Exercises Part 1

Delivery: Please for the sake of clarity, label the files that you hand in with this specific format:

T1 – name1_surname1&name2_surname2

1. In a study to investigate the effect of oyster density on seagrass biomass, researchers introduced oysters to thirty parcels of healthy seagrass. At the beginning of the study, the seagrass was clipped short in all plots. Next, 10 randomly chosen plots received a high density of oysters; 10, an intermediate density; and 10, a low density. As a control, an additional 10 randomly chosen clipped parcels received none. After 2 weeks, the belowground seagrass biomass was measured in each parcel (gr/m^2). The mean square error from the ANOVA table was 220.94.

Oyster density				
	None	Low	Intermediate	High
Mean	34.81	33.13	28.33	15.00

- Compute three Bonferroni-adjusted confidence intervals comparing parcels with Low, Intermediate and High density with the one with no oysters using α = 0.05.

- Repeat the previous item with Tukey's Honest Significant Difference.

- Which differences (among all possible) are significant?

2. Serum from two groups of subjects following streptococcal infection was assayed for neutralizing antibodies to streptolysin (*AS*). The results were as follows:

(Measured in Todd units)	
oup A	GroupB
324	558
275	108
349	291
604	863
566	303
810	640
340	358
295	503
357	646
580	689
344	250
655	540
380	630
503	190
314	

Test if the population medians are the same. Use a standard procedure and a resampling method. Compare results.

3. A study involving subjects with chronic back pain compared *conventional* therapy to *alternative* therapy. Only 2 out of 23 subjects assigned to the conventional therapy group suffered relapses in the first year of the study, compared to 8 of the 24 subjects assigned to the alternative therapy group.
Is this sufficient evidence to conclude, at the 0.05 level of significance, that the two types of therapies are not equally effective?
Use a standard procedure and a resampling method. Compare results.

4. A study examined 973 individuals who were in car accidents. It was found that of 247 drivers that wore seatbelts, 17 of them had a head injury. Of the rest of the drivers who did not wear seatbelts, 428 did not get a head injury.

(*i*) Make a contingency table for the data.

(*ii*) Without running any tests, does there appear to be a benefit to wearing a seatbelt?

(*iii*) What are the expected counts for the contingency table?

(*iv*) Test if wearing a seatbelt prevents head injuries. Use standard procedures and resampling methods.

5. The insulin pump is a device that delivers insulin to a diabetic patient at regular intervals. It presumably regulates insulin better than standard injections. However, data to establish this point are few, especially in children. The following study was set up to assess the effect of the insulin pump on HgbA1c, a long-term marker of compliance with insulin protocols. Data were collected on 256 diabetic patients for 1 year before and after using the insulin pump. A subset of the data for 10 diabetic patients is below:

ID	Mean HgbA1 c 1 year before (%}	Mean HgbAlc 1 year after (%)	Before - After (%)
1	6.7	7.0	-0.3
2	7.4	7.4	0
3	9.2	8.6	0.6
4	9.6	8.1	1.5
5	7.4	6.8	0.6
6	8.1	7.0	1.1
7	10.8	8.5	2.3
8	7.1	7.7	-0.6
9	7.9	9.7	-1.8
10	10.8	7.7	3.1

Compare the mean *HgbA1c* one year before versus the mean *HgbA1c* one year after the use of the insulin pump. Use standard procedures and resampling methods.

6. The following table shows caloric intake (*cal*/*day*/*kg*) and oxygen consumption VO_2 (*ml*/*min*/*kg*) in 10 infants.

Caloric Intake (X)	VO_2 (Y)
50	7.0
70	8.0
90	10.5
120	11.0
40	9.0
100	10.8
150	12.0
110	10.0
75	9.5
160	11.9

Test if the two variables are independent.

7. The file `dbp.txt`, contains diastolic blood pressure data from a small randomized clinical trial in 40 patients with hypertension. Diastolic blood pressure (*DBP*) was measured for 5 consecutive months in the supine position. Half of the patients received treatment *A* (new drug) or *B* (placebo). Also, the sex of the patient was recorded.

The aim was to test whether treatment *A* may be effective in lowering *DBP* as compared to *B*.

- Is the *DBP* of patients who took treatment *A* different from the ones who took treatment *B*?

- Is the effect of the two treatments similar over time?

8. Researchers studied the cellular telephone records of 699 persons who had automobile accidents. They determined that 170 of the 699 had made a cellular telephone call during the 10 minutes before their accident; this period is called the *hazard* interval. 37 persons had made a call during a corresponding 10-minute period on the day before their accident; this period is called the *control* interval. Finally, there were 13 who made calls both during the hazard interval and the control interval.
Do these data indicate that the use of a cellular telephone is associated with an increase in the accident rate?
Use standard procedures and resampling methods.

9. Summarize the main ideas about the *False Discovery Rate* (FDR).
Explain the *Benjamini-Hochberg* and the *q-Value* procedures.
Show examples of an application with R with a comparison between both methods (at **least two** pages).

10. Summarize the main ideas about the *ROC* curves. Show an example of an application with R (at **least two** pages).