COURSE 4

SAMPLE

EXAMINATION
Although a multiple choice format is not provided for most of the questions on this sample examination, the initial Course 4 Examination will consist entirely of multiple choice type questions. The solutions to the questions on this sample examination are at the end.

Appendices A and B of *Loss Models: From Data to Decisions*, and normal, chi-square, t, and F tables will be provided with the examination.
Question #1

You are given the following information about quarterly sales:

<table>
<thead>
<tr>
<th>Quarter</th>
<th>Sales</th>
<th>Four-quarter Moving Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15</td>
<td>24.00</td>
</tr>
<tr>
<td>2</td>
<td>15</td>
<td>25.00</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>26.25</td>
</tr>
<tr>
<td>4</td>
<td>50</td>
<td>25.50</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
<td>26.00</td>
</tr>
<tr>
<td>6</td>
<td>12</td>
<td>26.25</td>
</tr>
<tr>
<td>7</td>
<td>22</td>
<td>25.00</td>
</tr>
<tr>
<td>8</td>
<td>51</td>
<td>25.25</td>
</tr>
<tr>
<td>9</td>
<td>15</td>
<td>26.75</td>
</tr>
<tr>
<td>10</td>
<td>13</td>
<td>27.50</td>
</tr>
<tr>
<td>11</td>
<td>28</td>
<td>28.25</td>
</tr>
<tr>
<td>12</td>
<td>54</td>
<td>28.75</td>
</tr>
</tbody>
</table>

Determine the seasonally adjusted sales for Quarter 1.
Question #2

The number of employees leaving a company for all reasons is tallied by the number of months since hire. The following data was collected for a group of 50 employees hired one year ago:

<table>
<thead>
<tr>
<th>Number of Months Since Hire</th>
<th>Number Leaving the Company</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
</tr>
</tbody>
</table>

Determine the Nelson-Aalen estimate of the cumulative hazard at the sixth month since hire.

Note: Assume that employees always leave the company after a whole number of months.
Question #3

A fleet of cars has had the following experience for the last three years:

<table>
<thead>
<tr>
<th>Earned Car Years</th>
<th>Number of Claims</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>70</td>
</tr>
<tr>
<td>750</td>
<td>60</td>
</tr>
<tr>
<td>1,000</td>
<td>100</td>
</tr>
</tbody>
</table>

The Poisson distribution is used to model this process.

Determine the maximum likelihood estimate of the Poisson parameter for a single car year.
Question #4

An individual automobile insured has a claim count distribution per policy period that follows a Poisson distribution with parameter $\lambda$. For the overall population of automobile insureds, the parameter $\lambda$ follows a distribution with density function

$$f(\lambda) = 5\exp(-5\lambda), \; \lambda > 0.$$ 

One insured is selected at random from the population and is observed to have a total of one claim during two policy periods.

Determine the expected number of claims that this same insured will have during the third policy period.
Question #5

For the multiple regression model

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \epsilon_i,$$

you are given:

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Partial Correlation Coefficient</th>
<th>Standardized Coefficient</th>
<th>Elasticity</th>
</tr>
</thead>
<tbody>
<tr>
<td>X_2</td>
<td>0.64</td>
<td>0.50</td>
<td>0.20</td>
</tr>
<tr>
<td>X_3</td>
<td>-0.04</td>
<td>-0.01</td>
<td>-0.01</td>
</tr>
<tr>
<td>X_4</td>
<td>0.70</td>
<td>0.40</td>
<td>0.60</td>
</tr>
</tbody>
</table>

Which of the following is implied by this model?

(A) 16% of the variance of Y not accounted for by X_2 and X_3 is accounted for by X_4.

(B) An increase of 1 standard deviation in X_2 will lead to an increase of 0.64 standard deviations in Y.

(C) An increase of 1% in X_2 will lead to an increase of 0.20% in Y.

(D) An increase of 1 unit in X_3 will lead to a decrease of 0.04 units in Y.

(E) X_4 is a more important determinant of Y than X_2 is.
Question #6

The confidence interval for $S(t_M)$ within a linear 95% equal probability confidence band over the range $[t_L, t_U]$, where $t_L < t_M < t_U$, is $(0.360, 0.640)$.

