CAS Exam MAS-1
Practice Exam #1

These practice exams should be used during the month prior to your exam.

This practice exam contains **43 questions**, of equal value,
corresponding to about a **4 hour** exam.

Each problem is similar to a problem in my study guides, sold separately.
Solutions to problems are at the end of each practice exam.

prepared by
Howard C. Mahler, FCAS

Howard Mahler
hmahler@mac.com
www.howardmahler.com/Teaching

CAS Exam MAS-1, Practice Exam #1

1. You are given the following information:
- $x_1, \ldots, x_{20}$ is a random sample from a normal distribution with unknown parameters μ and σ.
- H_0: $\sigma = 7$
- H_1: $\sigma < 7$
- H_0 is rejected if $S^2 < k$, where S^2 is the unbiased sample variance.

Calculate the maximum value of k that results in a probability of Type I error of no more than 2.5%.
(A) Less than 20
(B) At least 20, but less than 25
(C) At least 25, but less than 30
(D) At least 30, but less than 35
(E) At least 35

2. A Generalized Linear Model uses a Normal Distribution.
One of the observations of the response variable is 805.
The corresponding fitted value is 740.
The fitted σ is 44.
Determine the corresponding Deviance residual.
A. 1.1 B. 1.3 C. 1.5 D. 1.7 E. 1.9

3. Two independent estimators, α and β, are available for estimating the parameter, τ, of a given distribution.
To test their performance, you have conducted 5000 simulated trials of each estimator, using $\tau = 3$, with the following results:

$$\sum_{i=1}^{5000} \alpha_i = 14{,}130, \quad \sum_{i=1}^{5000} \alpha_i^2 = 41{,}062, \quad \sum_{i=1}^{5000} \beta_i = 15{,}682, \quad \sum_{i=1}^{5000} \beta_i^2 = 49{,}975.$$

Consider the class of estimators of τ which are of the form: $w\alpha + (1-w)\beta$.
Determine the value of w that results in an estimator with the smallest mean squared error.
A. 40% B. 42% C. 44% D. 46% E. 48%

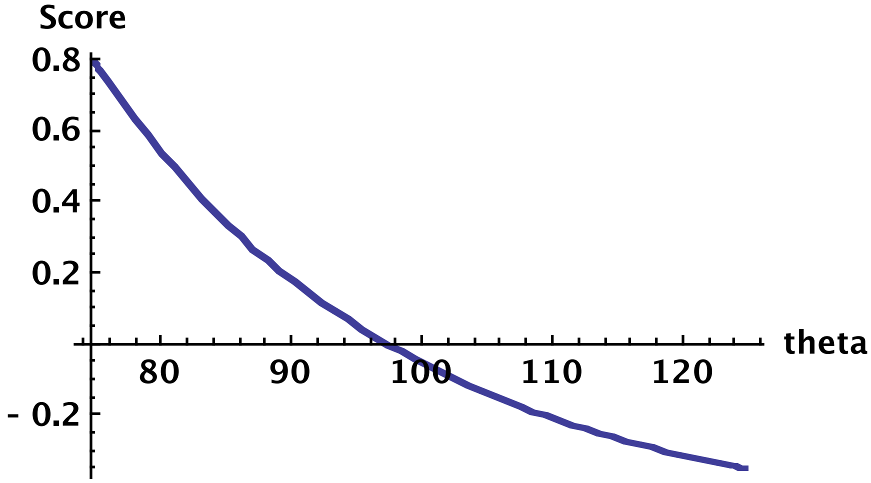
4. Use the following information:
- Using the Method of Maximum Likelihood, a Pareto Distribution has been fit to data.
- The fitted parameters are $\alpha = 3.0$ and $\theta = 1000$.
- The Inverse of the Information Matrix, with alpha first and theta second, is:

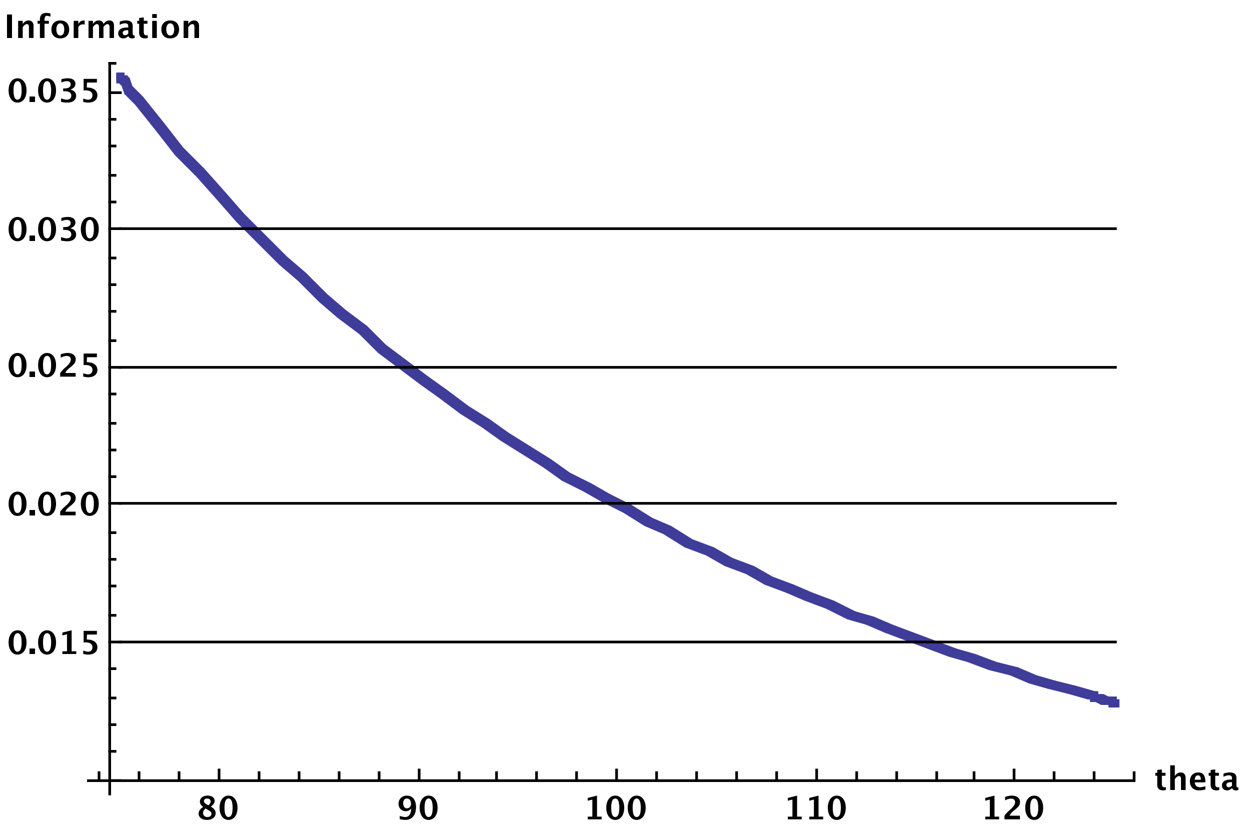
$$\begin{pmatrix} 0.029 & 12 \\ 12 & 5333 \end{pmatrix}$$

Estimate the standard deviation of the maximum likelihood estimate of the Survival Function at 5000, using the delta method.
A. less than 0.0005
B. at least 0.0005 but less than 0.0010
C. at least 0.0010 but less than 0.0015
D. at least 0.0015 but less than 0.0020
E. at least 0.0020

5. Each doctor makes medical errors via a Poisson Process with annual intensity λ.
λ varies across a group of doctors via a Gamma Distribution with $\alpha = 0.5$ and $\theta = 1.7$.
13% of medical errors result in the filing of a medical malpractice insurance claim.
Determine the probability that a doctor picked at random from this group will have exactly 4 medical malpractice insurance claims over the next 10 years.
A. 2.6% B. 2.8% C. 3.0% D. 3.2% E. 3.4%

6.
An actuary fits via maximum likelihood a distribution with one parameter θ to a sample of data. The actuary plots two quantities of interest across different values of θ.





Determine the upper end of a 95% confidence interval for θ.
A. Less than 115
B. At least 115, but less than 120
C. At least 120, but less than 125
D. At least 125, but less than 130
E. At least 130

7. You are given the following 14 losses (before the deductible is applied):

Loss	Number of Losses	Deductible	Maximum Covered Loss
500	5	0	10,000
2500	4	1000	25,000
≥ 10,000	3	0	10,000
≥ 25,000	2	1000	25,000

Past experience indicates that these losses follow a Pareto distribution with parameters α and θ = 10,000.

Determine the maximum likelihood estimate of α.

(A) Less than 1.4
(B) At least 1.4, but less than 1.5
(C) At least 1.5, but less than 1.6
(D) At least 1.6, but less than 1.7
(E) At least 1.7

8. You are given the following information:

- $X_1, ..., X_n$ is a random sample from a normal distribution with mean μ_x and standard deviation σ_x.
- $Y_1, ..., Y_n$ is a random sample from a normal distribution with mean μ_y and standard deviation σ_y.
- The two samples are independent.
- The unbiased sample variances S_x^2 and S_y^2 are 600 and 130, respectively.
- H_0: $\sigma_x^2 = \sigma_y^2$.
- H_1: $\sigma_x^2 \neq \sigma_y^2$.
- Use of the F test results in rejection of the null hypothesis at a significance level of 2%.

Calculate the minimum possible value of the sample size, n.

(A) 9 (B) 10 (C) 11 (D) 12 (E) 13

9. Use the following information about a Generalized Linear Model:
- The linear predictor, $\eta = \beta_0 + \beta_1 X_1 + \beta_2 X_2$.
- The errors follow a Gamma Distribution.
- The log link function is used.
- The fitted parameters are: $\beta_0 = 3.1$, $\beta_1 = 0.8$, and $\beta_2 = -0.5$
- The estimated covariance matrix is: $\begin{matrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{matrix} \begin{pmatrix} 0.0431 & -0.00264 & -0.00419 \\ -0.00264 & 0.00136 & 0.000278 \\ -0.00419 & 0.000278 & 0.00258 \end{pmatrix}$.

Determine the <u>upper</u> end of a 95% confidence interval for β_0.

