A1 Question 1 a) P21/3

a) i) <u>given</u>

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

Show

CDF of $|Z|$ is $g(x) = 2\Phi(x/\sigma) - 1$ where $\Phi(t)$ is the cdf $N(0,1)$

CDF of $N(0,1)$:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{t^2}{2}\right) dt$$

$$2\Phi\left(\frac{x}{\sigma}\right) - 1 = \frac{2}{\sqrt{2\pi}} \int_{-\infty}^{x/\sigma} \exp\left(\frac{-t^2}{2}\right) dt - 1$$

finding $G(x)$

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

pdf of $|Z|$:

$$\boxed{Z \geq 0}$$

$$\downarrow$$

$$f_{|Z|}(z, \sigma) = 2 f_Z(z; \sigma) = \frac{2}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-z^2}{2\sigma^2}\right)$$

$$= \left(\sqrt{\frac{2}{\pi}}\right)\left(\frac{1}{\sigma}\right) \exp\left(\frac{-z^2}{2\sigma^2}\right) \quad \boxed{z \geq c}$$

Pg 2/3

1a

CDF of $|Z|$

$$F_{|Z|}(z,\sigma) \int_{0}^{z} \left(\frac{1}{\sigma}\right)\left(\sqrt{\frac{2}{\pi}}\right) \exp\left(\frac{-z^2}{2\sigma^2}\right)$$

since we are taking $|Z|$ → 0

Use Change of variables

$x \to \sqrt{2}\,\sigma t$

$t \to \frac{x}{\sqrt{2}\sigma}$

$dx \to \sqrt{2}\sigma\, dt$

$$= \int_{0}^{\frac{z}{\sqrt{2}\sigma}} \frac{1}{\sigma}\sqrt{\frac{\cancel{2}}{\pi}} \exp\left(\frac{-\cancel{2\sigma^2}t^2}{\cancel{2\sigma^2}}\right) \cancel{\sqrt{2}}\sigma\, dt$$

$$= \frac{2}{\sqrt{\pi}} \int_{0}^{\frac{z}{\sqrt{2}\sigma}} \exp(-t^2)\, dt$$

$$G(x) = \frac{2}{\sqrt{\pi}} \int_{0}^{\frac{z}{\sqrt{2}\sigma}} \exp(-t^2)\, dt = \frac{2}{\sqrt{2\pi}} \int_{-\infty}^{x/\sigma} \exp\left(\frac{-t^2}{2}\right) dt - 1$$

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

1a

ii) <u>given</u>

from part 1) and given we know:

$$P(|Z| < G^{-1}(\tau)) = \tau \quad \tau \in (0,1)$$

$$G(x) = 2\phi(\tfrac{x}{\sigma}) - 1$$

Want to Show

τ quantile of the dist $|Z|$ is $G^{-1}(\tau) = \sigma\phi^{-1}(\frac{\tau+1}{2})$

Input $\sigma\phi^{-1}(\frac{\tau+1}{2})$ into $g(x)$

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

$$= 2\cancel{\phi(}\phi^{-1}(\tfrac{\tau+1}{2})) - 1$$

$$= 2(\tfrac{\tau+1}{2}) - 1$$

$$= \frac{2\tau+2}{2} - 1 = \tau+1-1 = \boxed{\tau}$$

A1 Q1 B)

Pg 2/2

B) given

$Z_1, \dots, Z_n$ are independent $N(0, \sigma^2)$

$W_i = |Z_i|$ $i = 1, \dots n$ with order statistics
$W_{(1)} \leq W_{(2)} \leq \dots \leq W_{(n)}$

$$\hat{\sigma}_k = \frac{W(k)}{\phi^{-1}((\gamma_k + 1)/2)}$$

We know $\sqrt{n}\left(\frac{k}{n} - \tau\right)$ $\tau \in (0, 1)$

$g(F^{-1}(\tau)) > 0$

and

$G^{-1}(\tau) = \sigma \phi\left(\frac{\tau + 1}{2}\right)$ (from Part a)