Determine the confidence interval for $S(t_M)$ within a log-transformed 95% equal probability confidence band over the range $[t_L, t_U]$. 
Summary statistics for a sample of 100 losses are:

<table>
<thead>
<tr>
<th>Interval</th>
<th>Number of Losses</th>
<th>Sum</th>
<th>Sum of Squares</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0, 2,000]</td>
<td>39</td>
<td>38,065</td>
<td>52,170,078</td>
</tr>
<tr>
<td>(2,000, 4,000]</td>
<td>22</td>
<td>63,816</td>
<td>194,241,387</td>
</tr>
<tr>
<td>(4,000, 8,000]</td>
<td>17</td>
<td>96,447</td>
<td>572,753,313</td>
</tr>
<tr>
<td>(8,000, 15,000]</td>
<td>12</td>
<td>137,595</td>
<td>1,628,670,023</td>
</tr>
<tr>
<td>(15,000, ∞)</td>
<td>10</td>
<td>331,831</td>
<td>17,906,839,238</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100</strong></td>
<td><strong>667,754</strong></td>
<td><strong>20,354,674,039</strong></td>
</tr>
</tbody>
</table>

**Question #7**

Determine the empirical limited expected value $E(\hat{X} \wedge 15,000)$.

**Question #8**

A Pareto distribution is fit to this data using the method of moments.

Determine the parameter estimates.

**Question #9**

When a similar study was conducted on a different data set, the estimated parameters were $\hat{\alpha} = 2.5$ and $\hat{\theta} = 10,000$.

Determine the chi-square statistic and number of degrees of freedom for a test (with five groups) to assess the acceptability of fit of the data above to these parameters.

**Question #10**

You are given the following values, where $\alpha$ is the first parameter and $\theta$ is the second parameter:

\[
\alpha = 1.20, \quad \theta = 3,000, \quad S = \begin{bmatrix} 2.716 \\ 0.172 \end{bmatrix}, \quad P = \begin{bmatrix} 59.769 \\ 0.0141 \end{bmatrix}, \quad .0000378
\]

Determine the next iterated values of the maximum likelihood estimates using the method of scoring.
Question #11

You are given:

- A portfolio consists of 150 independent risks.
- 100 of the risks each have a policy with a $100,000 per claim policy limit, and 50 of the risks each have a policy with a $1,000,000 per claim policy limit.
- The risks have identical claim count distributions.
- Prior to censoring by policy limits, the claim size distribution for each risk is as follows:

<table>
<thead>
<tr>
<th>Claim Size</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10,000</td>
<td>1/2</td>
</tr>
<tr>
<td>$50,000</td>
<td>1/4</td>
</tr>
<tr>
<td>$100,000</td>
<td>1/5</td>
</tr>
<tr>
<td>$1,000,000</td>
<td>1/20</td>
</tr>
</tbody>
</table>

- A claims report is available that shows actual claim sizes incurred for each policy after censoring by policy limits, but does not identify the policy limit associated with each policy.

The claims report shows exactly three claims for a policy selected at random. Two of the claims are $100,000, but the amount of the third is illegible.

Determine the expected value of this illegible number.
Question #12

To predict insurance sales using eight independent variables you fit two regression models based on 27 observations.

The first model contains all eight independent variables. For this model you are given:

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Degrees of Freedom</th>
<th>Sum of Squares</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>8</td>
<td>115,175</td>
</tr>
<tr>
<td>Error</td>
<td>18</td>
<td>76,893</td>
</tr>
</tbody>
</table>

The second model contains only the first two independent variables. For this model you are given:

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Degrees of Freedom</th>
<th>Sum of Squares</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>2</td>
<td>65,597</td>
</tr>
<tr>
<td>Error</td>
<td>24</td>
<td>126,471</td>
</tr>
</tbody>
</table>

Determine the F ratio to use to test the hypothesis that the coefficients for the third through the eighth independent variables are all equal to zero.
Question #13

The mortality of a group of people with a certain genetic defect is being compared to the mortality of a reference population using the relative mortality model. The following data is from the first 24 months of the study:

<table>
<thead>
<tr>
<th>Month</th>
<th>Number of Deaths During Month</th>
<th>Estimated Cumulative Relative Excess Mortality at End of Month</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>33.72</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>53.86</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>65.59</td>
</tr>
<tr>
<td>14</td>
<td>1</td>
<td>75.32</td>
</tr>
<tr>
<td>22</td>
<td>2</td>
<td>85.72</td>
</tr>
<tr>
<td>24</td>
<td>1</td>
<td>90.83</td>
</tr>
</tbody>
</table>

A review of the data reveals that two deaths recorded as occurring during month 50 actually occurred on the same day as the death in month 14.

