STA4507H1S Problem Set

Name: Tian Han Guan

Student ID 998478058

1. a) First consider the joint distribution function $G(\underline{x})$ of $(X_{(1)}, \ldots, X_{(n)})$,
where $G(\underline{x}) = \int g(\underline{x})\, d\underline{x} = \int \cdots \int g(x_1, \ldots x_n)\, dx_1 \cdots dx_n$, $x_1 < \cdots < x_n$

By definition, $G(\underline{x}) = P[\, X_{(1)} \leq x_1,\ X_{(2)} \leq x_2, \ldots, X_{(n)} \leq x_n \,]$, $x_1 < x_2 < \cdots < x_n$,

$= P[\, y_1 < X_{(1)} \leq x_1,\ y_2 < X_{(2)} \leq x_2, \ldots, y_n < X_{(n)} \leq x_n \,]$,

where $y_1 < x_1 < y_2 < x_2 < \cdots < y_n < x_n$ for some $y_1, \ldots, y_n$.

$$= P[\, \underbrace{X_{(1)} \in (y_1, x_1),\ X_{(2)} \in (y_2, x_2), \ldots, X_{(n)} \in (y_n, x_n)}_{\text{this is a union of disjoint events, let } \pi \text{ denote the permutations of this union of events}} \,]$$

$= P[\, \bigcup_{\pi} \{ X_{\pi(1)} \in (y_1, x_1), \ldots, X_{\pi(n)} \in (y_n, x_n) \} \,]$, where X_{π} is a permutation of the elements of $\underline{x}$

$= \sum_{\pi} P[\{ X_{\pi(1)} \in (y_1, x_1), \ldots, X_{\pi(n)} \in (y_n, x_n) \}]$, as it is disjoint union of events

$= \sum_{\pi} \int_{y_1}^{x_1} \cdots \int_{y_n}^{x_n} f(t_1, \ldots, t_n)\, dt_n \cdots dt_1 \quad (*)$

Differentiate with respect to $x_1, \ldots, x_n$ of both sides of the equation (*)

$$G(\underline{x}) = \sum_{\pi} \int_{y_1}^{x_1} \cdots \int_{y_n}^{x_n} f(t_1 \ldots t_n)\, dt_n \ldots dt_1 \quad (*)$$

*by Fundamental Thm of Calculus, $\frac{d}{dx_i}\int_a^{x_i} f(t_i)\,dt_i = f(x_i)$,

here the high dimension case can use Lebesgue Differentiation Thm as a generalization

$$\Rightarrow \quad g(\underline{x}) = \sum_{\pi} f(\pi \underline{x}) \quad \text{as required} \quad \square$$

b) First, show that $g_n(x) = \int \cdots \int g(x_1, \ldots x_{n-1}, x)\, dx_{n-1} \cdots dx_1$

$$= \frac{n!}{(n-1)!} \int \cdots \int f(x_1, \ldots, x_{n-1}, x)\, dx_{n-1} \cdots dx_1$$

~~proof~~ $g(x_1, \ldots, x_{n-1}, x)$

$= \binom{n}{n-1} f(x_1, \ldots, x_{n-1}, x)$, as given the realized value of $X_{(n)} = x$,

$= \frac{n!}{(n-1)!} f(x_1, \ldots, x_{n-1}, x)$. just need to select the rest $X_{(1)} \ldots X_{(n-1)}$ r.v.'s from the n r.v.'s

[ie. given $X_{(n)}$, we need to select $X_{(1)}, \ldots, X_{(n-1)}$ from the set of $\{X_1, \ldots, X_n\}$ $\Rightarrow$ which has $\binom{n}{n-1}$ ways to choose from]

$$\Rightarrow \int \cdots \int g(x_1, \ldots, x_{n-1}, x)\, dx_{n-1} \cdots dx_1 = \frac{n!}{(n-1)!} \int \cdots \int f(x_1, \ldots, x_{n-1}, x)\, dx_{n-1} \cdots dx_1$$

$\square$

Now, $\frac{n!}{(n-1)!} \int \cdots \int f(x_1, \ldots, x_{n-1}, x)\, dx_{n-1} \cdots dx_1$

$= n \int_{-\infty}^{x} \cdots \int_{-\infty}^{x} f(x_1, \ldots, x_{n-1}, x)\, dx_{n-1} \cdots dx_1$, x is fixed

$= n \int_{-\infty}^{x} \cdots \int_{-\infty}^{x} f(x_1, \ldots, x_{n-1} \mid X_n = x) \underbrace{f(x)}_{} \, dx_{n-1} \cdots dx_1$, $f(x) \overset{\shortparallel}{} P(X_n = x)$

$= n f(x) \int_{-\infty}^{x} \cdots \int_{-\infty}^{x} f(x_1, \ldots, x_{n-1} \mid X_n = x)\, dx_{n-1} \cdots dx_1$

$= n f(x)\, P\{X_1 \le x, \ldots, X_{n-1} \le x \mid X_n = x\}$ as required $\square$.

2. a). given $\lim_{x \to \infty} g(x) = a > 0$,

for any $t > 0$, $\lim_{x \to \infty} \left[\frac{g(tx)}{g(x)} \right] = \frac{\lim_{x \to \infty} g(tx)}{\lim_{x \to \infty} g(x)}$, provided that $\lim_{x \to \infty} g(x) = a > 0$

$= \frac{a}{a}$, as $tx \to \infty$ (since $t > 0$)

$= 1$

$\therefore g(x)$ is a slowly varying function $\square$

b). For any $t>0$, $a>0$, $\lim_{x\to\infty} \frac{[L(tx)]^a}{[L(x)]^a}$

$$= \lim_{x\to\infty} \left[\frac{L(tx)}{L(x)}\right]^a$$

$$= \left[\lim_{x\to\infty} \frac{L(tx)}{L(x)}\right]^a, \text{ by property of limits}$$

$$= 1^a, \text{ as } L(x) \text{ is slowly varying}$$

$$= 1.$$

$\therefore$ $[L(x)]^a$ is also slowly varying for any $a>0$ □

c). Assume $L_1(x)$ and $L_2(x)$ are positive and continuous functions,
$L_1: (0,\infty) \to (0,\infty)$, $L_2: (0,\infty) \to (0,\infty)$

Use the Karamata Representation of Slowly Varying Functions

① $L_1(x) = C_1(x) \exp\left\{\int_{x_0}^{x} \frac{\varepsilon_1(t)}{t} dt\right\}$.

for some $C_1(x)$, $\varepsilon_1(t)$ s.t. $\lim_{x\to\infty} C_1(x) = C_0 > 0$ (constant), $\lim_{x\to\infty} \varepsilon_1(x) = 0$, some $x_0 > 0$

② $L_2(x) = C_2(x) \exp\left\{\int_{y_0}^{x} \frac{\varepsilon_2(t)}{t} dt\right\}$,

for some $C_2(x)$, $\varepsilon_2(t)$ s.t. $\lim_{x\to\infty} C_2(x) = C_{2_0} > 0$, $\lim_{x\to\infty} \varepsilon_2(x) = 0$, some $y_0 > 0$.

