TEST OF MOBILE PHONES SCREEN DATA BASED ON THE MAHALANOBIS DISTANCE

F. Z. Okwonu
Department of Mathematics, Delta State University Abraka, Nigeria.
E-mail: fzokwonu_delsu@yahoo.com.

Accepted 21st May, 2016

A survey was conducted to obtain data set on scratched mobile smart and ordinary (conventional) phones screen, and to further investigate if the data set obtained are normally distributed or not. The Mahalanobis distance, the correlation coefficient, Chi squared and the quantile quantile plots are applied to determine if the data set obtained based on the categorization is normally distributed or otherwise. The techniques revealed that the data set are not normally distributed. The mean and the standard deviation approach based on the concept of data contamination validated the previous conclusions. At 10% level of significance, the hypothesis was rejected, implying non-normality of the data set. In general, the conclusions based on all techniques indicated that the data set are not normally distributed. This implies that though users of mobile phones are both diligent and non-diligent alike.

Key words: Mahalanobis distance, Chi squared, normal, bivariate, screen, phone.

INTRODUCTION

The data set $X = \{x_i\}, i = 1,...k,$ is often assumed to come from multivariate normal distribution. In some cases, the sample mean computed from the data set and the distance from the sample mean to the population mean may have been drawn from a normal distribution (Johnson and Wichern, 2008), that is

$$n\left(\overline{x}_i - \mu\right)` S_p^{-1} \left(\overline{x}_i - \mu\right), \quad (1)$$

In this case, the sample mean is distributed as $N_k\left(\mu, \delta\right)$ which by implication chi is squared distributed. It may be easy to conclude that the chi squared distribution is approximately the sampling distribution of Equation(1) which provided the sample mean that is normally distributed (Johnson and Wichern, 2008; Ramzan et al., 2013). For instance, suppose the sample mean is normally distributed, it implies that the population from which the mean is obtained is normal. From this analysis, if the population is not normally distributed, this implies that the computed mean is otherwise not from a normal distribution.

The concept of distance due to Mahalanobis has been extensively applied to detect influential observations by comparing it with well-defined cutoff point. It is often applied to detect and delete influential observations when the parent data is contaminated. Although, some robust techniques are proposed using the Mahalanobis distance measure (Rousseeuw and Leory, 1987; Alqallaf et al., 2002; Maronna and Yohai, 1998; Rousseeuw and Van Zomeren, 1991; Maronna and Zamar, 2002). In this consideration, we are not applying it to robustify the data set but as procedure to determine normality or non-normality of the mobile phone data set. Though the conventional influential observation detection approach can be invoked by comparing the distance value with the table chi squared value. For the purpose of this study, the distance value was plotted to investigate the points that fall within the 50% contour to determine normality or otherwise. The concept of correlation coefficient was coined by Fisher (1915) and Zimmerman et al. (2003). The test of normality has been discussed by Fisher, Pearson, Williams and Geary based on the third and fourth moment (Filliben, 1975). The Shapiro and Wilk test statistics and the correlation coefficient have been reported (Filliben, 1975; Shapiro and Wilk, 1968).

Basically, the correlation coefficient allows to interpret the near linearity of a plot due to test performed (Ryan and Joiner, 1976), it can be stated that the sample correlation coefficient is an unbiased estimator of the population correlation (Zimmerman et al., 2003). The bias of the sample correlation coefficient relies on the sample size and it is a feature of the location and the corresponding variation due to scale. Though, the correlation coefficient has been transformed to different fashions and names. In this discussion, the focus is to apply it to investigate the normality or non-normality of the mobile phone screen data set. Detailed discussion on different normality test has been discussed for review purposes (Farrell and Rogers-Stewart, 2006; Romao et al., 2013; Kim and Bickel, 2003; D' Agostino and Stephens, 1986; Cox and Small, 1978; Epps and Pulley, 1983; Leslie et al., 1986; Machado, 1983).

The concept of housing demand has been investigated based on whether housing distribution follows normal distribution (Min, 2007; Romao et al., 2013). The stock market data was studied to observed if the data set are normally distributed (Liesenfeld and Jung, 2000; Romao et al., 2013). Normality of data set of different sample sizes, sources and measurements have received detailed attention in recent times in the field of medicine (Schoder and Wilhelm, 2006; Surucu and Koc, 2007). The normality of quality control and gene expression data set has been researched extensively (Mathur and Dolo, 2008; Vamman and Albing, 2007; Muttlak and Al-Sabah, 2003; Oakland, 2008; Romao et al., 2013). In hydrological research for instance, in order to test for the shape or extreme value distribution, correlation coefficient was applied and this prove to be useful (Kinnison, 1989; Vogel, 1986).

The focus of this discussion is to determine whether the screen condition of mobile phones of different categories are normally distributed or not using the Mahalanobis distance and the correlation coefficient. The normality or otherwise of the data set will aid us to determine the nature of usage, care and maintenance of the mobile phones. This will enable us determine diligent users and non-diligent users. In general, if the data set is normal, it implies the combination of both categories of end users is true. This presentation will provide answers to questions such as: are the data set obtained normally distributed? Do the Q-Q plot or the chi squared plot revealed the characteristics of observations that is normally distributed?