By using the central limit theorem
we know

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

$$= \sqrt{n}\left(W_{(k)} - \sigma \phi\left(\frac{\tau+1}{2}\right)\right) \xrightarrow{d} N\left(0, \frac{\tau(1-\tau)}{g^2\left(\sigma \phi\left(\frac{\tau+1}{2}\right)\right)}\right)$$

$$= \sqrt{n}\left(\frac{W(k)}{\phi\left(\frac{\tau+1}{2}\right)} - \sigma\right) \xrightarrow{d} N\left(0, \frac{\tau(1-\tau)}{g^2\left(\sigma \phi\left(\frac{\tau+1}{2}\right)\right)}\right)$$

A1 Q1 c)

c) $Z_1, \ldots, Z_n$ are independent $N(0,1)$ random variables. Show that for any $\epsilon > 0$

$$P\left(\max_{1 \leq i \leq n} |Z_i| > (1+\epsilon)\sqrt{2\ln(n)}\right) \to 0$$

Using hint we know

$$P\left(\max_{1 \leq i \leq n} |Z_i| > x\right) = P\left(\bigcup_{i=1}^{n} [|Z_i| > x]\right)$$

Bonferroni inequalities

$$P\left(\bigcup_{i=1}^{n} E_i\right) \leq \sum_{i=1}^{n} P(E_i)$$

Using hint with Bonferroni inequality we have:

$$\sum_{i=1}^{n} P\left(|Z_i| > (1+\varepsilon)\sqrt{2\ln(n)}\right) \qquad \boxed{\epsilon > 0}$$

applying $P(|Z| > x) \approx \frac{2}{x\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$

we get

$$\approx \sum_{i=1}^{n} \left(\frac{2}{(1+\varepsilon)\sqrt{2\ln(n)}\sqrt{2\pi}} \exp\left(-\frac{\left[(1+\varepsilon)\sqrt{2\ln(n)}\right]^2}{2} \right) \right)$$

1 c)   Pg 2/2

$$\approx n\left(\frac{2}{(1+\varepsilon)\sqrt{2\ln(n)}\sqrt{2\pi}} \exp\left(\frac{-(1+\varepsilon)^2(2\ln(n))}{2}\right)\right.$$

$$\approx \frac{2n}{(1+\varepsilon)\sqrt{2\ln(n)}\sqrt{2\pi}} \exp\left(-(1+\varepsilon)^2(\ln(n))\right)$$

$$\approx \frac{2n}{(1+\varepsilon)\sqrt{2\ln(n)}\sqrt{2n}} \exp\left(\ln(n)^{-(1+\varepsilon)^2}\right)$$

$$\approx \frac{2n}{(1+\varepsilon)\,2n\sqrt{\ln(n)}}\left(n^{-(1+\varepsilon)^2}\right)$$

$$\approx \frac{n^{1-1-2\varepsilon-\varepsilon^2}}{(1+\varepsilon)\sqrt{\ln(n)}}$$

$$\approx \frac{1}{(1+\varepsilon)\sqrt{\ln(n)}\;n^{\varepsilon(2+\varepsilon)}}$$

$$\lim_{n\to\infty} \frac{1}{(1+\varepsilon)\sqrt{\ln(n)}\;n^{\varepsilon(2+\varepsilon)}} = 0 \quad [\varepsilon>0]$$

A1 Q1 d)

given

$n=1000$ $Z_1, \dots, Z_n$ are independent $N(0,1)$

evaluate

$$P\left[\max(|Z_1|, |Z_2|, \dots, |Z_{1000}|) > \sqrt{2\ln(1000)}\right]$$

we know

$$P(X_{(n)} \leq x) = F(x)^n$$

Thus

$$P(X_{(n)} > x) = 1 - F(x)^n$$

cdF of $|Z|$ is $Z \sim N(0,1)$

$$\sqrt{\frac{2}{\pi}}\int_0^x \exp\left(\frac{-t^2}{2}\right) dt$$

$$1 - F(\sqrt{2\ln 1000})^n$$

$$= 1 - \left(\sqrt{\frac{2}{\pi}}\int_0^{\sqrt{2\ln(1000)}} \exp\left(\frac{-t^2}{2}\right)\right)^{1000}$$

