A study the effect of competitive examination on student behavior using statistical methods

Prakash S. Chougule
ps_chougule@rediffmail.com
Rajarshi Chhatrapati Shahu College, Kolhapur, Maharashtra

Suresh V. Patil
sureshpatil2920@gmail.com
Rajarshi Chhatrapati Shahu College, Kolhapur, Maharashtra

Tejaswi S. Kurane
tejustat@gmail.com
Rajarshi Chhatrapati Shahu College, Kolhapur, Maharashtra

Prasad V. Shinde
prasadshinde2598@gmail.com
Rajarshi Chhatrapati Shahu College, Kolhapur, Maharashtra

Pratiksha S. Powar
pratikshapowar10@gmail.com
Rajarshi Chhatrapati Shahu College, Kolhapur, Maharashtra

ABSTRACT

We know that there are number of entrance exams like IIT JEE, Civil Services Exam, EAMCET, CAT, AIPMT, UPSC etc. An interesting fact about civil services exam is that it is considered as one of the toughest examination in the world with a success rate is below three percent. Parents who failed to meet their goals in their younger days want to fill that avoid through their child. But mechanism of these exams has both positive and negative effects on the takers. So for the purpose of this study we collect the primary data through systematic Questionaires containing several attributes and analysis is made through using various statistical tools and techniques. Our study shows the attributes stress and anxiety is dependent to each other. The proportion of arts students in competitive exam is forty percent. Most of students preparing for competitive exams from arts and science.

Keywords: Graphical Representation, Z-Test, Chi-Square Test, Level of Significance, ANOVA

1. INTRODUCTION

Today in a world of globalization and industrialization, it is very easy to earn money but it is very hard to earn prestige and respect in our society. The era has passed away when a simple graduation certificate used to help you in getting a prestigious job. Hence, the jobs are not available for large number of people. So the crowd is approaching towards competitive exams. Considering Maharashtra state, there are more than 4–5 lacks students appearing for competitive exams. The competitive exams include MPS, and other exams at state level, and the exams like UPSC, IBPS, SSC etc. at central level. For central level exams the number of competitors is increasing. At first it seems like quite acceptable. But as a statistician, the data regarding these aspirants dip into reality. The following project, consisting analysis of data of these aspirants provides us another point of view. This makes us think over the psychological problems faced by these aspirants. A large number of competitive exams aspirants face problems like depression, anxiety, stress. When we speak about psychological problems, it seems to be in two perspective. One is You and your competitors; Aspirants may go in depression or may lose their confidence as there is very high number of competition. Other is You and your family responsibilities or expectations. Some people are trying their luck in civil service exams even at the age of 30. At this level, family expectations can put them into a do or die situation. This study provides analysis of these aspirants from different angles by means of statistics. This analysis may be useful for government, coaching classes, aspirants and their family. We can evaluate or compute the stress or anxiety or depression of the students which are preparing for those exams by using different tests. We can see economical conditions of the students which are going to try competitive exam. We can correlate parents education with students education means we can say that those two things are independent or not independent. Similarly we can find the relation between study hours of students and their success in competitive exam. Our main aim is to examine the stress, depression and anxiety on the students.

2. OBJECTIVES

- To study which type of people (age wise/ gender wise/ area wise) preparing for competitive exams
- To study which source is mostly used competitive exam students.
- To study relation between previous marks of students and their parents education
- To study how many hours students study daily.
To study how many hours social media is used by aspirants.
To study how many years it takes to prepare for the competitive exams.
To study economical and health problem of students.
To study stress, depression and anxiety of the students.
To study an independence between different attributes

3. METHODOLOGY
For collection of primary data we prepare a systematic questionnaire and the questionnaire includes 49 questions regarding with some use of social media, study source, etc. For this study we collect the information from 105 students from some classes, colleges and study center from Kolhapur Districts, Arun Narke Foundation, A B Foundation, Study center and Shivaji university library.

