

Universidade Federal de Alagoas
Instituto de Computação
Laboratório de Computação Científica e Análise Numérica

Research report

Student: Danilo Fernandes Costa

Professor: Alejandro Frery

September
2019

1 Introduction

In this report, some results of the analysis of samples referring to plantation regions observed over time are shown. The first observation was made on 16 May 2016, which was followed by four others with a time interval of 24 days between adjacent observations. Those regions consist of three soybeans crops, three wheat, two oats and two canola and are shown in the figures 2(a) to 2(e). Those samples were obtained using the classification of the regions given in the figure 1.

The results consist of the graphic adjusting of the beta distribution to the histograms of the distances of the samples to the trihedral and random volume scatterers. In addition, the Kolmogorov-Smirnov good-of-fit test was performed to validate the hypothesis of the adjust. The choice of these scatterers is because they were more sensitive to vegetation variation in the analyzed regions, which can be verified in the figures 3(a) and 3(b) that contain respectively the pixel proportions in the sample that is more similar to trihedral and the random volume as a function of time.



Figura 1: Classification of the regions on the PolSAR image

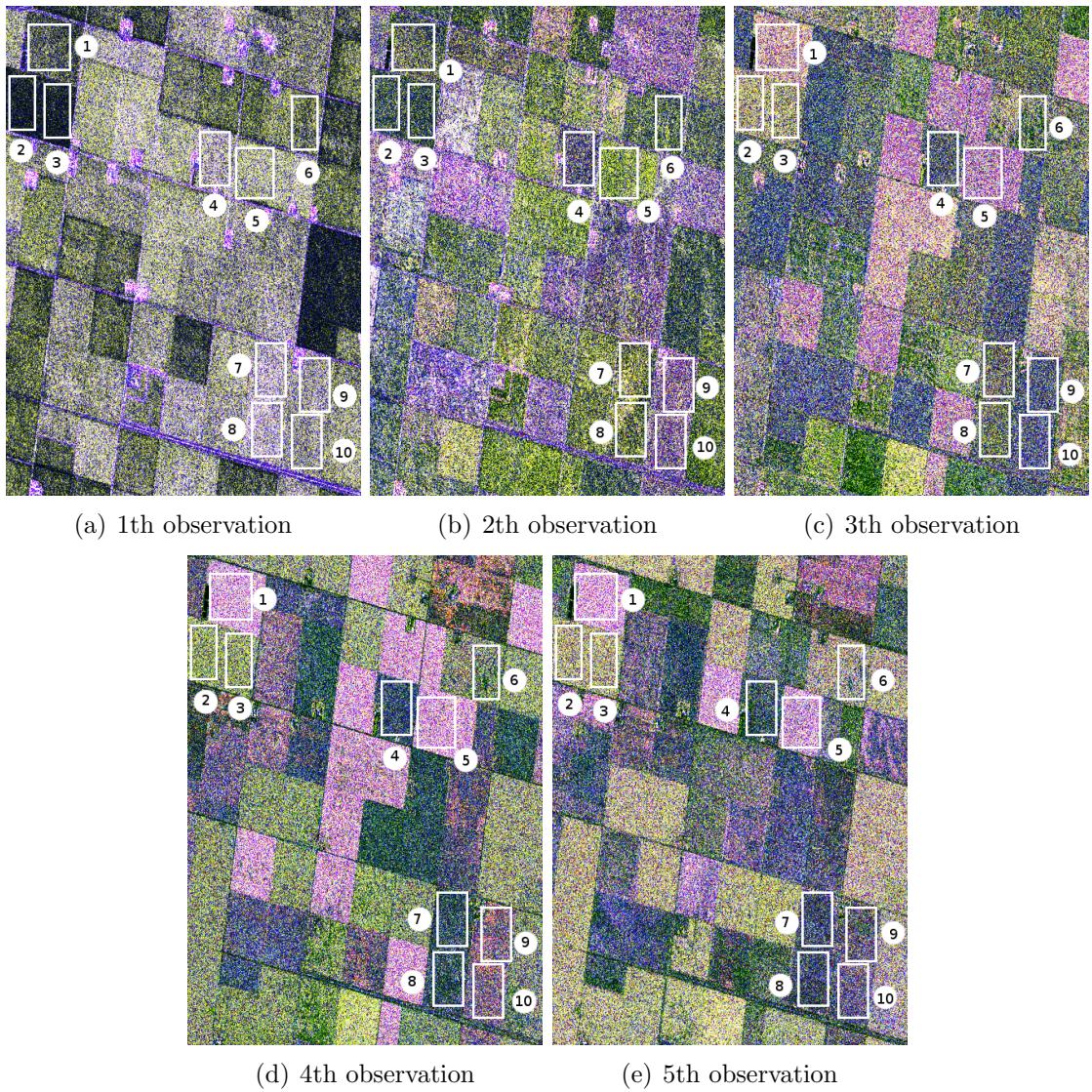


Figura 2: Samples analysed over time in where 1 to 10 corresponding respectively to Canola 43, Soybeans 231, Soybeans 232, Wheat 225, Canola 224, Soybeans 101, Oats 102, Oats 103, Wheat 105 and Wheat 104

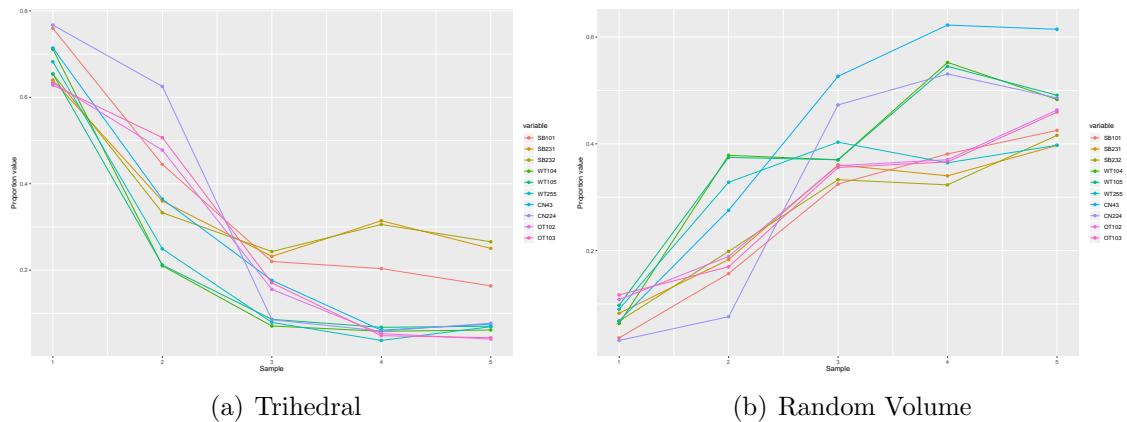


Figura 3: Pixel proportions on the analysed regions more similar to trihedral and random volume as function of time, where SB, WT, CN and OT indicate respectively Soybeans, Wheat, Canola and Oats

2 Adjust the histograms to the Beta Distribution

The figures 4 to 23 contain the histograms of the Geodesic Distances between the scatterer (random or trihedral volume) and the pixels of the sample most similar to this scatterer. The number of those pixels are in the table 1 which TR and RV indicates respectively trihedral and random volume. In addition, the histograms are graphically adjusted by the Beta Distribution, whose parameters were estimated from the distances analyzed.

Those adjusts was verified by Komolgorov-Smirnov good-of-fit test whose obtained p -values are in the table 2. In this table, the biggest and least p -values are in bold.

Tabela 1: Number of pixels more similar to trihedral and random volume

	First observation		Second observation		Third observation		Fourth observation		Fifth observation	
	TR	RV	TR	RV	TR	RV	TR	RV	TR	RV
SB 101	1481	71	867	306	429	633	397	743	319	829
SB 231	1285	148	656	381	449	706	615	649	508	775
SB 232	1275	132	649	387	474	649	596	630	518	811
WT 104	1618	144	478	861	161	842	133	1257	140	1099
WT 105	1488	222	482	852	196	842	154	1240	160	1117
WT 255	1552	205	567	746	180	917	85	829	157	904
CN 43	1964	155	1002	625	485	1195	168	1413	207	1395
CN 224	2106	87	1716	209	234	1298	162	1457	212	1334
OT 102	1441	246	1087	430	354	817	121	842	92	1054
OT 103	1429	266	1153	386	388	806	110	833	99	1045

Tabela 2: p -values of the Kolmogorov-Smirnov goodness-of-fit test of the distances to trihedral an random volume

	First observation		Second observation		Third observation		Fourth observation		Fifth observation	
	TR	RV	TR	RV	TR	RV	TR	RV	TR	RV
SB 101	0.065	0.517	0.947	0.758	0.059	0.495	0.452	0.109	0.401	0.144
SB 231	0.775	0.242	0.573	0.166	0.314	0.275	0.239	0.114	0.416	0.070
SB 232	0.244	0.340	0.968	0.328	0.713	0.070	0.422	0.357	0.163	0.630
WT 104	0.178	0.715	0.421	0.094	0.514	0.779	0.062	0.369	0.602	0.919
WT 105	0.231	0.090	0.069	0.139	0.557	0.613	0.108	0.195	0.192	0.252
WT 255	0.235	0.513	0.270	0.375	0.628	0.279	0.653	0.069	0.437	0.993
CN 43	0.238	0.406	0.217	0.202	0.930	0.318	0.623	0.732	0.262	0.747
CN 224	0.184	0.116	0.128	0.333	0.298	0.714	0.813	0.409	0.305	0.391
OT 102	0.289	0.191	0.243	0.532	0.384	0.212	0.710	0.370	0.928	0.396
OT 103	0.096	0.139	0.139	0.186	0.265	0.079	0.936	0.079	0.989	0.489

3 Parameters evaluation

When observing the regions referring to Soybeans 231 and 232 along its samples in the figure 2, which are respectively indexed by 2 and 3, it can be assumed that there was a gradual increase in the degree of vegetation of these regions.

In order to relate this to the variation of the information contained in the distances of its data to the trihedral, the graphs of the figures 24(a) and 24(b) were produced. Each graph contains, for each observation, a boxplot of the means of the distances between trihedral and the subregions generated by dividing a region into 45 subregions with dimensions 7x6. In addition, all boxplots were connected by the mean of their means.

In order to adjust the mean as a function of time for both regions, the following function has been proposed:

$$f(t) = -\frac{a}{bt + c} + d \quad (1)$$

which $a = 4.741$, $b = 2.415$, $c = 67.565$, $d = 0.276$ and t is the number of days since the first observation ($t = 0$). For these adjustments, the lack of fit tests were performed whose ANOVA tables are in tables 3 and 4. From these it can be concluded that the proposed model is acceptable at the significance level of 0.1.

Tabela 3: ANOVA for lack of fit on Soybeans 231

	Degree of freedom	Sum of squared errors	Mean squared error	Fisher statistics	p-value
Residual	223	0.1912	0.0008		
Lack od fit	3	0.0043	0.0014	1.7087	0.1661
Pure error	220	0.1859	0.0008		

Tabela 4: ANOVA for lack of fit on Soybeans 232

	Degree of freedom	Sum of squared errors	Mean squared error	Fisher statistics	p-value
Residual	223	0.1836	0.0008		
Lack od fit	3	0.0020	0.0007	0.7973	0.4965
Pure error	220	0.1819	0.0008		

4 Classifier for vegetation regions

In this section a classifier is proposed based on results of the analysis of a subregion of Soybeans 231 with dimensions 15x15. For that, a Beta distribution was fitted for its geodesic distances to left and right helix on the first and last

observation. The histograms of these distances are graphically adjusted in figures 25(a) and 25(a). In addition, the Kolmogorov-Smirnov test for goodness-of-fit was performed and returned p -values are in the table 5.