A. 3.5 B. 3.6 C. 3.7 D. 3.8 E. 3.9

10. Two independent populations X and Y have density functions $f(x) = \lambda x^{\lambda-1}$ for $0 < x < 1$, and $g(y) = \mu y^{\mu-1}$ for $0 < y < 1$, respectively.
Let $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_m$ be random samples from X and Y.
What is the form of the critical region for the best (Neyman-Pearson) test of H_0: $\lambda = 2$, $\mu = 4$ against the alternative H_1: $\lambda = 3$, $\mu = 6$?

(A) $\sum_{i=1}^{n} \ln X_i + 2\sum_{i=1}^{m} \ln Y_i \le c$

(B) $\sum_{i=1}^{n} \ln X_i + 2\sum_{i=1}^{m} \ln Y_i \ge c$

(C) $2\sum_{i=1}^{n} \ln X_i - \sum_{i=1}^{m} \ln Y_i \le c$

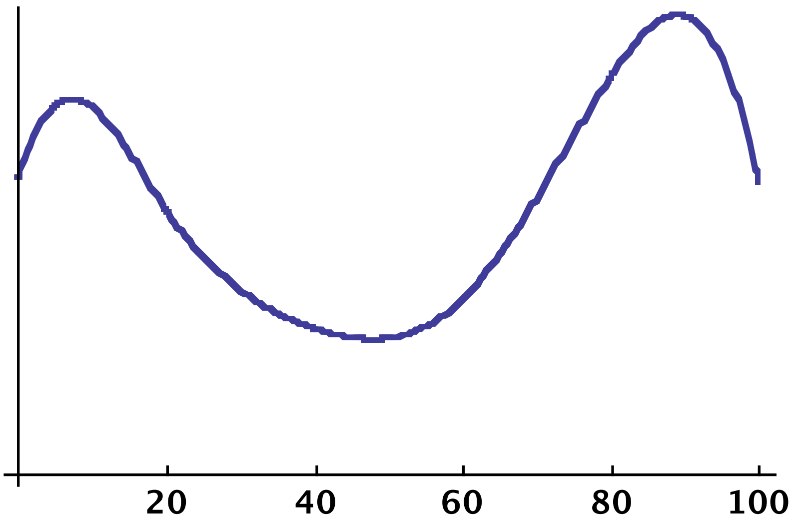
(D) $2\sum_{i=1}^{n} \ln X_i - \sum_{i=1}^{m} \ln Y_i \ge c$

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

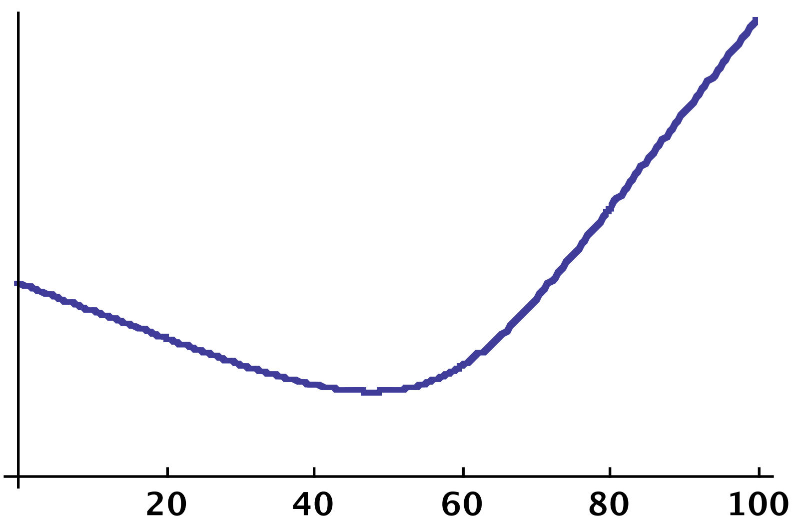
11. One has fit a regression model with 6 variables (5 independent variables plus the intercept), to 18 observations.
One is testing the hypothesis H_0: $\beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = 0$,
versus the alternative hypothesis that H_0 is false.
TSS is the total sum of squares.
RSS is the residual (error) sum of squares.
What is the critical region for a test at a 1% significance level?
A. TSS ≥ 3.1 RSS
B. TSS ≥ 3.2 RSS
C. TSS ≥ 3.3 RSS
D. TSS ≥ 3.4 RSS
E. TSS ≥ 3.5 RSS

12. Use the following information:
● A special 3-year annuity due that pays while at least one of them is alive is sold to lives (97) and (91).
● Their future life times are independent and each follow the illustrative life table.
● i = 8%
Determine the actuarial present value.
A. 2.50 B. 2.52 C. 2.54 D. 2.56 E. 2.58

13.
Which of the following regression splines with knots at 20, 40, 60 and 80 is a natural cubic spline?



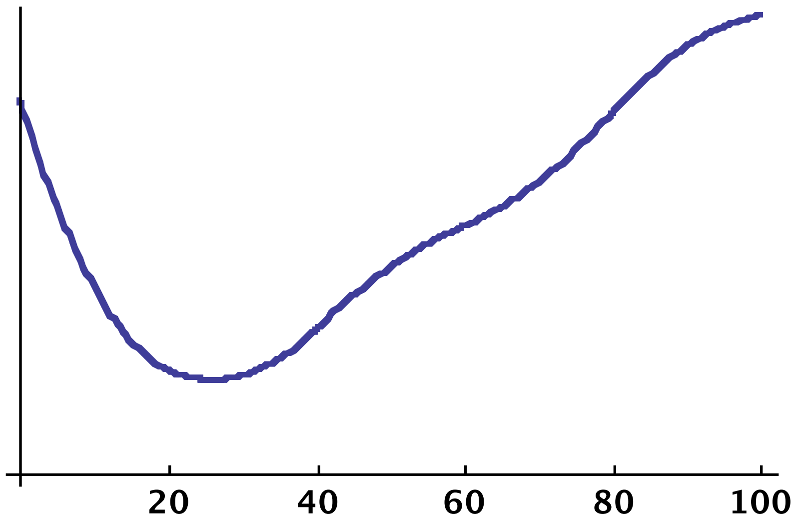
A.



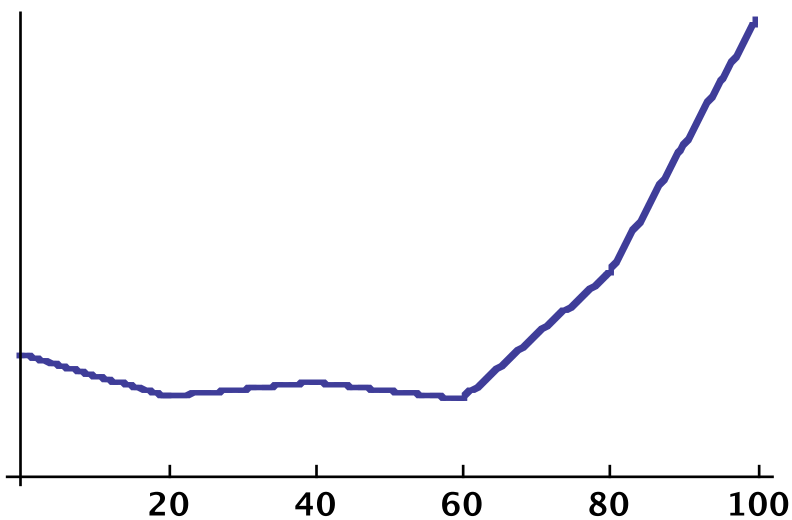
B.

Question continued on the next page.

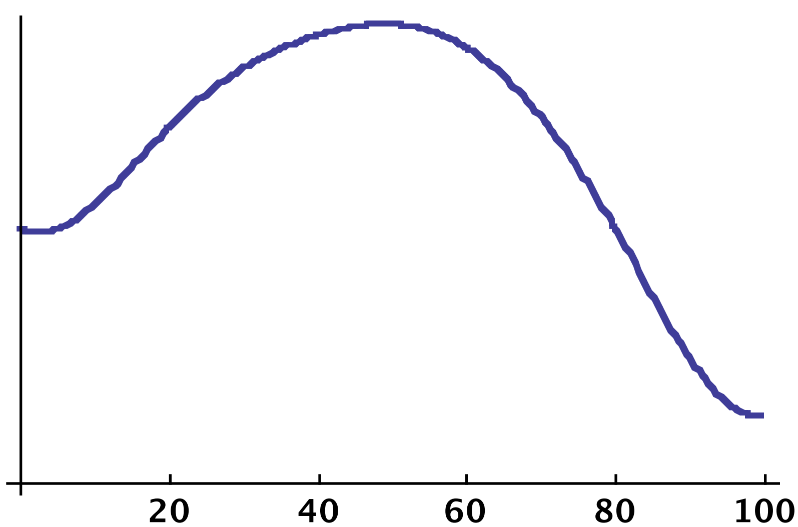
Question continued from the previous page.



C.



D.



E.

14. Ten claims have been observed: 65, 41, 51, 42, 26, 21, 43, 12, 104, 96.
A Gamma Distribution is fit to this data via the method of moments.
What is the third moment of the fitted Gamma Distribution?
A. less than 260,000
B. at least 260,000 but less than 270,000
C. at least 270,000 but less than 280,000
D. at least 280,000 but less than 290,000
E. at least 290,000

15. Take a random sample of size 13 from the uniform distribution on (0, 100).
Let Y be the fourth value in the sample from smallest to largest.
Determine the probability that $Y \geq 30$.
A. 34% B. 36% C. 38% D. 40% E. 42%

16. Which of the following is a minimal path set for a 4 out of 6 system?
A. {5} B. {1, 2} C. {2, 4, 6} D. {1, 3, 5, 6}
E. None of A, B, C, or D

17. Let $Y_1 < Y_2 < Y_3 < Y_4 < Y_5 < Y_6$ be the order statistics of six independent observations from a continuous distribution where $\Pi_{35\%}$ is the 35^{th} percentile for the cumulative distribution function.
Find Prob($Y_2 < \Pi_{35\%} < Y_4$).
A. Less than 0.50
B. At least 0.50, but less than 0.55
C. At least 0.55, but less than 0.60
D. At least 0.60, but less than 0.65
E. At least 0.65

18. You have fit a multiple regression with 6 slopes and one intercept to 34 data points.
Residual (Error) Sum of Squares = 504.
The 13^{th} diagonal entry of the hat matrix, $h_{13,13} = 0.218$.
The 13^{th} residual is 5.81.
Determine Cook's Distance, D_{13}.
A. 0.07 B. 0.08 C. 0.09 D. 0.10 E. 0.11

19. Which of the following are true with respect to high-dimensional situations?
1. Regularization or shrinkage plays a key role.
2. For ridge regression or the lasso, choosing the tuning parameter is crucial.
3. The test MSE increases as the dimensionality of the problem increases, unless the additional predictors are truly associated with the response.

