Estimation of Tail Indices of Heavy-Tailed Distributions with Application

Outline of Ph.D. Thesis

Amenah Al-Najafi

Supervisors:
Dr. Péter Kevei
Dr. László Viharos

Doctoral School of Mathematics and Computer Science
University of Szeged, Bolyai Institute

Szeged
2021
1 Weighted least squares estimators for the Parzen tail index

The results presented in this chapter are based on [ANV20].

We propose a class of weighted least squares (WLS) estimators for the Parzen tail index. Our approach is based on the method developed by Holan and McElroy [HM10]. We investigate consistency and asymptotic normality of the WLS estimators. Through a simulation study, we make a comparison with the Hill, Pickands, DEdH (Dekkers, Einmahl and de Haan) and ordinary least squares (OLS) estimators using the mean square error as criterion. The results show that in a restricted model some members of the WLS estimators are competitive with the Pickands, DEdH and OLS estimators.

1.1 The tail index estimation

In classical tail index estimation it is assumed that the tail of the distribution function is regularly varying at infinity with some positive index. Parzen [Par79, Par04] studied an alternative model for the tail of the distribution. Let $F$ be an absolutely continuous probability distribution function with density function $f$ and let $Q$ denote the corresponding quantile function defined as

$$Q(s) := \inf\{x : F(x) \geq s\}, \quad 0 < s \leq 1, \quad Q(0) := Q(0+) .$$

Parzen [Par79] used the density-quantile function $fQ(\cdot) = f(Q(\cdot))$ to classify probability distributions. Parzen [Par79] assumed that the limit

$$\nu_1 := \lim_{u \to 1} \frac{(1 - u)J(u)}{fQ(u)}$$

exists, where $J$ is the score function defined as $J(u) = -(fQ)'(u)$. Assumption (1) yields the following approximation for $u$ values near 1:

$$fQ(u) \approx C(1 - u)^{\nu_1},$$

for some positive constant $C$. Based on the parameter $\nu_1$, Parzen [Par79] classified the probability distributions. Heavy tailed distributions correspond to $\nu_1 > 1$.

Parzen [Par04] assumed that $fQ(\cdot)$ is regularly varying at 0 and 1:

$$fQ(u) = u^{\nu_0}L_0(u), \quad u \in [0, 1/2), \quad (2)$$

$$fQ(u) = (1 - u)^{\nu_1}L_1(u), \quad u \in (1/2, 1], \quad (3)$$

where $\nu_0, \nu_1 > 0$ are finite constants and $L_0$ and $L_1$ are slowly varying at zero. The parameters $\nu_0$ and $\nu_1$ are called the left and right tail exponents of the density-quantile function.

Using Karamata’s representation theorem for slowly varying functions ([BGT89, Theorem 1.3.1]), Holan and McElroy [HM10] proved the following result ([HM10]
Lemma 1): If \( K \) is a slowly varying function at infinity and \( L(x) = K(1/x) \) for \( x \in (0, 1) \), then \( \log L \) is square integrable. It follows that \( L_i \) can be expressed as

\[
L_i(u) = \exp \left\{ \theta_{i,0} + 2 \sum_{k=1}^{\infty} \theta_{i,k} \cos(2\pi ku) \right\}, \quad i = 0, 1.
\]

(4)

In order to estimate the tail exponents, Holan and McElroy [HM10] assumed that \( L_i \) satisfies the representation

\[
L_i(u) = L_i^{(p_i)}(u) = \exp \left\{ \theta_{i,0} + 2 \sum_{k=1}^{p_i} \theta_{i,k} \cos(2\pi ku) \right\}, \quad i = 0, 1,
\]

(5)

where \( p_i \) is fixed and unknown. In the representation [2] and [3] they considered \( f_q(u) \) for \( u \in (0, u_l) \) and \( u \in [u_r, 1] \), where \( u_l \leq 1/2 \) and \( u_r \geq 1/2 \) are chosen by the statistician, and they assumed that \( p_i < \tilde{p}_i \), where \( \tilde{p}_i \) is a prespecified integer. Using representation [5], we obtain the equations

\[
\log f_q(u) = \nu_0 \log u + \theta_{0,0} + 2 \sum_{k=1}^{p_0} \theta_{0,k} \cos(2\pi ku), \quad u \in (0, u_l],
\]

\[
\log f_q(u) = \nu_1 \log(1-u) + \theta_{1,0} + 2 \sum_{k=1}^{p_1} \theta_{1,k} \cos(2\pi k(1-u)), \quad u \in [u_r, 1).
\]

Based on some estimator \( \hat{f}_q(u) \) of the density-quantile \( f_q(u) \), this leads to the regression equations

\[
\log \hat{f}_q(u_j) = \nu_0 \log u_j + \theta_{0,0} + 2 \sum_{k=1}^{p_0} \theta_{0,k} \cos(2\pi ku_j) + \varepsilon(u_j),
\]

\[
\log \hat{f}_q(1-u_j) = \nu_1 \log u_j + \theta_{1,0} + 2 \sum_{k=1}^{p_1} \theta_{1,k} \cos(2\pi ku_j) + \varepsilon(1-u_j),
\]

where \( \varepsilon(u) = \log \left( \hat{f}_q(u)/f_q(u) \right) \) is the residual process, \( u_j = j/n, j = u_{[na]}, \ldots, u_{[nb]} \) and \( 0 < a < b < 1 \), so the percentiles \( u_j \) are chosen from a subset \( [a, b] \) of the interval \( (0, 1) \). Holan and McElroy [HM10] obtained some estimators \( \hat{\nu}_0 \) and \( \hat{\nu}_1 \) for the tail exponents \( \nu_0 \) and \( \nu_1 \) using ordinary least squares regression.