Tabela 5: p -values of the Kolmogorov-Smirnov goodness-of-fit test of the distances to left and right helix

	Left helix	Right helix
First sample	0.406	0.172
Last sample	0.940	0.817

Since the distances to the left helix and right helix are independent, because these are orthogonal, it is possible to define $d = 0.912$ as cutoff point for the densities Beta(21, 1.6) and Beta(10, 1.85) and obtain the joint probabilities shown in table 6, where D_{lh} and D_{rh} are respectively the distance to left helix and right helix.

Tabela 6: Joint probabilities for distance to left and right helix

	$D_{rh} \geq 0.912$	$D_{rh} > 0.912$	$D_{rh} \leq 0.912$	$D_{rh} > 0.912$
	$D_{rh} \leq 0.912$	$D_{rh} > 0.912$	$D_{rh} \leq 0.912$	$D_{rh} > 0.912$
First sample	0.09	0.21	0.21	0.49
Last sample	0.49	0.21	0.21	0.09

Assume the pixels in the first and last observation as poor and rich in vegetation, respectively. Then, it is possible classify pixels with $d_{lh} > 0.912$ and $d_{rh} > 0.912$ as poor in vegetation and those with $d_{lh} \leq 0.912$ and $d_{rh} \leq 0.912$ as rich in vegetation. In this, there is 0.09 of probability that classify a poor pixel in vegetation as rich and vice-versa. However, this approach allows classify only 56% of the both populations.

For classify the other 44% of the population, the components of the data in the direction of these elementary scatterers are removed as follows:

$$v'_{data} = v_{data} - \frac{\langle v_{data}, v_{lh} \rangle}{\|v_{lh}\|^2} v_{lh} - \frac{\langle v_{data}, v_{rh} \rangle}{\|v_{rh}\|^2} v_{rh}, \quad (2)$$

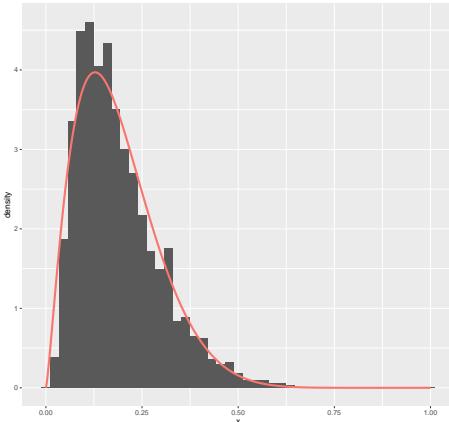
where v_{data} , v_{lh} and v_{rh} are respectively data, left helix and right helix in Kennaugh form. Then, the follow distance is calculate:

$$D_d = \frac{1}{\pi} \cos^{-1} \left(\frac{\langle v'_{data}, v_d \rangle}{\|v'_{data}\| \|v_d\|} \right) = \frac{1}{2} GD(v'_{data}, v_d), \quad (3)$$

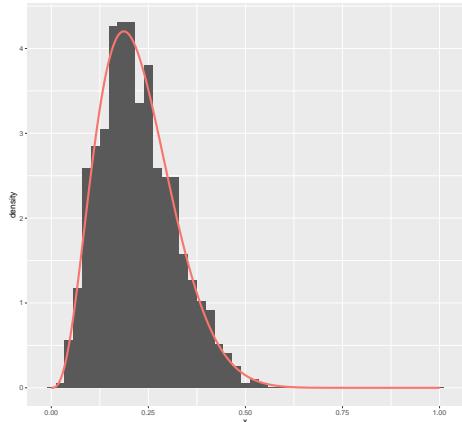
where v_d is dihedral in Kennaugh form. This distance was used because $GD(v'_{data}, v_d) \in [0, 2]$ for analysed subregion. The histograms of this distance between dihedral and analysed samples are shown in the figure 26 and the p -values from Kolmogorov-Smirnov goodness-of-fit test for first and last observation are respectively 0.127 and 0.105.

5 Conclusions

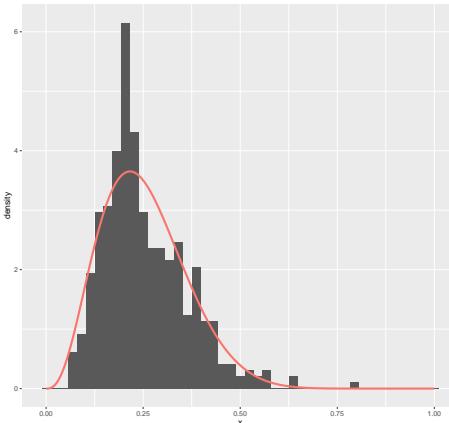
It can be concluded from the p -values table that the Beta distribution adjusts the distances of PolSAR data of the analyzed cultures to trihedral and random volume at the significance level of 0.05.



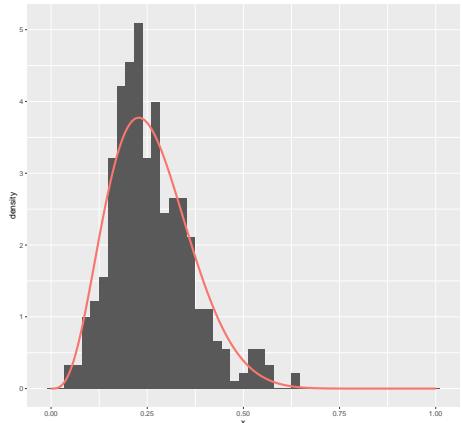
(a) 1th observation



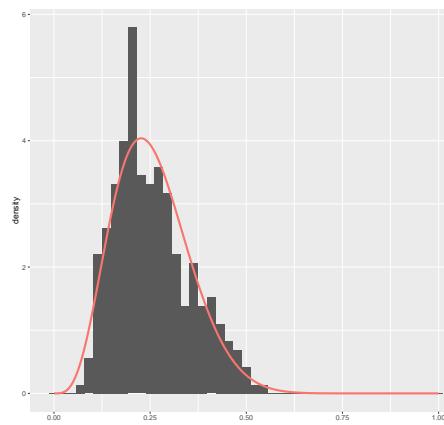
(b) 2th observation



(c) 3th observation



(d) 4th observation



(e) 5th observation

Figura 4: Histograms of the Geodesic Distances between trihedral and the pixels of the sample extracted from Soybeans 101 most similar to trihedral

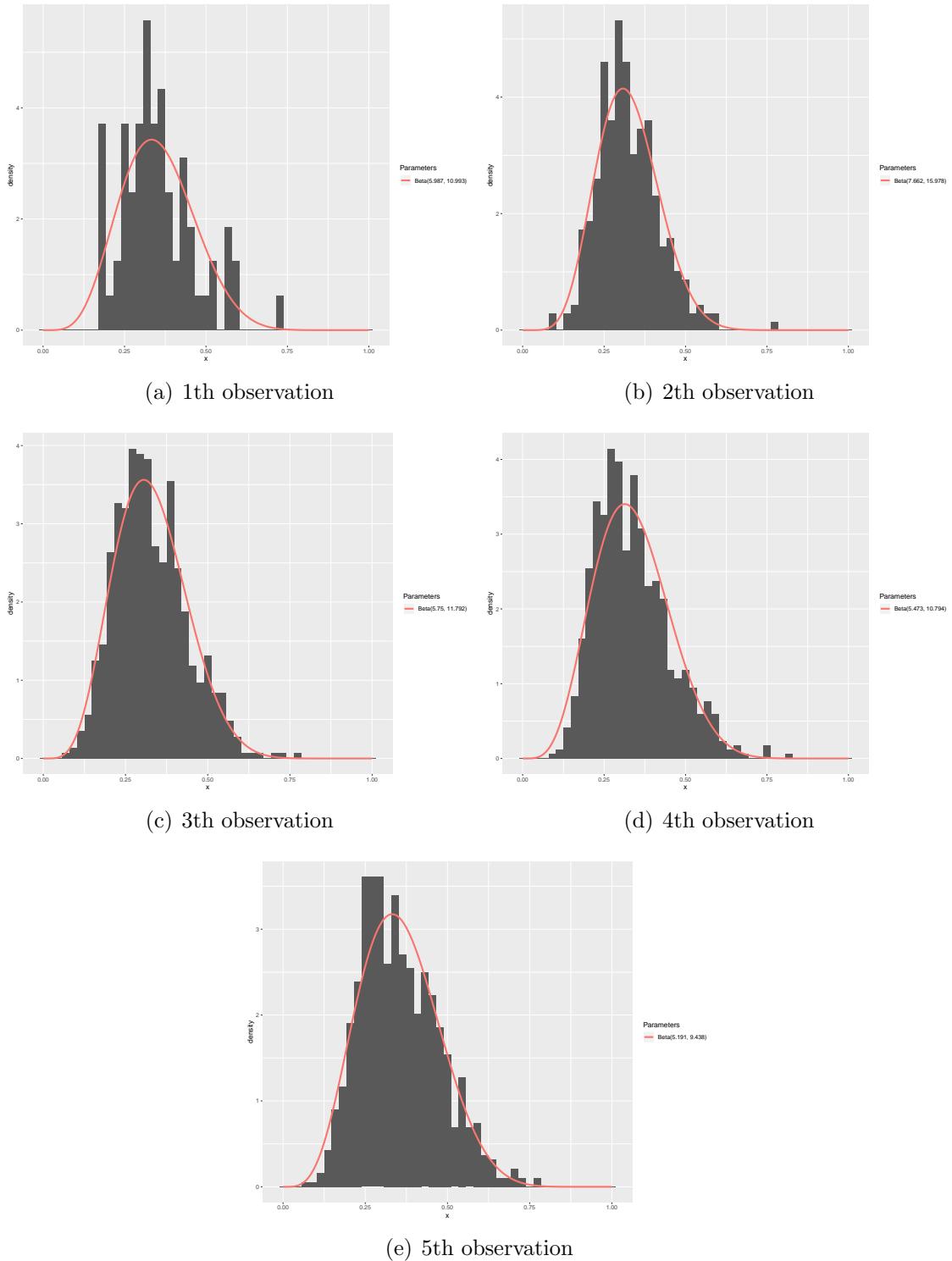


Figura 5: Histograms of the Geodesic Distances between random volume and the pixels of the sample extracted from Soybeans 101 most similar to random volume