→ Then, the composite $L(x) = L_2[L_1(x)]$

$$= C_2[L_1(x)] \exp\left\{ \int_{y_0}^{L_1(x)} \frac{\varepsilon_2(t)}{t}\, dt \right\}.$$

To show $L(x)$ is also a slowly varying function, for all $a>0$, trying to show that

$$\lim_{x\to\infty}\left[\frac{L(ax)}{L(x)}\right] = 1$$

$$\to \lim_{x\to\infty}\left[\frac{L(ax)}{L(x)}\right] = \lim_{x\to\infty}\left[\frac{C_2(L_1(ax))\exp\left\{\int_{y_0}^{L_1(ax)} \frac{\varepsilon_2(t)}{t}\, dt\right\}}{C_2(L_1(x))\exp\left\{\int_{y_0}^{L_1(x)} \frac{\varepsilon_2(t)}{t}\, dt\right\}}\right]$$

$$= \underbrace{\left[\lim_{x\to\infty}\frac{C_2(L_1(ax))}{C_2(L_1(x))}\right]}_{(*)}\left[\lim_{x\to\infty}\frac{\exp\left\{\int_{y_0}^{L_1(ax)} \frac{\varepsilon_2(t)}{t}\, dt\right\}}{\exp\left\{\int_{y_0}^{L_1(x)} \frac{\varepsilon_2(t)}{t}\, dt\right\}}\right]$$

Note: For (*):

Case 1: as $x\to\infty$, $L_1(x)=a>0$, then $(*) = \frac{C_2(a)}{C_2(a)} = 1$

Case 2: as $x\to\infty$, $L_1(x)\to\infty$, then $(*) = \frac{C_{20}}{C_{20}} = 1$ as $\lim_{x\to\infty} C_2(x) = C_{20} > 0$

$$\therefore \boxed{(*) = 1}$$

$$= \lim_{x\to\infty}\left[\exp\left\{\underbrace{\int_{y_0}^{L_1(ax)} \frac{\varepsilon_2(t)}{t}\, dt - \int_{y_0}^{L_1(x)} \frac{\varepsilon_2(t)}{t}\, dt}_{(**)}\right\}\right]$$

Assume $a>0$ s.t. $L_1(ax) > L_1(x)$, then $(**) =$ area under curve

$$= \lim_{x\to\infty}\left[\exp\left\{\int_{L_1(x)}^{L_1(ax)} \frac{\varepsilon_2(t)}{t}\right\}\right]$$

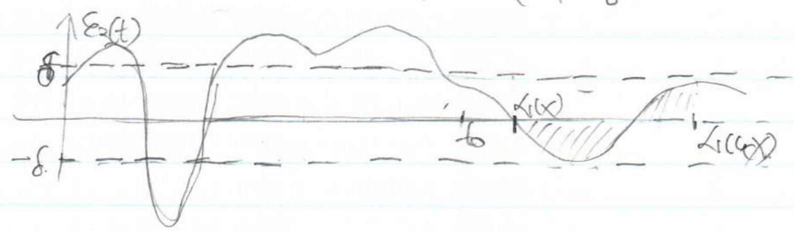
(If $L_1(x) > L_1(ax)$, just add a $(-)$ sign)

→ Consider only on $\int_{L_1(x)}^{L_1(ax)} \frac{\varepsilon_2(t)}{t}\,dt$.

→ Now use the fact that $\lim_{t\to\infty} \varepsilon_2(t) = 0$

$$\Updownarrow$$

for any $\delta > 0$, $\exists$ a t_0 s.t. $\forall\ t > t_0$, $|\varepsilon_2(t)| < \delta$.



$$\int_{L_1(x)}^{L_1(ax)} \frac{\varepsilon_2(t)}{t}\,dt \le \int_{L_1(x)}^{L_1(ax)} \frac{|\varepsilon_2(t)|}{t}\,dt$$

$$\le \int_{L_1(x)}^{L_1(ax)} \frac{\delta}{t}\,dt$$

$$= \delta \int_{L_1(x)}^{L_1(ax)} \frac{1}{t}\,dt$$

$\int_{L_1(x)}^{L_1(ax)} \frac{\varepsilon_2(t)}{t}\,dt \to 0$ as $x \to \infty$

$$= (\delta) \log\left[\frac{L_1(ax)}{L_1(x)}\right].$$

→ BUT since $L_1(x)$ is slowly varying, as $x \to \infty$, $\frac{L_1(ax)}{L_1(x)} = 1 \Rightarrow$ $= \delta \log(1) = 0$. (as $x \to \infty$) (indep of x)

$$\Rightarrow \lim_{x\to\infty} \frac{L(ax)}{L(x)} = \lim_{x\to\infty}\left[\exp\left\{\int_{L_1(x)}^{L_1(ax)} \frac{\varepsilon_2(t)}{t}\,dt\right\}\right] = 1 \quad \forall\ a > 0$$

$\therefore$ $L(x) = L_2[L_1(x)]$ is also a slowly varying function $\square$.

Case 2 as $x \to \infty$, $L_1(x) = a > 0$. $(a < \infty)$

then, the above expression $= \dfrac{L_2(a)}{L_2(a)} = 1$,

$\therefore$ In either case, $\displaystyle\lim_{x\to\infty} \frac{L_2[L_1(tx)]}{L_2[L_1(x)]} = 1$

therefore, $L_2[L_1(x)]$ is also a slowly varying function

---

3. a) $P[Z > x] = P[X + Y > x]$

$$= P[X > x - Y]$$

$$= \int_1^\infty \left[\int_{x-y}^\infty f_{X^*,Y}(x^*, y)\, dx^* \right] dy \quad , \quad \text{need } x - y > 1 \Rightarrow x > y + 1$$

$$= \int_1^\infty \left[\int_{x-y}^\infty f_X(x^*) f_Y(y)\, dx^* \right] dy, \quad \text{as } X \text{ and } Y \text{ are independent}$$

$$= \int_1^\infty \int_{x-y}^\infty \left[\alpha x^{*-\alpha-1}\, \alpha y^{-\alpha-1} \right] dx^*\, dy$$

$$= (\alpha^2) \int_1^\infty (y^{-\alpha-1}) \left[\int_{x-y}^\infty x^{*-\alpha-1}\, dx^* \right] dy$$

$$= (\alpha^2) \int_1^\infty (y^{-\alpha-1}) \left[\frac{x^{-\alpha}}{-\alpha} \Big|_{x-y}^{\infty} \right] dy$$

$$= \frac{(\alpha^2)}{\alpha} \int_1^\infty (y^{-\alpha-1}) \left[(x-y)^{-\alpha} - 0 \right] dy$$

as $\lim_{x \to \infty} x^{-\alpha} = 0$, given $\alpha > 0$

$$= \alpha \int_1^\infty \left[y^{-\alpha-1} (x-y)^{-\alpha} \right] dy$$

Now, multiply $x^{-\alpha} x^{\alpha}$ into the expression

$$= \int_1^\infty \left[\alpha (x^{-\alpha} x^{\alpha}) y^{-\alpha-1} (x-y)^{-\alpha} \right] dy$$

$$= \int_1^\infty (x^{-\alpha}) \, \alpha y^{-\alpha-1} \left(\frac{1}{x}\right)^{-\alpha} (x-y)^{-\alpha} dy$$

$$= \int_1^\infty (x^{-\alpha}) \, \alpha y^{-\alpha-1} \left(1 - \frac{y}{x}\right)^{-\alpha} dy$$

$$= x^{-\alpha} \underbrace{\int_1^\infty \alpha y^{-\alpha-1} \left(1 - \frac{y}{x}\right)^{-\alpha} dy}_{\text{call this integral } L(x)} = x^{-\alpha} L(x)$$

Now, show that $L(x)$ is a slowly varying function

For any $t>0$, $\lim_{x\to\infty} \frac{L(tx)}{L(x)}$

$$= \lim_{x\to\infty} \frac{\int_1^\infty \alpha y^{-\alpha-1}\left(1-\frac{y}{tx}\right)^{-\alpha} dy}{\int_1^\infty \alpha y^{-\alpha-1}\left(1-\frac{y}{x}\right)^{-\alpha} dy}$$

(*) $\lim_{x\to\infty}\left(1-\frac{y}{tx}\right)^{-\alpha}$ and $\lim_{x\to\infty}\left(1-\frac{y}{x}\right)^{-\alpha}$

$= 1^{-\alpha}$ $\qquad = 1^{-\alpha}$

$= 1$ $\qquad = 1$

$$\stackrel{*}{=} \frac{\int_1^\infty \alpha y^{-\alpha-1} dy}{\int_1^\infty \alpha y^{-\alpha-1} dy}$$

$$= 1$$

$\therefore$ $L(x)$ is a slowly varying function $\Rightarrow P(Z>x)=x^{-\alpha}L(x)$ as required $\square$

