University of Tehran
School of Electrical and Computer Engineering

Statistical Inference

Instructor: Dr. Hossein Vahabie

Assignment 3

Statistical Inference Tests

Teaching Assistants:

Mohammad H.K Nili [mohammad.nili@ut.ac.ir]

Amir Ali Soltani [aa.soltanitehrani@gmail.com]

Sarmad Zandi [sarmadzandi@ut.ac.ir]

Winter 2022

Homework 3

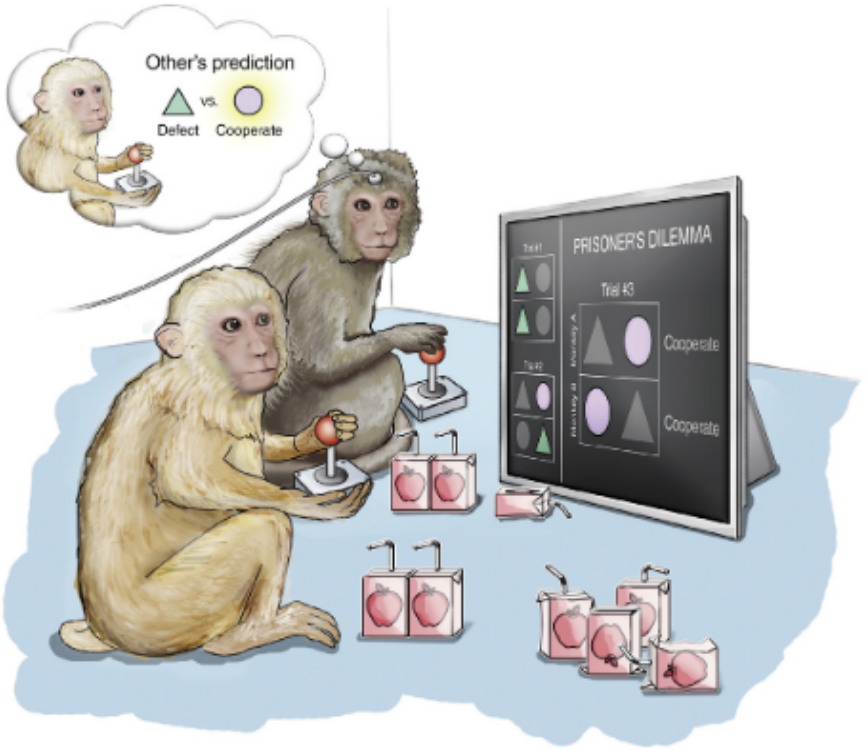
Statistical Inference, Winter 2022

Question 1- Gamer monkeys

A common approach in the comparative study of behavior and cognition is to develop standardized experimental paradigms that can be used across species, with the assumption that if the same task is being used, we can directly compare behavioral responses. This experimental approach rests on two underlying assumptions: first, that different species' perceptions of and effective responses to these paradigms are the same; and second, that behavioral and physiological (including endocrine and neural) responses to these paradigms are homologous; if either of these assumptions is not true, then the comparison becomes much less straightforward.[1]

There are such cognitive, and emotional paradigms in light of cooperative decision making. Affective correlates of cooperation are divided as below:

- **Physiological measures of effect** such as Cardiac function, Skin conductance, Facial skin temperature, Pupil mimicry, Hormones
- **Behavioral measures of effect** such as Cognitive Bias paradigms, Vocalizations, Facial displays, and body language



a) Design a Game Theoretic Test to find the significance of each feature based on cooperation and or each of bellow question assume a Test with a subset of features.
   i) Check the conditions and reject Null Hypothesis.
      (1) H_0: Perceptions and emotional reactions of different species to the mentioned paradigms are the same
      (2) H_0: behavioral and physiological (including endocrine and neural) responses to these paradigms are homologous
   ii) If we could reject both of the above Null hypotheses which type of parametric/Non-parametric test is more reliable to find out
   iii) Compare the distribution of rewards (apple juice) each monkey received.
   iv) Which distribution is better for reward assignment to increase coordination between monkeys?

