Practice Exam 1

1. Losses for an insurance coverage have the following cumulative distribution function:

$F(0) = 0$
$F(1{,}000) = 0.2$
$F(5{,}000) = 0.4$
$F(10{,}000) = 0.9$
$F(100{,}000) = 1$

with linear interpolation between these values.

Calculate the hazard rate at 9,000, $h(9{,}000)$.

(A) 0.0001 (B) 0.0004 (C) 0.0005 (D) 0.0007 (E) 0.0010

2. You are given the following data on loss sizes:

Loss Amount	Number of Losses
0– 1000	5
1000– 5000	4
5000–10000	3

An ogive is used as a model for loss sizes.

Determine the fitted median.

(A) 2000 (B) 2200 (C) 2500 (D) 3000 (E) 3083

3. In a mortality study on 5 individuals, death times were originally thought to be 1, 2, 3, 4, 5. It then turned out that one of these five observations was a censored observation rather than an actual death.

Determine the observation time of the five for which turning it into a censored observation would result in the lowest variance of the Nelson-Åalen estimator of $H(4)$.

(A) 1 (B) 2 (C) 3 (D) 4 (E) 5

4. For an insurance coverage, the number of claims per year follows a Poisson distribution. Claim size follows a Pareto distribution with $\alpha = 3$. Claim counts and claim sizes are independent.

The methods of classical credibility are used to determine premiums. The standard for full credibility is that actual aggregate claims be within 5% of expected aggregate claims 95% of the time. Based on this standard, 10,000 exposure units are needed for full credibility, where an exposure unit is a year of experience for a single insured.

Determine the expected number of claims per year.

(A) Less than 0.45
(B) At least 0.45, but less than 0.50
(C) At least 0.50, but less than 0.55
(D) At least 0.55, but less than 0.60
(E) At least 0.60

Exercises continue on the next page . . .

5. Which of the following statements is true?

(A) If data grouped into 7 groups are fitted to an inverse Pareto, the chi-square test of goodness of fit will have 5 degrees of freedom.

(B) The Kolmogorov-Smirnov statistic may be used to test the fit of a discrete distribution.

(C) The critical values of the Kolmogorov-Smirnov statistic do not require adjustment for estimated parameters.

(D) The critical values of the Kolmogorov-Smirnov statistic do not vary with sample size.

(E) The critical values of the Anderson-Darling statistic do not vary with sample size.

6–7. *Use the following information for questions 6 and 7:*

There are 2 classes of insureds. In Class A, the number of claims per year has a Poisson distribution with mean 0.1 and claim size has an exponential distribution with mean 500. In Class B, the number of claims per year has a Poisson distribution with mean 0.2 and claim size has an exponential distribution with mean 250. Each class has the same number of insureds.

An insured selected at random submits 2 claims in one year. Claim sizes are 200 and 400.

6. Calculate the probability that the insured is in class A.

(A) 0.01 (B) 0.04 (C) 0.15 (D) 0.19 (E) 0.27

7. Calculate the Bühlmann estimate of the aggregate losses for this insured in the following year.

(A) Less than 42

(B) At least 42, but less than 47

(C) At least 47, but less than 52

(D) At least 52, but less than 57

(E) At least 57

---

8. A class takes an exam. Half the students are good and half the students are bad. For good students, grades are distributed according to the probability density function

$$f(x) = \frac{4}{100}\left(\frac{x}{100}\right)^3 \qquad 0 \le x \le 100$$

For bad students, grades are distributed according to the probability density function

$$f(x) = \left(\frac{2}{100}\right)\left(\frac{x}{100}\right) \qquad 0 \le x \le 100$$

The passing grade is 65.

Determine the average grade on this exam for a passing student.

(A) 84.8 (B) 84.9 (C) 85.0 (D) 85.1 (E) 85.2

Exercises continue on the next page . . .

9. In a mortality study of 5 lives, death times were 6, 7, 9, 15, and 30. Using the empirical distribution, $S(10)$ is estimated as 0.4.

To approximate the mean square error of the estimate, the bootstrap method is used. 5 bootstrap samples are:

1. 6, 7, 7, 9, 30
2. 9, 6, 30, 7, 9
3. 30, 6, 6, 15, 7
4. 6, 15, 7, 9, 9
5. 30, 9, 6, 15, 15

Calculate the bootstrap approximation of the mean square error of the estimate.

(A) 0.032 (B) 0.034 (C) 0.036 (D) 0.038 (E) 0.040

10. An insurance coverage covers two types of insureds, A and B. There are an equal number of insureds in each class. Claim sizes in each class follow a Pareto distribution. Number of claims and claim sizes for insureds in each class have the following distributions:

Number of claims

	A	B
0	0.9	0.8
1	0.1	0.2

Size of claims (Pareto parameters)

	A	B
α	3	3
θ	50	60

Within each class, claim size and number of claims are independent.

Calculate the Bühlmann credibility to assign to 2 years of data.

(A) 0.01 (B) 0.02 (C) 0.03 (D) 0.04 (E) 0.05

11. An auto collision coverage is sold with deductibles of 500 and 1000. You have the following information for total loss size (including the deductible) on 86 claims:

Deductible 1000

Loss size	Number of losses
1000–2000	20
Over 2000	10

Deductible 500

Loss size	Number of losses
500–1000	32
Over 1000	24

Ground up underlying losses for both deductibles are assumed to follow an exponential distribution with the same parameter. You estimate the parameter using maximum likelihood.

Determine the fitted average total loss size (including the deductible) for claim payments on policies with a 500 deductible.

(A) 671 (B) 707 (C) 935 (D) 1171 (E) 1207

12. You are given the following data from a 2-year mortality study.

Duration j	Entries d_j	Withdrawals u_j	Deaths x_j
0	1000	100	33
1	500	100	c

Withdrawals and new entries are assumed to occur uniformly.

Based on this data, q_1 is estimated as 0.03.

Determine c.

(A) 26 (B) 32 (C) 35 (D) 38 (E) 41

Exercises continue on the next page . . .

13. The number of claims per year on an insurance coverage has a binomial distribution with parameters $m = 2$ and Q. Q varies by insured and is distributed according to the following density function:

$$f(q) = 42q(1-q)^5 \qquad 0 \le q \le 1$$

An insured submits 1 claim in 4 years.

Calculate the posterior probability that for this insured, q is less than 0.25.

(A) 0.52 (B) 0.65 (C) 0.70 (D) 0.76 (E) 0.78

14. You simulate a random variable with probability density function

$$f(x) = \begin{cases} -2x & -1 \le x \le 0 \\ 0 & \text{otherwise} \end{cases}$$

using the inversion method.

You use the following numbers random numbers from the uniform distribution on $[0, 1]$:

$$0.2 \qquad 0.4 \qquad 0.3 \qquad 0.7$$

Calculate the mean of the simulated observations.

(A) −0.7634 (B) −0.6160 (C) −0.2000 (D) 0.6160 (E) 0.7634

15. You are given a sample of 5 claims:

$$2, 3, 4, x_1, x_2$$

with $x_2 > x_1$. This sample is fitted to a Pareto distribution using the method of moments. The resulting parameter estimates are $\hat{\alpha} = 47.71$, $\hat{\theta} = 373.71$.

Determine x_1.

(A) 6.0 (B) 6.6 (C) 7.0 (D) 7.6 (E) 8.0

16. The number of claims per year on a policy follows a Poisson distribution with parameter Λ. Λ has a uniform distribution on $(0, 2)$.

An insured submits 5 claims in one year.