A. 1 only
B. 2 only
C. 3 only
D. 1, 2, and 3
E. The correct answer is not given by (A), (B), (C), or (D)

20. X is Normally distributed with mean 6 and standard deviation 0.4.
What is the standard deviation of e^X?

A. Less than 150
B. At least 150, but less than 160
C. At least 160, but less than 170
D. At least 170, but less than 180
E. At least 180

21. You are given that mortality follows the Illustrative Life Table and that you may use an interest rate, i = 6%.
An annuity-due is sold to a life aged 58, however the first twelve payments are guaranteed.
Determine the actuarial present value of the payments

A. 11.8 B. 12.0 C. 12.2 D. 12.4 E. 12.6

22. You simulate claims from a Pareto distribution with parameters $\alpha = 2.5$ and $\theta = 700$, and large random numbers corresponding to large claims.
Three random numbers from (0, 1): 0.103, 0.752, 0.450, are used to simulate three claims.
What is the sum of these three claims?

A. Less than 600
B. At least 600, but less than 650
C. At least 650, but less than 700
D. At least 700, but less than 750
E. 750 or more.

Use the following information about a branching process for the next 2 questions:
- The number of offspring for an individual are 0 with probability 25%, 1 with probability 40%, 2 with probability 25%, and 3 with probability 10%.
- Generation zero consists of 5 individuals.

23. What is the expected size of the eighth generation?
A. Less than 20
B. At least 20, but less than 30
C. At least 30, but less than 40
D. At least 40, but less than 50
E. At least 50

24. What is the variance of the size of the eight generation?
A. Less than 100
B. At least 100, but less than 200
C. At least 200, but less than 300
D. At least 300, but less than 400
E. At least 400

25. A sample of size 5 from a Normal Distribution with unknown mean μ and variance $\sigma^2 > 0$ has a sample standard deviation of 18.
Test the hypothesis that $\sigma = 11$ versus the alternative $\sigma > 11$.
A. Reject H_0 at 0.005.
B. Do not reject H_0 at 0.005; reject H_0 at 0.010.
C. Do not reject H_0 at 0.010; reject H_0 at 0.025.
D. Do not reject H_0 at 0.025; reject H_0 at 0.050.
E. Do not reject H_0 at 0.050.

26. A number of candidate models were fit using the following variables:
● An intercept term
● Variable A - a Yes/No indicator
● Variable B - a Yes/No indicator
● An interaction of Variables A and B
There are four observations, which were arranged into the following design matrix:

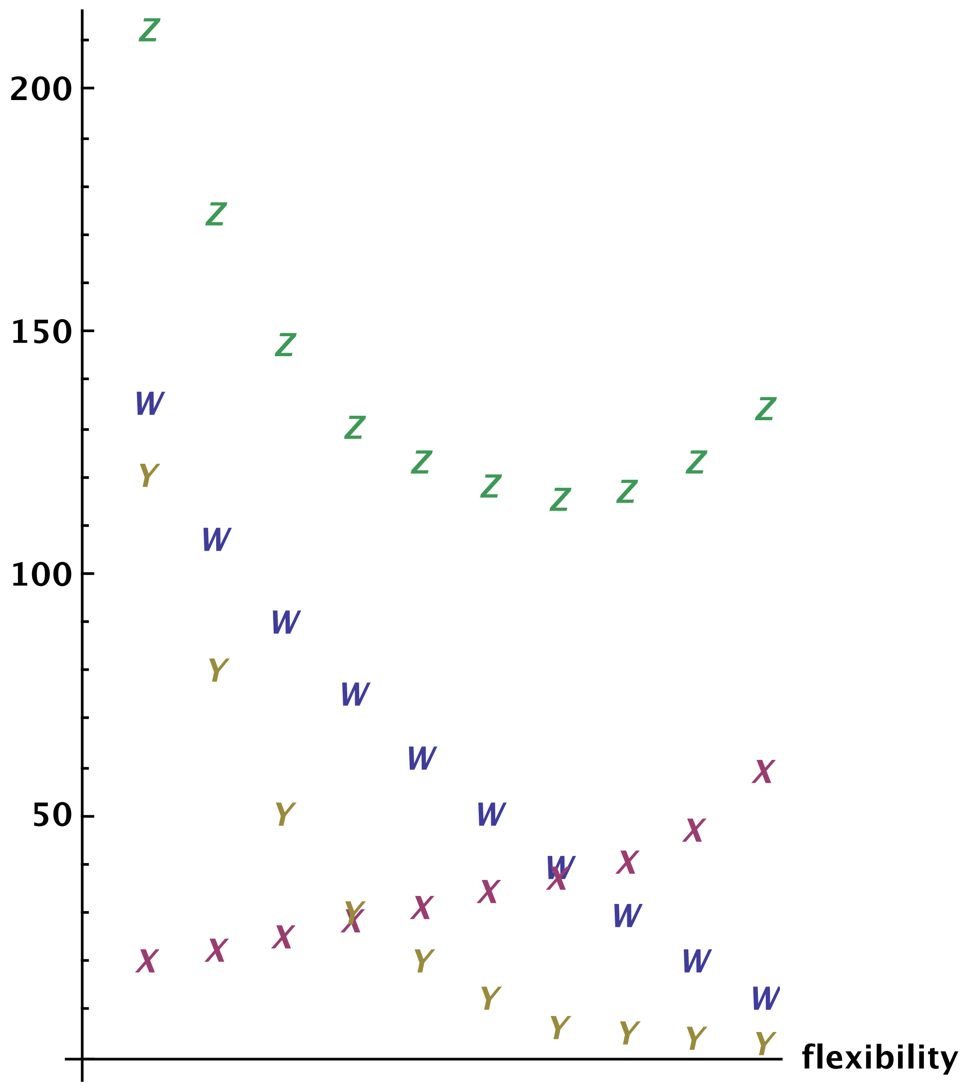
$$X = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

This data was fit using three different link functions:
I. Identity
II. Inverse
III. Log
The predicted values, given below, were the same under all three models:

$$\hat{Y} = \begin{pmatrix} 1 \\ 3 \\ 6 \\ 10 \end{pmatrix}$$

Determine for which of the above link functions the estimated interaction coefficient is positive.
A. Identity, Inverse, and Log
B. Identity and Inverse only
C. Identity and Log only
D. Inverse and Log only
E. The answer is not given by (A), (B), (C) or (D)

27. You have fit similar models to a large dataset and need to determine how flexible a model to use. Below is a chart of four statistics from this model valued for various amounts of flexibility: Training MSE, Test MSE, Squared Bias, and Variance.



Determine which of the following describes the above graph.

A. W is Squared Bias	X is Training MSE	Y is Variance
B. W is Variance	X is Squared Bias	Y is Training MSE
C. W is Training MSE	X is Variance	Y is Squared Bias
D. W is Variance	X is Training MSE	Y is Squared Bias

E. The correct answer is not given by (A), (B), (C), or (D)

28. Use the following information about a GLM:
- Adults were surveyed in a country about their relative attitudes towards abortion.
- There were three ordinal responses:
  1. "Negative" 2. "Mixed" 3. "Positive"
- A cumulative logit model has been fit.
- $x_1 = 1$ if a college graduate.
- $x_2 = 1$ if religious.
- $x_3 = x_1 x_2$
- The fitted parameters are: $\hat{\beta}_{01} = -0.7$, $\hat{\beta}_{11} = -0.56$, $\hat{\beta}_{21} = 1.45$, $\hat{\beta}_{31} = -0.28$,
  $\hat{\beta}_{02} = 0.6$, $\hat{\beta}_{12} = -1.25$, $\hat{\beta}_{22} = 1.84$, $\hat{\beta}_{32} = 0.23$.

For a religious college graduate, determine the probability of having a "mixed" attitude towards abortion.

A. 29% B. 31% C. 33% D. 35% E. 37%

29. Oil wells can be in one of three states: "Gusher," "Normal," or "Dry."
- A Gusher well makes a profit of 60,000 at the end of the year.
- A Normal well makes a profit of 20,000 at the end of the year.
- A Dry well makes no profit.

At the end of each year after paying the profit, the oil well is reclassified as follows:
- A Gusher well has a 60% chance of remaining a Gusher, a 30% chance of becoming Normal, and a 10% chance of becoming Dry.
- A Normal well has a 80% chance of remaining Normal and a 20% chance of becoming Dry.
- A Dry well remains Dry.

The interest rate is 8%.

Calculate the actuarial present value of a Gusher well at the beginning of the year.

(A) 150,000 (B) 155,000 (C) 160,000 (D) 165,000 (E) 170,000

30. For two independent random samples from Normal Distributions with the same variance you are given:

	Sample 1	Sample 2
Sample Size	25	20
Sample Mean	90	65
Sample Variance	800	1000

Test the hypothesis H_0: the means of the two distributions are equal, versus
H_1: the mean of the first distribution is larger than the mean of the second distribution.

A. Reject H_0 at 1/2%.
B. Do not reject H_0 at 1/2%. Reject H_0 at 1%.
C. Do not reject H_0 at 1%. Reject H_0 at 2.5%.
D. Do not reject H_0 at 2.5%. Reject H_0 at 5%.
E. Do not reject H_0 at 5%.

31. You fit a linear regression to 19 observations via least squares: $Y = \alpha + \beta X + \varepsilon$.
Let $\hat{Y}_i$ be the fitted values.

$\sum(X_i - \bar{X})^2 = 56.24$.

