CA – Exam 5

Exam Results

Nov 8th, 2024 at 7:05 PM

Back to Practice

Score 29/45	Difficulty	Earned Level	Total Time
64%	5.7	5.2 +0.0	4:56:32 of 4:00:00

Section Review

1 Probability Models	8/12	67%	1:53:26	27%
2 Statistics	6/9	67%	57:23	20%
3 Extended Linear Models	15/24	63%	2:05:43	53%

1/1 55% 1.1 7.0 6:01 3:12

1
2
3
4
5
6
7
8
9
10
11
12
13

You want to simulate a random variable U with a uniform distribution on $[0, 1]$. You are given a random variable X, which has a uniform distribution on $[-1, 1]$. You are considering the following three transformations of X:

I. $U = \frac{X+1}{2}$

II. $U = |X|$

III. $U = X^2$

Determine which of the above transformations result in U having a uniform distribution on $[0, 1]$.

3% A None of them

55% I and II only

9% C I and III only

9% D II and III only

24% E The correct answer is not given by (A), (B), (C), or (D).

1
2
3
4
5
6
7
8
9
10
11
12
13

For a collision coverage in an auto policy, loss amounts, X, are modeled using a Pareto distribution with cumulative distribution function:

$$F_X(x) = 1 - \left(\frac{\theta}{x+\theta}\right)^3$$

To study extreme losses, you calculate the expected value of the top 5% of losses, which is 10,000.

Calculate θ.

Incorrect Answer

9% A Less than 3,000

7% At least 3,000, but less than 3,250

54% At least 3,250, but less than 3,500

10% D At least 3,500, but less than 3,750

20% E At least 3,750

1
2
3
4
5
6
7
8
9
10
11
12
13

You are given the following information about waiting times at a subway station:

- Subway trains arrive at a Poisson rate of 20 per hour.
- 30% of the trains are Express and 70% are Local.
- The arrival times of each train are independent.
- An Express train gets you to work in 18 minutes, and a Local train gets you there in 30 minutes.
- You always take the first train to arrive and you get to the office in X_1 minutes from the time you arrive at the subway station.
- Your coworker always take the first Express train to arrive and he gets to the office in X_2 minutes from the time he arrives at the subway station.

Calculate the expected value of $X_1 - X_2$.

Incorrect Answer

5% A Less than -2.0

13% At least -2.0, but less than -1.0

8% C At least -1.0, but less than 0.0

6% D At least 0.0, but less than 1.0

67% At least 1.0

1
2
3
4
5
6
7
8
9
10
11
12
13

You are given the following:

- A Poisson process with $\lambda = 1.2$ has its 4th event occur at time 10.
- This process began operating at time 0.

Calculate the expected value of the time at which the 3rd event occurred.

Incorrect Answer

1% A Less than 2

23% At least 2, but less than 4

7% C At least 4, but less than 6

63% At least 6, but less than 8

6% E At least 8

1
2
3
4
5
6
7
8
9
10
11
12
13

You are given the following information:

- Lucy finds coins at a Poisson rate of 1 coin per 10 minutes.
- The denominations are randomly distributed as follows:
  - 65% of the coins are worth 1 each;
  - 20% of the coins are worth 5 each;
  - 15% of the coins are worth 10 each.

Calculate the probability that in the first 30 minutes she finds at least 1 coin worth 10 each and in the first hour finds at least 2 coins worth 10 each.

27% A Less than 0.165

6% B At least 0.165, but less than 0.175

50% C At least 0.175, but less than 0.185

3% D At least 0.185, but less than 0.195

14% E At least 0.195

1
2
3
4
5
6
7
8
9
10
11
12
13

The number of tow trucks needed for car accidents in a city follows a compound Poisson process. You are given:

- The expected number of accidents per hour $= \begin{cases} 5 & \text{for 7am to 7pm} \\ 2 & \text{for 7pm to 7am} \end{cases}$
- The number of tow trucks needed per accident follows the Poisson distribution with a mean of 1.5.
- The expected number of accidents per hour follows the Poisson distribution.

Calculate the probability that more than 150 tow trucks will be needed in a 24 hour period using a normal approximation.

