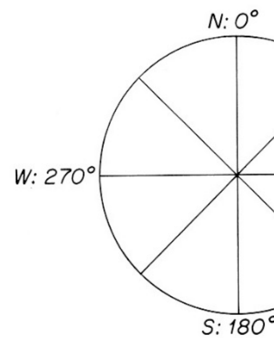
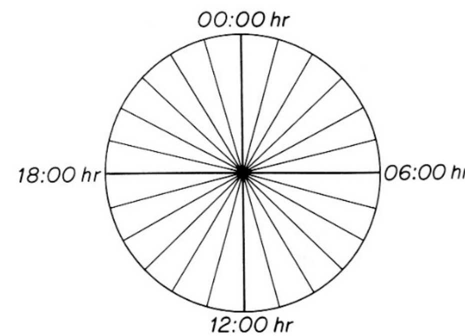
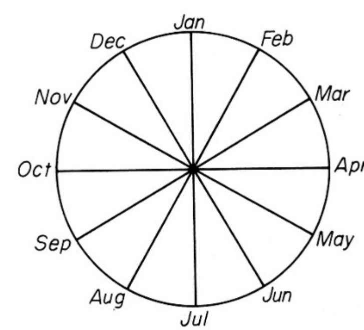
Directional (Circular) Statistics

Directional or circular distributions are those that have no true zero and any designation of high or low values is arbitrary:

- Compass direction
- Hours of the day
- Months of the year







Time can be converted to an angular measurement using the equation:

$$a = \frac{(360^0)(X)}{k}$$

where *a* is the angular measurement, *X* is the time period, and *k* is the number of time units on the circular measurement scale.

What is the angular measurement of 6:15 a.m. (6.25a.m.)?
(Remember to use a 24hr clock…)

$$a = \frac{(360^0)(6.25hr)}{24hrs} = 93.75^0$$

What is the angular measurement of February 14th?
(Remember to use total days…)

$$a = \frac{(360^0)(45th\ day)}{365\ days} = 44.38^0$$

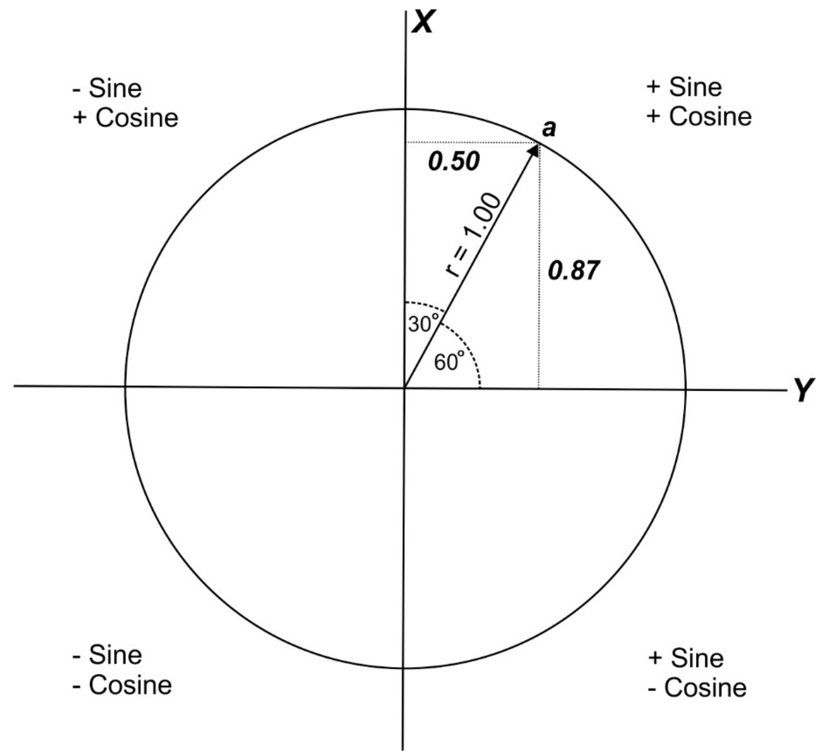
To analyze directional data they must first be transformed into rectangular polar coordinates.

- First, we specify a ‘unit circle’ that has a radius of 1.
- The polar location is then defined as the angular measurement and its intersection with the unit circle.
- The cosine and sine functions are then used to place this location (based on the angle and unit distance) into a standardized Cartesian space.

$$\cos a = \frac{x}{r} \qquad \sin a = \frac{y}{r}$$

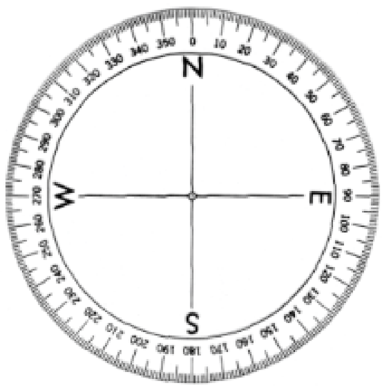
$$\cos 30 = 0.50 \qquad \sin 30 = 0.87$$

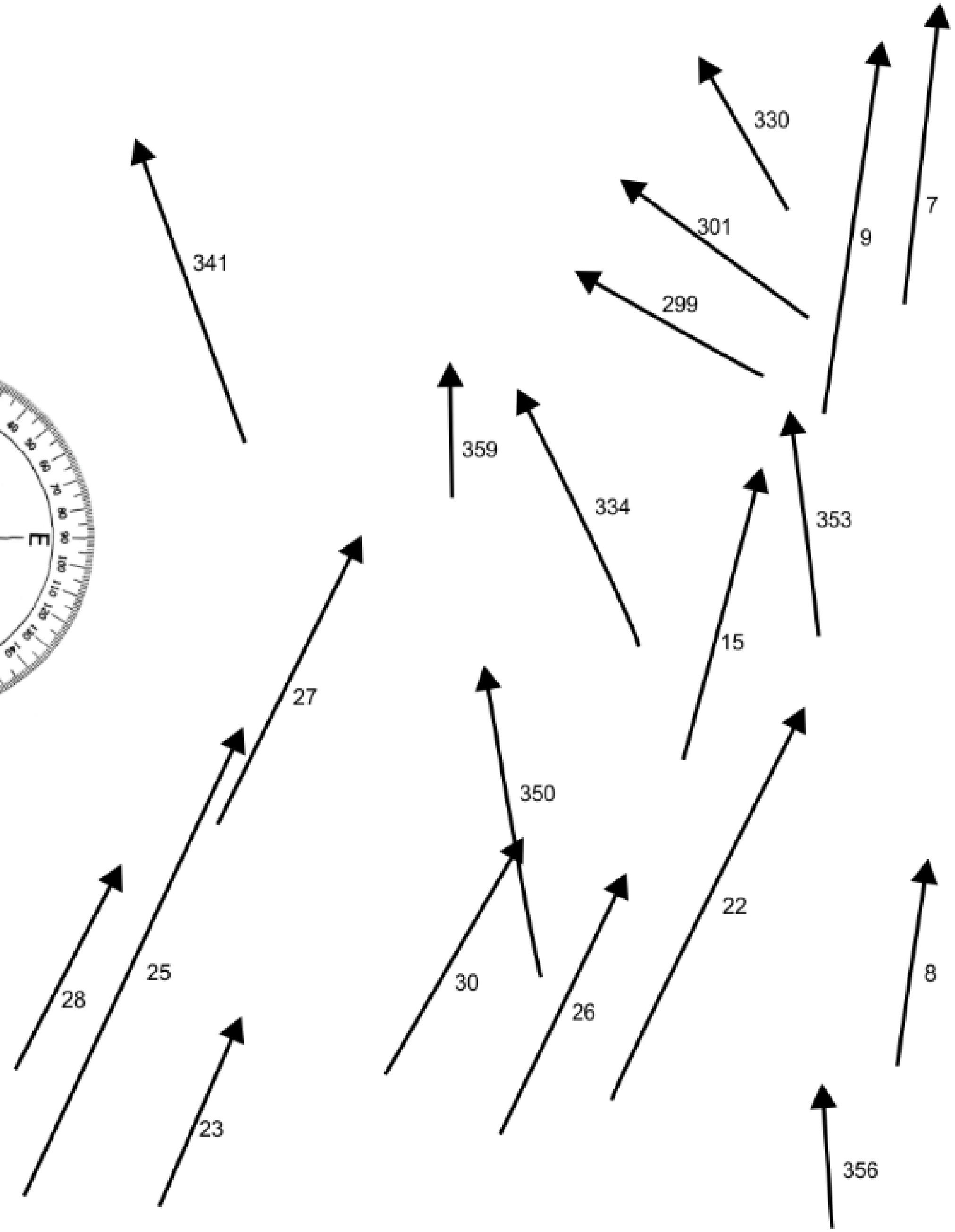
$$\cos 60 = 0.50 \qquad \sin 60 = 0.87$$



Note that the coordinates of opposite angles are identical. Also note that the x and y axes are opposite of the typical Cartesian plane.

Death Valley Moving Rocks, vectors (degrees)





Mean Angle (Azimuth)

The mean angle can not simply be the sum of the angles divided by the sample size, because the mean angle of 359º and 1º (north) would be 180º (south)! Therefore we use the following equations:

$$Y = \frac{\sum_{i=1}^{n} \sin_a}{n} \qquad X = \frac{\sum_{i=1}^{n} \cos_a}{n}$$

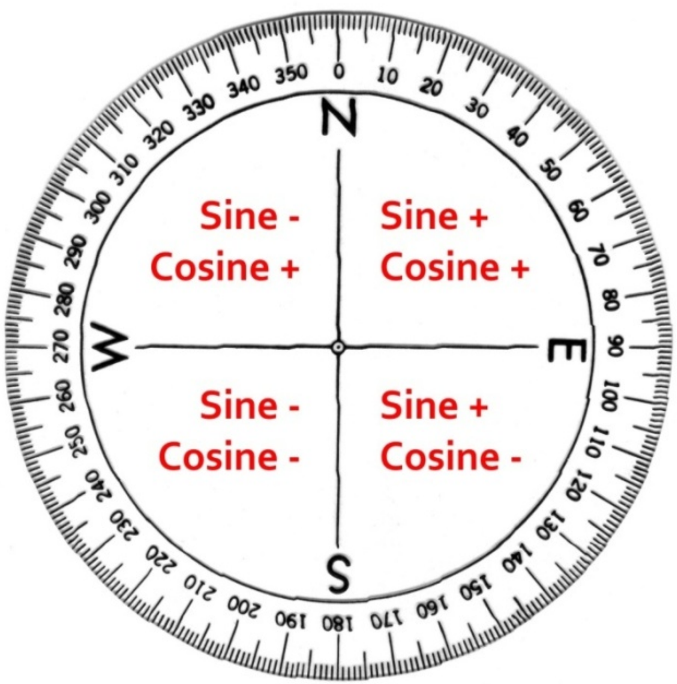
$$r = \sqrt{X^2 + Y^2}$$

$$\cos \bar{a} = \frac{X}{r} \qquad \sin \bar{a} = \frac{Y}{r} \qquad \theta_r = \arctan\left(\frac{\sin \bar{a}}{\cos \bar{a}}\right)$$

where *X* and *Y* are the rectangular coordinates of the mean angle, and *r* is the mean vector.

