PROVISIONAL ANSWER KEY
NAME OF THE POST
Assistant Professor Statistics, Class II, Advt No. 87/2016-17(AUQ)
Preliminary Test on 02-04-2017
Subeject Que : 101-300
Publish Date : 07-04-2017
Last Date to send suggestion(s) :15-04-2017
Note: Candidate must ensure the compliance to send all suggestion in the given format with reference to this paper with provisional answer key only.

101. The relationship between the mean and variance of χ^2 with n d.f is

(A) mean = 2variance (B) 2mean = variance

(C) mean = variance (D) none of the above

102. Circular Systematic sampling was first used by

(A) W.G. Cochran (B) M.H. Hansen

(C) D.B. Lahiri (D) P.C. Mahalanobis

103. The hypothesis that the population variance has a specified value can be tested by

(A) F-test (B) Z-test

(C) χ^2 test (D) none of the above

104. Moving average method suffers from

(A) loss of information

(B) element of subjectivity

(C) decision about the number of years in groups

(D) all of the above

105. Relation between expected value of R and S.D. σ with usual constant factors

(A) $E(R) = d_1\sigma$ (B) $E(R) = d_2\sigma$

(C) $E(R) = D_1\sigma$ (D) $E(R) = D_2\sigma$

106. The factors responsible for deciding the number of replications in an experiment are

(A) desire of the experimenter

(B) minimum degrees of freedom required for experimental error

(C) shape of experimental units

(D) all of the above

107. If there exists feasible solution to both the primal and dual then

(A) the primal may have an unbounded solution while the dual a bounded solution

(B) the primal may have a bounded solution while the dual a bounded soltion

(C) both primal and dual have unbounded solution

(D) both primal and dual must have bounded optimal solution

108. If each of X variate is divided by 5 and of Y by 10 , then b'_{YX} by coded value is :

(A) same as b_{YX} (B) half of b_{YX}

(C) twice of b_{YX} (D) none of the above

109. The point that lies on both the lines of regression for a bivariate distribution

(A) (0, 0) (B) $(\overline{X}, \overline{Y})$

(C) (σ_X, σ_Y) (D) (r_1, r_2)

110. The characteristic function of the binomial distribution for the binomial variate $X \sim b(n, p)$ is

(A) $(q + pe^{it})$ (B) $(p + qe^{it})^n$

(C) $(p + qe^{t})^n$ (D) $(q + pe^{it})^n$

111. The discrepency between the estimates and population parameters is known as

(A) human error (B) enumeration error

(C) sampling error (D) formula error

112. The degrees of freedom for statistic χ^2 in case of contingency table of order (2×2) is

(A) 3 (B) 4

(C) 2 (D) 1

113. Least square estimates of parameter of a trend line

(A) have minimum variance (B) are unbiased

(C) can exactly be obtained (D) all of the above

114. If μ and σ are the process mean and S.D. then the control limits $\mu \pm 3\sigma$ are known as

(A) modified control limits

(B) natural control limits

(C) specified control limits

(D) none of the above

115. Errors in a statistical experiment are always taken to be

(A) independent (B) distributed $N(0, \sigma_e^2)$

(C) both (A) and (B) (D) neither (A) nor (B)

116. In regular simplex method

(A) the iterations move towards feasibility maintaining optimality

(B) the iterations move towards optimality maintaining feasibility

(C) the iterations maintain both optimality and feasibility

(D) none of the above

117. If X and Y are independent variables then two lines of regression are

(A) X = 0, Y = 0 (B) X = 0, Y = constant

(C) X = constant, Y = 0 (D) X = constant, Y = constant

118. In a statistical test for equality of mean, such as $H_0 : \mu = 10$, if $\alpha = 0.05$

(A) 95% of the time we will make incorrect decisions

(B) 5% of the time we will say that there is a real difference when there is no difference

(C) 5% of the time we will say that there is no real difference when there is a difference

(D) 95% of the time the null hypothesis will be correct

119. If A and B are independent events then

(A) $P(A/B) = P(A) \cdot P(B)$ (B) $P(A/B) = P(B)$

(C) $P(A/B) = P(A)$ (D) None of these

120. Regarding the number of strata which statement is true?

(A) lesser the number of strata, better it is

(B) more the number of strata, poorer it is

(C) more the number of strata, better it is

(D) not more than ten times should be the there in the stratum

121. The best critical region consists of

(A) extreme positive values (B) extreme negative values

(C) both (A) and (B) (D) neither (A) nor (B)

122. If the value of a series at any time t is a function of its values at some previous time intervals such a time series is known as

(A) Autoregressive series (B) Fourier series

(C) Harmonic series (D) none of the above

123. Sampling inspection procedure by variables as compared to by attributes is

(A) more prevalent (B) not practised

(C) less prevalent (D) all of the above

124. A randomised block design has

(A) two way classification (B) one way classification

(C) three way classification (D) no classification

125. The set of basic optimal solutions to an LPP is

(A) finite (B) convex

(C) either singleton or infinte (D) none of the above

126. The ratio of the regression sum of squares to the total sum of square is called

(A) correlation index (B) coefficient of determination

(C) both (A) and (B) (D) neither (A) nor (B)

127. The degrees of freedom for χ^2 in case of dichotomised frequencies are

(A) 4 (B) 2

(C) 1 (D) 0

128. What is the probability to get two aces in succession (with replacement) from a deck of cards?

(A) 1/52 (B) 2/169

(C) 2/159 (D) 1/169

129. Formula for standard error of sample mean $\bar{x}$ based on sample size n having variances s^2, when population considered of N is