Determine the corrected estimate of the cumulative relative excess mortality at the end of month 14.
Question #14

You are given a tool that determines the expected prospective profitability of any insurance policy. You are asked to use this tool to estimate the average profitability of a large number of policies.

You want to be 95% certain that your estimate will not differ from the true value by more than 0.02 units. Your estimates of profitability, $X_i$, for the first 40 policies reviewed, together with the indicated statistics, are shown below.

<table>
<thead>
<tr>
<th>i</th>
<th>$X_i$</th>
<th>$\bar{X}_i$</th>
<th>$S_i^2$</th>
<th>$S_i$</th>
<th>$S_i / \sqrt{i}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.0795</td>
<td>1.0795</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1.0559</td>
<td>1.0677</td>
<td>0.00027908</td>
<td>0.0167</td>
<td>0.01181</td>
</tr>
<tr>
<td>3</td>
<td>1.1062</td>
<td>1.0806</td>
<td>0.00063266</td>
<td>0.0252</td>
<td>0.01452</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>20</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>21</td>
<td>1.1691</td>
<td>1.0830</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>22</td>
<td>1.0834</td>
<td>1.0830</td>
<td>0.001214</td>
<td>0.0307</td>
<td>0.01130</td>
</tr>
<tr>
<td>23</td>
<td>1.1272</td>
<td>1.0850</td>
<td>0.00276758</td>
<td>0.0526</td>
<td>0.01097</td>
</tr>
<tr>
<td>24</td>
<td>1.0722</td>
<td>1.0844</td>
<td>0.00265399</td>
<td>0.0515</td>
<td>0.01052</td>
</tr>
<tr>
<td>25</td>
<td>1.1373</td>
<td>1.0865</td>
<td>0.00265508</td>
<td>0.0515</td>
<td>0.01031</td>
</tr>
<tr>
<td>26</td>
<td>1.0428</td>
<td>1.0849</td>
<td>0.001214</td>
<td>0.0307</td>
<td>0.01130</td>
</tr>
<tr>
<td>27</td>
<td>1.0759</td>
<td>1.0845</td>
<td>0.00252463</td>
<td>0.0502</td>
<td>0.00967</td>
</tr>
<tr>
<td>28</td>
<td>1.1418</td>
<td>1.0866</td>
<td>0.00254843</td>
<td>0.0505</td>
<td>0.00954</td>
</tr>
<tr>
<td>29</td>
<td>1.1249</td>
<td>1.0879</td>
<td>0.00250808</td>
<td>0.0501</td>
<td>0.00930</td>
</tr>
<tr>
<td>30</td>
<td>1.2350</td>
<td>1.0928</td>
<td>0.00314274</td>
<td>0.0561</td>
<td>0.01024</td>
</tr>
<tr>
<td>31</td>
<td>1.0478</td>
<td>1.0914</td>
<td>0.00310330</td>
<td>0.0557</td>
<td>0.01001</td>
</tr>
<tr>
<td>32</td>
<td>1.0875</td>
<td>1.0912</td>
<td>0.00300365</td>
<td>0.0548</td>
<td>0.00969</td>
</tr>
<tr>
<td>33</td>
<td>1.1149</td>
<td>1.0919</td>
<td>0.00292673</td>
<td>0.0541</td>
<td>0.00942</td>
</tr>
<tr>
<td>34</td>
<td>1.1591</td>
<td>1.0939</td>
<td>0.00297075</td>
<td>0.0545</td>
<td>0.00935</td>
</tr>
<tr>
<td>35</td>
<td>1.0226</td>
<td>1.0919</td>
<td>0.00302872</td>
<td>0.0550</td>
<td>0.00930</td>
</tr>
<tr>
<td>36</td>
<td>0.9668</td>
<td>1.0884</td>
<td>0.00337693</td>
<td>0.0581</td>
<td>0.00969</td>
</tr>
<tr>
<td>37</td>
<td>1.1487</td>
<td>1.0900</td>
<td>0.00338141</td>
<td>0.0581</td>
<td>0.00956</td>
</tr>
<tr>
<td>38</td>
<td>1.1887</td>
<td>1.0926</td>
<td>0.00354638</td>
<td>0.0596</td>
<td>0.00966</td>
</tr>
<tr>
<td>39</td>
<td>1.1303</td>
<td>1.0936</td>
<td>0.00348950</td>
<td>0.0591</td>
<td>0.00946</td>
</tr>
<tr>
<td>40</td>
<td>1.0484</td>
<td>1.0925</td>
<td>0.00345104</td>
<td>0.0587</td>
<td>0.00929</td>
</tr>
</tbody>
</table>

