VCE & PDF
Pass4itSure.com

A00-240^{Q&As}

SAS Certified Statistical Business Analyst Using SAS 9: Regression and Modeling Credential

Pass SASInstitute A00-240 Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

https://www.pass4itsure.com/a00-240.html

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by SASInstitute Official Exam Center

Instant Download After Purchase

100% Money Back Guarantee

365 Days Free Update

800,000+ Satisfied Customers

QUESTION 1

This question will ask you to provide a missing option. Given the following SAS program:

```
proc corr data = MYDATA <insert option here> ;
   var x1 x2 x3 x4 x5;
   with Target;
run;
```

What option must be added to the program to obtain a data set containing Pearson statistics?

A. OUTPUT=estimates

B. OUTP=estimates

C. OUTSTAT=estimates

D. OUTCORR=estimates

Correct Answer: B

QUESTION 2

A marketing manager attempts to determine those customers most likely to purchase additional products as the result of a nation-wide marketing campaign.

The manager possesses a historical dataset (CAMPAIGN) of a similar campaign from last year.

It has the following characteristics:

1.

 Target variable Respond (0, 1)

2.

 Continuous predictor Income

3.

 Categorical predictor Homeowner(Y, N) Which SAS program performs this analysis?

```
C A.  proc logistic data=MYDIR.CAMPAIGN descending;
         class Homeowner;
         model Respond = Income Homeowner;
      run;


C B.  proc logistic data = MYDIR.CAMPAIGN descending;
         by Homeowner;
         model Respond = Income Homeowner;
      run;


C C.  proc logistic data = MYDIR.CAMPAIGN descending;
         model Respond = Income Homeowner;
      run;


C D.  proc logistic data = MYDIR.CAMPAIGN descending;
         class Income Homeowner;
         model Respond = Income Homeowner;
      run;
```

A. Option A

B. Option B

C. Option C

D. Option D

Correct Answer: A

QUESTION 3

A confusion matrix is created for data that were oversampled due to a rare target. What values are not affected by this oversampling?

A. Sensitivity and PV+

B. Specificity and PV

C. PV+ and PV

D. Sensitivity and Specificity

Correct Answer: D

QUESTION 4

Suppose training data are oversampled in the event group to make the number of events and non-events roughly equal. A logistic regression is run and the probabilities are output to a data set NEW and given the variable name PE. A

decision rule considered is, "Classify data as an event if probability is greater than 0.5." Also the data set NEW contains a variable TG that indicates whether there is an event (1=Event, 0= No event).