$$= 1 - (\text{Phalfnorm}(\sqrt{2\ln 1000}))^{1000}$$

$$= 0.1826475 = 0.183$$

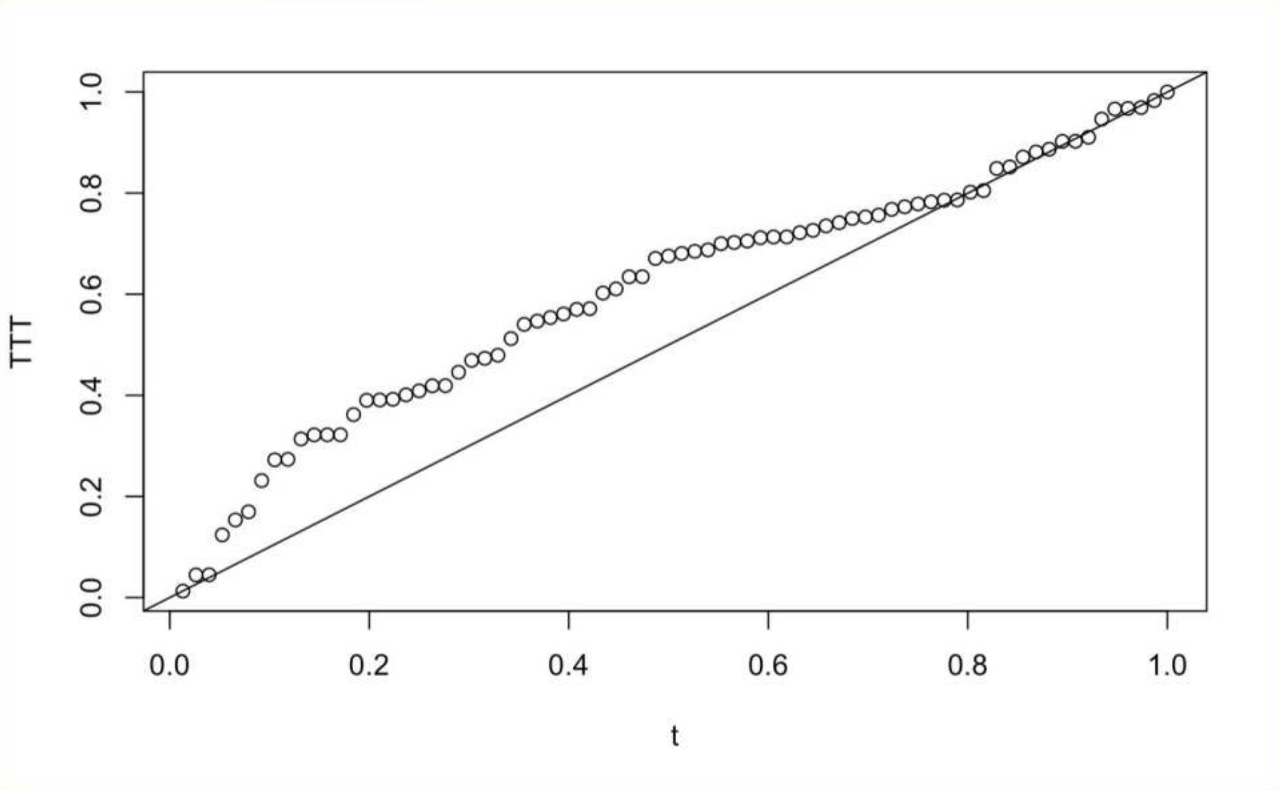
2E

~~~~
```{r}

kevlar <- scan("kevlar.txt")
x <- sort(kevlar) # order elements from smallest to largest
n <- length(x) # find length of x
d <- c(n:1)*c(x[1],diff(x))
plot(c(1:n)/n, cumsum(d)/sum(x), xlab="t", ylab="TTT")
abline(0,1) # add 45 degree line to plot

```
~~~~

R Console



Due to the fact that the points are mostly above the 45 degrees line, we know that the hazard function is increasing with time.