5% A Less than 1%

23% B At least 1%, but less than 5%

52% At least 5%, but less than 10%

15% D At least 10%, but less than 20%

5% E At least 20%

1
2
3
4
5
6
7
8
9
10
11
12
13

You are given the following information about a system of four components:

- The minimal path sets are $A_1 = \{1, 2\}$, $A_2 = \{2, 3\}$, and $A_3 = \{1, 4\}$.
- All components in the system are independent.
- Components 1 and 3 have a reliability of 0.8.
- Components 2 and 4 have a reliability of 0.75.

What is the reliability of the system?

6% A Less than 0.84

10% B At least 0.84, but less than 0.86

57% C At least 0.86, but less than 0.88

3% D At least 0.88, but less than 0.90

24% E At least 0.90

1/1 53% 1.5 6.8 13:09 4:24

1
2
3
4
5
6
7
8
9
10
11
12
13

You are given the minimal path sets of a system:

$$\{1, 2, 5\} \quad \{1, 3, 4\} \quad \{2, 3, 5\} \quad \{3, 4, 5\}$$

Determine the number of minimal cut sets this system has.

8% A Fewer than 4

10% B 4

16% C 5

53% D 6

13% E Greater than 6

1
2
3
4
5
6
7
8
9
10
11
12
13

You are given the following information about a system with four components:

- The minimal cut sets are $C_1 = \{1, 2\}$, $C_2 = \{3\}$, and $C_3 = \{4\}$.
- All components in the system are independent.
- Each component has a lifetime that is exponentially distributed with a mean of 2 years.

What is the probability that the system is functioning after one year of operation?

11% A Less than 0.25

5% B At least 0.25, but less than 0.30

68% C At least 0.30, but less than 0.35

5% D At least 0.35, but less than 0.40

10% E At least 0.40

0/1 66% 1.6 5.0 3:38 4:01

1
2
3
4
5
6
7
8
9
10
11
12
13

A coin is chosen at random from a hat containing two coins: one coin is a fair coin, while the other coin shows "head" with probability 0.6. Once the coin is chosen, a gambler who initially has 5 units of wealth flips the chosen coin repeatedly according to the following procedure:

- The gambler wins 1 unit if the coin shows "head".
- The gambler loses 1 unit if the coin shows "tail".
- The gambler will stop playing when he either has lost all of his units or he reaches 20 units.

Calculate the probability that the gambler reaches 20 units of wealth.

Incorrect Answer

4% A Less than 0.54

66% At least 0.54, but less than 0.58

8% C At least 0.58, but less than 0.62

15% At least 0.62, but less than 0.66

7% E At least 0.66

1
2
3
4
5
6
7
8
9
10
11
12
13

You are given the following probability transition matrix for a Markov chain with three states labeled 0, 1, and 2:

$$P = \begin{bmatrix} 0.2 & 0.3 & 0.5 \\ 0.1 & 0.5 & 0.4 \\ 0.0 & 0.2 & 0.8 \end{bmatrix}$$

At time $t = 0$, the Markov chain is in state 0.

Calculate the expected number of steps needed to return to state 0.

9% A Less than 20

7% B At least 20, but less than 22

12% C At least 22, but less than 24

5% D At least 24, but less than 26

67% At least 26

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15

You are given the following information for a policyholder age 65:

- A 3-year term insurance policy on (65) provides for a death benefit of 1,000 payable at the end of the year of death.
- This policy is purchased by a single premium, P, at time 0.
- If (65) lives to age 68, the single premium is returned without interest.
- Mortality rates are:

x	q_x
65	0.15
66	0.20
67	0.25

- $i = 0.10$

Calculate P using the equivalence principle.

14% A Less than 640

3% B At least 640, but less than 650

67% ✓ At least 650, but less than 660

2% D At least 660, but less than 670

14% E At least 670

1
2
3
4
5
6
7
8
9
10
11
12
13

The following losses (in thousands) are observed for an insurance policy:

$$5 \quad 5 \quad 12 \quad 16 \quad 22 \quad 23 \quad 25 \quad 25 \quad 30 \quad 30$$

Calculate the kernel estimate of $F(21)$ using a triangular kernel with bandwidth 5.

