ACSC/STAT 4703, Actuarial Models II
FALL 2015
Toby Kenney
Sample Midterm Examination
Model Solutions

This Sample examination has more questions than the actual midterm, in order to cover a wider range of questions. Estimated times are provided after each question to help your preparation.

Here are some values of the Gamma distribution function with $\theta = 1$ that will be needed for this examination:

x	α	$F(x)$
245	255	0.2697208
$\left(\frac{7.5}{12}\right)^3$	$\frac{4}{3}$	0.1117140
$\left(\frac{9.5}{12}\right)^3$	$\frac{4}{3}$	0.2507382
2.5	1	0.917915
2.5	2	0.7127025
2.5	3	0.4561869
2.5	4	0.2424239

1. *Loss amounts follow an exponential distribution with $\theta = 60,000$. The distribution of the number of losses is given in the following table:*

Number of Losses	*Probability*
0	*0.04*
1	*0.54*
2	*0.27*
3	*0.15*

Assume all losses are independent and independent of the number of losses. The insurance company buys excess-of-loss reinsurance on the part of the loss above \$150,000. Calculate the expected payment for this excess-of-loss reinsurance.

If the number of losses is n, then the aggregate loss follows a gamma distribution with $\alpha = n$ and $\theta = 60000$. The expected payment on the excess-of-loss insurance is therefore

$$\begin{aligned}
&\int_{150000}^{\infty}(x-150000)\frac{x^{n-1}e^{-\frac{x}{60000}}}{(n-1)!60000^n}\,dx\\
&=\int_{150000}^{\infty}\frac{x^{n}e^{-\frac{x}{60000}}}{(n-1)!6000^n}\,dx-150000\int_{150000}^{\infty}\frac{x^{n-1}e^{-\frac{x}{60000}}}{(n-1)!60000^n}\,dx\\
&=\int_{2.5}^{\infty}\frac{60000nu^ne^{-u}}{n!}\,du-150000\int_{2.5}^{\infty}\frac{u^{n-1}e^{-u}}{(n-1)!}\,du
\end{aligned}$$

This gives the following expected payments on the excess-of-loss reinsurance:

Number of Losses	Probability	Expected payment on excess-of-loss	product
0	0.04	0	0
1	0.54	$60000 \times 1 \times 0.2872975 - 150000 \times 0.0820850 = 4925.10$	2659.554
2	0.27	$60000 \times 2 \times 0.5438131 - 150000 \times 0.2872975 = 22162.95$	5983.996
3	0.15	$60000 \times 3 \times 0.7575761 - 150000 \times 0.5438131 = 54791.74$	8218.760

The total expected payment on the excess-of-loss reinsurance is therefore $2659.554 + 5983.996 + 8218.760 = \$16,862.31$.

2. *Aggregate payments have a compund distribution. The frequency distribution is negative binomial with $r = 4$ and $\beta = 12$. The severity distribution is a Gamma distribution with $\alpha = 8$ and $\theta = 3000$. Use a normal approximation to aggregate payments to estimate the probability that aggregate payments are more than \$200,000.*

   The frequency distribution has mean 48 and variance 624. The severity distribution has mean 24000 and variance 72000000.

   The mean of aggregate payments is therefore, $48 \times 24000 = 1152000$, and the variance is $624 \times 24000^2 + 48 \times 72000000 = 362880000000$, so the standard deviation is $\sqrt{362880000000} = 602395.2$. The probability of exceeding \$2,000,000 is therefore $1 - \Phi\left(\frac{2000000-1152000}{602395.2}\right) = 1 - Phi(1.407714) = 1 - 0.9203921 = 0.0796$.

3. *Claim frequency follows a negative binomial distribution with $r = 5$ and $\beta = 2.9$. Claim severity (in thousands) has the following distribution:*

Severity	*Probability*
0	*0*
1	*0.600*
2	*0.220*
3	*0.166*

Use the recursive method to calculate the exact probability that aggregate claims are at least 4.