b) Now change the game conditions to a league between about 70 monkeys (40 male, 30 female)
   i) How we can track any change in monkeys' game strategy after the second game?
   ii) Is the amount of apple juice received the same between men and women or not? which test can we use?
   iii) Is the proportion of monkeys with ages about 3 years equally between each gender group?

c) Assume There is n monkey and k joystick in the league. each joystick has a response delay, which affected consequence actions (collaboration or cooperation) if in this league each monkey plays with all of the joysticks, can we sort the joysticks by their response delay based on the monkey's decision without knowing anything about response delay distribution?

Question 2- Comparing more than two means

(R) The dataset "Diet" contains information about 79 people with different health backgrounds who began to undertake 3 different types of diets (referred to as diet A, B, and C). This exercise aims to see which diet is best for losing weight.

First and foremost look at the data, defining a new column weight loss, corresponding to the difference between the initial and final weights (respectively the corresponding to the columns' initial weight and final weight of the dataset).

a) Plot the three group's data using side-by-side boxplots.
b) Test whether there is significant evidence of a difference in the mean weight loss between groups using ANOVA.
c) Now repeat part b in R and compare the results.
d) Display and analyze the results.
e) Compare the mean of weight loss for every two groups and:
   i) Write the hypothesis.
   ii) Report the level of significance of the test and the decision about the hypothesis.
   iii) Estimate the size of the difference in the mean drop, Use α=0.05.
f) Now Compare the mean of weight loss for every two groups in R and compare it with the results of part d. (hint: You are allowed to use libraries for this part.)

Homework 3

Statistical Inference, Winter 2022

Question 3- Testing Goodness of Fit

Let X be a binomial random variable with n trials and probability p of success.

a) Suppose $n = 100$ and $X = 38$. Compute the Pearson chi-square statistic for testing the goodness of fit to the multinomial distribution with two cells with $H_0: p = p_0 = 0.5$

b) What is the approximate distribution of the test statistic in (a), under the null Hypothesis H_0.

c) What can you say about the P-value of the Pearson chi-square statistic in using the following table of percentiles for chi-square random variables?

(i.e., $P(\chi^2_3 \leq q.90 = 6.25) = .90$)

df	q.90	q.95	q.975	q.99	q.995
1	2.71	3.84	5.02	6.63	9.14
2	4.61	5.99	7.38	9.21	11.98
3	6.25	7.81	9.35	11.34	14.32
4	7.78	9.94	11.14	13.28	16.42

d) Consider the general case of the Pearson chi-square statistic in (a), where the outcome $X = x$ is kept as a variable (yet to be observed). Show that the Pearson chi-square statistic is an increasing function of $|x - \frac{n}{2}|$.

e) Suppose the rejection region of a test of H_0 is $\{X: |X - \frac{n}{2}| > k\}$ for some fixed known number k. Using the central limit theorem (CLT) as an approximation to the distribution of X, write an expression that approximates the significance level of the test for given k.

Question 4- Distributions comparison test (Parametric)

An analysis of the Iranian national football team speculates that Iranian Football players that are born early in the year have an advantage. This is because the birthdate cutoff for different levels of Football leagues in Iran is January 1st, so Football players who are born in January and February are slightly older than teammates born later in the year. Does this slight age advantage, in the beginning, lead to success later on? A 2018 study examined the birthdate distribution of players in the Persian Gulf Pro League (PGPL), and Iranian football youth league (ages 15-20), for the 2016-2017 season. The number of PGPL players born during the Jan-Mar, Apr-Jun, Jul-Sep, and Oct-Dec of the year is shown in the table below.

	Jan-Mar	Apr-Jun	Jul-Sep	Oct-Dec
Number of PGPL players	147	110	52	50
Percentage of Iranian birth	23.7%	25.9%	25.9%	24.5%

The overall percentage of Iranian birth is also provided for each Category. Is this evidence that the birthdate distribution for PGPL players differs significantly from the national distribution?

a) What are the null and alternative hypotheses?

b) Check the necessary conditions for the Chi-square test and see if the conditions are met?

c) What are the values of the chi-square statistic and P-value for this test?

d) **(R)** Calculate P-value and compare it with part c.

e) Talk about the concept of p-value and state the conclusion in context.

