

# قسم الإحصاء

أسئلة امتحان الكورس الثاني  
الدراسات العليا  
الماجستير

للعام الدراسي 2024 - 2025





**\*Remark : Answer only five questions, for each question (14) degrees**

**Q1: Correct the following statements.**

1. The average conditional variance reflects the extra variability introduced by the moving average term.
2. A more of data-suitable model may be the autoregressive moving average process. It has several of the properties of an autoregressive process, but can be more parsimonuous than a 'long'.
3. autocovariance in Moving Average Processes is are zero  $\gamma_s = 0$  for  $|s| > Q$ .
4. If the autocovariances after a certain lag are zero q, it may be appropriate to fit an MA(q) model to the time series. On the other hand, the autocovariances of any MR(p) process will only decay to zero as the lag increases.
5. Filters that eliminate high frequency cycles are known as high-pass filters.
6. If  $\phi > 1$ , the values are unchanged in absolute terms, it oscillates between (+, -).
7. Most of the economic and financial time series show some kind of seasonal variation.

**Q2\ Choose the correct answer to fill in the blanks the following.**

1. Quasi Maximum Likelihood estimator is the estimator Consists of when one likelihood is specified for the data they actually have ..... distribution.  
a. symmetrical                      b. equal                      c. different
2. Is taken on the .....is that the trend and the cyclic component are perfectly negatively correlation.  
a. Beveridge-Nelson                      b. Hodrick- Prescott Filler                      c. Baxter King Filter
3. The state in SETAR is determined by..... values of the dependent variable.  
a. lagged                      b. Static                      c. Unstable
4. If MA(1) is be Autocorrelation Function  $\rho_1 \neq 0, \rho_1 = 0, s > 1$  and Partial Autocorrelation Function is be.....  
a. non-zero through log p, 0 thereafter                      b. decays toward zero exponentially                      c. all 0
5. The null is rejected that yt contains a unit root if  $\hat{\gamma}_x$  is negative, is equivalent ..... in the original specification.  
a.  $\phi > 1$                       b.  $\phi < 1$                       c.  $\phi = 1$
6. ICs are often applied by estimating the largest model which is thought to correctly capture the dynamics and then dropping lags until the AIC or S/BIC fail to .....  
a. increase                      b. decrease                      c. constancy
7. An ARMA process is said to have .....phase when the roots of  $\phi(z)$  and  $\theta(z)$  both lie outside of the unit circle.  
a. minimum                      b. maximum                      c. equal





Q3: Choose the correct answer with clarification when choosing:

1. When it is  $\Sigma_k = \text{var}(X_k)$  and  $c_k = E(X_{k+1}X_k)$ , so  $P_{X_k}(X_n)$  equal

- a.  $\sum_{j=1}^k \Sigma_{k,k+1-j} X_j$       b.  $\sum_{j=1}^k \phi_{k,k+1-j} c_k$       c.  $\sum_{j=1}^k \phi_{k,k+1-j} X_j$       d. not what was mentioned

2. Differentiating a random walk produces a stationary series,

- a.  $\Delta y_t = \varepsilon_t + (I - L)\Theta(L)\eta_t$       b.  $\Delta y_t = \varepsilon_t - (I - L)\Theta(L)\eta_t$       c.  $\Delta y_t = \varepsilon_t - (I + L)\Theta(L)\eta_t$   
d. not what was mentioned

3. If the model is available with AR operations (1) and as follows  $X_t = 0.2X_{t-1} + \varepsilon_t$ , Thus process has the ACF

- a.  $c_X(k) = \frac{0.2^{k+1}}{1-0.2^2}$       b.  $c_X(k) = \frac{0.2^{2+|k|}}{1-0.2^2}$       c.  $c_X(k) = \frac{0.2^{2+|k|}}{1-0.2^2}$       d. not what was mentioned

4.  $y_t = 1.6y_{t-1} + 0.5y_{t-2} + x_t$  the characteristic roots is

- a. (1.17, 0.42)      b. (1.16, 0.2)      c. (0.21, 0.21)      d. not what was mentioned

Q4\ Let the zero mean AR(p) process  $\{X\}$  where  $X_t = \sum_{j=1}^p \phi_j X_{t-j} + \varepsilon_t$ , in the case that the roots of  $\phi(z)$

are not distinct, the solution is  $\sum_{j=1}^k \lambda_j^k P_{m_j}(k)$ , Prove that?