17% A Less than 0.45

58% At least 0.45, but less than 0.46

6% C At least 0.46, but less than 0.47

4% D At least 0.47, but less than 0.48

14% E At least 0.48

1
2
3
4
5
6
7
8
9
10
11
12
13

You are given the following:

- A random variable, X, is uniformly distributed on the interval $(0, \theta)$.
- θ is unknown.
- For a random sample of size n, an estimate of θ is given by:

$$\hat{\theta} = \frac{2}{n}\sum_{i=1}^{n} X_i$$

Calculate $\mathrm{Var}\left[\hat{\theta}\right]$.

9% A $\frac{4\theta^2}{3n}$

15% B $\frac{\theta^2}{3n^2}$

14% C $\frac{\theta^2}{12}$

7% D $\frac{\theta^2}{6}$

56% $\frac{\theta^2}{3n}$

1/1 71% 2.2 3.4 5:09 3:06

9
10
11
12
13
14
15
16
17
18
19
20
21

Let $Y_1, Y_2, \ldots, Y_n$ be a random sample of size n from a distribution with the probability density function.

$$f(y \mid \theta) = \begin{cases} \frac{2y}{\theta} e^{-(y^2/\theta)}, & 0 \leq y < \infty \\ 0, & \text{elsewhere} \end{cases}$$

Determine which of the following formulas is a sufficient statistic for θ using the factorization criterion.

12% A $\sum_{i=1}^{n} Y_i$

71% B $\sum_{i=1}^{n} Y_i^2$

3% C $\sum_{i=1}^{n} Y_i^4$

5% D $\sum_{i=1}^{n} e^{Y_i}$

8% E The distribution has no sufficient statistic for θ.

1/1 53% 2.3 6.5 :59 2:16

9
10
11
12
13
14
15
16
17
18
19
20
21

Determine which of the following statements is/are true about hypothesis testing.

I. Increasing the significance level would increase the probability of a Type I error.
II. Decreasing the significance level would increase the probability of a Type II error.
III. Increasing the significance level would decrease the power of a test.

4% A None

53% I and II only

22% C I and III only

8% D II and III only

14% E The answer is not given by (A), (B), (C), or (D).

9
10
11
12
13
14
15
16
17
18
19
20
21

$X_1, ..., X_{10}$ is a random sample drawn from a uniform distribution on the range $[0, \theta]$.

Perform the hypothesis test:

$$H_0 : \theta = 1 \quad \text{vs.} \quad H_A : \theta = 1.1$$

The rejection region for this hypothesis test is $RR : \{X_{(9)} > 0.9\}$, where $X_{(9)}$ is the second largest value observed in the sample.

Calculate the probability of a Type I error and the probability of a Type II error for this hypothesis test.

Probability of a Type I error

0.26 (round to the nearest 0.01)

Correct Answer: 0.26

Probability of a Type II error

0.43 (round to the nearest 0.01)

Correct Answer: 0.43

9
10
11
12
13
14
15
16
17
18
19
20
21

Let X be a random variable from a distribution with probability mass function $f(x|\theta)$ (PMF), for $x = 1, 2, \ldots, 7$, which depends on the parameter θ. A single observation from this distribution is being used to test $H_0 : \theta = 2$ vs. $H_A : \theta = 3$. The PMFs of X for the two parameter values $\theta = 2$ and $\theta = 3$ are given below:

x	1	2	3	4	5	6	7
$f(x;\theta=2)$	0.01	0.02	0.02	0.01	0.01	0.01	0.92
$f(x;\theta=3)$	0.01	0.10	0.04	0.06	0.08	0.03	0.68

Your colleague proposes using the rejection region $RR : \{X = 1, 3, 5, \text{or } 6\}$. Let α and β_1 be the probabilities of Type I and Type II errors, respectively, associated with this test. Let β_2 be the probability of a Type II error associated with the most powerful test of these hypotheses at significance level α.

Calculate the absolute difference of β_1 and β_2.

