Course 4

Fall 2002
Society of Actuaries

****BEGINNING OF EXAMINATION****

1. For a stationary AR(2) process, you are given:

$$r_1 = 0.5$$

$$r_2 = -0.2$$

Calculate ϕ_2.

(A) -0.8

(B) -0.6

(C) -0.2

(D) 0.6

(E) 0.8

2. You are given the following claim data for automobile policies:

200 255 295 320 360 420 440 490 500 520 1020

Calculate the smoothed empirical estimate of the 45th percentile.

(A) 358

(B) 371

(C) 384

(D) 390

(E) 396

3. You are given:

(i) The number of claims made by an individual insured in a year has a Poisson distribution with mean λ.

(ii) The prior distribution for λ is gamma with parameters $a = 1$ and $q = 1.2$.

Three claims are observed in Year 1, and no claims are observed in Year 2.

Using Bühlmann credibility, estimate the number of claims in Year 3.

(A) 1.35

(B) 1.36

(C) 1.40

(D) 1.41

(E) 1.43

4. In a study of claim payment times, you are given:

(i) The data were not truncated or censored.

(ii) At most one claim was paid at any one time.

(iii) The Nelson-Aalen estimate of the cumulative hazard function, $H(t)$, immediately following the second paid claim, was 23/132.

Determine the Nelson-Aalen estimate of the cumulative hazard function, $H(t)$, immediately following the fourth paid claim.

(A) 0.35

(B) 0.37

(C) 0.39

(D) 0.41

(E) 0.43

5. You fit the following model to eight observations:

$$Y = a + bX + e$$

You are given:

$$\hat{b} = 2.065$$

$$\sum \left(X_i - \overline{X} \right)^2 = 42$$

$$\sum \left(Y_i - \overline{Y} \right)^2 = 182$$

Determine R^2.

(A) 0.48

(B) 0.62

(C) 0.83

(D) 0.91

(E) 0.98

6. The number of claims follows a negative binomial distribution with parameters b and r, where b is unknown and r is known. You wish to estimate b based on n observations, where $\overline{x}$ is the mean of these observations.

Determine the maximum likelihood estimate of b.

(A) $\dfrac{\overline{x}}{r^2}$

(B) $\dfrac{\overline{x}}{r}$

(C) $\overline{x}$

(D) $r\overline{x}$

(E) $r^2\overline{x}$

7. You are given the following information about a credibility model:

First Observation	Unconditional Probability	Bayesian Estimate of Second Observation
1	1/3	1.50
2	1/3	1.50
3	1/3	3.00

Determine the Bühlmann credibility estimate of the second observation, given that the first observation is 1.

(A) 0.75

(B) 1.00

(C) 1.25

(D) 1.50

(E) 1.75

8. For a survival study, you are given:

(i) The Product-Limit estimator $\hat{S}(t_0)$ is used to construct confidence intervals for $S(t_0)$.

(ii) The 95% log-transformed confidence interval for $S(t_0)$ is $(0.695, 0.843)$.

Determine $\hat{S}(t_0)$.

(A) 0.758

(B) 0.762

(C) 0.765

(D) 0.769

(E) 0.779

9. You are given the following information about an AR(1) model with mean 0:

$$r_2 = 0.215$$

$$r_3 = -0.100$$

$$y_T = -0.431$$

Calculate the forecasted value of y_{T+1}.

(A) –0.2

(B) –0.1

(C) 0.0

(D) 0.1

(E) 0.2

10. A random sample of three claims from a dental insurance plan is given below:

225 525 950

Claims are assumed to follow a Pareto distribution with parameters $q = 150$ and a .

Determine the maximum likelihood estimate of a .

(A) Less than 0.6

(B) At least 0.6, but less than 0.7

(C) At least 0.7, but less than 0.8

(D) At least 0.8, but less than 0.9

(E) At least 0.9

11. An insurer has data on losses for four policyholders for 7 years. The loss from the i^{th} policyholder for year j is X_{ij}.

You are given:

$$\sum_{i=1}^{4}\sum_{j=1}^{7}\left(X_{ij}-\overline{X}_i\right)^2 = 33.60$$

$$\sum_{i=1}^{4}\left(\overline{X}_i-\overline{X}\right)^2 = 3.30$$

Using nonparametric empirical Bayes estimation, calculate the Bühlmann credibility factor for an individual policyholder.

