مجلّه‌ی پژوهش‌های آماری ایران

J. Statist. Res. Iran **14** (2017): 171–188

دوره‌ی ۱۴، شماره‌ی ۲، پاییز و زمستان ۱۳۹۶، صص ۱۷۱–۱۸۸

پژوهشکده‌ی آمار

DOI: 171.2.14.jsri/29252.10

Rayleigh Confidence Regions based on Record Data

Mousa Abdi[†,*] and Akbar Asgharzadeh[‡]

[†] Higher Education Complex of Bam
[‡] University of Mazandaran

Received: 8/8/2016 Approved: 10/16/2017

Abstract. This paper presents exact joint confidence regions for the parameters of the Rayleigh distribution based on record data. By providing some appropriate pivotal quantities, we construct several joint confidence regions for the Rayleigh parameters. These joint confidence regions are useful for constructing confidence regions for functions of the unknown parameters. Applications of the joint confidence regions using two environmental data sets are presented for illustrative purposes. Finally, a simulation study is conducted to study the performance of the proposed joint confidence regions.

Keywords. Joint confidence region; pivotal quantity; Rayleigh distribution; records.

MSC 2010: 62F25, 62E15.

1 Introduction

The Rayleigh distribution was proposed by Rayleigh (1880) in connection with a problem in the field of acoustics. A random variable X is said to have the Rayleigh distribution, if its probability density function (pdf) is given by

$$f(x;\mu,\lambda)=2\lambda(x-\mu)e^{-\lambda(x-\mu)^2},\qquad x>\mu,\quad \mu\in R,\quad \lambda>0,\qquad)(1$$

[*] Corresponding author

where μ and λ are the location and scale parameters, respectively. The Rayleigh cumulative distribution function (cdf) is

$$F(x;\mu,\lambda) = 1 - e^{-\lambda(x-\mu)^2}, \qquad x > \mu, \quad \mu \in R, \quad \lambda > 0. \tag{2}$$

An important characteristic of the Rayleigh distribution is that its hazard function is an increasing function of time. Therefore, this distribution can be used as a lifetime model in reliability and life testing studies. On the other hand, since the Rayleigh distribution has linearly increasing failure rate, it is an appropriate distribution for a study of components which may not have any manufacturing defects but age rapidly with time. The origin and other applications of this distribution are outlined by Johnson et al. (1994). Some recent references on the Rayleigh distribution are Raqab and Madi (2002), Wu et al. (2006), Soliman and Al-Aboud (2008), Khan et al. (2010), Dey et al. (2014) and Asgharzadeh and Azizpour (2016).

In many real life applications, only observations that exceed or only those that fall below the current extreme value are recorded and the complete data are not available. For example, an electronic component ceases to function in an environment of too high temperature, a battery dies under the stress of time and a wooden beam breaks when sufficient perpendicular force is applied to it (see Ahmadi and Arghami, 2003). Record data are commonly observed in many real life applications involving data relating to weather, sports, economics, environmental studies and life-tests. The statistical study of record values was started with Chandler (1952). Since then, extensive work has done on statistical inference for record data. For more details, see Nevzorov (1988), Ahsanullah (1995) and Arnold et al. (1998). For some recent references, see Asgharzadeh et al. (2013), Mirfarah and Ahmadi (2014) and Zakerzadeh et al. (2016).

Let $\{X_n, n \geqslant 1\}$ be a sequence of independent and identically distributed (iid) random variables with cdf $F(x)$ and pdf $f(x)$. We say an observation X_j is an upper record value of this sequence if its value exceeds all previous observations. Thus, X_j is an upper record if $X_j > X_i$ for all $i < j$. If $\{T_n, n \geqslant 1\}$ is defined as: $T_1 = 1$ and $T_n = \min\{j : j > T_{n-1}, X_j > X_{T_{n-1}}\}$ for $n \geqslant 2$, then the sequence $\{U_n = X_{T_n}, n \geqslant 1\}$ provides a sequence of upper record values from the original sequence $\{X_n, n \geqslant 1\}$. The sequence $\{T_n, n \geqslant 1\}$ represents the record times.

The construction of joint confidence regions for unknown parameters is important in statistical inference. Joint confidence regions are useful for constructing confidence bounds for any function of the unknown parameters as well as for testing hypotheses related to unknown parameters. Recently, some authors have discussed the construction of joint confidence regions of unknown parameters based on record data. See, for example, Asgharzadeh and Abdi (2011), Asgharzadeh et al. (2015), Asgharzadeh et al. (2016) and Kinanci et al. (2017). This paper considers joint confidence regions for two-parameter Rayleigh distribution based on record data. By providing some appropriate pivotal quantities, several joint confidence regions for the Rayleigh parameters are first constructed. Optimal joint confidence regions are then proposed.

The rest of the paper is organized as follows. In Section 2, we provide several joint confidence regions for the parameters μ and λ. In Section 3, we present illustrative examples using two environmental data sets. A Monte Carlo simulation study is given in Section 4 to compare different proposed joint confidence regions.

2 Joint Confidence Regions

In order to construct the joint confidence regions for the parameters μ and λ, we use the following lemmas. The proofs are easy and can be found in Arnold et al. (1998).

Lemma 1. *Let $U_1 < U_2 < \cdots < U_m$ be the first m upper record values from a population with cdf $F(\cdot)$. Define*

$$Y_i = -\ln[1 - F(U_i)], \qquad i = 1, 2, \ldots, m.$$

Then $Y_1 < Y_2 < \cdots < Y_m$ are the first m upper record values from a standard exponential distribution.