Incorrect Answer

11% A. Less than 0.06

16% B. At least 0.06, but less than 0.09

57% C. At least 0.09, but less than 0.12

7% D. At least 0.12, but less than 0.15

9% E. At least 0.15

9
10
11
12
13
14
15
16
17
18
19
20
21
22
23

You are given:

- The following sample losses:

$$4 \quad 6 \quad 7 \quad 10 \quad 11$$

- The inverse gamma distribution with $\alpha = 2$ is fitted to this data using the maximum likelihood method.
- The cumulative distribution function of an inverse gamma with $\alpha = 2$ is:

$$F(x) = e^{-\theta/x}\left(\frac{\theta}{x} + 1\right)$$

Calculate the Kolmogorov-Smirnov test statistic for the fitted distribution.

Incorrect Answer

3% A Less than 0.32

6% B At least 0.32, but less than 0.34

62% At least 0.34, but less than 0.36

4% D At least 0.36, but less than 0.38

24% At least 0.38

9
10
11
12
13
14
15
16
17
18
19
20
21

You are given a sample of six losses:

$$50 \quad 60 \quad 75 \quad 88 \quad 94 \quad 100$$

The losses follow the normal distribution with mean μ.

Calculate the upper bound of the 95% symmetric confidence interval for μ.

Incorrect Answer

24% Less than 94

4% B At least 94, but less than 96

4% C At least 96, but less than 98

53% At least 98, but less than 100

15% E At least 100

1/1 | 62% | 2.7 | 6.0 | 8:01 | 4:33

9
10
11
12
13
14
15
16
17
18
19
20
21

For a general liability policy, loss amounts, Y, follow the exponential distribution with probability density function:

$$f(y) = \frac{1}{\theta} e^{-y/\theta}, \quad \theta = 1{,}000, \quad 0 < y$$

For reinsurance purposes we are interested in the distribution of the median loss amount in a random sample of size 3, which is denoted by $Y_{(2)}$.

Calculate the probability that $Y_{(2)}$ is less than 2,000.

15% A Less than 0.78

3% B At least 0.78, but less than 0.83

7% C At least 0.83, but less than 0.88

13% D At least 0.88, but less than 0.93

62% At least 0.93

9
10
11
12
13
14
15
16
17
18
19
20
21

Stanley owns a department store that offers a wide selection of products. He plans to boost sales using a marketing campaign. He contracts the tasks of conducting the campaign to Angela, a local marketing analyst.

Angela uses a clustering method to group shoppers based on their shopping habits. She does this so that similar advertisements can be targeted to shoppers with similar shopping patterns.

Which of the following best describes the type of problem Angela is solving?

1% A Regression

34% B Classification

5% C Supervised

59% Unsupervised

0% E The correct answer is not given by (A), (B), (C), or (D).

17
18
19
20
21
22
23
24
25
26
27
28
29

You are given a random sample:

$$\{23, 1, 20, 5, 18\}$$

and would like to calculate the standard error of an estimate of the population median using a bootstrap procedure. The median is estimated using the third order statistic.

Your colleague disagrees with the use of the third order statistic. She instead uses the average of the second and the fourth order statistics as the estimate.

Let α be the maximum possible value of your standard error estimate, and β be maximum possible value of your colleague's standard error estimate. Both of you use only two bootstrap data sets.

Calculate the absolute difference of α and β.

53% Less than 0.5

10% B At least 0.5, but less than 1.0

18% C At least 1.0, but less than 1.5

7% D At least 1.5, but less than 2.0

11% E At least 2.0

17
18
19
20
21
22
23
24
25
26
27
28
29

You are given:

- A multiple linear regression model was fit to 20 observations.
- There are five explanatory variables with fitted values for β_1 through β_5.
- The following summarizes the fitted coefficients excluding the intercept:

	Point Estimate	Standard Error	t-Statistic
β_1	0.020000	0.012000	1.661
β_2	-0.004950	0.008750	-0.565
β_3	0.216000	0.043200	5.000
β_4	-0.034600	0.115000	-0.301
β_5	-0.000294	0.000141	-2.090

Determine the number of coefficients in the table above for the five explanatory variables that are not statistically different from zero at a significance level of $\alpha = 10\%$, based on a two-tailed test.

4% A 1

26% B 2

63% 3

6% D 4

1% E 5

17
18
19
20
21
22
23
24
25
26
27
28
29

You are given the following information about a linear model:

- $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$
- Three bootstrap samples were drawn from the data

Bootstrap Sample	Estimate for β_0
1	10.525
2	2.499
3	16.456

Calculate the standard error of these bootstrap estimates of β_0.

