**Approach -**

**Confidence Intervals for hit-rates**

Timestamps are binned into 6 categories (time slots)

Hit rates are calculated as percentages for each advertisement posted or sent

Plot distribution of hit-rates for each time slot

Impact of features such as Offer type, customer segment, mode of communication (SMS, email, etc) on hit rates in each time slot can be gauged using ANOVA

Tukey HSD post hoc test is performed where association is significant

This can be done for overall data as well as for each time slot separately

Find best fit distribution for each time slot with gaussian mixture models using expectation maximization algorithm.

Fit of the distribution can also be tested using QQ envelope technique where large samples are drawn from fitted distribution and QQ-plots are plotted against original sample.

QQ envelope thus obtained will show how well the fit reproduces the original data.

Parameters (means and variances) are point estimated from the components of the best-fit GMM.

Calculate 95% and 99% confidence intervals for mean and variance of hit rates using normal approximation.

Bootstrap sampling method is used as a non parametric approach to calculate 95% and 99% confidence intervals for mean and standard deviation.

Following is some EDA done on simulated data with 120 rows:

Sample from simulated data-

Time\_slot Mode Offer\_type Hit\_rates

0 6.0 2.0 1.0 3.96

1 5.0 3.0 2.0 42.17

2 5.0 2.0 3.0 6.94

3 3.0 2.0 2.0 47.50

4 1.0 2.0 3.0 46.49

5 2.0 1.0 2.0 44.79

6 3.0 3.0 3.0 28.91

7 1.0 3.0 1.0 37.39

8 3.0 1.0 1.0 19.10

9 5.0 2.0 1.0 12.46

Means per time slot

Time\_slot

1.0 24.552448

2.0 25.025956

3.0 26.687273

4.0 25.515426

5.0 24.342048

6.0 25.194323

Name: Hit\_rates, dtype: float64

Standard deviations per time slot

Time\_slot

1.0 15.109318

2.0 14.581392

3.0 13.999039

4.0 14.464514

5.0 13.574330

6.0 15.225088

Name: Hit\_rates, dtype: float64

95% and 99% Confidence intervals for mean hit rates assuming normality:

Time\_slot lower95 lower99 upper95 upper99

0 1 22.050708 21.292606 27.054187 27.812290

1 2 22.891739 22.245006 27.160174 27.806906

2 3 24.529422 23.875528 28.845124 29.499018

3 4 23.426658 22.793698 27.604193 28.237153

4 5 22.255975 21.623832 26.428121 27.060265

5 6 22.772964 22.039219 27.615681 28.349426

Time slot = 1

count 143.000000

mean 24.552448

std 15.109318

min 0.040000

25% 10.730000

50% 23.880000

75% 38.675000

max 49.620000

Name: Hit\_rates, dtype: float64

Time slot = 2

count 183.000000

mean 25.025956

std 14.581392

min 0.090000

25% 12.500000

50% 24.700000

75% 37.360000

max 49.850000

Name: Hit\_rates, dtype: float64

Time slot = 3

count 165.000000

mean 26.687273

std 13.999039

min 0.000000

25% 15.680000

50% 26.340000

75% 38.280000

max 49.890000

Name: Hit\_rates, dtype: float64

Time slot = 4

count 188.000000

mean 25.515426

std 14.464514

min 0.850000

25% 13.352500

50% 26.275000

75% 37.635000

max 49.870000

Name: Hit\_rates, dtype: float64

Time slot = 5

count 166.000000

mean 24.342048

std 13.574330

min 0.110000

25% 13.690000

50% 21.745000

75% 35.745000

max 49.670000

Name: Hit\_rates, dtype: float64

Time slot = 6

count 155.000000

mean 25.194323

std 15.225088

min 0.320000

25% 11.040000

50% 25.700000

75% 38.750000

max 49.910000

Name: Hit\_rates, dtype: float64

Association between Mode of communication and hit rates

OLS Regression Results

==============================================================================

Dep. Variable: Hit\_rates R-squared: 0.002

Model: OLS Adj. R-squared: -0.000

Method: Least Squares F-statistic: 0.8165

Date: Mon, 28 May 2018 Prob (F-statistic): 0.442

Time: 10:27:53 Log-Likelihood: -4089.4

No. Observations: 1000 AIC: 8185.

Df Residuals: 997 BIC: 8200.

