Advanced Short-Term Actuarial Mathematics

Sample Questions

This Study Note contains sample questions that are designed to help candidates in their preparation for the Advanced Short-Term Actuarial Mathematics exam.

Versions:

Nov. 9, 2022	Original set of 20 sample questions published for the ASTAM exam
Feb. 6, 2024	Added two new sample questions (Questions 21-22); Question 22 is a sample Excel question.

Current Version Dated Feb. 6, 2024

Question 1

Santiago Substandard Auto Company (SSAC) sells automobile insurance to drivers with bad driving records. Each policy only covers one automobile.

During 2021, SSAC had 10,000 policies. There were no deductibles, no upper limit, and no coinsurance, so that each policy covered 100% of all losses.

During 2021, the number of claims per insured automobile for SSAC was:

Number of claims in 2021	Number of Policies
0	2400
1	3400
2	2400
3	1500
4	300

(a) First, assume that the claim frequency is distributed as a Poisson random variable with mean parameter λ. Calculate the 90% linear confidence interval for λ based on the above data.

(b) Using the Chi-Square Test with a 99% significance level, SSAC wishes to use the data to test the following hypotheses:

H_0 : The data is from a Poisson distribution.

H_1 : The data is not from a Poisson distribution.

The data for 3 and 4 claims in 2021 are combined for this test.

(i) Calculate the Chi-Square test statistic.

(ii) Calculate the *p*-value for this test and state your conclusion with regard to the hypotheses.

(c) You are given the following claim severity sample from 2021:

200 1000 5000 10,000 100,000

Using the Kolmogorov-Smirnov Test with a 95% significance level, SSAC uses this data to test the following hypotheses:

H_0 : The data is from a Pareto distribution with $\alpha = 4$ and $\theta = 75{,}000$.

H_1 : The data is not from a Pareto distribution with $\alpha = 4$ and $\theta = 75{,}000$.

(i) Calculate the Kolmogorov-Smirnov Test statistic.

(ii) You are given that the 5% critical value for the Kolmogorov-Smirnov test is approximately $\dfrac{1.36}{\sqrt{n}}$.

State your conclusion.

(d) In selecting a model, there are two important concepts. The first concept is known as parsimony while the second concept does not have a specific name.

(i) State the principle of parsimony.

(ii) State the second concept.

(iii) Identify the score-based approach that is consistent with parsimony and state why it is consistent.

Question 2

(a) There are four essential objectives of Ratemaking. Describe each one and explain why these objectives are essential.

(b) You are the rate making actuary for ABC Auto. You are setting rates for the auto coverage, which is a short-term insurance product. You are given the following data:

Calendar Year	Earned Premium
2017	10,000
2018	12,000
2019	8,000

Assume that all policies are one-year policies and that policies are issued uniformly throughout the year.

The following rate changes have occurred:

Date	Rate Change
March 15, 2017	10% Increase
July 30, 2018	8% Increase
October 31, 2019	4% Decrease

(i) Graphically display the parallelogram that would be used to apply the parallelogram method.

(ii) Using the parallelogram method, calculate the earned premium for 2017, 2018, and 2019 based on current rates.

(c) There are two methods for determining the overall change in rates – the loss cost method and the loss ratio method. Prove that both methods will always provide the same answer.

(d) For collision coverage for ABC Auto, rate differentials for 2019 are developed using the following information:

Type	Proportion of Total Drivers	Claim Frequency Poisson Annual	Severity Gamma
Safe	0.5	0.04	$\alpha = 3, \theta = 1000$
Not So Safe	0.3	0.10	$\alpha = 4, \theta = 1000$
Reckless	0.2	0.25	$\alpha = 5, \theta = 1000$

(i) Calculate the indicated differentials for each Type, given that Safe is the base rate.

(ii) During 2019, ABC Auto experienced the following loss ratios for collision coverage based on the indicated differentials developed in Question 16:

Type	Loss Ratio
Safe	60%
Not So Safe	63%
Reckless	51%

Use the loss ratio method to determine the new indicated differentials for 2020 for the Not So Safe and Reckless categories.

Question 3

During 2021, AAA Insurance collects the following premium amounts:

Month	January	March	May	July
Premium Collected	1,600	1,800	1,200	1,200

All premiums are paid on the first day of the month and all premiums are annual premiums.

