SYSC 4005 L2

Discrete Simulation/Modelling

Winter 2021

Milestone #2

Group Members: Julian Mendoza (101067270)

Hamna Nimra Manzoor (101057437)

Colin Wallace (100872525)

Due: March 14th, 2021

**TABLE OF CONTENTS**

**Deliverable 2 3-20**

Sample Calculations 3-5

Input Modelling 5-19

* Inspector 1 – component 1 5-7
* Inspector 2 – component 2 8-10
* Inspector 2 – component 3 10-12
* Workstation 1 12-14
* workstation 2 14-16
* Workstation 3 17-19

data input generation 19-20

**Sample Calculations**

Since we repeat the process seven times, we will show sample calculations to determine the data generated for each table. These sample calculations are specifically for the **first or second row of table 1**.

**Sample Mean ()**

Where is a data point in the sample

Where N is the sample size

**Bin Size:**

(Applies to all models)

**Frequency:**

**Bin Low:**

Ex. Bin Low bin=2

**Bin High:**

**Occurrences (Oi)**

**Cumulative Density function (CDF)**

In this report, we noticed that the histograms looked like exponential distributions.

We will explain later in this report but, we can use our estimator

The CDF low corresponds to the CDF of the lower bin and the CDF high corresponds to the CDF of the higher bin.

Ex. CDF low

**Probability (Pi)**

We use Pi to denote the probability that the distribution falls under the two bin ranges (bin low and bin high)

**Expected Frequency (Ei)**

**Null hypothesis for chi-square test**

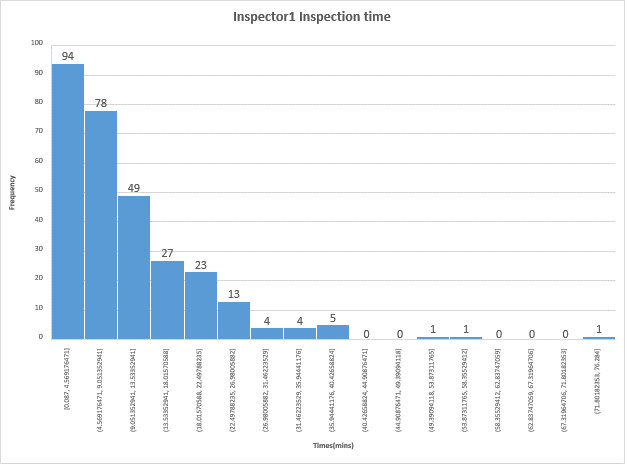
To compute the null hypothesis for chi-square, we take the partial value of each row:

Sample

**Input Modelling**

**Inspector 1 – Component 1**

Using the sample data, we computed a frequency of 4.4821. With a bin size of 17, we generated the following histogram:

*Figure 1: Histogram of Inspector 1's Inspection time*

From this histogram, we guessed that the distribution looks exponential.

We need to generate a Q-Q plot to verify this hypothesis.

CDF exponential:

We have to use an appropriate estimator for this distribution:

We can substitute with to get the inverse function:

We plot vs

We can determine the line of best fit using linear regression:

*Figure 2: Q-Q Plot of Yj and the inverse function*

Since the slope is approximately 1. The fit is appropriate.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Bin Low | Bin High | Oi | CDF low | CDF HIGH | Pi | Ei | (Oi-Ei)^2/Ei |
| 0.0870 | 4.5692 | 94 | 0.008364 | 0.35669 | 0.348326 | 105 | 1.152381 |
| 4.5692 | 9.0514 | 78 | 0.35669 | 0.582662 | 0.225972 | 68 | 1.470588 |
| 9.0514 | 13.5335 | 49 | 0.582662 | 0.729258 | 0.146596 | 44 | 0.568182 |
| 13.5335 | 18.0157 | 27 | 0.729258 | 0.82436 | 0.095102 | 29 | 0.137931 |
| 18.0157 | 22.4979 | 23 | 0.82436 | 0.886056 | 0.061696 | 19 | 0.842105 |
| 22.4979 | 26.9801 | 13 | 0.886056 | 0.92608 | 0.040024 | 13 | 0 |
| 26.9801 | 31.4622 | 4 | 0.92608 | 0.952046 | 0.025965 | 8 | 2 |
| 31.4622 | 35.9444 | 4 | 0.952046 | 0.96889 | 0.016845 | 6 | 0.666667 |
| 35.9444 | 40.4266 | 5 | 0.96889 | 0.979818 | 0.010928 | 4 | 0.25 |
| 40.4266 | 44.9088 | 0 | 0.979818 | 0.986907 | 0.007089 | 3 | 3 |
| 44.9088 | 49.3909 | 0 | 0.986907 | 0.991506 | 0.004599 | 2 | 2 |
| 49.3909 | 53.8731 | 1 | 0.991506 | 0.99449 | 0.002984 | 1 | 0 |
| 53.8731 | 58.3553 | 1 | 0.99449 | 0.996425 | 0.001936 | 1 | 0 |
| 58.3553 | 62.8375 | 0 | 0.996425 | 0.997681 | 0.001256 | 1 | 1 |
| 62.8375 | 67.3196 | 0 | 0.997681 | 0.998496 | 0.000815 | 1 | 1 |
| 67.3196 | 71.8018 | 0 | 0.998496 | 0.999024 | 0.000528 | 1 | 1 |
| 71.8018 | 76.2840 | 1 | 0.999024 | 0.999367 | 0.000343 | 1 | 0 |

*Table 1: Required values to compute Chi-Squared test.*

**Chi Square test**

We know that K=17 since it is our bin size.

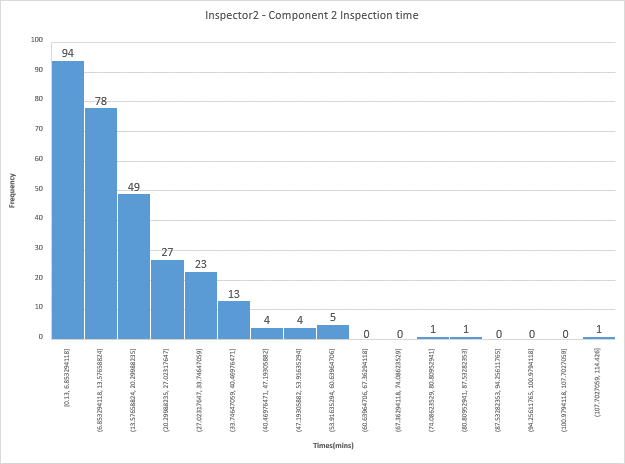
P=17-1-1 = 15 degrees of freedom

Using a significance level of 0.05, from the chi-squared table,

Since our test statistics from the observations is less than the chi-square value, our null hypothesis is accepted, and we can claim that the exponential distribution is a good fit.

**Inspector 2 - Component 2**

Using the sample data, we computed a frequency of 6.723294. With a bin size of 17, we generated the following histogram:

*Figure 3: Histogram of Inspector 2's Inspection time of component 2*

From this histogram, we guessed that the distribution looks exponential.

**Q-Q Plot**

Using the exact same process as in Inspector 1,

With

We plot vs

We can determine the line of best fit using linear regression:

*Figure 4: Q-Q Plot of Yj and the inverse function*

Since the slope is approximately 1. The fit is appropriate.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Bin Low | Bin High | Oi | CDF low | CDF HIGH | Pi | Ei | (Oi-Ei)^2/Ei |
| 0.1300 | 6.8533 | 94 | 0.008332 | 0.35667 | 0.348338 | 105 | 1.152381 |
| 6.8533 | 13.5766 | 78 | 0.35667 | 0.582649 | 0.225979 | 68 | 1.470588 |
| 13.5766 | 20.2999 | 49 | 0.582649 | 0.72925 | 0.146601 | 44 | 0.568182 |
| 20.2999 | 27.0232 | 27 | 0.72925 | 0.824355 | 0.095105 | 29 | 0.137931 |
| 27.0232 | 33.7465 | 23 | 0.824355 | 0.886053 | 0.061698 | 19 | 0.842105 |
| 33.7465 | 40.4698 | 13 | 0.886053 | 0.926078 | 0.040026 | 13 | 0 |
| 40.4698 | 47.1931 | 4 | 0.926078 | 0.952044 | 0.025966 | 8 | 2 |
| 47.1931 | 53.9164 | 4 | 0.952044 | 0.96889 | 0.016845 | 6 | 0.666667 |
| 53.9164 | 60.6396 | 5 | 0.96889 | 0.979818 | 0.010928 | 4 | 0.25 |
| 60.6396 | 67.3629 | 0 | 0.979818 | 0.986907 | 0.007089 | 3 | 3 |
| 67.3629 | 74.0862 | 0 | 0.986907 | 0.991506 | 0.004599 | 2 | 2 |
| 74.0862 | 80.8095 | 1 | 0.991506 | 0.99449 | 0.002984 | 1 | 0 |
| 80.8095 | 87.5328 | 1 | 0.99449 | 0.996425 | 0.001936 | 1 | 0 |
| 87.5328 | 94.2561 | 0 | 0.996425 | 0.997681 | 0.001256 | 1 | 1 |
| 94.2561 | 100.9794 | 0 | 0.997681 | 0.998496 | 0.000815 | 1 | 1 |
| 100.9794 | 107.7027 | 0 | 0.998496 | 0.999024 | 0.000528 | 1 | 1 |
| 107.7027 | 114.4260 | 1 | 0.999024 | 0.999367 | 0.000343 | 1 | 0 |

