Ibn Al-Haitham Journal for Pure and Applied Science

مجلة إبن الهيثم للعلوم الصرفة و التطبيقية

No.	1	Vol.	25	Year	2012	2012	السنة	25	المجلد	1	العدد

Statistical Analysis of Duration of Hospitalization

N.K. Dakhil
Department of Mathematics, College of Mathematics and Computer Sciences, University of Kufa.
Received in:7 June 2011, Accepted in: 7 December 2011

Abstract

This research is mostly concerned of exploration analysis of a random sample of data from Al-sadder hospital. We examine duration of hospital stay (DHS) and investigate any significant difference in duration between sex, age groups, occupation, patients' condition at admission, and patients' condition at discharge.

Keyword: Analysis , Duration , Hospitalization , Iraq

Introduction

Previous research examined the variations in hospital medical practice, indicated by the duration of hospital stay, and suggested that the variation in length of hospital stay, within hospitals is much smaller than the length-of-stay variation between different hospitals.[9]. Other studies concluded that shorter neonatal hospital stay was associated with increased readmission rates for conditions that may not give rise to symptoms or signs on days 1 to 3 of life [6].

Further studies evaluated the factors influencing the duration of hospitalization in patients with ruptured cerebral aneurysm managed surgically during the acute stage [3]. The study concluded, the duration of hospital stay can be reduced to differences in practice. More studies linked the duration of hospital stay with mortality [1,2,4].

Hence the objective of this study was to study the duration of hospitalization in Al-Sadder hospital. A random sample of data from Statistics department of Al-sadder hospital was selected. Al-Sadder hospital was built in 1984. It is the biggest and most importance teaching hospital in Al-Najaf city . It consisted of six floors with many specialized departments. It is a major referral hospital in Al-Najaf city and provided high standard health services to Al-Najaf province and its environs.

The data was summarized using tables and graphs. A number and percentages was used to summaries categorical data and descriptive statistics for continuous data. A t-test was used to examine whether there was any difference in duration time between female and male and ANOVA was used to see if there is any difference in duration between categories of occupation, age and patients' condition at admission and discharge hospital. SPSS statistical packages and Excel were used to analyze the data. Two tails hypothesis was used with 5% level of significant.

Statistical Methods

1. Inference About Two Means: Independent Samples

Test a claim about two independent population means or construct a confidence interval estimate of the difference between two independent population means.

For the population one we let, μ_1 be the *population* mean, $\bar{x}_1$ be the sample mean, σ_1 be the population standard deviation, S_1 be the samples standard deviation, and n_1 the size of the first sample . The corresponding notation $\mu_1, \bar{x}_2, S_2$ and n_2 are applied for population two.

The requirements for this test are σ_1 and σ_2 are unknown and it is not assumed that σ_1 and σ_2 are equal, the two samples are independent, both samples are simple random samples, and either or both of these conditions is satisfied: the two sample sizes are both large (with $n_1 > 30$ and $n_2 > 30$) or both come from populations having normal distributions (these methods are robust against departures from normality, so for small samples, the normality requirement is loose in the sense that the procedures perform well as long as there are no outliers and departures from normality are not too extreme.

The test statistics is a value in making a decision about the null hypothesis and denoted by Z, where

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \qquad (1)$$

Where $\mu_1 - \mu_2$ is often assumed to be(0)

When finding critical values or p-value, we use simple and conservative estimate df = smaller of $n_1 - 1$ and $n_2 - 1$. Refer to the t-distribution to find P-values and critical values, where p-value is the probability of getting a value of the test statistics that is at least as extreme as the one representing the sample data, assuming the null hypothesis is true. P-value can be founds after finding the area beyond the test statistics. While the critical region is the set of all values of the test statistics that cause us to reject the null hypothesis and the critical value is any value separates the critical region(where we reject the null hypothesis) from the values of the test statistics that do not lead to rejection of the null hypothesis.

The confidence interval estimate of the difference $\mu_1 - \mu_2$ is

$$(\bar{x}_1 - \bar{x}_2) - E < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + E \qquad (2)$$

$$E = t a/2 \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \qquad (3)$$

And the number of degree of freedom df is the smaller of $n_1 - 1$ and $n_2 - 1$.[7,8]

2. One-Way Analysis Of Variance

When testing for equality of three or more population means, use the method of one-way analysis of variance.

