Assignment 1 STA355

Q1.

a. i. $P(|Z| \leqslant x) = P(-x \leqslant Z \leqslant x)$

$\Phi(\frac{x}{\sigma}) \sim N(0, \sigma^2)$

$Z \sim N(0, \sigma^2)$

$= P(\frac{-x-0}{\sigma} \leqslant Z \leqslant \frac{x-0}{\sigma})$

$= P(Z \leqslant \frac{x}{\sigma}) - P(Z \leqslant \frac{-x}{\sigma})$

$= \Phi(\frac{x}{\sigma}) - [1 - P(Z \leqslant \frac{x}{\sigma})]$

$= \Phi(\frac{x}{\sigma}) - 1 + \Phi(\frac{x}{\sigma})$

$= 2\Phi(\frac{x}{\sigma}) - 1$

ii. $P(|Z| \leqslant G^{-1}(\tau)) = \tau$

$P(-G^{-1}(\tau) \leqslant Z \leqslant G^{-1}(\tau)) = \tau$

$P(-\frac{G^{-1}(\tau)}{\sigma} \leqslant Z \leqslant \frac{G^{-1}(\tau)}{\sigma}) = \tau$

$2\Phi(\frac{G^{-1}(\tau)}{\sigma}) - 1 = \tau$

$\Phi(\frac{G^{-1}(\tau)}{\sigma}) = \frac{\tau+1}{2}$

$\Phi^{-1}(\frac{\tau+1}{2}) = \frac{G^{-1}(\tau)}{\sigma}$

$G^{-1}(\tau) = \sigma \Phi^{-1}(\frac{\tau+1}{2})$

b. $\sqrt{n}(\hat{\sigma}_K - \sigma) \xrightarrow{d} N(0, r^2(\tau))$

$\text{var}[\sqrt{n}(\hat{\sigma}_K - \sigma)]$

$= \text{var}[\sqrt{n}(\frac{W(k)}{\Phi^{-1}(\frac{\tau_k+1}{2})}) - \sqrt{n}\sigma]$

$$= var\left(\sqrt{n}\left(\frac{W_{(k)}}{\Phi^{-1}\left(\frac{\tau_k+1}{2}\right)}\right)\right)$$

$$= \left[\frac{1}{\Phi^{-1}\left(\frac{\tau_k+1}{2}\right)}\right]^2 var(\sqrt{n}W_{(k)}) = r^2(\tau)$$

We know that: If $\{k_n\}$ is a sequence of integers with $\sqrt{n}\left(\frac{k_n}{n}-\tau\right) \to 0$ for some $\tau \in (0,1)$ and $f(F^{-1}(\tau)) > 0$

In this case $\sqrt{n}(W_{(k)} - G^{-1}(\tau)) \xrightarrow{d} N\left(0, \frac{\tau(1-\tau)}{g^2(G^{-1}(\tau))}\right)$

$$Var(\sqrt{n}(W_{(k)})) = \frac{\tau(1-\tau)}{g^2(G^{-1}(\tau))}$$

Thus, $var(\sqrt{n}(\hat{\sigma}_K - \sigma)) = \dfrac{\tau(1-\tau)}{\left[\Phi^{-1}\left(\frac{\tau_k+1}{2}\right)\right]^2 g^2(G^{-1}(\tau))}$

$$= \frac{\tau(1-\tau)}{\left[\Phi^{-1}\left(\frac{\tau_k+1}{2}\right)\right]^2 g^2\left(\sigma\Phi^{-1}\left(\frac{\tau+1}{2}\right)\right)}$$

from a) we get $g(x) = G'(x) = \left(2\Phi\left(\frac{x}{\sigma}\right)-1\right)' = 2\phi\left(\frac{x}{\sigma}\right)\cdot\frac{1}{\sigma}$

$$g\left(\sigma\Phi^{-1}\left(\frac{\tau+1}{2}\right)\right) = \frac{2}{\sigma}\phi\left[\frac{\sigma\Phi^{-1}\left(\frac{\tau+1}{2}\right)}{\sigma}\right] = \frac{2}{\sigma}\phi\left[\Phi^{-1}\left(\frac{\tau+1}{2}\right)\right]$$

$$var(\sqrt{n}(\hat{\sigma}_K - \sigma)) = \frac{\tau(1-\tau)}{\left[\Phi^{-1}\left(\frac{\tau_k+1}{2}\right)\right]^2\left(\frac{2}{\sigma}\phi\left[\Phi^{-1}\left(\frac{\tau+1}{2}\right)\right]\right)^2}$$