Df Model: 2

Covariance Type: nonrobust

==================================================================================

coef std err t P>|t| [0.025 0.975]

----------------------------------------------------------------------------------

Intercept 25.5285 0.795 32.101 0.000 23.968 27.089

C(Mode)[T.2.0] -1.0983 1.119 -0.982 0.327 -3.294 1.097

C(Mode)[T.3.0] 0.2406 1.125 0.214 0.831 -1.966 2.448

==============================================================================

Omnibus: 700.325 Durbin-Watson: 2.128

Prob(Omnibus): 0.000 Jarque-Bera (JB): 59.397

Skew: 0.006 Prob(JB): 1.27e-13

Kurtosis: 1.806 Cond. No. 3.74

==============================================================================

Warnings:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Association between offer type and hit rates

OLS Regression Results

==============================================================================

Dep. Variable: Hit\_rates R-squared: 0.005

Model: OLS Adj. R-squared: 0.003

Method: Least Squares F-statistic: 2.677

Date: Mon, 28 May 2018 Prob (F-statistic): 0.0693

Time: 10:27:53 Log-Likelihood: -4087.5

No. Observations: 1000 AIC: 8181.

Df Residuals: 997 BIC: 8196.

Df Model: 2

Covariance Type: nonrobust

========================================================================================

coef std err t P>|t| [0.025 0.975]

----------------------------------------------------------------------------------------

Intercept 24.4588 0.799 30.626 0.000 22.892 26.026

C(Offer\_type)[T.2.0] 2.2502 1.119 2.012 0.045 0.055 4.445

C(Offer\_type)[T.3.0] 0.0392 1.124 0.035 0.972 -2.167 2.246

==============================================================================

Omnibus: 736.013 Durbin-Watson: 2.125

Prob(Omnibus): 0.000 Jarque-Bera (JB): 60.028

Skew: 0.008 Prob(JB): 9.23e-14

Kurtosis: 1.800 Cond. No. 3.76

==============================================================================

Warnings:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Test results for time slot 1

Mode of communication:

OLS Regression Results

==============================================================================

Dep. Variable: Hit\_rates R-squared: 0.002

Model: OLS Adj. R-squared: -0.012

Method: Least Squares F-statistic: 0.1278

Date: Mon, 28 May 2018 Prob (F-statistic): 0.880

Time: 10:27:53 Log-Likelihood: -590.57

No. Observations: 143 AIC: 1187.

Df Residuals: 140 BIC: 1196.

Df Model: 2

Covariance Type: nonrobust

==================================================================================

coef std err t P>|t| [0.025 0.975]

----------------------------------------------------------------------------------

Intercept 25.2960 2.194 11.528 0.000 20.958 29.634

C(Mode)[T.2.0] -0.6363 3.173 -0.201 0.841 -6.910 5.637

C(Mode)[T.3.0] -1.5360 3.057 -0.502 0.616 -7.581 4.508

==============================================================================

Omnibus: 114.665 Durbin-Watson: 1.816

Prob(Omnibus): 0.000 Jarque-Bera (JB): 10.613

Skew: 0.092 Prob(JB): 0.00496

Kurtosis: 1.678 Cond. No. 3.73

==============================================================================

Warnings:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Offer type:

OLS Regression Results

==============================================================================

Dep. Variable: Hit\_rates R-squared: 0.025

Model: OLS Adj. R-squared: 0.011

Method: Least Squares F-statistic: 1.813

Date: Mon, 28 May 2018 Prob (F-statistic): 0.167

Time: 10:27:53 Log-Likelihood: -588.87

No. Observations: 143 AIC: 1184.

Df Residuals: 140 BIC: 1193.

Df Model: 2

Covariance Type: nonrobust

========================================================================================

coef std err t P>|t| [0.025 0.975]

----------------------------------------------------------------------------------------

Intercept 26.8730 2.291 11.729 0.000 22.343 31.403

C(Offer\_type)[T.2.0] -0.7240 3.259 -0.222 0.825 -7.168 5.720

C(Offer\_type)[T.3.0] -5.1972 3.023 -1.719 0.088 -11.174 0.780

==============================================================================

Omnibus: 112.532 Durbin-Watson: 1.817

Prob(Omnibus): 0.000 Jarque-Bera (JB): 10.557

Skew: 0.090 Prob(JB): 0.00510

Kurtosis: 1.681 Cond. No. 3.93

==============================================================================

Warnings:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Test results for time slot 2

Mode of communication:

OLS Regression Results

==============================================================================

Dep. Variable: Hit\_rates R-squared: 0.062

Model: OLS Adj. R-squared: 0.051

Method: Least Squares F-statistic: 5.902

Date: Mon, 28 May 2018 Prob (F-statistic): 0.00329

Time: 10:27:53 Log-Likelihood: -743.75

No. Observations: 183 AIC: 1493.