METHODOLOGY

The Mahalanobis distance

The Mahalanobis distance is denoted by M^2 and defined as

$$M^2 = \left(x_i - \overline{x}\right) S^{-1} \left(x_i - \overline{x}\right),$$

where $\overline{x} = \frac{\sum_{i=1}^{k} x_i}{n}$, and $S = \frac{\sum_{i=1}^{k}\left(x_i - \overline{x}\right)\left(x_i - \overline{x}\right)'}{n-1}$.

Analyzing the bivariate or multivariate data based on the Mahalanobis distance, a cutoff point is defined such that $M^2 > \chi_k^2(\alpha)$. Based on this definition, inliers and influential observations can be determined with 0.5 probability. This implies that 50% of the sample observations will be in the ellipse (Johnson and Wichern, 2008).

The correlation coefficient

The correlation coefficient in this respect can be applied to test the acceptance or rejection of the normality of the data set. This can be defined as

$$C_c = \frac{\sum_{i=1}^{k}\left(x_i - \overline{x}\right)\left(q_i - \overline{q}\right)}{\sqrt{\sum_{i=1}^{k}\left(x_i - \overline{x}\right)^2}\sqrt{\sum_{i=1}^{k}\left(q_i - \overline{q}\right)^2}}, i = 1, . \mathrm{k},$$

$$q_{\mathrm{i,k}}\left(\frac{m-\alpha}{k}\right) = \chi_k^2\left(\frac{k-m+\alpha}{k}\right),$$

where

$$\alpha = 0.5, \overline{q} = \frac{\sum_{i=1}^{k} q_i}{k}, m = 1, ..., k$$

In this case, since $\overline{q}$ is approximately zero, the aforementioned equation can be written as

$$C_c = \frac{\sum_{i=1}^{k}\left(x_i - \overline{x}\right)\left(q_i\right)}{\sqrt{\sum_{i=1}^{k}\left(x_i - \overline{x}\right)^2}\sqrt{\sum_{i=1}^{k}\left(q_i\right)^2}}.$$

In general, the correlation coefficient can be considered technically as invariant of the location and scale, and statistically independent of the sample mean and standard deviation. It can be posited that C_c rely only on the sample observation and the normal cumulative distribution function. This concept is applied to verify the straightness of the plot. The concept of correlation coefficient simply implies that normality tends to produce approximately linear normal probability plot which exhibit the characteristics of approximate unit values of the probability plot (Filliben, 1975).

The chi squared plot

The chi squared procedure simply involves the Mahalanobis distance and the chi squared percentiles. The plot depends strictly on the arrangement of the Mahalanobis distance, and the equivalent chi squared percentile values in ascending order of magnitude (Johnson and Wichern, 2008). The plot should have a linear relationship indicating normality otherwise the data set will not be normally distributed.

The quantile quantile (Q-Q) plot

Like the chi squared plot, the Q-Q plot depends on the ordered sample observations, the probabilities of the observations and standard normal quantiles. In general, the straightness of the plot depends on the ordered observations and the standard normal quantiles paired. Gnanadesikan gave details of the quartile quartile plots and other graphical techniques for analyzing the normality of data set (Gnanadesikan, 1977; Ramzan et al., 2013).

Data collection and analysis

This discussion is based on the conditions of the screen of mobile phones. A survey was conducted and questionnaire were designed and distributed in different places (in Delta State and Anambra State, Nigeria), and the targeted venues are supermarkets, fast food restaurants, bars, motor parks, tertiary institutions, hotels, and clubs e.t.c., the mobile phones were categorized as presented in Table 1. This exercise was carried out for the period of 30 days with defined age categories. Table 2 gives detail information of the exercise. Normal screen in this context implies the conventional mobile phone screen without dent on it while scratched screen means a mobile phone screen with crack, scratch or broken screen or surface (Table 3).

Table 1. Mobile phones categorization.

Ordinary (conventional) mobile phones	Smart mobile phones
Consist of phonebooks; Torch light, single or dual Sim port, keypad, FM radio	Consist of phonebooks, gallery, torch screen, camera, video recorder, 3G/4G, facebook, twitter, whatsapp, play store, carl care, facetime, google+, browser, google settings, instagram, operamini, voice search, youtube, palmchat, palmplay, maps, sound recoder, WPS office, flashshare, chrome, xclub, news and weather, apple store, stocks, yahoo, gmail, compass, itunes, safari, ipod, notes and others

Table 2. Data set for smart phones: scratched and normal screens.

15-39 years		
Targeted venue	**Scratched screen**	**Normal screen**
Supermarket	16	51
Fast food restaurant	9	24
Bar	10	23
Motor park	23	49
Tertiary institutions	37	126
Hotels	11	53
Clubs	15	54
Church premises	31	107
Market	29	98
Hospital	6	29
Auto workshop	15	47
Car wash	17	43
Ministry building (public building)	23	79
Tricycle operators	31	109