$\sum(X_i - \bar{X})(Y_i - \bar{Y}) = 21.03$.

$\sum(Y_i - \hat{Y}_i)^2 = 25.43$.

Let H_0 be the hypothesis that $\beta = 0$. H_1: $\beta \neq 0$.
Which of the following is true?

A. Reject H_0 at 1%.
B. Do not reject H_0 at 1%. Reject H_0 at 2%.
C. Do not reject H_0 at 2%. Reject H_0 at 5%.
D. Do not reject H_0 at 5%. Reject H_0 at 10%.
E. Do not reject H_0 at 10%.

32. For a claims process, you are given:
(i) The number of claims {N(t), t ≥ 0} is a nonhomogeneous Poisson process with intensity function:
$\lambda(t) = 50,\ 0 \le t < 2$ $\lambda(t) = 100,\ 2 \le t < 4$ $\lambda(t) = 150,\ 4 \le t.$
(ii) Claims amounts are independently and identically distributed random variables that are also independent of N(t).
(iii) Claims amounts are exponentially distributed with mean 1000.
(iv) The random variable P is the number of claims with claim amount less than 2000 by time t = 6.
(v) The random variable Q is the number of claims with claim amount greater than 2000 by time t = 6.
What is the conditional expected value of P, given Q = 50?
A. 480 B. 490 C. 500 D. 510 E. 520

33. You are given:
(i) A random sample of five hundred observations from a population are compared to a parametric distribution function.
(ii) You use the Kolmogorov-Smirnov test for testing the null hypothesis, H_0, that the distribution function for the population is the one to which the data was compared.
(iii) The test statistic is 0.075.
(iv) Critical values for the Kolmogorov-Smirnov test are:

Level of Significance:	0.20	0.10	0.05	0.01
Critical Value:	$1.07/\sqrt{n}$	$1.22/\sqrt{n}$	$1.36/\sqrt{n}$	$1.63/\sqrt{n}$

Determine the result of the test.
(A) Do not reject H_0 at the 0.20 significance level.
(B) Reject H_0 at the 0.20 significance level, but not at the 0.10 significance level.
(C) Reject H_0 at the 0.10 significance level, but not at the 0.05 significance level.
(D) Reject H_0 at the 0.05 significance level, but not at the 0.01 significance level.
(E) Reject H_0 at the 0.01 significance level.

34. You are given the following output from a GLM to estimate the probability of a claim:
● Distribution selected is Binomial.
● Link selected is Logit.

Parameter	β
Intercept	-1.6
Vehicle Body	
Coupe	-0.8
Roadster	-1.1
Sedan	-1.2
Station wagon	-1.0
Truck	-1.3
Utility	-1.4
Driver's Gender	
Male	0.1
Area	
B	-0.3
C	0.2
D	-0.1

Let V be the estimated probability of a claim for:
- Driver Gender: Male
- Vehicle Body: Coupe
- Area: C

Let W be the estimated probability of a claim for:
- Driver Gender: Female
- Vehicle Body: Sedan
- Area: B

Determine V / W.
A. 1.9 B. 2.0 C. 2.3 D. 2.5 E. 2.7

35. We have a sample of size n from the density: $f(x) = \lambda^2 x e^{-\lambda x}$, $x > 0$.
Determine a Minimum Variance Unbiased Estimator (MVUE) of λ.

A. $1 / \sum_{i=1}^{n} x_i$ B. $2 / \sum_{i=1}^{n} x_i$ C. $n / \sum_{i=1}^{n} x_i$ D. $2n / \sum_{i=1}^{n} x_i$

E. None of A. B, C, or D

Hint: The negative first moment of a Gamma Distribution is: $\theta^{-1} \Gamma[\alpha - 1] / \Gamma[\alpha] = \frac{1}{\theta (\alpha-1)}$.

36. A machine currently has a two engines each of whose time to failure has a constant force of mortality (hazard rate), μ.
The time to failure of these two engines are independent.
The machine fails when it no longer has an operating engine.
You are given two options to extend the time to failure of the machine.
1. Upgrade each engine, halving its force of mortality.
2. Upgrade one of the two engines, so that it has a constant force of mortality, $k\mu$.
You find that the machine's expected time to failure is the same under each of the two options.
Calculate k.
A. Less than 0.20
B. At least 0.20, but less than 0.25
C. At least 0.25, but less than 0.30
D. At least 0.30, but less than 0.35
E. At least 0.35

37. A dataset has n = 60 observations and p = 7 independent predictors.
Rick uses 5-fold cross validation to select from a variety of available models.
R is the number of times that the last observation will be included in the training dataset as part of his procedure.
Morty instead uses 10-fold cross validation.
M is the number of times that the first observation will be included in the training dataset as part of his procedure.
Determine M / R.
A. Less than 2.2
B. At least 2.2, but less than 2.4
C. At least 2.4, but less than 2.6
D. At least 2.6, but less than 2.8
E. At least 2.8

38. In 2016 losses follow an Exponential Distribution with mean 2000.
Inflation is 5% per year.
In 2019, there is deductible of size 1000, maximum covered loss of 5000,
and a coinsurance of 90%.
Determine the average payment per payment in 2019.
A. less than 1600
B. at least 1600 but less than 1700
C. at least 1700 but less than 1800
D. at least 1800 but less than 1900
E. at least 1900

39. Given the following random numbers from (0,1): 0.925, 0.419, 0.209, 0.685, 0.581, 0.112, 0.747, 0.564, use the rejection method to simulate a single random draw from a Standard Unit Normal. What is the value of that random draw?
A. less than -1
B. at least -1 but less than -0.3
C. at least -0.3 but less than 0.3
D. at least 0.3 but less than 1
E. at least 1

40. A special fund is established by collecting an amount of M from each of many independent lives age 70:
- The fund will pay $1000, payable at the end of the year of death, to each person who dies before age 73.
- The fund will pay M, payable at age 73, to each person who survives to age 73.
- Mortality follows the Illustrative Life Table.
- Interest rate i = 6%.

Calculate M using the equivalence principle.
A. Less than 350
B. At least 350, but less than 400
C. At least 400, but less than 450
D. At least 450, but less than 500
E. At least 500

41. Two different types of fertilizers are being tested on similar plots.
The first was applied to 33 test plots while the other type was applied to 29 test plots.
The amount of grain harvested from the different plots was recorded.

- H_0: The expected yields for the types are equal.
- H_1: The expected yields are not equal.

ANOVA was conducted and the resulting F-statistic was 6.19.
Which of the following is the correct conclusion?
A. H_0 will be rejected at the 0.01 significance level.
B. H_0 will be rejected at the 0.02 significance level, but not at the 0.01 level.
C. H_0 will be rejected at the 0.05 significance level, but not at the 0.02 level.
D. H_0 will be rejected at the 0.10 significance level, but not at the 0.05 level.
E. H_0 will not be rejected at the 0.10 significance level.

42. While riding down the highway, Justine is looking for cars with license plates from the state of Kenenbarby. Such cars become visible to Justine at a Poisson rate of 1 every 20 minutes. How many minutes must Justine ride down the highway so that there is at least a 95% chance of her seeing at least one such car?
A. 40 B. 45 C. 50 D. 55 E. 60

43. You are given the following:

- An insurance company provides a coverage which can result in only three loss amounts in the event that a claim is filed: \$500, \$1000 or \$5000.
- The probability, p, of a loss being \$500 is twice the probability of it being \$1,000.
- The following 6 claims are observed:

\$500 \$500 \$500 \$1,000 \$1000 \$5000

What is the maximum likelihood estimate of p?
A. Less than 0.30
B. At least 0.30, but less than 0.40
C. At least 0.40, but less than 0.50
D. At least 0.50, but less than 0.60
E. At least 0.60

END OF PRACTICE EXAM

This page intentionally left blank.

Solutions:

1. B. $(n-1)S^2/\sigma^2$ is Chi-Square with n - 1 degrees of freedom.
If the null hypothesis is true, then $19S^2/49$ is Chi-Square with 19 degrees of freedom.
For the Chi-Square with 19 degrees of freedom, the distribution function is 2.5% at 8.91.
2.5% = Prob[Type I error] = Prob[rejecting I H_0] = Prob[$19S^2/49 < 8.91$] = Prob[$S^2 < 23$].
Thus we can take k = **23**.
Comment: Similar to Q. 24.62 (CAS3L, 5/10, Q.22) in "Mahler's Guide to Statistics."
Since H_1: $\sigma < 7$, we reject H_0 when the statistic is small.

2. C. For the Normal $\delta_i^2 = \frac{1}{\sigma^2} (y_i - \hat{\mu}_i)^2 = \{(805 - 740)/44\}^2$.

Since 805 - 740 > 0, we take $\delta_i > 0$.

δ_i = (805 - 740)/44 = **1.477**.
Comment: Similar to Q. 12.8 in "Mahler's Guide to Generalized Linear Models."

3. B. $E[\alpha]$ = 14,130/5000 = 2.8260. $E[\beta]$ = 15,682/5000 = 3.1364.
$E[w\alpha + (1-w)\beta]$ = w2.8260 + (1-w)3.1364 = 3.1364 - 0.3104w.
$\text{Bias}[w\alpha + (1-w)\beta]$ = 3.1364 - 0.3104w - 3 = 0.1364 - 0.3104w.
$E[\alpha^2]$ = 41,062/5000 = 8.2124. $\text{Var}[\alpha] = 8.2124 - 2.8260^2 = 0.226124$.
$E[\beta^2]$ = 49,975/5000 = 9.995. $\text{Var}[\beta] = 9.995 - 3.1364^2 = 0.157995$.
Since α and β are independent: $\text{Var}[w\alpha + (1-w)\beta] = w^2\,\text{Var}[\alpha] + (1 - w)^2\text{Var}[\beta]$
$= w^2\,0.226124 + (1 - w)^2(0.157995) = 0.384119w^2 - 0.31599w + 0.157995$.
$\text{MSE}[w\alpha + (1-w)\beta] = \text{Var}[w\alpha + (1-w)\beta] + \text{Bias}[w\alpha + (1-w)\beta]^2 =$
$= 0.384119w^2 - 0.31599w + 0.157995 + (0.1364 - 0.3104w)^2$.
Setting the partial derivative of the MSE with respect w equal to zero:
0 = 0.768238w - 0.31599 + 2(-0.3104)(0.1364 - 0.3104w).
0 = 0.960934w - 0.40067. $\Rightarrow$ w = 0.40067 / 0.960934 = **0.417**.
Comment: See Q. 35.30 in "Mahler's Guide to Statistics."

