BRIGHAM YOUNG UNIVERSITY

STAT 446 ADVANCED SHORT-TERM ACTUARIAL MODELING

ASSIGNMENT PROBLEMS BY LECTURE

Contents

Lecture 1: Policy Modification **1**

Lecture 2: Aggregate Loss Models **4**

Lecture 3: Parametric Models: Estimation & Selection **7**

Lecture 4: Severity Models **11**

Lecture 5: Credibility **15**

Lecture 6: Ratemaking **21**

Lecture 7: Outstanding Claims Reserves **24**

Lecture 1: Policy Modification

1.1 The ground-up loss X on an insurance policy has the following distribution:

x	$\Pr(X = x)$
200	0.4
800	0.3
2000	0.2
5000	0.1

(a) Calculate $\mathbb{E}[X]$ and the loss elimination ratio for an ordinary deductible of 500. [1220; 0.3115]

(b) For an ordinary deductible of 500, calculate the expected payment per loss and the expected payment per payment. [840; 1400]

(c) For a franchise deductible of 500, calculate the expected payment per loss and the expected payment per payment. [1140; 1900]

(d) Determine the ordinary deductible that would produce a loss elimination ratio of 40%. [680]

(e) The policy is instead sold with a linearly disappearing deductible: no payment is made on a loss of 500 or less, full payment is made on a loss of 2500 or more, and the payment increases linearly between these two amounts. Calculate the expected payment per loss. [987.5]

(f) Next year all losses are expected to increase by 10%, while the ordinary deductible remains 500.

(i) Calculate the percentage increase in the expected payment per loss. [13.57%]

(ii) Explain why your answer to (i) is larger than 10%.

1.2 Ground-up losses follow a Pareto distribution with $\alpha = 3$ and $\theta = 2000$. An insurance policy has an ordinary deductible of 250, a maximum covered loss of 5000, and a coinsurance factor of 80%.

(a) Determine the maximum possible payment on a single loss. [3800]

(b) Calculate the expected payment per loss. [566.79]

(c) Calculate the expected payment per payment. [807.02]

(d) Next year, all losses increase uniformly by 6%, and the deductible, maximum covered loss, and coinsurance factor do not change.

(i) Calculate the expected payment per loss next year. [603.35]

(ii) Calculate the percentage increase in the expected payment per loss, and explain why it is not exactly 6%. [6.45%]

(e) The insurer also sells policies with no deductible and no coinsurance, using a basic limit of 5000. Assuming that claim frequency does not depend on the policy limit, calculate the increased limit factor for a policy limit of 10,000. [1.0586]

1.3 An insurer has the following data on 1,000 ground-up losses:

Ground-Up Loss	Number of Losses	Total Ground-Up Losses
0–1000	500	210,000
1000–2500	300	510,000
2500–5000	120	420,000
5000–10,000	60	390,000
10,000+	20	380,000
Total	1,000	1,910,000

Assume that the number of ground-up losses does not depend on the deductible or the policy limit. Use the empirical distribution of these data to answer the following.

(a) Calculate the expected ground-up loss. [1910]

(b) For an ordinary deductible of 1000, calculate

(i) the expected payment per loss, [1200]

(ii) the expected payment per payment, and [2400]

(iii) the loss elimination ratio. [0.3717]

(c) The insurer sets its base rate using a deductible of 1000. Determine the indicated deductible relativity for

(i) a policy with no deductible, and [1.5917]

(ii) a policy with a deductible of 2500. [0.575]

(d) The insurer buys excess-of-loss reinsurance with a retention of 5000 per loss on policies with no deductible. Calculate the reinsurer's expected cost per loss. [370]

(e) Using a basic limit of 5000 for policies with no deductible, calculate the increased limit factor for a policy limit of 10,000. [1.1234]

(f) Give one reason why the deductible relativities in (c) might not be appropriate in practice.

1.4 The number of losses N^L on a policy with no deductible belongs to the $(a, b, 0)$ class with

$$\frac{\Pr(N^L = 1)}{\Pr(N^L = 0)} = 0.6 \qquad \text{and} \qquad \frac{\Pr(N^L = 2)}{\Pr(N^L = 1)} = 0.45.$$

Loss amounts follow an exponential distribution with mean 1000 and are independent of N^L.

(a) Determine a and b, and identify the distribution of N^L and its parameters. [$a = b = 0.3$; negative binomial, $r = 2$, $\beta = 3/7$]

(b) An ordinary deductible of 500 is imposed. Identify the distribution of the number of payments, N^P, and its parameters. [negative binomial, $r = 2$, $\beta^* = 0.2599$]

(c) Calculate $\Pr(N^P = 0)$, $\mathbb{E}[N^P]$, and $\mathrm{Var}(N^P)$. [0.6299; 0.5199; 0.6550]

(d) Suppose instead that N^L follows a zero-modified distribution, with the same a and b as in (a) and with $p_0^M = 0.6$. The deductible of 500 still applies.

(i) Calculate p_0^{M*}, the probability of no payments. [0.7098]

(ii) Calculate $\Pr(N^P = 1)$. [0.2039]

(iii) Calculate $\mathbb{E}[N^P]$. [0.4078]

1.5 For a collective risk model, you are given:

(i) The number of losses follows a zero-modified Poisson distribution with $\lambda = 0.5$, and $p_0^M = 0.3$.

(ii) Loss size follows a single-parameter Pareto distribution with $\theta = 200$ and $\alpha = 1.5$, so $\Pr(X > x) = (200/x)^{1.5}$ for $x \geq 200$.

(iii) An insurance coverage has a deductible of 500.

(a) Calculate the probability of no paid claim. [0.7886]

(b) Calculate the probability of one paid claim. [0.1983]

(c) Calculate the mean of the number of payments. [0.2250]

(d) Calculate the variance of the number of payments. [0.2029]

Lecture 2: Aggregate Loss Models

2.1 A company sells three independent policies. For policy A, all claim payments are 10,000 or 20,000 with equal probabilities, and a single policy has a Poisson number of claims with mean 1. For policy B, all claim payments are 20,000 and a single policy has a Poisson number of claims with mean 0.5. For policy C, all claim payments are 40,000 and a single policy has a Poisson number of claims with mean 0.5.