Block industry	13	57
Water vendor	17	68
Newspaper vendor	10	51
Construction workers	46	153
Football field	8	34
Wood worker (carpenter/furniture makers)	19	58
Auto dealers	4	29
Wood dealers	11	61
Butchers	19	68
Scrap metal dealers/scavengers	33	119
Electronic dealers shops	5	37
Welder workshops	20	76
Fuel station attendants	18	59
Auto electricians workshops	15	48
Fruits dealers	32	114
Food vendors	43	147
Drycleaner shop	13	78
40-65 years		
Targeted venue	**Scratched screen**	**Normal screen**
Supermarket	5	34
Fast food restaurant	6	45
Bar	3	32
Motor park	10	56
Tertiary institutions	18	70
Hotels	7	41
Clubs	6	36
Church premises	24	174
Market	35	147
Hospital	5	48
Auto workshop	8	36
Car wash	9	36
Ministry building (public building)	14	65
Tricycle operators	19	81
Block industry	9	39
Water vendor	11	47
Newspaper vendor	8	69
Construction workers	38	121
Football field	8	38
Wood worker (carpenter/furniture makers)	12	61
Auto dealers	3	37
Wood dealers	16	73
Butchers	23	55
Scrap metal dealers/scavengers	28	74
Electronic dealers shops	11	56
Welder workshops	18	57
Fuel station attendants	4	21
Auto electricians workshops	9	30
Fruits dealers	7	87
Food vendors	21	95
Drycleaner shop	19	83

Table 3. Data set for ordinary (conventional) phones: Scratched and normal screens.

15-39 years		
Target venue	**Scratched screen**	**Normal screen**
Supermarket	5	33
Fast food restaurant	6	41
Bar	4	49
Motor park	10	67
Tertiary institutions	5	81
Hotels	5	31
Clubs	1	42
Church premises	16	141
Market	13	158
Hospital	4	41
Auto workshop	19	62
Car wash	11	66
Ministry building (public building)	20	179
Tricycle operators	24	113
Block industry	21	97
Water vendor	9	53
Newspaper vendor	21	134
Construction workers	53	268
Football field	3	25
Wood worker (carpenter/furniture makers)	12	89
Auto dealers	7	62
Wood dealers	19	51
Butchers	32	134
Scrap metal dealers/scavengers	41	161
Electronic dealers shops	3	70
Welder workshops	9	106
Fuel station attendants	6	44
Auto electricians workshops	17	48
Fruits dealers	22	98
Food vendors	48	278
Drycleaner shop	8	46
40-65 years		
Venue	**Scratched screen**	**Normal screen**
Supermarket	8	54
Fast food restaurant	2	15
Bar	6	67
Motor park	9	74
Tertiary institutions	13	88
Hotels	3	69
Clubs	3	19
Church premises	10	113
Market	21	102
Hospital	17	98
Auto workshop	12	69
Car wash	16	58
Ministry building (public building)	34	204
Tricycle operators	31	145
Block industry	18	76
Water vendor	26	82
Newspaper vendor	34	213

Construction workers	79	297
Football field	12	76
Wood worker (carpenter/furniture makers)	19	106
Auto Dealers	2	50
Wood Dealers	32	136
Butchers	45	187
Scrap Metal Dealers/Scavengers	56	208
Electronic Dealers Shops	12	77
Welder Workshops	10	134
Fuel Station Attendants	16	63
Auto Electricians Workshops	34	193
Fruits Dealers	30	211
Food Vendors	73	309
Drycleaner Shop	48	191

DISCUSSION

The investigators adduced reasons while smart phones screens often scratch or crack, most respondent revealed that smart phone screen is very fragile due to thin layer body cover. We also infer that screen protectors and phone jackets helps protect the screen from scratches due to free fail or involuntary actions.

Further investigations revealed that most of the smart phones with screen scratches or cracks do not have screen protectors or phone jackets or cover. Due to the sensitivity of the screen to other functions interacting with the screen, once the screen is scratched most functions often delay the sensitivity and connectivity of the phone. In this case, the screen touch often delays when dialing or composing short massages. On the other hand, the ordinary phones do not underperform when the screen scratch or crack; this is due to the fact that most of the functions depend strictly on the keypad. The ordinary or conventional mobile phones generally require the screen protectors as well as the phone jackets.

In Table 2, we observed that the age bracket between 15 to 39 years often use smart phones of different categorization than the age bracket 40 to 65 years. We also observed that the proportion of screen scratch or crack is higher for less than 39 years' category and less for 40 to 65 years' category. The study revealed that the proportion of normal phone user is higher for the age bracket between 40 to 65 years than 39 years' category. In general, this study revealed that the number of screen scratch depends on the actions (transactions, occupations) taking place in the targeted venue. Relying on this discussion, we note that the age bracket between 15 to 39 years use smart phones more than those of 40 to 65 years respectively. Based on the study, people between the age brackets 40 to 65 years use ordinary phones more than the younger group. Although in most cases, some people often use the ordinary or conventional mobile and the smart phones together due to their respective functions, purpose, durability and network service provider availability or service swap.

In group one (smart mobile phones), 3,500 questionnaires were administered and 5,500 for group two (ordinary mobile phones). For group one; 78.51% responded for scratched or cracked screen, and 67.37% for normal screen for age bracket between 15 to 39 years. In group two (ordinary mobile phones), 60.76% responded for scratched or cracked screen, and 82.1% for normal screen for age bracket between 40 to 65 years respectively.

The mean and standard deviation approach

Based on the computed means and standard deviations and invoking the concept of symmetric and asymmetric contamination model; that is $N_k(0,1)+N_k(19.32,69.32)$ we conclude that the data sets are not normally distributed in both categories (Tables 4, 5 and 6).