Calculate the Bühlmann credibility estimate of the number of claims for the following year.

(A) 1.6 (B) 1.8 (C) 2.0 (D) 2.5 (E) 3.5

17. According to the definitions in Klugman et al *Loss Models*, which of the following is true?

(A) $\tilde{\psi}(u,\tau) \ge \psi(u,\tau) \ge \psi(u)$
(B) $\psi(u) \ge \psi(u,\tau) \ge \tilde{\psi}(u,\tau)$
(C) $\psi(u,\tau) \ge \psi(u) \ge \tilde{\psi}(u)$
(D) $\tilde{\psi}(u) \ge \tilde{\psi}(u,\tau) \ge \psi(u,\tau)$
(E) $\psi(u,\tau) \ge \tilde{\psi}(u,\tau) \ge \tilde{\psi}(u)$

Exercises continue on the next page ...

18. A group has 100 lives. For each individual in this group, the mortality rate $q_x = 0.01$. Mortality for each individual is independent.

You are to simulate 3 years of mortality experience for this group using the inversion method. Use the following 3 numbers from the uniform distribution on $[0, 1]$: 0.12, 0.35, 0.68.

Calculate the total number of simulated deaths over 3 years.

(A) 0 (B) 1 (C) 2 (D) 3 (E) 4

19. A study is performed on the amount of time on unemployment. The records of 10 individuals are examined. 7 of the individuals are not on unemployment at the time of the study. The following is the number of weeks they were on unemployment:

5, 8, 10, 11, 17, 20, 26

3 individuals are still on unemployment at the time of the study. They have been unemployed for the following number of weeks:

5, 20, 26

Let T be the amount of time on unemployment.

Using the Kaplan-Meier estimator with exponential extrapolation past the last study time, estimate $\Pr(20 \le T \le 30)$.

(A) 0.17 (B) 0.21 (C) 0.28 (D) 0.32 (E) 0.43

20. Annual claim counts per risk are binomial with parameters $m = 2$ and Q. Q varies by risk uniformly on $(0.25, 0.75)$.

For a risk selected at random, determine the probability of no claims.

(A) 0.14 (B) 0.25 (C) 0.26 (D) 0.27 (E) 0.28

21. The distribution of auto insurance policyholders by number of claims submitted in the last year is as follows:

Number of claims	Number of insureds
0	70
1	22
2	6
3	2
Total	100

The number of claims for each insured is assumed to follow a Poisson distribution.

Use semi-parametric empirical Bayes estimation methods, with unbiased estimators for the variance of the hypothetical mean and the expected value of the process variance, to calculate the expected number of claims in the next year for a policyholder with 2 claims in the last year.

(A) Less than 0.52
(B) At least 0.52, but less than 0.57
(C) At least 0.57, but less than 0.62
(D) At least 0.62, but less than 0.67
(E) At least 0.67

Exercises continue on the next page . . .

22. Losses follow a lognormal distribution with $\mu = 3$, $\sigma = 0.5$.

Calculate the Value at Risk measure at $\alpha = 95\%$.

(A) Less than 40
(B) At least 40, but less than 45
(C) At least 45, but less than 50
(D) At least 50, but less than 55
(E) At least 55

23. You have the following experience for mortality in 3 classes of insureds.

	Class 1	Class 2	Class 3
Z	0	1	3
Number in class	5	3	2
Deaths at time y_1	1	1	0
Deaths at time y_2	0	1	1

There are no withdrawals.

A Cox model with covariate Z is used for this population, with values as indicated in the table. β is estimated as 0.1.

Calculate the Breslow estimate of the partial loglikelihood of this data.

(A) −8.7 (B) −8.4 (C) −8.1 (D) −7.8 (E) −7.5

24. For an insurance, the number of claims per year for each risk has a Poisson distribution with mean Λ. Λ varies by risk according to a gamma distribution with mean 0.5 and variance 1. Claim sizes follow a Weibull distribution with $\theta = 5$, $\tau = \frac{1}{2}$. Claim sizes are independent of each other and of claim counts.

Determine the variance of aggregate claims.

(A) 256.25 (B) 312.50 (C) 350.00 (D) 400.00 (E) 450.00

25. You are given the following claims data from an insurance coverage with claims limit 10,000:

1000, 2000, 2000, 2000, 4000, 5000, 5000

There are 3 claims for amounts over 10,000 which are censored at 10,000.

You fit this experience to an exponential distribution with parameter $\theta = 6{,}000$.

Calculate the Kolmogorov-Smirnov statistic for this fit.

(A) 0.11 (B) 0.13 (C) 0.15 (D) 0.18 (E) 0.19

Exercises continue on the next page . . .

26. For two stocks S and Q, prices are 60 and 75 respectively, and continuously compounded annual returns are 0.05 and 0.04 respectively. The covariance matrix is $\begin{pmatrix} 0.09 & 0.12 \\ 0.12 & 0.41 \end{pmatrix}$

Use the following standard normal random numbers to simulate the prices of S and Q respectively: 0.4, 0.2.

Determine the absolute difference between simulated prices of the stocks after one year.

(A) Less than 15
(B) At least 15, but less than 20
(C) At least 20, but less than 25
(D) At least 25, but less than 30
(E) At least 30

27. For an insurance coverage, you observe the following claims sizes:

$$400, 1100, 1100, 3000, 8000$$

You fit the loss distribution to a lognormal with $\mu = 7$ using maximum likelihood.

Determine the mean of the fitted distribution.

(A) Less than 2000
(B) At least 2000, but less than 2500
(C) At least 2500, but less than 3000
(D) At least 3000, but less than 3500
(E) At least 3500

28. In a mortality study performed on 5 lives, ages at death were

$$70, 72, 74, 75, 75$$

Estimate $S(75)$ using kernel smoothing with a uniform kernel with bandwidth 4.

(A) 0.2 (B) 0.3 (C) 0.4 (D) 0.5 (E) 0.6

29. For an insurance coverage, the number of claims per year follows a Poisson distribution with mean θ. The size of each claim follows an exponential distribution with mean 1000θ. Claim count and size are independent given θ.

You are examining one year of experience for four randomly selected policyholders, whose claims are as follows:

Policyholder #1	2000, 4000, 4000, 7000
Policyholder #2	4000
Policyholder #3	2000, 3000
Policyholder #4	1000, 4000, 5000

You use maximum likelihood to estimate θ.

Determine the variance of aggregate losses based on the fitted distribution.

(A) Less than 48,000,000
(B) At least 48,000,000, but less than 50,000,000
(C) At least 50,000,000, but less than 52,000,000
(D) At least 52,000,000, but less than 54,000,000
(E) At least 54,000,000

Exercises continue on the next page . . .

30. You are given:

(i) The annual number of claims for each risk follows a Poisson distribution with parameter Λ.
(ii) Λ varies by insured according to a gamma distribution with $\alpha = 3$ and $\theta = 0.1$.
(iii) Claims sizes follows a Pareto distribution with $\alpha = 3$ and $\theta = 20{,}000$.
(iv) Claim sizes are independent of claim counts.
(v) Your department handles only claims with sizes below 10,000.

Determine the variance of the annual number of claims handled per risk in your department.

(A) 0.159 (B) 0.163 (C) 0.226 (D) 0.232 (E) 0.330

31. Past data on aggregate losses for two group policyholders is given in the following table.

Group		Year 1	Year 2
A	Total losses	1000	1200
	Number of members	40	50
B	Total losses	500	600
	Number of members	20	40

Calculate the credibility factor used for Group A's experience using non-parametric empirical Bayes estimation methods.

