STATISTICS APPLIED TO BUSINESS ADMINISTRATION (BAD) - Second Year Academic Year 2022-23
STATISTICS APPLIED TO MARKETING (MD) - Second Year
STATISTICS APPLIED TO BUSINESS ADMINISTRATION (DD) - Third Year
First Call. May 31, 2023

INSTRUCTIONS

1. The exam contains multiple choice questions that must be answered in the orange code sheet we have provided you with, together with three exercises that must be answered in detail in separate sheets of paper.
2. To select an answer, all you need to do is mark in the orange code sheet, **filling the rectangle in which the selected answer is located appropriately**. Please make sure you know the answer you wish to mark before doing it. You can always cancel your mark by simply filling the rectangle below the answer you wish to cancel, and then fill the rectangle in which the new selected answer is located appropriately. Therefore, we advice you to first mark your selected answers in the exam and to use only the last ten minutes or so from the time assigned to the multiple choice questions-part of the exam to copy them into the code sheet.
3. In the multiple choice questions-part of the exam there is always **only one correct answer** for every question. **Every question correctly answered is worth 1 point, whereas each question incorrectly answered will not penalize your grade in any form. Therefore, you must answer all of the questions included in the exam**.
4. Each one of the exercises, A, B and C, must be answered in a separate sheet of paper. We will collect the different parts of the exam at the indicated times and in this order: first, we will collect the code sheet for the multiple choice questions part of the exam and, then, and in this order, exercises A, B and C will be collected.
5. The exam has six numbered sheets, going from 0.1 to 0.6. Please make sure that you have all sheets and contact your professor if this is not the case. There are different exam types. This exam is of type 0. Mark a 0 in the column labelled with I in your code sheet.
6. The maximum final grade for each of the parts of the exam (i.e., the multiple choice questions part and the exercises part) is 30 points. If you have participated in the on-going evaluation process, you will need to obtain 15 and 12 points in each part of the exam to pass it. Otherwise, 18 and 15 points in each part of the exam are required to pass it.
7. Please fill in your personal information in the appropriate places both in the code sheet and in the sheets provided for the exercises.

MULTIPLE CHOICE QUESTIONS (Time: 1 hour and 50 minutes)

1. FREE-QUESTION. The capital of Spain is:

   (A) Paris (B) Sebastopol (C) Madrid (D) London (E) Pekin

Questions 2 to 4 refer to the following exercise:

The weekly time (in hours) a family devotes to shopping in malls in a given city follows an exponential distribution with parameter $\lambda = 0.50$. It is established that the weekly shopping time for a given family may be considered excessive if it is larger than 3.2 hours. We assume independence in the weekly shopping time for the different families in that city. You need to round up to one decimal place the probability that the weekly shopping time for a given family is excessive.

2. If a random sample of 15 families in that city is taken, the probability that at least 4 of them have an excessive weekly shopping time is:

   (A) 0.7031 (B) 0.6482 (C) 0.3518 (D) 0.8358 (E) 0.1642

3. In the same sample of 15 families, the probability that exactly 12 families **do not** have an excessive weekly shopping time is:

   (A) 0 (B) 0.6482 (C) 0.1319 (D) 0.3980 (E) 0.2502

4. If we now take a random sample of 150 families in that city, the approximate probability that at most 35 of them have an excessive weekly shopping time is:

   (A) 0.1314 (B) 0.5871 (C) 0.8413 (D) 0.4129 (E) 0.8686

Questions 5 and 6 refer to the following exercise:

Let Z be a r.v. having a binomial distribution with mean equal to 3.5 and variance equal to 2.275.

5. $P(0 < Z < 3)$ is:

   (A) 0.2481 (B) 0.2616 (C) 0.0725 (D) 0.5003 (E) 0.1757

6. $P(Z < 1)$ is:

   (A) 0.0135 (B) 0.1346 (C) 0.0860 (D) 0.9865 (E) 0

Questions 7 to 10 refer to the following exercise:

The number of clients arriving, **every ten minutes**, at a shopping mall follows a Poisson distribution with $P(X = 2) = P(X = 1)$. We assume independence between the arrivals of the different clients at the shopping mall.