4. B. For the Pareto Distribution, $S(x) = \{\theta/(\theta+x)\}^\alpha$.

$$\frac{\partial S(x)}{\partial \alpha} = \ln[\theta/(\theta+x)]\ \{\theta/(\theta+x)\}^\alpha.$$

$$\frac{\partial S(5000)}{\partial \alpha} = \ln[1/(1+5)]\ \{1/(1+5)\}^3 = -0.00830.$$

$$S(x) = \frac{\theta^\alpha}{(\theta+x)^\alpha}.\quad \frac{\partial S(x)}{\partial \theta} = \frac{\frac{\partial\, \theta^\alpha}{\partial \alpha}(\theta+x)^\alpha - \theta^\alpha \frac{\partial\, (\theta+x)^\alpha}{\partial \alpha}}{(\theta+x)^{2\alpha}} = \frac{\alpha\theta^{\alpha-1}\,(\theta+x)^\alpha - \theta^\alpha\, \alpha(\theta+x)^{\alpha-1}}{(\theta+x)^{2\alpha}}$$

$= \alpha\theta^{\alpha-1}\{(\theta+x) - \theta\} / (\theta+x)^{\alpha+1} = \alpha\theta^{\alpha-1}x / (\theta+x)^{\alpha+1}$.

$$\frac{\partial S(5000)}{\partial \theta} = 3(1000^2)(5000) / (6000)^4 = 0.0000116.$$

Thus the gradient vector is: (-0.00830, 0.0000116).
Thus using the delta method, the variance of the estimated value of the Survival Function is:
(transpose of gradient vector) (Inverse of the information matrix) (gradient vector) =

$$(-0.00830,\ 0.0000116)\begin{pmatrix} 0.029 & 12 \\ 12 & 5333 \end{pmatrix}\begin{pmatrix} -0.00830 \\ 0.0000116 \end{pmatrix}$$

= (-0.00830, 0.0000116) (-0.0001015, -0.0378) = 0.00000040.

The standard deviation is: $\sqrt{0.00000040}$ = **0.00064**.

Comment: Similar to Q. 40.23 in "Mahler's Guide to Statistics."
The given variance-covariance matrix was calculated for 5000 data points.

$S(x) = \{\theta/(\theta+x)\}^\alpha = y^\alpha$, where $y = \theta/(\theta+x)$. $\frac{\partial\, y^\alpha}{\partial \alpha} = \ln[y]\ y^\alpha$.

5. E. Thinning, medical malpractice insurance claims are Poisson with annual intensity 0.13λ.
For a doctor, the number of claims over ten years is Poisson with mean: $(10)(0.13\lambda) = 1.3\lambda$.
1.3λ follows a Gamma Distribution with $\alpha = 0.5$ and $\theta = (1.3)(1.7) = 2.21$.
Therefore, the mixed distribution is Negative Binomial with $r = 0.5$ and $\beta = 2.21$.
The probability of 4 claims over ten years from a doctor picked at random is the density at four for this Negative Binomial: $\frac{(0.5)\,(1.5)\,(2.5)\,(3.5)}{4!}\ \frac{2.21^4}{3.21^{4.5}}$ = **3.43%**.

Comment: Similar to Q. 9.9 in "Mahler's Guide to Stochastic Models."

6. A. The expected value of the score function is zero, for the maximum likelihood fit.
(This is why the method of scoring converges when the score is zero.)
Thus a score of zero would correspond to the maximum likelihood estimate of θ.
Thus the maximum likelihood theta is approximately 97.
The corresponding information is approximately 0.021.
Thus $\text{Var}[\hat{\theta}] = 1/0.021 = 47.6$.

An approximate 95% confidence interval for θ is: $97 \pm 2\sqrt{47.6}$ = [83 , **111**].

Comment: Similar to Q. 2.17 (MAS-1, 5/18, Q.40)
in "Mahler's Guide to Generalized Linear Models."

7. E. $f(x) = \alpha(10000^\alpha)(10000 + x)^{-(\alpha+1)}$. $S(x) = \{10000/(10000 + x)\}^\alpha$.

Size of Loss	Deductible	Max. Cov. Loss	Contribution to the Likelihood per Loss	Number
500	None	10,000	$f(500) = \alpha 10{,}000^\alpha/10{,}500^{(\alpha+1)}$	5
2500	1000	25,000	$f(2500)/S(1000) = \alpha 11{,}000^\alpha/12{,}500^{(\alpha+1)}$	4
≥ 10,000	None	10,000	$S(10{,}000) = 10{,}000^\alpha/20{,}000^\alpha$	3
≥ 25,000	1000	25,000	$S(25{,}000)/S(1000) = 11{,}000^\alpha/35{,}000^\alpha$	2

Therefore, the loglikelihood is:

$5 \ln(\alpha) + 5\alpha \ln(10{,}000) - 5(\alpha+1)\ln(10{,}500) + 4 \ln(\alpha) + 4\alpha \ln(11{,}000) - 4(\alpha+1)\ln(12{,}500)$
$+ 3\alpha \ln(10{,}000) - 3\alpha \ln(20{,}000) + 2\alpha \ln(11{,}000) - 2\alpha \ln(35{,}000) =$

$9 \ln(\alpha) - 5.1496\alpha$ + constants.

Setting the derivative of the loglikelihood with respect to α equal to zero:

$9/\alpha - 5.1496 = 0. \Rightarrow \alpha = 9/5.1496 =$ **1.748**.

Alternately, for a Pareto with θ fixed, the fitted alpha via maximum likelihood is:

$$\hat{\alpha} = \frac{\text{Number of Uncensored Values}}{\sum \ln[1 + \text{payment}_i / (\theta + d_i)]} =$$

$$\frac{9}{5 \ln[1 + 0.5/10] + 4 \ln[1 + 1.5/11] + 3 \ln[1 + 10/10] + 2 \ln[1 + 24/11]} = 9/5.1496$$

= **1.748**.

<u>Comment</u>: Similar to Q. 6.128 (4, 5/05, Q.27) in “Mahler’s Guide to Statistics.”

8. D. Since H_1: $\sigma_X^2 \neq \sigma_Y^2$, we perform a two-sided test; we take the twice the probability in the righthand tail.

Rejecting at 2%. ⇔ less than 1% area in the righthand tail.
F = 600/130 = 4.615. We reject when this F statistic is greater than the 1% critical value.
The sample sizes of X and Y are given as the same.
If n = 11, then the F-statistic has 10 and 10 degrees of freedom.
The 1% critical value is 4.85, and we would not reject.
If n = **12**, then the F-statistic has 11 and 11 degrees of freedom.
The 1% critical value is 4.46 < 4.615, and we would reject.
<u>Comment</u>: Similar to Q. 27.27 (CAS3L, 5/09, Q.21) in “Mahler’s Guide to Statistics.”

9. A. From the covariance matrix, $\text{Var}[\hat{\beta}_0] = 0.0431$.

95% confidence interval for β_0: $3.1 \pm 1.960 \sqrt{0.0431}$ = (2.693, **3.507**).

<u>Comment</u>: Similar to Q. 7.1 in “Mahler’s Guide to Generalized Linear Models.”
Since we are not given the sample size, I have assumed that the sample size is big enough to use the Normal Approximation.

10. B. $f(x) = \lambda x^{\lambda-1}$. $\ln f(x) = \ln(\lambda) + (\lambda - 1)\ln(x)$. $g(y) = \mu y^{\mu-1}$. $\ln g(y) = \ln(\mu) + (\mu - 1)\ln(y)$.
Thus the loglikelihood is: $n \ln(\lambda) + (\lambda - 1)\sum\ln(x_i) + m \ln(\mu) + (\mu - 1)\sum\ln(y_i)$.
For H_0: $\lambda = 2$, $\mu = 4$, the loglikelihood is: $n \ln(2) + \sum\ln(x_i) + m \ln(4) + 3\sum\ln(y_i)$.
For H_1: $\lambda = 3$, $\mu = 6$, the loglikelihood is: $n \ln(3) + 2\sum\ln(x_i) + m \ln(6) + 5\sum\ln(y_i)$.
The difference between the loglikelihoods for H_1 and H_0 is:
$n \ln(3/2) + \sum\ln(x_i) + m \ln(6/4) + 2\sum\ln(y_i)$.
Thus the critical region for the best (Neyman-Pearson) test of H_0 against H_1 is:
$n \ln(3/2) + \sum\ln(x_i) + m \ln(6/4) + 2\sum\ln(y_i) \geq b$. $\Rightarrow$ $\boldsymbol{\sum\ln(x_i) + 2\sum\ln(y_i) \geq c}$.
Comment: Similar to Q. 30.24 (2, 5/88, Q.30) in "Mahler's Guide to Statistics."

11. A. F-Statistic = $\dfrac{(\text{TSS - RSS}) / (k-1)}{\text{RSS} / (N - k)} = \dfrac{(\text{TSS - RSS}) / (6-1)}{\text{RSS} / (18 - 6)}$ = 2.4 (TSS - RSS) / RSS.
F has k - 1 = 5, and N - 5 = 12 degrees of freedom, and the 1% critical value is 5.06.
Critical region is when we reject H_0, which is when F ≥ 5.06. ⇒
2.4 (TSS - RSS) / RSS ≥ 5.06. ⇒ **TSS ≥ 3.11 RSS**.
Comment: Similar to Q. 10.24 in "Mahler's Guide to Regression."
This F-Test is always one-sided; we reject when the F-Statistic is large.

12. C. The probability that they will both be dead in one year is:
$q_{97}\, q_{91}$ = (0.32834) (0.20493) = 0.067287.
Probability at least one of them is alive in one year is: 1 - 0.067287 = 0.93271.
The probability that they will both be dead in two years is:
${}_2q_{97}\, {}_2q_{91}$ = (1 - 64,617/148,832) (1 - 530,959/858,676) = 0.21595
Probability at least one alive in two years is: 1 - 0.21595 = 0.78405.
APV = 1 + 0.93271/ 1.08 + 0.78405 / 1.08^2 = **2.536**.
Comment: Similar to Q. 8.4 in "Mahler's Guide to Life Contingencies."