Determine the minimum number of policies you could have reviewed to meet your objective, using the method described in Ross.
Question #15

You are given:

- The number of claims per exposure follows a Poisson distribution with mean 0.01.
- Claim sizes follow a lognormal distribution with parameters $\mu$ (unknown) and $\sigma = 1$.
- The number of claims per exposure and claim sizes are independent.
- The method of limited fluctuation credibility is used, and the full credibility standard has been selected so that total claim dollars will be within 10% of expected total claim dollars 95% of the time.

Determine the number of exposures required for full credibility.
Question #16

The following autoregressive process of order 2 model is used to forecast inventory levels:

\[(y_t - \mu) - \phi_1(y_{t-1} - \mu) - \phi_2(y_{t-2} - \mu) = \epsilon_t\]

You are given:

\[\hat{\mu} = 25\]
\[\hat{\phi}_1 = 0.5\]
\[\hat{\phi}_2 = 0.2\]
\[y_{49} = 27.0\]
\[y_{50} = 28.0\]

Determine the origin-50 forecast for lead time three.
A life insurance company collected data on a group of 200 of its life insurance policies. Half of these insureds had a homeowners policy with an affiliated company. The data from these policies is shown below.

<table>
<thead>
<tr>
<th>Time</th>
<th>Number of Deaths</th>
<th>Homeowners Policy (Z = 1)</th>
<th>No Homeowners Policy (Z = 0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4 or greater</td>
<td>96</td>
<td>92</td>
<td></td>
</tr>
</tbody>
</table>

A proportional hazards model is used to model the difference in hazard rates between life insurance policies with and without associated homeowners policies. No other covariates are used.

Determine the value of Breslow’s partial log likelihood when the parameter $\beta$ is $-0.5$. 
Question #18

A group consisting of ten independent lives has a health policy with an ordinary deductible of 250, coinsurance of 80%, and a limit of 1,000 (before application of the deductible and coinsurance). In the past year, the following individual payments were made to members of the group:

40
120
160
280
600 (loss in excess of maximum)
600 (loss in excess of maximum)

Determine the likelihood function for estimating parameters of the ground-up loss distribution using \( f(x) \) to represent the probability density function and \( F(x) \) to represent the cumulative distribution function.
You are given:

- A portfolio of independent risks is divided into two classes of equal size.
- All of the risks in Class 1 have identical claim count and claim size distributions as follows:

<table>
<thead>
<tr>
<th>Number of Claims</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1/2</td>
</tr>
<tr>
<td>2</td>
<td>1/2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Claim Size</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>2/3</td>
</tr>
<tr>
<td>100</td>
<td>1/3</td>
</tr>
</tbody>
</table>

- All of the risks in Class 2 have identical claim count and claim size distributions as follows:

<table>
<thead>
<tr>
<th>Number of Claims</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2/3</td>
</tr>
<tr>
<td>2</td>
<td>1/3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Claim Size</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>1/2</td>
</tr>
<tr>
<td>100</td>
<td>1/2</td>
</tr>
</tbody>
</table>

- The number of claims and claim size(s) for each risk are independent.
- A risk is selected at random from the portfolio, and a pure premium of 100 is observed for the first exposure period.

Question #19

Determine the Bayesian estimate of the expected number of claims for this same risk for the second exposure period.

Question #20

A pure premium of 150 is observed for this risk for the second exposure period.

Determine the Buhlmann credibility estimate of the expected pure premium for this same risk for the third exposure period.
Question #21

The graphs shown on the following page show a time series, its first and second differences, and the sample autocorrelation functions for the series and its first and second differences.