For the negative binomial distribution, we have $a = \frac{\beta}{1+\beta} = \frac{2.9}{3.9}$ and $b = \frac{(r-1)\beta}{1+\beta} = \frac{4\times2.9}{3.9}$, so the recursive formula

$$f_S(x) = \frac{(p_1 - (a+b)p_0)f_X(x) + \sum_{i=1}^{x}\left(a + \frac{bi}{x}\right)f_X(i)f_S(x-i)}{1 - af_X(0)}$$

becomes

$$f_S(x) = \sum_{i=1}^{x} \frac{2.9}{3.9}\left(1 + \frac{4i}{x}\right) f_X(i)f_S(x-i)$$

Since the severity distribution has no probability at zero, the only way for the aggregate loss to be zero is if the frequency is zero, the probability of which is $\left(\frac{1}{1+\beta}\right)^r = \frac{1}{3.9}^5 = 0.00110835$. We now use the recurrence:

$$f_S(1) = \frac{2.9}{3.9} \times 5 \times 0.600 \times 0.00110835 = 0.002472473$$
$$f_S(2) = \frac{2.9}{3.9} \times (3 \times 0.600 \times 0.002472473 + 5 \times 0.220 \times 0.00110835) = 0.004215883$$
$$f_S(3) = \frac{2.9}{3.9} \times \left(\frac{7}{3} \times 0.600 \times 0.004215883 + \frac{11}{3} \times 0.220 \times 0.002472473 + 5 \times 0.166 \times 0.00110835\right) = 0.006555954$$

The probability that the aggregate payments exceed 4 is therefore $1-0.00110835-0.002472473-0.004215883-0.006555954 = 0.9856473$.

4. *Using an arithmetic distribution ($h = 1$) to approximate a Weibull distribution with $\tau = 3$ and $\theta = 12$, calculate the probability that the value is between 3.5 and 8.5, for the approximation using:*

   (a) The method of rounding.

   The method of rounding preserves this probability, since it assigns all values between 3.5 and 4.5 to 4, etc. Therefore this probability is $e^{-\left(\frac{3.5}{12}\right)^3} - e^{-\left(\frac{8.5}{12}\right)^3} = 0.2745978$.

   (b) The method of local moment matching, matching 1 moment on each interval. [$\Gamma\left(\frac{4}{3}\right) = 0.8929795$.]

   Using local moment matching, the probabilities of the intervals $[3.5, 5.5]$ and $[5.5, 7.5]$ are preserved, so the probability of these intervals is $e^{-\left(\frac{3.5}{12}\right)^3} - e^{-\left(\frac{7.5}{12}\right)^3} = 0.1921159$.

   For the interval $[7.5, 9.5]$, the probability of this interval is $e^{-\left(\frac{7.5}{12}\right)^3} - e^{-\left(\frac{9.5}{12}\right)^3} = 0.174517$, while the conditional mean times this probability is

$$\begin{aligned}
\int_{7.5}^{9.5} x\left(\frac{3x^2}{12^3}e^{-\left(\frac{x}{12}\right)^3}\right)dx &= \int_{\left(\frac{7.5}{12}\right)^3}^{\left(\frac{9.5}{12}\right)^3} 12\sqrt[3]{u}e^{-u}\,du \\
&= 12\int_{\left(\frac{7.5}{12}\right)^3}^{\left(\frac{9.5}{12}\right)^3} u^{\frac{1}{3}}e^{-u}\,du \\
&= 12\Gamma\left(\frac{4}{3}\right)(0.2507382 - 0.1117140) \\
&= 12 \times 0.8929795 \times (0.2507382 - 0.1117140) \\
&= 1.489749
\end{aligned}$$

   We are now trying to solve for p_8 and p_9 such that

$$\begin{aligned}
p_8 + p_9 &= 0.174517 \\
8p_8 + 9p_9 &= 1.489749 \\
p_8 &= 9 \times 0.174517 - 1.489749 = 0.080904
\end{aligned}$$

So the probability of the interval $[3.5, 8.5]$ is therefore $0.1921159 + 0.080904 = 0.273020$.

5. *An insurance company has the following portfolio of home insurance policies:*

Type of driver	*Number*	*Probability claim*	*mean of claim*	*standard deviation*
Good driver	*600*	*0.02*	*$2,500*	*$2,000*
Average driver	*1400*	*0.06*	*$3,800*	*$3,200*
Bad driver	*500*	*0.13*	*$7,000*	*$3,600*

Calculate the cost of reinsuring losses above $5,000,000, if the loading on the reinsurance premium is one standard deviation above the expected claim payment on the reinsurance policy, using a Pareto approximation for the aggregate losses on this portfolio.