11% A Less than 6

70% ✓ At least 6, but less than 8

15% C At least 8, but less than 10

2% D At least 10, but less than 12

2% E At least 12

1/1 49% 3.3 7.3 7:21 5:46

17
18
19
20
21
22
23
24
25
26
27
28
29

You are given the following dataset:

y	x_1	x_2
7.3	8.2	4.1
6.8	9.5	3.7
5.4	1.6	2.9

You perform an ordinary least squares regression using both predictors plus an intercept to predict the response.

Determine which of the following conclusions can be made on the resulting model.

I. The coefficient of determination is $R^2 = 1$.

II. The estimated coefficient for the intercept is $b_0 = 1.355$.

III. The residual for the first observation is 0.8.

16% A None

49% ✓ I and II only

8% C I and III only

18% D II and III only

9% E The answer is not given by (A), (B), (C), or (D)

17
18
19
20
21
22
23
24
25
26
27
28
29

You are given the following information for a one-way ANOVA:

- The ANOVA F-statistic for a sample is 3.75.
- The mean square for treatments is 606.69 with 2 degrees of freedom.
- The total sum of squares is 2,670.25.
- n is the number of observations in the sample.

Calculate n.

68% Less than 14

7% B At least 14, but less than 17

8% C At least 17, but less than 20

2% D At least 20, but less than 23

15% E At least 23

17
18
19
20
21
22
23
24
25
26
27
28
29

You are given the partial information on a one-factor ANOVA regression with 30 observations:

Source	Degrees of Freedom	Sum of Squares
Mean	1	76,246
Treatments		
Residuals	22	
Total		98,125

The F statistic to test the significance of the treatments is 3.3669.

Calculate the residual sum of squares.

6% A Less than 10,000

72% B At least 10,000, but less than 12,500

7% C At least 12,500, but less than 15,000

4% D At least 15,000, but less than 17,500

11% E At least 17,500

17
18
19
20
21
22
23
24
25
26
27
28
29

Determine which of the following statements is/are true.

I. The leverage for each observation in a linear model must be between $\frac{1}{n}$ and 1.

II. The leverage for each observation in a linear model must sum to the number of explanatory variables.

III. If an explanatory variable is uncorrelated with all other explanatory variables, the corresponding variance inflation factor would be zero.

Incorrect Answer

62% I only

8% B II only

8% C III only

10% I, II, and III

12% E The correct answer isn't given by A, B, C, or D.

17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36

You are given the following information for a fitted model:

Response variable	Rating
Response distribution	Normal
Link	Identity
Residual Std. Error	7.139

Parameter	df	$\hat{\beta}$
Intercept	1	11.011
Complaints	1	0.692
Privileges	1	-0.104
Learning	1	0.249
Raises	1	-0.033
Critical	1	0.015

- The first record in the data has the following values:

ID	Rating	Complaints	Privileges	Learning	Raises	Critical
1	43	51	30	39	61	92

- The corresponding hat matrix diagonal value is 0.3234.

Calculate the DFITS for the observation above.

Incorrect Answer

5% Less than -3

11% B At least -3, but less than -2

67% At least -2, but less than -1

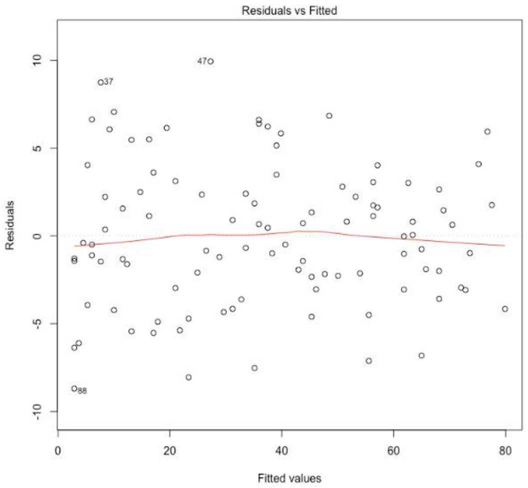
10% D At least -1, but less than 0

8% E At least 0

10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40

A simple linear regression is performed on a dataset. The summary output and the residual plot are shown below.