Lemma 2. *If $Y_1 < Y_2 < \cdots < Y_m$ are the first m upper record values from the standard exponential distribution. Then, the spacings $Y_1, Y_2 - Y_1, \ldots, Y_m - Y_{m-1}$ are iid standard exponential variables.*

Let $U_1 < U_2 < \cdots < U_m$ be the first m upper record values from the Rayleigh distribution. Here by providing some appropriate pivotal quantities, we construct several confidence regions for the Rayleigh parameters μ and λ. Let us define

$$Y_i = -\ln[1 - F(U_i)] = \lambda(U_i - \mu)^2; \qquad i = 1, 2, \ldots, m.$$

Then, we have from Lemma 1, that $Y_1 < Y_2 < \cdots < Y_m$ are the first m upper record values from a standard exponential distribution. Moreover, by Lemma 2, the spacings $Z_1 = Y_1,\ Z_2 = Y_2 - Y_1, \ldots, Z_m = Y_m - Y_{m-1}$ are iid random variables from the standard exponential distribution. Therefore, the random variable $V_j = 2\sum_{i=1}^{j} Z_i = 2Y_j$ follows a chi-square distribution with $2j$ degrees of freedom, i.e., $V_j \sim \chi^2_{2j}$, and $H_j = 2\sum_{i=j+1}^{m} Z_i = 2\ (Y_m - Y_j) \sim \chi^2_{2(m-j)}$, and also that V_j and H_j are independent, for any $j = 1, \ldots, m-1$. In order to derive the confidence regions for the parameters μ and λ, we consider the pivotal quantity $T_j \equiv T_j(\mu)$ defined by

$$T_j(\mu) = \frac{H_j/2(m-j)}{V_j/2j} = \frac{j}{m-j}\left(\frac{Y_m - Y_j}{Y_j}\right) = \frac{j}{m-j}\left[\frac{(U_m-\mu)^2}{(U_j-\mu)^2} - 1\right], \tag{3}$$

for $\mu < U_1$ and the pivotal quantity

$$S \equiv S(\mu,\lambda) = H_j + V_j = 2Y_m = 2\lambda(U_m - \mu)^2. \tag{4}$$

Clearly, T_j has an F distribution with $2(m-j)$ and $2j$ degrees of freedom for each $j = 1, \ldots, m-1$ and $S \sim \chi^2_{2m}$. Further, by Johnson et al. (1994) (see page 350), T_j and S are independent for each $j = 1, \ldots, m-1$.

Let for $0 < \alpha < 1$ that $\chi^2_{(\alpha),(\upsilon)}$ and $F_\alpha(\upsilon_1, \upsilon_2)$ be the upper α-quantiles of the $\chi^2_{(\upsilon)}$ and $F(\upsilon_1, \upsilon_2)$ distributions, respectively. Next theorem provides $m-1$ exact joint confidence regions for (μ, λ) based on the $m-1$ pairs of pivotal quantities $(T_1, S), \ldots, (T_{m-1}, S)$.

***Theorem* 1.** *Let $U_1 < U_2 < \cdots < U_m$ be the first m upper record values from the Rayleigh distribution in (1). Then, for any $j = 1, \ldots, m-1$, the $100(1-\alpha)\%$ joint confidence region for (μ, λ) is given by*

$$\begin{cases} \dfrac{U_j\sqrt{1+\frac{m-j}{j}F_{(\frac{1+\sqrt{1-\alpha}}{2}),(2(m-j),2j)}}-U_m}{\sqrt{1+\frac{m-j}{j}F_{(\frac{1+\sqrt{1-\alpha}}{2}),(2(m-j),2j)}}-1} < \mu < \dfrac{U_j\sqrt{1+\frac{m-j}{j}F_{(\frac{1-\sqrt{1-\alpha}}{2}),(2(m-j),2j)}}-U_m}{\sqrt{1+\frac{m-j}{j}F_{(\frac{1-\sqrt{1-\alpha}}{2}),(2(m-j),2j)}}-1}, \\ \dfrac{\chi^2_{(\frac{1+\sqrt{1-\alpha}}{2}),(2m)}}{2(U_m-\mu)^2} < \lambda < \dfrac{\chi^2_{(\frac{1-\sqrt{1-\alpha}}{2}),(2m)}}{2(U_m-\mu)^2}. \end{cases}$$

Proof. For any $j = 1, \ldots, m-1$, we have

$$P\left[F_{(\frac{1+\sqrt{1-\alpha}}{2}),(2(m-j),2j)} < T_j(\mu) < F_{(\frac{1-\sqrt{1-\alpha}}{2}),(2(m-j),2j)}\right] = \sqrt{1-\alpha}$$

and

$$P\left[\chi^2_{(\frac{1+\sqrt{1-\alpha}}{2}),(2m)} < S(\mu,\lambda) < \chi^2_{(\frac{1-\sqrt{1-\alpha}}{2}),(2m)}\right] = \sqrt{1-\alpha},$$

because $T_j \sim F_{2(m-j),2j}$ and $S \sim \chi^2_{2m}$. Since, for any $j = 1,\ldots,m-1$, T_j and S are independent, we conclude that

$$P[F_{(\frac{1+\sqrt{1-\alpha}}{2}),(2(m-j),2j)} < T_j < F_{(\frac{1-\sqrt{1-\alpha}}{2}),(2(m-j),2j)},$$
$$\chi^2_{(\frac{1+\sqrt{1-\alpha}}{2}),(2m)} < S < \chi^2_{(\frac{1-\sqrt{1-\alpha}}{2}),(2m)}] = 1-\alpha.$$

By solving the inequalities for μ and σ, the $m-1$ joint confidence regions for (μ,λ) can be obtained easily as given in Theorem 1. □