Finally we have:

$$r^2(\tau) = \left(\frac{\sigma^2}{4}\right) \frac{\tau(1-\tau)}{[\Phi^{-1}(\frac{\tau_k+1}{2})]^2(\phi[\Phi^{-1}(\frac{\tau+1}{2})])^2}$$

```
> var_mini<-function(x){
+   var<-(x*(1-x))/((2*dnorm(qnorm((x+1)/2))*qnorm((x+1)/2))^2)
+ }
> optimize(var_mini,c(0,1))
$minimum
[1] 0.8616666

$objective
[1] 0.7665833
```

I calaulate the τ by using R.

When $\tau=0.86166$, $r^2(\tau)$ minimized. and $r^2_{min}(\tau)=0.7665$

c. $P(|U| \leqslant x)$ $\qquad U \sim N(\mu_1, \sigma^2)$

$= P(-x \leqslant U \leqslant x)$ $\qquad u' \sim N(0,1)$

$= P\left(\frac{-x-\mu_1}{\sigma} \leqslant u' \leqslant \frac{x-\mu_1}{\sigma}\right)$

$= P\left(u' \leqslant \frac{x-\mu_1}{\sigma}\right) - \left[1 - p\left(u' \leqslant \frac{x+\mu_1}{\sigma}\right)\right]$

$= \Phi\left(\frac{x-\mu_1}{\sigma}\right) + \Phi\left(\frac{x+\mu_1}{\sigma}\right) - 1$

By same way, we have

$P(|v| \leqslant x) = \Phi\left(\frac{x-\mu_2}{\sigma}\right) + \Phi\left(\frac{x+\mu_2}{\sigma}\right) - 1$

Since the distribution of $|u|$ depends on $|\mu_1|$ so that we can assume that $\mu_1 > \mu_2 \geq 0$

If $x \sim N(\mu, \sigma^2)$ for $\mu \geq 0$

$$P(|x| \leq x) = \Phi\left(\frac{x-\mu}{\sigma}\right) + \Phi\left(\frac{x+\mu}{\sigma}\right) - 1$$

WTS : $P(|x| \leq x)$ decreases as μ increases

$$\frac{d}{d\mu} P(|x| \leq x) = \phi\left(\frac{x-\mu}{\sigma}\right) \cdot \left(-\frac{1}{\sigma}\right) + \phi\left(\frac{x+\mu}{\sigma}\right) \cdot \left(\frac{1}{\sigma}\right)$$

WTS : $\frac{d}{d\mu} P(|x| \leq x) < 0$

$$\Rightarrow \phi\left(\frac{x+\mu}{\sigma}\right) \frac{1}{\sigma} < \phi\left(\frac{x-\mu}{\sigma}\right) \frac{1}{\sigma}$$