Question 5- Distributions comparison test (Non-Parametric)

A professor of statistics at the Faculty of Electrical and Computer Engineering at the University of Tehran wants to investigate the claim that the IQ of Electrical engineering (EE) students is equivalent to that of Computer engineering (CE) students. He randomly selects a sample of Electrical engineering students and a sample of Computer engineering students at the University of Tehran, and after taking a test, calculates the IQ of individuals. These values are listed in the table below:

CE-IQ	125	127	126	126	105	128	127	126	79	124
EE-IQ	131	129	77	52	134	133	94	68	67	132

Investigate this claim using the non-parametric Mann-Whitney U test (Use 5% significant level).

a) What are the null and alternative hypotheses?

b) Calculate the rank of each of the values.

c) What are the values of the U statistic and P-value for this test?

d) **(R)** Calculate P-value and compare it with part c.

e) Talk about the concept of p-value and state the conclusion in context.

Homework 3

Statistical Inference, Winter 2022

Question 6- A Student one-sided test

Let $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$ are i.i.d random variables for some unknown parameter $(\mu, \sigma^2) \in R \times (0, \infty)$. We want to test the following hypotheses at a non-asymptotic level α (for some fixed $\alpha \in (0, 1)$):

$$H_0: \mu > 0 \quad \text{vs.} \quad H_1: \mu \leq 0$$

a) Recall the maximum likelihood estimator $(\hat{\mu}, \hat{\sigma}^2)$ of (μ, σ^2)

b) Let $S = \sqrt{n-1}\frac{\hat{\mu}-\mu}{\sqrt{\hat{\sigma}^2}}$. Prove that S is a Student random variable with $n-1$ degrees of freedom.

c) Propose a test with non-asymptotic level α. Prove your answer.

Question 7- Non Parametric test

a) Two different groups of individuals were compared with respect to how quickly they responded to a changing traffic light. Test using the two-sided Wilcoxon statistic the hypothesis that there is no difference between the two groups on the basis of the following data

Group 1	19.0	14.4	18.2	15.6	14.5	11.2	13.9	11.6
Group 2	12.1	19.1	11.6	21.0	16.7	10.1	18.3	20.5

b) Do the data suggest that the population variances differ? Carry out a Siegel-Tukey test at the 5% significance level.

c) Suppose that those same groups are compared at a later time. Indicate how you would combine the two data sets using the smooth embedding approach in order to define a single test.

Question 8- Non Parametric test

Alex Ozoto wants to become head of the University of Tehran. His fans made a poll from voters to know how many students in each faculty would vote for him before his talk in IRINN and 3 days after the talk. They do this poll in 6 faculties and here is the result.

Supporters (%)	Faculty A	Faculty B	Faculty C	Faculty D	Faculty E	Faculty F
Before	60	56	80	73	14	32
After	58	58	83	67	17	36

a) Using methods you have learned in the class, determine if there is a significant increase in the number of Alex Ozoto supporters. State your hypothesis, and report statistics, and p values.
b) Implement this test in R and compare the results.
c) Igor Cherishev is Ozoto's opponent. One day Igor revealed news about Ozoto which was about his bribe when he was head of the engineering faculty. Igor's fans took the same polls on social networks about Alex Ozoto. Answer to (a) and (b) based on the new poll answers.