(A) Less than 0.74

(B) At least 0.74, but less than 0.77

(C) At least 0.77, but less than 0.80

(D) At least 0.80, but less than 0.83

(E) At least 0.83

12. For the three variables Y, X_2 and X_3, you are given the following sample correlation coefficients:

$$r_{YX_2} = 0.6$$
$$r_{YX_3} = 0.5$$
$$r_{X_2 X_3} = 0.4$$

Calculate $r_{YX_2 \cdot X_3}$, the partial correlation coefficient between Y and X_2.

(A) 0.50

(B) 0.55

(C) 0.58

(D) 0.64

(E) 0.73

13. Losses come from an equally weighted mixture of an exponential distribution with mean m_1, and an exponential distribution with mean m_2.

Determine the least upper bound for the coefficient of variation of this distribution.

(A) 1

(B) $\sqrt{2}$

(C) $\sqrt{3}$

(D) 2

(E) $\sqrt{5}$

14. You are given the following information about a commercial auto liability book of business:

(i) Each insured's claim count has a Poisson distribution with mean λ, where λ has a gamma distribution with $\alpha = 1.5$ and $\theta = 0.2$.

(ii) Individual claim size amounts are independent and exponentially distributed with mean 5000.

(iii) The full credibility standard is for aggregate losses to be within 5% of the expected with probability 0.90.

Using classical credibility, determine the expected number of claims required for full credibility.

(A) 2165

(B) 2381

(C) 3514

(D) 7216

(E) 7938

15. An insurance company uses a proportional hazards model to investigate whether to have different premium rates for two different classes of drivers.

You are given:

(i) The model has a single covariate: $Z = 1$ if the driver is in class 1, $Z = 0$ if the driver is in class 2.

(ii) The model is $h(t|Z) = h_0(t)\exp(b\,Z)$, where $h_0(t)$ is an arbitrary baseline hazard rate and b is the parameter.

(iii) The estimated relative risk for drivers in class 1 compared to drivers in class 2 is 1.822.

(iv) The information matrix is $I(b) = 3.968$, where b is the partial maximum likelihood estimate of b.

You use Wald's test to test the hypothesis $b = 0$.

Determine the value of the test statistic.

(A) 0.7

(B) 0.9

(C) 1.4

(D) 2.2

(E) 5.7

16. Which of the following statements about stationary mixed autoregressive-moving average models is true?

(A) A necessary condition for stationarity is that each parameter f_i must have an absolute value less than 1.

(B) The autocorrelation function approaches f_1 as the displacement increases.

(C) The difference between adjacent forecasted values approaches d as the number of periods ahead increases.

(D) The forecasted values approach the mean as the number of periods ahead increases.

(E) These models are particularly well-suited to long forecasting horizons.

17. You are given:

(i) A sample of claim payments is:

29 64 90 135 182

(ii) Claim sizes are assumed to follow an exponential distribution.

(iii) The mean of the exponential distribution is estimated using the method of moments.

Calculate the value of the Kolmogorov-Smirnov test statistic.

(A) 0.14

(B) 0.16

(C) 0.19

(D) 0.25

(E) 0.27

18. You are given:

(i) Annual claim frequency for an individual policyholder has mean l and variance s^2.

(ii) The prior distribution for l is uniform on the interval [0.5, 1.5].

(iii) The prior distribution for s^2 is exponential with mean 1.25.

A policyholder is selected at random and observed to have no claims in Year 1.

Using Bühlmann credibility, estimate the number of claims in Year 2 for the selected policyholder.

(A) 0.56

(B) 0.65

(C) 0.71

(D) 0.83

(E) 0.94

19. You study the time between accidents and reports of claims. The study was terminated at time 3.

You are given:

Time of Accident	Time between Accident and Claim Report	Number of Reported Claims
0	1	18
0	2	13
0	3	9
1	1	14
1	2	10
2	1	11

Use the Product-Limit estimator to estimate the conditional probability that the time between accident and claim report is less than 2, given that it does not exceed 3.

