```
> #A WORKED EXAMPLE, supplement to A statistical approach to distinguish telomere elongation from error in longitudinal datasets
>
> #Read in data using comma seperated format. Each line is an individual with columns for the different timepoints of measurement. Note that in this example we assume that the
time between all measurements is equal, yet different times can be implemented using a seperate independent variable coding for time for each individual in the individual
regressions below. Simulated data are the result of a sample size of 300, average TL start of 100, with average decrease of TL of 3 per time and a SD of the slope of TL of 3,
and error SD of 2. Please refer to the simulation section of main manuscript for additional details. For help please email corresponding author, Mirre Simons at
mirresimons@gmail.com
>
> data=read.csv("simulateddata.csv")
>
> #Estimating error using individual regressions, Equation 4 in main manuscript
>
> matrixresi<-matrix(0,dim(data)[1],1)  # a matrix to put the residuals of the individual regressions in in
> x=1:dim(data)[2] #number of columns is the number of timepoints
> j=1
> while(j<(dim(data)[1]+1)) #loop the individual regressions for the amount of individuals in the dataset
+ {
+ fit<-lm(t(data[j,])~x)  #individual linear regression
+ matrixresi[j,]<-sum((residuals(fit))^2)/(length(x)-2) #residual sum of squares
+ j<-j+1
+ }
> sigma1<-mean(c(matrixresi)) #average residual sum of squares across the individuals
>
>
> #Estimating error under the assumption that TL cannot increase over time, Equation 5 in main manuscript
>
> #first we calculate the increases TL increases over time (between first(1) and last(3) timepoint)
> deltaTL=data[,3]-data[,1]
>
> #Next we determine which individuals increase in TL between the first and last timepoint
> indexTLincreases=which(deltaTL>0)
>
> #We create a new variable including only the data of individual increases
> TLincreases=deltaTL[indexTLincreases]
> #sigma2
> sigma2=0.5*sum(TLincreases^2)/(length(TLincreases))
>
>
> #compare both estimates of error variance (sigma1 and sigma2), Equation 6 in main manuscript
> vratio=sigma2/sigma1
> pvalue<-pf(vratio,length(TLincreases)-1,dim(data)[1]-1,lower.tail=F)
> print(pvalue)
[1] 1.020768e-09
>
>
> #designating a set of individuals who show TL increases with a set confidence interval, e.g. 97.5%, equation 7.
> upperlimit<-((dim(data)[1]-1)*sigma1)/qchisq(0.025,(dim(data)[1]-1)) #Note, change 0.025 in the qchisq to change the confidence at which individual increases are determined.
This upperlimit is the upper confidence of the variance determined by the individual regressions. This variance is the variance of the upper confidence limit of the underlying
normally distributed error function. Using this we can look up individual TL increases that are at the boundary (with 95% confidence) of this normal distribution (with standard
deviation equal to the upper confidence of sqrt(sigma1))
> upperTL=qnorm(0.05,0,sd=sqrt(upperlimit),lower.tail=F)
> outsideconfindex=which((0.5*TLincreases)>upperTL)  #because TL increases are a result from the addition of two equal error distributions divide by 2 (i.e. *0.5).
> print(indexTLincreases[outsideconfindex]) #initial row of individual in the dataset that shows TL increase beyond the set confidence interval
[1]  31  93 138 193 210 231 270 276
>
>
>
>
>
>
```