Remark 1. Based on the pivot T_j $(j = 1,\ldots,m-1)$, a $100(1-\alpha)\%$ confidence interval for μ is given by

$$\left(\frac{U_j\sqrt{1+\frac{m-j}{j}F_{(1-\frac{\alpha}{2}),(2(m-j),2j)}}-U_m}{\sqrt{1+\frac{m-j}{j}F_{(1-\frac{\alpha}{2}),(2(m-j),2j)}}-1}\ ,\ \frac{U_j\sqrt{1+\frac{m-j}{j}F_{(\frac{\alpha}{2}),(2(m-j),2j)}}-U_m}{\sqrt{1+\frac{m-j}{j}F_{(\frac{\alpha}{2}),(2(m-j),2j)}}-1}\right).$$

Based on the $m-1$ joint confidence regions mentioned in Theorem 1, one can select the optimal confidence region in term of the minimum confidence area. The obtained confidence regions depend only to the extreme records U_j and U_m. But it is reasonable that all of the sample information should be used to obtain confidence regions. For this reason, we consider here another pivotal quantity. Define

$$Q = \sum_{j=1}^{m-1} T_j(\mu) = \sum_{j=1}^{m-1} \frac{j}{m-j}\left[\frac{(U_m-\mu)^2}{(U_j-\mu)^2}-1\right], \qquad \mu < U_1. \tag{5}$$

It is easy to show that the distribution of Q does not depend on (μ,λ). Further, Q and S are independent. Now, using the joint pivot (Q,S), an exact joint confidence region for (μ,λ) can be constructed. To this, we need the following lemma.

Lemma 3. *Suppose that $\mu < c_1 < c_2 < \cdots < c_m$, where $c_1,\ldots,c_m$ are real* *constants. Let*

$$Q(\mu, m) = \sum_{j=1}^{m-1} \frac{j}{m-j} \left[\frac{(c_m - \mu)^2}{(c_j - \mu)^2} - 1 \right].$$

Then,

(a) $Q(\mu, m)$ *is strictly increasing in* μ *for any* $\mu \in (-\infty, c_1)$.

(b) For $0 < t < \infty$, *the equation,* $Q(\mu, m) = t$ *has a unique solution in* μ.

Proof. a) The first derivative of $Q(\mu, m)$ with respect to μ is

$$\frac{d}{d\mu} Q(\mu, m) = 2 \sum_{j=1}^{m-1} \frac{j}{m-j} \cdot \frac{(c_m - \mu)^2 - (c_m - \mu)(c_j - \mu)}{(c_j - \mu)^3}.$$

Since $\mu < c_1 < \cdots < c_m$, we have $(c_m - \mu)^2 \geqslant (c_m - \mu)(c_j - \mu)$ for any $j = 1, 2, \ldots, m-1$, and hence $\frac{d}{d\mu} Q(\mu, m) \geqslant 0$. This complete the proof.
b) Since the function $Q(\mu, m)$ is positive and increasing in μ with

$$\lim_{\mu \to -\infty} Q(\mu, m) = 0 \quad \text{and} \quad \lim_{\mu \to c_1} Q(\mu, m) = \sum_{j=1}^{m-1} \frac{j}{m-j} \left[\frac{(c_m - c_1)^2}{(c_j - c_1)^2} - 1 \right] = \infty,$$

then $Q(\mu, m) = t$ has a unique solution in μ, for any $0 < t < \infty$. $\square$

Now, let $q_{\alpha(m)}$ be the upper α-quantile of the distribution of the pivotal quantity Q. The following theorem provides an exact joint confidence region for (μ, λ) based on the pivot (Q, S).

Theorem 2. *Let* $U_1 < U_2 < \cdots < U_m$ *be the first* m *upper record values from the Rayleigh distribution. Then, the following inequalities determine a* $100(1-\alpha)\%$ *joint confidence region for* (μ, λ) *based on the pivot* (Q, S)*:*

$$\begin{cases} \psi(U_1, \ldots, U_m, q_{\frac{1+\sqrt{1-\alpha}}{2}(m)}) < \mu < \psi(U_1, \ldots, U_m, q_{\frac{1-\sqrt{1-\alpha}}{2}(m)}), \\ \frac{\chi^2_{(\frac{1+\sqrt{1-\alpha}}{2}),(2m)}}{2(U_m - \mu)^2} < \lambda < \frac{\chi^2_{(\frac{1-\sqrt{1-\alpha}}{2}),(2m)}}{2(U_m - \mu)^2}, \end{cases}$$