Df Residuals: 180 BIC: 1503.

Df Model: 2

Covariance Type: nonrobust

==================================================================================

coef std err t P>|t| [0.025 0.975]

----------------------------------------------------------------------------------

Intercept 23.7390 1.804 13.160 0.000 20.180 27.299

C(Mode)[T.2.0] -2.2379 2.551 -0.877 0.382 -7.272 2.796

C(Mode)[T.3.0] 6.3433 2.583 2.456 0.015 1.246 11.441

==============================================================================

Omnibus: 14.965 Durbin-Watson: 2.052

Prob(Omnibus): 0.001 Jarque-Bera (JB): 5.410

Skew: 0.020 Prob(JB): 0.0669

Kurtosis: 2.159 Cond. No. 3.71

==============================================================================

Warnings:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Offer type:

OLS Regression Results

==============================================================================

Dep. Variable: Hit\_rates R-squared: 0.006

Model: OLS Adj. R-squared: -0.005

Method: Least Squares F-statistic: 0.5189

Date: Mon, 28 May 2018 Prob (F-statistic): 0.596

Time: 10:27:53 Log-Likelihood: -749.03

No. Observations: 183 AIC: 1504.

Df Residuals: 180 BIC: 1514.

Df Model: 2

Covariance Type: nonrobust

========================================================================================

coef std err t P>|t| [0.025 0.975]

----------------------------------------------------------------------------------------

Intercept 24.4853 1.828 13.398 0.000 20.879 28.091

C(Offer\_type)[T.2.0] 2.0122 2.584 0.779 0.437 -3.088 7.112

C(Offer\_type)[T.3.0] -0.5426 2.688 -0.202 0.840 -5.847 4.762

==============================================================================

Omnibus: 82.929 Durbin-Watson: 2.147

Prob(Omnibus): 0.000 Jarque-Bera (JB): 11.046

Skew: -0.009 Prob(JB): 0.00399

Kurtosis: 1.797 Cond. No. 3.66

==============================================================================

Warnings:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

There is a significant association between Hit rates and mode of communication. Here is the post-hoc test for mode of communication:

Multiple Comparison of Means - Tukey HSD,FWER=0.05

=============================================

group1 group2 meandiff lower upper reject

---------------------------------------------

1.0 2.0 -2.2379 -8.2672 3.7914 False

1.0 3.0 6.3433 0.2379 12.4488 True

2.0 3.0 8.5812 2.4758 14.6867 True

Test results for time slot 3

Mode of communication:

OLS Regression Results

==============================================================================

Dep. Variable: Hit\_rates R-squared: 0.000

Model: OLS Adj. R-squared: -0.012

Method: Least Squares F-statistic: 0.03271

Date: Mon, 28 May 2018 Prob (F-statistic): 0.968

Time: 10:27:53 Log-Likelihood: -669.02

No. Observations: 165 AIC: 1344.

Df Residuals: 162 BIC: 1353.

Df Model: 2

Covariance Type: nonrobust

==================================================================================

coef std err t P>|t| [0.025 0.975]

----------------------------------------------------------------------------------

Intercept 26.6409 1.899 14.030 0.000 22.891 30.391

C(Mode)[T.2.0] 0.4232 2.711 0.156 0.876 -4.929 5.776

C(Mode)[T.3.0] -0.2593 2.662 -0.097 0.923 -5.516 4.997

==============================================================================

Omnibus: 37.148 Durbin-Watson: 2.286

Prob(Omnibus): 0.000 Jarque-Bera (JB): 8.309

Skew: -0.128 Prob(JB): 0.0157

Kurtosis: 1.931 Cond. No. 3.73

==============================================================================

Warnings:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Offer type:

OLS Regression Results

==============================================================================

Dep. Variable: Hit\_rates R-squared: 0.005

Model: OLS Adj. R-squared: -0.008

Method: Least Squares F-statistic: 0.3871

Date: Mon, 28 May 2018 Prob (F-statistic): 0.680

Time: 10:27:54 Log-Likelihood: -668.66

No. Observations: 165 AIC: 1343.

Df Residuals: 162 BIC: 1353.