We propose a more general class of estimators using weighted least squares regression. We choose some nonnegative weights of the form \( w_{j,n} = R(j/n) \) with some weight function \( R \). Set \( y_j := \log \hat{f}_q(u_j) \),

\[
y := (y_{[na]}, \ldots, y_{[nb]})',
\]

\[
W := \text{diag}(w_{[na],n}, \ldots, w_{[nb],n}),
\]

and let \( X := [G^*, G_0, 2G_1, \ldots, 2G_{\tilde{p}_0}] \), where

\[
G^* = \left( \log(u_{[na]}), \ldots, \log(u_{[nb]}) \right)',
\]

\[
G_k = \left( \cos(2\pi ku_{[na]}), \ldots, \cos(2\pi ku_{[nb]}) \right)', \quad k = 0, \ldots, \tilde{p}_0.
\]
Set \( \beta_{p_0} := (\nu_0, \theta_{0,0}, \theta_{0,1}, \ldots, \theta_{0,p_0})' \), where \( \theta_{0,j} = 0 \) if \( j > p_0 \). By minimizing the weighted sum of squares

\[
\sum_{j=\lfloor na \rfloor}^{\lceil nb \rceil} w_{j,n} (y_j - \nu_0 \log u_j - \theta_{0,0} - 2 \sum_{k=1}^{p_0} \theta_{0,k} \cos(2\pi k u_j))^2,
\]

we obtain the following estimator of \( \beta_{p_0} \):

\[
\hat{\beta}_{p_0} = (X'WX)^{-1}X'y.
\]

Then the weighted least squares estimator of \( \nu_0 \) can be written in the form

\[
\hat{\nu}_0 = e_1'(\hat{\beta}_{p_0}) = e_1'(X'WX)^{-1}X'y,
\]

where \( e_1 \) is the \( \tilde{p}_0 + 2 \) dimensional vector defined as \( e_1 = (1, 0, 0, \ldots, 0)' \). The right tail exponent \( \nu_1 \) can be estimated similarly.

A crucial point of this method is to choose a good estimator for the density-quantile \( f_Q(u) \). Letting \( q(u) := Q'(u) \) denote the quantile density function, and using the identity

\[
f_Q(u)Q'(u) = 1,
\]

one wish to estimate \( q(u) \) instead of \( f_Q(u) \). Given a sample \( X_1, \ldots, X_n \) with distribution function \( F \), let \( F_n \) denote its empirical distribution function and define \( Q_n := F_n^{-1} \) to be the empirical quantile function. Holan and McElroy \cite{HM10} used the kernel quantile estimator of \( q(u) \):

\[
\hat{q}_n(u) = \frac{d}{du} \int_0^1 Q_n(t)K_n(u,t)d\mu_n(t), \quad u \in (0,1),
\]

where the kernel function \( K_n(u,t) \) and the measure \( \mu_n \) satisfy the following conditions of Cheng \cite{Che95}:

- \((K_1)\) For every \( n, 0 < \mu_n([0,1]) < \infty \), and \( \mu_n(\{0,1\}) = 0 \).

- \((K_2)\) For every \( n \) and each \((u,t)\), \( K_n(u,t) \geq 0 \), and for every \( u \in [a,b], \int_0^1 K_n(u,t)d\mu_n(t) = 1 \).

- \((K_3)\) For every \( n \), \( \int_0^1 tK_n(u,t)d\mu_n(t) = u, u \in [a,b] \).

- \((K_4)\) There is a sequence \( \delta_n \downarrow 0 \) such that \( \sup_{u \in [a,b]} \mid \int_{u-\delta_n}^{u+\delta_n} K_n(u,t)d\mu_n(t) - 1 \downarrow 0 \) as \( n \uparrow \infty \).

Let \( S_n \) be the unique closed subset of \((0,1)\) such that \( \mu_n((0,1)\setminus S_n) = 0 \) and \( \mu_n((0,1)\setminus S_n') > 0 \) for any \( S_n' \subset S_n \).

For the sequence \( \delta_n \) in \((K_2)\), let \( I_n(u) = [u - \delta_n, u + \delta_n], I_n^c(u) = (0,1) \setminus I_n(u) \), for \( u \in [a,b] \). Define \( \Lambda(u; K_n) = \int_{I_n(u)} K_n(u,t)d\mu_n(t), u \in [a,b] \), and for a well-defined function \( g \) on \((0,1)\), let \( \Psi(g; K_n) = \sup_{u \in [a,b]} \int_{I_n^c(u)} \mid g(t)K_n'(u,t)d\mu_n(t) \). It is also assumed that the derivative \( K_n'(u,t) = \partial K_n(u,t)/\partial u \) satisfies the conditions \((K_5)\) – \((K_7)\) below:
(K5) For every $n$, $\sup_{u \in [a, b]} \int_0^1 |K_n'(u, t)| d\mu_n(t) < \infty$.

(K6) For every $n$ and each $u \in [a, b]$, $K_n(u, t) \equiv 0$, $t \in I_n^*(u)$; or $S_n \subset [\varepsilon, 1 - \varepsilon] \subset (0, 1)$, with $[a, b] \subset [\varepsilon, 1 - \varepsilon]$ for some $0 < \varepsilon < 1/2$.

(K7) For the sequence $\delta_n$ in (K4), $\delta_n \sup_{u \in [a, b]} \Lambda(u; K_n) \rightarrow 0$ and $\Psi(1; K_n) \rightarrow 0$ as $n \uparrow \infty$.

Similarly as in [HM10], in some cases we assume that the kernel function has the form $K_n(u, t) = K(h_n^{-1}(t - u)) h_n^{-1}$ and satisfies the condition

$$
(K_8) \sup_{u \in [a, b]} h_n^{-1} K(s - u) - h_n^{-1} K(t - u) \leq C_n |t - s|^\beta \quad \text{and} \quad |K''(x)| \leq C/|x|
$$

for some constants $C, \beta > 0$ and $|x|$ sufficiently large, and $C_n$ are positive constants such that $\sup_{n \geq 1} C_n < \infty$.

Moreover, Holan and McElroy [HM10] used the following assumptions of Cheng [Che95] on $q(u)$:

(Q1) The quantile density function is twice differentiable on $(0, 1)$.

(Q2) There exists a positive constant $\gamma$ such that $\sup_{u \in (0, 1)} u(1 - u)|J(u)|/fQ(u) \leq \gamma$, where $J$ is the score function in (1).

(Q3) Either $q(0) < \infty$ or $q(u)$ is nonincreasing in some interval $(0, u_*)$, and either $q(1) < \infty$ or $q(u)$ is nondecreasing in some interval $(u^*, 1)$.

We show that the limit matrix $M(a, b, R) := \lim_{n \rightarrow \infty} n^{-1} X' W X$ exists. Let $(v^*, v_0, \ldots, v_{\tilde{p}_1})$ be the first row of $M(a, b, R)^{-1}$, and set $G_R(u) := R(u)(v^* \log u + v_0 + 2 \sum_{k=1}^{\tilde{p}_1} v_k \cos(2\pi k u))$, $i = 0, 1$.

Finally, we assume that the weight function $R$ satisfies the following condition:

(R) $R$ is nonnegative and Riemann integrable on $[a, b]$.