(A) Less than 0.40
(B) At least 0.40, but less than 0.45
(C) At least 0.45, but less than 0.50
(D) At least 0.50, but less than 0.55
(E) At least 0.55

32. You study the relationship of gender and duration of illness to survival after radiation treatment for cancer.

(i) You model the relationship with a two-parameter Cox proportional hazards model. Gender is coded with an indicator variable $z_1 = 1$ for males. Duration is coded with the variable z_2 and is measured in years.
(ii) The resulting partial likelihood estimates of the coefficients are:

$$\hat{\beta}_1 = 0.1$$
$$\hat{\beta}_2 = 0.01$$

(iii) The covariance matrix of $\hat{\beta}_1$ and $\hat{\beta}_2$ is given by

$$\begin{pmatrix} 0.2 & -0.03 \\ -0.03 & 0.008 \end{pmatrix}$$

Construct a 95% confidence interval for the relative risk of a male who has had the disease for 5 years to a female who has had the disease for 2 years.

(A) $(0.54, 1.73)$ (B) $(0.59, 1.84)$ (C) $(0.59, 1.95)$ (D) $(0.63, 2.06)$ (E) $(0.84, 1.69)$

33. For an insurance coverage, claim size follows a Pareto distribution with parameters $\alpha = 4$ and θ. θ varies by insured and follows a normal distribution with $\mu = 3$ and $\sigma = 1$.

Determine the Bühlmann credibility to be assigned to a single claim.

(A) 0.05 (B) 0.07 (C) 0.10 (D) 0.14 (E) 0.20

Exercises continue on the next page ...

34. S_t is price of a stock at time t, with t expressed in years. You are given:

(i) S_t/S_0 is lognormally distributed.
(ii) The mean continuously compounded annual return on the stock is 5%.
(iii) The annual σ for the stock is 30%.
(iv) The stock pays no dividends.

Determine the probability that the stock will have a positive return at the end of 3 years.

(A) 0.49 (B) 0.51 (C) 0.54 (D) 0.59 (E) 0.61

35. Observations of a random variable X are fitted to a Pareto distribution using maximum likelihood. The estimated parameters are $\hat{\alpha} = 3$, $\hat{\theta} = 600$. The information matrix for $(\hat{\alpha}, \hat{\theta})$ is

$$\begin{pmatrix} 100 & 1 \\ 1 & 9 \end{pmatrix}$$

Calculate the width of a 95% confidence interval for the mean of X.

(A) 59 (B) 72 (C) 98 (D) 144 (E) 196

36. You are given the following information regarding loss sizes:

d	Mean excess loss $e(d)$	$F(d)$
0	3000	0.0
500	3500	0.2
10,000	2500	0.8

Determine the average payment per loss for a policy with a deductible of 500 and a maximum covered loss of 10,000.

(A) 1300 (B) 1500 (C) 2100 (D) 2300 (E) 2500

Exercises continue on the next page . . .

37. A claims adjustment facility adjusts all claims for amounts less than or equal to 10,000. Claims for amounts greater than 10,000 are handled elsewhere.

In 2002, the claims handled by this facility fell into the following ranges:

Size of Claim	Number of Claims
Less than 1000	100
1000– 5000	75
5000–10000	25

The claims are fitted to a parametric distribution using maximum likelihood.

Which of the following is the correct form for the likelihood function of this experience?

(A) $[F(1000)]^{100}\,[F(5000)-F(1000)]^{75}\,[F(10{,}000)-F(5000)]^{25}$

(B) $\dfrac{[F(1000)]^{100}\,[F(5000)-F(1000)]^{75}\,[F(10{,}000)-F(5000)]^{25}}{[F(10{,}000)]^{200}}$

(C) $\dfrac{[F(1000)]^{100}\,[F(5000)-F(1000)]^{75}\,[F(10{,}000)-F(5000)]^{25}}{[1-F(10{,}000)]^{200}}$

(D) $\dfrac{[1-F(1000)]^{100}\,[F(5000)-F(1000)]^{75}\,[F(10{,}000)-F(5000)]^{25}}{[F(10{,}000)]^{200}}$

(E) $\dfrac{[1-F(1000)]^{100}\,[F(5000)-F(1000)]^{75}\,[F(10{,}000)-F(5000)]^{25}}{[1-F(10{,}000)]^{200}}$

38. For a sample from an exponential distribution, which of the following statements is false?

(A) If the sample has size 2, the sample median is an unbiased estimator of the population median.

(B) If the sample has size 2, the sample median is an unbiased estimator of the population mean.

(C) If the sample has size 3, the sample mean is an unbiased estimator of the population mean.

(D) If the sample has size 3, 1.2 times the sample median is an unbiased estimator of the population mean.

(E) The sample mean is a consistent estimator of the population mean.

39. The median of a sample is 5. The sample is fitted to a mixture of two exponential distributions with means 3 and $x > 3$, using percentile matching to determine the weights to assign to each exponential.

Which of the following is the range of values for x for which percentile matching works?

(A) $3 < x < 4.3281$.

(B) $3 < x < 7.2135$.

(C) $4.3281 < x < 7.2135$.

(D) $x > 4.3281$.

(E) $x > 7.2135$.

40. A normal probability plot is drawn for the sample $\{2, 3, 6, 9, 20\}$.

Determine the value on the y-axis corresponding to 9.

(A) 0.430 (B) 0.524 (C) 0.667 (D) 0.700 (E) 0.758

Solutions to the above questions begin on page 1089.

Appendix A. Solutions to the Practice Exams

Answer Key for Practice Exam 1

1	C	11	E	21	E	31	E
2	A	12	B	22	C	32	D
3	D	13	D	23	A	33	A
4	E	14	A	24	D	34	B
5	E	15	C	25	D	35	A
6	D	16	C	26	A	36	D
7	C	17	B	27	A	37	B
8	D	18	B	28	B	38	A
9	A	19	D	29	E	39	E
10	A	20	D	30	C	40	D

Practice Exam 1

1. [**Lesson 1**] $F(9{,}000) = 0.4 + \left(\frac{9{,}000-5{,}000}{10{,}000-5{,}000}\right)(0.9-0.4) = 0.8$. Then $S(9{,}000) = 0.2$. The hazard rate is $\frac{f(x)}{S(x)}$. The density function is the derivative of F, which at 9,000 is the slope of the line from 5,000 to 10,000, which is $\frac{0.9-0.4}{10{,}000-5{,}000} = 0.0001$. The answer is

$$h(9{,}000) = \frac{f(9{,}000)}{S(9{,}000)} = \frac{0.0001}{0.2} = \boxed{\mathbf{0.0005}} \qquad \textbf{(C)}$$

2. [**Lesson 19**] The distribution function at 1000, $F(1000)$, is $\frac{5}{12}$, and $F(5000) = \frac{9}{12}$. By definition, the median is the point m such that $F(m) = 0.5$. The ogive linearly interpolates between 1000 and 5000. Thus we solve the equation

$$\begin{aligned} \frac{m-1000}{5000-1000} &= \frac{0.5-\frac{5}{12}}{\frac{9}{12}-\frac{5}{12}} \\ m - 1000 &= \frac{1}{4}(4000) = 1000 \\ m &= \boxed{\mathbf{2000}} \qquad \textbf{(A)} \end{aligned}$$