Based solely on this information, determine which of the following characterizations is most appropriate for the underlying stochastic process.

(A) The process is white noise.
(B) The process is a stationary process other than white noise.
(C) The process is a random walk.
(D) The process is a first-order homogeneous nonstationary process.
(E) The process is a second-order homogeneous nonstationary process.
Question #22

An insurance company wishes to estimate its four-year agent retention rate using data on all agents hired during the last six years. You are given:

- Using the Product-Limit estimator, the company estimates the proportion of agents remaining after 3.75 years of service as \( \hat{S}(3.75) = 0.25 \).
- One agent resigned between 3.75 and 4 years of service.
- Eleven agents have been employed longer than the agent who resigned between 3.75 and 4 years of service.
- Two agents have been employed for six years.

Determine the Product-Limit estimate of \( S(4) \).
Question #23

Forty observed losses have been recorded in thousands of dollars and are grouped as follows:

<table>
<thead>
<tr>
<th>Interval</th>
<th>Number of Losses</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1, 4/3)</td>
<td>16</td>
</tr>
<tr>
<td>[4/3, 2)</td>
<td>10</td>
</tr>
<tr>
<td>[2, 4)</td>
<td>10</td>
</tr>
<tr>
<td>[4, ∞)</td>
<td>4</td>
</tr>
</tbody>
</table>

The null hypothesis, $H_0$, is that the random variable $X$ underlying the observed losses, in thousands, has the density function

$$f(x) = \frac{1}{x^2}, \quad 1 < x < \infty.$$

Since exact values of the losses are not available, it is not possible to compute the exact value of the Kolmogorov-Smirnov statistic used to test $H_0$. However, it is possible to put bounds on the value of this statistic.

Based on the information above, determine the smallest possible value and the largest possible value of the Kolmogorov-Smirnov statistic used to test $H_0$. 
Question #24

Type A risks have each year’s losses uniformly distributed on the interval [0,1].
Type B risks have each year’s losses uniformly distributed on the interval [0,2].

A risk is selected at random with each type being equally likely. The first year’s losses equal L.

Let X be the Buhlmann credibility estimate of the second year’s losses.
Let Y be the Bayesian estimate of the second year’s losses.

Which of the following statements is true?

(A) If L<1, then X>Y.
(B) If L>1, then X<Y.
(C) If L=1/2, then X<Y.
(D) There are no values of L such that X=Y.
(E) There are exactly two values of L such that X=Y.
Question #25

A stationary ARMA (p, q) model is known to have zero partial autocorrelations at all lags greater than or equal to two and nonzero autocorrelation at lag one.

Determine (p, q).
Question #26

An insurance company collected the following data on the payment pattern for a group of 50 claims:

<table>
<thead>
<tr>
<th>Month</th>
<th>Number Unpaid</th>
<th>Number Paid at End of Month</th>
<th>Estimated Cumulative Hazard at End of Month</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50</td>
<td>5</td>
<td>0.1000</td>
</tr>
<tr>
<td>2</td>
<td>45</td>
<td>3</td>
<td>0.1667</td>
</tr>
<tr>
<td>3</td>
<td>42</td>
<td>2</td>
<td>0.2143</td>
</tr>
<tr>
<td>5</td>
<td>40</td>
<td>2</td>
<td>0.2643</td>
</tr>
<tr>
<td>7</td>
<td>38</td>
<td>2</td>
<td>0.2906</td>
</tr>
<tr>
<td>10</td>
<td>37</td>
<td>1</td>
<td>0.3176</td>
</tr>
<tr>
<td>12</td>
<td>36</td>
<td>2</td>
<td>0.3732</td>
</tr>
</tbody>
</table>

Determine the kernel-smoothed estimate of the hazard rate at the end of month six using a bandwidth of three months and the Epanechnikov kernel with

\[ K(x) = 0.75 \left( \frac{d}{2} - x^2 \right), \quad -1 \leq x \leq 1. \]
Question #27

A company wishes to analyze its usage of an office telephone line using a two-state continuous-time Markov chain. The process is in state 1 if the line is not busy and state 2 if the line is busy. The transition rates are $\mu_{12}$ and $\mu_{21}$.

