Fall 2003 Society of Actuaries

****BEGINNING OF EXAMINATION****

1. You are given the following information about a stationary AR(2) model:

(i) $\rho_1 = 0.5$

(ii) $\rho_2 = 0.1$

Determine ϕ_2.

(A) –0.2

(B) 0.1

(C) 0.4

(D) 0.7

(E) 1.0

2. You are given:

(i) Losses follow a loglogistic distribution with cumulative distribution function:

$$F(x) = \frac{(x/\theta)^{\gamma}}{1+(x/\theta)^{\gamma}}$$

(ii) The sample of losses is:

10 35 80 86 90 120 158 180 200 210 1500

Calculate the estimate of θ by percentile matching, using the 40^{th} and 80^{th} empirically smoothed percentile estimates.

(A) Less than 77

(B) At least 77, but less than 87

(C) At least 87, but less than 97

(D) At least 97, but less than 107

(E) At least 107

3. You are given:

(i) The number of claims has a Poisson distribution.

(ii) Claim sizes have a Pareto distribution with parameters $\theta = 0.5$ and $\alpha = 6$.

(iii) The number of claims and claim sizes are independent.

(iv) The observed pure premium should be within 2% of the expected pure premium 90% of the time.

Determine the expected number of claims needed for full credibility.

(A) Less than 7,000

(B) At least 7,000, but less than 10,000

(C) At least 10,000, but less than 13,000

(D) At least 13,000, but less than 16,000

(E) At least 16,000

4. You study five lives to estimate the time from the onset of a disease to death. The times to death are:

2 3 3 3 7

Using a triangular kernel with bandwidth 2, estimate the density function at 2.5.

(A) 8/40

(B) 12/40

(C) 14/40

(D) 16/40

(E) 17/40

5. For the model $Y_i = \alpha + \beta X_i + \varepsilon_i$, where $i = 1, 2, ..., 10$, you are given:

(i) $X_i = \begin{cases} 1, & \text{if the } i\text{th individual belongs to a specified group} \\ 0, & \text{otherwise} \end{cases}$

(ii) 40 percent of the individuals belong to the specified group.

(iii) The least squares estimate of β is $\hat{\beta} = 4$.

(iv) $\sum \left(Y_i - \hat{\alpha} - \hat{\beta} X_i \right)^2 = 92$

Calculate the t statistic for testing $H_0\!:\beta = 0$.

(A) 0.9

(B) 1.2

(C) 1.5

(D) 1.8

(E) 2.1

6. You are given:

(i) Losses follow a Single-parameter Pareto distribution with density function:

$$f(x) = \frac{\alpha}{x^{(\alpha+1)}}, \quad x > 1, \quad 0 < \alpha < \infty$$

(ii) A random sample of size five produced three losses with values 3, 6 and 14, and two losses exceeding 25.

Determine the maximum likelihood estimate of α.

(A) 0.25

(B) 0.30

(C) 0.34

(D) 0.38

(E) 0.42

7. You are given:

(i) The annual number of claims for a policyholder has a binomial distribution with probability function:

$$p(x|q) = \binom{2}{x} q^x (1-q)^{2-x}, \quad x = 0, 1, 2$$

(ii) The prior distribution is:

$$\pi(q) = 4q^3, \; 0 < q < 1$$

This policyholder had one claim in each of Years 1 and 2.

Determine the Bayesian estimate of the number of claims in Year 3.

(A) Less than 1.1

(B) At least 1.1, but less than 1.3

(C) At least 1.3, but less than 1.5

(D) At least 1.5, but less than 1.7

(E) At least 1.7

8. For a sample of dental claims $x_1, x_2, \ldots, x_{10}$, you are given:

(i) $\sum x_i = 3860$ and $\sum x_i^2 = 4{,}574{,}802$

(ii) Claims are assumed to follow a lognormal distribution with parameters μ and σ.

(iii) μ and σ are estimated using the method of moments.

Calculate $E\left[X \wedge 500\right]$ for the fitted distribution.

(A) Less than 125

(B) At least 125, but less than 175

(C) At least 175, but less than 225

(D) At least 225, but less than 275

(E) At least 275

9. You are given:

(i) Y_{tij} is the loss for the *j*th insured in the *i*th group in Year *t*.

(ii) $\bar{Y}_{ti}$ is the mean loss in the *i*th group in Year *t*.