(a) The aggregate claim payments of the company follow a compound Poisson distribution. Specify its primary and secondary distributions. [Poisson with $\lambda = 2$; 10,000, 20,000, 40,000 with probabilities 0.25, 0.5, 0.25]

(b) Calculate the expected value and standard deviation of the aggregate claim payments for the company. [45,000; 35,355]

(c) Calculate the probability that the aggregate claim payments are no more than 30,000. [0.4257]

2.2 An insurer has a portfolio of 400 independent policies. For each policy:

(i) The annual number of claims follows a Poisson distribution with mean 0.25.

(ii) Claim amounts follow a lognormal distribution with $\mu = 7$ and $\sigma = 1$.

(iii) Claim counts and claim amounts are independent.

Let S denote the total annual claims for the portfolio.

(a) Explain why S has a compound Poisson distribution, and state its Poisson parameter. [$\lambda = 100$]

(b) Calculate $\mathbb{E}[S]$ and the standard deviation of S. [180,804; 29,810]

(c) Using a normal approximation for S, calculate

(i) the 95% VaR of S, and [229,837]

(ii) the 95% Expected Shortfall (also called TVaR) of S. [242,293]

(d) The insurer instead approximates S with a lognormal distribution with parameters μ^* and σ^*, chosen so that its mean and variance match those of S.

(i) Show that $\sigma^* = 0.16$ to the nearest 0.01. You should calculate the value to the nearest 0.0001. [0.1638]

(ii) Calculate the 95% VaR of S using the lognormal approximation. [233,546]

(e) You are given that for a compound Poisson distribution with Poisson parameter λ,

$$\mathbb{E}\left[(S - \mathbb{E}[S])^3\right] = \lambda\,\mathbb{E}[X^3].$$

(i) Calculate the coefficient of skewness of S. [0.4482]

(ii) State with reasons which of the two approximations you would expect to give the more accurate estimate of the 95% VaR. [lognormal]

2.3 Let S denote the aggregate payments on a policy, in thousands. The number of losses is N, and the payment on each loss, after a per-loss deductible, is X (in thousands), so a loss can result in a payment of 0. You are given:

(i) N belongs to the $(a, b, 0)$ class with $\mathbb{E}[N] = 1.2$ and $\mathrm{Var}(N) = 0.96$.

(ii) X takes values in $\{0, 1, 2, \ldots\}$ and belongs to the $(a, b, 0)$ class with $\mathbb{E}[X] = 1.5$ and $\mathrm{Var}(X) = 3.75$.

(iii) N and the X's are independent.

(a) (i) Explain why N must have a binomial distribution, and determine its parameters. [$m = 6$, $q = 0.2$]

(ii) Identify the distribution of X and its parameters. [geometric, $\beta = 1.5$]

(b) Calculate $\mathbb{E}[S]$ and the standard deviation of S. [1.8; 2.5807]

(c) Determine the probability generating function of S. [$P_S(z) = \left(0.8 + \frac{0.2}{2.5 - 1.5z}\right)^6$]

(d) (i) Calculate $\Pr(S = 0)$. [0.4644]

(ii) Use the recursive formula to calculate $\Pr(S = 1)$ and $\Pr(S = 2)$. [0.1520; 0.1119]

(e) Calculate the net stop-loss premium for an aggregate deductible of

(i) 2 (that is, 2000), and [0.8808, or 880.8]

(ii) 1.5 (that is, 1500). [1.0726, or 1072.6]

2.4 For a collective risk model, you are given:

(i) The number of losses follows a binomial distribution with $m = 5$, and $q = 0.2$.

(ii) Loss size follows a lognormal distribution with $\mu = 5$ and $\sigma = 2$.

(a) With a span of $h = 100$, discretize the severity distribution of the losses using the method of rounding. You are required to calculate all values up to a discretized amount paid of 200. [$f_0 = 0.2932$, $f_1 = 0.2089$, $f_2 = 0.1007$]

(b) Use the recursive formula to calculate the probability that the discretized aggregate losses are no more than 200. [0.6461]

2.5 Loss amounts follow a Pareto distribution with $\alpha = 3$ and $\theta = 2000$, and the number of losses follows a Poisson distribution with mean 1.5. You discretize the loss amounts using a span of $h = 500$.

(a) Using local moment matching with one moment, calculate f_0, f_1, and f_2, the probabilities at 0, 500, and 1000. [0.2800, 0.3289, 0.1553]

(b) Using the method of rounding, calculate f_0, f_1, and f_2. [0.2977, 0.3177, 0.1516]

(c) Which of the two discretized distributions has the same mean as the loss amounts? Explain.

(d) Using the distribution from (a) and the recursive formula, calculate $\Pr(S = 0)$ and $\Pr(S = 500)$, where S is the discretized aggregate loss. [0.3396; 0.1675]

(e) Using your answer to (d), calculate the net stop-loss premium for an aggregate deductible of 1000. [923.36]

2.6 For a collective risk model:

(i) The number of losses has a negative binomial distribution with $\beta = 0.5$ and $r = 4$.

(ii) The common distribution of the individual losses is:

x	$f_X(x)$
200	0.4
400	0.4
600	0.2

There are two policies, A and B, available to insure these losses.

(a) Policy A has an ordinary deductible of 200 on each individual loss. Calculate the expected aggregate insurance payments from Policy A. [320]

(b) Policy B has an aggregate deductible of 400. Calculate the expected aggregate insurance payments from Policy B. [420.08]

Lecture 3: Parametric Models: Estimation & Selection

3.1 You are given:

(i) Four independent lives are observed from time $t = 0$ until death.

(ii) Deaths occur at times $t = 1, 2, 3$ and 4.

(iii) The lives are assumed to be subject to the probability density function

$$f(t) = \frac{te^{-t/c}}{c^2}, \qquad t > 0.$$

(a) Calculate the maximum likelihood estimate for c. [1.25]

(b) Calculate the variance of the maximum likelihood estimator for c. [0.1953]

(c) Determine the maximum likelihood estimate for the probability that death occurs before $t = 5$ for a randomly selected life, i.e., $F(5)$. [0.9084]

(d) Use the delta method to approximate the variance of the maximum likelihood estimator of $F(5)$. [0.01074]

3.2 A random sample of 6 claims is given:

$$39 \quad 150 \quad 247 \quad 388 \quad 705 \quad 1000^+$$

where the last claim is known to be greater than 1000.