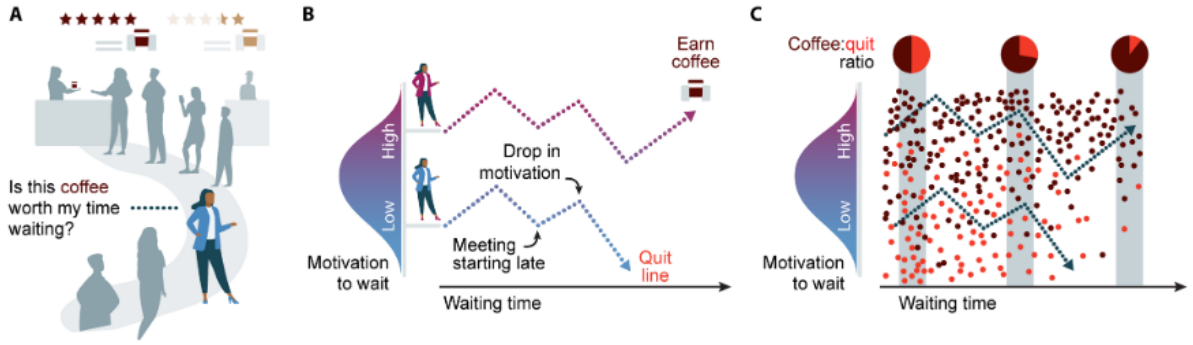
(%)	A	B	C	D	E	F	G	H	I
Before	78	50	40	40	20	50	50	50	50
After	65	20	50	8	16	44	38	40	45

d) Can you relate these two polls? Explain your answer.

Homework 3

Statistical Inference, Winter 2022

Question 9- Uncertainty on Decision Making Under Time Pressure



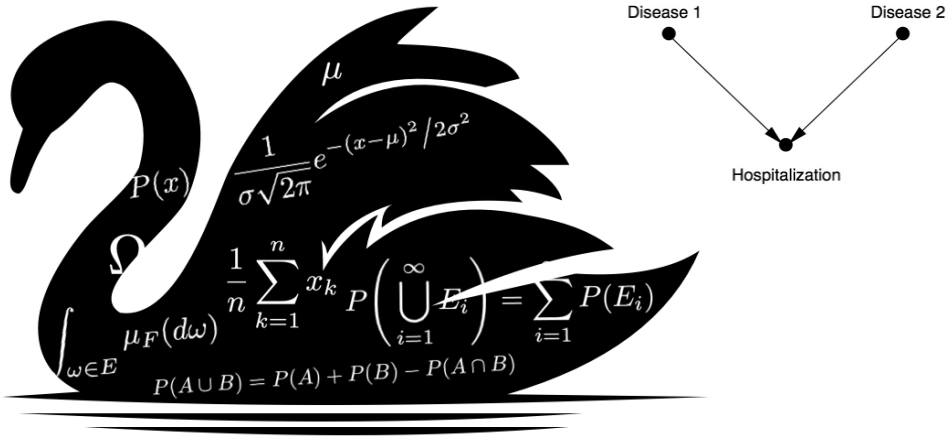
(R) Most of our daily activities in a highly mobile digital society require timely spatial decision-making. Much of such decision-making is supported by maps, however, is always subject to a multitude of uncertainties. If space-time decision-makers are not informed about potential uncertainties, misleading, or at worst, life-threatening outcomes might result from map-based decisions. hypothesized that uncertainty visualization would also have an effect on decision-making under time-critical and complex decision contexts. Using a map-based helicopter landing scenario in mountainous terrain, we found that neither time pressure nor uncertainty affected participants' decision-making accuracy. However, uncertainty affected participants' decision strategies, and time pressure affected participants' response times. the results as consistent with a loss-aversion heuristic. [2] "The data that you need has been attached (HW3_Q9.zip)"

a) Compare the confidence between the control group and time pressure group with a suitable test? plot the results.
b) Have all of the uncertain subgroups similar confidence mean?
c) Compare the proportion of correct answers between each of the uncertain subgroups of control versus time pressure.

Homework 3

Statistical Inference, Winter 2022

Question 10- Correlation and Causation



In 1946, Joseph Berkson, a biostatistician at the Mayo Clinic, pointed out a peculiarity of observational studies conducted in a hospital setting: even if two diseases have no relation to each other in the general population, they can appear to be associated among patients in a hospital.[3]
where two items that seem correlated to the general people are actually not correlated in reality. In statistical terms, it means that even when two values are statistically negatively correlated it may seem that they are positively correlated.[4] To understand Berkson's observation, let's start with a causal diagram. It's also helpful to think of a very extreme possibility: neither Disease 1 nor Disease 2 is ordinarily severe enough to cause hospitalization, but the combination is. In this case, we would expect Disease 1 to be highly correlated with Disease 2 in the hospitalized population.

