科目：統計學 適用：財金系

編號：353

考生注意：
1.依次序作答，只要標明題號，不必抄題。
2.答案必須寫在答案卷上，否則不予計分。
3.限用藍、黑色筆作答；試題須隨卷繳回。

1. **Multiple Choice Questions** (20 points, 1 point for each)

(1) Some hotels ask their guests to rate the hotel's services as excellent, very good, good, and poor. This is an example of the

(A) ordinal scale (B) ratio scale

(C) nominal scale (D) interval scale

(2) Quantitative data refers to data obtained with a(n)

(A) ordinal scale (B) nominal scale

(C) either interval or ratio scale (D) only interval scale

(3) A histogram is said to be skewed to the left if it has a

(A) longer tail to the right (B) shorter tail to the right

(C) shorter tail to the left (D) longer tail to the left

(4) Data that provide labels or names for categories of like items are known as

(A) qualitative data (B) quantitative data

(C) label data (D) category data

(5) If a data set has an even number of observations, the median

(A) cannot be determined (B) is the average value of the two middle items

(C) must be equal to the mean

(D) is the average value of the two middle items when all items are arranged in ascending order

(6) A continuous random variable may assume

(A) any value in an interval or collection of intervals

(B) only integer values in an interval or collection of intervals

(C) only fractional values in an interval or collection of intervals

(D) only the positive integer values in an interval

(7) When data are positively skewed, the mean will usually be

(A) greater than the median (B) smaller than the median

(C) equal to the median (D) positive

(8) Two events, A and B, are mutually exclusive and each have a nonzero probability. If event A is known to occur, the probability of the occurrence of event B is

(A) one (B) any positive value

(C) zero (D) any value between 0 to 1

(9) Bayes' theorem is used to compute

(A) the prior probabilities (B) the union of events

(C) intersection of events (D) the posterior probabilities

(10) A description of the distribution of the values of a random variable and their associated probabilities is called a

(A) probability distribution (B) random variance

(C) random variable (D) expected value

(11) Larger values of the standard deviation result in a normal curve that is

(A) shifted to the right (B) shifted to the left

(C) narrower and more peaked (D) wider and flatter

(12) If the mean of a normal distribution is negative,

(A) the standard deviation must also be negative

(B) the variance must also be negative

(C) a mistake has been made in the computations, because the mean of a normal distribution can not be negative

(D) None of these alternatives is correct

(13) The probability distribution of all possible values of the sample proportion $\bar{p}$ is the

(A) probability density function of $\bar{p}$

(B) sampling distribution of $\bar{x}$

(C) same as $\overline{p}$, since it considers all possible values of the sample proportion

(D) sampling distribution of $\overline{p}$

(14) As the sample size increases, the
(A) standard deviation of the population decreases
(B) population mean increases
(C) standard error of the mean decreases
(D) standard error of the mean increases

(15) From a population that is normally distributed, a sample of 25 elements is selected and the standard deviation of the sample is computed. For the interval estimation of μ, the proper distribution to use is the
(A) normal distribution
(B) t distribution with 25 degrees of freedom
(C) t distribution with 26 degrees of freedom
(D) t distribution with 24 degrees of freedom

(16) In the past, 75% of the tourists who visited Chattanooga went to see Rock City. The management of Rock City recently undertook an extensive promotional campaign. They are interested in determining whether the promotional campaign actually increased the proportion of tourists visiting Rock City. The correct set of hypotheses is
(A) H_0: $P > 0.75$ H_a: $P \leq 0.75$
(B) H_0: $P < 0.75$ H_a: $P \geq 0.75$
(C) H_0: $P \geq 0.75$ H_a: $P < 0.75$
(D) H_0: $P \leq 0.75$ H_a: $P > 0.75$

(17) A 95% confidence interval for a population mean is determined to be 100 to 120. If the confidence coefficient is reduced to 0.90, the interval for μ
(A) becomes narrower
(B) becomes wider
(C) does not change
(D) becomes 0.1

(18) If a hypothesis test leads to the rejection of the null hypothesis

(A) a Type II error must have been committed

(B) a Type II error may have been committed

(C) a Type I error must have been committed

(D) a Type I error may have been committed

(19) If we are interested in testing whether the proportion of items in population 1 is larger than the proportion of items in population 2, the

