Statistics for Social Sciences I

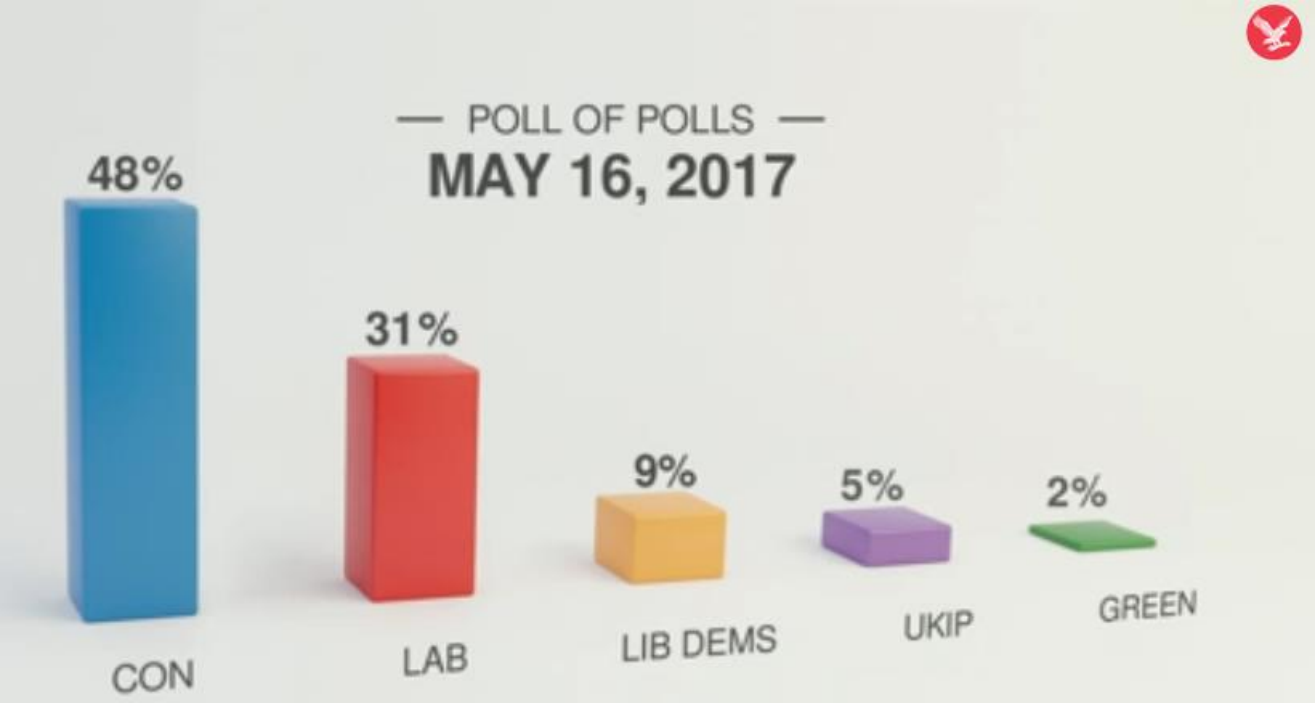
Exam: 31st May 2017

Name: _____

Group: _____ **Date:** _____

PROBLEM 1.

a) The following graphic taken from **The Independent** newspaper reflects the results of a latest poll of polls (May 16th 2017) for voting intentions in the next UK elections of June 8th 2017.



Assuming these results are representative of public opinion, what is the probability (to 4 decimal places) that two out of a set of three, independently selected, UK adults have the intention of voting Conservative?

i. 0.1198.
ii. 0.3594.
iii. 0.9600.
iv. 1.4800.

Solution (1 point)			
i)	**ii) X**	**iii)**	**iv)**

b) The UK Community Life Survey 2016-2017 included the following question.

> To what extent do you agree or disagree with the following statement?
>
> When people in this area get involved in their local community, they really can change the way that their area is run.
>
> 1. Definitely agree
> 2. Tend to agree
> 3. Neither agree nor disagree
> 4. Tend to disagree
> 5. Definitely disagree

In this case, the variable in the question is:

i. Quantitative and discrete.
ii. Qualitative and nominal.
iii. Qualitative and ordinal.
iv. None of the above.

