Moments of Order Statistics on Exponential Distribution

Sujatha.V
Assistant Professor
School of Advanced sciences,
Department of Mathematics,
VIT University, Vellore.

Dharanidharan.V
Data Analyst
Ausdauer, Dubai, UAE

Abstract - The entire population of the samples are characterized by one or more parameter, sample function of the statistic is called as an Estimator. In many Engineering and in Medicine problems, the random samples are unspecified by the population and they are characterized by their moments. The knowledge of moments are used to find a symmetric population and their expected values of their order statistics. Exponential distribution which is used to model the failure times of the manufactured items in production, In this paper, we simulate the data for univariate exponential distribution and we explores the assumption of normality for the distributed random variables and their accurate results.

Keywords - Exponential distribution, moments, Order Statistics, population parameter, Random Variables, simulated data

I. INTRODUCTION

Order statistics are said to be the functions of the random variables, statistical inferences are directly based on the order statistics, and their likelihood functions are given by the joint distribution functions of the ordered samples. Let $X_{k,n}$ denote the $k^{th}$ smallest of $(X_1,X_2, \ldots, X_n)$, then $X_{k,n}$ is a function of the sample variables, and hence the statistic, is called the $k^{th}$ order statistic. Our goal is to study the distribution of the order statistics, their properties and their applications. The extreme order statistics are the minimum and maximum values, $X_{1:n} = \min\{X_1,X_2, \ldots, X_n\}$, and $X_{n:n} = \max\{X_1,X_2, \ldots, X_n\}$.

If $X$ has cumulative distribution function $F(x) = 1 - e^{-\lambda x}$, $x > 0$, $\lambda > 0$, then

$1 - F_{1:n}(x) = P(X_{1:n} \geq x) = e^{-\lambda x}$ and $F_{n:n}(x) = P(X_{n:n} \leq x) = (1 - e^{-\lambda x})^n$.

Reliability Function

The Reliability function at time ‘t’ for two parameter exponential distribution is

$$S(t) = 1 - F(t) = 1 - [1 - e^{-\frac{(x-\eta)}{\theta}}] = e^{-\frac{(x-\eta)}{\theta}}$$

Hazard Function

The hazard function for the Exp $(\theta,\eta)$ is

$$h(t) = \frac{f(x)}{S(x)} = \frac{f(x)}{1 - F(x)} = \frac{1}{\theta}$$

For the two parameter Exponential distribution, the Expectation of $X$ is
\[ E(x) = [\theta + \eta] \]
\[ E(x^2) = \eta^2 + 2\eta \theta + \theta^2 \]
\[ V(x) = \theta^2 \]

If \( x-\eta = Y \), then \( Y \sim \exp(\theta) \), there is no need of separate statistical procedures on \( \text{Exp} (\theta, \eta) \) model. When \( \eta = 0 \), in two parameter exponential distribution, the random variable \( X \sim \exp(\theta) \)

\[
f(x; \theta; \eta) = \begin{cases} 
1 - \frac{(x-\eta)}{\theta} & ; x > \eta, \theta > 0, \eta > 0 \\
0 & ; \text{otherwise}
\end{cases}
\]

II PRELIMINARIES

2.1 Joint density of two order statistics –

The joint density function of most extreme exponential order statistics \( X_{1:n}, X_{n:n} \) for the function \( f(x) = e^{-x} \) is

\[ F_{1:n,n}(x, y) = P \{ X_{1:n} \leq x, X_{n:n} \leq y \} \]

Let \( x_{r:n}, x_{s:n}, 1 \leq r < s \leq n \) be the \( r^{th} \) and \( s^{th} \) order statistics of a random sample of size \( n \) arising from the exponential distribution \( f(x) = e^{-x} \). Then the joint probability density function of \( x_{r:n} \) and \( x_{s:n} \) is given by

\[
f_{r,s:n}(x, y) = (1 - e^{-x})^{r-1} (e^{-x} - e^{-y})^{s-r-1} e^{-x} e^{-y} \]

where \( 0 < x < y < \infty \)

**Joint probability density function of the range**

The joint probability density function of the range \( R = x_{(n)} - x_{(1)} \) is given by

When \( s=n, r=1 \)

\[
f(t, u) = n(n-1)(e^{(u-t)} - e^{-u})^{n-2} e^{(2u-t)} , \quad 0 \leq t < u < \infty
\]

The marginal distribution of the density function is given by

\[
f_r(t) = \int_0^\infty f(t, u) \, du = (n-1)e^{(e^t-1)^{n-2}}e^{-nt}
\]

Hence,

\[
f(t) = (n-1)e^{t}, e^{\eta}(e^{t} - 1)^{n-2} = (n-1)e^{t}(1-e^{t})^{n-2} \quad \text{where} \quad 0 < t < \infty.
\]

It is easily seen that, \( R \) and \( X_{(1)} \) are independently distributed, the independence of \( R \) and \( X_{(1)} \) characterizes the exponential distribution[5].