Df Model: 2

Covariance Type: nonrobust

========================================================================================

coef std err t P>|t| [0.025 0.975]

----------------------------------------------------------------------------------------

Intercept 25.7693 2.194 11.743 0.000 21.436 30.103

C(Offer\_type)[T.2.0] 2.1352 2.820 0.757 0.450 -3.433 7.703

C(Offer\_type)[T.3.0] 0.2779 2.838 0.098 0.922 -5.326 5.882

==============================================================================

Omnibus: 35.866 Durbin-Watson: 2.285

Prob(Omnibus): 0.000 Jarque-Bera (JB): 8.388

Skew: -0.155 Prob(JB): 0.0151

Kurtosis: 1.940 Cond. No. 4.27

==============================================================================

Warnings:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Test results for time slot 4

Mode of communication:

OLS Regression Results

==============================================================================

Dep. Variable: Hit\_rates R-squared: 0.003

Model: OLS Adj. R-squared: -0.007

Method: Least Squares F-statistic: 0.3169

Date: Mon, 28 May 2018 Prob (F-statistic): 0.729

Time: 10:27:54 Log-Likelihood: -768.22

No. Observations: 188 AIC: 1542.

Df Residuals: 185 BIC: 1552.

Df Model: 2

Covariance Type: nonrobust

==================================================================================

coef std err t P>|t| [0.025 0.975]

----------------------------------------------------------------------------------

Intercept 25.0169 1.906 13.124 0.000 21.256 28.778

C(Mode)[T.2.0] 1.5567 2.561 0.608 0.544 -3.497 6.610

C(Mode)[T.3.0] -0.3166 2.696 -0.117 0.907 -5.635 5.002

==============================================================================

Omnibus: 78.784 Durbin-Watson: 1.669

Prob(Omnibus): 0.000 Jarque-Bera (JB): 11.150

Skew: -0.061 Prob(JB): 0.00379

Kurtosis: 1.813 Cond. No. 3.87

==============================================================================

Warnings:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Offer type:

OLS Regression Results

==============================================================================

Dep. Variable: Hit\_rates R-squared: 0.015

Model: OLS Adj. R-squared: 0.005

Method: Least Squares F-statistic: 1.453

Date: Mon, 28 May 2018 Prob (F-statistic): 0.237

Time: 10:27:54 Log-Likelihood: -767.07

No. Observations: 188 AIC: 1540.

Df Residuals: 185 BIC: 1550.

Df Model: 2

Covariance Type: nonrobust

========================================================================================

coef std err t P>|t| [0.025 0.975]

----------------------------------------------------------------------------------------

Intercept 23.0057 1.928 11.931 0.000 19.202 26.810

C(Offer\_type)[T.2.0] 4.4040 2.587 1.702 0.090 -0.700 9.508

C(Offer\_type)[T.3.0] 2.6378 2.660 0.992 0.323 -2.610 7.886

==============================================================================

Omnibus: 70.737 Durbin-Watson: 1.640

Prob(Omnibus): 0.000 Jarque-Bera (JB): 10.731

Skew: -0.046 Prob(JB): 0.00468

Kurtosis: 1.833 Cond. No. 3.93

==============================================================================

Warnings:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Test results for time slot 5

Mode of communication:

OLS Regression Results

==============================================================================

Dep. Variable: Hit\_rates R-squared: 0.013

Model: OLS Adj. R-squared: 0.001

Method: Least Squares F-statistic: 1.059

Date: Mon, 28 May 2018 Prob (F-statistic): 0.349

Time: 10:27:54 Log-Likelihood: -666.93

No. Observations: 166 AIC: 1340.

Df Residuals: 163 BIC: 1349.

Df Model: 2

Covariance Type: nonrobust

==================================================================================

coef std err t P>|t| [0.025 0.975]

----------------------------------------------------------------------------------

Intercept 24.5784 1.782 13.794 0.000 21.060 28.097

C(Mode)[T.2.0] 1.3251 2.499 0.530 0.597 -3.609 6.259

C(Mode)[T.3.0] -2.4739 2.648 -0.934 0.352 -7.702 2.755

==============================================================================

Omnibus: 33.225 Durbin-Watson: 2.034

Prob(Omnibus): 0.000 Jarque-Bera (JB): 9.114

Skew: 0.250 Prob(JB): 0.0105

Kurtosis: 1.966 Cond. No. 3.67

==============================================================================

Warnings:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Offer type:

OLS Regression Results

==============================================================================

Dep. Variable: Hit\_rates R-squared: 0.012

Model: OLS Adj. R-squared: 0.000

Method: Least Squares F-statistic: 1.021

Date: Mon, 28 May 2018 Prob (F-statistic): 0.363

Time: 10:27:54 Log-Likelihood: -666.97

No. Observations: 166 AIC: 1340.