The following SAS program was used.

```
data NEW;
    set NEW;
    Solicit = PE > .5;
run;
proc means data=NEW(where = (TG=1)) mean;
    var Solicit;
run;
```
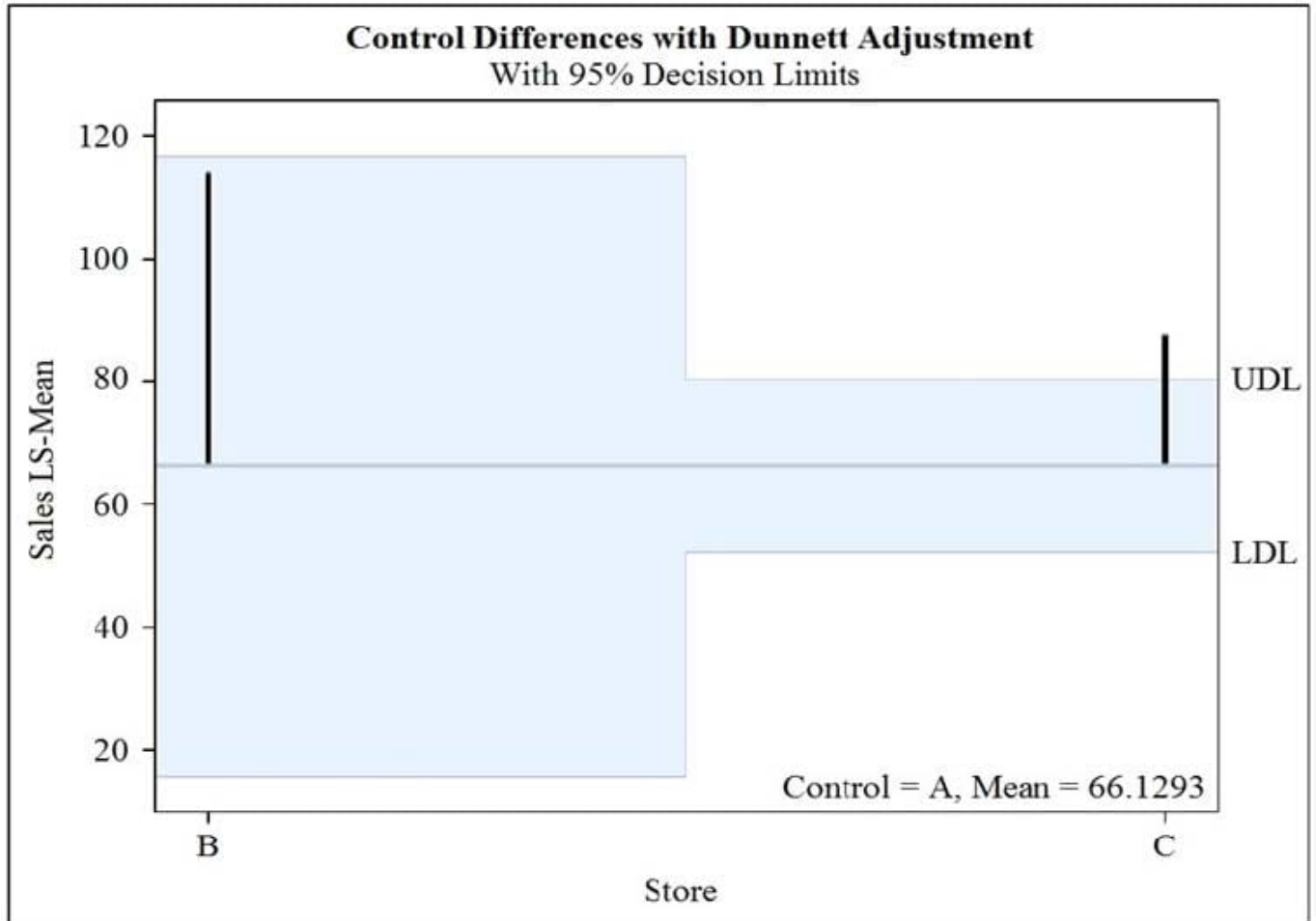
What does this program calculate?

A. Depth

B. Sensitivity

C. Specificity

D. Positive predictive value

Correct Answer: B

QUESTION 5

Refer to the exhibit.

Control Differences with Dunnett Adjustment
With 95% Decision Limits

Control = A, Mean = 66.1293

Which conclusion is justified concerning Sales, comparing stores A, B, and C?

A. Store B is significantly different from store A.

B. Store C is significantly different from Store A.

C. Store B is significantly different from store C.

D. There is no significant difference between stores.

Correct Answer: A

QUESTION 6

The total modeling data has been split into training, validation, and test data. What is the best data to use for model assessment?

A. Training data

B. Total data

C. Test data

D. Validation data

Correct Answer: D

QUESTION 7

Which SAS program will correctly use backward elimination selection criterion within the REG procedure?

```
A.  proc reg data=SASUSER.MLR;
        model y = x1-x10 /selection=backward sls=aic;
    run;
```

```
B.  proc reg data=SASUSER.MLR;
        model y = x1-x10 /selection=backward sls=0.15;
    run;
```

```
C.  proc reg data=SASUSER.MLR;
        model y = x1-x10 /selection=backward sle=cp;
    run;
```

```
D.  proc reg data=SASUSER.MLR;
        model y = x1-x10 /selection=backward sle=all;
    run;
```

A. Option A

B. Option B

C. Option C

D. Option D

Correct Answer: B

QUESTION 8

FILL BLANK

Refer to the confusion matrix:

		Predicted Outcome	
		0	1
Actual Outcome	0	345	155
	1	188	312

An analyst determines that loan defaults occur at the rate of 3% in the overall population. The above confusion matrix is from an oversampled test set (1 = default).