The Mahalanobis distance approach

Comparing the distance values with the cutoff points, that is $M^2 > \chi_2^2(0.5)=1.386$ the points within and outside the 50% contour is

Table 4. Mean and standard deviation for smart phones.

Category	Mean		Standard deviation	
	Scratched screen	Normal screen	Scratched screen	Normal screen
15-39 years	19.323	69.323	11.031	35.673
40-65 years	13.355	62.709	9.068	34.439

Table 5. Mean and standard deviation for ordinary mobile phones.

Category	Mean		Standard deviation	
	Scratched screen	Normal screen	Scratched screen	Normal screen
15-39 years	15.290	92.516	13.148	64.133
40-65 years	23.581	122.065	19.800	75.615

Table 6. Mahalonobis distance values and average.

Smart mobile phone screen		Ordinary mobile phone screen	
15-39 years		**15-39 years**	
M^2	M^2 / df	M^2	M^2 / df
0.5960843	0.2980421	0.862211	0.4311055
2.3068354	1.1534177	0.6461256	0.3230628
3.1190666	1.5595333	0.7515211	0.3757605
7.7048703	3.8524352	0.1721375	0.0860688
2.6171358	1.3085679	1.5548898	0.7774449
1.2084214	0.6042107	0.9271973	0.4635986
0.1866501	0.0933251	1.2656346	0.6328173
1.1485401	0.5742701	1.9461622	0.9730811
0.7765659	0.3882829	5.3233629	2.6616814
1.460241	0.7301205	0.7532184	0.3766092
0.7819223	0.3909611	2.0733263	1.0366632
2.8535359	1.4267679	0.1743244	0.0871622
0.1302266	0.0651133	4.3047466	2.1523733
1.2373811	0.6186906	0.6815805	0.3407903
0.7056767	0.3528384	0.5450106	0.2725053
0.2995156	0.1497578	0.3900729	0.1950364
1.5102177	0.7551088	0.4763584	0.2381792
5.87816	2.93908	8.5017642	4.2508821
1.0566925	0.5283462	1.1114305	0.5557153
0.8139143	0.4069572	0.1621569	0.0810784
2.260878	1.130439	0.414839	0.2074195
2.8117413	1.4058707	3.1385107	1.5692554
0.0017081	0.000854	2.3866202	1.1933101
1.9995625	0.9997813	5.2846771	2.6423385
2.6990123	1.3495061	1.6676641	0.8338321
0.165057	0.0825285	1.7225079	0.861254
0.3138734	0.1569367	0.5841288	0.2920644
0.6512089	0.3256044	2.5248068	1.2624034
1.5816722	0.7908361	0.7434509	0.3717255
4.810332	2.405166	8.3646561	4.182328
6.3133007	3.1566504	0.862211	0.4311055
40-65 years		**40-65 years**	
M^2	M^2 / df	M^2	M^2 / df

0.8800902	0.4400451	0.8219912	0.4109956
0.6976622	0.3488311	2.3072432	1.1536216
1.3039274	0.6519637	0.8412319	0.420616
0.1610065	0.0805033	0.5542164	0.2771082
0.3579438	0.1789719	0.2969942	0.1484971
0.507592	0.253796	1.5108723	0.7554361
0.7075887	0.3537943	2.1652912	1.0826456
15.333599	7.6667997	2.1973462	1.0986731
6.5507009	3.2753504	0.1561735	0.0780867
1.0801275	0.5400637	0.1114296	0.0557148
0.6024532	0.3012266	0.517405	0.2587025
0.6454104	0.3227052	1.7593006	0.8796503
0.005354	0.002677	2.6414525	1.3207263
0.3920615	0.1960307	0.1519237	0.0759618
0.4837101	0.2418551	0.8851362	0.4425681
0.2337153	0.1168576	2.7351204	1.3675602
1.4444643	0.7222321	3.674649	1.8373245
7.9022555	3.9511277	8.2946616	4.1473308
0.5166047	0.2583024	0.3748524	0.1874262
0.0343637	0.0171819	0.0535218	0.0267609
1.3637025	0.6818512	1.2049245	0.6024623
0.0977192	0.0488596	0.4641833	0.2320916
4.0737723	2.0368862	1.2949876	0.6474938
4.9309556	2.4654778	3.590688	1.795344
0.0676586	0.0338293	0.3637109	0.1818554
1.1086969	0.5543484	4.5798155	2.2899078
1.4832779	0.7416389	1.356927	0.6784635
1.0880286	0.5440143	1.6322938	0.8161469
4.6235312	2.3117656	5.1815524	2.5907762
0.90846	0.45423	6.4277089	3.2138545
-	-	1.852395	0.9261975

M^2 / df is distributed approximately as the t distribution (Hair 1998).

Table 7. Data point within and outside the 50% contour.