(A) null hypothesis should state $P_1 - P_2 < 0$

(B) null hypothesis should state $P_1 - P_2 \geq 0$

(C) alternative hypothesis should state $P_1 - P_2 > 0$

(D) alternative hypothesis should state $P_1 - P_2 < 0$

(20) The random variable for a chi-square distribution may assume

(A) any value between -1 to 1

(B) any value between - infinity to +infinity

(C) any negative value

(D) any value greater than zero

2. Pate's Pharmacy, Inc. operates a regional chain of 120 pharmacies. Each pharmacy's floor plan includes a greeting card department which is relatively isolated. Sandra Royo, Marketing Manager, feels that the level of lighting in the greeting card department may affect sales in that department. She chooses three levels of lighting (soft, medium, and bright) and randomly assigns six pharmacies to each lighting level. Analysis of Sandra's data yielded the following ANOVA table. (8 points, 2 points for each)

Source of Variation	SS	df	MS	F
Treatment	3608.333	2	1804.167	
Error	13591.67	15	906.1111	
Total	17200	17		

$F_{0.05}(2,15) = 3.68$, $F_{0.05}(2,17) = 3.59$, $F_{0.05}(15,2) = 19.43$, $F_{0.05}(15,17) = 2.31$

(1) Sandra's experimental design is a ________.

(A)factorial design

(B)random block design

(C)normalized block design

(D)completely randomized design

(2)In Sandra's experiment "lighting" is ________

(A)the dependent variable

(B)a concomitant variable

(C)a treatment variable

(D)a blocking variable

(3)Using $\alpha = 0.05$, the calculated F value is ________

(A)0.5022

(B)0.1333

(C)1.9911

(D)7.5000

(4)Using $\alpha = 0.05$, the appropriate decision is ________.

(A)do not reject the null hypothesis $\mu_1 \neq \mu_2 \neq \mu_3$

(B)do not reject the null hypothesis $\mu_1 = \mu_2 = \mu_3$

(C)reject the null hypothesis $\mu_1 \geq \mu_2 \geq \mu_3$

(D)reject the null hypothesis $\mu_1 \leq \mu_2 \leq \mu_3$

3. A market research firms conducts studies regarding the success of new products. The company is not always perfect in predicting the success. Suppose that there is a 50% chance that any new product would be successful (and a 50% chance that it would fail). In the past, for all new products that ultimately were successful, 80% were predicted to be successful (and the other 20% were inaccurately predicted to be failures). Also, for all new products that were ultimately failures, 70% were predicted to be failures (and the other

30% were inaccurately predicted to be successes). (6 points, 2 points for each)

(1) What is the a priori probability that a new product would be a success?

(A) 0.50

(B) 0.80

(C) 0.70

(D) 0.60

(2) For any randomly selected new product, what is the probability that the market research firm would predict it to be a success?

(A) 0.80

(B) 0.50

(C) 0.45

(D) 0.55

(3) If the market research predicted that the product would be a success, what is the probability that it actually would be a success?

(A) 0.27

(B) 0.73

(C) 0.80

(D) 0.24

4. Analysis of data for an autoregressive forecasting model produced the following tables. (6 points, 2 points for each)

	Coefficients	Standard Error	t Statistic	P-value
Intercept	4.85094	5.745787	0.84426	0.40299
y_{t-1}	-0.10434	0.062849	-1.66023	0.103822
y_{t-2}	0.962669	0.065709	14.65044	9.69E-19

科目：統計學 適用：財金系

編號：353

考生注意：
1.依次序作答，只要標明題號，不必抄題。
2.答案必須寫在答案卷上，否則不予計分。
3.限用藍、黑色筆作答；試題須隨卷繳回。

	df	*SS*	*MS*	*F*	*P*-value
Regression	2	135753.5	67876.76	107.3336	1.91E-17
Residual	43	27192.79	632.3904		
Total	45	162946.3			

(1) In an autoregressive forecasting model, the independent variable(s) is (are) ______.