Solution (1 point)			
i)	**ii)**	**iii) X**	**iv)**

c) If the covariance between two sets of data is equal to 1 then:

i. The data follow an exact, increasing linear relationship.
ii. The data follow an approximate, increasing linear relationship
iii. A regression line relating the two sets of data always has a positive intercept
iv. None of the above.

Solution (1 point)			
i)	**ii)**	**iii)**	**iv) X**

d) In the latest CIS barometer of April 2017, the President of the government, Mariano Rajoy achieved a mean rating of 2.91 with a standard deviation of 2. If the ratings are normally distributed with this mean and variance, what is the probability (rounding to 4 decimal places) that a randomly chosen person rates Mariano Rajoy at less than 2.91?

i. 0.0000.
ii. 0.1995.
iii. 0.5000.
iv. 0,9981.

Solution (1 point)			
i)	**ii)**	**iii) X**	**iv)**

e) The following table gives the numbers of foreign tourist arrivals in 2015 for the 10 most popular world tourist destinations.

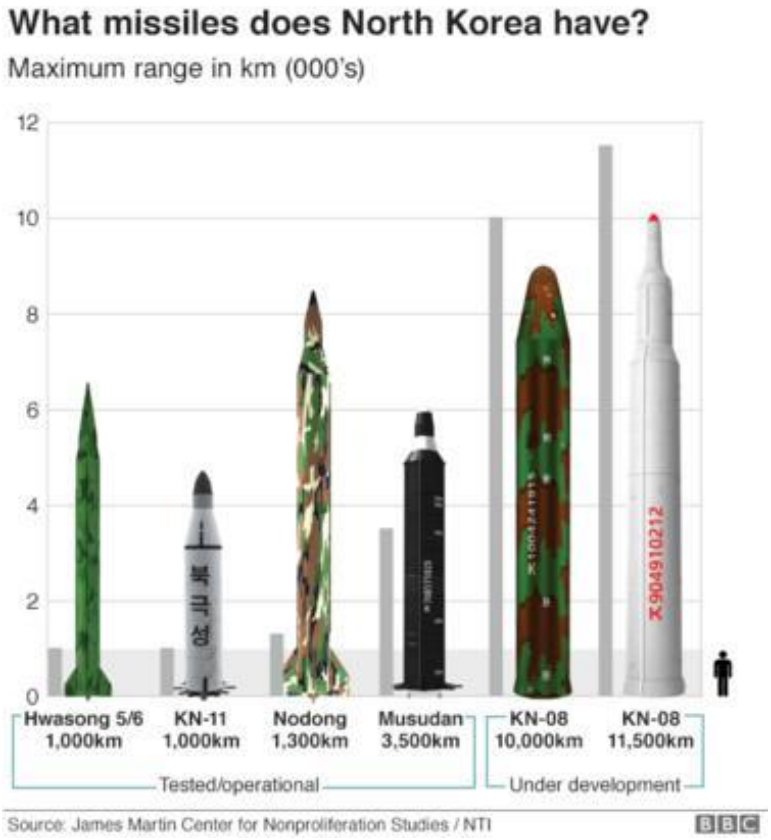
Ranking	Country	Continent	International tourist arrivals
1	France	Europe	84.5 million
2	United States	North America	77.5 million
3	Spain	Europe	68.2 million
4	China	Asia	56.9 million
5	Italy	Europe	50.7 million
6	Turkey	Europe	39.4 million
7	Germany	Europe	35.0 million
8	United Kingdom	Europe	34.4 million
9	Mexico	North America	32.1 million
10	Russia	Europe	31.3 million

For these data:

i. The median number of tourist arrivals is 45.05 millions and the range (in millions) is [31.3, 84.5].
ii. The mean number of tourist arrivals is 51 millions and the first quartile is 33.825 millions.
iii. The median number of tourist arrivals is 51 millions and range is 53.2 millions.
iv. None of the above.

Solution (1 point)			
i)	**ii) X**	**iii)**	**iv)**

PROBLEM 2. The following graphic appeared in an article on the North Korean missile programme that appeared on the BBC web page on 20th April 2017



Comment briefly on the adequacy of the graph.