(iii) $X_{ij} = \begin{cases} 0, & \text{if the } j\text{th insured is in the first group } (i = 1) \\ 1, & \text{if the } j\text{th insured is in the second group } (i = 2) \end{cases}$

(iv) $Y_{2ij} = \delta + \phi Y_{1ij} + \theta X_{ij} + \varepsilon_{ij}$, where $i = 1, 2$ and $j = 1, 2, \ldots, n$

(v) $\bar{Y}_{21} = 30$, $\bar{Y}_{22} = 37$, $\bar{Y}_{11} = 40$, $\bar{Y}_{12} = 41$

(vi) $\hat{\phi} = 0.75$

Determine the least-squares estimate of θ.

(A) 5.25

(B) 5.50

(C) 5.75

(D) 6.00

(E) 6.25

10. Two independent samples are combined yielding the following ranks:

Sample I: 1, 2, 3, 4, 7, 9, 13, 19, 20
Sample II: 5, 6, 8, 10, 11, 12, 14, 15, 16, 17, 18

You test the null hypothesis that the two samples are from the same continuous distribution.

The variance of the rank sum statistic is:

$$\frac{nm(n+m+1)}{12}$$

Using the classical approximation for the two-tailed rank sum test, determine the *p*-value.

(A) 0.015

(B) 0.021

(C) 0.105

(D) 0.210

(E) 0.420

11. You are given:

(i) Claim counts follow a Poisson distribution with mean θ.

(ii) Claim sizes follow an exponential distribution with mean 10θ.

(iii) Claim counts and claim sizes are independent, given θ.

(iv) The prior distribution has probability density function:

$$\pi(\theta) = \frac{5}{\theta^6}, \quad \theta > 1$$

Calculate Bühlmann's *k* for aggregate losses.

(A) Less than 1

(B) At least 1, but less than 2

(C) At least 2, but less than 3

(D) At least 3, but less than 4

(E) At least 4

12. You are given:

(i) A survival study uses a Cox proportional hazards model with covariates Z_1 and Z_2, each taking the value 0 or 1.

(ii) The maximum partial likelihood estimate of the coefficient vector is:

$$\left(\hat{\beta}_1, \hat{\beta}_2\right) = (0.71, 0.20)$$

(iii) The baseline survival function at time t_0 is estimated as $\hat{S}(t_0) = 0.65$.

Estimate $S(t_0)$ for a subject with covariate values $Z_1 = Z_2 = 1$.

(A) 0.34

(B) 0.49

(C) 0.65

(D) 0.74

(E) 0.84

13. You are given:

(i) Z_1 and Z_2 are independent N(0,1) random variables.

(ii) a, b, c, d, e, f are constants.

(iii) $Y = a + bZ_1 + cZ_2$ and $X = d + eZ_1 + f Z_2$

Determine $\mathrm{E}(Y|X)$.

(A) a

(B) $a + (b + c)(X - d)$

(C) $a + (be + cf)(X - d)$

(D) $a + [(be + cf)/(e^2 + f^2)]X$

(E) $a + [(be + cf)/(e^2 + f^2)](X - d)$

14. You are given:

(i) Losses on a company's insurance policies follow a Pareto distribution with probability density function:

$$f\left(x|\theta\right)=\frac{\theta}{\left(x+\theta\right)^{2}}, \quad 0<x<\infty$$

(ii) For half of the company's policies $\theta=1$, while for the other half $\theta=3$.

For a randomly selected policy, losses in Year 1 were 5.

Determine the posterior probability that losses for this policy in Year 2 will exceed 8.

(A) 0.11

(B) 0.15

(C) 0.19

(D) 0.21

(E) 0.27

15. You are given total claims for two policyholders:

	Year			
Policyholder	1	2	3	4
X	730	800	650	700
Y	655	650	625	750

Using the nonparametric empirical Bayes method, determine the Bühlmann credibility premium for Policyholder Y.

(A) 655

(B) 670

(C) 687

(D) 703

(E) 719

16. A particular line of business has three types of claims. The historical probability and the number of claims for each type in the current year are:

Type	Historical Probability	Number of Claims in Current Year
A	0.2744	112
B	0.3512	180
C	0.3744	138

You test the null hypothesis that the probability of each type of claim in the current year is the same as the historical probability.

Calculate the chi-square goodness-of-fit test statistic.