(A) s/n (B) $s/\sqrt{n-1}$

(C) $s/\sqrt{N-1}$ (D) $s/\sqrt{n}$

130. From a normal population $N \sim (\mu, \sigma^2)$

(A) A sample mean is not consistent estimator of μ

(B) A sample mean is not consistent estimator of σ^2

(C) A sample mean is consistent estimator of μ

(D) A sample mean is consistent estimator of σ^2

131. Area of critical region depends on

(A) size of Type I error (B) size of Type II error

(C) value of statistic (D) number of observations

132. Trend cannot be

(A) linaer (B) non-linear

(C) S-shaped in short duration (D) none of the above

133. The graph of proportion of defectives in the lot against average sample number is

(A) O.C. curve (B) A.S.N. curve

(C) power curve (D) all of the above

134. For a $k \times k$ Latin Square, the error d.f. in analysis of variance is equal to

(A) $(k-1)(k-2)$ (B) $k(k-1)(k-2)$

(C) (k^2-2) (D) (k^2-k-2)

135. Dual Simplex method is applicable if

(A) opimality condition is satisfied in the initial table

(B) opimality condition is satisfied in the starting table

(C) opimality condition is not satisfied

(D) none of the above

136. Given the two regression lines $X+2Y=5$ and $2X+3Y=8$ and $\sigma_Y^2 = 4$, the value of σ_X^2 is

(A) 12 (B) 27/4

(C) 6 (D) none of the above

137. The selected items of a sample resulted into the same values pertaining to a character. The variance of the sample is

(A) 1 (B) 0

(C) ∞ (D) not determinable

138. Let X be a normal random variable with mean zero and variance 9. If $a = P(X \geq 3)$ then $P(|X| \leq 3)$ equals

(A) a (B) $1 - a$

(C) $2a$ (D) $1 - 2a$

139. Under proportional allocation, the size of the sample from each stratum depends on

(A) total sample size (B) size of the stratum

(C) population size (D) all of the above

140. Neymann-Pearson lemma provides

(A) an unbiased test (B) a most powerful test

(C) an admissible test (D) minimax test

141. Personal income and business loans in forecasting are used as

(A) lead indicators (B) coincident indicators

(C) lag indicators (D) all of the above

142. The decision in a sequential sampling scheme is taken as

(A) after inspecting the sample as a whole

(B) after selection and inspection of items one by one

(C) both (A) and (B)

(D) none of the above

143. The missing value in an experiment is estimated by method of

(A) minimizing the error mean square (B) analysis of covariance

(C) both (A) and (B) (D) neither (A) nor (B)

144. Which of the following statements is correct?

(A) Every LPP admits an optimal solution

(B) Every LPP admits an unique optimal solution

(C) If an LPP admits two optimal solution, it has infinite number of optimal solutions

(D) The set of all feasible solutions to an LPP is not a convex set

145. If the two lines of regression are $X + 2Y - 5 = 0$ and $2X + 3Y - 8 = 0$, the means of X and Y are

(A) (–3, 4) (B) (2, 4)

(C) (1, 2) (D) none of the above

146. Circular Systematic sampling is used when

(A) N is a multiple of n (B) N is a whole number

(C) N is a not divisible by n (D) none of the above

147. The points of inflexion of *t*-distribution are

(A) $\pm\sqrt{\frac{n}{n+1}}$ (B) $\pm\left(\frac{n}{n-2}\right)^{1/2}$

(C) $\pm\left(\frac{n}{n+2}\right)^{1/2}$ (D) $\pm\sqrt{\frac{n+2}{n}}$

148. Which of the following statements is correct?

(A) Systematic sample is superior than stratified sample

(B) Simple random sample is inferior that systematic sample

(C) Stratified sample is superior than Systematic sample

(D) none of the above

149. Most of the non parametric methods utilize

(A) interval scale (B) ratio scale

(C) ordinal scale (D) nominal scale

150. Exponential smoothing method of forecasting can involve

(A) only one smoothing coefficient

(B) up to two smoothing coefficients

(C) an infinite number of smoothing coefficients

(D) none of the above

151. OC curve reveals the ability of the sampling plan to distinguish between

(A) good and bad lots (B) good and sampling plans

(C) good and bad products (D) all of the above

152. An experiment having several factors with equal number of levels is known as

(A) complex experiment (B) symmetrical factorial experiment

(C) asymmetrical experiment (D) all of the above

153. The statement 'The next arrival of a customer is independent of previous arrival of customer' is stated as

(A) memoryless property (B) memorable property

(C) shortage property (D) none on the above

154. A coefficient of any independent variable in a multiple linear regression equation is known as

(A) partial regression coefficient (B) multiple regression coefficient

(C) simple regression coefficient (D) none on the above

155. Probability of drawing a unit at each selection remains same in

(A) srswor

(B) srswr

(C) both (A) and (B)

(D) neither (A) nor (B)

156. If X is a random varaible with mean μ the $E\left(X-\mu\right)^r$ is called

(A) variance

(B) r^{th} raw moment

(C) r^{th} central moment

(D) none of the above

157. Supposing that in cluster sampling s_w^2 represents the variance within the clusters and s_b^2 between clusters . What is the relation between s_w^2 and s_b^2 ?

