SORBONNE UNIVERSITÉ
FACULTÉ DES SCIENCES ET INGÉNIERIE

MU5EEH04

Evaluation report

Human experimentation and statistics

Etudiants : Evan BOKOBZA 28614224 Julien JOYET 3800156 Laurent LIN 3808422

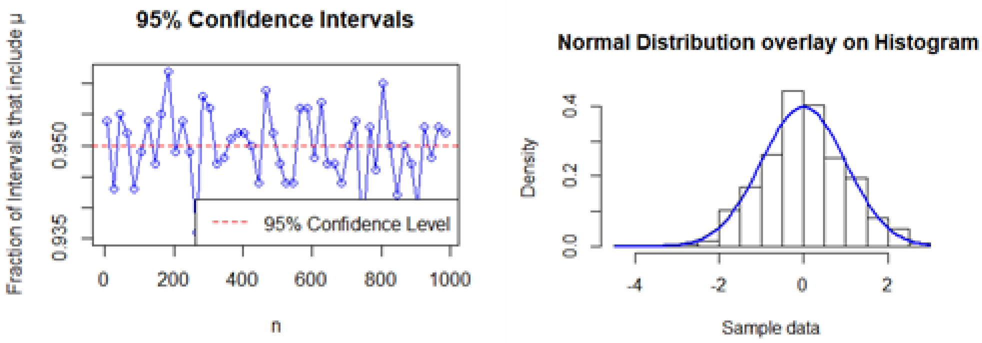
PARIS, JANUARY 2024

1 Statistical theory

1.1 Classical Statistics

1.1.1 Confidence interval

Classical statistical A confidence interval is a statistical tool that provides a range of values within which the population parameter is likely to fall, and this range is associated with a certain level of confidence. The confidence level represents the probability that the interval will contain the true parameter over time.



Clear interpretation of the meaning of a confidence interval

In the first figure , the level of confidence is set at 0.95. That meaningful information indicates that not all intervals include the population means(mu), but rather 95% of the interval. And it is possible to see that the fraction of interval fluctuates around the level of confidence. Over a long run the experiment suggests that the fraction of intervals would converge to 0.95. In the second figure , a normal distribution of the sample data is drawn to be compared to the standard normal distribution. The curve is similar to the standard distribution, therefore the data follow a Gaussian distribution

1.1.2 Frequential statistical test

Based on frequentist probability which focuses on the concept of providing evidence against a null hypothesis to highlight significant differences between various groups through the use of p-values. In order to display the frequential statistics, multivariate analysis of variance (MANOVA) is a method that determines whether multiple levels of independent variables on their own or mixed have a correlation effect with dependent variables. In the MANOVA test, the Pillai statistic is used as the decision variable to compare the differences among sample means.The risk of the first kind which is also called alpha is the probability to reject the true null hypothesis, therefore it showed there are no meaningful differences among the category means. Moreover the beta (risk of second kind) definition assesses the mistaken failure to reject the null hypothesis, indicating in the case of MANOVA a dissimilarity between the group means. The p value means the evidence of null hypothesis against the differences between species and a small p-value suggest strong evidence against the null hypothesis. The power of the MANOVA test depends on the value of the risk of the second kind by the formula power = (1 - β), suggesting the probability to correctly reject the null hypothesis. Furthermore the power of the test will increase the larger the effect size and sample size.

1.2 Bayesian Statistics

Bayesian statistics follow the Bayes theorem to make data analysis and parameter estimation. The fact that p-value is the variance between the data and a null hypothesis. If p-value is between 1 to 5% in Bayesian statistics, it suggests that, a modest Bayesian evidence against the null evidence. In Bayesian statistical point of view it gives insights of the precision of estimation for the posterior distribution. Frequential statistics may object to the null hypothesis H0, but it cannot prove the hypothesis to be true. In this case, Bayesian factors allow us to assume another hypothesis H1 which has been specified. The Bayes factor quantifies whether the data have increased or decreased the probability of H0 in competing with H1 by the Bayes theorem (superior or inferior to 1). $BF_{01} = \frac{P(\text{Data}|H_0)}{P(\text{Data}|H_1)}$

2 Data Analysis

2.1 Designing the experiment

Q1 Blood sugar, Blood pressure, Glucagon, Insulin, are all of them quantitative. Because these data have been generated by a random number generator, therefore they all have numerical representation. Blood pressure and blood sugar are variables of the experiment. As Glucagon and Insulin are factors of the variables, the subject would try to detect a correlation between variables and factors.

Q2 1. Does insulin affect blood sugar ? 2. Does insulin affect blood pressure ? 3. Does glucagon affect blood sugar ? 4. Does glucagon affect blood pressure ? 5. Do the two factors affect blood sugar independently ? 6. Do the two factors affect blood pressure independently ? 7. Are variables blood sugar and blood pressure independent ? 8. If blood sugar and blood pressure are dependent, what is their functional relation ? Relation entre S et I et G.