The density function of \( u \) is, \( f(u) = (n-1+1)e^{-(n-1+1)u} \)

\[ f(u) = X_{(0)} - X_{(1-1)} : n \]

If we let \( y_j = (n-1+1) (x_j - x_{(1-1)}) \), where \( 1, 2, 3, \ldots n \), with \( x_{(0)} = 0 \)
\[ E(X_r) = E\left( \sum_{l=1}^{r} \frac{y_l}{n-l+1} \right) = \sum_{l=1}^{r} \frac{E(y_l)}{n-l+1} \]

\[ = \sum_{l=1}^{r} \frac{(n-l+1)}{(n-l+1)} = \sum_{l=1}^{r} \frac{1}{n-l+1} \]

\[ V(X_r) = V\left[ \sum_{l=1}^{r} \frac{y_l}{n-l+1} \right] = \sum_{l=1}^{r} \frac{1}{(n-l+1)^2} V(y_l) \]

where \( f(u) = (n-l+1)e^{-u(n-l+1)} \)

**If \( X_i \sim \exp(\theta) \), then, \( M_\theta(t) = (1 - \frac{t}{\theta})^{-r} \)**

**Point Estimation (complete Samples and no censoring)**

Suppose \( x_1, x_2, \ldots, x_n \) denote random sample of size \( n \) from \( \exp(\theta) \). So that the joint density function or likelihood function is\(^2\)

\[ L(x) = f(x_1, \theta), \ldots, f(x_n, \theta) \]

\[ = -n \log(\theta) - \frac{x_1}{\theta} - \frac{x_2}{\theta} - \ldots - \frac{x_n}{\theta} = \left( \frac{1}{\theta} \right)^n e^{-\frac{1}{\theta} \sum x_i}; 0 < x_i < \infty \]

Maximum Likelihood estimation of the parameter

Now,

\[ \log L = n \log \left( \frac{1}{\theta} \right) - \frac{1}{\theta} \sum x_i = n \left[ \log 1 - \log \theta \right] - \frac{1}{\theta} \sum x_i \]

\[ \frac{\partial \log L}{\partial \theta} = -n + \frac{1}{\theta^2} \sum x_i \]

To find M.L.E \( \frac{\partial L}{\partial \theta} = 0 \), \( \Rightarrow \frac{1}{\theta^2} \sum x_i = \theta = \frac{1}{n} \sum x_i = \bar{x} \)

therefore,

\[ \theta = \frac{1}{n} \sum x_i \] is the M.L.E of \( \theta \)

\[ \Lambda \]

**hence, \( w(\theta) \) is the M.L.E of \( w(\theta) \)**
The reliability estimate at time $t_0$ is

$$R(t_0) = e^{-\frac{t_0}{\theta}}$$

similarly, for the hazard function,

$$h(t) = \frac{1}{\lambda}$$

**Confidence Interval Estimation**

The goal of statistical inference is to obtain information pertaining to the population from sample information. Confidence intervals provide a range within which the value of the parameter could lie with a certain probability\(^{[9]}\). The information from sampling distribution are used to construct the confidence intervals. From the confidence limits one can obtain within which the value of the parameter could possibly lie.\(^{[1]}\)

$$P\left[\frac{2\sum x_i}{\theta} \sim \chi(2n)\right] = \alpha$$

Then

$$P\left[\frac{2}{\chi_{\alpha_2} (2n)} \leq \frac{2\sum x_i}{\theta} \leq \frac{2}{\chi_{1-\alpha_2} (2n)}\right] = 1 - \alpha$$

$$\left(\frac{2\sum x_i}{\chi_{1-\alpha_2} (2n)}, \frac{2\sum x_i}{\chi_{\alpha_1} (2n)}\right)$$ is the confidence interval for the parameter $\theta$. 