*Table 2: Required values to compute Chi-Squared test.*

**Chi Square test**

We know that K=17 since it is our bin size.

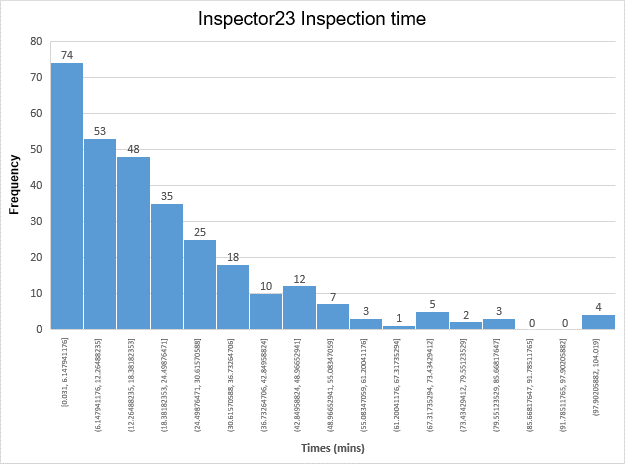
P=17-1-1 = 15 degrees of freedom

Using a significance level of 0.05, from the chi-squared table,

Since our test statistics from the observations is less than the chi-square value, our null hypothesis is accepted, and we can claim that the exponential distribution is a good fit.

**Inspector 2 – Component 3**

Using the sample data, we computed a frequency of 6.11694. With a bin size of 17, we generated the following histogram:

*Figure 5: Histogram of Inspector 2's Inspection time of component 3*

From this histogram, we guessed that the distribution looks exponential.

**Q-Q Plot**

Using the exact same process as in Inspector 1,

With

We plot vs

We can determine the line of best fit using linear regression:

*Figure 6: Q-Q Plot of Yj and the inverse function*

Since the slope is approximately 1. The fit is appropriate.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Bin Low | Bin High | Oi | CDF low | CDF HIGH | Pi | Ei | (Oi-Ei)^2/Ei |
| 0.0310 | 6.1479 | 74 | 0.001501 | 0.257676 | 0.256175 | 77 | 0.116883 |
| 6.1479 | 12.2649 | 53 | 0.257676 | 0.448127 | 0.190451 | 58 | 0.431034 |
| 12.2649 | 18.3818 | 48 | 0.448127 | 0.589716 | 0.141589 | 43 | 0.581395 |
| 18.3818 | 24.4988 | 35 | 0.589716 | 0.694978 | 0.105263 | 32 | 0.28125 |
| 24.4988 | 30.6157 | 25 | 0.694978 | 0.773235 | 0.078256 | 24 | 0.041667 |
| 30.6157 | 36.7326 | 18 | 0.773235 | 0.831414 | 0.058179 | 18 | 0 |
| 36.7326 | 42.8496 | 10 | 0.831414 | 0.874666 | 0.043253 | 13 | 0.692308 |
| 42.8496 | 48.9665 | 12 | 0.874666 | 0.906822 | 0.032156 | 10 | 0.4 |
| 48.9665 | 55.0835 | 7 | 0.906822 | 0.930728 | 0.023906 | 8 | 0.125 |
| 55.0835 | 61.2004 | 3 | 0.930728 | 0.9485 | 0.017773 | 6 | 1.5 |
| 61.2004 | 67.3174 | 1 | 0.9485 | 0.961713 | 0.013213 | 4 | 2.25 |
| 67.3174 | 73.4343 | 5 | 0.961713 | 0.971536 | 0.009823 | 3 | 1.333333 |
| 73.4343 | 79.5512 | 2 | 0.971536 | 0.978839 | 0.007303 | 3 | 0.333333 |
| 79.5512 | 85.6682 | 3 | 0.978839 | 0.984268 | 0.005429 | 2 | 0.5 |
| 85.6682 | 91.7851 | 0 | 0.984268 | 0.988304 | 0.004036 | 2 | 2 |
| 91.7851 | 97.9021 | 0 | 0.988304 | 0.991305 | 0.003001 | 1 | 1 |
| 97.9021 | 104.0190 | 4 | 0.991305 | 0.993536 | 0.002231 | 1 | 9 |