(A) time-lagged values of the dependent variable

(B) first-order differences of the dependent variable

(C) second-order, or higher, differences of the dependent variable

(D) first-order quotients of the dependent variable

(2) The results indicate that ______

(A) the first predictor, y_{t-1}, is significant at the 5% level

(B) the second predictor, y_{t-2}, is significant at the 5% level

(C) all predictor variables are significant at the 5% level

(D) none of the predictor variables are significant at the 5% level

(3) The actual values of this time series, y, were 228, 54, and 191 for May, June, and July, respectively. The predicted (forecast) value for July is ________.

(A) 36.91

(B) 83.67

(C) 218.71

(D) 174.41

5. A Cornell University study of wage differentials between men and women reported that one of the reasons wages for men are higher than wages for women is that men tend to have more years of work experience than women. Assume the following sample summaries show the years of experience for each group.

Men	**Women**
$n_1 = 100$	$n_2 = 64$
$\bar{x}_1 = 15$ years	$\bar{x}_2 = 10$ years
$\sigma_1 = 5$ years	$\sigma_2 = 4$ years

(1) At 95% confidence, what is the margin of error? (5 points)

(2) What is the 95% confidence interval estimate of the difference between the two population means? (5 points)

6. Using the following data (show your work).

x	y	$x-\bar{x}$	$(x-\bar{x})^2$	$y-\bar{y}$	$(x-\bar{x})(y-\bar{y})$
3	5	2	4	3	6
2	2	1	1	0	0
1	3	0	0	1	0
-1	2	-2	4	0	0
0	-2	-1	1	-4	4
$\sum x_i = 5$	$\sum y_i = 10$	$\sum(x_i-\bar{x})=0$	$\sum(x_i-\bar{x})^2=10$	$\sum(y_i-\bar{y})=0$	$\sum(x_i-\bar{x})(y_i-\bar{y})=10$

Note: $\hat{\sigma}^2 = 5$

$\alpha = 0.01$

$t(0.950, 2) = 2.92$ $t(0.950, 3) = 2.35$ $t(0.950, 4) = 2.13$

$t(0.975, 2) = 4.30$ $t(0.975, 3) = 3.18$ $t(0.975, 4) = 2.78$

$t(0.990, 2) = 6.97$ $t(0.990, 3) = 4.54$ $t(0.990, 4) = 3.75$

$t(0.995, 2) = 9.93$ $t(0.995, 3) = 5.84$ $t(0.995, 4) = 4.60$

(1) Compute a 99% prediction interval for y given $x_0=5$. (3 points)

(2) Compute a 99% prediction interval for y given $x=\bar{x}$. Compare the width of this interval to the one computed in part (1). (4 points)

(3) Compute a 99% confidence interval for $E(y|x_0=5)$. (3 points)

7. The following estimated regression equation relating sales to inventory investment and advertising expenditures was given.

$\hat{y} = 25 + 10x_1 + 8x_2$

The data used to develop the model came from a survey of 11 stores; for those data, SST = 16,000 and SSR = 12,000.

(1) For the estimated regression equation given, compute R^2. (3 points)

(2) Compute the adjusted R^2. (4 points)

(3) Does the model appear to explain a large amount of variability in the data? Explain. (3 points)

8. Let the probability density function of two random variables X_1 and X_2 be

$$f(x_1,x_2)=\begin{cases}\dfrac{x_1x_2}{k}, & x_1=1,2,3;\quad x_2=1,2;\ k>0.\\ 0, & otherwise.\end{cases}$$

(1) Find the value of k. (2 points)

(2) Find the marginal density function of X_1 and X_2. (4 points)

(3) Evaluate $P\{X_1X_2 \geq 4\}$. (4 points)

9. The random variable Y is said to have a lognormal distribution if $X = \log Y$ is $N(\mu,\sigma^2)$.

(1) Derive the probability density function of Y. (5 points)

(2) Find E(Y). (5 points)

10. Let $\{(x_i,y_i); i=1,2,3,\ldots,n\}$ be a sample from a bivariate normal distribution with means μ_1 and μ_2, variance 4 and 9, and correlation coefficient 1/6. Let $(\bar{x},\bar{y})$ denote the vector of sample means.

(1) Find the covariance between $\bar{x}$ and $\bar{y}$? (5 points)

(2) Given $(\bar{x},\bar{y})$ = (0.8, 0.6) and n = 99, find a 95% confidence interval for $\mu_1-\mu_2$? (5 points)