****BEGINNING OF EXAMINATION****

1. You are given:

(i) A random sample of five observations from a population is:

0.2 0.7 0.9 1.1 1.3

(ii) You use the Kolmogorov-Smirnov test for testing the null hypothesis, H_0, that the probability density function for the population is:

$$f(x) = \frac{4}{(1+x)^5}, \quad x > 0$$

(iii) Critical values for the Kolmogorov-Smirnov test are:

Level of Significance	0.10	0.05	0.025	0.01
Critical Value	$\frac{1.22}{\sqrt{n}}$	$\frac{1.36}{\sqrt{n}}$	$\frac{1.48}{\sqrt{n}}$	$\frac{1.63}{\sqrt{n}}$

Determine the result of the test.

(A) Do not reject H_0 at the 0.10 significance level.

(B) Reject H_0 at the 0.10 significance level, but not at the 0.05 significance level.

(C) Reject H_0 at the 0.05 significance level, but not at the 0.025 significance level.

(D) Reject H_0 at the 0.025 significance level, but not at the 0.01 significance level.

(E) Reject H_0 at the 0.01 significance level.

2. You are given:

(i) The number of claims follows a negative binomial distribution with parameters r and $\beta = 3$.

(ii) Claim severity has the following distribution:

Claim Size	Probability
1	0.4
10	0.4
100	0.2

(iii) The number of claims is independent of the severity of claims.

Determine the expected number of claims needed for aggregate losses to be within 10% of expected aggregate losses with 95% probability.

(A) Less than 1200

(B) At least 1200, but less than 1600

(C) At least 1600, but less than 2000

(D) At least 2000, but less than 2400

(E) At least 2400

3. You are given:

(i) A mortality study covers n lives.

(ii) None were censored and no two deaths occurred at the same time.

(iii) $t_k =$ time of the k^{th} death

(iv) A Nelson-Aalen estimate of the cumulative hazard rate function is $\hat{H}(t_2) = \dfrac{39}{380}$.

Determine the Kaplan-Meier product-limit estimate of the survival function at time t_9.

(A) Less than 0.56

(B) At least 0.56, but less than 0.58

(C) At least 0.58, but less than 0.60

(D) At least 0.60, but less than 0.62

(E) At least 0.62

4. Three observed values of the random variable X are:

1 1 4

You estimate the third central moment of X using the estimator:

$$g(X_1, X_2, X_3) = \frac{1}{3}\sum (X_i - \bar{X})^3$$

Determine the bootstrap estimate of the mean-squared error of g.

(A) Less than 3.0

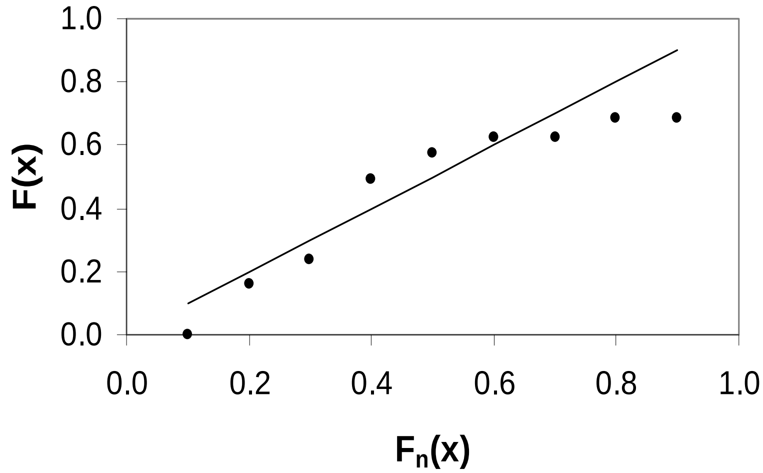
(B) At least 3.0, but less than 3.5

(C) At least 3.5, but less than 4.0

(D) At least 4.0, but less than 4.5

(E) At least 4.5

5. You are given the following *p*-*p* plot:



The plot is based on the sample:

1	2	3	15	30	50	51	99	100

Determine the fitted model underlying the *p*-*p* plot.