Q5\ Suppose  $\Delta y_t$  follows an MA(3) so that  $\{y_t\}$  is an I(1) series which has both a permanent and transitive component. Since  $\{y_t\}$  is I(1),  $\Delta y_t$  must be I(0) and so can be described by a stationary ARMA(P,Q) process.

Q6\ Explain the features of an ARCH?

GOODLUCK

Prof. Dr. Sahira H. Zain  
Lecturer

Ass. prof Dr. Ali N. Hussein  
Head of Dept.





Note: Answer only five questions

Q1) The daily expenditures on food  $x_1$  and clothing  $x_2$  for five people are shown in the table below. Use the hierarchical clustering analysis method, by the linkage method, to classify these expenditures.

$x_1$	2	8	9	1	8.5
$x_2$	4	2	3	5	1

----- (14 marks)

Q2) Let  $X \sim N_3 \left( \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 11 & -6 & 2 \\ -6 & 10 & -4 \\ 2 & -4 & 6 \end{pmatrix} \right)$

Find: 1)  $\mu_{x_3 | (x_1, x_2)}$  2)  $\rho_{3,12}$

----- (14 marks)

Q3) a) What is the difference between the discernment analysis and the factor analysis? (8 marks)

Q3) b) Is the following statement true or false? And correct it if it is wrong. (6 marks)

1- If  $x_1$  and  $x_2$  are the same size (both  $P \times 1$ ) and independent, then:

$$x_1 - x_2 \text{ is } N_p(\mu_{x_1} - \mu_{x_2}, \Sigma_{22} - \Sigma_{11})$$

2- The characteristic function of  $x$  distributed according to  $N(\mu, \Sigma)$  is  $\Phi(t) = e^{it'\mu - t'\Sigma t}$ .

3- The formula of Wilks's criterion is  $\Lambda = \frac{|S^2|}{|S_2^2|}$

----- (14 marks)

Q4) Choose the correct answer, explaining your choice:

1- If we have the following data

$y_1$	3	6	5	10
$y_2$	10	12	14	9

& the hypothesis  $H_0: \mu = (6 \ 11)'$ , so the  $T^2$  hotelling calculating is:

a) 0.15      b) 0.08      c) 0.06      d) 0.21

2- Suppose that the random variables  $x_1, x_2$  have the following covariance matrix

$$\Sigma = \begin{pmatrix} 6.67 & 3.42 \\ 3.42 & 6.24 \end{pmatrix}, \text{ and the principal component } CP_2 = 0.729X_1 + 0.6844X_2, \text{ so the}$$

$\rho_{CP_2, X_1}$  is:

a) 0.0842      b) 0.7335      c) 0.2431      d) 0.8874

----- (14 marks)





Q5) a) Find the maximum and minimum values of the function  $f(x, y) = x^2 + y^2$ , subject to the constraint  $x^2 + xy + y^2 = 3$ . (7 marks)

Q5) b) Let  $\underline{X} \sim N(\underline{\mu}, \underline{\Sigma})$ , &  $Q(\underline{X}) = 2X_1^2 + 3X_2^2 - 8X_1X_2 + 3 - 2X_1 + 4X_2$ , Find  $\underline{\Sigma}$  &  $\underline{\mu}$  (7 marks)

----- (14 marks)

Q6) Two samples are drawn with

$$n_1 = 37, n_2 = 12, \text{ and } \bar{X}_1 = \begin{pmatrix} 12.57 \\ 9.57 \\ 11.49 \\ 7.97 \end{pmatrix}, \bar{X}_2 = \begin{pmatrix} 8.75 \\ 5.33 \\ 8.50 \\ 4.75 \end{pmatrix}, S = \begin{pmatrix} 11.25 & 9.4 & 7.14 & 3.88 \\ & 13.53 & 7.83 & 2.55 \\ & & 11.51 & 2.61 \\ & & & 5.8 \end{pmatrix}$$

Test the hypothesis at  $(\alpha = 0.01)$  if:

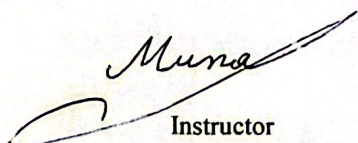
1)  $H_0: \mu_{11} = \mu_{12}$

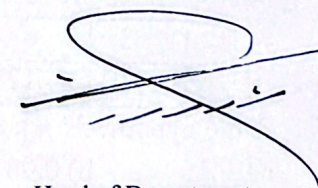
2)  $H_0: \mu_{11} + \mu_{21} = \mu_{12} + \mu_{22}$

Note:  $F(0.01, 4, 44) = 3.82$

----- (14 marks)

**GOOD LUCK**

  
Instructor  
Ass. prof Muna T. Ghafil

  
Head of Department  
Ass. Prof Dr. Ali N. Hussien





\* Answer six questions, including the third and sixth questions.