3. [**Lesson 23**] The variance is $\sum \frac{1}{r_i^2}$. To minimize the variance, maximize the r_i's. The r_i's for the first 3 deaths are 3 out of $\{5,4,3,2\}$. By making $\boxed{\mathbf{4}}$ the censored observation, the r_i's are 5, 4, and 3. (Making 5 the censored observation leads to a sum of *4* terms instead of 3, with the first 3 terms being the same as if 4 is censored, so it does not lead to minimal variance). **(D)**

4. [**Lesson 42**] The credibility formula in terms of expected number of claims, formula (42.2), requires $1+CV_s^2$, or the second moment divided by the first moment squared of the severity distribution. For a Pareto with $\alpha = 3$, this is

$$\frac{\frac{2\theta^2}{(2)(1)}}{\left(\frac{\theta}{2}\right)^2} = 4$$

The expected number of claims needed for full credibility is then $\left(\frac{1.96}{0.05}\right)^2(4) = 6146.56$, so we have

$$6146.56 = 10{,}000\lambda$$
$$\lambda = \boxed{\mathbf{0.614656}} \qquad \textbf{(E)}$$

5. [**Lessons 38, 39, 40**] (A) is false because the number of degrees of freedom is $n - 1$ minus the number of parameters estimated. Here $n = 7$ and the inverse Pareto has 2 parameters, so there are 4 degrees of freedom.

(B) is false, as indicated on page 555.

(C) is false, as indicated on page 555, where it says that the indicated critical values only work when the distribution is completely specified, not when parameters have been estimated.

(D) is false; in fact, the critical values get divided by $\sqrt{n}$.

(E) is true.

6. [**Lesson 45**] The likelihood that an insured in class A submits 2 claims for 200 and 400 is the product of the Poisson probability of 2 claims and the 2 densities under the exponential distribution with mean 500, or

$$f(200, 400|A) = e^{-0.1}\left(\frac{0.1^2}{2!}\right)\left(\frac{e^{-200/500}}{500}\right)\left(\frac{e^{-400/500}}{500}\right) = 5.4506 \times 10^{-9}$$

and similarly for an insured in class B:

$$f(200, 400|A) = e^{-0.2}\left(\frac{0.2^2}{2!}\right)\left(\frac{e^{-200/250}}{250}\right)\left(\frac{e^{-400/250}}{250}\right) = 2.3768 \times 10^{-8}$$

Since the classes are of equal size, these are the relative probabilities of the two classes. Then the probability that the insured is in class A is

$$\Pr(A|200, 400) = \frac{5.4506 \times 10^{-9}}{5.4506 \times 10^{-9} + 2.3768 \times 10^{-8}} = \boxed{\mathbf{0.1865}} \qquad \textbf{(D)}$$

7. [**Lesson 52**] Since the hypothetical means are the same in both classes—(0.1)(500) = 50 and (0.2)(250) = 50—$a = 0$ and there is no credibility. Thus, $\boxed{\mathbf{50}}$ is the credibility premium. **(C)**

8. [**Lesson 4**] We need to calculate $E[X|X > 65]$, where X is the grade. By definition

$$E[X|X > 65] = \frac{\int_{65}^{100} x f(x)\mathrm{d}x}{1 - F(65)}$$

$f(x)$ is an equally weighted mixture of the good and bad students, and therefore is

$$f(x) = \frac{1}{2}\left[\frac{4}{100}\left(\frac{x}{100}\right)^3 + \frac{2}{100}\left(\frac{x}{100}\right)\right]$$

First we calculate the denominator of $E[X|X > 65]$.

$$F(x) = \int_0^x f(u)\mathrm{d}u = 0.5\left[\left(\frac{x}{100}\right)^4 + \left(\frac{x}{100}\right)^2\right]$$
$$F(65) = 0.5(0.65^4 + 0.65^2) = 0.300503$$
$$1 - F(65) = 1 - 0.300503 = 0.699497$$

Then we calculate the numerator.

$$\begin{aligned}
\int_{65}^{100} x f(x)\mathrm{d}x &= \int_{65}^{100} 0.5\left[\frac{4x^4}{100^4} + \frac{2x^2}{100^2}\right]\mathrm{d}x \\
&= 0.5\left[\frac{4x^5}{5(100^4)} + \frac{2x^3}{3(100^2)}\right]_{65}^{100} \\
&= 0.5\left[\frac{400}{5} - \frac{4(65^5)}{5(100^4)} + \frac{200}{3} - \frac{2(65^3)}{3(100^2)}\right] \\
&= 0.5(80 - 9.2823 + 66.6667 - 18.3083) = 59.5380 \\
E[X|X > 65] &= \frac{59.5380}{0.699497} = \boxed{\mathbf{85.1155}} \qquad \textbf{(D)}
\end{aligned}$$

An alternative way to solve this problem is to use the tabular form for Bayesian credibility that we studied in lesson 45. The table would look like this:

	Good Students	Bad Students	
Prior probabilities	0.50	0.50	
Likelihood of experience	0.821494	0.5775	
Joint probabilities	0.410747	0.28875	0.699497
Posterior probabilities (p_i)	0.587203	0.412797	
65+e(65)	86.084252	83.737374	
$[65 + e(65)] \times p_i$	50.548957	34.566511	**85.1155**

The second line, the likelihood, is derived as follows:

$$F(65) = \int_0^{65} f(x)\,\mathrm{d}x = \begin{cases} (65/100)^4 & \text{for good students} \\ (65/100)^2 & \text{for bad students} \end{cases}$$

so the likelihood of more than 65 is $1-0.65^4 = 0.821494$ for good students and $1-0.65^2 = 0.5775$ for bad students.

The fifth line, the average grade of those with grades over 65, is derived as follows for good students:

$$\begin{aligned}
65 + e(65) &= 65 + \frac{\int_{65}^{100} S(x)\,\mathrm{d}x}{S(65)} \\
&= 65 + \int_{65}^{100}\left(1 - \left(\frac{x}{100}\right)^4\right)\mathrm{d}x \\
&= 65 + 35 - \left(\frac{x^5}{5(100^4)}\right)\Bigg|_{65}^{100} \\
&= 100 - \frac{100^5 - 65^5}{5(100^4)} = 86.084252
\end{aligned}$$

and for bad students, replace the 4's with 2's and the 5's with 3's to obtain $100 - \frac{100^3-65^3}{2(100^2)} = 83.737374$.

9. [**Lesson 62**] The estimates of $S(10)$ from these 5 samples are the proportion of numbers above 10, which are 0.2, 0.2, 0.4, 0.2, 0.6 respectively, so the bootstrap approximation is

$$\frac{(0.2-0.4)^2 + (0.2-0.4)^2 + (0.4-0.4)^2 + (0.2-0.4)^2 + (0.6-0.4)^2}{5} = \boxed{\mathbf{0.032}} \qquad \textbf{(A)}$$

10. [**Lesson 52**] Let μ_A be the hypothetical mean for A, v_A the process variance for A, and use the same notation with subscripts B for B. For process variance, we will use the compound variance formula.

$$\mu_A = 0.1(25) = 2.5 \qquad v_A = 0.1(2500 - 625) + 0.09(25^2) = 243.75$$
$$\mu_B = 0.2(30) = 6 \qquad v_B = 0.2(3600 - 900) + 0.16(30^2) = 684$$
$$a = (6 - 2.5)^2\left(\frac{1}{4}\right) = 3.0625 \qquad v = \frac{1}{2}(243.75 + 684) = 463.875$$

For 2 years of experience, the credibility is

$$Z = \frac{na}{na + v} = \frac{(2)(3.0625)}{(2)(3.0625) + 463.875} = \boxed{\mathbf{0.013032}} \qquad \textbf{(A)}$$