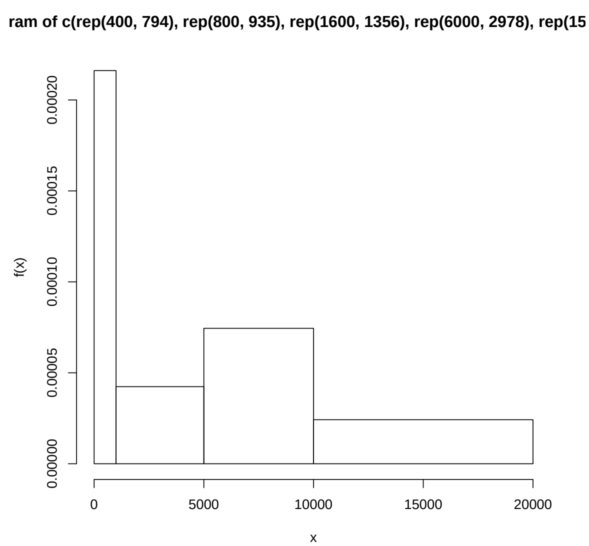
6. *For the following dataset:*

0.2 0.2 0.4 0.7 1.8 2.1 2.3 3.0 3.5 3.9 4.1 4.2 4.6 5.1 5.7 6.6 8.2 11.4

Calculate a Nelson-Åalen estimate for the probability that a random sample is more than 2.7.

The Nelson-Åalen estimate for $H(2.7)$ is $\frac{2}{18} + \frac{1}{16} + \frac{1}{15} + \frac{1}{14} + \frac{1}{13} + \frac{1}{12} = 0.4719628$, so the survival functions is $S(2.7) = e^{-0.4719628} = 0.6237767$.

7. *The histogram below is obtained from a sample of 8,000 claims.*



Which interval included most claims?

The probability of the first interval is approximately $0.00022 \times 1000 = 0.22$.

The probability of the second interval is approximately $0.00004 \times 4000 = 0.16$.

The probability of the third interval is approximately $0.00007 \times 5000 = 0.35$.

The probability of the first interval is approximately $0.00003 \times 10000 = 0.30$.

Therefore the interval 5000–10000 included most claims.

[The actual numbers are 1729, 1356, 2978 and 2037, with probabilities 0.216125, 0.169500, 0.372250 and 0.254625 respectively.]

8. *An insurance company collects the following data on insurance claims:*

Claim Amount	*Number of Policies*
Less than \$5,000	*232*
\$5,000–\$20,000	*147*
\$20,000–\$100,000	*98*
More than \$100,000	*23*

The policy currently has no deductible and a policy limit of \$100,000. The company wants to determine how much would be saved by introducing a deductible of \$2,000 and a policy limit of \$50,000. Using the ogive to estimate the empirical distribution, how much would the expected claim amount be reduced by the new deductible and policy limit?

Using the ogive, the current expected claim amount is $\frac{232\times2500+147\times12500+98\times60000+23\times100000}{500} = \$21,195.00$.

Using the ogive, the expected number of claims between 0 and 2000 is $\frac{2000}{5000}\times 232 = 92.8$. The number of claims above 50000 is $\frac{50000}{80000}\times 98 + 23 = 84.25$. The expected claim amount per loss is therefore $\frac{140.8\times1500+147\times10500+36.75\times33000+84.25\times50000}{500} = \$14,359.90$, so the reduction is $21195.00 - 14359.90 = \$6,835.10$.

9. *An insurance company collects the following claim data (in thousands):*

i	d_i	x_i	u_i	i	d_i	x_i	u_i	i	d_i	x_i	u_i
1	*0*	*0.8*	-	*8*	*0.5*	-	*5*	*15*	*2.0*	-	*5*
2	*0*	*1.3*	-	*9*	*1.0*	*1.2*	-	*16*	*2.0*	-	*10*
3	*0*	-	*20*	*10*	*1.0*	-	*15*	*17*	*2.0*	*2.4*	-
4	*0*	*4.4*	-	*11*	*1.0*	*1.8*	-	*18*	*2.0*	-	*5*
5	*0*	-	*10*	*12*	*1.0*	-	*10*	*19*	*2.0*	*11.6*	-
6	*0.5*	*1.4*	-	*13*	*1.0*	*6.3*	-	*20*	*5.0*	-	*15*
7	*0.5*	*1.8*	-	*14*	*2.0*	*4.9.*	-	*21*	*5.0*	*5.9*	-

Using a Kaplan-Meier product-limit estimator:

(a) estimate the probability that a random loss exceeds 3.

y_i	s_i	r_i
0.8	1	8
1.2	1	12
1.3	1	11
1.4	1	10
1.8	2	9
2.4	1	13

So the Kaplan-Meier estimator is $S(3) = \frac{7}{8}\times\frac{11}{12}\times\frac{10}{11}\times\frac{9}{10}\times\frac{7}{9}\times\frac{12}{13} = \frac{49}{104} = 0.4711538$.