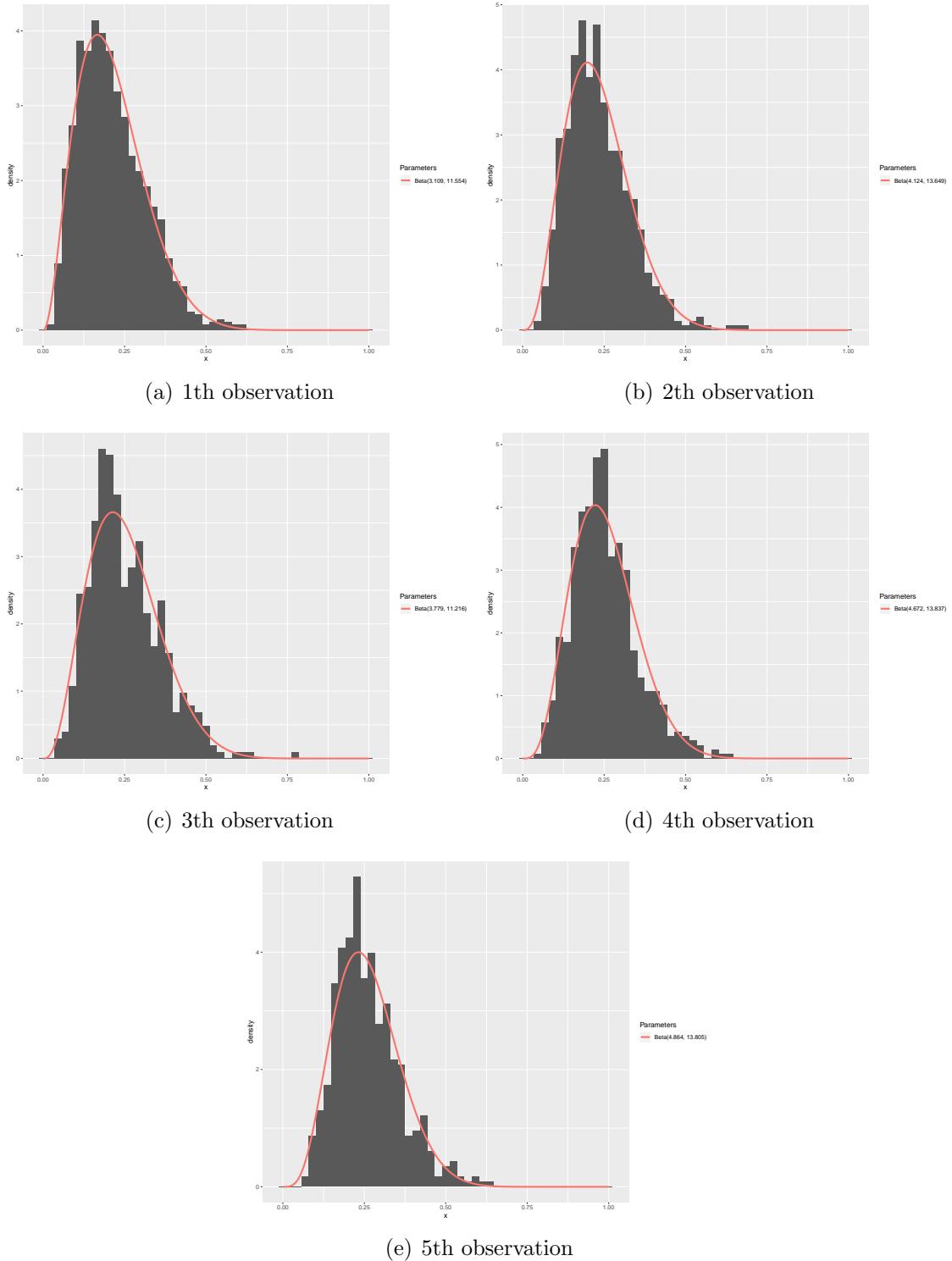


Figura 6: Histograms of the Geodesic Distances between trihedral and the pixels of the sample extracted from Soybeans 231 most similar to trihedral

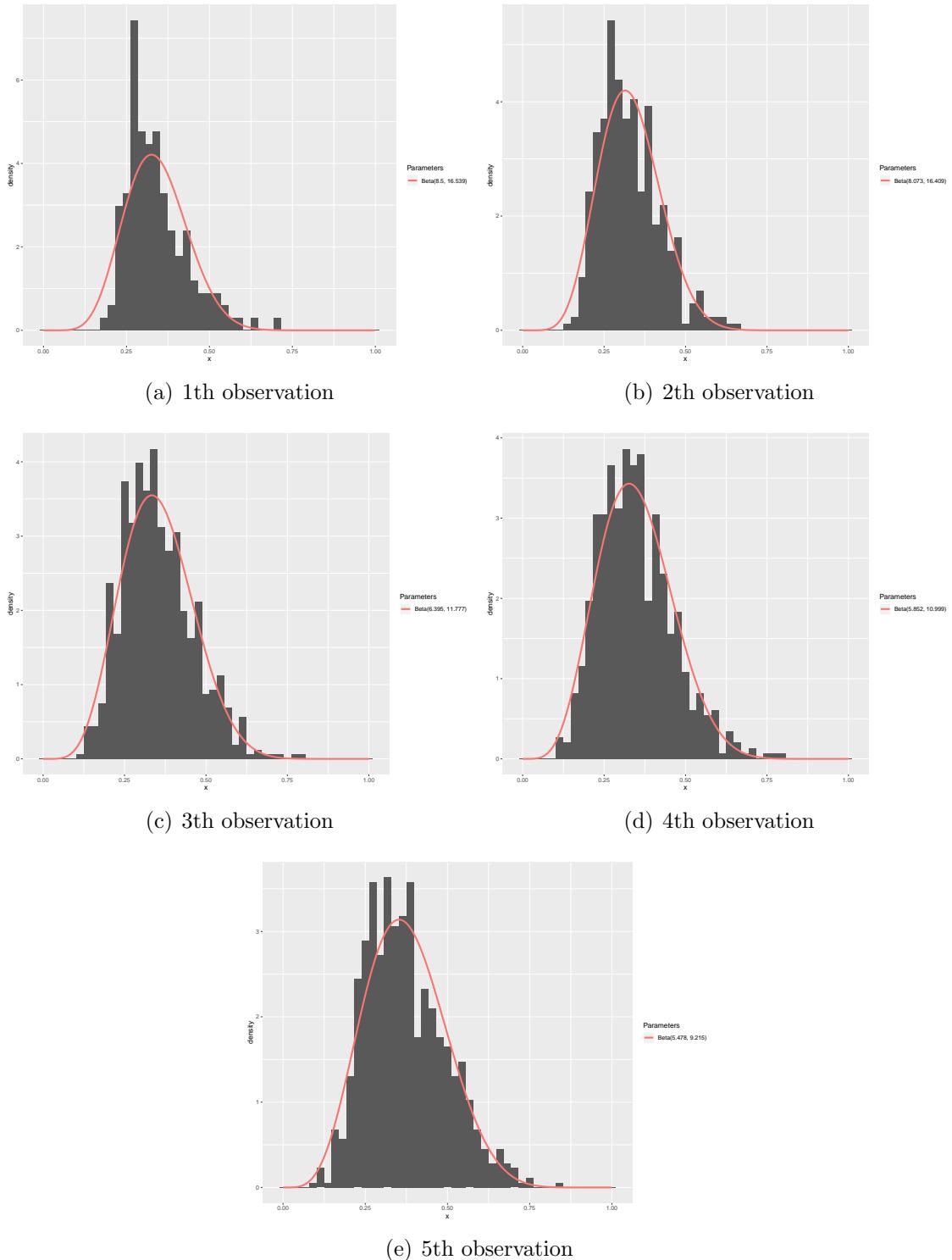
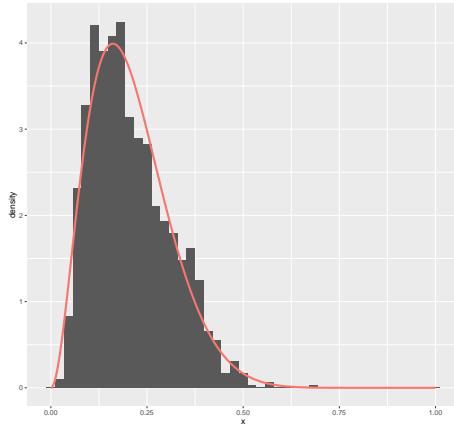
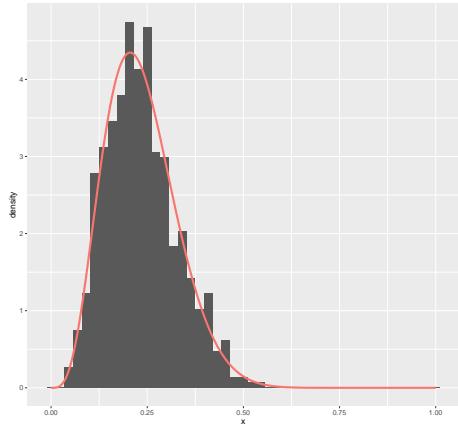


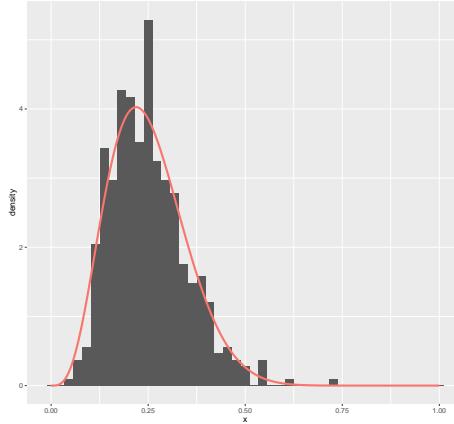
Figura 7: Histograms of the Geodesic Distances between random volume and the pixels of the sample extracted from Soybeans 231 most similar to random volume



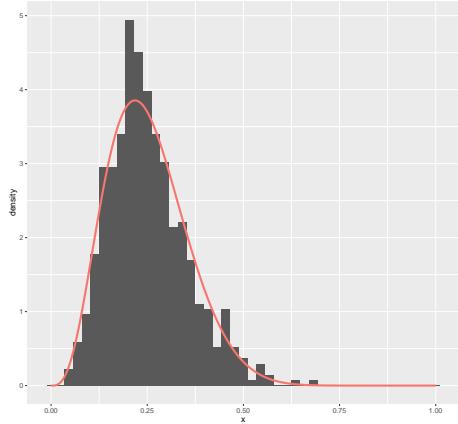
(a) 1th observation



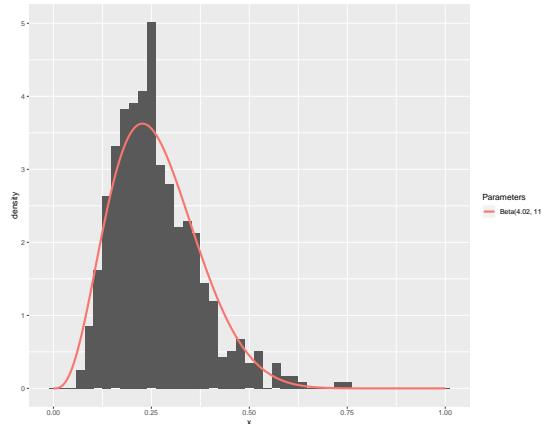
(b) 2th observation



(c) 3th observation



(d) 4th observation



(e) 5th observation

Figura 8: Histograms of the Geodesic Distances between trihedral and the pixels of the sample extracted from Soybeans 232 most similar to trihedral

