STATISTICAL TESTS FOR FREQUENCY DISTRIBUTION OF MEAN GRAVITY ANOMALIES

By

S. I. Agajelu
Department of Surveying
University of Nigeria, Enugu Campus.
(Manuscript received March 6, 1979 and in revised form April 20, 1980)

ABSTRACT

The hypothesis that a very large number of $1^{o}x\ 1^{0}$mean gravity anomalies are normally distributed has been rejected at 5% Significance level based on the X^2 and the unit normal deviate tests. However, the 5^0 equal area mean anomalies derived from the $1^{o}x\ 1^{0}$data, have been found to be normally distributed at the same level of significance. It is concluded that $1^{o}x\ 1^{0}$anomalies may not be treated as random variables without systematic errors, at a global and hemispherical extent; whereas 5^0equal area anomalies derived from them can be so treated.

1. INTRODUCTION

In recent years, Geodetic Engineers, Scientists and Geophysical analysts have turned attention to the application of statistical models in the determination of certain physical parameters of geodetic and geophysical interest. For example, using such models, observed gravity anomalies or gravity gradient tensors can be used to predict and/or estimate unknown gravity anomalies, geoid unodulations, deflections of the vertical and the correlation of these quantities with geologic and geophysical structures.

The large number of published works in this area points to the potentiality of this modern approach. Kaula [1,2] discussed the method of applying statistical techniques in the analysis and prediction of gravity data. Krarup [3] showed that least squares statistical prediction of the anomalous potential is nothing but the least squares adjustment in Hilbert space with a kernel function, and developed a general least squares theory for estimating any element of the earth's gravity field using discrete and heterogeneous data. This became known as the least squares collocation. Meissl [4] using Hilbert space functions on the unit sphere, made a study of the-covariance of isotropic stochastic process on unit sphere, and applied that to derive covariance functions related to the earth's disturbing potential. Moritz [5] made a systematic and comprehensive presentation of the theory of least squares collocation and its application. Tsherning and Rapp [6] developed closed covariance expressions for the components of the anomalous potential. Rapp and Agajelu [7] applied the least squares collocation to the upward continuation of gravity anomalies and found results which compared favourably with corresponding anomalies obtained from the determination Poisson Integral Equation. Pellinen [8] discussed the application of statistical modes in the estimation of the accuracy of astronomical leveling. Efforts have continued to be directed towards the improvement of statistical models for use in these estimations.

One basic measurement for use with the above models is the gravity anomaly. Clearly statistical models are developed for random quantities and the assumption has been that gravity anomalies or other elements of the anomalous potential are random quantities over the globe. Statistical solutions of geodetic problems will be successful to the extent that the assumptions on the observables are true. It is therefore relevant to examine the statistical behaviour of certain mean gravity anomalies which may be used with the statistical models. In the sequel, the frequency distributions of sets of $1^0 \times 1^0$ and 50 equal mean gravity anomalies are tested for normality. These anomalies are made up as follows:

$1^0 \times 1^0$ anomalies

(a) 24,608 gravity anomalies of Northern Hemisphere
(b) 11,54 gravity anomalies of Southern Hemisphere
(c) 36,149 gravity anomalies of the whole globe, made up from the sum of (a) and (b) above.

5^o Equal Area Anomalies:

(d) 813 gravity anomalies of Northern Hemisphere
(e) 669 gravity anomalies of Southern Hemisphere
(f) 1482 gravity anomalies of the whole globe, made up from the sum of (d) and (e) above.

The Chi-square test was made for each set in turn, using a batch interval of 8mgals for the $1^0 \times 1^0$ data and 4 mgals for the 5^o equal area data. In addition the "unit normal deviate form" of residuals (the mean anomalies themselves) was used to test the $1^0 \times 1^0$ set for normality.

2. THEORETICAL CONSIDERATIONS

Two statistical concepts have been applied in this study:

a) The Chi-square (X^2) as a goodness- of- fit test, and
b) The "Unit normal deviate" form of residual test for normality.