	Estimate	Std. Error
(Intercept)	1.348	0.752
x	0.785	0.014



You then use the bootstrap approach to estimate the standard error of the coefficient estimates, $\mathrm{SE}\left(\hat{\beta}_0\right)$ and $\mathrm{SE}\left(\hat{\beta}_1\right)$.

Determine which of the following conclusions can be made about those estimates based on the information above.

4% A The bootstrap estimate for $\mathrm{SE}\left(\hat{\beta}_0\right)$ is significantly lower than the estimate produced from the regression.

15% B The bootstrap estimate for $\mathrm{SE}\left(\hat{\beta}_1\right)$ is significantly higher than the estimate produced from the regression.

15% C The bootstrap estimate for $\mathrm{SE}\left(\hat{\beta}_1\right)$ is significantly lower than the estimate produced from the regression.

61% D The bootstrap estimate for $\mathrm{SE}\left(\hat{\beta}_1\right)$ is approximately the same as the estimate produced from the regression.

5% E The correct answer is not given by (A), (B), (C), or (D).

1/1 66% 3.6 5.1 2:05 2:21

10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29

An actuary has a dataset with one dependent variable, Y, and five independent variables (X_1, X_2, X_3, X_4, X_5). She is trying to determine which subset of the predictors best fits the data, and is using a Forward Stepwise Selection procedure with no stopping rule. Below is a subset of the potential models:

Model	Dependent variable	RSS	Independent variable	p-value
1	Y	9,823	X_1	0.0430
			X_2	0.0096
2	Y	7,070	X_1	0.0464
			X_2	0.0183
			X_3	0.0456
3	Y	6,678	X_1	0.0412
			X_2	0.0138
			X_4	0.0254
4	Y	4,800	X_1	0.0444
			X_2	0.0548
			X_5	0.0254
5	Y	3,475	X_1	0.0333
			X_2	0.0214
			X_3	0.0098
			X_4	0.0274
			X_5	0.0076

The procedure just selected Model 1 as the new candidate model.

Determine which of the following independent variable(s) will be added to the model in the next iteration of this procedure.

3% A No variables will be added

14% B X_3 only

10% C X_4 only

66% D X_5 only

8% E X_3, X_4 and X_5

23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42

You are given several models fitted to a data set with 100 observations. The following predictors plus an intercept are used:

- Income, Limit, and Rating are continuous variables.
- Education is a categorical variable with 16 levels.
- Married is a categorical variable with two levels.

The coefficient of determination for each model is given in the table below.

Model	Predictors	R^2
I	Income, Limit	0.8711
II	Income, Limit, Education	0.8762
III	Income, Limit, Rating	0.8762
IV	Income, Limit, Rating, Married	0.8772
V	Income, Limit, Rating, Education	0.8812

Determine which model is best based on adjusted R^2.

Incorrect Answer

5% Model I

3% B Model II

64% Model III

5% D Model IV

23% E Model V

23
24
25
26
27
28
29
30
31
32
33
34
35
36
37

Determine which of the following statements is/are true about subset selection.

I. Best subset selection results in a nested set of best models, each with different number of predictors.

II. Residual sum of squares is a suitable metric for selecting the best model among models with different number of predictors.

III. Forward stepwise selection cannot be used in high-dimensional settings.

Incorrect Answer

50% None

13% B I and II only

18% I and III only

6% D II and III only

14% E The answer is not given by (A), (B), (C), or (D)

23
24
25
26
27
28
29
30
31
32
33
34
35
36
37

Determine which of the following statements correctly demonstrates a difference between ridge regression and lasso regression. (select all that apply)

Incorrect Answer(s)

31% Both of them are regularized methods.

12% B Lasso regression has a tuning parameter that controls the relative impact of the regression coefficients, while ridge regression has a budget parameter.

65% Ridge regression uses an ℓ_2 penalty, while lasso regression uses an ℓ_1 penalty.

81% Both methods shrink coefficients towards zero, but lasso can force some of the estimates to be exactly equal to zero.

11% E Lasso regression tends to outperform ridge regression in terms of bias, variance, and MSE.