(A) $s_w^2 = s_b^2$

(B) $s_w^2 \geq s_b^2$

(C) $s_w^2 \leq s_b^2$

(D) none of the above

158. Randomness of a sequence through runs test is judged by comparing the observed number of run with

(A) two critical values

(B) one upper critical value

(C) one lower critical value

(D) none of the above

159. Econometric methods involve

(A) Economics and mathematics

(B) economics and statistics

(C) economics, statistics and mathematics

(D) none of the above

160. If all effects of the same order are confounded with incomplete block differences, it is said to be

(A) complete confounding

(B) partial confounding

(C) balanced confounding

(D) none of the above

161. The interarrival time is 10 min, then the arrival rate is

(A) 8

(B) 1/10

(C) 3

(D) 2

162. If r_1 and r_2 are regression coefficients of Y on X and X on Y respectively

(A) $r_1 + r_2 = 2r$

(B) $r_1 + r_2 < 2r$

(C) $r_1 + r_2 \geq 2r$

(D) none of the above

163. Non-response in surveys mean

(A) non-availability of respondents

(B) non-return of questionnaire by the respondents

(C) refusal to give information by the respondents

(D) all of the above

164. The distrubution for which moment generating function does not exist but moments exist

(A) Pareto distribution (B) *t*-distribution

(C) *F*-distribution (D) all of the above

165. Principle of optimisation in sampling methods is related to

(A) Cost and efficiency of sampling design

(B) validity of estimates

(C) asymptotic properties of estimates

(D) all of the above

166. The Wilcoxon's signed-rank test considers the differences $X_i - M_0$ by way of

(A) signs only (B) magnitudes only

(C) signs and magnitudes both (D) all of the above

167. Variables such as lifestyle in an economic system are known as

(A) endogenous variables (B) exogenous variables

(C) discrete variables (D) none of the above

168. The sign test in non parametric inference is similar to

(A) *t*-test (B) paired *t*-test

(C) *F*-test (D) chi squarred test

169. In a 3^3 factorial with factors A, B and C each with 3 levels, the interaction A^2BC^2 is same as the interaction

(A) ABC (B) AB^2C

(C) AB^2C^2 (D) A^2BC

170. Replacement problems are associated with

(A) setup cost (B) maintenance cost

(C) value of equipments (D) none of above

171. The relation $r = \sqrt{b_{YX}\ b_{XY}}$ is known as

(A) mean property of regression coefficients

(B) fundamental property of regression coefficients

(C) signature property of *r*

(D) none of above

172. Normal distribution was invented by

(A) Laplace (B) De-Moivre

(C) Gauss (D) all of the above

173. Double sampling is also known as

(A) two stage sampling (B) two phase sampling

(C) two directional sampling (D) all of the above

174. In non parametric statistic the confidence interval is usually found out for

(A) population median (B) population mean

(C) both (A) and (B) (D) neither (A) nor (B)

175. The lead-lag relationship theory of forecasting is also known as

(A) action-reaction theory (B) sequence theory

(C) cross-cut analysis theory (D) none of the above

176. A test is said to be unbiased when

(A) type I and type-II errrors are same

(B) type-I error is more than type -II error

(C) Power is more than the size

(D) Power is less than the size

177. In experimental designs, randomization is necessary to make the estimates

(A) valid (B) accurate

(C) precise (D) biased

178. M|M|3|FCFS|N , FCFS and N are

(A) Queue Discipline and arrival pattern

(B) Queue Discipline and buffer capacity

(C) arrival and departure distribution

(D) none of the above

179. If we transform the variables of simple regression line into stansard normal deviates then the two regression lines pass through the point

(A) (0,0) (B) (1,1)

(C) $\left(\overline{X},\overline{Y}\right)$ (D) $\left(X,Y\right)$

180. For an exponential distribution $f\left(x\right)=\frac{1}{2}e^{-x/2};x\geq 0$ its mean and variance are

(A) (1/2, 2) (B) (2, 1/4)

(C) (1/2, 1/4) (D) (2, 4)

181. Two stage sampling design is more efficient than single stage sampling if the correlation between the units in the first stage is

(A) negative (B) positve

(C) zero (D) none of the above

182. Fisher's exact test is used when

(A) cell frequency is small (B) all cell frequencies are small

(C) both (A) and (B) (D) none of the above

183. One of the main aspects of business forecasting is

(A) government policies (B) taxation

(C) historical analysis (D) all of the above

184. Mninimum Variance Unbiased Estimators (MVUE) can be obtained by

(A) Rao Blacwell theorem (B) Lehmann Scheffe theorem

(C) Cramer Rao theorem (D) Neyman theorem

185. The additional effect gained due to combined effect of two or more factors is known as

(A) main effect (B) interaction effect

(C) either(A) or (B) (D) neither (A) nor (B)

186. For M|M|1 model the steady state solution is

(A) $p_n = \rho^n(1-\rho)$ (B) $p_n = \rho^n$

(C) $p_n = (1-\rho)$ (D) none of the above

187. Let the correlation coefficient between two variables X and Y be unity. Then the relation between the regression coefficient β_{YX} and β_{XY} that always holds is

(A) $\beta_{YX} > \beta_{XY}$ (B) $\beta_{YX} < \beta_{XY}$

(C) $\beta_{YX} = \beta_{XY}$ (D) $\beta_{YX}\ \beta_{XY} = 1$

188. If a contingency table shows the gender and year of study of BBusSci students (ie: first year, second year, third year or fourth year) in your statistics class, which of the following statements is/are true?

(A) the events "male student" and "female student" are mutually exclusive events

(B) because of the fact that your stats class is a BBusSci course only, all the events are exhaustive, in other words, each student must fall into one of the classifications

(C) an example for the intersection of events (eg: A and B) would be of male students who are in first year

(D) all of the above statements are correct

189. In the notation below, X is the random variable, c is a constant, and V refers to the variance. Which of the following laws of variance is not correct?