11. [**Lesson 28**] The likelihood function (either using the fact that the exponential is memoryless, or else writing it all out and canceling out the denominators) is

$$L(\theta) = \left(1 - e^{-(1000/\theta)}\right)^{20}\left(e^{-(1000/\theta)}\right)^{10}\left(1 - e^{-(500/\theta)}\right)^{32}\left(e^{-(500/\theta)}\right)^{24}$$

Logging and differentiating:

$$\begin{aligned}
l(\theta) &= 20\ln\left(1 - e^{-1000/\theta}\right) + 32\ln\left(1 - e^{-500/\theta}\right) - \frac{10{,}000 + 12{,}000}{\theta}\\
\frac{dl}{d\theta} &= \frac{-20{,}000e^{-1000/\theta}}{\theta^2(1 - e^{-1000/\theta})} + \frac{-16{,}000e^{-500/\theta}}{\theta^2(1 - e^{-500/\theta})} + \frac{22{,}000}{\theta^2} = 0
\end{aligned}$$

Multiply through by $\frac{\theta^2}{1000}$, and set $x = e^{-500/\theta}$. We obtain

$$\begin{aligned}
22 - \frac{20x^2}{1 - x^2} - \frac{16x}{1 - x} &= 0\\
\frac{22(1 - x^2) - 20x^2 - 16x(1 + x)}{1 - x^2} &= 0\\
22 - 22x^2 - 20x^2 - 16x - 16x^2 &= 0\\
58x^2 + 16x - 22 &= 0\\
x = \frac{-16 + \sqrt{16^2 + 4(22)(58)}}{116} &= 0.493207\\
\theta = \frac{-500}{\ln x} &= 707.3875
\end{aligned}$$

An easier way to do this would be to make the substitution $x = e^{-500/\theta}$ right in $L(\theta)$, and to immediately recognize that $1 - x^2 = (1 - x)(1 + x)$. This would avoid the confusing differentiation step:

$$\begin{aligned}
l(\theta) &= 20\ln(1 - x^2) + 44\ln x + 32\ln(1 - x)\\
&= 20\ln(1 - x) + 20\ln(1 + x) + 44\ln x + 32\ln(1 - x)\\
&= 52\ln(1 - x) + 20\ln(1 + x) + 44\ln x\\
\frac{dl}{d\theta} &= -\frac{52}{1 - x} + \frac{20}{1 + x} + \frac{44}{x} = 0\\
&- 52(x)(1 + x) + 20(x)(1 - x) + 44(1 - x)^2 = 0\\
&- 52x^2 - 52x - 20x^2 + 20x + 44 - 44x^2 = 0
\end{aligned}$$

$$116x^2 + 32x - 44 = 0$$

which is double the quadratic above, and leads to the same solution for θ.

Using the fact that the exponential distribution is memoryless, the average total loss size for a 500 deductible is $500 + 707.3875 = \boxed{\mathbf{1207.3875}}$. **(E)**

12. [**Lesson 25**] The conditional probability of death at the second duration (note that 0 is the first duration and 1 is the second), q_1, is estimated by $\frac{s_2}{r_2}$, number of deaths over the risk set in duration 2. Duration 2 starts with $1000 - 100 - 33 = 867$ lives, and since withdrawals and new entries occur uniformly, we add half the new entries and subtract half the withdrawals to arrive at $r_2 = 867 + 0.5(500 - 100) = 1067$. Then

$$\begin{aligned} 0.03 = \hat{q}_1 &= \frac{c}{1067} \\ c &= \boxed{\mathbf{32}} \qquad \textbf{(B)} \end{aligned}$$

13. [**Lesson 49**] The prior distribution is a beta distribution with $a = 2$, $b = 6$. (In general, in a beta distribution, a is 1 more than the exponent of q and b is 1 more than the exponent of $1 - q$.) The number of claims is binomial, which means that 2 claims are possible each year. Of the 8 possible claims in 4 years, you received 1 and didn't receive 7. Thus $a' = 2 + 1 = 3$ and $b' = 6 + 7 = 13$ are the parameters of the posterior beta. The density function for the posterior beta is

$$\begin{aligned} \pi(q|\mathbf{x}) &= \frac{\Gamma(3+13)}{\Gamma(3)\Gamma(13)} q^2 (1-q)^{12} \\ &= 1365 q^2 (1-q)^{12} \end{aligned}$$

since $\frac{15!}{2!12!} = \frac{(15)(14)(13)}{2} = 1365$. We must integrate this function from 0 to 0.25 to obtain $\Pr(Q < 0.25)$. It is easier to integrate if we change the variable, by setting $q' = 1 - q$. Then we have

$$\begin{aligned} \Pr(Q < 0.25) &= 1365 \int_{0.75}^{1} (1-q')^2 q'^{12} \mathrm{d}q \\ &= 1365 \int_{0.75}^{1} (q'^{12} - 2q'^{13} + q'^{14}) \mathrm{d}q \\ &= \left. \frac{q'^{13}}{13} - \frac{2q'^{14}}{14} + \frac{q'^{15}}{15} \right|_{0.75}^{1} \\ &= 1365(0.0007326 - 0.0001730) = \boxed{\mathbf{0.7639}} \qquad \textbf{(D)} \end{aligned}$$

14. [**Lesson 59**] The distribution function is

$$F(x) = \int_{-1}^{x} -2u \, \mathrm{d}u = -u^2 \Big|_{-1}^{x} = 1 - x^2$$

Inverting,

$$\begin{aligned} u &= 1 - x^2 \\ 1 - u &= x^2 \\ x &= -\sqrt{1-u} \end{aligned}$$

It is necessary to use the negative square root, since the simulated observation must be between −1 and 0. So

$$x_1 = -\sqrt{1-0.2} = -0.8944$$
$$x_2 = -\sqrt{1-0.4} = -0.7746$$
$$x_3 = -\sqrt{1-0.3} = -0.8367$$
$$x_4 = -\sqrt{1-0.7} = -0.5477$$
$$\frac{-0.8944-0.7746-0.8367-0.5477}{4} = \boxed{\mathbf{-0.7634}} \qquad \textbf{(A)}$$

15. [**Lesson 26**] We write the moment equations for the first and second moments:

$$\frac{2+3+4+x_1+x_2}{5} = \frac{\theta}{\alpha-1}$$
$$9+x_1+x_2 = 5\left(\frac{373.71}{47.71-1}\right) = 40$$
$$\frac{2^2+3^2+4^2+x_1^2+x_2^2}{5} = \frac{2\theta^2}{(\alpha-1)(\alpha-2)}$$
$$29+x_1^2+x_2^2 = 5\left[\frac{2(373.71^2)}{(46.71)(45.71)}\right] = 654$$

We use the first equation to solve for x_2 in terms of x_1, and plug that into the second equation and solve.

$$x_2 = 31 - x_1$$
$$29 + x_1^2 + (31-x_1)^2 = 654$$
$$29 + 2x_1^2 - 62x_1 + 961 = 654$$
$$2x_1^2 - 62x_1 + 336 = 0$$
$$x_1^2 - 31x_1 + 168 = 0$$
$$x_1 = \frac{31 \pm \sqrt{31^2 - 4(168)}}{2}$$
$$= \frac{31 \pm 17}{2} = 7 \text{ or } 24$$

Since x_2 is higher than x_1, $x_1 = \boxed{7}$. **(C)**

16. [**Lesson 52**] The hypothetical mean is Λ. The expected hypothetical mean $\mu = E(\Lambda) = 1$ (the mean of the uniform distribution). The process variance is Λ. The expected process variance v, or the expected value of Λ, is 1. The variance of the hypothetical mean is $\mathrm{Var}(\Lambda)$. For a uniform distribution on $(0,\theta)$, the variance is $\frac{\theta^2}{12}$, so the variance is $a = \frac{1}{3}$. The Bühlmann k is therefore $\frac{1}{1/3} = 3$. $Z = \frac{1}{1+3} = 0.25$. The credibility premium is $\frac{1}{4}(5) + \frac{3}{4}(1) = \boxed{2}$. **(C)**

17. [**Lesson 17**] ψ is the probability of ruin. The probability of ruin if checking is continuous (ψ) is greater than if checking is only done at discrete intervals ($\tilde{\psi}$), and the probability of ever being ruined ($\psi(u)$) is greater than the probability of being ruined by time τ ($\psi(u,\tau)$). Only **(B)** satisfies these conditions.