One-way analysis of variance (ANOVA) is a method of testing the equality of three or more population means by analyzing sample variances. One-way analysis of variance is used with data categorized with one treatment (or factor), which is a characteristic that allows us to distinguish the different populations from one another.

The term treatment is used because early applications of analysis of variance involved agricultural experiments in which different plots of farmland were treated with different fertilizers, seed types, insecticides, and so on. The test requirement are:[7]

1. The populations have distributions that are approximately normal. This is a loose requirement, because the method works well unless a population has a distribution that is very far from normal. If a population does have a distribution that is far from normal, use the Kruskal-Wallis test.

2. The populations have the same variance σ^2 (or standard deviation σ). This is a loose requirement, because the method works well unless the population variances differ by large amounts. Statistician George E.P. Box showed that as long as the sample size are equal(or nearly equal),the variances can differ by amounts that make the largest up to nine times the smallest and the results of ANOVA will continue to be essentially reliable.)

3. The samples are simple random samples of quantitative data.

4. The samples are independent of each other.(The samples are not matched or paired in any way.)

5. The different samples are from populations that are categorized in only one way.

3. Estimations with Unequal Sample Sizes

While the calculations required for cases with equal sample sizes are reasonable, they become more complicated when the sample sizes are not all the same. The same basic reasoning applies because we calculate an F test statistic that is the ratio of two different estimates of the common population variance σ^2, but those estimates involve weighted measures that take the sample sizes into account, as shown below.

$$F = \frac{\text{variance between samples}}{\text{variance within sample}} = \frac{\left[\dfrac{\sum n_i(\bar{x}_i - \bar{\bar{x}})}{k-1}\right]}{\left[\dfrac{\sum (n_1-1)^2_{s1}}{\sum (n_1-1)}\right]} \quad \ldots\ldots\ldots\ldots\ldots\ldots(4)$$

Where

$\bar{\bar{X}}$ = mean of all sample values combined

K =number of all population means being compared

n_i = number of values in the i th sample

$\bar{X}_i$ = mean of value in the i th sample

S_i^2 =variance of values in the i th sample

The factor of n_j is included so that larger samples carry more weight. The denominator of the test statistic is simply the mean of the sample variances, but it is a weighted mean based on the sample sizes.

Because calculating this test statistic can lead to large rounding errors, the various software packages typically use a different (but equivalent) expression that involves SS (for sum of squares) and MS (for mean square) notation. Although the following notation and components are complicated and involved, the basic idea is the same: The test statistic F is a ratio with a numerator reflecting variation *between* the means of the samples and a denominator reflecting variation *within* the samples. If the populations have equal means, the F ratio tends to be small, but if the population means are not equal, the F ratio tends to be significantly large. Key components in our ANOVA method are described as follows.

SS(total), or total sum of squares, is a measure of the total variation (around $\bar{\bar{x}}$) of the sample data combined [8].

$$SS(total) = \sum (X - \bar{\bar{x}})^2 \ldots\ldots\ldots\ldots\ldots\ldots\ldots(5)$$

SS(total) can be broken down into the components of SS(treatment) and SS(error), described as follows.

SS(treatment), also referred to as SS(factor), SS(between groups), or SS(between samples), is a measure of the variation *between* the sample means.

$$SS(treatment) = n_1(x-\bar{\bar{x}})^2 + n_2(x_2-\bar{\bar{x}})^2 + \ldots + n_k(\bar{x}_k-\bar{\bar{x}})^2$$

$$= \sum n_i(\bar{x}i - \bar{\bar{x}})^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(6)$$

. If the population means $(\mu_1, \mu_2, \ldots, \mu_k)$ are equal, then the sample means $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k$ with all tend to be close together and also close to $\bar{\bar{x}}$ The result will be a relatively small value of SS(treatment). If the population means are not all equal, however, then at least one of $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k$ will tend to be far apart from the others and also far apart from $\bar{\bar{x}}$. The result will be a relatively large value of SS(treatment).