The unearned premium reserve at 31 Dec 2020 was 1,300.

The company expects a loss ratio of 80%.

During 2021, the company paid losses for claims incurred in 2021 of 2,500.

(a) Use the loss ratio method to determine the outstanding claim reserve for accident year 2021 for this coverage that should be held on 31 Dec 2021.

(b) The insurer has the following Paid Claims triangle for this coverage. There is no further development after year 5.

Cumulative Loss Payments by Development Year						
Accident Year	Development Year					
	0	1	2	3	4	5
2016	1,000	1,500	1,700	1,800	1,850	1,875
2017	1,100	1,750	1,775	1,825	1,870	
2018	1,200	1,900	2,200	2,350		
2019	1,500	2,200	2,500			
2020	2,000	2,900				
2021	2,500					

(i) The loss development factor for the period from year 0 to year 1 is 1.51 to two decimal places. Calculate the loss development factor to five decimal places.

(ii) Calculate the outstanding claim reserve on December 31, 2021.

(iii) Determine the total amount of claims paid in 2021.

(c) For the claims from accident year 2021, determine the reserves as of 31 Dec 2021 using the Bornhuetter-Ferguson method, where the prior estimate is the outstanding claim reserve using the loss ratio method.

(d) Explain briefly one advantage and one disadvantage of the Bornhuetter-Ferguson method, compared with the Chain Ladder approach.

Question 4

CBA Insurance Company splits drivers into three categories – Safe, Not So Safe, and Reckless. Within each category, claim frequency is Poisson distributed, and claim severity is Gamma distributed, but the parameters of the frequency and severity distributions differ, as shown in the following table.

Type	Proportion of Total Drivers	Frequency parameter	Severity parameters
Safe	0.5	λ=0.04	$\alpha = 3, \theta = 1000$
Not So Safe	0.3	λ=0.10	$\alpha = 4, \theta = 1000$
Reckless	0.2	λ=0.25	$\alpha = 5, \theta = 1000$

(a) The expected value of the process variance of the claim severities (for the observation of a single claim) under automobile coverage is 4 million to the nearest million. Calculate it to the nearest 1000.

(b) What is the variance of the hypothetical means for claim severities (for the observation of a single claim) under automobile coverage?

(c) Over several years, an individual driver with automobile coverage has single claim of 20,000. Use Bühlmann Credibility to estimate this driver's future average claim severity.

(d) What is the expected value of the process variance of the pure premium (for the observation of a single exposure) under this coverage?

Question 5

You are given that $X_1, X_2, ..., X_n$ are independent and identically distributed copies of a continuous loss random variable X, where X represents a firm's aggregate daily losses from operational risks. Let $M_n = \max(X_1, X_2, ..., X_n)$.

The *n-day Expected Maximum* (*nDEM*) risk measure is defined as $\rho(X) = \mathrm{E}[M_n]$.

Currently, the firm uses the 95% VaR risk measure, denoted $Q_{0.95}(X)$. A colleague suggests changing to the *nDEM* risk measure, using $n = 20$ for consistency with the current 1-in-20-day standard.

(a) Show that $\Pr[M_{20} > Q_{0.95}(X)] = 0.64$ to the nearest 0.01. You should calculate the probability to the nearest 0.001.

Now suppose that X is exponentially distributed with mean 1. Let $\alpha(n) = \dfrac{n-1}{n}$.

(b) Show that the $\alpha(n)$-VaR of X is $Q_{\alpha(n)}(X) = \log n$.

(c) Show that the $\alpha(n)$-Expected Shortfall of X is $ES_\alpha(X) = 1 + log\, n$. You may use without proof the memoryless property of the exponential distribution.

(d) You are given that

- The exponential distribution is in the Maximum Domain of Attraction (MDA) of the Gumbel distribution.
- The mean of the standard Gumbel distribution is 0.5772.
- The normalizing functions for the exp(1) distribution, under the Fisher-Tippett-Gnedenko theorem, are $c_n = 1$ and $d_n = \log n$.

(i) Explain what it means to say that the exponential distribution is in the MDA of the Gumbel distribution.

(ii) Assume that n is sufficiently large that the normalized n-day maximum loss can be approximated by the standard Gumbel distribution. Show that

$$Q_{\alpha(n)}(X) < E[M_n] < ES_{\alpha(n)}(X)$$

(e) Explain why the ordering in (d) is likely to hold more generally.