7. The probability that, in a half an hour period, at most 7 clients arrive at the shopping mall is:

   (A) 0.9489 (B) 0.0511 (C) 0.2560 (D) 0.6063 (E) 0.7440

8. The probability that, in a 40-minute period, at least 5 and at most 9 clients arrive at the shopping mall is:

   (A) 0.6170 (B) 0.5254 (C) 0.4929 (D) 0.0996 (E) 0.7166

9. The most likely number of clients expected to arrive at the shopping mall in a **one-hour period** is:

(A) 11 (B) 12 (C) 10 y 11 (D) 9 y 10 (E) 11 y 12

10. The approximate probability that, in a three-hour period, fewer than 40 clients arrive at the shopping mall is:

(A) 0.2810 (B) 0.4602 (C) 0.7881 (D) 0.7190 (E) 0.5398

Questions 11 and 12 refer to the following exercise:

Let X_1, X_2 and X_3 be three independent r.v., each having an exponential $\exp(\frac{1}{2})$ distribution.

11. If we define the r.v. $Y = X_1 + X_2 + X_3$, the value of k such that $P(k < Y < 14.4) = 0.95$ is:

(A) 0.216 (B) 12.6 (C) 0.352 (D) 1.24 (E) 1.64

12. If we define the r.v. $V = \frac{X_1+X_2+X_3}{3}$, the distribution of the r.v. V is:

(A) $\exp(\frac{3}{2})$ (B) $\gamma(\frac{3}{2}, 3)$ (C) $\gamma(\frac{2}{3}, 3)$ (D) χ^2_3 (E) $\gamma(3, 3)$

Questions 13 to 15 refer to the following exercise:

Let X_1, X_2, X_3, X_4 and X_5 be five independent r.v. such that their distributions are as follows: $X_1 \in N(-1, \sigma^2 = 1)$, $X_2 \in N(0, \sigma^2 = 4)$, $X_3 \in N(2, \sigma^2 = 9)$, $X_4 \in \chi^2_6$ and $X_5 \in \gamma(\frac{1}{2}, 2)$, respectively

13. The value of k such that $P(X_2^2 < k) = 0.75$ is:

(A) 5.28 (B) 0.102 (C) 1.32 (D) 3.84 (E) 2.77

14. If we define the r.v. $V = \dfrac{\sqrt{5}X_2}{2\sqrt{(X_1+1)^2 + X_5}}$, then $P(-0.92 \leq V \leq 3.36)$ is:

(A) 0.58 (B) 0.42 (C) 0.79 (D) 0.80 (E) 0.21

15. If we define the r.v. $W = \dfrac{\left[(X_1+1)^2 + \left(\frac{X_3-2}{3}\right)^2 + X_4\right]}{2\ X_5}$, the approximate value of k such that $P(W > k) = 0.99$ is:

(A) 7.01 (B) 14.8 (C) 0.14 (D) 0.07 (E) 0.36

Questions 16 and 17 refer to the following exercise:

Let X be a discrete r.v. with probability mass function given by:

$$P(X=0) = 1 - \frac{1}{\theta};\quad P(X=1) = \frac{1}{2\theta};\quad P(X=-1) = \frac{1}{2\theta}$$

In order to estimate the parameter θ a r.s. of size $n = 10$ has been taken, providing 4 zeroes.

16. The method of moments estimate of θ is:

(A) $\frac{5}{3}$ (B) $\frac{10}{3}$ (C) $\frac{1}{2}$ (D) $\frac{3}{10}$ (E) $\frac{3}{5}$

17. The maximum likelihood estimate of θ is:

(A) $\frac{3}{5}$ (B) $\frac{10}{3}$ (C) $\frac{5}{3}$ (D) $\frac{3}{10}$ (E) $\frac{1}{2}$

Questions 18 and 19 refer to the following exercise:

Let X be a r.v. having a uniform distribution, so that $X \in U(0, 5\theta)$. In order to estimate the parameter θ a r.s. of size n, $X_1, \ldots, X_n$, has been taken.

18. The method of moments estimator of θ is:

(A) $\frac{\overline{X}}{5}$ (B) $\frac{5}{2\overline{X}}$ (C) All false (D) $5\overline{X}$ (E) $\frac{2\overline{X}}{5}$

19. The maximum likelihood estimator of θ is:

(A) $5 \max(X_i)$ (B) $\frac{\max(X_i)}{5}$ (C) $\overline{X}$ (D) $\frac{\min(X_i)}{5}$ (E) All false

Questions 20 and 21 refer to the following exercise:

Let X be a r.v. having a Poisson $\mathcal{P}(\theta)$ distribution. In order to estimate the parameter θ, a r.s. of size n, $X_1, \ldots, X_n$, has been taken and $\hat{\theta} = (2X_1 + 4X_2 + \ldots + 4X_{n-1} + 3X_n)/(3n+1)$ is proposed as an estimator for θ.