At first sight, the graph appears to be a pictogram representing distances using the heights of missiles, which would of course be wrong. However, looking more closely, we can see that in reality this is two graphs in one. Distances are represented by the rather faint bars and missile sizes are represented by the comparison with the size of a typical human being as illustrated on the right. Thus, the graph provides information on both these aspects, although for clarity and to ease confusion it might be preferable to separate the graphs in two or represent the missiles below the x axis. On a positive note, in an article about the Korean missile program, the missile images bring the reader into the context of the article.

(1 point)

PROBLEM 3. The following table comes fom a survey by the Centre d'Estudis d'Opinió de la Generalitat de Cataluña en junio de 2016. The responses: yes (*si*), no, don't know (*no lo sap*) or no reply (*no contesta*) to the question Should Cataluña be an independent state are classified according to sex: male (*home*) or female (*dona*) and according to age (*edat*). The *base* represents the number of respondents in each category.

	Total	Sexe		Edat				
		Home	Dona	18 24	25 34	35 49	50 64	65+
31. I més concretament, "Vol que Catalunya esdevingui un Estat independent"?								
Base: Total entrevistes	**1500**	**712**	**788**	**117**	**215**	**451**	**345**	**372**
Sí	**47,7**	50,1	45,4	59,0	>58,1	48,3	43,8	40,9
No	**42,4**	40,3	44,3	<29,9	<32,6	39,5	47,2	>51,1
No ho sap	**8,3**	8,0	8,5	7,7	7,9	9,5	8,1	7,3
No contesta	**1,7**	1,5	1,8	3,4	1,4	2,7	0,9	0,8

a) If one of the people in the survey is chosen at random, what is the probability that they are male?

712/1500 ≈ 0,4747

(0, 5 points)

b) Assuming they are male, what is the probability that they think Cataluña should be independent?

0,501

(0. 5 points)

c) If the chosen person thinks that Cataluña should be independent, calculate the probability that they are male.

P(male|independent) = P(independent|male)P(male)/P(independent) = 0,501 x0,4747 / 0,477 ≈ 0,4985

(0. 5 points)

d) Thinking of the wider Catalan population, is there any evidence from the table that the two variables Age and Opinion about whether Cataluña should be independent are dependent? Comment your answer briefly but do not perform a formal statistical analysis.

There is a clear relationship between age and independence as we can see that the estimated probability of being in favour decreases from 0,59 in the 18-24 age group to 0,409 in the over 65 age group. If there were no relationship, we would expect the estimated probability of independence to be approximately constant over the different age groups.

(0. 5 points)

PROBLEM 4. The following headline and article concerning opinions on the Roma or gypsy community appeared in **The Local** an Italian online news organization in May 2015.

Nine in ten Italians don't like Roma: survey

A vast majority of Italians hold anti-Roma views, according to a new study published by the Pew Research Centre.

The study compared views on ethnic minority groups, including Muslims and Jews, across European countries. When asked whether they had favourable or unfavourable attitudes towards Roma, 86 percent of Italians said they were unfavourable.

Anti-Roma sentiment in Italy is significantly stronger than in any other European country surveyed. In Spain, the UK and Germany more than 50 percent of those surveyed expressed positive opinions towards Roma people.

Roma people in Italy face widespread discrimination, and many are forced to live in shanty towns in major cities.

At present, it is estimated that there are 40,000 Italian Roma living in such camps throughout Italy.

This figure emerges at a time when anti Roma feelings are running high in Italy. Last week, two young Roma boys killed a Filipino woman and injured eight more in a hit and run incident in Rome. The event has been covered widely by the Italian press.

In the wake of the event, Matteo Salvini, leader of Italy's far-right Lega Nord Party wrote on Facebook: “Once we're back in the government we will raze these damn Roma camps to the ground!”

Salvini's comments have drawn criticism from pro-Roma groups in Italy, but anti-Roma language is common. A Rome bakery last year drew criticism for a "no gypsies" sign on its door.

The Pew Research Centre report also asked Italians whether they had favourable views of Muslims and Jews. Sixty-one percent responded that they had unfavourable attitudes towards Muslims, while 21 percent were negative towards Jews.

According to Italian government statistics, Italy is home to more than 130,000 Roma people, compared to 650,000 in Spain, 120,000 in Germany and 90,000 in the UK.