Smart phones screen		**Ordinary phone screen**	
15-39 years (%)	40-65 years (%)	15-39 years (%)	40-65 years (%)
14 (45.2)	8 (25.8)	13 (42)	15 (48.4)
17 (54.8)	23 (74.2)	18 (58)	16 (51.6)

presented in Table 7. Bold indicates values or points outside 50% contour, and italics indicate values or points within the 50% contour. The proportion of data point within and outside the contour is uniquely competitive except for the smart phone screen category 40 to 65 years with eight points outside the contour. Suppose the data set comes from normal distribution, we would expect precisely 50% of the data set to be within the contour. Hence the conclusion is self-evident. This is also evident in the chi squared plot shown in Figure 1 to 4 (Table 8).

Quantile quantile (Q-Q) plot

The Q-Q plots (Figure 5) showed similar patterns, and we conclude that the normality of the data set is not accepted. This implies that the data set are not normally distributed. From Figures 1 to 5 presented, influential observations can be identified. To buttress this point further, that is rejecting the normality of the data set, we apply the correlation

Table 8. Mahalanobis distance and chi squared values.

Smart mobile phones screen		Ordinary mobile phones screen	
15-39 Years		**15-39 Years**	
M^2	$Q_{k,2}(p)$	M^2	$Q_{k,2}(p)$
0.0017081	0.032521	0.1621569	0.032521
0.1302266	0.0991939	0.1721375	0.0991939
0.165057	0.1681662	0.1743244	0.1681662
0.1866501	0.2396024	0.3900729	0.2396024
0.2995156	0.3136849	0.414839	0.3136849
0.3138734	0.3906175	0.4763584	0.3906175
0.5960843	0.4706282	0.5449066	0.4706282
0.6512089	0.5539736	0.5450106	0.5539736
0.7056767	0.6409438	0.5841288	0.6409438
0.7765659	0.7318685	0.6461256	0.7318685
0.7819223	0.8271246	0.6815805	0.8271246
0.8139143	0.9271455	0.7434509	0.9271455
1.0566925	1.0324329	0.7515211	1.0324329
1.1485401	1.1435726	0.7532184	1.1435726
1.2084214	1.2612536	0.862211	1.2612536
1.2373811	1.3862944	0.9271973	1.3862944
1.460241	1.5196771	1.1114305	1.5196771
1.5102177	1.662595	1.2656346	1.662595
1.5816722	1.8165171	1.5548898	1.8165171
1.9995625	1.9832803	1.6676641	1.9832803
2.260878	2.1652239	1.7225079	2.1652239
2.3068354	2.3653908	1.9461622	2.3653908
2.6171358	2.5878421	2.0733263	2.5878421
2.6990123	2.8381684	2.3866202	2.8381684
2.8117413	3.1243701	2.5248068	3.1243701
2.8535359	3.4584782	3.1385107	3.4584782
3.1190666	3.8598196	4.3047466	3.8598196
4.810332	4.3624485	5.2846771	4.3624485
5.87816	5.0353929	5.3233629	5.0353929
6.3133007	6.0570442	8.3646561	6.0570442
7.7048703	8.2542688	8.5017642	8.2542688
40-65 years		**40-65 years**	
M^2	$Q_{k,2}(p)$	M^2	$Q_{k,2}(p)$
0.005354	0.032521	0.0535218	0.032521
0.0343637	0.0991939	0.1114296	0.0991939
0.0676586	0.1681662	0.1519237	0.1681662
0.0977192	0.2396024	0.1561735	0.2396024
0.1610065	0.3136849	0.2969942	0.3136849
0.2337153	0.3906175	0.3637109	0.3906175
0.3579438	0.4706282	0.3748524	0.4706282
0.3920615	0.5539736	0.4641833	0.5539736
0.4135661	0.6409438	0.517405	0.6409438
0.4837101	0.7318685	0.5542164	0.7318685
0.507592	0.8271246	0.8219912	0.8271246
0.5166047	0.9271455	0.8412319	0.9271455
0.6024532	1.0324329	0.8851362	1.0324329
0.6454104	1.1435726	1.2049245	1.1435726
0.6976622	1.2612536	1.2949876	1.2612536

0.7075887	1.3862944	1.356927	1.3862944
0.8800902	1.5196771	1.5108723	1.5196771
0.90846	1.662595	1.6322938	1.662595
1.0801275	1.8165171	1.7593006	1.8165171
1.0880286	1.9832803	1.852395	1.9832803
1.1086969	2.1652239	2.1652912	2.1652239
1.3039274	2.3653908	2.1973462	2.3653908
1.3637025	2.5878421	2.3072432	2.5878421
1.4444643	2.8381684	2.6414525	2.8381684
1.4832779	3.1243701	2.7351204	3.1243701
4.0737723	3.4584782	3.590688	3.4584782
4.6235312	3.8598196	3.674649	3.8598196
4.9309556	4.3624485	4.5798155	4.3624485
6.5507009	5.0353929	5.1815524	5.0353929
7.9022555	6.0570442	6.4277089	6.0570442
15.333599	8.2542688	8.2946616	8.2542688