*Table 3: Required values to compute Chi-Squared test.*

**Chi Square test**

We know that K=17 since it is our bin size.

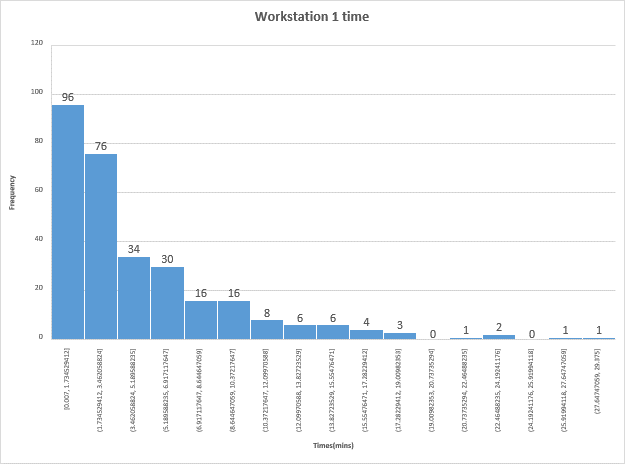
P=17-1-1 = 15 degrees of freedom

Using a significance level of 0.05, from the chi-squared table,

Since our test statistics from the observations is less than the chi-square value, our null hypothesis is accepted, and we can claim that the exponential distribution is a good fit for Inspector 2 component 3.

**Workstation 1**

Using the sample data, we computed a frequency of 1.727529. With a bin size of 17, we generated the following histogram:

*Figure 7: Histogram of Workstation 1's Production time*

**Q-Q Plot**

Using the exact same process as in Inspector 1,

With

We plot vs

We can determine the line of best fit using linear regression:

*Figure 8: Q-Q Plot of Yj and the inverse function*

Since the slope is approximately 1. The fit is appropriate.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Bin Low | Bin High | Oi | CDF low | CDF HIGH | Pi | Ei | (Oi-Ei)^2/Ei |
| 0.0070 | 1.7345 | 96 | 0.001519 | 0.313885 | 0.312366 | 94 | 0.042553 |
| 1.7345 | 3.4621 | 76 | 0.313885 | 0.52853 | 0.214645 | 65 | 1.861538 |
| 3.4621 | 5.1896 | 34 | 0.52853 | 0.676025 | 0.147495 | 45 | 2.688889 |
| 5.1896 | 6.9171 | 30 | 0.676025 | 0.777378 | 0.101353 | 31 | 0.032258 |
| 6.9171 | 8.6446 | 16 | 0.777378 | 0.847023 | 0.069645 | 21 | 1.190476 |
| 8.6446 | 10.3722 | 16 | 0.847023 | 0.894881 | 0.047857 | 15 | 0.066667 |
| 10.3722 | 12.0997 | 8 | 0.894881 | 0.927766 | 0.032886 | 10 | 0.4 |
| 12.0997 | 13.8272 | 6 | 0.927766 | 0.950364 | 0.022598 | 7 | 0.142857 |
| 13.8272 | 15.5548 | 6 | 0.950364 | 0.965892 | 0.015528 | 5 | 0.2 |
| 15.5548 | 17.2823 | 4 | 0.965892 | 0.976562 | 0.01067 | 4 | 0 |
| 17.2823 | 19.0098 | 3 | 0.976562 | 0.983895 | 0.007332 | 3 | 0 |
| 19.0098 | 20.7374 | 0 | 0.983895 | 0.988933 | 0.005038 | 2 | 2 |
| 20.7374 | 22.4649 | 1 | 0.988933 | 0.992395 | 0.003462 | 2 | 0.5 |
| 22.4649 | 24.1924 | 2 | 0.992395 | 0.994774 | 0.002379 | 1 | 1 |
| 24.1924 | 25.9199 | 0 | 0.994774 | 0.996409 | 0.001635 | 1 | 1 |
| 25.9199 | 27.6475 | 1 | 0.996409 | 0.997533 | 0.001123 | 1 | 0 |
| 27.6475 | 29.3750 | 1 | 0.997533 | 0.998304 | 0.000772 | 1 | 0 |