(A) Less than 0.4

(B) At least 0.4, but less than 0.5

(C) At least 0.5, but less than 0.6

(D) At least 0.6, but less than 0.7

(E) At least 0.7

20. You study the impact of education and number of children on the wages of working women using the following model:

$$Y = a + b_1 E + b_2 F + c_1 G + c_2 H + e$$

where $Y = ln(wages)$

$$E = \begin{cases} 1 & \text{if the woman has not completed high school} \\ 0 & \text{if the woman has completed high school} \\ -1 & \text{if the woman has post-secondary education} \end{cases}$$

$$F = \begin{cases} 1 & \text{if the woman has completed high school} \\ 0 & \text{if the woman has not completed high school} \\ -1 & \text{if the woman has post-secondary education} \end{cases}$$

$$G = \begin{cases} 1 & \text{if the woman has no children} \\ 0 & \text{if the woman has 1 or 2 children} \\ -1 & \text{if the woman has more than 2 children} \end{cases}$$

$$H = \begin{cases} 1 & \text{if the woman has 1 or 2 children} \\ 0 & \text{if the woman has no children} \\ -1 & \text{if the woman has more than 2 children} \end{cases}$$

Determine the expected difference between $ln(wages)$ of a working woman who has post-secondary education and more than 2 children and $ln(wages)$ of the average for all working women.

(A) $a - b_1 - b_2$

(B) $b_1 + b_2$

(C) $-b_1 - b_2$

(D) $a - b_1 - b_2 + c_2$

(E) $-b_1 - b_2 - c_1 - c_2$

21. You are given:

(i) The prior distribution of the parameter Θ has probability density function:

$$p(q) = \frac{1}{q^2}, \quad 1 < q < \infty$$

(ii) Given $\Theta = q$, claim sizes follow a Pareto distribution with parameters $a = 2$ and q.

A claim of 3 is observed.

Calculate the posterior probability that Θ exceeds 2.

(A) 0.33

(B) 0.42

(C) 0.50

(D) 0.58

(E) 0.64

22. You are given:

(i)

t	y_t
0	1.0
1	1.2
2	1.3

(ii) $y_t = 0$, for $t < 0$

(iii) $a = 0.6$

Use double exponential smoothing to determine $\tilde{\tilde{y}}_2$.

(A) 0.96

(B) 0.99

(C) 1.16

(D) 1.20

(E) 1.33

23. You are given:

(i) Losses follow an exponential distribution with mean q.

(ii) A random sample of 20 losses is distributed as follows:

Loss Range	Frequency
[0, 1000]	7
(1000, 2000]	6
(2000, ∞)	7

Calculate the maximum likelihood estimate of q.

(A) Less than 1950

(B) At least 1950, but less than 2100

(C) At least 2100, but less than 2250

(D) At least 2250, but less than 2400

(E) At least 2400

24. You are given:

(i) The amount of a claim, X, is uniformly distributed on the interval $[0,q]$.

(ii) The prior density of q is $p(q)=\dfrac{500}{q^2}, \quad q>500$.

Two claims, $x_1=400$ and $x_2=600$, are observed. You calculate the posterior distribution as:

$$f(q|x_1,x_2)=3\left(\frac{600^3}{q^4}\right), \quad q>600$$

Calculate the Bayesian premium, $\mathrm{E}(X_3|x_1,x_2)$.

(A) 450

(B) 500

(C) 550

(D) 600

(E) 650

25-26. *Use the following information for questions* 25 *and* 26.

The claim payments on a sample of ten policies are:

$$2 \quad 3 \quad 3 \quad 5 \quad 5^{+} \quad 6 \quad 7 \quad 7^{+} \quad 9 \quad 10^{+}$$

+ indicates that the loss exceeded the policy limit

25. Using the Product-Limit estimator, calculate the probability that the loss on a policy exceeds 8.

(A) 0.20

(B) 0.25

(C) 0.30

(D) 0.36

(E) 0.40

25-26. *(Repeated for convenience) Use the following information for questions* 25 *and* 26.