355 A1 Q2 a)

2a) $$h(x) = \frac{f(x)}{1-F(x)}$$

$$h(x) = \lim_{\delta \downarrow 0} \frac{1}{\delta} P(x \le X \le x+\delta \mid X \ge x)$$

$$H(x) = \int_0^x h(t)\, dt$$

Show it equals

$$H(x) = -\ln(1-F(x))$$

X is continous random variable.

we know $F'(x) = f(x)$

see lecture #6

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

$$H(x) = \int_0^x h(t)\, dt$$

$$= \int_0^x \frac{-d}{dt} (\ln(1-F(t)dt$$

Let $R(t) = \ln(1-F(t)$

$$= -\int_0^x R'(t)\, dt$$

$$= -[R(x) - R(0)]$$

2a Pg 212

$$= -\left[(\ln(1 - F(x)) - \ln(1 \overset{0}{\cancel{- F(0)}}))\right]$$

$$= -\ln(1 - F(x))$$

Thus we have showed that

$$H(x) = \int_0^x h(t)\,dt = -\ln(1 - F(x)$$

355 Question # 2 ( )

Given

non neg R.V

$$E(x) = \int_0^{\infty} (1 - F(x))\, dx$$

if $h(x)$ is hazard function Show

$$E(X) = \int_0^1 \frac{1}{h(F^{-1}(\Upsilon))}\, d\Upsilon$$

$$h(x) = \frac{f(x)}{1 - F(x)} \quad x \geq 0$$

Input $(F^{-1}(\Upsilon))$ into function $h(x)$

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

note $F(F^{-1}(x)) = x$

Thus:

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

2C Pg 2/12

as we have Shown

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

Taking the reciprocal we have

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

take the intrgrals

$$\int_0^1 \frac{1}{h(F^{-1}(\tau))} d\tau = \int_0^1 \frac{1-\tau}{f(F^{-1}(\tau))}$$

apply Change of Variables

$u \to F^{-1}(\tau)$ / $\tau \to F(u)$ / $d\tau \to f(u)\,du$

new bounds become $[0,\infty)$

thus

$$\int_0^\infty \frac{1-F(u)}{\cancel{f(u)}} \cancel{f(u)}\, du = \int_0^\infty 1-F(u)\, du$$

$$= E(u)$$

Sta355 q02 d) pg 1/2

given

$X_{(k)}$ is k^{th} order Statistic where $k \approx \tau n$ for some $(\tau \in (0,1))$.

define $D_k = X_{(k)} - X_{(k-1)}$

dist of nD_k is approx exponential with mean $\dfrac{1}{f(F^{-1}(\tau))}$

Want to Show

dist of $D_k(n-k+1)$ is approx exponential with mean $\dfrac{1}{h\,F^{-1}(\tau)}$

$k \approx \tau n$

$(n-k+1)D_k \approx (1-\tau)\,nD_k$ [using hint 2]

$$(n-k+1)D_k \xrightarrow{d} (1-\tau)n\,D_k$$

from hint 1 we know

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

2 d) Pg 2/2

Taking the reciprocal we have

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

Due to the fact we know that nD_k is approx exponential with mean $\frac{1}{f(F^{-1}(\tau))}$, using the equality above