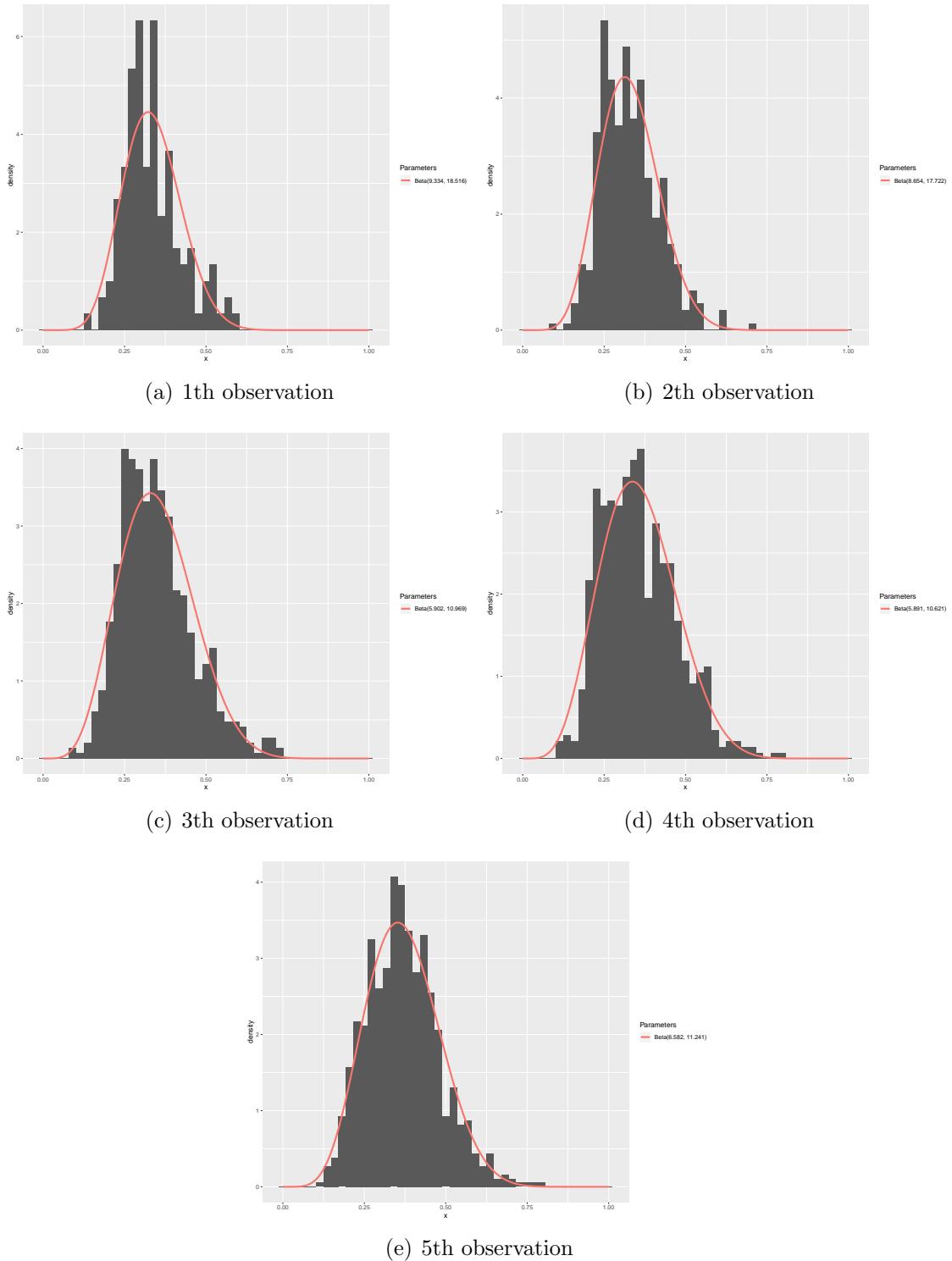


Figura 9: Histograms of the Geodesic Distances between random volume and the pixels of the sample extracted from Soybeans 232 most similar to random volume

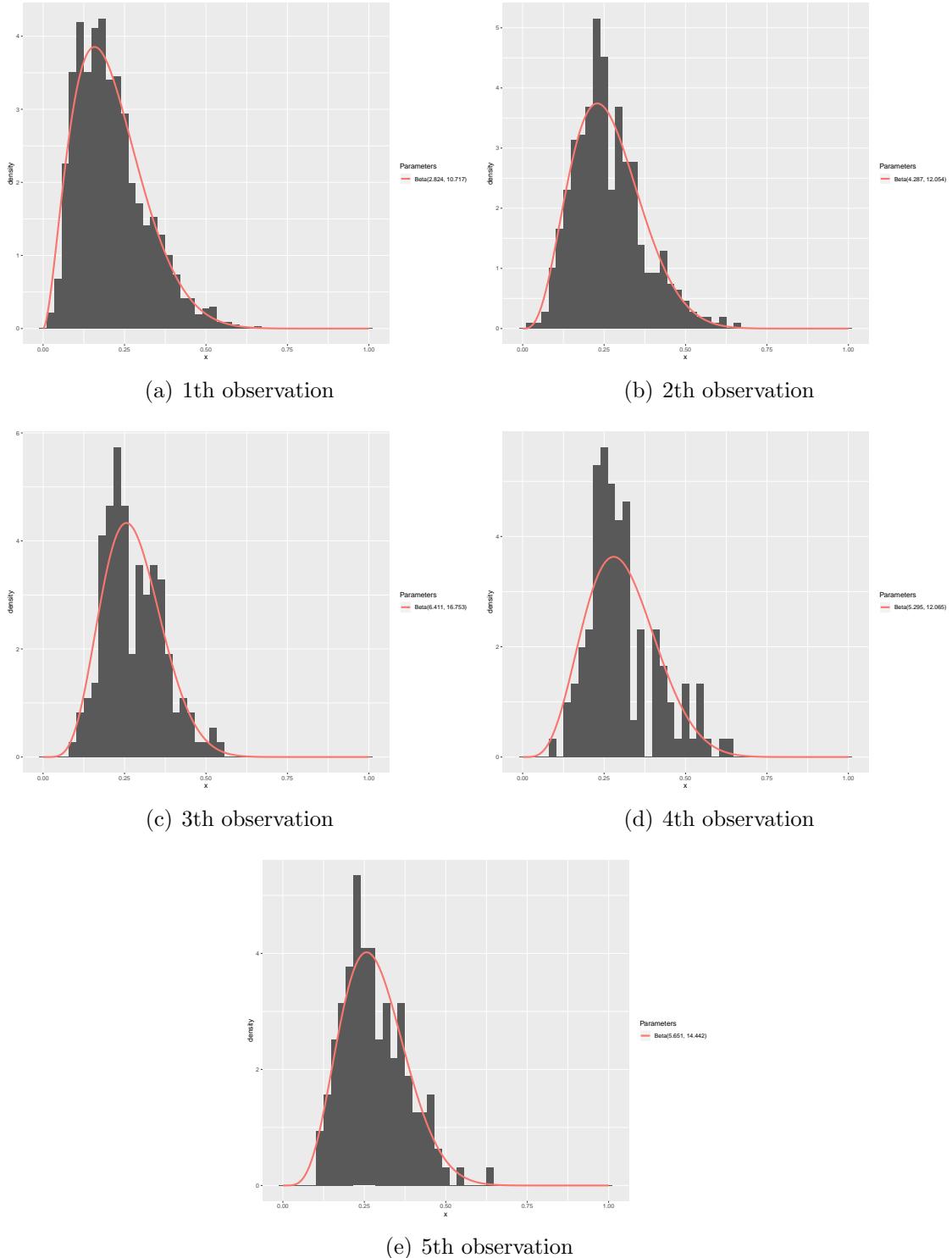
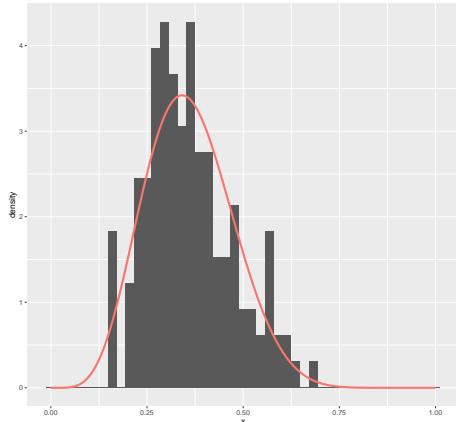
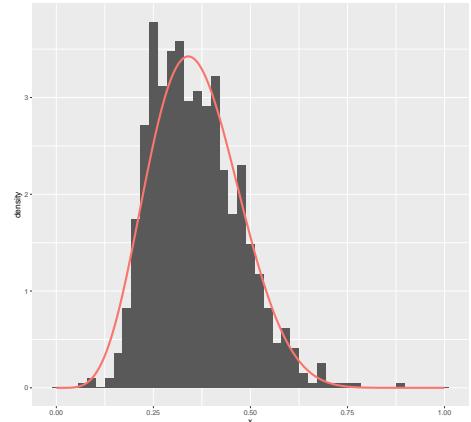


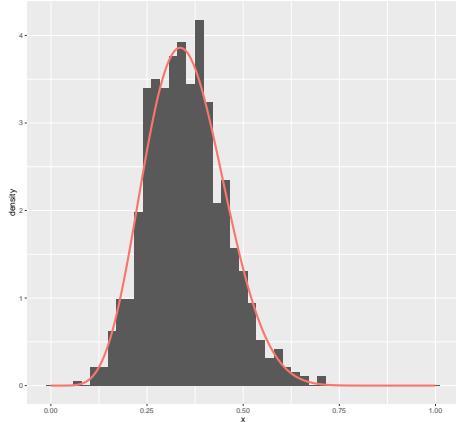
Figura 10: Histograms of the Geodesic Distances between trihedral and the pixels of the sample extracted from Wheat 104 most similar to trihedral



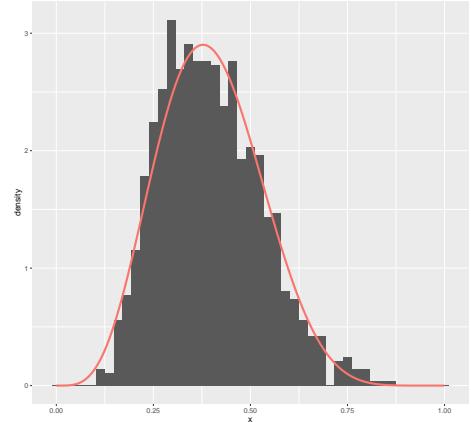
(a) 1th observation



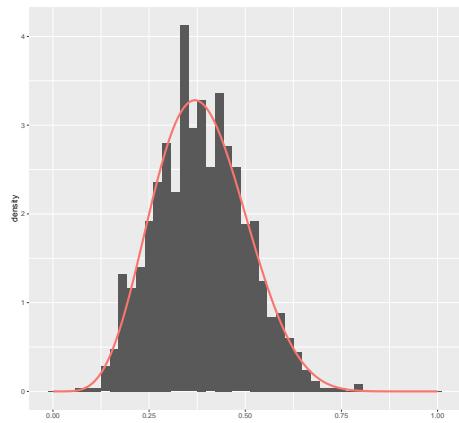
(b) 2th observation



(c) 3th observation

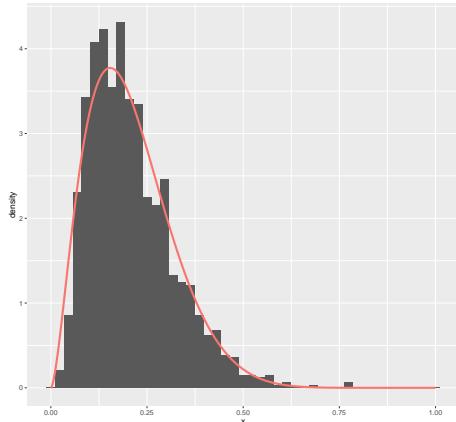


(d) 4th observation

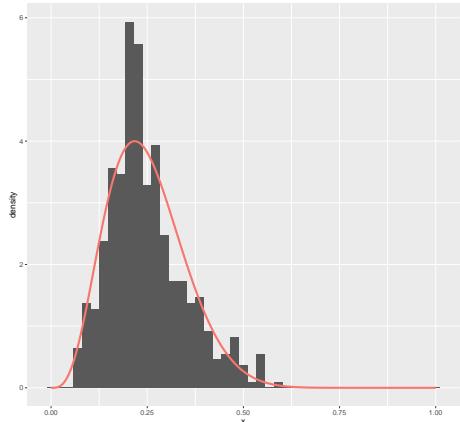


(e) 5th observation

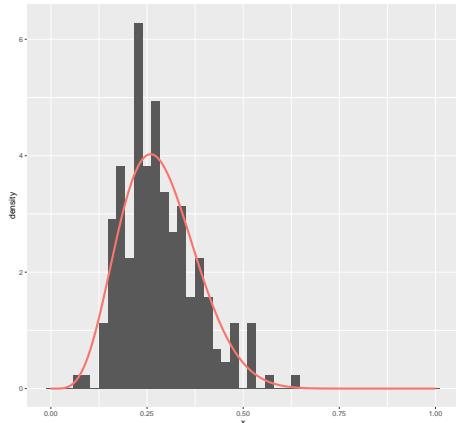
Figura 11: Histograms of the Geodesic Distances between random volume and the pixels of the sample extracted from Wheat 104 most similar to random volume