Let $\xrightarrow{P}$ denote convergence in probability, $\xrightarrow{D}$ denote convergence in distribution, and let $N(\mu, \sigma^2)$ stand for the normal distribution with mean $\mu$ and variance $\sigma^2$. Limiting and order relations are always meant as $n \rightarrow \infty$ if not specified otherwise. Our main results are contained in the following two theorems:

**Theorem 1.** Suppose that the conditions (Q1) – (Q3) are satisfied for the quantile density $q(u)$, and $\hat{q}(u)$ is a kernel smoothed estimator with kernel function satisfying (K1) – (K7), the weight function $R$ satisfies the condition (R), and the matrix $M(a, b, R)$ is invertible. Moreover, assume that the percentiles $u_j$ are chosen from a set $[a, b] \subset (0, 1)$ such that $u_j = j/n$, $j = \lfloor na \rfloor, \ldots, \lceil nb \rceil$, and $\tilde{p}_i > p_i$, $i = 0, 1$. Then $\hat{\nu}_i \xrightarrow{P} \nu_i$, $i = 0, 1$.  

4
Theorem 2. Assume that the conditions of Theorem 2 are satisfied, and suppose that the kernel function is symmetric and differentiable on \([-1,1]\), and satisfies the condition \((K_8)\). Suppose that the derivative \(g_R(u) := G'_R(u)\) exists, and \(g_R\) and \(G_R\) are uniformly bounded on \([a,b]\). Let \(h_n\) be a sequence such that \(nh_n^2 \to \infty\), \(nh_n^4 \to 0\) and \(h_n \to 0\), and assume that \(\tilde{p}_i > p_i\), \(i = 0,1\). Then

\[ \sqrt{n}(\tilde{\nu}_i - \nu_i) \overset{D}{\to} N(0,V), \quad i = 0,1, \]

where

\[ V = \int_a^b G^2_R(u)du + \int_a^b \int_a^b G_R(u)G_R(v) \left( 1 + [(u \wedge v) - uv] \frac{q'(u)q'(v)}{q(u)q(v)} \right) dudv. \tag{8} \]

In the special case when the weight function \(R\) is identically 1, the two theorems above reduces to Theorems 1 and 2 of [HM10].

1.2 Comparison of tail index estimators

1.2.1 Asymptotic variances

We evaluate the limiting variance \([8]\) for \(\tilde{p}_0 = 1\), different weight functions and tail indices to compare the WLS and the unweighted (ordinary least squares) estimators in the following submodel of \([4]\):

\[ L_0(u) = \exp \{ 2 \cos(2\pi u) \}, \quad u \in [a,b]. \]

The limiting variances are contained in Table 1. For the calculations we used numerical integration performed by the Wolfram Mathematica software. We see that in some cases the use of the weights makes the asymptotic variance smaller.

Table 1: Limiting variances for different weight functions and tail indices.

<table>
<thead>
<tr>
<th>(\nu_0 = 1.2)</th>
<th>(R(u))</th>
<th>unweighted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1 + \cos u)</td>
<td>(e^{-u})</td>
</tr>
<tr>
<td>(a = 0.1, b = 0.4)</td>
<td>821.232</td>
<td>816.812</td>
</tr>
<tr>
<td>(a = 0.1, b = 0.3)</td>
<td>1512.62</td>
<td>1513.46</td>
</tr>
<tr>
<td>(a = 0.2, b = 0.3)</td>
<td>269523</td>
<td>269655</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(\nu_0 = 1.8)</th>
<th>(R(u))</th>
<th>unweighted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1 + \cos u)</td>
<td>(e^{-u})</td>
</tr>
<tr>
<td>(a = 0.1, b = 0.4)</td>
<td>821.962</td>
<td>819.166</td>
</tr>
<tr>
<td>(a = 0.1, b = 0.3)</td>
<td>1521.58</td>
<td>1523.69</td>
</tr>
<tr>
<td>(a = 0.2, b = 0.3)</td>
<td>267666</td>
<td>267807</td>
</tr>
</tbody>
</table>
\[ \nu_0 = 1.667 \]

<table>
<thead>
<tr>
<th>(\nu_0 = 1.667)</th>
<th>(R(u))</th>
<th>unweighted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 + (u) e^{-u}</td>
<td>- \log u</td>
</tr>
<tr>
<td>(a = 0.1, b = 0.4)</td>
<td>819.423</td>
<td>816.278</td>
</tr>
<tr>
<td>(a = 0.1, b = 0.3)</td>
<td>1516.49</td>
<td>1518.31</td>
</tr>
<tr>
<td>(a = 0.2, b = 0.3)</td>
<td>268011</td>
<td>268151</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(\nu_0 = 2.25)</th>
<th>(R(u))</th>
<th>unweighted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 + (u) e^{-u}</td>
<td>- \log u</td>
</tr>
<tr>
<td>(a = 0.1, b = 0.4)</td>
<td>840.595</td>
<td>838.929</td>
</tr>
<tr>
<td>(a = 0.1, b = 0.3)</td>
<td>1551.91</td>
<td>1555.02</td>
</tr>
<tr>
<td>(a = 0.2, b = 0.3)</td>
<td>266776</td>
<td>266924</td>
</tr>
</tbody>
</table>

### 1.2.2 Simulation results

In order to make a comparison with existing proposals, simulations were done performed by the Matlab software. The samples were generated from the model \(L_0 \equiv 1\) using different tail indices \(\nu_0\). The Hill, Pickands, DEdH (Dekkers, Einmahl and de Haan) and the least squares estimators were included in the simulation study. Similarly as in [HM10], for the simulations we used the Bernstein polynomial estimator of \(Q(u)\). Let \(0 < \varepsilon < 1/2\) be a constant, and assume that \([a, b] \subset [\varepsilon, 1 - \varepsilon]\).

Set \(L_\varepsilon := 1 - 2\varepsilon\) and \(t_j := \varepsilon + (j/k)L_\varepsilon\), \(j = 0, 1, \ldots, k\). The Bernstein polynomial estimator is defined as

\[
\hat{Q}_n^B(u) = \frac{1}{L_\varepsilon^k} \sum_{j=0}^{k-1} \frac{Q_n(t_{j+1}) - Q_n(t_j)}{1/k} \left(\frac{k-1}{j}\right) (u-\varepsilon)^j (1 - \varepsilon - u)^{k-1-j}.
\]

This estimator belongs to the class \((\mathcal{K}_1) - (\mathcal{K}_7)\). We used the values \(k = n = 700\), \(\varepsilon = 0.001\), \(a = 0.001\) and \(b = 0.4\) for the regression estimators, and the weight function \(R(u) = u/300\) for the WLS estimator. Tables 2 and 3 contain the average simulated estimates (mean) and the calculated empirical mean square errors (MSE). We used the sample fraction size \(k_n = 100\) for the Hill, Pickands and DEdH estimators. All the simulations were repeated 200 times. We conclude that in the submodel \(L_0 \equiv 1\) for \(\alpha\) values between 0.8 and 1.5 the WLS estimator has better performance than the OLS estimator. Thus for thinner tails we propose the WLS estimator instead of the OLS estimator. The Hill estimator is the best among the examined estimators. This good performance is not surprising since the Hill estimator was obtained in the special case of \(1 - F(x) = x^{-1/\alpha_1} \ell_1(x)\), \(0 < x < \infty\) when the slowly varying function \(\ell_1(x)\) is constant for all \(x \geq x_{\alpha_1}\), for some threshold \(x_{\alpha_1}\). The Pickands estimator has also good performance. On the other hand, we emphasize that the WLS method can be applied not only for the estimation of the tail index but for the estimation of the slowly varying functions \(L_i\) in (2) and (3).
Table 2: Average simulated tail index estimates (Mean) for sample size \( n = 700 \) and for \( L_0 \equiv 1 \).