(A) $F(x) = 1 - x^{-0.25}$, $x \geq 1$

(B) $F(x) = x / (1 + x)$, $x \geq 0$

(C) Uniform on [1, 100]

(D) Exponential with mean 10

(E) Normal with mean 40 and standard deviation 40

6. You are given:

(i) Claims are conditionally independent and identically Poisson distributed with mean Θ.

(ii) The prior distribution function of Θ is:

$$F(\theta) = 1 - \left(\frac{1}{1+\theta}\right)^{2.6}, \quad \theta > 0$$

Five claims are observed.

Determine the Bühlmann credibility factor.

(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

7. Loss data for 925 policies with deductibles of 300 and 500 and policy limits of 5,000 and 10,000 were collected. The results are given below:

	Deductible		
Range	300	500	Total
(300 , 500]	50	–	50
(500 , 1,000]	50	75	125
(1,000 , 5,000)	150	150	300
(5,000 , 10,000)	100	200	300
At 5,000	40	80	120
At 10,000	10	20	30
Total	400	525	925

Using the Kaplan-Meier approximation for large data sets, with $\alpha = 1$ and $\beta = 0$, estimate $F(5000)$.

(A) 0.25

(B) 0.32

(C) 0.40

(D) 0.51

(E) 0.55

8. You are given the following knots:

(0,0) (2,0)

and derivative values:

$$f'(0) = -2$$
$$f'(2) = 2\,.$$

Determine the value of the squared norm measure of curvature for the cubic spline that satisfies these conditions.

(A) 16/15

(B) 8/3

(C) 4

(D) 8

(E) 16

9-10. *Use the following information for questions* 9 *and* 10.

The time to an accident follows an exponential distribution. A random sample of size two has a mean time of 6.

Let Y denote the mean of a new sample of size two.

9. Determine the maximum likelihood estimate of $\Pr(Y > 10)$.

(A) 0.04

(B) 0.07

(C) 0.11

(D) 0.15

(E) 0.19

9-10. *(Repeated for convenience) Use the following information for questions* 9 *and* 10.

The time to an accident follows an exponential distribution. A random sample of size two has a mean time of 6.

Let Y denote the mean of a new sample of size two.

10. Use the delta method to approximate the variance of the maximum likelihood estimator of $F_Y(10)$.

(A) 0.08

(B) 0.12

(C) 0.16

(D) 0.19

(E) 0.22

11. You are given:

(i) The number of claims in a year for a selected risk follows a Poisson distribution with mean λ.

(ii) The severity of claims for the selected risk follows an exponential distribution with mean θ.

(iii) The number of claims is independent of the severity of claims.

(iv) The prior distribution of λ is exponential with mean 1.

(v) The prior distribution of θ is Poisson with mean 1.

(vi) A priori, λ and θ are independent.

Using Bühlmann's credibility for aggregate losses, determine *k*.

(A) 1

(B) 4/3

(C) 2

(D) 3

(E) 4

12. A company insures 100 people age 65. The annual probability of death for each person is 0.03. The deaths are independent.

Use the inversion method to simulate the number of deaths in a year. Do this three times using:

$$u_1 = 0.20$$
$$u_2 = 0.03$$
$$u_3 = 0.09$$

Calculate the average of the simulated values.

(A) $\frac{1}{3}$

(B) 1

(C) $\frac{5}{3}$

(D) $\frac{7}{3}$

(E) 3

13. You are given claim count data for which the sample mean is roughly equal to the sample variance. Thus you would like to use a claim count model that has its mean equal to its variance. An obvious choice is the Poisson distribution.

Determine which of the following models may also be appropriate.

(A) A mixture of two binomial distributions with different means

(B) A mixture of two Poisson distributions with different means

(C) A mixture of two negative binomial distributions with different means

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

(E) All of (A), (B) and (C)

14. You are given:

(i) Annual claim frequencies follow a Poisson distribution with mean λ.