20. The proposed estimator is:

(A) Unbiased (B) Biased and asymptotically unbiased (C) It cannot be determined
(D) Unbiased and asymptotically biased (E) Biased and asymptotically biased

21. The variance of the proposed estimator is:

(A) $\dfrac{\theta(4n-3)}{(3n+1)^2}$ (B) $\dfrac{\theta}{n}$ (C) $\dfrac{\theta(16n-19)}{(3n+1)^2}$ (D) $\dfrac{\theta(16n+19)}{(3n+1)^2}$ (E) $\dfrac{16\theta}{(3n+1)^2}$

Questions 22 and 23 refer to the following exercise:

Let X be a r.v. with probability density function given by:

$$f(x, \theta) = \begin{cases} 1 + \theta(x^3 - \frac{1}{4}), & \text{si } x \in (0, 1) \\ 0, & \text{en otro caso} \end{cases}$$

In order to test the null hypothesis $\theta = 1$ against the alternative hypothesis $\theta = 0$, a r.s. of size $n = 1$ is taken.

22. The most powerful critical region for that observation and for a given significance level is of the form:

(A) $X \leq C$ (B) $X \in (C_1, C_2)^c$ (C) All false (D) $X \in (C_1, C_2)$ (E) $X \geq C$

23. If we decide to reject the null hypothesis if $X < 0.1$, the significance level for this test is:

(A) 0.405 (B) 0.205 (C) 0.925 (D) 0.105 (E) 0.075

Questions 24 and 25 refer to the following exercise:

Let X be a r.v. having a binary distribution with parameter p. In order to test the null hypothesis $H_0 : p \leq 0.40$ against the alternative hypothesis $H_1 : p > 0.40$, a r.s. of size $n = 10$ has been taken, and $Z = \sum_{i=1}^{10} X_i$ is used as the corresponding test statistic.

24. At the $\alpha = 0.10$ significance level, we reject the null hypothesis if:

(A) $Z \geq 7$ (B) $Z \geq 8$ (C) $Z \geq 9$ (D) $Z \leq 9$ (E) $Z \leq 8$

25. The probability of type II error for this test and $p = 0.80$ is:

(A) 0.1209 (B) 0.9672 (C) 0.8791 (D) 0.1074 (E) 0

Questions 26 and 27 refer to the folllowing exercise:

A r.s. of size $n = 18$ from a normal population provides the following sample values: $\overline{x} = 26.82$ and $s^2 = 61.33$, respectively.

26. A 90% confidence interval for the population mean is given by:

(A) $(26.02, 27.62)$ (B) $(23.52, 30.12)$ (C) $(24.92, 28.72)$ (D) $(22.81, 30.83)$ (E) $(23.71, 29.93)$

27. A 95% confidence interval for the population variance is given by:

(A) $(40.00, 127.33)$ (B) $(36.55, 146.02)$ (C) $(35.61, 152.16)$ (D) $(30.20, 150.40)$ (E) $(31.10, 110.12)$

Questions 28 and 29 refer to the following exercise:

A given firm buys eleven hard disks from the H1 brand and another eleven from the H2 brand. We assume that the hard disk prices follow a normal distribution, and that these distributions are independent from each other. These samples provided the corresponding sample mean prices and standard deviations as follows: $\overline{x}_{H1} = 60.50$, $\overline{x}_{H2} = 70.10$, $s_{H1} = 8.40$ y $s_{H2} = 10.20$.

28. A 90% confidence interval for the ratio of the variances, $\sigma^2_{H1}/\sigma^2_{H2}$ is:

(A) $(0.2276, 2.0210)$ (B) $(0.2923, 1.5734)$ (C) $(0.2380, 1.9329)$
(D) $(0.3550, 1.9106)$ (E) $(0.2764, 2.4541)$

29. If we assume that the variances are equal, a 95% confidence interval for the difference of the mean prices for the two brands of the hard disks, $m_{H1} - m_{H2}$, is:

(A) (-9.60 ± 2.17) (B) (9.60 ± 4.18) (C) (-9.60 ± 8.73) (D) (-9.60 ± 1.25) (E) (9.60 ± 1.22)

30. If we wish to estimate the variance of a given normal population by confidence interval and increase the interval's confidence level, without changing the sample size, the confidence interval will be:

(A) Narrower (B) It will not change (C) All false
(D) It cannot be determined (E) Wider

EXERCISES (Time: 65 minutes)

A. (10 points, 15 minutes)

The following table describes the probability mass function for the discrete random variable X under the null ($P_0(x)$) and alternative ($P_1(x)$) hypotheses.