You fit the data to the following density function

$$f(x) = \frac{\alpha\, 1000^\alpha}{(1000 + x)^{\alpha+1}}, \qquad x > 0.$$

(a) Determine $\hat{\alpha}$, the maximum likelihood estimate of α. [2.5597]

(b) Calculate the approximated variance of $\hat{\alpha}$. [1.3104]

(c) Construct an approximate 95% normal confidence interval for α. [(0.316, 4.803)]

(d) Construct an approximate 95% likelihood-based confidence interval for α, and compare it with the interval in (c). [(0.918, 5.502)]

3.3 Claim size X follows an inverse Pareto distribution, with $F(x) = \left(\frac{x}{x+\theta}\right)^\tau$ for $x > 0$. The maximum likelihood estimates of the two parameters τ and θ are $(2, 100)$, and the covariance matrix of these estimators is

$$\begin{bmatrix} 0.01 & -0.1 \\ -0.1 & 5 \end{bmatrix}.$$

Using the delta method, estimate the variance of $\Pr(X > 500)$. [0.0001285]

3.4 You are given the following random observations:

$$0.1 \quad 0.4 \quad 0.8 \quad 0.8 \quad 0.9$$

You test whether the sample comes from a distribution with probability density function:

$$f(x) = \frac{1+2x}{2}, \qquad 0 \le x \le 1.$$

Calculate the Kolmogorov-Smirnov test statistic. [0.320]

3.5 You are given the following:

- The observed number of claims for a group of 1000 risks has been recorded as follows:

Number of Claims	Number of Risks
0	729
1	242
2	29
3 or more	0

- The null hypothesis, H_0, is that the number of claims per risk follows a Poisson distribution.
- You use a chi-square test to measure the fit of a Poisson distribution. The minimum expected number of observations in any group should be 5. The maximum possible number of groups should be used.
- The estimated Poisson parameter is $\lambda = 0.3055$ from the sample.

Determine the chi-square statistic. [3.56]

3.6 You fit a Weibull distribution to a sample of 20 claim amounts $x_1, \ldots, x_{20}$ and use the likelihood ratio test to test the hypothesis

$$H_0 : \tau = 2 \quad \text{vs.} \quad H_1 : \tau \neq 2.$$

You are given that

$$\sum_{i=1}^{20} \ln x_i = 73.6177, \qquad \sum_{i=1}^{20} x_i^2 = 87{,}266,$$

and at the maximum likelihood estimate, the loglikelihood is -98.443.

Determine the result of the test. [test statistic 3.392 with 1 df: reject at 10%, do not reject at 5%]

Additional Practice Problems (not assigned)

3.7 A random sample of 5 claims is given:

$$51 \quad 120 \quad 240 \quad 265 \quad 310.$$

You fit the data to the following density function

$$f(x) = \frac{1}{\sqrt{2\pi\theta}}\, e^{-\frac{(x-250)^2}{2\theta}}, \qquad x \in \mathbb{R}.$$

(a) Determine $\hat{\theta}$, the maximum likelihood estimate of θ. [12,085.2]

(b) Calculate the approximated variance of $\hat{\theta}$. [5.842×10^7]

(c) Determine an approximate 95% confidence interval for θ. [$(-2896,\ 27{,}066)$]

3.8 You are given the following data for claim frequency:

Number of Claims	Number of Policies
0	134
1	45
2	15
3	5
4	1
5+	0

The data are fitted to a negative binomial distribution with $r = 2$ and β unknown.

(a) Determine the maximum likelihood estimate of p_0, the probability of 0 claims. [0.6556]

(b) Use the delta method to approximate the variance of the maximum likelihood estimator of p_0. [0.000818]

(c) Determine an approximate 95% confidence interval for p_0, using the obtained estimates. [(0.5996, 0.7117)]

3.9 You use the Kolmogorov-Smirnov test to assess the fit of the natural logarithms of $n = 200$ losses to a distribution with distribution function F^*. You are given:

(i) The largest value of $|F_n(x) - F^*(x)|$ occurs for some x between 4.26 and 4.42.

(ii) The distribution function values for observed x are

Observed x	$F^*(x)$	$F_n(x-)$	$F_n(x)$
4.26	0.584	0.505	0.510
4.30	0.599	0.510	0.515
4.35	0.613	0.515	0.520
4.36	0.621	0.520	0.525
4.39	0.636	0.525	0.530
4.42	0.638	0.530	0.535

(iii) Critical values for the Kolmogorov-Smirnov test are

Level of Significance	0.10	0.05	0.02	0.01
Critical Value	$\frac{1.22}{\sqrt{n}}$	$\frac{1.36}{\sqrt{n}}$	$\frac{1.52}{\sqrt{n}}$	$\frac{1.63}{\sqrt{n}}$

Determine the results of the test for the given significance levels. [$D = 0.111$: reject at 0.10, 0.05, and 0.02; do not reject at 0.01]

3.10 You are given the following:

- A random sample of 1000 observations from a loss distribution has been grouped into five intervals as follows:

Interval	Number of Observations
$[0, 3)$	180
$[3, 7.5)$	180
$[7.5, 15)$	235
$[15, 40)$	255
$[40, \infty)$	150
Total	1000

- The loss distribution is believed to be a two-parameter Pareto distribution and the estimated parameters based on the above sample are $\alpha = 3.5$ and $\theta = 50$.

You are to use the chi-square statistic to test the hypothesis, H_0, that the two-parameter Pareto distribution provides an acceptable fit.

Determine the test result. [$\chi^2 = 9.50$ with 2 df (p-value 0.009): reject H_0 at the 5% and 1% levels]

3.11 95 claim sizes are fitted to several models, with the following results:

Model	Loglikelihood
Exponential	-487.0
Inverse Exponential	-487.5
Gamma	-487.0
Inverse Gamma	-484.1
Burr	-482.0

(a) Using the Schwarz Bayesian Criterion, which model is selected? [inverse gamma]

(b) Using the Akaike Information Criterion, which model is selected? [Burr]

Lecture 4: Severity Models

4.1 An actuary has performed a longitudinal study on a group of patients and finds that the conditional hazard rate function of a patient's future lifetime is

$$h_{X|\Lambda}(x|\lambda) = 3\lambda x^2, \qquad x > 0,\ \lambda > 0,$$

where Λ is gamma distributed with $\alpha = 2$ and $\theta = 0.01$.

(a) Determine the conditional survival function of the future lifetime for a patient in the group. [$e^{-\lambda x^3}$]

(b) Calculate the probability that a randomly selected patient in the group survives more than 5 years. [0.1975]

4.2 A loss follows a spliced distribution. The density function below 500 is the same as the density function of an exponential distribution with mean $\theta = 250$. The density function of the loss beyond 500 is proportional to the density function of a Weibull distribution with parameters $\tau = 2$ and $\theta = 400$.