Technical data: 1000 face to face interviews with Italian adults

a) Calculate a 90% confidence interval for the true proportion of Italian adults who have unfavourable attitudes towards Muslims. Is there any evidence (at a 10% significance level) that this proportion is not equal to 50%?

The interval is (either using the Excel output or the formula for an interval for a proportion)

0,61 ± 0,0254 ≈ (0,585 0,635)

For the second part of the question, you can perform a formal hypothesis test, or equivalently just say that as 0,5 is not contained in the 90% c.i., then yes there is evidence at a 10% (=100-90) significance level that the true proportion is different to 50%.

(1 point)

b) Is there any evidence (at a 5% significance level) that less than 90% of Italian adults have unfavourable attitudes towards Roma people? Comment with respect to the headline of the article

p = true proportion of Italian adults with unfavourable attitudes to Roma people

Hypotheses:

H0: p = 0,9 **null hypothesis**
H1: p < 0,9 **experimental / alternative hypothesis**

Data

$\hat{p}$ **= 0,86** **sample proportion**
n = 1000 **sample size**
p_0 **= 0,9** **value of p assuming H0 is true**
α = 0,05 **significance level**

Test statistic

$$Z = \frac{\hat{p} - p0}{\sqrt{p0(1-p0)/n}}$$

Value of test statistic

z = -4,216 using the data numbers.

Rejection region.

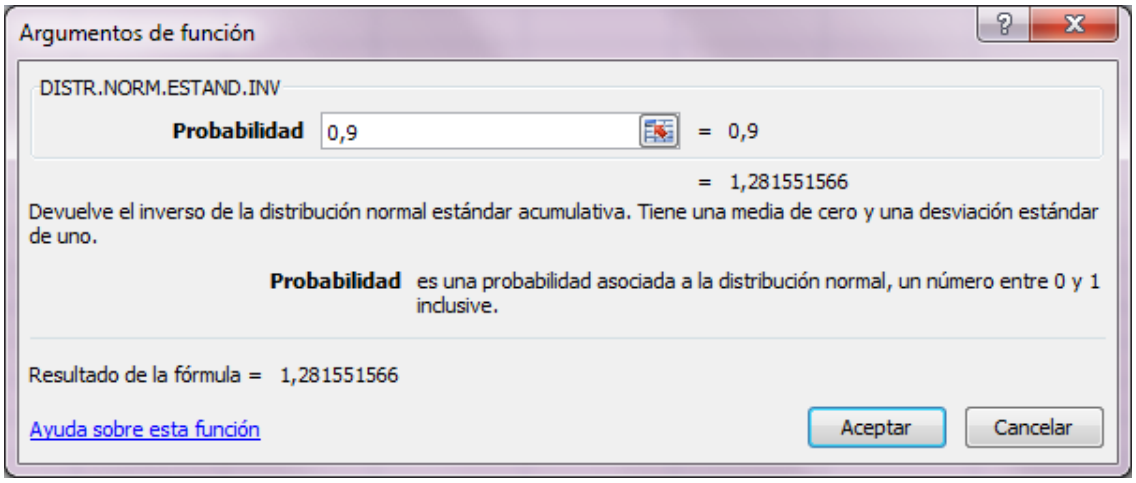
The critical point in this case is z(α = 0,05) = -1,645 so that if z < -1,645 we reject H0. (You can draw the critical region as we did in class). In our case, -4,216 < -1,645 and therefore we reject H0 in favour of H1 at a 5% significance level.