Df Residuals: 163 BIC: 1349.

Df Model: 2

Covariance Type: nonrobust

========================================================================================

coef std err t P>|t| [0.025 0.975]

----------------------------------------------------------------------------------------

Intercept 22.7682 1.683 13.524 0.000 19.444 26.092

C(Offer\_type)[T.2.0] 3.6692 2.568 1.429 0.155 -1.401 8.740

C(Offer\_type)[T.3.0] 1.5668 2.525 0.620 0.536 -3.420 6.553

==============================================================================

Omnibus: 29.522 Durbin-Watson: 2.028

Prob(Omnibus): 0.000 Jarque-Bera (JB): 9.044

Skew: 0.274 Prob(JB): 0.0109

Kurtosis: 1.997 Cond. No. 3.49

==============================================================================

Warnings:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Test results for time slot 6

Mode of communication:

OLS Regression Results

==============================================================================

Dep. Variable: Hit\_rates R-squared: 0.054

Model: OLS Adj. R-squared: 0.042

Method: Least Squares F-statistic: 4.380

Date: Mon, 28 May 2018 Prob (F-statistic): 0.0141

Time: 10:27:54 Log-Likelihood: -637.15

No. Observations: 155 AIC: 1280.

Df Residuals: 152 BIC: 1289.

Df Model: 2

Covariance Type: nonrobust

==================================================================================

coef std err t P>|t| [0.025 0.975]

----------------------------------------------------------------------------------

Intercept 28.4428 2.107 13.497 0.000 24.279 32.606

C(Mode)[T.2.0] -8.4979 3.027 -2.807 0.006 -14.479 -2.517

C(Mode)[T.3.0] -1.7950 2.876 -0.624 0.533 -7.477 3.886

==============================================================================

Omnibus: 56.065 Durbin-Watson: 2.101

Prob(Omnibus): 0.000 Jarque-Bera (JB): 9.112

Skew: -0.092 Prob(JB): 0.0105

Kurtosis: 1.827 Cond. No. 3.80

==============================================================================

Warnings:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Offer type:

OLS Regression Results

==============================================================================

Dep. Variable: Hit\_rates R-squared: 0.000

Model: OLS Adj. R-squared: -0.013

Method: Least Squares F-statistic: 0.006693

Date: Mon, 28 May 2018 Prob (F-statistic): 0.993

Time: 10:27:54 Log-Likelihood: -641.48

No. Observations: 155 AIC: 1289.

Df Residuals: 152 BIC: 1298.

Df Model: 2

Covariance Type: nonrobust

========================================================================================

coef std err t P>|t| [0.025 0.975]

----------------------------------------------------------------------------------------

Intercept 25.0112 2.012 12.430 0.000 21.036 28.987

C(Offer\_type)[T.2.0] 0.2748 2.927 0.094 0.925 -5.507 6.057

C(Offer\_type)[T.3.0] 0.3132 3.044 0.103 0.918 -5.701 6.328

==============================================================================

Omnibus: 149.579 Durbin-Watson: 2.077

Prob(Omnibus): 0.000 Jarque-Bera (JB): 11.665

Skew: -0.053 Prob(JB): 0.00293

Kurtosis: 1.660 Cond. No. 3.56

==============================================================================

Warnings:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

There is a significant association between Hit rates and mode of communication. Here is the post-hoc test for mode of communication:

Multiple Comparison of Means - Tukey HSD,FWER=0.05

==============================================

group1 group2 meandiff lower upper reject

----------------------------------------------

1.0 2.0 -8.4979 -15.6642 -1.3316 True

1.0 3.0 -1.795 -8.602 5.012 False

2.0 3.0 6.7029 -0.2198 13.6255 False

Time slot 1

Means of fitted gaussian components : [[13.87504826]

[37.98463412]]

Variance of fitted gaussian components : [[73.36567125]

[47.57000497]]

Standard deviation of fitted gaussian components : [[6.8971012 ]

[8.56537631]]

Time slot 2

Means of fitted gaussian components : [[38.4270832 ]

[11.18382738]]

Variance of fitted gaussian components : [[67.41423938]

[56.197648 ]]

Standard deviation of fitted gaussian components : [[8.21061748]