(a) Determine the density function of the loss. [proportionality constant $c = 0.6456$]

(b) Calculate the probability that the loss is no more than 600. [0.9320]

(c) Determine VaR of the loss at the 95% security level. [639.78]

4.3 You are given that the random variable X has probability density function

$$f(x) = (1 + 2x)e^{-2x}, \qquad x > 0.$$

(a) Calculate the mean of X. [0.75]

(b) Determine the survival function $S(x)$. [$(1 + x)e^{-2x}$]

(c) Determine the hazard rate function $h(x)$, and classify the tail of X based on $h(x)$. [$h(x) = \frac{1+2x}{1+x}$, increasing: light tail]

(d) Determine the mean excess loss function $e(d)$, and classify the tail of X based on $e(d)$. [$e(d) = \frac{0.75+0.5d}{1+d}$, decreasing: light tail]

(e) Using the limiting ratio of the survival functions, compare the relative tail weight of X to a Pareto random variable with parameters $\alpha = 3$ and $\theta = 0.25$. [limit is 0: X has a lighter tail]

4.4 Consider the Pareto distribution with distribution function

$$F_X(x) = 1 - \left(\frac{\theta}{\theta + x}\right)^{\alpha}.$$

(a) Show that this distribution is in the Fréchet MDA, using normalizing sequences $c_n = \frac{\theta}{\alpha} n^{1/\alpha}$ and $d_n = \theta\left(n^{1/\alpha} - 1\right)$.

(b) Given $\theta = 2000$ and $\alpha = 3.0$,

(i) Compare the exact probability that $M_n \leq m$ with the probability approximated by the GEV distribution, for $n = 25$ and $m = 5000$. [exact 0.5543; GEV 0.5582]

(ii) Repeat (i) for $m = 5000$ and $n = 100$. [exact 0.0944; GEV 0.0971]

(iii) Repeat (i) with $n = 25$ and $m = 10{,}000$. [exact 0.8905; GEV 0.8907]

(iv) Comment on your results.

4.5 A company is modelling losses from cyberattacks. The following table shows the largest 20 values of a sample of 1,000 observations of the losses, sorted in decreasing order:

196.1	148.2	79.8	35.8	27.1	22.9	21.8	20.9	16.7	15.8
15.5	13.7	13.0	12.5	11.5	10.3	9.7	9.4	8.3	8.0

(a) Estimate the 99% VaR of X. [15.80]

(b) Estimate the 99% Expected Shortfall of X. [58.51]

(c) Assume that the losses are from a distribution which is GPD for $Y = X - d \,|\, X > d$, where $d = 10$, with shape parameter $\xi = 0$, and scale parameter $\beta = 30$.

(i) Show that

$$\Pr(X > x) = \overline{F}_X(d) \left(e^{-(x-d)/\beta} \right)$$

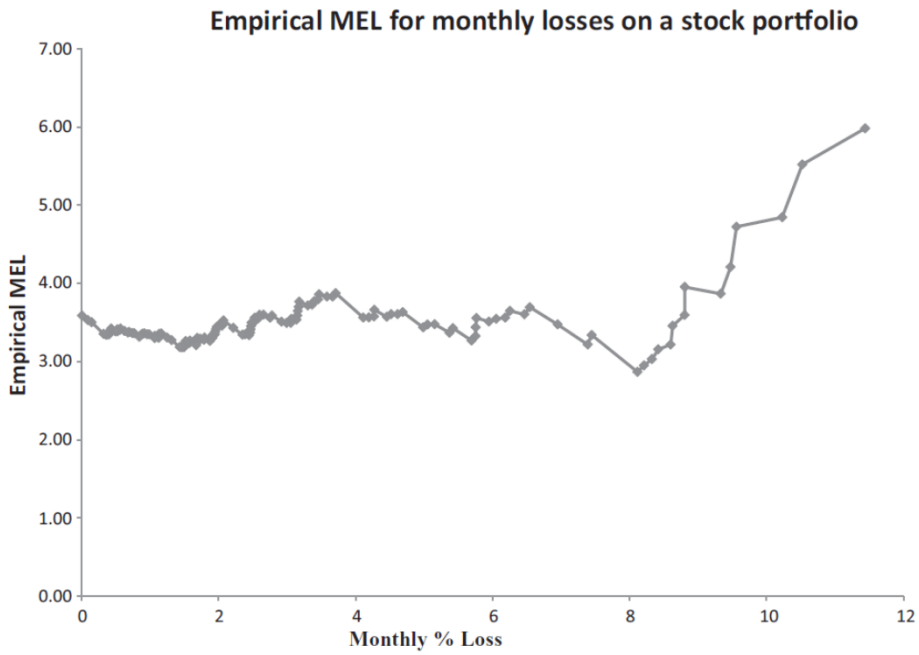
for $x > d$.

(ii) Estimate the 99% VaR and 99% Expected Shortfall using the GPD for the tail probabilities. [24.10; 54.10]

(iii) Comment on the differences between your estimates in (a) and (b), and your estimates using the GPD.

4.6 An investment firm is interested in fitting an extreme value distribution to portfolio losses. The data set comprises 300 values for the monthly percentage loss in portfolio value.

The analyst has assumed the data is a random sample of independent identically distributed observations. She has calculated the empirical MEL function, which is shown in the figure below.



You are also given the 20 largest values from the data, sorted from high to low:

24.3	18.4	15.6	11.4	10.5	10.2	9.6	9.5	9.3	8.8
8.8	8.6	8.6	8.4	8.3	8.2	8.1	7.4	7.4	6.9

(a) (i) Estimate the 98% Expected Shortfall from the data. [15.07]

(ii) Calculate the empirical MEL function, $\hat{e}(x)$ at $x = 10.0$. [5.07]

(b) The analyst decides to use the GPD with threshold $d = 8$. Explain why this choice appears reasonable.

(c) The MLE estimators for the GPD distribution, with $d = 8$, are $\hat{\xi} = 0.55$, $\hat{\beta} = 1.5$.

(i) Calculate the 98% Expected Shortfall using the GPD distribution. [16.02]

(d) Comment on the difference between the Expected Shortfall estimates from (a)(i) and (c)(i). Which method would you recommend?