4. METHOD OF DATA COLLECTION
For the project work, we have collected primary data from 105 students which is our sample size for area. The convenience (Non random) sampling method is used for collecting data from the study centers by using questionnaire method.

Statistical tools used:
- Testing of hypothesis
- Graphical Representation
- Measures of Central Tendency and dispersion

Software used:
- MS-Excel
- Microsoft Word
5. GRAPHICAL REPRESENTATION

<table>
<thead>
<tr>
<th>Gender</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of candidates</td>
<td>65</td>
<td>40</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Married</th>
<th>Unmarried</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>2</td>
<td>63</td>
</tr>
<tr>
<td>Female</td>
<td>9</td>
<td>31</td>
</tr>
<tr>
<td>Total</td>
<td>11</td>
<td>94</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Area</th>
<th>Rural</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of candidates</td>
<td>19</td>
<td>86</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Branch</th>
<th>Arts</th>
<th>Commerce</th>
<th>Science</th>
<th>Engineering</th>
<th>Other</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of candidates</td>
<td>35</td>
<td>6</td>
<td>39</td>
<td>20</td>
<td>5</td>
<td>105</td>
</tr>
</tbody>
</table>

Branchwise Distribution

<table>
<thead>
<tr>
<th>Branch</th>
<th>Number of students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arts</td>
<td>35</td>
</tr>
<tr>
<td>Commerce</td>
<td>6</td>
</tr>
<tr>
<td>Science</td>
<td>39</td>
</tr>
<tr>
<td>Engineering</td>
<td>20</td>
</tr>
<tr>
<td>Other</td>
<td>5</td>
</tr>
<tr>
<td>Parents Occupation</td>
<td>Business</td>
</tr>
<tr>
<td>--------------------</td>
<td>----------</td>
</tr>
<tr>
<td>Frequency</td>
<td>10</td>
</tr>
</tbody>
</table>

**Testing of Hypothesis:**

a) Test for Independence between the study hours and success:

Hypothesis are stated as

H₀: The study hours and success are independent.
H1: The study hours and success are dependent.

Observation Table:

<table>
<thead>
<tr>
<th>Study hours/Success</th>
<th>No</th>
<th>Preliminary</th>
<th>Mains</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-2</td>
<td>1</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>3-4</td>
<td>13</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>5 &amp; more</td>
<td>24</td>
<td>28</td>
<td>22</td>
</tr>
</tbody>
</table>

Notations: 
N= Population size =105

Test Statistics: Under H0 the test statistics,

\[ \chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = 13.9074 \]

Critical value: \[ \chi_{table}^2 = \chi_{(r-1)(c-1)}^2 = 9.4877 \] at 5% level of significance

Therefore, \[ \chi_{cal}^2 > \chi_{table}^2 \]

b) Test for Independence between the stress and anxiety:-

Hypothesis are stated as:

H0: The stress and anxiety are independent.

H1: The stress and anxiety are dependent.

Observation Table:

<table>
<thead>
<tr>
<th>Stress/Anxiety</th>
<th>NO</th>
<th>YES</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO</td>
<td>42</td>
<td>15</td>
</tr>
<tr>
<td>YES</td>
<td>26</td>
<td>22</td>
</tr>
</tbody>
</table>

Notations: 
N= Population size =105

Test Statistics: Under H0 the test statistics,

\[ \chi^2 = \frac{N(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)} = 4.350 \]

Critical value: \[ \chi_{table}^2 = \chi_{(r-1)(s-1)}^2 = 3.841 \] at 5% level of significance

Therefore, \[ \chi_{cal}^2 > \chi_{table}^2 \]

c) Test for Independence between the parent’s education and students marks:-

Hypothesis are stated as:

H0: The parent’s education and students marks are independent.

H1: The parent’s education and students marks are dependent.