Since $\phi(\cdot) \sim N(0,1)$ and we know that $x \geq 0$

i.e $\phi(a) = \frac{1}{\sqrt{2\pi}} e^{-\frac{a^2}{2}}$ for $a \geq 0$

larger a implies smaller $\phi(a)$

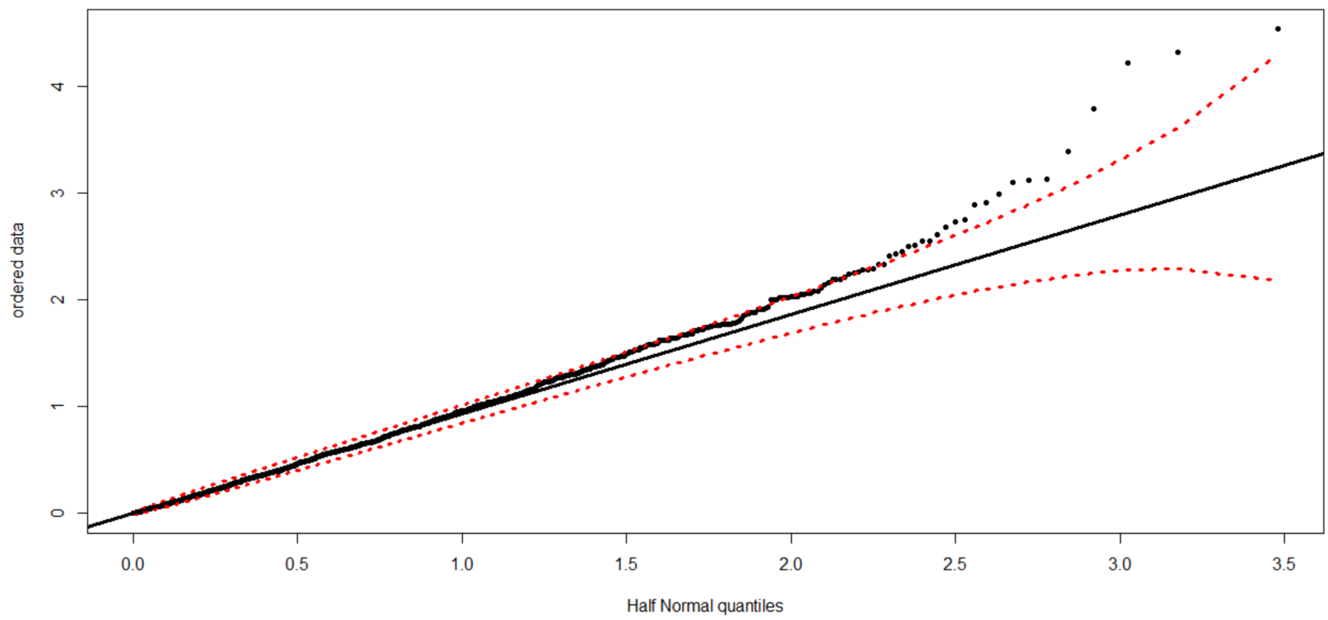
Thus, $\frac{x-\mu}{\sigma} < \frac{x+\mu}{\sigma}$ which implies $\phi\left(\frac{x+\mu}{\sigma}\right) < \phi\left(\frac{x-\mu}{\sigma}\right)$

Hence $\frac{d}{d\mu} P(|x| \leq x) < 0$ and $P(|x| \leq x)$ decreases as μ increases.

From question, $\mu_1 > \mu_2 \geq 0$, we get $P(|u| \leq x) \leq P(|v| \leq x)$

d.

```
> halfnormal <- function(x,tau=0.5,ylim) {
+   sigma <- quantile(abs(x),probs=tau)/sqrt(qchisq(tau,1))
+   n <- length(x)
+   pp <- ppoints(n)
+   qq <- sqrt(qchisq(pp,df=1))
+   # upper envelope
+   upper <- sigma*(qq + 3*sqrt(pp*(1-pp))/(2*sqrt(n)*dnorm(qq)))
+   # lower envelope
+   lower <- sigma*(qq - 3*sqrt(pp*(1-pp))/(2*sqrt(n)*dnorm(qq)))
+   # add upper and lower envelopes to plot
+   if (missing(ylim)) ylim <- c(0,max(c(upper,abs(x))))
+   plot(qq,sort(abs(x)),
+        xlab="Half Normal quantiles",
+        ylab="ordered data",pch=20,ylim=ylim)
+   lines(qq,lower,lty=3,lwd=3,col="red")
+   lines(qq,upper,lty=3,lwd=3,col="red")
+   abline(a=0,b=sigma,lwd=3)}
>
>
> x<-scan("data.txt")
Read 1000 items
> halfnormal(x)
```



There are about 20 points above the upper envelope.
Thus, there are 20 of 1000 means are non-zero.