(b) Use Greenwood's approximation to obtain a 95% confidence interval for the probability that a random loss exceeds 3, based on the Kaplan-Meier estimator, using a normal approximation.

Greenwood's approximation gives $\mathrm{Var}(\hat{S}(3)) = (\hat{S}(3))^2\sum_{i=1}^{6}\frac{s_i}{r_i(r_i-s_i)} = 0.4711538^2\left(\frac{1}{8\times7}+\frac{1}{12\times11}+\frac{1}{11\times10}+\frac{1}{10\times9}+\frac{2}{9\times7}+\frac{1}{13\times12}\right)$ 0.01860047.

Using a normal approximation, the confidence interval is $0.4711538 \pm 1.96\sqrt{0.01860047} = [0.2038421, 0.7384655]$.

(c) Use Greenwood's approximation to find a log-transformed confidence interval for the probability that a random loss exceeds 3.

The log-transformed inteval is $[S_n(x)^{\frac{1}{U}}, S_n(x)^{\frac{1}{U}}]$, where $U = e^{1.96\left(\frac{\sqrt{0.01860047}}{S_n(x)\log(S_n(x))}\right)}$. That is

$$\begin{aligned} U &= e^{1.96\left(\frac{\sqrt{0.01860047}}{0.4711538\log(0.4711538)}\right)} \\ &= 0.4705326 \end{aligned}$$

so the confidence interval is

$$[0.4711538^{\frac{1}{0.4705326}}, 0.4711538^{0.4705326}] = [0.2020174, 0.7017985]$$

10. *An insurance company records the following data in a mortality study:*

entry	*death*	*exit*	*entry*	*death*	*exit*	*entry*	*death*	*exit*
51.3	-	*58.4*	*56.5*	-	*58.2*	*55.3*	-	*59.9*
54.7	-	*59.7*	*54.7*	-	*59.8*	*53.3*	*59.1*	-
53.8	-	*58.5*	*57.9*	-	*61.3*	*56.7*	*58.4*	-
57.3	-	*58.3*	*58.0*	-	*59.3*	*52.4*	*58.9*	-
52.8	-	*60.6*	*58.4*	-	*59.8*	*57.7*	*58.8*	-
58.7	-	*59.5*	*53.0*	-	*58.3*	*58.3*	*60.4*	-
53.3	-	*62.4*	*53.1*	-	*60.1*	*58.1*	*58.4*	-

Estimate the probability of an individual currently aged exactly 58 dying within the next year using:

(a) the exact exposure method.

The exact exposure is $0.4+0.3+0.5+0.3+1+0.3+1+0.2+1+1+1+0.6+0.3+1+1+1+0.4+0.9+0.8+0.7+0.3 = 14$, and there are 4 deaths at age 58, so the hazard rate is $\frac{4}{14}$, and the probability of dying is therefore $1 - e^{-\frac{4}{14}} = 0.2485227$.

(b) the actuarial exposure method.

The actuarial exposure is $0.4+0.3+0.5+0.3+1+0.3+1+0.2+1+1+1+0.6+0.3+1+1+1+1+1+1+0.7+0.9 = 15.5$, so the probability of dying is $\frac{4}{15.5} = 0.2580645$.

11. *An insurance company observes the following claims (in thousands):*

3.1 4.6 17.2 6.5 3.8 2.0 5.7 9.2 8.3 7.5 9.8 3.2 6.1 5.8 9.2 3.7 4.4

using a kernel density estimate with a uniform kernel with bandwidth 0.8, estimate the expected payment per claim if the company introduces a deductible of 1.5 on each policy.