23
24
25
26
27
28
29
30
31
32
33
34
35
36
37

You have used least squares to fit a regression model with two continuous predictors. The resulting coefficient estimates are:

$$\hat{\beta}_0 = 3.4, \qquad \hat{\beta}_1 = 2.8, \qquad \hat{\beta}_2 = 6.5$$

When you fit a ridge regression model to the same data, the resulting coefficient estimates are:

$$\hat{\beta}_0^R = 4.1, \qquad \hat{\beta}_1^R = 2.1, \qquad \hat{\beta}_2^R = x$$

Calculate the maximum possible value for $\hat{\beta}_2^R$.

26% A Less than 6.4

9% B At least 6.4, but less than 6.5

13% C At least 6.5, but less than 6.6

2% D At least 6.6, but less than 6.7

51% E At least 6.7

23
24
25
26
27
28
29
30
31
32
33
34
35
36
37

Consider the following statements regarding the tuning parameter λ in the lasso model-fitting procedure:

I. As λ increases, the number of predictors in the chosen model will increase.

II. As λ increases, the squared bias of the parameters in the chosen model will increase.

III. As λ increases, the variance of the predictions made by the chosen model will increase.

Determine which of the above statements are true.

2% A I only

73% II only

18% C III only

2% D I, II, and III

5% E The answer is not given by (A), (B), (C) or (D).

23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42

An actuary fits two GLMs, M_1 and M_2, to the same data in order to predict the probability of a customer purchasing an automobile insurance product. You are given the following information about each model:

Model	Explanatory Variables Included in Model	Degrees of Freedom Used	Log Likelihood
M_1	• Offered Price • Number of Vehicles • Age of Primary Insured • Prior Insurance Carrier	10	−11,565
M_2	• Offered Price • Number of Vehicles • Age of Primary Insured • Gender of Primary Insured • Credit Score of Primary Insured	8	−11,562

The actuary wants to evaluate which of the two models is superior.

Determine which of the following is the best course of action for the actuary to take.

16% A Perform a likelihood ratio test

4% B Compute the F-statistic and perform an F-test

5% C Compute and compare the deviances of the two models

73% Compute and compare the AIC statistics of the two models

2% E Compute the Chi-squared statistic and perform a Chi-squared test

1/1 57% 3.8 6.8 1:27 1:54

22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37

One example of a generalized linear model is the normal linear model, where the link function is the identity function. If there is more than one explanatory variable, this may also be known as multiple linear regression. One potential concern with such models is collinearity.

Assume that a multiple linear regression model contains collinearity. Then, which of the following statements is FALSE?

10% A The columns of the design matrix, $\mathbf{X}$, may be nearly linearly dependent.

5% B It may be difficult to choose the best subset of explanatory variables.

22% C When choosing the best subset of explanatory variables, one may rely more on substantive knowledge of the model and variables than statistical grounds.

6% D The VIF (or Variance Inflation Factor) will be higher for models with more severe collinearity than for models with just mild collinearity.

57% The VIF (or Variance Inflation Factor) $= 1/\left(1 - R^2_{(j)}\right)$, where $R^2_{(j)}$ is obtained from regressing the response variable Y on the explanatory variable X_j.

29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44

The health status of an employee after a work-related injury (1 = no disability, 2 = temporary disability, 3 = permanent disability, and 4 = death) is modeled by a proportional-odds cumulative logit model.

The following two explanatory variables are used in the model:

- X_1: The age of the employee
- X_2: The number of years the employee has been with the company

Employee A is a 50-year-old employee who has worked with the company for 10 years, while Employee B is a 25-year-old employee who has worked with the company for 5 years.

It is given that the estimated coefficient for X_1, $\hat{\beta}_1$, is 0.02 and the estimated coefficient for X_2, $\hat{\beta}_2$, is 0.3.

The probability that Employee A has no disability is 0.32.

Calculate the probability that Employee B has no disability.

Incorrect Answer

64% Less than 0.1

10% At least 0.1, but less than 0.2

10% C At least 0.2, but less than 0.3

5% D At least 0.3, but less than 0.4

12% E At least 0.4

29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44

The health status of an employee (1 = no disability, 2 = temporary disability, 3 = permanent disability, and 4 = death) is modeled by a proportional-odds cumulative logit model. Only one explanatory variable is used in this model: if the employee experienced a work-related accident in the last ten days.

