Université Libanaise
Faculté d'Information

الجـامعـة اللبنـانيـة
كلية الاعلام

العام الجامعي2021\2022

المادة: statistics 2	المرحلة: إجازة
المدة: ساعة ونصف	السنة: 2022
اسم الأستاذ: د. حسين نصار	الاختصاص: علوم البيان

Section 1:

A researcher is investigating the fertility of four types of lots of land in a rural area. He recorded the yield of the lands and table-1 and table-2 summarize the results. To find out if there is a difference in the yields he conducted ANOVA test (table-3)

Table-1

Land-A	Land-B	Land-C	Land-D
8.9	8	9.8	8.6
8.7	8.3	8.8	8.9
8.6	8.2	9.2	8.1
9.3	9	9.7	7.8
9.1	9.1	10.4	8.1
9.1	8.3	8.9	8
9.6	8.8	10.4	8.4
9.7	9.1	9.7	9.1
9	9	9.5	8.8

Table-2

SUMMARY

Groups	Count	Sum	Average	Variance
Land-A	9	82	9.111111	0.138611
Land-B	9	77.8	8.644444	0.192778
Land-C	9	86.4	9.6	0.33
Land-D	9	75.8	8.422222	0.204444

Table-3

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	7.382222222	3	2.460741	11.3682	3.09749E-05	2.90112
Within Groups	6.926666667	32	0.216458			
Total	14.30888889	35				

Questions:

1) Why did the researcher use ANOVA?
2) Given the data in the ANOVA table make a conclusion regarding the yields in the lands (use the 6 steps hypothesis testing process)
3) Check to see if there is a difference between
   a. Land-A and Land-B
   b. Land-A and Land-C

Upon further investigation, the researcher learned that lands are managed by different farmers. Table-4 shows the yields of land relative to type and farmers and table-5 shows the corresponding ANOVA table.

Table-4

Farmer	Land-A	Land-B	Land-C	Land-D
Christian	8.9	8	9.8	8.6
Dave	8.7	8.3	8.8	8.9
Jenny	8.6	8.2	9.2	8.1
Alice	9.3	9	9.7	7.8
Norah	9.1	9.1	10.4	8.1
Fawzi	9.1	8.3	8.9	8
Sam	9.6	8.8	10.4	8.4
Dano	9.7	9.1	9.7	9.1
Mona	9	9	9.5	8.8

Table-5

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Rows	3.213889	8	0.401736	2.596888	0.03359	2.355081
Columns	7.382222	3	2.460741	15.90663	6.6E-06	3.008787
Error	3.712778	24	0.154699			
Total	14.30889	35				

Questions:

4) Name the dependent variable, treatment variable and blocking variable
5) What do you conclude from the ANOVA table?

HINTS:

CONFIDENCE INTERVAL FOR THE DIFFERENCE IN TREATMENT MEANS $(\bar{X}_1 - \bar{X}_2) \pm t\sqrt{MSE\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$

Section 2:

There are many restaurants in northeastern South Carolina. They serve beach vacationers in the summer, golfers in the fall and spring, and snowbirds in the winter. Bill and Joyce Tuneall manage several restaurants in the North Jersey area and are considering moving to Myrtle Beach, SC, to open a new restaurant. Before making a final decision, they wish to investigate existing restaurants and what variables seem to be related to profitability. They gather sample information where profit (reported in $000) is the dependent variable and the independent variables are:

X1 the number of parking spaces near the restaurant.

X2 the number of hours the restaurant is open per week.

X3 the distance from Peaches Corner, a landmark in Myrtle Beach.

X4 the number of servers employed.

X5 the number of years the current owner has owned the restaurant.