X	1	2	3	4	5	6
$P_0(x)$	0	0.05	0.05	0.10	0.40	0.40
$P_1(x)$	0.10	0.10	0	0.45	0.20	0.15

In order to test the null hypothesis $H_0 : P(x) = P_0(x)$ against the alternative hypothesis $H_1 : P(x) = P_1(x)$, a random sample of size $n = 1$ has been taken. We consider two possible critical regions: $\text{CR}_1 = \{1, 2, 3\}$ and $\text{CR}_2 = \{1, 4\}$.

i) For both critical regions, compute the significance level, the probability of type II error and the power of the test.

ii) Which of the two critical regions defined above is more appropriate for this test? **Remark**: The selection of the most appriate critical region should be adequately justified, which will allow us to discard or select the most appropriate critical region for this test.

B. (10 points, 25 minutes)

Let X be a random variable with probability density function given by:

$$f(x;\theta) = \begin{cases} \dfrac{1}{(\theta-1)}\, e^{-\frac{1}{(\theta-1)}x} & \text{if } x > 0, \quad \theta > 1 \\ 0 & \text{otherwise} \end{cases}$$

In addition, it is known that:

$$E(X) = \theta - 1$$
$$\text{Var}(X) = (\theta - 1)^2$$

In order to estimate the parameter θ, a random sample of size n, $X_1, \ldots, X_n$, has been taken.

i) Find, **providing all relevant details**, the maximum likelihood estimator of the parameter θ.

ii) Is this an unbiased estimator of θ? Is it consistent? Is it efficient? Provide all relevant details to justify your answers.

Remark: The Cramer-Rao lower bound for an unbiased and regular estimator of θ for a given r.s. is:

$$L_c = \frac{1}{n\text{E}\left[\frac{\partial \ln f(X,\theta)}{\partial\theta}\right]^2} == \frac{1}{-n\text{E}\left[\frac{\partial^2 \ln f(X,\theta)}{\partial\theta^2}\right]}$$

C. (10 points, 25 minutes)

A firm wishes to sell four types of mobile telephones and, for this purpose, it has been able to collect the following information:

Brand	A	B	C	D
Probabilities	θ^2	$(1-\theta)^2$	$\theta(1-\theta)$	$\theta(1-\theta)$

i) In order to be able to estimate these probabilities, a random sample of 80 individuals has been taken and has provided the following information: 12 individuals bought brand A mobile telephones; 38 individuals bought brand B mobile telephones; 16 individuals bought brand C mobile telephones and 14 individuals bought brand D mobile telephones. Find the maximum likelihood estimator of θ.

ii) At the 5% significance level, test the hypothesis that the probability distribution the firm has is the correct one.

SOLUTIONS TO MULTIPLE CHOICE QUESTIONS (exam type 0)

1: C	11: D	21: C
2: C	12: B	22: A
3: E	13: A	23: E
4: E	14: C	24: A
5: A	15: C	25: A
6: A	16: A	26: B
7: E	17: C	27: B
8: A	18: E	28: A
9: E	19: B	29: C
10: D	20: E	30: E

SOLUTIONS TO EXERCISES

Exercise A

We wish to test the null hypothesis that X is a discrete random variable with probability mass function $P_0(x)$ against the alternative that its probability mass function is $P_1(x)$:

X	1	2	3	4	5	6
$P_0(x)$	0	0.05	0.05	0.10	0.40	0.40
$P_1(x)$	0.10	0.10	0	0.45	0.20	0.15