The following information are given:

- The probability of no disability for an employee who experienced a work-related accident is 0.20.
- The probability of temporary disability for an employee who experienced a work-related accident is 0.25.
- The probability of permanent disability for an employee who experienced a work-related accident is 0.30.
- The probability of no disability or temporary disability for an employee who did not experience a work-related accident is 0.60.

Calculate the probability of permanent disability for an employee who did not experience a work-related accident in the last ten days.

10% A Less than 0.2

59% At least 0.2, but less than 0.4

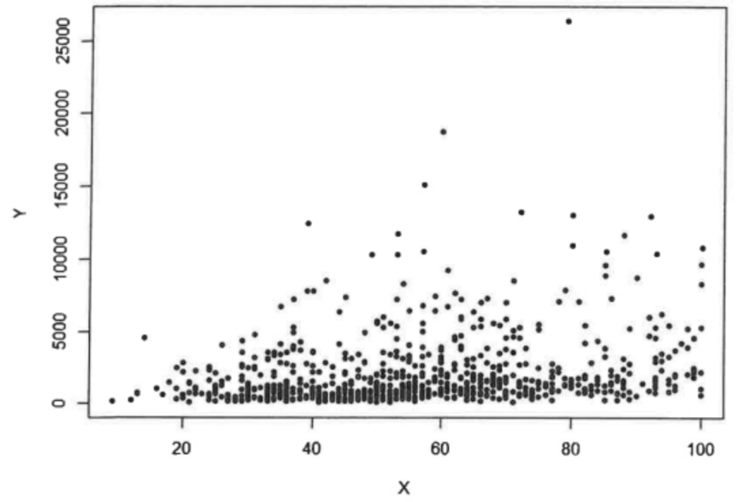
19% C At least 0.4, but less than 0.6

9% D At least 0.6, but less than 0.8

3% E At least 0.8

29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44

You are given the following data and want to perform a piecewise polynomial regression with knots at $X = \{20, 40, 60, 80\}$.



Determine which of the following models will use the most degrees of freedom.

11% A Cubic Spline

2% B Linear Spline

10% C Natural Cubic Spline

75% Piecewise Cubic Regression

3% E Piecewise Linear Regression

23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42

You are given:

- A set of 20 observations is fitted to a quadratic spline with one knot at x.

$$Y_i = \begin{cases} 10 + \beta_{11} x_i + \beta_{21} x_i^2 + \varepsilon_i, & x_i < x \\ 370 + \beta_{12} x_i + \beta_{22} x_i^2 + \varepsilon_i, & x_i \geq x \end{cases}$$

- At x, the fitted graph and its first derivative are continuous.
- Some observations and their fitted values are as follows:

i	x_i	$\hat{y}_i$
1	4.5	88.750
2	5.7	102.910
⋮	⋮	⋮
19	26.3	101.214
20	27.0	105.400

Calculate x.

Incorrect Answer

8% A Less than 12.5

9% B At least 12.5, but less than 13.5

16% C At least 13.5, but less than 14.5

62% D At least 14.5, but less than 15.5

6% E At least 15.5

23
24
25
26
27
28
29
30
31
32
33
34
35
36
37

You are considering fitting the following three different spline models to a data set:

- Model I: A linear spline model with k knots.
- Model II: A cubic spline model with k knots.
- Model III: A natural cubic spline model with k total knots, i.e. $k-2$ interior knots.

Determine the ordering of the number of degrees of freedom used for these models.

Incorrect Answer

5% A I > II > III

28% II > III > I

59% II > I > III

7% D III > II > I

2% E None of the above.

0/1 49% 3.11 7.6 19:49 5:15

29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44

You have fit a quadratic spline model to a set of data, the model having one knot at $x = 12$.

The fitted model has yielded the following predicted $\hat{y}_i$ values from the given x_i points:

$\hat{y}_i$	x_i
4.2	3
8.3	6
9.4	9
9.9	15
23.9	18
49.5	21

Calculate the predicted value of the model at $x = 10$.

Incorrect Answer

49% Less than 9.2

6% B At least 9.2, but less than 9.4

26% C At least 9.4, but less than 9.6

13% At least 9.6, but less than 9.8

5% E At least 9.8