k})(\frac{1}{55}) = \frac{1}{80}.
\]

Answer: C
18. \[ X_1 = \frac{(40)(200) + (50)(220)}{40+50} = 211.11 \text{ , } m_1 = 90 \]
\[ X_2 = \frac{(100)(200) + (120)(200) + (120)(150)}{100+120+120} = 182.35 \text{ , } m_2 = 340 \]
\[ X_3 = \frac{(50)(200) + (60)(250)}{50+60} = 227.27 \text{ , } m_3 = 110 \]
\[ \mu = \bar{X} = \frac{(40)(200) + (50)(220) + (100)(200) + (120)(200) + (120)(150) + (50)(200) + (60)(250)}{90+340+110} = 196.30 \]
\[ m = 540. \]
\[ \bar{v}_1 = \frac{1}{2} \cdot [40(200 - 211.11)^2 + 50(220 - 211.11)^2] = 8888.89, \]
\[ \bar{v}_2 = \frac{1}{3} \cdot [100(200 - 182.35)^2 + 120(200 - 182.35)^2 + 120(150 - 182.35)^2] = 97,058.82, \]
\[ \bar{v}_3 = \frac{1}{1+2+1} \cdot [50(200 - 227.27)^2 + 60(250 - 227.27)^2] = 68,181.82 \]
\[ \nu = \frac{\bar{v}_1 + 2\bar{v}_2 + \bar{v}_3}{1+2+1} = 67,797.1. \]
\[ \hat{a} = \frac{1}{m - \frac{1}{m} \sum_{i=1}^{r} m_i^2} \cdot \left[ \sum_{i=1}^{r} m_i (X_i - \bar{X})^2 - \bar{v}(r - 1) \right] \]
\[ = \frac{1}{540} \cdot \frac{1}{50} \cdot \left(90^2 + 340^2 + 110^2 \right) \times \left[ 90(211.11 - 196.3)^2 + 340(182.35 - 196.3)^2 + 110(227.27 - 196.3)^2 \right] - 67,797.1(2) \]
\[ = 193.5. \]

The credibility premium for policyholder 2 is
\[ \hat{Z}X_2 + (1 - \hat{Z})\hat{\mu} = \left(\frac{340}{340 + \frac{500}{103.5}}\right) \cdot (182.35) + \left(1 - \frac{340}{340 + \frac{500}{103.5}}\right)(196.3) = 189.4. \]

19. \( X \) is the random variable for annual loss. We are given that the conditional distribution of \( X \) given \( \theta \) is exponential with a mean of \( \theta \), where \( \theta \) has an unspecified distribution. Therefore, the hypothetical mean is \( HM = E[X|\theta] = \theta \) and the process variance is \( PV = Var[X|\theta] = \theta^2 \).

The expected hypothetical mean is \( \mu = EHM = E[E[X|\theta]] = E[\theta] = E[X] \)

(using the double expectation rule \( E[E[X|\theta]] = \theta \)).

The expected process variance is \( v = EPV = E[Var[X|\theta]] = E[\theta^2] \).

The variance of the hypothetical mean is
\[ a = VHM = Var[E[X|\theta]] = Var[\theta] = E[\theta^2] - (E[\theta])^2. \]

From this we see that \( v - a = (E[\theta])^2 = \mu^2. \)

In general, \( Var[X] = E[Var[X|\theta]] + Var[E[X|\theta]] = v + a. \)

From the data set we can use empirical estimation to estimate \( E[X] \):
\[ \hat{\mu} = \left(\frac{500}{1000}\right) \left(\frac{9+100}{2}\right) + \left(\frac{250}{1000}\right) \left(\frac{100+200}{2}\right) + \left(\frac{150}{1000}\right) \left(\frac{200+500}{2}\right) + \left(\frac{60}{1000}\right) \left(\frac{500+1000}{2}\right) + \left(\frac{40}{1000}\right) \left(\frac{1000+2000}{2}\right) = 220. \]