where $\psi(U_1, \ldots, U_m, t)$ *is the solution of* μ *for the equation*

Table 1. The upper α-quantiles $q_{\alpha(m)}$ of $Q(\mu, m)$

	α									
m	0.99	0.01	0.9873	0.0127	0.975	0.025	0.95	0.05	0.90	0.10
2	0.0095	99.6235	0.0126	83.3915	0.0259	40.7528	0.0524	19.5492	0.1144	9.0890
3	0.1225	104.4499	0.1373	86.5422	0.1984	48.0786	0.3092	25.9073	0.4997	13.6108
4	0.3185	110.5864	0.3564	92.6991	0.5033	52.5265	0.7128	28.8446	1.0349	16.6134
5	0.6623	115.3071	0.7215	96.2692	0.9291	53.4702	1.2095	31.7062	1.6459	19.2498
6	0.9329	114.5048	1.0154	92.3632	1.2925	54.4834	1.6854	33.4537	2.2518	21.4259
7	1.3339	138.2313	1.4132	110.5868	1.7975	62.2852	2.2611	37.3436	2.9450	23.3300
8	1.7990	130.7326	1.8845	103.4271	2.2811	60.6067	2.8465	37.3073	3.6175	24.8576
9	2.2972	126.3819	2.4371	104.6259	2.8304	64.2165	3.4337	39.6563	4.3004	26.7135
10	2.7129	118.5268	2.8487	100.6608	3.3410	62.4510	4.0277	41.1799	5.0616	28.7925
11	3.1652	114.5046	3.3151	99.7138	3.9176	62.3512	4.6887	41.3245	5.8162	29.3685
12	3.5793	143.5688	3.7625	113.8862	4.4965	68.4054	5.3484	44.6854	6.5063	31.2368
13	4.2893	134.5849	4.4793	114.5675	5.0818	70.2351	5.8937	46.5693	7.2046	33.1427
14	4.7887	133.9197	5.0214	103.4453	5.6831	66.9488	6.6637	46.0921	7.9878	34.0953
15	5.3406	137.5944	5.6308	114.1457	6.3912	70.8524	7.4715	48.9024	8.8551	35.4491
16	5.8045	128.6795	6.0754	111.9826	6.9524	73.8603	8.0119	49.8351	9.5002	36.9925
17	6.5094	137.0948	6.7653	112.5050	7.6366	71.7216	8.7618	50.7843	10.2887	37.8498
18	7.0301	124.2017	7.3233	104.2625	8.2129	71.4663	9.4661	51.6745	11.1041	39.0846
19	7.6386	130.0477	7.9189	108.5718	8.9345	74.0479	10.2628	53.0135	11.9386	40.9066
20	8.1163	140.4943	8.4362	116.1457	9.4578	77.0187	10.7905	55.2844	12.6170	42.5622
21	8.7996	142.9720	9.1562	121.8273	10.2679	77.3027	11.5815	56.7319	13.5124	43.9248
22	9.3456	144.5844	9.7339	121.5262	10.9180	79.8690	12.3567	57.7689	14.2724	44.7949
23	10.1114	140.0154	10.4640	117.9485	11.6041	79.7559	13.0181	58.5093	15.0332	46.1700
24	10.7637	146.8277	11.1783	126.1649	12.3699	80.9380	13.8670	60.4229	15.9385	47.7764
25	11.3202	131.0800	11.6991	112.9332	12.9891	79.3259	14.5160	60.7625	16.6909	48.7578
30	14.8479	141.9864	15.2383	124.6956	16.6703	86.6005	18.3592	67.2761	20.8426	54.4106
40	21.3058	168.6216	21.8866	143.6941	23.7145	99.7097	25.9825	79.9148	29.0420	67.0554
50	28.4495	168.4166	29.0802	148.1588	31.2146	111.0425	33.9035	92.6934	37.3505	79.0237

$$\sum_{j=1}^{m-1} \frac{j}{m-j} \left[\frac{(U_m-\mu)^2}{(U_j-\mu)^2} - 1 \right] = t.$$

Proof. We have

$$P\left[q_{\left(\frac{1+\sqrt{1-\alpha}}{2}\right)} < Q < q_{\left(\frac{1-\sqrt{1-\alpha}}{2}\right)} \right] = \sqrt{1-\alpha}$$

and

$$P\left[\chi^2_{\left(\frac{1+\sqrt{1-\alpha}}{2}\right),(2m)} < S < \chi^2_{\left(\frac{1-\sqrt{1-\alpha}}{2}\right),(2m)} \right] = \sqrt{1-\alpha}.$$

By independence of Q and S, we obtain

$$P\left[q_{(\frac{1+\sqrt{1-\alpha}}{2})} < Q < q_{(\frac{1-\sqrt{1-\alpha}}{2})},\ \chi^2_{(\frac{1+\sqrt{1-\alpha}}{2}),(2m)} < S < \chi^2_{(\frac{1-\sqrt{1-\alpha}}{2}),(2m)}\right] = 1-\alpha.$$

Now by Lemma 3, we can find the joint confidence region as described in Theorem 2. □

Remark 2. The $100(1-\alpha)\%$ confidence interval for μ based on the pivot Q is

$$\psi(U_1,\ldots,U_m,q_{1-\frac{\alpha}{2}(m)}) < \mu < \psi(U_1,\ldots,U_m,q_{\frac{\alpha}{2}(m)})$$

where $\psi(U_1,\ldots,U_m,t)$ is the solution of μ for the equation

$$\sum_{j=1}^{m-1} \frac{j}{m-j}\left[\frac{(U_m-\mu)^2}{(U_j-\mu)^2}-1\right] = t.$$

It should be mentioned here that since the exact distribution of the pivotal quantity Q is too hard to derive algebraically, we need to compute the upper α-quantile $q_{\alpha(m)}$ by using Monte Carlo simulation. In Table 1, we presented the upper α-quantile $q_{\alpha(m)}$ of Q for $m = 2, 3, \ldots, 25, 30, 40, 50$ and various values of α over 20000 replications.

3 Numerical Examples

In this section, two examples from wind speed data and annual rainfall data are given to illustrate the proposed joint confidence regions.

Example 1: Wind speed data

Let us first consider the wind speed data reported by Battacharya and Bhattacharjee (2010). The data are the average monthly wind speed (m/s) of Kolkata from 1st March 2009 to 31st March 2009.

```
0.56 0.28 0.56 0.56 1.11 0.83 1.11 1.94 1.11 0.83 1.11
1.39 0.28 0.56 0.28 0.28 0.28 0.83 1.39 1.11 1.11 0.83
0.56 0.83 1.67 1.94 1.39 0.83 2.22 1.67 2.22.
```

We have checked the validity of the Rayleigh distribution based on the parameters $\mu = 0.0225$ and $\lambda = 0.75911$, using the Kolmogorov-Smirnov (K–S)

Table 2. The areas of 95% confidence regions for (μ, λ) in Example 1.