(A) $V(c) = 0$ (B) $V(X + c) = V(X)$

(C) $V(X + c) = V(X) + c$ (D) $V(c\,X) = c2\ V(X)$

190. You are given the following: P(A and B) = 0.17, P($\overline{\text{A}}$ and E) = 0.32, P(A and $\overline{\text{E}}$) = 0.17 and P($\overline{\text{A}}$ and $\overline{\text{E}}$) = 0.34. Which of the following statements about A and E is correct?

(A) A and E are mutually exclusive and independent

(B) A and E are mutually exclusive and dependent

(C) A and E are not mutually exclusive but are independent

(D) A and E are not mutually exclusive and are dependent

191. If P(A) = 0.8, P(B) = 0.3 and P(A|B) = 0.6, what is P(A and B)?

(A) 0.18 (B) 0.24

(C) 0.03 (D) 0.3

192. Which of the following cannot generate a Poisson distribution?

(A) The number of cars arriving at a parking garage in a one hour time interval

(B) The number of telephone calls received in a ten minute interval

(C) The number of customers arriving at a petrol station

(D) The number of bacteria found in a cubic yard of soil

193. In a criminal trial, a Type II error is made when:

(A) a guilty defendant is acquitted (set free)

(B) an innocent person is convicted (sent to jail)

(C) a guilty defendant is convicted

(D) an innocent person is acquitted

194. Which of the following distributions is suitable to model the length of time that elapses before the first employee passes through the security door of a company?

(A) exponential (B) normal

(C) poisson (D) binomial

195. The degrees of freedom for statistic -t for paired t-test based on n pairs of observations is

(A) $2(n - 1)$ (B) $(n - 1)$

(C) $2n - 1$ (D) none of the above

196. Cyclic variations are interwoven with

(A) trend (B) seasonal variations

(C) irregular variations (D) all of the above

197. The ratio of the number of replications required in CRD and RBD for the same amount of information is

(A) 6:4 (B) 10:6

(C) 10:8 (D) 6:10

198. The constraints of a LPP including non-negativity restrictions are

(A) closed half spaces only

(B) hyperplanes only

(C) either closed half spaces or hyperplanes only

(D) None of the above

199. Regression coefficient is independent of the change of

(A) scale (B) origin

(C) both origin and scale (D) neither origin nor scale

200. Customarily the large variance in the variance ratio for F-statistic is taken

(A) in the denomiator (B) in the numerator

(C) either way (D) none of the above

201. Let X and Y be jointly distributed random variables having the joint pdf. Then $P(Y > \max(X, -X))$ is equal to

(A) 1/2 (B) 1/3

(C) 1/4 (D) 1/6

202. In a one tail test for the population mean, if the null hypothesis is rejected when the alternative hypothesis is not true, then

(A) a Type I error is committed

(B) a Type II error is committed

(C) a correct decision is made

(D) a two -tail test should be used instead of a one-tail test

203. In a popular shopping centre, the waiting time for an ABSA ATM machine is found to be uniformly distributed between 1 and 5 minutes. What is the probability of waiting between 2 and 3 minutes to use the ATM?

(A) 0.25 (B) 0.5

(C) 0.75 (D) 0.2

204. Sample mean as an estimator of population mean is

(A) Unbiased (B) sufficient

(C) consistent (D) All of the above

205. A random sample of 319 frontseat occupants involved in headon collisions in a certain region resulted in 100 who sustained no injuries. We wish to use this sample data to test whether the true proportion of uninjured occupants in headon collisions exceeds 0.25 or not. What would your conclusion be for this test of hypothesis, given a 5% significance level?

(A) I would conclude that the true proportion of uninjured occupants in headon collisions has increased

(B) I would conclude that the true proportion of uninjured occupants in headon collisions has decreased

(C) I would conclude that the true proportion of uninjured occupants in headon collisions has remained at 0.25

(D) I would conclude that there is too little information to make a correct decision

206. Method of MLE is not applicable for

(A) Normal distribution (B) Exponential distribution

(C) Cauchy distribution (D) Binomial distribution

207. An investor knows that his portfolio is equally likely to yield an annual return anywhere in the interval [5%, 35%]. What is the probability that he will earn more than 22.5% in the forthcoming year?

(A) 0.72 (B) 0.5

(C) 0.42 (D) 0.17

208. If P(A) = 0.6, P(B) = 0.3 and P(A|B) = 0.5, what is P(A and B)?

(A) 0.15 (B) 0.42

(C) 0.27 (D) 0.48

209. Using the confidence interval when conducting a twotail hypothesis test for the population mean, we do not reject the null hypothesis is the hypothesised value for the population mean:

(A) is to the left of the lower confidence limit

(B) is to the right of the upper confidence limit

(C) falls between the lower and upper confidence limits

(D) falls outside of the range of the confidence interval

210. Independent random samples of 10 observations each are drawn from two normal populations. The parameters of these populations are: $\mu = 280$, $\sigma = 30$ and $\mu = 270$ and $\sigma = 30$. The difference between the two sample means has a distribution which is equal to:

(A) N(10, 152.5) (B) N(20, 152.5)

(C) N(10, 180) (D) N(10, 185)

211. It is believed that 70% of first year statistics students got A's for their final matric exams. If 100 first year statistics students are randomly selected, what is the approximate probability that more than 70 of them got A's for matric?