You are given the following data:

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>process starts in state 1</td>
</tr>
<tr>
<td>3.6</td>
<td>process enters state 2</td>
</tr>
<tr>
<td>8.7</td>
<td>process enters state 1</td>
</tr>
<tr>
<td>9.5</td>
<td>process enters state 2</td>
</tr>
<tr>
<td>11.0</td>
<td>process enters state 1</td>
</tr>
<tr>
<td>15.6</td>
<td>process enters state 2</td>
</tr>
<tr>
<td>17.3</td>
<td>process enters state 1</td>
</tr>
<tr>
<td>19.2</td>
<td>process enters state 2</td>
</tr>
<tr>
<td>20.0</td>
<td>observation ceases with process still in state 2</td>
</tr>
</tbody>
</table>

Determine the maximum likelihood estimates of $\mu_{12}$ and $\mu_{21}$.
Question #28

Four urns contain balls marked either 1 or 3 in the following proportions:

<table>
<thead>
<tr>
<th>Urn</th>
<th>Marked 1</th>
<th>Marked 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$p_1$</td>
<td>$1-p_1$</td>
</tr>
<tr>
<td>2</td>
<td>$p_2$</td>
<td>$1-p_2$</td>
</tr>
<tr>
<td>3</td>
<td>$p_3$</td>
<td>$1-p_3$</td>
</tr>
<tr>
<td>4</td>
<td>$p_4$</td>
<td>$1-p_4$</td>
</tr>
</tbody>
</table>

An urn is selected at random (with each urn being equally likely) and balls are drawn from it in three separate rounds. In the first round, two balls are drawn with replacement. In the second round, one ball is drawn with replacement. In the third round, two balls are drawn with replacement.

After two rounds, the Buhlmann-Straub credibility estimate of the total of the values on the two balls to be drawn in the third round could range from 3.8 to 5.0 (depending on the results of the first two rounds).

Determine the value of Buhlmann-Straub’s k.
You wish to determine the nature of the relationship between sales (Y) and the number of radio advertisements broadcast (X). Data collected on four consecutive days is shown below.

<table>
<thead>
<tr>
<th>Day</th>
<th>Sales</th>
<th>Number of Radio Advertisements</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>30</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>40</td>
<td>3</td>
</tr>
</tbody>
</table>

Using the method of least squares, you determine the estimated regression line:

\[ \hat{Y} = -25 + 20X \]

**Question #29**

Determine the value of \( R^2 \) for this model.

**Question #30**

Determine the value of the Durbin-Watson statistic for this model.

**Question #31**

You perform an Empirical Bayes nonparametric credibility analysis by treating the first two days, on which two radio advertisements were broadcast, as one group, and the last two days, on which three radio advertisements were broadcast, as another group.

Determine the estimated credibility, \( \hat{Z} \), of the data from each group.
Question #32

A bank issued 50 loans at the beginning of 1996, 100 loans at the beginning of 1997, and 200 loans at the beginning of 1998. As of the end of 1998, 5 loans from each year had defaulted.

Determine the probability that a loan will default within two years of being issued using a modified Product-Limit estimator.
Question #33

You are given:

- The random variable $X$ has the density function

$$f(x) = \frac{1}{\lambda} \exp(-x / \lambda), \; 0 < x < \infty, \; \lambda > 0.$$ 

- $\lambda$ is estimated by the maximum likelihood estimator $\hat{\lambda}$ based on a large sample of data.

- The probability that $X$ is greater than $k$ is estimated by the estimator $\exp(-k / \hat{\lambda})$.

Determine the approximate variance of the estimator for the probability that $X$ is greater than $k$. 
Question #34

You are given:

- The number of claims for Risk 1 during a single exposure period follows a Bernoulli distribution with mean $p$.
- The prior distribution for $p$ is uniform on the interval $[0, 1]$.
- The number of claims for Risk 2 during a single exposure period follows a Poisson distribution with mean $\theta$.
- The prior distribution for $\theta$ has the density function
  \[
  f(\theta) = \beta \exp(-\beta \theta), \quad 0 < \theta < \infty, \quad \beta > 0.
  \]
- The loss experience of both risks is observed for an equal number of exposure periods.