*Table 4: Required values to compute Chi-Squared test.*

**Chi Square test**

We know that K=17 since it is our bin size.

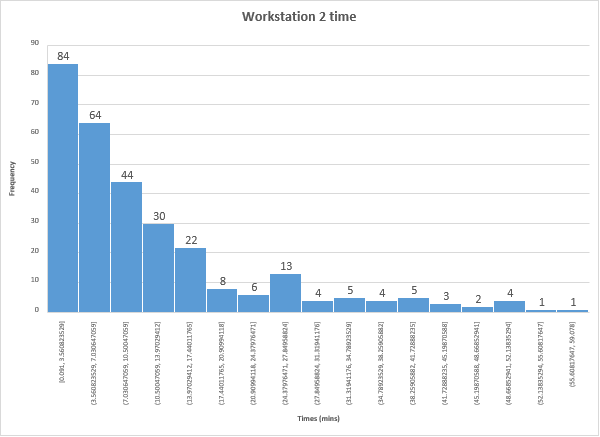
P=17-1-1 = 15 degrees of freedom

Using a significance level of 0.05, from the chi-squared table,

Since our test statistics from the observations is less than the chi-square value, our null hypothesis is accepted, and we can claim that the exponential distribution is a good fit for workstation 1.

**Workstation 2**

Using the sample data, we computed a frequency of 3.4698. With a bin size of 17, we generated the following histogram:

*Figure 9: Histogram of Workstation 2's production time*

From this histogram, we guessed that the distribution looks exponential.

**Q-Q Plot**

Using the exact same process as in Inspector 1,

With

We plot vs

We can determine the line of best fit using linear regression:

*Figure 10: Q-Q Plot of Yj and the inverse function*

Since the slope is approximately 1. The fit is appropriate.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Bin Low | Bin High | Oi | CDF low | CDF HIGH | Pi | Ei | (Oi-Ei)^2/Ei |
| 0.0910 | 3.5608 | 84 | 0.00817 | 0.274583 | 0.266413 | 80 | 0.2 |
| 3.5608 | 7.0306 | 64 | 0.274583 | 0.469436 | 0.194852 | 59 | 0.423729 |
| 7.0306 | 10.5005 | 44 | 0.469436 | 0.611949 | 0.142514 | 43 | 0.023256 |
| 10.5005 | 13.9703 | 30 | 0.611949 | 0.716183 | 0.104233 | 32 | 0.125 |
| 13.9703 | 17.4401 | 22 | 0.716183 | 0.792418 | 0.076236 | 23 | 0.043478 |
| 17.4401 | 20.9099 | 8 | 0.792418 | 0.848176 | 0.055758 | 17 | 4.764706 |
| 20.9099 | 24.3798 | 6 | 0.848176 | 0.888957 | 0.040781 | 13 | 3.769231 |
| 24.3798 | 27.8496 | 13 | 0.888957 | 0.918784 | 0.029827 | 9 | 1.777778 |
| 27.8496 | 31.3194 | 4 | 0.918784 | 0.940599 | 0.021815 | 7 | 1.285714 |
| 31.3194 | 34.7892 | 5 | 0.940599 | 0.956555 | 0.015955 | 5 | 0 |
| 34.7892 | 38.2591 | 4 | 0.956555 | 0.968224 | 0.01167 | 4 | 0 |
| 38.2591 | 41.7289 | 5 | 0.968224 | 0.97676 | 0.008535 | 3 | 1.333333 |
| 41.7289 | 45.1987 | 3 | 0.97676 | 0.983002 | 0.006243 | 2 | 0.5 |
| 45.1987 | 48.6685 | 2 | 0.983002 | 0.987568 | 0.004566 | 2 | 0 |
| 48.6685 | 52.1384 | 4 | 0.987568 | 0.990907 | 0.003339 | 2 | 2 |
| 52.1384 | 55.6082 | 1 | 0.990907 | 0.99335 | 0.002442 | 1 | 0 |
| 55.6082 | 59.0780 | 1 | 0.99335 | 0.995136 | 0.001786 | 1 | 0 |

*Table 5: Required values to compute Chi-Squared test.*

**Chi Square test**

We know that K=17 since it is our bin size.