With a bandwidth of 0.8, the only data point for which the kernel has some probability of being smaller than the deductible is 2.0. For this data point, after the deductible is applied, there is probability $\frac{0.3}{1.6}$ that the payment is zero, and conditional on the payment being non-zero, it is uniformly distributed on the interval $[0, 1.3]$, so the expected payment is $\frac{1.3}{1.6} \times 0.65 = 0.528125$. The total expected payment per loss is therefore

$$\frac{1.6+3.1+15.7+5.0+2.3+0.528125+4.2+7.7+6.8+6.0+8.3+1.7+4.6+4.3+7.7+2.2+2.9}{17} = \frac{84.628125}{17} = 4.978125$$

The probability of a claim is $\frac{16+\frac{1.3}{1.6}}{17}$, so the expected payment per claim is $\frac{84.628125}{16+\frac{1.3}{1.6}} = 5.033643$.

12. *Using the following table:*

Age	*No. at start*	*enter*	*die*	*leave*	*No. at next age*
48	*26*	*43*	*2*	*13*	*54*
49	*54*	*39*	*7*	*17*	*69*
50	*69*	*46*	*14*	*28*	*73*
51	*73*	*22*	*13*	*44*	*38*

Estimate the probability that an individual aged 49 withdraws from the policy within the next two years, conditional on surviving to the end of those two years.

Using exact exposure

For age 49, the exact exposure is $54 + \frac{39-7-17}{2} = 61.5$, and the number of withdrawls is 17, so the hazard rate is $\frac{17}{61.5}$. For age 50, the exact exposure is $69+\frac{46-14-28}{2} = 71$, and the number of withdrawls is 28, so the hazard rate is $\frac{28}{71}$. The probability of not withdrawing in the next two years is therefore $e^{-\frac{17}{61.5}-\frac{28}{71}} = 0.511305$, so the probability of withdrawing during the next two years is $1 - 0.511305 = 0.488695$.

Using actuarial exposure

For age 49, the actuarial exposure is $54+\frac{39-7}{2} = 70$, and the number of withdrawls is 17, so the probability of withdrawing is $\frac{17}{70}$. For age 50, the actuarial exposure is $69+\frac{46-14}{2} = 85$, and the number of withdrawls is 28, so the probability of withdrawl is $\frac{28}{85}$. The probability of not withdrawing in the next two years is therefore $\frac{53}{70} \times 5785 = 0.5164103$, so the probability of withdrawing during the next two years is $1 - 0.5164103 = 0.4835897$.

13. *An insurance company models number of claims an individual makes in a year as following a Poisson distribution with Λ an unknown parameter with prior distribution a gamma distribution with $\alpha = 3$ and $\theta = 0.12$.*

(a) What is the probability that a random individual makes exactly 2 claims?

For fixed $\Lambda = \lambda$, the probability of making 2 claims is $e^{-\lambda}\frac{\lambda^2}{2}$, so the marginal probability is

$$
\begin{aligned}
&\int_0^\infty e^{-\lambda}\frac{\lambda^2}{2}\frac{\lambda^2}{2\times 0.12^3}e^{-\frac{\lambda}{0.12}}\,d\lambda\\
&=\int_0^\infty e^{-\left(1+\frac{1}{0.12}\right)\lambda}\frac{\lambda^4}{2\times 2\times 0.12^3}\,d\lambda\\
&=\frac{4!\left(\frac{1}{1+\frac{1}{0.12}}\right)^5}{2\times 2\times 0.12^3}\\
&=\frac{6\left(\frac{0.12}{1.12}\right)^5}{0.12^3}\\
&=6\frac{0.12^2}{1.12^5}\\
&=0.04902568
\end{aligned}
$$

(b) The company observes the following claim frequencies:

Number of claims	*Frequency*
0	*234*
1	*104*
2	*44*
3	*12*
4	*6*

What is the posterior probability that $\Lambda > 0.6$*?*

The total number of claims observed is 252, and the total number of years observed is 400. The likelihood of these data is proportional to $\lambda^{252}e^{-400\lambda}$. The posterior density of *Lambda* is therefore proportional to $\lambda^{252}e^{-400\lambda}\lambda^2 e^{-\frac{\lambda}{0.12}} = \lambda^{254}e^{-\left(400+\frac{1}{0.12}\right)\lambda}$ so the posterior distribution is a Gamma distribution with $\alpha = 255$ and $\theta = \frac{0.12}{49}$.

The probability that $\Lambda > 0.6$ is therefore the probability that a gamma distribution with $\alpha = 255$ and $\theta = \frac{0.12}{49}$ is more than 0.6, which is the probability that a gamma distribution with $\alpha = 255$ and $\theta = 1$ is more than $\frac{0.6\times 49}{0.12}$, which is $1 - 0.2697208 = 0.7302792$.