<table>
<thead>
<tr>
<th>n = 1, \ldots, N</th>
<th>( \hat{\nu}_W ) Mean</th>
<th>( \hat{\nu}_L ) Mean</th>
<th>( \hat{\nu}_H ) Mean</th>
<th>( \hat{\nu}_D ) Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>1.0674</td>
<td>1.0607</td>
<td>1.0866</td>
<td>1.0761</td>
</tr>
<tr>
<td>700</td>
<td>1.0527</td>
<td>1.0476</td>
<td>1.0438</td>
<td>1.0496</td>
</tr>
</tbody>
</table>

Table 3: Empirical mean square errors (MSE) of tail index estimates for sample size \( n = 700 \) and for \( L_0 \equiv 1 \).

<table>
<thead>
<tr>
<th>n = 1, \ldots, N</th>
<th>( \hat{\nu}_W ) MSE</th>
<th>( \hat{\nu}_L ) MSE</th>
<th>( \hat{\nu}_H ) MSE</th>
<th>( \hat{\nu}_D ) MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>0.00022473</td>
<td>0.00048686</td>
<td>0.00174810</td>
<td>0.00198940</td>
</tr>
<tr>
<td>700</td>
<td>0.00007540</td>
<td>0.00032456</td>
<td>0.00085172</td>
<td>0.00093125</td>
</tr>
</tbody>
</table>

2 Regression estimators for the tail index

This chapter is based on \[ANSV\]. We propose a class of weighted least squares estimators for the tail index of a distribution function with a regularly varying upper tail. Our approach is based on the method developed by Holan and McElroy (2010) for the Parzen tail index. We prove asymptotic normality and consistency for the estimators under suitable assumptions. Through a simulation study, these and earlier estimators are compared in the Pareto and Hall models using the mean squared error as criterion. The results show that the weighted least squares estimator is better than the other estimators investigated.

2.1 Introduction and main result

Let \( X_1, X_2, \ldots \) be independent random variables with a common right-continuous distribution function \( F \), and for each \( n \in \mathbb{N} \), let \( X_1, X_2, \ldots, X_n \) denote the order statistics pertaining to the sample \( X_1, \ldots, X_n \). Let \( R_\alpha \) be the class of all distribution functions \( F \) such that \( 1 - F \) is regularly varying at infinity with index \(-1/\alpha\), that is,

\[
1 - F(x) = x^{-1/\alpha} \ell(x), \quad 1 < x < \infty.
\]
where \( \ell \) is some positive function on the half line \([1, \infty)\), slowly varying at infinity and \( \alpha > 0 \) is a fixed unknown parameter to be estimated. It is well known that \( F \in \mathcal{R}_\alpha \) if and only for some function \( L \) slowly varying at zero,

\[
Q(1-s) = s^{-\alpha}L(s), \quad 0 < s < 1.
\]

The asymptotic normality of Hill estimator was first considered by Hall (1982) \cite{Hal82} in the following submodel of \( \mathcal{R}_\alpha \):

\[
1 - F(x) = x^{-1/\alpha}C_1[1 + C_2x^{-\beta/\alpha}\{1 + o(1)\}], \quad \text{as } x \to \infty,
\]

for some constants \( C_1 > 0 \) and \( C_2 \neq 0 \). This is equivalent to

\[
Q(1-s) = s^{-\alpha}D_1[1 + D_2s^\beta \{1 + o(1)\}], \quad s \to 0,
\]

where \( D_1 = C_1^\alpha \) and \( D_2 = C_2/C_1^\beta \).

Following the idea of Holan and McElroy (2010) \cite{HM10}, we assume that the slowly varying function \( L \) in (9) admits the truncated orthogonal series expansion

\[
L(s) = \exp \left\{ \theta_0 + 2 \sum_{k=1}^{p} \theta_k \cos(2\pi ks) \right\},
\]

where \( p > 0 \) is a fixed integer, and \( \theta_0, \ldots, \theta_p \) are unknown parameters. We suppose that \( p \leq \tilde{p} \), where \( \tilde{p} \) is a prespecified integer. The knowledge of \( p \) is not assumed, condition \( p \leq \tilde{p} \) gives only an upper bound for \( p \). It follows that

\[
\log Q(1-s) = -\alpha \log s + \theta_0 + 2 \sum_{k=1}^{p} \theta_k \cos(2\pi ks).
\]

Let \( Q_n \) be the empirical quantile function defined as

\[
Q_n(s) = X_{k,n} \quad \text{if} \quad \frac{k-1}{n} < s \leq \frac{k}{n}, \quad k = 1, 2, \ldots, n.
\]

Based on the representation (11), we obtain the regression equations

\[
\log Q_n(1-s_j) = -\alpha \log s_j + \theta_0 + 2 \sum_{k=1}^{\tilde{p}} \theta_k \cos(2\pi ks_j) + \varepsilon(s_j),
\]

where

\[
\varepsilon(s) = \log \left( Q_n(1-s)/Q(1-s) \right)
\]

is the residual process, \( s_j = j/n, \; j = [na], \ldots, [nb], \; a < b \) are fixed constants taken from the interval \((0,1)\), and \( \theta_k = 0 \) for \( k > p \). The value \( \tilde{p} \) is chosen by the statistician. We propose a class of estimators for \( \alpha \) using weighted least squares. We choose some nonnegative weights of the form \( w_{j,n} = R(s_j) \) with some weight function \( R \). Set

\[
y_j := \log Q_n(1-s_j),
\]

\[
y := (y_{[na]}, \ldots, y_{[nb]})',
\]

8
\[ W := \text{diag}(w_{[na],n}, \ldots, w_{[nb],n}), \]

and let \( X := [G^*, G_0, 2G_1, \ldots, 2G_{\tilde{p}}], \) where
\[
G^* = ( -\log(s_{[na]}), \ldots, -\log(s_{[nb]}))', \quad G_k = ( \cos(2\pi ks_{[na]}), \ldots, \cos(2\pi ks_{[nb]}))', \quad k = 0, \ldots, \tilde{p}.
\]