$\rightarrow \lim_{x\to\infty} L(x) = \lim_{x\to\infty} \int_1^\infty \alpha y^{-\alpha-1}\left(1-\frac{y}{x}\right)^{-\alpha} dy$

$$\stackrel{(*)}{=} \int_1^\infty \alpha y^{-\alpha-1} dy$$

$$= \alpha \frac{y^{-\alpha}}{-\alpha}\Big|_1^\infty = y^{-\alpha}\Big|_\infty^1 = 1 - \overbrace{\lim_{y\to\infty} y^{-\alpha}}^{0 \text{ as } \alpha>0} = 1$$

$\therefore$ $\boxed{\lim_{x\to\infty} L(x) \text{ is finite}}$

3b) → Using the hint, find the distribution of. $n^{-\frac{1}{\alpha}} Z_{(n)}$, where $Z_{(n)} = \max\{Z_1, \ldots, Z_n\}$, Z_i's are iid with df. $F_Z(\cdot)$

$$\to P[n^{-\frac{1}{\alpha}} Z_{(n)} \leq x] = P[Z_{(n)} < n^{-\alpha} x]$$

$$= P^n[Z < n^{-\alpha} x]$$

* $a_n = n^{-\frac{1}{\alpha}}$, $b_n = 0$.

$$= F_Z^n[n^{-\frac{1}{\alpha}} x]$$

$$= \left\{ 1 - \frac{n[1 - F_Z(n^{-\frac{1}{\alpha}} x)]}{n} \right\}^n$$

→ In order to show Z is in the MDA of. Frechet (α), need to show:

$$n[1 - F_Z(n^{\frac{1}{\alpha}} x)] \to x^{-\alpha} \Leftrightarrow P(X+Y > n^{\frac{1}{\alpha}} x) = \frac{x^{-\alpha}}{n} + o\left(\frac{1}{n}\right)$$

$$\to P(X+Y > n^{\frac{1}{\alpha}} x) = \int_{-\infty}^{\infty} \int_{n^{\frac{1}{\alpha}} x - y}^{\infty} x^{*-\alpha-\beta} L_1(x^*) L_2(x^*) \, dx^* \, dy.$$

$\Rightarrow$ Given $n[1-F_Z(n^{\frac{1}{\alpha}}x)] \xrightarrow{d} \phi(x) = x^{-\alpha}$ [Z is in the MDA of Frechet(α)],

$$\Leftrightarrow \quad 1-F_Z(n^{\frac{1}{\alpha}}x) = \frac{x^{-\alpha}}{n} + o\left(\frac{1}{n}\right)$$

$$\Leftrightarrow \quad P(Z > n^{\frac{1}{\alpha}}x) = \frac{x^{-\alpha}}{n} + o\left(\frac{1}{n}\right)$$

$$\Rightarrow \quad P(Z > x) = \frac{(n^{\frac{1}{\alpha}}x)^{-\alpha}}{n} + o\left(\frac{1}{n}\right) \quad , \text{ as } n^{\frac{1}{\alpha}} \text{ is a constant}$$

$$= \frac{x^{-\alpha}}{n^2} + o\left(\frac{1}{n}\right)$$

$$= x^{-\alpha}\underbrace{\left(\frac{1}{n^2} + x^{\alpha}\, o\left(\frac{1}{n}\right)\right)}_{\text{call this function } L_3(x)}$$

Now, tries to show $L_3(x)$ is a slowly varying func.

$\rightarrow$ for all $t>0$, $$\lim_{x\to\infty} \frac{L_3(tx)}{L_3(x)} = \lim_{x\to\infty}\left[\frac{\frac{1}{n^2} + x^{\alpha}t^{\alpha}\, o\left(\frac{1}{n}\right)}{\frac{1}{n^2} + x^{\alpha}\, o\left(\frac{1}{n}\right)}\right]$$

$$= \frac{\frac{1}{n^2} + t^{\alpha}\underbrace{o\left(\frac{1}{n}\right)\left(\lim_{x\to\infty} x^{\alpha}\right)}_{=0}}{\frac{1}{n^2} + \underbrace{o\left(\frac{1}{n}\right)\left(\lim_{x\to\infty} x^{\alpha}\right)}_{=0}}$$

$$= 1$$

$\therefore$ Z has the form $P(Z > x) = x^{-\alpha}L_3(x)$, for some slowly varying $L_3(x)$ $\square$

#Question 4

First, write a function to simulate data from a Burr distribution using the "Inverse CDF" method.

1) Generate $U_1,U_2,...,U_n$ i.i.d. ~ Uniform(0,1) (using the runif(n) function)
2) Set $X_k = F^{-1}(U_k)$ where F is the distribution function of a Burr distribution

Then, $X_1,X_2,...,X_n$ i.i.d. ~ Burr(γ,β)

Note: $F^{-1}(U) = [(1-u)^{-\frac{1}{\beta}} - 1]^{\frac{1}{\gamma}}$

R code:

```
datburr <- function(n,gamma,beta) {
        U <- runif(n)
        X <- ((1-U)^(-1/beta)-1)^(1/gamma)
        return(X)
}
```

In this problem, the following values of γ are used: {4, 3, 2, 1, 0.5, 0.25, 0.1, and 0.05}. In addition, the following conditions are applied:

-The value of β is fixed to be 1 (therefore tail index α = γ)

-For each Burr distribution, 10,000 data points are simulated (n=10,000)

-The values of k are chosen such that a roughly horizontal line can be seen as k increases for both plots (k is smaller than n)

For each value of γ, a Pickands plot and a Hill plot is produced in order to compare the effectiveness of the two methods.

Next, write a function to compute the Pickands estimates (as a function of k, where k can be a vector of values)

```
#Function to compute the Pickands estimates
pickands <- function(x,k) {
        z <- rev(sort(x))
        const <- 1/log(2)
        ests <- NULL
        for (i in k) {
```

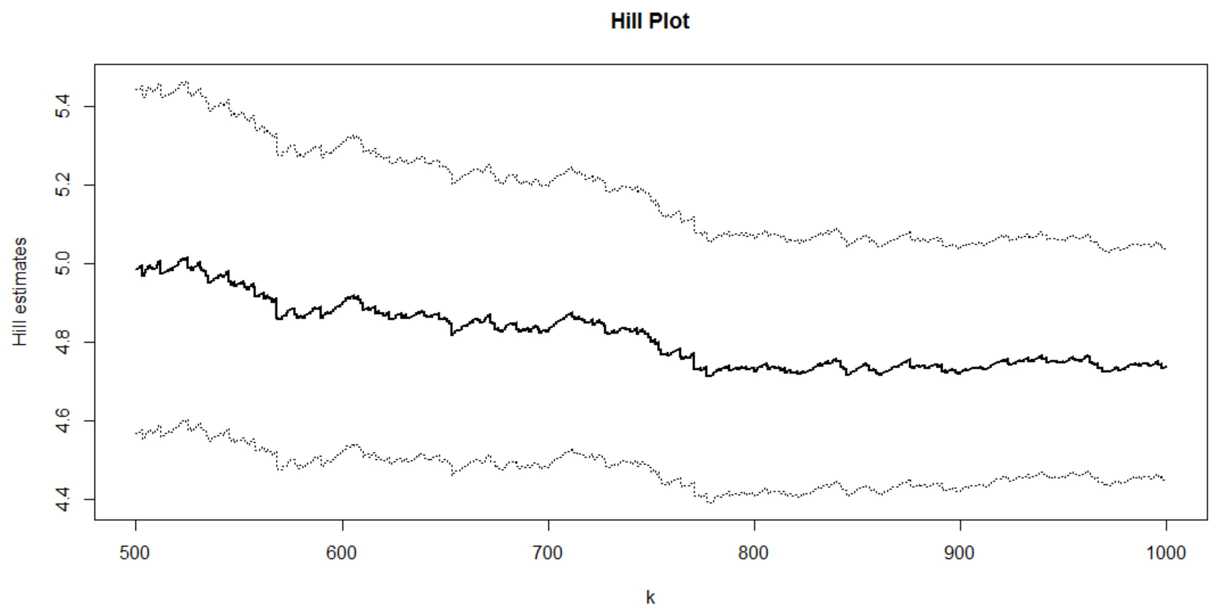
```
        gammak <- const*log( (z[i]-z[2*i])/(z[2*i]-z[4*i])  )
        ests <- c(ests,1/gammak)
    }
    r <- list(pickands=ests,k=k)
    return(r)
}
```