Determine all values of $\beta$ for which the Buhlmann credibility of the loss experience of Risk 2 will be greater than the Buhlmann credibility of the loss experience of Risk 1.
Question #35

To determine the relationship of salary (Y) to years of experience (X_2) for both men (X_3 = 1) and women (X_3 = 0) you fit the model

\[ Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{2i}X_{3i} + \epsilon_i \]

to a set of observations from a sample of 11 employees. For this model you are given:

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Degrees of Freedom</th>
<th>Sum of Squares</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>3</td>
<td>330.0117</td>
</tr>
<tr>
<td>Error</td>
<td>7</td>
<td>12.8156</td>
</tr>
</tbody>
</table>

You also fit the model

\[ Y_i = \beta_1^* + \beta_2^* X_{2i} + \epsilon_i^* \]

to the observations. For this model you are given:

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Degrees of Freedom</th>
<th>Sum of Squares</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>1</td>
<td>315.0992</td>
</tr>
<tr>
<td>Error</td>
<td>9</td>
<td>27.7281</td>
</tr>
</tbody>
</table>

Determine the F ratio to use to test whether the linear relationship between salary and years of experience is identical for men and women.
Question #36

For a stationary first-order autoregressive process, $\phi = 0.6$ and $\sigma^2 = 1$.

Determine the width of the symmetric 95 percent confidence interval around a five-period ahead forecast.
Question #37

Twenty widgets are tested until they fail. The failure times are distributed as follows:

<table>
<thead>
<tr>
<th>Interval</th>
<th>Number Failing</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0, 1]</td>
<td>2</td>
</tr>
<tr>
<td>(1, 2]</td>
<td>3</td>
</tr>
<tr>
<td>(2, 3]</td>
<td>8</td>
</tr>
<tr>
<td>(3, 4]</td>
<td>6</td>
</tr>
<tr>
<td>(4, 5]</td>
<td>1</td>
</tr>
<tr>
<td>(5, ∞)</td>
<td>0</td>
</tr>
</tbody>
</table>

The exponential survival function \( S(t) = \exp(-\lambda t) \) is used to model this process.

Determine the maximum likelihood estimate of \( \lambda \).
Question #38

You are given a random sample of four values from a distribution F:

\[ 4, 5, 9, 14 \]

You estimate \( \theta(F) = E(X) \) using the estimator \( g(X_1, X_2, X_3, X_4) = X_1 \).

Determine the bootstrap approximation to the mean square error.
Question #39

The number of losses arising from m+4 individual insureds over a single period of observation is distributed as follows:

<table>
<thead>
<tr>
<th>Number of Losses</th>
<th>Number of Insureds</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>m</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>3 or more</td>
<td>0</td>
</tr>
</tbody>
</table>

The number of losses for each insured follows a Poisson distribution, but the mean of each such distribution may be different for individual insureds.

You estimate the variance of the hypothetical means using Empirical Bayes semiparametric estimation.

Determine all values of m for which the estimate of the variance of the hypothetical means will be greater than 0.
Question #40

For a linear regression model, which of the following conditions can produce estimators of the slope parameters that are biased and inconsistent?

(A) Heteroscedasticity

(B) Serial correlation

(C) An error in measurement of the dependent variable

(D) Omission of a relevant independent variable

(E) Inclusion of an irrelevant independent variable
Solutions
1. 23
2. 0.1256
3. 0.102
4. 2/7
5. C
6. (0.354, 0.630)
7. 4.859
8. \(\hat{\alpha} = 2.78, \hat{\theta} = 11,884\)
9. 0.848, 4 d.f.
10. \(\alpha^* = 1.716, \theta^* = 5,379\)
11. $53,409
12. 1.93
13. 94.78
14. 31
15. 104,426
16. 26.155
17. –62.70
18. \(\frac{f(300) \cdot f(400) \cdot f(450) \cdot f(600) \cdot [1 - F(1,000)]^2}{[1 - F(250)]^6}\)
19. 41/29
20. 100
21. D
22. 0.229
23. Smallest possible value = 0.15, Largest possible value = 0.40
24. E
25. (1, 0)
26. 0.017
27. \(\mu_{12} = 40/109, \mu_{21} = 30/91\)
28. 7
29. 0.80
30. 3
31. 7/8
32. 0.05
33. \(\frac{k^2}{\lambda^4} \exp(-2k/\lambda) \text{Var}(\hat{\lambda})\)
34. \(\beta < 2\)
35. 4.07
36. 4.885
37. 0.397
38. 31/2
39. $m \geq 7$
40. D