Determining the Quadrant

- Sin +, Cos + : the mean angle is computed directly.
- Sin +, Cos - : the mean angle = 180 – θ_r.
- Sin -, Cos - : the mean angle = 180 + θ_r.
- Sin -, Cos + : the mean angle = 360 - θ_r.



Rocks Vectors		Sin (Azimuth)	Cos (Azimuth)
341		-0.32557	0.94552
330		-0.50000	0.86603
301		-0.85717	0.51504
299		-0.87462	0.48481
9		0.15643	0.98769
7		0.12187	0.99255
359		-0.01745	0.99985
334		-0.43837	0.89879
353		-0.12187	0.99255
15		0.25882	0.96593
27		0.45399	0.89101
28		0.46947	0.88295
25		0.42262	0.90631
23		0.39073	0.92050
350		-0.17365	0.98481
30		0.50000	0.86603
26		0.43837	0.89879
22		0.37461	0.92718
8		0.13917	0.99027
356		-0.06976	0.99756
	∑	0.34763	17.91415

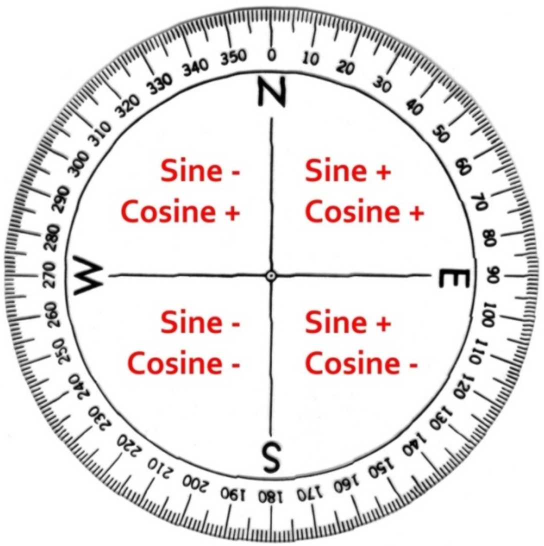
First take the sine and cosine of the angles (azimuths) and sum them.

In Excel the formula is*: =sin(radians(cell #))* and *=cos(radians(cell #)).*

$n = 20$

$\sum \sin_a = 0.34763 \qquad \sum \cos_a 17.91415$

$$Y = \frac{0.34763}{20} = 0.01738 \qquad X = \frac{17.91415}{20} = 0.89571$$



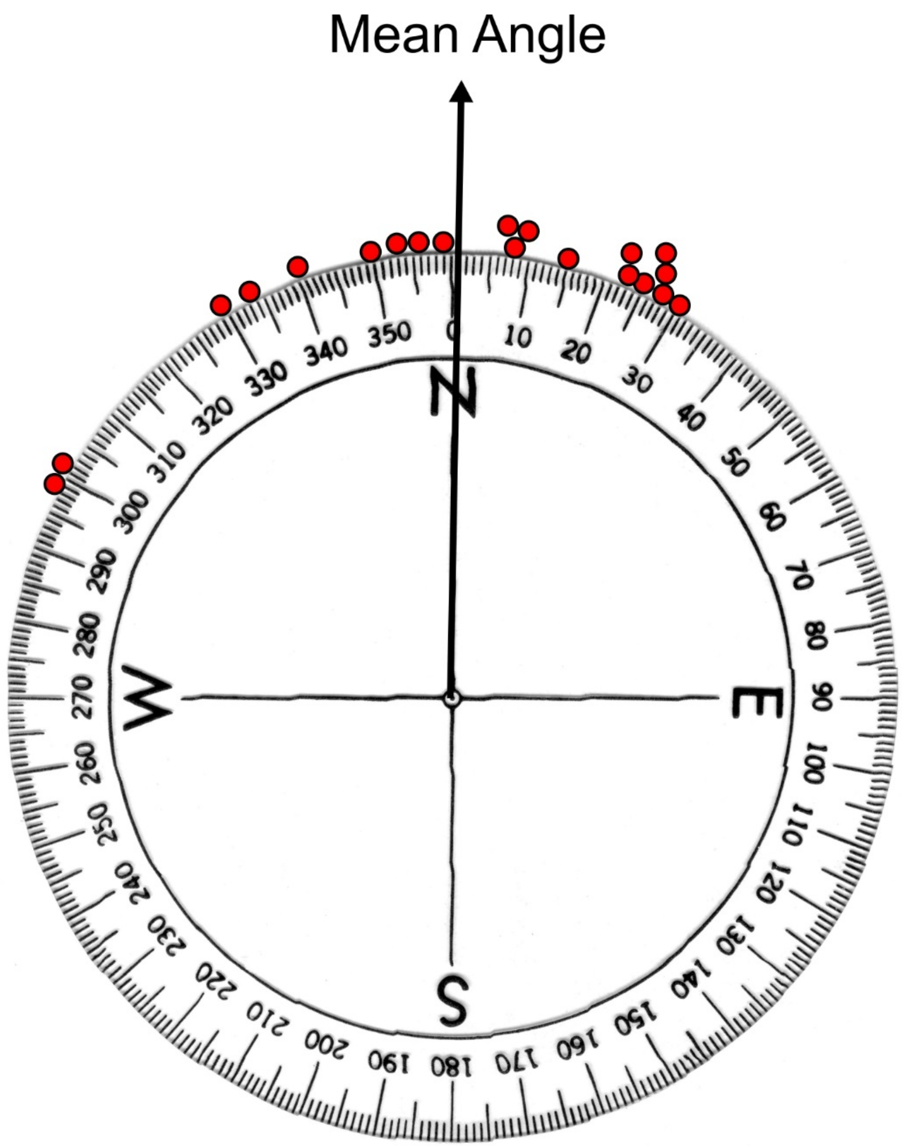
$$r = \sqrt{0.01738^2 + 089571^2} = \sqrt{0.00030 + 0.80229} = 0.8959$$

$$\sin \bar{a} = \frac{0.01738}{0.8959} = 0.0194 \qquad \cos \bar{a} = \frac{0.89571}{0.8959} = 0.9998$$

$$\theta_r = \arctan\left(\frac{0.0194}{0.9998}\right) = 1.11$$

Ignore the sign.

Since the sine is + and the cosine is +, we read the answer directly, ***regardless of the sign of*** Θ_r. Therefore the angle that corresponds to cos(0.9998), sin(0.01941) is 1.11^0... essentially due north.

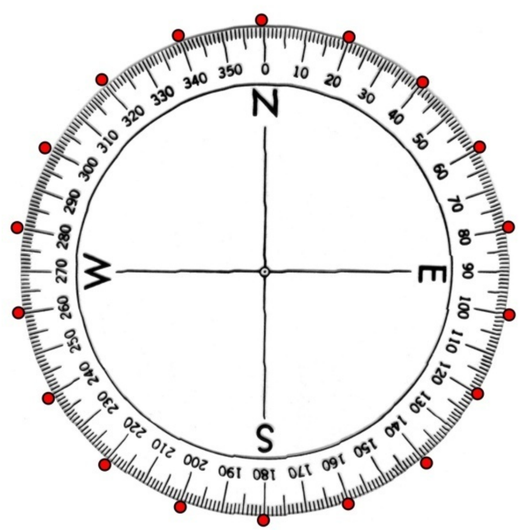

Mean Angle
N
E
S
W
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
210
220
230
240
250
260
270
280
290
300
310
320
330
340
350

The value of *r* is also a measure of angular dispersion, similar to the standard deviation with a few exceptions:

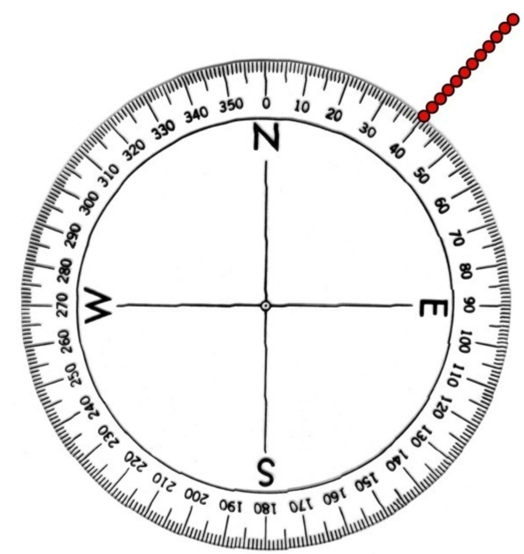
- Unlike the standard deviation it ranges from 0 – 1.
- A value of 0 means uniform dispersion.
- A value of 1 means complete concentration in one direction.