(c) Calculate the predictive probability that an individual makes no claims next year.

The conditional probability that an individual makes no claims is $e^{-\lambda}$, so the marginal probability is

$$
\begin{aligned}
\int_0^\infty \frac{e^{-\lambda}\lambda^{254}e^{-\frac{49}{0.12}\lambda}}{254!\left(\frac{0.12}{49}\right)^{255}}\,d\lambda &= \frac{254!\left(\frac{0.12}{49.12}\right)^{255}}{254!\left(\frac{0.12}{49}\right)^{255}}\\
&=\left(\frac{49}{49.12}\right)^{255}\\
&=0.5359436
\end{aligned}
$$

14. An insurance company models loss sizes as following a Weibull distribution with $\tau = 3$, and finds that the posterior distribution for Θ is a Pareto distribution with $\alpha = 4$ and $\theta = 1100$. Calculate the Bayes estimate for Θ based on a loss function:

(a) $l(\hat{\theta}, \theta) = (\hat{\theta} - \theta)^2$

The Bayes estimate for this loss function is the posterior mean, which is $\frac{1100}{4-1} = 366.667$.

(b) $l(\hat{\theta}, \theta) = |\hat{\theta} - \theta|^3$

(i) 422.35

(ii) 494.30

(iii) 560.87

(iv) 616.47

For this loss function, we are trying to maximise

$$\begin{aligned}
\mathbb{E}(|\hat{\theta} - \theta|^3) &= \int_0^\infty |\hat{\theta} - \theta|^3 \pi_{\Theta|X}(\theta)\,d\theta \\
&= \int_{\hat{\theta}}^\infty (\theta - \hat{\theta})^3 \pi_{\Theta|X}(\theta)\,d\theta - \int_0^{\hat{\theta}} (\theta - \hat{\theta})^3 \pi_{\Theta|X}(\theta)\,d\theta \\
&= \int_{\hat{\theta}}^\infty (\theta - \hat{\theta})^3 \frac{4 \times 1100^4}{(\theta + 1100)^5}\,d\theta - \int_0^{\hat{\theta}} (\theta - \hat{\theta})^3 \frac{4 \times 1100^4}{(\theta + 1100)^5}\,d\theta
\end{aligned}$$

Setting the derivative of this with respect to $\hat{\theta}$ to zero gives (after cancelling constants)

$$\int_{\hat\theta}^{\infty} \frac{(\theta-\hat\theta)^2}{(\theta+1100)^5}\,d\theta = \int_0^{\hat\theta} \frac{(\theta-\hat\theta)^2}{(\theta+1100)^5}\,d\theta$$

$$\int_{\hat\theta}^{\infty} \frac{\theta^2-2\hat\theta\theta+\hat\theta^2}{(\theta+1100)^5}\,d\theta = \int_0^{\hat\theta} \frac{\theta^2-2\hat\theta\theta+\hat\theta^2}{(\theta+1100)^5}\,d\theta$$

$$\int_{\hat\theta}^{\infty} \frac{1}{(\theta+1100)^3} - \frac{(2200+2\hat\theta)}{(\theta+1100)^4} + \frac{(1100+\hat\theta)^2}{(\theta+1100)^5}\,d\theta = \int_0^{\hat\theta} \frac{1}{(\theta+1100)^3} - \frac{(2200+2\hat\theta)}{(\theta+1100)^4} + \frac{(1100+\hat\theta)^2}{(\theta+1100)^5}\,d\theta$$

$$\left[\frac{1}{2(\theta+1100)^2} - \frac{(2200+2\hat\theta)}{3(\theta+1100)^3} + \frac{(1100+\hat\theta)^2}{4(\theta+1100)^4}\right]_0^{\hat\theta} = \left[\frac{1}{2(\theta+1100)^2} - \frac{(2200+2\hat\theta)}{3(\theta+1100)^3} + \frac{(1100+\hat\theta)^2}{4(\theta+1100)^4}\right]_{\hat\theta}^{\infty}$$