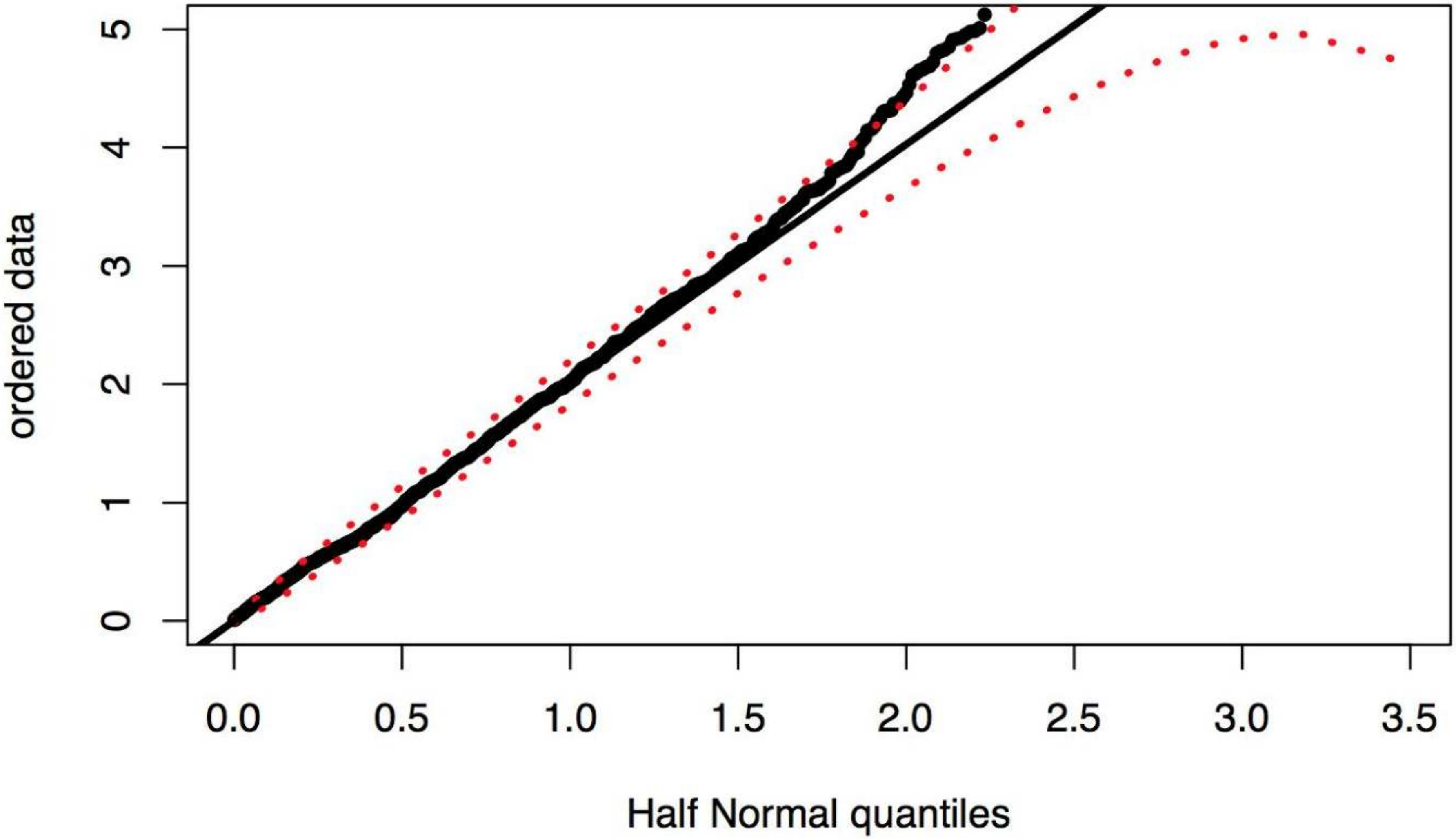
we know the distribution of $(1-\tau)nD_k = (n-k+1)D_k$ is approximately exponential with mean $\frac{1}{h(F^{-1}(\tau))}$

1E

```r
# modified the function halfnormal
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))
  CountNonZero <- 0

# 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)
  # add these code to estimate how many observations have non-zero mean.
  for (i in 1:length(x)){
    if (sort(abs(x))[i] > upper[i]){
      CountNonZero <- CountNonZero + 1
    }
  }
  CountNonZero
  }

data <- scan("data.txt")
halfnormal(data, ylim = c(0, 5))
```


ordered data
0
1
2
3
4
5
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
Half Normal quantiles

```
## [1] 60
```