From the data set we can use empirical estimation to estimate \( E[X^2] \):
\[ \left(\frac{500}{1000}\right) \left(\frac{100^2+0^2}{3(100-0)}\right) + \left(\frac{250}{1000}\right) \left(\frac{200^2-100^2}{3(200-100)}\right) + \left(\frac{150}{1000}\right) \left(\frac{500^2-200^2}{3(500-200)}\right) + \left(\frac{60}{1000}\right) \left(\frac{1000^2-500^2}{3(1000-500)}\right) + \left(\frac{40}{1000}\right) \left(\frac{2000^2-1000^2}{3(2000-1000)}\right) = 155,333. \]
The empirical estimate of the variance of $X$ is $\hat{\sigma}^2 = 155.333 - (220)^2 = 106,933$.

In the semiparametric empirical Bayes credibility model, we use the empirical estimate of $E[X]$ for $\mu$, so that $\hat{\mu} = 220$. We also know that $Var\{X\} = v + a$, so using the empirical estimate of $Var\{X\}$ gives $106,933 = \hat{\nu} + \hat\alpha$. But we also know that, for this model, $v - \alpha = \mu^2$, so using our sample estimate of $\mu$, we have $\hat{\nu} - \hat{\alpha} = (220)^2$. We can then solve the two equations $\hat{\nu} + \hat{\alpha} = 106,933$ and $\hat{\nu} - \hat{\alpha} = 48,400$ to get $\hat{\nu} = 77,667$ and $\hat{\alpha} = 29,267$.

We can now find the estimated loss in the 3rd year for a policy that had losses of $Y_1 = 150$ in the first year and $Y_2 = 0$ in the second year. The estimate is $\hat{Z}Y + (1 - \hat{Z})\hat{\mu}$, where $\hat{Z} = \frac{\hat{\nu}}{2 + \hat{\nu}} = .4298$ and $\hat{\mu} = 220$. The credibility premium is $(.4298)(75) + (.5702)(220) = 158$.

20. $\bar{X}_1 = \frac{(40)(200)+(50)(220)+(40)(250)}{40+50+40} = 223.08$, $m_1 = 130$, $n_1 = 3$

$\bar{X}_2 = \frac{(100)(200)+(120)(200)+(120)(150)}{100+120+120} = 182.35$, $m_2 = 340$

$\bar{X}_3 = \frac{(50)(200)+(60)(250)}{50+60} = 227.27$, $m_3 = 110$

$\hat{\mu} = \bar{X} = \frac{(40)(200)+(50)(220)+(40)(250)+(100)(200)+(120)(200)+(120)(150)+(50)(200)+(60)(250)}{130+340+110} = 200$, $m = 580$.

$\hat{\nu}_1 = \frac{1}{3-1} \cdot [40(200 - 223.08)^2 + 50(220 - 223.08)^2 + 40(250 - 223.08)^2] = 25,384.62$,

$\hat{\nu}_2 = \frac{1}{3-1} \cdot [100(200 - 182.35)^2 + 120(200 - 182.35)^2 + 120(150 - 182.35)^2] = 97,058.82$,

$\hat{\nu}_3 = \frac{1}{3-1} \cdot [50(200 - 227.27)^2 + 60(250 - 227.27)^2] = 68,181.82$,

$\hat{\nu} = \frac{\hat{\nu}_i + \hat{\nu}_2 + \hat{\nu}_3}{2+2+1} = 62,613.7$.

$\hat{\alpha} = \frac{1}{m-\frac{n}{m}} \cdot \left[ \sum_{i=1}^{r} m_i (\bar{X}_i - \bar{X})^2 - \hat{\nu}(r-1) \right]$

$= \frac{1}{580-\frac{340}{340+340+110}} \times \left[ 130(223.08 - 200)^2 + 340(182.35 - 200)^2 + 110(227.27 - 200)^2 \right] - 62,613.7(2)$

$= 398.4$.