	General Population			Hospitalized in Last Six Months		
Respiratory disease?	Bone disease?			Bone disease?		
	Yes	No	% Yes	Yes	No	% Yes
Yes	17	207	7.6	5	15	25.0
No(Control)	184	2,376	7.2	18	219	7.6

a) Which Test is better to show the relationship between the 2 groups in the above table?
b) Is this relationship a causal relationship? if no, in which condition we can conclude a causal relationship between them?

Appendix

[1] Haroush, K., & Williams, Z. M. (2015). Neuronal Prediction of Opponent's Behavior during Cooperative Social Interchange in Primates. In Cell (Vol. 160, Issue 6, pp. 1233–1245). Elsevier BV.https://doi.org/10.1016/j.cell.2015.01.045

[2] Korporaal, M., Ruginski, I. T., & Fabrikant, S. I. (2020). Effects of Uncertainty Visualization on Map-Based Decision Making Under Time Pressure. In Frontiers in Computer Science (Vol. 2). Frontiers Media SA. https://doi.org/10.3389/fcomp.2020.00032

[3] Pearl, J., & Mackenzie, D. (2019). *The book of why*. Penguin Books.

[4] Berkson, J. (1946). Limitations of the Application of Fourfold Table Analysis to Hospital Data. In Biometrics Bulletin (Vol. 2, Issue 3, p. 47). JSTOR. https://doi.org/10.2307/3002000

[5] Kühberger, A., Fritz, A., Lermer, E., & Scherndl, T. (2015). The significance fallacy in inferential statistics. In BMC Research Notes (Vol. 8, Issue 1). Springer Science and Business Media LLC. https://doi.org/10.1186/s13104-015-1020-4

Choosing a Statistical Test - [For your information]

How to Choose a Test?

Based on the five questions

- How many variables do you have?
- What is your statistical objective?
- What scales are the variables measured with?
- Is there a distinction between dependent and independent variables?
- Are the samples autocorrelated?

Based on the data

- Does data match an expected ratio?
- Is there an association between two variables?
- Do samples come from the same or different populations?

» One of the most important steps before running a statistical test is choosing the right test[s]. (If you need more information you can refer to the following links.)

- https://www.intro2r.info/unit3/which-test.html
- https://www.youtube.com/watch?v=rulIUAN0U3w
- https://www.lkouniv.ac.in/site/writereaddata/siteContent/202003241550010566rajeev_pandey_parametric_test.pdf
- https://www.scribbr.com/statistics/statistical-tests/
- https://jmyao17.github.io/Statistics/Parameter_Statistical_Tests.html

Taxonomy of Statistical Tests - [For your information]

Basic Statistical Tests:

- Parametric:
  - Descriptive Statistics
  - One-Sample T-Test
  - One-Sample Z-Test
  - Independent Samples T-Test
  - Paired Samples T-Test
  - One-Way ANOVA
  - Factorial ANOVA
  - MANOVA
- Non-parametric:
  - Sign-test
  - Wilcoxon
  - Mann-Whitney
  - Kruskal-Wallis
  - Mood's median test
  - Friedman-test

Advanced Statistical Tests:

- Two or Three Factor Crossed ANOVA
  - Post Hoc Tests
- ATI Analysis (ANCOVA with Interaction)
- MANOVA and Follow-up Discriminant Analysis with at Least Three Groups and Four Independent Variables
  - Post Hoc Tests

Post Hoc Tests:

- Bonferroni test
- Duncan Multiple Range test
- Holm-Bonferroni method
- Least significant Difference test (LSD)
- Newman-Keuls test
- Scheffe test
- Turkeys' or Honest significant Difference test
- Dunnetts' test
- Benjamini-Hochberg test