2.1 The chi-square as a goodness-of- fit test:

Following Freund [9] we write the Chi-square density function as;

$$f(x^2) = 2\frac{1}{\frac{\upsilon}{2}}\, \Gamma(\tfrac{\upsilon}{2})\, (x^2)^{\frac{\upsilon-2}{2}}\, e^{-\frac{x^2}{2}} \text{ for } x^2 > 0 \quad (1)$$

for $x^2 \leq 0$

where $\Gamma(\frac{\upsilon}{2})$ is the well-known gama function *[9];V* is the degree of freedom of the distribution. The values of x^2_υ, α such that the probability that *the* $x^2_\upsilon, x^2_{\upsilon,\alpha}$ is α, where α is the statistical level of significance, are usually tabulated in text books of statistics. If υ observations are made from a normal population with zero mean and unit variance, the sum of the squares of the observations is distributed as x^2 with υ degrees of freedom (Hamilton [10]). However, where the parameters of the distribution function are estimated from the observations, the degree of freedom will be reduced by the number of parameters so estimated. In this study, two parameters are estimated from the data and the degree of freedom is therefore reduced by two. This test is used to determine whether or not any group of observation belongs to any specified distribution, whatsoever. Usually, the observations are divided into a convenient number, of batch or class intervals, k. The number of observations, f_i lying in each interval is then found. Choosing a specific distribution (in our case, the normal distribution), the theoretical number of observations, f_i lying in the same interval as f. is then computed. Considering all the intervals the x^2 statistic is computed from the following:

$$x^2 = \sum_{i=1}^{k} [\frac{f_i - F_i}{F_i}] \qquad (2)$$

This is distributed as x^2 where υ = k-2 and k=23 in this study. The computation of F_i requires more explanation. We can write $F_i = p_i N$ where p_i is the probability that an observation falls in the class interval i, and N is the total number of observations. The p_i values are obtained by first evaluating the probabilities at the interval terminals, and finally taking the differences between adjacent values. The probabilities themselves are computed from an assumed normal distribution function with the mean equal to the weighted average $\hat{u}$ of the batch interval values of observations, and the variance equal to the weighted average variance $\hat{\sigma}^2$ of the interval values. These parameters were computed from equation (3) and (4) below:

$$\mu = \frac{\sum_{i=1}^{v} f_i \xi_i}{\sum_{i=1}^{k} f_i} \quad (3)$$

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{k} f_i(\xi_i - \hat{\mu})^2}{\sum_{i=1}^{k} f_i} \quad (4)$$

where k is the total number of intervals used, and ξ_i is the middle value of the ith intervals. In practice, probability evaluation was made for the standard normal distribution with $\hat{\mu}$= 0 and $\hat{\sigma}^2$ =1, using the IBM scientific NDTR. (IBM scientific subroutine package p.78). To obtain the theoretical variables at the interval terminals corresponding to this standard normal distribution, one normalizes his variables by

$$x_i^1 = \frac{x_i - \hat{\mu}}{\hat{\sigma}} \quad (5)$$

where $\hat{\mu}$ and $\hat{\sigma}$ are obtained from (3) and (4) above, and are the class interval terminal values chosen. Using (5), the P_i values are computed from (6) below

$$p_i = p(x_i^1) - p(x_{i-1}^1) \quad (6)$$

where $P(x^1)$ is the cumulative normal distribution function. which is evaluated by the scientific subroutine at the points x^1_{i-1}, x^1_i, x^1_{i+1} etc

2.2 The "Unit Normal Deviate" form of residuals test for normality:

Following Draper and Smith [11] we assume that the anomalies follow a normal distribution with zero mean and a population variance, σ^2 *ie* $.N(o,\sigma^2)$. This variance can be replaced by the sample variance, 0^2 estimated from the observations as follows:

$$\hat{\sigma}^2 = \frac{\sum \Delta g_i^2}{N-1} \quad (7)$$

where Δg_i are the observed gravity anomalies. If this were true, then the unit normal deviate from the residuals,