It is a $100(1-\alpha)\%$ two sided confidence interval for $\theta$ based on the statistic $\hat{\theta} = \frac{\sum x_i}{n}$

**Confidence Interval for Reliability**

Confidence intervals for monotonic functions of $\theta$ can be easily obtained from the confidence intervals for $\theta^{[4]}$

When $\alpha_1 = 0$, then

$$P\left[\frac{2n\bar{\theta}}{\chi^2_{1-\alpha}(2n)} \leq \theta \leq \infty\right] = 1-\alpha$$

$$= P[\theta_L \leq \theta] = P\left[\frac{1}{\theta_L} \geq \frac{1}{\theta}\right]$$

Since $f(x) = e^{-\frac{x}{\theta}}$, $P[R_{L} \leq R_{U}]$

i.e.

$$R_L = \exp\left[-t\frac{2}{\chi^2_{1-\alpha}(2n)}\right]$$ is a lower $1-\alpha$ confidence limit for the reliability at time $t^{[12]}$.

let $\alpha_2 = 0$, then,

$$1-\alpha = P\left[0 < \theta \leq \frac{2n\bar{\theta}}{\chi^2_{1-\alpha}(2n)}\right] = P\left[\theta \leq \theta_U\right]$$

$$= P\left[\frac{1}{\theta} \geq \frac{1}{\theta_U}\right] = P\left[\frac{-t}{e^{\theta}} \leq e^{-t}_{\theta_U}\right]$$

$$= P\left[R_{L(t)} \leq R_{U}\right]$$

i.e., $R_U = \exp\left[-t\frac{2}{\chi^2_{1-\alpha}(2n)}\right]$  

$R_U$ is a upper confidence limit for reliability, Two sided confidence limit for reliability$^{[13]}$
Confidence Interval for the estimate

We know that

\[
prob \left( \chi^2_{(2r), \frac{\alpha}{2}} \leq 2\lambda T \leq \chi^2_{(2r), 1-\frac{\alpha}{2}} \right) = 1 - \alpha
\]

Thus a two sided 1-\(\alpha\) confidence interval for the random variable 2\(\lambda T\) is

\[
prob \left( \chi^2_{(2r), \frac{\alpha}{2}} \leq 2\lambda T \leq \chi^2_{(2r), 1-\frac{\alpha}{2}} \right)
\]

III RESULTS AND DISCUSSIONS

Application of exponential order statistics for simulated data

Simulated data have been taken for exponential distribution with parameters \(\lambda=1\), for a random sample of size 20 and it is replicated for 12 times\(^\text{[7]}\). By using the simulated data the mean, variance, range, and midrange estimates are all estimated. The first four moments and the corresponding skewness and Kurtosis are also estimated so that we study the nature of the distribution\(^\text{[3]}\) Also their respective reliability estimates on 95% confidence limits were derived. To verify the statement on parameter \(\theta\) testing of hypothesis is carried out and the results are given at 5% level. The empirical estimates which we derived coincides with the theoretical estimates to a greater accuracy and they also become equal as the sample size goes to infinity\(^\text{[9]}\). These empirical estimates are presented in the following tables\(^\text{[10]}\).
Table 1  Sample statistics for simulated data from exponential distribution with $\lambda=1$