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

$$\pi(\lambda) = (0.4)\frac{1}{6}e^{-\lambda/6} + (0.6)\frac{1}{12}e^{-\lambda/12}, \qquad \lambda > 0$$

Ten claims are observed for an insured in Year 1.

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

(A) 9.6

(B) 9.7

(C) 9.8

(D) 9.9

(E) 10.0

15. Twelve policyholders were monitored from the starting date of the policy to the time of first claim. The observed data are as follows:

Time of First Claim	1	2	3	4	5	6	7
Number of Claims	2	1	2	2	1	2	2

Using the Nelson-Aalen estimator, calculate the 95% linear confidence interval for the cumulative hazard rate function $H(4.5)$.

(A) (0.189, 1.361)

(B) (0.206, 1.545)

(C) (0.248, 1.402)

(D) (0.283, 1.266)

(E) (0.314, 1.437)

16. For the random variable X, you are given:

(i) $\mathrm{E}[X] = \theta, \qquad \theta > 0$

(ii) $\mathrm{Var}(X) = \dfrac{\theta^2}{25}$

(iii) $\hat{\theta} = \dfrac{k}{k+1} X, \qquad k > 0$

(iv) $\mathrm{MSE}_{\hat{\theta}}(\theta) = 2\left[\mathrm{bias}_{\hat{\theta}}(\theta)\right]^2$

Determine k.

(A) 0.2

(B) 0.5

(C) 2

(D) 5

(E) 25

17. You are given:

(i) The annual number of claims on a given policy has a geometric distribution with parameter β.

(ii) The prior distribution of β has the Pareto density function

$$\pi(\beta) = \frac{\alpha}{(\beta+1)^{(\alpha+1)}}, \qquad 0 < \beta < \infty \ ,$$

where α is a known constant greater than 2.

A randomly selected policy had x claims in Year 1.

Determine the Bühlmann credibility estimate of the number of claims for the selected policy in Year 2.

(A) $\dfrac{1}{\alpha-1}$

(B) $\dfrac{(\alpha-1)x}{\alpha} + \dfrac{1}{\alpha(\alpha-1)}$

(C) x

(D) $\dfrac{x+1}{\alpha}$

(E) $\dfrac{x+1}{\alpha-1}$

18. You are given:

(i) The Cox proportional hazards model is used for the following data:

Gender (Z_1)	Age (Z_2)	Time to First Accident (X)
0	20	3
0	30	> 6
1	30	7
1	40	> 8

(ii) The baseline hazard rate function is a constant equal to $1/\theta$.

(iii) The regression coefficients for gender and age are β_1 and β_2, respectively.

Determine the loglikelihood function when $\theta = 18$, $\beta_1 = 0.1$, and $\beta_2 = 0.01$.

(A) Less than –10

(B) At least –10, but less than –7.5

(C) At least –7.5, but less than –5

(D) At least –5, but less than –2.5

(E) At least –2.5

19. Which of the following statements is true?

(A) For a null hypothesis that the population follows a particular distribution, using sample data to estimate the parameters of the distribution tends to decrease the probability of a Type II error.

(B) The Kolmogorov-Smirnov test can be used on individual or grouped data.

(C) The Anderson-Darling test tends to place more emphasis on a good fit in the middle rather than in the tails of the distribution.

(D) For a given number of cells, the critical value for the chi-square goodness-of-fit test becomes larger with increased sample size.

(E) None of (A), (B), (C) or (D) is true.

20. For a particular policy, the conditional probability of the annual number of claims given $\Theta = \theta$, and the probability distribution of Θ are as follows:

Number of claims	0	1	2
Probability	2θ	θ	$1-3\theta$

θ	0.05	0.30
Probability	0.80	0.20

Two claims are observed in Year 1.

Calculate the Bühlmann credibility estimate of the number of claims in Year 2.