SS(error), also referred to as SS(within groups) or SS(within samples), is a sum of squares representing the variation that is assumed to be common to all the populations being considered.

$$SS(error) = (n_1-1)s_1^2 + (n_2-1)s_2^2 + \ldots\ldots + (n_k-1)s_k^2$$

$$= \sum (ni-1)S_i^2 \quad \ldots\ldots\ldots(7)$$

Given the preceding expressions for SS(total), SS(treatment), and SS(error), the following relationship with always hold.

SS(total)=SS(treatment) + SS(error) ……..(8)

SS(treatment) and SS(error) are both sums of squares, and if we divide each by its corresponding number of degrees of freedom, we get mean squares. Some of the following expressions for mean squares include the notation N:

N = total number of values in all samples combined.

MS(treatment) is a mean square for treatment, obtained as follows:

$$Ms(treatment) = \frac{ss(treatment)}{k-1} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(9)$$

MS(error) is a mean square for error, obtained as follows:

$$Ms(error) = \frac{SS(error)}{N-K} \ldots\ldots\ldots\ldots\ldots\ldots\ldots(10)$$

MS(total) is a mean square for the total variation, obtained as follows:

$$Ms(total) = \frac{SS(total)}{N-1} \ldots\ldots\ldots\ldots\ldots\ldots\ldots(11)$$

In testing the null hypothesis $H_0 : \mu_1 - \mu_2 = \ldots \mu_k$ against the alternative hypothesis H_1: the mean are not all equal ,the last statistic

$$F = \frac{MS(treatment)}{MS(error)} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(12)$$

Has an F distribution (when the null hypothesis H_0 is true) with degrees of freedom given by

Numerator degrees of freedom $= \mathrm{k} - 1$

Denominator degrees of freedom $= \mathrm{N} - \mathrm{k}$

This test statistic is essentially the same as the one given earlier, and its interpretation is also the same as described earlier. The denominator

depends only on the sample variances that measure variation within the treatments and is not affected by the differences among the sample means. In contrast, the numerator is affected by differences among the sample means. If the differences among the sample means are excessively large, they will cause the numerator to be excessively large, so F will also be excessively large. Consequently, very large values of F suggest unequal means, and the ANOVA test is therefore right-tailed.

4. Identifying Which Means Are Different

After conducting an analysis of variance test, we might conclude that there is sufficient evidence to reject a claim of equal population means, but we cannot conclude from ANOVA that any *particular* means are different from the others. There are several formal and informal procedures that can be used to identify the specific means that are different. Here are two *informal* methods for comparing means:

1. Construct box-plots of the data sets to see if one or more of the data sets is very different from the others.
2. Construct confidence interval estimates of the means from the data sets, then compare those confidence intervals to see if one or more of them does not overlap with the others.

There are several formal procedures for identifying which means are different. Some of the tests, called range tests, allow us to identify subsets of means that are not significantly different from each other. Other tests, called multiple comparison tests, use pairs of means, but they make adjustments to overcome the problem of having a significance level that increases as the number of individual tests increases. There is no consensus on which test is best, but some of the more common tests are the Duncan test, Student-Newman-Keuls test (or SNK test), Tukey test (or Tukey honestly significant difference test), Scheffé test, Dunnett test, least significant difference test, and the Bonferroni test. In this paper, we will discuss the Bonferroni test in details.

5. Bonferroni Multiple Comparison Test

We have used the Bonferroni test to identify which means are different. Here are some steps to carry out the Bonferroni test:

Step1**.** Do a separate t test for each pair of samples, but make the adjustments described in the following steps.

Step2**.** For an estimate of the variance σ^2 that is common to all of the involved populations, use the value of MS(error), which uses all of the available sample data. The value of MS(error) is typically obtained when conducting the analysis of variance test. Using the value of MS(error), calculate the value of the test statistic t, as shown below. The particular Sample 2; change the subscripts and use another pair of samples until all of the different possible pairs of samples have been tested.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{MS(error).(\frac{1}{n_1} + \frac{1}{n_2})}}$$

Step3**.** After calculating the value of the test statistic t for a particular pair of samples, find either the critical t value or the P-value, but make the following adjustment so that the overall significance level does not increase.