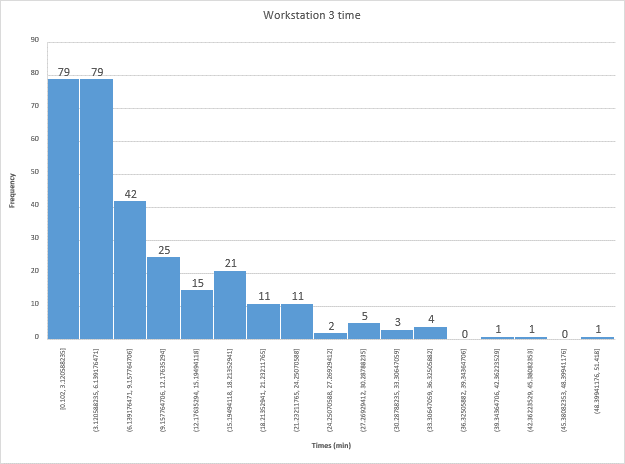
P=17-1-1 = 15 degrees of freedom

Using a significance level of 0.05, from the chi-squared table,

Since our test statistics from the observations is less than the chi-square value, our null hypothesis is accepted, and we can claim that the exponential distribution is a good fit for workstation 2.

**Workstation 3**

Using the sample data, we computed a frequency of 3.0185. With a bin size of 17, we generated the following histogram:

*Figure 11: Histogram of Workstation 3's production time*

From this histogram, we guessed that the distribution looks exponential.

**Q-Q Plot**

Using the exact same process as in Inspector 1,

With

We plot vs

We can determine the line of best fit using linear regression:

*Figure 12: Q-Q Plot of Yj and the inverse function*

Since the slope is approximately 1. The fit is appropriate.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Bin Low | Bin High | Oi | CDF low | CDF HIGH | Pi | Ei | (Oi-Ei)^2/Ei |
| 0.1020 | 3.1206 | 79 | 0.01153 | 0.29868 | 0.28715 | 87 | 0.735632 |
| 3.1206 | 6.1392 | 79 | 0.29868 | 0.502413 | 0.203733 | 62 | 4.66129 |
| 6.1392 | 9.1578 | 42 | 0.502413 | 0.646961 | 0.144549 | 44 | 0.090909 |
| 9.1578 | 12.1764 | 25 | 0.646961 | 0.749519 | 0.102557 | 31 | 1.16129 |
| 12.1764 | 15.1949 | 15 | 0.749519 | 0.822283 | 0.072765 | 22 | 2.227273 |
| 15.1949 | 18.2135 | 21 | 0.822283 | 0.87391 | 0.051627 | 16 | 1.5625 |
| 18.2135 | 21.2321 | 11 | 0.87391 | 0.910539 | 0.036629 | 11 | 0 |
| 21.2321 | 24.2507 | 11 | 0.910539 | 0.936527 | 0.025988 | 8 | 1.125 |
| 24.2507 | 27.2693 | 2 | 0.936527 | 0.954966 | 0.018439 | 6 | 2.666667 |
| 27.2693 | 30.2879 | 5 | 0.954966 | 0.968048 | 0.013082 | 4 | 0.25 |
| 30.2879 | 33.3065 | 3 | 0.968048 | 0.97733 | 0.009282 | 3 | 0 |
| 33.3065 | 36.3251 | 4 | 0.97733 | 0.983916 | 0.006586 | 2 | 2 |
| 36.3251 | 39.3436 | 0 | 0.983916 | 0.988588 | 0.004672 | 2 | 2 |
| 39.3436 | 42.3622 | 1 | 0.988588 | 0.991903 | 0.003315 | 1 | 0 |
| 42.3622 | 45.3808 | 1 | 0.991903 | 0.994255 | 0.002352 | 1 | 0 |
| 45.3808 | 48.3994 | 0 | 0.994255 | 0.995924 | 0.001669 | 1 | 1 |
| 48.3994 | 51.4180 | 1 | 0.990651 | 0.997108 | 0.006457 | 2 | 0.5 |

*Table 6: Required values to compute Chi-Squared test.*

**Chi Square test**

We know that K=17 since it is our bin size.