(A) Less than 1.68

(B) At least 1.68, but less than 1.70

(C) At least 1.70, but less than 1.72

(D) At least 1.72, but less than 1.74

(E) At least 1.74

21. You are given:

(i) The annual number of claims for a policyholder follows a Poisson distribution with mean Λ.

(ii) The prior distribution of Λ is gamma with probability density function:

$$f(\lambda) = \frac{(2\lambda)^5 e^{-2\lambda}}{24\lambda}, \qquad \lambda > 0$$

An insured is selected at random and observed to have $x_1 = 5$ claims during Year 1 and $x_2 = 3$ claims during Year 2.

Determine $\mathrm{E}\left(\Lambda \middle| x_1 = 5,\ x_2 = 3\right)$.

(A) 3.00

(B) 3.25

(C) 3.50

(D) 3.75

(E) 4.00

22. You are given the kernel:

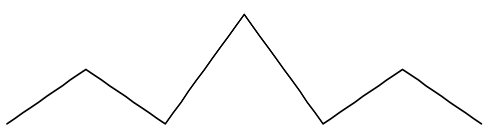
$$k_y(x) = \begin{cases} \dfrac{2}{\pi}\sqrt{1-(x-y)^2}\ , & y-1 \le x \le y+1 \\ \\ 0, & \text{otherwise} \end{cases}$$

You are also given the following random sample:

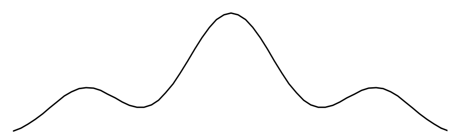
1 3 3 5

Determine which of the following graphs shows the shape of the kernel density estimator.

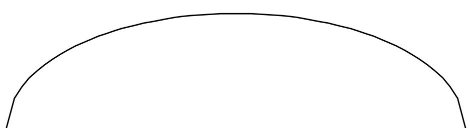
(A)



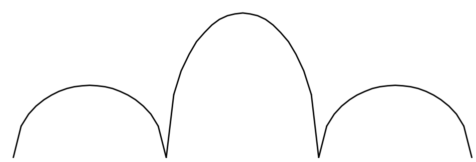
(B)



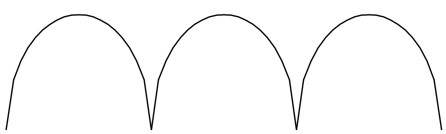
(C)



(D)



(E)



23. You are fitting a curve to a set of $n+1$ data points.

Which of the following statements is true?

(A) The collocation polynomial is recommended for extrapolation.

(B) Collocation polynomials pass through all data points, whereas cubic splines do not.

(C) The curvature-adjusted cubic spline requires that $m_0 = m_n = 0$ where $m_j = f''(x_j)$.

(D) For a cubic runout spline, the second derivative is a linear function throughout the intervals $[x_0, x_2]$ and $[x_{n-2}, x_n]$.

(E) If $f(x)$ is the natural cubic spline passing through the $n+1$ data points, and $h(x)$ is any function with continuous first and second derivatives that passes through the same points, then $\int_{x_0}^{x_n} [f''(x)]^2 dx > \int_{x_0}^{x_n} [h''(x)]^2 dx$.

24. The following claim data were generated from a Pareto distribution:

130   20   350   218   1822

Using the method of moments to estimate the parameters of a Pareto distribution, calculate the limited expected value at 500.

(A) Less than 250

(B) At least 250, but less than 280

(C) At least 280, but less than 310

(D) At least 310, but less than 340

(E) At least 340

25. You are given:

	Group	Year 1	Year 2	Year 3	Total
Total Claims	1		10,000	15,000	25,000
Number in Group			50	60	110
Average			200	250	227.27
Total Claims	2	16,000	18,000		34,000
Number in Group		100	90		190
Average		160	200		178.95
Total Claims					59,000
Number in Group					300
Average					196.67

You are also given $\hat{a}$ = 651.03.