[7.49650905]]

Time slot 3

Means of fitted gaussian components : [[13.36858112]

[35.53834899]]

Variance of fitted gaussian components : [[87.19825229]

[83.03741927]]

Standard deviation of fitted gaussian components : [[9.33800044]

[9.112487 ]]

Time slot 4

Means of fitted gaussian components : [[12.84313391]

[36.17047156]]

Variance of fitted gaussian components : [[73.737043 ]

[84.01830616]]

Standard deviation of fitted gaussian components : [[9.16615002]

[8.5870276 ]]

Time slot 5

Means of fitted gaussian components : [[14.48100243]

[36.7345549 ]]

Variance of fitted gaussian components : [[69.98192622]

[69.4646697 ]]

Standard deviation of fitted gaussian components : [[8.33454676]

[8.36552008]]

Time slot 6

Means of fitted gaussian components : [[34.05962103]

[14.81827899]]

Variance of fitted gaussian components : [[78.31280029]

[86.98999737]]

Standard deviation of fitted gaussian components : [[9.32684284]

[8.84945198]]

95% and 99% bootstrap confidence intervals (mean):

Time\_slot lower95mean lower99mean upper95mean upper99mean

0 1 22.127881 21.343776 27.008154 27.889102

1 2 22.852309 22.154711 27.157295 27.811150

2 3 24.558627 23.938168 28.802094 29.462557

3 4 23.409981 22.760653 27.561191 28.138877

4 5 22.319681 21.667948 26.422539 27.010992

5 6 22.776158 22.001768 27.586203 28.224883

95% and 99% bootstrap confidence intervals (standard deviation):

Time\_slot lower95sd lower99sd upper95sd upper99sd

0 1 22.127881 21.343776 27.008154 27.889102

1 2 22.852309 22.154711 27.157295 27.811150

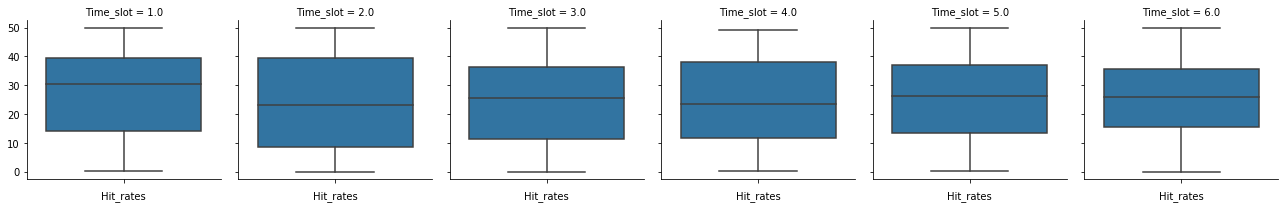
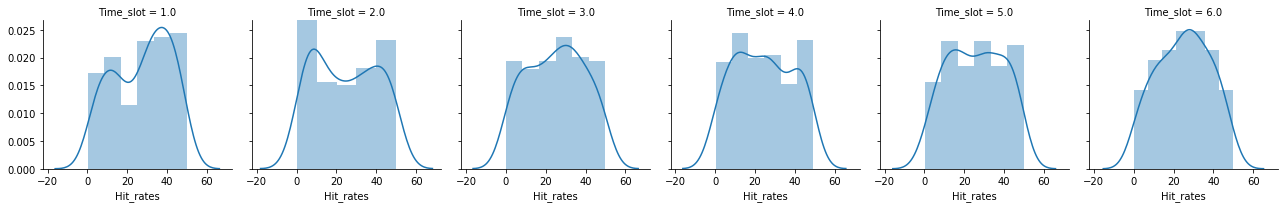
2 3 24.558627 23.938168 28.802094 29.462557

