Statistical and Risk Modelling – 2

Time: 2 hours
Total Marks: 60 marks

Note:

1. **The candidate has option to either attempt question 4A or question 4B. Rest all questions are mandatory.**
2. **Numbers to the right indicate full marks.**
3. **The candidates will be provided with the formula sheet and graph papers (if required) for the examination.**
4. **Use of approved scientific calculator is allowed.**

Q1.A (5 Marks)

Bornhuetter – Ferguson:

- The first accident year is fully run off.
- The estimated loss ratio is correct.
- Payments from each accident year will develop in the same way.
- Inflation is not allowed for explicitly, rather it is allowed for implicitly as a weighted average of past inflation.

Average Cost Per Claim:

- The first accident year is fully run off.
- The average cost per claim in each development year is a constant proportion in monetary terms of the ultimate average cost per claim for each accident year.
- The number of claims in each development year is a constant proportion in of the ultimate number of claims for each accident year.
- Inflation is not allowed for explicitly, rather it is allowed for implicitly as a weighted average of past inflation.

Statistical and Risk Modelling – 2

Time: 2 hours
Total Marks: 60 marks

Q1.B (5 Marks)

In usual notations, we start with:

$$\lambda + cR = \lambda M_x(R)$$
$$= \lambda \int_0^M \exp exp\, (Rx)\, f(x) dx$$
$$\leq \lambda \int_0^M \left(\frac{x}{M} \exp exp\, (RM) \;+\; 1 - \frac{x}{M} \right) f(x) dx$$

(using the given inequality)

$$\leq \frac{\lambda}{M} \exp exp\, (RM)\, m_1 + \lambda - \frac{\lambda}{M} m_1$$

$$\frac{c}{\lambda m_1} \leq \frac{1}{RM} (\exp exp\, (RM) \;-1) = 1 + \frac{RM}{2} + \frac{(RM)^2}{3!} + \ldots < 1 + \frac{RM}{2} + \frac{(RM)^3}{3!} + \ldots$$

Hence,

$$R > \frac{1}{M} log(\frac{c}{\lambda m_1})$$

Statistical and Risk Modelling – 2

Time: 2 hours
Total Marks: 60 marks

Q1.C (5 Marks)

Let X be a random variable representing the claim amounts from a portfolio of Fire Insurance policies. Let the pdf of X be given by:

$$f(x) = \frac{\alpha\beta\gamma^{\alpha}x^{\beta-1}}{\left(\gamma+x^{\beta}\right)^{\alpha+1}}\ ; x > 0$$

Derive an expression for the k^{th} raw moment of X.

$$E\left(X^k\right) = \int_0^{\infty} x^k \frac{\alpha\beta\gamma^{\alpha}x^{\beta-1}}{\left(\gamma+x^{\beta}\right)^{\alpha+1}}\, dx$$

Let,

$$\frac{\lambda}{\lambda+x^{\gamma}} = t$$

Then,

$$E\left(X^k\right) = \alpha\lambda^{\frac{k}{\gamma}} \int_0^1 (1 - t)^{\frac{k}{\gamma}}\, t^{\alpha-1-\left(\frac{k}{\gamma}\right)}\, dx$$

…(1)

Using Beta distribution, we have:

$$\int_0^{\infty} x^{\alpha-1} (1 - x)^{\beta-1}\, dx = (\Gamma\alpha + \Gamma\beta) / (\Gamma(\alpha + \beta))$$

Comparing it with (1), we have:

$$E\left(X^k\right) = \frac{\lambda^{\frac{k}{\gamma}}\left(\Gamma\left(\alpha-\frac{k}{\gamma}\right)*\Gamma\left(\frac{k}{\gamma}+1\right)\right)}{\Gamma\alpha}$$

Statistical and Risk Modelling – 2

Time: 2 hours
Total Marks: 60 marks

Q2.A (5 Marks)

Assumption:
Failure between vehicle(s) are independent of each other.

Given the information on the prior distribution:

$$\frac{\alpha}{\alpha+\beta} = 0.050$$

$$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} = 0.0015$$

Solving the above two equations, we get:

$\alpha = 1.53333$
$\beta = 29.1333$

Now, the prior PDF is proportional to:

$$p^{0.5333} * (1 - p)^{28.1333}$$

Likelihood is given by (based on Binomial Distribution):

$$L \propto p^{75} * (1 - p)^{56000-75}$$

Hence, the posterior PDF is proportional to:

$$p^{75.5333} * (1 - p)^{55953.1333}$$

This has the form of a beta distribution PDF with parameters alpha = 75.53333 and beta = 55,953.1333

Statistical and Risk Modelling – 2

Time: 2 hours
Total Marks: 60 marks

Q2.B (5 Marks)

Claims on a household insurance policy is said to follow Poisson distribution.

Creating a frequency table from the claims dataset, following is the information on the claims reported among the portfolio of 1,000 policies:

Claim Count	% of policies
0	26.6%
1	37.1%
2	21.2%
3	11.1%
4	3.2%
5	0.7%
6	0.1%

Perform a chi-square Goodness of Fit test to check of the statement made on the claim count distribution holds true.
You should clearly state the Hypothesis and your conclusion.