Use the test statistic t with $df = N - k$, where N is the total number of sample values and k is the number of samples, and find the P-value the usual way, but adjust the P-value by multiplying it by the number of different possible pairings of two samples. (For example, with three samples, there are three different possible pairings, so adjust the P-value by multiplying it by 3.) When finding the critical value, adjust the significance level α by dividing it by the number of different possible pairings of two samples. (For example, with three samples, there are three different possible pairings, so adjust the significance level by dividing it by 3.)

Note that in Step3 of the preceding Bonferroni procedure, either an individual test is conducted with a much lower significance level, or the P-value is greatly increased. Rejection of equality of means therefore requires differences that are much farther apart. This adjustment in Step3 compensates for the fact that we are doing several tests instead of only one test [8].

Statistical Analysis

1. Data

The Directorate of Health in Al-Najaf serves as the ethics committee for Al-Sadder Hospital. The Directorate of Health gave permission for the research. The main source of the collected data was the statistics department in the same hospital.

A random sample of 2008 admission into Al-sadder hospital was reviewed. The data consisted of 99 patient records, that contained sex, age, address, admission date, discharge date, diagnosis, operation, condition at admission, and condition at leaving hospital. The duration of hospital stay (duration) was calculated in days from subtracting discharged date from admission. Condition at admission was indicated by kind of operation.

Data were entered and coded. It was checked thoroughly then categorized according to categories that were suggested by statistics department in Al-sadder hospital. Sex was categorized to female and male, age in years was categorized to less than 20, 20 to 39, 40 to 49, and 60 and over. Occupation was categorized to child, employed, handicap, housewife, and student. Patients condition at leaving hospital was categorized to better, death, not known, and condition at admission was categorized to major, minor and chronic.

2. Results and Conclusion

The random sample of patients were analyzed using both descriptive and inferential statistics.

The percentage of male are higher than female indicated that the majority of patients who admitted to Al-sadder hospital in 2008 were male.

A total of 99 patients with (42%) female and (57%) male, and age range of 2-85 years, mean age of 39.36

Our sample showed that only 13% of patients admitted to the hospital were children. While 87% were adults. It is interesting to note that number of housewife patients (35.4%) were similar to employed ones (29.3%). Nevertheless only 18% were handicops and 4% were students.

Nearly (90%) of patients when leaving hospital were feeling better, and only 5% of them died at hospital, and 6% unknown.

We examine the requirement to carry out t-test and ANOVA. Firstly the samples were independent of each other as the samples are not matched or paired in any way. Secondly distributions were either approximately normal or was not very far-off from the normal distribution. Thirdly since the samples sizes were nearly equal, the results of ANOVA will continue to be essentially reliable

An z-test was used to test the following hypothesis:

H_0 = mean duration of male = mean duration of female

H_1 = mean duration of male ≠ mean duration of female

Critical regions are $Z_{0.025} < -1.96$ and $Z_{0.025} > 1.96$ at 5% level of significance

The test illustrated that there was no significant difference in duration $(z = 38.32 > 1.96)$ at 5% level of significance between the mean duration for female (3.58) and the mean duration of male (3.88). This could state that mean

duration of hospital stay were similar for male and female for 2008 year (see figure 1).

An ANOVA was used to test the following hypothesis:

H_0 = mean duration of all age groups are equal

H_1 = mean duration of all age groups are not equal

Critical regions are at 5% level of significance $F_{(3,93)} > 2.68$

Analysis of variance indicated that there was no significant difference in duration $(F_{(3,93)} = 0.311 < 2.68)$ at 5% level of significance between the mean of four age categories. This may specify that duration of hospital stay for age group were alike.

Out of 99 patients, 46 (46.5%) had chronic diseases , 27(27.3%) had acute (major) diseases, and 26(26.3%) has uncomplicated (mi ANOVA was used to test the following hypothesis:

An ANOVA was used to test the following hypothesis:

H_0 = mean duration of condition at admission are equal

H_1 = mean duration of condition at admission are not equal

Critical regions are at 5% level of significance $F_{(2,96)} > 3.07$

There differences in duration between condition at admission group was highly significant $(F_{(2,96)} = 5.05 > 3.07)$ at 5% level of significance.