13. B. A and E are each cubic splines, but a natural cubic spline is linear below the first knot at 20 and above the last knot at 80.
C is a quadratic spline. D is a linear spline.
Comment: Similar to Q.12.13 in "Mahler's Guide to Statistical Learning."

14. C. Compute the observed first two moments.

	Claim Size	Square of Claim Size
	65	4225
	41	1681
	51	2601
	42	1764
	26	676
	21	441
	43	1849
	12	144
	104	10816
	96	9216
Average	50.1	3341.3

Match the observed and theoretical means and variances:
$\alpha\theta = 50.1$, and $\alpha\theta^2 = 3341.3 - 50.1^2 = 831.29$.
Therefore, $\alpha = (\alpha\theta)^2 / (\alpha\theta^2) = 50.1^2 / 831.29 = 3.0194$, and $\theta = 50.1/3.0194 = 16.593$.
The third moment of a Gamma Distribution is:
$\theta^3(\alpha+2)(\alpha+1)\alpha = (16.593^3)(5.0194)(4.0194)(3.0194) =$ **278,297**.
Comment: Similar to Q.2.7-2.8 in "Mahler's Guide to Statistics.
The empirical third moment could be calculated in manner parallel to the empirical second moment.
However, that is not what we are asked to do in this question.
1. Fit a Gamma via method of moments. (Unless stated otherwise, since there are two parameters match the first two moments.)
2. For the fitted Gamma Distribution, calculate its third moment.
In general, the third moment of the fitted Gamma would differ from the empirical third moment.

15. E. $Y \geq 30 \Leftrightarrow$ At most 3 of the sample < 30.
Prob[$Y \geq 30$] = Prob[At most 3 of the sample < 30] =
Prob[0 of the sample < 30] + Prob[1 of the sample < 30] + Prob[2 of the sample < 30]
+ Prob[3 of the sample < 30] =
$0.7^{13} + 13(0.7^{12})(0.3) + \{(13)(12)/2\}(0.7^{11})(0.3^2) + \{(13)(12)(11)/6\}(0.7^{10})(0.3^3) =$ **0.4206**.
Comment: Similar to Q. 33.11 in "Mahler's Guide to Statistics."

16. D. A minimal path set is such that the system functions if all of the components of the set function, but the system would not function if any component in the set fails.
In this case, this is true for any subset of size four of the six components.
Comment: Similar to Q. 2.1 in "Mahler's Guide to Reliability Theory."

17. C. We want exactly 2 or 3 values in the sample to be less than $\Pi_{35\%}$.

Prob[exactly two values < $\Pi_{35\%}$] = $\binom{6}{2} F(\Pi_{35\%})^2 S(\Pi_{35\%})^4$ = (15) (0.35^2) (0.65^4) = 0.3280.

Prob[exactly three values < $\Pi_{35\%}$] = $\binom{6}{3} F(\Pi_{35\%})^3 S(\Pi_{35\%})^3$ = (20) (0.35^3) (0.65^3) = 0.2355.

Thus, Prob($Y_2 < \Pi_{35\%} < Y_4$) = 0.3280 + 0.2355 = **0.5635**.
Comment: Similar to Q. 34.39 (CAS3L, 5/13, Q.23) in "Mahler's Guide to Statistics."
If exactly 0 of the sample are less than $\Pi_{35\%}$, then $\Pi_{35\%} < Y_1 < Y_2$, No Good!
If exactly 1 of the sample is less than $\Pi_{35\%}$, then $\Pi_{35\%} < Y_2$, No Good!
If exactly 2 of the sample are less than $\Pi_{35\%}$, then $Y_2 < \Pi_{35\%} < Y_3 < Y_4$, OK
If exactly 3 of the sample are less than $\Pi_{35\%}$, then $Y_2 < Y_3 < \Pi_{35\%} < Y_4$, OK
If exactly 4 of the sample are less than $\Pi_{35\%}$, then $Y_4 < \Pi_{30} < Y_5$, No Good!

18. C. $\hat{\sigma}^2$ = RSS/(N-p-1) = 504 / (34 - 6 - 1) = 18.67.
The standard error of the 13th residual is: RSE $\sqrt{1 - h_{13,13}} = \sqrt{(18.67)(1 - 0.218)}$ = 3.82.

D_i is also: $\left(\frac{\hat{\varepsilon}_i}{\text{Standard Error of } \hat{\varepsilon}_i}\right)^2 \frac{h_{ii}}{(p+1)\ (1- h_{ii})}$.

Thus $D_{13} = (5.81/3.82)^2 \frac{0.218}{(6+1)\ (1 - 0.218)}$ = **0.092**.

Comment: Similar to Q. 15.12 in "Mahler's Guide to Regression."
Let $\hat{\beta}(i)$ be the vector of fitted coefficients, excluding observation i.
Let $\hat{Y}(i) = X \hat{\beta}(i)$, the fitted values, when observation i is excluded from the regression.

$$D_i = \frac{\sum_{j=1}^{N} (\hat{Y}_j - \hat{Y}_{j(i)})^2}{(k+1)\ s^2}.$$

19. D. All three statements are true.
Comment: Similar to Q.9.3 in "Mahler's Guide to Statistical Learning."

20. E. e^X is LogNormal with $\mu = 6$ and $\sigma = 0.4$.
Mean of LogNormal = $E[e^X] = \exp[\mu + \sigma^2/2] = \exp[6 + 0.4^2/2] = 437.03$.
$E[(e^X)^2] = E[e^{2X}]$ = Second moment of LogNormal = $\exp[2\mu + 2\sigma^2]$ =
$\exp[(2)(6) + (2)(0.4^2)] = 224{,}134$.
Variance of LogNormal = $224{,}134 - 437.03^2 = 33{,}139$.
Standard Deviation of LogNormal = $\sqrt{33{,}139}$ = **182.0**.
Comment: Similar to Q. 21.21 in "Mahler's Guide to Loss and Frequency Distributions."

21. C. $\ddot{a}_{\overline{12|}} = \frac{1 - v^{12}}{d} = \frac{1 - 1/1.06^{12}}{0.06/1.06} = 8.8869$.
APV = $\ddot{a}_{\overline{12|}} + v^{12}\ {}_{12}p_{58}\ \ddot{a}_{70} = 8.8869 + (6{,}616{,}155/8{,}389{,}826)\ (8.5693)\ /\ 1.06^{12} =$ **12.2453**.
Comment: Similar to Q. 5.10 in "Mahler's Guide to Life Contingencies."

22. D. Set $u = F(x) = 1 - \{\theta/(\theta+x)\}^{\alpha}$. Solving, $x = \theta\{(1-u)^{-1/\alpha} - 1\} = 700\ \{(1-u)^{-0.4} - 1\}$.
Thus the three simulated claim sizes are: $700\{(1-0.103)^{-0.4} - 1\} = 31$, $700\{(1-0.752)^{-0.4} - 1\} = 523$, and $700\{(1-0.450)^{-0.4} - 1\} = 189$. 31 + 523 + 189 = **743**.
Comment: Similar to Q. 4.5 in "Mahler's Guide to Simulation."

23. B. $\mu = (0)(25\%) + (1)(40\%) + (2)(25\%) + (3)(10\%) = 1.2$.
$E[X_8] = X_0\ \mu^8 = (5)(1.2^8) =$ **21.5**.
Comment: Similar to Q. 27.2 in "Mahler's Guide to Stochastic Models."

24. C. $(0^2)(25\%) + (1^2)(40\%) + (2^2)(25\%) + (3^2)(10\%) = 2.3$. Thus $\sigma^2 = 2.3 - 1.2^2 = 0.86$.
Since $\mu \neq 1$, $Var[X_8] = X_0\ \sigma^2\ \mu^{n-1}\ (1 - \mu^n)/\ (1-\mu) = (5)(0.86)(1.2^{8-1})\ (1 - 1.2^8)\ /\ (1 - 1.2) =$ **254.2**.
Comment: Similar to Q. 27.3 in "Mahler's Guide to Stochastic Models."

25. D. $(n - 1)S^2/\sigma^2$ has a Chi-Square Distribution with 5 - 1 = 4 degrees of freedom.
This Chi-Square Distribution is 95% at 9.49 and 97.5% at 11.14.
In other words, the survival function is 5% at 9.49 and 2.5% at 11.14.
H_1: $\sigma > 11$, therefore reject when S^2 is large.
The test statistic is: $(4)(18^2)/11^2 = 10.71$. Perform a one-sided test. $9.49 < 10.71 < 11.14$.
Reject at 5%, do not reject at 2.5%.
Comment: Similar to Q. 15.3 in "Mahler's Guide to Statistics."

26. B. For the Identity Link Function: $\beta_0 = 1$. $\beta_0 + \beta_A = 3$. $\Rightarrow \beta_A = 2$.

$\beta_0 + \beta_B = 6$. $\Rightarrow \beta_B = 5$. $\beta_0 + \beta_A + \beta_B + \beta_{AB} = 10$. $\Rightarrow \beta_{AB} = 10 - 1 - 2 - 5 = 2 > 0$.

For the Inverse Link Function: $\beta_0 = 1/1 = 1$. $\beta_0 + \beta_A = 1/3$. $\Rightarrow \beta_A = -2/3$.

$\beta_0 + \beta_B = 1/6$. $\Rightarrow \beta_B = -5/6$. $\beta_0 + \beta_A + \beta_B + \beta_{AB} = 1/10$.

$\Rightarrow \beta_{AB} = 1/10 - 1 - (-2/3) - (-5/6) = 0.6 > 0$.

For the Log Link Function: $\beta_0 = \ln(1) = 0$. $\beta_0 + \beta_A = \ln(3)$. $\Rightarrow \beta_A = \ln(3)$.

$\beta_0 + \beta_B = \ln(6)$. $\Rightarrow \beta_B = \ln(6)$.