Pivot	(T_1, S)	(T_2, S)	(T_3, S)	(Q, S)
Region	A_1	A_2	A_3	A_4
Area	4.31223	5.05170	5.25395	1.97945

test. It is observed that the K-S distance is K-S = 0.1257 with a corresponding p-value = 0.6655. This shows that the Rayleigh model fits well to the above data.

If only the upper record values have been observed, these are as follows:

```
0.56    1.11    1.94     2.22.
```

Now, we use the methods proposed in Section 2 to construct the exact joint confidence regions for (μ, λ). To obtain the 95% joint confidence regions, we need the following percentiles:

$$
\begin{aligned}
F_{0.0127(6,2)} &= 78.07254, & F_{0.9873(6,2)} &= 0.1014364,\\
F_{0.0127(4,4)} &= 14.02461, & F_{0.9873(4,4)} &= 0.07130322,\\
F_{0.0127(2,6)} &= 9.858393, & F_{0.9873(2,6)} &= 0.0128086,\\
q_{0.0127(4)} &= 92.6991, & q_{0.9873(4)} &= 0.3564\\
\chi^2_{0.0127(8)} &= 19.4347, & \chi^2_{0.9873(8)} &= 1.768713.
\end{aligned}
$$

By Theorem 1 and using the S-PLUS package, the 95% joint confidence regions for (μ, λ) based on the pivots $(T_1, S), (T_2, S)$ and (T_3, S) are given respectively by

$$
A_1 = \left\{ (\mu, \lambda) : -11.1249 < \mu < 0.4442, \quad \frac{1.76871}{2(2.22 - \mu)^2} < \lambda < \frac{19.4347}{2(2.22 - \mu)^2} \right\},
$$

$$
A_2 = \left\{ (\mu, \lambda) : -30.5701 < \mu < 0.5600, \quad \frac{1.76871}{2(2.22 - \mu)^2} < \lambda < \frac{19.4347}{2(2.22 - \mu)^2} \right\},
$$

$$
A_3 = \left\{ (\mu, \lambda) : -129.3617 < \mu < 0.5600, \quad \frac{1.76871}{2(2.22 - \mu)^2} < \lambda < \frac{19.4347}{2(2.22 - \mu)^2} \right\}.
$$

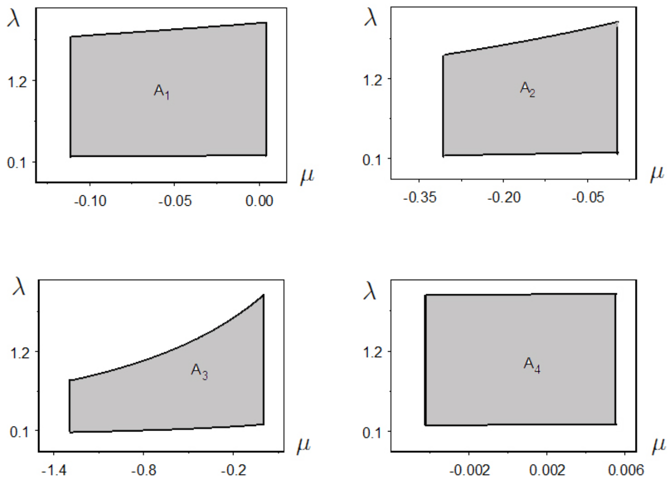

Figure 1. The 95% joint confidence region for (μ, λ) in Example 1.

By Theorem 2, the 95% joint confidence region based on the pivot (Q, S) is

$$A_4 = \left\{ (\mu, \lambda) : -0.4233 < \mu < 0.5600, \quad \frac{1.76871}{2(2.22 - \mu)^2} < \lambda < \frac{19.4347}{2(2.22 - \mu)^2} \right\},$$

In Table 2, we have presented the areas of the 95% joint confidence regions based on different pivots. From Table 2, it is observed that the region A_4 has the smallest area and hence, the pivot (Q, S) is the optimal pivotal quantity. Figure 1 shows the shapes of different Rayleigh confidence regions for μ and λ.

Example 2: Annual rainfall data

In this example, we present a data analysis of the amount of annual rainfall (in inches) recorded at the Los Angeles Civic Center for 50 years, from 1962 to 2012 (season July 1-June 30). See the website of Los Angeles Almanac: www.laalmanac.com/weather/we08aa.htm. The validity of the Rayleigh model to these data can be checked using the K-S test. For these data, the observed rainfall records are as follows:

8.38, 13.68, 20.44, 22.00, 27.47, 33.44, 37.96.

By Theorem 1, the 95% joint confidence regions for (μ, λ) based on the pivots

Table 3. The areas of 95% joint confidence regions for (μ, λ) in Example 2.

Pivot	(T_1, S)	(T_2, S)	(T_3, S)	(T_4, S)	(T_5, S)	(T_6, S)	(Q, S)
Region	B_1	B_2	B_3	B_4	B_5	B_6	B_7
Area	0.267077	0.315092	0.331864	0.362139	0.379825	0.394150	0.129316