(A) 0.011 (B) 0.006

(C) 0.4562 (D) 0.1151

212. A manufacturer claims that the market share of her product is 60%. What is the probability that in a random sample of 500 customers more than 55% use her product?

(A) 0.0113 (B) 0.1814

(C) 0.9887 (D) 0.9147

213. The length of time patients must wait to see a doctor at an emergency room of a large hospital is uniformly distributed between 40 minutes and 3 hours. What is the expected waiting time for patients to see a doctor?

(A) 110 minutes (B) 80 minutes

(C) 95 minutes (D) 60 minutes

214. Let X represent the amount of time it takes a student to find a parking space in the parking lot at a university. We know that the distribution of X can be modelled using an exponential distribution with a mean of 4 minutes. A student arrives at university 15 minutes before the scheduled start of her first lecture. What is the probability that it will take the student more than 15 minutes to find a parking space, causing her to be late for her lecture?

(A) 0.082 (B) 0.024

(C) 0.287 (D) 0.135

215. A small bank branch has a single teller to handle transactions with customers. Customers arrive at the bank at an average rate of one every three minutes. What is the probability that it will be more than 5 minutes before the first customer arrives for the day after the bank has opened at 8am?

(A) 0.036 (B) 0.189

(C) 0.368 (D) 0.097

216. If the true means of the k populations are equal, then MSTR/MSE should be:

(A) more than 1.00 (B) close to 1.00

(C) close to 0.00 (D) close to –1.00

217. The selection of the most appropriate multivariate analysis depends on

(A) how many variables are treated as dependent in a single analysis

(B) how variables , both dependent and independent are measured

(C) whether the values of variables are normally distributed

(D) both (A) and (B)

218. A hypothesis test is to be conducted to test whether a certain population mean is equal to or greater than 24.4. It is known that the population standard deviation is 7.6. A sample of size 60 is selected from the population and the sample mean is calculated as being 24.52. Assuming that the population follows a normal distribution, what is the pvalue of this test?

(A) 0.127 (B) 0.452

(C) 0.015 (D) 0.816

219. The purpose of simple linear regression analysis is to:

(A) Predict one variable from another variable

(B) Replace points on a scatter diagram by a straight-line

(C) Measure the degree to which two variables are linearly

(D) Obtain the expected value of the independent random variable for a given value of the dependent variable

220. According to the principal component analysis method , the variance of a variable can be partitioned into

(A) common variance, specific variance and error variance

(B) common variance and error variance

(C) common variance and specific variance

(D) variance of a variable cannot be partitioned and is 100% common variance

221. In a study, subjects are randomly assigned to one of three groups: control, experimental *A*, or experimental *B*. After treatment, the mean scores for the three groups are compared. The appropriate statistical test for comparing these means is:

(A) the correlation coefficient

(B) chi square

(C) *t*-test

(D) ANOVA

222. You are given the following probability distribution: $p(-1) = 0.4$, $p(2) = 0.15$, $p(3) = 0.25$ and $p(7) = 0.2$. What is $E(5X + 2)$?

(A) 12.5 (B) 11.8

(C) 14.5 (D) 12.3

223. The root mean square residual

(A) compares the current structure model to a specified independence model to determine the degree of improvement

(B) is the average of the residuals between individual observed and estimated covariance and variance terms

(C) measures badness-of fit representing the degree of model fit er estimated coefficient

(D) is the square root of all observed and estimated error terms

224. The main distinction between CFA and SEM is

(A) the mesurement part

(B) non-independent error variance

(C) independent relationships between latent factors

(D) neither (A) nor (B) nor (C)

225. A hypothesis test is conducted to test whether the mean age of clients at a certain health spa is equal to 25 or not. It is known that the population standard deviation of clients at the spa is 10. 36 clients are randomly selected, and their ages recorded, with the sample mean age being 27.8. What is your decision, at the 5% level of significance, regarding the null hypothesis that the mean age is equal to 25?

(A) reject the null hypothesis at the 5% level of significance and conclude that the mean age of clients at the spa is less than 25

(B) reject the null hypothesis at the 5% level of significance and conclude that the mean age of clients at the spa is not equal to 25

(C) reject the null hypothesis at the 5% level of significance and conclude that the mean age of clients at the spa is more than 25

(D) do not reject the null hypothesis at the 5% level of significance and conclude that the mean age of clients at the spa is 25

226. It is known that the average height of South African women is 1.65m with a standard deviation of 0.1m, and that the average height of South African men is 1.68m with a standard deviation of 0.12m. What is the probability that the average height of a sample of 75 South African women will exceed that of a random sample of 80 South African men?

(A) 0.0455 (B) 0.0119

(C) 0.9545 (D) 0.9881

227. As varaibility due to chance decreases the value of F will

(A) increase (B) stay the same

(C) decrease (D) can't tell

228. If $r_{xy} = 0$

(A) $b_{yx} = 0$ (B) $b_{xy} = 0$

(C) both (A) and (B) (D) b_{yx} not equal to b_{xy}

229. In ANOVA with 4 groups and a total sample size of 44, the computed F statistic is 2.33 In this case, the p-value is:

(A) exactly 0.05 (B) less than 0.05

(C) greater than 0.05 (D) cannot tell - it depends on what the SSE is

230. If a statistics professor tells his class: "All those who got 100 on the statistics test got 20 on the mathematics test, and all those that got 100 on the mathematics test got 20 on the statistics test", he is saying that the correlation between the statistics test and the mathematics test is:

(A) Negative (B) Positive

(C) Zero (D) Difficult to tell

231. A one sample, twotailed test of hypothesis about the population mean of a certain population is conducted, with the population standard deviation being known. The value of the test statistic is 2.89. Which of the following conclusions about this test would be correct?