Question 6

Individual losses from a risk follow a Pareto distribution with parameters α and λ. Insurance covers these losses up to a policy limit of *M*. There is no deductible.

(a) Show that the expected claim severity is

$$\frac{\lambda}{\alpha-1}\left(1-\left(\frac{\lambda}{\lambda+M}\right)^{\alpha-1}\right)$$

(b) A sample of 100 claims is available. Of these, three were above the policy limit of 20,000. The remaining 97 were below the policy limit and the precise values, $x_1, x_2, ..., x_{97}$, are known.

Derive and simplify as far as possible, expressions for the maximum likelihood estimators $\hat{\alpha}$ and $\hat{\lambda}$.

(c) You are given that $\hat{\lambda} = 10,683$, and that

$$\sum_{i=1}^{97}\log\left(\hat{\lambda}+x_i\right)=925.314; \quad \sum_{i=1}^{97}\frac{1}{\hat{\lambda}+x_i}=0.007150; \quad \sum_{i=1}^{97}\frac{1}{\left(\hat{\lambda}+x_i\right)^2}=5.488\times10^{-7}$$

Show that $\hat{\alpha}=3.4$ to the nearest 0.1. You should calculate the value to the nearest 0.01.

(d) Calculate the approximate asymptotic variance-covariance matrix for $\hat{\alpha}$ and $\hat{\lambda}$.

Question 7

Aggregate claims for a portfolio are assumed to follow a compound Poisson distribution, with Poisson parameter $\lambda = 50$. The individual claims distribution is lognormal, with parameters $\mu = 6.0, \quad \sigma^2 = 1.5$.

(a) Show that, if $f(x)$ is the density function for the lognormal (μ,σ) distribution, then

$$\int_d^\infty x^k f(x)\,dx = e^{k\mu + k^2\sigma^2/2}\left\{1 - \Phi\left(\frac{\log d - \mu - k\sigma^2}{\sigma}\right)\right\}$$

(b) The insurer shares the risk with a reinsurer under an excess of loss agreement with retention 3000.

(i) Calculate the expected number of claims involving the reinsurer.

(ii) Show that the expected value of an individual reinsurance payment, given that the loss involves the reinsurer, is 2722.

(iii) Show that the reinsurer's expected aggregate claim cost is 6,900 to the nearest 10.

(iv) Calculate the direct insurer's expected aggregate claims cost after reinsurance.

(c) Calculate the standard deviation of the reinsurer's aggregate claims under the excess of loss contract.

(d) The insurer is considering changing to a proportional reinsurance arrangement, under which the reinsurer pays $(1-\alpha)$ of every claim. You are given that the expected cost of claims under the new contract is the same as under the excess of loss contract. Show that the proportion retained by the insurer is $\alpha = 0.84$ to the nearest 0.01.

(e) Calculate the standard deviation of the reinsurer's claims under the proportional reinsurance contract.

(f) Comment briefly on the implications of the different standard deviations on the pricing of the two reinsurance contracts.

Question 8

The random variable X has distribution function $F(x)$ where

$$F(x) = 1 - \exp(-2x)\left(x^2 + x + 1\right) \qquad x > 0$$

(a) Derive the hazard rate function for this distribution.

(b) Comment on the tail behaviour of this distribution, compared with the exponential distribution with the same mean

(c) Derive the mean excess loss function for this distribution.

(d) The distribution lies in the MDA of the GEV distribution, with parameter ξ. State with reasons whether ξ is greater than 0, equal to 0, or less than 0.

Question 9

An insurer manages a portfolio of r commercial insurance policies.

Let X_{ij} denote the aggregate claims from insurance policy i in year *j,* for $j = 1, 2, \ldots, n$.

You are given the following information.

(i) Each X_{ij} is the average loss from $m_{i,j}$ different risks, that is $X_{ij} = \frac{1}{m_{i,j}} \sum_{l=1}^{m_{ij}} Y_{ijl}$

(ii) The individual risks, $Y_{i,j,l}$, are dependent on an unknown parameter, θ_i .

(iii) $Y_{i,j,l} \mid \theta_i$ are i.i.d for $l = 1, 2, ..., m_{i,j}$, and for $j = 1, 2, ..., n$.