A random sample of size $n = 1$ has been taken; that is, we observe X. In addition, we consider the two critical regions $\text{CR}_1 = \{1, 2, 3\}$ and $\text{CR}_2 = \{1, 4\}$. As we already know,

$$\alpha = P(I) = P(X \in CR|H_0)$$

$$\beta = P(II) = P(X \notin CR|H_1)$$

$$\text{Power} = P(X \in CR|H_1) = 1 - \beta$$

Therefore, if we compute these values for the two critical regions we are considering, we will have that:

$$\alpha_1 = P(X = 1, 2, 3|P_0) = 0 + 0.05 + 0.05 = 0.10$$

$$\alpha_2 = P(X = 1, 4|P_0) = 0 + 0.10 = 0.10$$

$$\beta_1 = P(X = 4, 5, 6|P_1) = 0.45 + 0.20 + 0.15 = 0.80$$

$$\beta_2 = P(X = 2, 3, 5, 6|P_1) = 0.10 + 0 + 0.20 + 0.15 = 0.45$$

$$\text{Power}_1 = P(X = 1, 2, 3|P_1) = 0.10 + 0.10 + 0 = 0.20$$

$$\text{Power}_2 = P(X = 1, 4|P_1) = 0.10 + 0.45 = 0.55$$

From the results we have obtained above, we conclude that, given that both tests have the same value for the probability of type I error and that the test for critical region CR_2 has higher power than the one for critical region CR_1 (and, thus, smaller probability of type II error), the test based on critical region CR_2 is more appropriate than the one based on critical region CR_1. In addition, we have to mention that the point $x = 1$ should be included in all critical regions because, as it has probability zero under the null hypothesis, it represents a clear point for the rejection of the null hypothesis. However, the point $x = 3$, as it has probability zero under the alternative hypothesis and positive probability under the null hypothesis, it should never be included in the critical region. Along these lines, the critical region RC_1 should be really discarded and not being selected for this test.

Exercise B

$$f(x;\theta) = \begin{cases} \dfrac{1}{\theta - 1}\, e^{-\frac{1}{\theta-1}x} & \text{if } x > 0, \quad \theta > 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\text{E}(X) = \theta - 1$$

$$\text{Var}(X) = (\theta - 1)^2$$

i) **Maximum likelihood estimator**

$$L(\vec{X};\theta) = f(X_1;\theta)\ldots f(X_n;\theta) = \frac{1}{(\theta-1)}\, e^{-\frac{1}{(\theta-1)}X_1} \ldots \frac{1}{(\theta-1)}\, e^{-\frac{1}{(\theta-1)}X_n} =$$

$$= \frac{1}{(\theta-1)^n}\ e^{-\frac{1}{(\theta-1)}\sum_{i=1}^n X_i}$$

$$\ln L(\vec{X};\theta) = -n\ln(\theta-1) - \frac{1}{(\theta-1)}\sum_{i=1}^n X_i$$

$$\frac{\partial \ln L(\vec{X},\theta)}{\partial\theta} = -\frac{n}{(\theta-1)} + \frac{\sum_{i=1}^n X_i}{(\theta-1)^2} = 0$$

$$\frac{n}{(\theta-1)} = \frac{\sum_{i=1}^n X_i}{(\theta-1)^2}$$

$$n(\theta-1) = \sum_{i=1}^n X_i$$

$$\hat{\theta}_{\text{ML}} = \frac{\sum_{i=1}^n X_i}{n} + 1 = \overline{X} + 1$$

ii) We wish to check if the estimator $\hat{\theta}_{\text{ML}} = \overline{X}+1$ is unbiased, consistent and efficient.

Unbiasedness.

$$\text{E}\left[\hat{\theta}_{\text{ML}}\right] = \text{E}\left(\overline{X}+1\right) =$$
$$= \text{E}(\overline{X}) + 1 = \text{E}(X) + 1 = (\theta-1)+1 = \theta$$

Thus, the estimator is unbiased.

Consistency. This is a consistent estimator because the two sufficient conditions hold. That is,

1) θ is an unbiased estimator and

2) $\lim\limits_{n\to\infty}\left[\text{Var}\left(\hat{\theta}_{\text{ML}}\right)\right] = \lim\limits_{n\to\infty}\frac{(\theta-1)^2}{n} = 0$, because:

$$\text{Var}\left(\hat{\theta}_{\text{ML}}\right) = \text{Var}\left(\overline{X}+1\right) = \text{Var}\left(\overline{X}\right) = \frac{\text{Var}(X)}{n} = \frac{(\theta-1)^2}{n}$$

Efficiency. To be able to verify if this estimator is efficient, we compute the Cramer-Rao lower bound.