P=17-1-1 = 15 degrees of freedom

Using a significance level of 0.05, from the chi-squared table,

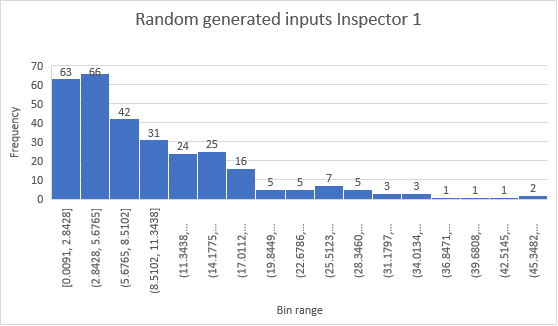
Since our test statistics from the observations is less than the chi-square value, our null hypothesis is accepted, and we can claim that the exponential distribution is a good fit for workstation 3.

**Data Input Generation**

Now that we know that our models can fit our hypothesis model, we can use these models to generate our input. To do this, we need to create a random number generator to produce random variates. We used Java’s Random library to generate a sequence of random numbers between the range 0 to 1 exclusively. These numbers will be used in our model’s respective inverse CDF to produce a random input.

To ensure that our random input stream is valid, we will use a frequency test between the random numbers generated to compare the distribution set of numbers generated to a uniform distribution.

Using the Chi-Square Test with sample size 300, we will validate the randomness in the number generation for seed 111 for inspector 1.



*Figure 13: Histogram of Randomly generated inputs for inspector 1*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Bin Low | Bin High | OI | CDF LOW | CDF HIGH | PI | EI | (Oi-Ei)^2/Ei |
| 0.009 | 2.842 | 63 | 0.000869 | 0.239957 | 0.239088 | 72 | 1.125 |
| 2.842 | 5.675 | 66 | 0.239957 | 0.421832 | 0.181875 | 55 | 2.2 |
| 5.675 | 8.508 | 42 | 0.421832 | 0.560185 | 0.138353 | 42 | 0 |
| 8.508 | 11.341 | 31 | 0.560185 | 0.665431 | 0.105246 | 32 | 0.03125 |
| 11.341 | 14.174 | 24 | 0.665431 | 0.745492 | 0.080061 | 25 | 0.04 |
| 14.174 | 17.007 | 25 | 0.745492 | 0.806395 | 0.060903 | 19 | 1.894737 |
| 17.007 | 19.84 | 16 | 0.806395 | 0.852724 | 0.046329 | 14 | 0.285714 |
| 19.84 | 22.673 | 5 | 0.852724 | 0.887966 | 0.035243 | 11 | 3.272727 |
| 22.673 | 25.506 | 5 | 0.887966 | 0.914775 | 0.026809 | 9 | 1.777778 |
| 25.506 | 28.339 | 7 | 0.914775 | 0.935169 | 0.020394 | 7 | 0 |
| 28.339 | 31.172 | 5 | 0.935169 | 0.950683 | 0.015514 | 5 | 0 |
| 31.172 | 34.005 | 3 | 0.950683 | 0.962484 | 0.011801 | 4 | 0.25 |
| 34.005 | 36.838 | 3 | 0.962484 | 0.971462 | 0.008977 | 3 | 0 |
| 36.838 | 39.671 | 1 | 0.971462 | 0.978291 | 0.006829 | 3 | 1.333333 |
| 39.671 | 42.504 | 1 | 0.978291 | 0.983486 | 0.005195 | 2 | 0.5 |
| 42.504 | 45.337 | 1 | 0.983486 | 0.987438 | 0.003952 | 2 | 0.5 |
| 45.337 | 48.17 | 2 | 0.987438 | 0.990444 | 0.003006 | 1 | 1 |

*Table 7: Required values to compute Chi-Squared test.*

Since our null hypothesis of random inputs is less than the chi-square value, we can accept the random number stream and say that the stream is acceptable.