Q1/ What are the things that the researcher should consider when choosing a research title? (12.5 mark)

Q2/ What are the ethics of scientific research, how many are they? (12.5 mark)

Q3/ Fill in the following blanks with appropriate words.

- 1- One of the most important purposes of scientific research is -----
- 2- Sources for obtaining the scientific research problem----- and-----  
and ----- and ----- and-----
- 3- Types of scientific research depending on its purpose-----and -----.
- 4- There are two ways to formulate the research problem, which are ----  
and -----.
- 5- Nonprobability samples include several types, including ----- and ---  
-----and -----.

(14 mark)

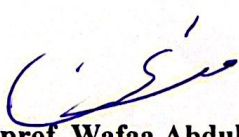
Q4/ Mention the most important characteristics that a successful researcher must have in the field of research. (12.5 mark)

Q5/ What are the main components of the research? (12.5 mark)

Q6/ Mention the steps for collecting data. (6 marks)

Q7/ What are the advantages and disadvantages of a case study? (12.5 mark)

**GOOD LUCK**

  
Ass .prof. Wafaa Abdul samed  
Lecturer

Ass. prof Dr. Ali N. Hussien  
Head of Dept.





Q1/A- If the probability density function of X is :  $f(x) = \frac{e^{-x}}{(1+e^{-x})^2}$  ,  $-\infty < x < \infty$

then what is the probability density function of :  $y = \frac{1}{1+e^{-x}}$  (5 Marks)

B- If the random variable X has a exponential distribution with parameters  $\lambda = 2$  find the probability density function of the random variable  $Y = e^X$ . (5 Marks)

Q2/ Extended the distribution of the Power function where x has : (10 Marks)

$$f(x) = \frac{B x^{B-1}}{\alpha^B} , 0 < x < \alpha , B, \alpha > 0 \text{ then find :}$$

1- E(X) 2-Let  $B = 2, \alpha = 1, \lambda = 2$  and find  $p(0 < x < 1)$

Q3/ If X has Exponential distribution:  $f(x) = \theta e^{-\theta x}$  ,  $x > 0$  (10 Marks)

Using T-X Family method to Find the( Exponential – Exponential) distribution.

Q4/ If  $f(x) = \theta e^{-\theta x}$  ,  $X > 0$  is Exponential distribution and :  $f(x) = \alpha \beta X^{\beta-1} e^{-\alpha x^\beta}$  ,  $x > 0$

Is Weibull distribution then : (10 Marks)

1- Find the Exponential - Weibull mixed distribution.

2- Find the mean of the mixed distribution.

3- IF  $\lambda = 2$  &  $w_1 = \frac{1}{2}$  find  $p(1 < X < 2)$ .

Q5- (A)- Find the median of p.d.f :  $f(x) = \begin{cases} ab^x, & x = 0, 1, 2, \dots \\ 0, & \text{o.w} \end{cases}$

where :  $a+b=1$  ,  $0 < a, b < 1$  (5 Marks)

( B)- If  $f(x) = \begin{cases} 12X^2(1-X) & 0 < X < 1 \\ 0 & \text{o.w} \end{cases}$  (5 Marks)

1- Find the mode of this distribution 2- Find pdf of X on the period [0,0.5]





Q6/ If X has Binomially distribution with parameters n and p and let  $E(x)=5$  and  $V(X)=4$ .