$$r = \sqrt{X^2 + Y^2}$$

r = 0



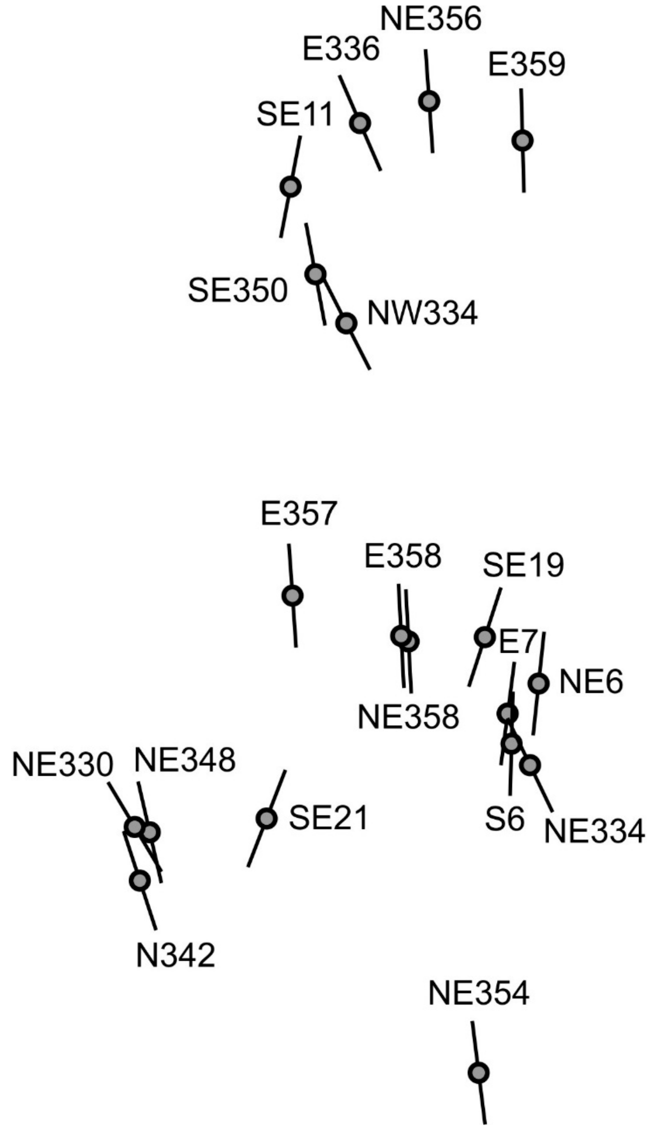
r = 1

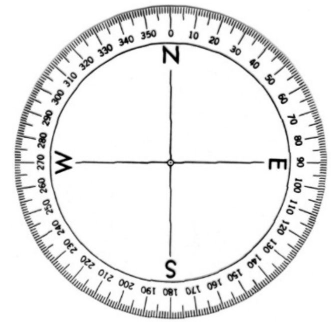


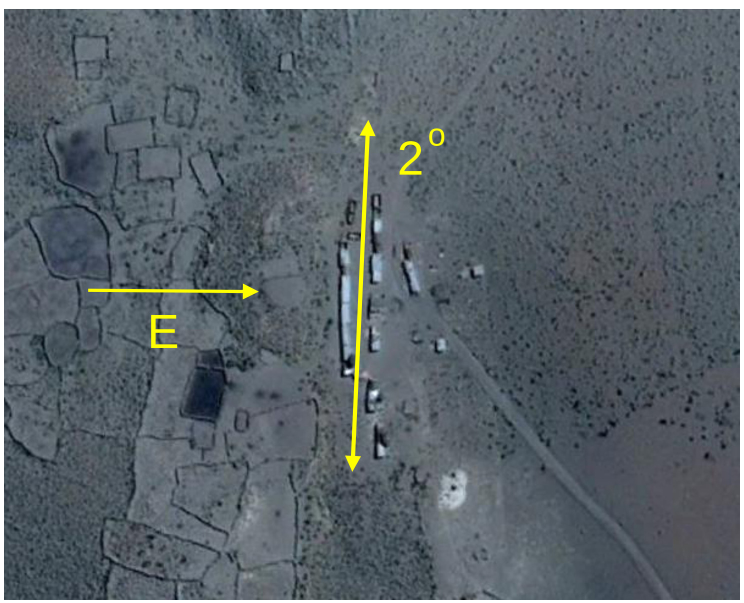
Andean Villages
Principal Street Azimuths
(labeled slope direction, degrees)

Note that these azimuths are not truly unidirectional but are bidirectional.

For example, azimuth 330 has an opposite azimuth of 330-180=150.

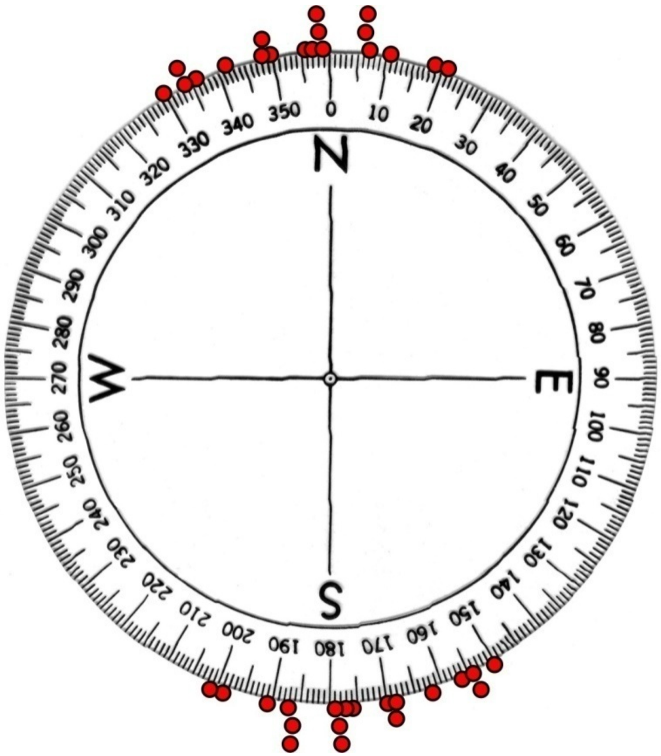




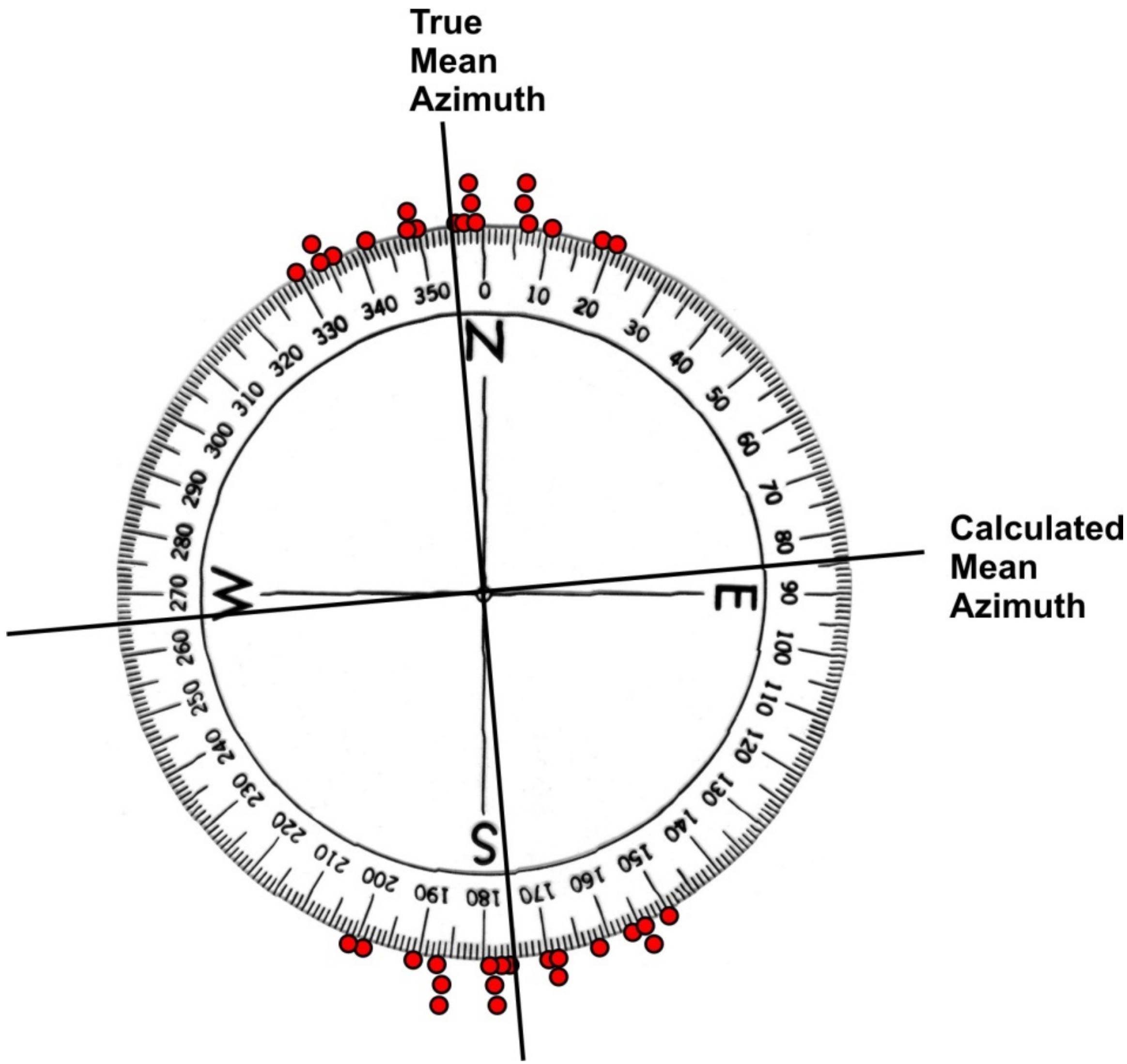


Bimodal Data

When angular data have opposite azimuths they are said to have *diametrically bimodal* circular distributions.



The mean angle of diametrically bimodal data is orthogonal (*at right angles*) to the true mean angle and a poor representation of the data.



One method of dealing with diametrically bimodal circular data is to use a procedure called *angle doubling*.

Data must be perfectly diametrical for this procedure to work.

<u>Angle Doubling Procedure:</u>

- Each angle (a_i) is doubled (e.g. 34 x 2 = 68).
- If the doubled angle is < 360º then it is recorded as $2a_i$.
- If the doubled angle is ≥ 360º then 360 is subtracted from it and the results recorded as $2a_i$.
- Then proceed as normal in calculating the mean angle.