Predictor	Coef	SE Coef	T	*p*-value
Constant	2.50	1.50	1.667	–
X_1	3.00	1.500	2.000	0.056
X_2	4.00	3.000	1.333	0.194
X_3	−3.00	0.20	−15.00	0.000
X_4	0.20	.05	4.00	0.000
X_5	1.00	1.50	0.667	0.511

Analysis of Variance

Source	DF	SS	MS	F	*p*-value
Regression	5	100	20	10	0.000
Residual Error	20	40	2		
Total	25	140			

1) What is the amount of profit for a restaurant with 40 parking spaces that is open 72 hours per week, is 10 miles from Peaches Corner, has 20 servers, and has been open 5 years?
2) Perform a global test of hypothesis to check if any of the regression coefficients are different from 0. What do you decide? Use the .05 significance level.
3) Which variables would you consider eliminating? Why?
4) What are the consequences that you can take away from the following correlation matrix

Correlation Coeffiecint Matrix

	x1	x2	x3	x4	x5
x1	1	0.1	-0.05	-0.1	0.02
x2		1	-0.101	0.2	0.33
x3			1	-0.76	0.21
x4				1	0.05
x5					1

Section 3

A researcher is interested in studying the association between gender and smoking. He conducted a survey of a sample of 100 individuals and used chi square to investigate the results. The following tables are the results from SPSS output file:

Gender * Smoking Crosstabulation

Count

		Smoking		Total
		yes	no	
Gender	male	30	20	50
	female	19	31	50
Total		49	51	100

	Value	df	p-value. Sig. (2-sided)
Pearson Chi-Square	4.842[a]	1	.028
N of Valid Cases	100		

Answer the following questions:

1) Why is chi-sqaure test considered appropriate in this case?
2) State the null and alternative hypothesis
3) What can you conclude from the tables regarding researcher's inquiry?

Part II

The researcher has collected data related to the stress level for the same sample. He was curious to know whether stress level has anything to do with the association between the gender and smoking. The following were the results:

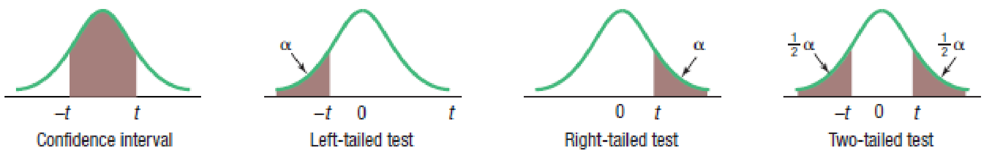
Gender * Smoking * StressLevel Crosstabulation

Count

StressLevel			Smoking: yes	Smoking: no	Total
high	Gender	male	24	7	31
		female	11	7	18
	Total		35	14	49
low	Gender	male	6	13	19
		female	8	24	32
	Total		14	37	51
Total	Gender	male	30	20	50
		female	19	31	50
	Total		49	51	100

StressLevel		Value	df	P-Value. Sig. (2-sided)
high	Pearson Chi-Square	1.484[c]	1	.223
	N of Valid Cases	49		
low	Pearson Chi-Square	.259[d]	1	.611
	N of Valid Cases	51		
Total	Pearson Chi-Square	4.842[a]	1	.028
	N of Valid Cases	100		

4) What do we call this test and why is it helpful ?
5) What conclusions can you make out of the tables above regarding the association between gender and smoking ?