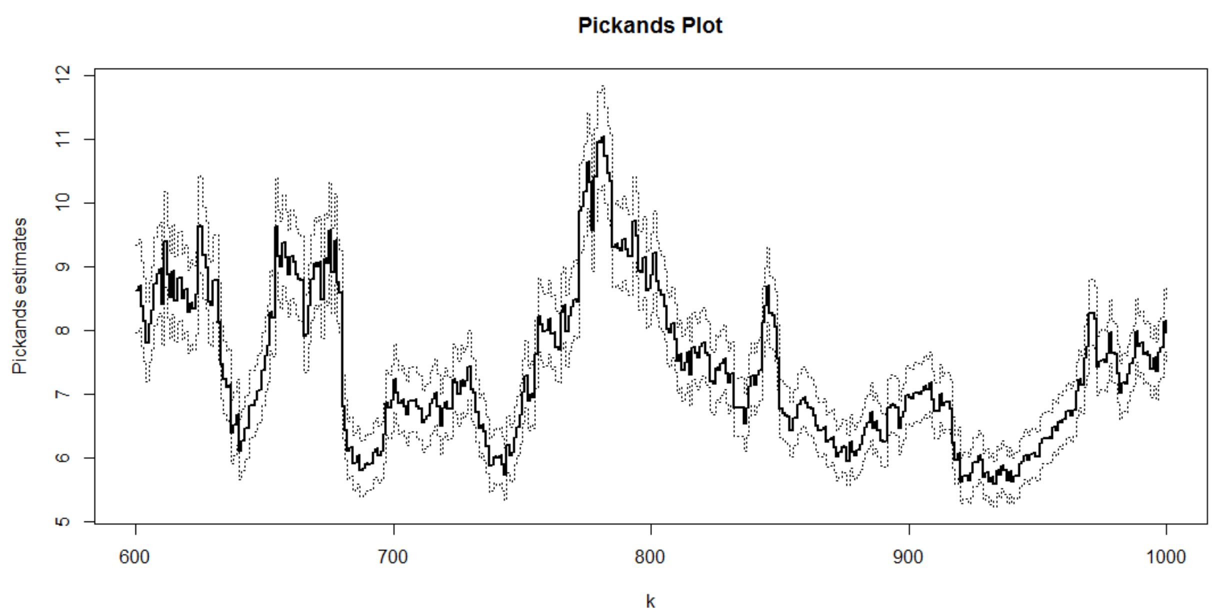
Finally, write a function to plot the Pickands estimates (also plot the approximate lower and upper 95% confidence limits, similar to the graphs in the article on Blackboard)

```
#Function to plot the Pickands estimates
pickandsplot <- function(x,start,end,plot=T) {
    k <- c(start:end)
    r <- pickands(x,k)
    if (plot) {
        upper <- r$pickands*exp(1.96/sqrt(r$k))
        lower <- r$pickands*exp(-1.96/sqrt(r$k))
        limits <- c(min(lower),max(upper))
        plot(r$k,r$pickands,main="Pickands Plot",xlab="k",ylab="Pickands
estimates",type="s",ylim=limits,lwd=2)
        lines(r$k,upper,lty=3,type="s")
        lines(r$k,lower,lty=3,type="s")
    }
    else r
}
```

This following section contains the Pickands and Hill plots for each value of γ

Case 1 γ=5





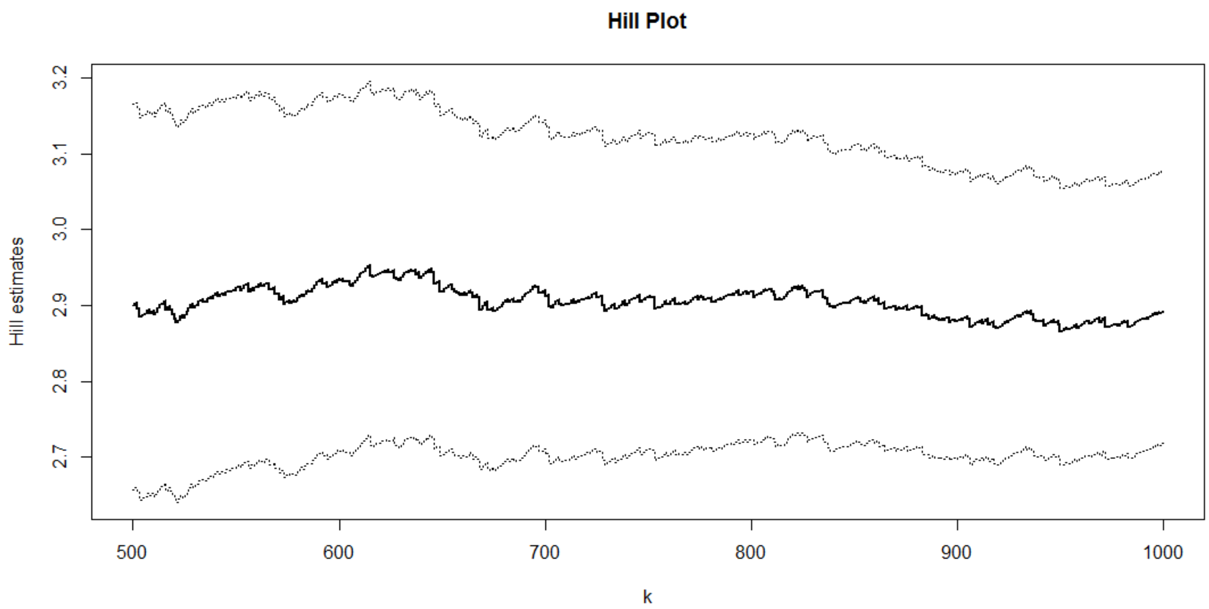
Conclusion:

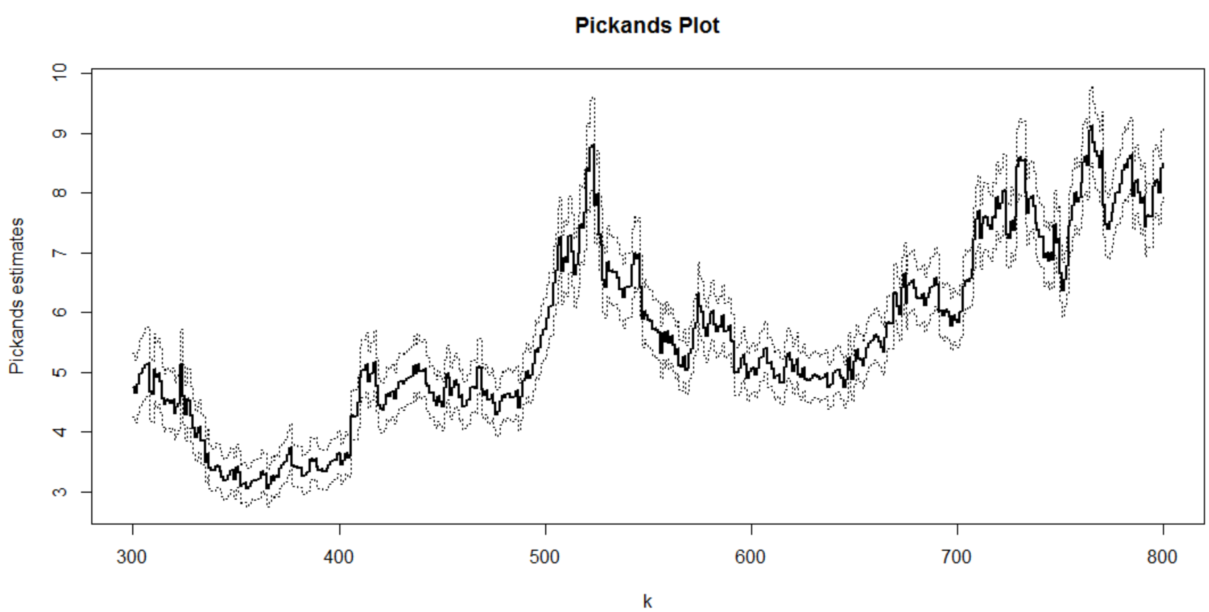
-Hill plot: estimated γ is around 4.8, which is pretty close to the true value of γ=5