Andean Village Subset (n=30) Principal Azimuths (diametrically bimodal data)

$356(2) - 360 = 352$

Village Azimuth (a_i)	$2a_i$	$\text{Sin}(2a_i)$	$\text{Cos}(2a_i)$
356	352	-0.1392	0.9903
359	358	-0.0349	0.9994
11	22	0.3746	0.9272
350	340	-0.3420	0.9397
334	308	-0.7880	0.6157
357	354	-0.1045	0.9945
358	356	-0.0698	0.9976
358	356	-0.0698	0.9976
19	38	0.6157	0.7880
7	14	0.2419	0.9703
6	12	0.2079	0.9781
6	12	0.2079	0.9781
334	308	-0.7880	0.6157
21	42	0.6691	0.7431
348	336	-0.4067	0.9135
176	352	-0.1392	0.9903
179	358	-0.0349	0.9994
191	22	0.3746	0.9272
170	340	-0.3420	0.9397
154	308	-0.7880	0.6157
177	354	-0.1045	0.9945
178	356	-0.0698	0.9976
178	356	-0.0698	0.9976
199	38	0.6157	0.7880
187	14	0.2419	0.9703
186	12	0.2079	0.9781
186	12	0.2079	0.9781
154	308	-0.7880	0.6157
201	42	0.6691	0.7431
168	336	-0.4067	0.9135
	Σ	**-0.8515**	**26.8976**

$$Y = \frac{-0.8515}{30} = -0.0284 \qquad X = \frac{26.8976}{30} = 0.8966$$

$$r = \sqrt{(-0.0284)^2 + (0.8966)^2} = \sqrt{0.8047} = 0.8970$$

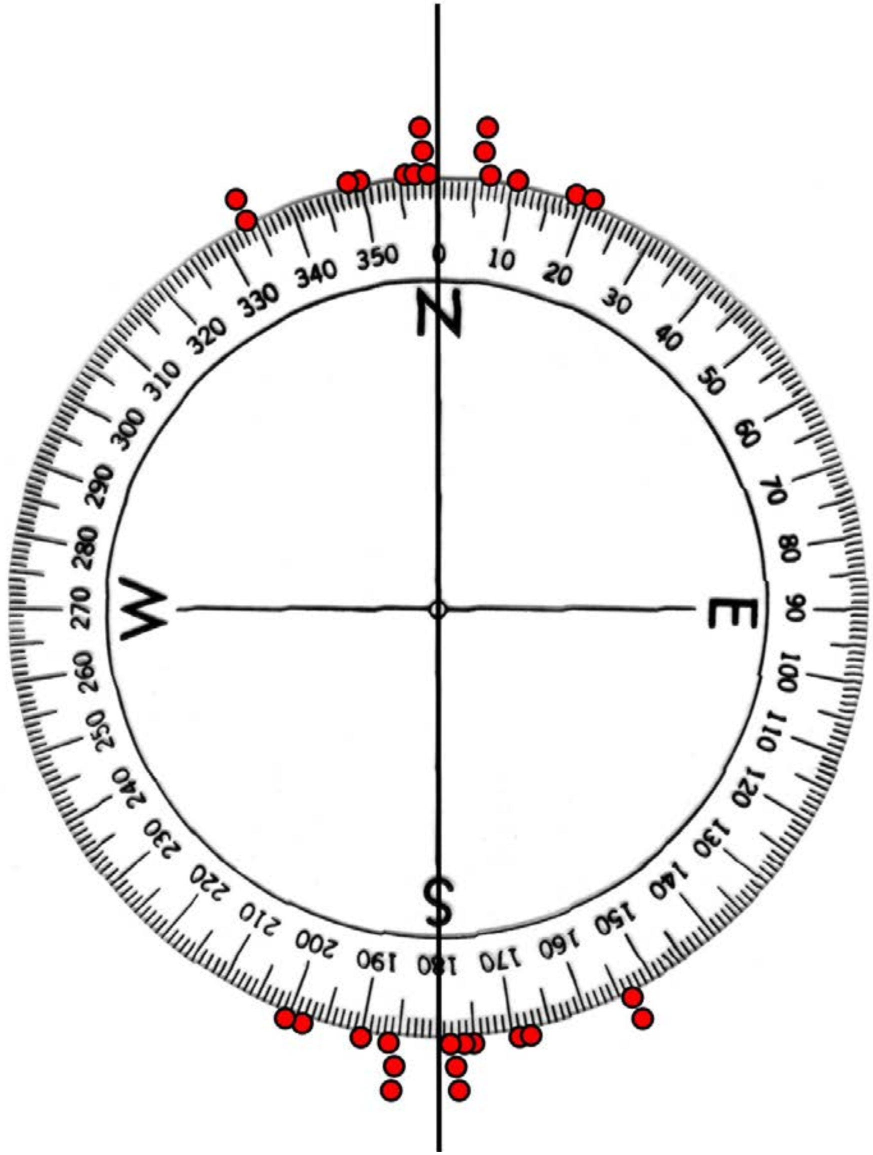
$$\cos 2a_i = \frac{0.8966}{0.8970} = 0.9996$$

$$\sin 2a_i = \frac{-0.0284}{0.8970} = -0.0317$$

$$\arctan\left(\frac{-0.0317}{0.9996}\right) = -0.0005$$

The mean angle is therefore 360-0.0005 or effectively 360º.

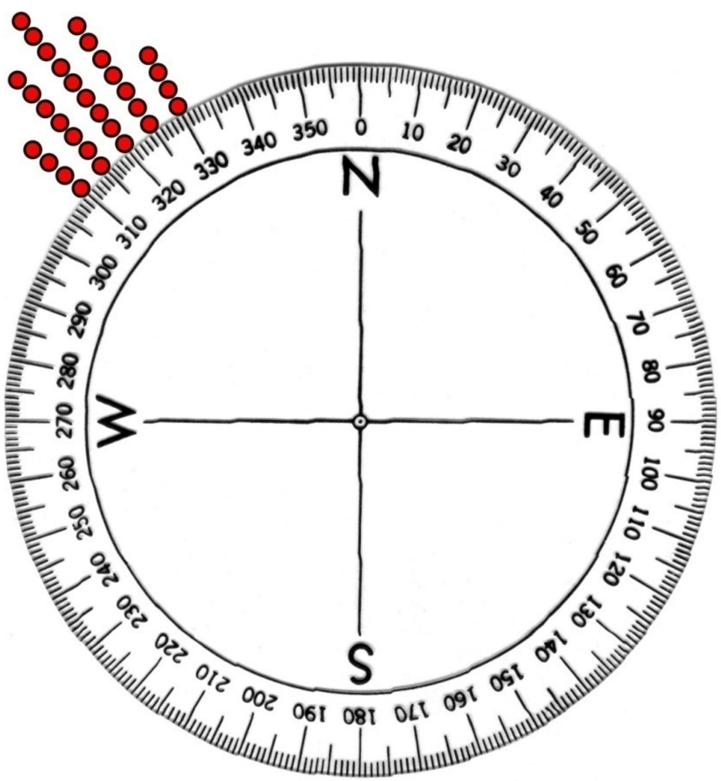
Mean Angle

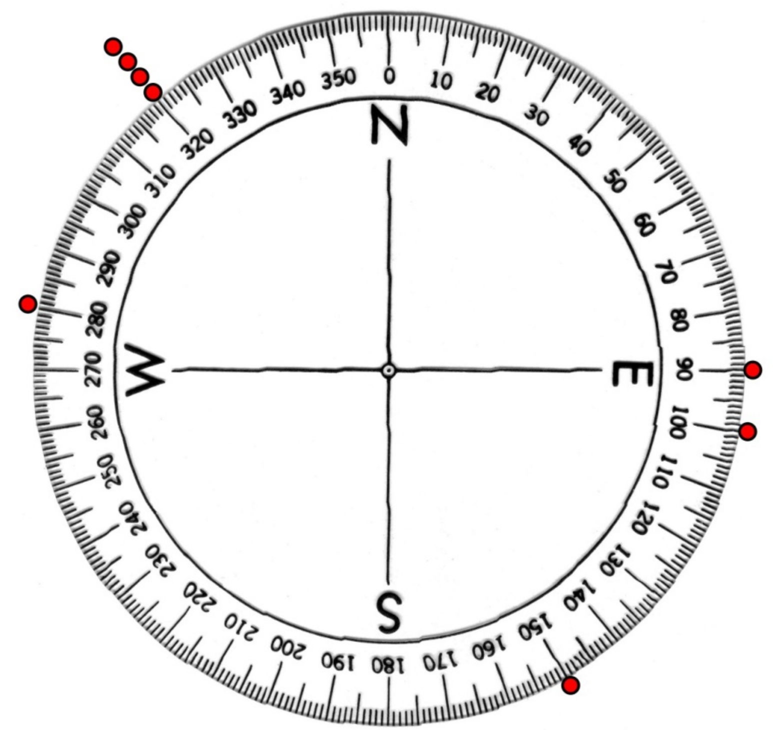


N
E
S
W
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
210
220
230
240
250
260
270
280
290
300
310
320
330
340
350

Testing the Significance of the Directional Mean

Directional statistics are more sensitive to small sample sizes and it is important to test the directional mean for significance… not something typically done with conventional measures.





Rayleigh z Test

We can use the Rayleigh z test to test the null hypothesis that there is no sample mean direction:

H_0: There is no sample mean direction.

H_a: There is a sample mean direction.

We determine the Rayleigh z statistic using the equation:

$$z = nr^2$$

where n is the sample size and r is taken from the mean angle equation.

The Rayleigh’s test has a few very important assumptions:

- The data are not diametrically bidirectional.
- The data are unimodal, meaning there are not more than one clustering of points around the circle.

So for the Andean village azimuth which are diametrically bidirectional we must use the data from the angle doubling procedure.