$\frac{v_i}{\hat{\sigma}}$ is distributed as a standard normal distribution with zero mean ands unit variance ie. $\frac{v_i}{\hat{\sigma}} \sim N(o,1)$ where the vi's are the residuals. At 5% significance level, $\frac{v_i}{\hat{\sigma}}$ will fall between the limits + 1.96 and -1.96 or roughly +2 and -2 as used in this study.

3. RESULTS

Table 1 shows the frequency distributions of both the 1^0xl^0 and 5^0 equal area mean anomalies for the Northern and Southern hemispheres and for the whole globe. The statistic computed form (2) above using this data, can now be compared with X.05,21= 32.671 (*[9]* p.438). The computed values for the six sets of data are given in Tables 2 and 3. Table 4 shows the percentages attained in the "unit normal deviate" form of residuals.

By comparing the figures in column 5 of Table 2 with X^2.05, 21= 32.671 one finds that the Chi-square values for the three sets of 10xl° anomalies are each much larger than 32.671. Further, out of the three sets, the global statistic is worse than the other two hemispherical sets.

On the other hand comparison of the figures in column 5 of Table 3 with x^2.05, 21 shows that the values are each less than 32.671.

In Table 4 which gives the result of the unit normal deviate-form of residual computation, for 1^0xl° anomalies only, the percentages of the whole data whose unit normal deviate follow (0,1) are given in column 5 and 6. Column 5 represents the case then the anomalies are taken to have zero mean directly. Column 6 represents the case when the mean of each set is subtracted from individual anomalies before the test was performed. From these two columns, it can be seen that no set has attained the 95% probability required for-normality.

4. CONCLUSION

From the results of this study one concludes as follows:

TABLE 1*

Batch interval (mgals)	Frequency of 1^0 x 1^0 anomalies frequency of 5^0 ea. Anomalies						
	Northern hemisphere	Southern hemisphere	GLOBAL	Batch interval (mgals)	Northern hemisphere	Southern hemisphere	GLOBAL
-84&less	204	98	302	-42& less	7	2	9
-84to -76	46	25	71	-42 to -38	6	1	7
-76 "- 68	75	60	135	-38" -34	15	5	20
-68 " -60	146	80	226	-34" -30	10	7	17
-60" -52	244	155	399	-30" -26	17	17	30
-52" -44	468	286	754	-26" --22	24	25	49
-44" -36	795	427	1222	-22" -18	44	33	77
-36"-28	1230	602	1832	-18" -14	46	50	96
-28" -20	2056	1054	3110	-14" -10	54	43	97
-20 " -12	2531	1201	3732	-10" -6	82	58	140
-12 " -4	2966	1441	4407	-6" -2	68	81	149
+4 " -4	3223	1494	4717	-2" -2	72	89	161
+4" -12	2971	1216	4187	2" -6	87	63	150
+12" +20	2341	958	3299	6" -10	77	52	129
+20" +28	1926	703	2629	10" -14	59	36	95
+28" +36	1216	515	1731	14" -18	40	31	71
+36" +44	810	372	1182	18" -22	34	24	58
+44" +52	475	236	711	22" -26	25	18	43
+52" +60	325	191	516	26" -30	18	14	32
+60" +68	161	117	278	30" -34	14	10	24
+68" +76	139	77	216	34" -38	7	4	11
+76" +84	79	50	129	38" -42	3	0	3
+84&over	181	183	364	+42+over	4	10	14

*The data was supplied by Prof. R.H. App, Department of Geodetic Science. The Ohio State University Columbus, ohio, U.S.A.