The credibility premium for policyholder 2 is

$\hat{Z}\bar{X}_2 + (1 - \hat{Z})\hat{\mu} = \left( \frac{340}{340+\frac{340}{208.4}} \right) \cdot (182.35) + (1 - \frac{340}{340+\frac{340}{208.4}})(200) = 187.9$. 
21. Hypothetical mean is \( E(X|\theta) = \int_0^\theta x \cdot \frac{2x}{\theta^2} \, dx = \frac{2\theta}{3} \).

Process variance is \( Var(X|\theta) = E(X^2|\theta) - [E(X|\theta)]^2 \).
\[
E(X^2|\theta) = \int_0^\theta x^2 \cdot \frac{2x}{\theta^2} \, dx = \frac{\theta^2}{2} \\
\Rightarrow \text{process variance} = Var(X|\theta) = \frac{\theta^2}{2} - \left( \frac{2\theta}{3} \right)^2 = \frac{\theta^2}{18}.
\]

Expected hypothetical mean is \( \mu = E[X] = E[E(X|\theta)] = E\left( \frac{2\theta}{3} \right) = \frac{2}{3} E[\theta] \),

Expected process variance \( = v = E[Var(X|\theta)] = E\left[ \frac{\theta^2}{18} \right] = \frac{1}{18} E[\theta^2] \).

Variance of hypothetical mean \( = a = Var[E(X|\theta)] = Var\left( \frac{2\theta}{3} \right) = \frac{4}{9} Var(\theta) = \frac{4}{9}[E(\theta^2) - (E(\theta))^2] \).

From the sample, we can estimate \( E(X) \) as \( \bar{X} = 2 \), so this is also the estimate of \( \frac{2}{3} E[\theta] \).

The estimate of \( E[\theta] \) is 3.

From the sample we can estimate \( Var(X) \) using the unbiased sample estimate,
\[
\frac{1}{99}[\Sigma X_i^2 - 100\bar{X}^2] = \frac{1}{99}[600 - 100(2^2)] = 2.02.
\]
But \( Var(X) = a + v = \frac{1}{18} E[\theta^2] + \frac{4}{9} [E(\theta^2) - (E(\theta))^2] = \frac{1}{2} E[\theta^2] - \frac{4}{9} (E(\theta))^2 \).

Using the estimated variance of \( X \) and the estimated mean of \( \theta \), we have \( 2.02 = \frac{1}{2} E[\theta^2] - \frac{4}{9} (3^2) \), so that the estimate of \( E[\theta^2] \) is 12.04.

Then, \( v = \frac{1}{18} E[\theta^2] \) is estimated to be 0.669, and
\[
a = \frac{4}{9} [E(\theta^2) - (E(\theta))^2] \text{ is estimated to be } 1.35.
\]

The estimate of losses in the 4th year is \( \hat{Z} \hat{Y} + (1 - \hat{Z})\hat{\mu} \)
where \( \hat{Z} = \frac{3}{3 + \hat{\sigma}} = \frac{3}{3 + .635} = .858 \), and \( \hat{\mu} = \bar{X} = 2 \),
so that \( \hat{Z} \hat{Y} + (1 - \hat{Z})\hat{\mu} = (.858)(\frac{4}{3}) + (.142)(2) = 1.43 \).

22. HM = \( E[S|\theta] = E[N|\theta] \cdot E[Y|\theta] = \theta^2 \),
PV = \( Var[S|\theta] = E[N|\theta] \cdot E[Y^2|\theta] = \theta(2\theta^2) = 2\theta^3 \).
EHM = \( \mu = E[\theta^2] = \int_0^\theta \theta^2 \, d\theta = \frac{7}{3} \)
EPV = \( v = E[2\theta^3] = \int_0^\theta 2\theta^3 \, d\theta = \frac{15}{2} \)
VHM = \( a = Var[\theta^2] = E[\theta^4] - (E[\theta^2])^2 = \frac{31}{5} - \left( \frac{7}{3} \right)^2 = \frac{34}{15} \).
\[
n = 1 \rightarrow Z = \frac{1}{1 + \frac{152}{31.45}} = .0915.
\]

Credibility premium is \( .0915S + .9085\left( \frac{7}{3} \right) = .0915S + 2.12 \).