1- Find n and p

(5 Marks)

B- Let  $y_1 < y_2 < y_3 < y_4$  be the order statistic of a random sample of size (4) from a distribution :  $f(x) = \begin{cases} 3x^2 & 0 < x < 1 \\ 0 & \text{O.W} \end{cases}$

Find mode of  $g(y_1)$

(5 Marks)

Q7/ Answer one only (A or B):

(10 Marks)

(A)- If  $x \sim \text{poisson}(m)$  and let  $m \sim \text{Gamma}(\alpha, \beta)$ .

find compound poisson distribution

(B)- If  $X \sim \text{Binomial}(n, p)$  and let  $n \sim \text{poisson}(m)$ .

Find compound binomial distribution.

GOOD LUCK

Ass. prof Dr. Raissan A. Zalan  
Lecturer

Ass. prof Dr. Ali N. Hussain  
Head of Dept.





**\*Remark : Answer five questions only**

**Q1) A-** Explain the reason for each of the following:

- 1- Why calculate statistical power?
- 2- Why is the standard error of skewness and kurtosis calculated?
- 3- Why are cells combined in a chi-square test when the frequency is less than 5?
- 4- Why the initial level of risk or  $\alpha$  must be adjusted In the kruskal wels test ?

**B-** Compare three levels (High, Medium, Low) of social interaction to study adolescents' self-confidence according to the data below.

$$R_H = 709.5 \text{ and } n_H = 20, R_M = 699 \text{ and } n_M = 23, R_L = 607.5 \text{ and } n_L = 20$$

Since there were ties involved in the ranking, there were (11) sets of ties with (2) values, (3) sets of ties with (3) values and (1) sets of ties with (4) values.

----- (A:8 marks, B:6 marks)

**Q2) A-** What are the sample characteristics for parametric assumptions?

**B-** Eight men were involved in a study to examine the resting heart rate regarding frequency of visits to the gym. The assumption is that the person who visits the gym more frequently for a workout will have a slower heart rate, the table below show the rank differences, use ordinal scales to illustrate the relationship and compare these values from one individual to another.

Rank differences	5	1	2	-4	3	4	-4	-7
------------------	---	---	---	----	---	---	----	----

----- (A:5 marks, B:9 marks)

**Q3) A-** If a p-value that exceeds the level of risk associated with the null hypothesis indicates that the observed sample approximated the empirical sample, so how to find the two-tailed probability estimate p-value.

**B-** In a survey study to know the relationship between the age of the individual and the degree of his interest in technology, whether he is (interested in technology or not interested in technology), where three age categories were taken (less than 30 years, 30-50 years, more than 50 years) for the purpose of achieving the goal of this study between the following:

- 1- Formulate appropriate hypotheses.
- 2- What is the appropriate statistical test for this study, detailing its procedures?

----- (A:5 marks, B:9 marks)

**Q4)** The counseling staff Basrah schools District has implemented a new program this year to reduce bullying in their elementary schools. The school district does not know if the new program resulted in improvement or deterioration. In order to evaluate the program's effectiveness, the school district has decided to compare the percentage of





successful interventions last year before the program began with the percentage of successful interventions this year with the program in place. In Table below show sign of differences, the 12 elementary school counselors, or participants, reported the percentage of successful interventions last year and the percentage this year.

Participant	1	2	3	4	5	6	7	8	9	10	11	12
Sign of differences	0	0	-	+	+	+	+	-	+	+	+	+

(14marks)

**Q5) A-** Explain why randomness is tested and what its statistical benefits are. Then, explain what the appropriate procedure would be if randomness does not exist.

**B-** The following data were collected from a study comparing two methods being used to teach reading recovery in the 4th grade. Method 1 was a pull-out program in which the children were taken out of the classroom for 30 min a day, 4 days a week. Method 2 was a small group program in which children were taught in groups of four or five for 45 min a day in the classroom, 4 days a week. Using the data in the table below to find the confidence interval for the difference between location parameters.

Method 1	48	40	39	50	41	38	53
Method 2	14	18	20	10	12	102	17

(A:5 marks, B:9 marks)

**Q6)A-** What are the parametric and non-parametric tests for the cases below, with a mention of their path in spss?

- 1- Comparing two unrelated samples.
- 2- comparing categorical data.
- 3- Comparing two rank-ordered variables.

**B-** Explain the reasons for using non-parametric regression, and what is the kernel smoothing regression method in detail with the conditions.

(A:6 marks, B:8 marks)

**Table values:**

$$r.s(0.05)=0.738$$

$$U(12,12,0.05)=42$$

$$H(2,0.05)=5.99$$

$$W_{\alpha/2} = 9$$

Ass.prof Dr. Asmaa Ayoob Yaqoob  
Lecturer

GOODLUCK

Ass. prof Dr. Ali N. Hussein  
Head of Dept.