H0: Claim Count follows Poisson Distribution
H1: Not H0

$\hat{\lambda} = 1.297$

X	Oi	Pi	Ei	(Oi - Ei)^2 / Ei
0	266	0.2734	273.35	0.1977
1	371	0.3545	354.54	0.7646
2	212	0.2299	229.92	1.3962
3	111	0.0994	99.40	1.3536
4	32	0.0322	32.23	0.0017
5	8	0.0106	10.57	0.6232
			Total	4.3368

Chi-Sq Test Statistic = 4.3368
Chi-Sq Tab Value = Chi-Sq (5%, 6 – 1 – 1) = 9.4877

Since, Chi-Sq Test Statistic < Chi-Sq Tab Value, we have insufficient evidence to Reject H0.

Hence, the claim holds true that the distribution of Claim Count follows Poisson Distribution.

Statistical and Risk Modelling – 2

Time: 2 hours
Total Marks: 60 marks

Q2.C (5 Marks)

1.) (2 Marks)

Lundberg's inequality states that:

$$\psi(U) \leq exp(-\ RU)$$

Where:

- U is the insurer's initial surplus.
- $\psi(U)$ is the probability of ultimate ruin.
- R is a parameter associated with a surplus process known as the adjustment coefficient. Its value depends upon the distribution of aggregate claims and on the rate of premium income.

2.) (3 Marks)

We need to solve the equation:
$\lambda\ +\ cR\ = \lambda M_x(R)$

Therefore, in usual notations:

$R\ = \frac{Rate^*\ \theta}{1+\theta}$

Where Rate: the parameter of exponential distribution

Substituting the values, we get:
R = 0.170689655

Statistical and Risk Modelling – 2

Time: 2 hours
Total Marks: 60 marks

Q3.A (5 Marks)

$\overline{X} = 157.175$

$E\left(s^2(\theta)\right) = 164.55$

$V(m(\theta)) = 887.59$

$Z = 0.981798$

Credibility Premium =
Risk 1: 153.8 * Z + (1 – Z) * 157.175 = 158.8614
Risk 2: 147.9 * Z + (1 – Z) * 157.175 = 148.0688
Risk 3: 127.9 * Z + (1 – Z) * 157.175 = 128.43285
Risk 4: 199.1 * Z + (1 – Z) * 157.175 = 198.3369

Statistical and Risk Modelling – 2

Time: 2 hours
Total Marks: 60 marks

Q3.B (5 Marks)

1.) (2 Marks)

Let Y = (1 + j%) * X

$F[Y \leq y] = F\left[X \leq \frac{y}{1+j\%}\right] = 1 - exp(-\lambda * \frac{y}{1+j\%})$

Therefore, Y ~ exp($\lambda/(1 + j\%)$)

If k = (1 + j%), then Y ~ exp(λ/k)

2.) (3 Marks)

$$L \propto \left(e^{-\lambda M}\right)^{34} \left(e^{-\frac{\lambda M}{k}}\right)^{41} \left(\lambda^{21}\right)\left(e^{-\lambda * 51,675}\right)\left(\frac{\lambda}{k}\right)^{9} (e^{-\frac{\lambda}{k} * 86,455})$$

Where M = $3,500

$$L \propto \lambda^{30} * \exp exp\left(- 34\lambda M - \frac{41\lambda M}{k} - \lambda * 51,675 - \frac{\lambda}{k} * 86,455\right)$$

Take log and differentiate w.r.t λ:

d Log L / dλ = 0

Solving the equation, we get:

$$\hat{\lambda} = \frac{30}{34M + \frac{41M}{k} + 51,675 + \frac{86,455}{k}}$$

Statistical and Risk Modelling – 2

Time: 2 hours
Total Marks: 60 marks

Q3.C (5 Marks)

1.) (2 Marks)

For Pareto distribution,

$$10,000 = \frac{\lambda}{\alpha-1}$$

$$30,000^2 = \frac{\alpha\lambda^2}{(\alpha-1)^2(\alpha-2)}$$

Solving the above two equations simultaneously, we get:

$$9 = \frac{\alpha}{\alpha-2}$$

$$9\alpha - 18 = \alpha$$
$$\alpha = \frac{18}{8}$$

$\alpha = 2.25$
$\lambda = 12,500$

2.) (3 Marks)

$V(S) = \lambda * m_2 = 5 * \left[V(X) + (E(X))^2\right] = (70,710.67812)^2$
$SD(S) = 70,710.68$

Statistical and Risk Modelling – 2

Time: 2 hours
Total Marks: 60 marks

Q4.A (15 Marks)

1.) (3 Marks)

Cumulative Incurred Triangle:

UW Year x DY	**0**	**1**	**2**	**3**
2019	1,496	2,461	3,276	4,049
2020	1,068	1,811	2,424	
2021	1,593	2,531		
2022	1,105			