(iv) $E\left[Y_{i,j,l} \mid \theta_i\right] = \mu(\theta_i)$ and $\text{Var}\left[Y_{i,j,l} \mid \theta_i\right] = v(\theta_i)$

(v) θ_i are i.i.d, with $\mu = E\left[\mu\left(\theta_i\right)\right]$, $v = E\left[v\left(\theta_i\right)\right]$, for $i = 1, 2, \cdots, r$.

(a)

(i) Write down the assumptions of the Bühlmann-Straub model, in terms of $\mu(\theta_i)$ and $v(\theta_i)$.

(ii) Show that the assumptions for the Bühlmann-Straub model are valid in this case.

(b)

(i) Write down the formula for Z_i, the credibility factor for insurance policy *i,* defining all the terms.

(ii) Explain why it is reasonable that a larger value of $\hat{a}$ generates a larger value for Z_i.

(iii) Explain why it is reasonable that a larger value of $\hat{v}$ generates a smaller value for Z_i.

(iv) Explain why it is reasonable that a larger value of m_i generates a larger value for Z_i.

(c)

The estimator for μ is $\hat{\mu} = \frac{\sum_{i=1}^{r} Z_i \overline{X}_i}{\sum_{i=1}^{r} Z_i}$

(i) Show that that the estimator for μ is unbiased.

(ii) The weights used in the estimator for μ are proportional to $\text{Var}[\overline{X}_i]$. Explain why this weighting is preferred to other possible weightings.

Question 10

For a portfolio of short-term insurance contracts, claim amounts are modeled with a density function

$$f(x;\lambda) = \lambda^2 x e^{-\lambda x} \quad x > 0,\ \lambda > 0$$

A random sample of n claim amounts, $x_1, x_2, \cdots, x_n$, is used to estimate the parameter λ.

(a) Derive and simplify as far as possible a formula for the maximum likelihood estimate of λ.

(b) Use the moment generating function to show that $Y = 2n\lambda\overline{X}$ has a chi-square distribution and state the degrees of freedom.

(c) Determine the Fisher information function, $I(\lambda)$ and use it to derive a formula for the asymptotic variance of $\widehat{\lambda}$.

(d) A random sample of 10 claim amounts resulted in a sample mean of $\overline{x} = 358$.

(i) Use the results in (b) to calculate an exact 95% confidence interval of λ.

(ii) Use the results in (c) to calculate an approximate (asymptotic) 95% confidence interval of λ.

(e) Comment briefly on the comparison of these confidence intervals.

Question 11

For a portfolio of 2000 policyholders of automobile insurance, you are given the observed number of claims per policyholder for a given period:

Number of claims	0	1	2	3	4	5+
Number of policyholders	1625	307	58	9	1	0

(a) Assume first that N, can be modeled as a Poisson random variable.

(i) Calculate the maximum likelihood estimate of the Poisson parameter.

(ii) Calculate the expected number of policyholders for each possible number of claims using the fitted Poisson distribution.

(iii) Calculate the chi-square test statistic for testing the goodness-of-fit of the Poisson model to the observed data.

(b) Now assume that the distribution of N is given by

$$\Pr\left[N = n\right] = \pi(1-\pi)^n \quad n = 0,1,2,\cdots \quad 0 < \pi < 1$$

(i) Calculate the maximum likelihood estimate of π.

(ii) Calculate the expected number of policyholders for each possible number of claims using the fitted distribution.

(iii) Calculate the chi-square test statistic for testing the goodness-of-fit of the given model to the observed data.

(c) Compare the suitability of the two models to the data.

Question 12

You are the actuary with the task of fitting a loss model to a set of 533 observed claims, over a two-year period, from a portfolio of liability insurance policies with very long tails. There is no censoring or truncation in the observed claims.

You are thinking of fitting a Burr XII model with distribution function of the form

$$F(x) = 1 - \left(\frac{\theta}{\theta + x^{\gamma}}\right)^{\alpha}, \ x > 0,$$

where the parameters θ, γ, α are all positive. Observe that the Pareto model is a special case of this model where the parameter γ is fixed at 1.