(T_j, S), $j = 1, \ldots, 6$ are respectively as:

$$B_1 = \left\{(\mu, \lambda) : -67.3112 < \mu < 6.9495, \quad \frac{4.88886}{2(37.96 - \mu)^2} < \lambda < \frac{28.37037}{2(37.96 - \mu)^2}\right\},$$

$$B_2 = \left\{(\mu, \lambda) : -105.529 < \mu < 8.3800, \quad \frac{4.88886}{2(37.96 - \mu)^2} < \lambda < \frac{28.37037}{2(37.96 - \mu)^2}\right\},$$

$$B_3 = \left\{(\mu, \lambda) : -142.525 < \mu < 8.3800, \quad \frac{4.88886}{2(37.96 - \mu)^2} < \lambda < \frac{28.37037}{2(37.96 - \mu)^2}\right\},$$

$$B_4 = \left\{(\mu, \lambda) : -299.6522 < \mu < 8.3800, \quad \frac{4.88886}{2(37.96 - \mu)^2} < \lambda < \frac{28.37037}{2(37.96 - \mu)^2}\right\},$$

$$B_5 = \left\{(\mu, \lambda) : -649.0486 < \mu < 8.3800, \quad \frac{4.88886}{2(37.96 - \mu)^2} < \lambda < \frac{28.37037}{2(37.96 - \mu)^2}\right\},$$

$$B_6 = \left\{(\mu, \lambda) : -4207.989 < \mu < 8.3800, \quad \frac{4.88886}{2(37.96 - \mu)^2} < \lambda < \frac{28.37037}{2(37.96 - \mu)^2}\right\}.$$

Also by Theorem 2, the 95% joint confidence region based on the pivot (Q, S) is

$$B_7 = \left\{(\mu, \lambda) : -5.91444 < \mu < 8.3800, \quad \frac{4.88886}{2(37.96 - \mu)^2} < \lambda < \frac{28.37037}{2(37.96 - \mu)^2}\right\}.$$

Table 3 shows the areas of different confidence regions for the parameters

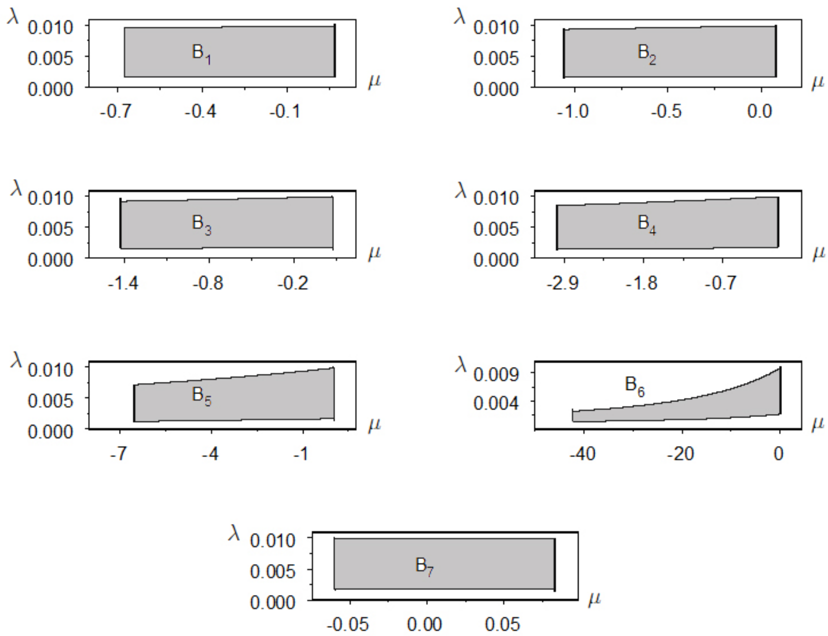

Figure 2. The 95% joint confidence region for (μ, λ) in Example 2.

μ and λ. As we can see from Table 3, the pivot (Q, S) provides again the smallest confidence area. The shapes of different joint confidence regions are given in Figure 2.

4 Simulation Study

In this section, we carry out a Monte Carlo simulation to compare the performances of the different confidence regions. Before progressing further, first we describe how we can generate an upper record sample from any continuous distribution. The upper record sample $U_1, U_2, \ldots, U_m$ from any continuous cdf $F(\cdot)$ can be generated using the following simulational algorithm:
step 1. Generate m independent Exp(1) observations $X_1, \ldots, X_m$.
step 2. Set $R_i = X_1 + X_2 + \cdots + X_i$ for $i = 1, 2, \ldots, m$.
step 3. Set $W_i = 1 - e^{-R_i}$ for $i = 1, 2, \ldots, m$. Then $W_1, W_2, \ldots, W_m$ is the required upper record sample from the Uniform (0,1) distribution.

Table 4. The average area and coverage probability of 95% joint confidence region for (μ, λ).

m	Pivots	CP	CA
5	(T_1, S)	0.950	7.0787
	(T_2, S)	0.949	7.6066
	(T_3, S)	0.945	8.1522
	(T_4, S)	0.953	8.6028
	(Q,S)	**0.975**	**6.2913**
10	(T_1, S)	0.946	3.9620
	(T_2, S)	0.952	4.1724
	(T_3, S)	0.946	4.3949
	(T_4, S)	0.948	4.6089
	(T_5, S)	0.951	4.8923
	(T_6, S)	0.953	5.1965
	(T_7, S)	0.954	5.4924
	(T_8, S)	0.951	5.8604
	(T_9, S)	0.949	5.9651
	(Q,S)	**0.971**	**3.6330**
15	(T_1, S)	0.948	2.9786
	(T_2, S)	0.952	3.1128
	(T_3, S)	0.953	3.2664
	(T_4, S)	0.951	3.4132
	(T_5, S)	0.950	3.5339
	(T_6, S)	0.948	3.6665
	(T_7, S)	0.948	3.8194
	(T_8, S)	0.948	3.9919
	(T_9, S)	0.949	4.1693
	(T_{10}, S)	0.951	4.3584
	(T_{11}, S)	0.946	4.6193
	(T_{12}, S)	0.949	4.9115
	(T_{13}, S)	0.948	5.2216
	(T_{14}, S)	0.950	5.4530
	(Q,S)	**0.970**	**2.9363**