Observation Table:

<table>
<thead>
<tr>
<th>Students marks/Parents education</th>
<th>Below 50</th>
<th>50-60</th>
<th>60-75</th>
<th>75-90</th>
<th>Above 90</th>
</tr>
</thead>
<tbody>
<tr>
<td>Illiterate</td>
<td>2</td>
<td>0</td>
<td>4</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>SSC</td>
<td>0</td>
<td>5</td>
<td>14</td>
<td>14</td>
<td>2</td>
</tr>
<tr>
<td>HSC</td>
<td>0</td>
<td>1</td>
<td>11</td>
<td>14</td>
<td>2</td>
</tr>
<tr>
<td>Graduate</td>
<td>0</td>
<td>1</td>
<td>7</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>Post-Graduate</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>

Notations: 
N= Population size =105

Test Statistics: Under H0 the test statistics,

\[ \chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = 25.871 \]

Critical value:

\[ \chi_{table}^2 = \chi_{(r-1)(s-1)}^2 = 26.296 \] at 5% level of significance

Therefore, \[ \chi_{cal}^2 < \chi_{table}^2 \]

Also the p-value is greater than the significance level alpha=0.05.

d) Test for Independence between the branch and success:-

Hypothesis are stated as :

H0: The branch and success are independent.
H1: The branch and success are dependent.

Observation Table:

<table>
<thead>
<tr>
<th>Branch/Success</th>
<th>No Preliminary</th>
<th>Mains</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arts</td>
<td>9</td>
<td>13</td>
</tr>
<tr>
<td>commerce</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Science</td>
<td>15</td>
<td>18</td>
</tr>
<tr>
<td>Medical</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Engineering</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>Other</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Notations: N= Population size =105

Test Statistics: Under H0 the test statistics,

\[ \chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = 7.5611 \]

Critical value: \( \chi^2_{table} = \chi^2_{(r-1)(c-1)} = 18.307 \) at 5% level of significance

Therefore \( \chi^2_{cal} < \chi^2_{table} \)

6. ANALYSIS OF VARIANCE (ONE WAY & TWO WAY):

i) The Testing of equality average of Different branch wise students depression

Hypothesis are stated as:

\[ H_0: \text{There is no significant difference Branchwise students depression} \]

i.e. \( H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu_6 \)

Where \( \mu_1 \): Arts students, \( \mu_2 \): Commerce students, \( \mu_3 \): Science students,
\( \mu_4 \): Medical students, \( \mu_5 \): Engineering students, \( \mu_6 \): other branch students

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>D f</th>
<th>MS</th>
<th>F</th>
<th>P-value</th>
<th>F crit.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>106.45</td>
<td>18</td>
<td>5.913889</td>
<td>1.24684</td>
<td>0.290648</td>
<td>1.972966</td>
</tr>
<tr>
<td>Within Groups</td>
<td>137.55</td>
<td>29</td>
<td>4.743103</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>244</td>
<td>47</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Here cal. value of F < tab value of F at 5% level of significance.alsothe p-value is greater than the significance level alpha=0.05

ii) The Testing of equality average of Different age wise anxiety Hypothesis are stated as:

\[ H_0: \text{There is no significant difference between anxieties of students according to their age} \]

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>D f</th>
<th>MS</th>
<th>F</th>
<th>P-value</th>
<th>F crit.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>40.41709</td>
<td>13</td>
<td>3.109007</td>
<td>1.686005</td>
<td>0.0928</td>
<td>1.917361</td>
</tr>
<tr>
<td>Within Groups</td>
<td>94.04444</td>
<td>51</td>
<td>1.844009</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>134.4615</td>
<td>64</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Here Cal value of F < tab value of F then Accept Ho at 5% level of significance

As the computed p-value is greater than the significance level alpha=0.05

iii) The Testing of equality average of Different branch wise study hours and area wise study hours.