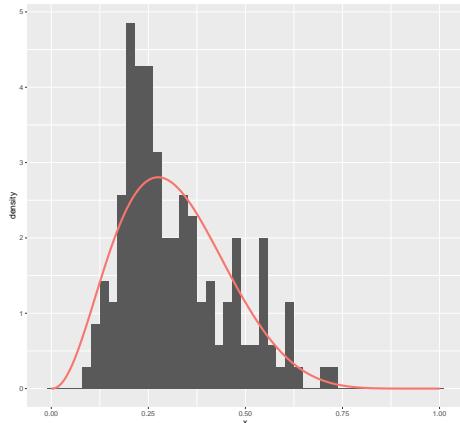
(a) 1th observation



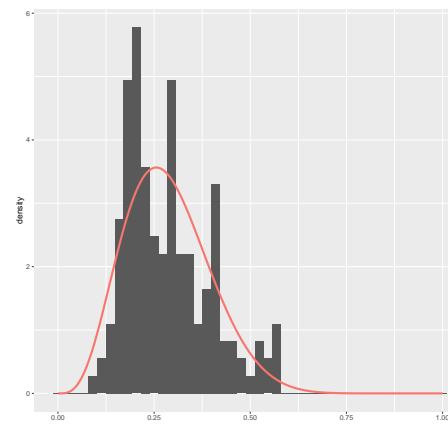
(b) 2th observation



(c) 3th observation



(d) 4th observation



(e) 5th observation

Figura 12: Histograms of the Geodesic Distances between trihedral and the pixels of the sample extracted from Wheat 105 most similar to trihedral

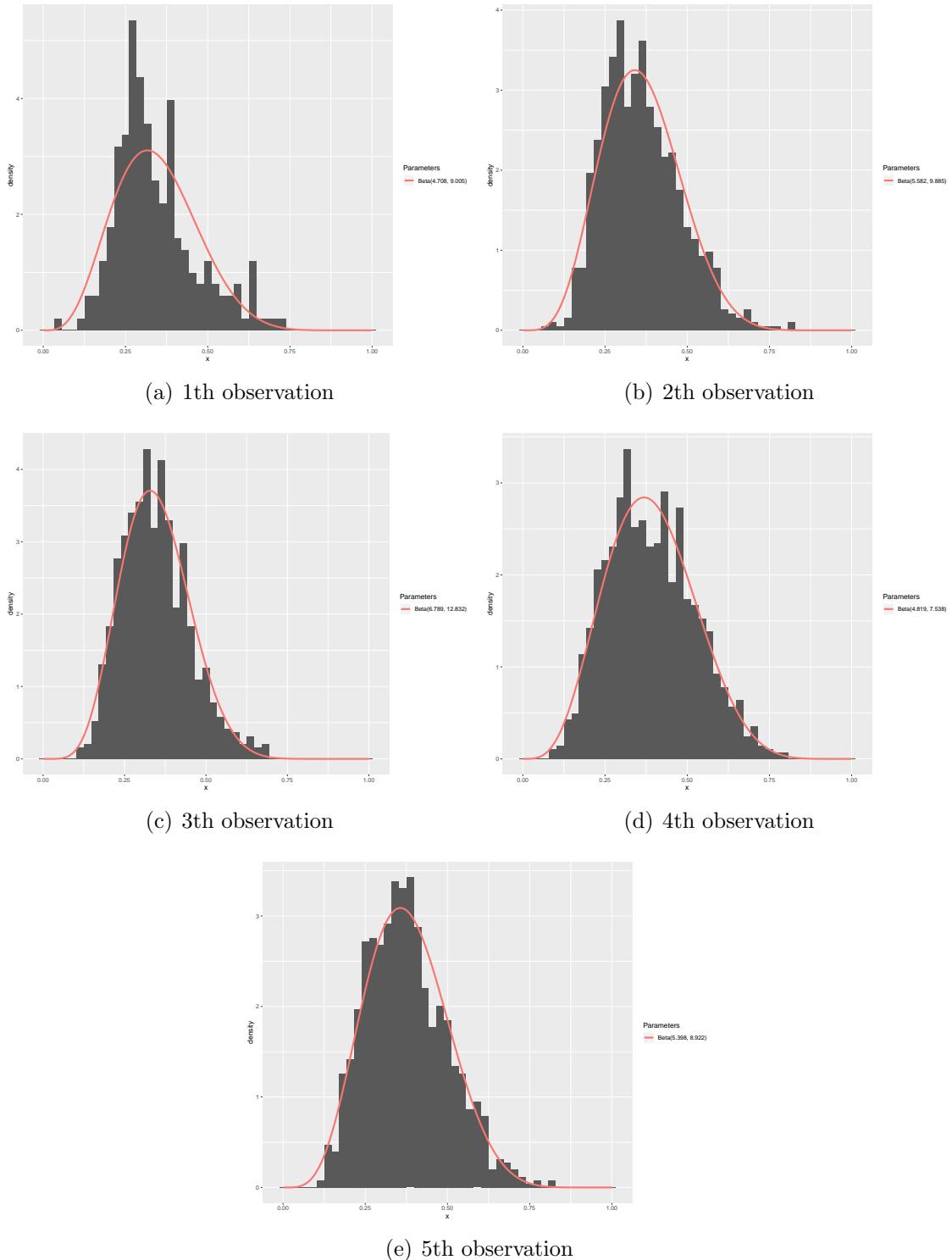
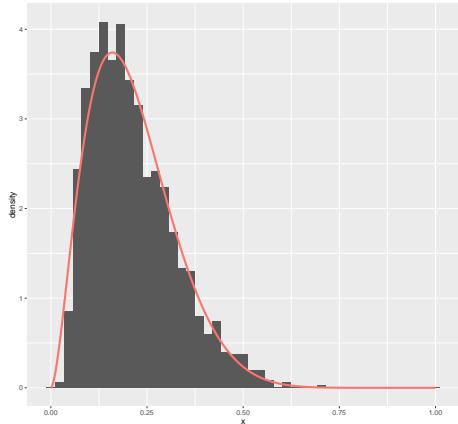
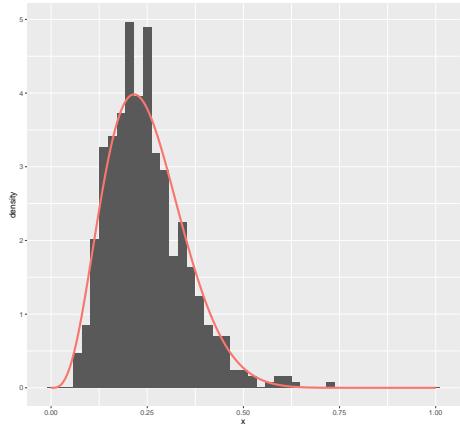


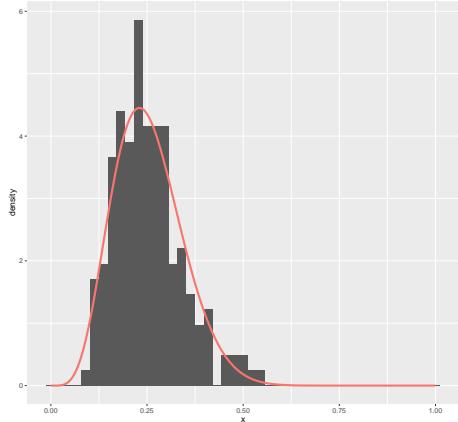
Figura 13: Histograms of the Geodesic Distances between random volume and the pixels of the sample extracted from Wheat 105 most similar to random volume



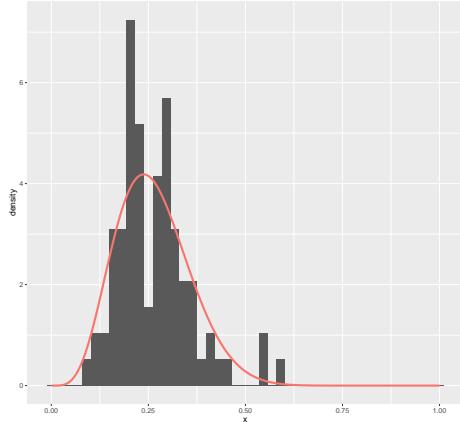
(a) 1th observation



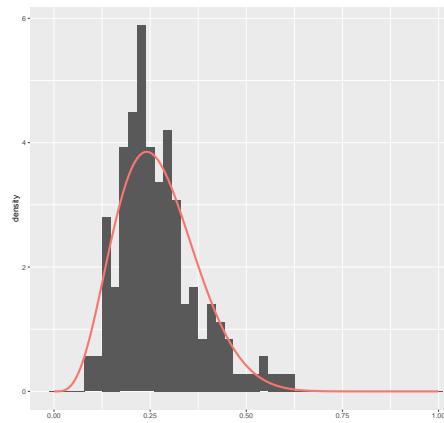
(b) 2th observation



(c) 3th observation

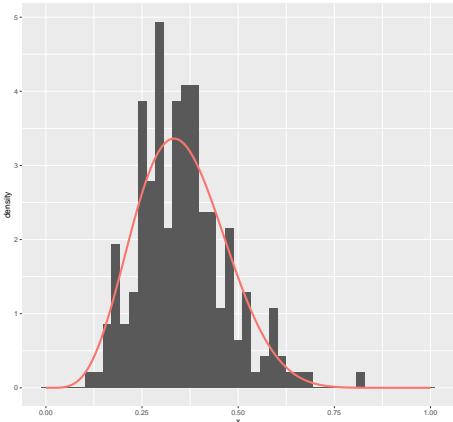


(d) 4th observation

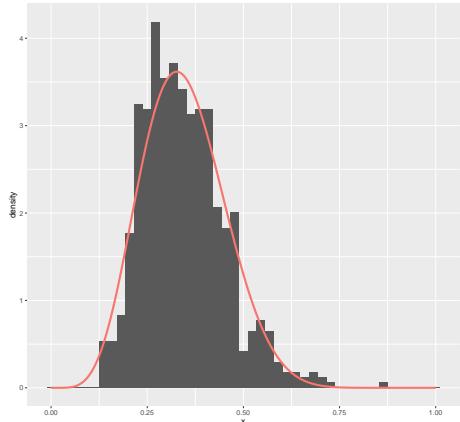


(e) 5th observation

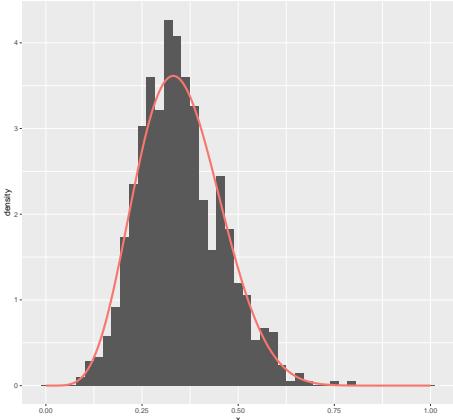
Figura 14: Histograms of the Geodesic Distances between trihedral and the pixels of the sample extracted from Wheat 255 most similar to trihedral



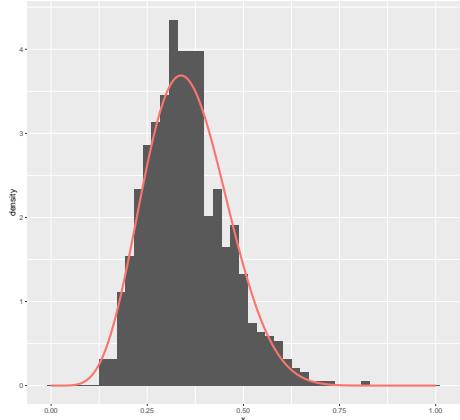
(a) 1th observation



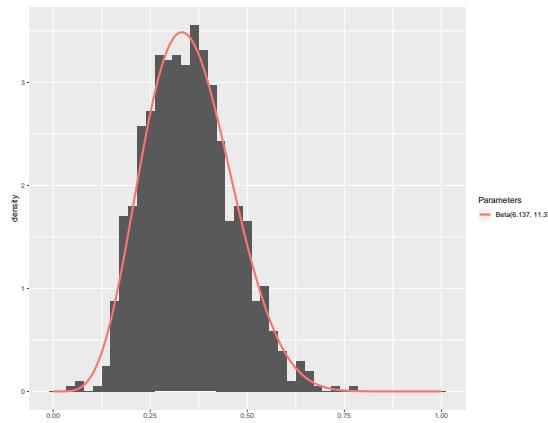
(b) 2th observation



(c) 3th observation



(d) 4th observation



(e) 5th observation

Figura 15: Histograms of the Geodesic Distances between random volume and the pixels of the sample extracted from Wheat 255 most similar to random volume