The claim payments on a sample of ten policies are:

$$2 \quad 3 \quad 3 \quad 5 \quad 5^+ \quad 6 \quad 7 \quad 7^+ \quad 9 \quad 10^+$$

+ indicates that the loss exceeded the policy limit

26. You use the log-rank test to test the hypothesis that losses follow a Weibull distribution with survival function:

$$S_0(x) = e^{-(x/5)^2}, \quad 0 < x < \infty$$

Determine the result of the test.

(A) Reject at the 0.005 significance level.

(B) Reject at the 0.010 significance level, but not at the 0.005 level.

(C) Reject at the 0.025 significance level, but not at the 0.010 level.

(D) Reject at the 0.050 significance level, but not at the 0.025 level.

(E) Do not reject at the 0.050 significance level.

27. For the multiple regression model $Y = b_1 + b_2 X_2 + b_3 X_3 + b_4 X_4 + b_5 X_5 + b_6 X_6 + e$, you are given:

(i) $N = 3{,}120$

(ii) $\text{TSS} = 15{,}000$

(iii) H_0: $b_4 = b_5 = b_6 = 0$

(iv) $R^2_{UR} = 0.38$

(v) $\text{RSS}_R = 5{,}565$

Determine the value of the F statistic for testing H_0.

(A) Less than 10

(B) At least 10, but less than 12

(C) At least 12, but less than 14

(D) At least 14, but less than 16

(E) At least 16

28. You are given the following observed claim frequency data collected over a period of 365 days:

Number of Claims per Day	Observed Number of Days
0	50
1	122
2	101
3	92
4+	0

Fit a Poisson distribution to the above data, using the method of maximum likelihood.

Regroup the data, by number of claims per day, into four groups:

0 1 2 3+

Apply the chi-square goodness-of-fit test to evaluate the null hypothesis that the claims follow a Poisson distribution.

Determine the result of the chi-square test.

(A) Reject at the 0.005 significance level.

(B) Reject at the 0.010 significance level, but not at the 0.005 level.

(C) Reject at the 0.025 significance level, but not at the 0.010 level.

(D) Reject at the 0.050 significance level, but not at the 0.025 level.

(E) Do not reject at the 0.050 significance level.

29. You are given the following joint distribution:

	Θ	
X	0	1
0	0.4	0.1
1	0.1	0.2
2	0.1	0.1

For a given value of Θ and a sample of size 10 for X:

$$\sum_{i=1}^{10} x_i = 10$$

Determine the Bühlmann credibility premium.

(A) 0.75

(B) 0.79

(C) 0.82

(D) 0.86

(E) 0.89

30. Which of the following is not an objection to the use of R^2 to compare the validity of regression results under alternative specifications of a multiple linear regression model?

(A) The F statistic used to test the null hypothesis that none of the explanatory variables helps explain variation of Y about its mean is a function of R^2 and degrees of freedom.

(B) Increasing the number of independent variables in the regression equation can never lower R^2 and is likely to raise it.

(C) When the model is constrained to have zero intercept, the ratio of regression sum of squares to total sum of squares need not lie within the range [0,1].

(D) Subtracting the value of one of the independent variables from both sides of the regression equation can change the value of R^2 while leaving the residuals unaffected.

(E) Because R^2 is interpreted assuming the model is correct, it provides no direct procedure for comparing alternative specifications.

31. You are given:

x	0	1	2	3
$\Pr[X = x]$	0.5	0.3	0.1	0.1

The method of moments is used to estimate the population mean, μ, and variance, σ^2, by $\bar{X}$ and $S_n^2 = \dfrac{\sum (X_i - \bar{X})^2}{n}$, respectively.

Calculate the bias of S_n^2, when $n = 4$.

(A) –0.72

(B) –0.49

(C) –0.24

(D) –0.08

(E) 0.00

32. You are given four classes of insureds, each of whom may have zero or one claim, with the following probabilities:

Class	Number of Claims	
	0	1
I	0.9	0.1
II	0.8	0.2
III	0.5	0.5
IV	0.1	0.9

A class is selected at random (with probability ¼), and four insureds are selected at random from the class. The total number of claims is two.

If five insureds are selected at random from the same class, estimate the total number of claims using Bühlmann-Straub credibility.