Hypothesis are stated as:

\[ H_{01}: \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = \alpha_6 \text{i.e. There is no significant difference Branch wise students study hours} \]
\[ H_{02}: \beta_1 = \beta_2 \text{i.e. There is no significant difference area wise students study hours} \]

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>D f</th>
<th>MS</th>
<th>F</th>
<th>P-value</th>
<th>F crit.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area</td>
<td>12.4286</td>
<td>1</td>
<td>12.4286</td>
<td>2.9446</td>
<td>0.14682</td>
<td>6.60789</td>
</tr>
<tr>
<td>Branch</td>
<td>12.83947</td>
<td>5</td>
<td>2.567894</td>
<td>1.6439</td>
<td>0.70060</td>
<td>5.05032</td>
</tr>
<tr>
<td>Error</td>
<td>21.10418</td>
<td>5</td>
<td>4.220835</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>46.37224</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Here Both Cal value of F < tab value of F at 5% level of significance

f) F-test for equality of two variance

a) Test for equality of variance of SSC marks and HSC marks

Hypothesis is are stated as \( H_0: \sigma_1 = \sigma_2 \text{Vs} H_1: \sigma_1 \neq \sigma_2 \)

Observation table:
Here calculated value of F > tabulated value of F

b) Test for equality of variance of UG marks of male and female candidates

Hypothesis is are stated as 

\[ H_0: \sigma_1 = \sigma_2 \]
\[ H_1: \sigma_1 \neq \sigma_2 \]

Observation table:

<table>
<thead>
<tr>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>3.138462</td>
</tr>
<tr>
<td>Variance</td>
<td>0.558654</td>
</tr>
<tr>
<td>Observations</td>
<td>65</td>
</tr>
<tr>
<td>DF</td>
<td>64</td>
</tr>
<tr>
<td>Calculated F</td>
<td>1.613889</td>
</tr>
<tr>
<td>P value</td>
<td>0.055213</td>
</tr>
<tr>
<td>F Critical</td>
<td>1.6379</td>
</tr>
</tbody>
</table>

Here calculated value of F < tabulated value of F. As the computed p-value is greater than the significance level alpha=0.05.

g) Z-test for equality of two means

i) Test for equality of study hours of male & female students

\[ H_0: \text{There is no significant difference between study hours of male and female} \]
\[ i.e. H_0: \mu_1 = \mu_2 \frac{V}{S} \]

\[ H_1: \text{There is significant difference between study hours of male and female} \]
\[ i.e. H_1: \mu_1 \neq \mu_2 \]

Test Statistics: Under \( H_0 \) the test statistics,

\[ |Z_0| = \frac{x_1 - x_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = 1.951 \]

<table>
<thead>
<tr>
<th>Mean</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5.184615</td>
<td>4.675</td>
</tr>
<tr>
<td></td>
<td>1.965917</td>
<td>1.519375</td>
</tr>
<tr>
<td></td>
<td>65</td>
<td>40</td>
</tr>
<tr>
<td>Z calculated</td>
<td>1.951</td>
<td></td>
</tr>
<tr>
<td>Ptwo-tail</td>
<td>0.051057</td>
<td></td>
</tr>
<tr>
<td>z Critical two-tail</td>
<td>1.96</td>
<td></td>
</tr>
</tbody>
</table>

Here calculated value of \(|Z|< tabulated value of Z\). As the computed p-value is greater than the significance level alpha=0.05.

ii) Test for equality of mean of depression of male & female students

\[ H_0: \text{There is no significant difference between mean of depression of male and female students} \]
\[ i.e. H_0: \mu_1 = \mu_2 \frac{V}{S} \]

\[ H_1: \text{There is significant difference between mean of depression of male and female students} \]
\[ i.e. H_1: \mu_1 \neq \mu_2 \]

Test Statistics: Under \( H_0 \) the test statistics,

\[ |Z_0| = \frac{x_1 - x_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = 2.8412 \]
Here calculated value of $|Z| >$ tab. value of Z. As the computed p-value is less than the significance level $\alpha = 0.05$

b) Test for Equality of Proportion of male and female candidates has stress:

$n_1$: Sample of male candidates = 65 and
$n_2$: Sample of female candidates = 40