(A) Less than 9

(B) At least 9, but less than 10

(C) At least 10, but less than 11

(D) At least 11, but less than 12

(E) At least 12

17. Which of the following is false?

(A) If the characteristics of a stochastic process change over time, then the process is nonstationary.

(B) Representing a nonstationary time series by a simple algebraic model is often difficult.

(C) Differences of a homogeneous nonstationary time series will always be nonstationary.

(D) If a time series is stationary, then its mean, variance and, for any lag *k*, covariance must also be stationary.

(E) If the autocorrelation function for a time series is zero (or close to zero) for all lags $k > 0$, then no model can provide useful minimum mean-square-error forecasts of future values other than the mean.

18. The information associated with the maximum likelihood estimator of a parameter θ is $4n$, where n is the number of observations.

Calculate the asymptotic variance of the maximum likelihood estimator of 2θ.

(A) $\frac{1}{2n}$

(B) $\frac{1}{n}$

(C) $\frac{4}{n}$

(D) $8n$

(E) $16n$

19. You are given:

(i) The probability that an insured will have at least one loss during any year is *p*.

(ii) The prior distribution for *p* is uniform on $[0, 0.5]$.

(iii) An insured is observed for 8 years and has at least one loss every year.

Determine the posterior probability that the insured will have at least one loss during Year 9.

(A) 0.450

(B) 0.475

(C) 0.500

(D) 0.550

(E) 0.625

20. At the beginning of each of the past 5 years, an actuary has forecast the annual claims for a group of insureds. The table below shows the forecasts (*X*) and the actual claims (*Y*). A two-variable linear regression model is used to analyze the data.

t	X_t	Y_t
1	475	254
2	254	463
3	463	515
4	515	567
5	567	605

You are given:

(i) The null hypothesis is $H_0: \alpha = 0,\ \beta = 1$.

(ii) The unrestricted model fit yields ESS = 69,843.

Which of the following is true regarding the *F* test of the null hypothesis?

(A) The null hypothesis is not rejected at the 0.05 significance level.

(B) The null hypothesis is rejected at the 0.05 significance level, but not at the 0.01 level.

(C) The numerator has 3 degrees of freedom.

(D) The denominator has 2 degrees of freedom.

(E) The *F* statistic cannot be determined from the information given.

21-22. *Use the following information for questions* 21 *and* 22.

For a survival study with censored and truncated data, you are given:

Time (t)	Number at Risk at Time t	Failures at Time t
1	30	5
2	27	9
3	32	6
4	25	5
5	20	4

21. The probability of failing at or before Time 4, given survival past Time 1, is ${}_3q_1$.

Calculate Greenwood's approximation of the variance of ${}_3\hat{q}_1$.

(A) 0.0067

(B) 0.0073

(C) 0.0080

(D) 0.0091

(E) 0.0105

21-22. (*Repeated for convenience*) *Use the following information for questions* 21 *and* 22.

For a survival study with censored and truncated data, you are given:

Time (t)	Number at Risk at Time t	Failures at Time t
1	30	5
2	27	9
3	32	6
4	25	5
5	20	4

22. Calculate the 95% log-transformed confidence interval for $H(3)$, based on the Nelson-Aalen estimate.

(A) (0.30, 0.89)

(B) (0.31, 1.54)

(C) (0.39, 0.99)

(D) (0.44, 1.07)

(E) (0.56, 0.79)

23. You are given:

(i) Two risks have the following severity distributions:

Amount of Claim	Probability of Claim Amount for Risk 1	Probability of Claim Amount for Risk 2
250	0.5	0.7
2,500	0.3	0.2
60,000	0.2	0.1

(ii) Risk 1 is twice as likely to be observed as Risk 2.

A claim of 250 is observed.

Determine the Bühlmann credibility estimate of the second claim amount from the same risk.

(A) Less than 10,200

(B) At least 10,200, but less than 10,400

(C) At least 10,400, but less than 10,600

(D) At least 10,600, but less than 10,800

(E) At least 10,800

24. You are given:

(i) A sample $x_1, x_2, \ldots, x_{10}$ is drawn from a distribution with probability density function:

$$\tfrac{1}{2}\left[\frac{1}{\theta}\exp(-\frac{x}{\theta}) + \frac{1}{\sigma}\exp(-\frac{x}{\sigma})\right], \quad 0 < x < \infty$$

(ii) $\theta > \sigma$

(iii) $\sum x_i = 150$ and $\sum x_i^2 = 5000$

Estimate θ by matching the first two sample moments to the corresponding population quantities.