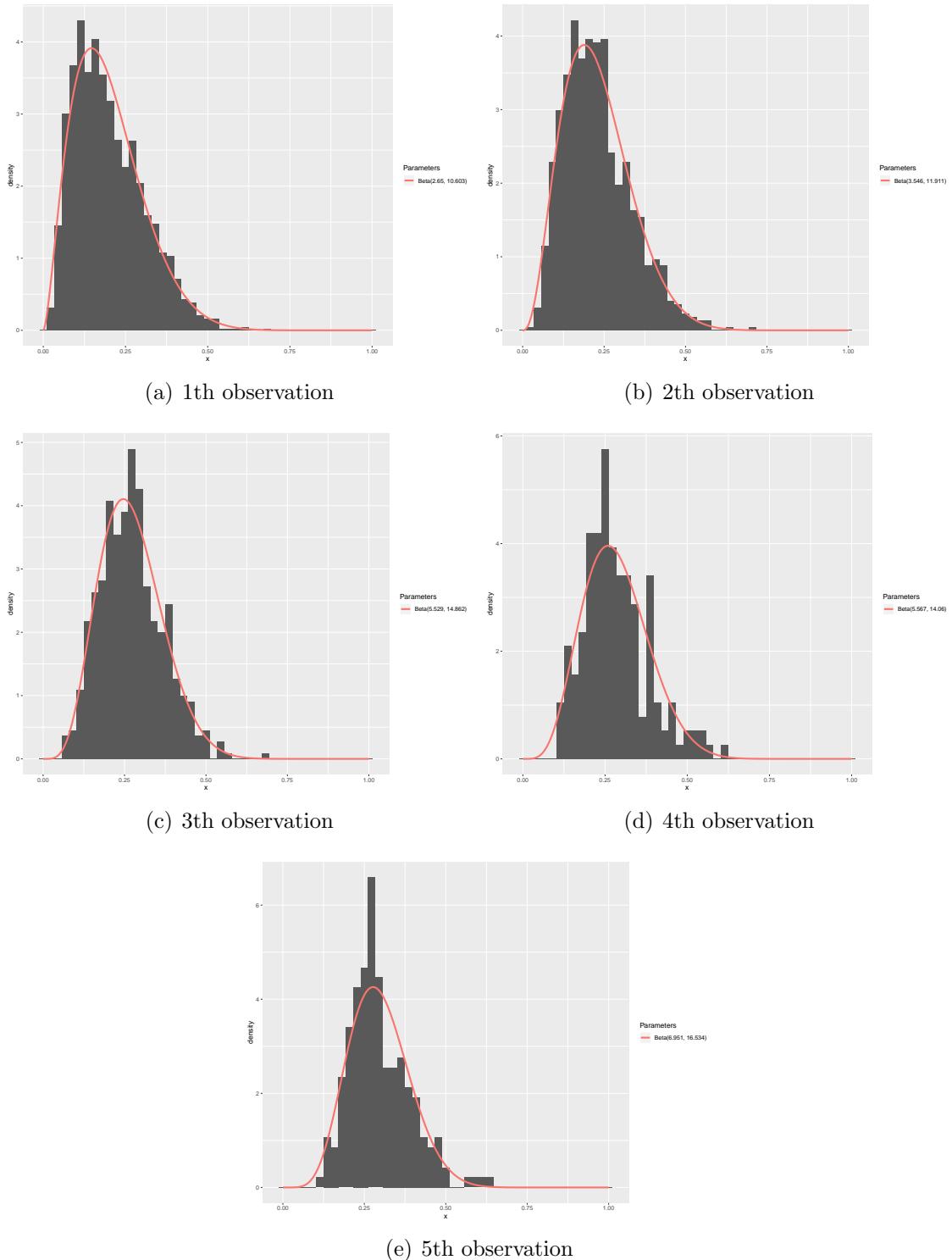


Figura 16: Histograms of the Geodesic Distances between trihedral and the pixels of the sample extracted from Canola 43 most similar to trihedral

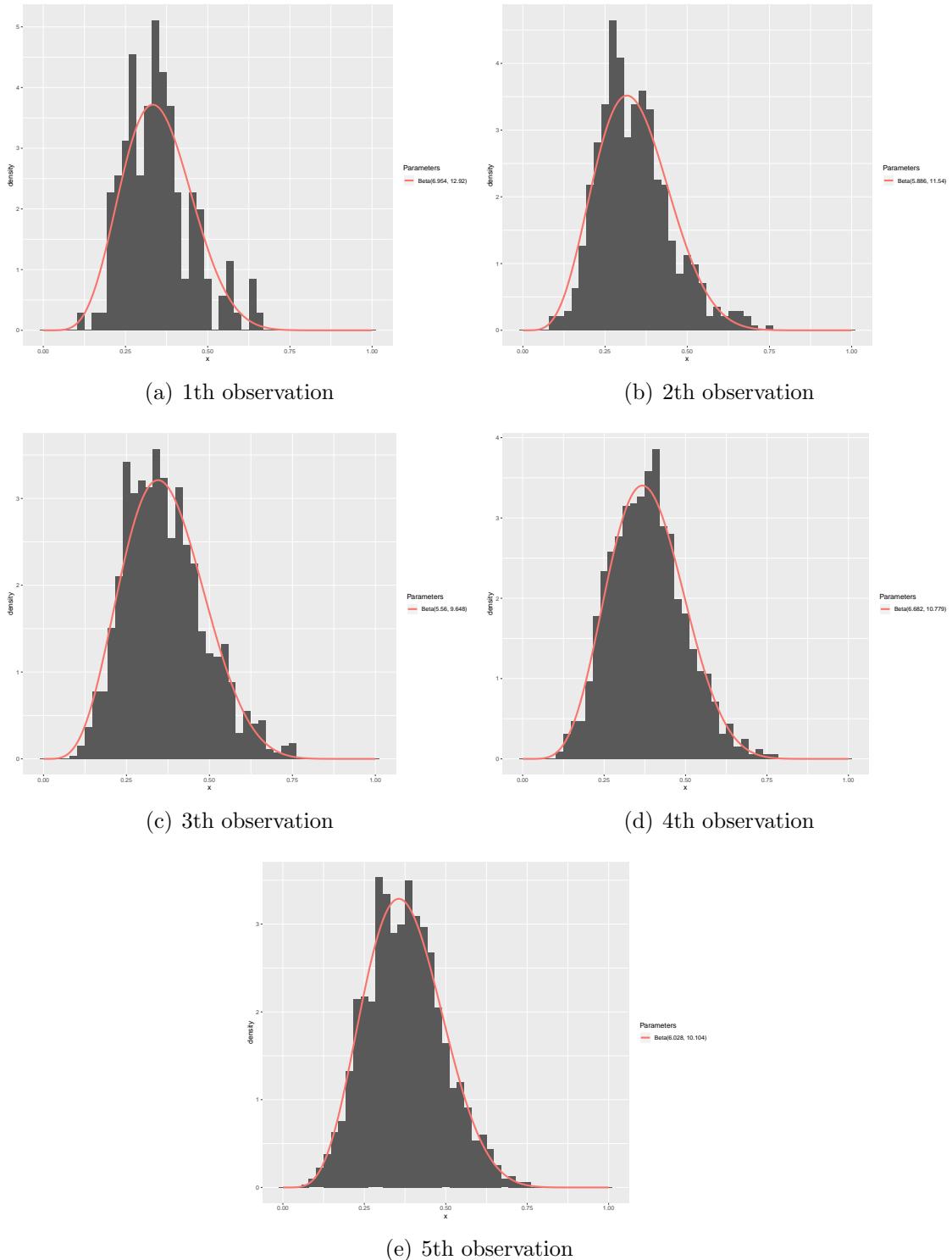
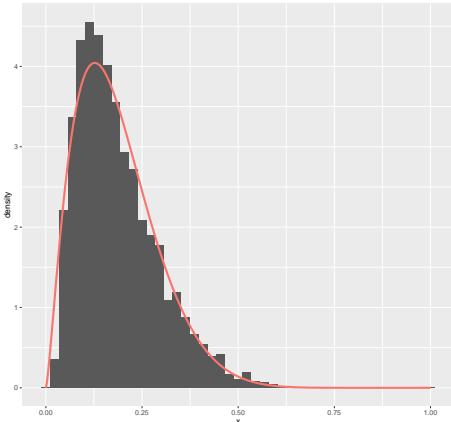
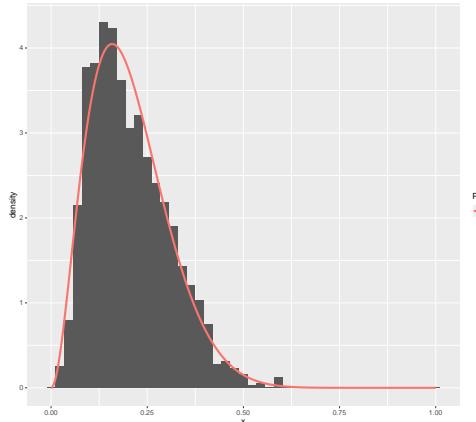


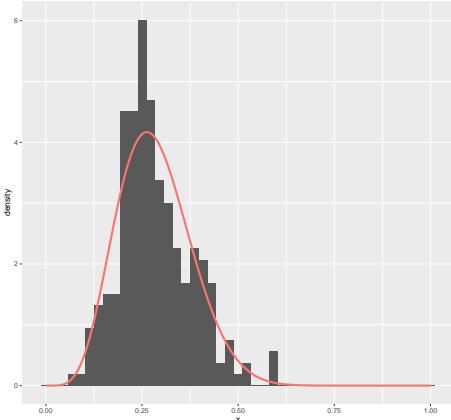
Figura 17: Histograms of the Geodesic Distances between random volume and the pixels of the sample extracted from Canola 43 most similar to random volume



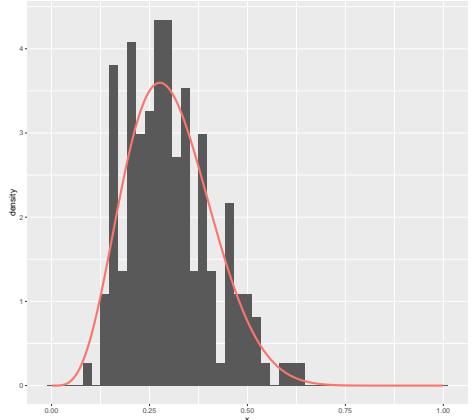
(a) 1th observation



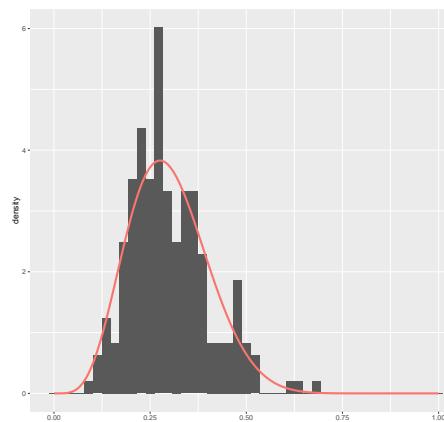
(b) 2th observation



(c) 3th observation



(d) 4th observation



(e) 5th observation

Figura 18: Histograms of the Geodesic Distances between trihedral and the pixels of the sample extracted from Canola 224 most similar to trihedral