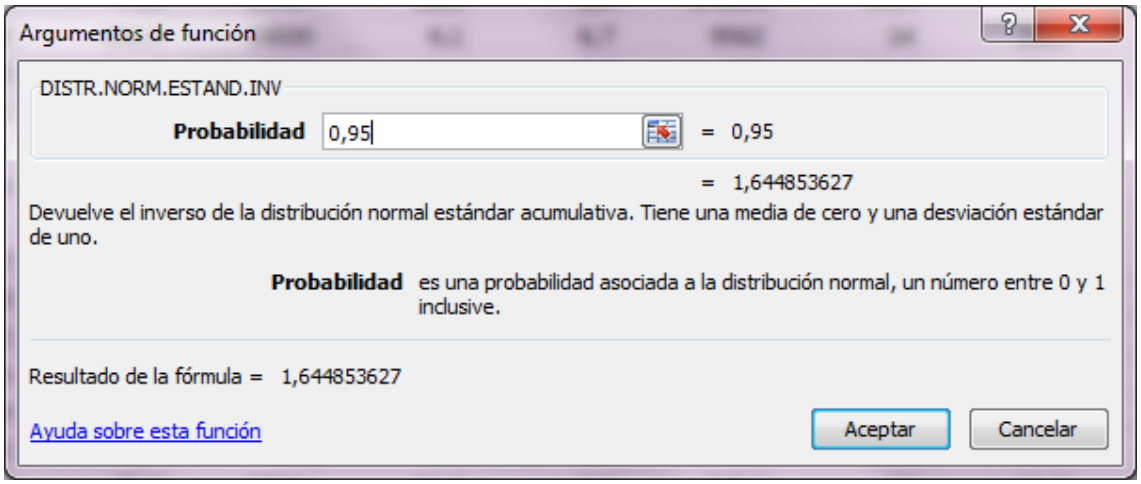
Real words

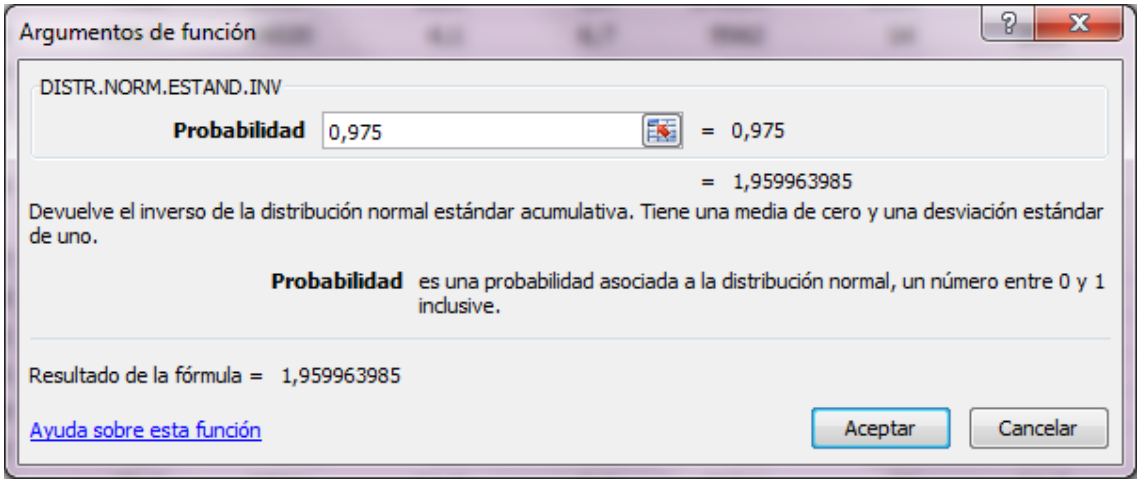
Yes we do have evidence that less than 90% of Italian adults have unfavourable attitudes to Roma people.

(1 point)