$$\left(\frac{1}{2(\hat\theta+1100)^2} - \frac{(2200+2\hat\theta)}{3(\hat\theta+1100)^3} + \frac{(1100+\hat\theta)^2}{4(\hat\theta+1100)^4}\right)$$
$$-\left(\frac{1}{2(1100)^2} - \frac{(2200+2\hat\theta)}{3(1100)^3} + \frac{(1100+\hat\theta)^2}{4(1100)^4}\right) = -\left(\frac{1}{2(\hat\theta+1100)^2} - \frac{(2200+2\hat\theta)}{3(\hat\theta+1100)^3} + \frac{(1100+\hat\theta)^2}{4(\hat\theta+1100)^4}\right)$$

$$2\left(\frac{1}{2(\hat\theta+1100)^2} - \frac{(2200+2\hat\theta)}{3(\hat\theta+1100)^3} + \frac{(1100+\hat\theta)^2}{4(\hat\theta+1100)^4}\right) = \left(\frac{1}{2(1100)^2} - \frac{(2200+2\hat\theta)}{3(1100)^3} + \frac{(1100+\hat\theta)^2}{4(1100)^4}\right)$$

$$\frac{1}{(\hat\theta+1100)^2}\left(1-\frac{4}{3}+\frac{2}{4}\right) = \left(\frac{1}{2(1100)^2} - \frac{(2200+2\hat\theta)}{3(1100)^3} + \frac{(1100+\hat\theta)^2}{4(1100)^4}\right)$$

We try the values given and find (multiplying both sides of the equation by 1100^4):

x	$\frac{1100^4}{6(\hat\theta+1100)^2}$	$\frac{1100^2}{2} - \frac{1100(2200+2\hat\theta)}{3} + \frac{(1100+\hat\theta)^2}{4}$
(i) 422.35	105290.81	67997.38
(ii) 494.30	96001.80	71294.79
(iii) 560.87	88460.26	76650.96
(iv) 616.47	82822.26	82822.65

We see that *(iv)* is the Bayes estimate.

15. *An insurance company models claim frequency per year as following a Poisson distribution with mean Λ, where the prior distribution for Λ is a Gamma distribution with $\alpha = 5$ and $\theta = 6$. Over a 10-year period, they observe a total of 374 claims.*

    (a) Calculate the posterior distribution of Λ.

    The likelihood of the observed data is proportional to $e^{-10\lambda}\lambda^{374}$. The posterior distribution is therefore proportional to $\lambda^4 e^{-\frac{\lambda}{6}} e^{-10\lambda}\lambda^{374} = \lambda^{384}e^{-\frac{61}{6}\lambda}$, so the posterior distribution is a gamma distribution with $\alpha = 385$ and $\theta = \frac{6}{61}$.

    You are given the following values of the distribution function of the posterior distribution for Λ. Calculate a 95% credibility interval for Λ.

x	$F(x)$
34.12	*0.0231*
34.18	*0.0250*
34.35	*0.0310*
34.48	*0.0364*
41.60	*0.9706*
41.68	*0.9731*
41.74	*0.9749*
41.98	*0.9810*

(b) Using an HPD interval.

Using an HPD interval, We want to find points that have the same posterior density. The posterior density is proportional to $x^{384}e^{-\frac{61}{6}x}$. For computational purposes we compute the logarithm of this $384\log(x) - \frac{61}{6}x$ for the values of x in the table

x	$384\log(x) - \frac{61}{6}x$
34.12	1008.589
34.18	1008.653
34.35	1008.830
34.48	1008.959
41.60	1008.657
41.68	1008.582
41.74	1008.524
41.98	1008.286

We see that the matching pair of values in the table are 34.12 and 41.68, and we see that the differ

(c) With equal probability above and below the interval.

With equal probability above and below the interval, we want to take the 2.5th and 97.5th percentiles. From the table these are 34.18 and 41.74 respectively. The credibility interval is therefore $[34.18, 41.74]$.

16. *Claim severity follows a Pareto distribution with* $\theta = 1000$ *and* α *unknown. Which of the following is a conjugate prior distribution for* α*? Justify your answer.*

(i) Log-normal distribution

(ii) Inverse Weibull distribution

(iii) Gamma distribution

(iv) Inverse Pareto distribution

The likelihood is $L(x|\alpha) = \frac{\alpha 1000^{\alpha}}{(1000+x)^{\alpha+1}}$, so the conjugate density for α should remain in the same parametric family upon being multiplied by $\alpha e^{-\frac{\alpha}{\theta}}$. That is the conjugate prior is a Gamma distribution.