m	pivots	CP	CA
30	(T_1, S)	0.960	1.9568
	(T_2, S)	0.962	2.0278
	(T_3, S)	0.954	2.0793
	(T_4, S)	0.959	2.1376
	(T_5, S)	0.959	2.1907
	(T_6, S)	0.958	2.2518
	(T_7, S)	0.960	2.2989
	(T_8, S)	0.964	2.3688
	(T_9, S)	0.959	2.4214
	(T_{10}, S)	0.958	2.4657
	(T_{11}, S)	0.961	2.5465
	(T_{12}, S)	0.961	2.5927
	(T_{13}, S)	0.957	2.6392
	(T_{14}, S)	0.963	2.7086
	(T_{15}, S)	0.960	2.7786
	(T_{16}, S)	0.965	2.8671
	(T_{17}, S)	0.964	2.9455
	(T_{18}, S)	0.962	3.0169
	(T_{19}, S)	0.959	3.1041
	(T_{20}, S)	0.957	3.2078
	(T_{21}, S)	0.965	3.3272
	(T_{22}, S)	0.956	3.4182
	(T_{23}, S)	0.960	3.5946
	(T_{24}, S)	0.962	3.7225
	(T_{25}, S)	0.968	3.8965
	(T_{26}, S)	0.956	4.0883
	(T_{27}, S)	0.961	4.3975
	(T_{28}, S)	0.960	4.4012
	(T_{29}, S)	0.966	4.4521
	(Q,S)	**0.969**	**1.2470**

Table 5. Comparison of average confidence area based on complete data and record data.

	Complete data	Record data
$m = 5$	3.0069	6.2913
$m = 10$	0.5339	3.6330
$m = 15$	0.5100	2.9363
$m = 30$	0.0244	1.2470

step 4. Finally, we set $U_i = F^{-1}(W_i)$ for $i = 1, 2, \ldots, m$. Then $U_1, U_2, \ldots, U_m$ is the required upper record sample from the distribution $F(\cdot)$.

4.1 Comparison of Confidence Regions

Using the above algorithm, we have randomly generated 10000 upper record sample $U_1, U_2, \ldots, U_m$ from the standard Rayleigh distribution (i.e., $\mu = 0$ and $\lambda = 1$) and then computed the 95% confidence regions using Theorems 1 and 2. We then compared the performances of these regions in terms of confidence area (CA) and coverage probability (CP). In Table 4, for different values of m, we presented the average confidence areas and coverage probabilities of different joint confidence regions over 10000 replications. All the computations are performed using SPLUS package. The simulation results show that when m increases, the areas are decreased. In addition, the coverage probabilities of the joint confidence regions for (μ, λ) are close to the desired level of 0.95 for different pivots and different record sample sizes. The coverage probabilities based on the joint pivot (Q, S) are slightly larger than the corresponding coverage probabilities based on the other pivots. Comparing the average areas of the confidence regions, it is observed that in all of cases considered, the joint pivot (Q, S) provides the first smallest area for (μ, λ) and the second smallest area for (μ, λ) is provided by the pivot (T_1, S). Based on the simulation results, overall speaking, we would recommend the use of pivot (Q, S) for constructing joint confidence region for the Rayleigh parameters μ and λ.

Let us now compare the confidence regions obtained based on first m record data by the confidence regions obtained based on a simple random sample $X_1, \ldots, X_m$ from the original distribution. In Section 2, the pivot (Q, S) was used for constructing a joint confidence region for (μ, λ) based on the first m upper record data. If (Q', S') is the corresponding pivot based

Table 6. The average confidence length (CL) and coverage probability (CP) of the 95% confidence intervals for μ.

m	Pivots	CP	CL
5	T_1	0.949	4.3941
	T_2	0.950	6.9294
	T_3	0.949	13.817
	T_4	0.952	75.7907
	Q	**0.922**	**1.4508**
10	T_1	0.943	2.9520
	T_2	0.947	3.5044
	T_3	0.952	4.1125
	T_4	0.955	4.9648
	T_5	0.957	6.0926
	T_6	0.951	8.0292
	T_7	0.951	11.4532
	T_8	0.947	21.5654
	T_9	0.950	114.1194
	Q	**0.924**	**1.5120**
15	T_1	0.947	2.6119
	T_2	0.953	2.9375
	T_3	0.947	3.2398
	T_4	0.946	3.5408
	T_5	0.951	3.8638
	T_6	0.949	4.3356
	T_7	0.949	4.8865
	T_8	0.952	5.6166
	T_9	0.952	6.4380
	T_{10}	0.949	7.8484
	T_{11}	0.946	10.2109
	T_{12}	0.950	14.6433
	T_{13}	0.950	26.9286
	T_{14}	0.952	142.3817
	Q	**0.934**	**1.5440**

m	pivots	CP	CL
30	T_1	0.951	2.2384
	T_2	0.950	2.3729
	T_3	0.952	2.5097
	T_4	0.953	2.626
	T_5	0.950	2.7474
	T_6	0.953	2.8489
	T_7	0.949	2.9707
	T_8	0.951	3.0917
	T_9	0.949	3.2353
	T_{10}	0.943	3.3629
	T_{11}	0.949	3.5734
	T_{12}	0.947	3.7468
	T_{13}	0.952	3.8811
	T_{14}	0.952	4.1047
	T_{15}	0.952	4.2471
	T_{16}	0.946	4.5696
	T_{17}	0.955	4.8355
	T_{18}	0.948	5.1708
	T_{19}	0.953	5.5137
	T_{20}	0.954	5.9125
	T_{21}	0.950	6.5241
	T_{22}	0.948	7.2477
	T_{23}	0.948	8.0891
	T_{24}	0.950	9.5706
	T_{25}	0.943	11.226
	T_{26}	0.952	14.7232
	T_{27}	0.942	21.0376
	T_{28}	0.950	38.9845
	T_{29}	0.932	206.399
	Q	**0.935**	**1.5564**

on the order statistics $X_{(1)}, \ldots, X_{(m)}$, then using methods similar to those discussed in Section 2, one can also construct a joint confidence region for (μ, λ). In Table 5, we have compared the confidence areas obtained based on ordered sample (complete) data and record data. As expected, the area obtained based on the complete data is smaller than that the area obtained based on record data.