$\beta_0 + \beta_A + \beta_B + \beta_{AB} = \ln(10)$. $\Rightarrow \beta_{AB} = \ln(10) - 0 - \ln(3) - \ln(6) = -0.588 < 0$.

Thus the estimated interaction coefficient is positive for the **Identity and Inverse only**.
Comment: Similar to Q. 5.31 (MAS-1, 11/19, Q.26)
in “Mahler’s Guide to Generalized Linear Models.”

27. C. Test MSE generally exhibits a U-shaped behavior as per Line Z.
Variance increases with flexibility as per Line X.
Bias Squared and Training MSE both decrease with flexibility as per Lines W and Y.
Also Test MSE is the sum of Variance + Bias Squared + Irreducible Error.
Thus Z - X - Bias Squared = Irreducible Error = Var[ε] = constant.
Z - X - Y = constant, while Z - X - W is not a constant.
Thus Y is the squared bias, while W is the training MSE.
Comment: Similar to Q. 1.12 (MAS-1 Sample Q.2) of “Mahler’s Guide to Statistical Learning.”
Reading values off the graph as best as possible:

W	X	Y	Z	Z - X - W	Z - X - Y
135	20	120	212	57	72
107	22	80	174	45	72
90	25	50	147	32	72
75	28	30	130	27	72
62	31	20	123	30	72
50	34	12	118	34	72
39	37	6	115	39	72
29	40	5	117	48	72
20	47	4	123	56	72
12	59	3	134	63	72

28. C. The form for the cumulative logit model:

$$\ln\left[\frac{\text{Prob[in category j or less]}}{\text{Prob[in category more than j]}}\right] = \ln\left[\frac{\pi_1 + ... + \pi_j}{\pi_{j+1} + ... + \pi_J}\right] = x^T \beta_j.$$

The model for j = 1: $\ln\left[\frac{\pi_1}{\pi_2 + \pi_3}\right] = -0.7 - 0.56 + 1.45 - 0.28 = -0.09$. $\Rightarrow \pi_1 = 0.91393 (\pi_2 + \pi_3)$

$\Rightarrow \pi_1 = 0.91393 (1 - \pi_1)$. $\Rightarrow \pi_1 = 0.47752$.

The model for j = 2: $\ln\left[\frac{\pi_1 + \pi_2}{\pi_3}\right] = 0.6 - 1.25 + 1.84 + 0.23 = 1.42$. $\Rightarrow \pi_1 + \pi_2 = 4.13712\ \pi_3$.

$\Rightarrow \pi_1 + \pi_2 = 4.13712 \{1 - (\pi_1 + \pi_2)\}$. $\Rightarrow \pi_1 + \pi_2 = 0.80534$. $\Rightarrow \pi_2 =$ **0.32782**. $\Rightarrow \pi_3 = 0.19466$.

Comment: Similar to Q. 16.3 in "Mahler's Guide to Generalized Linear Models."
For this individual there is: a 47.8% chance of negative, a 32.8% chance of mixed, and a 19.5% chance of positive. Not intended as a realistic model.

29. E. Let x be the actuarial present value of a "Gusher" at the beginning of the year.
Let y be the actuarial present value of a "Normal" at the beginning of the year.
Since a dry well remains dry, the actuarial present value of a "Dry" at the beginning of the year is zero.
The APV at the beginning of this year is the present value of the profits at the end of this year plus the APV of the remaining profits at the end of this year:
$x = 60{,}000/1.08 + 0.6x/1.08 + 0.3y/1.08. \Rightarrow 0.48x = 60{,}000 + 0.3y.$
$y = 20{,}000/1.08 + 0.8y/1.08. \Rightarrow y = 71{,}429. \Rightarrow x =$ **169,643**.
Alternately, if a well is Normal in a year, then it has a 80% chance of remaining Normal next year, and a 20% of becoming Dry next year and never paying a profit again.
Thus if a well is Normal this year, it has a 0.8 chance of being Normal next year, a 0.8^2 chance of being Normal in two years, a 0.8^3 chance of being Normal in 3 years, etc.
Therefore, if the well is Normal, then the APV of profits is:
$(20{,}000) \{1/1.08 + (0.8)(1/1.08^2) + (0.8^2)(1/1.08^3) + (0.8^3)(1/1.08^4) + ...\}$
$= (20{,}000)\ (1/1.08) / \{1 - (0.8/1.08)\} = 71{,}429.$
If a well is a Gusher in a year, then it has a 60% chance of remaining a Gusher next year, a 30% of becoming Normal next year and never being a Gusher again, and a 10% of becoming Dry next year and never being a Gusher again
Thus if a well is a Gusher this year, it has a 0.6 chance of being a Gusher next year, a 0.6^2 chance of being a Gusher in two years, a 0.6^3 chance of being a Gusher in 3 years, etc.
Therefore, if the well is a Gusher, then the APV of profits from Gusher years is:
$(60{,}000) \{1/1.08 + (0.6)(1/1.08^2) + (0.6^2)(1/1.08^3) + (0.6^3)(1/1.08^4) + ...\}$
$= (60{,}000)\ (1/1.08) / \{1 - (0.6/1.08)\} = 125{,}000.$
If a well is a Gusher, then there is a 30% chance it becomes Normal next year, a (0.6)(0.3) chance it first becomes Normal the year after that, a $(0.6^2)(0.3)$ chance it first becomes Normal the year after that, etc.
When the well becomes Normal the Actuarial Present Value of future profits is 71,429.
Therefore, if the well is a Gusher, then the APV of profits from Normal years is:
$(71{,}429) \{0.3/1.08 + (0.3)(0.6)(1/1.08^2) + (0.3)(0.6^2)(1/1.08^3) + (0.3)(0.6^3)(1/1.08^4) + ...\}$
$= (71{,}429)\ (0.3/1.08) / \{1 - (0.6/1.08)\} = 44{,}643.$
Thus, if the well is a Gusher, then the total APV of profits is:
125,000 + 44,643 = **169,643**.
Comment: Similar to Q. 19.46 (CAS3, 11/05, Q.39) in “Mahler’s Guide to Stochastic Models.”

30. A. The pooled sample variance is $s_p^2 = \{(24)(800) + (19)(1000)\} / (24 + 19) = 888.$
$t = (90 - 65) / \sqrt{888/25 + 888/20} = 2.796$, with 24 + 19 = 43 degrees of freedom.
Performing a one-sided test, $2.704 < 2.796$. **Reject at 1/2%**.
Comment: Similar to Q. 21.3 in “Mahler’s Guide to Statistics.”

31. C. $\hat{\beta} = \sum x_i y_i / \sum x_i^2 = \sum (X_i - \bar{X})(Y_i - \bar{Y}) / \sum (X_i - \bar{X})^2 = 21.03/56.24 = 0.3739$.

Error SS $= \sum \hat{\varepsilon}_i^2 = \sum (Y_i - \hat{Y}_i)^2 = 25.43$. σ^2 = ESS/(N - k) = 25.43/(19 - 2) = 1.4959.

$s_{\hat{\beta}}^2 = \sigma^2 / \sum (X_i - \bar{X})^2 = 1.4959/56.24 = 0.02660$. $s_{\hat{\beta}} = \sqrt{0.02660} = 0.1631$.

$t = \hat{\beta} / s_{\hat{\beta}} = 0.3739/0.1631 = 2.292$. For 19 - 2 = 17 degrees of freedom, for a two-tailed test using the t-distribution, the critical values are:

10%	5%	2%	1%
1.740	2.110	2.567	2.898

Since 2.110 < 2.292 < 2.567, we **reject H_0 at 5% and do not reject H_0 at 2%**.
Comment: Similar to Q. 7.16 in "Mahler's Guide to Regression."
Here I have used deviations form to get the fitted slope.

Another formula for the fitted slope: $\hat{\beta} = \dfrac{N \sum X_i Y_i - \sum X_i \sum Y_i}{N \sum X_i^2 - \left(\sum X_i\right)^2}$.

32. E. We are thinning a nonhomogeneous Poisson process.
P and Q are independent nonhomogeneous Poisson processes.
Thus the conditional expected value of P, given a certain value for Q, is just the unconditional mean of P.
For the original process, the mean number of claims by time 6 is:

$$m(6) = \int_0^6 \lambda(t)\, dt = (2)(50) + (2)(100) + (2)(150) = 600.$$

The probability of a claim being of less than 2000 is: $1 - e^{-2000/1000} = 1 - e^{-2} = 0.8647$.
Therefore, the expected value of P is: (0.8647)(600) = **519**.
Comment: Similar to Q. 16.14 (3, 5/01, Q.37) in "Mahler's Guide to Stochastic Models."

33. E. Critical values are:
$1.07/\sqrt{500} = 0.0479$, $1.22/\sqrt{500} = 0.0546$, $1.36/\sqrt{500} = 0.0608$, $1.63/\sqrt{500} = 0.0729$.

$0.075 > 0.0729$. $\Rightarrow$ **Reject at 1%**.
Comment: Similar to Q. 7.1 in "Mahler's Guide to Statistics."

34. D. For the first case: $\exp[\beta x] = \exp[-1.6 + 0.1 - 0.8 + 0.2] = e^{-2.1} = 0.12246$.

For the logit link function: $\mu = \frac{e^{\beta x}}{e^{\beta x} + 1}$ = 0.12246 / (0.12246 + 1) = 10.91%.

For the first case: $\exp[\beta x] = \exp[-1.6 + 0 - 1.2 - 0.3] = e^{-3.1} = 0.04505$.

For the logit link function: $\mu = \frac{e^{\beta x}}{e^{\beta x} + 1}$ = 0.04505 / (0.04505 + 1) = 4.31%.

10.91% / 4.31% = **2.53**.
Comment: Similar to Q. 15.34 (CAS S, 11/15, Q.33)
in "Mahler's Guide to Generalized Linear Models."
Not intended as a realistic model.

35. E. $\ln[f(x)] = 2 \ln[\lambda] + \ln[x] - \lambda x$.
This is an Exponential family with $p(\lambda) = \lambda$ and $K(x) = -x$.