**\*Remark : Answer 5 questions only ; The mark is distributed equally among the questions, and each question has 14 marks.**

Q1/ Following are the LP model and its associated optimal simplex tableau.

Maximize  $z = 3x_1 + 4x_2$

subject to

$2x_1 + 3x_2 \leq 1200$  (Resistors2)

$2x_1 + x_2 \leq 1000$  (Capacitors2)

$4x_2 \leq 800$  (Chips2)

$x_1, x_2 \geq 0$

Basic	X1	X2	X3	X4	X5	b
Z	0	0	5/4	1/4	0	1750
X1	1	0	-1/4	3/4	0	450
X5	0	0	-2	2	1	400
X2	0	1	1/2	-0.5	0	100

- Determine the status of each resource.
- In terms of the optimal revenue, determine the dual prices for the resistors, capacitors, and chips.
- If the available number of resistors is increased to 1300 units, find the new optimum solution.
- If the available number of chips is reduced to 350 units, will you be able to determine the new optimum solution directly from the given information? Explain.
- If the unit profit of first production is reduced to 2\$ units, will you be able to determine the new optimum solution directly from the given information? Explain.

Q2/(a) Explain your understanding of the relationship between the arrival rate and the average interarrival time. What are the units describing each parameter?

(b) Derive the expected number of customers in the system for m/m/1/N model.

c) community is served by two cab companies. Each company owns two cabs, and both share the market equally, with calls arriving at each company's dispatching office at the average rate of eight per hour. The average time per ride is 12 minutes. Calls arrive according to a Poisson distribution, and the ride time is exponential. The two companies have been bought by an investor and will be consolidated into a single dispatching office. Analyze the new owner's proposal.

Q3/ Find the optimum solution for the following L.P. Model

Maximize  $z = 3x_1 + 2x_2$

subject to

$2x_1 + 5x_2 \leq 18$

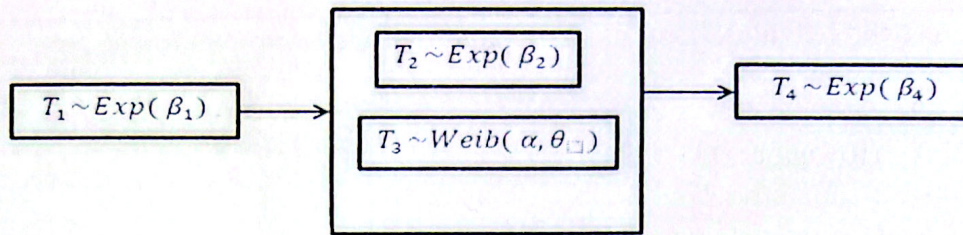
$4x_1 + 2x_2 \leq 18$

$x_1, x_2 \geq 0$  and integer





Q4/ derive reliability function for the following system:



Q5/ Maximize  $z = 3x_1 + 2x_2 + 5x_3$

subject to

$$x_1 + 2x_2 + x_3 \leq 430$$

$$3x_1 + 2x_3 \leq 460$$

$$x_1 + 4x_2 \leq 420$$

$$x_1, x_2, x_3 \geq 0$$

Compute the entire simplex tableau associated with the following basic solution, and check it for optimality and feasibility.

Basic variables =  $(x_2, x_3, x_6)$ ,

$$Inverse = \begin{pmatrix} \frac{1}{2} & -\frac{1}{4} & 0 \\ 0 & \frac{1}{2} & 0 \\ -2 & 1 & 1 \end{pmatrix}$$

Q6/ )A) True or False

(a) The optimum dual objective function value equal the optimum primary objective function value.

(b) the Reliability function can be take negative values

(c) If the primal constraint is of the type  $\leq$ , the corresponding dual variable will be nonnegative (nonpositive) if the primal objective is maximization (minimization).

(d) the duality simplex method is used when the solution not optimality

(e) An unrestricted primal variable will result in an equality dual constraint.

B) Explain the solution algorithm of the probabilistic EOQ model.

Bahaa A. Qasim  
Lecturer

GOODLUCK

Ass. prof Dr. Ali N. Hussein  
Head of Dept.