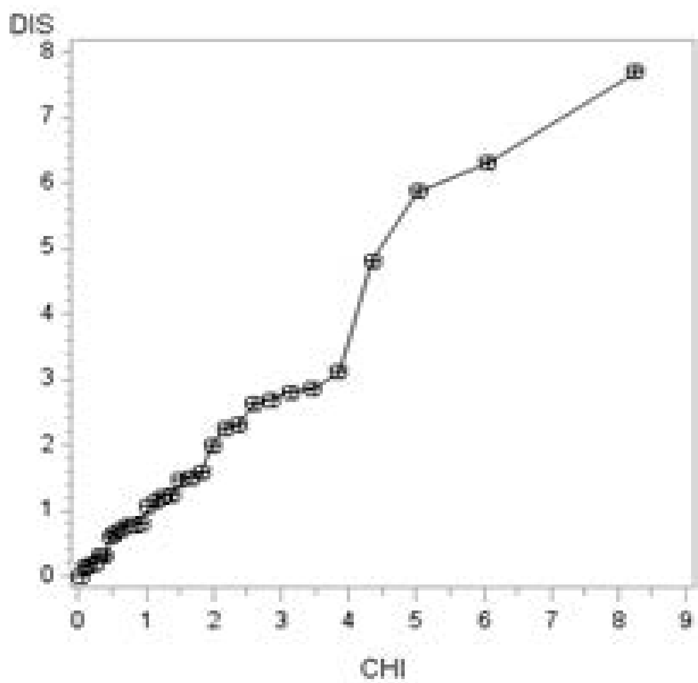

Figure 1. Chi squared plot for smart mobile phones screen (15-39 years).

coefficient to verify the aforementioned claims, and the straightness of the lines. Table 9 gives details of the computed values of the correlation coefficient C_c. At 10% level of significance with sample size $n = 31$, the table value is 0.972. This implies that the table value is greater than the computed value; hence we do not accept the hypothesis that the data set is normal. This is so since the computed value is less than the table value.

CONCLUSION

The results from the different graphical procedures and the test due to the correlation coefficient, and all techniques illustrates that the mobile phones screen data set are not normally distributed. This in general revealed that scratched

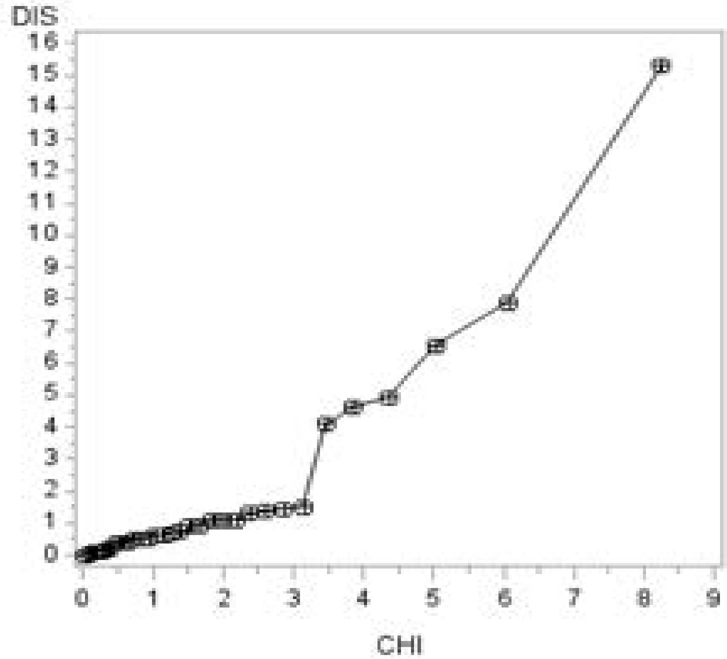

Figure 2. Chi squared plot for smart mobile phones screen (40-65 years).

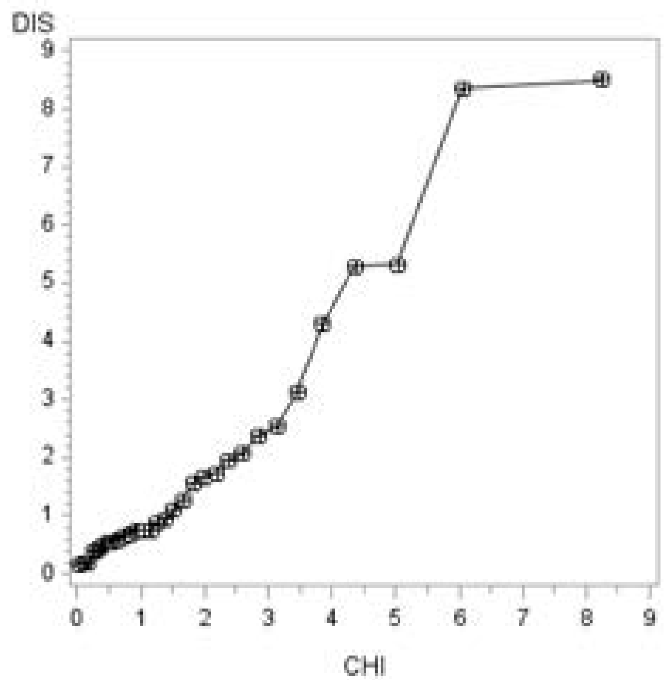

Figure 3. Chi squared plot for ordinary mobile phones screen (15-39 years).

and normal phone screens classified as smart mobile or ordinary mobile are not normally distributed. This discussion also adduces reasons for phone scratches but we are not in haste to conclude that occupation, target venue and personality profiles are factors (conditions) responsible for scratched