Additional Practice Problems (not assigned)

4.7 Suppose X follows an exponential distribution with mean $1/\lambda$. Find the density function of Y, where

(a) $Y = 3X^2$. [$f_Y(y) = \frac{\lambda}{2\sqrt{3y}}\, e^{-\lambda\sqrt{y/3}}, y > 0$]

(b) $Y = 1/\sqrt{X}$. [$f_Y(y) = \frac{2\lambda}{y^3}\, e^{-\lambda/y^2}, y > 0$]

(c) $Y = \ln X$. [$f_Y(y) = \lambda e^{y-\lambda e^y}, -\infty < y < \infty$]

4.8 Conditioning on the health status of an insured $\Lambda = \lambda$, the total loss from a medical policy for each insured follows an exponential distribution with mean λ. Λ follows a single-parameter Pareto distribution with $\alpha = 1$ and $\theta = 1000$. That is, the cumulative distribution function of Λ is given by

$$F_\Lambda(y) = 1 - \left(\frac{1000}{y}\right), \qquad y \geq 1000.$$

Calculate the probability that the total loss of the policy will be no more than 500. [0.2131]

4.9 You are given the following information about three random variables that are in the MDA of the GEV distribution. State with reasons whether $\xi < 0$, $\xi = 0$, or $\xi > 0$.

(a) The first random variable, X, is greater than 0, unbounded on the right side, and has finite kth moment for all $k = 1, 2, 3, \ldots$ [$\xi = 0$]

(b) The second random variable, Y, is equal to $-X$. [$\xi < 0$]

(c) The third random variable has a finite number of moments. [$\xi > 0$]

4.10 Consider an exponential distribution with distribution function

$$F(x) = 1 - e^{-x/\beta}, \qquad x \geq 0.$$

(a) Show that F is in the maximum domain of attraction of the Gumbel distribution, using normalizing sequences $c_n = \beta$ and $d_n = \beta \log n$.

(b) Let M denote the maximum of 80 observations of an exponential distribution with mean 100. Calculate the probability that $M > 1{,}000$ using (i) the exponential distribution and (ii) the Gumbel distribution. [(i) 0.0036255; (ii) 0.0036254]

Lecture 5: Credibility

5.1 Two eight-sided dice, A and B, are used to determine the number of claims for an insured. The faces of each die are marked with either 0 or 1, representing the number of claims for that insured for the year.

Die	$\Pr(\text{Claims} = 0)$	$\Pr(\text{Claims} = 1)$
A	1/4	3/4
B	3/4	1/4

Two spinners, X and Y, are used to determine claim cost. Spinner X has two areas marked 12 and c. Spinner Y only has one area marked 12.

Spinner	$\Pr(\text{Cost} = 12)$	$\Pr(\text{Cost} = c)$
X	1/2	1/2
Y	1	0

To determine the losses of the year, a die is randomly selected from A and B and rolled. If a claim occurs, a spinner is randomly selected from X and Y and spun. For subsequent years, the same die and spinner are used to determine losses.

Losses for the first year are 12. Based upon the first year, you determine that the expected losses for the second year are 10.

Calculate c. [36]

5.2 You are given:

(i) An individual automobile insured has an annual claim frequency distribution that follows a Poisson distribution with mean λ.

(ii) λ is assumed to follow a gamma distribution with parameters α and θ.

(iii) The first actuary assumes that $\alpha = 1$ and $\theta = 1/6$.

(iv) The second actuary assumes the same mean for the gamma distribution, but only half the variance.

(v) A total of one claim is observed for the insured over a three year period.

(vi) Both actuaries determine the Bayesian premium for the expected number of claims in the next year using their model assumptions.

You may use the fact that if Λ is gamma with integer α, then $2\Lambda/\theta$ is chi-square with 2α degrees of freedom. The 2.5th, 50th, and 97.5th percentiles of the chi-square distribution with 4 degrees of freedom are 0.4844, 3.3567, and 11.1433.

(a) Determine the ratio of the Bayesian premium that the first actuary calculates to the Bayesian premium that the second actuary calculates. $[10/9 = 1.111]$

(b) Using the first actuary's model, calculate the Bayes estimate of λ under absolute loss and under zero-one loss. [0.1865; 0.1111]

(c) Using the first actuary's model, calculate a 95% equal-tailed credibility interval for λ. [(0.0269, 0.6191)]

5.3 You are given the following information about workers' compensation coverage:

(i) The number of claims for an employee during the year follows a Poisson distribution with mean

$$(100 - p)/100$$

where p is the salary (in thousands) for the employee.

(ii) The distribution of p is uniform on the interval $(0, 100]$.

An employee is selected at random. During the last 4 years, the employee has had a total of 5 claims.

Determine the Bühlmann credibility estimate for the expected number of claims the selected employee will have next year. [0.8]

5.4 For a portfolio of insurance risks, aggregate losses per year per exposure follow a normal distribution with mean θ and standard deviation 1000, with θ varying by class as follows:

Class	θ	Percent of Risks in Class
X	2000	60%
Y	3000	30%
Z	4000	10%

A randomly selected risk has the following experience over three years:

Year	Number of Exposures	Aggregate Losses
1	24	24,000
2	30	36,000
3	26	28,000

Calculate the Bühlmann-Straub estimate of the mean aggregate losses per year per exposure in Year 4 for this risk. [1137.84]

5.5 In one year, 500 policyholders had the following numbers of claims:

Number of Claims	Number of Policyholders
0	300
1	140
2	45
3	12
4	3

Each policyholder's annual number of claims is assumed to be Poisson, with a mean that varies by policyholder. Use semiparametric empirical Bayes estimation.

(a) Explain why the Poisson assumption allows you to estimate v from the same data used to estimate μ.

(b) Estimate μ, v, and a. [0.556; 0.556; 0.0882]

(c) Calculate the credibility factor for one year of data. [0.1369]

(d) Calculate the credibility estimate of next year's number of claims for a policyholder who had 2 claims, and for a policyholder who had no claims. [0.7536; 0.4799]

(e) A policyholder from the same population had 4 claims in the last 3 years. Calculate the credibility estimate of the number of claims next year. [0.8066]

5.6 You are given the following experience for two insured groups:

Group		Year 1	Year 2	Year 3	Total
1	Number of members	8	12	5	25
	Average loss per member	96	91	113	97
2	Number of members	25	30	20	75
	Average loss per member	113	111	116	113
Total	Number of members				100
	Average loss per member				109

Determine the nonparametric empirical Bayes credibility premium for group 1. [98.26]

Additional Practice Problems (not assigned)

5.7 You are given the following:

(i) A portfolio consists of 150 independent risks.

(ii) 100 of the risks each have a policy with a \$100,000 per claim policy limit, and 50 of the risks each have a policy with a \$1,000,000 per claim policy limit.

(iii) The risks have identical claim count distributions.