(A) 9

(B) 10

(C) 15

(D) 20

(E) 21

25. You are given the following time-series model:

$$y_t = 0.8y_{t-1} + 2 + \varepsilon_t - 0.5\varepsilon_{t-1}$$

Which of the following statements about this model is false?

(A) $\rho_1 = 0.4$

(B) $\rho_k < \rho_1,\ k = 2,3,4,....$

(C) The model is ARMA(1,1).

(D) The model is stationary.

(E) The mean, μ, is 2.

26. You are given a sample of two values, 5 and 9.

You estimate Var(X) using the estimator $g(X_1, X_2) = \frac{1}{2}\sum (X_i - \bar{X})^2$.

Determine the bootstrap approximation to the mean square error of g.

(A) 1

(B) 2

(C) 4

(D) 8

(E) 16

27. You are given:

(i) The number of claims incurred in a month by any insured has a Poisson distribution with mean λ.

(ii) The claim frequencies of different insureds are independent.

(iii) The prior distribution is gamma with probability density function:

$$f(\lambda) = \frac{(100\lambda)^6 e^{-100\lambda}}{120\lambda}$$

(iv)

Month	Number of Insureds	Number of Claims
1	100	6
2	150	8
3	200	11
4	300	?

Determine the Bühlmann-Straub credibility estimate of the number of claims in Month 4.

(A) 16.7

(B) 16.9

(C) 17.3

(D) 17.6

(E) 18.0

28. You fit a Pareto distribution to a sample of 200 claim amounts and use the likelihood ratio test to test the hypothesis that $\alpha = 1.5$ and $\theta = 7.8$.

You are given:

(i) The maximum likelihood estimates are $\hat{\alpha}$ = 1.4 and $\hat{\theta}$ = 7.6.

(ii) The natural logarithm of the likelihood function evaluated at the maximum likelihood estimates is –817.92.

(iii) $\sum \ln(x_i + 7.8) = 607.64$

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.

29. You are given:

(i) The model is $Y_i = \beta X_i + \varepsilon_i$, i = 1, 2, 3.

(ii)

i	X_i	$\mathrm{Var}(\varepsilon_i)$
1	1	1
2	2	9
3	3	16

(iii) The ordinary least squares residuals are $\hat{\varepsilon}_i = Y_i - \hat{\beta} X_i$, i = 1, 2, 3.

Determine $E\left(\hat{\varepsilon}_1^2 \middle| X_1, X_2, X_3\right)$.

(A) 1.0

(B) 1.8

(C) 2.7

(D) 3.7

(E) 7.6

30. For a sample of 15 losses, you are given:

(i)

Interval	Observed Number of Losses
(0, 2]	5
(2, 5]	5
(5, ∞)	5

(ii) Losses follow the uniform distribution on $(0, \theta)$.

Estimate θ by minimizing the function $\sum_{j=1}^{3} \frac{(E_j - O_j)^2}{O_j}$, where E_j is the expected number of losses in the *j*th interval and O_j is the observed number of losses in the *j*th interval.

(A) 6.0

(B) 6.4

(C) 6.8

(D) 7.2

(E) 7.6

31. You are given:

(i) The probability that an insured will have exactly one claim is θ.

(ii) The prior distribution of θ has probability density function:

$$\pi(\theta) = \frac{3}{2}\sqrt{\theta},\ \ 0 < \theta < 1$$

A randomly chosen insured is observed to have exactly one claim.

Determine the posterior probability that θ is greater than 0.60.

(A) 0.54

(B) 0.58

(C) 0.63

(D) 0.67

(E) 0.72

32. The distribution of accidents for 84 randomly selected policies is as follows:

Number of Accidents	Number of Policies
0	32
1	26
2	12
3	7
4	4
5	2
6	1
Total	84

Which of the following models best represents these data?

(A) Negative binomial

(B) Discrete uniform

(C) Poisson

(D) Binomial

(E) Either Poisson or Binomial

33. A time series y_t follows an ARIMA(1,1,1) model with $\phi_1 = 0.7$, $\theta_1 = -0.3$ and $\sigma_\varepsilon^2 = 1.0$.

Determine the variance of the forecast error two steps ahead.

(A) 1

(B) 5

(C) 8

(D) 10

(E) 12

34. You are given:

(i) Low-hazard risks have an exponential claim size distribution with mean θ.

(ii) Medium-hazard risks have an exponential claim size distribution with mean 2θ.