Bonferroni multiple comparisons test was used to identify which mean are different. Table 4 illustrated that there was a significant difference in duration between major and minor conditions, and minor and chronic conditions(see also figure4).

In conclusion, the results suggest that duration of hospital stay is strongly linked with the condition at admission and not affected by sex, age. While previous study demonstrated that the duration of hospital stay was significantly prolonged in aged patients (more than 70 years old, and the duration of hospitalization was shortened in fatal cases mainly because death occurred in a relatively early stage after onset of subarachnoid hemorrhage [3].

The duration of hospital stay not only depends on clinical factors but also medical ,social and economic factors. The evaluation of various factors is necessary to shorten the duration of hospitalization[3]. We may need to extend this study to include social and economic factors which were unavailable from Al-Sadder hospital.

References

1. Caballero-Granado F.J., B. Becerril, L. Cuberos2, M. Bernabeu1, J. Cisneros1 and J. Pachón1(2001). Attributable Mortality Rate and Duration of Hospital Stay Associated with Enterococcal Bacteremia. the Infectious Diseases Society of America.
2. Garko S.B., and Ekweani C.N., and Anyiam C.A. Duration of Hospital Stay and Mortality the Medical Wards of Ahmadu Bello University Teaching Hospital, Kaduna. Annals of African Medicine. 2, No. 2; 2003: 68 – 71
3. Hiroshi Wanifuchi, Takashi Shimizu, Kosaku Amano, Shiho Harashima and Sakiko Suzuki. (2005) "Clinical Evaluation of Factors Influencing the Duration of Hospital Stay in Patients with Ruptured Cerebral Aneurysm Managed Surgically During the Acute Stage". Surgery for cerebral Stroke..33, No. 1, 26-29.
4. Landry SL, Kaiser DL, and Wenzel RP. Hospital stay and mortality attributed to nosocomial enterococcal bacteremia: a controlled study. Am J Infect Control 1989;17:323-9.
5. Larsen R.J.. and Max M.L. (2006) Mathematical Statistics and its applications, fiftrh edition Prentic Hall.
6. Lee KS, Perlman M, Ballantyne M, Elliott I, To T.Association between duration of neonatal hospital stay and readmission rate. J Pediatr. 1995 Nov;127(5):758-66.
7. Triola M.F. (2010) Elementary Statistics, Eleventh edition. Addison –Wesley .

8. Wayne W. D. (2010) Biostatistics. John Wiley & Sons, INC.
9. Westert, Gert P. and Nieboer, Anna P. and Groenewegen, Peter P. (1993) Variation in duration of hospital stay between hospitals and between doctors within hospitals. Social Science and Medicine, 37 (6).. 833-839.

Ppendix

Table (1): Duration by sex

Statistics				
Std. Deviation	Mean	N	Sex	
3.5769	3.286	42	F	Duration(days)
3.8788	3.939	57	M	

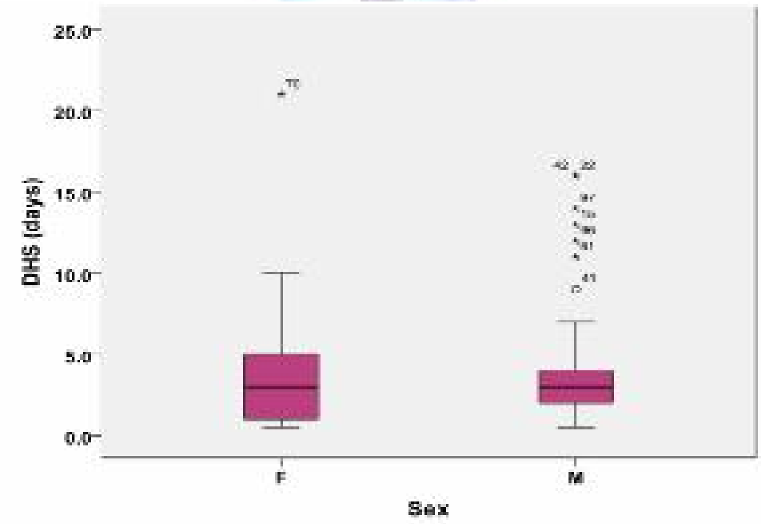

Fig.(1): Box plot DHS by sex

Table (2): Descriptive statistics of duration by age group

Descriptive Statistics					
Duration(days)					
Maximum	Minimum	Std. Deviation	Mean	N	Age groups
16.0	.5	4.5136	3.619	21	< 20
16.0	1.0	2.8632	3.258	31	20 – 39
21.0	.5	4.4944	4.159	22	40 – 59
14.0	.5	3.4769	4.043	23	60 and over
21.0	.5	3.7597	3.727	97	Sub-total
				1	Missing
				99	Total