2.2 Analyzing data in the SPIG_FS.csv file with frequential statistics

Q3 MANOVA means multivariate analysis of variance which conducts tests to compare multivariate sample means, with variables supposed depending on one or more factors. The problem of one way and two way ANOVA is the usage of multiple tests that would increase alpha. With MANOVA, one test is managed so alpha value (the risk of first kind) is weaker. Moreover MANOVA takes account of possible interactions between dependent variables which is not possible with one way or two way ANOVA.

Q4 MANOVA performed on the file named 281222_SPIG_FS.csv. It is possible to see a very small digit for the p value of each parameter (figure 2.1 2.2)

Q5 The result of a small p value, means that the null hypothesis is very incompatible with the blood dataset. In the statistical data grasped by MANOVA, the p value of Ins :Gln is very small, suggesting the hypothesis that the two factors to affect BP and BS independently is not realistic. In all of the hypotheses, two of them are invalid from the p value which are the one who suggest that Gln and Ins are independents.

Q6 Numerical factors and variables suggest calculating the Pearson correlation, because Pearson correlation is an index reflecting on the linear behavior of two continuous variables. Depending on the value of

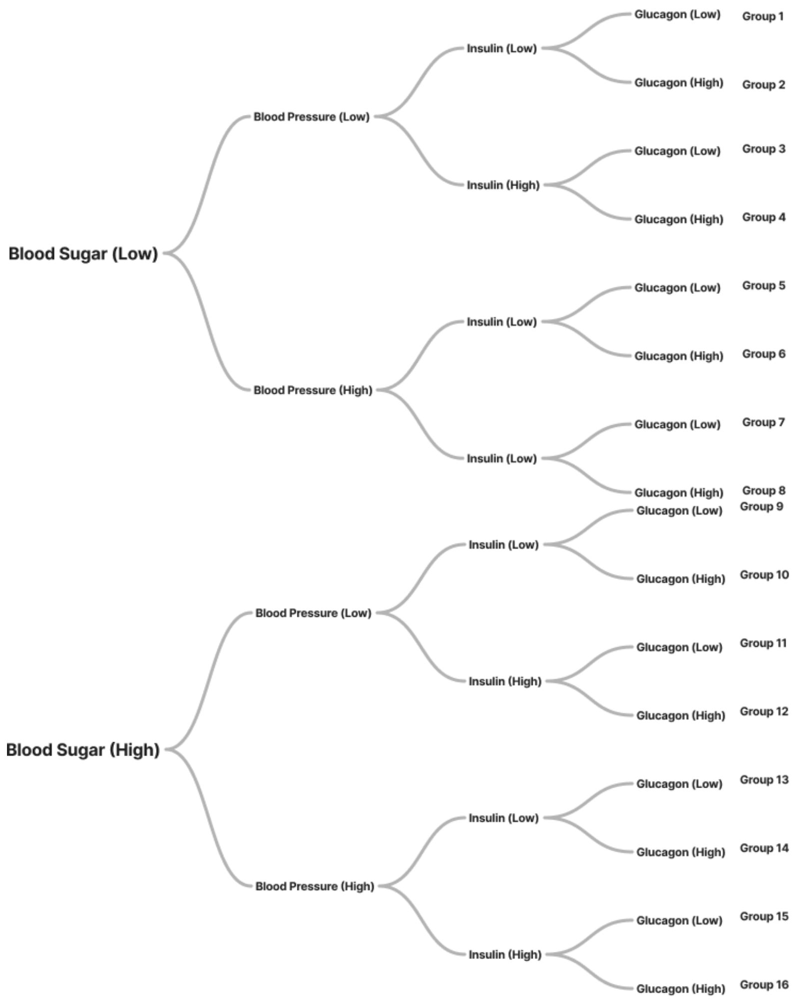


FIGURE 2.1 – Q4.1 Diagram of the 16 groups represented in the experimental design

```
          Df  Pillai approx F num Df den Df    Pr(>F)    
Ins        1 0.96142   498.41      2     40 < 2.2e-16 ***
Gln        1 0.82949    97.29      2     40 4.318e-16 ***
Ins:Gln    1 0.62056    32.71      2     40 3.826e-09 ***
Residuals 41                                             
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

FIGURE 2.2 – Q4.2 MANOVA involving the factors Insulin (Ins) and Glucagon (Gln) with dependent variables Blood Pressure (BP) and Blood Sugar (BS)

the correlation (negative or positive) it can tell if the variables have a proportionality behavior. A positive correlation suggests if one variable is increased, it implies the other one to escalate.Moreover the Pearson correlation permits to obtain the p value, in this case under the Pearson correlation is valuable to weight the null hypothesis H0. The correlation estimated is close to one (cor = 0.98), it is a positive correlation between blood pressure and blood sugar, and the probability associated with the null hypothesis is almost zero (p value $<$ 2.2e-16). It suggests that the blood pressure and the blood sugar are dependent on each other by a linear relation.

Q7 For a significant number of tests the p value is at 1, therefore it got no evidence to reject the null hypothesis. Pearson correlation the limit is set to -1,1, which demonstrates a positive or negative linear