Thus either $-\sum_{i=1}^{n} x_i$ or $\sum_{i=1}^{n} x_i$ is a sufficient statistic for λ.

$$\sum_{i=1}^{n} \ln[f(x_i)] = 2n \ln[\lambda] + \sum_{i=1}^{n} \ln[x_i] - \lambda \sum_{i=1}^{n} x_i .$$

Setting the partial derivative of the loglikelihood with respect lambda equal to zero:

$$0 = 2n/\lambda - \sum_{i=1}^{n} x_i . \Rightarrow \hat{\lambda} = 2n / \sum_{i=1}^{n} x_i .$$

However, as will be seen, this maximum likelihood estimator of lambda is biased.
Now each x_i is Gamma with $\alpha = 2$ and $\theta = 1/\lambda$.

Therefore, $\sum_{i=1}^{n} x_i$ is Gamma with $\alpha = 2n$ and $\theta = 1/\lambda$.

$E[1 / \sum_{i=1}^{n} x_i]$ is the negative first moment of this Gamma, which is:

$\theta^{-1}\, \Gamma[\alpha - 1] / \Gamma[\alpha] = \lambda\, (2n-2)! / (2n-1)! = \lambda / (2n-1)$, $n > 1$.

Thus, $\mathbf{(2n-1) / \sum_{i=1}^{n} x_i}$ is an unbiased estimator of λ which is a function of a sufficient statistic,

and therefore is a Minimum Variance Unbiased Estimator of λ.
Comment: Similar to Q. 42.7 in "Mahler's Guide to Statistics."

36. E. The minimum of two independent Exponentials has hazard rate equal to the sum of the two individual hazard rates.
minimum + maximum = sum. ⇒ E[maximum] = E[sum] - E[minimum] = $2/\mu - 1/(2\mu) = 1.5/\mu$.
Thus in the first option, the expected time to failure is: $1.5 / (\mu/2) = 3/\mu$.
In the second option, the probability that the machine has failed at time t is: $(1 - e^{-\mu t})\ (1 - e^{-k\mu t})$.
The survival function is: $1 - (1 - e^{-\mu t})\ (1 - e^{-k\mu t}) = e^{-\mu t} + e^{-k\mu t} - e^{-(k+1)\mu t}$.
Integrating over t from zero to infinity, the mean time to failure is: $1/\mu + 1/(k\mu) - 1/\{(k+1)\mu\}$.
Thus we want: $3/\mu = 1/\mu + 1/(k\mu) - 1/\{(k+1)\mu\}$. ⇒ $0 = 2 - 1/k + 1/(1+k)$.
⇒ $2(k+1)k - (k+1) + k = 0$. ⇒ $2k^2 + 2k - 1 = 0$.
Taking the positive root, $k = \frac{-2 + \sqrt{2^2 - (4)(2)(-1)}}{(2)(2)} = \frac{-2 + \sqrt{12}}{4}$ = **0.366**.
Comment: Similar to Q. 6.22 (CAS LC, 5/16, Q.5) in "Mahler's Guide to Reliability Theory."

37. B. Using 5-fold cross validation, the data is split into 5 subsets, and each subset in turn acts as the test data, while the remaining 4 subsets act as the training data.
Each data point is in the test set once and in the training set the remaining 5 - 1 = 4 times.
⇒ R = 4.
Using 10-fold cross validation, the data is split into 10 subsets, and each subset in turn acts as the test data, while the remaining 9 subsets act as the training data.
Each data point is in the test set once and in the training set the remaining 10 - 1 = 9 times.
⇒ M = 9.
M/R = 9/4 = **2.25**.
Comment: Similar to Q. 2.18 (MAS-1, 5/19, Q.34) in "Mahler's Guide to Statistical Learning."

38. C. In 2019 the losses are Exponential with $\theta = (1.05^3)(2000) = 2315$.
$E[X \wedge 1000] = (2315)(1 - e^{-1000/2315}) = 812$.
$E[X \wedge 5000] = (2315)(1 - e^{-5000/2315}) = 2048$.
Average payment per loss is:
$(0.9)\ (E[X \wedge 5000] - E[X \wedge 1000]) = (0.9)\ (2048 - 812) = 1112$.
$S(1000) = e^{-1000/2315} = 0.6492$.
Average payment per payment is: 1112 / 0.6492 = **1713**.
Alternately use the original Exponential and the formula for the average payment per payment:
$(1+r)\ c\ (E[X \wedge u/(1+r)] - E[X \wedge d/(1+r)]) / S(d/(1+r)) =$
$(1.05^3)\ (0.9)\ (E[X \wedge 5000/1.05^3] - E[X \wedge 1000/1.05^3]) / S(1000/1.05^3) =$
$(1.1576)(0.9)\ \{2000(1 - e^{-4319/2000}) - 2000(1 - e^{-864/2000})\} / e^{-864/2000}$ = **1713**.
Comment: Similar to Q. 34.44 in "Mahler's Guide to Loss and Frequency Distributions".

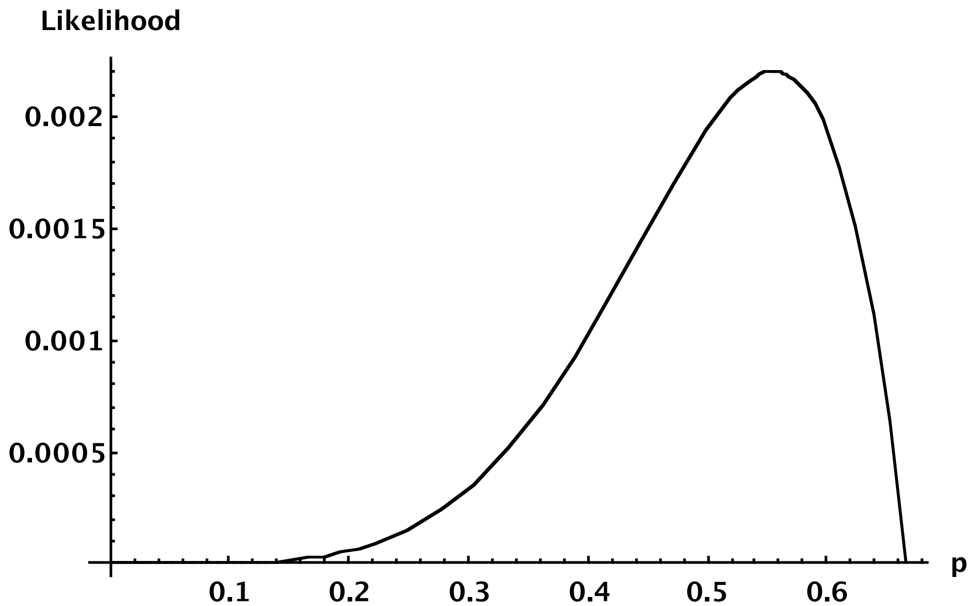
39. C. Simulate a random draw from an Exponential with mean 1: x = -ln(1 - 0.925) = 2.590.
$c = \sqrt{2e/\pi}$. Accept this loss with probability: f(x) / {cg(x)} =
$(2/\sqrt{2\pi})\exp(-x^2/2) / \{e^{-x}\sqrt{2e/\pi}\} = \exp(-(x-1)^2/2) = \exp[-(2.590-1)^2/2] = 0.2825$.
Since 0.419 > 0.2825 we reject this claim.
Simulate another random draw from an Exponential with mean 1: x = -ln(1 - 0.209) = 0.234.
Accept this loss with probability: $\exp(-(0.234-1)^2/2) = 0.746$.
Since 0.685 ≤ 0.746 we accept this loss.
Since 0.581 > 0.5, we take the random draw: **-0.234**, rather than +0.234.
Comment: Similar to Q. 12.4 in "Mahler's Guide to Simulation."

40. B. $q_{70} = 0.03318$. $q_{71} = 0.03626$. $q_{72} = 0.03962$.
The fund pays death benefits to someone who dies while they are 70, 71, or 72.
The actuarial present value of the death benefits:
($1000) {0.03318/1.06 + (1 - 0.03318)(0.03626)/1.06^2
+ (1 - 0.03318)(1 - 0.03626)(0.03962)/1.06^3} = $93.50.
Applying the equivalence principle:
M = $93.50 + M (1 - 0.03318) (1 - 0.03626) (1 - 0.03962) / 1.06^3.
⇒ M = **$376**.
Comment: Similar to Q. 6.13 (CAS LC, 11/15, Q.13) in "Mahler's Guide to Life Contingencies."

41. B. With two samples, ANOVA is equivalent to a (two-sided) two-sample t-test.
$t = \sqrt{F} = \sqrt{6.19} = 2.49$.
The number of degrees of freedom is: (33 - 1) + (29 - 1) = 60.
For this two-sided test, the 2% critical value is 2.390 and the 1% critical value is 2.660.
Since 2.390 < 2.49 < 2.660, **reject at 2% and not at 1%**.
Comment: Similar to Q. 28.4 in "Mahler's Guide to Statistics."
This is an F-Statistic with 2 - 1 = 1, and (33 - 1) + (29 - 1) = 60 degrees of freedom.
However, the tables attached to the exam do not extend that far.

42. E. Assuming Justine drives t minutes, the chance of seeing no such car arriving is $e^{-t/20}$.
Setting that equal to 5%: $0.05 = e^{-t/20}$. ⇒ t = **59.9 minutes**.
Comment: Similar to Q. 3.36 (CAS3, 11/04, Q.18) in "Mahler's Guide to Stochastic Models."
For a Poisson Process with rate λ, the probability of no events by time t is: $1 - e^{-\lambda t}$.

43. D. Prob[X = 1000] = Prob[X = 500]/2 = p/2.
Therefore, Prob[X = 5000] = 1 - 1.5p.
Therefore, $0 \le p \le 2/3$.
The likelihood of the observation is: p p p (p/2) (p/2) (1 - 1.5p) = $p^5/4 - 0.375p^6$.
Set the derivative with respect to p equal to zero:
$5p^4/4 - (6)(0.375)p^5 = 0. \Rightarrow p =$ **0.555**.
Comment: Similar to Q. 5.129 (CAS3L, 11/08, Q.6) in "Mahler's Guide to Statistics."
A graph of the likelihood as a function of p:



While these solutions are believed to be correct, anyone can make a mistake.
If you believe you've found something that may be wrong, send any corrections or comments to:
Howard Mahler, Email: hmahler@mac.com