(A) the null hypothesis can be rejected at the 1% level of significance

(B) the null hypothesis can be rejected at the 5% level of significance

(C) the null hypothesis can be rejected at the 10% level of significance

(D) both (A) and (B) are correct

232. Error deviations measure distances:

(A) within groups (B) between groups

(C) both (A) and (B) (D) none of the above

233. If $b_{yx} = -2$ and $r_{xy} = -1$ then b_{xy} is equal to

(A) –1 (B) –2

(C) 0.5 (D) –0.5

234. What is the function of a post-hoc test in ANOVA

(A) Determine if any statistically significant group differences have occurred.

(B) Describe those groups that have reliable differences between group means.

(C) Set the critical value for the *F* test (or chi-square).

(D) none of the above

235. Testing interaction effects within LISREL is possible by means of

(A) multigroup analysis

(B) continuous varaibe interaction

(C) both (A) and (B)

(D) addding a multiplicative term out of an independent and dependent variable

236. Flaws occur in telephone cabling at an average rate of 3.2 flaws per 1km of cable. What is the probability that the distance between two flaws exceeds 1km?

(A) 0.111 (B) 0.012

(C) 0.001 (D) 0.041

237. If the area to the right of a value of z (z has a standard normal distribution) is 0.0089, what is the value of z?

(A) –0.89 (B) 2.37

(C) 0.89 (D) –1.02

238. If $P(Z < z) = 0.9616$ what is the value of z (z has a standard normal distribution)?

(A) –2.12 (B) 2.12

(C) –1.77 (D) 1.77

239. Given that X is Normally distributed with a mean of 80 and a variance of 100, what is $p(75 < X < 95)$?

(A) 0.15 (B) 0.625

(C) 0.286 (D) 0.533

240. In an ANOVA, we find that the p-value is 0.003. We therefore conclude that:

(A) there is no statistical evidence that any population mean is different from any other

(B) no two population means are equal

(C) there is strong statistical evidence that not all the population means are equal

(D) no two variances are equal

241. The weights of newborn human babies are normally distributed with a mean of 3.2 kg and a standard deviation of 1.1 kg. What is the probability that a randomly selected newborn baby weighs less than 3.5 kg?

(A) 0.607 (B) 0.138

(C) 0.428 (D) 0.262

242. How many degrees of freedom are there in a test comparing five populations with each sample containing 37 observations?

(A) 185 (B) 4

(C) 36 (D) none of the above

243. Let X represent the amount of time it takes a student to find a parking space in the parking lot at a university. We know that the distribution of X can be modelled using an exponential distribution with a mean of 5 minutes. A student arrives at university 5 minutes before the scheduled start of her first lecture. What is the probability that it will take the student more than 5 minutes to find a parking space, causing her to be late for her lecture?

(A) 0.082 (B) 0.024

(C) 0.368 (D) 0.135

244. 60% of university students prefer MCQ tests to written tests. In a random sample of 300 students, what is the probability that more than 55% of them preferred MCQ tests?

(A) 0.7611 (B) 0.9616

(C) 0.2389 (D) 0.9934

245. Which of the following statements regarding the probability density function, $f(x)$, of the uniform distribution is correct?

(A) the height of the density function differs for different values of X

(B) the density function increases as the values of X increase

(C) the density function is roughly "bellshaped"

(D) the density function is constant for all values that X can assume

246. A confidence interval was used to estimate the proportion of statistics students that are females. A random sample of 72 statistics students generated the following 90% confidence interval: (0.438 ~ 0.642). Based on the interval above, is the population proportion of females equal to 0.60?

(A) No, and we are 90% sure that the population proportion of females is not equal to 0.60

(B) No. The proportion is 54.17%

(C) Maybe. 0.60 is a believable value for the population proportion based on the given information

(D) None of the above

247. The time until first failure of a brand of ink jet printers is normally distributed with a mean of 1500 hours and a standard deviation of 200 hours. A large company buys four such printers. What is the probability that the mean lifetime of the four printers is more than 1650 hours?

(A) 0.1587 (B) 0.3413

(C) 0.3085 (D) 0.0668

248. In a random sample of 400 electrical components, 95 are found to be defective. You wish to test the null hypothesis that the population proportion of defective components is 20% versus the alternative hypothesis that the population proportion is not 20%. You choose a significance level of 10%. What is your statistical decision in this case?

(A) Do not reject *H* at the 10% significance level

(B) Reject *H* at the 10% significance level

(C) Decision cannot be made at a 10% significance level

(D) More information is needed in order to be able to complete the hypothesis test

249. The recommended daily dietary allowance for zinc among males older than age 50 years is 15 mg/day. A study undertaken on a sample of 115 males aged between 65 and 74 years reports the average daily intake as 11.3 mg with a standard deviation of 6.43 mg. Researchers with to test whether the actual average daily zinc intake of males aged between 65 and 74 years falls below the recommended allowance. What is the conclusion of the test in this case?

(A) $p > 0.05$, we therefore cannot reject H0 , and conclude that the daily zinc intake of males between 65 and 74 equals the recommended daily allowance

(B) $p < 0.005$, we therefore reject H0 and conclude that the daily zinc intake of males between 65 and 74 is less than the recommended daily allowance

(C) $p > 0.05$, we therefore accept the alternative hypothesis, and conclude that the daily zinc intake of males between 65 and 74 is less than the recommended daily allowance

(D) $p < 0.005$, we therefore reject H0 and conclude that the daily zinc intake of males between 65 and 74 is the same as the recommended daily allowance

250. You are given the following: P(A and E) = 0.17, P($\overline{A}$ and E) = 0.36, P(A and $\overline{E}$) = 0.10 and P($\overline{A}$ and $\overline{E}$) = 0.37. What is P(E)?