18. [**Lesson 61**] This is a binomial distribution with $m = 100$, $q = 0.01$.

$$p_0 = 0.99^{100} = 0.366$$

$$p_1 = 100\left(0.99^{99}\right)(0.01) = 0.370$$

Together these add up to more than 0.68, the highest uniform number. 0.12 and 0.35 go to 0 and 0.68 goes to 1. $0 + 0 + 1 = \boxed{1}$. **(B)**

19. [**Lesson 21 and section 22.2**] First we calculate $\hat{S}(26)$.

y_i	r_i	s_i	$S_{10}(y_i)$
5	10	1	0.9
8	8	1	0.7875
10	7	1	0.675
11	6	1	0.5625
17	5	1	0.45
20	4	1	0.3375
26	2	1	0.16875

So $\hat{S}(26) = 0.16875$. To extrapolate, we exponentiate $\hat{S}(26)$ to the $\frac{30}{26}$ power, as discussed in example 21D on page 296:

$$\hat{S}(30) = 0.16875^{30/26} = 0.12834.$$

$\Pr(20 \leq T \leq 30) = S(20^-) - S(30)$, since the lower endpoint is included. But $S(20^-) = S(17) = 0.45$. So the answer is $0.45 - 0.12834 = \boxed{\mathbf{0.3217}}$. **(D)**

20. [**Section 2.3 and lesson 10**] The density of the uniform distribution is the reciprocal of the range (1/2), or 2. We integrate p_0 for the binomial, or $(1-q)^2$, over the uniform distribution.

$$\begin{aligned}
\Pr(N = 0) &= \int_{0.25}^{0.75} (2)(1-q)^2 \mathrm{d}q \\
&= -2\left(\frac{(1-q)^3}{3}\right)\Bigg|_{0.25}^{0.75} \\
&= \left(\frac{2}{3}\right)(0.75^3 - 0.25^3) = \boxed{\mathbf{0.270833}} \qquad \textbf{(D)}
\end{aligned}$$

21. [**Lesson 58**] We estimate μ, v, and a:

$$\hat{\mu} = \hat{v} = \bar{x} = \frac{22(1) + 6(2) + 2(3)}{100} = \frac{40}{100} = 0.4$$

We will calculate s^2, the unbiased sample variance, by calculating the second moment, subtracting the square of the sample mean (which gets us the empirical variance) and then multiplying by $\frac{n}{n-1}$ to turn it into the sample variance.

$$\begin{aligned}
\hat{a} + \hat{v} = s^2 &= \frac{100}{99}\left[\frac{22(1^2) + 6(2^2) + 2(3^2)}{100} - \bar{x}^2\right] \\
&= \frac{100}{99}(0.64 - 0.4^2) = 0.4848 \\
\hat{a} &= 0.4848 - 0.4 = 0.0848 \\
Z &= \frac{a}{a+v} = \frac{0.0848}{0.4848} \\
P_C &= 0.4 + \frac{0.0848}{0.4848}(1.6) = \boxed{\mathbf{0.68}} \qquad \textbf{(E)}
\end{aligned}$$

22. [**Lesson 66**] The 95th percentile of a normal distribution with parameters $\mu = 3$, $\sigma = 0.5$ is $3 + 1.645(0.5) = 3.8225$. Exponentiating, the 95th percentile of a lognormal distribution is $e^{3.8225} = \boxed{\mathbf{45.718}}$. **(C)**

23. [**Lesson 34**] We use expression (34.1), with the proportionality constants $c = e^{\beta Z}$.

$$\ln\left[\frac{e^{0.1}}{(5+3e^{0.1}+2e^{0.3})^2}\right] + \ln\left[\frac{e^{0.4}}{(4+2e^{0.1}+2e^{0.3})^2}\right] = \boxed{\mathbf{-8.6729}} \qquad \textbf{(A)}$$

24. [**Lessons 11 and 13**] Let N be claim counts, X claim size, S aggregate claims. N is a gamma mixture of a Poisson, or a negative binomial. The gamma has parameters α and θ (not the same θ as the Weibull) such that

$$\alpha\theta = 0.5$$
$$\alpha\theta^2 = 1$$

implying $\beta = \theta = 2$, $r = \alpha = 0.25$, so

$$E[N] = r\beta = 0.5$$
$$\mathrm{Var}(N) = r\beta(1+\beta) = 0.25(2)(3) = 1.5$$

The Weibull[1] has mean

$$E[X] = \theta\Gamma(1+2) = (5)(2!) = (5)(2) = 10$$

and second moment

$$E[X^2] = \theta^2\Gamma(1+2^2) = (5^2)(4!) = (25)(24) = 600$$

and therefore $\mathrm{Var}(X) = 600 - 10^2 = 500$. By the compound variance formula

$$\mathrm{Var}(S) = (0.5)(500) + (1.5)(10^2) = \boxed{\mathbf{400}} \qquad \textbf{(D)}$$

25. [**Lesson 38**] We set up a table for the empirical and fitted functions. Note that we do not know the empirical function at 10,000 or higher due to the claims limit.

x	$F_n(x^-)$	$F_n(x)$	$F(x)$	Largest difference
1000	0	0.1	$1 - e^{-1/6} = 0.1535$	0.1535
2000	0.1	0.4	$1 - e^{-1/3} = 0.2835$	0.1835
4000	0.4	0.5	$1 - e^{-2/3} = 0.4866$	0.0866
5000	0.5	0.7	$1 - e^{-5/6} = 0.5654$	0.1346
10,000	0.7		$1 - e^{-5/3} = 0.8111$	0.1111

Inspection indicates that the maximum difference occurs at 2000 and is $\boxed{\mathbf{0.1835}}$. **(D)**

26. [**Lesson 63**] The lognormal parameters for the 2 stocks (using subscript 1 for S and 2 for Q) are

$$\sigma_1^2 = 0.09 \qquad \mu_1 = 0.05 - 0.5(0.09) = 0.005$$
$$\sigma_2^2 = 0.41 \qquad \mu_2 = 0.04 - 0.5(0.41) = -0.165$$