-Pickands plot: it is difficult to decide what the estimated γ is as the plot is very volatile

In this case, the Hill plot performs much better than the Pickands plot.

Case 2: γ=3





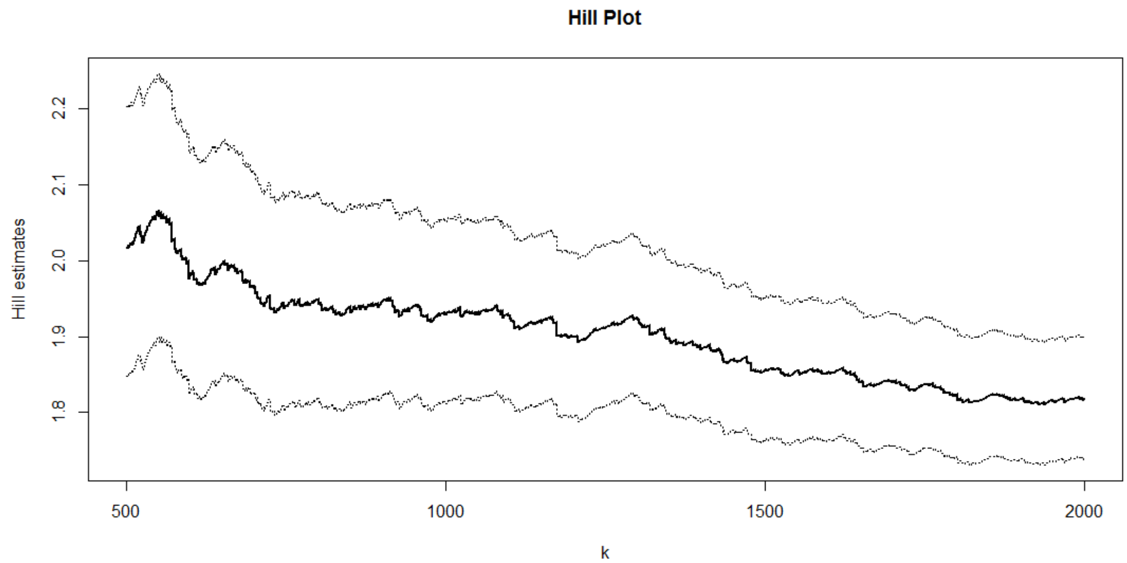
Conclusion:

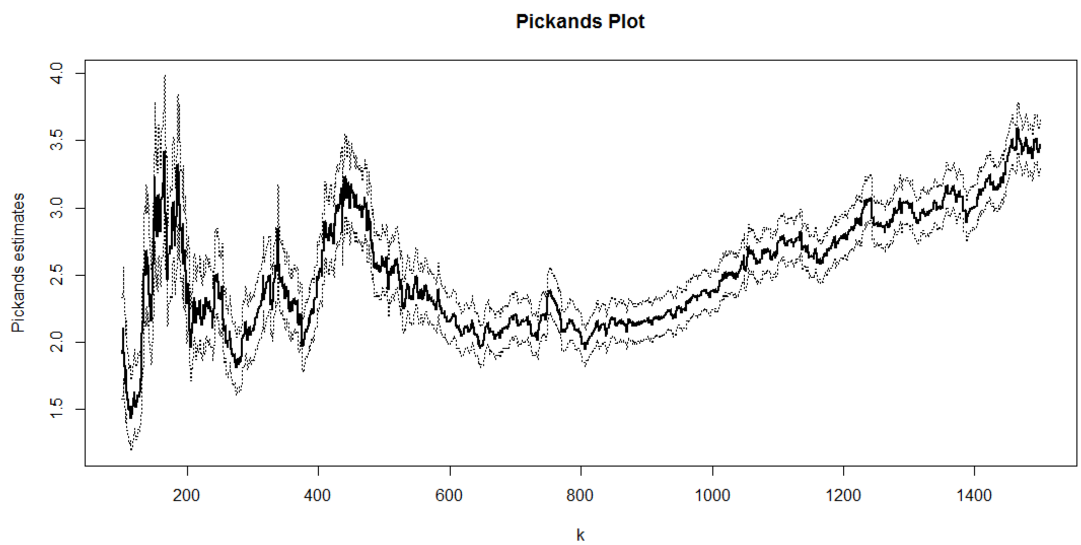
-Hill plot: estimated γ is around 2.9, which is pretty close to the true value of γ=3

-Pickands plot: estimated γ is around 3.5 (first horizontal line), which is not too accurate (but still acceptable). Also note that the estimates are very volatile.

In this case, the Hill plot performs much better than the Pickands plot.

Case 3: γ=2





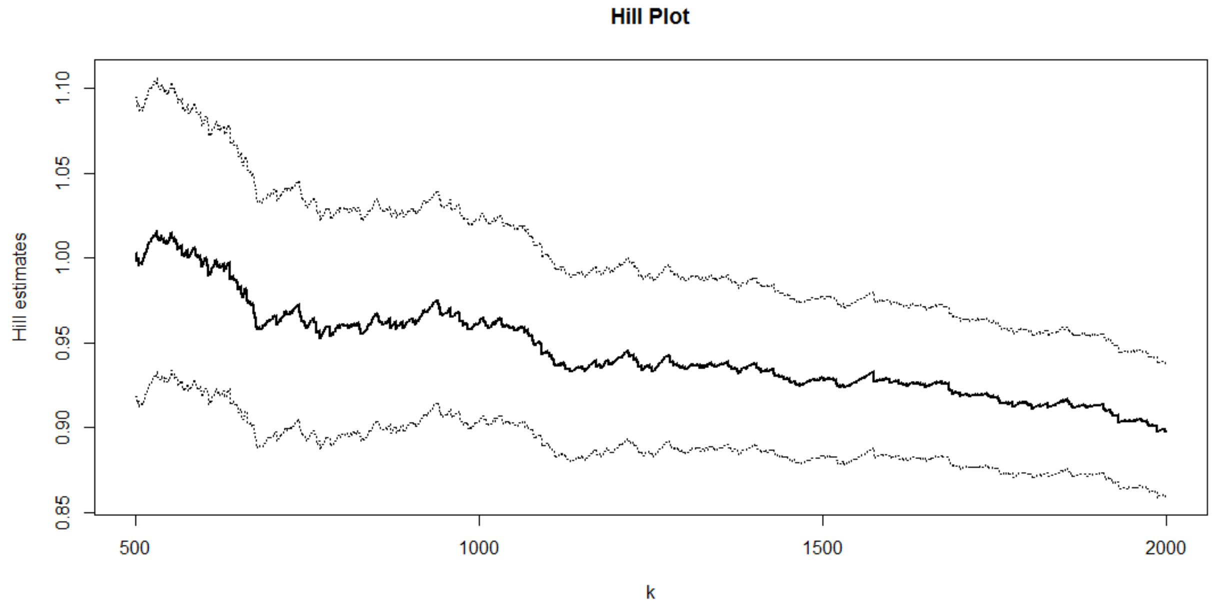
Conclusion:

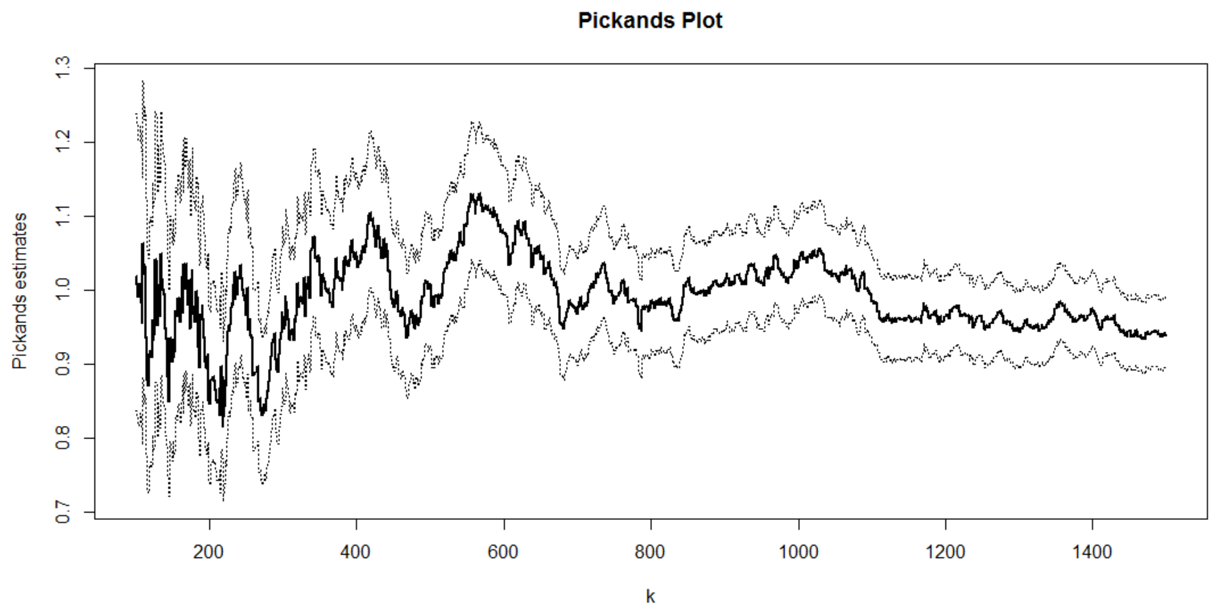
-Hill plot: estimated γ is around 1.9, which is pretty close to the true value of γ=2

-Pickands plot: estimated γ is around 2.2, which is also a good estimate of the true value of γ=2. However, note that the estimates are very volatile (and biased strongly upward after k=800)

In this case, both plots give a very good estimate for γ (while the Hill estimator performs slightly better)

Case 4: γ=1





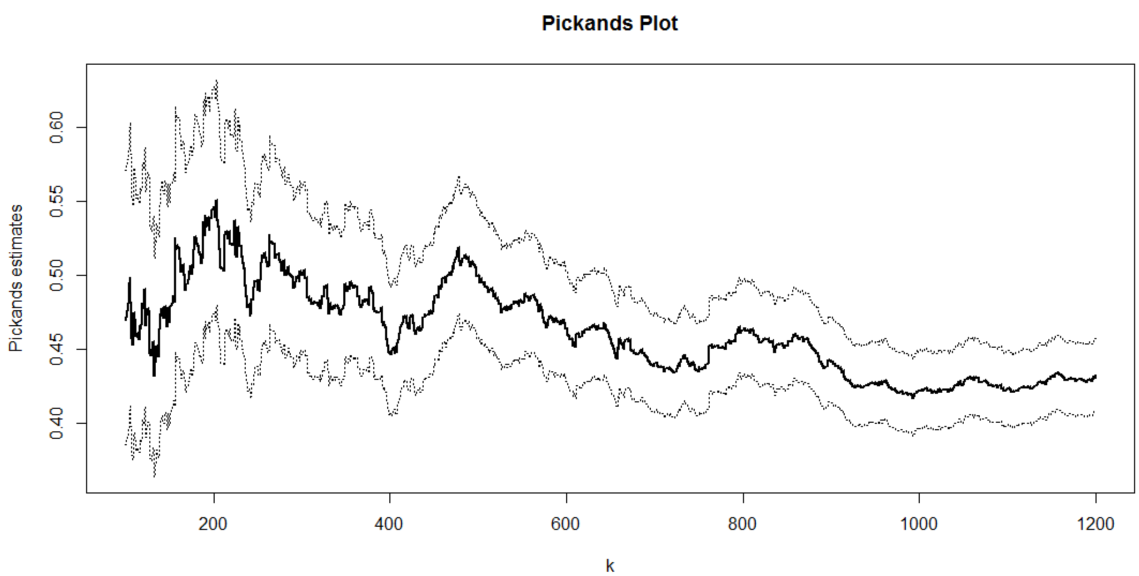
Conclusion:

-Hill plot: estimated γ is around 0.95, which is a very good estimate compared to the true value of γ=1

-Pickands plot: estimated γ is also around 0.95, which is a very good estimate

In this case, both plots give a very good estimate for γ

Case 5: γ=0.5



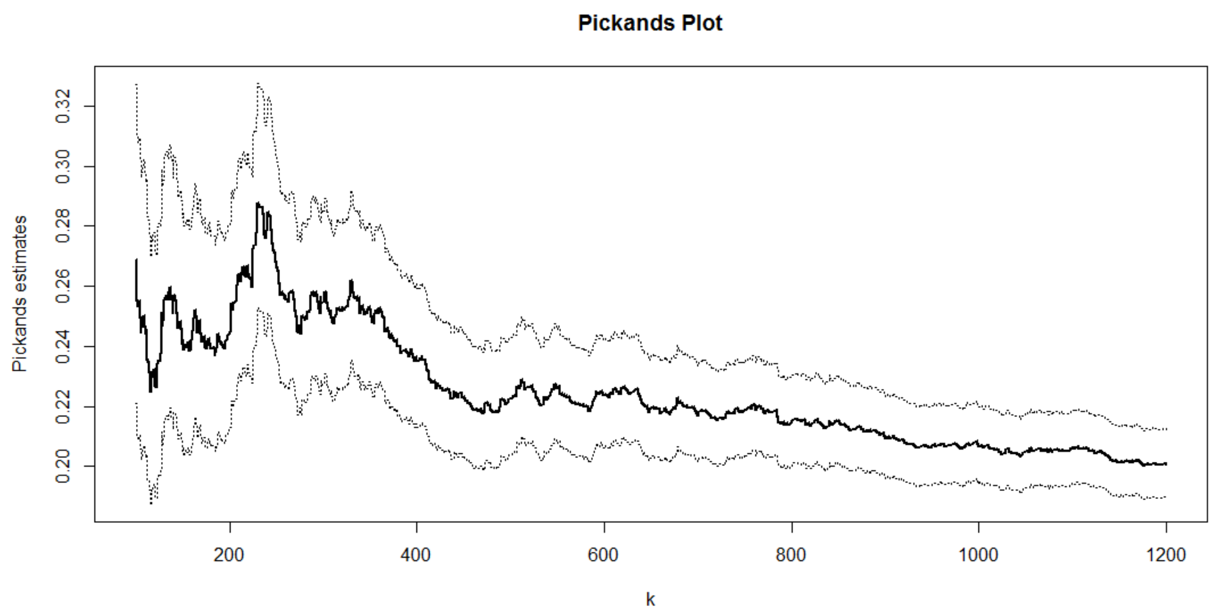
Conclusion:

-Hill plot: estimated γ is around 0.47, which is a very good estimate compared to the true value of γ=0.5

-Pickands plot: estimated γ is around 0.45, which is also a very good estimate compared to the true value of γ=0.5

In this case, both plots give a very good estimate for γ (while the Hill estimator performs slightly better)

Case 6: γ=0.25



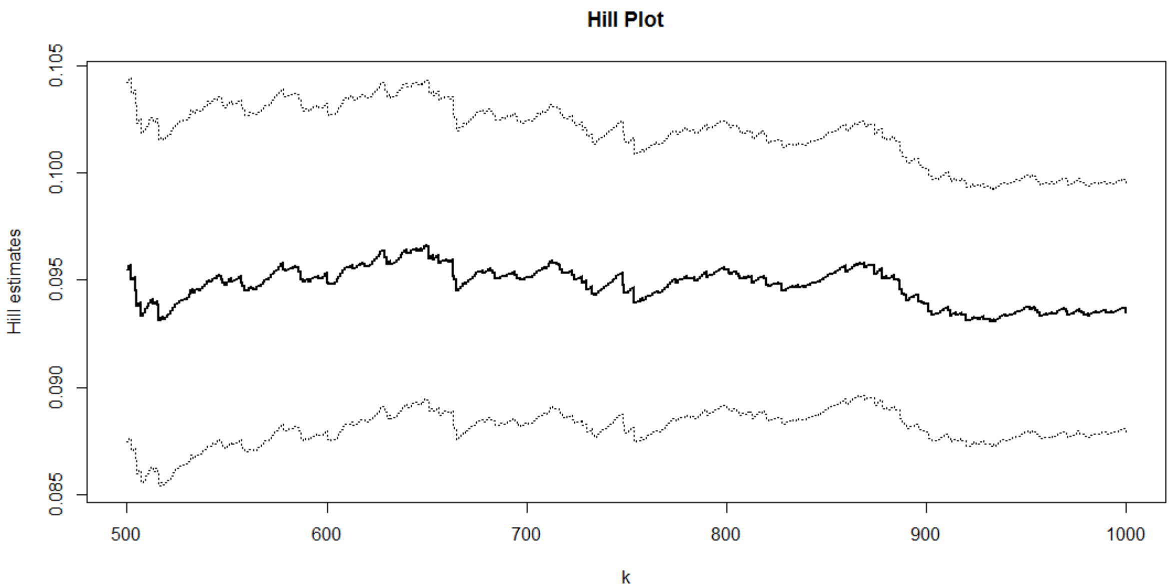
Conclusion:

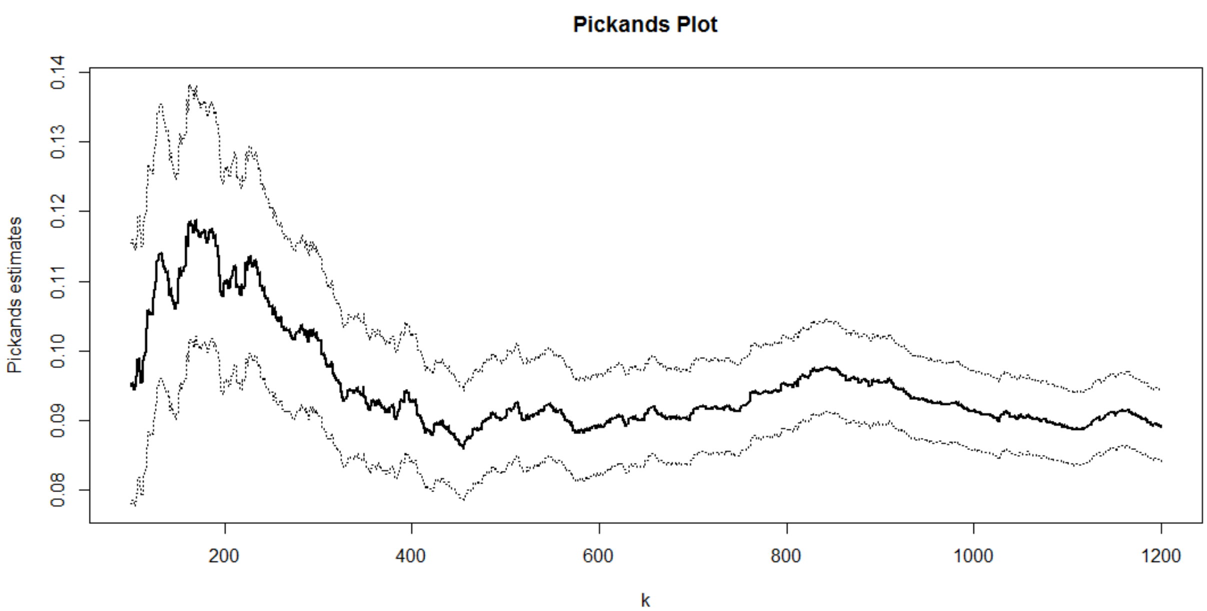
-Hill plot: estimated γ is around 0.235, which is a very good estimate compared to the true value of γ=0.25

-Pickands plot: estimated γ is also around 0.21, which is also a very good estimate compared to the true value of γ=0.25

In this case, both plots give a very good estimate for γ (while the Hill estimator performs slightly better)

Case 7: γ=0.1





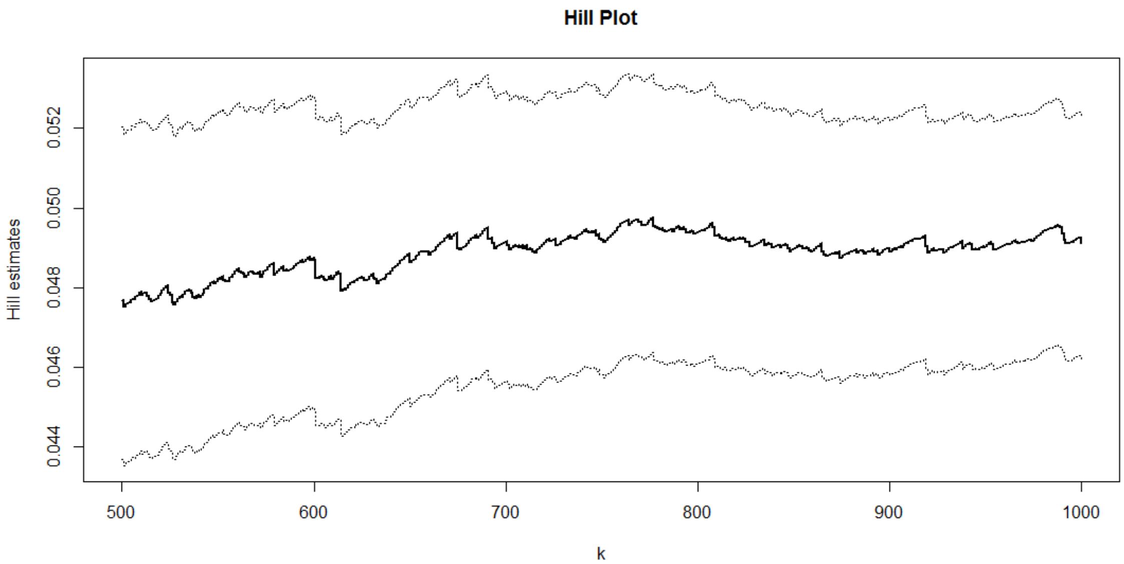
Conclusion:

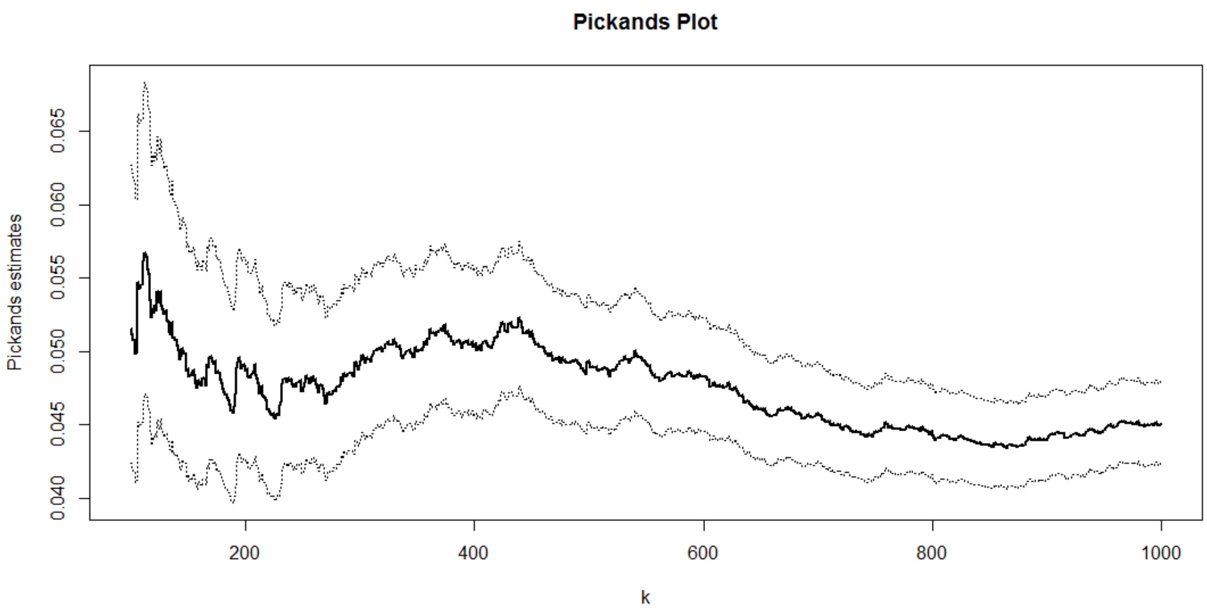
-Hill plot: estimated γ is around 0.095, which is a very good estimate compared to the true value of γ=0.1

-Pickands plot: estimated γ is also around 0.095, which is also a good estimate compared to the true value of γ=0.1

In this case, both plots give a very good estimate for γ

Case 8: γ=0.05





Conclusion:

-Hill plot: estimated γ is around 0.049, which is a very good estimate compared to the true value of γ=0.05

-Pickands plot: estimated γ is also around 0.046, which is also a good estimate compared to the true value of γ=0.05

In this case, both plots give a very good estimate for γ

Comments on the bias of the Pickands estimator:

When $\gamma = \frac{\alpha}{\beta}$ (which here in all cases β=1) decreases, it is noted from the Pickands plots that the estimates become more and more accurate compared to the true value, and therefore the bias decreases as γ decreases. In addition, the variance of the estimates also decreases as γ decreases and the plots become less and less volatile. This may imply that the Pickands estimator performs better when the distribution has a heavy tail (where γ or α are small).

Comments on the performance of Pickands estimator vs Hill estimator:

In general from all the cases above, the Hill estimator gives better estimates than the Pickands estimator, especially in cases where the tail index (γ or α) is not small. In addition, it can be seen from the plots that the Hill estimator is much less volatile than the Pickands estimator, and this can be explained by the fact that the Hill estimator depends on all data in the k indices compared to the Pickands estimator which only depends on three isolated order statistics.

#Question 5

5a)

The model is $Y_i = \mu+X_i$, where i=1,2,..,n, X_i's are i.i.d. standard GEV(γ). Here, assume the mean is 0 and therefore $Y_i = X_i$

First write a function to simulate data from a standard GEV (γ, μ=0, σ=0) using the "Inverse CDF" method.

3) Generate $U_1,U_2,...,U_n$ i.i.d. ~ Uniform(0,1) (using the runif(n) function)
4) Set $X_k = F^{-1}(U_k)$ where F is the distribution function of GEV(γ)
   Then, $X_1,X_2,...,X_n$ i.i.d. ~ GEV(γ)

Note:

Solve for $X_k = F^{-1}(U_k)$

Start with $U = \exp\{-(1+\gamma X)+^{-\frac{1}{\gamma}}\}$

⇨ $(-\ln(U))^{-\gamma} = \max\{0, 1+\gamma X\}$

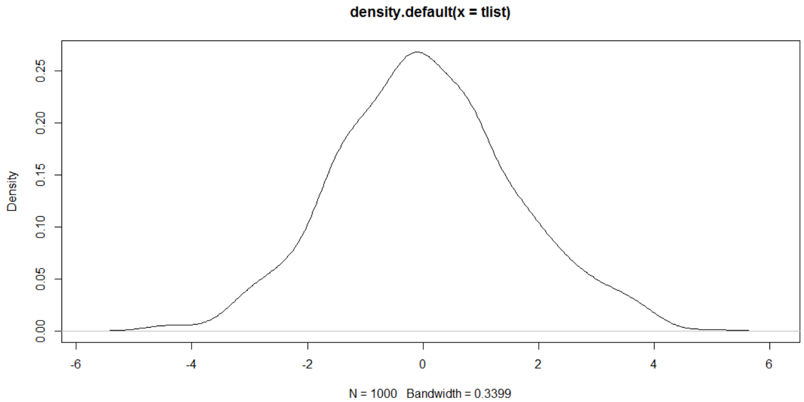
$* \ (-\ln(U))^{-\gamma} \geq 0 \ as \ 0 \leq U \leq 1$

⇨ $\boldsymbol{X = \frac{1}{\gamma}((-\ln(U))^{-\gamma} - 1)}$

R code:

```
#Function to simulate a standard GEV with a single parameter γ

gevscore <- function(n,gamma) {

    Tlist <- NULL

    for (i in 1:n) {

        U <- runif(n)

        X <- (1/gamma)*((-log(U))^(-gamma)-1)

        T <- (1/sqrt(n))*sum(0.5*(X^2)*(1-exp(-X))-X)

        Tlist <- c(Tlist,T)

    }

    return(Tlist)

}

#Test the code under H0 (gamma is very close to 0 and for large n)

tlist = gevscore(n=1000, gamma=0.0005)

mean(tlist) #output is -0.006466072
```

var(tlist) => 2.441276
plot(density(tlist)) => approximately normal



#Now use the function to estimate E(T) for n=10, 50, 100, and 500

For n = 10

tlist1 = gevscore(n=10, gamma=1/10)

mean(tlist1)

- ⇨ 1.659151

Therefore, when n=10, E(T) is estimated to be 1.659151

For n = 50

tlist2 = gevscore(n=50, gamma=1/10)

mean(tlist2)

- ⇨ 2.209361

Therefore, when n=50, E(T) is estimated to be 2.209361

For n = 100

tlist3 = gevscore(n=100, gamma=1/10)

mean(tlist3)

- ⇨ 3.297089

Therefore, when n=100, E(T) is estimated to be 3.297089

For n = 500

tlist4 = gevscore(n=500, gamma=1/10)

mean(tlist4)

- ⇨ 6.849899

Therefore, when n=500, E(T) is estimated to be 6.849899

5b)

Intuition as to why reject H_0 if T > some positive constant:

We know that one property of heavy-tailed data is that the sample mean is usually very misleading, as it usually underestimates the population mean (due to the huge variance and extremes/outliers of the heavy-tailed distribution). One additional consequence of that is as more and more samples are simulated, the sample mean is getting closer and closer to the population mean and therefore the sample mean has a positive correlation with the sample size.

From part a), we see that the sample mean E(T) always increases with the sample size n, and according to the above property, the distribution of T should be a heavy-tailed distribution which also implies Y_i's also come from a heavy-tailed distribution. Therefore, whenever we see T > some positive constant, we know that Y_i's come from a heavy-tailed distribution (which is the Frechet distribution), and therefore we would reject H_0: $\gamma = 0$ (Gumbel distribution).