```
        Pearson's product-moment correlation

data:  BS and BP
t = 32.221, df = 43, p-value < 2.2e-16
alternative hypothesis: true correlation is greater than 0
95 percent confidence interval:
 0.9668498 1.0000000
sample estimates:
      cor
0.9799134
```

FIGURE 2.3 – Q6 Pearson correlation between dependent variables

relation between the parameters. The Pearson correlation informs Insulin to have a negative correlation with the dependant variables BS and BP but they are not strong evidence to reject H_0 (p value = 1). With a low p-value for Glucagon, it implies a statistically significant positive linear correlation between Glucagon and the dependant variables BS and BP. The correlation also suggest a positive correlation between the dependant variables and no evidence for an Independence trend between the factors Ins and Gln in reference to a high p-value (=0.99).

TABLE 2.1 – Q7 Data sheet of the Pearson correlation on the experimental hypothesis

Hypothesis	correlation	p value	Observation
Does insulin affect blood sugar ?	-0.88	1	negative correlation Weak evidence against the null
Does insulin affect blood pressure ?	-0.85	1	negative correlation weak evidence against H0
Does glucagon affect blood sugar ?	0.38	0.0042	positive correlation strong evidence against H0
Does glucagon affect blood pressure ?	0.37	0.00621	positive correlation strong evidence against H0
Do insulin and glucagon affect blood sugar independently ?	-0.41	0.9974	negative correlation weak evidence against H0
Do the two factors affect blood pressure independently ?	-0.43	0.9983	negative correlation weak evidence against H0
Are variables blood sugar and blood pressure independent ?	0.98	$<$ 2.2e-16	positive correlation strong evidence against H0

2.3 Analyzing data with bayesian statistics

2.3.1 Analysis of the SPIG dataset

After running the ANOVA in R and Jasp for BS with the factor Ins and Gln, we obtain : (figure 2.4)

By comparing the p-value of both statistical programs, a small p value is obtained by the classical ANOVA. The outcome may be different but they propose the same idea, resulting in strong evidence to reject the null hypothesis. After running the MANOVA test with Pillai statistic on the variable BS and

```
           Df Sum Sq Mean Sq F value   Pr(>F)
Ins         1  4.910   4.910 1021.73  < 2e-16 ***
Gln         1  0.958   0.958  199.36  < 2e-16 ***
Ins:Gln     1  0.313   0.313   65.21 5.23e-10 ***
Residuals  41  0.197   0.005
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

ANOVA - BS ▼

Cases	Sum of Squares	df	Mean Square	F	p
Ins_dose	4.912	2	2.456	1394.824	7.822×10^{-35}
Gln_dose	1.027	2	0.513	291.580	5.768×10^{-23}
Ins_dose ∗ Gln_dose	0.376	4	0.094	53.355	1.212×10^{-14}
Residuals	0.063	36	0.002		

Note. Type III Sum of Squares

FIGURE 2.4 – II.4.1.1 Comparison of R and JASP on ANOVA of the variable BS with the factors Ins and Gln