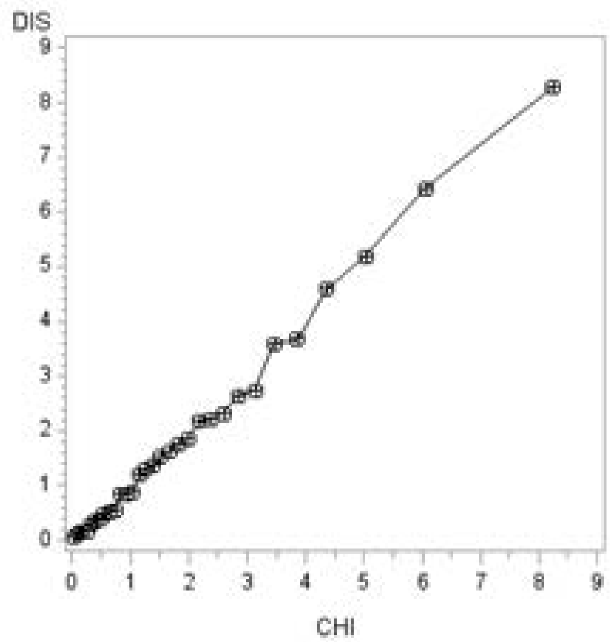

Figure 4. Chi squared plot for ordinary mobile phones screen (40-65 years).

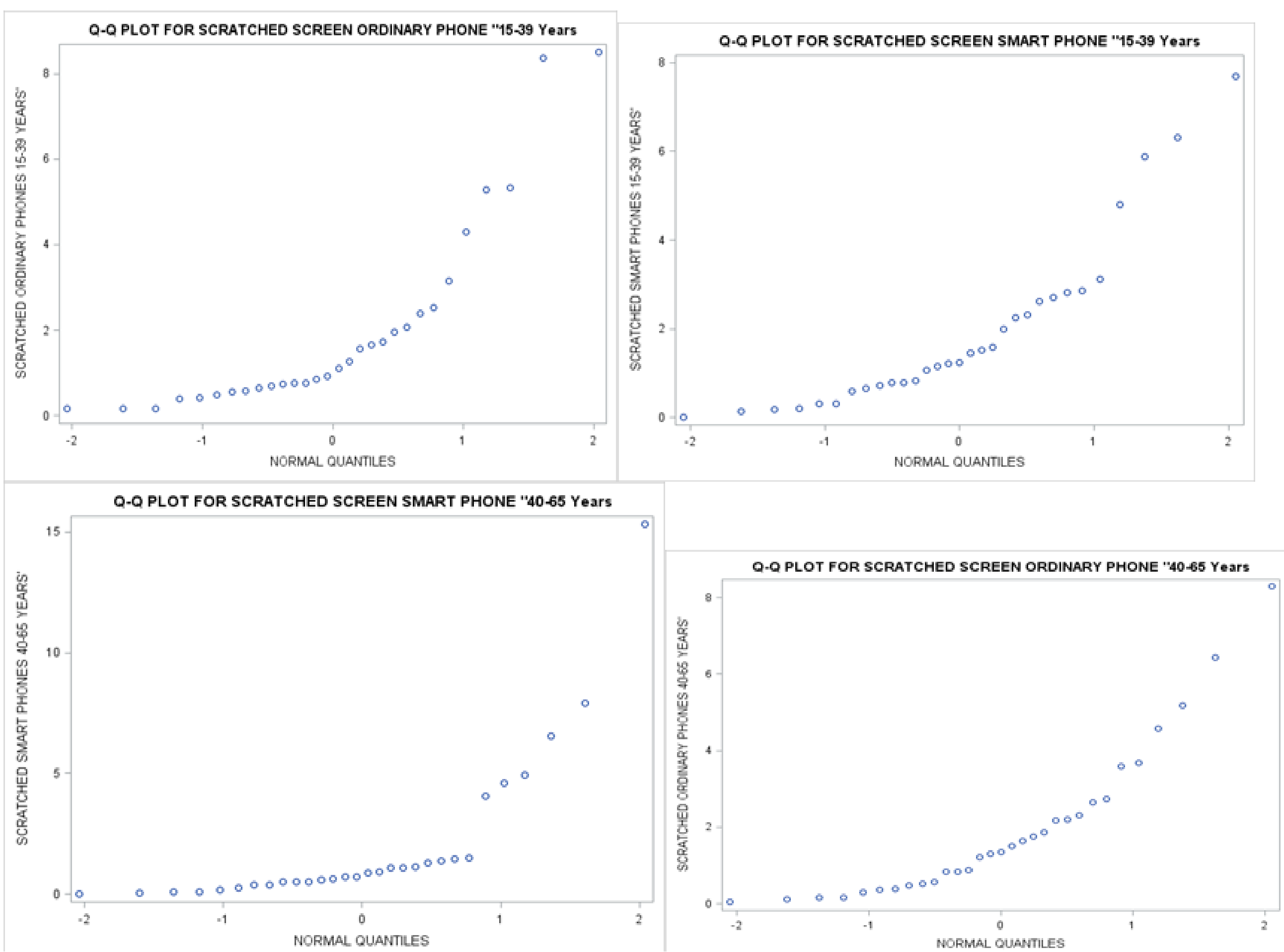

Figure 5. Q-Q plot for the mobile phone screen data.