What is the sensitivity adjusted for the population event probability?

Enter your answer in the space below. Round to three decimals (example: n.nnn).

Correct Answer: 0.617

Section: (none)

QUESTION 9

Consider scoring new observations in the SCORE procedure versus the SCORE statement in the LOGISTIC procedure.

Which statement is true?

A. The SCORE statement in the LOGISTIC procedure returns only predicted probabilities, whereas the SCORE procedure returns only predicted logits.

B. The SCORE statement in the LOGISTIC procedure returns only predicted logits, whereas the SCORE procedure returns only predicted probabilities.

C. Unlike the SCORE procedure, the SCORE statement in the LOGISTIC procedure produces both predicted probabilities and predicted logits.

D. The SCORE procedure and the SCORE statement in the LOGISTIC procedure produce the same output.

Correct Answer: A

QUESTION 10

This question will ask you to provide a missing option.

Complete the following syntax to test the homogeneity of variance assumption in the GLM procedure:

means Region / =levene ;

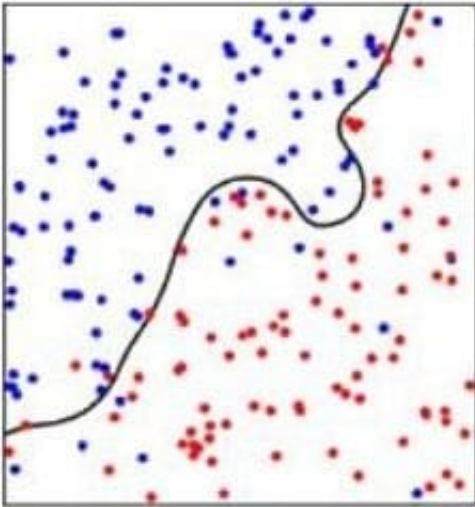A. test

B. adjust

C. var

D. hovtest

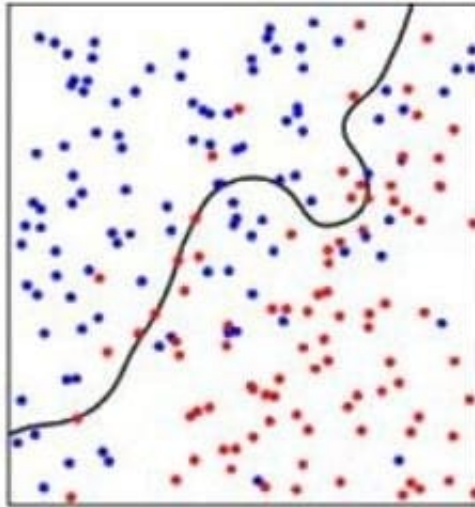Correct Answer: D

QUESTION 11

Refer to the exhibit:

Model A
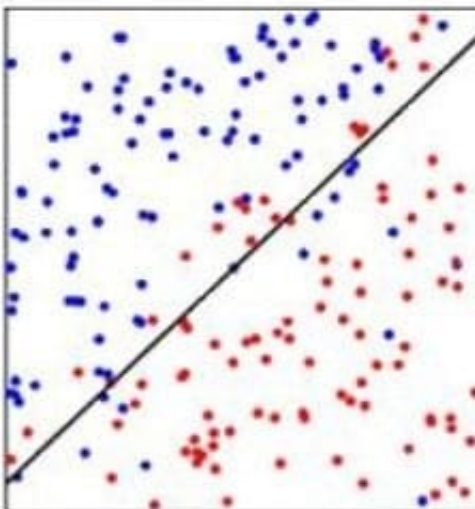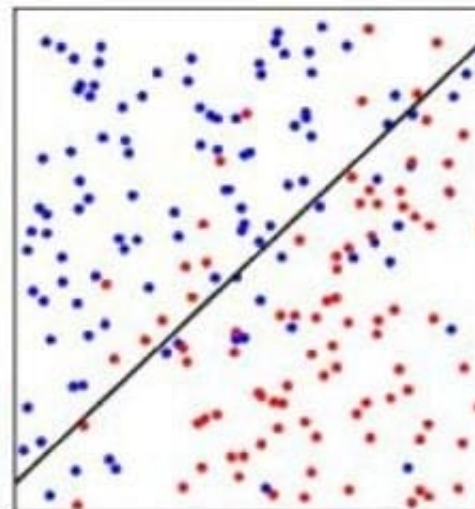

Model B

The plots represent two models, A and B, being fit to the same two data sets, training and validation.

Model A is 90.5% accurate at distinguishing blue from red on the training data and 75.5% accurate at doing the same on validation data. Model B is 83% accurate at distinguishing blue from red on the training data and 78.3% accurate at

doing the same on the validation data.

Which of the two models should be selected and why?

A. Model A. It is more complex with a higher accuracy than model B on training data.

B. Model A. It performs better on the boundary for the training data.

C. Model B. It is more complex with a higher accuracy than model A on validation data.

D. Model B. It is simpler with a higher accuracy than model A on validation data.

Correct Answer: D

QUESTION 12

Which SAS program will divide the original data set into 60% training and 40% validation data sets, stratified by county?

```
A.  proc surveyselect data=SASUSER.DATABASE samprate=0.6 out=sample;
        strata county;
    run;
```

```
B.  proc sort data=SASUSER.DATABASE;
        by county;
    run;
    proc surveyselect data=SASUSER.DATABASE samprate=0.6 out=sample outall;
    run;
```

```
C.  proc sort data=SASUSER.DATABASE;
        by county;
    run;
    proc surveyselect data=SASUSER.DATABASE samprate =0.6 out=sample outall;
        strata county;
    run;
```

```
D.  proc sort data=SASUSER.DATABASE;
        by county;
    run;
    proc surveyselect data=SASUSER.DATABASE samprate =0.6  out=sample;
        strata county;
    run;
```

A. Option A

B. Option B

C. Option C

D. Option D

Correct Answer: C

QUESTION 13

Within PROC GLM, the interaction between the two categorical predictors, Income and Gender, was shown to be significant. An item store was saved from the GLM analysis.

Which statement from PROC PLM would test the significance of Gender within each level of Income and adjust for multiple tests?