3 4 23.409981 22.760653 27.561191 28.138877

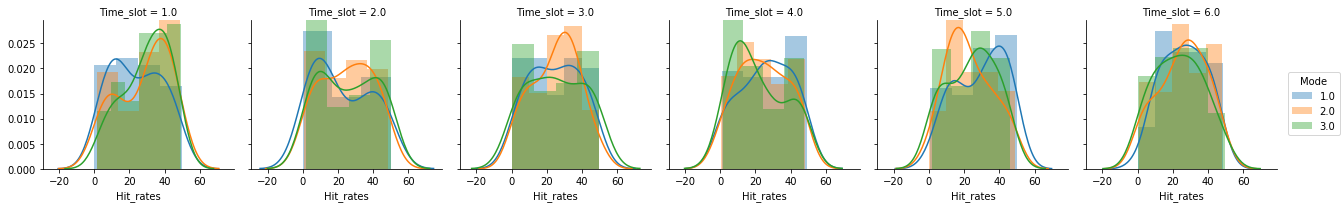
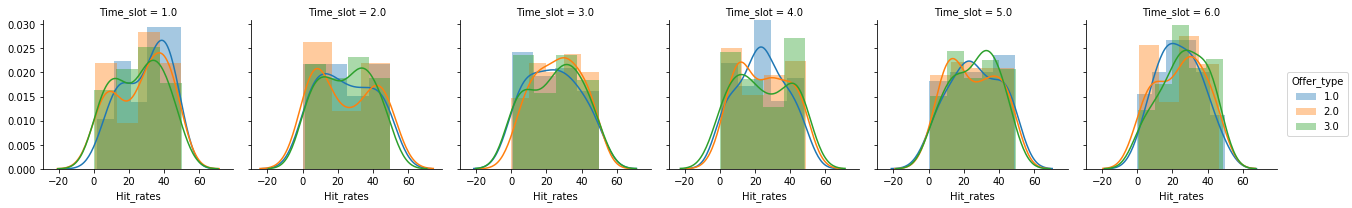
4 5 22.319681 21.667948 26.422539 27.010992

5 6 22.776158 22.001768 27.586203 28.224883

Distribution of hit-rates per time slot -



Segment by offer type per time slot:

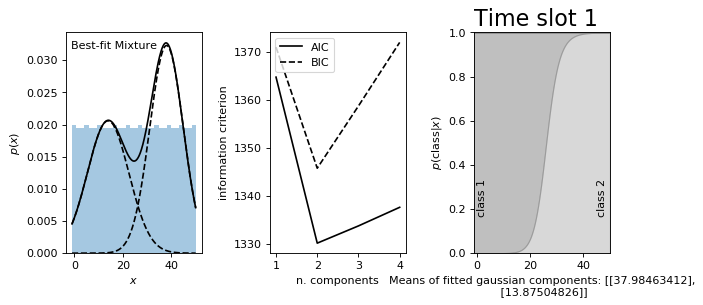
Segment by mode of communication per time slot:

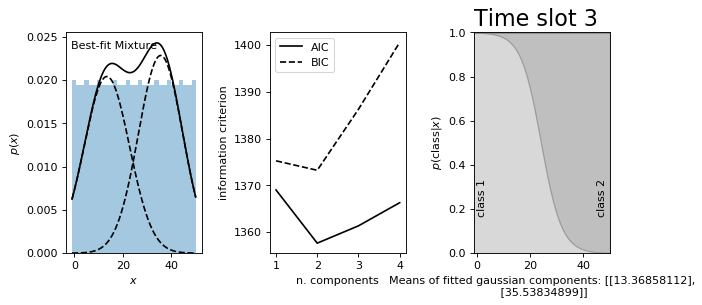
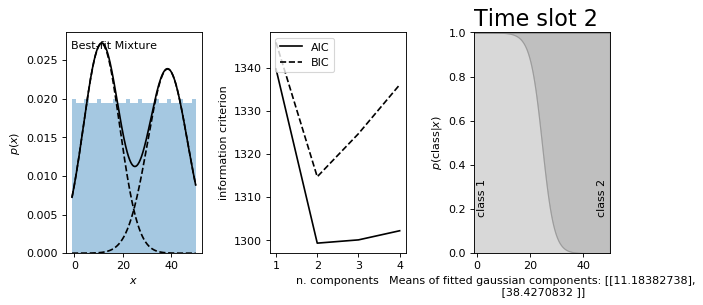
Find best fit GMM using Expectation Maximization:

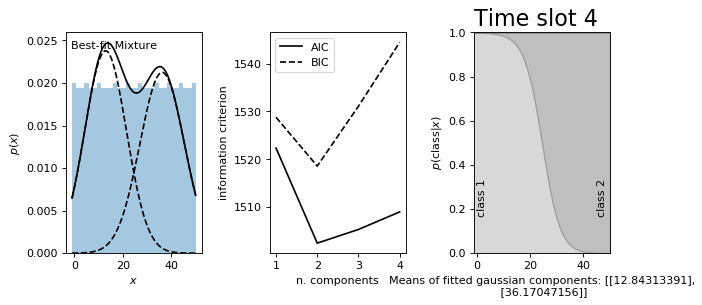
best fit is chosen on the basis of AIC

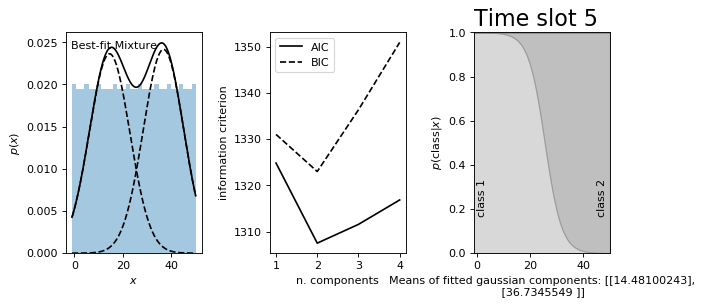
optimal number of gaussian models is found by plotting AIC and BIC scores against number of components

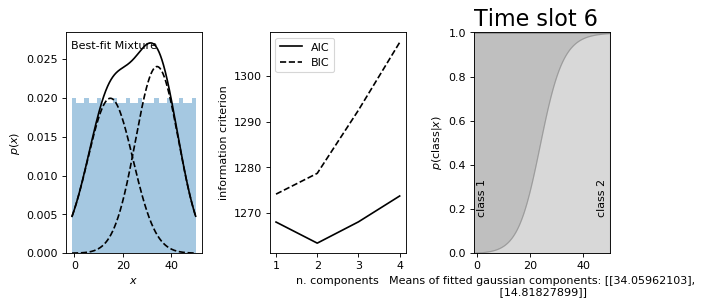
Posterior probabilities are also plotted for each component (class)



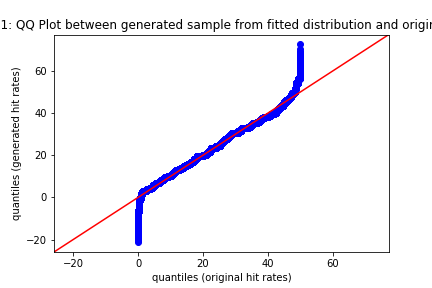


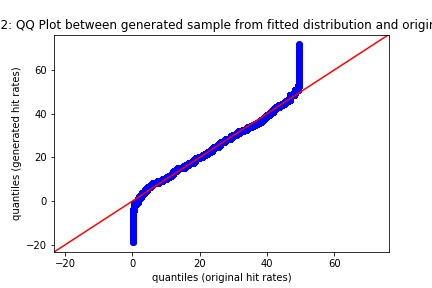


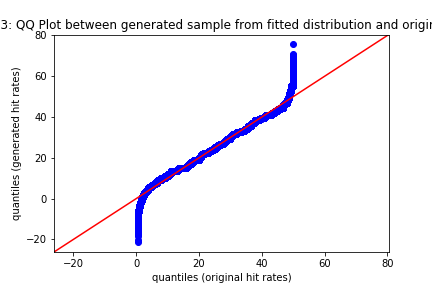


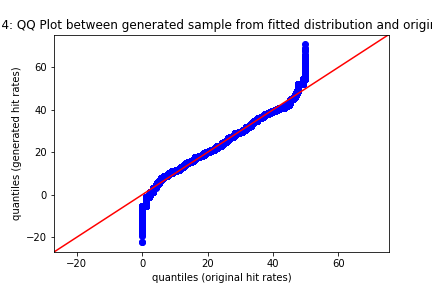


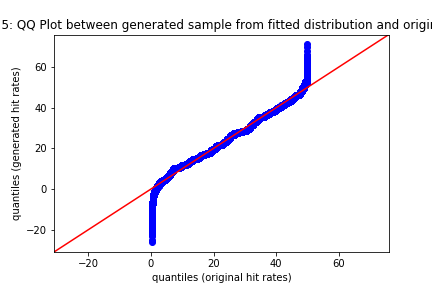
QQ Plots for testing goodness of fit-

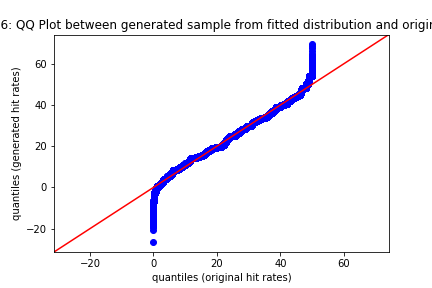












-Divyansh