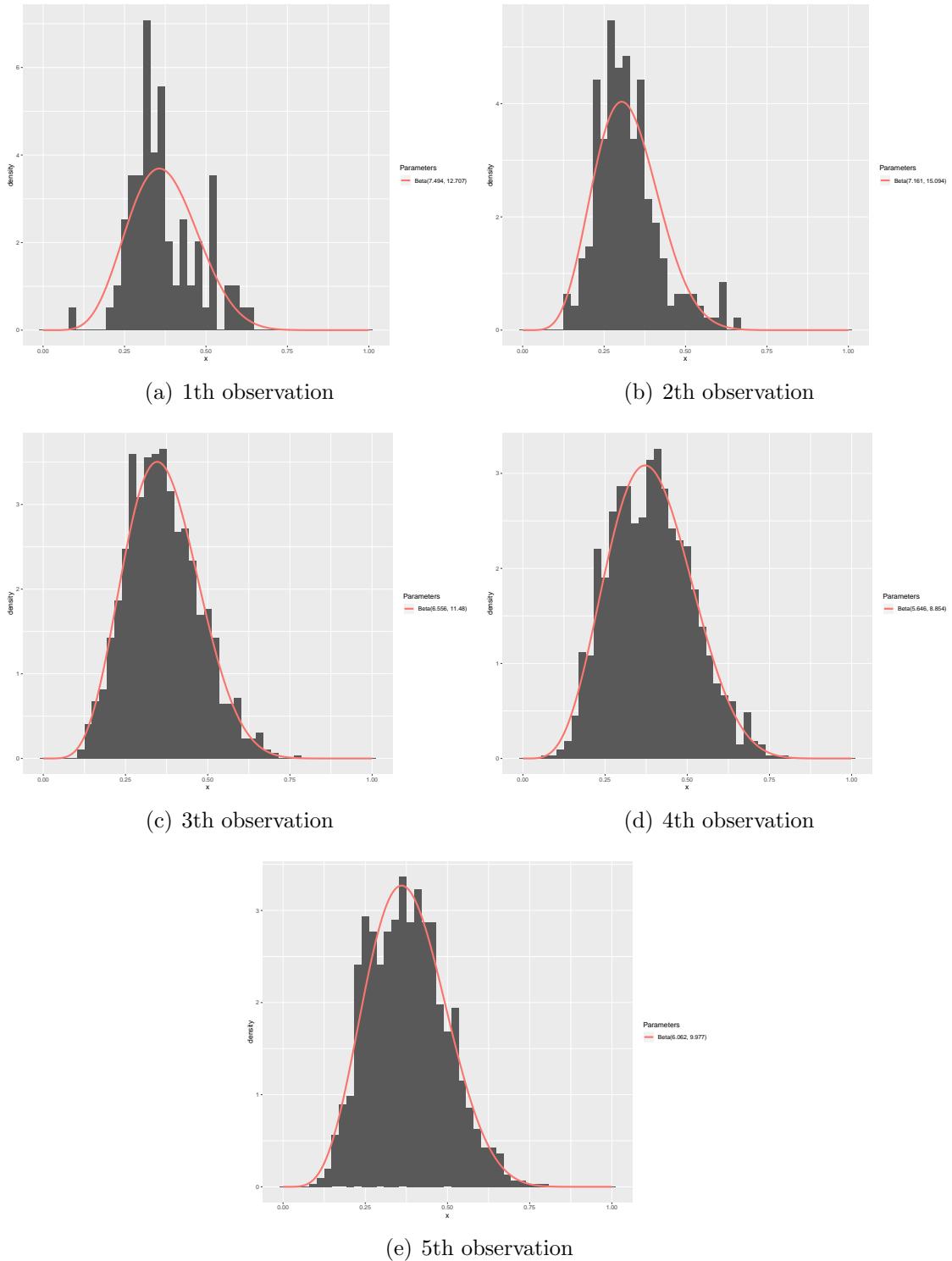


Figura 19: Histograms of the Geodesic Distances between random volume and the pixels of the sample extracted from Canola 224 most similar to random volume

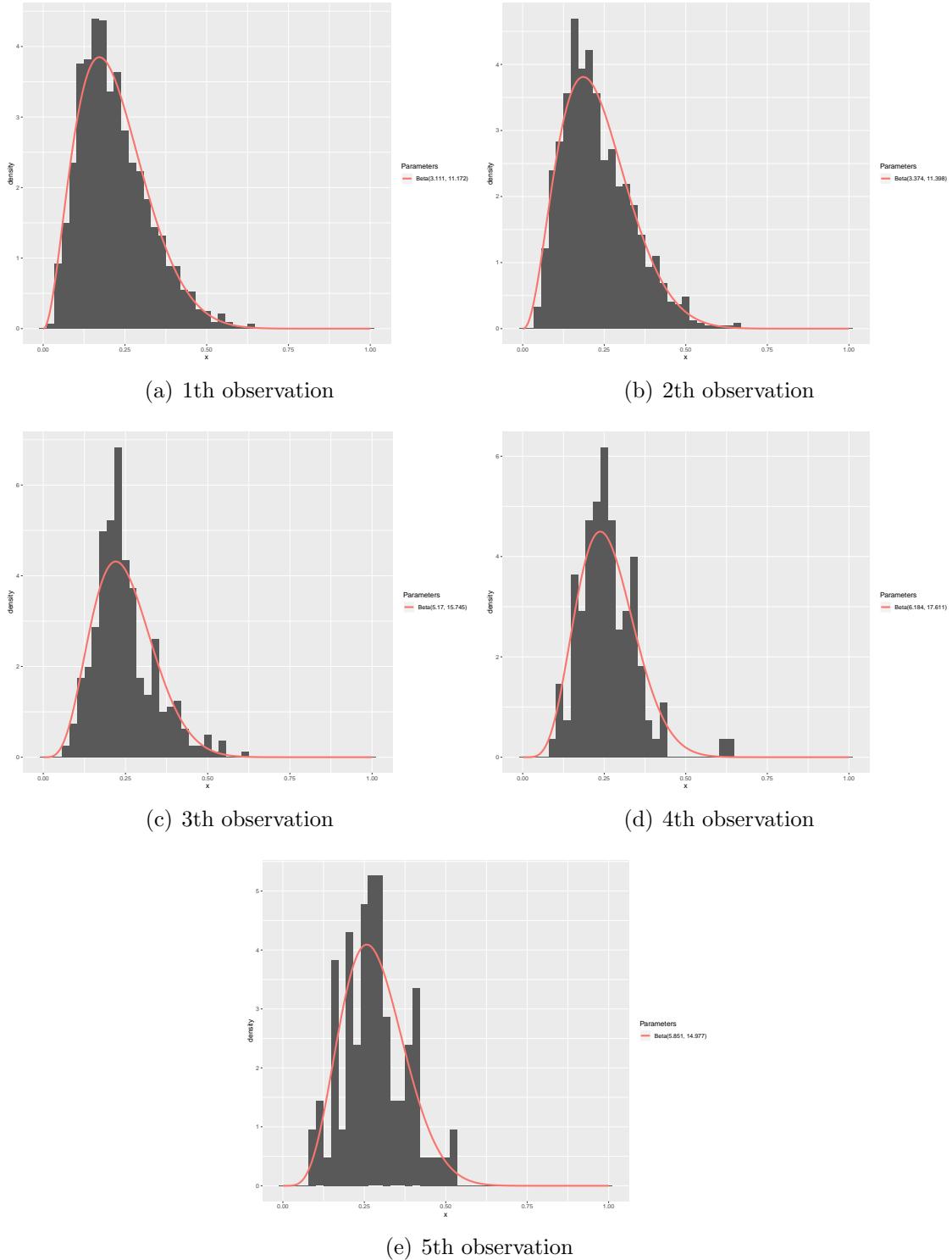


Figura 20: Histograms of the Geodesic Distances between trihedral and the pixels of the sample extracted from Oats 102 most similar to trihedral

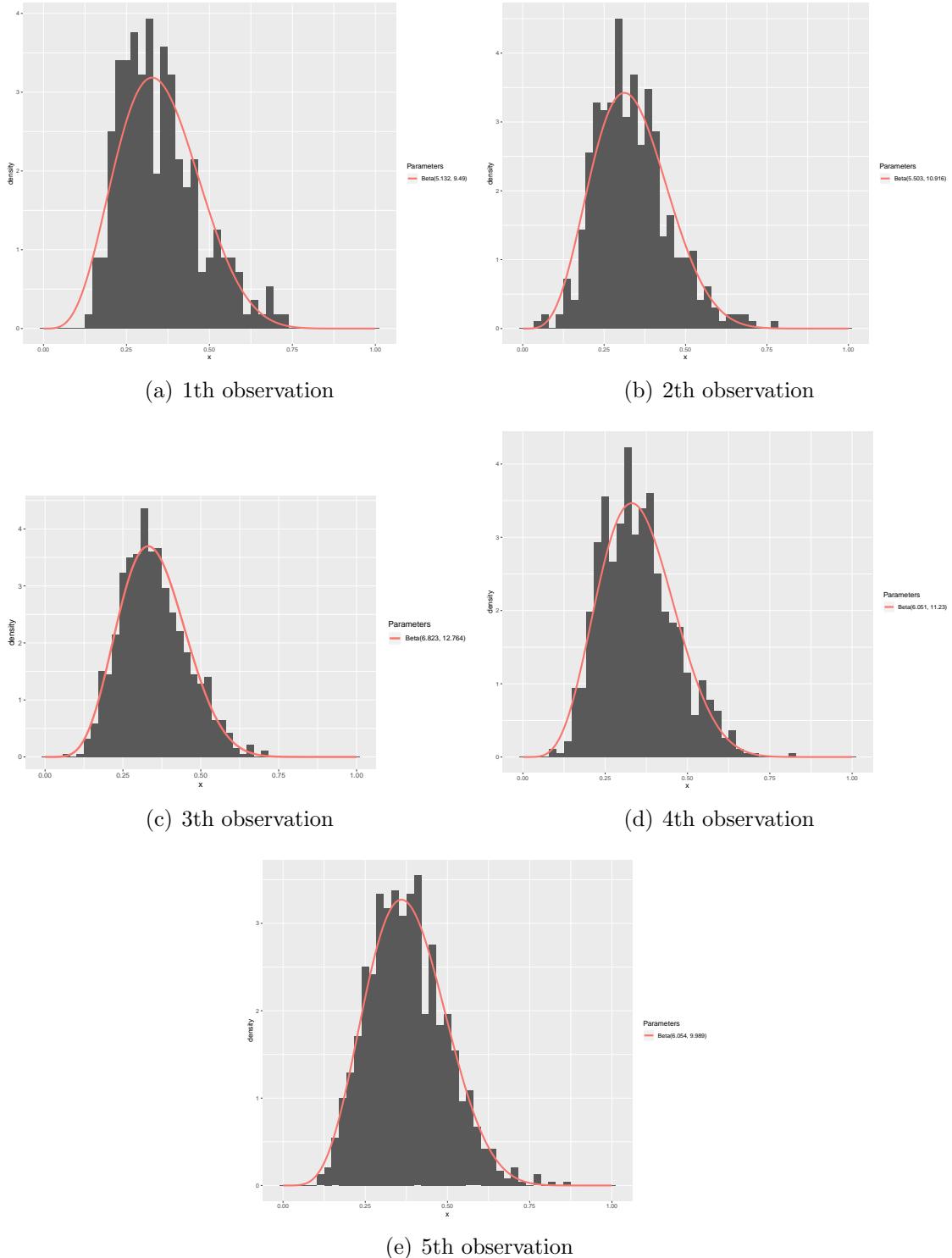
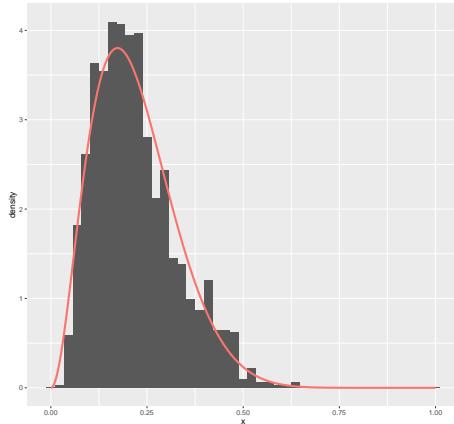
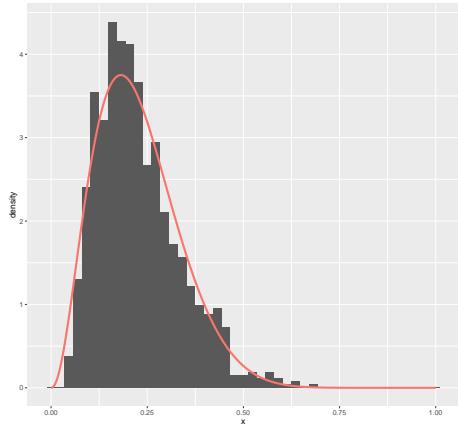