(iv) Prior to censoring by policy limits, the claim size distribution for each risk is as follows:

Claim Size	Probabilities
\$10,000	1/2
\$50,000	1/4
\$100,000	1/5
\$1,000,000	1/20

(v) A claims report is available that shows actual claim sizes incurred for each policy after censoring by policy limits, but does not identify the policy limit associated with each policy.

(vi) Claim counts and claim sizes are independent within each risk type.

The claims report shows exactly three claims for a policy selected at random. Two of the claims are \$100,000, but the amount of the third is illegible.

Determine the expected value of this illegible number. [53,409]

5.8 You are given the following information about workers' compensation coverage:

(i) The number of claims for an employee during the year follows a Poisson distribution with mean

$$(100 - p)/100$$

where p is the salary (in thousands) for the employee.

(ii) The distribution of p is uniform on the interval $(0, 100]$.

An employee is selected at random. No claims were observed for this employee during the year.

Determine the posterior probability that the selected employee has salary greater than 50 thousand. [0.6225]

5.9 You are given the following:

- The probability of observing exactly one claim in a trial is p, while the probability of no claim observed is $1 - p$.
- The prior distribution of p has density

$$\pi(p) = 280p^3(1 - p)^4, \qquad 0 < p < 1.$$

- Two claims are observed in 3 trials.

Determine the Bayesian estimate of the number of claims observed in the next trial. [0.5]

5.10 You are given:

(i) Losses in a given year follow a gamma distribution with parameters α and θ, where θ does not vary by policyholder.

(ii) The prior distribution of α has mean 10.

(iii) The Bühlmann credibility factor based on two years of experience is 0.25.

Calculate $\mathrm{Var}(\alpha)$. [5/3]

5.11 You are given the following:

(i) The respective claim count distributions for risks of Type A and Type B are

	Probabilities	
Number of Claims	Type A	Type B
0	4/9	1/9
1	4/9	4/9
2	1/9	4/9

(ii) The respective claim size distributions for risks of Type A and Type B are

Type A		Type B	
Claim Size	Probabilities	Claim Size	Probabilities
500	1/3	250	2/3
1235	2/3	328	1/3

(iii) Risks are equally likely to be Type A or Type B.

(iv) Claim counts and claim sizes are independent within each risk type.

(v) The variance of the total losses is 296,962.

A randomly selected risk is observed to have total annual losses of 500.

Determine the Bühlmann credibility premium for the next year for this same risk. [513]

5.12 You are given the following:

(i) The number of claims follows a distribution with mean λ and variance 2λ.

(ii) Claim sizes follow a distribution with mean μ and variance $2\mu^2$.

(iii) The number of claims and claim sizes are independent.

(iv) λ and μ have a prior probability distribution with joint density function

$$f(\lambda, \mu) = 1, \qquad 0 < \lambda < 1, \quad 0 < \mu < 1.$$

Determine the value of Bühlmann's K for aggregate losses. [$96/7 = 13.71$]

5.13 Three policyholders have the following claims experience over three months:

Policyholder	Month 1	Month 2	Month 3
I	4	6	5
II	8	11	8
III	5	7	6

Nonparametric empirical Bayes estimation is used to estimate the credibility premium in Month 4.

Calculate the credibility factor Z. $[34/39 = 0.8718]$

Lecture 6: Ratemaking

6.1 An insurer has the following loss costs (pure premiums, including LAE) for a line of business:

Accident Year	Loss Cost
2020	210.5
2021	221.8
2022	238.0
2023	247.6
2024	266.1

(a) Let $X =$ Accident Year $- 2020$. Fit the model $\ln(\text{Loss Cost}) = a + bX$ by least squares, and determine the indicated annual loss trend. [$a = 5.3481$, $b = 0.05788$; $+5.96\%$ per year]

(b) New rates will be in effect for one year beginning January 1, 2026, and all policies are annual. Using the AY 2024 loss cost and the trend from (a),

(i) determine the trend period, and [2.5 years]

(ii) calculate the projected loss cost. [307.53]

(c) Fixed expenses are 30 per exposure, variable expenses are 20% of premium, and the target underwriting profit is 4% of premium. Using the pure premium method, calculate the indicated average rate. [444.12]

(d) The current average rate is 420. Calculate the indicated overall rate change. [$+5.74\%$]

(e) Explain why the selected trend should not be based only on the fitted curve in (a), and give two factors that can drive loss trends.

6.2 An insurer is reviewing the rates for a line of business. You are given:

(i) The new rates will be in effect for one year beginning January 1, 2027.

(ii) The experience period is Accident Year (AY) 2025, with losses evaluated as of December 31, 2025.

(iii) The selected annual loss trend is 4%.

(iv) The following data for AY 2025:

Earned exposures	8,000
Earned premium at current rate level	4,000,000
Reported losses and LAE as of 12/31/2025	2,400,000
Loss development factor to ultimate	1.25
Projected fixed expenses	320,000

(v) Variable expenses are 15% of premium, and the target underwriting profit is 5% of premium.

(a) Determine the trend period and the trend factor if

(i) all policies are annual, [2.5 years; 1.1030]

(ii) all policies are six-month policies, and [2.25 years; 1.0923]

(iii) all policies are annual, but the experience period is Policy Year 2025 instead of AY 2025. [2 years; 1.0816]

(b) Assume all policies are annual. Calculate the projected ultimate losses and LAE for the AY 2025 experience. [3,309,060]

(c) Using the loss ratio method, calculate the indicated overall rate change. [$+13.41\%$]

(d) Using the pure premium method, calculate the indicated average rate, and show that it is consistent with your answer to (c). [567.04]

(e) Describe one situation in which the pure premium method is preferable and one situation in which the loss ratio method is preferable.

6.3 An insurer offers auto physical damage policies with a choice of three deductibles: 250 (the base level), 500, and 1000. You are given:

(i) Ground-up losses follow a Pareto distribution with $\alpha = 3$ and $\theta = 3000$.

(ii) The distribution of the number of ground-up losses is the same for every deductible option, and is independent of the loss amounts.

(a) (i) Calculate the expected payment per loss for each deductible option. [1278.11; 1102.04; 843.75]

(ii) Show that the indicated deductible relativities for the 500 and 1000 deductibles are 0.86 and 0.66 to the nearest 0.01. You should calculate the values to the nearest 0.0001. [0.8622; 0.6602]

The insurer currently uses the rounded relativities from (a)(ii), with a base rate of 150. The most recent experience, with losses developed and trended to the new rate period, is:

Deductible	Current Relativity	Earned Exposures	Losses
250	1.00	12,000	1,560,000
500	0.86	6,000	690,000
1000	0.66	2,000	160,000

(b) (i) Calculate the indicated relativities using the pure premium method. [1.00; 0.8846; 0.6154]

(ii) Calculate the indicated relativities using the loss ratio method. [1.00; 0.8846; 0.6154]

(iii) Explain why the two methods give the same answer here.