From the Andean Village example:

$n = 30$

$r = 0.897$

$z = (30)(0.897^2) = 24.14$

$z_{Critical} = 2.971$

Critical z Values for the Rayleigh's Test

Taken from Zar, 1981 Table B.32

n	α: 0.50	0.20	0.10	0.05	0.02	0.01	0.005	0.002	0.001
6	0.734	1.639	2.274	2.865	3.576	4.058	4.491	4.985	5.297
7	0.727	1.634	2.278	2.885	3.627	4.143	4.617	5.181	5.556
8	0.723	1.631	2.281	2.899	3.665	4.205	4.710	5.322	5.743
9	0.719	1.628	2.283	2.910	3.694	4.252	4.780	5.430	5.885
10	0.717	1.626	2.285	2.919	3.716	4.289	4.835	5.514	5.996
11	0.715	1.625	2.287	2.926	3.735	4.319	4.879	5.582	6.085
12	0.713	1.623	2.288	2.932	3.750	4.344	4.916	5.638	6.158
13	0.711	1.622	2.289	2.937	3.763	4.365	4.947	5.685	6.219
14	0.710	1.621	2.290	2.941	3.774	4.383	4.973	5.725	6.271
15	0.709	1.620	2.291	2.945	3.784	4.398	4.996	5.759	6.316
16	0.708	1.620	2.292	2.948	3.792	4.412	5.015	5.789	6.354
17	0.707	1.619	2.292	2.951	3.799	4.423	5.033	5.815	6.388
18	0.706	1.619	2.293	2.954	3.806	4.434	5.048	5.838	6.418
19	0.705	1.618	2.293	2.956	3.811	4.443	5.061	5.858	6.445
20	0.705	1.618	2.294	2.958	3.816	4.451	5.074	5.877	6.469
21	0.704	1.617	2.294	2.960	3.821	4.459	5.085	5.893	6.491
22	0.704	1.617	2.295	2.961	3.825	4.466	5.095	5.908	6.510
23	0.703	1.616	2.295	2.963	3.829	4.472	5.104	5.922	6.528
24	0.703	1.616	2.295	2.964	3.833	4.478	5.112	5.935	6.544
25	0.702	1.616	2.296	2.966	3.836	4.483	5.120	5.946	6.559
26	0.702	1.616	2.296	2.967	3.839	4.488	5.127	5.957	6.573
27	0.702	1.615	2.296	2.968	3.842	4.492	5.133	5.966	6.586
28	0.701	1.615	2.296	2.969	3.844	4.496	5.139	5.975	6.598
29	0.701	1.615	2.297	2.970	3.847	4.500	5.145	5.984	6.609
30	0.701	1.615	2.297	$\boxed{2.971}$	3.849	4.504	5.150	5.992	6.619
32	0.700	1.614	2.297	2.972	3.853	4.510	5.159	6.006	6.637
34	0.700	1.614	2.297	2.974	3.856	4.516	5.168	6.018	6.654
36	0.700	1.614	2.298	2.975	3.859	4.521	5.175	6.030	6.668
38	0.699	1.614	2.298	2.976	3.862	4.525	5.182	6.039	6.681
40	0.699	1.613	2.298	2.977	3.865	4.529	5.188	6.048	6.692
42	0.699	1.613	2.298	2.978	3.867	4.533	5.193	6.056	6.703
44	0.698	1.613	2.299	2.979	3.869	4.536	5.198	6.064	6.712
46	0.698	1.613	2.299	2.979	3.871	4.539	5.202	6.070	6.721
48	0.698	1.613	2.299	2.980	3.873	4.542	5.206	6.076	6.729
50	0.698	1.613	2.299	2.981	3.874	4.545	5.210	6.082	6.736
55	0.697	1.612	2.299	2.982	3.878	4.550	5.218	6.094	6.752
60	0.697	1.612	2.300	2.983	3.881	4.555	5.225	6.104	6.765
65	0.697	1.612	2.300	2.984	3.883	4.559	5.231	6.113	6.776
70	0.696	1.612	2.300	2.985	3.885	4.562	5.235	6.120	6.786
75	0.696	1.612	2.300	2.986	3.887	4.565	5.240	6.127	6.794
80	0.696	1.611	2.300	2.986	3.889	4.567	5.243	6.132	6.801
90	0.696	1.611	2.301	2.987	3.891	4.572	5.249	6.141	6.813
100	0.695	1.611	2.301	2.988	3.893	4.575	5.254	6.149	6.822
120	0.695	1.611	2.301	2.990	3.896	4.580	5.262	6.160	6.837
140	0.695	1.611	2.301	2.990	3.899	4.584	5.267	6.168	6.847
160	0.695	1.610	2.301	2.991	3.900	4.586	5.271	6.174	6.855
180	0.694	1.610	2.302	2.992	3.902	4.588	5.274	6.178	6.861
200	0.694	1.610	2.302	2.992	3.903	4.590	5.276	6.182	6.865
300	0.694	1.610	2.302	2.993	3.906	4.595	5.284	6.193	6.879
500	0.694	1.610	2.302	2.994	3.908	4.599	5.290	6.201	6.891
∞	0.6931	1.6094	2.3026	2.9957	3.9120	4.6052	5.2983	6.2146	6.9078

Since $24.14 > 2.971$ reject H_o.

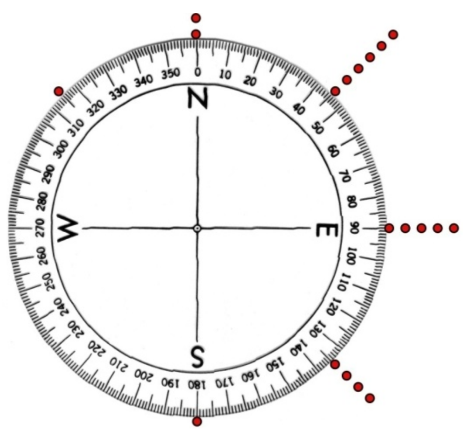
There is a mean direction of 360 (or 0) degrees in the principal azimuths of the Andean villages (Rayleigh $z_{24.14}$, $p < 0.001$).

Hypothesis Testing: Uniformity

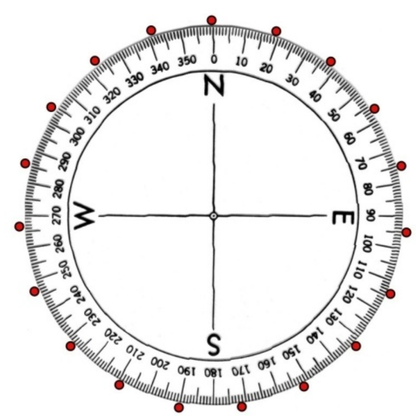
We can also test the hypothesis that the azimuths are not uniformly distributed (occur equally around the compass).

H_o: The distribution of slope aspects is not significantly different than uniform around the compass.

H_a: The distribution of slope aspects is significantly different than uniform around the compass.



Observed Distribution



Uniform Distribution

To test this hypothesis we can use the ratio of the observed slope aspects to the expected (uniform) slope aspects and χ^2.

- If the sample size is reasonable large (> 30) this technique works well.
- If possible, group the data such that no group has less than 4 observations.
- Sometimes grouping in this way is not possible.

First we need to calculate the expected values. Since we are using a uniform distribution, the expected values are:

$$\hat{f}_i = \frac{Sample\ Size}{Categories}$$

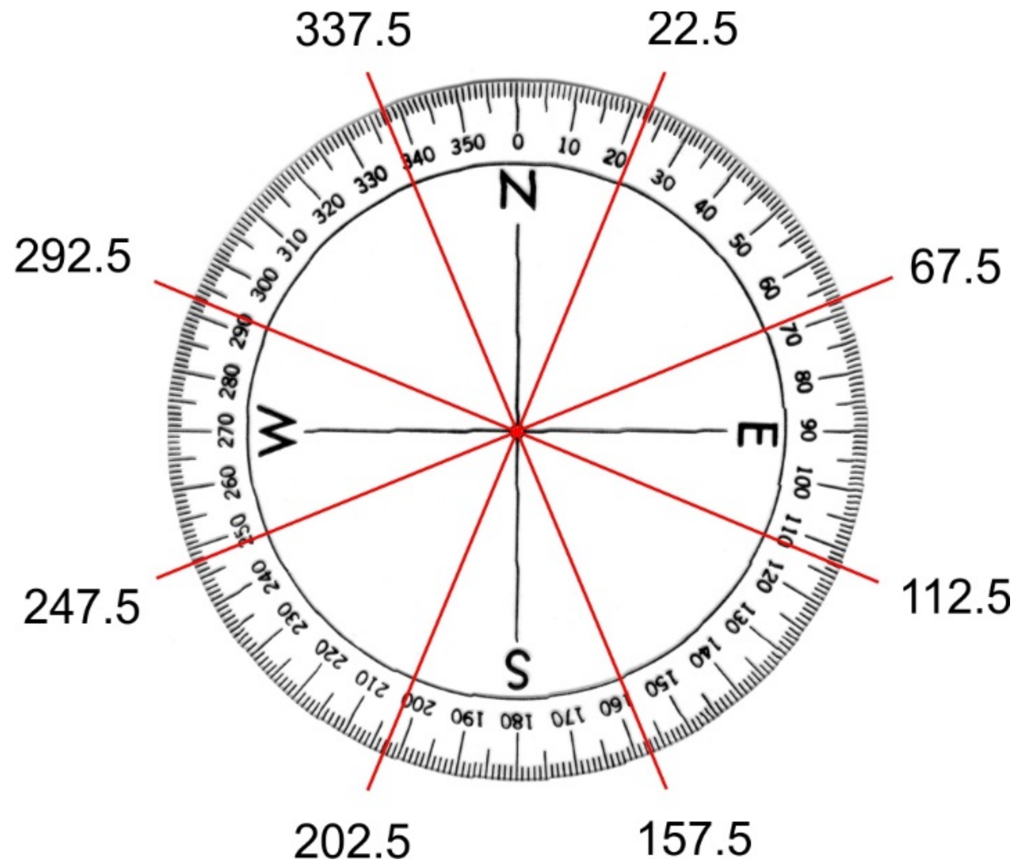
and all expected values for each group will be the same.

We then test the X^2 statistic, which is calculated as:

$$\chi^2 = \sum \frac{(f_i - \hat{f}_i)^2}{\hat{f}_i}$$

The azimuths can be categorized by sector (NE, SW, etc...) based on these compass directions:

N	337.5 – 22.5
NE	22.5 – 67.5
E	67.5 – 112.5
SE	112.5 – 157.5
S	157.5 – 202.5
SW	202.5 – 247.5
W	247.5 – 292.5
NW	292.5 – 337.5



$$\hat{f}_i = \frac{19}{8} = 2.375$$

	f_i	$\hat{f}_i$
East	**5**	**2.375**
North	**2**	**2.375**
Northeast	**6**	**2.375**
Northwest	**1**	**2.375**
South	**1**	**2.375**
Southeast	**4**	**2.375**
Southwest	**0**	**2.375**
West	**0**	**2.375**
n =	**19**	

$$\chi^2 = \frac{(5-2.375)^2}{2.375} + \frac{(2-2.375)^2}{2.375} + \frac{(6-2.375)^2}{2.375} + \frac{(1-2.375)^2}{2.375} + \frac{(1-2.375)^2}{2.375} + \frac{(4-2.375)^2}{2.375} + \frac{(0-2.375)^2}{2.375} + \frac{(0-2.375)^2}{2.375}$$

$$\chi^2 = 2.90 + 0.06 + 5.53 + 0.80 + 0.80 + 1.11 + 2.375 + 2.375$$

$$\chi^2 = 15.95$$

$$df = k - 1 = 8 - 1 = 7$$

$$\chi^2_{critical} = 14.067$$

$\nu \backslash \alpha$	.995	.975	.9	.5	.1	.05	.025	.01	.005	.001	α / ν
1	0.000	0.000	0.016	0.455	2.706	3.841	5.024	6.635	7.879	10.828	1
2	0.010	0.051	0.211	1.386	4.605	5.991	7.378	9.210	10.597	13.816	2
3	0.072	0.216	0.584	2.366	6.251	7.815	9.348	11.345	12.838	16.266	3
4	0.207	0.484	1.064	3.357	7.779	9.488	11.143	13.277	14.860	18.467	4
5	0.412	0.831	1.610	4.351	9.236	11.070	12.832	15.086	16.750	20.515	5
6	0.676	1.237	2.204	5.348	10.645	12.592	14.449	16.812	18.548	22.458	6
7	0.989	1.690	2.833	6.346	12.017	$\boxed{14.067}$	16.013	18.475	20.278	24.322	7
8	1.344	2.180	3.490	7.344	13.362	15.507	17.535	20.090	21.955	26.124	8
9	1.735	2.700	4.168	8.343	14.684	16.919	19.023	21.666	23.589	27.877	9
10	2.156	3.247	4.865	9.342	15.987	18.307	20.483	23.209	25.188	29.588	10
11	2.603	3.816	5.578	10.341	17.275	19.675	21.920	24.725	26.757	31.264	11
12	3.074	4.404	6.304	11.340	18.549	21.026	23.337	26.217	28.300	32.910	12
13	3.565	5.009	7.042	12.340	19.812	22.362	24.736	27.688	29.819	34.528	13
14	4.075	5.629	7.790	13.339	21.064	23.685	26.119	29.141	31.319	36.123	14
15	4.601	6.262	8.547	14.339	22.307	24.996	27.488	30.578	32.801	37.697	15
16	5.142	6.908	9.312	15.338	23.542	26.296	28.845	32.000	34.267	39.252	16
17	5.697	7.564	10.085	16.338	24.769	27.587	30.191	33.409	35.718	40.790	17
18	6.265	8.231	10.865	17.338	25.989	28.869	31.526	34.805	37.156	42.312	18
19	6.844	8.907	11.651	18.338	27.204	30.144	32.852	36.191	38.582	43.820	19
20	7.434	9.591	12.443	19.337	28.412	31.410	34.170	37.566	39.997	45.315	20
21	8.034	10.283	13.240	20.337	29.615	32.670	35.479	38.932	41.401	46.797	21
22	8.643	10.982	14.042	21.337	30.813	33.924	36.781	40.289	42.796	48.268	22
23	9.260	11.688	14.848	22.337	32.007	35.172	38.076	41.638	44.181	49.728	23
24	9.886	12.401	15.659	23.337	33.196	36.415	39.364	42.980	45.558	51.179	24
25	10.520	13.120	16.473	24.337	34.382	37.652	40.646	44.314	46.928	52.620	25
26	11.160	13.844	17.292	25.336	35.563	38.885	41.923	45.642	48.290	54.052	26
27	11.808	14.573	18.114	26.336	36.741	40.113	43.194	46.963	49.645	55.476	27
28	12.461	15.308	18.939	27.336	37.916	41.337	44.461	48.278	50.993	56.892	28
29	13.121	16.047	19.768	28.336	39.088	42.557	45.722	49.588	52.336	58.301	29
30	13.787	16.791	20.599	29.336	40.256	43.773	46.979	50.892	53.672	59.703	30
31	14.458	17.539	21.434	30.336	41.422	44.985	48.232	52.191	55.003	61.098	31
32	15.134	18.291	22.271	31.336	42.585	46.194	49.480	53.486	56.329	62.487	32
33	15.815	19.047	23.110	32.336	43.745	47.400	50.725	54.776	57.649	63.870	33
34	16.501	19.806	23.952	33.336	44.903	48.602	51.966	56.061	58.964	65.247	34
35	17.192	20.569	24.797	34.336	46.059	49.802	53.203	57.342	60.275	66.619	35
36	17.887	21.336	25.643	35.336	47.212	50.998	54.437	58.619	61.582	67.985	36
37	18.586	22.106	26.492	36.335	48.363	52.192	55.668	59.892	62.884	69.346	37
38	19.289	22.878	27.343	37.335	49.513	53.384	56.896	61.162	64.182	70.703	38
39	19.996	23.654	28.196	38.335	50.660	54.572	58.120	62.428	65.476	72.055	39
40	20.707	24.433	29.051	39.335	51.805	55.758	59.342	63.691	66.766	73.402	40
41	21.421	25.215	29.907	40.335	52.949	56.942	60.561	64.950	68.053	74.745	41
42	22.138	25.999	30.765	41.335	54.090	58.124	61.777	66.206	69.336	76.084	42
43	22.859	26.785	31.625	42.335	55.230	59.304	62.990	67.459	70.616	77.419	43
44	23.584	27.575	32.487	43.335	56.369	60.481	64.202	68.710	71.893	78.750	44
45	24.311	28.366	33.350	44.335	57.505	61.656	65.410	69.957	73.166	80.077	45
46	25.042	29.160	34.215	45.335	58.641	62.830	66.617	71.201	74.437	81.400	46
47	25.775	29.956	35.081	46.335	59.774	64.001	67.821	72.443	75.704	82.720	47
48	26.511	30.755	35.949	47.335	60.907	65.171	69.023	73.683	76.969	84.037	48
49	27.249	31.555	36.818	48.335	62.038	66.339	70.222	74.919	78.231	85.351	49
50	27.991	32.357	37.689	49.335	63.167	67.505	71.420	76.154	79.490	86.661	50

Since 15.95 > 14.067, reject H_o.

The distribution of slope aspects is significantly different than a uniform distribution around the compass ($X^2_{15.95}$, $0.05 > p > 0.025$).

Two-Sample Hypothesis Testing

We can test the hypothesis that two sets of azimuths are not significantly different in a procedure similar to the Mann-Whitney U test called ***Watson's U^2 test***.

H_o: The two groups of principal azimuths are not significantly different.

H_a: The two groups of principal azimuths are significantly different.

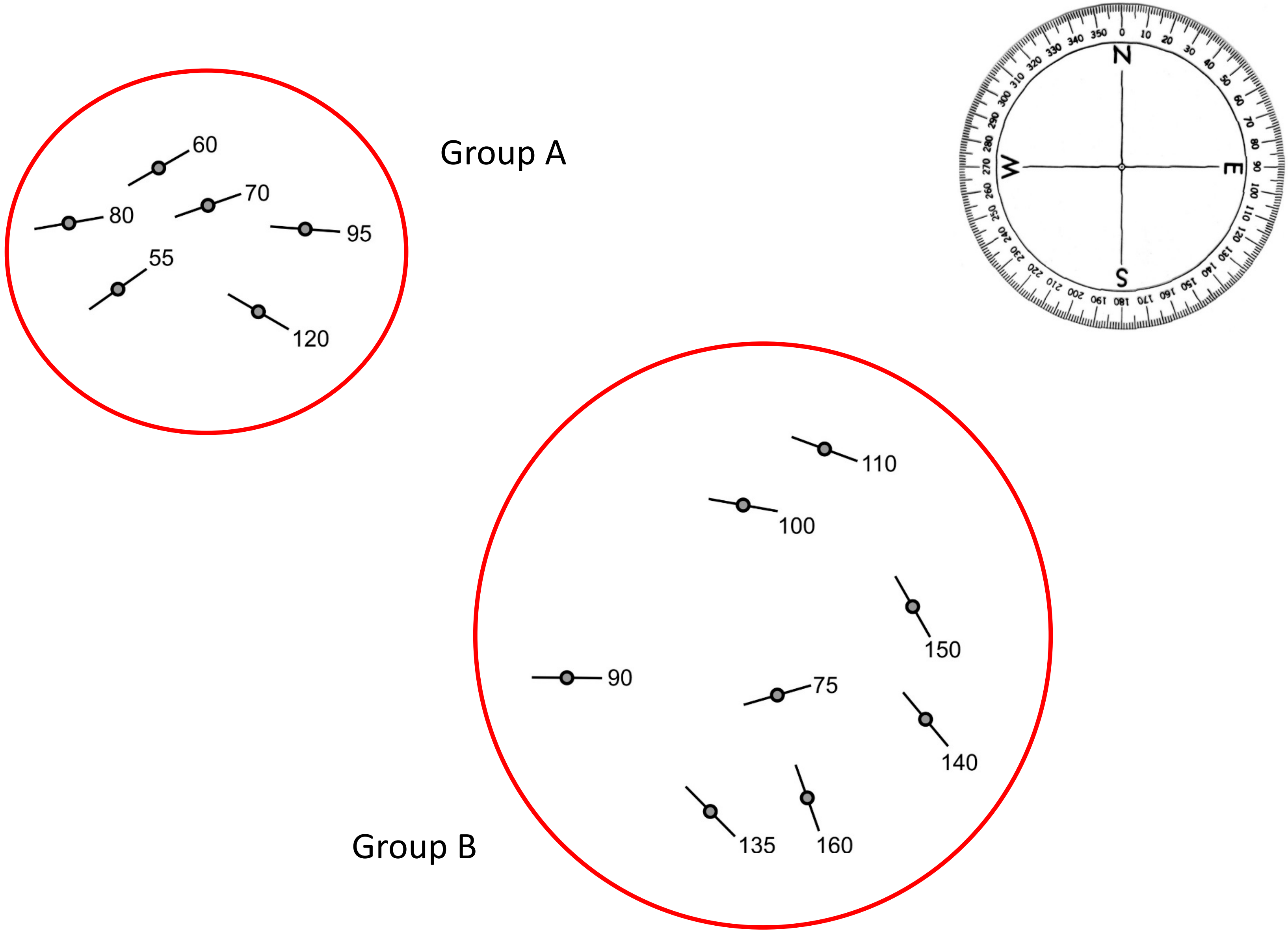
Watson's U2 test equation:

$$U^2 = \frac{n_1 n_2}{N^2}\left[\sum d_k^2 - \frac{\left(\sum d_k\right)^2}{N}\right]$$

where

$$N = n_1 + n_2$$

- First the data are ranked across groups
- n_1 and n_2 are the group sample sizes.
- The expected frequency (i/n_1 and j/n_2) are then calculated.