(A) 2.0

(B) 2.2

(C) 2.4

(D) 2.6

(E) 2.8

33. The following results were obtained from a survival study, using the Product-Limit estimator:

t	$\hat{S}(t)$	$\sqrt{\hat{V}[\hat{S}(t)]}$
17	0.957	0.0149
25	0.888	0.0236
32	0.814	0.0298
36	0.777	0.0321
39	0.729	0.0348
42	0.680	0.0370
44	0.659	0.0378
47	0.558	0.0418
50	0.360	0.0470
54	0.293	0.0456
56	0.244	0.0440
57	0.187	0.0420
59	0.156	0.0404
62	0.052	0.0444

Determine the lower limit of the 95% linear confidence interval for $x_{0.75}$, the 75th percentile of the survival distribution.

(A) 32

(B) 36

(C) 50

(D) 54

(E) 56

34. You fit an AR(2) model to a series of 100 observations.

You are given:

k	$\hat{r}_k$
1	–0.01
2	0.01
3	–0.02
4	0.04
5	–0.03
6	–0.13
7	–0.23
8	–0.05
9	–0.01
10	0.05
11	–0.04
12	0.10

Calculate the Box-Pierce Q statistic based on the first twelve residual autocorrelations.

(A) 9.0

(B) 9.3

(C) 9.6

(D) 9.9

(E) 10.2

35. With the bootstrapping technique, the underlying distribution function is estimated by which of the following?

(A) The empirical distribution function

(B) A normal distribution function

(C) A parametric distribution function selected by the modeler

(D) Any of (A), (B) or (C)

(E) None of (A), (B) or (C)

36. You are given:

Number of Claims	Probability	Claim Size	Probability
0	$1/5$		
1	$3/5$	25	$1/3$
		150	$2/3$
2	$1/5$	50	$2/3$
		200	$1/3$

Claim sizes are independent.

Determine the variance of the aggregate loss.

(A) 4,050

(B) 8,100

(C) 10,500

(D) 12,510

(E) 15,612

37. You are given:

(i) Losses follow an exponential distribution with mean q.

(ii) A random sample of losses is distributed as follows:

Loss Range	Number of Losses
(0 – 100]	32
(100 – 200]	21
(200 – 400]	27
(400 – 750]	16
(750 – 1000]	2
(1000 – 1500]	2
Total	100

Estimate q by matching at the 80^{th} percentile.

(A) 249

(B) 253

(C) 257

(D) 260

(E) 263

38. You fit a two-variable linear regression model to 20 pairs of observations.

You are given:

(i) The sample mean of the independent variable is 100.

(ii) The sum of squared deviations from the mean of the independent variable is 2266.

(iii) The ordinary least-squares estimate of the intercept parameter is 68.73.

(iv) The error sum of squares (ESS) is 5348.

Determine the lower limit of the symmetric 95% confidence interval for the intercept parameter.

(A) −273

(B) −132

(C) −70

(D) −8

(E) −3

39. You are given:

Class	Number of Insureds	Claim Count Probabilities				
		0	1	2	3	4
1	3000	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	0	0
2	2000	0	$\frac{1}{6}$	$\frac{2}{3}$	$\frac{1}{6}$	0
3	1000	0	0	$\frac{1}{6}$	$\frac{2}{3}$	$\frac{1}{6}$

A randomly selected insured has one claim in Year 1.

Determine the expected number of claims in Year 2 for that insured.

(A) 1.00

(B) 1.25

(C) 1.33

(D) 1.67

(E) 1.75

40. You are given the following information about a group of policies:

Claim Payment	Policy Limit
5	50
15	50
60	100
100	100
500	500
500	1000

Determine the likelihood function.

(A) $f(50)\, f(50)\, f(100)\, f(100)\, f(500)\, f(1000)$

(B) $f(50)\, f(50)\, f(100)\, f(100)\, f(500)\, f(1000) / [1-F(1000)]$

(C) $f(5)\, f(15)\, f(60)\, f(100)\, f(500)\, f(500)$

(D) $f(5)\, f(15)\, f(60)\, f(100)\, f(500)\, f(500) / [1-F(1000)]$

(E) $f(5)\, f(15)\, f(60)\, [1-F(100)]\, [1-F(500)]\, f(500)$

****END OF EXAMINATION****

Final Fall 2002 Course 4 Exam Answer Key

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