	Confidence Intervals, c							Confidence Intervals, c					
	80%	90%	95%	98%	99%	99.9%		80%	90%	95%	98%	99%	99.9%
	Level of Significance for One-Tailed Test, α							Level of Significance for One-Tailed Test, α					
df	0.10	0.05	0.025	0.01	0.005	0.0005	*df*	0.10	0.05	0.025	0.01	0.005	0.0005
	Level of Significance for Two-Tailed Test, α							Level of Significance for Two-Tailed Test, α					
	0.20	0.10	0.05	0.02	0.01	0.001		0.20	0.10	0.05	0.02	0.01	0.001
1	3.078	6.314	12.706	31.821	63.657	636.619	36	1.306	1.688	2.028	2.434	2.719	3.582
2	1.886	2.920	4.303	6.965	9.925	31.599	37	1.305	1.687	2.026	2.431	2.715	3.574
3	1.638	2.353	3.182	4.541	5.841	12.924	38	1.304	1.686	2.024	2.429	2.712	3.566
4	1.533	2.132	2.776	3.747	4.604	8.610	39	1.304	1.685	2.023	2.426	2.708	3.558
5	1.476	2.015	2.571	3.365	4.032	6.869	40	1.303	1.684	2.021	2.423	2.704	3.551
6	1.440	1.943	2.447	3.143	3.707	5.959	41	1.303	1.683	2.020	2.421	2.701	3.544
7	1.415	1.895	2.365	2.998	3.499	5.408	42	1.302	1.682	2.018	2.418	2.698	3.538
8	1.397	1.860	2.306	2.896	3.355	5.041	43	1.302	1.681	2.017	2.416	2.695	3.532
9	1.383	1.833	2.262	2.821	3.250	4.781	44	1.301	1.680	2.015	2.414	2.692	3.526
10	1.372	1.812	2.228	2.764	3.169	4.587	45	1.301	1.679	2.014	2.412	2.690	3.520
11	1.363	1.796	2.201	2.718	3.106	4.437	46	1.300	1.679	2.013	2.410	2.687	3.515
12	1.356	1.782	2.179	2.681	3.055	4.318	47	1.300	1.678	2.012	2.408	2.685	3.510
13	1.350	1.771	2.160	2.650	3.012	4.221	48	1.299	1.677	2.011	2.407	2.682	3.505
14	1.345	1.761	2.145	2.624	2.977	4.140	49	1.299	1.677	2.010	2.405	2.680	3.500
15	1.341	1.753	2.131	2.602	2.947	4.073	50	1.299	1.676	2.009	2.403	2.678	3.496
16	1.337	1.746	2.120	2.583	2.921	4.015	51	1.298	1.675	2.008	2.402	2.676	3.492
17	1.333	1.740	2.110	2.567	2.898	3.965	52	1.298	1.675	2.007	2.400	2.674	3.488
18	1.330	1.734	2.101	2.552	2.878	3.922	53	1.298	1.674	2.006	2.399	2.672	3.484
19	1.328	1.729	2.093	2.539	2.861	3.883	54	1.297	1.674	2.005	2.397	2.670	3.480
20	1.325	1.725	2.086	2.528	2.845	3.850	55	1.297	1.673	2.004	2.396	2.668	3.476
21	1.323	1.721	2.080	2.518	2.831	3.819	56	1.297	1.673	2.003	2.395	2.667	3.473
22	1.321	1.717	2.074	2.508	2.819	3.792	57	1.297	1.672	2.002	2.394	2.665	3.470
23	1.319	1.714	2.069	2.500	2.807	3.768	58	1.296	1.672	2.002	2.392	2.663	3.466
24	1.318	1.711	2.064	2.492	2.797	3.745	59	1.296	1.671	2.001	2.391	2.662	3.463
25	1.316	1.708	2.060	2.485	2.787	3.725	60	1.296	1.671	2.000	2.390	2.660	3.460
26	1.315	1.706	2.056	2.479	2.779	3.707	61	1.296	1.670	2.000	2.389	2.659	3.457
27	1.314	1.703	2.052	2.473	2.771	3.690	62	1.295	1.670	1.999	2.388	2.657	3.454
28	1.313	1.701	2.048	2.467	2.763	3.674	63	1.295	1.669	1.998	2.387	2.656	3.452
29	1.311	1.699	2.045	2.462	2.756	3.659	64	1.295	1.669	1.998	2.386	2.655	3.449
30	1.310	1.697	2.042	2.457	2.750	3.646	65	1.295	1.669	1.997	2.385	2.654	3.447
31	1.309	1.696	2.040	2.453	2.744	3.633	66	1.295	1.668	1.997	2.384	2.652	3.444
32	1.309	1.694	2.037	2.449	2.738	3.622	67	1.294	1.668	1.996	2.383	2.651	3.442
33	1.308	1.692	2.035	2.445	2.733	3.611	68	1.294	1.668	1.995	2.382	2.650	3.439
34	1.307	1.691	2.032	2.441	2.728	3.601	69	1.294	1.667	1.995	2.382	2.649	3.437
35	1.306	1.690	2.030	2.438	2.724	3.591	70	1.294	1.667	1.994	2.381	2.648	3.435