

Mock Exam 1 - Answers

TOPIC: QUANTITATIVE METHODS

TOTAL POINT VALUE OF THIS QUESTION SET IS 12 POINTS

Austin Clark, CFA, has been asked to analyze White Goods Corporation (WGC), a \$9 billion company that owns a nationwide chain of stores selling appliances and other electronic goods. As part of his analysis of WGC, Clark's supervisor, David Horvath, asks Clark to forecast WGC's 2019 sales using multiple regression analysis. The following model was developed:

$$\text{sales} = 20.1 + 0.001 \text{ GDP} + 1,000.6 \text{ TR} + 0.1 \text{ CC} - 3.2 \text{ PC} - 40.3 \text{ UR}$$

VIF (5.2) (3.1) (5.3) (1.9) (4.7)

Number of observations: 76

Standard error estimate: 15.67

Unadjusted R^2 : 0.96

Regression sum of squares: 412,522

Error sum of squares: 17,188

Independent Variable Descriptions

GDP = gross domestic product

TR = average rate on 5-year US Treasury securities

CC = most recent quarter-end consumer confidence index value

PC = previous year's sales of personal computers

UR = most recent quarter-end unemployment rate

Variable Estimates for 2019

GDP = 8,000

TR = 0.05

CC = 97

PC = 60,000

UR = 0.055

Dave Carington was recently hired by the same investment firm as a quantitative analyst. During Horvath's meeting with Carington, the two discussed the application of big data and machine learning to investing. Carington makes the following comments:

Comment 1: One has to be careful to avoid overfitting in these models. Overfitting occurs when noise in the data is perceived as relationships.

Comment 2: To reduce the risk of Type II error, choose a model with a high recall.

Question 1 of 88

Using his multiple linear regression, Clark's sales forecast for 2019 is *closest* to:

- A) **-\$191,914.** ✓
- B) \$180,502.
- C) \$192,090.

Explanation

2019 sales forecast = $20.1 + 0.001 \times 8,000 + 1,000.6 \times 0.05 + 0.1 \times 97 - 3.2 \times 60,000 - 40.3 \times 0.055 = -\$191,914.$

(Module 1.2, LOS 1.f)

Question 2 of 88

Based on the information provided, which of the following concerns about the model would be *most appropriate*?

- A) Low explanatory power.
- B) Type II error. ✓**
- C) Type I error.

Explanation

Unadjusted R^2 is very high at 96%, so the model explanatory power is not low. VIFs for variables GDP and CC are greater than 5, indicating that multicollinearity may be a problem. GDP and consumer confidence would also be reasonably expected to be highly correlated (in addition to the correlation to other variables included). Multicollinearity would result in Type II error. Type II errors may also be caused by negative serial correlation, but we don't have any supporting evidence for that. Type I errors can be caused by positive serial correlation as well as with conditional heteroskedasticity. Again, there is no evidence presented for that.

(Module 1.3, LOS 1.j)

Question 3 of 88

Regarding Carington's comments:

- A) **both comments are accurate.** ✓
- B) neither comments are accurate.
- C) only one comment is accurate.

Explanation

Comment 1 is correct. Overfitting means training a model to such a degree that noise or randomness is mistaken for patterns and relationships. Comment 2 is also correct. High precision is important when the cost of a Type I error is large, while high recall is needed when the cost of a Type II error is large.

(Module 4.3, LOS 4.c)

Question 4 of 88

What is the F -value that tests the hypothesis that all of the coefficients are equal to zero?

- A) 42.
- B) 101.
- C) **336** ✓

Explanation

The F -value is calculated as (mean regression sum of squares) / (mean squared error)
 $= (412,522 / 5) / (17,188 / 70) = 336$.

(Module 1.2, LOS 1.e)

TOPIC: QUANTITATIVE METHODS

TOTAL POINT VALUE OF THIS QUESTION SET IS 12 POINTS

William Shears, CFA, has been assigned the task of predicting sales for the specialty retail industry. Shears finds that sales have been increasing at a fairly constant rate over time and decides to estimate the linear trend in sales for the industry using quarterly data over the past 15 years, starting with Quarter 1 of 2004 and ending with Quarter 4 of 2018. On January 1, 2019, Shears estimates the following model:

$$\text{sales}_t = b_0 + b_1t + e_t \quad (1)$$

where:

sales = quarterly sales (measured in \$ millions) for the specialty retail industry

b_0 = intercept term

b_1 = slope

t = time variable (quarter number)

e = random error

Exhibit 1 provides the results of the linear trend regression.

Exhibit 1: Linear Trend Regression

	Coefficient	Standard Error
Intercept	10.0	3.50
Trend	16.0	6.55

Shears also estimates an autoregressive model of order one, AR(1), using the changes in quarterly sales data for the industry from the first quarter of 2004 through the fourth quarter of 2018. He obtains the following results for his AR(1) model:

$$\Delta \text{sales}_t = b_0 + b_1 \Delta \text{sales}_{t-1} + e_t$$

The following is found in Exhibit 2.

Exhibit 2: AR(1) Model for Changes in Industry Sales

	Coefficient	Standard Error
Intercept	20.00	2.15