Group A
60
70
80
95
55
120
Group B
110
100
150
90
75
140
135
160
N
E
S
W

i	Group A $a1_i$ (deg)	i/n_1	j	Group B $a2_i$ (deg)	j/n_2	$d_k = i/n_1 - j/n_2$	d^2_k
1	55	0.1667			0.0000	0.1667	0.0278
2	60	0.3333			0.0000	0.3333	0.1111
3	70	0.5000			0.0000	0.5000	0.2500
		0.5000	1	75	0.1250	0.3750	0.1406
4	80	0.6667			0.1250	0.5417	0.2934
		0.6667	2	90	0.2500	0.4167	0.1736
5	95	0.8333			0.2500	0.5833	0.3402
		0.8333	3	100	0.3750	0.4583	0.2100
		0.8333	4	110	0.5000	0.3333	0.1111
6	120	1.0000			0.5000	0.5000	0.2500
		1.0000	5	135	0.6250	0.3750	0.1406
		1.0000	6	140	0.7500	0.2500	0.0625
		1.0000	7	150	0.8750	0.1250	0.0156
		1.0000	8	160	1.0000	0.0000	0.0000

$n_1 = 6 \qquad n_2 = 8 \qquad \sum d_k = 4.9583 \qquad \sum d_k^2 = 2.1265$

$$U^2 = \frac{(6)(8)}{14^2}\left[2.1265 - \frac{(4.9583)^2}{14}\right] \qquad U^2 = \frac{n_1 n_2}{N^2}\left[\sum d_k^2 - \frac{\left(\sum d_k\right)^2}{N}\right]$$

$$U^2 = 0.0907$$

$$U^2_{Critical} = 0.1964$$

Critical Values for the Watson U^2 Test (cont.)
Taken from Zar, 1981 Table B.35

n_1	n_2	α = 0.50	0.20	0.10	0.05	0.02	0.01	0.005	0.002	0.001
5	28	0.0746	0.1188	0.1512	0.1802	0.2170	0.2417	0.2694	0.2937	0.3136
5	29	0.0743	0.1189	0.1510	0.1802	0.2171	0.2443	0.2666	0.2970	0.3153
5	30	0.0743	0.1189	0.1512	0.1802	0.2160	0.2419	0.2678	0.2979	0.3181
6	6	0.0880	0.1319	0.1713	0.2060	0.2639	------	------	------	------
6	7	0.0806	0.1209	0.1538	0.1941	0.2821	0.2821	------	------	------
6	8	0.0833	0.1265	0.1607	0.1964	0.2455	0.2976	0.2976	------	------
6	9	0.0815	0.1259	0.1556	0.1928	0.2321	0.2617	0.3111	------	------
6	10	0.0771	0.1260	0.1563	0.1896	0.2313	0.2479	0.3229	0.3229	------
5	11	0.0784	0.1212	0.1569	0.1872	0.2246	0.2620	0.2888	0.3333	------
6	12	0.0802	0.1242	0.1551	0.1829	0.2261	0.2593	0.2747	0.3426	0.3426
6	13	0.0769	0.1215	0.1538	0.1849	0.2213	0.2497	0.2780	0.3509	0.3509
6	14	0.0768	0.1220	0.1536	0.1839	0.2250	0.2506	0.2821	0.3196	0.3583
6	15	0.0762	0.1217	0.1524	0.1852	0.2201	0.2487	0.2730	0.3058	0.3651
5	15	0.0758	0.1212	0.1534	0.1823	0.2235	0.2500	0.2789	0.3073	0.3357
6	17	0.0750	0.1211	0.1526	0.1833	0.2199	0.2472	0.2745	0.3129	0.3427
5	18	0.0760	0.1211	0.1535	0.1840	0.2199	0.2461	0.2739	0.2998	0.3295
6	19	0.0751	0.1200	0.1523	0.1832	0.2204	0.2498	0.2744	0.3060	0.3298
6	20	0.0747	0.1196	0.1526	0.1824	0.2196	0.2490	0.2734	0.3077	0.3333
6	21	0.0758	0.1205	0.1523	0.1834	0.2205	0.2475	0.2734	0.3057	0.3369
6	22	0.0749	0.1204	0.1518	0.1824	0.2202	0.2473	0.2752	0.3036	0.3260
6	23	0.0745	0.1194	0.1514	0.1824	0.2194	0.2469	0.2729	0.3073	0.3273
6	24	0.0743	0.1194	0.1519	0.1826	0.2206	0.2484	0.2715	0.3056	0.3289
5	25	0.0744	0.1191	0.1514	0.1819	0.2202	0.2473	0.2731	0.3015	0.3277
6	26	0.0739	0.1188	0.1510	0.1815	0.2198	0.2464	0.2710	0.3047	0.3265
6	27	0.0741	0.1193	0.1515	0.1822	0.2200	0.2469	0.2731	0.3053	0.3281
6	28	0.0737	0.1190	0.1507	0.1821	0.2201	0.2467	0.2731	0.3039	0.3270
5	29	0.0736	0.1189	0.1511	0.1816	0.2200	0.2473	0.2719	0.3038	0.3258
6	30	0.0736	0.1193	0.1509	0.1823	0.2194	0.2471	0.2725	0.3045	0.3262
7	7	0.0791	0.1345	0.1578	0.1986	0.2511	0.3036	0.3036	------	------
7	8	0.0794	0.1198	0.1556	0.1817	0.2246	0.2722	0.3222	------	------
7	9	0.0786	0.1223	0.1560	0.1818	0.2215	0.2552	0.2909	0.3385	------
7	10	0.0773	0.1227	0.1546	0.1866	0.2269	0.2622	0.2773	0.3529	0.3529
7	11	0.0771	0.1219	0.1551	0.1839	0.2214	0.2532	0.2806	0.3225	0.3657
7	12	0.0764	0.1216	0.1541	0.1855	0.2256	0.2519	0.2757	0.3083	0.3772
7	13	0.0765	0.1216	0.1545	0.1842	0.2227	0.2523	0.2776	0.3150	0.3479
7	14	0.0761	0.1228	0.1568	0.1840	0.2248	0.2530	0.2744	0.3210	0.3337
7	15	0.0754	0.1213	0.1525	0.1845	0.2235	0.2503	0.2780	0.3118	0.3378
7	16	0.0753	0.1203	0.1530	0.1848	0.2236	0.2508	0.2772	0.3113	0.3432
7	17	0.0749	0.1204	0.1526	0.1827	0.2227	0.2500	0.2752	0.3109	0.3340
7	18	0.0749	0.1200	0.1524	0.1841	0.2235	0.2502	0.2768	0.3117	0.3346
7	20	0.0743	0.1198	0.1526	0.1832	0.2219	0.2499	0.2780	0.3081	0.3330
7	21	0.0751	0.1203	0.1534	0.1840	0.2224	0.2496	0.2782	0.3123	0.3336
7	22	0.0743	0.1196	0.1518	0.1832	0.2221	0.2512	0.2763	0.3090	0.3341
7	23	0.0739	0.1194	0.1522	0.1832	0.2226	0.2499	0.2780	0.3103	0.3327
8	8	0.0781	0.1250	0.1563	0.1836	0.2256	0.2500	0.2959	0.3438	------
8	9	0.0784	0.1225	0.1552	0.1863	0.2255	0.2582	0.2827	0.3627	0.3627
3	10	0.0775	0.1220	0.1546	0.1852	0.2220	0.2491	0.2796	0.3359	0.3796
8	11	0.0766	0.1220	0.1543	0.1842	0.2249	0.2524	0.2799	0.3194	0.3529
8	12	0.0766	0.1208	0.1557	0.1854	0.2229	0.2521	0.2807	0.3167	0.3396
8	13	0.0754	0.1212	0.1532	0.1853	0.2237	0.2531	0.2778	0.3135	0.3446

Since $0.0907 < 0.1964$ accept H_o.

The two groups of principal azimuths of Andean villages are not significantly different ($U^2_{0.0907}$, $0.50 > p > 0.20$).

Serial Randomness of Nominal Data on a Circle

Used to test the hypothesis that the occurrence of categorical data around a circle is random.

H_o: The distribution of data is not significantly different than random.

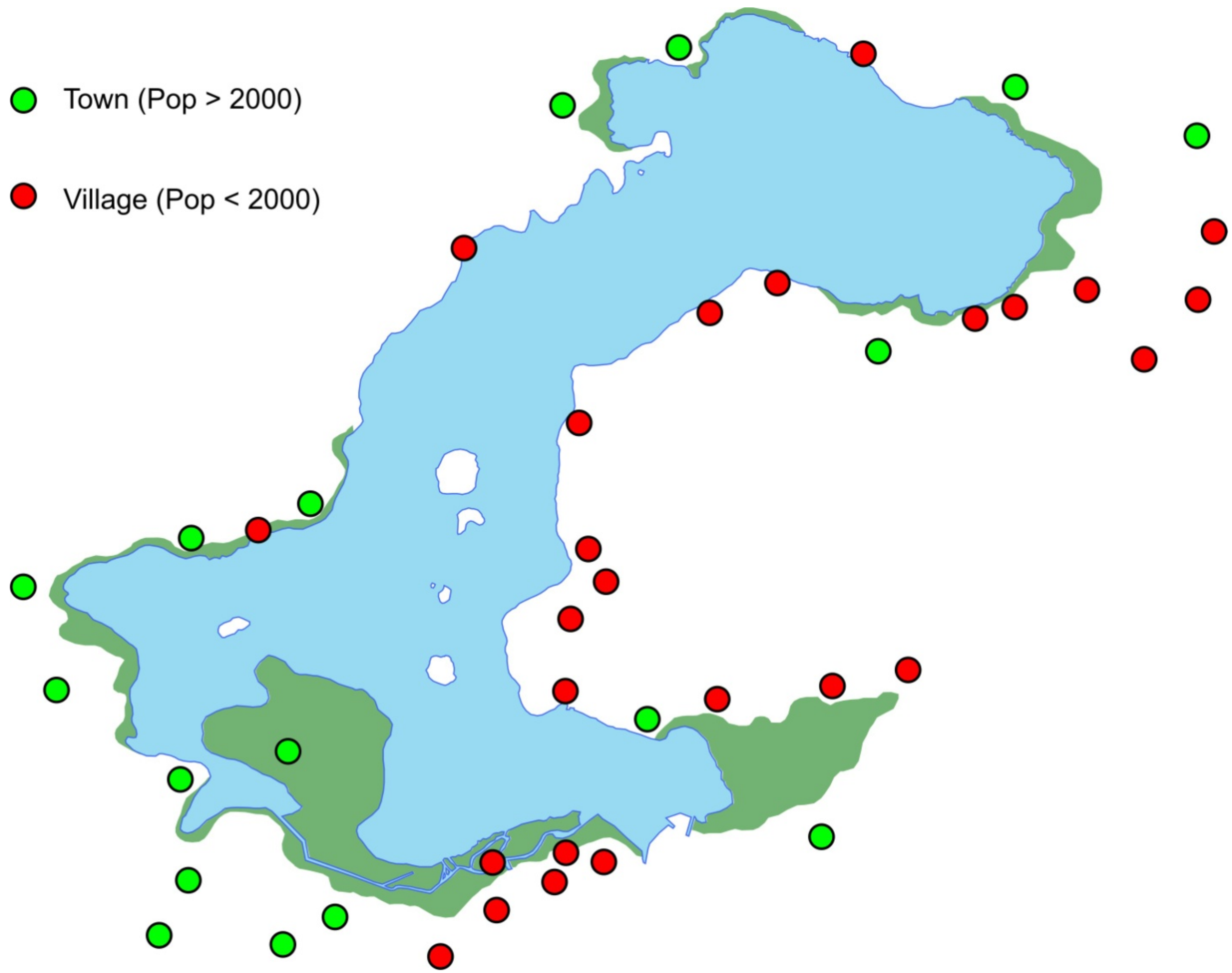
H_a: The distribution of data is significantly different than random.