Table 9. The values of the correlation coefficient

Smart mobile phone screens		Ordinary mobile phone screens	
15-39 years	40-65 years	15-39 years	40-65 years
0.803	0.669	0.663	0.769

mobile phone screens but in general, involuntary actions and domestic activities, actions may be responsible.

REFERENCES

Alqallaf, F.A., Konis, K.P., Martin, R.D., and Zamar, R.H. (2002). Scalablerobust covariance and correlation estimates for data mining, Proceedings of the Seventh ACM SIGKDD Internationa Conference on Knowledge Discovering and Data Mining, Edmonton, Alberta. pp.14-23.

Cox, D.R., and Small, N.J.H. (1978). Testing multivariate normality, Biometrika 65:263-272.

D' Agostino, R.B., and Stephens, M.A. (1986). Goodness of fit techniques (Marcel Dekker, New York).

Epps, T.W., and Pulley, I.B. (1983). A test for normality based on the emperical characteristic function. Biometrika 70:723-726.

Farrell, P. J., and Rogers-Stewart, K. (2006). Comprehensive study of tests for normality and symmetry: extending the Spiegelhalter test. J. Stat. Comput. Simul. 76:803-816.

Filliben, J. J. (1975). The probability plot correlation coefficient test for normality. Technometrics 17:111-117.

Fisher, R.A. (1915). Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. Biometrika 10:507-521.

Gnanadesikan, R. (1977). Methods for statistical analysis of multivariate data (John Wiley, New York).

Hair, J.F., Anderson, R.E., Tatham, R.L., and Black, W.C. (1998). Multivariate data analysis (Prentice Hall, New Jersey).

Johnson, R.A., and Wichern, D.W. (2008). Applied multivariate statistical analysis (Prentice -Hall International, Inc.

Kim, N., and Bickel, P. (2003). The limit distribution of a test statistic for bivariate normality. Statistica Sinica. pp. 13:327-349.

Kinnison, R. (1989). Correlation coefficient goodness of fit test for the extreme value distribution. Am. Stat. 43:98-100.

Leslie, J.R., Stephens, M.A., and Fotopoulos, S. (1986). Asymptotic distribution of the Shapiro-Wilk W for testing for normality. Ann. Stat. 14:1497-1506.

Liesenfeld, R., and Jung, R.C. (2000) Stochastic volatility models: conditional normality versus heavy tailed distributions. J. Appl. Econ. 15:137-160.

Machado, S.G. (1983). Two statistics for testing multivariate normality, Biometrika 70:713-718.

Maronna, R. A., and Yohai, V. (1998). Robust estimation of multivariate location and scatter, In "Encyclopedia of statistical sciences", Updated Volume 2 ((S. C. R Kotz and Banks , Eds), Wiley, New York).

Maronna, R. A., and Zamar, R. (2002). Robust estimation of location and dispersion for high dimensional datasets. Technometrics 44:307-317.

Mathur, S., and Dolo, S. (2008). A new efficient statistical test for detecting variability in the gene expression data, Statistical Methods Medical Research 17.

Min, I. (2007). A nonparamentric test of the conditional normality of housing demand. Appl. Econ. Lett. 14:105-109.

Muttlak, H.A., and Al-Sabah, W.S. (2003). Statistical quality control based on ranked set sampling. J. Appl. Stat. 30:1055-1078.

Oakland, J.S. (2008). Statistical process control (Butterworth Heinemann, UK).

Ramzan, S., Zahid, F.M., and Ramzan, S. (2013). Evaluating multivariate normality: A graphical approach. Middle-East J. Sci. Res. 13:254-263.

Romao, X., Delgado, R., and Costa, A. (2013). An empirical power comparison of univeriate goodness of fit test for normality. J. Stat. Comput. Simul. 80(5):545-591.

Rousseeuw, P.J., and Leory, A.M. (1987). Robust regression and outlier detection (Wiley-Interscience, New York).

Rousseeuw, P.J., and Van Zomeren, B.C. (1991). Robust distances: Simulation and cutoff values, In: "Directions in robust statistics and diagnostics" Part II (W. Stahel and Weisberge S., Eds) (Springer Verlag, New York).

Ryan, T.A., and Joiner, B.L. (1976). Normal probability plots and tests for normality, Statistics Department, The Pennsylvania State University.

Schoder, A.H., and Wilhelm, K.P. (2006). Preliminary testing for normality: Some statistical aspects of a common concept. Clin. Exp. Dermatol. 31:757-761.

Shapiro, S. S., and Wilk, M.B. (1968). A comparative study of various tests for normality. J. Am. Stat. Assoc. 63:1343-1372.

Surucu, B., and Koc, E. (2007). Assessing the validity of a statistical distribution:Some illustrative examples from dermatological research. Clin. Exp. Dermatol. 33:239-242.

Vamman, V., and Albing, M. (2007). Process capability indices for one sided specification intervals and skewed distributions. Q. Reliab. Eng. Intelligent 23:755-765.

Vogel, R. (1986). The probability plot correlation coefficient. Water Resour. Res. 22:587-590.

Zimmerman, D.W., Zumbo, B.D., and Williams, R.H. (2003). Bias in estimation and hypothesis testing of correlation. Psicologica 24:133-158.