(c) Give one reason why the indicated relativity for the 1000 deductible in (b) is lower than the value from the Pareto model in (a).

(d) The insurer adopts the relativities from (b) and implements an overall rate increase of 6%.

(i) Calculate the off-balance factor. [1.0032]

(ii) Calculate the new base rate. [158.50]

(e) Calculate the percentage change in premium for each deductible option, and comment on the result for the 1000 deductible. [$+5.67\%$; $+8.69\%$; -1.48%]

6.4 You are given the following information and asked to perform a ratemaking analysis for an insurance product:

- The new rates are to become effective February 1, Calendar Year (CY) 6 for 12 months, and all policies are one-year policies.
- The loss development factors are estimated based on the following run-off triangle.

Table 1: Reported Losses (including LAE) as of

Accident Year (AY)	12/31/AY2	12/31/AY3	12/31/AY4
1	1,000,000	1,050,000	1,050,000
2	1,100,000	1,210,000	1,270,500
3		1,300,000	1,430,000

- The loss trend factor is 8% per annum effective.
- The permissible loss ratio is 0.80.
- The new rates are determined based on the AY 4 claims experience:

Territory	Class	AY 4 Earned Exposure	AY 4 Earned Premium	AY 4 Reported L&LAE
A	1	10,000	900,000	700,000
A	2	4,000	460,000	400,000
B	1	2,000	180,000	200,000
B	2	2,000	250,000	270,000

- The current base rate is 100 with differentials:

Differentials	Territory A	Territory B
Class 1	1.00	1.10
Class 2	1.20	1.32

(a) Determine the indicated overall rate change. [$+40.8\%$]

(b) Determine the new differentials. (Note: both territory and class differentials will change.) [Territory B: 1.4372; Class 2: 1.4649]

(c) Determine the new rate for each territory and class. [A1: 120.52; A2: 176.55; B1: 173.21; B2: 253.73]

Lecture 7: Outstanding Claims Reserves

7.1 You are given the following cumulative paid claims (in thousands) for a portfolio of short-term policies. Assume that claims are fully developed by the end of Development Year (DY) 3, that all payments are made in the middle of each year, and that the valuation date is December 31, 2025.

	Cumulative Paid Claims by Development Year (DY)			
Accident Year (AY)	0	1	2	3
2022	1,200	1,980	2,280	2,350
2023	1,350	2,180	2,520	
2024	1,500	2,400		
2025	1,620			

(a) Calculate the chain ladder development factors $\hat{f}_j$ for $j = 0, 1, 2$. [1.6198; 1.1538; 1.0307]

(b) Using the chain ladder method, calculate the estimated ultimate claims and outstanding claims for each accident year, and the total outstanding claims. [ultimate: 2350, 2597.4, 2854.3, 3120.6; outstanding: 0, 77.4, 454.3, 1500.6; total 2032.3]

(c) Calculate the projected claim payments in each of calendar years 2026, 2027, and 2028. [1450.6; 488.7; 93.0]

(d) Using an annual effective interest rate of 4%, calculate the discounted value of the total outstanding claims as of December 31, 2025. [1967.5]

(e) The earned premium for AY 2025 is 3,300 and the expected loss ratio is 75%.

(i) Calculate the Bornhuetter-Ferguson estimate of the outstanding claims for AY 2025. [1190.2]

(ii) Show that the Bornhuetter-Ferguson estimate of the ultimate claims for AY 2025 is a credibility-weighted average of the chain ladder estimate and the prior estimate, and state the credibility factor. [$Z = 0.5191$]

(iii) Calculate the expected loss ratio estimate of the outstanding claims for AY 2025. [855]

(f) You learn that the AY 2025 payments in DY 0 include a single unusually large claim of 300 that has been fully settled.

(i) Recalculate the chain ladder and Bornhuetter-Ferguson estimates of the outstanding claims for AY 2025 with this claim removed from the data. [CL: 1222.8; BF: 1190.2]

(ii) Comment on the sensitivity of the two methods to unusual claims in the most recent accident year.

7.2 You are given the following incremental claim payments by development year. Assume all incremental payments are made at the middle of each year, and claims are fully developed by the end of Development Year (DY) 4.

	Incremental Claims by Development Year (DY)				
Accident Year (AY)	0	1	2	3	4
0	81,345	55,315	12,299	2,085	755
1	98,020	70,574	20,231	1,699	
2	94,987	56,042	16,613		
3	101,760	66,144			
4	115,734				

You are also given the following inflation rates r_t, measured from mid-year t to mid-year $t+1$:

Year t	Inflation Rate (Past)	Year t	Inflation Rate (Assumed)
0	2.0%	4	5.8%
1	3.5%	5	4.4%
2	5.7%	6	3.6%
3	6.3%	7	3.0%

(a) Calculate the claims inflation index, Q_k, using the inflation rates, for $k = 0, 1, \ldots, 8$. Set $Q_0 = 100.0$. [100, 102, 105.57, 111.59, 118.62, 125.50, 131.02, 135.74, 139.81]

(b) Using the inflation-adjusted chain ladder method, calculate the estimated total ultimate claims from all five accident years of the table, expressed in terms of mid AY 4 monetary units. [985,892]

(c) Calculate the inflation-adjusted outstanding claims reserve. [127,091]

7.3 An insurer records the cumulative number of settled claims and the incremental claim payments for a portfolio. All claims are settled by the end of Development Year (DY) 3.

Cumulative Number of Settled Claims

AY	0	1	2	3
0	40	70	85	90
1	45	80	96	
2	50	88		
3	55			

Incremental Claim Payments

AY	0	1	2	3
0	8,000	24,000	22,500	12,000
1	9,450	27,300	26,400	
2	11,000	31,200		
3	12,100			

(a) Using the chain ladder method, project the number of claims settled in each future development year for AY 1, AY 2, and AY 3. [AY 1: 5.647; AY 2: 18.187, 6.246; AY 3: 41.963, 20.039, 6.882]

(b) Calculate the average incremental payment per settled claim for each development year. [213.42, 800.97, 1577.42, 2400]

(c) Estimate the outstanding claims for each accident year and in total using the frequency-severity method. [13,553; 43,679; 81,739; total 138,971]

(d) Estimate the total outstanding claims by applying the chain ladder method to the cumulative payments, and compare. [146,252]

(e) Explain why the average payment per settled claim increases with the development year.