Set \( \beta_{\tilde{p}} := (\alpha, \theta_0, \theta_1, \ldots, \theta_{\tilde{p}})' \). By minimizing the weighted sum of squares
\[
\sum_{j=0}^{\tilde{p}} w_{j,n} (y_j + \alpha \log s_j - \theta_0 - 2 \sum_{k=1}^{\tilde{p}} \theta_k \cos(2\pi ks_j))^2,
\]
we obtain the following estimator of \( \beta_{\tilde{p}} \):
\[
\hat{\beta}_{\tilde{p}} = (X'WX)^{-1}X'y.
\]

Then the weighted least squares estimator of \( \alpha \) can be written in the form
\[
\hat{\alpha}_n^{(W)} := e'_1\hat{\beta}_{\tilde{p}} = e'_1(X'WX)^{-1}X'y,
\]
where \( e_1 \) is the \( \tilde{p} + 2 \) dimensional vector defined as \( e_1 = (1, 0, 0, \ldots, 0)' \).

We assume the following conditions on the underlying distribution:

\( (Q_1) \) The distribution function \( F \) is continuous and twice differentiable on \((a^*, b^*)\), where \( a^* = \sup \{ x : F(x) = 0 \} \), \( b^* = \inf \{ x : F(x) = 1 \} \), \(-\infty \leq a^* < b^* \leq \infty \) and \( f(x) := F'(x) \neq 0 \) on \((a^*, b^*)\).

\( (Q_2) \) \( \sup_{a^* < x < b^*} F(x)(1 - F(x))|f'(x)|/f^2(x) | < \infty. \)

\( (Q_3) \) \( \sup_{1 - b \leq s \leq 1 - a} 1/|Q(s)| < \infty, \sup_{1 - b \leq s \leq 1 - a} 1/fQ(s) < \infty \) and \( \sup_{1 - b \leq s \leq 1 - a} 1/|fQ(s)Q(s)| < \infty. \)

We show that the limit matrix \( M(a, b, R) := \lim_{n \to \infty} n^{-1}X'WX \) exists. Let \((v^*, v_0, \ldots, v_{\tilde{p}})\) be the first row of \( M(a, b, R)^{-1} \), and set \( G_R(u) := R(u)(-v^* \log u + v_0 + 2 \sum_{k=1}^{\tilde{p}} v_k \cos(2\pi ku)) \) for \( u \in (0, 1) \).

Moreover, we suppose the following conditions:

\( (R) \) The weight function \( R \) is nonnegative and Riemann integrable on \([a, b]\).

\( (M) \) The matrix \( M(a, b, R) \) is invertible.

Theorem 3. Assume that the conditions \( Q_1 - Q_5 \) are satisfied for the underlying distribution and suppose that the quantile function \( Q \) admits the representation (14). Moreover, assume the conditions \( R \) and \( M \), and assume also that the percentiles \( s_j \) are chosen from a closed set \( U = [a, b], 0 < a < b < 1, \) such that \( s_j = j/n, j = [na], \ldots, [nb], \) and \( p \leq \tilde{p} \). Then
\[
\sqrt{n} (\hat{\alpha}_n^{(W)} - \alpha) \xrightarrow{D} N(0, V),
\]
where
\[
V = \int_a^b \int_a^b G_R(s)G_R(t)\left( (1 - s) \wedge (1 - t) - (1 - s)(1 - t) \right) \frac{Q(1 - s)Q(1 - t)fQ(1 - s)fQ(1 - t)}{Q(1 - s)Q(1 - t)} \, ds \, dt.
\]
2.2 Asymptotics for $\tilde{p} \to \infty$

The estimation method proposed in previous section is heavily based on the assumption $p \leq \tilde{p}$. However, choosing $\tilde{p} < p$ incurs a bias. To overcome this difficulty, we adjust our method to study asymptotics when $\tilde{p} \to \infty$. In this section our investigation is based on the following series expansion:

$$\log L(s) \sim \sum_{k=0}^{\infty} \theta_k \varphi_k(s),$$

where

$$\varphi_0(s) = \frac{1}{\sqrt{(b-a)R(s)}},$$

$$\varphi_k(s) = \cos \left( \frac{\pi k}{b-a} \frac{s-a}{b-a} \right) \frac{1}{\sqrt{(b-a)R(s)/2}}, \quad k = 1, 2, \ldots,$$

and $\theta_k = \int_a^b \log L(x) \varphi_k(x) R(x) dx$. The sequence $\varphi_k \sqrt{R}$, $k = 0, 1, \ldots$, is a complete orthonormal system in $L^2[a, b]$. For convenience, in this section we use the percentiles $s_j = a + \frac{j}{n} (b-a)$, $j = 0, \ldots, n-1$. Similarly as in previous section, with $y_j := \log Q_n(1 - s_j)$ and $w_{j,n} = R(s_j)$ define

$$y := (y_0, \ldots, y_{n-1})',$$

$$W := \text{diag}(w_{0,n}, \ldots, w_{n-1,n}),$$

and let $X := [G^*, G_0, G_1, \ldots, G_{\tilde{p}}]$, where

$$G^* = (- \log s_0, \ldots, - \log s_{n-1})',$$

$$G_k = (\varphi_k(s_0), \ldots, \varphi_k(s_{n-1})), \quad k = 0, \ldots, \tilde{p}. \quad (16)$$

Set

$$b_{\tilde{p}}(s) := \log L(s) - \sum_{k=0}^{\tilde{p}} \theta_k \varphi_k(s). \quad (17)$$

Recall [12]. Then we have

$$\log Q_n(1 - s_j) = -\alpha \log s_j + \sum_{k=1}^{\tilde{p}} \theta_k \varphi_k(s_j) + b(s_j) + \varepsilon(s_j).$$

By minimizing the weighted sum of squares

$$\sum_{\lfloor na \rfloor}^{\lfloor nb \rfloor} w_{j,n} (y_j + \alpha \log s_j - \sum_{k=0}^{\tilde{p}} \theta_k \varphi_k(s_j))^2,$$

we obtain the following estimator of $\alpha$:

$$\hat{\alpha}_n^{(W)} = c'_1 (X'W X)^{-1} X'W y.$$
In order to formulate the result for $\hat{\alpha}_n^{(W)}$, we need the series expansion of the $-\log(\cdot)$ function:

$$-\log s \sim \sum_{j=0}^{\infty} c_j \varphi_j(s),$$

(18)

where $c_j = \int_a^b (-\log x) \varphi_j(x) R(x) dx$. We assume the following conditions on the sequences $\tilde{p}, \theta_n$ and $c_n$:

- $(P_1)$ $\tilde{p} \to \infty$ and $\tilde{p}/n \to 0$.
- $(P_2)$ For each $n$, $3(\tilde{p} + 1)/n < 1$.
- $(P_3)$ $n \sum_{i=\tilde{p}+1}^{\infty} c_i^2 \to \infty$.
- $(P_4)$ $\theta_n/c_n \to 0$.