TABLE 2

1^0 x1^0 mean anomalies				
Region	Total No.	Group mean	Standard Deviation	Chi-square
Northern Hemisphere Southern Hemisphere Global	24,608 11.541 . 36.149	0.29 -1.23 -0.19	25.79 27.35 26.31	356.08 313.23 623.93

TABLE 3

5^0 Equal Area Anomalies				
Region	Total no.	Group mean	Standard deviation	Chi-square
Northern hemisphere	813	-0.53	15.99	21.91
Southern hemisphere	669	-1.23	14.59	23.68
Gobal	1482	-0.85	15.38	22.43

TABLE 4

1^0 x1^0 mean anomalies					
Region	Total No	$\frac{v_i}{\hat{\sigma}} < 2$ for $\mu = 0$	$\frac{v_i}{\hat{\sigma}} < 2$ for $v = \Delta g - \mu$	Percentage(%) when mean = 0	% when mean is subtracted from all anomalies
Northern hemisphere	24608	19855	19690	80.68	80.01
Southern hemisphere	24851	20532	20532	82.62	82.62
Global	36149	28149	28726	79.46	79.46

1. None of the $1^{\circ}x1^{\circ}$ mean anomally sets satisfies the test criterion. This leads to the rejection at 5% significance level, of the null hypothesis that they are normally distributed. Again, none of these sets satisfies the "unit normal deviate" percentage criterion for normality.

2. Consequent upon 1 above, statistical models derived on the basis of normality may not be app lied to $1^{0}x1^{0}$ gravity anomalies treated as random variables. Further, these anomalies may not be considered as random variables, without systematic errors. If they are, their statistical distribution is yet unknown.
3. The 5^{0} equal area mean anomalies show that the null hypothesis that they are normally/distributed at 5% significance level cannot be rejected.
4. Statistical models can be freely applied to 5^{0} equal area mean anomalies taken as random variables, at global and hemispherical extent.

REFERENCES

1. Kaula,W.M. ,Statistical and Harmonic analysis of Gravity,J.G.R.Vol.64, pp 2401-2422,1959.

2. Kaula,W.M.,Determination of the earth's gravitational field. Reviews of Geophysics,Vol.1No.4,pp 507-551, 1963.

3. Krarup, T. ,A contribution to the mathematical foundation of physical geodesy.Publication No.44 Danish Geodetic Instit.Copenhagen, 1969.

4. Meissl,P., Advanced least Squares Methods. Reports of Dept. of Sci.No.151 O.S.U. Columbus, Ohio, 1971.
5. Moritz H., Advanced least Squares Methods. Reports of Dept of Geod. Sci. No.175, O.S.A. Columbus ,Ohio, 1972.

6. Tsherning C. & Rapp.R.H., Closed covariance expressions for gravity anomalies, geoid undulations, and deflections of the vertical implied by anomaly degree variance models. Reports of Dept. of Geodetic Science No. *208,* O.S.U. Columbus, Ohio, 1974.

7. Rapp,R.H. & Agajelu S.I., Comparison of upward continuation of gravity anomalies computed by the method of least squares collocation and Poisson's Integral, Reports of Dept. of Geodetic Science No.227. O.S.U. Columbus ,Ohio, 1975.

8. Pellinen,L.P., On accuracy of astronomical leveling, paper presented at the International symposium on geodetic measurements and computations, A.B.U Zaria, Nigeria, 1978.

9. Freund,J.E., Mathematical Statistics 2nd ed., Prentice-Hall,Inc. Englewood cliffs, New Jersey, U.S.A., p. 212, 1971.

10. Hamilton,W.C.,Statistics in Physical Science , the Ronald press Co. ,New York, U:S.A.,p.BI, 1964.

11. Draper ,&Smith., Applied Regression Analysis.John Wiley &Sons Inc.,p.88 1966.

12. Agajelu S.I., The Role of gravity information in national geodetic computation; paper. presented at the International symposium on geodetic computations and measurements, A.B.U.,Zaria,Nigeria,1978.