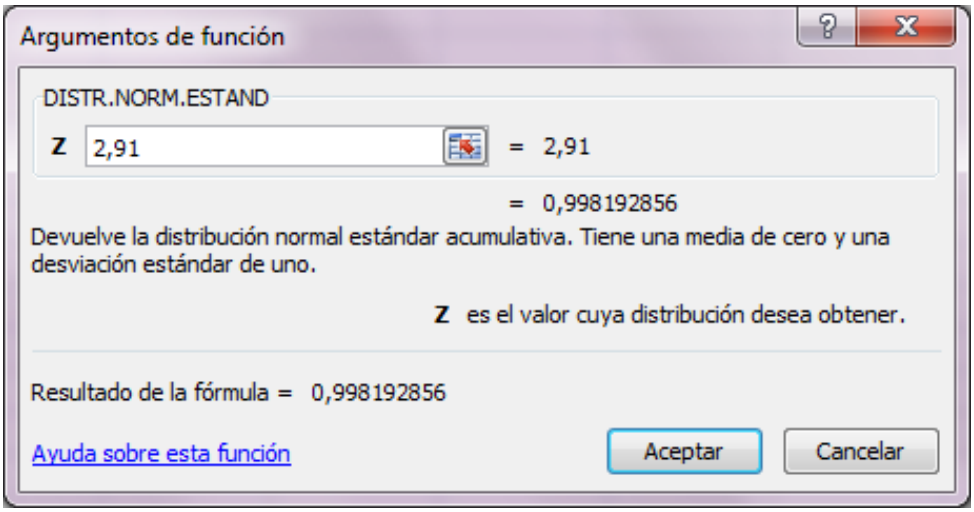
ANNEX: NORMAL PERCENTILES

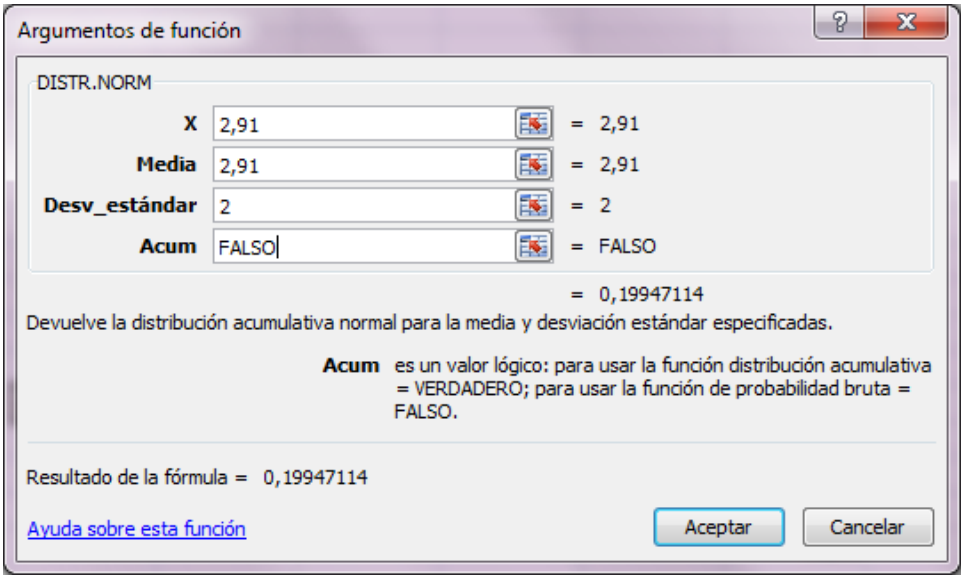


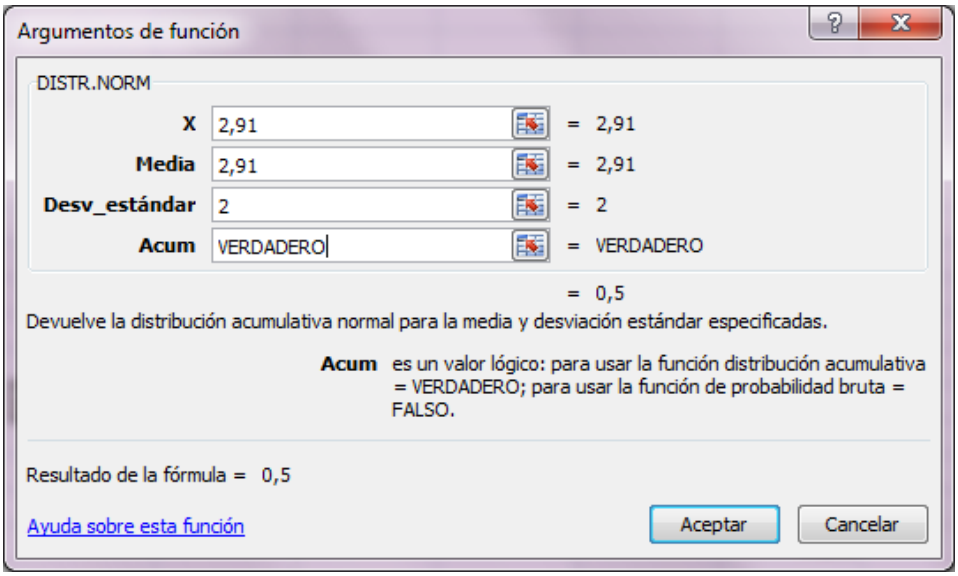




ANNEX: EXCEL OUTPUTS FOR Q 1d)







ANNEX: EXCEL OUTPUTS FOR Q 4