(A) 0.47
(B) 0.48
(C) 0.53
(D) 0.56

251. Which of the following statements is correct?

(A) The Exponential distribution is continuous and defined over the interval (∞, ∞)

(B) The mean of the Poisson distribution (with parameter μ) equals the mean of the Exponential distribution (with parameter λ) only when $\mu = \lambda = 1$

(C) It is impossible for a Normal distribution to have a negative population mean

(D) The Binomial distribution has equal mean and variance only when $p = 0.5$

252. Which of the following statements is correct?

(A) If X is normally distributed then the sample mean is skewed to the right

(B) If X is normally distributed then the sample mean is normally distributed with the same mean and variance as X.

(C) If X is not normally distributed then the sample mean is approximately normally distributed as long as the sample size is greater than 30

(D) If X is not normally distributed then the sample mean is not normally distributed

253. A random sample of 200 observations exhibits 36 successes. We wish to test at the 1% significance level whether the true proportion of successes in the population is less than 24%. What is the test statistic value for this hypothesis test?

(A) $t = 1.99$
(B) $t = -1.99$
(C) $z = 1.99$
(D) $z = -1.99$

254. Based on sample data, the 90% confidence interval for the population mean is (127.5, 154.3). If the 10% significance level was used in testing the alternate hypothesis that the true population mean is not equal to 163, the null hypothesis would

(A) be rejected
(B) not be rejected
(C) have to be revised
(D) need to be tested in a subsequent test

255. A study conducted at a certain university shows that 55% of the university's graduates obtain a job in their chosen field within one year after graduation. 15 graduates are selected at the university at random one year after graduation. What is the probability that exactly 7 of them will have found a job in their chosen field?

(A) 0.201
(B) 0.051
(C) 0.078
(D) 0.165

256. In a hypothesis test, the following random sample of six observations was selected from a normal distribution: 118, 105, 112, 119, 85, and 111. You are asked to conclude whether the population mean is different from 100 using the *p*value method. What is the estimated *p*value for this test?

(A) *p*value < 0.005 (B) *p*value < 0.100

(C) *p*value = 0.025 (D) *p*value< 0.200

257. In hypothesis testing, what level of significance would be most appropriate to choose if you knew that making a Type I error would be more costly than making a Type II error?

(A) 0.005 (B) 0.025

(C) 0.05 (D) 0.028

258. The *p*value of a test is the:

(A) smallest significance level at which the null hypothesis cannot be rejected

(B) largest significance level at which the null hypothesis cannot be rejected

(C) smallest significance level at which the null hypothesis can be rejected

(D) probability that no errors have been made in rejecting or not rejecting the null hypothesis

259. You are given the following: $P(A \text{ and } E) = 0.17$, $P(\overline{A} \text{ and } E) = 0.33$, $P(A \text{ and } \overline{E}) = 0.20$ and $P(\overline{A} \text{ and } \overline{E}) = 0.3$. What is $P(\overline{A} \text{ or } \overline{E})$?

(A) 0.67 (B) 0.69

(C) 0.6 (D) 0.7

260. _____ is another method to solve a given LPP involving some artificial variable ?

(A) Big M method (B) Method of penalties

(C) Two phase simplex method (D) None of the above

261. Consider a large population with a mean of 160 and a standard deviation of 30. A random sample of size 64 is taken from this population. What is the standard deviation of the sample mean?

(A) 3.125

(B) 2.5

(C) 3.75

(D) 5.625

262. You are given the following: $P(A \text{ and } E) = 0.16$, $P(\overline{A} \text{ and } E) = 0.32$, $P(A \text{ and } \overline{E}) = 0.12$ and $P(\overline{A} \text{ and } \overline{E}) = 0.40$. What is $P(\overline{A} \text{ or } \overline{E})$?

(A) 0.83 (B) 0.8

(C) 0.84 (D) 0.89

263. In ANOVA, the sample size is 500 and seven groups are compared. The total sum of squares is 10,000. The decision should be:

(A) do not reject the null hypothesis

(B) reject the null hypothesis immediately

(C) not enough information given to make a decision on the null hypothesis

(D) decide that the test must be repeated since only seven groups were involved

264. The dependent variable is also called:

(A) Regressand variable (B) Predictand variable

(C) Explained variable (D) All of the above

265. A spurious relationship between two latent factors is one that is

(A) real (B) significant

(C) non-significant (D) false

266. Please state which statement is incorrect.

(i) Linear programming was first formulated by an English economist L.V. Kantorovich

(ii) LP is generally used in solving maximization or minimization problems subject to certain assumptions.

(A) (ii) only (B) (i) only

(C) Both (i) and (ii) (D) Both are incorrect

267. If there are 'm' original variables and 'n' introduced variables, then there will be _____ columns in the simplex table

(A) $m + n$ (B) $m - n$

(C) $3 + m + n$ (D) $m + n - 1$

268. According to a recent survey of SA households, the probability that the residents of a household own two cars if their annual household income is over Rs. 1,50,000 is 60%. Of the households surveyed, 60% had incomes over Rs. 1,50,000 and 70% had two cars. What is the probability that the residents of a household own two cars and have an annual household income of over Rs. 1,50,000?