You fitted both the Burr XII and the Pareto models to the data and the maximum likelihood estimates together with other useful statistics are given:

Model	Parameters	Estimates (MLE)	Standard errors	Log-likelihood
Pareto	α	1.1819	0.1115	–3515.941
	θ	136.6418	21.1313	
Burr XII	α	1.1931	0.2019	–3526.253
	θ	138.6682	36.5152	
	γ	0.9956	0.0639	

(a) Briefly describe four steps you would follow to use the data to select a parametric loss model for the severity.

(b) Explain what the standard errors are in the table above, how these are to be used, and how they have been calculated.

(c) Conduct a hypothesis test to examine whether $\gamma = 1$ based on the model estimates in the Burr XII model.

(d) Calculate the Schwarz Bayesian Criterion (SBC) for each model. Explain what this criterion is used for and how you would choose between the Pareto and the Burr XII model based on this criterion.

(e) Briefly summarize your findings whether the Pareto or the Burr XII is a better model to the data.

Question 13

An insurance company offers telematics insurance to its automobile insurance policyholders. For a portfolio with n new policies, a total of M claims is recorded in the first t_0 months of cover. You may assume that each policy gives rise to a maximum of one claim and that claims arise independently of the others.

Assume that for each policy, the time (measured in months) until a claim occurs is modeled as a random variable T, with

$$\Pr\left[T \le t\right] = F(t) = 1 - e^{-\beta t} \quad t > 0, \quad \beta > 0$$

(a) Write down the likelihood of observing M claims in t_0 months.

(b) Determine the maximum likelihood estimate of β in terms of M, t_0, and n.

(c) Explain, but do not perform the calculations, how you could construct an approximate confidence interval for β.

(d) Explain carefully how you can use your estimated value of β to estimate the average time until a claim occurs, denoted μ.

(e) Explain carefully how you can estimate the variance of your estimate for μ.

Question 14

Two severity models, denoted by X_1 and X_2, are created to estimate a ground-up loss. You are given the following information.

(i) X_1 has a density function

$$f_1(x) = 0.001\beta e^{-0.001x} + 0.002(1-\beta)e^{-0.002x}, \qquad x > 0,$$

where β is a constant between 0 and 1.

(ii) The density of X_2 is proportional to the density of an exponential distribution with mean 500 in $(0,\ 500)$ and proportional to the density of a two-parameter Pareto distribution with $\alpha = 3$ and $\theta = 2000$ in $(500,\ \infty)$.

(iii) The probability that the loss is no more than 500 is the same under both models.

You are given that the mean of the ground-up loss is 800, and β is determined by matching the first moment.

(a) Show that $\beta = 0.6$.

(b) If X_1 is used to model the ground-up loss, calculate the mean excess loss over a deductible of 200.

(c) If X_1 is used to model the ground-up loss, calculate the Value at Risk of the loss at the 90% security level.

(d) Determine the density function of X_2

(e) If X_2 is used to model the ground-up loss, calculate the Value at Risk of the loss at the 90% security level.

(f) Compare the tails of the two severity models. Justify your answer.

Question 15

The ground-up loss severity random variable for a liability insurance risk is denoted Y. Conditioning on $\Lambda = \lambda$, Y follows an exponential distribution with mean $1/\lambda$, and Λ is gamma distributed with $\alpha = 4$, and $\theta = 0.002$.

The number of losses has a negative binomial distribution with parameters $r = 2$ and $\beta = 1.5$.

Suppose that an ordinary deductible of 100 is applied to each individual loss.

(a) Show that the unconditional distribution function of the ground-up loss is

$$F_Y(y) = 1 - \left(\frac{500}{500 + y}\right)^4, \qquad y \geq 0.$$

(b) Show that the distribution for the number of payments is negative binomial with $r = 2$, and $\beta^* = 0.72$ to the nearest 0.01. You should calculate the value to the nearest 0.001

(c) Show that the distribution of the amount paid per payment Y^P is given by

$$F_{Y^P}(y) = 1 - \left(\frac{600}{600 + y}\right)^4, \qquad y \geq 0.$$

(d) Using the method of rounding, with a span of $h = 50$, calculate the first three values of the probability function of the discretized distribution for Y^P.

(e) Use the recursive formula, with the discretized probabilities from (d), to estimate the probability that aggregate payments are less than or equal to 100.

Question 16

XYZ Insurer has two lines of insurance business with their aggregate claim payments denoted by S_1 and S_2, respectively.