The Choleski factorization of the covariance matrix is

$$a_{11} = \sqrt{0.09} = 0.3$$

[1] In general, $\Gamma(n) = (n-1)!$ for n an integer

$$a_{21} = \frac{0.12}{0.3} = 0.4$$
$$a_{22} = \sqrt{0.41 - 0.4^2} = 0.5$$

The stock returns are

$$\begin{pmatrix} 0.3 & 0 \\ 0.4 & 0.5 \end{pmatrix}\begin{pmatrix} 0.4 \\ 0.2 \end{pmatrix} + \begin{pmatrix} 0.005 \\ -0.165 \end{pmatrix} = \begin{pmatrix} 0.125 \\ 0.095 \end{pmatrix}$$
$$e^{0.125} = 1.11315 \qquad 60(1.113148) = 67.9889$$
$$e^{0.095} = 1.09966 \qquad 75(1.099659) = 82.4744$$

The difference is $82.4744 - 67.9889 = \boxed{\mathbf{14.4855}}$. **(A)**

27. [**Lesson 29**] See the discussion of transformations and the lognormal example 29B on page 436, and the paragraph before the example. *The shortcut for lognormals must be adapted for this question since μ is given; it is incorrect to calculate the empirical mean and variance as if μ were unknown.* We will log each of the claim sizes and fit them to a normal distribution. You may happen to know that for a normal distribution, the MLE's of μ and σ are independent, so given μ, the MLE for σ will be the same as if μ were not given. Moreover, the MLE for σ for a normal distribution is the sample variance divided by n (rather than by $n-1$). If you know these two facts, you can calculate the MLE for σ on the spot. If not, it is not too hard to derive. The likelihood function (omitting the constant $\sqrt{2\pi}$) is (where we let $x_i =$ the *log* of the claim size)

$$L(\sigma) = \frac{1}{\sigma^5}\prod_{i=1}^{5} e^{-\frac{(x_i-7)^2}{2\sigma^2}}$$
$$l(\sigma) = -5\ln\sigma - \sum_{i=1}^{5}\frac{(x_i-7)^2}{2\sigma^2}$$
$$\frac{\mathrm{d}l}{\mathrm{d}\sigma} = -\frac{5}{\sigma} + \frac{\sum_{i=1}^{5}(x_i-7)^2}{\sigma^3} = 0$$
$$\sigma^2 = \frac{\sum_{i=1}^{5}(x_i-7)^2}{5}$$

We calculate

$$\sigma^2 = \frac{(\ln 400 - 7)^2 + 2(\ln 1100 - 7)^2 + (\ln 3000 - 7)^2 + (\ln 8000 - 7)^2}{5} = 1.1958$$

The mean of the lognormal is

$$\exp\left(\mu + \sigma^2/2\right) = \exp(7 + 1.1958/2) = \boxed{\mathbf{1994}} \qquad \textbf{(A)}$$

28. [**Lesson 24**] The kernel survival function for a uniform kernel is a straight line from 1 to 0, starting at the observation point minus the bandwidth and ending at the observation point plus the bandwidth. From the perspective of 74, we reverse orientation; the kernel survival for 74 *increases* as the observation increases. Therefore, the kernels are 0 at 70, $\frac{1}{8}$ at 72 (which is $\frac{1}{8}$ of the way from 71 to 79), $\frac{3}{8}$ at 74 (which is $\frac{3}{8}$ of the way from 71 to 79), and $\frac{1}{2}$ at 75 (which is $\frac{1}{2}$ of the way from 71 to 79). Each point has a weight of $\frac{1}{n} = \frac{1}{5}$. We therefore have:

$$\frac{1}{5}\left[\frac{1}{8} + \frac{3}{8} + (2)\left(\frac{1}{2}\right)\right] = \boxed{\mathbf{0.30}} \qquad \textbf{(B)}$$

29. [**Lesson 28**] The likelihood function is the product of

$$e^{-\theta}\frac{\theta^{n_i}}{n_i!}$$

for the number of claims n_i for the 4 individuals, times the product of

$$\frac{1}{1000\theta}e^{-x_i/1000\theta}$$

for each of the 10 claim sizes x_i. These get multiplied together to form the likelihood function. We have

$$\sum n_i = 4 + 1 + 2 + 3 = 10$$

and

$$\sum \frac{x_i}{1000\theta} = \frac{2000 + 4000 + 4000 + 7000 + 4000 + 2000 + 3000 + 1000 + 4000 + 5000}{1000\theta} = \frac{36}{\theta}$$

If we ignore the constants, the likelihood function is:

$$\begin{aligned}
L(\theta) &= e^{-4\theta}\theta^{10}\frac{1}{\theta^{10}}e^{-36/\theta} \\
l(\theta) &= -4\theta - \frac{36}{\theta} \\
\frac{\mathrm{d}l}{\mathrm{d}\theta} &= -4 + \frac{36}{\theta^2} = 0 \\
\theta &= 3
\end{aligned}$$

To complete the problem, use equation (13.3), or better, since number of claims is Poisson, equation (13.5). Let S be aggregate losses. Using either formula, we obtain $\mathrm{Var}(S) = 3[2(3000^2)] = \boxed{\mathbf{54{,}000{,}000}}$ for the fitted distribution. **(E)**

30. [**Lessons 11 and 12**] The mixed number of claims for all risks is negative binomial with $r = 3$, $\beta = 0.1$. However, this must be adjusted for severity modification; only $F(10{,}000) = 1 - \left(\frac{20{,}000}{30{,}000}\right)^3 = \frac{19}{27}$ of claims are handled by your department, where F is the distribution function of a Pareto. The modification is to set $\beta = \frac{19}{27}(0.1)$. The variance is then $r\beta(1+\beta) = 3\left(\frac{1.9}{27}\right)\left[1 + \left(\frac{1.9}{27}\right)\right] = \boxed{\mathbf{0.2260}}$. **(C)**

31. [**Subsection 57.2**] We apply formulas (57.5) and (57.6).

$$\begin{aligned}
\bar{x}_1 &= \frac{1000 + 1200}{40 + 50} = 24\tfrac{4}{9} \\
\bar{x}_2 &= \frac{500 + 600}{20 + 40} = 18\tfrac{1}{3} \\
\bar{x} &= \frac{1000 + 1200 + 500 + 600}{40 + 50 + 20 + 40} = \frac{3300}{150} = 22 \\
\hat{v} &= \frac{40\left(\frac{1000}{40} - 24\frac{4}{9}\right)^2 + 50\left(\frac{1200}{50} - 24\frac{4}{9}\right)^2 + 20\left(\frac{500}{20} - 18\frac{1}{3}\right)^2 + 40\left(\frac{600}{40} - 18\frac{1}{3}\right)^2}{2} \\
&= 677\tfrac{7}{9} \\
\hat{a} &= \frac{1}{150 - \frac{1}{150}(90^2 + 60^2)}\left[90(24\tfrac{4}{9} - 22)^2 + 60(18\tfrac{1}{3} - 22)^2 - (677\tfrac{7}{9})(1)\right] \\
&= \frac{666\frac{2}{3}}{72} = 9.2593
\end{aligned}$$

$$\hat{k} = \frac{\hat{v}}{\hat{a}} = \frac{677\frac{7}{9}}{9.2593} = 73.2$$
$$\hat{Z}_1 = \frac{90}{90+\hat{k}} = \frac{90}{90+73.2} = \boxed{\mathbf{0.5515}} \qquad \textbf{(E)}$$

32. [**Section 33.2**] The estimated difference of the logarithms of the hazard rates is $0.1(1) + 0.01(3) = 0.13$. The variance is

$$\begin{aligned}\operatorname{Var}(\hat{\beta}_1 + 3\hat{\beta}_2) &= \operatorname{Var}(\hat{\beta}_1) + \operatorname{Var}(\hat{\beta}_2) + 2\operatorname{Cov}(\hat{\beta}_1, \hat{\beta}_2)\\ &= 0.2 + 0.008(3^2) - 2(0.03)(3) = 0.092\end{aligned}$$