(A) 0.48 (B) 0.4

(C) 0.36 (D) 0.52

269. Who originally called it 'Programming of interdependent activities in a linear structure' but shortened it to 'Linear Programming' ?

(A) Dantzig (B) Kantorovich

(C) Marshall (D) None of the above

270. Which of the following is a research hypothesis?

(A) Two population means are equal

(B) Intelligent people have more reading habit

(C) population correlation coefficient is zero

(D) two populations follow same distribution

271. In oreder to test the effect of certain diet on the weight gain of animals we can use

(A) *t*-test (B) paired *t*-test

(C) multiple comparison test (D) ANOVA

272. Split plot design is a sort of

(A) confounded design (B) partially nested design

(C) both (A) and (B) (D) neither (A) nor (B)

273. If there are 10 symbols of two types, equal in number, the maximum possible number of runs is

(A) 8 (B) 9

(C) 10 (D) none of the above

274. When using the '*ABC*' approach to stock categorisation. Which of the following describes class '*C*' items?

(A) High value, high risk (B) High value, low risk

(C) Low value, high risk (D) Low value, low risk

275. The cumulative distribution function $F(x,y)$ of two-dimensional random variables X and Y in terms of probability is equivalent to

(A) $F(x,y) = P(0 < X \le x, 0 < Y \le y)$ (B) $F(x,y) = P(-\infty < X \le \infty, -\infty < Y \le \infty)$

(C) $F(x,y) = P(-\infty < X \le x, -\infty < Y \le y)$ (D) none of the above

276. Let X be a random variable $U(0,1)$ then the variable $Y = -2\log X$ follows

(A) Log-normal distribution (B) Gamma distribution

(C) chi-sqaure distribution (D) exponential

277. A coin is tossed 400 times and it turns up head 216 times . The hypothesis that the coin is unbiased can be tested by

(A) chi-sqaure test (B) Z test

(C) both (A) or (B) (D) neither (A) nor (B)

278. If one regression coefficient of the two regression lines is greater than unity, the other will be

(A) >1 (B) 1

(C) <1 (D) 1/2

279. The maximum limit of percntage defectives in a finally accepted product is called

(A) AQL (B) AOQL

(C) LTPD (D) all the above

280. Drawing items from a lot without giving any heed to their quality is known as

(A) random sampling (B) purposive sampling

(C) systematic sampling (D) blind sampling

281. The effect which is utilised to divide a replicate into a fraction is called

(A) defining contrast (B) alias

(C) confounded effect (D) all the above

282. Out of a number of models fitted to a time series data, the best model can be adjudged by

(A) estimates of parameters (B) value of residual sum of squares

(C) shape of the curves (D) all the curves

283. Which if the following is/are restrictive assumptions of EOQ model?

(A) Demand is known with certainty and is relatively constant over time

(B) No shortages allowed

(C) lead time for rceipt of orders is constant

(D) all of the above statements

284. FNSD analysis is used

(A) to reduce cost (B) for budgeting control

(C) for inventory grading (D) none of the above

285. Laplace Distribution curve in respect to tails is

(A) not skewed (B) positively skewed

(C) negatively skewed (D) not definite

286. The maximum likelihood estimators are necessarily

(A) unbiased (B) sufficient

(C) most efficient (D) unique

287. If each observation of a set is divided by 10 the S.D. of the new observations is

(A) 1/10 of S.D. of original observations

(B) 1/100 of S.D. of original observations

(C) not changed

(D) 10 times S.D. of original observations

288. Th advantage of orthogonal polynomial is

(A) one can fit a polynomial of appropriate degree

(B) it saves time of comutation

(C) easy to fit in the equaton

(D) all the above

289. The shewhart control charts are meant

(A) to detect whether the process is under statistical quality control

(B) to find assignable causes

(C) to reflect the selection of samples

(D) all the above

290. The least square theory was propounded by

(A) Gauss in 1809 (B) Markov in 1900

(C) Fisher in 1920 (D) none of the above

291. A normal random variable has mean = 2 and variance = 4 . Its fourth central moment is

(A) 16 (B) 64

(C) 80 (D) 48

292. Ratio to trend method for seasonal indices provides good results if

(A) the periods are of long duration (B) the periods ar given six monthly

(C) the periods are of short duration (D) all of the above

293. If X is a standard normal variate the $(1/2)X^2$ is a gamma variate with parameters

(A) 1, 1/2 (B) 1/2, 1

(C) 1/2, 1/2 (D) 1, 1

294. The area under the standard normal curve beyond the lines $z = \pm 1.96$ is

(A) 95% (B) 90%

(C) 10% (D) 5%

295. In case of positive skewed distribution the extreme values lie in the

(A) left tail (B) right tail

(C) middle (D) anywhere

296. Chi-square distribution curve with respect to bulginess is

(A) mesokurtic (B) leptokurtic

(C) platykurtic (D) not definite

297. Sample median as an estimator of population mean is always

(A) unbiased (B) efficient

(C) sufficient (D) none of the above

298. Rao-Blackwell theorem enables us to obtain minimum variance unbiased estimator through

(A) unbiased estimators
(B) complete stastistics
(C) efficient statistics
(D) sufficient statistics

299. A zero in a cell of a contingency table having a finite expected value is termed as

(A) structural zero
(B) random zero
(C) false zero
(D) none of the above

300. In a non-parametric statistic, usually the confidence interval is found out for

(A) population median
(B) population mean
(C) both (A) and (B)
(D) neither (A) nor (B)