A. sliceby Gender / adjust=tukey;

B. slice Income*Gender / sliceby=Gender adjust=tukey;

C. slice Income*Gender / sliceby=Income adjust=tukey;

D. sliceby Income / adjust=tukey;

Correct Answer: D

QUESTION 14

One common approach for predicting rare events in the LOGISTIC procedure is to build a model that disproportionately over-re presents those cases with an event occurring (e.g. a 50-50 event/non-event split). What problem does this present?
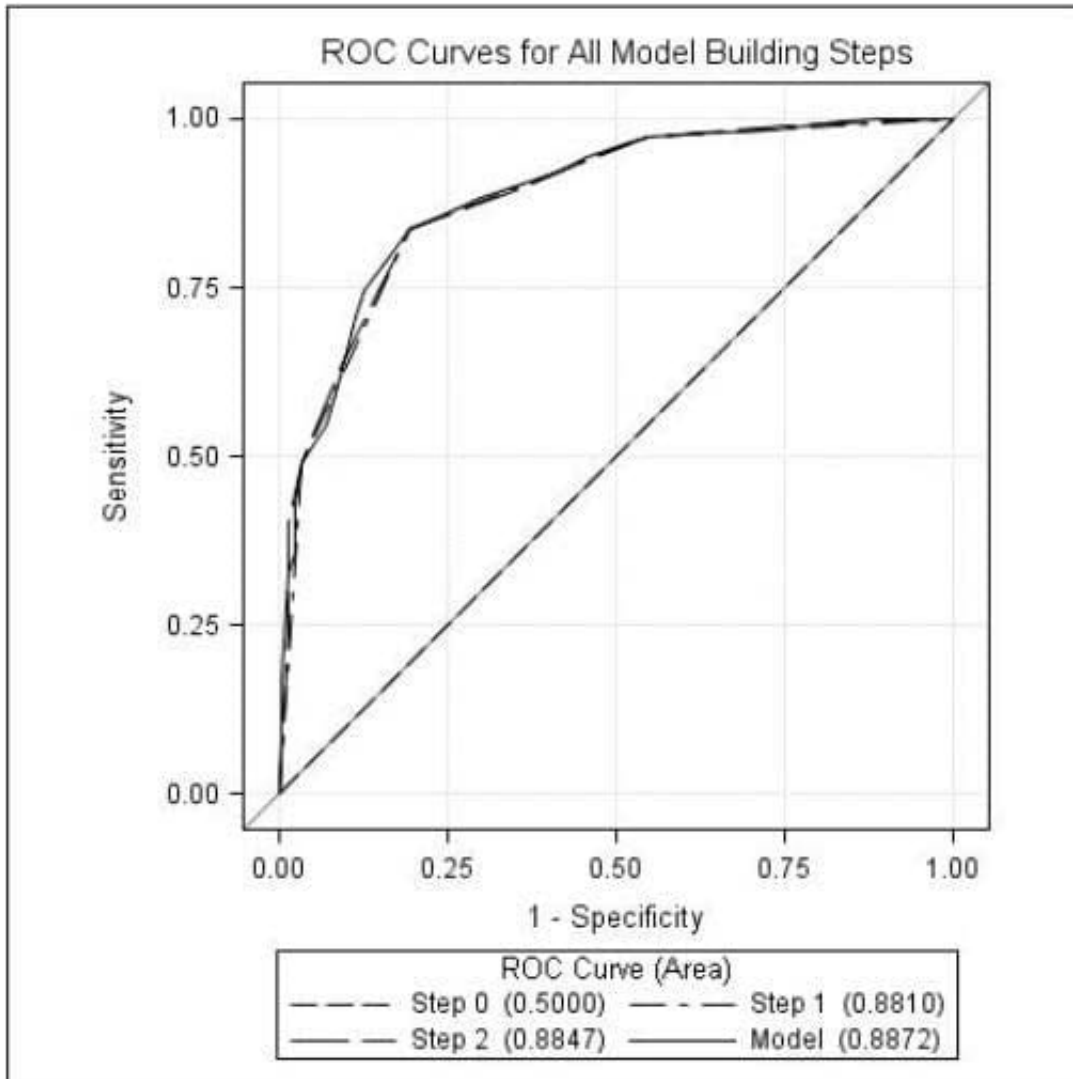
A. All parameter estimates are biased.

B. Only the intercept estimate is biased.

C. Only the non-intercept parameter estimates are biased.

D. Sensitivity estimates are biased.

Correct Answer: B

QUESTION 15

Refer to the exhibit:

An analyst examined logistic regression models for predicting whether a customer would make a purchase. The ROC curve displayed summarizes the models. Using the selected model and the analyst\\'s decision rule, 25% of the customers who did not make a purchase are incorrectly classified as purchasers.

What can be concluded from the graph?

A. About 25% of the customers who did make a purchase are correctly classified as making a purchase.

B. About 50% of the customers who did make a purchase are correctly classified as making a purchase.

C. About 85% of the customers who did make a purchase are correctly classified as making a purchase.

D. About 95% of the customers who did make a purchase are correctly classified as making a purchase.

Correct Answer: C