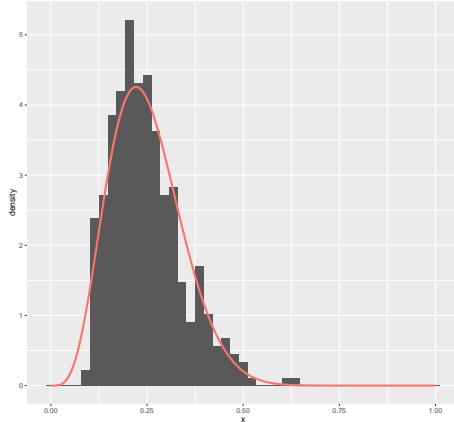
Figura 21: Histograms of the Geodesic Distances between random volume and the pixels of the sample extracted from Oats 102 most similar to random volume



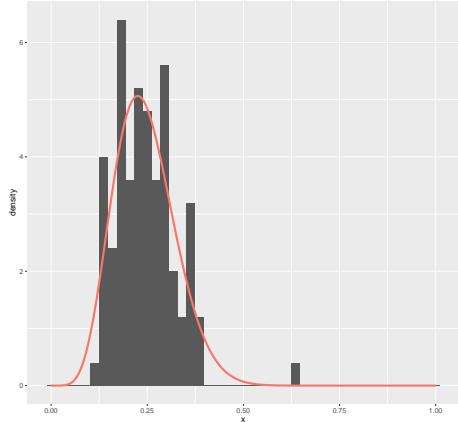
(a) 1th observation



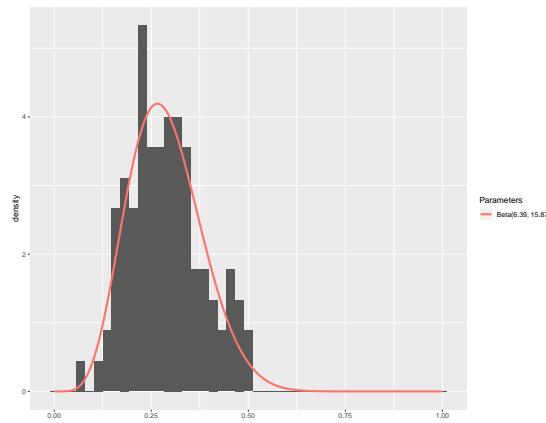
(b) 2th observation



(c) 3th observation

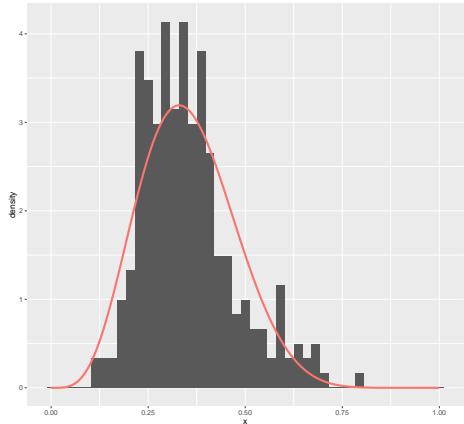


(d) 4th observation

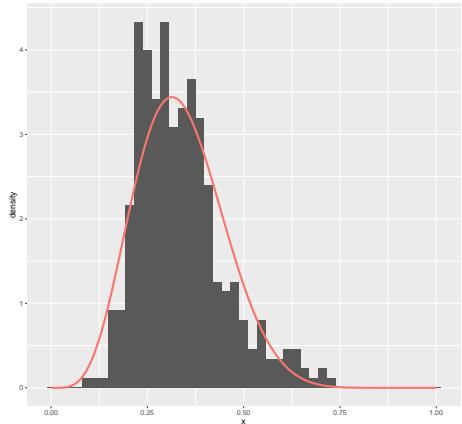


(e) 5th observation

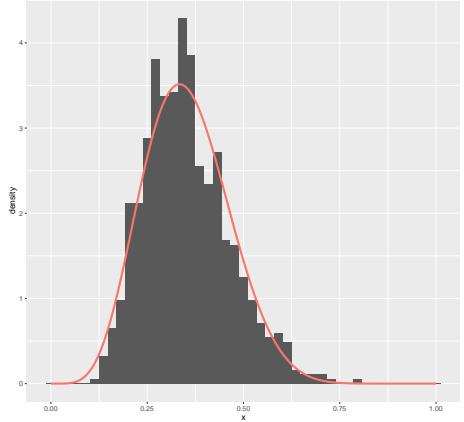
Figura 22: Histograms of the Geodesic Distances between trihedral and the pixels of the sample extracted from Oats 103 most similar to trihedral



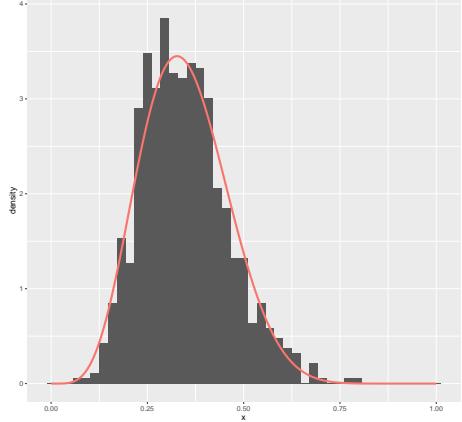
(a) 1th observation



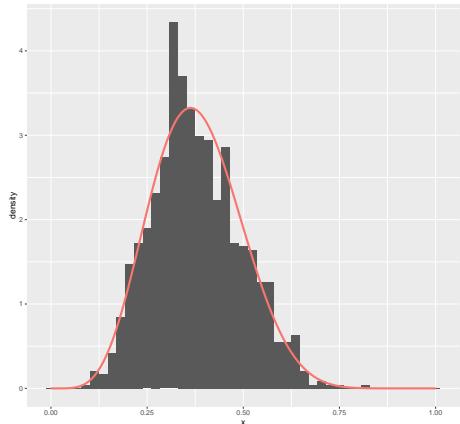
(b) 2th observation



(c) 3th observation



(d) 4th observation



(e) 5th observation

Figura 23: Histograms of the Geodesic Distances between random volume and the pixels of the sample extracted from Oats 103 most similar to random volume

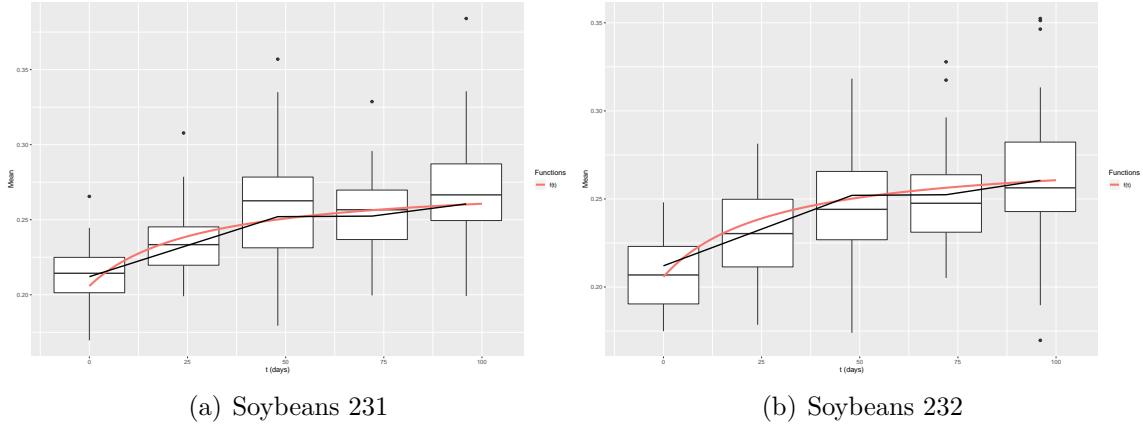


Figura 24: Mean of the distances between trihedral and samples extracted from Soybeans 231 and 232 over time

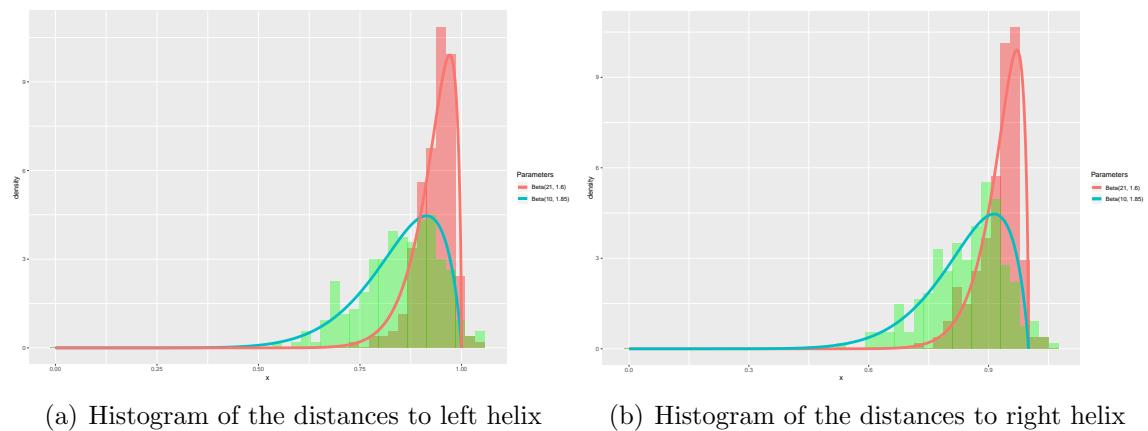


Figura 25: Histograms of the distances between subregion extracted from Soybeans 231 and elementary scatterers

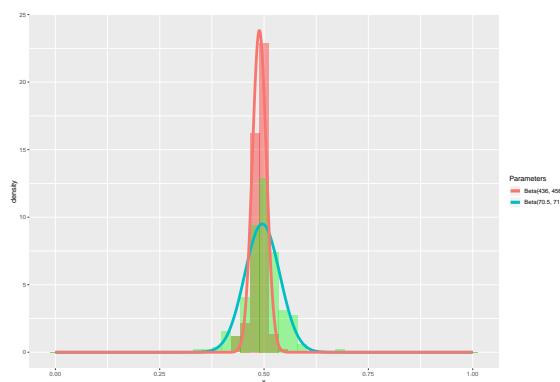


Figura 26: Modified geodesic distance between dihedral and samples