Table (3): Analysis of Variance for the data in Table 2.

Sig.	F	Mean Square	Df	Sum of Squares	Age group
.817	.311	4.491	3	13.473	Between Groups
		14.447	93	1343.538	Within Groups
			96	1357.010	Total

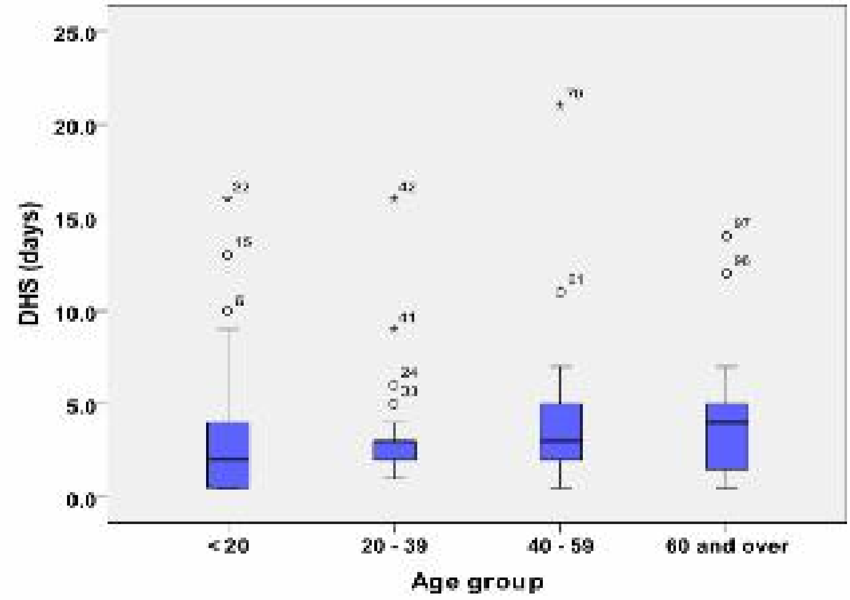

Fig.(2): Box plot DHS by Age Group

Table (4): Descriptive statistics of duration by condition at admission

Period(days)					
Maximum	Minimum	Std. Deviation	Mean	N	Category
21.0	.5	5.5492	4.704	27	Major
5.0	.5	1.2576	1.692	26	Minor
14.0	.5	2.9647	4.163	46	Chronic
21.0	.5	3.7490	3.662	99	Total

Table (5): Analysis of Variance for the data in Table 4.

Sig.	F	Mean Square	df	Sum of Squares	Condition at admission
.005	5.505	70.859	2	141.719	Between Groups
		12.872	96	1235.695	Within Groups
			98	1377.414	Total