Use the nonparametric empirical Bayes method to estimate the credibility factor for Group 1.

(A) 0.48

(B) 0.50

(C) 0.52

(D) 0.54

(E) 0.56

26. You are given the following information regarding claim sizes for 100 claims:

Claim Size	Number of Claims
0 - 1,000	16
1,000 - 3,000	22
3,000 - 5,000	25
5,000 - 10,000	18
10,000 - 25,000	10
25,000 - 50,000	5
50,000 - 100,000	3
over 100,000	1

Use the ogive to estimate the probability that a randomly chosen claim is between 2,000 and 6,000.

(A) 0.36

(B) 0.40

(C) 0.45

(D) 0.47

(E) 0.50

27. You are given the following 20 bodily injury losses (before the deductible is applied):

Loss	Number of Losses	Deductible	Policy Limit
750	3	200	∞
200	3	0	10,000
300	4	0	20,000
>10,000	6	0	10,000
400	4	300	∞

Past experience indicates that these losses follow a Pareto distribution with parameters α and $\theta = 10{,}000$.

Determine the maximum likelihood estimate of α.

(A) Less than 2.0

(B) At least 2.0, but less than 3.0

(C) At least 3.0, but less than 4.0

(D) At least 4.0, but less than 5.0

(E) At least 5.0

28. You are given:

(i) During a 2-year period, 100 policies had the following claims experience:

Total Claims in Years 1 and 2	Number of Policies
0	50
1	30
2	15
3	4
4	1

(ii) The number of claims per year follows a Poisson distribution.

(iii) Each policyholder was insured for the entire 2-year period.

A randomly selected policyholder had one claim over the 2-year period.

Using semiparametric empirical Bayes estimation, determine the Bühlmann estimate for the number of claims in Year 3 for the same policyholder.

(A) 0.380

(B) 0.387

(C) 0.393

(D) 0.403

(E) 0.443

29. For a study of losses on two classes of policies, you are given:

(i) The Cox proportional hazards model was used with baseline hazard rate function:

$$h_0(x) = \frac{2x}{\theta}, \quad 0 < x < \infty$$

(ii) A single covariate Z was used with $Z = 0$ for policies in Class A and $Z = 1$ for policies in Class B.

(iii) Two policies in Class A had losses of 1 and 3 while two policies in Class B had losses of 2 and 4.

Determine the maximum likelihood estimate of the coefficient β.

(A) – 0.7

(B) 0.5

(C) 0.7

(D) 1.0

(E) 1.6

30. A natural cubic spline is fit to $h(x) = x^5$ at the knots $x_0 = -2,\ x_1 = 0$ and $x_2 = 2$.

Determine $f''(-0.5)$.

(A) –1.0

(B) –0.5

(C) 0.0

(D) 0.5

(E) 1.0

31. Personal auto property damage claims in a certain region are known to follow the Weibull distribution:

$$F(x) = 1 - e^{-\left(\frac{x}{\theta}\right)^{0.2}}, \qquad x > 0$$

A sample of four claims is:

130 240 300 540

The values of two additional claims are known to exceed 1000.

Determine the maximum likelihood estimate of θ.

(A) Less than 300

(B) At least 300, but less than 1200

(C) At least 1200, but less than 2100

(D) At least 2100, but less than 3000

(E) At least 3000

32. For five types of risks, you are given:

(i) The expected number of claims in a year for these risks ranges from 1.0 to 4.0.