- S_1 and S_2 are independent and both follow compound Poisson processes.
- S_1 has a Poisson primary distribution with mean $\lambda_1 = 2$ and severity probabilities 0.3 on a payment of 0 and 0.7 on a payment of 2.
- S_2 has a Poisson primary with mean $\lambda_2 = 1$ and severity probabilities 0.6 on a payment of 0 and 0.4 on a payment of 4.

(a) Use the convolution formula to derive the probability function for S_1.

(b) Show that $S = S_1 + S_2$ is also a compound Poisson process. Identify the Poisson primary λ and the secondary distribution.

(c) Calculate the probability that the total claim payment amount from XYZ Insurer's all business is 4.

(d) XYZ Insurer is considering the purchase of stop-loss reinsurance, with an aggregate deductible of 4 on its entire business.

Determine the net stop-loss reinsurance premium.

Question 17

You are given the following data from a sample of 6 claims. The ground up loss for the risk is assumed to be exponential with mean θ.

Loss Amount	Modification
120	None
640	None
700	None
820	\$100 deductible
1220	\$100 deductible
1500	Censored at \$1500

(a) Write down the likelihood function for this sample and simplify as far as possible.

(b) Show that the Maximum Likelihood Estimate of the parameter θ is $\hat{\theta} = 960$.

(c) Estimate the standard deviation of $\hat{\theta}$ using Cramer-Rao lower bound.

(d) Let Y be the random variable representing the insured loss amount per claim with a limit of \$1000 and no deductible. Use your fitted model to estimate $\mathrm{E}[Y]$.

(e) Using the delta method, construct an approximate 80% confidence for $\mathrm{E}[Y]$.

Question 18

A claims study examines the frequency of automobile claims among 400 policies. An actuarial assistant has fit several frequency models (including Zero-Modified distributions) to the claims data using Maximum Likelihood Estimation. Summarized results are shown below.

	Poisson	***ZM Poisson***	***Geometric***	***ZM Geometric***	***Neg Binomial***	***ZM Neg Binomial***
Parameters	1	2	1	2	2	3
p_0		0.5625		0.5625		0.5625
λ	0.6175	0.7339				
r					2.1894	0.4963
β			0.6175	0.4114	0.2853	0.5715
Log-Likelihood function	-427.8	-426.2	-430.2	-423.2	-424.5	-423.1
SBC Statistic	-430.8	-432.2	-433.2	-429.2	-430.5	-432.0
χ^2 ***Statistic***	6.07	3.29	14.53	1.26	5.15	1.14
χ^2 ***Degrees of Freedom***	2	1	3	1	1	1
χ^2 ***P-value***						

(a) Describe how the SBC statistic is calculated and verify the calculation of the SBC for the Negative Binomial model.

(b) Perform a Likelihood ratio test to compare the ZM Poisson to the Poisson.

$$H_0: \ population\ is\ a\ Poison$$
$$H_1: \ population\ is\ a\ ZM\ Poison$$

(c) Give ranges for the χ^2test p-values for each model.

(d) What other criteria could be used to select a potential model?

(e) Using the above scores and any other criteria, pick a preferred model and justify your choice.

Question 19

You are fitting a lognormal distribution to a sample of size with complete individual data.

(a) Show that $\hat{\mu} = \frac{1}{n}\sum_{j=1}^{n} ln(x_j)$

(b) Show that $\hat{\sigma}^2 = \frac{1}{n}\sum_{j=1}^{n}\left(ln(x_j) - \mu\right)^2$

(c) Show that the estimate of the Covariance matrix of the lognormal parameters is

$$\widehat{Cov}(\hat{\mu}, \hat{\sigma}) = \begin{vmatrix} \frac{\hat{\sigma}^2}{n} & 0 \\ 0 & \frac{\hat{\sigma}^2}{2n} \end{vmatrix}$$

(d) Use the delta method to estimate the variance of mean of the population based on the lognormal model. State your answer in terms of the parameter estimates $\hat{\mu}$ and $\hat{\sigma}$.

Question 20

You are given the following run-off triangle of incremental claim payments, X_{ij},. Assume all claims are settled by the end of development year 4.