(iii) High-hazard risks have an exponential claim size distribution with mean 3θ.

(iv) No claims from low-hazard risks are observed.

(v) Three claims from medium-hazard risks are observed, of sizes 1, 2 and 3.

(vi) One claim from a high-hazard risk is observed, of size 15.

Determine the maximum likelihood estimate of θ.

(A) 1

(B) 2

(C) 3

(D) 4

(E) 5

35. You are given:

(i) $X_{\text{partial}} =$ pure premium calculated from partially credible data

(ii) $\mu = \mathrm{E}\left[X_{\text{partial}}\right]$

(iii) Fluctuations are limited to $\pm k\,\mu$ of the mean with probability P

(iv) Z = credibility factor

Which of the following is equal to P?

(A) $\Pr\left[\mu - k\mu \le X_{\text{partial}} \le \mu + k\mu\right]$

(B) $\Pr\left[Z\mu - k \le Z\,X_{\text{partial}} \le Z\mu + k\right]$

(C) $\Pr\left[Z\mu - \mu \le Z\,X_{\text{partial}} \le Z\mu + \mu\right]$

(D) $\Pr\left[1 - k \le Z\,X_{\text{partial}} + (1 - Z)\mu \le 1 + k\right]$

(E) $\Pr\left[\mu - k\mu \le Z\,X_{\text{partial}} + (1 - Z)\mu \le \mu + k\mu\right]$

36. For the model $Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \varepsilon_i$, you are given:

(i) $N = 15$

(ii)

$$(X'X)^{-1} = \begin{bmatrix} 13.66 & -0.33 & 2.05 & -6.31 \\ -0.33 & 0.03 & 0.11 & 0.00 \\ 2.05 & 0.11 & 2.14 & -2.52 \\ -6.31 & 0.00 & -2.52 & 4.32 \end{bmatrix}$$

(iii) ESS = 282.82

Calculate the standard error of $\hat{\beta}_3 - \hat{\beta}_2$.

(A) 6.4

(B) 6.8

(C) 7.1

(D) 7.5

(E) 7.8

37. You are given:

Claim Size (X)	Number of Claims
$(0, 25]$	25
$(25, 50]$	28
$(50, 100]$	15
$(100, 200]$	6

Assume a uniform distribution of claim sizes within each interval.

Estimate $\mathrm{E}(X^2) - \mathrm{E}\left[(X \wedge 150)^2\right]$.

(A) Less than 200

(B) At least 200, but less than 300

(C) At least 300, but less than 400

(D) At least 400, but less than 500

(E) At least 500

38. Which of the following statements about moving average models is false?

(A) Both unweighted and exponentially weighted moving average (EWMA) models can be used to forecast future values of a time series.

(B) Forecasts using unweighted moving average models are determined by applying equal weights to a specified number of past observations of the time series.

(C) Forecasts using EWMA models may not be true averages because the weights applied to the past observations do not necessarily sum to one.

(D) Forecasts using both unweighted and EWMA models are adaptive because they automatically adjust themselves to the most recently available data.

(E) Using an EWMA model, the two-period forecast is the same as the one-period forecast.

39. You are given:

(i) Each risk has at most one claim each year.

(ii)

Type of Risk	Prior Probability	Annual Claim Probability
I	0.7	0.1
II	0.2	0.2
III	0.1	0.4

One randomly chosen risk has three claims during Years 1-6.

Determine the posterior probability of a claim for this risk in Year 7.

(A) 0.22

(B) 0.28

(C) 0.33

(D) 0.40

(E) 0.46

40. You are given the following about 100 insurance policies in a study of time to policy surrender:

(i) The study was designed in such a way that for every policy that was surrendered, a new policy was added, meaning that the risk set, r_j, is always equal to 100.

(ii) Policies are surrendered only at the end of a policy year.

(iii) The number of policies surrendered at the end of each policy year was observed to be:

1 at the end of the 1^{st} policy year
2 at the end of the 2^{nd} policy year
3 at the end of the 3^{rd} policy year
⋮
n at the end of the n^{th} policy year

(iv) The Nelson-Aalen empirical estimate of the cumulative distribution function at time n, $\hat{F}(n)$, is 0.542.

What is the value of n?

(A) 8

(B) 9

(C) 10

(D) 11

(E) 12

****END OF EXAMINATION****

Course 4, Fall 2003

FINAL ANSWER KEY

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