$$Lc = \frac{1}{nE\left[\frac{\partial \ln f(X,\theta)}{\partial\theta}\right]^2}$$

$$f(x;\theta) = \frac{1}{(\theta-1)}\ e^{-\frac{1}{\theta-1}x}$$

$$\ln f(x;\theta) = -\ln(\theta-1) - \frac{1}{(\theta-1)}x$$

$$\frac{\partial \ln f(x;\theta)}{\partial\theta} = -\frac{1}{(\theta-1)} + \frac{x}{(\theta-1)^2}$$

$$E\left[\frac{\partial \ln f(X;\theta)}{\partial\theta}\right]^2 = E\left[-\frac{1}{(\theta-1)} + \frac{X}{(\theta-1)^2}\right]^2 =$$

$$= E\left\{\frac{1}{(\theta-1)^2}[X-(\theta-1)]\right\}^2 = \frac{1}{(\theta-1)^4}E[(X-(\theta-1))]^2 =$$

$$= \frac{1}{(\theta-1)^4}\quad \text{Var}(X) = \frac{1}{(\theta-1)^4}\quad (\theta-1)^2 = \frac{1}{(\theta-1)^2}$$

$$Lc = \frac{(\theta-1)^2}{n}$$

The variance of the estimator coincides with the Cramer-Rao lower bound, thus implying that this is an efficient estimator.

Exercise C It is a **goodness of fit test to a partially specified distribution.** The information we have to perform this test is given below:

Brand	A	B	C	D
Probabilities	θ^2	$(1-\theta)^2$	$\theta(1-\theta)$	$\theta(1-\theta)$

i) In order to be able to estimate these probabilities, a random sample of 80 individuals has been taken and has provided the following information: 12 individuals bought brand A mobile telephones; 38 individuals bought brand B mobile telephones; 16 individuals bought brand C mobile telephones and 14 individuals bought brand D mobile telephones. With this information and to find the maximum likelihood estimator of θ, we have to write the likelihood function:

$$L(\theta) = \left[\theta^2\right]^{12}\left[(1-\theta)^2\right]^{38}[\theta(1-\theta)]^{16}[\theta(1-\theta)]^{14} = \theta^{54}(1-\theta)^{106}$$

We now compute its natural logarithm to obtain:

$$\ln L(\theta) = 54\ln(\theta) + 106\ln(1-\theta)$$

If we take derivatives with respect to θ and equate this to zero, we will have that:

$$\frac{\partial \ln L(\theta)}{\partial\theta} = \frac{54}{\theta} - \frac{106}{(1-\theta)} = 0 \Longrightarrow 54(1-\theta) = 106\theta \Longrightarrow 54 = 160\theta \Longrightarrow \hat{\theta}_{ML} = \frac{54}{160} = 0.3375$$

b) Goodness of fit test to a partially specified distribution.

First of all, we have that the estimated probabilities, p_i, for each one of the mobile telephone brands will be:

$$P(A) = (0.3375)^2 = 0.1139; \qquad P(B) = (1-0.3375)^2 = 0.4389$$

$$P(C) = P(D) = (0.3375)(1-0.3375) = 0.2236$$

Given that we have estimated the parameter θ, we have that $h = 1$. Moreover, as we have four brands of mobile telephones, $k = 4$. With this information and in order to perform the test, we build the following table:

	n_i	$\hat{p}_i$	$n\hat{p}_i$	$\frac{(n_i - n\hat{p}_i)^2}{n\hat{p}_i}$
Brand A	12	0.1139	9.112	0.9153
Brand B	38	0.4389	35.112	0.2375
Brand C	16	0.2236	17.888	0.1993
Brand D	14	0.2236	17.888	0.8451
	$n = 80$	1	$n = 80$	$z = 2.1972$

Under the null hypothesis that the probability distribution the firm has is the correct one, the test statistic $\sum_i \frac{(n_i - n\hat{p}_i)^2}{n\hat{p}_i} \sim \chi^2_{k-h-1}$, where k is the number of brands of mobile telephones ($k = 4$) and h is the number of estimated parameters ($h = 1$).

The decision rule, at a 5% approximate significance level, will be to reject the null hypothesis if:

$$z > \chi^2_{k-h-1\,,\,0.05} = \chi^2_{2,0.05}$$

In this case:

$$z = 2.1972 < 5.99 = \chi^2_{2,0.05}$$

so that, at a 5% approximate significance level, we do not reject the null hypothesis that the probability distribution the firm has for the different brands of mobile telephones is the correct one.