7.4 An insurer records the following numbers of reported claims by accident year and development year. All claims are reported by the end of Development Year (DY) 3.

	Incremental Reported Claim Counts by Development Year (DY)			
Accident Year (AY)	0	1	2	3
0	120	60	25	10
1	130	70	30	
2	140	66		
3	150			

The insurer uses the Poisson model: the incremental counts $X_{i,j}$ are independent with

$$X_{i,j} \sim \text{Poi}\left(\mu_i \gamma_j\right), \qquad \mu_i, \gamma_j > 0, \qquad \sum_{j=0}^{3} \gamma_j = 1,$$

where μ_i and γ_j are estimated by maximum likelihood.

(a) Calculate the chain ladder development factors $\hat{f}_j$ for $j = 0, 1, 2$, based on the cumulative claim counts. [1.5026; 1.1447; 1.0488]

(b) Using the relationship between the Poisson model and the chain ladder method, determine $\hat{\gamma}_j$ for $j = 0, 1, 2, 3$ and $\hat{\mu}_i$ for $i = 0, 1, 2, 3$. [$\hat{\gamma}_j$: 0.5543, 0.2786, 0.1206, 0.0465; $\hat{\mu}_i$: 215, 241.22, 247.32, 270.59]

(c) The maximum likelihood estimates satisfy $\sum_{i=0}^{3-j} X_{i,j} = \hat{\gamma}_j \sum_{i=0}^{3-j} \hat{\mu}_i$ for each j. Verify that your estimates satisfy this equation for $j = 2$.

(d) Calculate the expected number of claims that will be reported

(i) in the next calendar year, and [116.42]

(ii) in total in all future years. [173.13]

(e) Treating the estimated parameters as known, calculate the standard deviation of the total number of claims still to be reported under the Poisson model. [13.16]

(f) Give two reasons why the standard deviation in (e) may understate the uncertainty in the number of claims still to be reported.

(g) The insurer instead uses the overdispersed Poisson model, with $\text{Var}(X_{i,j}) = \phi\, \mu_i \gamma_j$.

(i) Estimate ϕ using Pearson residuals. [0.157]

(ii) Calculate the process standard deviation of the total number of claims still to be reported, and comment on your result. [5.21]

7.5 You are given the following cumulative claims triangle. Claims are fully developed by the end of Development Year (DY) 2.

AY	0	1	2
0	1000	1500	1600
1	1200	1650	
2	900		

You use the Bühlmann-Straub credibility model to estimate the outstanding claims.

(a) Calculate the chain ladder estimates of $\hat{\gamma}_j$ and of the ultimate claims for each accident year. [0.6548, 0.2827, 0.0625; 1600, 1760, 1374.55]

(b) Calculate $\hat{s}_0^2$, $\hat{s}_1^2$, and $\hat{v}$. [5741.6, 11,483.3; 8612.4]

(c) Calculate $\hat{a}$. [23,533]

(d) Calculate the credibility factors Z_0, Z_1, Z_2 and the credibility-weighted mean $\hat{\mu}$. [0.7321, 0.7192, 0.6415; 1585.88]

(e) Calculate the Bühlmann-Straub estimates of the outstanding claims for AY 1 and AY 2, and compare them with the chain ladder estimates. [61.11, 550.32; CL: 110.00, 474.55]

(f) Give one advantage and one disadvantage of the Bühlmann-Straub approach compared with the Bornhuetter-Ferguson method.

7.6 Your actuarial team is given the following cumulative paid claims run-off triangle for determining the indicated outstanding claims reserves.

	Cumulative Paid Claims by Development Year (DY)				
Accident Year (AY)	0	1	2	3	4
0	81,345	136,660	148,959	151,044	151,800
1	98,020	168,594	188,826	190,525	
2	94,987	151,029	167,642		
3	101,760	167,904			
4	115,734				

(a) Team member A determines the total outstanding claims reserves using the chain ladder (CL) method.

(i) Calculate $f_{i,j}$ for $i, j \leq 3$. [AY 0: 1.680, 1.090, 1.014, 1.005; AY 1: 1.720, 1.120, 1.009; AY 2: 1.590, 1.110; AY 3: 1.650]

(ii) Estimate $\hat{f}_j$ for $j \leq 3$. [1.6596, 1.1077, 1.0112, 1.0050]

(iii) Calculate the CL estimate of $C_{i,4}$ for $i = 1, 2, 3, 4$. [191,478.6; 170,368.5; 189,013.0; 216,216.6]

(b) Team member B determines the total outstanding claims reserves using the Bornhuetter-Ferguson (BF) method with prior estimates of the expected total losses for each accident year as:

AY	1	2	3	4
μ_i	200,000	180,000	220,000	240,000

(i) Calculate $\hat{\beta}_j$ for $j \leq 3$. [0.5353, 0.8883, 0.9840, 0.9950]

(ii) Calculate the BF estimate of outstanding claims reserves for Accident Year i, $i = 1, 2, 3, 4$. [996.0; 2880.6; 24,569.6; 111,535.5]

(c) Team member C assumes that incremental claim payments $X_{i,j}$ follow the Poisson model:

$$X_{i,j} \sim \text{Poi}\left(\mu_i \gamma_j\right), \qquad \mu_i, \gamma_j > 0, \qquad \sum_{j=0}^{4} \gamma_j = 1,$$

where μ_i and γ_j are estimated using maximum likelihood based on the given run-off triangle data.

Using the results in (a), estimate $\hat{\mu}_i$ and $\hat{\gamma}_j$ for $i, j = 0, 1, \ldots, 4$. [$\hat{\mu}_i$: 151,800; 191,478.6; 170,368.5; 189,013.0; 216,216.6. $\hat{\gamma}_j$: 0.5353, 0.3531, 0.0957, 0.0110, 0.0050]

(d) Team member D further examines the uncertainty of outstanding claims reserves using Mack's model.

(i) Calculate $\hat{\sigma}_j^2$ for $j \leq 3$. [286.99, 34.575, 2.0813, 0.1253]

(ii) Calculate the process standard deviation for outstanding claims for AY 3, using Mack's approximation. [2531.8]

(iii) Calculate the square root of the estimation error for AY 3, using Mack's approximation. [1565.5]