	Development Year (DY), j				
Accident Year (AY) i	**0**	1	2	3	4
0	1023	132	188	43	10
1	1358	350	197	25	
2	1283	283	172		
3	1503	508			
4	1536				
$\hat{\beta}_j$	0.6931	?	0.9725	0.9928	1.000
$\hat{\gamma}_j$	0.6931	?	0.1086	0.0204	0.0072

(a)

(i) Show that $\hat{\beta}_1 = 0.86$ to the nearest 0.01. You should calculate the value to the nearest 0.001.

(ii) Show that $\hat{\gamma}_1 = 0.17$ to the nearest 0.01. You should calculate the value to the nearest 0.001.

(b) Calculate the estimated outstanding claims using the chain ladder method.

(c) Write down the three assumptions of the Bühlmann-Straub model of outstanding claims.

(d) Using the Bühlmann-Straub model,

(i) Calculate s_2^2.

(ii) You are given that $\hat{v} = 30,582$. Calculate $\hat{a}$.

(iii) Describe briefly what $\hat{v}$ and $\hat{a}$ are measuring.

(iv) You are given that $\hat{\mu} = 1926.2$. Calculate $\hat{C}_{2,4}^{\mathrm{BS2}}$.

(e) Explain briefly how the Bühlmann-Straub model can be seen as a special case of the Bornhuetter-Ferguson model.

Question 21

For all policies in an auto insurance portfolio, the number of accidents in a year follows a Poisson distribution with unknown parameter θ. The accident rate depends on the driving quality of the policyholder as described in the table below. The driving quality is not known in advance to the insurer.

Quality	Number of accidents per year	Proportion of population
Good	Poisson(1)	0.6
Fair	Poisson(2)	0.3
Poor	Poisson(5)	0.1

(a) For a random sample of 20 drivers from the population, calculate the expected number of accidents.

(b) Calculate the standard deviation of the aggregate claims cost per year for a policyholder selected at random, assuming that for all claims the severity distribution has mean 1000 and standard deviation 2000.

(c) A policyholder has a claims experience of 3 claims (in total) in the past 3 years. Determine the posterior distribution of θ for this policyholder.

(d) Calculate the Bayesian estimate of annual claim frequency for this policyholder.

(e) Identify the following functions/parameters used to calculate the Bühlmann credibility estimate of claim frequency, in the context of the parameters and distributions given above.

  (i) $\mu(\theta)$; $v(\theta)$

  (ii) μ; v; a

(f) Calculate the Bühlmann Credibility estimate of claim frequency for the policyholder in part (c).

(g) Your boss says, "It looks like this policyholder is a low-risk driver, so the correct risk premium should assume that they fall in the 'good driver' category." Comment on the appropriateness of this suggestion.

Question 22

This question is an Excel question. Please see the ASTAM Sample Excel Question file for details. The content of the Excel file for this question is provided below for reference.

You are given the following cumulative claims paid to date for 10 accident years in a particular line of business (LOB). Assume that claims are fully developed by the end of development year (DY) 9.

(a) Estimate the total aggregate outstanding claims from all open accident years, using the chain ladder method. Assume no discounting or inflation.

Cumulative Claims Paid											
AY	DY, j										$\hat{R}_i$
i	0	1	2	3	4	5	6	7	8	9	
0	10100	21056	26711	30594	32774	34094	35010	35153	35295	35320	
1	10258	20337	26108	30312	32877	34053	34731	34932	35043		
2	11474	22742	29816	34581	36827	37826	38414	38569			
3	11482	22675	29987	34063	36038	36841	37335				
4	11567	24416	31561	35351	37186	38383					
5	12755	25939	32694	36548	38618						
6	13575	26229	32698	36707							
7	12736	24471	31704								
8	12878	24721									
9	12621										
$\hat{f}_j$										X	X

(b) (i) Estimate the parameter $\hat{\beta}_5$ for this data.

(ii) Interpret the parameter $\hat{\beta}_5$.

(c) Calculate the development factors $f_{i,j}$ associated with this data, for $i + j \leq 8$ and $i \leq 7$ and $j \leq 6$.

(d) Calculate the Pearson correlations between vectors of development factors for successive development years, separately, up to DY 5 and DY 6.

(e) Calculate the appropriate test statistic and corresponding p-value for each of the correlations in part (d).

(f) State one reason why the *t*-test for Pearson correlation of development factors may not be suitable.