Theorem 4. Suppose the conditions $(P_1) - (P_4)$ are satisfied. Then $\hat{\alpha}_n^{(W)} \overset{P}{\to} \alpha$.

2.3 Simulation results

In order to make a comparison with existing proposals, simulations were performed by the Matlab software. The samples were generated from the strict Pareto model $L \equiv 1$ in (9) and from the Hall model (10). The Hill, Pickands, DEdH (Dekkers, Einmahl and de Haan) and the weighted least squares (WLS) estimators were included in the simulation study. We used the values $n = 5000$, $a = 0.001$, $b = 0.4$ and $\tilde{p} = 1, 2, 3$, and the weight function $R(s) = s/500$ for the WLS estimator. In case of $R \equiv 1$, we refer to as ordinary least squares (OLS) estimator. The tail indexes were chosen between 0.5 and 20. For the Hill, Pickands and DEdH estimators the simulations were done for sample size $n = 5000$ and sample fraction size $k_n = 200$. All the simulations were repeated 1000 times.

Tables 4 and 5 contains the empirical mean square errors (MSE) and the average simulated estimates (mean) for the strict Pareto model. We conclude that in the submodel $L \equiv 1$ for all $\alpha$ values, the WLS estimator performs better than the other estimators investigated.

Tables 6 and 7 presents the simulation results for the Hall model. Specifically, we used the parameters $D_1 = 0.4$, $D_2 = 1$ and $\beta = 0.01$. We see from Table 6 that the WLS estimator performs better than the other estimators, and the OLS estimator is competitive with the Hill estimator especially for $\tilde{p} = 3$.

Given the values of $[a, b]$, which determines the number of values taken from the simulation data, we experimented with some expanding intervals to find an appropriate range, and we stop when we obtain reasonable stability of the estimator of $\alpha$. Figure 1 shows the tail index estimates for WLS approach for different values of $(a)$ for the Pareto distribution with $\alpha = 1.8$ (left panel) and the $\alpha = 5$ (right panel), the values of the remaining $\alpha$ with both Pareto distribution and Hall model give fairly
Table 4: Empirical mean square errors (MSE) of tail index estimates for the Pareto model and for sample size $n = 5000$.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>WLS</th>
<th>OLS</th>
<th>Hill</th>
<th>Pickands</th>
<th>DeEdh</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p = 1$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>0.001485</td>
<td>0.002183</td>
<td>0.002017</td>
<td>0.002157</td>
<td>0.002356</td>
</tr>
<tr>
<td>0.7</td>
<td>0.001517</td>
<td>0.00253</td>
<td>0.002453</td>
<td>0.002732</td>
<td>0.002956</td>
</tr>
<tr>
<td>1</td>
<td>0.001548</td>
<td>0.002673</td>
<td>0.002618</td>
<td>0.002881</td>
<td>0.003121</td>
</tr>
<tr>
<td>1.2</td>
<td>0.00158</td>
<td>0.00282</td>
<td>0.002755</td>
<td>0.002998</td>
<td>0.003253</td>
</tr>
<tr>
<td>1.5</td>
<td>0.00162</td>
<td>0.00301</td>
<td>0.00295</td>
<td>0.00319</td>
<td>0.003435</td>
</tr>
<tr>
<td>2</td>
<td>0.00166</td>
<td>0.00321</td>
<td>0.00297</td>
<td>0.003276</td>
<td>0.003536</td>
</tr>
<tr>
<td>3</td>
<td>0.00172</td>
<td>0.00344</td>
<td>0.00312</td>
<td>0.003396</td>
<td>0.003701</td>
</tr>
</tbody>
</table>

Table 5: Average simulated tail index estimates (Mean) for sample size $n = 5000$ and for the Pareto model.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>WLS</th>
<th>OLS</th>
<th>Hill</th>
<th>Pickands</th>
<th>DeEdh</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p = 1$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>0.001823</td>
<td>0.002738</td>
<td>0.002651</td>
<td>0.002823</td>
<td>0.002973</td>
</tr>
<tr>
<td>0.7</td>
<td>0.00203</td>
<td>0.00308</td>
<td>0.003001</td>
<td>0.003201</td>
<td>0.003376</td>
</tr>
<tr>
<td>1</td>
<td>0.00216</td>
<td>0.00326</td>
<td>0.00317</td>
<td>0.003361</td>
<td>0.003556</td>
</tr>
<tr>
<td>1.2</td>
<td>0.00229</td>
<td>0.00345</td>
<td>0.00334</td>
<td>0.003544</td>
<td>0.003776</td>
</tr>
<tr>
<td>1.5</td>
<td>0.00242</td>
<td>0.00367</td>
<td>0.00347</td>
<td>0.003675</td>
<td>0.003936</td>
</tr>
<tr>
<td>2</td>
<td>0.00256</td>
<td>0.00393</td>
<td>0.00356</td>
<td>0.00382</td>
<td>0.00409</td>
</tr>
<tr>
<td>3</td>
<td>0.00271</td>
<td>0.00423</td>
<td>0.00358</td>
<td>0.00397</td>
<td>0.004315</td>
</tr>
</tbody>
</table>

Table 6: Empirical mean square errors (MSE) of tail index estimates for the Hall model and for sample size $n = 5000$.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>WLS</th>
<th>OLS</th>
<th>Hill</th>
<th>Pickands</th>
<th>DeEdh</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p = 1$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>0.001823</td>
<td>0.002738</td>
<td>0.002651</td>
<td>0.002823</td>
<td>0.002973</td>
</tr>
<tr>
<td>0.7</td>
<td>0.00203</td>
<td>0.00308</td>
<td>0.003001</td>
<td>0.003201</td>
<td>0.003376</td>
</tr>
<tr>
<td>1</td>
<td>0.00216</td>
<td>0.00326</td>
<td>0.00317</td>
<td>0.003361</td>
<td>0.003556</td>
</tr>
<tr>
<td>1.2</td>
<td>0.00229</td>
<td>0.00345</td>
<td>0.00334</td>
<td>0.003544</td>
<td>0.003776</td>
</tr>
<tr>
<td>1.5</td>
<td>0.00242</td>
<td>0.00367</td>
<td>0.00347</td>
<td>0.003675</td>
<td>0.003936</td>
</tr>
<tr>
<td>2</td>
<td>0.00256</td>
<td>0.00393</td>
<td>0.00356</td>
<td>0.00382</td>
<td>0.00409</td>
</tr>
<tr>
<td>3</td>
<td>0.00271</td>
<td>0.00423</td>
<td>0.00358</td>
<td>0.00397</td>
<td>0.004315</td>
</tr>
</tbody>
</table>
Table 7: Average simulated tail index estimates (Mean) for sample size \( n = 5000 \) and for the Hall model.