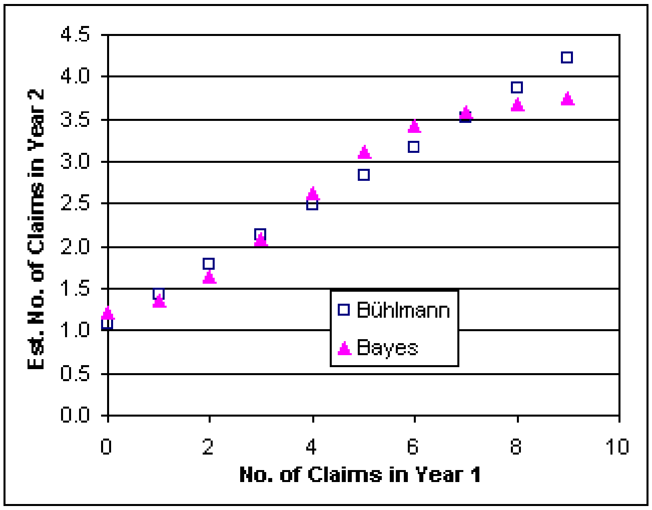
(ii) The number of claims follows a Poisson distribution for each risk.

During Year 1, n claims are observed for a randomly selected risk.

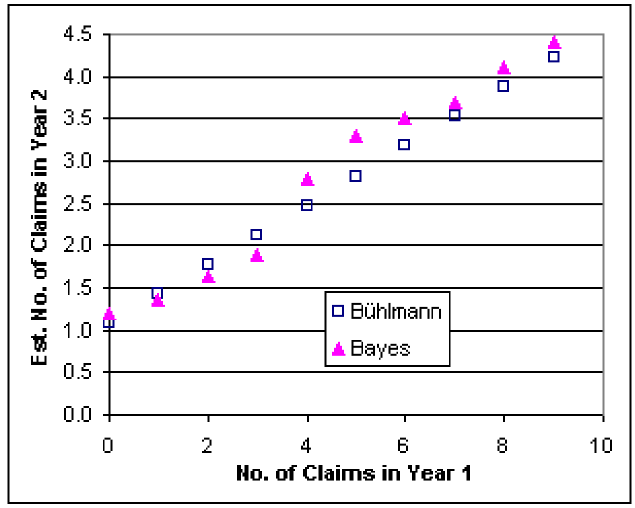
For the same risk, both Bayes and Bühlmann credibility estimates of the number of claims in Year 2 are calculated for $n = 0,1,2, \ldots ,9$.

Which graph represents these estimates?

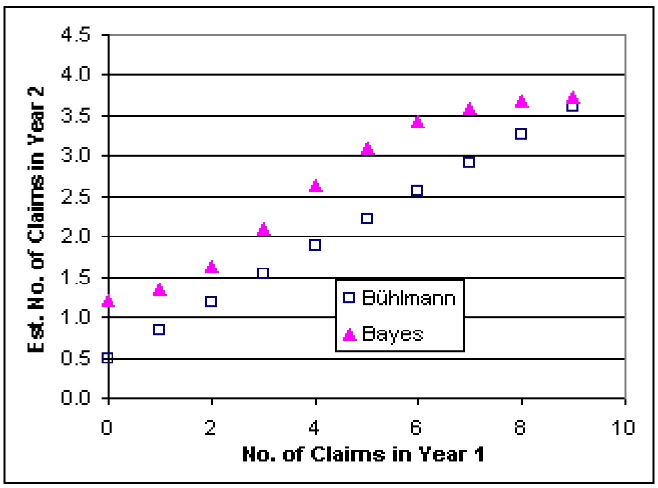
(A)


Est. No. of Claims in Year 2
No. of Claims in Year 1
Bühlmann
Bayes

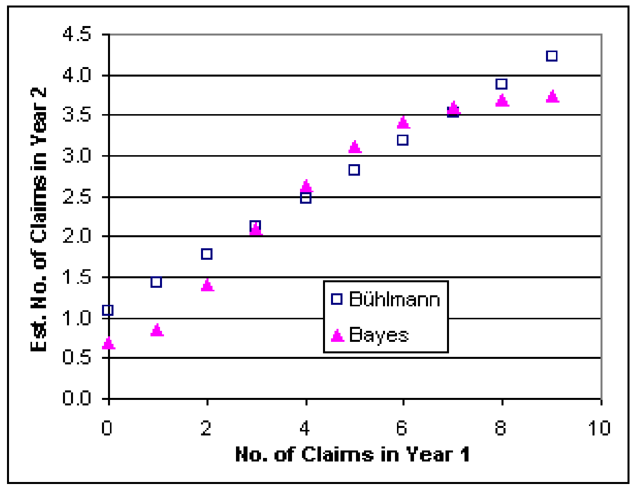
(B)