$p_1$: Sample proportion of male students has stress $= \frac{x_1}{n_1} = 0.3846$

$p_2$: Sample proportion of female students has stress $= \frac{x_2}{n_2} = 0.575$

Level of significance, $\alpha = 5\%$

Hypothesis:

$H_0$: $P_1 = P_2$  
$H_1$: $P_1 < P_2$

Calculation:

$\hat{P} = \frac{x_1}{n_1} = 0.3846$ and $\hat{Q} = 1 - \hat{P} = 0.5428$

Under $H_0$ the test statistic is,

$$Z_0 = \frac{P_1 - P_2}{\sqrt{\hat{P}\hat{Q}(\frac{1}{n_1} + \frac{1}{n_2})}} = \frac{0.3846 - 0.575}{\sqrt{(0.3846\cdot0.5428)(0.04038)}} = -1.9022$$

Critical value = $Z_{\alpha} = -1.64$

Calculated value of $Z_0 <$ critical value

i) Test for Proportion of arts students in competitive exams:

Let, $n$: Sample of competitive exams students = 105  
$x$: Number of arts students = 35

$p = \frac{x}{n} = 0.3334$

Level of significance, $\alpha = 5\%$

Hypothesis:

$H_0$: $P = 0.40$  
$H_1$: $P \neq 0.40$

$P_0 = 0.40$ and $Q_0 = 1 - P_0 = 1 - 0.40 = 0.60$

Under $H_0$ the test statistic is,

$$|Z_0| = \left| \frac{P - P_0}{\sqrt{\frac{P_0Q_0}{n}}} \right| = 1.3930$$

Critical value = $Z_{\alpha/2} = 1.96$

Calculated value of $|Z_0| <$ critical value

j) Test for Equality of Proportion of male and female students has anxiety:

Let, $n_1$: Sample of male students = 65 and $n_2$: Sample of female students = 40  
$x_1$: Number of male student has anxiety $= 24$ and

$x_2$: Number of female students has anxiety $= 13$

$p_1$: Sample proportion of male student has anxiety $= \frac{x_1}{n_1} = 0.3692$

$p_2$: Sample proportion of female students has anxiety $= \frac{x_2}{n_2} = 0.325$

Level of significance, $\alpha = 5\%$

Hypothesis:

$H_0$: $P_1 = P_2$  
$H_1$: $P_1 \neq P_2$

$$\hat{P} = \frac{n_1p_1 + n_2p_2}{n_1 + n_2} = 0.3523 \text{ and } \hat{Q} = 1 - \hat{P} = 0.6476$$

Under $H_0$ the test statistic is,
\[ |Z_0| = \frac{p_1 - p_2}{\sqrt{\hat{p}(\frac{1}{n_1} + \frac{1}{n_2})}} = 0.4604 \]

Critical value \( Z_{\alpha/2} = 1.96 \)

Calculated value of \( |Z_0| \) < critical value

6. MAJOR FINDING

- The maximum students parents education is SSC.
- Maximum students are uses Book for preparation of exams. Maximum students are read loksatta, pudhari and Maharashtra times newspaper.
- The most of students preparing for competitive exams from arts and science stream.
- The most of students have headache and eye problem.
- The most of students parents occupation is Farming.
- 14% students has job.
- The age wise Sample is random.
- The parent’s education and students marks are independent.
- The success is dependent on how many hour study daily.
- The stress and anxiety is dependent to each other.
- The Branch and success is independent.
- There is no significant difference between branch wise as well as age wise depression of students.
- There is no significant difference between branch wise and area wise study hours of students.
- The variation between SSC marks and HSC marks of students is in significant.
- The variation between UG marks of male and female candidates has no significance.
- There is no significance between study hours of male and female candidates.
- The proportion of male has stress is less than proportion of female has stress.
- The proportion of male has anxiety and proportion of female has anxiety has no significance.
- The proportion of arts students in competitive exam is 0.40.

7. REFERENCES