<table>
<thead>
<tr>
<th>( \alpha )</th>
<th>Mean</th>
<th>WLS</th>
<th>OLS</th>
<th>Hill</th>
<th>Pickands</th>
<th>DEdh</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \tilde{p} = 1 )</td>
<td>( \tilde{p} = 2 )</td>
<td>( \tilde{p} = 3 )</td>
<td>( \tilde{p} = 1 )</td>
<td>( \tilde{p} = 2 )</td>
<td>( \tilde{p} = 3 )</td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>0.49603</td>
<td>0.49630</td>
<td>0.49763</td>
<td>0.49657</td>
<td>0.49032</td>
<td>0.45841</td>
</tr>
<tr>
<td>0.8</td>
<td>0.797</td>
<td>0.79759</td>
<td>0.79873</td>
<td>0.79534</td>
<td>0.79959</td>
<td>0.79268</td>
</tr>
<tr>
<td>1.2</td>
<td>1.19672</td>
<td>1.19877</td>
<td>1.20103</td>
<td>1.19932</td>
<td>1.19436</td>
<td>1.18723</td>
</tr>
<tr>
<td>1.5</td>
<td>1.49791</td>
<td>1.5004</td>
<td>1.50308</td>
<td>1.49961</td>
<td>1.49485</td>
<td>1.48868</td>
</tr>
<tr>
<td>1.8</td>
<td>1.80074</td>
<td>1.80301</td>
<td>1.80587</td>
<td>1.80227</td>
<td>1.80175</td>
<td>1.79617</td>
</tr>
<tr>
<td>2</td>
<td>2.00136</td>
<td>2.00379</td>
<td>2.00781</td>
<td>1.99959</td>
<td>1.99945</td>
<td>1.98366</td>
</tr>
<tr>
<td>3</td>
<td>2.99923</td>
<td>2.99734</td>
<td>3.00277</td>
<td>3.00226</td>
<td>2.99773</td>
<td>2.99178</td>
</tr>
<tr>
<td>4</td>
<td>4.00142</td>
<td>4.00012</td>
<td>4.00452</td>
<td>4.00426</td>
<td>3.99827</td>
<td>3.98143</td>
</tr>
<tr>
<td>5</td>
<td>5.00308</td>
<td>5.00224</td>
<td>5.00465</td>
<td>5.00425</td>
<td>4.99837</td>
<td>4.98147</td>
</tr>
<tr>
<td>5.5</td>
<td>5.51692</td>
<td>5.51875</td>
<td>5.52059</td>
<td>5.51671</td>
<td>5.51191</td>
<td>5.49461</td>
</tr>
<tr>
<td>6</td>
<td>6.00572</td>
<td>6.00401</td>
<td>6.00637</td>
<td>6.00559</td>
<td>5.99775</td>
<td>5.98241</td>
</tr>
<tr>
<td>15</td>
<td>14.99548</td>
<td>15.01536</td>
<td>15.0485</td>
<td>15.0713</td>
<td>15.0849</td>
<td>15.0399</td>
</tr>
</tbody>
</table>

Figure 1: Tail index estimates for WLS approach with Pareto distribution in (left panel) from \( \alpha = 1.8 \) and in (right panel) from \( \alpha = 5 \).

### 3 Application

The results presented in this chapter are based on [IAN20, IAND20]. We study the prevalence of the COVID-19 pandemic in Iraq and Egypt using a generalised (SEIR) compartmental mathematical model, a logistic regression model, and a simple Gaussian model. The extreme value theory approach for finding and modeling Covid-19 peaks was studied, and one of the prime successes EVT is the return level idea.

#### 3.1 Forecast of the COVID–19 spread in Iraq and Egypt

The logistic growth takes the form:

\[
C(t) = \frac{K}{1 + be^{-rt}},
\]

where \( r > 0 \) is the rate of infection, \( K > 0 \) is the final epidemic size and \( b = \frac{K - C_0}{C_0} \) and \( C_0 \) is the initial population. Figure 2 shows the logistics growth model (19) fitted to
in (left panel) the cumulative number of infected cases from Iraq and in (right panel) the cumulative number of infected cases from Egypt with parameters given in Table 8. We note that the logistic model fitted the incidence data with a root mean square error (RMSE) of 5, 229.7, $R^2$ of 0.9981 for Iraq data and with (RMSE) of 1, 924.4, $R^2$ of 0.9980 for Egypt data, as shown in Tables 8. The logistic model gives a reasonable good fit for both countries.

Table 8: Estimated parameter results of the logistics model (19) to Iraq and Egypt.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Iraq</th>
<th>Egypt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated epidemic size $K$ (cumulative cases)</td>
<td>$490,900$ (478300, 503600)</td>
<td>$105,000$ (104500, 105900)</td>
</tr>
<tr>
<td>Growth Rate $r$</td>
<td>$0.03787$ (0.03685, 0.03889)</td>
<td>$0.05634$ (0.05546, 0.05721)</td>
</tr>
<tr>
<td>Estimated start of ending phase date</td>
<td>05/05/2021</td>
<td>04/11/2020</td>
</tr>
<tr>
<td>Goodness of fit ($R^2$)</td>
<td>$0.9981$</td>
<td>$0.9980$</td>
</tr>
<tr>
<td>Root Mean Square Error (RMSE)</td>
<td>5, 229.7</td>
<td>1, 924.4</td>
</tr>
</tbody>
</table>

We employed a simple Gaussian model, to model the time-dependent daily change of infections. Let $I(t)$ denotes the time-dependent Gaussian function and takes the following form:

$$I(t) = I_0 e^{-\left(\frac{t-\mu}{\sigma}\right)^2},$$

where $I_0$ denotes the maximum value at time $\mu$ and $\sigma$ controls the width. The Gaussian model was fitted to data from Iraq and Egypt with reproduction numbers 1.0659 and 1.0318, respectively. Figure 3 shows the Gaussian model fitted to in (left panel) the daily number of confirmed cases from Iraq, and in (right panel) the daily number of confirmed cases from Egypt with parameters given in Table 9. The model fits the actual data well with a root mean square error (RMSE) of 335.607, $R^2$ of 0.9614 for Iraq data and with (RMSE) of 110.33, $R^2$ of 0.9528 for Egypt data, as listed in Tables 9.
Figure 3: The Gaussian model fitted to the daily confirmed cases in Iraq (Left panel) and in Egypt (right panel).