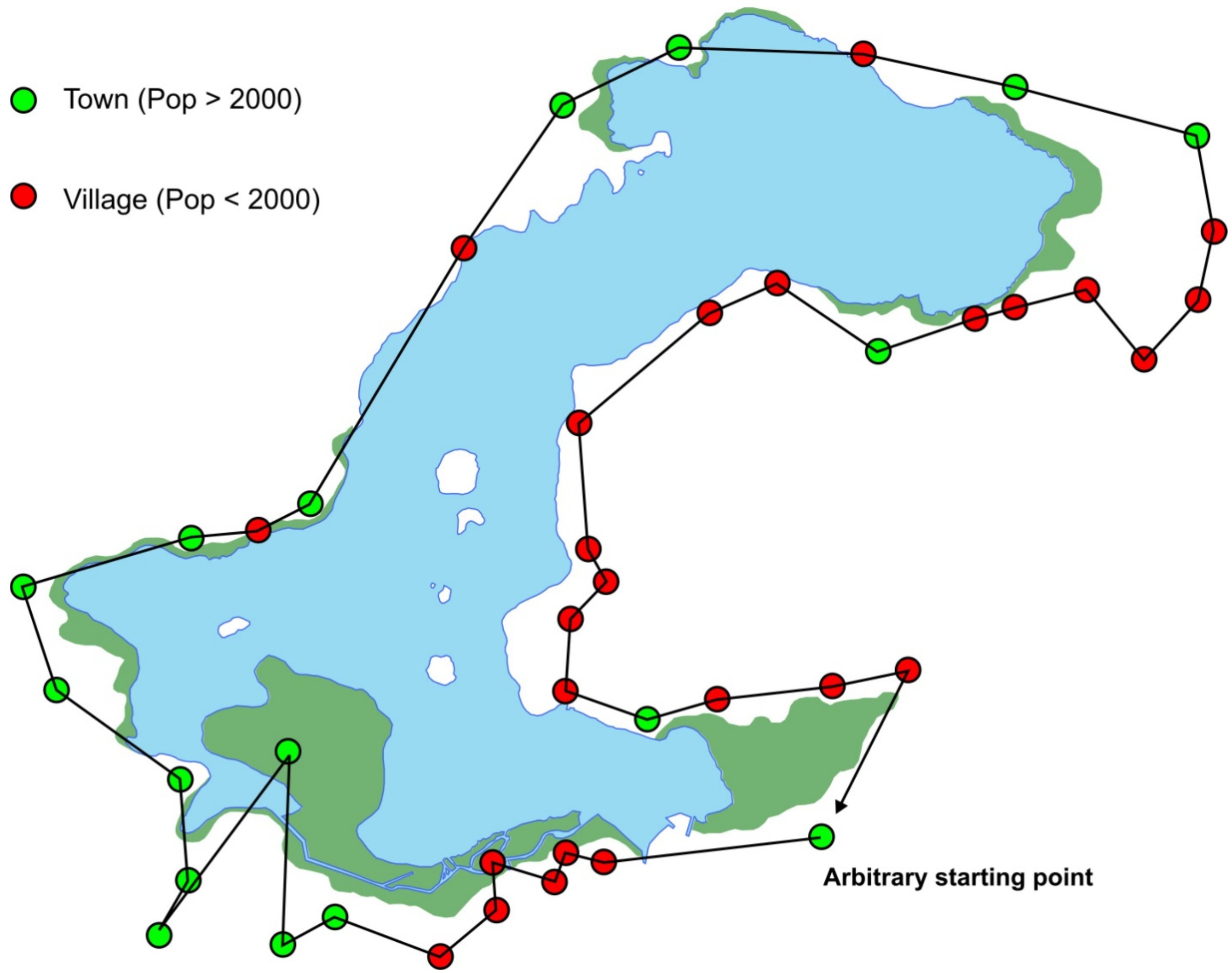
- Can only be used for 2 categories of data.
- Uses the Runs test statistic for sample sizes of 20 or less.
- Use the normal approximation for larger sample sizes.

Normal Approximation Equation

$$Z = \frac{u' + 1 - \frac{2n_1n_2 + N}{N}}{\sqrt{\frac{2n_1n_2(2n_1n_2 - N}{N^2(N-1)}}}$$

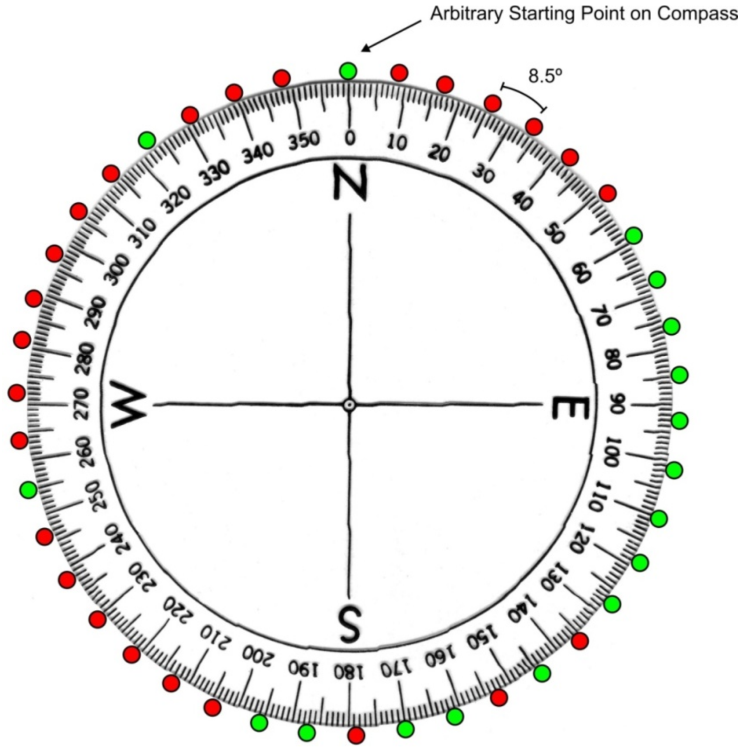
where *Z* is the normal approximation, *u'* is the total number of runs in the data set, n_1 and n_2 are the group sample sizes, and N is the total sample size ($n_1 + n_2$).


Town (Pop > 2000)
Village (Pop < 2000)


Town (Pop > 2000)
Village (Pop < 2000)
Arbitrary starting point

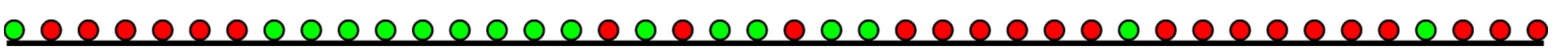
n = 42

360º/42 = 8.5º



Towns (n_1) = 17
Villages (n_2) = 25
Runs = 14

This could also simply be represented as dots along a line, similar to the Runs test for sequential data.



$n_1 = 17 \qquad n_2 = 25$

$u' = 14$

$$Z = \frac{14 + 1 - \dfrac{2(17)(25) + 42}{42}}{\sqrt{\dfrac{\left(2(17)(25)\right)\left[\left(2(17)(25)\right) - 42\right]}{42^2(42-1)}}} = \frac{15 - \dfrac{892}{42}}{\sqrt{\dfrac{(850)(808)}{72324}}}$$

$$Z = \frac{-6.238}{3.082} = -2.02$$

$p = 0.0217$

Since the probability is less than 0.05, we reject H_o.

Proportions of the Normal Curve: One-Tailed (*Z* Table)

Taken from Zar, 1984 Table B.2

Z	0	1	2	3	4	5	6	7	8	9	Z
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641	0.0
0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247	0.1
0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859	0.2
0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483	0.3
0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121	0.4
0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776	0.5
0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451	0.6
0.7	0.2420	0.2389	0.2358	0.2327	0.2297	0.2266	0.2236	0.2207	0.2177	0.2148	0.7
0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867	0.8
0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611	0.9
1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379	1.0
1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170	1.1
1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985	1.2
1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823	1.3
1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681	1.4
1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559	1.5
1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455	1.6
1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367	1.7
1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294	1.8
1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233	1.9
2.0	0.0228	0.0222	$\boxed{0.0217}$	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183	2.0
2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143	2.1
2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110	2.2
2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084	2.3
2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064	2.4
2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048	2.5
2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036	2.6
2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026	2.7
2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019	2.8
2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014	2.9
3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010	3.0
3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007	3.1
3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005	3.2
3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003	3.3
3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002	3.4
3.5	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	3.5
3.6	0.0002	0.0002	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	3.6
3.7	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	3.7
3.8	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	3.8

If a one-tailed test is used, u' values less than the lower of the u' critical values from the Runs table of circular data signify that the non-randomness is due to clustering.

If the u' values are greater than the higher of the u' critical values from the Runs table of circular data it signifies that the non-randomness is due to uniformity.

If the normal approximation is used, negative Z values signify that the non-randomness is due to clustering, while positive values signify uniformity.

Therefore in our example:

The distribution of towns and villages is significantly different than random *due to clustering* ($Z_{-2.02}$, $p = 0.0217$).