Est. No. of Claims in Year 2
No. of Claims in Year 1
Bühlmann
Bayes

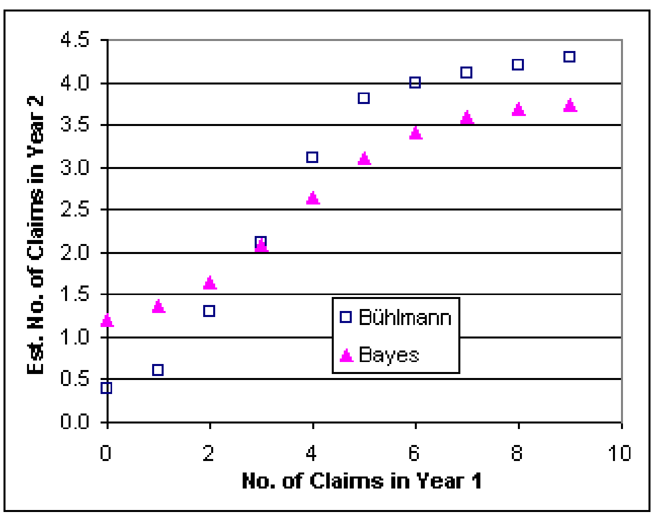
(C)


Est. No. of Claims in Year 2
No. of Claims in Year 1
Bühlmann
Bayes

(D)


Est. No. of Claims in Year 2
No. of Claims in Year 1
Bühlmann
Bayes

(E)


Est. No. of Claims in Year 2
No. of Claims in Year 1
Bühlmann
Bayes

33. You test the hypothesis that a given set of data comes from a known distribution with distribution function $F(x)$. The following data were collected:

Interval	$F(x_i)$	Number of Observations
$x < 2$	0.035	5
$2 \le x < 5$	0.130	42
$5 \le x < 7$	0.630	137
$7 \le x < 8$	0.830	66
$8 \le x$	1.000	50
Total		300

where x_i is the upper endpoint of each interval.

You test the hypothesis using the chi-square goodness-of-fit test.

Determine the result of the test.

(A) The hypothesis is not rejected at the 0.10 significance level.

(B) The hypothesis is rejected at the 0.10 significance level, but is not rejected at the 0.05 significance level.

(C) The hypothesis is rejected at the 0.05 significance level, but is not rejected at the 0.025 significance level.

(D) The hypothesis is rejected at the 0.025 significance level, but is not rejected at the 0.01 significance level.

(E) The hypothesis is rejected at the 0.01 significance level.

34. Unlimited claim severities for a warranty product follow the lognormal distribution with parameters $\mu = 5.6$ and $\sigma = 0.75$.

You use simulation to generate severities.

The following are six uniform (0, 1) random numbers:

0.6179 0.4602 0.9452 0.0808 0.7881 0.4207

Using these numbers and the inversion method, calculate the average payment per claim for a contract with a policy limit of 400.

(A) Less than 300

(B) At least 300, but less than 320

(C) At least 320, but less than 340

(D) At least 340, but less than 360

(E) At least 360

35. You are given:

(i) The annual number of claims on a given policy has the geometric distribution with parameter β.

(ii) One-third of the policies have $\beta = 2$, and the remaining two-thirds have $\beta = 5$.

A randomly selected policy had two claims in Year 1.

Calculate the Bayesian expected number of claims for the selected policy in Year 2.

(A) 3.4

(B) 3.6

(C) 3.8

(D) 4.0

(E) 4.2

****END OF EXAMINATION****

Exam C, Spring 2005

FINAL ANSWER KEY

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