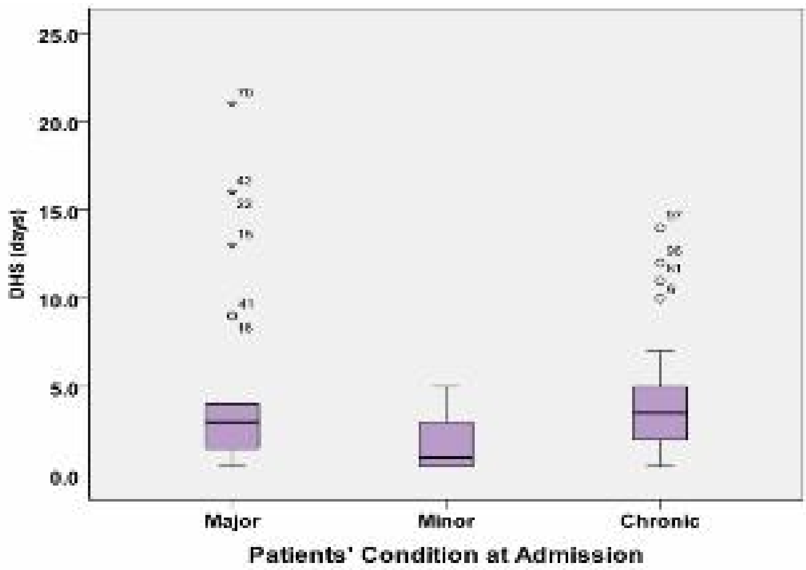

Fig.(3): Box plot DHS by condition at admission

Table (6): Bonferroni multiple comparison test

Multiple Comparisons: Duration (days)						
95% Confidence Interval		Sig.	Std. Error	Mean Difference (I-J)	(J) Condition at admission	(I) Condition at admission
Upper Bound	Lower Bound					
5.413	.609	.009	.9858	3.0114*	Minor	Major
2.660	-1.579	1.000	.8698	.5407	Chronic	
-.609	-5.413	.009	.9858	-3.0114*	Major	Minor
-.326	-4.616	.018	.8803	-2.4707*	Chronic	
1.579	-2.660	1.000	.8698	-.5407	Major	Chronic
4.616	.326	.018	.8803	2.4707*	Minor	
*. The mean difference is significant at the 0.05 level.						

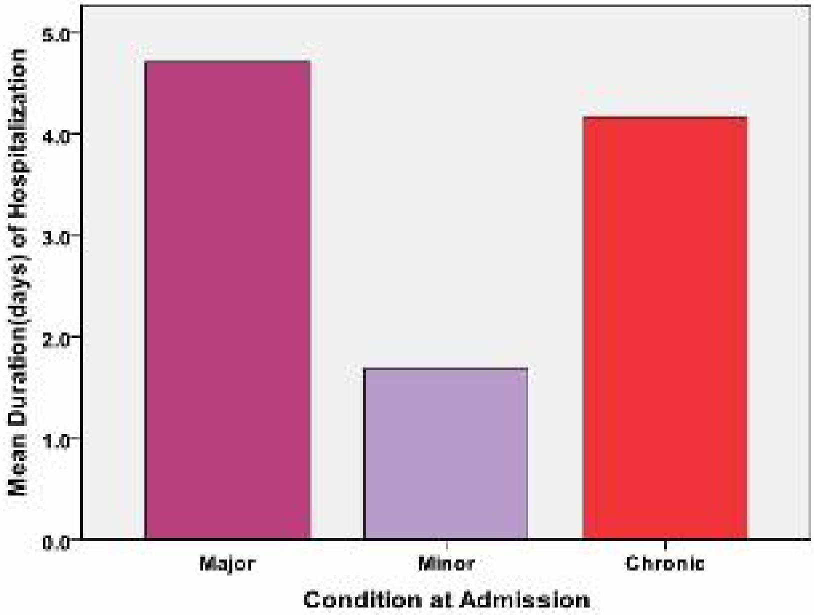

Fig.(4):Mean duration of condition at admission

مجلة إبن الهيثم للعلوم الصرفة و التطبيقية | Ibn Al-Haitham Journal for Pure and Applied Science
العدد 1 المجلد 25 السنة 2012 | No. 1 Vol. 25 Year 2012

التحليل الإحصائي لمدة المكوث في المستشفى

نزيرة خليل داخل
قسم الرياضيات، كلية الرياضيات وعلوم الحاسوب – جامعة الكوفة
استلم البحث في:7 حزيران 2011 ، قبل البحث في:7 كانون الاول 2011

الخلاصة

يتناول هذا البحث تحليلا استكشافيا إلى عينة عشوائية من مستشفى الصدر في النجف. قمنا بفحص مدة البقاء في المستشفى والتحقق من أهمية ودلالة الفرق في مدة البقاء بين النساء والرجال، فئات العمر ، فئات حالات المرضى عند دخول المستشفى.

الكلمات المفتاحية: تحليل ، فترة البقاء ، ادخال للمستشفى ، العراق