<table>
<thead>
<tr>
<th>mean</th>
<th>variance</th>
<th>True variance</th>
<th>True mean</th>
<th>est.mean</th>
<th>est.var</th>
<th>mu</th>
<th>mu 4</th>
<th>beta 1</th>
<th>beta2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0947</td>
<td>0.0043</td>
<td>0.0003</td>
<td>0.029</td>
<td>0.0067</td>
<td>0.0001</td>
<td>1.179</td>
<td>12.590</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0551</td>
<td>0.0095</td>
<td>0.0002</td>
<td>0.012</td>
<td>0.0018</td>
<td>0.0000</td>
<td>1.275</td>
<td>12.595</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.137</td>
<td>0.0112</td>
<td>0.0013</td>
<td>0.128</td>
<td>0.0021</td>
<td>0.0003</td>
<td>1.217</td>
<td>12.851</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0558</td>
<td>0.0104</td>
<td>0.0013</td>
<td>0.128</td>
<td>0.0017</td>
<td>0.0006</td>
<td>1.279</td>
<td>3.847</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0346</td>
<td>0.0142</td>
<td>0.0017</td>
<td>0.279</td>
<td>0.0133</td>
<td>0.0000</td>
<td>1.289</td>
<td>3.222</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0553</td>
<td>0.0131</td>
<td>0.0013</td>
<td>0.362</td>
<td>0.0201</td>
<td>0.0009</td>
<td>1.256</td>
<td>2.256</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0433</td>
<td>0.0128</td>
<td>0.0012</td>
<td>0.357</td>
<td>0.0251</td>
<td>0.0012</td>
<td>1.249</td>
<td>2.162</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.6059</td>
<td>0.0337</td>
<td>0.0431</td>
<td>0.495</td>
<td>0.0399</td>
<td>0.0017</td>
<td>0.951</td>
<td>2.129</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.3113</td>
<td>0.0317</td>
<td>0.0441</td>
<td>0.579</td>
<td>0.0393</td>
<td>0.0012</td>
<td>0.616</td>
<td>1.285</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.4092</td>
<td>0.0310</td>
<td>0.0434</td>
<td>0.683</td>
<td>0.0464</td>
<td>0.0017</td>
<td>0.755</td>
<td>1.427</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.7152</td>
<td>0.0346</td>
<td>0.0534</td>
<td>0.738</td>
<td>0.0534</td>
<td>0.0045</td>
<td>0.710</td>
<td>1.105</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.7948</td>
<td>0.0404</td>
<td>0.0717</td>
<td>0.599</td>
<td>0.0907</td>
<td>0.0064</td>
<td>0.676</td>
<td>1.295</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.9524</td>
<td>0.0522</td>
<td>0.0873</td>
<td>1.089</td>
<td>0.1072</td>
<td>0.0095</td>
<td>0.862</td>
<td>1.380</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.4930</td>
<td>0.0663</td>
<td>1.0177</td>
<td>1.232</td>
<td>0.0424</td>
<td>0.0149</td>
<td>0.991</td>
<td>1.522</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.3070</td>
<td>0.0567</td>
<td>1.1286</td>
<td>1.224</td>
<td>0.1323</td>
<td>0.0296</td>
<td>0.629</td>
<td>1.906</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.1764</td>
<td>0.1461</td>
<td>1.1720</td>
<td>1.544</td>
<td>0.1625</td>
<td>0.1374</td>
<td>0.845</td>
<td>1.906</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.8315</td>
<td>0.1991</td>
<td>1.2330</td>
<td>1.862</td>
<td>0.2310</td>
<td>0.2011</td>
<td>0.951</td>
<td>1.786</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.4238</td>
<td>0.1042</td>
<td>1.1872</td>
<td>2.372</td>
<td>0.2581</td>
<td>0.1975</td>
<td>0.940</td>
<td>1.115</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.7011</td>
<td>0.3620</td>
<td>0.5911</td>
<td>2.577</td>
<td>0.5951</td>
<td>0.5220</td>
<td>1.025</td>
<td>2.865</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.3359</td>
<td>0.0967</td>
<td>1.5911</td>
<td>2.977</td>
<td>1.5961</td>
<td>1.2551</td>
<td>1.328</td>
<td>1.2381</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

From the table 1, it is observed that the empirical data mean and the theoretical mean are nearly equal. Also the same holds good for the sample variance. The trend in Beta 1 and beta 2 values can be seen to decrease to some level of the order and starts increasing thereafter. The theoretical median, range and midrange for the above simulated data are 0.7188, 3.5477 and 1.8239 respectively. The 95% and 90% tolerance limits have been derived
and found to be 0.037593 and 0.077219 respectively. This indicate, when repeated samples are taken, in 90 out of 100 cases the true parameter will have value in the tolerance region\(^5\) with value 0.037593. From table 2, it is seen that nearly 86% of the objects survived beyond time 0.1688 units and nearly 12% of the objects survive beyond time 2.3 units. Also, the reliable estimates are derived and are exhibited with the corresponding 95% upper and lower limits\(^1\).

The parameter \(\theta\) is tested for the value of unity under the null hypothesis, with the test statistic value 1.746954, which is less than the corresponding the table value of 5.7585, leading to non-rejection of the null hypothesis\(^8\). Thus we conclude that the data comes from the exponential distribution with parameter one.

![Comparison of Theoretical mean and Sample mean for different iterations using exponential distribution of lambda=1, n=10](image1)

![Comparison of Theoretical Variance and Sample Variance for different iterations using exponential distribution of lambda=1](image2)

### References


4. ethesis.nitrkl.ac.in (online)