We exponentiate the estimated confidence interval for the logarithm of the relative hazard rate to obtain a confidence interval for the relative hazard rate.

$$\exp(0.13 \pm 1.96\sqrt{0.092}) = \boxed{\mathbf{(0.6285, 2.0637)}} \qquad \textbf{(D)}$$

33. [**Lesson 53**] For aggregate losses, the mean given θ is $\frac{\theta}{3}$ and the variance given θ is $\frac{2\theta^2}{6} - \left(\frac{\theta}{3}\right)^2 = \frac{2\theta^2}{9}$. Then

$$\begin{aligned}v &= \frac{2}{9}E[\theta^2] = \frac{2}{9}(\mu^2+\sigma^2) = \frac{2}{9}(3^2+1^2) = \frac{20}{9}\\ a &= \frac{1}{9}\operatorname{Var}(\theta) = \frac{1}{9}\sigma^2 = \frac{1}{9}\\ Z &= \frac{a}{a+v} = \boxed{\mathbf{\frac{1}{21}}} \qquad \textbf{(A)}\end{aligned}$$

34. [**Section 6.3**] We are given that the average return $\alpha = 0.05$, so the lognormal parameter $\mu = 0.05 - 0.5(0.3^2) = 0.005$. For a 3 year period, $\mu t = 0.015$ and $\sigma\sqrt{t} = 0.3\sqrt{3} = 0.5196$. For a positive return, we need the normal variable with these parameters to be greater than 0. The probability of that is $\Phi\left(\frac{0.015}{0.5196}\right) = \Phi(0.03) = \boxed{\mathbf{0.512}}$. **(B)**

35. [**Section 31.1**] We invert the information matrix to obtain the covariance matrix. The determinant of the matrix is $(100)(9) - (1)(1) = 899$, so the inverse matrix is

$$\frac{1}{899}\begin{pmatrix} 9 & -1 \\ -1 & 100 \end{pmatrix}.$$

The mean of a Pareto is $g(\alpha,\theta) = \frac{\theta}{\alpha-1}$. We use equation (31.2) to calculate the variance. In the 2-variable case, the equation reduces to

$$\operatorname{Var}[g(\alpha,\theta)] = \operatorname{Var}(\alpha)\left(\frac{\partial g}{\partial \alpha}\right)^2 + \operatorname{Var}(\theta)\left(\frac{\partial g}{\partial \theta}\right)^2 + 2\operatorname{Cov}(\alpha,\theta)\left(\frac{\partial g}{\partial \alpha}\right)\left(\frac{\partial g}{\partial \theta}\right)$$

so we have

$$\begin{aligned}\frac{\partial g}{\partial \alpha} &= -\frac{\theta}{(\alpha-1)^2} = -\frac{600}{2^2} = -150\\ \frac{\partial g}{\partial \theta} &= \frac{1}{\alpha-1} = \frac{1}{2}\\ \text{Var(estimated mean)} &= \frac{1}{899}\left[9(-150)^2 + 100\left(\frac{1}{2}\right)^2 + 2(-1)(-150)\left(\frac{1}{2}\right)\right]\\ &= 225.445\end{aligned}$$

The width of a 95% confidence interval is then $2(1.96)\sqrt{225.445} = \boxed{\mathbf{58.86}}$ **(A)**

36. [**Lesson 8**] Let X be loss size. Since $F(0) = 0$, $E[X] = e(0) = 3000$. Then

$$\begin{aligned}
E[X] &= E[X \wedge d] + e(d)[1 - F(d)] \\
3000 &= E[X \wedge 500] + e(500)[1 - F(500)] \\
&= E[X \wedge 500] + (3500)(0.8) \\
E[X \wedge 500] &= 200 \\
3000 &= E[X \wedge 10{,}000] + e(10{,}000)[1 - F(10{,}000)] \\
&= E[X \wedge 10{,}000] + (2500)(0.2) \\
E[X \wedge 10{,}000] &= 2500 \\
E[X \wedge 10{,}000] - E[X \wedge 500] &= 2500 - 200 = \boxed{\mathbf{2300}} \qquad \textbf{(D)}
\end{aligned}$$

37. [**Lesson 28**] The claims are truncated, not censored, at 10,000. The probability of seeing any claim is $F(10{,}000)$. Any likelihood developed before considering this condition must be divided by this condition.

The likelihood of each of the 100 claims less than 1000, if not for the condition, is $F(1000)$. The conditional likelihood, conditional on the claim being below 10,000, is $\frac{F(1000)}{F(10{,}000)}$.

The likelihood of each of the 75 claims between 1000 and 5000, if not for the condition, is $F(5000) - F(1000)$. The conditional likelihood is $\frac{F(5000)-F(1000)}{F(10{,}000)}$.

The likelihood of each of the 25 claims between 5000 and 10,000, if not for the condition, is $F(10{,}000)-F(5000)$. The conditional likelihood is $\frac{F(10{,}000)-F(5000)}{F(10{,}000)}$.

Multiplying all these 200 likelihoods together we get answer **(B)**.

38. [**Lesson 18**] The sample mean is an unbiased estimator of the population mean, and if the population variance is finite (as it is if it has an exponential distribution), the sample mean is a consistent estimator of the population mean. (C) and (E) are therefore true. For a sample of size 2, the sample median is the sample mean, so (B) is true. (D) is proved in *Loss Models*. That leaves (A). **(A)** is false, because (B) is true and the median of an exponential is not the mean. In fact, it is the mean times $\ln 2$. So the sample median, which is an unbiased estimator of the mean, and therefore has an expected value of θ, does not have expected value $\theta \ln 2$, the value of the median.

39. [**Lesson 27**] For a mixture F is the weighted average of the F's of the individual distribution. The median of the mixture F is then the number m such that

$$wF_1(x) + (1 - w)F_2(x) = 0.5$$

where w is the weight. Here, it is more convenient to use survival functions. (The median is the number m such that $S(m) = 0.5$) $m = 5$. We have:

$$\begin{aligned}
we^{-5/3} + (1 - w)e^{-5/x} &= 0.5 \\
w(e^{-5/3} - e^{-5/x}) &= 0.5 - e^{-5/x} \\
w &= \frac{0.5 - e^{-5/x}}{e^{-5/3} - e^{-5/x}}
\end{aligned}$$

In order for this procedure to work, w must be between 0 and 1. Note that since $x > 3$, $-\frac{5}{3} < -\frac{5}{x}$, so the denominator is negative. For $w > 0$, we need

$$\begin{aligned}
0.5 - e^{-5/x} &< 0 \\
e^{-5/x} &> 0.5 \\
-5/x &> \ln 0.5
\end{aligned}$$

$$
\begin{aligned}
5/x &< \ln 2 \\
x &> \frac{5}{\ln 2} = 7.2135
\end{aligned}
$$

For $w < 1$, we need

$$
\begin{aligned}
e^{-5/x} - 0.5 &< e^{-5/x} - e^{-5/3} \\
0.5 &> e^{-5/3} = 0.1889
\end{aligned}
$$

and this is always true. So percentile matching works when $\boxed{\boldsymbol{x > 7.2135}}$. **(E)**

40. [**Section 37.2**] 9 is the 4th point. With 5 points, each point is assigned the center of its 20% range (0%–20%, 20%–40%, etc.), so 9 goes to $\boxed{\textbf{70\%}}$. **(D)**