DY	0 to 1	1 to 2	2 to 3
DF	1.6365	1.3343	1.2360

UW Year x DY	**0**	**1**	**2**	**3**	**Ultimate**
2019	1,496	2,461	3,276	4,049	4,049
2020	1,068	1,811	2,424	2995.963	2,996
2021	1,593	2,531	3377.037	4173.877	4,174
2022	1,105	1808.351	2412.828	2982.155	2,982
				Total	**14,201.00**

2.) (5 Marks)

UW Year x DY	**0**	**1**	**2**	**3**
2019				
Actual	1,496	965	815	773
Fitted	1,496	952	818	771
Error	-	12.8	-3.4	2.1
2020				
Actual	1,068	743	613	
Fitted	1,068	680	584	
Error	-	63.2	28.7	
2021				

Actual	1,593	938	
Fitted	1,593	1,014	
Error	-	-75.9	

Overall, the error seems to be increasing in magnitude in the latest UW Years.

Statistical and Risk Modelling – 2

Time: 2 hours
Total Marks: 60 marks

3.) (3 Marks)

- Claims inflation may increase suddenly.
- Office expenses may also increase (if these are allowed for in the calculations).
- Weather patterns may change, altering the balance between short tail and long tail claims.
- The characteristics and behaviour of the insured lives in the portfolio may change over time.
- Catastrophe claims may occur.
- Underwriting procedures may change.

4.) (4 Marks)

DY	0 to 1	1 to 2	2 to 3
DF	1.6365	1.3343	1.236
Cumulative DF	2.6989	1.6492	1.236
% Developed	37.05 %	60.64 %	80.91 %

- Yes, modification to the claims incurred data is required for the pricing project.
- As we see from the above table, the cohort is not 100% developed even for UW Year 2019 as at end of 2022 (only 81% of the same is developed).
- Hence, it is less likely than policies belonging to UW Year 2018 – 2019 will be 100%.
- The adjustment required may be substantial and cannot be ignored.
- More Claims data from historical years needs to be evaluate to ensure more clarity on the adjustment required.

OR

Statistical and Risk Modelling – 2

Time: 2 hours
Total Marks: 60 marks

Q4.B (15 Marks)

1.) (0.5 Marks)

No, the statement '*Exponential Distribution is always the best distribution when compared to Weibull Distribution*' is NOT true.

2.) (3 Marks)

- Weibull reduces down to Exponential when $\gamma = 1$
- Weibull distribution provides more flexibility in the tails of the distribution.
- Given the losses are model for Motor Insurance, the tail maybe heavy / light, which the Exponential distribution may not account for.
- Hence, underestimating the large losses.

3.) (4.5 Marks)

CDF of Weibull distribution with parameters $c = 2$ and $\gamma = 0.25$
1 – F[5] = exp(-2*5^(0.25)) = 0.05025

CDF of Weibull distribution with parameters $c = 2$ and $\gamma = 1.25$
1 – F[5] = exp(-2*5^(1.25)) = 3.20467 * 10^(-7)

CDF of Weibull distribution with parameters $c = 2$ and $\gamma = 1$
1 – F[5] = exp(-2*5^(1)) = 4.53999 * 10^(-5)

When $\gamma = 1$, the PDF of Weibull reduces down to Exponential Distribution.

For different Insurance products, the tails can be heavy / light and will be of importance.
Here, we see that for $\gamma = 0.25$, the tail is heavier as compared to the tail when $\gamma = 1$
Here, we see that for $\gamma = 1.25$, the tail is lighter as compared to the tail when $\gamma = 1$

Hence, Weibull distribution provides more flexibility in the tails of the distribution.

4.) (5 Marks)
Using 'Methods Of Percentiles' find the parameter estimates given the summary information from the loss data:

- Mean (in millions): 0.484173
- Variance (in millions): 0.0647711
- First Quartile (in millions): 0.292387
- Second Quartile (in millions): 0.451039
- Third Quartile (in millions): 0.645018

Statistical and Risk Modelling – 2

Time: 2 hours
Total Marks: 60 marks

$F(0.292387) = 1 - \exp exp\left(- c * 0.292387^{\gamma}\right) = 0.25$
$F(0.645018) = 1 - \exp exp\left(- c * 0.645018^{\gamma}\right) = 0.75$

$1 - F(0.292387) = \exp exp\left(- c * 0.292387^{\gamma}\right) = 0.75$
$1 - F(0.645018) = \exp exp\left(- c * 0.645018^{\gamma}\right) = 0.25$

Taking log on both sides, we have:
$- c * 0.292387^{\gamma} = \log log\,(0.75)$ …(1)
$\left(- c * 0.645018^{\gamma}\right) = log(0.25)$…(2)

Dividing (1) by (2):

$0.20752 = 0.453300^{\gamma}$
$\gamma = 1.987527$
$c = 3.31387$

5.) (2 Marks)

$$Median = \left(\frac{\ln ln\,(2)}{c}\right)^{1/\gamma} = 0.4551$$

The empirical median = 0.451039.
This is close to the Median we calculated above.
Hence, Weibull distribution is possibly / may be a good fit for the given data.