Q2.

a. $h(x) = \frac{f(x)}{1-F(X)}$ for $x \geq 0$

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

$$E(X) = \int_0^\infty (1-F(X))\,dx$$
$$= \int_0^\infty \frac{f(X)}{h(X)}\,dx$$

Let $x = F^{-1}(\tau)$ so that $F(X) = \tau$ and $d\tau = f(x)\,dx$

$\therefore$ $dx = \frac{1}{f(x)}\,d\tau$

$$E(x) = \int_0^\infty \frac{f(x)}{h(x)} \cdot \frac{1}{f(x)}\,d\tau$$
$$= \int_0^\infty \frac{1}{h(x)}\,d\tau$$
$$= \int_0^\infty \frac{1}{h(F^{-1}(\tau))}\,d\tau$$

b. If $\frac{k_n}{n} \to \tau$ where $0<\tau<1$ and $f(F^{-1}(\tau)) > 0$

then $nD_k \xrightarrow{d}$ Exponential $(f(F^{-1}(\tau)))$

with $E(nD_k) = 1/f(F^{-1}(\tau))$

since $\frac{k}{n} \approx \frac{(k-1)}{n} \approx \tau$ $\quad (k-1) \approx \tau n$

So, $E((n-k+1)D_k) = E((n-(k-1))D_k)$

$$= E((n-\tau n)D_K)$$

$$= E((1-\tau)nD_K)$$

$$= (1-\tau)\ E(nD_K)$$

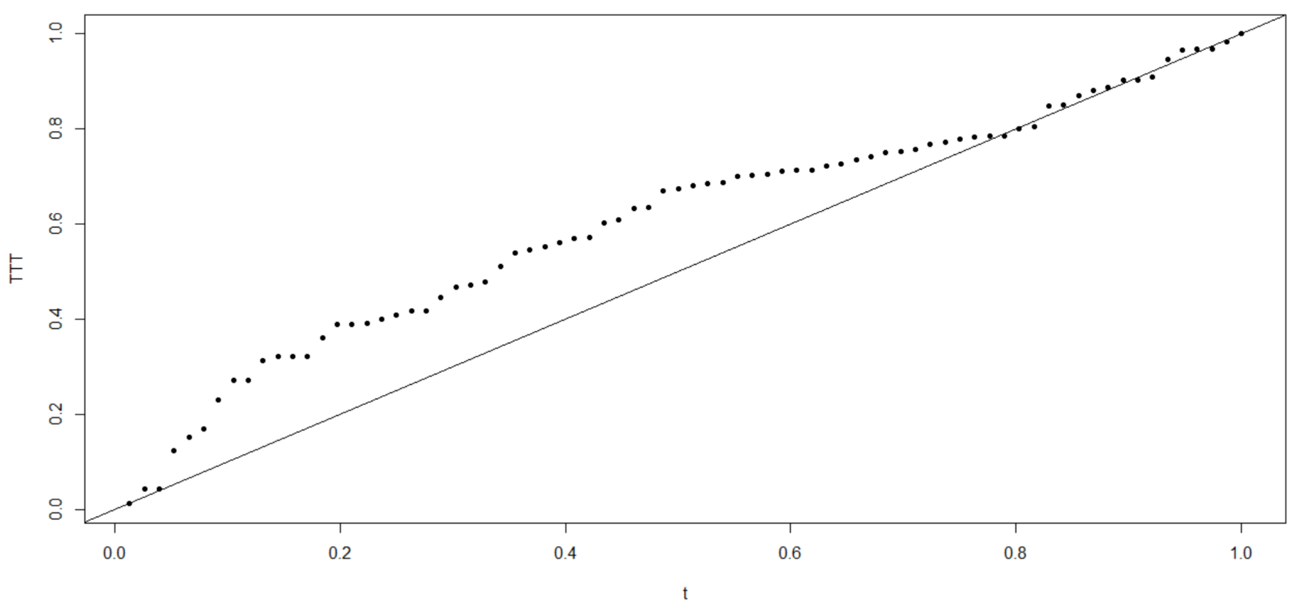
$$= \frac{(1-\tau)}{f(F^{-1}(\tau))}$$

Since $h(F^{-1}(\tau)) = f(F^{-1}(\tau))\ /(1-\tau)$

$$E((n-K+1)D_K) = \frac{1}{h(F^{-1}(\tau))}$$

C.

```
> #Q2-c
> kevlar <- scan("kevlar.txt",skip=1)
Read 76 items
> kevlar <- sort(kevlar)
> n <- length(kevlar)
> d <- c(n:1)*c(kevlar[1],diff(kevlar))
> plot(c(1:n)/n, cumsum(d)/sum(kevlar),xlab="t", ylab="TTT",pch=20)
> abline(0,1)
>
```



Since the shape of these point is concave.

The underlying hazard function $h(x)$ is increasing.