Table 9: Estimated parameter results of the Gaussian model to Iraq and Egypt.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Iraq</th>
<th>Egypt</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_0$</td>
<td>$1.0659$</td>
<td>$1.0318$</td>
</tr>
<tr>
<td>$C.1_{0.95}$</td>
<td>(4161, 4347)</td>
<td>(1493, 1574)</td>
</tr>
<tr>
<td>Estimated peak day cases $I_0$</td>
<td>4,254</td>
<td>1,534</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>80.16</td>
<td>34.99</td>
</tr>
<tr>
<td>Estimated peak date</td>
<td>$14/09/2020$</td>
<td>$16/06/2020$</td>
</tr>
<tr>
<td>Goodness of fit ($R^2$)</td>
<td>0.9614</td>
<td>0.9528</td>
</tr>
<tr>
<td>Root Mean Square Error (RMSE)</td>
<td>335.607</td>
<td>110.33</td>
</tr>
</tbody>
</table>

3.2 Compartmental model for COVID–19 transmission

We split the human population into seven compartments: susceptible $S(t)$, exposed $E(t)$, symptomatically infected $I_s(t)$, mildly infected $I_m(t)$, treated $H(t)$, recovered individuals $R(t)$, and $D(t)$ is the individuals who lose their lives due to the COVID–19. Hence, we consider the following SEIR model:

$$
S'(t) = -\frac{\beta_e E(t) + \beta_m I_m(t) + I_s(t) + \beta_h H(t)}{N(t) - D(t)} S(t),
$$

$$
E'(t) = \frac{\beta_e E(t) + \beta_m I_m(t) + I_s(t) + \beta_h H(t)}{N(t) - D(t)} S(t) - \nu E(t),
$$

$$
I_s'(t) = (1 - \theta)\nu E(t) + \sigma I_m(t) - \sigma_s I_s(t) - \delta_s I_s(t),
$$

$$
I_m'(t) = \theta \nu E(t) - \sigma_m I_m(t) - \sigma I_m(t),
$$

$$
H'(t) = \sigma_s I_s(t) - \sigma_h H(t) - \delta_h H(t),
$$

$$
R'(t) = \sigma_m I_m(t) + \sigma_h H(t),
$$

$$
D'(t) = \delta_s I_s(t) + \delta_h H(t).
$$

Figure 4 shows the model (20) fitted to the daily number of confirmed cases in (left panel) from Iraq, 22 February 2020 until 08 October 2020, and in (right panel) from Egypt, 05 March 2020 until 08 October 2020. Our model gives a reasonable good fit for both countries, predicting the peak in Iraq and showing the peak in Egypt. The fitting parameter results are listed in Table 10.
3.3 Prediction of the second wave of the COVID-19 epidemic

We assume that the observations are independent and identically distributed with common cdf $F$. For $y > u$, $F(y)$ is estimated, by $\hat{F}(u) = 1 - \hat{\zeta}_u (1 - \hat{G}(y - u))$, where $\hat{G}$ is the GPD and $\hat{\zeta}_u$ the empirical estimator of observations that exceed the threshold $u$. The return level estimate is the level expected to be exceeded by the maximum of $n$ observations with probability $1 - \alpha$ is estimated by $\hat{y}_\alpha$ of $\hat{F}(y)^n$. If $\gamma \neq 0$, we obtain $\hat{y}_\alpha$ as

$$\hat{y}_\alpha = \frac{\hat{\sigma}}{\gamma} \left[ \frac{1}{\hat{\zeta}_u} (1 - \alpha^{1/n}) - \gamma - 1 \right] + u \quad (21)$$

The mean excess function of $X$ denote the mean residual life function is

$$e(u) = E(X - u \mid X > u), \quad 0 \leq u < x^\ast. \quad (22)$$

The generalized Pareto distribution (GPD) of two-parameter was used to model exceedances over a threshold, the Maximum likelihood estimators was preferred, the
estimated parameters are gamma, sigma of the GPD, where $\gamma = -0.616$ and $\sigma = 686.19$ for Iraq and $\gamma = -0.648$ and $\sigma = 316.796$ for Egypt. Figure 5 shows pick the suitable threshold $u$ for infections, which are 4000 and 1300 for COVID-19 data in Iraq and Egypt, respectively, which gave two corresponding observations: 35 and 37 over the threshold. Hence the estimate of the exceedance probability $\hat{\xi}_u = 0.1003$ for Iraq and $\hat{\xi}_u = 0.1039$ for Egypt. Moreover, the mean excess plot with a downwards sloping line indicated thin tailed behaviour with $\gamma < 0$. We focus on estimate the return level during the following year and the following two years with two value of probability 0.1 and 0.01. These estimates were computed using Equation (21). The results indicate that there is a possibility 0.1 that the infection cases will exceed 5083 once during the next year and 5107 within two years for Iraq, while in Egypt the epidemic will exceed 1788 during the two years with probability 0.01, all results are presented in table 11.

Table 11: Estimated levels that the maximum of COVID-19 epidemic will exceed with probability 0.1 and 0.01 for the one year and two years for Iraq and Egypt.

<table>
<thead>
<tr>
<th>Probability</th>
<th>One year</th>
<th>Two year</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1 - $\alpha$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iraq</td>
<td>5083</td>
<td>5107</td>
</tr>
<tr>
<td>Egypt</td>
<td>1778</td>
<td>1787</td>
</tr>
</tbody>
</table>

References

[ANV20] A. Al-Najafi and L. Viharos. Weighted least squares estimators for the

[BGT89] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular variation*, vol-
ume 27 of *Encyclopedia of Mathematics and its Applications*. Cambridge

[Che95] C. Cheng. Uniform consistency of generalized kernel estimators of quantile

[Hal82] Peter Hall. On some simple estimates of an exponent of regular variation.


[IAN20] M.A. Ibrahim and A. Al-Najafi. Modeling, control, and prediction of the
spread of covid-19 using compartmental, logistic, and gauss models: A case

[IAND20] M.A. Ibrahim, A. Al-Najafi, and A. Dénes. Predicting the covid-19 spread
using compartmental model and extreme value theory with application to
Egypt and Iraq. *in press, Trends in Biomathematics: Chaos and Control


[Par04] Emanuel Parzen. Quantile probability and statistical data modeling.