4.2 Comparison of Confidence Intervals

We have also compared the confidence intervals obtained based on the usual F-distributed pivots $T_j(\mu)$ (for $j = 1, 2, \ldots, m-1$) and the one that obtained based on the additive pivotal quantity $Q = \sum_{j=1}^{m-1} T_j(\mu)$. Table 6 presents the average confidence lengths and coverage probabilities of the 95% confidence intervals for μ. As we can see from the table, the pivot $Q(\mu, m)$ provides the shortest confidence length. The coverage probabilities of the confidence intervals for μ are close to the desired level of 0.95 for different pivots and different record sample sizes. Also, the coverage probabilities based on the pivots $T_j(\mu)$ are slightly larger than the one that obtained based on the pivot Q.

Acknowledgement

The authors would like to thank the three referees for their valuable comments and suggestions which greatly improved the paper.

References

Ahmadi, J. and Arghami, N.R. (2003). Comparing the Fisher Information in Record Values and iid Observations. *Statistics*, **37**, 435-441.

Ahsanullah, M. (1995). *Introduction to Record Statistics*, NOVA Science Publishers Inc., Huntington, New York.

Arnold, B. C., Balakrishnan, N. and Nagaraja, H. N. (1998), *Records.* John Wiley & Sons, New York.

Asgharzadeh, A. and Abdi, M. (2011). Confidence Intervals and Joint Confidence Regions for the Two-parameter Exponential Distribution based on Records. *Communications of the Korean Statistical Society*, **18**, 102-110.

Asgharzadeh, A. and Azizpour, M. (2016). Bayesian Inference for Rayleigh Distribution under Hybrid Censoring. *International Journal of System Assurance Engineering and Management*, **7**, 239-249.

Asgharzadeh, A., Abdi, M. and Valiollahi, R. (2013). Analysis of Record Data from the Scaled Logistic Distribution, *Journal of Statistical Research of Iran*, **10**, 41-62.

Asgharzadeh, A., Abdi, M. and Wu, S.J. (2015). Interval Estimation for the Two-parameter Bathtub-Shaped Lifetime Distribution based on Records, *Hacettepe Journal of Mathematics and Statistics*, **44**, 399-416.

Asgharzadeh, A., Abdi, M. and Nadarajah, S. (2016). Interval Estimation for Gumbel Distribution Using Climate Records, *Bulletin of the Malaysian Mathematical Sciences Society*, **39**, 257-270.

Battacharya, P. and Bhattacharjee, R. (2010). A Study on Weibull Distribution for Estimating the Parameters. *Journal of Applied Quantitative Methods*, **5**, 234-241.

Chandler, K.N. (1952). The Distribution and Frequency of Record Values. *Journal of Royal Statist Soc.*, **B14**, 220-228.

Dey, S., Dey, T. and Kundu, D. (2014). Two-parameter Rayleigh Distribution: Different Methods of Estimation. *American Journal of Mathematical and Management Sciences*, **33**, 55-74.

Johnson, N.L., Kotz, S. and Balakrishnan, N. (1994). *Continuous Univariate Distributions*, John Wiley & Sons, New York.

Khan, H.M.R., Provost, S.B. and Singh, A. (2010). Predictive Inference from a Two-parameter Rayleigh Life Model Given a Doubly Censored Sample, *Communications in Statistics - Theory and Methods*, **39**, 1237-1246.

Kinanci, I., Wu, S.J. and Kus, C. (2017). Confidence Intervals and Regions for the Generalized Inverted Exponential Distribution based on Progressively Censored and Upper Record Data, *REVSTAT* (to appear).

Mirfarah, E. and Ahmadi, J. (2014). Pitman-Closeness of Preliminary Test and Some Classical Estimators based on Records from Two-parameter Exponential Distribution, *Journal of Statistical Research of Iran*, **11**, 73-96.

Nevzorov, V.B. (1988). Records, *Theory of Probability and its Applications,* **32**, 201-228.

Raqab, M.Z. and Madi, MT. (2002). Bayesian Prediction of the Total Time on Test Using Doubly Censored Rayleigh Data. *Journal of Statistical Computation and Simulation*, **72**, 781-789.

Rayleigh, J.W.S. (1880). On the Resultant of a Large Number of Vibrations of the Some Pitch and of Arbitrary Phase, *Philosophical Magazine*, 5-th Series, **10**, 73-78.

Soliman, A.A. and Al-Aboud F.M. (2008). Baysian Inference Using Record Values from Rayleigh Model with Application. *European Journal of Operational Research*, **185**, 252-272.

Wu, S., Chen, D. and Chen, S. (2006). Baysian Inference for Rayleigh Distribution under Progressive Censored Sample. *Applied Stochastic Models in Business and Industry*, **26**, 126-279.

Zakerzadeh, H., Jafari, A.A. and Karimi, M. (2016). Preliminary Test and Shrinkage Estimations of Scale Parameters for Two Exponential Distributions based on Record Values, *Journal of Statistical Research of Iran*, **13**, 43-58.

Mousa Abdi
Department of Statistics,
Higher Education Complex of Bam,
Bam, Iran.
email: *me.abdi@bam.ac.ir*

Akbar Asgharzadeh
Department of Statistics,
University of Mazandaran,
Babolsar, Iran.
email: *a.asgharzadeh@umz.ac.ir*