BP with the factors Ins and Gln we obtain similar results as for ANOVA. Both ANOVA and MANOVA demonstrates the same results for the experiment with interactions between the factors Insulin and Glucagon. (figure 2.5) There are strong evidences (weak p-value) to reject the null hypothesis that the proportion of factors are not equally distributed to the dependent variables.

```
          Df  Pillai approx F num Df den Df    Pr(>F)
Ins        1 0.96142   498.41      2     40 < 2.2e-16 ***
Gln        1 0.82949    97.29      2     40 4.318e-16 ***
Ins:Gln    1 0.62056    32.71      2     40 3.826e-09 ***
Residuals 41
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

MANOVA: Pillai Test

Cases	df	Approx. F	$Trace_{Pillai}$	Num df	Den df	p
(Intercept)	1	50770.365	1.000	2	35.000	2.526×10^{-61}
Ins_dose	2	19.420	1.038	4	72.000	7.113×10^{-11}
Gln_dose	2	17.612	0.989	4	72.000	4.042×10^{-10}
Ins_dose ∗ Gln_dose	4	7.056	0.879	8	72.000	7.671×10^{-7}
Residuals	36					

FIGURE 2.5 – II.4.1.2 Comparison of R and JASP on MANOVA of the variable BS with the factors Ins and Gln

After running the Bayesian test with the null model, depending on the observed data, the probability to obtain the null hypothesis is meaningless. Therefore the Bayes factor (BF10) suggest that the alternative hypothesis such as the model is valid because the alternative hypothesis weight H0 as $BF_{10} = \frac{P(\text{Data}|H_0)}{P(\text{Data}|H_1)}$.(figure 2.6)

Model Comparison

Models	P(M)	P(M|data)	BF_M	BF_{10}	error %
Null model	0.200	1.176×10^{-30}	4.703×10^{-30}	1.000	
Ins_dose + Gln_dose + Ins_dose ∗ Gln_dose	0.200	1.000	$2.176\times10^{+11}$	$8.505\times10^{+29}$	1.611
Ins_dose + Gln_dose	0.200	1.839×10^{-11}	7.354×10^{-11}	$1.564\times10^{+19}$	1.150
Ins_dose	0.200	1.537×10^{-19}	6.146×10^{-19}	$1.307\times10^{+11}$	4.440×10^{-6}
Gln_dose	0.200	2.986×10^{-30}	1.194×10^{-29}	2.539	0.008

FIGURE 2.6 – II.4.1.3 Bayesian ANOVA of BS with factors Ins and Gln

2.3.2 Analysis of SPIG_no_inter

After running the ANOVA program for the file without interactions between Gln and Ins we obtain for both cases a high p-value. The results mean that the data are not strong enough to have evidence to reject the null hypothesis such for Gln_dose. It is interesting to see that the interaction data suggest better evidence to reject the null hypothesis. (figure 2.7)

```
          Df  Sum Sq   Mean Sq F value Pr(>F)
Ins        1 0.00034 0.0003406   0.177  0.676
Gln        1 0.00025 0.0002453   0.128  0.723
Ins:Gln    1 0.00121 0.0012129   0.632  0.431
Residuals 41 0.07871 0.0019198
```

ANOVA no inter ▼

ANOVA - BS ▼

Cases	Sum of Squares	df	Mean Square	F	p
Gln_dose	2.564×10^{-4}	2	1.282×10^{-4}	0.073	0.930
Ins_dose	0.003	2	0.001	0.762	0.474
Gln_dose ∗ Ins_dose	0.014	4	0.004	2.013	0.113
Residuals	0.063	36	0.002		

Note. Type III Sum of Squares

FIGURE 2.7 – II.4.2.1Comparison of R and JASP on ANOVA of the variable BS with the factors Ins and Gln without interactions

As for the MANOVA test, the results are distinct between the programs. In R, the MANOVA test suggests a stronger weight of p-value compared to JASP. In R, the program needs more data to get better evidence to go against the null hypothesis, which in JASP the observed data are enough to conduct a conclusive test. (figure 2.8)

```
          Df    Pillai  approx F num Df den Df Pr(>F)
Ins        1 0.0045409  0.09123       2     40 0.9130
Gln        1 0.0065520  0.13190       2     40 0.8768
Ins:Gln    1 0.0155830  0.31659       2     40 0.7304
Residuals 41
```

MANOVA: Pillai Test

Cases	df	Approx. F	$Trace_{Pillai}$	Num df	Den df	p
(Intercept)	1	305370.725	1.000	2	35.000	5.865×10^{-75}
Ins_dose	2	1.600	0.163	4	72.000	0.184
Gln_dose	2	1.250	0.130	4	72.000	0.298
Ins_dose ∗ Gln_dose	4	2.157	0.387	8	72.000	0.041
Residuals	36					

FIGURE 2.8 – II.4.2.2 Comparison of R and JASP on MANOVA of the variable BS with the factors Ins and Gln without interactions

The Bayesian ANOVA without interaction indicates a Bayes factor of 0.645 for the null model, so the observed data are not suitable to obtain evidence to reject the null hypothesis. Therefore frequentist statistics relies on p-values, whereas Bayesian statistics employs Bayes factors. Then for both methods, the results show evidence to accept the null hypothesis or model.(figure 2.9)

Bayesian ANOVA no inter ▼

Model Comparison ▼

Models	P(M)	P(M|data)	BF_M	BF_{10}	error %
Null model	0.200	0.646	7.307	1.000	
Ins_dose	0.200	0.176	0.854	0.272	0.009
Gln_dose	0.200	0.111	0.501	0.172	0.008
Ins_dose + Gln_dose + Ins_dose ∗ Gln_dose	0.200	0.037	0.152	0.057	1.943
Ins_dose + Gln_dose	0.200	0.030	0.124	0.046	1.168

FIGURE 2.9 – II.4.2.3 Bayesian ANOVA without interactions between Insulin and Glucagon